

# 1

## Physik und Eigenschaften von Halbleitern – ein Überblick

### 1.1 Einleitung

Die Physik der Halbleiterbauelemente und Halbleitermaterialien hängt untrennbar miteinander zusammen. Das erste Kapitel dieses Buches bietet daher zunächst eine kompakte Zusammenfassung über die grundlegende Physik von Halbleitermaterialien und einen Überblick über die speziellen Eigenschaften von Halbleiterbauelementen. Wegen der sehr umfangreichen fachspezifischen Literatur werden aber nur Themen behandelt, die für den Betrieb von elektronischen Bauelementen wirklich relevant sind. Ausführliche Abhandlungen verschiedenster Themen der Halbleiterphysik findet der interessierte Leser dann in den Standardlehrbüchern und Nachschlagewerken von Moll [1], Smith [2], Moss [3], Böer [4], Seeger [5], Singh [6], Taur [7], Sze [8] und Hamaguchi [9], um nur einige zu nennen.

Um die große Menge an relevanten Informationen zusammenzufassen, wurden für dieses Kapitel über 30 Tabellen (einige im Anhang) und Abbildungen aus verschiedenen Originalquellen zusammengestellt. Die Eigenschaften der beiden wichtigsten Halbleiter, Silizium (Si) und Galliumarsenid (GaAs) werden besonders detailliert diskutiert. Silizium wird in praktisch allen kommerziellen Elektronikprodukten verwendet, Galliumarsenid hingegen ist hauptsächlich in der Glasfaserkommunikation und bei Höchstfrequenzanwendungen zu finden.

### 1.2 Kristallstrukturen

#### 1.2.1 Primitive Einheitszellen und Kristallebenen

Ein Kristall ist dadurch gekennzeichnet, dass er aus einer gut strukturierten periodischen Anordnung von Atomen im Raum besteht. Die kleinste Gruppierung von Atomen, die zum Aufbau des gesamten Kristalls periodisch im Raum angeordnet werden kann, wird als primitive Einheitszelle mit der Gitterkonstante  $a$  bezeichnet. Abbildung 1.1 zeigt die typischsten dieser primitiven Zellen.

Viele wichtige Halbleiter haben eine Diamant- oder Zinkblende-Gitterstruktur. Diese gehört zu den tetraedrischen Phasen, d. h., jedes Atom ist von vier äquidistanten Nachbarn umgeben, welche an den Ecken eines Tetraeders liegen. Die Bindung zwischen zwei nächsten Nachbarn wird durch zwei Elektronen mit entgegengesetztem Spin gebildet. Das Diamant- und das Zinkblende-Gitter können als zwei sich gegen-

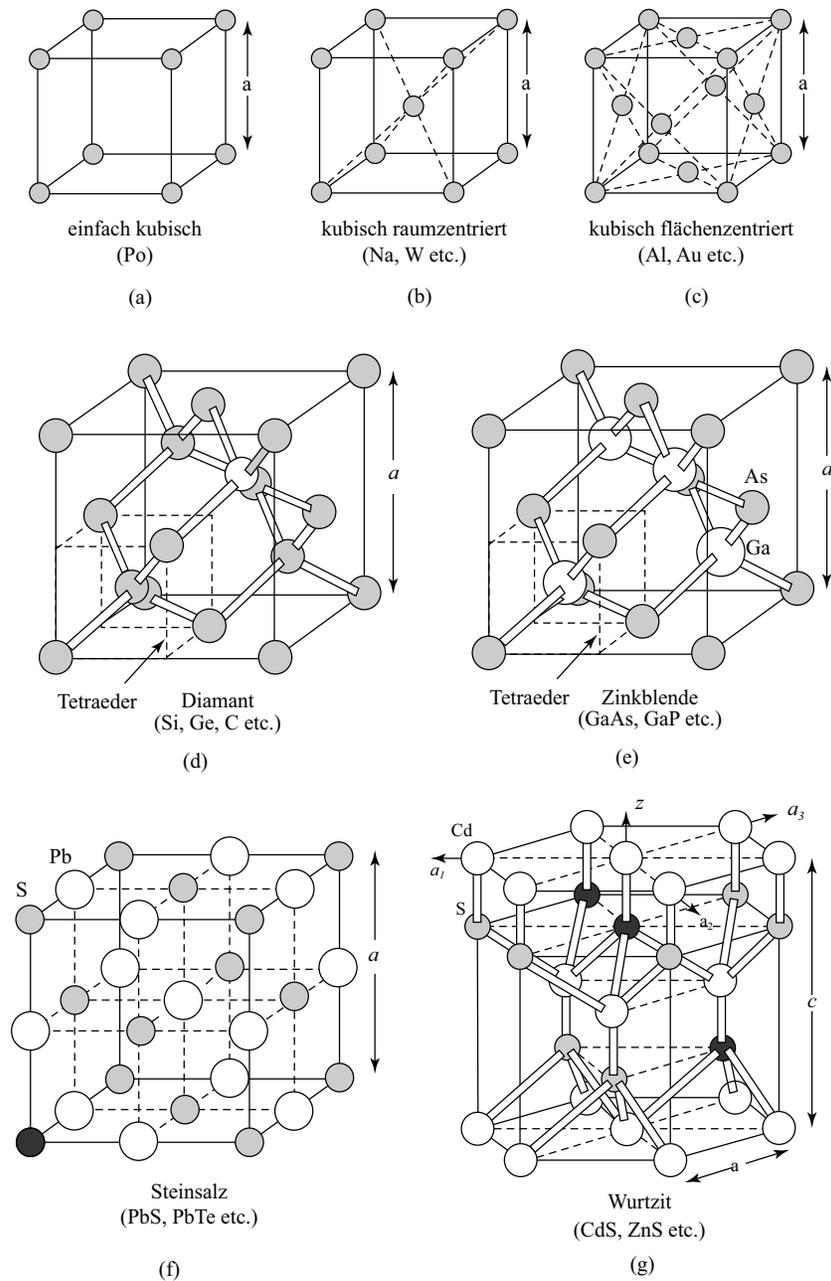


Abb. 1.1 Primitive Einheitszellen der wichtigsten Kristallgitter;  $a$  ist die Gitterkonstante.

seitig durchdringende kubisch-flächenzentrierte (engl. face-centered cubic, fcc) Gitter betrachtet werden. Im Diamant-Gitter, wie z. B. Silizium (Abb. 1.1d), sind alle Atome gleich, während in einem Zinkblende-Gitter, wie z. B. Galliumarsenid (Abb. 1.1e), eines der Gitter aus Gallium- und das andere aus Arsenatomen besteht. Galliumarsenid ist eine sogenannte III-V-Verbindung, da es aus Elementen der Gruppen III und V des Periodensystems gebildet wird.

Die meisten III-V-Verbindungen kristallisieren in der Zinkblende-Struktur [6, 9, 10], jedoch kristallisieren viele Halbleiter, einschließlich einiger III-V-Verbindungen, auch in Steinsalz- oder Wurtzit-Strukturen. Abbildung 1.1f zeigt das Steinsalz-Gitter, das wiederum als zwei sich gegenseitig durchdringende, kubisch flächenzentrierte Gitter betrachtet werden kann. In dieser Steinsalz-Struktur hat jedes Atom sechs nächste

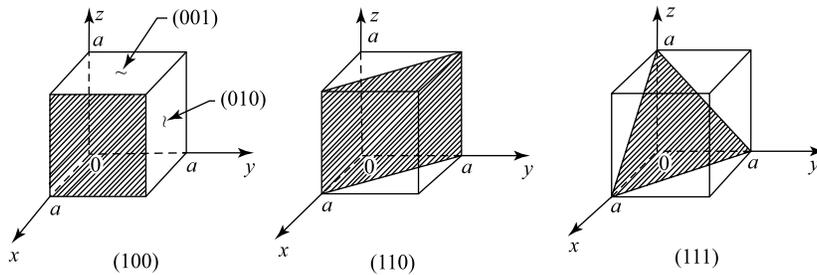


Abb. 1.2 Miller-Indizes einiger wichtiger Ebenen in einem kubischen Kristall.

Nachbarn. Abbildung 1.1g zeigt das Wurtzit-Gitter, das als Kombination von zwei, sich gegenseitig durchdringenden, hexagonal dicht gepackten Gittern (z. B. die Untergitter von Cadmium und Schwefel) betrachtet werden kann. In diesem Bild sind die beiden Ebenen benachbarter Schichten für jedes Untergitter (Cd oder S) horizontal so verschoben, dass der Abstand zwischen diesen beiden Ebenen für einen festen Abstand zwischen den Zentren zweier Atome minimal ist, daher auch der Name „dicht gepacktes Gitter“. Die Wurtzit-Struktur hat eine tetraedrische Anordnung von vier gleich weit voneinander entfernten Nachbarn, ganz ähnlich wie bei der Zinkblende-Struktur.

Anhang F enthält eine Zusammenfassung der Gitterkonstanten und Kristallstrukturen wichtiger Halbleiter [11, 12]. Beachten Sie, dass einige Verbindungen, wie Zink- und Cadmiumsulfid, sowohl in Zinkblende- als auch in Wurtzit-Strukturen kristallisieren können.

Da Halbleiterbauelemente immer auf oder nahe einer Halbleiteroberfläche hergestellt werden, sind die Orientierungen und Eigenschaften der Kristallebenen an der Oberfläche sehr wichtig. Eine bequeme Methode zur Definition der verschiedenen Ebenen in einem Kristall ist die Verwendung von Miller-Indizes. Die Miller-Indizes einer Kristallebene werden bestimmt, indem zunächst die Schnittpunkte der Ebene mit den drei Basisachsen in Form der Gitterkonstanten (oder primitiven Zellen) ermittelt werden und dann die Kehrwerte dieser Zahlen auf die kleinsten drei ganzen Zahlen mit gleichem Verhältnis reduziert werden. Das Ergebnis wird in (runde) Klammern ( $hkl$ ) gesetzt, und als Miller-Indizes für eine einzelne Ebene oder eine Reihe paralleler Ebenen  $\{hkl\}$  bezeichnet.<sup>1)</sup> Abbildung 1.2 zeigt die Miller-Indizes wichtiger Ebenen in einem kubischen Kristall. Einige andere Konventionen sind in Tab. 1.1 aufgeführt. Bei Silizium, einem Ein-Element-Halbleiter, sind die  $\{111\}$ -Ebenen die leichtesten Bruch- oder Spaltungsebenen. Im Gegensatz dazu spaltet sich Galliumarsenid, welches eine ähnliche Gitterstruktur hat, aber auch eine leichte ionische Komponente in den Bindungen aufweist, am besten auf den  $\{110\}$ -Ebenen.

Jede primitive Einheitszelle wird im Raum durch drei primitive Basisvektoren,  $\mathbf{a}$ ,  $\mathbf{b}$  und  $\mathbf{c}$  aufgespannt. Mit diesen Vektoren kann die Kristallstruktur eines Festkörpers so charakterisiert werden, dass die Kristallstruktur bei einer Translation durch jeden Vektor, der aus der Summe von ganzzahligen Vielfachen dieser Basisvektoren besteht, invariant bleibt. Mit anderen Worten, die Gitterplätze des Kristalls im Ortsraum können durch die Menge der Vektoren

$$\mathbf{R} = m\mathbf{a} + n\mathbf{b} + p\mathbf{c} \quad (1.1)$$

definiert werden [13], wobei  $m$ ,  $n$  und  $p$  ganze Zahlen sind. Anmerkung: Das Kristallgitter im Ortsraum wird auch direktes Gitter genannt.

<sup>1)</sup> Eckige Klammern bezeichnen normalerweise die Kristallrichtung, also den Vektor senkrecht auf diese Ebenen.

Tab. 1.1 Miller-Indizes von Kristalloberflächen und die zugehörigen Kristallebenen.

Miller-Indizes	Kristallebene oder Kristallrichtung
$(hkl)$	Ebene, die $1/h$ , $1/k$ , $1/l$ auf der $x$ -, $y$ - und $z$ -Achse schneidet
$(\bar{h}kl)$	Ebene, die die negative $x$ -Achse schneidet
$\{hkl\}$	Vollständiger Satz von Ebenen mit äquivalenter Symmetrie, wie z. B. $\{100\}$ für $(100)$ , $(010)$ , $(001)$ , $(\bar{1}00)$ , $(0\bar{1}0)$ und $(00\bar{1})$ in kubischer Symmetrie
$[hkl]$	Richtung eines Kristalls wie $[100]$ für die $x$ -Richtung
$\langle hkl \rangle$	Vollständiger Satz gleichwertiger Richtungen
$[hklm]$	Ebene in einem hexagonalen Gitter (wie z. B. Wurtzit), die $1/h$ , $1/k$ , $1/l$ , $1/m$ auf der $a_1$ -, $a_2$ -, $a_3$ - und $z$ -Achse schneidet (Abb. 1.1g)

### 1.2.2 Das reziproke Gitter

Für einen gegebenen Satz direkter Basisvektoren gibt es auch einen Satz reziproker Basisvektoren  $\mathbf{a}^*$ ,  $\mathbf{b}^*$ ,  $\mathbf{c}^*$  im Gitter, mit

$$\mathbf{a}^* \equiv 2\pi \frac{\mathbf{b} \times \mathbf{c}}{\mathbf{a} \cdot \mathbf{b} \times \mathbf{c}}, \quad (1.2a)$$

$$\mathbf{b}^* \equiv 2\pi \frac{\mathbf{c} \times \mathbf{a}}{\mathbf{a} \cdot \mathbf{b} \times \mathbf{c}} \quad \text{und} \quad (1.2b)$$

$$\mathbf{c}^* \equiv 2\pi \frac{\mathbf{a} \times \mathbf{b}}{\mathbf{a} \cdot \mathbf{b} \times \mathbf{c}}, \quad (1.2c)$$

sodass die Beziehungen  $\mathbf{a} \cdot \mathbf{a}^* = 2\pi$ ,  $\mathbf{a} \cdot \mathbf{b}^* = 0$  usw. gelten. Die Nenner bilden das von diesen Vektoren eingeschlossene Volumen und sind wegen der Beziehung  $\mathbf{a} \cdot \mathbf{b} \times \mathbf{c} = \mathbf{b} \cdot \mathbf{c} \times \mathbf{a} = \mathbf{c} \cdot \mathbf{a} \times \mathbf{b}$  identisch. Ein allgemeiner reziproker Gittervektor ist gegeben durch

$$\mathbf{G} = h\mathbf{a}^* + k\mathbf{b}^* + l\mathbf{c}^*, \quad (1.3)$$

wobei  $h$ ,  $k$  und  $l$  ganze Zahlen sind. Daraus folgt die wichtige Beziehung zwischen dem direkten Gitter und dem reziproken Gitter:

$$\mathbf{G} \cdot \mathbf{R} = 2\pi \cdot n, \quad n: \text{ganzzahlig}. \quad (1.4)$$

Mit dieser Beziehung steht jeder Vektor des reziproken Gitters senkrecht auf einer Menge von Ebenen im direkten Gitter. Das Volumen  $V_c^*$  einer primitiven Zelle des reziproken Gitters ist umgekehrt proportional zum Wert von  $V_c$  im direkten Gitter, d. h.,  $V_c^* = (2\pi)^3 / V_c$ , wobei  $V_c \equiv \mathbf{a} \cdot \mathbf{b} \times \mathbf{c}$  ist (Anmerkung: Daher auch der Name „reziprokes Gitter“).

Die primitive Zelle eines reziproken Gitters kann durch eine Wigner-Seitz-Zelle dargestellt werden. Die Wigner-Seitz-Zelle wird durch das Zeichnen von senkrechten Halbiebungsebenen im reziproken Gitter vom gewählten Zentrum zu den nächstgelegenen äquivalenten reziproken Gitterstellen konstruiert. Diese Technik kann auch auf ein direktes Gitter angewendet werden. Die Wigner-Seitz-Zelle im reziproken Gitter wird im Allgemeinen als erste Brillouin-Zone bezeichnet. Abbildung 1.3a zeigt ein typisches Beispiel für ein kubisch-raumzentriertes (engl. body-centered cubic, bcc) reziprokes Gitter [14]. Zieht man zunächst Linien vom Mittelpunkt ( $\Gamma$ ) zu den acht Ecken des Würfels und bildet dann die Bisektorebenen, so ergibt sich ein abgestumpfter Oktaeder innerhalb des Würfels – eine Wigner-Seitz-Zelle. Nach ein wenig Vektorrechnung zeigt sich [15], dass ein flächenzentriertes direktes kubisches (face-centered cubic, fcc) Gitter mit der Gitterkonstante  $a$  ein zugehöriges reziprokes bcc-Gitter mit der

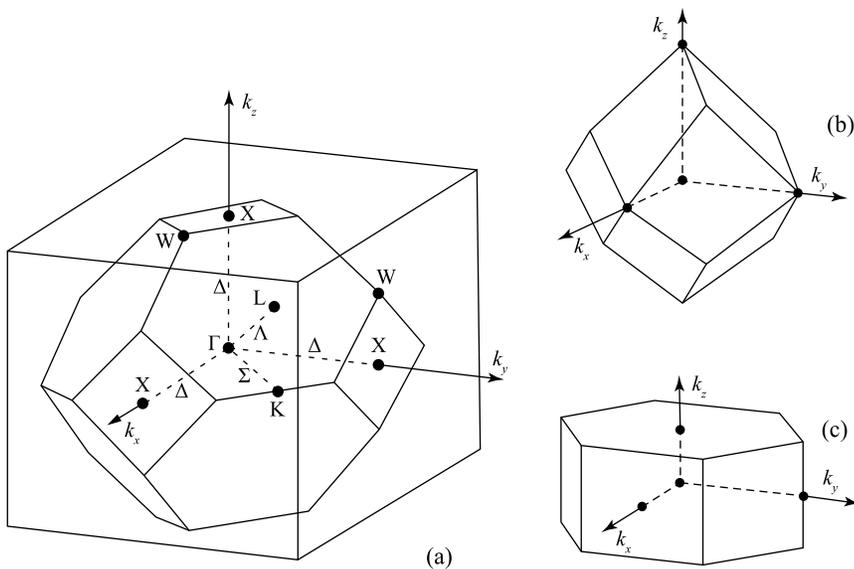


Abb. 1.3 Brillouin-Zonen für (a) fcc-, Diamant- und Zinkblende-Gitter, (b) bcc-Gitter und (c) Wurtzit-Gitter.

Gitterkonstante  $4\pi/a$  hat. Somit ist die in Abb. 1.3a gezeigte Wigner-Seitz-Zelle die primitive Zelle des reziproken (bcc-) Gitters für ein direktes fcc-Gitter. Die Wigner-Seitz-Zellen für direkte bcc- und hexagonale Gitter können auf ähnliche Weise konstruiert werden und sind in Abb. 1.3b,c dargestellt [16]. Man wird später sehen, dass das reziproke Gitter zur Darstellung der  $E$ - $k$ -Beziehung für Elektronen im Kristall dann sehr nützlich ist, wenn die Koordinaten der Wellenvektoren  $\mathbf{k}$  ( $|\mathbf{k}| = k = \lambda/\lambda$ ) auf die Koordinaten des reziproken Gitters abgebildet werden. Besonders wichtig ist hierbei die Brillouin-Zone des fcc-Gitters, weil die meisten Halbleitermaterialien, die hier von Interesse sind, ein fcc-Gitter haben. Die in Abb. 1.3a verwendeten Symbole werden später ausführlicher diskutiert.

### 1.3 Energiebänder und Bandlücken

In jedem Kristallgitter folgen die Ladungsträger ihrer eigenen und speziellen Energie-Impuls- $(E$ - $k)$ -Beziehung. Diese  $E$ - $k$ -Beziehung spielt an vielen Stellen eine entscheidende Rolle, wie z. B. bei den Wechselwirkungen der Ladungsträger mit Photonen und Phononen und auch bei Wechselwirkungen der Ladungsträger (Elektronen und Löcher) untereinander. Die  $E$ - $k$ -Beziehung bestimmt die Gruppengeschwindigkeit, die effektive Masse, und auch das Konzept der Bandlücken wird über die  $E$ - $k$ -Beziehung eingeführt.

Die Energie-Impuls- $(E$ - $k)$ -Beziehung eines kristallinen Festkörpers wird im gängigen Halbleiterjargon als Bandstruktur bezeichnet und normalerweise durch die Lösung der Schrödinger-Gleichung eines Ein-Elektronen-Problems erhalten. Das Bloch-Theorem, eines der wichtigsten Theoreme und Grundlage jeder Bandstruktur, besagt, dass wenn eine potentielle Energie  $V(\mathbf{r})$  im direkten Gitterraum (Ortsraum) periodisch ist, die Lösungen der Schrödinger-Gleichung [9, 15, 17]

$$\left[ -\frac{\hbar^2}{2m^*} \nabla^2 + V(\mathbf{r}) \right] \psi(\mathbf{r}, \mathbf{k}) = E(\mathbf{k}) \psi(\mathbf{r}, \mathbf{k}), \quad (1.5)$$

Tab. 1.2 Brillouin-Zone des fcc-, Diamant- und Zinkblenden-Gitters: Zonenränder und ihre entsprechenden Achsen ( $\Gamma$  ist das Zentrum).

Punkt	Entartung	Achse
$\Gamma, (0,0,0)$	1	
X, $2\pi/a(\pm 1, 0, 0), 2\pi/a(0, \pm 1, 0), 2\pi/a(0, 0, \pm 1)$	6	$\Delta, \langle 1, 0, 0 \rangle$
L, $2\pi/a(\pm 1/2, \pm 1/2, \pm 1/2)$	8	$\Lambda, \langle 1, 1, 0 \rangle$
K, $2\pi/a(\pm 3/4, \pm 3/4, 0), 2\pi/a(0, \pm 3/4, \pm 3/4), 2\pi/a(\pm 3/4, 0, \pm 3/4)$	12	$\Sigma, \langle 1, 1, 0 \rangle$

also die Wellenfunktion  $\psi(\mathbf{r}, \mathbf{k})$ , die Form einer Bloch-Funktion haben, welche so aussieht:

$$\psi(\mathbf{r}, \mathbf{k}) = \exp(j\mathbf{k} \cdot \mathbf{r})U_b(\mathbf{r}, \mathbf{k}). \quad (1.6)$$

Hier ist  $b$  der Bandindex, und  $\psi(\mathbf{r}, \mathbf{k})$  und  $U_b(\mathbf{r}, \mathbf{k})$  sind periodische Funktionen im direkten Gitter mit der Periode  $\mathbf{R}$ . Weil

$$\begin{aligned} \psi(\mathbf{r} + \mathbf{R}, \mathbf{k}) &= \exp[j\mathbf{k} \cdot (\mathbf{r} + \mathbf{R})]U_b(\mathbf{r} + \mathbf{R}, \mathbf{k}) \\ &= \exp(j\mathbf{k} \cdot \mathbf{r}) \exp(j\mathbf{k} \cdot \mathbf{R})U_b(\mathbf{r}, \mathbf{k}) \end{aligned} \quad (1.7)$$

gleich  $\psi(\mathbf{r}, \mathbf{k})$  ist, muss  $\mathbf{k} \cdot \mathbf{R}$  ein Vielfaches von  $2\pi$  sein. Das sieht man am besten dadurch, dass man  $\mathbf{k}$  durch  $\mathbf{G}$  ersetzt und dann Gl. (1.4) verwendet.

Mit dem Bloch-Theorem kann man auch zeigen, dass die Energiebeziehung  $E(\mathbf{k})$  nicht nur im Ortsraum, sondern auch im reziproken Gitter periodisch ist, d. h.  $E(\mathbf{k}) = E(\mathbf{k} + \mathbf{G})$ , wobei  $\mathbf{G}$  durch Gl. (1.3) gegeben ist. Für einen gegebenen Bandindex reicht es zur eindeutigen Kennzeichnung der Energie aus, nur  $\mathbf{k}$ -Werte in einer einzigen primitiven Zelle des reziproken Gitters zu verwenden, die aber leider nicht eindeutig definiert ist. Die Standardkonvention ist jedoch die Verwendung der Wigner-Seitz-Zelle im reziproken Gitter (Abb. 1.3), die auch (erste) Brillouin-Zone genannt wird [14]. Der Einfachheit halber wird dann jeder Impuls  $\mathbf{k}$  im reziproken Raum auf einen Punkt innerhalb der ersten Brillouin-Zone zurückgespiegelt, wo der zugehörige Energiezustand in diesem „reduzierten Zonenschema“ mit einem eigenen Index versehen werden kann.

Die Brillouin-Zone für Diamant- und Zinkblende-Gitter ist die gleiche wie die des fcc-Gitters und in Abb. 1.3a dargestellt. Tabelle 1.2 fasst die wichtigsten Symmetriepunkte und Symmetrielinien zusammen, wie die Mitte der Zone, die Zonenränder und ihre entsprechenden  $k$ -Achsen.

Die Energiebänder von Festkörpern wurden theoretisch mit einer Vielzahl von numerischen Methoden untersucht. Die am häufigsten verwendeten Methoden bei Halbleitern sind die orthogonalisierte Ebene-Wellen-Methode (engl. orthogonalized plane-wave method) [18, 19], die Pseudopotential-Methode [20] und die  $\mathbf{k} \cdot \mathbf{p}$ -Methode [6]. Abbildung 1.4 zeigt die damit berechneten Bandstrukturen von Si und GaAs. Beachten Sie, dass es für jeden Halbleiter einen verbotenen Energiebereich gibt, in der keine Energiezustände existieren. Die  $E$ - $k$ -Beziehungen oberhalb dieses verbotenen Bereichs werden als Leitungsbänder bezeichnet, die Bänder unterhalb des verbotenen Bereichs als Valenzbänder. Der Abstand zwischen der Unterkante des niedrigsten Leitungsbandes und der Oberkante des höchsten Valenzbandes wird als Bandlücke (engl. gap) mit der Energie  $E_g$  bezeichnet, die einer der wichtigsten Parameter in der Halbleiterphysik ist. In Abb.1.4 ist der untere Teil des Leitungsbandes mit  $E_C$  und der obere Teil des Valenzbandes mit  $E_V$  bezeichnet. Innerhalb der Bänder ist die Elektronenenergie konventioneller Weise so definiert, dass sie positiv ist, wenn sie von  $E_C$  nach oben gemessen wird, und die Löcherenergie ist positiv, wenn sie von  $E_V$  nach unten

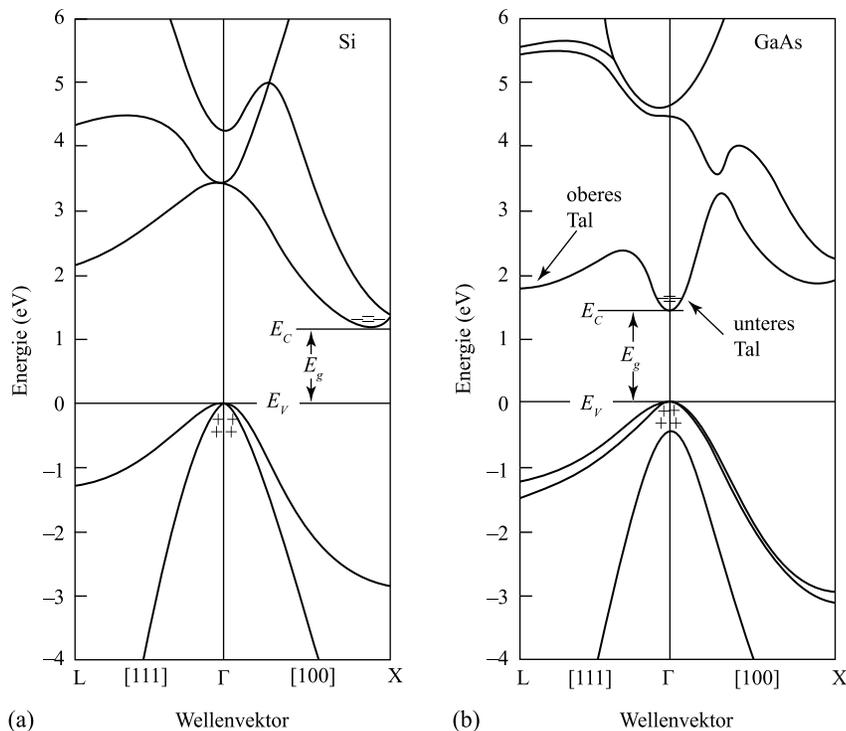


Abb. 1.4 Berechnete Bandstrukturen von (a) Si und (b) GaAs.  $E_g$  ist die Bandlücke, Pluszeichen (+) zeigen Löcher in den Valenzbändern und Minuszeichen (-) zeigen Elektronen in den Leitungsbandern an ([21], mit freundlicher Genehmigung von Springer Nature).

gemessen wird. Die Bandlücken einiger wichtiger Halbleiter sind in Anhang F aufgeführt.

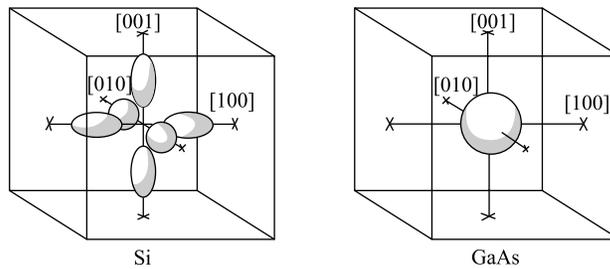
Das Valenzband der Zinkblende-Struktur, wie z. B. in GaAs (Abb. 1.4b), besteht aus vier Subbändern, wenn der Spin in der Schrödinger-Gleichung vernachlässigt wird. Unter Berücksichtigung des Spins spaltet sich jedes Subband nochmals in zwei Bänder auf. Drei der vier Bänder sind bei  $k = 0$  (Punkt  $\Gamma$ ) entartet und bilden den oberen Rand des Valenzbandes, das vierte Band bildet den unteren Rand (nicht dargestellt). Außerdem bewirkt die Spin-Bahn-Wechselwirkung eine Aufspaltung des Valenzbandes bei  $k = 0$ .

In der Nähe der Bandkanten, d. h. am Boden des Leitungsbandes  $E_C$  und an der Oberkante des Valenzbandes  $E_V$ , kann die  $E$ - $k$ -Beziehung durch eine quadratische Gleichung approximiert werden

$$E(k) = \frac{\hbar^2 k^2}{2m^*}, \quad (1.8)$$

wobei  $m^*$  die zugehörige effektive Masse ist. Wie aber in Abb. 1.4 gezeigt, haben die beiden oberen Valenzbänder entlang einer gegebenen Richtung unterschiedliche Krümmungen und werden als das schwere Löcherband (das breitere Band auf der  $k$ -Achse mit dem kleineren  $\partial^2 E / \partial k^2$ ) und das leichte Löcherband (das schmalere Band mit dem größeren  $\partial^2 E / \partial k^2$ ) bezeichnet. Die effektive Masse  $m^*$  ist im Allgemeinen ein Tensor  $m_{ij}^*$ , der über folgende Komponenten definiert ist:

$$\frac{1}{m_{ij}^*} \equiv \frac{1}{\hbar^2} \frac{\partial^2 E(k)}{\partial k_i \partial k_j}. \quad (1.9)$$



**Abb. 1.5** Flächen konstanter Energie für Elektronen in Si und GaAs. Für Si gibt es sechs Ellipsoide entlang der [100]-Achsen, wobei die Zentren der Ellipsoide bei etwa drei Viertel der Entfernung zwischen dem Zentrum der Brillouin-Zone und dem Zonenrand liegen. Bei GaAs ist die Oberfläche mit konstanter Energie eine Kugel im Zonenzentrum (verändert nach [22]).

Für wichtige Halbleiter sind die effektiven Massen in Anhang F aufgelistet. Im Bild der  $E$ - $k$ -Beziehungen haben sich bewegende Ladungsträger auch eine Gruppengeschwindigkeit  $v_g = (dE/dk)/\hbar$  und den Impuls  $p = \hbar k$ .

Wie man in Abb. 1.4 sieht, besteht auch das Leitungsband aus einer Reihe von Subbändern. Das Minimum des Leitungsbandes kann in der Mitte bei  $k = 0$  ( $\Gamma$ -Punkt) oder außerhalb der Mitte entlang verschiedener  $k$ -Achsen erscheinen. Symmetrieerwägungen allein bestimmen jedoch nicht die Lage des Leitungsbandminimums und alle Berechnungen zu diesem Thema sind sehr kompliziert. Sichere Aussagen zu diesem Thema bringen nur ein paar Experimente, und die sagen, dass das Leitungsbandminimum in Si außerhalb des Zentrums der Brillouin-Zone und entlang der [100]-Achse liegt ( $\Delta$ ), wogegen es in GaAs am untersten Punkt der  $E$ - $k$ -Beziehung bei  $k = 0$  ( $\Gamma$ -Punkt) zu finden ist. Halbleiter, in denen das Leitungsbandminimum und das Valenzbandmaximum den gleichen  $k$ -Wert haben, nennt man direkte Halbleiter; liegen das Leitungsbandminimum und das Valenzbandmaximum bei unterschiedlichen  $k$ -Werten, spricht man von indirekten Halbleitern. Direkte und indirekte Halbleiter verhalten sich sehr unterschiedlich für den Fall, dass Ladungsträger zwischen dem Valenzband und Leitungsband transferiert werden, denn für direkte Halbleiter bleibt der Impuls erhalten, bei indirekten Halbleitern jedoch nicht.

Abbildung 1.5 zeigt die Flächen konstanter Energie im  $k$ -Raum für Si und GaAs. Für Si gibt es sechs Ellipsoide entlang der [100]-Achsen, wobei die Zentren der Ellipsoide auf etwa drei Viertel der Strecke zwischen dem Zentrum der Brillouin-Zone und dem Zonenrand liegen. Für GaAs ist die Oberfläche mit konstanter Energie eine Kugel im Zonenzentrum. Durch Anpassung der experimentellen Ergebnisse an parabolische Bänder erhalten wir die effektiven Massen der Elektronen im Leitungsband; eine Masse  $m^*$  für GaAs und zwei verschiedene Massen für Si,  $m_l^*$  entlang der Symmetrieachsen und  $m_t^*$  quer zu den Symmetrieachsen. Im Anhang G lassen sich diese Werte in einer Tabelle finden.

Bei Raumtemperatur und normalem Atmosphärendruck betragen die Werte der Bandlücke 1.12 eV für Si und 1.42 eV für GaAs. Diese Werte gelten für hochreine Materialien. Bei hochdotierten Materialien werden die Bandlücken kleiner. Experimentelle Ergebnisse zeigen auch, dass die Bandlücken der meisten Halbleiter mit steigender Temperatur abnehmen. Abbildung 1.6 zeigt die Änderung der Bandlücken von Si und GaAs als Funktion der Temperatur. Die Bandlücke nähert sich bei  $T = 0$  K dem Wert 1.17 eV für Si bzw. 1.52 eV für GaAs. Die Variation der Bandlücke mit der Temperatur kann näherungsweise durch folgende semiempirische Funktion ausgedrückt werden:

$$E_g(T) \approx E_g(0) - \frac{\alpha T^2}{T + \beta}, \quad (1.10)$$

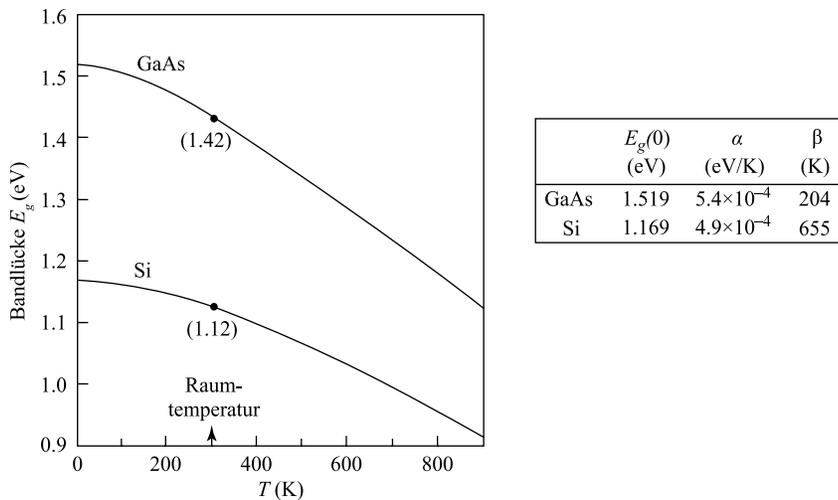


Abb. 1.6 Die Bandlücken von Si und GaAs als Funktion der Temperatur (nach [23, 24]).

wobei  $E_g(0)$ ,  $\alpha$  und  $\beta$  in der Tabelle in Abb. 1.6 angegeben sind. Der Temperaturkoeffizient  $dE_g/dT$  ist bei beiden Halbleitern negativ. Einige Halbleiter haben ein positives  $dE_g/dT$ ; z. B. steigt die Bandlücke von PbS (Anhang F) von 0.286 eV bei  $T = 0$  K auf 0.41 eV bei  $T = 300$  K. In der Nähe der Raumtemperatur nimmt die Bandlücke von GaAs mit steigendem hydrostatischem Druck  $P$  zu [25].  $dE_g/dP$  beträgt in etwa  $12.6 \times 10^{-6}$  eV cm<sup>2</sup>/N, während die Si-Bandlücke mit steigendem Druck abnimmt, wobei der Wert  $dE_g/dP = -2.4 \times 10^{-6}$  eV cm<sup>2</sup>/N beträgt.

