

MITTSCHRIEB ZUR VORLESUNG: HALBLEITERPHYSIK

Dr. habil. Hetterich und Dr. Schaadt

Vorlesung Sommersemester 2009

Letzte Aktualisierung und Verbesserung: 27. Juli 2009

Mitschrieb der Vorlesung HALBLEITERPHYSIK
von Herrn Dr. habil. HETTERICH und Herrn Dr. SCHAADT im Sommersemester 2009
von MARCO SCHRECK.

Dieser Mitschrieb erhebt keinen Anspruch auf Vollständigkeit und Korrektheit.
Kommentare, Fehler und Vorschläge und konstruktive Kritik bitte an Marco.Schreck@gmx.de.

Inhaltsverzeichnis

| | |
|---|-----------|
| 1 Die Boltzmann-Transportgleichung | 5 |
| 1.1 Zeitliche Inhomogenitäten bei räumlicher Homogenität | 8 |
| 1.1.1 Erzeugungsprozesse | 8 |
| 1.2 Bipolare Rekombination | 11 |
| 1.3 Nichtstrahlende Rekombination | 11 |
| 1.4 Räumliche Inhomogenitäten und Transportphänomene | 13 |
| 1.5 Drift von Ladungsträgern | 13 |
| 1.5.1 Hochfeldeffekte | 15 |
| 1.5.2 Negativer differentieller Widerstand | 16 |
| 1.5.3 Stoßionisation | 16 |
| 2 Die Differentialgleichungen der inneren Elektronik | 17 |
| 2.1 Die dielektrische Relaxationszeit und die Debye-Länge | 18 |
| 2.2 Die Diffusionslänge | 19 |
| 3 Metall-Halbleiter-Übergänge | 21 |
| 3.1 Schottky-Kontakt und ohmscher Kontakt | 21 |
| 3.2 Höhe der Schottky-Barriere | 23 |
| 3.3 Isolator-Halbleiter-Übergang | 24 |
| 4 Halbleiterübergänge im Nichtgleichgewicht/Bauelemente | 27 |
| 4.1 Der pn-Übergang bei angelegter Spannung | 27 |
| 4.1.1 Auswirkungen auf das elektrische Potential | 27 |
| 4.1.2 Mechanismus des Stromtransports | 28 |
| 4.1.3 Verlauf des chemischen Potentials | 29 |
| 4.1.4 Verlauf des Quasi-Fermi-Niveaus | 31 |
| 4.1.5 Diodenkennlinie | 31 |
| 4.1.6 Durchbruchverhalten | 32 |
| 4.1.7 Wechselstromverhalten der Diode | 33 |
| 4.2 Anwendungen der Diode | 34 |
| 4.3 Spezielle Bauelemente zur Mikrowellenerzeugung | 42 |
| 4.4 Der bipolare Transistor | 44 |
| 4.4.1 Funktionsweise | 44 |
| 4.4.2 Verstärkungsverhalten des Transistors | 45 |
| 4.4.3 Das Kennlinienfeld | 46 |
| 4.4.4 Betrieb des Transistors | 47 |
| 4.5 Der Thyristor | 48 |
| 4.6 Unipolare Bauelemente | 48 |
| 4.6.1 Junction Field Effect Transistor (JFET) | 48 |
| 4.6.2 Metal Semiconductor Field Effect Transistor (MESFET) | 49 |
| 4.6.3 Metal Insulator Semiconductor Diode (MIS-Diode) | 50 |
| 4.6.4 Metal Oxid Semiconductor Field Effect Transistor (MOSFET) | 51 |
| 5 Halbleitertechnologie | 53 |
| 5.1 Dotierungen | 54 |

