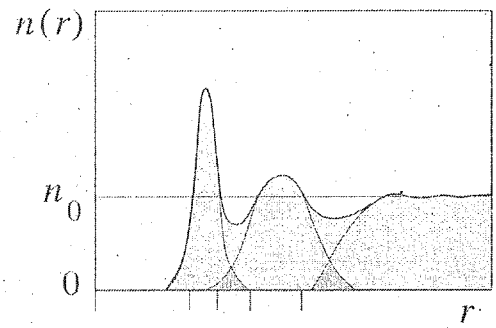
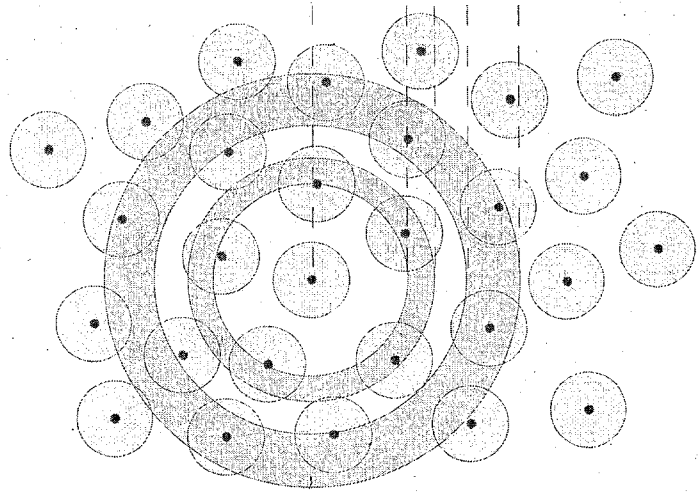
S. Hunklinger

Festkörperphysik



Skript zur Vorlesung
Universität Heidelberg SS 2002
Dozent: A. Pucci



Universität Heidelberg

Inhalt

1. Struktur kristalliner und amorpher Festkörper 1

1.1 Ordnung und Unordnung 1
1.2 Herstellung von Kristallen und amorphen Festkörpern 4
1.3 Struktur der Kristalle 8
1.4 Struktur amorpher Festkörper 24

2. Strukturbestimmung 31

2.1 Allgemeine Bemerkungen 31
2.2 Einfache Theorie der Streuung 33
2.3 Fourier-Entwicklung von Punktgittern 36
2.4 Streuung an Kristallen 43
2.5 Streuung an amorphen Substanzen 54
2.6 Experimentelle Methoden 59

3. Bindungsmodelle 67

3.1 Generelle Aspekte 67
3.2 Van der Waals-Bindung 71
3.3 Ionenbindung 75
3.4 Kovalente Bindung 78
3.5 Metallische Bindung 87
3.6 Wasserstoffbrückenbindung 89

4. Defekte 91

4.1 Punktdefekte 91
4.2 Ausgedehnte Defekte 106
4.3 Defekte in amorphen Materialien 118
4.4 Legierungen 120

5. Gitterdynamik 125

5.1 Schallausbreitung 125
5.2 Bewegungsgleichung der Gitteratome 128
5.3 Gitterschwingungen 131
5.4 Experimentelle Bestimmung von Dispersionskurven 141
5.5 Spezifische Wärmekapazität 155
5.6 Schwingungsspektrum amorpher Festkörper 168
5.7 Wärmekapazität von Gläsern bei sehr tiefer Temperatur 170

6. Anharmonische Gittereigenschaften 175

6.1 Thermische Expansion 175
6.2 Phonon-Phonon-Wechselwirkung 180
6.3 Wärmeleitfähigkeit dielektrischer Kristalle 190
6.4 Wärmeleitfähigkeit amorpher Festkörper 198

7. Elektronen im Festkörperpotential 201

7.1 Freies Elektronengas 201
7.2 Metalleigenschaften im Modell des freien Fermi-Gases 207
7.3 Elektronen im periodischen Potential 219
7.4 Einfluß der Brillouin-Zone 237

8. Bewegung der Elektronen im Festkörper 243

8.1 Bewegungsgleichung und effektive Masse 243
8.2 Elektrische und thermische Leitfähigkeit 251
8.3 Elektronenbahnen im Magnetfeld 267
8.4 Quantisierung der Elektronenbahnen im Magnetfeld 274
8.5 Hall-Effekt 282

9. Halbleiter 287

9.1 Intrinsische kristalline Halbleiter 287
9.2 Dotierte kristalline Halbleiter 297
9.3 Amorphe Halbleiter 307
9.4 Inhomogene Halbleiter 317
9.5 Halbleiterheterostrukturen und Übergitter 330

10. Magnetismus 337

10.1 Dia- und Paramagnetismus 337
10.2 Ferromagnetismus 344
10.3 Ferri- und Antiferromagnetismus 360
10.4 Spingläser 363

11. Dielektrische und optische Eigenschaften 367

11.1 Lokales Feld und Clausius-Mossotti-Beziehung 367
11.2 Elektrische Polarisation 370
11.3 Exzitonen 393
11.4 Optische Eigenschaften freier Ladungsträger 395
11.5 Übersicht über die Absorptionsprozesse 401

12. Supraleitung 403

12.1 Phänomenologische Beschreibung 403
12.2 Mikroskopische Beschreibung 411
12.3 Makroskopische Wellenfunktion 425
12.4 Supraleiter 2. Art 432

1. Struktur kristalliner und amorpher Festkörper

Die große Vielfalt, die man bei den Eigenschaften der Festkörper antrifft, beruht auf dem Zusammenwirken mehrerer, ganz unterschiedlicher Faktoren. Sicherlich spielen die atomaren Eigenschaften eine ganz wesentliche Rolle, doch ist die räumliche Anordnung der Atome und die Art der Bindung zwischen ihnen mindestens genauso wichtig. Um zu einer tieferen Einsicht in die Eigenheiten der Festkörper zu gelangen, erweist sich daher die Beschäftigung mit ihrem Aufbau und den Bindungsmechanismen als unumgänglich. Wir werden uns deshalb zunächst mit der *Struktur* der Festkörper und deren Bestimmung und anschließend mit der Bindung zwischen den atomaren Bausteinen auseinandersetzen.

Beim atomaren Aufbau von Festkörpern lassen sich zwei Grenzfälle unterscheiden: die streng periodische Aneinanderreihung der atomaren Bausteine bei idealen *Kristallen* und die völlig ungeordnete Anordnung bei idealen *amorphen Festkörpern.* Die Struktur realer Festkörper liegt zwischen diesen beiden Grenzfällen. Viele grundlegende Konzepte, die bei der Beschreibung der Festkörpereigenschaften Anwendung finden, setzen bei ihrer Herleitung den periodischen Aufbau der Kristalle voraus. Sie sind daher strenggenommen nur auf diese Substanzklasse anwendbar. Kristalle werden deshalb im Mittelpunkt der Betrachtungen zur Festkörperphysik stehen. Da aber in den letzten Jahren das Interesse an den Eigenschaften und Eigenheiten von komplexeren Strukturen und weniger regelmäßig aufgebauten Festkörpern stark gewachsen ist und verstärkt theoretische Konzepte für ihre Beschreibung entwickelt wurden, schließen wir, wenn auch in geringerem Umfang, amorphe Festkörper in unsere Diskussion ein.

Wir beginnen dieses Kapitel mit einer kurzen Veranschaulichung der Begriffe „Ordnung“ und „Unordnung“, da diese häufig bei einer Beschreibung der Struktur von Festkörpern auftreten. Daran schließen sich einige Anmerkungen zur Herstellung von Festkörpern an, insbesondere von Kristallen und Gläsern. Der folgende Hauptteil ist dann der systematischen Beschreibung von Kristallstrukturen und ihrer Klassifizierung gewidmet. Anschließend greifen wir eine kleine Zahl häufig vorkommender Kristallstrukturen heraus und betrachten diese näher. Der Strukturbeschreibung amorpher Festkörper wenden wir uns kurz am Ende dieses Kapitels zu. Hierbei steht die Erläuterung der prinzipiellen Vorgehensweise bei der Charakterisierung von ungeordneten Strukturen im Vordergrund.

1.1 Ordnung und Unordnung

Anhand einiger einfacher Beispiele von atomaren Strukturen wollen wir zunächst die Begriffe *Ordnung* und *Unordnung* erläutern, die eine zentrale Rolle bei der Unterscheidung zwischen kristallinen und amorphen Festkörpern spielen. Hierzu werfen wir einen Blick auf die Abb. 1.1 – 1.4, in denen einige tatsächlich vorkommende atomare Anordnungen skizziert sind. Der Einfachheit halber sind zweidimensionale Strukturen gezeichnet, doch können wir uns vorstellen, daß es sich jeweils um einen geeigneten Schnitt durch eine dreidimensionale Struktur handelt. So können wir die Abb. 1.1, 1.2a und 1.3a als Schnitte

durch wohlgeordnete Kristalle auffassen, während die restlichen Bilder Strukturen wiedergeben, in denen unterschiedliche Arten von Unordnung auftreten.

Packt man harte Kugeln möglichst dicht, so findet man eine Anordnung, wie sie in Abb. 1.1a zu sehen ist. Ein derartiger Aufbau ist typisch für Kristalle mit hochsymmetrischen Struktureinheiten und ungerichteten Bindungen. Deswegen findet man Kristalle mit so einfacher Struktur bei Edelgaskristallen und bei vielen Metallen. Oft erzwingt aber die chemische Bindung eine andere Anordnung. Bei Graphit, einer Modifikation des Kohlenstoffs, bilden Bindungen einen ebenen 120°-Stern, wodurch sich regelmäßige Sechserringe ausbilden. Die resultierende Anordnung der Kohlenstoffatome ist in Abb. 1.1b gezeigt. In realen Graphitkristallen sind derartige Strukturen schichtweise übereinandergestapelt; zwischen den Schichten herrschen nur schwache Bindungskräfte.

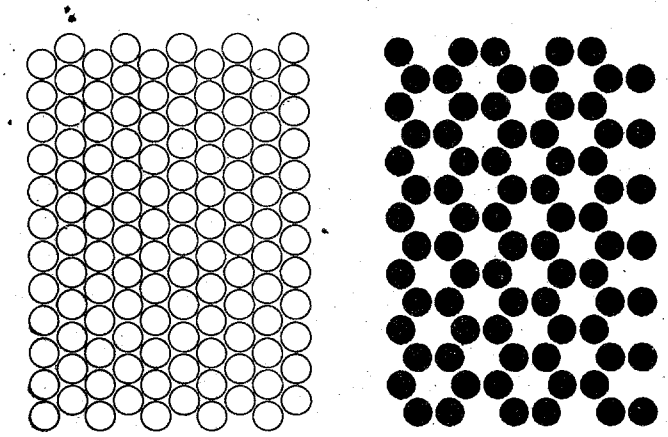


Abb. 1.1: Einfache Kristallstrukturen.
a) Einfacher, aus identischen Atomen aufgebauter Kristall.
b) Anordnung der Kohlenstoffatome in einer Graphitschicht.

In Abb. 1.2a sind ähnliche Ringstrukturen zu sehen, doch enthalten die Ringe nun zwei verschiedene Atomsorten A und B. Auch hier kann man sich vorstellen, daß Kristalle der chemischen Zusammensetzung A_2B_3 entstehen, wenn die dargestellten Schichten übereinandergestapelt werden. Wir wollen hier jedoch noch eine andere Möglichkeit erörtern: Wir können dieses Bild als einen Schnitt durch einen Quarzkristall mit der chemischen Zusammensetzung SiO_2 auffassen, in dem die dunkel gezeichneten vierwertigen Siliziumatome über zweiwertige Sauerstoffatome miteinander verbunden sind. Relativ starre SiO_4-Tetraeder bilden die Bausteine von Quarz, die in gewellten Sechserringen angeordnet sind. Im dargestellten Schnitt sind die vierten Sauerstoffatome, die jeweils zu einem Tetraeder gehören, nicht zu sehen, da sie nach oben oder nach unten aus der Zeichenebene herausragen. Sie sind bereits Bestandteil von weiteren Ringen und bewirken so die dreidimensionale Vernetzung.

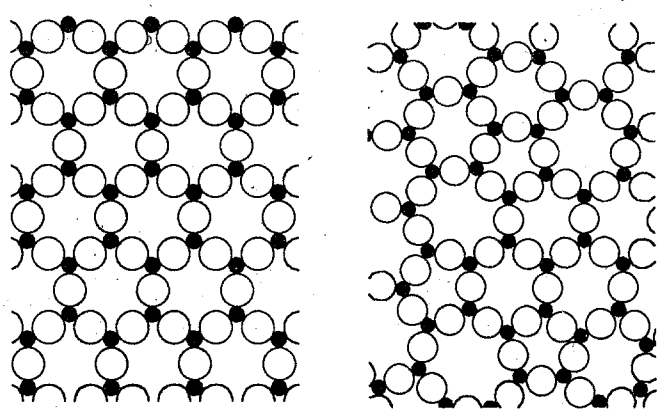


Abb. 1.2: Kristall und Glas mit gleicher chemischer Zusammensetzung.
a) A_2B_3-Kristall. Die schwarzen Punkte und offenen Kreise in den Ringen repräsentieren Atome unterschiedlicher Wertigkeit.
b) A_2B_3-Glas. Die Bindungswinkel zwischen den Atomen unterliegen kleinen Variationen. Dadurch entstehen irregulär geformte Ringe, und die Fernordnung ist unterdrückt.

Nun wenden wir uns der amorphen Struktur 1.2b zu, die aus den gleichen Bestandteilen aufgebaut ist wie die soeben diskutierte kristalline Struktur. Zweiwertige Atome sind mit zwei, dreiwertige Atome mit drei Nachbarn verbunden, so daß *Nahordnung* besteht. Dennoch treten anstelle der regelmäßigen, gleich aufgebauten Ringe nun irregulär geformte Ringe mit unterschiedlicher Zahl von Atomen auf. Der Grund hierfür sind die kleinen Abweichungen der Bindungswinkel von den wohldefinierten Winkeln bei Kristallen. Diese Abweichungen führen zur strukturellen Unordnung und zum Verlust der Ordnung über große Abstände hinweg, zerstören also die *Fernordnung*. Das Fehlen dieser Fernordnung ist charakteristisch für amorphe Festkörper. Im Falle von SiO_2 entspräche der Übergang von Abb. 1.2a zu 1.2b dem Übergang von kristallinem Quarz zum Quarzglas, die beide aus SiO_4-Tetraedern aufgebaut sind.

In Abbildung 1.3 sind zwei Kristalle mit der Zusammensetzung AB gezeigt. Während in Abb. 1.3a die beiden Atomsorten regelmäßig angeordnet sind, weist die Struktur in Abb. 1.3b bei gleicher nominaler Zusammensetzung substitutionelle Unordnung auf, denn die Atome A und B sind statistisch auf den vorhandenen Plätzen verteilt. Es gibt Systeme, die nicht nur in beiden Modifikationen angetroffen werden können, sondern bei denen auch noch ein reversibler Übergang dazwischen stattfinden kann. Ein Beispiel dafür ist die Legierung CuZn, die unter der Bezeichnung „β-Messing" bekannt ist. Wie wir in Abschnitt 4.4 sehen werden, findet im Falle von β-Messing bei der kritischen Temperatur $T_c = 735\,K$ ein sogenannter Ordnungs-Unordnungs-Übergang zwischen den beiden strukturell unterschiedlichen Phasen statt.

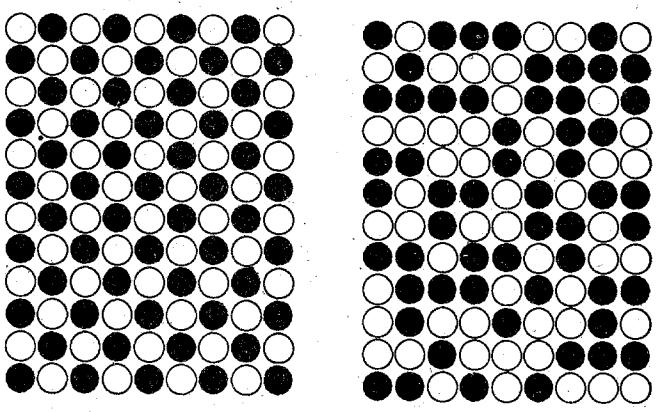


Abb. 1.3: Mögliche Strukturen der Legierung AB.
a) Regelmäßig aufgebauter Kristall AB.
b) Kristall mit substitutioneller Unordnung, bei dem die beiden Atomsorten regellos auf den verfügbaren Plätzen verteilt sind.

In den beiden letzten Abbildungen 1.4a und 1.4b sind die Bausteine selbst nicht vollkommen symmetrisch. Der Begriff Unordnung bezieht sich hier auf die Orientierung der Bausteine, deren Positionen aber einer völlig regelmäßigen Anordnung entsprechen. In Abb. 1.4a trägt eine Atomsorte ein magnetisches Moment, dessen Orientierung durch Pfeile angedeutet ist. In unserem zweidimensionalen Beispiel kann das magnetische Moment in vier verschiedene Richtungen weisen. Diese Art der magnetischen Unordnung findet man beispielsweise in paramagnetischen Salzen, wenn kein äußeres Magnetfeld anliegt. Natürlich handelt es sich hier nur um eine „Momentaufnahme", da die Orientierung der einzelnen magnetischen Momente ständigen Änderungen unterworfen ist. Bei derartigen Systemen tritt beim Abkühlen unter eine kritische Temperatur oft eine Ausrichtung der Momente auf, d.h. es findet ein Phasenübergang von einer paramagnetischen in eine ferromagnetische Phase statt.

In Abb. 1.4b ist ein Kristall dargestellt, der zur Hälfte aus elliptischen Molekülen bzw. Ionen besteht, deren Längsachsen statistisch in eine der beiden Diagonalenrichtungen zeigt. Ein Beispiel hierfür ist der Ionenkristall KCN, der aus kugelförmigen Kalium- und ellipsoidförmigen Cyanidionen aufgebaut ist, wobei die CN-Ionen regellos in Richtung der Raumdiagonalen eines Würfels orientiert sind. Oberhalb einer kritischen Temperatur $T_c = 168\,\mathrm{K}$ führen die Ellipsoide ständig Drehsprünge zwischen den Gleichgewichtslagen aus. Unterhalb von T_c sind sie alle in Richtung einer bestimmten Raumdiagonalen ausgerichtet, d.h. bei T_c findet eine Strukturumwandlung statt.

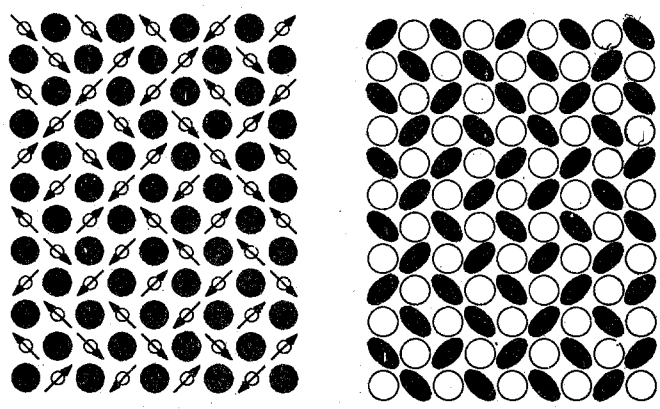


Abb. 1.4: Kristalle mit Orientierungsunordnung.
a) Magnetische Momente, die durch kleine Pfeile angedeutet sind, sind statistisch orientiert.
b) Ellipsoidförmige Moleküle besetzen die vorhandenen Gleichgewichtslagen ohne bevorzugte Orientierung.

Wir haben gesehen, daß auf der einen Seite der Idealkristall mit seinem streng periodischen Aufbau und auf der anderen Seite der amorphe Festkörper mit seiner weitgehend ungeordneten Struktur steht. Wie in den Abbildungen 1.1 – 1.4 zum Ausdruck kommt, liegt dazwischen ein breites Spektrum von Festkörperstrukturen, in denen eine weitgehende, aber keineswegs vollkommene Ordnung herrscht.

1.2 Herstellung von Kristallen und amorphen Festkörpern

Die meisten Festkörper entstehen beim Abkühlen ihrer Schmelze, wofür das Erstarren von Wasser zu Eis ein bekanntes Beispiel ist. Uns interessiert hier die Frage, unter welchen Bedingungen geordnete Kristalle und wann ungeordnete amorphe Festkörper gebildet werden. Trifft man keine besonderen Vorkehrungen, so liegen in den meisten Fällen nach dem Erstarren *polykristalline* Festkörper vor. Dies ist eine Ansammlung von regellos orientierten, meist kleinen Kristalliten, die über weitgehend ungeordnete Grenzschichten, die sogenannten *Korngrenzen* (vgl. Abschnitt 4.2), miteinander verbunden sind.

Unter gewissen Voraussetzungen erhält man aber *Einkristalle* mit durchgehend regulärer Atomanordnung, die im gesamten Volumen die gleichen physikalischen Eigenschaften aufweisen. Diese Eigenschaften sind richtungsabhängig, d.h. Einkristalle zeigen ein *anisotropes* Verhalten. Im Gegensatz hierzu besitzen polykristalline Materialien *isotrope* makroskopische Eigenschaften, die den Mittelwert über die anisotropen Eigenschaften aller regellos orientierten Kristallite widerspiegeln.

Einkristalle sind für die Festkörperphysik und die Materialwissenschaften von großer Bedeutung, da sich an ihnen die „reinen" Kristalleigenschaften besonders gut studieren lassen. Sie spielen aber auch in technischen Anwendungen eine große Rolle. So finden beispielsweise in Uhren und in Geräten der Nachrichtentechnik piezoelektrische Einkristalle eine breite Anwendung. Für die Halbleitertechnik werden sehr große, außerordentlich perfekte Einkristalle aus Silizium in großer Stückzahl produziert.

Die Herstellung („Zucht") von größeren Einkristallen ist jedoch meist aufwendig und erfordert viel Erfahrung. Ein wesentlicher Gesichtspunkt ist die Reinheit der Ausgangssubstanzen. Um ihre Reinheit zu gewährleisten, wird oft das *Zonenreinigungsverfahren* angewandt. Man legt dazu das Ausgangsmaterial in Form eines langen (polykristallinen) Stabs in einen Schmelztiegel und zieht diesen durch einen Ofen, der so gestaltet ist, daß nur ein schmaler Bereich („Zone") des Stabs aufgeschmolzen wird. So wird an der Vorderseite dieser Zone ständig neues Material flüssig und erstarrt auf der Rückseite. Bei diesem Verfahren macht man sich zunutze, daß der Einbau von Verunreinigungen in den meisten Fällen energetisch ungünstig ist. An der Erstarrungsfront unterscheidet sich daher die Konzentration an Verunreinigungen: Sie ist in der Schmelze höher als im neu entstandenen Kristall. Auf diese Weise werden die Verunreinigungen mit der geschmolzenen Zone mitgenommen und an das Ende des Tiegels transportiert.

Die Qualität der Kristalle wird vor allem durch die Zuchtbedingungen bestimmt. Um Einkristalle zu erhalten, darf die Kristallisation natürlich nicht an verschiedenen Stellen der Schmelze einsetzen, denn dies würde zu den oben erwähnten polykristallinen Festkörpern führen. In vielen Fällen wird daher im Gefäß eine definierte Temperaturverteilung eingestellt und bereits ein Kristallkeim vorgegeben, d.h. es wird ein geeignet ausgerichteter kleiner Kristall als Keim in die Wachstumszone eingebracht, der dann durch die Anlagerung von Atomen bzw. Molekülen aus der Schmelze wächst. Um weitgehend perfekte Kristalle zu erhalten, sollten in der Wachstumszone nur kleine Temperaturdifferenzen auftreten, d.h. der Wachstumsprozeß sollte möglichst nahe am thermischen Gleichgewicht ablaufen.

Am häufigsten werden Einkristalle aus Lösungen oder Schmelzen gezogen, doch gibt es noch viele andere Verfahren. Welches jeweils zur Herstellung bestimmter Kristalle am geeignetsten ist, hängt von den speziellen Kristalleigenschaften ab. Bei der Zucht aus der Lösung läßt man entweder das Lösungsmittel bei konstanter Temperatur langsam verdampfen oder erniedrigt allmählich die Temperatur, wobei man die Lösung in Sättigung hält. Eine sehr wichtige Methode zur Herstellung von Kristallen, auf die wir gleich noch kurz eingehen, ist das Ziehen aus der Schmelze. In den letzten Jahren haben epitaktische Verfahren an Bedeutung gewonnen. Dort werden aus der Gasphase oder aus Lösungen einkristalline Schichten lagenweise auf einkristallinen Unterlagen abgeschieden. Voraussetzung hierfür ist eine weitgehende Übereinstimmung der Atomabstände von Substrat und aufgebrachter Schicht. Über die vielen unterschiedlichen Zuchtverfahren existiert eine reichhaltige Literatur[1].

Wir wollen hier noch drei Zuchtverfahren erwähnen, die angewandt werden, um Kristalle aus ihren Schmelzen zu ziehen. Silizium- und Germaniumkristalle stellt man beispielsweise meist nach dem *Czochralski-* oder *Kyropoulos-Verfahren* her. Hierbei wird ein Keimkristall an einem Halter langsam aus dem geschmolzenen Material hochgezogen. Hält man die Temperatur des Keimlings ein paar Kelvin unter, die der Schmelze ein paar Kelvin über dem Schmelzpunkt, so wächst dabei der Keimling. Um Temperatur- und Druckgradienten weitgehend zu vermeiden, läßt man den Halter mit dem wachsenden Einkristall gleichzeitig langsam rotieren. Typische Ziehgeschwindigkeiten sind dabei einige Millimeter pro Stunde.

[1] Eine eingehende Darstellung der verschiedenen Techniken findet man bei: K.-Th. Wilke, J. Bohm, Kristallzüchtung, Harri Deutsch Verlag, Frankfurt/M.

Beim *Bridgman-Stockbarger-Verfahren* befindet sich das Grundmaterial in einem Behälter mit einem konischen Ende. Im ersten Schritt wird das Ausgangsmaterial im oberen Teil des Ofens geschmolzen. Anschließend wird der Tiegel in den unteren Teil des Ofens abgesenkt, dessen Temperatur einige Kelvin unter der Schmelztemperatur liegt. Dabei bildet sich an der Spitze des Konus meist ein einzelner Keim, von dem aus dann ein Einkristall in den Tiegel wächst.

Bei beiden Verfahren besteht die Gefahr, daß die Schmelze durch den Tiegel, der z.B. aus Graphit besteht, verunreinigt wird. Interessanterweise kann man das Prinzip des Zonenreinigungsverfahrens auch benutzen, um Kristalle tiegelfrei zu ziehen. Das Prinzip ist in Abb. 1.5 dargestellt. Am unteren Ende des polykristallinen Stabs befindet sich der Keimkristall. Die schmale geschmolzene Zone, die zwischen den beiden festen Enden der Probe befindet, wird durch die Oberflächenspannung der Schmelze gehalten. Langsam läßt man nun die Schmelzzone längs des Stabs nach oben wandern, wodurch von unten her ein Einkristall hochwächst. Bei diesem *tiegelfreien Zonenschmelzen* („floating-zone technique") wird also nicht nur ein Einkristall gezogen, sondern das Ausgangsmaterial gleichzeitig auch noch gereinigt.

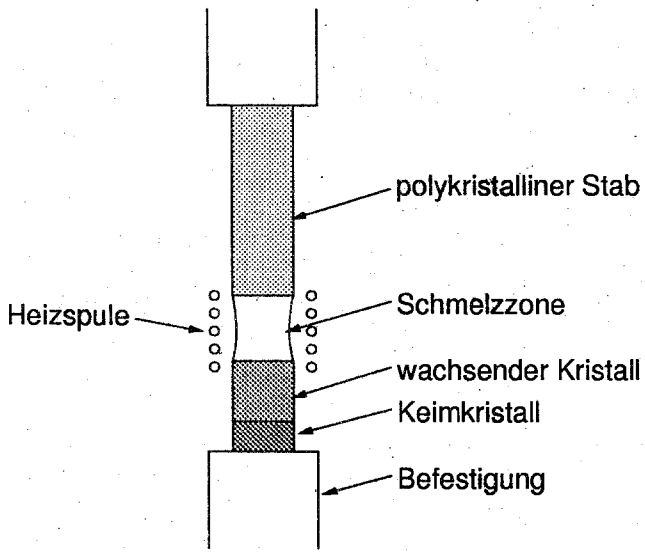


Abb. 1.5: Kristallzucht im tiegelfreien Zonenschmelzverfahren. Der Einkristall wächst vom Keimkristall ausgehend von unten nach oben.

Der Begriff „amorpher Festkörper" ist nicht exakt definiert. Er sagt im wesentlichen aus, daß sich die Atome nicht wie im Kristall auf regelmäßig angeordneten Plätzen befinden, wobei der Grad der Unordnung unterschiedlich groß sein kann. Neben den *Gläsern* zählen zu dieser Stoffklasse viele Dünnschichtsysteme, die durch Aufdampfen erzeugt wurden, Keramiken, Kunststoffe oder auch biologische Festkörper. Wir greifen aus der Vielzahl der amorphen Festkörper die Gläser heraus, da sie typische Vertreter dieser Substanzklasse sind. Der wesentliche Unterschied in der Herstellung von Kristallen und Gläsern besteht darin, daß bei der Kristallzucht schnelle zeitliche und räumliche Temperaturänderungen vermieden werden, während bei der Glasherstellung die Schmelze rasch abgekühlt wird.

Wie in Abb. 1.6 skizziert, treten beim Schmelzen und Erstarren von Festkörpern charakteristische Volumenänderungen auf, die bei Gläsern und Kristallen einen unterschiedlichen Verlauf aufweisen. Kühlt man eine Schmelze so langsam ab, daß die Ausbildung

geordneter Strukturen mit atomarer Fernordnung möglich ist, so tritt bei der Erstarrungstemperatur T_f Kristallisation ein. Dies ist ein Phasenübergang 1. Ordnung, der, wie in der Skizze angedeutet, von einer diskontinuierlichen Volumenänderung begleitet ist.

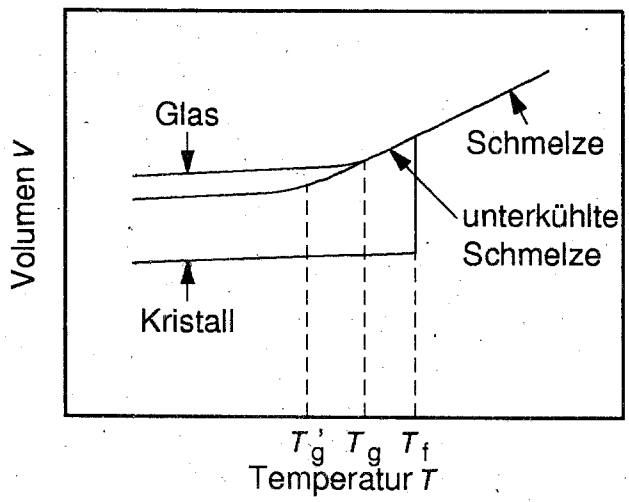


Abb. 1.6: Volumenänderung beim Erstarren bzw. bei der Kristallisation einer Schmelze. Letztere tritt bei der wohldefinierten Erstarrungstemperatur T_f ein. Der Glasübergang findet abhängig von der Kühlrate bei T_g oder T_g' statt. Die Volumenänderungen sind nicht maßstäblich gezeichnet.

Kühlt man dagegen eine Schmelze rasch ab, so kann die Kristallisationsstemperatur T_f unterschritten werden, ohne daß Verfestigung eintritt. Erst wesentlich unterhalb T_f erstarrt die Schmelze zu einem Glas. Anschaulich kann man sich vorstellen, daß in einer Schmelze ständig atomare Umlagerungen erfolgen, die mit abnehmender Temperatur immer langsamer ablaufen. Da das Aufbrechen von atomaren Bindungen in der Schmelze und die geordnete Anlagerung der Atome an der Kristalloberfläche aber Voraussetzungen für Kristallwachstum sind, wird die Kristallbildung in der unterkühlten Schmelze unterdrückt. Im V–T-Diagramm macht sich die Glasbildung als Knick in der Zustandskurve bei der sogenannten *Glasübergangstemperatur* T_g bemerkbar. Unterhalb dieser Temperatur ist die Anordnung der Atome innerhalb experimentell zugänglicher Zeitskalen eingefroren. Da sich die Struktur der Gläser kaum von der ihrer Schmelzen unterscheidet, bezeichnet man sie oft auch als „eingefrorene Flüssigkeiten". Wie der Skizze zu entnehmen ist, weisen Gläser ein größeres Volumen und somit eine geringere Dichte auf als die entsprechenden Kristalle.

Die Glasübergangstemperatur T_g liegt typischerweise bei etwa 2/3 der Erstarrungstemperatur des entsprechenden Kristalls, d.h. $T_g \approx 2T_f/3$. Für Quarzglas, das aus reinem SiO_2 besteht, liegt T_g bei 1450 K. Abhängig von der Zusammensetzung weisen optische Gläser oder „Fensterglas" Werte von T_g um 900 K auf. Bei der wäßrigen Lithiumchloridlösung $LiCl{\cdot}7H_2O$ liegt T_g weit unterhalb des Eispunktes, nämlich bei 137 K. Die Kühlraten, die für die Glasherstellung erforderlich sind, sind kleiner als 10^{-6} K/s bei guten Glasbildnern wie SiO_2 oder $LiCl{\cdot}7H_2O$, d.h. bei Substanzen, die leicht in den Glaszustand gebracht werden können. Bei den meisten metallischen Gläsern, wie zum Beispiel bei CuZr, müssen die Kühlraten dagegen über 10^6 K/s liegen, um eine Kristallisation zu vermeiden, da diese Substanzen schlechte Glasbildner sind. Wie in Abb. 1.6 angedeutet, kann T_g innerhalb enger Grenzen durch Variation der Kühlrate beeinflußt werden. Kleinere Kühlraten führen zu einer Reduktion der Übergangstemperaturen und einer etwas höheren Glasdichte. Dadurch werden die Struktur und folglich auch die physikalischen Eigenschaften etwas verändert.

„Tempert" man eine Glasprobe, d.h. hält man sie längere Zeit bei einer Temperatur, die nur knapp unterhalb von T_g liegt, so bilden sich in der Probe in vielen Fällen kleine kristalline Bereiche aus.

Aus der kurzen Schilderung der Vorgänge beim Glasübergang und aus der Tatsache, daß die Glasübergangstemperatur und die Glasstruktur von der Kühlrate abhängen, muß man schließen, daß sich Gläser bei Zimmertemperatur nicht im thermischen Gleichgewicht, sondern in einem metastabilen Zustand befinden. Bemerkenswert ist, daß amorphe Festkörper aufgrund ihrer strukturellen Unordnung auch am absoluten Nullpunkt eine hohe Restentropie aufweisen.

Die Vorgänge, die sich am Glasübergang abspielen, sind noch wenig verstanden. Ihre theoretische Beschreibung ist bis jetzt noch recht lückenhaft. Daher wird auf diesem Gebiet intentiv geforscht. Grundsätzlich lassen sich zwei extreme Vorstellungen vom Wesen des Glasübergangs unterscheiden: Entweder behandelt man den Glasübergang wie einen Phasenübergang im thermodynamischen Sinn, oder man sieht darin nur einen kinetischen Übergang, d.h. einen Vorgang, bei dem die molekularen Bewegungen beim raschen Abkühlen eingefroren werden. Nach dem gegenwärtigen Kenntnisstand handelt es sich wohl um eine Mischung aus beiden, so daß neuartige Konzepte und darauf zugeschnittene Experimente erforderlich sind.

1.3 Struktur der Kristalle

Wir werden zunächst die Struktur der Kristalle besprechen und im nächsten Abschnitt auf die Struktur amorpher Festkörper eingehen. Wie bereits erwähnt, wird die experimentelle Bestimmung der Festkörperstruktur im Kap. 2 behandelt.

Eine wichtige Rolle bei der Beschreibung der Kristalle spielen die verschiedenen Symmetrien des atomaren Aufbaus, die eine Unterteilung der Kristalle in Gruppen mit vergleichbarer Struktur ermöglichen. Eine gruppentheoretische Behandlung, die für ein tiefergehendes Verständnis dieser Symmetrieeigenschaften erforderlich wäre, würde den Rahmen dieser Einführung sprengen[2]. Wir diskutieren die Kristallstrukturen hier nur relativ kurz, da sich die für die Bearbeitung spezieller Fragen der Festkörperphysik erforderlichen kristallographischen Kenntnisse leicht durch Selbststudium erarbeiten lassen.

Translationsgitter und Kristallsysteme

Gut ausgebildete, natürlich gewachsene Kristalle zeichnen sich durch ihre regelmäßige, makroskopische Form („Tracht") aus. Vor der tatsächlichen Erforschung ihres atomaren Aufbaus hatte man daher bereits lange Zeit vermutet, daß ihre Bausteine regelmäßig angeordnet sind. 1912 konnte durch *von Laue* und seine Mitarbeiter *Friedrich* und *Knipping* der Beweis hierfür mit Hilfe der Röntgenbeugung an Kristallen erbracht werden.

Ein *idealer Kristall* setzt sich aus identischen, gleich orientierten Struktureinheiten zusammen, die in einer dreidimensionalen, unendlich ausgedehnten, streng periodischen Anordnung aneinandergereiht sind. Die periodisch wiederkehrenden Struktureinheiten, die das *Kristallgitter* aufbauen, bezeichnet man als *Basis*. Die Anzahl der Atome pro Basiseinheit schwankt substanzabhängig in weiten Bereichen: ein oder wenige Atome oder Moleküle in Metallen oder einfachen Dielektrika, 100 Atome in komplexen anorganischen

[2] Eine kurze, übersichtliche Darstellung der gruppentheoretischen Voraussetzungen findet man bei: H. Haken, H.Ch. Wolf, Molekülphysik und Quantenchemie, Springer-Verlag, Heidelberg.

Kristallen und mehr als 10^4 Atome in Proteinkristallen. Ordnen wir jeder derartigen Struktureinheit einen Punkt im Raum zu, so reduzieren wir das Kristallgitter auf ein *Punktgitter*, mit dessen Hilfe wir den Kristallaufbau in einfacher Weise mathematisch beschreiben können. Diese Vorgehensweise ist in Abb. 1.7 für ein zweidimensionales Gitter veranschaulicht. Natürlich kommt es nicht darauf an, welche Koordinate innerhalb der Basis den Punkten des Gitters zugeordnet wird, denn damit ist nur eine Verschiebung des Punktgitters als Ganzes verbunden.

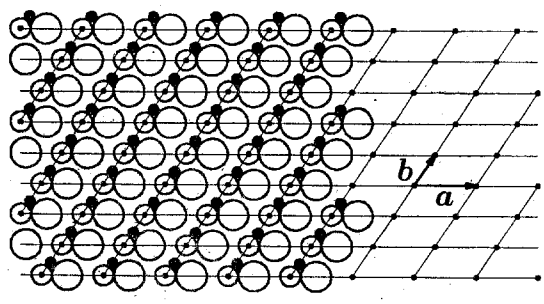


Abb. 1.7: Kristall- und Punktgitter. Die Wahl des Ursprungs der Basis ist für die Konstruktion des Punktgitters ohne Bedeutung. Die Vektoren $\boldsymbol{a}$ und $\boldsymbol{b}$ legen ein schiefwinkliges Koordinatensystem fest, das zur Beschreibung des Punktgitters besonders geeignet ist.

Punktgitter wie auch Kristallstrukturen lassen sich durch Symmetrieelemente charakterisieren, wobei man zwischen *Translations-* und *Punktsymmetrien* unterscheidet. Man findet diese Elemente dadurch, daß man Symmetrieoperationen auf sie wirken läßt, die die Punktgitter bzw. Kristalle in sie selbst überführen. Während bei Translationsoperationen die Kristallstruktur als Ganzes verschoben wird, bleibt bei Punktsymmetrieoperationen mindestens ein Punkt im Raum erhalten.

Zunächst lassen wir den Einfluß der Basis außer acht und betrachten die Punktgitter, die den Kristallen zugrunde liegen, d.h. wir gehen von einer punktförmigen Basis aus. Die auffälligste und wichtigste Symmetrie, die wir noch oft ausnutzen werden, ist die *Translationssymmetrie* der Kristalle. Wir greifen einen beliebigen Punkt im Kristall heraus und betrachten dessen Umgebung, die wir symbolisch mit $\mathcal{U}$ bezeichnen. Ein Blick auf den zweidimensionalen Kristall der Abb. 1.7 macht klar, daß sich diese Umgebung in regelmäßigen Abständen wiederholt. Wir drücken diesen Sachverhalt durch die Gleichung

$$\mathcal{U}(\boldsymbol{r}) = \mathcal{U}(\boldsymbol{r} + \boldsymbol{R}) \tag{1.1}$$

aus, wobei $\boldsymbol{r}$ für einen beliebigen Ortsvektor steht. Die Translation, die von einem Ort zu einem äquivalenten führt, wird durch den *Translations-* oder *Gittervektor* $\boldsymbol{R}$ beschrieben. Für ihn gilt

$$\boldsymbol{R} = n_1\boldsymbol{a} + n_2\boldsymbol{b} + n_3\boldsymbol{c}\,. \tag{1.2}$$

Die Vektoren $\boldsymbol{a}$, $\boldsymbol{b}$ und $\boldsymbol{c}$ definieren ein schiefwinkliges Koordinatensystem, das der Symmetrie des Punktgitters angepaßt ist. Für diese Vektoren findet man in der Literatur verschiedene Bezeichnungen, z.B. *fundamentale Translationsvektoren* oder *Basisvektoren.* Wir benutzen den in der Mathematik gebräuchlichen Ausdruck „Basisvektoren". Um Verwechslungen zu vermeiden, soll hier besonders betont werden, daß die Basisvektoren *nicht* durch die Lage der Basisatome festgelegen werden, sondern nur der Beschreibung des Punktgitters dienen. Die Gittervektoren $\boldsymbol{R}$, die zu äquivalenten Orten des Punktgitters führen, unterliegen der Bedingung, daß ihre Komponenten ganzzahlige Vielfache der Basisvektoren sind. Folglich müssen n_1, n_2 und n_3 ganze Zahlen sein. Die Längen a, b und c der Basisvektoren werden als *Gitterkonstanten* bezeichnet.

Im realen Raum spannen die Basisvektoren, die nicht koplanar gewählt werden dürfen, ein Parallelepiped auf, das durch Aneinanderreihung den Raum lückenlos, aber ohne Überschneidung erfüllen kann. Wie Abb. 1.8 für ein zweidimensionales Gitter veranschaulicht, lassen sich die Basisvektoren für ein Punktgitter nicht eindeutig festlegen. Die dort gewählten Basisvektoren führen zu *Elementarzellen* mit unterschiedlicher Gestalt.

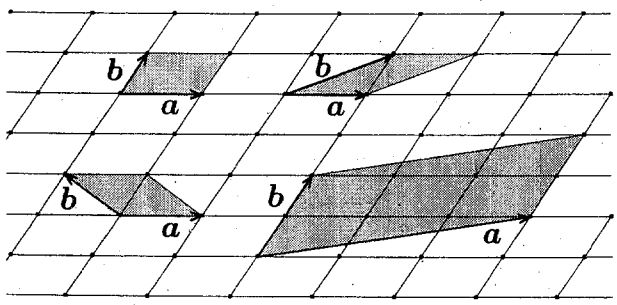


Abb. 1.8: Zweidimensionales Translationsgitter mit dunkel getönten Elementarzellen. Zu den kleinen Zellen gehört jeweils nur ein Gitterpunkt; sie sind daher primitiv. Die große, nicht-primitive Zelle enthält sechs Gitterpunkte.

Besonders hervorzuheben sind die *primitiven Elementarzellen*, die so gewählt sind, daß die Beziehung (1.1) für *alle* äquivalenten Raumpunkte gültig ist. Sie sind die kleinstmöglichen Zellen und enthalten nur *einen* Gitterpunkt, den man als Ursprung der Elementarzelle wählen kann. Wie Abb. 1.8 zeigt, ist auch dann die Wahl der Basisvektoren nicht eindeutig. Enthält eine Elementarzelle mehrere Gitterpunkte, ist also nicht-primitiv, so lassen sich, von einem Ortsvektor $\boldsymbol{r}$ ausgehend, mit der Relation (1.1) nicht *alle* Orte mit gleicher Umgebung aufsuchen.

Nun wenden wir uns den Punktsymmetrieoperationen zu, bei denen wir zwischen Drehung um eine Achse, Spiegelung und Inversion unterscheiden müssen. Wir greifen zunächst die *Drehung* um eine Achse heraus. Selbstverständlich ist, daß die Drehung eines Punktgitters um den Winkel 2π zur Deckungsgleichheit führt. Abhängig vom Punktgitter und der geeigneten Wahl der Drehachse können auch Drehungen um $2\pi/n$ zur Deckung führen. Für zweidimensionale Gitter kann man sich anhand von Abb. 1.9 leicht klar machen, daß aber nur die Werte $n = 1, 2, 3, 4$ und 6 auftreten können, wenn man Translationssymmetrie voraussetzt. Neben Parallelogrammen erlauben nur reguläre Dreiecke, Quadrate und reguläre Sechsecke eine flächendeckende Anordnung ohne Überschneidung. Die Werte $n = 5, 7, 8, \ldots$ sind also ausgeschlossen.

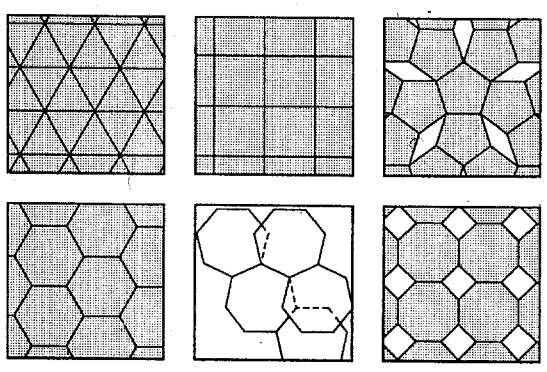


Abb. 1.9: Veranschaulichung der vollständigen Ausfüllung einer Ebene durch reguläre Vielecke. Die Symmetrie der entsprechenden Flächen spiegelt die Drehsymmetrie des verknüpften Punktgitters wider.

Die gleichen Argumente treffen auch auf den dreidimensionalen Raum zu, der mit den entsprechenden regulären Polyedern zu füllen ist. Die *Zähligkeit* n einer Drehachse gibt an, wie oft während einer Drehung des Gitters um 360° Deckungsgleichheit eintritt.

In der sogenannten *internationalen Bezeichnungweise* wird die Zähligkeit n einfach mit 1, 2, 3, 4 bzw. 6 bezeichnet. Konkrete Beispiele werden wir noch im Zusammenhang mit der Diskussion der Symmetrie der Basis kennenlernen.

Eine weitere Punktsymmetrieoperation ist die *Spiegelung* an einer Ebene, bei der nicht nur eine Achse, sondern eine ganze Ebene festgehalten wird. Das Vorhandensein einer Spiegelebene wird durch das Symbol m angedeutet. Spiegelebenen verlaufen entweder parallel zur Drehachse oder stehen darauf senkrecht. Die *Inversion* (auch Paritätstransformation genannt) ist die dritte der Punktsymmetrieoperationen, bei deren Anwendung der Vektor $\boldsymbol{r}$ in $-\boldsymbol{r}$ übergeht. In der internationalen Bezeichnungweise drückt man dies durch $x, y, z \to \overline{x}, \overline{y}, \overline{z}$ oder durch das Symbol $\overline{1}$ aus. Punktgitter besitzen grundsätzlich Inversionssymmetrie, doch kann diese bei realen Kristallen durch fehlende Symmetrie der Basis verloren gehen.

Mit der Drehung eng verwandt ist die *Drehinversion*, die auch als „uneigentliche" Drehung bezeichnet wird. Sie läßt sich als eine Aneinanderreihung von Drehung um $2\pi/n$ und anschließender Inversion darstellen und ist somit eine zusammengesetzte Punktsymmetrieoperation. Die Zähligkeit dieser Symmetrieoperation wird durch die Symbole $\overline{1}$, $\overline{2}$, $\overline{3}$, $\overline{4}$ oder $\overline{6}$ angegeben. An dieser Stelle soll noch erwähnt werden, daß die Punktsymmetrie $\overline{2}$ äquivalent zu einer Spiegelebene ist.

Die Punktgitter und die damit verknüpften Kristallstrukturen lassen sich mit Hilfe der Zähligkeit der Drehachsen klassifizieren. Damit ein Punktgitter einem bestimmten *Kristallsystem* zugeordnet werden kann, muß es einer *minimalen Symmetrieanforderung* genügen. Auf diese Weise werden 7 Kristallsysteme unterschieden, die in Tabelle 1.1 aufgelistet sind. Bei der Zähligkeit besagen die Angaben in Klammern, daß bei orthorhombischen Kristallen zwei 2-zählige und bei den kubischen vier 3-zählige Drehachsen vorausgesetzt werden. Die hier vorgenommene Einteilung mit Hilfe der Drehachsen ist jedoch in dieser Form noch zu eng, denn an die Stelle der einfachen Drehung kann auch die Drehinversion treten. Diese Erweiterung der Definition spielt vor allem dann eine wichtige Rolle, wenn die betrachteten Kristalle keine einfache Basis besitzen. So wird beispielsweise ein Kristall mit einer $\overline{6}$-Achse zu den hexagonalen Kristallen gezählt, obwohl nur eine 3-zählige Drehachse vorliegt.

Tabelle 1.1

Die sieben Kristallsysteme und ihre Basisvektoren. Die Angabe der Zähligkeit bezieht sich auf ihre Dreh- bzw. Drehinversionsachsen (siehe Text).

Kristallsystem	Gitterkonstanten	Winkel	Zähligkeit
triklin	$a \neq b \neq c$	$\alpha \neq \beta \neq \gamma$	1
monoklin	$a \neq b \neq c$	$\alpha = \gamma = 90°, \beta \neq 90°$	2
orthorhombisch	$a \neq b \neq c$	$\alpha = \beta = \gamma = 90°$	2 (zwei)
tetragonal	$a = b \neq c$	$\alpha = \beta = \gamma = 90°$	4
hexagonal	$a = b \neq c$	$\alpha = \beta = 90°, \gamma = 120°$	6
trigonal (rhomboedrisch)	$a = b = c$	$\alpha = \beta = \gamma \neq 90°$	3
kubisch	$a = b = c$	$\alpha = \beta = \gamma = 90°$	3 (vier)

Eine Kristallstruktur *kann* auch noch *weitere* Punktsymmetrieelemente aufweisen, doch sind diese dann für die Einordnung in ein Kristallsystem nicht maßgeblich. Bei der Angabe des Kristallsystems wird immer das System mit der höchsten Zahl von Symmetrieelementen gewählt. In der Reihe kubisch → tetragonal → orthorhombisch → monoklin → triklin besitzt jedes vorhergehende Kristallsystem auch die Symmetrieelemente des folgenden. Entsprechendes gilt für die „Seitenzweige" hexagonal → orthorhombisch, hexagonal → trigonal → monoklin, und kubisch → trigonal.

Viele Elemente und einfache Verbindungen kristallisieren mit kubischer oder hexagonaler Struktur, da diese besonders hohe Symmetrien aufweisen. Dies ist intuitiv verständlich, da einzelne Atome oder einfache Moleküle oft näherungsweise kugelförmig sind und sich zu hochsymmetrischen Gebilden zusammenlagern können. Ist die Basis jedoch aus weniger symmetrischen Molekülen aufgebaut, dann besitzt die Kristallstruktur meist eine geringere Symmetrie. Eine wichtige Rolle spielt dabei oft die Temperatur. Aufgrund der thermischen Drehbewegungen, die Moleküle bei hohen Temperaturen ausführen, sind die Details ihrer Gestalt meist weniger wichtig. Als Faustregel gilt, daß die Hochtemperaturphasen der entsprechenden Festkörper eine höhere Symmetrie aufweisen als die Tieftemperaturphasen. Auf den Einfluß der Basis werden wir im nächsten Abschnitt näher eingehen.

In vielen Fällen bringen primitive Elementarzellen die Symmetrie der Punktgitter nicht voll zum Ausdruck. Ein Beispiel hierfür stellt das in Abb. 1.10 gezeigte zweidimensionale zentrierte Rechteckgitter dar. Von einer primitiven Elementarzelle mit den Basisvektoren $\boldsymbol{a}'$ und $\boldsymbol{b}'$ ausgehend, ließe sich auf ein schiefwinkliges Punktgitter schließen. Läßt man aber eine Elementarzelle mit zwei Gitterpunkten zu, so kann eine Zelle mit wesentlich höherer Symmetrie definiert werden: Der Winkel zwischen den neuen Basisvektoren $\boldsymbol{a}$ und $\boldsymbol{b}$ beträgt dann 90°, und es treten zusätzliche Spiegelebenen auf.

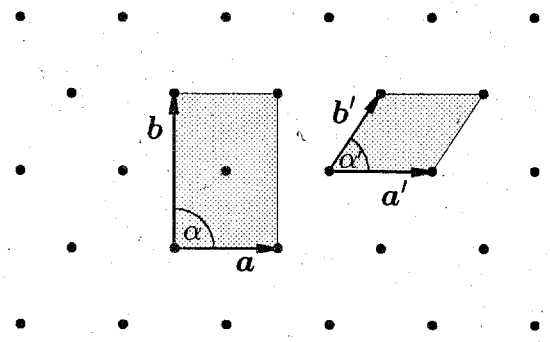


Abb. 1.10: Zentriertes Rechteckgitter. Die rechteckige, nicht-primitive Elementarzelle enthält zwei Gitterpunkte. Die primitive Elementarzelle spiegelt nicht die volle Symmetrie des Punktgitters wider.

Um die Symmetriebeziehungen voll auszuschöpfen, benutzt man daher oft auch bei realen Kristallen nicht-primitive Elementarzellen, deren Gestalt so gewählt wird, daß sie die höchstmögliche Anzahl von Punktsymmetrieelementen enthalten. Dies führt bei dreidimensionalen Punktgittern zu den 14 sogenannten *Bravais-Gittern*, die in der folgenden Abb. 1.11 dargestellt sind. Wie man dem Bild entnehmen kann, weist die Hälfte der Bravais-Gitter eine nicht-primitive Elementarzelle auf. Beim triklinen, trigonalen und hexagonalen Krystallsystem ist die primitive Elementarzelle identisch mit dem entsprechenden Bravais-Gitter. In allen anderen Fällen existieren jeweils mehrere Bravais-Gitter, die demselben Kristallsystem zugeordnet werden.

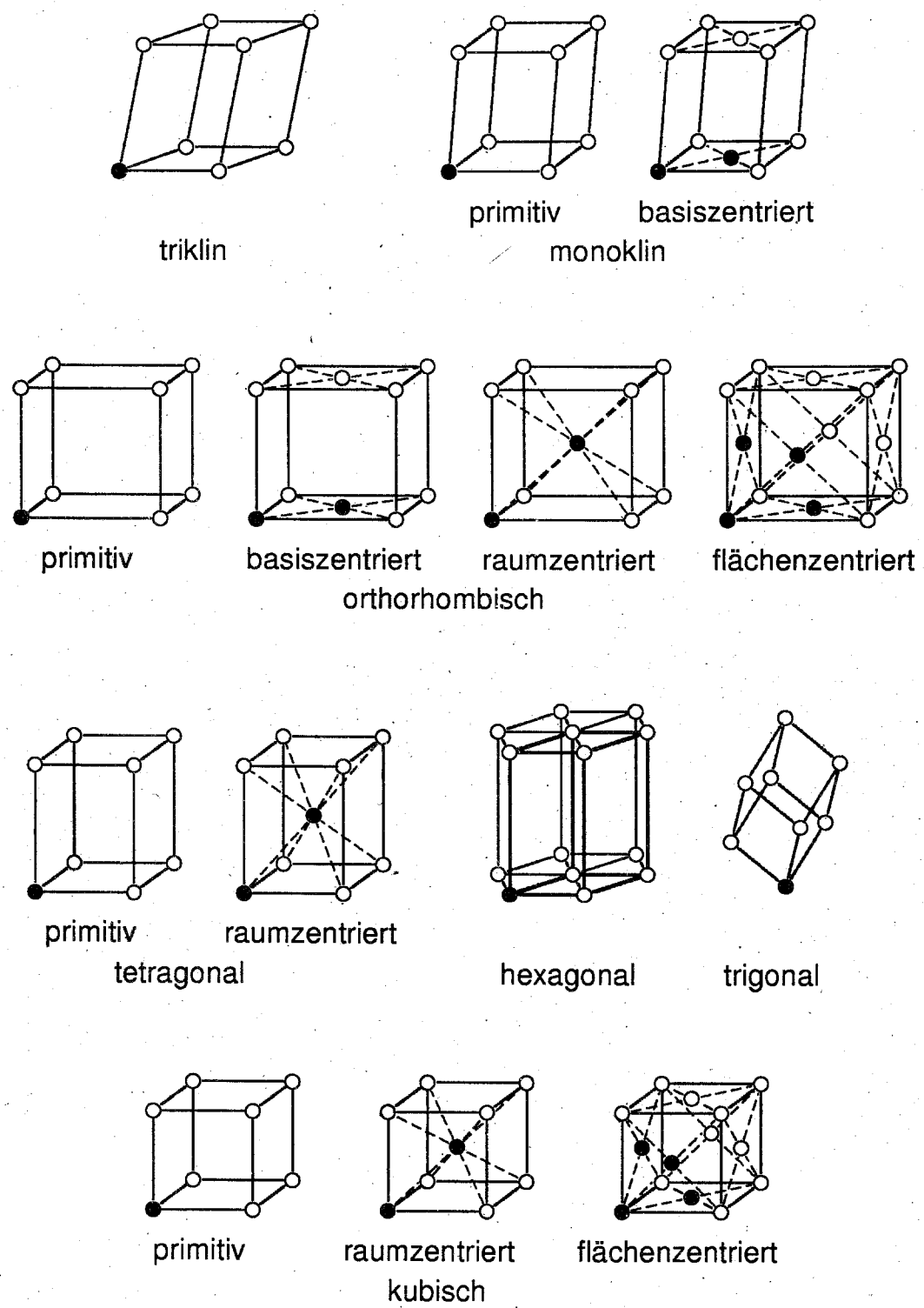


Abb. 1.11: Bravais-Gitter. Die Gitterpunkte, die zu den Elementarzellen gehören, sind durch schwarze Punkte dargestellt. Gitterpunkte, die bereits den benachbarten Elementarzellen zugerechnet werden, sind durch offene Kreise symbolisiert. Die Elementarzelle des hexagonalen Bravais-Gitters ist primitiv und umfaßt nur das stark ausgezogene reguläre Prisma.

Cluster und Quasikristalle

Wie wir im vorhergehenden Abschnitt gesehen haben, ist das Auftreten der 5-zähligen Symmetrie nicht mit der Forderung nach vollständiger Raumerfüllung mit identischen Struktureinheiten verträglich. Dennoch kann man in der Natur auch Festkörper mit dieser Symmetrie finden. So treten bei „sehr kleinen Festkörpern", bei den sogenannten „Clustern" oder Atomhaufen, 5-zählige Symmetrien auf. Bei Edelgasen lagern sich bevorzugt 13 Atome zu ikosaederförmigen[3] Cluster mit dieser Symmetrie zusammen, da diese Atomkonfiguration zu einer besonders dichten Packung und damit zu einer hohen Bindungsenergie führt. Bei Edelgasen wurden Cluster mit fünfzähliger Symmetrie beobachtet, die bis zu 1000 Atome enthielten. Atomhaufen mit ikosaederförmiger Struktur treten nicht nur in Form freier Cluster auf, sondern man findet derartige strukturelle Einheiten auch in unterkühlten Flüssigkeiten und Gläsern, da dort die Forderung nach vollständiger Raumerfüllung mit einem einzigen strukturellen Grundmuster nicht besteht. Periodisch aufgebaute makroskopische Kristalle mit 5-zähliger Symmetrie sind damit aber nicht möglich. Andere Clusterformen findet man bei Metallen. Beispielsweise ist von Natrium bekannt, daß dessen Cluster kugelförmige Gestalt besitzen, wobei das Mitwirken von freien Elektronen bei der Bindung von besonderer Bedeutung ist.

1984 wurde überraschenderweise eine Al-Mn-Legierung entdeckt, die das älteste und fundamentalste Theorem der Kristallographie über die Zähligkeit der erlaubten Drehachsen scheinbar außer Kraft setzte. Später wurden dann noch weitere Legierungen gefunden, die bei rascher Kühlung ihrer Schmelzen zu sogenannten *Quasikristallen* erstarren. Abb. 1.12 zeigt einen Quasikristall, der eindeutig die „verbotene" Gestalt besitzt. Trotz ihrer 5-zähligen Symmetrie erwiesen sich Quasikristalle in Strukturuntersuchungen als wohlgeordnet. Es konnte theoretisch gezeigt werden, daß Quasikristalle nicht wie normale Kristalle aus identischen Struktureinheiten aufgebaut sein können. 5-zählige Symmetrie bei vollständiger Raumerfüllung ist jedoch möglich, wenn man zwei unterschiedliche Typen von Rhomboedern (deformierte Würfel) mit jeweils 3-zähliger Drehachse so zusammenfügt, daß ikosaederförmige Struktureinheiten unterschiedlicher Größe entstehen. Obwohl solche Punktgitter eine definierte Drehsymmetrie aufweisen, fehlt ihnen trotz ihrer ausgeprägten Orientierungsordnung jedoch die für gewöhnliche Kristalle typische Translationsinvarianz[4].

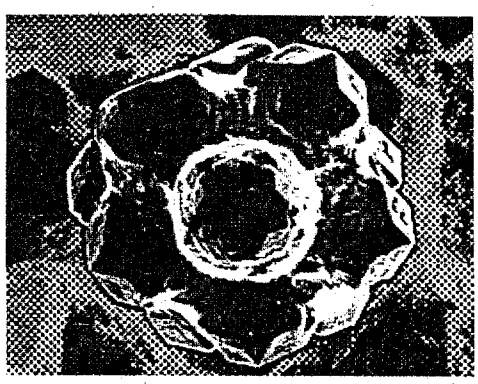


Abb. 1.12: Foto eines ikosaedrischen Quasikristalls mit der Zusammensetzung Al_6CuLi_3. Der etwa 1 mm große Quasikristall besitzt 2-, 3- und 5-zählige Drehachsen. (Aus J.M. Lang et al., J.Crystal Growth **83** (1987) 456).

[3] Ein Ikosaeder besteht aus 20 leicht deformierten Tetraedern, von denen jeweils eine Spitze im Zentrum liegt.

[4] In der Kristallographie wurden die Begriffe „Ordnung" und „Periodizität" meist als weitgehend äquivalent betrachtet, doch beinhalten sie tatsächlich völlig verschiedene Aussagen.

Der prinzipielle Aufbau von Quasikristallen läßt sich sehr schön in zwei Dimensionen veranschaulichen. In Abb. 1.13 ist ein sogenanntes *Penrose-Muster*[5] dargestellt, das sich aus zwei verschiedenen Rhomben mit gleicher Kantenlänge und den Winkeln von 36° bzw. 72° zusammensetzt. Damit läßt sich eine Ebene lückenlos füllen. Durch dunkle Tönung ist angedeutet, daß das Muster regelmäßige Zehnecke mit gleicher Orientierung enthält. Jedes dieser regelmäßigen Zehnecke enthält fünf kleine und fünf große Rhomben. Bei genauerer Betrachtung findet man auch eine gewisse Art von Translationssymmetrie. Dies läßt sich wie folgt zeigen: Man tönt alle Rauten dunkel, die Seiten parallel zu einer vorgegebenen Richtung aufweisen. Dadurch ergeben sich irreguläre Bänder, die sich aber, wie im Bild angedeutet, durch Geraden in regelmäßigen Abständen nähern lassen. Diese Linien entsprechen den Geraden, auf denen im Falle der Kristalle die Gitterpunkte liegen[6]. Wie man sich leicht klarmachen kann, existieren gerade fünf derartige Geradenscharen, die jeweils um 72° gegeneinander geneigt sind. Weiterhin läßt sich zeigen, daß die Zahl der kleinen zur Zahl der großen Rauten durch die irrationale Zahl $(1+\sqrt{5})/2$ („Goldener Schnitt") gegeben ist. Daraus läßt sich sofort der Schluß ziehen, daß das Penrose-Muster nicht durch eine einzige Elementarzelle beschrieben werden kann.

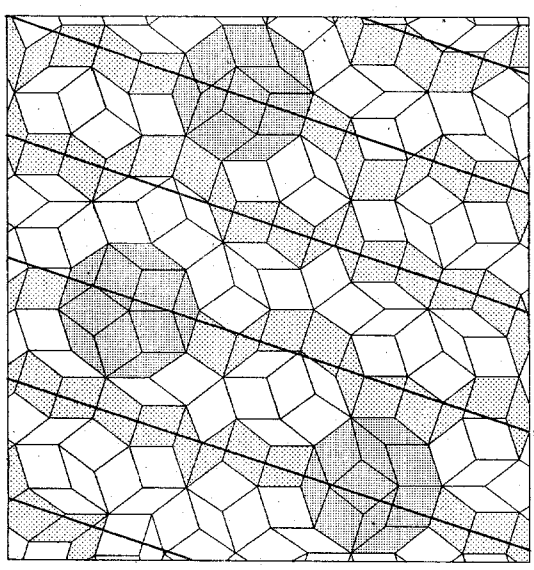


Abb. 1.13: Penrose-Muster. Drei reguläre Zehnecke sind besonders hervorgehoben. Man beachte, daß alle vorhandenen, sich überschneidenden Zehnecke gleich orientiert sind. Die dunklen, irregulären Streifen lassen sich durch die eingezeichneten Geraden repräsentieren.

Obwohl die prinzipielle Vorgehensweise bei der Beschreibung der Struktur von Quasikristallen klar ist, ist immer noch nicht endgültig geklärt, wie die Atome in derartigen Strukturen tatsächlich angeordnet sind, denn die beiden Rhomboeder sind nicht identisch mit Elementarzellen, wie sie von normalen Kristallen her bekannt sind. Es kann jedoch eine periodische Abbildung der ikosaederförmigen Strukturen in einem sechsdimensionalen Raum gefunden werden. Die tatsächliche Struktur läßt sich dann im Prinzip durch Projektion in den dreidimensionalen Raum finden. Dabei ist jedoch zu bedenken, daß Projektionen nur dann eindeutige Ergebnisse liefern, wenn der „Projektionswinkel" eindeutig festgelegt ist. Eine offene Frage ist auch noch, ob reale Quasikristalle von kleinen Löchern durchsetzt sind, die dadurch entstehen, daß ikosaederförmige Gebilde unterschiedlicher Größe so zusammengefügt sind, daß sie den Raum möglichst vollständig erfüllen. Einige

[5] 1974 zum ersten Mal vom Mathematiker Roger Penrose beschrieben.

[6] Bei dreidimensionalen Strukturen treten anstelle der Geraden natürlich Ebenen.

theoretische Überlegungen sagen die Existenz von Löchern mit wohldefinierter Größenverteilung voraus. Erste experimentelle Hinweise darauf konnten jüngst gefunden werden.

Quasikristalle sind gegenwärtig ein Gebiet äußerst intensiver Forschung. Seit ihrer Entdeckung konnten eine ganze Reihe quasikristalliner Materialen gefunden werden. Aus $Al_{65}Cu_{20}Fe_{12}$ lassen sich selbst bei moderaten Kühlraten 1 mm große Kristalle herstellen, die orientierungsabhängige Messung von verschiedenen physikalischen Parametern erlauben. Überraschenderweise weisen Quasikristalle charakteristische Eigenschaften auf, die sich von denen gewöhnlicher Metalle drastisch unterscheiden. Ohne weiter darauf einzugehen, wollen wir hier nur festhalten, daß Quasikristalle erstaunlich kleine elektrische und thermische Leitfähigkeiten besitzen und geringe Reibungs- und Benetzungskoeffizienten aufweisen. Weiterhin sind sie sehr hart und äußerst beständig gegen Korrosion. Der Zusammenhang zwischen den makroskopischen Eigenschaften und ihrem ungewöhnlichen Aufbau ist jedoch erst teilweise verstanden.

Einfluß der Basis, Nomenklatur

Bei unserer bisherigen Diskussion haben wir den Einfluß der Basis auf die Kristallstruktur außer acht gelassen. Solange die Basis nur aus einem kugelförmigen Atom besteht, ist eine tiefergehende Betrachtung auch nicht erforderlich. Die Situation ändert sich jedoch grundlegend bei einer komplizierten Basis. Dieser Sachverhalt läßt sich bereits Abb. 1.7 entnehmen, in der ein unsymmetrisches Molekül als Basis gewählt wurde. Im Gegensatz zum Punktgitter ist das Gitter *mit* Basis nicht mehr inversionssymmetrisch. Um einen Kristall einem bestimmten Kristallsystem zuordnen zu können, müssen Punktgitter *und* Basis den Mindestanforderungen an das Vorhandensein von Symmetrieelementen genügen. Die damit verbundene Problematik macht Abb. 1.14 noch deutlicher, in dem drei kubische Elementarzellen mit hypothetischen Basen dargestellt sind. Ein zusätzliches Atom im Zentrum des Würfels oder ein Molekül mit genügend hoher Symmetrie an den Würfelecken sind mit der Definition des kubischen Kristallsystems mit den vier 3-zähligen Achsen verträglich. Damit nicht vereinbar ist dagegen die zweiatomige Basis auf der rechten Seite, denn mit dieser Basis verliert die Elementarzelle die geforderten 3-zähligen Drehachsen.

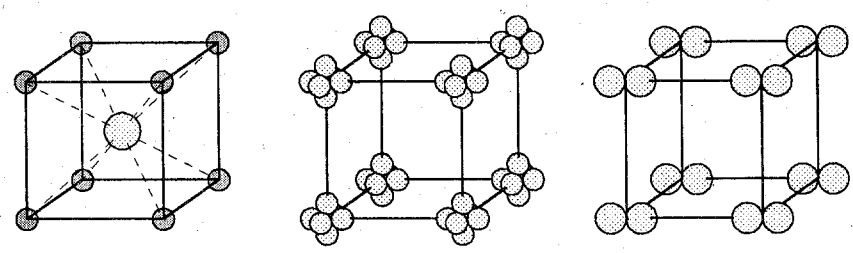


Abb. 1.14: Kubische Elementarzellen mit unterschiedlichen Basen. Die beiden Basiskonfigurationen auf der linken Seite sind mit der kubischen Symmetrie verträglich. Die rechte Anordnung erlaubt dagegen keine Einordnung in das kubische Kristallsystem.

Die Diskussion, welchen Einfluß die Gestalt der Basis hat, führt uns nochmals zu den Symmetrieoperationen zurück. Suchen wir in Abb. 1.14 nach den Symmetrieelementen des linken Würfels, so finden wir neben den vier 3-zähligen auch drei 4-zählige Drehachsen und Spiegelebenen, die bei der Definition des kubischen Kristallsystems nicht vorausgesetzt werden. Natürlich gibt es noch weitere Atomanordnungen mit anderen Symmetrieelementen, die ebenfalls vier 3-zählige Drehachsen aufweisen und somit als Basis bei

kubischen Kristallen auftreten können. Bei einer systematischen Suche finden sich fünf unterschiedliche Anordnungen, so daß man das kubische Kristallsystem in 5 *Kristallklassen* unterteilen kann. Entsprechende Überlegungen gelten bei den übrigen Kristallsystemen. Die mit dem jeweiligen Kristallsystem verträglichen Symmetrieelemente bezeichnet man als *Punktgruppen.* Insgesamt existieren 32 Punktgruppen bzw. Kristallklassen, die sich auf die sieben Kristallsysteme verteilen. Ihre Symmetrieelemente lassen sich mit Hilfe der *stereographischen Projektion*, die vor allem in der Kristallographie Anwendung findet, bildhaft darstellen, doch wollen wir hierauf nicht näher eingehen.

Zur Bezeichnung der Punktgruppen werden zwei verschiedene Notationen verwendet. Bei der *internationalen Bezeichnungsweise*, der wir uns bisher bedient haben, werden Drehachsen bzw. Drehinversionsachsen und Spiegelungen zur Kennzeichnung benutzt. Sie findet vor allem in der Kristallographie Anwendung, während die Bezeichnung nach *Schönflies* meist in der Gruppentheorie und in der Spektroskopie gebraucht wird. Anstelle der Drehinversion tritt bei Schönflies die *Drehspiegelung.* Darüberhinaus erfolgt bei Schönflies die Kennzeichnung mit Hilfe eines Hauptsymbols, das die Zähligkeit der (vertikalen) Drehachsen beinhaltet. Hierbei steht C_n für eine n-zählige vertikale Drehachse, S_n für eine n-zählige Drehspiegelachse, und D_n für eine n-zählige vertikale Drehachse, kombiniert mit einer senkrecht darauf stehenden 2-zähligen Drehachse. Für Spiegelebenen werden die Zusatzsymbole v und h eingeführt, je nachdem ob noch n vertikale (d.h. parallel zur Drehachse verlaufende) Spiegelebenen oder eine horizontale (also senkrecht zur Drehachse stehende) Spiegelebene vorhanden sind. Punktgruppen, die mehr als eine Symmetrieachse mit $n > 2$ besitzen, werden mit T („Tetraeder“) und O („Oktaeder“) abgekürzt. Wie die zugrundeliegenden regulären Polyeder besitzen diese Punktgruppen eine große Zahl von Symmetrieelementen. Wir gehen hier auf die einzelnen Punktgruppen und ihre Bezeichnung nicht näher ein, sondern wollen die Vorgehensweise anhand von drei Beispielen nur andeuten. Für ein eingehenderes Studium verweisen wir auf die einschlägigen Lehrbücher der Kristallographie.

Betrachten wir zunächst ein Wassermolekül, das in Abb. 1.15 dargestellt ist. Offensichtlich besitzt es eine 2-zählige Drehachse und zwei Spiegelebenen, durch die die Drehachse läuft. Die Bezeichnung in der internationalen Notation ist daher $2mm$. In der Schönflies-Notation wird die Punktgruppe C_{2v} genannt. Hierbei steht C_2 für die Drehachse, das Zusatzsymbol v gibt an, daß zusätzlich noch zwei Spiegelebenen vorhanden sind, die die Drehachse enthalten. Um die Darstellung der Punktgruppe zu verallgemeinern, kann man sich ein geometrisches Objekt suchen, das genau die erforderlichen Symmetrieelemente aufweist. Eine mögliche Realisierung ist ein Quader mit hellen und dunklen Flächen, wie er in Abb. 1.15 zu finden ist.

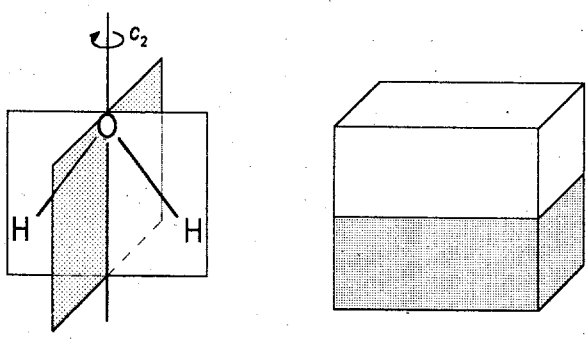


Abb. 1.15: Punktgruppe $2mm$ bzw. C_{2v}. Wassermolekül und Quader weisen die Symmetrieelemente dieser Punktgruppe auf.

Ein weiteres Beispiel ist Butadien in der ebenen, sogenannten trans-Konformation. Der Abbildung 1.16 können wir entnehmen, daß in diesem Fall wieder eine 2-zählige Drehachse vorliegt. Zusätzlich ist eine Spiegelebene vorhanden, die nun aber horizontal, also senkrecht zur Drehachse liegt. Die entsprechende Notation ist daher $2/m$ bzw. C_{2h}, wobei der Bruchstrich bei der internationalen Notation darauf hinweist, daß die Spiegelebene senkrecht zur Drehachse steht. Das äquivalente geometrische Objekt ist wieder ein Quader, mit zwei unterschiedlichen Paaren von Seitenflächen.

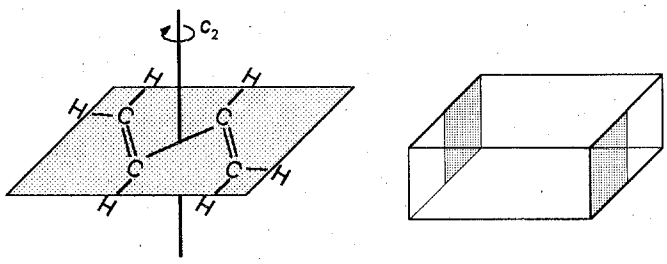


Abb. 1.16: Punktgruppe $2/m$ bzw. C_{2h}. Die Symmetrien von Butadien lassen sich durch den gezeichneten Quader darstellen.

Im letztes Beispiel wollen wir noch auf die Punktgruppe des Phosphorpentafluoridmoleküls eingehen, das in Abb. 1.17 zu sehen ist. Wie man leicht feststellen kann, besitzt das Molekül (und natürlich auch das entsprechende geometrische Analogon) eine 6-zählige Drehinversionsachse und senkrecht hierzu eine 2-zählige Drehachse. Weiterhin findet man noch eine vertikale Spiegelebene. Die entsprechende internationale Bezeichnung ist daher $\bar{6}2m$. In der Schreibweise von Schönflies heißt diese Punktgruppe D_{3h}. Dabei ist die 2-zählige Drehachse bereits in der Abkürzung D_n enthalten.

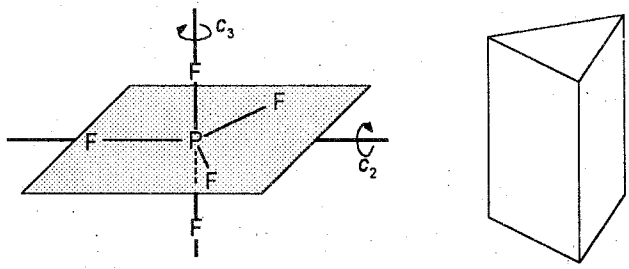


Abb. 1.17: Punktgruppe $\bar{6}2m$ bzw. D_{3h}. Die Symmetrien der dreikantigen Säule spiegeln die des Phosphorpentafluorids wider.

Neben den bereits angesprochenen Symmetrieoperationen lassen sich weitere finden, wenn man Translations- und Punktsymmetrieoperationen miteinander kombiniert. Zu diesen gemischten Symmetrieelementen zählen *Schraubenachsen* und *Gleitspiegelebenen*, deren Bedeutung sich bereits aus der Bezeichnung ablesen läßt. Berücksichtigt man diese zusätzlichen Symmetrieoperationen, so lassen sich die Punktgruppen weiter aufspalten, und es können insgesamt 230 kristallographische *Raumgruppen* gebildet werden. Wir wollen diese keineswegs triviale Thematik aber nicht weiter verfolgen, da sie für unsere folgenden Betrachtungen ohne Bedeutung ist.

Symmetrieoperationen dienen nicht nur der systematischen Klassifizierung der Kristalle, sie sind auch von großem Nutzen bei der Beschreibung der Festkörpereigenschaften, deren mathematische Behandlung sie erheblich vereinfachen. Symmetriebeziehungen sind von überragender Bedeutung in der Theorie der Gitterschwingungen und Elektronenzustände. Sie legen u. a. die Auswahlregel bei optischer Anregung fest und spielen eine entscheidende Rolle bei der Frage nach der Ankopplung von Defekten an ihre Umgebung.

Einfache Kristallgitter

In diesem Abschnitt werden wir uns mit einigen speziellen Kristallstrukturen beschäftigen, wobei wir ganz besonders auf kubische und hexagonale Kristalle eingehen, da diese einfach zu beschreiben und sehr häufig anzutreffen sind. Wir werden dabei wiederholt auf die Zahl der *nächsten Nachbarn* zu sprechen kommen, die man auch als *Koordinationszahl* bezeichnet. Dabei ist zu beachten, daß die Zahl der nächstgelegenen Gitterpunkte nur bei einfacher Basis mit der Zahl der physikalisch gleichwertigen Nachbarn übereinstimmt.

Kubische Kristalle

Wir entnehmen Abb. 1.11 mit den Bravais-Gittern, daß drei unterschiedliche kubische Punktgitter existieren: das primitive, das raum- und das flächenzentrierte. Sie unterscheiden sich durch die Zahl und Anordnung der Gitterpunkte in der Elementarzelle und durch ihr *Packungsverhältnis*. So bezeichnet man den Bruchteil des Raumes, der von identischen, sich berührenden Kugeln auf Atompositionen ausgefüllt wird. Die Gitterkonstante a ist jeweils durch die Kantenlänge der betrachteten würfelförmigen Elementarzellen gegeben.

Die Elementarzelle des *kubisch primitiven* Gitters (englische Abkürzung sc für „simple cubic") wurde bereits in Abb. 1.11 abgebildet. Das Bezugsatom im Koordinatenursprung (0,0,0) ist von sechs *nächsten Nachbarn* im Abstand a und zwölf *übernächsten Nachbarn* im Abstand $a\sqrt{2}$ umgeben. In dieser Konfiguration beträgt das oben erwähnte Packungsverhältnis 0,52, d.h. nur etwa die Hälfte des Raumes ist in dieser Anordnung von harten Kugeln ausgefüllt. Kubisch primitive Kristallgitter mit einatomiger Basis kommen in der Natur nicht vor, jedoch findet man sie oft bei Kristallen mit mehratomiger Basis.

Ein bekanntes Beispiel für ein kubisch primitives Gitter mit *zweiatomiger* Basis ist die *Cäsiumchloridstruktur*, die in Abb. 1.18 dargestellt ist. Die beiden Atome oder Ionen der Basis nehmen die Plätze (0,0,0) und $(\frac{1}{2},\frac{1}{2},\frac{1}{2})$ in der Elementarzelle ein, wodurch eine wesentlich höhere Raumerfüllung als bei einatomiger Basis möglich ist. Die Koordinationszahl, d.h. die Zahl der gleichartigen nächsten Nachbarn, ist acht. Diese Struktur findet man unter anderem bei einfach aufgebauten Ionenkristallen, wenn deren Ionen vergleichbare Radien aufweisen. Dies gilt insbesondere für Cäsiumchlorid mit den Radien $r_{Cs^+} = 1{,}65\,\text{Å}$ und $r_{Cl^-} = 1{,}81\,\text{Å}$.

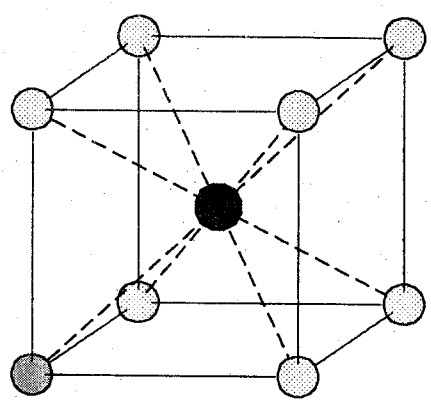


Abb. 1.18: Cäsiumchloridstruktur. Die Elementarzelle des kubisch primitiven Gitters enthält die beiden Atome bzw. Ionen auf den Plätzen (0,0,0) und $(\frac{1}{2},\frac{1}{2},\frac{1}{2})$. Atome bzw. Ionen benachbarten Elementarzellen sind etwas heller gezeichnet.

Ob tatsächlich eine kubische Struktur auftreten kann, hängt in erster Line von den Bindungsverhältnissen in der Basis und weniger stark von der Komplexität der Basis ab. Wie Abb. 1.19 zeigt, kristallisiert das relativ komplizierte Molekül $CaTiO_3$ in der kubischen

Perowskitstruktur. In der kubisch primitiven Elementarzelle sitzen neben dem Titanion am Ursprung noch das Kalziumion im Zentrum des Würfels und die drei Sauerstoffionen auf den Plätzen $(\frac{1}{2},0,0)$, $(0,\frac{1}{2},0)$ und $(0,0,\frac{1}{2})$. Mineralien mit Perowskitstruktur treten sehr häufig auf und sind ein wesentlicher Bestandteil der Erderinde. Im Zusammenhang mit der Diskussion der Eigenschaften der Ferroelektrika werden wir in Kap. 11.2 auf diese Struktur zurückkommen.

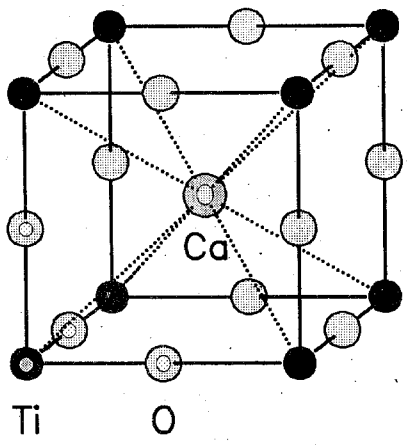


Abb. 1.19: Perowskitstruktur. Die Basis besteht aus dem Titanion im Ursprung der Zelle, dem Kalziumion im Zentrum des Würfels und den drei Sauerstoffionen auf den Würfelkanten. Diese Atome sind durch kleine zusätzliche Kreise markiert.

Kubisch raumzentrierte Gitter kommen in der Natur sehr häufig vor. So besitzen die Alkalimetalle, α-Eisen, Chrom, Barium, viele andere Metalle und Verbindungen dieses Gitter, für das oft die englische Abkürzung bcc („body centered cubic") benutzt wird. Wie der Darstellung der Bravais-Gitter in Abb. 1.11 zu entnehmen ist, enthält die Zelle zwei Gitterpunkte mit den Koordinaten (0,0,0) und $(\frac{1}{2},\frac{1}{2},\frac{1}{2})$. Die acht nächsten Nachbarn sind im Abstand $a\sqrt{3}/2$, die sechs übernächsten im Abstand a zu finden. Mit 0,68 ist das Packungsverhältnis wesentlich höher als beim primitiven Gitter. Die Abstände zwischen den nächsten und den übernächsten Nachbarn unterscheiden sich nur geringfügig.

Natürlich kann man anstelle der diskutierten kubischen Elementarzelle mit den beiden Gitterpunkten auch eine primitive wählen, die gemäß der Definition nur einen Gitterpunkt enthält. Statt des Würfels repräsentiert dann ein Rhomboeder mit geringerer Symmetrie, der Kantenlänge $a' = a\sqrt{3}/2$ und dem Winkel $\alpha = 109° 28'$ das Gitter. In Abb. 1.20 ist sowohl die primitive als auch die nicht-primitive Elementarzelle gezeichnet.

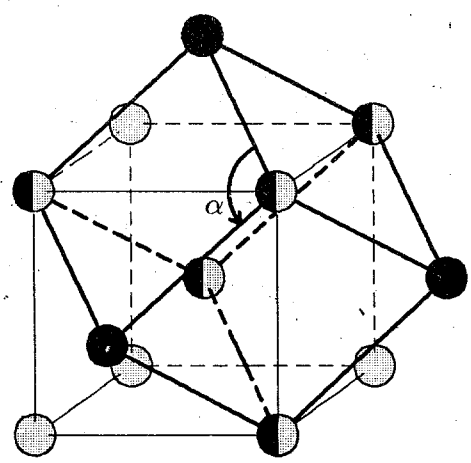


Abb. 1.20: Kubisch raumzentriertes Gitter. Die primitive Elementarzelle ist dick, die nicht-primitive kubische Zelle dünn gezeichnet. Der Winkel α zwischen den Achsen der primitiven Zelle beträgt $109° 28'$. Die Gitterpunkte, die zu beiden Elementarzellen gehören, sind zur Hälfte dunkel gezeichnet.

Das *kubisch flächenzentrierte Gitter* ist das letzte der drei kubischen Gitter und wird englisch mit fcc („face centered cubic") abgekürzt. Edelgaskristalle und viele Metalle, wie Kupfer, Silber, Gold oder Nickel besitzen ein kubisch flächenzentriertes Gitter mit einatomiger Basis. Die kubische Elementarzelle (vgl. Abb. 1.11) enthält vier Gitterpunkte mit den Koordinaten (0,0,0), $(\frac{1}{2},0,\frac{1}{2})$, $(\frac{1}{2},\frac{1}{2},0)$ und $(0,\frac{1}{2},\frac{1}{2})$. Die zwölf nächsten Nachbarn befinden sich im Abstand $a/\sqrt{2}$, die sechs übernächsten im Abstand a vom Aufatom. Zwölf ist die größtmögliche Anzahl von nächsten Nachbarn, die mit gleich großen, harten Kugeln realisiert werden kann. Kubisch flächenzentrierte Gitter weisen die dichteste Packung von Kugeln mit dem Packungsverhältnis 0,74 auf. In Abb. 1.21 sind die kubische und die primitive Elementarzelle dieses Gittertyps dargestellt. Wie beim kubisch raumzentrierten Gitter ist auch hier die primitive Elementarzelle rhomboedrisch. Sie weist eine Kantenlänge $a' = a/\sqrt{2}$ auf, der Winkel an den spitzen Ecken der Rauten beträgt 60°.

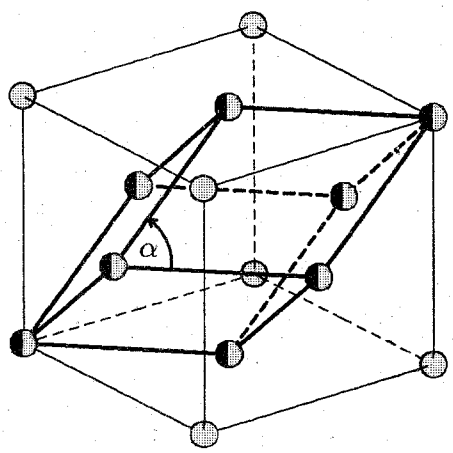


Abb. 1.21: Kubisch flächenzentriertes Gitter. Die primitive Elementarzelle ist dick, die nicht-primitive Zelle dünn gezeichnet. Zwischen den Achsen der primitiven Zelle tritt der Winkel $\alpha = 60°$ auf. Die Gitterpunkte, die zu beiden Elementarzellen gehören, sind zur Hälfte dunkel gezeichnet.

Ein bekanntes Beispiel für eine kubisch flächenzentrierte Struktur mit einer *zweiatomigen* Basis ist die *Natriumchlorid-* oder *Kochsalzstruktur* (vgl. Abb. 1.22), in der viele Alkalihalogenide kristallisieren. Die beiden unterschiedlichen Ionen auf den Plätzen (0,0,0) und $(\frac{1}{2},\frac{1}{2},\frac{1}{2})$ bilden die Basis. Man kann daher von zwei Untergittern sprechen, die um die halbe Raumdiagonale des Würfels gegeneinander verschoben sind. Diese Struktur mit der Koordinationszahl sechs tritt bevorzugt dann auf, wenn sich die Radien der beteiligten Ionen stark unterscheiden. Dies ist mit $r_{Na^+} = 0{,}98\,\text{Å}$ und $r_{Cl^-} = 1{,}81\,\text{Å}$ beim Kochsalz der Fall.

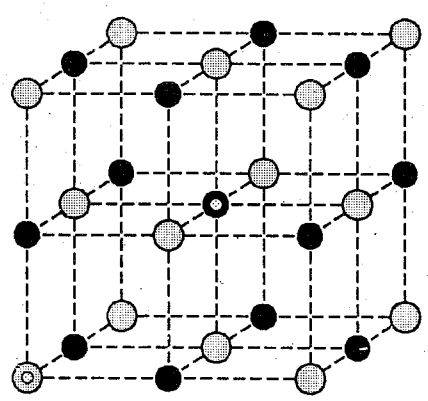


Abb. 1.22: Natriumchloridstruktur. Die besonders gekennzeichneten Ionen auf den Plätzen (0,0,0) und $(\frac{1}{2},\frac{1}{2},\frac{1}{2})$ bilden die Basis.

Weitere wichtige Beispiele für kubisch flächenzentrierte Gitter sind die von *Zinkblende* und *Diamant*, die beide in Abb. 1.23 dargestellt sind. Bei Zinkblende (ZnS) besteht die Basis aus einem Zink- und einem Schwefelatom mit den Koordinaten (0,0,0) und $(\frac{1}{4}, \frac{1}{4}, \frac{1}{4})$. Jedes Atom sitzt im Zentrum eines regulären Tetraeders, das von Atomen der anderen Sorte gebildet wird. Die Koordinationszahl ist nur vier, das Packungsverhältnis nur 0,34. Dennoch tritt diese Kristallstruktur häufig auf, wenn die Atome über gerichtete kovalente Bindungen (vgl. Kap. 3.4) mit vier Nachbarn tetraederförmig verbunden sind. Als weitere Beispiele für die Zinkblendestruktur seien die technologisch wichtigen Verbindungshalbleiter GaAs, GaP, CdS, InP und InSb erwähnt.

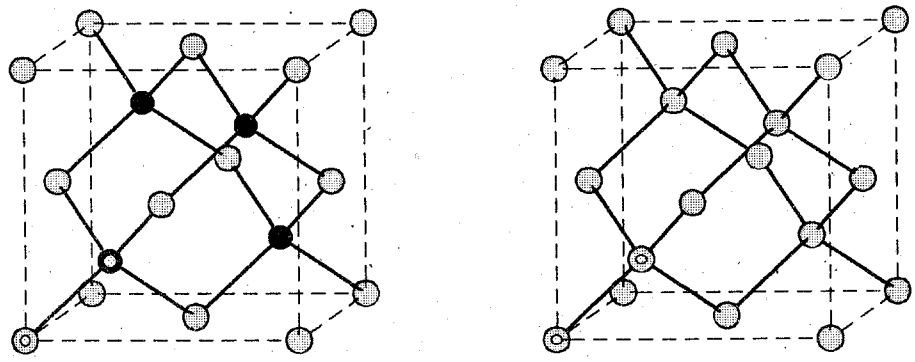


Abb. 1.23: Zinkblende- und Diamantstruktur. Die Basiseinheiten sind durch kleine Kreise besonders gekennzeichnet. Bei der Zinkblende sind die unterschiedlichen Atomsorten durch schwarze bzw. graue Kreise angedeutet. Obwohl beim Diamantgitter nur gleichartige Atome auftreten, ist die Basis ebenfalls zweiatomig.

Viele Elemente wie Silizium, Germanium, graues Zinn oder Kohlenstoff in Form von Diamant kristallisieren ebenfalls in dieser Struktur. Obwohl alle Plätze von gleichartigen Atomen besetzt werden, enthält die primitive Elementarzelle dieser Kristalle zwei Atome, d.h. bei dieser Struktur läßt sich keine Elementarzelle finden, die nur ein einziges Atom enthält. Die Diamantstruktur ist ein gutes Beispiel dafür, daß bei der Definition des kubischen Kristallsystems die 3-zähligen Drehachsen die entscheidenden Punktsymmetrien sind und nicht die scheinbar erforderlichen 4-zähligen Achsen.

Hexagonal dichteste Kugelpackung

Neben dem soeben diskutierten kubisch flächenzentrierten Gitter existiert noch eine weitere Anordnung von harten Kugeln, die ebenfalls zum höchstmöglichen Packungsverhältnis von 0,74 führt. In Abb. 1.24 sind die beiden Konstruktionsprinzipien erläutert. Legt man die Kugeln auf einer Ebene möglichst dicht aneinander, so erhält man die erste Lage A mit einer hexagonalen Kugelanordnung. Die Kugeln der darüberliegenden Lage B werden in die vorhandenen Mulden gelegt, die aber nur zur Hälfte in Anspruch genommen werden. Für die dritte Lage bestehen nun zwei Möglichkeiten der Anordnung. Wie in Abb. 1.24 angedeutet, können wir beim Aufbau dieser Lage die Mulden (a) oder (b) benutzen, die jedoch bezüglich der ersten Lage A nicht äquivalent sind. Wählen wir die Mulden (a), dann ist die Lage C gegenüber den Lagen A und B verschoben. Fahren wir mit dieser Stapelfolge ABC ABC ABC .... fort, so entsteht das bereits diskutierte kubisch flächenzentrierte Gitter. Die Ebenen stellen einen Schnitt durch den kubischen Kristall

dar, der durch die Gitterpunkte (0,0,0), (1,1,0) und (0,1,1) festgelegt wird. Wie wir im nächsten Kapitel erfahren werden, ist {111} die Bezeichnung für diese Ebenen.

Wählen wir beim Aufbau der dritten Lage dagegen die Mulden (b), dann ist diese Lage mit der ersten identisch. Die Stapelfolge AB AB AB ... führt zur *hexagonal dichtesten Kugelpackung* (englisch hcp = hexagonal close packed). Viele Metalle wie Beryllium, Magnesium, Zirkonium, Kadmium, Titan oder Kobalt weisen diese Struktur auf. Abstand und Anzahl der nächsten und übernächsten Nachbarn sind so groß wie beim kubisch flächenzentrierten Gitter. Erst bei den noch weiter entfernten Nachbarn zeigen sich Unterschiede. Deshalb unterscheiden sich die Bindungsenergien der beiden Strukturen auch kaum voneinander. Dies wird bei Edelgaskristallen deutlich, bei denen je nach Präparationsbedingungen beide Strukturen möglich sind. Darüberhinaus sind in diesen Systemen Phasenübergänge und Mischphasen beobachtbar.

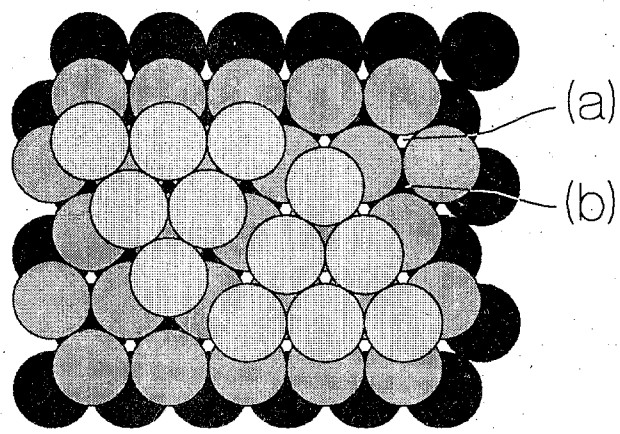


Abb. 1.24: Dichteste Kugelpackung. Die unterste Lage A ist schwarz, die mittlere Lage B dunkelgrau gezeichnet. Bei den Mulden (a) tritt ein kleiner Kanal auf, durch den man blicken kann. Unter den Mulden (b) liegen die Mittelpunkte der Kugeln der Lage A. Die Atome der dritten Lage (hellgrau) liegen links in den Mulden (a), rechts in (b). Die Stapelfolge ABC (links) führt zur kubischen, die Folge AB (rechts) zur hexagonalen Struktur.

Der Vollständigkeit halber ist in Abb. 1.25 die Elementarzelle der Kristalle mit hexagonal dichtester Kugelpackung dargestellt. Das dick gezeichnete Prisma der primitiven Elementarzelle weist den Winkel $\alpha = 120°$ und die Achsen $a = b$, $c = a\sqrt{8/3} \approx 1{,}63\,a$ auf. Die Atome bzw. Moleküle der Basis nehmen die Plätze (0,0,0) und $(\frac{2}{3}, \frac{1}{3}, \frac{1}{2})$ ein. Die 6-zählige Symmetrie wird erst nach dem Zusammenfügen von drei primitiven Zellen deutlich sichtbar.

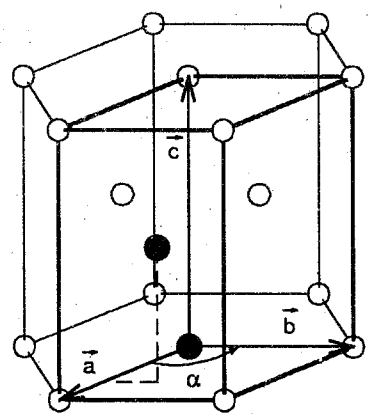


Abb. 1.25: Elementarzelle der hexagonal dichtesten Kugelpackung. Die dick gezeichnete, primitive Zelle enthält die beiden mit Kreisen markierten Atome. Die Symmetrie der Struktur kommt erst durch das Zusammenfügen von drei primitiven Elementarzellen klar zum Ausdruck.

Wigner-Seitz-Zelle

Die bisher betrachteten Elementarzellen haben die Form von Parallelepipeden, deren Kanten durch die Basisvektoren festgelegt werden. Diese Wahl erweist sich bei der Darstellung der Gitterperiodizität und bei der noch zu diskutierenden räumlichen Fourier-Zerlegung (vgl. Kap. 2.3) der Kristallstruktur als vorteilhaft. Bei der Behandlung von anderen Kristalleigenschaften, wie z.B. bei der Beschreibung der lokalen Elektronendichte, ist es aber oft sinnvoller das Zentrum der Basis als Aufpunkt der Elementarzelle zu wählen. Als Elementarzellen eignet sich dann ein symmetriegerechtes Polyeder, dessen Zentrum mit dem Gitterpunkt zusammenfällt.

Das Konstruktionsprinzip ist in Abb. 1.26 für den zweidimensionalen Fall angedeutet. Der herausgegriffene Gitterpunkt wird zunächst mit den benachbarten Gitterpunkten verbunden. Anschließend werden die Mittelsenkrechten auf diesen Strecken errichtet. Bei dreidimensionalen Gittern konstruiert man die entsprechende Ebenen. Die gesuchte Elementarzelle ist dann das Polyeder mit dem kleinsten Volumen, das von diesen Ebenen gebildet wird. Diese Zelle heißt *Wigner-Seitz-Zelle* und enthält alle Punkte, die dem *herausgegriffenen* Gitterpunkt am nächsten sind. Die so konstruierte Elementarzelle besitzt die volle Gittersymmetrie. Da sie nur einen Gitterpunkt aufweist, handelt sich um eine primitive Elementarzelle. Im Falle des schiefwinkligen Gitters der Abb. 1.26 ist die so erzeugte Zelle ein unregelmäßiges Sechseck mit zweizähliger Drehachse. In der folgenden Abb. 1.27 ist die Wigner-Seitz-Zelle des kubisch raumzentrierten Gitters als Beispiel für eine dreidimensionale Wigner-Seitz-Zelle dargestellt. Die Sechsecke sind gemeinsame Flächen mit den Zellen, die den nächsten, die Quadrate, die den übernächsten Nachbarn als Ursprung haben.

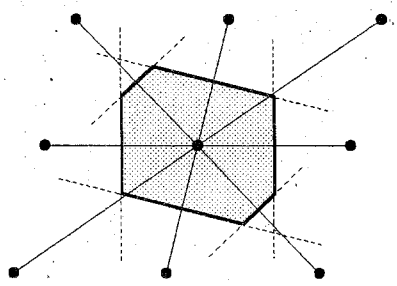


Abb. 1.26: Darstellung des Konstruktionsprinzips der Wigner-Seitz-Zelle anhand eines schiefwinkligen ebenen Gitters. Zwischen den verbundenen Gitterpunkten werden jeweils Mittelsenkrechte errichtet.

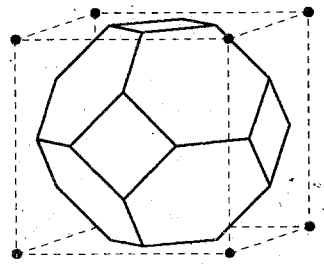


Abb. 1.27: Wigner-Seitz-Zelle des kubisch raumzentrierten Gitters. Die üblicherweise benutzte nicht-primitive Elementarzelle ist gestrichelt gezeichnet; der Gitterpunkt im Zentrum der Zelle wurde weggelassen.

1.4 Struktur amorpher Festkörper

In einem idealen Kristall ist die Lage aller Atome exakt vorgegeben, da nur eine einzige, wohldefinierte Konfiguration existiert. Die Koordinaten aller Atome können angegeben werden, wenn Größe und Inhalt einer Elementarzelle und die Gitterstruktur bekannt sind. In amorphen Festkörpern dagegen ist die Lage der einzelnen Atome wegen der fehlenden

Periodizität nicht so einfach festzulegen. Eigentlich müßte man in diesem Fall für jede Probe eine geeignete Liste der Atompositionen erstellen. Da dies nicht möglich ist, kann die Struktur nur mit statistischen Methoden in einer wesentlich weniger detaillierten Form als in Kristallen beschrieben werden.

Wichtige grundlegende Information enthält die *Teilchenzahldichte* $n(\boldsymbol{r})$, die über die Beziehung $\delta N(\boldsymbol{r}) = n(\boldsymbol{r})\,\mathrm{d}V$ definiert ist, wobei $\delta N(\boldsymbol{r})$ angibt, wie viele Teilchenmittelpunkte am Ort $\boldsymbol{r}$ im Volumenelement $\mathrm{d}V$ vorhanden sind. Eine grobe Angabe, die bereits einige Rückschlüsse auf die strukturellen Gegebenheiten erlaubt, läßt sich mit Hilfe des Mittelwerts der Teilchenzahldichte $n_0 = \langle n(\boldsymbol{r})\rangle = N/V$ machen, wobei N die Anzahl der Atome und V das Volumen des amorphen Festkörpers ist. Verglichen mit Kristallen sind amorphe Festkörper weniger dicht gepackt: Mittlere Teilchenzahldichte und mittlere Massendichte sind typischerweise um 1 – 10% kleiner.

Paarkorrelationsfunktion

Information über die lokale Anordnung der Atome läßt sich der *Paarkorrelationsfunktion* oder *Zweiteilchen-Korrelationsfunktion* entnehmen. Bei der Definition wird nicht zwischen Kristalle und amorphe Festkörper unterschieden, doch spielt sie nur für das Verständnis der Struktur amorpher Festkörper eine zentrale Rolle. Sie kann mit Beugungsexperimenten (vgl. Kap. 2.5) bestimmt und bei Kenntnis der Wechselwirkungspotentiale der Atome im Prinzip berechnet werden.

Nehmen wir der Einfachheit halber an, daß der betrachtete Festkörper nur aus *einer* Atomsorte aufgebaut ist, dann gibt die Paarkorrelationsfunktion Auskunft über die Wahrscheinlichkeit im Volumenelement $\mathrm{d}V'$ am Ort $\tilde{\boldsymbol{r}}'$ ein Teilchen zu finden, wenn im Volumenelement $\mathrm{d}V$ am Ort $\tilde{\boldsymbol{r}}$ bereits ein Teilchen vorhanden ist[7]. Da wir nicht an einer speziellen Konfiguration mit einem ausgewählten Atom am Ort $\tilde{\boldsymbol{r}}$ interessiert sind, kann eine Mittelung über alle Volumenelemente $\mathrm{d}V$ durchgeführt werden. Hierbei wird sinnvollerweise der Vektor $\boldsymbol{r} = \tilde{\boldsymbol{r}}' - \tilde{\boldsymbol{r}}$ festgehalten, der die relative Lage der Volumenelemente bezüglich des gewählten Aufpunkts bestimmt. Formal kann die Paarkorrelationsfunktion $g(\tilde{\boldsymbol{r}}, \tilde{\boldsymbol{r}}')$ mit Hilfe des Erwartungswerts der Teilchenzahldichten an den Orten $\tilde{\boldsymbol{r}}$ und $\tilde{\boldsymbol{r}}'$ definiert werden:

$$g(\tilde{\boldsymbol{r}}, \tilde{\boldsymbol{r}}') = g(\boldsymbol{r}) = \frac{1}{n_0^2} \Big\langle n(\tilde{\boldsymbol{r}})\, n(\tilde{\boldsymbol{r}} + \boldsymbol{r}) \Big\rangle \,. \tag{1.3}$$

Die Paarkorrelationsfunktion wird mit Hilfe der mittleren Teilchenzahldichte n_0 für große Abstände auf den Wert Eins normiert. Weiter ist zu bemerken, daß sich definitionsgemäß am Ort $\tilde{\boldsymbol{r}}$ bereits ein Teilchen befindet. Die Wahrscheinlichkeit, an diesem Ort ein weiteres Teilchen zu finden, muß verschwinden, da zwei Atome sich nicht gleichzeitig am gleichen Ort aufhalten können[8].

[7] Es ist zu beachten, daß das Auftreten von Werten größer als eins in der Paarkorrelationsfunktion nicht in Widerspruch mit der Interpretation als Wahrscheinlichkeit steht, da hierfür das Integral über das betrachtete Volumenelement entscheidend ist.

[8] Strenggenommen muß zwischen der Paar- und der Autokorrelationsfunktion unterschieden werden. Im ersten Fall wird das Teilchen am Aufpunkt nicht mitgerechnet, im zweiten Fall aber berücksichtigt. Um diesem Sachverhalt Rechnung zu tragen, müßte bei der formalen Definition der Paarkorrelationsfunktion der Ausdruck $g(\boldsymbol{r}) = [\langle n(\tilde{\boldsymbol{r}})\, n(\tilde{\boldsymbol{r}} + \boldsymbol{r})\rangle - n_0\, \delta(\boldsymbol{r})]/n_0^2$ anstelle von Gleichung (1.3) treten. Sieht man von Normierungsfaktoren ab, so kann daher (1.3) auch zur Definition der Autokorrelationsfunktion $R(\boldsymbol{r})$ herangezogen werden. Im Falle von punktförmigen Teilchen läßt sich die Autokorrelationsfunktion daher als die Summe einer Deltafunktion bei $r = 0$ und der Paarkorrelationsfunktion darstellen. In Beugungsexperimenten, auf die wir im nächsten Kapitel eingehen, wird die Autokorrelationsfunktion und nicht die Paarkorrelationsfunktion bestimmt.

Wenn auch die Umgebung jedes einzelnen Atoms unterschiedlich aussieht, so hängt bei einem ausreichend großen, statistisch homogenen System die Paarkorrelationsfunktion weder von der Wahl von $\tilde{\boldsymbol{r}}$ noch von der Richtung ab. In dem angesprochenen Fall eines amorphen Festkörpers mit gleichartigen Atomen spiegelt daher die Paarkorrelationsfunktion die Teilchenzahldichte wider, die man im Mittel im Abstand $r = |\boldsymbol{r}|$ von einem gewählten Bezugsatom antrifft, d.h.

$$g(r) = \frac{n(r)}{n_0} \,. \tag{1.4}$$

Um uns mit diesem Konzept vertraut zu machen, studieren wir zunächst den besonders einfachen Fall der Korrelationsfunktion eines eindimensionalen Systems. Wir verteilen N identische harte Kugeln mit dem Durchmesser a über eine Strecke L rein zufällig. Der Verlauf der Korrelationsfunktion hängt entscheidend vom Parameter $\ell = (L - Na)/N$ ab, der den mittleren Zwischenraum zwischen den Kugeln angibt. Wir greifen willkürlich eine Kugel heraus und stellen fest, in welchem Abstand sich die Mittelpunkte der anderen Kugeln befinden. Mitteln wir über viele Messungen - wobei wir unterschiedliche Kugeln herausgreifen - so erhalten wir nach Normierung mit der mittleren Teilchenzahldichte n_0 die Paarkorrelationsfunktion, die in Abb. 1.28 für verschiedene Fälle dargestellt ist.

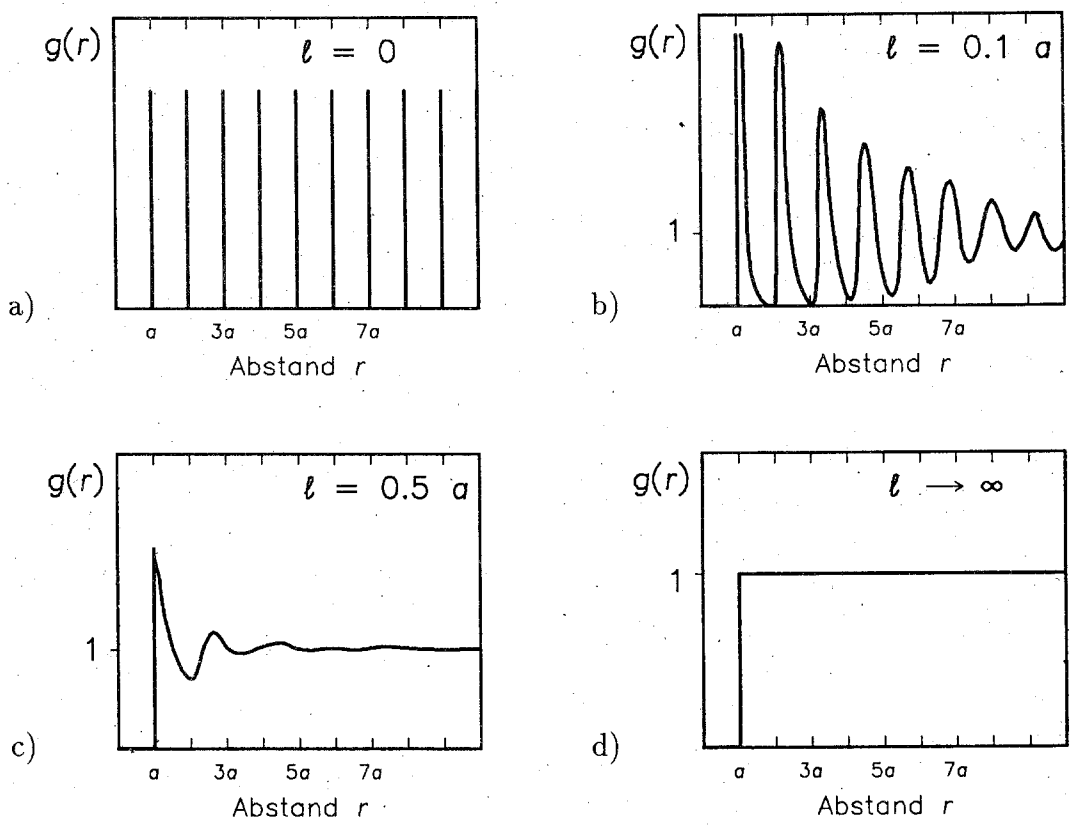


Abb. 1.28: Paarkorrelationsfunktion eines linearen Systems mit statistisch verteilten harten Kugeln. (Nach F. Zernike und J.A. Prins, Z. Physik **41** (1927) 184).
a) „Eindimensionaler Kristall" mit ($\ell = 0$),
b) „Flüssigkeit" mit $\ell = a/10$,
c) „Flüssigkeit" mit $\ell = a/2$,
d) „verdünntes Gas" ($\ell \to \infty$).

Für die Grenzfälle $\ell \to 0$ und $\ell \to L/N$ läßt sich die Paarkorrelationsfunktion unmittelbar angeben. Im ersten Fall tritt überhaupt kein Zwischenraum auf. Die Kugeln liegen dicht gepackt wie in einem „eindimensionalen Kristall" nebeneinander. Die Paarkorrelationsfunktion muß periodisch sein und nimmt an den Stellen $x = pa$ mit $p = 1, 2 \ldots$ die Form einer Deltafunktion an. Im zweiten Fall ist der mittlere Abstand zwischen den Kugeln groß im Verhältnis zu deren Durchmesser. Man könnte daher von einem „eindimensionalen, verdünnten Gas" sprechen. In diesem Fall ist die Wahrscheinlichkeit, in einem beliebigen Streckenelement dr einen Kugelmittelpunkt zu finden, relativ klein und unabhängig vom Abstand r zur herausgegriffenen Kugel. Da die Paarkorrelationsfunktion normiert ist, ist sie in diesem Fall für alle Abstände gleich Eins. Diesen Wert findet man natürlich nur für $r \geq a$, da sich harte Kugeln nur bis zu diesem minimalen Abstand nähern können. Der allgemeinere Fall einer „Flüssigkeit" oder aber – bei starrer Anordnung der Kugeln – der eines „amorphen eindimensionalen Festkörpers" liegt zwischen den beiden Grenzfällen. Abhängig vom Wert des Parameters ℓ treten mehr oder weniger ausgeprägte Maxima auf, die sich mit zunehmendem Abstand vom Bezugsatom aufgrund der Mittelung über viele Konfigurationen verwischen (vgl. Abb. 1.28). In der Flüssigkeit bzw. der amorphen Struktur deutet sich eine lokale Ordnung an, deren Reichweite umso größer ist, je enger die Kugeln gepackt sind.

In Abb. 1.29 ist der Zusammenhang zwischen der Struktur und der Teilchenzahldichte für ein zweidimensionales System schematisch gezeigt. Die Skizze macht deutlich, daß das erste Maximum durch die nächsten Nachbarn hervorgerufen wird und daher bei $r \approx a$ zu finden ist. Eine gründlichere Analyse zeigt erwartungsgemäß, daß die Fläche unter dem Maximum durch die Anzahl der nächsten Nachbarn bestimmt wird. Die folgende „Koordinationsschale", d.h. die übernächsten Nachbarn, verursachen das zweite Maximum usw. Mit zunehmendem Abstand werden aufgrund der fehlenden Fernordnung die Maxima und Minima der Teilchenzahldichte bzw. der Paarkorrelation immer stärker ausgeschmiert. Bei sehr großen Abständen wird schließlich der Wert n_0 bzw. $g(r) = 1$ erreicht. Die gleiche Argumentation gilt auch bei dreidimensionalen Proben.

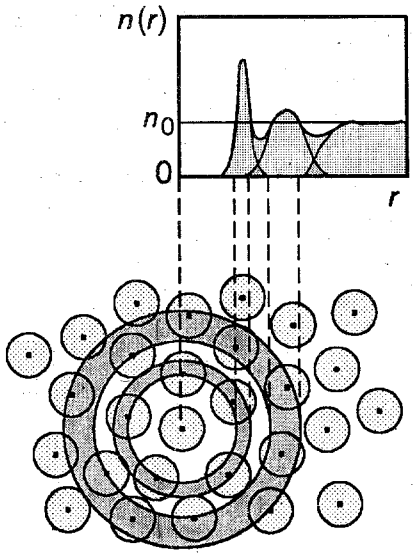


Abb. 1.29: Zweidimensionale Veranschaulichung des Zusammenhangs zwischen Struktur und Teilchenzahldichte. Wie bei der Paarkorrelationsfunktion ist der Beitrag des Aufatoms weggelassen.

In Abb. 1.30 wird die in Beugungeexperimenten (vgl. Kap. 2.5) bestimmte Paarkorrelationsfunktion eines amorphen Nickelfilms mit der einer Nickelschmelze verglichen. Deutlich ist zu erkennen, daß die Nahordnung in der amorphen Phase wesentlich stärker ausgeprägt ist als in der Flüssigkeit. In beiden Fällen nimmt die Ordnung mit dem Abstand vom Aufatom rasch ab.

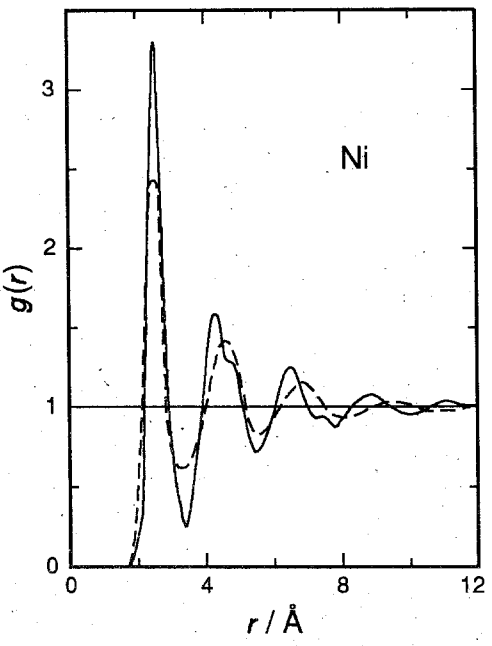


Abb. 1.30: Experimentell bestimmte Paarkorrelationsfunktion von amorphem (durchgezogen) und flüssigem (gestrichelt) Nickel. (Nach Y. Waseda, 1980).

Grundsätzlich wird die Beschreibung der Struktur von realen amorphen Festkörpern ganz erheblich durch die Tatsache vereinfacht, daß nur der Abstand r vom Bezugsatom für die Betrachtung herangezogen wird und keine Richtungsabhängigkeiten berücksichtigt werden müssen. Natürlich enthält die so gewonnene Paarkorrelationsfunktion der gemittelten lokalen Struktur der amorphen Festkörper wesentlich weniger Informationen, als man von kristallinen Festkörpern her gewohnt ist.

Einheitszelle im amorphen Festkörper

Nachdem wir die Problematik der Beschreibung der Struktur amorpher Festkörper ein wenig kennengelernt haben, wollen wir noch kurz die Frage stellen, ob auch in diesen Materialien eine Elementarzelle festgelegt werden kann. Einen Ansatzpunkt liefert die Definition der Wigner-Seitz-Zelle bei regulären Kristallgittern. Wie in Abb. 1.31 für zwei-dimensionale, monoatomare Systeme angedeutet, läßt sich im Falle dicht gepackter amorpher Festkörper das sogenannte *Voronoi-Polyeder* definieren, das große Ähnlichkeit mit der Wigner-Seitz-Zelle besitzt. Exakte Größe und Form dieser Polyeder variieren von Ort zu Ort. Die große Ähnlichkeit zur Wigner-Seitz-Zelle des entsprechenden Kristalls folgt unmittelbar aus der bestehenden Nahordnung.

Obwohl sich die Definition des Voronoi-Polyeders eng an die für kristalline Festkörper gebräuchliche Form für die Wigner-Seitz-Zelle anlehnt, ist ihr Nutzen bei der praktischen Durchführung von Berechnungen wesentlich geringer. Das Konzept ist dann nützlich, wenn physikalische Eigenschaften untersucht werden, die sich aus der Mittelung über viele Atome ergeben. Dagegen erhält man keine guten Ergebnisse, wenn lokale Eigenschaften

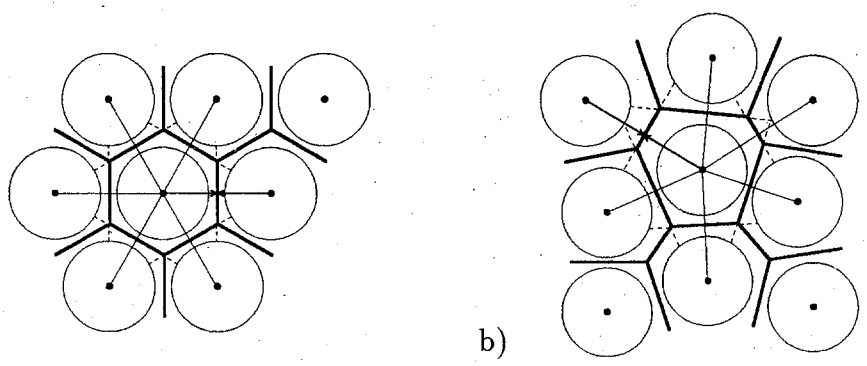


Abb. 1.31: Wigner-Seitz-Zellen und Voronoi-Polyeder zweidimensionaler Strukturen.
a) Wigner-Seitz-Zelle eines Kristalls,
b) Voronoi-Polyeder einer ungeordneten Struktur.

der amorphen Struktur Eingang finden oder wenn Defekte (vgl. Kap. 4 und 9) eine wichtige Rolle spielen, denn dort muß die starke Abweichung der lokalen Struktur vom „mittleren" Voronoi-Polyeder berücksichtigt werden.

2. Strukturbestimmung

Bei der Strukturuntersuchung von Festkörpern treten zwei ganz verschiedene Fragestellungen auf, die mit unterschiedlichen experimentellen Methoden angegangen werden. Es handelt sich einerseits um die Frage nach dem grundsätzlichen Aufbau, also nach der Struktur des Gitters und der Verteilung der Atome in der Elementarzelle, andererseits um die Frage nach Abweichungen vom idealen Aufbau. Wir beschränken uns in diesem Kapitel auf die erste der beiden Fragen, für deren Beantwortung Beugungsexperimente prädestiniert sind. Stehen dagegen Abweichungen von der idealen Struktur, Feinheiten der lokalen Anordnung von Atomen oder Defekte im Vordergrund, so sind andere Methoden besser geeignet. Hierzu zählen Untersuchungen, die z.B. auf spektroskopischen Methoden, auf dem Mößbauer-Effekt oder Elektronen- und Kernspinresonanz beruhen. Auf diese interessanten und ertragreichen Meßmethoden wollen wir hier aber nicht näher eingehen. Defekte in oberflächennahen Bereichen können, abhängig von ihrer Größe, mit unterschiedlichen abbildenden Verfahren, z.B. mit optischen, elektronen- oder rastermikroskopischen Techniken untersucht werden.

Wir beschäftigen uns zunächst mit der Beugung am Punktgitter und werden in diesem Zusammenhang den überaus wichtigen Begriff des reziproken Gitters kennenlernen. Nachdem wir auf die Bestimmung der Basis eingegangen sind, werden wir die Streuung an amorphen Festkörpern diskutieren. Am Ende dieses Kapitels werden wir uns noch mit den gebräuchlichen experimentellen Methoden der Strukturbestimmung beschäftigen.

2.1 Allgemeine Bemerkungen

Am „schönsten“ ist natürlich eine direkte Abbildung der atomaren Struktur. In den letzten Jahren ist atomare Auflösung von Oberflächenstrukturen mit Hilfe von Elektronen-, Feldionen-, Rasterkraft- oder Rastertunnelmikroskopie möglich geworden. Als Beispiel ist in Abb. 2.1 der Schnitt durch einen dünnen Film zu sehen, bei dem sich Silizium- und Germaniumschichten abwechseln. Diese hochauflösende elektronenmikroskopische Aufnahme ist in Transmission gemacht. Sie stellt aber nicht einfach eine geometrische Abbildung der Atome dar, wie sie von der gewöhnlichen Optik her bekannt ist. Es wirken auch Beugungseffekte mit, die die Germaniumschichten als dunkle Streifen in Erscheinung treten lassen. Dieser Effekt beruht darauf, daß die Germaniumatome aufgrund ihrer höheren Elektronenzahl in der Atomhülle die zur Abbildung benutzten Elektronen des Elektronenmikroskops stärker als die Siliziumatome streuen. In der Schicht wechseln sich 13 Atomlagen Silizium mit 2 Atomlagen Germanium ab. Die regelmäßige Aneinanderreihung von Schichten mit unterschiedlicher atomarer Zusammensetzung bezeichnet man als „Übergitter“. Die Herstellung derartiger Schichtstrukturen ist erst in der neueren Zeit möglich geworden. Auf einige der interessanten Eigenschaften dieser neuartigen Festkörper werden wir in Kap. 9.5 eingehen.

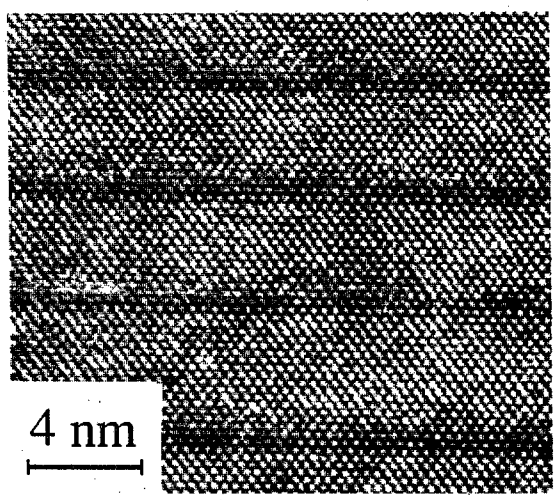


Abb. 2.1: Silizium-Germanium-Übergitter. Die Aufnahme wurde mit einem hochauflösenden Elektronenmikroskop in Transmission gemacht. Deutlich sind die einzelnen Atome zu erkennen. Es wechseln sich 13 Atomlagen Silizium mit 2 Atomlagen Germanium ab, die bei dieser Art der Abbildung etwas dunkler erscheinen. (Aus E. Müller et al., Phys. Rev. Lett. **63** (1989) 1819).

Rastertunnel- und Rasterkraftmikroskop wurden erst vor wenigen Jahren entwickelt und haben sofort große Bedeutung erlangt. Bei Messungen mit dem Rastertunnelmikroskop wird eine feine Metallspitze im Abstand von einigen Ångström über die Oberfläche einer leitenden Probe geführt. Bei kleinen Abständen überlappen die Wellenfunktionen von Probe und Spitze, und es können Elektronen ausgetauscht werden. Liegt eine elektrische Spannung an, so tritt ein Tunnelstrom auf, dessen Stärke ein Maß für den Abstand zwischen Probe und Spitze ist. Führt man die Spitze über die Oberfläche hinweg, so läßt sich aus dem lokalen Strom die Topographie bestimmen. Auf dieses Meßprinzip werden wir in Kap. 7.2 noch zu sprechen kommen.

Mit dem Rastertunnelmikroskop tastet man das Profil der Probenoberfläche ab und erhält daher keine Information über den inneren Aufbau des Festkörpers. Da auf die Atome an der Oberfläche andere Kräfte als im Inneren der Probe wirken, findet man dort oft Strukturen, die sich von der Struktur im Kristallinneren unterscheiden. Oberflächenabbildende Mikroskope sind daher nur bedingt für Strukturuntersuchungen geeignet, besitzen jedoch sehr große Bedeutung in der Oberflächenphysik. Die enorme Auflösung, die man mit einem Rastertunnelmikroskop erzielen kann, wird durch Abb. 2.2 demonstriert, in dem die Oberfläche eines Graphitkristalls mit atomarer Auflösung abgebildet ist.

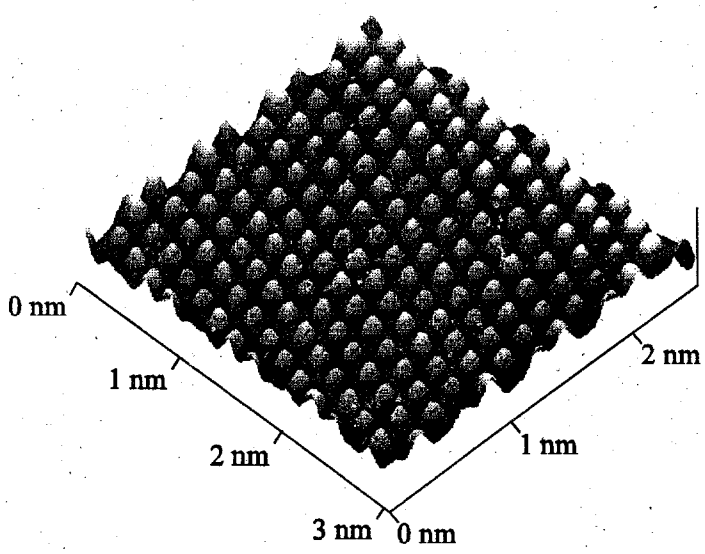


Abb. 2.2: Perspektivische Darstellung der Atome einer Graphitoberfläche. Die Messung wurde mit Hilfe eines Rastertunnelmikroskops durchgeführt. Aufgrund der Bindungsverhältnisse im Graphit ist nur jedes zweite Kohlenstoffatom der Probenoberfläche sichtbar. (C. Enss, D. Wehnes, Inst. Angw. Physik, Heidelberg).

Die detaillierten Kenntnisse, die wir über den Aufbau von Festkörpern besitzen, wurden in erster Linie durch Beugungs- bzw. Streuexperimente gewonnen. Hierbei bestrahlt man eine Probe beispielsweise mit kurzwelligen elektromagnetischen Wellen und beobachtet die richtungsabhängige Intensität der gebeugten Strahlung. Anstelle von Photonen werden oft auch Neutronen, Elektronen oder leichte Atome benutzt. Die Stärke der Wechselwirkung bestimmt die Eindringtiefe der Strahlung und somit den Teil des Festkörpers, der untersucht werden kann. Mit Ausnahme der Neutronen, die bei Atomen ohne magnetisches Moment ausschließlich an den Kernen gestreut werden, wechselwirken die übrigen Sonden mit den Elektronenhüllen. Aufgrund der starken Coulomb-Wechselwirkung eignen sich Elektronen- oder Atomstrahlen vor allem zur Untersuchung von Oberflächen und dünnen Filmen; Röntgen- und Neutronenstrahlen werden dagegen zur Untersuchung ausgedehnter Festkörper eingesetzt.

Aus der Optik ist bekannt, daß Beugung beobachtet wird, wenn die Lichtwellenlänge und die charakteristischen Abmessungen des beugenden Objekts vergleichbar sind. Bei der Bestimmung der Struktur von Festkörpern muß deshalb die Wellenlänge der einfallenden Strahlung kleiner als der Atomabstand sein. Je nach Natur der verwendeten Strahlung erfüllen Teilchen mit sehr unterschiedlicher Energie diese Bedingung. Bei einer Wellenlänge von 1 Å besitzen Röntgenquanten eine Energie von 12 keV, Elektronen 150 eV, Neutronen 80 meV und Heliumatome 20 meV.

Die beobachtete Streuwelle ist eine Superposition der Beiträge aller einzelnen Atome. Die atomare Anordnung läßt sich aber nur dann aus Beugungsexperimenten herleiten, wenn die einfallende Welle hohe Kohärenz besitzt und der Streuprozeß an gleichwertigen Atomen zur gleichen Phasenänderung führt, d.h. wenn die einfallende Welle in der Probe *kohärent* gestreut wird. Diese Art der Streuung wird auch als Beugung bezeichnet. Wir werden im weiteren Verlauf der Diskussion beide Ausdrücke nebeneinander benutzen. Bei *inkohärenter* Streuung hängt die Phasenverschiebung der gestreuten Welle vom Ort der Streuung ab. Aus einer inkohärenten Streuung lassen sich daher keine Rückschlüsse auf die Struktur der Probe ziehen. Ein bekanntes Beispiel hierfür ist die Streuung von Neutronen an Atomkernen mit Spin. Da die Kräfte zwischen Kern und Neutron von der relativen Orientierung der beiden Spins abhängen, variieren bei einer makroskopischen Probe mit ungerichteten Kernspins Phase und Amplitude der gestreuten Welle von Atom zu Atom.

Bei der Streuung an Festkörpern treten *elastische* und *inelastische* Prozesse auf. Bleiben die elektronischen oder vibronischen Zustände des Festkörpers beim Streuvorgang unbeeinflußt, ändert sich die Energie der gestreuten Strahlung nicht, und die Streuung ist elastisch. Bei inelastischen Streuprozessen dagegen wird zwischen dem gestreuten Teilchen und dem streuenden System Energie ausgetauscht. Beide Streuvorgänge spielen bei der Untersuchung der Festkörper eine wichtige Rolle, elastische Streuung nutzt man zur Strukturbestimmung und inelastische zur Untersuchung der Anregungszustände von Festkörpern.

2.2 Einfache Theorie der Streuung

Wir werden jetzt auf die Bestimmung der Struktur von Festkörpern mit Hilfe von Beugungsexperimenten eingehen. In diesem Zusammenhang betrachten wir eine Welle, die auf eine Probe trifft und von deren Atomen gestreut wird. Den Streuprozeß behandeln wir als rein klassischen Vorgang. Der quantenmechanische Aspekt tritt nur dadurch in

Erscheinung, daß auch Neutronen, Elektronen oder Atome wie elektromagnetische Strahlung als Welle behandelt werden. Um die Betrachtungen möglichst einfach zu halten, vernachlässigen wir den Einfluß von Absorption, Brechung und Änderungen der Intensität des einfallenden Strahls aufgrund von Streuprozessen. In unsere Überlegungen werden keine speziellen Eigenschaften der wechselwirkenden Strahlung eingehen, so daß sie für Wellen unterschiedlicher Natur gültig sind. Um die Streuung einer bestimmten Strahlung zu beschreiben, müssen die spezifischen Wechselwirkungsmechanismen berücksichtigt werden. Da neben der klassischen Methode der Strukturbestimmung mit Röntgenstrahlen auch die Neutronenstreuung sehr große Bedeutung erlangt hat, werden wir auf beide Meßmethoden etwa gleich stark eingehen.

Streuamplitude

Wie bereits erwähnt, soll mit Hilfe von Streuexperimenten die Verteilung der Streuzentren ermittelt werden. Mit der Neutronenstreuung läßt sich die Lage der Kerne bestimmen, die anderen Streuexperimente geben Auskunft über die Elektronenverteilung. Handelt es sich bei der einfallenden Welle um Röntgenstrahlung und beim streuenden Teilchen um ein Elektron, so bezeichnet man den Vorgang als *Thomson-Streuung.* Natürlich werden Röntgenstrahlen auch an Kernen gestreut, doch kann dieser Beitrag wegen der großen Kernmasse vernachlässigt werden.

In unserer Betrachtung beschreiben wir die einfallende Strahlung als eine ebene Welle mit der Kreisfrequenz ω_0, dem Wellenvektor $\boldsymbol{k}_0$ und der Feldstärke $A(t) = A_0 \mathrm{e}^{-\mathrm{i}(\omega_0 t - \boldsymbol{k}_0 \cdot \boldsymbol{r})}$. Jedes Streuzentrum der Probe wird selbst zum Ausgangspunkt einer Kugelwelle, deren Amplitude $A_\mathrm{s}(t)$ im Abstand $\tilde{r}$ von ihrem Entstehungsort sich durch

$$A_\mathrm{s}(t) = \frac{\widetilde{A}}{\tilde{r}} \mathrm{e}^{-\mathrm{i}(\omega_0 t - k_0 \tilde{r})} \tag{2.1}$$

ausdrücken läßt. Der Faktor $\widetilde{A}$ spiegelt die Streuwahrscheinlichkeit wider, ist also spezifisch für die verwendete Sonde und das Streuzentrum.

Die Aufgabe der Streutheorie besteht darin, den Zusammenhang zwischen der richtungsabhängigen Streuintensität und der *Streudichte* $\varrho(\boldsymbol{r})$, also der Verteilung der Streuzentren, herzustellen und daraus die Struktur der Probe herzuleiten. Wie bereits erwähnt, kann zwischen einfallender und gestreuter Welle eine Phasenverschiebung auftreten. Da wir kohärente Streuung voraussetzen, spielt ein konstanter Phasenfaktor keine Rolle und wird in unserer Diskussion nicht weiter beachtet. Von entscheidender Bedeutung ist dagegen der Phasenunterschied, der durch die unterschiedlichen Lagen der Streuzentren hervorgerufen wird und die Strukturinformation enthält.

Der Einfachheit halber nehmen wir an, daß sich der Beobachter B in großer Entfernung von der Probe befindet und die kugelförmigen Streuwellen nur innerhalb eines sehr kleinen Raumwinkels registriert. Diesen Teil der Wellenfront können wir näherungsweise als eben betrachten und ordnen ihm den Wellenvektor $\boldsymbol{k}$ zu. Ein Blick auf Abb. 2.3 zeigt uns, daß zwischen den Teilwellen, die von den beiden eingezeichneten Volumenelementen ausgehen und bezüglich des Beobachters durch die parallelen Strahlen 0 und 1 charakterisiert werden können, die Wegdifferenz $\Delta s = (\boldsymbol{k} \cdot \boldsymbol{r}/k) - (\boldsymbol{k}_0 \cdot \boldsymbol{r}/k_0)$ besteht. Da sich bei der elastischen Streuung der Betrag des Wellenvektors nicht ändert, dürfen wir $k = k_0$ setzen, so daß die Phasendifferenz $\Delta\varphi$ zwischen den beiden Strahlen durch $\Delta\varphi = (\boldsymbol{k} - \boldsymbol{k}_0) \cdot \boldsymbol{r}$ gegeben

ist. Damit können wir den Beitrag $\mathrm{d}A_\mathrm{s}(\boldsymbol{r},t)$ des Volumenelementes $\mathrm{d}V$ am Ort $\boldsymbol{r}$ zur Amplitude der Streustrahlung beim Beobachter B leicht angeben:

$$\mathrm{d}A_\mathrm{s}(\boldsymbol{r},t) = \frac{\widetilde{A}}{\tilde{r}_1}\varrho(\boldsymbol{r})\,\mathrm{e}^{-\mathrm{i}[\omega_0 t - k\tilde{r}_0 + (\boldsymbol{k}-\boldsymbol{k}_0)\cdot\boldsymbol{r}]}\,\mathrm{d}V\ . \tag{2.2}$$

Der Term $\varrho(\boldsymbol{r})\mathrm{d}V$ gibt an, mit welcher Wahrscheinlichkeit das Elektron oder der streuende Kern im Volumenelement $\mathrm{d}V$ angetroffen wird.

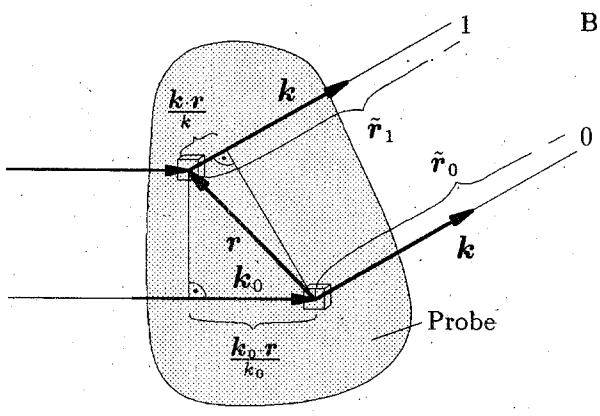


Abb. 2.3: Zur Herleitung der Streuamplitude. Es läuft eine ebene Welle mit dem Wellenvektor $\boldsymbol{k}_0$ ein. Die Wellenfront, die auf den weit entfernten Beobachter B zuläuft, läßt sich durch den Wellenvektor $\boldsymbol{k}$ charakterisieren. Die Vektoren $\tilde{\boldsymbol{r}}_0$ und $\tilde{\boldsymbol{r}}_1$ sind näherungsweise parallel und verbinden die herausgegriffenen Streuvolumina mit dem Beobachter B.

Durch dieses Vorgehen beziehen wir die Phase der Teilwellen auf die Streuwelle des Aufpunkts der Probe, d.h. auf den Ort $\boldsymbol{r} = 0$. Die Gesamtamplitude der Streustrahlung erhalten wir, kohärente Streuung vorausgesetzt, durch Integration über das Probenvolumen V_P. Ist die Ausdehnung der Probe klein gegen den Abstand des Beobachters, können wir im Nenner $\tilde{r}_1 \approx \tilde{r}_0$ setzen und erhalten

$$A_\mathrm{s}(t) = \frac{\widetilde{A}}{\tilde{r}_0}\,\mathrm{e}^{-\mathrm{i}(\omega_0 t - k\tilde{r}_0)} \int\limits_{V_\mathrm{P}} \varrho(\boldsymbol{r})\,\mathrm{e}^{-\mathrm{i}(\boldsymbol{k}-\boldsymbol{k}_0)\cdot\boldsymbol{r}}\mathrm{d}V\ . \tag{2.3}$$

Vor dem Integral stehen nur Größen, die nicht vom Ortsvektor $\boldsymbol{r}$ abhängen, sondern deren Wert durch die Versuchsführung bestimmt werden. Die Information über die Verteilung $\varrho(\boldsymbol{r})$ der Streuzentren und somit über die Struktur des untersuchten Festkörpers steckt im Integral. Man bezeichnet diese Größe als *Streuamplitude* $\mathcal{A}$, für die wir

$$\mathcal{A}(\boldsymbol{K}) = \int\limits_{V_\mathrm{P}} \varrho(\boldsymbol{r})\,\mathrm{e}^{-\mathrm{i}\boldsymbol{K}\cdot\boldsymbol{r}}\mathrm{d}V \tag{2.4}$$

schreiben können, wenn wir den *Streuvektor* $\boldsymbol{K} = \boldsymbol{k} - \boldsymbol{k}_0$ einführen. Ein Blick auf die Gleichung läßt unschwer erkennen, daß es sich hierbei um die Fourier-Transformierte der Streudichte $\varrho(\boldsymbol{r})$ handelt, die sich daher aus der Streuamplitude über die Beziehung

$$\varrho(\boldsymbol{r}) = \frac{1}{(2\pi)^3}\int \mathcal{A}(\boldsymbol{K})\,\mathrm{e}^{\mathrm{i}\boldsymbol{K}\cdot\boldsymbol{r}}\mathrm{d}^3K \tag{2.5}$$

gewinnen läßt.

2.3 Fourier-Entwicklung von Punktgittern

Aufgrund der Translationsinvarianz (2.6) der Kristalle ist $\varrho(\boldsymbol{r})$ periodisch in Richtungen der Basisvektoren. Es liegt daher nahe, diese Symmetrie bei der Auswertung der Streuspektren von Anfang an zu nutzen. Dazu entwickeln wir die Streudichte in Fourier-Reihen und untersuchen die einzelnen Fourier-Komponenten separat. Wie wir sehen werden, führt die Fourier-Entwicklung der Kristallgitter zu einer Reihe von interessanten und wichtigen Konzepten, von denen wir *reziprokes Gitter*, *Millersche Indizes* und *Brillouin-Zone* herausgreifen. Nach deren Behandlung werden wir auf unser ursprüngliches Problem, die Bestimmung der Kristallstruktur durch Streuexperimente, zurückkommen.

Reziprokes Gitter

Für die dreidimensionale periodische Streudichte $\varrho(\boldsymbol{r})$ hat die Entwicklung in eine Fourier-Reihe die Form

$$\varrho(\boldsymbol{r}) = \sum_{h,k,l} \varrho_{hkl}\, \mathrm{e}^{\mathrm{i}\boldsymbol{G}_{hkl}\cdot\boldsymbol{r}}\,, \tag{2.6}$$

wobei h, k und l unabhängige ganze Zahlen sind, über die summiert wird. Die Fourier-Koeffizienten ϱ_{hkl} sind wie üblich durch die Gleichung

$$\varrho_{hkl} = \frac{1}{V_\mathrm{Z}} \int_{V_\mathrm{Z}} \varrho(\boldsymbol{r})\, \mathrm{e}^{-\mathrm{i}\boldsymbol{G}_{hkl}\cdot\boldsymbol{r}}\, \mathrm{d}V \tag{2.7}$$

definiert. Die Integration erfolgt über die Periode der Funktion, in unserem Fall also über das Volumen V_Z der primitiven Elementarzelle.

Die Vektoren $\boldsymbol{G}_{hkl}$ stellen wir in der Form $\boldsymbol{G}_{hkl} = h\boldsymbol{b}_1 + k\boldsymbol{b}_2 + l\boldsymbol{b}_3$ dar, wobei die Vektoren $\boldsymbol{b}_1$, $\boldsymbol{b}_2$ und $\boldsymbol{b}_3$ ein neues, schiefwinkliges Koordinatensystem festlegen, das wir noch bestimmen müssen. Da h, k und l nur diskrete Werte annehmen können, repräsentiert jeder Vektor $\boldsymbol{G}_{hkl}$ einen Punkt eines neuen Gitters, das als *reziprokes Gitter* bezeichnet wird. Die Wahl des neuen Koordinatensystems mit den Basisvektoren $\boldsymbol{b}_1$, $\boldsymbol{b}_2$ und $\boldsymbol{b}_3$ und damit auch das von ihnen aufgespannte Gitter unterliegt Einschränkungen, die auf der Forderung nach Translationsinvarianz des realen Gitters beruhen. Die Periodizität der Streudichte $\varrho(\boldsymbol{r})$ läßt sich durch die Gleichung

$$\varrho(\boldsymbol{r}) = \varrho(\boldsymbol{r} + \boldsymbol{R}) \tag{2.8}$$

ausdrücken, wobei $\boldsymbol{R} = n_1\boldsymbol{a}_1 + n_2\boldsymbol{a}_2 + n_3\boldsymbol{a}_3$ ein Gittervektor des Kristalls ist. Zur Vereinfachung der Schreibweise benutzen wir in diesem Kapitel für die Basisvektoren die Bezeichnung $\boldsymbol{a}_1$, $\boldsymbol{a}_2$ und $\boldsymbol{a}_3$ anstelle von $\boldsymbol{a}$, $\boldsymbol{b}$ und $\boldsymbol{c}$. Da die Fourier-Koeffizienten ϱ_{hkl} nicht ortsabhängig sind, die Streudichte (2.6) aber translationsinvariant ist, trifft dies auch für den ortsabhängigen Faktor $\exp(\mathrm{i}\boldsymbol{G}_{hkl}\cdot\boldsymbol{r})$ zu, d.h. es muß $\exp[\mathrm{i}\boldsymbol{G}_{hkl}\cdot(\boldsymbol{r}+\boldsymbol{R})] = \exp(\mathrm{i}\boldsymbol{G}_{hkl}\cdot\boldsymbol{r})$ sein. Dies ist nur möglich, wenn $\exp(\mathrm{i}\boldsymbol{G}_{hkl}\cdot\boldsymbol{R}) = 1$ oder $\boldsymbol{G}_{hkl}\cdot\boldsymbol{R} = 2\pi p$ ist, wobei p eine ganze Zahl sein muß. Führt man das Skalarprodukt aus, so sieht man sofort, daß diese Bedingung nur dann für beliebige Koeffizienten n_i und h_j erfüllt sein kann, wenn der Zusammenhang

$$\boldsymbol{b}_i \cdot \boldsymbol{a}_j = 2\pi\, \delta_{ij} \tag{2.9}$$

besteht, wobei δ_{ij} das Kronecker-Symbol bezeichnet. Daraus folgen für die Basisvektoren des reziproken Gitters unmittelbar die Beziehungen

$$\boldsymbol{b}_1 = \frac{2\pi}{V_Z}(\boldsymbol{a}_2 \times \boldsymbol{a}_3) \ , \qquad \boldsymbol{b}_2 = \frac{2\pi}{V_Z}(\boldsymbol{a}_3 \times \boldsymbol{a}_1) \ , \qquad \boldsymbol{b}_3 = \frac{2\pi}{V_Z}(\boldsymbol{a}_1 \times \boldsymbol{a}_2) \ . \tag{2.10}$$

Das Volumen $V_Z = (\boldsymbol{a}_1 \times \boldsymbol{a}_2) \cdot \boldsymbol{a}_3$ der Elementarzelle des realen Gitters steht mit dem Volumen des reziproken Gitters über $(\boldsymbol{b}_1 \times \boldsymbol{b}_2) \cdot \boldsymbol{b}_3 = (2\pi)^3/V_Z$ in Zusammenhang. Für die Gitterkonstanten $b_i = |\boldsymbol{b}_i|$ im kubischen Gitter gilt nach (2.9) die Beziehung $b_i = 2\pi/a_i$. Durch (2.9) bzw. (2.10) ist das reale Gitter mit dem reziproken Gitter eindeutig verknüpft. An dieser Stelle soll betont werden, daß *ein* bestimmter reziproker Gittervektor G_{hkl} jeweils *einem* bestimmten Fourier-Koeffizienten ϱ_{hkl} der Streudichte zugeordnet ist.

Anhand einiger einfacher Beispiele wollen wir den Zusammenhang zwischen realem und reziprokem Gitter veranschaulichen. Bei einer *linearen Kette* mit der Gitterkonstanten a ist das reziproke Gitter ebenfalls linear und hat gemäß (2.9) die Gitterkonstante $2\pi/a$. Für ein schiefwinkliges *ebenes* Gitter ist die Verknüpfung der beiden Gitter in Abb. 2.4 dargestellt. Die Basisvektoren $\boldsymbol{b}_1$ und $\boldsymbol{b}_2$ des reziproken Gitters stehen senkrecht auf den Basisvektoren $\boldsymbol{a}_2$ und $\boldsymbol{a}_1$ des ursprünglichen Gitters.

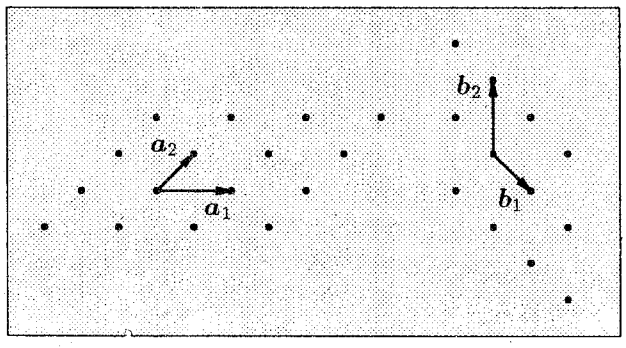


Abb. 2.4: Zusammenhang zwischen realem und reziprokem Gitter. Der Basisvektor $\boldsymbol{b}_1$ steht senkrecht auf $\boldsymbol{a}_2$, und $\boldsymbol{b}_2$ senkrecht auf $\boldsymbol{a}_1$.

Die Elementarzelle des *kubisch primitiven Gitters* im realen und im reziproken Raum ist in Abb. 2.5 zu sehen. Sind $\hat{\boldsymbol{x}}, \hat{\boldsymbol{y}}$ und $\hat{\boldsymbol{z}}$ die Einheitsvektoren des kartesischen Koordinatensystems und a die Gitterkonstante, so gilt für das reale Gitter $\boldsymbol{a}_1 = a\hat{\boldsymbol{x}}, \boldsymbol{a}_2 = a\hat{\boldsymbol{y}}$ und $\boldsymbol{a}_3 = a\hat{\boldsymbol{z}}$. Mit $(\boldsymbol{a}_1 \times \boldsymbol{a}_2) \cdot \boldsymbol{a}_3 = a^3$ findet man für die Basisvektoren des reziproken Gitters: $\boldsymbol{b}_1 = 2\pi(\boldsymbol{a}_2 \times \boldsymbol{a}_3)/a^3 = 2\pi\hat{\boldsymbol{x}}/a$ und entsprechende Ausdrücke für $\boldsymbol{b}_2$ und $\boldsymbol{b}_3$. Das reziproke Gitter des kubisch primitiven Gitters ist also ebenfalls kubisch primitiv. Die Gitterkonstante ist $2\pi/a$ und das Volumen der primitiven Elementarzelle $(2\pi/a)^3$.

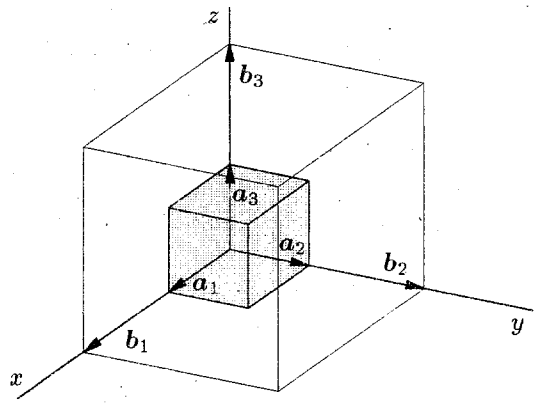


Abb. 2.5: Einheitszelle des realen (dunkel getönt) und des reziproken kubisch primitiven Gitters mit den Gitterkonstanten a bzw. $2\pi/a$.

Nicht ganz so einfach läßt sich der Zusammenhang zwischen realem und reziprokem Gitter bei den nicht-primitiven kubischen Gittern herstellen. Hier muß beachtet werden, daß sich die Definition des reziproken Gitters immer auf ein Gitter mit primitiver Elementarzelle bezieht. Eine etwas längere Rechnung zeigt, daß das *kubisch raumzentrierte* Gitter das reziproke Gitter des *kubisch flächenzentrierten* ist und umgekehrt.

Generell ist zu beachten, daß der reziproke Raum die Dimension m^{-3} besitzt. Der Betrag der reziproken Vektoren wird in Einheiten von m^{-1} gemessen. Die gleiche Dimension weisen auch die Wellenvektoren bzw. Wellenzahlen auf. Dies bedeutet, daß bei Darstellungen wie in Abb. 2.5 eigentlich zwei verschiedene Räume, der Ortsraum und der reziproke Raum, ineinander gezeichnet sind.

Millersche Indizes

Eine Ebene im Kristall, die durch Gitterpunkte aufgespannt wird, bezeichnet man als *Kristall-, Gitter-* oder *Netzebene.* Aufgrund der Translationsinvarianz der Kristalle gibt es jeweils (unendlich) viele äquivalente Ebenen, die parallel zueinander liegen. Am einfachsten wären diese Ebenen zu kennzeichnen, indem man ihre Schnittpunkte mit den Kristallachsen geeignet angibt. Es zeigt sich jedoch, daß diese Vorgehensweise bei der Analyse der Kristallstrukturen nicht besonders vorteilhaft ist. Es ist geschickter, wie bei der soeben diskutierten Fourier-Zerlegung des Gitters, reziproke Längen bei der Charakterisierung zu benutzen. Hierauf beruht die Indizierung der Kristallebenen mit Hilfe der sogenannten *Millerschen Indizes.* Die Vorgehensweise läßt sich anhand des einfachen Beispiels in Abb. 2.6 gut demonstrieren, in dem einige äquivalente Ebenen eines monoklinen Gitters eingezeichnet sind.

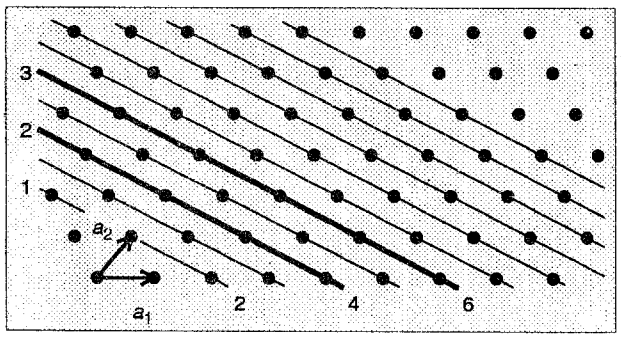


Abb. 2.6: Schnitt durch ein monoklines Gitter. Die $\boldsymbol{a}_3$-Achse und die eingezeichnete Netzebenenschar stehen senkrecht auf der Zeichenebene. Herausgehoben sind die Ebenen mit den Achsenabschnitten $s_1 = 4\boldsymbol{a}_1$ und $s_2 = 2a_2$ bzw. $s_1' = 6a_1$ und $s_2' = 3a_2$.

Wir starten mit der Angabe der Achsenabschnitte s_i der beiden hervorgehobenen Ebenen. In Einheiten der Basisvektoren finden wir die Werte $s_1/a_1 = \tilde{s}_1 = 4$ bzw. 6 in $\boldsymbol{a}_1$-Richtung und $s_2/a_2 = \tilde{s}_2 = 2$ bzw. 3 in $\boldsymbol{a}_2$-Richtung. Weiter nehmen wir an, daß die herausgegriffene Ebenenschar parallel zur nicht dargestellten $\boldsymbol{a}_3$-Achse verläuft, also senkrecht auf der Zeichenebene steht. Da dann kein Schnittpunkt auftritt, ist der Achsenabschnitt $s_3/a_3 = \tilde{s}_3 = \infty$. Nun bilden wir die Kehrwerte

$$h' = \frac{1}{\tilde{s}_1} = \frac{1}{4} \text{ bzw. } \frac{1}{6} \qquad k' = \frac{1}{\tilde{s}_2} = \frac{1}{2} \text{ bzw. } \frac{1}{3} \qquad l' = \frac{1}{\tilde{s}_3} = \frac{1}{\infty} = 0\,.$$

Die Zahlen h', k' und l' multiplizieren wir mit einem Faktor p, der so gewählt wird, daß sich drei ganze, normalerweise möglichst kleine Zahlen ergeben. Im vorliegenden Fall erhalten wir mit $p = 4$ (bzw. $p = 6$ für die zweite Ebene) die Millerschen Indizes (hkl)

der eingezeichneten Ebenen. In der Zeichnung haben wir offensichtlich die (120)-Ebenen herausgegriffen. Wie man sich leicht anhand eines Beispiels klarmachen kann, hängt die Bezeichnung der Gitterebenen von der Wahl der Basisvektoren ab.

Mit Hilfe der Millerschen Indizes lassen sich nicht nur die Gitterebenen, sondern auch Richtungen im Kristall angeben. Unter der $[hkl]$-*Richtung*, gekennzeichnet durch eckige Klammern, versteht man die Richtung des Vektors $[hkl] = h\boldsymbol{a}_1 + k\boldsymbol{a}_2 + l\boldsymbol{a}_3$. Auch hier benutzt man möglichst kleine Indizes. Bei kubischen Kristallen steht der Vektor $[hkl]$ senkrecht auf den (hkl)-Netzebenen. Negative Vorzeichen bei der Bezeichnung von Ebenen und Richtungen werden über die Zahl geschrieben, z.B. $(\bar{1}00)$ oder $[0\bar{1}0]$. Möchte man nicht eine spezielle Netzebenenschar (hkl) oder eine spezielle Richtung $[hkl]$ herausgreifen, sondern die Gesamtheit aller symmetriebedingten „gleichwertigen" Netzebenenscharen oder Kristallrichtungen angeben, so schreibt man $\{hkl\}$ bzw. $\langle hkl \rangle$. Die unterschiedlichen gleichwertigen Netzebenenscharen bzw. Richtungen lassen sich unter Ausnutzung von Drehsymmetrien ineinander überführen.

Meist sind Ebenen mit niedriger Indizierung von besonderem Interesse, da sie die grundlegenden Periodizitäten der Kristallgitter widerspiegeln. Dies sind neben den {100}-Netzebenen vor allem die {110}- und {111}-Ebenen. In Abb. 2.7 ist die Lage dieser Ebenen für kubische Kristalle angegeben. Gelegentlich findet man auch höher indizierte Ebenen, die nicht entsprechend dem Kriterium kleinstmöglicher Zahlen indiziert sind. Diese Ebenen haben durchaus physikalische Bedeutung, wenn die Basis komplizierter aufgebaut ist oder wenn die betrachtete Elementarzelle nicht primitiv ist und deswegen mehrere Gitterpunkte enthält. So verläuft die (200)-Ebene parallel zur (100)-Ebene, doch schneidet sie die Achse beim halben Gitterabstand. Bei kubisch raumzentrierten Kristallen liegt das Atom im Würfelzentrum der nicht-primitiven Elementarzelle (vgl. Bravais-Gitter Abb. 1.11) auf dieser Ebene.

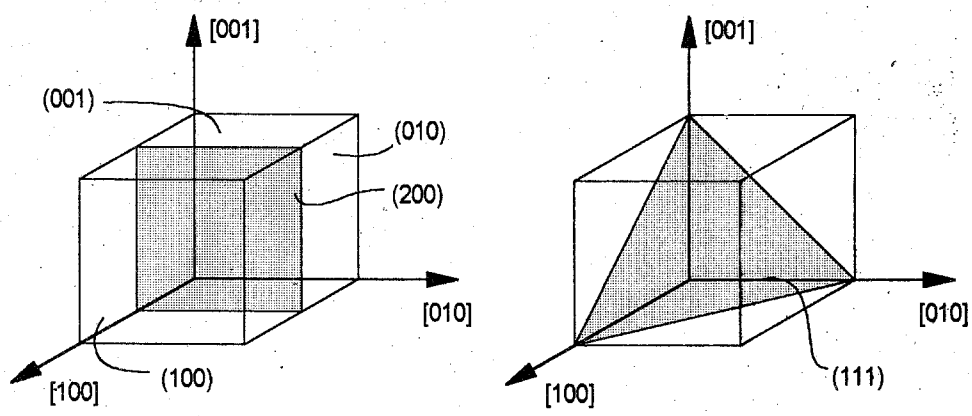


Abb. 2.7: Ausgewählte Gitterebenen eines kubischen Kristalls. Links sind von den gleichwertigen {100}-Ebenen die Ebenen (100), (010) und (001) bezeichnet. Dunkel dargestellt sind links die (200)- und rechts die (111)-Ebene.

Bei der Indizierung von Ebenen der hexagonalen Kristalle ist besondere Vorsicht geboten. Benutzt man das Bravais-Gitter, so bewirkt die spezielle Symmetrie dieses Kristallsystems, daß gleichwertige Ebenen mit der diskutierten Indizierungsvorschrift unterschiedlich bezeichnet werden. So lassen sich die beiden Ebenen (100) und $(1\bar{1}0)$ durch Drehung um 60° ineinander überführen, sind also gleichwertig, obwohl man dies aufgrund ihrer Indizierung nicht erwartet. Durch die Einführung eines zusätzlichen Index läßt sich

dieser Mangel beheben. Die Ebenen werden mit $(hkil)$ bezeichnet, und der zusätzliche Index i wird durch $i = -(h + k)$ definiert. In unserem Beispiel handelt es sich somit um die Ebenen $(10\bar{1}0)$ und $(1\bar{1}00)$, oder allgemeiner um $\{1\bar{1}00\}$-Ebenen.

Offensichtlich besteht ein Zusammenhang zwischen den Millerschen Indizes und dem reziproken Gitter. Um die Verknüpfung zu finden, greifen wir – wie in Abb. 2.8 skizziert – eine beliebige Netzebene heraus und legen ihre räumliche Lage durch die drei Vektoren $\boldsymbol{u}$, $\boldsymbol{v}$ und $\boldsymbol{w}$ fest, die sich mit Hilfe der Basisvektoren des betreffenden Gitters ausdrücken lassen: $\boldsymbol{u} = \tilde{h}\boldsymbol{a}_1$, $\boldsymbol{v} = \tilde{k}\boldsymbol{a}_2$, und $\boldsymbol{w} = \tilde{l}\boldsymbol{a}_3$. Wie bei der Herleitung der Millerschen Indizes gehen wir auch hier auf die Kehrwerte $h' = 1/\tilde{h}$, $k' = 1/\tilde{l}$ und $l' = 1/\tilde{l}$ der Achsenabschnitte über.

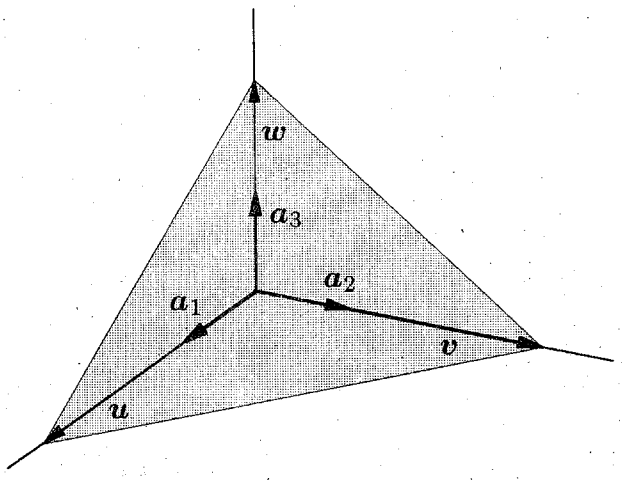


Abb. 2.8: Zur Herleitung des Zusammenhangs zwischen den Millerschen Indizes und dem reziproken Gitter. Die Vektoren $\boldsymbol{u}, \boldsymbol{v}$ und $\boldsymbol{w}$, die sich in Einheiten der Basisvektoren $\boldsymbol{a}_1$, $\boldsymbol{a}_2$ und $\boldsymbol{a}_3$ ausdrücken lassen, spannen eine Netzebene des Kristallgitters auf.

Da die Vektoren $(\boldsymbol{u}-\boldsymbol{v})$ und $(\boldsymbol{w}-\boldsymbol{v})$ in der Netzebene liegen, definiert das Kreuzprodukt $(\boldsymbol{u}-\boldsymbol{v})\times(\boldsymbol{w}-\boldsymbol{v})$ einen Vektor senkrecht zur betrachteten Ebene. Mit der Definition (2.10) gilt für diesen Vektor

$$(\boldsymbol{u}-\boldsymbol{v})\times(\boldsymbol{w}-\boldsymbol{v}) = (\boldsymbol{u}\times\boldsymbol{w}) - (\boldsymbol{v}\times\boldsymbol{w}) - (\boldsymbol{u}\times\boldsymbol{v})$$

$$= \tilde{h}\tilde{l}(\boldsymbol{a}_1\times\boldsymbol{a}_3) - \tilde{k}\tilde{l}(\boldsymbol{a}_2\times\boldsymbol{a}_3) - \tilde{h}\tilde{k}(\boldsymbol{a}_1\times\boldsymbol{a}_2)\,. \qquad (2.11)$$

Multiplizieren wir diese Gleichung mit $-2\pi h'k'l'/V_Z$, so erhalten wir

$$-\frac{2\pi}{V_Z}h'k'l'(\boldsymbol{u}-\boldsymbol{v})\times(\boldsymbol{w}-\boldsymbol{v}) = h'\boldsymbol{b}_1 + k'\boldsymbol{b}_2 + l'\boldsymbol{b}_3 = \frac{1}{p}\,\boldsymbol{G}_{hkl}\,. \qquad (2.12)$$

Der betrachtete Vektor und der reziproke Gittervektor verlaufen parallel zueinander, d.h. der reziproke Gittervektor $\boldsymbol{G}_{hkl}$ steht senkrecht auf den Netzebenen (hkl). Mit elementarer analytischer Geometrie läßt sich weiter zeigen, daß der *Abstand d* zwischen zwei benachbarten Netzebenen durch

$$d_{hkl} = 2\pi/|\boldsymbol{G}_{hkl}| \qquad (2.13)$$

direkt mit dem Betrag des reziproken Gittervektors verknüpft ist. Somit repräsentiert der reziproke Gittervektor $\boldsymbol{G}_{hkl}$ die Netzebenenschar (hkl): Er steht senkrecht auf diesen Ebenen, und sein Betrag wird durch den reziproken Ebenenabstand bestimmt. Jeder Netzebenenschar ist ein reziproker Gittervektor zugeordnet und umgekehrt. Diese

Wechselbeziehung wird verständlich, wenn man bedenkt, daß die reziproken Gittervektoren unmittelbar die Fourier-Komponenten der Streudichte widerspiegeln. Reziproke Gittervektoren mit kleinen Indizes werden durch die „grobe" Struktur des Kristallgitters bestimmt, die hoch indizierten durch die „Feinheiten" der Ladungsverteilung.

Brillouin-Zone

Wie im realen Gitter, so läßt sich auch eine Wigner-Seitz-Zelle (vgl. Kap. 1.3) des reziproken Gitters definieren. Sie wird als *erste Brillouin-Zone* bezeichnet und besitzt, wie wir in den folgenden Kapiteln sehen werden, fundamentale Bedeutung in der Festkörperphysik. Die Brillouin-Zone spielt eine wichtige Rolle in der Gitterdynamik, bei der Beschreibung der Bewegung von Elektronen und bei vielen anderen Phänomenen, bei denen die Gitterperiodizität im Vordergrund steht. Das Konstruktionsprinzip der Wigner-Seitz-Zelle läßt sich durch Hinzuziehen weiter entfernter Gitterpunkte erweitern und führt dann zu *Brillouin-Zonen höherer Ordnung.* Eine gründlichere geometrische Betrachtung zeigt, daß ihr Volumen unabhängig von der Ordnung ist.

An dieser Stelle wollen wir einige einfache Beispiele für Brillouin-Zonen höherer Ordnung anführen. In der Darstellung des reziproken Gitters werden wir die Achsen zukünftig meist mit dem Symbol k bezeichnen, das wir für den Wellenvektor bzw. die Wellenzahl reserviert hatten. Da die Gitterkonstanten b_i des reziproken Gitters mit den Gitterkonstanten a_j des realen Gitters über $\boldsymbol{b}_i \cdot \boldsymbol{a}_j = 2\pi\delta_{ij}$ zusammenhängen, ist diese Benennung gerechtfertigt. Weitere Argumente werden wir in den folgenden Abschnitten finden. Wie bereits erwähnt, ist das reziproke Gitter einer *linearen Kette* wieder eine Kette mit der Gitterkonstanten $2\pi/a$. Die 1. Brillouin-Zone liegt im Zentrum des reziproken Gitters, d.h. im Wellenvektorbereich $-\pi/a \leq k \leq \pi/a$. Die Zonen höherer Ordnung schließen sich außen an. In Abb. 2.9 ist die 2. Brillouin-Zone durch dunkle und die dritte durch helle Tönung angedeutet. Man beachte, daß jeweils zwei getrennte Hälften zu den Brillouin-Zonen höherer Ordnung gehören, die zusammen die gleiche Ausdehnung wie die 1. Brillouin-Zone besitzen. Der Aufpunkt des reziproken Gitters bzw. der Mittelpunkt der 1. Brillouin-Zone wird als Γ-Punkt bezeichnet.

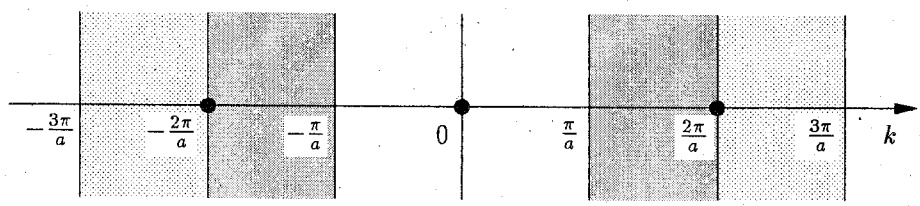


Abb. 2.9: Brillouin-Zonen einer linearen Kette. Die beiden Hälften der 2. Brillouin-Zone sind durch dunkle, die der 3. Brillouin-Zone durch helle Tönung gekennzeichnet. Der Γ-Punkt des reziproken Gitters liegt bei $k = 0$.

In Abb. 2.10 sind die ersten drei Brillouin-Zonen eines *Rechteckgitters* zu sehen. Die 1. Brillouin-Zone ist wieder ein Rechteck, doch mit zunehmender Ordnung zerfallen die Zonen in immer mehr Einzelteile. Durch die Verschiebung dieser Teile um *ganze* Basisvektoren des reziproken Gitters können sie jeweils wieder zu einem Rechteck mit der Fläche der 1. Brillouin-Zone zusammengefügt werden.

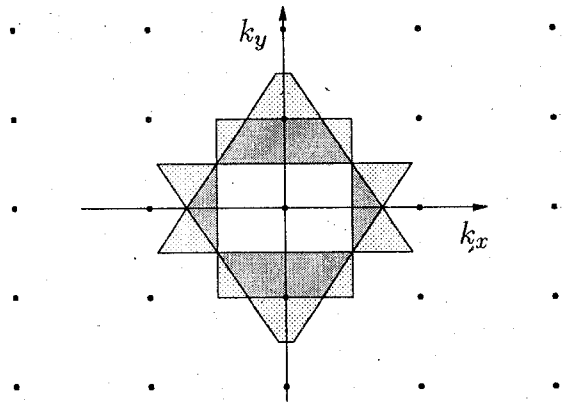


Abb. 2.10: Brillouin-Zonen eines Rechteckgitters. Die 2. Brillouin-Zone ist durch dunkle und die dritte durch helle Tönung gekennzeichnet.

Als Beispiel für dreidimensionale Brillouin-Zonen sind in Abb. 2.11 die 1. Brillouin-Zone des kubisch flächenzentrierten, kubisch raumzentrierten und hexagonalen Gitters gezeigt. Es ist üblich, Punkte hoher Symmetrie mit Abkürzungen aus der Gruppentheorie wie Γ, L, K, X usw. zu benennen. Wie bereits oben erwähnt, wird der Ursprung der 1. Brillouin-Zone als Γ-*Punkt* bezeichnet. Wie die Polyeder der Wigner-Seitz-Zelle den realen Raum, so füllen bei einer periodischen Wiederholung die Polyeder der Brillouin-Zone den reziproken Raum lückenlos. Auch die dreidimensionalen Brillouin-Zonen höherer Ordnung bestehen aus einer mit zunehmender Ordnung steigenden Anzahl von Einzelteilen, die sich durch Verschieben um einen reziproken Gittervekor wieder zu einem Polyeder mit der Größe und Gestalt der 1. Brillouin-Zone zusammenfügen lassen. Wir werden auf Brillouin-Zonen höherer Ordnung noch öfters zu sprechen kommen, insbesondere sind sie in Zusammenhang mit den elektrischen Eigenschaften von Festkörpern von Bedeutung. Ohne Beweis soll hier noch angefügt werden, daß die 1. Brillouin-Zone des kubisch flächenzentrierten Gitters gerade der Wigner-Seitz-Zelle des kubisch raumzentrierten Gitters entspricht und umgekehrt.

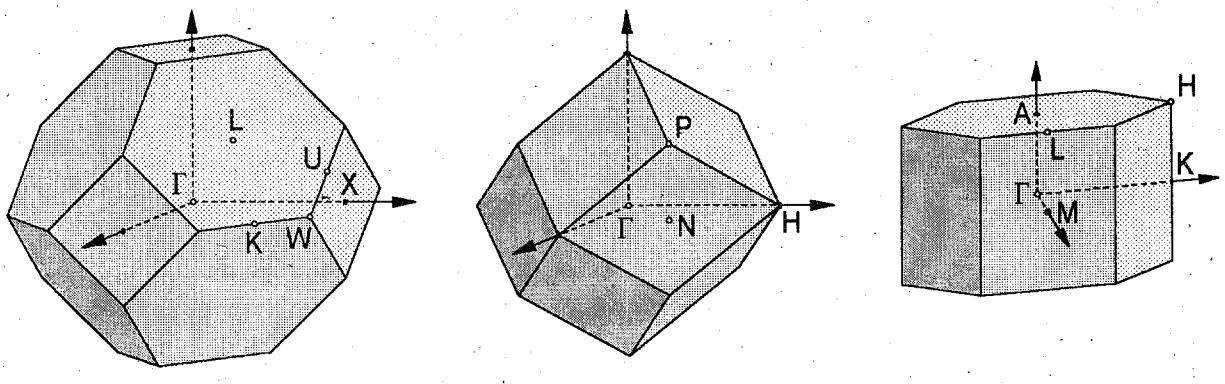


Abb. 2.11: 1. Brillouin-Zone des kubisch flächenzentrierten (links), des kubisch raumzentrierten (Mitte) und des hexagonalen (rechts) Gitters. Punkte hoher Symmetrie sind durch Buchstaben besonders gekennzeichnet.

2.4 Streuung an Kristallen

Bei der Diskussion des Streuvorganges haben wir gefunden, daß die Amplitude des Streusignals proportional zur Streuamplitude $\mathcal{A}(\boldsymbol{K})$ ist. Die beobachtete Streuintensität $I(\boldsymbol{K})$ ist also durch die Beziehung

$$I(\boldsymbol{K}) \propto |\mathcal{A}(\boldsymbol{K})|^2 = \left| \int_{V_\mathrm{P}} \varrho(\boldsymbol{r})\, \mathrm{e}^{-\mathrm{i}\boldsymbol{K}\cdot\boldsymbol{r}}\, \mathrm{d}V \right|^2 \tag{2.14}$$

gegeben. Die Integration erstreckt sich über das ganze Probenvolumen V_P.

Für die Strukturbestimmung ist nicht der Absolutwert der gestreuten Intensität, sondern deren Variation mit der Beobachtungsrichtung von Bedeutung. Diese Information steckt ausschließlich in $|\mathcal{A}(\boldsymbol{K})|^2$, das wir in der weiteren Diskussion betrachten. Der Einfachheit halber werden wir dennoch von der „Streuintensität" sprechen. Nun führen wir die Berechnung der „Streuintensität" mit Hilfe des Konzepts des reziproken Gitters durch. Wir setzen die Entwicklung (2.6) der Streudichte $\varrho(\boldsymbol{r})$ nach den Basisvektoren des reziproken Gitters in (2.14) ein und erhalten

$$|\mathcal{A}(\boldsymbol{K})|^2 = \left| \sum_{h,k,l} \varrho_{hkl} \int_{V_\mathrm{P}} \mathrm{e}^{\mathrm{i}(\boldsymbol{G}-\boldsymbol{K})\cdot\boldsymbol{r}}\, \mathrm{d}V \right|^2 . \tag{2.15}$$

Wie in dieser Gleichung werden wir auch im weiteren Verlauf der Diskussion die Indizes hkl beim reziproken Gittervektor meist weglassen. Da die Funktion $\exp[\mathrm{i}(\boldsymbol{G}-\boldsymbol{K})\cdot\boldsymbol{r}]$ oszilliert, mitteln sich die Beiträge weg, wenn die Integration über ein Streuvolumen erfolgt, dessen Ausdehnung groß gegen die Perioden der Oszillationen ist. Vom physikalischen Standpunkt betrachtet, bedeutet dies ein Auslöschen der Streubeiträge der einzelnen Atome durch Interferenz. Ausgenommen sind die speziellen Beobachtungsrichtungen, für die die *Streubedingung*

$$\boldsymbol{K} = \boldsymbol{G} \tag{2.16}$$

erfüllt sein muß. Die Exponentialfunktion in (2.15) nimmt für diese ausgezeichnete Richtung den Wert Eins an. Entlang dieser Richtung ist die Überlagerung der Streuwellen aufgrund der regelmäßigen Anordnung der Atome konstruktiv. Will man einen „Beugungsreflex", d.h. gestreute Strahlung beobachten, dann muß bei fester Kristallorientierung bezüglich des einfallenden Strahls die Beobachtungsrichtung so lange verändert werden, bis diese Streubedingung erfüllt ist. Formal läßt sich das geschilderte Ergebnis durch die Beziehung

$$\int_{V_\mathrm{P}} \mathrm{e}^{\mathrm{i}(\boldsymbol{G}-\boldsymbol{K})\cdot\boldsymbol{r}}\, \mathrm{d}V \simeq \begin{cases} V_\mathrm{P} & (\boldsymbol{K} = \boldsymbol{G}) \\ 0 & (\boldsymbol{K} \neq \boldsymbol{G}) \end{cases} \tag{2.17}$$

ausdrücken. Im Grenzfall eines unendlich großen Probenvolumens steht auf der linken Seite von (2.17) die Fourier-Darstellung der dreidimensionalen δ-Funktion. In diesem

Fall tritt nur dann Streustrahlung auf, wenn die Bedingung (2.16) streng erfüllt ist. Bei endlichem Probenvolumen ist diese Bedingung etwas „aufgeweicht". Auf die wichtigen Konsequenzen, die sich daraus ergeben, werden wir sogleich noch eingehen.

Berechnen wir nun die Streuintensität mit Hilfe von (2.15), so stellen wir fest, daß die Summation entfällt, denn jeder reziproke Gittervektor $\boldsymbol{G}_{hkl}$ repräsentiert gerade *einen* Fourier-Koeffizienten ϱ_{hkl} der Streudichte. Damit erhalten wir für die Streuintensität den einfachen Ausdruck

$$|\mathcal{A}(\boldsymbol{K} = \boldsymbol{G})|^2 = |\varrho_{hkl}|^2 \, V_\mathrm{P}^2 \, . \tag{2.18}$$

Die quadratische Abhängigkeit vom Probenvolumen ist auf den ersten Blick überraschend, da zu erwarten ist, daß die Intensität der gestreuten Strahlung proportional zur Anzahl der streuenden Atome ansteigt. Die Ursache für den scheinbaren Widerspruch ist offensichtlich: Wir haben bei der Herleitung dieses Ergebnisses zu starke Vereinfachungen gemacht. Generell läßt sich sagen, daß bei einem endlichen Volumen V_P die Streubedingung gelockert ist und mit abnehmendem Probenvolumen die Beugungsreflexe verbreitert werden. In Abhängigkeit vom Winkel durchläuft die Streuintensität eines Beugungsreflexes ein Maximum, dessen *Höhe* durch (2.18) gegeben ist, dessen Breite jedoch mit dem Probenvolumen abnimmt. Eine ausführliche Rechnung zeigt, daß die Breite des Reflexes proportional zu V_P^{-1} ist. Daraus folgt, daß die *integrale Intensität* des Beugungsreflexes wie erwartet linear mit dem Streuvolumen ansteigt, also proportional zur Zahl der Streuzentren ist. Aus der Verbreiterung der Streureflexe kann daher unter geeigneten Bedingungen auf die Größe der streuenden Kristallite geschlossen werden.

Untersucht man einen großen, perfekten Kristall, so sind die Beugungsreflexe dennoch nicht beliebig scharf. Im allgemeinen liegt dies nicht an der fehlenden Auflösung der Meßapparatur (oder an der begrenzten Kohärenz der einfallenden Strahlung), sondern an der endlichen Eindringtiefe des auftreffenden Strahls. Ist das Volumen, das zur Beugung beiträgt, relativ groß, so ist die anfangs gemachte Annahme einer vernachlässigbaren Intensität der gestreuten Strahlung nicht mehr erfüllt. Abhängig von der Streustärke der einzelnen Kristallebenen wird im Falle von Röntgenstrahlen an jeder Netzebene typischerweise etwa $10^{-5} - 10^{-3}$ der einfallenden Intensität gebeugt. Insgesamt führt dies zur vollständigen Reflexion des einfallenden Strahls, wobei die Reflexbreite durch die Eindringtiefe festgelegt ist. Erst wenn die untersuchten Kristalle kleiner als die Eindringtiefe sind, ist das Probenvolumen identisch mit dem Streuvolumen, über das dann die Reflexbreite Auskunft geben kann.

In einer vollständigen Theorie der Streuung muß auch berücksichtigt werden, daß die gebeugte Welle durch die Netzebenen teilweise wieder in Richtung des einfallenden Strahls gestreut wird, daß die Strahlung eine gewisse Absorption erfährt und der Brechungsindex sich von dem des Vakuums unterscheidet. In der sogenannten *dynamischen Theorie*, die wesentlich von *Ewald* vorangetrieben wurde, sind diese Effekte eingearbeitet. Natürlich ist sie wesentlich aussagekräftiger, aber auch komplizierter als die hier vorgeführte einfache Theorie der Beugung.

Läßt man Strahlung mit dem wohldefinierten Wellenvektoren $\boldsymbol{k}_0$ auf einen zufällig-orientierten Kristall fallen, so wird in den meisten Fällen keine Reflexion der Röntgenstrahlen auftreten, da $\boldsymbol{K} = \boldsymbol{G}$ für keine Beobachtungsrichtung erfüllt sein wird, d.h. der Strahl wird ohne besondere Wechselwirkung den Kristall durchqueren. Bei vorgegebener Richtung des einfallenden Strahls muß der Kristall passend ausgerichtet werden und die Beobachtung unter einer geeigneten Richtung erfolgen. Die erforderliche Orientierung des Kristalls in Bezug auf die einfallende Strahlung läßt sich durch die Konstruktion der

sogenannten *Ewald-Kugel* geometrisch veranschaulichen. Wie in Abb. 2.12 gezeigt, stellt man das reziproke Gitter dar und gibt $\boldsymbol{k}_0$ vor[1]. Dann schlägt man einen Kreis (bzw. eine Kugel im dreidimensionalen Fall) mit dem Radius k_0 um den Anfangspunkt des Wellenvektors. Die Streubedingung (2.16) ist erfüllt, wenn die Ewald-Kugel durch einen Punkt des reziproken Gitters geht, und ein Beugungsreflex kann beobachtet werden.

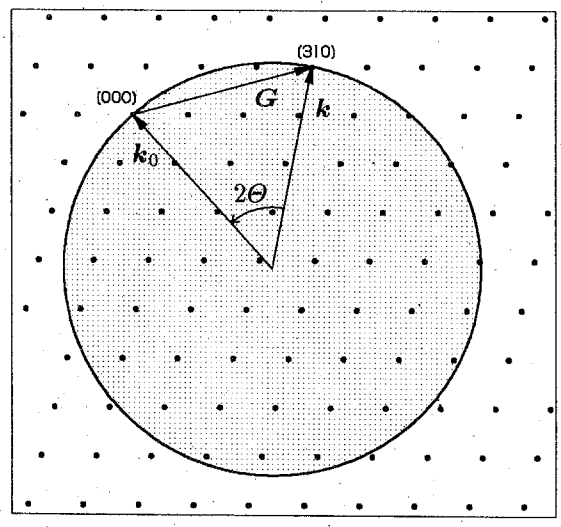


Abb. 2.12: Zweidimensionale Darstellung der dreidimensionalen Ewald-Kugel. Bei der vorgegebenen relativen Orientierung des Kristalls bezüglich des einfallenden Strahls erfolgt Streuung an den (310)-Netzebenen.

Bei der Fourier-Zerlegung des Gitters im letzten Abschnitt sind wir von einem unendlich ausgedehnten, perfekten Kristallgitter ausgegangen. Diese Vorgehensweise führte zu mathematisch scharfen reziproken Gitterpunkten. Betrachtet man aber Kristalle mit endlicher Abmessung, so sind die Punkte des reziproken Gitters „verschmiert". Je kleiner das beugende Streuvolumen, um so unschärfer werden die Gitterpunkte und um so breiter wird der beobachtete Reflex. In gleicher Weise ließe sich die begrenzte Eindringtiefe der Strahlung in der Abb. 2.12 miteinbeziehen. Auch die endliche Frequenzschärfe der einfallenden Strahlung führt zu einer Aufweichung der Streubedingung. Bei der Konstruktion der Ewald-Kugel ließe sich dieser Effekt durch eine endliche Strichbreite des Kreises bzw. der Dicke der Kugelschale berücksichtigen.

Wir wollen noch einen anderen Aspekt der erzielten Ergebnisse beleuchten. Der Experimentator gibt durch die räumliche Anordnung von Strahlungsquelle und Strahlungsdetektor und durch die Orientierung der Probe den Streuvektor $\boldsymbol{K}$ vor. Über die Streubedingung ist damit auch der mitwirkende reziproke Gittervektor $\boldsymbol{G}$ festgelegt, mit dem das Zahlentripel (hkl), die Fourier-Komponente der Streudichte und die dazugehörende Netzebenenschar verknüpft sind. Die beobachtete Streuung erfolgt also an einer periodischen Schwankung der Streudichte, deren Periode durch $d_{hkl} = 2\pi/|\boldsymbol{G}_{hkl}|$ gegeben ist. Der Zusammenhang zwischen Streuvektor und Periode der Dichteschwankung erlaubt eine anschauliche Interpretation des Streuexperimentes. Ist θ der Winkel zwischen der Einfallsrichtung der Welle und der streuenden Dichteschwankung (vgl. Abb. 2.13), so gilt wegen $k_0 = k$ für den Betrag des Streuvektors: $K = |\boldsymbol{k} - \boldsymbol{k}_0| = 2k_0 \sin\theta = (4\pi \sin\theta)/\lambda$.

[1] Bei dieser Art der Konstruktion legt man fest, daß die durchgehende Intensität vom reziproken Gittervektor $\boldsymbol{G} = (0,0,0)$ herrührt. Bei Experimenten, bei denen die Streuintensität in Rückwärtsrichtung überwiegt, ist eine leicht modifizierte Konstruktion üblich.

Setzt man diese Beziehung und Gleichung (2.13) in die Streubedingung (2.16) ein, so erhält man die bekannte *Bragg-Bedingung*:

$$2\,d_{hkl}\,\sin\theta = p\,\lambda\,. \tag{2.19}$$

Die ganze Zahl p auf der rechten Seite wurde noch eingefügt, weil konstruktive Interferenz auch bei ganzzahligen Vielfachen der Wellenlänge möglich ist, wie schon eine elementare Herleitung von (2.19) zeigt. Das steht nicht im Widerspruch zu unserem Ergebnis, da auch die Definition des reziproken Gitters (vgl. Herleitung von Gleichung (2.9)) einen derartigen Faktor enthalten kann. Den Winkel θ, unter dem der Beugungsreflex beobachtet wird, bezeichnet man oft als *Glanzwinkel*.

Wir wollen den wichtigen Zusammenhang zwischen Streuamplitude und der periodischen Schwankung der Streudichte anhand eines einfachen Beispiels verdeutlichen. In Abb. 2.13 ist die räumliche Verteilung einer gedachten Streudichte durch unterschiedlich dichte Punktierung angedeutet. Sie repräsentieren Netzebenen mit unterschiedlichen Millerschen Indizes. Die gezeichnete Streudichte weist zwei verschiedene Perioden auf und ist daher in der Lage bei vorgegebenem Wellenvektor der einfallenden Strahlung die Welle bereits für den Wert $p = 1$ in zwei Richtungen zu beugen. Die Perioden d_a bzw. d_b der Dichteschwankungen und der Wellenvektor $\boldsymbol{k}_0$ legen die Richtungen der gebeugten Wellen und somit die Vektoren $\boldsymbol{k}_a$ und $\boldsymbol{k}_b$ fest. In dieser Darstellung ist der Wellenvektor der einfallenden Strahlung so gewählt, daß die gebeugten Strahlen gleichzeitig auftreten. Beobachtet man eine ausreichend große Anzahl von Reflexen, so läßt sich die tatsächliche Streudichte eines Festkörpers durch Überlagerung der einzelnen Fourier-Komponenten der Streudichten rekonstruieren.

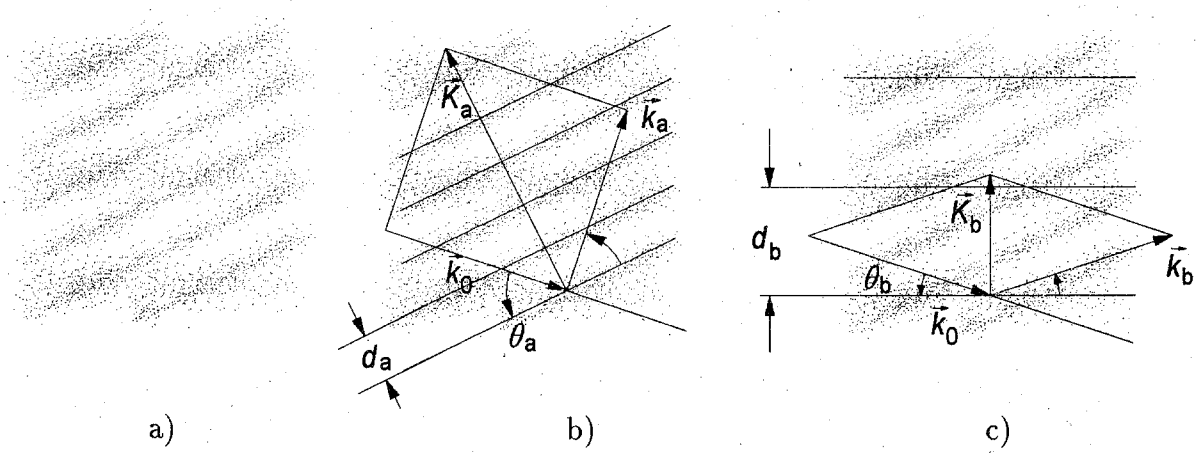


Abb. 2.13: Bragg-Reflexion an Streudichteschwankungen.
a) Streudichte,
b) Streuung an den schräg verlaufenden Dichteschwankungen,
c) Streuung an den waagerecht verlaufenden Dichteschwankungen.
Mit $\boldsymbol{k}_0$ sind der Wellenvektor der einfallenden, mit $\boldsymbol{k}_a$ bzw. $\boldsymbol{k}_b$ die Wellenvektoren der gestreuten Strahlung bezeichnet. Für Abstände der Streudichteschwankungen stehen d_a und d_b, für die Streuvektoren $\boldsymbol{K}_a$ und $\boldsymbol{K}_b$ und für die jeweiligen Einfallswinkel θ_a und θ_b.

Phasenproblem bei Streuexperimenten

Bei unserer bisherigen Diskussion der Streuung am periodischen Gitter wurde nicht berücksichtigt, daß die Streuamplitude $\mathcal{A}(\boldsymbol{K})$ nicht wirklich bestimmt werden kann, da die zur Verfügung stehenden Strahlungsdetektoren nicht in der Lage sind, die Phase der gestreuten Strahlung zu registrieren. Es kann daher nicht $\mathcal{A}(\boldsymbol{K})$, sondern nur die Intensität und damit $|\mathcal{A}(\boldsymbol{K})|^2$ gemessen werden. Dadurch geht wichtige Information verloren, und die Fourier-Transformation (2.5) kann nicht unmittelbar ausgeführt werden. Aus der gemessenen Streuintensität läßt sich daher nicht eindeutig die Struktur herleiten. Dies wird am folgenden einfachen Beispiel, der *Friedelschen Regel*, deutlich. Unter der Voraussetzung, daß die Absorption vernachlässigt werden kann, ist die Streudichte $\varrho(\boldsymbol{r})$ eine reelle Größe, und wir können nach (2.4) für sie $\mathcal{A}(\boldsymbol{K}) = \mathcal{A}^*(-\boldsymbol{K})$ schreiben. Für die Intensität gilt daher $|\mathcal{A}(\boldsymbol{K})|^2 = |\mathcal{A}^*(-\boldsymbol{K})|^2 = |\mathcal{A}(-\boldsymbol{K})|^2$, d.h. das Bild der Beugungsreflexe besitzt unabhängig von der tatsächlichen Struktur immer Inversionssymmetrie. Dem Experiment kann man nicht entnehmen, ob sich „Vorder-" oder „Rückseite" eines Kristalls unterscheiden. Diese Einschränkung stört kaum, wenn die Struktur von Kristallen mit relativ einfacher Basis ermittelt wird. Dort hilft in vielen Fällen „Raten" oder ein Vergleich mit bekannten Kristallen. Wie im übernächsten Abschnitt gezeigt wird, erlauben Beugungsexperimente nicht nur die Bestimmung des Punktgitters, sondern auch der Basis, d.h. die Bestimmung der Anordnung der Atome in der Elementarzelle. Bei der Untersuchung von komplex aufgebauten Elementarzellen macht sich der Verlust der Phaseninformation besonders nachteilig bemerkbar.

Es existiert die Möglichkeit, das Phasenproblem bei der Röntgenbeugung zu reduzieren, indem man sich einen zweiten, unabhängigen Datensatz beschafft. Bei großen organischen Molekülen führt oft die Methode der „isomorphen Substitution" zum Ziel. Dabei werden solche Schwermetallatome eingebaut, welche die molekulare Struktur nur unwesentlich verändern. Aufgrund ihrer hohen Elektronenzahl streuen diese Atome besonders stark und liefern so „Referenzwellen", die sich den Streuwellen der „normalen" Atome überlagern. Bei der Strukturbestimmung werden dann die Beugungsbilder mit und ohne Substitution verglichen. Auf diese Weise konnte die Struktur von sehr komplexen Proteinmolekülen aufgeklärt werden. Eine weitere Möglichkeit, die fehlende Phaseninformation zu ersetzen, besteht in der Messung bei zwei verschiedenen Wellenlängen, wobei eine Wellenlänge in der Nähe der resonanten Absorption einer Atomsorte liegen muß. In diesem Fall ist die Streudichte und damit auch die Streuamplitude nicht mehr reell, sondern komplex. Man erhält einen weiteren Datensatz, der zur Bestimmung der Struktur herangezogen werden kann. Bei der Neutronenstreuung kann gelegentlich Nutzen aus der Tatsache gezogen werden, daß die Isotope *eines* Elements meist ganz unterschiedliche Streueigenschaften besitzen. Messungen an Proben mit abweichender Isotopenzusammensetzung ergeben daher zwei unabhängige Datensätze und liefern verschiedene Strukturinformationen.

Debye-Waller-Faktor

Wir haben bereits erwähnt, daß endliche Kristalldimensionen und begrenzte Eindringtiefe der einfallenden Strahlung zu einer Verbreiterung der Beugungsreflexe führen. Interessant ist in diesem Zusammenhang die Frage, ob die Reflexbreite nicht auch von der Temperatur abhängig ist. Man könnte diese Frage zunächst bejahen, da die thermische Bewegung die Lage der Atome sozusagen verschmiert und dadurch eine Veränderung des reziproken Gitters bewirken könnte. Im zeitlichen Mittel wird dadurch jedoch nicht die Periodizität

des Gitters verändert. Experimentell findet man, daß die Temperaturbewegung die Reflexbreite *nicht* beeinflußt, aber deren *Intensität reduziert.* Die Temperaturabhängigkeit der Streuintensität läßt sich durch die Gleichung

$$I(T) = I_0 D = I_0 \, \mathrm{e}^{-\langle (\boldsymbol{K}\cdot\boldsymbol{u})^2 \rangle} \tag{2.20}$$

ausdrücken, wobei I_0 die Streuintensität des starren Gitters ist. Der Faktor D wird als *Debye-Waller-Faktor* bezeichnet und enthält neben dem Streuvektor $\boldsymbol{K}$ noch die Auslenkung $\boldsymbol{u}(T)$ der Atome aus ihrer Ruhelage (vgl. Kap. 5). Ist die Auslenkung der Atome isotrop, so erhält man für den Debye-Waller-Faktor den Ausdruck

$$D = \mathrm{e}^{-\frac{1}{3}K^2\langle u^2\rangle} = \mathrm{e}^{-K^2 B(T)} \, . \tag{2.21}$$

Der Faktor $1/3$ rührt hierbei von der Mittelung über alle Winkel zwischen Streuvektor und Auslenkung her. Bei Zimmertemperatur ist das mittlere Auslenkungsquadrat $\langle u^2 \rangle$ und damit auch $B(T)$ proportional zur Temperatur. Gehen wir zum absoluten Nullpunkt, so strebt B einem endlichen Wert zu und spiegelt somit die Nullpunktsbewegung der Atome wider.

Eine eingehende Untersuchung der Streueigenschaften von Festkörpern zeigt, daß überraschenderweise eine sehr schwache Streuintensität auch dann beobachtet wird, wenn die Streubedingung (2.16) nicht erfüllt ist. Dieser „Untergrund", d.h. die Streuintensität zwischen den Reflexen, steigt mit zunehmender Temperatur an. Die Ursache hierfür werden wir in Kap. 5.4 im Zusammenhang mit der inelastischen Streuung noch etwas genauer diskutieren.

Strukturfaktor

Mit Hilfe der bisherigen Überlegungen zur Streuung der Kristalle läßt sich voraussagen, welche Reflexe aufgrund der Gitterperiodizität auftreten *können*, doch ist damit noch keine Aussage über die Stärke dieser Reflexe verbunden. Die *Intensität* der Reflexe wird nach (2.18) durch die Fourier-Koeffizienten ϱ_{hkl} der Streudichte bestimmt und enthält somit Information über den Aufbau der Basis. Ist die Basis nicht einatomig, so führen Interferenzen zwischen den Beiträgen der Atome in der Elementarzelle zu einer Variation der Reflexintensitäten. Es liegt nahe, bei der Berechnung der Streuintensität die Beiträge der einzelnen Atome getrennt zu betrachten und sie dann aufzusummieren. Wie in Abb. 2.14 schematisch gezeigt, spalten wir hierzu den Ortsvektor $\boldsymbol{r}$ auf, in dem wir für ihn $\boldsymbol{r} = \boldsymbol{R}_m + \boldsymbol{r}_\alpha + \boldsymbol{r}'$ schreiben. $\boldsymbol{R}_m$ legt die Lage der betrachteten Elementarzelle fest und ist aufgrund der Translationsinvarianz des Kristalls für unsere weiteren Betrachtungen ohne Bedeutung. Die Anordnung der Atome α innerhalb der Elementarzelle wird durch $\boldsymbol{r}_\alpha$ definiert. Die Streueigenschaften der einzelnen Atome schließlich wird durch die Verteilung der Streudichte im Atom bestimmt. Dabei müssen wir über den Ortsvektor $\boldsymbol{r}'$ integrieren, der die Lage des Streuvolumens $\mathrm{d}V'$ im Atom angibt. Wir setzen $\boldsymbol{r}$ ein und erhalten anstelle von (2.7) den Ausdruck

$$\begin{aligned} \varrho_{hkl} = \frac{1}{V_\mathrm{Z}} \int\limits_{V_\mathrm{Z}} \varrho(\boldsymbol{r}) \, \mathrm{e}^{-\mathrm{i}\boldsymbol{G}\cdot\boldsymbol{r}} \mathrm{d}V &= \frac{1}{V_\mathrm{Z}} \sum_\alpha \mathrm{e}^{-\mathrm{i}\boldsymbol{G}\cdot\boldsymbol{r}_\alpha} \int\limits_{V_\alpha} \varrho_\alpha(\boldsymbol{r}') \, \mathrm{e}^{-\mathrm{i}\boldsymbol{G}\cdot\boldsymbol{r}'} \mathrm{d}V' \\ &= \frac{1}{V_\mathrm{Z}} \sum_\alpha \widetilde{a}_\alpha(\boldsymbol{G}) \, \mathrm{e}^{-\mathrm{i}\boldsymbol{G}\cdot\boldsymbol{r}_\alpha} \, . \end{aligned} \tag{2.22}$$

Die Integration der Streudichte erfolgt über das Volumen V_α der einzelnen Atome, die Summation läuft über die Zahl der Atome in der Basis. In der letzten Zeile haben wir als Abkürzung den *Atom-Strukturfaktor* $\widetilde{a}(\boldsymbol{G})$ eingeführt, der die spezifischen Eigenschaften des betrachteten Streuzentrums beschreibt, die wiederum von der einfallenden Strahlung abhängen. Wir wenden uns zunächst der Auswertung der Summe zu und diskutieren die Bedeutung des Atom-Strukturfaktors ausführlich im nächsten Abschnitt.

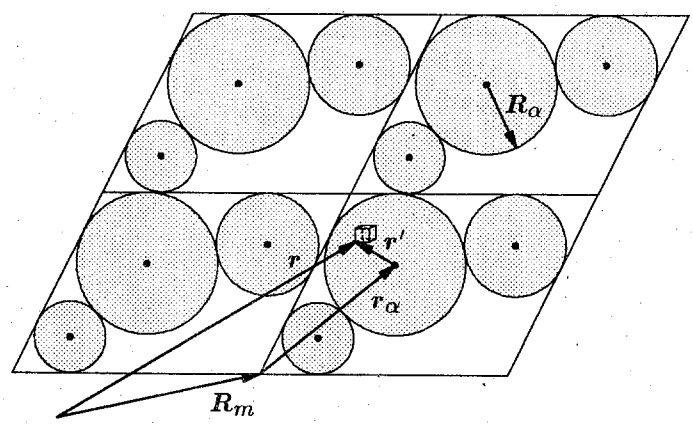


Abb. 2.14: Zur Ableitung von Struktur- und Atom-Strukturfaktor. Die Atome werden als kugelförmige Gebilde mit dem Radius R_α betrachtet. Die Lage der Elementarzelle wird durch $\boldsymbol{R}_m$, die der Atomzentren in der Zelle durch $\boldsymbol{r}_\alpha$ und das Streuvolumen $\mathrm{d}V'$ innerhalb des Atoms durch $\boldsymbol{r}'$ festgelegt.

Die Summation vereinfacht sich, wenn wir die Lage der Atome in der Elementarzelle mit Hilfe der Basisvektoren des realen Gitters ausdrücken. Für den Ortsvektor der Atome schreiben wir daher $\boldsymbol{r}_\alpha = u_\alpha \boldsymbol{a}_1 + v_\alpha \boldsymbol{a}_2 + w_\alpha \boldsymbol{a}_3$, wobei die Komponenten u_α, v_α und w_α durch die Struktur der Basis vorgegeben sind. Führen wir in (2.22) das Skalarprodukt aus, so erhalten wir für die *Strukturamplitude* oder den *Strukturfaktor* $\mathcal{S}_{hkl} = \varrho_{hkl} V_\mathrm{Z}$ das wichtige Ergebnis

$$\mathcal{S}_{hkl} = \sum_\alpha \widetilde{a}_\alpha(\boldsymbol{G})\, \mathrm{e}^{-2\pi \mathrm{i}(hu_\alpha + kv_\alpha + lw_\alpha)} \,. \tag{2.23}$$

Als einfache Beispiele für die Berechnung der auftretenden Streuintensität betrachten wir primitive und kubisch raumzentrierte Gitter. Ist die Basis des Gitters einatomig, dann entfällt die Summation, und wir erhalten $\mathcal{S}_{hkl} = \widetilde{a}(\boldsymbol{G})$, wenn wir den Ursprung der Elementarzelle in das Atomzentrum legen. Die Winkelabhängigkeit der Streuintensität wird dann nur durch den Atom-Strukturfaktor, d.h. durch die Streueigenschaften der einzelnen Atome bestimmt. Wählt man eine andere Koordinate für die Lage des Atoms in der Elementarzelle, so tritt ein Phasenfaktor auf, der bei der Betragsbildung zur Berechnung der Intensität wegfällt.

Für den Strukturfaktor des kubisch primitiven Cäsiumchlorids mit der zweiatomigen Basis, findet man durch Einsetzen:

$$\mathcal{S}_{hkl} = \begin{cases} \widetilde{a}_\mathrm{Cs} + \widetilde{a}_\mathrm{Cl} & h+k+l \text{ gerade} \\ \widetilde{a}_\mathrm{Cs} - \widetilde{a}_\mathrm{Cl} & h+k+l \text{ ungerade.} \end{cases} \tag{2.24}$$

Kristalle mit dieser Struktur verursachen zwei Kategorien von Reflexen mit stark unterschiedlicher Intensität. Da sich die Beiträge der beiden Ionensorten entweder addieren oder subtrahieren, treten abhängig von der Quersumme der Indizes starke und schwache Reflexe auf. Bei der Röntgenbeugung an Cäsiumjodid ist dieser Effekt besonders ausgeprägt, da Cäsium- und Jodionen die gleiche Anzahl von Elektronen besitzen. Wie wir

im nächsten Abschnitt sehen werden, ist der Atom-Strukturfaktor bei der Röntgenstreuung direkt mit der Elektronenzahl verknüpft, so daß bei CsJ die Atom-Strukturfaktoren der beiden Ionensorten nahezu gleich sind. Dies bedeutet, daß Reflexe mit einer ungeraden Summe der Millerschen Indizes kaum noch in Erscheinung treten.

Bei kubisch raumzentrierten Gittern mit einfacher Basis findet man

$$\mathcal{S}_{hkl} = \widetilde{a}\left[1 + e^{-i\pi(h+k+l)}\right] = \begin{cases} 2\widetilde{a} & h+k+l \text{ gerade,} \\ 0 & h+k+l \text{ ungerade.} \end{cases} \tag{2.25}$$

Dieses erstaunliche Ergebnis besagt, daß keine Reflexe mit ungerader Summe der Millerschen Indizes auftreten sollten. Insbesondere sollte auch der Reflex der (100)-Ebenen fehlen. Wie wir Abb. 2.15 entnehmen können, läßt sich hierfür eine anschauliche Erklärung finden. Zwischen den Strahlen, die an den benachbarten (100)-Ebenen gestreut werden, besteht der Gangunterschied 2π, wenn die Streubedingung (2.16) erfüllt ist, so daß sich die einzelnen Strahlen konstruktiv überlagern. Die Beiträge dieser Ebenen werden jedoch durch die dazwischen liegenden (200)-Ebenen kompensiert, da deren Streubeiträge bei dem vorgegebenen Streuwinkel gegenüber den (100)-Ebenen die Phasendifferenz π besitzen und die gleiche Intensität aufweisen. Der (200)-Reflex, der bei einem größeren Winkel auftritt, wird nicht gelöscht, da sich dort die Streubeiträge der (100)- und (200)-Ebenen konstruktiv überlagern.

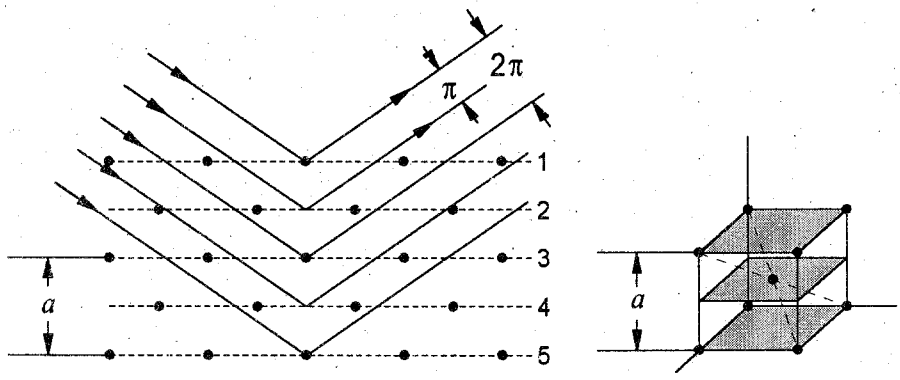


Abb. 2.15: Zum Verschwinden des (100)-Reflexes bei kubisch raumzentrierten Gittern. Die Streubeiträge der Ebenen $1, 3, \ldots$ verstärken sich durch konstruktive Interferenz, werden aber durch die Beiträge der Zwischenebenen $2, 4, \ldots$ wieder gelöscht. Rechts, zur Erinnerung, die entsprechende Elementarzelle.

Das vollständige Verschwinden von bestimmten Reflexen wollen wir noch etwas eingehender erläutern. Der Übergang vom realen zum reziproken Gitter wird über Transformationsvorschriften für die Basisvektoren vollzogen. Das Ergebnis der Transformation hängt von der Wahl der Ausgangselementarzelle ab. Dies soll anhand des zweidimensionalen quadratischen Gitters der Abb. 2.16 verdeutlicht werden. Im linken Gitter ist eine primitive Elementarzelle eingezeichnet, die zu dem darunter dargestellten reziproken Gitter führt. In der rechten Bildhälfte wurde eine Elementarzelle gewählt, die in Analogie zum kubisch raumzentrierten Gitter noch einen Gitterpunkt im Zentrum enthält. Die Anwendung der gleichen Transformationsvorschrift wie bei der primitiven Elementarzelle führt zu dem darunter gezeichneten reziproken Gitter, dessen Gitterpunkte nun entsprechend enger liegen. Die Punkte (10) und (01) beispielsweise oder allgemeiner alle Gitterpunkte mit ungeraden Werten von $h + k$ kommen auf der linken Seite nicht vor. Die reziproken

Gitterpunkte mit geraden Werten von $h + k$ dagegen haben Entsprechungen im reziproken Gitter des linken Bildes. Der (11)-Punkt des rechten Bildes entspricht zum Beispiel (10) im linken Bild. Gehen wir von einem nicht-primitiven Gitter aus, wählen also eine „zu große" Elementarzelle, dann treten zusätzliche reziproke Gitterpunkte auf. Da ihr Strukturfaktor verschwindet, können sie in Streuexperimenten nicht beobachtet werden. Die gleiche Ursache hat auch das oben diskutierte Fehlen des (100)-Reflexes bei kubisch raumzentrierten Gittern. Diese einfache Betrachtung zeigt, daß die Wahl der Elementarzelle erheblichen Einfluß auf die Zuordnung (und Bezeichnung) der Beugungsreflexe hat, daß aber die Meßergebnisse natürlich nicht von der theoretischen Beschreibung abhängen.

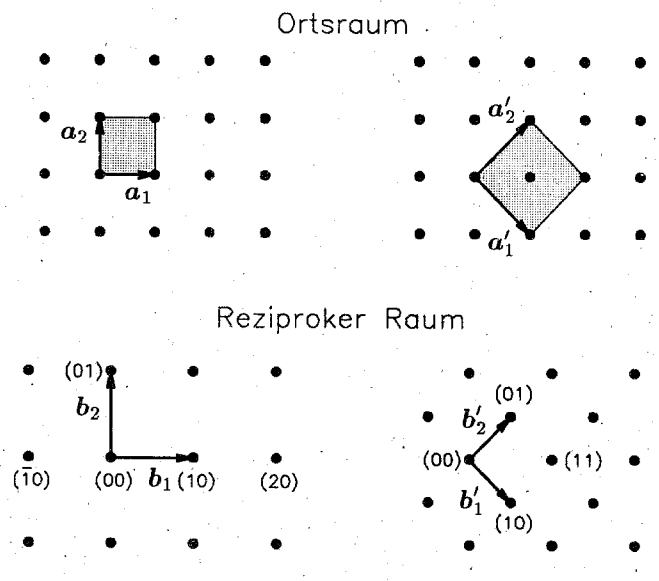


Abb. 2.16: Zum Zusammenhang zwischen reziprokem Gitter und der Form der Elementarzelle des realen Gitters. Links wurde von einer primitiven Elementarzelle ausgegangen und das darunter liegende reziproke Gitter konstruiert. Rechts dagegen wurde ein „zentriertes Quadratgitter" als Elementarzelle gewählt. Nun treten im reziproken Gitter zusätzliche Gitterpunkte auf, und die Benennung der Gitterpunkte verändert sich.

Atom-Strukturfaktor

Bei der Diskussion der Streuintensität haben wir bisher das *Streuvermögen* der *einzelnen Atome* außer acht gelassen. Nun sehen wir uns den Atom-Strukturfaktor

$$\widetilde{a}_\alpha(\boldsymbol{K}) = \int\limits_{V_\alpha} \varrho_\alpha(\boldsymbol{r})\, \mathrm{e}^{-\mathrm{i}\boldsymbol{K}\cdot\boldsymbol{r}}\, \mathrm{d}V \tag{2.26}$$

genauer an, der durch (2.22) definiert ist. Wir haben den entsprechenden Ausdruck nochmals hingeschrieben und dabei den reziproken Gittervektor $\boldsymbol{G}$ durch den Streuvektor $\boldsymbol{K}$ ersetzt. Für die Beschreibung der Beugungsreflexe bei Kristallen ist dieser Tausch ohne Belang, da dort $\boldsymbol{K} = \boldsymbol{G}$ erfüllt sein muß. Wie wir sehen werden, gilt diese Voraussetzung nicht für amorphe Festkörper. Dennoch spielt auch dort der Atom-Strukturfaktor eine wichtige Rolle, da er die Winkelabhängigkeit der Streuintensität bestimmt. Um etwas übersichtlichere Ausdrücke zu erhalten, benutzen wir für den Ortsvektor die Bezeichnung $\boldsymbol{r}$ anstelle von $\boldsymbol{r}'$.

Bevor wir uns der Diskussion der Bedeutung des Atom-Strukturfaktors bei der Röntgenbeugung zuwenden, wollen wir noch einen Blick auf die Neutronenstreuung werfen, da dort besonders einfache Verhältnisse vorliegen. Sieht man vom Spezialfall der Atome mit magnetischem Moment ab, so werden Neutronen ausschließlich an den Kernen gestreut. Diese sind klein im Vergleich zur Wellenlänge der gestreuten Strahlung, welche

bei Beugungsexperimenten vergleichbar mit dem Gitterabstand ist. Dies bedeutet, daß der Exponentialfaktor in (2.26) den Wert Eins annimmt. Die Integration führt somit zu einem konstanten, richtungsunabhängigen Atom-Strukturfaktor $\tilde{a}(\boldsymbol{K}) = -b$, dessen Wert die spezifischen Streueigenschaften des betrachteten Kerns widerspiegelt. Üblicherweise wird diese Konstante als *Streulänge* bezeichnet. Das Vorzeichen ist durch Konvention festgelegt und deutet an, daß die Wechselwirkung zwischen Neutron und Kern in den meisten Fällen attraktiv ist. Es spielt aber bei unseren Überlegungen keine Rolle, da die Streuintensität durch $|b|^2$ bestimmt wird. Sowohl das Vorzeichen als auch der Absolutwert der Streulänge schwanken nicht nur von Element zu Element, sondern auch von Isotop zu Isotop.

Völlig andere Verhältnisse herrschen bei der Röntgenstreuung. Da die Ausdehnung des Streuobjektes, d.h. die Elektronenwolke der Atome, vergleichbar mit der Wellenlänge der einfallenden Strahlung ist, treten innerhalb eines Atoms Interferenzen zwischen den Streubeiträgen der unterschiedlichen Volumenelemente der Elektronenverteilung auf. Gleichung (2.26) führt daher zu einem Ergebnis, das sich stark von der Neutronenstreuung unterscheidet. Im Falle der Röntgenstreuung bezeichnet man den Atom-Strukturfaktor $\tilde{a}(\boldsymbol{K})$ als *atomaren Streufaktor* oder als *Atomformfaktor* und kürzt ihn mit $f(\boldsymbol{K})$ ab. In der folgenden Diskussion nehmen wir zur Vereinfachung an, daß die Atome als kugelförmige Gebilde betrachtet werden können, d.h. daß ihre Ladung kugelsymmetrisch innerhalb eines bestimmten Radius verteilt ist. Die Ladungsdichte zwischen den Atomen wird als vernachlässigbar klein angenommen. Der Streufaktor hängt dann nur von $|\boldsymbol{K}|$ ab, so daß wir $\boldsymbol{K}\cdot\boldsymbol{r} = Kr\cos\vartheta$ schreiben können, wobei ϑ den von $\boldsymbol{K}$ und $\boldsymbol{r}$ eingeschlossenen Winkel bezeichnet. Nach der Umschreibung von (2.26) in Kugelkoordinaten erhalten wir für den atomaren Streufaktor $f_\alpha(K)$ die Gleichung

$$f_\alpha(K) = \int \varrho_\alpha(r)\,\mathrm{e}^{-\mathrm{i}\boldsymbol{K}\cdot\boldsymbol{r}}\mathrm{d}V = \int_0^{R_\alpha}\mathrm{d}r \int_0^{\pi}\mathrm{d}\vartheta \int_0^{2\pi}\mathrm{d}\varphi\,\varrho_\alpha(r)\,r^2\,\sin\vartheta\,\mathrm{e}^{-\mathrm{i}Kr\cos\vartheta}\,. \tag{2.27}$$

Die Integration über den Radiusvektor r erfolgt bis zum „Rand" R_α des Atoms. Durch die Kugelsymmetrie vereinfacht sich die Integration über die Polarwinkel ϑ und φ so stark, daß sie analytisch durchgeführt werden kann. Man erhält

$$f_\alpha(K) = \int_0^{R_\alpha} 4\pi r^2 \varrho_\alpha(r)\,\frac{\sin Kr}{Kr}\,\mathrm{d}r\,. \tag{2.28}$$

Für die vollständige Berechnung des atomaren Streufaktors muß die Ladungsdichteverteilung $\varrho(r)$ im Atom bekannt sein. Nur in den einfachsten Fällen läßt sich die Integration über den Radiusvektor analytisch ausführen, wie z.B. beim freien Wasserstoffatom im Grundzustand. In diesem Fall ist die Streudichte $\varrho(r)$ durch das Quadrat der bekannten Wellenfunktion $\psi_0(r) = (\pi a_0^3)^{-1/2}\,\exp(-r/a_0)$ bestimmt[2], d.h. es gilt

$$\varrho(r) = |\psi_0(r)|^2 = \frac{1}{\pi a_0^3}\,\mathrm{e}^{-2r/a_0}\,, \tag{2.29}$$

[2] Damit der atomare Streufaktor dimensionslos ist, wird die Ladung des Elektrons im Ausdruck für $f(\boldsymbol{K})$ üblicherweise nicht mitgenommen.

wobei a_0 für den Bohrschen Radius steht. Setzt man $\varrho(r)$ in (2.28) ein und läßt die Integrationsgrenze $R_\alpha \to \infty$ gehen, so findet man

$$f_{\mathrm{H}}(K) = \left[1 + \left(\frac{a_0 K}{2}\right)^2\right]^{-2} . \tag{2.30}$$

Diese Funktion und das Ergebnis einer numerischen Rechnung für freie Aluminiumatome ist in Abb. 2.17 gezeigt. Wie erwartet, hängt die Abstrahlung ausgedehnter kugelsymmetrischer Atome stark von der Richtung ab. Beim Aluminium sind zusätzlich auch experimentelle Werte eingetragen, die mit Hilfe von (2.22) aus Streuexperimenten an Kristallen ermittelt wurden. Offensichtlich ist im Festkörper die Abweichung von der Ladungsdichte der freien Atome relativ gering, doch liegen die gemessenen Werte systematisch unter den berechneten. Dies deutet darauf hin, daß selbst in Metallen die Ladung weitgehend, aber nicht vollständig in unmittelbarer Nähe der Kerne lokalisiert ist.

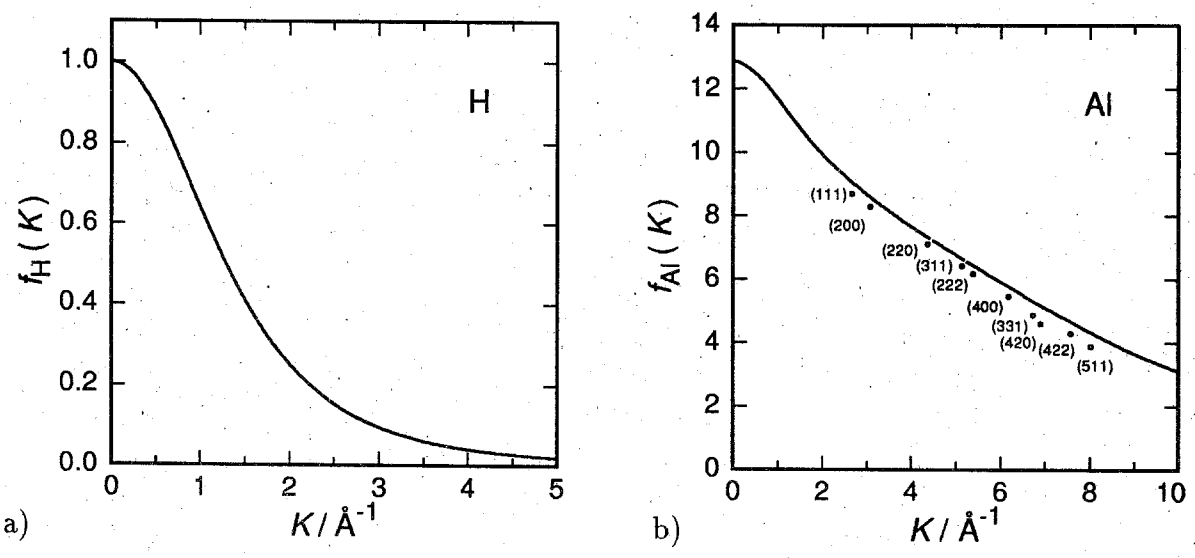


Abb. 2.17: Atomare Streufaktoren als Funktion des Streuvektors $K = (4\pi \sin\theta)/\lambda$.
a) Verlauf des atomaren Streufaktors beim freien Wasserstoffatom nach (2.30).
b) Numerisch berechneter Verlauf für Aluminiumatome. Die Kreise sind experimentelle Werte, die man aus Messungen der Streuintensität an Kristallen in unterschiedlichen Richtungen erhält. (Nach B.W. Batterman et al., Phys. Rev. **122** (1961) 68).

Wir wollen den Grenzfall $Kr \to 0$ noch etwas näher betrachten. Da in diesem Fall $(\sin Kr)/Kr \to 1$ geht, vereinfacht sich Gleichung (2.28) zu einer Integration über die Streudichte, die natürlich als Ergebnis die Gesamtzahl Z der Elektronen der streuenden Atome bzw. Ionen hat. Es gilt daher:

$$f(Kr \to 0) \approx \int_0^{R_\alpha} 4\pi r^2 \varrho(r)\, \mathrm{d}r = Z . \tag{2.31}$$

Dieser Grenzfall ist bei der Streuung in Vorwärtsrichtung realisiert, da dort $K \to 0$ geht. Da die Streuintensität proportional zu f^2 ist, ist sie bei kleinen Streuwinkeln proportional zu Z^2, d.h. sie ist ausschließlich durch die Elektronenzahl bestimmt und nimmt stark mit

der Zahl der streuenden Elektronen zu. Den gleichen Grenzfall findet man auch, wenn das streuende Atom klein ist gegen die Wellenlänge der einfallenden Strahlung, denn für $R_\alpha \ll \lambda$ ist $KR \ll 1$. In diesem Fall können wir das Atom als punktförmig betrachten und erhalten ein Ergebnis wie bei der Neutronenstreuung: Der atomare Streufaktor ist unabhängig vom Streuvektor und spiegelt nur die Streustärke des Atoms wider. Bei Röntgenstreuexperimenten gibt der Atomformfaktor somit die Streuamplitude an, die man aufgrund einer Abweichung der Ladungsverteilung des Atoms von der Punktform beobachtet. Die endliche Ausdehnung der Atome bewirkt, außer bei Vorwärtsstreuung, aufgrund der Interferenz der Streubeiträge, immer eine Reduktion der Streuintensität.

2.5 Streuung an amorphen Substanzen

Die Struktur amorpher Festkörper ist bisher noch nicht zufriedenstellend geklärt, doch werden große Anstrengungen unternommen, diese Lücke zu füllen. Wesentliche Beiträge zur Aufklärung der Struktur von Gläsern und Flüssigkeiten leisten dabei Streuexperimente. Zu deren Verständnis wollen wir die bisher diskutierte Streutheorie an die Erfordernisse der Strukturuntersuchung an diesen Substanzen anpassen. Die Vorgehensweise ist ähnlich wie bei der Herleitung des Strukturfaktors von Kristallen. Wir setzen kugelförmige Atome voraus, deren Streuvermögen durch den Atom-Strukturfaktor beschrieben wird. Damit läßt sich die Streuamplitude $\mathcal{A}(\boldsymbol{K})$, die ein Atom m mit dem Ortsvektor $\boldsymbol{r}_m$ hervorruft, in der Form $\mathcal{A}(\boldsymbol{K}) = \tilde{a}_m \exp(\mathrm{i}\boldsymbol{K}\cdot\boldsymbol{r}_m)$ schreiben. Dieser Streubeitrag interferiert mit dem aller anderen Atome n an dem Ort $\boldsymbol{r}_n$. Die Streuintensität der ganzen Probe, ausgedrückt durch $|\mathcal{A}(\boldsymbol{K})|^2$, erhalten wir, indem wir die Streuamplituden aller Atome aufsummieren und das Betragsquadrat bilden:

$$|\mathcal{A}(\boldsymbol{K})|^2 = \sum_m \tilde{a}_m(K)\,\mathrm{e}^{\mathrm{i}\boldsymbol{K}\cdot\boldsymbol{r}_m} \sum_n \tilde{a}_n(K)\,\mathrm{e}^{-\mathrm{i}\boldsymbol{K}\cdot\boldsymbol{r}_n} = \sum_{m,n} \tilde{a}_m(K)\,\tilde{a}_n(K)\,\mathrm{e}^{-\mathrm{i}\boldsymbol{K}\cdot\boldsymbol{r}_{mn}} \,. \quad (2.32)$$

Für das Auftreten von Interferenzen in der Streustrahlung sind nicht die Absolut-, sondern die Relativkoordinaten der Atome entscheidend, die im zweiten Ausdruck über den Vektor $\boldsymbol{r}_{mn} = \boldsymbol{r}_n - \boldsymbol{r}_m$ eingeführt wurden. Gleichung (2.32) gilt für beliebig strukturierte Proben. Für Kristalle läßt sich die Summation über die Vektoren $\boldsymbol{r}_{mn}$ mit Hilfe des reziproken Gitters ausführen, da dort die Translationsinvarianz des Gitters ausgenutzt werden kann. Im Falle amorpher Festkörper oder Flüssigkeiten ist diese Summation nicht geschlossen durchführbar.

Zur Vereinfachung nehmen wir bei der folgenden Diskussion zunächst an, daß die Substanz nur aus einer Atomsorte besteht. Eine Erweiterung auf mehrkomponentige Systeme werden wir noch anhängen. Für eine Atomsorte vereinfacht sich (2.32) zu

$$|\mathcal{A}(\boldsymbol{K})|^2 = \sum_m \tilde{a}^2(\boldsymbol{K}) + \sum_m \sum_{n\neq m} \tilde{a}^2(\boldsymbol{K})\,\mathrm{e}^{-\mathrm{i}\boldsymbol{K}\cdot\boldsymbol{r}_{mn}} \,, \quad (2.33)$$

wobei der Beitrag der Bezugsatome m von der zweiten Summe abgetrennt wurde. Nun gehen wir von der Summation über die Atome n zur Integration über die Teilchenzahldichte $n_m(\boldsymbol{r})$ um die Bezugsatome m über. Da bei amorphen Festkörpern im Mittel die Umgebung jedes Atoms isotrop ist, brauchen wir keine Richtungsabhängigkeit zu berücksichtigen. Wir ersetzen die Summe $\sum_n$ durch das Integral $\int n_m(r)\,\mathrm{d}V$ und erhalten

$$|\mathcal{A}(K)|^2 = \sum_m \tilde{a}^2(K) + \sum_m \tilde{a}^2(K) \int n_m(r)\,\mathrm{e}^{-\mathrm{i}\boldsymbol{K}\cdot\boldsymbol{r}}\,\mathrm{d}V \,. \quad (2.34)$$

Die gleiche Integration trat soeben bei der Berechnung des atomaren Streufaktors auf. Dort stand anstelle der Teilchenzahldichte $n(r)$ die Streudichte $\varrho(r)$ der Atome. Wir übernehmen das Ergebnis und erhalten

$$|\mathcal{A}(K)|^2 = \sum_m \widetilde{a}^2(K) + \sum_m \widetilde{a}^2(K) \int_{V_P} 4\pi r^2 n_m(r) \frac{\sin Kr}{Kr} \, \mathrm{d}r \,, \tag{2.35}$$

wobei im vorliegenden Fall die Integration über das gesamte Probenvolumen V_P erfolgt.

Wie bereits erwähnt, besitzen im Mittel alle Atome die gleiche Umgebung, so daß wir den Index m bei der Teilchenzahldichte fallen lassen können. Die Summation über den Index m ergibt die Zahl N der vorhandenen Atome, so daß sich (2.35) zu

$$|\mathcal{A}(K)|^2 = N\widetilde{a}^2(K) \left[1 + \int 4\pi r^2 n(r) \frac{\sin Kr}{Kr} \, \mathrm{d}r\right] \tag{2.36}$$

vereinfachen läßt. Die Größe $n(r)$ ist identisch mit der Teilchenzahldichte, wie sie bei der Definition der Paarkorrelationsfunktion $g(r)$ in Kap. 1.4 benutzt wurde. Wie wir dort gesehen haben, sind bei amorphen Festkörpern die beiden Größen über $g(r) = n(r)/n_0$ direkt miteinander verknüpft, wobei n_0 die mittlere Teilchenzahldichte ist.

Da sich das Integral über das ganze Probenvolumen erstreckt, enthält es nicht nur Beiträge von der Umgebung des „mittleren" Bezugsatoms, die hier von Interesse sind. Die Beiträge der Nachbarn lassen sich von denen aller übrigen Atome durch einen einfachen Trick abtrennen. Wir ziehen von $n(r)$ die mittlere Teilchenzahldichte n_0 ab und addieren diese wieder. Nach geeigneter Zusammenfassung der Terme und Division durch $N\widetilde{a}^2(K)$ findet man

$$\frac{|\mathcal{A}(K)|^2}{N\widetilde{a}^2(K)} = \underbrace{1}_{\text{Bezugsatom}} + \underbrace{\int 4\pi r^2 [n(r) - n_0] \frac{\sin Kr}{Kr} \, \mathrm{d}r}_{\text{lokale Struktur}} + \underbrace{\int 4\pi r^2 n_0 \frac{\sin Kr}{Kr} \, \mathrm{d}r}_{\text{entfernte Atome}} \,. \tag{2.37}$$

Wie bereits erwähnt, beschreibt der erste Term den Streubeitrag der Bezugsatome, also die Streuintensität, die beim Fehlen von Interferenz auftritt. Der Faktor $[n(r) - n_0]$ im zweiten Term verschwindet schon nach wenigen Atomabständen, da $n(r)$ für größere Abstände vom Bezugsatom dem Wert n_0 zustrebt (vgl. Kap. 1.4). Es tragen daher nur benachbarte Atome zu diesem Streuanteil bei. Der Interferenzterm spiegelt also die lokale Struktur um das (willkürlich herausgegriffene) Bezugsatom wider. Dagegen kommt der Hauptbeitrag zum dritten Term – wegen des Faktors r^2 – von Atomen mit großem Abstand vom Bezugsatom. Dieser Beitrag wird durch die endliche Probendimension, makroskopische Dichtefluktuationen oder andere makroskopische Inhomogenitäten bestimmt. Der Faktor $(\sin Kr)/Kr$ bewirkt, daß dieser Streubeitrag praktisch nur in Vorwärtsrichtung, d.h. bei $K = 0$ auftritt. Bereits bei relativ kleinen Werten des Streuvektors oszilliert der Integrand so rasch, daß das Integral verschwindet. In den meisten Streuexperimenten wird die Intensität in unmittelbarer Umgebung des durchgehenden Strahls ausgeblendet und nicht weiter aufgelöst. Da dieser Streubeitrag, den man als *Kleinwinkelstreuung* bezeichnet, für die Bestimmung der lokalen Ordnung bedeutungslos ist, lassen wir ihn in der weiteren Diskussion außer acht[3].