## 1.4 Ladungsträgerkonzentrationen im thermischen Gleichgewicht

Einer der wichtigsten Vorteile eines Halbleiters ist es, dass er mit verschiedenen Arten und Konzentrationen von Fremdstoffen dotiert werden kann. Auf diese Weise ist es möglich, durch eine Ionisierung (meist thermisch) dieser Fremdstoffe den spezifischen Widerstand des Halbleiters gezielt zu variieren. Wenn diese Fremdstoffe im Halbleiter ionisiert sind, sich aber nach der Ionisierung keine freien Ladungsträger mehr im dotierten Gebiet befinden (Verarmung), hinterlassen die ionisierten Fremdstoffe eine ortsfeste Raumladungszone, die zu einem elektrischen Feld und zu einer Potentialbarriere im Inneren des Halbleiters führt. Solche Eigenschaften gibt es bei einem Metall oder einem Isolator nicht.

Abbildung 1.7 zeigt die drei grundlegenden Situationen für die Bindungen zwischen den Atomen des Halbleiters. Abbildung 1.7a zeigt intrinsisches Silizium, das sehr rein ist und nur eine vernachlässigbar geringe Menge an Störstellen enthält. Jedes Siliziumatom teilt seine vier Valenzelektronen mit den vier benachbarten Atomen und bildet dabei vier kovalente Bindungen aus (siehe auch Abb. 1.1). Abbildung 1.7b zeigt ein  $n$ -Typ-Silizium, bei dem ein substitutionelles Phosphoratom mit fünf Valenzelektronen ein Siliziumatom ersetzt hat und ein negativ geladenes Elektron an das Gitter im Leitungsband abgegeben wird. Das Phosphoratom wird als Donator bezeichnet. Abbildung 1.7c zeigt in ähnlicher Weise, dass, wenn ein Boratom mit drei Valenzelektronen ein Siliziumatom ersetzt, ein positiv geladenes Loch im Valenzband entsteht und ein zusätzliches Elektron aufgenommen wird, um vier kovalente Bindungen um das Bor zu bilden. Dies ist der Fall eines  $p$ -Typ-Halbleiters, und Bor ist ein Akzeptor.

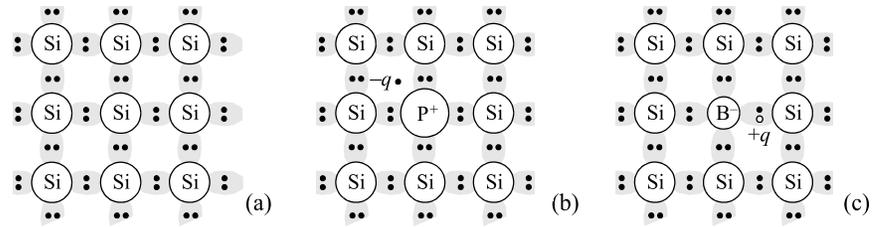


Abb. 1.7 Drei grundlegende Bindungsmöglichkeiten im Halbleiter. (a) Intrinsic Silizium ohne Fremdatome. (b) *n*-Typ-Silizium mit Donator (Phosphor). (c) *p*-Typ-Silizium mit Akzeptor (Bor).

Die Bezeichnungen „*n*- und *p*-Typ“ wurden in der Urzeit der Halbleiterforschung geprägt, als man Schottky-Dioden (siehe Kap. 3) dadurch erzeugte, dass Metallspitzen auf einen Halbleiter gepresst wurden. Für einen *p*-Typ-Halbleiter wurde beobachtet, dass eine positive Vorspannung auf dem Halbleiter erforderlich ist, um einen spürbaren Strom durch die Spitze zu erzeugen [26, 27]. Außerdem wurde bei Lichteinwirkung ein positives Potential in Bezug auf die Metallspitze erzeugt. Umgekehrt war bei einem *n*-Typ-Material eine negative Vorspannung erforderlich, um einen Stromfluss zu erzeugen und auch die Fotospannung war negativ.

#### 1.4.1 Ladungsträgerkonzentration und Fermi-Niveau

Wir betrachten zunächst den Fall eines undotierten (intrinsic) Halbleiters. Die Anzahl der Elektronen (= Anzahl der besetzten Zustände im Leitungsband) ergibt sich aus der energieabhängigen Dichte der Zustände  $N(E)$  multipliziert mit der Besetzung  $F(E)$ , integriert über die Energie im Leitungsband,

$$n = \int_{E_C}^{\infty} N(E)F(E)dE . \quad (1.11)$$

Die Zustandsdichte  $N(E)$  kann nahe dem Leitungsbandminimum (Anmerkung: parabolischer Bandverlauf) für niedrige Trägerdichten und Temperaturen folgendermaßen angenähert werden [2]:

$$N(E) = M_C \frac{\sqrt{2} m_{de}^{3/2} (E - E_C)^{1/2}}{\pi^2 \hbar^3} . \quad (1.12)$$

$M_C$  ist dabei die Anzahl der äquivalenten Minima im Leitungsband, und  $m_{de} = (m_1^* m_2^* m_3^*)^{1/3}$  ist die sogenannte effektive Zustandsdichtemasse der Elektronen [5], wobei  $m_1^*$ ,  $m_2^*$ ,  $m_3^*$  die effektiven Massen entlang der Hauptachsen der ellipsenförmigen Energiefläche sind. Als typisches Beispiel erwähnen wir hier Silizium mit  $m_{de} = (m_1^* m_2^*)^{1/3}$ . Die Besetzung  $F(E)$  ist eine Funktion von Temperatur und Energie und wird Fermi-Dirac-Verteilungsfunktion genannt,

$$F(E) = \frac{1}{1 + \exp[(E - E_F)/kT]} \quad (1.13)$$

wobei  $E_F$  das sogenannte Fermi-Niveau ist, das aus der Bedingung für die Ladungsneutralität (siehe Abschn. 1.4.3) bestimmt werden kann.

Rechnet man das Integral in Gl. (1.13) aus, bekommt man

$$n = N_C \frac{2}{\sqrt{\pi}} F_{1/2} \left( \frac{E_F - E_C}{kT} \right) , \quad (1.14)$$

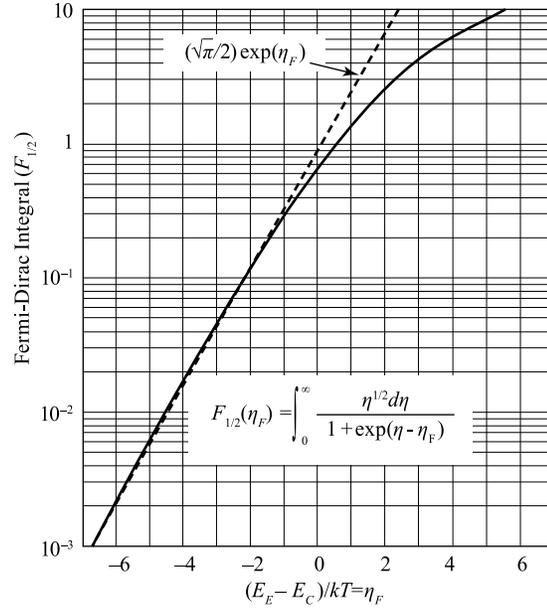


Abb. 1.8 Fermi-Dirac-Integral  $F_{1/2}$  als Funktion der Fermi-Energie (basierend auf [28]). Die gestrichelte Linie ist eine Näherung mit der Boltzmann-Statistik.

wobei  $N_C$  die effektive Zustandsdichte im Leitungsband ist

$$N_C \equiv 2 \left( \frac{2\pi m_{de} kT}{h^2} \right)^{3/2} M_C. \quad (1.15)$$

Das Fermi-Dirac-Integral ist nach einer Variablentransformation mit  $\eta \equiv (E - E_C)/kT$  und  $\eta_F \equiv (E_F - E_C)/kT$  gegeben durch die Funktion

$$\begin{aligned} F_{1/2} \left( \frac{E_F - E_C}{kT} \right) &\equiv F_{1/2}(\eta_F) = \int_{E_C}^{\infty} \frac{[(E - E_C)/kT]^{1/2} dE}{1 + \exp[(E - E_F)/kT]} \frac{1}{kT} \\ &= \int_0^{\infty} \frac{\eta^{1/2}}{1 + \exp(\eta - \eta_F)} d\eta, \end{aligned} \quad (1.16)$$

deren Werte in Abb. 1.8 dargestellt sind. Beachten Sie, dass für  $\eta_F \leq -1$  das Integral durch eine Exponentialfunktion approximiert werden kann. Bei  $\eta_F = 0$ , also wenn das Fermi-Niveau mit der Bandkante zusammenfällt, hat das Integral einen Wert von  $\approx 0.6$ , sodass  $n \approx 0.7N_C$ .

**Nicht degenerierte Halbleiter** Per Definition sind in nicht entarteten (nicht degenerierten) Halbleitern die Dotierungskonzentrationen kleiner als  $N_C$  und die Fermi-Niveaus liegen mehr als einige  $kT$  unterhalb von  $E_C$  (negatives  $\eta_F$ ). Das Fermi-Dirac-Integral kann dann mithilfe der Boltzmann-Statistik so geschrieben werden:

$$F_{1/2} \left( \frac{E_F - E_C}{kT} \right) = \frac{\sqrt{\pi}}{2} \exp \left( -\frac{E_C - E_F}{kT} \right). \quad (1.17)$$

Gleichung (1.17) wird dann zu

$$n = N_C \exp \left( -\frac{E_C - E_F}{kT} \right) \quad \text{oder} \quad E_C - E_F = kT \ln \left( \frac{N_C}{n} \right). \quad (1.18)$$

In ähnlicher Weise können wir bei  $p$ -Typ-Halbleitern die Löcherdichte und das Fermi-Niveau nahe der Oberkante des Valenzbandes erhalten;

$$p = N_V \frac{2}{\sqrt{\pi}} F_{1/2} \left( \frac{E_V - E_F}{kT} \right) \quad (1.19)$$

kann zu

$$p = N_V \exp\left(-\frac{E_F - E_V}{kT}\right) \quad \text{oder} \quad E_F - E_V = kT \ln\left(\frac{N_V}{p}\right) \quad (1.20)$$

vereinfacht werden, wobei  $N_V$  die effektive Dichte der Zustände im Valenzband

$$N_V \equiv 2 \left( \frac{2\pi m_{dh} kT}{h^2} \right)^{3/2} \quad (1.21)$$

und  $m_{dh}$  ist die effektive Zustandsdichtemasse im Valenzband ist [2]:

$$m_{dh} = \left( m_{lh}^{*3/2} + m_{hh}^{*3/2} \right)^{2/3}. \quad (1.22)$$

Die Indizes bezeichnen die leichten und schweren Löcher, die schon früher mit Gl. (1.9) eingeführt wurden.

**Entartete Halbleiter** Wie man in Abb. 1.8 sieht, muss für degenerierte Halbleiter, bei denen die  $n$ - oder  $p$ -Konzentrationen nahe oder jenseits der effektiven Zustandsdichten  $N_C$  oder  $N_V$  liegen, der exakte Wert des Fermi-Dirac-Integrals anstelle der vereinfachten Boltzmann-Statistik verwendet werden. Für  $\eta_F > -1$  hat das Integral aber eine relativ schwache Abhängigkeit von der Ladungsträgerkonzentration. Beachten Sie auch, dass für degenerierte Halbleiter die Fermi-Niveaus außerhalb der Bandlücke liegen. Eine nützliche Abschätzung des Fermi-Niveaus als Funktion der Ladungsträgerkonzentration für  $n$ -Typ-Halbleiter [29] bekommt man durch

$$E_F - E_C \approx kT \left[ \ln\left(\frac{n}{N_C}\right) + 2^{-3/2} \left(\frac{n}{N_C}\right) \right] \quad (1.23a)$$

und für  $p$ -Typ-Halbleiter durch

$$E_V - E_F \approx kT \left[ \ln\left(\frac{p}{N_V}\right) + 2^{-3/2} \left(\frac{p}{N_V}\right) \right]. \quad (1.23b)$$

**Intrinsische Ladungsträgerkonzentration** In intrinsischen Halbleitern kommt es bei endlichen Temperaturen zu einer kontinuierlichen thermischen Anregung von Elektronen aus dem Valenzband in das Leitungsband, welche eine gleiche Anzahl von Löchern im Valenzband hinterlässt. Dieser Prozess wird durch eine Rekombination der Elektronen im Leitungsband mit Löchern im Valenzband wieder ausgeglichen. Im stationären Zustand ist das Nettoergebnis  $n = p = n_i$ , wobei  $n_i$  die intrinsische Ladungsträgerdichte ist.

Das Fermi-Niveau für einen intrinsischen Halbleiter (der per Definition nicht entartet ist) erhält man durch Gleichsetzen von Gln. (1.21) und (1.23):

$$E_F = E_i = \frac{E_C + E_V}{2} + \frac{kT}{2} \ln\left(\frac{N_V}{N_C}\right) = \frac{E_C + E_V}{2} + \frac{3kT}{4} \ln\left(\frac{m_{dh}}{m_{de} M_C^{2/3}}\right). \quad (1.24)$$

Daher liegt das Fermi-Niveau eines intrinsischen Halbleiters  $E_i$  in der Regel sehr nahe bei, aber nicht genau in der Mitte der Bandlücke. Die intrinsische Trägerdichte  $n_i$  kann aus Gln. (1.18) oder (1.20) entnommen werden:

$$\begin{aligned} n_i &= N_C \exp\left(-\frac{E_C - E_i}{kT}\right) = N_V \exp\left(-\frac{E_i - E_V}{kT}\right) = \sqrt{N_C N_V} \exp\left(-\frac{E_g}{2kT}\right) \\ &= 4.9 \times 10^{15} \left(\frac{m_{de} m_{dh}}{m_0^2}\right)^{3/4} M_C^{1/2} T^{3/2} \exp\left(-\frac{E_g}{2kT}\right). \end{aligned} \quad (1.25)$$

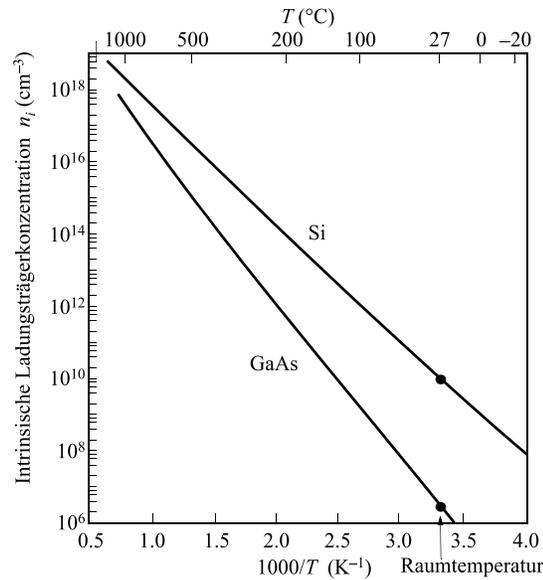


Abb. 1.9 Intrinsiche Ladungsträgerkonzentrationen von Si und GaAs als Funktion der reziproken Temperatur (nach [6, 23]).

Abbildung 1.9 zeigt die Temperaturabhängigkeit von  $n_i$  für Si und GaAs. Wie erwartet, ist die intrinsiche Trägerdichte umso geringer, je größer die Bandlücke ist [6, 31].

Bei nicht entarteten Halbleitern bekommt man für das Produkt aus den Konzentrationen von Majoritäts- und Minoritätsladungsträgern folgenden Ausdruck

$$pn = N_C N_V \exp\left(-\frac{E_g}{kT}\right) = n_i^2, \quad (1.26)$$

der (in Analogie zur Chemie) auch als Massenwirkungsgesetz bekannt ist. Hinweis: Für entartete Halbleiter gilt das nicht, denn dort wir haben eine Situation mit  $pn < n_i^2$ . Ebenfalls unter Verwendung von Gl. (1.28) bekommen wir mit  $E_i$  als Energienullpunkt zusätzliche alternative Gleichungen, nämlich für einen  $n$ -Typ-Halbleiter

$$n = n_i \exp\left(\frac{E_F - E_i}{kT}\right) \quad \text{oder} \quad E_F - E_i = kT \ln\left(\frac{n}{n_i}\right) \quad (1.27a)$$

und für  $p$ -Typ-Materialien

$$p = n_i \exp\left(\frac{E_i - E_F}{kT}\right) \quad \text{oder} \quad E_i - E_F = kT \ln\left(\frac{p}{n_i}\right). \quad (1.27b)$$

### 1.4.2 Donatoren und Akzeptoren

Wenn ein Halbleiter mit Donator- oder Akzeptoratomen dotiert wird, gibt es dazu gehörige Störstellenniveaus, die normalerweise energetisch innerhalb der Bandlücke liegen. Ein Donatoratom hat ein Donatorniveau, das als neutral definiert ist, wenn es mit einem Elektron gefüllt ist, und positiv, wenn es leer ist. Umgekehrt ist ein Akzeptorniveau neutral, wenn es leer ist, und negativ, wenn es mit einem Elektron gefüllt ist. Die Lage dieser Energieniveaus ist wichtig für die Berechnung der Konzentration der Dotierstoffe, welche ionisiert bzw. elektrisch aktiv sind (siehe Abschn. 1.4.3).

Um ein Gefühl für die Größe der Ionisationsenergie der Störstellen zu bekommen, verwenden wir ein einfaches Modell, das auf dem Modell des Wasserstoffatoms basiert. Die Ionisationsenergie für ein Wasserstoffatom im Vakuum beträgt

$$E_H = \frac{m_0 q^4}{32\pi^2 \epsilon_0^2 \hbar^2} = 13.6 \text{ eV}. \quad (1.28)$$

Die Ionisationsenergie für einen Donator ( $E_C - E_D$ ) in einem Gitter bekommen wir durch Ersetzen von  $m_0$  durch die effektive Leitfähigkeitsmasse der Elektronen [2]

$$m_{ce} = 3 \left( \frac{1}{m_1^*} + \frac{1}{m_2^*} + \frac{1}{m_3^*} \right)^{-1} \quad (1.29)$$

und durch Ersetzen von  $\varepsilon_0$  durch die Dielektrizitätskonstante des Halbleiters  $\varepsilon_0 \varepsilon_s$  in Gl. (1.28):

$$E_C - E_D = \left( \frac{\varepsilon_0}{\varepsilon_s} \right)^2 \left( \frac{m_{ce}}{m_0} \right) E_H. \quad (1.30)$$

Nach Gl. (1.30) beträgt die Ionisationsenergie für Donatoren 0.025 eV für Si und 0.007 eV für GaAs. Die Berechnung des Ionisierungsgrades für die Akzeptoren läuft analog wie für die Donatoren. Die berechnete Akzeptorionisationsenergie (gemessen von der Valenzbandkante,  $E_a \equiv (E_A - E_V)$ ) beträgt 0.05 eV für Si und GaAs.

Obwohl dieses einfache Wasserstoffmodell die Details der Ionisationsenergie, insbesondere für die tiefen Störstellenniveaus in Halbleitern [32–34], sicherlich nicht berücksichtigen kann, sagen die berechneten Werte die richtige Größenordnung der wahren Ionisationsenergien für flache Störstellen recht gut voraus. Diese berechneten Werte sind nachweislich viel kleiner als die Bandlücke und werden als flache Störstellen bezeichnet, weil sie nahe an den Bandkanten liegen. Da die Ionisationsenergien der flachen Störstellen mit der Wärmeenergie  $kT$  vergleichbar sind, führt dies in der Regel bei Raumtemperatur zu kompletter Ionisierung. Abbildung 1.10 zeigt die gemessenen Ionisierungsenergien für verschiedene Störstellen in Si und GaAs. Beachten Sie, dass es möglich ist, dass ein einzelnes Atom mehrere Niveaus hat; Gold in Silizium, z. B., hat sowohl ein Akzeptorniveau als auch ein Donatorniveau innerhalb der verbotenen Bandlücke.

### 1.4.3 Berechnung des Fermi-Niveaus

Das Fermi-Niveau für den intrinsischen Halbleiter (siehe Gl. (1.27)) liegt sehr nahe der Mitte der Bandlücke. Abbildung 1.11a illustriert die Situation und zeigt schematisch von links nach rechts das vereinfachte Banddiagramm, die Zustandsdichte  $N(E)$ , die Fermi-Dirac-Verteilung  $F(E)$  und die Ladungsträgerkonzentrationen. Die schattierten Bereiche im Leitungs- und Valenzband stellen die jeweiligen Elektronen- und Löcherkonzentrationen dar. Für den intrinsischen Fall ist ihre Anzahl gleich, d. h.,  $n = p = n_i$ .

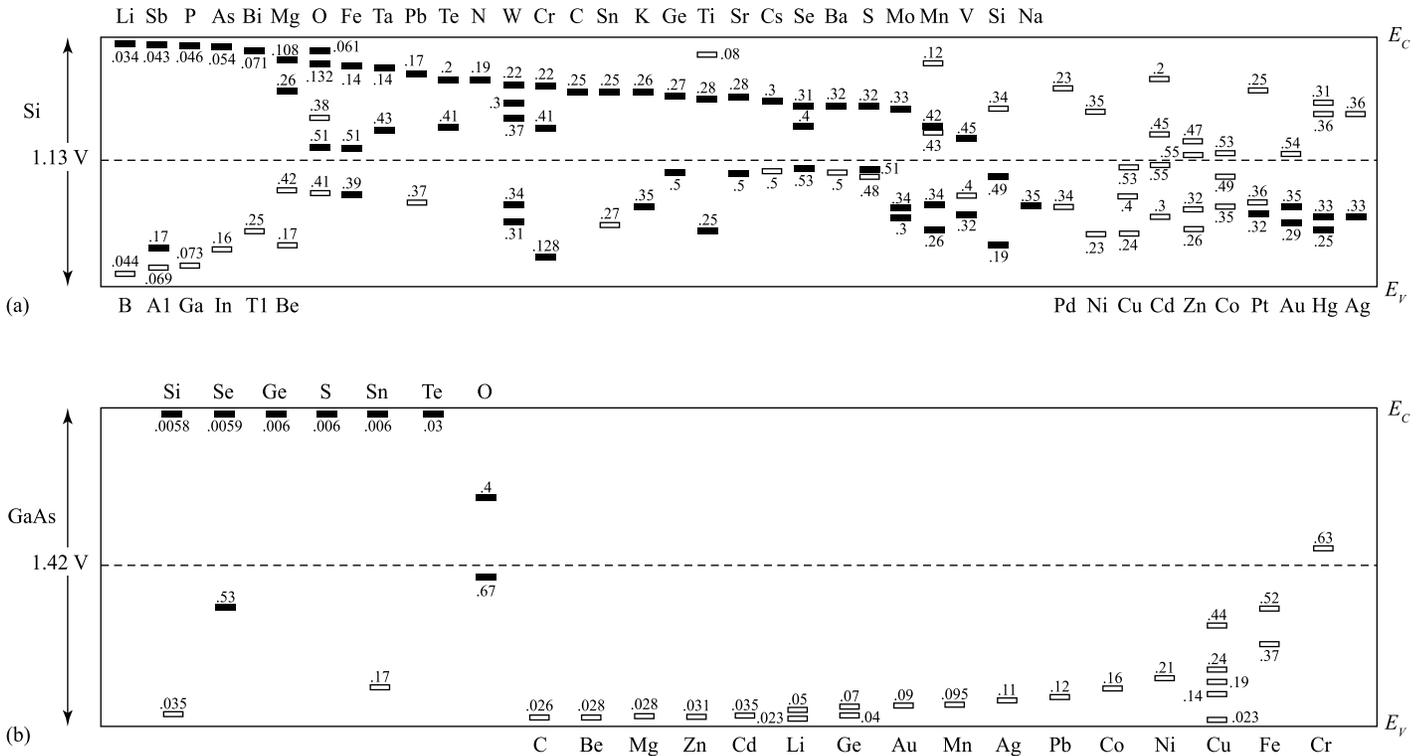
Wenn Störstellen in einen Halbleiter eingebracht werden, sind, abhängig von der Größe der Störstellenenergie und der Gittertemperatur, nicht unbedingt alle Störstellen ionisiert. Die Konzentration der ionisierten Donatoren ist durch [37]

$$N_D^+ = \frac{N_D}{1 + g_D \exp[(E_F - E_D)/kT]} \quad (1.31)$$

gegeben, wobei  $g_D$  die Entartung des Grundzustandes des Donators ist.  $g_D$  ist gleich zwei, weil ein Donatorniveau entweder Elektronen mit unterschiedlichem Spin oder gar kein Elektron aufnehmen kann. Wenn Akzeptoren der Konzentration  $N_A$  zu einem Halbleiter hinzugefügt werden, bekommt man einen ähnlichen Ausdruck für die ionisierten Akzeptoren

$$N_A^- = \frac{N_A}{1 + g_A \exp[(E_A - E_F)/kT]}, \quad (1.32)$$

wobei die Entartung des Grundzustandes den Wert  $g_A = 4$  hat. Der Wert ist vier, weil in den meisten Halbleitern jedes Akzeptorniveau ein Loch mit beiden Spins aufnehmen



**Abb. 1.10** Gemessene Ionisierungsenergien für unterschiedliche Störstellen in (a) Si und (b) GaAs innerhalb der Bandlücke. Energieniveaus unterhalb der Lückenmitte werden von  $E_V$  aus, Energieniveaus oberhalb der Lückenmitte werden von  $E_C$  aus gemessen. Die fetten Symbole repräsentieren die Donatorniveaus und die nicht ausgefüllten Symbole die Akzeptorniveaus (nach [30, 32, 35, 36]).

kann und das Störstellenniveau wegen der beiden degenerierten Valenzbänder bei  $k = 0$  doppelt entartet ist.

In Anwesenheit von Störstellen muss die Gesamtkonzentration der negativen Ladungen (Elektronen und ionisierte Akzeptoren) der Gesamtkonzentration aller positiven Ladungen (Löcher und ionisierte Donatoren) entsprechen. Dies wird durch die Formel für die Ladungsneutralität beschrieben

$$n + N_A^- = p + N_D^+ . \quad (1.33)$$

In dotierten, nicht entarteten Halbleitern gilt weiterhin das Massenwirkungsgesetz ( $pn = n_i^2$ ) aus Gl. (1.26) und das  $p$ - $n$ -Produkt ist weiterhin von der Störstellenkonzentration unabhängig.

Betrachten wir den in Abb. 1.11b gezeigten Fall, bei welchem der Kristall mit Donatoren in der Konzentration  $N_D$  ( $\text{cm}^{-3}$ ) dotiert wurde. Die Bedingung der Ladungsneutralität wird dann zu

$$n = N_D^+ + p \approx N_D^+ . \quad (1.34)$$

Nach der Substitution von  $n$  und  $N_D$  erhalten wir

$$N_C \exp\left(-\frac{E_C - E_F}{kT}\right) \approx \frac{N_D}{1 + 2 \exp[(E_F - E_D)/kT]} . \quad (1.35)$$

Mit obiger Formel kann für gegebene Werte von  $N_D$ ,  $E_D$ ,  $N_C$  und  $T$  das Fermi-Niveau  $E_F$  eindeutig, aber leider nur implizit, bestimmt werden. Nach der Bestimmung von  $E_F$  können dann die Ladungsträgerkonzentrationen  $n$  berechnet werden.

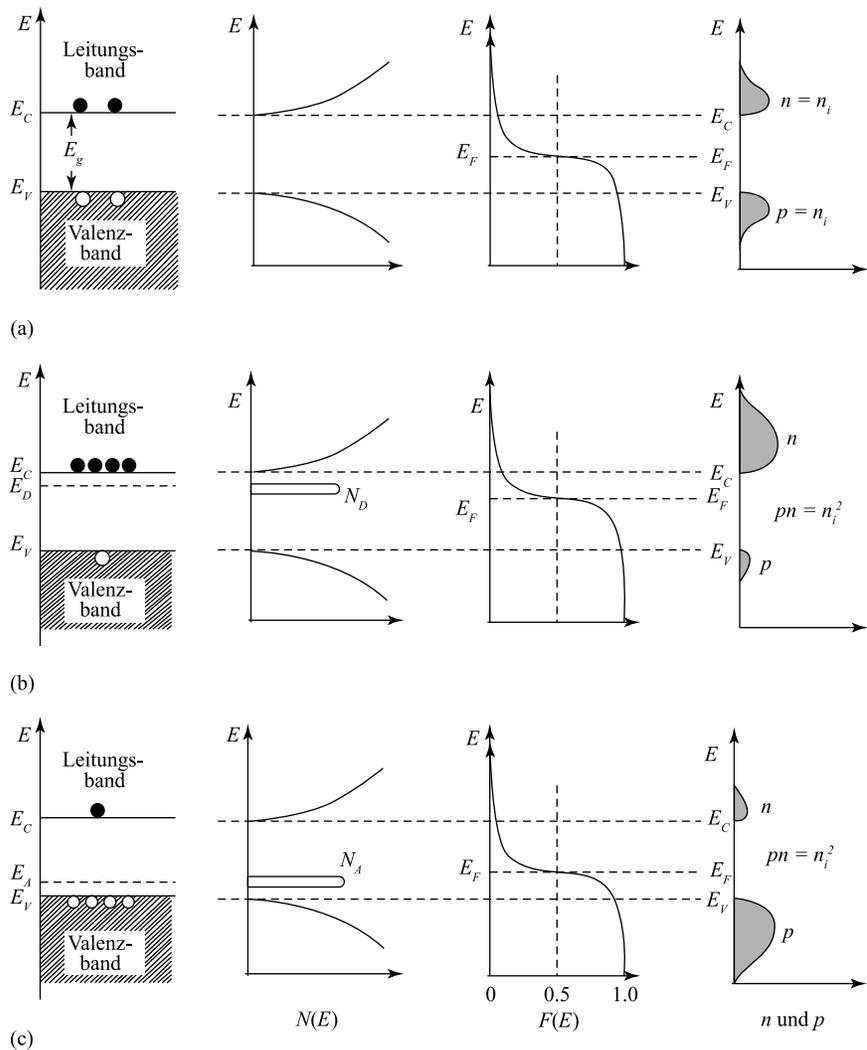


Abb. 1.11 Schematisches Banddiagramm, Zustandsdichte, Fermi-Dirac-Verteilung und Ladungsträgerkonzentrationen für (a) intrinsische, (b) *n*-Typ- und (c) *p*-Typ-Halbleiter bei thermischem Gleichgewicht. Beachten Sie, dass hier  $pn = n_i^2$  für alle drei Fälle gilt.

Gleichung (1.35) kann auch grafisch gelöst werden. In Abb. 1.12 sind die Werte von  $n$  und  $N_D^+$  als Funktion von  $E_F$  aufgetragen. Der Schnittpunkt der beiden Kurven bestimmt die Position von  $E_F$ . Auch ohne Gl. (1.35) zu lösen, kann gezeigt werden, dass für die Bedingung  $N_D \gg N_C \exp[-(E_C - E_D)/kT]/2 \gg N_A$  die Elektronenkonzentration durch folgende Formel angenähert werden kann [6]:

$$n \approx \sqrt{\frac{N_D N_C}{2}} \exp\left[-\frac{(E_C - E_D)}{2kT}\right]. \quad (1.36)$$

Für kompensiertes *n*-Typ-Material ( $N_D > N_A$ ) mit nicht vernachlässigbarer Akzeptorkonzentration, also wenn die Beziehung  $N_A \gg 1/2 N_C \exp[(E_C - E_D)/kT]$  erfüllt ist, gilt als Näherung für die Elektronendichte:

$$n \approx \left(\frac{N_D - N_A}{2N_A}\right) N_C \exp\left[-\frac{(E_C - E_D)}{kT}\right]. \quad (1.37)$$

Abbildung 1.13 zeigt ein typisches Beispiel, bei dem die Elektronendichte  $n$  als Funktion der reziproken Temperatur aufgetragen ist. Bei hohen Temperaturen befinden wir uns im intrinsischen Bereich, da  $n \approx p \approx n_i \gg N_D$ . In Worten ausgedrückt:

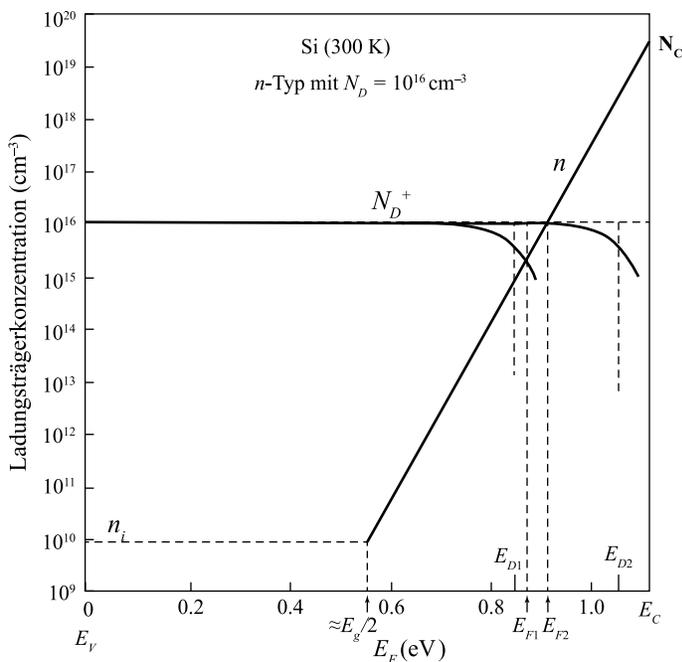


Abb. 1.12 Grafische Methode zur Bestimmung des Fermi-Niveaus  $E_F$  und der Elektronenkonzentration  $n$  bei nicht vollständiger Ionisierung und zwei verschiedenen Werten der Störstellenenergie  $E_D$ .