Kapitel 1

Die Boltzmann-Transportgleichung

Im Gleichgewicht hängt die Verteilungsfunktion $f(E)$ nur von der Energie ab und ist durch die Fermi-Dirac-Verteilung gegeben. Die Parameter, welche die Verteilungsfunktion festlegen, sind Temperatur T und Fermienergie E_F . Im Nichtgleichgewichtszustand sind T und E_F makroskopisch nicht definiert (Halbleiter mit Temperaturgradient, angelegte Spannung usw.). **Aber** es kann sich unter bestimmten Bedingungen trotzdem ein lokales thermodynamisches Gleichgewicht einstellen. Dann hängen die Zustandsgrößen vom Ort ab, also $T(\mathbf{r})$, $\eta(\mathbf{r})$. ($\eta(\mathbf{r})$ beinhaltet einerseits das chemische und andererseits das elektrische Potential.) Wir definieren uns nun eine Verteilungsfunktion für die Elektronen: $f(\mathbf{k}, \mathbf{r}, t)$. Diese gibt die Wahrscheinlichkeit an, Elektronen mit Wellenvektor \mathbf{k} am Ort \mathbf{r} zur Zeit t zu finden. Zu beachten ist jedoch, dass es sich um ein halbklassisches Modell handelt; wegen der Heisenbergschen Unbestimmtheitsrelation können nämlich Wellenvektor und Ort nicht gleichzeitig mit beliebiger Genauigkeit gemessen werden. Die zeitliche Änderung von $f(\mathbf{k}, \mathbf{r}, t)$ in einem infinitesimalen Zeitintervall δt . Wir wählen diese Zeitdauer so kurz, dass keine Stöße der Elektronen untereinander, mit Phononen, Störstellen usw. erfolgen. Dann ändern sich Ort \mathbf{r} und Impuls $\mathbf{p} = \hbar\mathbf{k}$ nur durch die freie Drift der Ladungsträger sowie durch beschleunigende äußere Felder. Wie sieht nun der Impuls $\mathbf{k}(t + \delta t)$ aus? Wir machen dazu eine Taylorentwicklung:

$$\mathbf{k}(t + \delta t) = \mathbf{k} + \dot{\mathbf{k}}\delta t + \mathcal{O}(\delta t^2), \quad (1.1a)$$

$$\mathbf{r}(t + \delta t) = \mathbf{r} + \dot{\mathbf{r}}\delta t + \mathcal{O}(\delta t^2). \quad (1.1b)$$

Die Verteilungsfunktion selbst ändert sich nicht, sondern nur deren Argumente:

$$f(\mathbf{k}, \mathbf{r}, t) = f(\mathbf{k} + \dot{\mathbf{k}}\delta t, \mathbf{r} + \dot{\mathbf{r}}\delta t, t + \delta t). \quad (1.2)$$

Wir bearbeiten diese Gleichung weiter, indem wir auf der rechten Seite eine lineare Taylorentwicklung nach δt durchführen:

$$f(\mathbf{k}, \mathbf{r}, t) = f(\mathbf{k}, \mathbf{r}, t) + \left[\nabla_{\mathbf{k}} f(\mathbf{k}, \mathbf{r}, t) \cdot \dot{\mathbf{k}} + \nabla_{\mathbf{r}} f(\mathbf{k}, \mathbf{r}, t) \cdot \dot{\mathbf{r}} + \frac{\partial}{\partial t} f(\mathbf{k}, \mathbf{r}, t) \right] \delta t. \quad (1.3)$$

Damit diese Gleichung erfüllt ist, muss der Ausdruck in der Klammer verschwinden:

$$\boxed{\nabla_{\mathbf{k}} f(\mathbf{k}, \mathbf{r}, t) \cdot \dot{\mathbf{k}} + \nabla_{\mathbf{r}} f(\mathbf{k}, \mathbf{r}, t) \cdot \dot{\mathbf{r}} + \frac{\partial}{\partial t} f(\mathbf{k}, \mathbf{r}, t) = 0,} \quad (1.4)$$

was wir auch in folgender Form schreiben können:

$$\boxed{\frac{df}{dt}(\mathbf{k}(t), \mathbf{r}(t), t) = 0.} \quad (1.5)$$

Beachte: Dass die totale Ableitung verschwindet, heißt **nicht**, dass f zeitlich konstant bleibt. (Das wäre für $\partial f / \partial t = 0$ der Fall.) Es bedeutet, dass sich f so entwickelt, wie man es ohne Stöße erwarten würde. Tatsächlich gibt es Stoßprozesse; das können Elektron-Elektron-Stoßprozesse vermittelt durch die Coulombwechselwirkung oder Stöße mit Phononen und Störstellen (Gitterdefekte, Fremdatome) sein. Diese ändern die Verteilungsfunktion **zusätzlich**.