[3] Natürlich gibt es auch Anwendungen, bei denen gerade größere Strukturen untersucht werden sollen. Für die dann erforderlichen Kleinwinkeluntersuchungen sind spezielle experimentelle Techniken verfügbar.

Ein Blick auf die Gleichungen (2.32) und (2.37) verdeutlicht, daß $|\mathcal{A}(K)|^2$ nicht direkt proportional zur Fourier-Transformierten der Paarkorrelationsfunktion ist. Der Grund hierfür ist, daß zur Streuung auch die Bezugsatome beitragen, für die Interferenzen zwischen verschiedenen Streubeiträgen keine Rolle spielen. Sie werden durch den ersten Term $N\widetilde{a}^2$ in (2.37) beschrieben. Gemäß dem *Patterson-Theorem* ist $|\mathcal{A}(K)|^2/\widetilde{a}^2(K)$ die Fourier-Transformierte der Autokorrelationsfunktion der Streudichte. Da $|\mathcal{A}(K)|^2$ eine Meßgröße ist und der Atom-Strukturfaktor $\widetilde{a}(K)$ aus anderen Messungen bekannt ist, läßt sich die Autokorrelationsfunktion und damit auch die Paarkorrelationsfunktion mit Hilfe einer Fourier-Transformation unmittelbar aus den Streudaten gewinnen.

Nun wollen wir noch die Erweiterung auf mehrkomponentige Systeme angeben. Hierzu schreiben wir die Gleichung (2.37) unter Vernachlässigung der Kleinwinkelstreuung wie folgt um:

$$\frac{|\mathcal{A}(K)|^2}{N} = \sum_i \widetilde{a}_i^2(K) + \sum_{i,j} \widetilde{a}_i(K)\,\widetilde{a}_j(K) \int 4\pi r^2 [n(r) - n_0]\,\frac{\sin Kr}{Kr}\,\mathrm{d}r\,. \tag{2.38}$$

Wir wollen nun die entwickelten Vorstellungen mit experimentellen Ergebnissen vergleichen. In Abb. 2.18 sind Röntgen- und Neutronenstreumessungen an Quarzglas gezeigt. In beiden Fällen ist die Streuintensität als Funktion des Streuvekturs K aufgetragen. Die gemessenen Intensitäten sind auf die „Intensität" I_0 bezogen, die in den beiden Fällen unterschiedliche Bedeutung hat. Bei der Röntgenstreuung ist die Streuintensität proportional zu $(Ze)^2$, d.h. I_0 hat dort die Einheit von e^2. Im Falle der Neutronenstreuung ist $I(K) \propto |b|^2$. Da die Streulänge b in fm gemessen wird, hat I_0 nun die Einheit $(\mathrm{fm})^2$.

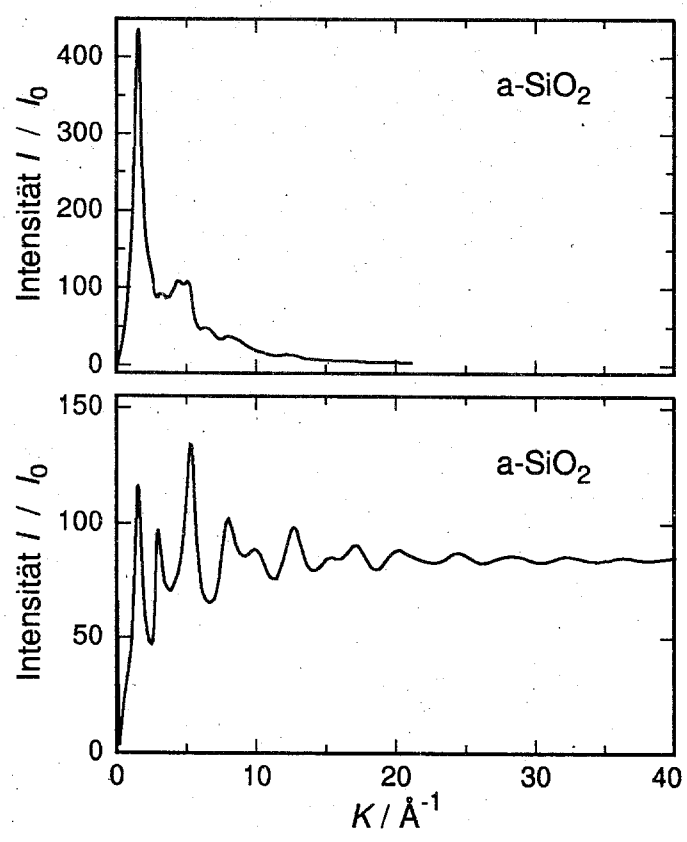


Abb. 2.18: Röntgen- und Neutronenstreuung an Quarzglas. Es ist jeweils die Streuintensität I als Funktlion des Streuvektors K aufgetragen. Die Bedeutung von I_0 ist im Text erläutert. (Nach R.L. Mozzi und B.E. Warren, J. Appl. Cryst., **2** (1969) 164 (oberes Bild), bzw. A.C. Wright, J. Non-Cryst. Solids **179** (1994) 84) (unteres Bild).

Es fällt auf, daß die Intensität der Röntgenstrahlung rasch mit dem Streuwinkel abfällt. Die Ursache hierfür ist die Winkelabhängigkeit des atomaren Streufaktors, wie sie in Abb. 2.17 zu sehen ist. Wie bereits erwähnt, hängt dagegen die Streulänge, die bei der Neutronenstreuung dem Atom-Strukturfaktor entspricht, nicht vom Winkel ab. Bei der Neutronenstreuung sind Daten bis $K = 40\,\text{Å}^{-1}$ gezeigt. Messungen bei derartig großen Werten des Streuvektors sind erst in den letzten Jahren möglich geworden und erweitern die Möglichkeiten der Strukturbestimmung ganz erheblich. Weiter fällt auf, daß die Oszillationen, die durch die Paarkorrelationsfunktion hervorgerufen werden, unterschiedlich stark ausgeprägt sind. Der Grund hierfür ist, daß sich Silizium- und Sauerstoffatome bezüglich Röntgenquanten und Neutronen unterschiedlich verhalten. Bei Röntgenstrahlen wird das Streuvermögen durch $f^2(K) \propto Z^2$ bestimmt. Deswegen streuen Siliziumatome etwa dreimal so stark wie Sauerstoffatome. Bei Neutronen ist dagegen $|b|^2$ entscheidend. Dort findet man gerade das umgekehrte Verhalten: Das Streusignal, das von einem Sauerstoffatom herrührt, ist etwa doppelt so groß wie das eines Siliziumatoms.

Wir wollen nun kurz auf die Auswertung der Streudaten eingehen. Wie bereits erwähnt, läßt sich in dem einfachen Fall, daß der amorphe Festkörper nur aus einer Atomsorte aufgebaut ist, die Autokorrelationsfunktion und damit die Paarkorrelationsfunktion $g(r)$ durch die Fourier-Transformation von $|\mathcal{A}(K)|^2$ gewinnen. Sehr häufig wird aber nicht $g(r)$, sondern die *radiale Dichteverteilungsfunktion* RDF angegeben, die über die Beziehung $\text{RDF} = 4\pi r^2 n(r)$ mit der Teilchendichte verknüpft ist. Die RDF gibt die Anzahl der Atome an, die in einer dünnen Kugelschale mit dem Radius r um das Bezugsatom im Mittel anzutreffen sind. Bei großen Abständen vom Aufatom ist $n(r) \approx n_0$, und der Wert der RDF nimmt mit dem Quadrat der Entfernung zu. In Abb. 2.19 ist als einfaches Beispiel die RDF von amorphem Germanium dargestellt und zum Vergleich auch der Kurvenverlauf für polykristallines Germanium eingezeichnet, das aufgrund der regellosen Orientierung seiner kleinen Kristallite im Mittel ebenfalls eine richtungsunabhängige Struktur besitzt.

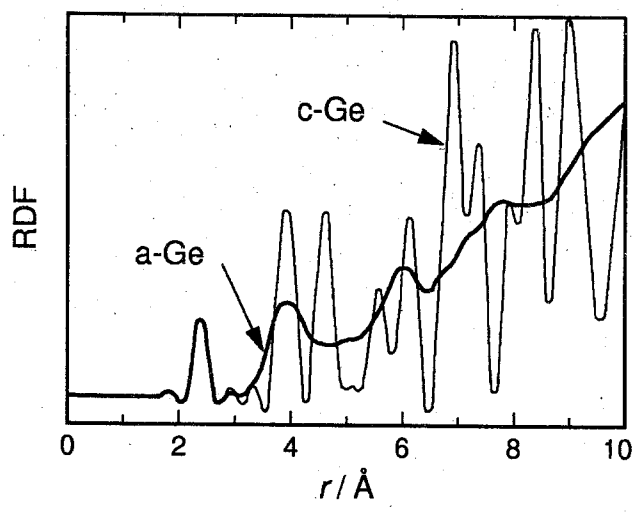


Abb. 2.19: Radiale Dichteverteilungsfunktion (RDF) von amorphem (dicke Kurve) und kristallinem (dünn gezeichnet) Germanium. (Nach R.J. Temkin et al., Adv. Phys. **22** (1973) 581)

Das erste Maximum tritt in beiden Fällen bei etwa 2,4 Å auf und hat das gleiche Aussehen. Dies bedeutet, daß in polykristallinem und amorphem Germanium der gleiche Abstand und die gleiche Anzahl von nächsten Nachbarn vorhanden ist. Somit ist die Nahordnung in der amorphen Phase ähnlich stark wie bei Kristallen ausgeprägt. Die

Verbreiterung des zweiten Maximums im Falle des amorphen Materials deutet auf Variationen des Bindungswinkels hin, die zu einer Verteilung im Abstand der übernächsten Nachbarn führen. Die weiteren Koordinationsschalen sind in amorphem Germanium nur schwach ausgeprägt. Bei großen Abständen geht die radiale Dichteverteilungsfunktion allmählich in die eines Kontinuums über. Offensichtlich sind die hier dargestellten Ergebnisse in guter Übereinstimmung mit den Vorstellungen von der Struktur von amorphen Festkörpern, wie wir sie in Kap. 1.4 entwickelt haben.

Es ist zu beachten, daß in Streuexperimenten immer eine Mittelung über alle existierenden Atomanordnungen erfolgt. Darüber hinaus beinhaltet die Paarkorrelationsfunktion oder die RDF grundsätzlich eine *eindimensionale Darstellung* einer *dreidimensionalen Struktur*. Doch bereitet die Bestimmung selbst dieser scheinbar einfachen Funktion erhebliche Schwierigkeiten. Theoretisch ist für die Fourier-Transformation die Kenntnis des exakten Verlaufs von $|\mathcal{A}(K)|^2$ erforderlich, doch experimentell kann die Streuintensität nicht für alle Werte des Streuvektors K bestimmt werden. Zusätzlich erweisen sich die Messungen bei sehr kleinen und sehr großen Streuvektoren als relativ stark fehlerbehaftet. Die unvermeidliche Einschränkung des Meßbereichs und die immer auftretenden Meßfehler können bei der Fourier-Transformation zusätzliche Strukturen in der Autokorrelationsfunktion bewirken, die kaum von „echten" Struktureigenheiten unterschieden werden können.

Eine weitere prinzipielle Schwierigkeit tritt bei Untersuchung von Mehrkomponentensystemen auf. Bei Substanzen der Zusammensetzung AB wie z.B. Quarzglas reicht zur Beschreibung der Struktur eine Funktion nicht aus, denn es gibt drei partielle Paarkorrelationsfunktionen: Neben den beiden Funktionen, welche die Korrelation zwischen den A- bzw. B-Atomen untereinander beschreiben, gibt es noch die gemischte RDF, die die Verteilung der einen Atomsorte bezüglich der anderen angibt. Für die Berechnung diese drei Funktionen sind drei unabhängige Sätze von Streudaten erforderlich. Wie Abb. 2.18 verdeutlichte, liefern Messungen mit Hilfe von Neutronen- und Röntgenstrahlen zwei unabhängige Datensätze. Elektronen- und Atomstrahlbeugung ergeben aber keine zusätzliche Information, da sie wie die Röntgenstrahlen an der Elektronenhülle gestreut werden. In Ausnahmefällen, z.B. beim metallischen Glas $Ni_{76}P_{24}$, erlaubt das unterschiedliche Streuverhalten verschiedener Isotope zwei voneinander unabhängige Messungen mit Neutronen. In diesem speziellen Fall ist es tatsächlich möglich, einen vollständigen Satz von partiellen radialen Dichteverteilungsfunktionen zu erstellen.

Wie geht man nun vor, um aus den Streudaten, d.h. aus $|\mathcal{A}(K)|^2$, auf die Struktur der Probe zu schließen? Bereits bei einkomponentigen Systemen ist hierfür erheblicher numerischer Aufwand erforderlich. Im allgemeinen geht man schrittweise vor. Man entwirft zunächst ein Strukturmodell, ermittelt dessen Paarkorrelationsfunktion bzw. Autokorrelationsfunktion und berechnet mit Hilfe der Fourier-Transformation die zu erwartende Streuintensität. Ein Vergleich mit der gemessenen Kurve wird im allgemeinen gewisse Abweichungen zeigen. Daraufhin verbessert man das Modell und durchläuft diese Prozedur so lange, bis die Übereinstimmung zwischen der gemessenen und der errechneten Kurve zufriedenstellend ist. Zusätzlich werden meist noch Daten herangezogen, die aus anderen strukturempfindlichen Experimenten stammen. Natürlich ist selbst bei bester Übereinstimmung noch nicht der Beweis dafür erbracht, daß das Strukturmodell tatsächlich mit der vorliegenden realen Struktur übereinstimmt, denn Paarkorrelationsfunktionen sind oftmals recht unspezifisch. Es hat sich gezeigt, daß sehr unterschiedliche Strukturmodelle unter Umständen zu sehr ähnlichen Korrelationsfunktionen führen. Liegen vom amorphen Festkörper mehrere unabhängige Streumessungen vor, so müssen erst mit Hilfe der

bekannten Atom-Strukturfaktoren die partiellen Korrelationsfunktionen ermittelt werden und deren Übereinstimmung mit dem Strukturmodell überprüft werden. In neuerer Zeit spielen bei der Auswertung der Streudaten zunehmend sogenannte Monte-Carlo-Techniken und molekulardynamische Simulationen eine wichtige Rolle, die mit enormem Rechenaufwand verbunden sind.

2.6 Experimentelle Methoden

Bereits am Anfang dieses Kapitels wurde darauf hingewiesen, daß Strukturuntersuchungen mit unterschiedlicher Strahlung durchgeführt werden. Zur Bestimmung der Struktur von ausgedehnten Festkörpern eignen sich vor allem die Röntgen- und Neutronenbeugung, auf deren experimentelle Realisierung wir etwas näher eingehen. Abhängig von der Meßmethode, die wir im folgenden Abschnitt diskutieren, erfordert die Messung ein kontinuierliches oder aber ein monochromatisches Strahlungsspektrum. Da eingehendere Untersuchungen im allgemeinen mit monochromatischer Strahlung erfolgen, ist in Abb. 2.20 der schematische Aufbau eines derartigen Streuexperiments skizziert. Um aus dem breiten Spektrum der Quelle ein enges Wellenlängenband herauszuschneiden, nutzt man die in diesem Kapitel ausführlich diskutierte Bragg-Reflexion aus. Man läßt den polychromatischen Röntgen- oder Neutronenstrahl auf einen *Monochromatorkristall* fallen. Der größte Teil des auftreffenden Strahls durchquert diesen Kristall praktisch ohne Ablenkung, doch der Teil mit der „richtigen" Wellenlänge wird unter dem Bragg-Winkel herausgebeugt. Dieser Teil des Strahls wird dann für monochromatische Streuuntersuchungen verwendet.

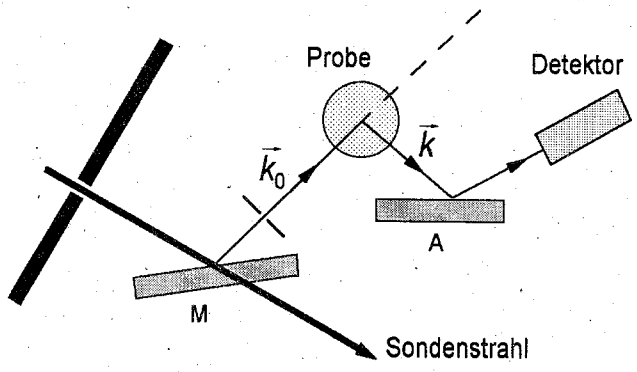


Abb. 2.20: Prinzipieller Aufbau eines Streuexperiments mit monochramatischer Strahlung. Die Einkristalle M und A werden als Monochromator bzw. als Analysator benutzt.

Die gestreute Strahlung läßt man entweder direkt in den Detektor laufen oder auf einen zweiten Einkristall fallen, an dem sie nochmals eine Bragg-Reflexion erfährt. Die Reflexion im sogenannten Analysatorkristall dient bei Untersuchungen der elastischen Streuung der Unterdrückung des Streuuntergrunds mit unterschiedlicher Wellenlänge. Bei inelastischer Streuung, auf die wir ausführlich in Kap. 5.4 zu sprechen kommen, wird der Analysatorkristall zur Messung der Energieverschiebung der gestreuten Strahlung benutzt.

Wegen des vergleichsweise geringen technischen Aufwands werden zur Strukturbestimmung meist Röntgenstrahlen aus Röntgenröhren verwendet, wobei je nach Meßverfahren das kontinuierliche Bremsspektrum oder die charakteristische Röntgenstrahlung benutzt wird. In den letzten Jahren gewann die Nutzung von Synchrotronstrahlung zur Strukturuntersuchung immer stärker an Bedeutung. Sie ist in einem breiten Wellenlängenbereich

verfügbar und zeichnet sich durch extrem hohe Intensität bei kleiner Strahldivergenz, hoher Kohärenz und vollständiger Polarisation aus. Der Nachweis von Röntgenstrahlen erfolgte früher ausnahmslos auf fotografischem Weg. In neuerer Zeit werden zunehmend Szintillationszähler, Proportionalzähler und vor allem Festkörperdetektoren (z.B. Siliziumdioden) verwendet.

Wie wir gesehen haben, ist die Intensität der Röntgenreflexe proportional zum Quadrat des atomaren Streufaktors. Da dessen Größe proportional zur Elektronenzahl ansteigt, nimmt die Streuung quadratisch mit der Kernladungszahl der streuenden Atome zu. Leichte Elemente sind daher insbesondere in Gegenwart von schweren kaum zu beobachten, bzw. erfordern eine besonders empfindliche Erfassung der Streuintensitäten. Wasserstoffatome konnten daher in vielen Fällen gar nicht nachgewiesen werden. Hier wurden in jüngster Zeit mit den modernen Synchrotronquellen wesentliche Fortschritte erzielt, so daß heute Wasserstoffverbindungen, die in der Polymerphysik und bei biologischen Untersuchungen eine große Rolle spielen, in Röntgenstreuexperimenten studiert werden können. Der Grund hierfür ist, daß sich mit den neuesten Quellen (z.B. mit der European Synchrotron Radiation Facility (ESRF) in Grenoble, Frankreich) auf der Probe eine spektrale Photonenflußdichte erzielen läßt, die viele Größenordnungen über der von konventionellen Röntgenquellen liegt. Bei Neutronenstreuung sind die Voraussetzungen für den Nachweis von leichten Atomen wesentlich günstiger, da der Absolutwert der Streulänge zwar von Kern zu Kern schwankt, aber für alle Elemente von gleicher Größenordnung ist. Allerdings muß man berücksichtigen, daß die meisten Kerne auch mehr oder weniger inkohärent streuen, somit einen Streuuntergrund verursachen und daher für Strukturuntersuchungen nicht gleich gut geeignet sind. Besonders groß ist der inkohärente Streuquerschnitt bei Wasserstoff. Dennoch wurde bis vor kurzem die Lage der Wasserstoffatome in Festkörpern vor allem mit Neutronen untersucht, wobei man sich den ebenfalls vorhandenen schwachen kohärenten Streubeitrag zunutze macht. Eine wesentliche Verbesserung des Signals erzielt man durch Ersetzen des leichten Wasserstoffs durch Deuterium, da das Verhältnis zwischen kohärentem und inkohärentem Streuquerschnitt dann wesentlich günstiger ist. Eine weitere Schwierigkeit bei der Röntgenbeugung bereitet die Unterscheidung von Elementen, die im Periodensystem benachbart liegen, da sich ihre atomaren Streufaktoren gleichen. Ein Beispiel hierfür haben wir in Kap. 2.4 kennengelernt: In Cäsiumjodid sind die beiden Ionensorten bei Röntgenbeugungsexperimenten kaum zu unterscheiden. Das Beugungsbild von CsJ gleicht daher dem eines kubisch raumzentrierten „Xenonkristalls“ mit einatomiger Basis[4]. Mit Neutronen lassen sich Atome benachbarter Elemente meist gut unterscheiden, da die Kerne unterschiedliche Streueigenschaften aufweisen. Ein weiteres wichtiges Anwendungsfeld der Neutronenstreuung stellt die Untersuchung von magnetischen Strukturen (vgl. Kap. 10.3) dar, die auf der Wechselwirkung von Neutronen mit den magnetischen Momenten beruht.

Geeignete Neutronenstrahlen erhält man aus Kernreaktoren, die so ausgelegt sind, daß sie einen besonders hohen Neutronenfluß besitzen, oder aus Spallationsquellen, bei denen Protonen mit einem Beschleuniger auf Schwermetalltargets geschossen werden. Da die hierbei erzeugten Neutronen für Strukturuntersuchungen eine zu große Energie bzw. eine zu kleine Wellenlänge aufweisen, müssen sie erst in einem Moderator abgebremst werden. Sie erlangen dort thermische Geschwindigkeiten und somit die passende Wellenlänge für Beugungsexperimente. Wie oben geschildert, wird dann mit Hilfe der Bragg-Reflexion

[4] Reale Xenonkristalle sind kubisch flächenzentriert.

aus der breiten Geschwindigkeitsverteilung ein enger Geschwindigkeitsbereich „herausgeschnitten", d.h. es wird dadurch ein „monochromatischer" Neutronenstrahl erzeugt. Der Nachweis der gestreuten Neutronen erfolgt mit geeigneten Zählrohren, z.B. mit BF_3-Proportionalzählern, in denen durch Neutroneneinfang ^{10}B-Kerne in ^{7}Li- und ^{4}He-Kerne umgewandelt werden, die das Gas im Zähler ionisieren. In jüngerer Zeit werden öfters ^{3}He-Detektoren benutzt, bei denen ^{3}He-Kerne Neutronen einfangen und dabei in ein ^{3}H und ein Proton zerfallen, das wiederum nachgewiesen wird.

Eine wesentliche Rolle bei der Wahl der Sonde spielt oft auch die Größe der verfügbaren Probe. So sind für Neutronenstreuexperimente relativ große Proben erforderlich, da die Wechselwirkung langsamer Neutronen mit den Kernen relativ schwach ist. Bei Röntgenstrahlen genügt dagegen zur Strukturbestimmung oft bereits eine relativ kleine Probe.

Meßverfahren

Die Streubedingung (2.16) enthält zwei Größen - Wellenlänge und Kristallorientierung - die im Experiment variiert werden können. Letztere bestimmt bei vorgegebener Richtung der einfallenden Welle den Streuwinkel. Will man Reflexe beobachten, so muß bei der Messung mindestens eine der beiden Größen verändert werden. Die meisten Experimente lassen sich daher, abhängig vom experimentellen Vorgehen, einer der drei im folgenden diskutierten Meßmethoden zuordnen, obwohl die Meßtechniken und der jeweilige apparative Aufwand je nach Natur der Strahlung grundverschieden sind.

Drehkristallverfahren

Beim Drehkristallverfahren, das man meist zur Ermittlung der Struktur komplex aufgebauter Kristalle heranzieht, wird mit monochromatischer Strahlung gearbeitet. Der Kristall wird dabei im Strahl um eine feste Achse gedreht. Bei dieser Rotation dreht sich das reziproke Gitter mit dem Kristall, und es treten immer dann Beugungsreflexe auf, wenn ein Punkt des reziproken Gitters auf die Ewald-Kugel trifft. Da dieser Fall bei einer vollständigen Drehung um 360° für bestimmte reziproke Gitterpunkte wiederholt auftreten kann, schränkt man meist den Drehwinkel ein. Es wurden verschiedene komplizierte Varianten des Drehkristallverfahrens ausgearbeitet mit dem Ziel, eine möglichst eindeutige Zuordnung der Reflexe zu gewährleisten. So wird beispielsweise beim *Weissenberg-Verfahren* der Film während der Kristalldrehung noch zusätzlich bewegt.

In allen Fällen vereinfacht sich die Auswertung der Meßdaten erheblich, wenn die Drehachse des untersuchten Kristalls mit einer Achse hoher Kristallsymmetrie zusammenfällt. Um dies zu erreichen, wird der Kristall vorher meist mit Hilfe des Laue-Verfahrens orientiert, auf das wir gleich noch zu sprechen kommen. In modernen Röntgenanlagen werden die Reflexe automatisch angefahren und ihre Intensität erfaßt. In Extremfällen werden bis zu 10^6 Reflexe registriert, so daß sich auch Feinheiten der Ladungsverteilung und damit sehr große Elementarzellen mit komplizierter Basis rekonstruieren lassen.

Pulververfahren (Debye-Scherrer-Verfahren)

Beim Pulververfahren wird eine monochromatische (Röntgen-)Strahlung auf feinkörniges, kristallines Pulver gerichtet. Wie in Abb. 2.21 schematisch dargestellt, fällt der Röntgenstrahl, eingeengt durch eine Blende, auf die Probe und wird entsprechend der Bragg-Bedingung an den Netzebenen gestreut. Im einfachsten Fall wird die gestreute Strahlung mit einem Röntgenfilm nachgewiesen.

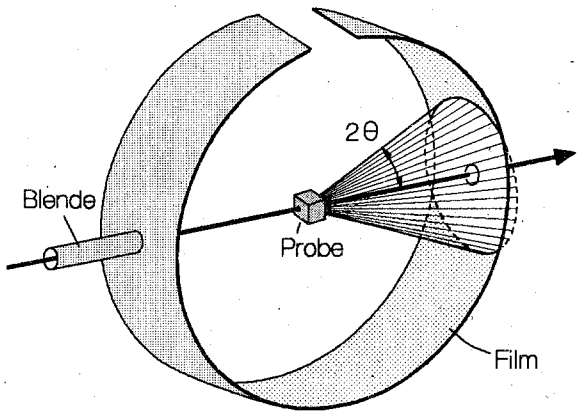


Abb. 2.21: Schematische Darstellung des Debye-Scherrer-Verfahrens. Monochromatische Röntgenstrahlung trifft auf die Probe in der Mitte der Anordnung. Die kegelförmig gestreute Röntgenstrahlung wird in der Skizze durch einen Filmstreifen aufgefangen.

Die auftretenden Reflexe lassen sich mit der Ewald-Kugel ermitteln, wenn man das reziproke Gitter um seinen Aufpunkt (000) in alle Richtungen dreht und die Durchstossungspunkte der Gitterpunkte durch die Kugel ermittelt. Da im Pulver alle Kristallorientierungen vorkommen, beobachtet man anstelle der punktförmigen Reflexe sogenannte „Debye-Scherrer-Ringe". Wie man sich anhand von Abb. 2.13 klarmachen kann, tritt eine zur Strahlachse symmetrische Streuung unter dem Winkel 2θ auf.

In Abb. 2.22 ist eine Debye-Scherrer-Aufnahme von kristallinem Quarz gezeigt. Oben ist die fotografische Aufnahme der Beugungsringe wiedergegeben, unten ist die Streuintensität als Funktion des Winkels aufgetragen, wie sie mit einem sogenannten „Pulverdiffraktometer" registriert wird.

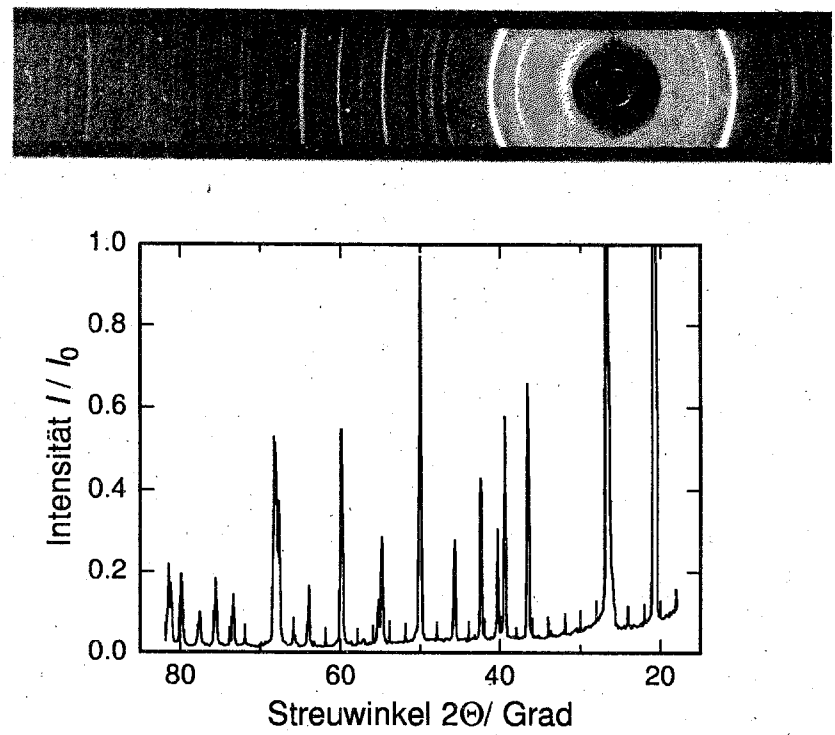


Abb. 2.22: Pulveraufnahme an α-Quarz. Aufnahme der Beugungsringe in einer Röntgen-Kamera (oben) und mit einem Pulverdiffraktometer (unten). Die Intensitätsmaxima lassen sich direkt den Ringen der fotografischen Aufnahme zuordnen.

Die Winkelabhängigkeit der Streuintensität bei Pulveraufnahmen ähnelt sehr stark dem Verlauf, der bei Beugungsexperimenten an amorphen Materialien auftritt. Bei Aufnahmen nach der Pulvermethode mittelt der Strahl über alle Kristallorientierungen, bei amorphen Festkörpern erfolgt eine Mittelung über alle lokalen Atomanordnungen. Es ist daher verständlich, daß eine Anhäufung von regellos orientierten Kristalliten eine ähnliche Streuung hervorruft, wie sie von den amorphen Festkörpern her bekannt ist. Dennoch besteht ein wesentlicher Unterschied zwischen diesen beiden Stoffklassen. Da die Kristallite nicht nur Nah-, sondern auch Fernordnung aufweisen, sind die beobachteten Ringe nicht verwaschen, sondern scharf, wie in Abb. 2.22 zu sehen ist.

Das Debye-Scherrer-Verfahren eignet sich besonders gut zur genauen Bestimmung von Gitterkonstanten. Die meisten verfügbaren Gitterdaten sind mit dieser Methode gewonnen worden. Der Vollständigkeit halber sei erwähnt, daß es ein anderes, wesentlich genaueres Verfahren gibt. Damit wurde in jüngster Zeit der Abstand der {220}-Ebenen von Silizium mit einer relativen Genauigkeit von $3 \cdot 10^{-8}$ bestimmt und der Wert $d_{220} = 192\,015,569 \pm 0,006\,\mathrm{fm}$ gefunden. Diese Messung wurde mit einem Röntgeninterferometer durchgeführt, das mit einem optischen Interferometer gekoppelt ist. Diese Verknüpfung ist für die hohe Präzision zwingend, da die Wellenlängen der Röntgenstrahlung nicht mit der erforderlichen Genauigkeit bekannt sind.

Natürlich läßt sich das Debye-Scherrer-Verfahren auch mit Neutronen durchführen. In Abb. 2.23 ist das Ergebnis einer Messung an einer $YBa_2Cu_3O_7$-Pulverprobe zu sehen. Dieses Material zählt zu den berühmten „Hochtemperatursupraleitern", auf die wir in Abschnitt 12.4 noch ausführlich eingehen.

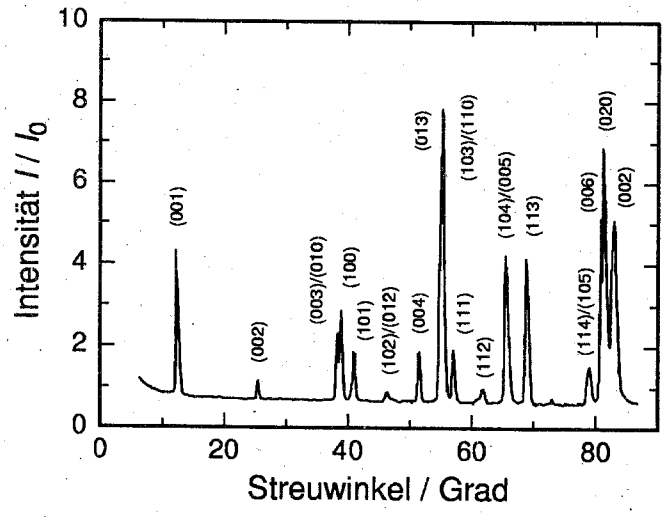


Abb. 2.23: Neutronendiffraktogramm einer $YBa_2Cu_3O_7$-Pulverprobe. Die Bezeichnung der Reflexe bezieht sich auf die orthorhomische Elementarzelle. (Nach G. Schatz, A. Weidinger, Nukleare Festkörperphysik, Teubner Studienbücher, Stuttgart 1992).

Laue-Verfahren

Beim Laue-Verfahren wird ein fest montierter Einkristall mit einem kontinuierlichen Spektrum bestrahlt, so daß viele Kristallebenen gleichzeitig die Streubedingung erfüllen und zu Reflexen Anlaß geben. Strahlt man entlang einer Symmetrieachse des Kristalls ein, so entsteht ein Beugungsbild mit den Symmetrieelementen der Achse. Daher macht

man bei einem Kristall mit unbekannter Struktur zunächst eine Aufnahme in beliebiger Richtung, um die Hauptsymmetrieachsen zu ermitteln. Anschließend strahlt man längs dieser Achsen ein und bestimmt so aus den Laue-Aufnahmen die Kristallsymmetrie. Häufig werden Laue-Aufnahmen zur Orientierung von bekannten Kristallen verwendet. In Abb. 2.24 ist als Beispiel die Laue-Aufnahme eines $[C(NH_2)_3]_4[UO_2O_2(CO_3)_2] \cdot 2H_2O$ Kristalls gezeigt, bei der der ungebeugte Strahl ausgeblendet ist. Deutlich zu sehen ist auch die Halterung, auf der der untersuchte Kristall montiert wurde.

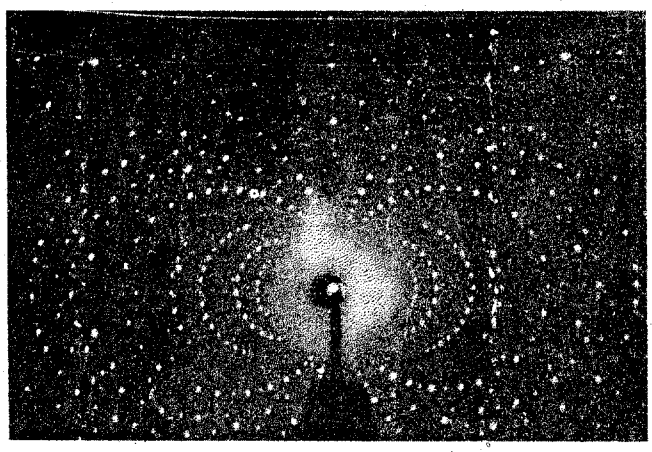


Abb. 2.24: Laue-Aufnahme. Der Kristall $[C(NH_2)_3]_4[UO_2O_2(CO_3)_2] \cdot 2H_2O$ besitzt eine zweizählige Symmetrie. Gut sichtbar ist die Kristallhalterung. (Aus B.K. Vainshtein, Modern Crystallography, in Solid State Sciences **15/I**, Springer-Verlag (Heidelberg), S. 288).

Untersuchung der Struktur von Oberflächen und dünnen Filmen

Die Wechselwirkung zwischen Festkörpern und Elektronen- oder Atomstrahlen ist besonders stark. Deshalb können bei Transmissionsexperimenten nur sehr dünne Schichten untersucht werden. Hierbei werden Elektronen, bevor sie auf die Probe treffen, auf etwa 50 – 100 keV beschleunigt. Aufgrund der kleinen Wellenlänge erfolgt dann die Streuung nur in einen kleinen Öffnungswinkel von etwa 3 – 5°. Anschließend werden die Elektronen mit einer Fotoplatte nachgewiesen. Als Beispiel für eine derartige Messung ist in Abb. 2.25 die Streuung an einer etwa 9 nm dicken Eisenschicht gezeigt.

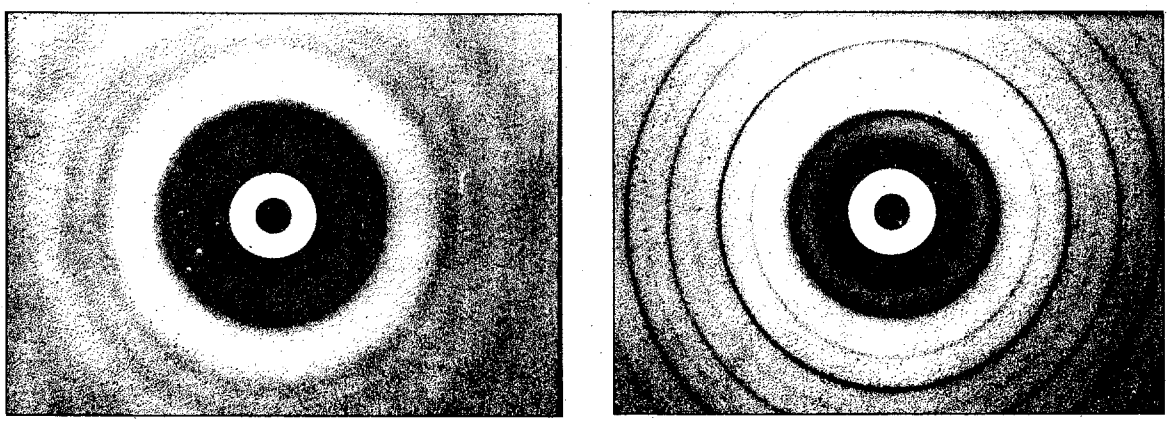


Abb. 2.25: Elektronenbeugungbilder eines 9 nm dicken Eisenfilms im amorphen und kristallinen Zustand. (Aus T. Ichikawa, Phys. Stat. Sol. A **19** (1973) 707).

Der Film wurde bei Heliumtemperatur, d.h. bei 4,2 K, durch Aufdampfen hergestellt und dann durchstrahlt. Beim Aufwärmen bis auf Zimmertemperatur wandelt sich der anfänglich amorphe Film in einen polykristallinen Film um, dessen Kristallite keine Vorzugsorientierung besitzen. Aufgrund dieser Phasenumwandlung werden die verschmierten Ringe so scharf, wie man sie von Debye-Scherrer-Aufnahmen an kristallinen Pulvern her kennt. Aus ähnlichen Daten wurde die in Abb. 1.31 gezeigte Paarkorrelationsfunktion von amorphem Nickel gewonnen.

Für die Untersuchung von Oberflächen oder oberflächennahen Schichten ist die Beugung von niederenergetischen Elektronen mit Energien zwischen 10 eV und 1000 eV besonders geeignet. Sie ist eine Standardmethode der Oberflächenphysik, die unter der Bezeichnung LEED (Low Energy Electron Diffraction) bekannt ist. In diesem Fall detektiert man nur Elektronen, die innerhalb der ersten Atomlagen der Probe rückgestreut wurden. Da die gestreuten Elektronen sehr empfindlich auf Oberflächenverunreinigungen reagieren, werden die Experimente üblicherweise im Ultrahochvakuum durchgeführt. Beobachtet wird meist das Fluoreszenzlicht, das die gestreuten Elektronen auf einem Leuchtschirm hervorrufen, der die Probe in Form einer Halbkugel umgibt.

An dieser Stelle soll noch auf einen wichtigen Unterschied zwischen LEED- und Röntgenstreumessungen hingewiesen werden. Die starke Streuung und Absorption der Elektronen hat zur Folge, daß die rückgestreuten Elektronen aus einer sehr dünnen Schicht an der Oberfläche kommen und das Probenvolumen deshalb näherungsweise als zweidimensional betrachtet werden kann. Die Streubedingung (2.16) muß dann nur noch von den Komponenten des Streuvektors erfüllt werden, die parallel zur Oberfläche stehen. Der Wellenvektor senkrecht zur Oberfläche ist praktisch bedeutungslos. Die Streubedingung wirkt sich daher weniger einschränkend aus, so daß selbst mit monoenergetischen Elektronen bei vorgegebener Streugeometrie mehrere Reflexe beobachtet werden, die von verschiedenen reziproken Gitterpunkten herrühren. LEED-Aufnahmen mit monoenergetischen Elektronen haben daher große Ähnlichkeit mit den oben diskutierten Laue-Aufnahmen.

Als Beispiel sind in Abb. 2.26 Aufnahmen gezeigt, die an 2 μm großen Al-Mn-Quasikristallen gewonnen wurden. Die beiden Aufnahmen unterscheiden sich durch den Winkel, unter dem sie aufgenommen wurden. Sie demonstrieren eindrucksvoll die Existenz einer ausgeprägten Orientierungsordnung in diesen eigenartigen Materialien. Deutlich ist zu erkennen, daß 2-, 3- und 5-zählige Drehsymmetrien auftreten, die für Ikosaeder typisch sind.

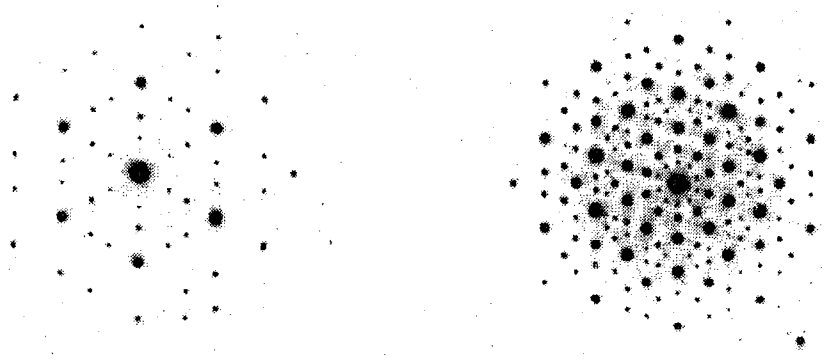


Abb. 2.26: Elektronenbeugung an einem Al-Mn-Quasikristall unter verschiedenen Winkeln. (Nach D. Shechtman et al., Phys. Rev. Lett. **53** (1984) 1951).

3. Bindungsmodelle

Im ersten Kapitel haben wir die Festkörper nach der geometrischen Anordnung ihrer Atome klassifiziert. Da man aber in allen Kristallsystemen und auch in den amorphen Festkörpern das gesamte Spektrum der optischen, elektrischen und mechanischen Eigenschaften findet, bedeutet dies, daß die *physikalischen Eigenschaften* erwartungsgemäß nicht ausschließlich von der Struktur bestimmt werden. Vor allem die außenliegenden Elektronen der Atome, also die Valenzelektronen, sind auf das Vorhandensein von benachbarten Atomen empfindlich und haben daher einen großen Einfluß auf die Festkörpereigenschaften. Hierbei spielt die Art der Bindung zwischen den Atomen eine ganz entscheidende Rolle. Diesen Gesichtspunkt wollen wir in diesem Kapitel näher untersuchen.

3.1 Generelle Aspekte

In vielen Fällen ist eine Einteilung der Festkörper sinnvoll, bei der gewisse physikalische Eigenschaften im Vordergrund stehen. Ziehen wir die elektrische Leitfähigkeit als Merkmal heran, so ist die Unterteilung nach *Isolatoren* und *Metallen* am augenfälligsten. Wie wir in Kap. 8 im Rahmen der „Bändertheorie" sehen werden, ist hierfür eigentlich die Verteilung der Elektronen im Impulsraum entscheidend. Ein grobes Verständnis erreicht man aber bereits, wenn man die Wellenfunktion der freien Atome im Ortsraum betrachtet. In Abb. 3.1 ist der Radialteil $r\psi(r)$ der Wellenfunktionen der besetzten Elektronenzustände von zwei freien Neon- bzw. Natriumatomen gezeichnet. Der Abstand der beiden Atome voneinander ist jeweils so gewählt, daß er dem Gitterabstand im Kristall entspricht.

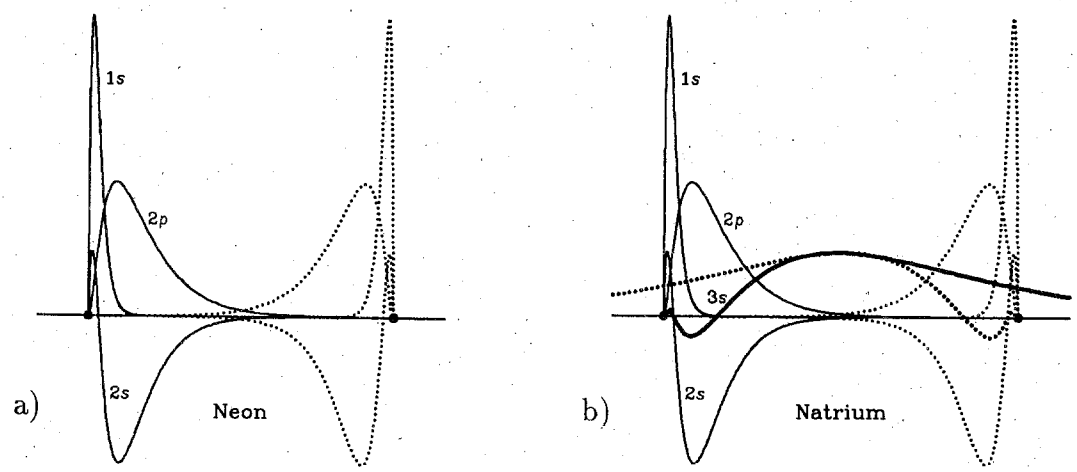


Abb. 3.1: Zum Überlapp von Wellenfunktionen. Für Neon und Natrium ist der Radialteil $r\psi(r)$ der Wellenfunktion von zwei Atomen (jeweils durchgezogene bzw. gepunktete Kurven) aufgetragen, die sich im Abstand nächster Nachbarn der entsprechenden Kristalle befinden. (Nach D.R. Hartree, W. Hartree, Proc. Roy. Soc. **A 193** (1948) 299).
a) Neon mit der Elektronenkonfiguration $1s^2\,2s^2\,2p^6$,
b) Natrium mit der Elektronenkonfiguration $1s^2\,2s^2\,2p^6\,3s^1$. Die $3s$-Wellenfunktionen sind durch dickere Kurven hervorgehoben.

Der Verlauf der Wellenfunktionen macht deutlich, daß beim Neon mit seiner abgeschlossenen Elektronenschale der Überlapp der Wellenfunktionen im Kristall eine untergeordnete Rolle spielt. Die außenliegenden Elektronen können auch im Festkörper den einzelnen Atomen zugeordnet werden. Völlig andere Verhältnisse liegen beim metallischen Natrium vor. Aufgrund des starken Überlapps der $3s$-Wellenfunktionen dürfen die Atome nicht mehr als unabhängige Einheiten betrachtet werden. Die Valenzelektronen können sich relativ leicht von einem Atom zum anderen bewegen und bewirken so die hohe elektrische Leitfähigkeit des Metalls.

Grundsätzlich lassen sich fünf verschiedene Bindungstypen unterscheiden, die jedoch in den meisten Fällen nicht in reiner, sondern in gemischter Form auftreten. Erwartungsgemäß findet man alle Bindungstypen sowohl in kristallinen wie auch in amorphen Festkörpern, so daß eine getrennte Behandlung der beiden Substanzklassen nicht erforderlich ist. Eine gewisse Sonderstellung nimmt die sogenannte *Wasserstoffbrückenbindung* ein, die in wasserstoffhaltigen Substanzen zu finden ist. Wir werden auf diese Art der Bindung am Ende dieses Kapitels eingehen. Bei den vier verbleibenden Bindungsarten kann man eine weitere Unterteilung treffen, je nachdem ob die Elektronenschale der Gitterbausteine abgeschlossen ist oder nicht. Im ersten Fall handelt es sich um Isolatoren, da die Elektronen am jeweiligen Atom oder Ion lokalisiert sind. Zu diesem Bindungstyp zählen die *Ionenkristalle*, bei denen ein Elektronentransfer zwischen den Bindungspartnern stattfindet, der den Abschluß der Elektronenschalen bewirkt. Die entgegengesetzt geladenen Ionen werden durch die starken Coulomb-Kräfte zusammengehalten. Wesentlich schwächer sind die Kräfte, wenn die Moleküle oder Atome bereits ohne Ladungstransfer eine abgeschlossene Schale besitzen, wie es bei Molekül- und Edelgaskristallen der Fall ist. Dann verbleiben als Bindungskräfte nur noch die relativ schwachen *Van der Waals-Kräfte*, auf die wir im folgenden Abschnitt eingehen.

Sind die Elektronenschalen der Gitterbausteine nicht abgeschlossen, so können im Extremfall die äußersten Elektronen im Kristall gleichmäßig verteilt sein. Diese starke Verschmierung der äußersten Elektronen ermöglicht nicht nur den Ladungstransport, sondern bewirkt auch Bindungskräfte zwischen den Atomrümpfen. In Kap. 3.5 werden wir auf diese *metallische Bindung* etwas näher eingehen. Der letzte verbleibende Bindungstyp ist die *kovalente Bindung*, bei der die Valenzelektronen *zwischen* den wechselwirkenden Atomen *lokalisiert* sind. In der hier benutzten stark vereinfachten Betrachtungsweise können diese Elektronen nicht zur elektrischen Leitfähigkeit beitragen. In vielen Isolatoren (und Halbleitern) werden die Atome durch diesen Mechanismus aneinandergebunden. Doch auch in zahlreichenen Metallen ist die kovalente Bindung von großer Bedeutung, da sie häufig die metallische Bindung ergänzt. Natürlich spielt die kovalente Bindung eine tragende Rolle in der Molekülphysik, doch würde die Behandlung dieses Aspekts weit über den hier gesteckten Rahmen hinausführen.

In Abb. 3.2 ist die angesprochene Verteilung der Valenzelektronen der verschiedenen Bindungstypen schematisch dargestellt. Beginnend beim metallischen Natrium mit völlig ausgeschmierten Valenzelektronen über das kovalent gebundene Silizium und die Van der Waals-Bindung beim festen Argon endet die Reihe beim Ionenkristall NaCl. Die Bereiche, in denen sich Valenzelektronen vorzugsweise aufhalten, wurden grau getönt. Im Falle von NaCl zählen wir in dieser Darstellung alle äußeren Elektronen des Chlors und das vom Natrium abgegebene Elektron zu den „Valenzelektronen". Natürlich liegt in der Regel nicht ein bestimmter Bindungstyp in reiner Form vor, sondern es treten gemischte Formen auf. So findet man einen allmählichen Übergang von der rein kovalenten zur rein ionischen Bindung, wenn man von vierwertigen Materialien wie Germanium über Verbindungen wie

GaAs und CaSe zu KBr geht. Bei der Diskussion der einzelnen Bindungsarten werden wir uns jedoch auf die einfachsten Fälle beschränken, d.h. wir werden nur Systeme betrachten, bei denen *ein* Bindungsmechanismus dominiert.

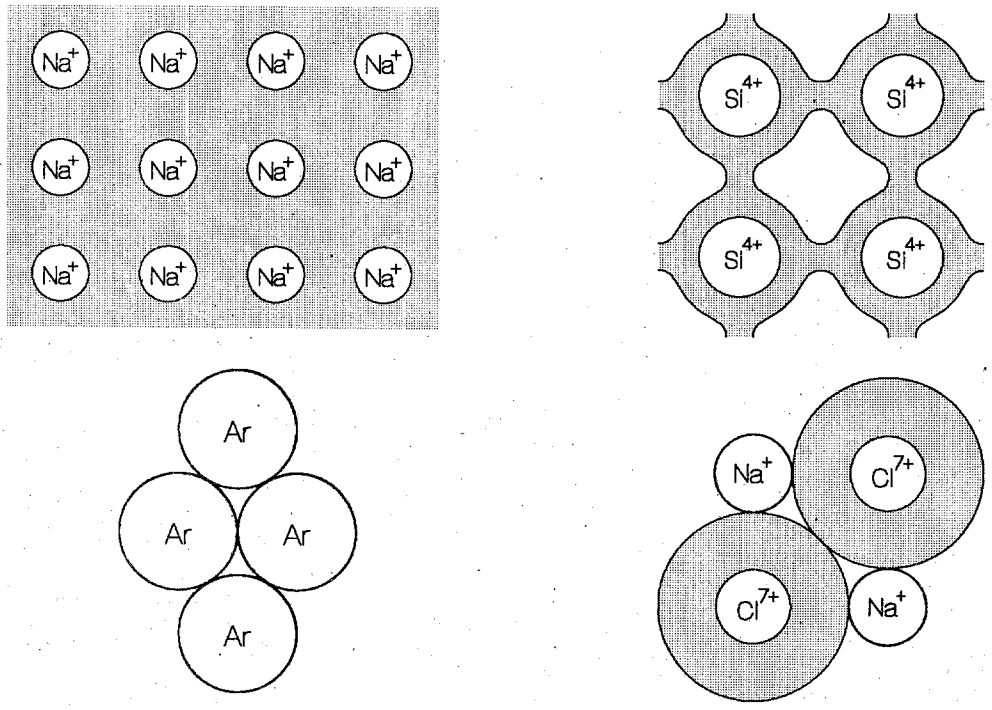


Abb. 3.2: Schematische Darstellung der Verteilung der Valenzelektronen in Festkörpern mit unterschiedlicher Bindung. Die dunklen Bereiche geben die bevorzugten Aufenthaltsorte der Valenzelektronen wieder. Im Falle von Chlor werden damit beispielsweise acht Elektronen angedeutet. Als Vertreter für die jeweiligen Bindungsarten wurden Natrium, Silizium, Argon und Natriumchlorid gewählt. Die Reihe führt von der metallischen über die kovalente und die Van der Waals- zur Ionenbindung.

Bindungsenergie

Die Kräfte zwischen den Atomen und Molekülen der Festkörper sind unabhängig vom Bindungstyp ausschließlich elektrostatischer Natur, d.h. sie beruhen auf den Coulomb-Kräften zwischen den Kernen und Elektronen. Ein Maß für die Stärke der Wechselwirkung und die Zahl der wechselwirkenden Nachbarn ist die *Bindungsenergie* oder *Gitterenergie.* Sie ist definiert als die Differenz zwischen der Energie der freien Atome oder Ionen und der Energie des Festkörpers am absoluten Nullpunkt. Beim Vergleich der Energien wird vorausgesetzt, daß sich Festkörper und Atome oder Ionen im Grundzustand befinden. Des weiteren müssen kinetische Energie und Wechselwirkungsenergie der freien Atome oder Ionen vernachlässigbar sein. Es ist bemerkenswert, daß beim Festkörper auch die Nullpunktsenergie in die Energiebilanz eingeht und die Bindungsenergie merklich reduzieren kann.

In Tabelle 3.1 sind die Bindungsenergien pro Atom für die Elemente der zweiten Reihe des Periodensystems angegeben. Geht man von links nach rechts, so nimmt der metallische Charakter der Bindung ab und der kovalente zu. Da die kovalente Bindung im allgemeinen stärker als die metallische ist, wächst die Bindungsenergie an. Während Lithium allein durch die metallische Bindung zusammengehalten wird, überwiegt beim

„Halbmetall" Bor bereits die kovalente Bindung. Diamant weist den Prototyp einer kovalenten Bindung auf. Die restlichen Elemente von Stickstoff bis Fluor sind in ihrer festen Form Molekülkristalle. Die hohe Bindungsenergie dieser Materialien beruht auf der großen Energie, die bei der Bildung der Moleküle frei wird. Die Bindung zwischen den Molekülen beruht auf der schwachen Van der Waals-Wechselwirkung und trägt vergleichsweise wenig zur Bindungsenergie bei. Es ist deshalb nicht verwunderlich, daß Stickstoff oder Sauerstoff schon bei relativ tiefen Temperaturen schmelzen bzw. sieden. Da beim festen Neon nur noch Van der Waals-Kräfte zum Tragen kommen, ist dort die Bindungsenergie äußerst klein.

Tabelle 3.1: Bindungsenergie der Elemente der zweiten Reihe des Periodensystems. Die Angaben beim Kohlenstoff beziehen sich auf Diamant. (Aus Ch. Kittel, Einführung in die Festkörperphysik, Oldenbourg, 1988).

	Li	Be	B	C	N	O	F	Ne
Bindungsenergie (eV/Atom)	1,6	3,3	5,8	7,4	4,9	2,6	0,8	0,02

Kristalle und amorphe Festkörper weisen leicht unterschiedliche Bindungsenergien auf, doch ist dies nicht verwunderlich. Wie wir sehen werden, tragen nicht nur die unmittelbar benachbarten, sondern auch weiter entfernte Atome zur Bindung eines herausgegriffenen Atoms bei. Dieser Beitrag wird durch die Fernordnung maximiert. Ihr Fehlen im Falle von amorphen Substanzen bewirkt daher eine Reduktion der Bindungsenergie. Wie groß diese Änderung ist, hängt vor allem von der Reichweite der Bindungskräfte ab, aber auch die Ausprägung der lokalen Ordnung im betrachteten amorphen Material spielt eine wesentliche Rolle.

Abstoßung

Nähern sich zwei Atome, so unterliegen sie Kräften, die in der Mehrzahl der Fälle bei relativ großen Abständen attraktiv, bei kleinen aber repulsiv sind. In Abb. 3.3 ist der generelle Verlauf des Wechselwirkungspotentials zu sehen. Der tatsächliche Verlauf hängt von der Art der Bindungskräfte ab. In dem angesprochenen Bild ist das sogenannte *Lennard-Jones-Potential* skizziert, das zwischen neutralen Atomen oder Molekülen mit abgeschlossener Elektronenschale auftritt. Wir werden auf diesen Potentialverlauf, beschrieben durch (3.4), noch ausführlich im nächsten Abschnitt im Zusammenhang mit den Edelgaskristallen eingehen.

Für unsere Diskussion ist hier nur von Bedeutung, daß bei kleinen Abständen die Anziehung in eine Abstoßung übergeht, wodurch eine Durchdringung der Atome verhindert wird. Man könnte glauben, daß die klassische Coulomb-Abstoßung der Elektronenwolken hierfür verantwortlich ist. Wie jedoch eine einfache Rechnung zeigt, führt dieser Effekt zu einer viel zu schwachen Abstandsabhängigkeit, d.h. die Atome verhielten sich wie relativ weiche Kugeln. Tatsächlich ist die Wirksamkeit des *Pauli-Prinzips* ausschlaggebend, da sich bei kleinen Abständen die Wellenfunktionen der Atome überlappen. In diesem Fall müssen wir die wechselwirkenden Atome als Einheit sehen, in der die Elektronen die verfügbaren Zustände gemeinsam einnehmen. Sind freie tiefliegende Zustände verfügbar,

dann werden diese besetzt. Die damit verbundene Veränderung der Aufenthaltswahrscheinlichkeit der Elektronen führt zu einer Absenkung der Gesamtenergie und möglicherweise zur kovalenten Bindung der Atome. Auf die Fragen, die damit verknüpft sind, werden wir in Abschnitt 3.4 ausführlich eingehen. Sind keine unbesetzten tiefliegenden Zustände vorhanden, dann verbietet das Pauli-Prinzip die Besetzung dieser Zustände durch weitere Elektronen, die dann nur unbesetzte Zustände höherer Energie einnehmen können. Dadurch erhöht sich die potentielle Energie der beiden Atome, woraus eine abstoßende Kraft resultiert. Das Pauli-Prinzip bewirkt also, daß sich die Atome fast wie „harte Kugeln" verhalten.

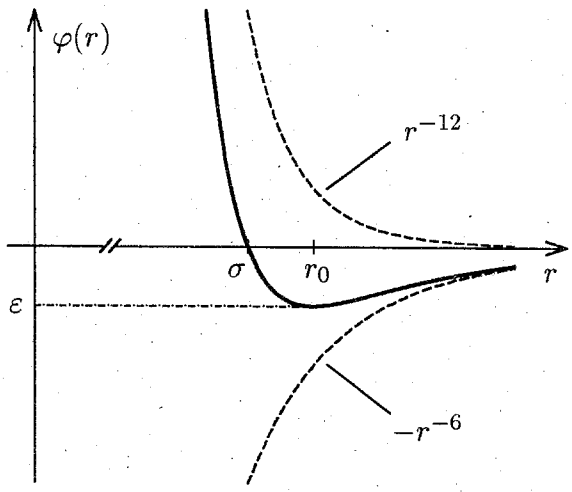


Abb. 3.3: Lennard-Jones-Potential. Der abstoßende Anteil des Potentials wird durch den r^{-12}-Term, das anziehende Van der Waals-Potential durch den r^{-6}-Term beschrieben. Man beachte, daß die Abszisse bei kleinen Abstandswerten unterbrochen ist. Der Parameter ε gibt die Muldentiefe an, σ bestimmt den Nulldurchgang des Potentials.

Zur empirischen Beschreibung der Abstoßung gebraucht man verschiedene analytische Ausdrücke. Häufig wird wie in Abb. 3.3 das Potential

$$\varphi(r) = \frac{A}{r^{12}} \tag{3.1}$$

benutzt, wobei die Konstante A die Potentialstärke angibt und die Steilheit des Potentialabfalls durch den Wert des Exponenten festgelegt wird[1]. Da es in den meisten Fällen nicht auf den exakten Kurvenverlauf ankommt, reicht eine näherungsweise Beschreibung des Potentials mit einer einfachen Funktion. Einen tieferen Grund für die Wahl gerade dieser Potenz gibt es nicht. Häufig wird auch ein exponentieller Verlauf für das abstoßende Potential angesetzt, eine Näherung, die auf *Born* und *Mayer* zurückgeht. Dieses hat dann die Form

$$\varphi(r) = A' \mathrm{e}^{-r/r_0} \,, \tag{3.2}$$

wobei r_0 die Reichweite des Potentials festlegt und A' wiederum eine Konstante ist.

3.2 Van der Waals-Bindung

Wir beginnen die Diskussion der Bindungskräfte mit den Kristallen, die am einfachsten aufgebaut sind, nämlich den Edelgaskristallen. Da Edelgasatome abgeschlossene Elektronenschalen besitzen, gehen sie normalerweise keine chemische Bindung mit Nachbarn ein

[1] Natürlich gilt der funktionale Zusammenhang nur im Bereich $r \approx \sigma$. Bei kleinen Werten von r flacht der Potentialverlauf ab.

und weisen auch im Festkörper in sehr guter Näherung Kugelgestalt auf. Edelgase bilden daher Kristalle mit hoher Symmetrie und einatomiger Basis. Zwischen den Atomen wirken nur die schwachen Van der Waals-Kräfte. Ähnlich verhalten sich, wie bereits erwähnt, auch viele Molekülkristalle. Obwohl dort die Annahme kugelförmiger Gitterbausteine meist nur näherungsweise zutrifft, lassen sich viele Eigenschaften der Edelgaskristalle auf die sehr wichtige Klasse der Molekülkristalle übertragen.

Van der Waals-Kräfte

Zwischen allen Atomen wirken Van der Waals-Kräfte, die auf der elektrischen *Dipol-Dipol-Wechselwirkung* beruhen. Bei Edelgasatomen ist die Existenz dieser Kraft zunächst etwas überraschend, da kugelsymmetrische Atome kein permanentes Dipolmoment aufweisen. Es ist jedoch zu bedenken, daß immer Nullpunktsfluktuationen der Ladungsverteilung auftreten. Die Verschiebungen der Elektronenwolke gegen den Kern ruft ein Dipolmoment p_1 hervor, das seinerseits eine Ladungsverschiebung im Nachbarn induziert. Die beiden Ladungsbewegungen sind so miteinander korreliert, daß sich die Dipole anziehen. Qualitativ gesehen verursacht ein Dipol p_1, der durch Ladungsfluktuation hervorgerufen wird, im Abstand r ein elektrisches Feld $\mathcal{E} \propto p_1/r^3$, wobei wir die Richtungsabhängigkeit vernachlässigt haben. Dieses Feld bewirkt im benachbarten Atom ein Dipolmoment p_2 proportional zur elektrischen Feldstärke. Da das Wechselwirkungspotential φ zweier parallel stehender Dipole durch $\varphi \propto -p_1 p_2/r^3$ gegeben ist, folgt für das Van der Waals-Potential

$$\varphi(r) = -\frac{B}{r^6} \,. \tag{3.3}$$

B ist eine für die wechselwirkenden Atome charakteristische Konstante. Mit Hilfe der quantenmechanischen Behandlung der van der Waals-Kräfte läßt sich zeigen, daß die geschilderte Wechselwirkung immer zu einer Energieabsenkung, also zu einer anziehenden Kraft führt, deren Größe durch die Polarisierbarkeit der wechselwirkenden Atome bestimmt wird.

Lennard-Jones-Potential

Addiert man zum abstoßenden Potential (3.1) das anziehend wirkende Van der Waals-Potential, so erhält man das bereits erwähnte *Lennard-Jones-Potential*

$$\varphi(r) = \frac{A}{r^{12}} - \frac{B}{r^6} \equiv 4\,\varepsilon \left[\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^{6} \right] , \tag{3.4}$$

dessen Verlauf in Abb. 3.3 zu sehen ist. Der Parameter ε gibt die Tiefe der Potentialmulde an, σ bestimmt den Nulldurchgang des Potentials. Das Potentialminimum tritt beim Abstand $r_0 = 2^{1/6}\sigma = 1{,}12\,\sigma$ auf.

In der Tabelle 3.2 sind eine Reihe von Materialparametern von Edelgasatomen und Edelgaskristallen aufgeführt, auf die wir noch eingehen werden. Die beiden ersten Zeilen enthalten Werte der hier diskutierten Parameter σ und ε. Diese Werte kann man beispielsweise durch die Untersuchung des Streuverhaltens von Edelgasatomstrahlen an Edelgasen gewinnen. Aus der Winkelabhängigkeit der Streuintensität lassen sich die Wechselwirkungspotentiale direkt ermitteln. Eine weitere Möglichkeit besteht in der Messung von thermodynamischen Zustandsgrößen der Edelgase, aus der sich die gewünschten Parameter unter Zuhilfenahme einer leicht modifizierten Van der Waals-Gleichung für reale Gase

herleiten lassen. Wir haben in die Tabelle jedoch die Werte eingetragen, die sich aus den Kristalleigenschaften ergeben. Zu diesen Eigenschaften zählen unter anderem der ebenfalls aufgeführte atomare Gleichgewichtsabstand R_0 der Atome, die Sublimationswärme und die Kompressibilität. Die Parametersätze aus der gasförmigen und der festen Phase unterscheiden sich nur geringfügig.

Tabelle 3.2: Wichtige Materialparameter der Edelgasatome und Edelgaskristalle. Die Größen werden im Text erläutert. (Die Daten wurden verschiedenen Lehrbüchern entnommen).

	Ne	Ar	Kr	Xe
σ (Å)	2,74	3,40	3,65	3,98
ε (10^{-2} eV)	0,31	1,04	1,41	2,00
R_0 (Å)	3,16	3,76	3,99	4,35
R_0/σ	1,15	1,11	1,09	1,09
$U_B/N = -8{,}60\,\varepsilon$ (eV)	- 0,026	- 0,089	- 0,127	- 0,174
U_0 (eV)	0,008	0,009	0,007	0,006
$[8{,}60\varepsilon + U_0]$ (eV)	- 0,018	- 0,080	- 0,120	- 0,168
Gemessene Bindungsenergie (eV)	- 0,019	- 0,081	- 0,116	- 0,166

Gitterenergie

Nun wollen wir die Bindungsenergie der Edelgaskristalle berechnen. Der Einfachheit halber berücksichtigen wir nur die potentielle Energie, obwohl diese Näherung bei den Edelgaskristallen nicht sehr gut ist, da die Nullpunktsenergie eine merkliche Reduktion der Bindungsenergie bewirkt. Am Ende dieses Abschnitts werden wir ihren Beitrag auf einfache Weise abschätzen und mit der Bindungsenergie vergleichen. Da die Nullpunktsenergie mit abnehmender Masse zunimmt, verursacht sie bei den leichten Edelgaskristallen erhebliche Korrekturen. Helium, das leichteste Edelgas, bleibt wegen der hohen Nullpunktsenergie auch am absoluten Nullpunkt noch flüssig. Verfestigung läßt sich nur durch Anwendung eines Drucks erreichen, der größer als 2,5 MPa ist.

Die Bindungsenergie U_B der Kristalle ergibt sich aus der Summe der Beiträge der wechselwirkenden Paare. Die Energie φ_m, mit der das einzelne Atom m gebunden ist, erhalten wir, indem wir alle Beiträge aufsummieren, die von der Wechselwirkung mit den restlichen Atomen n herrühren. Für diese Größe können wir daher $\varphi_m = \sum_{n \neq m} \varphi_{mn}$ schreiben, wobei φ_{mn} für die Wechselwirkungsenergie des Atompaars m, n steht[2]. Sie ist durch das Lennard-Jones-Potential (3.4) gegeben. Bezeichnen wir die Anzahl der Kristallatome mit N, so führt die Summation über die wechselwirkenden Paare zum Faktor $N/2$, und es ergibt sich

$$U_B = \frac{1}{2}\sum_m \varphi_m = \frac{N}{2}\varphi_m = 2N\varepsilon \sum_{m \neq n} \left[\left(\frac{\sigma}{r_{mn}}\right)^{12} - \left(\frac{\sigma}{r_{mn}}\right)^{6} \right] . \qquad (3.5)$$

[2] An dieser Stelle soll bemerkt werden, daß bei dieser Vorgehensweise keine Vielkörperkräfte berücksichtigt werden, d.h. die Veränderung der Wechselwirkung zwischen zwei Atomen durch die Gegenwart der Nachbaratome wird außer acht gelassen.

Nun drücken wir den Abstand r_{mn} der Atome in Einheiten von R, dem Abstand des nächsten Nachbarn aus und schreiben hierfür $r_{mn} = p_{mn}R$. Aus der kubisch flächenzentrierten Struktur der Edelgaskristalle folgen die Werte $p_{mn} = 1, \sqrt{2}, , 2 \ldots$ für die nächsten, übernächsten, drittnächsten Nachbarn, usw. Desweiteren müssen wir noch die Zahl der Atome in den jeweiligen Abständen berücksichtigen. Man findet 12 nächste, 6 übernächste Nachbarn und so fort. Damit läßt sich (3.5) schließlich in der Form

$$U_{\mathrm{B}} = 2N\varepsilon \left[\left(\frac{\sigma}{R}\right)^{12} \left(\frac{12}{1^{12}} + \frac{6}{(\sqrt{2})^{12}} + \ldots \right) - \left(\frac{\sigma}{R}\right)^{6} \left(\frac{12}{1^{6}} + \frac{6}{(\sqrt{2})^{6}} + \ldots \right) \right] \tag{3.6}$$

schreiben. Für die erste der sehr rasch konvergierenden Reihen erhält man den Wert 12,13 und für die zweite den Wert 14,45. Erwartungsgemäß überwiegt der anziehende Beitrag zur potentiellen Energie.

Hieraus können wir sofort die Reichweite σ des Potentials bestimmen, wenn wir die Tatsache ausnutzen, daß die Bindungsenergie beim Gleichgewichtsabstand R_0 der nächsten Nachbarn ein Minimum besitzt. Aus der Bedingung $[\mathrm{d}U_{\mathrm{B}}/\mathrm{d}R]_{R_0} = 0$, dem Stabilitätskriterium, ergibt sich

$$\frac{R_0}{\sigma} = 1{,}09 \,. \tag{3.7}$$

Diese Beziehung sollte für alle Edelgaskristalle gelten. Offensichtlich ist in Kristallen der Gleichgewichtsabstand der Atome etwas kleiner, als er sich aus dem Lennard-Jones-Potential (3.4) ergibt, denn dort hatten wir das Verhältnis $r_0/\sigma = 1{,}12$ gefunden. Die experimentell bestimmten Werte für die Kristalle sind in der 4. Zeile der Tabelle 3.2 aufgeführt. Während man bei den leichten Edelgasen kleine Abweichungen findet, ist die Übereinstimmung bei den schweren perfekt.

Setzt man (3.7) in (3.6) ein, so findet man für die Bindungsenergie

$$U_{\mathrm{B}}(R_0) = -8{,}60\, N\varepsilon \,. \tag{3.8}$$

Mit den Daten für ε aus Tabelle 3.2 errechnen sich die Bindungsenergien der Zeile 5, die höher sind als die in der letzten Zeile angegebenen experimentellen Werte. Die Ursache hierfür und auch für die Abweichungen des Verhältnisses R_0/σ vom idealen Wert liegt in der bereits erwähnten Nullpunktsbewegung. Diese tritt auf, weil sich die Atome in einem Käfig befinden, der durch die Nachbarn gebildet wird. Wir wollen die Nullpunktsenergie U_0, die in der Zeile 6 aufgelistet ist, hier nur abschätzen, denn eine exakte Rechnung setzt die Kenntnis des Spektrums der Gitterschwingungen voraus. Bei dieser Abschätzung gehen wir vom Lennard-Jones-Potential aus und nähern das Potentialminimum durch eine Parabel. Die Schwingungsfrequenz der Atome und ihre Nullpunktsenergie läßt sich leicht berechnen. Es ergibt sich die Beziehung $U_0 \propto (\varepsilon/R_0^2\, M)^{1/2}$, die zu einer guten qualitativen Übereinstimmung mit den Tabellenwerten führt. Die mangelnde quantitative Übereinstimmung ist nicht überraschend, da wir nur zwei wechselwirkende Atome betrachtet und die Dreidimensionalität der Nullpunktsbewegung außer acht gelassen haben. Wir wollen hier die Abschätzung nicht verbessern, da wir eine korrekte Berechnung der Nullpunktsenergie in Abschnitt 5.5 durchführen werden. Bei den in der vorletzten Zeile der Tabelle aufgeführten theoretischen Werten $(8{,}60\,\varepsilon + U_0)$ ist der Beitrag der Nullpunktsbewegung zur Bindungsenergie berücksichtigt. Ein Vergleich mit den experimentelle Bindungsenergien der letzten Zeile zeigt, daß eine sehr gute Übereinstimmung zwischen den theoretischen und experimentellen Resultaten besteht.

Aufgrund der Nullpunktsbewegung ist der Atomabstand selbst am absoluten Nullpunkt nicht exakt durch das Minimum der klassischen potentiellen Energie gegeben ist. Der Gleichgewichtsabstand wächst bei vorgegebenem Potential mit der Nullpunktsenergie und somit mit abnehmender Atommasse. Eine direkte Folge ist das Auftreten eines Isotopeneffekts bei den Gitterkonstanten. So findet man die Werte $R_0 = 4,4644\,\text{Å}$ bei ^{20}Ne und $R_0 = 4,4559\,\text{Å}$ bei ^{22}Ne.

3.3 Ionenbindung

Abschätzung und Messung der Bindungsenergie

Die *ionische* oder *heteropolare Bindung* ist durch den Elektronentransfer zwischen den Atomen gekennzeichnet, der zur Ausbildung von Ionen führt. Die starke Coulomb-Wechselwirkung zwischen den Ionen ist richtungsunabhängig und erlaubt wie die Van der Waals-Bindung eine hohe Raumerfüllung. Als Vertreter der Ionenkristalle greifen wir Natriumchlorid heraus, bei dem die Ionen gerade eine Elementarladung tragen. In Abb. 3.4 ist die aus Experimenten der Röntgenbeugung hergeleitete Elektronendichte wiedergegeben, aus der hervorgeht, daß die Ionen tatsächlich die erwartete Kugelsymmetrie besitzen. Nur am Rande der Cl^--Ionen findet man kleine Abweichungen. Bei der Bewertung der experimentellen Daten ist zu beachten, daß die Konturlinien einer logarithmischen Skala folgen, die Abweichungen von der Kugelsymmetrie also erst bei sehr kleinen Elektronendichten auftreten.

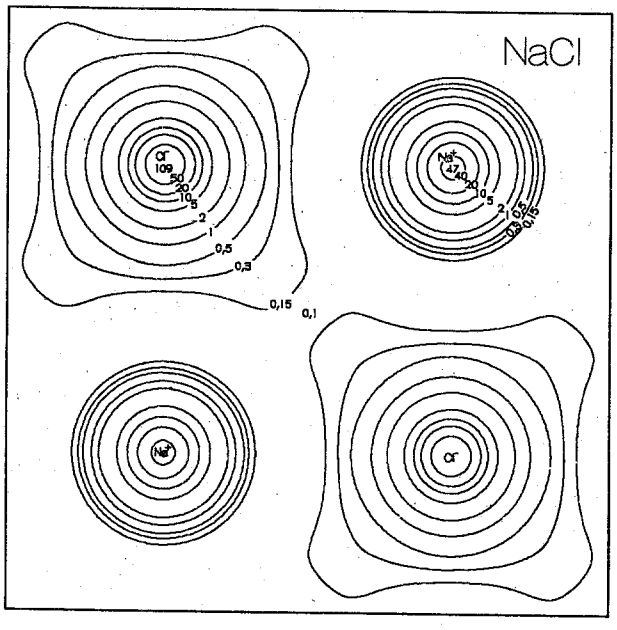


Abb. 3.4: Elektronendichte in der (100)-Ebene des Ionenkristalls NaCl. Die Elektronendichte der Ionen ist nahezu kugelsymmetrisch. Die Konturlinien geben die Dichte der Elektronen pro Å^3 an. (Nach S. Göttlicher, Acta Cryst. B **24** (1968) 122).

Die Bindungsenergie eines Paares kann sofort angegeben werden, da sie sich aus der Anziehung zwischen zwei geladenen Kugeln berechnen läßt. Mit $R_0 = 2,8\,\text{Å}$ für NaCl ergibt sich $\varphi = -e^2/4\pi\epsilon_0 R_0 = -5,1\,\text{eV}$. Da die Ionen im Kristall aber nicht nur mit *einem* Nachbarn wechselwirken, erwarten wir tatsächlich eine etwas stärkere Bindung, doch sollte die Größenordnung richtig getroffen sein. Die Größe der Bindungsenergie

erlaubt sofort den Schluß, daß bei Ionenkristallen die schwachen Van der Waals-Kräfte nicht beachtet werden müssen und die Nullpunktsenergie vernachlässigt werden kann.

Die Bindungsenergie der Ionenkristalle läßt sich nicht unmittelbar einer Messung entnehmen. Entsprechend der Definiton ist sie die Energie, die aufgewendet werden muß, um den Festkörper in seine Bestandteile zu zerlegen, geteilt durch die Zahl der Atome. Im Falle des Kochsalzes bedeutet dies, daß man ein verdünntes Gas aus Na^+- und Cl^--Ionen herstellen müßte, was natürlich nicht möglich ist. Um dennoch die Bindungsenergie zu ermitteln, geht man den „Umweg" über einen erdachten *Kreisprozeß*. Man stellt sich eine Reihe von Reaktionsschritten vor, bei denen ein NaCl-Kristall in seine Bestandteile zerlegt und dann wieder zusammengebaut wird, wobei alle in dieser Reihe auftretenden Energien mit Ausnahme der Bindungsenergie bekannt sind.

Wir wollen die erforderlichen Schritte anhand des *Born-Haber-Kreisprozesses* kurz erläutern. Man startet mit einem NaCl-Kristall und wendet die unbekannte Bindungsenergie U_B auf, um das geforderte verdünnte Gas aus Na^+- und Cl^--Ionen zu erzeugen. Im nächsten Schritt werden die Ionen durch Einfang bzw. Abgabe eines Elektrons neutralisiert. Im ersten Fall wird Energie frei, im zweiten muß Energie aufgewendet werden. Aus spektroskopischen Messungen sind die auftretenden Energien bekannt. Der folgende Schritt besteht in der Kondensation der Natriumatome zu Metall und in der Bildung von Cl_2-Molekülen. Die frei werdenden Energien sind thermochemischen Messungen zugänglich. Im letzten Schritt läßt man die beiden Materialien miteinander reagieren und registriert thermochemisch die Bildungswärme. Es entsteht ein NaCl-Kristall, und der Kreis des Gedankenexperiments schließt sich. Addiert man die Beiträge unter Berücksichtigung der Vorzeichen auf, so findet man $U_B = -8{,}15\,\text{eV}$ pro Ionenpaar. Das ist erwartungsgemäß etwas mehr als unsere Abschätzung von $-5{,}1\,\text{eV}$ für ein einzelnes Paar.

Berechnung der Bindungsenergie

Bei der Berechnung der Bindungsenergie der Ionenkristalle kann ähnlich wie bei den Edelgaskristallen verfahren werden. Für das abstoßende Potential benutzen wir wieder (3.1), obwohl bei Ionen der exponentielle Verlauf (3.2) die experimentellen Gegebenheiten etwas besser wiedergibt. Weiter berücksichtigen wir, daß der Coulomb-Term je nach Ladung der wechselwirkenden Partner unterschiedliche Vorzeichen hat. Für die Bindungsenergie pro Ion erhalten wir

$$\begin{aligned}\varphi_m &= \sum_{n\neq m} \frac{A}{r_{mn}^{12}} \pm \frac{e^2}{4\pi\varepsilon_0 r_{mn}} \\ &\approx \frac{zA}{R^{12}} - \sum_{n\neq m} \frac{\pm e^2}{4\pi\varepsilon_0 p_{mn} R}\,. \end{aligned} \tag{3.9}$$

Die zweite Zeile ergibt sich aus der ersten unter der Annahme, daß es wegen der kurzen Reichweite der Abstoßung genügt, nur die z nächsten Ionen im Abstand R zu berücksichtigen. Das negative Vorzeichen sagt aus, daß die Wechselwirkung mit den unmittelbar benachbarten Ionen attraktiv ist. Wie bei den Edelgaskristallen messen wir auch hier den Abstand in Einheiten des Abstands nächstbenachbarter Ionen und schreiben hierfür wieder $r_{mn} = p_{mn} R$.

Nun führen wir noch die dimensionslose *Madelung-Konstante* α ein, welche die Summation über die Coulomb-Energien beschreibt und durch

$$\alpha \equiv \sum_{n \neq m} \frac{\pm 1}{p_{mn}} \tag{3.10}$$

definiert ist. Im einfachen Fall einer linearen Kette, die aus abwechselnd positiven und negativen Ionen besteht, läßt sich das Ergebnis der Summation sofort hinschreiben:

$$\alpha = 2\left(1 - \frac{1}{2} + \frac{1}{3} - + \ldots\right) = 2\ln 2 \approx 1{,}39\,. \tag{3.11}$$

Der Faktor 2 berücksichtigt, daß das betrachtete Ion mit Ionen von beiden Seiten wechselwirkt. Ein ernsthaftes Problem tritt bei dreidimensionalen Gittern auf: Das Coulomb-Potential fällt zwar mit r^{-1} ab, die Zahl der zu berücksichtigenden Ionen steigt jedoch mit dem Abstand stärker an. Die Summe konvergiert daher nur bedingt, d.h. man findet unterschiedliche Resultate, je nachdem, wie die Glieder der Reihe zusammengefaßt werden. Es wurden mehrere Verfahren entwickelt, die eindeutige Werte für unendlich große Kristalle liefern. Eines dieser Verfahren besteht darin, den Kristall in elektrisch neutrale Zellen, sogenannte *Evjen-Zellen* zu unterteilen, die die Translationssymmetrie des Gitters besitzen. In Abb. 3.5 ist eine derartige Zelle für den Fall der CsCl-Struktur gezeigt. Nachdem man das Potential der Ladungen in der Zelle aufsummiert hat, berechnet man im nächsten Schritt das Potential, das von den weiter entfernt liegenden Zellen am Bezugsatom hervorgerufen wird. Da jede Zelle neutral ist und mehrere Ladungen mit unterschiedlichen Vorzeichen enthält, handelt es sich um Multipole, deren Potential relativ rasch mit dem Abstand abfällt. Die Ordnung des Multipols wird durch die Zahl der Ladungen, also durch die Größe der gewählten Zelle bestimmt. Eine andere Möglichkeit besteht darin, die neutrale Evjen-Zelle schrittweise zu vergrößern, bis die gewünschte Konvergenz eintritt. Dieses Verfahren liefert eine rasch konvergierende Reihe, dennoch hängt der erforderliche Rechenaufwand stark von der vorliegenden Struktur ab. Die Zahlenwerte der Madelung-Konstanten für verschiedene Strukturen sind vergleichbar, so findet man $\alpha = 1{,}748$ für die NaCl-, $\alpha = 1{,}763$ für die CsCl- und $\alpha = 1{,}638$ für die Zinkblende-Struktur. Diese Werte sind nicht sehr weit von $\alpha = 1{,}39$, dem Wert für den „eindimensionalen Kristall“ der Gleichung (3.11), entfernt.

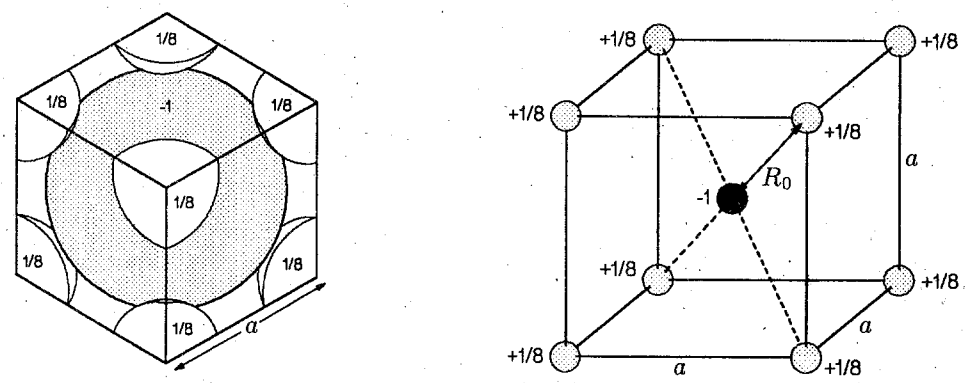


Abb. 3.5: Evjen-Zelle der Cäsiumchloridstruktur. Im Zentrum sitzt ein Cl^--Ion, die Ladungen der Cs^+-Ionen sind so aufgeteilt, daß die Zelle als Ganzes neutral ist. Dies ist ohne Verlust der Symmetrie nur möglich, wenn Ladungen „geteilt“ werden.

Um die Bindungsenergie des gesamten Festkörpers zu berechnen, multiplizieren wir die zweite Zeile von (3.9) mit der Anzahl N der Moleküle, d.h. hier mit der Anzahl der Ionenpaare, und erhalten das Zwischenergebnis

$$U_{\mathrm{B}} = N\varphi_m = N\left(\frac{zA}{R^{12}} - \frac{\alpha e^2}{4\pi\varepsilon_0 R}\right) . \tag{3.12}$$

Mit Hilfe des Stabilitätskriteriums läßt sich – wie im Fall der Edelgaskristalle – ein Parameter eliminieren. Wir wählen den Parameter zA und finden

$$U_{\mathrm{B}} = -\frac{N\alpha e^2}{4\pi\varepsilon_0 R_0}\left(1 - \frac{1}{12}\right) . \tag{3.13}$$

Der erste Term repräsentiert die anziehende „Madelung-Energie", der zweite die Abstoßung durch die nächsten Nachbarn. Dieses Ergebnis zeigt, daß die Abstoßung etwa 10% des Coulomb-Terms kompensiert. Damit errechnet sich für die Bindungsenergie von Natriumchlorid der Wert $U_{\mathrm{B}}/N = -8{,}25\,\mathrm{eV}$ pro Ionenpaar. Da der experimentelle Wert zu $U_{\mathrm{B}}/N \approx -8{,}15\,\mathrm{eV}$ bestimmt wurde, ist die Übereinstimmung befriedigend.

Offensichtlich wird der abstoßende Teil zur Bindungsenergie durch den Exponenten des Potentialabfalls bestimmt. Man könnte daher glauben, daß es vernünftiger wäre, zunächst den Exponenten in (3.1), den wir hier p nennen wollen, offenzulassen und diesen dann durch Anpassung von (3.13) zu bestimmen. Es zeigt sich aber, daß dieser Weg nicht sinnvoll beschritten werden kann, da die auftretenden Fehler zu groß sind. Besser geeignet ist die Bestimmung des Exponenten mit Hilfe einer weiteren Meßgröße. Hierfür kann man die Kompressibilität κ heranziehen, die mit der Bindungsenergie über deren zweite Ableitung nach dem Volumen verbunden ist[3]. Für Ionenkristalle findet man dabei Werte von p, die ein wenig kleiner als 10 sind, also auf einen etwas flacheren Verlauf des Potentials hindeuten als ursprünglich angenommen. Wie bereits erwähnt, verbessert sich die Übereinstimmung, wenn man von einem exponentiellen Verlauf des Potentials ausgeht.

Wir wollen nun noch die Stärke der Abstoßung bei NaCl mit der bei Edelgaskristallen vergleichen. Zunächst würde man erwarten, daß die abstoßungsbedingte Reduktion der Bindungsenergie vergleichbar ist. Setzt man die entsprechenden Parameter ein, so stellt man überraschenderweise fest, daß sich bei Kochsalz die Abstoßung wesentlich stärker auswirkt als beispielsweise bei festem Krypton. Die Erklärung hierfür ist recht einfach: Bei den Edelgaskristallen ist der Atomabstand viel größer als beim NaCl, so daß dort die Abstoßung weniger wirksam ist.

3.4 Kovalente Bindung

In vielen Festkörpern liegt *kovalente Bindung* vor, bei der die Elektronen nicht kugelsymmetrisch um den Kern verteilt, sondern verstärkt *zwischen* den Atomen lokalisiert sind. Diese Bindung tritt nicht nur in Festkörpern auf, sondern liegt in den meisten Fällen auch bei Molekülen vor. Besonders übersichtlich sind die Verhältnisse beim Wasserstoffmolekül, da dort insgesamt nur zwei Elektronen vorhanden sind. Wir wollen nun die Eigenheiten der kovalenten Bindung an diesem einfachen Beispiel studieren.

[3] Für Kristalle mit Natriumchloridstruktur ergibt eine einfache Rechnung unter Berücksichtigung der Beziehung $\kappa^{-1} = V(\mathrm{d}^2U/\mathrm{d}V^2)$ für den Exponenten p den Ausdruck $p = 1 + (72\pi\epsilon_0 R_0^4/\kappa\alpha e^2)$.

Wie wir sehen werden, beruht diese Art der Bindung auf einem quantenmechanischen Effekt, für den es in der klassischen Physik kein Analogon gibt. Ausgangspunkt der Beschreibung ist der Hamilton-Operator, der relativ viele Terme enthält. Um ihn zu vereinfachen, benutzen wir die *adiabatische* oder *Born-Oppenheimer-Näherung*, die auf der Tatsache beruht, daß sich im Atom die Elektronen viel rascher als die Kerne bewegen. Die Elektronen können daher jeder Kernbewegung praktisch unmittelbar folgen. Dies erlaubt eine Entkopplung von Kern- und Elektronenbewegung, so daß wir die Kernbewegung nicht weiter berücksichtigen müssen. Es verbleiben acht Terme, so daß der Hamilton-Operator des Wasserstoffmoleküls folgendes Aussehen hat:

$$\begin{aligned} H = &-\frac{\hbar^2}{2m}\Delta_1 - \frac{e^2}{4\pi\varepsilon_0 r_{a1}} - \frac{e^2}{4\pi\varepsilon_0 r_{b1}} + \frac{e^2}{4\pi\varepsilon_0 r_{ab}} \\ &-\frac{\hbar^2}{2m}\Delta_2 - \frac{e^2}{4\pi\varepsilon_0 r_{a2}} - \frac{e^2}{4\pi\varepsilon_0 r_{b2}} + \frac{e^2}{4\pi\varepsilon_0 r_{12}} . \end{aligned} \tag{3.14}$$

Die Bezeichnungen der auftretenden geometrischen Größen kann man Abb. 3.6 entnehmen: Die Indizes a und b benennen die beiden Kerne und die Zahlen 1 und 2 die Elektronen. Der erste Term in den Zeilen steht für die kinetische Energie der Elektronen 1 und 2. Die elektrische Anziehung bzw. Abstoßung zwischen den Elektronen und den Kernen drückt sich in den sechs Termen der potentiellen Energie aus. Im Prinzip lassen sich bei Kenntnis des Hamilton-Operators unter Zuhilfenahme der stationären Schrödinger-Gleichung die Eigenwerte und Eigenfunktionen des Wasserstoffmoleküls finden. Das Problem kann jedoch nicht analytisch gelöst werden. Es sind geeignete Vereinfachungen erforderlich, die gute Näherungslösungen erlauben. Wir gehen bei der Lösung schrittweise vor und betrachten zunächst das Wasserstoffatom, da sich anhand dieses einfachen und wohlbekannten Modellsystems die prinzipielle Vorgehensweise besonders gut erläutern läßt. Anschließend diskutieren wir das H_2^+-Molekülion, um schließlich zum Wasserstoffmolekül selbst zu kommen.

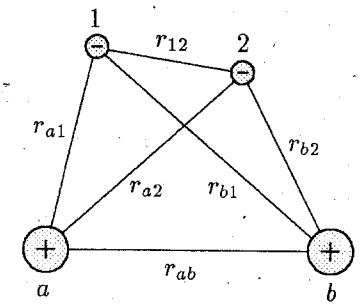


Abb. 3.6: Festlegung der Abstände, die in den Hamilton-Operator (3.14) des H_2-Moleküls eingehen.

Beim Wasserstoffatom wird nur *ein* Elektron von *einem* Proton angezogen. Der Hamilton-Operator besitzt daher eine sehr einfache Gestalt. Im Ausdruck (3.14) bleiben nur die beiden ersten Terme für die kinetische und potentielle Energie des Elektrons übrig. Wie in der Quantenmechanik üblich, bestimmen wir die Eigenwerte E der stationären Schrödinger-Gleichung $H\psi = E\psi$, indem wir diese Gleichung von der linken Seite mit ψ^* multiplizieren und über den ganzen Raum integrieren: $\int \psi^* H\psi \, \mathrm{d}V = \int \psi^* E\psi \, \mathrm{d}V$. Da

der Eigenwert E eine Zahl ist, die vor das Integral gezogen werden kann, erhalten wir für den Energieeigenwert E den bekannten Ausdruck

$$E = \frac{\int \psi^* H \psi \, \mathrm{d}V}{\int \psi^* \psi \, \mathrm{d}V} \,. \tag{3.15}$$

Die Wellenfunktion ψ_0 des Wasserstoffatoms im Grundstand setzen wir als bekannt voraus. Sie ist radialsymmetrisch und fällt exponentiell ab: $\psi_0(r) = (\pi a_0^3)^{-1/2} \exp(-r/a_0)$. In dieser Gleichung haben wir die Abkürzung $a_0 = (4\pi\varepsilon_0\hbar^2)/(me^2)$ für den Bohrschen Radius benutzt.

Da die Wellenfunktion ψ_0 reell ist, werden wir der Übersichtlichkeit halber in der weiteren Diskussion in den meisten Fällen nicht mehr zwischen ψ und ψ^* unterscheiden. Berücksichtigen wir, daß die Wellenfunktion des Wasserstoffatoms normiert ist, d.h. daß $\int \psi^* H \psi \, \mathrm{d}V = 1$ ist, so ergibt sich für die Energie des Grundzustands:

$$E_0 = \int \psi_0 H \psi_0 \, \mathrm{d}V = \int \psi_0(r) H \psi_0(r) \, 4\pi r^2 \mathrm{d}r = -\frac{\hbar^2}{2ma_0^2} = -13{,}6\,\mathrm{eV} \,. \tag{3.16}$$

Nun wollen wir den nächstkomplizierteren Schritt tun und den Grundzustand des H_2^+-Molekülions in ähnlicher Weise berechnen. Der Hamilton-Operator wird in diesem Fall durch die erste Zeile der Gleichung (3.14) wiedergegeben. Die Abstoßung zwischen den beiden Protonen hängt nur von deren Abstand r_{ab} und nicht von der Wellenfunktion des Elektrons ab. Wir lassen daher diesen Anteil zur Grundzustandsenergie der Einfachheit halber weg und fügen ihn wieder ein, nachdem wir die elektronischen Beiträge behandelt haben. Trotz der einfachen Struktur des Hamilton-Operators läßt sich das Eigenwertproblem nicht analytisch lösen. Wir setzen daher ein Variationsverfahren ein, das erlaubt, die Grundzustandsenergie des Wasserstoffmolekülions näherungsweise mit Hilfe eines geeigneten Lösungsansatzes zu berechnen. Ähnlich werden wir in späteren Kapiteln vorgehen, wenn exakte Lösungen nicht möglich sind. Wir gehen von zwei weit entfernten Kernen aus. In diesem Fall hält sich das Elektron entweder am Kern a oder Kern b auf und die zugehörigen Wellenfunktionen ψ_a und ψ_b sind durch Wellenfunktion ψ_0 des Wasserstoffatoms gegeben. Wir nehmen an, daß sich die Wellenfunktion ψ des Molekülions als Überlagerung der beiden Beiträge darstellen läßt und setzen die *Linearkombination*

$$\psi = \alpha \psi_a + \beta \psi_b \tag{3.17}$$

an, wobei wir die reellen Konstanten α und β noch bestimmen müssen. Mit Hilfe von Gleichung (3.15) läßt sich die Energie E des Grundzustandes berechnen. Wir finden

$$E = \frac{\int \psi^* H \psi \, \mathrm{d}V}{\int \psi^* \psi \, \mathrm{d}V} = \frac{\alpha^2 H_{aa} + \beta^2 H_{bb} + 2\alpha\beta H_{ab}}{\alpha^2 + \beta^2 + 2\alpha\beta S} \,. \tag{3.18}$$

Im letzten Ausdruck haben wir die Abkürzungen $H_{aa} = \int \psi_a^* H \psi_a \, \mathrm{d}V$, $H_{bb} = \int \psi_b^* H \psi_b \, \mathrm{d}V$, $H_{ab} = \int \psi_a^* H \psi_b \, \mathrm{d}V$ und $S = \int \psi_a^* \psi_b \, \mathrm{d}V$ benutzt, deren Bedeutung wir nun eingehender untersuchen werden[4].

[4] Bei der Festlegung der Abkürzungen wurde zwischen ψ und ψ^* unterschieden, da die Gleichung (3.18) nicht nur für den Grundzustand des Wasserstoffmolekülions Gültigkeit besitzt, sondern allgemeinerer Natur ist.

Sehen wir uns zunächst die Größe $S = \int \psi_a \psi_b \, \mathrm{d}V$ an. Sie beschreibt den Grad der Überlappung der beiden Wellenfunktionen ψ_a und ψ_b und besitzt einen endlichen Wert, da ψ_a und ψ_b nicht zueinander orthogonal sind.

Von den drei Gliedern des Hamilton-Operators, die zu den Energieerwartungswerten H_{aa} bzw. H_{bb} beitragen, haben wir die beiden ersten bereits beim Wasserstoffatom betrachtet und den Erwartungswert E_0 erhalten. Für H_{aa} können wir daher

$$H_{aa} = \int \psi_a H \psi_a \, \mathrm{d}V = E_0 + \int \psi_a(r_a) \left(-\frac{e^2}{4\pi\varepsilon_0 r_b} \right) \psi_a(r_a) \, \mathrm{d}V = E_0 + C \qquad (3.19)$$

schreiben[5]. Das rechte Integral, in dem die Elektronenladungsdichte $-e\psi_a^2$ auftritt und das wir mit C abgekürzt haben, beschreibt den Beitrag, den die Wechselwirkung zwischen dem Elektron und dem zweiten Kern zum Energieeigenwert leistet. Das negative Vorzeichen zeigt an, daß die Gegenwart des zweiten Kerns die potentielle Energie des Elektrons reduziert. Der numerische Wert von C hängt vom Kernabstand r_{ab} ab und verschwindet bei großen Abständen.

Die entscheidende Größe für die Bindungseigenschaften ist die Energie H_{ab}, deren Auftreten ausschließlich auf den quantenmechanischen Eigenschaften des Elektrons beruht. Schreiben wir die Integrale aus, so finden wir

$$H_{ab} = \int \psi_a H \psi_b \, \mathrm{d}V = E_0 S + \int \psi_a(r_a) \left(-\frac{e^2}{4\pi\varepsilon_0 r_a} \right) \psi_b(r_b) \, \mathrm{d}V = E_0 S + \mathcal{D} \, . \qquad (3.20)$$

In diesem Ausdruck finden wir statt der Elektronenladungsdichte $-e\psi_a^2$, die in Gleichung (3.19) auftritt, die sogenannte *Austauschdichte* $-e\psi_a\psi_b$. Dieser Beitrag zum Energieerwartungswert resultiert daraus, daß sich das Elektron teils im Zustand ψ_a und teils im Zustand ψ_b aufhält. Das entsprechende Integral bezeichnet man als *Austauschintegral* und die Größe $\mathcal{D}$ als *Austauschenergie*. Wie H_{aa} so ist auch $\mathcal{D}$ im allgemeinen negativ. In dem hier diskutierten Beispiel läßt sich dies mit einer einfachen, aber etwas längeren Rechnung zeigen.

Um den Energieeigenwert E tatsächlich angeben zu können, müssen erst die Konstanten α und β im Lösungsansatz bestimmt werden. Hier hilft die Tatsache, daß die exakte Lösung eines quantenmechanischen Problems mit den real auftretenden Wellenfunktionen immer zu kleineren Energieeigenwerten als die Näherungslösung führt. Wir kommen den wirklich vorliegenden Verhältnissen besonders nahe, wenn wir den Minimalwert von E bezüglich der Konstanten α und β aufsuchen, d.h. wenn wir $\partial E/\partial\alpha = 0$ und $\partial E/\partial\beta = 0$ als Bestimmungsgleichungen benutzen. Durch Ableiten von (3.18) findet man damit die Beziehungen

$$\begin{aligned} \alpha(H_{aa} - E) + \beta(H_{ab} - ES) &= 0 \\ \alpha(H_{ab} - ES) + \beta(H_{bb} - E) &= 0 \, . \end{aligned} \qquad (3.21)$$

Dieses einfache Gleichungssystem besitzt nichttriviale Lösungen, wenn die Koeffizientendeterminante verschwindet, d.h. wenn

$$(H_{aa} - E)(H_{bb} - E) - (H_{ab} - ES)^2 = 0 \qquad (3.22)$$

[5] Da nur ein Elektron vorhanden ist, wird hier der Index 1 bei den Elektronenkoordinaten weggelassen.

ist. Da die beiden betrachteten Kerne identisch sind, setzen wir $H_{aa} = H_{bb}$ und lösen nach E auf. Nun addieren wir noch die abstoßende Wechselwirkung der beiden Protonen und erhalten damit das Ergebnis

$$E_{\mathrm{g}} = \frac{H_{aa} + H_{ab}}{1+S} + \frac{e^2}{4\pi\varepsilon_0 r_{ab}} = E_0 + \frac{\mathcal{C}+\mathcal{D}}{1+S} + \frac{e^2}{4\pi\varepsilon_0 r_{ab}}$$

$$E_{\mathrm{u}} = \frac{H_{aa} - H_{ab}}{1-S} + \frac{e^2}{4\pi\varepsilon_0 r_{ab}} = E_0 + \frac{\mathcal{C}-\mathcal{D}}{1-S} + \frac{e^2}{4\pi\varepsilon_0 r_{ab}} \,. \tag{3.23}$$

Die Indizes g und u stehen für „gerade“ und „ungerade“. Wie wir gleich sehen werden, bezieht sich die Benennung auf die Wellenfunktionen, die mit den Energieerwartungswerten verknüpft sind. Der räumliche Überlapp der Wellenfunktionen ψ_a und ψ_b bewirkt das Auftreten der negativen Austauschenergie $\mathcal{D}$ und, damit verknüpft, die Aufhebung der Entartung der Grundzustände. Die beiden neuen Zustände E_{g} und E_{u} werden als *bindend* bzw. *antibindend* bezeichnet.

Mit Hilfe von (3.21) lassen sich bei bekanntem Wert von E die Konstanten α und β bestimmen. Für den bindenden Zustand E_{g} findet man $\alpha = \beta$ und für den antibindenden Zustand E_{u} die Beziehung $\alpha = -\beta$. Da auch die neue Wellenfunktion ψ normiert sein muß, gilt zusätzlich $|\alpha| = |\beta| \approx 1/\sqrt{2}$, wenn wir die Korrektur durch die Größe S vernachlässigen. Wir schreiben die so gefundene *gerade* und *ungerade* Lösung aus und erhalten

$$\psi_{\mathrm{g}} = \frac{1}{\sqrt{2}}\,(\psi_a + \psi_b) \qquad \text{und} \qquad \psi_{\mathrm{u}} = \frac{1}{\sqrt{2}}\,(\psi_a - \psi_b) \tag{3.24}$$

mit den Energiewerten E_{g} bzw. E_{u}.

In Abb. 3.7 ist der schematische Verlauf dieser Wellenfunktionen dargestellt. Für die Bindung ist entscheidend, daß sich die Amplituden der Wellenfunktionen ψ_a und ψ_b und nicht die Elektronendichten addieren. Beim bindenden Zustand führt das zu einer Elektronendichte, die zwischen den beiden Kernen doppelt so groß ist wie bei einer Addition der Einzelelektronendichten. Dagegen ist im antibindenden Zustand die Elektronendichte zwischen den Kernen stark reduziert; in der Mitte ist sie beim H_2^+-Molekülion sogar Null!

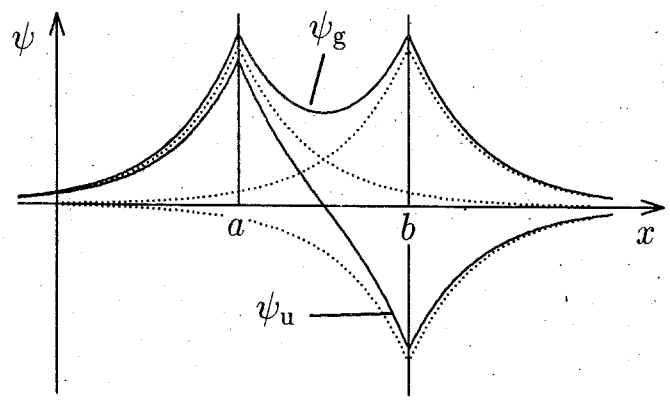


Abb. 3.7: Wellenfunktion des H_2^+-Molekülions. Die punktierten Kurven zeigen die Wellenfunktionen ψ_a, ψ_b bzw. $-\psi_b$. ψ_{g} ist die gerade, ψ_{u} die ungerade Gesamtwellenfunktion.

Abb. 3.8 zeigt den Verlauf der Energieerwartungswerte E_{g} und E_{u} als Funktion des Kernabstands. Bei sehr kleinen Abständen überwiegt in beiden Fällen die Coulomb-Abstoßung zwischen den Kernen. Bei größerem Kernabstand dagegen kann die Wechselwirkung attraktiv oder repulsiv sein. Die Energie E_{g} durchläuft ein Minimum, dessen Tiefe die Bindungsenergie widerspiegelt. Da E_{u} mit abnehmendem Abstand stetig

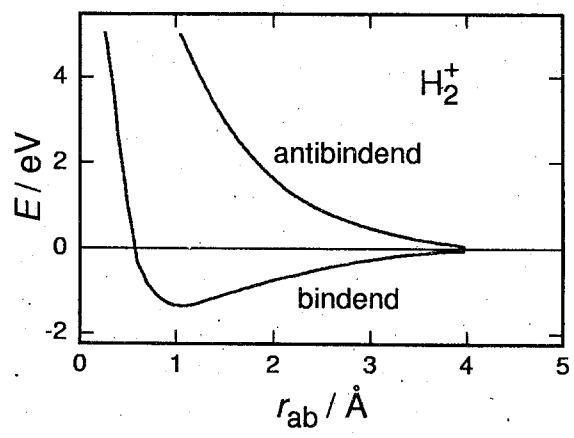


Abb. 3.8: Wechselwirkungsenergie des H_2^+-Molekülions als Funktion des Kernabstands r_{ab}.

anwächst, besteht in diesem Zustand zwischen den Wechselwirkungspartnern mit der Wellenfunktion ψ_u immer Abstoßung.

Wir wollen nun kurz auf die Frage eingehen, welche Veränderungen das *zweite Elektron* des Wasserstoffmoleküls hervorruft. Als erstes sei bemerkt, daß nun die *beiden* Elektronen durch eine *Zweiteilchen-Wellenfunktion* $\Psi(\boldsymbol{r}_1, \boldsymbol{r}_2)$ beschrieben werden. Wären die Elektronen völlig unabhängig voneinander, so wäre die Wellenfunktion ein Produkt der beiden Einelektronzustände. Die symmetrische Wellenfunktion beispielsweise hätte dann die Form $\Psi_g(\boldsymbol{r}_1, \boldsymbol{r}_2) \propto [\psi_a(\boldsymbol{r}_1) + \psi_b(\boldsymbol{r}_1)][\psi_a(\boldsymbol{r}_2) + \psi_b(\boldsymbol{r}_2)]$. In diesem Produkt kommen Terme vor, die den Zustand beschreiben, bei dem sich beide Elektronen in der Nähe von einem der beiden Kerne befinden. Die hohe Coulomb-Abstoßung zwischen den beiden Elektronen macht dies unwahrscheinlich. Die mathematische Behandlung des Problems vereinfacht sich, wenn man diese Terme, die nur bei hohen Anregungsenergien auftreten können, bei der Berechnung der Grundzustandsenergie erst gar nicht berücksichtigt. Dem Problem besser angepaßt ist daher der Ansatz

$$\Psi_g(\boldsymbol{r}_1, \boldsymbol{r}_2) \propto [\psi_a(\boldsymbol{r}_1)\psi_b(\boldsymbol{r}_2) + \psi_b(\boldsymbol{r}_1)\psi_a(\boldsymbol{r}_2)] \tag{3.25}$$

$$\Psi_u(\boldsymbol{r}_1, \boldsymbol{r}_2) \propto [\psi_a(\boldsymbol{r}_1)\psi_b(\boldsymbol{r}_2) - \psi_b(\boldsymbol{r}_1)\psi_a(\boldsymbol{r}_2)] \tag{3.26}$$

von *Heitler* und *London*, in dem die für den Grundzustand weniger wichtigen Terme fehlen. Die Indizes g und u stehen für „gerade" bzw. „ungerade" gegen den Austausch der Ortskoordinaten der Elektronen. Aus Gründen der Übersichtlichkeit wurde der Normierungsfaktor weggelassen.

Natürlich muß anstelle des bisher benutzten verkürzten Hamilton-Operators nun der volle Operator (3.14) berücksichtigt werden. Es ist bemerkenswert, daß der Spin der Elektronen nicht in den Hamilton-Operator eingeht und somit nicht direkt die Energie des Wasserstoffmoleküls bestimmt. Dies ist bei der Beschreibung der H_2^+-Molekülbindung unmittelbar einsichtig, denn bei einzelnen Elektronen im Grundzustand ist die Richtung des Spins bedeutungslos. Die Bindungsenergie wird daher nur durch die Ortswellenfunktion des Elektrons bestimmt. Dies ändert sich, wenn zwei Elektronen an der Bindung beteiligt sind. In Übereinstimmung mit dem Pauli-Prinzip muß die Gesamtwellenfunktion eines Systems aus mehreren Fermionen antisymmetrisch sein. Befinden sich die beiden an der Bindung beteiligten Elektronen beispielsweise in einem Zustand, der durch die Ortswellenfunktion $\Psi_u(\boldsymbol{r}_1, \boldsymbol{r}_2)$ beschrieben wird, so muß die Spinfunktion gerade sein. Dies ist der Fall, wenn die Spins der beiden Elektronen parallel zueinanderstehen, der Gesamtspin also Eins ist. Bezüglich eines äußeren Feldes existieren dann drei mögliche Einstellrichtungen.

Es handelt sich daher um den *Triplett-Zustand* des Wasserstoffmoleküls. Wird dagegen der Zustand durch die gerade Ortswellenfunktion $\Psi_g(\boldsymbol{r}_1, \boldsymbol{r}_2)$ beschrieben, dann muß die Spinfunktion ungerade sein. Dies ist der Fall, wenn die beiden Spins entgegengerichtet sind. Der Gesamtspin ist nun Null, das Wasserstoffmolekül befindet sich im *Singulett-Zustand.*

Mit den Ansätzen (3.25) bzw. (3.26) für die Wellenfunktion läßt sich der Erwartungswert der Energie berechnen, wobei eine große Zahl von Termen auftritt[6]. Das Ergebnis der Rechnung entspricht weitgehend dem Ergebnis beim Wasserstoffmolekülion. Auch beim Wasserstoffmolekül lassen sich die Beiträge zum Energieerwartungswert in zwei Klassen unterteilen, so daß wir vereinfachend

$$E_{g,u} \approx K \pm \mathcal{A} \tag{3.27}$$

schreiben können, wobei das positive Vorzeichen für die gerade Wellenfunktion gilt. K steht für die Glieder, die die klassische Wechselwirkung zwischen den elektrisch geladenen Teilchen widerspiegeln, während $\mathcal{A}$ die Beiträge berücksichtigt, die auf der Austauschwechselwirkung beruhen. Die Größe K setzt sich wie beim Molekülion zunächst aus $2E_0$, $2\mathcal{C}$ (der Faktor Zwei berücksichtigt jeweils, daß zwei Elektronen anwesend sind) und der Abstoßung $e^2/4\pi\varepsilon_0 r_{ab}$ der beiden Kerne zusammen. Hinzu kommt noch ein weiterer Beitrag, der die Abstoßung zwischen den beiden Elektronen beschreibt. Ähnlich wie beim Molekülion muß auch noch der Überlapp S der Wellenfunktionen berücksichtigt werden, doch ist er bei unserer einfachen Diskussion ohne Bedeutung.

Für die Austauschenergie $\mathcal{A}$ findet man

$$\mathcal{A} = \frac{2\mathcal{D} + \mathcal{E}}{1 \pm S^2}\,, \tag{3.28}$$

wobei $\mathcal{D}$ und $\mathcal{E}$ durch

$$\mathcal{D} = \int\int \psi_b(\boldsymbol{r}_1)\psi_a(\boldsymbol{r}_2)\left(\frac{-e^2}{4\pi\varepsilon_0 r_{b1}}\right)\psi_a(\boldsymbol{r}_1)\psi_b(\boldsymbol{r}_2)\,\mathrm{d}V_1\,\mathrm{d}V_2\,, \tag{3.29}$$

$$\mathcal{E} = \int\int \psi_b(\boldsymbol{r}_1)\psi_a(\boldsymbol{r}_2)\left(\frac{e^2}{4\pi\varepsilon_0 r_{12}}\right)\psi_a(\boldsymbol{r}_1)\psi_b(\boldsymbol{r}_2)\,\mathrm{d}V_1\,\mathrm{d}V_2\,. \tag{3.30}$$

gegeben sind[7]. $\mathcal{D}$ hat die gleiche Bedeutung wie in (3.20). $\mathcal{E}$ rührt von der Coulomb-Abstoßung zwischen den Elektronen her. Der numerische Wert von $\mathcal{D}$ bzw. $\mathcal{E}$ hängt von der speziellen Gestalt der Ortswellenfunktion ab. Da $\mathcal{D}$ negativ, die Größe $\mathcal{E}$ aber positiv ist, hängt es von der Summe der beiden Beiträge ab, ob es zu einer Bindung kommt oder nicht. Beim Wasserstoffmolekül ist $\mathcal{A}$ tatsächlich negativ und somit die Energiedifferenz $E_u - E_g \approx -2\mathcal{A}$ zwischen Triplett- und Singulett-Zustand positiv. Deshalb besitzt der Singulett-Zustand Ψ_g die niedrigere Energie, die bei nicht zu kleinen Kernabständen gegenüber der Energie der isolierten Atome abgesenkt ist.

Die Rechnung führt beim H_2-Molekül zu einer Abhängigkeit der Energie vom Kernabstand, wie wir sie in Abb. 3.8 bereits für das H_2^+-Molekül kennengelernt haben. Da an

[6] Eine ausführliche Diskussion aller Terme, die beim H_2^+-Molekül und beim Wasserstoffmolekül auftreten, findet man bei H. Haken, H.Ch. Wolf, *Atom- und Quantenphysik* (oder, *Molekülphysik und Quantenchemie*), Springer-Verlag, Heidelberg.

[7] Das positive Vorzeichen gehört zur geraden, das negative Vorzeichen zur ungeraden Wellenfunktion.

der Bindung nun zwei Elektronen beteiligt sind, ist das Minimum etwa 1,7 mal tiefer und der Bindungsabstand mit 0,74 Å etwas kleiner. Es muß jedoch bemerkt werden, daß die hier diskutierte Näherung zwar den generellen Verlauf der Energie mit dem Kernabstand richtig wiedergibt, aber zu einem theoretischen Wert der Bindungsenergie führt, der nur etwa 2/3 des experimentellen Wertes entspricht. Offensichtlich wird die Wellenfunktion bei der Näherung nach Heitler und London zu stark vereinfacht.

Es soll noch angefügt werden, daß auch bei einer klassischen Berechnung der potentiellen Energie ein Minimum auftritt, wenn man die Ladungswolken der isolierten Wasserstoffatome voraussetzt. Man erhält jedoch ein viel zu flaches Minimum, dessen Tiefe nur 1/15 des experimentellen Wertes besitzt. Die kovalente Bindung wird daher nur im Rahmen der Quantenmechanik verständlich.

Ein wichtiges Merkmal der kovalenten Bindung ist ihre starke Richtungsabhängigkeit. Da die Gestalt der Wellenfunktion der beteiligten Atome vorgegeben ist, ist ein maximaler Überlapp ihrer Wellenfunktionen, der zur größtmöglichen Bindungsenergie führt, außer bei s-Orbitalen nur in bestimmten Richtungen möglich. Experimentell läßt sich die Verteilung der Ladungsdichte mit Hilfe der Röntgenbeugung bestimmen, indem man die Formfaktoren der beteiligten Atome sehr genau ausmißt. Als Beispiel ist in Abb. 3.9 die Dichte der Valenzelektronen von Germanium in der $(1\bar{1}0)$-Fläche dargestellt. Offensichtlich sind die Germaniumatome nicht linear, sondern aufgrund der Richtungsabhängigkeit ihrer Bindung gewinkelt angeordnet. Deutliche zu sehen sind die Dichtemaxima zwischen den Atomrümpfen, die typisch für die kovalente Bindung sind.

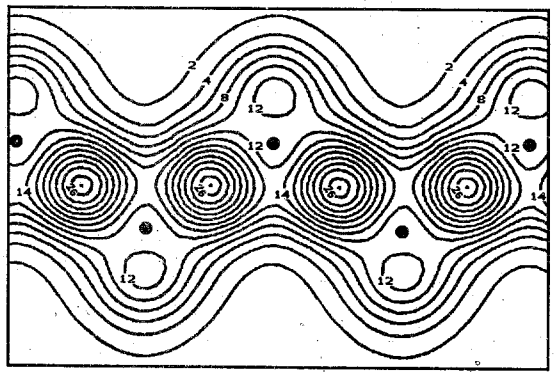


Abb. 3.9: Dichteverteilung der Valenzelektronen im Germanium. Die Zahlen an den Linien konstanter Dichte geben die Elektronenkonzentration pro primitiver Elementarzelle an. Die Lage der Ionenrümpfe ist durch große Punkte angedeutet. (Aus M.L. Cohen, Science **179** (1973) 1189).

Ein in der Festkörperphysik besonders interessanter und wichtiger Bindungstyp ist die *kovalente tetraedrische Bindung.* Man findet diese Art der Bindung unter anderem bei den Elementen der 4. Hauptgruppe wie z.B. beim Kohlenstoff, Silizium oder Germanium. Wir diskutieren diese spezielle Bindung am Beispiel des Kohlenstoffs, doch lassen sich die Argumente direkt auf die schwereren Elemente übertragen, denn zur Bindung tragen nur die außen liegenden Valenzelektronen bei. Die inneren Schalen sind daran nicht beteiligt. Freie Kohlenstoffatome besitzen die Elektronenkonfiguration $1s^2\,2s^2\,2p^2$. Die Ausbildung von sogenannten sp^3-Hybridorbitalen kann man sich folgendermaßen vorstellen: Zunächst entsteht aus der Konfiguration $2s^2\,2p^2$ der Valenzelektronen die Konfiguration $2s\,2p_x\,2p_y\,2p_z$, die allerdings energetisch ungünstiger ist, da die p-Zustände eine höhere Energie als die s-Zustände besitzen. Durch die Bildung der vier neuen Linearkombinationen wird maximaler Überlapp mit den Wellenfunktionen der nächsten Nachbarn ermöglicht. Die frei werdende Bindungsenergie überkompensiert den anfänglichen Energieverlust bei der Bildung der Hybridorbitale. Diamant weist daher

trotz schlechter Raumerfüllung und Koordinationszahl vier die hohe Bindungsenergie von 7,36 eV/Atom auf, bei Silizium sind es noch 4,64 eV/Atom, bei Germanium 3,87 eV/Atom.

Die vier Hybridorbitale lassen sich in sehr guter Näherung durch die lineare Überlagerung der ungestörten Zustände der freien Atome darstellen, d.h. die vier neuen Wellenfunktionen ψ_i haben die Form[8]:

$$\psi_i = \frac{1}{2}\left(\pm\psi_s \pm \psi_{p_x} \pm \psi_{p_y} \pm \psi_{p_z}\right) . \tag{3.31}$$

Wählt man nur positive (oder negative) Vorzeichen, so führt die Überlagerung der drei p-Wellenfunktionen zu einer p-artigen Wellenfunktion in [111]- und $[\bar{1}\,\bar{1}\,\bar{1}]$-Richtung. Die zusätzliche Überlagerung der s-Wellenfunktion bläht die Keule in [111]- auf und läßt sie in $[\bar{1}\,\bar{1}\,\bar{1}]$-Richtung schrumpfen, da dort die p-Wellenfunktion ein negatives (positives) Vorzeichen besitzt. Gleiches gilt für die drei anderen sp^3-Hybridorbitale, bei denen die Vorzeichen passend gewählt werden müssen. Man erhält dann die in Abb. 3.10 veranschaulichte Verteilung der Valenzelektronendichte.

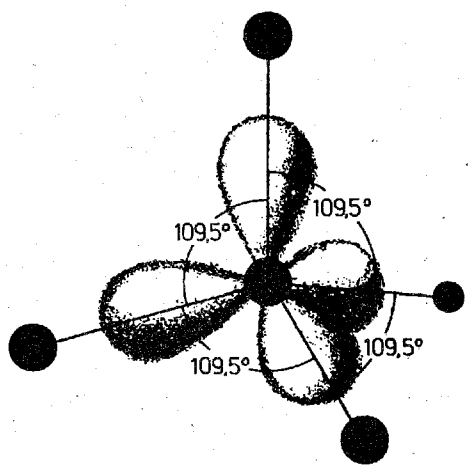


Abb. 3.10: Elektronendichteverteilung in den tetragonalen sp^3-Hybridorbitalen. Die Atommittelpunkte sind durch schwarze Kugeln angedeutet. (Aus H. Ibach, H. Lüth, *Festkörperphysik*, Springer-Verlag, Heidelberg 1988).

Statt des tetragonalen sp^3-Orbitals kann Kohlenstoff auch ein flächenhaftes trigonales sp^2-Hybridorbital bilden, das auf der Überlagerung einer $2s$- und zweier $2p$-Wellenfunktionen beruht. Die vierte, die $2p_z$-Wellenfunktion, steht senkrecht auf dem 120°-Stern des sp^2-Hybridorbitals und kann eine sogenannte π-Bindung eingehen. Ein bekanntes Beispiel hierfür ist das eingangs in Abschnitt 1.1 erwähnte Graphit (vgl. Abb. 1.1b), das aus übereinander gestapelten Schichten besteht, die durch van der Waals-Kräfte zusammengehalten werden. Die Kohlenstoffatome innerhalb einer Schicht bilden Sechsecke, die über die sp^2-Orbitale verknüpft sind. Zusätzliche Bindung wird durch die aus den Ebenen herausragenden π-Elektronen bewirkt, die vollständig delokalisiert sind. Der Atomabstand innerhalb der Schicht beträgt 1,4 Å, der Schichtabstand 3,35 Å.

Ein besonders schönes Beispiel für die sp^2-Hybridbindung stellt das Molekül C_{60} dar, das erst 1985 entdeckt wurde und als „Fulleren" bezeichnet wird. Wie Abb. 3.11 zeigt, besteht dieses Molekül aus 12 Fünfecken und 20 Sechsecken, enthält also 32 Ringe und hat

[8] Da das Vorzeichen der resultierenden Wellenfunktion ψ_i keine Rolle spielt, ergeben sich trotz der acht möglichen Vorzeichenkombinationen nur vier unterschiedliche Wellenfunktionen.

die Form eines Fußballs mit einem Durchmesser von 7,1 Å. Ähnlich wie beim Graphit ragen die π-Elektronen aus der Kugeloberfläche heraus und sind delokalisert. C_{60}-Moleküle können verschiedene Verbindungen eingehen und dreidimensionale Kristallstrukturen bilden. Bei Zimmertemperatur besitzen C_{60}-Kristalle eine kubisch flächenzentrierte Struktur mit einer Gitterkonstanten von 14,2 Å, wobei die „Kohlenstoffbälle" rasche Rotationsbewegungen um ihr Zentrum ausführen. Bei tieferen Temperaturen finden Übergänge zu zwei komplizierteren Phasen statt.

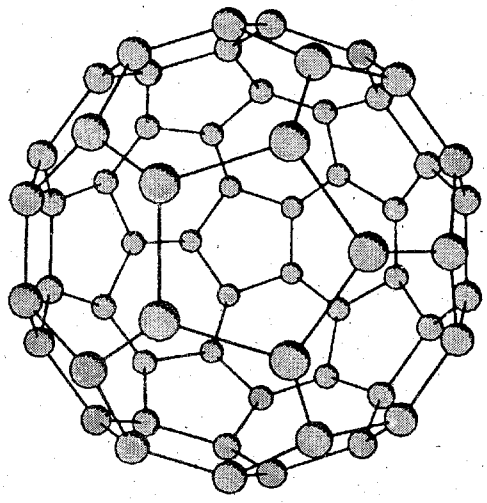


Abb. 3.11: Struktur des Moleküls C_{60}. (Nach H.W. Kroto et al., Nature **318** (1985) 162).

Kovalente Bindungen findet man auch bei den Elementen der 5. und 6. Gruppe, doch bilden sich in diesen Fällen Schicht- bzw. Kettenstrukturen aus. Natürlich treten auch Bindungen zwischen unterschiedlichen Elementen auf. Ein bekanntes Beispiel hierfür ist Bornitrid, das ein Diamantgitter besitzt, bei dem jedes Atom von vier Atomen der anderen Sorte umgeben ist. Jedes Boratom trägt drei, jedes Stickstoffatom fünf Elektronen zu den Bindungen bei. Wegen der unterschiedlichen Elektronegativität der beteiligten Elemente besitzt die Verbindung auch Ionencharakter.

3.5 Metallische Bindung

Wie bereits in Kap. 3.1 angedeutet, sind die Wellenfunktionen der Valenzelektronen metallbildender Atome bereits so ausgedehnt, daß ein starker Überlapp mit den Wellenfunktionen der Nachbarn auftreten würde, wenn sie im Gitterabstand angeordnet wären. In realen Metallen ist dieser Überlapp sogar noch wesentlich stärker ausgeprägt: Das Wesen der Metallbindung besteht darin, daß die Valenzelektronen weitgehend delokalisiert und somit fast gleichmäßig über den Kristall verschmiert sind. Da sich diese Elektronen dann relativ frei im Gitter der Ionenrümpfe bewegen können, spricht man von einem *Elektronengas*. Die metallische Bindung ist ungerichtet und ermöglicht eine optimale Raumerfüllung. Die Theorie dieses Bindungstyps ist jedoch wesentlich aufwendiger als bei der kovalenten Bindung, da eine einfache Überlagerung der Wellenfunktionen, wie wir sie in (3.25) und (3.26) durchgeführt haben, auch nicht mehr annähernd ausreicht.

Eine qualitative Betrachtung erlaubt aber bereits ein grobes Verständnis der Reduktion der Gesamtenergie bei der Ausbildung des Elektronengases. Unmittelbar einsichtig ist, daß die potentielle Energie der Elektronen am Bezugsatom durch das Coulomb-Potential der Nachbaratome abgesenkt wird. Diesen Sachverhalt haben wir bereits bei der kovalenten Bindung im Falle des H_2^+-Moleküls kennengelernt. Er ist nochmals in Abb. 3.12 verdeutlicht, in dem die Coulomb-Energie gestrichelt eingezeichnet ist. Daß auch die kinetische Energie vermindert wird, folgt bereits aus der Unschärferelation, denn die Elektronen sind im Metall weit schwächer als im freien Atom lokalisiert. Diese Schlußfolgerung läßt sich auch aus der Skizze 3.12 ablesen. Vergleicht man die eingezeichnete 3s-Wellenfunktion isolierter Atome mit der Wellenfunktion, die sich aus der Überlagerung ergibt, so sieht man, daß die Krümmung der Funktion in der Mitte zwischen den beiden Atomen, d.h. in der Nähe der Begrenzung der Wigner-Seitz-Zelle, wesentlich geringer ist. Dies bedeutet, daß die kinetische Energie dieser Elektronen reduziert ist, denn entsprechend der Schrödinger-Gleichung ist die kinetische Energie proportional zu $-\Delta\psi$, d.h. sie wird durch die Krümmung der Wellenfunktion bestimmt.

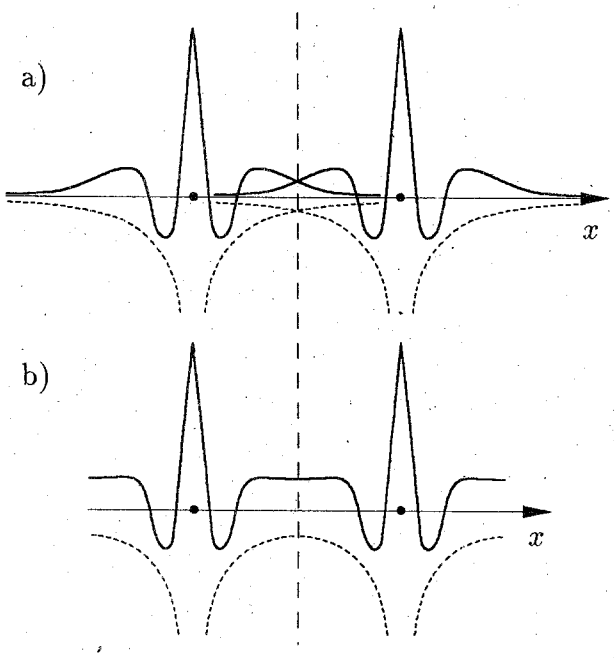


Abb. 3.12: 3s-Wellenfunktion von Natrium. Der Abstand zwischen den Kernen entspricht dem Gitterabstand im Metall. Das Coulomb-Potential der Ionenrümpfe ist gestrichelt angedeutet.
a) Wellenfunktion von isolierten Atomen,
b) überlagerte Wellenfunktionen.

Wir wollen die auftretenden Größen für den Fall des Natriums noch etwas eingehender betrachten. Im freien Atom sind die 3s-Elektronen mit der Energie $E_{3s} = -5{,}14\,\mathrm{eV}$ an die Ionenrümpfe gebunden. Die Absenkung der kinetischen und potentiellen Energie können wir hier mit den uns zur Verfügung stehenden Methoden noch nicht berechnen, doch läßt sich zeigen[9], daß der energetisch tiefste Zustand eines freien Elektrons im Natriummetall die Energie $E_0 = -8{,}2\,\mathrm{eV}$ besitzt. Nun muß noch berücksichtigt werden, daß die Elektronen dem Pauli-Prinzip unterworfen sind. Die daraus folgenden Konsequenzen werden wir ausführlich in Kap. 7.2 diskutieren. Wie man sich leicht vorstellen kann, können nicht alle Elektronen in den tiefsten Zustand mit der Energie E_0 gehen. Wie die Rechnung in Abschnitt 7.2 zeigen wird, führt die Berücksichtigung des Pauli-Prinzips zu einer inneren Energie U_0 der freien Elektronen, die bei Natrium den Wert

[9] Die Berechnung geht auf Wigner und Seitz zurück.

$U_0 = 1{,}95\,\text{eV/Atom}$ annimmt. Somit erwarten wir beim Natrium eine Bindungsenergie von $U_B = E_0 + U_0 - E_{3s} = -1{,}1\,\text{eV/Atom}$. Wie wir der Tabelle 3.3 entnehmen, stimmt dieser Wert exakt mit dem experimentell ermittelten überein. Man sollte jedoch diese gute Übereinstimmung nicht überbewerten, denn bei den meisten anderen Metallen ist sie nicht so überzeugend.

Zur metallischen Bindung können s- und p-Elektronen beitragen. In den *Alkalimetallen* sind neben den Rumpfelektronen nur s-Elektronen vorhanden. Dort ermöglicht das Modell des „freien Elektronengases" (vgl. Kap. 7) eine gute Beschreibung des elektronischen Verhaltens. Reine metallische Bindung ist dadurch gekennzeichnet, daß der Abstand zum nächsten Nachbarn viel größer als der Ionendurchmesser ist. In *Übergangsmetallen* treten neben den delokalisierten s-Elektronen auch stark lokalisierte d-Elektronen auf, die zusätzliche kovalente Bindungen eingehen, den Abstand der Ionen bestimmen und einen wesentlichen Beitrag zur Bindungsenergie liefern, aber kaum zur Leitfähigkeit beitragen. Ionenradius und Ionenabstand werden bei diesen Metallen weitgehend von den d-Elektronen festgelegt.

In Tabelle 3.3 sind hierfür einige Beispiele aufgeführt. Dort ist der Ionendurchmesser gemeinsam mit dem Gitterabstand für je drei Alkalimetalle mit kubisch raumzentriertem und drei Übergangsmetalle mit kubisch flächenzentriertem Gitter wiedergegeben. Darunter sind die Bindungsenergien aufgeführt, die bei den Übergangsmetallen aufgrund des kovalenten Bindungsanteils und der höheren Koordinationzahl wesentlich größer sind.

Tabelle 3.3: Ionendurchmesser, Abstand der nächsten Nachbarn und Bindungsenergie einiger Alkali- und Übergangsmetalle.

	Li	Na	K	Cu	Ag	Au
Ionendurchmesser (Å)	1,20	1,90	2,66	1,92	2,52	2,74
Abstand des nächsten Nachbarn (Å)	3,02	3,66	4,52	2,56	2,90	2,88
Bindungsenergie (eV/Atom)	1,63	1,11	0,93	3,49	2,95	3,81

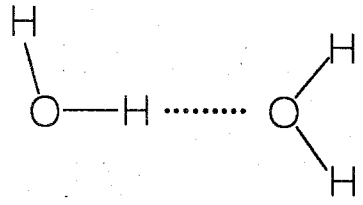
3.6 Wasserstoffbrückenbindung

Wasserstoffatome bilden oft Bindungen aus, die nicht den beschriebenen Bindungstypen zugeordnet werden können. Diese sogenannten *Wasserstoffbrückenbindungen* weisen einige Eigenheiten auf, die für die besonderen Eigenschaften von Wasser und Eis und für die Stabilität der Konfiguration komplexer Proteinmoleküle wesentlich sind. Qualitative Betrachtungen lassen erwarten, daß Wasserstoff ähnliche ionische Bindungen wie die Alkalimetalle eingehen kann, doch ist die Ionisationsenergie von Wasserstoff 13,6 eV im Vergleich zu 5,1 eV beim Natrium. Es ist deshalb schwierig, das Elektron vollständig vom Wasserstoffatom zu trennen. Ionenkristalle mit H^+-Ionen treten daher wegen dieses hohen Ionisationspotentials nicht auf.

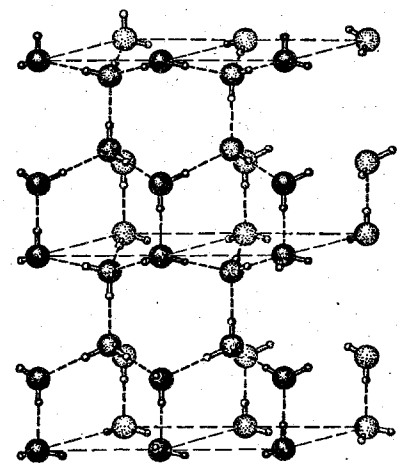
Da Wasserstoffatome nur ein Elektron besitzen, können sie im Gegensatz zu den meisten anderen Atomen nur mit *einem* Nachbarn eine echte *kovalente* Bindung eingehen. Von einer Wasserstoffbrücke spricht man, wenn ein Wasserstoffatom an *zwei* Atome gebunden ist. Dies ist dann möglich, wenn das Wasserstoffatom eine kovalente Bindung mit einem

stark elektronegativen Atom eingeht. Das Bindungselektron geht dann weitgehend, aber nicht vollständig auf den Bindungspartner über, und das fast nackte Proton sitzt auf dem negativ geladenen Ion. Das positive Proton wirkt nun anziehend auf einen weiteren negativ geladenen Partner. Wegen des geringen Durchmessers des Protons von nur 1 fm ist dies aus räumlichen Gründen nur mit *einem* weiteren Partner möglich, so daß eine Brücke zwischen zwei elektronegativen Atomen auftritt. In den meisten Fällen ist die Bindung unsymmetrisch, d.h. sie ist vom Typ A-H $\cdots$ B, wobei das Proton einen größeren Abstand zum Atom B als zum Atom A aufweist. Die typischen Bindungsenergien liegen bei etwa 0,1 eV/Bindung.

Abb. 3.13a zeigt schematisch die Wasserstoffbrückenbindung zwischen zwei Wassermolekülen. Wie oben angedeutet, existieren für das Proton zwischen den Sauerstoffatomen prinzipiell zwei Gleichgewichtslagen. Abhängig vom Ladungszustand des benachbarten Sauerstoffatoms, der wiederum von der Orientierung der Nachbarmoleküle abhängt, wird vom Proton die eine oder andere Lage bevorzugt. Der Übergang zwischen den beiden Lagen erfolgt meist durch thermisch aktivierte Prozesse, doch können in anderen Systemen mit Wasserstoffbrückenbindung auch Tunnelprozesse auftreten. Die Struktur des Eises ist in Abb. 3.13b zu sehen. Die Protonen verknüpfen jeweils zwei Sauerstoffatome, doch ist die Ordnung auch in Einkristallen von Eis durchaus nicht perfekt.



a)



b)

Abb. 3.13: Wasserstoffbrückenbindung in H_2O.
a) Schematische Darstellung der Wasserstoffbrückenbindung. Die Brücke ist durch Punkte angedeutet. Der Abstand zwischen zwei Sauerstoffatomen ist etwa 2,7 Å, die Länge der punktierten Brücke etwa 1,7 Å.
b) Anordnung der Moleküle im Eiskristall. Die Wassermoleküle nehmen im Eiskristall keine einheitliche Orientierung ein. (Aus L. Pauling, *Die Natur der chemischen Bindung*, Verlag Chemie 1964).

4. Defekte

In den bisherigen Betrachtungen standen immer ideale Kristalle oder „ideale“ amorphe Festkörper im Blickpunkt. Tatsächlich ist es jedoch so, daß viele physikalische Eigenschaften von realen Festkörpern durch die immer vorhandenen Defekte geprägt sind. Bereits aus Gründen der Entropie ist eine fehlerlose Anordnung der etwa 10^{22} Atome, aus denen ein makroskopischer Kristall aufgebaut ist, praktisch ausgeschlossen. Darüber hinaus stellt die endliche Ausdehnung, also die Begrenzung des Festkörpers, bereits einen überaus schwerwiegenden Defekt dar. Während in alten Lehrbüchern der Festkörperphysik die Defekte immer relativ ausführlich behandelt wurden, fehlt ihre Beschreibung oft in modernen Büchern. Es werden dort meist nur noch spezielle Defekte im Zusammenhang mit der Dotierung von Halbleitern (vgl. Kap. 9) abgehandelt. Wir wollen hier einen Zwischenweg einschlagen: Wir werden die wichtigsten Defektarten kurz beschreiben und einige ihrer speziellen Eigenschaften herausgreifen. Nicht eingehen werden wir auf die experimentellen Methoden wie Elektronen- und Kernspinresonanz oder die Messung der dielektrischen und optischen Eigenschaften, die bei ihrer Untersuchung angewandt werden. Es würde auch den vorgegebenen Rahmen sprengen, wollten wir hier die interessanten materialwissenschaftlichen Aspekte diskutieren, die für praktische Anwendungen von überragender Bedeutung sind.

Besonders gut ausgearbeitete Konzepte existieren für Kristalle, auf die wir zunächst eingehen. Dort können Defekte mit Abweichungen von der periodischen Anordnung der Gitterbausteine gleichgesetzt werden. Neben den Punktdefekten, deren typische Größe vergleichbar mit atomaren Dimensionen ist, treten in Kristallen auch makroskopisch ausgedehnte Defekte, die sogenannten Linien- und Flächendefekte, auf. Naturgemäß ist in amorphen Festkörpern die Definition von Defekten wesentlich schwieriger. Man findet Punktdefekte mit ähnlichen Eigenschaften wie in Kristallen, aber auch eine neue Defektart, nämlich ungesättigte oder aufgebrochene Bindungen, die in Kristallen normalerweise nicht stabil sind und daher kaum auftreten. Linien- und Flächendefekte in Analogie zu denen in Kristallen scheint es aber nicht zu geben. Am Ende dieses Kapitels wollen wir noch auf Legierungen eingehen, bei denen abhängig von der Temperatur die unterschiedlichen Atome geordnet oder statistisch verteilt eingebaut sind. Derartige Ordnungsphänomene spielen in vielen Bereichen der Festkörperphysik eine wichtige Rolle.

4.1 Punktdefekte

Von erheblicher Bedeutung für die Eigenschaften von Festkörpern sind Punktdefekte, die dadurch entstehen, daß einzelne oder wenige Atome dem Gitter entnommen, hinzugefügt oder ausgetauscht werden. Von „Punktdefekten“ spricht man auch dann, wenn sie einen etwas komplizierteren Aufbau besitzen, vorausgesetzt, die Störung des Gitters ist auf wenige Atomabstände beschränkt. Ausgedehnte Defekte dagegen erstrecken sich meist über Millionen von Gitterabständen. Grundsätzlich unterscheidet man bei Punktdefekten zwischen *Leerstellen*, d.h. fehlenden Atomen, *Zwischengitteratomen*, das sind Atome, die sich

an Orten zwischen den regulären Gitterplätzen aufhalten, und *Fremdatomen*, die reguläre Atome ersetzen bzw. sich zwischen regulären Atomen befinden. Punktdefekte sind immer dann von Bedeutung, wenn der Festkörper relativ wenige große Defekte enthält, d.h. wenn er relativ „perfekt" aufgebaut ist. Sie bestimmen dann u.a. die Diffusion von Atomen und die elektrische Leitfähigkeit von Halbleitern (siehe Kap. 9) sowie Metallen (vgl. Kap. 7) bei tiefen Temperaturen. Wir wollen uns hier mit ihnen anhand einiger Beispiele vertraut machen, da ihr Verhalten aufgrund ihres einfachen Aufbaus besonders leicht zu verstehen ist.

Leerstellen

Ein fehlendes Gitteratom nennt man *Leerstelle* oder *Schottky-Defekt*. Dieses ist der denkbar einfachste Defekt, dessen theoretische Beschreibung keinen allzu großen Aufwand erfordert. Er tritt vor allem in einfach aufgebauten Festkörpern auf und spielt bei Metallen und Ionenkristallen eine große Rolle.

Mit Hilfe der Rastertunnelmikroskopie lassen sich neuerdings auch Leerstellen an Kristalloberflächen sichtbar machen. Allerdings eignen sich hierfür keine gewöhnlichen Metalloberflächen, da dort die Leerstellendichte zu gering ist. In Abb. 4.1 sind Bleiatome abgebildet, die als monoatomare Schicht auf einen Kupfereinkristall aufgedampft wurden. Die Bleiatome ordnen sich auf der Oberfläche regelmäßig an, doch ist die Bindung der Bleiatome untereinander und an die Kupferatome relativ schwach. Wie wir sogleich sehen werden, steigt mit abnehmender Bindungsenergie die Dichte der Leerstellen drastisch an. Es treten dann viele Leerstellen auf, die direkt abgebildet werden können.

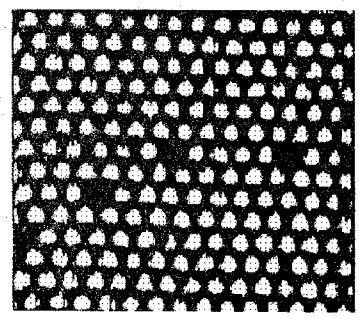


Abb. 4.1: Leerstellen in einer monoatomaren Schicht von Bleiatomen, die auf einen Kupferkristall aufgedampft wurde. Die Abbildung erfolgte mit einem Rastertunnelmikroskop. (Mit freundlicher Genehmigung von M. Schmid und P. Varga; Inst. Allgem. Physik, TU Wien).

Leerstellenkonzentration

Wir wollen nun mit Hilfe einer einfachen thermodynamischen Betrachtung herausfinden, wie groß die Leerstellenkonzentration sein kann und durch welche Parameter sie bestimmt wird. Die Erzeugung oder Vernichtung von Leerstellen läßt sich ähnlich wie eine chemische Reaktion behandeln. Wie bei der chemischen Reaktion ist die entscheidende Größe die freie Enthalpie $G = U - TS + pV = F + pV$, wobei U für die innere Energie, S für die Entropie, p für den Druck und F für die freie Energie steht. Unter der Voraussetzung, daß Druck und Temperatur konstant gehalten werden, gilt für die Änderung $\mathrm{d}G$ der freien Enthalpie: $\mathrm{d}G = \mathrm{d}U - T\mathrm{d}S + p\,\mathrm{d}V$. Solange die Leerstellen das Volumen der Probe kaum verändern, kann der Beitrag von $p\,\mathrm{d}V$ zur freien Enthalpie vernachlässigt werden, und es genügt, sich auf die Betrachtung der freien Energie F zu beschränken. Weiter setzen wir voraus, daß die Konzentration der Leerstellen N_L so klein ist, daß sie nur isoliert auftreten und die Wechselwirkung zwischen ihnen außer acht gelassen werden kann, d.h. es gilt $N_\mathrm{L} \ll N$, wobei N für die Zahl der vorhandenen Atome steht.

Der Beitrag δU_L der Leerstellen zur inneren Energie läßt sich sofort angeben. Ist E_L die Energie, die aufgewendet werden muß, um eine Leerstelle zu erzeugen, so gilt $\delta U_\mathrm{L} = N_\mathrm{L} E_\mathrm{L}$. Nun müssen wir noch berücksichtigen, daß die Leerstellen in zweifacher Hinsicht zur Entropie beitragen. Zunächst bedingt jede Leerstelle eine Reduktion der Gitterschwingungen in ihrer unmittelbaren Umgebung und verändert so die thermische Bewegung des Gitters. Dieser Effekt wird berücksichtigt, wenn man jeder geschaffenen Leerstelle die Entropie S_L zuschreibt[1]. Zusätzlich tritt noch die Konfigurationsentropie auf, die darauf beruht, daß eine Vielzahl von möglichen Anordnungen der N_L Leerstellen auf den $(N + N_\mathrm{L})$ Gitterplätzen existiert. Dieser Beitrag, der proportional zum Logarithmus der Zahl der möglichen Anordnungen ist, hat entscheidenden Einfluß auf die Leerstellenkonzentration. Damit ergibt sich für den Beitrag δF_L der Leerstellen zur freien Energie:

$$\delta F_\mathrm{L} = N_\mathrm{L} E_\mathrm{L} - T\,\delta S = N_\mathrm{L} E_\mathrm{L} - N_\mathrm{L} T S_\mathrm{L} - k_\mathrm{B} T \ln\left[\frac{(N+N_\mathrm{L})!}{N! N_\mathrm{L}!}\right] . \tag{4.1}$$

Im thermodynamischen Gleichgewicht ist die freie Energie minimal, d.h. es stellt sich eine Leerstellendichte ein, für die $\partial F/\partial N_\mathrm{L} = 0$ ist. Unter Ausnutzung der Stirling-Formel $\ln X! \simeq X(\ln X - 1)$ und der oben erwähnten Voraussetzung $N_\mathrm{L} \ll N$ folgt mit (4.1) aus der Gleichgewichtsbedingung die Beziehung $N_\mathrm{L} = N\,\mathrm{e}^{S_\mathrm{L}/k_\mathrm{B}}\,\mathrm{e}^{-E_\mathrm{L}/k_\mathrm{B}T}$. Dividieren wir diese Gleichung durch das Volumen V der Probe, so erhalten wir für die Leerstellendichte $n_\mathrm{L} = N_\mathrm{L}/V$ den Ausdruck

$$n_\mathrm{L} = n\,\mathrm{e}^{S_\mathrm{L}/k_\mathrm{B}}\,\mathrm{e}^{-E_\mathrm{L}/k_\mathrm{B}T} . \tag{4.2}$$

Wir erwarten also, daß die Leerstellenkonzentration im thermischen Gleichgewicht exponentiell mit der Temperatur zunimmt. Die entscheidende Größe ist dabei die Energie E_L, die erforderlich ist, um den Kristall um einen Gitterplatz zu vergrößern. Im Prinzip entsteht eine Leerstelle dadurch, daß ein Atom aus dem Kristallinneren an die Oberfläche gebracht wird. Somit sollte E_L mit der Bindungsenergie U_B vergleichbar sein, die bei vielen Materialien im Bereich von einigen eV/Atom liegt. Wie wir anhand einer Messung an Natrium noch sehen werden, sind für die Leerstellenentropie Werte um $S_\mathrm{L}/k_\mathrm{B} \approx 5$ typisch. Setzt man diesen Wert und $E_\mathrm{L} = 1\,\mathrm{eV}$ (vgl. Tabelle 4.1) ein, so errechnet sich eine Leerstellenkonzentration $n_\mathrm{L}/n \approx 10^{-3}$ bei 1000 K, aber nur $n_\mathrm{L}/n \approx 10^{-15}$ bei Zimmertemperatur.

Wir wollen kurz auf die experimentelle Bestätigung des exponentiellen Anstiegs der Leerstellenkonzentration und dem tatsächlichen Wert von E_L eingehen. Mit der Erzeugung von Leerstellen ist eine Volumenzunahme verbunden. Bei der Auswertung der gemessenen Volumenänderung muß berücksichtigt werden, daß neben der Erzeugung von Leerstellen auch die üblicherweise beobachtete thermische Expansion (vgl. Kap. 6.1) dazu beiträgt. Dieser Beitrag läßt sich dadurch getrennt ermitteln, daß man die Änderung der Gitterabstände mit Hilfe der Röntgenbeugung sehr genau bestimmt. Zieht man von der gemessenen makroskopischen Volumenänderung $\Delta V/V$ den Beitrag $(3\Delta a/a)$ durch die Gitteraufweitung ab[2], so erhält man den Anteil der Leerstellen an der Volumenänderung. Eine meßbare Volumenänderung ist nur bei hoher Leerstellenkonzentration, d.h. in

[1] In vielen Lehrbüchern wird dieser Beitrag zur Entropie bei der Herleitung der Leerstellendichte nicht berücksichtigt, denn er beeinflußt nicht den Temperaturverlauf der Leerstellendichte. Ohne diesen Beitrag läßt jedoch sich keine befriedigende Übereinstimmung zwischen der Vorhersage und dem gemessenen Wert der Leerstellenkonzentration erzielen.

[2] Die einfache Beziehung $\Delta V/V = 3\Delta a/a$ gilt natürlich nur für kubische Kristalle, da deren thermische Expansion nicht von der Richtung abhängt.

der Nähe des Schmelzpunktes zu erwarten. Da die Schmelztemperatur vor allem durch den Wert der Bindungsenergie bestimmt wird, tritt bei niedrig schmelzenden Materialien bereits bei Zimmertemperatur eine meßbare Leerstellenkonzentration auf.

In Abb. 4.2 ist das Ergebnis einer Messung zu sehen, die an Natrium durchgeführt wurde, das bereits bei 371 K schmilzt und daher in einem experimentell gut zugänglichen Temperaturbereich eine hohe Leerstellendichte aufweist. Aus den Meßdaten ergeben sich die Werte $E_L = 0{,}42\,\text{eV}$ für die Bildungsenergie und $S_L/k_B = 5{,}8$ für den Entropiefaktor, während man (aus anderen Messungen) für die Bindungsenergie 1,11 eV/Atom findet. Das Ergebnis zeigt, daß Bildungsenergie und Bindungsenergie nicht gleichgesetzt werden dürfen, wenn auch die Größenordnungen vergleichbar sind. Es muß berücksichtigt werden, daß sich das Gitter in der Nähe der geschaffenen Leerstelle etwas verändert, um sich den neuen Gegebenheiten anzupassen. Diese sogenannte *Relaxation des Gitters* ist mit einem Energiegewinn verbunden und bewirkt so eine merkliche Absenkung der Bildungsenergie E_L gegenüber der Bindungsenergie U_B. Bei Natrium erniedrigt die Gitterrelaxation die Bildungsenergie für Leerstellen gegenüber der Bindungsenergie fast um den Faktor 3. Zum Vergleich von Bindungs- und Bildungsenergie sind in der Tabelle 4.1 die Werte für einige Festkörper aufgeführt.

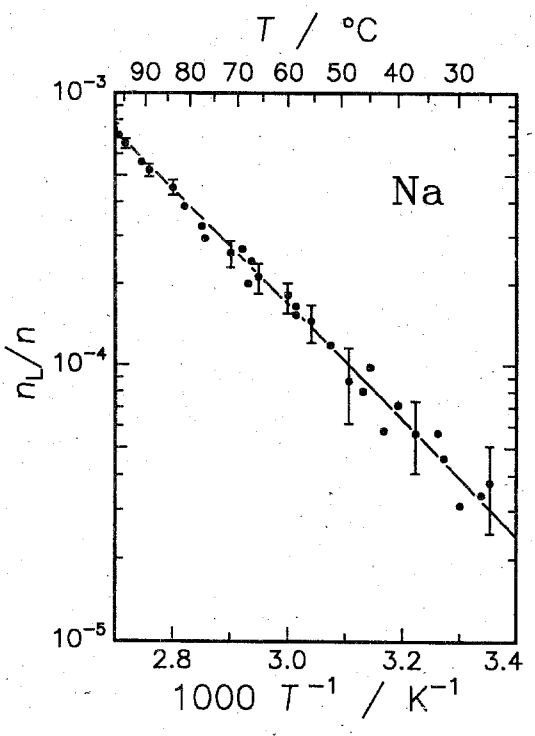


Abb. 4.2: Leerstellenkonzentration n_L/n in Natrium. Durch die Auftragung von $\log(n_L/n)$ als Funktion der reziproken Temperatur wird der exponentielle Zusammenhang unmittelbar sichtbar. (Aus R. Feder und H.P. Charbnau, Phys. Rev. **149** (1966) 464).

Bei Materialien mit höherem Schmelzpunkt ist die Bindungsenergie U_B und damit auch die Energie E_L für die Erzeugung von Leerstellen größer. Nach Gleichung (4.2) nimmt dann die Leerstellenkonzentration bei Zimmertemperaturen wesentlich kleinere Werte als bei Natrium an. Damit wird auch verständlich, warum in Abb. 4.1 beim Nachweis von Leerstellen mit dem Rastertunnelmikroskop nicht einfach die Oberfläche des Kupferkristalls abgebildet, sondern erst noch eine Bleischicht aufgedampfte wurde. Die schwache Bindung der Bleiatome bewirkt eine hohe, mit dem Mikroskop beobachtbare Leerstellenkonzentration.

Tabelle 4.1: Bindungsenergie U_B, Bildungsenergie E_L für Leerstellen und Aktivierungsenergie E_D für Volumendiffusion. Um die Vergleichbarkeit der Zahlenwerte zu gewährleisten, sind bei den Ionenkristallen für U_B und E_L die halben Werte angegeben, die man für Ionen*paare* erhält. Weiterhin beziehen sich bei den Ionenkristallen die Werte von E_D auf positive Ionen.

	Al	Cu	Zn	Au	LiF	LiCl	NaCl	KCl
U_B (eV)	3,39	3,49	1,35	3,81	5,36	4,38	4,07	3,68
E_L (eV)	0,67	1,28	0,42	0,97	1,34	1,06	1,01	1,15
E_D (eV)	0,62	0,71	0,40	0,74	0,65	0,41	0,86	0,89

Eine Besonderheit, die hier kurz erwähnt werden soll, findet man bei Ionenkristallen. Da in diesen Materialien die Leerstellen gegenüber der Umgebung geladen sind, tritt zwischen gleichartigen Leerstellen eine hohe Coulomb-Abstoßung auf, die zu einer Erhöhung der freien Energie führt. Durch lokale Ladungsneutralität wird dieser zusätzliche Energieaufwand minimiert, so daß sich immer Leerstellen*paare* mit entgegengesetzter Ladung bilden. Wir wollen kurz betrachten, welche Auswirkungen dieses Phänomen auf die Leerstellendichte hat. Hierzu berechnen wir die Leerstellendichten n_L^+ und n_L^- der beiden Ionensorten mit den Dichten n^+ und n^- zunächst separat mit Hilfe von (4.2). Dann multiplizieren wir die beiden Gleichungen miteinander und erhalten für die Defektpaare den Ausdruck

$$n_L^+ n_L^- = n^+ n^- \, e^{(S_L^+ + S_L^-)/k_B} \, e^{-(E_L^+ + E_L^-)/k_B T} \,, \tag{4.3}$$

wobei die Bildungsenergien E_L^+ und E_L^-, und die dazu gehörigen Entropiefaktoren S_L^+ und S_L^- durchaus verschiedene Werte besitzen können. Nun bringen wir die Forderung nach Ladungsneutralität ein und setzen $n_L^+ = n_L^-$. Dadurch erhalten wir

$$n_L^+ = n_L^- = \sqrt{n^+ n^-} \, e^{S_P/2k_B} e^{-E_P/2k_B T} \,, \tag{4.4}$$

wobei $S_P = S_L^+ + S_L^-$ für die Entropie und $E_P = E_L^+ + E_L^-$ für die Bildungsenergie eines Defektpaares stehen. Die Forderung nach Ladungsneutralität bewirkt, daß die Wahrscheinlichkeit für die Bildung von Defekten durch die Summe der individuellen Bildungsenergien bestimmt wird. Dies führt häufig zu Defektkonzentrationen, die sich wesentlich von denen unterscheiden, die bei unabhängiger Defektbildung aufträten. Die Zahlenwerte für die Bildungsenergien E_L von Ionenkristallen, die in der Tabelle 4.1 angegeben sind, sind daher nicht die individuellen, sondern die halben Bildungsenergien pro Ionenpaar, also $E_P/2$.

Leerstellendiffusion

Aufgrund von Gleichung (4.2) sollten höherschmelzende Metalle wie Eisen mit der Bindungsenergie $U_B = 4{,}3\,\text{eV/Atom}$ bei Zimmertemperatur eine verschwindend kleine Leerstellenkonzentration aufweisen. Tatsächlich ist in den meisten Fällen dennoch eine merkliche Konzentration vorhanden. Dies liegt daran, daß zunächst bei der Herstellung des Festkörpers am Schmelzpunkt viele Leerstellen gebildet werden. Während des Abkühlens vermindert sich ihre Zahl in erster Linie dadurch, daß sie an die Oberfläche[3]

[3] Wie wir in Kap. 4.2 sehen werden, existiert noch ein weiterer Mechanismus: Leerstellen können auch im Kristallinneren, an sogenannten Stufenversetzungen, verschwinden, doch müssen sie in diesem Fall erst zu den Versetzungen diffundieren.

wandern. Da der Weg dorthin von den Leerstellen durch *Diffusion* zurückgelegt werden muß, hängt es von der verfügbaren Zeit ab, ob sie die Oberfläche erreichen oder im Innern der Probe gefangen werden.

Ein Elementarschritt dieses Diffusionsprozesses ist in Abb. 4.3a angedeutet: Atome auf regulären Gitterplätzen können in benachbarte Leerstellen springen, wobei sich die Lage der Leerstellen in entgegengesetzter Richtung um einen Gitterplatz verschiebt. Da die Sprünge statistisch erfolgen, gewinnt man von außen den Eindruck, daß die Leerstellen im Gitter umherwandern. Mit abnehmender Temperatur werden die Sprungraten immer kleiner und bei Temperaturen wesentlich unter dem Schmelzpunkt sind die von den Leerstellen zurückgelegten Strecken so klein, daß sie die Oberfläche nicht mehr erreichen und sich ihre Konzentration auf experimentell zugänglichen Zeitskalen nicht mehr ändern kann. Um dieses Verhalten quantitativ zu erfassen, wollen wir kurz auf die Leerstellendiffusion eingehen, da der zugrundeliegende Sprungmechanismus auch in vielen anderen Bereichen der Physik eine wichtige Rolle spielt.

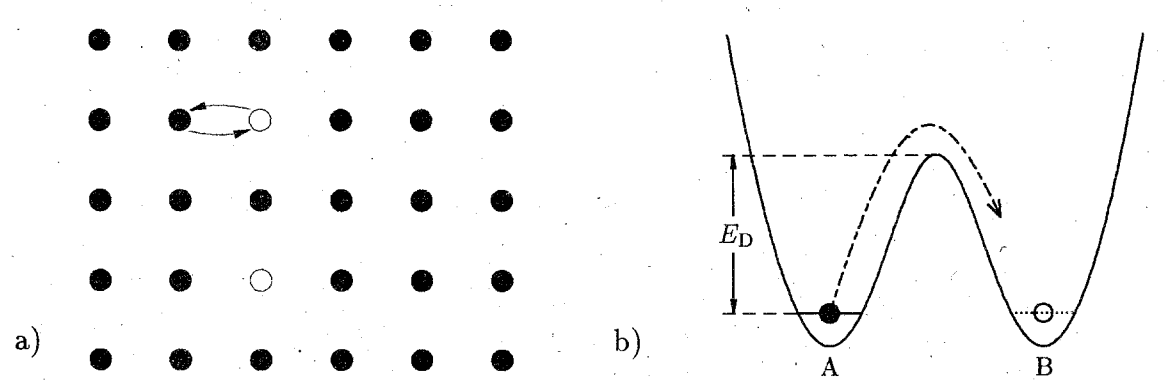


Abb. 4.3: Schematische Darstellung der Leerstellenwanderung.
a) Der Sprung eines Atoms bewirkt eine Verschiebung der Leerstelle in entgegengesetzter Richtung. Da sich dieser Vorgang statistisch wiederholt, wandert die Leerstelle im Gitter regellos umher.
b) Schematischer Potentialverlauf beim Sprung eines Atoms vom Gitterplatz A in die Leerstelle B. Die Größe der Aktivierungsenergie E_D ergibt sich aus der Differenz zwischen dem Potentialmaximum und der Nullpunktsenergie des Atoms.

In Abb. 4.3b ist der Verlauf des Potentials angedeutet, in dem sich ein Atom bewegt, wenn der benachbarte Gitterplatz unbesetzt ist. Für das Atom existiert dann ein weiterer, energetisch äquivalenter Platz, von dem es durch eine Potentialbarriere getrennt ist. Das Atom kann auf den leeren Platz gelangen, wenn thermische Energie zur Verfügung steht, die ein Überspringen des Potentialbergs erlaubt. Die Wahrscheinlichkeit, daß dieser Fall tatsächlich eintritt, ist durch den Boltzmann-Faktor $\exp(-E_D/k_BT)$ gegeben, der neben der Temperatur als wichtige Größe die *Aktivierungsenergie* E_D enthält. Schwingt das Atom mit der „Versuchsfrequenz" ν_0 in seiner Potentialmulde, so ist die Sprungfrequenz ν, also die Anzahl der pro Zeiteinheit „erfolgreichen" Versuche, durch die Gleichung

$$\nu = \nu_0 \, e^{-E_D/k_BT} \tag{4.5}$$

gegeben. In guter Näherung kann die Versuchsfrequenz der Schwingungsfrequenz der Atome im Kristallpotential gleichgesetzt werden und ist daher in den meisten Fällen von der Größenordnung $\nu_0 \approx 10^{13}\,s^{-1}$ (vgl. Kap. 5). Vorgänge, bei denen eine gewisse Potentialschwelle mit Hilfe thermischer Energie überwunden wird, bezeichnet man als *thermisch aktiviert.*

Entsprechend dem *Fickschen Gesetz* ist der Diffusionsstrom $\boldsymbol{j}$, d.h. die Teilchenstromdichte, durch die *Diffusionskonstante* D und den Gradienten der Leerstellendichte n_L gegeben:

$$\boldsymbol{j} = -D \operatorname{grad} n_\mathrm{L} \,. \tag{4.6}$$

Die Diffusionskonstante[4] wird durch die Häufigkeit und die Weite der Sprünge bestimmt. In dem von uns betrachteten Fall ist die Sprungweite durch den Gitterabstand a und die Sprungrate durch (4.5) gegeben, so daß wir

$$D \approx a^2 \nu = a^2 \nu_0 \, \mathrm{e}^{-E_\mathrm{D}/k_\mathrm{B}T} = D_0 \, \mathrm{e}^{-E_\mathrm{D}/k_\mathrm{B}T} \tag{4.7}$$

schreiben können. Der Wert der Aktivierungsenergie E_D ist meist vergleichbar mit der Energie E_L, die zur Bildung von Leerstellen in der betrachteten Substanz erforderlich ist, und liegt daher ebenfalls bei Energien um 1 eV. Setzen wir diesen Zahlenwert in (4.7) ein, so finden wir für die Diffusionskonstante bei Zimmertemperatur einen Wert um 10^{-18} cm^2/s. Da der mittlere Diffusionsweg $\overline{x}$ bei dreidimensionalen Proben durch die Gleichung $\overline{x} = \sqrt{6Dt}$ gegeben ist, wird verständlich, warum Leerstellen bei Temperaturen wesentlich unter dem Schmelzpunkt im Probeninneren eingefroren werden.

Ladungstransport in Ionenkristallen

Die Beschreibung der *elektrischen Leitfähigkeit* von Ionenkristallen ist ein eindrucksvolles Anwendungsbeispiel für das soeben diskutierte Konzept. In diesen Materialien erfolgt der elektrische Ladungstransport nicht durch Elektronen, sondern häufig durch die Wanderung von Leerstellen, deren Konzentration und Beweglichkeit mit zunehmender Temperatur drastisch anwächst. Natürlich bewegen sich beim Ladungstransport die verschieden geladenen Ionen nicht paarweise, sondern unabhängig voneinander.

Die elektrische Leitfähigkeit σ ist durch die allgemeine Beziehung

$$\sigma_\mathrm{el} = n_q \, q \, \mu \tag{4.8}$$

definiert, wobei n_q für die Ladungsträgerdichte, q für deren Ladung und μ für ihre Beweglichkeit[5] steht. Bei diffundierenden Ionen sind Beweglichkeit und Diffusionskoeffizient über die *Einstein-Beziehung* $\mu k_\mathrm{B} T = qD$ miteinander verknüpft. Zur Vereinfachung nehmen wir an, daß sich die Aktivierungsenergie E_D der Ionensorten nicht unterscheidet. In diesem einfachen Fall besitzen alle Ionen nach (4.7) die gleiche Beweglichkeit. Setzen wir die Einstein-Beziehung in (4.8) ein und ersetzen n_q und D durch die Ausdrücke (4.4) und (4.7), so erhalten wir

$$\sigma_\mathrm{el} = \frac{n q^2 D_0}{k_\mathrm{B} T} \, \mathrm{e}^{S_\mathrm{P}/2k_\mathrm{B}} \, \mathrm{e}^{-E_\mathrm{P}/2k_\mathrm{B}T} \, \mathrm{e}^{-E_\mathrm{D}/k_\mathrm{B}T} \,, \tag{4.9}$$

wobei wir die Abkürzung $n = n^+ + n^-$ für die Ionendichte benutzt haben. Die Exponentialfunktionen bewirken eine außerordentlich starke Temperaturabhängigkeit der Leitfähigkeit. Ionenkristalle sind daher bei Zimmertemperatur gute Isolatoren, bei hohen Temperaturen aber gute Leiter. Der Übergang vom Isolator zum Leiter erfolgt in einem relativ schmalen Temperaturintervall.

[4] In Kristallen hängt die Diffusionsgeschwindigkeit von der Richtung ab, d.h. die Diffusionskonstante ist ein Tensor. In unser einfachen Betrachtung behandeln wir sie als skalare Größe.

[5] Der Begriff „Beweglichkeit" wird ausführlicher in Kap. 8.2 diskutiert, wenn die elektrische Leitfähigkeit der Metalle behandelt wird.

Besonders deutlich wird das Mitwirken der Leerstellen beim Ladungstransport im Falle der *Ladungskompensation*, die in Ionenkristallen auftritt, wenn Fremdionen mit höherer Ladung als die der Wirtsionen eingebaut werden. Dotiert man beispielsweise NaCl mit $CaCl_2$, so beobachtet man bereits bei relativ niedrigen Temperaturen eine gut meßbare Leitfähigkeit, während sie in reinem NaCl in diesem Temperaturbereich noch verschwindend gering ist. Darüberhinaus findet man, daß sich beim Einbau der schwereren Kalziumatome überraschenderweise die Kristalldichte erniedrigt. Beide Beobachtungen werden verständlich, wenn man bedenkt, daß für jedes eingebaute Ca^{2+}-Ion im Kristall eine zusätzliche Na^+-Leerstelle geschaffen werden muß, um die Ladungsneutralität zu wahren. In Abb. 4.4 ist dieser Effekt schematisch dargestellt. Die tatsächlich vorhandene Leerstellendichte wird nicht mehr durch die Temperatur, sondern durch die Dichte n_{Ca} der Kalziumionen bestimmt. Durch die Dotierung ist also bereits bei Zimmertemperatur eine große Zahl von Leerstellen vorhanden, die für den Stromtransport zur Verfügung steht. Obwohl der Strom proportional zu n_{Ca} ist, wird die Ladung durch die Natriumionen transportiert[6], die an der Kathode abgeschieden werden. Die Kalziumionen selbst tragen dazu nicht direkt bei. Das Anwachsen der Leitfähigkeit mit der Temperatur wird durch die Aktivierungsenergie E_D für die Diffusion der Natriumionen bestimmt, d.h. es gilt $\sigma_{el} \propto n_{Ca} \exp(-E_D^{Na}/k_B T)$.

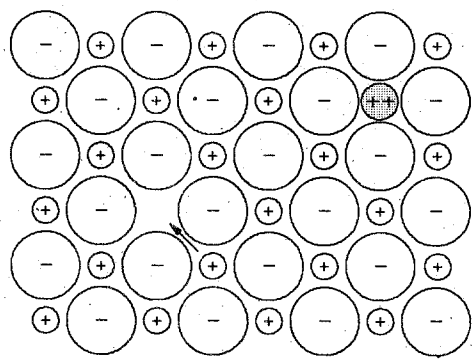


Abb. 4.4: Ca^{2+}-Ion im NaCl-Gitter. Das eingebaute Kalziumion bewirkt das Auftreten einer Leerstelle, wodurch die Bewegung von Na^+-Ionen ermöglicht wird. Ein möglicher derartiger Sprung ist durch den Pfeil angedeutet.

Farbzentren

In der Entwicklung der Festkörperphysik nehmen die Defekte in Alkalihalogenidkristallen eine herausragende Stellung ein. Das Studium der sogenannten *Farbzentren* war der Ausgangspunkt für die Untersuchung von Punktdefekten in Festkörpern. Reine Alkalihalogenidkristalle sind im sichtbaren Spektralbereich transparent. Durch Verunreinigung oder durch Bestrahlen mit Röntgen-, Gamma- oder energiereicher Teilchenstrahlung verfärben sich die Kristalle. Als Beispiel ist in Abb. 4.5 das Absorptionsspektrum eines bestrahlten KCl-Kristalls gezeigt. Es treten Absorptionsbanden auf, die unterschiedlichen Typen von Punktdefekten zugeschrieben werden. Das einfach aufgebaute *F-Zentrum*, das die stärkste Absorptionsbande verursacht, wollen wir hier kurz diskutieren. Daneben existiert eine Reihe komplizierterer Zentren, an denen mehrere Leerstellen oder Ionen beteiligt sind. Für sie werden die Abkürzungen F_A, M, N, R, V_K, ... verwandt. Entsprechende Defekte mit ähnlichen Absorptionsspektren findet man auch bei den anderen Alkalihalogenidkristallen, wenngleich dort die Banden bei anderen Wellenlängen liegen.

[6] Man kann sich auch eine andere Argumentation zu eigen machen: Da die Leerstellen bezogen auf einen regulär besetzten Gitterplatz negativ geladen sind, werden sie vom elektrischen Feld durch die Probe getrieben. In Kap. 8 und 9 werden wir diese Argumentation auf Elektronen anwenden und ausführlich diskutieren.

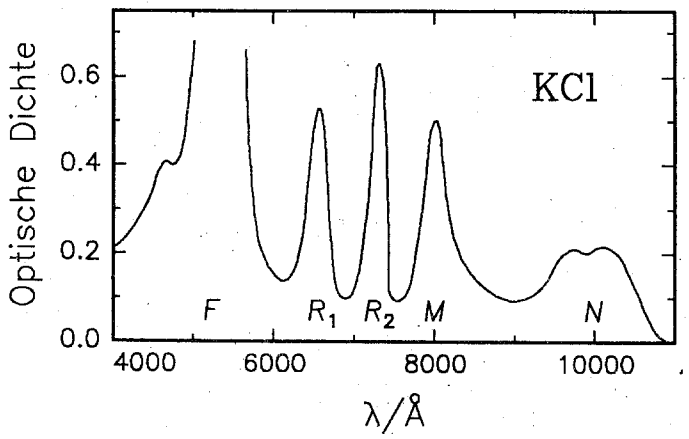


Abb. 4.5: Absorptionsspektrum von KCl nach Bestrahlung mit Röntgenlicht. Die verschiedenen Banden werden unterschiedlichen Defektarten zugeordnet und tragen die Bezeichnungen F, R_1, R_2, M und N. (Aus R.H. Silsbee, Phys. Rev. **138** (1969) A 180).

Besonders übersichtliche Verhältnisse treten auf, wenn man F-Zentren dadurch erzeugt, daß man einen NaCl-Kristall in Natriumdampf erwärmt und anschließend abschreckt. Der ursprünglich klare Kristall zeigt dann eine gelb-braune Färbung. Bei dieser Vorgehensweise lagern sich Natriumatome aus dem Dampf auf der Oberfläche an und verändern die Stöchiometrie des Festkörpers. Der behandelte NaCl-Kristall enthält überschüssige Natriumatome, die ihr Valenzelektron abgeben wollen. Die Struktur wird dadurch stabilisiert, daß Chlorleerstellen gebildet werden, die das Valenzelektron der Natriumatome aufnehmen und in das Kristallinnere diffundieren. Da das magnetische Moment dieser Elektronen nicht durch das entgegengerichtete Moment eines weiteren Elektrons kompensiert wird, läßt es sich leicht mit Elektronenspinresonanz-Messungen nachweisen. Während man bei defektfreien Kristallen kein ESR-Signal findet, machen sich die F-Zentren bei derartigen Messungen deutlich bemerkbar. Aufgrund der Unschärferelation können die Elektronen mit ihrer kleinen Masse nicht auf den kleinen Raum lokalisiert sein, den die Chlorionen einnehmen. Sie „verteilen" sich daher vor allem über die benachbarten positiven Ionen. Eine schematische Darstellung der Ladungsverteilung des F-Zentrums zeigt Abb. 4.6.

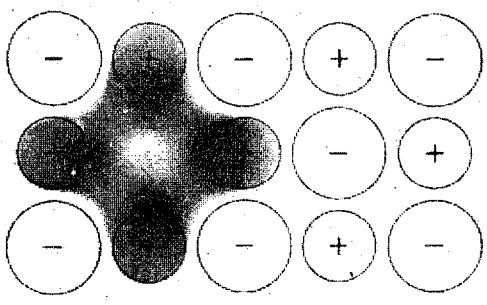


Abb. 4.6: Bildhafte Darstellung der Ladungsverschmierung des Elektrons eines F-Zentrums in einem Alkalihalogenidkristall.

Anschaulich gesprochen bewegt sich das „überschüssige" Elektron um die Leerstelle, die aufgrund der lokalen Ladungskonfiguration positiv erscheint. Hierbei ist aber zu beachten, daß das effektive elektrische Feld, das durch die Nachbarionen hervorgerufen wird, keinesfalls dem einer Punktladung entspricht, sondern kubische Symmetrie besitzt. Wie die Elektronen gewöhnlicher Atome können auch die Elektronen der F-Zentren unterschiedliche Eigenzustände einnehmen. Durch Lichteinstrahlung lassen sich Übergänge zwischen den Zuständen anregen, wobei die auftretende Absorptionsbande charakteristisch für den betreffenden Alkalihalogenidkristall ist. Es tritt aber keine enge Linie, sondern eine relativ

breite Bande auf, da die Elektronenbewegung stark an die Schwingungen der benachbarten Ionen ankoppelt. Dadurch ähnelt das Spektrum dem von Molekülen, bei denen die breiten Absorptionsbanden ebenfalls durch die Kopplung der elektronischen Übergänge an Schwingungszustände zustande kommen. Da die Absorption meist im Bereich des sichtbaren Spektrums erfolgt, kommt es zur Färbung der Kristalle. Wie bereits erwähnt, werden weitere Absorptionsbanden von komplexer aufgebauten Defekten verursacht. So besteht das M-Zentrum aus zwei unmittelbar benachbarten Chlorleerstellen, in denen sich jeweils ein Elektron zur Ladungskompensation befindet. Das R-Zentrum setzt sich sogar aus drei F-Zentren zusammen, die direkt nebeneinander liegen.

Wie erwähnt und in Abb. 4.7a für eine Reihe von Alkalichloriden dargestellt, tritt die Absorptionsbande bei unterschiedlichen Wellenlängen auf. Offensichtlich verschiebt sich die Bande mit zunehmender Masse des Kations zu größeren Wellenlängen. Auf den ersten Blick scheint diese Beobachtung anzudeuten, daß entgegen unserer einfachen Vorstellung vom Aufbau der F-Zentren die Kationen eine wichtige Rolle spielen. Ein Blick auf Abb. 4.7b macht jedoch sofort klar, daß nicht die Masse, sondern die Größe der Ionen und damit die Gitterkonstante ausschlaggebend ist. Die Wellenlänge $\lambda_{\max}$, bei der das Maximum auftritt, hängt über die einfache Beziehung $\lambda_{\max} = AR_0^2$ mit dem Abstand R_0 der nächsten Nachbarn zusammen. Bei Alkalichloriden mit NaCl-Struktur besitzt die Proportionalitätskonstante den Wert $A = 6 \cdot 10^{12}\,\mathrm{m}^{-1}$. Dieser Zusammenhang mit R_0 ist unmittelbar einsichtig: Das Elektron des F-Zentrums bewegt sich in grober Näherung in einem Kastenpotential, das durch die Nachbarn gebildet wird und dessen Abmessung daher mit dem Abstand der nächsten Nachbarn ansteigt. Mit elementare Quantenmechanik findet man, daß der Abstand der Energieniveaus eines Teilchens in einem Potentialkastens umgekehrt proportional zum Quadrat der Kastengröße anwächst. Nur am Rand sei vermerkt, daß sich mit dieser einfachen Vorstellung sogar die Lage der Banden relativ gut vorhersagen läßt.

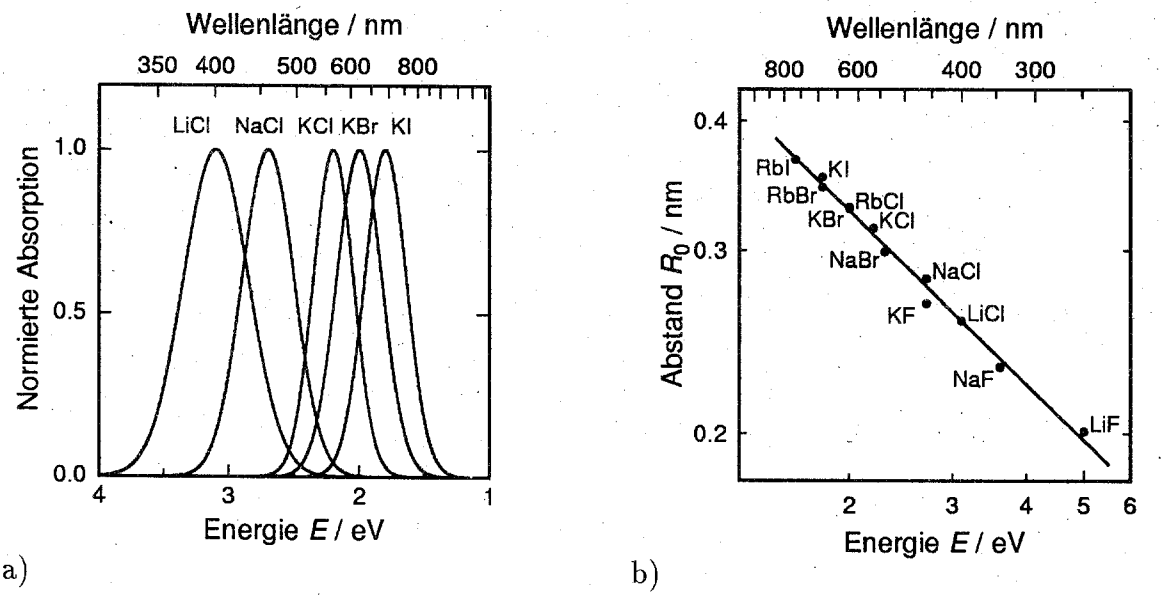


Abb. 4.7: Optische Absorption durch F-Zentren in Alkalihalogeniden.
a) Absorptionsbanden der F-Zentren in verschiedenen Alkalihalogeniden,
b) Photonenenergie bzw. Wellenlänge am Maximum der Absorptionsbanden in Abhängigkeit vom Abstand R_0 der nächsten Nachbarn im logarithmischen Maßstab.
(Nach G. Miessner und H. Pick, Z. f. Physik **134** (1953) 604).

Ein weiterer interessanter Aspekt ist die starke Verschiebung der Lumineszenzbande gegenüber der Anregungswellenlänge. Während beispielsweise die F-Zentren von KCl bei etwa 560 nm absorbieren, beobachtet man ihre Lumineszenzstrahlung bei etwa 1000 nm. Der Grund hierfür ist die Umordnung des umgebenden Gitters während der Zeit, in der das F-Zentrum angeregt ist. Die Anregung geschieht aus einem F-Zentrum, bei dem sich das Gitter an die Elektronkonfiguration angepaßt hat, d.h. aus einem energetisch relativ tief liegenden Grundzustand. Die Rekombination des Elektrons erfolgt dagegen in eine nicht relaxierte Leerstelle mit höherer Grundzustandsenergie. Die Energie des emittierten Photons ist daher kleiner als die des absorbierten.

In den letzten Jahren haben Kristalle mit Farbzentren auch technische Bedeutung erlangt. Sie werden in durchstimmbaren Lasern, die im nahen Infrarot arbeiten, als aktives Medium verwendet. Hierbei nutzt man die soeben erwähnte Verschiebung der Lumineszenz- gegenüber der Absorptionsbande aus und erreicht auf einfachem Weg eine Besetzungsinversion des emittierenden Niveaus. Allerdings werden nicht die einfachen F-Zentren ausgenutzt, sondern etwas komplizierter aufgebaute Farbzentren.

Zwischengitteratome

Neben den Leerstellen findet man noch eine weitere Klasse von sehr einfachen Punktdefekten, nämlich *Zwischengitteratome*. In diesem Fall nehmen Atome einen Platz *zwischen* den regulären Gitteratomen ein. Das Einschieben eines zusätzlichen Atoms bewirkt natürlich in der Umgebung eine sehr starke Verzerrung des Gitters. Wie in Abb. 4.8 bildhaft gezeigt, paßt sich das Gitter dieser neuen Gegebenheit durch eine lokale Verschiebung der benachbarten Atome weitgehend an, d.h. es tritt auch hier eine strukturelle Relaxation auf. Aus Platzgründen tritt dieser Defekt vor allem in Materialien mit offener Struktur auf, d.h. in Festkörpern mit kleinem Packungsverhältnis. Ein wesentlicher Teil der Bildungsenergie E_Z wird für die Verzerrung des umliegenden Gitters aufgewandt, weshalb E_Z meist größer als E_L, die Bildungsenergie für Leerstellen, ist. Aus diesem Grund werden beim Erhitzen von Metallen praktisch keine Zwischengitteratome, sondern Leerstellen erzeugt. In Ionenkristallen dagegen sind die beiden Bildungsenergien vergleichbar. So findet man in Alkalihalogeniden überwiegend Schottky-Defekte, in Silberhalogeniden dagegen Zwischengitteratome.

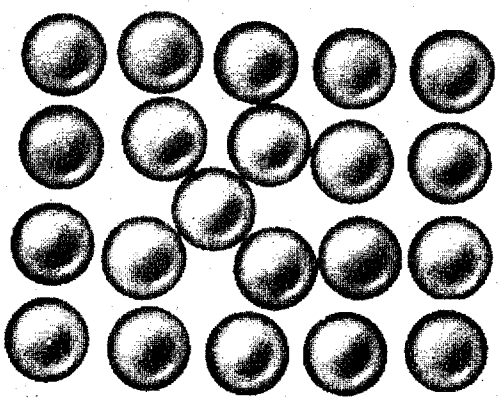


Abb. 4.8: Schematische Darstellung eines Zwischengitteratoms. Die Umgebung hat sich dem eingeschobenen Atom angepaßt, d.h. die Relaxation des Gitters ist bereits erfolgt.

In den meisten Fällen werden Zwischengitteratome dadurch geschaffen, daß Atome auf regulären Gitterplätzen in benachbarte Zwischenräume des Gitters springen bzw. geschoben werden. Wenn Leerstellen und Zwischengitteratome eng benachbart sind, so spricht man von *Frenkel-Defekten*. Diese Art von Defekten läßt die Stöchiometrie der Probe

unangetastet. Da ihre Erzeugung auch lokale Ladungsneutralität garantiert, treten sie in Ionenkristallen häufig anstelle von einfachen Zwischengitteratomen auf. Ihre Gleichgewichtskonzentration läßt sich genauso wie bei den Schottky-Defektpaaren in Ionenkristallen herleiten, und man erhält ein ähnliches Ergebnis.

Sind Zwischengitteratome erst erzeugt, so bewegen sie sich meist mit Leichtigkeit durch den Festkörper und bestimmen so das Diffusionsverhalten, d.h. die Aktivierungsenergie für Diffusionsbewegungen ist relativ klein. So erfolgt beispielsweise der Stromtransport in Silberhalogeniden über Zwischengitterplätze und nicht, wie vorher geschildert, über Leerstellenwanderung.

Wenn auch im thermischen Gleichgewicht in Metallen kaum Zwischengitteratome auftreten, so lassen sie sich doch in großer Zahl durch den Beschuß mit energiereichen Teilchen erzeugen, da bei deren Abbremsung Atome aus ihren Gitterplätzen geschlagen werden. Man spricht dann von *Strahlenschäden*, die in der Reaktortechnik eine wichtige Rolle spielen. In Metallen teilen sich in diesem Fall Zwischengitteratom und ein benachbartes „reguläres" Atom häufig den verfügbaren Platz und bilden eine *Hantel-Konfiguration*. Diese Atomanordnung ist im Gitter besonders beweglich, denn bereits kleine lokale Umlagerungen der umgebenden Atome führen zu einer scheinbaren Verschiebung der Hantel auf einen benachbarten Gitterplatz, was letztendlich der Wanderung eines Zwischengitteratoms gleichkommt. Die typischen Aktivierungsenergien E_D für eine derartige Diffusion sind mit 0,1 eV im Vergleich zur Bildungsenergie außerordentlich klein.

Fremdatome

In realen Festkörpereigenschaften spielen die unvermeidbaren Verunreinigungen eine wichtige Rolle. Sie können abhängig von der Temperatur, ihrer Größe und ihren chemischen Eigenschaften entweder einen regulären Gitterplatz (*substitutionelles Fremdatom*) oder aber auch einen Zwischengitterplatz (*interstitielles Fremdatom*) einnehmen. Substitutionelle Verunreinigungen werden gezielt in Halbleiter eingebracht und spielen eine zentrale Rolle bei deren technischer Anwendung. Wir werden auf dieses Thema noch ausführlich in Kap. 9.2 zurückkommen, wenn wir die elektrischen Eigenschaften dieser Materialklasse diskutieren. Hier beschäftigen wir uns mit ihrem Einbau und ihren ebenfalls bemerkenswerten optischen Eigenschaften. Wir greifen als einfache Beispiele die vierwertigen Elemente Silizium und Germanium heraus. Wie in Kap. 3.4 diskutiert, weisen die Atome in diesen Kristallen eine sp^3-Hybridbindung auf. Bringen wir nun fünfwertige P-, As- oder Sb-Atome ein, so erzwingen die Siliziumnachbarn eine vierfache Koordination, obwohl bei den eingebrachten Atomen eine s^2p^3-Konfiguration vorliegt. Die Fremdatome besitzen ein Elektron mehr als für die tetraedrische Bindung erforderlich ist. Dieses Elektron kann nicht an einer Bindung teilnehmen, bleibt jedoch bei tiefer Temperatur aufgrund der Coulomb-Wechselwirkung mit dem Kern am positiven Rumpf gebunden. In Abb. 4.9a ist der Einbau eines Phosphoratoms bildhaft dargestellt. Wie wir weiter unten sehen werden, ist die Bindungsenergie des überzähligen Elektrons so gering, daß es sich bereits bei Zimmertemperatur vom Ionenrumpf entfernen und zur elektrischen Leitfähigkeit beitragen kann. Fünfwertige Fremdatome in Kristallen mit vierfacher Bindung bezeichnet man daher als *Donatoren*, da sie bei Zimmertemperatur Elektronen abgeben.

Auf den ersten Blick verhalten sich dreiwertige Fremdatome wie B, Al, Ga oder In ganz anders. Zwar werden auch die dreiwertigen Fremdatome auf regulären Gitterplätzen eingebaut, doch fehlt ihnen ein Elektron zur Vervollständigung der vierten Bindung. Dennoch bleibt die kubische Symmetrie erhalten: das fehlende Elektron wird „abwechselnd"

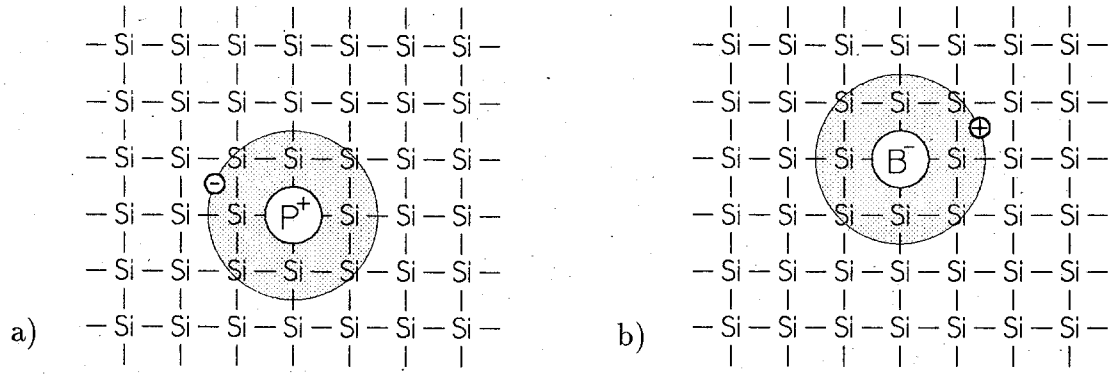


Abb. 4.9: Fremdatome in Silizium.
a) Donator: Das ungebundene Elektron umkreist, bildlich gesprochen, den positiv geladenen Rumpf des Phosphoratoms.
b) Akzeptor: Der negativ geladene Rumpf des Boratoms ist scheinbar von einer positiven Ladung umgeben.

einem der vier Siliziumnachbarn „entzogen", und das eingebaute Atom lädt sich auf. Nach außen sieht es so aus, als ob eine positive Ladung um den negativ geladenen Rumpf des Fremdatoms kreisen würde. Eine bildhafte Darstellung dieser Konfiguration ist in Abb. 4.9b für den Fall eines Boratoms zu sehen. Bei Zimmertemperatur wird die positive Ladung an das Gitter abgegeben oder, anders ausgedrückt, ein Elektron auf Dauer aufgenommen. Dreiwertige Fremdatome wirken deshalb als *Akzeptoren.* Eine tiefergehende Betrachtung, auf die wir in Kap. 9 eingehen, führt zu bemerkenswerten Ähnlichkeiten zwischen drei- und fünfwertigen Fremdatomen.

Nun stellt sich die Frage, welche neuen Eigenschaften das nicht direkt gebundene Elektron der Donatoratome besitzt. Den sehr wichtigen Einfluß der Fremdatome auf die elektrischen Eigenschaften werden wir in Kap. 9 bei der Behandlung der Halbleiter näher erläutern. Auf die optische Absorption der Donatoren gehen wir hier kurz ein. In einfachster Näherung läßt sich die Bewegung eines Elektrons um den positiv geladenen Rumpf eines Fremdatoms mit Hilfe des sogenannten „Wasserstoffatommodells" beschreiben. Die Energieeigenwerte E_ν sind in dieser Näherung durch den bekannten Ausdruck

$$E_\nu = -\frac{1}{2}\frac{m^* e^4}{(4\pi\epsilon_r\epsilon_0)^2\hbar^2}\cdot\frac{1}{\nu^2} \tag{4.10}$$

gegeben, wobei ν für die Hauptquantenzahl steht[7]. Diese Gleichung unterscheidet sich in zwei Punkten vom Ausdruck für das Wasserstoffatom. Im Zähler steht anstelle der Masse m des freien Elektrons die sogenannte *effektive Masse* m^*, auf deren Bedeutung wir aber erst in Kap. 8 eingehen können. Es sei nur vermerkt, daß bei Silizium $m^* \approx 0{,}3\,m$ gilt. Desweiteren wird durch die Dielektrizitätskonstante ϵ_r im Nenner die Abschirmung des Coulomb-Potentials (vgl. Abschnitt 7.2) durch die benachbarten Halbleiteratome berücksichtigt. Für Silizium und Germanium gilt $\epsilon_{r,Si} = 11{,}7$ und $\epsilon_{r,Ge} = 15{,}8$. Dies bedeutet, daß verglichen mit dem freien Wasserstoffatom die Energieeigenwerte beim Silizium um den Faktor $\epsilon_r^2\, m/m^* \approx 450$ reduziert sind. Bemerkenswert ist in diesem Zusammenhang, daß der effektive Bohrsche Radius $a_0 = 4\pi\epsilon_r\epsilon_0\hbar^2/m^*e^2$ bei Donatoratomen in Silizium um den Faktor 40 größer ist als beim Wasserstoffatom. Innerhalb der Elektronenbahn liegen daher etwa 1000 Gitteratome!

[7] Üblicherweise wird die Hauptquantenzahl mit n bezeichnet, doch diese Abkürzung benutzen wir bereits für die Teilchenzahldichte.

Wir nutzen (4.10), um die Ionisierungsenergie zu berechnen. Setzen wir $\nu = 1$, so finden wir bei Silizium $E_1 = 30\,\mathrm{meV}$, also einen Wert, der etwa der thermischen Energie bei Zimmertemperatur entspricht. Bei dieser Temperatur sind die Donatoren daher, wie bereits erwähnt, ionisiert. Dies gilt auch für Akzeptoren, auf die wir hier nicht weiter eingehen. In Analogie zum Wasserstoff erwartet man weitere Energieeigenwerte bei höheren Quantenzahlen. Abbildung 4.10 macht deutlich, daß diese Erwartung berechtigt ist. Sie zeigt das Infrarotspektrum von antimondotiertem Germanium, bei dem mehrere Absorptionslinien zu erkennen sind. Die Messung wurde bei tiefer Temperatur durchgeführt, damit sich die nicht-bindenden Elektronen der Antimonatome im Grundzustand befinden. Das beobachtete Spektrum ist jedoch nicht so einfach strukturiert, wie es (4.10) erwarten läßt. Da sich die Elektronen der Donatoren im Kristallfeld mit kubischer Symmetrie und nicht in einem kugelsymmetrischen Coulomb-Feld bewegen, ist dies nicht überraschend. Oberhalb der Ionisierungsenergie von 9,6 meV – für Germanium gilt $\epsilon_r^2\, m/m^* \approx 1500$ – erfolgt der Übergang in kontinuierlich verteilte Zustände. Wie wir in Kap. 9 sehen werden, handelt es sich hierbei um Zustände des sogenannten „Leitungsbandes". Ähnlich aussehende Spektren beobachtet man auch beim Einbau von anderen Fremdatomen und anderen Halbleitern als Wirtskristall.

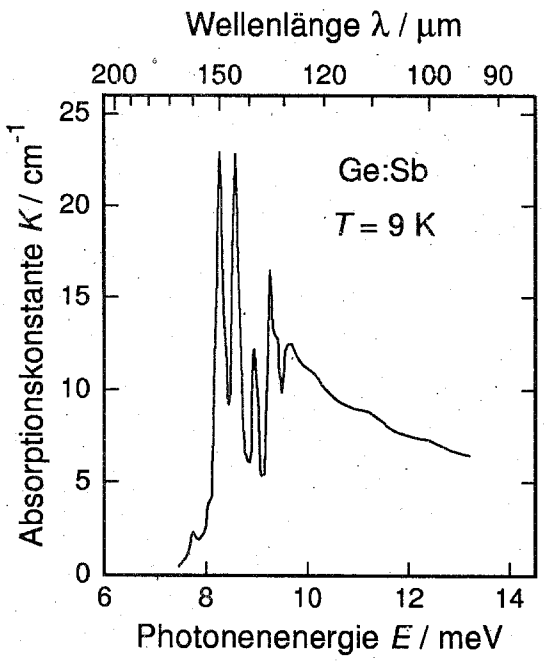


Abb. 4.10: Infrarotabsorption von Germanium, dotiert mit $7 \cdot 10^{26}\,\mathrm{m}^{-3}$ Antimonatomen. Es tritt eine Reihe diskreter Linien auf. (Nach J.H. Reuszer und P. Fisher, Phys. Rev. **135** (1964) A 1125).

Die theoretische Beschreibung der elektronischen Eigenzustände der Fremdatome läßt sich durch die Berücksichtigung von einigen festkörperspezifischen Eigenheiten wesentlich verbessern. So ist eine korrekte Behandlung des tatsächlich am Elektron herrschenden elektrischen Feldes erforderlich, denn es verursacht eine merkliche Aufspaltung des Grundzustandes und bewirkt Auswahlregeln bei den optischen Übergängen, die sich von denen des Wasserstoffatoms unterscheiden. Legt man Wert auf eine quantitative Beschreibung der Lage der Energieniveaus und vergleicht verschiedene Verunreinigungen und Wirtsgitter, so müssen die Eigenarten der betreffenden kovalenten Bindung berücksichtigt und der Einfluß der effektiven Massen und deren Richtungsabhängigkeit (vgl. die Abschnitte 8.1 und 9.2) in Rechnung gestellt werden. Die Übereinstimmung mit den experimentellen Daten wird dadurch besser, doch sind solche Feinheiten für das hier angestrebte Grundverständnis ohne Bedeutung.

Wird der substitutionelle Einbau der Atome nicht durch die kovalenten Bindungskräfte erzwungen, so befinden sich Fremdatome häufig auf Zwischengitterplätzen, wenn ihr Durchmesser kleiner als der der Wirtsatome ist. Als Beispiel einer typischen Eigenschaft von Zwischengitteratomen greifen wir das bereits erwähnte Diffusionsverhalten heraus. Da sich die Atome beim Sprung von einem Zwischengitterplatz zum anderen zwischen den Wirtsatomen auf den regulären Plätzen „durchzwängen" müssen, bewegen sie sich in einem Potential, das weitgehend dem bei Leerstellendiffusion (vgl. Abb. 4.3b) entspricht. Wir können daher (4.7) auch zur Berechnung des Diffusionskoeffizienten von Zwischengitteratomen heranziehen und erwarten deshalb einen exponentiellen Anstieg dieser Meßgröße mit der Temperatur. Die Daten in Abb. 4.11 für interstitielle Stickstoffatome in Eisenkristallen zeigen tatsächlich das erwartete Verhalten. Aus der Meßkurve lassen sich die Zahlenwerte $E_D = 0{,}85\,\mathrm{eV}$ und $D_0 = 0{,}05\,\mathrm{cm}^2/\mathrm{s}$ ablesen. Dies bedeutet, daß bei 1100 K, also weit entfernt von der Schmelztemperatur $T_f \approx 1800\,\mathrm{K}$, ein Stickstoffatom sich etwa einmal pro Nanosekunde um einen Zwischengitterplatz verschiebt, während es bei Zimmertemperatur mehrere Minuten am gleichen Platz verharrt.

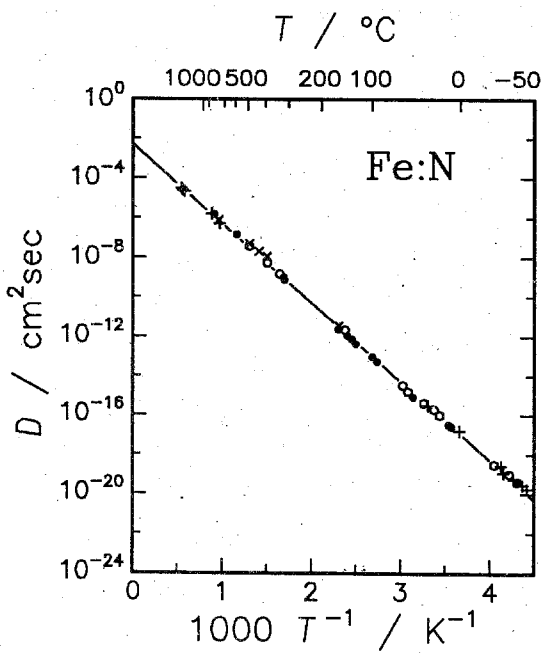


Abb. 4.11: Temperaturabhängigkeit des Diffusionskoeffizienten von Stickstoff in Eisen. Die Daten wurden mit verschiedenen Meßmethoden erzielt und stammen von verschiedenen Autoren. Es ist bemerkenswert, daß sich im untersuchten Temperaturintervall der Diffusionskoeffizient um etwa 16 Größenordnungen ändert! (Nach A.E. Lord, Jr. und D.N. Beskers, Acta Met. **14** (1966) 1659).

Die starke Temperaturabhängigkeit des Diffusionsprozesses macht man sich beim „Dotieren" von Bauelementen in der Halbleitertechnologie zunutze. Dort läßt man häufig von der Oberfläche her die bereits diskutierten Fremdatome wie Phosphor oder Arsen in den Halbleiter eindiffundieren, wobei das „Dotierungsprofil" über Zeit und Temperatur gesteuert werden kann. Nach dem Abkühlen auf Raumtemperatur sind die eindiffundierten Fremdatome praktisch unbeweglich. Ein weiteres interessantes und für zukünftige Anwendung in der Energietechnik möglicherweise wichtiges Beispiel ist die Einlagerung von Wasserstoff in Metallen. In sehr vielen Metallen wie z.B. Lithium, Magnesium, Palladium oder Titan und in deren Legierungen lassen sich sehr große Mengen von Wasserstoff einbauen. Die Wasserstoffatome geben ihr Elektron an das Metall ab und gehen in Form von Ionen, d.h. als Protonen, auf Zwischengitterplätze. Dabei kann die Zahl der Zwischengitteratome größer sein als die Zahl der Atome des regulären Wirtmetalls. Natürlich treten

beim Einbau von so großen Wasserstoffmengen auf Zwischengitterplätzen erhebliche Gitterverzerrungen auf, die zu Volumenvergrößerungen bis zu 10% führen. Bei sehr hohen Konzentrationen bilden sich neue chemische Verbindungen, sogenannte Metallhydride, aus. Die Wasserstoffatome sind dann ein Teil der Basis von Kristallen mit geänderten Gitterkonstanten.

Bei Kristallen mit einer Basis, die aus einer größeren Anzahl von Atomen besteht, ist in vielen Fällen die Defektstruktur komplexer. Bereits bei zweiatomiger Basis kann es durch den Platzwechsel von zwei benachbarten, chemisch unterschiedlichen Atomen geschehen, daß reguläre Gitterplätze mit „falschen" Atomen besetzt sind. Diese Defektart hat große Bedeutung bei den Verbindungshalbleitern (vgl. Kap. 9), da sie zu einer ungewollten Dotierung führt. Diese „Anti-site-Defekte" treten besonders dann häufig auf, wenn die beteiligten Atome vergleichbare Durchmesser aufweisen, wie dies beispielsweise beim Halbleiter Indiumantimonid der Fall ist.

4.2 Ausgedehnte Defekte

Im Gegensatz zu den elektrischen und optischen Eigenschaften, die sich durch Punktdefekte stark verändern lassen, werden die *mechanischen Eigenschaften* der Festkörper kaum von diesen Defekten, sondern vor allem durch ausgedehnte Defekte beeinflußt. Hierzu zählen die *Versetzungen*, die man durch eine Linie beschreiben kann, längs der die Gitterstörungen aufgereiht sind. Außerdem findet man in polykristallinen Materialien *Korngrenzen*, die zwischen unterschiedlich orientierten Kristallen auftreten. Diese beiden Defektarten wollen wir hier kurz betrachten.

Mechanische Eigenschaften

Nimmt man einen gewöhnlichen Aluminiumdraht und zieht leicht daran, so wird er sich dem Hookeschen Gesetz entsprechend verlängern. Zieht man „wesentlich kräftiger", so tritt eine irreversible Verformung des Drahtes ein. Der Verlauf der Drahtdehnung $\delta l/l$ als Funktion der angelegten Zugspannung σ für ein typisches polykristallines Material ist in Abb. 4.12a schematisch dargestellt. Bis zur Spannung σ_f ist die auftretende Dehnung reversibel. Daran schließt sich ein Bereich an, der durch „Fließvorgänge" gekennzeichnet ist. Da die erforderlich Spannung auch in diesem Bereich zunimmt, spricht man von „Fließverfestigung". Bei σ_b wird die Bruchfestigkeit erreicht. Bei noch größeren Werten der Dehnung verformt sich der Draht lokal und reißt schließlich. Da bei den irreversiblen Vorgängen auch die zeitliche Entwicklung von großer Bedeutung ist, werden derartige Zug-Dehnung-Diagramme meist bei konstanter Verformungsgeschwindigkeit aufgenommen[8].

Das Zug-Dehnung-Diagramm eines Aluminiumeinkristalls ist in Abb. 4.12b skizziert. Da Kristalle anisotrope mechanische Eigenschaften aufweisen, spielt ihre Orientierung bezüglich der Zugrichtung eine wesentliche Rolle. Allgemein gilt, daß bei Einkristallen irreversible Vorgänge wie die „plastische Deformation" bereits bei wesentlich geringeren Spannungen als bei polykristallinen Proben einsetzen. Wie in Abb. 4.12b zu sehen ist, verformt sich dabei der Kristall praktisch bei konstanter Zugspannung. Daran schließt sich ein Bereich an, in dem sich die Probe durch die Ausbildung von Versetzungen ähnlich wie in polykristallinen Materialien wieder verfestigt, worauf wir am Ende dieses Abschnitts noch kurz eingehen. Bei der plastischen Deformation gleiten Teile des Kristalls als Ganzes

[8] Typische Verformungsraten $\mathrm{d}(\delta l/l)/\mathrm{d}t$ liegen bei derartigen Experimenten um 10^{-3}/s.

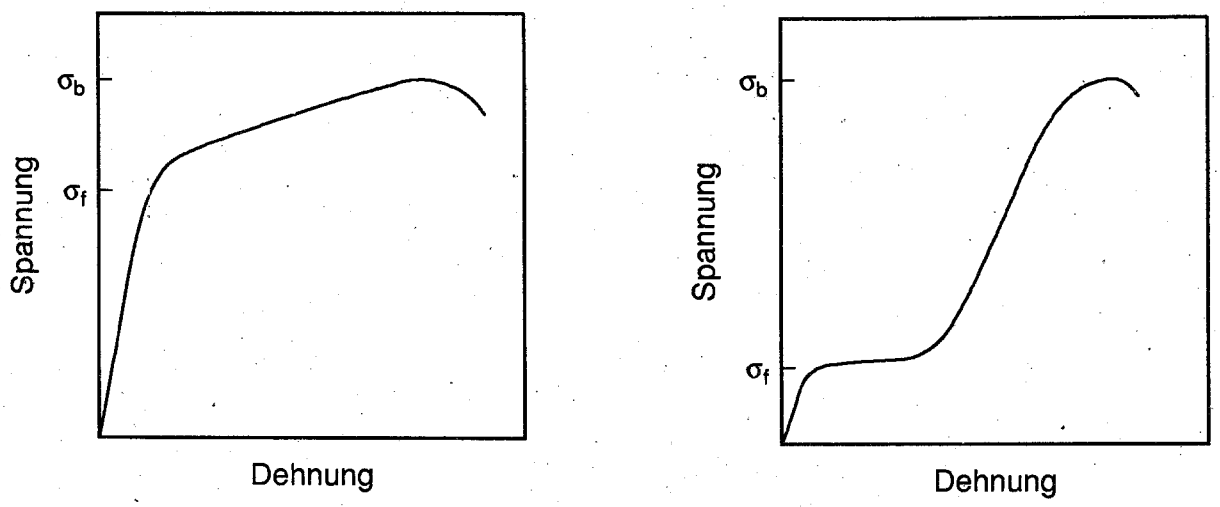


Abb. 4.12: Zug-Dehnung-Diagramme von polykristallinen und einkristallinen Proben.
a) Deformation eines polykristallinen Werkstücks unter mechanischer Zugspannung. Die Spannungen σ_f und σ_b geben die Fließ- bzw. Bruchgrenze an.
b) Deformation eines Einkristalls unter mechanischer Zugspannung. Ein typisches Merkmal ist das Auftreten von plastischer Deformation, bei der die Probe unter nahezu konstanter Spannung gedehnt wird.

über angrenzende Teile. Wie in der Schemazeichnung 4.13a zu sehen ist, bilden sich dabei leicht verschobene Bänder aus, die bemerkenswerterweise die gleiche Perfektion aufweisen wie der ursprüngliche Einkristall. Abbildung 4.13b zeigt in einer lichtmikroskopischen Aufnahme, wie die ursprünglich glatte Oberfläche eines Aluminiumkristalls nach einem Zugversuch aussieht, bei dem plastische Deformation aufgetreten ist. Deutlich erkennbar sind die Bänder, die bei der Deformation entstanden sind.

Abb. 4.13: Plastische Deformation von Einkristallen.
a) Unter dem Einfluß der Scherspannung, die aufgrund der Orientierung des Kristalls auftritt, gleiten Teile des Kristalls beim Ziehen längs wohldefinierter Netzebenen ab.
b) Lichtmikroskopische Aufnahme der Oberfläche eines Aluminiumeinkristalls nach plastischer Deformation. Durch die Verformung sind deutlich sichtbare Bänder entstanden. (Nach B. Jaoul, Bull. Soc. Fr. Mecan. **2** (1955) 69).

Kritische Schubspannung

Bevor wir auf den Gleitmechanismus zu sprechen kommen, wollen wir die *kritische Schubspannung* abschätzen, bei deren Überschreitung das Gitter instabil werden und Gleiten eintreten sollte. Die einfache Abschätzung wird zeigen, daß das Konzept des idealen Kristalls bei der Erklärung der plastische Deformation von Einkristallen zu völlig falschen Vorhersagen führt.

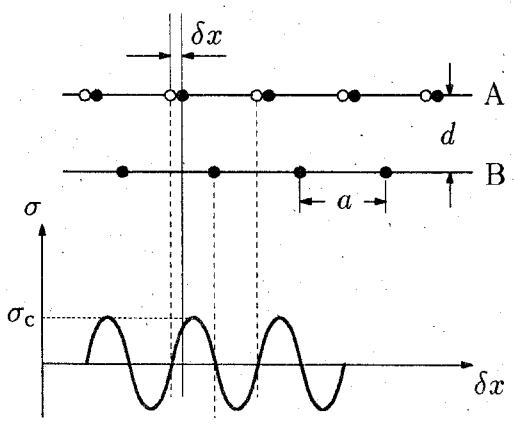


Abb. 4.14: Zur Abschätzung der kritischen Schubspannung. Die Netzebene A ist um die Strecke δx gegen die Gleichgewichtslage (offene Kreise) verschoben. Die erforderliche mechanische Spannung σ ist in der unteren Bildhälfte dargestellt. Liegt die Schubspannung oberhalb ihres kritischen Wertes σ_c, so gleiten die beiden Netzebenen übereinander hinweg.

In der schematischen Zeichnung 4.14 sind die für die Abschätzung wichtigen Größen dargestellt. Bei kleinen Auslenkungen δx besteht nach dem Hookeschen Gesetz $\sigma = \mathcal{G}e$ zwischen der mechanischen Schubspannung σ und der verursachten Verzerrung $e = \delta x/d$ ein linearer Zusammenhang, wobei $\mathcal{G}$ für den Schubmodul steht. Bei Vergrößerung der Spannung tritt irgendwann der Fall ein, daß die Atome der beiden benachbarten Netzebenen A und B übereinander zu liegen kommen. Wird nun die Kraft wieder weggenommen, dann bewegt sich die Netzebene A entweder zurück in die Ausgangslage, oder sie rastet in eine neue stabile Lage ein, die um eine Gitterkonstante a gegen die ursprüngliche verschoben ist. Bei unserer weiteren einfachen Betrachtung nähern wir die Schubspannung durch die Sinusfunktion

$$\sigma = \frac{\mathcal{G}a}{2\pi d} \sin\left(\frac{2\pi\,\delta x}{a}\right) . \tag{4.11}$$

Der Vorfaktor $\mathcal{G}a/2\pi d$ ergibt sich aus der Forderung, daß für kleine Auslenkungen das Hookesche Gesetz gilt. Die kritische Schubspannung σ_c ist erreicht, wenn der Faktor $\sin(2\pi\delta x/a) \approx 1$ ist, denn dann reicht die Kraft aus, die Netzebenen übereinander hinwegzuschieben. Mit $a \approx d$ können wir daher in grober Näherung für die kritische Schubspannung

$$\sigma_c \approx \frac{\mathcal{G}a}{2\pi d} \approx \frac{\mathcal{G}}{2\pi} \tag{4.12}$$

schreiben. Ein Vergleich zwischen den geschätzten und tatsächlich an Aluminiumproben gemessenen Werten, die in Tabelle 4.2 aufgeführt sind, macht deutlich, daß eine überraschend große Diskrepanz besteht.

Die gemessenen kritischen Schubspannungen sind viel kleiner als die berechneten! Die Abweichung zwischen den beiden Werten ist entgegen der Erwartung besonders groß beim Einkristall und wesentlich geringer beim polykristallinen Material. Dem errechneten Wert am nächsten kommt das in der Technik verwendete Aluminium, das Verunreinigungen enthält. Das Ergebnis legt den Schluß nahe, daß plastische Verformung von Einkristallen

nicht, wie bei der Abschätzung der kritischen Schubspannung angenommen, in einem einfachen Übereinandergleiten von Gitterebenen besteht, sondern daß andere Mechanismen das Gleiten erleichtern. Tatsächlich beruht der Fließvorgang bei kristallinen Festkörpern auf der Bewegung von sogenannten Versetzungen, die wir nun eingehend betrachten.

Tabelle 4.2: Schubmodul und kritische Schubspannung von unterschiedlichen Aluminiumproben.

	$\mathcal{G}\left(\frac{\mathrm{N}}{\mathrm{m}^2}\right)$	$\sigma_c^{\mathrm{exp}}\left(\frac{\mathrm{N}}{\mathrm{m}^2}\right)$	$\mathcal{G}/\sigma_c^{\mathrm{exp}}$
Einkristall	$2{,}5 \cdot 10^{10}$	$4 \cdot 10^5$	60 000
Polykristallines Material	$\approx 2{,}5 \cdot 10^{10}$	$2{,}6 \cdot 10^7$	1000
Duraluminium[9]	$\approx 2{,}5 \cdot 10^{10}$	$3{,}6 \cdot 10^8$	70

Stufen- und Schraubenversetzungen

Es existieren zwei Grundtypen von Versetzungen, deren mikroskopischen Aufbau wir kurz beschreiben wollen. Zur Veranschaulichung einer *Stufenversetzung* kann man sich vorstellen, daß man einen Kristall einschneidet und in den Schnitt eine zusätzliche Gitterebene einzwängt. Das untere Ende dieser Ebene bezeichnet man als *Versetzungslinie*. Sie steht in Abb. 4.15 senkrecht auf der Zeichenebene und erstreckt sich im einfachsten Fall bis an die Oberflächen des Kristalls. In der Umgebung der Versetzungslinie herrscht bis zu einer Entfernung von mehreren Atomabständen eine starke, nach außen abfallende Verzerrung.

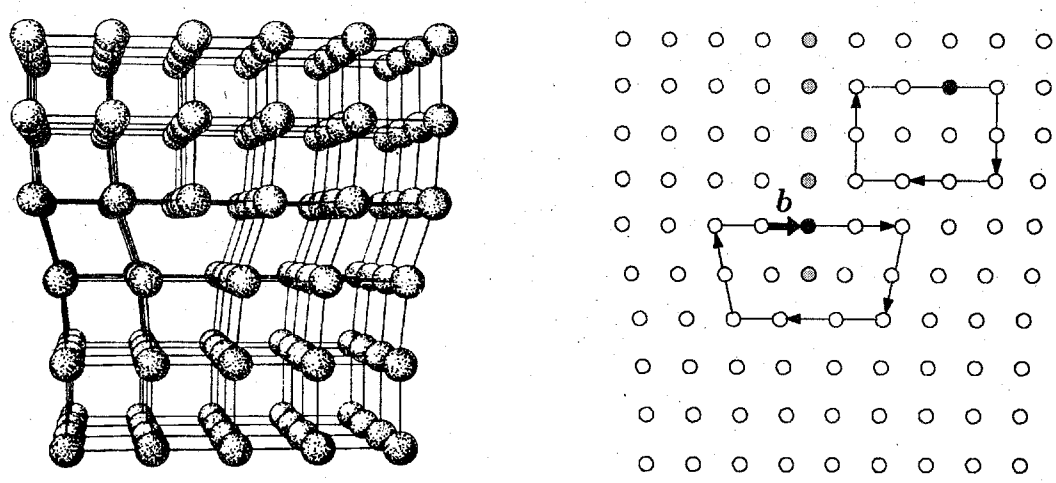


Abb. 4.15: Stufenversetzung.
a) Anschauliche Darstellung einer Stufenversetzung in einem kubischen Kristall.
b) Zur Definition des Burgers-Vektors. Die grauen Atome markieren die zusätzlich vorhandene Gitterebene. Die beiden eingezeichneten Pfade werden, jeweils vom schwarz markierten Gitterpunkt beginnend, im Uhrzeigersinn umlaufen. Der Burgers-Vektor $\boldsymbol{b}$ der Versetzung ist ebenfalls eingezeichnet.

[9] Bei technischen Anwendungen wird häufig Duraluminium verwendet, das auch Beimengungen anderer Metalle enthält.

Der Versetzungstyp läßt sich durch die in Abb. 4.15b gezeigte Prozedur charakterisieren: Zunächst durchläuft man im perfekten Teil des Kristalls einen geschlossenen Pfad, indem man von Gitterplatz zu Gitterplatz springt. Führt man nun die gleiche Schrittfolge im Bereich der Versetzungslinie durch, so endet man einen Gitterabstand entfernt vom Startpunkt. Der Vektor, der den zusätzlich erforderlichen Schritt repräsentiert, wird als *Burgers-Vektor* $\boldsymbol{b}$ bezeichnet. Legt man den Umlaufsinn der Bewegung fest, so ist die Richtung des Burgers-Vektor unabhängig vom Weg, auf dem man die Versetzungslinie umläuft. Wie dem Bild zu entnehmen ist, steht im Fall der Stufenversetzung der Burgers-Vektor senkrecht auf der Versetzungslinie.

Eine *Schraubenversetzung* erhält man, wie in Abb. 4.16 veranschaulicht, durch folgendes Vorgehen: Man schneidet einen Kristall ein und verschiebt die beiden Schnittflächen in der Nähe des Schnittendes parallel zur Schnittkante um einen Netzebenenabstand. Die Schnittkante markiert die Versetzungslinie. Der Burgers-Vektor läßt sich genauso wie bei der Stufenversetzung ermitteln, doch verläuft er nun parallel zur Versetzungslinie. Umfährt man die Versetzungslinie auf einer senkrecht dazu stehenden Gitterebene, so bewegt man sich auf einer Schraubenfläche entlang der Versetzungslinie. Dies erklärt den Namen dieser Versetzung.

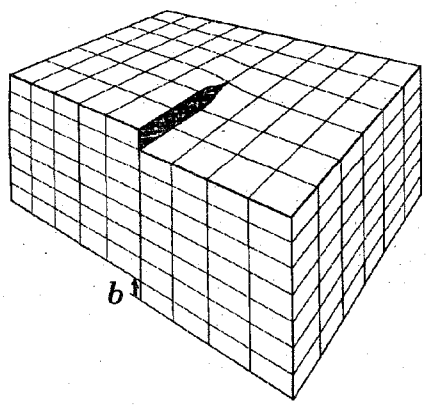


Abb. 4.16: Schematische Darstellung einer Schraubenversetzung in einem tetragonalen Gitter. Versetzungslinie und Burgers-Vektor $\boldsymbol{b}$ verlaufen parallel zu einander.

Grundsätzlich sind auch Anordnungen denkbar, bei denen der Betrag des Burgers-Vektors kleiner als ein Gitterabstand ist. Bei derartigen Konfigurationen fällt die Verzerrung nicht in der unmittelbaren Umgebung der Versetzungslinie ab, sondern die Verschiebung der Gitterebenen gegeneinander erstreckt sich über weitere Bereiche. Dadurch erhöht sich die innere Energie sehr stark, so daß solche Konfigurationen nicht stabil sind.

Im einfachsten Fall läuft die Versetzungslinie von einer Oberfläche zur anderen durch den ganzen Kristall. Weitaus häufiger treten aber *Versetzungsringe* auf, die in sich geschlossen sind und vollständig im Kristallinnern verlaufen. Das einfachste Beispiel ist eine eingeschobene Netzebene, deren Ausdehnung so klein ist, daß sie nirgendwo bis an den Kristallrand heranreicht. Entlang der Begrenzung dieser Ebene läuft eine Stufenversetzung in der Form eines geschlossenen Versetzungsrings. Dagegen ist mit Schraubenversetzungen eine derartige Konstruktion nicht möglich, denn die Versetzungslinie verläuft immer längs einer Geraden. Dennoch sind Schraubenversetzungen oft Teilstücke eines Versetzungsrings, in dem sie sich mit Stufenversetzungen abwechseln.

Bisher haben wir streng zwischen Stufen- und Schraubenversetzungen unterschieden und von zwei Grundtypen gesprochen. Es ist aber richtiger, darin nicht verschiedene Versetzungstypen zu sehen, sondern sie als Versetzungen mit einer speziellen Orientierung der Versetzungslinie bezüglich des Burgers-Vektors zu betrachten, nämlich mit einem eingeschlossenen Winkel von 90° bzw. 0°. Es läßt sich zeigen, daß der Burgers-Vektor längs einer beliebig geformten Versetzungslinie immer konstant ist. Läuft man entlang einer Versetzungslinie, so kann sich aber der *Charakter* der Versetzung ändern, der durch den Winkel zwischen Versetzungslinie und Burgers-Vektor bestimmt ist.

Wir wollen noch einmal einen Blick auf die Abbildungen. 4.15 und 4.16 werfen, denn sie machen deutlich, daß Stufenversetzungen nicht aber Schraubenversetzungen, eine Quelle bzw. Senke für Leerstellen und Zwischengitteratome bilden. Diffundiert bespielsweise ein Atom vom Ende der eingeschobenen Ebene in Abb. 4.15 nach unten in den Kristall, so entsteht ein Zwischengitteratom, und die Versetzungslinie wandert nach oben. Die erforderliche Bildungsenergie für das Zwischengitteratom ist bei dieser Art der Erzeugung wesentlich geringer als im idealen Gitter, da in der Umgebung der Versetzung das Gitter bereits passend verzerrt ist. Diffundiert umgekehrt ein Zwischengitteratom zur Stufenversetzung, so wird es dort „vernichtet", denn der geschilderte Prozeß läuft nun im umgekehrter Richtung ab. Auf gleiche Weise können Leerstellen entstehen oder verschwinden. Damit wird auch klar, daß unsere Diskussion der Ursache für das Auftreten von Nichtgleichgewichtskonzentrationen bei den Leerstellendichten in Abschnitt 4.1 zu stark vereinfacht war. In Wirklichkeit müssen die Leerstellen beim Abkühlen nicht immer bis an die Oberfläche diffundieren, um zu verschwinden, denn die Existenz von Stufenversetzungen verkleinert den erforderlichen Diffusionsweg der Leerstellen und damit gleichzeitig die tatsächlich vorhandene Konzentration.

Theoretisch interessant und von großer praktischer Bedeutung ist die Beantwortung der Frage nach der Zahl der Versetzungen, die in einem Kristall anzutreffen sind. Um diese Frage beantworten zu können, müßten wir den Beitrag der Versetzungen zur inneren Energie der Probe kennen, d.h. wir müßten wissen, welche Energie in der elastischen Verzerrung des Gitters gespeichert ist, die durch die Versetzung hervorgerufen wird. Eine einfache Rechnung zeigt, daß sie mit etwa 5 eV pro Atom der Versetzungslinie immer wesentlich größer als die Bildungsenergie von Leerstellen und vergleichbar mit der Bildungsenergie für Zwischengitteratome ist. Die Konfigurationsentropie dagegen ist, wieder verglichen mit der von Leerstellen, verschwindend klein. Dies wird verständlich, wenn man bedenkt, daß eine Versetzung durch eine lückenlose Aneinanderreihung von Punktdefekten gebildet wird. Die Forderung nach einer lückenlosen Reihe verkleinert die Anzahl der denkbaren Anordnungen der Punktdefekte erheblich, so daß die Konfigurationsentropie der Versetzungen wesentlich kleiner ist als die einer entsprechenden Anzahl von Punktdefekten. Der Beitrag der Entropie zur freien Energie ist daher zu vernachlässigen. Die freie Energie, die bekanntlich im thermischen Gleichgewicht ein Minimum aufweist, hat daher ihren Minimalwert dann, wenn die innere Energie minimal ist, d.h. wenn keine Versetzungen vorhanden sind. Tatsächlich beobachtet man aber typischerweise eine Dichte von etwa 10^8 Versetzungen pro cm^2, wobei man unter der *Versetzungsdichte* die Zahl der Versetzungslinien versteht, die eine Einheitsfläche im Kristall durchsetzen. Die Versetzungen bilden sich bei der Herstellung der Kristalle und stellen einen eingefrorenen, metastabilen Zustand dar. Tatsächlich spielen Versetzungen beim Kristallwachstum eine wichtige Rolle. Wie gezeigt werden konnte, führt ihre Gegenwart zu einer Erhöhung der Wachstumsgeschwindigkeit. In „guten" Einkristallen findet man etwa 10^5 Versetzungen pro cm^2, bei sehr sorgfältig gezogenen Kristallen kann die Versetzungsdichte bis

auf einige Versetzungen/cm^2 zurückgehen. Dagegen erreicht die Versetzungsdichte bei „kaltgehärteten“ Metallen, d.h. bei Metallstücken, die bei niedriger Temperatur verformt wurden, Werte um $10^{12}\,\mathrm{cm}^{-2}$.

Abbildung von Versetzungen

Versetzungen lassen sich mit verschiedenen Methoden sichtbar machen. Eine der ersten Nachweismöglichkeiten beruhte auf chemischem Ätzen von Probenoberflächen. Hierbei wird ausgenutzt, daß die Ätzgeschwindigkeit in der Umgebung einer Versetzung erhöht ist, da die Atome in der stark verzerrten Struktur schwächer gebunden sind. Stößt eine Versetzung an die Oberfläche, so bildet sich daher an dieser Stelle beim Ätzen eine kleine Grube aus, die sich leicht mit dem Mikroskop beobachten läßt. Ein Beispiel hierfür (vgl. Abb. 4.24) werden wir noch im Zusammenhang mit der Diskussion der sogenannten Kleinwinkelkorngrenzen kennenlernen. Diese relativ einfache Methode wird im allgemeinen angewandt, wenn die Versetzungsdichte bestimmt werden soll.

Bei durchsichtigen Kristallen kann unter bestimmten Voraussetzungen sogar der Verlauf der Versetzungen im Kristallinnern sichtbar gemacht werden. Dies ist möglich, weil sich Zwischengitteratome entlang von Versetzungen besonders leicht bewegen können. So läßt man zunächst Silber- oder Kupferatome in die Probe eindiffundieren. Durch „geeignete“ thermische (und/oder optische) Behandlung bilden sich dann Präzipitate (Ausfällungen), die im Mikroskop sichtbar sind. In Abb. 4.21 ist die Aufnahme einer sogenannten Frank-Read-Quelle zu sehen, die mit dieser Technik gewonnen wurde.

In dünnen Filmen (siehe Abb. 4.19) lassen sich Versetzungen mit Hilfe der Transmissionselektronenmikroskopie sichtbar machen, wobei die Bilderzeugung auf der lokalen Dichtevariation beruht, d.h. auf den Verzerrungen, die mit der Existenz der Versetzungen verbunden sind. In dickeren Proben sind ähnliche Untersuchungen mit Hilfe der sogenannten Röntgentopographie möglich. Dabei wird ausgenutzt, daß Bragg-Streuung lokal unter leicht modifiziertem Winkel auftreten kann, wenn Verzerrungen die Gitterabstände etwas verändert. Wir wollen jedoch hier keine Beschreibung der etwas komplizierten Grundlagen dieser Technik liefern.

Versetzungen an der Oberfläche lassen sich in jüngerer Zeit auch mit dem Rastertunnelmikroskop abbilden. Als Beispiel ist in Abb. 4.17 eine Schraubenversetzung an der Oberfläche eines $Pt_{25}Ni_{75}$-Kristalls zu sehen. Die Ähnlichkeit mit der schematischen Darstellung Abb. 4.16 ist unverkennbar.

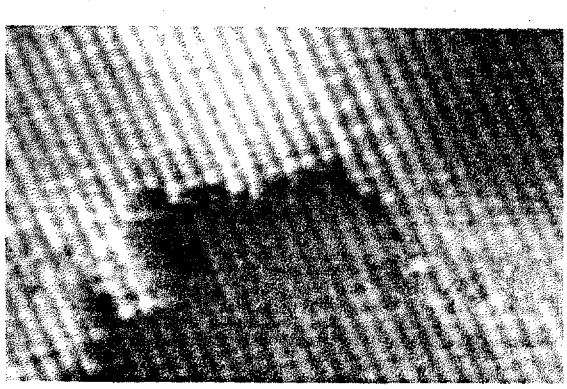


Abb. 4.17: Schraubenversetzung an der Oberfläche eines $Pt_{25}Ni_{75}$-Kristalls. Die Abbildung mit atomarer Auflösung wurde mit einem Rastertunnelmikroskop gemacht. (Mit freundlicher Genehmigung von M. Schmid und P. Varga; Inst. Allgem. Physik, TU Wien).

Plastische Deformation durch Gleiten

Wir kommen nochmals auf die Beobachtung zurück, daß die kritische Schubspannung für die plastische Verformung von einkristallinen Metallen wesentlich kleiner ist, als wir mit Hilfe von Gleichung (4.12) abgeschätzt haben. Die Ursache hierfür macht Abb. 4.18 für den Fall der Stufenversetzung plausibel.

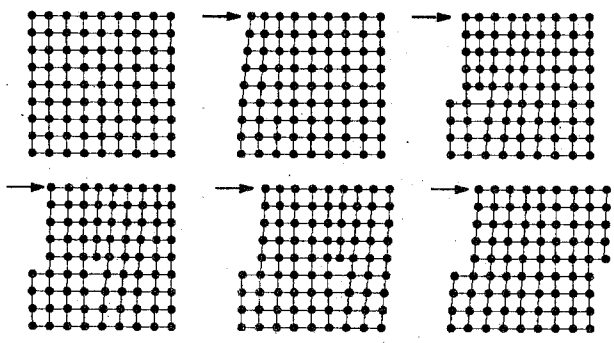


Abb. 4.18: Schervorgang unter Ausbildung einer Stufenversetzung. Die Versetzungslinie wandert durch den Kristall, der an der Unterlage festgehalten wird. Der Pfeil deutet die Schubrichtung an.