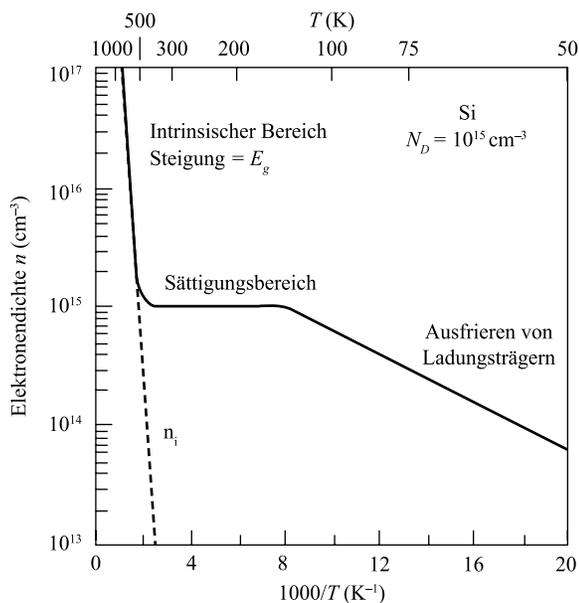
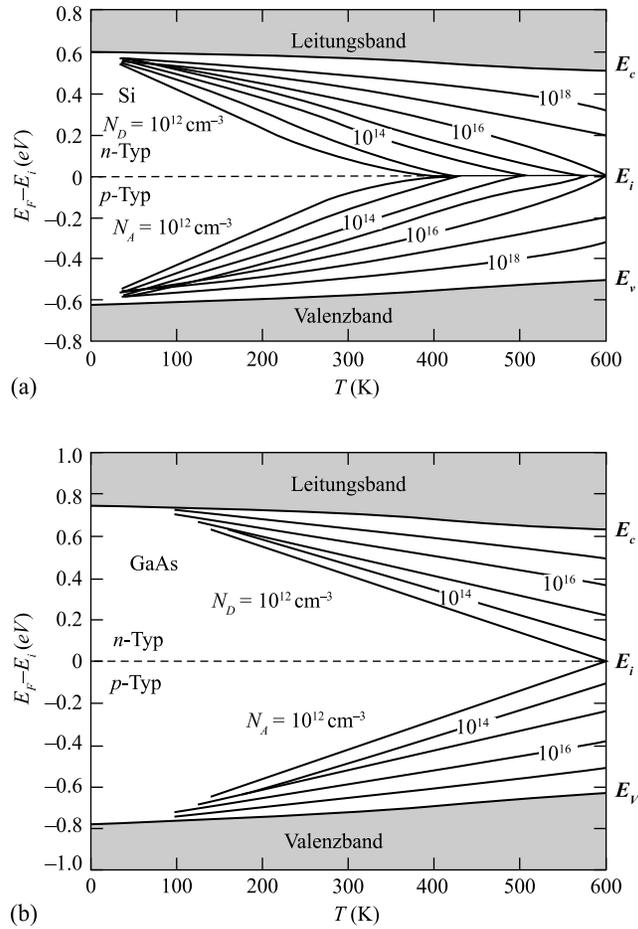


Abb. 1.13 Elektronendichte einer Si-Probe mit einer Donatorkonzentration von  $10^{15} \text{ cm}^{-3}$  als Funktion der Temperatur (nach [8]). Der Sättigungsbereich wird gerne auch als Störstellenreserve bezeichnet.

Auf thermischem Wege werden so viele Ladungsträger über die Bandlücke generiert, dass die paar zusätzlichen Ladungsträger von den ionisierten Dotierstoffatomen völlig egal sind. Bei eher niedrigen Temperaturen ist es genau umgekehrt. Es gibt kaum Ladungsträger, die über die Bandlücke angeregt werden, und die Dotierstoffatome spielen mit ihren geringen Ionisierungsenergien die dominante Rolle. Die Steigung ist je nach Kompensationsbedingung durch Gln. (1.36) oder (1.37) gegeben. Bei mittleren Temperaturen ist  $n \approx N_D$ , und die Elektronendichte bleibt über einen weiten Temperaturbereich im Wesentlichen konstant ( $T \approx 100 \text{ K}$  bis  $T \approx 500 \text{ K}$ ). Dieser Bereich ist auch unter dem Namen Störstellenreserve bekannt. Die beste anschauliche Erklärung hierfür kommt aus der Chemie. Jeder kennt eine gepufferte Säure. Die kann man verdünnen oder mehr konzentrierte Säure zugeben – der pH-Wert bleibt immer gleich. Jetzt ersetzen wir das Wort Säure durch das Wort Elektronen und Puffer



**Abb. 1.14** Fermi-Niveau für (a) Si und (b) GaAs als Funktion der Temperatur und der Störstellenkonzentration. Die Abhängigkeit der Bandlücke von der Temperatur ist ebenfalls eingezeichnet ([38], mit freundlicher Genehmigung von John Wiley & Sons).

durch Dotierung und dann noch das Wort pH-Wert durch Elektronenkonzentration. Und wenn man genau hinschaut – man glaubt es kaum –, ist der Formalismus im Halbleiter und in der Chemie tatsächlich extrem ähnlich. Nicht umsonst gibt es ja auch noch das Massenwirkungsgesetz für Elektronen und Löcher.

Abbildung 1.14 zeigt das Fermi-Niveau für Si und GaAs als Funktion der Temperatur und der Störstellenkonzentration sowie die Abhängigkeit der Bandlücke von der Temperatur (siehe Abb. 1.6). Bei relativ hohen Temperaturen werden die meisten Donatoren und Akzeptoren ionisiert, sodass die Neutralitätsbedingung nun durch

$$n + N_A = p + N_D \quad (1.38)$$

angenähert werden kann. Die Gln. (1.26) und (1.38) können verwendet werden, um die Konzentrationen von Elektronen und Löchern zu erhalten. In einem  $n$ -Typ-Halbleiter bekommt man mit  $|N_D - N_A| \gg n_i$  oder  $N_D \gg N_A$

$$n_{n0} = \frac{1}{2} \left[ (N_D - N_A) + \sqrt{(N_D - N_A)^2 + 4n_i^2} \right] \approx N_D \quad (1.39)$$

$$\Rightarrow p_{n0} = \frac{n_i^2}{n_{n0}} \approx \frac{n_i^2}{N_D}. \quad (1.40)$$

Das Fermi-Niveau lässt sich dann über

$$n_{n0} = N_D = N_C \exp\left(-\frac{E_C - E_F}{kT}\right) = n_i \exp\left(\frac{E_F - E_i}{kT}\right) \quad (1.41)$$

berechnen. In ähnlicher Weise wie oben berechnen sich die Ladungsträgerkonzentrationen in einem  $p$ -Typ-Halbleiter mit  $|N_A - N_D| \gg n_i$  oder  $N_A \gg N_D$  zu

$$n_{n0} = \frac{1}{2} \left[ (N_A - N_D) + \sqrt{(N_A - N_D)^2 + 4n_i^2} \right] \approx N_A \quad (1.42)$$

$$\Rightarrow n_{p0} = \frac{n_i^2}{p_{p0}} \approx \frac{n_i^2}{N_A}, \quad (1.43)$$

wobei

$$p_{p0} = N_A = N_V \exp\left(-\frac{E_F - E_V}{kT}\right) = n_i \exp\left(\frac{E_i - E_F}{kT}\right). \quad (1.44)$$

In den obigen Formeln beziehen sich die Indizes  $n$  und  $p$  auf die Art der Halbleiter und der Index 0 auf die Bedingung des thermischen Gleichgewichts. Bei  $n$ -Typ-Halbleitern wird das Elektron als Majoritätsladungsträger und das Loch als Minoritätsladungsträger bezeichnet, da die Elektronenkonzentration größer als die Löcherkonzentration ist. Bei den  $p$ -Typ-Halbleitern ist es genau umgekehrt.

## 1.5 Ladungsträgertransportphänomene

### 1.5.1 Drift und Beweglichkeit

Bei niedrigen elektrischen Feldern ist die Driftgeschwindigkeit  $v_d$  proportional zur elektrischen Feldstärke  $\mathcal{E}$ , also  $v_d = \mu\mathcal{E}$ . Die Proportionalitätskonstante wird als Beweglichkeit  $\mu$  ( $\text{cm}^2/\text{V}\cdot\text{s}$ ) bezeichnet. Bei nicht polaren Halbleitern, wie Ge und Si, führt die Anwesenheit von akustischen Phononen (siehe Abschn. 1.6.1) und ionisierten Störstellen zu Streuprozessen, welche die Beweglichkeit der Elektronen und Löcher erheblich beeinträchtigen. Die Beweglichkeit aus der Wechselwirkung mit den akustischen Phononen des Gitters,  $\mu_l$  (der Index  $l$  steht für engl. lattice), ist gegeben durch [39]

$$\mu_l = \frac{\sqrt{8\pi q \hbar^4} C_l}{3E_{ds}^2 m_c^{*5/2} (kT)^{3/2}} \propto \frac{1}{m_c^{*5/2} T^{3/2}}, \quad (1.45)$$

wobei  $C_l$  die mittlere longitudinale elastische Konstante des Halbleiters,  $E$  die Verschiebung der Bandkante pro Einheitsausdehnung des Gitters und  $m_c^*$  die effektive Leitfähigkeitsmasse des Halbleiters ist. Wie man in Gl. (1.49) sieht, nimmt die Beweglichkeit bei dominanter Phononenstreuung mit steigender Temperatur und steigender effektiver Masse ab. Die Beweglichkeit durch ionisierte Störstellen  $\mu_i$  kann durch die Formel

$$\mu_i = \frac{64\sqrt{\pi}\epsilon_s^2 (2kT)^{3/2}}{N_I q^3 m^{*1/2}} \left\{ \ln \left[ 1 + \left( \frac{12\pi\epsilon_s kT}{q^2 N_I^{1/3}} \right)^2 \right] \right\}^{-1} \propto \frac{T^{3/2}}{N_I m^{*1/2}} \quad (1.46)$$

beschrieben werden [40], wobei  $N_I$  die Dichte der ionisierten Störstellen ist. Bei dominanter Störstellenstreuung nimmt die Beweglichkeit mit der effektiven Masse ebenfalls ab, aber im Gegensatz zur Phononenstreuung nimmt die Beweglichkeit bei dominanter Störstellenstreuung bei steigender Temperatur zu, weil Ladungsträger mit höherer thermischer Geschwindigkeit durch die Coulomb-Streuung weniger abgelenkt werden. Die kombinierte Beweglichkeit, die man aus beiden oben genannten Streumechanismen bekommt, ist durch die Matthiessen-Regel gegeben:

$$\mu = \left( \frac{1}{\mu_l} + \frac{1}{\mu_i} \right)^{-1}. \quad (1.47)$$

Zusätzlich zu den oben diskutierten Streumechanismen wirken sich auch andere Mechanismen auf die tatsächliche Beweglichkeit aus. Zum Beispiel (1) die „Intravalley“-Streuung (also innerhalb eines Tales im Leitungsband), bei der ein Elektron innerhalb eines Energieellipsoids gestreut wird (Abb. 1.5) und nur langwellige Phononen (akustische Phononen) beteiligt sind, und (2) die „Intervalley“-Streuung (Streuung zwischen den Tälern), bei der ein Elektron aus der Nähe eines Leitungsbandminimums zu einem anderen Minimum im Leitungsband gestreut wird. Bei diesem Prozess wird aber ein optisches Phonon benötigt, das dann für die Energie- und Impulserhaltung sorgt. Diese polar-optische Phononenstreuung ist besonders bei polaren Halbleitern wie GaAs von Bedeutung.

Da die Beweglichkeit durch Streuung kontrolliert wird, kann sie auch mit der mittleren Lebensdauer  $\tau_m$  oder der mittleren freien Weglänge  $\lambda_m$  in Beziehung gesetzt werden:

$$\mu = \frac{q\tau_m}{m^*} = \frac{q\lambda_m}{\sqrt{3kTm^*}}. \quad (1.48)$$

Für die mittlere freie Weglänge gilt die Beziehung  $\lambda_m = v_{th} \tau_m$ , wobei  $v_{th}$  die thermische Geschwindigkeit ist, die durch die Formel

$$v_{th} = \sqrt{\frac{3kT}{m^*}} \quad (1.49)$$

gegeben ist. Für multiple Streumechanismen wird die effektive mittlere Lebensdauer aus den einzelnen mittleren Lebensdauern von Streueignissen durch die Gleichung

$$\frac{1}{\tau_m} = \frac{1}{\tau_{m1}} + \frac{1}{\tau_{m2}} + \dots \quad (1.50)$$

bestimmt, womit wir dann als Verallgemeinerung von Gl. (1.47) den Ausdruck  $\mu = (1/\mu_{m1} + 1/\mu_{m2} + 1/\mu_{m2} \dots)^{-1}$  bekommen.

Abbildung 1.15 zeigt die gemessenen Beweglichkeiten von Si und GaAs als Funktion der Störstellenkonzentration bei  $T = 300$  K. Mit zunehmender Störstellenkonzentration (bei Raumtemperatur sind die meisten flachen Störstellen ionisiert) nimmt die

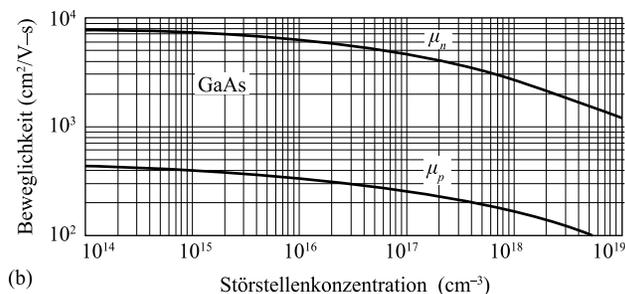
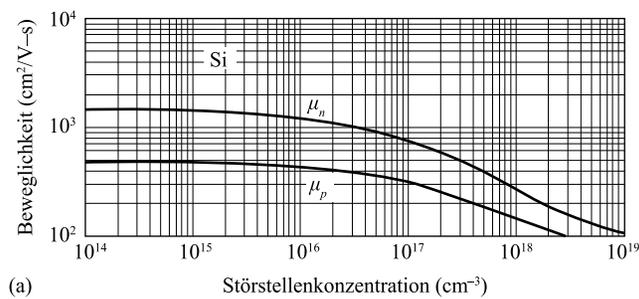
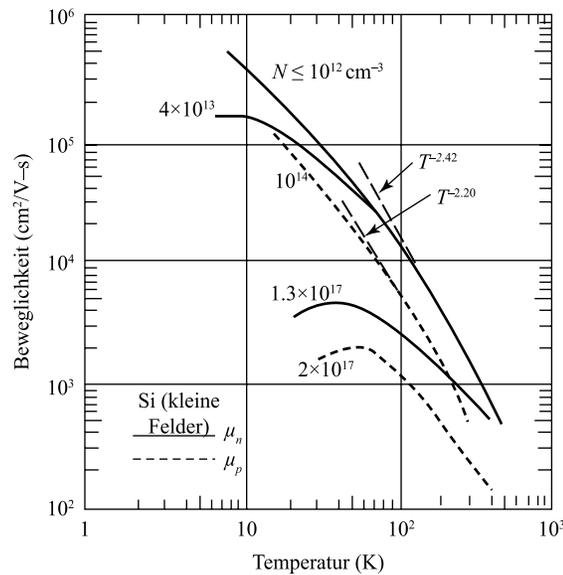


Abb. 1.15 Driftbeweglichkeit von (a) Si (basierend auf [8, 41]) und (b) GaAs bei  $T = 300$  K als Funktion der Störstellenkonzentrationen (nach [8, 12]).



**Abb. 1.16** Beweglichkeiten von Elektronen und Löchern in Si als Funktion der Temperatur (basierend auf [42]).

Beweglichkeit, wie durch Gl. (1.50) vorhergesagt, ab. Auch für größere effektive Massen  $m^*$  nimmt  $\mu$  ab; daher sind für eine gegebene Störstellenkonzentration auch die Elektronenbeweglichkeiten in Si und GaAs größer als die Löcherbeweglichkeiten in diesen Materialien. Die detaillierten Werte für die effektiven Massen finden sich in den Anhängen F und G.

Abbildung 1.16 zeigt den Temperatureinfluss auf die Beweglichkeit für  $n$ - und  $p$ -Typ-Silizium. Für niedrigere Störstellenkonzentrationen ist die Beweglichkeit durch die Phononenstreuung begrenzt und nimmt, wie durch Gl. (1.45) vorhergesagt, mit steigender Temperatur ab. Die gemessenen Steigungen haben aber aufgrund anderer Streuungsmechanismen nicht mehr den Wert  $-3/2$ . Im Bereich der Raumtemperatur variiert die Beweglichkeit für diese reinen Materialien mit  $T^{-2.42}$  und  $T^{-2.20}$  für  $n$ - und  $p$ -Typ-Silizium und mit  $T^{-1.0}$  und  $T^{-2.1}$  für  $n$ - und  $p$ -Typ-GaAs (nicht dargestellt).

Abbildung 1.17 zeigt die Elektronen- und Löcherbeweglichkeiten der gängigen Halbleiter bei Raumtemperatur und niedrigen elektrischen Feldern. Sowohl die Elektronen- als auch die Löcherbeweglichkeiten nehmen mit zunehmender Bandlücke  $E_g$  ab. Die Beweglichkeit von Materialien mit direkter Bandlücke ist relativ gesehen größer als die von Materialien mit indirekter Bandlücke; insbesondere ist die Elektronenbeweglichkeit von GaAs sechsmal größer als die von Si, sodass GaAs gerne für Hochfrequenzbauelementtechnologien, wie z. B. Transistoren mit hoher Elektronenbeweglichkeit, verwendet wird (siehe Kap. 8).

Hinweis: Die oben diskutierten Beweglichkeiten sind die Leitfähigkeitsbeweglichkeiten, die nachweislich den Driftbeweglichkeiten entsprechen [35]. Sie unterscheiden sich jedoch von den im nächsten Abschnitt betrachteten Hall-Beweglichkeiten, hängen aber mit diesen direkt zusammen.

### 1.5.2 Spezifischer Widerstand und Hall-Effekt

Bei Halbleitern mit Elektronen und Löchern ist der Driftstrom bei einem angelegten elektrischen Feld gegeben durch  $J = \sigma \mathcal{E} = q(\mu_n n + \mu_p p)\mathcal{E}$ , wobei  $\sigma = 1/\rho = q(\mu_n n + \mu_p p)$  die Leitfähigkeit und  $\rho$  der spezifische Widerstand ist. Wenn wie bei  $n$ -Typ-Halbleitern  $n \gg p$  ist, ist die Formel für den spezifischen Widerstand

$$\rho = \frac{1}{q\mu_n n} \quad (1.51)$$

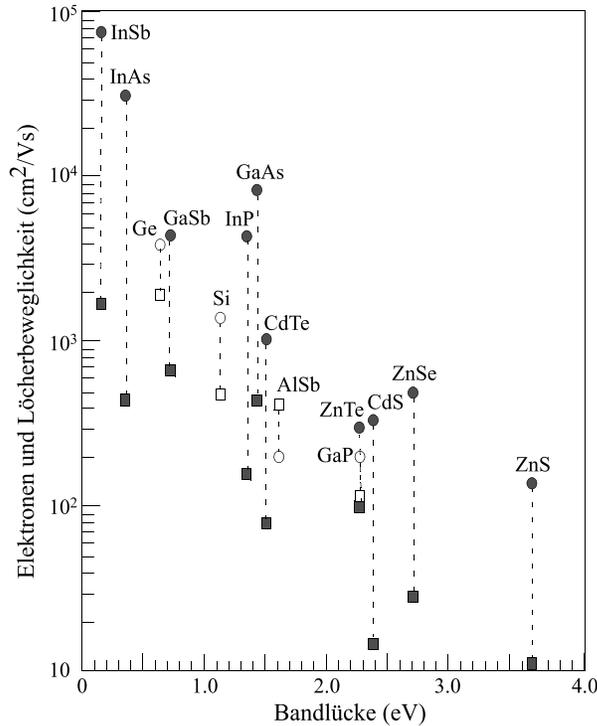


Abb. 1.17 Elektronen- (Kreise) und Löcherbeweglichkeiten (Quadrate) als Funktion der Bandlücke für direkte (gefüllte Datenpunkte) und indirekte Halbleiter (offene Datenpunkte) (nach [43, 44]).

und für die Leitfähigkeit

$$\sigma = q\mu_n n . \tag{1.52}$$

Die gebräuchlichste Methode zur Messung des spezifischen Widerstandes ist eine Vier-Punkt-Messung (kleines Bild in Abb. 1.18) [45, 46]. Ein kleiner, konstanter Strom wird durch die beiden äußeren Kontakte eingepreßt und die Spannung zwischen den beiden inneren Kontakten wird gemessen. Für einen dünnen Wafer mit einer Dicke  $W$ , die viel kleiner ist als  $a$  oder  $d$ , ist der Schichtwiderstand  $R_{\square}$  durch die Formel

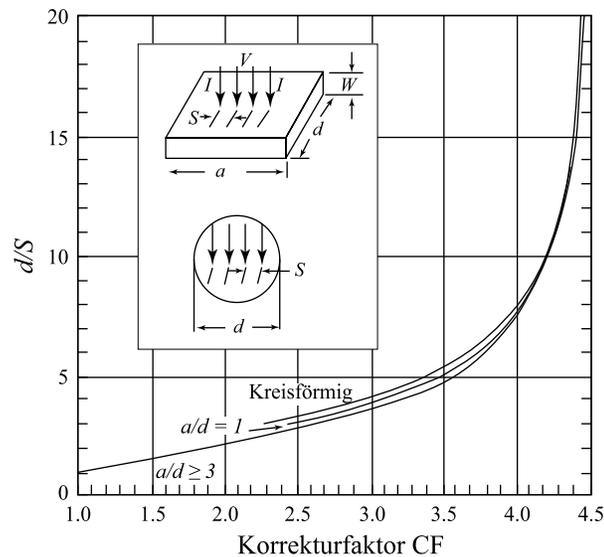
$$R_{\square} = \frac{V}{I} \cdot CF \quad \Omega/\square \tag{1.53a}$$

gegeben, wobei  $CF$  der in Abb. 1.18 dargestellte Korrekturfaktor ist. Der spezifische Widerstand beträgt dann

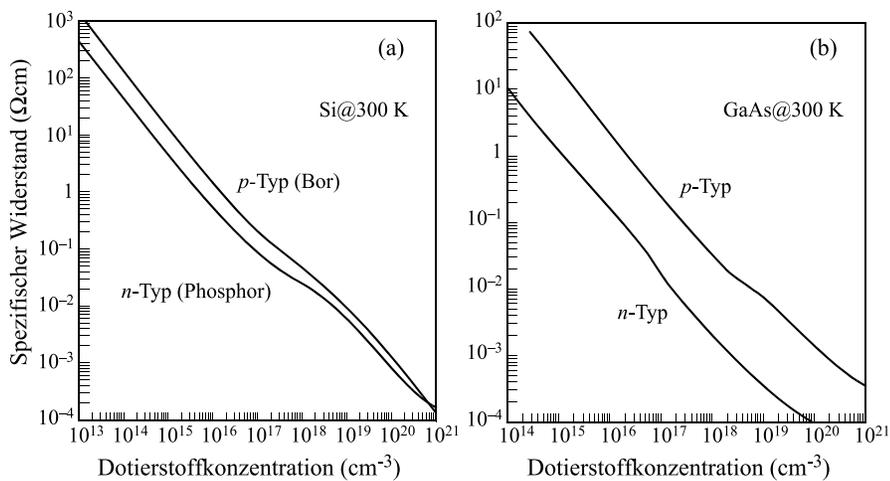
$$\rho = R_{\square} W \quad \Omega\text{-cm} . \tag{1.53b}$$

Im Grenzfall von  $d \gg S$ , wobei  $S$  der Abstand zwischen den Kontakten ist, wird der Korrekturfaktor zu  $\pi / \ln 2$  ( $= 4.54$ ).

Abbildung 1.19a zeigt den gemessenen spezifischen Widerstand ( $T = 300\text{ K}$ ) als Funktion der Störstellenkonzentration ( $n$ -Typ, Phosphor und  $p$ -Typ, Bor) für Silizium. Der Widerstand ist keine lineare Funktion der Konzentration, da die Beweglichkeit nicht konstant ist und normalerweise mit zunehmender Konzentration abnimmt. Abbildung 1.19b zeigt die gemessenen Widerstände für GaAs. Aus dieser Abbildung können wir somit die Störstellenkonzentration eines Halbleiters erhalten, wenn sein spezifischer Widerstand bekannt ist, und umgekehrt. Beachten Sie, dass sich die Störstellenkonzentration aufgrund unvollständiger Ionisierung von der Ladungsträgerkonzentration unterscheiden kann. In  $p$ -Typ-Silizium mit  $10^{17}\text{ cm}^{-3}$  Galliumstörstellen (Akzeptoren) haben beispielsweise die nicht ionisierten Akzeptoren bei Raumtemperatur einen Anteil von etwa 23 %, aus Gl. (1.32), Abb. 1.10 und 1.14; mit anderen Worten, die Trägerkonzentration beträgt nur  $7.7 \times 10^{16}\text{ cm}^{-3}$ .



**Abb. 1.18** Korrekturfaktor für die Messung des spezifischen Widerstandes mit der Vier-Punkt-Methode ([45], mit freundlicher Genehmigung von John Wiley & Sons).



**Abb. 1.19** Spezifischer Widerstand als Funktion der Störstellenkonzentration bei  $T = 300\text{ K}$  für (a) Silizium und (b) GaAs (nach [36, 41]).

**Hall-Effekt** Die Messung des Widerstandes gibt nur das Produkt der Beweglichkeit und der Trägerkonzentration an. Um jeden Parameter direkt zu messen, nutzt die gebräuchlichste Methode den Hall-Effekt [47]. Er ist auch heute noch eines der faszinierendsten Phänomene im Bereich der Halbleiter und ist sowohl grundsätzlich als auch praktisch interessant. Jüngste Beispiele sind die Untersuchung des fraktionierten Quanten-Hall-Effekts und die Anwendungen in Magnetfeldsensoren. Der Hall-Effekt wird in der gängigen Praxis zur Messung der Ladungsträgerkonzentration, auch bei niedrigen Niveaus in der Größenordnung von  $10^{12}\text{ cm}^{-3}$ , und zur Bestimmung der Beweglichkeit und des Typs ( $n$  oder  $p$ ) der Ladungsträger verwendet. Er ist ein wichtiges Analysewerkzeug, da bei einer einfachen Leitwertmessung nur das Produkt von Konzentration und Beweglichkeit geliefert wird und die Art der Ladungsträger unbekannt bleibt.

Abbildung 1.20 zeigt den grundlegenden Aufbau der Messung, bei der ein elektrisches Feld entlang der  $x$ -Achse und ein magnetisches Feld entlang der  $z$ -Achse angelegt wird [48]. Betrachten Sie eine  $p$ -Typ-Probe. In diesem Fall ist die Lorentz-Kraft eine nach unten gerichtete Kraft auf die Löcher. Es gilt die Formel Lorentz-Kraft =

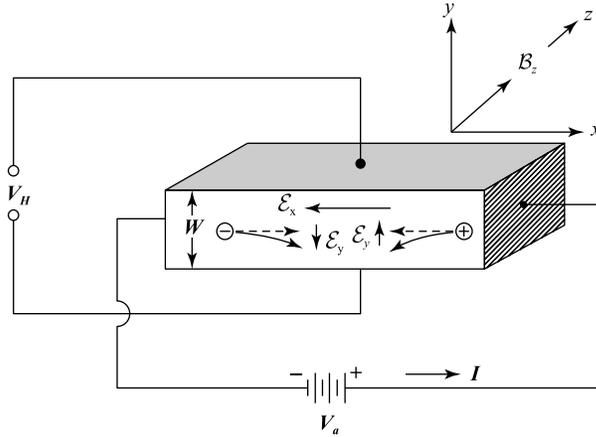


Abb. 1.20 Aufbau einer Hall-Messung zur Bestimmung der Ladungsträgerkonzentration.

$qv_x \times B_z$ , wobei der nach unten gerichtete Strom eine Anhäufung von Löchern an der Unterseite der Probe bewirkt, die wiederum ein elektrisches Feld  $\mathcal{E}_y$  erzeugt. Da es im stationären Zustand keinen Nettostrom entlang der  $y$ -Richtung gibt, gleicht das elektrische Feld entlang der  $y$ -Achse (Hall-Feld) die Lorentz-Kraft genau aus, sodass sich die Ladungsträger parallel zum angelegten Feld  $\mathcal{E}_x$  bewegen. Bei  $n$ -leitenden Materialien häufen sich die Elektronen an der Unterseite an und bauen damit eine Spannung entgegengesetzter Polarität auf. Die Trägergeschwindigkeit  $v$  hängt mit der Stromdichte über die Formel  $J_x = qv_x p$  zusammen. Für jeden Ladungsträger ist die Lorentz-Kraft gleich der Kraft, welche das Hall-Feld ausübt:

$$q\mathcal{E}_y = qv_x B_z. \quad (1.54a)$$

Die dadurch entstandene Hall-Spannung kann extern gemessen werden und ist gegeben durch:

$$V_H = \mathcal{E}_y W = \frac{J_x B_z W}{qp}. \quad (1.54b)$$

Wenn man Streuung berücksichtigen will, wird die Hall-Spannung zu  $V_H = R_H J_x B_z W$ , wobei  $R_H$  der Hall-Koeffizient ist:

$$R_H = \frac{r_H}{qp} \quad p \gg n, \quad (1.55a)$$

$$R_H = -\frac{r_H}{qn} \quad n \gg p. \quad (1.55b)$$

Der Hall-Faktor  $r_H$  berücksichtigt die mittleren Streuzeiten und lautet:

$$r_H \equiv \frac{\langle \tau_m^2 \rangle}{\langle \tau_m \rangle^2}. \quad (1.56)$$

Mit obigem Formelwerk können die Ladungsträgerkonzentration und der Ladungsträgertyp (Elektronen oder Löcher) aus der Hall-Spannung und deren Polarität gewonnen werden, vorausgesetzt, dass ein Ladungsträgertyp dominiert und  $r_H$  bekannt ist.

Die Gln. (1.55a) und (1.55b) gehen jeweils von der Stromleitung durch einen einzigen Trägertyp aus. Eine verallgemeinerte Lösung für beide Typen von Ladungsträgern ist gegeben durch [6]:

$$R_H = \frac{r_H}{q} \frac{\mu_p^2 p - \mu_n^2 n}{(\mu_p p + \mu_n n)^2}. \quad (1.57)$$

In Gl. (1.57) ist zu erkennen, dass das Vorzeichen von  $R_H$  und damit  $V_H$  den Typ der Majoritätsladungsträger der Halbleiterprobe anzeigt. Die Hall-Beweglichkeit  $\mu_H$

ist definiert als das Produkt aus dem Hall-Koeffizienten und der Leitfähigkeit  $\mu_H = |R_H|\sigma$ . Die Hall-Beweglichkeit unterscheidet sich von der Driftbeweglichkeit  $\mu_n$  (oder  $\mu_p$ ) lediglich durch den Hall-Faktor  $r_H$ . Der Zusammenhang zwischen der Driftbeweglichkeit und der Hall-Beweglichkeit ist damit durch  $\mu_H = r_H\mu$  hergestellt. Der Parameter  $\tau_m$  im Hall-Faktor ist die mittlere freie Zeit zwischen Ladungsträgerkollisionen, die auch gerne Streuzzeit genannt wird. Vorsicht:  $\tau_m$  hängt von der Energie ab. Für Halbleiter mit kugelförmigen Oberflächen konstanter Energie gilt z. B., dass  $\tau_m \propto E^{-1/2}$  für Phononenstreuung und  $\tau_m \propto E^{3/2}$  für ionisierte Störstellenstreuung ist. Im Allgemeinen gilt  $\tau_m = C_1 E^{-s}$ , wobei  $C_1$  und  $s$  Konstanten sind. Aus der Boltzmann-Verteilung für nicht entartete Halbleiter ist der Durchschnittswert der  $n$ -ten Potenz von  $\tau_m$ :

$$\langle \tau_m^n \rangle = \frac{\int_0^\infty \tau_m^n E^{3/2} \exp\left(-\frac{E}{kT}\right) dE}{\int_0^\infty E^{3/2} \exp\left(-\frac{E}{kT}\right) dE}, \quad (1.58)$$

sodass wir mit der allgemeinen Form von  $\tau_m$  die Beziehungen

$$\langle \tau_m^2 \rangle = \frac{C_1^2 (kT)^{-2s} \Gamma\left(\frac{5}{2} - 2s\right)}{\Gamma\left(\frac{5}{2}\right)} \quad (1.59)$$

und

$$\langle \tau_m \rangle = \frac{C_1 (kT)^{-s} \Gamma\left(\frac{5}{2} - s\right)}{\Gamma\left(\frac{5}{2}\right)} \quad (1.60)$$

erhalten, wobei  $\Gamma(n)$  eine Gammafunktion ist, die über  $\Gamma(n) \equiv \int_0^\infty x^{n-1} e^{-x} dx$  definiert ist, und für die  $\Gamma(n+1) = n\Gamma(n)$ ,  $\Gamma(1/2) = \sqrt{\pi}$  gilt. Aus dem obigen Ausdruck bekommen wir  $r_H = 3\pi/8 = 1.18$  für die Phononenstreuung und  $r_H = 315\pi/512 = 1.93$  für die Streuung an ionisierten Störstellen. Im Allgemeinen liegt  $r_H$  im Bereich von 1–2. Bei sehr hohen Magnetfeldern nähert sich  $r_H$  einem Wert knapp unter eins.

In der vorhergehenden Diskussion wurde angenommen, dass das angelegte Magnetfeld so klein ist, dass im Magnetfeld keine Veränderung im spezifischen Widerstand der Probe auftritt. Unter starken Magnetfeldern wird jedoch eine signifikante Erhöhung des spezifischen Widerstandes beobachtet, der sogenannte Magnetowiderstandseffekt, der dadurch entsteht, dass sich die Ladungsträger auf einer Bahn bewegen, die von der Richtung des angelegten elektrischen Feldes abweicht. Für sphärische Oberflächen mit konstanter Energie im  $\mathbf{k}$ -Raum ist das Verhältnis der Änderung des spezifischen differentiellen Widerstands im Magnetfeld zum spezifischen Gesamtwiderstand ohne Magnetfeld gegeben durch [2]:

$$\frac{\Delta\rho}{\rho_0} = \left\{ \left[ \frac{\Gamma^2\left(\frac{5}{2}\right)\Gamma\left(\frac{5}{2} - 3s\right)}{\Gamma^3\left(\frac{5}{2} - s\right)} \right] \left( \frac{\mu_n^3 n + \mu_p^3 p}{\mu_n n + \mu_p p} \right) - \left[ \frac{\Gamma\left(\frac{5}{2}\right)\Gamma\left(\frac{5}{2} - 2s\right)}{\Gamma^2\left(\frac{5}{2} - s\right)} \right]^2 \left( \frac{\mu_n^2 n - \mu_p^2 p}{\mu_n n + \mu_p p} \right)^2 \right\} \mathcal{B}_z^2. \quad (1.61)$$

Das Verhältnis ist proportional zum Quadrat der Magnetfeldkomponente senkrecht zur Richtung des Stromflusses. Für  $n \gg p$  ist  $(\Delta\rho/\rho_0) \propto \mu_n^2 \mathcal{B}_z^2$ . Ein ähnliches Ergebnis kann für den Fall  $p \gg n$  erzielt werden.

### 1.5.3 Transport bei hohen elektrischen Feldern

In den vorhergehenden Abschnitten haben wir uns um den Transport von Ladungsträgern in Halbleitern bei niedrigen elektrischen Feldern gekümmert. In diesem Abschnitt behandeln wir kurz einige besondere Effekte im Elektronentransport bei mittleren und hohen elektrischen Feldern.

Wie in Abschn. 1.5.1 gezeigt, hängt bei niedrigen elektrischen Feldern die Driftgeschwindigkeit über die Beweglichkeit linear mit dem elektrischen Feld im Halbleiter zusammen. Bei niedrigen Feldern ist die Beweglichkeit konstant. Bei höheren elektrischen Feldern werden jedoch Nichtlinearitäten in der Beweglichkeit und in einigen Fällen sogar eine Sättigung der Driftgeschwindigkeit beobachtet. Bei noch größeren Feldern kann Stoßionisation auftreten. Betrachten wir zunächst die nicht lineare Beweglichkeit.