$$\boxed{\frac{df}{dt}(\mathbf{k}, \mathbf{r}, t) = \frac{\partial f}{\partial t} \Big|_{\text{Stöße}}.} \quad (1.6)$$

Dabei handelt es sich um die Boltzmann-Transportgleichung in linearer Näherung. Aus (1.4) ergibt sich weiter:

$$\boxed{\frac{\partial f}{\partial t} + \dot{\mathbf{k}} \cdot \nabla_{\mathbf{k}} f + \dot{\mathbf{r}} \cdot \nabla_{\mathbf{r}} f = \frac{\partial f}{\partial t} \Big|_{\text{Stöße}}} . \quad (1.7)$$

Eine Änderung von \mathbf{k} durch äußere Felder lautet dann mit $\mathbf{F} = \dot{\mathbf{p}} = \hbar \dot{\mathbf{k}}$:

$$\dot{\mathbf{k}} = -\frac{e}{\hbar}(\mathbf{E} + \dot{\mathbf{r}} \times \mathbf{B}) . \quad (1.8)$$

Der Ortsvektor ändert sich durch Teilchenbewegung. Wir beschreiben das Teilchen durch ein Wellenpaket, das sich mit der Gruppengeschwindigkeit $\dot{\mathbf{r}} = 1/\hbar \nabla_{\mathbf{k}} E(\mathbf{k})$ ausbreitet. Hier taucht nun die Dispersionsstruktur $E(\mathbf{k})$ des Halbleiters auf.

Durch Lösen der Boltzmann-Transportgleichung ist dann $f(\mathbf{k}, \mathbf{r}, t)$ bekannt. Die Mittelwerte für die physikalischen Größen (Teilchendichte, Energiedichte) sind dann leicht berechenbar. Im Nichtgleichgewicht gibt es zusätzlich Ströme. Wir definieren die Stromdichte:

$$\mathbf{j} = -\frac{e}{\hbar} \int \nabla_{\mathbf{k}} E \cdot f(\mathbf{k}, \mathbf{r}, t) D(\mathbf{k}) d^3 k , \quad (1.9)$$

mit der Gruppengeschwindigkeit $1/\hbar \nabla_{\mathbf{k}} E(\mathbf{k})$ und der Zustandsdichte $D(\mathbf{k})$ pro Volumen im realen Raum als Funktion von \mathbf{k} . Nun definieren wir noch eine Energiestromdichte durch:

$$\omega = \frac{1}{\hbar} \int \nabla_{\mathbf{k}} E \cdot E(\mathbf{k}) f(\mathbf{k}, \mathbf{r}, t) D(\mathbf{k}) d^3 k . \quad (1.10)$$

Das Problem ist jedoch, dass $f(\mathbf{k}, \mathbf{r}, t)$ schwer zu berechnen ist, insbesondere der Stoßterm auf der rechten Seite ist problematisch. Ganz allgemein ist der Stoßterm durch folgenden Ausdruck gegeben:

$$\frac{\partial f}{\partial t} \Big|_{\text{Stöße}} = \int \{f(\mathbf{k}') [1 - f(\mathbf{k})] w(\mathbf{k}', \mathbf{k}) - f(\mathbf{k}) [1 - f(\mathbf{k}')] w(\mathbf{k}, \mathbf{k}')\} d^3 k' . \quad (1.11)$$

Hierbei ist $w(\mathbf{k}', \mathbf{k})$ die Streurrate von \mathbf{k}' zu \mathbf{k} , wenn \mathbf{k}' besetzt und \mathbf{k} leer ist. $A \equiv f(\mathbf{k}') [1 - f(\mathbf{k})]$ ist die Wahrscheinlichkeit, dass der Zustand mit \mathbf{k}' besetzt und \mathbf{k} leer ist, $B \equiv f(\mathbf{k}) [1 - f(\mathbf{k}')]$ die Wahrscheinlichkeit, dass der Zustand mit \mathbf{k} besetzt und mit \mathbf{k}' leer ist (Rückstreuung). Weiterhin nehmen wir an, dass Streuprozesse auf mikroskopischer Ebene reversibel verlaufen: $w(\mathbf{k}, \mathbf{k}') = w(\mathbf{k}', \mathbf{k})$. Damit gilt:

$$\frac{\partial f}{\partial t} \Big|_{\text{Stöße}} = \int [f(\mathbf{k}') - f(\mathbf{k})] w(\mathbf{k}, \mathbf{k}') d^3 k' . \quad (1.12)$$