Unter dem Einfluß der Schubspannung bildet sich eine Versetzung aus, die durch den Kristall wandert. Von außen gesehen gleitet die obere Kristallhälfte über die untere hinweg. Da sich die einzelnen Gitterebenen „nacheinander" bewegen, muß nur eine „lokale" kritische Spannung aufgebracht werden. Die tatsächlich auftretende kritische Spannung σ_c ist bei diesem Mechanismus wesentlich geringer als der abgeschätzte Wert, bei dem die gleichzeitige Bewegung aller Gitterebenen vorausgesetzt wurde. Abhängig von den angreifenden Kräften und der Art der vorhandenen oder erzeugten Versetzung erfolgt die Scherbewegung senkrecht oder parallel zur Versetzungslinie. Da die plastische Deformation über die Bewegung von Versetzungen erfolgt, ist die tatsächliche Festigkeit eines Materials im wesentlichen durch die Mechanismen bestimmt, welche die Bewegung von Versetzungen blockieren. Dies bewirken vor allem Gitterfehler, die durch Verunreinigungen verursacht werden. Beim „Härten" von Stahl beispielsweise entstehen Eisenkarbid-Teilchen, von denen die Versetzungen festgehalten werden. Eine andere Möglichkeit stellt die Ausbildung von Versetzungsnetzwerken dar. Da die gegenseitige Durchdringung von Versetzungen Energie kostet[10], wird die Bewegung der Versetzungslinien mit zunehmender Anzahl von Versetzungen stark behindert. „Kalthärten" beruht auf diesem Effekt, denn bei der plastischen Deformation werden zusätzliche Versetzungen erzeugt, die sich dann nicht einfach gegenseitig durchdringen können. In allen Fällen wirkt die Diffusion der Atome der Verankerung der Versetzungen entgegen, so daß die Festigkeit mit zunehmender Temperatur abnimmt.

Ein Beispiel für die Verankerung von Versetzungen ist in Abb. 4.19 in einer elektronenmikroskopischen Aufnahme zu sehen. Unter dem Einfluß der Scherung wandern die Versetzungen, werden aber an Defekten festgehalten, so daß Knicke im Verlauf der Versetzungen auftreten.

[10] Dies ist eine Konsequenz der Tatsache, daß sich Versetzungen aufgrund ihres Verzerrungsfeldes abstoßen.

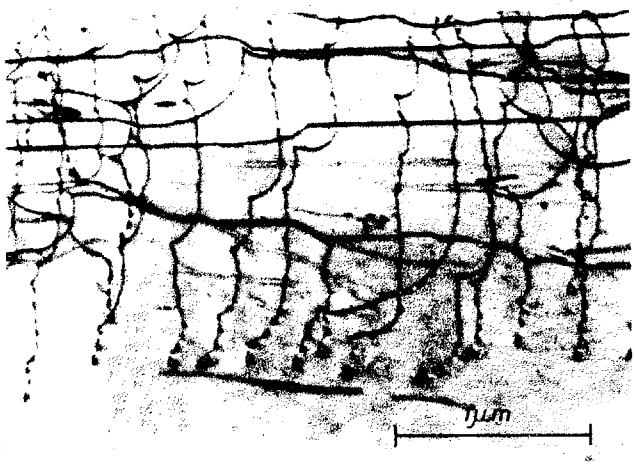


Abb. 4.19: Elektronenmikroskopische Aufnahme von Versetzungen in Magnesiumoxid nach Anwendung von mechanischer Spannung. An den Verankerungsstellen tritt ein Knick im Verlauf der Versetzungslinie auf. (Aus B.K. Kardashev et al., Phys. Stat. Solidi A **91** (1985) 79).

Zum Schluß soll hier noch ein quantitatives Problem angesprochen werden. Läuft eine Versetzung durch den Kristall, wie in Abb. 4.18 skizziert, so verschieben sich die beiden Kristallhälften nur um eine Gitterkontante gegeneinander. Eine makroskopische Verschiebung ist nur möglich, wenn während der Verformung ständig neue Versetzungen erzeugt werden. Ein Beispiel dafür, wie eine bereits bestehende Versetzung als Ausgangspunkt für weitere Versetzungen wirken kann, ist die sogenannte *Frank-Read-Quelle*, deren Wirkungsweise wir hier kurz erläutern wollen.

Der Mechanismus, auf dem dieser Quelle beruht, ist schematisch in Abb. 4.20 skizziert. Von der ursprünglichen Versetzung K ist nur das gerade Teilstück 1 sichtbar, das zwischen den Ankerpunkten A und B verläuft. Der restliche Teil der Versetzung (von B nach A) soll unterhalb der Zeichenebene verlaufen und ist deswegen nicht dargestellt. Unter dem Einfluß einer Scherspannung beult sich die Versetzung K zwischen den Ankerpunkten aus und durchläuft mit zunehmender Spannung die Stadien 1 bis 6. Zwischen den beiden „Momentaufnahmen" 5 und 6 berührten sich die Ausbeulungen. Dadurch entstand der neuer Versetzungsring K', und die ursprüngliche Versetzung K befindet sich wieder im Anfangszustand 1. Dieser Vorgang, bei dem immer eine neue Versetzung erzeugt wird, kann sich beliebig oft wiederholen.

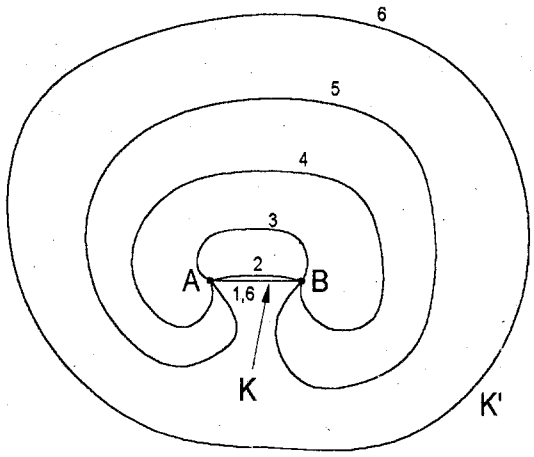


Abb. 4.20: Frank-Read-Mechanismus zur Erzeugung von Versetzungen. Unter dem Einfluß der Scherspannung durchläuft die Versetzung K die Stadien 1 bis 6. Am Ende des Zyklus ist die ursprüngliche Versetzung K wieder im Anfangszustand und zusätzlich ist die Versetzung K' entstanden.

Abb. 4.21 zeigt die Aufnahme einer Frank-Read-Quelle in Silizium. Die Versetzungen sind mit Kupferpräzipitaten dekoriert und so sichtbar gemacht. Da Silizium im sichtbaren Wellenlängenbereich undurchlässig ist, erfolgte die Abbildung mit Infrarotlicht. Neben den neu entstandenen Versetzungen ist auch der Teil der ursprünglichen Versetzung zu erkennen, der weiter im Innern des Kristalls verläuft.

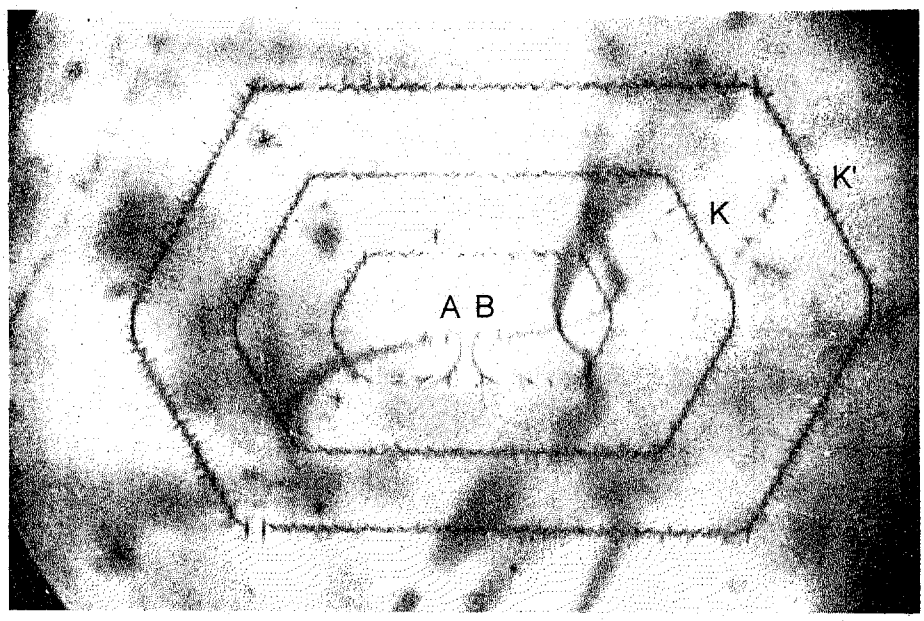


Abb. 4.21: Frank-Read-Quelle in Silizium. Die Versetzungen sind mit Kupferpräzipitaten dekoriert. Die Abbildung erfolgte unter Infrarotbeleuchtung, da Silizium im Sichtbaren undurchlässig ist. (Aus W.C. Dash, J. Appl. Phys. **27** (1958) 1193).

Korngrenzen

Wie bereits in Abschnitt 1.2 erwähnt, muß man beim Ziehen von Einkristallen besondere Sorgfalt walten lassen. Die meisten Festkörper sind polykristallin, denn ohne besondere Vorkehrungen treten beim Abkühlen von Schmelzen Keime an verschiedenen Stellen auf, von denen aus Kristallite zu wachsen beginnen. Wie in Abb. 4.22 zu sehen ist, entsteht dadurch ein Festkörper, der aus vielen kleinen Kristalliten zusammengesetzt ist.

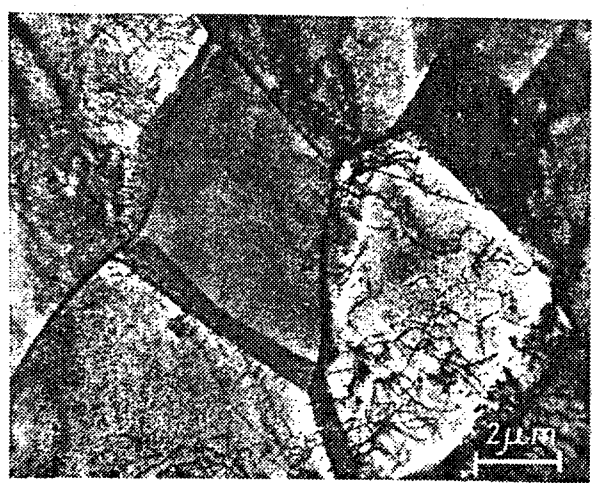


Abb. 4.22: Polykristallines Natriumchlorid. In der elektronenmikroskopischen Aufnahme sind die Korngrenzen zwischen den Kristalliten klar zu erkennen. Versetzungen machen sich als dunkle „Fäden" bemerkbar. (Aus E.M. Nadgorngi und H.P. Strunk, Phys. Stat. Solidi A **104** (1987) 193).

Die aneinander stoßenden Kristallite sind bei polykristallinen Materialien unterschiedlich orientiert. Zwischen ihnen liegen Bereiche, die man als *Korngrenzen* bezeichnet. Sie stellen zweidimensionale Defekte dar, die einen starken Einfluß auf die Eigenschaften von polykristallinem Material ausüben. Besonders stark betroffen ist hiervon das mechanische Verhalten, denn Korngrenzen reduzieren die Festigkeit der Festkörper. Beim Tempern heilen die Korngrenzen zum Teil aus, da die stark erhöhte Diffusionsgeschwindigkeit in der Umgebung der Korngrenzen die Umlagerung von Atomen in energetisch günstigere Anordnungen erleichtert. Dabei beobachtet man, daß die großen Kristallite auf Kosten der kleinen wachsen.

Die Eigenschaften von Korngrenzen hängen stark von der Orientierung der aneinander grenzenden Kristallite ab. Bei einem kleinen Unterschied in der Orientierung stellt die Korngrenze eine regelmäßige Anordnung von Versetzungen dar. Bei stark unterschiedlicher Orientierung ist die Struktur der Korngrenze „unkristallographisch", d.h. es handelt sich um relativ ungeordnete „innere Oberflächen", die aus einer Anhäufung von Punktdefekten und Versetzungen bestehen. In diesen flächenhaften Defekten lassen sich bei mechanischer Beanspruchung relativ leicht einzelne Atome bewegen. Dies ist auch die Ursache dafür, daß atomare Diffusion in den Korngrenzen im Vergleich zur Diffusion im Kristallinnern sehr erleichtert ist.

Seit einigen Jahren kann man die atomare Struktur von Korngrenzen mit Hilfe hochauflösenden Elektronenmikroskopie studieren. Eindrucksvolle Beispiele hierfür sind in Abb. 4.23 zu sehen, in der Korngrenzen in kubischen NiO-Kristallen wiedergegeben sind. Bei der hochauflösender Elektronenmikroskopie werden die Atomreihen häufig in Form dunkler Bereiche abgebildet, die hellen Gebiete repräsentieren die Zwischenräume. In beiden Bildern blicken wir auf die (100)-Ebenen von Kristallen, die unterschiedlich stark gegeneinander verkippt sind. Die schwarzen Gebiete stellen also die Projektion der Atomreihen in [100]-Richtung dar.

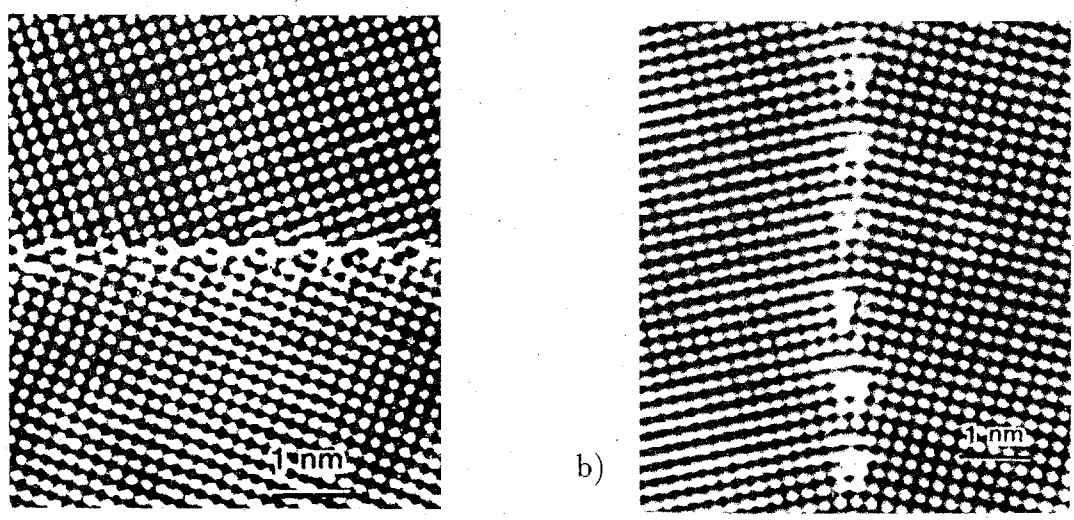


Abb. 4.23: Hochauflösende elektronenoptische Aufnahme von Korngrenzen in NiO mit unterschiedlicher Verkippung der Kristalle. (Aus K.L. Merkle und D.J. Smith, Phys. Rev. Lett. **59** (1987) 2887).
a) Korngrenze bei einem Verkippungswinkel von 36,9°,
b) Korngrenze bei einem Verkippungswinkel von 22,6°.

Die Korngrenze in Abb. 4.23a ist durch die Verkippung der beiden Kristalle um 36,9° entstanden. In dieser Korngrenze wiederholen sich nach drei aufeinander folgenden Gitterebenen relativ lockere Strukturen, in denen die atomare Dichte auf etwa 70% der Gitterdichte reduziert ist. Bei genauerer Betrachtung lassen sich zwei unterschiedliche atomare Anordnungen unterscheiden, die wir aber nicht weiter diskutieren wollen. Verkleinert sich der Orientierungsunterschied, so rücken die lockeren Strukturen weiter auseinander. In Abb. 4.23b treten bei einer Verkippung von 22,6° diese Strukturen nach jeweils fünf Atomlagen auf. Wie eine genauere Analyse zeigt, handelt es sich hierbei um Gebilde, die eine große Ähnlichkeit mit Stufenversetzungen besitzen und deren Versetzungslinien senkrecht zur Abbildungsebene verlaufen.

Bestimmt man in relativ guten Einkristallen die Ausrichtung von identischen Gitterebenen an zwei verschiedenen Stellen sehr exakt, so stellt man oft fest, daß diese geringfügig, d.h. um Winkel $< 1°$, gegeneinander verkippt sind. Die Korngrenzen zwischen diesen sogenannten *Mosaikblöcken* bezeichnet man bei so geringen Verkippungswinkeln als *Kleinwinkelkorngrenzen.* Wie in Abb. 4.23b für größere Winkel zu sehen war, stellen Kleinwinkelkorngrenzen eine Aneinanderreihung von Stufenversetzungen[11] dar, die sich längs der Korngrenzen in relativ großem Abstand periodisch wiederholen. Bei geeigneter Belastung wandern derartige Korngrenzen senkrecht zu den Versetzungslinien als Ganzes durch den Kristall. Ätzt man Kristalle mit Kleinwinkelkorngrenzen an, so werden die Versetzungen, wie bereits diskutiert, als kleine Ätzgruben sichtbar. In Abb. 4.24 ist ein derartig behandelter Germaniumkristall mit einer Kleinwinkelkorngrenze zu sehen. Bemerkenswert ist, wie schön die erwartete regelmäßige Anordnung der Versetzungen zu erkennen ist. Aus dem Abstand der Ätzgruben läßt sich unmittelbar die Verkippung zwischen den Mosaikblöcken ermitteln.

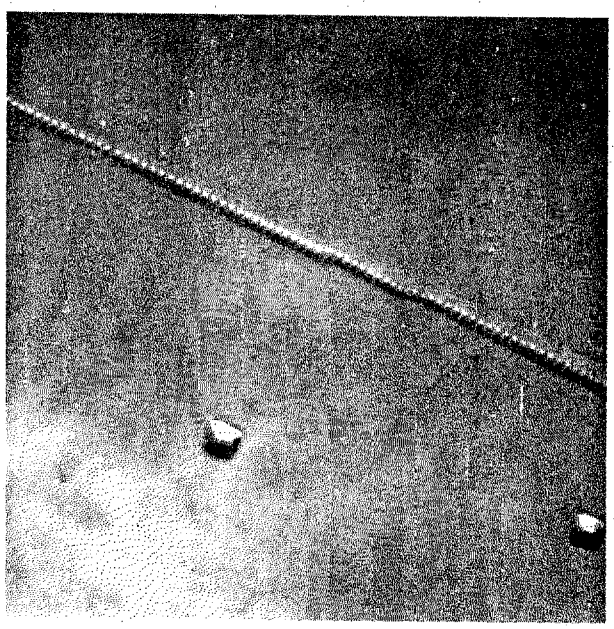


Abb. 4.24: Versetzungen längs einer Kleinwinkelkorngrenze. Die an der Oberfläche endenden, regelmäßig angeordneten Versetzungen sind als Ätzgruben deutlich sichtbar. (Aus F.L. Vogel, Acta Met. **3** (1957) 245).

[11] Gelegentlich sind die Teilkristalle leicht gegeneinander verdreht. In diesem Fall treten anstelle der Stufenversetzungen natürlich Schraubenversetzungen auf.

4.3 Defekte in amorphen Materialien

Naturgemäß ist die Definition von Defekten in amorphen Festkörpern wesentlich schwieriger als in Kristallen, wo Defekte als Abweichungen von der periodischen Anordnung der Atome aufgefaßt werden können. Ein extremer, aber zu stark vereinfachender Standpunkt wäre, im amorphen Zustand lediglich eine Anhäufung von Defekten zu sehen, da die Atome nicht auf wohldefinierten Plätzen zu finden sind. Dieser Standpunkt wäre gerechtfertigt, wenn nur die geometrische Anordnung der Atome von Bedeutung wäre und nicht berücksichtigt werden müßte, daß die wesentlichen Eigenschaften der Festkörper durch die Topologie der chemischen Nahordnung bestimmt werden.

In amorphen Festkörpern tritt wegen der „lockeren" Packung, die durch die Abweichung von der strengen periodischen Anordnung der Atome hervorgerufen wird, ein Dichteunterschied gegenüber den Kristallen auf. Aufgrund dessen existieren in der Probe Hohlräume, die eine große Ähnlichkeit mit den Leerstellen der Kristalle besitzen. Die Größe dieser Hohlräume ist jedoch nicht einheitlich, sondern variiert über einen weiten Bereich. So findet man auch Hohlräume, deren Volumen größer als das einzelner Atome ist. Ähnlich wie die Leerstellen der Kristalle hängt die Anzahl der Hohlräume, die bei Zimmertemperatur vorhanden ist, von der Kühlrate bei der Glasherstellung bzw. den Herstellbedingungen des amorphen Festkörpers ab. Wie Leerstellen in Kristallen so haben auch die Hohlräume in amorphen Festkörpern einen großen Einfluß auf die Diffusion von Fremdatomen und bestimmen weitgehend die Zähigkeit der Gläser in der Nähe des Glasübergangs.

Ein Defekttyp, der normalerweise nicht in Kristallen zu finden ist, ist die aufgebrochene oder *ungesättigte Bindung* (engl. *dangling bond*). In den technisch verwendeten Gläsern werden diese Defekte vor allem durch die beigemischten Metallionen oder durch andere Verunreinigungen hervorgerufen, die das Netzwerk aufbrechen. Sie treten aber auch in reinen amorphen Materialien auf und sind dort sehr eingehend studiert worden. Sie machen sich in Elektronenspinresonanz-Untersuchungen bemerkbar, denn die *ungepaarten* Elektronen verursachen aufgrund des magnetischen Moments ihrer Spins ein leicht erkennbares Signal. Erstaunlicherweise hängt die Anzahl dieser Defekte stark von der Koordinationszahl des amorphen Festkörpers ab. So findet man in den tetraedrisch gebundenen Halbleitern a-Si oder a-Ge etwa $10^{19} - 10^{20}$ offene Bindungen pro cm^3. In den Chalkogenidgläsern wie a-As_2S_3 oder a-Se liegt dagegen die Zahl der ungepaarten Elektronen unter der Nachweisgrenze. Diese um viele Größenordnungen unterschiedliche Konzentration von ungesättigten Bindungen in den beiden Substanzklassen war lange Zeit ein unverstandenes Problem.

Der in den letzten Jahren am intensivsten untersuchte Defekt dieser Art ist die ungesättigte Bindung in amorphem Silizium. Stellt man Schichten aus reinem amorphen Silizium z.B. durch Aufdampfen her, so kann das Auftreten einer sehr großen Zahl von ungesättigten Bindungen nicht vermieden werden. Dies liegt daran, daß die starre sp^3-Hybridbindung des Siliziums kaum Verkippungen der Atome gegeneinander erlaubt. Da Verkippungen aber Voraussetzung für eine regellose Anordnung der Atome in einem Netzwerk sind, muß eine große Zahl von Bindungen offen bleiben, da sonst die auftretenden mechanischen Spannungen zu groß wären. Die Existenz von ungesättigten Bindungen ist daher charakteristisch für vierfach koordinierte amorphe Netzwerke. Wie wir in Abschnitt 9.3 sehen werden, bestimmen offene Bindungen weitgehend die elektrischen

Eigenschaften von amorphen Halbleitern: Sie verursachen eine so hohe Leitfähigkeit, so daß Dotieren des halbleitenden Materials, wie es bei kristallinen Halbleitern üblich ist, keinen merklichen Einfluß auf die elektrischen Eigenschaften hat. Schichten aus reinem amorphen Silizium sind daher für technische Anwendungen ungeeignet. Dennoch werden Solarzellen aus diesem Material in großer Stückzahl produziert. Um für die technische Anwendung die elektrischen Eigenschaften grundlegend zu verbessern, baut man bereits bei der Herstellung der Schichten größere Mengen Wasserstoff ein. Aufgrund seiner Einwertigkeit führt er zu einer Absättigung der freien Bindungen, ohne selbst neue ungesättigte Bindungen hervorzurufen. Dabei entstehen neue Defektstrukturen, wie sie in Abb. 4.25 für den Fall gezeigt sind, daß noch keine vollständige Absättigung der Bindungen durch Wasserstoff stattgefunden hat.

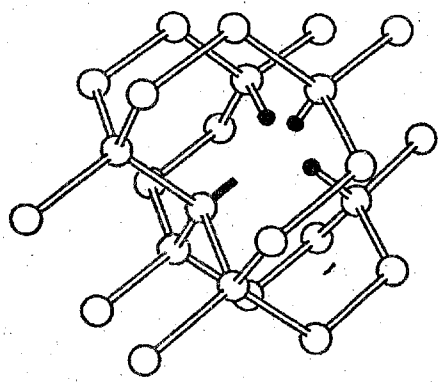


Abb. 4.25: Modell eines Defekts in amorphem Silizium. Drei offene Bindungen sind durch schwarz gezeichnete Wasserstoffatome abgesättigt. Eine Bindung ist jedoch noch ungesättigt. (Nach Y. Bar-Yam und J.D. Joannopoulos, Phys. Rev. Lett. **56** (1986) 2203).

Wie bereits erwähnt, lassen sich in den Chalkogenidgläsern keine ungepaarten Elektronen nachweisen, obwohl ebenfalls strukturelle Defekte auftreten. Wir wollen zur Diskussion der Defekte in dieser Substanzklasse als besonders einfaches Beispiel amorphes Selen herausgreifen. Wie auf der linken Seite in Abb. 4.26 zu sehen ist, liegen in diesem Material Ketten von Selenatomen mit s^2p^4-Konfiguration vor. Die Bindung innerhalb der Ketten wird von p-Elektronen vermittelt, doch werden hierzu nur zwei benötigt. An jedem Atom tritt daher noch ein nichtbindendes p-Elektronenpaar auf.

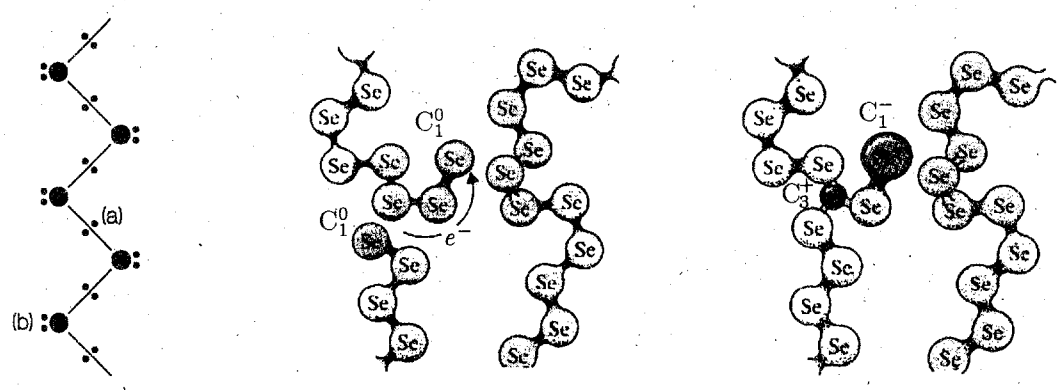


Abb. 4.26: Ausbildung eines valenz-alternierenden Paares in amorphem Selen. Links ist ein Teil einer Selenkette mit bindenden (a) und nicht-bindenden (b) Elektronen skizziert. In der Mitte sind drei Ketten veranschaulicht, wovon zwei im Bild enden und so zwei C_1^0-Defekte hervorrufen. Daraus kann sich unter Energiegewinn ein C_1^- – C_3^+-Paar (rechts) bilden. Die dritte Kette wirkt bei der Ausbildung des geladenen Defektpaars nicht mit.

An den Enden der unterschiedlich langen Ketten des amorphen Selens treten zunächst ungepaarte Elektronen auf. Das Endatom wird je nach benutzter Nomenklatur als D^0- oder als C_1^0-Defekt bezeichnet. Hierbei steht D^0 für „neutraler Defekt" und C_1^0 deutet an, daß es sich um ein neutrales Chalkogenatom handelt, das nur einfach koordiniert ist. Im mittleren Bild sind zwei Kettenenden, also zwei C_1^0-Defekte, und eine fortlaufende Kette dargestellt. Die beiden ersten Ketten reagieren miteinander: Sie bilden einen dreifach koordinierten, positiv geladenen C_3^+-Defekt und gleichzeitig einen negativ geladenen, einfach koordinierten C_1^--Defekt. Beide Defekte besitzen nur Elektronen*paare*, sind also ohne freien Spin und in Elektronenspinresonanzexperimenten nicht mehr nachweisbar.

Bei der Erzeugung des C_1^--Defekts muß Energie aufgewandt werden, da das überschüssige Elektron von den bereits vorhandenen Elektronen abgestoßen wird. Die Ausbildung der dritten Bindung beim C_3^+-Defekt ist andererseits mit einem Energiegewinn verbunden. Insgesamt ist die Reaktion $2C_1^0 \rightarrow C_3^+ + C_1^-$ exotherm, so daß sich diese für Chalkogenidgläser typische Defektstruktur ausbildet. Man bezeichnet die beiden Defektzustände zusammen als *valenz-alternierendes Paar* (VAP). Ähnliche Defekte findet man auch in anderen amorphen Festkörpern, doch ist die tatsächliche Defektstruktur meist komplizierter. Dies trifft in besonderem Maße für amorphe Verbindungen zu.

Eine zusätzliche Schwierigkeit bei der Beschreibung von Defektstrukturen besteht darin, daß bei der Bildung oder Zerstörung dieser Defektkonfigurationen das amorphe Netzwerk in der Umgebung der Defekte stark relaxiert, wobei die potentielle Energie durch lokale strukturelle Umlagerungen abgesenkt wird. Diese Relaxation ist in amorphen Festkörpern stärker ausgeprägt als in kristallinen, da das amorphe Netzwerk eine größere Variabilität in der räumlichen Anordnung der Atome besitzt. Strukturelle Relaxationen laufen aber relativ langsam ab, da eine größere Zahl von Atomen beteiligt ist. Daher reagieren Defekte in amorphen Materialien auf thermische und optische Anregung häufig unterschiedlich, da diese auf verschiedenen Zeitskalen ablaufen.

Die üblichen flächenhaften Defekte wie Korngrenzen können in amorphen Materialien nicht auftreten. Dies hat eine interessante und überaus wichtige Eigenart der Gläser zur Folge: Sie sind in den meisten Fällen durchsichtig. Zwar sind auch Einkristalle, die im sichtbaren Spektralbereich keine Absorption aufweisen, optisch transparent, doch die technisch meistens verwendeten polykristallinen Materialien sind milchig oder weiß. Die Ursache hierfür sind die Lichtreflexionen, die an den Oberflächen der unterschiedlich orientierten Kristalliten auftreten. Das Fehlen von Korngrenzen ist daher für den täglichen Gebrauch von Gläsern von entscheidender Bedeutung.

4.4 Legierungen

In Kap. 4.1 wurde der Einbau von Fremdatomen in das Kristallgitter angesprochen. Bei diesen Betrachtungen gingen wir von einer geringen Konzentration von Fremdatomen aus. Was geschieht, wenn die Zahl der „Fremdatome" vergleichbar mit der Anzahl der „Wirtsatome" ist? Dieser Fall tritt häufig bei den technisch genutzten metallischen Werkstoffen auf. Je nach Art des Einbaus der Fremdatome spricht man in diesem Fall von *Substitutionslegierungen* oder von *interstitiellen Legierungen*. Wir wollen hier einige Fragen zur Struktur der Legierungen aufgreifen, gehen aber sonst auf die Eigenschaften dieser interessanten Stoffklasse, die ein sehr wichtiges Gebiet der Materialforschung darstellen, nicht weiter ein.

Die Atome der Legierung sitzen zwar auf regulären Gitterplätzen, doch können diese Plätze, wie bereits in Abb. 1.3b gezeigt, statistisch mit unterschiedlichen Atomsorten besetzt sein. Dieser Zustand tritt in den meisten Fällen dann auf, wenn die Schmelze rasch abgekühlt wird. Ein bekanntes Beispiel hierfür sind die Kupfer-Gold-Legierungen, die sich in beliebiger Zusammensetzung herstellen lassen und bei denen der ungeordnete Zustand „eingefroren" werden kann. Wie wir in Kap. 8.2 sehen werden, ist der elektrische Widerstand bei tiefen Temperaturen, der sogenannte *Restwiderstand*, ein Maß für die Abweichungen vom idealen Kristall, da die Elektronen auf ihrem Weg durch das Kristallgitter an den Fremdatomen gestreut werden. „Falsch" angeordnete Atome erhöhen daher den Widerstand. In Abb. 4.27 ist der Restwiderstand als Funktion der Goldkonzentration dargestellt. Erwartungsgemäß findet man ein Widerstandsmaximum bei maximaler Unordnung, d.h., wenn die Legierung aus gleichen Teilen Gold und Kupfer besteht. Nur der Vollständigkeit halber sei erwähnt, daß der Widerstand der sogenannten *Nordheim-Regel* folgt. Diese besagt, daß der Restwiderstand proportional zu $x(1-x)$ ist, wobei mit x die Konzentration einer der beiden Komponenten gemeint ist. Die Übereinstimmung zwischen den experimentellen Daten und dieser einfachen Beschreibung ist beeindruckend. Kühlt man die Schmelze genügend langsam ab oder tempert man die Legierung längere Zeit bei einer Temperatur knapp unter der Schmelztemperatur, so bilden sich bei bestimmten Zusammensetzungen geordnete Phasen aus. Dadurch sinkt der Grad der Unordnung, und der Restwiderstand fällt. Dieser Effekt läßt sich besonders schön anhand der Kupfer-Gold-Legierungen demonstrieren, deren Verhalten in Abb. 4.27b gezeigt ist. In diesen Legierungen bilden sich beim Tempern die Metallverbindungen Cu_3Au bzw. CuAu.

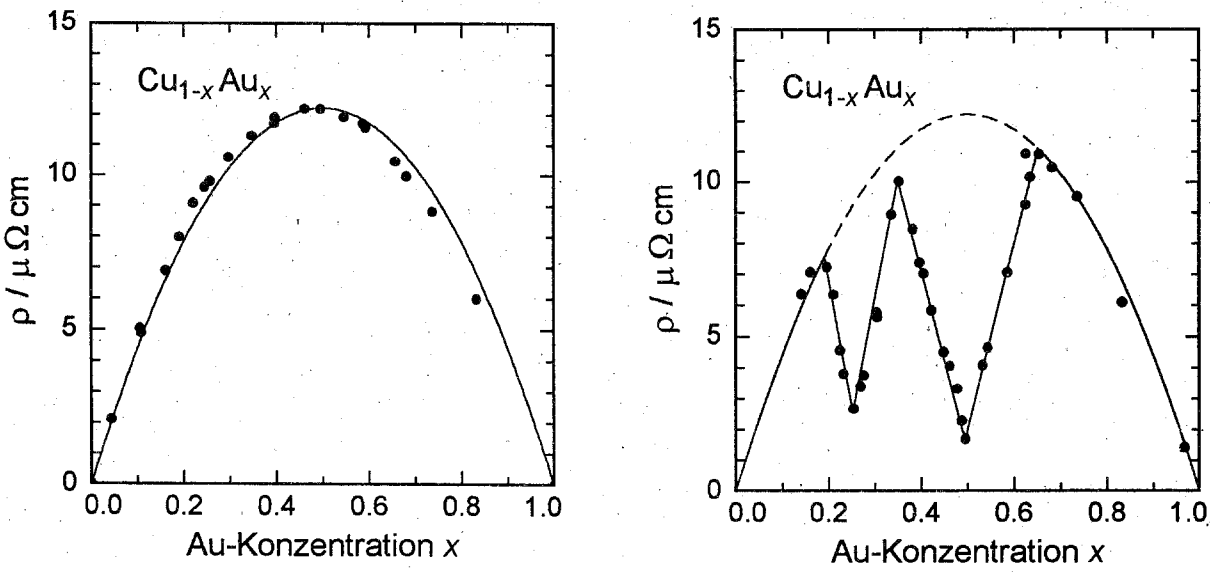


Abb. 4.27: Restwiderstand von Kupfer-Gold-Legierungen. Der Widerstand der reinen Metalle ist abgezogen. Die eingezeichneten Parabeln geben die Vorhersage der Nordheim-Regel wieder. (Nach C.H. Johanson und J.O. Linde, Ann. Phys. **25** (1936) 1).
a) Widerstand von ungeordneten Legierungen.
b) Widerstand von getemperten Proben mit deutlichen Minima bei der Zusammensetzung von Metallverbindungen.

Es stellt sich nun die Frage, wie sich die Struktur beim Übergang von der ungeordneten in die geordnete Phase verändert. Wir greifen als besonders instruktives Beispiel das Verhalten der Legierung CuZn heraus, die aus gleichen Teilen Kupfer und Zink besteht und unter der Bezeichnung „β-Messing" bekannt ist. Bei Zimmertemperatur ist diese Legierung wohlgeordnet, ihr liegt ein kubisch primitives Gitter mit einer zweiatomigen Basis zugrunde. Die Würfelecken werden von Kupfer-, die Würfelzentren von Zinkatomen eingenommen, d.h. es liegt eine CsCl-Struktur vor. Erwärmt man die Probe, so erfolgt bei der kritischen Temperatur von etwa 735 K – der genaue Wert hängt von der exakten Zusammensetzung ab – ein Übergang in eine vollständig ungeordnete Phase. Nun sind die Plätze statistisch besetzt, und die Legierung besitzt im Mittel ein kubisch raumzentriertes Gitter. Dieser Vorgang ist reversibel, d.h. beim Abkühlen tritt wieder Ordnung ein.

Der Übergang von der geordneten in die ungeordnete Phase kann je nach System von 1. oder 2. Ordnung sein. Ist der Phasenübergang von 1. Ordnung, so tritt beim Übergang latente Wärme, die Umwandlungswärme auf. Beim Übergang 2. Ordnung wird ein Maximum in der spezifischen Wärme, aber keine latente Wärme beobachtet. Bei den oben erwähnten Kupfer-Gold-Legierungen liegt beispielsweise ein Ordnungs-Unordnungs-Übergang 1. Ordnung vor, während bei der soeben diskutierten CuZn-Legierung der Übergang von 2. Ordnung ist. Die spezifische Wärme, deren Verlauf am Phasenübergang in Abb. 4.28 gezeigt ist, weist die hierfür typische λ-Form auf.

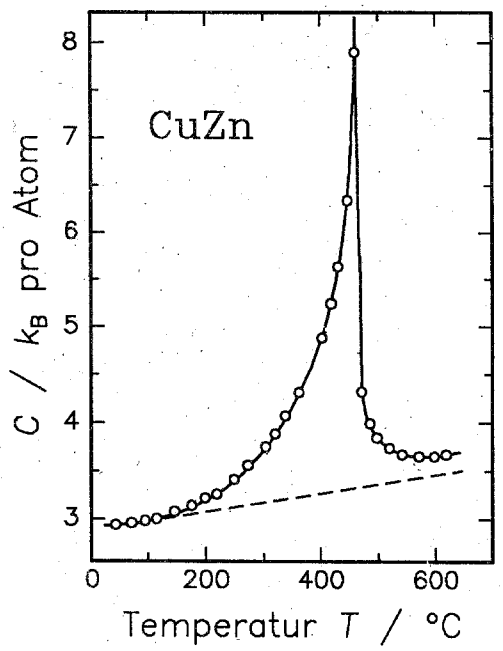


Abb. 4.28: Verlauf der spezifischen Wärme von β-Messing. Das λ-förmige Maximum ist ein charakteristisches Merkmal eines Phasenübergangs 2. Ordnung. (Nach F.C. Nix und W. Shocklex, Rev. Mod. Phys. **10** (1938) 1).

Um den Ordnungsgrad zu charakterisieren, führen wir den Parameter s der Fernordnung ein. Besteht die Probe aus N Atomen und befinden sich R Atome auf „richtigen" und F Atome auf „falschen" Plätzen, so definieren wir

$$s = \frac{2R - N}{N} = \frac{N - 2F}{N} \,. \tag{4.13}$$

Bei vollständiger Ordnung sitzen alle N Atome auf den richtigen Plätzen, d.h. $R = N$ und $s = 1$. Bei vollständiger Unordnung nehmen die Hälfte der Atome falsche Plätze

ein, d.h. $s = 0$. Der schematische Temperaturverlauf dieses Parameters ist für die beiden Phasenübergänge unterschiedlicher Ordnung in Abb. 4.29 gezeigt. Während beim Phasenübergang 2. Ordnung die Fernordnung bei Annäherung an die kritische Temperatur T_c stetig gegen Null geht, tritt beim Übergang 1. Ordnung ein Sprung auf. Der unterschiedliche Verlauf des Ordnungsparameters macht auch klar, warum nur im Falle des Übergangs 1. Ordnung eine Umwandlungswärme zu erwarten ist und beim Übergang 2. Ordnung nicht.

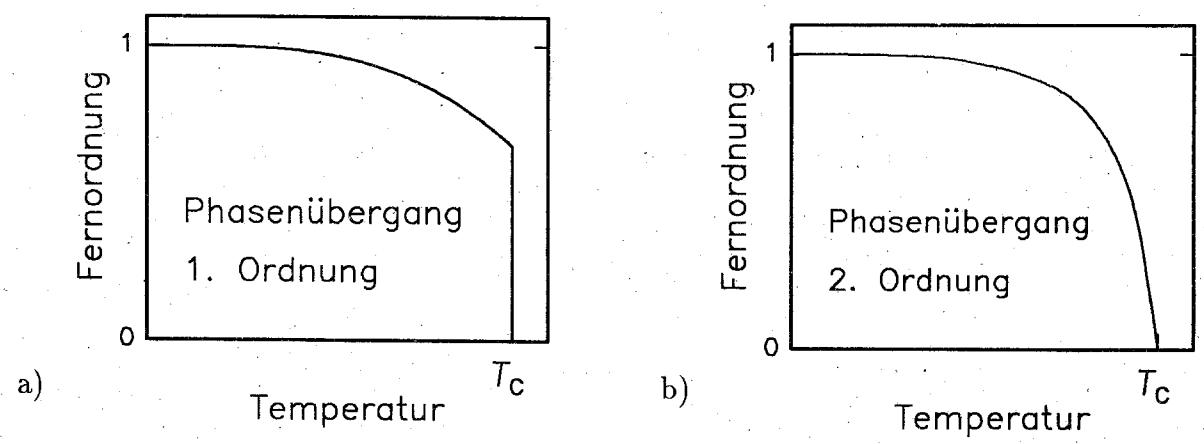


Abb. 4.29: Schematischer Temperaturverlauf des Ordnungsparameters.
a) Phasenübergang 1. Ordnung. Bei T_c macht der Ordnungsparameter einen Sprung.
b) Phasenübergang 2. Ordnung. Der Ordnungsparameter ändert sich stetig.

Mit Hilfe der Röntgenbeugung läßt sich der Ordnungsgrad verfolgen. Wir wollen die Beobachtungen für den Fall der CuZn-Legierung etwas genauer erläutern. Wie bereits erwähnt, liegt in der geordneten Phase des β-Messing ein kubisch primitives Gitter mit zweiatomiger Basis vor. Wie beim Cäsiumchlorid (vgl. Kap. 2.3) ist der Strukturfaktor durch $\mathcal{S}_{hkl} = \widetilde{a}_{\mathrm{Cu}} + \widetilde{a}_{\mathrm{Zn}} \exp[-\mathrm{i}\pi(h + k + l)]$ gegeben. Es treten alle Reflexe auf, wenn auch, wie von (2.24) vorhergesagt, mit unterschiedlicher Intensität. In der ungeordneten Phase wird die Streuintensität durch einen gemittelten Strukturfaktor bestimmt. Mit dem mittleren Atomformfaktor $\langle\widetilde{a}\rangle = (\widetilde{a}_{\mathrm{Cu}} + \widetilde{a}_{\mathrm{Zn}})/2$ können wir in diesem Fall für den Strukturfaktor $\langle\mathcal{S}_{hkl}\rangle = \langle\widetilde{a}\rangle\{1+\exp[-\mathrm{i}\pi(h+k+l)]\}$ schreiben. Die Legierung verhält sich im Mittel wie ein kubisch raumzentrierter Kristall, d.h., wie in (2.25) vorhergesagt, die Reflexe mit der ungeraden Summe $h + k + l$ entfallen. Die zusätzlichen Linien, die man bei Debye-Scherrer-Aufnahmen an der geordneten Phase beobachtet, bezeichnet man als *Überstrukturlinien.*

5. Gitterdynamik

Die meisten typischen Festkörpereigenschaften lassen sich entweder auf die Bewegung der Atome um ihre Gleichgewichtslage oder auf Elektronenbewegungen zurückführen. Im ersten Fall spricht man von den gitterdynamischen, im zweiten von den elektronischen Eigenschaften. Diese Aufteilung ist möglich, weil die typischen Geschwindigkeiten, mit denen sich Elektronen bewegen, viel größer sind als die Geschwindigkeiten der Atomkerne. Werden Atome aus ihren Gleichgewichtslagen verschoben, so stellt sich „instantan" eine neue Elektronenverteilung ein, wobei sich die Gesamtenergie des Systems erhöht. Kehren diese Atome in die Ausgangslage zurück, so wird die aufgewandte Energie vollständig zurückgewonnen. Die Elektronen werden bei diesem Vorgang nicht angeregt, sondern verbleiben unabhängig von den augenblicklichen Kernkoordinaten im Grundzustand. Wie bereits erwähnt, wird die separate Behandlung dieser Untersysteme (siehe auch H_2-Molekül in Kap. 3.4) als *adiabatische* oder *Born-Oppenheimer*-Näherung bezeichnet.

Wir wollen die Dynamik des Gitters in zwei Schritten diskutieren. In Kap. 5 betrachten wir die verschiedenen Schwingungszustände, die in Kristallen und amorphen Festkörpern auftreten können. Dabei nehmen wir an, daß die Auslenkung der Atome aus ihrer Ruhelage mit einer harmonischen Variation der potentiellen Energie des Festkörpers verbunden ist. Diese Näherung erlaubt bereits das Verständnis einer grundlegenden thermodynamische Größe, nämlich der spezifischen Wärme. In amorphen Strukturen treten zusätzliche Bewegungsmöglichkeiten auf, auf die wir im letzten Abschnitt dieses Kapitels eingehen, obwohl ihre theoretische Behandlung bisher nur in Ansätzen verstanden ist. In Kapitel 6 werden wir dann Eigenschaften wie Wärmeleitfähigkeit oder thermische Ausdehnung diskutieren, bei denen die schwache Anharmonizität des Potentials eine wesentliche Rolle spielt.

5.1 Schallausbreitung

In unseren bisherigen Überlegungen gingen wir fast immer davon aus, daß Festkörper starre Gebilde sind, deren Atome sich im Minimum der lokalen Potentialmulden aufhalten. Wir wollen uns jetzt der Auslenkung der Atome aus ihrer Ruhelage zuwenden. Im einfachsten Fall befinden sich zu einem vorgegebenen Zeitpunkt alle Atome im Grundzustand, nur ein einziges wird angestoßen und schwingt mit einer Amplitude entsprechend seiner Anregung um seine Gleichgewichtslage. Wollen wir diesen Vorgang genauer betrachten, so müssen wir die Kopplung der Atome untereinander berücksichtigen. Sie bewirkt, daß zunächst die Nachbaratome mitschwingen und daß sich nach einiger Zeit die Schwingungsenergie über das ganze Gitter verteilt. Die Ausbreitung der Schwingungsenergie kann dabei wie in Gasen oder Flüssigkeiten mit der Geschwindigkeit des Schalls erfolgen.

Die Schallwellen sind eine besonders einfache Form der Gitterschwingungen, ihre Eigenschaften werden mit Hilfe der *elastischen Kontinuumstheorie* beschrieben[1], bei der der Aufbau der Festkörper aus Atomen unberücksichtigt bleibt. Wie man sich leicht vorstellen kann, ist diese Vereinfachung nur dann gerechtfertigt, wenn die Wellenlänge groß gegen die atomaren Abstände ist. Bei sehr kleiner Wellenlänge beobachtet man erhebliche Abweichungen von den vertrauten Eigenschaften des Schalls: Die Gruppengeschwindigkeit und damit die Geschwindigkeit, mit der der Energietransport erfolgt, ist reduziert, und Quantenphänomene treten in den Vordergrund.

Neben den *longitudinalen* Schallwellen, die auch in Gasen und Flüssigkeiten auftreten, findet man in Festkörpern aufgrund ihrer Schersteifigkeit auch *transversale* Schallwellen. Bezeichnen wir die Auslenkung eines herausgegriffenen kleinen Volumenelements des Festkörpers aus der Ruhelage mit $\boldsymbol{u}$ und den Wellenvektor der Schallwelle mit $\boldsymbol{q}$, so gilt $\boldsymbol{u} \parallel \boldsymbol{q}$ für longitudinale und $\boldsymbol{u} \perp \boldsymbol{q}$ für transversale Wellen. In elastisch *isotropen Festkörpern*[2] hängt das Ausbreitungsverhalten der Schallwellen nicht von der Richtung ab. Läuft eine Schallwelle in einem derartigen Festkörper in x-Richtung, dann lauten die Wellengleichungen für die drei möglichen Auslenkungen[3] u_x, u_y und u_z:

$$\rho \frac{\partial^2 u_x}{\partial t^2} = c_{11} \frac{\partial^2 u_x}{\partial x^2} \,, \tag{5.1}$$

$$\rho \frac{\partial^2 u_y}{\partial t^2} = c_{44} \frac{\partial^2 u_y}{\partial x^2} \qquad \text{und} \qquad \rho \frac{\partial^2 u_z}{\partial t^2} = c_{44} \frac{\partial^2 u_z}{\partial x^2} \,, \tag{5.2}$$

wobei ρ für die Dichte der Probe steht. Die Elastizitätsmoduln c_{11} und c_{44} sind makroskopische Parameter[4] zur Beschreibung der Steifigkeit des Materials, auf deren Verknüpfung mit den Gittereigenschaften wir noch eingehen werden. Die Indizes deuten bereits darauf hin, daß im allgemeinen eine größere Anzahl von Konstanten auftritt. So sind im Extremfall, beim triklinen Kristallsystem, 21 unterschiedliche Elastizitätsmoduln zur vollständigen Beschreibung der elastischen Eigenschaften erforderlich. Bei Kristallen mit ihren häufig sehr ausgeprägten anisotropen elastischen Eigenschaften treten daher oft recht unübersichtliche Verhältnisse auf. Isotrope Festkörper, auf die wir uns im Augenblick beschränken, stellen einen besonders einfachen Spezialfall dar. Dort sind nur zwei Moduln zur Charakterisierung der elastischen Eigenschaften erforderlich.

Der Elastizitätsmodul c_{44}, der in die Wellengleichung (5.2) für transversale Wellen eingeht, ist identisch mit dem in Kap. 4.2 benutzten Schub- oder Schermodul $\mathcal{G}$. Der Modul c_{11} in (5.1) für longitudinale Wellen ist eng verwandt mit dem linearen Elastizitätsmodul E, der über die Dehnung eines Stabes definiert ist. Es gilt der Zusammenhang $c_{11} = E\,(1-\mu)/[(1+\mu)(1-2\mu)]$, wobei μ für den Poissonschen Querkontraktionskoeffizienten steht.

---

[1] Eine umfassende und übersichtliche Darstellung der elastischen Eigenschaften fester Körper findet man z.B. in L.D. Landau, E.M. Lifschitz, Lehrbuch der Theoretischen Physik, Band VII, *Elastizitätstheorie*, Verlag Harri Deutsch, Frankfurt.

[2] Elastisch isotrop verhalten sich beispielsweise Gläser, polykristalline Materialien, wenn die regellos angeordneten Kristallite wesentlich kleiner sind als die Wellenlänge des Schalls, oder auch Quasikristalle.

[3] Wir benutzen hier kartesische Koordinaten.

[4] Die Elastizitätsmoduln sind Elemente des Elastizitätstensors, der eigentlich ein Tensor 4. Stufe mit 81 Einträgen ist. Aufgrund von Symmetriebeziehungen reduziert sich die Zahl der unabhängigen Einträge jedoch auf 21. Wir benutzen hier die verkürzte Schreibweise mit zwei Indizes, bei der die Elastizitätsmoduln in einer 6×6 - Matrix untergebracht werden. Wir werden jedoch auf eine systematische Behandlung der elastischen Eigenschaften von Festkörpern verzichten.

Die Form der Wellengleichungen wird verständlich, wenn man sich die Bewegungsgleichung eines kleinen Volumenelements des elastischen Kontiuums überlegt. Auf der linken Seite stehen die Beschleunigungskräfte und auf der rechten die rücktreibenden Kräfte, die durch die innere Spannung hervorgerufen werden. Zusätzlich wird noch die Gültigkeit des Hookeschen Gesetzes vorausgesetzt.

Ebene monochromatische Wellen stellen besonders einfache Lösungen der obigen Wellengleichungen dar. Mit dem Ansatz $u_i = U_i e^{-i(\omega t - qx)}$, wobei der Index i die kartesischen Komponenten der Auslenkung angibt, finden wir durch Einsetzen in die entsprechende Gleichung die einfache *Dispersionsrelation*[5]

$$\omega = vq\,. \tag{5.3}$$

Es besteht ein linearer Zusammenhang zwischen Frequenz und Wellenvektor, d.h. die Schallwellen breiten sich unabhängig von der Frequenz mit der gleichen Geschwindigkeit v aus[6]. In elastischen Kontinua ist die Schallausbreitung also dispersionsfrei. Je nach Polarisation der Schallwelle ist die Ausbreitungsgeschwindigkeit durch

$$v_\mathrm{l} = \sqrt{\frac{c_{11}}{\rho}} \qquad \text{oder} \qquad v_\mathrm{t} = \sqrt{\frac{c_{44}}{\rho}} \tag{5.4}$$

gegeben. Grob gesprochen ist in Festkörpern die Geschwindigkeit v_l der longitudinalen Schallwellen etwa doppelt so groß wie die transversale Schallgeschwindigkeit v_t. Schallgeschwindigkeitmessungen erlauben also die Bestimmung der elastischen Konstanten und machen somit wichtige Aussagen über die elastischen bzw. mechanischen Eigenschaften der Festkörper.

Ebene Schallwellen lassen sich in Festkörpern mit relativ einfachen Versuchsaufbauten erzeugen. Wie in der Prinzipskizze 5.1 zu sehen ist, werden zwei piezoelektrische Schallwandler an den Probenenden angebracht. Mit einer sehr dünnen Schicht einer zähen Flüssigkeit wird der akustische Kontakt zwischen dem Schallgeber und der Probe hergestellt.

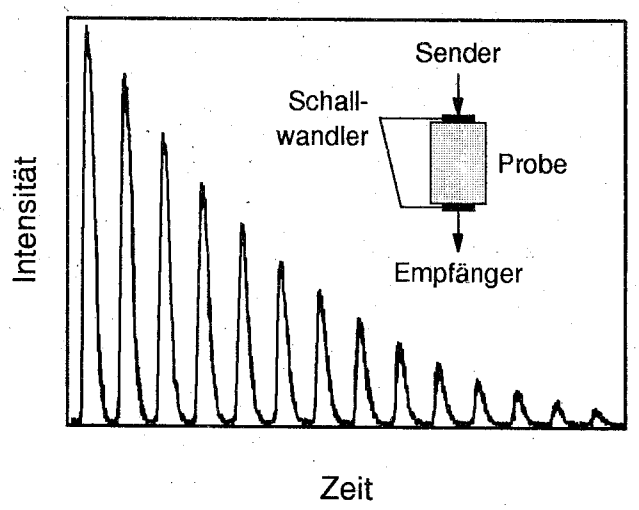


Abb. 5.1: Ultraschallechos in einer Glasprobe. In der Oszilloskopaufnahme sind nur die Einhüllenden der Schallpulse und nicht die hochfrequenten 1 GHz-Schwingungen zu sehen. Rechts oben ist das Schema eines Meßaufbaus eingezeichnet.

[5] Unter der Dispersionsrelation verstehen wir den Zusammenhang zwischen der Frequenz der Welle und ihrer Wellenzahl bzw. ihrem Wellenvektor.

[6] Wir unterscheiden nicht zwischen der Phasengeschwindigkeit $v_\mathrm{p} = \omega/q$ und der Gruppengeschwindigkeit $v_\mathrm{g} = \partial\omega/\partial q$, da sie hier identisch sind.

Der Sendeschallwandler[7] wird mit einem Hochfrequenzimpuls von beispielsweise 1 GHz und einer Dauer von 1 μs beaufschlagt und so kurzfristig zu Schwingungen angeregt. Dadurch wird ein Schallimpuls erzeugt, der bei Proben mit planparallelen Stirnflächen wiederholt in der Probe hin und her läuft und bei jeder Reflexion an der Empfangsseite registriert werden kann. Bei Proben von 1 cm Länge liegt der zeitliche Abstand zwischen zwei Echos typischerweise bei etwa 5 μs. Bei bekannter Probenlänge läßt sich aus der zeitlichen Folge der Echos die Schallgeschwindigkeit bestimmen. Die beobachtete Signalamplitude nimmt aufgrund der Ultraschalldämpfung mit der Laufstrecke x bzw. mit der Zeit ab. Für den Verlauf der Schallintensität findet man den Zusammenhang $I = I_0 \exp(-x/l) = I_0 \exp(-vt/l)$, wobei l für die *freie Weglänge*[8] und I_0 für die Intensität zum Zeitpunkt $t = 0$ steht. Einige Mechanismen, die zur Dämpfung von Schallwellen führen, werden wir im nächsten Kapitel besprechen.

Durch geeignete Wahl der Orientierung der piezoelektrischen Schallwandler lassen sich longitudinale oder transversale Wellen erzeugen und nachweisen und so die beiden Konstanten c_{11} und c_{44} getrennt bestimmen. Bei anisotropen Kristallen muß die Schallausbreitung in verschiedenen Richtungen und mit unterschiedlicher Polarisation gemessen werden, wenn alle elastischen Konstanten mit dieser Technik bestimmt werden sollen.

5.2 Bewegungsgleichung der Gitteratome

Wir wollen nun die Näherung des elastischen Kontinuums verlassen, die Dynamik der Kristalle etwas genauer betrachten und ihre Eigenschwingungen berechnen. Im ersten Schritt stellen wir die Bewegungsgleichung für die einzelnen Atome auf, wobei wir besonders auf die Kopplung zwischen den Atomen achten. Dabei treten leider viele Indizes zur eindeutigen Kennzeichnung der Lage der einzelnen Atome und deren Auslenkungen auf, wodurch die an sich einfachen Gleichungen etwas unübersichtlich werden. Im zweiten Schritt wird die Translationssymmetrie des Gitters ausgenutzt, um zu einer wesentlich einfacheren Struktur der Gleichungen zu kommen. Schließlich machen wir wie bei der einführenden Diskussion der Schallwellen von der Tatsache Gebrauch, daß ebene Wellen besonders einfache Lösungen der Bewegungsgleichung darstellen.

Die potentielle Energie der Atome eines Festkörpers ist aufgrund ihrer Kopplung untereinander eine Funktion der augenblicklichen Lage *aller* Atome. Wir führen daher die Potentialfunktion ϕ ein, die für die potentielle Energie des gesamten Festkörpers steht. Für die meisten Fragestellungen reicht es, diese Funktion in einer Taylor-Reihe bis zur zweiten Ordnung der atomaren Auslenkungen aller Atome zu entwickeln:

$$\begin{aligned}\phi &= \phi_0 + \sum_{\boldsymbol{R}\alpha i} \left.\frac{\partial \phi}{\partial u_{\boldsymbol{R}\alpha i}}\right|_0 u_{\boldsymbol{R}\alpha i} + \frac{1}{2} \sum_{\boldsymbol{R}\alpha i} \sum_{\boldsymbol{R}'\alpha' i'} \left.\frac{\partial^2 \phi}{\partial u_{\boldsymbol{R}\alpha i}\, \partial u_{\boldsymbol{R}'\alpha' i'}}\right|_0 u_{\boldsymbol{R}\alpha i}\, u_{\boldsymbol{R}'\alpha' i'} + \cdots \\ &\approx \frac{1}{2} \sum_{\boldsymbol{R}\alpha i} \sum_{\boldsymbol{R}'\alpha' i'} \left.\frac{\partial^2 \phi}{\partial u_{\boldsymbol{R}\alpha i}\, \partial u_{\boldsymbol{R}'\alpha' i'}}\right|_0 u_{\boldsymbol{R}\alpha i}\, u_{\boldsymbol{R}'\alpha' i'} \,. \end{aligned} \tag{5.5}$$

[7] In vielen Experimenten wird der Sendeschallwandler gleichzeitig auch als Empfangsschallwandler benutzt und elektronisch zwischen den beiden Betriebsarten geschaltet.

[8] Meist steht anstelle der freien Weglänge in den entsprechenden Gleichungen der Dämpfungskoeffizient α. Bei geeigneter Wahl der Einheiten gilt die Beziehung $\alpha \equiv l^{-1}$.

Die Größe $u_{\boldsymbol{R}\alpha i}$ steht für die kartesische Komponente i der Auslenkung des Atoms α in der Elementarzelle, die durch den Gittervektor $\boldsymbol{R}$ festgelegt ist. Die Konstante ϕ_0 der potentiellen Energie ist für die weitere Diskussion ohne Bedeutung, und wir dürfen zur Vereinfachung $\phi_0 = 0$ setzen. Der lineare Term verschwindet ebenfalls, da im Gleichgewicht keine Kräfte auf die Atome wirken. Somit bleibt in der von uns benutzten Näherung nur der letzte Term bzw. die zweite Zeile übrig.

Die Ähnlichkeit mit dem Potential eines harmonischen Oszillators ist nicht zu übersehen. Bewegt sich ein Teilchen in dem eindimensionalen harmonischen Potential der Form $\widetilde{\phi} = \widetilde{C}x^2/2$, so wirkt auf dieses Teilchen die rücktreibende Kraft $\widetilde{F} = -\widetilde{C}x$, wobei die Kraftkonstante durch $\widetilde{C} = \mathrm{d}^2\widetilde{\phi}/\mathrm{d}x^2$ gegeben ist. Anhand dieses sehr einfachen Beispiels kann man sich die Bedeutung von (5.5) und das weitere Vorgehen bei der Aufstellung der Bewegungsgleichung der Atome klar machen. Zunächst definieren wir die verallgemeinerten Kraftkonstanten $C_{\boldsymbol{R}\alpha i,\boldsymbol{R}'\alpha' i'}$ mit Hilfe der Gleichung

$$C_{\boldsymbol{R}\alpha i,\boldsymbol{R}'\alpha' i'} \equiv \frac{\partial^2\phi}{\partial u_{\boldsymbol{R}\alpha i}\,\partial u_{\boldsymbol{R}'\alpha' i'}} \,. \tag{5.6}$$

Die Konstanten lassen sich zu einem kartesischen Tensor, dem Kraftkonstantentensor $[\boldsymbol{C}]$, zusammenfassen. Um in den folgenden Gleichungen die Vektorschreibweise benutzen zu können, fassen wir die drei Raumrichtungen, die mit den Indizes i bzw. i' gekennzeichnet sind, zusammen und benutzen für die Komponenten des Kraftkonstantentensors die verkürzte Schreibweise $\boldsymbol{C}_{\boldsymbol{R}\alpha,\boldsymbol{R}'\alpha'}$.

Die Kraft, die auf das Atom α in der Elementarzelle mit dem Gittervektor $\boldsymbol{R}$ wirkt, wenn das Atom α' mit dem Gittervektor $\boldsymbol{R}'$ die Auslenkung $\boldsymbol{u}_{\boldsymbol{R}'\alpha'}$ aufweist, läßt sich durch die Beziehung

$$\boldsymbol{F}_{\boldsymbol{R}\alpha} = -\boldsymbol{C}_{\boldsymbol{R}\alpha,\boldsymbol{R}'\alpha'} \cdot \boldsymbol{u}_{\boldsymbol{R}'\alpha'} \tag{5.7}$$

darstellen. Trotz der formalen Ähnlichkeit mit der einfachen Beziehung $\widetilde{F} = -\widetilde{C}x$ bedarf diese Gleichung der Erläuterung. Sie besagt, daß auf das herausgegriffene Atom eine Kraft wirkt, wenn das Atom α' in der Zelle $\boldsymbol{R}'$ ausgelenkt wird. Es wird also grundsätzlich die Kopplung zwischen allen Atomen berücksichtigt. Wird das Aufatom selbst ausgelenkt, so gilt ebenfalls (5.7), wobei $\boldsymbol{R}' = \boldsymbol{R}$ und $\alpha' = \alpha$ gesetzt wird.

Die Translationssymmetrie der Kristalle erlaubt eine wesentliche Vereinfachung des Kraftkonstantentensors. Da keine Elementarzelle ausgezeichnet ist, kann der Wert des Tensors $[\boldsymbol{C}]$ nicht von der absoluten Lage der betrachteten Zelle abhängen. Die Komponenten des Tensors sind also lediglich eine Funktion der relativen Position $\boldsymbol{h} = \boldsymbol{R}' - \boldsymbol{R}$ der wechselwirkenden Atome. Es gilt daher:

$$\boldsymbol{C}_{\boldsymbol{R}\alpha,\boldsymbol{R}'\alpha'} = \boldsymbol{C}_{\boldsymbol{h}\alpha,\alpha'} \,. \tag{5.8}$$

Da die Kopplung mit dem Abstand rasch abnimmt, reicht es oft, nur die nächsten Nachbarn in die Rechnung einzubeziehen. Anschaulich gesprochen ist dann jedes Atom nur noch über Federn mit den Nachbaratomen verbunden. Darüber hinaus ist bei Berechnungen wichtig, daß die Zahl der unabhängigen Komponenten des Kraftkonstantentensors von der Kristallsymmetrie abhängt. Die Berücksichtigung vorhandener Symmetrien führt daher oft zu erheblichen Vereinfachungen des Problems.

Mit (5.7) und (5.8) läßt sich die auf ein Atom wirkende Gesamtkraft ausdrücken, und wir können die Bewegungsgleichung des Atoms α mit der Masse M_α in der Elementarzelle mit dem Gittervektor $\boldsymbol{R}$ angeben:

$$M_\alpha \, \ddot{\boldsymbol{u}}_{\boldsymbol{R}\alpha} = -\sum_{\boldsymbol{h}\,\alpha'} \boldsymbol{C}_{\boldsymbol{h}\,\alpha,\alpha'} \cdot \boldsymbol{u}_{\boldsymbol{R}+\boldsymbol{h},\alpha'} \,. \tag{5.9}$$

Sind $\mathcal{N}$ Elementarzellen mit je p Atomen vorhanden, so ergeben sich $3p\mathcal{N} = 3N$ gekoppelte Differentialgleichungen, wobei N wieder für die Gesamtzahl der Atome der Probe steht. Eine Entkopplung in unabhängige Gleichungen läßt sich auf einfache Weise durch einen Lösungsansatz mit ebenen Wellen erreichen. In diesem Ansatz lassen wir unterschiedliche Amplituden für die verschiedenen Atome innerhalb einer Elementarzelle zu[9]. Wir schreiben also

$$\boldsymbol{u}_{\boldsymbol{R}\alpha}(t) = \boldsymbol{U}_\alpha \, \mathrm{e}^{-\mathrm{i}[\omega t - \boldsymbol{q}\cdot\boldsymbol{R}]} \,, \tag{5.10}$$

wobei $\boldsymbol{U}_\alpha$ die Amplitude des Atoms α in der Zelle mit den Gittervektoren $\boldsymbol{R} = 0$ ist. $\boldsymbol{U}_\alpha$ enthält einen Phasenfaktor, den wir im Lösungsansatz nicht besonders hervorheben, der aber vom Platz des Atoms α in der herausgegriffenen Elementarzelle abhängt. Setzen wir diesen Ansatz in (5.9) ein, so erhalten wir

$$\begin{aligned} \omega^2 M_\alpha \boldsymbol{U}_\alpha &= \sum_{\alpha'} \left[\sum_{\boldsymbol{h}} \boldsymbol{C}_{\boldsymbol{h}\,\alpha,\alpha'} \, \mathrm{e}^{\mathrm{i}\boldsymbol{q}\cdot\boldsymbol{h}} \right] \boldsymbol{U}_{\alpha'} \\ &= \sum_{\alpha'} \boldsymbol{D}_{\alpha,\alpha'}(\boldsymbol{q}) \, \boldsymbol{U}_{\alpha'} \,. \end{aligned} \tag{5.11}$$

Der Ausdruck in der Klammer ist die Fourier-Transformierte des Kraftkonstantentensors $[\boldsymbol{C}]$, für die wir in der zweiten Zeile die Abkürzung

$$\boldsymbol{D}_{\alpha,\alpha'}(\boldsymbol{q}) \equiv \sum_{\boldsymbol{h}} \boldsymbol{C}_{\boldsymbol{h}\,\alpha,\alpha'} \, \mathrm{e}^{\mathrm{i}\boldsymbol{q}\cdot\boldsymbol{h}} \tag{5.12}$$

geschrieben haben. $\boldsymbol{D}_{\alpha,\alpha'}$ steht für die Einträge der *dynamischen Matrix* $[\boldsymbol{D}]$. Sie enthält die gesamte Information, die zur Beschreibung der elastischen Eigenschaften eines Kristalls erforderlich ist. Das lineare Gleichungssystem (5.11) ist homogen und nur noch von der Ordnung $3p$. Für einen Kristall mit einatomiger Basis ist $p = 1$, so daß für jeden Wellenvektor $\boldsymbol{q}$ nur noch ein System von drei Gleichungen für die drei Raumrichtungen zu lösen ist. Die resultierenden drei Lösungen beschreiben Wellen mit unterschiedlicher Polarisationsrichtung.

Die Eigenwerte lassen sich wie in der gewöhnlichen Schwingungstheorie finden: Das Gleichungssystem (5.11) hat nur dann nicht-triviale Lösungen, wenn die Koeffizientendeterminante verschwindet, also

$$\det \left[\boldsymbol{D}_{\alpha,\alpha'}(\boldsymbol{q}) - \omega^2 M_\alpha \boldsymbol{I}_{\alpha,\alpha'} \right] = 0 \tag{5.13}$$

ist. Hierbei steht $\boldsymbol{I}_{\alpha,\alpha'}$ für die Einträge des Einheitstensor $[\boldsymbol{I}]$. Für jeden vorgegebenen Wellenvektor $\boldsymbol{q}$ findet man $3p$ Eigenfrequenzen $\omega_{\boldsymbol{q},j}$. Der Index j dient der Unterscheidung

[9] Dieser Ansatz ist äquivalent zu einer Fourier-Entwicklung der Atomauslenkungen und stellt einen Spezialfall des *Bloch-Theorems* dar, mit dem wir uns in Kap. 7.3 noch eingehender beschäftigen.

der $3p$ Eigenfrequenzen[10]. Wie bereits bei der Diskussion der Schallwellen erwähnt, wird der Zusammenhang zwischen den Eigenfrequenzen $\omega_{\boldsymbol{q},j}$ und den Wellenvektoren $\boldsymbol{q}$ als *Dispersionsrelation*[11] bezeichnet. Sind die Lösungen für $\omega_{\boldsymbol{q},j}$ bekannt, lassen sich mit (5.11) unmittelbar die Amplituden $\boldsymbol{U}_\alpha(\boldsymbol{q},j)$ berechnen.

Die entscheidenden Schritte bei der Herleitung dieses überaus wichtigen Ergebnisses waren die Ausnutzung der Translationssymmetrie des Gitters und die anschließende Entkopplung des Gleichungsystems (5.9). Ein Blick auf (5.13) zeigt, daß die Gleichungen die Eigenschaften von harmonischen Oszillatoren widerspiegeln. Jede Normalschwingung, gekennzeichnet durch $\omega_{\boldsymbol{q},j}$ und die dazugehörenden Auslenkungen $\boldsymbol{U}_\alpha(\boldsymbol{q},j)$, entspricht formal einem harmonischen Oszillator. Wie wir im Laufe dieses Kapitels noch sehen werden, besitzen diese harmonischen Oszillatoren erwartungsgemäß quantenmechanische Eigenschaften.

5.3 Gitterschwingungen

Das Konzept zur Berechnung der Gitterschwingungen wollen wir nun auf einfache Beispiele anwenden. Ein wichtiges Ergebnis wird sein, daß erhebliche Abweichungen vom bekannten Verhalten der Schallwellen auftreten, wenn die Wellenlängen vergleichbar mit den Abständen im Gitter werden. Darüber hinaus wird sich zeigen, daß reziproke Gittervektoren in der Gitterdynamik eine wesentliche Rolle spielen.

Wir werden die Dynamik von Kristallen mit ein- und zweiatomiger Basis gesondert untersuchen, da diese beiden Kristalltypen ein qualitativ unterschiedliches Verhalten aufweisen. Eine Erweiterung der Überlegungen auf Systeme mit größeren Basiseinheiten ist dann ohne größere Probleme möglich, erfordert aber umfangreiche Rechnungen. Um die mathematische Handhabung so einfach wie möglich zu gestalten, machen wir eine sehr weitgehende Vereinfachung: Wir betrachten einen „eindimensionalen Kristall", d.h. eine lineare Kette. Im ersten Augenblick könnte man glauben, daß die so hergeleiteten Resultate nur eine stark begrenzte Aussagekraft für dreidimensionale Probleme haben. Wie wir sehen werden, kommen die Ergebnisse der Realität aber durchaus nahe.

Wählen wir in einem realen Kristall als Ausbreitungsrichtung einer ebenen Welle eine Richtung mit hoher Symmetrie, so kompensieren sich aus Symmetriegründen alle Kräfte, die nicht in Auslenkungsrichtung wirken. Dies kann man sich leicht anhand der Abb. 5.2 klarmachen, in der ein Schnitt längs der Grundfläche, d.h. längs einer (001)-Ebene, eines orthorhombischen Kristalls gezeichnet ist. Breitet sich eine longitudinale Welle in [100]-Richtung aus, so wirken auf die Atome nur Kräfte längs dieser Richtung. Bei kubischen Kristallen beispielsweise gilt diese Vereinfachung auch für die ⟨110⟩- und ⟨111⟩-Richtung.

Bei den eingezeichneten Auslenkungen der Gitterebenen üben zum Beispiel die Atome $1, 2', 2'', \ldots$ der Ebene $(s-2)$ Kräfte auf das Atom A aus, die mit zunehmendem Abstand rasch abnehmen. Die resultierende Kraft weist in Richtung der Auslenkung, da sich die Kräfte senkrecht zur Auslenkungsrichtung aus Symmetriegründen exakt kompensieren. Entsprechende Argumente lassen sich auch bei Ausbreitung von transversalen Wellen in Richtungen hoher Symmetrie finden. Die Auslenkung der Atome besitzt daher in diesen

[10] Die Indizes $\boldsymbol{q}$ und j benutzen wir vor allem dann, wenn sonst Verwechslungen möglich sind.

[11] Die Dispersionsrelation gewöhnlicher Schallwellen ist durch Gleichung (5.3) gegeben, die wir in Kap. 5.1 kennengelernt haben.

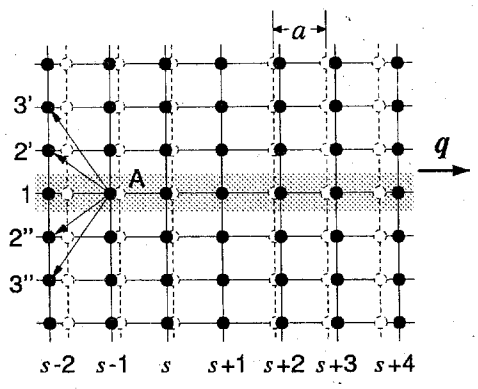


Abb. 5.2: Wellenausbreitung in einem orthorhombischen Kristall. Atomauslenkung und Wellenausbreitung erfolgt in [100]-Richtung. Die augenblickliche Lage der Atome ist durch Punkte, ihre Ruheposition durch offene Kreise wiedergegeben. Die Pfeile deuten die Richtung der Kräfte an, die auf das Atom A durch die Auslenkung der Ebene $(s-2)$ wirken. Die herausgegriffene lineare Kette ist dunkel getönt.

bevorzugten Richtungen einen rein longitudinalen oder rein transversalen Charakter, wodurch sich die mathematische Behandlung auf ein eindimensionales Problem reduziert. Es reicht aus, eine *lineare Kette* herauszugreifen. In die Bewegungsgleichung gehen aber *effektive Kräfte* ein, zu denen nicht nur die Atome auf der Kette, sondern auch die übrigen Nachbaratome beitragen, von denen jeweils die Komponente in Richtung der Auslenkung wirksam ist.

Erfolgt die Wellenausbreitung in eine beliebige Richtung, so entfallen die Symmetrieargumente. Auf die Atome wirken dann Kräfte in Raumrichtungen, die nicht notwendigerweise mit der Ausbreitungsrichtung zusammenfallen. Die Atome beschreiben nun Bahnkurven, die nicht mehr durch die Bewegung einer linearen Kette genähert werden können. Die Wellen besitzen dann eine *gemischte Polarisation* und werden je nachdem, welche Auslenkung überwiegt, als *quasi-longitudinal* oder *quasi-transversal* bezeichnet.

Um die Eigenwerte und die dazu gehörenden Amplituden zu ermitteln, gehen wir von (5.13) aus. Für einen „linearen Kristall" mit zweiatomiger Basis existieren $3p = 6$ Gleichungen. Da die Atome nur in eine Richtung ausgelenkt werden und die effektiven Kräfte nur in diese Richtung wirken, sind die Bewegungen in den beiden übrigen Raumrichtungen entkoppelt und können jeweils separat behandelt werden. Es sind daher nur zwei Gleichungen für unsere Fragestellung von Bedeutung, und (5.13) reduziert sich zur einfachen Determinantengleichung

$$\begin{vmatrix} D_{11} - \omega^2 M_1 & D_{12} \\ D_{21} & D_{22} - \omega^2 M_2 \end{vmatrix} = 0 \,. \tag{5.14}$$

Wie bereits erwähnt, steckt die erforderliche Information über die elastischen Kristalleigenschaften in der dynamischen Matrix $[\boldsymbol{D}]$. Um eine übersichtliche Behandlung der grundlegenden Fragestellungen zu ermöglichen, werden wir die Matrix noch weiter vereinfachen. Die Ergebnisse werden trotz der Vereinfachungen die wesentlichen dynamischen Eigenschaften der Kristalle widerspiegeln.

Gitter mit einatomiger Basis

Als einfachstes Beispiel wählen wir eine lineare Kette mit *einatomiger Basis*, für die sich (5.14) weiter vereinfacht. Bei einer einatomigen Basis erübrigt sich die Indizierung der Atome, und wir erhalten unter Berücksichtigung von (5.12) die einfach aussehende Gleichung

$$\omega^2 M = D = \sum_{\boldsymbol{h}} C_{\boldsymbol{h}} \, \mathrm{e}^{\mathrm{i}\boldsymbol{q}\cdot\boldsymbol{h}} = \sum_{s} C_s \, \mathrm{e}^{\mathrm{i}qsa} \,. \tag{5.15}$$

Hierbei wurde im letzten Ausdruck der Abstand der Gitterebenen bzw. der Atomabstand in der linearen Kette mit a bezeichnet und für den Betrag des Gittervektors $|\boldsymbol{h}| = sa$ geschrieben, wobei s eine ganze Zahl ist. An dieser Stelle müssen wir noch auf die Bedeutung der einzelnen Glieder in der Summe eingehen, denn sie spiegeln die Kräfte wider, die auf das Bezugsatom $s = 0$ wirken. Für die weitere Diskussion stellen wir uns vor, daß alle Atome der Kette durch Federn miteinander verbunden sind[12]. Die Kraft zwischen zwei herausgegriffenen Atomen hängt von deren Abstand und somit von deren Relativbewegung ab. Wir spalten nun diese Bewegung und damit auch die wirkenden Kräfte in zwei Teile auf: in die Auslenkung des Bezugsatoms und in die Auslenkung der übrigen Atome. Wird das Bezugsatom selbst ausgelenkt, so verändert sich der Abstand zu allen anderen Atomen, und es werden alle Federn beansprucht, die es mit den übrigen Atomen verbinden. Es tritt eine rücktreibende Kraft auf, die von der Wechselwirkung mit allen anderen Atomen herrührt. Die Kraftkonstante C_0, die die auftretenden Kräfte beschreibt, erhalten wir durch Aufsummieren aller einzelnen Beiträge, also durch die Summation über alle Federkonstanten C_s. Somit gilt

$$C_0 = \sum_{s\neq 0} C_s \,. \tag{5.16}$$

Wird ein beliebiges Atom s in die gleiche Richtung wie das vorhin betrachtete Bezugsatom ausgelenkt, so wird nur die Feder beansprucht, die das Atom s mit dem Bezugsatom verbindet. Die Kraft wirkt in Richtung der Auslenkung des Atoms s. Diese Kraft ist also nicht rücktreibend, sondern versucht, das Aufatom auszulenken. Benutzen wir die übliche Definition für das Vorzeichen der Kraftkonstanten, so müssen wir in (5.15) für diesen Beitrag $-C_s \exp(iqsa)$ schreiben. Fassen wir alle auftretenden Kraftkonstanten zusammen, so nimmt (5.15) die Form

$$\omega^2 M = C_0 - \sum_{s\neq 0} C_s \mathrm{e}^{iqsa} = \sum_{s\neq 0} C_s \left(1 - \mathrm{e}^{iqsa}\right) \tag{5.17}$$

an. Berücksichtigen wir noch weiter, daß aufgrund der Translationssymmetrie des Gitters $C_s = C_{-s}$ sein muß, so läßt sich die Summation auf positive Werte von s beschränken:

$$\omega^2 M = \sum_{s>0} C_s \left(2 - \mathrm{e}^{iqsa} - \mathrm{e}^{-iqsa}\right) = 2 \sum_{s>0} C_s \left[1 - \cos\left(qsa\right)\right] \,. \tag{5.18}$$

Die Zwei in der ersten Summe tritt auf, weil die Einschränkung des Summationsindex s auf positive Werte auch in (5.16) bei der Berechnung von C_0 berücksichtigt werden muß.

In den meisten Festkörpern fällt die Wechselwirkung zwischen den Atomen so rasch ab, daß die Summation in (5.18) nach dem nächsten oder übernächsten Nachbarn abgebrochen werden kann, ohne daß dadurch ein größerer Fehler verursacht wird. Im ersten Fall tritt nur die Federkonstante C_1 auf, und wir finden

$$\omega^2 = \frac{2C_1}{M}(1 - \cos\, qa) = \frac{4C_1}{M} \sin^2\left(\frac{qa}{2}\right) \,,$$

$$\omega = 2\sqrt{\frac{C_1}{M}} \left|\sin\left(\frac{qa}{2}\right)\right| \,. \tag{5.19}$$

[12] Hierbei ist zu beachten, daß im allgemeinen Fall jedes Atom mit jedem anderen verbunden ist, wobei die Federn mit dem Abstand der verbundenen Atome rasch schwächer werden.

Berücksichtigen wir auch die Wechselwirkung mit dem übernächsten Nachbarn durch die Federkonstante C_2, so ergibt sich

$$\omega^2 = \frac{4C_1}{M}\left[\sin^2\left(\frac{qa}{2}\right) + \frac{C_2}{C_1}\sin^2(qa)\right] . \tag{5.20}$$

Die entsprechenden Dispersionskurven sind in Abb. 5.3 dargestellt. Im linken Bild finden wir den Verlauf, den man erhält, wenn nur die unmittelbaren Nachbarn wechselwirken. Diese Näherung ist, wie bereits die Berechnung der Bindungsenergie in Kap. 3.2 zeigte, gut erfüllt bei Molekül- und Edelgaskristallen. Spielt dagegen auch die Wechselwirkung mit den übernächsten Nachbarn eine wichtige Rolle, so beobachtet man oft eine Dispersionskurve, wie sie auf der rechten Seite der Abb. 5.3 skizziert ist.

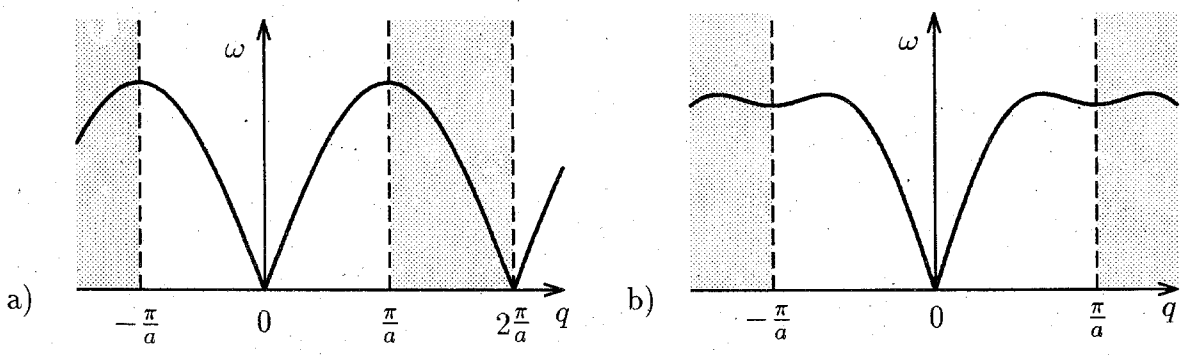


Abb. 5.3: Dispersionskurven einer linearen Kette.
a) Die Wechselwirkung beschränkt sich auf die nächsten Nachbarn.
b) Die übernächsten Nachbarn tragen bei. In der Skizze ist für die Kraftkonstante C_2 der Wert $C_2 = C_1/2$ gewählt worden.

Obwohl die soeben durchgeführte Rechnung für longitudinale und transversale Wellen gültig ist, unterscheiden sich die entsprechenden Dispersionskurven in realen Materialien oft beträchtlich. So findet man für die longitudinalen Schwingungen von Metallen oft Dispersionskurven, die stark von der Kraftkonstanten C_2 mitbestimmt wird. Transversale Schwingungen besitzen dagegen auch in Metallen meist eine Dispersion, die Abb. 5.3a sehr ähnlich sieht. Dies liegt vor allem daran, daß die Wechselwirkung mit den übernächsten Nachbarn bei Scherbewegungen kaum eine Rolle spielt.

Um die Gitterschwingungen besser zu verstehen, wollen wir uns noch mit einigen Besonderheiten vertraut machen. Zunächst sehen wir uns die Phase der Schwingung benachbarter Atome an. Schreiben wir wie in (5.15) $|\boldsymbol{h}| = sa$, so ergibt sich aus (5.10) für die lineare Kette

$$\frac{u_{s+1}}{u_s} = \frac{U\mathrm{e}^{-\mathrm{i}\omega t}\,\mathrm{e}^{\mathrm{i}q(s+1)a}}{U\mathrm{e}^{-\mathrm{i}\omega t}\,\mathrm{e}^{\mathrm{i}qsa}} = \mathrm{e}^{\mathrm{i}qa} . \tag{5.21}$$

Da der Phasenunterschied zwischen benachbarten Gitterpunkten nur zwischen 0 und 2π liegen kann, läßt sich der Wellenvektor ohne Verlust der Allgemeingültigkeit durch die Beziehung

$$-\pi < qa \leq \pi \qquad \text{oder} \quad . \qquad -\frac{\pi}{a} < q \leq \frac{\pi}{a} \tag{5.22}$$

einschränken. Das ist gerade der Bereich der 1. Brillouin-Zone des linearen Gitters oder eines primitiven Gitters in Richtung der Gitterachsen. Liegt ein Wellenvektor q' außerhalb

der 1. Brillouin-Zone, so läßt er sich durch Addition eines reziproken Gittervektors $2\pi p/a$ wieder in diese zurückführen:

$$q' = q \pm \frac{2\pi p}{a} \qquad (p = \text{ganze Zahl}) \,. \tag{5.23}$$

Man bezeichnet diese Vorgehensweise als „Reduktion auf die 1. Brillouin-Zone". Die Gültigkeit dieser Beziehung läßt sich leicht durch Einsetzen von (5.23) in (5.21) zeigen. Für die Phase der Auslenkung benachbarter Gitterpunkte ergeben sich folgende Beziehungen: $u_{s+1}/u_s = \exp(\mathrm{i}q'a) = \exp(\mathrm{i}qa) \cdot \exp(\pm 2\pi \mathrm{i}p) = \exp(\mathrm{i}qa) \cdot 1$. Gleichung (5.23) spiegelt die Tatsache wider, daß der Wellenvektor einer Schwingung von gekoppelten Massepunkten durch deren momentane Position definiert ist. Es ist sinnlos, eine Welle zwischen den Atomen zu konstruieren. Diese Aussage wird in Abb. 5.4 veranschaulicht, in dem eine vorgegebene Konfiguration von Punkten in der Mitte der Bilder durch eine Welle mit zwei Wellenlängen, nämlich $\lambda < 2a$ und $\lambda > 2a$ beschrieben wird.

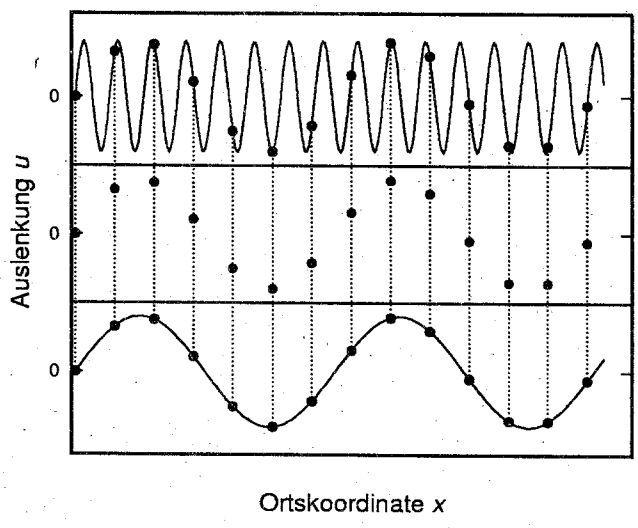


Abb. 5.4: Periodische Auslenkung einer linearen Atomkonfiguration. Die augenblickliche Lage der Atome (Mitte) läßt sich formal durch eine Welle mit $\lambda < 2a$ (oben) oder aber mit $\lambda > 2a$ (unten) beschreiben. Die Wellenzahlen der beiden eingezeichneten Wellen unterscheiden sich gerade um $2\pi/a$, d.h. um einen reziproken Gittervektor.

Wichtig ist in diesem Zusammenhang, daß sich bei der Addition des reziproken Gittervektors die Ausbreitungsrichtung der Welle scheinbar ändert: Eine Welle, die in Richtung der positiven x-Achse läuft, und deren Wellenvektor in der 2. Brillouin-Zone liegt, besitzt nach der Subtraktion des reziproken Gittervektors[13] $g = 2\pi/a$ einen negativen Wert. Dies bedeutet, daß die Phasengeschwindigkeit ihr Vorzeichen wechselt. Wie man sich leicht anhand einer Skizze überzeugen kann, wird die Gruppengeschwindigkeit und damit der Energietransport davon aber nicht beeinflußt. Tatsächlich sind beide Beschreibungen der atomaren Auslenkungen äquivalent, die Ausbreitungsrichtung der Welle läßt sich anhand der atomaren Bewegungen nicht unterscheiden. Dieses Resultat wird in Kap. 6 bei der Diskussion des Wärmewiderstandes von entscheidender Bedeutung sein.

Es lohnt sich, die Gitterschwingungen in den Grenzfällen sehr großer und sehr kleiner Wellenlängen näher zu untersuchen. Im *langwelligen Grenzfall*, d.h. für $q \to 0$, sind in der Entwicklung der Cosinus-Terme in (5.18) nur die beiden ersten Glieder von Bedeutung,

[13] Wir kürzen hier und in den folgenden Kapiteln den kleinsten reziproken Gittervektor $2\pi/a$ mit dem Buchstaben g ab.

und die Gleichung vereinfacht sich zu

$$\omega^2 = \frac{2}{M}\sum_{s>0} C_s \left[1 - \cos(qsa)\right] \approx \frac{q^2 a^2}{M}\sum_{s>0} s^2 C_s \,. \tag{5.24}$$

Offenbar besteht in diesem Fall ein linearer Zusammenhang zwischen ω und q, wie wir ihn in Form von (5.3) bereits bei den Schallwellen kennengelernt haben. Die resultierende Schallgeschwindigkeit folgt unmittelbar aus der Definition $v = \omega/q$. Damit haben wir auch einen mikroskopischen Ausdruck für die Elastizitätsmoduln gefunden. Wir übertragen unser Resultat auf einen kubisch primitiven Kristall, dessen Dichte durch $\rho = M/a^3$ gegeben ist. Ein Vergleich von (5.24) und (5.4) führt zu dem Ergebnis

$$c_{11} = \sum_{s>0} \frac{s^2}{a} C_s^{\mathrm{l}} \qquad \text{und} \qquad c_{44} = \sum_{s>0} \frac{s^2}{a} C_s^{\mathrm{t}} \,. \tag{5.25}$$

Bei den Kraftkonstanten wurde noch der Index l bzw. t angefügt, um longitudinale und transversale Auslenkungen zu unterscheiden. An dieser Stelle sei nochmals betont, daß alle Betrachtungen, die wir hier durchgeführt haben, auch für transversale Wellen gelten. Von Ausnahmen abgesehen sind die Kraftkonstanten C_s^{t} und damit die rücktreibenden Kräfte für Scherbewegungen jedoch immer kleiner als C_s^{l}. Es gilt deshalb die allgemeine Beziehung $v_{\mathrm{l}} \geq v_{\mathrm{t}}$.

Für Gitterschwingungen mit Wellenvektoren an der *Grenze der Brillouin-Zone* ist die Phasenverschiebung zwischen der Auslenkung benachbarter Netzebenen besonders interessant. Aus (5.21) folgt für $q \to \pm\pi/a$ unmittelbar $u_{s+1}/u_s = \exp(\pm \mathrm{i}\,\pi) = -1$. Die benachbarten Gitterebenen schwingen also in Gegenphase, d.h. Wellen mit dem Wellenvektor $\pm\pi/a$ und der Wellenlänge $\lambda = 2a$ sind stehende Wellen. Dieses Resultat macht eine einfache Betrachtung verständlich. Wie Röntgenwellen sind auch elastische Wellen der Bragg-Reflexion unterworfen. Dies folgt unmittelbar aus der Streutheorie, die in Kap. 2 behandelt wurde, denn dort wurde keine Einschränkung bezüglich der Natur der einfallenden Wellen gemacht. Wir setzen die Gitterkonstante a in (2.19) als Periode der Dichteschwankung ein und erhalten $2a \sin\theta = \lambda$. Breiten sich die Wellen senkrecht zu den Gitterebenen aus, dann ist $\theta = 90°$, und wir erwarten Bragg-Reflexion für die Wellenlänge $\lambda = 2a$. Einlaufende und reflektierte Welle überlagern sich, und es bildet sich eine stehende Welle aus.

Das Auftreten von stehenden Wellen an der Grenze der Brillouin-Zone drückt auch die Gruppengeschwindigkeit $v_{\mathrm{g}} = \partial\omega/\partial q$ aus. Durch Ableiten von (5.18) findet man, daß für alle Werte der Kraftkonstanten C_s die Gruppengeschwindigkeit für $q \to \pi/a$ gegen Null geht. Diese Bemerkung gilt nicht allgemein, sondern ist nur für Richtungen mit hoher Gittersymmetrie gültig, die wir implizit in unserer Betrachtung immer vorausgesetzt haben. Im allgemeinen Fall kann die Dispersionskurve an der Grenze der Brillouin-Zone eine endliche Steigung aufweisen.

Ohne Beweis soll hier noch angefügt werden, daß die übrigen in diesem Abschnitt gemachten Aussagen über die Ausbreitung von Gitterwellen in Richtung hoher Symmetrie in qualitativer Hinsicht allgemein gültig sind. Die Reduktion auf die 1. Brillouin-Zone ist in allen Fällen möglich, die Dispersionskurven zeigen auch in beliebiger Richtung vergleichbare Verläufe. Beispiele hierfür werden wir in Kap. 5.4 kennenlernen.

Gitter mit mehratomiger Basis

Enthält die primitive Elementarzelle mehrere Atome, so treten neue Effekte auf. Für die Erläuterung der Grundprinzipien genügt es auch hier, das besonders einfache Verhalten von Gitterwellen bei Ausbreitung in Richtungen hoher Symmetrie zu studieren. Qualitativ treffen auch hier die Aussagen wieder für beliebige Richtungen zu. Wir betrachten der Einfachheit halber ein Gitter mit zweiatomiger Basis und der Gitterkonstanten a, bei dem die Atome so, wie in Abb. 5.5 angedeutet, angeordnet sind. Wie beim Gitter mit einatomiger Basis können wir uns vorstellen, daß es sich um eine (001)-Ebene eines orthorhombischen Kristalls handelt. Zur näherungsweisen Beschreibung des Verhaltens dreidimensionaler Kristalle greifen wir wieder auf eine lineare Kette zurück. In unserer weiteren Diskussion bezeichnen wir einfachheitshalber die schwarzen Punkte als Atome A und die offenen Kreise als Atome B. Ihre Lage in der linearen Kette wird durch den Index der Netzebenen gekennzeichnet. Wie bei den Kristallen mit einatomiger Basis gilt auch hier, daß sich die Kräfte senkrecht zur Ausbreitungsrichtung der Welle kompensieren. Die Näherung in Form der linearen Kette erweist sich daher auch dann als gut, wenn die Atome in benachbarten Ebenen gegeneinander versetzt sind.

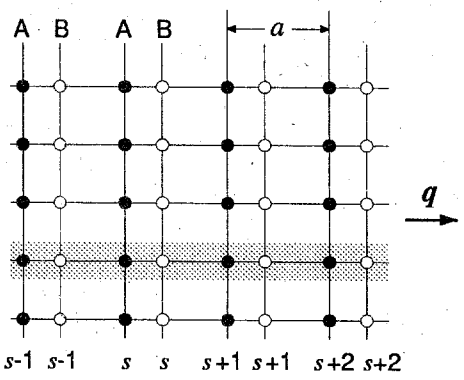


Abb. 5.5: Zur Herleitung der Gitterschwingung bei zweiatomiger Basis. Die einzelnen Ebenen enthalten jeweils nur gleichartige Atome. Die Punkte bezeichnen wir als die Atome A, die offenen Kreise als Atome B. Die herausgegriffene lineare Kette ist dunkel getönt.

Bevor wir die Koeffizientenmatrix (5.14) lösen, überlegen wir uns auch hier die Bedeutung der Kraftkonstanten und vereinfachen die auftretenden Ausdrücke. Wir nehmen an, daß die Kräfte so rasch mit dem Abstand abfallen, daß nur die Wechselwirkung mit den unmittelbar benachbarten Atomen berücksichtigt werden muß. Stellen wir uns vor, daß das Atom A_s in der linearen Kette ausgelenkt wird, so bewirken die beiden benachbarten Atome B_s bzw. B_{s-1} eine rücktreibende Kraft, die wir mit den Kraftkonstanten C' und C'' charakterisieren. Die gleiche Argumentation gilt auch bei einer Auslenkung des Atoms B_s, d.h. wir finden $D_{11} = D_{22} = C' + C''$. Umgekehrt erfährt das Atom A_s eine Kraft, die es aus seiner Gleichgewichtslage treibt, wenn das Atom B_s ausgelenkt wird. Wählen wir das Vorzeichen der Kraftkonstanten wie bei der Beschreibung des Gitters mit einatomiger Basis, so tritt nun ein negatives Vorzeichen auf. Da nur die nächsten Nachbarn berücksichtigt werden, finden wir die Ausdrücke $D_{12} = -[C' + C''\exp(-iqa)]$ und $D_{21} = -[C' + C''\exp(iqa)]$. Mit der Abkürzung $C' + C'' = 2C$ erhalten wir schließlich die einfache Koeffizientenmatrix

$$\begin{vmatrix} 2C - \omega^2 M_1 & -(C' + C''\mathrm{e}^{-\mathrm{i}qa}) \\ -(C' + C''\mathrm{e}^{\mathrm{i}qa}) & 2C - \omega^2 M_2 \end{vmatrix} = 0 \qquad (5.26)$$

und die Eigenwerte

$$\omega_{\mathrm{a,o}}^2 = C\left(\frac{1}{M_1}+\frac{1}{M_2}\right) \pm C\sqrt{\left(\frac{1}{M_1}+\frac{1}{M_2}\right)^2 - \frac{4C'C''}{M_1 M_2 C^2}\sin^2\left(\frac{qa}{2}\right)}\,. \tag{5.27}$$

Hierbei stehen die Indizes a und o für *akustisch* und *optisch* und sind den Vorzeichen − und + der Wurzel zugeordnet. Der Grund für diese Bezeichnung wird auf der nächsten Seite erläutert. Der schematische Verlauf der Dispersionskurven ist in Abb. 5.6 dargestellt. Da zwei getrennte Kurven auftreten, spricht man von zwei *Zweigen*. Die Gitterschwingungen einer linearen Kette mit ein- und zweiatomiger Basis unterscheiden sich also bereits in der Anzahl der auftretenden Zweige.

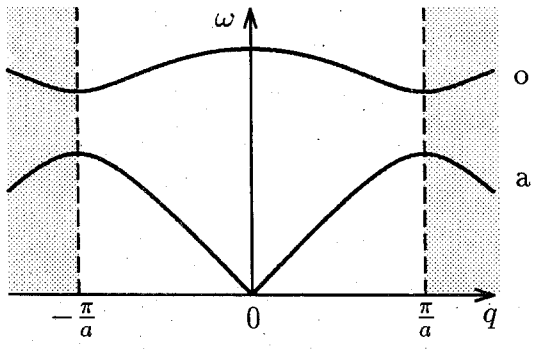


Abb. 5.6: Verlauf der Dispersionskurven in Kristallen mit zweiatomiger Basis. Der optische Zweig liegt über dem akustischen, dazwischen tritt eine „verbotene" Zone auf.

Um übersichtliche Ausdrücke zu erhalten, setzen wir in der weiteren Diskussion $C' = C'' = C$, d.h. wir nehmen an, daß zwischen benachbarten Gitterbausteinen die gleichen Kräfte herrschen. Diese Annahme ist beispielsweise bei den Ionenkristallen sehr gut erfüllt. Bei Kristallen bestehend aus zweiatomigen Molekülen ist sie natürlich eine schlechte Näherung. Der akustische Zweig mit dem Minuszeichen in (5.27) besitzt die gleichen Eigenschaften wie bei einer Kette mit einatomiger Basis. Dies läßt sich leicht für den speziellen Fall $M = M_1 = M_2$ zeigen, wenn man berücksichtigt, daß der Basisvektor bei einer zweiatomigen Basis doppelt so groß ist, d.h. wenn in der obigen Gleichung der Gitterabstand a durch $2a$ ersetzt wird. Da wir die akustischen Schwingungen schon ausführlich diskutiert haben, beschäftigen wir uns hier nicht weiter mit dieser Lösung.

Nun untersuchen wir das Verhalten des optischen Zweiges, für den in (5.27) vor der Wurzel das positive Vorzeichen steht. Für $q \to 0$ erhalten wir die Dispersionsrelation $\omega_{\mathrm{o}}^2 = 2C/\mu$, wobei die reduzierte Masse μ durch $\mu^{-1} = M_1^{-1} + M_2^{-1}$ gegeben ist. Die Frequenz des optischen Zweiges hängt, wie auch in Abb. 5.6 zu sehen ist, in der Nähe des Γ-Punktes kaum vom Wellenvektor ab. Die Tatsache, daß akustische und optische Gitterschwingungen bei gleichem Wellenvektor zwei völlig verschiedene Frequenzen aufweisen können, hat ihre Ursache in den unterschiedlichen Phasenbeziehungen zwischen den Auslenkungen der beiden Atomsorten. Um diesen Aspekt näher zu untersuchen, gehen wir zur Koeffizientenmatrix (5.26) zurück und schreiben, weiterhin unter der Annahme $C' = C'' = C$, die Gleichung an, die der ersten Zeile zugrunde liegt:

$$(2C - \omega^2 M_1)\,u_1 - C\,(1 + \mathrm{e}^{-\mathrm{i}qa})\,u_2 = 0\,. \tag{5.28}$$

Betrachten wir zunächst den Fall $q \to 0$, so ergibt sich

$$\frac{u_1}{u_2} \approx \frac{2C}{2C - \omega^2 M_1}\,. \tag{5.29}$$

Für den akustischen Zweig ist $\omega \approx 0$ und somit $u_1 \approx u_2$, d.h. die beiden Atomsorten schwingen mit der gleichen Phase. Im Gegensatz hierzu finden wir für den optischen Zweig $u_1/u_2 = -M_2/M_1$, da dort $\omega_o^2 = 2C/\mu$ ist. Dies bedeutet, daß alle Atome mit der Masse M_1 in Gegenphase zu den Atomen mit der Masse M_2 schwingen. Diese Art der Auslenkung läßt sich besonders einfach für transversale Gitterschwingungen darstellen, obwohl die Feststellungen natürlich auch für longitudinale Wellen gültig sind. Die Abb. 5.7 veranschaulicht die Auslenkungen der Atom bei den unterschiedlichen Wellentypen für den Fall großer Wellenlängen.

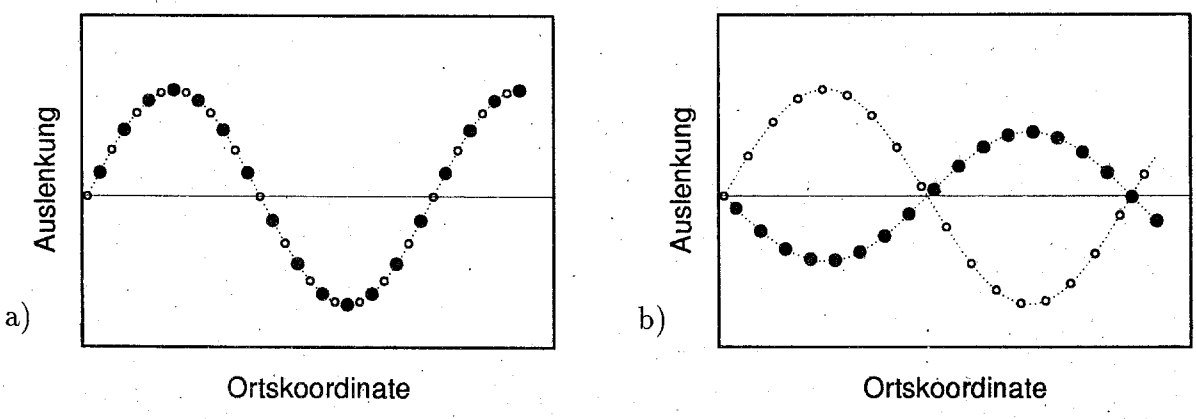


Abb. 5.7: Atomare Auslenkung bei langwelligen Gitterschwingungen. Offene Kreise und Punkte kennzeichnen die unterschiedlichen Atomsorten. Die Amplituden sind im Vergleich zur Wellenlänge stark vergrößert dargestellt.
a) Akustische Gitterschwingung; die benachbarten Atome besitzen fast die gleiche Phase.
b) Optische Gitterschwingung; die beiden Atomsorten schwingen gegeneinander.

Warum unterscheidet man zwischen akustischen und optischen Gitterwellen? Offensichtlich sind akustische Gitterwellen bei langen Wellenlängen, d.h. für $q \to 0$ identisch mit den üblichen Schallwellen. Bei den optischen Gitterwellen dagegen ist in vielen Fällen mit der Auslenkung der Untergitter gegeneinander ein oszillierendes elektrisches Dipolmoment verknüpft. Dadurch können diese Schwingungen an elektromagnetische Wellen ankoppeln und bestimmen weitgehend die optischen Eigenschaften im Infrarotbereich (vgl. Kap. 11.2). Das bekannteste Beispiel hierfür sind die Ionenkristalle, bei denen die gegeneinander schwingenden Untergitter entgegensetzte elektrische Ladungen tragen. Kristalle mit derartigen Eigenschaften bezeichnet man als *infrarot-aktiv*. Anders liegen die Verhältnisse, wenn die Basis beispielsweise aus zwei identischen Atomen besteht, wie es bei Diamant oder Silizium der Fall ist. Die Schwingungen der Untergitter gegeneinander bewirken hier kein Dipolmoment, folglich wird auch keine Infrarotabsorption beobachtet. Wie wir im folgenden Abschnitt 5.4 sehen werden, lassen sich diese Schwingungen mit Hilfe der Raman-Streuung nachweisen. Hier sei noch vermerkt, daß sich die Schwingungszustände von isolierten Molekülen beim Übergang vom einzelnen Molekül zum ausgedehnten Festkörper zu optischen Gitterwellen entwickeln.

Für $q \to \pi/a$ finden wir für die Dispersionsrelation unter der Annahme $M_1 < M_2$ die Ausdrücke $\omega_o^2 = 2C/M_1$ bzw. $\omega_a^2 = 2C/M_2$, die beide nicht explizit von der Wellenzahl abhängen. Setzen wir diese Beziehungen wieder in (5.29) ein, so ergeben sich die erstaunlichen Ergebnisse $u_2/u_1 = 0$ bzw. $u_1/u_2 = 0$. Abhängig vom betrachteten Zweig ist also entweder das Untergitter der schweren oder der leichten Atome in Ruhe, während der andere Teil des Gitters schwingt. Für Gitterwellen mit Wellenvektoren nahe der

Brillouin-Zone verschwindet das oszillierende Dipolmoment aufgrund der gegenphasigen Schwingung der Ionen in benachbarten primitiven Elementarzellen. Dies bedeutet, daß auch bei optisch aktiven Festkörpern die Gitterschwingungen bei diesen Wellenlängen nicht an elektromagnetische Wellen ankoppeln.

Ist $M_1 \neq M_2$ oder treten asymmetrische Bindungskräfte auf, d.h. ist $C' \neq C''$, so öffnet sich an der Grenze der Brillouin-Zone eine *Frequenzlücke*. In der *verbotenen Zone* existieren keine Eigenschwingungen des Gitters. Regt man elastische Wellen mit einer derartigen Frequenz an, so klingen sie im Gitter *exponentiell* innerhalb weniger Gitterkonstanten ab. Die Breite der verbotenen Zone wird durch das Verhältnis der Massen oder der Kopplungskonstanten der beteiligten Atome bestimmt.

Wir wenden uns nun noch kurz der Frage zu, welchen Einfluß die Wahl der Elementarzelle auf die Ergebnisse unserer Rechnungen hat. Wir starten mit einer linearen Kette bestehend aus identischen Atomen im Abstand a. Die uns bereits bekannte Dispersionsrelation dieser Kette ist in Abb. 5.8a zu sehen. Nun aber fassen wir bei der gleichen Kette in Gedanken je zwei Atome zusammen, d.h. die Basis ist nun zweiatomig und der Basisvektor $a' = 2a$. Damit verbunden ist eine Halbierung der Brillouin-Zone und das zusätzliche Auftreten eines optischen Zweiges, wie in Abb. 5.8b gezeigt ist. Durch die „ungeschickte" Wahl des Basisvektors a' werden die Gitterwellen, die im Bereich $\pi/a' < q \leq \pi/a$ liegen, im Bereich $-\pi/a' < q \leq 0$ dargestellt. Entsprechendes gilt für die Wellenvektoren mit entgegengesetztem Vorzeichen, die dann im Bereich $0 < q \leq \pi/a$ anzutreffen sind. Dieser Eigentümlichkeit der Darstellung im reziproken Raum werden wir in Kap. 7.3 bei der Beschreibung der Elektronen wieder begegnen und sie dort ausführlich diskutieren. Natürlich gelten die oben gemachten Aussagen über die Eigenschaften der Gitterschwingungen an der Grenze der Brillouin-Zone nicht für die „künstliche" Grenze bei $q = \pi/a'$, da dort die Dispersionskurve nicht horizontal verläuft.

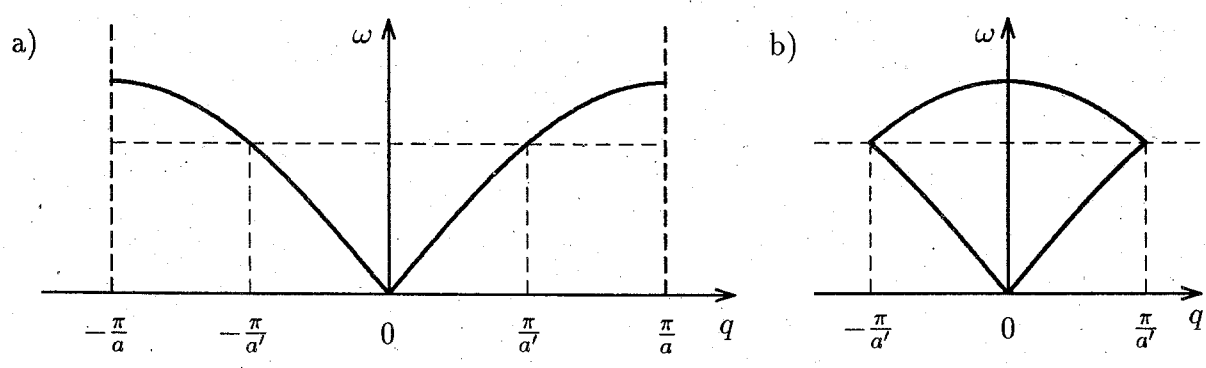


Abb. 5.8: Einfluß der Elementarzellengröße auf das Schwingungsspektrum.
a) Verlauf eines akustischen Zweiges bei einatomiger Basis.
b) Schwingungsspektrum nach „künstlicher" Halbierung der Brillouin-Zone durch die Wahl einer Elementarzelle mit zweiatomiger Basis.

Die Ergebnisse, die wir anhand von linearen Ketten erzielt haben, lassen sich weitgehend auf reale Kristalle übertragen, wenn die Ausbreitung der Gitterwelle in einer Richtung hoher Symmetrie erfolgt. Natürlich ist die Übereinstimmung mit dem gemessenen Verlauf der Dispersionskurven nicht in allen Fällen zufriedenstellend. Dies gilt insbesondere für optische Zweige, da dort aufgrund der gegenphasigen Auslenkung meist mehrere Kraftkonstanten berücksichtigt werden müssen. In allen Fällen findet man drei akustische

Zweige, nämlich einen longitudinalen und zwei transversale, deren Polarisationsrichtungen aufeinander senkrecht stehen. In Richtungen mit hoher Symmetrie des Gitters können Zweige zusammenfallen. Dies gilt selbstverständlich bei isotropen Medien, deren beide transversale Zweige immer entartet sind. In der Nähe des Γ-Punkts liegt der longitudinale Zweig in fast allen Fällen über den transversalen, nie aber darunter.

In Abb. 5.9 ist der generelle Verlauf der Dispersionskurven für einen Kristall mit zweiatomiger Basis nochmals schematisch dargestellt. Es soll hier erneut betont werden, daß im allgemeinen Fall, d.h. wenn man nicht gerade eine Ausbreitungsrichtung mit hoher Kristallsymmetrie betrachtet, die Zweige nicht entartet sind und keinen „reinen" Charakter besitzen, d.h. daß man dann nur vor quasi-longitudinalen bzw. quasi-transversalen Schwingungen sprechen kann.

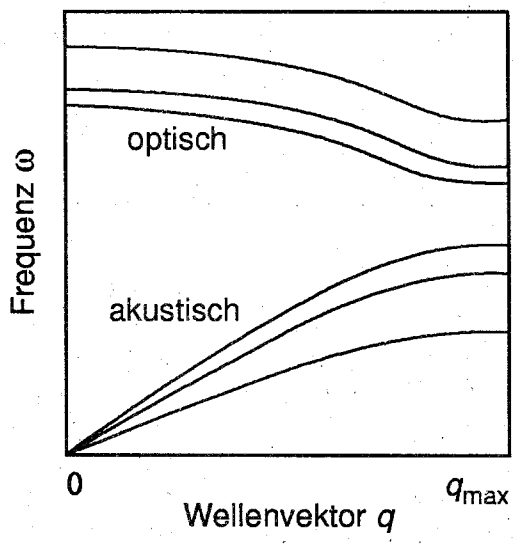


Abb. 5.9: Schematische Darstellung der sechs Phononenzweige in Kristallen mit zweiatomiger Basis. Sieht man von der Ausbreitung in Richtung mit hoher Symmetrie ab, so sind die verschiedenen Phononenzweige nicht entartet.

Enthält eine primitive Elementarzelle mehr als nur ein Atom, so treten zusätzlich zu den akustischen noch optische Zweige auf. Allgemein gilt, daß bei einer Basis mit p Atomen $3p$ Zweige existieren, drei akustische und $(3p-3)$ optische. Wie bei den akustischen Zweigen gibt es auch bei den optischen doppelt so viele transversale wie longitudinale Zweige. Mit zunehmender Größe der Elementarzelle wird das Schwingungsspektrum komplexer und unübersichtlicher, da eine Vielzahl von optischen Zweigen auftritt. Im folgenden Abschnitt werden wir auf die experimentelle Bestimmung der Dispersionskurven eingehen und den tatsächlichen Verlauf anhand einiger einfacher Beispiele betrachten.

5.4 Experimentelle Bestimmung von Dispersionskurven

Streuung am dynamischen Gitter, Quantisierung der Gitterschwingungen

Aus der kohärenten elastischen Streuung am starren Gitter läßt sich über den *statischen* Strukturfaktor die geometrische Anordnung der Atome herleiten. Um Auskunft über die Dynamik zu bekommen, müssen wir im Streuvorgang die zeitlichen Veränderungen des Gitters berücksichtigen, die zu inelastischen Streuprozessen führen. Wir benutzen bei der

Herleitung der Streuamplitude die Ausdrücke (2.3) und (2.4), die wir in Kap 2.2 für die Feldstärke $A_s(t)$ der gestreuten Strahlung abgeleitet haben:

$$A_s(t) = \frac{\widetilde{A}}{\tilde{r}_0} e^{-i(\omega_0 t - k\tilde{r}_0)} \mathcal{A}(\boldsymbol{K}) \propto e^{-i\omega_0 t} \mathcal{A}(\boldsymbol{K}) . \tag{5.30}$$

Der Einfachheit halber lassen wir in der weiteren Diskussion die unwichtigen Vorfaktoren weg und betrachten nur den Ausdruck $\exp(-i\omega_0 t)\,\mathcal{A}(\boldsymbol{K})$. Wir spalten die Streuamplitude $\mathcal{A}(\boldsymbol{K})$ wie in (2.22) in Strukturamplitude und Atom-Strukturfaktor auf, erlauben eine Zeitabhängigkeit der Atomkoordinaten $\boldsymbol{r}_m$ und erhalten

$$A_s(t) \propto e^{-i\omega_0 t} \sum_m e^{-i\boldsymbol{K}\cdot\boldsymbol{r}_m(t)}\, \widetilde{a}_m(\boldsymbol{K}) . \tag{5.31}$$

Nehmen wir zur Vereinfachung an, daß der Kristall ein primitives Gitter mit einatomiger Basis besitzt, so bewirkt der Atom-Strukturfaktor $\widetilde{a}_m(\boldsymbol{K})$ nur eine Winkelabhängigkeit der Streuintensität, die für die weitere Betrachtung unerheblich ist. In dem hier betrachteten einfachen Fall ist die Lage der Atome durch den Gittervektor $\boldsymbol{R}$ und die Auslenkung $\boldsymbol{u}_{\boldsymbol{R}}(t)$ gegeben, so daß wir $\boldsymbol{r}_m$ durch $\boldsymbol{r}_{\boldsymbol{R}}(t) = \boldsymbol{R} + \boldsymbol{u}_{\boldsymbol{R}}(t)$ ersetzen können. Die Teilchenauslenkung wiederum setzt sich aus den Beiträgen aller Eigenschwingungen zusammen:

$$\boldsymbol{u}_{\boldsymbol{R}}(t) = \sum_{\boldsymbol{q}>0,j} \boldsymbol{U}(\boldsymbol{q},j) \left[e^{-i(\boldsymbol{q}\cdot\boldsymbol{R}-\omega_{q,j}t)} + e^{+i(\boldsymbol{q}\cdot\boldsymbol{R}-\omega_{q,j}t)} \right] . \tag{5.32}$$

Die Auswertung dieser Summe werden wir erst im folgenden Abschnitt 5.5 im Zusammenhang mit der spezifischen Wärme besprechen. Für die Diskussion der inelastischen Streuung ist die Kenntnis der Summe nicht erforderlich. Ersetzen wir $\boldsymbol{r}_m(t)$ in (5.31) durch $\boldsymbol{r}_{\boldsymbol{R}}(t)$, so ergibt sich unter Einbeziehung von (5.32)

$$\begin{aligned} A_s(t) &\propto e^{-i\omega_0 t} \sum_{\boldsymbol{R}} \exp\left\{ -i\boldsymbol{K}\cdot\left[\boldsymbol{R} + \sum_{\boldsymbol{q}>0,j} \boldsymbol{U}(\boldsymbol{q},j) \left[e^{-i(\boldsymbol{q}\cdot\boldsymbol{R}-\omega_{q,j}t)} + e^{+i(\boldsymbol{q}\cdot\boldsymbol{R}-\omega_{q,j}t)} \right] \right] \right\} \\ &\propto e^{-i\omega_0 t} \sum_{\boldsymbol{R}} e^{-i\boldsymbol{K}\cdot\boldsymbol{R}} \prod_{\boldsymbol{q}>0,j} \exp\left[-i\boldsymbol{K}\cdot\boldsymbol{U}(\boldsymbol{q},j) \left[e^{-i(\boldsymbol{q}\cdot\boldsymbol{R}-\omega_{q,j}t)} + e^{+i(\boldsymbol{q}\cdot\boldsymbol{R}-\omega_{q,j}t)} \right] \right] . \end{aligned} \tag{5.33}$$

In den meisten Fällen ist $\boldsymbol{K}\cdot\boldsymbol{U}_{\boldsymbol{q}} \ll 1$, d.h. wir können den Teil der Exponentialfunktion, der die Auslenkung enthält, entwickeln und finden:

$$\begin{aligned} &\exp\left\{ -i\boldsymbol{K}\cdot\boldsymbol{U}(\boldsymbol{q},j) \left[e^{-i(\boldsymbol{q}\cdot\boldsymbol{R}-\omega_{q,j}t)} + e^{+i(\boldsymbol{q}\cdot\boldsymbol{R}-\omega_{q,j}t)} \right] \right\} \\ &\quad \approx 1 - i\boldsymbol{K}\cdot\boldsymbol{U}(\boldsymbol{q},j) \left[e^{-i(\boldsymbol{q}\cdot\boldsymbol{R}-\omega_{q,j}t)} + e^{+i(\boldsymbol{q}\cdot\boldsymbol{R}-\omega_{q,j}t)} \right] - \\ &\quad - \left| \boldsymbol{K}\cdot\boldsymbol{U}(\boldsymbol{q},j) \right|^2 - \frac{1}{2} \left[\boldsymbol{K}\cdot\boldsymbol{U}(\boldsymbol{q},j) \right]^2 \left[e^{-2i(\boldsymbol{q}\cdot\boldsymbol{R}-\omega_{q,j}t)} + e^{+2i(\boldsymbol{q}\cdot\boldsymbol{R}-\omega_{q,j}t)} \right] - \ldots . \end{aligned} \tag{5.34}$$

Wir setzen die Entwicklung in (5.33) ein und multiplizieren das Produkt aus. Von den vielen Termen berücksichtigen wir nur den konstanten und die Glieder der Entwicklung, die linear in der Auslenkung sind. Auf die Bedeutung der quadratischen Terme gehen wir

etwas später ein. Unter dieser vereinfachenden Voraussetzung läßt sich das Produkt in eine Summe umschreiben, und wir finden

$$A_s(t) \propto \sum_{\boldsymbol{R}} e^{-i\boldsymbol{K}\cdot\boldsymbol{R}} \left\{ 1 - \sum_{\boldsymbol{q}>0,j} i\boldsymbol{K}\cdot\boldsymbol{U}(\boldsymbol{q},j) \left[e^{-i(\boldsymbol{q}\cdot\boldsymbol{R}-\omega_{q,j}t)} + e^{+i(\boldsymbol{q}\cdot\boldsymbol{R}-\omega_{q,j}t)} \right] \right\} e^{-i\omega_0 t}$$

$$\propto \underbrace{\sum_{\boldsymbol{R}} e^{-i\boldsymbol{K}\cdot\boldsymbol{R}}\, e^{-i\omega_0 t}}_{\text{elastisch}} - \underbrace{\sum_{\boldsymbol{R}} \sum_{\boldsymbol{q}>0,j} i\boldsymbol{K}\cdot\boldsymbol{U}(\boldsymbol{q},j)\, e^{-i(\boldsymbol{K}\pm\boldsymbol{q})\cdot\boldsymbol{R}}\, e^{-i(\omega_0\pm\omega_{q,j})t}}_{\text{inelastisch}} . \qquad (5.35)$$

Den ersten Term haben wir bereits in Kap. 2.3 im Zusammenhang mit der elastischen Streuung behandelt. Dort wurde gefunden, daß die Summation über die Gittervektoren $\boldsymbol{R}$ nur dann zu endlichen Beiträgen führt, wenn die Streubedingung $\boldsymbol{K} = \boldsymbol{G}$ erfüllt ist. Die gleichen Argumente gelten auch bei inelastischer Streuung, die durch den zweiten Term beschrieben wird. Die modifizierte Streubedingung lautet dann $\boldsymbol{K} \pm \boldsymbol{q} = \boldsymbol{G}$. Bildet man das Zeitmittel über die gestreute Intensität, so ist dieses nur von Null verschieden, wenn die Streustrahlung zusätzlich die Bedingung $\omega = \omega_0 \pm \omega_{\boldsymbol{q},j}$ erfüllt. Klassisch betrachtet wird die Streustrahlung durch die Oszillation der Atome moduliert, wodurch Seitenbanden der abgestrahlten Welle erzeugt werden.

Die beiden Streubedingungen lassen sich durch Multiplikation mit dem Planckschen Wirkungsquantum als *Energie-* und *Impulserhaltung* bei der Wechselwirkung von Teilchen darstellen:

$$\hbar\omega = \hbar\omega_0 \pm \hbar\omega_{\boldsymbol{q},j} , \qquad (5.36)$$

$$\hbar\boldsymbol{k} = \hbar\boldsymbol{k}_0 \pm \hbar\boldsymbol{q} + \hbar\boldsymbol{G} . \qquad (5.37)$$

Dieses Ergebnis bedarf der Erläuterung. Zunächst ist zu bemerken, daß wir beim Term $\hbar\boldsymbol{G}$ aus „ästhetischen" Gründen einen Vorzeichenwechsel vorgenommen haben. Diese Änderung ist erlaubt, weil der reziproke Gittervektor nicht auf ein Vorzeichen beschränkt ist. Den Energiesatz können wir ähnlich wie bei atomaren Streuprozessen, z.B. der Raman-Streuung an Molekülen, interpretieren: Ein einfallendes Röntgenquant, Neutron oder Elektron wechselwirkt mit dem Gitter und erzeugt oder vernichtet ein Schwingungsquant des Festkörpers, ein „Schallteilchen". In Analogie zu den Schwingungsquanten des elektromagnetischen Feldes, den Photonen, bezeichnet man die Schwingungsquanten des elastischen Feldes als *Phononen*[14].

Entsprechend besitzt ein Phonon die Energie $\hbar\omega_{\boldsymbol{q},j}$ und den Impuls $\hbar\boldsymbol{q}$. Auf die Eigenschaften dieser „Teilchen" werden wir in den folgenden Abschnitten eingehen. Die Erhaltungssätze (5.36) und (5.37) sind ein Ausdruck dafür, daß beim Streuprozeß entweder ein Phonon erzeugt (− Zeichen) oder ein Phonon vernichtet (+ Zeichen) wird. Der Impulssatz hat jedoch eine ungewöhnliche Form: Zum Impuls der beteiligten Teilchen kann noch der Impuls $\hbar\boldsymbol{G}$ addiert werden, wobei $\boldsymbol{G}$ ein beliebiger reziproker Gittervektor ist. Dies kann man sich so vorstellen, daß beim Streuprozeß ein Phonon erzeugt oder vernichtet wird und daß gleichzeitig eine Bragg-Reflexion erfolgt. Der Impuls $\hbar\boldsymbol{G}$ wird an den Kristall als Ganzes übertragen. Man bezeichnet daher die impulsähnliche Größe $\hbar\boldsymbol{q}$ als *Quasiimpuls* oder *Kristallimpuls*. Eigentlich sollte man nicht von der Impulserhaltung, sondern eher von der Erhaltung des Wellenvektors sprechen.

[14] Der Begriff *Phonon* wurde erstmals 1932 von *Frenkel* in seinem Buch „*Wave Mechanics, Elementary Theory*" benutzt.

In Abb. 5.10 ist der soeben diskutierte Streuprozeß schematisch für den Fall der Phononenvernichtung dargestellt. Ein Teilchen, z.B. ein Neutron, mit dem Impuls $\hbar \boldsymbol{k}_0$ stößt mit einem Phonon und vernichtet es. Dadurch erhöht sich die Energie des gestreuten Teilchens, der Betrag des Wellenvektors vergrößert sich und kommt dadurch außerhalb der Ewald-Kugel zu liegen. Der Impulsübertrag $\hbar \boldsymbol{K}$ enthält den Anteil $\hbar \boldsymbol{G}$, den das gesamte Gitter aufnimmt, und den Quasiimpuls $\hbar \boldsymbol{q}$ des vernichteten Phonons. Wäre ein Phonon erzeugt worden, dann läge der Wellenvektor des gestreuten Neutrons innerhalb der Ewald-Kugel.

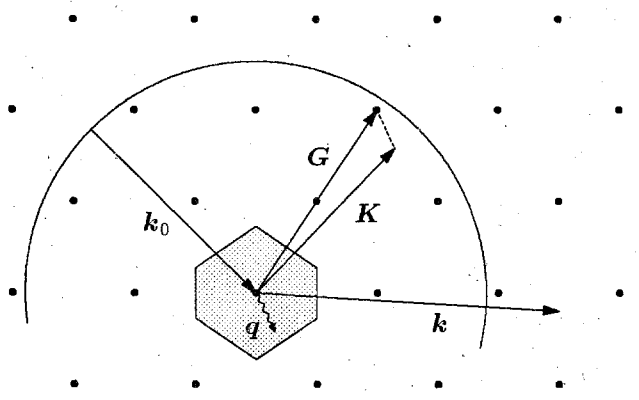


Abb. 5.10: Inelastische Streuung. Das einfallende Teilchen mit dem Impuls $\hbar \boldsymbol{k}_0$ vernichtet ein Phonon mit dem Impuls $\hbar \boldsymbol{q}$. Der beteiligte reziproke Gittervektor ist mit $\boldsymbol{G}$, der Streuvektor mit $\boldsymbol{K}$ bezeichnet. Die 1. Brillouin-Zone und die Ewald-Kugel sind ebenfalls eingezeichnet.

Es ist ein interessantes Ergebnis der Diskussion des Streuvorgangs, daß die Gitterwellen auch Teilchencharakter besitzen und daß ihre Energie und somit auch ihre Amplitude quantisiert sind. Diese Eigenschaft tritt aber nicht völlig unerwartet auf. Die Behandlung der Bewegungsgleichung des Gitters führte zu einem Ausdruck, der die Schwingungen von harmonischen Oszillatoren beschreibt. Natürlich ist die Energie dieser Oszillatoren quantisiert, d.h. es treten die Eigenwerte $E_{\boldsymbol{q},j} = (n_{\boldsymbol{q}} + \frac{1}{2})\,\hbar\omega_{\boldsymbol{q},j}$ auf. Jedes Phonon repräsentiert ein Energiequant $\hbar\omega_{\boldsymbol{q},j}$. Die Quantenzahl $n_{\boldsymbol{q}}$ gibt an, wieviele Phononen mit dem Wellenvektor $\boldsymbol{q}$ im Festkörper anzutreffen sind. Für das Verständnis der Phononeneigenschaften ist wichtig, sich klar zu machen, daß die hier angesprochenen harmonischen Oszillatoren nicht den lokalisierten Schwingungen einzelner Atome zugeschrieben werden können. *Jedes* Atom des Festkörpers trägt zu *jedem* Schwingungszustand $\omega_{\boldsymbol{q},j}$, d.h. zu jedem Phonon, bei. Ein Phonon ist die Anregung des *gesamten* Gitters.

Im Gegensatz zu den Photonen tragen die Phononen keinen echten Impuls. Es läßt sich zeigen, daß der wahre Impuls der Phononen exakt Null ist und daher mit den Gitterschwingungen kein Massentransport verknüpft ist. Phononen benehmen sich so, *als ob* sie den Impuls $\boldsymbol{p} = \hbar \boldsymbol{q}$ tragen würden. Im Gegensatz zu den Photonen, den Quanten des elektromagnetischen Feldes, sind die Phononen, die Quanten des Schallfeldes, keine fundamentalen Teilchen. Man bezeichnet die Phononen deshalb auch oft als *Quasiteilchen*, denn sie sind nur eine Konsequenz der Einführung von Normalkoordinaten. Es lohnt sich, nochmals auf das Streuexperiment in Abb. 5.10 einzugehen. Tatsächlich ändert sich der Impuls der gestreuten Teilchen, wenn ein Phonon erzeugt oder vernichtet wird. Den Impulsübertrag übernimmt, wie bei der elastischen Streuung, der Festkörper als Ganzes. Ist die Probe beweglich, so bewegt sich nach dem Streuprozeß ihr Schwerpunkt. Die angeregten oder vernichteten Schwingungen selbst, d.h. die Phononen, tragen zu dessen Impuls nicht bei.

Wir wollen uns noch mit einigen Zahlenwerten vertraut machen. Als erstes schätzen wir die Teilchenauslenkung ab, die ein *einzelnes* Phonon in einer Probe hervorruft. Die Energie $\hbar\omega$ dieses Phonons ist gleichmäßig über die N Atome der einkristallinen Probe verteilt, jedes Atom besitzt also die Energie $\hbar\omega/N$. Da die Atome näherungsweise in einem harmonischen Potential schwingen, ist andererseits ihre Energie $E = M\omega^2U^2/2$. Somit verursacht ein einzelnes Phonon eine Auslenkung mit der Amplitude

$$U = \sqrt{\frac{2\hbar}{\omega\rho V}} \,. \tag{5.38}$$

Wir betrachten bei unserer Abschätzung zunächst die Zahlenwerte für ein Phonon am Rande der Brillouin-Zone mit der Wellenlänge $\lambda = 4\,\text{Å}$. Mit einer Schallgeschwindigkeit von $v = 5\cdot10^3\,\text{m/s}$ errechnet sich mit Hilfe von (5.19) eine Frequenz von etwa $\nu = 10^{13}\,\text{Hz}$. Setzen wir einen Eisenwürfel mit der Kantenlänge 1 cm und der Dichte $\rho = 7{,}8\,\text{g/cm}^3$ voraus, so finden wir die Amplitude $U \approx 2\cdot10^{-21}\,\text{cm}$. Dies ist ein extrem kleiner Wert! Der Kern dieser Aussage ändert sich auch nicht, wenn wir zu großen Wellenlängen übergehen. Für $\lambda = 4\,\text{cm}$ liefert die Abschätzung $U \approx 2\cdot10^{-17}\,\text{cm}$.

Eine weitere interessante Größe ist die Zahl der angeregten Phononen, z.B. bei Zimmertemperatur. Setzen wir die Gültigkeit des Dulong-Petit-Gesetzes bis $T = 0$ voraus (der dadurch auftretende Fehler wird in Kap. 5.5 diskutiert und ist hier ohne Bedeutung), so sind in einem Eisenwürfel mit 1 cm Kantenlänge etwa 1000 J gespeichert. Mit der Energie $\hbar\omega = 6\cdot10^{-21}\,\text{J}$ für kurzwellige Phononen finden wir, daß die Zahl der Phononen mit $2\cdot10^{23}$ pro cm^3 vergleichbar mit der Zahl der Atome ist. Würde die Auslenkung der Atome in Phase erfolgen, so ergäbe sich eine Amplitude $NU \approx 4\,\text{m}$! Da die thermischen Phononen voneinander unabhängig sind, die von den einzelnen Phononen bewirkten Auslenkungen also keine festen Phasenbeziehungen zueinander besitzen, ergibt sich die wesentlich kleinere mittlere Auslenkungsamplitude von $\sqrt{N\,\overline{U^2}} \approx 0{,}1\,\text{Å}$.

Kohärente inelastische Neutronenstreuung an Kristallen

Phononendispersionskurven lassen sich am einfachsten mit Hilfe von kohärenten inelastischen Streuexperimenten bestimmen. In Kap. 2 haben wir bereits im Zusammenhang mit der elastischen Streuung Meßsonden kennengelernt, die zur Strukturbestimmung verwandt werden können. Sind alle dort diskutierten Strahlen auch für inelastische Streuexperimente geeignet? Um einen gut meßbaren Effekt zu erhalten, sollte die relative Energieänderung des gestreuten Teilchens möglichst groß sein. Da typische Phononenenergien im Bereich $10^{-2}\,\text{eV}$ liegen, ist für thermische Neutronen die relative Energieänderung $\delta E/E \approx 10^{-1}$ vergleichsweise groß. Die Phononendispersionskurven werden daher in den meisten Fällen mit Neutronenstreuexperimenten bestimmt. Die relative Energieänderung, die Röntgenquanten erfahren, ist dagegen nur von der Größenordnung $\delta E/E = \omega_{\boldsymbol{q}}/\omega_0 \approx 10^{-6}$. Messungen der Phononendispersion mit Hilfe der inelastischen Röntgenstreuung erfordern daher eine extrem hohe Auflösung. In den letzten Jahren wurden auf diesem Gebiet enorme Fortschritte erzielt. Es wurden nicht nur neue, äußerst intensive Synchrotronquellen gebaut, sondern auch die Entwicklung der Spektrometer im Röntgenbereich vorangetrieben. Es stehen nun beide Experimentiertechniken zur Verfügung und ergänzen sich gegenseitig. Der Nachweis von akustischen Phononen in der Nähe des Zentrums der Brillouin-Zone mit sehr kleiner Wellenzahl ist allerdings in

beiden Fällen nicht möglich, da der Energieübertrag entsprechend klein ist. Hierfür eignen sich Ultraschallmessungen, wie wir sie in Kap. 5.1 diskutiert haben.

In Kapitel 2 wurde in Abb. 2.20 bereits der schematische Aufbau eines Streuexperiments skizziert. In Abb. 5.11 ist nun etwas eingehender der experimentelle Aufbau eines sogenannten Drei-Achsen-Spektrometers für Neutronenstreuung gezeigt. Nachdem die Neutronen des Kernreaktors durch den Moderator auf thermische Energien abgebremst wurden, fallen sie auf einen Einkristall, der durch Bragg-Reflexion einen monoenergetischen Neutronenstrahl erzeugt. Durch Drehen des Kristalls läßt sich der geeignete Reflex und damit die gewünschte Energie einstellen. Nach der Streuung in der Probe fallen die Neutronen erneut auf einen Einkristall und erfahren eine weitere Bragg-Reflexion. Infolge der Energieänderung beim inelastischen Streuprozeß in der Probe tritt eine kleine Änderung des Bragg-Winkels am Analysatorkristall auf, die gemessen wird. Um den Streuwinkel selbst zu variieren, werden der Analysatorkristall und der ^{3}He-Detektor um die Probe geschwenkt. Natürlich werden auch noch andere Meßtechniken angewandt, bei denen sich die Versuchsaufbauten von dem hier skizzierten unterscheiden. Ein bekanntes Beispiel sind die Flugzeitspektrometer („TOF“ vom Englischen „Time of Flight“), bei denen aus dem kontinuierlichen Neutronenstrahl durch „Zerhacker“ (engl. „Chopper“) Neutronenpakete mit wohldefinierter Energie erzeugt werden. Nachdem die Neutronen in der Probe Energie gewonnen oder verloren haben, fliegen sie anschließend mit unterschiedlicher Geschwindigkeit weiter und treffen zu unterschiedlichen Zeitpunkten an den Detektoren ein. Aus der Flugzeit läßt sich unmittelbar die Energieänderung ermitteln.

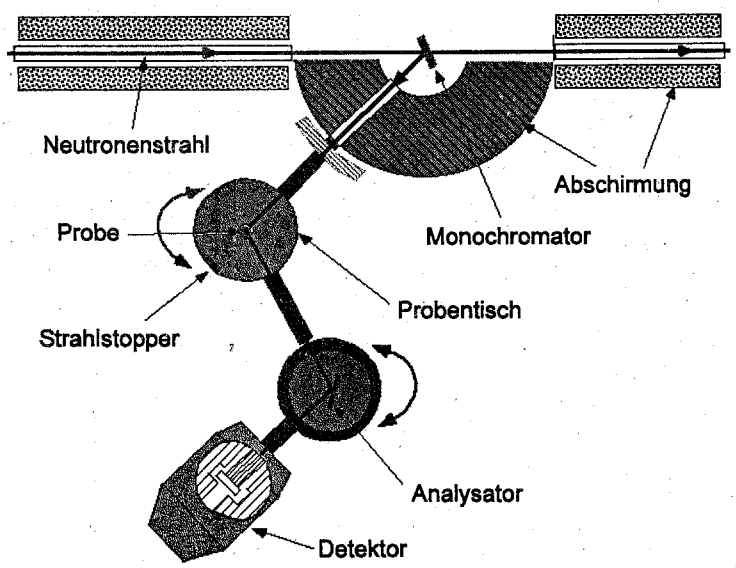


Abb. 5.11: Aufbau eines Drei-Achsen-Neutronenspektrometers. (Die Zeichnung wurde nach Angaben des ILL, Grenoble, für das Instrument IN3 erstellt).

Bei gegebener Beobachtungsrichtung werden Neutronen detektiert, die an unterschiedlichen reziproken Gittervektoren unter Beteiligung von wohldefinierten, aber unterschiedlichen Phononen gestreut wurden. Die beobachteten Neutronen besitzen daher verschiedene diskrete Energien, die mit Hilfe des Analysatorkristalls bzw. der Flugzeit aufgelöst und den entsprechenden Phononen zugeordnet werden können. In Abb. 5.12 wird dieser Aspekt der Neutronenstreuung nochmals im Impulsraum dargestellt. Bei der vorgegebenen Geometrie erfolgt elastische Streuung an den $(\bar{1}40)$-Ebenen, die uns aber nicht weiter interessiert, denn die gestreuten Neutronen werden in einer davon abweichenden Richtung

beobachtet. Es sind drei inelastische Streuprozesse eingezeichnet, an denen die reziproken Gittervektoren (020), (030) und (130) beteiligt sind. Von diesen Vektoren ist nur der reziproke Gittervektor (020) eingezeichnet. Bei der inelastischen Streuung gewinnen oder verlieren die Neutronen an Energie, je nachdem ob ein Phonon vernichtet oder erzeugt wird. Endet der Wellenvektor innerhalb der Ewald-Kugel wie bei den Prozessen 2 und 3, so wurde Energie verloren und somit ein Phonon erzeugt. Beim Streuprozeß 4 dagegen wurde ein Phonon vernichtet.

Die bisherige Diskussion könnte den Eindruck vermitteln, daß die gestreuten Teilchen ein breites Energiespektrum aufweisen, da die Wechselwirkung mit Phononen unterschiedlicher Energie erfolgt. Tatsächlich ist dies jedoch nicht der Fall: Durch die Beobachtungsrichtung wird bei festem reziproken Gittervektor nach (5.37) der Wellenvektor $\boldsymbol{q}$ des beteiligten Phonons festgelegt. Da auch die Energieerhaltung (5.36) erfüllt sein muß, verursacht jeder Phononenzweig eine wohldefinierte Energieänderung des gestreuten Neutrons, die abhängig vom Streuprozeß positiv oder negativ sein kann. Zusätzlich muß noch berücksichtigt werden, daß an der Streuung unterschiedliche reziproke Gittervektoren beteiligt sein können.

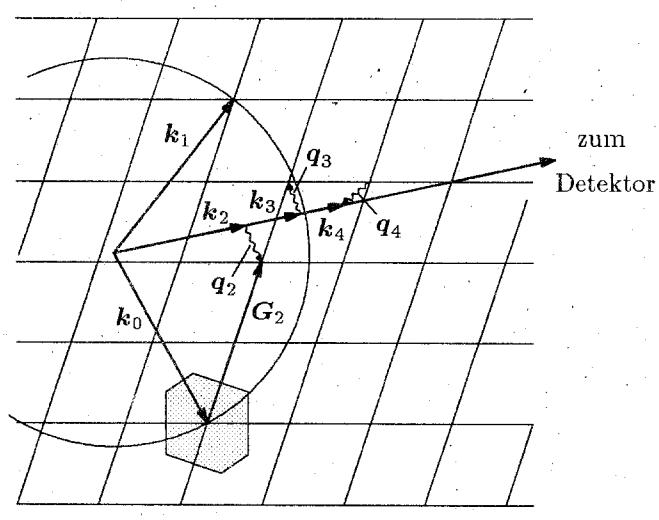


Abb. 5.12: Inelastische Streuung. Bei den Ereignissen 2 und 3 werden Phononen erzeugt, beim Streuereignis 4 wird ein Phonon vernichtet. Die Brillouin-Zone ist grau gezeichnet.

Ursache des Debye-Waller-Faktors

Neben der hier diskutierten *Ein-Phonon-Streuung* gibt es noch Streuprozesse höherer Ordnung, bei denen mehrere Phononen mitwirken. Es läßt sich zeigen, daß die Potenz, mit der die Auslenkung in der Entwicklung (5.34) vorkommt, identisch mit der Zahl der am Streuprozeß beteiligten Phononen ist. Da die Reihe rasch konvergiert, sind nur die ersten Terme von Bedeutung. Die Wahrscheinlichkeit für das Auftreten von *Mehr-Phononen-Streuprozessen* nimmt daher sehr rasch mit der Zahl der beteiligten Phononen ab. Die direkte Beobachtung von diesen Prozessen ist kaum möglich, doch sind sie für die Intensitätsabnahme der Bragg-Reflexe mit zunehmender Temperatur verantwortlich. In Kap. 2.4 haben wir bereits auf die überraschende Temperaturabhängigkeit der Streuintensität hingewiesen, die durch den Debye-Waller-Faktor (2.21) beschrieben wurde. Nun können wir eine Erklärung hierfür geben: Vom gestreuten Teilchen werden Impuls und Energie auf die Phononen übertragen. Während bei Ein-Phonon-Prozessen der Übertrag einen wohldefinierten Wert besitzt, können bei Mehr-Phononen-Prozessen Energie und

Quasiimpuls beliebig zwischen den mitwirkenden Phononen verteilt werden. Die für die Ein-Phonon existierende eindeutige Zuordnung von Streuvektor und Impuls eines Phonons verschwindet damit. Die Streuung unter Einbeziehung von mehreren Phononen trägt daher nur zum „Streuuntergrund" bei.

Wir wollen hier nur den Einfluß von Zwei-Phononen-Prozessen auf die Streuintensität etwas genauer betrachten, da Streuprozesse, an denen mehr als zwei Phononen beteiligt sind, wegen $\boldsymbol{K} \cdot \boldsymbol{U} \ll 1$ keine merkliche Rolle spielen. Hierzu setzen wir die Entwicklung (5.34) in den Ausdruck (5.33) für die Streuamplitude ein und bilden den zeitlichen Mittelwert des Amplitudenquadrats. Nach der Mittelungsprozedur bleibt im wesentlichen nur der Faktor $\prod_{\boldsymbol{q},j}(1 - |\boldsymbol{K} \cdot \boldsymbol{U}(\boldsymbol{q},j)|^2)$ übrig, mit dem die Intensität der gestreuten Strahlung multipliziert werden muß. Benutzt man ein Standardtheorem der Analysis, so läßt sich das Produkt in eine Summe überführen, die der Entwicklung in eine Potenzreihe entspricht, und man findet $\prod_{\boldsymbol{q},j}(1 - |\boldsymbol{K} \cdot \boldsymbol{U}(\boldsymbol{q},j)|^2) = \exp[-\sum_{\boldsymbol{q},j} |\boldsymbol{K} \cdot \boldsymbol{U}(\boldsymbol{q},j)|^2]$. Der zusätzliche Faktor aufgrund der Zwei-Phononen-Prozesse bewirkt also eine Reduktion der elastischen und inelastischen Streuintensität gemäß der Gleichung (2.21), wobei im Falle richtungsunabhängiger Auslenkungsamplituden der Zusammenhang

$$\sum_{\boldsymbol{q},j} |\boldsymbol{K} \cdot \boldsymbol{U}(\boldsymbol{q},j)|^2 = K^2 B(T) \qquad (5.39)$$

besteht. Die Summation erstreckt sich über alle Wellenvektoren und Zweige der thermisch angeregten Phononen. Weil die Zahl der Phononen von der Temperatur abhängt, ist auch $B(T)$ eine Funktion der Temperatur, die wir bereits in Kap. 2.4 angesprochen haben. Da die Summation erst im nächsten Abschnitt behandelt wird, verzichten wir auf die Herleitung der Temperaturabhängigkeit des Debye-Waller-Faktors.

Experimentelle Dispersionskurven

Anhand von experimentell mit Hilfe der inelastischen Neutronenstreuung ermittelten Dispersionskurven von Kupfer und von Lithiumfluorid wollen wir die vorausgehenden theoretischen Überlegungen überprüfen. Kristalle mit einatomiger Basis weisen nur akustische Zweige auf. Als Beispiel hierfür sind in Abb. 5.13 die Dispersionskurven von Kupfer gezeigt. Aufgetragen sind die Phononzweige in [010]-, [110]- und [111]-Richtung des kubisch flächenzentrieren Gitters. Die Kurvenverläufe machen deutlich, daß für diese ausgezeichneten Richtungen die Näherung mit Hilfe einer linearen Kette bereits zu guten Ergebnissen führt. In [010]- und [111]-Richtung sind die beiden senkrecht zueinander polarisierten transversalen Zweige entartet. Erfolgt die Ausbreitung der Gitterwellen jedoch nicht in eine Richtung mit hoher Symmetrie, so treten wie z.B. in [110]-Richtung drei unterschiedliche Zweige auf. Die durchgezogenen Kurven in der Abbildung sind theoretische Dispersionskurven, deren Übereinstimmung mit den experimentellen Daten beeindruckend ist. Bei ihrer Berechnung wurde die Wechselwirkung bis zu den viertnächsten Nachbarn und die kubisch flächenzentrierte Gitterstruktur von Kupfer berücksichtigt.

Bei der Darstellung fällt auf, daß die Dispersionkurven abhängig von der Ausbreitungsrichtung der Gitterwellen bei verschiedenen Werten des Wellenvektors enden. Der Grund für diese Art der Darstellung soll anhand von Abb. 5.14 erläutert werden. Dort sind zwei Brillouin-Zonen eines kubisch flächenzentrierten Gitters gezeichnet, die um einen primitiven Translationsvektor des reziproken Giters gegeneinander verschoben sind. Diese Aneinanderreihung ist sinnvoll, weil der reziproke Raum periodisch ist. Ebenfalls eingezeichnet sind zwei Elementarzellen des reziproken Gitters, das zum nicht-primitiven

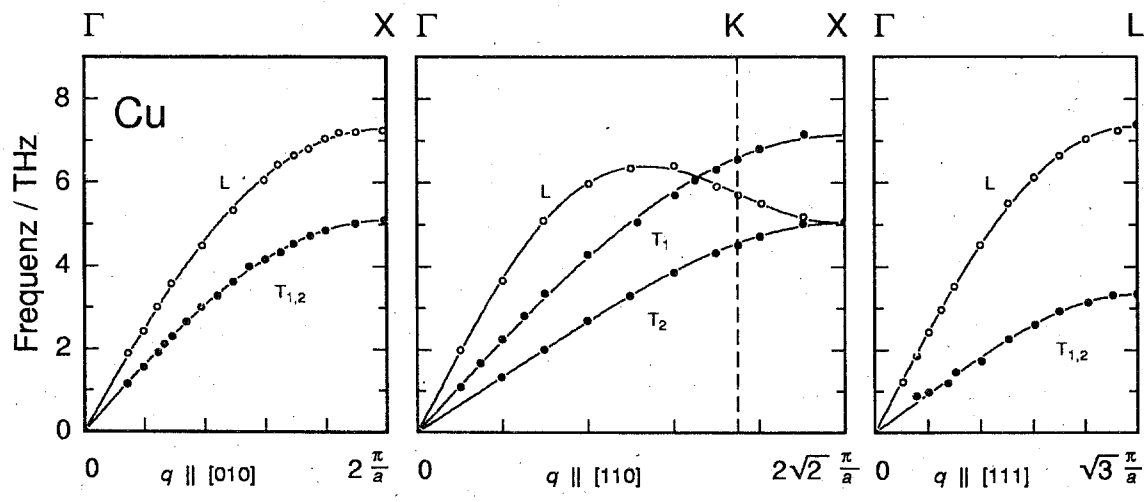


Abb. 5.13: Phonon-Dispersionskurven von Kupfer. Die durchgezogenen Kurven geben berechnete Dispersionsverläufe wieder, die Meßwerte sind durch Kreise und Punkte dargestellt. Mit L, T_1 und T_2 sind die longitudinalen und die beiden transversalen Zweige gekennzeichnet. Die gestrichelte Linie gibt das Ende der Brillouin-Zone in [110]-Richtung an. (Nach E.C. Svensson, B.N. Brockhouse, J.M. Rowe, Phys. Rev. **155** (1967) 619).

kubisch flächenzentrierten Gitter gehört. Der Mittelpunkt der ersten Elementarzelle fällt mit dem Γ-Punkt zusammen, die zweite ist um den (nicht-primitiven) Reziprokengittervektor $(\frac{1}{2}, \frac{1}{2}, -\frac{1}{2})$ dagegen verschoben. Die Kantenlänge des Würfels der reziproken Elementarzelle ist $4\pi/a$. Geht man nun, wie in Abb. 5.14 angedeutet, in [010]-Richtung, so trifft man nach der Strecke $2\pi/a$ am X-Punkt auf die Grenze der Brillouin-Zone. In [111]-Richtung erreicht man den L-Punkt und damit die Grenze der Brillouin-Zone beim Wert $\sqrt{3}\pi/a$. In [110]-Richtung verläßt man die Brillouin-Zone am K-Punkt. Die Dispersionskurven in Abb. 5.13 sind bis zum X-Punkt der wiederholten Brillouin-Zone fortgesetzt, bei dem der Wellenvektor dann den Wert $2\sqrt{2}\pi/a$ besitzt. An der Grenze der Brillouin-Zone, beim K-Punkt, tritt in dieser Richtung kein Extremum der Dispersionskurve auf, weil die Brillouin-Zone an einer Kante verlassen wird, die mehreren Zonen angehört.

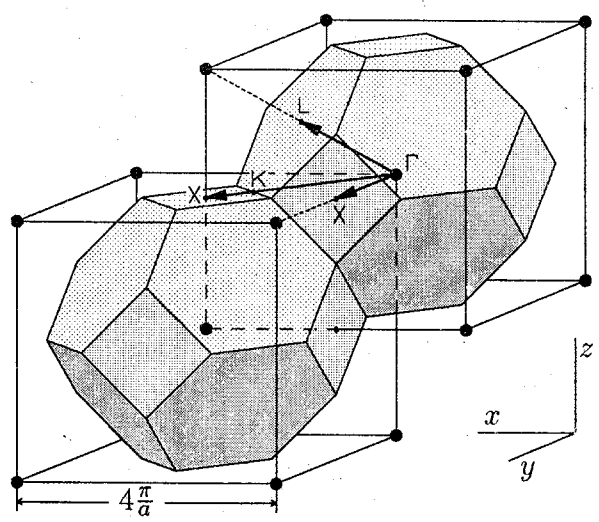


Abb. 5.14: Brillouin-Zone eines kubisch flächenzentrierten Kristalls. Es sind zwei benachbarte Brillouin-Zonen und weiter zwei nicht-primitive Einheitszellen dargestellt. Die Pfeile geben die Richtungen an, längs derer die in Abb. 5.13 gezeichneten Kurven gemessen wurden.

In Abb. 5.15 sind die theoretischen und experimentellen Dispersionskurven des kubisch flächenzentrierten Lithiumfluorids zu finden, dessen Basis aus den beiden Ionen Li^+ und F^- besteht. Bemerkenswert ist wieder die gute Übereinstimmung zwischen Theorie und Experiment, insbesondere, da nur die Wechselwirkung bis zu den übernächsten Nachbarn berücksichtigt wurde. Neben den akustischen treten aufgrund der zweiatomigen Basis noch drei optische Phononenzweige auf. Auffällig ist die starke Aufspaltung der optischen Zweige bei $\boldsymbol{q} = 0$, die für Ionenkristalle typisch ist. Der longitudinale Zweig liegt weit über dem transversalen. Diese Aufspaltung ist eine Folge der Coulomb-Wechselwirkung zwischen den Untergittern. Ihre lange Reichweite führt bei den longitudinalen Schwingungen zu einer makroskopischen Polarisation, die wir eingehend in Kap. 11.2 behandeln werden. Besitzt dagegen die Wechselwirkung zwischen den Atomen nur kurze Reichweite, so sind bei kubischen Kristallen am Γ-Punkt alle optischen Schwingungen entartet. Die hohe Symmetrie bewirkt, daß die Kräfte unabhängig von der Auslenkungsrichtung sind. Ein Beispiel hierfür ist das Schwingungsspektrum von Diamant, das in Abb. 5.24 im Zusammenhang mit der Diskussion der spezifischen Wärme gezeigt ist.

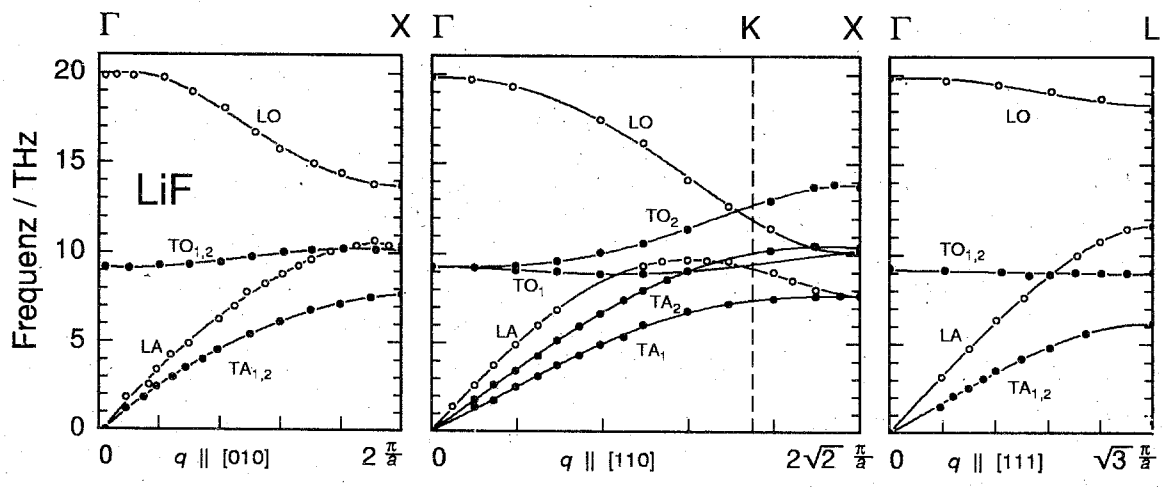


Abb. 5.15: Phononen-Dispersionskurve von Lithiumfluorid. Es sind die theoretischen und experimentellen Verläufe wiedergegeben. Die Phononenzweige sind mit LA, TA, LO und TO bezeichnet, wobei der erste Buchstabe für die Polarisation und der zweite für „akustisch" oder „optisch" steht. (Nach G. Dolling, H.G. Smith, R.M. Nicklow, P.R. Vijayaraghavan, M.K. Wilkinson, Phys. Rev. **168** (1968) 970).

Lichtstreuung

Bisher sind wir immer davon ausgegangen, daß bei Streuexperimenten die Wellenlänge der einfallenden Strahlung kleiner als der Gitterabstand ist. Erwartet man auch Streuung im umgekehrten Fall? Wie wir sehen werden, ist dies durchaus möglich. Für *sichtbares* Licht, mit dem wir uns hier befassen, läßt sich die Streubedingung (5.37) vereinfachen. Da die Wellenzahl k_0 des sichtbaren Lichts etwa $10^5\,\mathrm{cm}^{-1}$, der kleinste reziproke Gittervektor $g \approx 10^8\,\mathrm{cm}^{-1}$ ist, gilt in allen Fällen die Beziehung $k_0 \ll g$. Daher spielt bei Lichtstreuexperimenten das reziproke Gitter keine Rolle, d.h. die Impulserhaltung (5.37) kann nur erfüllt werden, wenn $\boldsymbol{G} = 0$ ist. In einer klassischen Beschreibung des Streuvorgangs (vgl. Kap. 2.2) ruft die einfallende Welle in jedem Atom ein oszillierendes Dipolmoment hervor, das dann selbst eine Welle abstrahlt, deren Intensität nach einer relativ einfachen elektrodynamischen Rechnung proportional zu ω^4 ansteigt.

Rayleigh-Streuung

Findet die Wechselwirkung des Lichts mit dem Festkörper ohne Phononenbeteiligung statt, so dürfen wir in (5.37) zusätzlich $\hbar\boldsymbol{q} = 0$ setzen. Diesen *elastischen* Streuprozeß, bei dem keine Frequenzverschiebung auftritt, bezeichnet man als *Rayleigh-Streuung*. Aus (5.37) folgt unmittelbar die einfache, aber erstaunliche Beziehung $\boldsymbol{k}_0 = \boldsymbol{k}$, d.h. es tritt nur Streuung in *Vorwärtsrichtung* auf. Die Lichtwelle läuft durch den Kristall in Einfallsrichtung mit einer Geschwindigkeit, die durch den Brechungsindex bestimmt wird, doch sollte in den übrigen Raumrichtungen kein Streulicht zu beobachten sein. Oder anders ausgedrückt: in idealen Kristallen löschen sich die Streubeiträge der einzelnen Atome, außer in Vorwärtsrichtung, exakt durch Interferenz aus. Die Quasiimpulserhaltung (5.37) beruht auf der Translationssymmetrie, die aber durch Gitterdefekte gestört wird. Sie wirken als Streuzentren, deren Streuwellen sich nicht wegmitteln, so daß in realen Kristallen die Rayleigh-Streuung dennoch beobachtet werden kann. Entsprechendes gilt für amorphe Festkörper, da dort die Erhaltung der Wellenvektoren überhaupt nicht streng gültig ist.

Raman- und Brillouin-Streuung

Ist $\boldsymbol{q} \neq 0$, so werden beim Streuprozeß Phononen erzeugt oder vernichtet, wie im Diagramm der Abb. 5.16 zu sehen ist. Verschiebt sich die Wellenlänge des gestreuten Lichts zu größeren Wellenlängen, so spricht man von einem *Stokes-Prozeß*, wird es kurzwelliger, so handelt es sich um einen *Anti-Stokes-Prozeß*. Die Streuung unter Mitwirkung von *akustischen* Phononen bezeichnet man als *Brillouin-* und die Streuung an *optischen* Phononen als *Raman-Streuung*.

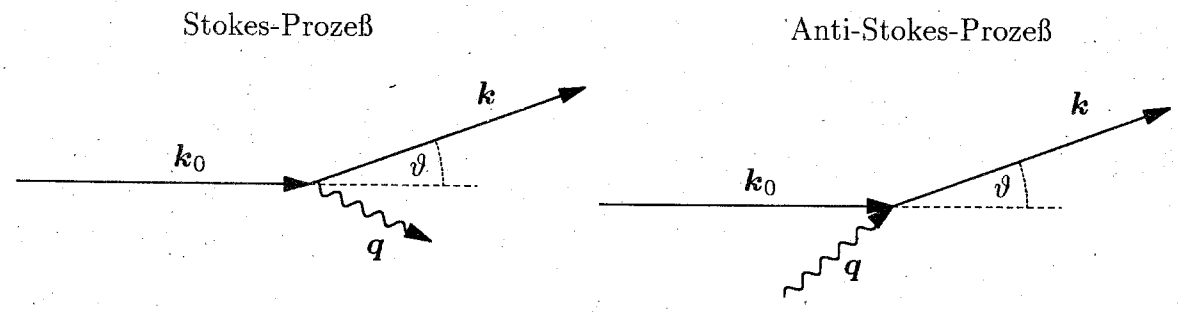


Abb. 5.16: Darstellung der Wellenvektoren bei Lichtstreuung. Links wird ein Phonon erzeugt, rechts dagegen ein Phonon vernichtet.

Gehen wir vereinfachend von einem isotropen Medium aus, so läßt sich der Impuls durch $q^2 = k_0^2 + k^2 - 2kk_0 \cos\vartheta$ ausdrücken. Weil die Energie der Phononen gegenüber der Photonenenergie vernachlässigt werden kann, ändert der Streuprozeß die Energie und den Betrag des Impulses der Photonen kaum. Somit dürfen wir $k_0 \approx k$ setzen, woraus $q^2 \approx 2k_0^2(1 - \cos\vartheta) = 4k_0^2 \sin^2(\vartheta/2)$ folgt. Für die Wellenzahl der Phononen finden wir schließlich

$$q = 2k_0 \sin\left(\frac{\vartheta}{2}\right) . \tag{5.40}$$

Aufgrund des kleinen Wellenvektors k_0 von Licht sind nur Phononen in der unmittelbaren Nachbarschaft des Γ-Punktes an der Lichtstreuung beteiligt. Bei der inelastischen Neutronenstreuung sind dagegen alle Wellenvektoren zugänglich. Die kleine Frequenzänderung $\Delta\omega = \omega - \omega_0 = \pm\omega_{\boldsymbol{q}}$ wird mit hochauflösenden Spektrometern nachgewiesen.

Betrachten wir zunächst die Lichtstreuung an optischen Phononen, also die *Raman-Streuung*, etwas genauer. Da die Dispersionskurve dieser Phononen am Γ-Punkt ein Extremum aufweist, hängt die Frequenz der wechselwirkenden Phononen kaum vom Wellenvektor ab, und die beobachtete Linienverschiebung ist von der Beobachtungsrichtung unabhängig. Typische Frequenzänderungen $\Delta\nu = \Delta\omega/2\pi$ liegen im Bereich von einigen hundert bis etwa $1000\,\mathrm{cm}^{-1}$, d.h. sie sind von der Größenordnung 10 THz. Bei Raman-Streuexperimenten durchstrahlt man meist die Probe mit einem Laser und analysiert das unter einem bestimmten Winkel gestreute Licht mit einem Doppelmonochromator.

In den beiden folgenden Abbildungen sind Messungen der Raman-Streuung an Germanium und Silizium gezeigt. Diese Kristalle besitzen eine zweiatomige Basis und weisen somit drei optische Phononenzweige auf. Wie bei allen nicht-polaren kubischen Kristallen fallen auch bei diesen Kristallen am Γ-Punkt die Dispersionskurven der longitudinalen und transversalen optischen Phononen zusammen, so daß man nur eine Stokes- bzw. Anti-Stokes-Linie beobachtet. In der Abb. 5.17 sind die Raman-Linien von vier verschiedenen Germaniumkristallen zu sehen. Einer der Kristalle besteht aus natürlichem Germanium, während die drei übrigen isotopenrein sind. Wie Gleichung (5.27) erwarten läßt, ist die Frequenz der optischen Phononen am Γ-Punkt und damit auch die Raman-Verschiebung proportional zu $M^{-1/2}$. Allerdings muß man bei einer sehr genauen Analyse zusätzlich berücksichtigen, daß leicht unterschiedliche Anharmonizitäten und die Isotopenunordnung beim natürlichen Germanium ebenfalls Einfluß auf die Lage der Phononenfrequenzen besitzen.

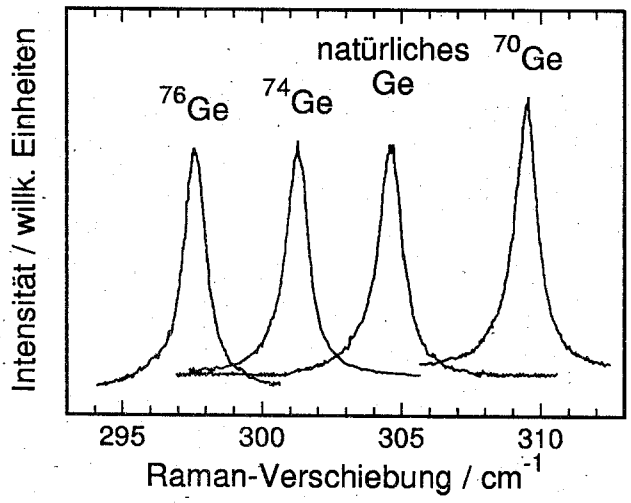


Abb. 5.17: Raman-Spektren von isotopenreinem und natürlichem Germanium bei 80 K. Die Verschiebung der Raman-Linie ist näherungsweise proportional zu $M^{-1/2}$. Die unverschobene Rayleigh-Linie ist nicht zu sehen, da der Nullpunkt der Frequenzverschiebung unterdrückt ist. (Nach T. Ruf, H.D. Fuchs und M. Cardona, Phys. Bl. **52** (1996) 1115).

In Abb. 5.18 ist die Messung des Raman-Spektrums an Silizium bei unterschiedlichen Temperaturen gezeigt. Aufgrund der leichten Anharmonizität des Gitters (vgl. Kap. 6.1) hängt sowohl die Lage als auch die Breite dieser Linien von der Temperatur ab. Die Streuintensität, gegeben durch die Fläche der Kurven, zeigt deutlich erkennbar eine Asymmetrie bezüglich der Stokes- und Anti-Stokes-Linien[15]. Da 15 THz einer thermischen Energie von etwa 700 K entsprechen, sind selbst bei 770 K nur recht wenige Phononen mit dieser hohen Frequenz angeregt, die für den Anti-Stokes-Prozeß erforderlich sind. Bei

[15] Eine eingehendere Betrachtung der Lichtstreuung führt zu dem Ergebnis, daß die Intensität der Anti-Stokes-Linien proportional zum Bose-Einstein-Faktor $\langle n(\omega,T)\rangle$ (vgl. Abschnitt 5.5), die Intensität der Stokes-Linien aber proportional zu $\langle n(\omega,T)\rangle + 1$ ist.

20 K ist diese Linie überhaupt nicht mehr beobachtbar, da bei so tiefen Temperaturen Phononen dieser Energie vollständig fehlen.

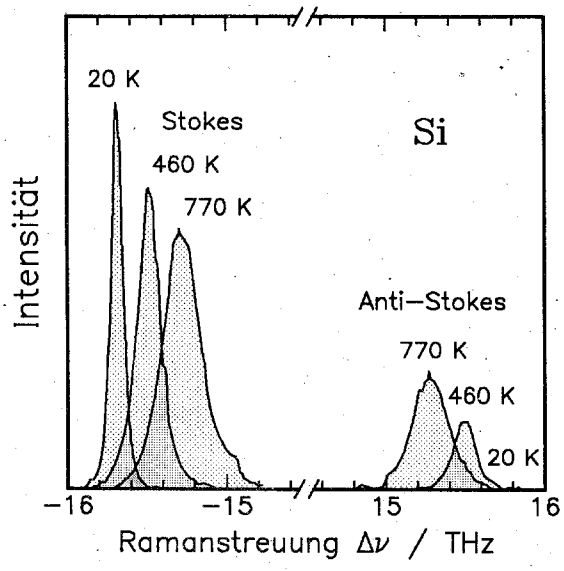


Abb. 5.18: Raman-Spektrum von Silizium bei unterschiedlichen Temperaturen. Der Nullpunkt der Frequenzverschiebung ist unterdrückt. (Nach T.R. Hart, R.L. Aggarwal, B. Lax, Phys. Rev. B 1 (1970) 683).

In den Raman-Spektren werden aber nicht alle Streulinien beobachtet, die aufgrund von Energie- und Impulserhaltung (5.36) und (5.37) erlaubt wären. Voraussetzung für ihr Auftreten ist eine nicht verschwindende Kopplung zwischen Phononen und elektromagnetischen Wellen. Wie bereits erwähnt, ruft die einfallende Lichtwelle über die elektrische Suszeptibilität (vgl. auch Kap. 11.2) der Atome eine mit der Lichtfrequenz oszillierende Polarisation hervor. Zur Raman-Streuung kommt es, wenn die Gitterschwingungen eine Modulation dieser Polarisation bewirken. Klassisch betrachtet ist damit die Abstrahlung einer elektromagnetischen Welle mit der Summen- bzw. Differenzfrequenz der beiden Schwingungen verbunden, wobei die gestreute Welle eine charakteristische Richtungs- und Polarisationsabhängigkeit aufweist. Bei der Infrarotabsorption erfolgt die Kopplung über das oszillierende Dipolmoment, das von den Phononen hervorgerufen wird. Mit Hilfe von Symmetrieüberlegungen lassen sich für beide Effekte die Auswahlregeln herleiten. Für Kristalle mit Inversionssymmetrie gilt streng, daß sich Infrarotabsorption und Raman-Streuung gegenseitig ausschließen. So beobachtet man bei Ionenkristallen mit NaCl-Struktur eine starke Infrarotabsorption, aber keine Raman-Streuung, während es bei Silizium gerade umgekehrt ist.

Eine modifizierte $\boldsymbol{k}$-Erhaltung (5.37) ist gültig, wenn zwei Phononen am Streuprozeß beteiligt sind. Die Bedingung für die Wellenvektoren lautet dann $\boldsymbol{k} - \boldsymbol{k}_0 = \boldsymbol{q}_1 + \boldsymbol{q}_2$. Sind die Wellenvektoren der beiden Phononen nahezu entgegengerichtet, d.h. ist $\boldsymbol{q}_1 \approx -\boldsymbol{q}_2$, so kann der Quasiimpuls für die Lichtstreuung erfüllt werden, selbst wenn am Streuprozeß Phononen mit großen Wellenzahlen beteiligt sind. Die Einschränkung des Streuprozesses auf die engste Umgebung des Γ-Punktes existiert dann nicht mehr, doch sind bei Beteiligung von zwei Phononen die auftretenden Streuintensitäten aus Gründen, die wir bei der Diskussion des Debye-Waller-Faktors angeführt haben, vergleichsweise gering. Dennoch ist die Raman-Streuung 2. Ordnung beobachtbar und erlaubt eine Bestimmung der Zustandsdichte (vgl. Kap. 5.5) der Phononen.

Da die Quasiimpulserhaltung (5.37) auf der Translationsinvarianz der Kristalle beruht, gilt sie nicht für amorphe Festkörper, so daß in diesen Materialien alle Phononen

am Raman-Prozeß 1. Ordnung teilnehmen können. Folglich treten keine wohldefinierten, engen Linien wie in Kristallen auf, sondern stark verbreiterte Maxima. Zusätzlich entfallen die Symmetrieargumente, die bei vielen Kristallen das gleichzeitige Auftreten von Infarotabsorption und Raman-Streuung ausschließen. In amorphen Festkörpern kann man daher beide Effekte gleichzeitig beobachten.

Wir betrachten nun die *Brillouin-Streuung* etwas genauer. Aus Gründen der Quasiimpulserhaltung können auch bei dieser Streuung nur langwellige akustische Phononen beteiligt sein. Da für diese nach (5.24) die Dispersionsrelation $\omega = vq$ gilt, hängt bei der Brillouin-Streuung die Frequenz der streuenden Phononen vom Streuwinkel ϑ ab. Wie aus (5.40) folgt und Abb. 5.19 zeigt, liegen die typischen Frequenzänderungen $\Delta\nu$ bei dieser Streuung im Bereich um 20 GHz.

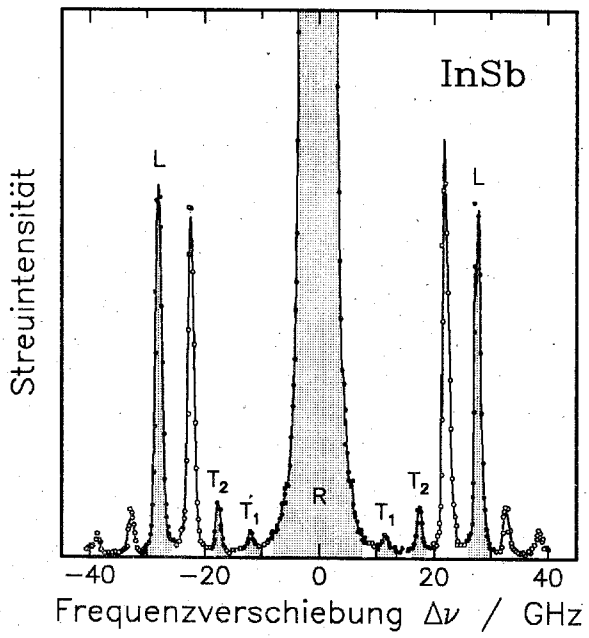


Abb. 5.19: Brillouin-Spektrum von Indiumantimonid. Die Meßwerte, die zu einem „Liniensatz" gehören, sind grau markiert und mit R, L bzw. T bezeichnet. (Nach J.R. Sandercock, in Proc. 2nd Int. Conf. Light Scattering in Solids, M. Balkanski (ed.) Flammarion Paris 1971, p. 9).

Die relativ kleine Frequenzänderung stellt hohe Anforderungen an die Frequenzstabilität des Lasers und vor allem an die Auflösung des verwendeten Spektrometers. Für derartige Messungen benutzt man Fabry-Perot-Interferometer, die in hoher Ordnung betrieben werden. Die Spektren wiederholen sich deshalb periodisch als Funktion des Plattenabstands und überschneiden sich bei den meisten Experimenten teilweise. Von Interesse ist aber immer nur ein „Satz von Linien", der zu einer Ordnung gehört und in Abb. 5.19 durch dunkle Tönung hervorgehoben ist. In dem gezeigten Experiment wurden neben der Rayleigh-Linie (R), die durch experimentelle Vorkehrungen um viele Größenordnungen unterdrückt ist, alle drei akustischen Phononenzweige L, T_1 und T_2 beobachtet.

Es besteht ein enger formaler Zusammenhang zwischen der Brillouin-Streuung und der Bragg-Reflexion, wie ein kurzer Vergleich sofort zeigt. Entsprechend der Festlegung des Streuwinkels θ in der Röntgenbeugung und ϑ bei der Lichtstreuung, gilt $\vartheta/2 = \theta$. Damit läßt sich Gleichung (5.40), die für die Lichtstreuung gültig ist, wie folgt umschreiben:

$$q = 2k_0 \sin\theta \qquad \text{oder} \qquad \lambda_0 = 2\lambda_q \sin\theta\,. \tag{5.41}$$

Hierbei ist λ_0 die Licht-[16] und λ_q die Schallwellenlänge. Der Vergleich mit der Bragg-

[16] Es ist zu beachten, daß mit λ_0 die Wellenlänge des Lichts in der Probe gemeint ist, so daß also noch der Brechungsindex des Festkörpers berücksichtigt werden muß.

Streuung (2.17) zeigt, daß bei der Lichtstreuung die Variationen der Streudichte durch die Gitterwellen hervorgerufen werden, die somit den Gitterebenen bei der Bragg-Streuung entsprechen. In diesem einfachen Bild wird die Frequenzverschiebung des Lichts durch den Doppler-Effekt an den laufenden Schallwellen hervorgerufen. Die Schallwellen modulieren periodisch den Brechungsindex, wodurch es zur Beugung des einfallenden Lichts kommt. Da im allgemeinen die Phononen nur kleine Schwankungen des Brechungsindexes hervorrufen, ist die Intensität des gestreuten Lichts meist sehr gering.

5.5 Spezifische Wärmekapazität

Die spezifische Wärme[17] ist eine wichtige thermodynamische Größe, die in einfachen dielektrischen Festkörpern durch die Phononen bestimmt wird. Ihre korrekte Beschreibung ist ein Prüfstein für die Theorie der Gitterschwingungen. Aus experimentellen Gründen wird die spezifische Wärme von Festkörpern meist bei konstantem Druck gemessen, d.h. es wird C_p bestimmt. Zur theoretischen Beschreibung geeigneter ist jedoch die spezifische Wärme bei konstantem Volumen, die über die Gleichung $C_V = (\partial U/\partial T)_V$ direkt mit der inneren Energie U verknüpft ist. Die beiden wichtigen Größen sind über die thermodynamische Beziehung $C_p - C_V = 9\alpha^2 VT/\kappa$ miteinander verbunden, wobei α der lineare thermische Ausdehnungskoeffizient und κ die Kompressibilität ist. Während in Gasen die Differenz $(C_p - C_V)$ relativ groß ist, unterscheiden sich C_V und C_p im Festkörper wegen der geringen thermischen Ausdehnung kaum.

Die Temperaturabhängigkeit wie auch der Absolutwert der spezifischen Wärme weisen einige charakteristische Merkmale auf. Der typische Verlauf ist in Abb. 5.20 festgehalten, die Messungen an Diamant aus dem letzten Jahrhundert wiedergibt. Von tiefen Temperaturen kommend steigt die spezifische Wärme bei allen Kristallen zunächst rasch an und nähert sich dann einem konstanten Wert, der durch das *Dulong-Petit-Gesetz* $C_V = 3N_A k_B = 3R_m \approx 25\,\mathrm{J\,mol^{-1}\,K^{-1}}$ gegeben ist. Hierbei ist N_A die Avogadro- und R_m die Gaskonstante, d.h. die spezifische Wärme ist auf die Stoffmenge bezogen.

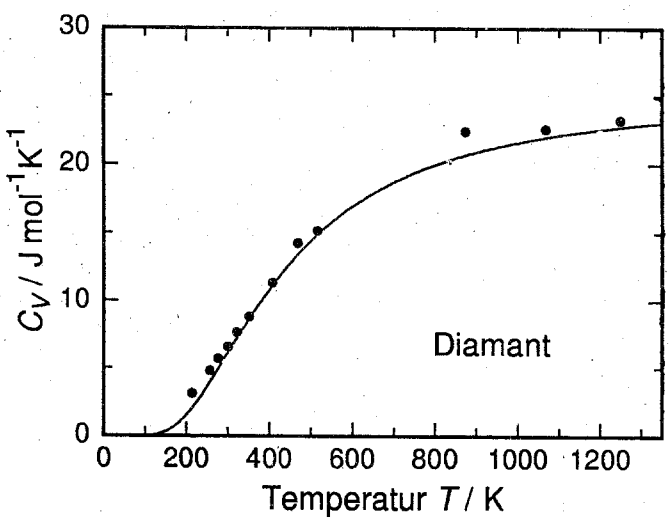


Abb. 5.20: Molare Wärmekapazität von Diamant als Funktion der Temperatur. Die durchgezogene Kurve gibt den von Einstein errechneten Verlauf wieder. (Nach A. Einstein, Ann. Phys. **22** (1907) 180).

[17] Wir unterscheiden hier nicht explizit zwischen der *spezifischen* und *molaren* Wärmekapazität, da sich die jeweilige Bedeutung unmittelbar aus dem Zusammenhang ergibt. Weiterhin beziehen wir die molare Wärmekapazität auf die Teilchenmenge, also auf die Zahl der Atome, und nicht auf die Zahl der Moleküle. Sie hat daher die Dimension $\mathrm{J\,mol^{-1}\,K^{-1}}$.

Der prinzipielle Verlauf, insbesondere die beobachtete starke Abnahme der spezifischen Wärme mit abnehmender Temperatur, wurde 1907 von *Einstein* erklärt. Er nahm an, daß (gleiche) Atome als identische, ungekoppelte, harmonische Oszillatoren aufgefaßt werden können. Entscheidend ist dabei, daß sich die Energie der Oszillatoren nicht dem Gleichverteilungssatz der klassischen Physik folgend kontinuierlich ändern kann, sondern gequantelt ist. Die spezifische Wärme muß bei tiefen Temperaturen rasch abnehmen, denn dort wird die thermische Energie $k_{\mathrm{B}}T$ klein im Vergleich zur Schwingungsenergie der Oszillatoren. Wir wollen hier jedoch nicht näher auf die Einsteinsche Theorie eingehen, die bei tiefsten Temperaturen einen zu kleinen Wert der spezifischen Wärme vorhersagt, sondern gleich berücksichtigen, daß die Atome des Festkörpers untereinander gekoppelt sind. Wie wir in Kap. 5.2 gesehen haben, bewirkt diese Kopplung eine Erweiterung des Frequenzspektrums zu kleinen Frequenzwerten hin, deren Berücksichtigung eine deutliche Verbesserung im Vergleich zum Einstein-Modell bewirkt.

Phononenzustandsdichte

Die bisherige Diskussion in diesem Kapitel könnte den Eindruck erweckt haben, daß die Frequenzen der Gitterschwingungen kontinuierlich verteilt und nur den Dispersionsrelationen unterworfen sind. Die endliche Ausdehnung einer Probe und die damit verbundene endliche Anzahl von Atomen erlauben jedoch nur eine begrenzte Anzahl von Eigenschwingungen, für deren Berechnung die Randbedingungen festgelegt werden müssen. Ein unmittelbar einsichtiges, wenn auch nicht sehr realistisches Vorgehen besteht darin, den Bewegungszustand der Randatome fest vorzugeben, sie z.B. als unbeweglich oder frei beweglich zu betrachten. Wir wollen aber nicht diese sogenannten *festen Randbedingungen*, sondern die *periodischen Randbedingungen* benutzen. Ausgangspunkt ist in diesem Fall ein makroskopischer Kristall endlicher Größe mit der Form eines Parallelepipeds. Wir stellen uns vor, daß sich dieser Kristall in alle Richtungen periodisch wiederholt. Dabei soll in jedem dieser „Kopien" zu jedem Zeitpunkt dieselbe momentane Atomanordnung und somit derselbe Schwingungszustand herrschen, so daß sich alle Eigenschaften im Raume periodisch wiederholen. Dadurch entsteht ein unendlich ausgedehnter Gesamtkristall mit der vollen Translationssymmetrie eines idealen Kristalls. Der Einfluß der endlichen Ausdehnung des ursprünglich betrachteten Kristall auf die physikalischen Eigenschaften ist dadurch in der theoretischen Beschreibung unterdrückt.

Bei isotropen oder kubischen Festkörpern ist ein Würfel mit der Kantenlänge L und $\mathcal{M}$ Elementarzellen pro Kantenlänge ein einfaches und geeignetes Parallelepiped. Die atomare Auslenkung $\boldsymbol{u}(\boldsymbol{r},t) = \boldsymbol{U}\,\mathrm{e}^{-\mathrm{i}(\omega_q t - \boldsymbol{q}\cdot\boldsymbol{r})}$, die jedes Phonon hervorruft, unterliegt der periodischen Randbedingung:

$$u(x,y,z,t) = u(x+L,y,z,t) = u(x,y+L,z,t) = u(x,y,z+L,t)\,. \tag{5.42}$$

Setzt man in diese Gleichungen den Ausdruck für die Auslenkung der Atome ein, so erhält man beispielsweise für die x-Richtung

$$\mathrm{e}^{\mathrm{i}\boldsymbol{q}\cdot\boldsymbol{r}} = \mathrm{e}^{\mathrm{i}\boldsymbol{q}\cdot(x+L,y,z)} = \mathrm{e}^{\mathrm{i}\boldsymbol{q}\cdot\boldsymbol{r}}\cdot\mathrm{e}^{\mathrm{i}q_x L}\,. \tag{5.43}$$

Mit den beiden anderen Richtungen ergeben sich daraus für die Wellenvektoren $\boldsymbol{q}$ die Einschränkungen

$$q_x = \frac{2\pi}{L}\,m_x\,, \qquad q_y = \frac{2\pi}{L}\,m_y\,, \qquad q_z = \frac{2\pi}{L}\,m_z\,, \tag{5.44}$$

wobei m_x, m_y und m_z ganze Zahlen sind. Diese Ausdrücke gelten für akustische und optische Phononen unabhängig von ihrer Polarisationsrichtung, also für alle Zweige.

Wie wir in Kap. 5.2 gefunden haben, gibt es bei $\mathcal{M}^3 = \mathcal{N}$ Elementarzellen mit jeweils p Atomen insgesamt $3p\mathcal{N} = 3N$ Eigenschwingungen. Jeder Phononenzweig enthält also $\mathcal{N}$ Schwingungszustände. Ist $|m^{\max}|$ dem Betrag nach die größte erlaubte Quantenzahl, so muß bei einem Würfel mit $\mathcal{N}$ Elementarzellen die Beziehung $\mathcal{N} = 2m_x^{\max} \cdot 2m_y^{\max} \cdot 2m_z^{\max}$ erfüllt sein. Die Faktoren 2 berücksichtigen, daß positive und negative Quantenzahlen erlaubt sind. Hieraus folgt $m^{\max} = \pm\mathcal{M}/2$, so daß die Quantenzahl m_i der Einschränkung

$$-\frac{\mathcal{M}}{2} < m_{x,y,z} \leq \frac{\mathcal{M}}{2} \tag{5.45}$$

unterliegen. In einem Festkörper sind also genau soviele unterschiedliche Wellenvektoren möglich, wie Atome vorhanden sind. Exakt $\mathcal{N}$ Schwingungszustände pro Zweig erhält man, wenn von den beiden Wellenvektoren mit $m^{\max} = \pm\mathcal{M}/2$ nur einer mitgezählt wird. Diese Vorgehensweise ist dadurch begründet, daß beide Vorzeichen zu identischen Auslenkungen führen. Um dies zu zeigen, betrachten wir eine ebene Welle, die sich in einem primitiven Gitter in x-Richtung bewegt, und greifen die Auslenkung der Elementarzelle s mit der Koordinate $x = sa$ zu einem festen Zeitpunkt heraus. Für die Auslenkung bei den beiden maximalen Wellenzahlen gilt:

$$\begin{aligned} U\mathrm{e}^{\mathrm{i}qx} = U\mathrm{e}^{\mathrm{i}qsa} &= U\exp\left(\pm\frac{\mathrm{i}\pi\,\mathcal{M}}{L}\,sa\right) = U\mathrm{e}^{\pm\mathrm{i}\pi s} \\ &= U\Big[\cos(s\pi) \pm \mathrm{i}\,\sin(s\pi)\Big] = U\cos(s\pi)\,. \end{aligned} \tag{5.46}$$

Die Auslenkung hängt nicht vom Vorzeichen der maximalen Wellenzahl ab.

Der wesentliche Unterschied zwischen festen und periodischen Randbedingungen besteht darin, daß in dem einen Fall stehende und im anderen Fall laufende Wellen zugelassen werden. Da bei stehenden Wellen die Periodizität durch die halbe Wellenlänge vorgegeben ist, liegen die erlaubten Zustände bei festen Randbedingungen pro Raumrichtung doppelt so dicht. Allerdings ist in diesem Fall nur ein Vorzeichen für die erlaubten Wellenvektoren sinnvoll. Da es aber, wie wir gleich sehen werden, in erster Linie auf die Zahl der Zustände in einem Frequenzintervall ankommt, ist diese Unterscheidung zwischen den beiden Randbedingungen bei makroskopischen Proben unerheblich. Ist die Zahl der Atome in einer Probe jedoch klein, dann spielen die exakten Randbedingungen für die Gitterschwingungen eine wichtige Rolle. Bei derartigen Systemen erweist sich meist die exakte Spezifizierung der Randbedingungen bereits als problematisch.

Nun übertragen wir unsere Überlegungen auf nicht-kubische Gitter. Ausgangspunkt ist wieder ein endlicher Kristall mit dem Volumen V und $\mathcal{N}$ Elementarzellen. Dieser Kristall soll in den Richtungen der drei Basisvektoren jeweils $\mathcal{M}$ Elementarzellen aufweisen. Um den periodischen Randbedingungen zu genügen, müssen die Wellenvektoren die zu (5.43) analoge Bedingung

$$\exp\left[\mathrm{i}\mathcal{M}\boldsymbol{q}(\boldsymbol{a}_1 + \boldsymbol{a}_2 + \boldsymbol{a}_3)\right] \equiv 1 \tag{5.47}$$

erfüllen. Ein Vergleich mit der Definition des reziproken Gitters macht klar, daß für die erlaubten Wellenvektoren

$$\boldsymbol{q} = \sum_{i=1}^{3} \frac{m_i\,\boldsymbol{b}_i}{\mathcal{M}} \tag{5.48}$$

gelten muß. Hierbei steht $\boldsymbol{b}_i$ für die Basisvektoren des reziproken Gitters und m_i wieder für eine ganze Zahl mit dem Maximalwert $|m_i^{\max}| = \mathcal{M}/2$. Damit verknüpft ist der maximale Wellenvektor $\boldsymbol{q}_i^{\max} = \pm\boldsymbol{b}_i/2$ in Richtung der Achsen des reziproken Gitters. Die Basisvektoren $\boldsymbol{b}_i$ definieren eine Elementarzelle des reziproken Gitters, die wir so legen, daß ihr Zentrum mit dem Gittervektor (000) zusammenfällt. In diesem Parallelepiped sind die erlaubten Zustände, wie in Abb. 5.21 veranschaulicht, *gleichmäßig verteilt.*

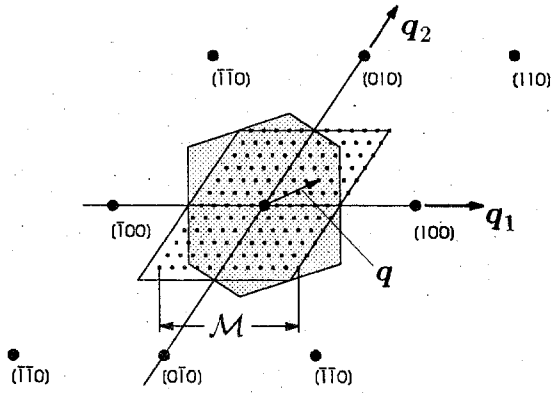


Abb. 5.21: Schnitt durch das reziproke Gitter eines monoklinen Kristalls mit 12 Elementarzellen pro Kantenlänge. Die großen Punkte geben das reziproke Gitter, die kleinen die erlaubten $\boldsymbol{q}$-Vektoren wieder. Zusätzlich sind der Wellenvektor (3,2,0) und die dunkel getönte Brillouin-Zone eingezeichnet.

Es ändert sich nichts an der physikalischen Aussage, wenn wir eine andere Zelle derselben Größe wählen. Für die meisten Betrachtungen erweist sich die Wahl der 1. Brillouin-Zone als besonders vorteilhaft, die in Abb. 5.21 ebenfalls eingezeichnet ist. Da alle erlaubten Wellenvektoren in *einer* Elementarzelle des reziproken Gitters liegen, ist ihre Dichte ϱ_{q} im reziproken Raum durch das Verhältnis Anzahl N der erlaubten Zustände zu Volumen der Elementarzelle des reziproken Gitters gegeben. Wir erhalten daher die für die Dichte wichtige Beziehung:

$$\varrho_{\mathrm{q}} = \frac{N}{(2\pi)^3/V_{\mathrm{Z}}} = \frac{V}{(2\pi)^3} \,. \tag{5.49}$$

Es ist bemerkenswert, daß im reziproken Raum die Dichte der erlaubten Zustände nur vom Probenvolumen abhängt, die Kristallstruktur selbst aber keine Rolle spielt.

Im nächsten Schritt wollen wir die *Zustandsdichte* $D(\omega)$ berechnen, die angibt, wieviele Zustände innerhalb eines gewissen *Frequenzintervalls*[18] $\mathrm{d}\omega$ anzutreffen sind. Sie ist die entscheidende Größe bei der Berechnung der spezifischen Wärme, denn diese wird durch die Zahl der Anregungen bestimmt, deren Energie vergleichbar mit der thermischen Energie ist. Gesucht ist also, wie in Abb. 5.22 gezeigt, die Zahl der erlaubten Zustände im Frequenzintervall $\mathrm{d}\omega$, d.h. zwischen den beiden Flächen $S(\omega)$ bzw. $S(\omega + \mathrm{d}\omega)$ mit konstanter, aber leicht unterschiedlicher Frequenz. Im Prinzip müssen wir alle Zustände in der Schale aufsummieren, doch können wir die Vorgehensweise vereinfachen und die Summation durch eine Integration ersetzen, da die Zustände im reziproken Raum sehr dicht liegen. Wir erhalten dann die Anzahl der in Frage kommenden Zustände durch Multiplikation der Dichte ϱ_{q} mit dem eingeschlossenen Volumen d^3q, das in der Abbildung

[18] Wir ermitteln hier die Anzahl der Zustände in einem Frequenzintervall. In den meisten Fällen jedoch versteht man unter dem Begriff „Zustandsdichte" die Zahl der Zustände in einem Energieintervall.

schwarz gezeichnet ist. Formal läßt sich die Vorgehensweise durch die Gleichung

$$D(\omega)\,\mathrm{d}\omega = \varrho_\mathrm{q} \int\limits_{\omega=\mathrm{const.}}^{\omega+\mathrm{d}\omega=\mathrm{const.}} \mathrm{d}^3q \tag{5.50}$$

ausdrücken. Um diese allgemeine Formel tatsächlich nutzen zu können, drücken wir d^3q durch das Integral $\mathrm{d}^3q = \int \mathrm{d}q_\perp \mathrm{d}S_\omega$ aus. Hierbei steht $\mathrm{d}S_\omega$ für das Element der Fläche $S(\omega)$ und $\mathrm{d}q_\perp$ für den Abstand zwischen den beiden Flächen $S(\omega)$ und $S(\omega+\mathrm{d}\omega)$. Die Integration erstreckt sich dabei über die geschlossene Fläche $S(\omega)$. Um $\mathrm{d}q_\perp$ in geeigneter Weise auszudrücken, benutzen wir die Definition der Gruppengeschwindigkeit. Für sie gilt $v_\mathrm{g} = |\mathrm{d}\omega/\mathrm{d}\boldsymbol{q}| = |\mathrm{grad}_{\boldsymbol{q}}\,\omega| = |\mathrm{d}\omega/\mathrm{d}q_\perp|$ und somit $|\mathrm{d}q_\perp| = \mathrm{d}\omega/v_\mathrm{g}$. Damit folgt aus (5.50) für die Zustandsdichte der Ausdruck:

$$D(\omega)\,\mathrm{d}\omega = \varrho_\mathrm{q}\,\mathrm{d}\omega \int\limits_{\omega=\mathrm{const.}} \frac{\mathrm{d}S_\omega}{|\mathrm{grad}_{\boldsymbol{q}}\omega|} = \frac{V}{(2\pi)^3}\,\mathrm{d}\omega \int\limits_{\omega=\mathrm{const.}} \frac{\mathrm{d}S_\omega}{v_\mathrm{g}}\,. \tag{5.51}$$

Bei kubischen Gittern (und auch bei isotropen Festkörpern) ist die Fläche konstanter Frequenz im reziproken Raum die Oberfläche einer Kugel mit dem Radius q, auf der die Gruppengeschwindigkeit konstant ist. In diesem speziellen Fall findet man für die Zustandsdichte den einfachen Zusammenhang

$$D(\omega)\mathrm{d}\omega = \varrho_\mathrm{q}\,\mathrm{d}\omega\,\frac{4\pi q^2}{v_\mathrm{g}} = \frac{V}{2\pi^2}\,\frac{q^2}{v_\mathrm{g}}\,\mathrm{d}\omega\,. \tag{5.52}$$

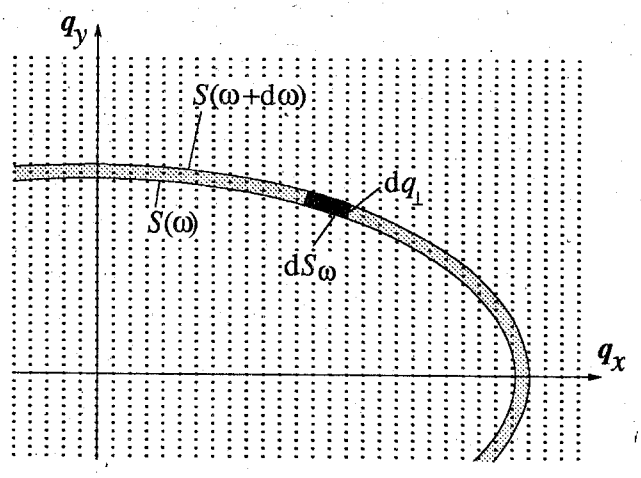


Abb. 5.22: Zur Berechnung der Zustandsdichte eines orthorhombischen Kristalls. Die Zustandsdichte erhält man, indem man die Zustände zwischen den Flächen $S(\omega)$ und $S(\omega + \mathrm{d}\omega)$ aufsummiert. Das herausgegriffene schwarze Element besitzt das Volumen $\mathrm{d}S_\omega\,\mathrm{d}q_\perp$.

Im allgemeinen Fall muß jedoch die Richtungsabhängigkeit berücksichtigt werden. Von besonderer Bedeutung ist die Tatsache, daß die Zustandsdichte umso größer ist, je flacher der Verlauf der Dispersionskurve $\omega_{\boldsymbol{q}}$ ist, da damit eine Reduktion der Gruppengeschwindigkeit einhergeht. Im Phononenspektrum gibt es immer Frequenzbereiche, in denen die Dispersionskurve waagerecht verläuft (vgl. Kap. 5.4), die Gruppengeschwindigkeit sogar verschwindet. Man spricht dann von *kritischen Punkten* oder *van-Hove-Singularitäten*. Bei einer linearen Kette, also bei eindimensionalen Systemen, divergiert die Zustandsdichte an dieser Stelle. Wie wir nun zeigen wollen, ist dies in realen, dreidimensionalen Festkörpern jedoch nicht der Fall. Hierzu setzen wir vereinfachend einen isotropen

Festkörper voraus, dessen Dispersionskurve $\omega(q)$ bei den kritischen Werten ω_c bzw. q_c ein lokales Maximum aufweist. Nun entwickeln wir die Dispersionsrelation in eine Taylor-Reihe, brechen die Reihen nach dem ersten nicht-verschwindenden Glied ab und finden $(\omega - \omega_c) \propto (q - q_c)^2$. Die Gruppengeschwindigkeit v_g der Phononen in der Umgebung des Extremums ist daher durch $v_g = \partial\omega/\partial q \propto (q - q_c)$ gegeben. Da die Gruppengeschwindigkeit in einem isotropen Festkörper unabhängig von der Richtung ist, kann sie in (5.51) vor das Integral gezogen werden. Mit der Fläche konstanter Frequenz $S_\omega = 4\pi(q - q_c)^2$ ergibt sich für die Zustandsdichte die Beziehung

$$D(\omega) \propto \frac{S_\omega}{v_g} \propto \frac{(q - q_c)^2}{(q - q_c)} \propto \sqrt{\omega - \omega_c}\,. \tag{5.53}$$

Im dreidimensionalen Fall divergiert also nicht die Zustandsdichte, sondern deren Ableitung $\mathrm{d}D(\omega)/\mathrm{d}\omega$. Bei anisotropen Kristallen kann der Verlauf der Dispersionskurve natürlich wesentlich komplizierter sein als bei isotropen Festkörpern. So treten neben lokalen Extrema der Dispersionskurve auch noch Sattelpunkte auf. Eine entsprechende Taylor-Entwicklung um einen Sattelpunkt führt zu dem Ergebnis, daß auch dort eine Unstetigkeit der Tangente der Zustandsdichte zu finden ist. In Fig. 5.23 ist der schematische Verlauf der Zustandsdichte für akustische Phononen aufgetragen, deren Dispersionskurve neben einem Maximum an der Grenze der Brillouin-Zone noch zwei Sattelpunkte mit unterschiedlichen Krümmungsvorzeichen aufweist.

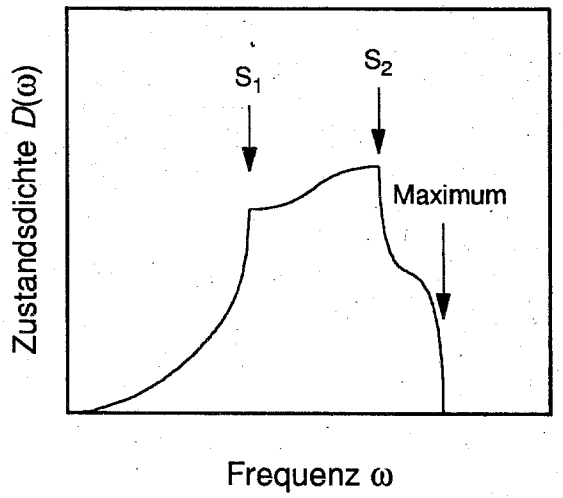


Abb. 5.23: Beispiel für die Zustandsdichte mit van-Hove-Singularitäten. Bei den Sattelpunkten S_1 und S_2 ist die Dispersionskurve unterschiedlich gekrümmt.

Die Zustandsdichte läßt sich aus einer Reihe von Experimenten (z.B. durch Raman-Streuung 2. Ordnung oder durch inkohärente inelastische Neutronenstreuung) direkt ermitteln oder aber bei vollständiger Kenntnis der Dispersionskurven durch Mittelung über alle Zweige und Raumrichtungen errechnen. In Abb. 5.24 sind die Phononen-Dispersionskurven für Diamant[19] und in Abb. 5.25 die daraus errechnete Zustandsdichte dargestellt. Es ist zu beachten, daß es sich bei dieser Zustandsdichte um eine Überlagerung des Beitrags von sechs Phononenzweigen handelt, und die Zuordnung von besonderen Merkmalen in Zustandsdichte und Dispersionskurve nicht immer einfach ist. Klar zu

[19] Die eingezeichneten Kurven wurden mit einem Modell errechnet, bei dem nur die Wechselwirkung mit den nächsten Nachbarn berücksichtigt wurde.

erkennen ist die Ähnlichkeit bei den niederfrequenten Phononen zwischen der tatsächlichen Zustandsdichte und der Prinzipskizze 5.23. Die hohe Zustandsdichte oberhalb von 30 THz rührt von den optischen Phononen her, die nur ein relativ enges Frequenzintervall überstreichen. In Abb. 5.25 ist zusätzlich die Zustandsdichte eingezeichnet, die sich aus der sogenannten Debye-Näherung ergibt, die wir nun besprechen wollen.

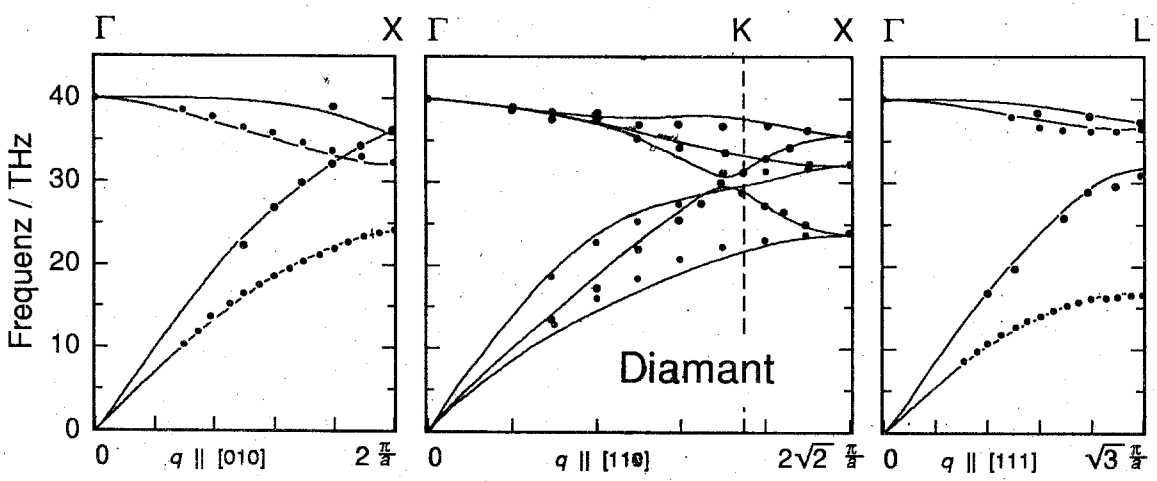


Abb. 5.24: Phononen-Dispersionskurven von Diamant. Die durchgezogenen Kurven geben berechnete Dispersionsverläufe wieder, die Meßwerte sind durch Punkte dargestellt. (Nach J.L. Warren, J.L. Yarnell, G. Dollinger, R.A. Cowley, Phys. Rev. **158** (1967) 805).

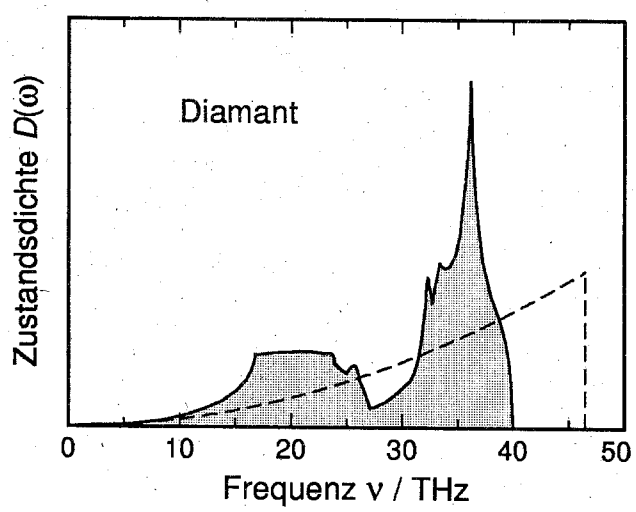


Abb. 5.25: Phononen-Zustandsdichte von Diamant. Die gestrichelte Kurve gibt die Dichte der Zustände nach dem Debye-Modell wieder. (Nach G. Dolling und R.A. Cowley, Proc. Phys. Soc. London **88** (1966) 463).

Spezifische Wärme in der Debye-Näherung

Wir wollen jetzt das Konzept der Zustandsdichte benutzen, um die spezifische Wärme von Festkörpern in der *Debye-Näherung* zu berechnen. Diese Näherung bezieht sich in ihrer ursprünglichen Form auf *isotrope* Festkörper mit einatomiger Basis, d.h. auf Festkörper, die nur akustische Phononen besitzen. Wir setzen für alle Wellenvektoren die Relation $\omega = vq$ voraus, also die Beziehung, die eigentlich nur für große Wellenlängen gültig ist.

Unter diesen Voraussetzungen vereinfacht sich Gleichung (5.52) für die Zustandsdichte zu

$$\mathcal{D}(\omega)\,\mathrm{d}\omega = \frac{V}{(2\pi)^3}\,4\pi q^2\,\mathrm{d}q = \frac{V}{2\pi^2}\,\frac{\omega^2}{v^3}\,\mathrm{d}\omega\,, \tag{5.54}$$

wobei $\mathcal{D}(\omega)$ für die Zustandsdichte *pro* Phononenzweig steht. Sie steigt quadratisch mit der Frequenz an und muß aufgrund der begrenzten Anzahl von Schwingungszuständen eine obere *Abschneidefrequenz* besitzen. In der Debye-Näherung wird diese Abschneidefrequenz über die Gesamtzahl $\mathcal{N}$ der pro Zweig erlaubten Schwingungszustände definiert und nicht direkt auf (5.45) zurückgegriffen. Es wird dazu die Beziehung

$$\mathcal{N} = \int_0^{\omega_{\max}} \frac{V}{2\pi^2}\frac{\omega^2}{v^3}\,\mathrm{d}\omega \tag{5.55}$$

benutzt. Setzen wir $V/\mathcal{N} = a^3$, wobei a für die Gitterkonstante steht, so ergibt sich hiermit für die Abschneidefrequenz $\omega_{\max}$ der Zusammenhang

$$\omega_{\max} = v\sqrt[3]{\frac{6\pi^2\mathcal{N}}{V}} = \frac{v}{a}\sqrt[3]{6\pi^2}\,. \tag{5.56}$$

Diese Grenzfrequenz weicht etwas von dem Wert ab, den man aus der Bedingung für $m^{\max}$ herleiten würde. Dies ist verständlich, da in der Debye-Näherung die erlaubten Zustände in einer Kugel im $\boldsymbol{q}$-Raum untergebracht sind, während sie bei den ursprünglich gewählten Randbedingungen einen Würfel im reziproken Raum ausfüllen. Da die Dichte der Zustände im $\boldsymbol{q}$-Raum in beiden Fällen gleich ist, müssen Kugel und Würfel das gleiche Volumen aufweisen.

Berücksichten wir, daß es im isotropen Festkörper einen longitudinalen und zwei (entartete) transversale Zweige gibt, so können wir für die *Debyesche Zustandsdichte*

$$D(\omega) = \frac{V\omega^2}{2\pi^2}\left(\frac{1}{v_{\mathrm{l}}^3} + \frac{2}{v_{\mathrm{t}}^3}\right) = \frac{3V}{2\pi^2}\,\frac{\omega^2}{v_{\mathrm{D}}^3} \tag{5.57}$$

schreiben, wenn wir als Abkürzung die *Debye-Geschwindigkeit* v_{D} über die Gleichung

$$\frac{3}{v_{\mathrm{D}}^3} = \frac{1}{v_{\mathrm{l}}^3} + \frac{2}{v_{\mathrm{t}}^3} \tag{5.58}$$

einführen. Der Faktor 3 in der Definition der Debye-Geschwindigkeit berücksichtigt die Tatsache, daß $3\mathcal{N}$ Schwingungszustände vorhanden sind. Für viele Festkörper können wir näherungsweise $v_{\mathrm{l}} \approx 2v_{\mathrm{t}}$ setzen, wodurch v_{D} den Wert $v_{\mathrm{D}} \approx 1{,}1\,v_{\mathrm{t}}$ annimmt. Die Zustandsdichte im Debye-Modell wird bei kleinen Frequenzen also in erster Linie durch die transversalen Gitterwellen bestimmt.

Obwohl es für longitudinale und transversale Phononen aufgrund ihrer unterschiedlichen Schallgeschwindigkeit eine unterschiedliche obere Grenzfrequenz geben muß, ist es üblich, in der hier diskutierten Näherung nur eine Abschneidefrequenz, die *Debye-Frequenz* ω_{D}, einzuführen. Der Wert der Debye-Frequenz errechnet sich daher aus Gleichung (5.56), wenn man die Schallgeschwindigkeit v durch die Debye-Geschwindigkeit v_{D} ersetzt. Diese einfache Theorie wird oft auch auf Kristalle mit mehratomiger Basis angewendet, ohne zwischen akustischen und optischen Phononen zu unterscheiden. Auch

in diesem Fall wird die Debye-Frequenz einfach durch die Zahl der vorhandenen Atome festgelegt.

Ist die Zustandsdichte der Phononen bekannt, so läßt sich sofort der temperaturabhängige Beitrag der Atomschwingungen zur inneren Energie U und somit zur spezifischen Wärme des Festkörpers berechnen. Für die innere Energie gilt der einfache Ausdruck

$$U = \int_0^{\omega_D} \hbar\omega \, D(\omega) \langle n(\omega, T) \rangle \mathrm{d}\omega \; . \tag{5.59}$$

$\langle n(\omega, T) \rangle = (\mathrm{e}^{\hbar\omega/k_B T} - 1)^{-1}$ ist der Bose-Einstein-Faktor, der die mittlere Besetzung der Eigenzustände wiedergibt. Beim harmonischen Oszillator gibt er an, wieviele Energieniveaus im Mittel besetzt sind. Im vorliegenden Fall sagt uns $\langle n(\omega_{\boldsymbol{q}}, T) \rangle$, wieviele Phononen mit der Frequenz ω und dem Wellenvektor $\boldsymbol{q}$ im thermischen Mittel im betrachteten Festkörper vorhanden sind.

Bei der Integration von (5.59) berücksichtigen wir die Nullpunktsenergie nicht, da sie keinen Beitrag zur spezifischen Wärme leistet. Auf ihre Größe gehen wir im nächsten Abschnitt gesondert ein. Für die innere Energie folgt somit die Beziehung

$$U(T) = \frac{9N}{\omega_D^3} \int_0^{\omega_D} \frac{\hbar\omega^3}{\mathrm{e}^{\hbar\omega/k_B T} - 1} \mathrm{d}\omega \; . \tag{5.60}$$

Die einzige materialabhängige Größe ist die Debye-Frequenz ω_D. Nun führen wir noch die *Debye-Temperatur* Θ über die Beziehung $k_B\Theta = \hbar\omega_D$ ein und benutzen die Abkürzungen $x = \hbar\omega/k_B T$ bzw. $x_D = \hbar\omega_D/k_B T = \Theta/T$. Damit erhalten wir für die spezifische Wärme das Endresultat

$$C_V = \left(\frac{\partial U}{\partial T}\right)_V = 9Nk_B \left(\frac{T}{\Theta}\right)^3 \int_0^{x_D} \frac{x^4 \mathrm{e}^x}{(\mathrm{e}^x - 1)^2} \mathrm{d}x \; . \tag{5.61}$$

Dies ist die berühmte Formel von Debye, die eine universelle Darstellung der Temperaturabhängigkeit der spezifischen Wärme mit nur *einem* materialspezifischen Parameter erlaubt. Die folgende Tabelle enthält die Debye-Temperaturen Θ von einigen Kristallen.

Tabelle 5.1: Debye-Temperaturen einiger ausgewählter Substanzen.

	C	Si	Fe	Al	NaCl	Au	Ar(fest)	He (fest)
Θ (K)	2230	640	470	430	320	164	92	25

Wir wollen die beiden Grenzfälle $T \to 0$ und $T \to \infty$ in der Debye-Näherung genauer untersuchen. Wir erwarten einen raschen Anstieg der spezifischen Wärme bei tiefen Temperaturen, doch sollte dieser Anstieg weniger stark ausgeprägt sein als in der Einsteinschen Theorie. Bei hohen Temperaturen sollte die Debye-Theorie mit den Ergebnissen der klassischen Thermodynamik übereinstimmen, d.h. sie sollte das Dulong-Petit-Gesetz reproduzieren.

Für *hohe Temperaturen* geht $x \to 0$. Wir können daher die Exponentialfunktion im Zähler durch eine Eins ersetzen und die Entwicklung der Exponentialfunktion im Nenner nach dem linearen Term abbrechen. Das Integral läßt sich dann wie folgt auswerten:

$$\int_0^{x_D} \frac{x^4 e^x}{(e^x-1)^2} dx \approx \int_0^{x_D} \frac{x^4 \cdot 1}{(1+x-1)^2} dx = \int_0^{x_D} x^2 dx = \frac{1}{3}\left(\frac{\Theta}{T}\right)^3 . \qquad (5.62)$$

Setzen wir dieses Ergebnis in (5.61) ein, so finden wir übereinstimmend mit dem Dulong-Petit-Gesetz für die molare Wärmekapazität den Ausdruck $C_V = 3N_A k_B = 3R_m$, wobei R_m wieder für die Gaskonstante steht.

Für *tiefe Temperaturen* geht die obere Integrationsgrenze $x_D \to \infty$. Unter dieser Voraussetzung kann man das Integral analytisch lösen und findet:

$$\int_0^{\infty} \frac{x^4 e^x}{(e^x-1)^2} dx = \frac{4\pi^4}{15} \quad \to \quad C_V = \frac{12\pi^4}{5} N k_B \left(\frac{T}{\Theta}\right)^3 . \qquad (5.63)$$

Dies ist das berühmte T^3-Gesetz für die spezifische Wärme von Festkörpern bei tiefen Temperaturen. Die gute Übereinstimmung zwischen der theoretischen Beschreibung (5.63) und dem Experiment wird in Abb. 5.26 deutlich. Ähnlich gute Übereinstimmung findet man auch bei anderen reinen Kristallen.

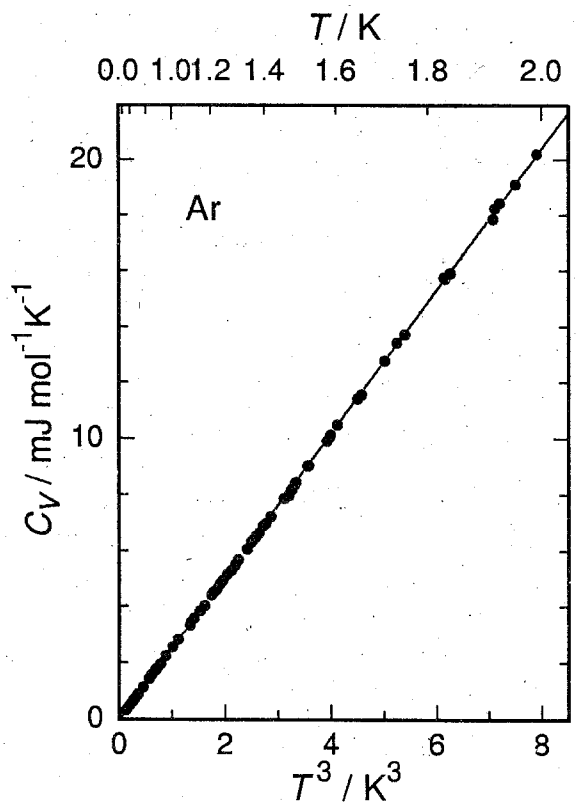


Abb. 5.26: Spezifische Wärme von kristallinem Argon bei tiefen Temperaturen, aufgetragen als Funktion von T^3. (Nach L. Finegold und N. E. Phillips, Phys. Rev. **177** (1964) 383).

Da die Schallgeschwindigkeiten der Festkörper in weiten Bereichen variieren kann und die Debye-Temperatur $\Theta \propto v$ ist, unterscheiden sich bei tiefen Temperaturen die spezifischen Wärmen sehr stark. Die Extremfälle sind Diamant und festes Helium, deren Debye-Temperaturen nach Tabelle 5.1 sich um etwa zwei Größenordnungen und die spezifischen Wärmen daher um etwa sechs Größenordnungen unterscheiden!

Wie gut ist die Debye-Näherung? Abb. 5.26 demonstriert eine beeindruckend gute Übereinstimmung bei tiefen Temperaturen. Der Vergleich zwischen Theorie und Experiment bei höheren Temperaturen wird in Abb. 5.27 gezeigt und ist ebenfalls sehr überzeugend. Rückt man von einer gemeinsamen Darstellung der spezifischen Wärme von vielen Materialien ab und führt eine genaue Analyse der Daten einzelner Proben durch, so wird offenbar, daß im Übergangsbereich zwischen tiefen und hohen Temperaturen Diskrepanzen auftreten. Im allgemeinen ist die Debye-Näherung gut für $T < \Theta/100$ bei tiefen und für $T > \Theta/5$ bei hohen Temperaturen. Offenbar wird bei niedrigen Frequenzen, die bei tiefen Temperaturen wichtig sind, die Zustandsdichte der Phononen durch die quadratische Frequenzabhängigkeit von (5.57) sehr gut beschrieben. Bei hohen Temperaturen sind die Details des Spektrums unwichtig, da alle Schwingungen angeregt sind. Vergleicht man in Abb. 5.25 im Falle des Diamants die tatsächliche Zustandsdichte mit der Zustandsdichte in der Debye-Näherung, so ist die gute Übereinstimmung von Theorie und Experiment dennoch überraschend. Im Prinzip läßt sich eine Verbesserung der Theorie bei höheren Temperaturen durch die getrennte Behandlung von akustischen und optischen Phononen erzielen. So kann man für die akustischen Phononen weiter die Debye-Näherung benutzen, zur Beschreibung der optischen Phononen aber zum Einstein-Modell übergehen, da dort die Annahme von festen Oszillatorfrequenzen nicht sehr weit von den realen Gegebenheiten entfernt ist.

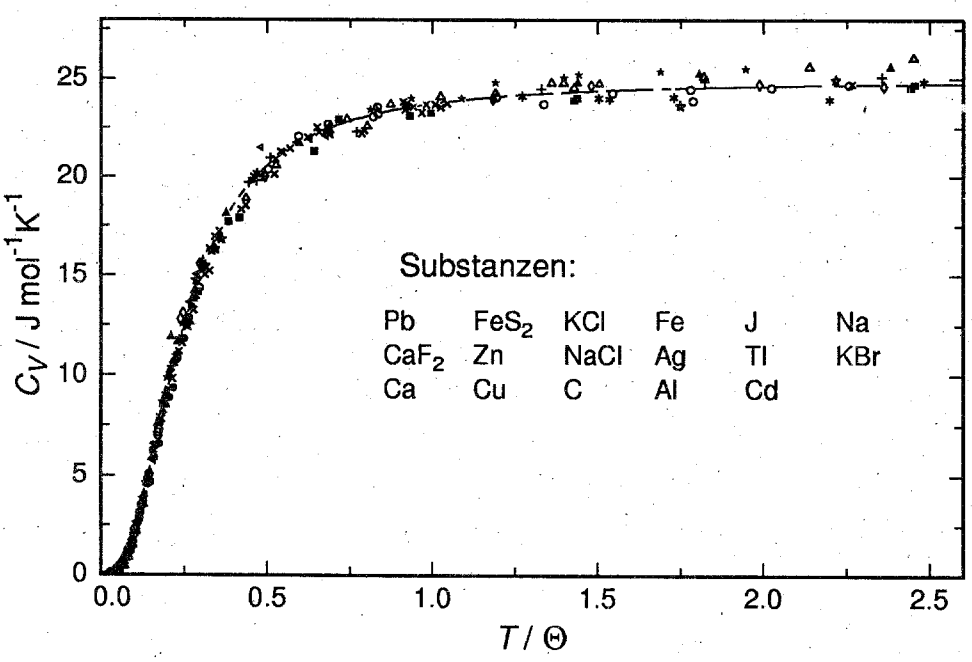


Abb. 5.27: Molare Wärmekapazität einer Reihe von Substanzen als Funktion der reduzierten Temperatur T/Θ. Durch die Normierung auf die Debye-Temperatur fallen alle Meßdaten auf eine Kurve. (Nach E. Schrödinger, *Handbuch der Physik*, Band X, H. Geiger und K. Scheel (eds.), J. Springer Berlin, 1926, p. 275).

Spezifische Wärme niederdimensionaler Systeme

Nachdem wir die spezifische Wärme der Kristalle diskutiert haben, wollen wir noch kurz niederdimensionale Systeme betrachten, bevor wir etwas ausführlicher auf amorphe Substanzen eingehen. Die ω^2-Abhängigkeit der Zustandsdichte bei kleinen Frequenzen ist

typisch für dreidimensionale Systeme. Wollen wir die spezifische Wärme von ein- oder zweidimensionalen Systemen berechnen, so müssen wir deren Zustandsdichte in das Integral für die innere Energie einsetzen. Allgemein gilt, daß im $\boldsymbol{q}$-Raum die Dichte der erlaubten Zustände durch $\varrho_q = (L/2\pi)^d$ gegeben ist, wobei d für die Dimensionalität des Systems steht. Abhängig von der Dimensionalität muß bei der Berechnung der Zustandsdichte isotroper Materialien die Zahl der Zustände in einer Kugelschale mit dem Radius q innerhalb eines Kreisrings mit diesem Radius bzw. einer Strecke bestimmt werden. Das Ergebnis läßt sich leicht voraussagen. Da $\omega \propto q$ ist, findet man

$$D(\omega) \propto \omega^{d-1} \qquad \text{und} \qquad C_V \propto T^d \,. \tag{5.64}$$

Aufgrund dieser Beziehung erwartet man für zweidimensionale Systeme bei tiefen Temperaturen einen quadratischen Temperaturverlauf von C_V. Ein schönes Beispiel hierfür liefern Gasatome, die auf Graphit kondensiert wurden. Die Van der Waals-Kräfte binden die Atome an die Unterlage, erlauben aber eine Bewegung längs der Substratoberfläche. So bilden sich zweidimensionale Kristalle aus, die sowohl mit Streumethoden als auch durch Messungen der spezifischen Wärme untersucht werden können. Das Ergebnis einer derartigen Meßreihe an ^{3}He-Schichten mit einer Dicke von weniger als einer Monolage ist in Abb. 5.28 gezeigt. Die spezifische Wärme verläuft tatsächlich in allen Fällen proportional zu T^2. Der Abstand zwischen den Atomen verkleinert sich mit anwachsender Belegung, und die Wechselwirkung zwischen den Atomen wird stärker. Mit zunehmender Dichte wird der „zweidimensionale Kristall", der sich auf der Substratoberfläche ausbildet, daher steifer. Damit verbunden nimmt die Debye-Temperatur, die im Bereich von $\Theta \approx 15-35\,\mathrm{K}$ liegt, ebenfalls zu.

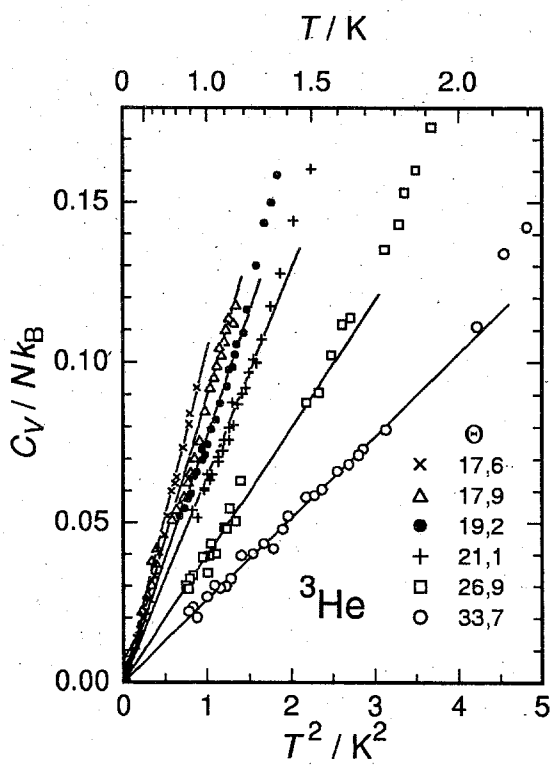


Abb. 5.28: Spezifische Wärme von ^{3}He-Filmen auf Graphit als Funktion von T^2. Die Dichte der Heliumatome steigt links beginnend von 0,078 auf 0,092 pro Å^2 an. Damit nimmt auch die Debye-Temperatur Θ zu, die ebenfalls angegeben ist. (Nach S.V. Hering, S.W. Van Sciver, O.E. Vilches, J. Low Temp. Phys. **25** (1976) 793).

Nullpunktsenergie, Zahl der angeregten Phononen

Ehe wir auf die spezifische Wärme der amorphen Festkörper eingehen, wollen wir noch den Wert der Nullpunktsenergie berechnen und herausfinden, wie die Zahl der thermisch angeregten Phononen von der Temperatur abhängt. Die Frage nach der Größe der Nullpunktsenergie U_0 tauchte bereits im Zusammenhang mit der Bindungsenergie von Edelgaskristallen auf. Dort hatten wir sie nur abgeschätzt, nun wollen wir sie im Rahmen der Debye-Näherung berechnen. Nullpunktsphononen besitzen bei gleicher Frequenz die halbe Energie von thermisch angeregten Phononen. Da sie unabhängig von der Temperatur immer vorhanden sind, besitzt die mittlere Besetzungszahl jeder Eigenschwingung den Wert eins. Für die innere Energie der Nullpunktsschwingungen gilt daher anstelle von (5.59) der Ausdruck

$$U_0 = \int_0^{\omega_\mathrm{D}} \frac{\hbar\omega}{2} D(\omega)\,\mathrm{d}\omega = \frac{9}{8} N k_\mathrm{B} \Theta \,. \tag{5.65}$$

Die Nullpunktsenergie steigt linear mit der Debye-Temperatur an. Die Zahlenwerte für Edelgaskristalle sind in Tabelle 3.2 eingetragen. Bemerkenswert an diesen Werten ist zunächst, daß die Nullpunktsenergie U_0 nicht monoton mit der Atommasse abnimmt. Vom jetzigen Standpunkt aus betrachtet ist dies nicht verwunderlich, denn es kommt nicht auf die Masse der einzelnen Atome an, sondern auf die Debye-Temperatur, in die auch die Dichte und die Stärke der Kopplung zwischen den Atomen eingeht. Mit $\Theta = 92\,\mathrm{K}$ besitzt festes Argon die höchste Debye-Temperatur aller Edelgaskristalle und damit auch die größte Nullpunktsenergie.

Wie vergleicht sich die thermische Energie mit der Nullpunktsenergie? Ziehen wir die thermische Energie bei der Temperatur heran, die der Debye-Temperatur entspricht, so gilt grob $U(T = \Theta) \approx 2\,U_0$. Die Nullpunktsenergie und die thermische Energie sind also bei Zimmertemperatur in den meisten Fällen vergleichbar!

Nun wollen wir uns noch überlegen, wie die Zahl N_ph der thermisch angeregten Phononen mit der Temperatur ansteigt. Im Rahmen der Debye-Näherung gilt für diese Größe der Zusammenhang

$$N_\mathrm{ph} = \int_0^{\omega_\mathrm{D}} D(\omega)\langle n(\omega,T)\rangle\,\mathrm{d}\omega \propto T^3 \int_0^{x_\mathrm{D}} \frac{x^2}{e^x - 1}\,\mathrm{d}x \,. \tag{5.66}$$

Für den Grenzfall $T \ll \Theta$ geht $x_\mathrm{D} \to \infty$, und das Integral wird konstant. Die Auswertung für $T \gg \Theta$ erfolgt wie bei Gleichung (5.62), und man findet einen Wert proportional zu $(\Theta/T)^2$. Beide Ergebnisse lassen sich durch

$$N_\mathrm{ph} \propto \begin{cases} T^3 & \text{für } T \ll \Theta \\ T & \text{für } T \gg \Theta \end{cases} \tag{5.67}$$

zusammenfassen. Die Zahl der Phononen steigt bei tiefen Temperaturen rasch an. Um $T \approx \Theta$ flacht sich dann der Anstieg ab und geht in einen linearen Zusammenhang über. In vielen Phänomenen, die wir noch besprechen werden, spielt die Zahl der Phononen eine wichtige Rolle.

5.6 Schwingungsspektrum amorpher Festkörper

Wir wenden uns nun dem Schwingungsspektrum amorpher Festkörper zu, das wesentlich weniger gut bekannt und schwerer zu verstehen ist als das von Kristallen. Bei amorphen Festkörpern stellt sich die Frage, wie überhaupt die Dispersionskurven aussehen? In Abb. 5.29 ist der Versuch unternommen, das Schwingungsspektrum amorpher Festkörper zu veranschaulichen.

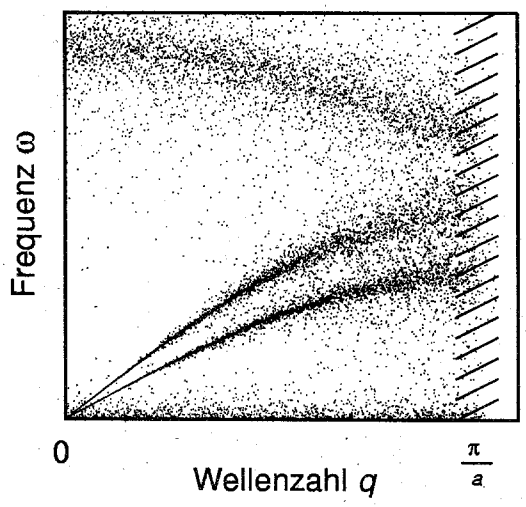


Abb. 5.29: Veranschaulichung der Phononendispersion in amorphen Materialien. Mit zunehmender Frequenz werden die Phononenzweige unscharf. Bei sehr kleinen Energien treten weiche, lokalisierte Schwingungszustände auf. Die Grenze der Brillouin-Zone ist nicht eindeutig definiert.