Im thermischen Gleichgewicht ist die Energieverteilung der Ladungsträger eine Maxwell-Boltzmann-Verteilung, und die Ladungsträger emittieren und absorbieren die Phononen mit einer Nettoenergieaustauschrate von null. In Gegenwart eines moderaten elektrischen Feldes nehmen die Ladungsträger Energie aus dem Feld auf und verlieren sie wieder an akustische Phononen. Im Durchschnitt werden jetzt zwar mehr Phononen emittiert als absorbiert, aber dennoch haben die Ladungsträger im Durchschnitt mehr Energie als im thermischen Gleichgewicht. Wenn das Feld zunimmt, steigt auch die durchschnittliche Energie der Ladungsträger. Dieser Energie kann man eine effektive Temperatur  $T_e$  zuordnen, die über der Gittertemperatur  $T$  liegt. Wenn die Rate, mit der die Energie vom Feld auf die Ladungsträger übertragen wird, durch die Rate des Energieverlusts an das Gitter ausgeglichen wird, erhält man aus der Ratengleichung für Si (und andere Halbleiter ohne Elektronentransfer in andere Täler, also ohne „transferred electron effect“ [1]):

$$\frac{T_e}{T} = \frac{1}{2} \left[ 1 + \sqrt{1 + \frac{3\pi}{8} \left( \frac{\mu_0 \mathcal{E}}{c_s} \right)^2} \right] \quad (1.62)$$

und

$$v_d = \mu_0 \mathcal{E} \sqrt{\frac{T}{T_e}}, \quad (1.63)$$

wobei  $\mu_0$  die Beweglichkeit im niedrigen Feld und  $c_s$  die Schallgeschwindigkeit ist. Bei mäßiger Feldstärke, also wenn  $c_s$  mit  $\mu_0 \mathcal{E}_s$  vergleichbar ist, weicht die Trägergeschwindigkeit  $v_d$  von der linearen Abhängigkeit vom angelegten Feld um den Faktor  $\sqrt{T/T_e}$  ab. Schließlich treten dann bei ausreichend hohen Feldern Streuprozesse der Ladungsträger mit optischen Phononen auf und Gl. (1.62) ist nicht mehr korrekt. Stattdessen hängen die Driftgeschwindigkeiten für Ge und Si in diesem Bereich immer weniger vom angelegten Feld ab und nähern sich einer Sättigungsgeschwindigkeit von

$$v_s = \sqrt{\frac{8E_p}{3\pi m_0}} \approx 10^7 \text{ cm/s}, \quad (1.64)$$

wobei  $E_p$  die Energie der optischen Phononen im jeweiligen Material ist (ca. 30–50 meV, für genaue Werte siehe Anhang H).

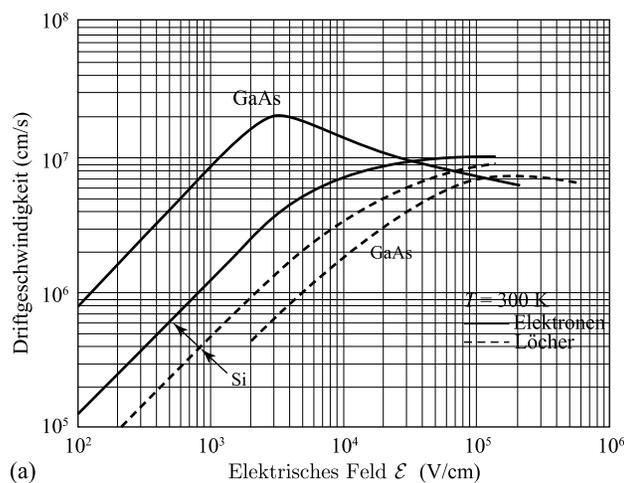
Um Lücken zwischen den von Gln. (1.62)–(1.64) abgedeckten Bereichen zu vermeiden, wird oft eine einzige empirische Formel verwendet, die das Verhalten im gesamten Bereich zwischen der Driftgeschwindigkeit bei geringem Feld und der Sättigungsgeschwindigkeit beschreibt [49]:

$$v_d = \frac{\mu_0 \mathcal{E}}{[1 + (\mu_0 \mathcal{E}/v_s)c_2]^{1/c_2}}. \quad (1.65)$$

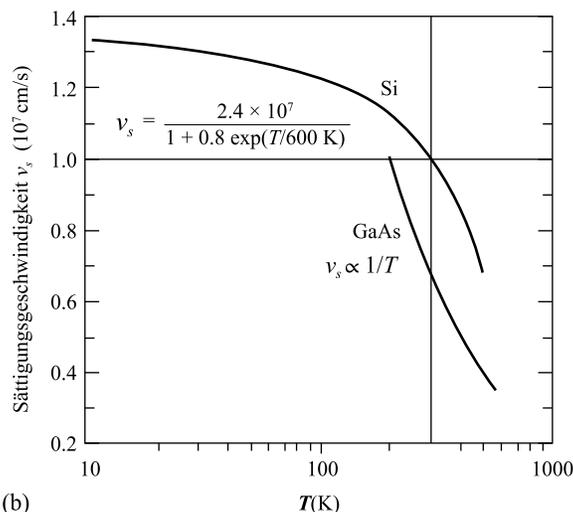
Die Konstante  $C_2$  ist eine Funktion der Temperatur und hat einen Wert von ca. zwei für Elektronen und eins für Löcher.

Die Geschwindigkeits-Feld-Beziehung bei GaAs ist komplizierter, und wir müssen die Bandstruktur berücksichtigen. Wie man in Abb. 1.4 sieht, befindet sich ein Tal mit hoher Beweglichkeit ( $\mu \approx 4000\text{--}8000\text{ cm}^2/\text{V}\cdot\text{s}$ ) im Zentrum der Brillouin-Zone, und ein niederbewegliches Satellitaltal ( $\mu \approx 100\text{ cm}^2/\text{V}\cdot\text{s}$ ) entlang der  $\langle 111 \rangle$ -Achsen [50], das auf der Energieskala etwa 0.3 eV oberhalb des Leitungsbandminimums liegt. Der Unterschied in der Beweglichkeit ist auf die unterschiedlichen effektiven Massen der Elektronen (Gl. (1.48)) zurückzuführen:  $0.063m_0$  im unteren Tal und etwa  $0.55m_0$  im oberen Tal. Wenn das Feld zunimmt, können die Elektronen im unteren Tal in das normalerweise unbesetzte obere Tal gestreut werden, was zu einem negativen differentiellen Widerstand im GaAs führt. Der Transfermechanismus zwischen den verschiedenen Tälern im Leitungsband (engl. transferred electron effect) und die zugehörige Geschwindigkeits-Feld-Beziehung werden in Kap. 10 näher diskutiert.

Abbildung 1.21a zeigt die bei Raumtemperatur gemessenen Driftgeschwindigkeiten für hochreines (niedrige Dotierstoffkonzentration  $< 10^{15}\text{ cm}^{-3}$ ) Si und GaAs als Funktion des elektrischen Feldes. Bei niedrigen Feldern und hoher Störstellenkonzentration wird die Driftgeschwindigkeit und die Beweglichkeit aufgrund der Streuung mit den Störstellen verringert. Die Geschwindigkeit bei hohen Feldern ist jedoch im Wesentlichen unabhängig von der Störstellenkonzentration und nähert sich einem Sätti-



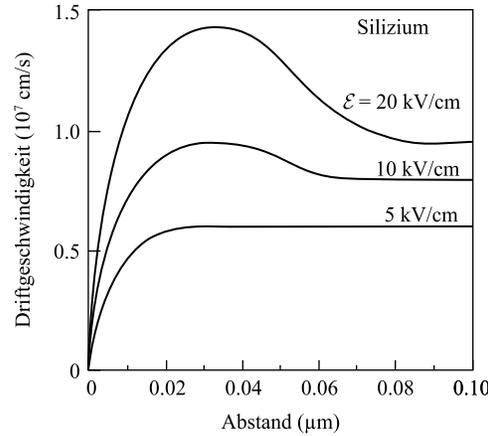
(a)



(b)

**Abb. 1.21** (a) Gemessene Ladungsträgergeschwindigkeit als Funktion des elektrischen Feldes für hochreines Si und GaAs. Bei hochdotierten Proben sind die niedrigen Geschwindigkeiten (Beweglichkeiten) geringer als hier angegeben. Im Hochfeldbereich ist die Geschwindigkeit jedoch im Wesentlichen unabhängig von den Dotierungen (nach [42, 51–53]). (b) Sättigungsgeschwindigkeit der Elektronen in Abhängigkeit von der Temperatur in Si und GaAs (nach [42, 54]).

Abb. 1.22 „Velocity overshoot“ in ultrakurzen Entfernungen vom Startpunkt. Für ultrakurze Zeiträume nach dem Einschalten des elektrischen Feldes sieht das Bild sehr ähnlich aus. Das Beispiel gilt für Silizium (verändert nach [56]).



gungswert [55]. Für Si betragen die Sättigungsgeschwindigkeiten für Elektronen und Löcher etwa  $1 \times 10^7$  cm/s. Für GaAs gibt es einen weiten Bereich mit negativer differentieller Beweglichkeit für Felder über  $3 \times 10^3$  V/cm, und die Sättigungsgeschwindigkeit bei hohen Feldern nähert sich dann dem Wert von  $6 \times 10^6$  cm/s. Abbildung 1.21b zeigt die Temperaturabhängigkeit der Sättigungsgeschwindigkeit für Elektronen. Mit steigender Temperatur nimmt die Sättigungsgeschwindigkeit sowohl für Si als auch für GaAs ab.

Die bisher diskutierten Driftgeschwindigkeiten gelten für stationäre Zustände, in denen die Ladungsträger genügend Streueignisse durchlaufen, um ihren Gleichgewichtszustand zu erreichen. Bei modernen Bauelementen wird die kritische Länge, über welche die Ladungsträger durchgeschaltet werden, immer kleiner. Wenn diese Länge mit der mittleren freien Weglänge vergleichbar oder kürzer wird, treten ballistische Transporteffekte auf.

Abbildung 1.22 zeigt die Driftgeschwindigkeit der Ladungsträger in Abhängigkeit der zurückgelegten Distanz. Ohne Streuung nimmt die Geschwindigkeit mit der Zeit (und der Entfernung) gemäß  $\approx q\mathcal{E}t/m^*$  zu. Bei hohen Feldern kann die Driftgeschwindigkeit auf kurzen Distanzen (in der Größenordnung der mittleren freien Weglänge) und in kurzen Zeiträumen (in der Größenordnung der mittleren Lebensdauer) einen höheren Wert erreichen als im stationären Zustand. Dieses Phänomen kann im Deutschen, wie im Straßenverkehr, als Überschreitung der zulässigen Höchstgeschwindigkeit (engl. velocity overshoot) bezeichnet werden. Vorsicht: Hier kann beim Studium der weiterführenden Literatur Verwirrung entstehen, da die in Abb. 1.21a gezeigte Maximalgeschwindigkeit, welche im GaAs durch den Transfer von Ladungsträgern in höhere Täler entsteht, ebenfalls als „velocity overshoot“ bezeichnet wird. Bei niedrigen Feldern ist die Beschleunigung der Ladungsträger gering und die erreichte Geschwindigkeit zu niedrig, als dass es innerhalb der Streuzzeit zu einem „velocity“ overshoot kommen kann.

Betrachten wir als Nächstes die Stoßionisation (engl. impact ionization). Wird das elektrische Feld in einem Halbleiter über einen bestimmten Wert hinaus erhöht, gewinnen die Ladungsträger genügend Energie, um Elektronen-Loch-Paare durch Stoßionisation zu erzeugen. Die kinetische Minimalenergie der Ladungsträger muss für diesen Prozess natürlich größer als die Bandlücke sein. Dieser Multiplikationsprozess der Ladungsträger ist durch eine Ionisierungsrate  $\alpha$  gekennzeichnet, die als die Anzahl der (sekundären) Elektron-Loch-Paare definiert ist, die von einem primären Ladungsträger (Elektron oder Loch) pro Längeneinheit erzeugt werden (Abb. 1.23). Für ein Primärelektron mit einer Geschwindigkeit  $v_n$  gilt also

$$\alpha_n = \frac{1}{n} \frac{dn}{d(tv_n)} = \frac{1}{nv_n} \frac{dn}{dt} \tag{1.66}$$

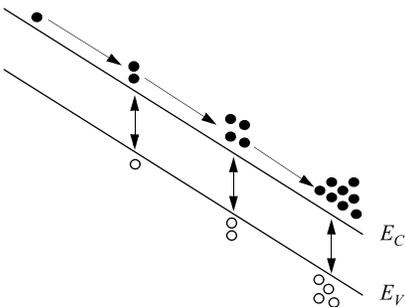


Abb. 1.23 Multiplikation von Elektronen und Löchern durch Stoßionisation, in diesem Beispiel aufgrund von Elektronen ( $\alpha_p = 0$ ).

Unter Berücksichtigung sowohl von Elektronen als auch von Löchern ist die Erzeugungsrate durch Stoßionisation an jedem festen Ort gegeben durch

$$\begin{aligned} \frac{dn}{dt} &= \frac{dp}{dt} = \alpha_n n v_n + \alpha_p p v_p \\ &= \frac{\alpha_n J_n}{q} + \frac{\alpha_p J_p}{q}. \end{aligned} \quad (1.67)$$

Umgekehrt variiert die Trägerdichte oder der Strom zu jedem Zeitpunkt mit der Entfernung und folgt den folgenden Differenzialgleichungen:

$$\frac{dJ_n}{dx} = \alpha_n J_n + \alpha_p J_p, \quad (1.68a)$$

$$\frac{dJ_p}{dx} = -\alpha_n J_n - \alpha_p J_p. \quad (1.68b)$$

Die Gesamtstromdichte für Elektronen und Löcher ( $J_n + J_p$ ) bleibt als Funktion der Entfernung konstant und es gilt  $dJ_n/dx = -dJ_p/dx$ .

Die Ionisierungsraten  $\alpha_n$  und  $\alpha_p$  sind stark vom elektrischen Feld abhängig. Die Ionisierungsrate ist gegeben durch [57]

$$\alpha(\mathcal{E}) = \frac{q\mathcal{E}}{E_I} \exp\left\{-\frac{\mathcal{E}_I}{\mathcal{E}[1 + (\mathcal{E}/\mathcal{E}_p)] + \mathcal{E}_T}\right\}, \quad (1.69)$$

wobei  $E_I$  die effektive Ionisationsenergie bei hohen Feldern ist und  $\mathcal{E}_T$ ,  $\mathcal{E}_p$  und  $\mathcal{E}_I$  sind die minimalen elektrischen Felder, bei denen Ladungsträger die abbremsenden Effekte der thermischen, optischen Phononen- und Streuung an ionisierten Störstellen überwinden können. Für Si wird ein  $E_I$ -Wert von 3.6 eV für Elektronen und 5.0 eV für Löcher gefunden. Über einen begrenzten Feldbereich kann Gl. (1.69) auf folgende Form reduziert werden:

$$\alpha(\mathcal{E}) = \exp\left(-\frac{\mathcal{E}_I}{\mathcal{E}}\right) \quad \text{für } \mathcal{E}_p > \mathcal{E} > \mathcal{E}_T \quad (1.70a)$$

oder

$$\alpha(\mathcal{E}) = \frac{q\mathcal{E}}{E_I} \exp\left(-\frac{\mathcal{E}_I \mathcal{E}_p}{\mathcal{E}^2}\right) \quad \text{für } \mathcal{E} > \mathcal{E}_p \text{ und } \mathcal{E} > \sqrt{\mathcal{E}_p \mathcal{E}_T}. \quad (1.70b)$$

Abbildung 1.24a zeigt die experimentellen Ergebnisse der Ionisierungsraten für Ge, Si, SiC und GaN und Abb. 1.24b die gemessenen Ionisierungsraten von GaAs und einigen anderen binären und ternären Verbindungen. Diese Ergebnisse wurden durch die Verwendung von Photomultiplikationsmessungen (engl. photomultiplication measurements) an  $p$ - $n$ -Übergängen erzielt. Man beachte, dass bei bestimmten Halbleitern, wie z. B. GaAs, die Ionisierungsrate eine Funktion der Kristallorientierung ist. Des Weiteren nimmt die Ionisierungsrate mit zunehmender Bandlücke tendenziell ab, weshalb Materialien mit höheren Bandlücken im Allgemeinen eine höhere Durchbruchspannung haben. Beachten Sie, dass Gl. (1.70) für die meisten der in Abb. 1.24 gezeigten Halbleiter gilt. Die Ausnahmen sind GaAs und GaP, für diese Halbleiter gilt Gl. (1.72). Bei einem gegebenen elektrischen Feld nimmt die Ionisierungsrate mit steigender Temperatur ab. Abbildung 1.25 zeigt als Beispiel die theoretisch vorhergesagten Ionisierungsraten für Elektronen in Silizium zusammen mit den experimentellen Ergebnissen bei drei verschiedenen Temperaturen.

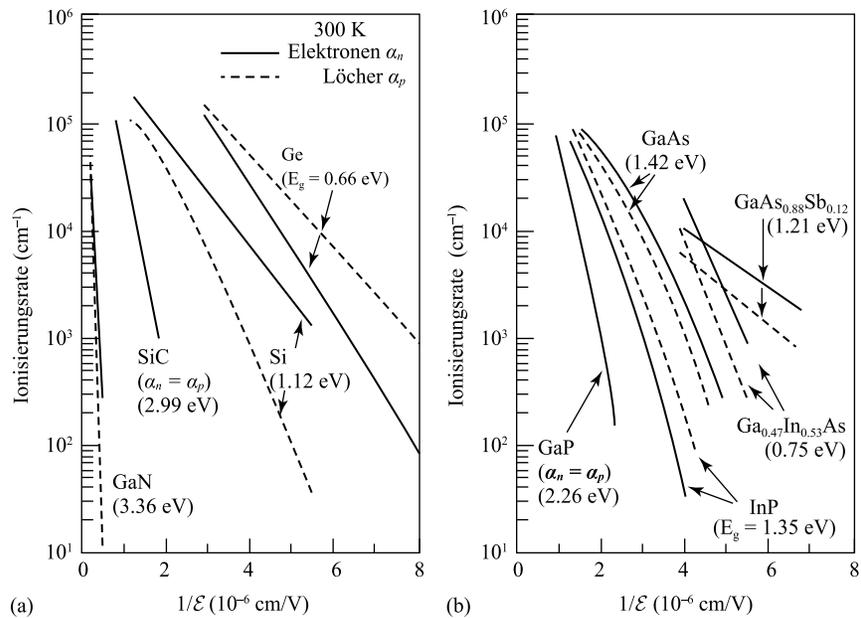


Abb. 1.24 Ionisationsraten bei 300 K in Abhängigkeit vom reziproken elektrischen Feld für Si, GaAs und einige IV-IV- und III-V-Verbindungshalbleiter (nach [58–68]).

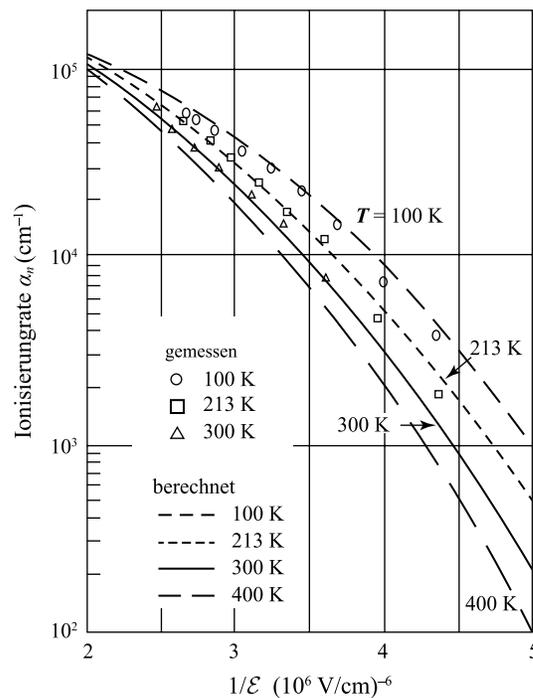
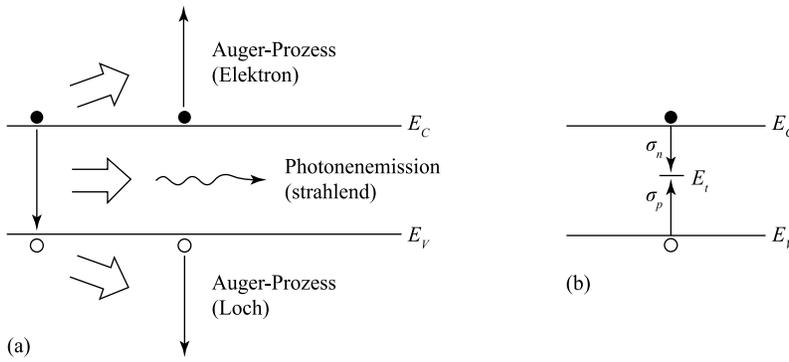


Abb. 1.25 Ionisierungsrate für Elektronen in Abhängigkeit vom reziproken elektrischen Feld in Si für vier verschiedene Temperaturen (verändert nach [69]).

### 1.5.4 Rekombination, Generation und Ladungsträgerlebensdauer

Immer wenn der thermische Gleichgewichtszustand eines Halbleitersystems gestört ist, d. h., wenn  $pn \neq n_i^2$  ist, gibt es Prozesse, die das System wieder ins Gleichgewicht bringen (d. h.,  $pn = n_i^2$ ). Diese Prozesse sind Rekombinationsprozesse, wenn  $pn > n_i^2$  ist, und die thermische Generation, wenn  $pn < n_i^2$  ist. Abbildung 1.26a veranschaulicht eine Band-zu-Band-Elektron-Loch-Rekombination. Die Energie eines Elektrons im Übergang vom Leitungsband zum Valenzband wird durch die Emission eines Photons (Strahlungsprozess) oder durch die Übertragung der Energie auf ein anderes frei-



**Abb. 1.26** Rekombinationsprozesse. (a) Band-zu-Band-Rekombination. Die Energie wird in einem Strahlungs- oder Auger-Prozess umgewandelt. (b) Strahlungsfreie Rekombination über Rekombinationszentren innerhalb der Bandlücke.

es Elektron oder Loch (Auger-Prozess) erhalten. Der erstgenannte Prozess ist die Umkehrung der direkten optischen Absorption, und der zweite Prozess ist die Umkehrung der Stoßionisation.

Band-zu-Band-Übergänge sind häufig in Halbleitern mit direkter Bandlücke (meistens III-V-Verbindungen) zu finden. Bei dieser Art von Übergang ist die Rekombinationsrate proportional zum Produkt der Elektronen- und Löcherkonzentration, die gegeben ist durch

$$R_e = R_{ec}pn. \quad (1.71)$$

Der Begriff  $R_{ec}$ , der als Rekombinationskoeffizient bezeichnet wird, ist mit der thermischen Generationsrate  $G_{th}$  durch die Gleichung  $R_{ec} = G_{th}/n_i^2$  verknüpft.  $R_{ec}$  ist eine Funktion der Temperatur, die auch von der Bandstruktur des Halbleiters abhängt. Ein Halbleiter mit direkter Bandlücke hat ein viel größeres  $R_{ec}$  ( $\approx 10^{-10} \text{ cm}^3/\text{s}$ ) als ein Halbleiter mit indirekter Bandlücke ( $\approx 10^{-15} \text{ cm}^3/\text{s}$ ). Im thermischen Gleichgewicht ist die Nettoübergangsrate  $U = R_e - G_{th}$  gleich null, da  $pn = n_i^2$  ist, und damit  $R_e = G_{th}$  gilt. Bei niedrigen Injektionsraten, also im Fall, dass die Konzentration der überschüssigen Ladungsträger  $\Delta p = \Delta n$  geringer als die Konzentration der Majoritätsladungsträger ist, ist die Nettoübergangsrate bei einem  $n$ -Typ-Material ( $p_n = p_{n0} + \Delta p$  und  $n_n \approx N_D$ ) gegeben durch

$$U = R_e - G_{th} = R_{ec}(pn - n_i^2) \approx R_{ec}\Delta p N_D \equiv \frac{\Delta p}{\tau_p}, \quad (1.72)$$

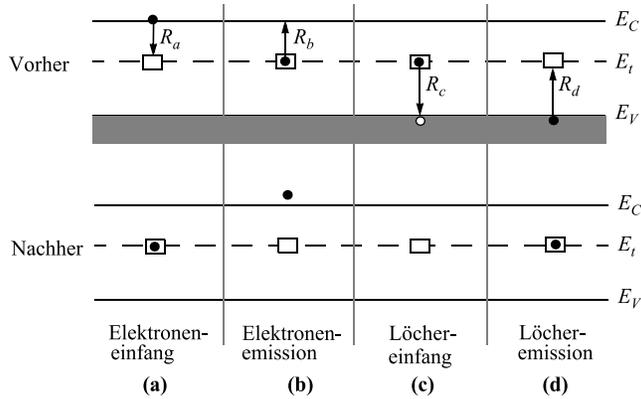
wobei die Lebensdauer für Löcher

$$\tau_p = \frac{1}{R_{ec}N_D} \quad (1.73a)$$

und für Elektronen in  $p$ -Typ-Material

$$\tau_n = \frac{1}{R_{ec}N_A} \quad (1.73b)$$

ist. In Halbleitern mit indirekter Bandlücke, wie Si und Ge, sind die dominierenden Übergänge jedoch die indirekte Rekombination/Generation über Störstellenniveaus, die mit einer gewissen Dichte  $N_t$  und Energie  $E_t$  innerhalb der Bandlücke vorhanden sind (Abb. 1.26b). Die einstufige Rekombination kann durch zwei Prozesse beschrieben werden, nämlich den Elektroneneinfang und den Löchereinfang. Die Nettoübergangsrate kann durch die Shockley-Read-Hall-Statistik beschrieben werden [70–72].



**Abb. 1.27** Indirekte Generations- und Rekombinationsprozesse: (a) Elektroneneinfang, (b) Elektronenemission, (c) Löchereinfang und (d) Löcheremission im thermischen Gleichgewicht.

Verschiedene Übergänge, die bei der Rekombination über Rekombinationszentren auftreten, sind in Abb. 1.27 schematisch dargestellt. Wenn die Konzentration der Zentren im Halbleiter  $N_t$  ist, ist die Konzentration der unbesetzten Zentren durch  $N_t(1 - F)$  gegeben, wobei  $F$  die Fermi-Verteilungsfunktion im Gleichgewicht ist

$$F = \frac{1}{1 + \exp\left(\frac{E_t - E_F}{kT}\right)}. \quad (1.74)$$

$E_t$  ist das Energieniveau des Rekombinationszentrums und  $E_F$  das Fermi-Niveau. Die Einfangrate  $R_a$  eines Elektrons durch ein Rekombinationszentrum (Abb. 1.27a) ist proportional zu

$$R_a \approx nN_t(1 - F). \quad (1.75)$$

Wenn wir der Proportionalitätskonstante einfach mal so und ohne nachzudenken das Produkt  $v_{th}\sigma_n$  zuordnen, bekommen wir für die Einfangrate folgende Formel

$$R_a = v_{th}\sigma_n nN_t(1 - F). \quad (1.76)$$

Das Produkt  $v_{th}\sigma_n$  kann aber auch als das Volumen betrachtet werden, aus welchem ein Elektron von einem Rekombinationszentrum mit einem Einfangquerschnitt  $\sigma_n$  pro Zeiteinheit eingefangen wird. Die Emission von Elektronen aus dem Zentrum (Abb. 1.27b) ist die Umkehrung des Einfangprozesses. Die Emissionsrate ist proportional zur Konzentration der von Elektronen besetzten Zentren, d. h.  $v_s \propto 1/T$ . Damit bekommen wir dann

$$R_b = e_n N_t F. \quad (1.77)$$

Die Proportionalitätskonstante  $e_n$  ist die Emissionswahrscheinlichkeit. Im thermischen Gleichgewicht müssen die Elektroneneinfang- und -Emissionsraten aber gleich sein, d. h.,  $R_a = R_b$ . Die Emissionswahrscheinlichkeit kann also über Gl. (1.76) folgendermaßen ausgedrückt werden:

$$e_n = \frac{v_{th}\sigma_n n(1 - F)}{F}. \quad (1.78)$$

Da die Elektronenkonzentration im thermischen Gleichgewicht durch

$$e_n = v_{th}\sigma_n n_i \exp\left(\frac{E_t - E_i}{kT}\right) \quad (1.79)$$

gegeben ist, bekommen wir

$$e_n = v_{th}\sigma_n n_i \exp\left(\frac{E_t - E_i}{kT}\right). \quad (1.80)$$

Die Übergänge zwischen dem Rekombinationszentrum und dem Valenzband sind analog zu denen, die oben beschrieben wurden. Die Einfangrate eines Lochs durch ein besetztes Rekombinationszentrum (Abb. 1.27c) ist gegeben durch

$$R_c = v_{th}\sigma_p p N_t F, \quad (1.81)$$

wobei  $\sigma_p$  der Einfangquerschnitt für Löcher ist. Mit ähnlichen Argumenten wie bei der Elektronenemission ist die Rate der Löcheremission (Abb. 1.27d)

$$R_d = e_p N_t (1 - F). \quad (1.82)$$

Die Emissionswahrscheinlichkeit  $e_p$  eines Lochs kann in Form von  $v_{th}$  und  $\sigma_p$  ausgedrückt werden, indem die thermische Gleichgewichtsbedingung, für die  $R_c = R_d$  berücksichtigt wird:

$$e_p = v_{th}\sigma_p n_i \exp\left(\frac{E_i - E_t}{kT}\right). \quad (1.83)$$

Betrachten wir nun den Nichtgleichgewichtsfall, in dem ein  $n$ -leitender Halbleiter gleichmäßig beleuchtet wird, um eine Generationsrate  $G_L$  zu erhalten. Zusätzlich zu dem in Abb. 1.27 dargestellten Prozess werden also Elektronen-Loch-Paare durch Licht erzeugt. Im stationären Zustand muss die Anzahl der im Leitungsband generierten und rekombinierenden Elektronen gleich groß sein. Dies nennt man das Prinzip des detaillierten Gleichgewichts, und es ergibt sich

$$\frac{dn_n}{dt} = G_L - (R_a - R_b) = 0. \quad (1.84)$$

In ähnlicher Weise führt das detaillierte Gleichgewicht der Löcher im Valenzband im stationären Zustand zu

$$\frac{dp_n}{dt} = G_L - (R_c - R_d) = 0. \quad (1.85)$$

Unter Gleichgewichtsbedingungen sind  $G_L = 0$ ,  $R_a = R_b$  und  $R_c = R_d$ . Im Nichtgleichgewichtszustand ist jedoch  $R_a \neq R_b$  und  $R_c \neq R_d$ . Aus Gln. (1.84) und (1.85) bekommen wir dann

$$G_L = R_a - R_b = R_c - R_d \equiv U. \quad (1.86)$$

Die Nettorekombinationsrate [8]  $U$  können wir aus Gln. (1.76), (1.77), (1.81) und (1.82) bekommen:

$$U = \frac{\sigma_n \sigma_p v_{th} N_t (p_n n_n - n_i^2)}{\sigma_n \left[ n_n + n_i \exp\left(\frac{E_i - E_t}{kT}\right) \right] + \sigma_p \left[ p_n + n_i \exp\left(\frac{E_i - E_t}{kT}\right) \right]}, \quad (1.87)$$

wobei  $\sigma_n$  und  $\sigma_p$  die Einfangquerschnitte für Elektronen und Löcher sind. Die Nettoübergangsrate ist proportional zu  $p_n n_n - n_i^2$ , und das Vorzeichen bestimmt, ob es sich um eine Nettorekombination oder Erzeugung handelt. Zweitens wird  $U$  maximal, wenn  $E_t = E_i$  ist, was darauf hindeutet, dass nur die Störstellenniveaus in der Mitte der Bandlücke effektive Rekombinations- und Generationszentren sind. Betrachtet man nur diese Niveaus, so reduziert sich Gl. (1.87) auf

$$U = \frac{\sigma_n \sigma_p v_{th} N_t (p_n n_n - n_i^2)}{\sigma_n (n_n + n_i) + \sigma_p (p_n + n_i)}. \quad (1.88)$$

Auch bei der schwachen Injektion in  $n$ -Typ-Halbleitern wird die Nettorekombinationsrate zu

$$U = \frac{\sigma_n \sigma_p v_{th} N_t [(p_{n0} + \Delta p) n_n - n_i^2]}{\sigma_n n} \approx \sigma_p v_{th} N_t \Delta p_n \equiv \frac{\Delta p_n}{\tau_p}, \quad (1.89)$$

wobei

$$\tau_p = \frac{1}{\sigma_p v_{th} N_t} \quad (1.90a)$$

ist. In ähnlicher Weise ist die Lebensdauer der Elektronen in einem  $p$ -Typ-Halbleiter durch

$$\tau_n = \frac{1}{\sigma_n v_{th} N_t} \quad (1.90b)$$

gegeben. Wie erwartet, ist die Lebensdauer bei indirekten Übergängen umgekehrt proportional zur Störstellendichte  $N_t$ , während im vorherigen Fall die Lebensdauer bei direkten Übergängen umgekehrt proportional zur Störstellenkonzentration bzw. der Dotierung ist (Gln. (1.73a) und (1.73b)).

Bei Störstellen mit mehreren Energieniveaus haben die Rekombinationsprozesse grob das gleiche Verhalten wie bei Störstellen mit nur einem Energieniveau. Im Detail gibt es jedoch Unterschiede, insbesondere im Fall der starken Injektion, wo sich  $\Delta n = \Delta p$  der Konzentration der Majoritätsladungsträger nähert. Hier ist die asymptotische Lebensdauer ein Mittelwert der Lebensdauern, die zu allen positiv geladenen, negativ geladenen und neutralen Störstellenniveaus gehören.

Für die starke Injektion ( $\Delta n = \Delta p > n$  und  $p$ ) wird die Formel für die Lebensdauer im Falle der Band-zu-Band-Rekombination zu

$$\tau_n = \tau_p = \frac{1}{R_{ec} \Delta n}. \quad (1.91)$$

Die aus den Einfangzentren resultierende Lebensdauer kann aus Gl. (1.88) abgeleitet werden und beträgt

$$\tau_n = \tau_p = \frac{\sigma_n + \sigma_p}{\sigma_n \sigma_p v_{th} N_t}. \quad (1.92)$$

Vergleicht man Gl. (1.92) mit Gln. (1.90a) und (1.90b), so ist die Lebensdauer bei starker Injektion tatsächlich höher. Interessant ist, dass die Lebensdauer aufgrund der Band-zu-Band-Rekombination mit dem Injektionsniveau abnimmt, während die Lebensdauer aufgrund der Rekombination über Rekombinationszentren mit dem Injektionsniveau zunimmt.

Die Gln. (1.90a) und (1.90b) wurden experimentell unter Verwendung von Festkörperdiffusion und hochenergetischer Strahlung verifiziert. Viele Störstellen haben ein Energieniveau nahe der Mitte der Bandlücke (Abb. 1.10). Diese Störstellen sind effiziente Rekombinationszentren. Ein typisches Beispiel ist Gold in Silizium [73]; die Lebensdauer der Minoritätsladungsträger sinkt linear mit der Goldkonzentration über den Bereich von  $10^{14}$  auf  $10^{17} \text{ cm}^{-3}$  ab, wobei  $\tau$  von etwa  $2 \times 10^{-6}$  auf  $2 \times 10^{-9} \text{ s}$  abnimmt. In einigen Hochgeschwindigkeitsanwendungen, besonders wenn eine kurze Lebensdauer zur Reduzierung der Ladungsspeicherzeit wünschenswert ist, ist dieser Effekt durchaus vorteilhaft. Eine weitere Methode zur Verkürzung der Lebensdauer der Minoritätsladungsträger ist die Bestrahlung des Halbleiters mit hochenergetischen Teilchen, die eine teilweise Verschiebung der Gitteratome und damit eine Schädigung des Gitters bewirkt. Die wiederum führt zu zusätzlichen Energieniveaus in der

Bandlücke. Zum Beispiel führt eine Elektronenbestrahlung in Si zu einem Akzeptorniveau bei 0.4 eV oberhalb des Valenzbandes und einem Donatorniveau bei 0.36 eV unterhalb des Leitungsbandes. Auch Neutronenbestrahlung erzeugt ein Akzeptorniveau bei 0.56 eV, und eine Bestrahlung mit Deuterium erzeugt einen Zwischengitterzustand (engl. interstitial state) mit einem Energieniveau von 0.25 eV über dem Valenzband. Ähnliche Ergebnisse werden für Ge, GaAs und andere Halbleiter erzielt. Im Gegensatz zu den durch Festkörperdiffusion eingebrachten Störstellen können die strahlungsinduzierten Störstellen und Einfangzentren bei relativ niedrigen Temperaturen getempert (engl. annealed) werden, wodurch sie sich wieder in Richtung Nirwana verabschieden.