Die Berechnung von $w(\mathbf{k}, \mathbf{k}')$ erfolgt quantenmechanisch. Die Boltzmann-Transportgleichung exakt zu lösen, erweist sich als schwierig. Wir möchten damit mit folgenden Vereinfachungen und Näherungen arbeiten:

- 1.) Wir betrachten einen stationären Zustand (beispielsweise einen konstanten Stromfluss durch den Halbleiter). Somit ist

$$\frac{\partial f}{\partial t}(\mathbf{k}, \mathbf{r}, t) = 0 \Rightarrow \boxed{\frac{\partial f}{\partial t} \Big|_{\text{Stöße}} = \dot{\mathbf{k}} \cdot \nabla_{\mathbf{k}} f + \dot{\mathbf{r}} \cdot \nabla_{\mathbf{r}} f} . \quad (1.13)$$

Dabei handelt es sich um die Boltzmannsche Stationaritätsbedingung, welche die Grundlage der Transporttheorie ist.

- 2.) Relaxationszeitnäherung für den Stoßterm:

- i.) Wir betrachten ein **lokales** thermodynamisches Gleichgewicht. Das System an sich befindet sich im Nichtgleichgewicht; die Verteilungsfunktion f wird aber trotzdem **räumlich und zeitlich lokal** näherungsweise durch eine Gleichgewichtsfunktion $f_0(\mathbf{k}, \mathbf{r}, t)$ beschrieben, die nur noch von der Energie abhängt.

$$f_0(\mathbf{k}, \mathbf{r}, t) = \frac{1}{\exp\left(\frac{E(\mathbf{k}) - \zeta(\mathbf{r}, t)}{k_B T(\mathbf{r}, t)}\right) + 1} . \quad (1.14)$$

Hierbei handelt es sich um die Fermi-Dirac-Verteilung mit dem Unterschied, dass die Temperatur und das chemische Potential **orts- und zeitabhängig**, also nicht global definiert sind. Die gesamte Nichtgleichgewichtsverteilung ist dann

$$f(\mathbf{k}, \mathbf{r}, t) = f_0(\mathbf{k}, \mathbf{r}, t) + \delta f(\mathbf{k}, \mathbf{r}, t) , \quad (1.15)$$

wobei $f_0(\mathbf{k}, \mathbf{r}, t)$ das Nichtgleichgewicht in nullter Näherung beschreibt. Der zusätzliche Term $\delta f(\mathbf{k}, \mathbf{r}, t)$ sei eine kleine Abweichung von der lokalen Gleichgewichtsverteilung f_0 .

- ii.) Zur Erreichung des lokalen Gleichgewichts müssen genügend Stoßprozesse auf einer Längenskala bzw. Zeitskala auftreten, innerhalb der sich T und ζ nicht signifikant ändern. Dies führt zu einer Entkopplung verschiedener Orten \mathbf{r} und Zeiten t . Die mittlere freie Weglänge zwischen Stößen sei l_f und die mittlere freie Flugzeit zwischen Stößen t_f . Kennt man diese beiden Parameter, kann man quantifizieren, wann diese Näherung erfüllt ist, was für räumliche Entkopplung auf die Bedingungen

$$|\nabla T| \cdot l_f \ll T, \quad |\zeta| \cdot l_f \ll \zeta, \quad (1.16)$$

führt. Für zeitliche Entkopplung müssen

$$\left| \frac{\partial T}{\partial t} \right| \cdot t_f \ll T, \quad \left| \frac{\partial \zeta}{\partial t} \right| \cdot t_f \ll \zeta, \quad (1.17)$$

erfüllt sein. Typische Werte sind $l_f(e, h) \sim 10 \text{ nm}$ und $t_f(e, h) \sim 10^{-13} \text{ s}$; also sind für gewöhnliche Halbleiterbauelemente (Mikroelektronik) diese Bedingungen erfüllt. Falls ein lokales Gleichgewicht vorliegt und die Abweichung δf davon sehr klein ist, kann man auf die **Relaxationszeitnäherung** zurückgreifen:

$$\boxed{\left. \frac{\partial f}{\partial t} \right|_{\text{Stöße}} = -\frac{\delta f}{\tau(E)} = -\frac{1}{\tau(E)} [f - f_0].} \quad (1.18)$$