Natürlich können in Gläsern Schallwellen angeregt werden. So wurde das in Abb. 5.1 dargestellte Ultraschallexperiment an Glasproben durchgeführt. Doch die ursprünglich benutzte Definition der Phononen kann in amorphen Festkörpern nur für große Wellenlängen gültig sein. Die Behandlung der Phononen mit kleinen Wellenlängen in den Abschnitten 5.2 und 5.3 setzte die Existenz eines periodischen Gitters voraus. Messungen[20] an dünnen amorphen Filmen haben gezeigt, daß wohldefinierte Phononen mit Frequenzen bis 400 GHz existieren, d.h. Phononen mit Wellenlängen von 120 Å sind experimentell nachgewiesen. Mindestens bis dorthin ist die beobachtete Dispersionskurve linear und damit die Schallgeschwindigkeit frequenzunabhängig. Die untersuchten Wellenvektoren entsprechen aber nur etwa 1/50 der „Brillouin-Zone". Inwieweit die phononenartigen Schwingungen bei höheren Frequenzen räumlich lokalisiert sind oder sich wie Phononen in Kristallen fortbewegen können, ist noch weitgehend ungeklärt, und ist Gegenstand intensiver experimenteller und theoretischer Untersuchungen.

Bei der Definition der Brillouin-Zone stellen sich weitere grundlegende Fragen. In Kristallen ist der Rand der Brillouin-Zone durch den Gitterabstand bestimmt. Versucht man in amorphen Materialien eine entsprechende Größe zu definieren, so muß diese mit dem atomaren Abstand schwanken, d.h., die Ausdehnung der „Brillouin-Zone" wird von Ort

[20] Die Messung erfolgte mit supraleitenden Tunneldioden, auf die wir in Kap. 12.2 kurz eingehen. Mit supraleitenden Tunneldioden lassen sich sehr hochfrequente Phononen erzeugen und nachweisen. Unter geeigneten Bedingungen erlauben sie auch eine relativ gute Auflösung der Phononenfrequenzen.

zu Ort variieren[21]. Zusätzlich zu den lokalen Dichtefluktuationen treten auch lokale Variationen der Federkonstanten und somit auch der Schwingungsfrequenzen auf. Darüber hinaus konnte durch theoretische Überlegungen und numerische Simulationen gezeigt werden, daß in Ausnahmefällen die lokalen rücktreibenden Kräfte für die schwingenden Atome außerordentlich klein werden können. Dadurch bilden sich „weiche" Schwingungszustände aus, die räumlich lokalisiert sind, sehr kleine Schwingungsenergien besitzen, sich aber nicht wie Phononen im Festkörper ausbreiten können. In Abb. 5.29 wurde versucht, das Auftreten dieser „weichen" Schwingungszustände und die Unschärfe der Dispersionskurven durch „Punktwolken" darzustellen.

In amorphen Festkörpern ist nicht nur die Analyse der kohärenten elastischen Röntgen- oder Neutronenstreuung sehr schwierig, sondern auch die inelastische Streuung ist schwer zu interpretieren. Gehen wir noch einmal zurück zur inelastischen Neutronenstreuung. Natürlich ist der Ausdruck für die Energieerhaltung bei der Streuung an Phononen weiterhin gültig. Da aber die Gitterperiodizität fehlt, wird der Erhaltungssatz für den Quasiimpuls „aufgeweicht". Darüberhinaus läßt sich kein eindeutiger reziproker Gittervektor definieren. Falls bei der Neutronenstreuung mit festem Energieübertrag kein Wellenvektor bevorzugt wird, folgt aus (5.35) für die inelastische Streuung eine K^2-Abhängigkeit der Streuintensität. Tatsächlich wurde ein solcher Verlauf experimentell ermittelt.

In den letzten Jahren wurden große Anstrengungen unternommen, das Spektrum der elastischen Anregungen, also das Schwingungsspektrum amorpher Festkörper, zu bestimmen. Hierzu haben neben Messungen der kohärenten inelastischen Neutronen- und der Raman-Streuung vor allem Untersuchungen der inkohärenten Neutronenstreuung beigetragen. In Abb. 5.30 ist das neue Ergebnis für die Zustandsdichte von Quarzglas gezeigt, wie sie sich aus verschiedenen Neutronenstreumessungen bei tiefen und sehr hohen Temperaturen herleiten läßt. Um einen Vergleich mit dem Debye-Modell zu ermöglichen ist die normierte Zustandsdichte $D(\nu)/\nu^2$ als Funktion der Frequenz ν aufgetragen. In dieser Art der Auftragung wird die Debyesche Zustandsdichte durch eine waagerechte Gerade repräsentiert. Die eingezeichnete Debyesche Zustandsdichte wurde mit Hilfe der Gleichung (5.57) aus den elastischen Daten errechnet.

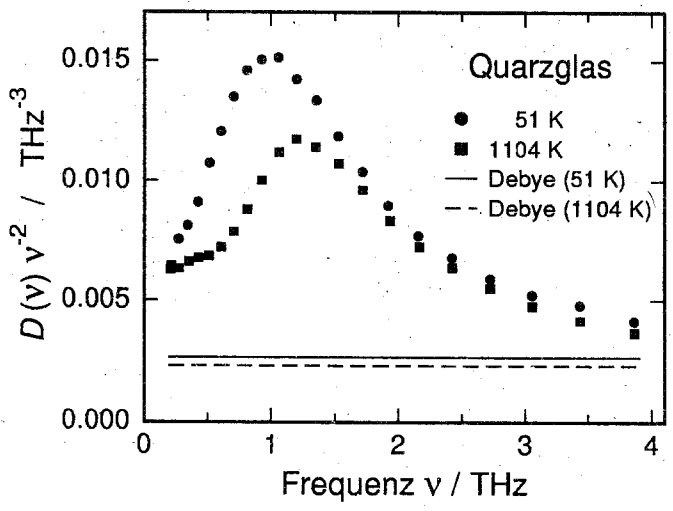


Abb. 5.30: Normierte Zustandsdichte $D(\nu)/\nu^2$ der Schwingungsanregungen von Quarzglas bei relativ kleinen Frequenzen, hergeleitet aus Neutronenstreumessungen. Es sind die Ergebnisse für 51 K bzw. 1104 K eingetragen. Ebenfalls dargestellt sind die errechneten Debyeschen Zustandsdichten für diese Temperaturen. Bemerkenswert ist die starke Temperaturabhängigkeit des Schwingungsspektrums. (Nach A. Wischnewski, Doktorabeit Köln, 1998).

[21] Entsprechendes gilt im realen Raum, wie wir bei der Einführung der Voronoi-Polyeder in Kap. 1.4 gesehen haben.

Offensichtlich findet man bei niedrigen Frequenzen eine Zustandsdichte, die im ganzen gezeigten Frequenzbereich erheblich über der des Debye-Modells liegt. Sie ist damit auch wesentlich höher als in den entsprechenden kristallinen Materialien. Deutlich zu erkennen ist, daß bei etwa 1 THz ein Maximum auftritt, das als „Bosonpeak" bezeichnet wird und das bei allen amorphen Festkörpern zu finden ist. Dieses Maximum wird gegenwärtig sehr intensiv untersucht, doch ist die mikroskopische Ursache dieser „weichen" Schwingungszustände noch umstritten und Gegenstand vieler Spekulationen. Bei sehr kleinen Energien steigt die normierte Zustandsdichte $D(\nu)/\nu^2$ mit abnehmender Energie nochmals steil an, doch ist dieser Energiebereich Neutronenstreumessungen nicht zugänglich und wird daher in der Abbildung nicht gezeigt. Auf diesen Teil des Schwingungsspektrums werden wir im nächsten Abschnitt eingehen.

Überraschend ist die starke Temperaturabhängigkeit der gemessenen Zustandsdichte, denn in der hier diskutierten Näherung der Schwingungszustände der Festkörper sollte sie überhaupt nicht von der Temperatur abhängen. Die Ursache hierfür sind Anharmonizitäten, die offensichtlich in Gläsern besonders stark ausgeprägt sind. Wir werden auf die anharmonischen Eigenschaften von Festkörpern in Kap. 6 eingehen.

5.7 Wärmekapazität von Gläsern bei sehr tiefer Temperatur

Wie wir gesehen haben, lassen sich die elastischen Eigenschaften eines Festkörpers für Gitterschwingungen mit großer Wellenlänge sehr gut auf die eines elastischen Kontinuums zurückführen. Das Konzept der Phononen sollte sich problemlos auf amorphe Festkörper übertragen lassen, solange die Wellenlänge so groß ist, daß das Medium als homogen angesehen werden kann. Wie im vorausgehenden Abschnitt erwähnt, ist die Existenz wohldefinierter Phononen im langwelligen Grenzfall bis 400 GHz experimentell nachgewiesen. Folglich kann man davon ausgehen (wenn es auch nicht bewiesen ist), daß auch für amorphe Festkörper bei tiefen Temperaturen die Debyesche Theorie eine gute Beschreibung der spezifischen Wärme liefert, die durch die Phononen hervorgerufen wird.

Wie wir bei der Diskussion der thermischen Eigenschaften von Kristallen gesehen haben, lassen sich aus dem Temperaturverlauf der spezifischen Wärme Rückschlüsse auf das Schwingungsspektrum des betreffenden Festkörpers ziehen. Bei hohen Temperaturen wird die spezifische Wärme von Gläsern wie bei den Kristallen durch das Dulong-Petit-Gesetz beschrieben, das die Tatsache ausdrückt, daß alle Schwingungsfreiheitsgrade angeregt sind. Mit abnehmender Temperatur fällt die spezifische Wärme jedoch in Gläsern weniger steil ab. Dies bedeutet, daß in amorphen Strukturen zusätzlich zu den Phononen noch weitere Schwingungsfreiheitsgrade existieren müssen. Diese Beobachtung steht in Einklang mit der in Abb. 5.30 dargestellten Zustandsdichte der Gitterschwingungen von Quarzglas.

Besonders deutlich wird der Beitrag der zusätzlichen Freiheitsgrade bei tiefen Temperaturen. Um dies zu zeigen ist in Abb. 5.31 die Wärmekapazität von Quarzglas mit der von kristallinem Quarz bei tiefen Temperaturen verglichen. Man stellt fest, daß unterhalb von 1 K die spezifische Wärme der beiden Modifikationen einen völlig unterschiedlichen Verlauf aufweist. Man beobachtet beim Glas nicht die erwartete T^3-Abhängigkeit, sondern einen näherungsweise linearen Anstieg. Bei der tiefsten Meßtemperatur von etwa 25 mK ist die spezifische Wärme des Glases mehr als 1000 mal größer als die des Quarzkristalls!

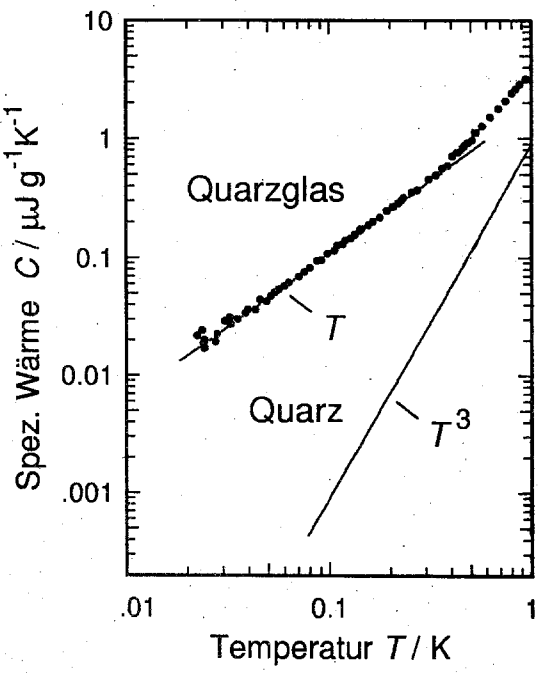


Abb. 5.31: Tieftemperaturverlauf der spezifischen Wärme von Quarzglas und kristallinem Quarz. (Nach S. Hunklinger, Adv. Solid State.Physics XVII, J. Treusch (ed.), Vieweg, Braunschweig 1977, p. 1).

Der lineare Term von C_V wird durch das phänomenologische *Tunnelmodell* beschrieben, auf das wir hier kurz eingehen. In diesem Modell wird angenommen, daß in amorphen Netzwerken auch bei sehr tiefen Temperaturen, d.h. bis unter 1 mK, noch strukturelle Umlagerungen auftreten, die auf quantenmechanischen Tunnelprozessen beruhen. Dies ist möglich, wenn einzelne Atome oder kleine Atomgruppen keine eindeutige Gleichgewichtslage besitzen. In Abb. 5.32a wurde versucht, diese Annahme für Quarzglas plausibel zu machen. Die beweglichen Atomgruppen, die wir als „Teilchen" bezeichnen, bewegen sich in einfachster Näherung in einem Doppelmuldenpotential, wie in Abb. 5.32b skizziert.

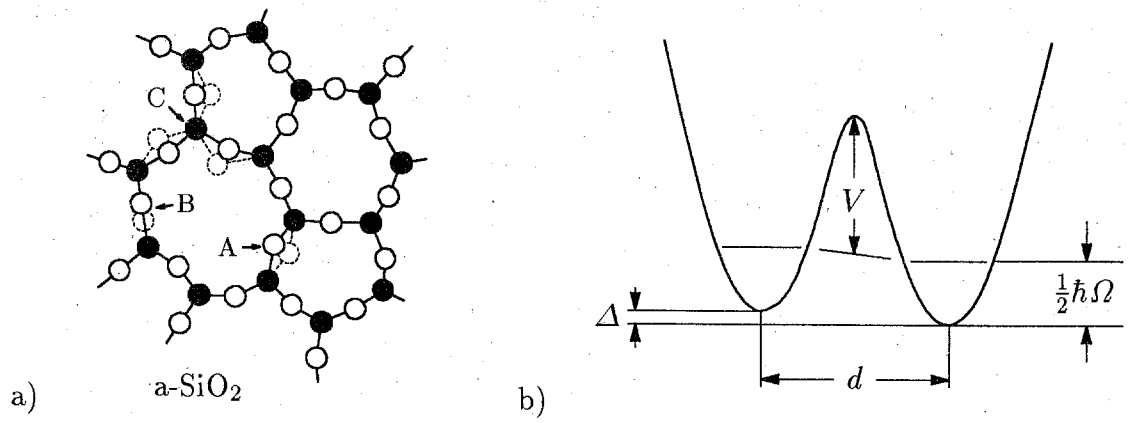


Abb. 5.32: Tunnelsysteme in Quarzglas.
a) Zweidimensionale Veranschaulichung der Struktur von Quarzglas. Einzelne Atome oder kleine Atomgruppen (mit A, B und C gekennzeichnet) nehmen keine eindeutige Gleichgewichtslage ein.
b) Doppelmuldenpotential. Das Potential der tunnelnden „Teilchen" ist gekennzeichnet durch die Asymmetrieenergie Δ, den Muldenabstand d und die Potentialbarriere V. Die Grundzustandsenergie (Nullpunktsenergie) $\hbar\Omega/2$ der isolierten Mulden ist ebenfalls eingezeichnet.

Die „Teilchen“ springen bei hoher Temperatur thermisch aktiviert über die Barriere. Bei tiefen Temperaturen können sie jedoch zwischen den beiden Mulden nur noch tunneln. Die Tunnelbewegung bewirkt, daß diese „Teilchen“ auch bei tiefen Temperaturen nicht in einer Mulde lokalisiert sind. Aus den Grundzuständen der beiden Einzelmulden entsteht ein gemeinsamer Zustand, der durch den Tunnelprozeß aufgespalten ist. Die Masse des tunnelnden „Teilchens“ muß relativ groß sein, da sonst die Tunnelaufspaltung zu groß für eine Erklärung der spezifischen Wärme bei sehr tiefen Temperaturen wäre. Damit scheiden Elektronen als Tunnelteilchen aus.

Die Aufspaltung des Grundzustands, die für die spezifische Wärme bei tiefen Temperaturen entscheidend ist, läßt sich wie bei der Behandlung der kovalenten Bindung in Kap. 3.4 in guter Näherung durch eine Linearkombination der Wellenfunktion ψ_a und ψ_b der ungekoppelten Zustände in Mulde a oder Mulde b berechnen. Wir benutzen daher Gleichung (3.22), die wir nochmals anschreiben:

$$(H_{aa} - E)(H_{bb} - E) - (H_{ab} - ES)^2 = 0 \,. \tag{5.68}$$

Mit H_{aa} und H_{bb} bezeichnen wir die Grundzustandsenergie, die das Teilchen jeweils einnähme, wenn beide Potentialmulden so weit voneinander entfernt wären, daß kein Tunneln möglich wäre[22]. Wir legen den Energienullpunkt in die Mitte zwischen die beiden Potentialminima. Dann ist $H_{aa} = (\hbar\Omega + \Delta)/2$ und $H_{bb} = (\hbar\Omega - \Delta)/2$. Weiterhin nehmen wir an, daß der Überlapp der Wellenfunktionen nur gering ist, d.h. wir dürfen das Überlappungsintegral $S \approx 0$ setzen. Dann ergibt sich

$$E_{\pm} = \frac{1}{2}\left(\hbar\Omega \pm \sqrt{\Delta^2 + 4H_{ab}^2}\right) \,. \tag{5.69}$$

Der Grundzustand spaltet also auf, wobei die Energiedifferenz zwischen den beiden Zuständen durch

$$E = E_+ - E_- = \sqrt{\Delta^2 + 4H_{ab}^2} = \sqrt{\Delta^2 + \Delta_0^2} \tag{5.70}$$

gegeben ist. Mit Hilfe der Störungstheorie läßt sich H_{ab} berechnen. Es ergibt sich näherungsweise $-2H_{ab} = \Delta_0 = \hbar\Omega\,\mathrm{e}^{-\lambda}$, wobei Δ_0 als Tunnelaufspaltung bezeichnet wird. Für den Tunnelparameter λ, der durch die Potentialform und die Masse m der tunnelnden „Teilchen“ bestimmt wird, findet man näherungsweise[23] den Ausdruck $\lambda^2 = mVd^2/2\hbar^2$. Offen ist hierbei noch immer die Frage: Welches „Teilchen“ tunnelt? Sicherlich ist Abb. 5.31a, in dem das Tunneln einzelner „isolierter“ Atome suggeriert wird, zu stark vereinfacht, denn es führt zu Werten der Grundzustandsaufspaltung E, die zu groß sind. Darüberhinaus können sich einzelne Atomgruppen aufgrund der Kopplung der Atome nicht völlig unabhängig voneinander bewegen. Die Tunnelbewegung des „Teilchens“ besteht wahrscheinlich in einer gemeinsamen Bewegung mehrerer Atome, wobei die Amplitude der Auslenkung vom Zentrum des Tunnelsystems ausgehend mit der Entfernung rasch abfällt. Die Bewegung selbst wird keine reine Translations- oder Rotationsbewegung sein. Man faßt deshalb den Muldenabstand d als eine sogenannte *Konfigurationskoordinate* und m als „effektive Masse“ auf, deren Bedeutung von der Natur der Teilchenbewegung

[22] Die einfachste Behandlung geht von parabelförmigen Potentialen aus, so daß H_{aa} und H_{bb} die Grundzustandsenergien von harmonischen Potentialen darstellen.

[23] Die Rechnung kann man leicht mit der sogenannten WKB-Methode durchführen, die nach Wentzel, Kramers und Brillouin benannt ist.

abhängt[24]. Die konkrete Lösung des Problems für reale Systeme stößt jedoch auf Schwierigkeiten, die bis jetzt noch nicht überwunden sind. Der Grund hierfür liegt nicht nur in der mangelnden Kenntnis der amorphen Struktur, sondern auch daran, daß die fehlende Symmetrie der amorphen Strukturen eine Behandlung der lokalen Bewegung der Atome außerordentlich erschwert.

Die Parameter der Doppelmuldenpotentiale sind aufgrund der Irregularitäten der amorphen Struktur nicht für alle Tunnelsysteme gleich. Es ist naheliegend, eine Verteilung einzuführen, die die statistischen Schwankungen der Parameter berücksichtigt. Im sogenannten *Tunnelmodell* wird angenommen, daß die Parameter λ und Δ unabhängige Größen sind, die gleichförmig verteilt sind. Dies läßt sich durch die einfache Verteilungsfunktion

$$P(\Delta,\lambda)\,\mathrm{d}\Delta\,\mathrm{d}\lambda = \overline{P}\,\mathrm{d}\Delta\,\mathrm{d}\lambda \tag{5.71}$$

ausdrücken, wobei $\overline{P}$ eine Konstante ist. Unter Berücksichtigung von (5.70) folgt daraus die Verteilungsfunktion $P(E,\lambda)$:

$$P(E,\lambda)\,\mathrm{d}E\,\mathrm{d}\lambda = P(\Delta,\lambda)\,\frac{\mathrm{d}\Delta}{\mathrm{d}E}\,\mathrm{d}E\,\mathrm{d}\lambda = \overline{P}\,\frac{E}{\sqrt{E^2-(\hbar\Omega\,\mathrm{e}^{-\lambda})^2}}\,\mathrm{d}E\,\mathrm{d}\lambda\,. \tag{5.72}$$

Die spezifische Wärme der Tunnelsysteme wird wie bei den Phononen durch die Zustandsdichte $D(E)$ der relevanten Anregungen bestimmt. Um die Zustandsdichte zu berechnen, integrieren wir über die auftretenden λ-Werte und erhalten

$$D(E) = \int_0^{\lambda_{\max}} P(E,\lambda)\,\mathrm{d}\lambda = \overline{P}\,\lambda_{\max}\,\ln\frac{\hbar\Omega}{2E}\,. \tag{5.73}$$

Um in diesem einfachen Modell das Auftreten einer unphysikalischen Divergenz der Zustandsdichte zu vermeiden, wurde für den Tunnelparameter eine obere Schranke $\lambda_{\max}$ eingeführt. In realen Systemen bricht natürlich die Verteilungsfunktion $P(\lambda)$ nicht abrupt ab, sondern geht allmählich gegen Null. Sehr große Werte von λ sind schon deswegen ohne physikalische Relevanz, da bei diesen „Teilchen" die Tunnelwahrscheinlichkeit so klein wäre, daß sie praktisch in einer Mulde lokalisiert bleiben und die Dynamik des Netzwerks nicht mehr beeinflussen können.

Obwohl der Zahlenwert der Grundzustandsenergie $\hbar\Omega$ nicht bekannt ist, können wir davon ausgehen, daß $\hbar\Omega \gg 2E$ ist. In erster Näherung kann $D(E)$ daher als konstant betrachtet werden, da die Energie E nur in relativ engen Grenzen variiert und die logarithmische Abhängigkeit somit nur eine schwache Variation der Zustandsdichte mit der Energie darstellt. Wie in Kap. 5.5 bei den niederdimensionalen Systemen diskutiert, führt eine konstante Zustandsdichte zu einem linearen Temperaturverlauf der spezifischen Wärme[25], wie er tatsächlich in Quarzglas und praktisch auch in allen anderen amorphen Festkörper zu sehen ist.

[24] Diese Vorstellung wird auch durch neuere „Simulationsexperimente" bestätigt.

[25] Bei der Berechnung der spezifischen Wärme von Zwei-Niveau-Systemen muß berücksichtigt werden, daß die thermische Besetzung der beiden Niveaus *nicht* durch den Bose-Einstein-Faktor wiedergegeben wird, bei dessen Herleitung das Termschema eines harmonischen Oszillators vorausgesetzt wird. Für Zwei-Niveau-Systeme findet man für die mittlere thermische Besetzung $f(E,T)$ den Ausdruck $f(E,T) = [1+\exp(E/k_{\mathrm{B}}T)]^{-1}$. Dieser Faktor entspricht der Fermi-Dirac-Verteilung, auf die wir in Kap. 7.1 eingehen. Auf den Temperaturverlauf der spezifischen Wärme haben die unterschiedlichen Verteilungen keinen Einfluß, da beide Funktionen als Argument jeweils die Größe $E/k_{\mathrm{B}}T$ enthalten.

Von großem Interesse ist natürlich die Frage nach der Anzahl der Tunnelsysteme. Eindeutige quantitative Aussagen sind nicht möglich, da oberhalb einiger Kelvin der Beitrag der Tunnelsysteme zur spezifischen Wärme zunehmend durch andere Beiträge überdeckt wird, die nicht separiert werden können. Grob läßt sich jedoch durch Vergleich mit den experimentellen Daten sagen, daß im Bereich $0 < E/k_B \leq 1\,\mathrm{K}$ etwa $10^{17} - 10^{18}$ Tunnelsysteme/cm^3 in amorphen Festkörpern vorhanden sind. Obwohl die Anzahl der Tunnelsysteme klein ist im Vergleich zur Gesamtzahl der Atome, bestimmen sie die meisten Eigenschaften der amorphen Festkörper bei Temperaturen unter 1 K. Bei höheren Temperaturen spielen wahrscheinlich lokalisierte Anregungen eine Rolle, die mit den soeben diskutierten Tunnelsystemen eng verwandt sind. Im Prinzip könnte es sich hierbei um vergleichbare Atomkonfigurationen handeln, denen jedoch die Barriere fehlt, welche die beiden Gleichgewichtslagen voneinander trennt. Das „Teilchen" schwingt dann in einem leicht gekrümmten Potential und ist nur schwachen rücktreibenden Kräften ausgesetzt. Dies bedeutet, daß die Schwingungsfrequenzen dieser Atomgruppen klein gegen die für kristalline Festkörper typische Debye-Frequenz sind. Ob diese Zustände auch für den „Bosonpeak" verantwortlich sind, ist noch nicht geklärt.

Wir haben hier unter der Vielzahl der amorphen Festkörper das Quarzglas als Beispiel herausgegriffen, da an dieser Substanz viele Untersuchungen durchgeführt worden sind. Doch gelten die Ausführungen, die hier gemacht wurden, mit leichten Modifikationen praktisch für alle amorphen Festkörper.

6. Anharmonische Gittereigenschaften

Bisher haben wir bei der Beschreibung der Gittereigenschaften von Kristallen harmonische Schwingungen vorausgesetzt, d.h. wir hatten angenommen, daß sich die Atome in einem streng harmonischen Potential bewegen. In diesem Fall können die resultierenden gekoppelten Schwingungen nach unabhängigen ebenen Wellen zerlegt werden[1]. In diesem stark idealisierenden Konzept findet keine Wechselwirkung zwischen den einzelnen Schwingungsmoden statt. Träfe dies tatsächlich zu, dann könnte eine Nichtgleichgewichtsverteilung der Phononen nicht abgebaut werden. Erst die Anharmonizität des Kristallgitters ermöglicht eine Einstellung des thermischen Gleichgewichtes und ruft den Wärmewiderstand der Dielektrika hervor. Die Anharmonizität äußert sich in vielen weiteren Festkörpereigenschaften, beispielsweise in der thermischen Expansion, den unterschiedlichen Werten von adiabatischen und isothermen Konstanten oder dem leichten Ansteigen der spezifischen Wärme bei Temperaturen oberhalb der Debye-Temperatur.

Im Prinzip lassen sich solche Nichtlinearitäten durch Berücksichtigung von Termen höherer Ordnung in der Entwicklung der Potentialfunktion (5.5) des Gitters beschreiben. Es ergeben sich jedoch Ausdrücke, die so kompliziert und unübersichtlich sind, daß sich damit nur schwerlich konkrete Eigenschaften berechnen lassen. Darüber hinaus wären die vielen dabei auftretenden Materialkonstanten auch experimentell kaum zu ermitteln. Um dennoch ein gewisses „Gefühl" für die wichtige Eigenschaft „Anharmonizität" zu entwickeln, wollen wir hier eine phänomenologische Beschreibung der thermischen Expansion und der Wechselwirkung zwischen den Phononen vornehmen, ohne den analytischen Weg über die Potentialentwicklung zu höherer Ordnung zu gehen.

6.1 Thermische Expansion

In der thermischen Expansion spiegelt sich direkt die Abweichung des Gitterpotentials von der rein harmonischen Form wider, denn ohne Anharmonizität wäre das Volumen eines Festkörpers temperaturunabhängig. Da die Amplitude der Gitterschwingungen mit der Temperatur zunimmt, und die Atome immer stärker die Anharmonizität des Gitterpotentials fühlen, ist ein Anwachsen des thermischen Expansionskoeffizienten mit der Temperatur zu erwarten. Wir wollen diesen Zusammenhang herleiten und übernehmen in Ermangelung adäquater mikroskopischer Modelle einen phänomenologischen Ansatz, der von *Grüneisen* vorgeschlagen wurde. In diesem Ansatz wird eine Verknüpfung zwischen der thermischen Expansion, der Kompressibilität und der spezifischen Wärme hergestellt.

Es liegt nahe, bei der Suche nach dem Zusammenhang zwischen thermodynamischen Variablen von einer Zustandsgleichung auszugehen. Wir leiten diese Gleichung für einen defektfreien Festkörper her und berücksichtigen dabei, daß die Anharmonizität des Gitters zu einer Volumenabhängigkeit der freien Energie führt. Wie wir gesehen haben, kann jeder Gitterwelle mit dem Wellenvektor $\boldsymbol{q}$ formal ein harmonischer Oszillator mit der

[1] Wie weit diese Vorgehensweise auch auf Gläser übertragbar ist, ist für kleine Wellenlängen noch nicht geklärt.

Schwingungsfrequenz $\omega_{\boldsymbol{q}}$ zugeordnet werden. Wir gehen bei der Herleitung der Zustandsgleichung zunächst von der mittleren freien Energie F eines harmonischen Oszillators aus, für die man mit Hilfe der Zustandssumme Z eines Oszillators den Ausdruck

$$F = -k_{\mathrm{B}}T \ln Z = k_{\mathrm{B}}T \ln\left(1 - \mathrm{e}^{-\hbar\omega/k_{\mathrm{B}}T}\right) + \frac{1}{2}\hbar\omega \tag{6.1}$$

findet[2]. Der zweite, temperaturunabhängige Term in (6.1) berücksichtigt die Nullpunktsenergie. Die freie Energie aller Phononen erhalten wir, indem wir die Beiträge aller Oszillatoren addieren, d.h. indem wir über alle Wellenvektoren $\boldsymbol{q}$ und alle Phononenzweige j summieren. Hinzu kommt noch die elastische Energie, die mit einer Volumenänderung $\delta V = (V - V_0)$ des Festkörpers verbunden ist, die in erster Näherung nicht von der Temperatur abhängt. Bezeichnen wir mit κ die Kompressibilität, so ist die freie Energie F des Festkörpers durch

$$F = \frac{V_0}{2\kappa}\left(\frac{\delta V}{V_0}\right)^2 + \sum_{\boldsymbol{q},j}\left[k_{\mathrm{B}}T \ln\left(1 - \mathrm{e}^{-\hbar\omega_{q,j}/k_{\mathrm{B}}T}\right) + \frac{1}{2}\hbar\omega_{\boldsymbol{q},j}\right] \tag{6.2}$$

gegeben. Durch Ableiten nach dem Volumen erhalten wir den gesuchten Zusammenhang zwischen Druck und Volumen:

$$p = -\left(\frac{\partial F}{\partial V}\right)_T = -\frac{1}{\kappa}\left(\frac{\delta V}{V}\right) - \hbar\sum_{\boldsymbol{q},j}\frac{\partial\omega_{\boldsymbol{q},j}}{\partial V}\left[\frac{1}{\mathrm{e}^{\hbar\omega_{q,j}/k_{\mathrm{B}}T} - 1} + \frac{1}{2}\right]. \tag{6.3}$$

Da nur kleine Volumenänderungen auftreten, wurde in dieser Gleichung V_0 durch V ersetzt. Die Anharmonizität macht sich dadurch bemerkbar, daß die Volumenänderung zu einer Änderung der Oszillatorfrequenzen führt. Verhielte sich der Festkörper vollkommen harmonisch, so wäre $(\partial\omega_{\boldsymbol{q},j}/\partial V) = 0$.

Um die Summation tatsächlich ausführen zu können, ist eine sehr starke Vereinfachung erforderlich. Sie besteht in der Annahme, daß die relative Frequenzänderung $\delta\omega_{\boldsymbol{q},j}/\omega_{\boldsymbol{q},j}$ der Gitterschwingungen proportional zur relativen Volumenänderung $\delta V/V$ ist, aber unabhängig von der Frequenz und vom Phononenzweig. Diese Annahme läßt sich durch

$$\frac{\delta\omega_{\boldsymbol{q},j}}{\omega_{\boldsymbol{q},j}} = -\gamma\frac{\delta V}{V} \tag{6.4}$$

zusammenfassen, wobei der dimensionslose *Grüneisen-Parameter* γ eine charakteristische Konstante des betrachteten Festkörpers ist. In differentieller Form können wir hierfür

$$\gamma = -\frac{\partial(\ln\omega_{\boldsymbol{q},j})}{\partial(\ln V)} \tag{6.5}$$

schreiben. Bei Zimmertemperatur liegen typische Werte der Grüneisen-Parameter bei etwa $\gamma \approx 1 - 2$, d.h. die Phononenfrequenzen und das Volumen ändern sich etwa gleich stark. Die Annahme, daß $\delta\omega_{\boldsymbol{q},j}/\omega_{\boldsymbol{q},j}$ nicht von der Frequenz abhängt, setzt eine gleichmäßige Änderung aller Komponenten des Kraftkonstantentensors (5.6) mit dem Volumen voraus. Dies kann nur eine Näherung der tatsächlichen Gegebenheiten sein. Noch problematischer ist die Annahme, daß sich alle Phononenzweige gleich verhalten, denn longitudinale und transversale Schwingungen reagieren auf Volumenänderungen unterschiedlich. Eine verbesserte Übereinstimmung zwischen Theorie und Experiment erreicht man, indem man für die unterschiedlichen Zweige individuelle Grüneisen-Parameter γ_j einführt.

[2] Die Bedeutung des Begriffes „Zustandssumme" und ihre Verknüpfung mit den thermodynamischen Größen wird erst in der *Statistischen Mechanik* erläutert. Wir benutzen hier dieses Konzept im Vorgriff, da die Herleitung der freien Energie eines Festkörpers sonst etwas langwierig ist.

Die Einführung des Grüneisen-Parameters ermöglicht die Vereinfachung von (6.3):

$$p = -\frac{1}{\kappa}\left(\frac{\delta V}{V}\right) + \frac{\gamma}{V}\sum_{\boldsymbol{q},j}\hbar\omega_{\boldsymbol{q},j}\left[\frac{1}{e^{(\hbar\omega_{q,j}/k_B T)}-1}+\frac{1}{2}\right] . \tag{6.6}$$

Man erkennt, daß die Summation nun über alle Phononenenergien, multipliziert mit der Besetzungswahrscheinlichkeit, erfolgt. Dies ist gerade die Energie aller angeregten Phononen und somit die innere Energie U einschließlich der Nullpunktsenergie. Multiplizieren wir diese Gleichung noch mit dem Volumen, so erhalten wir die *Zustandsgleichung*

$$pV = -\frac{\delta V}{\kappa} + \gamma U(T) . \tag{6.7}$$

Leiten wir die Zustandsgleichung bei konstantem Volumen nach der Temperatur ab und berücksichtigen, daß die spezifische Wärme durch $C_V = (\partial U/\partial T)_V$ gegeben ist, so ergibt sich $(\partial p/\partial T)_V = \gamma C_V/V$. Mit der Kettenregel $(\partial p/\partial T)_V \cdot (\partial T/\partial V)_p \cdot (\partial V/\partial p)_T = -1$, der Definition des linearen thermischen Expansionskoeffizienten $\alpha = (3V)^{-1}(\partial V/\partial T)_p$ und der Kompressibilität $\kappa = -(V)^{-1}(\partial V/\partial p)_T$ folgt der Zusammenhang $(\partial p/\partial T)_V = 3\alpha/\kappa$. Damit folgt aus (6.7) schließlich die *Grüneisen-Beziehung*

$$\alpha = \frac{\gamma\kappa C_V}{3V} . \tag{6.8}$$

Der thermische Expansionskoeffizient wird in unserer Näherung durch den Grüneisen-Parameter, durch die spezifische Wärme und die Kompressibilität bestimmt. Der Temperaturverlauf des Expansionskoeffizienten folgt fast ausschließlich dem von C_V, weil die Kompressibilität meist nur eine sehr schwache Temperaturabhängigkeit aufweist. Bei hohen Temperaturen, d.h. bei Zimmertemperatur oder darüber, ist daher der Expansionskoeffizient α praktisch konstant, da dort das Dulong-Petitsche Gesetz Gültigkeit besitzt. Bei tiefen Temperaturen dagegen hängt die spezifische Wärme und somit auch die thermische Expansion stark von der Temperatur ab. In Abb. 6.1 ist das Tieftemperaturverhalten der spezifischen Wärme und der thermischen Expansion für Kupferchlorid gezeigt. Tatsächlich verlaufen beide Kurven parallel und streben, der Debyeschen Theorie der spezifischen Wärme folgend, unter 3 K einem kubischen Temperaturverlauf zu.

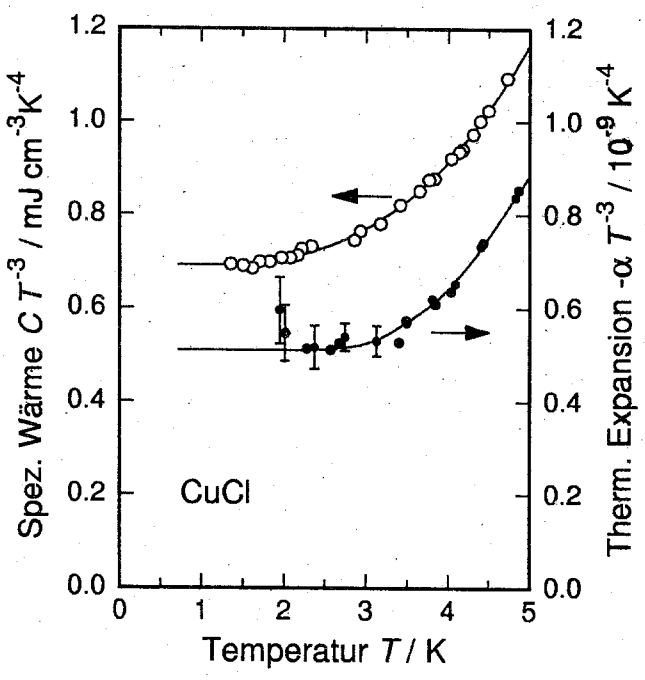


Abb. 6.1: Thermische Expansion und spezifische Wärme von Kupferchlorid. Es sind C/T^3 bzw. $-\alpha/T^3$ als Funktion der Temperatur aufgetragen. (Nach T.H.K. Barron, J.A. Birch, G.K. White, J. Phys. C **10** (1977) 1617).

Wie bereits oben erwähnt, basiert dieses Modell auf sehr starken Vereinfachungen. Da die Gitterpotentiale in den Kristallen meist stark anisotrop sind, wirken sich Volumenänderungen auf verschiedene Phononenzweige im allgemeinen unterschiedlich aus. So findet man für longitudinale Wellen gewöhnlich einen viel größeren Wert des Grüneisen-Parameters als für transversale. Der Grüneisen-Parameter kann auch negative Werte annehmen und unterschiedliches Vorzeichen für verschiedene Zweige und Kristallrichtungen besitzen. Dies wirkt sich natürlich auch auf die thermische Expansion aus. Beispielsweise ist in Abb. 6.1 $-\alpha/T^3$ aufgetragen, d.h. die thermische Expansion von Kupferchlorid ist bei tiefen Temperaturen negativ! Da die Kompressibilität bekannt ist, läßt sich aus den gezeigten Daten der Grüneisen-Parameter ermitteln. Man findet für $T \to 0$ den Wert $\gamma \approx -2{,}3$.

Entgegen der üblichen Vorstellung beobachtet man des öfteren einen negativen Grüneisen-Parameter. So zeigen viele Kristalle mit Zinkblendestruktur wie zum Beispiel Silizium zwar bei Zimmertemperatur einen positiven Grüneisen-Parameter, doch bei tiefen Temperaturen ist dieser dann negativ. Bemerkenswert ist auch das Verhalten von Tellur, das ein hexagonales Gitter besitzt und in Richtung der c-Achse einen negativen, in den anderen Richtungen dagegen einen positiven Ausdehnungskoeffizienten aufweist.

Obwohl wir Gleichung (6.8) nur für Phononen hergeleitet haben, hängt ihre Anwendbarkeit nicht davon ab, ob die freie Energie von Phononen oder von anderen Anregungen herrührt. Beispielsweise reagieren Gitterdefekte meist sehr empfindlich auf äußere Einwirkungen und ändern ihre Niveauschemata stark mit dem Volumen. Infolgedessen üben Gitterdefekte oft einen erheblichen Einfluß auf die anharmonischen Eigenschaften von Festkörpern aus. Wir betrachten in diesem Zusammenhang die thermische Expansion von *amorphen Substanzen*. In Kap. 5.7 haben wir gesehen, daß in diesen Materialien bei Temperaturen unter 1 K die spezifische Wärme und damit die innere Energie in erster Linie Tunnelsystemen und nicht Phononen zuzuschreiben ist. Ändert sich die Energieaufspaltung E der Tunnelsysteme analog zu (6.4), d.h. ist $\delta E/E = -\gamma\,(\delta V/V)$, so sollten diese Systeme auch die thermische Expansion bei tiefen Temperaturen mitbestimmen. Zur thermischen Ausdehnung tragen nun Phononen *und* Tunneldefekte bei, wenn auch ihre Grüneisen-Parameter unterschiedlich sind. Gleichung (6.8) sagt aus, daß eine Proportionalität zwischen dem Ausdehnungskoeffizienten und der spezifischen Wärme besteht. Da sich bei tiefen Temperaturen die spezifische Wärme amorpher Festkörper durch die Summe eines linearen und eines kubischen Terms nähern läßt, erwarten wir für die thermische Expansion die Beziehung $\alpha = aT + bT^3$, wobei a und b materialspezifische Konstanten sind, die durch die jeweiligen Werte von γ und C_V bestimmt werden. Wie Abb. 6.2 zeigt, ist dieser Temperaturverlauf tatsächlich bei amorphen Festkörpern gefunden worden. In diesem Bild ist die Größe α/T als Funktion von T^2 aufgetragen, um den linearen und den kubischen Term graphisch voneinander zu trennen. In allen Fällen findet man bei dieser Art der Auftragung innerhalb der Meßgenauigkeit eine Gerade: Der Achsenabschnitt wird durch die Konstante a, die Steigung durch b bestimmt.

Bemerkenswert ist der Verlauf der thermischen Expansion im Falle des sehr reinen Quarzglases, denn dort findet man den Wert $a = -1 \cdot 10^{-9}\,\mathrm{K}^{-2}$. Hieraus folgt für den Beitrag der Tunnelzentren zum Grüneisen-Parameter, der sich aus dem linearen Term der spezifischen Wärme ergibt, der sehr große negative Wert $\gamma = -65$. Die Energieaufspaltung der Tunnelzentren und damit ihre innere Energie reagiert offensichtlich äußerst empfindlich auf Volumen- bzw. Druckänderungen. Ergänzend soll noch bemerkt werden, daß die thermische Expansion von amorphen Materialien trotz des großen Absolutwertes ihres Grüneisen-Parameters bei Temperaturen unter 1 K extrem klein ist, denn die

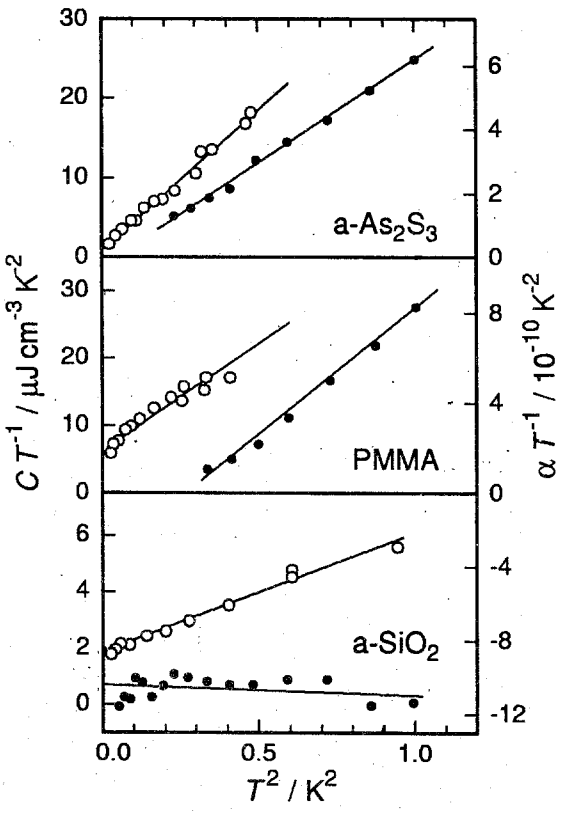


Abb. 6.2: Spezifische Wärme (offene Kreise) und linearer thermischer Expansionskoeffizient (Punkte) amorpher Materialien. Es ist C/T bzw. α/T als Funktion von T^2 aufgetragen. a-As_2S_3 ist ein halbleitendes Glas, PMMA (Plexiglas) ein Polymer. a-SiO_2 steht für ein Quarzglas hoher Reinheit. (Nach D.A. Ackerman, A.C. Anderson, E.J. Cotts, J.N. Dobbs, W.M. MacDonald, F.J. Walker, Phys. Rev. **29** (1984) 966).

spezifische Wärme ist, verglichen mit ihrem Wert bei Zimmertemperatur, ebenfalls verschwindend klein. Dies erklärt auch die starke Streuung der Meßpunkte in Abb. 6.2 trotz einer Auflösung der Meßanordnung von $2 \cdot 10^{-4}$ Å.

Aus unseren Ausführungen könnte man den Schluß ziehen, daß Grüneisen-Parameter üblicherweise durch die Messung der thermischen Expansion und der spezifischen Wärme bestimmt werden. Das ist jedoch nicht der Fall, denn der Grüneisen-Parameter einzelner Phononenzweige läßt sich auf „direktem Wege" messen, indem man das Volumen und damit nach (6.4) auch die Phononenfrequenzen der Probe ändert. So kann man beispielsweise den Grüneisen-Parameter optischer Phononen relativ einfach über die Druckverschiebung von Raman- oder Infrarotspektren bestimmen. Ultraschallexperimente erlauben das Studium der Anharmonizität von akustischen Zweigen. Dabei mißt man die Druckabhängigkeit der Schallgeschwindigkeit und berechnet hieraus unmittelbar den entsprechenden Grüneisen-Parameter. Beide Methoden erlauben aber nur die Untersuchung langwelliger Phononen. Ist man am Verhalten kurzwelliger Phononen interessiert, so muß im allgemeinen auf die Druckabhängigkeit der inelastischen Neutronenstreuung zurückgegriffen werden.

An dieser Stelle wollen wir noch kurz auf die *Temperaturabhängigkeit* der Phononenspektren eingehen. Bei den meisten Festkörpern, aber nicht bei allen, nimmt bei festem Wellenvektor die Frequenz mit zunehmender Temperatur ab. Bei den akustischen Phononen folgt daraus eine Reduktion der Schallgeschwindigkeit, d.h. die Festkörper werden weicher. Wie man dem Raman-Spektrum der Abb. 5.17 entnehmen kann, beobachtet man eine derartige Frequenzabnahme auch bei den optischen Phononen von Silizium. Wir wollen anhand dieser Daten den Grüneisen-Parameter dieses Phononenzweigs abzuschätzen.

Eine quantitative Beschreibung der Temperaturabhängigkeit des Phononspektrums erfordert größeren theoretischen Aufwand. Ein qualitatives Verständnis aber läßt sich leicht mit Hilfe der Beziehung (6.4) und der Definition des thermischen Expansionskoeffizienten erzielen, denn für den Temperaturgang der Phononenfrequenzen finden wir damit den einfachen Ausdruck

$$\frac{\partial\omega}{\partial T} = \left(\frac{\partial\omega}{\partial V}\right)\left(\frac{\partial V}{\partial T}\right) = -3\gamma\alpha\omega \; . \tag{6.9}$$

Bei positivem Wert von γ nehmen die Phononenfrequenzen mit der Temperatur ab. Entsprechend verringert sich die Schallgeschwindigkeit bzw. die Frequenzverschiebung der Raman-Linie verkleinert sich. Da bei hohen Temperaturen der thermische Expansionskoeffizient näherungsweise konstant ist, nimmt die Frequenzverschiebung proportional zur Temperatur ab. Wie bereits oben vermerkt, kehrt sich bei tiefen Temperaturen das Vorzeichen des Grüneisen-Parameters von Silizium um, d.h. die Frequenzverschiebung wird dann mit abnehmender Temperatur kleiner. Diese Beobachtung läßt sich aber Abb. 5.17 nicht entnehmen, da die beobachteten Effekte bei tiefen Temperaturen nur sehr klein sind.

Wir wollen noch den Grüneisen-Parameter der optischen Phononen von Silizium abschätzen. Dazu nehmen wir an, daß der Expansionskoeffizient α in etwa konstant ist. Damit können wir Gleichung (6.9) umschreiben und finden in erster Näherung: $\delta\omega/\omega \approx -3\gamma\alpha\,\delta T$. Mit der Frequnzverschiebung aus der Abb. 5.17 und der Kenntnis, daß sich zwischen 460 K und 770 K die relative Änderung des Volumens $\delta V/V \approx 0{,}001$ beträgt, finden wir den relativ hohen Wert von $\gamma \approx 3{,}5$. Zum Schluß wollen wir noch einen Blick auf die Temperaturabhängigkeit der Zustandsdichte von Quarzglas werfen, die in Abb. 5.30 dargestellt war. Dort haben wir gesehen, daß sich das Maximum bei 0,94 THz, der Bosonpeak, zwischen 51 K und 1104 K nach 1,28 THz verschiebt. Gleichzeitig tritt eine Volumenänderung $\delta V/V \approx 0{,}0019$ auf. Dies bedeutet, daß der Grüneisen-Parameter negativ ist und den enormen Wert von $\gamma \approx -180$ besitzt. Eine Erklärung für diese Beobachtung gibt es gegenwärtig nicht.

6.2 Phonon-Phonon-Wechselwirkung

Drei-Phononen-Prozeß

Wie bereits am Anfang dieses Kapitels erwähnt, tritt in der harmonischen Näherung keine Wechselwirkung zwischen den Normalschwingungen des Gitters auf. Die Existenz nicht-quadratischer Terme im Gitterpotential zerstört jedoch die Unabhängigkeit der Gitterschwingungen. Die Wechselwirkung der Phononen untereinander ist deshalb eine besonders wichtige Konsequenz der Anharmonizität des Gitterpotentials. Wegen der Komplexität der Verhältnisse gibt es jedoch keine voll ausgearbeitete Theorie. Wir wollen daher hier nur einige wichtige Prozesse herausgreifen und qualitativ diskutieren.

Der einfachste und wichtigste Wechselwirkungsprozeß ist der *Drei-Phononen-Prozeß*. Sein Auftreten läßt sich eindrucksvoll in einem Ultraschallexperiment demonstrieren, dessen prinzipieller Aufbau in Abb. 6.3 zu sehen ist. Die von den Schallwandlern (1) und (2) erzeugten Wellenbündel wechselwirken in der Überlagerungszone und rufen aufgrund der

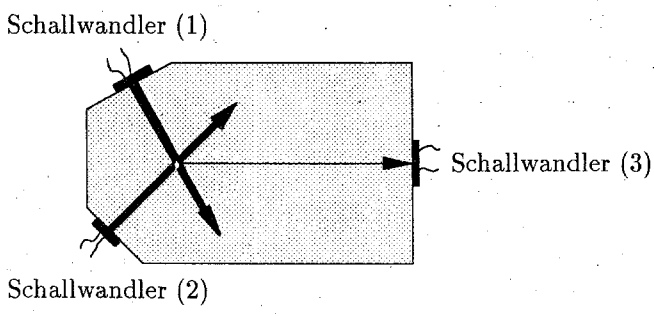


Abb. 6.3: Meßanordnung zum Nachweis der Phonon-Phonon-Wechselwirkung mit Ultraschalltechnik. Der Schallwandler (1) erzeugt transversale Wellen mit einer Frequenz von 10 MHz. Die Schallwandler (2) und (3) arbeiten bei 15 MHz und 25 MHz und erzeugen longitudinale Schallwellen bzw. weisen diese nach. (Nach F.R. Rollins, Jr. und L.H. Todd, Jr., Phys. Rev. **136** (1964) A597).

Gitteranharmonizität eine neue Welle hervor, die mit dem Schallwandler (3) nachgewiesen wird. Bei dem skizzierten Experiment haben die wechselwirkenden Schallwellen unterschiedliche Frequenzen und Polarisationen. Auf die Gründe hierfür werden wir im folgenden Abschnitt kurz eingehen.

Aus den gewählten Frequenzen und aus der Geometrie des experimentellen Aufbaus kann man bei diesem Experiment ablesen, daß Energie- und Quasiimpulserhaltung für die wechselwirkenden Phononen erfüllt sind, d.h. daß die Gleichungen

$$\hbar\omega_1 + \hbar\omega_2 = \hbar\omega_3 \qquad \text{und} \qquad \hbar\boldsymbol{q}_1 + \hbar\boldsymbol{q}_2 = \hbar\boldsymbol{q}_3 \tag{6.10}$$

gelten. Dieses Experiment zeigt nicht nur in eindrucksvoller Weise, daß Phononen untereinander wechselwirken, sondern auch, daß Ultraschallexperimente besonders geeignet sind, die Wechselwirkung zwischen den Phononen zu untersuchen.

Ultraschalldämpfung in Kristallen

Die Ausbreitung von Phononen wird durch Streuprozesse behindert. Bei dielektrischen Kristallen tragen hierzu vor allem zwei Mechanismen bei: die Streuung an Defekten und die Wechselwirkung der Phononen untereinander. In Ultraschallexperimenten machen sich die Streuprozesse durch das Auftreten von Dämpfung bemerkbar. In Abb. 5.1 haben wir die prinzipielle Meßanordnung für derartige Experimente bereits skizziert. Aus dem Abfall der Signalamplitude mit zunehmender Laufstrecke bzw. mit der Laufzeit läßt sich direkt der Dämpfungskoeffizient bzw. die freie Weglänge l der Ultraschallwellen bestimmen.

Bei der Behandlung der Phononenstreuung benutzen wir das Teilchenbild, mit dem sich die auftretenden Prozesse sehr anschaulich beschreiben lassen[3]. Wir gehen von dem bei Streuprozessen allgemein gültigen Zusammenhang

$$l = \frac{1}{\tilde{n}\,\sigma} \qquad \text{bzw.} \qquad l^{-1} = \tilde{n}\,\sigma \tag{6.11}$$

aus. Die freie Weglänge der stoßenden Teilchen ist umgekehrt proportional zur Dichte $\tilde{n}$ der Streuzentren und zum totalen Streu- oder Stoßquerschnitt σ. Der Streuquerschnitt selbst ist proportional zum Quadrat $|\mathcal{A}|^2$ der Streuamplitude, die wir in Abschnitt 2.2 eingeführt haben.

[3] Bei der Ultraschallausbreitung handelt es sich um ein Transportphänomen, das theoretisch am besten mit Hilfe der „Boltzmannschen Transportgleichung" beschrieben wird. Wir wollen uns hier wie auch bei der Diskussion der Wärmeleitung mit einer vereinfachten Betrachtungsweise begnügen. Auf die Transportgleichung gehen wir in Kap. 8.2 im Zusammenhang mit der elektrischen Leitfähigkeit ein, da dort eine anschaulichere Einführung möglich ist.

Wir wollen zunächst die elastische *Streuung an Punktdefekten* betrachten, an der sich die prinzipielle Vorgehensweise besonders gut verdeutlichen läßt. Anschließend beschäftigen wir uns mit der *Phonon-Phonon-Streuung*, die inelastisch ist und durch die Anharmonizität des Gitters hervorgerufen wird. Wir werden die hier erarbeiteten Ergebnisse nochmals in Kap. 6.3 bei der Untersuchung der Wärmeleitfähigkeit verwenden.

Nehmen wir an, daß der untersuchte Kristall eine größere Zahl von Punktdefekten, z.B. Leerstellen oder Zwischengitteratome, enthält. Die Ultraschallwelle wird dann an diesen Defekten elastisch gestreut. Bezeichnen wir die Phononen vor und nach der Streuung mit den Indizes 1 und 2, so gelten Energie- und Quasiimpulssatz: $\omega_1 = \omega_2$ bzw. $\boldsymbol{q}_1 = \boldsymbol{q}_2 + \boldsymbol{K}$. Bei der Streuung wird der Quasiimpuls $\hbar\boldsymbol{K}$ auf den Defekt übertragen. Für den Betrag des übertragenen Impulses gilt $0 < K \leq 2q$, wobei das Gleichheitszeichen bei Rückstreuung auftritt. Wie bei der Streuung von Licht an kleinen Teilchen läßt sich auch hier die Streuamplitude bzw. der Streuquerschnitt mit Hilfe der Theorie von Rayleigh berechnen. Unter der Annahme, daß die Wellenlänge groß im Vergleich zum störenden Defekt ist, d.h. für $\lambda \gg a$, findet man den Zusammenhang $\mathcal{A}^2 \propto \sigma \approx \pi a^2 (aq)^4$, wobei a für die Ausdehnung des Defekts steht, die bei atomaren Punktdefekten von der Größenordnung des Atomabstands ist. Der genaue Wert des Streuquerschnitts hängt von den elastischen Eigenschaften des Kristalls und den speziellen Eigenschaften des Defekts ab.

Wir interessieren uns hier nicht näher für den Absolutwert des Effektes, sondern nur für dessen Frequenz- und Temperaturabhängigkeit. Da für langwellige Phononen $\omega \propto q$ und die Streuzentrendichte $\tilde{n}$ durch die Punktdefektdichte n_D gegeben ist, ergibt sich mit (6.11) für die inverse freie Weglänge der Ultraschallphononen

$$l^{-1} \propto n_\mathrm{D}\, \omega^4 \,. \tag{6.12}$$

Dieser Gleichung entnehmen wir zunächst, daß Defekte vor allem bei hohen Frequenzen wichtig werden. Darüber hinaus ist die rein geometrische Streuung unabhängig von der Temperatur, d.h. Defekte liefern bei einer Untersuchung des Temperaturverlaufs der Ultraschalldämpfung nur einen konstanten Untergrund. Bei stark verunreinigten Kristallen kann die Streuung an Defekten sogar die restlichen Dämpfungsbeiträge überwiegen.

Sehen wir uns nun den in Abb. 6.4 gezeigten Temperaturverlauf der Ultraschalldämpfung in einem Germaniumkristall an, der für reine dielektrische Kristalle[4] typisch ist. Bei tiefen Temperaturen ist die Dämpfung verschwindend klein, steigt dann aber rasch an. Verursacht wird sie durch Drei-Phononen-Prozesse: die Ultraschallphononen stoßen mit thermischen Phononen und werden dabei vernichtet. Die unterschiedliche Stärke der Dämpfung von longitudinalen und transversalen Wellen beruht in erster Linie auf ihren unterschiedlichen Ausbreitungsgeschwindigkeiten und soll uns hier nicht weiter interessieren.

Die Behandlung der Phononenstreuung ist ein sehr anspruchsvolles theoretisches Problem, das umfangreiche Rechnungen erfordert. Wir wollen hier nur die Grundzüge der Theorie ansprechen. Zur Vereinfachung setzen wir einen isotropen Festkörper mit linearer Dispersionskurve voraus. Wichtig für die weiteren Betrachtungen ist, daß die entscheidende Größe für die Stärke der Phonon-Phonon-Wechselwirkung und damit für den Streuquerschnitt *nicht* die *Auslenkung* der Atome aus ihrer Ruhelage ist. Werden nämlich alle Atome gemeinsam ausgelenkt, so ändert sich der Abstand zwischen den Atomen nicht, und Anharmonizitäten spielen keine Rolle. Entscheidend sind die Abstandsänderungen

[4] Bei Metallen dominiert meist Dämpfung aufgrund der Wechselwirkung mit den freien Elektronen.

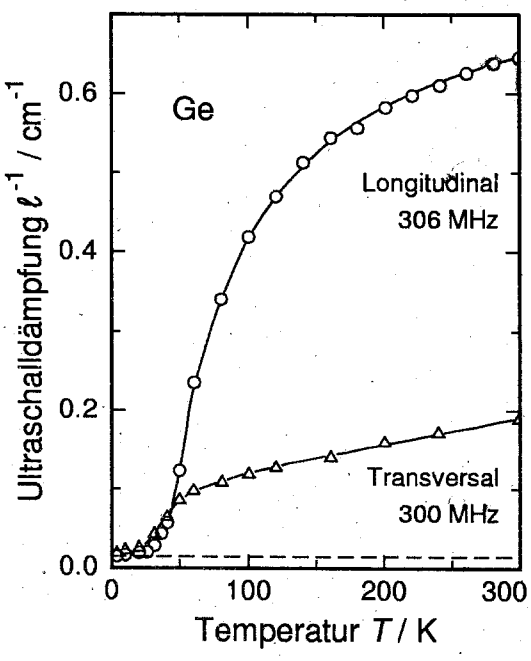


Abb. 6.4: Temperaturabhängigkeit der Ultraschalldämpfung in Germanium bei 300 MHz. (Nach W.P. Mason in Physical Acoustics Band III/B (W.P. Mason ed.), Academic Press, New York (1965) 235).

zwischen den Atomen, die durch den *Verzerrungstensor* $[e]$ beschrieben werden und dessen Komponenten durch die Beziehung $e_{ij} = (\partial u_i/\partial x_j + \partial u_j/\partial x_i)/2$ definiert sind, wobei die Indizes i und j für die kartesischen Komponenten x, y, z stehen. Die Verzerrung, die *ein* Phonon verursacht, hängt davon ab, welche atomare Auslenkung es bewirkt und welche Frequenz es besitzt. Um die Herleitung dieses Zusammenhangs möglichst einfach zu gestalten, betrachten wir eine longitudinale dispersionslose ebene Welle, die in x-Richtung läuft. Da in diesem Fall die Teilchenauslenkung $u_x = U\sin(\omega t - qx)$ in Ausbreitungsrichtung erfolgt, besitzt der Verzerrungstensor nur die Komponente $e_{xx} = \partial u_x/\partial x$, und somit ergibt sich für die Verzerrungsamplitude $e_0 = qU$. Bei gleicher Amplitude der Teilchenauslenkung nimmt somit die Verzerrung linear mit dem Wellenvektor zu. Die Schwingungsamplitude eines einzelnen Phonons ist nach Gleichung (5.38) durch $U^2 = 2\hbar/(\varrho V\omega)$ gegeben. Bei linearer Dispersion führen diese Zusammenhänge zu einer Verzerrungsamplitude $e_0 \propto \sqrt{\omega}$. Es soll hier noch bemerkt werden, daß diese Beziehung für alle Phononen gilt, also nicht auf den herausgegriffenen Spezialfall beschränkt ist.

Ohne die erforderlichen Schritte nachzuvollziehen, wollen wir festhalten, daß die Streuamplitude $\mathcal{A}$ proportional zur Verzerrung aller wechselwirkenden Phononen ist. Dieser Zusammenhang liegt nahe, da die Anharmonizität des Gitterpotentials sich bei einer Abweichung des Atomabstands vom Gleichgewichtswert bemerkbar macht. Für den Streuquerschnitt gilt daher beim Drei-Phononen-Prozeß die Proportionalität:

$$\sigma \propto |\mathcal{A}|^2 \propto (e_{0,\mathrm{us}}\, e_{0,2}\, e_{0,3})^2 \propto \omega_{\mathrm{us}}\, \omega_2\, \omega_3 \,. \tag{6.13}$$

Wir schreiben hier den Index „us“ anstelle der in (6.10) eingeführten 1, da die einfallenden Teilchen Ultraschallphononen sind. Der Ausdruck für den Absolutwert des Streuquerschnitts enthält Faktoren mit der zweiten und dritten Ableitung der Potentialfunktion (5.5) des Gitters. Die auftretenden Konstanten hängen von der Kristallsymmetrie ab und unterscheiden sich für die verschiedenen Phononenzweige.

Für $\tilde{n}$ können wir nicht einfach eine Zahl einsetzen, wie wir es bei der Punktdefektstreuung getan haben, denn bei der Wechselwirkung zwischen Phononen kommen mehrere Aspekte ins Spiel. Offensichtlich ist, daß bei der Bestimmung der Zahl der wirksamen thermischen Phononen das ganze Frequenzspektrum berücksichtigt werden muß. Es muß also über die Dichte der thermischen Phononen integriert werden, die durch die Zustandsdichte und durch die mittlere Besetzungzahl, also durch $D(\omega)\langle n(\omega,T)\rangle$ bestimmt wird. Weniger offensichtlich ist, daß sich hinter den Erhaltungssätzen (6.10) sechs unterschiedliche Prozesse verbergen. Wir wollen hiervon zwei herausgreifen, die verbleibenden vier kann man sich leicht überlegen, wenn man die Reihenfolge der Phononen beim Wechselwirkungsprozeß vertauscht. Der Stoßprozeß, der bisher angesprochen wurde, ist der Stoß des Ultraschallphonons $\boldsymbol{q}_{\mathrm{us}}$ mit dem thermischen Phonon $\boldsymbol{q}_2$, wobei das Phonon $\boldsymbol{q}_3$ erzeugt wird. Dieser Prozeß kann auch in umgekehrter Richtung ablaufen, wobei ein Ultraschallphonon erzeugt wird: Ein thermisches Phonon $\boldsymbol{q}_3$ zerfällt unter dem Einfluß der Ultraschallwelle in ein zusätzliches Ultraschallphonon $\boldsymbol{q}_{\mathrm{us}}$ und ein thermisches Phonon $\boldsymbol{q}_2$. Wären gleich viele Phononen mit den Wellenvektoren $\boldsymbol{q}_2$ und $\boldsymbol{q}_3$ vorhanden, so liefen beide Prozesse mit der gleichen Wahrscheinlichkeit ab, und ihr Auftreten würde die Ultraschallwelle weder schwächen noch verstärken. In den Ausdruck für die inverse freie Weglänge geht aber nicht die Dichte der thermischen Phononen bei der entsprechenden Frequenz, sondern die Differenz $D(\omega_2)\langle n(\omega_2)\rangle - D(\omega_3)\langle n(\omega_3)\rangle$ ein. Da sich die beiden Frequenzen ω_2 und ω_3 nur geringfügig um die Ultraschallfrequenz $\omega_{\mathrm{us}} = \omega_3 - \omega_2$ voneinander unterscheiden, können wir in erster Näherung $D(\omega_3) \approx D(\omega_2)$ setzen und die mittlere Besetzungszahl in einer Taylor-Reihe entwickeln:

$$D(\omega_2)\langle n(\omega_2)\rangle - D(\omega_3)\langle n(\omega_3)\rangle \approx D(\omega_2)\left\{\langle n(\omega_2)\rangle - \left[\langle n(\omega_2)\rangle + \left.\frac{\partial\langle n(\omega)\rangle}{\partial\omega}\right|_{\omega_2} \omega_{\mathrm{us}} + \ldots\right]\right\}$$

$$\approx -\omega_{\mathrm{us}}\, D(\omega_2) \left.\frac{\partial\langle n(\omega)\rangle}{\partial\omega}\right|_{\omega_2} . \tag{6.14}$$

Wir wollen hier die Herleitung des Dämpfungskoeffizienten nicht weiter ausführen, sondern nur den Weg und das Ergebnis verständlich machen. Zunächst müssen die restlichen Streuprozesse mit einbezogen werden, was auf analoger Weise geschieht. Anschließend wird über alle Richtungen und Streuwinkel unter Berücksichtigung der Energieerhaltung integriert. Hierbei ergibt sich ein Faktor $1/\omega_{\mathrm{us}}$, der den Faktor ω_{us} in (6.14) kompensiert. Setzen wir $\omega_2 \approx \omega_3 = \omega$, so läßt sich das Ergebnis der Rechnungen in folgenden Ausdrücken zusammenfassen:

$$l^{-1} \propto \int \sigma(\omega)\, D(\omega)\, \frac{\partial\langle n(\omega,T)\rangle}{\partial\omega}\, \mathrm{d}\omega \propto \omega_{\mathrm{us}} \int \omega^4\, \frac{\partial\langle n(\omega,T)\rangle}{\partial\omega}\, \mathrm{d}\omega$$

$$\propto \omega_{\mathrm{us}}\, T^4 \int_0^{x_{\mathrm{D}}} x^4\, \frac{\partial\langle n(x)\rangle}{\partial x}\, \mathrm{d}x \propto \omega_{\mathrm{us}}\, T^4 \int_0^{x_{\mathrm{D}}} \frac{x^4 \mathrm{e}^x}{(\mathrm{e}^x-1)^2}\, \mathrm{d}x\,, \tag{6.15}$$

wobei wir wie bei der Berechnung der spezifischen Wärme in Abschnitt 5.5 die Abkürzung $x = \hbar\omega/k_{\mathrm{B}}T$ bzw. $x_{\mathrm{D}} = \hbar\omega_{\mathrm{D}}/k_{\mathrm{B}}T$ benutzt haben. Dort ist das gleiche Integral aufgetreten, so daß wir die Auswertung von dort direkt übernehmen können. Bei tiefen Temperaturen geht $x_{\mathrm{D}} \to \infty$, und das Integral liefert nur eine Konstante. Wir erhalten daher für die Frequenz- und Temperaturabhängigkeit:

$$l^{-1} \propto \omega_{\mathrm{us}}\, T^4\,. \tag{6.16}$$

Dieser Verlauf der Dämpfung wird experimentell bei tiefen Temperaturen beobachtet und als *Landau-Rumer-Dämpfung* bezeichnet.

Wie in Abb. 6.4 zu sehen ist, flacht der Dämpfungsanstieg bei hohen Temperaturen, d.h. in der Nähe der Debye-Temperatur, ab. Bei diesen Temperaturen macht sich bereits die Begrenzung des Phononspektrums durch die Debye-Frequenz ω_D bemerkbar. Wie bei der spezifischen Wärme steigt das Maximum in der spektralen Phononenverteilung nicht mehr proportional zur Temperatur an. Da das thermisch angeregte Spektrum nun in erster Linie aus Phononen mit der Debye-Frequenz besteht, wird der Streuquerschnitt (6.13) näherungsweise konstant: $\sigma \propto \omega_{us}\,\omega_D^2$. Die Gesamtzahl der Phononen steigt nach (5.66) nur noch proportional zur Temperatur an, so daß auch für die inverse freie Weglänge dieser Verlauf zu erwarten ist. Geht man bei der Auswertung des Integrals in (6.15) wie in (5.61) vor, so findet man wie erwartet

$$l^{-1} \propto \omega_{us}\, T\,. \tag{6.17}$$

Bei einer genaueren Betrachtung der Ultraschalldämpfung bei Zimmertemperatur muß jedoch noch berücksichtigt werden, daß neben dem Drei-Phononen-Prozeß weitere Dämpfungsmechanismen auftreten können, die zu einer Erhöhung der beobachteten Dämpfung führen, den prinzipiellen Temperaturverlauf aber nicht verändern.

Gleichung (6.16) gilt nicht für longitudinale Schallwellen, für deren Dämpfung man bei tiefen Temperaturen eine noch stärkere Temperaturabhängigkeit beobachtet. Der Grund hierfür ist, daß Energie- und Quasiimpulserhaltung gewisse „Auswahlregeln" erzwingen, auf die wir bereits bei der Diskussion der Abb. 6.3 hinwiesen, die wir aber bei der Herleitung von (6.16) nicht beachtet haben. Schon eine einfache Überlegung zeigt, daß Prozesse mit Beteiligung von Phononen nur *eines* Phononzweiges weitestgehend unterdrückt werden. Die beiden Erhaltungssätze (6.10) sind nur dann erfüllbar, wenn die Dispersionskurve strikt linear ist und der Stoß kollinear erfolgt. Diese Forderung nach der Kollinearität schränkt die Zahl der möglichen Stoßpartner erheblich ein. Darüber hinaus haben wir in in Kap. 5.3 gesehen, daß die Dispersionskurven immer gekrümmt sind, so daß die Forderung nach der gleichzeitigen Erfüllung der beiden Erhaltungssätze den Stoß zwischen Phononen des gleichen Zweiges unmöglich macht. Die Wechselwirkung erfolgt daher zwischen Phononen unterschiedlicher Zweige, denn Energie- und Impulssatz können in diesem Fall durch geeignete Wahl des Streuwinkels leicht gleichzeitig erfüllt werden, da Phononen unterschiedlicher Zweige bei gleicher Frequenz verschiedene Wellenvektoren besitzen. Die Auswahlregeln müssen natürlich bei der Herleitung der Ultraschalldämpfung berücksichtigt werden und bewirken Modifikationen bei den oben angestellten einfachen Überlegungen. Eine weitere Aufweichung der stark einschränkenden Auswahlregeln kommt dadurch zustande, daß die Dispersionskurven aufgrund der endlichen Lebensdauer der Phononen bereits eine gewisse Breite aufweisen. Dieser Effekt erlaubt im begrenzten Umfang die Wechselwirkung zwischen Phononen des gleichen Zweigs und spielt bei der Dämpfung longitudinaler Schallwellen bei tiefen Temperaturen eine wichtige Rolle.

Spontaner Phononenzerfall

Wir wollen noch die interessante Frage diskutieren, ob ein Phonon in einem defektfreien Kristall am absoluten Nullpunkt unendlich lange lebt. Dies könnte man zunächst vermuten, da keine thermischen Phononen vorhanden sind, mit denen das Phonon in Wechselwirkung treten könnte. Die Antwort ist etwas überraschend: Das Phonon erfährt einen spontanen Zerfall. In anderen Worten ausgedrückt: In der Nähe des absoluten Nullpunkts

macht sich die Wechselwirkung des Phonons mit den Nullpunktsschwingungen des Gitters bemerkbar und führt zu einer endlichen Lebensdauer. Abb. 6.5 zeigt, daß die Lebensdauer hochfrequenter Phononen in CaF_2 bei 2 K aufgrund dieses Zerfallsprozesses relativ kurz ist und stark von der Phononenfrequenz abhängt[5].

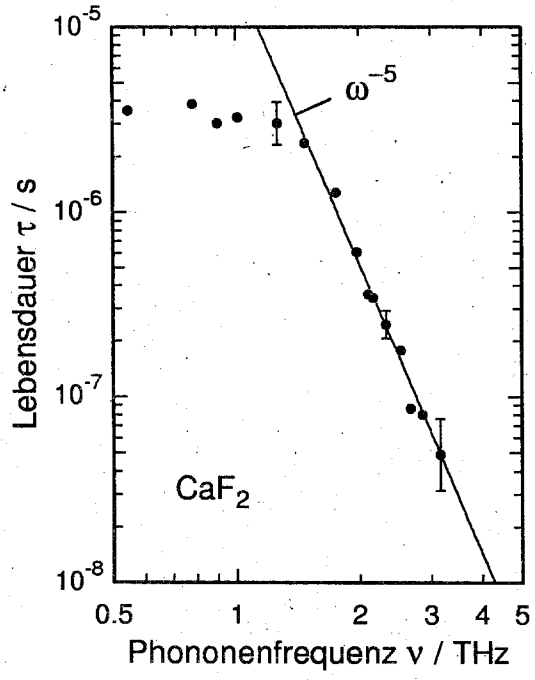


Abb. 6.5: Frequenzabhängigkeit der Lebensdauer von THz-Phononen in CaF_2, gemessen bei 2 K. (Nach K.F. Renk, *Excited-State Spectroscopy in Solids*, XCVII Corso (1987) 214).

Grob gesprochen, zerfällt ein Phonon mit der Frequenz ω in zwei Phononen mit etwa der halben Frequenz, da der Phasenraum für diesen Endzustand am größten ist. Da es sich hier wieder um einen Drei-Phononen-Prozeß handelt, läßt sich die freie Weglänge auf ganz ähnliche Weise berechnen, wie bei der normalen Ultraschalldämpfung. Man erhält $l^{-1} \propto \omega^5$ und für die Lebensdauer $\tau = vl$ die Beziehung

$$\tau \propto \omega^{-5} \,. \tag{6.18}$$

Die starke Frequenzabhängigkeit des Prozesses bewirkt, daß der spontane Zerfall nur bei sehr hochfrequenten Phononen und nicht bei den üblichen Ultraschallfrequenzen einen merklichen Beitrag zum Dämpfungskoeffizienten liefert. In der Messung, die in Abb. 6.5 gezeigt ist, war $\hbar\omega \gg k_B T$, so daß bereits bei der relativ hohen Temperatur von 2 K der Zerfall und nicht die Streuung der Phononen dominiert. Das Abknicken der Kurve unterhalb von 1,5 THz beruht auf einem zusätzlichen Effekt, der die Lebensdauer verkürzt. Es ist die resonante Absorption der Phononen durch Eu^{2+}-Verunreinigungen, die zur Phononenerzeugung benutzt werden. Auf diesen Wechselwirkungsmechanismus gehen wir im folgenden Abschnitt im Zusammenhang mit der Ultraschalldämpfung von amorphen Festkörpern ein.

[5] In der dargestellten Messung wurden die Phononen über die Fluoreszenz von Eu^{2+}-Verunreinigungen optisch erzeugt und nachgewiesen.

Ultraschalldämpfung in amorphen Festkörpern

Kristalle und amorphe Festkörper zeigen einen völlig unterschiedlichen Temperaturverlauf der Ultraschalldämpfung. Obwohl die Dämpfungsmechanismen in amorphen Materialien bei höheren Temperaturen noch nicht völlig verstanden sind, ist sicher, daß Phonon-Phonon-Streuprozesse, die in reinen dielektrischen Kristallen die Dämpfung bestimmen, nur eine untergeordnete Rolle spielen.

Wir wollen hier die Ultraschallabsorption von Gläsern bei tiefen Temperaturen etwas näher betrachten. Während in reinen Kristallen, wie soeben diskutiert, die Ultraschalldämpfung bei tiefen Temparaturen extrem klein wird, tritt in amorphen Materialien eine sehr starke Dämpfung auf. Die Ursache hierfür ist die resonante Wechselwirkung zwischen der Schallwelle und den Tunnelzentren, die wir bereits in Kap. 5.7 bei der Diskussion der spezifischen Wärme kennengelernt haben. Prinzipiell läuft dieser Prozeß wie folgt ab: Stimmt die Energie $\hbar\omega$ der Ultraschallphononen mit der Energieaufspaltung E von Tunnelzentren überein, so werden Übergänge zwischen den beiden Niveaus induziert, wie in Abb. 6.6 schematisch dargestellt.

Abb. 6.6: Schematische Darstellung der Übergänge bei resonanter Wechselwirkung zwischen Phononen und Tunnelsystemen (Zwei-Niveau-Systemen).

Auch hier verläuft der Streuprozeß in beide Richtungen: Entweder wird ein Ultraschallphonon vernichtet oder erzeugt. Wie effektiv der jeweilige Prozeß ist, hängt von der Besetzungsdifferenz der beiden Niveaus ab. Bei $T \to 0$ wird die freie Weglänge minimal, da sich alle Tunnelsysteme im Grundzustand befinden und somit nur der Prozeß übrigbleibt, bei dem Phononen vernichtet werden. Ist $k_\mathrm{B}T \gg E$, so sind die beiden Niveaus näherungsweise gleich besetzt, und es werden Übergänge in beiden Richtungen mit gleicher Häufigkeit induziert. Die Schallwelle erfährt dann im Mittel durch die Anwesenheit der Tunnelsysteme keine resonante Dämpfung. In die Berechnung der freien Weglänge geht also analog zum Drei-Phononen-Prozeß die Differenz der Besetzungszahlen der beiden Niveaus ein. Ist n die Dichte der betrachteten Tunnelsysteme mit der Energieaufspaltung E, n_1 und n_2 die Dichte der Systeme im oberen bzw. unteren Niveau, so gilt $n = n_1 + n_2$ und $n_1/n_2 = \exp(-E/k_\mathrm{B}T)$. Damit errechnet sich für die Besetzungszahldifferenz $\delta n = n_2 - n_1$ der Ausdruck

$$\delta n = n \tanh \frac{E}{2k_\mathrm{B}T} \,. \tag{6.19}$$

Da beim resonanten Prozeß $E = \hbar\omega_\mathrm{us}$ ist, erhalten wir mit (6.11) für die inverse freie Weglänge

$$l^{-1} = \tilde{n}\,\sigma = \delta n\,\sigma = C'\omega_\mathrm{us} \tanh \frac{\hbar\omega_\mathrm{us}}{2k_\mathrm{B}T} \,. \tag{6.20}$$

Für die Konstante C' findet man $C' = \overline{P}\tilde{\gamma}^2/\rho v^3$. Neben Dichte und Schallgeschwindigkeit, bestimmen die Zustandsdichte $\overline{P}$ der Tunnelsysteme und der Parameter $\tilde{\gamma}$ den Absolutwert der Dämpfung. Der letztere Parameter gibt die Stärke der Kopplung zwischen den Phononen und Tunnelsystemen an und ist proportional zum Grüneisen-Parameter der Tunnelsysteme. Die Proportionalität zur Ultraschallfrequenz beruht darauf, daß am Absorptionsprozeß nur das Ultraschallphonon beteiligt ist, d.h. es gilt $\sigma \propto e_0^2 \propto \omega_{\mathrm{us}}$. Von der spezifischen Wärme der amorphen Festkörper (vgl. Kap. 5.7) ist bekannt, daß eine breite Verteilung der Energieaufspaltungen der Tunnelsysteme existiert. Wir erwarten also bei allen Meßfrequenzen eine Dämpfung, deren Temperaturverlauf, wie in Abb. 6.7 skizziert, durch die thermische Population der beiden Niveaus gegeben ist.

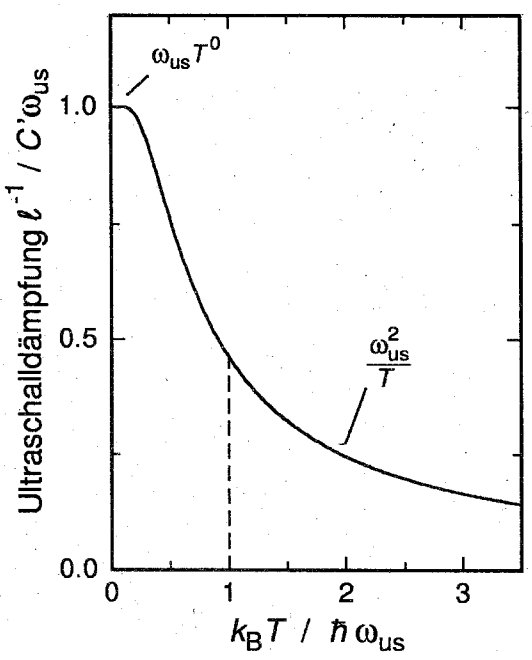


Abb. 6.7: Verlauf der resonanten Ultraschalldämpfung in Gläsern, normiert auf den Wert bei $T = 0$, als Funktion der reduzierten Temperatur $k_B T/\hbar\omega_{\mathrm{us}}$.

Die Grenzfälle für hohe und tiefe Temperaturen lassen sich leicht angeben. Bei sehr tiefen Temperaturen ist $k_B T \ll \hbar\omega_{\mathrm{us}}$ und der Besetzungsfaktor $\tanh(\hbar\omega_{\mathrm{us}}/2k_B T) \approx 1$. Die Dämpfung hat daher bei tiefen Temperaturen ihren größten, nahezu konstanten Wert, d.h. $l^{-1} \propto \omega_{\mathrm{us}} T^0$. Bei hohen Temperaturen ist $\hbar\omega_{\mathrm{us}} \ll k_B T$, und der Hyperbeltangens läßt sich durch sein Argument $(\hbar\omega_{\mathrm{us}}/2k_B T)$ nähern. Der Dämpfungsverlauf zeigt dann eine ausgeprägte Frequenz- und Temperaturabhängigkeit und ist durch $l^{-1} \propto \omega_{\mathrm{us}}^2/T$ gegeben.

Die Schalldämpfung durch Tunnelsysteme besitzt noch eine weitere, sehr charakteristische Eigenschaft. Nach (6.20) ist die Dämpfung proportional zur Besetzungszahldifferenz δn, doch gilt die Beziehung (6.19) für δn nur im thermischen Gleichgewicht. Der Absorptionsprozeß selbst bewirkt eine Umbesetzung: Mit zunehmender Schallintensität J nimmt die Besetzung des oberen Niveaus zu, die Besetzungszahldifferenz δn erringert sich, und damit auch die Dämpfung durch den resonanten Prozeß. Eine experimentelle Bestätigung dieses *Sättigungseffekts* ist in Abb. 6.8 anhand einer Messung an einem Borsilikatglas bei etwa 0,5 K und 1 GHz zu sehen. Der Verlauf der inversen freien Weglänge läßt sich sehr gut mit den vorhandenen, hier allerdings nicht diskutierten Theorien beschreiben.

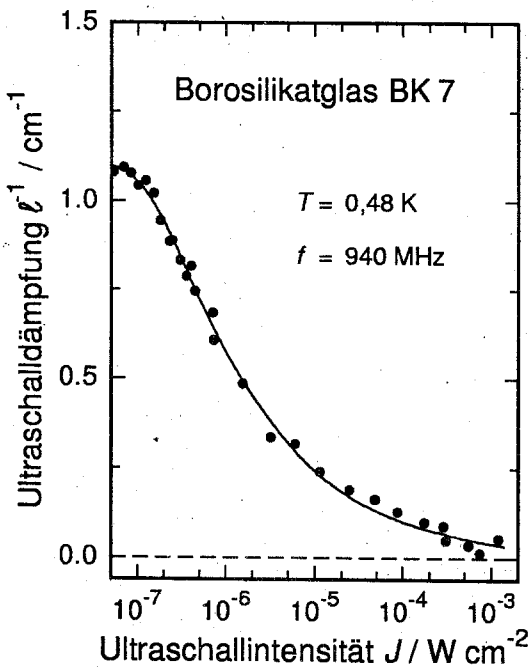


Abb. 6.8: Ultraschalldämpfung in Borsilikatglas als Funktion der eingestrahlten Schallintensität. Die durchgezogene Kurve gibt den theoretischen Verlauf wieder. (Nach S. Hunklinger, Cryogenics **28** (1988) 224).

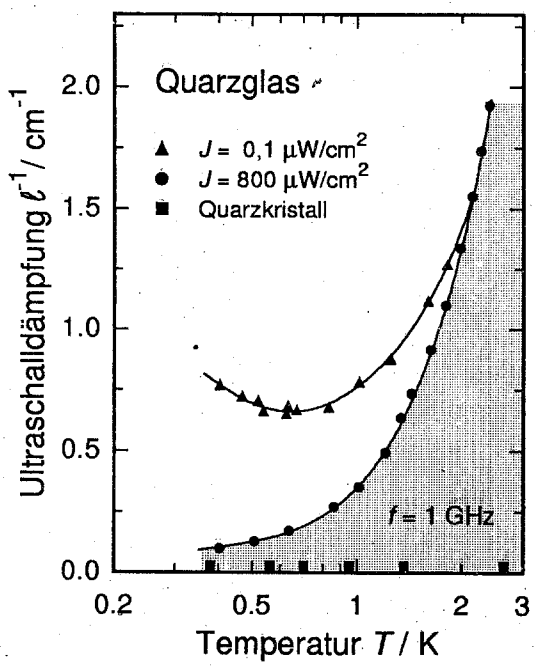


Abb. 6.9: Temperaturabhängigkeit der Ultraschalldämpfung von Quarzglas bei 1 GHz und zwei unterschiedlichen Schallintensitäten J. In kristallinem Quarz ist die Dämpfung in diesem Temperaturbereich verschwindend klein. (Nach S. Hunklinger, Adv. Solid State Physics XVII, J. Treusch (ed.), Vieweg, Braunschweig (1977) 1).

In Abb. 6.9 sind Messungen zur Temperaturabhängigkeit der Ultraschalldämpfung in Quarzglas bei relativ hohen und sehr kleinen Schalleistungen wiedergegeben. Neben der gerade besprochenen resonanten Dämpfung bei Temperaturen unter 1 K findet man zusätzlich noch einen starken Anstieg zu höheren Temperaturen. Die Ursache hierfür ist das Auftreten eines weiteren Dämpfungsmechanismus, der sogenannten *Relaxationsabsorption*. Diese Dämpfung beruht darauf, daß Schallwellen nicht nur mit Tunnelzentren geeigneter Energieaufspaltung resonant wechselwirken, sondern auch die Energieaufspaltung aller Systeme modulieren und somit deren dem thermischen Gleichgewicht entsprechende Besetzung stören. Dieser Mechanismus bewirkt einen starken Anstieg der Absorption oberhalb von 1 K. Da der Beitrag der resonanten Wechselwirkung bei höheren (absolut gesehen jedoch relativ kleinen) Schallintensitäten durch den Sättigungseffekt verschwindet, bleibt bei höheren Intensitäten nur der Relaxationsbeitrag übrig, der in Abb. 6.9 dunkel dargestellt ist. Durch Messung der Dämpfung bei sehr kleiner und bei relativ großer Schallamplitude läßt sich daher der Beitrag der resonanten Dämpfung von dem des Relaxationseffektes trennen. Wir wollen hier aber die Diskussion der Relaxationsdämpfung nicht weiter verfolgen, da wir Relaxationseffekte ausführlicher in Kap. 11.2 im Zusammenhang mit den dielektrischen Eigenschaften der Festkörper besprechen werden.

Ergänzend sei noch bemerkt, daß die resonante Wechselwirkung der Phononen nicht nur bei Gläsern von großer Bedeutung ist, sondern auch in Kristallen mit Punktdefekten beobachtet wird, die dort zu Tunnelzentren Anlaß geben können. Bei der Diskussion der Wärmeleitung werden wir nochmals auf diese Art der Wechselwirkung stoßen.

6.3 Wärmeleitfähigkeit dielektrischer Kristalle

In dielektrischen Kristallen wird die Wärme durch Gitterschwingungen, d.h. durch Phononen transportiert. Ohne Gitteranharmonizitäten oder Defekte würden Phononen ungehindert durch die Probe laufen. Der Wärmewiderstand wäre dann verschwindend klein und die Wärmeleitfähigkeit unendlich groß. Der endliche Wärmewiderstand macht daher deutlich, daß beim Wärmetransport Gitteranharmonizitäten ins Spiel kommen.

Experimente

Die Ausbreitung von thermischen Phononen läßt sich mit verschiedenen experimentellen Techniken untersuchen. In der Prinzipskizze 6.10 sind zwei typische Probengeometrien dargestellt. In klassischen Wärmeleitungsexperimenten benutzt man einen langen dünnen Stab, an dessen freiem Ende eine konstante Heizleistung eingespeist wird. Das Wärmebad nimmt diese Energie auf und legt gleichzeitig die Meßtemperatur T_0 fest. Mit zwei Thermometern wird die Temperaturdifferenz δT bestimmt, die bei derartigen Messungen klein gegen die Meßtemperatur sein muß. Rechts wird der Wärmetransport durch eine dünne Platte gemessen. Der Heizer wird für kurze Zeit angeschaltet und der zeitliche Verlauf der Temperatur auf der anderen Seite mit Hilfe eines Bolometers registriert.

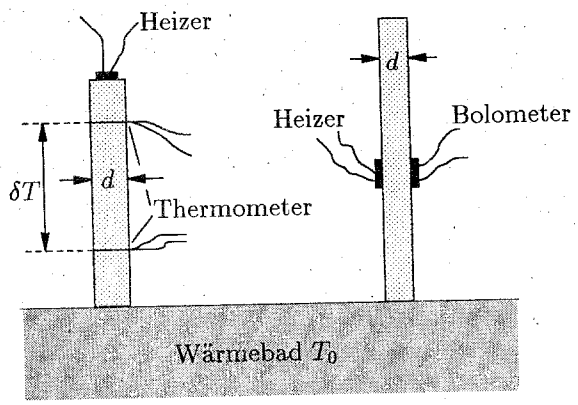


Abb. 6.10: Anordnung zur Messung der Wärmeleitfähigkeit bzw. zur Untersuchung der ballistischen Ausbreitung von Phononen. Links wird der Wärmetransport längs eines langen, dünnen Stabes, rechts der Transport durch eine dünne Platte gemessen.

Wir beschäftigen uns zunächst mit der „klassischen" Anordnung zur Wärmeleitfähigkeitsmessung und gehen auf die Besonderheiten der sogenannten „ballistischen" Phononenausbreitung in einer dünnen Platte etwas später ein.

Laufen die Phononen nicht wie in Ultraschallexperimenten einfach ballistisch, d.h. fast stoßfrei, durch die Probe, sondern diffundieren, so wird der Energietransport durch den Temperaturgradienten bestimmt. Unter stationären Bedingungen gilt dann die Gleichung $\boldsymbol{j} = -\Lambda \operatorname{grad} T$, wobei $\boldsymbol{j}$ die Wärmestromdichte und Λ der Koeffizient der Wärmeleitfähigkeit ist.

Für das Auftreten von Wärmewiderstand bzw. Wärmeleitung im üblichen Sinn müssen zwei wesentliche Bedingungen erfüllt sein:

i) In der Probe muß lokales thermisches Gleichgewicht herrschen, d.h. Abweichungen von der Gleichgewichtsbesetzung der Phononen müssen sich lokal abbauen. Dies geschieht durch inelastische Phononenstöße.

ii) Zusätzlich müssen Stoßprozesse auftreten, die den Gesamtimpuls des Phononengases verändern und somit den Energietransport reduzieren.

In Abb. 6.11 ist der Temperaturverlauf des Wärmeleitfähigkeitskoeffizienten von Natriumfluorid gezeigt, der für kristalline Dielektrika typisch ist. Der berichtete maximale Leitfähigkeitswert von 240 $\mathrm{Wcm^{-1}K^{-1}}$ bei etwa 15 K ist einer der höchsten, die je in einem Festkörper gemessen wurden. Erst in jüngster Zeit wurde er durch den Leitfähigkeitswert von künstlich gezogenen Diamanten „überboten". In weniger reinen Kristallen ist das Leitfähigkeitsmaximum jedoch weniger stark ausgeprägt. Es überrascht ein wenig, daß diese hohe Wärmeleitfähigkeit von Metallen nicht erreicht wird. Zum Vergleich: die Wärmeleitfähigkeit von reinem Kupfer ist bei Zimmertemperatur mit etwa 4 $\mathrm{Wcm^{-1}K^{-1}}$ zwar höher als in Dielektrika, doch erreicht sie im Maximum bei tiefen Temperaturen nur etwa 50 $\mathrm{Wcm^{-1}K^{-1}}$.

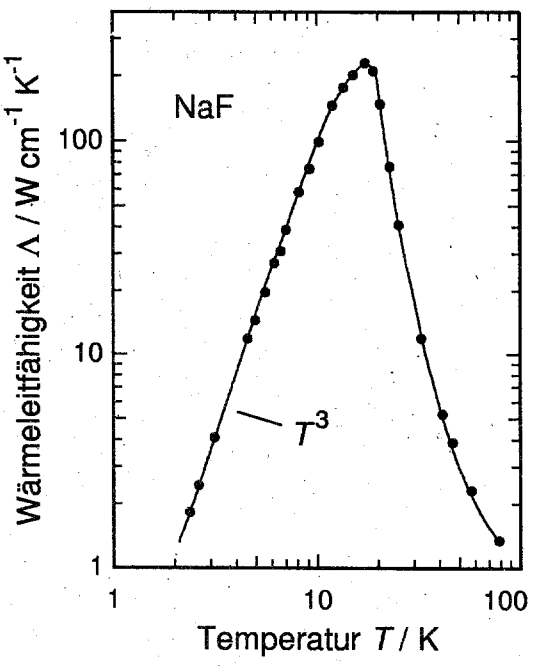


Abb. 6.11: Wärmeleitfähigkeit von hochreinem Natriumfluorid als Funktion der Temperatur. (Nach H.E. Jackson, C.T. Walker, T.F. Nelly, Phys. Rev. Lett. **25** (1970) 26).

Der Einfachheit halber stützen wir uns bei der Beschreibung der Wärmeleitung auf die von Debye erstmals benutzte Analogie zur kinetischen Gastheorie und betrachten die Phononen näherungsweise als ideales Gas. Für den Koeffizienten der Wärmeleitfähigkeit[6] Λ gilt dann:

$$\Lambda = \frac{1}{3} C v l \,. \tag{6.21}$$

C steht für die spezifische Wärme des Phononengases, v für die Schallgeschwindigkeit und l für die mittlere freie Weglänge der Phononen. Im Prinzip müßte berücksichtigt werden, daß alle drei Größen von der Frequenz abhängen und daß deshalb die verschiedenen Phononenzweige unterschiedlich stark zur Wärmeleitung beitragen. Korrekterweise müßte daher über alle Phononenfrequenzen integriert und über alle Phononenzweige j summiert

[6] Im weiteren Verlauf wird der Koeffizienten der Wärmeleitfähigkeit oft nur als *Wärmeleitung* bezeichnet.

werden. Anstelle der einfachen Gleichung (6.21) träte dann

$$\Lambda = \frac{1}{3} \sum_j \int_0^{\omega_{\max}} c_j(\omega)\, v_j(\omega)\, l_j(\omega)\, \mathrm{d}\omega \,, \tag{6.22}$$

wobei $c_j(\omega) = \mathrm{d}C_j/\mathrm{d}\omega$ die „spektrale spezifische Wärme" repräsentiert. Meist vereinfacht man diese Gleichung in zweifacher Hinsicht. Man „erspart" sich die Summation über die Phononenzweige, indem man, dem Debye-Modell folgend, einen „effektiven" Phononenzweig mit linearer Dispersion $v = \omega/q = \text{const.}$ einführt. Die Integration wird durch die Anwendung der „dominanten Phononennäherung" vermieden. Hierbei berücksichtigt man nicht das gesamte Frequenzspektrum, sondern betrachtet nur den Bereich des Phononenspektrums, der am meisten zum Wärmetransport beiträgt. In den meisten Fällen sind es grob gesprochen diejenigen Phononen, deren Energie vergleichbar mit $k_\mathrm{B}T$ ist. Dieses Vorgehen ist schon deswegen zweckmäßig, weil in den wenigsten Fällen der genauere Frequenzverlauf der in (6.22) aufgetretenen Größen bekannt ist.

Die dominante Phononennäherung wird nicht nur zur Beschreibung der Wärmeleitung herangezogen, sondern auch in vielen anderen Fällen verwendet. So analysiert man oft die gemessene spezifische Wärme oder die Ultraschalldämpfung in dieser Näherung. Eine eingehendere Betrachtung zeigt, daß der Zusammenhang $\hbar\overline{\omega} = p\, k_\mathrm{B}T$ zwischen der Frequenz $\overline{\omega}$ der dominanten Phononen und der Temperatur besteht. Der Proportionalitätsfaktor p hängt vom betrachteten Phänomen ab und liegt bei Werten zwischen 1 und 3. Berechnen läßt sich der Wert von p über das erste Moment der betrachteten Größe.

Wie wir gleich sehen werden, treten bei thermischen Phononen in dielektrischen Kristallen zwei wichtige Streumechanismen auf: Entweder wechselwirken die Phononen untereinander, oder sie werden an Defekten gestreut. Sind mehrere Streumechanismen A, B, C... gleichzeitig und unabhängig voneinander wirksam, so addieren sich die Einzelstreuraten bzw. die inversen freien Weglängen. Die effektive inverse freie Weglänge l^{-1} ergibt sich somit zu:

$$l^{-1} = l_\mathrm{A}^{-1} + l_\mathrm{B}^{-1} + l_\mathrm{C}^{-1} + \ldots \tag{6.23}$$

Dies bedeutet, daß immer der Prozeß den Wärmewiderstand bestimmt, der zur stärksten Streuung führt bzw. zur kürzesten freien Weglänge Anlaß gibt.

Phonon-Phonon-Stöße

In Abschnitt 6.2 beschäftigten wir uns bereits mit dem Drei-Phononen-Prozeß, der für die Ultraschalldämpfung verantwortlich ist. Es liegt nahe, diesen inelastischen Prozeß auch für den Wärmewiderstand verantwortlich zu machen. Wie bei der Ultraschalldämpfung muß die Energie- und Quasiimpulserhaltung gelten:

$$\hbar\omega_1 \pm \hbar\omega_2 = \hbar\omega_3 \qquad \text{und} \qquad \hbar\boldsymbol{q}_1 \pm \hbar\boldsymbol{q}_2 = \hbar\boldsymbol{q}_3 + \hbar\boldsymbol{G}\,. \tag{6.24}$$

Je nach Vorzeichen wird beim Stoß ein Phonon erzeugt oder vernichtet. Wie wir in Abschnitt 5.4 bei der Diskussion der inelastischen Streuung von Wellen in Kristallen gesehen haben, kann in der Gleichung für die Quasiimpulserhaltung noch ein reziproker Gittervektor $\boldsymbol{G}$ auftreten. In (6.10) wurde dieser weggelassen, da reziproke Gittervektoren groß im Vergleich zum Wellenvektor q_us der Ultraschallphononen sind. Sie treten daher in der Impulsbilanz nicht auf. Dagegen kann der Wellenvektor von thermischen Phononen durchaus mit dem reziproken Gittervektor vergleichbar sein, so daß nun reziproke

Gittervektoren eine wichtige Rolle spielen. Je nachdem, ob ein reziproker Gittervektor am Streuprozeß beteiligt ist oder nicht, spricht man von *Umklapp-* oder *U*-Prozessen bzw. von *Normal-* oder *N*-Prozessen. In Abb. 6.12 sind diese beiden Prozesse für den Fall der Phononenvernichtung dargestellt. Sind die Wellenvektoren der beteiligten Phononen relativ klein, so läuft der Streuprozeß innerhalb der 1. Brillouin-Zone ab, und die Summe der Quasiimpulse aller beteiligten Phononen bleibt erhalten. Dies gilt auch beim umgekehrten Prozeß, bei dem ein thermisches Phonon in zwei Phononen zerfällt. Endet der resultierende Vektor $\boldsymbol{q}_3$ außerhalb der 1. Brillouin-Zone, so läßt sich durch Addition eines reziproken Gittervektors $\boldsymbol{G}$ der Wellenvektor $\boldsymbol{q}_3'$ finden, der in der 1. Brillouin-Zone liegt. Dabei ändert sich die Summe der Quasiimpulse der beteiligten Phononen. Wie im vorhergehenden Abschnitt erwähnt, existieren für die Phonon-Phonon-Streuung „Auswahlregeln". Sie bewirken, daß im wesentlichen nur Phononen unterschiedlicher Polarisation und damit unterschiedlicher Geschwindigkeit am Streuprozeß beteiligt sind, doch sind die Auswirkungen dieser „Feinheiten" bei der folgenden quantitativen Diskussion ohne Bedeutung.

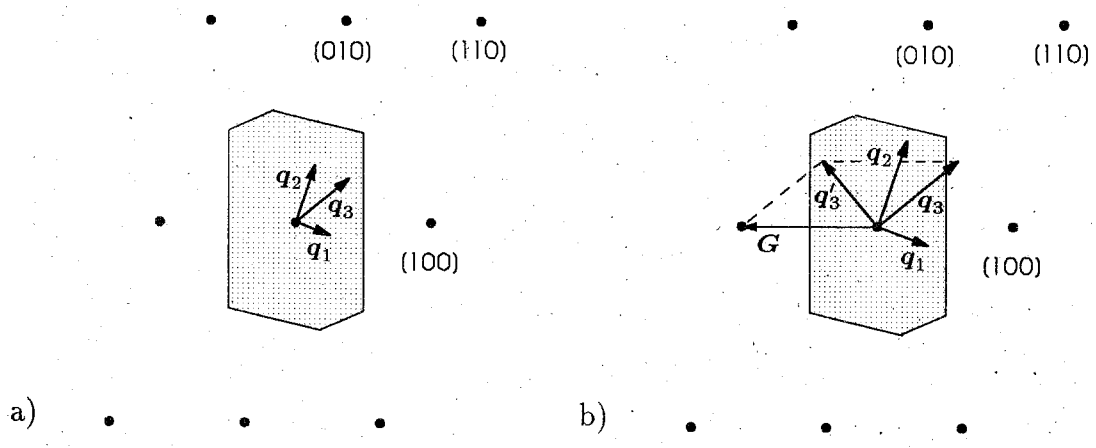


Abb. 6.12: Drei-Phononen-Prozeß.
a) Normal-Prozeß. Alle Wellenvektoren liegen in der dunkel dargestellten 1. Brillouin-Zone.
b) Umklapp-Prozeß. Der Betrag der Wellenvektoren $\boldsymbol{q}_1$ und $\boldsymbol{q}_2$ wurde gegenüber der Abbildung a) um 50 % vergrößert. Dadurch kommt der resultierende Vektor $\boldsymbol{q}_3$ außerhalb der 1. Brillouin-Zone zu liegen. Nun findet ein Umklapp-Prozeß statt.

Normalprozeß

Bei *N*-Prozessen bleibt die Summe der Quasiimpulse der beteiligten Phononen und damit auch der Quasiimpuls $\boldsymbol{P}$ aller Phononen erhalten, d.h.

$$\boldsymbol{P} = \sum_{\boldsymbol{q}} N_{\boldsymbol{q}}\, \hbar \boldsymbol{q} = \text{const.} \tag{6.25}$$

$N_{\boldsymbol{q}}$ steht hierbei für die Zahl der Phononen mit dem Wellenvektor $\boldsymbol{q}$. Da durch diese Stöße weder der Impulsfluß noch der damit verknüpfte Energietransport beeinträchtigt werden, tragen diese Stöße nicht zum Wärmewiderstand bei. An dieser Argumentation ändern auch die bei höheren Temperaturen auftretenden Prozesse höherer Ordnung, z.B. Vier-Phononen-Prozesse, nichts, da auch sie den Gesamtimpuls erhalten. Für die Einstellung des lokalen thermischen Gleichgewichts sind *N*-Prozesse jedoch von Bedeutung, da sich bei diesen Stößen die Frequenzen der stoßenden Phononen ändern.

Umklapp-Prozeß

Ganz anders sieht die Impulsbilanz bei Umklapp-Prozessen aus, obwohl auch daran drei Phononen beteiligt sind. Voraussetzung für ihr Auftreten ist, daß nach dem Stoß mindestens einer der Wellenvektoren außerhalb der 1. Brillouin-Zone zu liegen kommt. Nach der Addition eines reziproken Gittervektors liegt der resultierende Wellenvektor $\boldsymbol{q}_3'$ wieder in der 1. Brillouin-Zone, wobei sich das Vorzeichen der Phasengeschwindigkeit ändert. Jeder derartige Stoß ändert daher den Gesamtquasiimpuls $\boldsymbol{P}$. Man kann den Vorgang auch etwas anders betrachten: Ein Blick auf Abb. 5.3 verdeutlicht, daß mit dem Überschreiten der Grenze der Brillouin-Zone im allgemeinen ein Vorzeichenwechsel der Gruppengeschwindigkeit verbunden ist. Gerade dieser Übergang von einer Zone zur nächsten wird beim Umklapp-Prozeß durch die Addition der beiden Wellenvektoren bewirkt.

Bei hohen Temperaturen dominieren Umklapp-Prozesse, da die Frequenz der überwiegenden Zahl der angeregten Phononen mit der Debye-Frequenz ω_D vergleichbar ist. Praktisch jeder Stoß führt zu einem Endzustand außerhalb der 1. Brillouin-Zone und ist somit ein Umklapp-Prozeß. Wir können also die Überlegungen zur Frequenz- und Temperaturabhängigkeit aus Kap. 6.2 für den Drei-Phononen-Prozeß bei Ultraschallfrequenzen direkt übertragen. Ersetzen wir in (6.17) die Ultraschall- durch die Debye-Frequenz, so ergibt sich unmittelbar $l^{-1} \propto \omega_\mathrm{D} T$. Da bei hohen Temperaturen C_V konstant ist, sollte die Wärmeleitfähigkeit umgekehrt proportional zur Temperatur abnehmen. Die Vorhersage $\Lambda \propto T^{-1}$ stimmt mit Beobachtungen gut überein.

Bei mittlerer Temperatur, d.h. unterhalb der Debye-Temperatur, hängt die Zahl der Phononen, deren Impuls für einen Umklapp-Prozeß ausreicht, stark von der Temperatur ab. Damit der Endzustand eines Streuprozesses außerhalb der 1. Brillouin-Zone liegt, muß in grober Näherung $\hbar\omega > \hbar\omega_\mathrm{D}/2$ sein. Die Wahrscheinlichkeit, derartige Phononen zu finden, folgt aus der Besetzungszahl. Für $\hbar\omega_\mathrm{D} > k_\mathrm{B}T$ gilt näherungsweise: $\langle n(\omega,T)\rangle = [\exp(\hbar\omega_\mathrm{D}/2k_\mathrm{B}T) - 1]^{-1} \approx \exp(-\hbar\omega_\mathrm{D}/2k_\mathrm{B}T) = \exp(-\Theta/2T)$. Da Streuquerschnitt und spezifische Wärme vergleichsweise schwach von der Temperatur abhängen, können wir für die Wärmeleitfähigkeit näherungsweise

$$\Lambda \propto l(T) \propto \mathrm{e}^{\Theta/2T} \tag{6.26}$$

schreiben. Wie man Abb. 6.11 entnehmen kann, steigt unterhalb der halben Debye-Temperatur die Wärmeleitfähigkeit tatsächlich näherungsweise exponentiell mit abnehmender Temperatur an. Wir wollen hier jedoch auf eine quantitative Analyse des Temperaturverlaufs verzichten und uns unmittelbar dem Tieftemperaturverlauf zuwenden.

Streuung an Defekten

Bei tiefen Temperaturen verliert der Umklapp-Prozeß vollständig seine Bedeutung, weil der Impuls aller beteiligten Phononen innerhalb der 1. Brillouin-Zone liegt. Es muß daher einen weiteren Streumechanismus geben, der die Wärmeleitfähigkeit bei tiefen Temperaturen begrenzt, denn diese nimmt bei tiefen Temperaturen sogar wieder ab.

Die Streuung der Phononen an Defekten und an der Oberfläche begrenzt bei tiefen Temperaturen die freie Weglänge. Welche Bedeutung die Oberfläche für den Wärmetransport hat, hängt von der experimentellen Anordnung (vgl. Abb. 6.10) bei der Messung ab. Eine wesentliche Rolle spielt sie beim klassischen Meßverfahren, da dort die Phononen auf ihrem Weg von einem zum anderen Probenende an der Oberfläche wiederholt gestreut werden können. Hierbei ändert sich ihr Gesamtimpuls ähnlich wie bei den Umklapp-Prozessen. In der Gleichung für den Koeffizienten der Wärmeleitfähigkeit

übernimmt der Probendurchmesser d die Funktion der freien Weglänge, d.h. wir setzen in Gleichung (6.21) $l \approx d$. Für die Wärmeleitfähigkeit finden wir dann

$$\Lambda \approx C_V v d \propto T^3 d \,. \tag{6.27}$$

Diesen Temperaturbereich, in dem die Wärmeleitfähigkeit von der Probengeometrie abhängt, bezeichnet man als den *Casimir-Bereich*. Der Einfluß der Probendimension ist in Abb. 6.13 zu sehen, die die Wärmeleitfähigkeit verschiedener LiF-Kristallen mit unterschiedlichem Querschnitt zeigt. Es ist deutlich zu erkennen, daß die Wärmeleitfähigkeit mit dem Probenquerschnitt abnimmt.

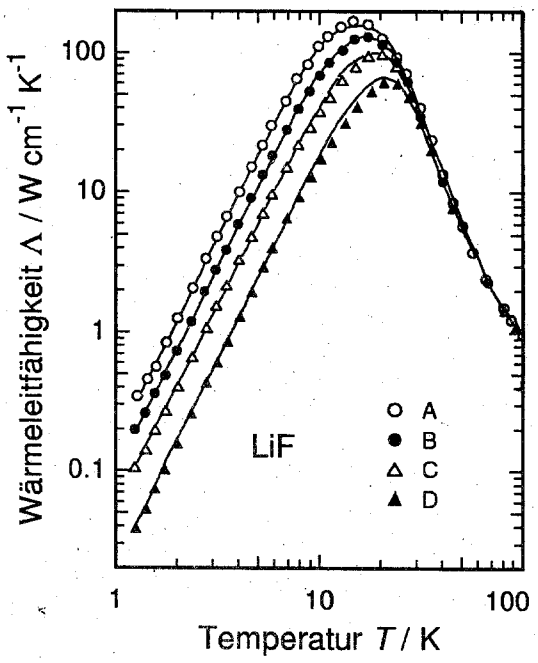


Abb. 6.13: Wärmeleitfähigkeit von LiF-Kristallen mit nahezu quadratischen Querschnitten als Funktion der Temperatur. Die Kantenlänge nimmt von (A) 7,3 mm auf (D) 1,1 mm ab. (Nach P.D. Thacher, Phys. Rev. **156** (1967) 975).

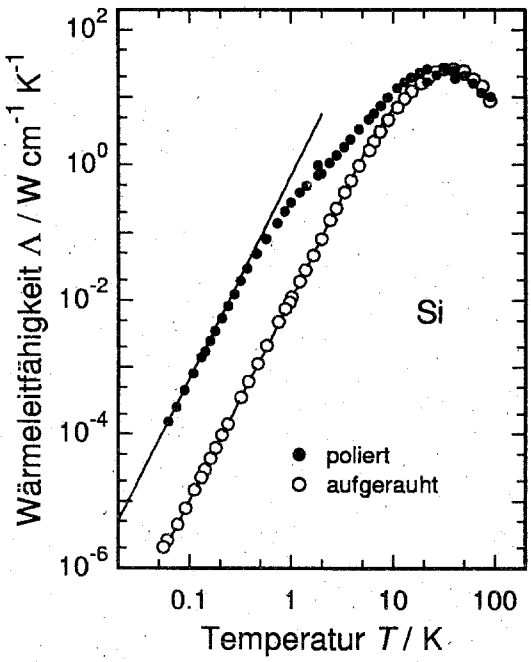


Abb. 6.14: Wärmeleitfähigkeit eines Siliziumkristalls mit sehr gut polierter (•) bzw. aufgerauhter (○) Oberfläche. (Nach V. Röhring, private Mitteilung).

In Wirklichkeit hängt die maximale freie Weglänge auch von der Oberflächenbehandlung, z.B. von der Rauhigkeit der Probe ab. Bei gut polierten Oberflächen tritt spiegelnde Reflexion auf, durch die sich die Komponente des Phononenimpulses parallel zur Oberfläche nicht ändert. Die effektive Weglänge ist daher größer als die Probendicke. Die Kristalldimension gibt deshalb nur die Größenordnung der freien Weglänge an. Der Einfluß der Probenpräparation ist in Abb. 6.14 klar zu erkennen. Die Oberfläche einer Probe wurde durch Sandstrahlen und chemisches Ätzen angerauht, während die andere Probe eine sehr gut polierte Oberfläche aufwies. Die Wärmeleitfähigkeiten unterscheiden sich unter 1 K etwa um den Faktor 500. Die Meßkurve des polierten Kristalls zeigt im Vergleich zu dem des aufgerauhten Kristalls im Temperaturbereich 1 K – 20 K einen veränderten Kurvenverlauf. Unter 1 K ist die effektive freie Weglänge bei der polierten Probe extrem groß: Sie beträgt etwa 7 cm und ist somit durch die *Probenlänge* bestimmt. Dies bedeutet,

daß sehr gut polierte Oberflächen den Wärmetransport kaum behindern. An dieser Stelle soll noch einmal daraufhin gewiesen werden, daß man aus der Wärmeleitfähigkeit nur Information über die freie Weglänge bezüglich quasiimpulszerstörender Prozesse erhält. Normalprozesse haben keinen Einfluß auf die Zahlenwerte der freien Weglängen, die man aus Wärmeleitungsmessungen erhält.

Bei welcher Temperatur beeinflussen die in Kap. 4 diskutierten Punktdefekte oder Versetzungen den Verlauf der Wärmeleitfähigkeit? Es zeigt sich, daß beide Defektarten im allgemeinen erst bei sehr hoher Defektkonzentration merklich zum Wärmewiderstand beitragen. Eine Ausnahme hiervon stellen Kristalle mit Defekten dar, die mit Phononen resonant wechselwirken können. Bereits bei der Diskussion der Ultraschalldämpfung von amorphen Festkörpern haben wir gesehen, daß diese Wechselwirkung die Phononenausbreitung sehr effektiv behindert. Wir werden auf diesen Mechanismus mit Hinblick auf die Wärmeleitung amorpher Festkörper kurz im folgenden Abschnitt eingehen. Jetzt betrachten wir nur den Einfluß einer größeren Konzentration von Punktdefekten, die auf dem Weg der Phononen wie geometrische Hindernisse wirken. Wie bereits bei der Diskussion der Ultraschalldämpfung erwähnt, erfahren Gitterwellen an Punktdefekten ähnlich wie elektromagnetische Wellen an kleinen Streuzentren Rayleigh-Streuung. Nach (6.12) ist die freie Weglänge $l^{-1} \propto \omega^4$. Da entsprechend der dominanten Phononennäherung $\overline{\omega} \propto T$ ist, können wir für die freie Weglänge $l \propto T^{-4}$ schreiben. Punktdefekte streuen also besonders effektiv bei hohen Temperaturen, doch dominiert in diesem Temperaturbereich im allgemeinen die Phonon-Phonon-Wechselwirkung aufgrund der großen Zahl von angeregten Phononen. Dennoch kann man in günstigen Fällen den Einfluß der Streuung an Punktdefekten auf die Wärmeleitung beobachten, nämlich dann, wenn die Umklapp-Prozesse noch relativ selten sind, d.h. bei Temperaturen in der Nähe des Wärmeleitungsmaximums.

Ein besonders einprägsamer Spezialfall der Punktdefektstreuung ist die sogenannte „Isotopenstreuung", deren Auswirkung in Abb. 6.15 zu sehen ist.

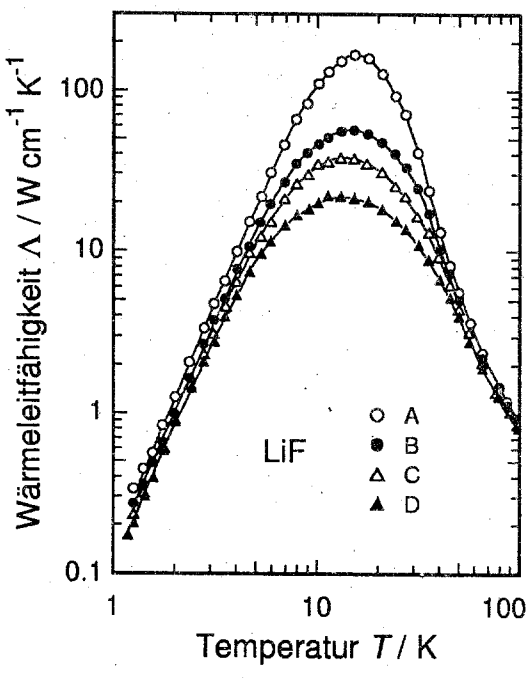


Abb. 6.15: Wärmeleitfähigkeit in LiF mit verschiedenen Anteilen an ^{6}Li und ^{7}Li als Funktion der Temperatur. Die Konzentration des ^{7}Li-Isotops betrug (A) 99,99%, (B) 97%, (C) 93% und (D) 51%. (Nach P.D. Thacher, Phys. Rev. **156** (1967) 975).

Ausgehend von einem nahezu isotopenreinen ^{7}LiF-Kristall, reduziert sich die Wärmeleitfähigkeit durch Beimischung von ^{6}Li im Maximum um etwa einen Faktor zehn! Die kleinen Unterschiede im Casimir-Bereich werden durch die etwas unterschiedlichen Probengeometrien verursacht. Besonders intensiv wurde der Einfluß der Isotopenstreuung auf die Wärmeleitfähigkeit bei künstlichen Diamantschichten studiert. Die bereits erwähnte sehr hohe Leitfähigkeit von Diamant findet man bei Proben, die kaum ^{13}C-Verunreinigungen enthalten.

Experimente bei sehr tiefen Temperaturen

Bei sehr tiefen Temperaturen sterben auch die Normalprozesse aus. Zwischen den Phononen stellt sich dann auch kein lokales thermisches Gleichgewicht mehr ein. Somit ist die Wärmeleitfähigkeit im üblichen Sinn nicht mehr definiert. Bei geringer Defektkonzentration breiten sich in diesem Fall die Phononen ballistisch aus, so daß *Wärmepulsexperimente* durchgeführt werden können. Der schematische Versuchsaufbau für derartige Experimente wurde bereits am Anfang dieses Abschnittes in Abb. 6.10 gezeigt. Im Heizer werden für kurze Zeit Phononen aller Polarisationsrichtungen erzeugt, laufen in den Kristall und werden auf der anderen Seite detektiert. In Abb. 6.16 ist das Ergebnis eines derartigen Experiments an Sauerstoff und Chrom dotierten GaAs-Kristallen dargestellt.

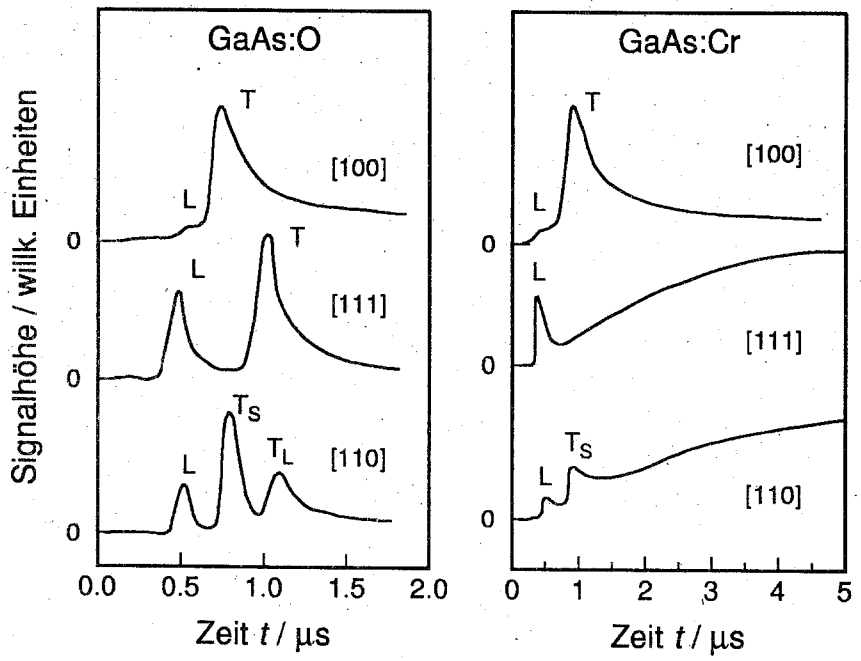


Abb. 6.16: Ausbreitung von ballistischen Phononen in Sauerstoff und Chrom dotierten GaAs-Kristallen. Longitudinale und transversale Phononen lassen sich anhand ihrer Laufzeit unterscheiden. (Nach V. Narayanamurti, *Phonon Scattering in Condensed Matter*, H.J. Maris (ed.) Plenum Press, New York (1980) 385).

Vom Bolometer werden longitudinale und transversale Phononen detektiert, die sich anhand ihrer Ankunftszeit unterscheiden lassen. In [100]- und [111]-Richtung wird nur ein Puls mit transversalen Phononen beobachtet, da die beiden transversalen Zweige T_S und T_L entartet sind. Die beiden Proben sind mit Sauerstoff bzw. Chrom dotiert, d.h. sie

enthalten unterschiedliche Punktdefekte. Deutlich zu erkennen ist, daß die Phononensignale durch die Defekte unterschiedlich beeinflußt werden. Eingehende Untersuchungen haben gezeigt, daß nicht die Rayleigh-Streuung dominiert, sondern daß resonante Wechselwirkung zwischen Phononen und Defekten zur starken Dämpfung Anlaß gibt. Diesen Mechanismus haben wir bereits in Kap. 6.2 in Zusammenhang mit der Ultraschalldämpfung in amorphen Festkörpern diskutiert. Der Absorptionsquerschnitt für diesen Prozeß hängt in Kristallen von der Ausbreitungsrichtung und der Polarisation der Phononen ab. Daher lassen sich aus derartigen Messungen Rückschlüsse auf Auswahlregeln, Kopplung der Defekte an Gitterschwingungen und somit letztendlich auf die Natur der Defekte ziehen. Neben den ballistischen werden noch „diffundierende" Phononen beobachtet. Dies sind Phononen, die von den Chromdefekten wiederholt absorbiert und reemittiert wurden und auf „Umwegen" das Bolometer zu einem späteren Zeitpunkt erreichen.

6.4 Wärmeleitfähigkeit amorpher Festkörper

Der Temperaturverlauf der Wärmeleitfähigkeit von Kristallen unterscheidet sich grundlegend von dem von amorphen Festkörpern. Dies macht der Vergleich zwischen Quarzglas und Quarzkristall in Abb. 6.17 deutlich. Beide Kurvenverläufe sind charakteristisch für diese beiden Substanzklassen. In reinen dielektrischen Kristallen durchläuft die Leitfähigkeit ein Maximum, dessen Ursache wir in Kap. 6.3 diskutiert haben. In amorphen Festkörpern dagegen nimmt das thermische Leitvermögen stetig mit der Temperatur zu, liegt aber bei allen Temperaturen wesentlich unter dem der entsprechenden Kristalle. Wie in Abb. 6.17 angedeutet, lassen sich für Gläser drei Temperaturbereiche unterscheiden, die wir mit Hilfe der dominanten Phononennäherung kurz diskutieren wollen.

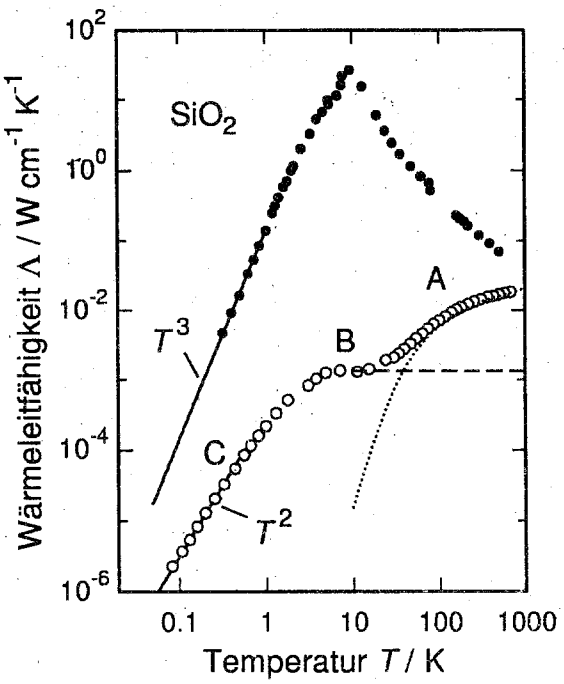


Abb. 6.17: Vergleich der Wärmeleitfähigkeit von Quarzkristall (•) und Quarzglas (○). Bei Quarzglas ist die Unterteilung in drei Temperaturbereiche A, B und C angedeutet. (Nach R.C. Zeller und R.O. Pohl, Phys. Rev. **B4** (1971) 2029).

Bereich hoher Temperaturen:

Bei hohen Temperaturen liegt die Frequenz der dominanten Phononen um 10^{13} Hz und ihre Wellenlänge bei 10 Å. Aus den Wärmeleitungsdaten ergeben sich für die freie Weglänge Werte von etwa 5 Å, d.h. Wellenlänge und freie Weglänge sind vergleichbar. Die Phononen sind, soweit sie in amorphen Festkörpern bei diesen Wellenlängen überhaupt existieren, überdämpft und im Sinne von Elementaranregungen nicht mehr richtig definiert. Die Beschreibung der Wärmeleitung im Phononenbild ist unter diesen Umständen nicht sinnvoll. Besser geeignet ist die Vorstellung, daß die Schwingungsenergie von Atom zu Atom diffundiert, wie man sie in einer alten Arbeit von Einstein aus dem Jahre 1911 findet. Erst in jüngster Zeit wurden Theorien zur Wärmeleitung von amorphen Materialien für den Bereich höherer Temperaturen entwickelt. Gegenwärtig ist jedoch noch nicht klar, welche dieser Theorien den Transportmechanismus korrekt wiedergibt.

Bereich mittlerer Temperaturen (Plateaubereich):

Im Temperaturbereich um 10 K findet man in allen amorphen Festkörpern ein „Plateau". Dieses Phänomen hat jedoch nichts mit dem Maximum in der Wärmeleitfähigkeit der Kristalle zu tun, obwohl die Abbildung 6.17 dies suggerieren könnte. Während die Lage des Maximums bei Kristallen von der Probendimension abhängt, hat die Probengeometrie keinen Einfluß auf die Plateaulage. Das Auftreten einer nahezu temperaturunabhängigen Wärmeleitfähigkeit bedeutet, daß die freie Weglänge der Phononen dort stark temperatur- oder frequenzabhängig sein muß, um den raschen Anstieg der spezifischen Wärme in diesem Temperaturbereich zu kompensieren. Dafür sind verschiedene Erklärungen vorgeschlagen worden: z.B. Streuung an Punktdefekten oder an lokalisierten Schwingungen, oder die Lokalisierung der Phononen selbst. Eine allgemein anerkannte Erklärung gibt es auch hier noch nicht. Es ist noch nicht einmal klar, ob im Bereich A und B der gleiche Transportmechanismus wirksam ist.

Bereich tiefer Temperaturen:

Zwei Eigenschaften charakterisieren das Verhalten unter 1 K: Der Temperaturverlauf der Wärmeleitfähigkeit ist quadratisch und der Absolutwert der Leitfähigkeit verschiedener amorpher Substanzen vergleichbar. Wie wir in Kap. 5.6 gesehen haben, existieren langwellige Phononen auch in amorphen Festkörpern. Wie in kristallinen Dielektrika wird daher auch in amorphen die Wärme bei tiefen Temperaturen von Phononen transportiert. Doch erreicht man den Casimir-Bereich bei den üblichen Probengeometrien selbst bei sehr tiefen Temperaturen nicht. Die Phononenstreuung in diesem Temperaturbereich erfolgt nicht wie bei Kristallen an der Oberfläche, sondern innerhalb der Probe durch die resonante Wechselwirkung zwischen Phononen und Tunnelzentren, die wir in Kap. 6.2 für den Fall der Ultraschallausbreitung diskutiert haben. Dort haben wir für die inverse freie Weglänge den Ausdruck (6.20) erhalten:

$$l^{-1} = \delta n\, \sigma = C' \omega \tanh \frac{\hbar\omega}{2k_{\mathrm{B}}T} \,. \tag{6.28}$$

Wie bereits betont, tritt in amorphen Festkörpern die resonante Wechselwirkung bei allen Phononenfrequenzen auf, da Tunnelzentren mit unterschiedlichen Energieaufspaltungen mit gleicher Wahrscheinlichkeit existieren. Um die Temperaturabhängigkeit der Wärmeleitfähigkeit zu verstehen, betrachten wir auch hier das Verhalten der dominanten Phononen, für die wieder $\hbar\overline{\omega} \propto k_{\mathrm{B}}T$ gilt. Somit ist die Besetzungszahldifferenz $\tanh(\hbar\overline{\omega}/2k_{\mathrm{B}}T)$ der Tunnelzentren, die mit den dominanten Phononen wechselwirken, konstant. Damit

ergibt die Rechnung für die inverse freie Weglänge $l^{-1} \propto \overline{\omega} \propto T$. Da die spezifische Wärme der Phononen entsprechend dem Debye-Modell bei tiefen Temperaturen proportional zu T^3 ist, finden wir schließlich

$$\Lambda \propto T^2 \tag{6.29}$$

in guter Übereinstimmung mit den experimentellen Ergebnissen.

7. Elektronen im Festkörperpotential

In den beiden vorangegangenen Kapiteln haben wir uns mit der Bewegung der Atome um ihre Gleichgewichtslage beschäftigt. Die Elektronen spielten dabei keine Rolle, da sie der Bewegung der Kerne „instantan" folgen können. Nun wollen wir die Elektronenbewegung herausgreifen und die elektronischen Eigenschaften der Festkörper studieren. Bei dieser Diskussion setzen wir wieder die Gültigkeit der adiabatischen Näherung voraus, denn das Gitter bewegt sich im Vergleich zu den Elektronen so langsam, daß Rückwirkungen in guter Näherung vernachlässigt werden können. Für unsere Diskussion halten sich deshalb die Elektronen in einem quasi-starren Gitter auf. Wir greifen ein beliebiges Elektron heraus und nehmen an, daß dieses Elektron ein effektives Potential $\widetilde{V}(\boldsymbol{r})$ sieht, das zeitlich nahezu konstant ist und von allen anderen Elektronen und Atomkernen hervorgerufen wird. Dieses Potential soll bereits die Wechselwirkung zwischen den ausgelenkten Atomrümpfen und die Wechselwirkung zwischen Rumpf- und Valenzelektronen enthalten. Diese Vorgehensweise, bei der nur das Verhalten eines Elektrons untersucht wird, und die restlichen Ladungen nur zum Potential beitragen, in dem sich das herausgegriffene Elektron bewegt, wird als *Einelektron-Näherung* bezeichnet. Die Auswirkungen der Gitterschwingungen werden in dieser Beschreibung nachträglich als *Elektron-Gitter-Wechselwirkung* bei den Transportphänomenen berücksichtigt.

Wir diskutieren die grundlegenden Fragen der Elektronenbewegung in zwei Schritten. Zunächst gehen wir auf einige Eigenschaften der Festkörper ein, die unmittelbar mit der elektronischen Zustandsdichte verknüpft sind. Wir werden sie in diesem Kapitel zusammen mit anderen wichtigen Größen ausführlich diskutieren. In Kap. 8 behandeln wir die Bewegung der Elektronen und die damit verbundenen Transporteigenschaften, z.B. die elektrische oder thermische Leitfähigkeit von Metallen.

7.1 Freies Elektronengas

Ausgangspunkt unserer Betrachtungen sind Metalle, deren elektronische Eigenschaften sich erstaunlich gut auf das Verhalten *freier Elektronen* zurückführen lassen. Auf den ersten Blick scheint es sich hierbei um eine äußerst grobe und zu weit gehende Vereinfachung der tatsächlichen Gegebenheiten zu handeln. Die anschauliche Darstellung des Potentials in Abb. 7.1 macht die Annahme jedoch verständlich. Zwar treten bei einer Bewegung der Elektronen längs einer Geraden, auf der die Ionenrümpfe liegen, tiefe Mulden im Potential auf, doch ist die Modulation des Potentials relativ schwach, wenn der Weg zwischen den Ionenrümpfen liegt. Wir gehen zunächst noch weiter in der Vereinfachung und benutzen die von *Sommerfeld* vorgeschlagene Näherung: Die Elektronen bewegen sich *frei* in einem *konstanten Potential*, sind aber dem Pauli-Prinzip unterworfen. Es stellt sich heraus, daß diese starken Vereinfachungen bei den Alkalimetallen oder den einfachen Metallen wie Kupfer, Silber und Gold weitgehend gerechtfertigt sind. Dort finden sich neben den delokalisierten, weitgehend freien s-Elektronen nur Elektronen in abgeschlossenen Schalen, die für die elektronischen Festkörpereigenschaften keine Rolle spielen. Dagegen ist

die Annahme quasi-freier Elektronen bei Übergangsmetallen nur schlecht erfüllt. Neben den s-Elektronen treten bei diesen Metallen auch noch d- oder f-Elektronen auf, deren Orbitale sich im Festkörper überlappen, wodurch sie die charakteristischen Eigenschaften eines freien Elektronengases weitgehend verlieren.

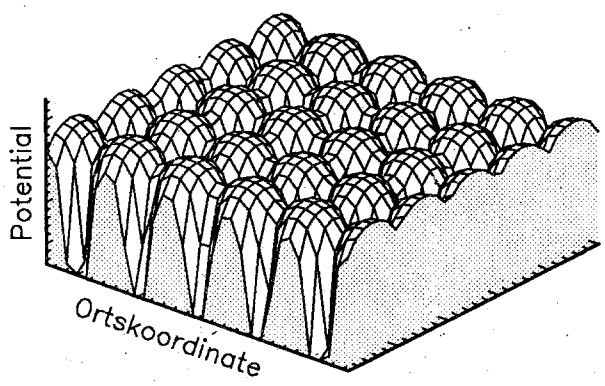


Abb. 7.1: Anschauliche Darstellung des Kristallpotentials. Die positiven Ionenrümpfe sitzen im Zentrum der tiefen Mulden. Die Bildbegrenzung ist so gewählt, daß ein Schnitt entlang von Atomkernen erfolgt, der andere dagegen zwischen den Rümpfen liegt.

Zustandsdichte

Im ersten Schritt wollen wir die Zustandsdichte der freien Elektronen herleiten und damit ihre spezifische Wärme berechnen, deren kleiner Wert lange Zeit nicht erklärt werden konnte. Durch die Annahme eines ortsunabhängigen Potentials wird die Anisotropie der Kristalle vernachlässigt, und wir können beim Abzählen der erlaubten Zustände wie bei den Phononen in elastisch isotropen Medien vorgehen. Wir greifen einen Würfel mit der Kantenlänge L heraus, in dem sich N freie Elektronen befinden. Den Nullpunkt des Potentials legen wir so, daß im Würfel das Potential $\widetilde{V}_0 = 0$ herrscht. Unter diesen Voraussetzungen besitzen die Elektronen nur *kinetische Energie*, und die stationäre Schrödinger-Gleichung für ein Elektron nimmt die einfache Form

$$-\frac{\hbar^2}{2m}\Delta\psi(\boldsymbol{r}) = E\,\psi(\boldsymbol{r}) \tag{7.1}$$

an. Als Lösungen setzen wir für die Wellenfunktionen der Elektronen ebene Wellen der Gestalt

$$\psi(\boldsymbol{r}) = \frac{1}{\sqrt{V}}\,\mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}} \tag{7.2}$$

an, die durch ihren Wellenvektor $\boldsymbol{k}$ charakterisiert sind. Das Würfelvolumen V geht aus Gründen der Normierung in den Ausdruck für die Wellenfunktion ein. Mit diesem Ansatz erhalten wir aus der Schrödinger-Gleichung die Energieeigenwerte freier Elektronen:

$$E = \frac{\hbar^2 k^2}{2m}\,. \tag{7.3}$$

Die Wellenfunktion der Elektronen unterliegt aber Randbedingungen, wodurch die Zahl der erlaubten Wellenvektoren $\boldsymbol{k}$ eingeschränkt wird. Wie bei der Berechnung der Phononenzustandsdichte in Kap. 5.5 benutzen wir die periodische Randbedingungen

$$\psi(x,y,z) = \psi(x+L,y,z) = \psi(x,y+L,z) = \psi(x,y,z+L)\,. \tag{7.4}$$

Für die Komponenten der Wellenvektoren folgt hieraus

$$k_x = \frac{2\pi}{L} m_x, \qquad k_y = \frac{2\pi}{L} m_y, \qquad k_z = \frac{2\pi}{L} m_z \,, \tag{7.5}$$

wobei m_x, m_y und m_z ganze Zahlen sind. Wie bei den Phononen sind auch hier die erlaubten Wellenvektoren gleichmäßig im Impulsraum verteilt und haben die gleiche Dichte $\varrho_k = V/(2\pi)^3$.

Zur Berechnung der Zustandsdichte $\mathcal{D}(E)$ benutzen wir die Gleichungen (5.50) und (5.51) und schreiben

$$\mathcal{D}(E)\,\mathrm{d}E = \varrho_k \int\limits_{E(k)}^{E+\mathrm{d}E} \mathrm{d}^3k = \frac{\varrho_k}{\hbar}\,\mathrm{d}E \int \frac{\mathrm{d}S_E}{v_\mathrm{g}} \,. \tag{7.6}$$

Die Gruppengeschwindigkeit v_g der Elektronen hängt beim freien Elektronengas nicht von der Richtung ab, und es gilt $v_\mathrm{g} = \partial E/\partial(\hbar k) = \hbar k/m$. Damit ergibt das Oberflächenintegral eine Kugelfläche im k-Raum, d.h. es gilt $\int \mathrm{d}S_E = 4\pi k^2$. Somit folgt für die Zustandsdichte $\mathcal{D}(E)$ die Beziehung

$$\mathcal{D}(E) = \frac{V}{(2\pi)^3\hbar}\,\frac{m}{\hbar k}\,4\pi k^2 = \frac{(2m)^{3/2}\sqrt{E}}{4\pi^2\hbar^3}\,V \,. \tag{7.7}$$

Entsprechend dem Pauli-Prinzip kann jeder Zustand mit zwei Elektronen unterschiedlicher Spinrichtung besetzt werden. Deshalb erhalten wir schließlich für die Zustandsdichte pro Volumen $D(E) = 2\mathcal{D}(E)/V$ das wichtige Endergebnis

$$D(E) = \frac{(2m)^{3/2}\sqrt{E}}{2\pi^2\hbar^3} \,. \tag{7.8}$$

Obwohl die Dichte der Zustände im reziproken Raum für Elektronen und Phononen gleich ist, besitzt die Zustandsdichte der Elektronen, bedingt durch die unterschiedliche Dispersionsrelation, eine andere Energieabhängigkeit als die der Phononen. Der Absolutwert der Zustandsdichte wird beim freien Elektronengas nur durch die Masse der Elektronen bestimmt.

Wir wollen hier auch noch die Zustandsdichte eines zweidimensionalen Elektronengases herleiten, da wir auf dieses in Kap. 8.4 noch ausführlich eingehen werden. In Kap. 5.5 wurde bereits erwähnt, daß die Zustandsdichte im Impulsraum von der Dimensionalität des betrachteten Systems abhängt. Demnach gilt $\varrho_k = (L/2\pi)^d$, wobei L für die charakteristische Länge des Systems und d für die Dimensionalität steht. Die Integration über die Oberfläche reduziert sich bei zweidimensionalen Systemen auf ein Linienintegral, für das wir im k-Raum den Kreisumfang $2\pi k$ erhalten. Für die Zustandsdichte $\mathcal{D}_2(E)$ eines zweidimensionalen Elektronengases gilt daher:

$$\mathcal{D}_2(E) = \left(\frac{L}{2\pi}\right)^2 \frac{m}{\hbar^2 k}\,2\pi k = \frac{L^2 m}{2\pi\hbar^2} \,. \tag{7.9}$$

Daraus folgt nach Berücksichtigung der erlaubten Besetzung der Zustände durch zwei Elektronen für die Zustandsdichte $D_2(E)$ pro Fläche der Ausdruck

$$D_2(E) = \frac{m}{\pi\hbar^2} \,. \tag{7.10}$$

Bemerkenswerterweise ist die Zustandsdichte eines zweidimensionalen Elektronengases energieunabhängig, also konstant!

Fermi-Energie, Fermi-Kugel

In einem Ensemble von Teilchen mit halbzahligem Spin, d.h. bei Fermionen, wird das Pauli-Prinzip wirksam. Die Wahrscheinlichkeit, mit der ein quantenmechanischer Zustand im thermischen Gleichgewicht besetzt ist, drückt die Fermi-Dirac-Funktion

$$f(E) = \frac{1}{\mathrm{e}^{(E-\mu)/k_\mathrm{B}T} + 1} \tag{7.11}$$

aus[1]. Die Fermi-Dirac-Verteilungsfunktion geht in die klassische Boltzmann-Verteilung über, wenn die Wahrscheinlichkeit für die Besetzung der betrachteten Zustände sehr viel kleiner als Eins ist. Dieser Fall tritt ein, wenn $\exp[(E-\mu)/k_\mathrm{B}T] \gg 1$ ist.

Das *chemische Potential* μ ist eine thermodynamische Größe, die den Zusammenhang zwischen der freien Energie F und der Teilchenzahl N herstellt und durch

$$\mu = \left(\frac{\partial F}{\partial N}\right)_{T,V} \tag{7.12}$$

definiert wird. Am absoluten Nullpunkt nimmt die Fermi-Dirac-Verteilung die Werte

$$f(E, T=0) = \begin{cases} 1 & \text{für} \quad E < \mu \\ \frac{1}{2} & \text{für} \quad E = \mu \\ 0 & \text{für} \quad E > \mu \end{cases} \tag{7.13}$$

an. Bei $T = 0$ sind also alle Zustände mit $E < \mu$ besetzt, wobei entsprechend dem Pauli-Prinzip jeweils nur zwei Elektronen mit unterschiedlicher Spinrichtung pro Zustand erlaubt sind. Die Energie, bis zu der die Zustände lückenlos gefüllt sind, bezeichnet man als *Fermi-Energie* E_F. Da das chemische Potential die kleinste Energie angibt, die aufzuwenden ist, um ein zusätzliches Elektron in das Fermi-Gas einzubringen, und dies bei $T = 0$ nur bei der Fermi-Energie geschehen kann, sind chemisches Potential und Fermi-Energie E_F bei dieser Temperatur identisch, d.h. $E_\mathrm{F} = \mu\,(T = 0)$.

Die Fermi-Energie ist durch die Elektronendichte $n = N/V$ festgelegt, denn integriert man über alle besetzten Zustände, so erhält man gerade die Teilchenzahl, d.h. es muß

$$n = \frac{N}{V} = \int_0^\infty D(E) f(E, T=0)\,\mathrm{d}E = \int_0^{E_\mathrm{F}} D(E)\,\mathrm{d}E = \frac{(2m)^{3/2}}{2\pi^2\hbar^3}\,\frac{2E_\mathrm{F}^{3/2}}{3} \tag{7.14}$$

gelten. Wenn man die Gleichung nach E_F aufgelöst, sieht man, daß die Fermi-Energie nur durch die Masse und Konzentration der Elektronen bestimmt ist:

$$E_\mathrm{F} = \frac{\hbar^2}{2m}(3\pi^2 n)^{2/3}\,. \tag{7.15}$$

[1] Diese Verteilungsfunktion bestimmt auch die thermische Besetzung von Zwei-Niveau-Systemen, die wir bei der Diskussion der thermischen Eigenschaften amorpher Festkörper in Abschnitt 5.7 kennengelernt haben. Es besteht jedoch nur eine formale Analogie: Durch die beiden möglichen Spineinstellungen hat jedes Elektron zwei erlaubte Zustände, die jeweils nur einmal besetzt werden können.

In Abb. 7.2 ist der parabelförmige Verlauf der Zustandsdichte dargestellt. Die Zustände sind am absoluten Nullpunkt bis zur Fermi-Energie E_F mit Elektronen besetzt, die darüberliegenden Zustände sind leer.

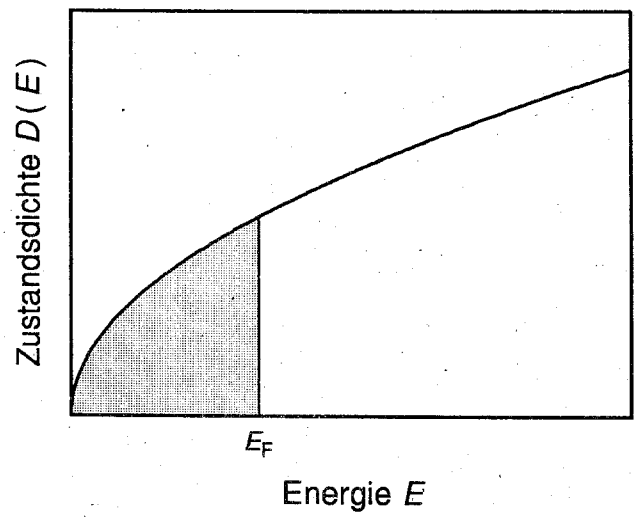


Abb. 7.2: Zustandsdichte des freien Elektronengases. Bei $T = 0$ sind alle Zustände bis zur Fermi-Energie E_F besetzt, oberhalb E_F sind die Zustände leer.

In Verbindung mit der Fermi-Energie lassen sich der *Fermi-Wellenvektor* $\boldsymbol{k}_F$ über den *Fermi-Impuls* $\hbar\boldsymbol{k}_F$ der Elektronen, die *Fermi-Geschwindigkeit* $\boldsymbol{v}_F$, und die *Fermi-Temperatur* T_F definieren. Eine weitere häufig auftretende Größe ist die Zustandsdichte $D(E_F)$ an der „Fermi-Kante". Für den Betrag der jeweiligen Größe gilt:

$$k_F = (3\pi^2 n)^{1/3} , \qquad v_F = \frac{\hbar}{m}(3\pi^2 n)^{1/3} , \tag{7.16}$$

und weiter:

$$T_F = \frac{E_F}{k_B} , \qquad D(E_F) = \frac{3}{2}\frac{n}{E_F} . \tag{7.17}$$

Tabelle 7.1: Elektronendichte, Betrag des Fermi-Vektors, Fermi-Geschwindigkeit, Fermi-Energie und Fermi-Temperatur von ausgewählten Metallen. Bei der Berechnung der einzelnen Größen wurde die Elektronenmasse und die Elektronenkonzentration als bekannt vorausgesetzt. (Aus R.W.G. Wyckoff, Crystal Structures, Interscience, New York, 1966).

	$n\,(10^{28}\,\mathrm{m}^{-3})$	$k_F\,(\mathrm{Å}^{-1})$	$v_F\,(10^6\,\mathrm{m/s})$	$E_F\,(\mathrm{eV})$	$T_F\,(\mathrm{K})$
Li	4,7	1,1	1,3	4,7	55 000
Na	2,7	0,9	1,0	3,2	38 000
Al	1,8	1,8	2,0	11,7	136 000
Cu	8,5	1,4	1,6	7,0	82 000
Ag	5,9	1,2	1,4	5,5	64 000

In allen Fällen ist die Fermi-Temperatur T_F sehr viel größer als die Schmelztemperatur des betreffenden Metalls. Bei allen zugänglichen Temperaturen benehmen sich deshalb die Elektronen so, als befände sich das Metall bei einer Temperatur in der Nähe des absoluten Nullpunktes!

Bei der Temperatur $T = 0$ ist die Besetzungswahrscheinlichkeit durch eine Kastenfunktion gegeben. Bei endlicher Temperatur werden unterhalb der Fermi-Energie E_F Zustände frei und oberhalb dieser Energie Zustände besetzt, die am absoluten Nullpunkt unbesetzt waren. Dies bedeutet, daß die Zustände entsprechend Gleichung (7.11) umbesetzt werden und eine Aufweichung der Fermi-Kante um etwa $2k_BT$ erfolgt. Somit wird von allen Elektronen nur ein Bruchteil der Größenordnung T/T_F thermisch angeregt. Alle anderen Elektronen bleiben nach wie vor im Zustand mit der geringsten Energie und spielen für die überwiegende Zahl von Festkörperphänomenen keine Rolle. In Abb. 7.3 ist der Verlauf der Fermi-Dirac-Funktion für die Temperaturen $T = 0\,\mathrm{K}$ bzw. $3\,000\,\mathrm{K}$ bei einer vorgegebenen Fermi-Temperatur von $T_F = 50\,000\,\mathrm{K}$ veranschaulicht[2].

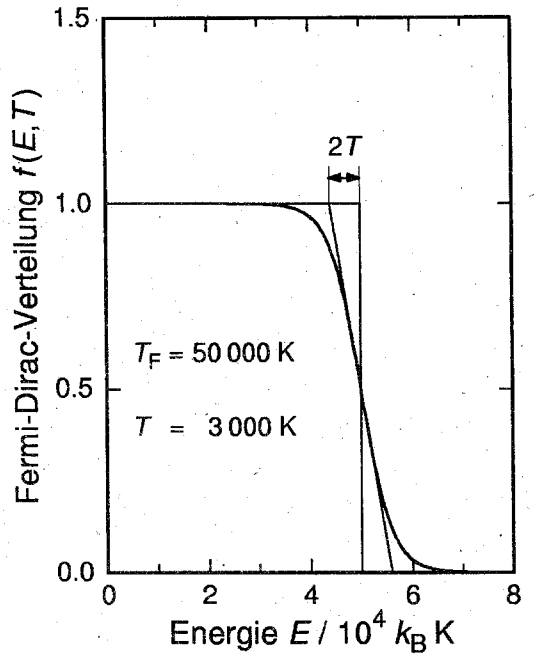


Abb. 7.3: Fermi-Dirac-Verteilung als Funktion der Energie bei den Temperaturen $T = 0\,\mathrm{K}$ und $T = 3\,000\,\mathrm{K}$. In diesem Beispiel wurde als Fermi-Temperatur $T_F = 50\,000\,\mathrm{K}$ angenommen.

Der Wert des chemischen Potentials ist durch die Bedingung $f(E,T) = \frac{1}{2}$ festgelegt und nimmt mit steigender Temperatur leicht ab. Solange $T \ll T_F$ ist, läßt sich die Temperaturabhängigkeit näherungsweise durch

$$\mu(T) \approx E_F \left[1 - \frac{\pi^2}{12}\left(\frac{T}{T_F}\right)^2\right] \tag{7.18}$$

ausdrücken. Wie man anhand dieser Gleichung und der Tabellenwerte für T_F sofort erkennt, ist in Metallen die Verschiebung des chemischen Potentials für alle experimentell zugänglichen Temperaturen unerheblich. In Kap. 9 werden wir finden, daß aber bei Halbleitern die Verschiebung merkliche Ausmaße annehmen kann.

In unserem einfachen Modell freier Elektronen bewegen sich die Elektronen in einer isotropen Umgebung, d.h. der Betrag des Fermi-Wellenvektors $\boldsymbol{k}_F$ besitzt einen festen richtungsunabhängigen Wert. Die Elektronen sind daher bei $T = 0$ im Impulsraum innerhalb einer Kugel, der sogenannten *Fermi-Kugel*, lokalisiert, deren Radius durch die Dichte

[2] Es wurden so hohe Temperaturen gewählt, weil sich bei einer Darstellung der Verhältnisse bei Zimmertemperatur die resultierende Kurve kaum von der bei $T = 0$ unterscheiden ließe.

der Elektronen bestimmt wird. Die Oberfläche dieser Kugel, die in der weiteren Diskussion der elektronischen Eigenschaften von Metallen eine zentrale Rolle spielt, bezeichnet man als *Fermi-Fläche*. Bei endlichen Temperaturen ist die Oberfläche der Fermi-Kugel „aufgeweicht".

7.2 Metalleigenschaften im Modell des freien Fermi-Gases

Spezifische Wärme

Wir berechnen nun die innere Energie des Fermi-Gases, um die spezifische Wärme der Metallelektronen herzuleiten. Sehen wir uns zunächst die innere Energie u_0 pro Volumen am absoluten Nullpunkt an, so finden wir den Wert

$$u_0 = \int_0^\infty E\,D(E)\,f(E,T=0)\,\mathrm{d}E = \int_0^{E_\mathrm{F}} E\,D(E)\,\mathrm{d}E = \frac{3n}{5}\,E_\mathrm{F} = \frac{3n}{5}k_\mathrm{B}T_\mathrm{F}\,. \tag{7.19}$$

Wegen der hohen Fermi-Temperatur ist selbst bei $T = 0$ die innere Energie der Elektronen sehr viel größer als die eines klassischen Gases bei Zimmertemperatur!

Für die spezifische Wärme entscheidend ist jedoch nicht der Absolutwert der inneren Energie, sondern der temperaturabhängige Anteil $\delta u(T) = u(T) - u_0$, für den wir zunächst eine Abschätzung vornehmen wollen. Wie wir Abb. 7.3 entnehmen, ist der Bruchteil der Elektronen, der thermische Energie aufnehmen kann, grob durch T/T_F gegeben. Der Energiezuwachs pro Elektron ist dabei von der Größenordnung $k_\mathrm{B}T$, und somit ist $\delta u \approx nk_\mathrm{B}T^2/T_\mathrm{F}$. Damit erhalten wir für den Beitrag der Elektronen C_V^el zur spezifischen Wärme (pro Volumen) näherungsweise

$$C_V^\mathrm{el} = \left(\frac{\partial u}{\partial T}\right)_V \approx \frac{2nk_\mathrm{B}T}{T_\mathrm{F}}\,. \tag{7.20}$$

Verglichen mit einem klassischen Gas tritt, bedingt durch den kleinen Bruchteil der beteiligten Elektronen, eine drastische Reduktion der spezifischen Wärme um etwa den Faktor T/T_F auf. Bei einer genaueren Rechnung müßte wir das Fermi-Dirac-Integral

$$u = \int_0^\infty E\,D(E)\,f(E,T)\,\mathrm{d}E = \frac{1}{2\pi^2}\left(\frac{2m}{\hbar^2}\right)^{3/2}\int_0^\infty \frac{E^{3/2}}{\mathrm{e}^{(E-\mu)/k_\mathrm{B}T}+1}\,\mathrm{d}E \tag{7.21}$$

auswerten, das jedoch analytisch nicht lösbar ist. Eine Näherungslösung für $k_\mathrm{B}T \ll E_\mathrm{F}$ führt zu

$$u \approx u_0 + \frac{\pi^2}{6}D(E_\mathrm{F})(k_\mathrm{B}T)^2 \tag{7.22}$$

und damit zu

$$C_V^\mathrm{el} \approx \frac{\pi^2}{3}D(E_\mathrm{F})\,k_\mathrm{B}^2T = \frac{mk_\mathrm{B}^2T}{\hbar^2}\left(\frac{\pi^2 n}{9}\right)^{1/3} = \frac{\pi^2}{3}\frac{T}{T_\mathrm{F}}\,\frac{3nk_\mathrm{B}}{2} = \gamma T\,. \tag{7.23}$$

Dieses Ergebnis unterscheidet sich nicht wesentlich von unserer Abschätzung (7.20). Für alle zugänglichen Temperaturen steigt die spezifische Wärme proportional zur Temperatur

an. Der Faktor $\pi^2 T/3T_F$ gibt die Reduktion gegenüber der spezifischen Wärme eines klassischen Gases an. Der Proportionalitätsfaktor γ wird durch die Elektronendichte an der Fermi-Kante und somit durch die Konzentration und die Masse der Elektronen bestimmt[3].

Die *gesamte spezifische Wärme* eines Metalls setzt sich aus den Beiträgen der Elektronen und des Gitters zusammen und wird bei hohen bzw. tiefen Temperaturen näherungsweise durch

$$C_V = \gamma T + \begin{cases} 3n_A k_B & \text{für } T > \Theta \\ \beta T^3 & \text{für } T \ll \Theta \end{cases} \tag{7.24}$$

beschrieben. Bei *hohen Temperaturen* ($T > \Theta$) dominiert der Gitterbeitrag, der durch das Dulong-Petit-Gesetz genähert werden kann. Um Verwechslungen zu vermeiden, ist die atomare Teilchenzahldichte mit n_A bezeichnet. Die Elektronen liefern in diesem Temperaturbereich keinen wesentlichen Beitrag zur Wärmekapazität. Bei *tiefen Temperaturen* trägt das Gitter den Term βT^3 zur spezifischen Wärme bei, wobei die Konstante β durch Gleichung (5.63) gegeben ist, wenn wir dort die Zahl N der Atome durch die atomare Teilchezahldichte n_A ersetzen. Aufgrund der schwächeren Temperaturabhängigkeit überwiegt bei sehr tiefen Temperaturen immer der Beitrag der Elektronen.

Die Aufspaltung in die beiden Anteile ist in Abb. 7.4 für die Tieftemperaturdaten von Kupfer durchgeführt. Offensichtlich läßt sich die gemessene spezifische Wärme sehr gut als Summe der Beiträge von Elektronen und Gitter darstellen, die bei 4 K ungefähr gleich groß sind. Oberhalb dieser Temperatur dominieren die Phononen, deren Zahl wesentlich rascher mit der Temperatur anwächst als die der angeregten Elektronen.

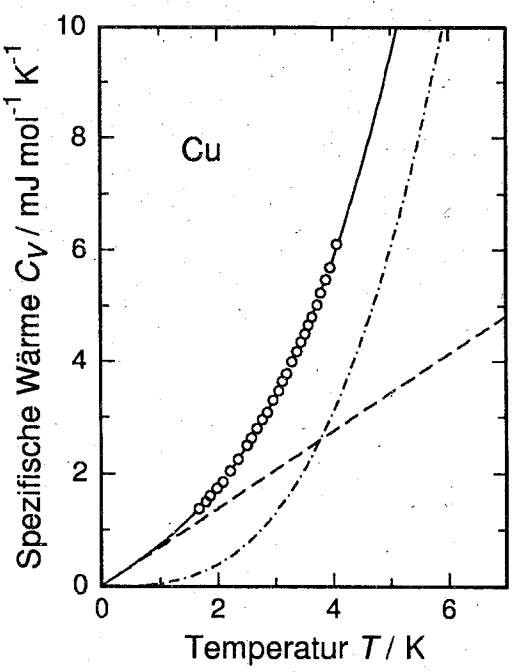


Abb. 7.4: Typischer Temperaturverlauf der spezifischen Wärme eines Metalls (hier von Kupfer) bei tiefer Temperatur. Die gemessene spezifische Wärme (Kurve mit Meßpunkten) setzt sich aus dem linearen Anteil der Elektronen und dem T^3-Beitrag der Phononen zusammen. (Nach J.A. Rayne, Austral. J. Phys. **9** (1956) 189).

[3] Es ist üblich, den Vorfaktor bei der spezifischen Wärme der Elektronen mit γ zu bezeichnen. Er sollte jedoch nicht mit dem Grüneisen-Parameter verwechselt werden.

Um die beiden Beiträge zu trennen, trägt man am besten C_V/T als Funktion von T^2 auf. Der Wert von γ läßt sich dann der Zeichnung unmittelbar als Achsenabschnitt entnehmen, aus der Steigung ergibt sich der Wert von β. Diese Art der Auftragung ist in Abb. 7.5 für die in Abb. 7.4 dargestellten Daten von Kupfer zu sehen und macht noch einmal die gute Übereinstimmung mit (7.24) deutlich.

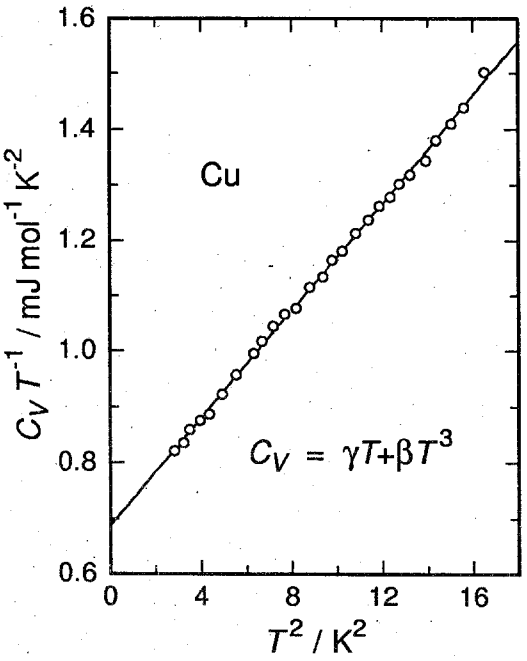


Abb. 7.5: Spezifische Wärme C_V/T von Kupfer als Funktion von T^2. Man beachte, daß der Nullpunkt der Ordinate unterdrückt ist. (Nach J.A. Rayne, Austral. J. Phys. 9 (1956) 189).

Es stellt sich nun die wichtige Frage: Wie gut ist die Näherung des freien Fermi-Gases für die Beschreibung der spezifischen Wärme? Die Werte in Tabelle 7.2 zeigen, daß bei einfachen Metallen die Übereinstimmung relativ gut ist, denn das Verhältnis vom experimentellen zum theoretischen Wert $\gamma_{\text{exp}}/\gamma_{\text{theo}}$ liegt nahe bei Eins.

Tabelle 7.2: Vergleich der experimentellen Werte der spezifischen Wärme (pro Mol) einiger Metalle mit den Werten aus dem Modell freier Elektronen.

	Li	Na	Al	Cu	Ag
γ_{exp} (mJ mol^{-1}K^{-2})	1,63	1,38	1,35	0,70	0,65
$\gamma_{\text{exp}}/\gamma_{\text{theo}}$	2,18	1,26	1,48	1,38	1,00

Aber warum ist die Näherung bei den verschiedenen Metallen unterschiedlich gut? Der berechnete Wert von γ ist proportional zur Masse und zur Konzentration der Elektronen. Da beide Größen genau bestimmt werden können, müssen die Ursachen für die unterschiedlich gute Übereinstimmung tiefer liegen. Der Hauptgrund für die Abweichung beruht darin, daß die Elektronen als freie Teilchen behandelt wurden; in Wirklichkeit ändert aber die Wechselwirkung mit dem Gitter und den übrigen Elektronen ihr Verhalten. Dies wird formal am einfachsten durch die Einführung einer pauschalen *effektiven*

Masse m^* berücksichtigt[4]. Das Konzept der effektiven Masse wird in Kap. 8 eingehend diskutiert. Im vorliegenden Fall spricht man von der effektiven *thermischen Masse* $m^*_{\rm th}$, in der die Wechselwirkung der Elektronen mit dem Gitter und ihre Wechselwirkung untereinander „versteckt" wird. Definiert ist diese Größe durch den einfachen Zusammenhang $m^*_{\rm th}/m = \gamma_{\rm exp}/\gamma_{\rm theo}$.

Bei den Übergangsmetallen ist das Verhältnis $m^*_{\rm th}/m$ relativ groß. So findet man bei Nickel $m^*_{\rm th}/m \approx 15$. Die Ursache hierfür ist jedoch nicht allein in der Wechselwirkung der Elektronen zu suchen, sondern es liegt an den teilweise gefüllten d-Schalen, die bei diesen Metallen neben den s-Elektronen vorhanden sind. Für sie trifft die Näherung des freien Elektronengases in keiner Weise zu, da d-Wellenfunktionen Vorzugsrichtungen besitzen, an kovalenten Bindungen beteiligt sind und den Kristall nicht isotrop ausfüllen. Diese Elektronen bewirken eine hohe Zustandsdichte bei der Fermi-Energie und tragen besonders stark zur spezifischen Wärme bei. Die errechnete und an experimentelle Daten angepaßte Zustandsdichte $D(E)$ der Leitungselektronen von Nickel in Abb. 7.6 macht dies deutlich. An der Fermi-Kante dominieren die d-Elektronen, während in unsere Näherung nur die s-Elektronen eingehen. Auf die etwas komplizierten Verhältnisse, die in vielen Metallen herrschen, werden wir bei der Diskussion der „Bandstruktur" in diesem und den folgenden Kapiteln noch etwas ausführlicher eingehen[5].

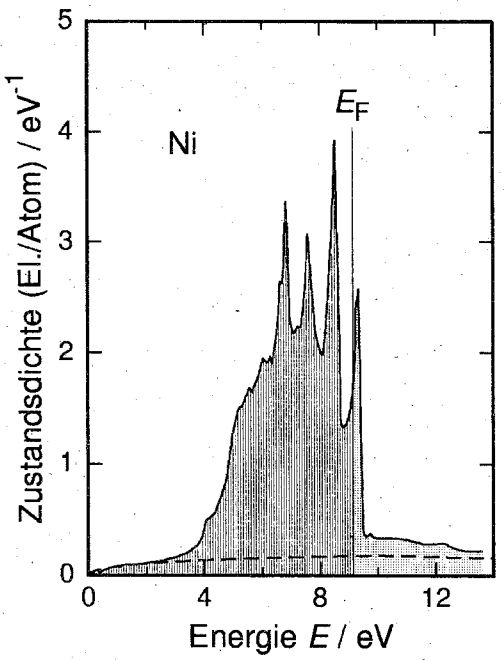


Abb. 7.6: Elektronische Zustandsdichte von Nickel. Der Beitrag der d-Elektronen überlagert das s-Band, das gestrichelt gezeichnet ist. Die besetzten Zustände sind durch dunkle Tönung hervorgehoben. (Nach J. Callaway und C.S. Wang, Phys. Rev. B 7 (1973) 1096).

[4] Für das scheinbar willkürliche Einführen einer neuen, der Vorstellung nicht unmittelbar zugänglichen Größe, gibt es durchaus gute theoretische Gründe. Es handelt sich hierbei um ein häufig benutztes Konzept, bei dem wechselwirkende Teilchen mit Hilfe von freien Teilchen beschrieben werden, deren Eigenschaften jedoch durch die Wechselwirkung modifiziert sind. Die neuen Teilchen, die im Experiment beobachtet werden, bezeichnet man als „Quasiteilchen".

[5] Besonders interessante Eigenschaften weisen die „Schweren Fermion-Systeme" wie z.B. $CeAl_3$ auf. In diesen Metallen bewirkt die starke Wechselwirkung der Elektronen untereinander bei tiefen Temperaturen eine spezifische Wärme, die so groß ist, daß sie mit einer thermischen Elektronenmasse bis zu $m^*_{\rm th} \approx 1000\, m$ beschrieben werden muß. Eine Diskussion dieser eigenartigen Metalle würde jedoch den hier vorgegebenen Rahmen sprengen.

Paulische Spinsuszeptibilität

Ein unlösbares Problem der „klassischen" Theorie der Metalle der Jahrhundertwende war der *Paramagnetismus* der einfachen Metalle. Erst die Fermi-Statistik erlaubt zusammen mit der Kenntnis des magnetischen Moments der Elektronen eine Erklärung der temperaturunabhängigen magnetischen Suszeptibilität. Bringt man ein Elektron in ein Magnetfeld $\mathcal{B}$, so spalten die beiden entarteten Spinzustände auf. Es entsteht ein Zwei-Niveau-System mit einer Energieaufspaltung $\delta E = g_0 \mu_B \mathcal{B}$, wobei μ_B für das Bohrsche Magneton steht; für den elektronischen g-Faktor werden wir im weiteren Verlauf den Wert $g_0 = 2$ benutzen. Im thermischen Gleichgewicht ist die Differenz $\delta n'$ in der Besetzung der beiden Niveaus durch Gleichung (6.19)

$$\delta n' = n \tanh \frac{\delta E}{2k_B T} = n \tanh \frac{\mu_B \mathcal{B}}{k_B T} \tag{7.25}$$

gegeben. Wäre das Pauli-Prinzip unwirksam, so wäre mit der Besetzungszahldifferenz eine Magnetisierung $M' = \delta n' \mu_B$ verbunden, da die magnetischen Momente der Elektronen im unteren Niveau in Richtung des Feldes weisen, im oberen dem Feld aber entgegengerichtet sind. Da außer bei sehr tiefen Temperaturen $\mu_B \mathcal{B} \ll k_B T$ ist, können wir die Entwicklung des Hyperbeltangens nach dem ersten Glied der Reihe abbrechen und für die Magnetisierung M' bzw. für die Suszeptibilität $\chi' = \mu_0 M'/\mathcal{B}$ näherungsweise

$$M' \simeq \frac{n\mu_B^2 \mathcal{B}}{k_B T} \qquad \text{oder} \qquad \chi' \simeq \mu_0 \frac{n\mu_B^2}{k_B T} \tag{7.26}$$

schreiben. Dieser Zusammenhang entspricht dem Curie-Gesetz paramagnetischer Festkörper, das im Kap. 10.1 diskutiert wird und für Systeme mit lokalisierten, unabhängigen Spins Gültigkeit besitzt. Für Natrium erhält man mit dieser Formel bei Zimmertemperatur den Wert $\chi'_{Na} = 6{,}9 \cdot 10^{-4}$. Der experimentelle Wert ist jedoch nur $\chi_{Na} = 8{,}3 \cdot 10^{-6}$ und darüber hinaus auch noch temperaturunabhängig! Ein Vergleich mit den experimentellen Werten anderer Metalle zeigt, daß die auf diese Weise errechnete paramagnetische Suszeptibilität immer um Größenordnungen zu hoch ist.

Tatsächlich verhindert die Fermi-Statistik wie bei der spezifischen Wärme die thermische Besetzung der beiden Zustände nach (7.25), da diese Gleichung nur für unabhängige Elektronenspins Gültigkeit besitzt, deren Wellenfunktionen sich nicht überlappen. Da bei Metallen jedoch ein Überlapp der Wellenfunktion der freien Elektronen besteht, wird die Fermi-Statistik wirksam. Um deren Auswirkung zu verdeutlichen, spalten wir, wie in Abb. 7.7 gezeigt, die Zustandsdichte in zwei Teile auf. Der eine Teil umfaßt die Elektronen, deren Spin in z-Richtung zeigt, während der Spin des zweiten Teils der z-Richtung entgegengerichtet ist. Ohne Magnetfeld heben sich die Beiträge der beiden Subsysteme zur Magnetisierung gerade auf. Durch das Feld wird der Energienullpunkt der beiden Teildichten um $\delta E = g_0 \mu_B \mathcal{B} \approx 2\mu_B \mathcal{B}$ gegeneinander verschoben, doch besitzt im Gleichgewicht die Fermi-Energie in der Probe überall denselben Wert. Da die beiden Zustandsdichten bis zur gemeinsamen Fermi-Energie mit Elektronen besetzt sind, weisen im Magnetfeld unterschiedlich viele Spins in die jeweilige Richtungen.

Wegen $\mu_B \mathcal{B} \ll E_F$ kann die Zustandsdichte $D(E_F \pm \mu_B \mathcal{B})$ in der Nähe der Fermi-Energie als konstant betrachtet werden. Die Elektronenkonzentration δn, deren magnetisches Moment nicht kompensiert wird, ist dann näherungsweise

$$\delta n = \frac{1}{2} D(E_F)\, 2\mu_B \mathcal{B} \,. \tag{7.27}$$

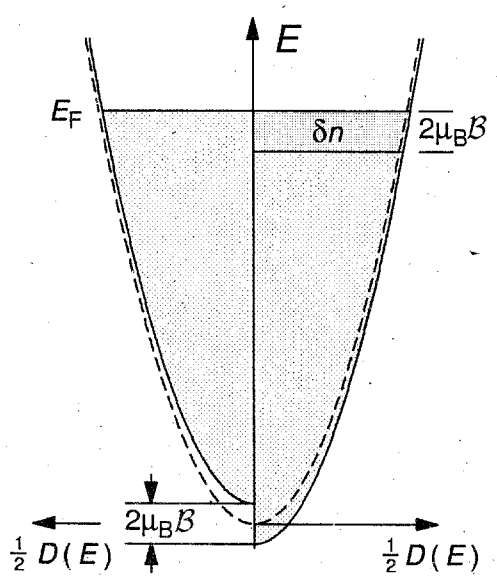


Abb. 7.7: Aufspaltung der Zustandsdichte nach den beiden Spinrichtungen. Links steht das magnetische Moment der Elektronen antiparallel, rechts dagegen parallel zum Magnetfeld. Die Elektronen in der dunklen Fläche bewirken die beobachtete Magnetisierung. Die Zustandsdichte ohne Feld ist gestrichelt gezeichnet.

Magnetisierung und Suszeptibilität, die wir auf diese korrekte Weise errechnen, sind temperaturunabhängig und relativ klein:

$$M = \delta n \mu_B = D(E_F)\,\mu_B^2 \mathcal{B} = \frac{3n\mu_B^2 \mathcal{B}}{2k_B T_F}\,, \tag{7.28}$$

$$\chi_P = \mu_0 D(E_F)\,\mu_B^2\,. \tag{7.29}$$

Wir haben hier bei der Bezeichnung der Suszeptibilität den Index p angehängt um anzudeuten, daß es sich um die paramagnetische Suszeptibilität des Elektronengases handelt. Wie wir in Abschnitt 8.3 sehen werden, treten im freien Elektronengas zusätzlich diamagnetische Beiträge auf. Vergleichen wir die beiden für die paramagnetische Suszeptibilität errechneten Werte, so finden wir

$$\frac{M}{M'} = \frac{\chi_P}{\chi'} = \frac{3T}{2T_F}\,. \tag{7.30}$$

Aufgrund der Wirksamkeit der Fermi-Statistik tritt wie bei der spezifischen Wärme eine Reduktion der Suszeptibilität um den Faktor T/T_F ein, da nur der Bruchteil der Spins an der Fermi-Kante zur Magnetisierung beiträgt, der nicht durch entgegengesetzt ausgerichtete Spins kompensiert wird.

Die von den Elektronenspins hervorgerufene Magnetisierung wird als *Paulische Spinsuszeptibilität* oder *Pauli-Paramagnetismus* bezeichnet. Bei der Beurteilung der Übereinstimmung mit den experimentellen Daten muß berücksichtigt werden, daß die Leitungselektronen gleichzeitig auch Diamagnetismus zeigen, der durch ihre Bahnbewegung hervorgerufen wird. Addiert man die beiden entgegengesetzt wirkenden Beiträge, so stimmen die Vorhersagen bei einfachen Metallen gut mit den Beobachtungen überein. Wie bei der spezifischen Wärme findet man aber erhebliche Abweichungen bei den Übergangsmetallen. Die Ursache hierfür ist in beiden Fällen die gleiche, nämlich die Wechselwirkung der Elektronen mit den umgebenden Teilchen und die Abweichung der tatsächlichen Zustandsdichte der Metalle von der Zustandsdichte des freien Elektronengases.

Abgeschirmtes Coulomb-Potential, Mottscher Metall-Isolator-Übergang

Das Elektronengas reagiert auf lokale elektrische Ladungsänderungen, wie sie beispielsweise durch geladene Punktdefekte hervorgerufen werden. Je nach Vorzeichen der Ladung werden Elektronen angezogen oder abgestoßen, bauen auf diese Weise eine Raumladung auf und bewirken so eine Abschirmung des elektrischen Feldes der Störstelle. Die Abschirmung spielt natürlich auch bei ungestörten Festkörpern eine wesentliche Rolle, denn sie verändert auch das Potential der Ionenrümpfe. Wir werden sehen, daß die Abschirmung mitverantwortlich dafür ist, ob ein Festkörper metallische oder isolierende Eigenschaften besitzt.

Für eine stark vereinfachte Behandlung des Abschirmeffektes stellen wir uns vor, daß eine zusätzliche Ladung e am Ort $\boldsymbol{r}_0$ in den Festkörper eingebracht wird, die das Störpotential $\phi(\boldsymbol{r})$ verursacht. Die zusätzliche Ladung bewirkt eine lokale Verschiebung des Nullpunkts der Energie der freien Elektronen um den Betrag $e\Phi(\boldsymbol{r})$. Wie in Abb. 7.8b dargestellt, entsteht bei einer positiven Störladung eine Mulde in der Besetzung der Zustände, die durch Zufluß von Elektronen aus dem restlichen Festkörper aufgefüllt wird. Entsprechend bewirkt eine negative Ladung eine Anhebung und einen Elektronenabfluß.

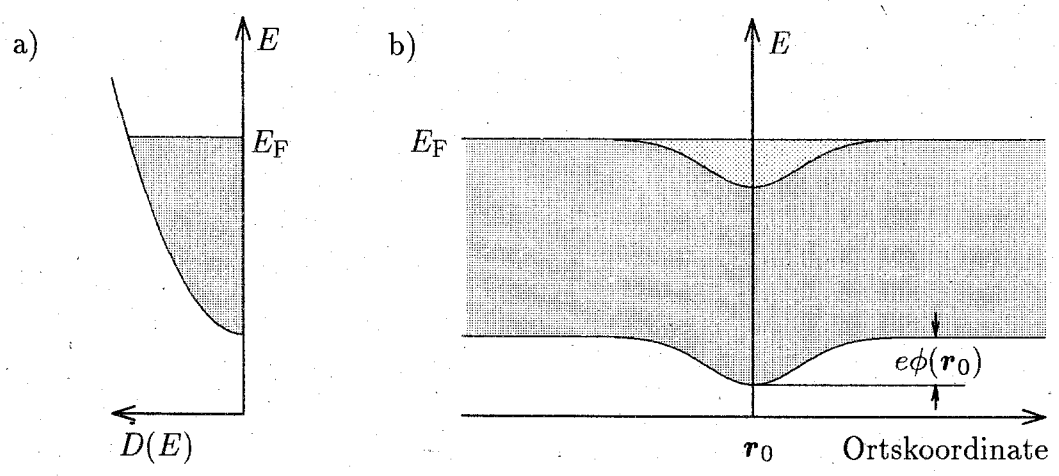


Abb. 7.8: Lokale Veränderung der Zustandsdichte durch eine zusätzliche positive Ladung. Die besetzten Zustände sind jeweils durch dunkle Tönung gekennzeichnet.
a) Zustandsdichte in größerer Entfernung vom Störpotential.
b) Energetische Lage der besetzten Zustände in der Nachbarschaft der positiven Ladung. Die hellgraue Mulde wird durch Elektronen des restlichen Festkörpers aufgefüllt.

In der Umgebung der Störung wird durch die Elektronenverschiebung deren Konzentration um den Betrag

$$\Delta n(\boldsymbol{r}) = -D(E_\mathrm{F})\, e\phi(\boldsymbol{r}) \tag{7.31}$$

gegenüber dem Gleichgewichtswert verschoben, d.h. die Ladungsdichte ändert sich lokal. Den Zusammenhang zwischen dem Potential $\phi(\boldsymbol{r})$ und der dadurch aufgebauten Raumladung stellt die Poisson-Gleichung

$$\nabla^2\phi(\boldsymbol{r}) = \frac{e}{\varepsilon_0}\left[\Delta n(\boldsymbol{r}) - \delta(\boldsymbol{r}-\boldsymbol{r}_0)\right] \tag{7.32}$$

her. Die zusätzlich eingebrachte Punktladung haben wir durch die Deltafunktion $\delta(\boldsymbol{r} - \boldsymbol{r}_0)$ berücksichtigt. Diese partielle Differentialgleichung besitzt eine kugelsymmetrische Lösung der Form

$$\phi(r) = \frac{-e}{4\pi\varepsilon_0} \frac{\mathrm{e}^{-r/r_{\mathrm{TF}}}}{r} , \tag{7.33}$$

wobei r_{TF} als Abkürzung für die *Thomas-Fermi-Abschirmlänge*

$$r_{\mathrm{TF}} = \sqrt{\frac{\varepsilon_0}{e^2 D(E_{\mathrm{F}})}} , \tag{7.34}$$

steht. Die Abschirmung bewirkt, daß in der Gleichung für den Potentialverlauf neben der typischen $1/r$-Abhängigkeit des Coulomb-Potentials zusätzlich ein exponentieller Faktor auftritt. Das Ergebnis läßt sich unmittelbar auf die Abschirmung einer negativen Ladung übertragen. In Abb. 7.9 ist das reine Coulomb-Potential mit einem abgeschirmten verglichen. Das *abgeschirmte Coulomb-Potential* hat die gleiche Form wie das in der Kernphysik benutzte *Yukawa-Potential* und wurde erstmals von *Debye* und *Hückel* abgeleitet, um die elektrostatische Wechselwirkung von Ionen in Elektrolyten zu beschreiben.

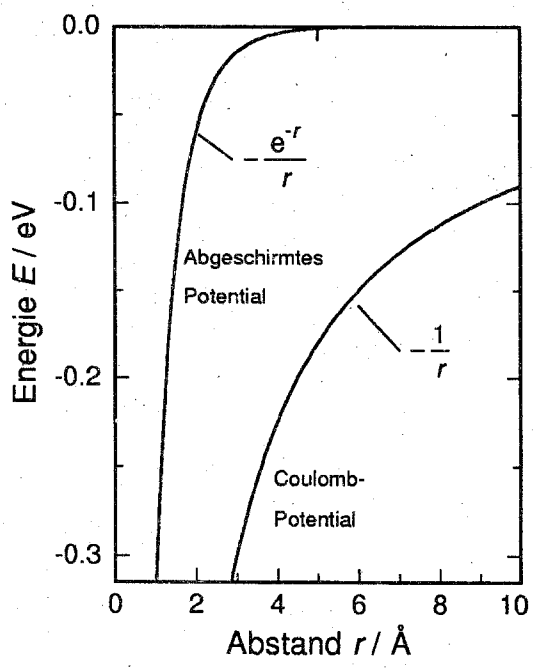


Abb. 7.9: Vergleich zwischen einem reinen und einem abgeschirmten Coulomb-Potential einer Elementarladung. Für die Abschirmlänge wurde der Wert $r_{\mathrm{TF}} = 1$ Å gewählt.

Abschirmeffekte spielen bei vielen Festkörpereigenschaften eine wichtige Rolle. So sind sie bei der quantitativen Beschreibung von Defekteigenschaften von erheblicher Bedeutung, denn durch die Abschirmung werden die Reichweiten der Coulomb-Felder stark reduziert. Elektronen, die weiter als die Abschirmlänge r_{TF} von einer Störladung entfernt sind, werden kaum noch von deren Feld beeinflußt. Um ein Gefühl für die Reichweite der Abschirmung zu bekommen, setzen wir die Elektronendichte $n = 8{,}5 \cdot 10^{28}\ \mathrm{m}^{-3}$ von Kupfer ein und erhalten den sehr kleinen Wert $r_{\mathrm{TF}} \approx 0{,}55$ Å, d.h. in Metallen ist die Abschirmung aufgrund der hohen Elektronendichte äußerst effektiv.

Die Abschirmung ist auch bei regulären Ionenrümpfen wirksam und verkleinert die Reichweite ihres Potentials. Wir können das Konzept benutzen, um eine einfache Einordnung von Materialien in die Gruppe der Metalle oder Isolatoren vorzunehmen. Neben

der Elektronenkonzentration muß für die Unterscheidung nur noch der Gitterabstand der Atome bekannt sein. Je wirksamer die Abschirmung aufgrund einer hohen Elektronendichte ist, umso enger werden die Potentiale (vgl. Abb. 7.9), und umso stärker werden die Rumpfelektronen räumlich lokalisiert. Bedingt durch die Unschärferelation steigt durch die Lokalisation die kinetische Energie der Elektronen an, und die Zustände werden energetisch angehoben. Bei sehr hoher Elektronenkonzentration liegen die Zustände der äußeren Elektronen so hoch, daß die Elektronen nicht mehr gebunden, sondern frei sind.

Eine numerische Lösung der Schrödinger-Gleichung für ein Elektron in einem abgeschirmten Potential der Form (7.33) führt nur für die Abschirmlänge $r_{\mathrm{TF}} > 0{,}84\,a_0$ zu gebundenen Eigenzuständen. Hierbei steht a_0 wieder für den Bohrschen Radius. Um die kritische Elektronendichte bei gegebenem Gitterabstand abzuschätzen, formen wir (7.34) um und erhalten

$$\frac{1}{r_{\mathrm{TF}}^2} = \frac{3ne^2}{2\varepsilon_0 E_{\mathrm{F}}} = \frac{4(3\pi^2)^{1/3}}{\pi}\,\frac{n^{1/3}}{a_0} \approx \frac{4n^{1/3}}{a_0}\,. \tag{7.35}$$

Demnach kann eine Substanz nur dann ein Isolator sein, wenn zwischen der Elektronendichte und dem Bohrschen Radius die folgende Ungleichung besteht:

$$n < \frac{0{,}05}{a_0^3}\,. \tag{7.36}$$

In Oxiden von Übergangsmetallen sowie in Gläsern, festen amorphen und flüssigen Halbleitern lassen sich Elektronendichten über die Materialzusammensetzung verändern. Die beobachteten abrupten Leitfähigkeitsänderungen bei bestimmten Konzentrationen werden auf den hier diskutierten Abschirmeffekt zurückgeführt. Besonders eindrucksvolle Beispiele für diesen sogenannten *Mottschen Metall-Isolator-Übergang* findet man bei stark dotierten Halbleitern (vgl. Kap. 4.1 und 9.2), bei denen sich die Konzentration der freien Elektronen über weite Bereiche verändern läßt. In Abb. 7.10 ist die elektrische Leitfähigkeit von Silizium mit unterschiedlicher Dotierung und damit unterschiedlicher Elektronenkonzentration dargestellt.

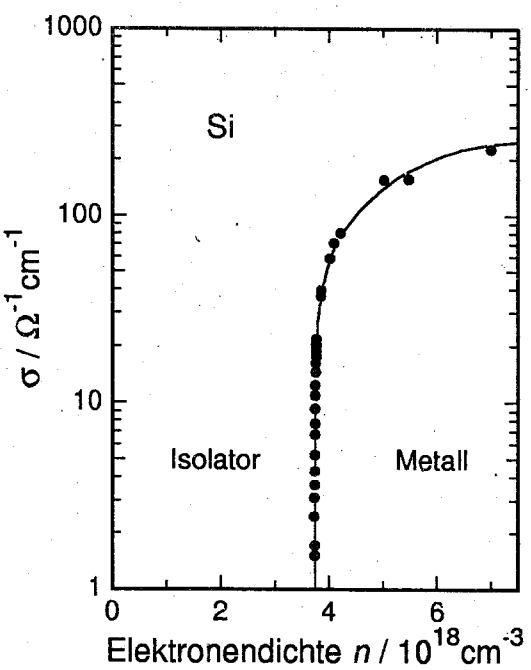


Abb. 7.10: Elektrische Leitfähigkeit von stark dotierten Silizium-Proben. Der Übergang vom Isolator zum Metall erfolgt bei einer Elektronenkonzentration von $n = 3{,}7 \cdot 10^{24}\,\mathrm{m}^{-3}$. (Nach H.F. Hess, K. DeCouch, T.F. Rosenbaum, G.A. Thomas, Phys. Rev. B **25** (1982) 5578).

In der Nähe des Übergangs hängt die elektrische Leitfähigkeit außerordentlich stark von der Elektronendichte ab. Bei der Dichte $n = 3{,}7 \cdot 10^{24}\,\mathrm{m}^{-3}$ findet der Übergang vom Isolator zum Metall statt. Bei dem hier diskutierten System muß man jedoch den Bohrschen Radius in (7.36) durch den wesentlich größeren *effektiven* Bohrschen Radius ersetzen, da das Coulomb-Potential auch in der nichtmetallischen Phase abgeschirmt wird. Diesen Gesichtspunkt der Abschirmung von geladenen Störstellen in Halbleitern und ihre Beschreibung mit Hilfe der Dielektrizitätskonstanten haben wir bereits in Kap. 4.1 kennengelernt.

Glühemission

Die *endliche Tiefe* des Potentials, das die Elektronen im Festkörper festhält, macht sich bei Metallen zum Beispiel im Auftreten der Glühemission bemerkbar. Mit Hilfe eines elektrischen Feldes wird bei einem derartigen Experiment aus der geheizten metallischen Kathode einer Elektronenröhre ein Strom gezogen. Mit zunehmender Feldstärke steigt die Stromdichte an und erreicht schließlich einen temperaturabhängigen Sättigungswert j_s, bei dem alle Elektronen, die die Kathode verlassen können, zum Stromfluß beitragen. Wie wir sehen werden, läßt sich die Abhängigkeit der Sättigungsstromdichte von der Temperatur und dem Kathodenmaterial mit unserem einfachen Modell des freien Elektronengases erklären.

Da die Metallelektronen gebunden sind, liegt die Fermi-Energie unter dem Potential E_Vak des Vakuums. Die Energiedifferenz $\Phi = E_\mathrm{Vak} - E_\mathrm{F}$, die von Elektronen beim Verlassen des Metalls aufgewandt werden muß, wird als *Austrittsarbeit* bezeichnet. Bei der Durchführung der Messung bewirkt das in x-Richtung angelegte elektrische Feld $\mathcal{E}$ im Außenraum eine Potentialvariation $-ex\mathcal{E}$, wie in Abb. 7.11 eingezeichnet. Zusätzlich muß noch berücksichtigt werden, daß eine Ladung vor einer Metalloberfläche eine Bildkraft hervorruft. Dadurch wird eine Absenkung $\delta\Phi$ der effektiven Austrittsarbeit bewirkt, die von der angelegten Feldstärke abhängt. Zwischen der tatsächlichen und der effektiven Austrittsarbeit Φ_eff findet man leicht den Zusammenhang

$$\Phi_\mathrm{eff} = \Phi - \delta\Phi = \Phi - \sqrt{\frac{e^3\mathcal{E}}{4\pi\varepsilon_0}}\,. \tag{7.37}$$

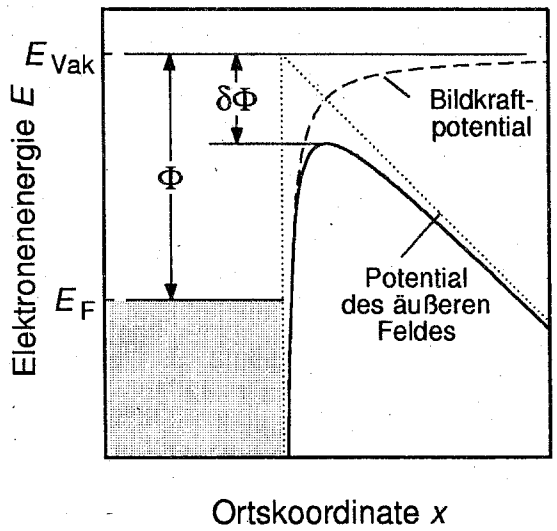


Abb. 7.11: Verlauf des Potentials an der Grenzfläche Metall/Vakuum. Eingezeichnet ist das Potential des angelegten elektischen Feldes, der Bildkraft und das tatsächliche Potential (durchgezogene Kurve). Die Bildkraft erniedrigt die Austrittsarbeit Φ um den Wert $\delta\Phi$.

Wir wollen nun die Sättigungsstromdichte j_s berechnen. Dazu müssen wir uns überlegen, welcher Bruchteil der Elektronen, die gegen die Begrenzung des Metalls laufen, tatsächlich aus diesem austreten können. Haben alle Elektronen eines Metalls die gleiche Driftgeschwindigkeit, so ist die Stromdichte durch $\boldsymbol{j} = -en\boldsymbol{v}$ gegeben. Bei der Berechnung des Stroms aus der Probe müssen wir beachten, daß die Metallelektronen unterschiedliche Wellenvektoren und somit unterschiedliche Geschwindigkeiten $\boldsymbol{v} = \hbar\boldsymbol{k}/m$ besitzen. Wir summieren daher die Beiträge der Elektronen mit den unterschiedlichen Geschwindigkeiten $\boldsymbol{v}(\boldsymbol{k})$ auf. Bei der Berechnung des Stroms gehen wir von einer Probe aus, deren Begrenzung senkrecht zur Feldrichtung verläuft, so daß wir beim Strom nur die x-Komponente

$$j_x = -\frac{e}{V}\sum_{\boldsymbol{k}} v_x(\boldsymbol{k}) \tag{7.38}$$

berücksichtigen müssen. Die Rechnung läßt sich leichter ausführen, wenn wir von der Summe $\sum_{\boldsymbol{k}}$ zum Integral $\int \rho_k \mathrm{d}^3k$ übergehen. Die Dichte der erlaubten Zustände im Impulsraum $\rho_k = 2V/(2\pi)^3$ haben wir bereits in Kap. 7.1 hergeleitet, wobei der Faktor 2 im Zähler die beiden Spineinstellmöglichkeiten berücksichtigt. Da nicht alle erlaubten Zustände besetzt sind, müssen wir noch mit der Fermi-Dirac-Funktion multiplizieren. Weiter berücksichtigen wir, daß die Geschwindigkeit und der Wellenvektor der Elektronen über die Beziehung $m\boldsymbol{v} = \hbar\boldsymbol{k}$ zusammenhängen und erhalten somit

$$j_x = \frac{-2e\hbar}{(2\pi)^3 m}\int_{-\infty}^{\infty} \mathrm{d}k_y\,\mathrm{d}k_z \int_{k_{x,\min}}^{\infty} k_x\, f[E(\boldsymbol{k}),T]\,\mathrm{d}k_x\,. \tag{7.39}$$

In der Fermi-Dirac-Funktion haben wir $E(\boldsymbol{k})$ anstelle von E geschrieben, um zu verdeutlichen, daß die Funktion vom Wellenvektor abhängt. Da nur jene Elektronen das Metall verlassen können, deren kinetische Energie in x-Richtung größer als $E_\mathrm{F} + \Phi_\mathrm{eff}$ ist, ist die untere Integrationsgrenze durch $k_{x,\min} = [2m(E_\mathrm{F} + \Phi_\mathrm{eff})/\hbar^2]^{1/2}$ gegeben. Negative Werte von k_x kommen nicht vor, da nur Elektronen zum Stromfluß beitragen, die sich auf die Oberfläche zu bewegen.

Die Austrittsarbeit ist wesentlich größer als die mittlere thermische Energie, so daß $E - E_\mathrm{F} \gg k_\mathrm{B}T$ ist. Unter dieser Voraussetzung kann die Fermi-Dirac-Verteilung näherungsweise durch die Boltzmann-Verteilung ersetzt werden, und wir erhalten:

$$j_x = \frac{-e\hbar}{4\pi^3 m}\,\mathrm{e}^{E_\mathrm{F}/k_\mathrm{B}T}\int_{-\infty}^{\infty}\mathrm{e}^{-\hbar^2k_z^2/2m\,k_\mathrm{B}T}\mathrm{d}k_z\int_{-\infty}^{\infty}\mathrm{e}^{-\hbar^2k_y^2/2m\,k_\mathrm{B}T}\mathrm{d}k_y\int_{k_{x,\min}}^{\infty}k_x\,\mathrm{e}^{-\hbar^2k_x^2/2m\,k_\mathrm{B}T}\mathrm{d}k_x\,. \tag{7.40}$$

Alle auftretenden Integrationen können analytisch durchgeführt werden. Bei genügend hohen Feldstärken ist die berechnete Stromdichte j_x identisch mit der Sättigungsstromdichte j_s, und wir erhalten als Endergebnis die *Richardson-Dushman*-Formel

$$j_\mathrm{s} = -\frac{emk_\mathrm{B}^2}{2\pi^2\hbar^3}\,T^2\mathrm{e}^{-\Phi_\mathrm{eff}/k_\mathrm{B}T}\,, \tag{7.41}$$

wobei der universelle Vorfaktor $emk_\mathrm{B}^2/2\pi^2\hbar^3$ den Wert $120\,\mathrm{AK^{-2}cm^{-2}}$ besitzt.

In der Herleitung haben wir implizit angenommen, daß alle Elektronen, die mit der Geschwindigkeit $v_x > [2(E_\mathrm{F}+\Phi_\mathrm{eff})/m]^{1/2}$ auf die Oberfläche treffen, das Metall verlassen, alle

anderen aber reflektiert werden. Diese Annahme ist jedoch nicht richtig. Berechnet man das Transmissionsverhalten von Elektronen an einer Potentialschwelle endlicher Dicke mit Hilfe der Quantenmechanik, so findet man ein modifiziertes Verhalten. Einerseits können Elektronen die Potentialbarriere durchtunneln, wenn $k_x < k_{x,\min}$ ist, andererseits werden Elektronen reflektiert, für die $k_x > k_{x,\min}$ ist. Die Berücksichtigung dieses Effekts führt zu einem zusätzlichen, schwach temperaturabhängigen Faktor, der die Sättigungsstromdichte leicht erniedrigt.

Bei der experimentellen Ermittlung der Austrittsarbeit Φ muß die gemessene Sättigungsstromdichte j_s auf die Feldstärke Null extrapoliert werden. Der Wert von Φ hängt nicht nur vom Material, sondern auch von der Kristallorientierung und der Oberflächenbeschaffenheit ab. Typische Werte sind in der Tabelle 7.3 zu finden.

Tabelle 7.3: Austrittsarbeit einiger Metalle, gemessen an polykristallinen Proben.

	Li	Na	Cs	Cu	Au	Pt
Φ (eV)	2,4	2,35	1,81	4,65	5,1	5,65

Tunnelstrom aus Metallen

Zu Beginn von Kap. 2 wurde auf die Abbildung von Oberflächen auf atomarer Skala mit Hilfe des Raster-Tunnelmikroskopes hingewiesen, dessen Meßprinzip hier kurz erläutert werden soll. Ähnlich wie bei der Glühemission werden bei dieser Technik Elektronen aus der Meßspitze oder der Probe herausgezogen, doch geschieht dies hier nur mit Hilfe eines elektrischen Feldes, ohne daß die Probe zusätzlich aufgeheizt wird. Das Meßprinzip eines derartigen Mikroskops läßt sich an der schematischen Darstellung des Aufbaus in Abb. 7.12 erkennen.

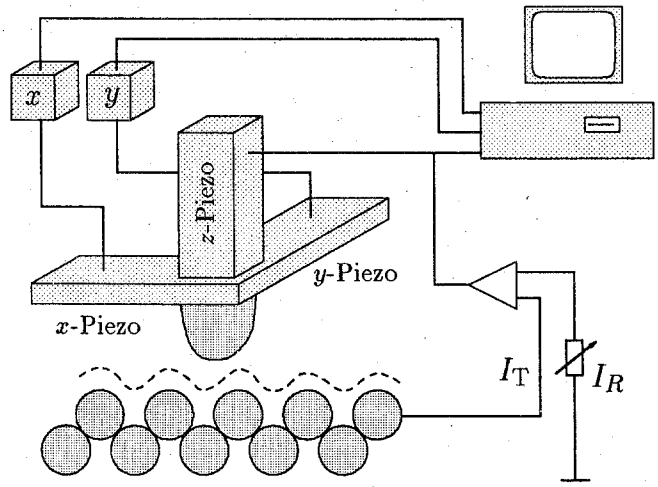


Abb. 7.12: Schematische Darstellung der Funktionsweise des Raster-Tunnelmikroskops. Um reproduzierbare kleine Positionsänderungen in x- bzw. y-Richtung vorzunehmen, wird mit den beiden Generatoren x und y Spannung an die Stellelemente angelegt. Das z-Piezoelement hält während des Rasterns den Tunnelstrom I_T und damit den Spitzenabstand konstant sofern sich nicht die Austrittsarbeit lokal verändert. I_R gibt den Sollwert des Tunnelstroms und somit den Spitzenabstand vor.

Eine feine Metallspitze, die im Idealfall in einem einzelnen Atom endet, wird durch piezoelektrische Stellelemente über die Oberfläche geführt. Ist der Abstand d zwischen Spitze und Probe genügend klein, so tritt ein Tunnelstrom auf. Da die atomaren Wellenfunktionen nach außen hin exponentiell abklingen, hängen der Überlapp der Wellenfunktionen

von Probe und Spitze und somit Tunnelwahrscheinlichkeit und Tunnelstrom empfindlich vom Abstand ab. Mit Hilfe des z-Stellelements wird beim Rastern der Oberfläche der Abstand so eingestellt, daß der Tunnelstrom I_T konstant bleibt. Aus der Regelspannung am Stellelement läßt sich die Höhenvariation der Metallspitze und damit die Lage der Atome auf der Oberfläche rekonstruieren. Neben dieser Betriebsart („Konstantstrom-Modus") des Raster-Tunnelmikroskops gibt es noch weitere, die wir aber nicht weiter erläutern.

Während bei der Glühemission die Elektronen aus dem Potential des Metalls ins Vakuum gehen, tunneln sie beim Raster-Tunnelmikroskop durch das Vakuum von einem Metall zum anderen. Hierzu ist keine Energie erforderlich, so daß der Prozeß nicht thermisch aktiviert ist. Voraussetzung für das Auftreten des Tunnelstroms ist, daß Elektronen aus besetzten Zustände in freie tunneln können, die durch das Anlegen der Spannung U zugänglich gemacht werden, denn die Spannung verschiebt die beiden Fermi-Energien um den Betrag $-eU$. In Abb. 7.13 sind die Potentialverhältnisse an der Tunnelspitze stark vereinfacht für den Fall dargestellt, daß Probe und Spitze aus demselben Metall bestehen.

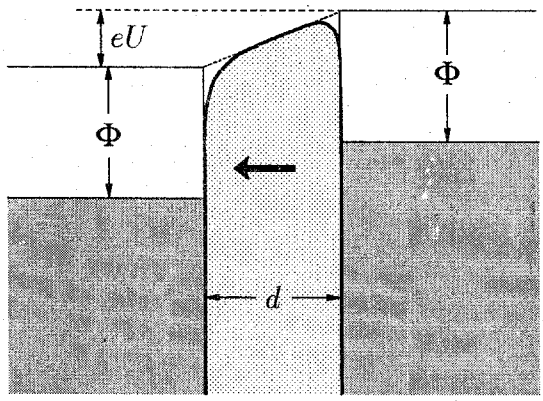


Abb. 7.13: Potentialverlauf an der Tunnelbarriere beim Anliegen einer Spannung. Der Tunnelstrom der Elektronen von den besetzten in die freien Zustände ist durch einen dicken Pfeil angedeutet.

Für die Tunnelwahrscheinlichkeit der Elektronen und damit für den Tunnelstrom I_T findet man mit Hilfe der WKB-Methode den Ausdruck

$$I_T \propto \frac{U}{d} \mathrm{e}^{-d\sqrt{2m\Phi_{\mathrm{eff}}/\hbar^2}} , \tag{7.42}$$

wobei Φ_{eff} hier die effektive Höhe der Potentialbarriere, also grob die Austrittsarbeit Φ, ist. Sieht man vom exakten Zahlenwert ab, der durch die Form des Potentialbergs bestimmt wird, so haben wir diesen Ausdruck bereits in Kap. 5.7 bei der Diskussion der Tunnelsysteme kennengelernt. Der genauere Verlauf des Potentials, der Einfluß der Bildladung und weitere experimentelle Gegebenheiten wurden in der Gleichung (7.42) grob durch die Einführung einer effektiven Potentialbarriere Φ_{eff} berücksichtigt.

7.3 Elektronen im periodischen Potential

Das Modell der freien Elektronen kann eine ganze Reihe von Eigenschaften erklären, u.a. auch Transportvorgänge, auf die wir erst im nächsten Kapitel eingehen. Es treten aber auch Abweichungen auf, die deutlich machen, daß dieses einfache Modell nicht immer realistisch sein kann. Hierzu zählt der Hall-Effekt, der eine Bestimmung der Elektronendichte

erlaubt und ebenfalls im nächsten Kapitel diskutiert wird. So mißt man beim Natrium erwartungsgemäß, daß pro Atom ein Elektron als frei betrachtet werden kann. Beim Beryllium findet man jedoch anstelle der zwei Elektronen pro Atom, die sich aus dem Periodensystem ergeben, 0,2 *positive* Ladungsträger. Es sieht so aus, als würden nicht die Elektronen, sondern positive Ladungsträger den Stromtransport bewirken! Offensichtlich müssen wir die Beschreibung der elektronischen Eigenschaften in diesem Fall stark modifizieren: Es muß berücksichtigt werden, daß sich die Elektronen in einem *periodischen Potential* bewegen.

Wellengleichung eines Elektrons in einem periodischen Potential

In der folgenden Herleitung spalten wir die stationäre Schrödinger-Gleichung eines Elektrons im Gitter in einen Satz von linearen Gleichungen auf, die sich näherungsweise lösen lassen. Dabei berücksichtigen wir in

$$H\psi(\boldsymbol{r}) = \left[-\frac{\hbar^2}{2m}\Delta + \widetilde{V}(\boldsymbol{r})\right]\psi(\boldsymbol{r}) = E\,\psi(\boldsymbol{r})\;, \tag{7.43}$$

daß das Potential $\widetilde{V}(\boldsymbol{r}) = \widetilde{V}(\boldsymbol{r}+\boldsymbol{R})$ die Translationssymmetrie des Gitters besitzt, wobei $\boldsymbol{R}$ für einen Gittervektor steht. Das Potential läßt sich daher wie die Streudichte in Kap. 2.3 in eine Fourier-Reihe

$$\widetilde{V}(\boldsymbol{r}) = \sum_{\boldsymbol{G}} \widetilde{V}_{\boldsymbol{G}}\,\mathrm{e}^{\mathrm{i}\boldsymbol{G}\cdot\boldsymbol{r}} \tag{7.44}$$

nach reziproken Gittervektoren $\boldsymbol{G}$ entwickeln. Die Fourier-Koeffizienten $\widetilde{V}_{\boldsymbol{G}}$ sind durch das Gitterpotential vorgegeben. Es ist naheliegend, auch die Wellenfunktion selbst nach ebenen Wellen zu entwickeln, da sich diese Näherung beim Modell freier Elektronen bereits als geeignet erwiesen hat:

$$\psi(\boldsymbol{r}) = \sum_{\boldsymbol{k}} c_{\boldsymbol{k}}\,\mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}}\;. \tag{7.45}$$

Die Entwicklungskoeffizienten $c_{\boldsymbol{k}}$ sind zunächst noch nicht festgelegt und werden erst im Laufe der Rechnung bestimmt.

Ein ähnlicher Ansatz wurde bereits für die Gitterschwingungen in Kap. 5.2 gemacht, um die Schwingungen der einzelnen Atome zu „entkoppeln“, doch besteht hier nicht die Forderung, daß der Wellenvektor in der 1. Brillouin-Zone liegt. Der Wellenvektor $\boldsymbol{k}$ muß aber wie beim freien Elektronengas zusätzlich den periodischen Randbedingungen genügen. Nachdem wir die Entwicklungen des Potentials und der Wellenfunktion in die Schrödinger-Gleichung eingesetzt haben, finden wir den Ausdruck

$$\sum_{\boldsymbol{k}} \frac{\hbar^2 k^2}{2m} c_{\boldsymbol{k}}\,\mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}} + \sum_{\boldsymbol{k}',\boldsymbol{G}} c_{\boldsymbol{k}'}\widetilde{V}_{\boldsymbol{G}}\,\mathrm{e}^{\mathrm{i}(\boldsymbol{k}'+\boldsymbol{G})\cdot\boldsymbol{r}} = E\sum_{\boldsymbol{k}} c_{\boldsymbol{k}}\,\mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}}\;. \tag{7.46}$$

Dabei haben wir bei der zweiten Summe den Wellenvektor nicht mit $\boldsymbol{k}$, sondern mit $\boldsymbol{k}'$ bezeichnet, um nach der Umbenennung der Summationsindizes $\boldsymbol{k}' + \boldsymbol{G} \to \boldsymbol{k}$ zu einem einfachen Ausdruck zu gelangen. Diese etwas verwirrende Vorgehensweise führt zum richtigen Ergebnis, weil die Summation über *alle* Werte der Wellenvektoren $\boldsymbol{k}$ und der reziproken Gittervektoren $\boldsymbol{G}$ erfolgt. Durch die Umbenennung ändert sich nicht der Wert

der Glieder in der Summe, sondern nur die Reihenfolge der Addition. Wir hätten in (7.45) anstelle von $\boldsymbol{k}$ den Wellenvektor $\boldsymbol{k}+\boldsymbol{G}'$ wählen können, denn auch mit Wellenvektoren, die um einen beliebigen reziproken Gittervektor verschoben sind, ändert sich nichts an der Argumentation. Nach Umbenennung der Summationsindizes können wir vereinfachend

$$\sum_{\boldsymbol{k}} \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}} \left[\left(\frac{\hbar^2 k^2}{2m} - E \right) c_{\boldsymbol{k}} + \sum_{\boldsymbol{G}} \widetilde{V}_{\boldsymbol{G}}\, c_{\boldsymbol{k}-\boldsymbol{G}} \right] = 0 \tag{7.47}$$

schreiben. Diese Gleichung muß für jeden Ortsvektor $\boldsymbol{r}$ gültig sein. Dies ist nur möglich, wenn der Ausdruck in der eckigen Klammer für jeden Wellenvektor $\boldsymbol{k}$ getrennt verschwindet. Somit gilt:

$$\left(\frac{\hbar^2 k^2}{2m} - E \right) c_{\boldsymbol{k}} + \sum_{\boldsymbol{G}} \widetilde{V}_{\boldsymbol{G}}\, c_{\boldsymbol{k}-\boldsymbol{G}} = 0\,. \tag{7.48}$$

Dies ist die stationäre Schrödinger-Gleichung eines Elektrons, das sich in einem periodischen Potential bewegt, dargestellt im $\boldsymbol{k}$-Raum.

Wir können nun für jeden Wellenvektor $\boldsymbol{k}$ getrennt eine Lösung finden. Die periodischen Randbedingungen erlauben bei $\mathcal{N}$ Elementarzellen $\mathcal{N}$ Werte für $\boldsymbol{k}$ und somit ebenfalls $\mathcal{N}$ Lösungen mit einem Wellenvektor in der 1. Brillouin-Zone. Jeweils $\mathcal{N}$ weitere Lösungen existieren in jeder der höheren Brillouin-Zonen. Wie wir gleich sehen werden, besitzen jedoch die Wellenvektoren der 1. Brillouin-Zone eine besondere Bedeutung. Zu jedem erlaubten Wellenvektor gehört ein Eigenwert $E_{\boldsymbol{k}}$. Dies bedeutet, daß das ursprünglich äußerst komplexe Problem für die Wellenvektoren der 1. Brillouin-Zone und alle weiteren in jeweils $\mathcal{N}$ getrennte Probleme aufgespalten wurde. Die auftretenden Koeffizienten $c_{\boldsymbol{k}}$ sind nicht mehr mit allen anderen Entwicklungskoeffizienten verknüpft, sondern nur noch mit solchen, die sich um reziproke Gittervektoren unterscheiden, d.h. mit $c_{\boldsymbol{k}-\boldsymbol{G}}$, $c_{\boldsymbol{k}-\boldsymbol{G}'}$, $c_{\boldsymbol{k}-\boldsymbol{G}''}$, ... Dennoch gibt es im Prinzip für die Wellenfunktion jedes Wellenvektors eine unendliche Zahl von Entwicklungskoeffizienten $c_{\boldsymbol{k}-\boldsymbol{G}}$, die bestimmt werden müssen. In der praktischen Durchführung reicht es aber meist, sich auf relativ wenige Koeffizienten zu beschränken. Dadurch läßt sich das Gleichungssystem (7.48) oft mit vergleichsweise einfachen Methoden lösen.

Bevor wir mit der Berechnung der Koeffizienten $c_{\boldsymbol{k}-\boldsymbol{G}}$ fortfahren, wollen wir schon einmal kurz die Struktur und einige Eigenschaften der Wellenfunktionen betrachten. Während im Lösungsansatz (7.46) für die Wellenfunktion über alle Wellenvektoren summiert wurde, die die Randbedingungen erfüllen, ist wegen der periodischen Gitterstruktur nur noch eine Summation über reziproke Gittervektoren erforderlich. Damit vereinfacht sich die Entwicklung (7.45) der Wellenfunktion zu

$$\psi_{\boldsymbol{k}}(\boldsymbol{r}) = \sum_{\boldsymbol{G}} c_{\boldsymbol{k}-\boldsymbol{G}}\, \mathrm{e}^{\mathrm{i}(\boldsymbol{k}-\boldsymbol{G})\cdot\boldsymbol{r}}\,. \tag{7.49}$$

Da das reziproke Gitter die Fourier-Entwicklung der Streudichte des realen Gitters widerspiegelt, ist zu erwarten, daß im allgemeinen die Summation bereits nach einigen Gliedern abgebrochen werden kann. Wie oben bereits erwähnt, gelten die hier durchgeführten Überlegungen und damit auch die Lösung (7.49) nicht nur für einen bestimmten Wellenvektor $\boldsymbol{k}$ in der 1. Brillouin-Zone, sondern für alle, die sich von diesem um reziproke Gittervektoren unterscheiden.

Bringen wir die Wellenfunktion (7.49) in die Form

$$\psi_{\boldsymbol{k}}(\boldsymbol{r}) = \left(\sum_{\boldsymbol{G}} c_{\boldsymbol{k}-\boldsymbol{G}} \, \mathrm{e}^{-\mathrm{i}\boldsymbol{G}\cdot\boldsymbol{r}} \right) \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}} \, , \tag{7.50}$$

so wird die Struktur der Lösung deutlicher. Der Ausdruck in der Klammer ist die Fourier-Entwicklung einer gitterperiodischen Funktion. Die Lösungen der Schrödinger-Gleichung für ein periodisches Potential sind also *ebene Wellen mit einem gitterperiodischen Modulationsfaktor*, den wir $u_{\boldsymbol{k}}(\boldsymbol{r})$ nennen. Damit läßt sich die Wellenfunktion eines Elektrons in einem periodischen Potential durch die *Bloch-Funktion*

$$\psi_{\boldsymbol{k}}(\boldsymbol{r}) = u_{\boldsymbol{k}}(\boldsymbol{r}) \, \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}} \tag{7.51}$$

ausdrücken, wobei sich die Gitterperiodizität in der Beziehung

$$u_{\boldsymbol{k}}(\boldsymbol{r}) = u_{\boldsymbol{k}}(\boldsymbol{r} + \boldsymbol{R}) \tag{7.52}$$

manifestiert. Ein einfaches Beispiel für eine Bloch-Funktion ist in Abb. 7.14 zu sehen.

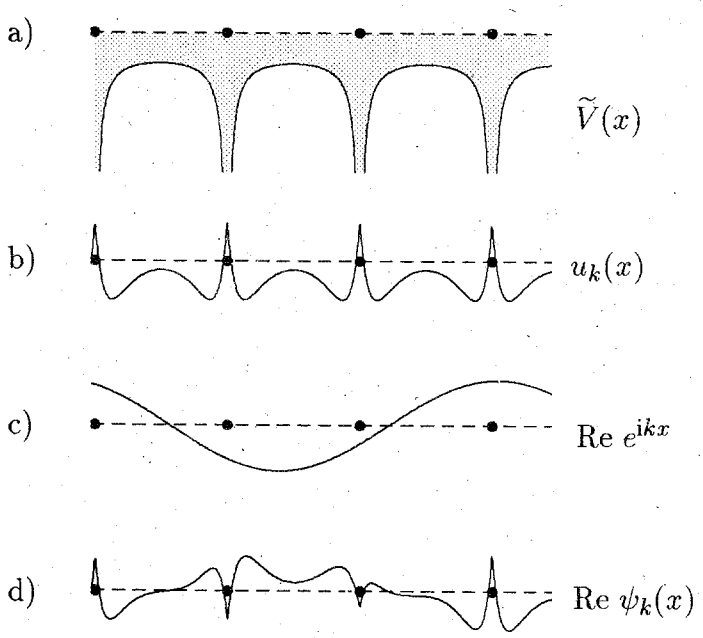


Abb. 7.14: Beispiel für eine Bloch-Funktion eines eindimensionalen Gitters. In dieser einfachen Darstellung kann nur der Realteil der komplexen Funktion wiedergegeben werden. Die Punkte geben die Lage der Atomrümpfe an.
a) Potentialverlauf $\widetilde{V}(x)$,
b) Verlauf der willkürlich gewählten gitterperiodischen Funktion $u_k(x)$,
c) Realteil des Phasenfaktors $\mathrm{e}^{\mathrm{i}kx}$,
d) Realteil der Bloch-Funktion $\psi_k(x)$.

Die Gleichungen (7.51) und (7.52) sind ein Spezialfall des von *Bloch* für Kristalle formulierten und bewiesenen Theorems, das folgendes besagt: Für jede beliebige Wellenfunktion, welche die Schrödinger-Gleichung erfüllt, existiert ein derartiger Wellenvektor $\boldsymbol{k}$, daß die Translation um einen Gittervektor $\boldsymbol{R}$ gleichwertig mit der Multiplikation mit dem Phasenfaktor $\mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{R}}$ ist. Durch Einsetzen kann man sich leicht davon überzeugen, daß die von uns gefundene Lösung diese Bedingung erfüllt:

$$\psi_{\boldsymbol{k}}(\boldsymbol{r} + \boldsymbol{R}) = u_{\boldsymbol{k}}(\boldsymbol{r} + \boldsymbol{R}) \, \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}} \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{R}} = \psi_{\boldsymbol{k}}(\boldsymbol{r}) \, \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{R}} \, . \tag{7.53}$$

Im allereinfachsten Fall freier Elektronen ist $u(\boldsymbol{r})$ eine Konstante. Auch die elastischen Gitterwellen, die wir in Kap. 5 behandelt haben, sind in diesem Sinn Bloch-Wellen, mit dem denkbar einfachsten gitterperiodischen Vorfaktor $U(\boldsymbol{q}) = \text{const}$.

Die Bloch-Funktionen $\psi_{\boldsymbol{k}}(\boldsymbol{r})$ haben zwei wichtige Eigenschaften, auf die wir hier besonders hinweisen. Bloch-Funktionen, deren Wellenvektoren sich um einen reziproken Gittervektor unterscheiden, sind gleich! Dieses überraschende Ergebnis haben wir oben bereits erwähnt; es folgt unmittelbar aus (7.50). Addieren wir zum Wellenvektor $\boldsymbol{k}$ den reziproken Gittervektor $\boldsymbol{G}'$, so erhalten wir

$$\psi_{\boldsymbol{k}+\boldsymbol{G}'}(\boldsymbol{r}) = \sum_{\boldsymbol{G}} c_{\boldsymbol{k}+\boldsymbol{G}'-\boldsymbol{G}}\, \mathrm{e}^{\mathrm{i}(\boldsymbol{k}+\boldsymbol{G}'-\boldsymbol{G})\cdot\boldsymbol{r}} \,. \tag{7.54}$$

Durch die Umbenennung der Summationsindizes $\boldsymbol{G} - \boldsymbol{G}' \to \boldsymbol{G}''$ folgt dann

$$\psi_{\boldsymbol{k}+\boldsymbol{G}'}(\boldsymbol{r}) = \underbrace{\sum_{\boldsymbol{G}''} c_{\boldsymbol{k}-\boldsymbol{G}''}\, \mathrm{e}^{-\mathrm{i}\boldsymbol{G}''\cdot\boldsymbol{r}}}_{u_{\boldsymbol{k}}(\boldsymbol{r})}\, \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}} \tag{7.55}$$

und somit

$$\psi_{\boldsymbol{k}+\boldsymbol{G}'}(\boldsymbol{r}) = \psi_{\boldsymbol{k}}(\boldsymbol{r}) \,. \tag{7.56}$$

Aber auch die Eigenwerte $E_{\boldsymbol{k}}$ sind im reziproken Raum periodisch! Für Eigenwerte, deren Wellenvektoren sich um reziproke Gittervektoren unterscheiden, lautet die Schrödinger-Gleichung:

$$H\psi_{\boldsymbol{k}} = E_{\boldsymbol{k}}\psi_{\boldsymbol{k}} \qquad \text{bzw.} \qquad H\psi_{\boldsymbol{k}+\boldsymbol{G}} = E_{\boldsymbol{k}+\boldsymbol{G}}\psi_{\boldsymbol{k}+\boldsymbol{G}} \,. \tag{7.57}$$

Mit (7.56) muß also

$$H\psi_{\boldsymbol{k}} = E_{\boldsymbol{k}+\boldsymbol{G}}\psi_{\boldsymbol{k}} \qquad \text{und somit} \qquad E_{\boldsymbol{k}} = E_{\boldsymbol{k}+\boldsymbol{G}} \tag{7.58}$$

sein. Die Wellenfunktionen der Elektronen und ihre Eigenwerte wiederholen sich im $\boldsymbol{k}$-Raum periodisch. Es reicht daher aus, sich auf die Lösungen in der *1. Brillouin-Zone zu beschränken*, da die Lösungen in diesem Teil des $\boldsymbol{k}$-Raums bereits die ganze Information enthalten. Diese Vorgehensweise wird als *Reduktion auf die 1. Brillouin-Zone* und die Darstellung als *reduziertes Zonenschema* bezeichnet. Wir werden in den folgenden Abschnitten noch ausführlich darauf eingehen.

Näherung für quasi-freie Elektronen

Wir wollen uns nun mit der Gleichung (7.48) anhand einiger besonders einfacher Fälle vertraut machen. Zunächst stellen wir das Gleichungssystem für ein Potential auf, das nur zwei Fourierkomponenten mit der Periodizität $\boldsymbol{g}$ und $2\boldsymbol{g}$ enthält, wobei $\boldsymbol{g}$ ein Basisvektor des reziproken Gitters ist. Den Nullpunkt des Potentials wählen wir so, daß der konstante Anteil $\widetilde{V}_0$ verschwindet. Setzen wir ein inversionssymmetrisches Potential voraus, so folgt für die Fourier-Koeffizienten der Entwicklung (7.44) der Zusammenhang $\widetilde{V}_{\boldsymbol{g}} = \widetilde{V}_{-\boldsymbol{g}}$. Mit der vereinfachten Schreibweise $\widetilde{V}_1 = \widetilde{V}_{\boldsymbol{g}} = \widetilde{V}_{-\boldsymbol{g}}$ und $\widetilde{V}_2 = \widetilde{V}_{2\boldsymbol{g}} = \widetilde{V}_{-2\boldsymbol{g}}$ bekommt dann das Potential die einfache Form

$$\begin{aligned}
\widetilde{V} &= \widetilde{V}_{\boldsymbol{g}} e^{i\boldsymbol{g}\cdot\boldsymbol{r}} + \widetilde{V}_{-\boldsymbol{g}} e^{-i\boldsymbol{g}\cdot\boldsymbol{r}} + \widetilde{V}_{2\boldsymbol{g}} e^{2i\boldsymbol{g}\cdot\boldsymbol{r}} + \widetilde{V}_{-2\boldsymbol{g}} e^{-2i\boldsymbol{g}\cdot\boldsymbol{r}} \\
&= \widetilde{V}_1 e^{i\boldsymbol{g}\cdot\boldsymbol{r}} + \widetilde{V}_1 e^{-i\boldsymbol{g}\cdot\boldsymbol{r}} + \widetilde{V}_2 e^{2i\boldsymbol{g}\cdot\boldsymbol{r}} + \widetilde{V}_2 e^{-2i\boldsymbol{g}\cdot\boldsymbol{r}} \\
&= 2\widetilde{V}_1 \cos(\boldsymbol{g}\cdot\boldsymbol{r}) + 2\widetilde{V}_2 \cos(2\,\boldsymbol{g}\cdot\boldsymbol{r}) \,.
\end{aligned} \tag{7.59}$$

Setzen wir dieses Potential unter Benutzung der Abkürzung $\eta_{\boldsymbol{k}+\boldsymbol{G}} = \hbar^2|\boldsymbol{k}+\boldsymbol{G}|^2/2m$ in die Schrödingergleichung ein, so ergibt sich:

$$\begin{array}{ccccccccccc}
\vdots & & \vdots & & \vdots & & \vdots & & \vdots & & \vdots \\
(\eta_{\boldsymbol{k}-\boldsymbol{g}}-E)\,c_{\boldsymbol{k}-\boldsymbol{g}} & + & \widetilde{V}_1\,c_{\boldsymbol{k}-2\boldsymbol{g}} & + & \widetilde{V}_1\,c_{\boldsymbol{k}} & + & \widetilde{V}_2 c_{\boldsymbol{k}-3\boldsymbol{g}} & + & \widetilde{V}_2 c_{\boldsymbol{k}+\boldsymbol{g}} & = & 0 \\
(\eta_{\boldsymbol{k}}-E)\,c_{\boldsymbol{k}} & + & \widetilde{V}_1\,c_{\boldsymbol{k}-\boldsymbol{g}} & + & \widetilde{V}_1\,c_{\boldsymbol{k}+\boldsymbol{g}} & + & \widetilde{V}_2 c_{\boldsymbol{k}-2\boldsymbol{g}} & + & \widetilde{V}_2 c_{\boldsymbol{k}+2\boldsymbol{g}} & = & 0 \\
(\eta_{\boldsymbol{k}+\boldsymbol{g}}-E)\,c_{\boldsymbol{k}+\boldsymbol{g}} & + & \widetilde{V}_1\,c_{\boldsymbol{k}} & + & \widetilde{V}_1\,c_{\boldsymbol{k}+2\boldsymbol{g}} & + & \widetilde{V}_2 c_{\boldsymbol{k}-\boldsymbol{g}} & + & \widetilde{V}_2 c_{\boldsymbol{k}+3\boldsymbol{g}} & = & 0 \\
\vdots & & \vdots & & \vdots & & \vdots & & \vdots & & \vdots
\end{array} \tag{7.60}$$

Dieses lineare Gleichungssystem hat nicht-triviale Lösungen, wenn die Determinante der Koeffizientenmatrix gleich Null ist, wobei die unendlich große Determinante folgendes Aussehen hat:

$$\begin{vmatrix}
 & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \\
\cdots & \widetilde{V}_1 & \eta_{\boldsymbol{k}-\boldsymbol{g}}-E & \widetilde{V}_1 & \widetilde{V}_2 & 0 & 0 & \cdots \\
\cdots & \widetilde{V}_2 & \widetilde{V}_1 & \eta_{\boldsymbol{k}}-E & \widetilde{V}_1 & \widetilde{V}_2 & 0 & \cdots \\
\cdots & 0 & \widetilde{V}_2 & \widetilde{V}_1 & \eta_{\boldsymbol{k}+\boldsymbol{g}}-E & \widetilde{V}_1 & \widetilde{V}_2 & \cdots \\
\cdots & 0 & 0 & \widetilde{V}_2 & \widetilde{V}_1 & \eta_{\boldsymbol{k}+2\boldsymbol{g}}-E & \widetilde{V}_1 & \cdots \\
 & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots &
\end{vmatrix} \tag{7.61}$$

Abhängig von der geforderten Genauigkeit muß ein mehr oder minder großer Teil der Determinante bearbeitet werden. Enthält das Potential wie im vorliegendem Fall nicht sehr viele Fourierkomponenten, so reicht es aus, einen Teil der Determinante mit nur wenigen Zeilen gleich Null zu setzen.

In einer äußerst weitgehenden Vereinfachung nehmen wir zunächst an, daß die Amplitude des periodischen Potentials so klein ist, daß wir $\widetilde{V}_{\boldsymbol{G}} \approx 0$ setzen dürfen. Die Symmetrie des Gitters soll aber nach wie vor periodische Lösungen erzwingen! Man spricht dann von einem *leeren Gitter*. Unter dieser Annahme folgt aus den Gleichungen (7.60) oder (7.48) der einfache Ausdruck $(\eta_{\boldsymbol{k}+\boldsymbol{G}}-E)=0$, wobei alle reziproken Gittervektoren erlaubt sind. Setzen wir den entsprechenden Ausdruck für η ein, so ergibt sich für die Energieeigenwerte die Lösung

$$E_{\boldsymbol{k}+\boldsymbol{G}} = \frac{\hbar^2}{2m}|\boldsymbol{k}+\boldsymbol{G}|^2\,. \tag{7.62}$$

Offensichtlich handelt es sich um Parabeln, die im $\boldsymbol{k}$-Raum jeweils um den Wellenvektor $\boldsymbol{k} = -\boldsymbol{G}$ gegeneinander verschoben sind.

Wir untersuchen zunächst das eindimensionale Gitter mit der Gitterkonstanten a und dem kleinsten reziproke Gittervektor $g = 2\pi/a$ etwas näher. Die Lösung (7.62) ist in Abb. 7.15 veranschaulicht. Wie bereits erwähnt, erlaubt die Periodizität der Lösung eine Darstellung aller Energieeigenwerte im Wellenvektorbereich $-\pi/a = k \leq \pi/a$. Bei dieser

„Reduktion" auf die 1. Brillouin-Zone verschiebt man Teile der Kurve um reziproke Gittervektoren so, daß sie innerhalb der 1. Brillouin-Zone zu liegen kommen. Man erhält eine Darstellung der Energieeigenwerte $E_{\boldsymbol{k}}$ wie in Abb. 7.16 zu sehen ist, wobei bei gleichem Wellenvektor $\boldsymbol{k}$ mehrere Energiewerte auftreten. Bei den hier diskutierten sehr einfachen Verhältnissen kommt man zum gleichen Ergebnis, wenn man aus der Abb. 7.15 einfach den Bereich der 1. Brillouin-Zone herausschneidet.

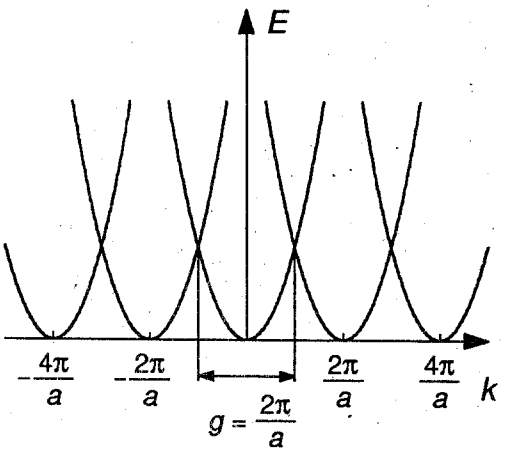


Abb. 7.15: Darstellung der Lösung der Schrödinger-Gleichung eines eindimensionalen „leeren" Gitters mit der Gitterkonstanten a. Die Energieeigenwerte wiederholen sich im $\boldsymbol{k}$-Raum mit der Periode $g = 2\pi/a$.

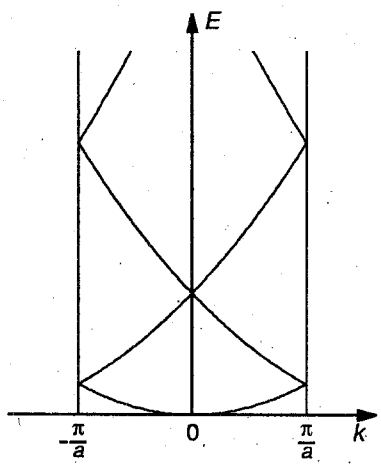


Abb. 7.16: Reduktion der Energieeigenwerte eines eindimensionalen Gitters auf die 1. Brillouin-Zone. Alle erlaubten Energieeigenwerte werden nach Addition eines geeigneten reziproken Gittervektors im Bereich $-\pi/a < k \leq \pi/a$ dargestellt.

Im dreidimensionalen Raum ist der Sachverhalt etwas komplizierter. Für ein kubisches Gitter läßt sich (7.62) in der Form

$$E_{\boldsymbol{k}} = \frac{\hbar^2}{2m}|\boldsymbol{k}+\boldsymbol{G}|^2 = \frac{\hbar^2}{2m}\left[(k_x+G_x)^2+(k_y+G_y)^2+(k_z+G_z)^2\right] \qquad (7.63)$$

schreiben. In Abb. 7.17 ist diese Dispersionsrelation für ausgezeichnete Richtungen eines kubisch flächenzentrierten Gitters gezeichnet.