Liegt die Ladungsträgerkonzentration im thermischen Gleichgewicht, d. h., wenn wir uns im Bereich  $pn < n_i^2$  befinden, kommt es eher zur Erzeugung von Ladungsträgern als zur Rekombination von überschüssigen Trägern. Die Erzeugungsrate kann aus Gl. (1.88) ermittelt werden

$$U = -\frac{\sigma_p \sigma_n v_{th} N_t n_i}{\sigma_p [1 + (p_n/n_i)] + \sigma_n [1 + (n_n/n_i)]} \equiv -\frac{n_i}{\tau_g}, \quad (1.93)$$

wobei die Lebensdauer der generierten Ladungsträger  $\tau_g$  so hingeschrieben werden kann:

$$\tau_g = \frac{1 + (n_n/n_i)}{\sigma_p v_{th} N_t} + \frac{1 + (p_n/n_i)}{\sigma_n v_{th} N_t} = \left(1 + \frac{n_n}{n_i}\right) \tau_p + \left(1 + \frac{p_n}{n_i}\right) \tau_n. \quad (1.94)$$

Abhängig von den Elektronen- und Löcherkonzentrationen kann die Generationslebensdauer viel länger als die Rekombinationslebensdauer sein. Typischerweise hat sie einen Mindestwert von etwa dem Doppelten der Rekombinationslebensdauer, wenn sowohl  $n$  als auch  $p$  viel kleiner als  $n_i$  sind.

Die Lebensdauer der Minoritätsladungsträger  $\tau$  wird im Allgemeinen unter Verwendung des Photoleitungseffekts (engl. photoconductive effect, PC-Effekt) [74] oder des photoelektromagnetischen (PEM)-Effekts [75] gemessen. Die grundlegende Gleichung für den PC-Effekt ist gegeben durch

$$J_{PC} = q(\mu_n + \mu_p) \Delta n \mathcal{E} = q(\mu_n + \mu_p) \frac{G_e}{\tau} \mathcal{E}, \quad (1.95)$$

wobei  $J_{PC}$  die Änderung der Stromdichte als Ergebnis der Beleuchtung mit der Erzeugungsrate  $G_e$  und  $\mathcal{E}$  das elektrische Feld entlang der Probe ist. Die Größe  $\Delta n$  ist die zusätzliche Ladungsträgerdichte bzw. die Anzahl der durch die Beleuchtung erzeugten Elektron-Loch-Paare pro Volumeneinheit, die dem Produkt aus der Erzeugungsrate  $G_e$  und der Lebensdauer  $\tau$  entspricht ( $\Delta n = \tau G_e$ ). Für den PEM-Effekt messen wir den Kurzschlussstrom, der entsteht, wenn ein konstantes Magnetfeld  $B_z$  senkrecht zur Richtung der einfallenden Strahlung angelegt wird. Die zugehörige Stromdichte ist gegeben durch

$$J_{PEM} = q(\mu_n + \mu_p) \mathcal{B}_z \frac{D}{L_d} \tau G_e = q(\mu_n + \mu_p) \mathcal{B}_z \sqrt{D\tau} G_e, \quad (1.96)$$

wobei  $D$  und  $L_d$  ( $\equiv (D\tau)^{1/2}$ ) der Diffusionskoeffizient und die Diffusionslänge sind, die im nächsten Abschnitt diskutiert werden. Ein anderer Ansatz zur Messung der Ladungsträgerlebensdauer wird in Abschn. 1.8.2 erörtert.

### 1.5.5 Diffusion

Im vorhergehenden Abschnitt hatten wir die Situation, dass überschüssige Ladungsträger gleichmäßig über den Halbleiter verteilt waren. In diesem Abschnitt erörtern

wir die Situation, in der überschüssige Ladungsträger örtlich injiziert oder generiert werden, was zu einem räumlichen Nichtgleichgewicht in der Ladungsträgerkonzentration führt. Beispiele sind die Injektion von Trägern über einen lokalen elektrischen Kontakt und eine lokale Beleuchtung der Probe. Wann immer ein Gradient in der Ladungsträgerkonzentration auftritt, findet ein Diffusionsprozess statt, durch den die Ladungsträger aus dem Bereich hoher Konzentration in den Bereich niedriger Konzentration wandern, und so das System wieder in den Gleichgewichtszustand bringen. Dieser Fluss von Ladungsträgern ist proportional zum Konzentrationsgradienten und wird (hier für Elektronen) durch das Ficksche Gesetz geregelt

$$\left. \frac{d\Delta n}{dt} \right|_x = -D_n \frac{d\Delta n}{dx}. \quad (1.97)$$

Die Proportionalitätskonstante wird als Diffusionskoeffizient oder Diffusivität  $D_n$  bezeichnet. Dieser Fluss von Ladungsträgern stellt einen Diffusionsstrom dar, der gegeben ist durch

$$J_n = qD_n \frac{d\Delta n}{dx}, \quad \text{und} \quad J_p = -qD_p \frac{d\Delta p}{dx}. \quad (1.98)$$

Physikalisch gesehen ist die Diffusion sowohl auf die zufällige thermische Bewegung der Ladungsträger als auch auf Streuprozesse zurückzuführen, und es gilt die Beziehung  $D = v_{th} \tau_m$ . Auch die Diffusionskoeffizienten und die Beweglichkeit hängen zusammen. Um diese Beziehung abzuleiten, betrachten wir einen  $n$ -Typ-Halbleiter mit inhomogener Dotierstoffkonzentration, aber ohne extern angelegtes elektrisches Feld. Da der Nettostrom gleich null ist, ist es notwendig, dass der Driftstrom und der Diffusionsstrom exakt gleich groß sind, also die Beziehung

$$qn\mu_n \mathcal{E} = -qD_n \frac{dn}{dx} \quad (1.99)$$

gilt. In diesem Fall wird das elektrische Feld  $\mathcal{E} = dE_C/qdx$  durch die inhomogene Dotierung erzeugt. Bei Verwendung von Gl. (1.18) erhält man für  $n$

$$\frac{dn}{dx} = \frac{-q\mathcal{E}}{kT} N_C \exp\left(-\frac{E_C - E_F}{kT}\right) = \frac{-q\mathcal{E}}{kT} n. \quad (1.100)$$

Wird dies in Gl. (1.99) eingebaut, so ergibt sich die Beziehung

$$D_n = \left(\frac{kT}{q}\right) \mu_n. \quad (1.101)$$

In ähnlicher Weise kann man für  $p$ -Typ-Halbleiter Folgendes ableiten:

$$D_p = \left(\frac{kT}{q}\right) \mu_p. \quad (1.102)$$

Diese Gleichungen sind als Einstein-Beziehungen bekannt und gelten nur für nicht degenerierte Halbleiter. Bei  $T = 300$  K ist  $kT/q = 0.0259$  V, die Werte von  $D$  sind dann mit den in Abb. 1.15 gezeigten Beweglichkeiten leicht zu ermitteln.

Ein weiterer Parameter, der eng mit der Diffusion zusammenhängt, ist die Diffusionslänge

$$L_d = \sqrt{D\tau}. \quad (1.103)$$

Bei üblichen Diffusionsproblemen, die sich aus einer festen Injektionsquelle als Randbedingung ergeben, fällt das resultierende Konzentrationsprofil exponentiell mit einer charakteristischen Länge  $L_d$  ab. Diese Diffusionslänge kann auch als die Distanz betrachtet werden, über welche die Ladungsträger innerhalb der ihrer Lebensdauer diffundieren können, ehe sie vernichtet werden.

### 1.5.6 Thermionische Emission

Ein weiterer Stromleitungsmechanismus ist die thermionische Emission. Hier handelt es sich um einen Strom von Majoritätsladungsträgern, der immer mit der Existenz einer Potentialbarriere verbunden ist. Beachten Sie, dass dabei der kritische Parameter die Höhe der Barriere ist, die Form der Barriere hingegen spielt dabei so gut wie keine Rolle. Das gebräuchlichste Bauelement, in dem thermionische Emission eine Rolle spielt, ist die Schottky-Diode oder allgemein ausgedrückt, jeder beliebige Metall-Halbleiter-Kontakt (siehe Kap. 3). Wie man in Abb. 1.28 sieht, ist die thermionische Emission dann der dominante Prozess, wenn Streuung oder Drift-Diffusions-Prozesse innerhalb der Sperrschicht vernachlässigbar sind. Entsprechend muss die Breite der Barriere immer schmaler als die mittlere freie Weglänge sein. Im Falle einer dreieckigen Barriere sollte die Neigung der Barriere relativ steil sein und die Beziehung  $q\Delta V_b(x) > kT/\lambda_m$  erfüllen ( $\lambda_m$  ist die mittlere freie Weglänge). Darüber hinaus darf nach der Injektion der Ladungsträger über die Barriere der Diffusionsstrom in dieser Region nicht der begrenzende Faktor sein, weshalb der Bereich hinter der Barriere ein weiterer  $n$ -Typ-Halbleiter oder eine Metallschicht sein muss.

Aufgrund der Fermi-Dirac-Statistik nimmt bei endlichen Temperaturen die Dichte der Elektronen in Abhängigkeit von ihrer Energie über der Leitungsbandkante exponentiell ab. Von besonderem Interesse ist dabei die integrierte Anzahl von Ladungsträgern oberhalb der Barrierenkante. Dieser Teil der thermisch erzeugten Ladungsträger wird nicht mehr durch die Barriere zurückgehalten und trägt damit zum thermionischen Emissionsstrom bei. Der gesamte Strom über die Barriere (siehe Kap. 3) ist durch die Formel

$$J = A^* T^2 \exp\left(-\frac{q\phi_B}{kT}\right) \quad (1.104)$$

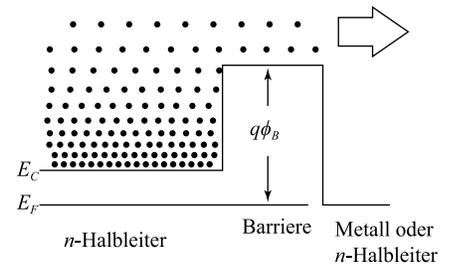
gegeben, wobei  $\phi_B$  die Höhe der Barriere ist. Die Konstante

$$A^* \equiv \frac{4\pi q m^* k^2}{h^3} \quad (1.105)$$

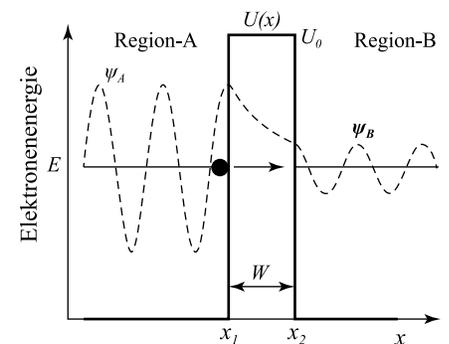
wird die effektive Richardson-Konstante genannt und ist eine Funktion der effektiven Masse.  $k$  ist die Boltzmann-Konstante. Hinweis: Der Ausdruck für  $A^*$  kann durch die Berücksichtigung quantenmechanischer Tunnel- und Reflexionsprozesse weiter modifiziert werden.

### 1.5.7 Tunnelprozesse

Tunnelprozesse sind ein quantenmechanisches Phänomen. Im klassischen Bild verhalten sich durch Potentialwände eingeschlossene Ladungsträger wie Zuckerkrystalle in einer Zuckerdose. Nur diejenigen Ladungsträger, die energetisch über der Barriere liegen, können aus dem Potentialtopf entkommen, oder man könnte – auf die Zuckerdose umgelegt – sagen, dass die Dose übertoll ist, und der Zucker deswegen über den Rand rinnt. In der Quantenmechanik kann ein Elektron durch seine Wellenfunktion dargestellt werden. Diese Wellenfunktion endet aber nicht abrupt an einer Wand mit endlicher Potentialhöhe und kann durchaus etwas in die Barriere eindringen (Abb. 1.29) oder diese Barriere sogar durchdringen. Die Wahrscheinlichkeit des Elektronentransfers durch eine Barriere endlicher Höhe und Breite ist also nicht null (Tunneleffekt). Im Bild der Zuckerdose hätte die Wand einer halbvollen Dose weiter unten in der Nähe des Bodens irgendein Loch, durch das der Zucker herausrinnt, was zwar möglich, in einer ordentlichen Küche aber eher unwahrscheinlich ist (Zucker hat also eine eher geringe Tunnelwahrscheinlichkeit). Zur Berechnung der Tunnelwahrscheinlichkeit muss die Wellenfunktion  $\psi$  aus der Schrödinger-Gleichung bestimmt



**Abb. 1.28** Thermionische Emission von Elektronen über eine Barriere. Man beachte, dass die Form der Barriere, hier als Rechteck dargestellt, keine Rolle spielt.



**Abb. 1.29** Wellenfunktion und Tunnelprozess von Elektronen durch eine rechteckige Barriere.

werden

$$\frac{d^2\psi}{dx^2} + \frac{2m^*}{\hbar^2}[E - U(x)]\psi = 0. \quad (1.106)$$

Im Falle einer einfachen rechteckigen Barriere der Höhe  $U_0$  und der Breite  $W$  hat die Wellenfunktion  $\psi$ , bis auf irgendwelche Normierungskonstanten, die allgemeine Form  $\psi = \exp(\pm ikx)$ , wobei  $k = \sqrt{2m^*(E - U_0)}/\hbar$  ist. Beachten Sie, dass beim Tunneln die Energie  $E$  unterhalb der Barrierrhöhe  $U_0$  liegt, sodass der Term innerhalb der Quadratwurzel negativ und  $k$  damit imaginär ist. Die Lösung der Wellenfunktionen und die Tunnelwahrscheinlichkeiten werden wie folgt berechnet:

$$T_t = \frac{|\psi_B|^2}{|\psi_A|^2} = \left[ 1 + \frac{U_0^2 \sinh^2(|k|W)}{4E(U_0 - E)} \right]^{-1} \\ \approx \frac{16E(U_0 - E)}{U_0^2} \exp\left(-2\sqrt{\frac{2m^*(U_0 - E)}{\hbar^2}}W\right). \quad (1.107)$$

Für kompliziertere Barrierenformen kann, wenn das Potential  $U(x)$  nicht zu schnell variiert, zur Lösung der Schrödinger-Gleichung die WKB-Näherung (Wentzel-Kramers-Brillouin, WKB) verwendet werden. Die Wellenfunktion hat nun eine allgemeine Form von  $\exp(\int ik(x)dx)$ , und die Tunnelwahrscheinlichkeit kann durch die Formel

$$T_t = \frac{|\psi_B|^2}{|\psi_A|^2} \approx \exp\left\{-2 \int_{x_1}^{x_2} |k(x)| dx\right\} \\ \approx \exp\left\{-2 \int_{x_1}^{x_2} \sqrt{\frac{2m^*}{\hbar^2}[U(x) - E]} dx\right\} \quad (1.108)$$

berechnet werden. Zusammen mit der bekannten Tunnelwahrscheinlichkeit ergibt sich der Tunnelstrom  $J_t$  aus dem Produkt der Anzahl der verfügbaren Ladungsträger in der Ausgangsregion  $A$  (Abb. 1.29) und der Anzahl der Leerzustände in der Zielregion  $B$ :

$$J_t = \frac{qm^*}{2\pi^2\hbar^3} \int F_A N_A T_t (1 - F_B) N_B dE, \quad (1.109)$$

wobei  $F_A$ ,  $F_B$ ,  $N_A$  und  $N_B$  die Fermi-Dirac-Verteilungen und Zustandsdichten in den entsprechenden Regionen sind.

### 1.5.8 Raumladungseffekte

Die Raumladung ist eine ortsfeste Ladungsverteilung in einem Halbleiter, wie z. B. in einem  $p$ - $n$ -Übergang, und wird sowohl von der Dotierungs- als auch von der Konzentration der freien Ladungsträger bestimmt. Die Formel für die Raumladungsdichte lautet also  $\rho = q(p - n + N_D - N_A)$ . Im neutralen Bereich eines Halbleiters ist  $n = N_D$  und  $p = N_A$ , sodass die Raumladungsdichte null ist. In der Nähe einer Grenzschicht, die aus verschiedenen Materialien, Dotierungsarten oder Dotierstoffkonzentrationen besteht, können  $n$  und  $p$  durchaus kleiner oder größer als  $N_D$  bzw.  $N_A$  sein. Bei der Verarmungsnäherung (engl. depletion approximation) werden  $n$  und  $p$  als null angenommen, sodass die Raumladung gleich dem Dotierungsniveau der Majoritätsladungsträger ist. Unter Vorspannung können die Ladungsträgerkonzentrationen  $n$  und

$p$  über ihre Gleichgewichtswerte hinaus erhöht werden. Wenn die injizierten Konzentrationen  $n$  oder  $p$  größer als ihr Gleichgewichtswert sowie größer als die Dotierungskonzentration sind, spricht man vom Raumladungseffekt. Die injizierten Ladungsträger steuern somit die Raumladungszone und das elektrische Feldprofil. Dies führt zu einem Rückkopplungsmechanismus, bei dem das Feld den Strom antreibt, der wiederum das Feld aufbaut. Der Raumladungseffekt ist bei leicht dotierten Materialien häufiger und er kann auch außerhalb der Verarmungsregion auftreten.

Wenn bei Vorliegen eines Raumladungseffekts der Strom von der Driftkomponente der injizierten Ladungsträger dominiert wird, spricht man von einem raumladungsbegrenzten Strom. Da es sich um einen Driftstrom handelt, ist die Driftkomponente der Stromdichte im Falle der Elektroneninjektion gegeben durch

$$J = qnv. \quad (1.110)$$

Die Raumladung wiederum wird durch die injizierten Ladungsträger bestimmt, woraus sich die Poisson-Gleichung der Form

$$\frac{d^2\psi_i}{dx^2} = \frac{qn}{\epsilon_s} \quad (1.111)$$

ergibt.

Die Ladungsträgergeschwindigkeit  $v$  hängt je nach Feldstärke auf verschiedene Art mit dem elektrischen Feld zusammen, und bei niedrigen Feldern gilt  $v = \mu\mathcal{E}$ . Im Bereich der Sättigungsgeschwindigkeit  $v_s$  ist die Geschwindigkeit unabhängig vom elektrischen Feld. Im Grenzfall von ultrakurzen Proben oder Zeitskalen befinden wir uns im ballistischen Bereich, bei dem es keine Streuung gibt, und erhalten die Beziehung

$$v = \sqrt{\frac{2qV}{m^*}}. \quad (1.112)$$

Aus Gln. (1.110)–(1.112) kann der raumladungsbegrenzte Strom im Beweglichkeitsbereich berechnet werden (Mott-Gurney-Gesetz, siehe [5]):

$$J = \frac{9\epsilon_s\mu V^2}{8L^3}. \quad (1.113)$$

Im Bereich der Sättigungsgeschwindigkeit bekommt man

$$J = \frac{2\epsilon_s v_s V}{L^2}, \quad (1.114)$$

und im ballistischen Bereich erhält man das Child-Langmuir-Gesetz

$$J = \frac{4\epsilon_s}{9L^2} \left(\frac{2q}{m^*}\right)^{1/2} V^{3/2}. \quad (1.115)$$

Dabei ist  $L$  die Länge der Probe in Richtung des Stromflusses. Beachten Sie, dass die Spannungsabhängigkeit des Stromes in den genannten Bereichen unterschiedlich ist.

## 1.6 Phononen, optische und thermische Eigenschaften

Im vorhergehenden Abschnitt haben wir verschiedene Transportmechanismen in Halbleitern betrachtet. In diesem Abschnitt betrachten wir kurz einige andere Effekte und Eigenschaften von Halbleitern, die für den Betrieb von Halbleiterbauelementen wichtig sind.

### 1.6.1 Phononenspektren

Phononen sind quantisierte Gitterschwingungen, die hauptsächlich aus der thermischen Gitterenergie resultieren. Ähnlich wie Photonen und Elektronen haben

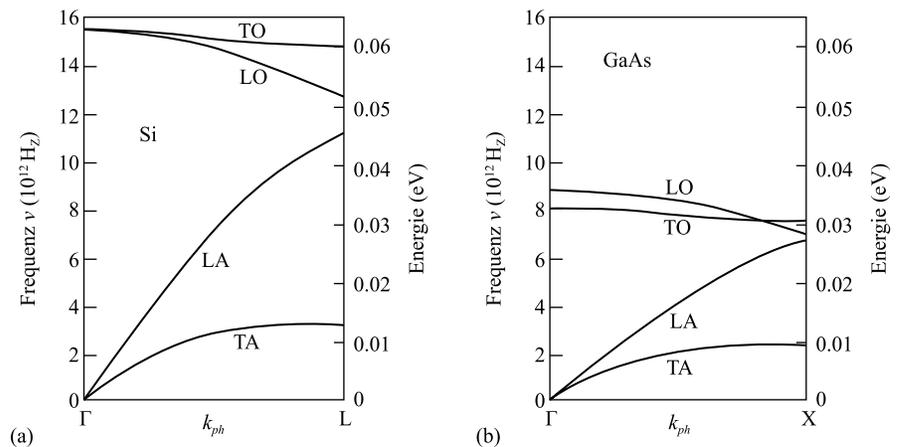
sie charakteristische Frequenzen (oder Energien) und Wellenzahlen (Impulse oder Wellenlängen). In einem eindimensionalen Gittermodell, bei dem nur die Kopplung zwischen benachbarten Atomen berücksichtigt wird und zwei verschiedene Massen  $m_1$  und  $m_2$  abwechselnd angeordnet sind, sind die Schwingungsfrequenzen gegeben durch [4]

$$\nu_{\pm} = \sqrt{\alpha_f} \left[ \left( \frac{1}{m_1} + \frac{1}{m_2} \right) \pm \sqrt{\left( \frac{1}{m_1} + \frac{1}{m_2} \right)^2 - \frac{4 \sin^2(k_{ph}a/2)}{m_1 m_2}} \right]^{1/2}, \quad (1.116)$$

wobei  $\alpha_f$  die Kraftkonstante des hookeschen Gesetzes,  $k_{ph}$  die Wellenzahl der Phononen und  $a$  der Gitterabstand ist. In der Nähe von  $k_{ph} = 0$  ist die Frequenz  $\nu_-$  proportional zu  $k_{ph}$ . Diesen Zweig nennt man akustischen Zweig, da es sich um die langwellige Schwingung des Gitters handelt und die Geschwindigkeit  $\omega/k$  nahe der Schallgeschwindigkeit in diesem Medium liegt. Die Frequenz  $\nu_+$  ist bei  $k_{ph} \approx 0$  in etwa konstant ( $\approx [2\alpha_f(1/m_1 + 1/m_2)]^{1/2}$ ). Diesen, vom akustischen Fall deutlich getrennten Zweig der Dispersionsrelation, nennt man optischen Zweig, da die Frequenz  $\nu_+$  im Allgemeinen im infrarotoptischen Bereich liegt. Bei den akustischen Moden bewegen sich die beiden Untergitter der Atome mit unterschiedlichen Massen in die gleiche Richtung, während sie sich bei den optischen Moden in entgegengesetzte Richtungen bewegen.

Die Gesamtzahl der akustischen Moden ist gleich der Dimension (1-, 2-, 3-dimensionales Gitter) mal der Anzahl der Atome pro Zelle. Für ein realistisches dreidimensionales Gitter mit einem Atom pro Primitivzelle, wie z. B. ein einfaches kubisches, raumzentriertes oder flächenzentriertes kubisches Gitter, existieren nur drei akustische Moden. Für ein dreidimensionales Gitter mit zwei Atomen pro primitiver Einheitszelle, wie z. B. Si und GaAs, existieren drei akustische und drei optische Moden. Longitudinal polarisierte Moden sind Moden, bei denen der Verschiebungsvektor jedes Atoms entlang der Richtung des Wellenvektors liegt. Wir haben also eine longitudinale akustische Mode (LA) und eine longitudinale optische Mode (LO). Moden mit Atomen, die sich in den Ebenen senkrecht zum Wellenvektor bewegen, werden als transversal polarisierte Moden bezeichnet. Es gibt für unseren Fall also zwei transversal akustische Moden (TA) und zwei transversal optische Moden (TO).

Abbildung 1.30 zeigt experimentelle Ergebnisse für Si und GaAs in einer der Kristallrichtungen. Der Bereich von  $k_{ph} = \pm\pi/a$  definiert die Brillouin-Zone, außerhalb der sich das Frequenz- $k_{ph}$ -Schema lediglich wiederholt. Beachten Sie, dass die Ener-



**Abb. 1.30** Gemessene Phononenspektren in (a) Si (basierend auf [76]) und (b) GaAs (basierend auf [77]). TO und LO stehen für transversale und longitudinale optische Moden, und TA und LA für transversale und longitudinale akustische Moden.

gien (oder Frequenzen) bei kleinen Werten von  $k_{ph}$  sowohl für die LA-Mode als auch für die TA-Moden proportional zu  $k_{ph}$  sind. Die longitudinal optische (LO-) Phononenenergie bei  $k_{ph} = 0$  ist die Raman-Streuenergie erster Ordnung mit den Werten von 0.063 eV für Si und 0.035 eV für GaAs. In Anhang G sind diese Werte zusammen mit weiteren anderen wichtigen Parametern aufgelistet.

### 1.6.2 Optische Eigenschaften

Optische Messungen stellen das wichtigste Mittel zur Bestimmung der Bandstruktur von Halbleitern dar. Photoneninduzierte elektronische Übergänge können zwischen verschiedenen Bändern oder innerhalb eines einzelnen Bandes, wie z. B. im Falle der freien Trägerabsorption, auftreten. Optische Messungen können auch zur Untersuchung von Gitterschwingungen (Phononen) verwendet werden. Charakterisiert werden die optischen Eigenschaften von Halbleitern durch den komplexen Brechungsindex,

$$\bar{n} = n_r - ik_e. \quad (1.117)$$

Der Realteil des Brechungsindex  $n_r$  bestimmt die Ausbreitungsgeschwindigkeit ( $v$  und Wellenlänge  $\lambda$ ) im Medium gemäß

$$n_r = \frac{c}{v} = \frac{\lambda_0}{\lambda}. \quad (1.118)$$

$\lambda_0$  ist die Wellenlänge und  $c$  die Lichtgeschwindigkeit im Vakuum. Der Imaginärteil  $k_e$ , der sogenannte Extinktionskoeffizient, bestimmt wiederum den Absorptionskoeffizienten

$$\alpha = \frac{4\pi k_e}{\lambda}. \quad (1.119)$$

In Halbleitern hängt der Absorptionskoeffizient stark von der Wellenlänge bzw. der Photonenenergie ab. In der Nähe der Absorptionskante kann der Absorptionskoeffizient ausgedrückt werden als [6]

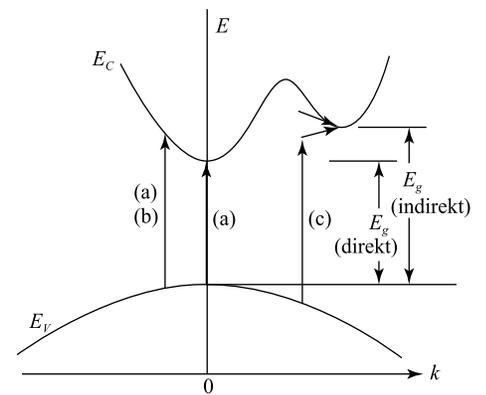
$$\alpha \propto (h\nu - E_g)^\gamma, \quad (1.120)$$

wobei  $\nu$  die Frequenz,  $h\nu$  die Photonenenergie und  $\gamma$  eine Konstante ist. Es gibt zwei Arten von Band-zu-Band-Übergängen: erlaubte und verbotene, wobei verbotene Übergänge den kleinen, aber endlichen Impuls der Photonen berücksichtigen und damit viel weniger wahrscheinlich sind. Bei Materialien mit direkter Bandlücke treten Übergänge meist zwischen zwei Bändern bei gleichem  $k$ -Wert auf, wie z. B. die in Abb. 1.31 gezeigten Übergänge (a) und (b). Während erlaubte direkte Übergänge in allen  $k$  Werten auftreten können, können verbotene direkte Übergänge nur bei  $k \neq 0$  auftreten.

In der Ein-Elektronen-Näherung ist  $\gamma$  gleich 1/2 für erlaubte bzw. 3/2 für verbotene direkte Übergänge. Beachten Sie, dass für  $k = 0$ , bei dem die Bandlücke definiert ist, nur der zulässige Übergang ( $\gamma = 1/2$ ) auftritt und daher zur experimentellen Bestimmung der Bandlücke verwendet wird. Bei indirekten Übergängen, wie Übergang (c) in Abb. 1.31, sind Phononen beteiligt, um den Impuls zu erhalten. Bei diesen Übergängen werden Phononen mit der Energie  $E_p$  entweder absorbiert oder emittiert, und die Formel für den Absorptionskoeffizienten ändert sich zu

$$\alpha \propto (h\nu - E_g \pm E_p)^\gamma. \quad (1.121)$$

Hier ist die Konstante  $\gamma$  für erlaubte und verbotene indirekte Übergänge gleich zwei bzw. drei.



**Abb. 1.31** Optische Übergänge im Halbleiter: (a) erlaubte, (b) verbotene direkte Übergänge und (c) indirekte Übergänge mit Phononenemission (oberer Pfeil) und Phononenabsorption (unterer Pfeil).

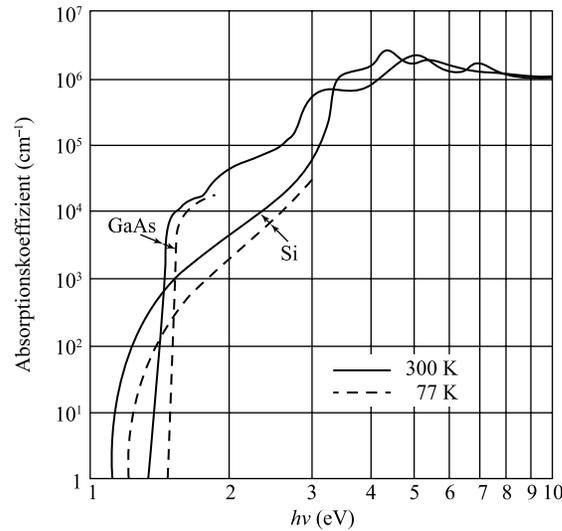


Abb. 1.32 Gemessene Absorptionskoeffizienten nahe und oberhalb der fundamentalen Absorptionskante für Si und GaAs (nach [78–81]).

Darüber hinaus können zusätzliche Absorptionsmaxima und Stufen in den Absorptionsspektren auf die Bildung von Exzitonen zurückzuführen sein. Exzitonen sind aneinander gebundene Elektron-Loch-Paare, die sich aber als Ganzes durch das Kristallgitter bewegen können. In der Nähe der Absorptionskante, wo der Wert von  $(E_g - h\nu)$  mit der Bindungsenergie eines Exzitons vergleichbar wird, muss die Coulomb-Wechselwirkung zwischen dem freien Elektron und dem Loch berücksichtigt werden, und die zur Absorption erforderliche Photonenenergie wird dann durch diese Bindungsenergie vermindert. Für  $h\nu \lesssim E_g$  geht die Absorption kontinuierlich in die fundamentale Absorption über. Ist  $h\nu \gg E_g$ , sind höhere Energiebänder an den Übergangsprozessen beteiligt, und die komplizierte Bandstruktur spiegelt sich in zusätzlichen Strukturen im Absorptionskoeffizienten wider.

Abbildung 1.32 zeigt den experimentellen Absorptionskoeffizienten  $\alpha$  für Si und GaAs in der Nähe der fundamentalen Absorptionskante. Die Verschiebung der Kurven zu höheren Photonenenergien bei niedrigerer Temperatur hängt mit der Temperaturabhängigkeit der Bandlücke zusammen (Abb. 1.6). Ein  $\alpha$  von  $10^4 \text{ cm}^{-1}$  bedeutet, dass 63% des Lichts bereits  $1 \mu\text{m}$  unter der Halbleiteroberfläche absorbiert werden.

Durch die Absorption von Licht im Halbleiter kommt es zur Erzeugung von Elektron-Loch-Paaren ( $G_e$ ), womit die Lichtintensität  $P_{op}$  mit der Distanz gemäß der Beziehung

$$\frac{dP_{op}(x)}{dx} = -\alpha P_{op}(x) = G_e h\nu \quad (1.122)$$

abnimmt. Die Lösung des obigen Problems ergibt einen exponentiellen Intensitätsabfall

$$P_{op}(x) = P_0(1 - R) \exp(-\alpha x), \quad (1.123)$$

wobei  $P_0$  die von außen einfallende Lichtintensität und  $R$  die Reflexion an der Halbleiteroberfläche bei normalem Einfall ist:

$$R = \frac{(1 - n_r)^2 + k_e^2}{(1 + n_r)^2 + k_e^2}. \quad (1.124)$$

In einer Halbleiterprobe der Dicke  $W$  treten bei einem nicht zu großen Produkt von  $\alpha W$  zwischen den beiden Grenzflächen Mehrfachreflexionen auf. Wenn man alle

Lichtkomponenten in Rückwärtsrichtung zusammenzählt, ergibt sich der totale Reflexionskoeffizient zu

$$R_{\Sigma} = R \left[ 1 + \frac{(1 - R)^2 \exp(-2\alpha W)}{1 - R^2 \exp(-2\alpha W)} \right], \quad (1.125)$$

und der Transmissionskoeffizient ist gegeben durch

$$T_{\Sigma} = \frac{(1 - R)^2 \exp(-\alpha W)}{1 - R^2 \exp(-2\alpha W)}. \quad (1.126)$$

Der Transmissionskoeffizient  $T_{\Sigma}$  und der Reflexionskoeffizient  $R_{\Sigma}$  sind zwei wichtige Größen, die an vielen Stellen benötigt werden. Durch die Analyse der Daten von  $T_{\Sigma}(\lambda)$  oder  $R_{\Sigma}(\lambda)$  bei normalem Einfallswinkel oder durch die Messung von  $R_{\Sigma}$  oder  $T_{\Sigma}$  für verschiedene Einfallswinkel bei konstantem  $\lambda$ , können sowohl  $n_r$  als auch  $k_e$  erhalten und mit der Übergangsenergie zwischen den Bändern in Beziehung gesetzt werden.

### 1.6.3 Thermische Eigenschaften

Wenn in einem Halbleiter zusätzlich zu einem angelegten elektrischen Feld ein Temperaturgradient besteht, beträgt die Gesamtstromdichte (in einer Dimension) [2]:

$$J = \sigma \left( \frac{1}{q} \frac{dE_F}{dx} - \mathcal{P} \frac{dT}{dx} \right), \quad (1.127)$$

wobei der Name für die thermoelektrische Leistung  $\mathcal{P}$  daher kommt, dass bei offenem Stromkreis der Nettostrom null ist und durch den Temperaturgradienten ein elektrisches Feld erzeugt wird. Für einen nicht degenerierten Halbleiter mit einer mittleren Streuzeit  $\tau_m \propto E^{-s}$  ist die thermoelektrische Leistung gegeben durch

$$\mathcal{P} = -\frac{k}{q} \left( \frac{\left[ \frac{5}{2} - s + \ln(N_C/n) \right] n\mu_n - \left[ \frac{5}{2} - s - \ln(N_V/p) \right] p\mu_p}{n\mu_n + p\mu_p} \right) \quad (1.128)$$

( $k$  ist die Boltzmann-Konstante). Diese Gleichung zeigt, dass die thermoelektrische Leistung bei  $n$ -Typ-Halbleitern negativ und bei  $p$ -Typ-Halbleitern positiv ist, was dann gerne zur Bestimmung des Leitungstyps im Halbleiter verwendet wird. Die thermoelektrische Leistung kann auch zur Bestimmung des spezifischen Widerstandes und der Position des Fermi-Niveaus relativ zu den Bandkanten verwendet werden. Bei Raumtemperatur steigt die thermoelektrische Leistung von  $p$ -Typ-Silizium mit dem spezifischen Widerstand. Typische Werte sind 1 mV/K für eine 0.1  $\Omega$ -cm-Probe und 1.7 mV/K für eine 100  $\Omega$ -cm-Probe. Ähnliche Ergebnisse (abgesehen vom Vorzeichen für  $\mathcal{P}$ ) können für  $n$ -Typ-Siliziumproben erzielt werden.