Die zeitliche Änderung der Verteilungsfunktion ist somit proportional zur Abweichung von der Gleichgewichtsverteilung und wird durch diese Abweichung getrieben. Die Lösung der Differentialgleichung (1.18) ist

$$f = f_0 + [f(t=0) - f_0] \exp\left(-\frac{t}{\tau(E)}\right), \quad (1.19)$$

wobei f_0 zeitunabhängig ist. Es handelt sich also um einen exponentiellen Zerfall der Abweichung vom Gleichgewicht.

Wir wollen nun empirische Gleichungen aus der Boltzmann-Transportgleichung in Relaxationszeitnäherung herleiten:

$$\dot{\mathbf{k}} \cdot \nabla_{\mathbf{k}} f + \dot{\mathbf{r}} \cdot \nabla_{\mathbf{r}} f = -\frac{\delta f}{\tau(E)}, \quad (1.20)$$

und zwar das Ohmsche Gesetz als Beispiel. Das äußere Feld sei ein homogenes elektrisches Feld, womit $\dot{\mathbf{k}} = -e/\hbar \mathbf{E}$ und $\nabla_{\mathbf{r}} f = 0$ gilt.

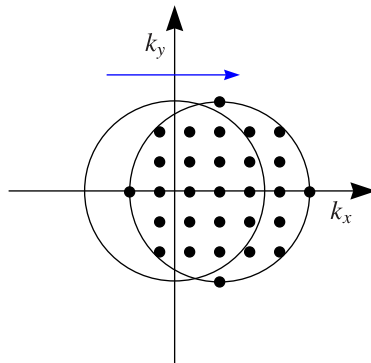
$$-\frac{e}{\hbar} \mathbf{E} \cdot \nabla_{\mathbf{k}} f = -\frac{1}{\tau} (f - f_0) \Rightarrow f(\mathbf{k}) = f_0(\mathbf{k}) + \frac{e}{\hbar} \tau \mathbf{E} \cdot \nabla_{\mathbf{k}} f(\mathbf{k}). \quad (1.21)$$

Der zweite Summand entspricht der Abweichung δf vom Gleichgewicht; sie führt zum elektrischen Strom. Zur Lösung dieser Differentialgleichung machen wir eine Iteration derart, dass wir $f(\mathbf{k})$ auf der rechten Seite durch $f_0(\mathbf{k})$ ersetzen. Daraus resultieren Beiträge zu f , die **linear** im elektrischen Feld sind. (Bei Iteration (erneutem Einsetzen) erhält man Beiträge zu f , die proportional zu \mathbf{E}^2 usw. sind. Das Ganze geht somit über die vereinfachte klassische Drude-Theorie, welche zum Ohmschen Gesetz führt, hinaus.) Wir begnügen uns zunächst mit dem linearen Term:

$$f(\mathbf{k}) \approx f_0(\mathbf{k}) + \frac{e}{\hbar} \tau \mathbf{E} \cdot \nabla_{\mathbf{k}} f_0(\mathbf{k}). \quad (1.22)$$

Bemerkung: Man kann (1.22) als Entwicklung um \mathbf{k} auffassen:

$$f(\mathbf{k}) \approx f_0\left(\mathbf{k} + \frac{e}{\hbar} \tau \mathbf{E}\right). \quad (1.23)$$



Die gesamte Verteilung verschiebt sich proportional zum elektrischen Feld, woraus ein Nettostrom (proportional zum elektrischen Feld) resultiert, falls das Band nicht leer oder komplett besetzt ist. Wir können nun die Stromdichte berechnen:

$$\mathbf{j} = -\frac{e}{8\pi^3} \int d^3k \mathbf{v}(\mathbf{k}) f(\mathbf{k}). \quad (1.24)$$