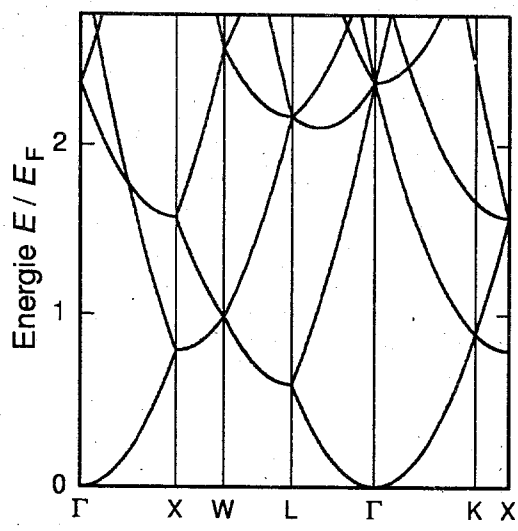


Abb. 7.17: E_k-Kurven von freien Elektronen in einem leeren Gitter mit der Symmetrie eines kubisch flächenzentrierten Kristalls nach der Reduktion auf die 1. Brillouin-Zone. Die Energie wurde auf die Fermi-Energie E_F eines Metalls mit drei freien Elektronen pro Atom normiert.

Wie bei den Phonon-Dispersionskurven in Kap. 5.4 startet man auch bei der Darstellung der Elektronendispersionskurven am Γ-Punkt, geht zunächst zur Grenze der Brillouin-Zone und bewegt sich dann längs der Begrenzung. Der Weg läßt sich leicht anhand der Abb. 2.11 und 5.14 nachvollziehen. Am X-, W- und L-Punkt findet man jeweils einen Knick in den Kurvenverläufen. Dies liegt daran, daß sich an diesen Punkten die Richtung des Weges ändert, längs dessen die Energieeigenwerte dargestellt sind. In Abb. 7.17 ist die Energie auf die Fermi-Energie E_F normiert, wobei wir angenommen haben, daß drei Elektronen pro Atom berücksichtigt werden müssen. Wie man sich anhand von Gleichung (7.63) leicht überzeugen kann, tritt bei der Reduktion auf die 1. Brillouin-Zone eine zunehmende Entartung der Dispersionskurven auf, d.h. gleiche Teile der E_k-Kurven gehören zu unterschiedlichen reziproken Gittervektoren.

Zum Vergleich mit den tatsächlichen Gegebenheiten ist in Abb. 7.18 der Verlauf der E_k-Kurven von Aluminium wiedergegeben, wobei zu beachten ist, daß sich die Energieskala von der in Abb. 7.17 unterscheidet. Die Übereinstimmung zwischen den beiden Abbildungen ist überraschend gut, denn man findet alle Teilkurven des freien Elektronengases des leeren Gitters bei den aus dem freien Elektronengas hergeleiteten Energien wieder. Offensichtlich ist die Näherung $\widetilde{V}_G \approx 0$ für Aluminium nicht schlecht. Bei näherem Hinsehen bemerkt man jedoch einen wichtigen qualitativen Unterschied: An den Grenzen der Brillouin-Zone treten *Energielücken* auf, die durch die endliche Stärke des Gitterpotentials bewirkt werden und die wir nun eingehend untersuchen wollen.

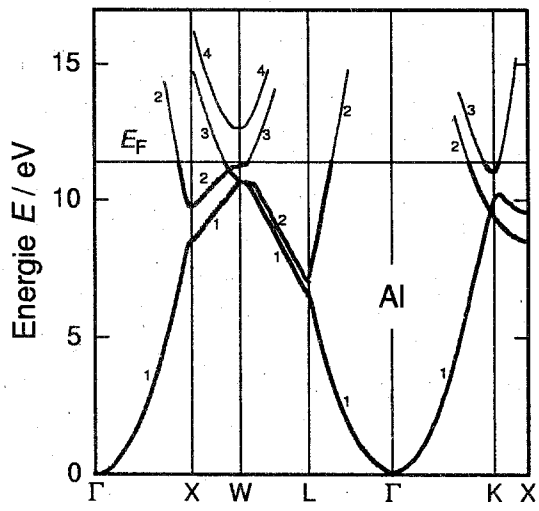


Abb. 7.18: E_k-Kurven von Aluminium. Die Zahlen an den Kurven geben die Ordnung der Brillouin-Zonen an, in denen die gekennzeichneten Teile der Dispersionskurve vor der Reduktion lagen. Die Fermi-Energie E_F ist ebenfalls eingezeichnet. (Nach B. Segall, Phys. Rev. **124** (1961) 1792).

Wir verlassen nun das leere Gitter und studieren den Einfluß eines schwachen, aber endlichen Potentials. Das Ergebnis wird sein, daß die einzelnen Gleichungen für die verschiedenen „Parabeln" nicht mehr wie beim leeren Gitter entkoppelt sind. Dadurch entstehen die soeben angesprochenen Lücken, die Zustände mit gleichen Wellenvektoren selbst an der Grenze der Brillouin-Zone voneinander trennen. Die Eigenzustände sind daher in *Energiebändern* untergebracht. Um das prinzipielle Verhalten herauszuarbeiten, diskutieren wir etwas eingehender die Lösung der Schrödinger-Gleichung für quasi-freie Elektronen mit dem Wellenvektor $k = g/2 = +\pi/a$. Wir wollen dabei herausfinden, welche Größen die Aufspaltung bestimmen und wodurch sich die Wellenfunktion bei diesem Wert des Wellenvektors von der eines freien Teilchens unterscheidet.

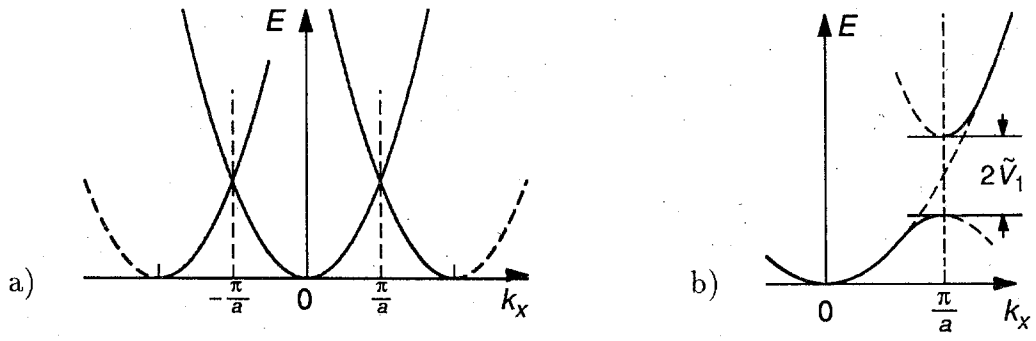


Abb. 7.19: Dispersionskurven von quasi-freien Elektronen.
a) Das periodische Potential ist so schwach, daß es vernachlässigt werden kann.
b) Das periodische Potential bewirkt bei $k_x = \pi/a$ die Aufspaltung $2\widetilde{V}_1$.

Wie der Abb. 7.19a zu entnehmen ist, schneiden sich bei $k_x = \pi/a$ zwei Parabeln, d.h. für das freie Elektronengas des leeren Gitters existieren zwei entartete Lösungen, die zu den Koeffizienten $c_{\boldsymbol{k}}$ und $c_{\boldsymbol{k}-\boldsymbol{g}}$ gehören. Ist die periodische Störung des freien Elektronengases schwach, so sollte der Hauptbeitrag zur Wellenfunktion ebenfalls von diesen beiden Lösungen stammen. Wellenfunktionen, die von „weiter entfernten" Parabeln herrühren, spielen aufgrund ihrer kleinen $c_{\boldsymbol{k}}$-Werte dann keine Rolle. Um dies zu zeigen, schreiben wir die Schrödinger-Gleichung (7.48) für den Wellenvektor $(\boldsymbol{k} - \boldsymbol{G})$ an, lösen nach dem Entwicklungskoeffizienten $c_{\boldsymbol{k}-\boldsymbol{G}}$ auf und finden

$$c_{\boldsymbol{k}-\boldsymbol{G}} = \frac{\sum_{\boldsymbol{G}'} \widetilde{V}_{\boldsymbol{G}'}\, c_{\boldsymbol{k}-\boldsymbol{G}-\boldsymbol{G}'}}{E - \frac{\hbar^2}{2m}|\boldsymbol{k}-\boldsymbol{G}|^2} = \frac{\sum_{\boldsymbol{G}''} \widetilde{V}_{\boldsymbol{G}''-\boldsymbol{G}}\, c_{\boldsymbol{k}-\boldsymbol{G}''}}{E - \frac{\hbar^2}{2m}|\boldsymbol{k}-\boldsymbol{G}|^2} \,. \tag{7.64}$$

Den zweiten Ausdruck erhalten wir durch Umbenennung der reziproken Gittervektoren gemäß $\boldsymbol{G}'' = \boldsymbol{G}' + \boldsymbol{G}$. Der Nenner weist für ein freies Elektronengas Nullstellen auf, wenn die Bedingung $k^2 = |\boldsymbol{k} - \boldsymbol{G}|^2$ erfüllt ist. Neben dem reziproken Gittervektor $\boldsymbol{G} = 0$ erfüllt für $k = \pi/a$ diese Bedingung auch der kleinste reziproke Gittervektor $\boldsymbol{G} \doteq \boldsymbol{g}$, wobei wir die Abkürzung $g = 2\pi/a$ benutzt haben. Für alle anderen Werte von $\boldsymbol{G}$ tritt keine Nullstelle auf, die Koeffizienten bleiben daher vergleichsweise klein und werden in unserer Diskussion vernachlässigt. Man spricht daher von einer „Zwei-Komponenten-Näherung". Bei dieser Vorgehensweise wird nur der erste Fourier-Koeffizient des periodischen Potentials berücksichtigt, d.h. nach (7.59), daß das Potential durch $\widetilde{V} = 2\widetilde{V}_1 \cos(\boldsymbol{g}\cdot\boldsymbol{r})$ gegeben ist. Da auf die Elektronen ein anziehendes Potential wirkt, ist der Fourier-Koeffizient $\widetilde{V}_1$ auf alle Fälle negativ. Dies bedeutet, daß das Potential am Ort der Atomrümpfe ein Minimum aufweist. Wie wir sehen werden, ist die erste Bandlücke in der hier diskutierten Näherung auf die Komponente $\widetilde{V}_1$ zurückzuführen. Für die Lücken bei höherer Energie sind Fourier-Komponenten höherer Ordnung verantwortlich.

Da in der Zwei-Komponenten-Näherung nur $\widetilde{V}_1 = \widetilde{V}_{\boldsymbol{g}} = \widetilde{V}_{-\boldsymbol{g}}$ berücksichtigt wird, ergeben sich aus (7.60) die beiden Gleichungen

$$\left(\frac{\hbar^2 k^2}{2m} - E\right) c_{\boldsymbol{k}} + \widetilde{V}_1 c_{\boldsymbol{k}-\boldsymbol{g}} = 0 \,,$$

$$\left(\frac{\hbar^2 |\boldsymbol{k}-\boldsymbol{g}|^2}{2m} - E\right) c_{\boldsymbol{k}-\boldsymbol{g}} + \widetilde{V}_1 c_{\boldsymbol{k}} = 0 \,, \tag{7.65}$$

die nur die beiden Fourier-Koeffizienten $c_{\boldsymbol{k}}$ und $c_{\boldsymbol{k}-\boldsymbol{g}}$ enthalten.

Die Energieeigenwerte des leeren Gitters, die wir mit dem Index 0 versehen, benutzen wir wie folgt als Abkürzungen:

$$E^0_{\boldsymbol{k}} = \frac{\hbar^2 k^2}{2m} \qquad \text{und} \qquad E^0_{\boldsymbol{k}-\boldsymbol{g}} = \frac{\hbar^2 |\boldsymbol{k}-\boldsymbol{g}|^2}{2m} \,. \tag{7.66}$$

Damit lassen sich die Eigenwerte der Gleichungen (7.65) in der folgenden Form schreiben:

$$E_{\mathrm{g,u}} = \frac{1}{2}(E^0_{\boldsymbol{k}-\boldsymbol{g}} + E^0_{\boldsymbol{k}}) \pm \sqrt{\frac{1}{4}(E^0_{\boldsymbol{k}-\boldsymbol{g}} - E^0_{\boldsymbol{k}})^2 + \widetilde{V}_1^2} \,. \tag{7.67}$$

Wie in Kap. 3.4 benutzen wir auch hier die Indizes g und u, da die Eigenwerte zu den Wellenfunktionen $\psi_{\boldsymbol{k},\mathrm{g}}$ und $\psi_{\boldsymbol{k},\mathrm{u}}$ mit gerader bzw. ungerader Parität gehören. Dabei ist das negative Vorzeichen E_g, das positive E_u zugeordnet.

An der Grenze der Brillouin-Zone ist $\boldsymbol{k} = \boldsymbol{g}/2$ und $E^0_{\boldsymbol{k}-\boldsymbol{g}} = E^0_{\boldsymbol{k}}$, wodurch sich aus (7.66) und (7.67) die einfachen Ausdrücke

$$E_{\mathrm{g,u}} = E^0_{\boldsymbol{k}} \pm |\widetilde{V}_1| \qquad \text{und} \qquad E_\mathrm{u} - E_\mathrm{g} = \delta E = 2|\widetilde{V}_1| \tag{7.68}$$

ergeben. Die Größe der Energieaufspaltung bei $k = \pi/a$ wird in dieser einfachen Näherung durch die erste Fourier-Komponente des Potentials $\widetilde{V}(\boldsymbol{r})$ bestimmt. E_u hat die höhere Energie und gehört daher zum zweiten Energieband. An dieser Stelle soll noch betont werden, daß die hier durchgeführte Rechnung nur für die Grenze der Brillouin-Zone bei $k = \pi/a$ gilt. Für $k = -\pi/a$ müssen die Koeffizienten $c_{\boldsymbol{k}}$ und $c_{\boldsymbol{k}+\boldsymbol{g}}$ berücksichtigt werden, die zu den gleichen Energiewerten führen.

Nun wollen wir uns noch überlegen, wodurch sich in der Umgebung der Grenze der Brillouin-Zone die Wellenfunktion $\psi = c_{\boldsymbol{k}} \exp(\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}) + c_{\boldsymbol{k}-\boldsymbol{g}} \exp[\mathrm{i}(\boldsymbol{k}-\boldsymbol{g})\cdot\boldsymbol{r}]$ der beiden Energiebänder unterscheidet. Hierzu bestimmen wir die Koeffizienten $c_{\boldsymbol{k}}$ bzw. $c_{\boldsymbol{k}-\boldsymbol{g}}$ beim Wellenvektor $\boldsymbol{k} = \boldsymbol{g}/2$ durch Einsetzen der beiden Energieeigenwerte E_g und E_u in (7.65). Wir finden

$$\begin{aligned} \frac{c_{\boldsymbol{k}-\boldsymbol{g}}}{c_{\boldsymbol{k}}} &= \frac{E - \hbar^2 k^2/2m}{\widetilde{V}_1} \,, \\ \frac{c_{-\boldsymbol{g}/2}}{c_{\boldsymbol{g}/2}} &= \frac{E^0_k \pm |\widetilde{V}_1| - E^0_k}{\widetilde{V}_1} = \pm 1 \,. \end{aligned} \tag{7.69}$$

Die zweite Zeile gibt das Verhältnis der Entwicklungskoeffizienten an der Grenze der Brillouin-Zone wieder, das die Werte ± 1 annimmt. Bei diesem Wellenvektor ist die Wellenfunktion also entweder eine symmetrische oder antisymmetrische Überlagerung der Einzelfunktionen. Da die erste Fourier-Komponente $\widetilde{V}_1$ immer negativ ist, ist beim unteren, ersten Band mit dem Energieeigenwert E_g das Verhältnis positiv. Somit wird die Gesamtwellenfunktion des tiefer liegenden Bandes durch die Summe der beiden Einzelfunktionen dargestellt. An der Grenze der Brillouin-Zone tragen beide Lösungen gleichermaßen zur Wellenfunktion bei. Entfernt man sich von diesem Punkt, so nimmt das Verhältnis der Entwicklungskoeffizienten ab. Wie in Abb. 7.20 zu sehen ist, überwiegt im ersten Band $c_{\boldsymbol{k}}$ und im zweiten $c_{\boldsymbol{k}-\boldsymbol{g}}$.

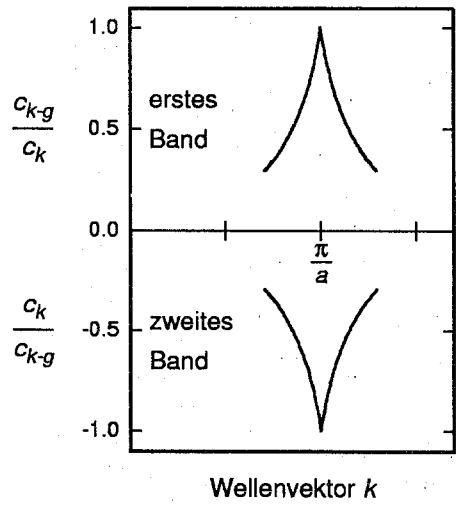


Abb. 7.20: Verhältnis der Entwicklungskoeffizienten c_k bzw c_{k-g} in der Umgebung von $k = \pi/a$. An der Zonengrenze sind beide Koeffizienten dem Betrag nach gleich. Entfernt man sich von der Grenze, so ändert sich ihr Verhältnis.

Setzen wir weiter vereinfachend konstante Entwicklungskoeffizienten voraus, d.h. nehmen wir ebene Wellen an, so gilt in der Zwei-Komponenten-Näherung in der unmittelbaren Umgebung von $k = g/2 = \pi/a$ für das 1. und 2. Band:

$$\psi_{k,\mathrm{g}}(x) = \frac{1}{\sqrt{2V}}\left[\mathrm{e}^{\mathrm{i}kx} + \mathrm{e}^{\mathrm{i}(k-g)x}\right] = \frac{1}{\sqrt{2V}}\left[\mathrm{e}^{\mathrm{i}gx/2} + \mathrm{e}^{-\mathrm{i}gx/2}\right] = \sqrt{\frac{2}{V}}\cos\left(\frac{\pi x}{a}\right),$$

$$\psi_{k,\mathrm{u}}(x) = \frac{1}{\sqrt{2V}}\left[\mathrm{e}^{\mathrm{i}kx} - \mathrm{e}^{\mathrm{i}(k-g)x}\right] = \frac{1}{\sqrt{2V}}\left[\mathrm{e}^{\mathrm{i}gx/2} - \mathrm{e}^{-\mathrm{i}gx/2}\right] = \sqrt{\frac{2}{V}}\,\mathrm{i}\sin\left(\frac{\pi x}{a}\right) \quad (7.70)$$

Dies sind stehende Wellen, für deren Auftreten sich eine anschauliche Erklärung finden läßt. Läuft eine Welle durch ein variables Potential, so wird sie ständig gestreut. Ist ihre Wellenlänge gerade gleich der doppelten Potentialperiode, so tritt „Bragg-Reflexion" auf. Durch die Überlagerung der ursprünglichen Welle mit der reflektierten bildet sich eine stehende Welle aus, bei der die Lage der Extrema und damit die Energie des Elektrons von der Phase bezüglich des periodischen Potentials abhängt. Für die Ladungsdichten der stehenden Wellen $\psi_{k,\mathrm{g}}$ und $\psi_{k,\mathrm{u}}$ erhalten wir

$$\varrho_{\mathrm{g}} = e|\psi_{k,\mathrm{g}}|^2 = \frac{2e}{V}\cos^2\left(\frac{\pi x}{a}\right) \qquad \text{und} \qquad \varrho_{\mathrm{u}} = e|\psi_{k,\mathrm{u}}|^2 = \frac{2e}{V}\sin^2\left(\frac{\pi x}{a}\right). \quad (7.71)$$

Im Gegensatz hierzu ist die Ladungsdichte einer laufenden Welle konstant:

$$\varrho = e|\psi|^2 = \frac{e}{V}\mathrm{e}^{-\mathrm{i}kx}\,\mathrm{e}^{\mathrm{i}kx} = \text{const.} \quad (7.72)$$

In Abb. 7.21 ist neben dem Potential der Verlauf der Elektronendichte schematisch dargestellt. Sie ist bei der Lösung $\psi_{\boldsymbol{k},\mathrm{g}}$ in Kernnähe erhöht, während sie bei $\psi_{\boldsymbol{k},\mathrm{u}}$ dort reduziert ist. In Analogie zur Atomphysik bezeichnet man die symmetrische Lösung $\psi_{\boldsymbol{k},\mathrm{g}}$ oft als s-artig, die antisymmetrische Lösung $\psi_{\boldsymbol{k},\mathrm{u}}$ als p-artig. Im Vergleich zur laufenden Welle führt dies zu einer Absenkung bzw. Erhöhung der Elektronenenergie. Das Bild macht das Ergebnis der formalen Diskussion plausibel, daß Lösungen der Schrödinger-Gleichung bei gleichem Wellenvektor unterschiedliche Energiewerte haben können.

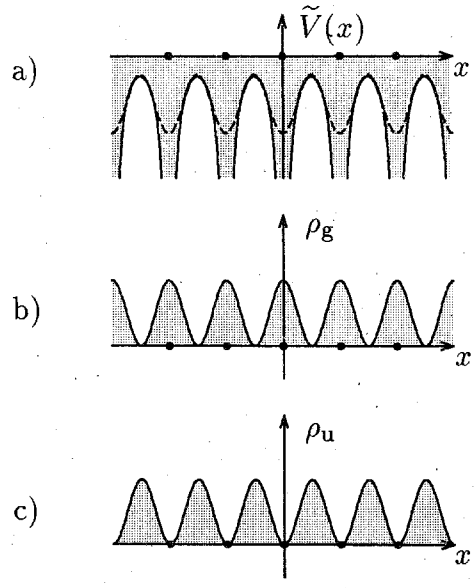


Abb. 7.21: Ladungsdichteverteilung der Elektronen mit $\boldsymbol{k} = \boldsymbol{g}/2$. Die Lage der Atomrümpfe ist jeweils durch Punkte gekennzeichnet.
a) Qualitativer Verlauf der potentiellen Energie. Die Überlagerung der Coulomb-Potentiale ist dick gezeichnet, gestrichelt ist der kosinusförmige Potentialverlauf der Zwei-Komponenten-Näherung.
b) Ladungsdichte ϱ_g der geraden Wellenfunktion.
c) Ladungsdichte ϱ_u der ungeraden Wellenfunktion.

Wie verläuft die $E_{\boldsymbol{k}}$-Kurve in der Nähe der Zonengrenze? Diese Frage läßt sich mit einer Entwicklung der Energieeigenwerte (7.67) um den Punkt $\boldsymbol{k} = \boldsymbol{g}/2$ beantworten, doch führen wir diese einfache, aber längliche Rechnung hier nicht durch. Wie man auch leicht erraten kann, ergibt sich ein Ausdruck der Form $E_{g,u} \propto |\delta \boldsymbol{k}|^2$, wobei $\delta \boldsymbol{k} = \boldsymbol{k} - \boldsymbol{g}/2$ als Abkürzung für die Abweichung des Wellenvektors von seinem Wert an der Grenze der Brillouin-Zone steht. Die Dispersionskurven $E_{\boldsymbol{k}}$ verlaufen in der Umgebung der Grenze der Brillouin-Zone parabelförmig. Diese Aussage gilt nicht nur in dem betrachteten speziellen Fall sondern für beliebige Bänder. In Abb. 7.22 ist der schematische Verlauf der $E_{\boldsymbol{k}}$-Kurven im periodisch wiederholten Zonenschema für den einfachen Fall eines eindimensionalen Gitters dargestellt.

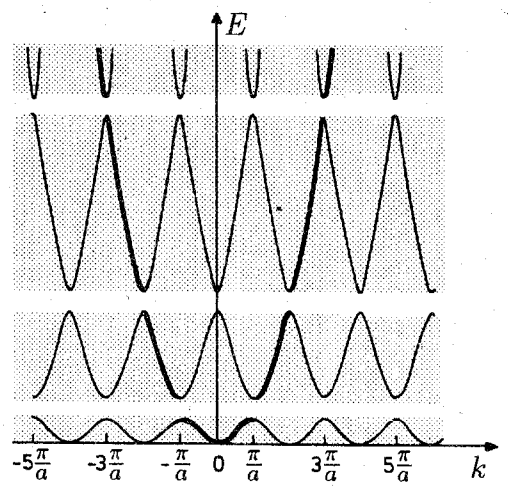


Abb. 7.22: E_k-Kurven eines eindimensionalen Gitters im periodisch wiederholten Zonenschema. Die erlaubten Energiebänder sind dunkel getönt, die verbotenen Bereiche hell gelassen. Die Teile der Kurve, die der Energieparabel eines freien Elektrons entsprechen, sind dicker gezeichnet.

Die Einführung eines periodischen Potentials in die Schrödinger-Gleichung führt in natürlicher Weise zu erlaubten Energiebändern und zu verbotenen Bereichen, die dazwischen liegen. Die hier durchgeführte Behandlung der Elektronen in Festkörpern ist besonders gut geeignet für die Beschreibung von Metallen. Ausgehend von freien Elektronen lassen sich die elektronischen Eigenschaften durch die Berücksichtigung von einer zunehmenden Zahl von Fourier-Koeffizienten der Potentialentwicklung (7.44) immer besser beschreiben. Im folgenden Abschnitt wollen wir aber einen anderen Weg gehen, der für kristalline Isolatoren und auch für amorphe Festkörper besser geeignet ist.

„Stark gebundene" Elektronen im Festkörper

Bisher sind wir davon ausgegangen, daß sich die Elektronen im Festkörper weitgehend frei bewegen können. Nun wollen wir das Problem vom entgegengesetzten Standpunkt aus beleuchten. Der Ausgangspunkt sind nun *quasi-isolierte Atome*, deren Elektronen stark an die Ionenrümpfe gebunden sind und für deren Wellenfunktion wir eine lineare Superposition der Atomeigenfunktionen ansetzen können. In dieser Näherung stellt der Kristall eine Anordnung von schwach wechselwirkenden Einzelatomen dar. Molekülkristalle sind ein Beispiel für die Anwendbarkeit dieser Betrachtungsweise. Der Einfachheit halber betrachten wir hier jedoch nur Systeme mit einatomiger Basis.

Wir nehmen an, daß die Schrödinger-Gleichung

$$H_{\mathrm{A}}\tilde{\psi}_i = E_i\tilde{\psi}_i \tag{7.73}$$

für isolierte Atome bereits gelöst sei und die Wellenfunktionen $\tilde{\psi}_i$ und die Eigenwerte E_i somit bekannt sind. H_{A} steht für den Hamiltonoperator der freien Atome, der Index i für die unterschiedlichen Energieniveaus. Zur Vereinfachung der Ausdrücke lassen wir den Index i zunächst weg und benutzen ihn erst wieder, wenn er erforderlich ist.

In unseren Überlegungen beschränken uns wieder auf eine *Einelektron-Näherung.* Zunächst stellen wir den Hamiltonoperator H eines Elektrons im Gesamtpotential aller Atome auf. Das Potential läßt sich in einen Anteil $\tilde{V}_{\mathrm{A}}$ aufspalten, der dem des freien Atoms entspricht, und in eine Störung H_{S} durch die Nachbaratome, die in der Umgebung der Aufatome klein gegen $\tilde{V}_{\mathrm{A}}$ sein soll[6]. Für den Hamiltonoperator können wir also

$$H = H_{\mathrm{A}} + H_{\mathrm{S}} = -\frac{\hbar^2}{2m}\Delta + \tilde{V}_{\mathrm{A}}(\boldsymbol{r} - \boldsymbol{R}_m) + H_{\mathrm{S}}(\boldsymbol{r} - \boldsymbol{R}_m) \tag{7.74}$$

schreiben. Die Lage des Bezugsatoms ist durch den Gittervektor $\boldsymbol{R}_m$ festgelegt.

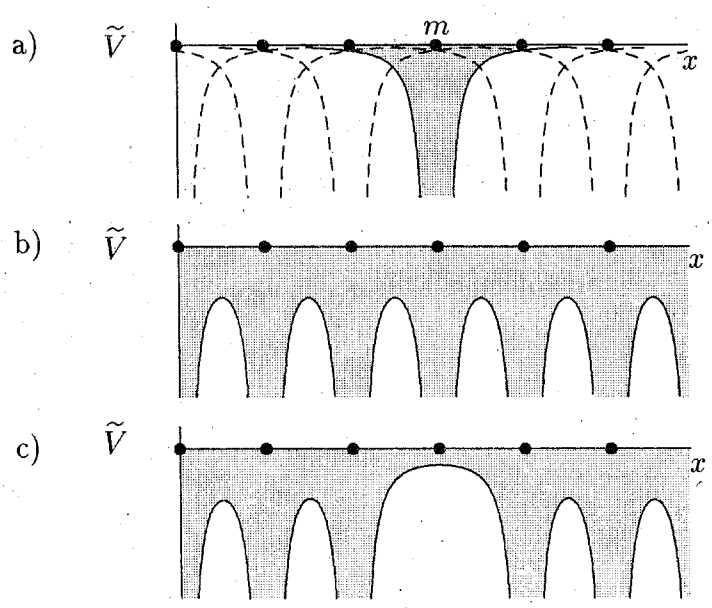


Abb. 7.23: Zum Potentialverlauf im Modell des stark gebundenen Elektrons.
a) Potential der Einzelatome. Das Potential des Bezugsatoms ist durch dunkle Tönung hervorgehoben.
b) Überlagerung der Einzelpotentiale zum Gitterpotential,
c) Resultierendes Störpotential.

[6] Da in der hier diskutierten Näherung die Nachbaratome nur das Potential am Aufatom verändern, wäre naheliegend hierfür $\tilde{V}_{\mathrm{s}}$ zu schreiben. Aus „Symmetriegründen" haben wir hier jedoch die Abkürzung H_{s} gewählt

Das Störpotential H_S (siehe Abb. 7.23) ist die Summe der atomaren Potentiale aller *anderen* Atome mit den Gittervektoren $\boldsymbol{R}_n \neq \boldsymbol{R}_m$:

$$H_S(\boldsymbol{r} - \boldsymbol{R}_m) = \sum_{n \neq m} \widetilde{V}_A(\boldsymbol{r} - \boldsymbol{R}_n) \,. \tag{7.75}$$

Mit Hilfe des Hamiltonoperators H und der Wellenfunktion $\psi_{\boldsymbol{k}}$ läßt sich der Energiewert $E'_{\boldsymbol{k}}$ des herausgegriffenen Elektrons berechnen:

$$E'_{\boldsymbol{k}} = \frac{\int \psi^*_{\boldsymbol{k}} \, H \, \psi_{\boldsymbol{k}} \, \mathrm{d}V}{\int \psi^*_{\boldsymbol{k}} \, \psi_{\boldsymbol{k}} \, \mathrm{d}V} \,. \tag{7.76}$$

Da die exakte Wellenfunktion $\psi_{\boldsymbol{k}}$ nicht bekannt ist, kann (7.76) nur näherungsweise gelöst werden. Wir ersetzen $\psi_{\boldsymbol{k}}$ in unserer Rechnung durch die *Näherungsfunktion* $\Phi_{\boldsymbol{k}}$, die wir durch Überlagerung der Atomeigenfunktionen bilden. In Kap. 3.4 wurde bereits darauf hingewiesen, daß der tatsächliche Eigenwert $E'_{\boldsymbol{k}}$ immer kleiner als der genäherte Wert $E_{\boldsymbol{k}}$ ist, d.h. es gilt

$$E'_{\boldsymbol{k}} \leq E_{\boldsymbol{k}} = \frac{\int \Phi^*_{\boldsymbol{k}} \, H \, \Phi_{\boldsymbol{k}} \, \mathrm{d}V}{\int \Phi^*_{\boldsymbol{k}} \, \Phi_{\boldsymbol{k}} \, \mathrm{d}V} \,. \tag{7.77}$$

Der richtige Eigenwert kann daher durch das Minimum des obigen Ausdrucks nur angenähert werden. Wir wollen hier diese Rechnung nicht tatsächlich durchführen sondern nur die Vorgehensweise diskutieren. Als Näherungsfunktion $\Phi_{\boldsymbol{k}}$ setzen wir eine lineare Überlagerung der atomaren Eigenfunktionen $\widetilde{\psi}(\boldsymbol{r} - \boldsymbol{R}_m)$ der Form

$$\psi_{\boldsymbol{k}} \approx \Phi_{\boldsymbol{k}} = \sum_m a_m \widetilde{\psi}(\boldsymbol{r} - \boldsymbol{R}_m) \tag{7.78}$$

an. Entsprechend dem Bloch-Theorem muß bei einem periodischen Potential jede Wellenfunktion $\psi_{\boldsymbol{k}}$ eine Bloch-Funktion sein, d.h. einen gitterperiodischen Anteil und einen Phasenfaktor enthalten. Die Gitterperiodizität ergibt sich automatisch durch die Überlagerung der Wellenfunktion $\widetilde{\psi}(\boldsymbol{r} - \boldsymbol{R}_m)$, also muß der Phasenfaktor in den Koeffizienten a_m zu finden sein. Wir setzen daher $a_m = \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{R}_m}$ und benutzen in der weiteren Rechnung die Näherungsfunktion

$$\Phi_{\boldsymbol{k}} = \sum_m \widetilde{\psi}(\boldsymbol{r} - \boldsymbol{R}_m) \, \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{R}_m} \,. \tag{7.79}$$

Es soll hier betont werden, daß der Ansatz einer gitterperiodischen Funktion nicht zwingend erforderlich ist, doch vereinfachen sich damit die weiteren Berechnungen ganz wesentlich.

Wir gehen mit dieser Funktion nun in (7.77) ein und betrachten Zähler und Nenner separat. Der Nenner stellt die Normierung der Wellenfunktion dar und hat die Form:

$$\int \Phi^*_{\boldsymbol{k}} \, \Phi_{\boldsymbol{k}} \, \mathrm{d}V = \sum_{m,n} \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot(\boldsymbol{R}_m - \boldsymbol{R}_n)} \int \widetilde{\psi}^*(\boldsymbol{r} - \boldsymbol{R}_n) \, \widetilde{\psi}(\boldsymbol{r} - \boldsymbol{R}_m) \, \mathrm{d}V \,. \tag{7.80}$$

Wie eingangs erwähnt, nehmen wir an, daß die Elektronen stark lokalisiert sind; dann ist die Wellenfunktion $\Phi_{\boldsymbol{k}}$ nur am Bezugsatom merklich von Null verschieden. Bei der Berechnung der Normierung berücksichtigen wir daher den Beitrag der Nachbarn am

Aufatom nicht, d.h. wir setzen $m = n$. Da die Wellenfunktion der isolierten Atome auf Eins normiert ist, erhalten wir schließlich

$$\int \Phi_{\boldsymbol{k}}^{*}\, \Phi_{\boldsymbol{k}}\, \mathrm{d}V \approx \sum_{m=0}^{N} \int \widetilde{\psi}^{*}(\boldsymbol{r} - \boldsymbol{R}_m)\widetilde{\psi}(\boldsymbol{r} - \boldsymbol{R}_m)\, \mathrm{d}V = N\,, \tag{7.81}$$

da die Summation über die N Atome des Kristalls läuft. Unter Berücksichtigung dieses Ergebnisses können wir für die Energieeigenwerte

$$E_{\boldsymbol{k}} \approx \frac{1}{N} \sum_{m,n} \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot(\boldsymbol{R}_m - \boldsymbol{R}_n)} \int \widetilde{\psi}^{*}(\boldsymbol{r} - \boldsymbol{R}_n)\, [H_{\mathrm{A}} + H_{\mathrm{S}}(\boldsymbol{r} - \boldsymbol{R}_m)]\, \widetilde{\psi}(\boldsymbol{r} - \boldsymbol{R}_m)\, \mathrm{d}V \tag{7.82}$$

schreiben. Wir gehen zunächst auf die Auswertung des Integrals ein: Da bei isolierten Atomen kein Überlapp der Wellenfunktion mit den Nachbaratomen zu berücksichtigen ist, ist das Integral $\int \widetilde{\psi}^{*} H_{\mathrm{A}} \widetilde{\psi}\, \mathrm{d}V$ nur für $m = n$ auszuwerten, und wir erhalten den Eigenwert E, den wir als bekannt vorausgesetzt haben. Das Integral $\int \widetilde{\psi}^{*} H_S \widetilde{\psi}\, \mathrm{d}V$, das den Störhamiltonoperator $H_{\mathrm{S}}(\boldsymbol{r} - \boldsymbol{R}_m)$ enthält, spalten wir in zwei Teile auf. Mit der Abkürzung α bezeichnen wir die Verschiebung des Energieeigenwertes des Bezugsatoms aufgrund der Potentialänderung, die von den Nachbarn hervorgerufen wird:

$$\alpha = -\int \widetilde{\psi}^{*}(\boldsymbol{r} - \boldsymbol{R}_m) H_{\mathrm{S}}(\boldsymbol{r} - \boldsymbol{R}_m) \widetilde{\psi}(\boldsymbol{r} - \boldsymbol{R}_m)\, \mathrm{d}V\,. \tag{7.83}$$

Die zweite Abkürzung

$$\beta_n = -\int \widetilde{\psi}^{*}(\boldsymbol{r} - \boldsymbol{R}_n) H_{\mathrm{S}}(\boldsymbol{r} - \boldsymbol{R}_m) \widetilde{\psi}(\boldsymbol{r} - \boldsymbol{R}_m)\, \mathrm{d}V \tag{7.84}$$

steht für die Energieänderung, die durch den Überlapp der Wellenfunktion des Aufatoms mit den übrigen Atomen zustande kommt. Da die Rechnungen für die verschiedenen Eigenwerte E_i der Atome gelten, führen wir nun den Index i wieder ein und schreiben für die Energieeigenwerte

$$E_{\boldsymbol{k},i} \approx \frac{1}{N} \sum_{m,n} \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot(\boldsymbol{R}_m - \boldsymbol{R}_n)} (E_i - \alpha_i - \beta_{n,i}) = E_i - \alpha_i - \sum_{n} \beta_{i,n}\, \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot(\boldsymbol{R}_m - \boldsymbol{R}_n)}\,. \tag{7.85}$$

Im letzten Ausdruck wurde berücksichtigt, daß bei E_i und α_i alle Beiträge ausschließlich von der Wellenfunktion $\widetilde{\psi}_m$ herrühren. Wir setzen deshalb $m = n$, so daß die Summation über n entfällt. Die Summation über m führt zum Faktor N, der sich gerade mit der Normierung weghebt. Im Falle der Größe $\beta_{i,n}$ geht der Überlapp der Wellenfunktion mit den Wellenfunktionen der benachbarten Atomen explizit ein. Hier muß die Summation über n ausgeführt werden, wobei für stark gebundene und damit stark lokalisierte Elektronen nur die nächsten Nachbarn Berücksichtigung finden müssen.

Da Wellenfunktionen im allgemeinen nicht kugelsymmetrisch sind, ist ihr Überlapp mit Wellenfunktionen der Nachbaratome unterschiedlich stark ausgeprägt. Deshalb ergeben sich unterschiedliche Konstanten β, selbst wenn nur die nächsten Nachbarn in die Betrachtung einbezogen werden. Zur weiteren Vereinfachung nehmen wir deshalb kugelsymmetrische Zustände, d.h. s-Wellenfunktionen an. In diesem Falle hängt β nur vom Abstand und nicht von der Richtung ab und besitzt für jeden wechselwirkenden Nachbarn

den gleichen Wert. Betrachten wir als einfaches Beispiel ein kubisch primitives Gitter, in dem sechs nächste Nachbarn mit $\boldsymbol{R}_m - \boldsymbol{R}_n = (\pm a, 0, 0), (0, \pm a, 0)$, und $(0, 0, \pm a)$ auftreten. Für die Energieeigenwerte folgt dann aus (7.85)

$$E_{\boldsymbol{k},i} \approx E_i - \alpha_i - 2\beta_i \left[\cos(k_x a) + \cos(k_y a) + \cos(k_z a)\right] . \tag{7.86}$$

Durch die Wechselwirkung mit den Nachbarn ist aus dem diskreten Energieniveau E_i der isolierten Atome ein leicht abgesenktes Band der Breite $12\beta_i$ entstanden! In Abb. 7.24 ist der Übergang von diskreten Atomniveaus zu den Bändern in Festkörpern für den diskutierten Fall schematisch dargestellt.

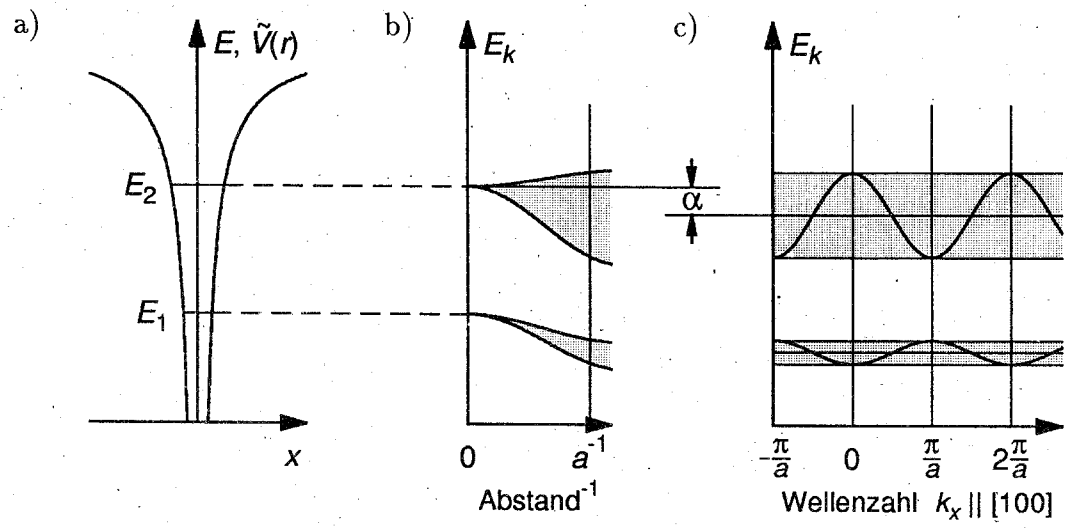


Abb. 7.24: Veranschaulichung des Modells stark gebundener Elektronen im Festkörper.
a) Lage der Niveaus E_1 und E_2 im Potential $\tilde{V}_A(r)$ isolierter Atome,
b) Absenkung und Aufspaltung der Niveaus abhängig vom reziproken Atomabstand,
c) Dispersionskurve E_k für Wellenzahlen k_x längs der [100]-Richtung

Einige wichtige Punkte sollen noch einmal betont werden. Da das Störpotential negativ ist und das Integral (7.83) nur die Wellenfunktion des Aufatoms enthält, ist die Größe α positiv. Sie beschreibt die mittlere Absenkung der Energieniveaus mit abnehmendem Atomabstand. Diese Absenkung wird durch die Potentialänderung hervorgerufen, die von den benachbarten Atomen am Ort des Aufatoms verursacht wird. Die Stärke des Überlapps der Wellenfunktionen der Nachbarn legt den Wert von β_i und damit die Breite des Bandes fest. Da Elektronen in tiefer liegenden Atomniveaus i stärker lokalisiert sind, „sehen" sie weniger von ihrer Umgebung, und die Bandaufspaltung ist entsprechend geringer. Dagegen kann es bei den Bändern, die von den äußeren Schalen herrühren, sogar zu einer Überlappung der Bänder kommen. Das Vorzeichen der Größe β_i bestimmt, ob am Γ-Punkt ein Maximum oder ein Minimum der Dispersionskurve vorliegt. Da in (7.84) explizit der Überlapp der Wellenfunktion mit den Nachbarn eingeht, hängt das Vorzeichen von der Struktur der Wellenfunktion selbst ab.

Die Gleichung (7.86) für die Dispersionsrelation des kubisch primitiven Gitters läßt sich am Γ-Punkt für kleine Wellenzahlen entwickeln. Mit der Abkürzung $k^2 = k_x^2 + k_y^2 + k_z^2$ für den Betrag des Wellenvektors erhält man

$$E_{\boldsymbol{k},i} \approx E_i - \alpha_i - 6\beta_i + \beta_i a^2 k^2 . \tag{7.87}$$

Sieht man vom Nullpunkt der Energie ab, so entspricht diese Dispersionsrelation der eines freien Elektronengases, d.h. Elektronen in kubischen Kristallen verhalten sich bei kleinen Wellenzahlen wie in einem isotropen Medium. Vergleicht man (7.87) mit der Dispersionsrelation $E = \hbar^2 k^2/2m$ für freie Elektronen, so sieht man, daß wir dem Elektron im Kristall die *effektive Masse*

$$m_i^* = \frac{\hbar^2}{2\beta_i a^2} \tag{7.88}$$

zuschreiben müssen. Je nach Vorzeichen von β_i kann die effektive Masse positiv oder negativ sein! Die effektive Masse ist proportional zu $1/\beta_i$ und nimmt daher mit abnehmender Bandbreite zu. Die physikalische Bedeutung dieser Aussage werden wir in Kap. 8.3 eingehend erläutern. Hier soll nur noch darauf hingewiesen werden, daß eine entsprechende Entwicklung auch im Maximum der Dispersionskurve möglich ist. Man findet dann gerade ein Vorzeichen der effektiven Masse, das dem von (7.88) entgegengesetzt ist.

Die Berechnung der Bandstruktur mit Hilfe des Modells stark gebundener Elektronen zeigt, daß das Auftreten von Energielücken nicht an die Periodizität des Gitters gebunden ist. Nur aus Gründen der mathematischen Einfachheit wurden im Ansatz für die Wellenfunktionen $\Phi_{\boldsymbol{k}}$ Bloch-Funktionen benutzt. Da nur Atome in der Nachbarschaft des Bezugsatoms in die Betrachtung einbezogen werden müssen, läßt sich die Rechnung auch für amorphe Materialien durchführen und führt zu ähnlichen Resultaten. Auf diese Tatsache soll ganz besonders hingewiesen werden, denn sie macht verständlich, warum auch in den nicht-periodisch aufgebauten amorphen Festkörpern elektronische Bänder auftreten.

Am Ende dieser Überlegungen zum Bändermodell wollen wir die Frage aufwerfen, ob unser Vorgehen, einzelne Energieniveaus isolierter Atome herauszugreifen und die Bandstrukturrechnung durchzuführen, nicht doch zu stark vereinfacht. Tatsächlich verlieren bei einer Wechselwirkung der Atome untereinander die Eigenzustände ihre Orthogonalität. Deshalb war es eigentlich nicht erlaubt den Index i vorübergehend wegzulassen und dann wieder einzuführen. Es zeigt sich jedoch, daß eine verbesserte Behandlung nur zu unerheblichen Korrekturen gegenüber der hier diskutierten Vorgehensweise führt.

Als Beispiele für die diskutierte Näherung sollen die Ergebnisse für den Ionenkristall KCl und die etwas komplizierteren Verhältnisse bei den tetraedrisch gebundenen Halbleitern gezeigt werden. Bei Ionenkristallen ist zu erwarten, daß die Aufspaltungen, d.h. die resultierenden Bandbreiten, im Vergleich zu den Niveauabständen sehr klein sind, da diese Kristalle abgeschlossene Elektronenschalen besitzen und ihre Wellenfunktionen kaum überlappen. Wie Abb. 7.25 für den Fall von Kaliumchlorid zeigt, wird diese Vorstellung durch die Rechnung bestätigt. Erst bei kleinen Ionenabständen wird die Bandaufspaltung merklich. So ist in der gezeigten Abbildung beim Gleichgewichtsabstand R_0 nur das oberste Niveau deutlich verbreitert.

Bei kovalenter Bindung ist der Überlapp der Wellenfunktionen groß. Dies führt zu ausgeprägten Bändern, wie das Beispiel eines tetraedrisch koordinierten Kristalls in Abb. 7.26 zeigt. Wie in Abschnitt 3.4 diskutiert, wird bei Annäherung der Atome aus den s- und p-Zuständen der sp^3-Hybrid geformt, der bindende und antibindende Niveaus aufweist (vgl. Kap. 3.4). Beim Gleichgewichtsabstand R_0 sitzen die vier pro Atom verfügbaren Elektronen im *Valenzband*, das durch eine Energielücke vom *Leitungsband* getrennt ist.

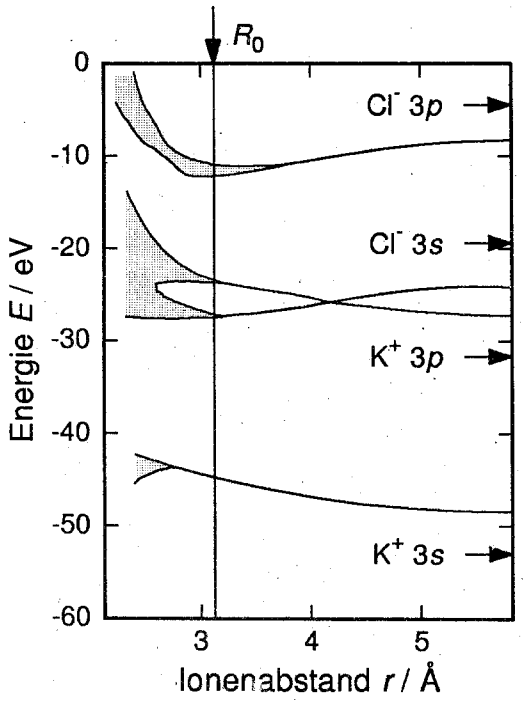


Abb. 7.25: Abstandsabhängigkeit der vier höchsten noch besetzten Energiebänder von KCl. Die Energie der Zustände der freien Ionen ist jeweils durch einen Pfeil markiert. Der Gleichgewichtsabstand $R_0 = 3{,}15$ Å ist ebenfalls eingetragen. (Nach L.P. Howland, Phys. Rev. **109** (1958) 1927).

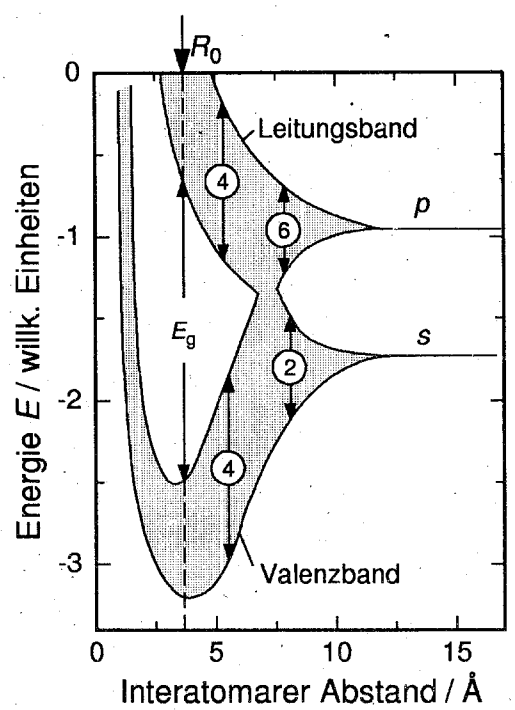


Abb. 7.26: Abstandsabhängigkeit der Energiebänder von Diamant. Die Zahlen in den Kreisen geben die Anzahl der besetzbaren Zustände pro Atom an. Der Gleichgewichtsabstand zwischen den Kohlenstoffatomen beträgt $R_0 = 3{,}57$ Å. (Nach A.H. Wilson, *The Theory of Metals*, Cambridge Univ. Press, London, 1965).

Durch Wechselwirkung werden weder Zustände erzeugt noch vernichtet. So findet man für N wechselwirkende Atome wie beim quasi-freien Elektronengas N Zustände pro Band. Wegen des Spinfreiheitsgrades kann jeder dieser Zustände jeweils zweifach besetzt werden. Dies bedeutet, daß in einem primitiven Gitter mit einem Valenzelektron pro Atom das oberste Band nur zur Hälfte besetzt ist. Im Fall der sp^3-Hybridbindung sind beim Gleichgewichtsabstand R_0 nur volle und leere Bänder vorhanden. Wie wir im nächsten Abschnitt sehen werden, ist für die Einordnung der Festkörper in Metalle oder Isolatoren der Besetzungsgrad der Bänder entscheidend. Materialien, bei denen nur vollbesetzte und leere Bänder vorliegen, sind nicht-leitend. Eine spezielle Gruppe hiervon sind die Halbleiter wie z.B. Silizium, bei denen die Energielücke relativ klein ist (vgl. Kap. 9). Sie sind bei tiefen Temperaturen ebenfalls Isolatoren.

Natürlich sind die Näherungen, die wir hier besprochen haben, für die verschiedenen Substanzen unterschiedlich gut geeignet. Tritt eine sehr starke Überlappung der Wellenfunktionen auf, wie dies bei den Metallen der Fall ist, so ist die Beschreibung mit Hilfe von quasi-freien Elektronen in einem periodischen Potential eine gute Näherung. Bei kovalenter Bindung oder bei Ionenkristallen, also immer dann, wenn die Valenzelektronen stark lokalisiert sind, ist die Näherung über stark gebundene Elektronen wesentlich geeigneter. Die reale Situation liegt oft zwischen stark gebundenen und quasi-freien Elektronen. Um dieser Tatsache Rechnung zu tragen, sind verbesserte Verfahren entwickelt worden, mit deren Hilfe sich die Bandstruktur realer Festkörper zuverlässig berechnen läßt.

7.4 Einfluß der Brillouin-Zone

Fermi-Flächen

Um die elektronischen Eigenschaften eines Kristalls verstehen zu können, müssen im Prinzip alle $E_{\boldsymbol{k}}$-Flächen, d.h. Flächen mit konstanter Energie, im $\boldsymbol{k}$-Raum bekannt sein. Von besonderer Bedeutung für die elektronischen Eigenschaften sind hierbei die Ober- und Unterkanten der Bänder, die beispielsweise bei der Diskussion von Halbleitereigenschaften in Kap. 9 eine große Rolle spielen. Bei den Metallen, die hier im Vordergrund stehen, kommt es hauptsächlich auf die Lage und Form der Fermi-Fläche an. Auf ihre Gestalt hat die Form der Brillouin-Zone einen wesentlichen Einfluß. Anhand eines zweidimensionalen, quadratischen Gitters, dessen ersten drei Brillouin-Zonen in Abb. 7.27 zu sehen sind, soll dieser Zusammenhang zwischen Struktur und elektronischen Eigenschaften dargelegt werden.

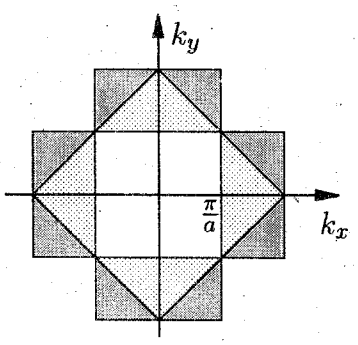


Abb. 7.27: Die ersten drei Brillouin-Zonen eines Quadratgitters. Das weiße Quadrat stellt die 1. Brillouin-Zone mit den Grenzen $k_x = k_y = \pi/a$ dar. Die 2. Brillouin-Zone ist leicht, die dritte stärker getönt.

Wir füllen nun den $\boldsymbol{k}$-Raum mit freien Elektronen auf. In unserem einfachen Beispiel ist die Fermi-Fläche eine Kugel mit dem Radius $k_{\mathrm{F}} = (3\pi^2 n)^{1/3}$, der durch die Dichte n der Elektronen bestimmt wird. Liegt die Fermi-Kugel vollständig innerhalb der 1. Brillouin-Zone, wie in Abb. 7.28a dargestellt, dann ist die Begrenzung durch die Brillouin-Zone von untergeordneter Bedeutung. Ist $k_{\mathrm{F}} > \pi/a$ wie in in Abb. 7.28b, dann ragt die Fermi-Kugel über die 1. Brillouin-Zone hinaus. Dies bedeutet, daß sich ein Teil der besetzten Zustände in der 2. Brillouin-Zone und damit im nächsthöheren Band befindet.

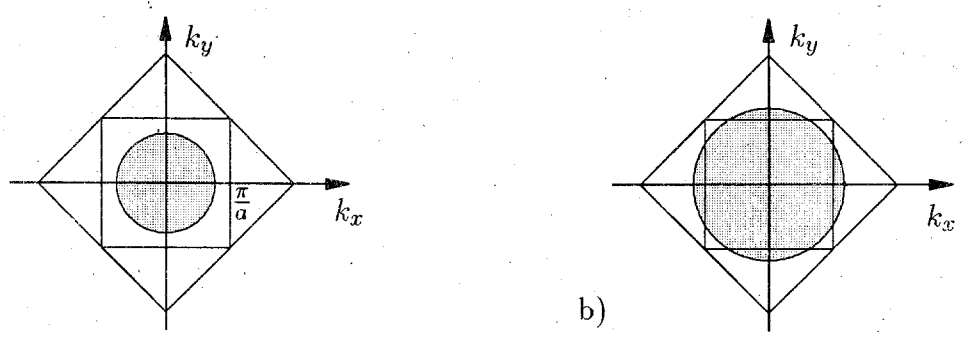


Abb. 7.28: Brillouin-Zonen eines Quadratgitters und Fermi-Kugel freier Elektronen.
a) Die Fermi-Kugel liegt innerhalb der 1. Brillouin-Zone, da $k_{\mathrm{F}} < \pi/a$ ist.
b) Die Fermi-Kugel ragt in die 2. Brillouin-Zone hinein.

Die Reduktion auf die 1. Brillouin-Zone führt zu Abb. 7.29. Die 1. Brillouin-Zone ist noch nicht vollständig mit Elektronen gefüllt, denn in ihren „Ecken" sind noch freie Plätze. Die Elektronen aus der 2. Brillouin-Zone bilden nun ein teilweise besetztes 2. Band, dessen Lage an den Rändern der Brillouin-Zone in Abb. 7.29b zu sehen ist. Da die Wellenfunktionen von Elektronenzuständen in unterschiedlichen Bändern aufeinander orthogonal stehen, „sehen" sich die Elektronen nicht, selbst wenn sie nach der Rückfaltung die gleichen Wellenvektoren besitzen.

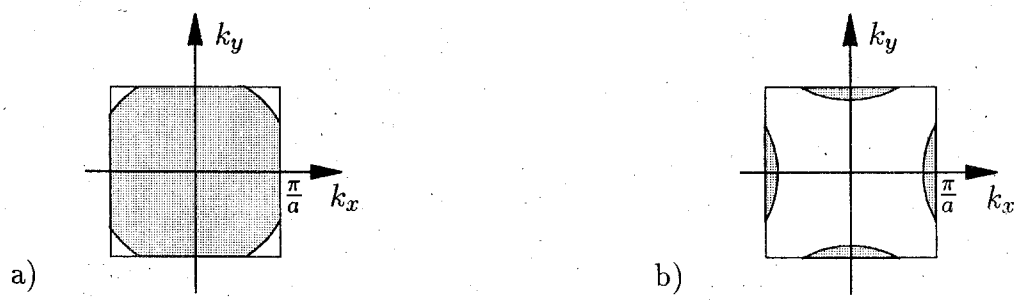


Abb. 7.29: Darstellung einer Fermi-Kugel mit $k_F > \pi/a$ im reduzierten Zonenschema. Man beachte, daß sich die dunkel getönten Elektronen in den beiden Bildern in unterschiedlichen Bändern befinden.
a) Die 1. Brillouin-Zonen ist nur teilweise mit Elektronen gefüllt.
b) Nach der Rückfaltung der Elektronen der 2. Brillouin-Zone sind „Elektronentaschen" entstanden.

In einem realen Metall mit „fast freien" Elektronen tritt aufgrund der Wechselwirkung der Elektronen mit dem Gitter an der Grenze der Brillouin-Zone eine Energielücke auf. Dies bedeutet, daß, wie in Abb. 7.30 dargestellt, die Oberfläche der Fermi-Kugel in der Nähe der Begrenzung deformiert ist. Die $E_{\boldsymbol{k}}$-Flächen treffen nun senkrecht auf die Begrenzung der Brillouin-Zone. Die Verbiegung der Fermi-Fläche in der hier skizzierten Form ermöglicht den Einbau einiger zusätzlicher Elektronen in die 1. Brillouin-Zone. Da die Elektronen dort eine kleinere Energie als in der nächsthöheren Brillouin-Zone besitzen, wird durch die Wechselwirkung mit dem Gitter die Gesamtenergie der Elektronen abgesenkt. In der Abbildung besitzen der Kreis und die dunkel gezeichneten Bereiche die gleiche Fläche, da durch die Wechselwirkung der Elektronen mit dem Gitter nicht die Zustandsdichte im $\boldsymbol{k}$-Raum beeinflußt wird.

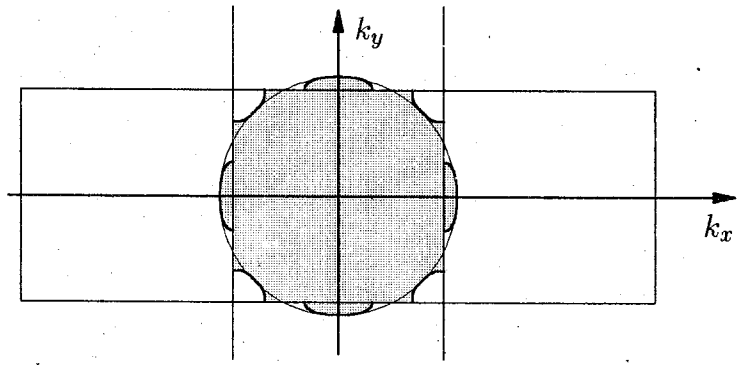


Abb. 7.30: Fermi-Fläche unter Berücksichtigung der Wechselwirkung der Elektronen mit dem Gitter. An der Begrenzung der quadratischen Brillouin-Zone tritt eine Lücke auf.

Als Beispiel für das Aussehen der Fermi-Fläche eines dreidimensionalen Systems betrachten wir ein Metall mit kubisch flächenzentrierter Struktur und drei freien Elektronen. Infolge der komplizierten Struktur der verschiedenen Brillouin-Zonen (vgl. Abb. 7.31) ist die Oberfläche der Fermi-Kugel keine einfach zusammenhängende Fläche. Wie in Abb. 7.32 zu sehen ist, liegen bei einem dreiwertigen Metall die besetzten Zustände der Elektronen in den ersten vier Brillouin-Zonen. Die 1. Brillouin-Zone ist vollständig mit Elektronen gefüllt. Die Zustände in den folgenden drei sind nur noch zum Teil besetzt. Um eine Darstellung mit einer zusammenhängenden Fermi-Fläche zu erhalten, wurde in dieser Abbildung die 3. Brillouin-Zone so zusammengesetzt, daß der X-Punkt im Zentrum der Darstellung liegt. In Abb. 7.17 war bereits die Dispersionskurve für ein freies Elektronengas in einem kubisch flächenzentrierten Gitter in ausgezeichneten Richtungen gezeigt worden. Ein Vergleich mit den Kurven von Aluminium in Abb. 7.18 zeigt, daß die Wechselwirkung der Elektronen mit den Rumpfpotentialen und die damit verbundene Deformation der Fermi-Kugel eine Entleerung der 4. Brillouin-Zone bewirkt. Die kleinen „Zwickel" bzw. Punkte verschwinden, diese Brillouin-Zone ist daher im realen Aluminium leer.

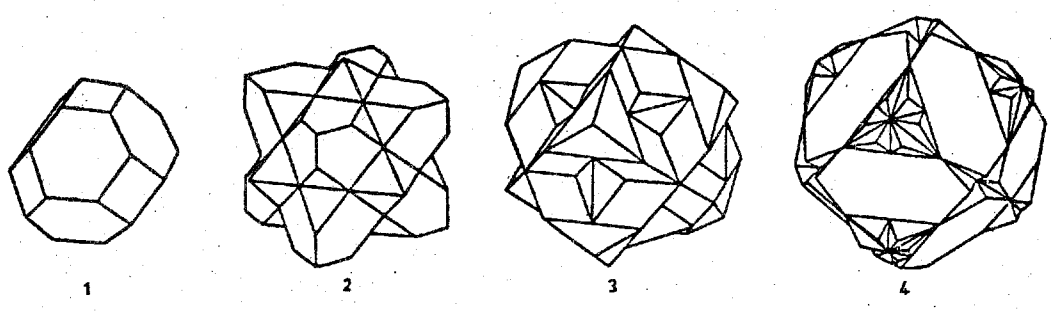


Abb. 7.31: Die ersten vier Brillouin-Zonen eines kubisch flächenzentrierten Gitters.

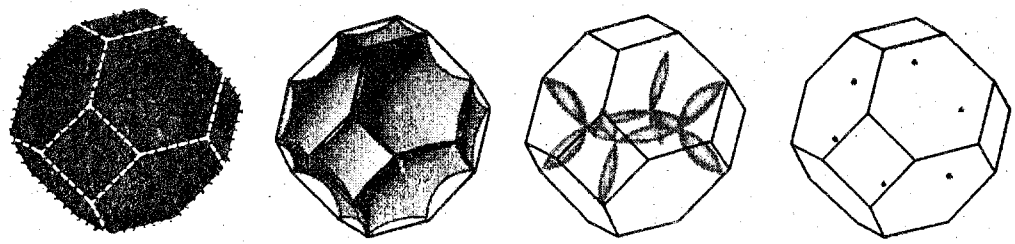


Abb. 7.32: Fermi-Flächen in den ersten vier Brillouin-Zonen eines dreiwertigen Metalls, dargestellt im reduzierten Zonenschema. Die Zustände der höheren Brillouin-Zonen sind nur teilweise mit Elektronen besetzt. Es ist zu beachten, daß in der zweiten Brillouin-Zone die Zustände am Rand besetzt und die im Innern frei sind. Bei der Darstellung der 3. Brillouin-Zone liegt der X-Punkt im Zentrum des Polyeders.

Besetzung der Bänder, Zustandsdichte

Die Bänder werden entsprechend der Fermi-Dirac-Statistik von unten mit Elektronen aufgefüllt. Volle Bänder tragen, wie wir in Kap. 8 herleiten werden, nicht zur elektrischen Leitfähigkeit bei. Deswegen sind Festkörper mit vollen elektronischen Bändern Isolatoren. Dies ist im Bild des freien Elektronengases unverständlich. Wir wollen aber im Vorgriff diese Kenntnis verwenden und qualitativ einige typische Bandstrukturen diskutieren.

Pro Band haben bei Kristallen mit einatomiger Basis jeweils $2N$ Elektronen Platz. Sind die Bänder innerhalb der gesamten Brillouin-Zone vollständig getrennt, und sind zwei Valenzelektronen vorhanden (Abb. 7.33a), so ist das *Valenzband* gefüllt. Das Band darüber, das *Leitungsband*, ist leer. Kristalle mit dieser Band- und Elektronenkonfiguration sind Isolatoren. Ähnlich liegen die Verhältnisse im Falle der Halbleiter Silizium oder Germanium. Da sich zwei vierwertige Atome in der Elementarzelle befinden, stehen bei diesen Materialien acht Valenzelektronen pro Elementarzelle zur Verfügung. Der in Abb. 7.26 dargestellte schematische Verlauf der Bandaufspaltung als Funktion des Abstands zeigt, daß das Valenzband mit $2 \cdot 4$ Elektronen (bzw. mit vier Elektronen pro Atom) vollständig gefüllt, das Leitungsband also leer ist. Am absoluten Nullpunkt sind diese Materialien daher ebenfalls Isolatoren. Ihre halbleitenden Eigenschaften, die wir in Kap. 9 ausführlich diskutieren werden, beruhen auf der Kleinheit ihrer Energielücke.

In der Skizze 7.33b ist angenommen, daß der repräsentierte Kristall drei Valenzelektronen besitzt. Dies bedeutet, daß sein Valenzband mit zwei Elektronen pro Atom voll besetzt ist. Die Elektronen in diesem Band tragen nicht zur Leitfähigkeit bei. Das darüber liegende Leitungsband ist mit einem Elektron pro Atom halb gefüllt. Festkörper mit einer derartigen Elektronenzahl und Bandstruktur sind Metalle. Noch einfacher liegen die Verhältnisse bei den Alkali- und Edelmetallen, die nur ein Valenzelektron aufweisen, so daß das Leitungsband ebenfalls nur halb gefüllt ist.

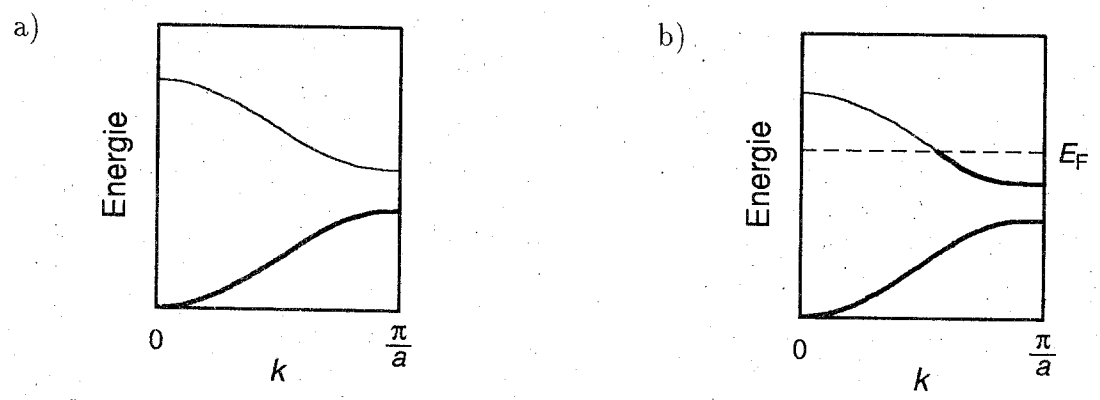


Abb. 7.33: Zur Besetzung der Bänder von (eindimensionalen) Kristallen mit einatomiger Basis.
a) Das Valenzband ist mit zwei Valenzelektronen pro Atom voll besetzt, das Leitungsband ist leer. Es handelt sich daher um einen Isolator.
b) Es wird angenommen, daß drei Valenzelektronen pro Atom zur Verfügung stehen. Das Leitungsband des Metalls ist zur Hälfte gefüllt.

Auf den ersten Blick ist erstaunlich, daß es Erdalkalielemente mit metallischen Eigenschaften gibt, denn sie besitzen 2 Elektronen pro Elementarzelle. Man würde aufgrund der bisherigen Argumentation erwarten, daß diese Elemente Isolatoren sind. Bei der Übertragung von Betrachtungen von ein- auf dreidimensionalen Systemen muß man jedoch Vorsicht walten lassen. Oft unterschlagen eindimensionale Modelle Aspekte, die für das Verständnis von realen Kristallen entscheidend sind. Wie in Abb. 7.34 angedeutet, kann die $E_{\boldsymbol{k}}$-Fläche für verschiedene Kristallrichtungen einen ganz unterschiedlichen Verlauf aufweisen. Dadurch können sich Bänder überlappen, die, in wohldefinierte Richtungen betrachtet, getrennt sind. Da die Fermi-Energie für alle Bänder die gleiche vorgegebene Größe ist, wird in diesem Beispiel das Band 3 teilweise entleert, und es werden Zustände in Band 4 besetzt. Ist der Überlapp nur schwach, so stehen nur wenige Elektronen für den Stromtransport zur Verfügung, und man spricht von *Halbmetallen*.

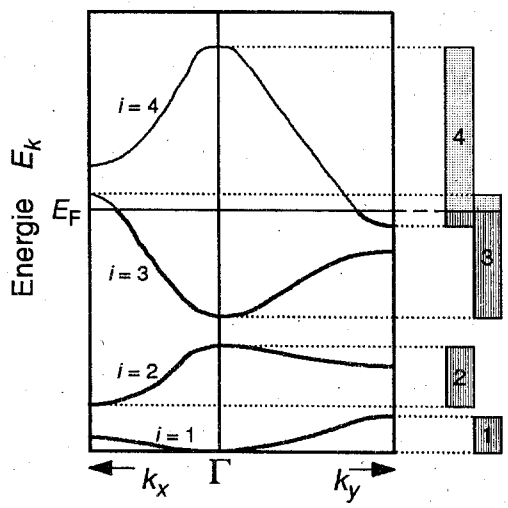


Abb. 7.34: Richtungsabhängigkeit der elektronischen Dispersion. Mit k_x und k_y sind die Wellenvektoren in unterschiedliche Richtungen bezeichnet. Band 3 und 4 überlappen. Sie sind bis zur gemeinsamen Fermi-Energie mit Elektronen besetzt.

Für viele Betrachtungen, insbesondere bei Metallen, genügt die Kenntnis der Zustandsdichte $D(E)$ anstelle der Gesamtinformation, die in den $E_{\boldsymbol{k}}$-Flächen steckt. Nach (7.6) läßt sich die Zustandsdichte mit Hilfe der Gleichung

$$D(E)\,\mathrm{d}E = \frac{2}{(2\pi)^3}\int_{E}^{E+\mathrm{d}E} \mathrm{d}^3k \tag{7.89}$$

durch Integration der $E_{\boldsymbol{k}}$-Kurven errechnen. Wie bei den Phononen kommen die wichtigsten Beiträge zur Zustandsdichte von Bereichen im $\boldsymbol{k}$-Raum, in denen die Energieflächen waagerecht verlaufen. Es treten daher ebenfalls van Hove-Singularitäten oder kritische Punkte auf. Die starke Korrelation zwischen Maxima in der Zustandsdichte und Extrema in der Bandstruktur macht Abb. 7.35 deutlich. Gut zu erkennen ist der bereits öfter angesprochene unterschiedliche Verlauf der Zustandsdichte von s- und d-Elektronen.

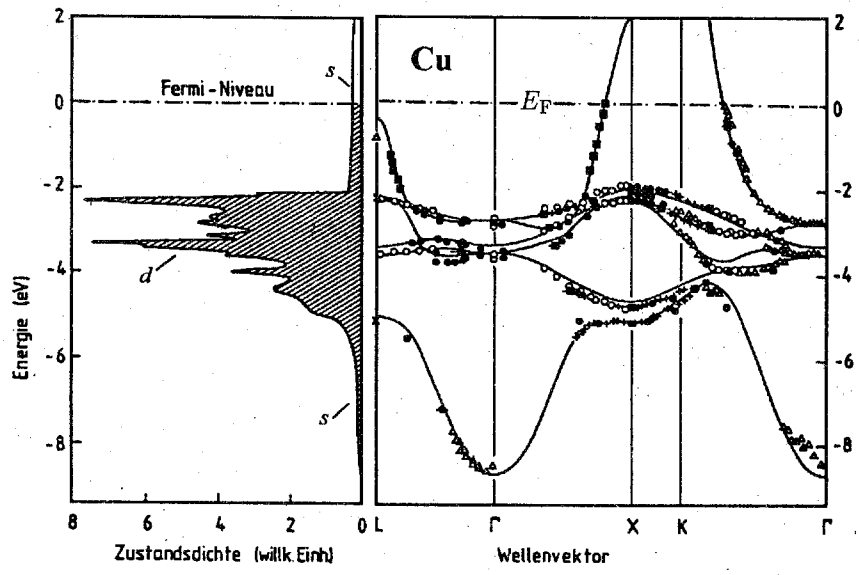


Abb. 7.35: Zustandsdichte und Bandstruktur von Kupfer. Beide Größen wurden mit Hilfe der Photoemissionsspektroskopie gemessen. (Nach H. Ibach, H. Lüth, *Festkörperphysik*, Springer, Heidelberg 1990).

Experimentell lassen sich Zustandsdichten zum Beispiel mit Hilfe der *Photoemissionsspektroskopie* messen. Hierbei wird der Festkörper mit ultraviolettem Licht bestrahlt und die Anzahl der austretenden Photoelektronen energieaufgelöst gemessen. Wie in Abb. 7.36 angedeutet, werden hierbei die Elektronen aus den besetzten Zuständen in quasi-kontinuierliche Zustände gehoben. Ist $\hbar\omega$ die Energie der einfallenden Photonen, Φ die Austrittsarbeit, E_{kin} die kinetische Energie der emittierten Elektronen und E_{b} deren Bindungsenergie bezüglich der Fermi-Energie, so muß die Energiebilanz $\hbar\omega = \Phi + E_{\mathrm{kin}} + E_{\mathrm{b}}$ gelten. Gemessen wird die Zahl der emittierten Elektronen $N(E_{\mathrm{kin}})$ als Funktion ihrer kinetischen Energie. Da die Anregungswahrscheinlichkeit durch ein Photon innerhalb eines Bandes im allgemeinen nur schwach von der energetischen Lage des Zustands abhängt, spiegelt die Zahl der registrierten Elektronen die Zustandsdichte wider. Dem Spektrum überlagert ist ein Untergrund von Elektronen, die bereits vor dem Austritt aus dem Festkörper inelastische Stöße erlitten haben.

Ohne näher darauf einzugehen soll noch erwähnt werden, daß sich mit Hilfe von winkelaufgelösten Messungen der Photoemission die elektronische Bandstruktur selbst ermitteln läßt. Ein Beispiel hierfür ist in Abb. 7.35b zu sehen.

Da die Elektronen aus den obersten 5 Å austreten, hängt das Meßergebnis stark von der Oberflächenbeschaffenheit ab. Damit wohldefinierte Verhältnisse vorliegen, werden derartige Messungen im Ultrahochvakuum durchgeführt. Typischerweise ist für die Anregung der Elektronen eine Energie über 20 eV und damit sehr kurzwelliges ultraviolettes Licht erforderlich. In diesem Spektralbereich absorbieren alle Festkörper sehr stark, so daß es keine geeigneten Fenster für den Durchtritt des Lichtes in die UHV-Apparatur gibt. Die anregenden Lichtquanten müssen deshalb in einem speziellen Teil der Vakuumkammer mit Hilfe von Gasentladung oder Lichtbogen bei höheren Drücken erzeugt werden.

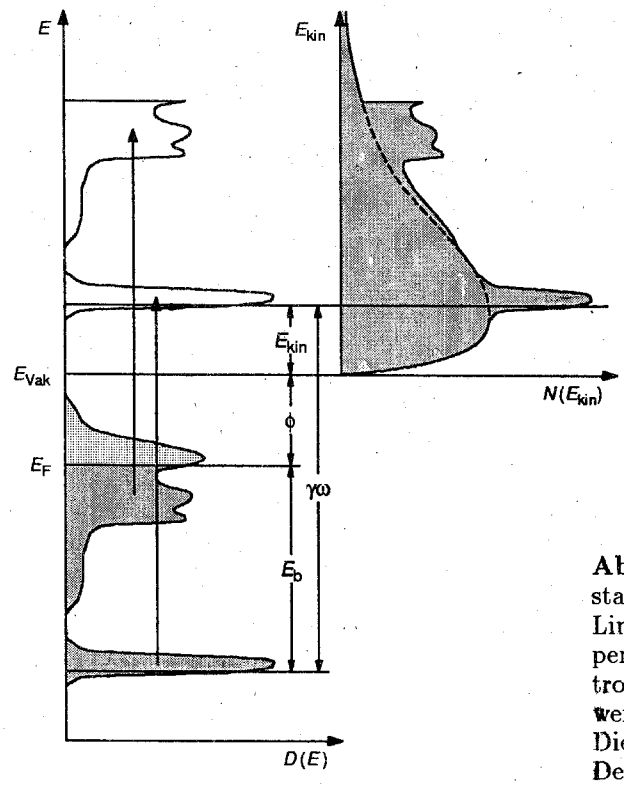


Abb. 7.36: Bestimmung der elektronischen Zustandsdichte durch Photoemissionsspektroskopie. Links sind die Zustandsdichte und die für das Experiment relevanten Energien dargestellt. Die Elektronen aus den dunkelgetönten besetzten Zuständen werden in quasi-kontinuierliche Zustände gehoben. Die rechte Seite zeigt das beobachtete Spektrum. Dem Meßsignal überlagert ist ein Untergrund, der von inelastisch gestreuten Elektronen herrührt.

8. Bewegung der Elektronen im Festkörper

8.1 Bewegungsgleichung und effektive Masse

Die Kenntnis der elektronischen Zustandsdichte, die Anwendung der Fermi-Statistik und die Berücksichtigung der Wechselwirkung der Elektronen mit dem periodischen Potential des Gitters ermöglichte die Erklärung einer Reihe von grundlegenden Festkörpereigenschaften, speziell bei Metallen. Nun wenden wir uns den Transporteigenschaften zu, bei denen die Bewegung der Elektronen durch den Festkörper von entscheidender Bedeutung ist. Wir werden sehen, daß das periodische Potential und die damit verbundene Bandstruktur erstaunliche Konsequenzen haben. Zwei wichtige Resultate wurden bereits im vorhergehenden Kapitel erwähnt: Ein vollbesetztes Band liefert keinen Beitrag zur elektrischen Leitfähigkeit, und die effektive Masse der Elektronen kann negativ sein. Neben dem Konzept der effektiven Masse untersuchen wir in diesem Kapitel den Begriff des „Loches" näher. Beide Größen werden im weiteren Verlauf der Diskussion von großer Bedeutung sein. Anschließend gehen wir auf die Transporteigenschaften am Beispiel der elektrischen und thermischen Leitfähigkeit ein. Eine Besonderheit stellt die Bewegung der Elektronen dar, wenn sich der betrachtete Festkörper im Magnetfeld befindet. Da sich Elektronen in Magnetfeldern auf Flächen konstanter Energie bewegen, erlauben derartige Experimente die Bestimmung der tatsächlichen Gestalt von Fermi-Flächen. Darüber hinaus werden wir sehen, daß Magnetfelder zur Quantisierung der Elektronenbahnen führen, also die Bewegungsfreiheit der Elektronen einschränken. Besonders deutlich wird dies beim Quanten-Hall-Effekt sichtbar.

Wellenpakete

Zur Beschreibung der Elektronenbewegung sind die über den ganzen Festkörper ausgedehnten Bloch-Funktionen nicht geeignet, denn Elektronen können nicht gleichzeitig im Orts- und Impulsraum streng lokalisiert sein. Da die Ungleichung $\delta k\, \delta r > 1$ immer erfüllt sein muß, können der Ortsvektor $\boldsymbol{r}$ und der Wellenvektor $\boldsymbol{k}$ nicht gleichzeitig exakt festgelegt werden. Eine begrenzte räumliche Lokalisierung der Wellenfunktion läßt sich durch die Überlagerung von Zuständen mit unterschiedlichem Wellenvektor erreichen. Diese Vorgehensweise ist wohlbekannt bei der Wellenfunktion $\psi(\boldsymbol{r}, t)$ *freier* Elektronen, die sich durch Überlagerung von ebenen Wellen als Wellenpaket darstellen lassen:

$$\psi(\boldsymbol{r}, t) = \sum_{\boldsymbol{k}'} g(\boldsymbol{k}')\, \mathrm{e}^{\mathrm{i}(\boldsymbol{k}'\cdot\boldsymbol{r} - \hbar k'^2 t/2m)} \,. \tag{8.1}$$

Sind die Entwicklungskoeffizienten $g(\boldsymbol{k}')$ innerhalb eines Intervalls δk geeignet, z.B. gaußförmig, verteilt, so führt dies zu der gewünschten räumlichen Lokalisierung. Genauso können wir bei der Beschreibung von lokalisierten Elektronen im Festkörper vorgehen.

Anstelle der ebenen Wellen treten nun Bloch-Wellen, und wir erhalten statt (8.1) den Ausdruck

$$\psi^n(\boldsymbol{r},t) = \sum_{\boldsymbol{k}'} g(\boldsymbol{k}')\,\psi^n_{\boldsymbol{k}'}(\boldsymbol{r})\,\mathrm{e}^{-\mathrm{i}E^n_{k'}t/\hbar}\,, \tag{8.2}$$

wobei $\psi^n_{\boldsymbol{k}'}(\boldsymbol{r})$ für die Bloch-Funktionen steht. Der Index n bezeichnet das Energieband, in dem die beschriebenen Zustände liegen. Gilt für $g(\boldsymbol{k}')$ wieder die gleiche Bedingung wie bei der Darstellung der freien Elektronen durch ein Wellenpaket, so wird auch hier eine Lokalisierung erreicht.

Unter diesen Voraussetzungen erlauben Wellenpakete eine einfache und anschauliche Beschreibung der Elektronenbewegung mit Hilfe der *semiklassischen Bewegungsgleichung.* Da der Wellenvektor der Elektronen relativ gut definiert sein sollte, müssen die Wellenpakete so konstruiert werden, daß ihre Unschärfe im $\boldsymbol{k}$-Raum wesentlich kleiner als die Ausdehnung der Brillouin-Zone ist. Daher ist die semiklassische Beschreibung nur anwendbar, wenn die räumliche Ausdehnung des Wellenpaketes δx wesentlich größer als der Gitterabstand a ist, die Elektronen also über mehrere Elementarzellen ausgeschmiert sind. Gleichzeitig muß die Wellenlänge λ eines von außen angelegten Störfeldes groß gegen die Ausdehnung des Wellenpaketes sein, damit dieses wie ein kompaktes Teilchen auf das Feld reagiert. Die charakteristischen Größenverhältnisse sind in Abb. 8.1 schematisch dargestellt, wobei zu beachten ist, daß die tatsächlich auftretenden Unterschiede zwischen den Längenskalen wesentlich stärker ausgeprägt sind als die Zeichnung wiedergeben kann.

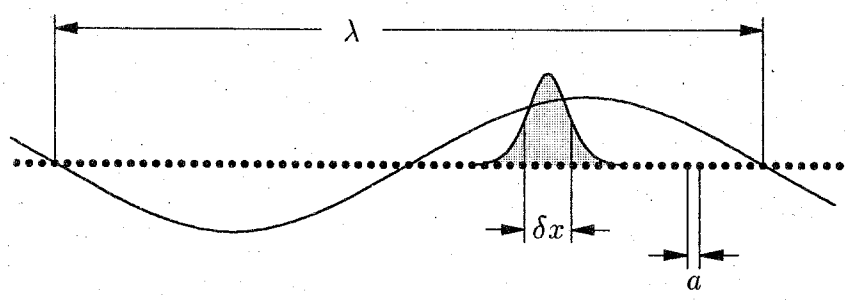


Abb. 8.1: Typische Längenskalen, die bei der semiklassischen Näherung auftreten. Es muß die Beziehung Gitterkonstante $a \ll$ Ausdehnung δx der Wellenpakete $\ll$ Wellenlänge λ der angelegten äußeren Störung gelten. Die Größenverhältnisse sind nicht maßstabsgerecht wiedergegeben.

Semiklassische Bewegungsgleichung und effektive Masse

Die semiklassische Näherung beschreibt die zeitliche Entwicklung der Position $\boldsymbol{r}$ und des Wellenvektors $\boldsymbol{k}$ eines Elektrons in Gegenwart eines äußeren elektrischen oder magnetischen Feldes $\boldsymbol{\mathcal{E}}$ bzw. $\boldsymbol{\mathcal{B}}$. Die Geschwindigkeit, mit der sich das Wellenpaket und damit das Elektron bewegt, ist durch die Gruppengeschwindigkeit festgelegt, die aus der Dispersionsrelation $E^n(\boldsymbol{k})$ berechnet werden kann[1]:

$$\dot{\boldsymbol{r}} = \boldsymbol{v}^n(\boldsymbol{k}) = \frac{1}{\hbar}\nabla_{\boldsymbol{k}} E^n(\boldsymbol{k}) = \frac{1}{\hbar}\frac{\partial E^n(\boldsymbol{k})}{\partial \boldsymbol{k}}\,. \tag{8.3}$$

[1] Die Notation $E(\boldsymbol{k})$ soll andeuten, daß es sich nicht um die Eigenwerte $E_{\boldsymbol{k}}$ von Bloch-Funktionen handelt, sondern daß die Energie der Wellenpakete gemeint ist.

Die Geschwindigkeit ist in eindeutiger Weise mit der Energiefläche $E^n(\boldsymbol{k})$ verknüpft, weitere Parameter gehen nicht ein. Für ein freies Elektron, das sich auf der parabelförmigen Energiefläche $E(\boldsymbol{k}) = \hbar^2 k^2/2m$ bewegt, liefert diese Definition den geläufigen Zusammenhang $\boldsymbol{v} = \hbar\boldsymbol{k}/m = \boldsymbol{p}/m$ für die Geschwindigkeit.

Wirkt auf ein Elektron die Kraft $\boldsymbol{F}$, so ändert sich sein Wellenvektor und damit der *Quasiimpuls* $\hbar\boldsymbol{k}$ gemäß der Gleichung

$$\hbar\dot{\boldsymbol{k}} = \boldsymbol{F} = -e\Big[\boldsymbol{\mathcal{E}}(\boldsymbol{r},t) + \boldsymbol{v}^n(\boldsymbol{k}) \times \boldsymbol{B}(\boldsymbol{r},t)\Big] \,. \tag{8.4}$$

Diese Gleichung ist von fundamentaler Bedeutung für die Beschreibung der Bewegung von Kristallelektronen. Für ein freies Elektron ist auch diese Gleichung wohlbekannt. Ihre Herleitung für den Fall, daß das Elektron mit dem Gitter wechselwirkt, ist jedoch aufwendig und findet sich in Lehrbüchern der theoretischen Festkörperphysik. Wie bei den ursprünglich eingeführten Bloch-Funktionen ist der Wellenvektor nur bis auf einen reziproken Gittervektor $\boldsymbol{G}$ bestimmt, weswegen $\hbar\boldsymbol{k}$ nicht als Impuls, sondern als „Quasiimpuls" bezeichnet wurde. In der diskutierten Näherung ändert sich der Bandindex n unter dem Einfluß der äußeren Felder nicht, d.h. *Interbandübergänge* werden ausgeklammert. In der weiteren Diskussion werden wir daher den Index n weglassen.

Die Definition (8.3) der Gruppengeschwindigkeit läßt sich unter Benutzung von (8.4) umschreiben:

$$\dot{\boldsymbol{v}} = \frac{\partial}{\partial t}\left(\frac{1}{\hbar}\frac{\partial E(\boldsymbol{k})}{\partial \boldsymbol{k}}\right) = \frac{1}{\hbar}\frac{\partial}{\partial \boldsymbol{k}}\left(\frac{1}{\hbar}\frac{\partial E(\boldsymbol{k})}{\partial \boldsymbol{k}}\right)\hbar\dot{\boldsymbol{k}} = \frac{1}{\hbar^2}\frac{\partial^2 E(\boldsymbol{k})}{\partial \boldsymbol{k}\partial \boldsymbol{k}}\boldsymbol{F} \,. \tag{8.5}$$

Für die kartesischen Komponenten $\dot{v}_i$ der Geschwindigkeitsänderung erhalten wir:

$$\dot{v}_i = \frac{1}{\hbar^2}\sum_j \frac{\partial^2 E(\boldsymbol{k})}{\partial k_i\, \partial k_j} F_j = \sum_j \left(\frac{1}{m^*}\right)_{ij} F_j \qquad \text{mit} \qquad \left(\frac{1}{m^*}\right)_{ij} = \frac{1}{\hbar^2}\frac{\partial^2 E(\boldsymbol{k})}{\partial k_i\, \partial k_j} \,. \tag{8.6}$$

Damit wird die *effektive Masse* m^* definiert, mit deren Hilfe die Verbindung zur klassischen Bewegungsgleichung $\dot{\boldsymbol{v}} = \boldsymbol{F}/m$ hergestellt wird. Die reziproke effektive Masse ist durch die *Krümmung* der Energiefläche bestimmt. In ihr ist die Wechselwirkung der Elektronen versteckt, die sie bei ihrer Bewegung im periodischen Potential erfahren. Man spricht daher auch von der *dynamischen Masse*. Es ist bemerkenswert, daß die Kenntnis der $E(\boldsymbol{k})$-Fläche des Kristalls zur Beschreibung der Elektronenbewegung ausreicht, und keine weiteren Einzelheiten des Kristallpotentials eingehen. Das hier entwickelte Konzept zur Beschreibung der Elektronenbewegung im Kristall wird als *effektive Massennäherung* bezeichnet.

Im Gegensatz zur freien Masse besitzt die effektive Masse aufgrund der Richtungsabhängigkeit der Wechselwirkung Tensorcharakter, d.h., die Beschleunigung erfolgt im allgemeinen nicht parallel zur Kraft. Darüber hinaus hängt die effektive Masse selbst vom Wellenvektor ab, so daß bei vorgegebener Feldstärke der Betrag der Beschleunigung für verschiedene Elektronen unterschiedlich ist. Aus dem Gesagten folgt, daß die Beziehung $\hbar\boldsymbol{k} = m\boldsymbol{v}$ nicht mehr gültig sein kann, denn sie setzt eine konstante, richtungsunabhängige Masse voraus.

Wie man der Definition (8.6) unmittelbar entnehmen kann, sind die Tensoren der reziproken Masse $(1/m^*)_{ij}$ und der Massentensor m^*_{ij} symmetrisch und lassen sich auf Hauptachsen transformieren. Damit reduziert sich die Anzahl der unabhängigen Komponenten auf höchstens drei. In besonders einfachen Fällen sind alle Komponenten gleich

groß, dann ist die effektive Masse $m^*(k) = \hbar^2/(\mathrm{d}^2E(k)/\mathrm{d}k^2)$ eine skalare Größe, und der Festkörper besitzt isotrope elektronische Eigenschaften, wenn auch m^* durchaus noch eine Funktion der Wellenzahl k sein kann. Eine von der Wellenzahl unabhängige effektive Masse tritt bei parabelförmigen Bändern auf. Diesen einfachen Fall trifft man häufig in der Umgebung von Bandextrema an, da sich dort die Energieflächen besonders gut durch Parabeln nähern lassen.

Der prinzipielle Zusammenhang zwischen effektiver Masse und Wellenvektor ist in Abb. 8.2 dargestellt. Von $k = 0$ ausgehend, nimmt die Bandkrümmung zunächst mit dem Wellenvektor ab und somit die effektive Masse zu. An den Wendepunkten der Dispersionskurve wird m^* unendlich groß und anschließend sogar negativ. Die effektive Masse der Elektronen ist also positiv in den Bandminima und negativ in den Maxima. In der Nähe der Extrema ist sie darüberhinaus näherungsweise konstant.

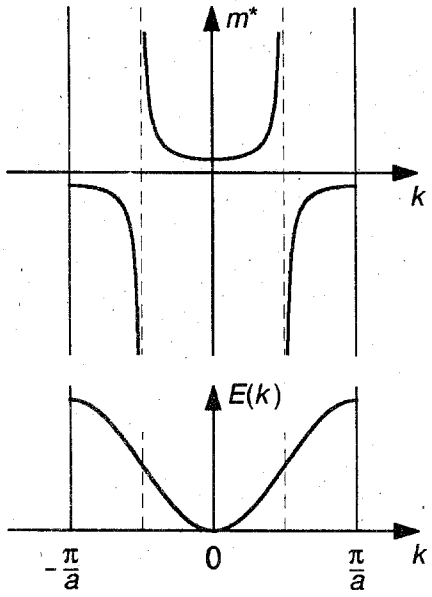


Abb. 8.2: Zusammenhang zwischen der effektiven Masse m^* und der Dispersionsrelation $E(k)$ in einer eindimensionalen Darstellung.

Es ist nicht überraschend, daß die effektive Masse vom Bewegungszustand der Elektronen abhängt, denn diese „Beobachtung" macht man auch bei klassischen Systemen. Stellen wir uns vor, daß auf eine Kugel in einer Flüssigkeit eine konstante Kraft wirkt. Beim Einsetzen der Kraft zum Zeitpunkt $t = 0$ spielt die Reibung keine Rolle, die Beschleunigung der Kugel wird durch die träge Masse bestimmt. Nach einiger Zeit wird die Geschwindigkeit der Kugel konstant, denn Reibungskraft und äußere Kraft halten sich dann die Waage. Läßt man bei der Beschreibung die Wechselwirkung der Kugel mit der Flüssigkeit unberücksichtigt und wendet das Konzept der effektiven Masse an, so wächst diese mit der Geschwindigkeit an, bis der Kugel schließlich die Masse „unendlich" zugeordnet werden muß, denn trotz der angreifenden Kraft tritt keine Beschleunigung mehr auf.

Die eigenartige Energieabhängigkeit der effektiven Elektronenmasse führt zu überraschenden Folgerungen. Legt man an einen idealen Kristall ein elektrisches Gleichfeld $\boldsymbol{\mathcal{E}}$ an, so wirkt auf die Elektronen eine konstante Kraft $\boldsymbol{F} = -e\boldsymbol{\mathcal{E}}$. Diese bewirkt aufgrund von Gleichung (8.4) eine gleichmäßige Bewegung der Elektronen im $\boldsymbol{k}$-Raum. Da aber die

Wellenfunktionen und die Energieeigenwerte periodisch verlaufen, resultiert daraus eine periodische Geschwindigkeitsänderung. Damit verbunden ist eine oszillatorische Bewegung der Elektronen, wie sie in Abb. 8.3 dargestellt ist. Im Idealfall stoßfreier Elektronenbewegung gibt es also bei einem unendlich großen Kristall keine Gleichstromleitfähigkeit! In einem perfekten Kristall würden die Elektronen bei $T = 0$ nur sogenannte *Bloch-Oszillationen* ausführen. Anschaulich kann man sich vorstellen, daß die Elektronen bei Annäherung an die Grenze der Brillouin-Zone eine „Bragg-Reflektion" erfahren und sich ihre Bewegungsrichtung umkehrt, obwohl die Kraft nach wie vor in die gleiche Richtung wirkt.

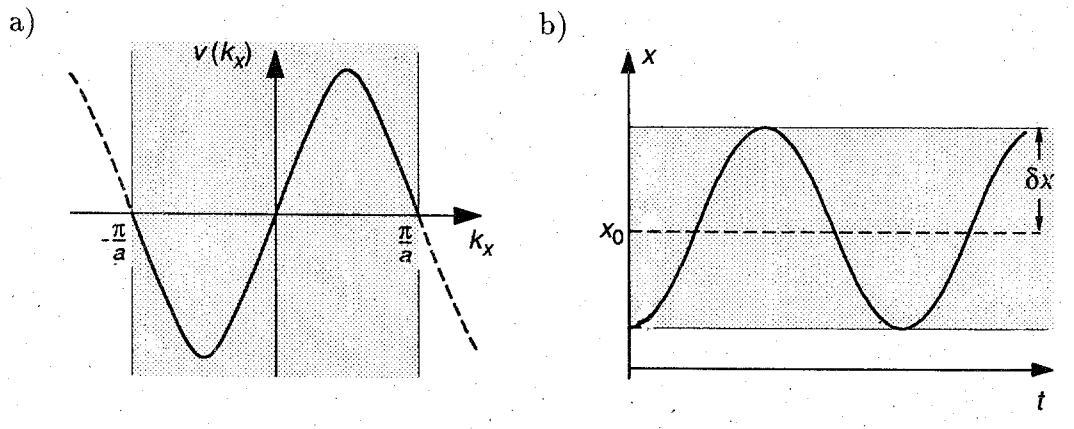


Abb. 8.3: Elektronenbewegung unter dem Einfluß eines konstanten elektrischen Feldes $-\mathcal{E}_x$.
a) Geschwindigkeit als Funktion des Wellenzahl: An der Grenze der Brillouin-Zone „springen" die Elektronen im reduzierten Zonenschema von π/a nach $-\pi/a$.
b) Zeitlicher Verlauf der räumlichen Auslenkung: Es tritt eine oszillatorische Bewegung um den Mittelwert x_0 auf.

Da sich die Elektronen im $\boldsymbol{k}$-Raum mit der Geschwindigkeit $|\dot{k}| = e\mathcal{E}/\hbar$ durch die Brillouin-Zone mit der Ausdehnung $2\pi/a$ bewegen, ist ihre Schwingungsdauer T_B durch

$$T_\mathrm{B} = \frac{2\pi}{a} \Big/ \frac{e\mathcal{E}}{\hbar} = \frac{h}{ae\mathcal{E}} \tag{8.7}$$

gegeben. Mit $\mathcal{E} = 1\,\mathrm{V/cm}$ und $a = 2\,\text{Å}$ ergibt sich der Zahlenwert $T_\mathrm{B} \approx 200\,\mathrm{ns}$ bzw. die Frequenz $\nu_\mathrm{B} \approx 5\,\mathrm{MHz}$. Zum Abschätzen der Amplitude δx der Auslenkung setzen wir als mittlere Geschwindigkeit $\overline{v}$ die Fermi-Geschwindigkeit, also etwa $\overline{v} \approx 10^6\,\mathrm{m/s}$, ein. Damit ergibt sich für die Amplitude der Auslenkung in etwa $\delta x \approx \overline{v}\,T_\mathrm{B}/4 \approx 5\,\mathrm{cm}$.

Während bei sehr großen Kristallen die Leitfähigkeit verschwinden sollte, würde man bei genügend kleiner Probendimension aber eine sehr große Leitfähigkeit erwarten, da die Elektronen über kurze Distanzen ungehindert durch die Probe laufen könnten. Beide Grenzfälle werden nicht beobachtet. Offensichtlich haben wir die Idealisierung zu weit getrieben, denn in realen Materialien durchlaufen die Elektronen nicht ungehindert die Brillouin-Zone, sondern stoßen mit Verunreinigungen, Phononen und anderen Elektronen zusammen. Da diese Stöße relativ häufig auftreten, verändern sie das Verhalten der Elektronen drastisch. Die typischen Stoßzeiten (vgl. Kap. 8.2) liegen größenordnungsmäßig bei etwa $10^{-14}\,\mathrm{s}$, d.h. Elektronen legen nur etwa 10 nm stoßfrei zurück. Oszillationsbewegungen sind daher grundsätzlich nur in sehr, sehr reinen und fehlerfreien Materialien bei tiefen Temperaturen zu erwarten. In neuesten Messungen an besonders präparierten, sehr

reinen Halbleiterübergitter[2] konnten Bloch-Oszillationen nachgewiesen werden; allerdings nicht, wie hier geschildert, über das Verhalten der Gleichstromleitfähigkeit.

Ladungstransport in Bändern

Wenn auch die Bloch-Oszillationen nicht direkt in Erscheinung treten, so hat das eigenartige Verhalten der Elektronen im periodischen Potential dennoch erhebliche Konsequenzen für eine Reihe von Eigenschaften, insbesondere für den Stromtransport. Ein Elektron mit der Ladung $-e$, das sich an der unteren Bandkante befindet, läuft beim Anlegen eines Feldes wie ein freies Elektron dem Feld *entgegen*. Ein Elektron über dem Wendepunkt der $E(\boldsymbol{k})$-Kurve dagegen bewegt sich aufgrund seiner negativen effektiven Masse wie eine positive Ladung *in* Feldrichtung.

Wir berechnen nun die Stromdichte, die ein elektrisches Gleichfeld hervorruft. Der Beitrag der Elektronen mit einem Wellenvektor zwischen $\boldsymbol{k}$ und $\boldsymbol{k}+\mathrm{d}\boldsymbol{k}$ zur Stromdichte $\boldsymbol{j}$ ist durch

$$\mathrm{d}\boldsymbol{j} = \frac{-e}{4\pi^3}\,\boldsymbol{v}(\boldsymbol{k})\,\mathrm{d}^3k = \frac{-e}{4\pi^3\hbar}\nabla_{\boldsymbol{k}}E(\boldsymbol{k})\,\mathrm{d}^3k \tag{8.8}$$

gegeben. $(4\pi^3)^{-1}$ ist die Zustandsdichte der Elektronen im $\boldsymbol{k}$-Raum. Die Stromdichte erhält man wie in Kap. 7.2 bei der Diskussion der Glühemission durch Integration über die Wellenvektoren. Um die auftretenden Ausdrücke zu vereinfachen, führen wir die Rechnung für $T = 0$ durch. Am absoluten Nullpunkt besitzt die Fermi-Dirac-Verteilung die Form einer Stufenfunktion, die besetzte von unbesetzten Zuständen trennt. Natürlich gelten die folgenden Überlegungen auch bei endlichen Temperaturen, wenn wir die thermische Anregung der Elektronen berücksichtigen. Die Integration erstreckt sich über die besetzten Zustände und führt zu dem Ergebnis

$$\boldsymbol{j} = \frac{-e}{4\pi^3}\int\limits_{\text{besetzt}} \boldsymbol{v}(\boldsymbol{k})\,\mathrm{d}^3k = \frac{-e}{4\pi^3\hbar}\int\limits_{\text{besetzt}} \nabla_{\boldsymbol{k}}E(\boldsymbol{k})\,\mathrm{d}^3k\,. \tag{8.9}$$

Für die folgende Betrachtung ist entscheidend, daß es ohne Feld immer Paare von Elektronen mit entgegengesetzter Geschwindigkeit gibt. Dies läßt sich einfach bei Kristallen mit Inversionssymmetrie zeigen: Da das reziproke Gitter die Punktsymmetrie des realen Gitters besitzt, gilt in diesem Fall $E(\boldsymbol{k}) = E(-\boldsymbol{k})$. Damit folgt für die Geschwindigkeit eines Elektrons mit dem Wellenvektor $-\boldsymbol{k}$:

$$\boldsymbol{v}(-\boldsymbol{k}) = \frac{1}{\hbar}\nabla_{-\boldsymbol{k}}E(-\boldsymbol{k}) = -\frac{1}{\hbar}\nabla_{\boldsymbol{k}}E(\boldsymbol{k}) = -\boldsymbol{v}(\boldsymbol{k})\,. \tag{8.10}$$

Elektronen mit entgegengesetzten Wellenvektoren laufen in entgegengesetzten Richtungen. Ohne Feld ist zu jedem Elektron mit Wellenvektor $+\boldsymbol{k}$ eines mit Wellenvektor $-\boldsymbol{k}$ zu finden. Daher existiert für jedes Elektron mit der Geschwindigkeit $\boldsymbol{v}(\boldsymbol{k})$ eines mit $\boldsymbol{v}(-\boldsymbol{k}) = -\boldsymbol{v}(\boldsymbol{k})$, und das Integral (8.9) nimmt den Wert Null an, d.h. $\boldsymbol{j} \equiv 0$. Legt man ein elektrisches Feld an, so ändert sich bei einem *vollen* Band nichts an der Argumentation. In diesem Fall werden die Zustände nur umbesetzt werden, und nach außen hin ist keine Veränderung zu sehen. Der räumliche und zeitliche Mittelwert der Elektronengeschwindigkeit über alle Elektronen im Band verschwindet.

Diese Beziehung läßt sich auf Strukturen ohne Inversionssymmetrie erweitern, wenn man die beiden erlaubten Spinrichtungen der Unterbänder mit einbezieht. Ohne näher

[2] Diese spezielle Art von Gittern werden wir in Abschnitt 9.5 kennenlernen.

darauf einzugehen, sei hier erwähnt, daß dort $E(\boldsymbol{k})\uparrow = E(-\boldsymbol{k})\downarrow$ gilt und deshalb ein analoges Vorgehen möglich ist. Volle Bänder tragen also nicht zum Stromfluß bei. Damit wird verständlich, warum Isolatoren existieren, obwohl diese auch Bänder besitzen, in denen sich die Elektronen bewegen können.

Ist das Band nicht voll, so bewirkt das angelegte Feld eine Veränderung der Geschwindigkeitsverteilung innerhalb des Bandes. Ohne Stöße würden die Elektronen die oben geschilderten Bloch-Oszillationen ausführen. Die Stöße bewirken jedoch, wie wir im nächsten Abschnitt sehen werden, daß es nur zu geringfügigen Verschiebungen der Elektronen im $\boldsymbol{k}$-Raum kommt. Wegen der Auszeichnung durch die Feldrichtung ist die Besetzung der Zustände in einem teilweise gefüllten Band nicht mehr inversionssymmetrisch, wodurch $\boldsymbol{j} \neq 0$ wird. In Kap. 8.2 werden wir die Integration durchführen und eine geeignete Näherung für die elektrische Leitfähigkeit finden.

Die Integration über die besetzten Zustände in (8.9) läßt sich auf unterschiedliche Weise ausführen, denn man kann entweder die besetzten oder die unbesetzten Zustände betrachten:

$$\boldsymbol{j} = \frac{-e}{4\pi^3} \int\limits_{\text{besetzt}} \boldsymbol{v}(\boldsymbol{k})\,\mathrm{d}^3k = \frac{-e}{4\pi^3}\Bigg[\underbrace{\int\limits_{\text{BZ}} \boldsymbol{v}(\boldsymbol{k})\,\mathrm{d}^3k}_{\equiv 0} - \int\limits_{\text{leer}} \boldsymbol{v}(\boldsymbol{k})\,\mathrm{d}^3k\Bigg]$$

$$= \frac{+e}{4\pi^3} \int\limits_{\text{leer}} \boldsymbol{v}(\boldsymbol{k})\,\mathrm{d}^3k\,. \tag{8.11}$$

In dieser Gleichung haben wir das Integral über die besetzten Zustände aufgespalten in ein Integral über die gesamte Brillouin-Zone (BZ) und die leeren Zustände. Wie wir soeben diskutiert haben, liefert die Brillouin-Zone bei vollem Band keinen Beitrag zum Strom. In der neuen Betrachtungsweise erfolgt der Stromtransport durch die leeren Zustände, die nun aber *positiv* geladen sind. Man nennt diese Quasiteilchen *Defektelektronen* oder *Löcher*. Das Einführen dieses Konzepts mag zunächst ein wenig „künstlich" erscheinen. Es wird sich aber zeigen, daß es viele Vorteile bietet, wenn das Band *fast* gefüllt ist. Will man beispielsweise Eigenschaften für diesen Fall berechnen, die mit der Elektronenbewegung verknüpft sind, so muß nur die Energiefläche in der Nähe des Maximums bekannt sein. Dort kann die Dispersionskurve meist in sehr guten Näherung durch eine Parabel beschrieben werden und die Rechnung vereinfacht sich erheblich.

Elektronen und Löcher

Die Löcher haben Eigenschaften, die stark an positive Ladungsträger erinnern. Dennoch darf man in ihnen nicht einfach positiv geladene Teilchen sehen. Wir wollen nun ihre wichtigsten Eigenschaften kurz erläutern. Um zwischen Elektronen und Löchern zu unterscheiden, kennzeichnen wir die auftretenden Größen in der weiteren Beschreibung mit dem Index n bzw. p.

Die Summe der Wellenvektoren aller Elektronen in einem vollen Band verschwindet, d.h. $\sum \boldsymbol{k} = 0$. Wir nehmen an, daß *ein* Elektron mit dem Wellenvektor $\boldsymbol{k}_\mathrm{n}$ aus dem vollen Valenzband ins Leitungsband gehoben wurde. Wie in Abb. 8.4a veranschaulicht, könnte dieser Vorgang durch optische Absorption bewirkt worden sein[3]. Da im Valenzband nun ein Elektron fehlt, gilt jetzt $\sum \boldsymbol{k} = -\boldsymbol{k}_\mathrm{n}$. Wir können daher dem fehlenden

[3] Dieser Effekt wird in Kap. 9.1 im Zusammenhang mit den optischen Eigenschaften von Halbleitern ausführlich diskutiert.

Elektron bzw. dem Loch den Wellenvektor $\boldsymbol{k}_\mathrm{p} = -\boldsymbol{k}_\mathrm{n}$ zuordnen[4]. Bei der Interpretation der Darstellung 8.4a ist jedoch Vorsicht geboten. Der Übergang des Elektrons erfolgt vom Punkt A zum Punkt B. Das Loch besitzt jedoch nicht den Wellenvektor $\boldsymbol{k}_\mathrm{n}$ des fehlenden Elektrons sondern $-\boldsymbol{k}_\mathrm{n}$. Es wird daher durch den Punkt C repräsentiert und besitzt den Wellenvektor des Elektrons, das diese Stelle der Dispersionsrelation belegt.

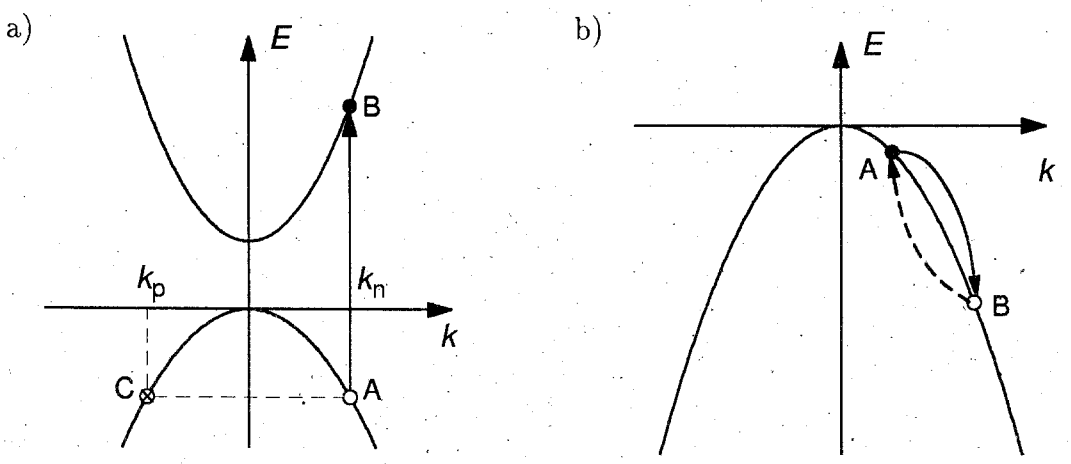


Abb. 8.4: Zum Konzept des Lochs.
a) Ein Elektron wird, z.B. durch optische Absorption, aus dem vollen Valenzband ins leere Leitungsband angehoben. Das Elektron im Leitungsband ist durch einen Punkt, das fehlende Elektron durch einen offenen Kreis dargestellt. Das resultierende Loch in C mit dem Wellenvektor $\boldsymbol{k}_\mathrm{p} = -\boldsymbol{k}_\mathrm{n}$ ist durch einen Kreis mit Kreuz angedeutet.
b) Durch den Sprung eines Elektrons von A in den leeren Zustand B wird das entsprechende Loch angehoben. Bei diesem Vorgang wird die Energiedifferenz zwischen den beiden Zuständen frei.

Wie in Abb. 8.4b angedeutet, kann ein Loch aus dem Inneren des Valenzbandes dadurch an die Bandoberkante gelangen, daß ein Elektron aus dem Zustand A den leeren Zustand B besetzt. Bei diesem Vorgang wird Energie frei. Umgekehrt muß Energie aufgebracht werden um ein Loch von der Valenzbandoberkante nach „unten" zu verschieben, denn dazu muß ein Elektron aus dem Bandinneren an die Valenzbandkante gehoben werden. Anhand dieser Überlegungen wird verständlich, daß zwischen der Elektronenenergie $E_\mathrm{n}(\boldsymbol{k})$ und der Lochenergie $E_\mathrm{p}(\boldsymbol{k})$ die Beziehung

$$E_\mathrm{p}(\boldsymbol{k}) = -E_\mathrm{n}(\boldsymbol{k}) \tag{8.12}$$

besteht.

Die Beziehung zwischen den effektiven Massen der beiden Ladungsträgersorten folgt unmittelbar aus der Definition (8.6) und der Inversionssymmetrie der Bänder. Sie besitzen entgegengesetztes Vorzeichen, denn es gilt

$$\left[\frac{1}{m^*}\right]_\mathrm{p} = \frac{1}{\hbar^2}\left[\frac{\partial^2 E(\boldsymbol{k})}{\partial \boldsymbol{k}\,\partial \boldsymbol{k}}\right]_\mathrm{p} = \frac{1}{\hbar^2}\left[\frac{-\partial^2 E(\boldsymbol{k})}{(-\partial \boldsymbol{k})\,(-\partial \boldsymbol{k})}\right]_\mathrm{n} = -\left[\frac{1}{m^*}\right]_\mathrm{n},$$

$$m_\mathrm{p}^* = -m_\mathrm{n}^*\,. \tag{8.13}$$

[4] Eine analoge Betrachtung gilt auch für den Spin, den die Löcher tragen.

Unter dem Einfluß eines elektrischen Feldes bewegen sich alle Elektronen im $\boldsymbol{k}$-Raum mit der gleichen Geschwindigkeit. Den Elektronen folgt natürlich auch der unbesetzte Zustand, den wir hier als Loch beschreiben. Dies läßt sich leicht zeigen, da bei der Berechnung der Geschwindigkeit der Löcher sowohl bei der Gradientenbildung als auch bei der Lochenergie ein negatives Vorzeichen auftritt:

$$\nabla_{\boldsymbol{k}} E_{\mathrm{p}}(\boldsymbol{k}) = \nabla_{\boldsymbol{k}} E_{\mathrm{n}}(\boldsymbol{k}) \qquad \rightarrow \qquad \boldsymbol{v}_{\mathrm{p}}(\boldsymbol{k}_{\mathrm{p}}) = \boldsymbol{v}_{\mathrm{n}}(\boldsymbol{k}_{\mathrm{n}}) \; . \tag{8.14}$$

Die Bewegungsgleichungen nehmen somit folgende Gestalt an:

$$\hbar \dot{\boldsymbol{k}}_{\mathrm{n}} = -e(\boldsymbol{\mathcal{E}} + \boldsymbol{v}_{\mathrm{n}} \times \boldsymbol{B}) \; , \tag{8.15}$$

$$\hbar \dot{\boldsymbol{k}}_{\mathrm{p}} = +e(\boldsymbol{\mathcal{E}} + \boldsymbol{v}_{\mathrm{p}} \times \boldsymbol{B}) \; . \tag{8.16}$$

Das Loch wirkt nach außen wie ein Teilchen mit positiver Ladung!

Hier soll noch auf eine wichtige Folgerung hingewiesen werden: Beim Stromtransport bewegen sich Elektronen und Löcher in entgegengesetzter Richtung, da sich die Elektronen an der Unterkante des Leitungsbandes, die Löcher aber an der Oberkante des Valenzbandes aufhalten. Infolge ihrer entgegengesetzten Ladung addiert sich der Beitrag der Elektronen und der Löcher zum Stromtransport.

8.2 Elektrische und thermische Leitfähigkeit

Bei der Diskussion der *elektrischen Leitfähigkeit* von Metallen beschränken wir uns hier auf die Gleichstromleitfähigkeit. Die „klassische" Theorie hierzu wurde bereits 1898 von *Riecke* und 1900 von *Drude* entwickelt und gab erstaunlich gut die experimentellen Beobachtungen wieder. Die Erklärung beruhte auf der Annahme, daß sich die Bewegung von Elektronen mit Hilfe der kinetischen Gastheorie beschreiben läßt. Die Elektronen werden wie freie Teilchen behandelt, die sich mit der thermischen Geschwindigkeit $\boldsymbol{v}_{\mathrm{th}}$ bewegen und ständig untereinander und mit den Ionenrümpfen stoßen. Da diese Stöße statistisch erfolgen, verschwindet der Mittelwert ihrer Geschwindigkeit, d.h. $\overline{v}_{\mathrm{th}} = 0$. Die Zeit τ zwischen zwei derartigen Stößen wird als *mittlere Stoßzeit* oder *Relaxationszeit* bezeichnet. Legt man ein elektrisches Feld $\boldsymbol{\mathcal{E}}$ an, so gewinnen die Elektronen zwischen den Stößen die zusätzliche Geschwindigkeit $-e\boldsymbol{\mathcal{E}}t/m$. Die grundlegende Annahme ist, daß auch im elektrischen Feld nach jedem Stoß die Bewegungsrichtung willkürlich verteilt ist. Die Elektronen erfahren daher durch das angelegte Feld nur eine Driftgeschwindigkeit $\boldsymbol{v}_{\mathrm{d}}$ entgegen der Feldrichtung, die durch die Stärke des anliegenden Feldes und durch die Relaxationszeit τ bestimmt wird:

$$\boldsymbol{v}_{\mathrm{d}} = -e\frac{\boldsymbol{\mathcal{E}}\tau}{m} = -\mu\boldsymbol{\mathcal{E}} \; . \tag{8.17}$$

Als Abkürzung haben wir die *Beweglichkeit* $\mu = e\tau/m$ eingeführt. Damit ist die Stromdichte nach Drude durch

$$\boldsymbol{j} = -en\boldsymbol{v}_{\mathrm{d}} = \frac{ne^2\tau\boldsymbol{\mathcal{E}}}{m} = ne\mu\boldsymbol{\mathcal{E}} \tag{8.18}$$

gegeben. Obwohl gute Übereinstimmung mit den experimentellen Resultaten gefunden wurde, ist dieses Konzept nicht richtig. Bei der Behandlung des Problems wurde von der falschen Annahme ausgegangen, daß alle Elektronen stoßen können. Dies verbietet jedoch das Pauli-Prinzip. Die Konsequenzen, die sich daraus ergeben, werden wir nun ausführlich diskutieren.

Sommerfeld-Modell

Eine wesentlich verbesserte Theorie wurde von *Sommerfeld* vorgeschlagen. Sie beschreibt nicht nur die elektrische Leitfähigkeit, auf die wir hier eingehen, sondern auch die elektronischen Eigenschaften, die wir bereits in Kap. 7 behandelt haben. Die Elektronen werden zwar weiterhin als quasi-freie Teilchen betrachtet, doch gehorchen sie der Schrödinger-Gleichung und unterliegen dem Pauli-Prinzip. Für Energie und Impuls gilt

$$E = \frac{\hbar^2 k^2}{2m^*} = \frac{p^2}{2m^*} \qquad \text{und} \qquad \boldsymbol{p} = \hbar \boldsymbol{k} \,. \tag{8.19}$$

Unter der Einwirkung einer äußeren Kraft $\boldsymbol{F}$ bzw. eines elektrischen Feldes $\boldsymbol{\mathcal{E}}$ entwickelt sich der Wellenvektor entsprechend Gleichung (8.4):

$$\hbar \dot{\boldsymbol{k}} = \boldsymbol{F} = -e\boldsymbol{\mathcal{E}} \,. \tag{8.20}$$

Welche Auswirkungen hat das Anlegen eines elektrischen Feldes auf die Fermi-Kugel? Wie in Abb. 8.5a gezeigt, fallen ohne äußeres Feld das Zentrum der Fermi-Kugel und der Ursprung des $\boldsymbol{k}$-Raumes, der Γ-Punkt, zusammen. Legen wir ein elektrisches Feld an, so verschiebt sich entsprechend (8.20) jeder Wellenvektor und damit auch die Fermi-Kugel als Ganzes um den Betrag

$$\delta \boldsymbol{k} = \frac{-e\boldsymbol{\mathcal{E}}\delta t}{\hbar} \,, \tag{8.21}$$

wobei δt die Zeit ist, die nach dem Anlegen des elektrischen Feldes verstrichen ist. Die Verschiebung der Fermi-Kugel ist in Abb. 8.5b veranschaulicht.

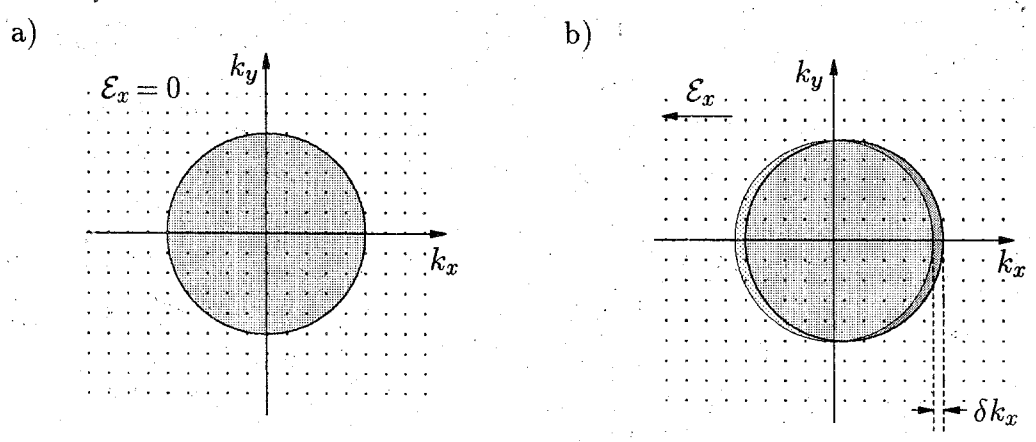


Abb. 8.5: Verschiebung der Fermi-Kugel im elektrischen Feld. Die Punkte repräsentieren die erlaubten Wellenvektoren.
a) Fermi-Kugel ohne elektrisches Feld. Koordinatenursprung und Kugelmittelpunkt fallen zusammen.
b) Unter dem Einfluß des elektrischen Feldes $-\mathcal{E}_x$ haben sich die erlaubten Wellenvektoren um den Betrag δk_x verschoben. Durch Stöße werden Elektronen aus dem dunklen Bereich der Vorderseite in den hellen Bereich der Rückseite transportiert.

Streuprozesse, auf die wir im nächsten Abschnitt etwas ausführlicher eingehen, bewirken einen Transport der Elektronen von der Vorder- auf die Rückseite der Fermi-Kugel. Kurze Zeit nach dem Anlegen des Feldes stellt sich so ein dynamisches Gleichgewicht

zwischen der Verschiebung der Fermi-Kugel und den Umlagerungen aufgrund der Streuprozesse ein, und die Kugel kommt zum Stillstand. Es soll hier betont werden, daß sich die Fermi-Kugel aufgrund der Streuprozesse nur geringfügig verschiebt, obwohl die Wellenvektoren der Elektronen gemäß (8.21) einer ständigen Veränderung unterliegen. Da die geringe Verschiebung der Fermi-Kugel nur kleine Energieänderungen hervorruft und die Elektronen nach dem Stoß einen freien Zustand vorfinden müssen, können nur Elektronen stoßen, die sich in unmittelbarer Umgebung der Fermi-Fläche befinden.

Während sich ohne Feld der Stromfluß in positive und negative x-Richtung gerade aufheben, gilt dies nicht mehr in Gegenwart eines Feldes. Der resultierende Strom wird durch die schnellen Elektronen des dunkel getönten Gebietes bzw. durch die leeren Zustände im hellen Bereich hervorgerufen. Im Gegensatz zu der klassischen Betrachtungsweise, bei der alle Elektronen mit der kleinen Driftgeschwindigkeit $\boldsymbol{v}_\mathrm{d}$ zum Stromfluß beitragen, sind es im Modell von Sommerfeld nur die wenigen, aber schnellen Elektronen an der Fermi-Fläche. Offen bleibt bei den bisherigen Überlegungen, wie man die „Stoßzeit“ bzw. die Zeit δt in Gleichung (8.21) einführt, die zwischen dem Anschalten des Feldes und dem Einstellen der stationären Verschiebung verstreicht. Im folgenden Abschnitt werden wir diese Frage behandeln und den korrekten Ausdruck für die Leitfähigkeit ableiten. Anschließend werden wir auf die Streumechanismen eingehen, die die Streuzeiten bestimmen.

Boltzmann-Gleichung, Relaxationszeit

Ohne äußeres elektrisches Feld ist die Besetzung der elektronischen Zustände durch die Fermi-Dirac-Verteilungsfunktion $f_0(\boldsymbol{k})$ gegeben, wobei wir den Gleichgewichtswert durch den Index 0 kennzeichnen. Wird ein Feld angelegt, so werden die Elektronen beschleunigt, doch nach kurzer Zeit stellt sich aufgrund von Stößen ein stationärer Nichtgleichgewichtswert $f(\boldsymbol{k}, \boldsymbol{r}, t)$ ein.

Grundsätzlich gibt es drei Ursachen für Änderungen der Verteilungsfunktion: die Diffusion aufgrund von Schwankungen der räumlichen Elektronendichte, das Wirken von äußeren Feldern und Streuprozesse. Die zeitliche und räumliche Entwicklung der Verteilungsfunktion läßt sich mit Hilfe der *Transport-* oder *Boltzmann-Gleichung* beschreiben. Zur Herleitung dieser Gleichung betrachten wir die Wirkung einer Zeitverschiebung $\mathrm{d}t$ auf die Verteilungsfunktion. Nach dem *Liouvilleschen Satz* der klassischen Mechanik bleibt die Dichte im Phasenraum, d.h. die Verteilungsfunktion, in Abwesenheit von Stößen erhalten:

$$f(\boldsymbol{k} + \mathrm{d}\boldsymbol{k}, \boldsymbol{r} + \mathrm{d}\boldsymbol{r}, t + \mathrm{d}t) = f(\boldsymbol{k}, \boldsymbol{r}, t) \,. \tag{8.22}$$

Für kleine Änderungen der Variablen läßt sich diese Gleichung wie folgt umschreiben:

$$f(\boldsymbol{k} + \mathrm{d}\boldsymbol{k}, \boldsymbol{r} + \mathrm{d}\boldsymbol{r}, t + \mathrm{d}t) - f(\boldsymbol{k}, \boldsymbol{r}, t) = \frac{\partial f}{\partial \boldsymbol{k}} \cdot \mathrm{d}\boldsymbol{k} + \frac{\partial f}{\partial \boldsymbol{r}} \cdot \mathrm{d}\boldsymbol{r} + \frac{\partial f}{\partial t}\,\mathrm{d}t = 0 \,. \tag{8.23}$$

Beim Auftreten von Stößen muß ein „Korrekturterm“ hinzugefügt werden, der dann summarisch den Beitrag aller Stöße enthält. Damit entsteht aus (8.23) die Boltzmann-Gleichung

$$\dot{\boldsymbol{k}} \cdot \frac{\partial f}{\partial \boldsymbol{k}} + \dot{\boldsymbol{r}} \cdot \frac{\partial f}{\partial \boldsymbol{r}} + \frac{\partial f}{\partial t} = \left.\frac{\partial f}{\partial t}\right|_{\mathrm{Streu}} \,. \tag{8.24}$$

Die beiden ersten Terme auf der linken Seite sind der *Feld-* bzw. der *Diffusionsterm*, auf der rechten Seite steht der *Streu-* oder *Stoßterm*. Bei der Diskussion der elektrischen Leitfähigkeit spielt der Diffusionsterm keine Rolle, da ein elektrisches Feld die homogene

räumliche Elektronenverteilung nicht verändert. Wir lassen daher in der weiteren Diskussion diesen Term weg[5].

Für den Stoßterm benutzen wir den einfachen *Relaxationszeit-Ansatz*

$$\left.\frac{\partial f(\boldsymbol{k})}{\partial t}\right|_{\text{Streu}} = -\frac{f(\boldsymbol{k}) - f_0(\boldsymbol{k})}{\tau(\boldsymbol{k})}\,, \tag{8.25}$$

der sich näherungsweise aus der Streutheorie herleiten läßt, doch wollen wir hier die Herleitung nicht nachvollziehen. Die grundsätzliche Bedeutung der Relaxationszeit $\tau(\boldsymbol{k})$ bzw. der Streurate τ^{-1} haben wir bereits in Abschnitt 6.2 bei der Diskussion der Phononenstreuung kennengelernt. Im vorliegenden Fall wird die Nichtgleichgewichtsverteilung $f(\boldsymbol{k}, \boldsymbol{r}, t)$ durch Streuung der Elektronen mit dem Wellenvektor $\boldsymbol{k}$ in Zustände mit dem Wellenvektor $\boldsymbol{k}'$ abgebaut. Wie bei der Phononenstreuung muß auch hier berücksichtigt werden, daß Streuprozesse grundsätzlich umkehrbar sind, d.h. daß auch Streuung von $\boldsymbol{k}'$ nach $\boldsymbol{k}$ erfolgt, wodurch das Nichtgleichgewicht sogar verstärkt wird. Die Wahrscheinlichkeit für das Auftreten eines Streuereignis ist in beiden Richtungen gleich groß. Die beiden Prozesse heben sich jedoch nicht auf, weil die Besetzung des jeweiligen Ausgangszustandes für die Vor- und Rückreaktion aufgrund der unterschiedlichen Verteilungsfunktionen $f(\boldsymbol{k}, \boldsymbol{r}, t)$ bzw. $f_0(\boldsymbol{k}, \boldsymbol{r}, t)$ verschieden groß ist. Die Streuwahrscheinlichkeit läßt sich für eine Reihe von Streumechanismen berechnen, doch ist hierfür meist ein größerer mathematischer Aufwand erforderlich, den wir hier vermeiden werden.

Die Bedeutung der Relaxationszeit beim Ladungstransport und der Zusammenhang mit der Verschiebung der Fermi-Kugel läßt sich an einem einfachen Beispiel verdeutlichen: Nehmen wir an, daß sich unter dem Einfluß eines elektrischen Feldes eine stationäre Nichtgleichgewichtsverteilung $\overline{f}(\boldsymbol{k})$ eingestellt hat und schalten das Feld abrupt ab, dann erfolgt ein Übergang von $\overline{f}(\boldsymbol{k})$ zur Gleichgewichtsverteilung $f_0(\boldsymbol{k})$. Ohne äußeres Feld nimmt (8.24) mit dem Relaxationszeit-Ansatz (8.25) die einfache Form

$$\frac{\partial f(\boldsymbol{k})}{\partial t} = -\frac{f(\boldsymbol{k}) - f_0(\boldsymbol{k})}{\tau(\boldsymbol{k})} \tag{8.26}$$

an. Mit der Anfangsbedingung $f(\boldsymbol{k}, t = 0) = \overline{f}(\boldsymbol{k})$ folgt daraus

$$f(\boldsymbol{k}) - f_0(\boldsymbol{k}) = \left[\, \overline{f}(\boldsymbol{k}) - f_0(\boldsymbol{k}) \,\right] \mathrm{e}^{-t/\tau}\,. \tag{8.27}$$

Die Abweichung von der Gleichgewichtsverteilung klingt nach dem Abschalten des Feldes mit der charakteristischen Zeitkonstante τ exponentiell ab. Entsprechend stellt sich mit der gleichen Zeitkonstante beim Einschalten eines elektrischen Feldes die stationäre Verschiebung der Fermi-Kugel ein. τ ist also die Zeit, in der sich die Verteilungsfunktion auf eine plötzliche Änderung des äußeren Feldes einstellt, also die Zeit, in der ein neuer stationärer Zustand erreicht wird.

[5] Hätten wir die Glühemission von Elektronen in Kap. 7.2 mit Hilfe der Boltzmann-Gleichung diskutiert, so hätten wir diesen Term betrachten müssen.

Ladungstransport

Nun wollen wir die elektrische Gleichstromleitfähigkeit von Metallen diskutieren. Hierzu berechnen wir mit Hilfe der Boltzmann-Gleichung die Stromdichte $\boldsymbol{j}$, die von einem konstanten elektrischen Feld $\boldsymbol{\mathcal{E}}$ hervorgerufen wird. Wir lassen Einschaltvorgänge außer acht und betrachten nur den stationären Zustand, d.h. der Term $\partial f/\partial t$, der die explizite Zeitabhängigkeit beschreibt, verschwindet in Gleichung (8.24). Die Verteilungsfunktion hängt in dem betrachteten einfachen Fall weder vom Ort noch von der Zeit ab, so daß die Boltzmann-Gleichung mit $\dot{\boldsymbol{k}} = -e\boldsymbol{\mathcal{E}}/\hbar$ die einfache Gestalt

$$-\frac{e}{\hbar}\,\boldsymbol{\mathcal{E}} \cdot \nabla_{\boldsymbol{k}} f(\boldsymbol{k}) = -\frac{f(\boldsymbol{k}) - f_0(\boldsymbol{k})}{\tau(\boldsymbol{k})} \tag{8.28}$$

annimmt. Löst man nach $f(\boldsymbol{k})$ auf, so folgt daraus

$$f(\boldsymbol{k}) = f_0(\boldsymbol{k}) + \frac{e\tau(\boldsymbol{k})}{\hbar}\,\boldsymbol{\mathcal{E}} \cdot \nabla_{\boldsymbol{k}} f(\boldsymbol{k})\,. \tag{8.29}$$

Diese Gleichung können wir für *kleine Abweichungen* der Verteilungsfunktion vom Gleichgewichtswert lösen, wenn wir den Gradienten der tatsächlichen Verteilungsfunktion durch den Gradienten der Gleichgewichtsverteilung ersetzen. Wir erhalten dann die *linearisierte Boltzmann-Gleichung*

$$f(\boldsymbol{k}) \approx f_0(\boldsymbol{k}) + \frac{e\tau(\boldsymbol{k})}{\hbar}\,\boldsymbol{\mathcal{E}} \cdot \nabla_{\boldsymbol{k}} f_0(\boldsymbol{k})\,. \tag{8.30}$$

Um die Stromdichte zu berechnen, muß der Beitrag $-e\boldsymbol{v}(\boldsymbol{k})$ aller Elektronen mit den unterschiedlichen Wellenvektoren aufsummiert werden. Wir gehen von Gleichung (8.9) aus, die bei $T = 0$ gültig ist, und fügen die Verteilungsfunktion wie in (7.37) bei der Behandlung der Glühemission ein. Für die Stromdichte schreiben wir daher

$$\boldsymbol{j} = -\frac{e}{4\pi^3}\int \boldsymbol{v}(\boldsymbol{k})\, f(\boldsymbol{k})\, \mathrm{d}^3k\,. \tag{8.31}$$

Nehmen wir hier der Einfachheit halber an, daß es sich um einen isotropen oder kubischen Festkörper handelt, dann ist $j_y = j_z = 0$, wenn das elektrische Feld in x-Richtung anliegt. Nun setzen wir $f(\boldsymbol{k})$ aus der linearisierten Boltzmann-Gleichung (8.30) in (8.31) ein. Die Gleichgewichtsverteilung $f_0(\boldsymbol{k})$ trägt nicht zum Stromfluß bei, da der Mittelwert über alle Geschwindigkeiten $\boldsymbol{v}(\boldsymbol{k})$ verschwindet. Der Term mit dem Gradienten liefert dagegen den interessierenden Stromfluß. Da $\partial f_0(\boldsymbol{k})/\partial k_x = [\partial f_0(\boldsymbol{k})/\partial E]\,\hbar v_x$ ist, läßt sich für die Stromdichte

$$j_x = -\frac{e^2\mathcal{E}_x}{4\pi^3}\int v_x^2\,\tau(\boldsymbol{k})\,\frac{\partial f_0(\boldsymbol{k})}{\partial E}\,\mathrm{d}^3k \tag{8.32}$$

schreiben. Bei der Auswertung dieses Integrals berücksichtigen wir, daß die Fermi-Dirac-Funktion bei der Fermi-Energie einen sehr steilen Abfall aufweist (vgl. Abb. 7.3), sonst aber näherungsweise konstant ist. Die Ableitung dieser Funktion ist daher nur in der unmittelbaren Umgebung der Fermi-Energie merklich von Null verschieden. Da sie dort sehr große Werte annimmt, ersetzen wir die Ableitung durch eine Deltafunktion, d.h. wir setzen $[\partial f_0(\boldsymbol{k})/\partial E] \approx -\delta(E - E_\mathrm{F})$. Auf das Volumenelement d^3k sind wir bereits

ausführlich in Abschnitt 5.5 und 7.2 bei der Herleitung der Zustandsdichte von Phononen und Elektronen eingegangen. Wir benutzen (5.51) und erhalten für die elektrische Leitfähigkeit

$$\sigma = \frac{j_x}{\mathcal{E}} \approx \frac{e^2}{4\pi^3\hbar} \int \frac{v_x^2(\boldsymbol{k})}{v(\boldsymbol{k})}\, \tau(\boldsymbol{k})\, \delta(E - E_\mathrm{F})\, \mathrm{d}E\, \mathrm{d}S_\mathrm{E}$$

$$\approx \frac{e^2}{4\pi^3\hbar} \int \frac{v_x^2(\boldsymbol{k}_\mathrm{F})}{v(\boldsymbol{k}_\mathrm{F})}\, \tau(\boldsymbol{k}_\mathrm{F})\, \mathrm{d}S_\mathrm{F}\,. \tag{8.33}$$

Wie in Gleichung (5.51) steht $\mathrm{d}S_\mathrm{E}$ für ein Element der Fläche konstanter Energie im $\boldsymbol{k}$-Raum. Die zweite Zeile ergibt sich nach Ausführung der Integration über die Energie, wobei $\mathrm{d}S_\mathrm{F}$ das Flächenelement der Oberfläche der Fermi-Kugel repräsentiert. Bemerkenswert ist, daß zum Stromtransport und damit zur Leitfähigkeit nur Elektronen an der Fermi-Fläche beitragen, denn nur dort ist der Gradient der Verteilungsfunktion deutlich von Null verschieden.

Viele Metalle mit kubischem Gitter besitzen eine näherungsweise kugelförmige Fermi-Fläche. Die Fermi-Geschwindigkeit ist dann richtungsunabhängig, und wir dürfen vereinfachend $v_x^2 = v^2/3$ setzen. Weist auch die Relaxationszeit $\tau(\boldsymbol{k}_\mathrm{F})$ keine Richtungsabhängigkeit auf, so kann man all diese Größen aus dem Integral ziehen, und das Oberflächenintegral vereinfacht sich zu $\int \mathrm{d}S_\mathrm{F} = 4\pi k_\mathrm{F}^2$. Für die elektrische Leitfähigkeit findet man dann:

$$\sigma = \frac{1}{3}\, e^2\, D(E_\mathrm{F})\, v_\mathrm{F}^2\, \tau(E_\mathrm{F})\,. \tag{8.34}$$

Setzen wir noch die Zustandsdichte an der Fermi-Kante $D(E_\mathrm{F}) = 3n/(2E_\mathrm{F})$ und die Fermi-Geschwindigkeit $v_\mathrm{F}^2 = 2E_\mathrm{F}/m^*$ ein, so vereinfacht sich dieser Ausdruck weiter zu

$$\sigma = \frac{n\, e^2}{m^*}\, \tau(E_\mathrm{F})\,. \tag{8.35}$$

Diese Gleichung entspricht der klassischen Formel (8.18) von Drude, wenn man die effektive Masse m^* durch die freie Masse m der Elektronen ersetzt. Diese Formel suggeriert, daß alle Elektronen am Ladungstransport teilnehmen. Tatsächlich taucht die Elektronendichte n nur deswegen auf, weil wir in (8.34) die Größen $D(E_\mathrm{F})$ und v_F durch (7.14) und (7.15) ersetzt haben. Ein weiterer Unterschied besteht darin, daß in die „korrekte" Formel (8.34) die Relaxationszeit $\tau(E_\mathrm{F})$ der Elektronen an der Fermi-Fläche eingeht, in der Theorie von Drude aber die nicht näher spezifizierte Stoßzeit τ eher als Stoß der Elektronen gegen benachbarte Atome interpretiert werden konnte.

Hier noch einige typische Zahlenwerte. Für die Stoßzeit findet man in reinem Kupfer den Wert $\tau \approx 2 \cdot 10^{-14}\,\mathrm{s}$ bei Zimmertemperatur und $\tau \approx 2 \cdot 10^{-10}\,\mathrm{s}$ bei 4 K. Da nur Elektronen an der Oberfläche der Fermi-Kugel stoßen und sich diese Elektronen mit der Fermi-Geschwindigkeit $v_\mathrm{F} = 1{,}6 \cdot 10^6\,\mathrm{m/s}$ bewegen, läßt sich aus der Stoßzeit die freie Weglänge $l = v_\mathrm{F}\,\tau$ mit $l \approx 30\,\mathrm{nm}$ bzw. $l \approx 0{,}3\,\mathrm{mm}$ angeben. Bei Gallium wurden sogar freie Weglängen von $l \approx 10\,\mathrm{cm}$ beobachtet. Bei derartig großen Werten wird die freie Weglänge oft durch die Probengeometrie bestimmt. Interessant ist auch eine Abschätzung der relativen Verschiebung $\delta k/k$ der Fermi-Kugel. Setzen wir die Stoßzeit von reinem Kupfer bei Zimmertemperatur und eine elektrische Feldstärke von $\mathcal{E} = 1\,\mathrm{V/cm}$ in (8.21) ein, so erhalten wir $\delta k \approx 3 \cdot 10^3\,\mathrm{m}^{-1}$ und somit $\delta k/k \approx 10^{-7}$. Die Fermi-Kugel verschiebt sich also nur geringfügig.

Streuung von Elektronen

Wir wollen nun die Prozesse kurz betrachten, die zur Streuung der Elektronen und damit zu einer endlichen Stoß- bzw. Relaxationszeit τ Anlaß geben. In Kap. 7.3 haben wir zur Beschreibung der Elektronen im Rahmen einer Einelektron-Näherung die Bloch-Wellen eingeführt. Dies sind die Eigenfunktionen der Elektronen in einem periodischen Potential. Da Bloch-Wellen aufeinander senkrecht stehen, können in einem streng periodischen Potential keine Elektronstreuprozesse auftreten. Defekte und Phononen verursachen jedoch Abweichungen von der periodischen Kristallstruktur und rufen somit Streuprozesse hervor. Diese Aussage läßt sich ohne Einschränkung auf die Wellenpakete der semiklassischen Näherung übertragen. Wie wir am Ende dieses Abschnitts sehen werden, können auch Elektronen untereinander stoßen, doch sind diese Elektron-Elektron-Stöße für die meisten Transporteigenschaften ohne größere Bedeutung.

In der folgenden, rein qualitativen Behandlung der Streuprozesse wollen wir nicht auf die quantenmechanische Herleitung der Relaxationsraten eingehen, sondern nur einige allgemeingültige Überlegungen anstellen. Der Einfachheit halber behandeln wir die Elektronen als freie Teilchen. Würden wir sie durch Bloch-Funktionen darstellen, so ergäben sich gegenüber dem einfachen Bild freier Elektronen quantitative Korrekturen, doch sind die grundlegenden Aussagen davon nicht beeinträchtigt.

a) Streuung an Defekten

Zunächst wollen wir uns mit statischen Gitterstörungen wie Punktdefekten, Fremdatomen oder Versetzungen beschäftigen. Nach der Streutheorie, die in ihren Grundzügen in den Abschnitten 2.3 und 5.4 behandelt wurde, werden Wellen an statischen Streuzentren elastisch gestreut. Beim elastischen Stoß der Elektronen, die wir in der semiklassischen Näherung als Wellenpakete beschreiben, ändert sich zwar die Richtung, nicht aber der Betrag ihres Wellenvektors, d.h. die Energie der Elektronen bleibt beim Streuprozeß erhalten. Bezeichnen wir im weiteren den Elektronenimpuls vor bzw. nach dem Stoß mit $\hbar\boldsymbol{k}$ und $\hbar\boldsymbol{k}'$, so läßt sich die Erhaltung des Wellenvektors durch $\boldsymbol{k} = \boldsymbol{k}' + \boldsymbol{K}$ ausdrücken, wobei $\boldsymbol{K}$ ein beliebiger Wellenvektor ist, für dessen Betrag natürlich $K \leq 2k$ gilt. Der damit verbundene Impulsübertrag $\hbar\boldsymbol{K}$ wird vom Gitter als Ganzes aufgenommen. In Abb. 8.6 ist der elastischen Streuprozeß schematisch dargestellt. Neben der Fermi-Kugel ist auch die Brillouin-Zone eines kubisch primitiven Gitters eingezeichnet, wobei die Größenverhältnisse in etwa denen eines einfachen Metalles entsprechen.

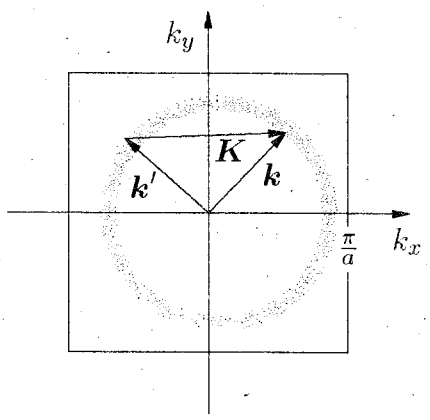


Abb. 8.6: Elektronstreuung an einem Defekt. Die Wellenvektoren $\boldsymbol{k}$ und $\boldsymbol{k}'$ des Elektrons liegen vor und nach dem Stoß innerhalb der aufgeweichten Fermi-Fläche. Das Gitter nimmt den Impuls $\hbar\boldsymbol{K}$ auf. Das Zentrum der Fermi-Kugel ist nach rechts verschoben.

Durch das elektrische Feld ist die Fermi-Kugel um den Betrag δk_x verschoben, durch die Temperaturbewegung ist sie „aufgeweicht". Die Auswirkungen dieser beiden Effekte sind in dieser Darstellung stark übertrieben. Da das Elektron bei der elastischen Streuung keine Energie gewinnt und nach dem Streuprozeß einen freien Zustand vorfinden muß, können nur Elektronen innerhalb des aufgeweichten Bereiches der Fermi-Kugel stoßen.

b) Streuung an Phononen

Im Gegensatz zur Defektstreuung ist die Elektron-Phonon-Streuung inelastisch und läßt sich analog zu der in Kap. 5.4 diskutierten Neutronenstreuung behandeln. Da in allen Fällen $E_F \gg \hbar\omega_D$ ist, wird die Elektronenenergie durch die Erzeugung oder Vernichtung eines Phonons kaum verändert. Für die Wellenzahlen gilt daher auch hier in guter Näherung $k \approx k'$, so daß wieder nur Elektronen an der Fermi-Fläche am Stoßprozeß teilnehmen können. Für den Wellenvektor selbst gilt der Erhaltungssatz $\boldsymbol{k} - \boldsymbol{k}' \pm \boldsymbol{q} = \boldsymbol{G}$, wobei $\boldsymbol{G}$ ein reziproker Gittervektor ist. Das Vorzeichen von $\boldsymbol{q}$ hängt davon ab, ob ein Phonon erzeugt oder vernichtet wird. Wie bei der Phonon-Phonon-Streuung können auch hier zwei Prozesse unterschieden werden: der Normalprozeß ohne Beteiligung eines reziproken Gittervektors und der Umklapp-Prozeß, bei dem $\boldsymbol{G} \neq 0$ ist.

α) Normalprozeß: Normalprozesse wirken sich auf die elektrische Leitfähigkeit je nach Probentemperatur unterschiedlich stark aus. Wie in Abb. 8.7 dargestellt[6], hängt der mittlere Streuwinkel ϑ der Elektronen von der Temperatur ab, da der mittlere Impulsübertrag von der Frequenz der dominanten Phononen bestimmt wird. Die Streuung um kleine Winkel bei tiefen Temperaturen trägt kaum zum elektrischen Widerstand bei, da viele Streuereignisse erforderlich sind, um ein Elektron von der Vorderseite der Fermi-Kugel auf die Rückseite zu bringen. Die „effektive" Stoßzeit, d.h. die Zeit, die verstreicht, bis dieser Vorgang abgeschlossen ist, und die nach Gleichung (8.34) die elektrische Leitfähigkeit bestimmt, ist in diesem Fall relativ groß. Mit zunehmender Temperatur wächst der Streuwinkel an, und die Normalprozesse werden immer effektiver.

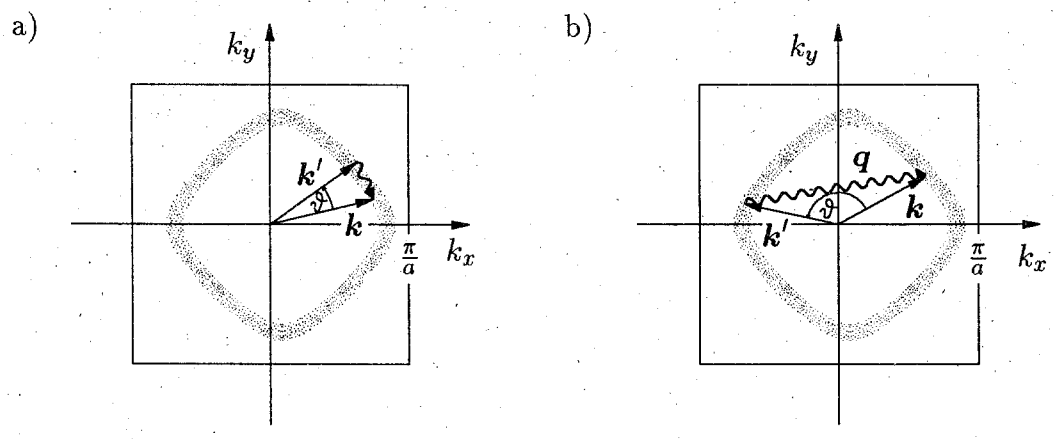


Abb. 8.7: Normalprozesse, bei denen ein Phonon erzeugt wird. Das Zentrum der Fermi-Flächen ist jeweils gegenüber dem Koordinatenursprung nach rechts verschoben.
a) Streuprozeß bei tiefen Temperaturen mit kleinem Impulsübertrag,
b) Streuung von der Vorder- auf die Rückseite der Fermi-Kugel durch *ein* Streuereignis bei hohen Temperaturen.

[6] In dieser und den folgenden Zeichnungen wurde nicht eine Fermi-Kugel gewählt, sondern eine von der Kugelgestalt abweichende Fermi-Fläche gezeichnet, wie sie bei vielen Metallen tatsächlich auftritt.

β) Umklapp-Prozeß: Zu einem Umklapp-Prozeß kommt es, wenn das Elektron durch die Erzeugung oder Vernichtung eines Phonons aus der 1. Brillouin-Zone hinausgestreut wird. Ähnlich wie beim Phonon-Umklapp-Prozeß ist hierfür ein Mindestimpuls des beteiligten Phonons erforderlich. Da die Fermi-Fläche in vielen Fällen in der Nähe der Brillouin-Zonengrenze liegt, kann der Impuls bzw. die Energie des benötigten Phonons, abhängig von der Form der Fermi-Fläche, relativ klein sein. Wie Abb. 8.8 zu entnehmen ist, bringt die Addition eines reziproken Gittervektors das Elektron wieder auf die Fermi-Fläche zurück. Umklapp-Prozesse können zu großen Richtungsänderungen des Elektronenimpulses selbst dann führen, wenn der Impuls des beteiligten Phonons so klein ist, daß Normalprozesse nur kleine Winkeländerungen bewirken würden. Diese Prozesse sind deshalb ein besonders effektiver Mechanismus zur Einstellung des stationären Gleichgewichts beim elektrischen Ladungstransport. In Abb. 8.8 sind zwei verschiedene Darstellungsweisen des gleichen Vorgangs gezeigt, nämlich die Darstellung im reduzierten und im periodisch wiederholten Zonenschema. Bei der letzteren kommen die physikalischen Vorgänge, die beim Streuprozeß auftreten, deutlicher zum Ausdruck.

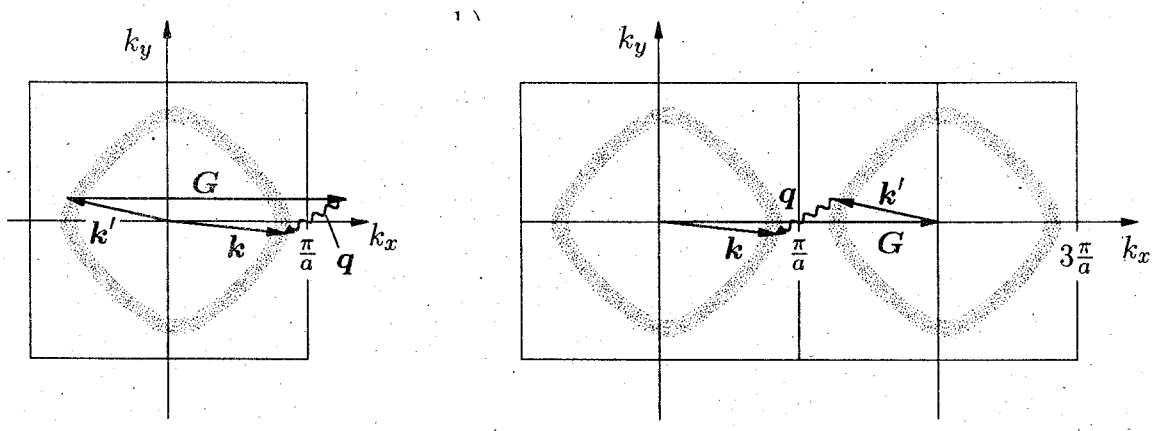


Abb. 8.8: Umklapp-Prozeß. Das wechselwirkende Elektron erzeugt ein Phonon und wird unter Mitwirkung eines reziproken Gittervektors auf die „Rückseite" der Fermi-Kugel gestreut. Das Zentrum der Fermi-Flächen ist jeweils gegenüber dem Koordinatenursprung nach rechts verschoben.
a) Darstellung im reduzierten Zonenschema,
b) Darstellung im periodisch wiederholten Zonenschema.

c) Elektron-Elektron-Streuung

Bisher haben wir die Wechselwirkung *zwischen* den Elektronen immer außer acht gelassen, da ohne Störung der Periodizität des Potentials zwischen den orthogonalen Bloch-Wellenfunktionen keine Wechselwirkung auftritt. Diese Argumentation ist nur im Rahmen der Einelektron-Näherung gültig. Geht man zu einer Vielteilchenbeschreibung über, so ist durchaus Wechselwirkung zwischen Elektronen möglich. Sie führt zu einer Streuung aus oder in die Zustände der Einteilchennäherung. Bei der Stärke des Coulomb-Potentials und der hohen Elektronendichte in Metallen von etwa 1 Elektron pro Elementarzelle ist zu befürchten, daß die Streurate möglicherweise sogar höher als bei den bisher diskutierten Mechanismen ist. Es gibt aber zwei Gründe für die starke Unterdrückung der Elektron-Elektron-Streuung. Zum einen bewirkt die Abschirmung (vgl. Kap. 7.2) eine starke Reduktion der Reichweite des Coulomb-Potentials, zum anderen ist das Pauli-Verbot wirksam, dessen Konsequenzen wir uns kurz ansehen wollen.

Um die Wirkung des Pauli-Verbots zu zeigen, greifen wir den einfachsten und häufigsten Stoßprozeß, die Wechselwirkung zwischen zwei Elektronen, heraus. Wir betrachten der Einfachheit halber den Streuvorgang bei $T = 0$ und nehmen an, daß nur ein Elektron angeregt ist. Es hat also die Energie $E_1 > E_F$, während das zweite wechselwirkende Elektron die Energie $E_2 < E_F$ besitzt. Da die Streuung in unbesetzte Zustände erfolgt, muß die Energie E_1' bzw. E_2' der beiden Elektronen nach dem Stoß größer als die Fermi-Energie sein, d.h. es muß $E_1', E_2' > E_F$ gelten. Zusätzlich verlangt die Energieerhaltung

$$E_1 + E_2 = E_1' + E_2' \,. \tag{8.36}$$

Daraus ergibt sich die Einschränkung, daß E_2 und E_1' nahe der Fermi-Fläche liegen müssen, nämlich innerhalb einer Kugelschale, deren Dicke durch $|E_1 - E_F|$ vorgegeben ist. Dies bedeutet, daß von allen möglichen Zuständen nur der Bruchteil $(E_1 - E_F)^2/E_F^2$ als Endzustand zur Verfügung steht. Die Einschränkung für E_2' ist bereits in (8.36) enthalten und stellt keine weitere Einengung dar. Gehen wir zu endlichen Temperaturen, so besteht eine Aufweichung der Fermi-Kugel von der Größenordnung $k_B T$. Die obige Argumentation läßt sich auf diesen Fall übertragen, wobei an die Stelle von $(E_1 - E_F)$ der Bereich der thermischen Aufweichung tritt. Es ist daher zu erwarten, daß die Streurate $1/\tau_E$ für Elektron-Elektron-Stöße um den Faktor $(k_B T/E_F)^2$ gegenüber der Streurate reduziert ist, die ohne Gültigkeit des Pauli-Verbots für abgeschirmte Punktladungen auftreten würde.

Die Streurate läßt sich unter Berücksichtigung der Abschirmung berechnen. Man findet

$$\frac{1}{\tau_E} \approx \frac{1}{\hbar} \frac{(k_B T)^2}{E_F} \,. \tag{8.37}$$

Die Rate ist nicht wie erwartet proportional zu E_F^{-2} sondern zu E_F^{-1}. Dies liegt daran, daß der Streuquerschnitt und damit auch die Streurate von abgeschirmten Elektronen proportional zu r_{TF}^2, dem Quadrat der Thomas-Fermi-Abschirmlänge (vgl. Kap. 7.2), ist. Da weiterhin $r_{TF}^2 \propto E_F$ gilt, erhält man das oben angegebene Ergebnis. Setzen wir typische Zahlenwerte ein, so finden wir für die Relaxationszeit bei Zimmertemperatur den Wert $\tau_E \approx 10^{-10}$ s, also einen Wert, der um etwa vier Größenordnungen über der experimentell beobachteten Stoßzeit liegt. Wir werden deshalb die Elektron-Elektron-Streuung bei unserer weiteren Diskussion vernachlässigen. Wenn überhaupt, so spielt die Elektron-Elektron-Wechselwirkung erst bei tiefen Temperaturen eine merkliche Rolle.

Temperaturabhängigkeit der elektrischen Leitfähigkeit

Wir beginnen die Diskussion der elektrischen Leitfähigkeit mit einem Blick auf die Temperaturvariation der Leitfähigkeit von Kupfer. Wie in Abb. 8.9 zu sehen, ist sie bei tiefer Temperatur nahezu konstant, nimmt dann sehr rasch ab und fällt bei Zimmertemperatur näherungsweise mit $1/T$ ab. Eingezeichnet ist auch die Leitfähigkeit der Legierung „Manganin“. Sie hängt nur sehr schwach von der Temperatur ab und ist immer sehr viel kleiner als die von Kupfer. Betrachten wir die Gleichung für die elektrische Leitfähigkeit (8.34), so sehen wir, daß die Stoßzeit $\tau(k_F)$ die einzige Größe ist, die eine nennenswerte Temperaturabhängigkeit aufweisen kann. Um den Temperaturverlauf der Leitfähigkeiten verstehen zu können, müssen wir deshalb die Temperaturabhängigkeit der soeben diskutierten Streuprozesse etwas näher ansehen.

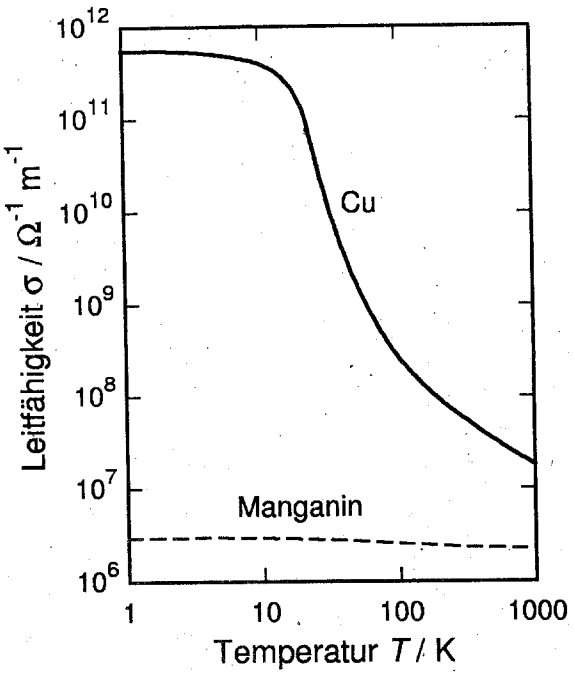


Abb. 8.9: Temperaturabhängigkeit der elektrischen Leitfähigkeit von reinem Kupfer (durchgezogene Kurve) und der Legierung „Manganin" (gestrichelt).

Wie bereits erwähnt, wird die Relaxationszeit τ durch die Streuung der Leitungselektronen an Defekten und Gitterschwingungen bestimmt, die Elektron-Elektron-Streuung kann vernachlässigt werden. Da die beiden wirksamen Streumechanismen unabhängig voneinander sind, ergibt sich die wirksame Streurate aus der Summe der Einzelraten:

$$\frac{1}{\tau} = \frac{1}{\tau_{\mathrm{D}}} + \frac{1}{\tau_{\mathrm{G}}(T)} \,. \tag{8.38}$$

Der elektrische Widerstand $\varrho = 1/\sigma$ kann deswegen in zwei Teile aufgespalten werden. Ein Teil beruht auf der Defektstreuung, der andere wird durch die thermische Bewegung des Gitters hervorgerufen:

$$\varrho = \frac{m^*}{ne^2\tau} = \varrho_{\mathrm{D}} + \varrho_{\mathrm{G}} = \frac{m^*}{ne^2\tau_{\mathrm{D}}} + \frac{m^*}{ne^2\tau_{\mathrm{G}}(T)} \,. \tag{8.39}$$

Dieser empirische Zusammenhang wird als *Matthiesensche Regel* bezeichnet.

Als Bestätigung der Matthiesensche Regel ist in Abb. 8.10 der Temperaturverlauf des elektrischen Widerstandes von drei Natriumproben gezeigt. Für $T \to 0$ tritt ein temperaturunabhängiger *Restwiderstand* auf, dessen Größe durch die in der Probe vorhandenen Defekte bestimmt wird. Da die Phononenstreuung bei tiefen Temperaturen ausstirbt, wird der Widerstand nur noch von den Defekten in den Kristallen hervorgerufen. Das Widerstandsverhältnis $\varrho\,(300\,\mathrm{K})/\varrho\,(4{,}2\,\mathrm{K})$ ist ein Maß für die Anzahl der streuenden Defekte und somit auch für die Reinheit und Qualität der untersuchten Kristalle. In elementaren Metallen läßt sich leicht ein Verhältnis 1000:1 erreichen. Dagegen ist bei stark legierten und amorphen Metallen $\varrho\,(300\,\mathrm{K})/\varrho\,(4{,}2\,\mathrm{K}) \approx 1$, d.h. es tritt nur eine schwache Temperaturabhängigkeit auf. Dort bewirkt die strukturelle und substitutionelle Unordnung eine so starke Defektstreuung, daß die Phononstreuung bedeutungslos wird. Typische Widerstandswerte für amorphe (und auch flüssige) Metalle liegen im Bereich von 1 bis $5\,\mu\Omega\mathrm{m}$.

Als Beispiel hierfür ist in Abb. 8.9 die Leitfähigkeit der Legierung Manganin[7] eingezeichnet. Der Widerstand weist eine sehr schwache Temperaturabhängigkeit auf und besitzt bei Zimmertemperatur den Wert $\varrho \approx 4{,}2\,\mu\Omega\mathrm{m}$.

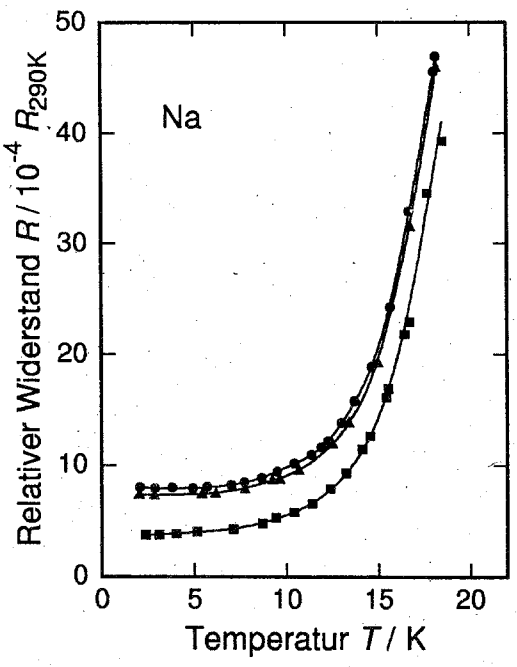


Abb. 8.10: Temperaturverlauf des elektrischen Widerstands von drei Natriumproben mit unterschiedlichem Restwiderstand. (Nach D.K.C. MacDonald und K. Mendelssohn, Proc. Roy. Soc. London **202** (1950) 103).

Bei hohen Temperaturen ($T > \Theta$) dominiert in reinen Metallen die Streuung der Elektronen an Phononen. Für die inverse freie Weglänge der gestreuten Elektronen gilt die bereits wiederholt benutzte Beziehung[8] $l^{-1} = n\sigma_{\mathrm{st}}$. Die wechselwirkenden Elektronen weisen die Fermi-Energie auf und werden vorwiegend an Phononen mit der Debye-Frequenz gestreut, die bei hohen Temperaturen in großer Zahl angeregt sind. Der Streuquerschnitt σ_{st} ist deshalb für diesen Prozeß näherungsweise konstant. Die inverse freie Weglänge steigt mit der Phononendichte n (vgl. Kap. 5.5) proportional zur Temperatur an, d.h. $l^{-1} \propto T$. Da die Stoßzeit über $l = \tau v_{\mathrm{F}}$ direkt mit der freien Weglänge verbunden ist, folgt für sie und die elektrische Leitfähigkeit bei Temperaturen $T > \Theta$ die Proportionalität

$$\sigma \propto \tau_{\mathrm{G}} \propto T^{-1}\,. \tag{8.40}$$

Betrachten wir nun den Verlauf der elektrischen Leitfähigkeit beim Abkühlen eines reinen Metalls. Wie in Abb. 8.9 für Kupfer gezeigt, beobachtet man nahe der Zimmertemperatur das soeben diskutierte typische $1/T$-Verhalten. Mit abnehmender Temperatur, d.h. im Bereich zwischen 100 K und 10 K steigt aber dann die Leitfähigkeit wesentlich steiler an. Die Ursache hierfür ist nicht nur die kleiner werdende Zahl der Phononen, die als Stoßpartner in Frage kommen. Mit abnehmender Temperatur nimmt auch der Streuquerschnitt ab, und die Effektivität der Streuprozesse verringert sich. Dies läßt sich folgendermaßen verstehen: Bei tieferen Temperaturen, d.h. bei $T < \Theta$, sind Phononen mit einer Frequenz dominant, die unterhalb der Debye-Frequenz liegt. Wie bei der Phonon-Phonon-Streuung (vgl. Kap. 6.2) ist der Streuquerschnitt σ_{st} bei der Elektron-Phonon-Streuung proportional zur Frequenz der beteiligten Phononen, so daß mit abnehmender

[7] Manganin besteht aus 86% Kupfer, 12% Mangan und 2% Nickel.

[8] Um Verwechslungen mit der Leitfähigkeit σ zu vermeiden, ist bei der Bezeichnung des Streuquerschnitts der Index st angefügt, d.h. wir schreiben hier σ_{st}.

Temperatur die Frequenz der dominanten Phononen und damit auch der Streuquerschnitt kleiner wird. Hinzu kommt, daß mit einer kleineren Phononenfrequenz ein kleinerer Impuls verbunden ist, wodurch die Winkeländerungen, die ein Stoß bewirken kann, ebenfalls kleiner werden. Der Übergang des elektronischen Wellenvektors von der Vorder- auf die Rückseite der Fermi-Kugel erfordert deshalb eine zunehmende Zahl von Stößen. Formal läßt sich diese Abhängigkeit durch die Beziehung $\tau_{\mathrm{G}}^{-1} \propto n\,\eta\,\sigma_{\mathrm{st}}$ ausdrücken, wobei η die „Effektivität" des Streuprozesses, d.h. die bewirkte Winkeländerung widerspiegelt. Eine eingehende Analyse, bei der alle drei Effekte berücksichtigt werden, führt zu dem Zusammenhang $\tau_{\mathrm{G}}^{-1} \propto (T/\Theta)^5$, wobei Θ wieder für die Debye-Temperatur steht. Die gleiche Temperaturabhängigkeit erwartet man auch für den Widerstand. Dieses sogenannte *Bloch-Grüneisen*-Gesetz wird bei vielen reinen Metallen beobachtet. Zieht man beispielsweise von den in Abb. 8.10 gezeigten Daten für Natrium den Restwiderstand ab, so findet man tatsächlich den vorhergesagten Temperaturverlauf.

Thermische Leitfähigkeit der Metalle, Wiedemann-Franz-Gesetz

Nun wenden wir uns der thermischen Leitfähigkeit zu. Offen ist zunächst die Frage, ob der Wärmetransport in erster Linie auf dem Beitrag der Elektronen oder dem der Phononen beruht. Bei Zimmertemperatur dominieren offenbar Elektronen, da die tägliche Erfahrung zeigt, daß die Wärmeleitfähigkeit der Metalle fast immer die der Dielektrika übertrifft. Bei ganz tiefen Temperaturen bestreiten die Elektronen ebenfalls den Hauptanteil an der Wärmeleitung. Dies läßt sich aus der Temperaturabhängigkeit der spezifischen Wärme folgern. Da die elektronische spezifische Wärme linear mit der Temperatur anwächst, der Beitrag des Gitters dagegen proportional zu T^3 ist, gibt es bei sehr tiefen Temperaturen immer mehr angeregte Elektronen als Phononen, die für den Energietransport in Frage kommen. Sieht man von den beiden Grenzfällen ab, so ist es für die meisten Materialien sehr schwer, die Beiträge der Phononen und Elektronen voneinander zu trennen.

Wir betrachten zunächst nur den Wärmetransport durch die Elektronen, d.h. wir gehen davon aus, daß $\Lambda_{\mathrm{el}} \gg \Lambda_{\mathrm{ph}}$ ist. Bei der Berechnung dieses Beitrags zur thermischen Leitfähigkeit benutzen wir wieder die Analogie zur kinetischen Gastheorie, wie wir es bereits in Kap. 6.3 beim Gitter getan haben. Da der Wärmetransport – wie der Ladungstransport – nur durch Elektronen an der Fermi-Oberfläche erfolgt, setzen wir in (6.21) die spezifische Wärme C_V^{el} und die Fermi-Geschwindigkeit v_{F} ein und finden

$$\Lambda_{\mathrm{el}} = \frac{1}{3} C_V^{\mathrm{el}}\, v\, l = \frac{1}{3}\,\frac{\pi^2 n k_{\mathrm{B}}^2 T}{m^* v_{\mathrm{F}}^2}\, v_{\mathrm{F}}\, l\,. \tag{8.41}$$

Unter der Annahme, daß sowohl die elektrische als auch die thermische Leitfähigkeit durch die Elektronen bewirkt wird, kann man (8.41) und (8.35) vergleichen und findet

$$\frac{\Lambda_{\mathrm{el}}}{\sigma} = \frac{\pi^2}{3}\left(\frac{k_{\mathrm{B}}}{e}\right)^2 T = LT\,. \tag{8.42}$$

Dies ist das *Wiedemann-Franz-Gesetz*. Das Verhältnis der beiden Transportgrößen sollte durch die Temperatur und die universelle *Lorenz-Zahl* $L = 2{,}5{\cdot}10^{-8}\ \mathrm{V^2\,K^{-2}}$ bestimmt sein. Dieses bemerkenswerte Gesetz ist nur gültig, wenn die freien Weglängen in beiden Fällen durch den gleichen Streumechanismus bestimmt werden. In Abb. 8.11 ist zu sehen, daß bei hohen Temperaturen das Wiedemann-Franz-Gesetz erfüllt ist und auch der erwartete Wert der Lorenz-Zahl gefunden wird. Diese Feststellung gilt nicht nur für Kupfer, sondern

für die meisten Metalle. Mit abnehmender Temperatur nimmt die Lorenz-Zahl L ab und wird bei tiefen Temperaturen wieder konstant. Grund für die Temperaturabhängigkeit ist die unterschiedliche Wichtung der Streuprozesse beim elektrischen und thermischen Transport, die wir noch näher betrachten.

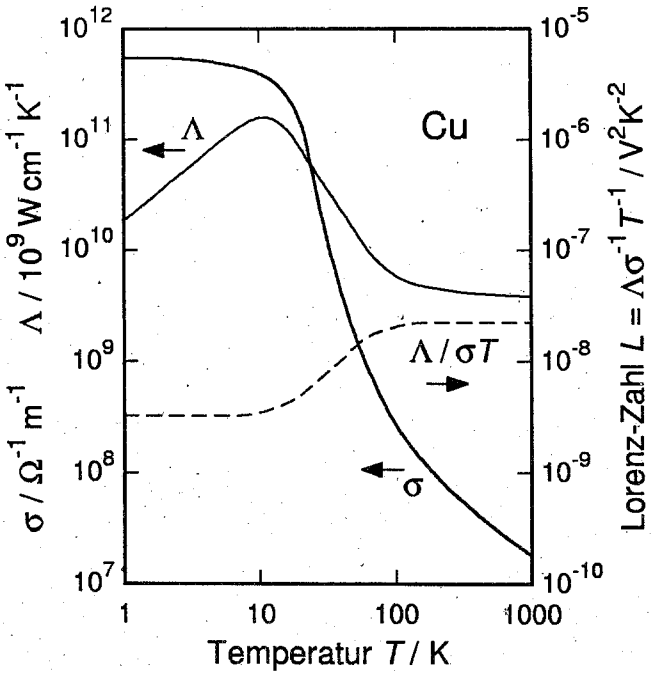


Abb. 8.11: Temperaturabhängigkeit der elektrischen (stark ausgezogene Kurve) und der thermischen (schwach ausgezogene Kurve) Leitfähigkeit von Kupfer. Die gestrichelte Kurve gibt den Temperaturverlauf der Lorenz-Zahl wieder. (Nach J. Blakemore, *Solid State Physics*, Saunders 1974).

Liegt ein elektrisches Feld an, so führt dieses nach (8.21) im stationären Gleichgewicht zu einer mittleren Änderung der Wellenvektoren um $\delta\boldsymbol{k} = -e\tau\boldsymbol{\mathcal{E}}/\hbar$ und damit zu einer Verschiebung der Fermi-Kugel, wie wir sie bereits anhand von Abb. 8.5 diskutiert haben. Formal folgt dies auch aus Gleichung (8.30), die verdeutlicht, daß $f(\boldsymbol{k})$ als der Beginn einer Taylor-Entwicklung nach dem Wellenvektor $\boldsymbol{k}$ um die Gleichgewichtsverteilung $f_0(\boldsymbol{k})$ aufgefaßt werden kann. Ein elektrisches Gleichfeld verändert nicht die Form der Fermi-Dirac-Verteilungsfunktion, sondern transformiert den Gleichgewichtswert $f_0(\boldsymbol{k})$ in den neuen Wert $f(\boldsymbol{k})$, wobei sich beide Funktionen nur in ihren Argumenten unterscheiden. Bezogen auf den Wellenvektor bzw. die Energie der Elektronen läßt sich diese Aussage durch den Zusammenhang

$$f(\boldsymbol{k}) = f_0\left(\boldsymbol{k} + \frac{e\tau\boldsymbol{\mathcal{E}}}{\hbar}\right) \qquad \text{bzw.} \qquad f(\boldsymbol{k}) = f_0\Big(E(\boldsymbol{k}) + e\tau\boldsymbol{v}(\boldsymbol{k})\cdot\boldsymbol{\mathcal{E}}\Big) \tag{8.43}$$

ausdrücken. Hierbei haben wir benutzt, daß die Änderung des Wellenvektors näherungsweise die Energieänderung $\delta E \approx -e\tau\boldsymbol{v}(\boldsymbol{k})\cdot\boldsymbol{\mathcal{E}}$ bewirkt.

Laufen die Elektronen dem Feld entgegen, so gewinnen sie (wegen ihrer negativen Ladung) Energie, während die auf der „Rückseite" der Fermi-Kugel gebremst werden. Insgesamt verschiebt sich der Nullpunkt der Fermi-Kugel, und die Energiewerte verändern sich, wie in Abb. 8.12 dargestellt, um den Betrag $\pm\delta E$, abhängig von der Bewegungsrichtung. Tatsächlich handelt es sich aber nicht um eine statische Verschiebung, sondern um einen dynamischen Prozeß, bei dem die Elektronen so lange im Impulsraum verschoben werden, bis sie entlang der Fermi-Fläche von der Vorder- auf die Rückseite gestreut werden.

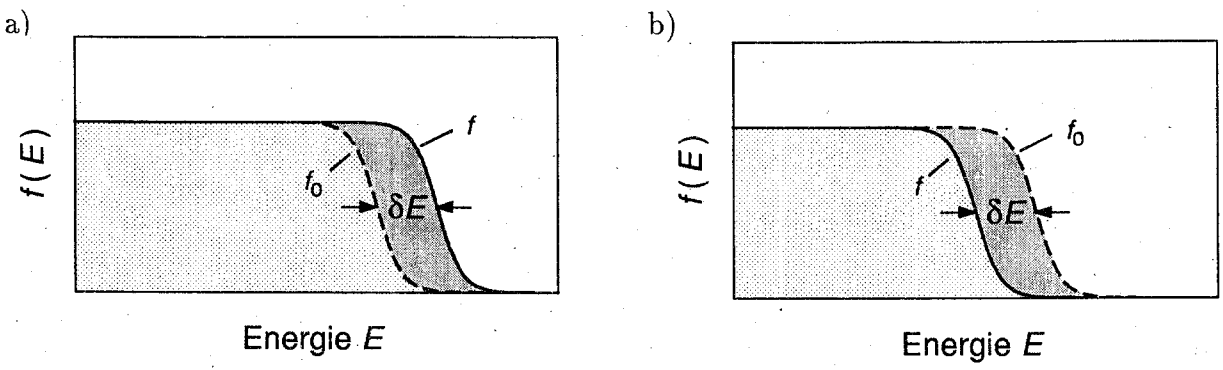


Abb. 8.12: Fermi-Dirac-Verteilung $f(E)$ bei Anliegen eines elektrischen Feldes. Die Gleichgewichtsverteilung $f_0(E)$ ohne Feld ist gestrichelt eingezeichnet, die Differenz zwischen den beiden Verteilungen ist dunkel schattiert hervorgehoben.
a) Die Energie der Elektronen auf der „Vorderseite" erhöht sich um den Betrag δE.
b) Auf der „Rückseite" vermindert sich die Elektronenenergie um Wert δE.

Fließt Wärme durch die Probe, d.h. herrscht ein Temperaturgradient, dann hängt die Geschwindigkeitsverteilung der Elektronen von der Bewegungsrichtung ab. Die Elektronen, die vom warmen Ende kommend sich in Richtung des Wärmestromes bewegen, sind schneller als diejenigen, die in entgegengesetzte Richtung laufen, da ihre „Temperatur" etwas höher ist. Um die zu (8.43) analoge Transformation der Fermi-Dirac-Verteilungsfunktion zu finden, geht man von (8.23) aus, betrachtet aber nun den Diffusionsterm, der die räumliche Variation der Verteilungsfunktion berücksichtigt. Nach einigen Rechenschritten findet man ein zu (8.43) analoges Ergebnis:

$$f(T) = f_0\Big(T - \tau \boldsymbol{v}(\boldsymbol{k}) \cdot \nabla T\Big) \,. \tag{8.44}$$

Die Elektronen sind also aufgrund des vorhandenen Wärmestroms je nach Bewegungsrichtung um die Temperatur $\delta T = -\tau \boldsymbol{v}(\boldsymbol{k}) \cdot \nabla T$ „wärmer" bzw. „kälter". Die Verteilung der Elektronen, die vom „kalten" Ende der Probe kommen, besitzt eine steilere, die vom „warmen" Ende eine flachere Flanke. Diese Veränderung der Verteilungsfunktion ist für beide Fälle in Abb. 8.13 skizziert.

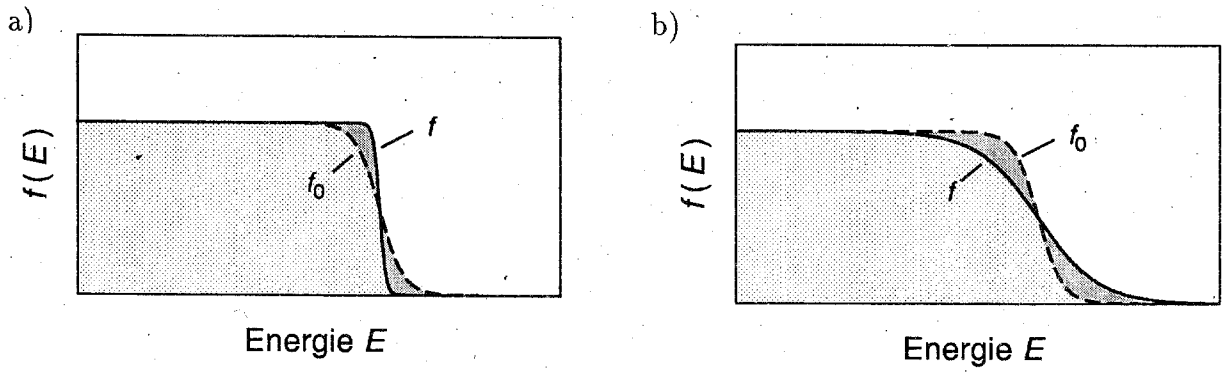


Abb. 8.13: Fermi-Dirac-Verteilung $f(E)$ bei Anliegen eines Temperaturgradienten. Die Gleichgewichtsverteilung $f_0(E)$ ohne Feld ist gestrichelt eingezeichnet, die Differenz zwischen den beiden Verteilungen ist besonders hervorgehoben.
a) Die Verteilung der Elektronen, die vom „kalten" Ende der Probe kommen, besitzt eine steilere Flanke als im thermischen Gleichgewicht.
b) Die Verteilung der Elektronen vom „warmen" Ende verläuft flacher.

Beim Anliegen eines elektrischen Feldes stellt sich das Gleichgewicht dadurch ein, daß Elektronen von der „Vorder-“ auf die „Rückseite“ der Fermi-Kugel gestreut werden. Die wirksamen Streuprozesse müssen daher vor allem den Impuls der Elektronen verändern. Beim Wärmetransport wird ein stationärer Zustand dadurch erreicht, daß „heiße“ Elektronen ihre überschüssige Energie durch Stöße verlieren und so zur Fermi-Fläche zurückkehren. Hier spielt die Bewegungsrichtung eine untergeordnete Rolle, denn es kommt in erster Linie auf die „richtige“ Temperatur, d.h. auf die Energie bzw. den Betrag des Impulses der Elektronen, an. Wie bereits in Abb. 8.7 zu sehen war, aber nicht besonders hervorgehoben wurde, kann man daher zwischen „vertikalen“ und „horizontalen“ Prozessen unterscheiden, die dort links bzw. rechts dargestellt sind. Als „vertikal“ bezeichnet man jene Prozesse, bei denen nur eine kleine Impulsänderung auftritt. Sie schwächen in erster Linie den Wärmetransport, denn sie kühlen die „heißen“ Elektronen innerhalb der aufgeweichten Fermi-Fläche. Bei vertikalen Prozessen erfolgen daher in erster Linie Übergänge von Zuständen, die außerhalb der Fermi-Kugel liegen, zu solchen, die sich näher an der Fermi-Kante befinden. Bei „horizontalen“ Prozessen erfolgt ein Übergang mit großem Impulsübertrag. Es wird mehr oder weniger die ganze Fermi-Kugel durchquert. Sie reduzieren sowohl die elektrische als auch die thermische Leitfähigkeit ganz erheblich. Es sind also die horizontalen Prozesse, auf denen das Wiedemann-Franz-Gesetz beruht, denn sie bestimmen beide Transportgrößen, wenn sie die Streuprozesse dominieren. Da sie nur bei hohen Temperaturen auftreten, ist es nicht verwunderlich, daß nur dort die Lorenz-Zahl den vorhergesagten Wert besitzt.

Wie bereits erwähnt, ist die Wärmeleitung der Metalle bei Zimmertemperatur meist wesentlich größer als die der Dielektrika. Als Beispiel ist in Abb. 8.14 nochmals der Temperaturverlauf der Wärmeleitung von Kupfer gezeigt, der eine gewisse Ähnlichkeit mit dem von dielektrischen Kristallen aufweist. Bemerkenswert ist, daß Metalle zwar bei Raumtemperatur eine höhere Wärmeleitfähigkeit als dielektrische Materialien aufweisen, der Maximalwert ihrer Wärmeleitfähigkeit aber kleiner als in reinen dielektrischen Kristallen ist.

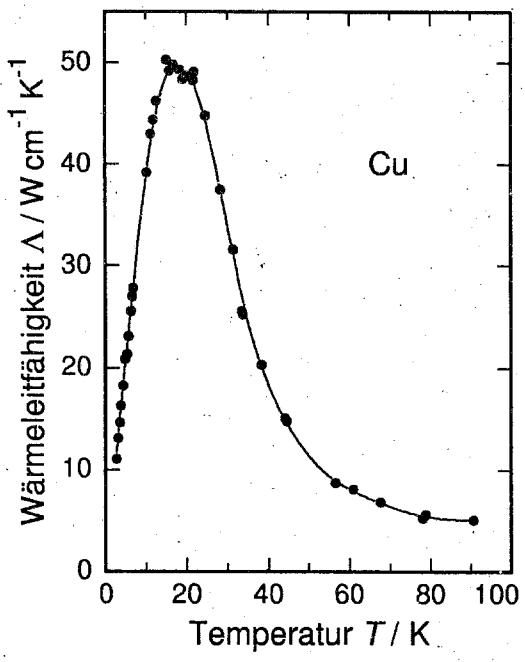


Abb. 8.14: Wärmeleitung von Kupfer. Ähnlich wie in dielektrischen Kristallen findet man auch in Metallen ein ausgeprägtes Maximum im Temperaturverlauf. (Nach R. Berman und D.K.C. MacDonald, Proc. Roy. Soc. London **211** (1952) 122).

Grob läßt sich der Verlauf bereits verstehen, wenn man nur den Beitrag der Elektronen betrachtet. Bei tiefen Temperaturen dominiert die Streuung der Elektronen an Verunreinigungen. Damit ist die freie Weglänge konstant, und die spezifische Wärme allein bestimmt nach (8.41) den Temperaturverlauf, d.h. die Wärmeleitfähigkeit steigt linear mit der Temperatur an. Mit zunehmender Temperatur tritt verstärkt Elektron-Phonon-Wechselwirkung auf, da die Zahl und die Frequenz der dominanten thermischen Phononen anwächst. Der Rückgang der freien Weglänge überkompensiert den Anstieg der spezifischen Wärme, und die Leitfähigkeit nimmt wieder ab. Oberhalb von 100 K wird die Wärmeleitfähigkeit von Metallen nahezu konstant. Bei den hohen Temperaturen sind Phononen mit der Debye-Frequenz dominant, deren Zahl näherungsweise proportional zur Temperatur ansteigt. Die freie Weglänge nimmt daher mit der inversen Temperatur ab und kompensiert so gerade den Anstieg der spezifischen Wärme in (8.41).

Auf den Beitrag der Phononen zur Wärmeleitung wollen wir nur kurz eingehen. Zunächst stellt sich die Frage, warum der Beitrag der Phononen zur Wärmeleitung in Metallen überhaupt so gering ist. Der Grund ist, daß in Metallen die Phononen nicht nur durch die Wechselwirkung untereinander gestreut werden, sondern vor allem durch die Wechselwirkung mit den Elektronen. Dieser Streumechanismus, der ja auch die elektrische Leitfähigkeit begrenzt, ist außerordentlich effektiv. Die in dielektrischen Festkörpern dominierende Phonon-Phonon-Streuung ist daher in Metallen meist ohne Bedeutung. Der verbleibende Beitrag der Phononen zur Wärmeleitung ist im allgemeinen schwer zu beobachten, da er vom Wärmetransport durch die Elektronen überdeckt ist. Eine Ausnahme stellen Legierungen oder stark verunreinigte Metalle dar, in denen die Elektronen an Fremdatomen und Defekten so stark gestreut werden, daß sie nur wenig zur Wärmeleitfähigkeit beitragen können. Ihre thermische Leitfähigkeit ist somit relativ klein.

8.3 Elektronenbahnen im Magnetfeld

Wir wollen uns nun dem Einfluß von Magnetfeldern auf die Elektronenbewegung in Metallen zuwenden. Von besonderem Interesse ist die Vermessung von Fermi-Flächen, die in Magnetfeldern möglich ist, und auf die wir ausführlich eingehen werden. Darüber hinaus gibt die zusätzliche Quantisierung der Bahnbewegung der Elektronen durch Magnetfelder zu überraschenden Effekten Anlaß.

Fermi-Flächen im periodischen Zonenschema

Bevor wir auf Meßmethoden zur Bestimmung der Fermi-Flächen eingehen, wollen wir einführend die zu erwartende Form dieser Flächen und die Art der Darstellung etwas genauer ansehen. Die Diskussion wird zeigen, daß bei der Behandlung des Magnetfeldeinflusses auf die Elektronenbewegung nicht die Reduktion auf die erste Brillouin-Zone, sondern die periodische Wiederholung des Zonenschemas besonders zweckmäßig ist. Diese Art der Darstellung haben wir bereits in Kap. 5.4 kennengelernt, als wir in Abb. 5.14 die Brillouin-Zone wiederholten, um die Phononenenergie längs bestimmter Richtungen darzustellen. Analog kann die Fermi-Fläche im $\boldsymbol{k}$-Raum periodisch wiederholt werden, wie es in Abb. 8.15 zu sehen ist. Bei dieser Darstellung kommen die Strukturen der Fermi-Fläche an der Grenze der Brillouin-Zone besonders stark zur Geltung.

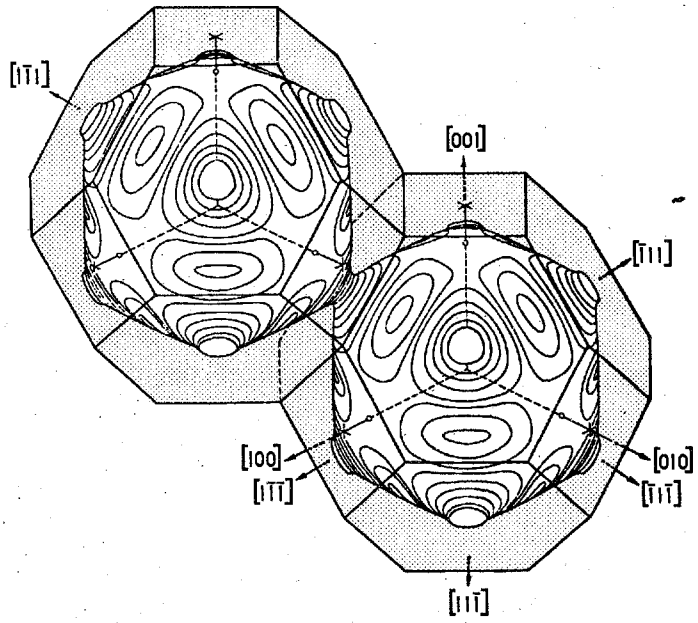


Abb. 8.15: Fermi-Fläche und Brillouin-Zone von Kupfer im periodisch wiederholten Zonenschema. Die Linien auf der Fermi-Fläche sind Kurven konstanter Wellenzahl. Man beachte das Auftreten von Ausstülpungen in den ⟨111⟩-Richtungen. (Nach W. Harrison, M.B. Webb (ed.), *The Fermi Surface*, New York 1960).

Die Verteilung der erlaubten Zustände im $\boldsymbol{k}$-Raum wird durch die Randbedingungen festgelegt, denen die Wellenfunktion der Elektronen unterliegt. Wie wir in Abschnitt 7.1 gefunden haben, ist ihre Dichte konstant. Da die Wechselwirkung der Elektronen mit dem Gitter diese Verteilung nicht ändert, ist das Volumen der „Fermi-Kugel" durch die Elektronendichte festgelegt. Die Form der Fermi-Fläche dagegen kann sehr unterschiedlich aussehen, denn sie wird gerade weitgehend von der Wechselwirkung bestimmt. Wie wir in Kap. 7.4 gesehen haben, bewirkt die Elektron-Gitter-Wechselwirkung beträchtliche Deformationen der Dispersionskurven an der Grenze der Brillouin-Zone. Symmetrieüberlegungen zeigen, daß die $E(\boldsymbol{k})$-Flächen, also insbesondere auch die Fermi-Flächen, die Grenzen senkrecht schneiden, wenn keine Entartung vorliegt. Es ist daher nicht verwunderlich, daß vor allem in der Nähe der Grenze der Brillouin-Zone Abweichungen von der Kugelgestalt besonders ausgeprägt sind.

Weitgehend kugelförmig ist die Fermi-Fläche nur bei den Alkalimetallen. So weist sie im Falle von Kalium eine Anisotropie von weniger als 1% auf, besitzt also fast exakt Kugelgestalt. Die Fermi-Fläche der Alkalimetalle liegt vollständig innerhalb der ersten Brillouin-Zone und ist weit von deren Grenze entfernt. Bandverbiegungen durch die Wechselwirkung mit dem Gitter spielen daher keine Rolle. In Abb. 8.15 ist die Fermi-Fläche von Kupfer gezeigt, die für Edelmetalle typisch ist. Sie hat die Gestalt einer Kugel, auf die noch Ausbuchtungen aufgesetzt sind. Da sich bei den Edelmetallen bereits im Modell freier Elektronen die Fermi-Fläche in den ⟨111⟩-Richtungen in der Nähe der Grenze der Brillouin-Zone befindet, verursacht die Wechselwirkung diese charakteristischen Ausstülpungen. In der periodischen Darstellung werden daraus *Hälse*, die die Fermi-Flächen der verschiedenen Brillouin-Zonen verbinden. Durch diese Verbiegung der Fermi-Fläche werden Zustände bei Energien geschaffen, die etwas tiefer liegen als es bei „regulärem" Einbau der Fall wäre. Die Wechselwirkung mit dem Gitter verursacht also, verglichen mit der Energie des freien Elektronengases, eine Energieabsenkung (vgl. Kap. 7.4).

Es gibt eine Reihe von experimentellen Methoden, die Fermi-Fläche zu untersuchen. Hierzu gehören u.a. die Messung des anomalen Skineffekts, der Zyklotronresonanz, des Magnetowiderstandes sowie des De Haas-Van Alphen-Effekts und der Ultraschallausbreitung im Magnetfeld. Hier sollen Zyklotronresonanz und De Haas-Van Alphen-Effekt näher diskutiert werden.

Zyklotronresonanz

Ein „klassisches" Experiment, um Auskunft über die Elektronen an der Fermi-Fläche zu erhalten, ist die Zyklotronresonanz. Wirkt auf ein Elektron ein Magnetfeld, so folgt aus den beiden Gleichungen (8.3) und (8.4) für die Lorentz-Kraft der Ausdruck

$$\dot{\boldsymbol{k}} = \frac{-e}{\hbar^2}\left[\nabla_{\boldsymbol{k}} E(\boldsymbol{k}) \times \boldsymbol{\mathcal{B}}\right] . \tag{8.45}$$

Das Elektron erfährt eine Kraft, die senkrecht zum Magnetfeld und im $\boldsymbol{k}$-Raum senkrecht zum Gradienten der $E(\boldsymbol{k})$-Fläche steht. Es bewegt sich somit auf einer Fläche konstanter Energie. Dies kann man natürlich auch aus Gründen der Energieerhaltung schließen, da ein zeitlich konstantes Magnetfeld keine Energieänderung des abgelenkten Teilchens bewirkt. Insbesondere bleibt ein Elektron bei seiner Bewegung auf der Fermi-Fläche.

Aufgrund der Symmetrie der Fermi-Fläche erfolgt die Bewegung oft längs einer geschlossenen Bahn, wie dies in Abb. 8.16 für eine $E(\boldsymbol{k})$-Fläche mit der willkürlich gewählten Gestalt eines Rotationsellipsoids dargestellt ist. Die Änderung $\mathrm{d}\boldsymbol{k}$ des Wellenvektors eines herausgegriffenen Elektrons steht senkrecht auf dem Magnetfeld und verläuft in Richtung der Kurve konstanter Energie. Der Vektor $\mathrm{d}\boldsymbol{k}$ ist proportional zu $\nabla_{\boldsymbol{k}} E(\boldsymbol{k}) \times \boldsymbol{\mathcal{B}}$, wobei wegen des Kreuzproduktes nur die Komponente $[\nabla_{\boldsymbol{k}} E(\boldsymbol{k})]_\perp$ des Gradienten wirksam ist, die senkrecht zum Magnetfeld steht. Führen wir noch den Vektor $\mathrm{d}\boldsymbol{k}_\perp$ ein, der senkrecht auf $\mathrm{d}\boldsymbol{k}$ und $\boldsymbol{\mathcal{B}}$ steht und in Abb. 8.16b ebenfalls eingezeichnet ist, so läßt sich $[\nabla_{\boldsymbol{k}} E(\boldsymbol{k})]_\perp$ vereinfacht als $\mathrm{d}E/\mathrm{d}k_\perp$ ausdrücken.

Wir berechnen nun die Zeit $\widetilde{T}$, die ein Elektron für einen Bahnumlauf benötigt. Dazu lösen wir (8.45) nach der Zeit auf, integrieren die Gleichung und erhalten

$$\widetilde{T} = \oint \mathrm{d}t = \frac{\hbar^2}{e\mathcal{B}} \oint \frac{|\mathrm{d}\boldsymbol{k}|}{|\mathrm{d}E/\mathrm{d}k_\perp|} . \tag{8.46}$$

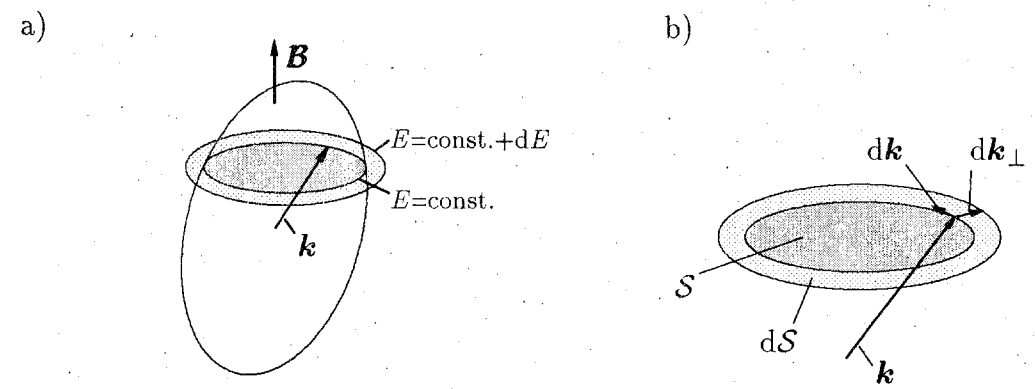


Abb. 8.16: Elektronenbewegung im Magnetfeld.
a) Umlaufbahn eines Elektrons auf einer Fläche konstanter Energie mit der Gestalt eines Rotationsellipsoids.
b) Bahnebene mit der Querschnittsfläche $\mathcal{S}$ und der Festlegung von $\mathrm{d}\boldsymbol{k}$ und $\mathrm{d}\boldsymbol{k}_\perp$.

Aus der Skizze 8.16b läßt sich unmittelbar ablesen, daß für die Querschnittsfläche $\mathcal{S}$ im $\boldsymbol{k}$-Raum die Beziehung $\mathrm{d}\mathcal{S} = \mathrm{d}k_\perp \oint |\mathrm{d}\boldsymbol{k}|$ gilt. Umlaufintegral und Querschnittsfläche sind deshalb durch die Gleichung

$$\oint \left| \frac{\mathrm{d}k_\perp}{\mathrm{d}E} \right| |\mathrm{d}\boldsymbol{k}| = \frac{\mathrm{d}\mathcal{S}}{\mathrm{d}E} \tag{8.47}$$

miteinander verbunden. Setzen wir diese Beziehung in (8.46) ein, so erhalten wir schließlich für die Umlaufzeit

$$\widetilde{T} = \frac{\hbar^2}{e\mathcal{B}} \frac{\mathrm{d}\mathcal{S}}{\mathrm{d}E} . \tag{8.48}$$

Die Umlaufzeit wird durch die Energieabhängigkeit der Schnittfläche bestimmt, die im $\boldsymbol{k}$-Raum von der Bahn eingeschlossen wird. Im Modell des freien Elektronengases findet man unabhängig vom Wellenvektor für diese Größe $\mathrm{d}\mathcal{S}/\mathrm{d}E = 2\pi m/\hbar^2$. Dies bedeutet, daß alle Elektronen dieselbe Umlaufzeit besitzen. Die Umlauffrequenz ω_c ist dann durch die klassische Zyklotronfrequenz $\omega_\mathrm{c} = e\mathcal{B}/m$ gegeben. Im allgemeinen werden aber verschiedene Elektronen unterschiedliche Umlaufzeiten aufweisen. Um den Ausdruck für die Zyklotronfrequenz auch im allgemeinen Fall benutzen zu können, ersetzt man die Masse m des Elektrons durch die effektive Masse m^*. Dabei ist zu beachten, daß diese sogenannte *Zyklotronmasse* nicht immer mit der dynamischen Masse der Elektronen identisch ist. Die Zyklotronmasse wird durch die Lage der betreffenden Bahn auf der Fermi-Fläche und nicht durch einen einzelnen elektronischen Zustand gemäß der Definitionsgleichung (8.6) bestimmt. Mit dieser Verallgemeinerung für die Zyklotronmasse des Elektrons gilt für die *Zyklotronfrequenz* weiterhin die wohlbekannte Formel

$$\omega_\mathrm{c} = \frac{e\mathcal{B}}{m^*} . \tag{8.49}$$

Bei der experimentellen Untersuchung der Zyklotronresonanz bringt man die Probe in ein Magnetfeld und strahlt zusätzlich Mikrowellen ein. Gemessen wird dann im allgemeinen die Absorption R der Mikrowellen oder deren Ableitung $\mathrm{d}R/\mathrm{d}\mathcal{B}$ als Funktion der Magnetfeldstärke[9]. Untersucht man Metalle, so benutzt man häufig eine Anordnung, wie in Abb. 8.17 skizziert, bei der Magnetfeld und elektrische Feldstärke der eingestrahlten Mikrowelle senkrecht aufeinander stehen.

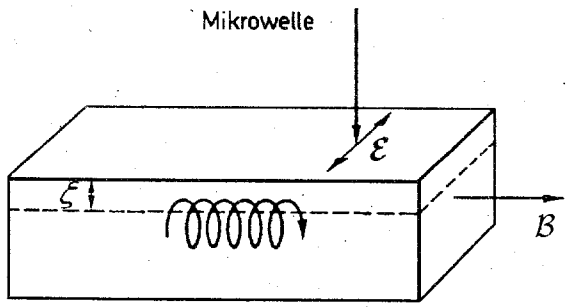


Abb. 8.17: Meßanordnung bei Zyklotronresonanzmessungen an Metallen. Magnetfeld und elektrisches Feld der eingestrahlten Mikrowellen sind durch Pfeile angedeutet. ζ ist die Eindringtiefe der Mikrowelle, die Spirale symbolisiert eine Elektronenbahn.

[9] Es liegt nahe, das Experiment bei konstantem Magnetfeld und variabler Mikrowellenfrequenz durchzuführen, doch ist dies meßtechnisch wesentlich aufwendiger.

Bei derartigen Experimenten muß man beachten, daß elektromagnetische Wellen kaum in Metalle eindringen können. Die Eindringtiefe ζ ist bei hohen Frequenzen meist kleiner als der Durchmesser der Elektronenbahnen, weshalb die umlaufenden Elektronen nur an der Oberfläche mit dem elektrischen Feld wechselwirken können und ihnen nur dort Energie zugeführt oder entzogen werden kann. Absorption tritt auf, wenn die Elektronen bei jedem Eintritt in die Wechselwirkungszone durch das elektrische Feld phasengerecht beschleunigt werden. Dies bedeutet, daß eine ganzzahlige Anzahl von Perioden der elektromagnetischen Schwingungen zwischen zwei Durchläufen verstreichen muß. Man beobachtet daher das Auftreten einer Resonanz bei der Frequenz ω_{mi} der eingestrahlten Mikrowellen, wenn diese ein ganzzahliges Vielfaches p der Zyklotronfrequenz ist, d.h. wenn

$$\omega_{\mathrm{mi}} = p\,\omega_{\mathrm{c}} = p\,\frac{e\mathcal{B}}{m^*} \tag{8.50}$$

ist. Bei konstanter Mikrowellenfrequenz werden mit zunehmender Magnetfeldstärke die Resonanzen immer ausgeprägter, und der zugehörige Wert von p immer kleiner. Aus der Lage der Resonanzen läßt sich unmittelbar die effektive Masse m^* der Elektronen bestimmen. Ergebnisse eines Zyklotronresonanzexperiments an Kupfer sind in Abb. 8.18 gezeigt. Die Messung wurde an einer (110)-Kupferoberfläche bei 35 GHz und einem statischen Magnetfeld in [100]-Richtung durchgeführt. Deutlich ist die erwartete Oszillation von $\mathrm{d}R/\mathrm{d}\mathcal{B}$ zu erkennen.

Scharfe Resonanzen erhält man nur, wenn die Elektronen mehrere Umläufe ungestört vollenden können, d.h. wenn die Bedingung $\omega_{\mathrm{c}}\tau \gg 1$ erfüllt ist, wobei τ für die mittlere Stoßzeit steht. Es ist daher günstig, die Experimente zur Zyklotronresonanz bei hohen Frequenzen und damit bei hohen Magnetfeldern durchzuführen. Um Streuung an Phononen und Defekten zu vermeiden und lange Stoßzeiten zu erreichen, muß bei tiefen Temperaturen und mit sehr reinen Substanzen gearbeitet werden. Werte für die Stoßzeiten von Elektronen in Kupfer wurden bereits in Kap. 8.2 angegeben. Aus Abb. 8.18 folgt, daß beim Anwachsen der Temperatur von 4,2 K auf 12,9 K die Phononen bereits eine erhebliche Reduktion der elektronischen Stoßzeit verursachen und damit eine Ausschmierung der Meßkurven bewirken.

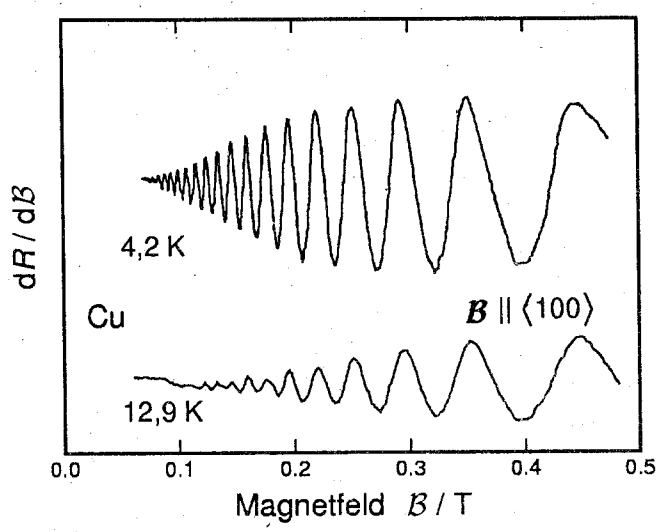


Abb. 8.18: Differentielle Mikrowellenabsorption $\mathrm{d}R/\mathrm{d}\mathcal{B}$ von Kupfer als Funktion des statischen Magnetfelds bei 4,2 K und bei 12,9 K. Mit zunehmender Temperatur werden die Oszillationen immer stärker verwaschen. (Nach P. Häussler und S.J. Welles, Phys. Rev. **152** (1966) 675).

Ist das untersuchte Energieband nicht parabolisch, so unterscheiden sich die effektiven Massen der Elektronen und damit auch ihre Umlaufzeit auf Flächen konstanter Energie. Da sich die Beiträge der einzelnen Elektronen überlagern, sollten bei realen Metallen – von wenigen Ausnahmen abgesehen – keine scharfen Resonanzen in der Absorption beobachtbar sein. Um zu verstehen, daß dennoch scharfe Maxima auftreten, müssen zwei Tatsachen berücksichtigt werden. Erstens können überhaupt nur Elektronen an der Fermi-Fläche Energie aufnehmen und zur Absorption beitragen, da nur dort unbesetzte Zustände zur Verfügung stehen. Doch diese Einschränkung reicht noch nicht aus, denn auch die effektive Masse der Elektronen auf der Fermi-Fläche variiert in den meisten Fällen. Der zweite wichtige Punkt ist, daß nur Elektronen auf sogenannten *Extremalbahnen* beobachtet werden können. Ganz allgemein findet man derartige Bahnen, wenn $\partial^2 S/(\partial k_\parallel \partial E) = 0$ ist. Dabei ist $k_\parallel$ die Komponente des Wellenvektors $\boldsymbol{k}$ in Richtung des Magnetfelds. Wie sich zeigen läßt, ist für diese Elektronen die Größe $\mathrm{d}S/\mathrm{d}E$ und damit die effektive Masse näherungsweise konstant. Innerhalb einer gewissen Bandbreite besitzen diese Elektronen daher nahezu die gleiche Umlauffrequenz und absorbieren somit bei der gleichen Mikrowellenfrequenz bzw. beim gleichen Magnetfeld. In Abb. 8.19 sind Beispiele für Extremalbahnen bei einer einfachen Fermi-Fläche gezeigt.

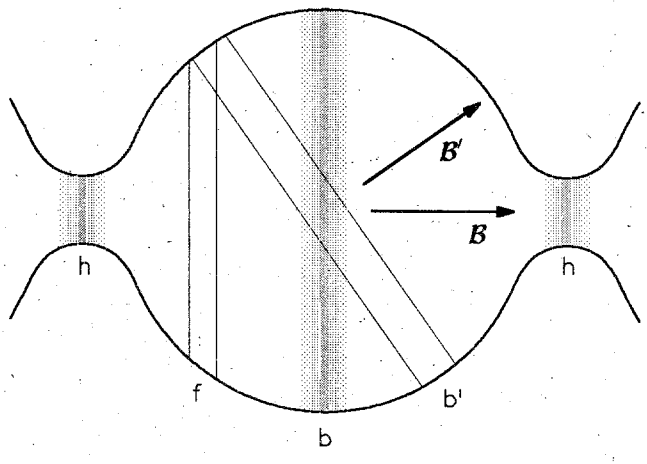


Abb. 8.19: Extremalbahnen. Liegt das Magnetfeld $\boldsymbol{B}$ an, so treten die *Hals-* und *Bauchbahnen* h und b auf. Die Bauchbahn b′ gehört zum Magnetfeld $\boldsymbol{B}'$. Die Bahn f ist keine Extremalbahn und ist daher nicht beobachtbar.

Verläuft das Magnetfeld parallel zur Rotationsachse, so beobachtet man Signale, die von einer *Hals-* und einer *Bauchbahn* herrühren. Ist das Magnetfeld etwas verkippt, so ist nur eine Bauchbahn beobachtbar. Die Elektronen auf den übrigen Bahnen absorbieren bei unterschiedlichen Feldern und tragen daher nur zu einem Untergrundsignal bei. Trägt man, wie in Abb. 8.18 geschehen, die differentielle Absorption auf, so wird der Untergrund unterdrückt. In der gezeigten Messung wird übrigens eine Bauchbahn von Kupfer beobachtet.

Elektronenbahnen

Die Bahnen, längs derer sich die Elektronen bewegen können, werden durch die Topologie der Fermi-Flächen bestimmt. Man unterscheidet zwischen *Elektronen-* und *Lochbahnen* und zwischen *offenen* und *geschlossenen* Bahnen. Die Bedeutung dieser Begriffe verdeutlichen die Abbildungen 8.20 bis 8.22.

In Abb. 8.20a bewegen sich die Elektronen auf einer „normalen“ Fermi-Kugel, ihre Bahn im Magnetfeld schließt die besetzten Zustände niedriger Energie ein. Wie bei freien Elektronen ist der Umlaufsinn der Bahn dem Uhrzeigersinn entgegengerichtet, und man

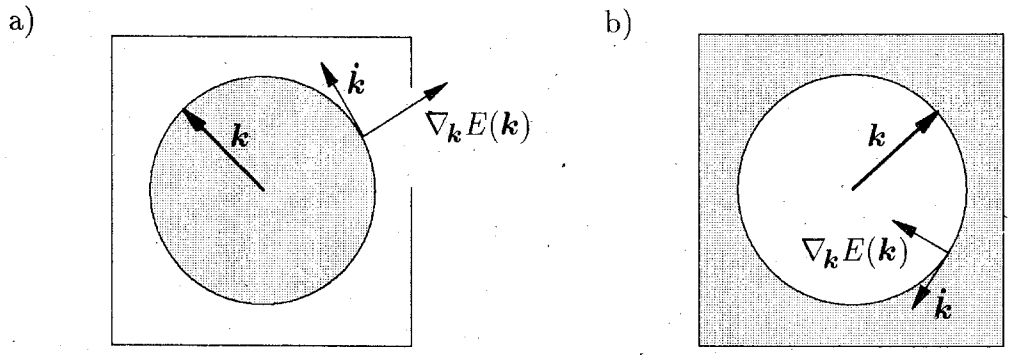


Abb. 8.20: Schematische Darstellung der Elektronenbewegung auf der Fermi-Fläche unter dem Einfluß eines Magnetfelds, das in beiden Fällen aus der Zeichenebene heraustritt. In den grauen Flächen sind mit Elektronenzustände besetzt.
a) Elektronbahn,
b) Lochbahn.

spricht von Elektronenbahnen. Liegen wie in Abb. 7.32 bei der 2. Brillouin-Zone die besetzten Zustände außen, wie es durchaus bei höher liegenden Bändern vorkommen kann, so kehrt sich die Richtung von $\nabla_{\boldsymbol{k}} E(\boldsymbol{k})$ und damit auch die Elektronenbewegung um. Solche Bahnen werden als Lochbahnen bezeichnet. In mehrfach zusammenhängenden Fermi-Flächen können Elektronen- und Lochbahnen gleichzeitig auftreten, wie dies die dreidimensionalen Darstellung in Abb. 8.21 zeigt.

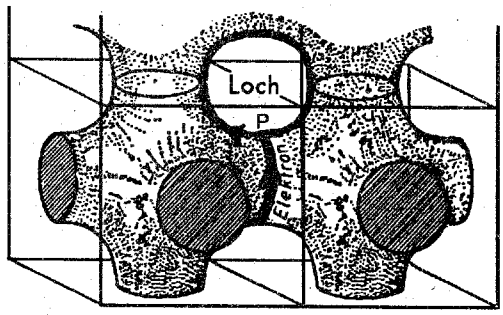


Abb. 8.21: Geschlossene Elektron- und Lochbahnen. Elektronen am Punkt P können je nach Magnetfeldrichtung beide Bahntypen durchlaufen. (Aus J.M. Ziman, *Prinzipien der Festkörpertheorie*, Akademieverlag, 1979).

Daneben gibt es auch offene Bahnen, die topologisch zwischen Elektronen- und Lochbahnen liegen. Ein Beispiel hierfür ist in Abb. 8.22 zu sehen. Bei offenen Bahnen treten natürlich keine Resonanzphänomene auf. Dies bedeutet, daß sie nicht in Experimenten zur Zyklotronresonanz beobachtet werden können. Sie tragen jedoch ganz wesentlich zum *Magnetowiderstand* bei, den wir hier nicht behandeln.

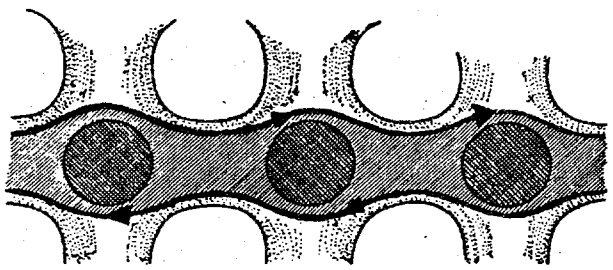


Abb. 8.22: Schnitt durch die oben gezeigte Fermi-Fläche. Es sind zwei offene Bahnen eingezeichnet, die neben den geschlossenen auftreten. (Aus J.M. Ziman, *Prinzipien der Festkörpertheorie*, Akademieverlag, 1979).

8.4 Quantisierung der Elektronenbahnen im Magnetfeld

Landau-Niveaus

Wir wollen nun auf die Quantisierung der Elektronenbahnen eingehen, die wir bisher ignoriert haben. Hierzu betrachten wir die Bewegung freier Elektronen in einem konstanten Magnetfeld $\boldsymbol{\mathcal{B}}$, das an einen Festkörper in z-Richtung anliegt. Bei der Bahnbeschreibung benutzen wir anstelle von Bloch-Wellenfunktionen das Modell des freies Elektronengases, um die Formeln nicht zu umfangreich werden zu lassen. Die Ergebnisse lassen sich aber unmittelbar auf Bloch-Wellen übertragen. Zur weiteren Vereinfachung lassen wir bei unserer Diskussion den Spin außer acht, da dessen magnetisches Moment die Bahnbewegung nicht beeinflußt, sondern nur einen Zusatzterm in der Energie verursacht. Diesen Effekt haben wir bereits in Kap. 7.2 bei der Pauli-Suszeptibilität separat behandelt.

Bei klassischer Beschreibung bewegen sich freie Elektronen auf Spiralbahnen entlang des Magnetfelds. Bei der quantenmechanischen Beschreibung wird das Magnetfeld über das Vektorpotential $\boldsymbol{A}$ im Hamilton-Operator berücksichtigt. Die stationäre Schrödinger-Gleichung nimmt dann die Form

$$\frac{1}{2m}(-\mathrm{i}\hbar\nabla + e\boldsymbol{A})^2\psi = E\psi \,. \tag{8.51}$$

an. Setzt man das Magnetfeld $\boldsymbol{\mathcal{B}} = (0,0,\mathcal{B})$ bzw. das Vektorpotential $\boldsymbol{A} = (0,\mathcal{B}x,0)$ ein und schreibt die Komponenten in Gleichung (8.51) aus, so findet man

$$\frac{\partial^2\psi}{\partial x^2} + \frac{\partial^2\psi}{\partial z^2} + \left(\frac{\partial}{\partial y} + \frac{\mathrm{i}e\mathcal{B}x}{\hbar}\right)^2\psi + \frac{2mE}{\hbar^2}\psi = 0 \,. \tag{8.52}$$

Wir wählen den von *Landau* stammenden Lösungsansatz $\psi = \widetilde{\psi}(x)\,\exp\left[\mathrm{i}(k_y y + k_z z)\right]$ für die Wellenfunktion, der für die y- und z-Richtung ebene Wellen beibehält, setzen den Ansatz ein und erhalten

$$\frac{\partial^2\widetilde{\psi}}{\partial x^2} + \left[\frac{2mE}{\hbar^2} - k_z^2 - \left(k_y + \frac{e\mathcal{B}x}{\hbar}\right)^2\right]\widetilde{\psi} = 0 \,. \tag{8.53}$$

Benutzen wir die Abkürzungen

$$E' = E - \frac{\hbar^2 k_z^2}{2m} \qquad \text{und} \qquad x' = x + \frac{\hbar k_y}{e\mathcal{B}} \,, \tag{8.54}$$

so vereinfacht sich die Gleichung mit der Zyklotronfrequenz $\omega_\mathrm{c} = e\mathcal{B}/m$ zu der Form

$$\frac{\partial^2\widetilde{\psi}}{\partial x'^2} + \frac{2m}{\hbar^2}\left[E' - \frac{1}{2}m\omega_\mathrm{c}^2 x'^2\right]\widetilde{\psi} = 0 \,. \tag{8.55}$$

Dies ist die Schrödinger-Gleichung eines linearen harmonischen Oszillators. Die Eigenfunktionen $\widetilde{\psi}(x)$ und Eigenwerte $E' = (\ell + 1/2)\,\hbar\omega_\mathrm{c}$ sind aus der Quantenmechanik wohlbekannt. Drücken wir die Energie der Elektronen mit den ursprünglichen Größen aus, so erhalten wir mit $\ell \geq 0$:

$$E = \left(\ell + \frac{1}{2}\right)\hbar\omega_\mathrm{c} + \frac{\hbar^2 k_z^2}{2m} \,. \tag{8.56}$$

Die elektronischen Bänder des Festkörpers spalten unter Einwirkung eines Magnetfelds in Subbänder auf, die man als *Landau-Niveaus* bezeichnet. In Abb. 8.23 ist die Elektronenenergie als Funktion der z-Komponente des Wellenvektors aufgetragen. Die parabolische Dispersionskurve der freien Elektronen spaltet in Subbänder auf, deren Eigenwerte sich um $\delta E = \hbar\omega_c$ unterscheiden. Legt man ein Feld von 2 Tesla an, so beträgt die Aufspaltung etwa $\delta E/k_B \approx 1\,\mathrm{K}$. Bei einer typischen Fermi-Temperatur $T_F \approx 10^5\,\mathrm{K}$ ist also eine sehr große Anzahl von Landau-Niveaus besetzt. Wie bereits erwähnt, erfahren die Niveaus durch die Einstellung der Elektronenspins eine zusätzliche Aufspaltung, die in Abb. 8.23 auf der rechten Seite ebenfalls eingezeichnet ist.

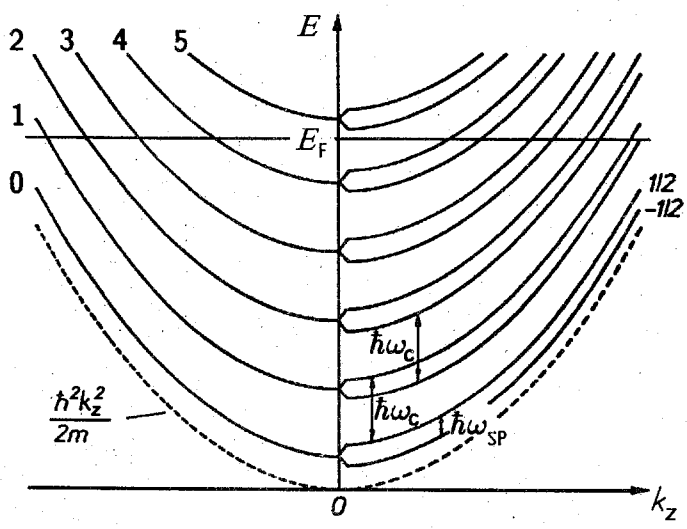


Abb. 8.23: Elektronenenergie im Magnetfeld als Funktion von der Wellenzahl k_z in Feldrichtung. Auf der rechten Seite ist die zusätzliche Aufspaltung $\hbar\omega_{sp}$ der Landau-Niveaus durch den Elektronenspin eingezeichnet. Die gestrichelte Kurve entspricht der Energie freier Elektronen ohne Magnetfeld.

Das Magnetfeld bewirkt eine Quantisierung der Elektronenenergie in der xy-Ebene, während in z-Richtung nichts verändert wird. Welche Konsequenzen ergeben sich daraus für die Bewegung der Elektronen? Zur anschaulichen Beschreibung nutzen wir das Korrespondenzprinzip, das besagt, daß im Grenzfall großer Quantenzahlen ℓ die klassische und quantenmechanische Beschreibung ineinander übergehen. Für die beobachtbaren Elektronen bei der Fermi-Energie sind die klassischen Bahnen sicherlich eine sehr gute Näherung, da diese Elektronen extrem große Quantenzahlen aufweisen. Die Bewegung der Elektronen auf einer Kreisbahn mit dem Radius r in der xy-Ebene können wir mit $x = x_0 + r\cos(\omega_c t)$ und $y = y_0 + r\sin(\omega_c t)$ beschreiben. Der Mittelpunkt der Kreisbewegung ist bei großen Quantenzahlen ℓ durch die Koordinaten x_0 und y_0 festgelegt. Aus (8.54) folgt unmittelbar die Koordinate $x_0 = (\hbar k_y/m\omega_c)$. Ohne Beweis setzen wir aus Symmetriegründen für y_0 den analogen Ausdruck $y_0 = (\hbar k_x/m\omega_c)$ ein. Diese Vorgehensweise wird plausibel, wenn man bedenkt, daß durch den Lösungsansatz die x-Koordinate ausgezeichnet wurde. Wir hätten genausogut die y-Koordinate wählen können und hätten dann den entsprechenden Ausdruck für y_0 erhalten. Den klassischen Bahnradius erhalten wir aus (8.55), wenn wir berücksichtigen, daß die Amplitude der linearen Auslenkung identisch mit dem Radius r_ℓ der Kreisbewegung ist. Es ergibt sich

$$r_\ell^2 = \left(\ell + \frac{1}{2}\right)\frac{2\hbar}{m\omega_c}\,. \tag{8.57}$$

Offensichtlich sind die Energieeigenwerte mit ganz bestimmten Bahnradien verknüpft. Aus (8.57) folgt unmittelbar, daß auch der von der Bahn umschlossene Magnetfluß Φ_ℓ quantisiert ist[10]:

$$\Phi_\ell = \pi r_\ell^2 \mathcal{B} = \left(\ell + \frac{1}{2}\right) \frac{h}{e} . \tag{8.58}$$

Auch im $\boldsymbol{k}$-Raum durchlaufen die Elektronen Kreisbahnen mit quantisierten Radien k_ℓ und schließen wohldefinierte Flächen $\mathcal{S}_\ell$ ein. Für den Betrag des Wellenvektors in der xy-Ebene folgt sofort

$$k_\ell^2 = k_x^2 + k_y^2 = \left(\frac{m}{\hbar}\right)^2 (\dot{x}^2 + \dot{y}^2) = \left(\ell + \frac{1}{2}\right) \frac{2m\omega_\mathrm{c}}{\hbar} \tag{8.59}$$

und für die im $\boldsymbol{k}$-Raum umschlossene Fläche $\mathcal{S}_\ell$

$$\mathcal{S}_\ell = \left(\ell + \frac{1}{2}\right) \frac{2\pi e\mathcal{B}}{\hbar} . \tag{8.60}$$

Da im Ortsraum die Fläche $A_\ell = \pi r_\ell^2 = (2\ell + 1)\pi\hbar/e\mathcal{B}$ umschlossen wird, gilt die Beziehung

$$A_\ell = \left(\frac{\hbar}{e\mathcal{B}}\right)^2 \mathcal{S}_\ell . \tag{8.61}$$

Mit steigendem Magnetfeld verkleinern sich die Kreisbahnen im Ortsraum, während sie im Impulsraum größer werden. Die Beziehung zwischen den in den verschiedenen Räumen umschlossenen Flächen gilt nicht nur für die Kreisbahnen der freien Elektronen, sondern ganz allgemein für geschlossene Bahnen.

Eine eingehendere Inspektion von (8.52) und des Lösungsansatzes zeigt, daß weder die Operatoren für x_0 und y_0 noch die für k_x und k_y vertauschen und ihre Erwartungswerte damit nicht mehr unabhängig voneinander festgelegt werden können. Dies bedeutet, daß das Magnetfeld einen Zusammenbruch des bisher verwendeten Quantisierungsschemas der Wellenvektoren verursacht. Die Größen k_x und k_y verlieren ihre Bedeutung als „gute" Quantenzahlen. Ohne Magnetfeld sind die erlaubten Wellenvektoren gleichmäßig im $\boldsymbol{k}$-Raum verteilt. Mit Magnetfeld befinden sie sich auf Kreisen, wie in Abb. 8.24b zu sehen ist, in der wir noch die alten Quantenzahlen k_x und k_y benutzen. Die Zustände auf den Kreisen sind nicht wirklich durch Punkte darstellbar. Sie rotieren mit der Kreisfrequenz ω_c, da sich bei der Kreisbewegung die Wellenvektoren k_x und k_y ständig ändern. Jeder Quantenzahl ℓ und damit auch jedem Landau-Niveau ist ein Kreis zugeordnet.

Stellen wir die erlaubten Zustände dreidimensional dar, so liegen sie im $\boldsymbol{k}$-Raum auf *Landau-Röhren*. Die Form der Röhren wird von der Energiefläche $E(\boldsymbol{k})$ bestimmt. Beim freien Elektronengas sind dies, wie Abb. 8.25 zeigt, einfache Zylinder. Die im konventionellen Quantisierungsschema erlaubten Zustände „kondensieren" auf der nächstliegenden Röhre. Alle Zustände, die sich auf einer Röhre befinden, sind bezüglich k_x und k_y entartet und besitzen die gleiche Quantenzahl ℓ. Der Wellenvektor k_z dagegen wird nach wie vor durch die geometrischen Randbedingungen vorgegeben.

[10] Die Gleichung (8.58) ist nur für freie Elektronen oder parabolische Bänder gültig. Im allgemeinen Fall steht in der Klammer anstelle von 1/2 ein „Phasenfaktor" γ.

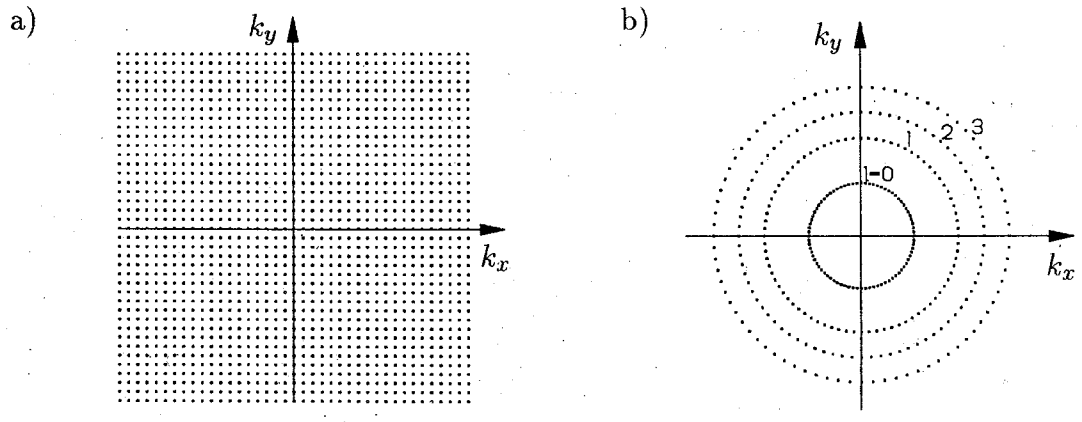


Abb. 8.24: Einfluß eines Magnetfelds auf die Eigenwerte der Wellenvektoren freier Elektronen in x- und y-Richtung. Die Komponenten k_z in Feldrichtung werden nicht beeinflußt.
a) Erlaubte Wellenvektoren ohne Magnetfeld,
b) erlaubte Wellenvektoren mit Magnetfeld.

Die Zahl der Zustände, die auf einem derartigen Zylinder Platz finden, ist begrenzt. Betrachten wir eine würfelförmige Probe mit der Kantenlänge L, so ist der Wellenvektor k_y in Einheiten von $2\pi/L$ quantisiert, da das Magnetfeld keinen Einfluß auf die Randbedingungen für k_y besitzt. Es gibt aber noch eine weitere Einschränkung: Die Mittelpunktkoordinate x_0 muß sich innerhalb der Probe befinden, d.h. es gilt die Bedingung $0 < x_0 < L$. Setzen wir $x_0 = (\hbar k_y/m\omega_c)$ dort ein, so ergibt sich ein maximaler Wellenvektor $k_y^{\max} \leq (m\omega_c L/\hbar)$. Auf jedem Kreis der Abb. 8.24b können daher höchstens $k_y^{\max}/(2\pi/L)$ Zustände Platz finden. Unabhängig vom Wellenvektor k_z und von der Bahnquantenzahl ℓ weist also jedes Landau-Niveau den folgenden Entartungsgrad auf:

$$G = \frac{L}{2\pi}\frac{m\omega_c}{\hbar}L = \frac{e}{h}L^2\mathcal{B}\,. \tag{8.62}$$

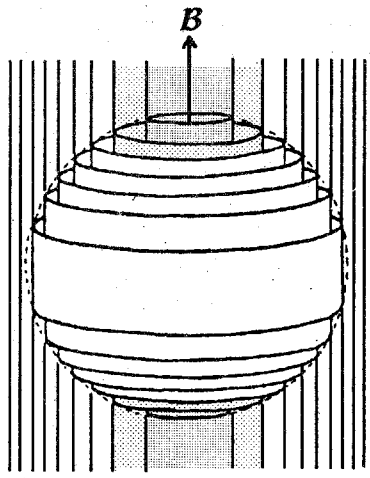


Abb. 8.25: Landau-Röhren des freien Elektronengases. Der innerste Zylinder ist grau, der zweite hellgrau getönt. Wie weit die Zylinder besetzt sind, hängt von der Größe der Fermi-Kugel ab, deren Schnitt mit den Zylindern ebenfalls gezeichnet ist.