Ein weiterer wichtiger thermischer Effekt ist die Wärmeleitung. Es handelt sich um einen Diffusionsprozess, bei dem der Wärmestrom  $Q$  durch den Temperaturgradienten bestimmt wird, und es gilt die Gleichung

$$Q = -\kappa \frac{dT}{dx}. \quad (1.129)$$

Die Wärmeleitfähigkeit  $\kappa$  hängt hauptsächlich von der Wärmeleitung durch Phononen  $\kappa_L$  und der gemischten Wärmeleitung durch freie Ladungsträger (Elektronen und Löcher)  $\kappa = \kappa_L + \kappa_M$  ab. Der Gitterbeitrag der Wärmeleitung erfolgt durch Diffusion und Streuung von Phononen. Diese Streueignisse umfassen viele Varianten, wie

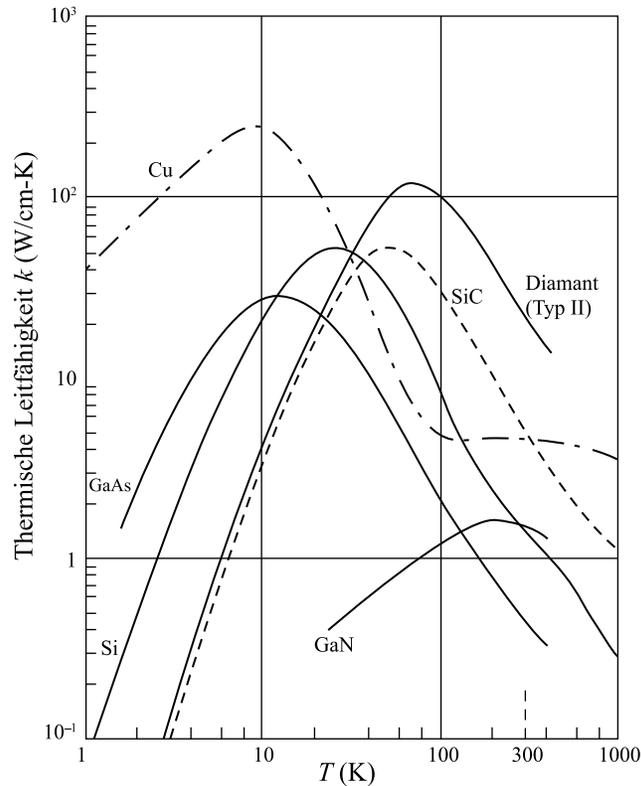


Abb. 1.33 Gemessene Wärmeleitfähigkeit über der Temperatur für reines Si, GaAs, SiC, GaN, Cu und Diamant (Typ II) (nach [82–86]).

z. B. Phonon-Phonon-Streuung, Phonon-Defekt-Streuung, Streuung zwischen Phononen und Ladungsträgern, Grenz- und Oberflächenstreuung usw. Die Gesamtwirkung kann geschrieben werden als

$$\kappa_L = \frac{1}{3} C_p v_{ph} \lambda_{ph}, \quad (1.130)$$

wobei  $C_p$  die spezifische Wärme,  $v_{ph}$  die Phononengeschwindigkeit und  $\lambda_{ph}$  die freie Weglänge des Phonons ist. Der Beitrag durch gemischte Träger ( $\tau_m \propto E^{-s}$ ) gilt sowohl für Elektronen- als auch für Löcherstreuung und ist gegeben durch

$$\kappa_M = \frac{\left(\frac{5}{2} - s\right) k^2 \sigma T}{q^2} + \frac{k^2 \sigma T [5 - 2s + (E_g/kT)]^2 n p \mu_n \mu_p}{q^2 (n \mu_n + p \mu_p)^2}. \quad (1.131)$$

Abbildung 1.33 zeigt die gemessene Wärmeleitfähigkeit als Funktion der Gittertemperatur für Si und GaAs. In Anhang G sind die Raumtemperaturwerte aufgeführt. Die Beiträge der Leitungsträger zur Wärmeleitfähigkeit sind im Allgemeinen recht gering, sodass die allgemeine Temperaturabhängigkeit der von  $\kappa_L$  folgt und eine umgekehrte V-Form aufweist. Bei niedrigen Temperaturen hat die spezifische Wärme eine  $T^3$ -Abhängigkeit und  $\kappa$  steigt stark an. Bei hohen Temperaturen dominiert die Phonon-Phonon-Streuung und  $\lambda_{ph}$  (und  $\kappa_L$ ) folgt einem  $1/T$ -Verhalten. Abbildung 1.33 zeigt auch die Wärmeleitfähigkeiten für Cu, Diamant, SiC und GaN. Kupfer ist das meistverwendete Metall für die Wärmeleitung in Bauelementen mit  $p$ - $n$ -Übergängen; Diamant hat die höchste bisher bekannte Wärmeleitfähigkeit bei Raumtemperatur und ist als Wärmesenke für Halbleiterlaser und IMPATT-Oszillatoren nützlich. SiC und GaN sind wichtige Halbleiter für Leistungsbaulemente.

## 1.7 Heteroübergänge und Nanostrukturen

Ein Heteroübergang ist ein Übergang, der zwischen zwei unterschiedlichen Halbleitermaterialien gebildet wird. Bei Anwendungen in Halbleiterbauelementen bietet der damit verbundene Unterschied in der Bandlücke viele interessante Möglichkeiten, weswegen Heteroübergänge in verschiedenen Bauelementen inzwischen weit verbreitet sind. Der erfolgreiche Einsatz von Heteroübergängen in verschiedenen Bauelementen ist auf die technologischen Möglichkeiten in der Epitaxie zurückzuführen, welche es ermöglichen, gitterangepasste Halbleitermaterialien praktisch ohne Grenzflächendefekte übereinander aufzuwachsen. Die zugrunde liegende Physik der epitaktischen Heteroübergänge ist die Anpassung der Gitterkonstanten zwischen den verwendeten Materialien. Eine starke Gitterfehlpassung führt zu Versetzungen an der Grenzfläche und damit zu elektronischen Defekten und Störstellenniveaus innerhalb der Bandlücke. Die Gitterkonstanten einiger gebräuchlicher Halbleiter finden sich zusammen mit ihren jeweiligen Bandlücken in Abb. 1.34.

Eine gute Kombination für Bauelemente mit Heteroübergängen sind zwei Materialien mit ähnlichen Gitterkonstanten, die sich jedoch in der Bandlücke  $E_g$  unterscheiden, wie dies z. B. im GaAs/AlGaAs (oder GaAs/AlAs)-System der Fall ist. Im  $\text{Al}_x\text{Ga}_{1-x}\text{As}$ /GaAs-System gibt es eine nahezu perfekte Gitteranpassung unabhängig vom Aluminiumgehalt  $x$ . Für  $x = 0.4$  hat  $\text{Al}_{0.4}\text{Ga}_{0.6}\text{As}$  eine Gitterkonstante von  $a = 5.6570 \text{ \AA}$ , welche praktisch identisch mit der Gitterkonstante von GaAs ist.

Abbildung 1.35 zeigt die Bandlücke von  $\text{Al}_x\text{Ga}_{1-x}\text{As}$  als Funktion des Aluminiumgehalts  $x$ . Die punktierte Linie zeigt die Lage des X-Tals. Man erkennt, dass für einen Aluminiumgehalt oberhalb von  $x = 0.5$  das X-Tal energetisch unter das  $\Gamma$ -Tal fällt und  $\text{Al}_x\text{Ga}_{1-x}\text{As}$  damit zum indirekten Halbleiter wird. Den Bandverlauf einer  $\text{Al}_{0.4}\text{Ga}_{0.6}\text{As}$ -GaAs-Heterostruktur sieht man in Abb. 1.36.

Sind die Gitterkonstanten nicht zu verschieden und ist die Dicke der epitaktischen Schichten klein genug, so ist noch immer eine Heteroepitaxie in guter Qualität möglich. Dies kann mithilfe von Abb. 1.37 erklärt werden, wo man erkennt, dass das Ausmaß der Gitterfehlpassung und die maximal zulässige Dicke der Epita-

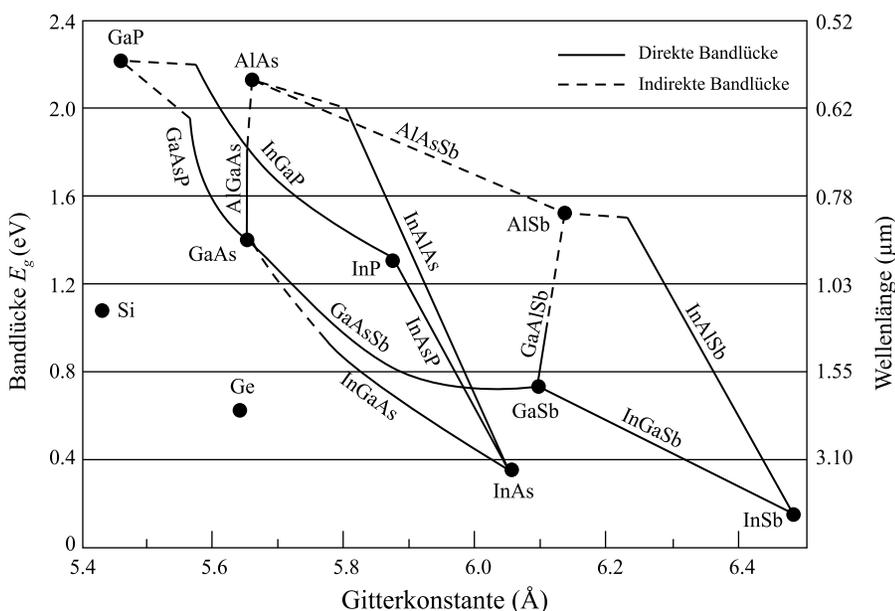
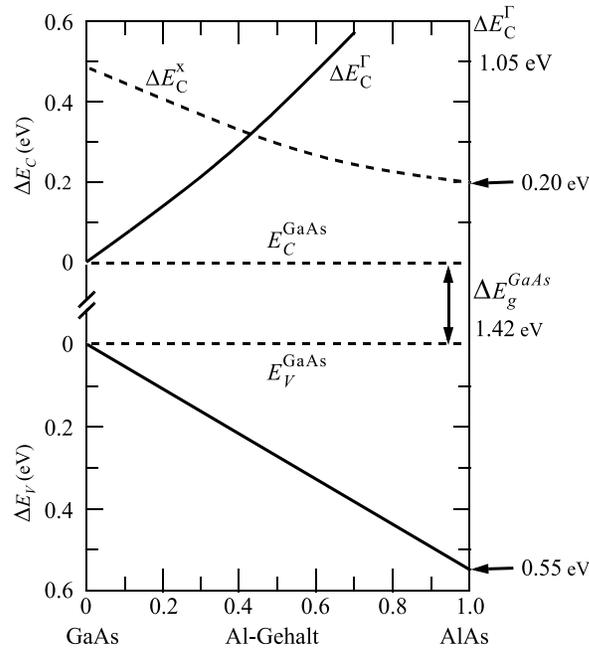
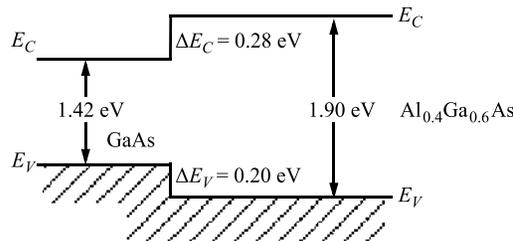


Abb. 1.34 Bandlücke als Funktion der Gitterkonstante für einige gängige elementare und binäre Halbleiter.



**Abb. 1.35** Die Bandlücke von  $\text{Al}_x\text{Ga}_{1-x}\text{As}$  von GaAs als Funktion des Aluminiumgehalts  $x$ . Die gepunktete Linie zeigt die Lage des X-Tals.



**Abb. 1.36** Der Bandverlauf von  $\text{Al}_{0.4}\text{Ga}_{0.6}\text{As}$  auf GaAs.

gieschicht in direktem Zusammenhang stehen. Für unverspannte, dicke, heteroepitaktische Schichten sind Versetzungen aufgrund der physikalischen Fehlanpassung der terminierenden Bindungen an der Grenzfläche unvermeidlich. Ist die epitaktisch aufgewachsene Schicht jedoch dünn genug, kann diese Schicht physikalisch so stark verspannt werden, dass ihre Gitterkonstante mit der des Substrats übereinstimmt (Abb. 1.37c). Wenn dies geschieht, können Versetzungen vermieden werden.

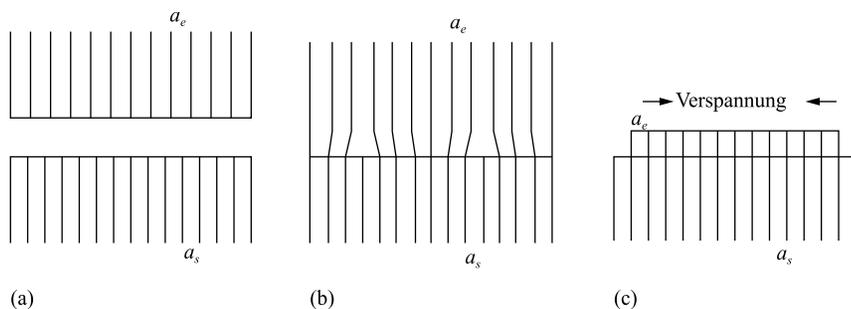
Um die kritische Dicke dieser verspannten Schicht abzuschätzen, betrachten wir den heteroepitaktischen Prozess mit zunehmendem Abstand vom Substrat. Anfangs folgt die Epitaxieschicht dem Gitter des Substrats, aber mit zunehmender Filmdicke baut sich eine steigende Dehnungsenergie auf. Schließlich hat sich im Film zu viel Spannung aufgebaut, um sie aufrechtzuerhalten, und er geht in einen entspannten Zustand, d. h. von Abb. 1.37c zu Abb. 1.33b, über. Die Gitterfehlanpassung ist definiert als

$$\Delta \equiv \frac{|a_e - a_s|}{a_e}, \tag{1.132}$$

wobei  $a_e$  und  $a_s$  die Gitterkonstanten der Epitaxieschicht bzw. des Substrats sind. Die kritische Dicke folgt einer empirischen Formel, die durch

$$t_c \approx \frac{a_e}{2\Delta} \approx \frac{a_e^2}{2|a_e - a_s|} \tag{1.133}$$

gegeben ist. Ein typischer Wert für die kritische Dicke liegt bei einer Fehlanpassung von 2% und einem  $a_e$  von  $5 \text{ \AA}$  bei etwa 10 nm. Diese Technik der graduell verspann-



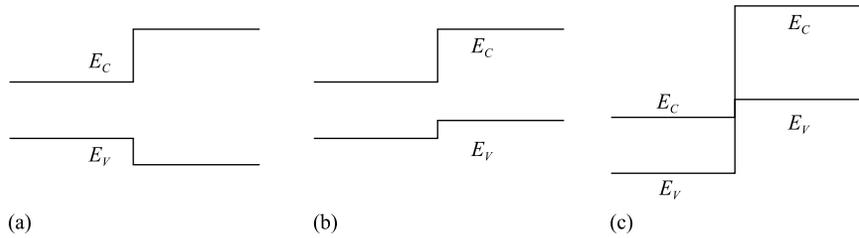
**Abb. 1.37** (a) Zwei getrennte Materialien mit leicht unterschiedlichen Gitterkonstanten  $a_s$  und  $a_e$  ( $s$  steht für Substrat,  $e$  für Epitaxie). (b) Heteroepitaxie dieser Materialien mit dicker, entspannter Epitaxieschicht mit Versetzungen an der Grenzfläche. (c) Mit einer dünnen, verspannten Epitaxieschicht ohne Versetzungen. Die epitaktische Gitterkonstante  $a_e$  wird so verspannt, dass sie der Gitterkonstante des Substrats  $a_s$  folgt.

ten Heteroepitaxie bringt einen zusätzlichen Freiheitsgrad und erlaubt die Verwendung einer breiteren Palette von Materialien. Das hatte in der Vergangenheit bereits große Auswirkungen auf die Anwendungen von Heterostrukturen in neuartigen Bauelementen und auf die Verbesserung ihrer Leistungen.

Zusätzlich zu den unterschiedlichen Bandlücken sind auch die Elektronenaffinitäten in diesen Halbleitern unterschiedlich und müssen bei Bauelementanwendungen berücksichtigt werden. Dies führt zu unterschiedlichen Möglichkeiten für die relative Lage der Leitungs- und Valenzbänder ( $E_C$  und  $E_V$ ) an der Grenzfläche. Entsprechend ihrer Bandanordnung können, wie in Abb. 1.34 dargestellt, Heteroübergänge in drei Gruppen eingeteilt werden: (1) Typ-I- oder gespreizter Heteroübergang, (2) Typ-II- oder gestaffelter Heteroübergang und (3) Typ-III-Heteroübergang mit unterbrochener Bandlücke. In einem gespreizten (engl. straddled) Typ-I-Heteroübergang hat ein Material sowohl ein niedrigeres Leitungsband  $E_C$  als auch ein höheres Valenzband  $E_V$  und damit eine kleinere Bandlücke. Bei einem Typ-II-Heteroübergang (gestaffelt) sind die Orte mit niedrigerem Leitungsband  $E_C$  und höherem Valenzband  $E_V$  verschoben, sodass Elektronen, die bei niedrigerem  $E_C$  gesammelt werden, und Löcher, die bei höherem  $E_V$  gesammelt werden, sich an unterschiedlichen Orten aufhalten. Ein Typ-III-Heteroübergang (engl. broken gap) ist ein Sonderfall des Typ-II-Übergangs, bei dem das Leitungsband  $E_C$  auf einer Seite niedriger liegt als das Valenzband  $E_V$  auf der anderen Seite. Das Leitungsband überlappt also das Valenzband an der Grenzfläche, daher der Name „unterbrochene Bandlücke“ oder broken gap.

**Quantentöpfe und Übergitter** Eine wichtige Anwendung von Heteroübergängen ist die Nutzung von  $\Delta E_C$  und  $\Delta E_V$  zum Aufbau von Barrieren für Ladungsträger. Ein Quantentopf (engl. quantum well), oder auch Quantentrog genannt, wird durch zwei Heteroübergänge oder drei Materialschichten gebildet, sodass die mittlere Schicht das am tiefsten liegende Leitungsband  $E_C$  für einen Elektronentopf oder das am höchsten liegende Valenzband  $E_V$  für einen Löchertopf besitzt. Ist der Elektronentopf sehr dünn, wird von einem 2-D-System gesprochen, dessen elektronische Eigenschaften sich von einem 3-D-System deutlich unterscheiden. Da sich Elektronen in einem 3-D-Volumenhalbleiter in alle Richtungen frei bewegen können, ist ihre Energie über der Leitungsbandkante kontinuierlich verteilt, was zu folgender Beziehung zwischen Energie und Impuls führt, siehe Gl. (1.8):

$$E - E_C = \frac{\hbar^2}{2m_e^*} (k_x^2 + k_y^2 + k_z^2) . \quad (1.134)$$



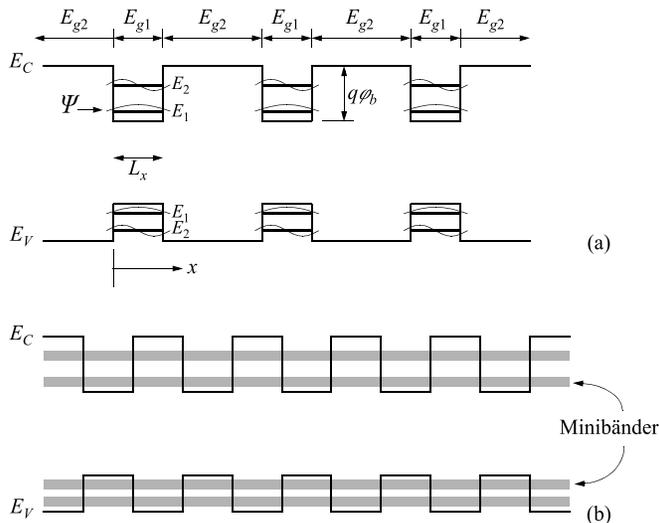
**Abb. 1.38** Klassifizierung der Heteroübergänge. (a) Typ-I- oder gespreizter Heteroübergang. (b) Typ-II- oder gestaffelter Heteroübergang. (c) Typ-III- oder „broken-gap“-Heteroübergang.

Salopp ausgedrückt könnte man auch sagen, dass sich ein 3-D-Elektron durch nichts von einem Tennisball unterscheidet, denn der Tennisball folgt exakt der gleichen Energie-Impuls-Beziehung wie das Elektron. In einem Quantentopf ist die Bewegung der Ladungsträger in einer Richtung, z. B. in der  $x$ -Richtung, komplett unterbunden, in den beiden anderen Dimensionen können sich die Elektronen aber noch frei bewegen. Dennoch ist die Energie des Elektrons innerhalb dieses Topfes in Bezug auf die  $x$ -Richtung nicht null, sondern endlich und quantisiert, und eine Nullpunktsenergie gibt es noch dazu. Den Tennisball kann man zwar auch in irgendeinen Kochtopf stecken, aber irgendwelche Energien hat dieser dort drinnen ganz sicher nicht. Er bleibt einfach mit der Energie null am Boden des Topfes liegen, weil es keinen Grund gibt, warum er im Topf durch die Gegend schweben sollte und das noch dazu auf irgendwelchen fix vorgegebenen „Flugflächen“ (2-D-Subbänder, siehe Gl. (1.136)).

Die wichtigsten Parameter für einen Quantentopf sind die Breite  $L_x$  und die Tiefe des Topfes  $\phi_b$ . Das Banddiagramm in Abb. 1.39a zeigt, dass sich die Potentialbarrieren aus dem Leitungsband- und dem Valenzbandversatz ( $\Delta E_C$  und  $\Delta E_V$ ) ergeben. Die Lösung für die Wellenfunktion der Schrödinger-Gleichung im Quantentopf ist

$$\psi(x) = \sin\left(\frac{i\pi x}{L_x}\right), \tag{1.135}$$

wobei  $i$  eine ganze Zahl ist. Man beachte, dass an den Topfgrenzen  $\psi$  nur dann wirklich null ist, wenn  $\phi_b$  unendlich groß ist. Außerdem ist zu beachten, dass ganz unbedingt die Beziehung  $0 \leq x \leq L_x$  gelten muss. Ist das nicht der Fall, sieht die Wellenfunktion etwas komplizierter aus. Bei endlichem  $\phi_b$  können die Ladungsträger mit endlicher Wahrscheinlichkeit aus dem Topf durch Tunneleffekte entkommen. Dies spielt für die Bildung eines Übergitters eine wichtige Rolle und wird später im Detail diskutiert.



**Abb. 1.39** Banddiagramm einer (a) Heterostruktur aus multiplen Quantentöpfen und (b) einem sogenannten Übergitter aus unendlich vielen Quantentöpfen.

Die Forderung, dass die Wellenfunktion an den Wänden des Potentialtopfs null sein muss, führt zur Quantisierung der Energie in  $x$ -Richtung, welche (bezogen auf die Bandkanten) folgende Werte besitzt:

$$E_i = \frac{\hbar^2 \pi^2 i^2}{2m^* L_x^2}. \quad (1.136)$$

$i$  ist eine ganze Zahl. Vorsicht, diese Lösungen gelten für endlich tiefe Töpfe nur sehr bedingt und zeigen für höhere  $i$ -Werte in der Nähe der Topfoberkante deutliche Abweichungen von der exakten Lösung. Des Weiteren kann mit  $L_x$  als einziger Variable ein Quantentopf im realen Halbleiter nur ungenügend gut definiert werden. Als Mindestanforderung für reale Quantentöpfe im Halbleiter muss die Quantisierungsenergie  $\hbar^2 \pi^2 / (2m^* L_x^2)$  auf jeden Fall viel größer als  $kT$  sein. Zusätzlich sollte  $L_x$  kleiner sein als die mittlere freie Weglänge und auch kleiner als die de Broglie-Wellenlänge. Hinweis: Man beachte, dass die de Broglie-Wellenlänge  $\lambda = h / (2m^* E)^{1/2}$  eine ähnliche Form wie  $L_x$  von Gl. (1.136) hat. Interessant ist auch, dass sich die Ladungsträger im 2-D-System nicht mehr an den Bandkanten  $E_C$  oder  $E_V$ , sondern nur noch in den Subbändern mit den Energien aus Gl. (1.136) befinden. Dadurch wird die effektive Bandlücke für Interbandübergänge (also Übergänge zwischen Valenz- und Leitungsband) innerhalb des Quantentopfes größer als die Bandlücke  $E_g$  im Volumenmaterial.

Wenn die Quantentöpfe durch dicke Barrieren voneinander getrennt sind, gibt es keine Kommunikation zwischen den Töpfen und das System kann als Kette identischer Quantentöpfe beschrieben werden. Wenn jedoch die Barrieren zwischen den Töpfen dünner werden, beginnen die Wellenfunktionen sich zu überlappen, wodurch ein sogenanntes Heterostrukturübergitter entsteht. Das Übergitter hat zwei wesentliche Unterschiede zu einer einfachen Kette aus Quantentöpfen: (1) die diskreten Niveaus im einzelnen Topf erweitern sich zu Minibändern (Abb. 1.39b), welche durch die ganze Struktur reichen, und (2) Minibänder gibt es auch oberhalb der Barrieren und dadurch haben die Elektronen auch oberhalb der Barrieren keine kontinuierliche Energieverteilung. Dieser Effekt verwäscht sich allerdings mit zunehmender Energie und wird bei hohen Energien schließlich ganz vernachlässigbar.

Der Übergang von mehreren Quantentöpfen zu einem Übergitter ist analog zur Bildung eines regelmäßigen Kristallgitters durch eine geordnete Anordnung von Atomen mit geringem Abstand im Raum. Die isolierten Atome haben diskrete Niveaus, während sich in einem Kristallgitter diese diskreten Niveaus in das kontinuierliche Leitungsband und das Valenzband des Halbleiters verwandeln.

Ein weiterer Ansatz zur Bildung von Quantentöpfen und Übergittern ist die räumliche Modulation der Dotierung [87], wobei die Potentialbarrieren durch Raumladungszonen gebildet werden (Abb. 1.40a). Die Form der Barrieren ist in diesem Fall eher parabolisch als rechteckig. Es gibt zwei interessante Merkmale in dieser modulierten  $n$ - $i$ - $p$ - $i$ -Mehrfach-Quantentopf-Struktur. Zunächst einmal sind das Leitungsbandminimum und das Valenzbandmaximum gegeneinander verschoben, was bedeutet, dass sich Elektronen und Löcher an verschiedenen Orten ansammeln. Dies führt zu einer minimalen Elektron-Loch-Rekombination und einer sehr langen Ladungsträgerlebensdauer, die, ganz ähnlich wie bei einer Typ-II-Heterostruktur, um viele Größenordnungen höher ist als im regulären Material. Zweitens ist die effektive Bandlücke, die zwischen dem ersten quantisierten Niveau für die Elektronen und Löcher liegt, im Vergleich zum Grundmaterial (und auch im Vergleich zu Heterostrukturübergittern) reduziert. Diese durch die Dotierung abstimmbare effektive Bandlücke ermöglicht eine Lichtemission und Absorption bei längeren Wellenlängen. Das ist insofern einzigartig, weil wir hier eine indirekte Bandlücke im „realen Raum“ haben, und nicht im  $k$ -Raum, wie es sonst der Fall ist. Liegen die Dotierungsquantentöpfe

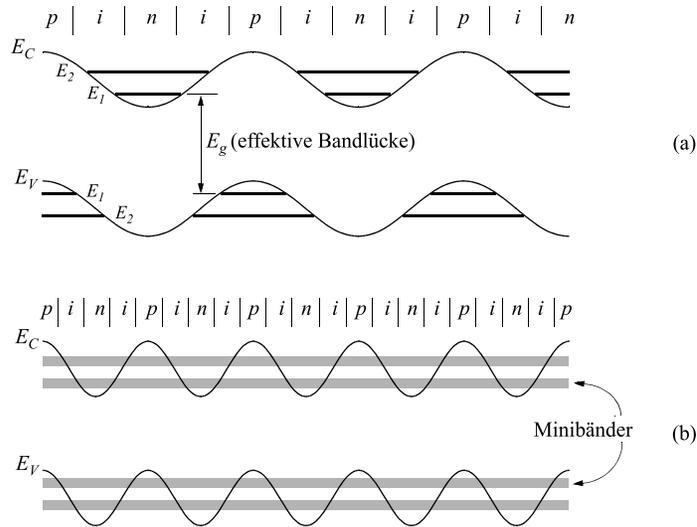


Abb. 1.40 Banddiagramm für (a) multiple *n-i-p-i*-Quantentöpfe und (b) *n-i-p-i*-Übergitter.

nahe beieinander, bildet sich, ähnlich wie bei Heterostrukturen, ein Dotierungs-*(n-i-p-i)*-Übergitter (Abb. 1.40b).

**Quantendrähte und Quantenpunkte** Die physikalischen Dimensionen eines Halbleiters haben erhebliche Auswirkungen auf die elektronischen Eigenschaften eines Bauelements, wenn die Dimensionen auf die Größenordnung der de Broglie-Wellenlänge reduziert sind. Der Einschluss von 3-D- und 2-D-Ladungsträgern kann weiter auf Linien (1-D-Systeme) und kleine Inseln (0-D-Systeme) ausgedehnt werden. Im üblichen Jargon sind diese Systeme auch als Quantendrähte und Quantenpunkte bekannt. Eine der wichtigsten Auswirkungen ist die veränderte Zustandsdichte als Funktion der Energie,  $N(E)$ . Je nach Dimension des Systems hat  $N(E)$  sehr unterschiedliche Formen. Die qualitativen Formen von  $N(E)$  für Volumenhalbleiter, Quantentopf, Quantendraht und Quantenpunkt sind in Abb. 1.41 dargestellt. Für ein 3-D-System wurde die Zustandsdichte bereits früher ausgerechnet (Gl. (1.12)) und man erhielt

$$N(E) = \frac{m^* \sqrt{2m^* E}}{\pi^2 \hbar^3}. \quad (1.137)$$

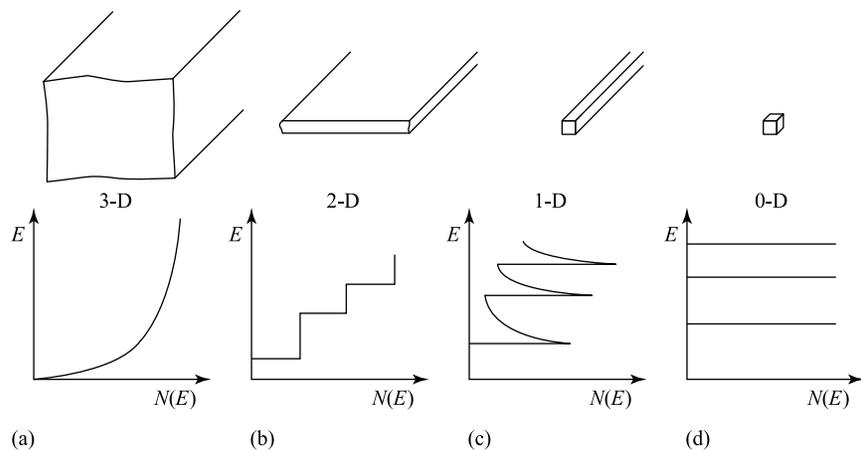


Abb. 1.41 Zustandsdichte  $N(E)$  für (a) Volumenhalbleiter (3-D), (b) Quantentöpfe (2-D), (c) Quantendraht (1-D) und (d) Quantenpunkt (0-D).

Die Zustandsdichte in einem 2-D-System (Quantentopf) ist eine Stufenfunktion der Form

$$N(E) = \frac{m^*i}{\pi\hbar^2L_x}. \quad (1.138)$$

Die Zustandsdichte in einem 1-D-System (Quantendraht) lautet

$$N(E) = \frac{\sqrt{2m^*}}{\pi\hbar L_x L_y} \sum_{i,j} (E - E_{i,j})^{-1/2}, \quad (1.139)$$

wobei

$$E_{i,j} = \frac{\hbar^2\pi^2}{2m^*} \left( \frac{i^2}{L_x^2} + \frac{j^2}{L_y^2} \right). \quad (1.140)$$

Die Zustandsdichte in einem 0-D-System (Quantenpunkt) wird mit einer  $\delta$ -Funktion beschrieben,

$$N(E) = \frac{2}{L_x L_y L_z} \sum_{i,j,k} \delta(E - E_{i,j,k}), \quad (1.141)$$

wobei

$$E_{i,j,k} = \frac{\hbar^2\pi^2}{2m^*} \left( \frac{i^2}{L_x^2} + \frac{j^2}{L_y^2} + \frac{k^2}{L_z^2} \right). \quad (1.142)$$

Da die Ladungsträgerkonzentration und ihre Energieverteilung durch das Integral über dem Produkt aus Zustandsdichte und der Fermi-Dirac-Verteilung gegeben ist, gewinnen die Zustandsdichtefunktionen für den Betrieb von Halbleiterbauelementen wesentlich an Bedeutung, sobald die physikalischen Abmessungen eines Bauelements in irgendeiner Dimension die Größenordnung der de Broglie-Wellenlänge ( $\approx 20$  nm) erreicht.

Werden Elektronen im Halbleiter in allen drei Raumrichtungen eingesperrt, kommt es zur Bildung von Quantenpunkten oder „künstlichen Atomen“. Die Abb. 1.42a,b zeigen rasterelektronenmikroskopische Bilder (REM-Bilder) eines Arrays von Si-Nanoscheiben auf einem  $\text{SiO}_2$ -Substrat in zwei verschiedenen Vergrößerungen. Verglichen mit Quantentrögen und Quantendrähten bieten Quantenpunkte einen zusätzlichen Freiheitsgrad zur Modulation der Energieniveaus und der energetischen Breite der Minibänder. Zum Beispiel wird die Quantisierungsenergie in Quantendrähten nur durch den Radius bestimmt, wogegen in Quantenpunkten sowohl der

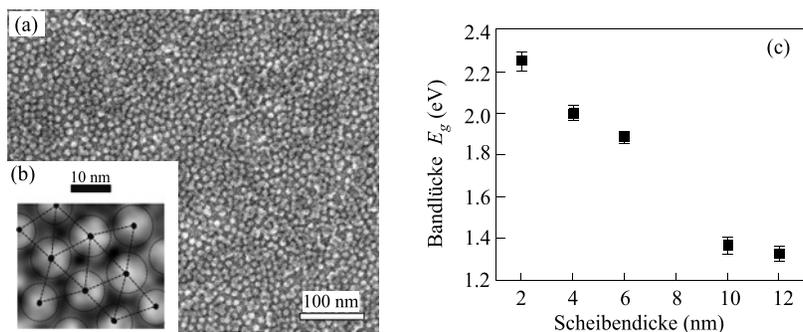
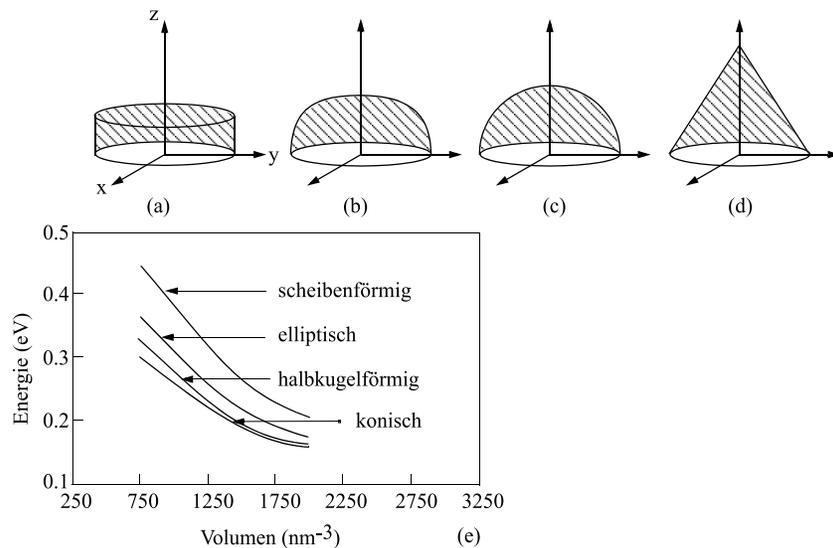


Abb. 1.42 (a) REM-Bild eines Si-Nanodisk-Arrays. (b) Vergrößerte Ansicht. (c) Abhängigkeit der Bandlücke  $E_g$  von der Dicke der Nanoscheibe (basierend auf [88]).



**Abb. 1.43** InAs/GaAs-Quantenpunkte in vier verschiedenen Formen: (a) Scheibe, (b) elliptische Linse, (c) Schnittkugellinse und (d) konische Form. Alle Quantenpunkte bestehen aus InAs und sind in einer GaAs-Matrix eingebettet. (e) Energieniveaus des Grundzustands als Funktion des Volumens der Quantenpunkte (basierend auf [93]).

Radius als auch die Dicke die Energieniveaus beeinflussen [88, 89]. Abbildung 1.42c zeigt die Abhängigkeit der Bandlücke  $E_g$  von der Dicke der Si-Nanoscheibe, wobei der Durchmesser der Si-Nanoscheibe immer 10 nm beträgt. Die Bandlücke  $E_g$  nimmt mit steigender Dicke der Quantenpunkte aufgrund der schlechteren Quantisierung ab [88–94].

Abbildung 1.43 zeigt die Energieniveaus von InAs/GaAs-Quantenpunkten als Funktionen des Volumens. Der Radius der Quantenpunkte an der Basis ist für alle Formen gleich 10 nm. Das Volumen der Quantenpunkte und der Radius der Basis wurden aus verfügbaren experimentellen Daten entnommen. Von den Quantenpunkten in den Abb. 1.43a–d reagieren die Quantenscheiben am empfindlichsten auf die Variation des Volumens, am wenigsten empfindlich sind die konischen Quantenpunkte. Dies ist keine Überraschung, da die Wellenfunktionen der Elektronen bei fixem Volumen und Radius der Quantenscheiben nur auf die Scheibendicke reagieren. Wenn das Volumen der Quantenpunkte zunimmt, konvergieren die Grundzustandsenergien der verschiedenen Quantenpunkte. Das ist auch keine Überraschung, da sich deren Geometrie mit steigender Dicke mehr und mehr dem Grenzfall für Quantendrähte mit gleichem Durchmesser nähert. Die Abhängigkeit der Energieniveaus vom Volumen ist allerdings ganz anders als das allgemein zitierte  $V^{-2/3}$ -Verhalten. Der Exponent variiert über einen weiten Bereich und hängt stark von der Form der Quantenpunkte ab [93].

## 1.8 Halbleitergrundgleichungen und Anwendungsbeispiele

### 1.8.1 Halbleitergrundgleichungen

Die Grundgleichungen für den Betrieb von Halbleiterbauelementen beschreiben das statische und dynamische Verhalten von Ladungsträgern in Halbleitern unter äußeren Einflüssen, welche eine Abweichung vom thermischen Gleichgewicht verursachen, wie z. B. ein angelegtes Feld oder optische Anregung [37]. Die grundlegenden Gleichungen können in drei Gruppen eingeteilt werden: Die Poisson-Gleichung, Stromdichtegleichungen und Kontinuitätsgleichungen. Um die Auswirkung der ungleichmäßigen Gittertemperatur zu berücksichtigen, muss die Wärmeflussgleichung

zusammen mit den oben genannten partiellen Differenzialgleichungen (engl. partial differential equations, PDEs) gelöst werden

**Poisson-Gleichung** Um uns ein wenig Hintergrundwissen über die Poisson-Gleichung zu verschaffen, beschäftigen wir uns am besten zuerst mit den Weisheiten des Herrn Maxwell und seiner Gleichungen. Beginnen wir mit dem gaußschen Gesetz

$$\nabla \cdot \mathcal{D} = \rho(x, y, z), \quad (1.143)$$

mit der dielektrischen Verschiebung  $\mathcal{D}$  und  $\rho$ , der dreidimensionalen Ladungsdichte. Gehen wir davon aus, dass die dielektrische Verschiebung mit dem elektrischen Feld  $\mathcal{E}$  durch ein skalares homogenes Dielektrikum mit der Dielektrizitätskonstante  $\varepsilon_s$  ( $s$  steht für engl. semiconductor) verbunden ist und dass damit die Gleichung  $\mathcal{D} = \varepsilon_s \mathcal{E}$  gilt. Die Beziehung zwischen der magnetischen Flussdichte  $B$  und dem elektrischen Feld ist in den Maxwell-Gleichungen durch

$$\nabla \times \mathcal{E} = -\frac{\partial B}{\partial t} \quad (1.144)$$

gegeben. In der Halbleitertechnik gibt es normalerweise keine Magnetfelder, und damit gibt es deren zeitliche Ableitung natürlich auch nicht, also nehmen wir an, dass

$$\frac{\partial B}{\partial t} \approx 0. \quad (1.145)$$

Aus Gln. (1.144) und (1.145) können wir erkennen, dass die Rotation des elektrischen Feldes null ist. Die Rotation des Gradienten ist auch null, sodass wir ein elektrostatisches Potential für das elektrische Feld einführen können:

$$\mathcal{E} = -\nabla\psi. \quad (1.146)$$

Aus Gln. (1.143) und (1.146) erhalten wir nun die Poisson-Gleichung in einem homogenen Material

$$\nabla \cdot (\varepsilon_s \nabla\psi) = -\rho, \quad (1.147)$$

wobei die Dielektrizitätskonstante  $\varepsilon_s = 11.9\varepsilon_0$  für Si ist (siehe Anhang F). Die elektrische Ladungsdichte in einem Halbleiter ist die Summe der Elektronendichten  $n$ , der Löcherdichten  $p$  und der Dotierungsdichte  $N^*$  für ionisierte Störstellen – als Formel ausgedrückt also  $\rho = q(n - p + N^*)$ . Dabei ist  $q$  die Elementarladung und  $N^*$  die Summe der ionisierten Akzeptoren und Donatoren  $N^* = N_A - N_D$ . Die Poisson-Gleichung (Gl. (1.147)) wird damit zu

$$\nabla^2\psi = \frac{q(n - p + N_A - N_D)}{\varepsilon_s}. \quad (1.148)$$

Mit dem Laplace-Operator in kartesischen Koordinaten erhält man

$$\nabla^2\psi(x, y, z) = \frac{\partial^2\psi}{\partial x^2} + \frac{\partial^2\psi}{\partial y^2} + \frac{\partial^2\psi}{\partial z^2}. \quad (1.149)$$

Die Gl. (1.148) wird gerne zusammen mit den Gleichungen für den Ladungsträgertransport gelöst, die weiter unten (Gln. (1.151) bis (1.152)) diskutiert werden. Gleichung (1.148) berücksichtigt auch lang- und kurzreichweitige Coulomb-Wechselwirkungen. Die dielektrische Verschiebung ist, wie schon erwähnt, mit dem elektrischen Feld durch eine skalare und homogene dielektrische Permittivität verbunden, die aber für piezoelektrische Phänomene, ferroelektrische Phänomene und nicht lineare Optik unbedingt modifiziert werden muss.

Für ein eindimensionales Problem (also in der Praxis fast immer) reduziert sich die Poisson-Gleichung auf die deutlich nützlichere Form

$$\frac{d^2\psi_i}{dx^2} = -\frac{d\mathcal{E}}{dx} = -\frac{\rho}{\epsilon_s} = \frac{q(n - p + N_A - N_D)}{\epsilon_s} \quad (1.150)$$

( $\psi_i \equiv -E_i/q$  ist das Potential;  $E_i$  ist eine potentielle Energie und nicht das elektrische Feld). Diese Gleichung wird häufig in Bauelementen verwendet, um die Potential- und Feldverteilung zu bestimmen, die durch eine Ladungsdichte  $\rho$  innerhalb einer Raumladungszone verursacht wird (siehe z. B.  $p$ - $n$ -Übergang in Kap. 2).

**Stromdichtegleichungen** Zwei Prozesse dominieren die Stromleitung im Halbleiter: Drift und Diffusion. Drift wird durch das elektrische Feld verursacht; die Diffusionskomponente kommt vom Konzentrationsgradienten der Ladungsträger im Material. Die Stromdichtegleichungen lauten:

$$\mathbf{J}_n = q\mu_n n\mathcal{E} + qD_n \nabla n, \quad (1.151a)$$

$$\mathbf{J}_p = q\mu_p p\mathcal{E} - qD_p \nabla p \quad (1.151b)$$

und

$$\mathbf{J}_{\text{cond}} = \mathbf{J}_n + \mathbf{J}_p, \quad (1.152)$$

wobei  $\mathbf{J}_n$  und  $\mathbf{J}_p$  die Elektronen- bzw. Löcherstromdichten sind. Vorsicht: Die fett gedruckten Stromdichten sind hier Vektoren! Die Werte der Elektronen- und Löcherbeweglichkeiten ( $\mu_n$  und  $\mu_p$ ) wurden schon in Abschn. 1.5.1 behandelt. Für Ladungsträger in nicht entarteten Halbleitern werden die Diffusionskonstanten ( $D_n$  und  $D_p$ ) und die Beweglichkeiten ( $\mu_n$  und  $\mu_p$ ) aus den Einstein-Beziehungen,  $D_n = (kT/q)\mu_n$  etc., berechnet. Damit reduzieren sich die Gln. (1.151a) und (1.151b) im eindimensionalen Fall auf

$$J_n = q\mu_n n\mathcal{E} + qD_n \frac{dn}{dx} = q\mu_n \left( n\mathcal{E} + \frac{kT}{q} \frac{dn}{dx} \right) = \mu_n n \frac{dE_{Fn}}{dx}, \quad (1.153a)$$

$$J_p = q\mu_p p\mathcal{E} - qD_p \frac{dp}{dx} = q\mu_p \left( p\mathcal{E} - \frac{kT}{q} \frac{dp}{dx} \right) = \mu_p p \frac{dE_{Fp}}{dx}, \quad (1.153b)$$

wobei  $E_{Fn}$  und  $E_{Fp}$  Quasi-Fermi-Niveaus für Elektronen bzw. Löcher sind. Diese Gleichungen sind für niedrige elektrische Felder gültig. Bei ausreichend hohen Feldern sollte der Ausdruck  $\mu_p \mathcal{E}$  oder  $\mu_n \mathcal{E}$  durch die Sättigungsgeschwindigkeit  $v_s$  ersetzt werden, wobei dann auch die letzten Gleichungen über  $E_{Fn}$  und  $E_{Fp}$  nicht mehr gelten. Man beachte, dass diese Gleichungen auch nicht den Effekt eines extern angelegten Magnetfeldes beinhalten, bei dem der Strom durch den magnetoresistiven Effekt reduziert wird.

**Kontinuitätsgleichungen** Während die obigen Stromdichtegleichungen für stationäre Bedingungen gelten, befassen sich die Kontinuitätsgleichungen mit zeitabhängigen Phänomenen wie Injektion, Generation und Rekombination auf niedrigem Niveau. Qualitativ gesehen ist die Nettoveränderung der Ladungsträgerkonzentration die Differenz zwischen Erzeugung und Rekombination, zuzüglich des Nettostroms, der in die interessante Region hinein- und wieder herausfließt. Die Kontinuitätsgleichungen sind:

$$\frac{\partial n}{\partial t} = G_n - U_n + \frac{1}{q} \nabla \cdot \mathbf{J}_n \quad \text{und} \quad (1.154a)$$

$$\frac{\partial p}{\partial t} = G_p - U_p - \frac{1}{q} \nabla \cdot \mathbf{J}_p. \quad (1.154b)$$

$G_n$  und  $G_p$  sind die Elektronen- bzw. Löchergenerationsraten ( $\text{cm}^{-3}\cdot\text{s}^{-1}$ ), die durch äußere Einflüsse wie durch optische Anregung oder Stoßionisation bei großen elektrischen Feldern verursacht werden. Die Rekombinationsraten,  $U_n = \Delta n / \tau_n$  und  $U_p = \Delta p / \tau_p$ , wurden bereits in Abschn. 1.5.4 diskutiert. Für den eindimensionalen Fall der schwachen Injektion bekommt man dann aus den Gln. (1.154a) und (1.154b) die Beziehungen

$$\frac{\partial n_p}{\partial t} = G_n - \frac{n_p - n_{p0}}{\tau_n} + n_p \mu_n \frac{\partial \mathcal{E}}{\partial x} + \mu_n \mathcal{E} \frac{\partial n_p}{\partial x} + D_n \frac{\partial^2 n_p}{\partial x^2} \quad (1.155a)$$

und

$$\frac{\partial p_n}{\partial t} = G_p - \frac{p_n - p_{n0}}{\tau_p} - p_n \mu_p \frac{\partial \mathcal{E}}{\partial x} - \mu_p \mathcal{E} \frac{\partial p_n}{\partial x} + D_p \frac{\partial^2 p_n}{\partial x^2}. \quad (1.155b)$$

Sobald die Abmessungen elektronischer Bauelemente in den tiefen Submikrometerbereich schrumpfen, verletzen relativ große elektrische Felder und die damit verbundene starke räumliche Variation der Ladungsträgerkonzentrationen die Annahme eines lokalen thermischen Gleichgewichts im Drift-Diffusions-Modell. Die Ladungsträger in diesen Bauelementen werden durch das angelegte  $\mathcal{E}$ -Feld so stark aufgeheizt, dass es zu einer Nichtgleichgewichtsverteilung der Ladungsträgerenergie kommt. Große Gradienten im  $\mathcal{E}$ -Feld führen zusätzlich zu nicht lokalen Effekten, da die Träger nicht genügend Zeit oder zu wenig Abstand voneinander haben, um ihren Gleichgewichtszustand zu erreichen. Inzwischen wurde jedoch viel Arbeit in fortgeschrittene Transportmodelle investiert, die Effekte mit heißen Elektronen in Strukturen im tiefen Submikrometerbereich genau berücksichtigen können. Entwickelt wurden diese Modelle auf Basis der halbklassischen Boltzmann-Transportgleichung (engl. Boltzmann transport equation, BTE) und dem hydrodynamischen Modell (HD), das aus einer Reihe von Erhaltungsgleichungen für die Teilchenzahl, den mittleren Impuls und die mittlere Energie besteht [95, 96]. Die Ableitungen des BTE- und HD-Modells sind in Anhang I aufgeführt. Im Gegensatz zum Drift-Diffusions-Modell sind die moderneren Modelle in der Lage, nicht stationäre Transportphänomene zu beschreiben; so kann z. B. die mittlere Energie zur Modellierung von heißen Trägerprozessen verwendet werden, was genauer ist als der Zugang über lokale elektrische Felder.

**Gittertemperaturgleichung** Für den Entwurf von Halbleiterleistungsbaulementen ist es notwendig, die Wechselwirkung elektrothermischer Phänomene zu berücksichtigen. Zum Beispiel ist das thermische Davonlaufen (engl. thermal runaway) eine der wichtigen thermischen Fragen in Halbleiterbaulementen. Um Selbsterhitzungseffekte und inhomogene Temperaturverteilungen in Halbleiterbaulementen zu beschreiben, muss die Wärmeflussgleichung, die auch als Gittertemperaturgleichung bekannt ist, gelöst werden [9]:

$$\rho_m C_p \frac{\partial T}{\partial t} - \nabla \cdot (\kappa \nabla T) = Q, \quad (1.156)$$

wobei  $Q$  die Wärmequelle,  $\rho_m$  die Massendichte,  $C_u$  die spezifische Wärme und  $\kappa$  ( $\text{W}/\text{cm}\cdot\text{K}$ ) die Wärmeleitfähigkeit von Halbleitermaterialien ist. Die Werte von  $\rho_m$  und  $C_p$  sind in Tab. 1.3 aufgelistet. Die Ausdrücke der Wärmeleitfähigkeit von Si und GaAs sind durch  $(1.65 \times 10^{-6} T^2 + 1.56 \times 10^{-3} T + 0.03)^{-1}$  und  $0.44 \times (T/300)^{-1.25}$  gegeben. Die Wärmeerzeugung unter Berücksichtigung der Auswirkung der Reduktion

**Tab. 1.3** Spezifische Wärme und Massendichte für Si und GaAs bei 300 K. Warum die Autoren des Buches die Einheit der Dichte ( $\text{kg m}^{-3}$ ) unbedingt in der Form ( $\text{V As}^{-3} \text{m}^{-5}$ ) hinschreiben wollten, konnte Ihr Übersetzer aber leider nicht herausfinden.

Halbleiter	$C_p$ ( $\text{m}^2 \text{s}^{-2} \text{K}^{-1}$ )	$\rho_m$ ( $\text{V As}^3 \text{m}^{-5}$ )
Si	703	2328
GaAs	351	5316

der Bandlücke berechnet sich durch

$$Q = \frac{1}{q} J_n \cdot \nabla E_C + \frac{1}{q} J_p \cdot \nabla E_V + E_g(U - G), \quad (1.157)$$

wobei die zusätzlichen Terme  $qnD_n \nabla T / 2T$  und  $-qpD_p \nabla T / 2T$  in die Gln. (1.151a) und (1.151b) eingefügt werden müssen. Gleichung (1.155) kann mit dem Drift-Diffusions-Modell gelöst werden, um die sinkende Beweglichkeit der Ladungsträger aufgrund der Gittertemperatur zu berücksichtigen. Zur Messung der Gittertemperatur von Si wird gerne die klassische Flugzeitmethode während der gepulsten Laserbestrahlung verwendet [97, 98].

### 1.8.2 Anwendungsbeispiele

In diesem Abschnitt werden Beispiele zur Verwendung der Kontinuitätsgleichungen zur Beschreibung der zeit- und ortsabhängigen Konzentration von Überschussladungsträgern diskutiert. Überschüssige Ladungsträger können durch optische Anregung oder elektrische Injektion von einem nahegelegenen Injektionspunkt aus erzeugt werden. In den folgenden Beispielen verwenden wir der Einfachheit halber aber lieber die optische Anregung.

**Zerfall der überschüssigen Ladungsträger mit der Zeit** Betrachten wir eine  $n$ -Typ-Probe, die von oben beleuchtet wird (Abb. 1.38a), wobei überall in der gesamten Probe Elektron-Loch-Paare mit einer gleichmäßigen Generationsrate  $G_p$  erzeugt werden. Die Probendicke sei viel kleiner als die Absorptionslänge  $1/\alpha$ , womit eine räumliche Abhängigkeit der Erzeugung nicht vorhanden ist. Die Randbedingungen sind  $\mathcal{E} = \partial \mathcal{E} / \partial x = 0$  und  $\partial p_n / \partial x = 0$ . Damit bekommen wir aus Gl. (1.155b)

$$\frac{dp_n}{dt} = G_p - \frac{p_n - p_{n0}}{\tau_p}. \quad (1.158)$$

Im stationären Gleichgewicht, also bei  $\partial p_n / \partial t = 0$  und  $\mathcal{E} = \partial \mathcal{E} / \partial x = 0$  ist

$$p_n - p_{n0} = \tau_p G_p = \text{konstant}. \quad (1.159)$$

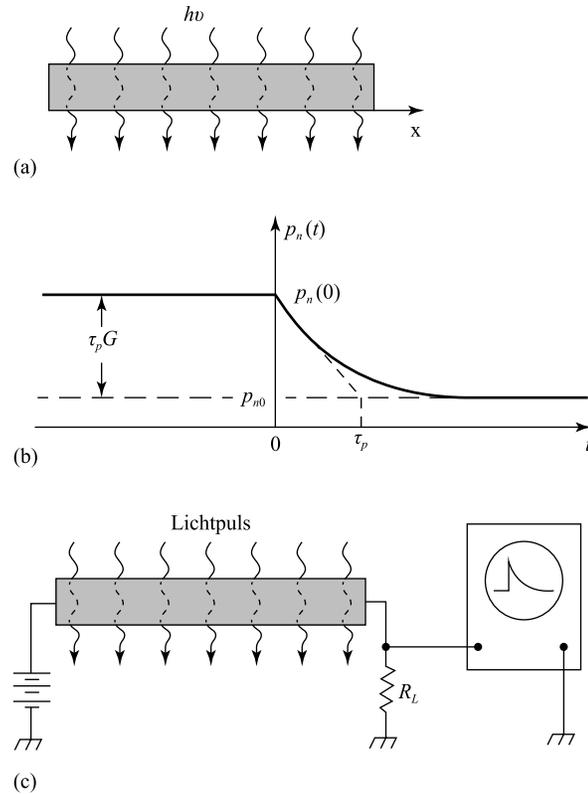
Wenn zu einem beliebigen Zeitpunkt, sagen wir bei  $t = 0$ , das Licht plötzlich ausgeschaltet wird, lautet die Differenzialgleichung

$$\frac{dp_n}{dt} = -\frac{p_n - p_{n0}}{\tau_p}. \quad (1.160)$$

Mit den Randbedingungen  $p_n(t = 0) = p_{n0} + \tau_p G_p$  aus Gl. (1.159) und  $p_n(\infty) = p_{n0}$ , lautet die Lösung

$$p_n(t) = p_{n0} + \tau_p G_p \exp\left(-\frac{t}{\tau_p}\right). \quad (1.161)$$

Abbildung 1.44b zeigt den zeitlichen Verlauf von  $p_n$ .



**Abb. 1.44** Zerfall von optisch angeregten Ladungsträgern. (a)  $n$ -Typ-Probe unter konstanter Beleuchtung. (b) Zerfall der Minoritätsladungsträger (Löcher) mit der Zeit. (c) Schematischer Versuchsaufbau zur Messung der Lebensdauer von Minoritätsladungsträgern (basierend auf [74]).

Das obige Beispiel stellt das wesentliche Prinzip der Stevenson-Keyes-Methode zur Messung der Lebensdauer von Minoritätsladungsträgern dar [74]. Abbildung 1.44c zeigt den schematischen Aufbau des Experiments. Die durch die Lichtpulse gleichmäßig in der gesamten Probe erzeugten überschüssigen Ladungsträger verursachen einen momentanen Anstieg der Leitfähigkeit und damit des Stroms. Während der düsteren Zeiten, d. h. zwischen den einzelnen Lichtpulsen, kann auch bei moderater Hintergrundbeleuchtung das Abklingen der Photoleitfähigkeit auf einem Oszilloskop direkt beobachtet werden. Die Auswertung dieser Daten liefert dann die Lebensdauer der Minoritätsladungsträger.

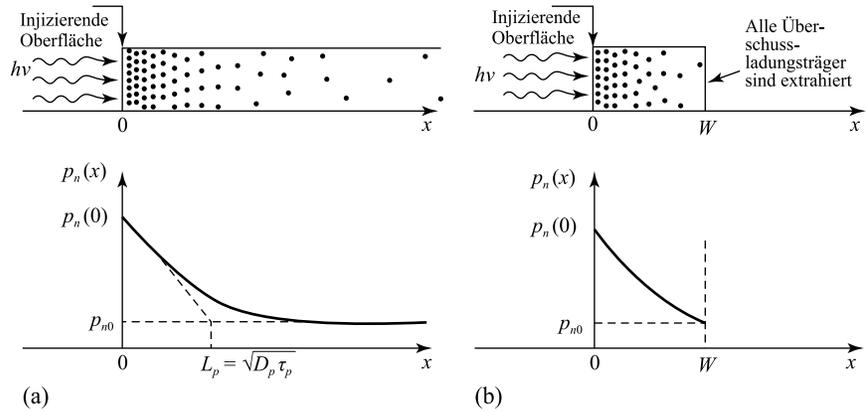
**Zerfall der überschüssigen Ladungsträger mit der Entfernung** Abbildung 1.45a zeigt ein weiteres einfaches Beispiel, bei dem überschüssige Ladungsträger von einer Seite der Probe injiziert werden. Dies kann z. B. durch hochenergetische Photonen erreicht werden, die nur an der Oberfläche Elektron-Loch-Paare erzeugen, oder in anderen Worten, die Absorptionslänge liegt jetzt weit unter der Probendicke. In Abb. 1.32 sieht man, dass für  $h\nu = 3.5 \text{ eV}$  der Absorptionskoeffizient etwa  $10^6 \text{ cm}^{-1}$  beträgt, d. h., die Lichtintensität nimmt auf einer Länge von 10 nm um den Faktor  $e$  ab.

Im stationären Zustand gibt es damit einen Konzentrationsgradienten nahe der Oberfläche. Die Differenzialgleichung für eine  $n$ -Typ-Probe ohne angelegte Spannung lautet gemäß Gl. (1.155b)

$$\frac{\partial p_n}{\partial t} = 0 = -\frac{p_n - p_{n0}}{\tau_p} + D_p \frac{\partial^2 p_n}{\partial x^2}. \quad (1.162)$$

Die Randbedingungen in unserem Problem sind hier  $p_n(x=0) = \text{konstant}$  und  $p_n(\infty) = p_{n0}$ . Die Lösung von  $p_n(x)$  ist

$$p_n(x) = p_{n0} + [p_n(0) - p_{n0}] \exp\left(-\frac{x}{L_p}\right), \quad (1.163)$$



**Abb. 1.45** Konstante Ladungsträgerinjektion von einer Seite: (a) einseitig unendlich lange Probe, (b) Probe mit der Länge  $W$ .

wobei  $L_p = (D_p \tau_p)^{1/2}$  die Diffusionslänge ist (Abb. 1.45a). Die Maximalwerte von  $L_p$  und  $L_n$  liegen in der Größenordnung von 1 cm in Silizium, aber nur in der Größenordnung von  $10^{-2}$  cm in GaAs.

Besonders wichtig ist der Fall, bei dem als zweite Randbedingung alle überschüssigen Ladungsträger am Ende der Probe bei  $x = W$  extrahiert werden, oder in anderen Worten:  $p_n(W) = p_{n0}$ . Das sieht zwar jetzt nur wie eine unnötige Rechenübung aus, aber etwas weiter hinten im Text wird man erkennen, dass Gl. (1.165) dringend zur Berechnung der Stromverstärkung in bipolaren Transistoren benötigt wird (Kap. 5). Zuerst aber brauchen wir für Gl. (1.162) eine neue Lösung, nämlich

$$p_n(x) = p_{n0} + [p_n(0) - p_{n0}] \left\{ \frac{\sinh[(W-x)/L_p]}{\sinh(W/L_p)} \right\}. \quad (1.164)$$

Dieses Ergebnis ist in Abb. 1.45b dargestellt. Die Stromdichte bei  $x = W$  wird schließlich mit Gl. (1.153b) berechnet:

$$J_p = -qD_p \left. \frac{dp}{dx} \right|_W = \frac{qD_p [p_n(0) - p_{n0}]}{L_p \sinh(W/L_p)}. \quad (1.165)$$

Diese Gleichung bitte nicht vergessen, sie wird bei der Berechnung der Stromverstärkung des Bipolartransistors wieder gebraucht.

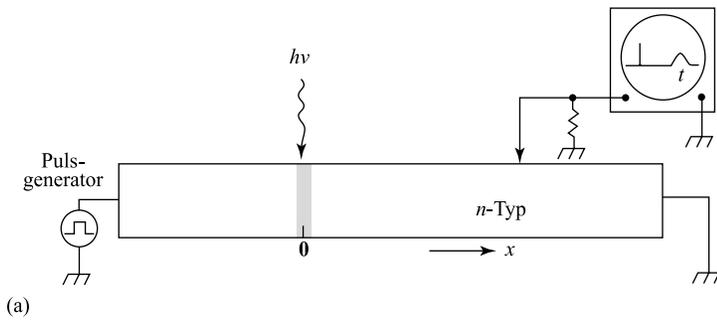
**Zerfall der überschüssigen Ladungsträger mit der Zeit und der Entfernung** Wenn man mit Lichtpulsen lokal überschüssige Ladungsträger in einem Halbleiter erzeugt (Abb. 1.46a), lautet die Transportgleichung nach dem Impuls ohne Vorspannung (Randbedingung:  $G_p = \mathcal{E} = \partial \mathcal{E} / \partial x = 0$ ):

$$\frac{\partial p_n}{\partial t} = -\frac{p_n - p_{n0}}{\tau_p} + D_p \frac{\partial^2 p_n}{\partial x^2}. \quad (1.166)$$

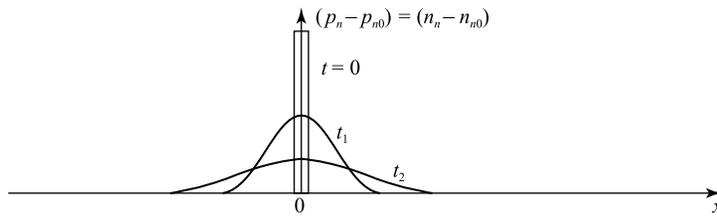
Als Lösung erhält man

$$p_n(x, t) = \frac{N'}{\sqrt{4\pi D_p t}} \exp\left(-\frac{x^2}{4D_p t} - \frac{t}{\tau_p}\right) + p_{n0}, \quad (1.167)$$

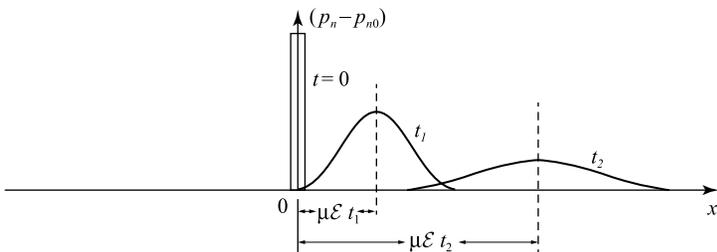
wobei  $N'$  die Anzahl der anfangs pro Flächeneinheit erzeugten Elektronen oder Löcher ist. Abbildung 1.46b zeigt diese Lösung, wobei die Ladungsträger vom Injektionspunkt weg diffundieren und auch rekombinieren. Der Rekombinationsfall ist leicht zu erkennen, denn hier verringert sich die Fläche unter der Kurve.



(a)



(b)



(c)

**Abb. 1.46** Transiente und stationäre Trägerdiffusion nach einem lokalisierten Lichtpuls: (a) experimenteller Aufbau, (b) Ladungsverteilung ohne angelegtes Feld und (c) mit angelegtem Feld.

Wenn ein elektrisches Feld entlang der Probe angelegt wird, hat die Lösung die gleiche Form, wobei aber in Gl. (1.167)  $x$  durch den Ausdruck  $(x - \mu_p \mathcal{E} t)$  ersetzt wird. Dadurch bewegt sich das gesamte Paket der überschüssigen Ladungsträger mit einer Driftgeschwindigkeit  $\mu_p \mathcal{E}$  zum negativen Ende der Probe (Abb. 1.46c). Gleichzeitig diffundieren die Ladungsträger nach außen und rekombinieren wie im feldfreien Fall.

Das obige Beispiel ähnelt dem berühmten Haynes-Shockley-Experiment zur Messung der Ladungsträgerdriftbeweglichkeit in Halbleitern [85]. Bei bekannter Probenlänge, angelegtem Feld und bei bekannter Zeitverzögerung zwischen den angelegten Signalen (Spannung ein und Licht aus) und dem detektierten Signal am Probenende (beides auf dem Oszilloskop sichtbar) kann mit der Formel  $\mu = x / \mathcal{E} t$  die Driftbeweglichkeit berechnet werden.

**Oberflächenrekombination** Ist Oberflächenrekombination an einem Ende der Halbleiterprobe der dominante Prozess (Abb. 1.47), so wird die Randbedingung bei  $x = 0$  durch die Gleichung

$$qD_p \left. \frac{dp_n}{dx} \right|_{x=0} = qS_p [p_n(0) - p_{n0}] \quad (1.168)$$

bestimmt, die besagt, dass die Minoritätsladungsträger, welche die Oberfläche erreichen, dort sofort rekombinieren. Die Konstante  $S_p$  mit der Einheit cm/s ist definiert als die Oberflächenrekombinationsgeschwindigkeit für Löcher. Die Randbedingung bei  $x = \infty$  ist durch Gl. (1.159) gegeben. Ohne Vorspannung und im stationären Zu-

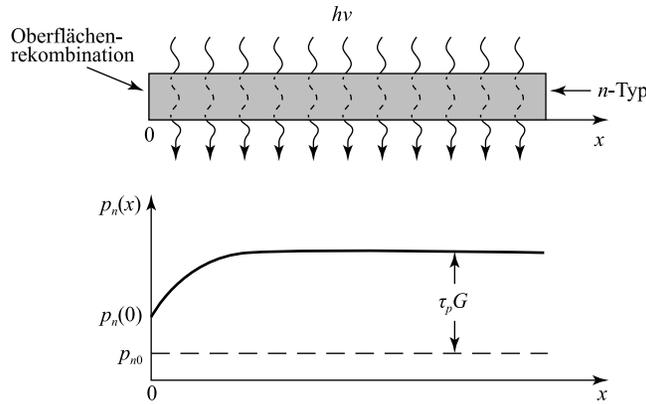


Abb. 1.47 Oberflächenrekombination bei  $x = 0$ . Die räumliche Verteilung der Minoritätsträger nahe der Oberfläche wird durch die Oberflächenrekombinationsgeschwindigkeit beeinflusst.

stand lautet die Differenzialgleichung,

$$0 = G_p - \frac{p_n - p_{n0}}{\tau_p} + D_p \frac{d^2 p_n}{dx^2}. \quad (1.169)$$

Löst man Gl. (1.169) unter den obigen Randbedingungen, bekommt man für  $p_n(x)$  den Ausdruck

$$p_n(x) = p_{n0} + \tau_p G_p \left[ 1 - \frac{\tau_p S_p \exp(-x/L_p)}{L_p + \tau_p S_p} \right], \quad (1.170)$$

der in Abb. 1.47 für endliche Werte von  $S_p$  aufgetragen ist. Wenn  $S_p \rightarrow 0$  geht, dann geht gemäß Gl. (1.162)  $p_n(x) \rightarrow p_{n0} + \tau_p G_p$ . Geht  $S_p \rightarrow \infty$ , dann läuft  $p_n(x) \rightarrow p_{n0} + \tau_p G_p [1 - \exp(-x/L_p)]$ , und die Minoritätsträgerdichte an der Oberfläche nähert sich ihrem thermischen Gleichgewichtswert  $p_{n0}$ . Analog zum Rekombinationsprozess im Volumen, bei dem die Rekombinationsrate bei schwacher Injektion zu  $(1/\tau) = \sigma_p v_{th} N_t$  bestimmt wurde (Gl. (1.90a)), erhält man für die Oberflächenrekombinationsgeschwindigkeit die Formel  $S_p = \sigma_p v_{th} N'_{st}$ .  $N'_{st}$  ist hier die Anzahl der Oberflächenrekombinationszentren pro Flächeneinheit in der Grenzregion.

## Literatur

- 1 J.L. Moll, *Physics of Semiconductors*, McGraw-Hill, New York, 1964.
- 2 R.A. Smith, *Semiconductors*, 2. Aufl., Cambridge University Press, London, 1979.
- 3 T.S. Moss (Hrsg.), *Handbook on Semiconductors*, Bde. 1–4, North-Holland, Amsterdam, 1980.
- 4 K.W. Böer, *Survey of Semiconductor Physics*, Van Nostrand Reinhold, New York, 1990.
- 5 K. Seeger, *Semiconductor Physics*, 7. Aufl., Springer, Berlin, 1999.
- 6 J. Singh, *Physics of Semiconductors and Their Heterostructures*, Ting Lung Book Co., Taichung, 2007.
- 7 Y. Taur, *Fundamentals of Modern VLSI Devices*, 2. Aufl., Cambridge University Press, New York, 2009.
- 8 S.M. Sze und M.K. Lee, *Semiconductor Devices Physics and Technology*, 3. Aufl., Wiley, New York, 2012.
- 9 C. Hamaguchi, *Basic Semiconductor Physics*, 3. Aufl., Springer, New York, 2017.
- 10 R.K. Willardson und A.C. Beer (Hrsg.), *Semiconductors and Semimetals*, Bd. 2, *Physics of III-V Compounds*, Academic, New York, 1966.
- 11 W.B. Pearson, *Handbook of Lattice Spacings and Structure of Metals and Alloys*, Pergamon, New York, 1967.
- 12 H.C. Casey und M.B. Panish, *Heterostructure Lasers*, Academic, New York, 1978.