Voraussetzung für die Ausbildung der Quantisierung, d.h. für eine zeitliche und räumliche Festlegung der Phase der Wellenfunktion, sind mehrere ungestörte Umläufe. Die an sich unendlich scharfen Zylinder werden durch Stöße der Elektronen mit dem Gitter verbreitert. Wie bei der Zyklotronresonanz muß die Bedingung $\omega_c\tau \gg 1$ erfüllt werden, wobei τ wieder die mittlere Stoßzeit ist. Wird ein Elektron angeregt, z.B. bei der Zyklotronresonanz, so erfolgt ein Übergang von einem Landau-Rohr zum nächsten bzw. zu einem noch höheren. Dies ist natürlich nur an der Fermi-Fläche möglich, da nur dort freie Zustände vorhanden sind, die energetisch zugänglich sind.

De Haas-van Alphen-Effekt

In diesem Abschnitt wollen wir den Einfluß untersuchen, den die Quantisierung der Elektronenbewegung und die Entartung der Zustände im Magnetfeld auf die Zustandsdichte und Magnetisierung von Metallen besitzen. Betrachten wir nochmals Gleichung (8.56), die die Energieeigenwerte freier Elektronen im Magnetfeld beschreibt, so können wir uns bereits ein Bild von der Zustandsdichte von zwei- und dreidimensionalen Systemen im Magnetfeld machen. Das Magnetfeld führt in der xy-Ebene zu einer Quantisierung, bei der nur noch die Energiewerte $(\ell + 1/2)\,\hbar\omega_c$ auftreten. In zweidimensionalen Systemen, deren Zustandsdichte nach (7.10) konstant ist, resultiert daraus vollständige Quantisierung. In diesem Fall besteht die Zustandsdichte wie bei harmonischen Oszillatoren aus einer Aneinanderreihung von Deltafunktionen, wie sie in Abb. 8.26a zu sehen ist.

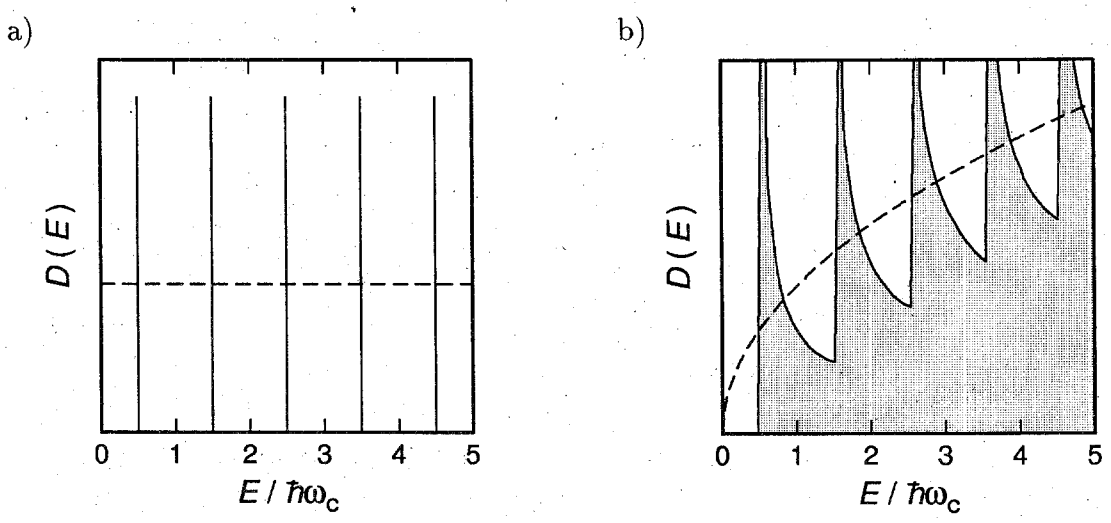


Abb. 8.26: Elektronische Zustandsdichte mit und ohne (gestrichelt) Magnetfeld. Die Spinaufspaltung wurde bei dieser Darstellung nicht berücksichtigt.
a) Zustandsdichte eines zweidimensionalen Systems,
b) Zustandsdichte eines dreidimensionalen Systems.

Im dreidimensionalen Raum wird k_z durch ein Magnetfeld in z-Richtung nicht quantisiert und kann daher jeden Wert annehmen, der durch die geometrischen Randbedingungen erlaubt ist. Betrachtet man die Darstellung der Elektronenenergie der Subbänder als Funktion von k_z in Abb. 8.23, so wird der in Abb. 8.26b gezeigte Verlauf der Zustandsdichte offensichtlich. Unterhalb der Energie $E = \hbar\omega_c/2$ existieren keine Zustände. Bei dieser Energie verläuft die Dispersionskurve waagerecht, so daß eine van Hove-Singularität auftritt. Diese Singularitäten wiederholen sich, immer wenn die Dispersionskurve waagerecht verläuft. Anschließend verringert sich die Zustandsdichte infolge der zunehmenden Steigung der $E(\boldsymbol{k})$-Kurve wieder. Damit wird der Kurvenverlauf in Abb. 8.26b verständlich. Da sich die Anzahl der Zustände durch das Magnetfeld nicht verändert, sind die Flächen unter den Zustandsdichtefunktionen ohne (gestrichelte Linie) und mit Feld (durchgezogene Linie) gleich.

Um die innere Energie der Elektronen und damit die Magnetisierung berechnen zu können, müssen wir die Magnetfeldabhängigkeit der Energieaufspaltung $\hbar\omega_c$ und des Entartungsgrades G berücksichtigen. In Abb. 8.27 sind die Zustandsdichten ohne und mit Magnetfeld für den zweidimensionalen Fall wiedergegeben.

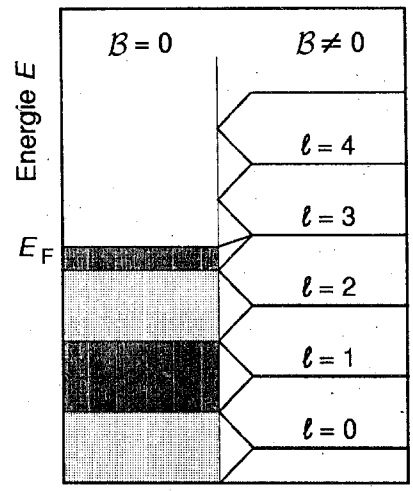


Abb. 8.27: Besetzung der Zustände in einem zweidimensionalen Elektronengas ohne und mit Magnetfeld. In dieser Darstellung sind die Zustände bis $\ell = 2$ vollständig, der Zustand $\ell = 3$ nur teilweise besetzt, die darüberliegenden Zustände sind leer. Die unterschiedliche Graufärbung gibt an, welche Zustände beim Anlegen des Magnetfelds $\mathcal{B}$ jeweils gemeinsam auf einem Niveau zu liegen kommen.

Die Fermi-Energie springt bei Änderungen des Feldes von Niveau zu Niveau. Immer, wenn die Fermi-Energie ohne Feld die Bedingung $E_F(\mathcal{B} = 0) = (\ell + 1/2)\,\hbar\,\omega_c$ erfüllt, gibt es (bei $T = 0$) keine teilbesetzten Niveaus. Dabei werden gleich viele Zustände energetisch angehoben wie abgesenkt, so daß die innere Energie der Elektronen mit und ohne Feld gleich ist. Da unter realen Bedingungen sehr viele Landau-Niveaus besetzt sind, können wir $(\ell + 1/2)$ näherungsweise durch die Anzahl p der besetzten Niveaus ersetzen, d.h. $E_F \approx p\hbar\omega_c$ schreiben. Setzen wir (8.49) für die Zyklotronfrequenz ein, so ergibt sich

$$p = \frac{mE_F}{\hbar e} \cdot \frac{1}{\mathcal{B}} \,. \tag{8.63}$$

Eine Magnetfeldänderung $\delta\mathcal{B}$ bewirkt daher eine Änderung in der Zahl der besetzen Niveaus, die durch

$$\delta p = -p\,\frac{\delta\mathcal{B}}{\mathcal{B}} \tag{8.64}$$

gegeben ist. Die Periode der vollständigen Besetzung der Niveaus ist daher proportional zu $1/\mathcal{B}$. Damit unmittelbar verbunden ist eine Oszillation der inneren Energie, deren Variation in Abb. 8.28 schematisch dargestellt ist. Sieht man von den speziellen Feldstärken ab, bei denen nur vollbesetzte Niveaus vorhanden sind, so ist die innere Energie mit Feld immer größer als ohne Feld.

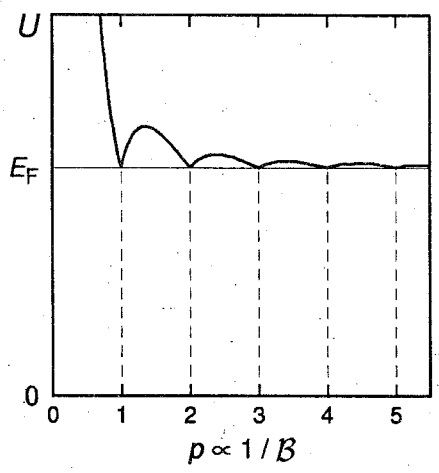


Abb. 8.28: Schematischer Verlauf der inneren Energie als Funktion von $1/\mathcal{B}$. p gibt die Zahl der vollbesetzten Landau-Niveaus an. Es ist zu beachten, daß in realen Metallen $p > 10^4$ ist.

Die Modulation der inneren Energie führt auch zu einer Modulation der spezifischen Wärme im Magnetfeld, da C_V durch die Ableitung der inneren Energie gegeben ist. In Abb. 8.29 ist das Ergebnis einer Messung der spezifischen Wärme im Magnetfeld an Beryllium bei 1,5 K gezeigt. Bei Magnetfeldern um 2 T bewirkt bereits eine relativ kleine Änderung des Feldes eine Reihe von Maxima und Minima.

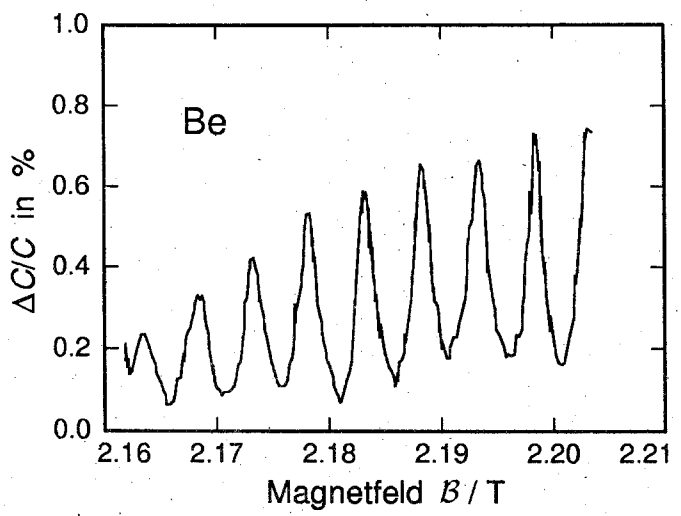


Abb. 8.29: Variation der spezifischen Wärmen von Beryllium bei 1,5 K als Funktion des anliegenden Magnetfelds. (Nach P.F. Sullivan und G. Seidel, Phys. Rev. **173** (1968) 679).

Teil- und vollbesetzte Landau-Niveaus haben völlig verschiedene Eigenschaften, die sich im Tieftemperaturverhalten des elektrischen Widerstands, der magnetischen Suszeptibilität, der thermischen Leitfähigkeit oder der spezifischen Wärme im Magnetfeld bemerkbar machen. Kurz herausgegriffen seien hier die Oszillationen der elektrischen Leitfähigkeit. Elektronen in vollbesetzten Landau-Niveaus (wie auch in „normalen" vollbesetzten Bändern) tragen nicht zur Leitfähigkeit bei. Der elektrische Stromfluß wird daher nur von denjenigen Elektronen bewirkt, die sich in teilbesetzten Niveaus aufhalten. Da sich deren Anzahl periodisch mit der Magnetfeldstärke ändert, treten in der elektrischen Leitfähigkeit senkrecht zum angelegten Magnetfeld charakteristische Oszillationen, die sogenannten *Shubnikov-de Haas-Oszillationen*, auf. Bei hohen Temperaturen verschwindet dieser Effekt, da die Landau-Niveaus durch Elektron-Phonon-Stöße ausgeschmiert werden.

Eingehender betrachten wir die Magnetisierung M, die am absoluten Nullpunkt durch $M = -(\partial U/\partial \mathcal{B})$ gegeben ist. Da die innere Energie, wie in Abb. 8.28 skizziert, mit einer Periode proportional zu $1/\mathcal{B}$ schwankt, erwarten wir ein ähnliches Verhalten auch für die Magnetisierung. Bemerkenswert ist, daß die Ableitung der inneren Energie positive und negative Werte annehmen kann. Ist $M > 0$, so ist der Beitrag der Bahnbewegung der Elektronen paramagnetisch, für $M < 0$ dagegen diamagnetisch. Diese *De Haas-van Alphen-Oszillation* der Magnetisierung ist in Abb. 8.30 schematisch dargestellt. Beim Durchgang durch ein vollbesetztes Landau-Niveau macht die Magnetisierung jeweils einen Sprung.

In verunreinigten Substanzen und mit zunehmender Temperatur werden die Landau-Niveaus und damit auch die mit ihnen verknüpften Effekte verschmiert. Daher sind bei höheren Temperaturen die Spitzen der De Haas-van Alphen-Oszillationen abgerundet und bei Zimmertemperatur ganz verschwunden. Jedoch mittelt sich der Effekt nicht vollständig weg. Es verbleibt der sogenannte *Landau-Diamagnetismus*, der auf der hier diskutierten Bahnbewegung der freien Elektronen beruht. Eine weitergehende Rechnung

zeigt, daß gerade ein Drittel des Pauli-Spinmagnetismus kompensiert wird. Wir werden diesen Punkt nochmals in Kap. 10.1 streifen.

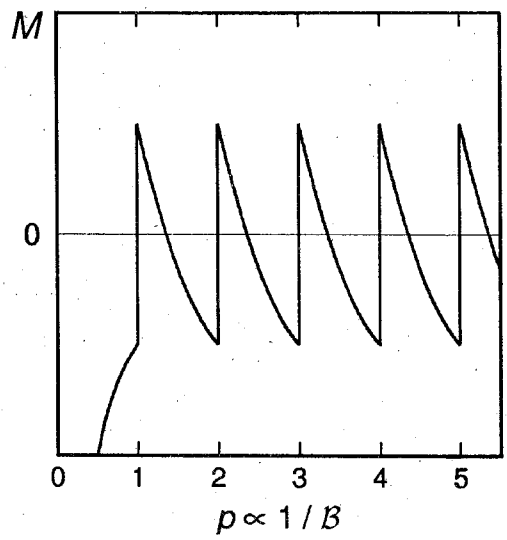


Abb. 8.30: De Haas-van Alphen-Oszillation. Die Magnetisierung der Leitungselektronen oszilliert bei Variation des Magnetfelds. Der Sprung tritt den Feldern auf, bei denen die Landau-Niveaus vollbesetzt sind.

Wir wollen nun zeigen, daß sich aus der Periode der De Haas-van Alphen-Oszillationen direkt die Querschnittsfläche der Fermi-Fläche ablesen läßt. Für unsere Betrachtungen sind nur die Elektronen mit der Fermi-Energie von Interesse. Wir drücken mit Hilfe von (8.59) und (8.60) die Fermi-Energie E_F aus und finden

$$E_\mathrm{F} = \frac{\hbar^2 k_\mathrm{F}^2}{2m} = \frac{\hbar^2 \mathcal{S}_\mathrm{F}}{2\pi m}, \tag{8.65}$$

wobei $\mathcal{S}_\mathrm{F}$ für die von der Elektronenbahn im $\boldsymbol{k}$-Raum eingeschlossene Fläche steht. Auch hier gelten die Gleichungen (8.63) und (8.64), die wir in (8.65) einsetzen. So finden wir für die Periode der De Haas-van Alphen-Oszillationen

$$\delta\mathcal{B} = \frac{2\pi e}{\hbar \mathcal{S}_\mathrm{F}} \mathcal{B}^2 \qquad \text{oder} \qquad \delta\left(\frac{1}{\mathcal{B}}\right) = \frac{2\pi e}{\hbar \mathcal{S}_\mathrm{F}}. \tag{8.66}$$

Die im Experiment gemessene Periode der Oszillation ist also proportional zu $1/\mathcal{S}_\mathrm{F}$, der inversen Querschnittsfläche im $\boldsymbol{k}$-Raum. Aus den gleichen Gründen wie bei der Zyklotronresonanz tragen auch hier praktisch nur Elektronen auf Extremalbahnen zum Meßsignal bei.

In Abb. 8.31a ist ein Meßergebnis an Gold wiedergegeben, bei dem das Feld in [111]-Richtung angelegt war. Deutlich zu erkennen ist, daß die Oszillation der Magnetisierung mit zwei unterschiedlichen Perioden erfolgt, die auf eine Hals- bzw. eine Bauchbahn zurückgeführt werden können. Aus den beiden Perioden der De Haas-van Alphen-Oszillationen ergibt sich, daß sich die Querschnittsflächen beim Gold wie 1:29 verhalten. Man findet die Zahlenwerte $\mathcal{S}_\mathrm{F} = 1{,}5 \cdot 10^{15}\,\mathrm{cm}^{-2}$ für die Hals- und $\mathcal{S}_\mathrm{F} = 4{,}3 \cdot 10^{16}\,\mathrm{cm}^{-2}$ für die Bauchbahn. Wie in Abb. 8.31b angedeutet, beobachtet man im Falle eines Feldes in [100]-Richtung nur eine Oszillation, die von der entsprechenden Bauchbahn herrührt. Führt man derartige Messungen mit unterschiedlicher Orientierung der Probe durch, so läßt sich die Form der Fermi-Fläche sehr gut rekonstruieren.

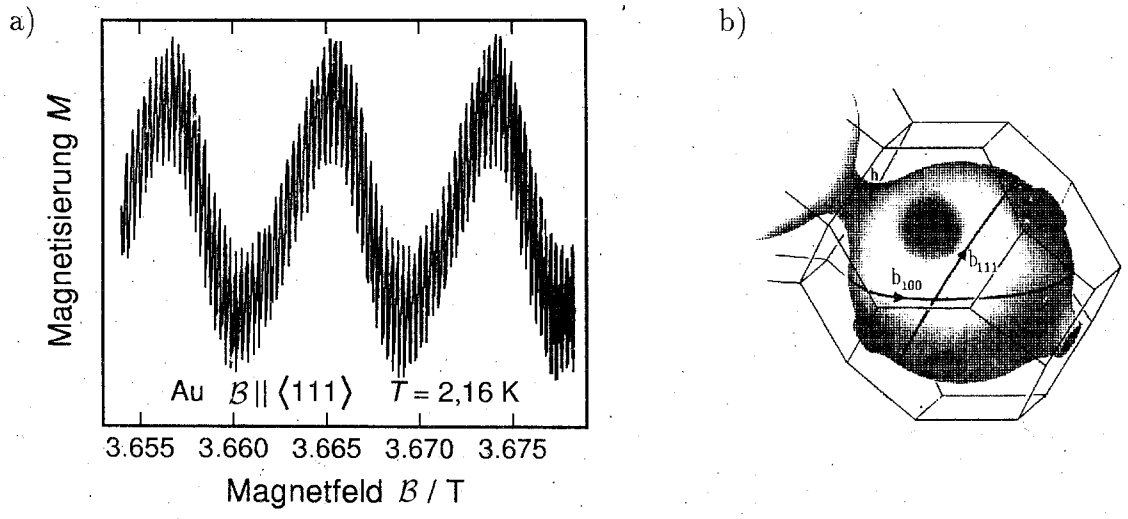


Abb. 8.31: De Haas-Van Alphen-Effekt an Gold.
a) De Haas-Van Alphen-Effekt bei Magnetfeld in [111]-Richtung. Die niederfrequente Oszillation wird der Halsbahn, die hochfrequente der Bauchbahn zugeschrieben (nach I.M. Templeton).
b) Fermi-Fläche von Gold. Liegt ein Magnetfeld in [111]-Richtung an, so treten die Halsbahn h und die Bauchbahn b_{111} auf.

8.5 Hall-Effekt

In diesem Abschnitt beschäftigen wir uns zunächst mit dem klassischen *Hall-Effekt*, der eine Bestimmung der Ladungsträgerkonzentration erlaubt. Die abgeleiteten Gleichungen werden wir dann bei der Diskussion des *Quanten-Hall-Effektes* benutzen, der in zweidimensionalen Elektronensystemen beobachtet wird.

Klassischer Hall-Effekt

Zur Beschreibung der Elektronenbewegung müssen wir die in Kap. 8.2 benutzten Gleichungen etwas erweitern. Wir gehen von der Bewegungsgleichung der Elektronen in elektrischen und magnetischen Feldern aus und berücksichtigen wie bei der Diskussion der elektrischen Leitfähigkeit, daß die Elektronen Stöße erleiden, die zum Ohmschen Widerstand führen:

$$\dot{\boldsymbol{v}}_\mathrm{d} = \frac{-e}{m^*}(\boldsymbol{\mathcal{E}} + \boldsymbol{v}_\mathrm{d} \times \boldsymbol{\mathcal{B}}) - \frac{\boldsymbol{v}_\mathrm{d}}{\tau} \,. \tag{8.67}$$

τ ist die mittlere Stoßzeit der Elektronen, wie sie in Kap. 8.2 eingeführt wurde. Wir haben hier wieder die Driftgeschwindigkeit $\boldsymbol{v}_\mathrm{d}$ benutzt, die ja entsprechend (8.17) und (8.21) mit der stationären Verschiebung $\delta\boldsymbol{k} = -e\boldsymbol{\mathcal{E}}\tau/\hbar = m^*\boldsymbol{v}_\mathrm{d}/\hbar$ der Fermi-Kugel verknüpft ist. Überraschenderweise ist die Verwendung der Driftgeschwindigkeit auch für die Lorentz-Kraft korrekt. Da die Elektronen nach jedem Stoß ihre Richtung ändern, bewegen sie sich auf komplizierten Bahnen, auf denen stets die Lorentzkraft $\boldsymbol{v}(\boldsymbol{k}) \times \boldsymbol{\mathcal{B}}$ wirkt. Aufgrund der irregulären Bewegung mittelt sich jedoch der Beitrag zum Hall-Effekt bis auf den Term $\boldsymbol{v}_\mathrm{d} \times \boldsymbol{\mathcal{B}}$ weg. Die Verwendung der effektiven Masse m^* soll betonen, daß wir es nicht mit freien Elektronen, sondern mit Bloch-Elektronen zu tun haben, die wie in Kap. 8.1

als Wellenpakete beschrieben werden können. Im Gleichgewicht ist die Driftgeschwindigkeit $\boldsymbol{v}_\mathrm{d}$ konstant, d.h. $\dot{\boldsymbol{v}}_\mathrm{d} = 0$. Wir erhalten dann aus (8.67)

$$\begin{aligned} 0 &= -e\mathcal{E}_x - \omega_\mathrm{c} m^* v_{\mathrm{d},y} - \frac{m^* v_{\mathrm{d},x}}{\tau} \\ 0 &= -e\mathcal{E}_y + \omega_\mathrm{c} m^* v_{\mathrm{d},x} - \frac{m^* v_{\mathrm{d},y}}{\tau} \,, \end{aligned} \tag{8.68}$$

da bei der hier diskutierten und in Abb. 8.32 dargestellten Probengeometrie keine Komponenten des elektrischen Feldes in z-Richtung auftreten.

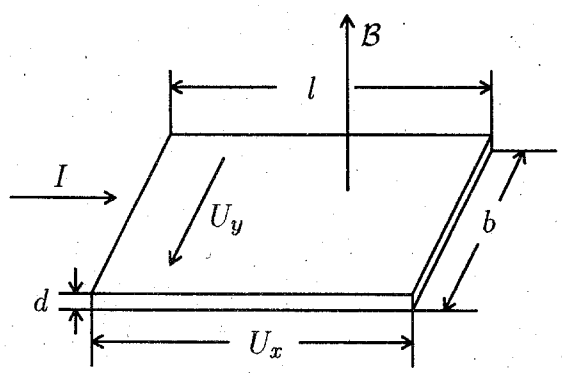


Abb. 8.32: Probengeometrie bei der Messung des Hall-Effektes. Die Probe besitzt die Länge l, die Breite b und die Dicke d. Gemessen werden die Längsspannung U_x und die Hallspannung U_y bei eingeprägtem Strom I.

Mit den beiden Abkürzungen

$$\varrho_{xx} = \frac{\mathcal{B}}{ne}\frac{1}{\omega_\mathrm{c}\tau} = \frac{m^*}{ne^2\tau} \qquad \text{und} \qquad \varrho_{xy} = \frac{\mathcal{B}}{ne} \tag{8.69}$$

für die auftretenden Widerstände läßt sich (8.68) in die Form

$$\begin{pmatrix} \mathcal{E}_x \\ \mathcal{E}_y \end{pmatrix} = \begin{pmatrix} \varrho_{xx} & \varrho_{xy} \\ -\varrho_{xy} & \varrho_{xx} \end{pmatrix} \begin{pmatrix} j_x \\ j_y \end{pmatrix} \tag{8.70}$$

bringen. Die Komponente ϱ_{xx} entspricht dem üblichen Ausdruck für den spezifischen Widerstand, während ϱ_{xy} mit dem *Hall-Widerstand* verknüpft ist. Es sei bemerkt, daß im allgemeinen Fall des elektrischen Stromtransports in anisotropen Materialien in (8.70) die Größen ϱ_{yy} und ϱ_{yx} auftreten sollten. Für isotrope Materialien und bei der gewählten Richtung des Magnetfelds ist jedoch $\varrho_{xx} = \varrho_{yy}$ und $\varrho_{yx} = -\varrho_{xy}$.

Lösen wir die Gleichung (8.70) nach dem Strom auf, so erhalten wir

$$\begin{pmatrix} j_x \\ j_y \end{pmatrix} = \begin{pmatrix} \sigma_{xx} & \sigma_{xy} \\ -\sigma_{xy} & \sigma_{xx} \end{pmatrix} \begin{pmatrix} \mathcal{E}_x \\ \mathcal{E}_y \end{pmatrix} \tag{8.71}$$

mit den Leitwerten

$$\sigma_{xx} = \frac{ne}{\mathcal{B}}\frac{\omega_\mathrm{c}\tau}{1+\omega_\mathrm{c}^2\tau^2} \qquad \text{und} \qquad \sigma_{xy} = -\frac{ne}{\mathcal{B}}\frac{\omega_\mathrm{c}^2\tau^2}{1+\omega_\mathrm{c}^2\tau^2} \,. \tag{8.72}$$

Da bei der in Abb. 8.32 dargestellten Geometrie kein Stromfluß in y-Richtung auftreten kann, vereinfacht sich Gleichung (8.70) zu:

$$\mathcal{E}_x = \varrho_{xx}\, j_x \qquad \text{und} \qquad \mathcal{E}_y = -\varrho_{xy}\, j_x \,. \tag{8.73}$$

Hieraus folgt für den Längs- und Querwiderstand:

$$R_x = \varrho_{xx}\frac{l}{bd} \qquad \text{und} \qquad R_y = \left|\frac{\varrho_{xy}}{d}\right| . \tag{8.74}$$

Für die *Hall-Konstante* R_H ergibt sich damit der Ausdruck

$$R_\mathrm{H} = \frac{\mathcal{E}_y}{j_x \mathcal{B}} = -\frac{1}{ne} . \tag{8.75}$$

Da $\mathcal{E}_y$, j_x und $\mathcal{B}$ Meßgrößen sind, läßt sich R_H und somit die Elektronendichte direkt bestimmen. Für Natrium findet man die erwartete Dichte von 1 e^-/Atom, doch bei Beryllium ergeben sich 0,2 e^+/Atom. Dies bedeutet, daß in Beryllium in erster Linie nicht Elektronen, sondern Löcher für den Stromtransport verantwortlich sind. Auf die besondere Bandstruktur der Erdalkalimetalle haben wir bereits in Kap. 7.4 hingewiesen.

Wird der Strom durch unterschiedliche Ladungsträger, z.B. durch Elektronen in unterschiedlichen Bändern oder durch Elektronen und Löchern transportiert, dann muß die hier durchgeführte Ableitung etwas modifiziert werden. Die Existenz der beiden Ladungsträgersorten muß bereits in (8.68) berücksichtigt werden. Hier soll nur das Ergebnis für den Ladungstransport durch Elektronen und Löcher angegeben werden, ohne auf die Herleitung einzugehen. Vernachlässigen wir bei der Berechnung der Stromdichte j_x den Beitrag des Hall-Felds $\mathcal{E}_y$, da dieser proportional zu $\mathcal{B}^2$ ist, so findet man für die Hall-Konstante die Gleichung

$$R_\mathrm{H} = \frac{p\mu_\mathrm{p}^2 - n\mu_\mathrm{n}^2}{e(p\mu_\mathrm{p} + n\mu_\mathrm{n})^2} . \tag{8.76}$$

Hierbei stehen n und p für die Dichte, μ_n und μ_p für die Beweglichkeit von Elektronen bzw. Löchern. Diese Beziehung wird in Kap. 9.2 bei der Diskussion der Halbleiter noch Anwendung finden.

Quanten-Hall-Effekt

Nun wollen wir den Hall-Effekt zweidimensionaler Systeme betrachten, bei denen eine Besonderheit auftritt. Elektronen lassen sich auf relativ einfache Weise in dünnen Schichten lokalisieren, so daß sich ein zweidimensionales Elektronengas bildet. Eine Realisierungsmöglichkeit bieten beispielsweise Feldeffekttransistoren oder Halbleiterheterostrukturen, die in Kap. 9.4 und 9.5 kurz angesprochen werden. Wichtig für unsere weiteren Betrachungen ist nur, daß sich die Elektronen nach wie vor frei in x- und y-Richtung bewegen können, in z-Richtung aber festgehalten werden. Wird nun ein Magnetfeld mit der Stärke $\mathcal{B}$ in z-Richtung angelegt, so wird, wie wir in Kap. 8.4 ausführlich diskutiert haben, die Bahnbewegung der Elektronen in der xy-Ebene zusätzlich quantisiert. Es treten Landau-Niveaus auf, in denen Elektronen bei voller Besetzung nicht gestreut werden können. Dies bedeutet, daß die Stoßzeit τ unendlich groß wird, und ϱ_{xx} nach (8.69), aber auch σ_{xx} nach (8.72) den Wert Null annimmt! Der Strom wird dann nicht durch das elektrische Längsfeld $\mathcal{E}_x$, das wegen $\sigma_{xx} = 0$ ebenfalls den Wert Null hat, durch die Probe getrieben, sondern über ϱ_{xy} durch die Hall-Spannung und damit durch das Magnetfeld.

In jedem Landau-Niveau haben nach (8.64) und entsprechend der Probengeometrie $G = (e/h)\, b\, l\, \mathcal{B}$ Elektronen Platz. Ist p die Zahl der besetzten Niveaus und $\tilde{n}$ die Zahl der Elektronen pro Fläche, so sind immer dann nur vollbesetzte Niveaus vorhanden, wenn

$$\tilde{n} = nd = p\frac{G}{bl} = p\frac{e}{h}\mathcal{B} \tag{8.77}$$

ist. Variiert man das Magnetfeld, so sollte, wenn (8.77) erfüllt ist, $\varrho_{xx} = 0$ sein. Der Hall-Widerstand $R_y = |\varrho_{xy}/d|$ sollte linear mit der Magnetfeldstärke ansteigen und bei den durch (8.77) vorgegebenen Magnetfeldern exakt den Wert

$$R_y = \frac{\varrho_{xy}}{d} = \frac{h}{p\,e^2} = \frac{25812,8\,\Omega}{p} \tag{8.78}$$

annehmen, wobei $p = 1, 2, 3, \ldots$ eine positive ganze Zahl ist.

In Abb. 8.33 ist das Ergebnis einer derartigen Messung an einer GaAS-(AlGa)As-Heterostruktur zu sehen.

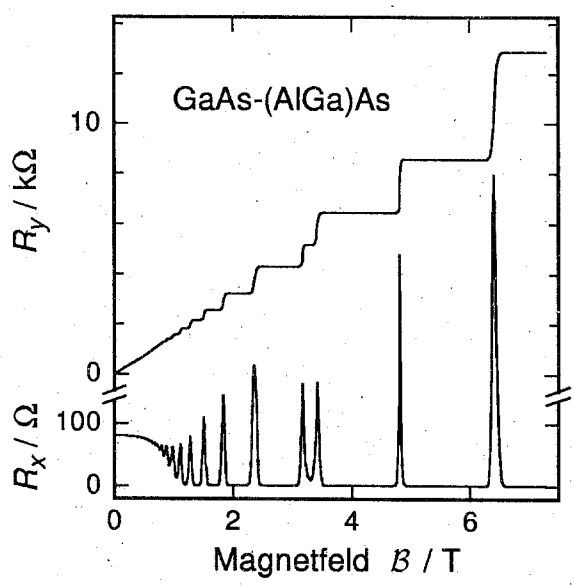


Abb. 8.33: Hall-Widerstand R_y und Längswiderstand R_x, gemessen an einer GaAs-(AlGa)As-Heterostruktur bei sehr tiefen Temperaturen. (Aus K. von Klitzing, Physica **126 B** (1984) 242).

Das Experiment bestätigt diese Voraussagen. Bei den vorausgesagten Werten des Magnetfelds verschwindet ϱ_{xx}, und R_y nimmt Werte nach (8.78) an. Das eigentlich Erstaunliche bei diesen Experimenten sind aber die Plateaus, die der Hall-Widerstand R_y aufweist. Gleichzeitig verschwindet in diesem Bereich, und nicht nur an den Stellen, bei denen alle Elektronen in vollbesetzten Niveaus untergebracht sind, der Längswiderstand R_x. Dieses erstaunliche Verhalten wurde von *von Klitzing* entdeckt und wird als *Quanten-Hall-Effekt* bezeichnet. Eine vollständige Erklärung des Effektes steht noch aus. Grundsätzlich sollte die Variation der Dichte der freien Elektronen eine Modulation des Widerstands R_x, die bereits erwähnten Shubnikov-de Haas Oszillationen, hervorrufen. Es ist jedoch nicht zu erwarten, daß der Widerstand über einen weiten Bereich verschwindet. Man nimmt an, daß ein Teil der Elektronen nicht wirklich frei, sondern durch Streuung an zufällig verteilten Störstellen (vgl. Kap. 9.2) oder durch den Einfluß der Ränder des Strompfades auf noch nicht völlig verstandene Art und Weise lokalisiert ist. Diese Elektronen besitzen nicht eine den Landau-Niveaus der freien Elektronen entsprechende Energie, so daß zwischen den Niveaus ihre Zustandsdichte nicht verschwindet. Eine bildhafte Darstellung hierfür ist in Skizze 8.34 zu sehen. Die lokalisierten Elektronen liefern keinen Beitrag zum Stromtransport, haben jedoch Einfluß auf die Lage des Fermi-Niveaus. Wird das Magnetfeld langsam erhöht, so kann die Fermi-Energie nicht, wie bisher diskutiert, von einem Landau-Niveau zum nächsten springen, sondern wird von den lokalisierten Zuständen zwischen den Niveaus festgehalten. Während die Fermi-Energie bei einer Variation des Magnetfelds durch

diesen Energiebereich wandert, ist die Stoßzeit τ der freien, beweglichen Elektronen in den Landau-Niveaus tatsächlich unendlich und somit gilt für die Widerstände $R_x = 0$ und $R_y = \text{const}$. Die lokalisierten Elektronen beeinflussen den Stromtransport nicht. Damit läßt sich das Auftreten der überraschenden Plateaus in einem phänomenologischen Modell zumindest qualitativ erklären. Darüberhinaus kann gezeigt werden, daß die Plateauwerte von R_y dem Wert des Hall-Widerstands völlig freier Elektronen entsprechen, wie man es bei vollständiger Besetzung der Landau-Niveaus auch erwartet. Unabhängig von Art und Herkunft der verwendeten Proben wurden mit einer Genauigkeit von 10^{-7} in den Plateaus Widerstandswerte gefunden, die durch (8.78) vorhergesagt wurden.

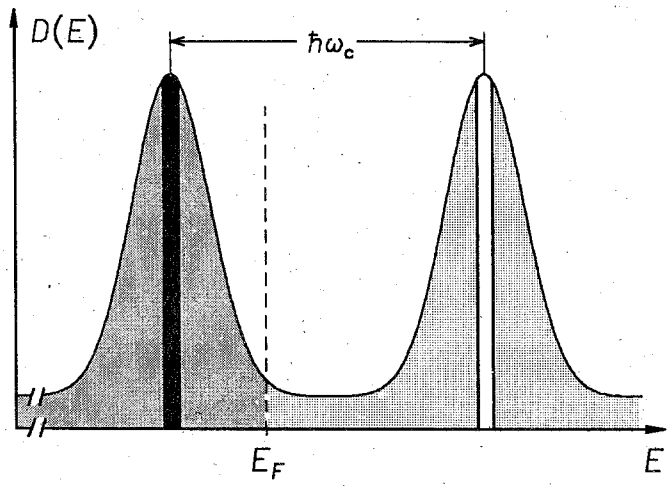


Abb. 8.34: Modell der elektronischen Zustandsdichte beim Quanten-Hall-Effekt. Die besetzten lokalisierten Zustände sind dunkel, die unbesetzten hellgrau markiert. Das letzte volle Landau-Niveau ist schwarz, das erste freie weiß gezeichnet. Es ist angenommen, daß die Fermi-Energie im Bereich der lokalisierten Zustände liegt.

Nach (8.78) ist der Wert des Hall-Widerstands R_y eng mit der Feinstrukturkonstanten α verknüpft, die durch

$$\alpha = \frac{e^2}{h}\frac{1}{2\varepsilon_0 c} \approx \frac{1}{137} \tag{8.79}$$

gegeben ist. Der Quanten-Hall-Effekt erlaubt daher eine Präzisionsbestimmung dieser fundamentalen Konstanten. Setzt man andererseits diese Größe als bekannt voraus, so läßt sich mit Hilfe dieses Effektes ein Eichnormal für die Widerstandseinheit „Ohm" aufbauen.

Bei sehr tiefen Temperaturen und noch höheren Magnetfeldern wird der *fraktionale Quanten-Hall-Effekt* beobachtet, bei dem gebrochene Quantenzahlen gefunden werden. Trotz der auffälligen Ähnlichkeit mit dem ganzzahligen Quanten-Hall-Effekt ist der physikalische Ursprung ein anderer. Dies macht sofort die Tatsache deutlich, daß durchaus Werte $p < 1$ auftreten. In diesem Fall ist nur noch das unterste Landau-Niveau bevölkert, das aber nicht voll besetzt sein kann. Zur Beschreibung dieses Phänomens reicht die Einteilchen-Näherung nicht mehr aus, die wir bisher benutzt haben. Man muß vom quasifreien Elektronengas zu stark wechselwirkenden Elektronen („Quantenflüssigkeit") übergehen, die sich korreliert bewegen. Es treten dann effektive Ladungen $e^* = p\,e$ auf, die kleiner als die Elementarladung e sein können. Ein Behandlung dieses Effekts geht jedoch über den hier gesteckten Rahmen der Vorlesung hinaus[11].

[11] Einen neuen Artikel von G. Abstreiter über die Entdeckung und die Beschreibung des fraktionalen Quanten-Hall-Effekts findet man in den Physikalischen Blättern, **54** (1998) 1098.

9. Halbleiter

Die Bezeichnung *Halbleiter* deutet darauf hin, daß bei dieser Substanzklasse die elektrische Leitfähigkeit zwischen derjenigen von Metallen und Isolatoren liegt. Gute Metalle haben einen spezifischen Widerstand um $10^{-7}\,\Omega$m, gute Isolatoren dagegen über $10^{10}\,\Omega$m. Bei Halbleitern variiert der spezifische Widerstand über weite Bereiche, etwa von $10^{-4}\,\Omega$m bis $10^{7}\,\Omega$m. Halbleitende Elemente mit einfacher Kristallstruktur findet man in der vierten Hauptgruppe des Periodensystems, wobei Silizium und Germanium die beiden wichtigsten sind. Diamant gehört eher in die Klasse der Isolatoren und Blei zu den Metallen. Zinn besitzt sowohl eine halbleitende als auch eine metallische Modifikation. Weitere halbleitende Elemente sind roter Phosphor, Bor, Selen und Tellur, die alle kovalent gebunden sind und eine komplexe Kristallstruktur besitzen. Zusätzlich gibt es eine große Anzahl von Verbindungshalbleitern. Einige davon wie z.B. die zu den III-V-Verbindungen gehörenden Halbleiter GaAs und InSb haben Zinkblendestruktur und eine teilweise ionische Bindung. Natürlich gibt es auch II-VI-Verbindungen mit halbleitendem Charakter wie ZnS oder CdS, aber auch IV-IV-Halbleiter wie SiC.

Am absoluten Nullpunkt verhalten sich fehlerfreie Halbleiter wie Isolatoren. Da ihre Bänder entweder vollständig mit Elektronen besetzt oder völlig leer sind, kann kein Ladungstransport auftreten. Bei endlichen Temperaturen werden Valenzelektronen angeregt, in ein leeres Band gehoben und machen so das Material leitfähig. Für technische Anwendungen ist die gezielte Erzeugung von Defekten, das sogenannte *Dotieren*, von entscheidender Bedeutung, denn dadurch können die elektrischen Eigenschaften des Ausgangsmaterials drastisch verändert und den technischen Bedürfnissen angepaßt werden.

Wir werden in diesem Kapitel zunächst reine, sogenannte *intrinsische* Halbleiter betrachten und dann auf die Auswirkungen des Dotierens eingehen, das zur *Störstellenleitung* führt. Anschließend werfen wir einen Blick auf amorphe Halbleiter, deren technische Bedeutung ständig zunimmt, und zum Schluß besprechen wir die Wirkungsweise von einigen elektronischen Bauelementen und sogenannte Heterostrukturen.

9.1 Intrinsische kristalline Halbleiter

Bandlücke und optische Absorption

Bei Halbleitern bezeichnet man das energetisch höchste voll besetzte Band als *Valenzband* und das erste leere darüberliegende als *Leitungsband*. Die Energie der Valenzbandoberkante werden wir mit E_{V}, die Leitungsbandunterkante mit E_{L} bezeichnen. Die *Bandlücke* $E_{\mathrm{g}} = E_{\mathrm{L}} - E_{\mathrm{V}}$ zwischen diesen Bändern ist für viele elektronische Eigenschaften von entscheidender Bedeutung. Ihre Größe hängt schwach von der Temperatur ab. Von tiefen Temperaturen kommend nimmt die Bandlücke zunächst quadratisch und dann linear mit der Temperatur ab. Die gesamte Änderung bis Zimmertemperatur beträgt etwa 10% und beruht auf der thermischen Expansion und Effekten der Elektron-Phonon-Wechselwirkung, die wir hier nicht diskutieren werden.

In der Tabelle 9.1 sind für eine Reihe von Halbleitern die Bandlücken bei Zimmertemperatur und am absoluten Nullpunkt aufgeführt. Zusätzlich ist angegeben, ob eine *direkte* oder *indirekte Bandlücke* vorliegt. Liegen Valenzbandmaximum und das tiefstliegende Leitungsbandminimum am Γ-Punkt, so spricht man von einem *direkten Halbleiter*, bzw. von einer direkten Bandlücke. Liegen die Extrema im $\boldsymbol{k}$-Raum nicht am selben Ort, so handelt es sich um einen *indirekten Halbleiter*.

Tabelle 9.1: Bandlücke E_g und Hinweis auf die Natur der Lücke einiger Halbleiter.

	$E_g(T = 300\,\mathrm{K})$ / (eV)	$E_g(T = 0\,\mathrm{K})$ / (eV)	Natur der Lücke
Diamant	5,47	5,48	indirekt
Si	1,12	1,17	indirekt
Ge	0,66	0,75	indirekt
GaP	2,26	2,32	indirekt
GaAs	1,43	1,52	direkt
InSb	0,18	0,24	direkt
InP	1,35	1,42	direkt
CdS	2,42	2,58	direkt
PbTe	0,30	0,19	direkt

In der Abb. 9.1 sind die Bandstrukturen der beiden technisch wichtigen und wissenschaftlich interessanten Halbleiter Galliumarsenid und Germanium dargestellt. Es handelt sich hierbei um die Ergebnisse von Bandstrukturrechnungen, die an spektroskopische Messungen angepaßt wurden. Eine ganz ähnliche Bandstruktur wie Germanium weist Silizium auf, der Halbleiter mit der größten technischen Bedeutung.

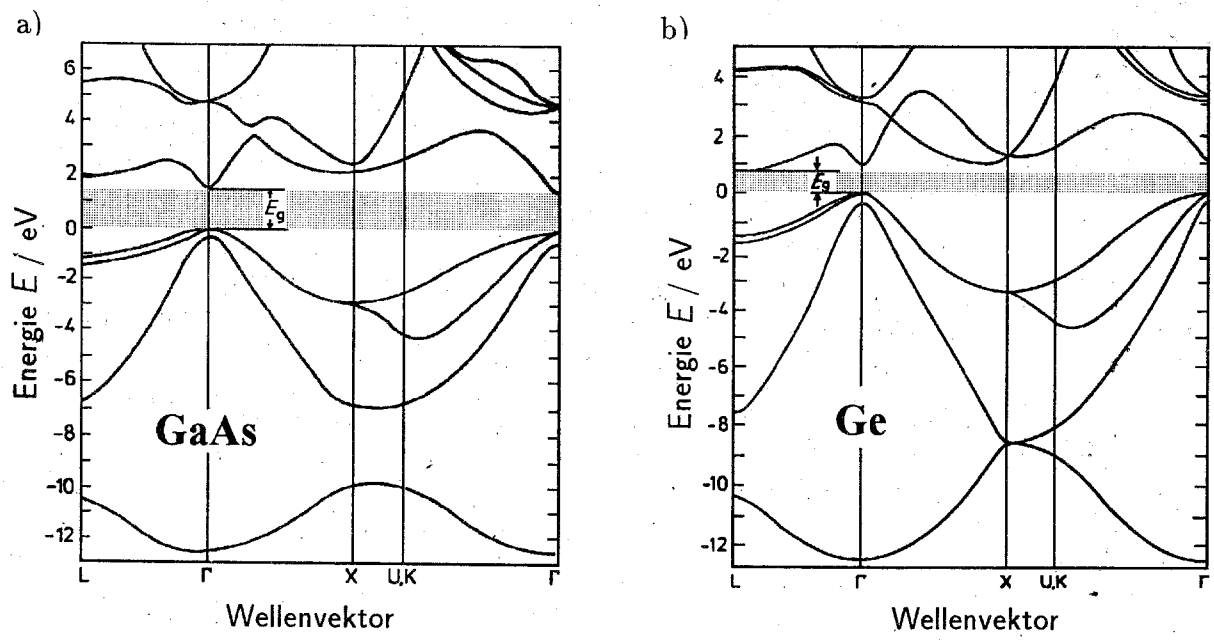


Abb. 9.1: Bandstruktur von Galliumarsenid und Germanium. Die Energielücke E_g ist grau markiert. (Nach J.R. Chelikowsky und M.L. Cohen, Phys. Rev. B **14** (1976) 556).
a) Das Valenzbandmaximum und das tiefste Leitungsbandminimum stehen sich am Γ-Punkt gegenüber: Galliumarsenid ist ein direkter Halbleiter.
b) Beim indirekten Halbleiter Germanium befindet sich das Valenzbandmaximum ebenfalls am Γ-Punkt, das tiefste Leitungsbandminimum aber am L-Punkt .

Wie in Abb. 9.1a zu sehen ist, ist Galliumarsenid ein direkter Halbleiter, denn Valenzbandmaximum und tiefstliegendes Leitungsbandminimum liegen im $\boldsymbol{k}$-Raum an derselben Stelle. Beim Germanium findet man das Valenzbandmaximum ebenfalls am Γ-Punkt, doch besitzt das darüberliegende Leitungsbandminimum nicht den kleinsten Energieabstand zum Valenzband. Wie in Abb. 9.1b zu erkennen ist, treten längs der $\langle 111 \rangle$-Richtungen an den L-Punkten tiefer liegende Leitungsbandminima auf.

Die Bandlücke läßt sich auf besonders einfache Weise mit Hilfe der optischen Absorption bestimmen. Wie in Abschnitt 8.1 bereits angesprochen, wird bei der Absorption eines Photons ein Elektron ins Leitungsband gehoben, während im Valenzband ein Loch zurückbleibt. Wir werden auf derartige *Interbandübergänge* noch einmal in Kap. 11 in einem allgemeineren Zusammenhang zu sprechen kommen. In Abb. 9.2a sind zwei derartige Absorptionsübergänge in Halbleitern durch Pfeile dargestellt. Der Übergang erfolgt im Bandschema vertikal, da der Betrag des Photonenimpulses $\hbar k_\mathrm{l}$ im Vergleich zu typischen Werten des Elektronenimpulses vernachlässigbar klein ist. Da die Bandlücke überwunden werden muß, kann ein Photon nur absorbiert werden, wenn dessen Energie den Minimalwert $\hbar\omega_\mathrm{l} = E_\mathrm{L} - E_\mathrm{V} = E_\mathrm{g}$ übertrifft. Halbleiter sind daher bei großen Wellenlängen durchsichtig. Oberhalb der Schwellenenergie, die durch die Bandlücke vorgegeben wird, steigt dann die optische Absorption mit zunehmender Frequenz steil an. Ein Beispiel hierfür ist in Abb. 9.2b zu sehen. Dort ist der Absorptionskoeffizient von Indiumantimonid logarithmisch als Funktion der Photonenenergie aufgetragen. Der steile Anstieg an der Absorptionkante ist charakteristisch für direkte Halbleiter.

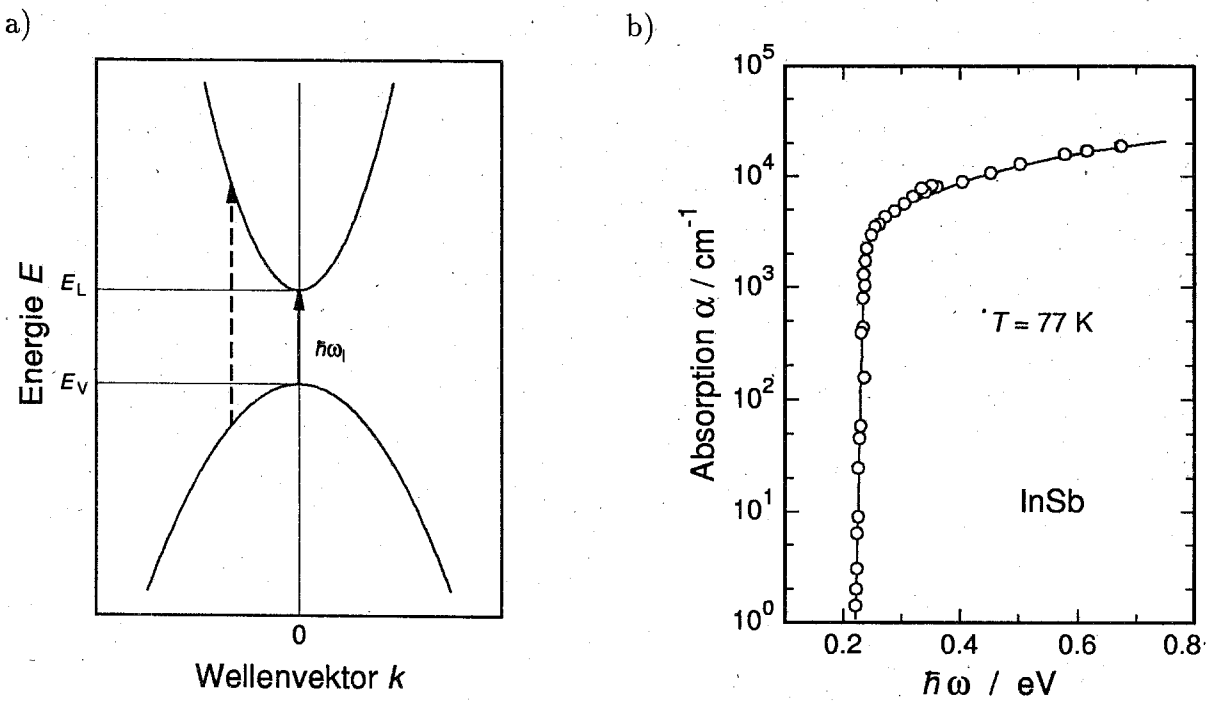


Abb. 9.2: Optische Absorption bei Halbleitern mit direkter Bandlücke.
a) Schematische Darstellung des Absorptionsvorgangs. Der ausgezogene Pfeil markiert den Übergang mit der kleinstmöglichen Energie. Bei höheren Photonenenergien (gestrichelter Pfeil) werden tiefer liegende Elektronen angeregt.
b) Optischer Absorptionskoeffizient von Indiumantimonid im logarithmischen Maßstab als Funktion der Energie der eingestrahlten Photonen. (Nach G.W. Gobeli und H.Y. Fan, Phys. Rev. **119** (1960) 613).

Bei indirekten Halbleitern ist der Absorptionsmechanismus etwas komplizierter. Wie in Abb. 9.3a angedeutet, setzt die Absorption bereits bei einer Photonenenergie ein, die kleiner ist als der direkte Bandabstand $(E'_{\mathrm{L}} - E_{\mathrm{V}})$. Liegt das tiefste Leitungsbandminimum bei $\boldsymbol{k}_{\mathrm{m}}$, so kann aufgrund des kleinen Photonenimpulses ein Interbandübergang nicht durch die Absorption eines Photons allein erfolgen. Da Energie und Quasiimpuls erhalten bleiben, ist die Mitwirkung eines Phonons erforderlich. Bezeichnen wir mit ω_q und $\boldsymbol{q}$ dessen Kreisfrequenz bzw. dessen Wellenvektor, so müssen bei der Erzeugung eines Elektron-Loch-Paars mit der kleinstmöglichen Photonenenergie die Bedingungen

$$\hbar\omega_{\mathrm{l}} \pm \hbar\omega_q = E_{\mathrm{g}} \qquad \text{und} \qquad \hbar\boldsymbol{k}_{\mathrm{l}} \pm \hbar\boldsymbol{q} = \hbar\boldsymbol{k}_{\mathrm{m}} \tag{9.1}$$

erfüllt sein. Da $\hbar\omega_q \ll E_{\mathrm{g}}$ und $|k_{\mathrm{l}}| \ll |\boldsymbol{k}_{\mathrm{m}}|$ ist, liefert das Photon die Energie und das Phonon den für den Übergang erforderlichen Impuls. Da das beteiligte Elektron nicht nur an das Photon, sondern auch an das Phonon ankoppeln muß, ist die Wahrscheinlichkeit für das Eintreten dieses Prozesses wesentlich kleiner als beim direkten Prozeß und die damit verbundene Absorption deshalb wesentlich schwächer. Die experimentellen Kurven für die Absorption von Germanium in Abb. 9.3b machen dies deutlich. Der Absorptionskoeffizient ist bei der Grenzenergie relativ klein und wächst mit der Photonenenergie an. Der Anstieg der Absorptionskurve verstärkt sich merklich, sobald der in Abb. 9.3a durch den gestrichelten Pfeil angedeutete direkte Prozeß einsetzt. Die wiedergegebenen Meßdaten zeigen darüber hinaus, daß die Lage der Absorptionskante von der Temperatur abhängt, denn wie bereits erwähnt, vergrößert sich die Bandlücke mit abnehmender Temperatur.

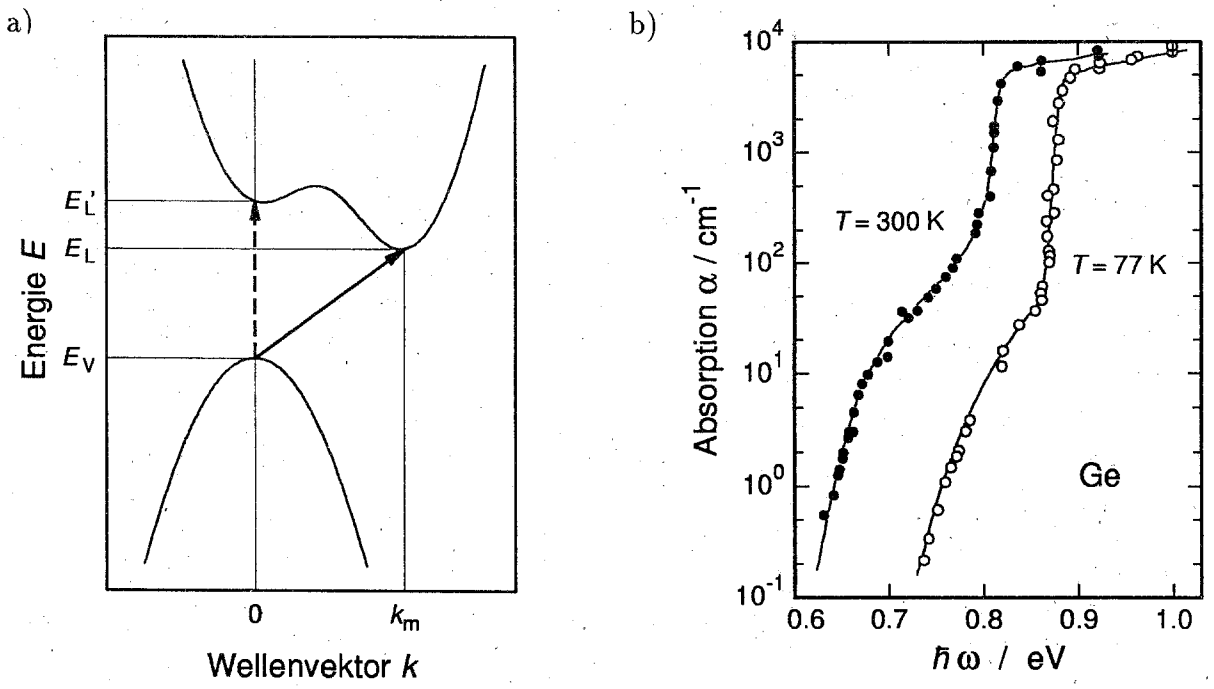


Abb. 9.3: Optische Absorption bei indirekter Bandlücke.
a) Schematische Darstellung des Prozesses. Die Energie des Leitungsbandminimums am Γ-Punkt ist mit E'_{L} bezeichnet. Der durchgezogene Pfeil markiert den Übergang mit der kleinstmöglichen Energie, bei dem ein Phonon mitwirkt. Gestrichelt gezeichnet ist der vertikale direkte Übergang mit der kleinsten Energie.
b) Optischer Absorptionskoeffizient von Germanium als Funktion der Energie der eingestrahlten Photonen. Die schwächere indirekte Absorption ist der direkten vorgelagert. (Nach W.C. Dash und R. Newman, Phys. Rev. **99** (1955) 1151).

Effektive Masse von Elektronen und Löchern

Die elektrischen Eigenschaften von Halbleitern werden vor allem durch Elektronen im Leitungsbandminimum und durch Löcher im Valenzbandmaximum bestimmt. Wir müssen daher die Bandstruktur in diesen speziellen Energiebereichen etwas näher kennenlernen. In Kap. 8.1 haben wir gefunden, daß die effektive Masse von Elektronen und Löchern durch die Bandkrümmung bestimmt wird. In der Umgebung der Bandextrema ist die Krümmung näherungsweise konstant und damit auch die dynamische effektive Masse, die dort sehr gut mit der Zyklotronmasse (vgl. Abschnitt 8.3) übereinstimmt. Messungen der Zyklotronresonanz bieten sich daher zur Massenbestimmung an. Bei Halbleitern ist die Eindringtiefe von Mikrowellen relativ groß, so daß man für eine bestimmte effektive Masse meist nur *eine* Resonanz beobachtet.

Experimente zur Zyklotronresonanz werden bei Mikrowellenfrequenzen durchgeführt. Aus $\omega_c = e\mathcal{B}/m$ errechnet sich für ein Feld von 1 Tesla eine Resonanzfrequenz von 28 GHz. Abb. 9.4 zeigt eine Messung an Germanium bei 23 GHz. Wird das Magnetfeld durchgestimmt, so treten eine Reihe von Resonanzen auf, die Elektronen und Löchern in unterschiedlichen Bändern zugeordnet werden können. Die Bezeichnung und Bedeutung der einzelnen Maxima werden wir sogleich diskutieren. Bei dem gezeigten Experiment wurde der Germaniumkristall im Magnetfeld so orientiert, daß alle unterschiedlichen effektiven Massen beobachtet werden können. Um gute Meßsignale zu erhalten, müssen die Elektronen ihre Bahn mehrmals stoßfrei umlaufen, d.h. es muß $\omega_c\tau \gg 1$ sein. Da die Stoßzeiten τ bei Zimmertemperatur um 10^{-13} s liegen, sind derartige Experimente nur an sehr reinen Proben bei tiefen Temperaturen erfolgreich. Dabei erweist sich die kleine Zahl von verfügbaren Ladungsträgern als problematisch. Wie wir im vorhergehenden Abschnitt gesehen haben, läßt sich jedoch durch Einstrahlung von Licht mit einer Energie größer als die Bandlücke deren Konzentration künstlich erhöhen.

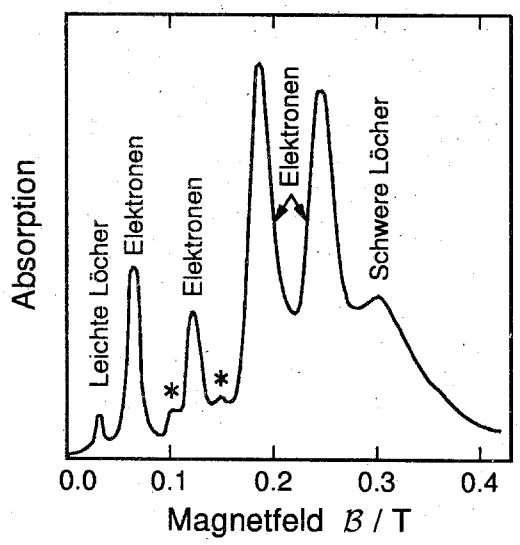


Abb. 9.4: Zyklotronresonanz an Germanium bei 4 K und 23 GHz. Bei der gewählten Probenorientierung treten als Funktion des angelegten Magnetfelds alle Sorten von Ladungsträgern in Erscheinung. Die mit „Elektronen" bezeichneten Maxima stammen von unterschiedlichen Extremalbahnen der Elektronen im Leitungsband. Bei den mit * gekennzeichneten Maxima handelt es sich um Oberschwingungen „schwerer" Löcher. (Nach R.N. Dexter, H.J. Zeiger, B. Lax, Phys. Rev. **104** (1956) 637).

Durch geeignete Wahl von Magnetfeldstärke und -richtung können die Massen der unterschiedlichen Elektronen und Löcher bestimmt werden. Daraus lassen sich Schlüsse auf die Bandstruktur ziehen, die zu folgenden Vorstellungen führen: Bei den Halbleitern mit direkter Bandlücke wie GaAs (vgl. Abb. 9.1a) sind für das *Leitungsband* die elektronischen Eigenschaften am Γ-Punkt von entscheidender Bedeutung. Bei den meisten

direkten Halbleitern hängt die Energie kaum von der Richtung ab, so daß für Elektronen[1] nur eine effektive Masse m_n^* auftritt und die Dispersionsrelation die einfache Form

$$E_n = E_L + \frac{\hbar^2 k^2}{2m_n^*} \tag{9.2}$$

hat. In Tabelle 9.2 sind die effektiven Elektronenmassen einiger Halbleiter aufgeführt. Bemerkenswerterweise sind sie meist wesentlich kleiner als die Masse freier Elektronen.

Tabelle 9.2: Effektive Massen von Elektronen im Leitungsband von Halbleitern.

	m_n^*/m	m_l^*/m	m_t^*/m
GaAs	0,067	—	—
GaSb	0,047	—	—
InSb	0,015	—	—
InP	0,073	—	—
Si	—	0,98	0,19
Ge	—	1,64	0,082

Bei Halbleitern mit indirekter Bandlücke wie Germanium oder Silizium ist die Bandstruktur komplizierter. In Germanium liegen die Leitungsbandmimima in ⟨111⟩-Richtung an den L-Punkten. In Silizium findet man die Minima in ⟨100⟩-Richtung in der Nähe der X-Punkte. Diese Energieminima sind nicht isotrop, sondern besitzen die Gestalt von Rotationsellipsoiden. Zur Veranschaulichung sind in Abb. 9.5 Flächen konstanter Energie des Leitungsbands von Germanium dargestellt. Die Größe der acht Halbellipsoide im Vergleich zur Brillouin-Zone hängt von der gewählten Energie der dargestellten Fläche ab. Sie hat nichts mit der Besetzung des Leitungsbands zu tun, da gewöhnlich nur wenige Elektronen angeregt sind. Im periodisch wiederholten Zonenschema sind die Energieflächen vollständige Ellipsoide, auf denen die Elektronen bei Zyklotronresonanzmessungen umlaufen. Abhängig von der Magnetfeldrichtung tragen unterschiedliche Extremalbahnen zum Signal bei. Durch Verkippen der Proben läßt sich daher die Form der Energieflächen ausmessen.

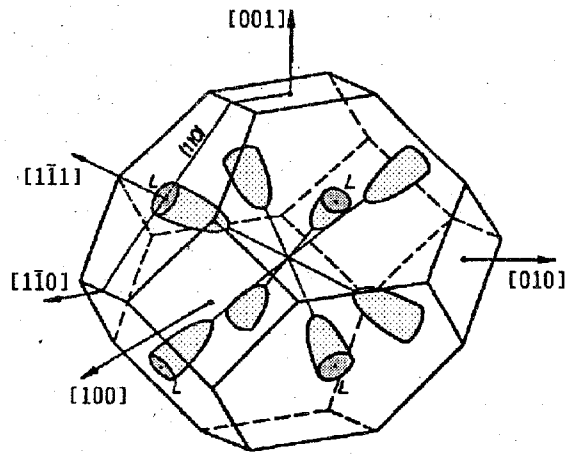


Abb. 9.5: Flächen konstanter Energie der Leitungselektronen von Germanium. Die Minima des Leitungsbands findet man in ⟨111⟩-Richtung an den acht L-Punkten.

[1] Wie in Abschnitt 8.1 benutzen wir zur Charakterisierung der effektiven Massen von Elektronen und Löchern den Index n bzw. p.

Die effektive Masse wird durch die Angabe von zwei Größen beschrieben, durch die *longitudinale Masse* m_l in ⟨111⟩- bzw. ⟨100⟩-Richtung und die *transversale Masse* m_t, die das Verhalten senkrecht dazu charakterisiert. Die Energiefläche läßt sich dann in der Nähe der Minima durch die Gleichung

$$E_\mathrm{n} = E_\mathrm{L} + \hbar^2 \left[\frac{k_1^2 + k_2^2}{2m_\mathrm{t}^*} + \frac{k_3^2}{2m_\mathrm{l}^*} \right] \tag{9.3}$$

angeben, wobei die Richtung von k_i durch die jeweiligen Hauptachsen der Ellipsoide vorgegeben ist. Abhängig von der Richtung unterscheiden sich die Krümmungen der Energiefläche beträchtlich und damit auch die auftretenden effektiven Massen. Für Silizium und Germanium sind sie in Tabelle 9.2 aufgeführt. Die vier Absorptionsmaxima, die in Abb. 9.4 auftreten und von den Elektronen im Leitungsband herrühren, stammen von verschiedenen Extremalbahnen, da die Probe nicht längs einer bevorzugten Achse ausgerichtet war.

Nun kommen wir zu den Löchern im Valenzband. Eine genauere Betrachtung der Abb. 9.1 läßt erkennen, daß bei den dort dargestellten Halbleitern im Valenzbandmaximum zwei Bänder zusammenfallen. Aufgrund ihrer unterschiedlichen Krümmung besitzen die Löcher unterschiedliche effektive Massen und werden als *schwere* und *leichte Löcher* bezeichnet. Wie in Abb. 9.4 zu sehen ist, findet man beide Sorten von Löchern bei Messungen der Zyklotronresonanz. Ein weiteres Band ist aufgrund der Spin-Bahn-Kopplung um die Energie Δ abgesenkt. Die Ladungsträger dieses Bands werden als *abgespaltene Löcher* bezeichnet und besitzen die effektive Masse m_Δ^*. Das dazugehörende Band ist in Abb. 9.1 knapp unterhalb des Valenzbandmaximums zu erkennen. In grober Näherung können bei fast allen Halbleitern die Valenzbänder am Γ-Punkt als kugelförmig angesehen werden, so daß *eine* effektive Masse pro Band zur Charakterisierung ausreicht. In Tabelle 9.3 sind für einige Halbleiter die effektiven Massen von leichten und schweren Löchern sowie die Energie Δ der Spin-Bahn-Kopplung zusammengestellt.

Tabelle 9.3: Effektive Massen von Löchern und Energie Δ einiger Halbleiter. Die Indizes s und l stehen für „schwer“ und „leicht“.

	$m_\mathrm{p,s}^*/m$	$m_\mathrm{p,l}^*/m$	m_Δ^*/m	Δ (eV)
GaAs	0,45	0,082	0,17	0,34
GaSb	0,3	0,06	0,14	0,80
InSb	0,39	0,021	0,11	0,82
InP	0,4	0,078	0,15	0,11
Si	0,49	0,16	0,23	0,044
Ge	0,28	0,044	0,075	0,29

Ladungsträgerdichte

In Halbleitern tragen sowohl Elektronen als auch Löcher zum Stromtransport bei. Für die Leitfähigkeit schreiben wir daher in Erweiterung von (8.18)

$$\sigma = e(n\mu_\mathrm{n} + p\mu_\mathrm{p}) \,, \tag{9.4}$$

wobei n und p für die Dichte der freien Elektronen und Löcher stehen. Beim Beitrag der Löcher zum Stromtransport überwiegt der Anteil, der durch die leichten Löcher zustande kommt. Da die Ladungsträgerdichten überaus stark von der Temperatur abhängen, ändert sich die Leitfähigkeit ebenfalls stark mit der Temperatur. Wir diskutieren erst die Ladungsträgerdichten und gehen anschließend auf die Beweglichkeiten μ_n und μ_p ein.

Zunächst betrachten wir die *Eigenleitung* oder *intrinsische Leitfähigkeit* von halbleitenden Kristallen, die in sehr reinen Kristallen beobachtet werden kann. Durch thermische Anregung werden Elektronen ins Leitungsband gehoben und gleichzeitig Löcher im Valenzband erzeugt. Bei der Ermittlung der Ladungsträgerkonzentration brauchen wir nicht zwischen direkten und indirekten Halbleitern zu unterscheiden, da im thermischen Gleichgewicht die Dynamik der Anregung keine Rolle spielt.

Die Elektronenkonzentration n im Leitungsband erhalten wir durch Integration über das Produkt aus Zustandsdichte und Besetzung der Zustände. Da die Fermi-Funktion $f(E,T)$ mit zunehmender Energie rasch abnimmt, sind nur die Zustände an der unteren Leitungsbandkante von Bedeutung, und wir dürfen als obere Integrationsgrenze anstelle der Energie der Oberkante den Wert Unendlich einsetzen. Die gleiche Argumentation gilt auch für die Berechnung der Löcherkonzentration p. Somit können wir

$$n = \int_{E_L}^{\infty} D_L(E)\, f(E,T)\, \mathrm{d}E \qquad \text{und} \qquad p = \int_{-\infty}^{E_V} D_V(E)\, [1 - f(E,T)]\, \mathrm{d}E \tag{9.5}$$

schreiben. Die Besetzungswahrscheinlichkeit der Löcher ist $[1 - f(E,T)]$, denn sie entstehen durch unbesetzte Elektronenzustände. Da in der Nähe der Bandextrema die Dispersionskurven parabelförmig verlaufen, können wir für die Zustandsdichten D_L und D_V der beiden Bänder die des freien Elektronengases benutzen, wenn wir in Gleichung (7.8) die Masse der freien Elektronen durch die effektive Masse m^* ersetzen:

$$D_L(E) = \frac{(2m_n^*)^{3/2}}{2\pi^2\hbar^3}\sqrt{E - E_L} \qquad \text{für} \qquad E > E_L\,, \tag{9.6}$$

$$D_V(E) = \frac{(2m_p^*)^{3/2}}{2\pi^2\hbar^3}\sqrt{E_V - E} \qquad \text{für} \qquad E < E_V\,. \tag{9.7}$$

Die beiden Bänder sind durch eine Energielücke voneinander getrennt, d.h. im Bereich $E_V < E < E_L$ existieren keine Zustände.

Wie wir gleich sehen werden, liegt das Fermi-Niveau bei Eigenleitung ungefähr in der Mitte der Energielücke. Dies bedeutet, daß sein Abstand von den Bandkanten, also $|E_L - E_F|$ bzw. $|E_V - E_F|$, immer sehr viel größer als die thermische Energie $k_B T$ ist. Bei der Integration von Gleichung (9.5) dürfen wir deshalb die Fermi-Funktion durch eine Exponentialfunktion ersetzen, denn es gilt

$$f(E,T) = \frac{1}{\mathrm{e}^{(E-E_F)/k_B T} + 1} \approx \mathrm{e}^{-(E-E_F)/k_B T}\,. \tag{9.8}$$

Auf die Ladungsträger von Halbleitern ist die Boltzmann-Statistik anwendbar, weil nur die Ausläufer der Fermi-Verteilung bis in die Bänder reichen. Die verfügbaren Zustände sind dann nur schwach besetzt, und die Annahmen der Boltzmann-Statistik sind gerechtfertigt. Elektronen im Leitungsband und Löcher im Valenzband bewegen sich wie die Atome klassischer Gase und können mit Hilfe der kinetischen Gastheorie beschrieben werden. Einschränkend muß natürlich gesagt werden, daß diese Aussage nur Gültigkeit besitzt, solange das Fermi-Niveau genügend weit von der Bandkante entfernt ist. Wie wir noch sehen werden, gilt diese *Näherung der Nichtentartung* bei intrinsischen Halbleitern und bei solchen mit kleiner Störstellendichte. Bei starker Dotierung schiebt sich das Fermi-Niveau in die Nähe der Bandkanten oder liegt sogar in einem Band, und die Näherung bricht zusammen. Es liegt dann ein *entarteter Halbleiter* vor; auf diese Situation werden wir noch kurz am Ende dieses Kapitels eingehen.

An dieser Stelle soll darauf hingewiesen werden, daß für die Fermi-Energie unterschiedliche Bezeichnungen zu finden sind. Oft wird nur das in Kap. 7.1 eingeführte chemische Potential bei $T = 0$ als Fermi-Energie bezeichnet. Wir werden hier die Lage des chemischen Potentials $\mu(T)$ mit dem Begriff *Fermi-Niveau* und – nicht ganz konsequent – mit der Abkürzung E_F belegen, um eine Verwechslung mit der Beweglichkeit μ zu vermeiden.

Nach dem Übergang von der Fermi- zur Boltzmann-Verteilung läßt sich die Ladungsträgerkonzentration auf einfache Weise berechnen:

$$n = \int D_\mathrm{L}(E)\, f(E)\, \mathrm{d}E = \frac{(2m_\mathrm{n}^*)^{3/2}}{2\pi^2\hbar^3}\, \mathrm{e}^{E_\mathrm{F}/k_\mathrm{B}T} \int_{E_\mathrm{L}}^{\infty} \sqrt{E - E_\mathrm{L}}\, \mathrm{e}^{-E/k_\mathrm{B}T}\, \mathrm{d}E\,. \tag{9.9}$$

Das Integral läßt sich leicht analytisch lösen, und wir erhalten für die beiden Ladungsträgertypen

$$n = 2\left(\frac{m_\mathrm{n}^* k_\mathrm{B}T}{2\pi\hbar^2}\right)^{3/2} \mathrm{e}^{-(E_\mathrm{L}-E_\mathrm{F})/k_\mathrm{B}T} = N_\mathrm{L}\mathrm{e}^{-(E_\mathrm{L}-E_\mathrm{F})/k_\mathrm{B}T}\,, \tag{9.10}$$

$$p = 2\left(\frac{m_\mathrm{p}^* k_\mathrm{B}T}{2\pi\hbar^2}\right)^{3/2} \mathrm{e}^{(E_\mathrm{V}-E_\mathrm{F})/k_\mathrm{B}T} = N_\mathrm{V}\mathrm{e}^{(E_\mathrm{V}-E_\mathrm{F})/k_\mathrm{B}T}\,. \tag{9.11}$$

Die hier eingeführten *effektiven Zustandsdichten* N_L und N_V sind, verglichen mit dem Exponentialfaktor, nur „schwach" temperaturabhängig. In dieser Näherung sieht es aus, als wären nicht breite Bänder sondern zwei Niveaus bei E_L bzw. E_V vorhanden.

Die Ladungsträgerdichten n und p werden hauptsächlich durch die Lage des Fermi-Niveaus bestimmt, nicht aber das Produkt $n \cdot p$, denn bei der Produktbildung fällt E_F heraus. Berücksichtigen wir, daß die Bandlücke durch die Differenz $E_\mathrm{g} = E_\mathrm{L} - E_\mathrm{V}$ gegeben ist, so erhalten wir aus (9.10) und (9.11) die Beziehung

$$n \cdot p = N_\mathrm{L}N_\mathrm{V}\, \mathrm{e}^{-E_\mathrm{g}/k_\mathrm{B}T} = 4\left(\frac{k_\mathrm{B}T}{2\pi\hbar^2}\right)^3 \left(m_\mathrm{n}^* m_\mathrm{p}^*\right)^{3/2} \mathrm{e}^{-E_\mathrm{g}/k_\mathrm{B}T}\,. \tag{9.12}$$

Das Produkt der Ladungsträgerdichten ist vollständig charakterisiert durch die effektiven Massen und den Bandabstand und hat bei vorgegebener Temperatur für jeden nichtentarteten Halbleiter einen charakteristischen Wert. In Anlehnung an die Thermodynamik wird die Beziehung $n \cdot p = \text{const.}$ oft auch als *Massenwirkungsgesetz* bezeichnet.

In die bisherigen Überlegungen sind keine Annahmen eingegangen, die unsere Diskussion auf Eigenleiter eingeschränkt hätten. Wir werden die hier abgeleiteten Beziehungen daher später auch bei den dotierten Halbleitern benutzen. Neben (9.10) und (9.11) wird besonders das Massenwirkungsgesetz eine wichtige Rolle spielen. Im Rest dieses Abschnitts wollen wir uns aber auf intrinsische Halbleiter beschränken. Da bei ihnen die Leitungselektronen allein aus dem Valenzband stammen, gilt die wichtige Beziehung $n_\mathrm{i} = p_\mathrm{i}$. Den Index i werden wir in diesem Kapitel ausschließlich zur Charakterisierung der intrinsischen Größen nutzen. Aus (9.12) folgt für die Eigenleiter

$$n_\mathrm{i} = p_\mathrm{i} = \sqrt{N_\mathrm{L}N_\mathrm{V}}\, \mathrm{e}^{-E_\mathrm{g}/2k_\mathrm{B}T} = 2\left(\frac{k_\mathrm{B}T}{2\pi\hbar^2}\right)^{3/2} \left(m_\mathrm{n}^* m_\mathrm{p}^*\right)^{3/4} \mathrm{e}^{-E_\mathrm{g}/2k_\mathrm{B}T}\,. \tag{9.13}$$

Typische Zahlenwerte für die Energielücke und die intrinsischen Ladungsträgerdichte sind in Tabelle 9.4 aufgeführt. In Silizium und Galliumarsenid ist die Dichte bei Zimmertemperatur sehr gering. Es ist daher äußerst schwer bzw. unmöglich, Eigenleitung bei Raumtemperatur zu beobachten, denn im allgemeinen verursachen Verunreinigungen eine Ladungsträgerdichte, die höher ist als die intrinsische.

Tabelle 9.4: Bandlücke und Ladungsträgerkonzentration einiger wichtiger Halbleiter bei 300 K.

	E_g (eV)	n_i (m^{-3})
Germanium	0,66	$2{,}4 \cdot 10^{19}$
Silizium	1,12	$1{,}5 \cdot 10^{16}$
Galliumarsenid	1,43	$1{,}8 \cdot 10^{12}$

Die Temperaturabhängigkeit der *Lage* des *Fermi-Niveaus* läßt sich aus der Forderung $n_i = p_i$ für Eigenleitung unmittelbar aus (9.10) und (9.11) bestimmen. Man findet

$$E_F = \frac{E_L + E_V}{2} + \frac{k_B T}{2} \ln\left(\frac{N_V}{N_L}\right) = \frac{E_L + E_V}{2} + \frac{3}{4} k_B T \ln\left(\frac{m_p^*}{m_n^*}\right) . \qquad (9.14)$$

Bei $T = 0$ liegt das Fermi-Niveau also in der Mitte der Energielücke. Sind die effektiven Massen von Elektronen und Löchern gleich, d.h. sind Valenz- und Leitungsband gleich stark gekrümmt, dann ändert sich die Lage auch nicht mit zunehmender Temperatur. Unterscheiden sich die Massen, verschiebt sich das Fermi-Niveau, doch ist der Temperaturgang, verglichen mit der Größe der Bandlücke, nur gering.

In Abb. 9.6 sind die wichtigsten Ergebnisse anschaulich zusammengefaßt. Auf der rechten Seite der Bilder ist jeweils der Verlauf der Elektronen- und Löcherdichte dargestellt. Bei kleiner Energie wird dieser Verlauf durch die parabolische Zustandsdichte, bei großer Energie durch die abfallende Besetzungswahrscheinlichkeit bestimmt. Da bei intrinsischen Halbleitern gleich viele Elektronen und Löcher vorhanden sind, sind die beiden getönten Flächen jeweils gleich groß. Dies ist nur möglich, wenn sich das Fermi-Niveau, wie im rechten Bild zu sehen ist, zum Band mit der geringeren Zustandsdichte schiebt. Bei gleich großer Zustandsdichte in beiden Bändern verläuft das Fermi-Niveau unabhängig von der Temperatur in der Mitte der Lücke.

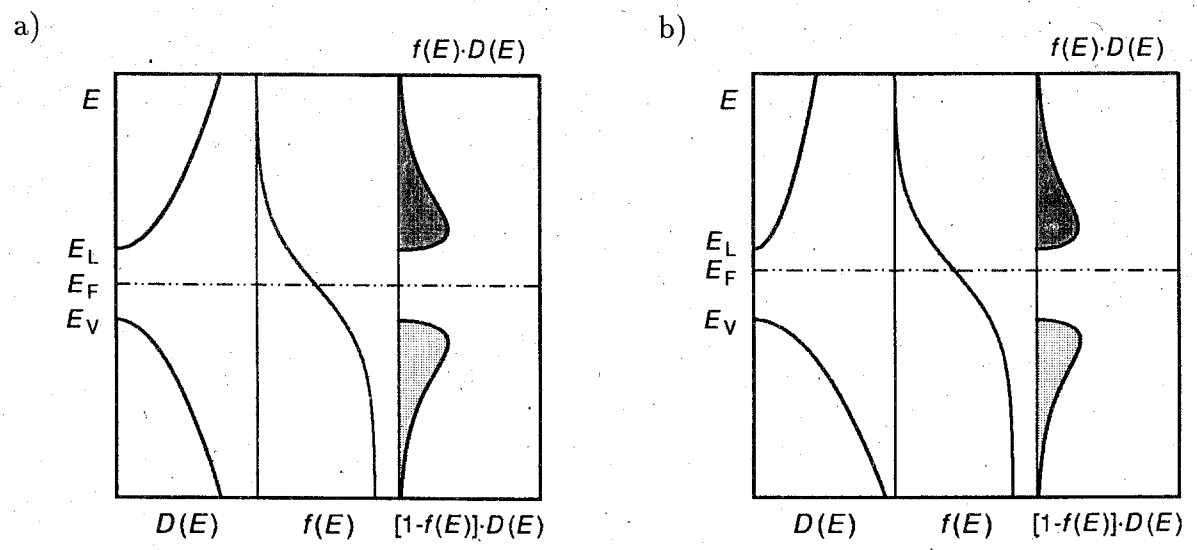


Abb. 9.6: Zustandsdichten $D(E)$, Fermi-Funktion $f(E)$ und das Produkt $f(E) \cdot D_L(E)$ bzw. $[1 - f(E)] \cdot D_V(E)$ in Valenz- und Leitungsband für $T > 0$. Die Elektronen im Leitungsband sind dunkel, die Löcher etwas heller getönt.
a) Halbleiter mit gleicher Zustandsdichte in Valenz- und Leitungsband, d.h. $N_L = N_V$,
b) Halbleiter mit $N_V > N_L$, also mit der größeren Zahl von Zuständen im Valenzband.

9.2 Dotierte kristalline Halbleiter

Dotierung

In realen Kristallen treten immer Verunreinigungen auf, deren Einbau bei der Kristallzucht nicht vermieden werden kann. In der Regel sind sie „*elektrisch aktiv*", d.h. sie verursachen zusätzliche Ladungsträger in den Bändern. So findet man bei guten GaAs-Kristallen bei Raumtemperatur eine Ladungsträgerkonzentration von $n \approx 10^{22}\,\mathrm{m}^{-3}$ statt des intrinsischen Werts $n_\mathrm{i} \approx 1{,}8 \cdot 10^{12}\,\mathrm{m}^{-3}$. Für Anwendungszwecke wäre die elektrische Leitfähigkeit des intrinsischen Materials wegen der kleinen Ladungsträgerdichte ohnehin zu klein, denn setzt man Zahlenwerte in Gleichung (9.3) für die Leitfähigkeit ein, so findet man den kleinen Wert $\sigma_\mathrm{i} \approx n_\mathrm{i} e \mu_\mathrm{n} \approx 3 \cdot 10^{-8}\,(\Omega\mathrm{m})^{-1}$.

Für technische Anwendungen werden in Halbleitern gezielt elektrisch aktive Verunreinigungen eingebaut, d.h. die Kristalle werden *dotiert*. Grundsätzlich unterscheidet man zwischen zwei verschiedenen Sorten von elektrisch aktiven Störstellen, auf die wir bereits kurz in Abschnitt 4.1 eingegangen sind und die in Abb. 4.8 veranschaulicht wurden. *Donatoren* sind Störstellen, die Elektronen ins Leitungsband abgeben können, wodurch ein n-Halbleiter entsteht. Dagegen können *Akzeptoren* Elektronen aus dem Valenzband aufnehmen. Dadurch entstehen Löcher im Valenzband, und der dotierte Halbleiter wird p-leitend.

Wie bereits bei der Diskussion der Fremdatome in Abschnitt 4.1 erwähnt, entstehen Donatoren durch den Einbau von Atomen, die eine höhere Wertigkeit als die Atome des Wirtsgitters haben. So sind die Elemente P, As oder Sb geeignete Kandidaten für vierfach-koordinierte Halbleiter. Anstelle der sp^3-Konfiguration des Siliziums liegt bei atomarem Phosphor eine s^2p^3-Konfiguration vor. Phosphoratome weisen daher bei einem Einbau in eine tetraedrische Bindung, wie sie durch die Siliziumnachbarn erzwungen wird, ein überschüssiges Elektron auf. Dieses Elektron kann nicht an der Bindung teilnehmen und ist nur schwach an den positiven Donatorrumpf gebunden. Einen Akzeptor erhält man in vierfach-koordinierten Halbleitern durch Dotieren mit dreiwertigen Atomen, wie z.B. mit B, Al, Ga oder In. Diese Atome haben ein Elektron weniger als die Atome des Wirtsgitters, dies entspricht einem Loch, das schwach an das negativ geladene Fremdatom gebunden ist. Bei genügend hohen Temperaturen sind die Störstellen ionisiert, d.h. die Donatoren geben ihr schwach gebundenes Elektron ans Leitungsband, die Akzeptoren ihr Loch ans Valenzband ab.

Die erlaubten Eigenwerte dieser Störstellen entsprechen in einfachster Näherung den Zuständen eines Wasserstoffatoms. Wir schreiben (4.10) nochmals an:

$$E_\nu = \frac{m^* e^4}{32\pi^2\hbar^2\epsilon_\mathrm{r}^2\epsilon_0^2} \cdot \frac{1}{\nu^2} \,. \tag{9.15}$$

Die Abschirmung des Coulomb-Felds durch die Ladungen der übrigen Halbleiteratome wird hier pauschal durch die makroskopische Dielektrizitätskonstante ϵ_r berücksichtigt. Je nach Art der Störstelle muß für die Masse m^* die effektive Elektronen- oder Löchermasse eingesetzt werden. Für $\nu = 1$ erhält man die Ionisationsenergie, die in Abb. 9.7 mit E_d bzw. E_a bezeichnet wird. Für Donatoren in Silizium findet man mit $\epsilon_{\mathrm{r,Si}} = 11{,}7$ den Wert $E_\mathrm{d} \approx 30\,\mathrm{meV}$. Das Leitungsband entspricht dem Vakuum beim Wasserstoffatom, d.h.

der Grundzustand E_D des Donators liegt 30 meV unter E_L. Als effektive Masse wurde bei der Abschätzung der Wert $m_n^* \approx 0{,}3\,m$ benutzt. Ähnlich liegen die Verhältnisse bei Germanium. Da dort $\epsilon_{r,Ge} = 15{,}8$ und $m_n^* \approx 0{,}15\,m$ ist, ist die Ionisationsenergie mit $E_d \approx 9$ meV noch kleiner. Die Donator- und Akzeptoratome selbst sind nicht beweglich und tragen nicht direkt zur Leitfähigkeit bei. Bei tiefen Temperaturen befinden sich die Störstellen im Grundzustand, sind also neutral. Bei Raumtemperatur sind sie aufgrund ihrer kleinen Ionisationsenergie positiv bzw. negativ geladen.

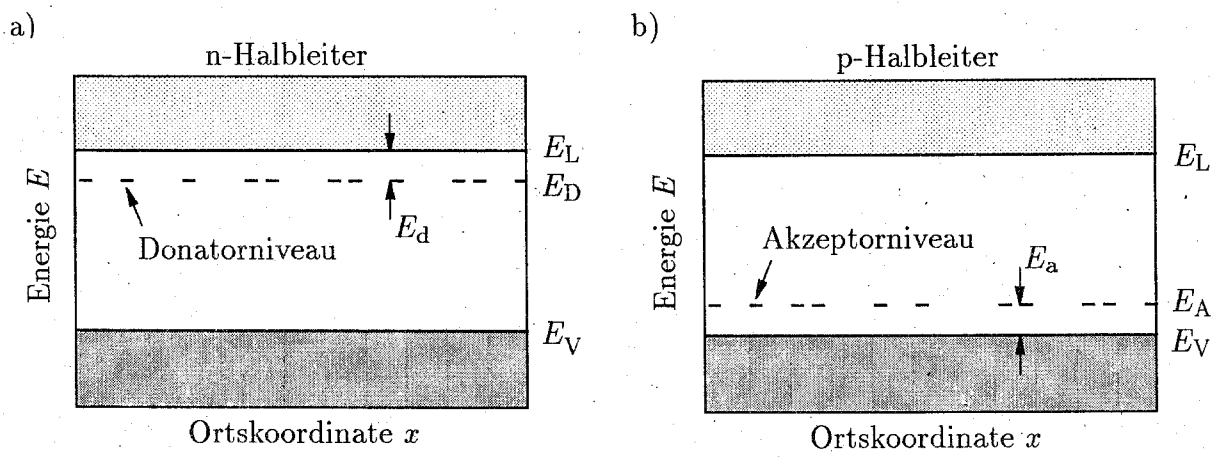


Abb. 9.7: „Ortsdarstellung" der typischen Energien in Halbleitern.
a) Der Grundzustand der Donatoren E_D liegt unmittelbar unter der Leitungsbandkante E_L, ihre Ionisationsenergie ist E_d.
b) Die Akzeptorniveaus mit der Energie E_A liegen knapp über der Valenzbandkante E_V. Die Ionisationsenergie der Löcher bezeichnen wir mit E_a.

Im Wasserstoffmodell ist die Lage der Störstellenniveaus von der Natur der Donator- oder Akzeptoratome unabhängig. In Tabelle 9.5 sind für Silizium und Germanium die experimentellen Werte der Ionisationsenergien einiger Störstellen angegeben. Bedenkt man, wie grob wir bei der Berechnung vorgegangen sind, so ist die Übereinstimmung mit den Werten der einfachen Rechnung bemerkenswert. Offensichtlich läßt sich die Abschirmung gut mit der Dielektrizitätskonstanten beschreiben. Der Grund hierfür ist der große Radius $r = 4\pi\epsilon_r\epsilon_0\hbar^2/m^*e^2$ der „Wasserstoffatome". Der Bahnradius des gebundenen Störelektrons ist gegenüber dem Radius beim wirklichen Wasserstoffatoms um den Faktor $\epsilon_r\, m/m^* \approx 40$ vergrößert.

Tabelle 9.5: Ionisationsenergien (in meV) von Donatoren und Akzeptoren in Silizium und Germanium.

Donator	P	As	Sb	Bi
Silizium	45	49	39	69
Germanium	12	12,7	9,6	—
Akzeptor	**B**	**Al**	**Ga**	**In**
Silizium	45	57	65	16
Germanium	10,4	10,2	10,8	11,2

In Abb. 9.8 ist die Lage der Störstellenniveaus in Germanium für eine Reihe von unterschiedlichen Dotierelementen dargestellt. Offensichtlich läßt sich die Lage der Störstellen bei den Elementen, die nicht aus der dritten oder fünften Hauptgruppe stammen, nicht in das einfache Bild einordnen. Insbesondere können wie bei Zink, Chrom oder Kupfer mehrere Ladungszustände mit unterschiedlicher Energie auftreten. Gold kann in Germanium vier verschiedene Niveaus bilden und sowohl als Donator als auch als Akzeptor wirken.

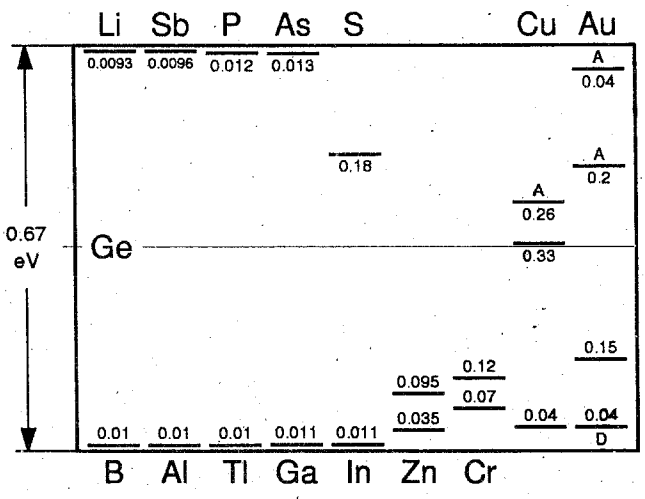


Abb. 9.8: Störstellenniveaus in Germanium. Die Zahlen geben den Abstand von der nächstgelegenen Bandkante an. Die Buchstaben A und D stehen für „Donator" bzw. „Akzeptor", wenn sich die Zuordnung der Zustände nicht unmittelbar aus ihrer Lage ergibt. (Nach S.M. Sze, *Physics of Semiconductor Devices*, J. Wiley, New York 1981).

Ladungsträgerdichte und Fermi-Niveau in dotierten Halbleitern

Wir wollen nun die Ladungsträgerdichten in dotierten Halbleitern berechnen und beschränken uns dabei auf nicht-entartete Halbleiter. Die Wechselwirkung zwischen den Störstellen vernachlässigen wir, da sie erst bei hohen Konzentrationen wichtig wird.

In realen Halbleitern sind durch unkontrollierbare Verunreinigung immer ungewollte Donatoren und Akzeptoren vorhanden. Dies hat zur Folge, daß die energetisch höherliegenden Donatoren Elektronen an die tieferliegenden Akzeptoren abgeben, doch ist dieser Elektronentransfer nicht mit dem Auftreten von freien Ladungsträgern verbunden. Die Wirkung der beiden Sorten von Störstellen hebt sich gerade auf, ein Effekt, den man als *Kompensation* bezeichnet. Bei technischen Materialien nutzt man diesen Kompensationseffekt bei der Herstellung von hochohmigen Halbleitern aus. Dabei werden die meist bekannten, aber unvermeidbaren Verunreinigungen durch gezielte Zugabe von Dotiermaterial kompensiert.

Die energetische Lage von Bandzuständen hängt im wesentlichen nicht von der Besetzung der Zustände ab. Sie ist unabhängig davon, ob ein Zustand leer, mit einem Elektron oder zwei Elektronen (mit entgegengesetztem Spin) besetzt ist. Dies ist verständlich, da die Elektronen in diesen Zuständen delokalisiert, d.h.über den ganzen Kristall verschmiert sind. Sie „sehen" sich daher gegenseitig nicht; die Einelektron-Näherung hat also Bestand. Elektronen an Störstellen sind dagegen räumlich lokalisiert. Sehen wir von den komplizierter aufgebauten Störstellen ab, die Zustände bei verschiedenen Energien besitzen, so kann im Prinzip jede Störstelle wie bei den Bandzuständen unbesetzt oder mit einem oder zwei Elektronen besetzt sein. Bei zweifacher Besetzung erhöht die Coulomb-Abstoßung zwischen den Elektronen die Energie jedoch erheblich. Doppelt besetzte Zustände liegen

daher sehr häufig im Leitungsband und sind für die Ladungsträgerbilanz ohne Bedeutung. Im Normalfall treten daher nur der einfach geladene und der neutrale Zustand auf.

Sind Störstellen nicht doppelt geladen, so sind sie entweder neutral oder ionisiert. Wir können daher bei der Berechnung der Anzahl der freien Ladungsträger ihre Konzentrationen n_D und n_A in einen neutralen und einen geladenen Anteil aufspalten:

$$n_D = n_D^0 + n_D^+ \qquad \text{und} \qquad n_A = n_A^0 + n_A^- \tag{9.16}$$

Die Besetzung des Grundzustands, d.h. die Wahrscheinlichkeit dafür, daß eine Störstelle nicht ionisiert ist, läßt sich mit Hilfe der Fermi-Dirac-Verteilung ausdrücken:

$$\frac{n_D^0}{n_D} = \left[\frac{1}{g_D}\,\mathrm{e}^{(E_D-E_F)/k_BT} + 1\right]^{-1} \quad \text{und} \quad \frac{n_A^0}{n_A} = \left[g_A\,\mathrm{e}^{(E_F-E_A)/k_BT} + 1\right]^{-1} . \tag{9.17}$$

Die Wichtungsfaktoren g_D und g_A berücksichtigen die Entartung der Störstellenniveaus, wodurch der übliche Ausdruck für die Fermi-Statistik etwas modifiziert wird. Bei den einfachen Donatoren, die wir hier betrachten, kann ein Elektron mit dem Spin nach oben oder unten eingebaut werden. Diese „Wahlmöglichkeit" gibt dem Zustand ein doppeltes statistisches Gewicht und führt zum Wichtungsfaktor $g_D = 2$. Bei Akzeptoren liegen die Verhältnisse etwas komplizierter. Dort muß bei den gebräuchlichen Halbleitern auch noch die Entartung des Valenzbands berücksichtigt werden. Dadurch bekommt der Grundzustand ebenfalls ein höheres Gewicht und erniedrigt g_A wieder. Obwohl der Entartungsgrad bei numerischen Rechnungen von Bedeutung ist, werden wir ihn bei der weiteren Diskussion zur Vereinfachung der Formeln weglassen, da wir nur am prinzipiellen Verhalten interessiert sind.

Freie Ladungsträger mit der Dichte n bzw. p entstehen durch Anregung der Elektronen vom Valenz- ins Leitungsband und durch Ionisation der Störstellen. Im vorhergehenden Abschnitt haben wir bereits die entsprechenden Ausdrücke (9.10) und (9.11) hergeleitet. Da dort keine Annahme einging, die eine Einschränkung auf intrinsische Ladungsträger bewirkt, gelten die beiden Gleichungen auch für dotierte Halbleiter, und wir können sie auch hier nutzen. Hinzu kommt noch die Gleichung

$$n + n_A^- = p + n_D^+ , \tag{9.18}$$

welche die Forderung nach Ladungsneutralität widerspiegelt.

Trotz dieser stark vereinfachenden Annahmen ist die Situation relativ komplex, da sieben Variablen auftreten. Im folgenden wollen wir den einfacheren, aber auch realistischeren Fall betrachten, daß eine Störstellensorte vorherrscht. Wir diskutieren einen *n-Halbleiter*, in dem viele Donatoren und einige wenige Akzeptoren vorhanden sind. Der Übersichtlichkeit halber stellen wir die nun relevanten Gleichungen nochmals zusammen, die wir nicht exakt lösen, sondern nur qualitativ diskutieren.

$$n = N_L \mathrm{e}^{-(E_L-E_F)/k_BT} , \tag{9.19}$$

$$n_D = n_D^0 + n_D^+ , \tag{9.20}$$

$$\frac{n_D^0}{n_D} = \frac{1}{\mathrm{e}^{(E_D-E_F)/k_BT} + 1} , \tag{9.21}$$

$$n + n_A^- = n_D^+ + p . \tag{9.22}$$

Gleichung (9.19) gilt natürlich nur für nicht-entartete Halbleiter, welche die Annahme $E_L - E_F \gg k_B T$ erfüllen. Weiterhin soll die Dotierung so groß sein, daß ihr Beitrag zur Konzentration der freien Ladungsträger überwiegt und somit *Störstellenleitung* vorliegt. Dann ist $n_D^+ \gg n_i$, und die Dichte n_i der intrinsisch angeregten Elektronen kann gegenüber dem Beitrag der Störstellen vernachlässigt werden. Aus dem Massenwirkungsgesetz (9.12) $n \cdot p = n_i \cdot p_i$ folgt unmittelbar, daß in einem n-Leiter praktisch keine Löcher existieren. Wir sehen dies, wenn wir $p = p_i(n_i/n) \approx p_i(n_i/n_D^+) \ll p_i$ schreiben. Eine weitere Vereinfachung folgt aus der oben gemachten Annahme $n_D \gg n_A$. In diesem Fall fangen alle Akzeptoren ein Elektron ein, das ursprünglich zu einem Donator gehörte. Da praktisch keine neutralen Akzeptoren mehr vorkommen, können wir n_A^0 vernachlässigen, und n_A^- durch n_A ersetzen. Damit nimmt (9.22) die einfache Form $n + n_A = n_D^+$ an. In Abb. 9.9 ist die hier geschilderte Situation schematisch dargestellt. In dem nur zum geringen Teil kompensierten n-Halbleiter sind Valenzband und Akzeptorniveaus mit Elektronen besetzt. Ein Teil der Donatoren ist ionisiert und hat die Elektronen in die Akzeptorniveaus und ins Leitungsband abgegeben.

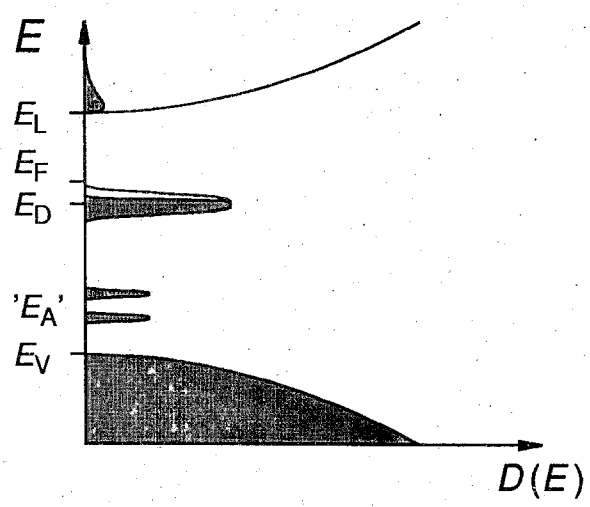


Abb. 9.9: Zustandsdichte und Besetzung der Zustände eines teilkompensierten n-Halbleiters. Mit Elektronen besetzte Zustände sind grau dargestellt. Stellvertretend für die Zustandsdichte von unterschiedlichen Verunreinigungen sind zwei Niveaus eingezeichnet und mit $'E_A'$ bezeichnet. Alle Akzeptoren haben ein Elektron eingefangen.

Mit diesen Vereinfachungen ergeben sich aus (9.20) – (9.22) die folgenden Beziehungen:

$$n_D^+ = n_D - n_D^0 = n + n_A$$

$$n_D^+ = n_D \left[1 - \frac{1}{e^{(E_D - E_F)/k_B T} + 1}\right] . \tag{9.23}$$

Mit Hilfe von (9.19) eliminieren wir noch E_F und erhalten das Endergebnis:

$$\frac{n(n_A + n)}{n_D - n_A - n} = N_L \, e^{-E_d/k_B T} . \tag{9.24}$$

Diese Gleichung wollen wir benutzen, um die Dichte der freien Elektronen in den verschiedenen Temperaturbereichen zu diskutieren. Der qualitative Temperaturverlauf ist in Abb. 9.10 skizziert. Die Lage des Fermi-Niveaus ergibt sich bei Kenntnis der Elektronendichte n unmittelbar aus (9.19) und ist ebenfalls dargestellt.

Bei sehr *tiefen Temperaturen* (Bereich α), d.h. für $k_\mathrm{B}T \ll E_\mathrm{d}$, gilt die Ungleichung $n \ll n_\mathrm{A} \ll n_\mathrm{D}$. Damit lassen sich (9.24) und (9.19) durch die Gleichungen

$$n \approx \frac{N_\mathrm{L} n_\mathrm{D}}{n_\mathrm{A}} \mathrm{e}^{-E_\mathrm{d}/k_\mathrm{B}T} , \tag{9.25}$$

$$E_\mathrm{F} \approx E_\mathrm{L} - E_\mathrm{d} + k_\mathrm{B}T \ln\left(\frac{n_\mathrm{D}}{n_\mathrm{A}}\right) \tag{9.26}$$

nähern. Bei $T \to 0$ wird die Lage des Fermi-Niveaus durch die Donatoren bestimmt, die aufgrund des Kompensationseffekts geladen sind. Es liegt zunächst bei der Donatorenergie und steigt mit der Temperatur an, da dann neutrale Donatoren Elektronen ans Leitungsband abgeben. Gleichzeitig nimmt die Ladungsträgerdichte exponentiell zu.

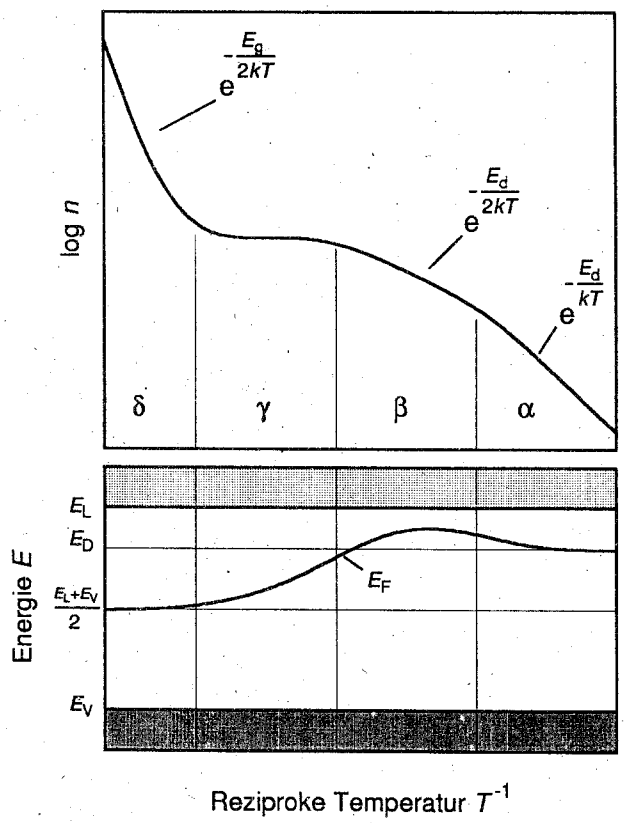


Abb. 9.10: Elektronendichte im Leitungsband eines n-Halbleiters (oben) und Lage des Fermi-Niveaus (unten) als Funktion der reziproken Temperatur. Die unterschiedlichen Temperaturbereiche $\alpha - \delta$ werden im Text diskutiert.

Die Elektronendichte n wächst mit der Temperatur rasch an und übertrifft schließlich die Akzeptordichte, die jedoch immer wesentlich kleiner als die Donatordichte ist. Bei diesen *tiefen Temperaturen* (Bereich β) gilt daher die Ungleichung $n_\mathrm{A} \ll n \ll n_\mathrm{D}$. Damit ergeben sich für die gesuchten Größen die Ausdrücke

$$E_\mathrm{F} \approx E_\mathrm{L} - \frac{E_\mathrm{d}}{2} - \frac{k_\mathrm{B}T}{2} \ln\left(\frac{N_\mathrm{L}}{n_\mathrm{D}}\right) , \tag{9.27}$$

$$n \approx \sqrt{n_\mathrm{D} N_\mathrm{L}}\, \mathrm{e}^{-E_\mathrm{d}/2k_\mathrm{B}T} . \tag{9.28}$$

Das Fermi-Niveau liegt etwa in der Mitte zwischen Leitungsband und Donatorniveau. Der exponentielle Anstieg der Leitungselektronendichte verlangsamt sich daher, verglichen mit dem Anstieg bei sehr tiefen Temperaturen, um den Faktor 2. Anschaulich ausgedrückt übernehmen die Donatoren die Funktion des Valenzbands im Falle der intrinsischen Ladungsträgerkonzentration als Elektronenquelle. Da in diesem Temperaturintervall nur ein Bruchteil der Donatoren ionisiert ist, spricht man von der *Störstellenreserve*.

Bei *Raumtemperatur* (Bereich γ) ist $k_B T \approx E_d$, und wir können in guter Näherung $\exp(-E_d/k_B T) \approx 1$ setzen. Da $n \ll N_L$ ist, folgt aus (9.24) die einfache Beziehung $n_D - n \approx 0$. Die Dichte der freien Ladungsträger wird ausschließlich durch die Zahl der Störstellen bestimmt und ist temperaturunabhängig. Vernachlässigen wir n_A, so gilt

$$E_F \approx E_L - k_B T \ln\left(\frac{N_L}{n_D}\right) , \tag{9.29}$$

$$n \approx n_D = \text{const} . \tag{9.30}$$

Das Fermi-Niveau bewegt sich kontinuierlich nach unten. Da alle Störstellen ionisiert sind, spricht man vom *Erschöpfungszustand.*

Bei *hohen Temperaturen* (Bereich δ) setzt Eigenleitung ein. Das Fermi-Niveau liegt etwa in der Mitte der Bandlücke, und die Konzentration der freien Ladungsträger nimmt mit der Temperatur wieder stark zu. Es gelten die Gleichungen (9.10) und (9.14), die wir für intrinsische Halbleiter hergeleitet haben.

In Abb. 9.11 sind die experimentell bestimmten Dichten der freien Elektronen von n-dotiertem Germanium dargestellt. Die Donatorkonzentration der Proben läßt sich in dieser Abbildung direkt aus der Ladungsträgerdichte bei höheren Temperaturen ablesen.

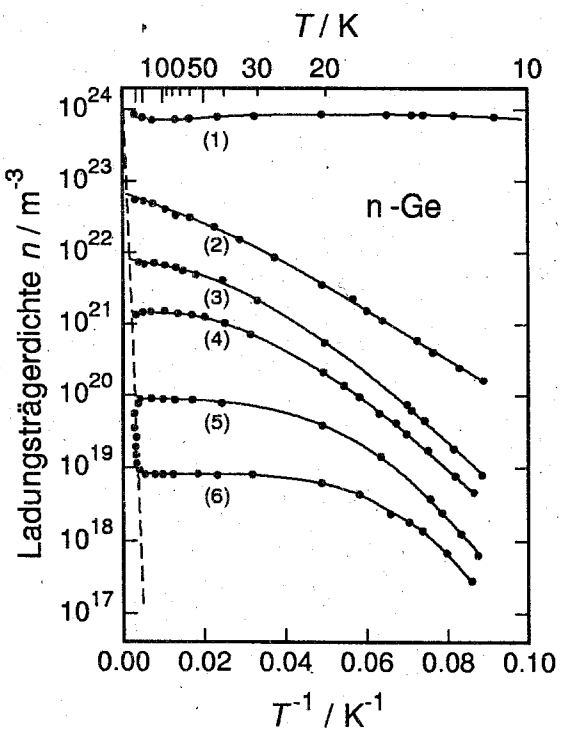


Abb. 9.11: Dichte der Ladungsträger in n-Germanium als Funktion der reziproken Temperatur, gemessen mit Hilfe des Hall-Effekts. Der Bereich der Eigenleitung ist gestrichelt dargestellt. Die Arsendotierung nimmt ungefähr in Zehnerschritten zu und überstreicht den Bereich von $10^{19}\,\mathrm{m}^{-3}$ bis $10^{24}\,\mathrm{m}^{-3}$. (Nach E.M. Conwell, Proc. I.R.E. **40** (1952) 1327).

Deutlich zu erkennen sind die mehr oder weniger ausgeprägten Bereiche der Eigenleitung, des Erschöpfungszustands und der Störstellenreserve. Ein abweichendes Verhalten weist die Probe (1) mit der hohen Donatorkonzentration von $n_D = 10^{24}\,\mathrm{m}^{-3}$ auf, denn die Ladungsträgerdichte ist entgegen den Erwartungen weitgehend temperaturunabhängig. Bei dieser hohen Störstellenkonzentration ist der Abstand zwischen den Donatoren so klein, daß sich ihre Wellenfunktionen überlappen. Dies führt dazu, daß die Elektronen der neutralen Donatoren von einem Donator zum nächsten tunneln können. Es bildet sich bei der Grundzustandsenergie der Donatoren ein *Störstellenband* aus, in dem sich die

Elektronen bewegen. Da jede Störstelle höchstens mit einem Elektron besetzt sein kann, ist dieses Band halbvoll, so daß ähnliche Verhältnisse wie bei den Metallen vorliegen. Da der Abstand zwischen den Störstellen verglichen mit dem Atomabstand sehr groß ist, ist das Band sehr schmal und somit die effektive Masse der Elektronen in diesem Band sehr groß. Gleiches gilt für Akzeptoren: Bei hoher Konzentration treten Störstellenbänder in der Nähe des Valenzbandes auf, in denen Löcherleitung vorliegt.

Die Konzentration der freien Ladungsträger läßt sich am einfachsten mit dem Hall-Effekt bestimmen. Bei Halbleitern tragen beide Ladungsträgersorten bei, und es gilt Gleichung (8.76), die folgendes Aussehen hat:

$$R_{\mathrm{H}} = \frac{p\mu_{\mathrm{p}}^2 - n\mu_{\mathrm{n}}^2}{e(p\mu_{\mathrm{p}} + n\mu_{\mathrm{n}})^2} \,. \tag{9.31}$$

Das Vorzeichen von R_{H} wird von den Ladungsträgern bestimmt, die den Hauptbeitrag zum Stromtransport leisten. Dominiert eine Ladungsträgersorte, so kann man aus der Hall-Spannung sofort die Konzentration der freien Ladungsträger angeben. In diesem Fall ist die Konstante R_{H} unabhängig von den Werten der Beweglichkeit.

In Abb. 9.12 sind Hall-Messungen an n- und p-dotierten InSb-Proben zu sehen. Im Erschöpfungsbereich dominiert eine Ladungsträgersorte, damit ergibt sich die Vereinfachung $R_{\mathrm{H}} \approx -1/en \approx -1/en_{\mathrm{D}}$ bzw. $R_{\mathrm{H}} \approx 1/ep \approx 1/en_{\mathrm{A}}$. Die Ladungsträgerdichte ist konstant, und die Beweglichkeit hat keinen Einfluß auf die Hall-Spannung. Bei hohen Temperaturen fällt die Hall-Konstante exponentiell ab, da bei Eigenleitung die Dichte der freien Ladungsträger exponentiell ansteigt. Aufgrund ihrer höheren Beweglichkeit (vgl. Tabelle 9.6) tragen bei Eigenleitung meist die Elektronen den Hauptteil zur Leitfähigkeit und zur Hall-Konstanten bei. Dies führt zu einem bemerkenswerten Temperaturverlauf bei den p-dotierten Proben: Beim Übergang von der p-Leitung zur Eigenleitung verschwindet nach (9.31) die Hall-Konstante und wechselt das Vorzeichen.

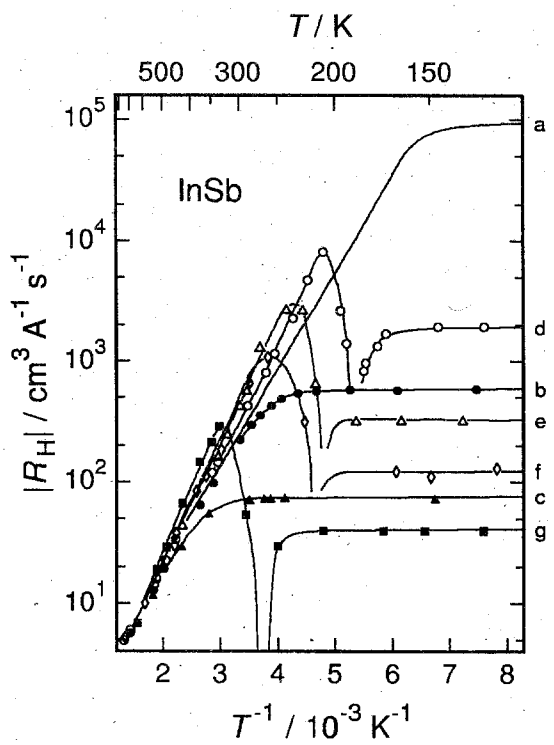


Abb. 9.12: Hall-Konstante $|R_{\mathrm{H}}|$ von Indiumantimonid als Funktion der reziproken Temperatur. Die Proben (a) bis (c) sind n-dotiert, die von (d) bis (g) p-dotiert. (Nach O. Madelung und H. Weiss, Z. Naturf. **9a** (1954) 527).

Beweglichkeit der freien Ladungsträger und elektrische Leitfähigkeit

Zu einem vollständigen Verständnis der Leitfähigkeit fehlt noch die Diskussion der Beweglichkeit $\mu = e\tau/m^*$, für die einige typische Werte in Tabelle 9.6 aufgelistet sind. Neben der effektiven Masse beeinflussen vor allem die Stoßzeiten die Beweglichkeit. Die Elektronen werden in Halbleitern wie in Metallen vor allem an Defekten und Phononen gestreut. Die Elektron-Elektron-Streuung spielt schon wegen der vergleichsweise geringen Elektronendichte keine Rolle. Positive und negative Ladungsträger verhalten sich bezüglich der Stoßzeiten qualitativ gleich, da sich die Streuung der Löcher letztendlich auf die Streuung der Elektronen zurückführen läßt.

Tabelle 9.6: Beweglichkeit μ_n und μ_p von Elektronen und Löchern in verschiedenen Halbleitern bei Raumtemperatur.

	C	Si	Ge	GaAs	InSb	Cu
μ_n (cm^2/Vs)	1 800	1 500	3 900	8 500	80 000	35
μ_p (cm^2/Vs)	1 200	450	1 900	400	1 250	–

In Abb. 9.13 ist die Beweglichkeit der Elektronen in aluminiumdotiertem Mg_2Ge dargestellt. Probe (1) wurde nicht absichtlich dotiert, weist aber dennoch eine Störstellenkonzentration $n_D - n_A = 1{,}3 \cdot 10^{22}\,m^{-3}$ auf. Die Aluminiumkonzentration der Probe (2) liegt bei $4{,}2 \cdot 10^{22}\,m^{-3}$ und bei der am stärksten dotierten Probe (3) schließlich bei $8{,}2 \cdot 10^{23}\,m^{-3}$. Sieht man von Probe (3) ab, in der sich aufgrund der hohen Donatorkonzentration ein Störstellenband ausgebildet hat, so findet man einen starken Anstieg der Beweglichkeit bei tiefen Temperaturen, ein Maximum im Bereich von 50 K bis 100 K und anschließend einen steilen Abfall.

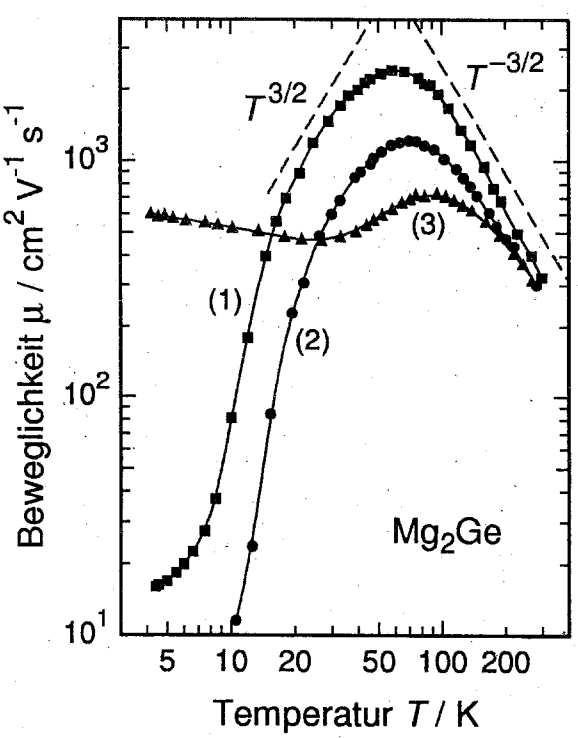


Abb. 9.13: Beweglichkeit der Elektronen in aluminiumdotiertem Mg_2Ge, gemessen mit Hilfe des Hall-Effekts. Die Proben enthalten $1{,}3 \cdot 10^{22}$, $4{,}2 \cdot 10^{22}$ und $8{,}2 \cdot 10^{23}$ Störstellen pro m^3. Bei der Probe (3) liegt Störstellenbandleitfähigkeit vor. (P.W. Li, S.N. Lee, G.C. Danielson, Phys. Rev. B **6** (1972) 442).

Wir wollen zuerst den Temperaturbereich unterhalb des Beweglichkeitsmaximums diskutieren. Bei tiefen Temperaturen dominiert die Streuung an Störstellen, die durch die Dotierung selbst hervorgerufen werden. Da die Elektronendichte in Halbleitern nicht sehr groß ist, ist die Abschirmung wesentlich weniger ausgeprägt als in Metallen. Wir müssen daher berücksichtigen, daß die Störstellen geladen sind, und behandeln den Streuprozeß wie die Rutherford-Streuung. Nach der bekannten Formel aus der Kernphysik ist der totale Streuquerschnitt $\sigma_{\mathrm{st}} \propto v^{-4}$, woraus sich für die reziproke freie Weglänge der Ausdruck $l^{-1} = n_{\mathrm{st}}\sigma_{\mathrm{st}} \propto n_{\mathrm{st}}/\overline{v}^4$ ergibt. Die mittlere Geschwindigkeit $\overline{v} = (3k_{\mathrm{B}}T/m^*)^{1/2}$ läßt sich mit Hilfe der kinetischen Gastheorie ausdrücken, da die Ladungsträger in nicht-entarteten Halbleitern der Boltzmann-Statistik gehorchen. Nehmen wir zur weiteren Vereinfachung an, daß die Zahl der geladenen Störstellen temperaturunabhängig ist, so erhalten wir

$$\frac{1}{\tau} = \frac{\overline{v}}{l} \propto \frac{\overline{v}\, n_{\mathrm{st}}}{\overline{v}^4} \propto \frac{n_{\mathrm{st}}}{T^{3/2}}\,, \tag{9.32}$$

$$\mu = \frac{e\tau}{m^*} \propto \frac{T^{3/2}}{n_{\mathrm{st}}}\,. \tag{9.33}$$

Die tatsächlich vorhandene Zahl der Streuzentren hängt von den Details der Kompensation der Probe ab, die bei der in Abb. 9.13 wiedergegebenen Messung nicht bekannt sind. Die Annahme einer konstanten Zahl ist deswegen fragwürdig, und die beobachtete Temperaturabhängigkeit der Beweglichkeit weicht von der Vorhersage (9.33) auch ein wenig ab. Wir wollen aber die Daten benutzen, um den Streuquerschnitt der Aluminiumionen abzuschätzen. Nehmen wir die Beweglichkeit der Probe (2) bei 30 K, so finden wir den Wert $\mu \approx 500\,\mathrm{cm}^2/\mathrm{Vs}$. Damit ergibt sich mit $\sigma_{\mathrm{st}} = e/(n_{\mathrm{st}}m^*\mu\overline{v})$ der Wert $\sigma_{\mathrm{st}} \approx 10^{-11}\,\mathrm{cm}^2$. Die Aluminiumionen besitzen somit einen Streudurchmesser von etwa 300 Å, bewirken also eine außerordentlich starke Streuung[2].

Oberhalb des Maximums überwiegt die Streuung an Phononen. Um die Temperaturabhängigkeit der Beweglichkeit in diesem Bereich herzuleiten, benutzen wir wieder den Zusammenhang zwischen freier Weglänge und Streuquerschnitt. Bei Raumtemperatur dominieren die Phononen mit der Debye-Frequenz ω_{D}, so daß wir von einem temperaturunabhängigen Streuquerschnitt σ_{st} ausgehen können. Nach Gleichung (5.67) steigt die Phononendichte n_{ph} in diesem Temperaturbereich proportional zur Temperatur an. Für die Stoßrate erhalten wir daher

$$\frac{1}{\tau} = \frac{\overline{v}}{l} = \overline{v}\, n_{\mathrm{ph}}\sigma_{\mathrm{st}} \propto \overline{v}\, n_{\mathrm{ph}} \propto T^{3/2} \qquad \text{oder} \qquad \mu \propto T^{-3/2}\,. \tag{9.34}$$

Ein Blick auf Abb. 9.13 macht die gute Übereinstimmung mit den experimentellen Resultaten deutlich. Der besondere Charakter des Ladungstransports in den stark dotierten Proben (3) kommt in der geringen Temperaturabhängigkeit und dem hohen Wert der Beweglichkeit bei tiefen Temperaturen zum Ausdruck. In erster Näherung verhalten sich derartig stark dotierte Halbleiter wie Metalle.

Am Ende dieses Abschnitts wollen wir nochmals auf die Leitfähigkeit zurückkommen, deren Temperaturverlauf in Abb. 9.14 für n-dotiertes Germanium zu sehen ist. Von tiefen Temperaturen kommend steigt die Leitfähigkeit entsprechend der Ladungsträgerkonzentration exponentiell an. Die Temperaturabhängigkeit der Beweglichkeit und der effektiven Zustandsdichte der Bänder spielt eine untergeordnete Rolle, da sie relativ schwach

[2] In Metallen streuen geladene Punktdefekte wesentlich weniger effektiv, da ihr Coulomb-Potential durch die hohe Konzentration freier Elektronen sehr stark abgeschirmt wird.

ist. Im Erschöpfungsbereich, in dem die Ladungsträgerkonzentration näherungsweise konstant ist, fällt die Leitfähigkeit mit zunehmender Temperatur aufgrund der Temperaturabhängigkeit der Beweglichkeit wieder ab. Bei hohen Temperaturen setzt Eigenleitung ein.

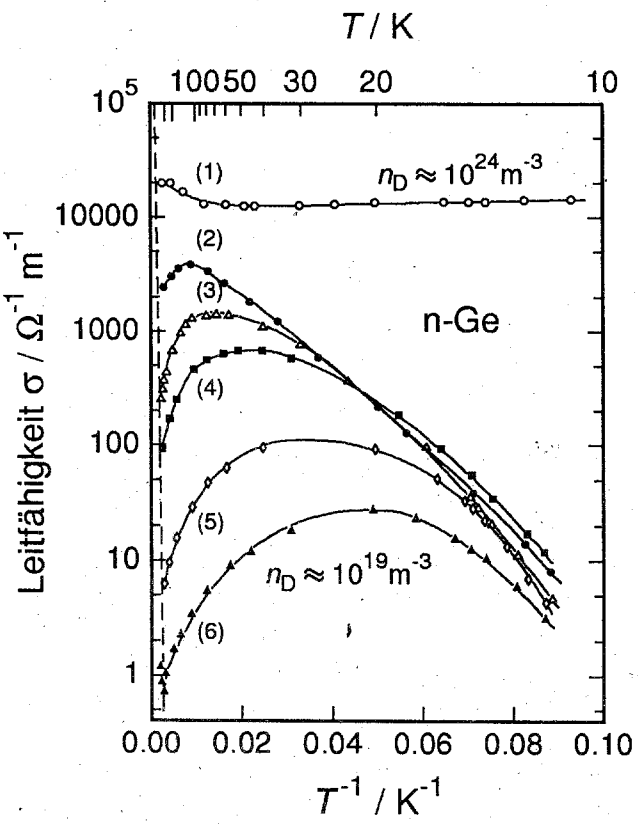


Abb. 9.14: Elektrische Leitfähigkeit von n-Germanium. Die Proben sind identisch mit denen, die bei der Bestimmung der Ladungsträgerdichte in Abb. 9.11 benutzt wurden. Die Leitfähigkeit der Probe (1) beruht auf der Ausbildung eines Störstellenbandes. (Nach E.M. Conwell, Proc. I.R.E. **40** (1952) 1327).

Einen Sonderfall stellt die Probe (1) mit der hohen Dotierung von $n_D = 10^{24}\,m^{-3}$ dar, bei der sich die Leitfähigkeit im Bereich der Störstellenleitung nur geringfügig ändert. Wie wir bereits in Abb. 9.11 gesehen haben, ist bei dieser Probe auch die Konzentration der beweglichen Ladungsträger weitgehend temperaturunabhängig. Die Ursache hierfür liegt in der erwähnten Ausbildung eines Störstellenbandes, in dem sich die Donatorelektronen ähnlich wie in Metallen im ganzen Temperaturbereich relativ frei bewegen können. Der Leitfähigkeitswert selbst ist kleiner als der von Legierungen oder amorphen Metallen.

9.3 Amorphe Halbleiter

In diesem Abschnitt gehen wir auf die amorphen Halbleiter ein, deren elektrische Eigenschaften ein aktuelles Forschungsgebiet der Festkörperphysik sind. Obwohl die Transporteigenschaften noch nicht vollständig verstanden sind, lohnt es sich, auf die Mechanismen des Ladungstransports näher einzugehen.

Wichtig ist zunächst, daß auch in amorphen Halbleitern elektronische Bänder auftreten, denn deren Existenz ist nicht an eine periodische Anordnung der Atome gebunden. Darauf haben wir bereits bei der Diskussion der Bänder im Modell der stark gebundenen Elektronen in Abschnitt 7.3 hingewiesen. Eine ähnliche Folgerung läßt sich auch aus der Existenz von Störstellenbändern bei stark dotierten kristallinen Halbleitern ziehen. Diese Bänder bilden sich aus, obwohl Störstellen ungeordnet im Kristall eingebaut werden.

Wichtig ist offensichtlich nur, daß sich die Wellenfunktionen der Störstellen genügend stark überlappen. Der ursprünglich für kristalline Festkörper hergeleitete Begriff des Bands kann daher weiter verwendet werden.

Die ungeordnete Struktur bewirkt eine Veränderung der elektronischen Eigenschaften der amorphen Halbleiter im Vergleich zu denen der kristallinen Halbleiter in mehrfacher Hinsicht: Sie modifiziert den Bandabstand, die Eigenschaften der Bandkanten, das Verhalten der Elektronen im Band, und sie ruft Defektzustände in der Bandlücke hervor. In Abb. 9.15 ist der Versuch gemacht, einige Besonderheiten der Bandstruktur amorpher Halbleiter zu veranschaulichen.

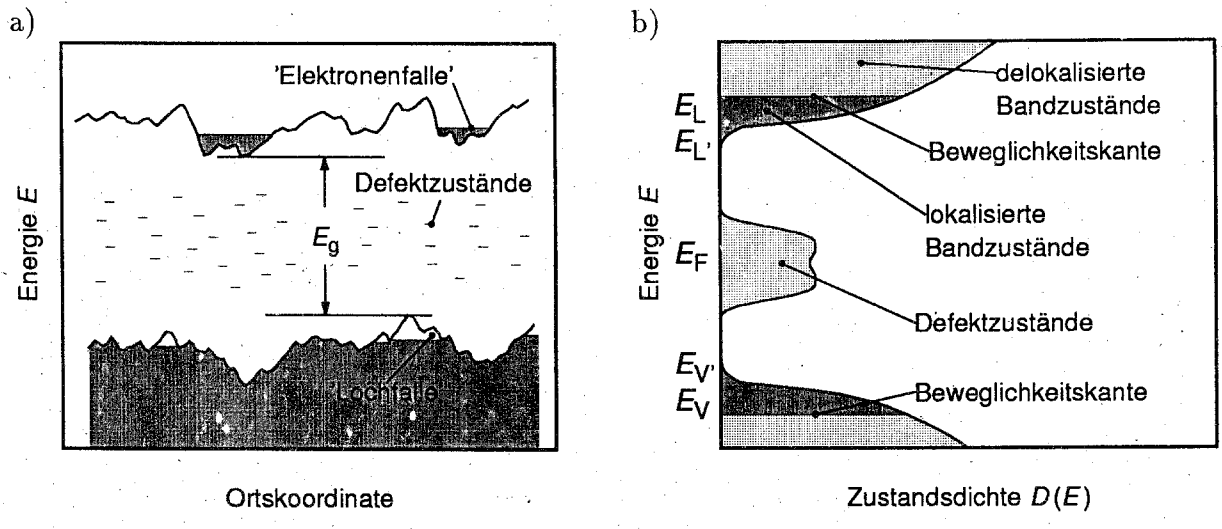


Abb. 9.15: Schematische Darstellung der elektronischen Zustände in amorphen Halbleitern.
a) „Ortsdarstellung" der auftretenden Zustände. Zustände, die mit Elektronen besetzt sind, sind grau dargestellt.
b) Zustandsdichte amorpher Halbleiter. Lokalisierte Bandzustände sind durch dunklere Färbung hervorgehoben.

Lokale Fluktuationen in den Atomabständen und Potentialen rufen eine räumliche Fluktuation der Bandkanten hervor. Dadurch entstehen Bereiche, in denen Elektronen oder Löcher gefangen werden und nur unter Energiezufuhr entweichen können: Sie sind lokalisiert, obwohl sich die Zustände im Band befinden. Um bewegliche und unbewegliche Ladungsträger in den Bändern zu unterscheiden, führte *Mott* den Begriff der *Beweglichkeitskante* ein, die in Abb. 9.15b eingezeichnet ist.

Die zweite Eigenart der elektronischen Zustandsdichte amorpher Halbleiter besteht darin, daß bereits in *undotierten* Substanzen Defektzustände auftreten. Sie werden von den intrinsischen Defekten verursacht, auf die wir bereits in Abschnitt 4.3 eingegangen sind. Grob gesprochen liegen diese Zustände, wie in Abb. 9.15 angedeutet, in der Mitte der Energielücke, da es sich weder um bindende noch um antibindende Zustände handelt. Sie besitzen aber einen nachhaltigen Einfluß auf die elektrische Leitfähigkeit und auf die Lage des Fermi-Niveaus.

Hier soll noch eine kurze Bemerkung zur optischen Absorption angefügt werden, die besonders empfindlich auf den Bandabstand reagiert. Aufgrund der nicht-periodischen Anordnung der Atome ist die $\boldsymbol{k}$-Auswahlregel bei amorphen Halbleitern aufgehoben. Die Übergänge vom Valenz- ins Leitungsband müssen daher bei der Absorption von Photonen auch dann nicht-vertikal erfolgen, wenn keine Phononen beteiligt sind. Ein Blick auf

Abb. 9.15a macht deutlich, daß in amorphen Halbleitern optische Übergänge bei kleineren Energien möglich sind als bei kristallinen Halbleitern der gleichen Zusammensetzung. Eine Bestätigung hierfür liefert Abb. 9.16. Dort ist der Imaginärteil ϵ'' der dielektrischen Funktion[3] von kristallinem und amorphem Silizium aufgetragen. Im amorphen Silizium ist die Absorptionskurve verschmiert und deutlich zu kleineren Photonenenergien hin verschoben. Dies bewirkt, daß dünne Schichten aus amorphem Silizium wesentlich weniger transparent sind als kristalline. Dieser Effekt wird beim Bau von Solarzellen (vgl. Abschnitt 9.4) ausgenutzt, denn bei Verwendung von amorphem anstelle von kristallinem Silizium wird der Hauptteil des Sonnenlichts bereits von wesentlich dünneren Schichten absorbiert. Daraus ergibt sich die Möglichkeit, beträchtliche Mengen des relativ teureren kristallinen Siliziums einzusparen.

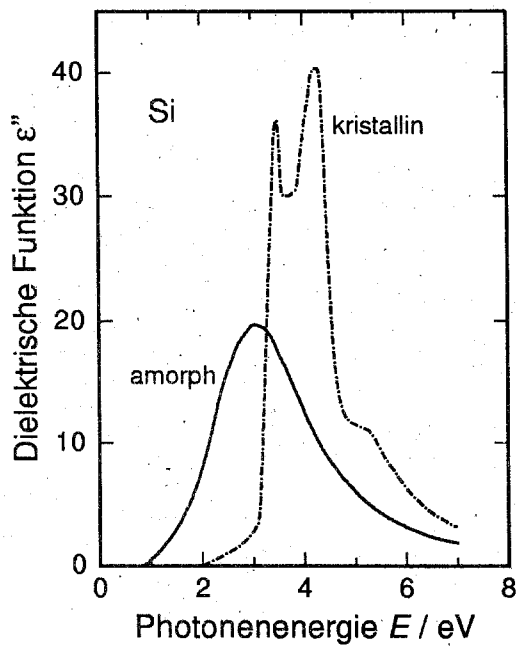


Abb. 9.16: Imaginärteil ϵ'' der dielektrischen Funktion von amorphem und kristallinem Silizium als Funktion der Energie der eingestrahlten Photonen. (Nach J. Stuke in Proc. 10th Int. Conf. Physics of Semiconductors, S.P. Keller, J.C. Hensel, F. Stern (eds.), US Atomic Energy Comm. Washington (1970) 14).

Elektrische Leitfähigkeit

Im folgenden betrachten wir die Eigenleitfähigkeit amorpher Halbleiter, bei denen Beweglichkeitskante und Defektzustände in der Bandlücke zu neuen Leitfähigkeitsphänomenen führen. Bei der Diskussion gehen wir davon aus, daß das Fermi-Niveau unabhängig von der Temperatur in der Mitte der Energielücke festliegt. Die Gründe für dieses Verhalten werden wir am Ende dieses Abschnitts kennenlernen. Zunächst ist zu berücksichtigen, daß Elektronen in amorphen Halbleitern nicht durch Bloch-Wellen beschrieben werden können, da diese ein periodisches Gitter voraussetzen. Elektronen werden daher auch in idealen defektfreien amorphen Netzwerken stark gestreut. Dies hat zur Folge, daß die Beweglichkeit der Ladungsträger im Vergleich zu kristallinen Materialien stark reduziert ist. Bei hoher Temperatur weisen deshalb amorphe Halbleiter eine vergleichsweise geringe Eigenleitfähigkeit auf. Andererseits beruht auf den intrinsischen Defekten in der

[3] ϵ'' ist über die Beziehung $\kappa = \epsilon''/2\tilde{n}$ mit dem optischen Extinktionskoeffizienten κ verknüpft, wobei $\tilde{n}$ für den Brechungsindex steht. Wie wir in Abschnitt 11.2 erläutern werden, spiegelt die Größe ϵ'' grob die Zustandsdichte der Bänder wider.

Bandlücke ein zusätzlicher Transportmechanismus, der zu einer vergleichsweise hohen elektrischen Leitfähigkeit bei tiefen Temperaturen führt. Dieser Ladungstransport ähnelt stark der Leitung im Störstellenband eines stark dotierten kristallinen Halbleiters. Da in den meisten Fällen die Elektronen den Hauptbeitrag zur Leitfähigkeit leisten, werden wir bei der Diskussion der Leitungsmechanismen nur Elektronen berücksichtigen.

Wie wir sehen werden, lassen sich vier Mechanismen des Ladungstransports unterscheiden, die in unterschiedlichen Temperaturbereichen vorherrschen. Die angegebenen Temperaturen und Beweglichkeitswerte sind typisch für gebräuchliche amorphe Halbleiter wie amorphes Silizium.

Bei *hohen Temperaturen* erfolgt der Ladungstransport in den delokalisierten Bandzuständen, die oberhalb der Beweglichkeitskante E_L anzutreffen sind. Dieser Transportmechanismus unterscheidet sich nicht prinzipiell von der Eigenleitung der kristallinen Halbleiter. Man erwartet, wie bei intrinsischen kristallinen Halbleitern, einen exponentiellen Zusammenhang zwischen Leitfähigkeit und Temperatur:

$$\sigma = \sigma_1 \mathrm{e}^{-(E_\mathrm{L}-E_\mathrm{F})/k_\mathrm{B}T} \,. \tag{9.35}$$

Der Faktor $\exp[-(E_\mathrm{L} - E_\mathrm{F})/k_\mathrm{B}T]$ spiegelt wie bei der Eigenleitung kristalliner Halbleiter die Temperaturabhängigkeit der Ladungsträgerkonzentration wider. Es ist jedoch zu beachten, daß hier mit E_L die Energie der Beweglichkeitskante gemeint ist. Die Temperaturabhängigkeit der Beweglichkeit ist vergleichsweise gering und wurde nicht explizit in die Gleichung aufgenommen. Aus dem experimentellen Wert der Konstanten σ_1 läßt sich bei Kenntnis der Zustandsdichte bei E_L die Beweglichkeit ermitteln. Typische Werte liegen um $10\,\mathrm{cm}^2/\mathrm{Vs}$ und sind somit etwa einen Faktor 1000 kleiner als bei Kristallen.

Bei *Raumtemperatur* reicht die thermische Energie nur noch aus, die Elektronen in die lokalisierten Bandzustände zu heben. Die Beweglichkeit der Elektronen in diesen Zuständen ist drastisch reduziert im Vergleich zu denen in Zuständen weiter oben im Leitungsband. Der Ladungstransport basiert auf Sprüngen der Elektronen von einem lokalisierten Zustand zu einem benachbarten. Man bezeichnet diesen Transportmechanismus als *Hopping* - oder *Hüpfleitfähigkeit*. Bei diesem Vorgang nehmen die Elektronen thermische Energie auf und tunneln anschließend zu benachbarten lokalisierten Zuständen. Dieser Vorgang soll anhand von Abb. 9.17 etwas näher erläutert werden.

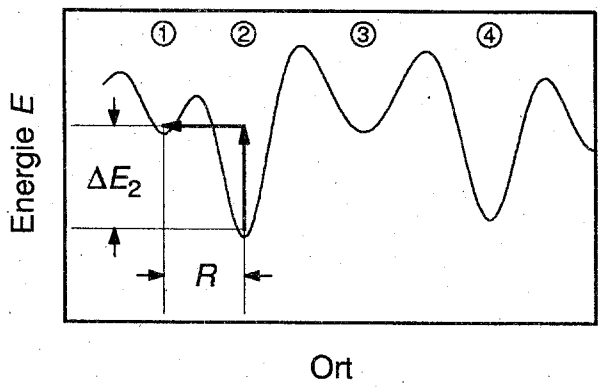


Abb. 9.17: Schematische Darstellung des Hüpfprozesses. Das Elektron in der Mulde 2 wird thermisch angehoben und tunnelt zum nächstgelegenen Potentialminimum. R ist der Abstand des nächstgelegenen Zustands und ΔE_2 die Aktivierungsenergie für den eingezeichneten Sprung.

Der Sprung setzt sich aus zwei Teilvorgängen zusammen. Im ersten Schritt nimmt das Elektron mindestens so viel thermische Energie aus dem Phononenbad auf, wie es zur Überwindung des Energieunterschieds zwischen den beiden Zuständen braucht. Die Wahrscheinlichkeit für das Eintreten diese Vorgangs läßt sich durch den Boltzmann-Faktor

ausdrücken. Anschließend erfolgt die Tunnelbewegung, bei der der Überlapp der Wellenfunktionen in den beiden Mulden die Wahrscheinlichkeit bestimmt, mit der dieser Prozeß auftritt. Die Sprungrate ν läßt sich daher näherungsweise durch

$$\nu = \nu_0 \, \mathrm{e}^{-\Delta E_2/k_\mathrm{B}T} \, \mathrm{e}^{-2\alpha R} \tag{9.36}$$

wiedergeben. Die beiden ersten Faktoren auf der rechten Seite kennen wir bereits von der Behandlung der thermisch aktivierten Diffusion in Abschnitt 4.1. Hier stellt jedoch das Anheben des Elektrons nur den ersten Schritt des Hüpfprozesses dar. Die „Versuchsfrequenz" ν_0 mit der das Elektron gegen die Potentialbarriere anläuft, ist mit $10^{13}\,\mathrm{s}^{-1}$ vergleichbar mit der bei der atomaren Diffusion. Tatsächlich handelt es sich um eine typische Gitterfrequenz, denn die Bewegung des Elektrons in der Potentialmulde wird durch die Schwingung der benachbarten Atome verursacht. Der Faktor $\exp(-2\alpha R)$ beschreibt den Überlapp der Wellenfunktion mit dem benachbarten Zustand im Abstand R und drückt somit die Tunnelwahrscheinlichkeit aus. Da die tatsächliche Wellenfunktion der lokalisierten Elektronen nicht bekannt ist, haben wir in diesem Ausdruck angenommen, daß sie näherungweise durch $\psi \propto \exp(-\alpha r)$ dargestellt werden kann, wobei sich α^{-1} als „Lokalisierungslänge" der Elektronen interpretieren läßt[4]. Da die Tiefe der Potentialmulden von Fall zu Fall schwankt, tritt anstelle der individuellen Aktivierungsenergie ΔE_2 ihr Mittelwert, den wir mit ΔW_2 bezeichnen. Die für die Bewegung der Elektronen erforderliche Aktivierungsenergie ist daher $(E_{\mathrm{L}'} + \Delta W_2)$, so daß der Exponent in Gleichung (9.35) durch $\exp\left[-(E_{\mathrm{L}'} + \Delta W_2 - E_\mathrm{F})/k_\mathrm{B}T\right]$ ersetzt werden muß. Da die Tunnelwahrscheinlichkeit nicht von der Temperatur abhängt, kann sie in den Vorfaktor des Ausdrucks für die Leitfähigkeit gezogen werden. Unter diesen Voraussetzungen ergibt sich für die Hüpfleitfähigkeit der Zusammenhang

$$\sigma = \sigma_2 \, \mathrm{e}^{-(\Delta W_2 + E_{\mathrm{L}'} - E_\mathrm{F})/k_\mathrm{B}T} \, . \tag{9.37}$$

Drückt man die Leitfähigkeit mit Hilfe der Beweglichkeit aus, so findet man ungefähr $0{,}01\,\mathrm{cm}^2/\mathrm{Vs}$ und damit einen wesentlich kleineren Wert als beim Ladungstransport in delokalisierten Zuständen. Es ist bemerkenswert, daß bei den amorphen Halbleitern nicht nur die Ladungsträgerkonzentration stark mit der Temperatur abnimmt, sondern auch die Beweglichkeit, da sich die Transportmechanismen in den verschiedenen Temperaturbereichen unterscheiden.

Unterhalb der Raumtemperatur stirbt die Hüpfbewegung im Band aus, und die Defektzustände in der Bandlücke gewinnen für den Ladungstransport an Bedeutung, da sie in der Nähe des Fermi-Niveaus liegen und somit thermisch besetzt sind. Die Elektronen hüpfen nun ähnlich wie beim soeben diskutierten Leitungsmechanismus von Defekt zu Defekt. Der formale Ausdruck für die Leitfähigkeit sieht wie bei der Raumtemperaturleitfähigkeit aus, nur die Konstanten unterscheiden sich. Die mittlere Aktvierungsenergie ΔW_3 ist durch den energetischen Abstand der Defektzustände vom Fermi-Niveau gegeben, entspricht also, wie wir Abb. 9.15b entnehmen können, in etwa der halben Defektbandbreite. Für die Leitfähigkeit können wir

$$\sigma = \sigma_3 \mathrm{e}^{-\Delta W_3/k_\mathrm{B}T} \tag{9.38}$$

[4] Formal ist der Ausdruck (9.36) identisch mit (4.5), denn $\nu_0 \exp(-2\alpha R)$ kann zu einer temperaturabhängigen Konstanten zusammengefaßt werden.

schreiben. Da die Defekte räumlich relativ weit voneinander entfernt sind, müssen beim Tunneln größere Abstände überwunden werden als beim Ladungstransport an der Leitungsbandkante. Die Tunnelwahrscheinlichkeit ist wesentlich geringer und somit auch die Leitfähigkeit, d.h. $\sigma_3 \ll \sigma_2$, obwohl der Exponentialfaktor größer als in (9.37) ist.

Bei tiefen Temperaturen, wenn $k_B T < \Delta W_3$ ist, ist der Hüpfprozeß etwas modifiziert. Es tritt sogenanntes *variable range hopping* zwischen den Defektzuständen auf. Die Energiedifferenz zwischen räumlich benachbarten Zuständen ist wesentlich größer als die thermische Energie und kann daher nicht mehr aufgebracht werden. Für die Elektronen ist es dann einfacher, über größere Abstände zu Zuständen zu tunneln, deren Energie sich vom Ausgangsdefekt nicht so stark wie die der unmittelbaren Nachbarn unterscheidet. Zur Veranschaulichung können wir nochmals die Skizze 9.17 heranziehen, wenn wir das gezeichnete „Potentialgebirge" nun den Defektzuständen zuschreiben[5]. Im gezeichneten Fall würde das Tieftemperaturhüpfen wahrscheinlich von Mulde (2) zur Mulde (4) und nicht zur Mulde (1) erfolgen. Für diesen Mechanismus ergibt die hier nicht durchgeführte Rechnung:

$$\sigma = \sigma_4 \exp\left[-\left(\frac{T_0}{T}\right)^{1/4}\right] . \tag{9.39}$$

T_0 ist eine Konstante, in der alle wichtigen Parameter wie Zustandsdichte, Lokalisierungslänge usw. zusammengefaßt sind. Die Potenz im Exponenten hängt von der Form der Defektzustandsdichte ab. Nur wenn die Zustandsdichte am Fermi-Niveau näherungsweise konstant und die untersuchte Probe dreidimensional ist, ergibt sich der numerische Wert 1/4 für den Exponenten. In Abb. 9.18 ist der Widerstand $R \propto 1/\sigma$ zweier Schichten aus amorphem Silizium aufgetragen, die durch Kathodenzerstäubung hergestellt wurden. Beide Proben weisen den erwarteten Temperaturverlauf des Widerstands auf.

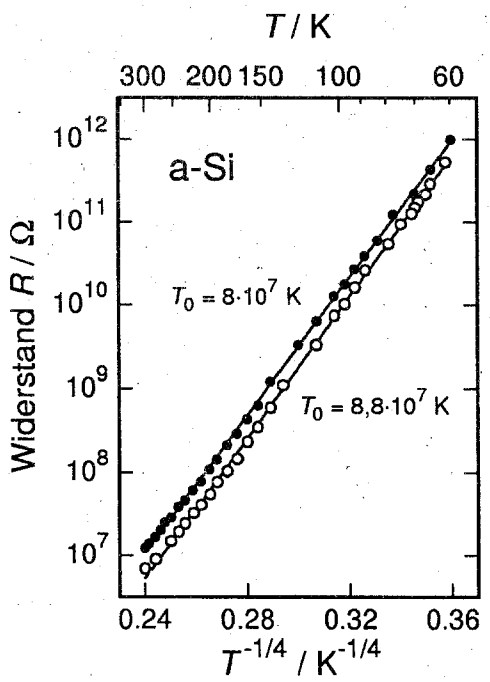


Abb. 9.18: Widerstand von zwei Schichten aus amorphem Silizium als Funktion von $T^{-1/4}$. (Nach J.J. Hauser, Phys. Rev. B **8** (1973) 3817).

[5] Hier soll nochmals betont werden, daß es sich in Wirklichkeit um unterschiedliche Potentialgebirge handelt. Bei den Defektzuständen sind die Potentialschwankungen kleiner, die Abstände zwischen den Mulden aber größer als im Leitungsband.

Damit haben wir qualitativ die elektrische Leitfähigkeit von amorphen Halbleitern diskutiert. Sieht man von hohen und tiefen Temperaturen ab, so erweist sich eine experimentelle Trennung der unterschiedlichen Transportmechanismen als außerordentlich schwierig, da der Übergang von einem zum anderen Transportmechanismus nicht abrupt, sondern gleitend geschieht. Als weitere Komplikation kommt hinzu, daß die Eigenschaften der Proben nicht nur von ihrer Zusammensetzung abhängen, sondern daß auch die Probenpräparation eine ganz wesentliche Rolle spielt.

Defektzustände

Die Diskussion der elektrischen Leitfähigkeit und der optischen Absorption amorpher Halbleiter machte deutlich, daß bereits in reinen Materialien, d.h. ohne Dotierung, Defekte vorhanden sind. Die einfachsten und wichtigsten Defekte sind nicht-abgesättigte chemische Bindungen. In kovalent gebundenen Kristallen entstehen sie, wenn man aus einer Bindung ein Elektron entfernt oder ein Atom entnimmt. Dadurch werden auch die Bindungseigenschaften der Nachbaratome in Mitleidenschaft gezogen, doch läßt sich immer eine vollständige Absättigung aller Bindungen durch Rekombination bzw. durch Einfügen des fehlenden Atoms erreichen. In amorphen Festkörpern werden derartige Defekte bereits während des Herstellungsprozesses isoliert eingebaut, ohne daß dadurch bei benachbarten Atomen ebenfalls eine Bindung offen bleibt. Eine nachträgliche Absättigung aller Bindungen ist daher aus topologischen Gründen nicht mehr möglich. Es würde eine vollkommen neue Anordnung des Netzwerks erfordern, d.h. es müßten viele Bindungen aufgebrochen und viele Atome umgelagert werden.

Wie wir in Abschnitt 4.3 bereits diskutiert haben, hängen Natur und Eigenschaften dieser Defekte von der Koordinationszahl des amorphen Halbleiters ab. Defekte in tetraedrischen Halbleitern wie a-Si verhalten sich völlig anders als in Chalkogenidgläsern wie a-As_2S_3 oder a-Se. So findet man in reinem amorphen Silizium etwa $10^{25} - 10^{26}\,\mathrm{m}^{-3}$ offene Bindungen, die sich mit Hilfe von Elektronenspinresonanzmessungen leicht als lokalisierte, ungepaarte Spins nachweisen lassen. In Chalkogenidgläsern liegt dagegen die Zahl der ungepaarten Elektronen unter der Nachweisgrenze dieser Meßmethode. Dieses völlig unterschiedliche Verhalten der beiden Klassen von amorphen Substanzen war lange Zeit unverstanden.

Mit den Donatoren und Akzeptoren kristalliner Halbleiter haben wir bereits spezielle Defekte mit ungesättigten Bindungen kennengelernt. In den meisten Fällen liegt ihre Energie in der Nähe der Bandkanten. Handelt es sich aber nicht um Fremdatome, sondern um aufgebrochene oder nicht abgesättigte Bindungen, so findet man die Zustände bei starrer Umgebung ungefähr in der Mitte der Energielücke, da weder ein bindendes noch antibindendes Orbital existiert. Durch „Relaxation", d.h. durch Umlagerung der benachbarten Atome, wird die Energie dieser Zustände aber etwas abgesenkt. Ungeladene Defekte besitzen die Eigenart, daß sie sowohl als Donatoren wirken, also das nicht-bindende Elektron abgeben, als auch ein zusätzliches Elektron aufnehmen, ein nicht-bindendes Elektronpaar bilden, und so als Akzeptoren fungieren können. Unter „normalen" Umständen, d.h. bei den üblichen Störstellen in kristallinen Halbleitern, wird im zweiten Fall die Energie des Zustands stark erhöht, da sich die beiden Elektronen des Defekts am gleichen Ort befinden und sich somit abstoßen. Die Energiedifferenz zwischen dem neutralen und dem negativ geladenen Zustand wird als *Korrelations-* oder *Hubbard-Energie* bezeichnet und ist in kristallinen Halbleitern typischerweise von der Größenordnung eines Elektronenvolts.

Bei der Diskussion der Störstellen in kristallinen Halbleitern haben wir daher angenommen, daß die doppelt besetzten Zustände im Leitungsband liegen und nicht weiter beachtet werden müssen. In amorphen Festkörpern sind lokale Strukturänderungen relativ leicht möglich. Das Netzwerk um einen geladenen Defekt ist zu starken Veränderungen fähig und senkt so die Energie des Defekts sehr weit ab. Defekte, die mit zwei Elektronen besetzt sind, spielen daher bei amorphen Halbleitern eine sehr wichtige Rolle.

Was die Korrelationsenergie U anbetrifft, existieren nun zwei Möglichkeiten. Entweder liegt der doppelt besetzte Zustand über dem neutralen bzw. positiv geladenen Zustand, oder aber er liegt darunter. Im ersten Fall ist die Korrelationsenergie wie in Kristallen positiv, wenn auch wesentlich kleiner. Im zweiten Fall ist sie negativ. In dem hier diskutierten einfachen Konzept treten, wie in der Prinzipskizze 9.19 angedeutet, innerhalb der Energielücke zwei Maxima in der Zustandsdichte auf. Die neutralen Defektzustände S^0 mit je einem ungepaarten Elektron liegen bei der Energie E_s, wobei s als Abkürzung für „Störstelle“ steht. Diese Zustände verursachen das starke Elektronspinresonanzsignal. Gibt ein derartiger Defekt sein ungepaartes Elektron ab, so entsteht der Zustand S^+, ohne daß sich die Energie wesentlich verschiebt. Das abgegebene Elektron kann von einem anderen Defekt aufgenommen werden, wobei der Defekt S^- gebildet wird. Während die Energie von S^+ weiter bei E_s liegt, befindet sich die Energie von S^- bei $E_s + |U|$ bzw. $E_s - |U|$, je nach dem ob die Korrelationsenergie positiv oder negativ ist. In der weiteren Diskussion müssen wir nun die Unterscheidung zwischen den beiden Fällen treffen, da sich daraus unterschiedliche Konsequenzen ergeben.

Betrachten wir zunächst amorphe Halbleiter wie a-Si oder a-Ge, bei denen die Korrelationsenergie positiv ist, die S^--Zustände daher über den neutralen oder positiv geladenen Störstellen liegen. Der Elektronentransfer, der zur Paarbildung führt, kostet Energie und wird normalerweise nicht eintreten, so daß man nur S^0-Defekte finden wird. Damit wird das beobachtete starke Elektronenspinresonanzsignal der tetraederförmig koordinierten Halbleiter verständlich. Versucht man diese amorphen Halbleiter zu dotieren, so verhindern die intrinsischen Defekte, die in großer Zahl vorhanden sind, eine merkliche Verschiebung des Fermi-Niveaus. In Abb. 9.19c ist das Verhalten derartiger Halbleiter gegenüber Dotierung in Abhängigkeit von der mittleren Elektronenzahl N_D pro Defekt dargestellt. Ausgangspunkt ist der intrinsische Zustand mit der Elektronenzahl $N_D = 1$, bei dem nur S^0-Defekte vorhanden sind. Bringt man nun Akzeptoren ein, so wird zwar die mittlere Elektronenzahl pro Defekt reduziert, doch verändert sich dadurch nicht die Lage des Fermi-Niveaus, da S^+- und S^0-Defekte ungefähr die gleiche Energie besitzen. Die S^0-Defekte wirken wie Donatoren, und es kommt zu einem Kompensationseffekt. Nun dotieren wir den amorphen Halbleiter mit Donatoren und vergrößern damit die Elektronenzahl in den Defektzuständen über den Wert $N_D = 1$. Nun werden doppelt besetzte Zustände bevölkert, d.h. es werden S^- Defekte gebildet. Die Fermi-Energie „springt“ zwar vom Wert E_s nach $E_s + U$, doch liegt dieser ebenfalls in der Nähe der Energielückenmitte. Die hohe Konzentration an intrinsischen Störstellen und ihre Lage in der Nähe der Energielückenmitte bewirkt, daß Dotieren ohne nachhaltigen Einfluß auf das Fermi-Niveau bleibt! Die Zahl der freien Ladungsträger läßt sich durch Dotieren nicht wesentlich ändern.

Es soll betont werden, daß wir hier bei der Darstellung der Zustandsdichte die Ein- und Zweielektronenzustände vermischt haben. Die oberen, mit zwei Elektronen besetzten Zustände können nur auftreten, wenn die unteren bereits voll sind. Sie werden also durch die Besetzung mit zwei Elektronen erst bei dieser Energie erzeugt. Ihre Dichte ist daher in Abb. 9.19 punktiert gezeichnet.

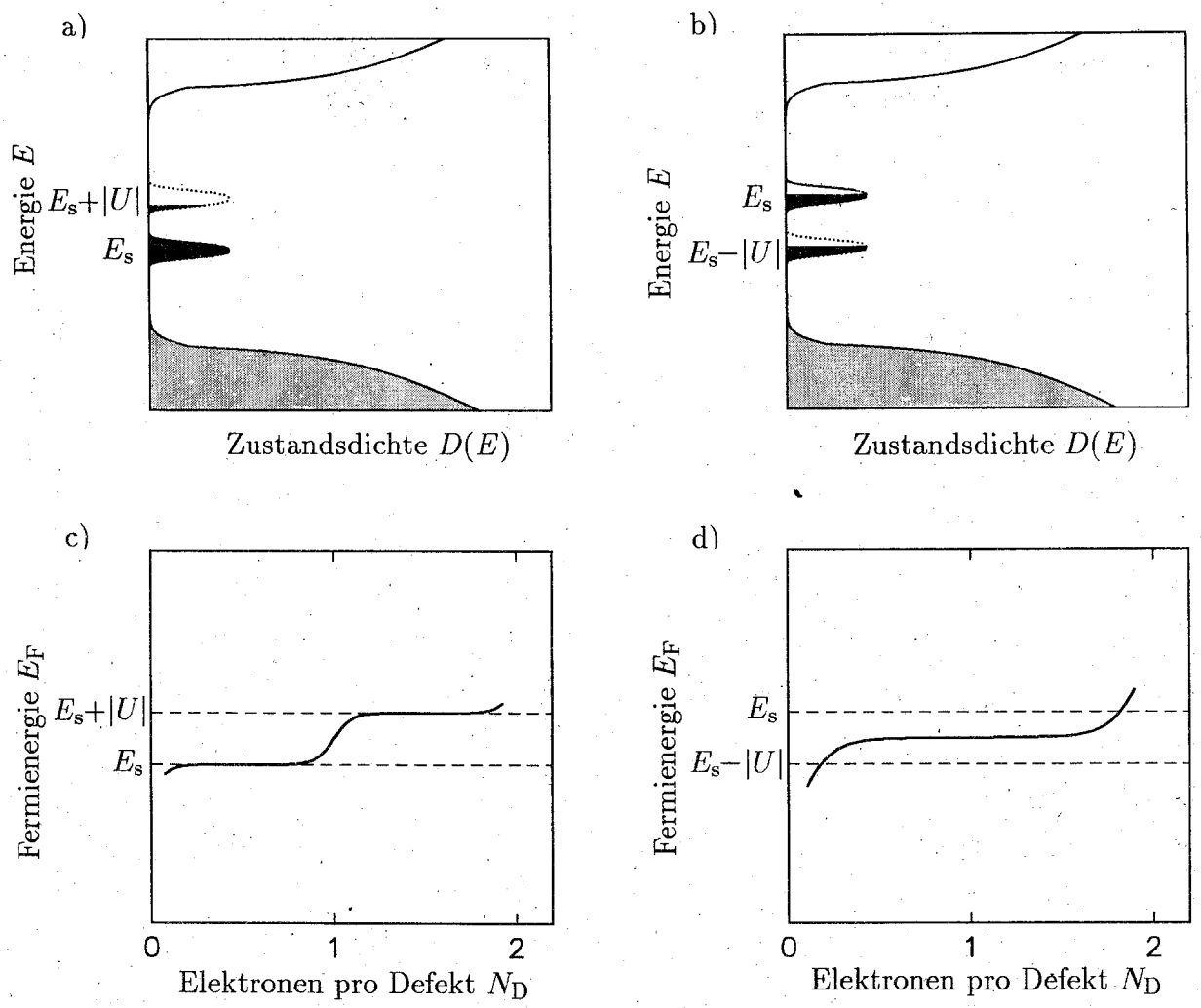


Abb. 9.19: Besetzung von Defektzuständen und ihr Einfluß auf die Lage des Fermi-Niveaus. Besetzte Defektzustände sind schwarz gezeichnet, die Zustandsdichte der doppelt besetzten Defekte ist gepunktet. Die gezeichnete Auffüllung der Zustände entspricht einer mittleren Elektronendichte $N_D = 1{,}2$ pro Atom.
a) Besetzung bei positiver Korrelationsenergie,
b) Besetzung bei negativer Korrelationsenergie,
c) Lage des Fermi-Niveaus bei positiver Korrelationsenergie,
d) Lage des Fermi-Niveaus bei negativer Korrelationsenergie.

Eine gezielte und wirkungsvolle Dotierung von amorphem Silizium kann nur durchgeführt werden, wenn die Defektzustände „unschädlich" gemacht werden. Dies erreicht man durch chemisches Absättigen der offenen Bindungen. Hierzu scheidet man im Falle des amorphen Siliziums den Film in wasserstoffhaltiger Atmosphäre ab, so daß sich Wasserstoff an die nicht abgesättigten Defekte anlagern kann. Da Silizium und Wasserstoff eine starke Bindung eingehen, liegt die Energie der nun auftretenden Silizium-Wasserstoffdefektzustände tief im Valenzband. Wie bei den kristallinen Halbleitern haben diese Defekte keinen Einfluß auf die elektrischen Eigenschaften. Über die Wasserstoffkonzentration, die von den Herstellungsbedingungen abhängt, läßt sich die Anzahl der ungesättigten Bindungen variieren. Bei einer Wasserstoffkonzentration im Bereich von 10% verschwindet das Signal der Elektronenspins, und es ändern sich die optischen

Eigenschaften. Derartig präpariertes amorphes Silizium kann zur Herstellung von elektronischen Bauelementen verwendet werden[6]. Derzeit wird a-Si vorzugsweise zur Herstellung billiger Solarzellen verwendet, da die optische Absorption, wie in Abb. 9.16 gezeigt, im Spektralbereich des Sonnenlichts wesentlich größer ist als die des kristallinen Siliziums.

In Abb. 4.25 war das Modell einer atomaren Konfiguration mit einer ungesättigten Bindung zu sehen, die auftritt, wenn der Wasserstoffgehalt nicht zur vollständigen Absättigung ausreicht. Den drei Wasserstoffatomen steht eine ungesättigte Bindung gegenüber. Hier muß noch ergänzt werden, daß die Defekte nicht alle die gleiche Energie besitzen, denn aufgrund der verschiedenen Defektumgebungen ist natürlich auch die strukturelle Relaxation unterschiedlich effektiv. Dies führt zu einer Verbreiterung der Zustandsdichten, die in Abb. 9.19 der Übersichtlichkeit halber zu schmal gezeichnet ist. Häufig überlagern sich die beiden Maxima wie in der schematischen Darstellung 9.17 und füllen einen wesentlichen Teil der Energielücke aus.

Bemerkenswert ist auch, daß in amorphen Halbleitern der Dotierprozeß weniger effektiv ist als in kristallinen. Dies liegt daran, daß die Dotieratome, die eine andere Wertigkeit als das Wirtsmaterial besitzen, beim Einbau in das amorphe Netzwerk nicht notwendigerweise die Koordination des Wirtmaterials annehmen. So kann Phosphor in amorphem Silizium durchaus als drei- oder fünfwertiges Atom eingebaut werden. Um dennoch eine merkliche Dotierung zu erreichen, sind in amorphen Halbleitern daher wesentlich höhere Konzentrationen des Dotiermaterials erforderlich als in kristallinen Festkörpern.

Bei den Chalkogenidgläsern ist die Korrelationsenergie negativ. Eine anschauliche Erklärung hierfür haben wir bereits in Kap. 4.3 bei der Diskussion der Defektkonfigurationen in amorphen Festkörpern kennengelernt. In den Chalkogenidgläsern bildet sich unter Energiegewinn aus zwei neutralen S^0-Defekten ein geladenes Defektpaare S^+ und S^-, bei dem keine ungepaarten Elektronen auftreten. Im intrinischen Zustand, also bei $N_D = 1$, sind deswegen keine S^0-Defekte vorhanden. Damit ist zu verstehen, warum in diesen Materialien keine freien lokalisierten Spins beobachtet werden können.

Die negative Korrelationsenergie hat eine interessante Konsequenz für die Dotierbarkeit des Halbleiters, auf die wir anhand von Abb. 9.19b und 9.19d näher eingehen wollen. Bringen wir ein Elektron in den Halbleiter ein, so besetzt es den Zustand bei E_s, das zweite Elektron geht jedoch in den Zustand mit der Energie $E_s - |U|$, da dieser tiefer liegt als ein weiterer Zustand bei E_s. Die Zustände werden also immer paarweise besetzt, so daß keine überschüssigen einfach besetzten Zustände auftreten. Das Fermi-Niveau liegt daher immer zwischen den Energien der beiden Zustände und ist somit festgehalten. Dabei treten praktisch keine unabgesättigten Spins auf, die in Elektronenspinresonanzmessungen nachgewiesen werden könnten. Ohne Dotierung, also beim intrinsischen Zustand mit $N_D = 1$, existieren bereits mit zwei Elektronen besetzte Defekte. Erst wenn alle S^--Zustände besetzt sind, kann das Fermi-Niveau verschoben werden.

Bisher haben wir so argumentiert, als wären in allen amorphen Halbleitern isolierte aufgebrochene Bindungen die wichtigste Defektsorte. Für das grundlegende Verständnis, auch bei negativer Korrelationsenergie, ist diese Betrachtungsweise nützlich. In realen Systemen wie den Chalkogenidgläsern sind die Defekte jedoch komplexer aufgebaut. Am relativ einfachen Beispiel des amorphen Halbleiters Selen haben wir das Zustandekommen der Defektzustände in Abschnitt 4.3 bereits angesprochen.

[6] Es gibt ein neueres Verfahren, die sogenannte „hot wire deposition", bei dem wesentlich weniger Wasserstoff zur Kompensation der intrinsischen Defekte benötigt wird.

9.4 Inhomogene Halbleiter

In diesem Abschnitt werden wir auf die physikalischen Grundlagen der technischen Anwendung von Halbleitern eingehen und einige einfache Bauelemente beschreiben. Von „inhomogenen" Halbleitern spricht man in diesem Zusammenhang, weil die Dotierung in diesen Halbleitern räumlich variiert. Ausgangspunkt unserer Betrachungen wird der p-n-Übergang sein. Daran schließt sich die Diskussion einiger Bauelemente an, die unmittelbar auf der Anwendung des p-n-Übergangs beruhen. Als „historisches" Beispiel für ein verstärkendes Bauelement werden wir den p-n-p-Transistor kennenlernen und kommen dann auf den MOS-Feldeffekttransistor zu sprechen. Zum Schluß gehen wir noch kurz auf Heterostrukturen, Übergitter und den Halbleiter-Injektionslaser ein.

p-n-Übergang ohne äußere Spannung

Für die technische Anwendung von Halbleitern in der Festkörperelektronik ist eine räumliche Variation der Störstellenkonzentration eine unumgängliche Voraussetzung. Diese erreicht man, indem man die Dotieratome in bestimmte Bereiche des Ausgangsmaterials eindiffundieren läßt oder implantiert. Zur lateralen Begrenzung der Dotierungsgebiete mit Dimensionen im Mikrometerbereich oder darunter wird heute vorwiegend die Photolithographie angewandt. Wir wollen aber hier auf diese interessante Technik nicht näher eingehen.

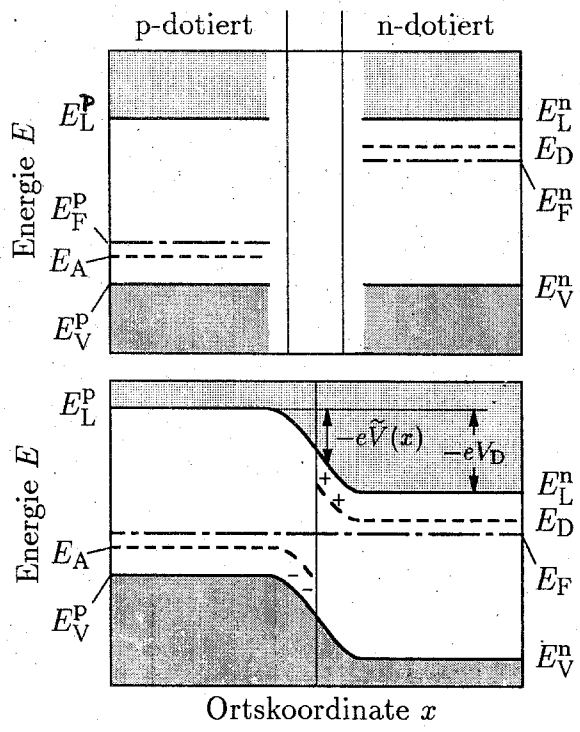


Abb. 9.20: Lage der Störstellenniveaus, der Bandkanten und des Fermi-Niveaus beim p-n-Übergang.
a) Lage der Energieniveaus in getrennten p- und n-dotierten Kristallen.
b) p-n-Übergang im Gleichgewicht. Die Bandkanten sind um die Diffusionsspannung V_D gegeneinander verschoben. Die auftretenden Raumladungen sind ebenfalls durch „+ +" bzw. „− −" angedeutet.

In den folgenden Betrachtungen werden wir „abrupte" Übergänge voraussetzen, bei denen die Änderung der Dotierungen sprunghaft erfolgt. In Abb. 9.20a ist die Ausgangssituation dargestellt. E_L^p und E_L^n bezeichnen die Lage der Leitungsbandkanten, E_V^p und E_V^n die Kanten der Valenzbänder im p- bzw. n-Leiter bei genügend großer

Entfernung vom Übergang. E_F^p und E_F^n stehen für die Energie der Fermi-Niveaus in den getrennten Kristallen. Das Fermi-Niveau liegt je nach Dotierung bei Raumtemperatur knapp unter dem Donator- bzw. über dem Akzeptorniveau. Werden nun p- und n-dotierter Halbleiter in Kontakt gebracht, so gleichen sich die bei unterschiedlicher Energie liegenden Fermi-Niveaus an. Dies geschieht durch Diffusion von Elektronen und Löchern aus den jeweiligen Gebieten hoher in die niedriger Konzentration. Wie in Abb. 9.20b dargestellt, baut sich dadurch eine Spannung auf, die die chemischen Potentiale so lange gegeneinander verschiebt, bis in der ganzen Probe ein einheitliches Potential herrscht. Der Vorgang erzwingt eine Bandverbiegung, die durch ein ortsabhängiges, eindimensionales makroskopisches Potential $\widetilde{V}(x)$ beschrieben werden kann, das der Diffusion entgegenwirkt.

Zunächt wollen wir die *Diffusionsspannung* V_D, d.h. die Potentialdifferenz zwischen den unterschiedlich dotierten Halbleitern berechnen, die eine der charakteristischen Größen eines p-n-Übergangs ist. Sie wird durch die ursprüngliche Lage der beiden Fermi-Niveaus und damit im wesentlichen durch die Bandlücke des Halbleiters bestimmt. Der Abbildung 9.20b läßt sich unmittelbar entnehmen, daß die Diffusionsspannung durch die Differenz $E_F^n - E_F^p$ der beiden Fermi-Niveaus der dotierten Kristalle gegeben ist. Mit Hilfe von (9.29) für die Lage des Fermi-Niveaus im Erschöpfungszustand erhält man

$$eV_D = E_L - k_B T \ln\left(\frac{N_L}{n_D}\right) - E_V - k_B T \ln\left(\frac{N_V}{n_A}\right) = E_g - k_B T \ln\left(\frac{N_L N_V}{n_D n_A}\right) , \quad (9.40)$$

so daß wir in erster Näherung $eV_D \simeq E_g$ setzen können.

Der Verlauf der Dichte der freien Ladungsträger ist in Abb. 9.21a schematisch im logarithmischen Maßstab dargestellt. Hierbei wurde angenommen, daß sich die Halbleiter im Erschöpfungsbereich befinden, in dem die Störstellen vollständig ionisiert sind.

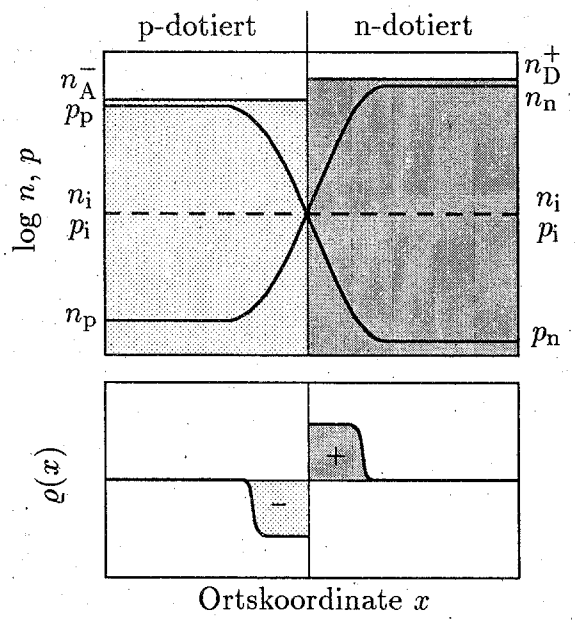


Abb. 9.21: Dichte der Störstellen, der freien Ladungsträger und der Raumladung beim p-n-Übergang.
a) Konzentration der ionisierten Störstellen und freien Ladungsträger im logarithmischen Maßstab. Die dunkel getönte Donatordichte überwiegt hier die Dichte der Akzeptoren (hellgrau getönt).
b) Raumladung. Die nahezu rechteckige Form wird durch die rasche Abnahme der Dichte der freien Ladungsträger der Raumladungszone bewirkt.

Ladungsträger, die sich in dem von der Dotierung vorgegebenen Gebiet aufhalten, d.h. Elektronen im n- und Löcher im p-Gebiet, werden als *Majoritätsladungsträger* bezeichnet. Diffundieren sie in das entgegengesetzt dotierte Gebiet, so heißen sie dort *Minoritätsladungsträger*. Wir versehen in der weiteren Diskussion die Konzentrationen der

Ladungsträger und der übrigen wichtigen Größen mit Indizes, die den Bereich angeben, in dem sie sich aufhalten. Für die Majoritätsladungsträger schreiben wir daher n_n bzw. p_p, für die Minoritätsladungsträger n_p bzw. p_n. In größerer Entfernung vom Übergang ist $n_\mathrm{n} \simeq n_\mathrm{D}^+ \simeq n_\mathrm{D}$ bzw. $p_\mathrm{p} \simeq n_\mathrm{A}^- \simeq n_\mathrm{A}$, wobei sich die Konzentration der Störstellen gewöhnlich unterscheidet. Das Massenwirkungsgesetz $n_\mathrm{p} \cdot p_\mathrm{p} = n_\mathrm{n} \cdot p_\mathrm{n} = n_\mathrm{i} \cdot p_\mathrm{i}$ bewirkt, daß bei den üblichen Konzentrationen der Dotierung die Majoritätsladungsträgerdichte sehr viel größer, die der Minoritätsladungsträger dagegen sehr viel kleiner als die Konzentration der intrinsischen Ladungsträger ist.

Die Diffusion der Majoritätsladungsträger verursacht eine starke Reduktion der Konzentration der freien Ladungsträger in der Umgebung des Übergangs. Folglich werden im Übergangsbereich die Ladungen der ionisierten Donatoren und Akzeptoren nicht mehr durch freie Ladungsträger kompensiert. Es entsteht eine *Raumladungszone*, wie sie in Abb. 9.21b schematisch dargestellt ist. Für die Raumladungsdichte können wir

$$\varrho\,(x > 0) = \quad e[n_\mathrm{D}^+ - n_\mathrm{n}(x) + p_\mathrm{n}(x)] \qquad \text{n-Gebiet}\,, \tag{9.41}$$

$$\varrho\,(x < 0) = -e[n_\mathrm{A}^- + n_\mathrm{p}(x) - p_\mathrm{p}(x)] \qquad \text{p-Gebiet} \tag{9.42}$$

schreiben, wobei $x = 0$ die Grenzen zwischen den beiden Bereichen kennzeichnet. Potentialverlauf $\widetilde{V}(x)$ und Raumladung $\varrho(x)$ sind über die Poisson-Gleichung

$$\frac{\partial^2 \widetilde{V}(x)}{\partial x^2} = -\frac{\varrho(x)}{\epsilon_\mathrm{r}\epsilon_0} \tag{9.43}$$

miteinander verbunden. Die Lösung dieser Gleichung stößt auf Schwierigkeiten, da die Raumladungsdichte vom Potentialverlauf selbst abhängt. Die Lösung müßte daher in einem selbstkonsistenten Verfahren erfolgen. Wir benutzen die Tatsache, daß in der Raumladungszone die Konzentration der freien Ladungsträger (vgl. Abb. 9.21a) sehr gering ist und dort in erster Näherung vernachlässigt werden kann. Der tatsächliche Verlauf von $\varrho(x)$ kann deshalb durch einen rechteckigen ersetzt werden. Diese Vereinfachung bezeichnet man als *Schottky-Modell* der Raumladungszone. Die Fläche der Rechtecke muß so gewählt werden, daß die Raumladungszone insgesamt neutral ist. Für die Raumladungdichte werden wir also folgenden Ausdruck benutzen:

$$\varrho(x) = \begin{cases} 0 & \text{für} \quad x < -d_\mathrm{p} \\ -en_\mathrm{A} & \text{für} \quad -d_\mathrm{p} < x < 0 \\ en_\mathrm{D} & \text{für} \quad 0 < x < d_\mathrm{n} \\ 0 & \text{für} \quad x > d_\mathrm{n}\,. \end{cases} \tag{9.44}$$

Hierbei sind d_n und d_p die Dicken der jeweiligen Raumladungszone, die wir nun berechnen wollen. Mit dieser vereinfachenden Annahme für die Raumladung kann die Poisson-Gleichung stückweise integriert werden. Für den n-leitenden Teil der Raumladungszone, d.h. für $0 < x < d_\mathrm{n}$, ergibt sich damit:

$$\frac{\partial^2 \widetilde{V}(x)}{\partial x^2} \simeq -\frac{en_\mathrm{D}}{\epsilon_\mathrm{r}\epsilon_0}\,, \tag{9.45}$$

$$\mathcal{E}_x = -\frac{\partial \widetilde{V}}{\partial x} = -\frac{e}{\epsilon_\mathrm{r}\epsilon_0} n_\mathrm{D}(d_\mathrm{n} - x)\,, \tag{9.46}$$

$$\widetilde{V}(x) = \widetilde{V}_\mathrm{n}(\infty) - \frac{en_\mathrm{D}}{2\epsilon_\mathrm{r}\epsilon_0}(d_\mathrm{n} - x)^2\,. \tag{9.47}$$

Entsprechende Gleichungen gelten für den p-leitenden Teil des p-n-Überganges. Mit Hilfe der Neutralitätsbedingung $n_D d_n = n_A d_p$ und der Bedingung für die Kontinuität des Potentials $\widetilde{V}(x)$ am Ort $x = 0$,

$$\frac{e}{2\epsilon_r\epsilon_0}(n_D d_n^2 + n_A d_p^2) = \widetilde{V}_n(\infty) - \widetilde{V}_p(-\infty) = V_D \ , \tag{9.48}$$

lassen sich die Dicken d_n und d_p der Raumladungen berechnen. Man findet

$$d_n = \sqrt{\frac{2\epsilon_r\epsilon_0 V_D}{e}\,\frac{n_A/n_D}{n_A + n_D}} \qquad \text{und} \qquad d_p = \sqrt{\frac{2\epsilon_r\epsilon_0 V_D}{e}\,\frac{n_D/n_A}{n_A + n_D}} \ . \tag{9.49}$$

Setzt man mit $eV_D \simeq E_g \approx 1\,\text{eV}$ und $n_A \approx n_D \simeq 10^{20} \ldots 10^{24}\,\text{m}^{-3}$ typische Zahlenwerte für Bandlücke und Störstellenkonzentration ein, so findet man für die Dicke der Raumladungszone $d_n \approx d_p \approx 1\,\mu\text{m} \ldots 10\,\text{nm}$ und damit elektrische Feldstärken im Bereich $\mathcal{E} \approx 10^6\,\text{V/m} \ldots 10^8\,\text{V/m}$.

In Abb. 9.22 ist der schematische Verlauf der wichtigen Größen $\varrho(x)$, $\mathcal{E}_x$ und $\widetilde{V}(x)$ am p-n-Übergang zusammenfassend dargestellt.

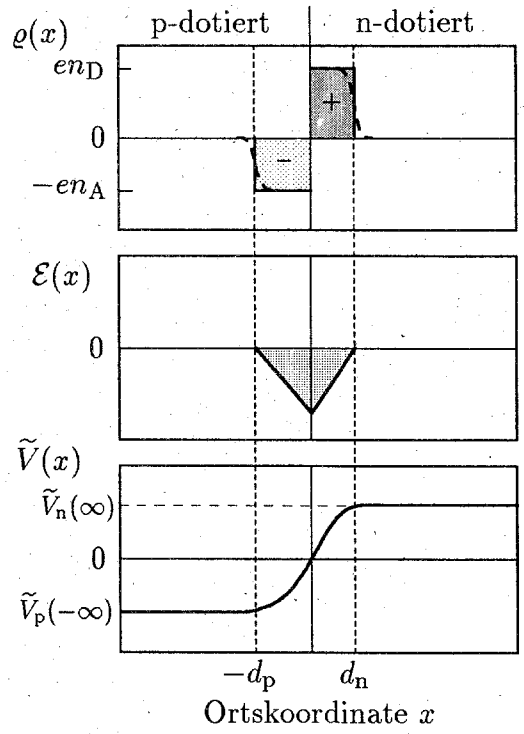


Abb. 9.22: Schottky-Modell der Raumladungszone.
a) Rechtecknäherung der Raumladungszonen. Der realistische Verlauf ist gestrichelt eingezeichnet.
b) Elektrische Feldstärke $\mathcal{E}(x)$,
c) Potentialverlauf $\widetilde{V}(x)$.

Um die Funktionsweise von Bauelementen verstehen zu können, müssen wir als erstes die Ströme im p-n-Übergang betrachten, die bereits ohne äußere Spannung fließen. An dieser Stelle wollen wir noch vorausschicken, daß in einem Halbleiter aufgrund des geringen Bandabstands an jeder Stelle durch die thermische Bewegung ständig Elektron-Loch-Paare erzeugt werden und durch Rekombination wieder verschwinden; darüber hinaus werden Störstellen ionisiert, die dann wieder rekombinieren.

Der *Diffusionsstrom* wird durch den Konzentrationsunterschied der Ladungsträger auf den beiden Seiten des p-n-Übergangs hervorgerufen. Dadurch baut sich, wie eben diskutiert, die Diffusionsspannung auf. Mit dieser Spannung ist das elektrische Feld (9.46) in

der Raumladungszone verbunden, das einen *Feldstrom* bewirkt, der dem Diffusionsstrom entgegengerichtet ist. Im thermischen Gleichgewicht kompensieren sich beide Ströme gerade. Bezeichnen wir aus Gründen, die wir sofort diskutieren werden, die Dichte des Feldstroms mit j^{E} und die des Diffusionsstroms mit j^{R}, so gilt im Gleichgewicht:

$$j^{\mathrm{E}} + j^{\mathrm{R}} = 0 \,. \tag{9.50}$$

Teilen wir Feld- und Diffusionsstromdichte mit Hilfe der Gleichungen $j^{\mathrm{E}} = j_{\mathrm{n}}^{\mathrm{E}} + j_{\mathrm{p}}^{\mathrm{E}}$ und $j^{\mathrm{R}} = j_{\mathrm{n}}^{\mathrm{R}} + j_{\mathrm{p}}^{\mathrm{R}}$ in die Teilströme auf, die von den Elektronen und Löchern herrühren, so gilt diese Argumentation auch separat für beide Sorten von Ladungsträgern. Elektronen- und Löcherstrom heben sich auf, da sich weder Elektronen noch Löcher in irgendeinem Gebiet akkumulieren dürfen. Daraus folgt:

$$j_{\mathrm{n}}^{\mathrm{E}} + j_{\mathrm{n}}^{\mathrm{R}} = 0 \qquad\qquad j_{\mathrm{p}}^{\mathrm{E}} + j_{\mathrm{p}}^{\mathrm{R}} = 0 \,. \tag{9.51}$$

Wir wollen nun die Teilströme näher betrachten. Die Minoritätsladungsträger werden durch das elektrische Feld am p-n-Übergang in das Gebiet der Majoritätsladungsträger gezogen. So laufen Elektronen aus dem p-Gebiet in das n-Gebiet und Löcher in umgekehrter Richtung. Nehmen wir zur Vereinfachung an, daß die Raumladungszone nur wenig ausgedehnt ist, so dürfen wir die Rekombination in der Raumladungszone vernachlässigen. Dies bedeutet, daß alle Ladungsträger, die in das Feld der Raumladungszone geraten, in den entgegengesetzt dotierten Bereich des Halbleiters gelangen. Im Rahmen dieser Vereinfachung hängt die Stärke dieses Stroms nicht vom Potentialverlauf ab. Da der Feldstrom von den Minoritätsladungsträgern verursacht wird, die im gesamten Halbleiter ständig neu erzeugt werden, wird der Feldstrom oft auch als *Erzeugungsstrom* j^{E} bezeichnet.

Die Majoritätsladungsträger, z.B. Elektronen aus dem n-Gebiet, die in das p-Gebiet diffundieren, müssen *gegen* die Potentialdifferenz an der Grenzschicht anlaufen. Entsprechendes gilt in entgegengesetzter Richtung für die Löcher. Die Höhe der Barriere ist durch die Größe eV_{D}, also durch die Diffusionsspannung festgelegt. Der Boltzmann-Faktor bestimmt den Bruchteil der „erfolgreichen" Ladungsträger, d.h. $j^{\mathrm{R}} = a(T)\exp(-eV_{\mathrm{D}}/k_{\mathrm{B}}T)$. Dieser Strom j^{R} wird meist als *Rekombinationsstrom* bezeichnet. Der Ausdruck weist darauf hin, daß die Ladungsträger nach Durchlaufen der als verlustfrei angenommenen Raumladungszone durch Rekombination verschwinden. Der Vorfaktor $a(T)$ hängt schwach von der Temperatur ab, doch kann diese Abhängigkeit im Vergleich zum exponentiellen Faktor vernachlässigt werden. Unter Berücksichtigung von Gleichung (9.50) finden wir für den Betrag der Stromdichten:

$$|j^{\mathrm{E}}| = |j^{\mathrm{R}}| = a(T)\,\mathrm{e}^{-eV_{\mathrm{D}}/k_{\mathrm{B}}T} \,. \tag{9.52}$$

p-n-Übergang unter äußerer Spannung, Gleichrichtung

Der p-n-Übergang wirkt als Diode und läßt sich zur Stromgleichrichtung nutzen, denn der resultierende Strom durch die Kontaktstelle hängt von der Richtung der äußeren Spannung ab. Wir wollen den p-n-Übergang als einfaches Beispiel für ein Halbleiterbauelement etwas näher betrachten. Durch das Anlegen einer äußeren elektrischen Spannung wird das geschilderte thermische Gleichgewicht der Ströme gestört, und die Ergebnisse der Gleichgewichtsthermodynamik sind nicht mehr anwendbar. Ist der Zustand des p-n-Übergangs jedoch stationär und vom thermischen Gleichgewicht nicht allzu weit entfernt,

so führt eine relativ einfache Betrachtungsweise zum Ziel. Die angelegte Spannung U fällt in erster Linie in der Raumladungszone ab, die an freien Ladungsträgern verarmt ist und somit einen großen Widerstand besitzt. Der restliche Halbleiter ist weitgehend feldfrei. Wir schreiben daher $\widetilde{V}_{\mathrm{n}}(\infty) - \widetilde{V}_{\mathrm{p}}(-\infty) = V_{\mathrm{D}} - U$, dabei haben wir das Vorzeichen der anliegenden Spannung so gewählt, daß eine positive Spannung der Diffusionsspannung entgegengerichtet ist und die Potentialdifferenz reduziert. In Hinblick auf die Gleichrichtereigenschaften von p-n-Übergängen spricht man beim Anliegen einer positiven Spannung von der „Durchlaßrichtung" und von der „Sperrichtung" im umgekehrten Fall.

In Abb. 9.23 sind die Auswirkungen einer äußeren Spannung auf den Verlauf der Bänder eines p-n-Übergangs dargestellt. Das Fermi-Niveau ist nicht eingezeichnet, da die Verhältnisse nun etwas komplizierter sind. Befindet sich das betrachtete System nicht im thermischen Gleichgewicht, und das ist der Fall, wenn ein resultierender Strom durch den p-n-Übergang fließt, so existiert kein gemeinsames Fermi-Niveau für Elektronen und Löcher. Sind die einzelnen Ladungsträgersorten untereinander im Gleichgewicht, so kann man ihnen zwei getrennte, sogenannte *Quasi-Fermi-Niveaus* zuschreiben und diese getrennt betrachten. Wir wollen hier aber diese Problematik nicht weiter vertiefen, da sie für unsere Diskussion bedeutungslos ist. Eine interessante Konsequenz unterschiedlicher Quasi-Fermi-Niveaus werden wir aber am Ende dieses Kapitels beim Injektionslaser kennenlernen.

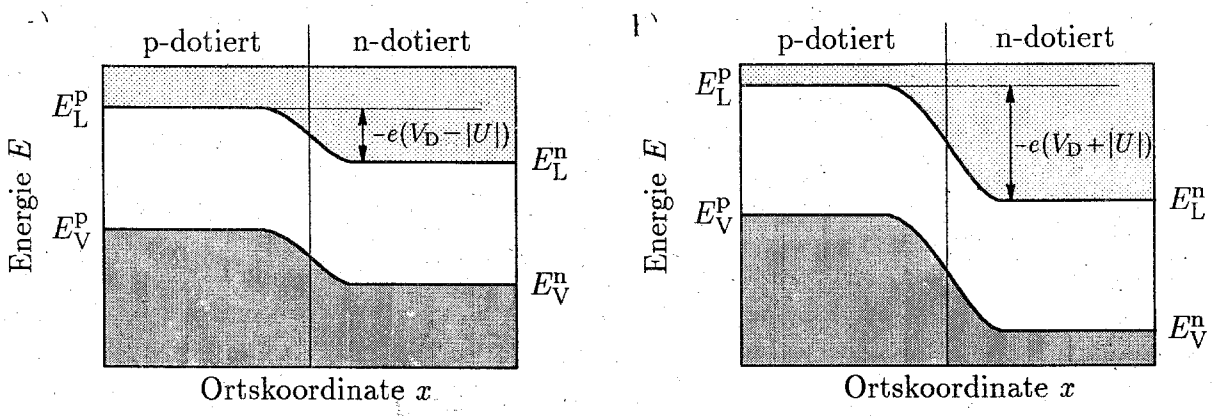


Abb. 9.23: p-n-Übergang beim Anlegen einer äußeren Spannung.
a) Spannung in Durchlaßrichtung reduziert die Potentialdifferenz,
b) Spannung in Sperrichtung vergrößert die Potentialdifferenz am Übergang.

Die Spannung verändert nicht nur den Bandverlauf, sondern wirkt auch direkt auf die Dicke der Raumladungszone ein. Benutzen wir wieder das Schottky-Modell, so können wir in Formel (9.49) die Diffusionsspannung V_{D} durch $V_{\mathrm{D}} - U$ ersetzen. Bezeichnen wir die Dicken ohne äußere Spannung mit $d_{\mathrm{n}}(0)$ und $d_{\mathrm{p}}(0)$, so erhält man

$$d_{\mathrm{n}}(U) = d_{\mathrm{n}}(0)\sqrt{1 - \frac{U}{V_{\mathrm{D}}}} \qquad \text{und} \qquad d_{\mathrm{p}}(U) = d_{\mathrm{p}}(0)\sqrt{1 - \frac{U}{V_{\mathrm{D}}}}\,. \tag{9.53}$$

Bei Kenntnis der Spannungsabhängigkeit der Dicke der Raumladungszonen läßt sich deren *Kapazität* leicht berechnen. Ist A die Querschnittsfläche des p-n-Überganges, dann ist die gespeicherte Ladung $Q = en_{\mathrm{D}}\, d_{\mathrm{n}}(U)\, A$. Mit (9.53) ist die Kapazität der Raumladung somit durch

$$C = \frac{\mathrm{d}Q}{\mathrm{d}U} = en_{\mathrm{D}}\, A \left[\frac{\mathrm{d}}{\mathrm{d}U} d_{\mathrm{n}}(U)\right] = \frac{A}{2}\sqrt{\frac{n_{\mathrm{A}} n_{\mathrm{D}}}{n_{\mathrm{A}} + n_{\mathrm{D}}} \frac{2e\epsilon_{\mathrm{r}}\epsilon_0}{(V_{\mathrm{D}} - U)}} \tag{9.54}$$

gegeben. Da die Kapazität der Raumladungszone von der Zahl der Störstellen abhängt, wird die Spannungsabhängigkeit der „Sperrschichtkapazität“ häufig zur experimentellen Bestimmung der Störstellenkonzentrationen herangezogen. Darüber hinaus dienen p-n-Übergänge in elektronischen Schaltungen häufig als einstellbare Kapazitäten.

Kehren wir zur zentralen Frage nach dem Einfluß einer anliegenden Spannung auf den Stromfluß in p-n-Übergängen zurück. Der Erzeugungsstrom wird durch die angelegte Spannung in erster Näherung nicht beeinflußt. Jeder Ladungsträger, ob Elektron oder Loch, wird unabhängig von der Stärke des Felds abgesaugt und durchquert die Raumladungszone. Wir können daher

$$j^{\mathrm{E}}(U) \approx j^{\mathrm{E}}(0) \tag{9.55}$$

schreiben. Für die Majoritätsträger ändert sich jedoch die Barrierenhöhe, denn sie laufen nun nicht mehr gegen die Spannung V_{D}, sondern gegen $V_{\mathrm{D}} - U$ an. Der ungleiche Einfluß der Spannung auf die beiden Teilströme hat zur Folge, daß sich Erzeugungsstrom und Rekombinationsstrom nicht mehr kompensieren. Wenn wir noch die Tatsache benutzen, daß $j^R(0) = j^E(0)$ ist, so erhalten wir

$$j^{\mathrm{R}}(U) = a(T)\,\mathrm{e}^{-e(V_{\mathrm{D}}-U)/k_{\mathrm{B}}T} = j^{\mathrm{R}}(0)\,\mathrm{e}^{eU/k_{\mathrm{B}}T} = j^{\mathrm{E}}(0)\,\mathrm{e}^{eU/k_{\mathrm{B}}T}\,. \tag{9.56}$$

Berücksichtigen wir weiterhin, daß die beiden Ströme in entgegengesetzte Richtung fließen, so läßt sich die resultierende Elektronenstromdichte durch

$$j = j^{\mathrm{R}} - j^{\mathrm{E}} = j^{\mathrm{E}}\,\mathrm{e}^{eU/k_{\mathrm{B}}T} - j^{\mathrm{E}} = j^{\mathrm{E}}\left(\mathrm{e}^{eU/k_{\mathrm{B}}T} - 1\right) \tag{9.57}$$

ausdrücken. Da bei hoher Sperrspannung der Erzeugungsstrom j^{E} mit dem Strom in Sperrichtung identisch ist, bezeichnen wir j^{E} als j_{s} und kommen so zum Endergebnis

$$j(U) = j_{\mathrm{s}}\left(\mathrm{e}^{eU/k_{\mathrm{B}}T} - 1\right)\,. \tag{9.58}$$

Dieses Resultat ist in Abb. 9.24 dargestellt. Das ausgeprägt nichtlineare Verhalten verleiht dem p-n-Übergang die Eigenschaften eines Gleichrichters. Liegt die Spannung in Durchlaßrichtung an, so kann der Strom „beliebig“ stark anwachsen. Bei Polung in Sperrrichtung wirkt die Grenzschicht als Sperre für den Strom. Bei größerer Spannung fließt dann nur noch der spannnungsunabhängige Erzeugungsstrom.

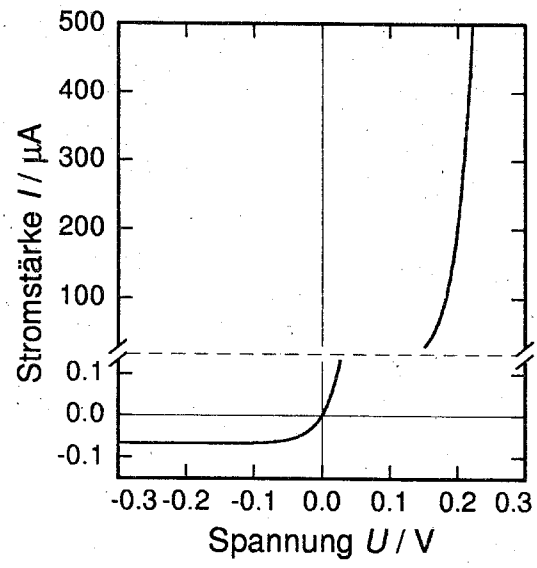


Abb. 9.24: Strom-Spannungskennlinie einer p-n-Diode. Man beachte die unterschiedlichen Skalen in Sperr- und Durchlaßrichtung.

Erhöht man die Sperrspannung an einem p-n-Übergang, so setzt plötzlich ein sehr hoher Stromfluß in Sperrichtung ein. Im Falle von Siliziumdioden liegt die kritische Spannung zwischen 1 V und 2 V. Der Grund hierfür ist der *Zener-Durchbruch*, der beobachtet wird, wenn die Leitungsbandkante des n-Leiters unterhalb der Kante des p-Leiters zu liegen kommt. Elektronen aus dem Valenzband können dann durch die Bandlücke ins weitgehend leere Leitungsband des n-Leiters tunneln. Da dieser starke Stromfluß sehr plötzlich einsetzt, nutzt man diesen Effekt oft in sogenannten Zener-Dioden aus, um in elektronischen Schaltkreisen eine konstante Referenzspannung zu erzeugen.

Solarzelle und Photodiode

Solarzelle

Häufig sind p-n-Übergänge das Herzstück von Solarzellen. Wird ein Photon in einem Halbleiter absorbiert, so wird ein Elektron-Loch-Paar erzeugt. Erfolgt diese Absorption in der Raumladungszone, so werden die beiden Ladungsträger durch das dort herrschende elektrische Feld getrennt. Dadurch entsteht ein zusätzlicher Strom I_L, der die Richtung des Feldstroms besitzt. Dieser Strom addiert sich zum Strom ohne Beleuchtung. Wir ersetzen daher Gleichung (9.58) durch

$$I = I_\mathrm{s}\left(\mathrm{e}^{eU/k_\mathrm{B}T} - 1\right) - I_\mathrm{L}\,, \tag{9.59}$$

wobei wir anstelle der Stromdichte die Gleichung für die Stromstärke angeschrieben haben. In Abb. 9.25 ist die Strom-Spannungscharakteristik einer etwa $4\,\mathrm{cm}^2$ großen Siliziumsolarzelle bei Lichteinfall dargestellt. Zur Berechnung der Kennlinie wurden die Parameterwerte $I_\mathrm{L} = 100\,\mathrm{mA}$ und $I_\mathrm{s} = 1\,\mathrm{nA}$ benutzt, die für derartige Bauelemente typisch sind. Ebenfalls eingezeichnet ist der Prinzipschaltkreis mit dem Verbraucherwiderstand R_L, dessen optimale Größe durch die folgenden Überlegungen bestimmt ist.

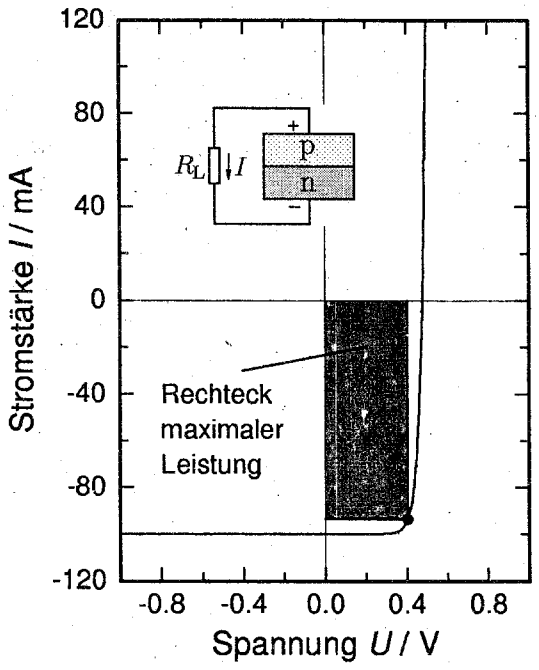


Abb. 9.25: Strom-Spannungscharakteristik einer Siliziumsolarzelle bei Beleuchtung. Der optimale Arbeitspunkt ist gefunden, wenn die Fläche des dunklen Rechtecks maximal ist. Eingezeichnet ist auch der Prinzipschaltkreis mit Verbraucherwiderstand R_L.

Der Stromkreis wird durch zwei Größen charakterisiert: durch die Leerlaufspannung und den Kurzschlußstrom. Wird eine Solarzelle im Leerlauf, d.h. bei $I = 0$, betrieben, so stellt sich nach Gleichung (9.59) am p-n-Übergang die Spannung $U \approx (k_B T/e) \ln(I_L/I_s)$ ein. Die Leerlaufspannung liegt meist im Bereich von 0,5 V. Im Falle des Kurzschlusses, also bei $U \doteq 0$, ist der Strom durch I_L gegeben. Er wird ausschließlich durch den lichtinduzierten Beitrag bestimmt und ist direkt proportional zur Beleuchtungsstärke. Die Sonnenenergie wird dann optimal genutzt, wenn die an einen Verbraucher abgegebene Leistung $P = U \cdot I$ ihren maximalen Wert besitzt. Die Betriebsbedingungen müssen daher so gewählt werden, daß die Fläche des grauen Rechtecks in der Abbildung 9.25 möglichst groß ist. Bei den üblichen Kennlinien ist dies der Fall, wenn die Arbeitsspannung bei etwa 80% der Leerlaufspannung liegt. Um einen hohen Wirkungsgrad zu erzielen, muß deshalb der Verbraucherwiderstand R_L den Parametern der Solarzelle angepaßt werden.

Warum ist der erzielbare Wirkungsgrad von Solarzellen relativ klein? Ausschlaggebend ist, daß einerseits Photonen mit einer Energie kleiner als die Bandlücke keinen Beitrag leisten, andererseits der Energieüberschuß der energiereichen Photonen verloren geht, da sie ebenfalls nur ein Elektronen-Loch-Paar erzeugen. Der Wirkungsgrad hängt daher von der Energielücke des verwendeten Halbleiters und der Form des Beleuchtungsspektrums ab. In Abb. 9.26 ist der ideale Wirkungsgrad von Solarzellen eingetragen, der die relative Lage von Sonnenspektrum und Bandlücke des betreffenden Bauelements widerspiegelt. In der Darstellung ist angenommen, daß sich die Sonne 45° über dem Horizont befindet. Dem Schaubild läßt sich entnehmen, daß ohne weitere Vorkehrungen, wie Lichtkonzentration oder ein komplizierter Aufbau aus einer Aneinanderreihung von Halbleitern mit verschiedenen Energielücken, ein maximaler Wirkungsgrad von 31% erzielt werden kann. Dieser Wirkungsgrad wird erreicht, wenn die Energielücke mit der Photonenenergie im Maximum des Sonnenspektrums zusammenfällt.

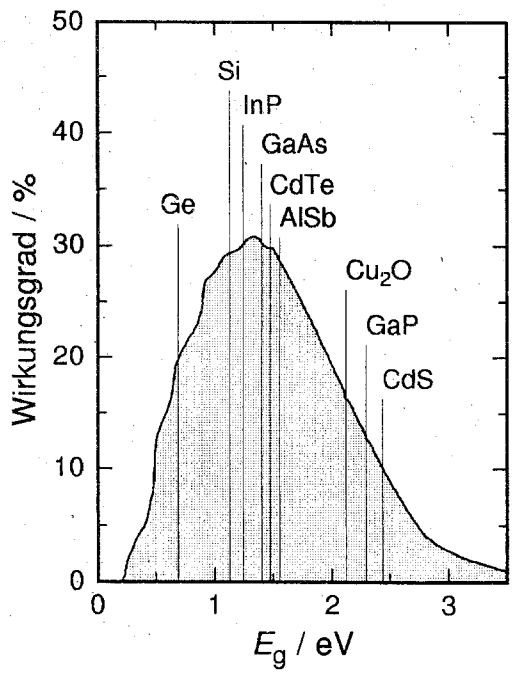


Abb. 9.26: Idealer Wirkungsgrad von Solarzellen. Der Wirkungsgrad ergibt sich als Schnitt der Linie für die Energielücke des Halbleiters und dem grau getönten Sonnenspektrum. Bei der Darstellung des Sonnenspektrums wurde angenommen, daß die Sonne 45° über dem Horizont steht.

Photodiode

Auch in Photodioden sind p-n-Übergänge oft das zentrale Element. Wie wir gesehen haben, bewirkt die Einstrahlung von Licht in einen p-n-Übergang einen Stromfluß I_L. Da mit der Absorption eines Photons jeweils ein Elektron-Loch-Paar erzeugt wird, ist bei vorgegebener Wellenlänge der Strom im p-n-Übergang proportional zur Intensität des einfallenden Lichts. Wie in Abb. 9.27 angedeutet, verschiebt sich dadurch die Kennlinie nach unten, wodurch sich der Spannungsabfall am Arbeitswiderstand R_L verändert. Um ein Signal zu erhalten, das weitgehend unabhängig von der Vorspannung ist, betreibt man die Photodiode, wie in der Abbildung durch Punkte angedeutet, in Sperrichtung.

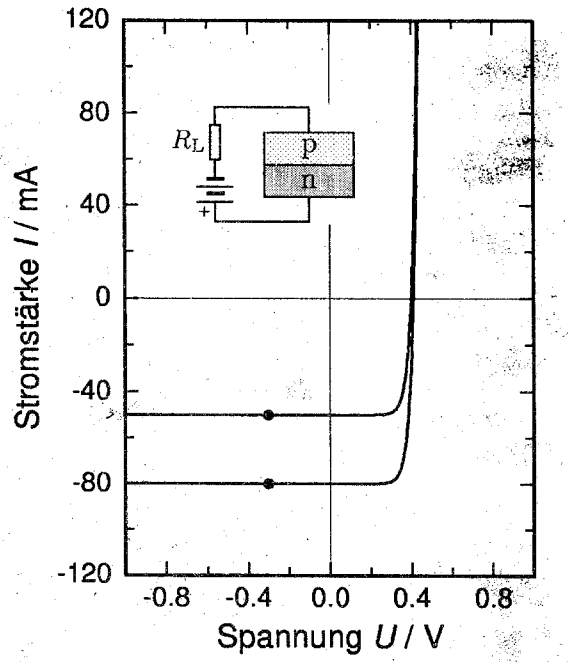


Abb. 9.27: Strom-Spannungscharakteristik einer Photodiode bei zwei Lichtintensitäten. Mit zunehmender Intensität verschiebt sich die Kennlinie nach unten. Die Punkte markieren typische Arbeitspunkte. Der Prinzipschaltkreis mit dem Arbeitswiderstand R_L ist ebenfalls eingezeichnet.

Leuchtdiode (LED)

Leuchtdioden[7] benutzt man häufig bei Anzeigen in Meßgeräten oder Rechnern. Zur Lichterzeugung nutzt man den Effekt, daß in einer Diode im Durchlaßbereich die Majoritätsladungsträger nach der Durchquerung der Raumladungszone rekombinieren. Bei Halbleitern mit direkter Bandlücke erfolgt dabei häufig ein Übergang des Elektrons aus dem Leitungsband in ein Loch des Valenzbands unter Emission eines Photons. Dieser Prozeß ist also die Umkehrung des optischen Absorptionsprozesses, der in Abschnitt 9.1 diskutiert wurde. Bei Halbleitern mit indirekter Bandlücke läuft die Rekombination meist *strahlungslos* ab, d.h. die Rekombinationsenergie wird vollständig in Form von Phononen an das Gitter abgegeben. Um auch derartige Halbleiter als Leuchtdioden nutzen zu können, baut man geeignete *Rekombinationszentren* ein. Das sind Störstellen mit einem Energieniveau in der Nähe des Valenzbandes, zu dem ein strahlender Übergang erfolgt. Ein sehr häufig benutzter Halbleiter zur Herstellung von Leuchtdioden ist der Mischkristall $GaAs_{1-x}P_x$, der für $0 \leq x < 0{,}45$ ein direkter Halbleiter und für größere Werte von x ein indirekter Halbleiter ist. Bei reinem Galliumarsenid liegt die emittierte Strahlung im Infraroten, da die Bandlücke nur 1,42 eV beträgt. Bei $x = 0{,}45$ ist $E_g = 1{,}98\,\text{eV}$, und

[7] Die Abkürzung LED kommt aus dem Englischen und steht für light-emitting diode.

die Dioden leuchten rot. Um Emission im gelbgrünen Spektralbereich zu erhalten, wird Stickstoff als Rekombinationszentrum eingebaut, da dann der Halbleiter eine indirekte Bandlücke aufweist und der oben erwähnte „Umweg" gegangen werden muß. Seit kurzem gibt es auch blauleuchtende Dioden, die aus GaN bestehen und mit Magnesium dotiert sind, das als Rekombinationszentrum wirkt.

Transistoren

Der Name „Transistor" ist eine Abkürzung aus „transferred resistor", was andeutet, daß Strom von einem Schaltkreis auf einen anderen übertragen wird. Wir betrachten hier zunächst kurz den bipolaren p-n-p-Transistor, der im Grunde eine Kombination aus zwei p-n-Übergängen darstellt. Eine umfassendere Theorie dieses Transistors ist relativ aufwendig, doch wollen wir uns hier nur mit der prinzipiellen Wirkungsweise beschäftigen. Anschließend beschreiben wir kurz die Arbeitsweise eines MOS-Feldeffekttransistors.

Bipolarer Transistor

Wie Abb. 9.28 zeigt, besteht ein p-n-p-Transistor aus Emitter, Basis und Kollektor. Für den Aufbau des Transistors ist wichtig, daß die Dicke der Basis so gering ist, daß dort die Rekombination der Löcher keine Rolle spielt.

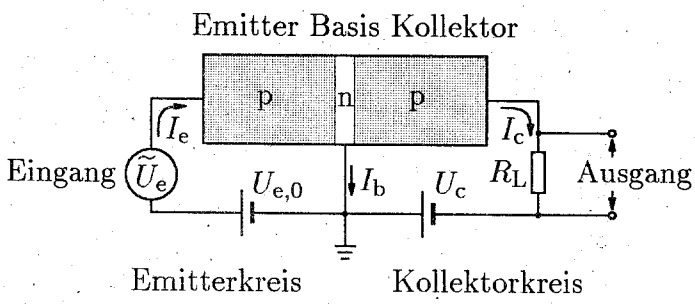


Abb. 9.28: Schematischer Aufbau und elektrische Beschaltung des „klassischen" p-n-p-Transistors. Die relevanten Spannungen und Teilströme sind ebenfalls angegeben.

Der in Durchlaßrichtung gepolte Emitter-Basis-Übergang „emittiert" Löcher in die Basis, in der sie diffundieren, bis sie an den in Sperrichtung gepolten Basis-Kollektor-Übergang gelangen. Dort sind die Löcher Minoritätsladungsträger und können ungehindert passieren. Legt man beispielsweise eine positive Signalspannung an den Emitter, so werden zunächst Löcher am linken p-n-Übergang erzeugt. Nachdem diese Löcher durch die Basis diffundiert sind, werden sie am Kollektor aufgesammelt und bewirken eine Stromänderung am Lastwiderstand R_L des Kollektorkreises. Entscheidend ist, daß die Löcher, deren Erzeugung von der Emitter-Basis-Spannung U_e kontrolliert wird, am Basisanschluß „vorbei" zur Kollektorsperrschicht diffundieren. Der Basisstrom I_b ist deshalb sehr klein, und der Emitterstrom fließt weitgehend in den Kollektor. Emitter- und Kollektorstrom sind also in etwa gleich groß und somit auch unabhängig von U_c und damit vom Wert des Widerstands R_L. Die Spannung U_L am Lastwiderstand kann deshalb viel größer sein als das eingangs benutzte Signal im Emitterkreis, d.h. das Signal wird verstärkt.

Grob läßt sich die Verstärkung wie folgt ermitteln: Sind $U_e = U_{e,0} + \tilde{U}_e$ und I_e Spannung und Strom im Emitterkreis, wie in Abb. 9.28 angegeben, so folgt aus unseren Betrachtungen zum p-n-Übergang aus Gleichung (9.58)

$$I_e = I_e^s \left(e^{eU_e/k_BT} - 1\right) \approx I_e^s \, e^{eU_e/k_BT} , \tag{9.60}$$

wobei I_e^s den Sättigungsstrom des Emitters in Sperrichtung darstellt. Bezeichnen wir mit α den Bruchteil der Löcher, der durch die Basis diffundiert, nicht rekombiniert und am Kollektor ankommt, so können wir für den Kollektorstrom schreiben:

$$I_c = I_c^s + \alpha I_e \approx \alpha I_e \,. \tag{9.61}$$

Die letzte Beziehung gilt, weil die Basis-Kollektor-Diode in Sperrichtung gepolt und der Sättigungsstrom I_c^s deshalb sehr klein ist. Damit erhalten wir für die Spannung U_L am Lastwiderstand:

$$U_L = R_L I_c = \alpha R_L I_e \,. \tag{9.62}$$

Mit diesen Gleichungen läßt sich der Verstärkungsfaktor sofort angeben:

$$\frac{dU_L}{dU_e} = \frac{\alpha R_L I_e e}{k_B T} \,. \tag{9.63}$$

Mit den typischen Zahlenwerten $I_e = 10\,\text{mA}$, $k_B T/e \approx 0{,}025$ V, $\alpha \approx 1$ und $R_L = 1\,\text{k}\Omega$ finden wir für den Verstärkungsfaktor den Wert 400.

MOSFET

Voraussetzung für das Verständnis eines Metall-Oxid-Halbleiter-Feldeffekttransistors, abgekürzt MOSFET nach den Anfangsbuchstaben der englischen Bezeichnung, ist die Kenntnis der Wirkungsweise einer MOS-Grenzfläche. In Abb. 9.29 sind der Aufbau einer derartigen Grenzfläche und die Energieniveaus für den Fall eines p-dotierten Halbleiters skizziert.

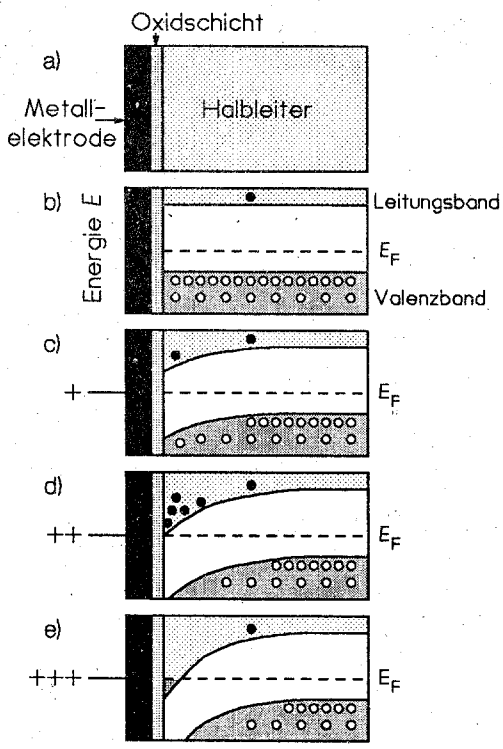


Abb. 9.29: Zur Wirkungsweise einer Metall-Oxid-Halbleiter-Grenzschicht. Die Elektronen sind durch Punkte bzw. durch dunkle Tönung, die Löcher durch offene Kreise symbolisiert. Die Gegenelektrode ist nicht dargestellt.
a) Grenzschicht, bestehend aus Halbleiter, Oxidschicht und Metallelektrode,
b) Energieniveaus ohne Feld,
c) Verarmungsschicht, verursacht durch eine kleine positive Vorspannung,
d) n-Inversionsschicht, bei großer positiver Vorspannung,
e) Inversionsschicht mit entartetem Elektronengas.

Legt man an die Elektrode eine positive Vorspannung, so werden Löcher abgestoßen und Elektronen angezogen. Wegen der p-Dotierung tritt ähnlich wie beim p-n-Übergang eine *Verarmung* an freien Ladungsträgern ein, da mehr Löcher ab- als Elektronen zufließen. Da die Ladung der Akzeptoren nicht mehr durch Löcher kompensiert wird, entsteht dabei eine hochohmige Raumladungszone. Wie in Abb. 9.29 angedeutet, ist im Gleichgewicht das chemische Potential in der Probe konstant. Das Potential, dessen genauer Verlauf hier ohne Bedeutung ist, läßt sich durch numerische Integration der Poisson-Gleichung errechnen. Wird die Vorspannung weiter erhöht, so senken sich die Bänder an der Grenzfläche weiter ab, und die Elektronendichte nimmt dort zu. Da in dieser dünnen Schicht der Leitungsmechanismus von der p- zur n-Leitung übergeht, spricht man von einer *Inversionsschicht*. Sie ist durch die isolierend wirkende Verarmungszone vom p-leitenden Substrat getrennt. Bei weiterer Erhöhung des Potentials taucht die Leitungsbandkante schließlich unter das Fermi-Niveau, und es bildet sich ein entartetes Fermi-Gas mit metallischem Charakter aus. Da dieser gut leitende Kanal relativ schmal ist, stellt die Inversionsschicht eine experimentelle Realisierung eines zweidimensionalen Elektronengases dar. Der Quanten-Hall-Effekt, den wir in Abschnitt 8.5 behandelt haben, wurde zum ersten Mal an einer derartigen Struktur beobachtet.

Nun sind wir in der Lage, die Schalteigenschaften eines MOSFET, dessen Aufbau in Abb. 9.30 schematisch dargestellt ist, zu verstehen. Zwei hoch dotierte n-Bereiche, bedeckt mit einer dünnen Oxidschicht (meist a-SiO_2), sind in p-leitendes Material eingebettet. Die Oxidschicht wird von zwei Kontakten durchbrochen, die die Verbindung zu der n-dotierten Elektronenquelle („Source") und der ebenfalls n-dotierten Elektronensenke („Drain") herstellen. Der dritte Kontakt zum Tor („Gate") bleibt durch eine etwa 100 nm dicke Oxidschicht vom p-dotierten Substrat getrennt. Dieser Kontakt bildet die soeben diskutierte MOS-Grenzfläche.

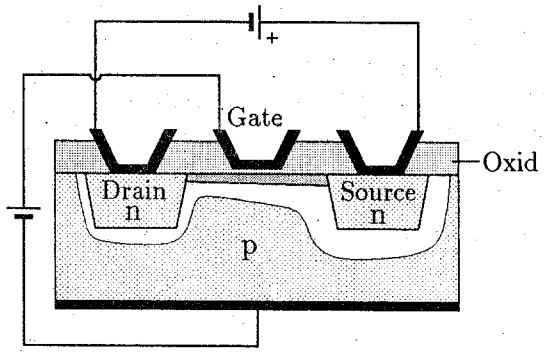


Abb. 9.30: Schematische Darstellung eines MOSFET. Die dunklen Metallelektroden stehen in direktem Kontakt mit den n-Bereichen von Source und Drain. Die positive Gatespannung ruft den dunkel gezeichneten leitfähigen Kanal hervor. Die helle Verarmungszone trennt Kanal und die n-leitenden Bereiche vom p-leitenden Substrat.

Ohne Vorspannung am Gate ist einer der beiden p-n-Übergänge gesperrt und verursacht einen hohen Widerstand. Legt man nun eine genügend große positive Gatespannung an, so bildet sich der oben beschriebene niederohmige, metallisch leitende Kanal aus. Kanal, Source und Drain sind durch Verarmungszonen vom p-leitenden Substrat getrennt. Die Breite der Verarmungsschicht variiert mit dem lokalen Potential. Der Transistor kann über die Gatespannung leicht in den stromdurchflossenen Zustand „an" oder den stromlosen Zustand „aus" gebracht werden.

9.5 Halbleiterheterostrukturen und Übergitter

Mit Hilfe der Molekularstrahlepitaxie oder der metallorganischen Gasphasenepitaxie lassen sich unterschiedliche Halbleiter mit nahezu perfekter durchgehender kristalliner Struktur in Schichten aufeinander abscheiden, die man als *Heterostrukturen* bezeichnet. Voraussetzung hierfür ist, daß sich die Gitterparameter der beiden Halbleiter kaum, d.h. weniger als 1%, unterscheiden. Dies ist beispielsweise bei den Systemen GaP/Si, GaAs/Ge oder InAs/GaSb der Fall. Hinzu kommen noch ternäre oder quaternäre Legierungen, wie z.B. $Al_xGa_{1-x}As$ oder $Ga_xIn_{1-x}As_yP_{1-y}$. Durch ein entsprechendes Mischungsverhältnis läßt sich die Bandlücke den Erfordernissen anpassen. Beim System $Al_xGa_{1-x}As$ kann man so die Energielücke zwischen 1,4 eV (GaAs) und 2,2 eV (AlAs) kontinuierlich einstellen.

Eine der vielen interessanten Fragen, die sich bei diesen neuartigen Systemen stellt, ist die Frage nach dem Aussehen der Bänderschemata. Bringt man zwei Halbleiter mit unterschiedlicher Bandlücke zusammen, wie in Abb. 9.31 schematisch dargestellt, so bilden sich eine *Banddiskontinuität* und eine *Bandverbiegung* aus. Der Übergang von einem zum anderen Bandabstand kann, wie in dieser Abbildung gezeichnet, sehr abrupt erfolgen, nämlich innerhalb eines atomaren Abstands. Die Banddiskontinuität tritt am Valenz- und am Leitungsband auf. Beim System GaAs/Ge findet man beispielsweise die Werte $\Delta E_V = 0{,}49\,\text{eV}$ und $\Delta E_L = 0{,}18\,\text{eV}$. Die theoretische Herleitung dieser Zahlenwerte ist aufwendig und soll hier nicht nachvollzogen werden. Natürlich stellt sich im thermischen Gleichgewicht im gesamten Halbleiter das gleiche chemische Potential ein. Dies hat zur Folge, daß Bandverbiegungen wie beim p-n-Übergang auftreten, die sich je nach Ladungsträgerdichte über Abstände von einigen hundert Ångström erstrecken. Bei ihrer Beschreibung geht man wie bei klassischen p-n-Übergängen vor. Zusätzlich muß berücksichtigt werden, daß die beiden Halbleiter unterschiedliche Dielektrizitätskonstanten besitzen und die Kontinuität der dielektrischen Verschiebung beachtet werden muß.

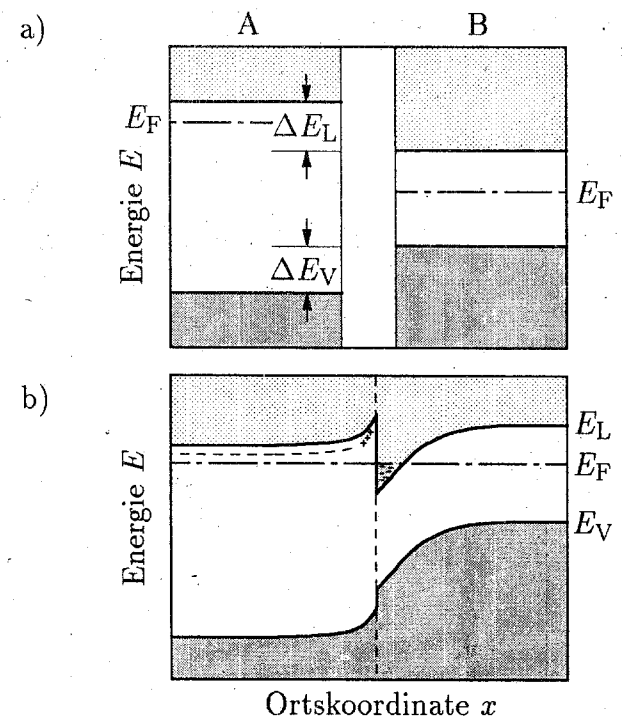


Abb. 9.31: Bänderschema einer Heterostruktur, gebildet aus zwei n-dotierten Halbleitern mit unterschiedlichen Bandlücken.
a) Es treten die beiden Banddiskontinuitäten ΔE_L und ΔE_V auf. Weiter wurde angenommen, daß der Halbleiter A stark, der Halbleiter B nur schwach dotiert ist. Entsprechend verschieden liegen die beiden Fermi-Niveaus.
b) Die Halbleiter sind in Kontakt. Die Fermi-Niveaus haben sich angeglichen, und die Bandverbiegungen haben sich ausgebildet.

Wir wollen uns hier kurz den sogenannten *isotypen Heteroübergängen* zuwenden. Hierbei handelt es sich um zwei unterschiedliche Halbleiter mit gleichartiger Dotierung. Dieser Fall wurde bereits in Abb. 9.31 dargestellt. Da im thermischen Gleichgewicht das Fermi-Niveau im ganzen Halbleiter eine feste Lage besitzen muß, tritt im Halbleiter B mit der kleineren Bandlücke in der Nähe des Übergangs eine Anreicherung von freien Elektronen auf. Unter geeigneten Bedingungen liegt am Übergang das Fermi-Niveau sogar im Leitungsband, so daß dort ein entarteter Halbleiter vorliegt. Dem steht eine Verarmungsraumladungszone im Halbleiter A gegenüber. Die Elektronen sind vom Halbleiter A zum Halbleiter B übergewechselt, da dort ein energetisch tiefer liegender Potentialtopf vorhanden ist. Diese Argumentation ist auch gültig, wenn der Halbleiter B nicht dotiert, also intrinsisch ist und kaum Störstellen aufweist. Es tritt dann der überraschende Fall auf, daß sich im undotierten Halbleiter nahezu alle freien Elektronen aufhalten.

Der Elektronentransfer führt zu der bemerkenswerten Situation, daß eine hohe Elektronendichte räumlich getrennt von den Dotieratomen vorkommt, während in klassischen Halbleitern eine hohe Ladungsträgerkonzentration immer mit einer hohen Störstellendichte verknüpft ist. Die Beweglichkeit der Elektronen ist dann bei tiefen Temperaturen durch die Streuung an geladenen Störstellen begrenzt. Es ist daher zu erwarten, daß in den hier diskutierten Heterostrukturen besonders hohe Beweglichkeitswerte auftreten, da sich die Elektronen in den Bereichen geringer Störstellenkonzentration aufhalten. Tatsächlich wurden bei tiefen Temperaturen enorme Werte für die Elektronenbeweglichkeit gefunden. In Abb. 9.32 ist gezeigt, wie mit der Weiterentwicklung der Molekularstrahlepitaxie die Proben verbessert wurden. Die Probe mit der höchsten hier gezeigten Beweglichkeit besaß eine relativ komplexe Schichtstruktur, die wir hier nicht näher erläutern wollen. Die Dotierung erfolgte mit Hilfe von Silizium in der $Al_{0,35}Ga_{0,65}As$-Schicht, das eine größere Bandlücke als GaAs besitzt. Die Überschußelektronen der Donatoratome werden an das energetisch tiefer liegende Leitungsband des angrenzenden intrinsischen GaAs-Materials abgegeben. Verglichen mit GaAs-Kristallen wurden eine Zunahme der Beweglichkeit um etwa vier Größenordnungen beobachtet, da epitaktisch hergestelltes GaAs weitgehend störungsfrei ist.

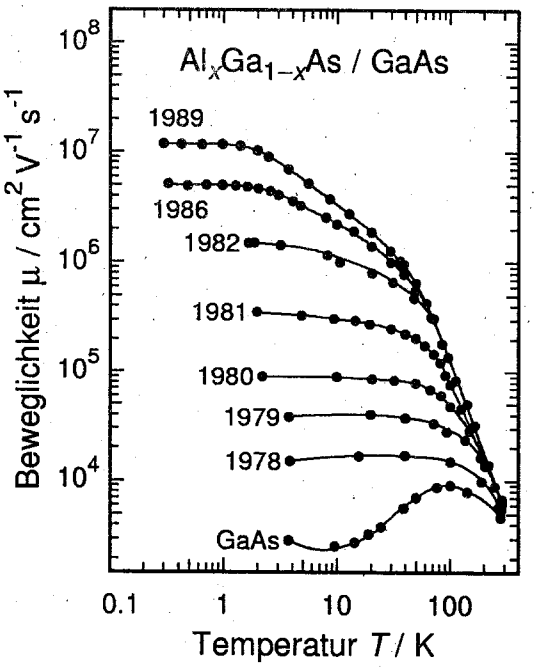


Abb. 9.32: Entwicklung der Elektronenbeweglichkeit in GaAs/AlGaAs-Heterostrukturen durch Verbesserung der Herstellbedingungen. Die Zahlen an den Kurven geben das Jahr der Herstellung an. Die höchsten Beweglichkeitswerte wurden mit einem $Al_{0,35}Ga_{0,65}As$/GaAs-System erreicht. (Nach L.Pfeiffer, K.W. West, H.L. Störmer, K.W. Baldwin, Appl. Phys. Lett. **55** (1989) 1886).

Reiht man die unterschiedlichen Materialien periodisch aneinander, so kommt man zu *Übergittern*. Im angesprochenen Beispiel handelt es sich um *Kompositionsübergitter*, da unterschiedliche Halbleiter zusammengefügt werden. Die Aneinanderreihung der soeben betrachteten Heterostruktur führt daher zu einem „dotierungsmodulierten Kompositionsübergitter". Wie in Abb. 9.33 gezeigt ist, wechseln sich nun Schichten des Halbleiters A mit starker n-Dotierung mit Schichten des Halbleiters B ab, die nahezu intrinsisch sind. Aufgrund der vorgegebenen neuen Bandstruktur wandern die Elektronen in die Potentialtöpfe des undotierten Halbleiters B. Dadurch entsteht in den stark dotierten Schichten des Halbleiters A jeweils eine hochohmige Verarmungszone.