- 13 C. Kittel, *Introduction to Solid State Physics*, 8. Aufl., Wiley, New York, 2005.
- 14 L. Brillouin, *Wave Propagation in Periodic Structures*, 2. Aufl., Dover, New York, 1963.
- 15 J.M. Ziman, *Principles of the Theory of Solids*, Cambridge University Press, London, 1964.
- 16 M.L. Cohen, „Pseudopotential Calculations for II-VI Compounds“, in (Hrsg. D.G. Thomas), *II-VI Semiconducting Compounds*, W.A. Benjamin, New York, 1967, p. 462.
- 17 C. Kittel, *Quantum Theory of Solids*, Wiley, New York, 1963.
- 18 L.C. Allen, „Interpolation Scheme for Energy Bands in Solids“, *Phys. Rev.*, **98**, 993 (1955).
- 19 F. Herman, „The Electronic Energy Band Structure of Silicon and Germanium“, *Proc. IRE*, **43**, 1703 (1955).
- 20 J.C. Phillips, „Energy-Band Interpolation Scheme Based on a Pseudopotential“, *Phys. Rev.*, **112**, 685 (1958).
- 21 M.L. Cohen und J.R. Chelikowsky, *Electronic Structure and Optical Properties of Semiconductors*, 2. Aufl., Springer, Berlin, 1988.
- 22 J.M. Ziman, *Electrons and Phonons*, Clarendon, Oxford, 1960.
- 23 C.D. Thurmond, „The Standard Thermodynamic Function of the Formation of Electrons and Holes in Ge, Si, GaAs and GaP“, *J. Electrochem. Soc.*, **122**, 1133 (1975).
- 24 V. Alex, S. Finkbeiner und J. Weber, „Temperature Dependence of the Indirect Energy Gap in Crystalline Silicon“, *J. Appl. Phys.*, **79**, 6943 (1996).
- 25 W. Paul und D.M. Warschauer (Hrsg.), *Solids under Pressure*, McGraw-Hill, New York, 1963.
- 26 R.S. Ohl, „Light-Sensitive Electric Device“, US Patent 2,402,662. Filed May 27, 1941. Granted June 25, 1946.
- 27 M. Riordan und L. Hoddeson, „The Origins of the *pn* Junction“, *IEEE Spectrum*, **34-6**, 46 (1997).
- 28 J.S. Blackmore, „Carrier Concentrations and Fermi Levels in Semiconductors“, *Electron. Commun.*, **29**, 131 (1952).
- 29 W.B. Joyce und R.W. Dixon, „Analytic Approximations for the Fermi Energy of an Ideal Fermi Gas“, *Appl. Phys. Lett.*, **31**, 354 (1977).
- 30 O. Madelung (Hrsg.), *Semiconductors – Basic Data*, 2. Aufl., Springer, Berlin, 1996.
- 31 R.N. Hall und J.H. Racette, „Diffusion and Solubility of Copper in Extrinsic and Intrinsic Germanium, Silicon, and Gallium Arsenide“, *J. Appl. Phys.*, **35**, 379 (1964).
- 32 A.G. Milnes, *Deep Impurities in Semiconductors*, Wiley, New York, 1973.
- 33 J. Hermanson und J.C. Phillips, „Pseudopotential Theory of Exciton and Impurity States“, *Phys. Rev.*, **150**, 652 (1966).
- 34 J. Callaway und A.J. Hughes, „Localized Defects in Semiconductors“, *Phys. Rev.*, **156**, 860 (1967).
- 35 E.M. Conwell, „Properties of Silicon and Germanium, Part II“, *Proc. IRE*, **46**, 1281 (1958).
- 36 S.M. Sze und J.C. Irvin, „Resistivity, Mobility, and Impurity Levels in GaAs, Ge, and Si at 300 K“, *Solid-State Electron.*, **11**, 599 (1968).
- 37 W. Shockley, *Electrons and Holes in Semiconductors*, D. Van Nostrand, Princeton, New Jersey, 1950.
- 38 A.S. Grove, *Physics and Technology of Semiconductor Devices*, Wiley, New York, 1967.
- 39 J. Bardeen und W. Shockley, „Deformation Potentials and Mobilities in Nonpolar Crystals“, *Phys. Rev.*, **80**, 72 (1950).
- 40 E. Conwell und V.F. Weisskopf, „Theory of Impurity Scattering in Semiconductors“, *Phys. Rev.*, **77**, 388 (1950).
- 41 C. Bulucea, „Recalculation of Irvin’s Resistivity Curves for Diffused Layers in Silicon Using Updated Bulk Resistivity Data“, *Solid-State Electron.*, **36**, 489 (1993).
- 42 C. Jacoboni, C. Canali, G. Ottaviani und A.A. Quaranta, „A Review of Some Charge Transport Properties of Silicon“, *Solid-State Electron.*, **20**, 77 (1977).

- 43 O. Oda, *Compound Semiconductor Bulk Materials and Characterizations*, World Scientific, Singapore, 2007.
- 44 O. Madelung, *Semiconductors: Data Handbook*, 3. Aufl., Springer, Berlin, 2004.
- 45 W.E. Beadle, J.C.C. Tsai und R.D. Plummer (Hrsg.), *Quick Reference Manual for Silicon Integrated Circuit Technology*, Wiley, New York, 1985.
- 46 F.M. Smits, „Measurement of Sheet Resistivities with the Four-Point Probe“, *Bell Syst. Tech. J.*, **37**, 711 (1958).
- 47 E.H. Hall, „On a New Action of the Magnet on Electric Currents“, *Am. J. Math.*, **2**, 287 (1879).
- 48 L.J. Van der Pauw, „A Method of Measuring Specific Resistivity and Hall Effect of Disc or Arbitrary Shape“, *Philips Res. Rep.*, **13**, 1 (1958).
- 49 D.M. Caughey und R.E. Thomas, „Carrier Mobilities in Silicon Empirically Related to Doping and Field“, *Proc. IEEE*, **55**, 2192 (1967).
- 50 D.E. Aspnes, „GaAs Lower Conduction-Band Minima: Ordering and Properties“, *Phys. Rev. B*, **14**, 5331 (1976).
- 51 P. Smith, M. Inoue und J. Frey, „Electron Velocity in Si and GaAs at Very High Electric Fields“, *Appl. Phys. Lett.*, **37**, 797 (1980).
- 52 J.G. Ruch und G.S. Kino, „Measurement of the Velocity-Field Characteristics of Gallium Arsenide“, *Appl. Phys. Lett.*, **10**, 40 (1967).
- 53 K. Brennan und K. Hess, „Theory of High-Field Transport of Holes in GaAs and InP“, *Phys. Rev. B*, **29**, 5581 (1984).
- 54 B. Kramer und A. Mircea, „Determination of Saturated Electron Velocity in GaAs“, *Appl. Phys. Lett.*, **26**, 623 (1975).
- 55 K.K. Thornber, „Relation of Drift Velocity to Low-Field Mobility and High Field Saturation Velocity“, *J. Appl. Phys.*, **51**, 2127 (1980).
- 56 J.G. Ruch, „Electron Dynamics in Short Channel Field-Effect Transistors“, *IEEE Trans. Electron Devices*, **ED-19**, 652 (1972).
- 57 K.K. Thornber, „Applications of Scaling to Problems in High-Field Electronic Transport“, *J. Appl. Phys.*, **52**, 279 (1981).
- 58 R.A. Logan und S.M. Sze, „Avalanche Multiplication in Ge and GaAs *p-n* Junctions“, *Proc. Int. Conf. Phys. Semicond.*, Kyoto, and *J. Phys. Soc. Jpn. Suppl.*, **21**, 434 (1966).
- 59 W.N. Grant, „Electron and Hole Ionization Rates in Epitaxial Silicon at High Electric Fields“, *Solid-State Electron.*, **16**, 1189 (1973).
- 60 G.H. Glover, „Charge Multiplication in Au-SiC (6H) Schottky Junction“, *J. Appl. Phys.*, **46**, 4842 (1975).
- 61 T.P. Pearsall, F. Capasso, R.E. Nahory, M.A. Pollack und J.R. Chelikowsky, „The Band Structure Dependence of Impact Ionization by Hot Carriers in Semiconductors GaAs“, *Solid-State Electron.*, **21**, 297 (1978).
- 62 I. Umebu, A.N.M. Choudhury und P.N. Robson, „Ionization Coefficients Measured in Abrupt InP Junction“, *Appl. Phys. Lett.*, **36**, 302 (1980).
- 63 R.A. Logan und H.G. White, „Charge Multiplication in GaP *p-n* Junctions“, *J. Appl. Phys.*, **36**, 3945 (1965).
- 64 T.P. Pearsall, „Impact Ionization Rates for Electrons and Holes in  $\text{Ga}_{0.47}\text{In}_{0.53}\text{As}$ “, *Appl. Phys. Lett.*, **36**, 218 (1980).
- 65 T.P. Pearsall, R.E. Nahory und M.A. Pollack, „Impact Ionization Rates for Electrons and Holes in  $\text{GaAs}_{1-x}\text{Sb}_x$  Alloys“, *Appl. Phys. Lett.*, **28**, 403 (1976).
- 66 L.W. Cook, G.E. Bulman und G.E. Stillman, „Electron and Hole Impact Ionization Coefficients in InP Determined by Photomultiplication Measurements“, *Appl. Phys. Lett.*, **40**, 589 (1982).
- 67 I.H. Oguzman, E. Bellotti, K.F. Brennan, J. Kolnik, R. Wang und P.P. Ruden, „Theory of Hole Initiated Impact Ionization in Bulk Zinblend and Wurtzite GaN“, *J. Appl. Phys.*, **81**, 7827 (1997).
- 68 M.R. Brozel und G.E. Stillman (Hrsg.), *Properties of Gallium Arsenide*, 3. Aufl., INSPEC, London, 1996.

- 69 C.R. Crowell und S.M. Sze, „Temperature Dependence of Avalanche Multiplication in Semiconductors“, *Appl. Phys. Lett.*, **9**, 242 (1966).
- 70 C.T. Sah, R.N. Noyce und W. Shockley, „Carrier Generation and Recombination in  $p$ - $n$  Junction and  $p$ - $n$  Junction Characteristics“, *Proc. IRE*, **45**, 1228 (1957).
- 71 R.N. Hall, „Electron–Hole Recombination in Germanium“, *Phys. Rev.*, **87**, 387 (1952).
- 72 W. Shockley und W.T. Read, „Statistics of the Recombination of Holes and Electrons“, *Phys. Rev.*, **87**, 835 (1952).
- 73 W.M. Bullis, „Properties of Gold in Silicon“, *Solid-State Electron.*, **9**, 143 (1966).
- 74 D.T. Stevenson und R.J. Keyes, „Measurement of Carrier Lifetime in Germanium and Silicon“, *J. Appl. Phys.*, **26**, 190 (1955).
- 75 W.W. Gartner, „Spectral Distribution of the Photomagnetic Electric Effect“, *Phys. Rev.*, **105**, 823 (1957).
- 76 S. Wei und M.Y. Chou, „Phonon Dispersions of Silicon and Germanium from First-Principles Calculations“, *Phys. Rev. B*, **50**, 2221 (1994).
- 77 C. Patel, T.J. Parker, H. Jamshidi und W.F. Sherman, „Phonon Frequencies in GaAs“, *Phys. Stat. Sol. (b)*, **122**, 461 (1984).
- 78 W.C. Dash und R. Newman, „Intrinsic Optical Absorption in Single-Crystal Germanium and Silicon at 77 °K and 300 °K“, *Phys. Rev.*, **99**, 1151 (1955).
- 79 H.R. Philipp und E.A. Taft, „Optical Constants of Silicon in the Region 1 to 10 eV“, *Phys. Rev. Lett.*, **8**, 13 (1962).
- 80 D.E. Hill, „Infrared Transmission and Fluorescence of Doped Gallium Arsenide“, *Phys. Rev.*, **133**, A866 (1964).
- 81 H.C. Casey, D.D. Sell und K.W. Wecht, „Concentration Dependence of the Absorption Coefficient for  $n$ - and  $p$ -type GaAs between 1.3 and 1.6 eV“, *J. Appl. Phys.*, **46**, 250 (1975).
- 82 C.Y. Ho, R.W. Powell und P.E. Liley, *Thermal Conductivity of the Elements – A Comprehensive Review*, Am. Chem. Soc. and Am. Inst. Phys., New York, 1975.
- 83 M.G. Holland, „Phonon Scattering in Semiconductors from Thermal Conductivity Studies“, *Phys. Rev.*, **134**, A471 (1964).
- 84 B.H. Armstrong, „Thermal Conductivity in SiO<sub>2</sub>“, in (Hrsg. S.T. Pantelides), *The Physics of SiO<sub>2</sub> and Its Interfaces*, Pergamon, New York, 1978.
- 85 G.A. Slack, „Thermal Conductivity of Pure and Impure Silicon, Silicon Carbide, and Diamond“, *J. Appl. Phys.*, **35**, 3460 (1964).
- 86 E.K. Sichel und J.I. Pankove, „Thermal Conductivity of GaN, 25–360 K“, *J. Phys. Chem. Solids*, **38**, 330 (1977).
- 87 G.H. Dohler, „Doping Superlattices – Historical Overview“, in (Hrsg. P. Bhattacharya), *III-V Quantum Wells and Superlattices*, INSPEC, London, 1996.
- 88 C.H. Huang, X.Y. Wang, M. Igarashi, A. Murayama, Y. Okada, I. Yamashita und S. Samukawa, „Optical Absorption Characteristic of Highly Ordered and Dense Two-Dimensional Array of Silicon Nanodiscs“, *Nanotechnology*, **22**, 105301 (2011).
- 89 M.F. Budiman, W. Hu, M. Igarashi, R. Tsukamoto, T. Isoda, K. M. Itoh, I. Yamashita, A. Murayama, Y. Okada und S. Samukawa, „Control of Optical Bandgap Energy and Optical Absorption Coefficient by Geometric Parameters in Sub-10 nm Silicon-Nanodisc Array Structure“, *Nanotechnology*, **23**, 065302 (2012)
- 90 Y.C. Tsai, M.Y. Lee, Y. Li und S. Samukawa, „Design and Simulation of Intermediate Band Solar Cell with Ultradense Type-II Multilayer Ge/Si Quantum Dot Superlattice“, *IEEE Trans on Electron Devices*, **64**, 4547 (2017).
- 91 Y. Cao, J. Xu, Z. Ge, Y. Zhai, W. Li, X. Jiang und K. Chen, „Enhanced Broadband Spectral Response and Energy Conversion Efficiency for Hetero-Junction Solar Cells with Graded- Sized Si Quantum Dots/SiC Multilayers“, *J. Mater. Chem. C*, **3**, 12061 (2015).
- 92 X. Liu, Y. Zhang, T. Yu, X. Qiao, R. Gresback, X. Pi und D. Yang, „Optimum Quantum Yield of the Light Emission from 2 to 10 nm Hydrosilylated Silicon Quantum Dots“, *Part. Part. Syst. Charact.*, **33**, 44 (2015).

- 93 Y. Li, O. Voskoboynikov, C.P. Lee, S.M. Sze und O. Tretyak, „Electron Energy State Dependence on the Shape and Size of Semiconductor Quantum Dots“, *J. Appl. Phys.*, **90**, 6416 (2001).
- 94 D. Alima, Y. Estrin, D.H. Rich und I. Bar, „The Structural and Optical Properties of Super-Continuum Emitting Si Nanocrystals Prepared by Laser Ablation in Water“, *J. Appl. Phys.*, **112**, 114312 (2012).
- 95 D. Vasileska, S.M. Goodnick und G. Klimeck, *Computational Electronics: Semiclassical and Quantum Device Modeling and Simulation*, CRC Press, New York, 2010.
- 96 T. Grasser, T.W. Tang, H. Kosina und S. Selberherr, „A Review of Hydrodynamic and Energy-Transport Models for Semiconductor Device Simulation“, *Proc. IEEE*, **91**, 251 (2003).
- 97 B. Stritzker, A. Pospieszczyk und J.A. Tagle, „Measurement of Lattice Temperature of Silicon during Pulsed Laser Annealing“, *Phys. Rev. Lett.*, **47**, 356 (1981).
- 98 J.M. Poate, G. Foti und D.C. Jacobson, *Surface Modification and Alloying: by Laser, Ion, and Electron Beams*, Plenum Press, New York, 1983.
- 99 J.R. Haynes und W. Shockley, „The Mobility and Life of Injected Holes and Electrons in Germanium“, *Phys. Rev.*, **81**, 835 (1951).

## Übungsbeispiele

---

### 1.1

- (a) Finden Sie den maximalen Anteil eines konventionellen Einheitszellenvolumens, das durch identische harte Kugeln in einem Diamantgitter gefüllt werden kann.
- (b) Ermitteln Sie die Anzahl der Atome pro Quadratcentimeter in Silizium in der (111)-Ebene bei 300 K.

1.2 Berechnen Sie den tetraedrischen Bindungswinkel, d. h. den Winkel zwischen einem beliebigen Paar der vier Bindungen. (Hinweis: Stellen Sie die vier Bindungen als gleich lange Vektoren dar. Wie groß ist die Summe der vier Vektoren? Nehmen Sie Komponenten dieser Vektorgleichung entlang der Richtung eines dieser Vektoren.)

1.3 Für eine kubisch-flächenzentrierte Zelle ist das Volumen einer konventionellen Einheitszelle  $a^3$ . Finden Sie das Volumen einer primitiven fcc-Einheitszelle mit drei Basisvektoren:  $(0, 0, 0 \rightarrow a/2, 0, a/2)$ ,  $(0, 0, 0 \rightarrow a/2, a/2, 0)$  und  $(0, 0, 0 \rightarrow 0, a/2, a/2)$ .

### 1.4

- (a) Berechnen Sie einen Ausdruck für die Bindungslänge  $d$  im Diamantgitter als Funktion der Gitterkonstante  $a$ .
- (b) In einem Siliziumkristall hat eine Ebene Schnittpunkte bei 10.86, 16.29 und 21.72 Å mit den drei kartesischen Koordinatenachsen. Finden Sie die Miller-Indizes der Ebene.

1.5 Zeigen Sie, dass

- (a) jeder Vektor des reziproken Gitters normal zu einer Menge von Ebenen im direkten Gitter ist und
- (b) das Volumen einer Einheitszelle des reziproken Gitters umgekehrt proportional zum Volumen einer Einheitszelle des direkten Gitters ist.

1.6 Zeigen Sie, dass das reziproke Gitter eines kubisch-raumzentrierten (bcc-) Gitters mit der Gitterkonstante  $a$  ein kubisch-flächenzentriertes (fcc-) Gitter ist, wobei die Seite der kubischen Zelle  $4\pi/a$  sein muss.

Tipp: Verwenden Sie einen symmetrischen Satz von Basisvektoren für das bcc-Gitter:

$$\mathbf{a} = \frac{a}{2}(\mathbf{y} + \mathbf{z} - \mathbf{x}), \quad \mathbf{b} = \frac{a}{2}(\mathbf{z} + \mathbf{x} - \mathbf{y}), \quad \mathbf{c} = \frac{a}{2}(\mathbf{x} + \mathbf{y} - \mathbf{z}),$$

wobei  $a$  die Gitterkonstante einer konventionellen primitiven Zelle ist und  $\mathbf{x}$ ,  $\mathbf{y}$ ,  $\mathbf{z}$  die Einheitsvektoren in kartesischen Koordinaten sind. Nehmen Sie für das fcc-Gitter

$$\mathbf{a} = \frac{a}{2}(\mathbf{y} + \mathbf{z}), \quad \mathbf{b} = \frac{a}{2}(\mathbf{z} + \mathbf{x}), \quad \mathbf{c} = \frac{a}{2}(\mathbf{x} + \mathbf{y}).$$

**1.7** In der Nähe der Leitungsbandminima kann die Energie ausgedrückt werden als

$$E = \frac{\hbar^2}{2} \left( \frac{k_x^2}{m_x^*} + \frac{k_y^2}{m_y^*} + \frac{k_z^2}{m_z^*} \right).$$

In Si gibt es sechs zigarrenförmige Minima entlang der [100]-Richtung. Ermitteln Sie das Verhältnis der effektiven Masse in Längsrichtung  $m_l^*$  zur effektiven Masse in Querrichtung  $m_t^*$ , wenn das Verhältnis der Achsen des Ellipsoids mit konstanter Energie 5:1 beträgt.

**1.8** Im Leitungsband eines Halbleiters gebe es ein unteres Tal in der Mitte der Brillouin-Zone und sechs höhere Täler an der Zonengrenze entlang der [100]-Richtung. Ermitteln Sie das Verhältnis der effektiven Zustandsdichte der oberen Täler zu der des unteren Tals, wenn die effektive Masse für das untere Tal  $0.1m_0$  und die für die oberen Täler  $1.0m_0$  beträgt.

**1.9** Leiten Sie die Dichte der Zustände im Leitungsband gemäß Gl. (1.14) ab. (Hinweis: Die Wellenlänge  $\lambda$  einer stehenden Welle hängt mit der Länge des Halbleiters  $L$  durch die Beziehung  $L/\lambda = n_x$  zusammen, wobei  $n_x$  eine ganze Zahl ist. Die Wellenlänge kann durch die de Broglie-Hypothese  $\lambda = h/p_x$  ausgedrückt werden. Betrachten Sie einen dreidimensionalen Würfel der Seitenlänge  $L$ .)

**1.10** Berechnen Sie die durchschnittliche kinetische Energie von Elektronen im Leitungsband eines  $n$ -Typ-Halbleiters ohne Entartung. Die Zustandsdichte ist durch Gl. (1.12) gegeben.

**1.11** Finden Sie die Dichte ionisierter Donatoren bei  $T = 77$  K, wenn eine Siliziumprobe mit  $10^{16}$  Phosphorstörstellen/cm<sup>3</sup> dotiert ist. Nehmen Sie an, dass die Ionisierungsenergie für Donatorstörstellen (Phosphor) und die effektive Masse der Elektronen unabhängig von der Temperatur ist. (Hinweis: Die Lösung muss iterativ gefunden werden. Wählen Sie dazu zuerst als Startwert einen beliebigen  $N_D^+$ -Wert zur Berechnung des Fermi-Niveaus und suchen Sie dann das entsprechende  $N_D^+$  mithilfe der angegebenen Formel. Wenn die Werte nicht übereinstimmen, wählen Sie einen anderen  $N_D^+$ -Wert und wiederholen Sie den Vorgang solange, bis ein konsistenter  $N_D^+$ -Wert erreicht ist.)

**1.12** Benutzen Sie eine grafische Methode zur Bestimmung des Fermi-Niveaus für eine bordotierte Siliziumprobe mit einer Störstellenkonzentration von  $10^{15}$  cm<sup>-3</sup> bei 300 K (Anmerkung:  $n_i = 9.65 \times 10^9$  cm<sup>-3</sup>).

**1.13** Finden Sie bei 300 K die Position des Fermi-Niveaus in Bezug auf die Leitungsbandunterkante ( $E_C - E_F$ ) für eine Siliziumprobe für eine Dotierung von  $2 \times 10^{10}$  cm<sup>-3</sup>. Nehmen Sie an, dass alle Donatoren vollständig ionisiert sind.

**1.14** Gold in Si hat zwei Energieniveaus in der Bandlücke:  $E_C - E_A = 0.54 \text{ eV}$  und  $E_D - E_V = 0.29 \text{ eV}$ . Angenommen, das dritte Niveau  $E_D - E_V = 0.35 \text{ eV}$  ist inaktiv.

- Wie wird der Ladungszustand der Goldniveaus in Si, das mit einer hohen Konzentration von Boratomen dotiert ist, sein? Warum?
- Welche Auswirkungen hat Gold auf die Elektronen- und Löcherkonzentrationen?

**1.15** Bewerten und bestimmen Sie aus Abb. 1.13, welche Art von Störstellen zur Dotierung der Si-Probe verwendet wird.

**1.16** Für eine  $n$ -Typ-Siliziumprobe, die mit  $2.86 \times 10^{16} \text{ cm}^{-3}$  Phosphoratomen dotiert ist, finden Sie das Verhältnis der neutralen zu den ionisierten Donatoren bei 300 K. Hinweis: Die Ionisierungsenergie des Donators beträgt  $E_C - E_D = 0.045 \text{ eV}$ .

**1.17**

- Angenommen, das Beweglichkeitsverhältnis  $\mu_n/\mu_p \equiv b$  in Si ist eine von der Störstellenkonzentration unabhängige Konstante. Finden Sie den maximalen spezifischen Widerstand  $\rho_m$  in Bezug auf den intrinsischen Widerstand  $\rho_i$  bei 300 K. Berechnen Sie  $\rho_i$  und  $\rho_m$ , wenn  $b = 3$  und die Löcherbeweglichkeit von intrinsischem Si  $450 \text{ cm}^2/\text{V}\cdot\text{s}$  beträgt,
- Ermitteln Sie die Elektronen- und Löcherkonzentration, Beweglichkeit und den spezifischen Widerstand einer GaAs-Probe bei 300 K mit  $5 \times 10^{15}$  Zinkatomen/ $\text{cm}^3$ ,  $10^{17}$  Schwefelatomen/ $\text{cm}^3$  und  $10^{17}$  Kohlenstoffatomen/ $\text{cm}^3$ .

**1.18** Betrachten Sie ein kompensiertes  $n$ -Typ-Silizium bei  $T = 300 \text{ K}$ , mit einer Leitfähigkeit von  $\sigma = 16 \text{ S/cm}$  und einer Akzeptorkonzentration von  $10^{17} \text{ cm}^{-3}$ . Bestimmen Sie die Donatorkonzentration und die Elektronenbeweglichkeit. (Anmerkung: Ein kompensierter Halbleiter ist ein Halbleiter, der sowohl Donator- als auch Akzeptoratome in der gleichen Region enthält.)

**1.19** Ermitteln Sie bei 300 K den spezifischen Widerstand für eine Siliziumprobe, die mit  $1.0 \times 10^{14} \text{ cm}^{-3}$  Phosphoratomen,  $8.5 \times 10^{12} \text{ cm}^{-3}$  Arsenatomen und  $1.2 \times 10^{13} \text{ cm}^{-3}$  Boratomen dotiert ist. Nehmen Sie an, die Störstellen seien vollständig ionisiert und die Beweglichkeiten betragen, unabhängig von den Verunreinigungskonzentrationen,  $\mu_n = 1500 \text{ cm}^2/\text{V}\cdot\text{s}$ ,  $\mu_p = 500 \text{ cm}^2/\text{V}\cdot\text{s}$ .

**1.20** Ein Halbleiter hat einen spezifischen Widerstand von  $1.0 \Omega\cdot\text{cm}$  und einen Hall-Koeffizienten von  $-1250 \text{ cm}^2/\text{C}$ . Berechnen Sie die Ladungsträgerdichte und Ladungsträgerbeweglichkeit unter der Annahme, dass nur ein Trägertyp vorhanden ist und die mittlere Lebensdauer proportional zur Energie der Ladungsträger ist, d. h.,  $\tau \propto E$ .

**1.21** Die Rekombinationsrate ist durch Gl. (1.87) gegeben. Unter schwachen Injektionsbedingungen kann  $U$  als  $(p_n - p_{n0})/\tau_r$  ausgedrückt werden, wobei  $\tau_r$  die Rekombinationslebensdauer ist. Finden Sie die Werte von  $(E_t - E_i)$ , bei denen die Rekombinationslebensdauer  $\tau_r$  zu  $2\tau_{r0}$  wird, wenn  $\sigma_n = \sigma_p = \sigma_o$ ,  $n_{n0} = 10^{15} \text{ cm}^{-3}$  und  $\tau_{r0} \equiv (v_{th}\sigma_o N_t)^{-1}$ .

**1.22** Ermitteln Sie für die einstufige Rekombination mit identischem Einfangquerschnitt für Elektronen und Löcher die Anzahl der Störstellen pro Volumeneinheit und die Generationsrate unter der Bedingung einer vollständigen Verarmung der La-

dungsträger. Nehmen Sie an, die Störstellen befinden sich in der Mitte der Bandlücke,  $\sigma = 2 \times 10^{-16} \text{ cm}^2$  und  $v_{th} = 10^7 \text{ cm/s}$ .

**1.23** In einem Halbleiterbereich, der vollständig an Ladungsträgern verarmt ist (d. h.,  $n \ll n_i$ ,  $p \ll n_i$ ), werden Elektron-Loch-Paare durch abwechselnde Emission von Elektronen und Löchern durch Generationszentren erzeugt. Leiten Sie die durchschnittliche Zeit ab, in der solche Emissionsvorgänge stattfinden (Annahme:  $\sigma_n = \sigma_p = \sigma$ ). Finden Sie auch die durchschnittliche Zeit für  $\sigma = 2 \times 10^{-16} \text{ cm}^2$ ,  $v_{th} = 10^7 \text{ cm/s}$  und  $E_t = E_i$  ( $T = 300 \text{ K}$ ).

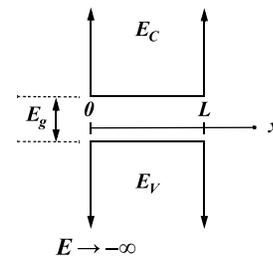
**1.24** Finden Sie für einen einstufigen Rekombinationsprozess die durchschnittliche Zeit, die zwischen den einzelnen Rekombinationsprozessen in einem Bereich einer Siliziumprobe stattfindet, in dem  $n = p = 10^{13} \text{ cm}^{-3}$ ,  $\sigma_n = \sigma_p = 2 \times 10^{-16} \text{ cm}^2$ ,  $v_{th} = 10^7 \text{ cm/s}$ ,  $N_t = 10^{16} \text{ cm}^{-3}$  und  $(E_t - E_i) = 5kT$ .

**1.25** Finden Sie die optische Phononenenergie für einen Si-Kristall mit  $m_1 = m_2$  und  $\sqrt{\alpha_f/m_1} = 7.63 \times 10^{12} \text{ Hz}$  an der Grenze der Brillouin-Zone. Die Kraftkonstante ist  $\alpha_f$ .

**1.26** Angenommen,  $\text{Ga}_{0.5}\text{In}_{0.5}\text{As}$  ist bei  $500^\circ\text{C}$  mit dem InP-Substrat gitterangepasst. Finden Sie die Gitterfehlangepassung zwischen den Schichten, wenn die Probe auf  $27^\circ\text{C}$  gekühlt wird.

**1.27** Finden Sie das Verhältnis der Leitungsbanddiskontinuität des Heteroübergangs  $\text{Al}_{0.4}\text{Ga}_{0.6}\text{As}/\text{GaAs}$  zur  $\text{Al}_{0.4}\text{Ga}_{0.6}\text{As}$ -Bandlücke.

**1.28** In der Abbildung unten sieht man ein eindimensionales Bandprofil eines Quantentopfes in einem Substratmaterial.  $E_g$  ist die Bandlücke des Substrats,  $L$  ist die Länge des Quantentopfes. Die Barrieren im Valenz- und Leitungsband seien unendlich hoch. Die effektiven Massen der Elektronen und Löcher werden mit  $m_e^*$  und  $m_h^*$  bezeichnet. Leiten Sie die Emissionsenergie des Quantentopfes ab und zeigen Sie, dass die Emissionsenergie mit der Größe des Quantentopfes skaliert.



Zu Aufgabe 1.28

**1.29** Bei thermischem Gleichgewicht und für nicht entartete Halbleiter mit einer parabolischen Bandstruktur hängen die mobilen Elektronen- und Löcherkonzentrationen  $n$  und  $p$  mit dem elektrostatischen Potential durch die Boltzmann-Statistik nach Gln. (1.27a) und (1.27b) zusammen. Zeigen Sie, dass sich Gl. (1.150) als folgende nicht lineare Poisson-Gleichung hinschreiben lässt ( $V_{th} = kT/q$ ):

$$\frac{d^2\psi}{dx^2} = \frac{qn_i}{\epsilon_s} \left[ \exp\left(\frac{\psi}{V_{th}}\right) - \exp\left(\frac{-\psi}{V_{th}}\right) + \frac{N_A - N_D}{n_i} \right].$$

**1.30** Das Drift-Diffusions-Modell besteht aus der Poisson-Gleichung, der Elektronenstromkontinuitätsgleichung und der Löcherstromkontinuitätsgleichung für das elektrostatische Potential und der Elektronenkonzentration bzw. der Löcherkonzentration. Schreiben Sie für stationäre Bedingungen die drei gekoppelten partiellen Differenzialgleichungen als Funktion der Variablen  $\psi$ ,  $n$  und  $p$  in zwei und drei Dimensionen an.

**1.31** Zeigen Sie, dass bei Bauelementen mit gleichmäßiger Bandlücke die Wärmequelle in Gl. (1.157) sich in Form des elektrischen Feldes ausdrücken lässt.

**1.32** In einem Haynes-Shockley-Experiment unterscheiden sich die Maximalamplituden der Minoritätsladungsträger bei  $t_1 = 25 \mu\text{s}$  und  $t_2 = 100 \text{ ps}$  um den Faktor zehn. Finden Sie die Lebenszeit der Minoritätsladungsträger.

**1.33** Finden Sie die Halbwertsbreite des Trägerimpulses bei  $t = 1 \text{ s}$  aus dem Ausdruck, der die Drift und Diffusion der Ladungsträger im Haynes-Shockley-Experiment beschreibt. Nehmen Sie an, der Diffusionskoeffizient beträgt  $10 \text{ cm}^2/\text{s}$ .

**1.34** Überschüssige Ladungsträger werden auf einer Oberfläche ( $x = 0$ ) einer dünnen Schicht aus  $n$ -Typ-Silizium ( $N_D = 3 \times 10^{17} \text{ cm}^{-3}$ ) mit der Dicke  $W = 0.05 \text{ mm}$  injiziert und an der gegenüberliegenden Rückseite entnommen, wobei  $p_n(W) = p_{n0}$  ist. Ermitteln Sie den Anteil des injizierten Stroms, der durch Diffusion die Rückseite erreicht, wenn die Lebensdauer der Ladungsträger  $50 \mu\text{s}$  beträgt.

**1.35** Eine  $n$ -Typ-GaAs-Probe mit einer Dotierung von  $N_D = 5 \times 10^{15} \text{ cm}^{-3}$  wird beleuchtet. Das gleichmäßig absorbierte Licht erzeugt  $10^{17}$  Elektron-Loch-Paare/ $\text{cm}^3\text{-s}$ . Die Lebensdauer  $\tau_p$  beträgt  $10^{-7} \text{ s}$ ,  $L_p = 1.93 \times 10^{-3} \text{ cm}$ , die Oberflächenrekombinationsgeschwindigkeit  $S_p$  beträgt  $10^5 \text{ cm/s}$ . Ermitteln Sie die Anzahl der an der Oberfläche rekombinierenden Löcher pro Flächeneinheit und Zeiteinheit.

**1.36** Ein  $n$ -Typ-Halbleiter hat eine überschüssige Löcherkonzentration von  $10^{14} \text{ cm}^{-3}$ , eine Lebensdauer der Minoritätsladungsträger von  $10^{-6} \text{ s}$  im Volumen der Probe und eine Lebensdauer der Minoritätsladungsträger von  $10^{-7} \text{ s}$  an der Oberfläche. Angenommen, das angelegte elektrische Feld ist null und  $D_p = 10 \text{ cm}^2/\text{s}$ . Bestimmen Sie die stationäre überschüssige Ladungsträgerkonzentration als Funktion des Abstands von der Oberfläche ( $x = 0$ ) des Halbleiters.