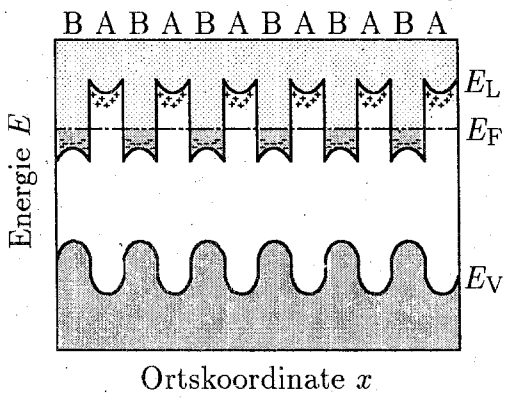


Abb. 9.33: Kompositionsübergitter, bestehend aus einer Aneinanderreihung von Strukturen der Art, wie sie bereits in Abb. 9.31 dargestellt sind. Die alternierend angeordneten Halbleiter A und B sind jeweils unterschiedlich dotiert.

Das Verhalten der modulationsdotierten Kompositionsübergitter ist auch in anderer Hinsicht bemerkenswert. So besitzen die in den Potentialtöpfen gefangenen Elektronen längs und senkrecht zu den Grenzflächen ganz unterschiedliche Eigenschaften. Parallel zu den Grenzflächen hat die Wellenfunktion der Elektronen den Charakter von ausgedehnten Bloch-Wellen. Senkrecht dazu, wir bezeichnen diese Richtung im folgenden als z-Richtung, ist die Bewegung durch den engen Potentialtopf stark eingeschränkt. Zur Diskussion der Eigenwerte der Elektronen greifen wir einen Potentialtopf heraus und nähern ihn in z-Richtung durch ein Kastenpotential. Sind die „Wände" dieses eindimensionalen Kastens unendlich hoch, so besitzen die Elektronen Eigenwerte ε_ν, die durch

$$\varepsilon_\nu = \frac{h^2}{8m_z^*}\frac{\nu^2}{d^2} \tag{9.64}$$

gegeben sind. Hierbei steht $\nu = 1, 2, 3, \ldots$ für die Quantenzahl der stehenden Elektronenwelle, d für die Breite des Potentialtopfs und m_z^* für die effektive Masse bei Bewegung in z-Richtung. Betrachten wir das Übergitter senkrecht zur z-Achse als unendlich ausgedehnt, so unterliegen die Elektronen bei ihrer Bewegung in x- oder y-Richtung keiner zusätzlichen Quantisierung und besitzen daher die Eigenwerte

$$E_\nu(k_x, k_y) = \frac{\hbar^2(k_x^2 + k_y^2)}{2m_{xy}^*} + \varepsilon_\nu \,. \tag{9.65}$$

Hier bezeichnet m_{xy}^* die effektive Masse, die mit der Elektronenbewegung in der xy-Ebene verbunden ist. Stellen wir E_ν als Funktion von k_x oder k_y dar, so erhalten wir eine Parabelschar. In Kap. 8.4 haben wir bei der Beschreibung der Landau-Niveaus einen analogen Sachverhalt (vgl. Abb. 8.23) kennengelernt. Der Unterschied besteht darin, daß das Magnetfeld eine Quantisierung der Elektronenbewegung in der xy-Ebene bewirkt, während hier das Auftreten von diskreten *Subbändern* durch die Einschränkung der Bewegung in z-Richtung verursacht wird.

Da sich die Elektronen nur längs der xy-Ebene bewegen können, stellen diese Heterostrukturen die Realisierung eines zweidimensionalen Elektronengases[8] dar. Die Zustandsdichte ist konstant und durch Gleichung (7.10) gegeben. Für jedes Subband ν gilt daher der einfache Ausdruck $D_2(E) = m^*_{xy}/\pi\hbar^2$. Die Zustandsdichte aller freien Elektronen ergibt sich aus den Beiträgen der einzelnen Bänder und ist daher eine Treppenfunktion. Die Stufen treten jeweils bei der Energie ε_ν auf, d.h. immer dann, wenn gerade ein neues Subband besetzt wird. Die Existenz dieser Zustandsdichte läßt sich mit Hilfe optischer Absorptions- oder Photolumineszenzexperimente nachweisen. Eine interessante Anwendung derartiger Heterostrukturen haben wir bereits in Abschnitt 8.5 kennengelernt: Die in Abb. 8.33 gezeigten Messungen zum Quanten-Hall-Effekt sind an einer $GaAs/Ga_{0,71}Al_{0,29}As$-Heterostruktur durchgeführt worden.

An dieser Stelle soll noch auf einen weiteren interessanten Effekt hingewiesen werden, den man an derartigen Übergittern beobachtet. Die Energieniveaus (9.64) sind natürlich nur dann scharf, wenn die einzelnen Potentialtöpfe des Übergitters genügend weit voneinander entfernt sind. Ist der Abstand zwischen den Potentialtöpfen kleiner als 100 Å, so wird der Überlapp zwischen den Wellenfunktionen merklich und bewirkt eine Aufspaltung der Einzelniveaus in Bänder. Dieser Vorgang ist völlig analog zur Aufspaltung der Atomniveaus im Festkörper, wie wir sie in Abschnitt 7.3 im Rahmen des Modells der „stark gebundenen" Elektronen diskutiert haben. Bei Übergittern besteht zusätzlich die Möglichkeit, den Abstand zwischen den Potentialtöpfen und damit den Überlapp der Wellenfunktionen über weite Bereiche zu variieren.

Einen weiteren interessanten Typus von Übergittern stellen die *Dotierungsübergitter* dar, deren Aufbau in Abb. 9.34 skizziert ist.

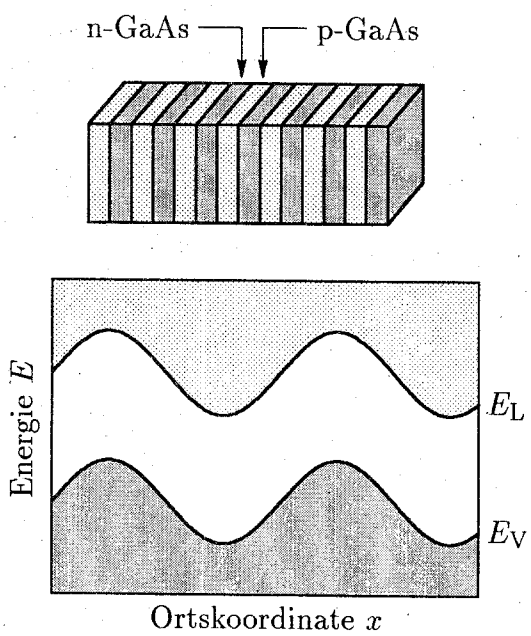


Abb. 9.34: Dotierungsübergitter.
a) Räumliche Struktur des Übergitters,
b) Modulation der Bänder im thermischen Gleichgewicht.

Sie bestehen aus einem Halbleiter, der abwechselnd n- und p-dotiert ist. Die Periode kann in weiten Bereichen variieren, liegt aber typischerweise bei einigen hundert Ångström. Da zwischen den n- und p-dotierten Bereichen auch ein schmaler intrinsischer Bereich auftritt oder auch absichtlich abgeschieden wird, bezeichnet man diese Übergitter auch als

[8] Eine entsprechende Betrachtung kann auch für die Löcher in den Maxima des Valenzbands durchgeführt werden.

„nipi-Strukturen". Wie Abb. 9.34 andeutet, besitzen diese Gitter wellenförmige Bandkanten. Dieser Verlauf wird wieder durch das Fermi-Niveau erzwungen, das im gesamten Halbleiter einen konstanten Wert haben muß. Durch die wechselnde Dotierung muß abwechselnd die Leitungsband- und Valenzbandkante näher am Fermi-Niveau liegen. Dies hat z.B. zur Folge, daß angeregte freie Elektronen in den Minima des Leitungsbands und Löcher in den Valenzbandmaxima sitzen. Die Ladungsträgertypen sind also räumlich voneinander getrennt. Dies wiederum führt dazu, daß die Lebensdauer von Elektronen und Löchern gegen Rekombination außerordentlich groß wird. Zusätzlich bewirkt die Modulation der Bänder – ähnlich wie bei den amorphen Halbleitern – eine Reduktion der effektiven optischen Bandlücke. Ein weiterer Effekt beruht auf der Tatsache, daß die Bandmodulation durch die Raumladungen bewirkt wird. So entstehen die „Täler" in den n-leitenden Gebieten, da dort positiv geladene Donatoren existieren. Im Bereich der „Berge" hat die Raumladung gerade das umgekehrte Vorzeichen. Erzeugt man durch intensive optische Beleuchtung viele freie Elektronen und Löcher, so reduzieren sie die effektive Raumladung. Dadurch vermindert sich auch die Modulationsstärke, d.h. die effektive optische Bandlücke kann durch Lichteinstrahlung verändert werden.

Halbleiter-Injektionslaser

Eine interessante, technisch sehr wichtige Anwendung von Heterostrukturen ist der Halbleiter-Injektionslaser. Die Wirkungsweise soll anhand von Abb. 9.35 erläutert werden.

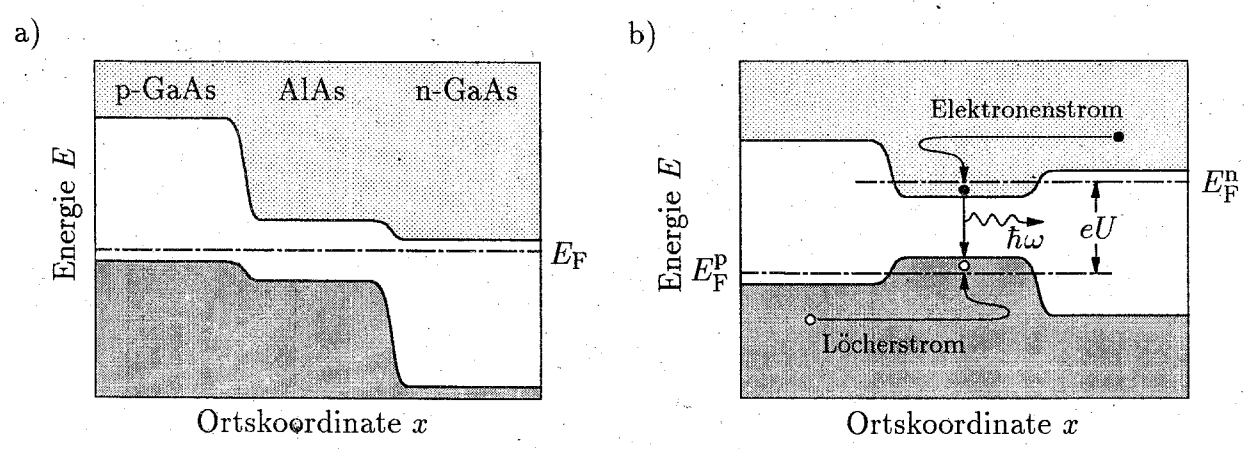


Abb. 9.35: Injektionslaser mit doppelter Heterostruktur.
a) Bandschema der Heteroübergänge im thermischen Gleichgewicht.
b) Bandstruktur bei der Betriebsspannung. Es treten, getrennt nach Ladungsträgertyp, die Quasi-Fermi-Niveaus E_F^n und E_F^p auf. (Nach Ch. Kittel, *Festkörperphysik*, Oldenbourg 1988).

Die aktive Schicht in der Mitte besteht aus AlAs und ist von GaAs umgeben, das eine größere Bandlücke aufweist. Das Galliumarsenid auf der linken Seite ist p-dotiert, das auf der rechten Seite n-dotiert. Ohne Spannung stellt sich ein Bandverlauf ein, wie er in Abb. 9.35a zu sehen ist. Legt man in Durchlaßrichtung eine genügend große Spannung an, so bilden sich für Elektronen und Löcher im Bereich des Aluminiumarsenids Potentialmulden aus. Aus dem p-dotierten Bereich fließen Löcher, aus dem n-dotierten Bereich Elektronen in diese Schicht. Die Potentialbarrieren verhindern ein Abfließen der Ladungsträger, und es baut sich eine Besetzungsinversion auf. Die Elektronen im n-dotierten Bereich befinden sich im diffusiven Gleichgewicht mit den Elektronen in der

aktiven Schicht. Entsprechendes gilt für die Löcher, doch herrscht wegen der angelegten Spannung kein thermisches Gleichgewicht „zwischen" Elektronen und Löchern. Ihre Besetzung in der aktiven Zone wird daher durch zwei getrennte Fermi-Niveaus $E_{\mathrm{F}}^{\mathrm{n}}$ und $E_{\mathrm{F}}^{\mathrm{p}}$ beschrieben, die man, wie bereits erwähnt, als Quasi-Fermi-Niveaus bezeichnet. Da diese Niveaus in den Bändern liegen, sind dort die Ladungsträger entartet. Die Elektronen rekombinieren mit den Löchern vorzugsweise in der aktiven Schicht und senden dabei Photonen aus. Inversion und damit Lasertätigkeit erreicht man, wie in Abb. 9.35b dargestellt, beim Anlegen einer Spannung U in Durchlaßrichtung, die größer ist, als der Energielücke der aktiven Schicht entspricht. Der Laserresonator wird ohne zusätzliche Maßnahmen durch die stark reflektierenden Grenzflächen zwischen Halbleiter und Luft gebildet.

10. Magnetismus

Elektronen und Kerne erzeugen Magnetfelder durch ihre magnetischen Momente und durch die Bewegung ihrer Ladung. Der Beitrag der Kerne ist vergleichsweise klein und spielt nur in besonderen Fällen eine Rolle. Hierzu zählen Gebiete der Tieftemperaturphysik und die Kernspinspektroskopie, doch werden wir diese Bereiche – abgesehen von einem speziellen Beispiel – hier ausklammern.

Ehe wir mit der Diskussion der magnetischen Eigenschaften beginnen, wollen wir noch einige wichtige Bezeichnungen in Erinnerung rufen. Wir schreiben für die magnetische Induktion $\boldsymbol{B} = \boldsymbol{B}_\mathrm{a} + \mu_0 \boldsymbol{M} = \mu_0(\boldsymbol{\mathcal{H}} + \boldsymbol{M})$, wobei der Index a für „außen“ steht und $\boldsymbol{\mathcal{H}} = \boldsymbol{B}_\mathrm{a}/\mu_0$ somit das äußere Magnetfeld angibt. Um den Sprachgebrauch zu „vereinfachen“, wird im folgenden oft auch die Induktion $\boldsymbol{B}$ als „Magnetfeld“ bezeichnet. Die Magnetisierung $\boldsymbol{M}$ ist als das magnetische Moment $\boldsymbol{m}$ pro Volumen definiert. Es besteht daher der Zusammenhang $\boldsymbol{M} = \boldsymbol{m}/V = N\overline{\boldsymbol{\mu}}/V = n\overline{\boldsymbol{\mu}}$, wobei N für die Zahl der magnetischen Dipole, n für ihre Dichte und $\overline{\boldsymbol{\mu}}$ für ihr mittleres Dipolmoment steht. Die magnetischen Eigenschaften einer Probe werden durch die magnetische Suszeptibilität $[\chi] = \boldsymbol{M}/\boldsymbol{\mathcal{H}} = \mu_0 \boldsymbol{M}/\boldsymbol{B}$ bestimmt. Aus der Definition folgt, daß die Suszeptibilität eine tensorielle Größe ist. Um die Gleichungen zu vereinfachen, werden wir sie als Skalar behandeln und die Richtungsabhängigkeit nur berücksichtigen, wenn dies erforderlich ist.

Wir gehen hier nicht auf die Frage ein, wie groß das Magnetfeld ist, das tatsächlich an einem Dipol im Festkörper herrscht. Es soll nur bemerkt werden, daß benachbarte Dipole durchaus einen merklichen Einfluß auf die *lokale Feldstärke* ausüben können. Wir werden dieses Phänomen in Kap. 11 bei den dielektrischen Eigenschaften diskutieren, wo bei der Behandlung der elektrischen Dipole ein ähnliches Problem auftritt.

Makroskopische Dipolmomente findet man in den meisten Materialien nur, wenn ein äußeres Feld anliegt. Je nachdem, ob das Vorzeichen der magnetischen Suszeptibilität positiv oder negativ ist, bezeichnet man sie als *para-* oder *diamagnetisch*. Eine Sonderstellung nehmen *ferro-* und *ferrimagnetische* Festkörper ein, denn diese Substanzen besitzen eine *spontane Magnetisierung*, d.h. die magnetischen Dipole weisen auch ohne äußeres Feld eine Vorzugsrichtung auf. Wir werden auf diese Klasse von magnetischen Materialien und auf die eng verwandten *Antiferromagneten* in den Abschnitten 10.2 und 10.3 eingehen. Im letzten Abschnitt werden wir uns kurz mit den *Spingläsern* beschäftigen, in denen die Spins beim Abkühlen statistisch orientiert eingefroren werden.

10.1 Dia- und Paramagnetismus

Bei paramagnetischen Materialien sind angelegtes und induziertes Feld gleichgerichtet, bei diamagnetischen wirken die beiden Felder in entgegengesetzte Richtungen und verursachen so eine Verstärkung bzw. Schwächung des äußeren Feldes. Para- und Ferromagnetismus beruhen auf der Ausrichtung vorhandener Spin- und/oder Bahndrehimpulse, Diamagnetismus dagegen ist eine Folge der Änderung des Betrags des Bahndrehimpulses durch das anliegende Feld.

Diamagnetismus

Eine äußerst simple Erklärung für die Existenz von Diamagnetismus läßt sich mit Hilfe der Lenzschen Regel geben. Beim Einschalten des Magnetfelds ändert sich der von den „Elektronenbahnen" der Atome umschlossene Magnetfluß. Es wird ein Strom induziert, der ein Magnetfeld bewirkt, das dem ursprünglichen Feld entgegengerichtet ist. Da der angeworfene Strom in Atomen ungedämpft ist, kommt es zu einer dauerhaften Schwächung des äußeren Feldes. Dieser Effekt tritt bei allen Atomen oder Ionen auf; Diamagnetismus ist somit keine typische Festkörpereigenschaft.

Klassische und quantenmechanische Behandlung des Diamagnetismus führen zum gleichen Resultat. Für Substanzen, aufgebaut aus isotropen Atomen oder Ionen, findet man

$$\chi_{\mathrm{d}} = -\frac{n\mu_0 e^2}{6m_{\mathrm{e}}} Z\overline{r^2}, \tag{10.1}$$

wobei Z für die Elektronenzahl, $\overline{r^2}$ für den Mittelwert des Abstandsquadrats und m_{e} für die Elektronenmasse[1] steht. In dieser Näherung ist die diamagnetische Suszeptibilität der Festkörper richtungs- und temperaturunabhängig. Bei Ionenkristallen kann man die Beiträge der unterschiedlichen Ionen einfach addieren und findet dann wie auch bei den Edelgaskristallen gute Übereinstimmung mit experimentellen Daten. Gleichung (10.1) ist nicht auf kovalente oder gemischt kovalent-ionische Festkörper anwendbar, obwohl diese meist diamagnetisch sind. Die einfache Rechnung, die zu (10.1) führt, setzt isotrope Atome voraus und verliert ihre Gültigkeit, wenn sich die Bindungselektronen bevorzugt zwischen benachbarten Atomrümpfen aufhalten.

In Tabelle 10.1 sind experimentelle Werte der molaren Suszeptibilität für eine Reihe von Edelgasatomen und Ionen aufgeführt[2]. Bemerkenswert ist der weite Bereich, über den die Werte der Suszeptibilität verteilt sind.

Tabelle 10.1: Molare Suszeptibilität χ_{d} (in $10^{-6}\,\mathrm{cm}^3/\mathrm{mol}$) von Edelgasen und einigen ausgewählten Ionen. (Nach W.R. Myrers, Rev. Mod. Phys. **24** (1952) 1).

He	−24	Li^+	−88	F^-	−1181
Ne	−85	Na^+	−767	Cl^-	−3042
Ar	−246	K^+	−1835	Br^-	−4337
Kr	−362	Rb^+	−2765	J^-	−6360
Xe	−552	Cs^+	−4412	Mg^{2+}	−541

Diamagnetismus beobachtet man auch beim Elektronengas der Metalle, mit dem wir uns bereits in Abschnitt 8.4 beschäftigt haben. Die Kreisbewegung der Elektronen im Magnetfeld wird bei Zimmertemperatur zwar durch Stöße gestört, führt aber dennoch zum Aufbau eines Gegenfelds. Ohne auf die Rechnung einzugehen, wollen wir hier das Ergebnis festhalten:

$$\chi_{\mathrm{d}} = -\frac{n\mu_0\mu_{\mathrm{B}}^2}{2E_{\mathrm{F}}}\left(\frac{m_{\mathrm{e}}}{m^*}\right)^2 . \tag{10.2}$$

[1] Wir versehen hier ausnahmsweise die Elektronenmasse mit einem Index, um eine Verwechslung mit dem magnetischen Moment zu vermeiden.

[2] Beim Vergleich mit Literaturwerten ist zu beachten, daß die Suszeptibilität in SI- und CGS-Einheiten die gleiche Dimension hat, die SI-Werte jedoch um den Faktor 4π größer sind.

m^* steht wieder für die effektive Masse der Elektronen. Ist $m^* = m_e$, so kompensiert der *Landau-Diamagnetismus* zu einem Drittel den Pauli-Paramagnetismus der Elektronen. Beim Vergleich mit experimentellen Werten muß neben der effektiven Masse noch der Beitrag (10.1) der Ionenrümpfe berücksichtigt werden. Beide Effekte zusammen können bewirken, daß der Diamagnetismus sogar überwiegt, wie es beispielsweise beim metallischen Cäsium der Fall ist.

Paramagnetismus

Die magnetischen Eigenschaften ändern sich grundlegend, sobald die betrachtete Probe permanente magnetische Dipole enthält. Diese treten bei Atomen und Molekülen mit ungerader Elektronenzahl oder mit nur teilweise gefüllten inneren Schalen auf. Aber auch Gitterdefekte tragen oft ein magnetisches Moment und sind daher bei der Untersuchung von magnetischen Eigenschaften von Bedeutung. Wie aus der Atomphysik bekannt, setzt sich das magnetische Moment $\boldsymbol{\mu}$ aus den Beiträgen von Spin- und Bahnmomenten zusammen:

$$\boldsymbol{\mu} = -g\mu_B \boldsymbol{J} \quad \text{mit} \quad g = 1 + \frac{J(J+1) + S(S+1) - L(L+1)}{2J(J+1)} . \tag{10.3}$$

Hierbei ist g der Landé-Faktor, $\hbar\boldsymbol{J}$ der Gesamtdrehimpuls, L und S sind die Bahndrehimpuls- und Spinquantenzahlen.

Während der Diamagnetismus weitgehend temperaturunabhängig ist, nimmt die Suszeptibilität paramagnetischer Materialien meist, dem *Curie-Gesetz* folgend, mit $1/T$ ab. Eine wichtige Voraussetzung für die Gültigkeit dieses Gesetzes ist, daß die Wechselwirkung zwischen den magnetischen Momenten vernachlässigt werden kann. Ein Beispiel für dieses typische Curie-Verhalten zeigt Abb. 10.1, in dem die reziproke Suszeptibilität eines Dysprosiumsalzes als Funktion der Temperatur aufgetragen ist.

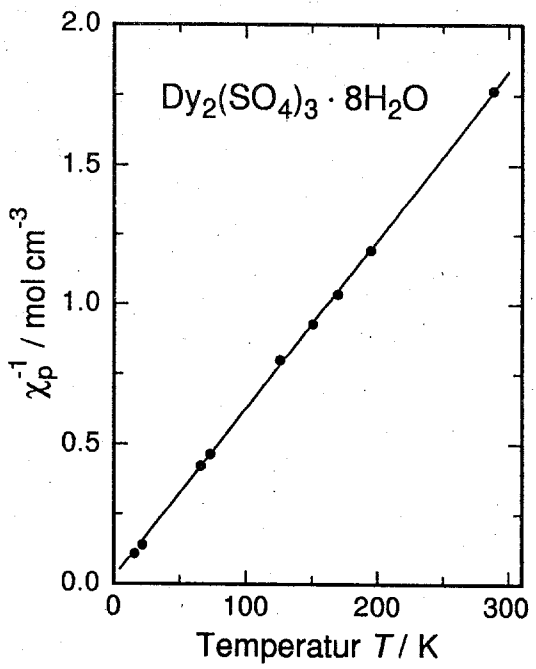


Abb. 10.1: Reziproke magnetische Suszeptibilität χ_p^{-1} von Dysprosium-Sulfat-Oktahydrat $Dy_2(SO_4)_3 \cdot 8\ H_2O$ als Funktion der Temperatur. (Nach L.C. Jackson, Proc. Phys. Soc. (London) **48** (1936) 741).

Wir berechnen zunächst die Magnetisierung in klassischer Näherung, wie sie ursprünglich von *Langevin* für ein ideales magnetisches Gas hergeleitet wurde. Legen wir an eine Probe mit permanenten Dipolen ein Feld an, so bewirkt es eine teilweise Ausrichtung der Dipole, da die potentielle Energie $U = -\boldsymbol{\mu} \cdot \boldsymbol{\mathcal{B}} = -\mu\mathcal{B}\cos\theta$ vom Winkel θ zwischen Feld und Dipolmoment abhängt. Sehen wir von tiefen Temperaturen oder sehr hohen Feldstärken ab, so kommt es nur zu einer partiellen Ausrichtung, da unter den üblichen experimentellen Bedingungen $\mu\mathcal{B} \ll k_\mathrm{B}T$ ist. Der Mittelwert $\langle\cos\theta\rangle = \mu\mathcal{B}/3k_\mathrm{B}T$ läßt sich leicht mit Hilfe der Thermodynamik errechnen, und man erhält bei einer Dipoldichte n für die Magnetisierung den Ausdruck:

$$M = n\mu\,\langle\cos\theta\rangle = n\mu\,\frac{\mu\mathcal{B}}{3k_\mathrm{B}T}\,. \tag{10.4}$$

Daraus ergibt sich für die Suszeptibilität unmittelbar der Zusammenhang

$$\chi_\mathrm{P} = \frac{\mu_0 n\mu^2}{3k_\mathrm{B}T}\,. \tag{10.5}$$

Nun wollen wir die quantenmechanische Behandlung erläutern, die berücksichtigt, daß im Magnetfeld die Energieniveaus der Atome in $(2J+1)$ Niveaus aufspalten, die jeweils den Abstand $g\mu_\mathrm{B}\mathcal{B}$ voneinander besitzen. Der Temperaturverlauf der Suszeptibilität wird durch die thermische Besetzung der Energieniveaus bestimmt. Den besonders einfachen Fall $J = 1/2$, d.h. die thermische Besetzung von Zweiniveausystemen, haben wir bereits in Abschnitt 7.2 diskutiert. Im allgemeinen Fall erhalten wir für die Magnetisierung einen analogen Ausdruck:

$$M = ng\mu_\mathrm{B}JB(x)\,. \tag{10.6}$$

Die Abhängigkeit der Magnetisierung von Temperatur und Magnetfeld ist im Argument $x = g\mu_\mathrm{B}J\mathcal{B}/k_\mathrm{B}T$ der *Brillouin-Funktion*

$$B(x) = \frac{2J+1}{2J}\coth\left[\frac{(2J+1)\,x}{2J}\right] - \frac{1}{2J}\coth\left(\frac{x}{2J}\right) \tag{10.7}$$

versteckt. Für sehr tiefe Temperaturen geht $x \to \infty$ und damit $B(x) \to 1$, da dann nur noch der Grundzustand besetzt ist und alle Momente in Feldrichtung ausgerichtet sind.

In den meisten Fällen allerdings ist die Aufspaltung klein gegen $k_\mathrm{B}T$, d.h. $x \ll 1$. Dann läßt sich die Brillouin-Funktion entwickeln, und man erhält

$$\chi_\mathrm{P} = \frac{\mu_0 n g^2 J(J+1)\mu_\mathrm{B}^2}{3k_\mathrm{B}T} = \frac{\mu_0 n p^2 \mu_\mathrm{B}^2}{3k_\mathrm{B}T} = \frac{C}{T}\,. \tag{10.8}$$

Dies ist das bereits erwähnte Curie-Gesetz, wobei C für die *Curie-Konstante* und die Größe p für die „effektive Magnetonenzahl" $p = g\sqrt{J(J+1)}$ steht. Offensichtlich stimmen der quantenmechanische und der klassische Ausdruck, d.h. Gleichung (10.8) und (10.5) überein, wenn wir $\mu = p\mu_\mathrm{B}$ setzen. Das magnetische Moment μ wird dadurch mit dem maximalen magnetischem Moment in z-Richtung identifiziert. Die Gültigkeit des Curie-Gesetzes wurde für paramagnetische Materialien in vielen Experimenten bestätigt. Trägt man wie in Abb. 10.1 die reziproke Suszeptibilität auf, so läßt sich aus der Steigung der Geraden unmittelbar der Wert der Curie-Konstanten C und damit die effektive Magnetonenzahl p ablesen.

In Abb. 10.2 ist die Temperatur- und Feldabhängigkeit der Magnetisierung von paramagnetischen Salzen mit reinem Spinmagnetismus gezeigt. Die Messungen wurden bei tiefen Temperaturen und Magnetfeldern bis 5 T durchgeführt. Bei den höchsten Feldern ist die magnetische Sättigung erreicht. Es ist deutlich zu sehen, daß die gemessenen Magnetisierungen exakt den eingezeichneten Brillouin-Kurven folgen.

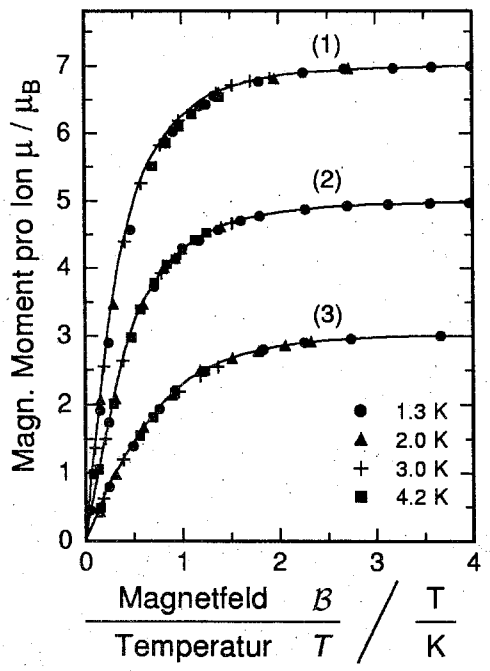


Abb. 10.2: Mittleres magnetisches Moment pro Ion abhängig von B/T. Gadolinium-Sulfat-Oktahydrat (1), Eisen-Ammonium-Alaun (2) und Kalium-Chrom-Alaun (3) besitzen der Reihe nach die Spinquantenzahl $S = 7/2$, $5/2$ und $3/2$. Die Kurven geben jeweils die Brillouin-Funktion für diese Werte wieder. (Nach W.E. Henry, Phys. Rev. **88** (1952) 559).

Am eingehendsten untersucht sind die magnetischen Eigenschaften der Eisengruppe und der Seltenen Erden. Bei den Ionen der Seltenen Erden werden die magnetischen Momente von den $4f$-Elektronen verursacht, die von den weiter außenliegenden $5s$- und $5p$-Elektronen abgeschirmt werden. Bei diesen Materialien passen die aus der Atomphysik bekannten und die gemessenen Werte der effektiven Magnetonenzahl gut zusammen. Eine derartige Übereinstimmung findet man bei den Ionen der Übergangsmetalle nicht. Dort werden die Beiträge der Bahnbewegung „ausgelöscht", und nur die Spinbeiträge bleiben wirksam. Dies läßt sich wie folgt verstehen: Im Festkörper herrscht ein von Ort zu Ort variierendes, starkes inhomogenes elektrisches Feld, das sogenannte *Kristallfeld*, das von den benachbarten Atomen hervorgerufen wird. Im Falle der $4f$-Elektronen der Seltenen Erden hat dieses Feld keinen großen Einfluß auf ihre Bewegung, da sie sich tief in den Ionenrümpfen aufhalten und das Kristallfeld durch die äußeren Elektronen abgeschirmt wird. Die $3d$-Elektronen der Ionen der Eisengruppe dagegen bilden die Oberfläche der Ionenrümpfe. Ihre Bewegung wird durch das Kristallfeld so stark verändert, daß sogar die LS-Kopplung aufgehoben wird. Im stark inhomogenen Kristallfeld bleibt zwar der Betrag des Bahndrehimpulses erhalten, doch ist die z-Komponente keine Konstante der Bewegung mehr. Mit dem zeitlichen Mittelwert der z-Komponente verschwindet auch der Beitrag der Bahnbewegung zum magnetischen Moment. Wie Abb. 10.2 zu entnehmen ist, sind dann nur die Spinbeiträge der d-Elektronen in die Betrachtung einzubeziehen, d.h. $p \to 2\sqrt{S(S+1)}$. Beispielsweise werden die paramagnetischen Messungen an Chrom-Alaun am besten mit den Quantenzahlen $L = 0$, $S = 3/2$ wiedergegeben, obwohl sich die Cr^{3+}-Ionen im ${}^4F_{3/2}$-Zustand befinden.

Wesentlich kompliziertere Verhältnisse liegen vor, wenn die magnetischen Ionen mit ihren d- oder f-Elektronen nicht in Isolatoren, sondern in Metalle mit freien Elektronen eingebaut werden. Es tritt dann eine Reihe interessanter Effekte auf, die wir hier natürlich nicht alle betrachten können. Wir wollen daher nur auf eine typische Beobachtung etwas näher eingehen. Die „verdünnte Legierung" **CuMn**, d.h. metallisches Kupfer mit einer geringen Konzentration an Manganionen, besitzt paramagnetische Eigenschaften. Die magnetische Suszeptibilität gehorcht dem Curie-Gesetz, aus dem sich die hohe effektive Magnetonenzahl $p = 4.9$ für die Manganionen ableiten läßt. Dies ist nicht unerwartet, da freie Manganatome mit der Elektronenkonfiguration $3d^5 4s^2$ und dem Grundzustand ${}^6S_{5/2}$ stark magnetisch sind. Überraschenderweise verschwinden die magnetischen Eigenschaften der Manganionen aber vollständig, wenn Aluminium als Wirtsmetall dient. Die Ursache hierfür ist, daß die lokalisierten d-Elektronen stark mit den delokalisierten s-Elektronen des Wirtsmetalls wechselwirken. Dies führt zu einer Hybridisierung der Zustände und so zu einer Verbreiterung der $3d$-Niveaus. Ist die Wechselwirkung wie bei Aluminium aufgrund der hohen Dichte an freien Elektronen sehr stark, so führt die große Verbreiterung zu einer Unterdrückung der lokalen magnetischen Momente. Den Einfluß der Elektronendichte verdeutlicht folgende Beobachtung: Die Ionen der Eisengruppe von Vanadium bis Kobalt zeigen in Gold die bekannten magnetischen Eigenschaften. Werden diese Metalle in geringer Konzentration in Zink mit der etwas höheren Dichte an freien Elektronen eingebracht, so treten nur noch bei Chrom und Mangan magnetische Momente auf. Bei Aluminium als Wirtsmetall wird der Magnetismus sogar in allen Fällen vollständig unterdrückt.

Die Magnetisierung paramagnetischer Salze wird in der Tieftemperaturphysik häufig zur Messung der Temperatur herangezogen, denn die Anwendbarkeit des Curie-Gesetzes erleichtert die Kalibrierung derartiger „Thermometer" ganz erheblich. In Abb. 10.3 ist der Temperaturverlauf der Suszeptibilität von Cer-Magnesium-Nitrat (CMN) bei sehr tiefen Temperaturen gezeigt. Deutlich zu sehen ist, daß bei dieser „klassischen Substanz der Tieftemperaturphysik" die $1/T$-Abhängigkeit der Suszeptibilität bis zu wenigen Millikelvin beobachtet wird.

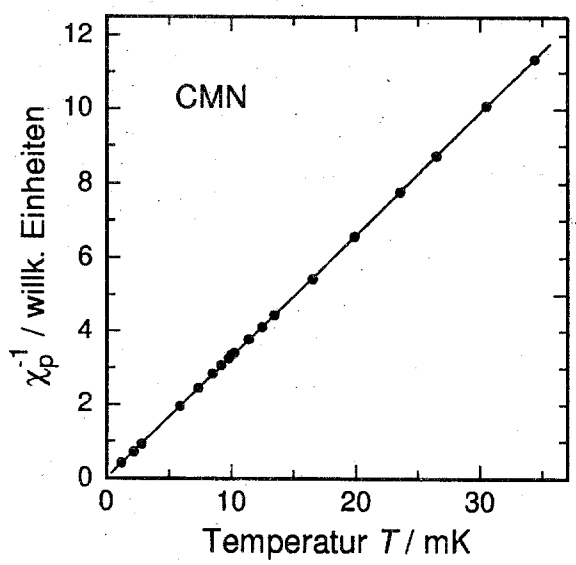


Abb. 10.3: Inverse magnetische Suszeptibilität χ_p^{-1} von Cer-Magnesium-Nitrat als Funktion der Temperatur. (Nach J.M. Parpia, W.P. Kirk, P.S. Kobiela, Z. Olejniczak, J. Low Temp. Phys. **60** (1985) 54).

Bei Tieftemperaturexperimenten tritt sehr häufig das Problem auf, daß sich das thermische Gleichgewicht zwischen Probe und Thermometer erst nach längerer Zeit einstellt. Da bei tiefen Temperaturen die thermische Leitfähigkeit von Metallen viel höher als die von Nichtmetallen ist (vgl. Abschnitt 8.2), sind Metalle als Meßfühler wesentlich geeigneter als Isolatoren. In den letzten Jahren werden daher zunehmend Metalle als Thermometer herangezogen, die mit geringen Mengen spintragender Substanzen dotiert sind. In Tieftemperaturexperimenten werden häufig die Systeme **Pd**Fe und **Au**Er zur Temperaturmessung benutzt. Die Dotierung mit Eisen bzw. Erbium liegt dabei im Konzentrationsbereich von 10 – 100 ppm, um die Wechselwirkung zwischen den magnetischen Momenten und damit Abweichungen vom Curie-Verhalten zu vermeiden.

Als Beispiel hierfür ist in Abb. 10.4 eine Messung der Suszeptibilität von **Au**Er gezeigt, die bis zu Temperaturen von 250 μK (!) durchgeführt wurde. Erst bei Temperaturen unter 0,5 mK treten in diesem System Abweichungen vom Curie-Verhalten auf, die in der hier gewählten Darstellung kaum zu erkennen sind. Sie werden durch die schwache Wechselwirkung zwischen den Spins der Erbiumatome hervorgerufen, die auf der sogenannten RKKY-Wechselwirkung beruht, auf die wir im nächsten Abschnitt kurz eingehen.

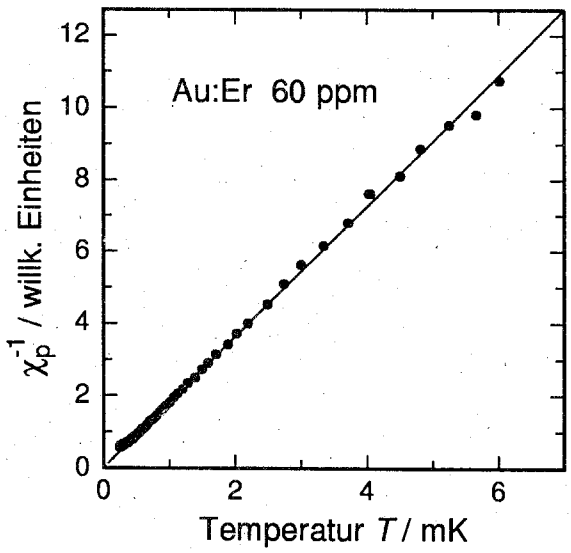


Abb. 10.4: Inverse magnetische Suszeptibilität χ_p^{-1} von 60 ppm Erbium in Gold gemessen bis 250 μK. (Nach R. König, T. Herrmannsdörfer, C. Enss (Daten nicht veröffentlicht)).

Wie am Anfang dieses Kapitels erwähnt, spielen Kerne in der Festkörperphysik meist keine Rolle da ihr magnetisches Moment etwa 1000 mal kleiner als das der Elektronen ist. Dennoch können auch Kerne zur Magnetisierungsthermometrie herangezogen werden. Speziell bei Temperaturen unter 1 mK ist der Kernmagnetismus zur Temperaturbestimmung geeignet[3]. So hat sich die Messung der Suszeptibilität von Platin mit Hilfe der Kernspinresonanz zu einer Standardmethode der Tieftemperaturthermometrie entwickelt. Mit modernen hochauflösenden SQUID-Magnetometern[4] lassen sich Suszeptibilitätmessungen mit hoher Genauigkeit selbst bei höheren Temperaturen durchführen.

Das Ergebnis einer derartigen Messung in einem Magnetfeld von 0,25 mT an einer sehr reinen Kupferprobe ist in Abb. 10.5 zu sehen. Diese Daten zeigen, daß man nicht nur

[3] Allerdings muß man dabei längere thermische Relaxationszeiten in Kauf nehmen, da Elektronen und Kernspins nur schwach gekoppelt sind.

[4] Das Prinzip der SQUID-Magnetometer werden wir in Abschnitt 12.4 kennenlernen.

ein perfektes „Curie-Verhalten“, sondern auch einen Absolutwert findet, der hervorragend mit den theoretischen Werten übereinstimmt.

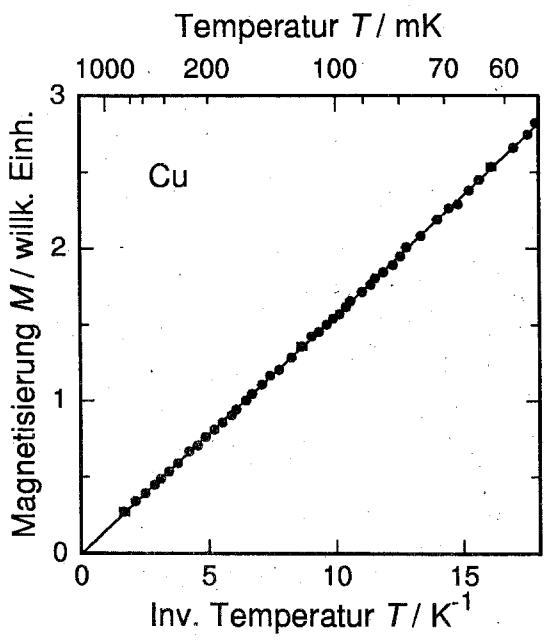


Abb. 10.5: Magnetisierung von reinem Kupfer als Funktion der inversen Temperatur $1/T$. Diese Messung wurde in einem Magnetfeld von 0,25 mT durchgeführt. (Nach R.A. Buhrman, W.P. Halprin, S.W. Schwenterly, J. Reppy, R.C. Richardson, W.W. Webb, Proc. 12th Int. Conf. Low Temp. Phys. (E. Kanda ed.) Academic Press, Tokyo 1971, p. 831).

Bei derartigen Messungen ist im allgemeinen Vorsicht geboten, da geringste Verunreinigungen mit magnetischen Ionen einen merklichen temperaturabhängigen Beitrag zur Magnetisierung liefern können. Beispielsweise würde bei kleinen Feldern eine Verunreinigung durch 1 ppm Eisen die gleiche Magnetisierung verursachen wie die Kupferkernspins selbst. Daher wird die Messung häufig bei möglichst hohen Feldern durchgeführt, da dann der Beitrag der Elektronenspins gesättigt ist, d.h. da dort ihre Magnetisierung bereits temperaturunabhängig ist.

Natürlich ist die Annahme der Unabhängigkeit der magnetischen Momente nur so lange aufrechtzuerhalten, wie die Wechselwirkung zwischen den Momenten vernachlässigt werden kann. Ist die Wechselwirkungsenergie nicht mehr klein verglichen mit der thermischen Energie, so treten Abweichungen vom Curie-Gesetz (10.8) auf. Diese können mit dem modifizierten Ausdruck $\chi_{\mathrm{p}} = C/(T+\Delta)$ berücksichtigt werden, wobei der Korrekturfaktor Δ die Stärke der Wechselwirkung widerspiegelt. Typische Werte für diese Größe liegen im Bereich $\Delta/k_{\mathrm{B}} < \pm 10\,\mathrm{K}$.

10.2 Ferromagnetismus

Während beim Dia- und Paramagnetismus die einzelnen magnetischen Momente als unabhängig voneinander betrachtet werden, beruht der Ferromagnetismus auf kollektiven Phänomenen. Ferromagnetische Materialien zeichnen sich durch eine spontane Magnetisierung aus, d.h. magnetische Momente sind ohne äußeres Feld ausgerichtet und in Kristallen regelmäßig angeordnet. Je nach Anordnung der magnetischen Momente unterscheidet man zwischen *ferro-*, *antiferro-* und *ferrimagnetischen* Systemen. Die Bedeutung dieser Begriffe verdeutlicht Abb. 10.6, in der die jeweilige Anordnung der magnetischen Momente skizziert ist.

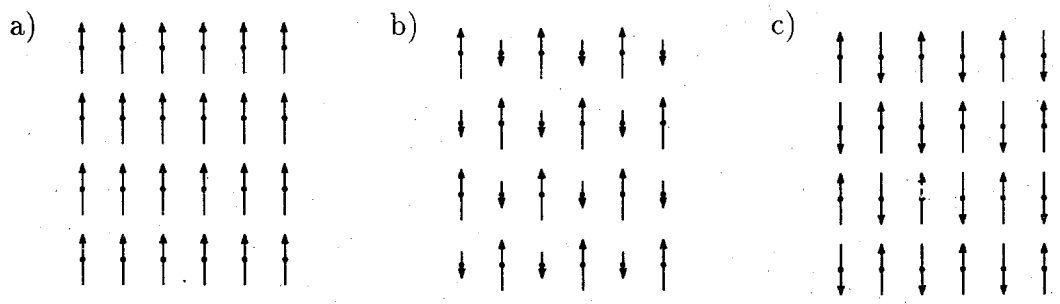


Abb. 10.6: Schematische Darstellung der drei Typen magnetischer Ordnung.
a) Ferromagnetismus: Alle magnetischen Momente sind gleich gerichtet.
b) Ferrimagnetismus: Die magnetischen Momente der zwei Untergitter unterscheiden sich in Größe und Richtung.
c) Antiferromagnetismus: Die Magnetisierung der Untergitter hebt sich auf.

Die Formulierung einer mikroskopischen Theorie des Ferromagnetismus erweist sich als eine sehr schwierige Aufgabe, da Einelektron- und Mehrelektronenaspekte berücksichtigt werden müssen. Obwohl bei verschiedenen Ferromagneten das gleiche phänomenologische Verhalten vorliegt, können die mikroskopischen Ursachen hierfür unterschiedlich sein: Der Ferromagnetismus kann auf der Austauschwechselwirkung zwischen delokalisierten $3d$-Elektronen und somit auf Bandeigenschaften beruhen, er kann aber auch durch das kooperative Verhalten von magnetischen Momenten hervorgerufen werden. Den ersteren der beiden Mechanismen findet man bei den „klassischen" Ferromagneten wie Eisen, Nickel und Kobalt. Der zweite Mechanismus ist dagegen bei den meisten ferromagnetischen Verbindungen und den $4f$-Übergangsmetallen wirksam.

Molekularfeldnäherung

Wir wollen im ersten Schritt die Magnetisierung ferromagnetischer Materialien mit Hilfe der *Molekularfeldnäherung* beschreiben, die auf *Weiss* (1907) zurückgeht. Da sie nicht auf mikroskopische Phänomene zurückgreift, ist sie auf beide Ferromagnettypen anwendbar. Ausgangspunkt der Molekularfeldnäherung ist die Annahme, daß auf jeden Dipol neben dem äußeren Feld $\boldsymbol{B}_\mathrm{a}$ zusätzlich noch das starke *Molekularfeld* $\boldsymbol{B}_\mathrm{M}$ wirkt, das proportional zur Magnetisierung $\boldsymbol{M}$ ist. Das Molekularfeld ist aber kein reales Magnetfeld, sondern ersetzt in der Beschreibung die (nicht magnetische) Wechselwirkung des Aufatoms mit allen anderen Atomen. Damit wird das komplexe Vielteilchenproblem auf ein (modifiziertes) Einspinproblem reduziert, nämlich auf das Verhalten eines herausgegriffenen Moments in dem mittleren „Feld", das von den Nachbarn hervorgerufen wird. Für das effektive Feld $\boldsymbol{B}_\mathrm{eff}$ machen wir den Ansatz

$$\boldsymbol{B}_\mathrm{eff} = \boldsymbol{B}_\mathrm{a} + \boldsymbol{B}_\mathrm{M} = \boldsymbol{B}_\mathrm{a} + \lambda\mu_0\boldsymbol{M} \; . \tag{10.9}$$

Die *Molekularfeldkonstante* λ wird als temperaturunabhängig angenommen. Lokale Fluktuationen der Magnetisierung werden in dieser Beschreibung außer acht gelassen, obwohl sie in der Nähe des Übergangs von der ferro- zur paramagnetischen Phase überaus wichtig sind.

Ziel unserer theoretischen Betrachtungen ist die Beschreibung der spontanen Magnetisierung $M_\mathrm{s}(T)$. Wie in Abb. 10.7 anhand experimenteller Daten für Nickel und Eisen zu sehen, ist die Magnetisierung bei tiefen Temperaturen nahezu konstant und fällt zur Übergangstemperatur oder *ferromagnetischen Curie-Temperatur* T_c hin steil ab. Oberhalb

von T_c verschwindet die spontane Magnetisierung, und es liegt die magnetisch ungeordnete paramagnetische Phase vor.

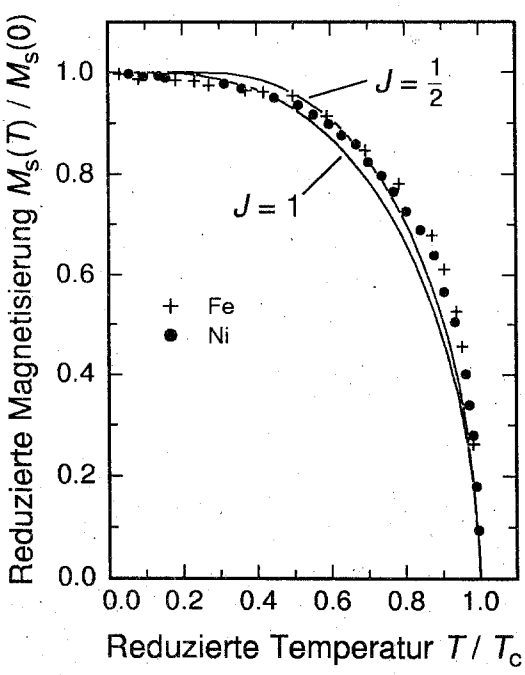


Abb. 10.7: Spontane Magnetisierung von Nickel und Eisen als Funktion der Temperatur, aufgetragen in normierten Einheiten. Die durchgezogenen Kurven erhält man in der Molekularfeldnäherung für die Werte $J = 1/2$ und $J = 1$. (Nach P. Weiss und R. Forrer, Ann. Phys. **5** (1926) 153; H. Potter, Proc. Roy. Soc. (London) **A 146** (1934) 362).

Weiterhin erwarten wir von der Theorie Aussagen über den Zusammenhang zwischen den experimentellen Parametern, die den Temperaturverlauf der Magnetisierung und die Curie-Temperatur bestimmen. Die zentrale Größe dabei ist die Molekularfeldkonstante λ. Um diese zu ermitteln, nutzen wir die Kenntnis, daß das Molekularfeld auch in der paramagnetischen Phase wirksam ist und das äußere Feld verstärkt. Wir modifizieren deshalb das Curie-Gesetz (10.8), um diesen Rückkoppeleffekt zwischen effektivem Feld und Magnetisierung einzubauen. Anstelle von $\mu_0 M = C\mathcal{B}/T$ schreiben wir

$$\mu_0 M = \frac{C}{T}\,(\mathcal{B}_a + \mathcal{B}_M) = \frac{C}{T}\,(\mathcal{B}_a + \lambda\mu_0 M)\ . \tag{10.10}$$

Damit erhalten wir unmittelbar das für ferromagnetische Materialien in der paramagnetischen Phase gültige *Curie-Weiss-Gesetz*[5]

$$\chi_p = \frac{\mu_0 M}{\mathcal{B}_a} = \frac{C}{T - \lambda C} = \frac{C}{T - \Theta}\ . \tag{10.11}$$

Ähnlich wie bei den normalen paramagnetischen Substanzen erwarten wir einen steilen Anstieg der Suszeptibilität mit abnehmender Temperatur, die aber schließlich in einer Singularität bei der *paramagnetischen Curie-Temperatur* $\Theta = C\lambda$ endet. In der Molekularfeldnäherung sollte bei Θ der ferromagnetische Phasenübergang stattfinden, d.h. die beiden charakteristischen Temperaturen Θ und T_c sollten identisch sein. In realen Systemen ist Θ jedoch stets größer als T_c.

[5] Eine entsprechende Modifikation des Curie-Gesetzes haben wir bereits bei der Diskussion der paramagnetischen Materialien eingeführt. Hier wie dort spiegelt die Modifikation die Wechselwirkung der magnetischen Momente wider.

In Abb. 10.8 ist die inverse Suszeptibilität von Nickel in der paramagnetischen Phase wiedergegeben. Bei hohen Temperaturen, also weit entfernt vom Phasenübergang, ist die Übereinstimmung mit (10.11) sehr gut, doch in der Nähe der Curie-Temperatur treten deutliche Abweichungen hiervon auf.

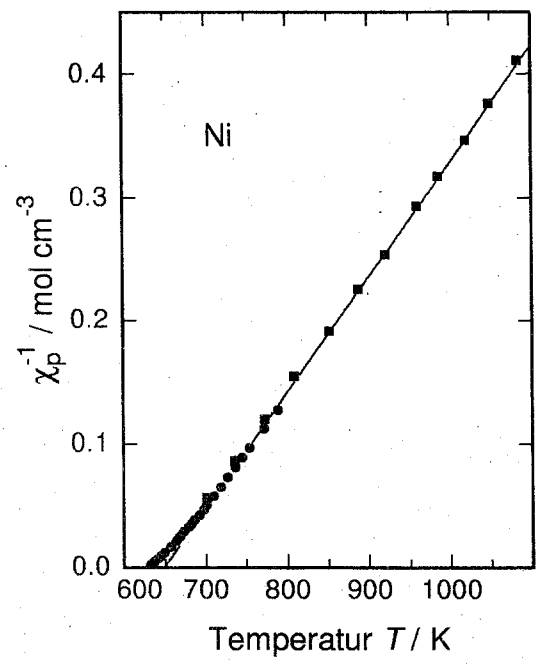


Abb. 10.8: Inverse Suszeptibilität χ_p^{-1} von Nickel in seiner paramagnetischen Phase. Die paramagnetische Curie-Temperatur von Nickel liegt bei $\Theta = 649\,\mathrm{K}$. (Nach P. Weiss und R. Forrer, Ann. Phys. **5** (1926) 153 (Punkte); W. Sucksmith, R.R. Pearce, Proc. Roy. Soc. (London) **A 167** (1938) 189 (Quadrate)).

Von hohen Temperaturen kommend steigt die Suszeptibilität weniger rasch an, als vom Curie-Weiss-Gesetz vorausgesagt. Extrapoliert man die χ_p^{-1}-Kurve von hohen Temperaturen her, so findet man $\Theta = 649$ K. Der tatsächliche Phasenübergang findet aber erst bei $T_c = 630$ K statt. Diese Abweichungen werden vor allem durch lokale Fluktuationen der Magnetisierung verursacht, die bei einem Phasenübergang 2. Ordnung auftreten und in der Molekularfeldnäherung nicht berücksichtigt werden. Theorien zum Phasenübergang sagen in Übereinstimmung mit dem Experiment für den ferromagnetischen Phasenübergang $\chi \propto |T - T_c|^{-\eta}$ voraus, wobei η die Werte 4/3 und 1/3 annimmt, je nachdem, ob man sich dem kritischen Punkt von hohen oder tiefen Temperaturen her nähert[6].

Mit Hilfe von (10.8) läßt sich aus der Übergangstemperatur der Wert der Molekularfeldkonstante abschätzen. Durch Einsetzen der Konstanten C finden wir:

$$\lambda = \frac{\Theta}{C} = \frac{3k_B\Theta}{n\mu_0 p^2 \mu_B^2} \,. \tag{10.12}$$

Wie bereits erwähnt, ist bei den Metallen der Eisengruppe der Beitrag des Bahndrehimpulses zum magnetischen Moment unterdrückt. Es geht nur der Spinbeitrag in die Berechnung der effektiven Magnetonenzahl $p^2 = g^2 S(S+1)$ ein. Mit Hilfe der Daten von Tabelle 10.2 und der Teilchenzahldichte $n \approx 8{,}5 \cdot 10^{28}$ m^{-3} errechnet sich für die Molekularfeldkonstante von Eisen der Wert $\lambda \approx 1000$. Mit der spontanen Magnetisierung von $1{,}75 \cdot 10^6$ A/m ergibt sich damit ein Molekularfeld von $\mathcal{B}_M = \lambda\mu_0 M_s \approx 1000$ T. Diese Feldstärke liegt weit über 30 T, der Feldstärke, die man heute mit Magneten im

[6] In der Molekularfeldnäherung besitzt η die Werte 1 bzw. 1/2.

kontinuierlichen Betrieb erreicht. Berechnet man dagegen das Dipolfeld, das am Ort des Bezugsatoms von benachbarten Spins hervorgerufen wird, so findet man $\mu_0\mu_B/a^3 \approx 0{,}1$ T. Das Molekularfeld ist also wesentlich größer als das Magnetfeld, das die benachbarten magnetischen Momente bewirken. Wie bereits eingangs betont, dient das Molekularfeld der pauschalen Beschreibung der Wechselwirkung zwischen den Atomen. Die Kräfte, die die Ausrichtung der magnetischen Momente bewirken, haben ihren Ursprung aber *nicht* in der magnetischen Wechselwirkung zwischen den Spins.

Tabelle 10.2: Experimentelle Werte der spontanen Magnetisierung M_s bei $T = 0$, Übergangstemperatur T_c und effektive Magnetonenzahl p einiger ferromagnetischer Elemente. (Aus *American Institute of Physics Handbook* (D.W. Gray, ed.) McGraw-Hill, New York 1963).

	M_s (10^6 A/m)	T_c (K)	p
Eisen	1,75	1 043	2,22
Kobalt	1,45	1 404	1,72
Nickel	0,51	631	0,60
Gadolinium	2,00	289	7,12
Terbium	1,44	230	4,95
Dysprosium	2,01	85	6,84
Holmium	2,55	20	8,52

Die Molekularfeldnäherung erlaubt auch eine Beschreibung des Temperaturverlaufs der spontanen Magnetisierung, der in Abb. 10.7 für Nickel und Eisen dargestellt ist. Hierzu ersetzen wir einfach in der Brillouin-Funktion der Gleichung (10.6) das äußere Feld $\mathcal{B}_a$ durch das effektive Feld $\mathcal{B}_{eff}$. Da in der ferromagnetischen Phase $\mathcal{B}_M \gg \mathcal{B}_a$ ist, können wir das äußere Feld ganz vernachlässigen und schreiben

$$M_s = ng\mu_B J B(x) \qquad \text{mit} \qquad x = \frac{gJ\mu_B\lambda\mu_0 M_s}{k_B T} . \tag{10.13}$$

Es ist zu beachten, daß die Magnetisierung nun auch im Argument der Brillouin-Funktion enthalten ist. Der Temperaturverlauf von M_s kann daher nur numerisch bestimmt werden. In Abb. 10.7 ist neben den experimentellen Werten auch der theoretisch erwartete Verlauf für $J = 1/2$ bzw. $J = 1$ eingetragen. Bedenkt man, wie grob die Näherungen der Molekularfeldtheorie sind, so ist die Übereinstimmung durchaus befriedigend.

Betrachtet man den Temperaturverlauf etwas eingehender, so stellt man signifikante Abweichungen nicht nur in der Nähe der Curie-Temperatur, sondern auch bei tiefen Temperaturen fest. Dort beobachtet man experimentell eine wesentlich stärkere Variation der Magnetisierung mit der Temperatur, als die einfache Molekularfeldnäherung vorhersagt. Eine Erklärung hierfür werden wir gegen Ende dieses Abschnitts geben.

Austauschwechselwirkung zwischen lokalisierten magnetischen Momenten

Der Ferromagnetismus gehört zu den Phänomenen der Festkörperphysik, für die noch keine einheitliche mikroskopische Theorie entwickelt werden konnte. Wir müssen uns hier deswegen auf einige grundlegende Aspekte beschränken. Besonders wichtig ist die Beantwortung der Frage, warum sich die Elektronenspins ohne äußeres Feld ausrichten, d.h. warum überhaupt Ferromagnetismus auftritt.

Wir stellen die Behandlung der wohlbekannten Ferromagneten wie Eisen oder Nickel zurück und betrachten zunächst die spontane Ausrichtung von lokalisierten Spins, wie

man sie vor allem bei ferromagnetischen Isolatoren wie EuO, EuS oder $CrBr_3$ antrifft. Das einfachste System, anhand dessen sich das Phänomen der Ausrichtung lokalisierter Spins diskutieren läßt, besteht aus zwei Elektronen. Die wesentlichen Fragen, die dabei auftauchen, haben wir bereits in Abschnitt 3.2 im Zusammenhang mit der chemischen Bindung des Wasserstoffmoleküls behandelt. Abhängig von der Ausrichtung der beiden Spins existieren ein Singulett- und ein Triplettgrundzustand. Deren Energien E_g bzw. E_u unterscheiden sich um den Betrag

$$\mathcal{J} = E_\mathrm{g} - E_\mathrm{u} = 4|\mathcal{N}|^2 \int \psi_a^*(\boldsymbol{r}_1)\,\psi_b^*(\boldsymbol{r}_2)\,H\,\psi_b(\boldsymbol{r}_1)\,\psi_a(\boldsymbol{r}_2)\,\mathrm{d}V_1\,\mathrm{d}V_2\,. \tag{10.14}$$

Die Konstante $\mathcal{N}$ sorgt für die Normierung der Wellenfunktionen. Wie bereits beim Wasserstoffmolekül diskutiert, wird die Energiedifferenz durch das Austauschintegral bestimmt, in das neben dem Hamilton-Operator H nur die Ortswellenfunktionen eingehen. Bei einem positiven Wert des *Austauschkoeffizienten* $\mathcal{J}$ ist die parallele Ausrichtung der Spins energetisch bevorzugt, bei negativem $\mathcal{J}$ die antiparallele. Das Vorzeichen von $\mathcal{J}$ hängt von der Form der Wellenfunktionen und vom Verlauf der potentiellen Energie ab, zu der mehrere Terme beitragen. Betrachten wir zunächst die Abstoßung zwischen den beiden Elektronen im Abstand r_{12}, für die wir $U_\mathrm{ee} = e^2/4\pi\varepsilon_0 r_{12}$ schreiben können. Sie bewirkt immer einen positiven Beitrag zum Austauschkoeffizienten, d.h. die Coulomb-Wechselwirkung zwischen den Elektronen versucht, die Spins auszurichten. Einen weiteren Beitrag zur potentiellen Energie liefert die anziehende Wechselwirkung zwischen den Elektronen und den Ionen. Da sie attraktiv ist, ist ihr Beitrag zu $\mathcal{J}$ negativ und führt so zu einer Energieabsenkung bei antiparallel gerichteten Spins. Je dichter die Atome gepackt sind, um so stärker ist der Überlapp der Wellenfunktionen, und die Energieabsenkung nimmt zu. Welches Vorzeichen der Austauschkoeffizient $\mathcal{J}$ tatsächlich annimmt, hängt von der relativen Größe der entgegengesetzt wirkenden Beiträge ab.

Berücksichtigt man, daß aufgrund des Pauli-Prinzips mit den Ortswellenfunktionen in eindeutiger Weise Spinfunktionen verknüpft sind, so kann man die auftretenden Energien auch mit Hilfe der Spinzustände ausdrücken. Bezeichnen wir mit $\boldsymbol{s}_1$ und $\boldsymbol{s}_2$ die Spinoperatoren der beiden Elektronen, so gilt für den Gesamtspin $\boldsymbol{S}$ die Beziehung $\boldsymbol{S}^2 = |\boldsymbol{s}_1 + \boldsymbol{s}_2|^2 = 3/2 + 2\,\boldsymbol{s}_1 \cdot \boldsymbol{s}_2$. Der Operator $\boldsymbol{s}_1 \cdot \boldsymbol{s}_2$ nimmt die Eigenwerte $-3/4$ für den Singulettzustand mit $S = 0$ und $+1/4$ für den Triplettzustand mit $S = 1$ an. Dieses Ergebnis nutzen wir, um einen modifizierten Hamilton-Operator zur Beschreibung der beiden Zustände zu formulieren, der nur auf die Spinfunktion der beiden Elektronen wirkt:

$$H_\mathrm{Spin} = \frac{1}{4}(E_\mathrm{g} + 3E_\mathrm{u}) - (E_\mathrm{g} - E_\mathrm{u})\,\boldsymbol{s}_1 \cdot \boldsymbol{s}_2 = \frac{1}{4}(E_\mathrm{g} + 3E_\mathrm{u}) - \mathcal{J}\,\boldsymbol{s}_1 \cdot \boldsymbol{s}_2\,. \tag{10.15}$$

Durch Einsetzen der Eigenwerte des Operators $\boldsymbol{s}_1 \cdot \boldsymbol{s}_2$ kann man sich leicht davon überzeugen, daß dieser Hamilton-Operator die Energieeigenwerte E_g für den Singulett- und E_u für den Triplettzustand ergibt. Formal hängen die Eigenwerte nur noch von der Spinrichtung ab. Der Term $(E_\mathrm{g} + 3E_\mathrm{u})/4$ in Gleichung (10.15) ist für die weiteren Überlegungen unwichtig, denn er verschwindet bei geeigneter Wahl des Energienullpunkts, während der zweite Term die Energiedifferenz zwischen den unterschiedlichen Spinausrichtungen festlegt[7]. Die Einführung dieses Hamilton-Operators mag hier etwas künstlich erscheinen

[7] Viele Autoren schreiben für den letzten Teil des Hamilton-Operators auch $-2\mathcal{J}\,\boldsymbol{s}_1 \cdot \boldsymbol{s}_2$, wobei dann $\mathcal{J}$ die halbe Energiedifferenz widerspiegelt.

und ist auch beim vorliegenden Problem aufwendig, doch läßt sich diese Vorgehensweise, wie wir gleich sehen werden, in nützlicher Weise verallgemeinern.

Die Gleichung (10.15) läßt sich auf beliebige Spinoperatoren $\boldsymbol{S}_i$ und $\boldsymbol{S}_j$ und unterschiedliche Austauschkoeffizienten $\mathcal{J}_{ij}$ übertragen. Dies ist der Ausgangspunkt des sogenannten *Heisenberg-Modells* des Ferromagnetismus, in dem der spinabhängige Hamilton-Operator folgendes Aussehen hat:

$$H_{\text{Spin}} = -\sum_i \sum_{j \neq i} \mathcal{J}_{ij} \boldsymbol{S}_i \cdot \boldsymbol{S}_j \,. \tag{10.16}$$

Die Summation erfolgt über alle vorhandenen Atome i und alle Nachbarn j. Eine strenge theoretische Begründung dieses Hamilton-Operators ist nicht möglich. Der Heisenberg-Operator ist nichtlinear; die Durchführung von Rechnungen in konkreten Fällen ist daher trotz seines einfachen Aussehens selbst bei groben Näherungen nur begrenzt möglich[8].

Wir wollen nun noch den Zusammenhang zwischen der Molekularfeldkonstanten λ und dem Austauschkoeffizienten $\mathcal{J}$ herstellen. Hierzu berechnen wir die potentielle Energie, die ein Spin $\boldsymbol{S}_i$ in der Molekularfeldnäherung besitzt, und vergleichen sie mit der Austauschenergie. Der Einfachheit halber nehmen wir an, daß die Probe nur aus gleichartigen Atomen besteht und nur Wechselwirkungen mit den z nächsten Nachbarn auftreten. In der Molekularfeldtheorie besitzt ein Spin mit dem magnetischen Moment $\boldsymbol{\mu}$ die potentielle Energie

$$U = -\boldsymbol{\mu} \cdot \boldsymbol{B}_{\text{eff}} = g\mu_{\text{B}} \boldsymbol{S}_i \cdot \boldsymbol{B}_{\text{eff}} \,. \tag{10.17}$$

Nun ermitteln wir die Energie U mit Hilfe von Gleichung (10.16). Bei der Berechnung der Wechselwirkungsenergie ersetzen wir den Operator $\boldsymbol{S}_j$ durch seinen Erwartungswert $\langle \boldsymbol{S}_j \rangle$. Diese Vorgehensweise ist äquivalent zu der in der Molekularfeldtheorie, in der $\boldsymbol{B}_{\text{eff}}$ für die mittlere Stärke der Wechselwirkung steht. Berücksichtigen wir, daß die z nächsten Nachbarn, die mit dem Spin $\boldsymbol{S}_i$ wechselwirken, identisch sind, so können wir $U = -z\mathcal{J}\boldsymbol{S}_i \cdot \langle \boldsymbol{S}_j \rangle$ schreiben. Weiter nutzen wir den Zusammenhang zwischen $\langle \boldsymbol{S}_j \rangle$ und der Magnetisierung $\boldsymbol{M} = -ng\mu_{\text{B}} \langle \boldsymbol{S}_j \rangle$ und finden so für die Austauschenergie:

$$U = g\mu_B \boldsymbol{S}_i \cdot \frac{z\mathcal{J}}{ng^2\mu_{\text{B}}^2} \boldsymbol{M} \,. \tag{10.18}$$

Vernachlässigen wir in Gleichung (10.9) das äußere Feld, dann kann $\boldsymbol{B}_{\text{eff}} \approx \lambda\mu_0 \boldsymbol{M}$ gesetzt werden, und wir erhalten durch Vergleich der beiden Gleichungen (10.17) und (10.18) für die Molekularfeldkonstante den Ausdruck

$$\lambda = \frac{z\mathcal{J}}{\mu_0 n g^2 \mu_{\text{B}}^2} \,. \tag{10.19}$$

Der Zusammenhang zwischen der Molekularfeldkonstanten und der Curie-Temperatur Θ läßt sich unmittelbar über Gleichung (10.12) herstellen, und wir finden $\mathcal{J} = 3k_{\text{B}}\Theta/z$: Erwartungsgemäß ist die Austauschenergie vergleichbar mit der thermischen Energie beim Phasenübergang.

Nicht in allen Fällen wechselwirken die magnetischen Ionen direkt miteinander, d.h. es tritt nicht immer ein Überlappen der Elektronenhüllen der beteiligten Ionen auf. Es

[8] Eine sehr bekannte und oft benutzte Vereinfachung des Heisenberg-Modells ist das „Ising-Modell", bei dem nur die z-Komponenten berücksichtigt werden, so daß die Spins nur in zwei Richtungen orientiert sein können.

gibt durchaus Ferromagnete, bei denen die magnetischen Ionen durch ein unmagnetisches Ion voneinander getrennt sind. In diesem Fall wird die Austauschwechselwirkung von dem dazwischenliegenden diamagnetischen Ion vermittelt. Diesen *Superaustausch* findet man in vielen ferromagnetischen Oxiden und Verbindungen der Übergangselemente. Ein bekanntes Beispiel hierfür ist MnO, bei dem das O^{2-}-Ion die Wechselwirkung vermittelt.

Bei den Seltenen Erden wird der sogenannte *indirekte Austausch* beobachtet. Bei diesen Materialien tragen die $4f$-Elektronen die magnetischen Momente, doch ist der Überlapp ihrer Wellenfunktionen gering. Die Kopplung erfolgt über die Leitungselektronen. Das magnetische Moment des Aufatoms richtet die Spins der Leitungselektronen in der Umgebung aus, und diese wiederum orientieren die Momente der benachbarten Ionen. Diese nach *Rudermann, Kittel, Kasuya* und *Yosida* benannte RKKY-Wechselwirkung ist proportional zu $(1/r^3)\cos 2k_{\mathrm{F}}r$, hat also eine lange Reichweite und oszillatorischen Charakter. Abhängig vom Abstand der Ionen bewirkt diese Wechselwirkung daher eine parallele oder antiparallele Ausrichtung der magnetischen Momente.

Austauschwechselwirkung in Bändern

Die oben andiskutierten Konzepte der Austauschwechselwirkung zwischen lokalisierten magnetischen Momenten eignen sich zur Beschreibung des Ferromagnetismus bei vielen Systemen, jedoch nicht bei den bekanntesten ferromagnetischen Substanzen wie Eisen, Kobalt oder Nickel. Dort spielen kollektive Eigenschaften der nichtlokalisierten $3d$-Elektronen in den Bändern eine entscheidende Rolle. Es läßt sich leicht zeigen, daß überraschenderweise bereits im freien Elektronengas die Tendenz zur Ausrichtung der Spins existiert, wenn Korrelationseffekte berücksichtigt werden. Obwohl die Überlegungen von einem derartig vereinfachten System nicht direkt auf reale Ferromagnetika übertragen werden können, wollen wir dennoch einen kurzen Blick auf diesen Effekt werfen, um die grundlegenden Ideen zu erörtern. Wir greifen zwei Elektronen mit parallel ausgerichtetem Spin aus dem freien Elektronengas heraus und stellen, um dem Pauli-Prinzip zu genügen, ihre gemeinsame Wellenfunktion durch eine antisymmetrische Überlagerung von ebenen Wellen dar:

$$\begin{aligned}\Psi(\boldsymbol{r}_1,\boldsymbol{r}_2) &= \mathcal{N}\left(\mathrm{e}^{\mathrm{i}\boldsymbol{k}_1\cdot\boldsymbol{r}_1}\,\mathrm{e}^{\mathrm{i}\boldsymbol{k}_2\cdot\boldsymbol{r}_2} - \mathrm{e}^{\mathrm{i}\boldsymbol{k}_1\cdot\boldsymbol{r}_2}\,\mathrm{e}^{\mathrm{i}\boldsymbol{k}_2\cdot\boldsymbol{r}_1}\right)\\ &= \mathcal{N}\mathrm{e}^{\mathrm{i}(\boldsymbol{k}_1\cdot\boldsymbol{r}_1+\boldsymbol{k}_2\cdot\boldsymbol{r}_2)}\left[1-\mathrm{e}^{-\mathrm{i}(\boldsymbol{k}_1-\boldsymbol{k}_2)\cdot(\boldsymbol{r}_1-\boldsymbol{r}_2)}\right]\,. \end{aligned} \tag{10.20}$$

Der Faktor $\mathcal{N}$ dient hier wieder der Normierung der Wellenfunktion. Die Wahrscheinlichkeit, das Elektron 1 im Volumenelement $\mathrm{d}V_1$ und das Elektron 2 im Volumenelement $\mathrm{d}V_2$ zu finden, ist dann:

$$|\Psi(\boldsymbol{r}_1,\boldsymbol{r}_2)|^2\,\mathrm{d}V_1\,\mathrm{d}V_2 = |\mathcal{N}|^2\left\{1-\cos\left[(\boldsymbol{k}_1-\boldsymbol{k}_2)\cdot(\boldsymbol{r}_1-\boldsymbol{r}_2)\right]\right\}\mathrm{d}V_1\,\mathrm{d}V_2\,. \tag{10.21}$$

Dieses Ergebnis ist bemerkenswert: Die Wahrscheinlichkeit, zwei Elektronen mit gleichem Spin am gleichen Ort zu finden, verschwindet für *beliebige* Wellenvektoren, ohne daß die abstoßende Coulomb-Wechselwirkung berücksichtigt wurde! Greifen wir ein Elektron mit vorgegebener Spinrichtung heraus, so ist die Aufenthaltswahrscheinlichkeit von Elektronen mit der gleichen Spinrichtung in der Nachbarschaft stark reduziert. Die Elektronendichte wird durch diese Korrelation reduziert und damit die Abschirmlänge vergrößert. Diesen Effekt haben wir in Abschnitt 7.2 bei der Diskussion der Coulomb-Abschirmung

nicht berücksichtigt. Elektronen können das Potential der Atomrümpfe nicht so gut abschirmen, wie dies ohne Berücksichtigung der Korrelation der Fall wäre. Die schwächere Abschirmung des Atomrumpfpotentials bewirkt eine Reduktion der Elektronenenergie: Die Parallelstellung der Spins bewirkt eine Absenkung der Energie der freien Elektronen, die von der Elektronenkonzentration abhängt. Eine ausführlichere Rechnung ergibt, daß für $n < (5{,}5\,a_\mathrm{B})^{-3}$ das Elektrongas eigentlich ferromagnetisch sein sollte. Cäsium und auch eine Reihe von metallischen Verbindungen erfüllen diese Bedingung, ohne Ferromagnetismus zu zeigen. Offensichtlich vereinfacht das Konzept des freien Elektronengases zu stark, um realistische Vorhersagen machten zu können.

Im *kollektiven Elektronenmodell* von *Stoner* und *Slater*, das wir hier kurz behandeln wollen, wird dieser Tendenz zur Spinausrichtung dadurch Rechnung getragen, daß jedem Elektronenpaar mit entgegengesetztem Spin ein positiver Beitrag I zur Austauschenergie zugeschrieben wird. Darüberhinaus findet die Tatsache Berücksichtigung, daß sich die Elektronen der Festkörper in Bändern aufhalten. Für die Energie der Elektronen in den Subbändern mit unterschiedlicher Spinstellung bietet sich daher folgender Ansatz an:

$$E_\uparrow(\boldsymbol{k}) = E(\boldsymbol{k}) + I\frac{n_\downarrow}{n} - \mu_\mathrm{B}\mathcal{B}\,, \qquad E_\downarrow(\boldsymbol{k}) = E(\boldsymbol{k}) + I\frac{n_\uparrow}{n} + \mu_\mathrm{B}\mathcal{B}\,. \tag{10.22}$$

Hierbei steht $E(\boldsymbol{k})$ für den Energieeigenwert in der Einelektronnäherung und $E_\uparrow(\boldsymbol{k})$ bzw. $E_\downarrow(\boldsymbol{k})$ für die Energie nach Berücksichtigung der Austauschwechselwirkung. Zusätzlich wurde noch die Verschiebung der Energieniveaus durch die Pauli-Suszeptibilität (vgl. Abschnitt 7.2) mitgenommen. Die Variablen $n_\downarrow$ und $n_\uparrow$ bezeichnen die Dichten der Elektronen mit den angezeigten Spinrichtungen. Der obige Ansatz geht von einer besetzungsabhängigen Energie aus, die wiederum Rückwirkung auf die Besetzung selbst hat. Im Grunde sind die Annahmen ähnlich wie bei der bereits diskutierten Molekularfeldnäherung, denn jedes Spinpaar wirkt auf das gesamte Subband.

Wir verschieben nun den Energienullpunkt um $I/2$ und schreiben $\widetilde{E}(\boldsymbol{k}) = E(\boldsymbol{k}) + I/2$. Für den relativen Überschuß einer Spinsorte führen wir die Abkürzung $r = \left(n_\uparrow - n_\downarrow\right)/n$ ein und erhalten damit anstelle von Gleichung (10.22)

$$E_\uparrow(\boldsymbol{k}) = \widetilde{E}(\boldsymbol{k}) - \frac{Ir}{2} - \mu_\mathrm{B}\mathcal{B}\,, \qquad E_\downarrow(\boldsymbol{k}) = \widetilde{E}(\boldsymbol{k}) + \frac{Ir}{2} + \mu_\mathrm{B}\mathcal{B}\,. \tag{10.23}$$

In unserer einfachen Betrachtung ist die Energieaufspaltung zwischen den beiden Subbändern unabhängig vom Wellenvektor.

Im thermischen Gleichgewicht darf zwischen den beiden Subbändern keine Differenz des chemischen Potentials auftreten, d.h. die beiden Teilbänder besitzen das gleiche, durch die Teilchenzahldichte n festgelegte Fermi-Niveau. Der Gleichung (7.14) folgend schreiben wir für $r = \left(n_\uparrow - n_\downarrow\right)/n$:

$$r = \frac{1}{n}\int \frac{D(E)}{2}\Big\{f[E_\uparrow(\boldsymbol{k})] - f[E_\downarrow(\boldsymbol{k})]\Big\}\mathrm{d}E\,. \tag{10.24}$$

Der Faktor $1/2$ tritt auf, weil die Zustandsdichte auf die Subbänder bezogen wird. Wir entwickeln die beiden Fermi-Funktionen um E_F nach $(Ir/2 + \mu_\mathrm{B}\mathcal{B})$ und brechen die Entwicklung jeweils nach dem linearen Glied ab. Die Beiträge der beiden Fermi-Funktionen lassen sich zusammenfassen, und wir finden

$$r = -\int \frac{D(E)}{2n}\frac{\partial f(\boldsymbol{k})}{\partial \widetilde{E}(\boldsymbol{k})}(Ir + 2\mu_\mathrm{B}\mathcal{B})\,\mathrm{d}E\,. \tag{10.25}$$

Wie in Abschnitt 8.2 nähern wir die Ableitung der Fermi-Funktion durch eine Deltafunktion, d.h. wir setzen $-\partial f/\partial \widetilde{E} \approx \delta(\widetilde{E} - E_\mathrm{F})$ und erhalten

$$r = \int \frac{D(E)}{2n}(Ir + 2\mu_\mathrm{B}\mathcal{B})\,\delta[\widetilde{E}(\boldsymbol{k}) - E_\mathrm{F}]\,\mathrm{d}E = \frac{D(E_\mathrm{F})}{2n}\,(Ir + 2\mu_\mathrm{B}\mathcal{B})\,. \tag{10.26}$$

Mit diesem Ergebnis läßt sich die Magnetisierung $M = rn\mu_\mathrm{B}$ und daraus die Suszeptibilität

$$\chi = \frac{\mu_0\mu_\mathrm{B}^2 D(E_\mathrm{F})}{1 - [ID(E_\mathrm{F})/2n]} = \frac{\chi_\mathrm{P}}{1 - [ID(E_\mathrm{F})/2n]} \tag{10.27}$$

errechnen, wobei χ_P für die Pauli-Spinsuszeptibilität steht, die durch Gleichung (7.29) gegeben ist. Offensichtlich verstärkt der Rückkoppelmechanismus das bereits vorhandene paramagnetische Verhalten des Elektronengases.

Der Stoner-Parameter I läßt sich näherungsweise aus der Elektronenkonfiguration der beteiligten Atome unter Berücksichtigung der Austauschwechselwirkung berechnen. Das Ergebnis einer derartigen Rechnung ist in Abb. 10.9 zusammen mit der auf die Elektronendichte normierten Zustandsdichte $D(E_\mathrm{F})/n$ und dem Produkt $ID(E_\mathrm{F})/n$ für eine Reihe von Elementen dargestellt. Für die Metalle Kalzium, Scandium und Palladium ergeben sich besonders hohe Verstärkungsfaktoren. So findet man für χ/χ_P die Werte 5,4 (Ca), 6,1 (Sc) und 4,5 (Pd). Nach einer Korrektur für die noch vorhandenen Bahnmomente stimmen diese Werte gut mit dem Experiment überein.

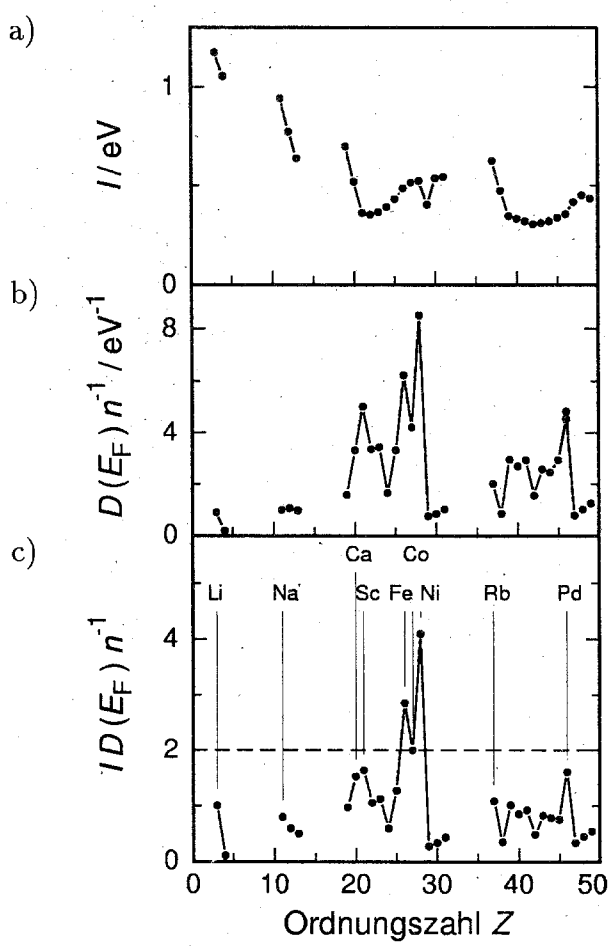


Abb. 10.9: Stoner-Kriterium.
a) Stoner-Parameter I für Elemente bis zur Ordnungszahl $Z = 49$,
b) Normierte Zustandsdichte $D(E_\mathrm{F})/n$ an der Fermi-Kante,
c) Produkt $ID(E_\mathrm{F})/n$. Bei Werten über zwei tritt Ferromagnetismus auf.
(Nach J.F. Janak, Phys. Rev. **B16** (1977) 255).

Mit zunehmendem Wert von $ID(E_F)$ wird der Pauli-Paramagnetismus immer mehr verstärkt. Ist $ID(E_F)/n > 2$, so bildet sich aufgrund der Rückkopplung eine spontane Magnetisierung aus. Ein Blick auf Abb. 10.9 zeigt, daß die geschilderte Rechnung den Ferromagnetismus von Eisen, Kobalt und Nickel tatsächlich richtig voraussagt. Einschränkend soll jedoch bemerkt weren, daß es bisher keine Theorie gibt, die in der Lage ist, die Curie-Temperaturen dieser drei Ferromagneten zu berechnen.

Die Abbildung 10.9c macht auch deutlich, daß die oben erwähnten Metalle Ca, Sc und Pd an der Grenze zum Ferromagnetismus stehen. Dies wird eindrucksvoll durch die verdünnte Legierung **Pd**Fe demonstriert: Bereits die Zugabe von 0,15% Fe führt dazu, daß die Legierung ferromagnetisch wird.

Wir wollen uns noch kurz der Frage zuwenden, welche Voraussetzungen die Bandstrukturen erfüllen müssen, um Ferromagnetismus zu ermöglichen. Die Frage läßt sich anhand der Eigenschaften von Nickel und Kupfer besonders gut diskutieren: Die beiden Metalle sind Nachbarn im Periodensystem, aber nur Nickel ist ferromagnetisch. In den Abbildungen 10.10 und 10.11 sind die berechneten elektronischen Zustandsdichten dargestellt. Der generelle Verlauf ist sehr ähnlich: $4s$- und $3d$-Band überlappen sich, wobei die Unterkante des $4s$-Bandes wesentlich tiefer liegt. Dies hat keine Konsequenz für die magnetischen Eigenschaften von Kupfer mit der atomaren Elektronenkonfiguration $3d^{10}4s^1$, bei dem das $3d$-Band vollständig mit 10 Elektronen pro Atom und das $4s$-Band durch das verbleibende Elektron zur Hälfte gefüllt wird. Das Fermi-Niveau liegt deshalb hoch im s-Band, und die Austauschwechselwirkung verändert diesen Zustand nicht.

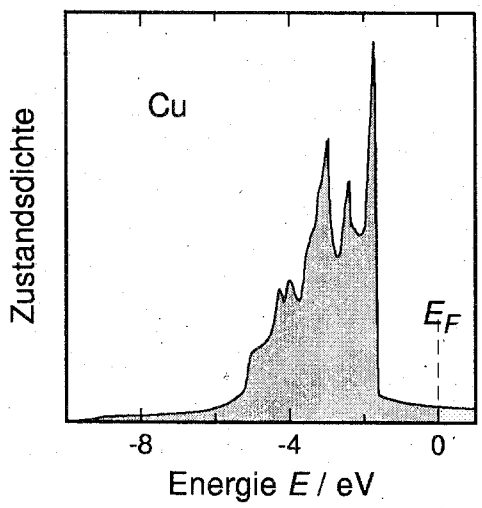


Abb. 10.10: Zustandsdichte $D(E)$ von Kupfer. Das d-Band ist vollständig, das s-Band bis zur Fermi-Energie gefüllt. (Nach P. Steiner und S. Hüfner, Solid State Commun. **44** (1982) 619).

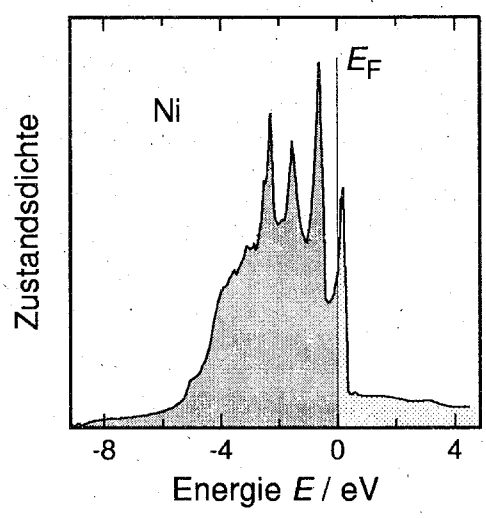


Abb. 10.11: Zustandsdichte von Nickel oberhalb der Curie-Temperatur. Die Fermi-Energie liegt im d-Band. (Nach J. Callaway und C.S. Wang, Phys. Rev. B **7** (1973) 1096).

Wesentlich subtilere Verhältnisse liegen beim Nickel vor, das als freies Atom die Konfiguration $3d^8 4s^2$ besitzt, bei dem aber im Metall nur 0,54 der ursprünglich vorhandenen beiden $4s$-Elektronen im s-Band verbleiben. Das Fermi-Niveau liegt bei Temperaturen oberhalb der Curie-Temperatur in der Nähe der Oberkante des $3d$-Bandes (vgl. Abbildung 10.11). Um das $3d$-Band voll zu besetzen, müßten alle vorhandenen s-Elektronen in dieses Band überwechseln. Da aber das untere Ende des s-Bandes tiefer als das

d-Band liegt, wird es teilweise von Elektronen besetzt, die nun zur vollständigen Besetzung des d-Bandes fehlen. Wird die Curie-Temperatur unterschritten, so verschieben sich die beiden Subbänder mit entgegengesetzter Spinorientierung gegeneinander aufgrund der Austauschwechselwirkung, und die Verhältnisse ändern sich grundlegend. Um diesen Unterschied herauszuarbeiten, betrachten wir, wie in Abb. 10.12 angedeutet, die beiden Subbänder getrennt. Die Austauschwechselwirkung, oder anders ausgedrückt, das Molekularfeld, hebt bzw. senkt die beiden Subbänder um mehr als 0,5 eV. Die Zahl der Elektronen im s-Band wird dadurch nicht verändert, doch ist ein Subband des $3d$-Bandes nun vollständig aufgefüllt, während das andere noch schwächer besetzt ist. Es bildet sich eine spontane Magnetisierung aus, die den 0,54 Löchern pro Atom im $3d$-Band zugeschrieben werden kann.

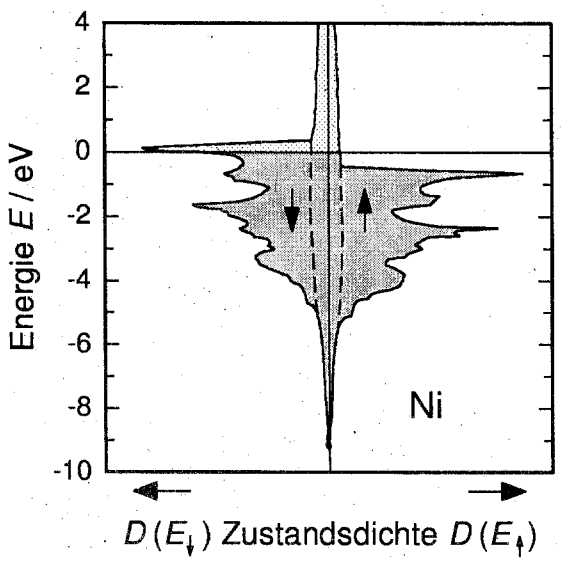


Abb. 10.12: Zustandsdichten $D(E_\uparrow)$ und $D(E_\downarrow)$ der Subbänder von Nickel unterhalb der Curie-Temperatur. Die beiden $3d$-Subbänder sind abhängig von der Spinrichtung vollständig bzw. nur teilweise gefüllt. Die gestrichelte Linie deutet die Zustandsdichte des s-Bandes an. (Nach J. Callaway und C.S. Wang, Phys. Rev. B **7** (1973) 1096).

Dieses Beispiel zeigt einerseits, daß für das Auftreten von Ferromagnetismus eine ganz besondere Konstellation bezüglich der Anzahl der vorhandenen Elektronen und der Bandstruktur vorhanden sein muß. Andererseits können wir daraus ersehen, daß beim Auftreten von Bandferromagnetismus die effektive Magnetonenzahl nicht unmittelbar mit der Spinquantenzahl der Einzelatome korreliert werden kann, selbst wenn kein resultierender Bahndrehimpuls berücksichtigt werden muß.

Spinwellen

In ferromagnetischen Materialien gibt es neben den bisher besprochenen elektronischen Anregungen und Gitterschwingungen noch *Spinwellen* oder *Magnonen*. Zunächst würde man erwarten, daß die energetisch niedrigste magnetische Anregung das Umklappen einzelner Spins sein müßte. Nach Gleichung (10.18) ist der erforderliche Energieaufwand $\delta E = z\mathcal{J}S^2$. Bedenkt man, daß $\mathcal{J} \approx k_B T_c$ ist, so wird verständlich, daß derartige Einspin-Anregungen weit unterhalb der Curie-Temperatur thermisch kaum erzeugt werden. Energetisch günstiger ist die Anregung einer kollektiven Präzessionsbewegung, wie sie in Abb. 10.13 schematisch angedeutet ist. Ihre mathematische Beschreibung kann in völliger Analogie zur Behandlung der Gitterschwingungen erfolgen.

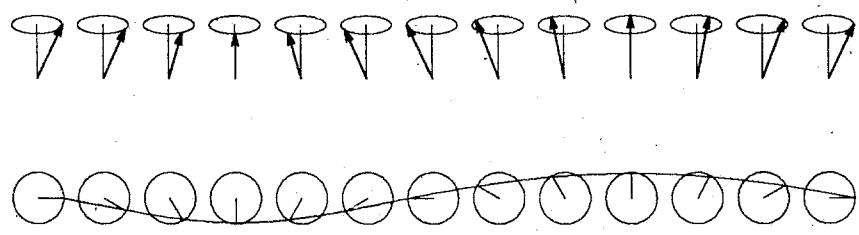


Abb. 10.13: Schematische Darstellung einer Spinwelle (Magnon) längs einer Kette in perspektivischer Darstellung und in Aufsicht.

Wir leiten die Bewegungsgleichung der Spins in einer halbklassischen Betrachtungsweise her, d.h. wir ersetzen die Spinvektoren durch klassische Vektoren. Anschließend quantisieren wir die Auslenkung wie bei den Phononen und ordnen jedem *Magnon* die Energie $E = \hbar\omega$ zu. Der Einfachheit halber betrachten wir statt eines dreidimensionalen Festkörpers eine lineare Kette. Die Plätze im Abstand a sind mit identischen Atomen besetzt, deren Spin $\boldsymbol{S}$ in z-Richtung orientiert ist. Wird ein Spin mit dem magnetischen Moment $\boldsymbol{\mu}$ aus der z-Richtung ausgelenkt, so wirkt auf ihn aufgrund des Molekularfelds $\boldsymbol{B}_\mathrm{M}$ ein Drehmoment $\boldsymbol{\mu} \times \boldsymbol{B}_\mathrm{M}$, das über die Beziehung $\boldsymbol{\mu} = -g\mu_\mathrm{B}\boldsymbol{S}$ mit dem Drehimpuls verbunden ist. Da die zeitliche Ableitung des Drehimpulses $\hbar\boldsymbol{S}_m$ gleich dem angreifenden Drehmoment ist, folgt daraus die Bewegungsgleichung $\mathrm{d}\boldsymbol{S}/\mathrm{d}t = -g\mu_\mathrm{B}\boldsymbol{S} \times \boldsymbol{B}_\mathrm{M}$.

Nun wollen wir noch das Molekularfeld mit Hilfe der Gleichungen (10.16) und (10.17) mit der Austauschenergie in Verbindung bringen. Berücksichtigen wir nur die Wechselwirkung mit den nächsten Nachbarn, so ergibt sich aus (10.16) für die Austauschenergie bzw. für die potentielle Energie U_m des Spins $\boldsymbol{S}_m$ der Ausdruck $U_m = -\mathcal{J}\boldsymbol{S}_m(\boldsymbol{S}_{m-1} + \boldsymbol{S}_{m+1})$. Andererseits läßt sich (10.17) in Abwesenheit eines äußeren Feldes mit $\boldsymbol{B}_\mathrm{eff} = \boldsymbol{B}_\mathrm{M}$ in der Form $U_m = -\boldsymbol{\mu}\mathcal{B}_\mathrm{M} = g\mu_\mathrm{B}\boldsymbol{S}_m\mathcal{B}_\mathrm{M}$ schreiben. Durch Gleichsetzen der beiden Ausdrücke für die Energie folgt schließlich für das Molekularfeld $\boldsymbol{B}_{\mathrm{M},m} = (-\mathcal{J}/g\mu_B)(\boldsymbol{S}_{m-1} + \boldsymbol{S}_{m+1})$. Setzen wir dieses Ergebnis in die oben hergeleitete Bewegungsgleichung ein, so ergibt sich die Beziehung

$$\frac{\mathrm{d}\boldsymbol{S}_m}{\mathrm{d}t} = \frac{\mathcal{J}}{\hbar}\boldsymbol{S}_m \times (\boldsymbol{S}_{m-1} + \boldsymbol{S}_{m+1}) \,. \tag{10.28}$$

Für kleine Auslenkungen in x- und y-Richtung eignet sich ein Lösungsansatz der Form

$$\begin{aligned} S_{m,x} &= A\cos(mqa - \omega t) \,, \\ S_{m,y} &= A\sin(mqa - \omega t) \,, \\ S_{m,z} &= \sqrt{S^2 - A^2} \,, \end{aligned} \tag{10.29}$$

der eine laufende Welle mit der Amplitude A und dem Wellenvektor q vorausetzt. Der Wellenvektor unterliegt den Randbedingungen, die in Abschnitt 5.5 für Phononen diskutiert wurden. Setzt man den Lösungsansatz (10.29) in die Bewegungsgleichung (10.28) ein, so erhält man für die Spinwellen die Dispersionsbeziehung

$$\omega = \frac{2\mathcal{J}S}{\hbar}\Big[1 - \cos(qa)\Big] = \frac{4\mathcal{J}S}{\hbar}\sin^2\left(\frac{qa}{2}\right) \,. \tag{10.30}$$

Für kleine Wellenvektoren können wir näherungsweise

$$\omega \approx \frac{\mathcal{J}S}{\hbar}a^2q^2 \tag{10.31}$$

schreiben, d.h. bei langen Wellenlängen steigt die Frequenz der Spinwellen proportional zu q^2 an; bei den Phononen war dieser Zusammenhang linear.

Führt man diese Rechnung nicht für eine lineare Kette, sondern für einen kubischen Kristall durch, so findet man eine ganz ähnliche Dispersionsrelation:

$$\omega = \frac{\mathcal{J}S}{\hbar} \sum_m \Big[1 - \cos(\boldsymbol{q} \cdot \boldsymbol{r}_m)\Big] \,. \tag{10.32}$$

In dieser Gleichung wird über alle nächsten Nachbarn summiert, wobei die Ortsvektoren $\boldsymbol{r}_m$ das Bezugsatom mit den nächsten Nachbarn verbinden. Im Grenzfall $qa \ll 1$ kommt man zum Ausdruck (10.31) zurück, wobei a nun die Gitterkonstante des kubischen Kristalls ist.

In Abb. 10.14 ist die Dispersionskurve der Spinwellen von Nickel gezeigt, wie sie mit Hilfe der inelastischen Neutronstreuung bestimmt wurde. Bei kleinen Energien findet man den erwarteten q^2-Verlauf. Bei höheren Energien treten Abweichungen auf, da dann auch Einspin-Anregungen, d.h. Umklappen einzelner Spins, möglich wird. Die starke Wechselwirkung zwischen den beiden Anregungstypen verändert den Verlauf der Dispersionskurve bei Wellenvektoren um 0,6 Å^{-1}. Gleichzeitig bewirkt diese Wechselwirkung eine starke Reduktion der Magnonenlebensdauer und somit eine Verbreiterung der Dispersionskurve, die in Abb. 10.14 ebenfalls angedeutet ist.

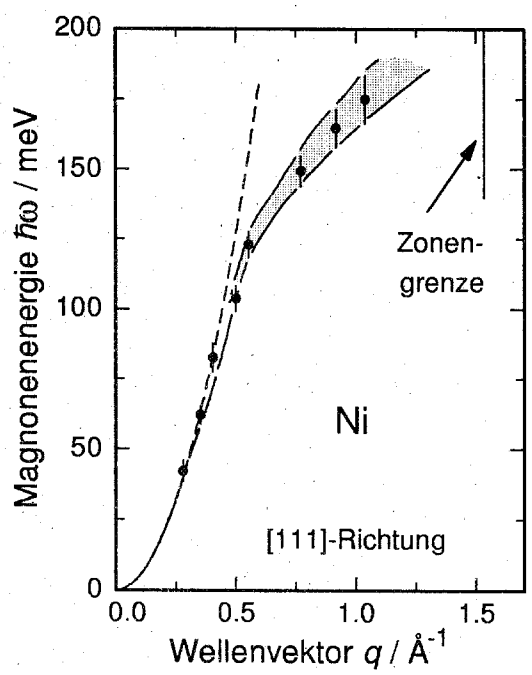


Abb. 10.14: Magnonendispersion in Nickel in [111]-Richtung. Der graue Bereich gibt die Verbreiterung der Dispersionskurve aufgrund der Reduktion der Magnonenlebensdauer an. Die parabelförmige Näherung ist gestrichelt eingezeichnet. (Nach H.A. Mook und D.McK. Paul, Phys. Rev. Lett. **54** (1985) 227).

Thermodynamik der Magnonen

Da Magnonen eine Abweichung der Spins von ihrer perfekten Ausrichtung darstellen, bestimmen sie bei tiefen Temperaturen den Temperaturlauf der spontanen Magnetisierung. Bei $T = 0$ läßt sich die Magnetisierung mit Hilfe von Gleichung (10.13) angeben. Setzen wir $J = S$, so gilt $M_\mathrm{s}(0) = ng\mu_\mathrm{B}S$. Bei endlicher Temperatur sind Magnonen angeregt, die aufgrund ihrer Präzessionsbewegung den effektiven Gesamtspin der Probe erniedrigen. Wie man zeigen kann, reduziert jedes angeregte Magnon unabhängig von seiner Energie $\hbar\omega$ den Gesamtspin um den Betrag $\hbar$ und damit die Magnetisierung gerade um

den Betrag $g\mu_B$. Wir können deshalb den Temperaturverlauf der spontanen Magnetisierung angeben, wenn wir die Zahl N_{mag} der angeregten Magnonen kennen, denn es gilt: $M_s(T) = M_s(0) - (g\mu_B N_{mag}/V)$.

Bei der Berechnung von N_{mag} gehen wir wie bei der Berechnung der Zahl der Phononen in Abschnitt 5.6 vor. Analog zu Gleichung (5.66) ist die Magnonenzahl durch $N_{mag} = \int D(\omega)\langle n(\omega,T)\rangle \mathrm{d}\omega$ gegeben. Hierbei gibt der Bose-Einstein-Faktor $\langle n(\omega,T)\rangle$ die mittlere Besetzung der Magnonenzustände an und $D(\omega)$ ihre Zustandsdichte, die wir noch berechnen müssen. Der Einfachheit halber nehmen wir an, daß die Zustandsdichte isotrop ist. Dann können wir Gleichung (5.52) benutzen, die wir ursprünglich für Phononen in isotropen Festkörpern hergeleitet haben.

Für mäßig hohe Temperaturen sind nur kleine Wellenvektoren von Bedeutung, und wir können die Dispersionsrelation (10.31) benutzen. Damit folgt mit (5.52) für die Zustandsdichte der Ausdruck

$$D(\omega) = \frac{V}{2\pi^2}\,\frac{q^2}{v_g} = \frac{V}{4\pi^2}\left(\frac{\hbar}{\mathcal{J}Sa^2}\right)^{3/2}\sqrt{\omega}\,, \tag{10.33}$$

wobei wir die Gruppengeschwindigkeit v_g eingesetzt haben, die sich aus (10.31) ergibt.

Da bei tiefen Temperaturen nur Spinwellen mit kleinen Energien angeregt sind, können wir bei der Integration die obere Grenze nach Unendlich legen und finden mit $x = \hbar\omega/k_BT$ für die Magnonenzahl

$$N_{mag} = \frac{V}{4\pi^2}\left(\frac{k_BT}{\mathcal{J}Sa^2}\right)^{3/2}\int_0^\infty \frac{\sqrt{x}}{e^x - 1}\,\mathrm{d}x\,, \tag{10.34}$$

wobei das Integral den Wert $4\pi^2 \cdot 0{,}059$ annimmt. Dies führt zum Endergebnis für die Temperaturabhängigkeit der spontanen Magnetisierung:

$$\frac{M_s(0) - M_s(T)}{M_s(0)} = \frac{0{,}059}{nSa^3}\left(\frac{k_BT}{\mathcal{J}S}\right)^{3/2}. \tag{10.35}$$

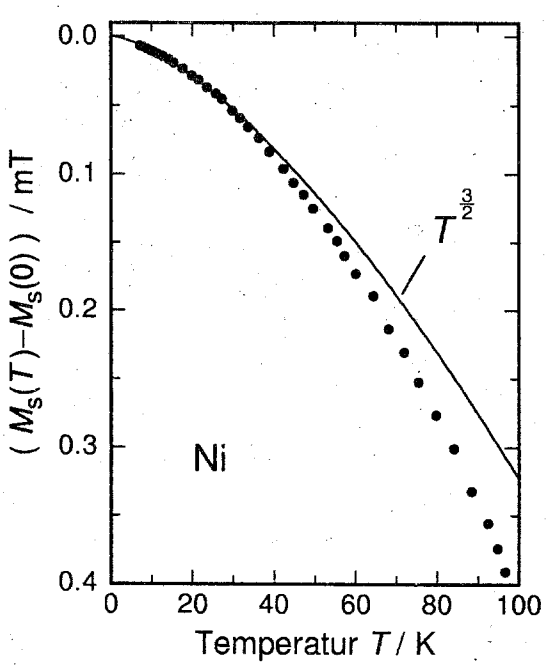


Abb. 10.15: Spontane Magnetisierung von Nickel als Funktion der Temperatur. Bei tiefen Temperaturen folgt die Magnetisierung dem erwarteten $T^{3/2}$-Verlauf. (Nach B.E. Argyle, S.H. Charap, E.W. Pugh, Phys. Rev. **132** (1963) 2051).

Dies ist das *Blochsche $T^{3/2}$-Gesetz* für die Magnetisierung von Ferromagneten, das bei tiefer Temperatur in guter Übereinstimmung mit dem Experiment steht. In Abb. 10.15 werden die experimentellen Daten von Nickel mit der theoretischen Vorhersage verglichen. Die Abweichungen bei höheren Temperaturen rühren in erster Linie von der Verwendung der Näherung (10.31) für Magnonen mit kleinem Wellenvektor her.

Ferromagnetische Domänen

Obwohl ferromagnetische Materialien eine spontane Magnetisierung aufweisen, ist ohne äußeres Feld das Dipolmoment einer makroskopischen Probe in den meisten Fällen zu vernachlässigen. Die Proben enthalten Bereiche, sogenannte *Domänen*, deren Magnetisierung in unterschiedliche Richtungen weist. In Abb. 10.16 sind die Domänengrenzen eines 50 μm breiten Eiseneinkristalls mit Hilfe eines feinen magnetischen Pulvers sichtbar gemacht, das sich an den Stellen der stärksten Magnetfelder ablagert.

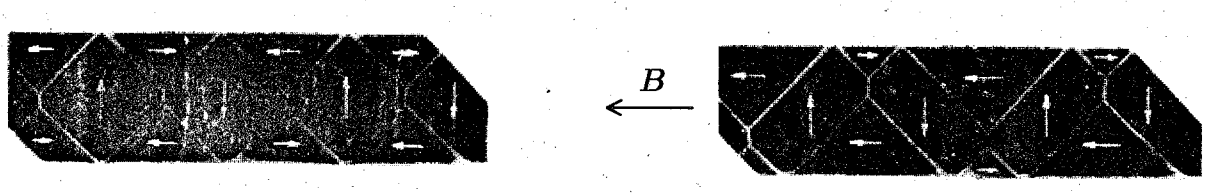


Abb. 10.16: Richtung der spontanen Magnetisierung in den Domänen eines 50 μm breiten Eiseneinkristalls. Die Domänenwände sind mit Hilfe eines feinen magnetischen Pulvers sichtbar gemacht. Das linke Bild wurde ohne, das rechte mit Magnetfeld in angegebener Richtung aufgenommen. (Nach R.W. DeBois C.D. Graham und J. Appl. Phys. **29** (1958) 931).

Es treten ganz bestimmte *Richtungen leichter Magnetisierung* auf, da das Überlappungsintegral und damit die Austauschenergie in realen Kristallen nicht isotrop, sondern richtungsabhängig ist. Bei kubisch-raumzentrierten Kristallen wie Eisen sind die $\langle 100 \rangle$-Richtungen bevorzugt. Die Domänengrenzen sind nicht scharf, denn der Übergang der Magnetisierung von der Richtung der einen zur Richtung der anderen Domäne erfolgt stetig innerhalb eines engen Bereiches. In diesem Bereich, der sogenannten *Bloch-Wand*, ist die Richtung eines Spins gegenüber dem benachbarten jeweils ein wenig gedreht. So wechselt die Magnetisierung innerhalb einiger hundert Spins ihre Richtung. Beim Anlegen eines äußeren Feldes bildet sich durch Verschieben der Bloch-Wände eine makroskopische Magnetisierung aus. Dieser Vorgang ist in Abb. 10.16 deutlich zu erkennen: Die ursprünglich vorhandenen Domänen, die in Feldrichtung zeigten, sind auf Kosten der anderen gewachsen.

Magnetische Domänen entstehen, weil sie die magnetische Feldenergie reduzieren. Dies kann man sich leicht anhand von zwei Stabmagneten klar machen. Deren Energie ist viel geringer, wenn zwei entgegengesetzte Pole und nicht die gleichen benachbart sind. Eine Ausrichtung der Domänen, wie sie in Abb. 10.16 zu sehen ist, minimiert daher die magnetische Feldenergie. Andererseits kostet die Ausbildung von Bloch-Wänden Energie, selbst wenn der Winkel zwischen benachbarten Spins sehr klein ist, die Bloch-Wände also sehr dick sind. Der Grund hierfür ist, daß die Spins der Wand nicht in Richtung der leichten Magnetisierung zeigen. Die freie Energie der Probe wird nicht nur durch die Bildung von Domänen minimiert, es stellt sich auch ein optimaler Winkel zwischen den Richtungen der benachbarten Spins und somit eine bestimmte Wandstärke ein.

10.3 Ferri- und Antiferromagnetismus

Bei ferri- und antiferromagnetischen Materialien sind die magnetischen Momente in (mindestens) zwei Untergittern angeordnet, deren Spinausrichtung entgegengesetzt ist. Ein Festkörper ist antiferromagnetisch, wenn sich die magnetischen Momente der Untergitter gerade kompensieren und ferrimagnetisch, wenn dies nur teilweise der Fall ist. Unterhalb einer kritischen Temperatur treten spontan magnetisierte Domänen auf, oberhalb dieser Temperatur ist diese Substanzklasse paramagnetisch.

Ferrimagnetismus

Der Name „Ferrimagnetismus" kommt von der Bezeichnung *Ferrite* für Materialien mit der Zusammensetzung $MO \cdot Fe_2O_3$. Hierbei steht M für ein zweiwertiges Metall, wie z.B. Kupfer, Blei, Magnesium, Kobalt oder auch Eisen. Im letzten Fall handelt es sich um Magnetit Fe_3O_4, in dem das Eisen zwei- und dreiwertig vorkommt. Die Ferrite besitzen die Struktur des Minerals Spinell ($MgAl_2O_4$) mit einer relativ kompliziert aufgebauten kubischen Elementarzelle, die je nach der chemischen Zusammensetzung eine Ausdehnung von 6 bis 9 Å aufweist. Die Ionen der beiden Metalle sitzen in der Elementarzelle auf nicht äquivalenten Gitterplätzen. Ihr Spin ist bei tiefen Temperaturen parallel zu den Würfelkanten ausgerichtet.

Ferrimagnetismus tritt auf, wenn die Austauschenergie zwischen zwei magnetischen nächsten Nachbarn A und B und damit die Molekularfeldkonstante λ_{AB} negativ ist, denn dann sind die Spins der einen Metallionensorte denen der anderen Ionensorte entgegengerichtet. In vielen Fällen ist auch die Austauschenergie zwischen den Ionen *eines* Untergitters negativ. Doch gilt meist $\lambda_{AB} \gg \lambda_A, \lambda_B$, so daß die Spins innerhalb eines Untergitters dennoch in die gleiche Richtung weisen. Wir wollen nun kurz untersuchen, wie die Magnetisierung von der Temperatur abhängt. Die Gesamtmagnetisierung setzt sich aus den Beiträgen der Untergitter A und B zusammen, d.h. $M = M_A + M_B$. Das Molekularfeld $\mathcal{B}_A$ am Ort eines A-Ions wird in erster Linie durch die Magnetisierung der benachbarten B-Ionen hervorgerufen. Der Einfachheit halber vernachlässigen wir den Beitrag der weiter entfernten A-Ionen, der natürlich ebenfalls in Rechnung gestellt werden müßte. Es wird dann $\mathcal{B}_A = \mathcal{B}_a - \lambda\mu_0 M_B$ und ein entsprechender Ausdruck für $\mathcal{B}_B$ gelten. Damit können wir die Magnetisierung der Untergitter mit Hilfe von Gleichung (10.13) darstellen. Wir vernachlässigen das äußere Feld $\mathcal{B}_a$ und finden

$$M_A = n_A g_A \mu_B S_A B(x_B) \qquad \text{und} \qquad M_B = n_B g_B \mu_B S_B B(x_A) \;, \tag{10.36}$$

wobei $x_B = -g_A \mu_B S_A \lambda \mu_0 M_B / k_B T$ ist und ein entsprechender Ausdruck für x_A gilt. Die Gesamtmagnetisierung läßt sich numerisch berechnen, und man findet einen Temperaturverlauf, der dem von Ferromagneten sehr ähnlich ist.

Oberhalb der Curie-Temperatur Θ verschwindet die spontane Magnetisierung und die Probe ist paramagnetisch. In dieser Phase gelten Gleichungen, die analog zu (10.10) sind:

$$M_A = \frac{C_A}{\mu_0 T}(\mathcal{B}_a - \lambda\mu_0 M_B) \;, \qquad M_B = \frac{C_B}{\mu_0 T}(\mathcal{B}_a - \lambda\mu_0 M_A) \;, \tag{10.37}$$

wobei C_A und C_B die Curie-Konstanten der Untergitter sind. Hieraus ergibt sich für die Suszeptibilität χ_p in der paramagnetischen Phase der Ausdruck

$$\chi_p = \frac{\mu_0 M}{\mathcal{B}_a} = \frac{(C_A + C_B)T - 2\lambda C_A C_B}{T^2 - \Theta^2} . \tag{10.38}$$

Die Curie-Temperatur ist durch $\Theta = \lambda\sqrt{C_A C_B}$ gegeben. Für hohe Temperaturen geht Gleichung (10.38) asymptotisch in das Curie-Weiss-Gesetz (10.11) über. Eine kleine algebraische Umformung zeigt, daß die paramagnetische Curie-Temperatur negativ ist. Wie bei den Ferromagneten wird auch hier in der Molekularfeldnäherung nicht zwischen ferrimagnetischer und paramagnetischer Curie-Temperatur unterschieden, obwohl experimentell durchaus unterschiedliche Werte beobachtet werden.

Antiferromagnetismus

Kompensiert sich die Magnetisierung der beiden Untergitter exakt, so spricht man von Antiferromagnetismus. Obwohl bei tiefen Temperaturen keine spontane Magnetisierung auftritt, sind die Spins wohlgeordnet. Bei der kritischen Temperatur, die hier als *Néel-Temperatur* bezeichnet wird, verschwindet diese Ordnung und die Substanz wird paramagnetisch. Eine einfache antiferromagnetische Substanz ist das tetragonale Manganfluorid, dessen Struktur in Abb. 10.17 zu sehen ist. Die beiden gleichwertigen ferromagnetischen Manganuntergitter A und B sind antiparallel orientiert.

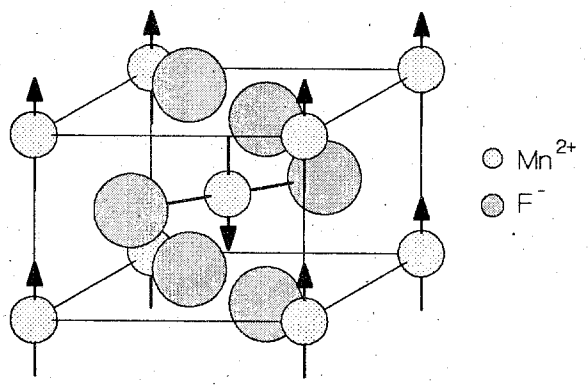


Abb. 10.17: Elementarzelle des antiferromagnetischen MnF_2. Der Übersichtlichkeit halber sind die Ionen verkleinert dargestellt. Die Spinrichtungen sind durch Pfeile angedeutet.

Die Néel-Temperatur errechnen wir wie bei den Ferrimagneten aus der paramagnetischen Phase mit den Gleichungen (10.36) und (10.37), in denen wir $C_A = C_B = C/2$ setzen. Der Faktor 1/2 tritt auf, weil wir die Konstante auf *alle* Atome mit magnetischen Momenten und nicht nur auf die Atome eines Untergitters beziehen. Durch Einsetzen finden wir mit der Néel-Temperatur $T_N = \lambda C/2$ für die Suszeptibilität

$$\chi_p = \frac{C}{T + T_N} . \tag{10.39}$$

In der paramagnetischen Phase verläuft die Suszeptibilität χ_p wie bei Ferromagneten. Sie divergiert jedoch nicht bei der Néel-Temperatur, da das Vorzeichen im Nenner der Gleichung positiv ist, die Divergenz also erst bei $-T_N$ eintreten sollte.

In Abb. 10.18 ist der Temperaturverlauf der Suszeptibilität von MnF_2 zu sehen. Auffallend ist, daß sie unterhalb der Néel-Temperatur von der Richtung des angelegten Feldes

abhängt. Verläuft das Feld in Richtung der spontanen Magnetisierung der Untergitter, so ist am absoluten Nullpunkt $\chi_{\parallel} = 0$, da die beiden Untergitter starr miteinander verbunden sind. Mit zunehmender Temperatur steigt die Suszeptibilität zunächst exponentiell an, denn thermisch aktiviert können Spins umgeklappt werden und weisen dann bevorzugt in Feldrichtung. Bei $T = T_N$ wird die paramagnetische Suszeptibilität erreicht. Ein Feld senkrecht zur Magnetisierung versucht die beiden Untergitter in Feldrichtung zu drehen. Dieser Effekt hängt kaum von der Temperatur ab, so daß $\chi_{\perp} \approx$ const. ist.

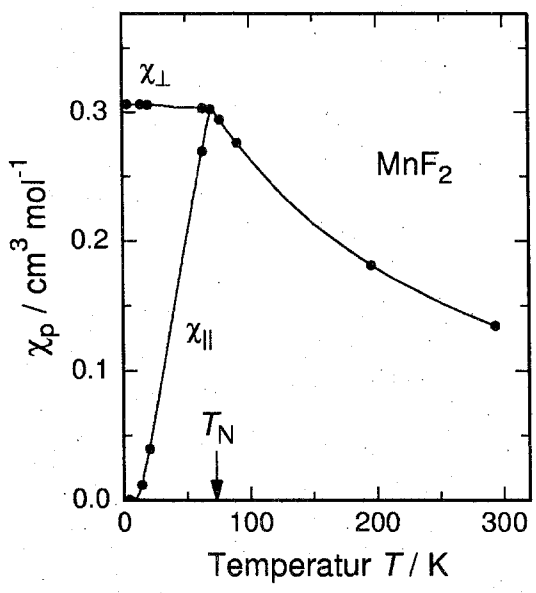


Abb. 10.18: Suszeptibilität von MnF_2. Unterhalb der Néel-Temperatur T_N hängt die Suszeptibilität von der Magnetfeldrichtung ab. (Nach H. Bizette und B. Tsai, Comptes rendus **238** (1954) 1575).

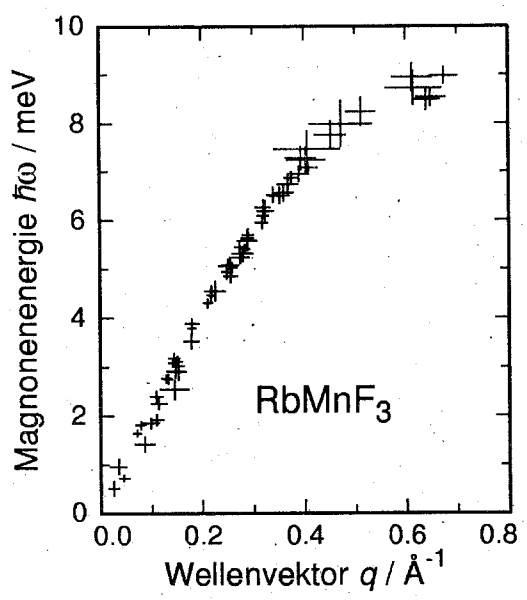


Abb. 10.19: Dispersion der Magnonen im Antiferromagneten $RbMnF_3$, gemessen mit Hilfe der inelastischen Neutronenstreuung. (Nach C.G. Windsor und R.W.H. Stevenson, Proc. Phys. Soc. **87** (1960) 501).

Auch in der antiferromagnetischen Phase existieren Magnonen, die sich jedoch anders als in Ferromagneten verhalten. Eine Ableitung der Bewegungsgleichung der Spins wie bei den Ferromagneten, jedoch mit der zusätzlichen Komplikation, daß zwei gekoppelte Untergitter existieren, führt zur Dispersionsrelation

$$\omega = \frac{2S|\mathcal{J}|}{\hbar} |\sin qa| \,, \tag{10.40}$$

die der Phononendispersionsbeziehung sehr ähnlich sieht. Auch hier besteht bei kleinen Wellenvektoren ein linearer Zusammenhang zwischen Frequenz und Wellenvektor. Dies bestätigen Neutronenstreumessungen am kubischen $RbMnF_3$. In Abb. 10.19 sind Daten dargestellt, die für die Ausbreitung der Magnonen in [100]-, [110]- und [111]-Richtung gefunden wurden. In diesem Kristall unterscheiden sich die Dispersionskurven für die unterschiedlichen Richtungen kaum, so daß hier die Datenpunkte in einem Bild zusammengefaßt wurden.

10.4 Spingläser

In einigen verdünnten Legierungen findet man bei tiefen Temperaturen einen ungewöhnlichen Zustand, den man zwischen dem paramagnetischen und ferromagnetischen Zustand einordnen kann. In diesen Materialen frieren die Spins beim Abkühlen regellos orientiert ein und bilden so ein *Spinglas*. Die prominentesten Vertreter sind **Cu**Mn, **Ag**Mn, **Au**Fe und EuS legiert mit Strontium. Wir wollen das Verhalten der Spingläser anhand von EuS diskutieren, in dem die Austauschwechselwirkung zwischen benachbarten Eu^{2+}-Ionen positiv, mit den übernächsten Nachbarn aber negativ ist. Je nach Entfernung vom herausgegriffenen Spin besteht daher eine Tendenz zur parallelen bzw. antiparallelen Ausrichtung der benachbarten Spins. Bezeichnen wir mit $\mathcal{J}_1$ bzw. $\mathcal{J}_2$ die damit verbundenen Austauschenergien, so gilt bei EuS näherungsweise $\mathcal{J}_1/\mathcal{J}_2 \approx -2$, d.h. die Wechselwirkung mit den nächsten Nachbarn ist etwa doppelt so stark wie mit den übernächsten. Da die Wechselwirkung mit dem nächsten Nachbarn überwiegt, stellt sich in EuS unterhalb von 16,5 K eine ferromagnetische Phase ein. Ersetzt man nun Eu^{2+}-Ionen durch unmagnetische Sr^{2+}-Ionen, so verschiebt sich der Phasenübergang vom paramagnetischen zum ferromagnetischen Zustand zu tieferen Temperaturen hin. Das Phasendiagramm von $Eu_xSr_{1-x}S$ in Abb. 10.20 macht den Verlauf der kritischen Temperatur deutlich. Wie dieser Abbildung weiter zu entnehmen ist, tritt unterhalb von 2 K die oben erwähnte Spinglasphase auf, wenn die Eu^{2+}-Ionen durch Sr^{2+}-Ionen verdünnt sind.

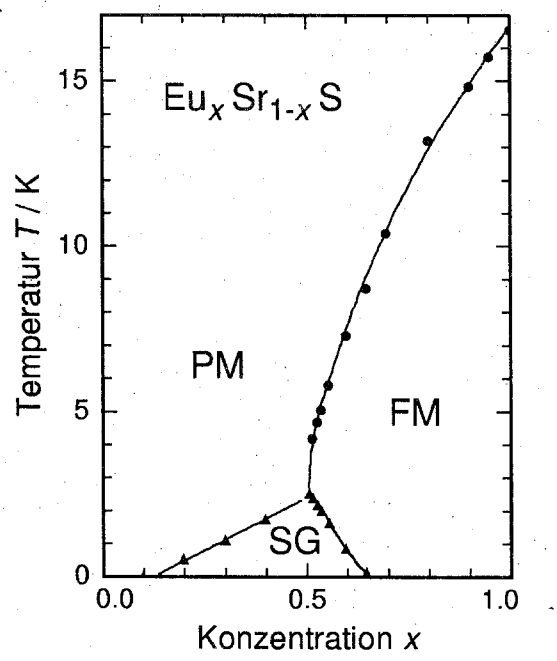


Abb. 10.20: Magnetisches Phasendiagramm von $Eu_xSr_{1-x}S$. Die Abkürzungen PM und FM stehen für die paramagnetische bzw. ferromagnetische Phase und SG für die Spinglasphase. (Nach H. Maletta, J. Appl. Phys. **53** (1982) 2185).

Die Ursache für die Ausbildung dieses Zustands unterhalb der sogenannten *Spinglastemperatur* T_f läßt sich bereits anhand eines zweidimensionalen Modells finden. Zur Vereinfachung der Verhältnisse gehen wir davon aus, daß die Spins nur zwei Einstellmöglichkeiten besitzen. Schränkt man weiter die Reichweite der Wechselwirkung auf die nächsten und übernächsten Nachbarn ein, so läßt sich im Rahmen dieses sogenannten *Ising-Modells*

das Vorzeichen der Austauschenergie, die ein einzelner Spin bei vorgegebenen Austauschkonstanten und vorgegebener Konfiguration der Umgebung besitzt, leicht berechnen. Betrachten wir zunächst die Spinkonfiguration in Abb. 10.21a, bei der die beiden eingerahmten Spins der ferromagnetischen Ordnung entgegengerichtet sind. Dennoch besitzt diese Anordnung bei der vorgegebenen Verteilung der unmagnetischen Ionen den tiefsten Energiewert.

Sind die unmagnetischen Ionen jedoch so wie in den Abbildungen 10.21b und 10.21c angeordnet, ist die Summe der Austauschenergien der beiden Spins unabhängig von ihrer Orientierung. Die beiden dargestellten Konfigurationen sind energetisch entartet. In beiden Konfigurationen wird die optimale Einstellung einer der beiden Spins durch die Anwesenheit des zweiten Spins verhindert. Man spricht in diesem Fall von einer *Frustration* der Wechselwirkungen. Mit zunehmender Konzentration der unmagnetischen Ionen nimmt die Zahl der Konfigurationen zu, die eine einheitliche ferromagnetische Ordnung stören. Bei tiefer Temperatur und genügend hoher Konzentration der Sr^{2+}-Ionen wird daher eine regellos orientierte Konfiguration eingefroren. Dieser Zustand ist oberflächlich betrachtet dem paramagnetischen sehr ähnlich, doch besteht ein wichtiger Unterschied: Im paramagnetischen Zustand wird die Energiedifferenz zwischen den beiden Spinrichtungen durch das Magnetfeld bestimmt und besitzt daher einen wohldefinierten Wert. Die thermische Bewegung reduziert den Ausrichtungsgrad der Spins, die Nachbarn spielen dabei im Idealfall keine, in realen Systemen nur eine untergeordnete Rolle. In Spingläsern dagegen hängt die Energie eines Spins und damit seine Orientierung in komplizierter Weise von der Spinkonfiguration seiner Umgebung ab und wird von der thermischen Bewegung kaum beeinflußt. Die Argumentation ändert sich nicht, wenn man zu einem dreidimensionalen Modell übergeht und weitere Spinorientierungen erlaubt.

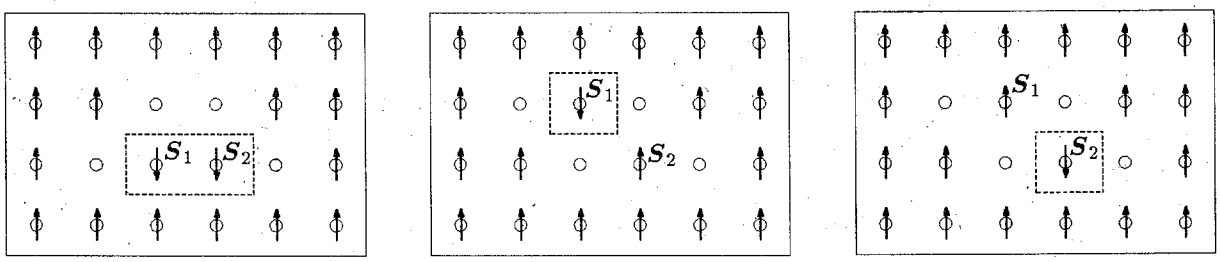


Abb. 10.21: Zweidimensionales Gittermodell für den Frustrationseffekt. Die offenen Kreise repräsentieren die unmagnetischen Ionen. Die Spins der eingerahmten Atome stehen jeweils antiparallel zur ferromagnetischen Ordnung.
a) Die beiden Spins S_1 und S_2 nehmen die energetisch tiefste Lage ein.
b) Die Anwesenheit von Spin S_1 verhindert die anti-ferromagnetische Ausrichtung von Spin S_2.
c) Diese Anordnung ist energetisch äquivalent zur Anordnung b), obwohl sich die Ausrichtung der beiden Spins gedreht hat.

Zum Schluß wollen wir noch kurz das Verhalten der magnetischen Suszeptibilität von Spingläsern schildern. Von hoher Temperatur kommend, folgt die Suszeptibilität der Probe in der paramagnetischen Phase zunächst einem Curie Weiß-Gesetz. Kühlt man die Probe weiter ab, so hängt das Ergebnis von den experimentellen Vorgaben ab. Mißt man die Suszeptibilität mit einem Wechselfeld, so durchläuft sie ein Maximum bei der

Spinglastemperatur T_f. Ein Beispiel hierfür ist in Abb. 10.22 zu sehen. Die offenen Kreise geben eine Wechselfeldmessung an einer Kupferprobe wieder, die mit 1% Mangan dotiert war.

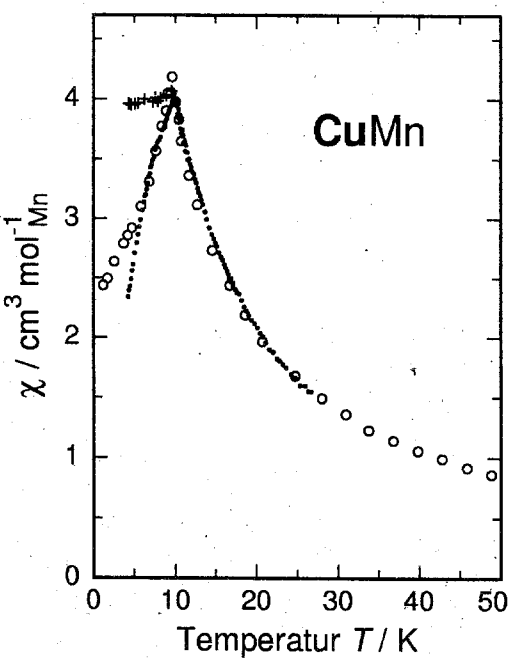


Abb. 10.22: Temperaturverlauf der Suszeptibilität von **CuMn** mit 1% Mangan. Kreise geben die Wechselfeld-, Kreuze die Gleichfelddaten wieder, die während des Abkühlens aufgenommen wurden. Durch Punkte sind die Werte dargestellt, die bei einer Gleichfeldmessung während des Aufwärmens erhalten wurden. (Nach C.A.M. Muller, J.A. van Duyneveldt, J.A. Mydosh, Phys. Rev. **B23** (1981) 1384; S. Nagata, P.H. Keesom, H.R. Harrison, Phys. Rev. **B19** (1979) 1633).

Ermittelt man die Suszeptibilität dagegen durch eine Gleichfeldmessung und führt die Messung von hohen Temperaturen kommend durch, so bleibt χ unterhalb der Spinglastemperatur nahezu konstant. Diese Meßwerte sind durch Kreuze wiedergegeben. Kühlt man dagegen die Probe zunächst bis zur tiefsten Temperatur ohne Feld ab und mißt dann die Suszeptibilität wieder mit einem Gleichfeld, so erhält man, wie die Punkte zeigen, einen ähnlichen Verlauf wie bei der Wechselfeldmessung. Stellt man unterhalb von T_f für längere Zeit eine konstante Temperatur ein, so steigt die Gleichfeldsuszeptibilität langsam an und erreicht schließlich den temperaturunabhängigen Wert der Suszeptibilität. Bei diesem Experiment beobachtet man unterhalb der Spinglastemperatur eine remanente Magnetisierung, die mit der Zeit verschwindet.

Das geschilderte Verhalten der Suszeptibilität läßt sich qualitativ verstehen, wenn man annimmt, daß sich die freie Energie der Spingläser nur dann im Minimum befindet, wenn die Probe im Gleichfeld abgekühlt wurde. Kühlt man dagegen die Probe ohne Feld ab, so wird eine der vielen metastabilen Konfigurationen eingefroren, die aufgrund des Frustrationseffektes existieren. Erst durch eine Änderung der lokalen Struktur an vielen Stellen kann das Spinglas in das absolute Minimum der freien Energie gelangen. Da bei den Übergängen von einer Spinkonfiguration zur anderen Potentialbarrieren zu überwinden sind, können sich neue Konfigurationen, selbst wenn sie energetisch günstiger sind, nur langsam einstellen.

11. Dielektrische und optische Eigenschaften

Unter dem Einfluß eines elektrischen Feldes werden positive und negative Ladungen in entgegengesetzte Richtung getrieben und erzeugen nun ihrerseits ein elektrisches Feld. Wie ein Festkörper auf elektrische Felder reagiert, hängt einerseits davon ab, ob freie oder gebundene Ladungen vorhanden sind und andererseits, auf welche Weise die Feldänderungen erfolgen. So schirmen die freien Metallelektronen statische Felder durch Ausbildung von Oberflächenladungen ab, doch wird dieser Abschirmmechanismus bei hohen Frequenzen unwirksam. Wie wir in Kap. 7.2 gefunden haben, ist die statische Abschirmung nicht bis zu beliebig kleinen Längenskalen möglich, da selbst in Metallen charakteristische Abschirmlängen existieren. In Isolatoren dagegen können Elektronen nur über sehr kleine Distanzen verschoben werden. Es baut sich eine elektrische Polarisation auf, die aber nicht in der Lage ist, statische Felder vollständig abzuschirmen. Generell läßt sich sagen, daß die Reaktion von Isolatoren und Metallen auf elektrische Felder in komplizierter Weise von der Frequenz und der Wellenlänge der Störung, also von der Zeit- und Längenskala abhängt. Verantwortlich hierfür sind eine Vielfalt von dielektrischen und optischen Prozessen, von denen wir die wichtigsten in diesem Kapitel diskutieren wollen. In den ersten drei Abschnitten beschäftigen wir uns mit den dielektrischen und optischen Eigenschaften von Isolatoren, wobei die Ionenkristalle eine herausragende Rolle einnehmen. Im letzten Abschnitt gehen wir auf das Verhalten der freien Elektronen von Metallen ein.

11.1 Lokales Feld und Clausius-Mossotti-Beziehung

Zunächst wollen wir der Frage nachgehen, welches Feld an einem bestimmten Atom eines Isolators herrscht, wenn wir ein äußeres Feld an die Probe anlegen. Diese Frage haben wir im vorausgehenden Kapitel ausgeklammert, da sie bei den elektrischen Eigenschaften wesentlich größere Bedeutung als bei den magnetischen hat. Bringen wir Atome in ein elektrisches Feld, so werden diese polarisiert. Bildlich gesprochen verschiebt sich das Zentrum der Elektronenwolke relativ zum Atomkern. Dadurch wirken die Atome nun ihrerseits als Dipole mit dem Moment

$$\boldsymbol{p} = \epsilon_0[\boldsymbol{\alpha}]\boldsymbol{\mathcal{E}} \; . \tag{11.1}$$

$\boldsymbol{\mathcal{E}}$ ist das elektrische Feld am Ort des betrachteten Atoms und $[\boldsymbol{\alpha}]$ die atomare Polarisierbarkeit [1], die bei Molekülen mit nicht-sphärischer Geometrie eine Tensorgröße ist. In den meisten Fällen gilt dies auch für die Polarisierbarkeit der Atome im Festkörper, da dort der anisotrope Aufbau und gerichtete Bindungen eine Richtungsabhängigkeit bewirken. Wir werden die Polarisierbarkeit jedoch als Skalar behandeln, um die Gleichungen übersichtlich zu halten. Tatsächlich ist diese Vereinfachung für kubische Kristalle und amorphe

[1] Gelegentlich wird die Größe $\epsilon_0\alpha$ als Polarisierbarkeit definiert. Dies hat zur Folge, daß in Gleichungen wie (11.8) oder (11.9) der Faktor ϵ_0 auftritt.

Festkörper berechtigt, da sich diese „elektrisch isotrop" verhalten. Eine Berücksichtigung der Richtungsabhängigkeit in den übrigen Fällen ist aber ohne großen Aufwand möglich.

Wendet man Gleichung (11.1) auf Gase an, so ist $\boldsymbol{\mathcal{E}}$ identisch mit dem angelegten äußeren Feld $\boldsymbol{\mathcal{E}}_\mathrm{a}$. In Festkörpern wirken jedoch aufgrund der kleinen Atomabstände auf ein Atom nicht nur das äußere Feld, sondern auch Felder der benachbarten Atome. Die Summe aus beiden, also das auf ein Atom tatsächlich wirkende Feld, wird als *lokales Feld* $\boldsymbol{\mathcal{E}}_\mathrm{lok}$ bezeichnet. Der Beitrag der Atome bewirkt eine starke Schwankung der elektrischen Feldstärke auf atomarer Skala. Das makroskopische Feld $\boldsymbol{\mathcal{E}}$, das in die Maxwell-Gleichungen eingeht, ist der Mittelwert über die lokal variierenden Felder.

Wir wollen nun die Verknüpfung zwischen lokalem und äußerem Feld herstellen. Hierzu berechnen wir die Polarisation $\boldsymbol{P}$, die ein anliegendes Feld in der mikroskopischen und makroskopischen Beschreibung hervorruft, und vergleichen die beiden Ergebnisse. Sehen wir von Substanzen mit permanenten Dipolen ab, auf die wir noch getrennt eingehen, so weisen alle induzierten Dipolmomente in Feldrichtung und tragen gleichermaßen zur Polarisation bei, d.h. es gilt $\boldsymbol{P} = n\boldsymbol{p} = n\epsilon_0\alpha\boldsymbol{\mathcal{E}}_\mathrm{lok}$. Andererseits wird in der Elektrodynamik die Polarisation mit dem makroskopisch wirksamen Feld über die elektrische Suszeptibilität $[\boldsymbol{\chi}]$ verknüpft, d.h. es gilt $\boldsymbol{P} = \epsilon_0[\boldsymbol{\chi}]\boldsymbol{\mathcal{E}}$. Die Suszeptibilität besitzt tensoriellen Charakter, doch behandeln wir sie hier zur Vereinfachung wie die Polarisierbarkeit als Skalar. Mit dieser Vereinfachung sind makroskopisches und lokales Feld durch die einfache Beziehung $\chi\boldsymbol{\mathcal{E}} = n\alpha\boldsymbol{\mathcal{E}}_\mathrm{lok}$ miteinander verbunden.

Den Zusammenhang zwischen dem makroskopischen Feld $\boldsymbol{\mathcal{E}}$ und dem lokalen Feld $\boldsymbol{\mathcal{E}}_\mathrm{lok}$ werden wir in der *Lorentz-Näherung* herleiten, in der das lokale Feld als die Summe von vier Einzelbeiträgen

$$\boldsymbol{\mathcal{E}}_\mathrm{lok} = \boldsymbol{\mathcal{E}}_\mathrm{a} + \boldsymbol{\mathcal{E}}_\mathrm{D} + \boldsymbol{\mathcal{E}}_\mathrm{L} + \boldsymbol{\mathcal{E}}_\mathrm{K} \tag{11.2}$$

dargestellt wird, deren Bedeutung wir nun getrennt untersuchen.

Bei der folgenden Diskussion gehen wir von einer homogen polarisierten dielektrischen Probe aus, z.B. von einer Scheibe in einem Plattenkondensator. Die Scheibe wird gedanklich in zwei Bereiche zerlegt, einen sehr kleinen, kugelförmigen um das Aufatom herum, in dem die exakte Dipol-Dipol-Wechselwirkung betrachtet wird, und in einen restlichen Bereich, der mit den Mittelwerten der Felder beschrieben wird. In Abb. 11.1 ist dieser Fall schematisch skizziert.

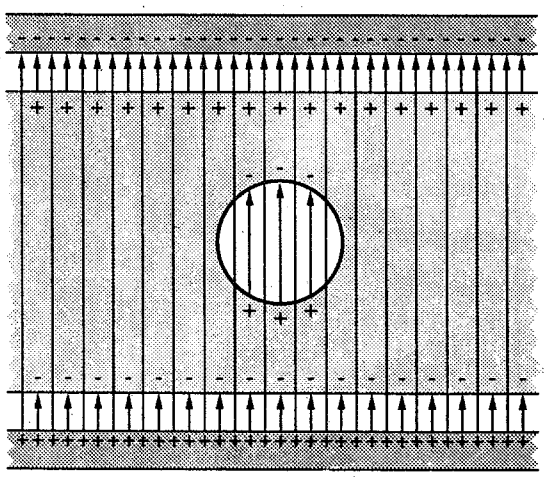


Abb. 11.1: Zur Herleitung des lokalen elektrischen Feldes. Querschnitt durch eine dünne dielektrische Scheibe. Die dunkler gezeichneten Elektroden tragen Ladungen. Das äußere Feld $\boldsymbol{\mathcal{E}}_\mathrm{a}$ im „Luftspalt" wird durch Polarisationsladungen an der Probenoberfläche reduziert. Das Feld im Lochinnern ist größer als das im umgebenden Medium.

Die induzierten Ladungen auf der Probenoberfläche bewirken ein Feld $\boldsymbol{\mathcal{E}}_\mathrm{D} = -f\boldsymbol{P}/\epsilon_0$, das dem äußeren entgegen gerichtet ist und dessen Stärke von der Geometrie der Probe abhängt. Für die dargestellte dünne Scheibe ist der geometrieabhängige Depolarisationsfaktor $f = 1$, bei einer Kugel dagegen $f = 1/3$. Liegt das Feld längs eines dünnen Zylinders oder längs einer dünnen Scheibe, so verschwindet das Gegenfeld, d.h. $f = 0$. Das Auftreten dieses Gegenfeldes ist wohl bekannt: Schiebt man ein Dielektrikum in einen geladenen Kondensator, so fällt die elektrische Feldstärke und damit die Spannung an den Kondensatorplatten ab. Um zu geometrieunabhängigen Aussagen zu kommen, fassen wir das angelegte äußere Feld $\boldsymbol{\mathcal{E}}_\mathrm{a}$ und das *Depolarisations-* oder *Entelektrisierungsfeld* $\boldsymbol{\mathcal{E}}_\mathrm{D}$ zum makroskopisch wirksamen Feld $\boldsymbol{\mathcal{E}} = \boldsymbol{\mathcal{E}}_\mathrm{a} + \boldsymbol{\mathcal{E}}_\mathrm{D}$ zusammen.

Nun kommen wir auf das in Abb. 11.1 skizzierte Loch zurück. Wir stellen uns vor, daß eine kleine Kugel mit dem Radius R aus der Probe herausgeschnitten wurde. Die Kugel soll eine ausreichend große Anzahl von induzierten Dipolen enthalten, damit die Mittelung über die atomaren Felder zum gewünschten makroskopischen Feld führt. Auf der Oberfläche des kugelförmigen Lochs befinden sich Ladungen aufgrund der Polarisation der Scheibe. Bezogen auf die Polarisationsrichtung ist deren Flächenladungsdichte ρ_p bei homogener Polarisation der Probe aus geometrischen Gründen $\rho_\mathrm{p} = -P\cos\theta$. Auf der dunkel getönten Ringfläche der Abb. 11.2 sitzt daher die Ladung $\mathrm{d}q = -P\cos\theta \cdot 2\pi R\sin\theta \cdot R\,\mathrm{d}\theta$. Aus Symmetriegründen ruft die Oberflächenladung nur ein Feld in Polarisationsrichtung hervor. Die Ringfläche trägt zum Feld im Zentrum des Lochs den Teil

$$\mathrm{d}\mathcal{E}_\mathrm{L} = -\frac{1}{4\pi\epsilon_0}\,\frac{\mathrm{d}q}{R^2}\,\cos\theta \tag{11.3}$$

bei. Das resultierende Feld $\mathcal{E}_\mathrm{L}$, das durch die Polarisation auf der Lochoberfläche hervorgerufen wird, erhalten wir durch Integration:

$$\mathcal{E}_\mathrm{L} = \frac{P}{2\epsilon_0}\int_0^\pi \cos^2\theta\,\sin\theta\,\mathrm{d}\theta = \frac{P}{3\epsilon_0}\,. \tag{11.4}$$

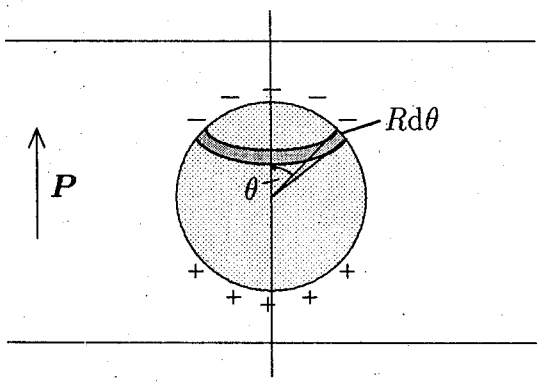


Abb. 11.2: Zur Herleitung des lokalen elektrischen Feldes. Die Ladung auf der Lochoberfläche rührt von der homogenen Polarisation der Probe her. Auf dem dunkler gezeichneten Streifen sitzt die Ladung $\mathrm{d}q$.

Nun müssen wir noch den Beitrag der Nachbaratome in der herausgeschnittenen Kugel zum Feld am Ort des herausgegriffenen Atoms diskutieren. Da sich der Beitrag der benachbarten Atome weitgehend kompensiert, ist dieses Feld $\mathcal{E}_\mathrm{K}$ relativ klein. Es hängt von der Kristallstruktur ab und verschwindet sogar, wenn die benachbarten Atome kubisch

angeordnet sind. Um dies zu zeigen, summieren wir die Beiträge aller atomaren Dipolfelder in der Kugel auf. Nehmen wir an, daß das elektrische Feld in z-Richtung anliegt, so gilt:

$$\mathcal{E}_{\mathrm{K}} = \frac{1}{4\pi\epsilon_0} \sum_m p_m \frac{3z_m^2 - r_m^2}{r_m^5} \,. \tag{11.5}$$

Aus der Symmetrie des kubischen Gitters und der herausgeschnittenen Kugel folgt:

$$\sum_m \frac{x_m^2}{r_m^5} = \sum \frac{y_m^2}{r_m^5} = \sum \frac{z_m^2}{r_m^5} \,. \tag{11.6}$$

Setzen wir diesen Ausdruck in (11.5) ein, so erhalten wir das aus geometrischen Gründen plausible einfache Ergebnis $\mathcal{E}_{\mathrm{K}} = 0$. Bei einer kubischen Anordnung der benachbarten Atome heben sich die Dipolfelder am Ort des Aufatoms gerade auf. Damit bekommen wir als Endergebnis für das lokale Feld in einer kubischen Umgebung die *Lorentz-Beziehung*

$$\boldsymbol{\mathcal{E}}_{\mathrm{lok}} = \boldsymbol{\mathcal{E}} + \frac{\boldsymbol{P}}{3\epsilon_0} \,. \tag{11.7}$$

Obwohl das Feld der herausgeschnittenen Kugel strenggenommen nur für kubische Gitter verschwindet, ist die Lorentz-Beziehung auch in den meisten anderen Fällen brauchbar.

Abhängig von der Form der Probe und dem damit verbundenen Depolarisationsfeld wirkt die Polarisation verstärkend oder schwächend auf das äußere Feld. Steht das elektrische Feld senkrecht auf einer dünnen Scheibe wie in Abb. 11.1, so finden wir für das lokale Feld $\boldsymbol{\mathcal{E}}_{\mathrm{lok}} = \boldsymbol{\mathcal{E}}_{\mathrm{a}} - 2\boldsymbol{P}/3\epsilon_0$. Verläuft es dagegen parallel, so ergibt sich $\boldsymbol{\mathcal{E}}_{\mathrm{lok}} = \boldsymbol{\mathcal{E}}_{\mathrm{a}} + \boldsymbol{P}/3\epsilon_0$. Bei kugelförmigen Proben heben sich $\boldsymbol{\mathcal{E}}_{\mathrm{D}}$ und $\boldsymbol{\mathcal{E}}_{\mathrm{L}}$ gerade auf.

Wir können nun über die Beziehung $\chi\boldsymbol{\mathcal{E}} = n\alpha\boldsymbol{\mathcal{E}}_{\mathrm{lok}}$ in der Lorentz-Näherung auch die hier als isotrop angenommene Suszeptibilität berechnen und finden mit Hilfe von (11.7)

$$\chi = \frac{n\alpha}{1 - n\alpha/3} \,. \tag{11.8}$$

Ohne Berücksichtigung des lokalen Feldes wäre $\chi = n\alpha$. Für die Dielektrizitätskonstante $\epsilon = 1 + \chi$ ergibt sich daraus

$$\frac{\epsilon - 1}{\epsilon + 2} = \frac{n\alpha}{3} \,. \tag{11.9}$$

Diese Verknüpfung zwischen der experimentell ermittelbaren Dielektrizitätskonstanten und der Polarisierbarkeit ist die bekannte *Clausius-Mossotti-Beziehung.* Mit Hilfe dieser Beziehung und Gleichung (11.7) läßt sich bei bekanntem äußeren Feld das lokale Feld errechnen, das am Ort eines herausgegriffenen Atomes tatsächlich herrscht und das beispielsweise bei der Interpretation von Meßdaten die entscheidende Größe ist.

11.2 Elektrische Polarisation

Im vorhergehenden Abschnitt sind wir davon ausgegangen, daß elektrische Felder in Dielektrika die Ausbildung einer Polarisation bewirken, deren Ursache wir nun etwas systematischer untersuchen. Zur Veranschaulichung ist in Abb. 11.3 eine schematische Übersicht des Frequenzgangs der Dielektriziätskonstanten bzw. des Realteils der dielektrischen Funktion eines polaren Kristalls, d.h. eines Kristalls mit ionischer Bindung, skizziert. Es

lassen sich drei Beiträge unterscheiden, die bei wachsender Frequenz jeweils in einem charakteristischen Frequenzbereich ihre Bedeutung verlieren. Moleküle mit permanentem Dipolmoment rufen bei tiefen Frequenzen die *dipolare Polarisation* hervor. Wie wir sehen werden, wird sie spätestens im Mikrowellenbereich unwirksam. Die *Ionenpolarisation* beruht auf der Verschiebung geladener Ionen im elektrischen Feld und besitzt seine stärkste Ausprägung bei den typischen Ionenkristallen. Der ionische Beitrag zur Dielektrizitätskonstanten verschwindet im nahen Infrarotbereich des Spektrums. Bei noch höheren Frequenzen tritt nur noch *elektronische Polarisation* auf, die man als Verschiebung der Elektronenwolke bezüglich der Kerne auffassen kann. Im Röntgengebiet verschwindet schließlich der Beitrag der Atomhüllen, und die Dielektrizitätskonstante nähert sich dem Grenzwert $\epsilon = 1$. Die Gesamtpolarisation ist die Summe der jeweiligen Einzelbeiträge, die wir nun getrennt diskutieren werden.

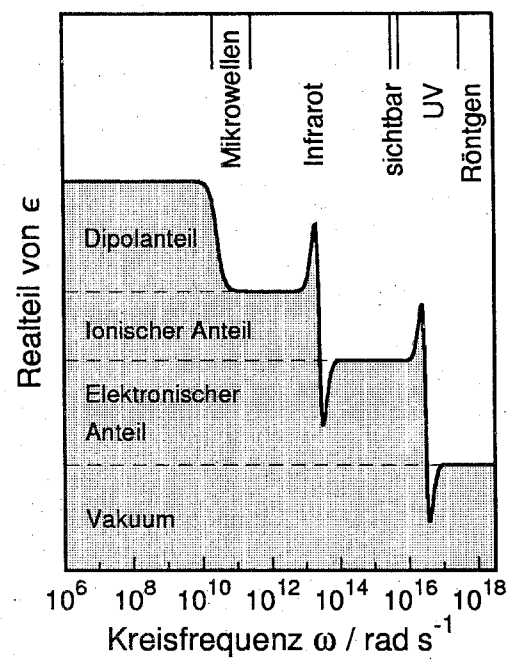


Abb. 11.3: Schematische Darstellung der Frequenzabhängigkeit der Dielektrizitätskonstanten eines polaren Kristalls. Die Größe der einzelnen Beiträge hängt von den spezifischen Eigenschaften des betrachteten Festkörpers ab.

Ionenpolarisation

Die schematische Darstellung 11.3 legt nahe, daß die ionische und die elektronische Polarisation durch Resonanzvorgänge hervorgerufen werden. Zu ihrer Beschreibung ziehen wir ein äußerst einfaches *Oszillatormodell* heran, das sich in der Grundaussage auf beide Effekte anwenden läßt, und beginnen mit der Diskussion der Ionenpolarisation.

Die Wechselwirkung zwischen den Ionen eines Gitters und elektrischen Wechselfeldern bei Frequenzen der Infrarotstrahlung läßt sich auf verschiedene Weisen beschreiben. Vom mikroskopischen, quantenmechanischen Standpunkt handelt es sich um die Wechselwirkung zwischen optischen Phononen und Photonen. Da bei Kristallen der Quasiimpuls erhalten bleibt, können nur Phononen mit sehr kleinen Wellenvektoren daran teilnehmen. Wie wir im Abschnitt 5.4 gesehen haben, verläuft die Dispersionskurve der optischen Phononen in der Nähe des Γ-Punktes waagrecht. Alle an der Wechselwirkung beteiligten Phononen eines Zweiges haben daher in sehr guter Näherung die gleiche Frequenz.

Bei Kristallen mit zweiatomiger Basis, auf die wir uns hier beschränken, erfolgt die Bewegung der Ionen eines Untergitters in Gegenphase zu den Ionen des entgegengesetzt geladenen Untergitters. Zur Beschreibung der optischen Phononen vom makroskopischen Standpunkt aus reicht es daher, ein Ionenpaar herauszugreifen und dessen Reaktion auf die elektromagnetische Welle zu diskutieren, da alle Ionenpaare die gleiche Bewegung ausführen.

Auf die Ionen der beiden Untergitter mit den Ladungen q bzw. $-q$ wirkt neben den elastischen Kräften das lokale elektrische Feld $\mathcal{E}_{\text{lok}}$, das durch die Ionen selbst hervorgerufen wird. Bezeichnen wir mit u_1 und u_2 die Auslenkungen der Ionen mit den Massen M_1 und M_2, so erhalten wir anstelle der Gleichungen, die der Determinanten (5.26) zugrunde liegen, die Ausdrücke

$$\begin{aligned} M_1\ddot{u}_1 + 2\mathcal{C}u_1 - 2\mathcal{C}u_2 &= \quad q\mathcal{E}_{\text{lok}} \,, \\ M_2\ddot{u}_2 + 2\mathcal{C}u_2 - 2\mathcal{C}u_1 &= -q\mathcal{E}_{\text{lok}} \,. \end{aligned} \tag{11.10}$$

Hierbei haben wir $\mathcal{C}' = \mathcal{C}'' = \mathcal{C}$ gesetzt, da sich in zweiatomigen Ionenkristallen die beiden Kraftkonstanten nicht unterscheiden. Weiter wurde die Vektorkennzeichnung weggelassen, weil in unserer einfachen Näherung Auslenkungen und lokales Feld parallel verlaufen. Da wir nur an der Relativbewegung der beiden Untergitter interessiert sind, führen wir die Verschiebung $u = u_1 - u_2$ der Untergitter gegeneinander und die reduzierte Masse μ der Ionenpaare ein und erhalten

$$\mu\ddot{u} + \mu\omega_0^2 u = q\mathcal{E}_{\text{lok}} \,, \tag{11.11}$$

wobei wir die Abkürzung $\omega_0^2 = 2\mathcal{C}/\mu$ für die Resonanzfrequenz benutzen, die das System ohne elektrische Wechselwirkung, also aufgrund der elastischen Eigenschaften besitzt. Diese Bewegungsgleichung ist identisch mit der eines linearen harmonischen Oszillators, auf den eine äußere Störung wirkt. Wir werden die dielektrische Funktion eines Ensembles von derartigen Oszillatoren herleiten, um damit die Ionenpolarisation zu erklären.

In Gleichung (11.11) fügen wir noch einen Dämpfungsterm ein, der pauschal die endliche Lebensdauer der optischen Phononen berücksichtigt. Die Ursache hierfür ist die Wechselwirkung der Phononen untereinander und die Kopplung an elektromagnetische Wellen. Bezeichnen wir die Dämpfungskonstante mit γ, so tritt anstelle von (11.11) die Gleichung

$$\mu\ddot{u} + \mu\gamma\dot{u} + \mu\omega_0^2 u = q\mathcal{E}_{\text{lok}} \,. \tag{11.12}$$

Wirkt das elektrische Wechselfeld $\mathcal{E}_{\text{lok}}(t) = \mathcal{E}_{\text{lok}}^0 \exp(-i\omega t)$ auf die Ladung q, so lautet die stationäre Lösung dieser Differentialgleichung:

$$u(t) = \frac{q}{\mu}\,\frac{1}{\omega_0^2 - \omega^2 - i\gamma\omega}\,\mathcal{E}_{\text{lok}}(t) \,. \tag{11.13}$$

Die Bewegung der Ionen gegeneinander ruft das oszillierende Dipolmoment $p(t) = qu(t)$ hervor. Hinzu kommt noch der elektronische Beitrag zur Polarisation α, von dem wir hier annehmen, daß sein Wert nicht von der Frequenz abhängt. Die Rechtfertigung für diese Annahme werden wir im nächsten Abschnitt finden. In der weiteren Diskussion fassen wir die Beiträge der positiven und negativen Ionen zur elektronischen Polarisierbarkeit $\alpha = \alpha_+ + \alpha_-$ zusammen. Damit läßt sich die Polarisation der Probe durch

$$P(t) = nqu(t) + n\epsilon_0\alpha\mathcal{E}_{\text{lok}}(t) \tag{11.14}$$

ausdrücken. Um die dielektrische Funktion $\epsilon(\omega) = 1 + \chi = 1 + P/\epsilon_0\mathcal{E}$ zu erhalten, eliminieren wir $\mathcal{E}_{\text{lok}}(t)$ und $u(t)$ in Gleichung (11.14) mit Hilfe der Gleichungen (11.7) und (11.13). So finden wir den Ausdruck

$$\epsilon(\omega) = 1 + \frac{n\alpha}{1 - n\alpha/3} + \frac{nq^2}{\epsilon_0\mu}\left(\frac{1}{1 - n\alpha/3}\right)^2 \left[\omega_0^2 - \frac{nq^2}{3\epsilon_0\mu(1 - n\alpha/3)} - \omega^2 - \mathrm{i}\gamma\omega\right]^{-1}, \tag{11.15}$$

der verdeutlicht, daß die Kräfte zwischen den Ionen aufgrund der elektrischen Felder zu einer Verschiebung der Resonanzfrequenz führen, die bei einer rein elastischen Wechselwirkung bei ω_0 liegt. Die neue Resonanzfrequenz ω_t läßt sich direkt der eckigen Klammer entnehmen. Es gilt:

$$\omega_t^2 = \omega_0^2 - \frac{nq^2}{3\epsilon_0\mu}\,\frac{1}{1 - n\alpha/3}\,. \tag{11.16}$$

Die etwas unübersichtliche Gleichung (11.15) läßt sich auf die einfache Form

$$\epsilon(\omega) = \epsilon_\infty + \frac{\omega_t^2(\epsilon_{\text{st}} - \epsilon_\infty)}{\omega_t^2 - \omega^2 - \mathrm{i}\gamma\omega} \tag{11.17}$$

bringen, wenn man neben der neuen Resonanzfrequenz (11.16) noch die Abkürzungen ϵ_∞ und ϵ_{st} einführt, die für die Grenzwerte der dielektrischen Funktion bei hohen und tiefen Frequenzen stehen. Der Index ∞ weist nicht auf unendlich hohe Frequenzen hin, sondern auf Frequenzen, die groß gegen die Frequenz der Ionenresonanz sind, d.h. ϵ_∞ spiegelt die elektronische Polarisierbarkeit wider. Dagegen steht ϵ_{st} für den „statischen Wert" der Dielektrizitätskonstanten, also für ihren Wert bei Frequenzen wesentlich unterhalb der Ionenresonanz. Eine genauere Inspektion der Gleichung (11.15) zeigt, daß der Wert der statischen Dielektrizitätskonstanten anwächst, wenn die Resonanzfrequenz abnimmt. Bei Substanzen ohne permanente Dipolmomente entspricht ϵ_{st} tatsächlich dem Wert der dielektrischen Funktion bei Gleichfeld. Der Faktor $(1 - n\alpha/3)$, der an verschiedenen Stellen von (11.15) und (11.16) auftritt, rührt von der Korrektur für das lokale Feld her.

Nun zerlegen wir die komplexe dielektrische Funktion $\epsilon = \epsilon' + \mathrm{i}\epsilon''$ noch in Real- und Imaginärteil:

$$\epsilon'(\omega) = \epsilon_\infty + \frac{(\epsilon_{\text{st}} - \epsilon_\infty)\,\omega_t^2\,(\omega_t^2 - \omega^2)}{(\omega_t^2 - \omega^2)^2 + \gamma^2\omega^2}\,, \tag{11.18}$$

$$\epsilon''(\omega) = \frac{(\epsilon_{\text{st}} - \epsilon_\infty)\,\omega_t^2\,\gamma\,\omega}{(\omega_t^2 - \omega^2)^2 + \gamma^2\omega^2}\,. \tag{11.19}$$

In Abb. 11.4 ist der schematische Frequenzverlauf dieser beiden Funktionen wiedergegeben. Charakteristisch für resonante Anregungen ist, daß der Imaginärteil ϵ'' nur in der Umgebung der Resonanzfrequenz merklich von Null verschieden ist. In diesem Frequenzbereich treten auch starke Änderungen im Realteil ϵ' auf. ϵ' nimmt zweimal den Wert Null an: in der Nähe von ω_t und bei ω_l. Den Zusammenhang dieser Frequenzen mit den longitudinalen und transversalen optischen Phononen werden wir sogleich diskutieren.

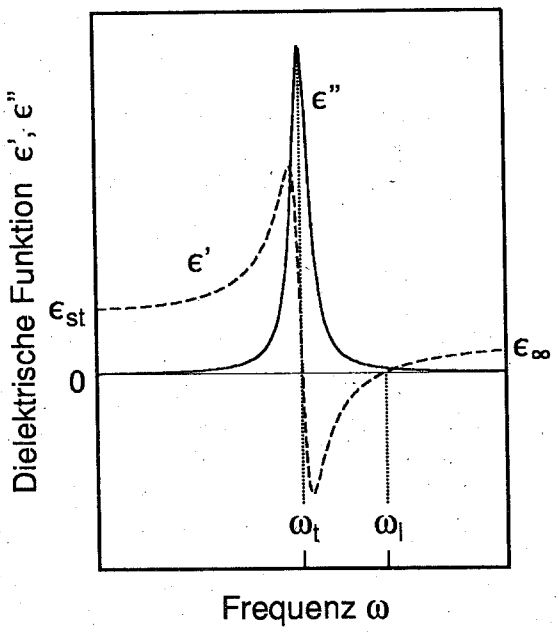


Abb. 11.4: Frequenzgang des Real- und Imaginärteils der dielektrischen Funktion. Die Nullstellen von ϵ' liegen bei ω_t, also in der Nähe des Maximums von ϵ'', und bei der Frequenz ω_l.

Optische Phononen in Ionenkristallen

An dieser Stelle wollen wir uns im Rahmen der Diskussion der dielektrischen Funktion kurz der Frage zuwenden, auf welchen Schwingungstyp sich die Resonanzfrequenz ω_t in Gleichung (11.16) bezieht. Wie wir in Kap. 5 am Beispiel von Diamant (vgl. Abb. 5.24) sahen, besitzen longitudinale und transversale optische Phononen[2] nicht-polarer kubischer Kristalle am Γ-Punkt die gleiche Frequenz. Ein Blick auf die Dispersionskurve von Lithiumfluorid in Abb. 5.15 jedoch lehrt, daß bei Ionenkristallen zwischen ω_l und ω_t eine Lücke klafft, die von den elektrostatischen Kräften zwischen den Ionen verursacht wird.

Wir wollen nun die Phononenfrequenzen ω_l und ω_t berechnen; dabei gehen wir von der Bewegungsgleichung (11.12) aus und vernachlässigen zur Vereinfachung den Dämpfungsterm, der hier nur einen unbedeutenden Einfluß hat. Zunächst müssen wir klären, worin sich longitudinale und transversale Gitterwellen unterscheiden, um diese Kenntnis dann in geeigneter Weise bei unserer Rechnung zu berücksichtigen. Die entscheidenden Punkte sind die unterschiedliche Orientierung der elektrischen Dipole bezüglich des Wellenvektors der Phononen und die damit verbundenen unterschiedlichen mittleren (makroskopischen) Felder.

Greifen wir eine ebene longitudinale optische Gitterwelle heraus, die sich in x-Richtung ausbreitet, so können wir die damit verbundene Polarisation $\boldsymbol{P}_l$ durch die Gleichung $\boldsymbol{P}_l = \hat{\boldsymbol{x}} P_0 \exp[-\mathrm{i}(\omega t - q_l x)]$ ausdrücken. Bilden wir hiervon die Divergenz, so sehen wir, daß $\operatorname{div} \boldsymbol{P}_l \neq 0$ ist. Nach den Maxwell-Gleichungen verschwindet die Divergenz der dielektrischen Verschiebung bei einer neutralen, raumladungsfreien Probe, d.h. es gilt

$$\operatorname{div} \boldsymbol{D} = \operatorname{div} [\epsilon_0 \epsilon(\omega)\, \boldsymbol{\mathcal{E}}] = \epsilon_0 \operatorname{div} \boldsymbol{\mathcal{E}} + \operatorname{div} \boldsymbol{P} = 0 \,. \tag{11.20}$$

Da $\operatorname{div} \boldsymbol{P}$ einen endlichen Wert besitzt, muß auch $\operatorname{div} \boldsymbol{\mathcal{E}} \neq 0$ sein. Gleichung (11.20) kann nur dann Gültigkeit besitzen, wenn die dielektrische Funktion $\epsilon(\omega)$ bei der Frequenz

[2] Im folgenden werden wir wiederholt die Abkürzungen LO und TO für „longitudinal optisch" bzw. „transveral optisch" benutzen.

der Schwingung verschwindet und somit $\epsilon(\omega_\mathrm{l}) = 0$ ist. Dies hat zur Folge, daß bei ω_l zwischen den elektrischen Größen die Beziehung $\boldsymbol{D}_\mathrm{l} = \epsilon_0 \boldsymbol{\mathcal{E}}_\mathrm{l} + \boldsymbol{P}_\mathrm{l} = 0$ und $\boldsymbol{\mathcal{E}}_\mathrm{l} = -\boldsymbol{P}_\mathrm{l}/\epsilon_0$ bestehen muß. Mit Gleichung (11.7) folgt daraus für das lokale Feld der interessante Zusammenhang $\boldsymbol{\mathcal{E}}_\mathrm{lok} = -2\boldsymbol{P}_\mathrm{l}/3\epsilon_0$, d.h. die beiden Größen variieren gegenphasig. Das lokale Feld wirkt also der relativen Auslenkung $u = u_1 - u_2$ der Untergitter entgegen und verursacht so eine zusätzliche rücktreibende Kraft.

Transversale Phononen mit der Polarisation $\boldsymbol{P}_\mathrm{t}$ verursachen kein makroskopisches Feld in der Probe, so daß das lokale Feld $\boldsymbol{\mathcal{E}}_\mathrm{lok} = \boldsymbol{P}_\mathrm{t}/3\epsilon_0$ gleich dem Lorentz-Feld ist. Das lokale Feld, das die atomare Verschiebung mitbestimmt, hat also bei LO- und TO-Phononen entgegengesetztes Vorzeichen. Bei TO-Phononen wirkt es den rücktreibenden elastischen Kräften entgegen und macht das Material weicher. Mit der Kenntnis des lokalen Feldes, der Polarisation (11.14) und der Auslenkung (11.13) läßt sich die Frequenz der longitudinalen und transversalen Phononen leicht finden. Für die TO-Phononen erhält man die Beziehung (11.16) und für den longitudinalen Zweig

$$\omega_\mathrm{l}^2 = \omega_0^2 + \frac{2nq^2}{3\mu\epsilon_0} \cdot \frac{1}{1 + 2n\alpha/3} \,. \tag{11.21}$$

Damit wird verständlich, daß sich die beiden Phononenzweige in ihrer Frequenz unterscheiden und immer $\omega_\mathrm{l} > \omega_\mathrm{t}$ ist. In einem Fall hebt das elektrische Feld die Phononenfrequenz an, im anderen verursacht es eine Absenkung. Gitterverzerrung und Ionenpolarisation sind also eng verknüpft.

Es lohnt sich, die Wechselwirkung zwischen elektromagnetischen Wellen und optischen Phononen noch kurz zu betrachten. Es ist offensichtlich, daß eine wirkungsvolle Kopplung nur erfolgen kann, wenn beide Wellen in die gleiche Richtung laufen. Bei LO-Phononen stehen unter dieser Voraussetzung die elektrischen Felder der beiden Wellen aufeinander senkrecht. Wechselwirkung ist daher, unabhängig von der Ausbreitungsrichtung der beiden Wellen, ausgeschlossen, so daß longitudinale optische Phononen nicht durch Infrarotstrahlung angeregt werden können. Möglich ist ihr Nachweis dagegen mit Hilfe der Raman-Streuung (vgl. Abb. 11.6). Inelastische Elektronenstreuung ist eine weitere wichtige Methode zur Aufklärung der Anregungsspektren polarer Medien. Elektronen werden beim Durchgang durch den Festkörper vor allem an optischen Phononen gestreut.

Die Nullstelle der Gleichung (11.15) für die dielektrische Funktion findet man bei der Frequenz der longitudinalen optischen Phononen, die durch Gleichung (11.21) gegeben ist. Bilden wir das Verhältnis $\omega_\mathrm{l}^2/\omega_\mathrm{t}^2$ unter Berücksichtigung von Gleichung (11.17), so ergibt eine kurze Rechnung

$$\frac{\omega_\mathrm{l}^2}{\omega_\mathrm{t}^2} = \frac{\epsilon_\mathrm{st}}{\epsilon_\infty} \,. \tag{11.22}$$

Dies ist die *Lyddane-Sachs-Teller*-Beziehung, die hervorragend mit experimentellen Ergebnissen übereinstimmt. Sie zeigt, daß dielektrische und elastische Eigenschaften sehr eng miteinander verbunden sind. Ein Beispiel hierfür werden wir in Abb. 11.21 kennenlernen.

Phonon-Polaritonen

Transversale optische Phononen können direkt an elektromagnetische Wellen ankoppeln. Diese Wechselwirkung führt zu einer drastischen Änderung des Spektrums in der Nähe des Γ-Punktes, weil sich eine Mischung der beiden Wellentypen ausbildet. Bildlich gesprochen ruft eine elektromagnetische Welle eine Polarisation und damit eine Gitterverzerrung

hervor, die mit der Welle mitläuft. Umgekehrt wird eine TO-Gitterwelle von einer elektromagnetischen Welle begleitet.

Um auf möglichst einfache Art eine gemeinsame Lösung für die Schwingung der elektromagnetischen Welle und der TO-Gitterwelle zu finden, nehmen wir an, daß beide in x-Richtung laufen und in y-Richtung polarisiert sind. Für die Polarisation, die unsere Gitterwelle charakterisiert, schreiben wir daher: $\boldsymbol{P}_\mathrm{t} = \hat{\boldsymbol{y}} P_0 \exp[-\mathrm{i}(\omega t - q_\mathrm{t} x)]$. Die Ausbreitung elektromagnetischer Wellen läßt sich mit Hilfe der Wellengleichung

$$c^2 \Delta \boldsymbol{\mathcal{E}} = \epsilon(\omega) \ddot{\boldsymbol{\mathcal{E}}} \tag{11.23}$$

beschreiben, wobei c für die Lichtgeschwindigkeit im Vakuum steht. Setzen wir eine ebene Welle mit der elektrischen Feldstärke $\boldsymbol{\mathcal{E}} = \hat{\boldsymbol{y}} \mathcal{E}_0 \exp[-\mathrm{i}(\omega t - q_\mathrm{t} x)]$ als Lösung[3] an, so erhalten wir die bekannte Dispersionsrelation

$$\omega^2 = \frac{1}{\epsilon(\omega)} c^2 q_\mathrm{t}^2 \,, \tag{11.24}$$

deren Frequenzgang im wesentlichen durch die Frequenzabhängigkeit der uns bereits bekannten dielektrischen Funktion bestimmt wird. Wir vernachlässigen die Dämpfung, setzen (11.17) in diese Gleichung ein und erhalten

$$\omega^2 \left[\epsilon_\infty + \frac{\omega_\mathrm{t}^2 (\epsilon_\mathrm{st} - \epsilon_\infty)}{\omega_\mathrm{t}^2 - \omega^2} \right] = c^2 q_\mathrm{t}^2 \,. \tag{11.25}$$

Diese Dispersionsbeziehung, deren Verlauf in Abb. 11.5 schematisch dargestellt ist, hat folgende wichtige Grenzfälle: Ist $\omega \ll \omega_\mathrm{t}$, so vereinfacht sich die Gleichung zu $\omega = c q_\mathrm{t} / \sqrt{\epsilon_\mathrm{st}}$. Die elektromagnetische Welle breitet sich bei kleinen Frequenzen mit einer Geschwindigkeit aus, die durch die statische Dielektrizitätskonstante festgelegt wird. Ähnliches gilt bei optischen Frequenzen $\omega \gg \omega_\mathrm{t}$, bei denen man $\omega = c q_\mathrm{t} / \sqrt{\epsilon_\infty}$ findet, denn dort bestimmt die elektronische Polarisierbarkeit die Wellenausbreitung. Besonders interessant sind die Fälle $\omega \to \omega_\mathrm{t}$ und $\omega \to \omega_\mathrm{l}$. Im ersten Fall findet man, daß $q_\mathrm{t} \to \infty$, im zweiten dagegen, unter Berücksichtigung von Gleichung (11.22), daß $q_\mathrm{t} \to 0$ geht.

Die starke Kopplung zwischen Photonen und transversal optischen Phononen führt zu Mischzuständen, die man als *Polaritonen* bezeichnet. Fährt man, vom Wert Null kommend, die Frequenz durch, so erfolgt, wie in Abb. 11.5 dargestellt, ein stetiger Übergang von rein elektromagnetischen Wellen zu reinen Gitterschwingungen. Daran schließt sich ein verbotenes Frequenzband an, das durch die resonante Kopplung hervorgerufen wird und dessen Breite gerade $(\omega_\mathrm{l} - \omega_\mathrm{t})$ ist. Oberhalb der Lücke findet man bei der Frequenz ω_l wieder optische Phononen mit rein transversaler Auslenkung. Bei sehr hohen Frequenzen gehen die Anregungen dann allmählich in rein elektromagnetische Wellen über. Wie die schematische Darstellung der Polaritonendispersion in Abb. 11.5 zeigt, gibt es für jeden Wellenvektor zwei Frequenzwerte. Der verbotene Frequenzbereich wird nicht durch die Periodizität des Gitters hervorgerufen, wie es bei der Frequenzlücke der reinen Gitterschwingungen der Fall ist, sondern ist eine Folge der resonanten Kopplung, bei der die Entartung von elektromagnetischen und elastischen Wellen aufgehoben wird. In unpolaren Festkörpern ist $\epsilon(\omega)$ in diesem Frequenzbereich näherungsweise konstant, und die Dispersionskurven von Licht- und Gitterwellen kreuzen sich ohne Besonderheiten.

[3] Wir schreiben hier für den Wellenvektor q_t anstelle der üblichen Bezeichnung k_x. Der Grund hierfür ist, daß wir eine gemeinsame Lösung für $\boldsymbol{\mathcal{E}}$ und $\boldsymbol{P}$ suchen und den Wellenvektor der „Polarisationswelle" bereits mit q_t bezeichnet haben.

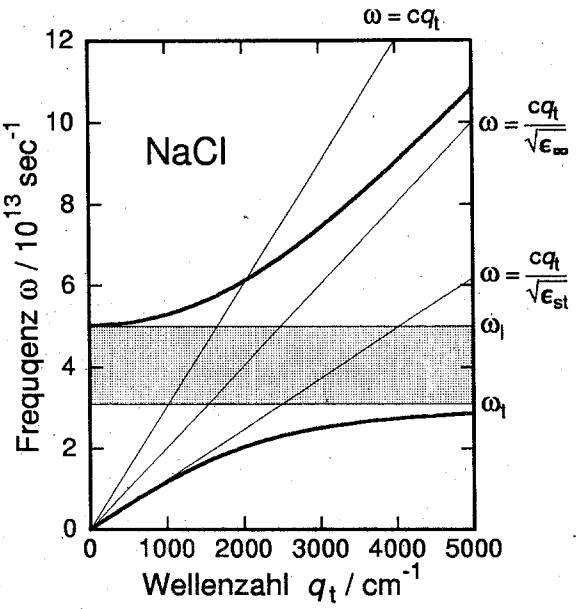


Abb. 11.5: Schematische Darstellung der Polaritondispersion. Dünn gezeichnet sind die Dispersionskurven des Lichts im Vakuum und in den beiden erwähnten Grenzfällen sowie der optischen Phononen ohne Kopplung. Zur Berechnung der ausgezogenen Kurven wurden die Zahlenwerte von Kochsalz $\omega_t = 3{,}1 \cdot 10^{13}\,s^{-1}$ und $\omega_l = 5{,}0 \cdot 10^{13}\,s^{-1}$ für die Frequenzen, und $\epsilon_{st} = 5{,}9$ und $\epsilon_\infty = 2{,}25$ für die Dielektrizitätskonstanten gewählt.

In Abb. 11.6 sind Daten der Raman-Streuung[4] an GaP wiedergegeben, mit deren Hilfe sich die Dispersionskurve von Phonon-Polaritonen verfolgen läßt. Die Meßpunkte, die zu den Polaritonen gehören, sind durch offene Kreise dargestellt. Daneben sind noch Werte für LO-Phononen eingezeichnet, die als Punkte dargestellt sind, die ebenfalls mit Hilfe der Raman-Streuung ermittelt wurden. In dem hier untersuchten kleinen Bereich hängt die Phononenfrequenz nicht von der Wellenzahl ab.

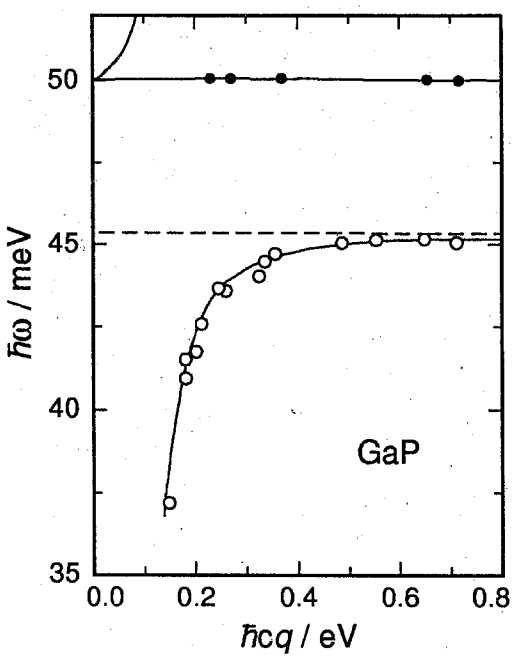


Abb. 11.6: Polaritonen (offene Kreise) in GaP, gemessen mit Hilfe der Raman-Streuung. Die Dispersionskurve der ungekoppelten TO-Phononen ist strichliert. Zusätzlich sind die Meßwert für LO-Phononen als Punkte eingezeichnet. (Nach C.H. Henry und J.J. Hopfield, Phys. Rev. Lett. **15** (1965) 964).

Ergänzend sei noch bemerkt, daß wir bei unseren Betrachtungen von großen Proben ausgegangen sind, also Oberflächeneffekte außer acht gelassen haben. Natürlich ändern

[4] Man beachte, daß bei diesem Experiment nur kleine Wellenzahlen untersucht werden können.

sich die Verhältnisse erheblich, wenn wir zu dünnen Schichten oder kleinen Körnern gehen, denn dort liegen andere elektrische Randbedingungen vor. Es ist deshalb nicht verwunderlich, daß sich Schwingungsfrequenzen und Kopplungsstärken ändern, wenn die Wellenlänge der anregenden Strahlung vergleichbar mit der Probendimension wird.

Optische Eigenschaften

Die Existenz von Polaritonen spiegelt sich im Verlauf der dielektrischen Funktion (11.17) und damit in den Infraroteigenschaften von Ionenkristallen wider, die wir anhand der experimentellen Daten von Kadmiumsulfid diskutieren wollen. Da in optischen Experimenten oder Messungen im Infrarotbereich nicht unmittelbar die dielektrische Funktion bestimmt werden kann, wollen wir vor der Betrachtung der experimentellen Daten noch kurz den Zusammenhang der dielektrischen Funktion mit den optischen Größen erläutern. Bezeichnen wir mit $\tilde{n}$ den *Brechungsindex* und mit κ den *Extinktionskoeffizienten*, so gilt

$$\epsilon' + \mathrm{i}\epsilon'' = (\tilde{n} + \mathrm{i}\kappa)^2, \tag{11.26}$$

und damit

$$\epsilon' = \tilde{n}^2 - \kappa^2 \qquad \text{und} \qquad \epsilon'' = 2\tilde{n}\kappa\,. \tag{11.27}$$

Messungen von $\tilde{n}$ und κ erweisen sich bei starker Absorption, z.B. im interessanten Bereich der Resonanz, als besonders schwierig. Deshalb wird oft die Reflektivität R bestimmt, für die bei senkrechtem Einfall die Beziehung

$$R = \left|\frac{\sqrt{\epsilon} - 1}{\sqrt{\epsilon} + 1}\right|^2 = \frac{(\tilde{n} - 1)^2 + \kappa^2}{(\tilde{n} + 1)^2 + \kappa^2} \tag{11.28}$$

gilt. Die beiden interessierenden Größen $\tilde{n}$ und κ lassen sich nicht aus dem Intensitätsverhältnis R von reflektiertem zu einfallendem Strahl ermitteln. Eine Reflexionsmessung wäre jedoch nur ausreichend, wenn neben diesem Verhältnis noch die Phasenverschiebung zwischen einlaufenden und reflektierten Wellen gemessen werden könnte. Einen Ausweg aus dieser Schwierigeit bietet die *Kramers-Kronig-Beziehung*, die den Realteil der Antwortfunktion eines linearen, passiven Systems mit deren Imaginärteil, d.h. $\tilde{n}$ mit κ bzw. ϵ' mit ϵ'', verbindet. Um die hierfür erforderliche numerische Integration tatsächlich ausführen zu können, muß die Meßgröße R einer unendlich dicken Probe bei „allen" Frequenzen, d.h. in der Praxis über einen sehr weiten Frequenzbereich, bekannt sein.

Obwohl in einem realen Experiment der Ausgangspunkt die gemessene Reflektivität ist, wollen wir in unserem Beispiel die Infraroteigenschaften von Kadmiumsulfid in „umgekehrter" Reihenfolge diskutieren. Die Abbildungen 11.7 und 11.8 zeigen den aus den Reflexionsdaten mit Hilfe der Kramers-Kronig-Beziehung berechneten Frequenzgang des Imaginär- und Realteils der dielektrischen Funktion von Kadmiumsulfid.

In Übereinstimmung mit Gleichung (11.19) tritt ein ausgeprägtes Maximum von ϵ'' bei der Wellenzahl $240\,\mathrm{cm}^{-1}$, d.h. bei der Frequenz $\omega_\mathrm{t} = 8{,}1 \cdot 10^{12}\,\mathrm{s}^{-1}$ der TO-Phononen auf. Das vorgelagerte kleine Maximum läßt sich mit unserer einfachen Theorie nicht verstehen. Es beruht auf anharmonischen Effekten und ist für unsere Diskussion linearer Zusammenhänge ohne Bedeutung. Vom statischen Wert $\epsilon_\mathrm{st} = 8{,}4$ steigt der Realteil ϵ' der Dielektrizitätskonstanten, wie von Gleichung (11.1) voraussagt, zunächst an, durchläuft ein Maximum und schneidet die Frequenzachse nahe ω_t. Anschließend ist ϵ' negativ und kreuzt die Achse ein zweites Mal bei $300\,\mathrm{cm}^{-1}$ bzw. bei $\omega_\mathrm{l} = 1{,}0 \cdot 10^{13}\,\mathrm{s}^{-1}$. Im weiteren Verlauf strebt ϵ' dem Wert $\epsilon_\infty = 5{,}3$ zu.

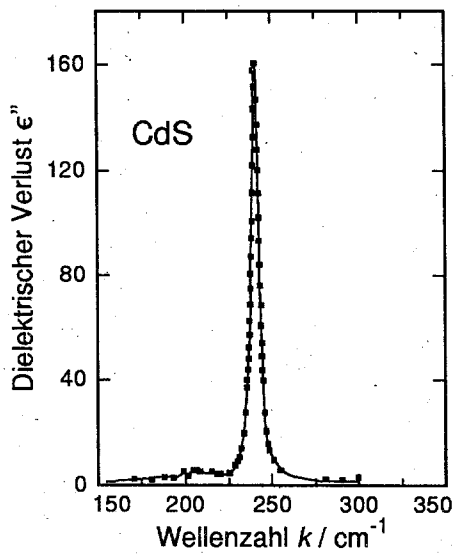


Abb. 11.7: Imaginärteil ϵ'' der dielektrischen Funktion von Kadmiumsulfid. (Nach M. Balkanski, in *Optical Properties of Solids* (F. Abelès ed.), North-Holland Publishing, Amsterdam, 1972).

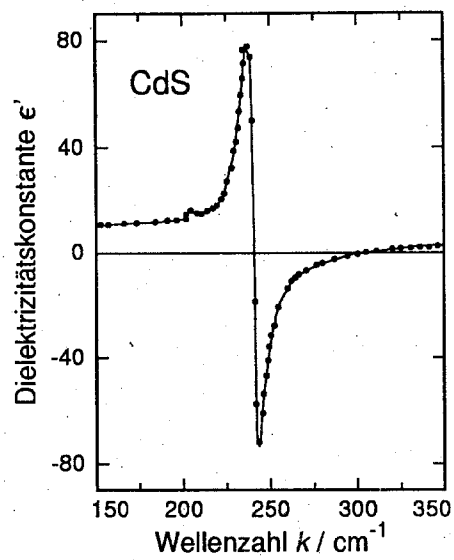


Abb. 11.8: Realteil ϵ' der dielektrischen Funktion von Kadmiumsulfid. (Nach M. Balkanski, in *Optical Properties of Solids* (F. Abelès ed.), North-Holland Publishing, Amsterdam, 1972).

Aus den beiden Kurven läßt sich der Verlauf des Brechungsindexes $\tilde{n}$ und des Extinktionskoeffizienten κ mit Hilfe von Gleichung (11.27) berechnen. In Abb. 11.9 ist das Ergebnis für κ dargestellt, das einen ähnliche Frequenzverlauf wie ϵ'' aufweist. Wie in 11.10 zu sehen ist, steigt der Brechungsindex $\tilde{n}$ von kleinen Frequenzen kommend zunächst mit wachsendem ϵ' an. Zwischen ω_t und ω_l, also im verbotenen Frequenzbereich, ist ϵ' negativ. Wäre in diesem Gebiet $\epsilon'' = 0$, so würde $\tilde{n}$ ebenfalls verschwinden.

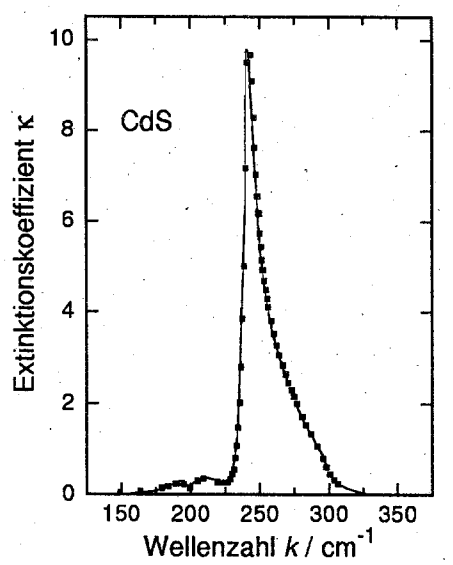


Abb. 11.9: Extinktionskoeffizient von Kadmiumsulfid. (Nach M. Balkanski, in *Optical Properties of Solids* (F. Abelès ed.), North-Holland Publishing, Amsterdam, 1972).

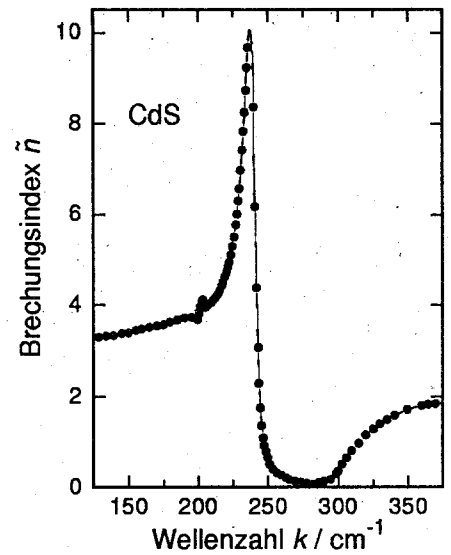


Abb. 11.10: Brechungsindex von Kadmiumsulfid. (Nach M. Balkanski, in *Optical Properties of Solids* (F. Abelès ed.), North-Holland Publishing, Amsterdam, 1972).

Die gemessene Reflektivität, aus der die gezeigten Daten hergeleitet wurden, ist in Abb. 11.11 zu finden. Sie steigt mit zunehmender Frequenz zunächst an, da der Brechungsindex zunimmt. Im verbotenen Frequenzbereich liegt die Reflektivität nahe bei eins. Die endliche Dämpfung bewirkt einen endlichen Wert von $\tilde{n}$ und damit eine Reflektivität kleiner als eins. Bei der Frequenz $1{,}2 \cdot 10^{13}\,\mathrm{s}^{-1}$, d.h. bei $320\,\mathrm{cm}^{-1}$, durchlaufen ϵ' und $\tilde{n}$ den Wert eins, und das Reflexionsvermögen verschwindet nach (11.28) annähernd. Oberhalb dieser Frequenz strebt der Brechungsindex dem optischen Grenzwert $\sqrt{\epsilon_\infty}$ zu, und das Reflexionsvermögen steigt entsprechend an. Es soll nochmals betont werden, daß man experimentell in umgekehrter Reihenfolge vorgeht. Es wird zunächst die Reflektivität sehr genau gemessen. Anschließend werden unter Zuhilfenahme der hier nicht diskutierten Kramers-Kronig-Beziehung die übrigen Größen berechnet.

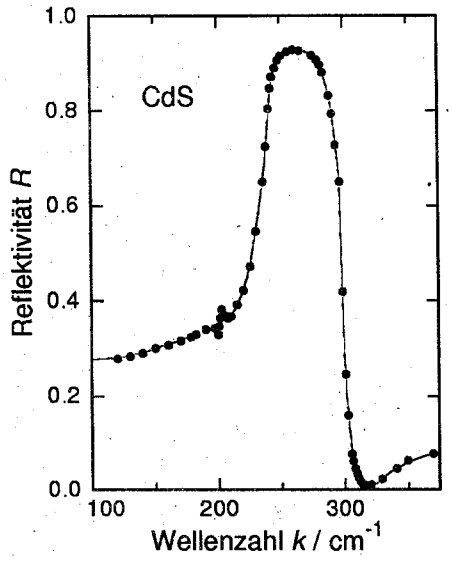


Abb. 11.11: Experimentell bestimmter Verlauf der Reflektivität von Kadmiumsulfid im Infraroten. Das elektrische Feld der einfallenden Strahlung stand bei dieser Messung senkrecht zur c-Achse. (Nach M. Balkanski, in *Optical Properties of Solids* (F. Abelès ed.), North-Holland Publishing, Amsterdam, 1972).

Früher wurden die soeben diskutierten dielektrischen Eigenschaften von Ionenkristallen zur Erzeugung von Infrarotstrahlung mit relativ enger Frequenzverteilung ausgenutzt. Läßt man einen Infrarotstrahl mit einem breiten Frequenzspektrum wiederholt an einem Ionenkristall reflektieren, so bleibt nach einigen Reflexionen nur Strahlung des Frequenzintervalls übrig, in dem eine hohe Reflektivität auftritt. Man bezeichnete daher dieses Vorgehen bei der Erzeugung eines schmalen Frequenzspektrums im Infraroten als *Reststrahlmethode.*

Elektronische Polarisierbarkeit

Bereits 1907 gelang es *Lorentz*, die elektronische Polarisierbarkeit mit Hilfe des *Oszillatormodells* sehr gut zu beschreiben. Die Elektronen werden dabei als negative Ladungswolken behandelt, die den Kern umgeben und durch die einfallende elektromagnetische Welle zu harmonischen Schwingungen angeregt werden. Wir benutzen hier dieses Modell, das uns erlaubt, die Ergebnisse, die wir bei der Behandlung der Ionenpolarisation erhalten haben, nach kleinen Modifikationen unmittelbar auf den Beitrag der Elektronen zu übertragen. In einfachster Näherung wird nur eine einzige elektronische Resonanz betrachtet, und wir können Gleichung (11.15) benutzen. Da dann kein weiterer Beitrag zur Polarisierbarkeit von Systemen berücksichtigt werden muß, deren Resonanz bei noch höheren Frequenzen liegt, setzen wir in dieser Gleichung $\alpha = 0$. Zur Unterscheidung zwischen der Ladung

und Masse der Ionen und der der Elektronen benutzen wir anstelle von q und μ nun die Abkürzungen e bzw. m und finden

$$\epsilon(\omega) = 1 + \frac{ne^2}{\epsilon_0 m} \frac{1}{\omega_0^2 - [ne^2/3\epsilon_0 m] - \omega^2 - \mathrm{i}\gamma\omega} . \tag{11.29}$$

Das lokale Feld bewirkt auch hier eine Frequenzerniedrigung, d.h. die Zustände der freien Atome werden durch die Nachbarn energetisch abgesenkt. Für unsere weitere Diskussion ist diese Änderung jedoch ohne Bedeutung, da eine starke Verschiebung der atomaren Resonanzfrequenzen beim Einbau der Atome in den Festkörper bereits durch andere Effekte (vgl. Abschnitt 7.3) verursacht wird. Bringen wir Gleichung (11.29) auf die Form von (11.17), so stellen wir fest, daß nun der hochfrequente Grenzwert der Dielektrizitätskonstanten erwartungsgemäß wegen $\alpha = 0$ den Vakuumwert eins annimmt. Die Bedeutung des Ausdrucks „statische Dielektrizitätskonstante" hat sich ebenfalls geändert, denn nun ist damit der niederfrequente Grenzwert des elektronischen Beitrags gemeint. Dies bedeutet, daß der statische Wert des elektronischen Beitrags bei der Diskussion der Ionenpolarisation mit der Abkürzung ϵ_∞ belegt wurde.

Wie gut stimmt diese einfache Betrachtung mit quantenmechanischen Überlegungen überein? Die Resonanzstelle entspricht dem Übergang zwischen zwei quantenmechanischen Niveaus, deren Energie sich um $\Delta E = \hbar\omega$ unterscheidet. Aus dem angeregten Zustand kehrt das Elektron anschließend unter Lichtemission in das Ausgangsniveau zurück. Bei diesem Vorgang bleibt normalerweise die Strahlungsenergie erhalten, d.h. es tritt keine Dissipation auf. Mit einer gewissen, meist geringen Wahrscheinlichkeit kann die absorbierte Energie an das Phononensystem oder an Störstellen[5] abgegeben werden. Diese Verluste werden in der hier gewählten makroskopischen Beschreibung durch die Dämpfungskonstante γ berücksichtigt. Für die Elektronen im Atom bzw. im Festkörper existieren nicht nur eine, sondern mehrere Resonanzfrequenzen, die sich unterschiedlich stark auf die Polarisation auswirken. Um dieser Tatsache Rechnung zu tragen, führt man für die einzelnen Übergänge die Oszillatorstärken f_i ein, deren Wert den experimentellen Gegebenheiten angepaßt wird[6]. Man summiert über alle auftretenden Resonanzen und findet schließlich für die dielektrische Funktion den Ausdruck

$$\epsilon(\omega) = 1 + \frac{ne^2}{\epsilon_0 m} \sum_i f_i \frac{1}{\omega_i^2 - \omega^2 - \mathrm{i}\gamma_i\omega} . \tag{11.30}$$

Einen wichtigen Festkörpereffekt, nämlich die Existenz von Energiebändern, haben wir bisher außer acht gelassen. Die Bänder beeinflussen die optischen Eigenschaften in zweifacher Hinsicht. Zum einen können bei nicht vollgefüllten Bändern elektronische Übergänge innerhalb eines Bandes erfolgen. Diese *Intrabandübergänge* spielen eine wesentliche Rolle bei den optischen Eigenschaften von Metallen und bewirken den endlichen Wert der elektrischen Leitfähigkeit. Bei *Interbandübergängen* verursachen die Bänder eine starke Verbreiterung der Absorptionslinien und eine Frequenzverbreiterung des Dispersionsbereiches. Die Absorption, die auf dem Übergang vom Valenz- ins Leitungsband beruht und somit den kleinsten Energieaufwand erfordert, wird als *Fundamentalabsorption* bezeichnet.

[5] Diese Vorgänge bei der Rekombination haben wir bereits in Abschnitt 9 bei der Diskussion der Leuchtdiode kennengelernt.

[6] Die hier benutze Definition der Oszillatorstärke f_i bezeichnet man als „Kramers-Heisenberg-Form."

In Abb. 11.12 ist der Verlauf von Real- und Imaginärteil der dielektrischen Funktion von Germanium über einen weiten Bereich der Photonenenergie zu sehen. Da Germanium weder Ionen- noch Dipolpolarisation aufweist, wird die dielektrische Funktion im gesamten Frequenzbereich nur durch den Beitrag der elektronischen Polarisation bestimmt. Offensichtlich treten in der Absorption mehrere Maxima auf, die sich überlagern, zu unterschiedlichen elektronischen Übergängen gehören und Anlaß zu anomaler Dispersion geben. Wie wir in Abschnitt 9.1 gesehen haben, bleibt beim Übergang zwischen den Bändern der Quasiimpuls erhalten. Da der Wellenvektor der Photonen sehr klein ist im Vergleich zu dem der Elektronen, ändert sich beim Interbandübergang der Wellenvektor der Elektronen praktisch nicht. Der Übergang erfolgt daher im $\boldsymbol{k}$-Raum „senkrecht". Die Übergangswahrscheinlichkeit und somit die Absorption hängen daher von der elektronischen Zustandsdichte des Anfangs- und Endzustands beim gleichen Wellenvektor ab. Dies bedeutet, daß die $\epsilon''(\omega)$-Kurve die *kombinierte Zustandsdichte* widerspiegelt. Aus Messungen der Reflektivität läßt sich ϵ'' unter Zuhilfenahme der Kramers-Kronig-Beziehung berechnen. Aus dem Verlauf von ϵ'' wiederum lassen sich Schlußfolgerungen über kritische Punkte (vgl. Kap. 7.4), die Zustandsdichte und damit über die Bandstruktur selbst ziehen. Die ausgeprägten Strukturen im Spektrum, wie in Abb. 11.12 zu sehen, sind direkt mit dem Auftreten von kritischen Punkten in der Zustandsdichte verknüpft.

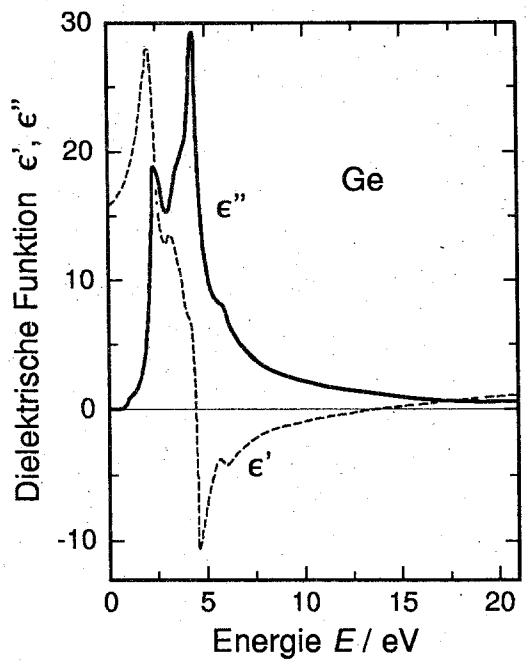


Abb. 11.12: Real- und Imaginärteil der dielektrischen Funktion von Germanium in einem weiten Energiebereich. (Nach H.R. Philipp und H. Ehrenreich, Phys. Rev. **129** (1963) 1550).

Mit zunehmender Photonenenergie nehmen die Oszillatorstärken rasch ab, da die Rumpfelektronen kaum zur Polarisierbarkeit beitragen. Die Absorption im Bereich des tiefen Ultraviolett und der Röntgenstrahlen ist daher sehr schwach und beeinflußt die optischen Eigenschaften bei tieferen Frequenzen kaum. Berechnet man den Brechungsindex aus den Daten von Abb. 11.12 mit Hilfe von Gleichung (11.27), so findet man, daß er bei einer Photonenenergie über 7 eV kleiner als eins wird, obwohl noch weitere Resonanzen im Röntgengebiet auftreten. Oberhalb von 14 eV wird ϵ' wieder positiv, und der Brechungsindex nähert sich dem Wert eins.

Orientierungspolarisation

Statische Polarisation

Bisher sind wir davon ausgegangen, daß das elektrische Feld Dipolmomente *induziert*. Besteht jedoch bereits ein *permanentes* Moment, so bewirkt ein anliegendes Feld eine *Vorzugsorientierung* der Momente, da die potentielle Energie $U = -\boldsymbol{p} \cdot \boldsymbol{\mathcal{E}} = -p\mathcal{E}\cos\theta$ vom Winkel θ zwischen Feld und Dipolmoment abhängt. Die gleiche Argumentation haben wir bereits in Abschnitt 10.1 bei der Diskussion des Paramagnetismus kennengelernt. Wie dort erreicht man, von sehr tiefen Temperaturen abgesehen, nur eine partielle Ausrichtung, da $p\mathcal{E} \ll k_\mathrm{B}T$ ist. Sind die Moleküle frei drehbar, wie es beispielsweise in Flüssigkeiten oder Gasen der Fall ist, so läßt sich der Mittelwert $\langle\cos\theta\rangle = p\mathcal{E}/3k_\mathrm{B}T$ wie beim Paramagnetismus ausdrücken, und man erhält für die Orientierungspolarisation P_0 im Falle eines Gleichfeldes die *Langevin-Debye*-Gleichung

$$P_0 = np\langle\cos\theta\rangle = np\frac{p\mathcal{E}}{3k_\mathrm{B}T}\,. \tag{11.31}$$

Dieser Ausdruck ist völlig analog zu Gleichung (10.4). Der Temperaturverlauf der damit verbundenen Suszeptibilität entspricht dem Curie-Gesetz (10.8) für die Magnetisierung. Wie bei den paramagnetischen Substanzen wurde auch hier vorausgesetzt, daß die Wechselwirkung zwischen den Dipolen vernachlässigt werden kann, so daß die Suszeptibilität erst am absoluten Nullpunkt divergieren sollte. Ein Beispiel für die statische Dielektrizitätskonstante von frei drehbaren Molekülen finden wir in Abb. 11.13. Von hoher Temperatur kommend, steigt die Dielektrizitätskonstante der polaren Flüssigkeit Nitromethan CH_3NO_2 bis zum Erstarrungspunkt an. Beim Übergang zum Festkörper springt die Dielektrizitätskonstante auf einen kleineren Wert, der nahezu temperaturunabhängig ist. Dies bedeutet, daß die CH_3NO_2-Moleküle im festen Zustand nicht mehr frei drehbar sind, nur feste Gleichgewichtslagen einnehmen und somit nicht mehr zur Orientierungspolarisation beitragen. Es gibt aber viele Festkörper, in denen sich Moleküle auch in der festen Phase noch drehen können. Dies ist meist dann der Fall, wenn keine kovalenten Bindungen vorliegen und die Moleküle näherungsweise kugelförmig sind.

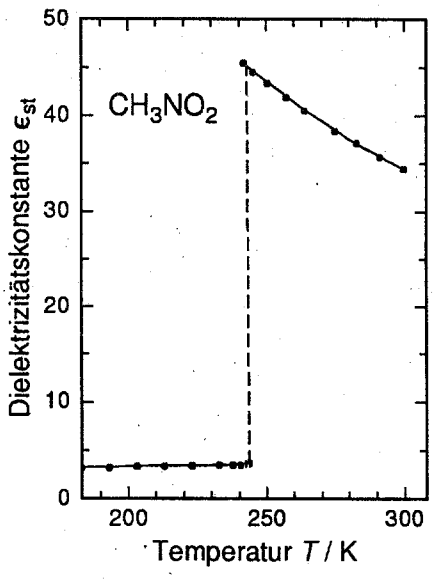


Abb. 11.13: Dielektrizitätskonstante der polaren Flüssigkeit CH_3NO_2 mit dem Erstarrungspunkt bei 243 K. (Nach C.P. Smyth und W.S. Walls, J. Am. Chem. Soc. **54** (1932) 3230).

In Festkörpern können die Struktureinheiten nicht wirklich frei rotieren, sondern nehmen bevorzugte diskrete Orientierungen ein. In der folgenden Diskussion setzen wir vereinfachend voraus, daß die Dipole aufgrund ihrer Gestalt und der Kristallstruktur nur zwei Gleichgewichtslagen einnehmen können, die einer Drehung um 180° entsprechen. Weiterhin gehen wir davon aus, daß das Feld in Richtung der Verbindungslinie der beiden Gleichgewichtslagen anliegt. Die potentielle Energie der Dipole als Funktion des Drehwinkels verläuft dann wie in Abb. 11.14 skizziert. Ohne Feld sind die beiden gestrichelt gezeichneten Gleichgewichtslagen, zwischen denen die Potentialbarriere V_a besteht, gleichberechtigt.

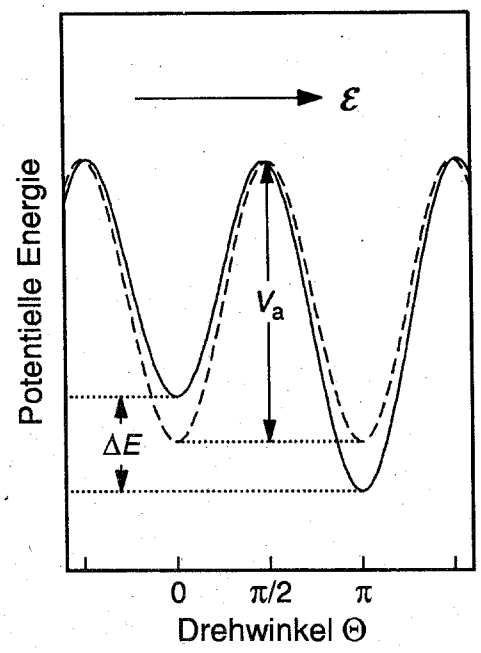


Abb. 11.14: Verlauf der potentiellen Energie eines elektrischen Dipols in Abhängigkeit vom Winkel zwischen Dipol und anliegendem Feld. Es ist der Verlauf mit und ohne (strichlierte Kurve) elektrisches Feld dargestellt.

Liegt das elektrische Feld $\mathcal{E}$ in Richtung der beiden Einstellmöglichkeiten, so tritt, ohne Berücksichtigung der Korrektur für das lokale Feld, zwischen den beiden Gleichgewichtslagen die Energiedifferenz $\Delta E = 2p\mathcal{E}$ auf. Die Polarisation P_0 wird durch die Differenz in der Besetzung der beiden Richtungen festgelegt. Bereits in Abschnitt 6.2 haben wir die thermische Besetzung von Zweiniveausystemen diskutiert. Wir benutzen dieses Ergebnis und erhalten für die Polarisation und die statische Dielektrizitätskonstante

$$P_0 = np \tanh \frac{p\mathcal{E}}{k_B T} \approx \frac{np^2\mathcal{E}}{k_B T} , \tag{11.32}$$

und

$$\epsilon_{st} - 1 \approx \frac{np^2}{\epsilon_0 k_B T} . \tag{11.33}$$

Bis auf den Faktor 1/3 entspricht dieses Ergebnis dem Resultat für frei rotierende Moleküle in Flüssigkeiten oder Festkörpern. Das Ergebnis läßt sich auf kompliziertere Verhältnisse übertragen, ohne daß dabei die Ergebnisse qualitative Änderungen erfahren. In Abb. 11.15 ist die statische Dielektrizitätskonstante von flüssigem und festem HCl dargestellt. Klar zu erkennen ist, daß sich der Anstieg von ϵ_{st} mit abnehmender Temperatur auch unterhalb der Erstarrungstemperatur mit geänderter Steigung fortsetzt. Der Sprung bei der Verfestigung überrascht nicht, denn dort erfolgt eine Einschränkung der Einstellmöglichkeiten.

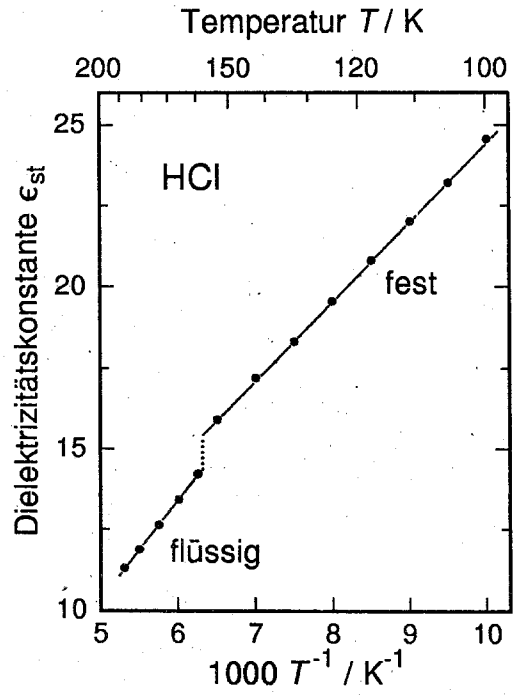


Abb. 11.15: Statische Dielektrizitätskonstante ϵ_{st} von HCl als Funktion der reziproken Temperatur. Die sprunghafte Änderung tritt bei der Verfestigung auf. (Nach R.W. Swenson und R.H. Cole, J. Chem. Phys. **22** (1954) 284).

Es stellt sich die interessante Frage nach dem Verhalten derartiger Dipole bei sehr tiefen Temperaturen, denn dort treten quantenmechanische Aspekte immer stärker in den Vordergrund. Wie wir bereits in Abschnitt 5.7 gesehen haben, weisen amorphe Festkörper Tunnelsysteme auf, die die Tieftemperatureigenschaften solcher Materialien weitgehend bestimmen. Auch in Kristallen mit Punktdefekten findet man häufig Tunnelsysteme, die wir nun etwas genauer betrachten wollen. Ein Beispiel hierfür stellen Lithiumatome dar, die in Kaliumchloridkristallen als substitutionelle Fremdatome eingebaut werden. Es zeigt sich, daß die energetisch günstigste Lage für die Li^+-Ionen, deren Durchmesser weniger als die Hälfte der Kaliumionen beträgt, nicht in der Mitte des K^+-Gitterplatzes ist, sondern auf „Off-center"-Positionen etwas verschoben in den acht ⟨111⟩-Richtungen. Zwischen diesen acht energetisch äquivalenten Positionen können die Lithiumionen auch bei tiefsten Temperaturen noch tunneln. Dies führt zu einem System mit acht Eigenzuständen, dessen Termschema in Abb. 11.16 als Teilbild eingezeichnet ist. Der Abstand zwischen den Niveaus, die „Tunnelaufspaltung", ist mit Δ_0 bezeichnet. Insgesamt treten nur vier unterschiedliche Energiewerte auf, da die beiden mittleren Niveaus jeweils dreifach entartet sind. Mit dem Aufenthalt der Li^+-Ionen in den Off-center-Positionen ist aufgrund der Ladungsverschiebung ein vergleichsweise großes Dipolmoment p verbunden. Die resultierende Polarisation in einem statischen elektrischen Feld läßt sich mit Hilfe der Thermodynamik aus der (negativen) Ableitung der freien Energie F nach dem elektrischen Feld berechnen. Für $p\mathcal{E} \ll \Delta_0$ ergibt die Rechnung, ähnlich wie bei der Orientierungspolarisation klassischer Dipole, das einfache Ergebnis

$$P_0 = -\frac{\partial F}{\partial \mathcal{E}} = \frac{2}{3}\frac{np^2\mathcal{E}}{\Delta_0}\tanh\left(\frac{\Delta_0}{2k_\mathrm{B}T}\right) . \tag{11.34}$$

Interessanterweise erhält man das Ergebnis (11.34) auch für die Polarisation einfacher Zwei-Niveau-Systeme mit der Tunnelaufspaltung Δ_0. Dies ist ein Ausdruck der hohen

Symmetrie der Lithiumdefekte im kubischen KCl. Für „hohe" Temperaturen $T > \Delta_0/k_B$ dürfen wir den Hyperbeltangens entwickeln, und Gleichung (11.34) geht in die klassische Langevin-Debye-Formel (11.31) über. Die Polarisation ist in diesem Bereich proportional zu T^{-1} und unabhängig von der Tunnelaufspaltung.

Abbildung 11.16 zeigt die Temperaturabhängigkeit der dielektrischen Suszeptibilität χ_{Li}, die von sehr geringen Konzentrationen ^{6}Li bzw. ^{7}Li in KCl-Kristallen hervorgerufen wird. Sie ist über $\chi_{Li} = P_0/\epsilon_0\mathcal{E}$ mit der Polarisation (11.34) verknüpft. Die durchgezogenen Linien in Abb. 11.16 sind auf diese Weise unter Berücksichtigung der unterschiedlichen Tunnelaufspaltungen der beiden Isotope berechnet. Die Übereinstimmung zwischen Theorie und Experiment ist bei beiden Kristallen perfekt. Besonders schön läßt sich bei dieser Messung der Einfluß der unterschiedlichen Isotope erkennen: Der Gesamteffekt wächst proportional zu $1/\Delta_0$ an, und die Kurve des leichteren Isotops ist zu höheren Temperaturen hin verschoben.

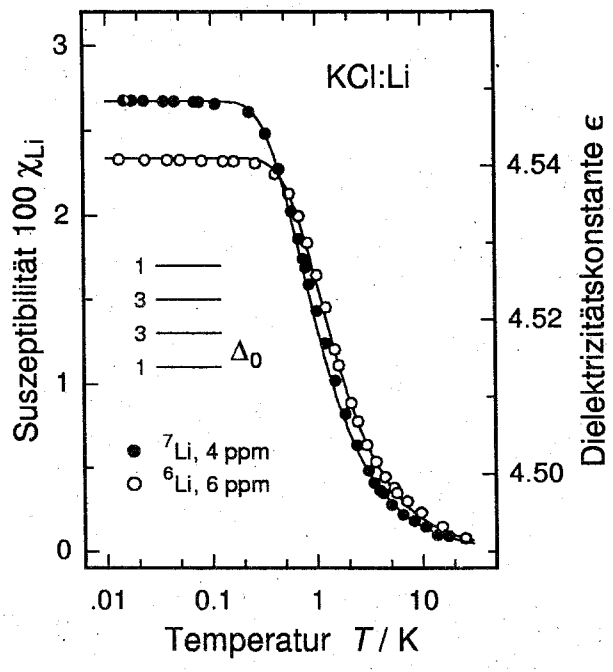


Abb. 11.16: Dielektrische Suszeptibilität χ_{Li} für zwei KCl:Li-Proben als Funktion der Temperatur. Die durchgezogenen Kurven spiegeln die Gleichung (11.34) wider. Ebenfalls eingezeichnet ist das Termschema der Lithiumdefekte, in dem auch der Grad der Entartung angegeben ist. Die Skala auf der rechten Seite bezieht sich auf die Dielektrizitätskonstante der gesamten Probe. (Nach C. Enss, Physica B **219&220** (1996) 239).

Relaxationseffekte

Nach diesem kurzen Blick auf die Quanteneffekte bei sehr tiefen Temperaturen wenden wir uns wieder klassischen Dipolen zu. Interesssante Phänomene treten auf, wenn anstelle des statischen Feldes ein Wechselfeld angelegt wird, da dann die Dynamik des Orientierungsprozesses die Reaktion der Dipole bestimmt. Nach Gleichung (11.32) wächst die statische Gleichgewichtspolarisation P_0 proportional zum Feld an. Da die Energiedifferenz ΔE im Wechselfeld periodisch schwankt, muß ständig eine Umbesetzung der Orientierungsrichtung und damit der beiden Energieniveaus erfolgen. Den Annäherungvorgang an den Gleichgewichtszustand, der in ähnlicher Form in vielen Bereichen der Physik eine wichtige Rolle spielt, bezeichnet man als *Relaxation*. Einen speziellen Fall haben wir bereits bei der elektrischen Leitfähigkeit in Abschnitt 8.2 kennengelernt. Wir benutzen die Relaxationsgleichung (8.26) und setzen anstelle der Verteilungsfunktion $f(E,T)$ die Polarisation P

als relaxierende Größe ein:

$$\frac{\mathrm{d}P}{\mathrm{d}t} = -\frac{P - P_\mathrm{g}}{\tau} \,. \tag{11.35}$$

Hierbei bezeichnet τ die *Relaxationszeit* und $P_\mathrm{g} = np^2\mathcal{E}/k_\mathrm{B}T$ den *augenblicklichen Gleichgewichtswert* der Polarisation, der durch die gerade herrschende elektrische Feldstärke bestimmt wird und sich einstellen würde, wenn man beliebig lange warten würde. Legt man beispielsweise sprungartig ein elektrisches Feld an, so errechnet man mit Hilfe der Relaxationsgleichung erwartungsgemäß ein exponentielles Anwachsen der Polarisation, bis der neue Gleichgewichtswert P_g erreicht ist: $P(t) = P_\mathrm{g}[1 - \exp(-t/\tau)]$. Mit der Relaxationszeit τ haben wir uns bereits in Abschnitt 4.1 bei der Diskussion der Leerstellendiffusion beschäftigt. Mit τ wurde dort die Zeit bezeichnet, die zwischen zwei Platzwechselvorgängen verstreicht. Auch im vorliegenden Fall erfolgt der „Platzwechsel", d.h. die Reorientierung der Dipole, thermisch aktiviert. Wie in Abb. 11.14 schematisch dargestellt, ist hierzu ein Sprung über die Potentialbarriere der Höhe V_a erforderlich[7]. Da formal die gleichen Umstände wie bei der Diffusion vorliegen, können wir die in Abschnitt 4.1 eingeführte Sprungrate ν der reziproken Relaxationszeit τ^{-1} gleichsetzen und finden den Zusammenhang

$$\tau^{-1} = \tau_0^{-1}\mathrm{e}^{-V_\mathrm{a}/k_\mathrm{B}T} \,. \tag{11.36}$$

Für die „Versuchsfrequenz" τ_0^{-1} setzen wir die Debye-Frequenz ν_D ein, da bei höheren Temperaturen die Moleküle mit ν_D in den jeweiligen Potentialmulden schwingen.

Wie hängt die Orientierung der permanenten Dipole im periodischen Wechselfeld $\mathcal{E}(t) = \mathcal{E}_0 \exp(-\mathrm{i}\omega t)$ von der Frequenz ab? Vereinfachend setzen wir $\mathcal{E} = \mathcal{E}_\mathrm{lok}$, obwohl diese Näherung in vielen Fällen nicht gut erfüllt ist. Sie führt jedoch zu übersichtlichen Ausdrücken, ohne die auftretenden Effekte qualitativ zu verändern. Wir gehen von einem linearen Zusammenhang zwischen der Polarisation und der elektrischen Feldstärke aus und berücksichtigen, daß die Suszeptibilität im allgemeinen frequenzabhängig ist, d.h. es gilt $P(t) = \epsilon_0\chi_\mathrm{d}(\omega)\mathcal{E}(t)$. Die Gleichgewichtspolarisation P_g wird durch den statischen Wert $\chi_\mathrm{d}(0)$ der Suszeptibilität bestimmt. Durch Einsetzen dieses Ansatzes in Gleichung (11.35) erhalten wir somit für die dipolare Suszeptibilität

$$\chi_\mathrm{d}(\omega) = \frac{\chi_\mathrm{d}(0)}{1 - \mathrm{i}\omega\tau} \,. \tag{11.37}$$

Die dielektrische Funktion des Festkörpers ergibt sich, wenn wir die Beiträge der Ionen und Elektronen addieren: $\epsilon(\omega) = 1 + \chi_\mathrm{d} + \chi_\mathrm{i} + \chi_\mathrm{e}$. Da der Beitrag der permanenten Dipole nur bis zu Frequenzen im Mikrowellenbereich von Bedeutung ist, können wir χ_e und χ_i als konstant annehmen und erhalten für Mikrowellenfrequenzen

$$\epsilon(\omega) = 1 + \chi_\mathrm{i} + \chi_\mathrm{e} + \frac{\chi_\mathrm{d}(0)}{1 - \mathrm{i}\omega\tau} = \epsilon_\infty + \frac{\epsilon_\mathrm{st} - \epsilon_\infty}{1 - \mathrm{i}\omega\tau} \,. \tag{11.38}$$

Die Grenzwerte ϵ_∞ und ϵ_st sind, wenn auch mit geänderten Grenzfrequenzen, wie im vorhergehenden Abschnitt definiert. Anstelle des Resonanznenners, wie er z.B. in Gleichung (11.17) zu finden ist, tritt der typische Relaxationsnenner auf. Spalten wir die

[7] Auch an Tunnelsystemen können Relaxationseffekte beobachtet werden, doch wollen wir diese Frage hier nicht weiter verfolgen.

dielektrische Funktion wieder in Real- und Imaginärteil auf, so ergeben sich die bekannten *Debye-Gleichungen*:

$$\epsilon'(\omega) = \epsilon_\infty + \frac{\epsilon_{st} - \epsilon_\infty}{1 + \omega^2\tau^2}\,, \tag{11.39}$$

$$\epsilon''(\omega) = \frac{(\epsilon_{st} - \epsilon_\infty)\omega\tau}{1 + \omega^2\tau^2}\,. \tag{11.40}$$

Trägt man die beiden Größen ϵ' und ϵ'' als Funktion von $\omega\tau$ auf, so findet man den in Abb. 11.17 skizzierten Verlauf. Der Realteil nimmt stetig mit $\omega\tau$ ab, weist also nicht das in Abb. 11.4 dargestellte Resonanzverhalten auf. Ähnlich wie bei einer Resonanz durchläuft der Imaginärteil ein Maximum, das jedoch vergleichsweise breit ist. Der steilste Abfall von ϵ' bzw. das Maximum von ϵ'' tritt bei $\omega\tau = 1$ auf, wenn Schwingungsdauer und Relaxationszeit vergleichbar sind. Wie der Gleichung (11.40) bzw. der Abbildung 11.17 zu entnehmen ist, verringert sich ϵ'' nur um den Faktor 5, wenn sich die Meßfrequenz (oder die Relaxationszeit) von der Bedingung $\omega\tau = 1$ eine Dekade entfernt.

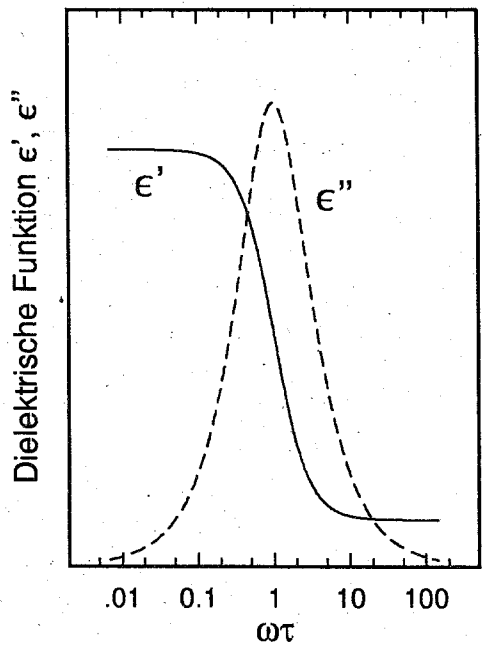


Abb. 11.17: Real- und Imaginärteil der dielektrischen Funktion bei Relaxation in Abhängigkeit von $\omega\tau$. Der Maßstab der Abszisse ist logarithmisch, der der Ordinate dagegen liear. Die wesentlichen Änderungen erfolgen bei $\omega\tau \approx 1$.

Als Beispiel zeigen wir in Abb. 11.18 dielektrische Messungen an Eis bei einer konstanten Temperatur von 226 K. Die Messung überstreicht vier Größenordnungen der Frequenz und bestätigt weitgehend die Erwartung. Tatsächlich wurde die gezeigte Messung jedoch an einer Eisemulsion durchgeführt. Der Grund hierfür ist, daß Eis bei niedrigen Frequenzen eine relativ hohe Leitfähigkeit aufweist, die von Defekten und Verunreinigungen hervorgerufen wird, wodurch die Messung des Relaxationseffektes verfälscht wird. Das isolierende Öl, das die Eiskriställchen umgibt, bewirkt eine Unterdrückung dieser störenden „Gleichstromleitfähigkeit", zeigt aber selbst bei der Meßtemperatur keine frequenzabhängigen Relaxationseffekte. Somit spiegeln die beiden Bilder das typische Relaxationsverhalten von Eis wider. Dies gilt allerdings nicht für den Absolutwert von ϵ', da zur Dielektrizitätskonstanten einer Mischung natürlich beide Bestandteile beitragen.

Hier überwiegt der Beitrag von Öl, das selbst nur eine kleine Dielektrizitätskonstante aufweist[8].

Eine genauere Analyse ergibt jedoch für die Meßgrößen eine schwächere Frequenzabhängigkeit, als sie durch Gleichung (11.39) bzw. (11.40) vorhergesagt wird. Diese „Verschmierung" wird durch eine Verteilung der Relaxationszeiten hervorgerufen, die darauf zurückzuführen ist, daß Eis nicht perfekt geordnet ist (vgl. auch Abschnitt 3.6 und Abb. 3.13). Aus der Lage des Maximums des Imaginärteils der dielektrischen Funktion läßt sich über die Beziehung $\omega\tau = 1$ direkt die mittlere Relaxationszeit ablesen. Bemerkenswert ist die niedrige Frequenz, bei der das Maximum auftritt. Sie zeigt an, daß die Relaxation der Wassermoleküle in Eis sehr langsam vonstatten geht[9].

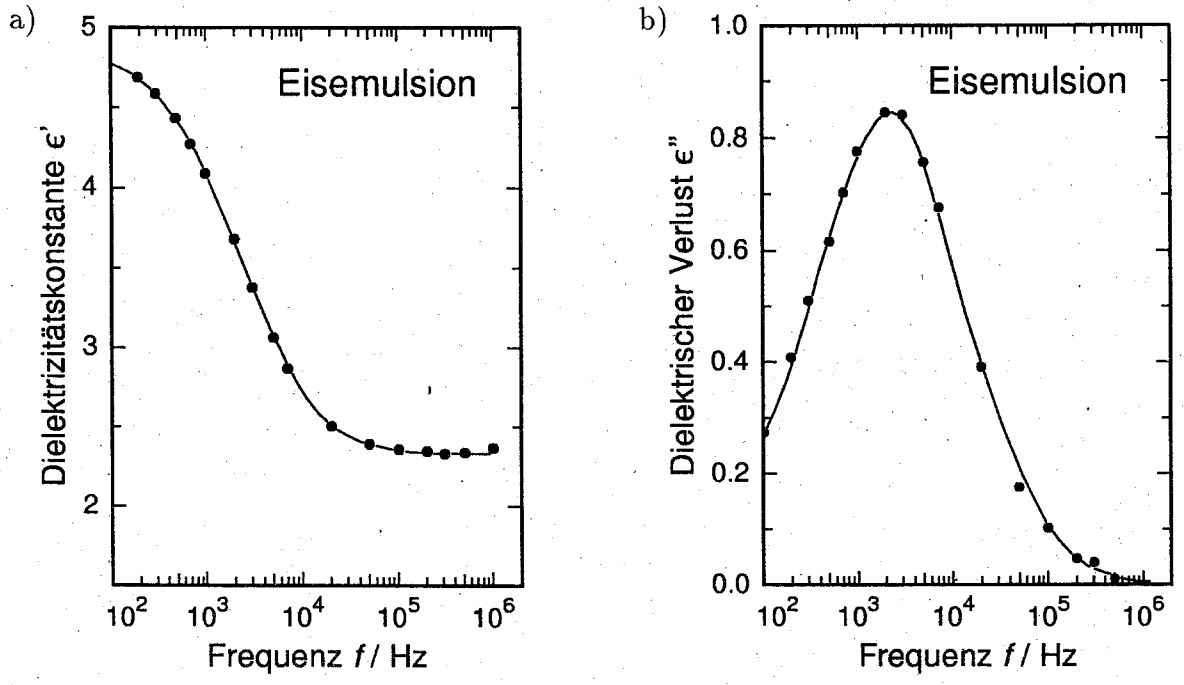


Abb. 11.18: Frequenzabhängigkeit der dielektrischen Funktion von Eis bei 226 K. (Nach G. Evrard in *Physics and Chemistry of Ice*, E. Whalley, S.J. Jones, L.W. Gold (eds.), Royal Society of Canada, Ottawa, 1973).
a) Realteil,
b) Imaginärteil.

In vielen dielektrischen Experimenten wird die Frequenz festgehalten und die Temperatur verändert. Da nach Gleichung (11.36) die Relaxationszeit stark von der Temperatur abhängt, kann auf diese Weise das Produkt $\omega\tau$ über weite Bereiche variiert werden. Daraus lassen sich Aussagen über die Temperaturabhängigkeit der Relaxationszeit und ihre möglicherweise vorhandene Verteilung gewinnen.

[8] Bekanntlich besitzt Wasser bei Gleichspannung die Dielektrizitätskonstante $\epsilon' \approx 81$. Bei Eis nimmt die Dielektrizitätskonstante sogar den Wert 105 an.

[9] In Wasser erfolgt die Relaxation sehr viel rascher als in Eis. Bei Zimmertemperatur ist die Relaxationszeit der Wassermoleküle im Bereich von $\tau \approx 10\,\text{ps}$, d.h. das Maximum wird bei Messungen im Mikrowellenbereich beobachtet. Die kurze Relaxationszeit ist einer der Gründe, weshalb Mikrowellenherde bei relativ hohen Frequenzen arbeiten.

Ferroelektrizität

In Abschnitt 10.2 haben wir die ferromagnetischen Festkörper kennengelernt, für die das Auftreten einer spontanen Magnetisierung charakteristisch ist. Ein ähnliches Verhalten zeigen auch die *pyroelektrischen* Materialien, bei denen sich bereits ohne äußeres Feld eine makroskopische elektrische Polarisation einstellt. Eine derartige *spontane Polarisation* kann sich nur dann ausbilden, wenn der Kristall *eine polare Achse* besitzt. Sind mehrere polare Achsen vorhanden, so ist der Kristall lediglich *piezoelektrisch*, d.h. eine mechanische Deformation ruft eine elektrische Polarisation hervor. Da polare Achsen nur auftreten können, wenn keine strukturelle Inversionssymmetrie vorliegt, findet man keine amorphen Substanzen in dieser Kategorie von Festkörpern. Als *ferroelektrisch* bezeichnet man einen Kristall, wenn die spontane Polarisation durch ein genügend starkes, der Polarisation entgegengerichtetes Feld umgeklappt werden kann.

Erwärmt man ferroelektrische Kristalle, so verschwindet die Polarisation bei der kritischen Temperatur T_c. Oberhalb dieser Temperatur sind die Kristalle *paraelektrisch* und besitzen eine statische Dielektrizitätskonstante, die mit einer Gleichung analog dem Curie Weiss-Gesetz (10.11) beschrieben werden kann:

$$\epsilon_{st} = \frac{C}{T - \Theta} \,. \tag{11.41}$$

Hierbei ist C eine materialspezifische Konstante und Θ die paraelektrische Curie-Temperatur. Als Beispiel ist in Abb. 11.19 der Temperaturverlauf der statischen Dielektrizitätskonstanten ϵ_{st} von $BaTiO_3$ gezeigt. Bemerkenswert sind die äußerst großen Werte von ϵ_{st}, die in der Nähe von Θ auftreten.

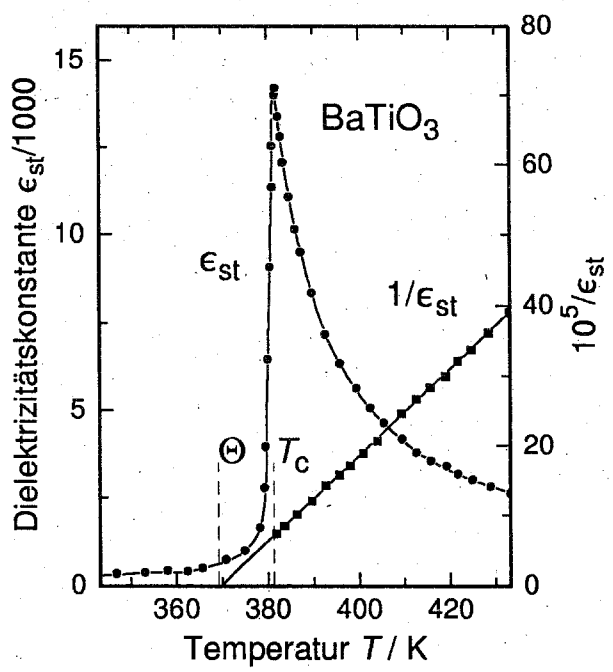


Abb. 11.19: Dielektrizitätskonstante ϵ_{st} von $BaTiO_3$, gemessen in Richtung der tetragonalen c-Achse. Zusätzlich ist der Verlauf von ϵ_{st}^{-1} in der paraelektrischen Phase aufgetragen, um die Übereinstimmung mit dem Curie Weiss-Gesetz zu verdeutlichen. (Nach W.J. Merz, Phys. Rev. **91** (1953) 513).

Die Namensgebung weist auf die große phänomenologische Ähnlichkeit mit den Ferromagneten hin, die über die hier diskutierten Effekte hinausgeht. Es scheint daher naheliegend, Ferroelektrika wie Ferromagneten zu beschreiben und in dieser Theorie die magnetischen Dipole durch permanente elektrische zu ersetzen. Dieses Vorgehen ist jedoch

zum Scheitern verurteilt, weil die beiden Phänomene auf ganz unterschiedlichen mikroskopischen Ursachen beruhen.

Der Übergang von der para- zur ferroelektrischen Phase zeigt, verglichen mit ferromagnetischen Übergängen, ein relativ komplexes Erscheinungsbild. Während ferromagnetische Phasenübergänge immer Übergänge 2. Ordnung sind, können ferroelektrische je nach Substanz von 1. *oder* 2. Ordnung sein. Abhängig von der Natur des Übergangs weist der charakteristische Ordnungsparameter (vgl. Abschnitt 4.4) daher bei T_c entweder einen Sprung oder einen stetigen Verlauf auf. Als Ordnungsparameter können wir hier die spontane Polarisation heranziehen, die oberhalb von T_c verschwindet und unterhalb von T_c mit abnehmender Temperatur ansteigt.

Als Beispiel ist in Abb. 11.20 die Temperaturabhängigkeit der spontanen Polarisation an einem ferroelektrischen Übergang 1. Ordnung gezeigt. Die Polarisation von $BaTiO_3$ springt beim Abkühlen bei der Übergangstemperatur T_c auf einen endlichen Wert und nimmt dann weiter zu. Wie dem Bild zu entnehmen ist, treten im Fall von $BaTiO_3$ bei tieferen Temperaturen noch zwei weitere Sprünge auf. Sie rühren von Richtungsänderungen der polaren Achse her, die durch strukturelle Phasenübergänge hervorgerufen werden. Darüberhinaus beobachtet man Hystereseeffekte, die für Phasenübergänge 1. Ordnung typisch sind, auf die wir aber hier nicht weiter eingehen.

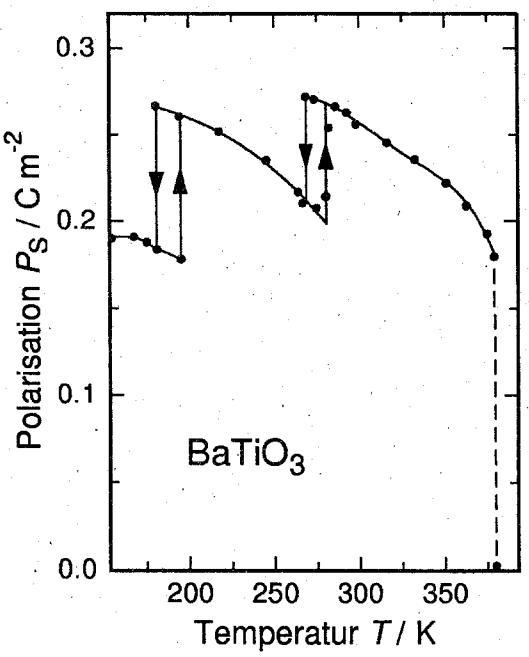


Abb. 11.20: Spontane Polarisation von $BaTiO_3$. Bei $T_c = 383\,K$ erfolgt eine diskontinuierliche Änderung der Polarisation. Die Sprünge bei tieferen Temperaturen sind eine Folge weiterer Phasenübergänge. (Nach W.J. Merz, Phys. Rev. **91** (1953) 513); H.H. Wieder, Phys. Rev. **99** (1955) 1161).

Man kennt zwei verschiedene Mechanismen, die zur Ausbildung einer spontanen Polarisation bei den Ferroelektrika führen, die wir hier nur sehr kurz ansprechen wollen. Bei einer Reihe von Substanzen ist der ferroelektrische Übergang mit einem Ordnungs-Unordnungs-Übergang (vgl. Abschnitt 4.4) verbunden. Dies bedeutet, daß beim Erwärmen am Phasenübergang die Ordnung der elektrischen Dipolmomente verloren geht, die Dipole selbst aber nicht verschwinden. Zu dieser Substanzklasse zählt das ferroelektrische Kalium-Dihydrogen-Phosphat (KH_2PO_4), bei dem Protonen in Wasserstoffbrücken (vgl. Abschnitt 3.6) zwischen den negativ geladenen PO_4-Ionen die Dipolmomente hervorrufen. Ähnlich wie beim Eis ist mit der Wasserstoffbrückenbindung ein Dipolmoment verknüpft, dessen Stärke und Richtung durch die Lage des Protons

bezüglich der Bindungspartner bestimmt wird. Oberhalb der kritischen Temperatur von 123 K nehmen die Protonen statistisch eine der beiden möglichen Gleichgewichtslagen zwischen benachbarten PO_4-Ionen ein, und der Mittelwert der Polarisation verschwindet. Beim Durchgang durch den ferroelektrischen Übergang tritt eine Bevorzugung einer bestimmten Ausrichtung ein, und es baut sich eine makroskopische Polarisation auf.

Bei der zweiten Gruppe der Ferroelektrika verschieben sich beim Übergang zur ferroelektrischen Phase zwei Untergitter gegeneinander. Mit anderen Worten: Beim Phasenübergang werden nicht bereits vorhandene Dipolmomente geordnet, sondern durch Verschiebung von Ionen erzeugt. Typische Vertreter hierfür sind Ionenkristalle mit Perowskit-Struktur[10], zu denen auch $BaTiO_3$ gehört. Der Verlauf der Dielektrizitätskonstanten und der spontanen Polarisation in der Nähe des Phasenübergangs wurde bereits anhand der Abbildungen 11.19 bzw. 11.20 diskutiert. Bei der kritischen Temperatur $T_c = 388\,\mathrm{K}$, die oft durch Verunreinigungen etwas reduziert wird, verschieben sich die negativ geladenen Sauerstoffionen bleibend um 0,1 Å gegen die positiven Metallionen. Dies führt zur Ausbildung eines Dipolmomentes von etwa 10^{-29} Cm pro Elementarzelle.

Verschiebungsübergänge lassen sich auf verschiedene Weisen beschreiben, nämlich mit Hilfe der Clausius-Mossotti- oder über die Lyddane-Sachs-Teller-Beziehung. Betrachten wir zunächst die Clausius-Mossotti-Beziehung (11.9). Hierzu lösen wir nach der statischen Dielektrizitätskonstanten ϵ_{st} auf und finden:

$$\epsilon_{st} = \frac{1 + \frac{2}{3}\sum_i n_i \alpha_i}{1 - \frac{1}{3}\sum_i n_i \alpha_i} \,. \tag{11.42}$$

Hierbei steht α_i für die Summe aus elektronischer und ionischer Polarisierbarkeit. Durch die Summation wird berücksichtigt, daß Ferroelektrika aus verschiedenen Ionensorten bestehen, die jeweils zur Dielektrizitätskonstanten beitragen. Nähert sich in der paraelektrischen Phase $\sum n_i \alpha_i$ beim Abkühlen dem Wert drei, dann strebt $\epsilon_{st} \to \infty$. Dieser Anstieg kann bereits durch die übliche Verkleinerung der Gitterkonstanten beim Abkühlen ausgelöst werden, falls $\sum n_i \alpha_i$ bei höheren Temperaturen bereits genügend nahe am kritischen Wert liegt. Es bildet sich eine bleibende Auslenkung und damit eine spontane Polarisation aus, weil die auslenkende Kraft, die auf ein Ion aufgrund des lokalen elektrischen Feldes wirkt, schneller mit der Verschiebung ansteigt als die lineare elastische Rückstellkraft. Man bezeichnet diesen Vorgang als *Polarisationskatastrophe*. Auslenkung und Polarisation bleiben aber endlich, weil bei großer Auslenkung zusätzliche nichtlineare Gitterkräfte der Verzerrung entgegenwirken.

Wie bereits erwähnt, lassen sich das Ansteigen der statischen Dielektrizitätskonstanten und die Ausbildung einer spontanen Polarisation bei den Verschiebungsferroelektrika auch auf ganz andere Weise beschreiben. Wir wollen diese Vorgehensweise, die auch auf nicht-ferroelektrische Materialien übertragbar ist, am Beispiel von $SrTiO_3$ verdeutlichen, das ebenfalls Perowskitstruktur besitzt. Wie Abb. 11.21 zeigt, nimmt von hohen Temperaturen kommend die Dielektrizitätskonstante von $SrTiO_3$, wie von den Ferroelektrika bekannt, dem Curie Weiss-Gesetz folgend zu. Bei Strontiumtitanat erfolgt jedoch in der Nähe der Curie-Temperatur von etwa 35 K kein ferroelektrischer Übergang, sondern ein verlangsamtes Ansteigen der Dielektrizitätskonstanten (bzw. ein Abflachen der $1/\epsilon_{st}$-Kurve).

[10] Das Mineral Perowskit ($CaTiO_3$) hat kubische Struktur und wurde bereits in Kap. 1 beschrieben. Die Einheitszelle ist in Abb. 1.19 dargestellt.

Die Verknüpfung zwischen der statischen Dielektrizitätskonstante ϵ_{st} und der Frequenz ω_t der transversal optischen Phononen, die wiederum nach Gleichung (11.16) von der elektrischen Polarisation abhängt, erfolgt durch die Lyddane-Sachs-Teller-Beziehung. Setzen wir das Curie Weiss-Gesetz (11.41) in Gleichung (11.22) ein, so erhalten wir

$$\omega_t^2 = \epsilon_\infty \epsilon_{st}^{-1} \omega_l^2 \propto (T - \Theta) \,. \tag{11.43}$$

Weil ϵ_∞ und ω_l kaum von der Temperatur abhängen, sollte ω_t^2 proportional zur reduzierten Temperatur $(T - \Theta)$ verlaufen. Die Abnahme von ω_t beim Kühlen bezeichnet man als „Weichwerden" der transversalen Kristallschwingung. In dieser Sichtweise könnte man die spontane Polarisation der Ferroelektrika als eine „eingefrorene" Transversalschwingung bezeichnen.

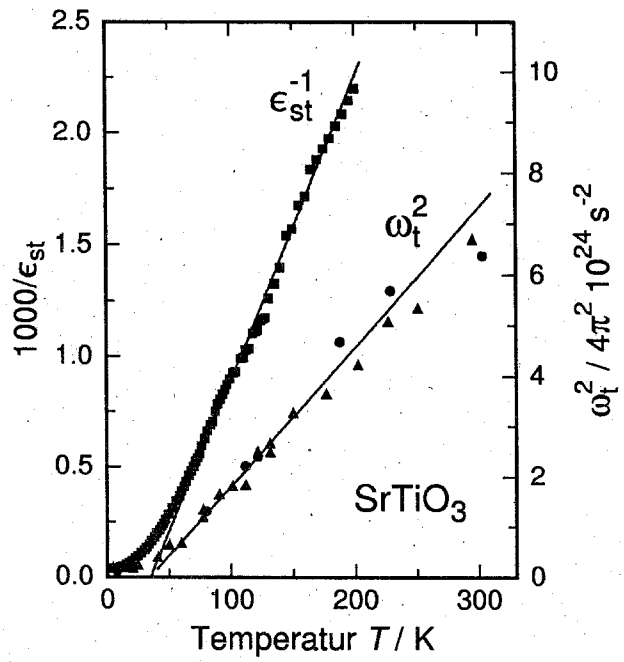


Abb. 11.21: Temperaturverlauf von $1/\epsilon_{st}$ und ω_t^2 von $SrTiO_3$. Die Phononenfrequenzen wurden mit Hilfe von Neutronen- und Raman-Streuung gemessen. (Nach T. Sakudo und H. Unoki, Phys. Rev. Lett. **26** (1971) 851; Y. Yamada und G. Shirane, J. Phys. Soc. Japan **26** (1969) 396).

11.3 Exzitonen

In diesem Kapitel haben wir bislang drei Mechanismen diskutiert, die in Isolatoren Anlaß zur Polarisation und Absorption von elektromagnetischen Wellen geben. Sieht man von speziellen Substanzklassen wie der ferromagnetischen ab, so existieren noch zwei weitere wichtige Absorptionsmechanismen, von denen wir einen, nämlich die Absorption durch Störstellen, im Zusammenhang mit den Punktdefekten in Abschnitt 4.1 kennengelernt haben. So verbleibt für die Diskussion hier noch die Absorption durch Exzitonen, auf die wir nun eingehen wollen.

Bei der bereits erwähnten Fundamentalabsorption bewirkt jedes absorbierte Photon das Anheben eines Elektrons ins Leitungsband, wobei ein Loch im Valenzband zurückbleibt. Bisher haben wir Elektronen im Leitungsband und Löcher im Valenzband als unabhängige Teilchen behandelt. Dies ist jedoch nicht mehr erlaubt, wenn die Energie der eingestrahlten Photonen etwas kleiner als die Bandlücke ist. Dann kommt es zu einer speziellen, elektrisch neutralen Anregung, bei der Elektron und Loch nicht vollständig

voneinander getrennt werden, sondern aufgrund der Coulomb-Anziehung eine Einheit, ein *Exziton*, bilden. Voraussetzung für die Erzeugung dieser Anregung ist, daß sich beide Ladungsträger mit der gleichen Gruppengeschwindigkeit bewegen, denn sonst werden sie räumlich voneinander getrennt. Da der Übergang des Elektrons vom Valenz- ins Leitungsband nahezu senkrecht erfolgt, ist diese Voraussetzung für die Bildung von Exzitonen nur erfüllt, wenn die Gruppengeschwindigkeit der beiden Ladungsträgersorten verschwindet. Dies ist an den kritischen Punkten (vgl. Abschnitt 7.4) der Elektronenbänder der Fall. Durch thermische Energie oder zusätzliche Einstrahlung elektromagnetischer Wellen (Photoionisation) kann ein bestehendes Elektron-Loch-Paar aufgebrochen werden. Dann entstehen freie Ladungsträger mit den bereits ausführlich diskutierten Eigenschaften. Natürlich kann ein Exziton auch durch Rekombination zerfallen, wobei das Loch durch das Elektron aufgefüllt wird. Hierbei wird entweder Lumineszenzstrahlung mit einer Frequenz emittiert, die der Energiedifferenz entspricht, oder aber die Rekombination erfolgt strahlungslos, wobei die Anregungsenergie direkt an das Gitter abgegeben wird.

Man unterscheidet stark und schwach gebundene Exzitonen. Stark gebundene oder *Frenkel-Exzitonen* treten vorwiegend in Molekül-, Edelgas- oder Ionenkristallen auf. Aufgrund der relativ starken Coulomb-Wechselwirkung ist die Bindung zwischen Elektron und Loch stark, so daß die Bindungsenergie bei dieser Sorte von Exzitonen relativ groß ist, nämlich etwa 1 eV oder sogar noch größer. Damit unterscheiden sich die Energie für Exzitonenanregung und Band-Band-Übergänge deutlich, so daß Frenkel-Exzitonen sehr gut mit optischen Methoden, z.B. durch Messung der optischen Absorption, untersucht werden können. Loch und Elektron sitzen bei den meisten Exzitonen auf ein- und demselben Atom oder Molekül und können somit auch als angeregter Zustand eines Moleküls betrachtet werden. Exzitonen haben die Fähigkeit, von Atom zu Atom zu tunneln und somit im Kristall zu diffundieren und ihre Anregungsenergie durch den Festkörper zu transportieren.

Schwach gebundene oder *Mott-Wannier-Exzitonen* findet man vor allem bei Materialien mit kleinem Bandabstand. Sie sind deshalb typisch für Halbleiter. Da die Bindungsenergie der Exzitonen wesentlich kleiner als der Bandabstand sein muß, folgt bereits aus der Unschärferelation, daß der räumliche Abstand zwischen Elektron und Loch relativ groß sein muß. Zum Beispiel beträgt in Germanium die Bindungsenergie der Exzitonen etwa 2,5 meV und der Elektron-Loch-Abstand etwa 5 nm. Die Eigenzustände von Exzitonen lassen sich in einfachster Näherung durch ein Wasserstoffatommodell mit einer modifizierten Rydberg-Formel darstellen:

$$E_\nu = E_\mathrm{g} - \frac{\mu^* e^4}{32\pi^2\hbar^2\epsilon_\mathrm{r}^2\epsilon_0^2}\,\frac{1}{\nu^2}\,. \tag{11.44}$$

Hierbei ist E_g die Energie der Bandlücke, ν die Hauptquantenzahl und μ^* die reduzierte Masse, die durch $1/\mu^* = 1/m_\mathrm{n}^* + 1/m_\mathrm{p}^*$ gegeben ist.

Der starke Einfluß der Exzitonen auf die optische Absorption soll anhand von zwei Beispielen, nämlich von GaAs und Cu_2O, gezeigt werden. In Abb. 11.22 ist eine Messung der optischen Absorption von GaAs bei tiefen Temperaturen dargestellt. Deutlich ist das Maximum zu erkennen, das der Fundamentalabsorption vorgelagert ist und durch die Erzeugung von Exzitonen hervorgerufen wird. Die einzelnen Linien, die man aufgrund von Gleichung (11.44) erwartet, lassen sich in diesem Material jedoch nicht auflösen. Bei Raumtemperatur ist die thermische Energie größer als die „Ionisationsenergie" der Exzitonen, d.h. es gilt $k_\mathrm{B}T \gg (E_\mathrm{g} - E_1)$. In diesem Fall erfolgt der optische Übergang unter Mitwirkung von Phononen direkt ins Leitungsband, und das Maximum verschwindet.

Eine Paradesubstanz, in der man das Auftreten von einzelnen Exzitonenlinien beobachten kann, ist Cu_2O. In der Darstellung des Absoptionskoeffizienten in Abb. 11.23 sind die Übergänge in die Exzitonenniveaus mit den Quantenzahlen $\nu = 2$ bis $\nu = 5$ klar zu erkennen. In Messungen bei Heliumtemperatur konnten sogar Zustände bis $\nu = 11$ beobachtet werden. Der Übergang, der auf der Erzeugung eines Exzitons im Grundzustand basiert, wäre in der Abb. 11.23 auch dann nicht zu sehen, wenn die Energieachse zu kleineren Werten hin fortgesetzt wäre. Dieser Übergang ist als elektrischer Dipolübergang verboten, doch tritt ein Quadrupolübergang auf, der ebenfalls nachgewiesen werden kann, aber zu einer wesentlich schwächeren Absorptionslinie führt.

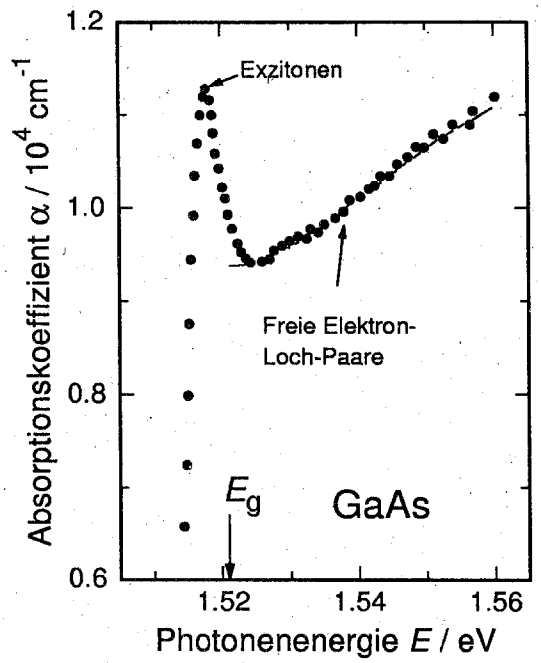


Abb. 11.22: Optische Absorption von GaAs an der Absorptionskante, gemessen bei 21 K. Bereits vor dem Einsetzen der Fundamentalabsorption tritt ein Maximum auf, das von der Erzeugung von Exzitonen herrührt. (Nach M.D. Sturge, Phys. Rev. **127** (1962) 768).

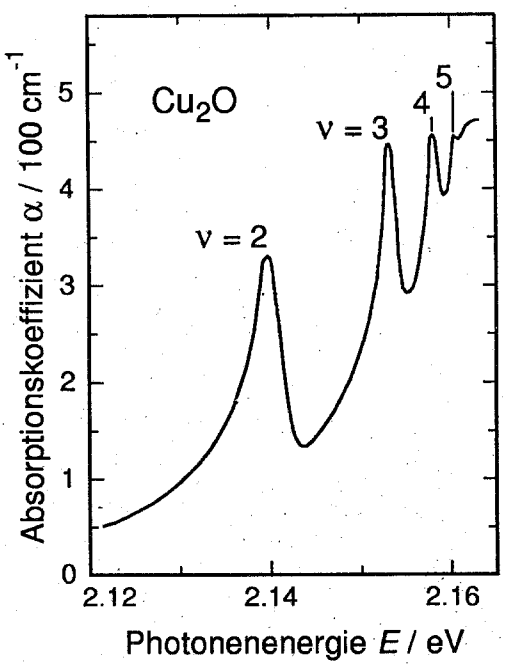


Abb. 11.23: Optische Absorption von Cu_2O bei 77 K als Funktion der eingestrahlten Photonenenergie. Es wird eine ganze Serie von Exzitonenlinien beobachtet. (Nach P.W. Baumeister, Phys. Rev. **121** (1961) 359).

11.4 Optische Eigenschaften freier Ladungsträger

In diesem Abschnitt beschäftigen wir uns mit der Bewegung der Elektronen unter dem Einfluß elektromagnetischer Wellen in den teilweise gefüllten Bändern von Metallen und stark dotierten Halbleitern. Die optischen Eigenschaften dieser Materialien werden in erster Linie durch kollektive Anregungen des freien Elektronengases und nicht durch Anregung einzelner Elektronen bestimmt. Die dielektrische Funktion läßt sich wie bei den gebundenen Elektronen herleiten. Wir benutzen die Bewegungsgleichung (11.12), berücksichtigen aber, daß $\omega_0 = 0$ gesetzt werden kann, da im freien Elektronengas keine rücktreibenden Kräfte herrschen:

$$m^*\ddot{u} + \frac{m^*\dot{u}}{\tau} = -e\mathcal{E}(t)\,. \tag{11.45}$$

Ladung und effektive Masse der Elektronen wurden mit e bzw. m^* bezeichnet. Anstelle der Dämpfungskonstanten γ, die in Gleichung (11.12) benutzt wurde, haben wir die Stoßzeit τ eingeführt, die sich aus der elektrischen Leitfähigkeit (vgl. Abschnitt 8.2) ermitteln läßt. Im vorliegenden Fall sind Korrekturen für das lokale Feld unwesentlich, und die zu (11.15) analoge Lösung für die dielektrische Funktion nimmt die einfache Gestalt

$$\epsilon(\omega) = 1 + n\alpha - \frac{ne^2}{\epsilon_0 m^*} \frac{1}{\omega^2 + \mathrm{i}\omega/\tau} \tag{11.46}$$

an. Da die Stoßzeit der Elektronen in Metallen mittlerer Reinheit bei Zimmertemperatur über 10^{-14} s liegt (vgl. Abschnitt 8.2), spielt die Dämpfung bei optischen Frequenzen keine wesentliche Rolle und wird im folgenden vernachlässigt. Bei kleineren Frequenzen ist dies nicht erlaubt, und das Verhalten des freien Elektronengases ist wesentlich komplexer. Für den Fall optischer Frequenzen schreiben wir für (11.46) vereinfachend

$$\epsilon(\omega) = \epsilon_\infty - \frac{ne^2}{\epsilon_0 m^*} \cdot \frac{1}{\omega^2} = \epsilon_\infty \left(1 - \frac{\omega_\mathrm{p}^2}{\omega^2}\right) , \tag{11.47}$$

wobei der Beitrag der gebundenen Elektronen in ϵ_∞ enthalten ist. Zusätzlich haben wir als Abkürzung noch die *Plasmafrequenz*

$$\omega_\mathrm{p}^2 = \frac{ne^2}{\epsilon_0 \epsilon_\infty m^*} \tag{11.48}$$

eingeführt, deren Bedeutung wir im folgenden noch kennenlernen werden. Wir wollen dieses Ergebnis nun nutzen, um die optischen Eigenschaften von Metallen und stark dotierten Halbleitern zu diskutieren.

Ausbreitung elektromagnetischer Wellen in Metallen

Die sehr einfache dielektrische Funktion (11.47) erlaubt bereits ein weitgehendes Verständnis der optischen Eigenschaften von Metallen. Wir gehen von der Wellengleichung (11.23) für elektromagnetische Wellen aus und setzen als Lösung wieder eine ebene Welle an, für deren elektrische Feldstärke wir $\mathcal{E} = \mathcal{E}_0 \exp[-\mathrm{i}(\omega t - kx)]$ schreiben. Die Wahl der x-Richtung schränkt die Gültigkeit der Lösung nicht ein, da ein freies Elektronengas isotrop ist. Wir erhalten damit die bekannte Beziehung $\epsilon(\omega)\omega^2 = c^2k^2$. Setzen wir dort die dielektrische Funktion (11.47) für Metalle ein, so finden wir für elektromagnetische Wellen in Metallen die Dispersionsrelation

$$\epsilon_\infty \omega^2 \left(1 - \frac{\omega_\mathrm{p}^2}{\omega^2}\right) = c^2k^2 . \tag{11.49}$$

Abhängig von der Frequenz treten zwei Bereiche mit völlig unterschiedlichen Eigenschaften des Elektronengases auf. Unterhalb der Plasmafrequenz, d.h. für $\omega < \omega_\mathrm{p}$, ist $k^2 < 0$ und k somit imaginär. Langwellige elektromagnetische Wellen können sich in Metallen weder ausbreiten noch können sie in diese eindringen. Wie wir bereits bei der Diskussion der polaren Medien gesehen haben, tritt bei $\epsilon < 0$ Totalreflexion auf. Ist $\omega > \omega_\mathrm{p}$, dann ist $\epsilon(\omega) > 0$, und die Wellengleichung beschreibt die Ausbreitung einer elektromagnetischen Welle mit der Dispersionsbeziehung

$$\omega^2 = \omega_\mathrm{p}^2 + \frac{c^2k^2}{\epsilon_\infty} \tag{11.50}$$

für *Plasmon-Polaritonen.* Bei großen Wellenvektoren nähert sich die Gruppengeschwindigkeit, wie von den Isolatoren her bekannt, dem Wert $v_g = c/\sqrt{\epsilon_\infty}$. In Abb. 11.24 ist der prinzipielle Verlauf der Dispersionskurve skizziert. Bei tiefen Frequenzen sind Metalle für elektromagnetische Strahlung undurchlässig, aber oberhalb der Plasmafrequenz werden sie transparent.

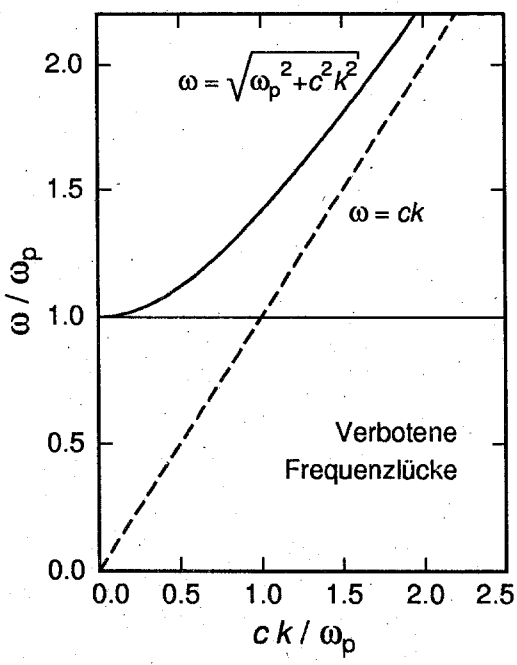


Abb. 11.24: Dispersion von Plasmon-Polaritonen im freien Elektronengas. Unterhalb der Plasmafrequenz ω_p können sich elektromagnetische Wellen nicht ausbreiten; es tritt ein verbotener Frequenzbereich auf. In dieser schematichen Darstellung wurde $\epsilon_\infty = 1$ gesetzt.

In Tabelle 11.1 sind für einige Elektronendichten unter der Annahme $\epsilon_\infty = 1$ die Plasmafrequenzen und die dazugehörenden Wellenlängen λ_p angegeben. Daraus läßt sich schließen, daß einfache Metalle bereits im Ultravioletten durchsichtig werden. Die gute Übereinstimmung zwischen der einfachen Rechnung und den experimentell ermittelten Werten macht Tabelle 11.2 deutlich. Bei Übergangsmetallen spielen die bereits angesprochenen Interbandübergänge eine wichtige Rolle und modifizieren die optischen Eigenschaften.

Tabelle 11.1: Plasmafrequenz ω_p und Plasmawellenlänge λ_p bei verschiedenen Elektronendichten.

$n\,(\mathrm{cm}^{-3})$	10^{22}	10^{18}	10^{14}	10^{10}
$\omega_p\,(\mathrm{s}^{-1})$	$5{,}7 \cdot 10^{15}$	$5{,}7 \cdot 10^{13}$	$5{,}7 \cdot 10^{11}$	$5{,}7 \cdot 10^{9}$
$\lambda_p\,(\mathrm{cm})$	$3{,}3 \cdot 10^{-5}$	$3{,}3 \cdot 10^{-3}$	$3{,}3 \cdot 10^{-1}$	$3{,}3 \cdot 10^{1}$

Tabelle 11.2: Errechnete und gemessene Plasmawellenlänge λ_p einiger Alkalimetalle.

	Li	Na	K	Rb
$\lambda_p\,(\mu\mathrm{m})$ (errechnet)	0,155	0,209	0,287	0,322
$\lambda_p\,(\mu\mathrm{m})$ (experimentell)	0,155	0,21	0,315	0,34

Als Beispiele für den Verlauf des Reflexionsvermögens des freien Elektronengases sind in Abb. 11.25 die experimentellen Daten für Aluminium dargestellt. Ein Vergleich mit den Reflexionseigenschaften von polaren Kristallen macht die Bedeutung der Plasmafrequenz nochmals klar. Bei Ionenkristallen tritt zwischen den beiden Frequenzen ω_t und ω_l sehr hohe Reflektivität auf, die dann steil abfällt (vgl. Abb. 11.11). Dieser Abfall bei ω_l ist mit dem Nulldurchgang des Realteils ϵ' der dielektrischen Funktion verbunden. Beim Elektronengas verschwindet das Reflexionsvermögen oberhalb der Plasmafrequenz, bei der in diesem Fall ebenfalls $\epsilon' = 0$ ist. Dies legt den Schluß nahe, daß mit der Plasmafrequenz longitudinale Schwingungen verbunden sind. Wir werden auf die experimentelle Beobachtung dieser Schwingung im folgenden Abschnitt eingehen. Während bei polaren Kristallen die Reflektivität zu niedrigen Frequenzen hin jenseits der Frequenz der transversalen optischen Phononen wieder stark abnimmt, bleibt sie bei Metallen bis zu Gleichfeldern sehr groß. Das unterschiedliche Verhalten bei kleinen Frequenzen beruht auf der fehlenden Schersteifigkeit des Elektronengases, so daß für Metalle $\omega_t = 0$ gesetzt werden darf. Der „Einbruch" der Reflektivität von Aluminium bei etwa 1,5 eV und das gegenüber dem idealen Wert $R = 1$ reduzierte Reflexionsvermögen im verbotenen Frequenzbereich rührt vom endlichen Wert von ϵ'' aufgrund von Interbandübergängen her.

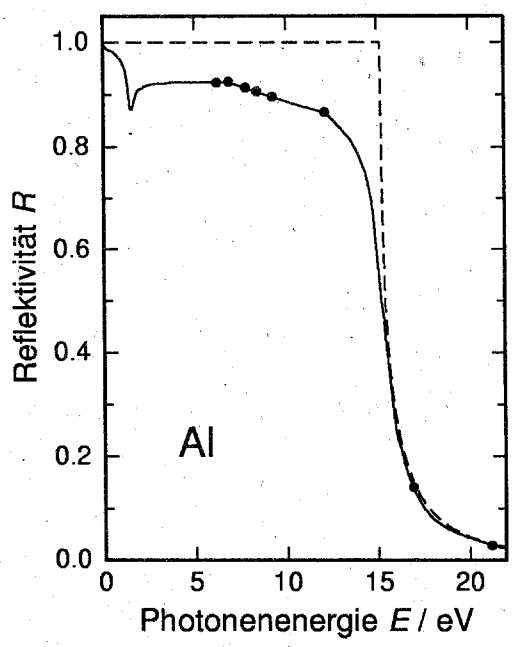


Abb. 11.25: Reflektivität von Aluminium. Die durchgezogene Kurve gibt den tatsächlichen, die gestrichelte Kurve den mit Gleichung (11.47) und den Konstanten $\epsilon_\infty = 1$ und $\hbar\omega_p = 15{,}3\,\text{eV}$ errechneten Verlauf wieder. (Nach H. Ehrenreich, H.R. Philipp, B. Segall, Phys. Rev. **132** (1963) 1913).

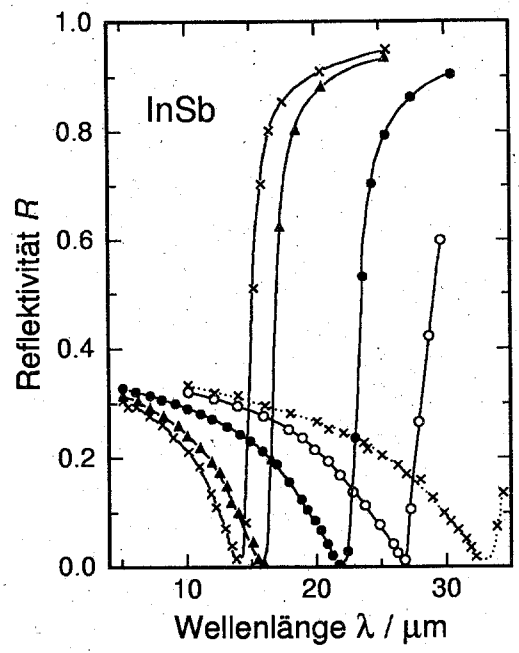


Abb. 11.26: Reflektivität von tellurdotiertem Indiumantimonid an der Plasmakante. Mit zunehmender Dotierung verschiebt sich die Kante zu kleineren Wellenlängen. Die Dichte der Ladungsträger reicht von $3{,}5 \cdot 10^{23}\,\text{m}^{-3}$ bis $4{,}0 \cdot 10^{24}\,\text{m}^{-3}$. (Nach W.G. Spitzer und H.Y. Fan, Phys. Rev. **106** (1957) 882).

Die Abhängigkeit der Plasmafrequenz von der Ladungsträgerkonzentration ist in Abb. 11.26 am Beispiel der Reflektivität des Halbleiters InSb zu sehen. Die starke Dotierung der Proben bewirkt eine hohe Konzentration von Elektronen im Leitungsband. Die Halbleiterprobe ist entartet, so daß die Elektronen wie ein freies Elektronengas behandelt

werden können. Die Plasmafrequenz nimmt mit der Elektronenkonzentration zu, und die Reflexionskante verschiebt sich zu kürzeren Wellenlängen. Dies gilt auch für das Minimum der Reflektivität, das bei $\epsilon = 1$ und somit nach (11.49) bei $\omega = \omega_p \sqrt{\epsilon_\infty/(\epsilon_\infty - 1)}$ auftritt. Das Ansteigen der Reflektivität zu kleinen Wellenlängen hin beruht auf dem großen Realteil der Dielektrizitätskonstanten von Indiumantimonid bei diesen Frequenzen. Da sich aus dem Minimum die Plasmafrequenz und daraus wiederum mit Hilfe von Gleichung (11.48) die effektive Masse der Elektronen ermitteln läßt, werden Reflexionsmessungen bei Halbleitern häufig zur Bestimmung von m^* herangezogen.

Als Beispiel dafür, wie man die Variation der Plasmafrequenz durch Dotierung nutzen und die Reflexions- und Absorptionseigenschaften von Materialien verändern kann, wollen wir noch zinndotierte In_2O_3-Schichten betrachten, deren Verhalten in Abb. 11.27 dargestellt ist.

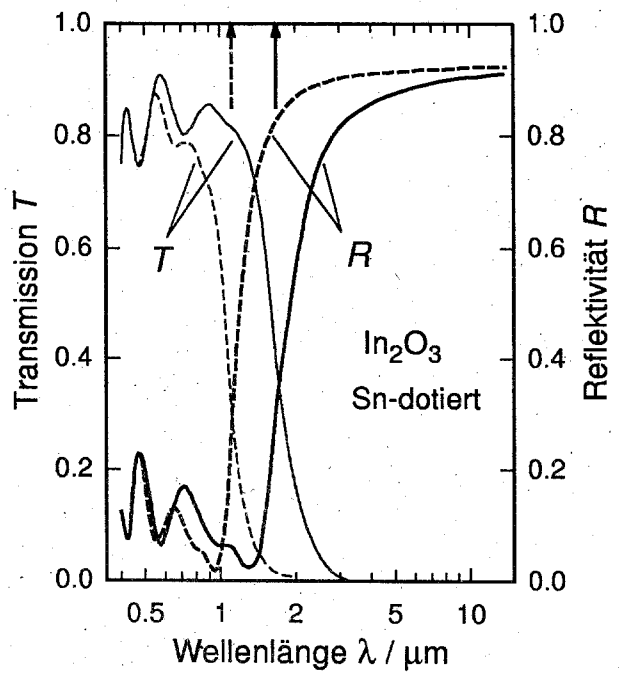


Abb. 11.27: Transmission (T) und Reflektivität (R) von zwei 0,3 μm dicken In_2O_3-Schichten. Die Pfeile kennzeichnen die Lage der Plasmawellenlängen, die aufgrund der unterschiedlichen Zinndotierungen bei verschiedenen Wellenlängen auftreten. Der wellenförmige Verlauf wird durch Interferenzen in der dünnen Schicht verursacht. (Nach G. Franck, E. Kauer, H. Köstlin, Phys. Bl. **34** (1978) 106).

Durch unterschiedlich starkes Dotieren mit Zinn läßt sich die Absorptionskante von In_2O_3 fast nach Belieben einstellen. In dem gezeigten Beispiel liegt die Plasmawellenlänge abhängig von der Elektronendichte von $5 \cdot 10^{26}\,m^{-3}$ bzw. $1{,}3 \cdot 10^{27}\,m^{-3}$ knapp unter 2 μm oder knapp über 1 μm. Da in Indiumoxid die Interbandübergänge erst bei 2,8 eV einsetzen, ist die dotierte Schicht optisch transparent, aber im Infraroten reflektierend. Solche Schichten werden zur Verminderung der Wärmestrahlungsverluste bei Fenstern oder der Wärmeabstrahlung bei Natriumdampflampen eingesetzt. Darüber hinaus können diese elektrisch leitenden Schichten als transparente Elektroden, z.B. bei LC-Anzeigen oder Solarzellen, dienen.

Longitudinale Schwingungen des Elektronengases: Plasmonen

Im Gas freier Elektronen sind longitudinale Schwingungen möglich, die allerdings nicht an elektromagnetische Wellen ankoppeln können. Der Grund für die fehlende Kopplung ist der gleiche wie bei den LO-Phononen. Während bei den transversalen Schwingungen des Elektronengases keine rücktreibenden Kräfte auftreten, sind mit longitudinalen

Schwingungen derartige Kräfte verknüpft. Im Grenzfall $k \to 0$ kann man sich deren Stärke leicht klar machen, denn dort schwingen alle Elektronen einheitlich gegen alle Ionen. Dadurch baut sich bei einer Auslenkung u, wie in Abb. 11.28 angedeutet, an der Begrenzung der Probe eine Flächenladung mit der Dichte $\sigma = neu$ auf, die das elektrische Feld $\mathcal{E} = neu/\epsilon_0\epsilon_\infty$ zur Folge hat.

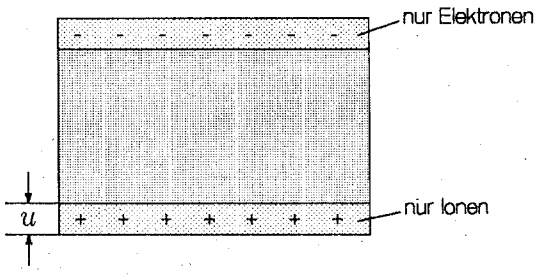


Abb. 11.28: Schematische Darstellung einer Plasmaschwingung mit dem Wellenvektor $k \simeq 0$. Alle Elektronen schwingen in Phase gegen das starre Ionengitter. Die Auslenkung u der Elektronen bewirkt eine Flächenladungsdichte $\sigma = neu$.

Der Faktor ϵ_∞ berücksichtigt auch hier die Polarisierbarkeit der gebundenen Elektronen. Somit lautet die Bewegungsgleichung der freien Elektronen ohne Dämpfung

$$nm\ddot{u}(t) = -ne\mathcal{E}(t) = -\frac{n^2e^2u(t)}{\epsilon_0\epsilon_\infty} \tag{11.51}$$

oder

$$\ddot{u} + \omega_p^2 u = 0\,, \tag{11.52}$$

wobei die Plasmafrequenz ω_p wieder durch Gleichung (11.48) gegeben ist. Dies ist die Bewegungsgleichung eines harmonischen Oszillators mit der Eigenfrequenz ω_p. Die gleiche Schwingungsfrequenz erhält man auch aus der Bedingung, daß die dielektrische Funktion (11.47) bei der Frequenz der longitudinalen Eigenschwingung verschwindet, d.h. daß $\epsilon(\omega_l = \omega_p) = 0$ ist. Wie bei den Phonon-Polaritonen, so stellt auch hier die Frequenz der longitudinalen Schwingungen eine untere Grenze für die Ausbreitung von elektromagnetischen Wellen dar. Ähnlich wie die Phononen eine kohärente Bewegung aller Atome des Gitters darstellen, so handelt es sich bei den Plasmaschwingungen um kohärente, kollektive Anregungen aller Elektronen des Fermi-Gases. Die Amplitude des harmonischen Oszillators, der diese Plasmaschwingung beschreibt, ist quantisiert. Die zugeordneten Anregungen besitzen die Energie $\hbar\omega_p$ und werden *Plasmonen* genannt. Daneben gibt es natürlich noch die Einelektronanregungen, die der Bewegung einzelner, unabhängiger Elektronen zugeordnet werden. Geht man zu endlichen Wellenlängen, d.h. zu $k > 0$, über, so ist die Elektronendichte räumlich moduliert. Aufgrund des Gradienten in der Ladungsdichte erhöhen sich in diesem Fall die Rückstellkräfte, die auf die Elektronen wirken, und die Frequenz der Plasmonen steigt mit dem Wellenvektor an.

Experimentell lassen sich Plasmonen mittels Reflexion von Photonen hoher Energie oder durch Elektronenstreuung untersuchen. Beispielsweise können Metallfilme mit Elektronen durchstrahlt werden, wobei das Elektronengas angeregt wird und so Plasmonen erzeugt werden. In Abb. 11.29 ist das Verlustspektrum gezeigt, das beim Durchtritt von 20 keV Elektronen durch einen 258 nm dicken Aluminiumfilm auftritt. Es ist klar zu erkennen, daß die wechselwirkenden Elektronen einen Energieverlust erfahren, dem ein ganzzahliges Vielfaches der Plasmonenenergie $\hbar\omega_p \approx 15{,}3\,\text{eV}$ entspricht.

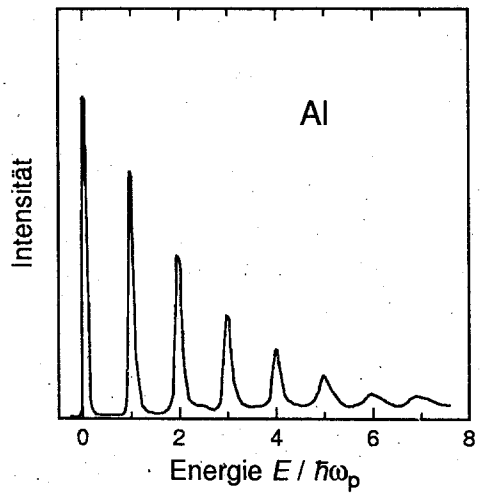


Abb. 11.29: Elektronenverlustspektrum von 20 keV-Elektronen nach der Durchstrahlung einer Aluminiumschicht. Der Energieverlust ist in Einheiten der Plasmonenenergie $\hbar\omega_p = 15{,}3\,\mathrm{eV}$ angegeben. (Nach L. Marton, J.A. Simpson, H.A. Fowler, N. Swanson, Phys. Rev. **126** (1962) 182).

In vielen Streuexperimenten erhält man jedoch kein so eindeutiges Signal, wofür es zwei Gründe gibt. Einerseits treten bei der gleichen Energie oft Interbandübergänge auf, die zu einer starken Reduktion der Plasmonenlebensdauer und somit zu einer starken Verbreiterung des Signals führen. Andererseits treten neben den hier diskutierten *Volumenplasmonen* auch *Oberflächenplasmonen*, bei denen die kollektive Elektronenbewegung an der Oberfläche lokalisiert ist. Diese Schwingungsform ist ebenfalls stabil, besitzt aber eine kleinere Energie, da die begleitenden elektrischen Felder teilweise im Vakuum verlaufen und dort keine Polarisationen hervorrufen können. Bei vielen Metallen können in den Verlustspektren beide Anregungen beobachtet werden.

Im Zusammenhang mit der Existenz der Oberflächenplasmonen sei noch darauf hingewiesen, daß deren Eigenschaften sehr stark von der Beschaffenheit der Oberfläche selbst abhängen. Diese Tatsache nutzt man experimentell und technisch aus. So lassen sich mit Hilfe von Oberflächenplasmonen die dielektrischen Eigenschaften von dünnen Schichten sehr genau vermessen, wenn diese auf einer Metallprobe aufgebracht sind. Oberflächenplasmonen reagieren sehr stark auf Adsorbate und werden daher in der Sensorik zum empfindlichen Nachweis von Substanzen in Gasen und Flüssigkeiten genutzt.

11.5 Übersicht über die Absorptionsprozesse

Zum Schluß dieses Kapitels wollen wir noch einen Blick auf den in Abb. 11.30 gezeigten schematischen Verlauf des Absorptionskoeffizienten eines polaren, stark dotierten Halbleiters werfen. Sehen wir von der Absorption durch permanente Dipole ab, so können in solchen Materialen alle hier diskutierten Absorptionsmechanismen auftreten. Natürlich sind die einzelnen Mechanismen für unterschiedliche Materialien und Stoffklassen mehr oder weniger wichtig. Die Lage der Störstellenabsorption hängt stark von der Natur der Störstellen ab. Hier wurde die Absorption eingezeichnet, die durch Dotieratome (vgl. Abschnitt 4.1) hervorgerufen wird. In der Abbildung sind natürlich nicht alle möglichen Mechanismen berücksichtigt, wie z.B. Magnonenabsorption oder Zyklotronresonanz, dennoch verdeutlicht sie die eingangs gemachte Behauptung, daß die Reaktion der Festkörper auf elektromagnetische Wellen relativ komplex ist.

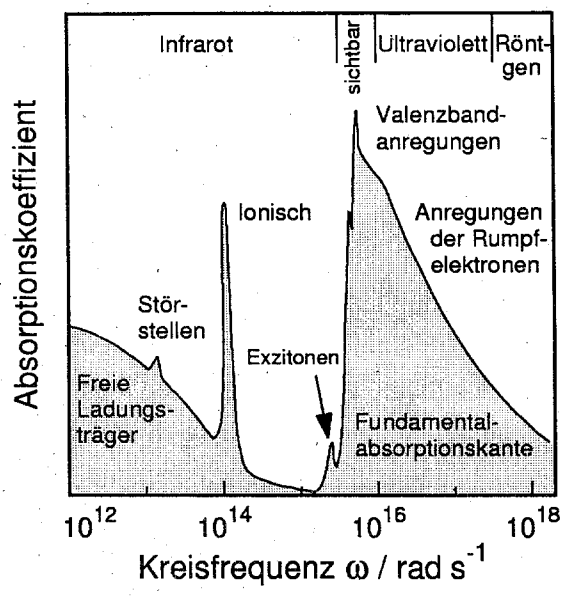


Abb. 11.30: Schematische Darstellung des Absorptionsspektrums eines polaren Halbleiters. (Nach J.R. Hook, H.E. Hall, *Solid State Physics*, Wiley & Sons, New York, 1993).

12. Supraleitung

In Kap. 8 haben wir uns mit der Leitfähigkeit von Metallen befaßt. Da sie durch die Streuung der Elektronen untereinander, an Kristalldefekten und durch Stöße mit Phononen begrenzt wird, ist zu erwarten, daß der ohmsche Widerstand nur in perfekten Kristallen und erst am absoluten Nullpunkt verschwindet. Erstaunlicherweise tritt dieser scheinbar hypothetische Fall bei vielen Metallen bereits bei endlicher, nicht sehr tiefer Temperatur auf. Man bezeichnet dieses Phänomen als *Supraleitung.* Sie wurde bereits 1911 von *H.K. Onnes* entdeckt, als er den Widerstand eines Quecksilberfadens bei tiefer Temperatur bestimmte. Wie in Abb. 12.1 zu sehen ist, nahm der Widerstand beim Kühlen zunächst stetig ab und fiel bei etwa 4,2 K sprungartig auf einen unmeßbar kleinen Wert. Die gleiche Beobachtung, wenn auch bei unterschiedlichen Temperaturen, wurde später bei vielen anderen Substanzen gemacht.

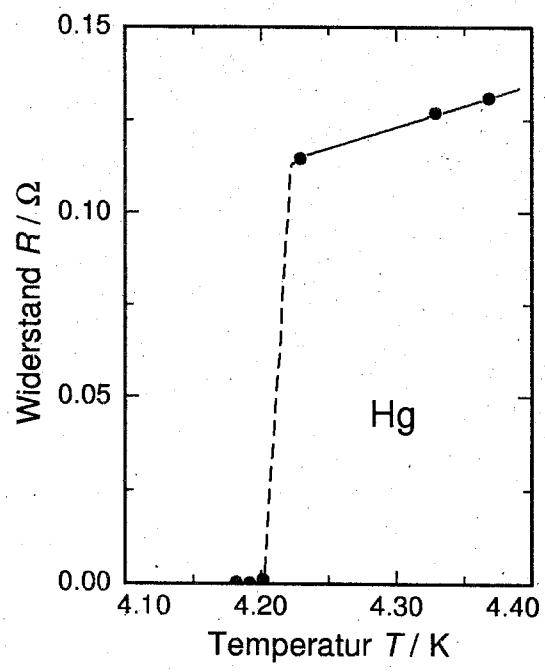


Abb. 12.1: Elektrischer Widerstand eines Quecksilberfadens als Funktion der Temperatur. (Nach K.H. Onnes, Leiden Commun. **124c** (1911)).

12.1 Phänomenologische Beschreibung

Für eine theoretische Behandlung der Supraleitung ist die Beantwortung der Frage sehr wichtig, ob unterhalb der *Sprungtemperatur* T_c der Widerstand tatsächlich Null oder nur sehr klein wird. Eine elegante und sehr empfindliche Prüfung dieser Frage erlauben Dauerstromversuche: Man bringt einen Ring aus supraleitendem Material oberhalb der Sprungtemperatur in ein statisches Magnetfeld, erniedrigt die Temperatur und schaltet das Magnetfeld unterhalb von T_c ab. Bei endlichem Widerstand R wird der dabei induzierte Strom entsprechend der Gleichung $I(t) = I_0 \exp(-Rt/L)$ abklingen, wobei L die

Induktivität des Rings ist. Innerhalb der Meßgenauigkeit bleibt der Strom bei Supraleitern jedoch konstant, und die empfindlichsten Messungen erlauben den Schluß, daß der Widerstand bei T_c um mindestens 14 Größenordnungen fällt. Man geht daher davon aus, daß im supraleitenden Zustand tatsächlich kein elektrischer Widerstand auftritt.

Es sind einige allgemeine empirische Zusammenhänge bekannt, denen entnommen werden kann, in welchen Materialien Supraleitung zu erwarten ist und welchen Wert der Sprungpunkt annehmen sollte. So weiß man, daß vorzugsweise Elemente mit kleinem Atomvolumen supraleitend werden, oder daß bei Legierungen die Elektronendichte eine wichtige Rolle spielt. Quantitative Voraussagen lassen sich hieraus jedoch nicht ableiten.

Überraschenderweise werden die meisten Metalle supraleitend, wobei ihre strukturelle Ordnung ohne Bedeutung ist, d.h. bei den Supraleitern kann es sich um reine Kristalle, um Legierungen oder auch um amorphe Metalle handeln. Daß die Fernordnung keine entscheidende Rolle spielt, sieht man am Verhalten von Wismut, bei dem die amorphe, nicht aber die kristalline Phase supraleitend wird. Eine besonders interessante neue Klasse von Supraleitern stellen die oxidischen Hochtemperatursupraleiter dar, auf die wir am Ende dieses Kapitels eingehen.

Bei elementaren Metallen liegen die Sprungtemperaturen (vgl. Tabelle 12.1) unter 10 K. Bei speziellen Legierungen und metallischen Verbindungen findet man Werte über 20 K. Bei Hochtemperatursupraleitern, deren Sprungtemperatur für vier Vertreter am Tabellenende aufgeführt ist, findet man sogar Werte bis 133 K. Auf den ersten Blick erstaunlich ist die Tatsache, daß „gute" Metalle mit hoher Leitfähigkeit wie Ag, Au, Cu oder Na nicht supraleitend werden. Wichtig für das Verständnis ist weiter, daß magnetische Metalle ebenfalls normalleitend bleiben. Verunreinigungen haben, abgesehen von magnetischen, nur einen schwachen Einfluß auf die Sprungtemperatur. Obwohl die theoretischen Grundlagen der Supraleitung sehr gut verstanden sind, ist es bis heute nicht möglich, Sprungtemperaturen zu berechnen, selbst wenn die Struktur und die Materialparameter der betreffenden Substanzen bekannt sind.

Tabelle 12.1: Sprungtemperatur T_c der Elemente und von einigen Verbindungen und Hochtemperatursupraleitern.

Element	T_c (K)	Element	T_c (K)	Verbindung	T_c (K)
Al	1,19	Pb	7,2	$BaBi_3$	5,69
Be	0,03	Re	1,7	MoC	14,3
Cd	0,55	Ru	0,5	$NbSn_2$	2,6
Ga	1,09	Sn	3,72	Nb_3Sn	18,0
Hf	0,13	Ta	4,39	Nb_3Au	11,3
Hg	4,15	Tc	7,8	Nb_3Ge	23,2
In	3,40	Th	1,37	UPt	0,54
Ir	0,14	Ti	0,39	V_3Si	17,1
La	4,8	Tl	2,39	$Cr_{0,1}Ti_{0,3}V_{0,6}$	5,6
Mo	0,92	U	0,2	$PbTl_{0,27}$	6,43
Nb	9,2	V	5,3	$La_{1,8}Sr_{0,8}CuO_4$	36,2
Np	0,08	W	0,01	$YBa_2Cu_3O_7$	92
Os	0,65	Zn	0,9	$Tl_2Ca_2Ba_2Cu_3O_{10}$	122
Pa	1,3	Zr	0,55	$HgBa_2Ca_2Cu_3O_{8+x}$	133

Meißner-Effekt und London-Gleichung

Dauerstromversuche scheinen zu zeigen, daß sich Supraleiter wie *ideale Leiter* verhalten. Es stellt sich die Frage, ob die elektromagnetischen Eigenschaften von Supraleitern mit den Maxwell-Gleichungen beschrieben werden können, wenn man einfach von einem verschwindenden Widerstand ausgeht. Eine Antwort darauf geben Experimente im Magnetfeld. Wir führen ein Gedankenexperiment mit einer Probe durch, die bei der Sprungtemperatur T_c ideal leitend wird, und vergleichen das Ergebnis mit dem Verhalten, das man bei realen Supraleitern findet. Betrachten wir hierzu eine geschlossene Kurve im Innern unserer Proben, so besagt das Induktionsgesetz, daß eine zeitliche Änderung des Magnetflusses mit dem Auftreten eines elektrisches Feld längs dieser Kurve verbunden ist. Da aber in idealen Leitern elektrische Felder perfekt kurzgeschlossen werden, kann sich kein derartiges Feld aufbauen und somit kann sich auch der eingeschlossene Magnetfluß nicht ändern. Übertragen wir diese Überlegung auf die gesamte Probe, so ergeben sich Konsequenzen, wie sie in den beiden Abbildungen 12.2 und 12.3 bildhaft dargestellt sind.

Im oberen Teil der beiden Abbildungen ist jeweils die angenommene zeitliche Variation von Temperatur und Magnetfeld dargestellt. Darunter sind die (ellipsoidförmigen) Proben und die auftretenden Feldlinien skizziert. In Abb. 12.2 wird angenommen, daß beide Proben zunächst ohne Magnetfeld auf eine Temperatur unterhalb von T_c abgekühlt werden und anschließend ein Magnetfeld angelegt wird. Idealer Leiter und Supraleiter reagieren auf gleiche Weise: Das Magnetfeld kann nicht eindringen, da an der Probenoberfläche Ströme induziert werden, die das Feld abschirmen. Da wegen des fehlenden Widerstands die Ströme nicht gedämpft werden, klingen die Abschirmströme nicht ab, und das Probeninnere bleibt dauerhaft magnetfeldfrei. Schalten wir das Feld ab, so wird in beiden Fällen der Ausgangszustand wieder hergestellt. Offensichtlich erlaubt die hier geschilderte Versuchsführung keine Unterscheidung zwischen idealen Leitern und Supraleitern.

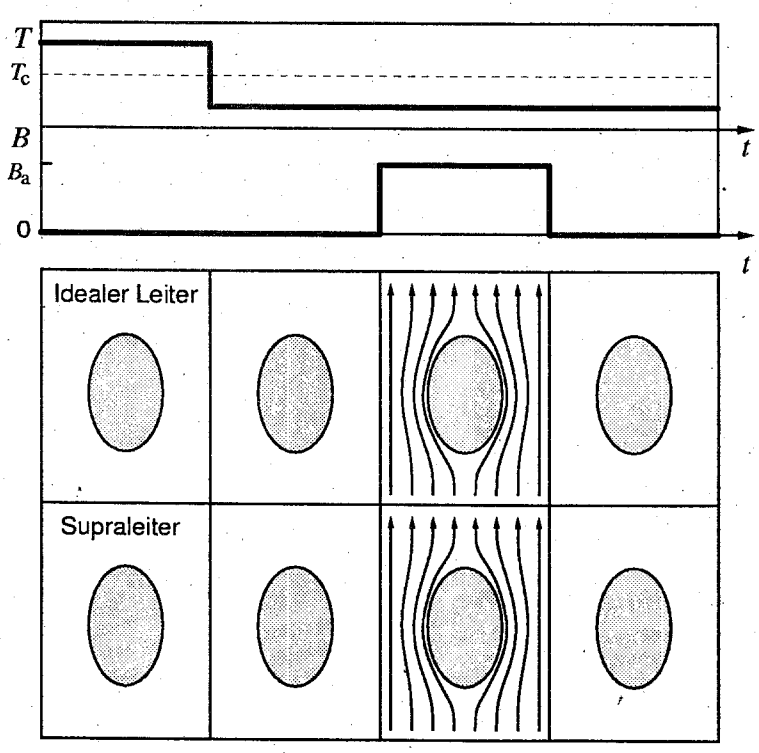


Abb. 12.2: Idealer Leiter und Supraleiter im Magnetfeld. Im oberen Bildteil ist der zeitliche Verlauf von Temperatur und angelegtem Magnetfeld dargestellt. Das Magnetfeld B_a wird unterhalb der Sprungtemperatur T_c angelegt. Während des gesamten Experiments bleiben beide Proben feldfrei.

Nun legen wir, wie in Abbildung 12.3 dargestellt, das Magnetfeld vor dem Unterschreiten der Sprungtemperatur T_c an. Beide Proben werden vom Feld durchsetzt, da bei endlichem Widerstand die Abschirmströme abklingen. Beim Durchgang durch T_c ändert sich dieser Zustand im idealen Leiter nicht; der Magnetfluß bleibt konstant und die Probe vom Magnetfeld durchdrungen. Schalten wir nun das Feld im ideal leitenden Zustand ab, so werden entsprechend der Lenzschen Regel Ströme angeworfen, die eine Änderung des Magnetflusses in der Probe verhindern. Der ideale Leiter besitzt nun ein permanentes magnetisches Moment. Ein Supraleiter verhält sich dagegen völlig anders: Beim Unterschreiten des Sprungpunkts wird das Feld aus der Probe verdrängt. Nicht nur die Feldänderung verschwindet, wie es die Maxwell-Gleichungen fordern, sondern im Probeninnern sogar das Feld selbst. Schalten wir das anliegende Feld ab, so ist der Ausgangszustand wieder hergestellt. Abhängig von der Versuchsführung treten beim idealen Leiter zwei verschiedene Endzustände auf, während beim Supraleiter der Endzustand eindeutig ist. Die Magnetfeldverdrängung am Sprungpunkt, der sogenannte *Meißner-Effekt*, wurde 1933 von *Meißner* und *Ochsenfeld* entdeckt.

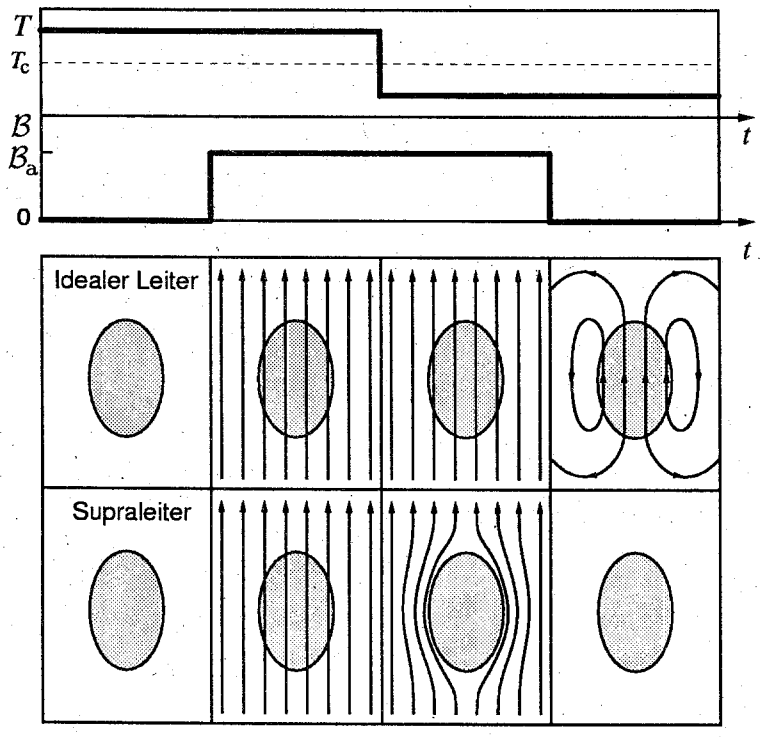


Abb. 12.3: Idealer Leiter und Supraleiter im Magnetfeld. Im oberen Bildteil ist der zeitliche Verlauf von Temperatur und angelegtem Feld dargestellt. Das Magnetfeld wird oberhalb der Sprungtemperatur angelegt und dringt in beide Proben ein. Beim Unterschreiten von T_c bleibt das Magnetfeld im idealen Leiter gefangen, während es aus dem Supraleiter verdrängt wird. Nach dem Abschalten des Felds stellen sich zwei unterschiedliche Endzustände ein.

Aus dem Verschwinden des Felds $\boldsymbol{B}_i$ im Probeninnern läßt sich der Zusammenhang zwischen Magnetfeld und Magnetisierung herleiten: Aus $\boldsymbol{B}_i = \mu_0(\boldsymbol{H} + \boldsymbol{M}) = 0$ folgt, daß die Magnetisierung $\boldsymbol{M}$ des Supraleiters das angelegte Magnetfeld $\boldsymbol{H}$ exakt kompensiert. Mit $\boldsymbol{H} = -\boldsymbol{M}$ ergibt sich daraus die Suszeptibilität $\chi = M/H = -1$. Daraus folgt, daß Supraleiter nicht nur ideale Leiter sondern auch *ideale Diamagneten* sind.

Wie könnte man die ideale Leitfähigkeit und den idealen Diamagnetismus in den Maxwell-Gleichungen berücksichtigen? Die erste Forderung ist leicht durch Weglassen des Stoßterms in der Bewegungsgleichung der Elektronen zu erfüllen, denn nach Gleichung (8.67) gilt $m\dot{\boldsymbol{v}} = -e\boldsymbol{\mathcal{E}} - m\boldsymbol{v}/\tau$. Setzen wir nun $\tau = \infty$, so ergibt sich für die

Zeitableitung der Stromdichte

$$\frac{\mathrm{d}\boldsymbol{j}}{\mathrm{d}t} = \frac{n_\mathrm{s} e^2}{m} \boldsymbol{\mathcal{E}} \,. \tag{12.1}$$

Nicht die Stromdichte, wie beim Ohmschen Gesetz, sondern ihre zeitliche Änderung ist proportional zur elektrischen Feldstärke. Die Gleichung erweckt den Eindruck, daß die Stromdichte beliebig anwachsen könnte. Die unendlich hohe Gleichstromleitfähigkeit der Supraleiter hat aber zur Folge, daß im stationären Fall in der Probe keine Potentialdifferenzen auftreten und somit die zeitliche Änderung des Stroms verschwindet. Der Stromfluß ist daher konstant und durch die Stromquelle vorgegeben. Es soll noch angemerkt werden, daß in der Gleichung nicht die Elektronendichte n, sondern die Dichte n_s der supraleitenden Ladungsträger steht. Damit lassen wir hier offen, ob alle freien Elektronen zum Suprastrom beitragen. Die Bedeutung von n_s wird bei der Diskussion der mikroskopischen Ursache der Supraleitung in Abschnitt 12.2 angesprochen.

Nun setzen wir Gleichung (12.1) in die Maxwell-Gleichung $\mathrm{rot}\,\boldsymbol{\mathcal{E}} = -\partial \boldsymbol{\mathcal{B}}/\partial t$ ein und erhalten

$$\frac{\partial}{\partial t}\left(\mathrm{rot}\,\boldsymbol{j} + \frac{n_\mathrm{s} e^2}{m}\boldsymbol{\mathcal{B}}\right) = 0 \,. \tag{12.2}$$

Diese Gleichung gilt für alle idealen Leiter und besagt, daß der Magnetfluß durch eine beliebige Fläche innerhalb der Probe zeitlich unveränderlich ist. Da nach dem Meißner-Effekt in einem Supraleiter aber der Fluß und nicht nur seine zeitliche Ableitung verschwindet, muß der Klammerausdruck selbst verschwinden. Damit erhalten wir die *London-Gleichung*

$$\mathrm{rot}\,\boldsymbol{j} = -\frac{n_\mathrm{s} e^2}{m}\boldsymbol{\mathcal{B}} \,. \tag{12.3}$$

Sie verknüpft den Stromfluß mit der magnetischen Induktion im Supraleiter.

Bei der Beschreibung des Meißner-Effekts haben wir so argumentiert, als würde das Magnetfeld vollständig aus dem Supraleiter ferngehalten. Dies würde aber an der Probenoberfläche unendlich hohe Abschirmströme erfordern. Tatsächlich dringt das Magnetfeld etwas ein. Durch Einsetzen von (12.3) in die Maxwell-Gleichung $\mathrm{rot}\,\boldsymbol{\mathcal{B}} = \mu_0 \boldsymbol{j}$ lassen sich Stromdichteverteilung und Magnetfeldverlauf in der Probe berechnen. Nimmt der Supraleiter, wie in Abb. 12.4 angedeutet, den Halbraum $x > 0$ ein und liegt ein Magnetfeld in z-Richtung an, dann ergibt die Rechnung

$$j_y(x) = j_0\, \mathrm{e}^{-x/\lambda_\mathrm{L}} \,, \tag{12.4}$$

$$\mathcal{B}_z(x) = \mathcal{B}_0\, \mathrm{e}^{-x/\lambda_\mathrm{L}} \,, \tag{12.5}$$

wobei die *Londonsche Eindringtiefe* λ_L durch

$$\lambda_\mathrm{L} = \sqrt{\frac{m}{\mu_0 n_\mathrm{s} e^2}} \tag{12.6}$$

gegeben ist. Sowohl Stromdichte als auch Magnetfeld fallen im Supraleiter exponentiell ab, wobei λ_L als charakteristische Länge auftritt. Mit Hilfe der anfangs benutzten Maxwell-Gleichung finden wir für die maximale Stromdichte $j_0 = \mathcal{B}_0/\mu_0\lambda_\mathrm{L}$.

Wie wir noch sehen werden, nimmt die Dichte n_s der supraleitenden Ladungsträger mit zunehmender Temperatur ab und geht am Sprungpunkt gegen den Wert Null. Die Eindringtiefe wird deshalb bei T_c unendlich groß. Nehmen wir für eine einfache Abschätzung

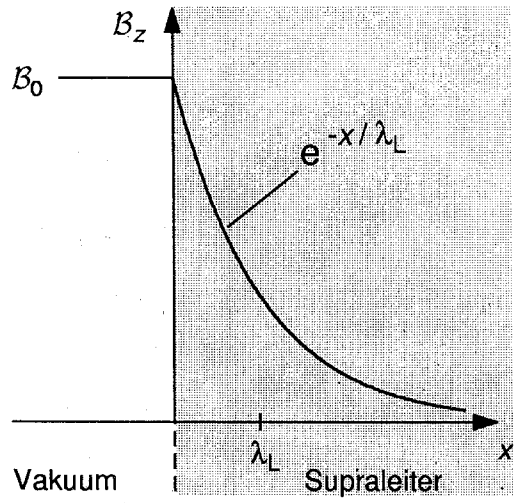


Abb. 12.4: Abfall des Magnetfelds im Innern eines Supraleiters, der den Halbraum $x > 0$ einnimmt. Der exponentielle Abfall der Feldstärke wird durch die Londonsche Eindringtiefe λ_L bestimmt.

an, daß am absoluten Nullpunkt alle Elektronen zur Supraleitung beitragen und setzen $n_s = 10^{29}$ m^{-3}, so erhalten wir für die Eindringtiefe den Wert $\lambda_L \approx 15$ nm. Das Magnetfeld fällt also im Supraleiter sehr rasch ab. Die Eindringtiefe läßt sich durch Messung der Suszeptibilität[1] bestimmen. Überraschenderweise findet man aber in vielen Fällen einen größeren Wert als erwartet. Der Grund hierfür ist nicht ein grundsätzlicher Fehler in der Londonschen Theorie, sondern beruht darauf, daß sich Strom und Magnetfeld innerhalb sehr kleiner Abstände drastisch ändern. Ein ähnliches Problem taucht auch bei der Anwendung des Ohmschen Gesetzes bei hohen Frequenzen auf. Sobald dort die freie Weglänge der Elektronen vergleichbar mit der Wellenlänge der elektromagnetischen Welle wird, muß das elektrische Feld bei der Berechnung des Widerstands über einen Bereich von der Größenordnung der freien Weglängen gemittelt werden. Dieser Gedanke einer *nicht-lokalen* Beschreibung, in der die Stromdichte nicht durch den lokalen Wert des Magnetfelds sondern durch einen gemittelten bestimmt wird, wurde von *Pippard* ausgearbeitet und führt zu guter Übereinstimmung mit dem Experiment.

Kritisches Magnetfeld und thermodynamische Eigenschaften

In hohen Magnetfeldern werden die supraleitenden Metalle wieder normalleitend. Bereits eine einfache thermodynamische Überlegung zeigt, daß es eine *kritische Magnetfeldstärke* $\mathcal{B}_c$ geben muß. Maßgeblich für die magnetischen Eigenschaften ist die *freie Enthalpie* $G(T, p, \mathcal{B}) = U - TS + pV - V\boldsymbol{M}\cdot\boldsymbol{\mathcal{B}}$. Betrachten wir zunächst den Übergang von der Normal- zur Supraleitung ohne Magnetfeld[2]. Da der supraleitende Zustand unterhalb von T_c stabiler als der normalleitende Zustand ist, muß $G_n(\mathcal{B} = 0, T) > G_s(\mathcal{B} = 0, T)$ sein. Die freie Enthalpie ändert sich jedoch mit dem Magnetfeld. Für die Änderung $\mathrm{d}G$ finden wir unter Berücksichtigung der inneren Energie $\mathrm{d}U = T\,\mathrm{d}S - p\,\mathrm{d}V + V\mathcal{B}\,\mathrm{d}M$ den Ausdruck

$$\mathrm{d}G = -S\,\mathrm{d}T + V\mathrm{d}p - VM\,\mathrm{d}\mathcal{B}\,. \tag{12.7}$$

[1] Um ein möglichst günstiges Verhältnis von feldfreiem Volumen zu magnetfelddurchdrungener Oberflächenschicht zu erhalten, führt man solche Messungen häufig an kleinen Kügelchen durch.

[2] Da die Magnetisierung $\boldsymbol{M}$ stets parallel oder antiparallel zu $\boldsymbol{\mathcal{B}}$ verläuft, können wir auch in diesem Kapitel die Vektorkennzeichnung weglassen.

In der weiteren Diskussion setzen wir konstanten Druck voraus, so daß der Term $V\mathrm{d}p$ verschwindet. Um die Komplikation durch das Entmagnetisierungsfeld zu umgehen, dessen Auswirkung wir im Falle des elektrischen Felds in Abschnitt 11.1 diskutiert haben, untersuchen wir das Verhalten eines langen Zylinders, dessen Achse parallel zum Magnetfeld verläuft, da in diesem Fall das Entmagnetisierungsfeld verschwindet. Setzen wir die Suszeptibilität $\chi = -1$ für den Supraleiter ein, vereinfacht sich Gleichung (12.7) bei konstantem Druck zu

$$\mathrm{d}G_\mathrm{s} = -S\,\mathrm{d}T + \frac{V}{\mu_0}\mathcal{B}\,\mathrm{d}\mathcal{B}\,. \tag{12.8}$$

Die Gleichung läßt sich leicht integrieren, und wir erhalten

$$G_\mathrm{s}(\mathcal{B},T) = G_\mathrm{s}(0,T) + \int_0^{\mathcal{B}} \frac{V}{\mu_0}\,\mathcal{B}'\,\mathrm{d}\mathcal{B}' = G_\mathrm{s}(0,T) + \frac{V\mathcal{B}^2}{2\mu_0}\,. \tag{12.9}$$

Das Verdrängen des Magnetfelds erhöht offensichtlich die freie Enthalpie des Supraleiters.

Für unmagnetische normalleitende Metalle ist der magnetfeldabhängige Anteil der freien Enthalpie sehr klein (vgl. Abschnitt 10.1), da die freien Elektronen nur über die Pauli-Spinsuszeptibilität und den Landau-Diamagnetismus dazu beitragen. Wir vernachlässigen diesen Anteil und schreiben $G_\mathrm{n}(\mathcal{B},T) = G_\mathrm{n}(0,T)$. Mit zunehmendem Feld steigt die freie Enthalpie G_s des Supraleiters an und erreicht schließlich den Wert G_n des Normalleiters. Die Energie, die durch die Ausbildung der supraleitenden Phase gewonnen wurde, wird dann gerade durch die Energie kompensiert, die für die Feldverdrängung aufzubringen ist. Die supraleitende Phase wird dann instabil, und es erfolgt ein Übergang zur Normalleitung. In Abb. 12.5 ist der schematische Verlauf der Magnetisierung einer supraleitenden Probe als Funktion des äußeren Felds aufgetragen. Bis zur *kritischen Feldstärke* $\mathcal{B}_\mathrm{c}$ ist $M = -\mathcal{B}/\mu_0$, dann bricht die Supraleitung zusammen, und die Magnetisierung fällt auf den kleinen Wert des Normalleiters. Da am Übergang vom Supra- zum Normalleiter die freien Enthalpien der beiden Phasen gleich sind, d.h. $G_\mathrm{s}(\mathcal{B}_\mathrm{c},T) = G_\mathrm{n}(\mathcal{B}_\mathrm{c},T) = G_\mathrm{n}(0,T)$ ist, folgt aus Gleichung (12.9)

$$G_\mathrm{n}(0,T) - G_\mathrm{s}(0,T) = \frac{V\mathcal{B}_\mathrm{c}^2}{2\mu_0}\,. \tag{12.10}$$

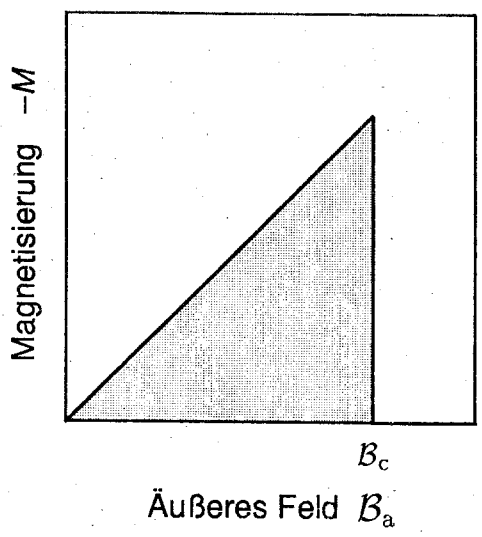


Abb. 12.5: Magnetisierung einer supraleitenden Probe als Funktion des äußeren Magnetfeldes $\mathcal{B}_\mathrm{a}$. Bei der kritischen Feldstärke $\mathcal{B}_\mathrm{c}$ wird die Probe normalleitend, und die Magnetisierung verschwindet.

Der Wert des kritischen Felds wird durch die Differenz der freien Enthalpien bestimmt, die beim Übergang vom normal- zum supraleitenden Zustand bei verschwindender Feldstärke auftritt. Diese Energiedifferenz spielt eine wichtige Rolle in der Theorie der Supraleitung und wird aus Gründen, die wir noch diskutieren werden, als *Kondensationsenergie* G_{kon} bezeichnet. Wie wir in Abschnitt 12.2 noch sehen werden, ist mit dem kritischen Magnetfeld eine kritische Dichte der Abschirmströme verbunden, so daß auch ein zu hoher Strom die Supraleitung in einem Draht zerstören kann.

Die kritische Feldstärke hängt von der Temperatur ab. In Abb. 12.6 ist der Verlauf für eine Reihe von Metallen aufgetragen, der sich näherungsweise durch die Gleichung

$$\mathcal{B}_{\mathrm{c}}(T) = \mathcal{B}_{\mathrm{c}}(0)\left[1 - \left(\frac{T}{T_{\mathrm{c}}}\right)^2\right] \tag{12.11}$$

ausdrücken läßt. Wie dem Bild zu entnehmen ist, ist in vielen Fällen $\mathcal{B}_{\mathrm{c}}(0) \propto T_{\mathrm{c}}$, wobei die Proportionalitätskonstante im Bereich von 0,01 TK^{-1} liegt. Damit errechnet sich als typische Kondensationsenergie der kleine Wert von $G_{\mathrm{kon}}/V \approx 40\,T_{\mathrm{c}}^2\ \mathrm{Jm}^{-3}$.

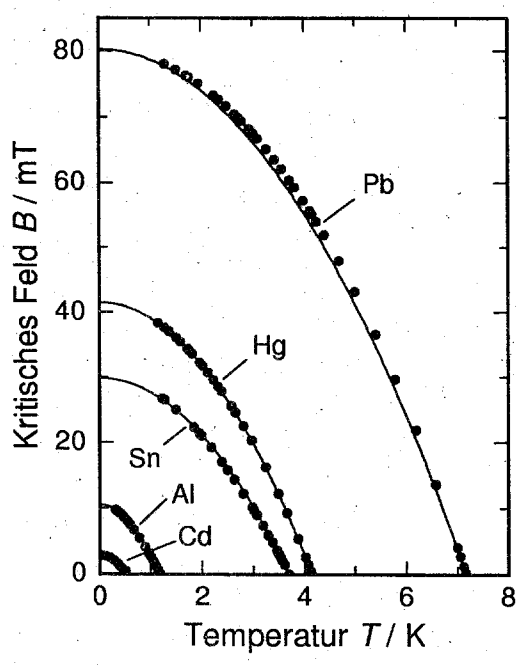


Abb. 12.6: Kritische Magnetfeldstärken als Funktion der Temperatur. Die eingezeichneten Linien geben den Verlauf wieder, den man nach Gleichung (12.11) erwartet. (Die Daten wurden verschiedenen Veröffentlichungen entnommen).

Aus der Enthalpiedifferenz (12.10) lassen sich über die Beziehungen $S = -(\partial G/\partial T)$ und $C = T\,(\partial S/\partial T)$ sofort Aussagen zum unterschiedlichen Verhalten von Entropie und spezifischer Wärme im normal- und supraleitenden Zustand machen. Die Differentiationen führen zu den Ausdrücken

$$\Delta S = S_{\mathrm{s}} - S_{\mathrm{n}} = \frac{V\mathcal{B}_{\mathrm{c}}}{\mu_0}\frac{\mathrm{d}\mathcal{B}_{\mathrm{c}}}{\mathrm{d}T} \tag{12.12}$$

und

$$\Delta C = C_{\mathrm{s}} - C_{\mathrm{n}} = \frac{VT}{\mu_0}\left[\mathcal{B}_{\mathrm{c}}\frac{\mathrm{d}^2\mathcal{B}_{\mathrm{c}}}{\mathrm{d}T^2} + \left(\frac{\mathrm{d}\mathcal{B}_{\mathrm{c}}}{\mathrm{d}T}\right)^2\right]. \tag{12.13}$$

Aus einem Vergleich zwischen den experimentellen Daten der Abb. 12.6 und den beiden Gleichungen lassen sich einige wichtige Schlüsse ziehen.

Für $T \to T_c$ verschwindet das kritische Feld und damit die Entropiedifferenz. Da bei der kritischen Temperatur $d\mathcal{B}_c/dT \neq 0$ ist, weisen bei T_c die spezifischen Wärmen von Normal- und Supraleiter unterschiedliche Werte auf, d.h. es tritt eine Diskontinuität auf. In Abb. 12.7 ist die spezifische Wärme von Aluminium im normal- und supraleitenden Zustand gezeigt. Wir werden in Abschnitt 12.2 nochmals auf dieses Bild zurückkommen. Augenblicklich ist nur wichtig, daß die spezifische Wärme bei T_c einen deutlichen Sprung macht. Weiterhin entnehmen wir der Gleichung (12.12), daß beim Übergang die Entropiedifferenz verschwindet, d.h. daß keine latente Wärme $\Delta Q = T_c \Delta S$ auftritt. Der Phasenübergang ist daher von 2. Ordnung. Dies ist der Ausgangspunkt einer phänomenologischen Theorie der Supraleitung von *Ginzburg* und *Landau*, auf die wir aber nicht näher eingehen, wenn wir auch einige Ergebnisse dieser Theorie später benutzen werden.

Im Einklang mit dem 3. Hauptsatz der Thermodynamik verschwinden für $T \to 0$ die beiden ersten Ableitungen von $\mathcal{B}_c$ nach der Temperatur und damit auch die Größen ΔS und ΔC. Im Bereich $0 < T < T_c$ ist $d\mathcal{B}_c/dT < 0$ und damit $\Delta S < 0$. Das Vorzeichen der Entropiedifferenz besagt, daß die Ordnung im supraleitenden Zustand größer als im normalleitenden ist. Zusätzlich kann man daraus schließen, daß beim Übergang im Magnetfeld $\Delta S \neq 0$ ist und daher latente Wärme auftritt. Dies bedeutet, daß im Magnetfeld der Phasenübergang von 1. Ordnung ist.

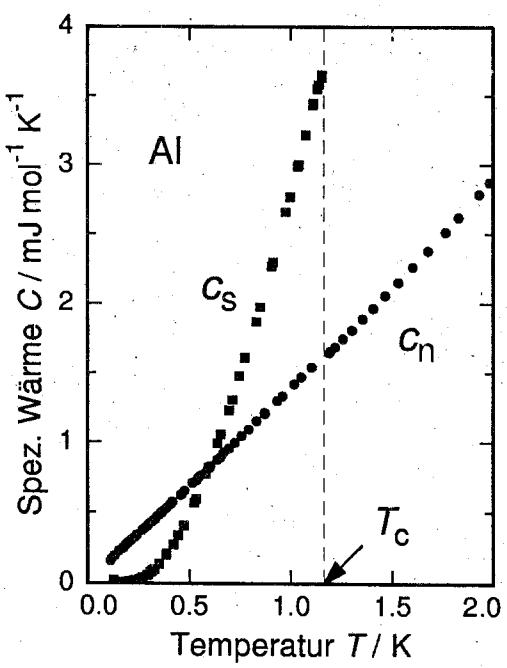


Abb. 12.7: Spezifische Wärme von Aluminium im normal- und supraleitenden Zustand. Durch ein äußeres Magnetfeld von etwa 0,05 T wurde zur Messung des Verhaltens im normalleitenden Zustand die Supraleitung zerstört. Der Beitrag des Gitters spielt nur eine untergeordnete Rolle. (Nach N.E. Phillips, Phys. Rev. **114** (1959) 676).

12.2 Mikroskopische Beschreibung

Wir wollen nun auf die 1957 von *Bardeen*, *Cooper* und *Schrieffer* formulierte mikroskopische Beschreibung der Supraleitung, die sogenannte BCS-Theorie, eingehen. Vorher existierten bereits phänomenologische Theorien, die vor allem in den 30er Jahren von den Brüdern *F.* und *H. London*, von *Gorter* und *Casimir* und um 1950 von *Ginzburg* und *Landau* entwickelt wurden.

Cooper-Paare

Die Ausarbeitung einer mikroskopischen Theorie der Supraleitung erwies sich als sehr schwierig. Natürlich erfolgte ihre Entwicklung in mehreren Schritten, deren wichtigste wir hier in kurzer Form nachvollziehen wollen.

Welche Anforderung muß an die Theorie der Supraleitung gestellt werden? Natürlich muß die Erklärung eines Phänomens, das bei völlig unterschiedlich aufgebauten Festkörpern zu finden ist, auf sehr grundlegenden Annahmen beruhen muß. Bei Normalleitern kommt keine unendlich hohe Leitfähigkeit zustande, weil das Anregungsspektrum der freien Elektronen den Transfer von beliebig kleinen Energien zwischen den Elektronen und dem Gitter erlaubt. Es liegt daher die Vermutung nahe, daß das Elektronenspektrum der Supraleiter gegenüber dem der Normalleiter modifiziert ist. Ein möglicher Mechanismus wurde 1956 von *Cooper* vorgeschlagen. Er konnte zeigen, daß bei einer (beliebig kleinen) attraktiven Wechselwirkung zwischen zwei Elektronen der Grundzustand des Fermi-Gases nicht mehr stabil ist und die Energie des Paars abgesenkt wird. Mit anderen Worten ausgedrückt, bedeutet dies folgendes: Bringt man zwei Elektronen, die nicht wechselwirken, auf die Oberfläche des „Fermi-Sees", d.h. bei der Energie E_F in den Festkörper ein, so bleiben sie dort, weil die tiefer liegenden Plätze bereits besetzt sind. Herrscht dagegen zwischen den beiden eine attraktive Wechselwirkung, so können die beiden Elektronen ihre Energie absenken und bildlich gesprochen „in den See eintauchen". Aufgrund der Coulomb-Abstoßung erscheint aber eine anziehende Wechselwirkung zwischen zwei Elektronen zunächst nur schwer vorstellbar.

Seit 1950 war bekannt, daß die Sprungtemperatur von der Masse der Atome und somit von den Gittereigenschaften abhängt. Die Beobachtung des *Isotopeneffektes* war für das mikroskopische Verständnis der Supraleitung von grundlegender Bedeutung. Ein besonders gutes Beispiel hierfür sind Messungen von T_c an unterschiedlichen Gemischen von Zinnisotopen, die in Abb. 12.8 wiedergegeben sind. Die Sprungtemperatur ist proportional zur Wurzel aus der Ionenmasse und damit proportional zur Debye-Frequenz der untersuchten Proben (vgl. Gleichungen (5.19) und (5.56) in Kap. 5).

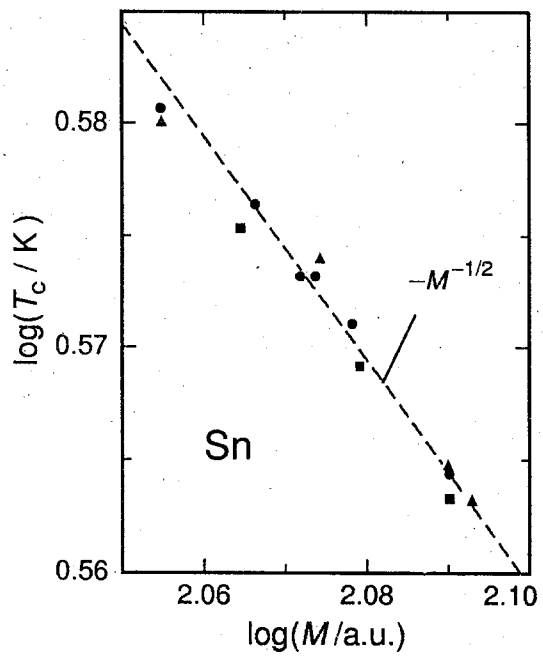


Abb. 12.8: Isotopeneffekt von Zinn. Die Sprungtemperatur ist als Funktion der Isotopenmasse im logarithmischen Maßstab aufgetragen. Die Daten stammen aus verschiedenen Veröffentlichungen und enthalten Messungen an isotropenreinen und gemischten Proben.

Daß das Gitter eine Anziehung zwischen Elektronen vermitteln kann, läßt sich in einem groben Bild folgendermaßen veranschaulichen: Die Elektronen fliegen durch das Gitter und ziehen im Vorbeiflug die positiv geladenen Ionen an. Dadurch bildet sich im „Kielwasser" eine positive Ladungswolke, die ihrerseits andere Elektronen anziehen kann. Der Kraftstoß beim Vorbeiflug eines Elektrons treibt zwar die Ionen in Richtung der Elektronenbahn, aber erst nach einer viertel Schwingungsperiode der Ionen bildet sich die höchste positive Ladungsdichte aus. In der Zwischenzeit hat das erste Elektron etwa 100 nm zurückgelegt. Die „verzögerte" Reaktion der Ionen hat zur Folge, daß die wechselwirkenden Elektronen relativ weit voneinander entfernt sind und damit die Coulomb-Abstoßung relativ schwach ist.

Bei der theoretischen Beschreibung des Wechselwirkungspotentials benutzt man das Konzept des Teilchenaustausches. Die beiden wechselwirkenden Elektronen mit den Wellenvektoren $\boldsymbol{k}_1$ und $\boldsymbol{k}_2$ tauschen *virtuelle Phononen* mit dem Wellenvektor $\boldsymbol{q}$ aus. Nach dem Austauschprozeß besitzen die wechselwirkenden Elektronen die Wellenvektoren $\boldsymbol{k}_1' = \boldsymbol{k}_1 + \boldsymbol{q}$ bzw. $\boldsymbol{k}_2' = \boldsymbol{k}_2 - \boldsymbol{q}$. Der Gesamtimpuls $\boldsymbol{K}$ bleibt jedoch erhalten, und es gilt $\boldsymbol{k}_1 + \boldsymbol{k}_2 = \boldsymbol{k}_1' + \boldsymbol{k}_2' = \boldsymbol{K}$. Am absoluten Nullpunkt sind alle Zustände unterhalb der Fermi-Energie besetzt. Daher sind für die beiden wechselwirkenden Elektronen nur Zustände oberhalb von E_F zugänglich, und die Wechselwirkung spielt sich zwischen den beiden Energien E_F und $E_\mathrm{F} + \hbar\omega_\mathrm{D}$ ab, wobei $\hbar\omega_\mathrm{D}$ für die Energie der Phononen mit der Debye-Frequenz steht. Im $\boldsymbol{k}$-Raum entspricht dieser Energiebereich einer Kugelschale mit der Dicke $\delta k = (m\omega_\mathrm{D}/\hbar k_\mathrm{F})$. Diese Einschränkung ist in Abb. 12.9 skizziert. Da nur Elektronenpaare die Impulserhaltung erfüllen, deren Wellenvektoren im dunkelgetönten Übergangsbereich anfangen bzw. enden, ist anschaulich klar, daß für $\boldsymbol{K} = 0$ der Phononenaustausch mit größtmöglicher Wahrscheinlichkeit erfolgt. Dann ist den beteiligten Elektronen nicht nur ein kleiner Ausschnitt, sondern die ganze Kugelschale zugänglich. Damit ergibt sich die wichtige Aussage, daß für die Wellenvektoren der beiden Elektronen der *Cooper-Paare* die Bedingung $\boldsymbol{k}_1 = -\boldsymbol{k}_2 = \boldsymbol{k}$ erfüllt sein muß. Im weiteren Verlauf der Diskussion werden wir ein derartiges Paar häufig mit dem Symbol $(\boldsymbol{k}, -\boldsymbol{k})$ kennzeichnen.

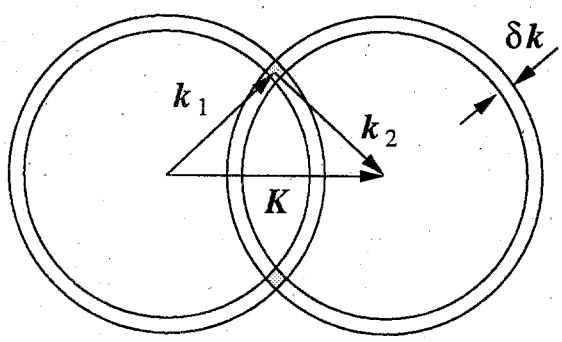


Abb. 12.9: Skizze zur Veranschaulichung der Impulserhaltung bei der Paarwechselwirkung. Die Impulserhaltung ist nur erfüllt, wenn die Wellenvektoren der beteiligten Elektronen innerhalb der beiden Kugelschalen liegen.

Wie beim Wasserstoffmolekül, an dem ebenfalls zwei Elektronen beteiligt sind, ist zur Beschreibung auch hier eine Zweiteilchen-Wellenfunktion $\Psi(\boldsymbol{r}_1, \boldsymbol{r}_2)$ erforderlich, für welche die stationäre Schrödinger-Gleichung das folgende Aussehen hat:

$$\left[-\frac{\hbar^2}{2m}(\Delta_1 + \Delta_2) + \widetilde{V}(\boldsymbol{r}_1, \boldsymbol{r}_2)\right] \Psi(\boldsymbol{r}_1, \boldsymbol{r}_2) = E\,\Psi(\boldsymbol{r}_1, \boldsymbol{r}_2)\,. \tag{12.14}$$

Das Potential $\widetilde{V}(\boldsymbol{r}_1, \boldsymbol{r}_2)$ ist hauptsächlich durch die soeben diskutierte attraktive Wechselwirkung der Elektronen gegeben. Hinzu kommt noch die Coulomb-Abstoßung zwischen den beiden Elektronen, die durch Abschirmeffekte wesentlich erniedrigt wird. Der tatsächliche Potentialverlauf ist nur in groben Zügen bekannt, ist aber auch für unsere weitere Diskussion ohne große Bedeutung. Da es sich nur um eine kleine Störung des freien Elektronengases handelt, stellen wir im Lösungsansatz die Wellenfunktion $\Psi(\boldsymbol{r}_1, \boldsymbol{r}_2)$ als eine Linearkombination von Einteilchen-Funktionen dar, für die wir näherungsweise ebene Wellen ansetzen. Der Übergang zu Bloch-Funktionen wäre ohne konzeptionelle Komplikationen möglich, führt aber zu unübersichtlichen Ausdrücken. Da die Elektronen des Paars entgegengesetzte Impulse besitzen, der Schwerpunkt also in Ruhe ist, gehen in die Beschreibung nur Relativkoordinaten ein. Die Struktur der Wellenfunktion wird deutlich, wenn wir das Paar mit dem Wellenvektor $\boldsymbol{k} = \boldsymbol{k}_1 = -\boldsymbol{k}_2$ herausgreifen, es durch ebene Wellen darstellen und die Zweiteilchen-Funktion bilden: $\Psi = A \exp(\mathrm{i}\boldsymbol{k}_1 \cdot \boldsymbol{r}_1) \cdot \exp(\mathrm{i}\boldsymbol{k}_2 \cdot \boldsymbol{r}_2) = A \exp[\mathrm{i}\boldsymbol{k}\,(\boldsymbol{r}_1 - \boldsymbol{r}_2)] = A \exp(\mathrm{i}\boldsymbol{k} \cdot \boldsymbol{r})$. Hierbei haben wir im letzten Ausdruck die Relativkoordinate mit $\boldsymbol{r} = \boldsymbol{r}_1 - \boldsymbol{r}_2$ abgekürzt.

Durch Wechselwirkung mit dem Gitter werden die Paare ständig in neue Zustände mit unterschiedlichen Wellenvektoren gestreut. Wir setzen daher als Lösung von (12.14) eine Überlagerung derartiger Paarzustände an:

$$\Psi(\boldsymbol{r}) = \sum_{\boldsymbol{k}} A_{\boldsymbol{k}}\, \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}}\ , \tag{12.15}$$

wobei der Entwicklungskoeffizient $A_{\boldsymbol{k}}$ nicht vom Ort abhängt. Beschränken wir uns auf den Grenzfall $T = 0$, so liegt die Wellenzahl der wechselwirkenden Elektronen im Bereich $k_\mathrm{F} < k < [2m(E_\mathrm{F} + \hbar\omega_\mathrm{D})/\hbar^2]^{1/2}$. Der Entwicklungskoeffizient $A_{\boldsymbol{k}}$ verschwindet daher für alle anderen Werte von k.

Die Schrödinger-Gleichung (12.14) lösen wir mit den üblichen Methoden der Quantenmechanik: Wir setzen den Lösungsansatz (12.15) ein, multiplizieren mit $\exp(-\mathrm{i}\boldsymbol{k}' \cdot \boldsymbol{r})$ und integrieren über das Probenvolumen. Wie bei der Ableitung der Streubedingung in Abschnitt 2.4 verschwindet das Integral über $\exp[\mathrm{i}(\boldsymbol{k} - \boldsymbol{k}') \cdot \boldsymbol{r}]$, wenn $\boldsymbol{k} \neq \boldsymbol{k}'$ ist und ist gleich dem Probenvolumen V_P für $\boldsymbol{k} = \boldsymbol{k}'$. Machen wir darüber hinaus noch die plausible Annahme, daß die Wechselwirkung nur vom Abstand abhängt und wir deshalb $\widetilde{V}(\boldsymbol{r}_1, \boldsymbol{r}_2) = \widetilde{V}(\boldsymbol{r})$ schreiben dürfen, so erhalten wir die Lösung

$$\frac{\hbar^2\boldsymbol{k}^2}{m} A_{\boldsymbol{k}} + \frac{1}{V_\mathrm{P}} \sum_{\boldsymbol{k}'} A_{\boldsymbol{k}'} \int \widetilde{V}(\boldsymbol{r})\, \mathrm{e}^{\mathrm{i}(\boldsymbol{k}-\boldsymbol{k}')\cdot\boldsymbol{r}}\, \mathrm{d}V = E\, A_{\boldsymbol{k}}$$

$$\left(\frac{\hbar^2\boldsymbol{k}^2}{m} - E\right) A_{\boldsymbol{k}} = -\frac{1}{V_\mathrm{P}} \sum_{\boldsymbol{k}'} A_{\boldsymbol{k}'} \widetilde{V}_{\boldsymbol{k}\,\boldsymbol{k}'}\ . \tag{12.16}$$

Hierbei wurde zur Vereinfachung der Darstellung für das Matrixelement die Abkürzung

$$\widetilde{V}_{\boldsymbol{k}\,\boldsymbol{k}'} = \int \widetilde{V}(\boldsymbol{r})\, \mathrm{e}^{\mathrm{i}(\boldsymbol{k}-\boldsymbol{k}')\cdot\boldsymbol{r}}\, \mathrm{d}V \tag{12.17}$$

eingeführt. Da $\widetilde{V}(\boldsymbol{r})$ nicht näher bekannt ist, macht man meist eine weitere, sehr starke Vereinfachung. Man nimmt an, daß für den oben definierten Bereich der erlaubten Wellenzahlen das Matrixelement den konstanten Wert $\widetilde{V}_{\boldsymbol{k}\,\boldsymbol{k}'} = -\widetilde{V}_0$ annimmt, sonst aber

verschwindet. Da wir eine attraktive Wechselwirkung voraussetzen, ist $\widetilde{V}_0$ eine positive Konstante. Mit dieser starken Vereinfachung wird die rechte Seite von Gleichung (12.16) unabhängig von $\boldsymbol{k}$ und konstant. Lösen wir nach $A_{\boldsymbol{k}}$ auf, so finden wir

$$A_{\boldsymbol{k}} = \frac{\widetilde{V}_0}{V_\mathrm{P}} \frac{1}{(\hbar^2 \boldsymbol{k}^2/m - E)} \sum_{\boldsymbol{k}'} A_{\boldsymbol{k}'} \,. \tag{12.18}$$

Nun summieren wir diese Gleichung noch über alle $\boldsymbol{k}$-Werte. Da das Ergebnis der Summation nicht von der Benennung der Wellenvektoren abhängt, gilt $\sum_{\boldsymbol{k}'} A_{\boldsymbol{k}'} = \sum_{\boldsymbol{k}} A_{\boldsymbol{k}}$. Damit läßt sich die Gleichung (12.18) stark vereinfachen:

$$1 = \frac{\widetilde{V}_0}{V_\mathrm{P}} \sum_{\boldsymbol{k}} \frac{1}{(\hbar^2 \boldsymbol{k}^2/m - E)} \,. \tag{12.19}$$

Wir ersetzen die verbleibende Summation über die Wellenvektoren durch eine Integration über die Energie, wie wir es bereits wiederholt in den vorausgegangenen Kapiteln gemacht haben. Unter Verwendung von $\eta = \hbar^2 \boldsymbol{k}^2/2m$ lautet die Gleichung dann wie folgt

$$1 = \widetilde{V}_0 \frac{D(E_\mathrm{F})}{2} \int_{E_\mathrm{F}}^{E_\mathrm{F}+\hbar\omega_\mathrm{D}} \frac{\mathrm{d}\eta}{2\eta - E} \,. \tag{12.20}$$

In diesem Ausdruck haben wir die Zustandsdichte in der Umgebung der Fermi-Kante als konstant angenommen und vor das Integral gezogen. Der Faktor 1/2 tritt auf, weil sich Gleichung (12.14) auf Paare bezieht, die Zustandsdichte $D(E)$ jedoch die Dichte in der Einelektronnäherung beschreibt. Die Integration kann analytisch durchgeführt werden. Anschließend lösen wir nach der Energieänderung auf, die ein Elektronenpaar im Fermi-See durch die attraktive Wechselwirkung erfährt, und erhalten für die Energieänderung:

$$\delta E = E - 2E_\mathrm{F} = \frac{-2\hbar\omega_\mathrm{D}}{\mathrm{e}^{4/[\widetilde{V}_0 D(E_\mathrm{F})]} - 1} \approx -2\hbar\omega_\mathrm{D}\, \mathrm{e}^{-4/[\widetilde{V}_0 D(E_\mathrm{F})]} \,. \tag{12.21}$$

Die Energieänderung δE_F ist negativ, d.h. die Energie der Cooper-Paare wird durch die Wechselwirkung der beiden Elektronen untereinander reduziert. Der letzte Ausdruck gilt nur im Grenzfall $4/\widetilde{V}_0 D(E_\mathrm{F}) \gg 1$, doch ist diese Voraussetzung, wie wir im folgenden Abschnitt sehen werden, relativ gut erfüllt. An der Oberfläche des Fermi-Sees formieren sich Zweielektronen-Zustände, deren Energie um den Wert δE gegenüber der Energie der freien Elektronen bei $T = 0$ abgesenkt ist. Diese Instabilität des Fermi-Sees bewirkt einen Übergang vom Grundzustand des Normalleiters zu einem neuen Zustand, dem BCS-Grundzustand, den wir im nächsten Kapitel kurz diskutieren werden.

Die Gleichung (12.21) erlaubt eine Erklärung der scheinbar paradoxen Beobachtung, daß gute Metalle wie Silber oder Kupfer nicht supraleitend werden. Diese Metalle besitzen bei Zimmertemperatur eine hohe Leitfähigkeit, weil die Elektronen nur sehr schwach an Phononen koppeln und daher kaum gestreut werden. Die schwache Elektron-Phonon-Kopplung bewirkt aber auch, daß der Austausch von virtuellen Phononen nicht sehr ausgeprägt ist. Die attraktive Kopplung der Elektronen über die Phononen reicht nicht aus, die Coulomb-Abstoßung zu überwinden. Eine Ausbildung von Cooper-Paaren ist daher in diesen Metallen nicht möglich bzw. wurde zumindest noch nicht beobachtet.

Natürlich müssen wir noch berücksichtigen, daß die beiden Elektronen ununterscheidbare Fermionen sind. Die Gesamtwellenfunktion des Cooper-Paars muß deshalb antisymmetrisch sein. Die Wellenfunktion des Lösungsansatzes (12.15) ist symmetrisch bezüglich des Austausches der beiden Elektronen. Daher ist der nicht dargestellte Spinanteil der Wellenfunktion antisymmetrisch, d.h. die Spins müssen entgegengerichtet sein. Wir deuten dies symbolisch durch $(\boldsymbol{k}\uparrow, -\boldsymbol{k}\downarrow)$ an und sprechen von einem Singulett-Paar, da der Gesamtdrehimpuls Null ist. Nach außen hin benehmen sich die Cooper-Paare wie Bosonen. Insbesondere können sie einen gemeinsamen quantenmechanischen Zustand einnehmen, was Fermionen aufgrund der antisymmetrischen Wellenfunktion verwehrt ist. Die Konsequenzen daraus werden wir in den folgenden Kapiteln diskutieren. Bei komplizierterer Wechselwirkung ist auch eine gleichsinnige Spinausrichtung möglich, wodurch Triplett-Paare entstehen. In diesem Fall muß die Ortswellenfunktion der Paare antisymmetrisch, also eine p-Wellenfunktion, sein. Tatsächlich findet man eine derartige Paarbildung beim suprafluiden ^{3}He, dessen ungewöhnliche (natürlich nicht-metallischen) Eigenschaften, die man bei Temperaturen um 1 mK beobachtet, auf dem Auftreten derartiger Cooper-Paare beruhen. Auch bei einer Reihe von intermetallischen Verbindungen, die man als „Schwere Fermion-Systeme" bezeichnet, und bei den Hochtemperatursupraleitern ist die Struktur der Cooper-Paare komplexer. Mit sehr großer Wahrscheinlichkeit bilden sich dort Paare mit d-Wellenfunktionen aus, die einen resultierenden Gesamtdrehimpuls besitzen.

Berechnet man mit Hilfe von Gleichung (12.15) den mittleren Abstand zwischen den Elektronen eines Paars, so findet man einen Wert, der sich leicht auch auf einfache Weise abschätzen läßt. Aus der Energieunschärfe δE eines Cooper-Paars folgt für die Unschärfe der Wellenzahl $\delta k \approx (m\delta E/\hbar^2 k_\mathrm{F})$ und somit eine Mindestausdehnung $\delta x \approx 1/\delta k$ der Cooper-Paare von der Größenordnung $\delta x \approx (\hbar^2 k_\mathrm{F}/m\,\delta E)$. Da die Energieunschärfe nicht größer als die Bindungsenergie der Cooper-Paare sein kann und diese vergleichbar mit der thermischen Energie am Sprungpunkt sein muß, setzen wir in grober Näherung $\delta E \approx k_\mathrm{B} T_\mathrm{c}$. Damit ergibt sich, daß die Ausdehnung eines Cooper-Paars im Bereich von 100 bis 1000 nm liegt! Cooper-Paare sind also relativ groß. Zwischen den beiden Elektronen eines Paars halten sich Millionen anderer Elektronen auf.

BCS-Grundzustand

Im vorausgehenden, weitgehend theoretischen Kapitel wurde gezeigt, daß eine attraktive Wechselwirkung zwischen den freien Elektronen zur Paarbildung führt, die von einer Energieabsenkung begleitet ist. Die theoretische Beschreibung des Gesamtsystems ist mathematisch aufwendiger als die Behandlung eines einzelnen Cooper-Paars, da das Verhalten *aller* freien Elektronen eingeht. Natürlich wollen wir hier nicht die Theorie in ihren mathematischen Details nachvollziehen. Wir werden daher nur die Ergebnisse wiedergeben und versuchen, sie plausibel zu machen.

Der Ausgangspunkt der theoretischen Beschreibung ist der passende Ausdruck für die Wellenfunktion der Cooper-Paare. Man startet nicht mit der sonst üblichen Überlagerung der Wellenfunktionen einzelner Elektronen, sondern konstruiert sie aus den Paar-Zuständen. Die freie Energie des BCS-Grundzustands, an dem alle Cooper-Paare beteiligt sind, ist nicht allein durch die kinetische Energie der Elektronen bestimmt. Davon abzuziehen ist die potentielle Energie, die durch die Paarbildung verursacht wird. Aus der Bedingung, daß die freie Energie im thermodynamischen Gleichgewicht ein Minimum besitzt, läßt sich die Wahrscheinlichkeit $w_{\boldsymbol{k}}$ für die Besetzung der Paarzustände bestimmen.

Die Rechnung führt bei $T = 0$ zu der Beziehung

$$w_{\boldsymbol{k}} = \frac{1}{2}\left(1 - \frac{z_{\boldsymbol{k}}}{\sqrt{z_{\boldsymbol{k}}^2 + \Delta^2}}\right) , \tag{12.22}$$

wobei wir für die kinetische Energie eines Elektrons die Abkürzung $z_{\boldsymbol{k}} = \hbar^2\boldsymbol{k}^2/2m - E_\mathrm{F}$ geschrieben haben. Für die Konstante Δ findet man unter der Annahme $\widetilde{V}_0 D(E_\mathrm{F}) < 1$ den Ausdruck

$$\Delta \approx 2\hbar\omega_\mathrm{D}\, \mathrm{e}^{-2/[\widetilde{V}_0 D(E_\mathrm{F})]} . \tag{12.23}$$

Es fällt die große Ähnlichkeit mit Gleichung (12.21) auf, in der die Energieabsenkung bei Paarbildung beschrieben wird. Offensichtlich ist die Größe Δ ein Maß für die Bindungsenergie der Cooper-Paare. Die genaue Bedeutung von Δ, das sich experimentell bestimmen läßt, werden wir noch ausführlich diskutieren. Mit der bekannten Debye-Frequenz der Supraleiter folgen aus den Meßwerten von Δ für $\widetilde{V}_0 D(E_\mathrm{F})$ Werte um 0,6.

In Abb. 12.10 ist die Funktion (12.22), die die Besetzung $w_{\boldsymbol{k}}$ der Paarzustände bei der Temperatur $T = 0$ beschreibt, mit der Fermi-Funktion $f(E, T)$ bei der Temperatur $T = T_\mathrm{c}$ verglichen, die die Besetzung der Einelektronzustände bei dieser Temperatur angibt. Bemerkenswert ist der ähnliche Verlauf der beiden Kurven. Der Abbildung entnehmen wir, daß im Supraleiter auch am absoluten Nullpunkt die Besetzung der Zustände im Bereich $E_\mathrm{F} \pm \Delta$ aufgeweicht ist. Dies war zu erwarten, da die Wechselwirkung zwischen Elektronen und Phononen nur in der Umgebung der Fermi-Kante erfolgt kann. Somit ist, über alle Leitungselektronen summiert, die kinetische Energie der Elektronen im Supraleiter höher als im Normalleiter. Es mag überraschen, daß Gleichung (12.22) die Besetzung auch bei großen negativen $z_{\boldsymbol{k}}$-Werten korrekt wiedergibt. Dies liegt daran, daß das Pauli-Prinzip bereits beim Ansatz der Wellenfunktion für den BCS-Grundzustand berücksichtigt wurde. Diese Maßnahme führt bei Energien $E \ll E_\mathrm{F}$ unmittelbar zur Fermi-Verteilung.

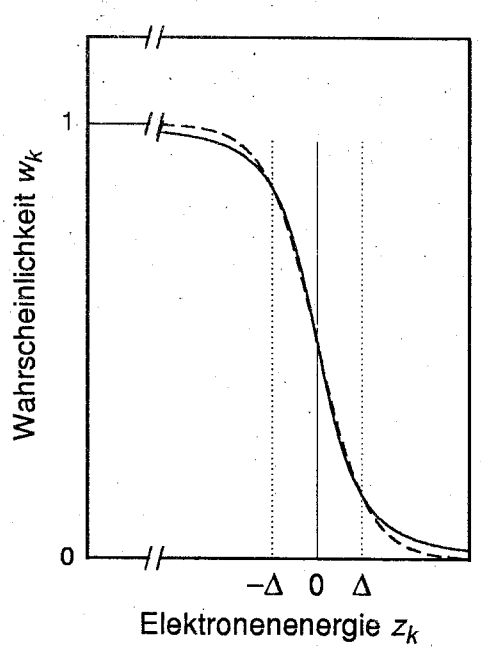


Abb. 12.10: Besetzungswahrscheinlichkeit der Paarzustände bei $T = 0$ als Funktion der Einteilchenenergie $z_{\boldsymbol{k}}$. Sie ähnelt sehr der Fermi-Funktion bei der Temperatur $T = T_\mathrm{c}$, die gestrichelt gezeichnet ist.

Wenn auch die kinetische Energie der Elektronen des Supraleiters erhöht wird, so tritt aufgrund der Absenkung der potentiellen Energie dennoch ein Energiegewinn beim Übergang von der normal- zur supraleitenden Phase auf. In der BCS-Theorie findet man für diese sogenannte *Kondensationsenergie* $G_{\mathrm{kon}} = G_{\mathrm{n}} - G_{\mathrm{s}}$ den Zusammenhang

$$G_{\mathrm{kon}} = \frac{V}{4} D(E_{\mathrm{F}}) \Delta^2 \,. \tag{12.24}$$

Diese Energie wird bei der „Kondensation" der Leitungselektronen zu Cooper-Paaren frei. Sie ist identisch mit der Differenz der freien Enthalpie von normal- und supraleitendem Zustand, für die wir in Abschnitt 12.1 den Ausdruck $G_{\mathrm{kon}} = V B_{\mathrm{c}}^2 / 2\mu_0$ mit Hilfe von thermodynamischen Betrachtungen hergeleitet haben.

Welche Energie ist erforderlich, um den BCS-Zustand anzuregen? Der einfachste Anregungsvorgang ist das Aufbrechen *eines* Cooper-Paars, wenn wir die Energie ε zuführen. Hierbei wird ein Elektron aus dem Zustand $\boldsymbol{k}$ gestreut, und ein ungepaartes mit dem Wellenvektor $-\boldsymbol{k}$ bleibt zurück. Bezeichnen wir mit $E_{\boldsymbol{k}}$ die auf die Fermi-Energie bezogene Energie eines einzelnen, ungepaarten Elektrons im Supraleiter, so muß zum Aufbrechen des Cooper-Paars die Energie $\varepsilon = 2E_{\boldsymbol{k}}$ aufgewandt werden. Eine etwas längere Rechnung führt zu dem einfachen Resultat:

$$E_{\boldsymbol{k}} = \sqrt{z_{\boldsymbol{k}}^2 + \Delta^2} \,. \tag{12.25}$$

Dies ist die Dispersionsrelation für die ungepaarten Elektronen im Supraleiter. Sie ersetzt die Beziehung $E = \hbar^2 k^2 / 2m = z_{\boldsymbol{k}} + E_{\mathrm{F}}$ für freie Elektronen im Normalleiter. Die kleinste Anregungsenergie erhält man für Elektronen, die nach dem Aufbrechen des Cooper-Paars die kinetische Energie $z_{\boldsymbol{k}} = 0$, bzw. den Wellenvektor $\boldsymbol{k}_{\mathrm{F}}$ besitzen. Dann ist die Energie $\varepsilon_{\mathrm{min}} = 2\Delta$ aufzubringen. Im Gegensatz zum Normalleiter, bei dem Elektronen mit beliebig kleinen Energien angeregt werden können, existiert im Supraleiter die *Energielücke* Δ, die bei der Anregung überwunden werden muß. Der Faktor zwei im Ausdruck für die erforderliche minimale Energie drückt die Tatsache aus, daß beim Aufbrechen eines Cooper-Paars immer *zwei* ungepaarte Elektronen erzeugt werden.

Die Interpretation der Gleichung (12.25) ist nicht ganz einfach. Wir betrachten zur Veranschaulichung das Aufbrechen eines Cooper-Paars. Nehmen wir an, daß das Elektron mit dem Wellenvektor $\boldsymbol{k}$ in den Zustand $\boldsymbol{k}'$ gestreut wurde. Dann bleibt ein Loch bei $\boldsymbol{k}$ zurück, mit dem nun das zweite Elektron mit dem Wellenvektor $-\boldsymbol{k}$ wechselwirkt. Umgekehrt konnte der Zustand $\boldsymbol{k}'$ nur besetzt werden, da vor dem Streuprozeß kein Paarzustand $(\boldsymbol{k}', -\boldsymbol{k}')$ existierte, also kein Elektron mit dem Wellenvektor $-\boldsymbol{k}'$ vorhanden war. Durch den Streuprozeß ist also jetzt ein weiteres Elektron-Loch-Paar mit dem Wellenvektor $\boldsymbol{k}'$ und $-\boldsymbol{k}'$ gebildet worden[3]. Die Elektronen, die sich nicht in Paarzuständen befinden, bezeichnet man als normalleitende Elektronen oder als *Quasiteilchen*, um den soeben geschilderten komplizierten Mischzustand zu charakterisieren. Für große Werte von $z_{\boldsymbol{k}}$, also weiter entfernt von der Fermi-Fläche, geht die Dispersionsrelation (12.25) in die von freien Elektronen über. Wie bei den Normalleitern existiert bei Elektronenzuständen tiefer in der Fermi-Kugel zu jedem Elektron mit dem Wellenvektor $\boldsymbol{k}$ ein weiteres Elektron mit dem Wellenvektor $-\boldsymbol{k}$. Im Gegensatz zu den Cooper-Paaren wechselwirken sie nicht miteinander, d.h. ihre potentielle Energie ist gegenüber dem normalleitenden Zustand nicht abgesenkt.

[3] Es besteht eine große Ähnlichkeit zum bereits diskutierten direkten Übergang bei der optischen Absorption, bei dem ebenfalls Elektron-Loch-Paare mit entgegengesetzten Wellenvektoren erzeugt werden.

Wie sieht nun die Zustandsdichte $D_s(E_k)$ des Supraleiters für Quasiteilchen aus? Da jedem Zustand im Supraleiter umkehrbar eindeutig ein Zustand im Normalleiter zugeordnet werden kann, können wir

$$D_s(E_k) = D_n(z_k)\frac{dz_k}{dE_k} = D_n(z_k)\frac{E_k}{\sqrt{E_k^2 - \Delta^2}} \tag{12.26}$$

schreiben. Die Zustandsdichte der Quasiteilchen läßt sich beispielsweise mit Hilfe der Tunnelkontaktspektroskopie messen, auf die wir im nächsten Abschnitt eingehen. In Abb. 12.11 ist die so gemessene Zustandsdichte von Blei wiedergegeben. Deutlich zu erkennen ist die Existenz einer Energielücke und das starke Anwachsen der Zustandsdichte bei $E_k \to \Delta$, denn dort wird die Zustandsdichte $D_s(E_k)$ singulär. Von außen eingebrachte Elektronen finden nur Zustände bei Energien $E_k \geq \Delta$ vor. Durch Paarbildung können sie in den BCS-Grundzustand bei $E_k = 0$ übergehen.

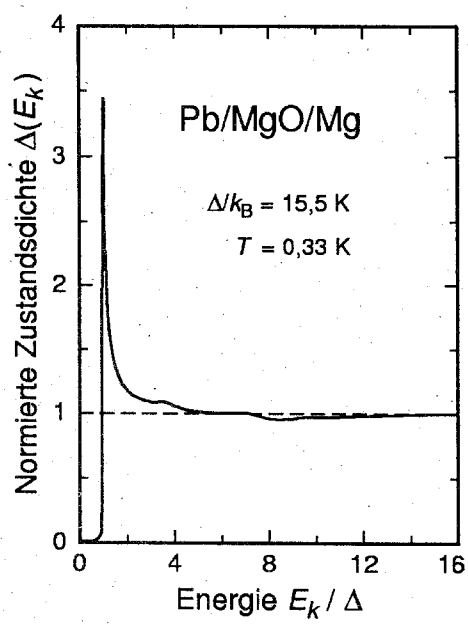


Abb. 12.11: Zustandsdichte der Quasiteilchen in Blei als Funktion der Einteilchenenergie (in Einheiten von Δ), bezogen auf die Zustandsdichte bei der Fermi-Energie. Diese Messung wurde mit einem Pb/MgO/Mg - Tunnelkontakt durchgeführt. (Nach I. Giaever, H. Hart, K. Megerle, Phys. Rev. **126** (1962) 941).

BCS-Zustand bei endlicher Temperatur

Bei endlicher Temperatur sind nicht alle Elektronen an der Fermi-Fläche gepaart, denn durch thermische Anregung werden Quasiteilchen geschaffen. Sie besetzen Zustände, die dann den wechselwirkenden Cooper-Paaren nicht mehr zugänglich sind und behindern so den Austausch von virtuellen Phononen. Damit reduziert sich die Bindungsenergie des BCS-Zustands, und die Energielücke $\Delta(T)$ nimmt mit der Temperatur ab, bis sie schließlich bei T_c ganz verschwindet. $\Delta(T)$ läßt sich ganz ähnlich wie die Energielücke bei $T = 0$ berechnen, doch muß nun auch der Beitrag der Quasiteilchen zur freien Energie berücksichtigt werden. Es ergibt sich ein Ausdruck, der Gleichung (12.20) sehr ähnlich sieht, aber nur numerisch auswertbar ist. Der resultierende Temperaturverlauf der Energielücke ist in Abb. 12.12 zusammen mit experimentellen Ergebnissen zu sehen.

Die Punkte sind das Ergebnis von Messungen mit unterschiedlichen Techniken und stimmen erstaunlich gut mit der Vorhersage der BCS-Theorie überein. Dies ist besonders bemerkenswert, da kein freier Parameter zur Kurvenanpassung verwendet wird. Das

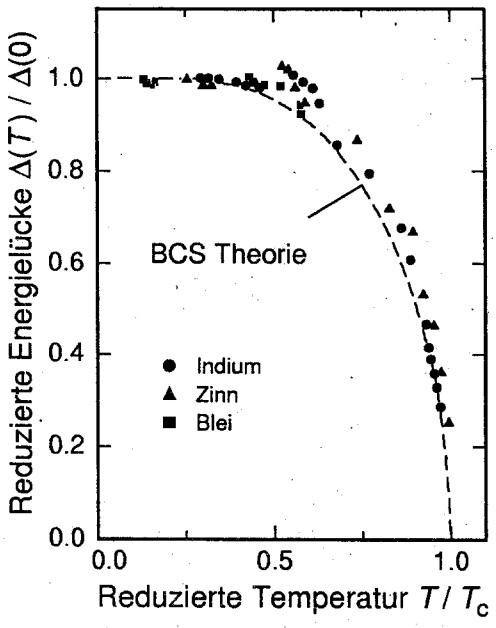


Abb. 12.12: Verlauf der normierten Energielücke $\Delta(T)/\Delta(0)$ in Abhängigkeit von der reduzierten Temperatur T/T_c. Neben der theoretischen Kurve sind die experimentellen Werte für Indium, Zinn und Blei eingezeichnet. (Nach I. Giaever und K. Megerle, Phys. Rev. **122** (1961) 1001).

Auftreten von kleinen Abweichungen ist nicht überraschend, da unsere Annahme eines in einem eingeschränkten Bereich konstanten Kopplungsparameters $\widetilde{V}_0$ sicherlich nur eine sehr grobe Näherung sein kann. In einer realistischen Rechnung müßten die Coulomb-Wechselwirkung der Elektronen, das tatsächliche Phononenspektrum einschließlich der Anisotropie und die Frequenzabhängigkeit der Kopplung in Rechnung gestellt werden.

Berücksichtigt man bei der oben erwähnten numerischen Integration, daß die Energielücke bei T_c verschwindet, so erhält man die Beziehung

$$k_B T_c = 1{,}14\,\hbar\omega_D\, e^{-2/[\widetilde{V}_0\, D(E_F)]}\,, \tag{12.27}$$

und unter Berücksichtigung von Gleichung (12.23) das Verhältnis

$$\frac{\Delta(0)}{k_B T_c} = 1{,}76\,. \tag{12.28}$$

In Tabelle 12.2 ist für eine Reihe von Supraleitern der gemessene Wert von $2\Delta(0)/k_B T_c$ eingetragen. In den meisten Fällen ist die Übereinstimmung gut, doch bei einigen Supraleitern ist dieses Verhältnis wesentlich größer als 3,52. Man spricht dann von stark koppelnden Supraleitern, für die die BCS-Theorie etwas modifiziert werden muß.

Tabelle 12.2: Verhältnis von Energielücke zu Sprungtemperatur von einigen Supraleitern.

Supraleiter	Al	Cd	Hg	In	Nb	Pb	Nb_3Ge	V_3Si
$2\Delta(0)/k_B T_c$	3,4	3,2	4,6	3,6	3,7	4,4	4,2	3,6

Nachweis der Energielücke

Ein wesentliches Charakteristikum der Supraleitung ist die Existenz einer Energielücke. Sie bestimmt die Zahl der vorhandenen Quasiteilchen, die sich zum Beispiel unmittelbar

in der spezifischen Wärme bemerkbar macht. Erhöht man die Temperatur eines Supraleiters, so muß zur Erzeugung von Quasiteilchen bzw. zum Aufbrechen von Cooper-Paaren Energie aufgewendet werden. Wegen der Energielücke $\Delta(T)$ ist die Wahrscheinlichkeit für die Besetzung der angeregten Zustände und damit auch die spezifische Wärme dem Boltzmann-Faktor $\exp[-\Delta(T)/k_\mathrm{B}T]$ proportional. Trägt man beispielsweise die Daten für die spezifische Wärme C_s von Aluminium in Abb. 12.7 logarithmisch als Funktion der reziproken Temperatur auf, so findet man die erwartete Gerade. Nur in der Nähe des Sprungpunkts treten Abweichungen von diesem Verlauf auf, die durch die Temperaturabhängigkeit der Energielücke bedingt sind. Für den Sprung der spezifischen Wärme bei T_c sagt die BCS-Theorie den Wert $(C_\mathrm{s} - C_\mathrm{n})/C_\mathrm{n} = 1{,}43$ voraus. In den meisten Fällen stimmen die experimentellen Werte mit dieser Zahl überein, doch treten bei den stark koppelnden Supraleitern auch hier Abweichungen auf.

Die Zahl der Quasiteilchen spiegelt sich auch in vielen Transporteigenschaften wider. Cooper-Paare befinden sich im quantenmechanischen Grundzustand, tragen daher keine Entropie und fallen somit für den Wärmetransport aus. Mit der Zahl der Quasiteilchen nimmt daher die Wärmeleitfähigkeit von Supraleitern unterhalb von T_c drastisch ab. Ein weiters Beispiel ist die Ultraschallabsorption. Die freie Weglänge der Ultraschallphononen wird in reinen Metallen durch Stöße mit den freien Elektronen begrenzt. Da Cooper-Paare nur mit Phononen wechselwirken können, deren Energie zur Paarbrechung ausreicht, können Stöße nur mit den thermisch angeregten Quasiteilchen stattfinden. Mit den Quasiteilchen verschwindet daher beim Abkühlen auch die Ultraschalldämpfung.

Wie läßt sich die Energielücke aber „direkt" messen? Unmittelbar einleuchtend sind Mikrowellen- und Infrarotexperimente, mit deren Hilfe sich die Lücke wie bei der optischen Absorption an Halbleitern bestimmen läßt. Reicht bei derartigen Experimenten die Energie der eingestrahlten Photonen nicht zum Brechen der Cooper-Paare aus, so tritt keine Absorption auf. Ist jedoch $\hbar\omega > 2\Delta$, so können pro Photon zwei Quasiteilchen erzeugt werden, und die Strahlung wird stark absorbiert. Derartige Messungen sind relativ aufwendig und nicht ganz einfach zu analysieren, da in Metallen noch zusätzliche Absorptionsmechanismen existieren (vgl. Abschnitt 11.4).

Wir wollen hier auf die konzeptionell außerordentlich einfache und elegante *Tunnelkontaktspektroskopie* näher eingehen. Der schematische Aufbau und das Prinzipschaltbild eines derartigen Tunnelkontaktexperiments sind in den beiden Bildern der Abb. 12.13 dargestellt.

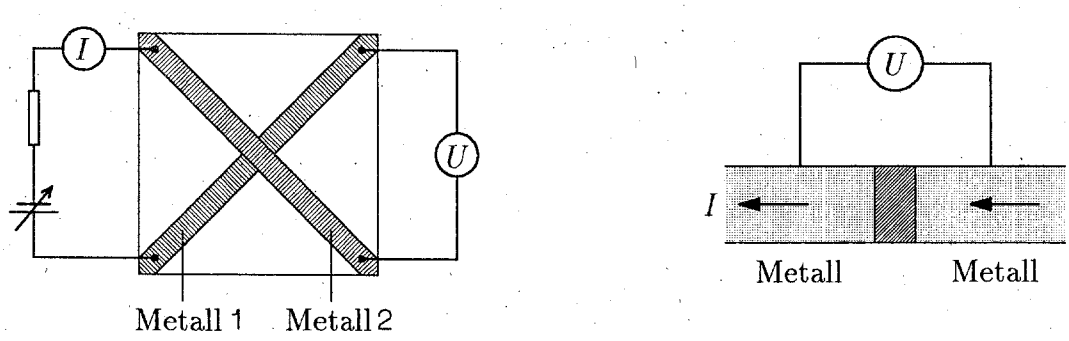


Abb. 12.13: Tunnelkontakt.
a) Die kreuzweise aufgedampften Metallstreifen sind durch eine dünne Isolatorschicht voneinander getrennt.
b) Prinzipschaltbild. Bei vorgegebenem Strom I wird der Spannungsabfall an den beiden Metallstreifen gemessen.

Die beiden Metallstreifen werden nacheinander auf das isolierende Substrat aufgedampft und sind durch eine dünne, etwa 3 nm dicke Isolatorschicht voneinander getrennt. Ähnlich wie beim Rastertunnelmikroskop (vgl. Abschnitt 7.2) können Elektronen durch die Isolatorschicht von einem Metall zum anderen tunneln. Bei Anliegen einer Spannung U sind die Fermi-Niveaus der beiden Metalle um den Betrag eU gegeneinander verschoben, und Elektronen tunneln, wie es in Abb. 12.14 für zwei Normalleiter angedeutet ist, aus den besetzten Zuständen des einen Metalls in die unbesetzten des anderen. Die beobachtete Strom-Spannungs-Charakteristik zeigt im Fall von zwei normalleitenden Metallen ein Ohmsches Verhalten, d.h. $I \propto U$.

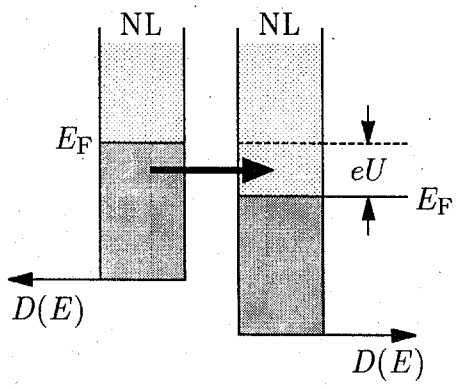


Abb. 12.14: Zustandsdichte zweier Normalleiter bei $T = 0$ in der Nähe der Fermi-Energie. Durch die Spannung U sind die Fermi-Niveaus um die Energie eU gegeneinander verschoben.

Besteht der Tunnelkontakt aus einem Normal- und einem Supraleiter, so kann bei $T = 0$ und der Spannung $U < \Delta/e$ kein Strom fließen, da in der Energielücke des Supraleiters keine freien Zustände vorhanden sind. Die dazugehörende Zustandsdichte ist in Abb. 12.15 dargestellt. Stromfluß tritt erst auf, wenn die kritische Spannung $U_k = \Delta/e$ überschritten ist. Diese Schlußfolgerung leuchtet unmittelbar ein, wenn der Normalleiter auf höherem Potential liegt (vgl. Abb. 12.15a). Elektronen aus den besetzten Zuständen unterhalb E_F können durch die Isolatorschicht erst dann in die leeren Quasiteilchenzustände des Supraleiters tunneln, wenn $U > U_k$ ist.

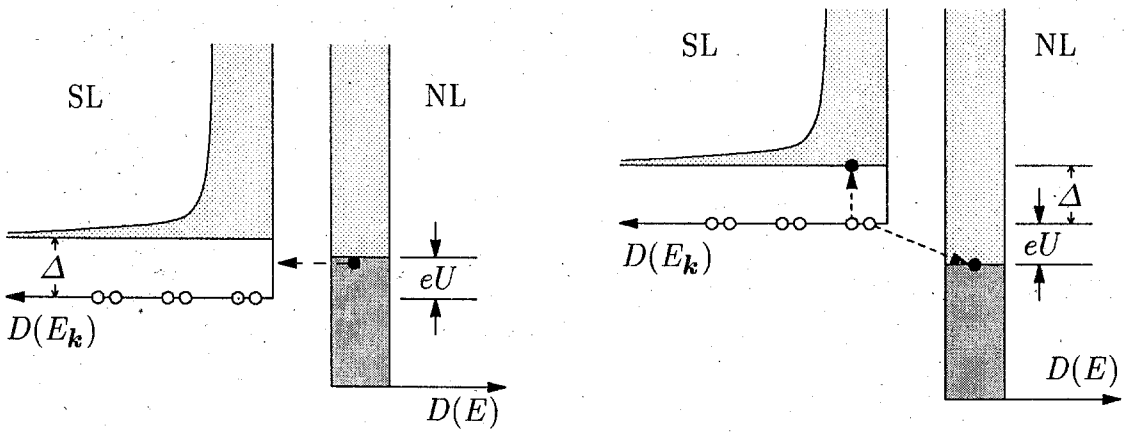


Abb. 12.15: Zustandsdichtediagramm eines Supraleiter-Normalleiter-Kontakts bei Anliegen einer Spannung $|U| < \Delta/e$. Cooper-Paare sind durch zwei offene Kreise, Quasiteilchen bzw. ungepaarte Elektronen durch ausgefüllte Kreise angedeutet.

a) Das Fermi-Niveau des Normalleiters ist angehoben, doch finden seine Elektronen keine freien Zustände im Supraleiter.

b) Das Fermi-Niveau des Normalleiters ist abgesenkt, der Energiegewinn eU des tunnelnden Elektrons reicht jedoch nicht aus, um das im Supraleiter verbleibende Elektron in das Band der Quasiteilchen zu heben. Der angedeutete Prozeß ist daher ebenfalls verboten.

Bei umgekehrter Polung (vgl. Abb. 12.15b) entsteht der Eindruck, daß Elektronen bereits bei kleineren Spannungen den BCS-Grundzustand verlassen und in den Normalleiter überwechseln können. Dies ist jedoch nicht richtig, denn ein Cooper-Paar muß aufgebrochen und das zurückbleibende Elektron in das Band der Quasiteilchen gehoben werden. Die erforderliche Energie steht aber nicht zur Verfügung, da $eU < \Delta$ ist. Ist dagegen die anliegende Spannung $U > U_k$, so kann ein Cooper-Paar aufgebrochen werden: Ein Elektron tunnelt durch die Isolatorschicht in den Normalleiter, und die frei werdende Energie wird genutzt, um das zweite Elektron in einen Quasiteilchenzustand zu heben.

Wie man sich leicht vorstellen kann, treten zwei entgegengesetzt gerichtete Teilströme auf, deren Stärken durch die Zustandsdichten und die Besetzungswahrscheinlichkeiten bestimmt werden. Der resultierende Strom ist dann durch

$$I \propto \int D_s(E_{\boldsymbol{k}}) D_n(E)[f(E) - f(E + eU)]\,\mathrm{d}E \qquad (12.29)$$

gegeben. Da im Normalleiter die Zustandsdichte in der Nähe des Fermi-Niveaus konstant ist und bei $T = 0$ die Besetzungswahrscheinlichkeit eine Sprungfunktion darstellt, folgt daraus die einfache Beziehung $\mathrm{d}I/\mathrm{d}U \propto D_s(E_{\boldsymbol{k}} = eU)$. Die Messung der Strom-Spannungs-Charakteristik erlaubt daher die Bestimmung der Zustandsdichte, wofür bereits in Abb. 12.11 ein Beispiel gegeben wurde.

Der schematische Verlauf der Strom-Spannungscharakteristik eines Supraleiter-Normalleiter-Kontakts ist in Abb. 12.16 zu sehen. Der steile Anstieg beim Einsetzen des Stromflusses spiegelt die Singularität der Quasiteilchenzustandsdichte bei $E_F + \Delta$ wider. Bei endlichen Temperaturen sind in Normal- und Supraleitern angeregte Elektronen vorhanden. Tunnelprozesse und somit Stromfluß treten daher bereits bei $U < U_k$ auf. Natürlich kann das geschilderte Experiment auch mit zwei supraleitenden Metallen durchgeführt werden. Wie man sich leicht überlegen kann, gilt dann für die kritische Spannung die Gleichung $eU_k = \Delta_1 + \Delta_2$.

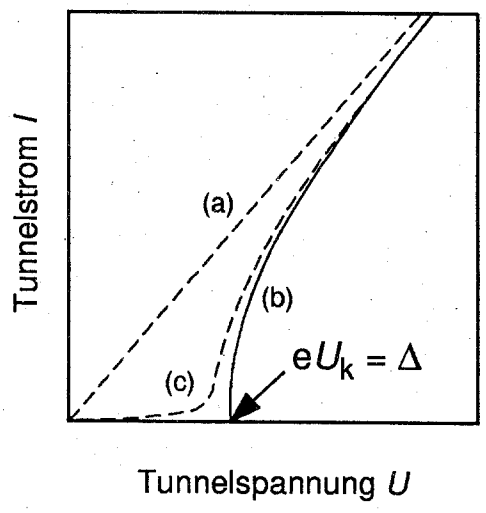


Abb. 12.16: Strom-Spannungs-Charakteristik eines Supraleiter-Normalleiter-Kontakts bei unterschiedlichen Temperaturen.
a) Bei $T > T_c$ sind beide Metalle normalleitend; es wird eine Ohmsche Strom-Spannungscharakteristik beobachtet.
b) Bei $T = 0$ setzt bei der Spannung $U_k = \Delta/e$ ein steiler Stromanstieg ein.
c) Bei $0 < T < T_c$ tritt ein schwacher Stromfluß durch angeregte Quasiteilchen bereits bei $U < \Delta/e$ auf.

Auf eine weitere, äußerst interessante Anwendung der Tunnelkontaktspektroskopie kommen wir in Abschnitt 12.3 zu sprechen. Am Rande sei erwähnt, daß die hier geschilderten Experimente häufig durch eine modifizierte Darstellung der Zustandsdichten, die sogenannte „Halbleiterdarstellung", beschrieben werden, auf die wir hier aber nicht eingehen.

Kritischer Strom und kritisches Magnetfeld

In Abschnitt 12.1 haben wir gefunden, daß bereits aus thermodynamischen Gründen ein kritisches Magnetfeld existiert. In diesem Abschnitt wollen wir zeigen, daß damit auch eine *kritische Stromdichte* verknüpft ist, oberhalb der die Supraleitung zerstört wird.

Bei Stromfluß ist der Bewegung der einzelnen Elektronen eines Paars eine Schwerpunktsbewegung überlagert. Anstelle der Cooper-Paare $(\boldsymbol{k}\uparrow, -\boldsymbol{k}\downarrow)$ treten nun solche, die durch $(\boldsymbol{k} + \delta\boldsymbol{K}\uparrow, -\boldsymbol{k} + \delta\boldsymbol{K}\downarrow)$ charakterisiert sind. Über die Gleichung $m\boldsymbol{v} = \hbar\,\delta\boldsymbol{K}$ ist die mit der Schwerpunktsgeschwindigkeit $\boldsymbol{v}$ verknüpfte Änderung $\delta\boldsymbol{K}$ der Wellenvektoren festgelegt. Der Vollständigkeit halber sei erwähnt, daß die Schwerpunktsbewegung der Paare weder $\widetilde{V}_{\boldsymbol{k}\,\boldsymbol{k}'}$ noch Δ selbst verändert. Die Energielücke bewegt sich also mit der Fermi-Kugel im $\boldsymbol{k}$-Raum.

Es ist zu erwarten, daß die Supraleitung zusammenbricht, wenn die zusätzliche kinetische Energie der Elektronen die Kondensationsenergie (12.24) übersteigt. Wir können daher für die kritische Schwerpunktsgeschwindigkeit $\boldsymbol{v}_\mathrm{c}$ die Bedingung

$$\frac{1}{2}nmv_\mathrm{c}^2 = \frac{1}{4}D(E_\mathrm{F})\Delta^2 \tag{12.30}$$

aufstellen. Setzen wir $D(E_\mathrm{F})$ ein, so finden wir

$$v_\mathrm{c} = \sqrt{\frac{3}{2}}\,\frac{\Delta}{mv_\mathrm{F}}\,. \tag{12.31}$$

Berücksichtigen wir, daß die Dichte des Suprastroms durch $\boldsymbol{j}_\mathrm{s} = -n_\mathrm{s}e\boldsymbol{v}$ gegeben ist, so folgt für die kritische Stromdichte

$$j_\mathrm{c} = \sqrt{\frac{3}{2}}\,\frac{en_\mathrm{s}\Delta}{\hbar k_\mathrm{F}}\,. \tag{12.32}$$

Mit typischen Zahlenwerten für die auftretenden Größen ergibt sich eine kritische Stromdichte von etwa $j_\mathrm{c} \approx 10^7\mathrm{A/cm}^2$.

Mit dem Strom ist ein Magnetfeld $\boldsymbol{H}$ verknüpft, dessen Stärke wir für den Fall des kritischen Stromflusses in einem Draht berechnen wollen. Für das Feld an der Drahtoberfläche gilt $2\pi rH = \int \boldsymbol{j}\cdot\mathrm{d}\boldsymbol{f}$, wobei r der Drahtradius und $\mathrm{d}\boldsymbol{f}$ ein Flächenelement des Drahtquerschnitts ist. Unter der Voraussetzung $r \gg \lambda_\mathrm{L}$ erhalten wir mit Hilfe von Gleichung (12.4) $\int \boldsymbol{j}\cdot\mathrm{d}\boldsymbol{f} = 2\pi j_0\lambda_\mathrm{L}r$. Setzen wir für die Stromdichte j_0 an der Oberfläche die kritische Stromdichte ein, so finden wir für das kritische Feld

$$H_\mathrm{c} = \lambda_\mathrm{L}\,j_\mathrm{c} = \sqrt{\frac{3}{2}}\,\frac{en_\mathrm{s}\lambda_\mathrm{L}\Delta}{\hbar k_\mathrm{F}}\,. \tag{12.33}$$

Diese Beziehung ergibt sich auch, wenn man die Ausdrücke (12.10) und (12.24) für die Kondensationsenergie aus der thermodynamischen und der mikroskopischen Betrachtung gleichsetzt. Kritische Magnetfeldstärke und kritischer Strom sind also direkt miteinander verknüpft, unabhängig davon, ob die Stromdichte durch Abschirm- oder Transportströme hervorgerufen wird. Eine entsprechende Hypothese wurde bereits 1916 von *Silsbee* aufgestellt.

Nun wird auch verständlich, warum im Supraleiter der elektrische Widerstand verschwindet. Im Normalleiter wird der Widerstand durch Streuung der Elektronen an Defekten und Phononen verursacht. Da alle Cooper-Paare die gleiche gemeinsame Bewegung ausführen, die durch den zusätzlichen Wellenvektor $\delta \boldsymbol{K}$ charakterisiert ist, muß bei einer Streuung das Cooper-Paar die gemeinsame Wellenfunktion und somit den BCS-Zustand verlassen. Dies ist nur unter Aufwendung der Bindungsenergie möglich, wodurch elastische Streuprozesse von vornherein ausgeschlossen sind. Natürlich sind Stöße mit Phononen möglich, wenn deren Energie groß genug ist. Entweder wird ein Phonon absorbiert und das Cooper-Paar zerstört, oder ein Phonon wird bei der Bildung eines Paars emittiert. Da im thermischen Gleichgewicht immer Quasiteilchen vorhanden sind, finden beide Prozesse ständig statt. Im Mittel wird hierbei jedoch der Ladungstransport nicht beeinflußt, da die gebildeten Cooper-Paare nur in den bestehenden BCS-Zustand kondensieren können, aus dem sie vorher herausgestreut wurden. Zusätzlich stellt sich die Frage, was die Quasiteilchen machen, da diese durchaus Verluste verursachen könnten. Die Antwort ist einfach: Beim Fluß eines Gleichstroms tritt im Supraleiter kein elektrisches Feld auf, da dieses durch die Cooper-Paare „kurzgeschlossen" wird. Die Quasiteilchen können am Stromtransport gar nicht teilnehmen, bzw. ihr Beitrag mittelt sich weg. Dies gilt nicht mehr beim Anliegen einer Wechselspannung. Nach Gleichung (12.2) werden hierbei die Cooper-Paare zwar verlustfrei beschleunigt, doch das vorhandene elektrische Feld beschleunigt auch die Quasiteilchen, die ihrerseits mit dem Gitter wechselwirken und Verluste hervorrufen. Verlustfreie Leitfähigkeit tritt daher nur bei Gleichstrom auf.

12.3 Makroskopische Wellenfunktion

In der bisherigen Diskussion haben wir zwar an einigen Stellen von der *makroskopischen Wellenfunktion* der Supraleiter gesprochen, aber davon nicht explizit Gebrauch gemacht. Nun wollen wir auf diese Wellenfunktion und ihre Bedeutung näher eingehen. Cooper-Paare tragen keinen Gesamtspin, verhalten sich daher wie Bosonen und können in einen *gemeinsamen Zustand* kondensieren, der durch die makroskopische Wellenfunktion

$$\Psi = \Psi_0 \, \mathrm{e}^{\mathrm{i}\varphi(\boldsymbol{r})} = \sqrt{n_\mathrm{s}} \, \mathrm{e}^{\mathrm{i}\varphi(\boldsymbol{r})} \tag{12.34}$$

charakterisiert ist. Der Betrag der Wellenfunktion ist durch die Dichte der Cooper-Paare $\Psi\Psi^\star = |\Psi_0|^2 = n_\mathrm{s}$ gegeben. Die reelle Funktion $\varphi(\boldsymbol{r})$ beschreibt die Phase der Wellenfunktion und besitzt bei Supraleitern über makroskopische Entfernungen einen wohldefinierten Wert. Die Existenz einer makroskopischen Wellenfunktion hat erhebliche Konsequenzen für das Verhalten von Supraleitern, die besonders augenfällig beim Anlegen eines Magnetfelds werden. Wir wollen hier zunächst auf die *Flußquantisierung* und anschließend auf den sogenannten *Josephson-Effekt* näher eingehen.

Flußquantisierung

Bringen wir einen mehrfach zusammenhängenden Supraleiter (vgl. Abb. 12.17) in ein Magnetfeld und kühlen ihn unter die Sprungtemperatur, so bleibt der innere Bereich des Rings vom Magnetfeld durchdrungen. Beim Abschalten des äußeren Felds bleibt der Fluß im Innenbereich gefangen. Der supraleitende Ring selbst ist aber bis auf eine dünne Schicht an der Probenoberfläche feldfrei.

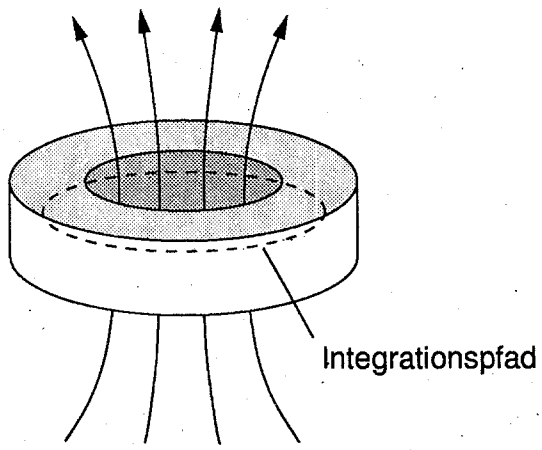


Abb. 12.17: Supraleitender Ring mit umschlossenem Magnetfluß. Abschirmströme fließen nur an der Oberfläche des Supraleiters. Punktiert ist ein Integrationspfad angedeutet, der im stromfreien inneren Bereich des Supraleiters verläuft.

Die Wellenfunktion von Supraleitern hat an jeder Stelle einen wohldefinierten Wert, so daß auch die Phase $\varphi(\boldsymbol{r})$ eindeutig festgelegt ist. Bewegen wir uns vom Ort 1 zu einem Ort 2, so gilt für die Phasendifferenz $\varphi_2 - \varphi_1 = \int_1^2 \operatorname{grad}\varphi(\boldsymbol{r}) \cdot \mathrm{d}\boldsymbol{s}$. Verläuft der Weg auf einer geschlossenen Bahn innerhalb des Supraleiters um die Ringöffnung, so muß für die Phasendifferenz die Bedingung $\varphi_2(\boldsymbol{r}_2 = \boldsymbol{r}_1) - \varphi_1(\boldsymbol{r}_1) = 2\pi p$ gelten, wobei p eine ganze Zahl ist.

Diese Quantisierungsbedingung hat Konsequenzen für den Stromfluß, den wir nun betrachten. Wir benutzen den quantenmechanischen Ausdruck für den Strom im Magnetfeld

$$\boldsymbol{j} = -\mathrm{i}\frac{\hbar q}{2M}\left(\Psi^\star\nabla\Psi - \Psi\nabla\Psi^\star\right) - \frac{q^2}{M}\boldsymbol{A}\,\Psi^\star\Psi \tag{12.35}$$

und ersetzen in dieser Gleichung q durch $-2e$ und M durch $2m$. Mit der Wellenfunktion (12.34) ergibt sich hieraus

$$\mu_0\lambda_\mathrm{L}^2\boldsymbol{j} = -\left(\frac{\hbar}{e}\nabla\varphi + 2\boldsymbol{A}\right) . \tag{12.36}$$

Nun bilden wir das Umlaufintegral über die Stromdichte:

$$\mu_0\lambda_\mathrm{L}^2 \oint \boldsymbol{j}\cdot\mathrm{d}\boldsymbol{s} = -\frac{\hbar}{e}\oint \nabla\varphi\cdot\mathrm{d}\boldsymbol{s} - 2\oint \boldsymbol{A}\cdot\mathrm{d}\boldsymbol{s} . \tag{12.37}$$

Wählen wir einen geschlossenen Integrationsweg in der Ringmitte, dann ist $\boldsymbol{j} = 0$, und das Integral auf der linken Seite verschwindet. Das Integral über den Gradienten haben wir bereits diskutiert. Das Linienintegral über das Vektorpotential läßt sich mit Hilfe des Stokesschen Satzes in ein Flächenintegral über die magnetische Induktion umformen; man erhält $\oint \boldsymbol{A}\cdot\mathrm{d}\boldsymbol{s} = \int \boldsymbol{B}\cdot\mathrm{d}\boldsymbol{f} = \Phi$, wobei Φ der umschlossene Magnetfluß ist. Damit erhalten wir das wichtige Ergebnis

$$\Phi = p\,\frac{h}{2e} = p\,\Phi_0 . \tag{12.38}$$

Der Magnetfluß durch eine geschlossene supraleitende Schleife ist quantisiert. Es kann nur ein Vielfaches des *Flußquants* $\Phi_0 = h/2e = 2{,}07\cdot 10^{-15}\,\mathrm{Tm}^2$ auftreten, das oft auch

als *Fluxoid* bezeichnet wird. Der Fluß Φ_0 ist extrem klein. Soll beispielsweise ein Hohlzylinder im Erdfeld nur ein Flußquant umschließen, so darf sein Durchmesser nur etwa $5\,\mu$m betragen.

Von Bedeutung ist nicht nur die Quantisierung des Flusses, auf die *London* bereits 1950 hinwies, sondern auch die Tatsache, daß die Ladung des Cooper-Paars $2e$ auftritt. Die Bestimmung des Fluxoids im Jahre 1961 durch *Doll* und *Nähbauer* bzw. *Deaver* und *Fairbank* war daher ein starker Hinweis auf die Existenz von Cooper-Paaren. In den Experimenten wurden dünne supraleitende Zylinder zunächst in einem sehr schwachen Magnetfeld abgekühlt. Anschließend wurde das Magnetfeld abgeschaltet und das magnetische Moment des Zylinders bestimmt. Die Messungen ergaben, daß der eingefrorene Fluß tatsächlich quantisiert war und die Einheit des Flusses durch Gleichung (12.38) gegeben ist.

In Abb. 12.18 ist das Ergebnis einer neueren Messung[4] an einem Zinnzylinder mit einem Durchmesser von $56\,\mu$m gezeigt, der in einem Magnetfeld abgekühlt wurde. Es ist offensichtlich, daß der eingefrorene Magnetfluß nicht der stetigen Änderung des Feldes folgt, sondern quantisiert ist. Die auftretenden Rundungen beruhen darauf, daß unter den experimentellen Gegebenheiten ein Flußquant nicht immer auf der ganzen Länge innerhalb des Zylinders verläuft.

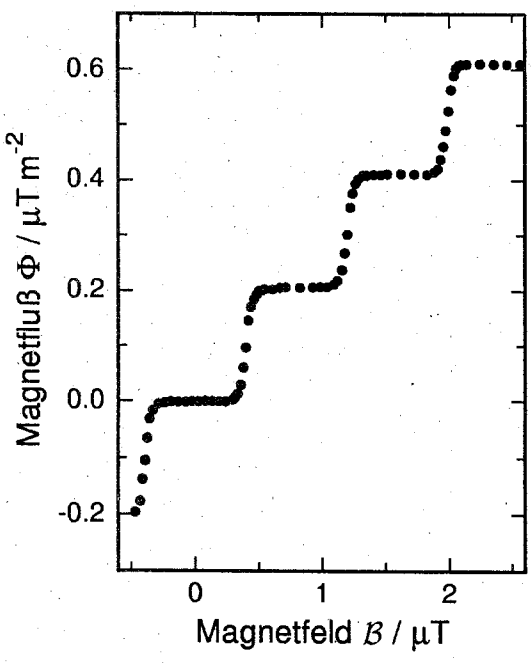


Abb. 12.18: Magnetfluß in einem dünnen supraleitenden Zinnzylinder (Länge 24 mm, Durchmesser $56\,\mu$m) als Funktion des Magnetfelds, in dem der Zylinder abgekühlt wurde. (Nach W.L. Goodman, W.D. Willis, D.A. Vincent, B.S. Deaver, Phys. Rev. B **4** (1971) 1530).

Die Quantisierung des Magnetflusses zieht die Quantisierung des Stroms in einer Stromschleife nach sich. Da sich die Phase der Wellenfunktion nur um ein ganzzahliges Vielfaches von 2π ändern kann, sind keine kontinuierlichen Variationen des Stroms sondern nur Sprünge erlaubt. Eine Änderung der Phase der Wellenfunktion setzt aber eine vorübergehende Zerstörung der Kohärenz der Wellenfunktion voraus, bei der alle Cooper-Paare ihre Kondensationsenergie verlieren würden. Da dieser Effekt nicht auftritt, kann kein Fluxoid die supraleitende Schleife verlassen, und Dauerströme sind stabil.

[4] Die Messung wurde mit einem SQUID-Magnetometer durchgeführt, auf das am Ende dieses Abschnitts hingewiesen wird.

Josephson-Effekt

In Abschnitt 12.2 haben wir bereits Experimente mit Tunnelkontakten betrachtet und das Tunneln von Quasiteilchen diskutiert. Wenn wir die Dicke der Isolationsschicht zwischen den beiden Supraleitern auf etwa 1 nm reduzieren, reicht die Wellenfunktion des einen Supraleiters in den Bereich des anderen. Mit gewisser Wahrscheinlichkeit können nun auch Cooper-Paare durch die Isolatorschicht tunneln, wodurch die Wellenfunktionen der beiden Supraleiter gekoppelt werden. Dies gibt Anlaß zu einigen überraschenden Effekten, die bereits 1962 von *Josephson* vorausgesagt wurden.

Wir wollen einige theoretische Überlegungen nachvollziehen und schreiben deshalb zunächst die zeitabhängige Schrödinger-Gleichungen für die beiden Supraleiter an:

$$\mathrm{i}\hbar\dot{\Psi}_1 = \mu_1\Psi_1 + \mathcal{K}\Psi_2 \qquad \text{und} \qquad \mathrm{i}\hbar\dot{\Psi}_2 = \mu_2\Psi_2 + \mathcal{K}\Psi_1 \,. \tag{12.39}$$

Die Variablen μ_1 und μ_2 sind die chemischen Potentiale der beiden Supraleiter und spiegeln die Lösungen der ungekoppelten Supraleiter wider, die als bekannt vorausgesetzt werden. Die Konstante $\mathcal{K}$ drückt die Stärke der Kopplung zwischen den beiden aus. Liegt keine Spannung zwischen den beiden Kontakten an, dann ist das chemische Potential natürlich in der ganzen Probe gleich, d.h. $\mu_1 = \mu_2$. Fällt die Spannung U ab, so verschieben sich die Fermi-Niveaus der beiden Supraleiter gegeneinander, und $\mu_2 - \mu_1 = -2eU$. In unserer Rechnung nehmen wir an, daß die Supraleiter aus dem gleichen Material bestehen und die gleiche Cooper-Paardichte $n_{\mathrm{s1}} = n_{\mathrm{s2}} = n_{\mathrm{s}}$ aufweisen. Wir benutzen die Wellenfunktion (12.34) und berücksichtigen, daß in dem hier diskutierten Experiment die Dichte der Cooper-Paare und die Phase von der Zeit abhängen können. Setzen wir (12.34) in die obigen Gleichungen ein, so erhalten wir nach Trennung von Real- und Imaginärteil:

$$\dot{n}_{\mathrm{s1}} = \frac{2\mathcal{K}}{\hbar}\, n_{\mathrm{s}} \sin(\varphi_2 - \varphi_1) = -\dot{n}_{\mathrm{s2}} \,, \tag{12.40}$$

$$\hbar\,(\dot{\varphi}_2 - \dot{\varphi}_1) = -(\mu_2 - \mu_1) \,. \tag{12.41}$$

Betrachten wir zunächst den Fall, daß zwischen den Tunnelkontakten keine Spannung herrscht. Nach Gleichung (12.41) ist in diesem Fall die Phasendifferenz $(\varphi_1 - \varphi_2)$ zwischen den beiden Wellenfunktionen zeitunabhängig. Da nach Gleichung (12.40) $\dot{n}_{\mathrm{s}} \neq 0$ ist, bedeutet dies, daß in dem Tunnelkontakt ein Gleichstrom I_{J} fließt und dennoch in dem dünnen Isolator keine Spannung abfällt. Dieser Strom ist proportional zu $\dot{n}_{\mathrm{s1}}$, so daß wir für ihn

$$I_{\mathrm{J}} = I_{\mathrm{max}} \sin(\varphi_2 - \varphi_1) \tag{12.42}$$

schreiben können. Dieses erstaunliche Phänomen bezeichnet man als *Josephson-Gleichstrom-Effekt*. In den Ausdruck für den maximalen Strom I_{max} geht neben der Dichte n_{s} der Cooper-Paare und der Fläche A (typischer Wert etwa 0,1 mm^2) des Tunnelkontakts vor allem die Kopplungsstärke $\mathcal{K}$ und damit die Dicke der Isolatorschicht ein, d.h. $I_{\mathrm{max}} \propto n_{\mathrm{s}}\mathcal{K}A$. Solange $I_{\mathrm{J}} < I_{\mathrm{max}}$ ist, bestimmt der „Nachschub" aus der Stromquelle (vgl. Abb. 12.13a) die Stromstärke und damit auch die Phasendifferenz $(\varphi_2 - \varphi_1)$ zwischen den beiden makroskopischen Wellenfunktionen. Die Strom-Spannungs-Charakteristik eines Pb/PbO/Pb-Josephsonkontakts ist in Abb. 12.19 wiedergegeben.

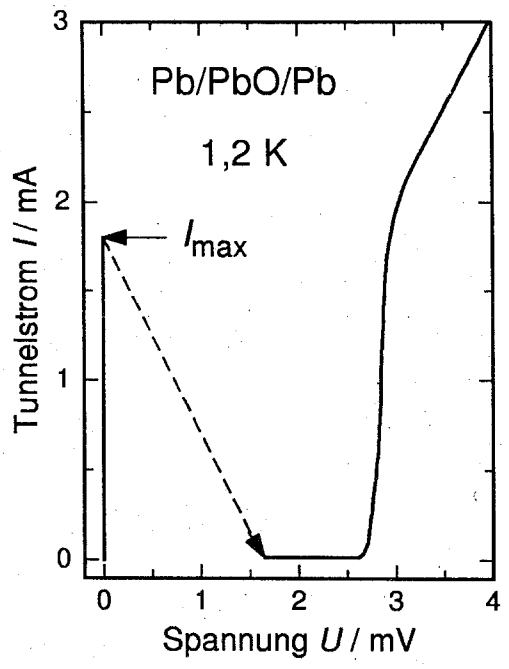


Abb. 12.19: Strom-Spannungscharakteristik eines Pb/PbO/Pb-Tunnelkontakts. Bei Vergrößerung des Tunnelstroms durch den Kontakt springt der Arbeitspunkt von I_{max} auf die Quasiteilchenkennlinie. (Nach D.N. Langenberg, D.J. Scalapino, B.N.Taylor, Proc. IEEE **54** (1966) 560).

Vergrößert man den Stromfluß im Kreis, so springt bei I_{max} die Spannung am Kontakt auf einen endlichen Wert, der durch die Quasiteilchenkennlinie bestimmt wird. Nun ist $\mu_2 - \mu_1 = -2eU$, und die Phasendifferenz wächst linear mit der Zeit an. Die Integration der Gleichung (12.41) ergibt

$$\varphi_2 - \varphi_1 = \frac{2eU}{\hbar} t + \varphi_0 = \omega_J t + \varphi_0 \,. \tag{12.43}$$

Setzen wir dieses Ergebnis in Gleichung (12.42) ein, so erhalten wir

$$I = I_{max} \sin(\omega_J t + \varphi_0) \,. \tag{12.44}$$

Es tritt ein Wechselstrom mit der Kreisfrequenz $\omega_J = 2eU/\hbar$ auf. Dies ist der *Wechselstrom-Josephson-Effekt.* In der Gleichstromcharakteristik der Abb. 12.19 ist dieser Strom natürlich nicht zu sehen. Die auftretenden Frequenzen sind relativ hoch, denn eine Spannung von 100 μV führt bereits zu einer Frequenz von 50 GHz. Da Spannung und Frequenz nur über das Verhältnis von e/h verknüpft sind, läßt sich dieses Verhältnis mit Hilfe des Josephson-Effekts sehr genau bestimmen. Andererseits kann man bei Kenntnis der Konstanten e/h Spannungen über eine Frequenzmessung mit hoher Präzision bestimmen. Der Josephson-Effekt wird daher zur Herstellung neuartiger Spannungsnormale ausgenutzt.

Der direkte Nachweis des hochfrequenten Josephson-Wechselstroms ist schwierig und für praktische Anwendungen nicht geeignet, da die auftretende Mikrowellenleistung sehr klein ist und Auskopplung und Nachweis der Wellen äußerst schwierig sind. Es gibt jedoch andere Meßmethoden, die sich relativ einfach realisieren lassen und bei denen der Effekt indirekt beobachtet wird. Eine Methode besteht darin, den Strom durch den Josephson-Kontakt so einzustellen, daß Spannung abfällt, und den Kontakt gleichzeitig mit Mikrowellen der Frequenz ω zu bestrahlen. Dadurch fließt ein frequenzmodulierter Strom mit Komponenten $\omega_J \pm p\omega$ durch den Kontakt, wobei p eine ganze Zahl ist. Immer dann, wenn die Josephson-Frequenz ω_J mit einem Vielfachen der eingestrahlten Frequenz übereinstimmt, tritt ein Seitenband mit der Frequenz $\omega = 0$, also ein Gleichstrom auf,

der leicht nachgewiesen werden kann. Auf diesen und ähnlichen Prinzipien beruhen meist die indirekten Messungen des Josephson-Wechselstrom-Effekts.

Wir betrachten nun den Einfluß eines Magnetfelds auf das Tunneln der Cooper-Paare. Da die Ableitung der relevanten Gleichungen im Falle einzelner Tunnelkontakte etwas mühevoll ist, diskutieren wir hier zwei parallel geschaltete, identische Josephson-Kontakte, wie sie in Abb. 12.20 schematisch dargestellt sind. Die Kontaktstellen sollen so klein sein, daß ihre Ausdehnung bei den folgenden Überlegungen vernachlässigt werden kann.

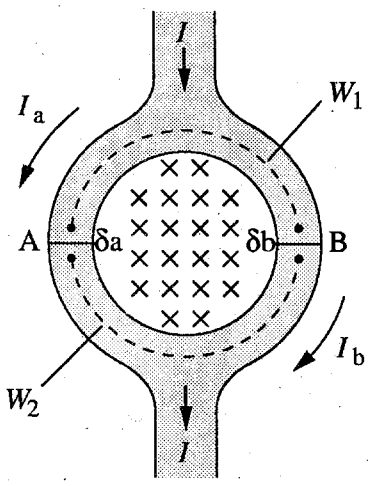


Abb. 12.20: Stromfluß durch zwei parallel geschaltete, identische Josephson-Kontakte A und B mit eingeschlossenem Magnetfluß (durch Kreuze angedeutet). An den Kontaktstellen besteht die Phasendifferenz δ_a bzw. δ_b. Die Integration der Phasen erfolgt längs der strichliert eingezeichneten Wege W_1 und W_2.

Der Gesamtstrom I setzt sich aus den Teilströmen durch die beiden einzelnen Kontakte zusammen:

$$I = I_{\max}(\sin\delta_a + \sin\delta_b) = 2\,I_{\max}\cos\left(\frac{\delta_a - \delta_b}{2}\right)\sin\left(\frac{\delta_a + \delta_b}{2}\right) . \tag{12.45}$$

Hierbei wurde mit $\delta_a = (\varphi_{a1} - \varphi_{a2})$ und $\delta_b = (\varphi_{b1} - \varphi_{b2})$ die Phasendifferenz an den Kontakten A und B bezeichnet. Wie wir gleich sehen werden, wird die Differenz $(\delta_a - \delta_b)$ vom Magnetfluß Φ durch den Ring bestimmt. Bei der Berechnung der Phase gehen wir wie bei der Diskussion der Flußquantisierung vor und legen den Integrationsweg wieder in den supraleitenden Bereich der Probe, in dem kein Stromfluß auftritt. Für die Phasendifferenz längs der Wege W_1 und W_2 folgt dann nach Gleichung (12.37):

$$\varphi_{a1} - \varphi_{b1} = \frac{2e}{\hbar}\int\limits_{W_1} \boldsymbol{A}\cdot\mathrm{d}\boldsymbol{s} \quad \text{und} \quad \varphi_{b2} - \varphi_{a2} = \frac{2e}{\hbar}\int\limits_{W_2} \boldsymbol{A}\cdot\mathrm{d}\boldsymbol{s} . \tag{12.46}$$

Addieren wir die beiden Gleichungen auf, so ergibt sich

$$\delta_a - \delta_b = \frac{2e}{\hbar}\oint \boldsymbol{A}\cdot\mathrm{d}\boldsymbol{s} = \frac{2e\Phi}{\hbar} . \tag{12.47}$$

Hierbei haben wir die Beiträge des Felds in den Kontakten vernachlässigt. Setzen wir Gleichung (12.47) in (12.45) ein, so finden wir

$$I = 2I_{\max}\sin\left(\frac{\delta_a + \delta_b}{2}\right)\cos\left(\frac{\pi\Phi}{\Phi_0}\right) . \tag{12.48}$$

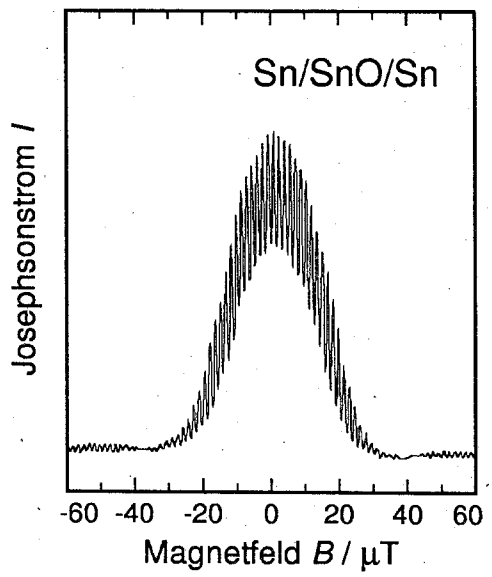


Abb. 12.21: Strom durch zwei parallel geschaltete Josephson-Kontakte als Funktion des Magnetfelds. (Nach R.C. Jaklevic, J. Lambe, J.E. Mercereau, A.H. Silver, Phys. Rev. **140** (1965) A 1628).

Der eingeprägte Strom I bestimmt den Wert des Phasenwinkels $(\delta_a + \delta_b)/2$, auf den das Magnetfeld keinen Einfluß hat. Der Kosinus-Term beschreibt die Oszillation des Stroms mit dem Magnetfeld. Der experimentelle Stromverlauf durch einen Doppelkontakt ist in Abb. 12.21 gezeigt. Deutlich zu sehen ist, daß der Stromfluß auf extrem kleine Magnetfeldänderungen reagiert. Die Einhüllende des Tunnelstroms wird, wie wir gleich sehen werden, von der endlichen Dimension der einzelnen Josephson-Kontakte hervorgerufen.

Untersucht man die Magnetfeldabhängigkeit des Tunnelstroms durch *einen* Kontakt, so muß der Magnetfluß *im* Josephson-Kontakt berücksichtigt werden. Man findet dann

$$I = I_{\max} \left| \frac{\sin(\pi\Phi/\Phi_0)}{(\pi\Phi/\Phi_0)} \right| , \tag{12.49}$$

in Übereinstimmung mit experimentellen Resultaten, wie sie in Abb. 12.22 gezeigt sind.

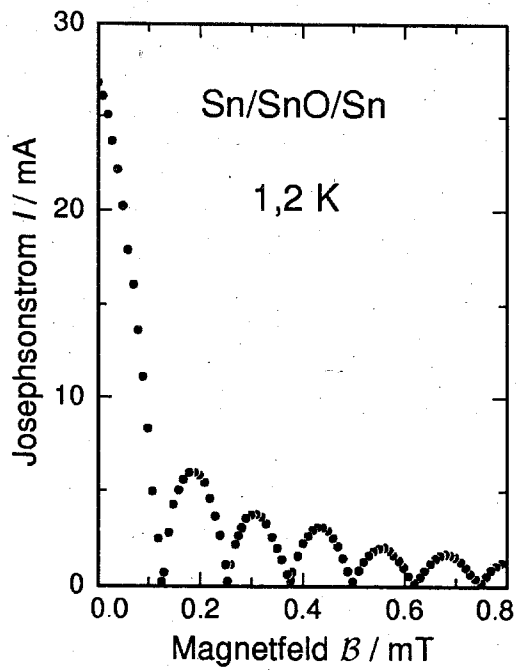


Abb. 12.22: Magnetfeldabhängigkeit des Stroms durch einen einzelnen Josephson-Kontakt. Das anliegende Magnetfeld verläuft parallel zur Isoalatorschicht des Sn/SnO/Sn-Tunnelkontakts. (Nach D.N. Langenberg, D.J. Scalapino, B.N. Taylor, Proc. IEEE **54** (1966) 560).

Die geschilderten Gleichungen und die gezeigten experimentellen Ergebnisse demonstrieren, daß große Ähnlichkeit mit Phänomenen aus der Optik besteht. Offenbar existiert eine formale Verwandtschaft mit der Lichtbeugung am einfachen und am Doppelspalt. Dies wird besonders deutlich, wenn wir den Stromverlauf durch den Doppelkontakt noch einmal betrachten. Wie beim optischen Doppelspalt die Einhüllende der Lichtintensität durch die Spaltdimensionen bestimmt wird, so wird hier die Einhüllende des Tunnelstroms durch Abmessungen der einzelnen Josephson-Kontakte festgelegt. Die gemeinsame phänomenologische Grundlage der Beobachtungen auf ganz unterschiedlichen Gebieten ist die Interferenz von Wellen.

Die Interferenz der makroskopischen Wellenfunktion der Supraleiter wird in kommerziell erhältlichen Meßgeräten benutzt. Diese sogenannten SQUIDs (Superconductive Quantum Interference Device) eignen sich zum Nachweis kleinster Magnetfeldänderungen bis zu 10^{-14} T und dienen darüberhinaus als überaus empfindliche Ampère- und Voltmeter. Im wesentlichen handelt es sich hierbei um supraleitende Ringe oder Zylinder mit einem oder zwei Bereichen der „schwachen Kopplung", die wie Josephson-Kontakte wirken und durch die das Magnetfeld in den Ring eintreten kann.

12.4 Supraleiter 2. Art

Bisher sind wir davon ausgegangen, daß die Cooper-Paardichte in der Probe konstant ist. Zwei einfache Argumente sprechen jedoch gegen die Gültigkeit dieser Annahme in der Nähe von Grenzflächen. Betrachten wir die Grenzfläche zwischen einem Supraleiter und einem Normalleiter bzw. Isolator. Im Normalleiter sind keine Cooper-Paare vorhanden, im Supraleiter ist ihre Dichte n_s. Da aber Cooper-Paare eine relativ große räumliche Ausdehnung aufweisen, kann sich ihre Dichte nicht sprunghaft ändern, sondern es muß ein stetiger Verlauf auftreten, wie er beispielsweise in Abb. 12.23 skizziert ist. Entsprechendes gilt für die Cooper-Paardichte an der Oberfläche einer supraleitenden Probe.

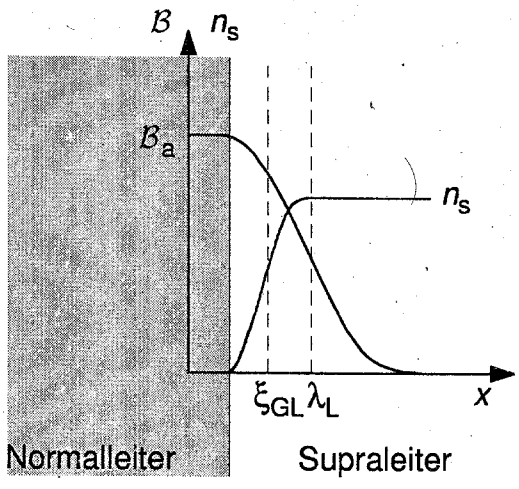


Abb. 12.23: Schematischer Verlauf des Magnetfelds $\mathcal{B}$ und der Dichte n_s der Cooper-Paare im Grenzbereich Normalleiter-Supraleiter. Der Verlauf des Magnetfelds wird durch die Eindringtiefe λ_L, der Verlauf der Cooper-Paardichte durch die Kohärenzlänge ξ_{GL} bestimmt. Der Proximity-Effekt wurde hier nicht berücksichtigt.

Das zweite Argument beruht auf der Tatsache, daß die Krümmung der Wellenfunktion die lokale kinetische Energie des Systems widerspiegelt. Wie ein Blick auf Gleichung (12.34) verdeutlicht, ist der Betrag der Wellenfunktion mit der Cooper-Paardichte verknüpft. Um eine Minimierung der freien Energie zu erreichen, vermeidet der Supraleiter deshalb eine sprunghafte Änderung, d.h. der Betrag der Wellenfunktion verändert

sich an Grenzflächen stetig. Formal kann der Anstieg der Cooper-Paardichte durch die *Kohärenzlänge* ξ_{GL} beschrieben werden, die immer größer als die Ausdehnung der Cooper-Paare (vgl. Abschnitt 12.2) ist. Der Index GL weist auf *Ginzburg* und *Landau* hin, in deren makroskopischer Theorie der Supraleitung diese Größe eine zentrale Rolle spielt.

Der Vollständigkeit halber sei erwähnt, daß auch im Normalleiter eine kleine Zahl von Cooper-Paaren in unmittelbarer Nachbarschaft zur Grenzfläche anzutreffen ist („Proximity-Effekt"), die zu einer Reihe interessanter Effekte Anlaß geben. Ein spezielles Beispiel hierfür ist der bereits diskutierte Josephson-Effekt, bei dessen Auftreten man auch Cooper-Paare im Isolator antrifft.

Grenzflächenenergie und Supraleiter 2. Art

Grenzflächen verkleinern die Kondensationsenergie, da dort die Cooper-Paardichte reduziert ist. Für eine grobe Abschätzung der Bedeutung dieses Effekts nehmen wir an, daß eine Schicht mit der Dicke der Kohärenzlänge frei von Cooper-Paaren ist. Mit Hilfe von Gleichung (12.10) läßt sich die fehlende Kondensationsenergie berechnen, und wir finden $\Delta E_{\text{kon}} = A\xi_{GL}\mathcal{B}_c^2/2\mu_0$, wobei A die Fläche der Grenzschicht ist. Supraleiter werden daher den Aufbau von Grenzflächen möglichst vermeiden. Legen wir ein Magnetfeld an, so müssen die Supraleiter nach Gleichung (12.9) die Verdrängungsenergie $E_B = V\mathcal{B}^2/2\mu_0$ aufbringen. Das Auftreten von Grenzflächen reduziert die aufzuwendende Energie und begünstigt somit deren Ausbildung unter dem Einfluß eines Magnetfelds. In grober Näherung schreiben wir für die gewonnene Verdrängungsenergie $\Delta E_B = A\lambda_L\mathcal{B}^2/2\mu_0$, wobei λ_L für die Londonsche Eindringtiefe steht. Für die Energieänderung, die eine Grenzfläche hervorruft, finden wir somit den Ausdruck

$$\Delta E_G = \Delta E_{\text{kon}} - \Delta E_B = \frac{A}{2\mu_0}\mathcal{B}_c^2\left(\xi_{GL} - \lambda_L\frac{\mathcal{B}^2}{\mathcal{B}_c^2}\right). \tag{12.50}$$

Ist $\xi_{GL} > \lambda_L$, so bleibt ΔE_G für alle Magnetfelder positiv, d.h. die Ausbildung einer Grenzfläche, zusätzlich zur Oberfläche, wird unterdrückt. Ist dagegen $\xi_{GL} < \lambda_L$, so ist bereits bei $\mathcal{B} < \mathcal{B}_c$ die Ausbildung von weiteren Grenzflächen energetisch günstiger.

Tatsächlich beobachtet man diesen Effekt bei den *Supraleitern 2. Art.* Die typische Reaktion dieser Klasse von Supraleitern auf ein äußeres Magnetfeld ist in Abb. 12.24 skizziert.

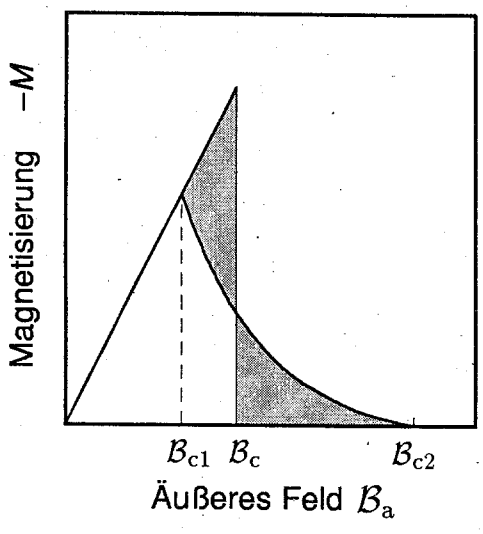


Abb. 12.24: Verlauf der Magnetisierung bei Supraleitern 1. und 2. Art. Die Magnetisierung steigt bei Supraleitern 1. Art bis $\mathcal{B}_c$ an und bricht dann zusammen. Bei Supraleitern 2. Art nimmt sie im Bereich $\mathcal{B}_{c1} < \mathcal{B} < \mathcal{B}_{c2}$ langsam ab, da das Feld teilweise eindringt. Die beiden dunklen Bereiche sind flächengleich.

Sieht man von der Oberfläche ab, so bleibt bei kleinen Feldern wie bei den bisher diskutierten *Supraleitern 1. Art* das Probeninnere feldfrei, und der Supraleiter verhält sich wie ein idealer Diamagnet. Man bezeichnet diesen Bereich, in dem sich die beiden Typen von Supraleiter nicht unterscheiden, als *Meißner-Phase*. Bei der Feldstärke $\mathcal{B}_{c1}$ wechselt die Oberflächenenergie ihr Vorzeichen, und das Magnetfeld dringt in den Supraleiter ein, d.h. äußere Magnetfelder werden nur noch teilweise abgeschirmt. Man bezeichnet diesen Zustand als die *Shubnikov-Phase*. Erst beim Überschreiten der noch höheren kritischen Magnetfeldstärke $\mathcal{B}_{c2}$ wird die Probe vollständig normalleitend.

Für die kritischen Feldstärken gilt $\mathcal{B}_{c1} < \mathcal{B}_c < \mathcal{B}_{c2}$, wobei aus der thermodynamischen Definition von $\mathcal{B}_c$ folgt, daß die beiden dunklen Flächen in Abb. 12.24 gleich groß sind. Wie in Abb. 12.25 für eine Indium-Wismut-Legierung zu sehen ist, weisen die kritischen Feldstärken einen ähnlichen Temperaturverlauf auf; in erster Näherung unterscheiden sich nur deren Vorfaktoren. Das thermodynamische kritische Feld $\mathcal{B}_c$ ist nach wie vor durch Gleichung (12.10) über die Kondensationsenergie definiert.

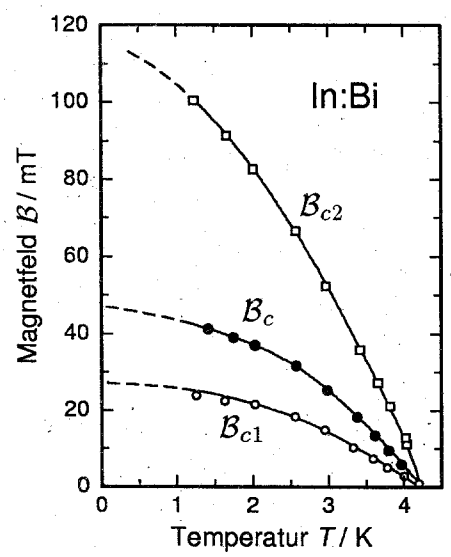


Abb. 12.25: Kritische Magnetfeldstärken einer Indium-Wismut-Legierung. Neben den beiden Feldstärken $\mathcal{B}_{c1}$ und $\mathcal{B}_{c2}$ ist auch die thermodynamisch kritische Feldstärke $\mathcal{B}_c$ eingetragen. (Nach T. Kinsel, E.A. Lyntor, B. Serin, Rev. Mod. Phys. **36** (1964) 105).

Typische Vertreter der Supraleiter 2. Art sind Legierungen, Übergangsmetalle, metallische Gläser und auch die neuartigen Oxidsupraleiter, auf die wir sogleich kurz eingehen. Der Wert von $\mathcal{B}_{c2}$ kann sehr viel größer sein als das kritische Feld von Supraleitern 1. Art. So findet man für den Hochfeldsupraleiter $PbMo_6S_8$, der zu den Chevrel-Verbindungen[5] gehört, bei $T = 0$ den Wert $\mathcal{B}_{c2} = 60\,\mathrm{T}$. Technisch von großer Bedeutung ist Nb_3Sn, das sich relativ gut bearbeiten läßt und zu den sogenannten A15-Verbindungen[6] gehört. Es besitzt die hohen kritischen Konstanten $T_c = 18{,}7\,\mathrm{K}$ und $\mathcal{B}_{c2}(T = 0) = 25\,\mathrm{T}$.

Tabelle 12.3: Sprungtemperatur T_c und kritische Feldstärke $\mathcal{B}_{c2}$ einiger Supraleiter 2. Art.

Supraleiter	MoC	Nb_3Sn	Nb_3Ge	Nb_3Ga	V_3Si	$PbMo_6S_8$	$PbMo_6Se_8$
T_c (K)	14,3	18,7	23,2	20,3	17,1	15,3	6,7
$\mathcal{B}_{c2}$ (T)	9,8	25	38	34	25	60	7

[5] Chevrel-Verbindungen besitzen die Zusammensetzung MMo_6X_8, wobei M für ein Metall der Seltenen Erden und X für Schwefel oder Selen steht.

[6] A15-Verbindungen weisen die Zusammensetzung A_3B auf und kristallisieren in der β-Wolframstruktur. Die Niobatome sind kettenförmig angeordnet und besitzen dort einen Abstand, der kleiner ist als im metallischen Niob.

Es stellt sich eine sehr interessante Frage: Wie verteilt sich der Magnetfluß in den Supraleitern 2. Art? In Abb. 12.26 ist der Zwischenzustand, die Shubnikov-Phase anschaulich dargestellt. Das Magnetfeld tritt durch kleine normalleitende Inseln in Form von *Flußschläuchen* durch. Jeder Flußschlauch trägt ein Fluxoid, da die in Abschnitt 12.3 vorgebrachten Argumente für einen supraleitenden Ring auch hier gelten. In perfekten Kristallen ordnen sich die Flußschläuche regelmäßig in einer *Abrikosov-Struktur* an. Nach „Dekoration" mit feinen Eisenkolloidteilchen konnte dieses Flußgitter mit Hilfe eines Elektronenmikroskops sichtbar gemacht werden.

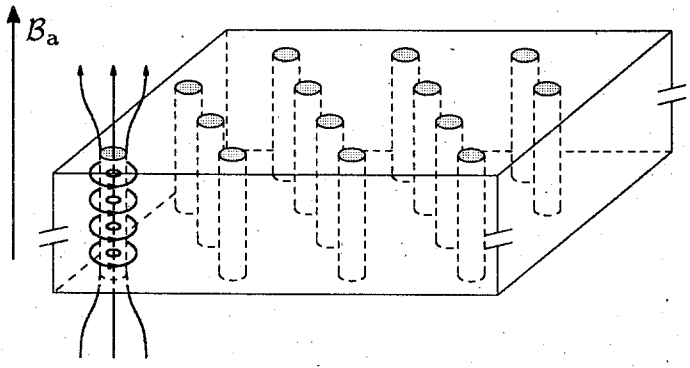


Abb. 12.26: Anordnung der Flußschläuche in der Shubnikov-Phase. Für einen Flußschlauch sind Magnetfeldverlauf und Abschirmströme eingezeichnet.

Seit einiger Zeit können derartige Gitter auch mit dem Rastertunnelmikroskop untersucht werden. Ein Beispiel hierfür zeigt Abb. 12.27 mit der Wiedergabe der Flußschläuche in $NbSe_2$. Der Kontrast bei der Aufnahme beruht auf der unterschiedlichen Austrittsarbeit und damit auf der unterschiedlichen Tunnelstromcharakteristik von normal- und supraleitenden Bereichen. Tatsächlich kann in derartigen Messungen sogar die elektronische Zustandsdichte ortsaufgelöst in und um die Flußschläuche ausgemessen werden.

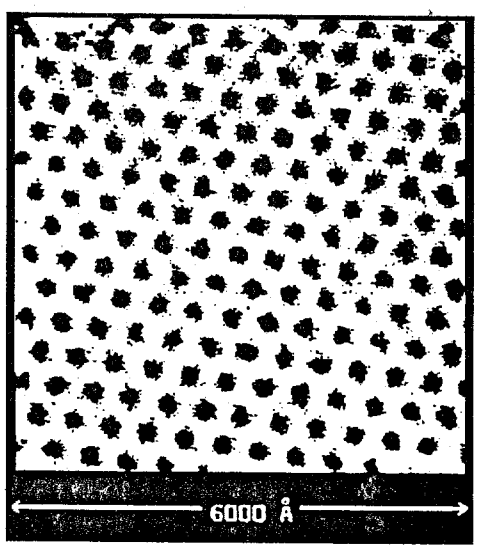


Abb. 12.27: Abbildung der Abrikosov-Struktur mit einem Rastertunnelmikroskop. Das Flußgitter wurde an $NbSe_2$ bei 1,8 K und einem Magnetfeld $\mathcal{B} = 1\,\mathrm{T}$ aufgenommen. (Aus H.F. Hess, R.B. Robinson, R.C. Dynes, R.M. Valles Jr., J.V. Waszczak, Phys. Rev. Lett. **62** (1989) 214).

Hochtemperatursupraleiter

Es ist immer versucht worden, Supraleiter mit besonders hoher Sprungtemperatur zu entdecken bzw. zu entwickeln, wobei die „richtige" Vorgehensweise nicht bekannt war. Wenn nicht bei Zimmertemperatur, so sollte T_c doch zumindest über 77 K, dem Siedepunkt des flüssigen Stickstoffs liegen. Ein entscheidender Schritt gelang 1986 *Bednorz* und *Müller*, als sie das Mischsystem Ba-La-Cu-O untersuchten und eine Sprungtemperatur von etwa 30 K fanden, die deutlich über den bis dahin bekannten Werten lag. Etwas später wurden an Systemen der Zusammensetzung Ba-Y-Cu-O Werte bis 92 K, an Tl-Ca-Ba-Cu-O bis 125 K und vor kurzem an Hg-Ba-Ca-Cu-O bis 132,5 K erreicht.

Wie Tabelle 12.1 zu entnehmen ist, weisen die Hochtemperatursupraleiter eine relativ komplizierte Zusammensetzung auf. Es ist schwierig, Einkristalle zu züchten oder stöchiometrische Proben herzustellen, so daß viele veröffentlichte Messungen an Mischphasensystemen durchgeführt wurden und nur begrenzt interpretierbar sind. Diese Problematik wurde bereits bei der ersten Veröffentlichung von Bednorz und Müller deutlich. Der Übergang vom Normal- zum Supraleiter erfolgte nicht sprungartig, sondern in einem Temperaturbereich von etwa 10 K.

Das bestuntersuchte System ist $YBa_2Cu_3O_{7-x}$, das oft mit YBCO oder wegen der relativen Anzahl der Metallatome auch mit Y123 bezeichnet wird. Der Index x deutet an, daß die Zusammensetzung dieser Verbindung in weiten Bereichen variieren kann. Der Wert von x bestimmt die elektrischen Eigenschaften. So ist YBCO bei $x = 1$ ein Isolator, in dem die Kupferatome antiferromagnetisch geordnet sind. In der Nähe von $x = 0{,}7$ tritt ein Isolator-Metall-Übergang auf. In der sich anschließenden metallischen Phase wird die Leitfähigkeit von Löchern hervorgerufen, deren Zahl mit abnehmendem x ansteigt und bei $x = 0{,}1$ schließlich 0,2 Löcher pro Kupferatom erreicht. Knapp unter $x = 0{,}7$ setzt Supraleitung ein, und die Sprungtemperatur steigt mit zunehmender Sauerstoffkonzentration von $T_c = 40$ K auf 92 K an. Die „besten" supraleitenden Eigenschaften beobachtet man bei $x = 0{,}08$. Abb. 12.28 zeigt die für die Hochtemperatursupraleiter typische Struktur von YBCO.

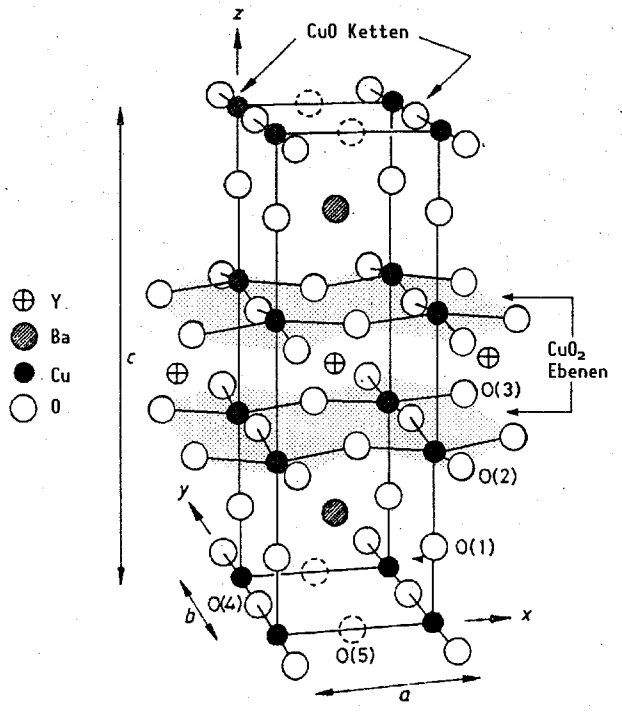


Abb. 12.28: Kristallstruktur des Hochtemperatursupraleiters $YBa_2Cu_3O_{7-x}$. Die Sauerstoffatome können fünf verschiedene Plätze einnehmen, die mit O(1) bis O(5) bezeichnet sind. Bei $x \approx 0{,}1$ sind die O(4)-Plätze fast voll, die O(5)-Plätze dagegen leer.

Das wichtigste Merkmal stellen CuO-Ebenen dar, die für die Ausbildung der Supraleitung verantwortlich sind. Diese Ebenen sind hier durch Yttriumatome bzw. BaO-Lagen voneinander getrennt. Die verschiedenen Hochtemperatursupraleiter unterscheiden sich in der Anzahl und im Aufbau dieser Zwischenschichten. Aufgrund ihrer Schichtstruktur besitzen diese neuen Materialien stark anisotrope Eigenschaften.

Der anisotrope Aufbau spiegelt sich beispielsweise auch im elektrischen Widerstand von Einkristallen im normalleitenden Zustand wider, der in Abb. 12.29 für Messungen an YBCO in den drei Kristallrichtungen dargestellt ist. Wie man aufgrund der Struktur erwartet, tritt der höchste Widerstand in Richtung der c-Achse, also senkrecht zu den Schichten auf.

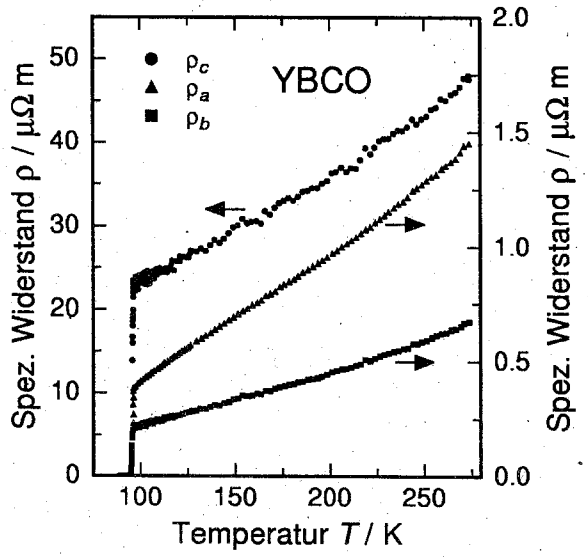


Abb. 12.29: Widerstand eines YBCO-Einkristalls in Richtung der Kristallachsen a, b und c. Zu beachten sind die unterschiedlichen Skalen für ρ_a, ρ_b bzw. ρ_c. (Nach T.A. Friedmann, M.V. Rabin, J. Giapintzakis, J.P. Rice, D.M. Ginzberg, Phys. Rev. B **42** (1990) 6217).

Auch die supraleitenden Eigenschaften weisen starke Anisotropien auf. In Abb. 12.30 sind Messungen der kritischen Feldstärke $\mathcal{B}_{c1}$ von YBCO in verschiedenen Kristallrichtungen wiedergegeben. Das kritische Feld in Richtung der c-Achse ist deutlich größer als senkrecht zu diesen Achsen.

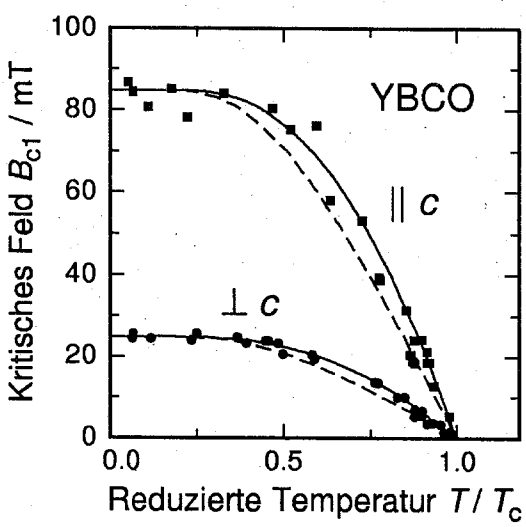


Abb. 12.30: Kritische Feldstärke $\mathcal{B}_{c1}$ von YBCO parallel und senkrecht zur c-Achse. Es tritt eine deutliche Anisotropie im kritischen Verhalten auf. (Nach D.H. Wu und S. Sridhar, Phys. Rev. Lett. **65** (1990) 2074).

Es läßt sich zeigen, daß bei Supraleitern 2. Art $\mathcal{B}_{c1}$ über die Beziehung

$$\mathcal{B}_{c1} \approx \frac{\Phi_0}{4\pi\lambda_L^2} \tag{12.51}$$

mit der Eindringtiefe des Magnetfelds zusammenhängt. Der relativ kleine Wert von $\mathcal{B}_{c1}$ zeigt uns, daß λ_L groß ist und das Magnetfeld weit eindringt. So findet man bei $T = 0$ mit $\lambda_L \approx 150\,\text{nm}$ einen Wert, der wesentlich größer ist als bei den bisher bekannten Supraleitern. Wegen $\lambda_L^2 \propto n_s^{-1}$ folgt, daß in den Hochtemperatursupraleitern die Dichte der Cooper-Paare relativ gering ist. Um die Richtungsabhängigkeit zu beschreiben, führt man in die Ginzburg-Landau-Theorie richtungsabhängige effektive Massen ein. Wir wollen jedoch auf die Anisotropieeffekte und ihre Beschreibung nicht weiter eingehen.

Für das kritische Feld $\mathcal{B}_{c2}$ liefert die Ginzburg-Landau-Theorie einen Ausdruck ähnlich wie Gleichung (12.51):

$$\mathcal{B}_{c2} \approx \frac{\Phi_0}{2\pi\xi_{GL}^2} \,. \tag{12.52}$$

Die großen Werte von $\mathcal{B}_{c2}$ sind also mit einer kleinen Kohärenzlänge ξ_{GL} verknüpft. So ergeben sich bei $T = 0$ Werte von ξ_{GL} von nur etwa 1,5 nm in Schichtrichtung und 0,3 nm in Richtung der c-Achse. Damit verbunden ist eine außerordentlich geringe Ausdehnung der Cooper-Paare über nur wenige Atome. Vergleicht man die Größenordnungen von λ_L und ξ_{GL}, so wird klar, daß es sich bei den Hochtemperatursupraleitern um extreme Supraleiter 2. Art handelt.

In Abb. 12.31 sind Messungen von $\mathcal{B}_{c2}$ gezeigt. Von klassischen Supraleitern 2. Art ist bekannt, daß in grober Näherung $\mathcal{B}_{c1}\mathcal{B}_{c2} \approx \mathcal{B}_c^2$ ist, und entsprechendes gilt auch bei Hochtemperatursupraleitern für eine vorgegebene Richtung. In Übereinstimmung hiermit findet man in Richtung der c-Achse eine wesentlich kleinere kritische Feldstärke $\mathcal{B}_{c2}$ als senkrecht dazu. Zu tieferen Temperaturen hin kann der Verlauf nicht verfolgt werden, da bereits wenige Kelvin unterhalb von T_c kritische Feldstärken erreicht werden, die experimentell nicht mehr zugänglich sind. Dennoch läßt sich aus diesen Daten der Wert von $\mathcal{B}_{c2}(T = 0)$ errechnen. Mit einer Näherung, die wir nicht diskutiert haben, findet man für die beiden Richtungen die extrem hohen Werte von 120 T bzw. 620 T.

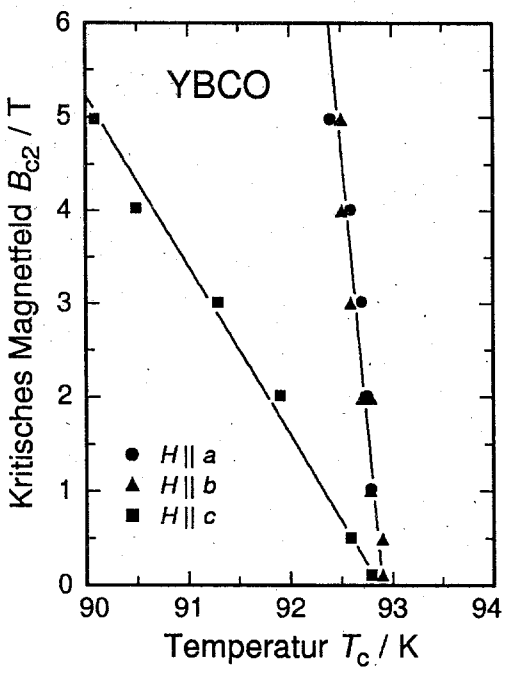


Abb. 12.31: Kritische Feldstärke $\mathcal{B}_{c2}$ von YBCO in Richtung der drei Kristallachsen. Parallel zur c-Achse ist die kritische Feldstärke wesentlich kleiner als in der ab-Ebene. (Nach U. Welp, M. Grimsditch, H. You, W.K. Kwok, M.M. Fang, G.W. Crabtree, J.Z. Liu, Physica C **161** (1989) 1).

Über den mikroskopischen Mechanismus der Hochtemperatursupraleitung läßt sich gegenwärtig keine eindeutige Aussage machen. Wir haben Größen aus der London- und Ginzburg-Landau-Theorie bei der Beschreibung einiger Beobachtungen benutzt. Dies ist natürlich möglich, weil diese Theorien nicht auf mikroskopische Ursachen eingehen. Ein wichtiger Punkt ist, daß im Unterschied zu klassischen Supraleitern der elektrische Ladungstransport, wie bereits oben erwähnt, über Löcher erfolgt. Messungen haben gezeigt, daß auch in den Hochtemperatursupraleitern der Strom von Cooper-Paaren getragen wird und Flußquantisierung mit dem Fluxiod $\Phi_0 = h/2e$ auftritt. Mit verschiedenen Methoden wurde die Energielücke gemessen und mit dem Wert $2\Delta(0)/k_B T_c = 3{,}52$ der BCS-Theorie verglichen. Die gefundenen Werte schwanken zwischen 5 und 7. Dies sind Werte, die bei stark koppelnen Supraleitern nicht ungewöhnlich sind. Aus dem Sprung der spezifischen Wärme bei T_c sollte auch eine direkte Aussage zur Übereinstimmung mit der BCS-Theorie möglich sein. Doch auch hier ergeben sich experimentelle Schwierigkeiten. Da der Phasenübergang bei so hohen Temperaturen auftritt, ist der Beitrag des Gitters dominant. Darüber hinaus läßt sich die Referenzmessung im Normalzustand wegen der hohen kritischen Feldstärke nur in einem sehr engen Temperaturbereich durchführen. In Abb. 12.32 ist eine Messung der spezifischen Wärme von YBCO wiedergegeben. Ein Vergleich mit den Aluminiumdaten in Abb. 12.7 macht deutlich, daß der Gitterbeitrag stark überwiegt. Weiter fällt auf, daß die spezifische Wärme beim Übergang vom Normal- zum Supraleiter keinen abrupten Sprung macht. Diese „Abrundung" der Daten ist auf die starke Fluktuation zurückzuführen, der die Dichte der Cooper-Paare in der Nähe des Sprungpunktes unterliegt. Dies bedeutet, daß knapp oberhalb von T_c bereits Cooper-Paare existieren, deren Zahl jedoch zeitlich und räumlich stark schwankt.

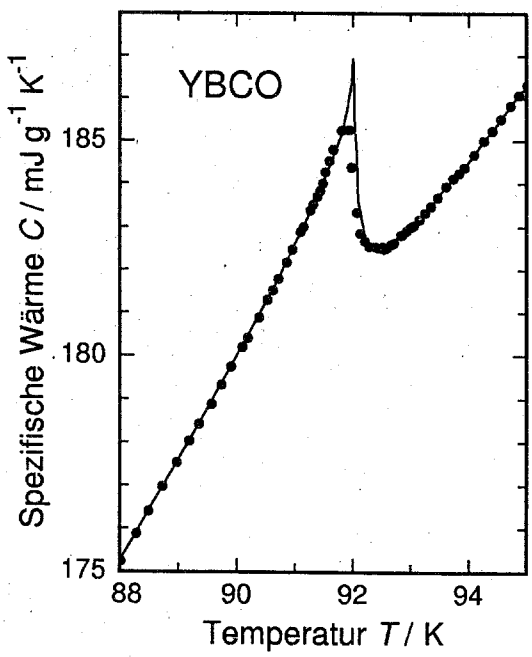


Abb. 12.32: Spezifische Wärme von YBCO in der Nähe des Sprungpunktes. Man beachte, daß der Nullpunkt der Ordinate unterdrückt ist. (Nach N. Overend, M.A. Howson, I.D. Lawrie, Phys. Rev. Lett. **72** (1994) 3238).

Mit großer Wahrscheinlichkeit besitzen die Cooper-Paare der Hochtemperatursupraleiter eine d-Wellenfunktion. Noch nicht vollständig geklärt ist der Mechanismus, der die Kopplung bewirkt. Da an einigen Proben auch ein Isotopeneffekt beobachtet wurde, liegt der Schluß nahe, daß Phononen eine wichtige Rolle spielen. Ob aber der Austausch

virtueller Phononen ausreicht, d.h. ob die klassische BCS-Theorie auch Hochtemperatursupraleiter beschreiben kann, ist noch ungeklärt. Allen diesen Supraleitern gemeinsam ist eine Perowskit-Schichtstruktur mit CuO-Ebenen, in denen die Kupferatome in einem gemischtwertigen Cu^{2+}-Cu^{3+}-Zustand vorliegen und ein magnetisches Moment tragen. Möglicherweise hat auch die antiferromagnetische Ordnung einen Einfluß auf die Ausbildung der Supraleitung.

Zum Schluß sei noch auf die technische Anwendung hingewiesen. Sicherlich ist die Nutzung der Supraleitung in SQUIDs oder ähnlichen meßtechnischen Einrichtungen bei relativ hohen Temperaturen ein vielversprechendes Anwendungsfeld. Für den Bau von leistungsfähigen Magneten oder verlustlosen Stromkabeln sind sie gegenwärtig noch ungeeignet. Der Grund hierfür ist, daß Hochtemperatursupraleiter mit größerer Abmessung aus keramischem Material bestehen, also aus supraleitenden Körnern. Während in dünnen einkristallinen Filmen bei 4 K bereits Stromdichten von $5 \cdot 10^7$ A/cm^2 erreicht wurden, sind die erreichbaren Ströme in keramischem Material um viele Größenordnungen geringer. Die Ginzburg-Landau-Theorie sagt für Einkristalle wesentlich höhere Werte voraus, als bisher beobachtet wurden. Gegenwärtig wird der kritische Strom durch die Bewegung der Flußschläuche begrenzt, die zu Verlusten im Supraleiter führt.

Sachwortregister

Abrikosov-Struktur 435
Abschirmlänge 214
Abschirmung, elektrostatische 213
Abschneidefrequenz 162
Abstoáung, atomare 70
–, Wechselwirkungspotential 71
Adiabatische Näherung 79, 125
Aktivierungsenergie
–, bei Leerstellendiffusion 96
–, Elektronen in amorphen Halbleitern 310
Akustischer Zweig 138
Akzeptor 103, 297
Alkalimetalle 89
Amorpher Festkörper 6
–, Schwingungsspektrum 168
–, Struktur 24
–, Strukturbestimmung 54
Anharmonizitt, Gitterpotential 175
Antibindender Zustand 82
Antiferrimagnetismus 361
Atom-Strukturfaktor 49, 51
–, bei Neutronenstreuung 51
–, bei Rntgenstreuung 51
Atomformfaktor 51
Auslenkung, atomare 128
Áustausch
–, indirekter 351
– -energie 81, 84
– -integral 81
– -koeffizient 349
– -wechselwirkung 348
Austrittsarbeit 216
Autokorrelationsfunktion 25
Band
– -aufspaltung 234
– -besetzung 239
– -diskontinuität 330
– -verbiegung 318, 330
Bandlücke 287
–, direkte 288
–, indirekte 288
Bandstruktur, Halbleiter 291
–, Metalle 240
Basis, Kristalle 8, 16
Basisvektor 9
Bauchbahn 272
BCS-Theorie 411
BCS-Grundzustand 416
BCS-Zustand, endliche Temperatur 419
Besetzung
–, Cooper-Paare 417
–, Zwei-Niveau-Zustnde 187, 211
Beugungsexperimente 31
Beweglichkeit, elektronische
–, Halbleiter 305
–, Metalle 251
Beweglichkeitskante, amorpher Halbleiter 308
Bewegungsgleichung, semiklassische 244
Bildungsenergie
–, Leerstellen 93
–, Zwischengitteratome 101
Bindender Zustand 82
Bindung 67
–, heteropolare 75
–, ionische 75
–, kovalente 78
–, metallische 87
–, van der Waals 71
Bindungsenergie 69
Bipolarer Transistor 327
Bloch
– -Funktion 222
– -Grüneisen-Gesetz 262
– -Oszillation 247
– -Theorem 222
– -Wand 359
Blochsches-$T^{3/2}$-Gesetz 359
Bohrscher Radius 103
Boltzmann-Gleichung 253
–, linearisierte 255
Born-Haber-Kreisprozeß 76
Born-Mayer-Potential 71
Born-Oppenheimer Näherung 79, 125
Bose-Einstein-Faktor 163
Bragg-Bedingung 46
Bravais-Gitter 12
Brechungsindex 378
Bridgman-Verfahren 6
Brillouin
– -Funktion 340
– -Streuung 151,154
– -Zone 41, 237
Burgers-Vektor 110
Cäsiumchloridstruktur 19
Casimir-Bereich 195
Chemisches Potential 204
Clausius-Mossotti-Beziehung 370
Cluster 14
Cooper-Paar 412
–, Ausdehnung 416
Coulomb-Potential, abgeschirmtes 213
Curie
– -Gesetz 339
– -Konstante 340

Curie-Temperatur
–, ferromagnetische 345
–, paramagnetische 346
Curie Weiss-Gesetz
–, magnetisches 346
–, elektrisches 390
Czochralski-Verfahren 5
Dauerstromversuch 403
De Haas van Alphen
– -Effekt 278
– -Oszillation 280
Debye
– -Frequenz 162
– -Geschwindigkeit 162
– -Gleichungen 388
– -Näherung 161
– -Scherrer-Verfahren 61
– -Temperatur 163
– -Waller-Faktor 47, 147
Defekte 91
–, ausgedehnte 106
–, in Halbleiter 313
Defektelektron 249
Defektstreuung
–, von Phononen 194
–, von Elektronen 257
Defektzustände
–, amorpher Festkrper 118
–, amorpher Halbleiter 308, 313
Deformation, plastische 106, 113
Depolarisationsfeld 369
Diamagnetismus 338
Diamagnet, idealer 406
Diamantstruktur 22
Dichteverteilungsfunktion, radiale 57
Dielektrische Eigenschaften 367
Dielektrische Funktion
–, Ionenpolarisation 371
–, elektronische 380
–, Orientierungspolarisation 383
–, freie Ladungstrger 395
Diffusion
–, von Leerstellen 95
–, von Zwischengitteratomen 105
Diffusionskonstante 97
Diffusionsspannung 318
Diffusionsstrom, p-n-Übergang 320
Diffusionsterm, Boltzmann-Gleichung 253
Direkter Halbleiter 288
Dispersionsrelation
–, Elektronen 244
–, Gitterschwingungen 131
–, Magnonen 356, 362
–, Schallwellen 127
–, Supraleiter 418
Domänen, ferromagnetische 359
Dominante Phononen 192
Donator 102,297
Dotieren 105
Dotierung 297
Dotierungsübergitter 333
Drehachse 10
Drehinversion 11
Drehkristallverfahren 61
Drehung 10
Drei-Phononen-Prozeß 180
Dulong-Petit-Gesetz 155, 164
Durchlaßrichtung, p-n-bergang 322
Dynamische Masse 245
Dynamische Matrix 130
Edelgaskristalle 73
Effektive Elektronenmasse 235, 245
–, dynamische 210
–, Halbleiter 291
–, thermische 210
Effektive Massennäherung 245
Effektive Zustandsdichte in Halbleitern 295
Eigenleitung
–, amorphe Halbleiter 309
–, kristalline Halbleiter 294
Ein-Phonon-Streuung 147
Einelektron-Näherung 201, 231
Einkristall 4
Einstein-Beziehung 97
Elastisch isotroper Festkrper 126
Elastische Streuung 33
Elastisches Kontinuum 126
Elastizitätsmodul 126
Elektrische Leitfähigkeit 260
–, amorpher Halbleiter 309
–, Drude 251
–, Ionenkristalle 97
–, Metalle 252, 260
Elektronen
– -bahn im Magnetfeld 272
– -beugung 64
– -gas 87, 201
–, effektive Masse 210
– -mikroskopie 32
–, im Magnetfeld 267
–, quasi-freie 223
–, stark gebundene 231
Elektron-Elektron-Streuung 259
Elektron-Phonon-Streuung 258
Elektronstreuung
–, an Defekten 257
–, an Phononen 258
Elektronische Bänder, amorpher Halbleiter 307
Elektronische Zustandsdichte 202
Elementarzelle 10

Elementarzelle, primitive 10
Emitter 327
Energieband 226
Energieerhaltung, bei Streuung 143
Energielücke
–, in der Elektronenenergie 226
–, in Supraleitern 418
–, in Supraleitern, Nachweis 420
Entarteter Halbleiter 294
Entelektrisierungfeld 369
Entropie
–, Leerstellen 93
–, Supraleiter 410
Erschöpfungszustand 303
Erzeugungsstrom, p-n-Übergang 321
Evjen-Zelle 77
Ewald-Kugel 45
Expansionskoeffizient 175
–, amorpher Festkörper 178
–, von Kristallen 175
Extinktionskoeffizient 378
Extremalbahnen 272
Exziton 393
Fabry-Perot-Interferometer 154
Farbzentrum 98
Feldeffekttransistor 328
Feldstärke, kritische 409
Feldstrom, p-n-Übergang 321
Feldterm, Boltzmann-Gleichung 253
Fermi
– -Energie 204
– -Fläche 207, 237, 267
– -Funktion 204
– -Geschwindigkeit 205
– -Impuls 205
– -Temperatur 205
– -Wellenvektor 205
Fermi-Kugel 206
–, Verschiebung 252
Fermi-Niveau
–, dotierter Halbleiter 299
–, amorpher Halbleiter 315
Fernordnung 3
Ferrimagnetismus 360
Ferroelektrizität 389
Ferromagnetismus 344
–, Stoner und Slater 352
–, Heisenberg-Modell 350
Festkörper, amorpher 1, 6
–, kristalliner 1
Ficksches Gesetz 97
Fließen 106
Fluáquant 426
Fluáquantisierung 425
Fluáschlauch 435
Fluxoid 426
Frank-Read-Quelle 114
Freie Weglänge 128
–, von Phononen 182
–, von Elektronen 256
Freies Elektronengas 201
Fremdatom 92, 102, 105
Frenkel-Defekt 101
Frenkel-Defekt 101
Frenkel-Exzitonen 394
Frequenzlücke 140
Friedelsche Regel 47
Frustration 364
Fulleren 86
Fundamentalabsorption 381
Gamma-Punkt 42
Ginzburg Landau-Theorie 411
Gitter
– -anharmonizitt 175
– -dynamik 125
– -ebenenabstand 40
– -ebene 38
– -energie 69
– -konstante 9
–, leeres 224
– -relaxation 94
–, reziprokes 36
Gitterdynamik 125
–, einatomige Basis 132
–, mehratomige Basis 137
Glühemission 216
Glanzwinkel 46
Glasübergang 7
Glasherstellung 6
Gleichrichtung 321
Gleichstromleitfähigkeit 255
Gleitspiegelebene 18
Grenzflächenenergie 433
Grüneisen
– -Beziehung 177
– -Parameter 176
Halbleiter 287
–, amorpher 307
–, direkter 288
–, entarteter 294
–, indirekter 288
–, inhomogener 317
–, intrinsische 287
–, kristalline 287
Halbmetalle 240
Hall-Effekt 282
–, Halbleiter 304
–, klassischer 282
Hall-Konstante 284
Hall-Widerstand 283

Halsbahn 272
Heisenberg-Modell, Ferromagnetismus 350
Heteroübergang, isotyper 331
Heteropolare Bindung 75
Heterostruktur 330
Hexagonal dichteste Kugelpackung 23
Hochtemperatursupraleiter 436
Horizontale Stáe 266
Hüpfleitfähigkeit 310
Hybridbindung 85
H_2^+-Moleklion 79
Impulserhaltung bei Streuung 143
Indirekte Bandlücke 288
Inelastische
– Ein-Phonon-Streuung 147
– Mehr-Phononen-Streuung 147
– Neutronenstreuung 145
– Streuung 33, 142
Infrarotabsorption 104,153
Infrarotaktivitt 139
Inhomogene Halbleiter 317
Injektionslaser 334
Inkohärente Streuung 33
Innere Energie
–, Gitterschwingungen 155
–, Fermi-Gas 207
Interbandübergang 245, 381
–, in Halbleitern 289
Internationale Notation 17
Interstitielle Legierung 120
Intrabandübergang 381
Intrinsische
– Halbleiter 287
– Leitfähigkeit 294
Inversion 11
Inversionsschicht 329
Ionenbindung 75
Ionenpolarisation 371
Ionisationsenergie, Strstellen 102, 298
Ising-Modell 364
Isolator 67
Isotopeneffekt, Supraleitung 412
Isotopenstreuung, Phononen 196
Josephson-Effekt 428
– -Gleichstrom-Effekt 428
– -Wechselstrom-Effekt 429
Kleinwinkelkorngrenze 117
Kleinwinkelstreuung 55
Kohärente Streuung 33
Kohärenzlänge, Supraleiter 433
Kollektor 327
Kompensation 299
Kompensationübergitter 332
Kompressibilität 176
Kondensationsenergie 410, 418
Konfigurationsentropie 93
Koordinationszahl 19
Korngrenze 4, 106, 116
Korrelationsenergie, bei Strstellen 299, 313
Kovalente Bindung 78
Kraftkonstantentensor 129
Kramers-Kronig-Beziehung 378
Kreisprozeß 76
Kristall
–, kubischer 19
– -ebene 38
– -feld 341
– -gitter 8
– -impuls 143
– -klasse 17
– -richtung 39
– -struktur 8
– -system 11
– -zucht 5
Kritische Stromdichte 424
Kritischer Punkt 159
Kritisches Magnetfeld 408, 424, 434
Kubischer Kristall 19
Kubisch flächenzentriertes Gitter 21
Kubisch primitives Gitter 19
Kubisch raumzentriertes Gitter 20
Ladungskompensation
–, in Halbleitern 299
–, in Ionenkristallen 98
Ladungsträgerdichte
–, in dotierten Halbleitern 299
–, in intrinsischen Halbleitern 293
Ladungstransport
–, in Bändern 248
–, in Ionenkristallen 97
–, Beschreibung mit Boltzmann-Gleichung 255
Landau
– -Diamagnetismus 280, 339
– -Niveaus 274
– -Röhren 276
– -Rumer-Dämpfung 185
Landé-Faktor 339
Langevin-Debye-Gleichung 383
Langwelliger Grenzfall 135
Laue-Verfahren 63
LEED 65
Leeres Gitter 224
Leerstelle 91
–, Bildungsenergie 93
– -diffusion 95
– -entropie 93
– -konzentration 92
Legierung 120
Leichtes Loch, Halbleiter 293
Leiter, idealer 405

Leitfähigkeit
–, von dotierten Halbleitern 307
–, von Halbleitern 293
–, intrinsische 294
–, von Ionenkristallen 97
–, von Metallen 260
Leitungsband 235, 240
–, von Halbleitern 287
Lennard-Jones-Potential 70, 72
Leuchtdiode 326
Lichtstreuung 150
Lochbahn im Magnetfeld 272
Löcher 249
Löcherkonzentration 295
Lokales Feld 368
Lokalisierte Zustände amorpher Halbleiter 308
London-Gleichung 407
Londonsche Eindringtiefe 407
Longitudinale
– Elektronenmasse, Halbleiter 293
– Schallwelle 126
– Schwingung im Elektronengas 399
– Schwingung in Ionenkristallen 164
Lorentz
– -Beziehung 370
– -Näherung 368
Lorenz-Zahl 263
Lyddane-Sachs-Teller-Beziehung 375
Madelung-Konstante 77
Magnetfeld, kritisches 408, 424
Magnetische Suszeptibilität 211
Magnetisierung 337
–, paramagnetische 340
–, spontane 348
Magnetismus 337
Magnetonenzahl, effektive 340
Magnon 355, 362
Majoritätsladungsträger 318
Makroskopische Wellenfunktion 425
Masse, effektive 235
Matthiesensche Regel 261
Mehr-Phononen-Streuprozeß 147
Meißner-Effekt 405
Meißner-Phase 434
Metall-Isolator-Übergang 215
Metall 67
Metallbindung 87
Millersche Indizes 38
Minoritätsladungsträger 318
Molekularfeld 345
– -näherung 345
Monochromatorkristall 59
Mosaikblock 117
MOSFET 328
Mott-Wannier-Exzitonen 394
n-Halbleiter 300
Nachbar, nächster 19
Näherung der Nichtentartung 294
Nahordnung 3
Natriumchloridstruktur 21
Néel-Temperatur 361
Netzebene 38
Neutronenspektrometer 146
Neutronenstreuung, inelastische 145
Normalprozeá
–, Elektronen 258
–, Phononen 193
Notation, Symmetrieelemente 17
Nullpunktsenergie, Gitter 167
Oberflächenplasmonen 401
Oberflächenstreuung, Phononen 194
Optische Absorption
–, in amorphen Halbleitern 308
–, durch Exzitonen 394
–, in Halbleitern 289
Optische Eigenschaften 367, 378
–, freie Ladungstrger 395
Optische Phononen, Ionenkristall 374
Optischer Zweig 138
Ordnung 1
Ordnungsparameter 122, 390
Ordnungs-Unordnungs-Übergang 122, 391
Orientierungspolarisation 383
Oszillatormodell
–, bei der Ionenpolarisation 371
–, bei optischen bergngen 380
p-n-Übergang 317
p-n-Diode 323
Paarkorrelationsfunktion 25
Packungsverhältnis 19
Paraelektrizität 390
Paramagnetismus 339
Patterson-Theorem 56
Pauli-Paramagnetismus 212
Penrose-Muster 15
Periodische Randbedingungen 202
Periodisches Potential
–, Elektronen 219
–, Schrödinger-Gleichung 221
Periodisches Zonenschema fr Elektronen 267
Perowskitstruktur 20, 392, 440
Phase bei Gitterschwingungen 134, 138
Phasenproblem bei Streuexperimenten 47
Phasenübergang 122
–, ferroelektrischer 390
–, ferromagnetischer 347
–, supraleitender 410
Phonon-Dispersionskurven
–, von Kupfer 148
–, von Lithiumfluorid 150

Phononen 143
– -amplitude 145
– -austausch im Supraleiter 413
–, ballistische 197
– -zahl 167
– -zerfall, spontaner 185
– -zustandsdichte 156
Phonon-Polariton 375
Phonon-Streuung
–, an Defekten 194
–, an Elektronen in Metallen 258
–, an Elektronen in Halbleitern 306
–, an der Oberfläche 194
–, an Phononen 182
Photodiode 326
Photoemissionsspektroskopie 242
Piezoelektrizität 389
Plasmafrequenz 396
Plasmon 399
Plasmon-Polaritonen 396
Plastische Deformation 106, 113
Polarisation
–, elektrische 370
–, gemischte (bei Phononen) 132
–, von Tunnelsystemen 385
Polarisationskatastrophe 392
Polarisierbarkeit
–, atomare 367
–, elektronische 380
Polariton 375
Polykristall 4, 115
Potentialfunktion 128
Proximity-Effekt 433
Pulververfahren 61
Punktdefekt 91
– -streuung 182
Punktgitter 9
Punktgruppe 17
Punktsymmetrieoperation 10
Pyroelektrizität 389
Quanten-Hall-Effekt 284
Quanten-Hall-Effekt, fraktionaler 286
Quasi-Fermi-Niveau 322, 335
Quasi-freie Elektronen 223
Quasiimpuls
–, von Elektronen 245
–, von Phononen 143
Quasikristall 14
Quasiteilchen 144, 418
Radiale Dichteverteilungsfunktion 57
Raman-Streuung 151
–, 2. Ordnung 153
–, in amorphen Festkörpern 154
Randbedingungen
–, feste 156
–, periodische 156, 202
Rastertunnelmikroskopie 32, 218
Raumgruppe 18
Raumladungszone 319
Rayleigh-Streuung
–, von Licht 151
–, von Phononen 182
Reduktion, 1. Brillouin-Zone 135, 223
Reduziertes Zonenschema 223
Reflektivität 378
Reflektivität, Metalle 398
Reflexbreite 44
Rekombinationsstrom, p-n-Übergang 321
Rekombinationszentren 326
Relaxation, dielektrische 386
Relaxationsabsorption 189
Relaxationszeit
– -ansatz 254
–, dielektrische 387
–, von Elektronen 251
Resonante Wechselwirkung 187, 199
Reststrahlmethode 380
Restwiderstand 121, 261
Reziprokes Gitter 36
Richardson-Dushman-Formel 217
RKKY-Wechselwirkung 351
Schönflies-Notation 17
Schallausbreitung 125
Schallwellen in Gläsern 168
Schottky-Defekt 92
Schottky-Modell der Raumladungszone 319
Schraubenachse 18
Schraubenversetzung 110, 112
Schubspannung, kritische 108
Schweres Loch, Halbleiter 293
Schwingungsquant 143
Schwingungsspektrum
–, amorpher Festkörper 168
–, kristalliner Festkörper 168
Semiklassische Bewegungsgleichung 244
Shubnikov-de Haas-Oszillation 280
Shubnikov-Phase 434
Singulett-Paar, Supraleiter 416
Solarzelle 324
Sommerfeld-Näherung 201, 252
Sperrichtung, p-n-bergang 322
Spezifische Wärme
–, von Metallen 207
–, von Glsern 107
–, nach Einstein 156
–, nach Debye 161
–, im Magnetfeld 280
–, von niederdimensionalen Systemen 165
–, von Supraleiter 411, 421
–, von bergangsmetalle 207

Spiegelung 11
Spinglas 363
Spinglastemperatur 363
Spins, lokalisierte 348
Spinwellen 355
Spontane Magnetisierung 337, 345
Spontane Polarisation 389
Sprungtemperatur, Supraleiter 403
SQUID 432
Stabilitätskriterium 74
Stark gebundene Elektronen 231
Störstellen
– -band 303
– -leitung 301
– -reserve 302
Stoßterm, Boltzmann-Gleichung 253
Stoßzeit, von Elektronen in Metallen 251
Stoßzeit, von Ladungstrgern in Halbleitern 306
Stoner-Parameter 353
Streu
– -amplitude 35
– -bedingung 43
– -dichte 34
– -experimente 59, 145
– -faktor, atomarer 52
– -länge 52
– -querschnitt 181
– -term, Boltzmann-Gleichung 253
– -theorie 33
– -vektor 35
Streuung 54
–, an amorphen Substanzen 54
–, am dynamischen Gitter 141
–, an Kristallen 43
–, experimentelle Methoden 59
–, Strukturbestimmung 33
Stromdichte, kritische 424
Struktur, amorphe Festkörper 54
Strukturamplitude 49
Strukturbestimmung 31
Strukturfaktor
–, dynamischer 141
–, statischer 48
Strukturrelaxation 101
Stufenversetzung 109
Subband, magnetisches 355
Substitutionslegierung 120
Superaustausch 351
Supraleitung 403
–, mikroskopische Beschreibung 411
Supraleiter 1. Art 434
Supraleiter 2. Art 432
Suszeptibilität
–, antiferromagnetische 361
–, elektrische 370
–, von Kernen 343
–, paramagnetische 211, 340
–, von Spinglas 364
Symmetrieoperation 10
Teilchenzahldichte 25
Temperaturmessung 342
Tempern 7, 121
Tetraedrische Bindung 85
Thermische Expansion 175
–, amorpher Festkrper 179
–, bei tiefen Temperaturen 177
Thermische Leitfähigkeit
–, von dielektrischen Kristallen 190
–, von Metallen 263
–, von amorpher Festkrpern 198
Thomas-Fermi-Abschirmlänge 214
Transistor 327
Translationssymmetrie 9
Translationsvektor, fundamentaler 9
Transportgleichung 253
Transversale Elektronenmasse, Halbleiter 293
Transversale Schallwellen 126
Triplett-Paar, Supraleiter 416
Tunnel
– -kontaktspektroskopie 421
– -modell, amorpher Festkrper 171
– -strom, aus Metallen 218
– -strom, in Supraleitern 218
– -System, amorpher Festkrper 422, 428
T^3-Gesetz, spezifische Wärme 164
Übergangsmetalle 89
Übergitter 31, 332
Ultraschalldämpfung 127
–, amorphe Festkörper 187
–, Kristallen 181
Ultraschallsättigung 188
Umklapp-Prozeá
–, Elektronen 259
–, Phononen 194
Ungesättigte Bindung 118
Unordnung 1
Valenz-alternierendes Paar 120
Valenzband 235, 240
Valenzband von Halbleitern 287
Van der Waals-Kräfte 72
Van-Hove-Singularität 159
Variable range hopping 312
Variationsverfahren 80
Verarmungzone 329
Verschiebungsübergang 391
Versetzung 106
Versetzungsdichte 111
Versetzungslinie 109
Vertikaler Prozeá 266
Verzerrungstensor 183

Volumenplasmonen 401
Voronoi-Polyeder 28
Wärmeleitfähigkeit
–, amorpher Festkörper 198
–, dielektrischer Kristalle 190
–, Messung 190
–, von Metallen 263, 266
Wärmepulsexperiment 197
Wasser, Struktur 90
Wasserstoffatommodell
–, bei Fremdatomen 103
–, bei Exzitonen 394
Wasserstoffbrückenbindung 89
Wasserstoffmolekül 79
Wechselstrom-Josephson-Effekt 429
Wechselstromwiderstand im Supraleiter 425
Wechselwirkung, resonante 187
Wechselwirkungspotential 70
Wellenfunktion, makroskopische 425
Wellengleichung 126
Wellenpaket 243
Wiedemann-Franz-Gesetz 263
Wigner-Seitz-Zelle 24
Wirkungsgrad, Solarzelle 325
YBCO 436
Zähligkeit 10
Zener-Durchbruch 324
Zinkblendestruktur 22
Zonenreinigung 5
Zonenschema, reduziertes 223
Zonenschmelzen 6
Zustandsdichte 158
–, Debyesche 162
–, Elektronen im Magnetfeld 278
–, an der Fermi-Kante 205
–, freier Elektronen 202
–, im Halbleiter 294
–, im Impulsraum 203
–, von Magnonen 358
–, niederdimensionaler Systeme 165
–, von Quasiteilchen in Supraleitern 419
–, im reziproker Raum 158
Zwei-Komponenten-Näherung 227
Zweiteilchen-Wellenfunktion 83
Zwischengitteratom 91, 101, 105
Zyklotron
– -masse 270
– -resonanz, in Halbleitern 291
– -resonanz, in Metallen 269