Dabei liefert $f_0(\mathbf{k})$ keinen Beitrag, weil sie inversionssymmetrisch bezüglich $\mathbf{k} = \mathbf{0}$ ist. Damit reicht es aus, den nächsten Term zu betrachten. Für $\mathbf{E} \parallel \mathbf{e}_x$, dann gilt $j_y = j_z = 0$ (falls der Halbleiter isotrop ist). Die Stromdichte in x -Richtung lautet

$$j_x = -\frac{e}{8\pi^3} \int d^3k v_x(\mathbf{k}) \frac{e}{\hbar} \tau E_x \frac{\partial f_0}{\partial k_x}(\mathbf{k}), \quad (1.25)$$

also

$$j_x = \sigma E_x, \quad \sigma = -\frac{e^2 \langle \tau \rangle}{8\pi^3 \hbar} \int d^3k v_x(\mathbf{k}) \frac{\partial f_0}{\partial k_x}(\mathbf{k}). \quad (1.26)$$

Vergleich mit Drude-Theorie:

$$\sigma = \frac{e^2 n \tau}{m^*}. \quad (1.27)$$

1.1 Zeitliche Inhomogenitäten bei räumlicher Homogenität

Zustandsgrößen hängen von der Zeit ab:

$$n = n(t) \neq n_0, \quad p = p(t) = p_0, \quad (1.28)$$

$$n_0 = N_L \exp\left(\frac{E_F - E_L}{k_B T}\right), \quad p_0 = N_V \exp\left(-\frac{E_F}{k_B T}\right), \quad (1.29)$$

wobei p_0 und n_0 die Gleichgewichtsdichten sind. Die zeitliche Veränderung der Ladungsträgerkonzentrationen n und p rührt von der Erzeugung und Vernichtung freier Ladungsträger her. Die Gesamtheit der Elektronen im Halbleiter bleibt dabei erhalten!

- Die Erzeugung und Vernichtung von freien Ladungsträgern erfolgt durch Austausch mit gebundenen Ladungsträgern. Donatoren und Akzeptoren spielen kaum eine Rolle.
- „flache“ Störstellenniveaus, die normalerweise ($T \sim 300$ K) ionisiert sind

Damit läuft die Erzeugung und Vernichtung durch **tiefe Störstellen**, beispielsweise atomare Punktdefekte im Gitter (Vakanz, interstitieller Defekt). Im thermodynamischen Gleichgewicht sind diese im Grundzustand, da sich so weit weg vom Leitungsband sind. Im Nichtgleichgewicht (beispielsweise durch Anlegen einer äußeren Spannung) wird die Bandstruktur verbogen dies kann zu einem Austausch von Elektronen bzw. Löchern mit Bändern zu angeregten Zuständen führen. **Elektron-Loch-Paare** entstehen, indem Elektronen aus dem Valenzband ins Leitungsband gehoben werden, wobei dann gleichzeitig eines freies Loch im Valenzband erzeugt wird. Solche Paare können sich wieder paarweise vernichten; man spricht von einem **bipolaren** Erzeugungs- bzw. Vernichtungsprozess. Dies steht im Gegensatz zur **unpolaren** Erzeugung bzw. Vernichtung, bei der nur eine Sorte von beweglichen Ladungsträgern vorliegt (beispielsweise tiefe Störstellen).

1.1.1 Erzeugungsprozesse

Die Erzeugung oder Vernichtung freier Ladungsträger erfolgt aus einem tiefen Niveau bzw. aus Bändern und erfordert Energie. Diese notwendige Energie kann auf alle möglichen Arten zur Verfügung gestellt werden, beispielsweise durch Licht, γ -Quanten (beim Szintillationszähler), Elektronen- oder Ionenstrahlen oder eine elektrische Spannung. Wir möchten zunächst das Prinzip der optischen Anregung (durch γ -Quanten, analog Elektronen- oder Ionenstrahlen) verstehen.



1.1. ZEITLICHE INHOMOGENITÄTEN BEI RÄUMLICHER HOMOGENITÄT

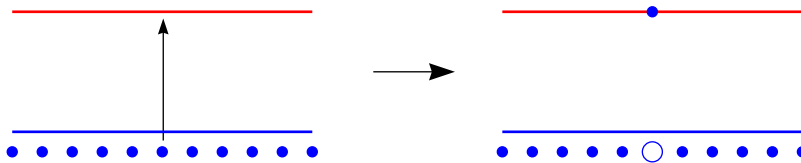
Da im Valenzband keine freien Löcher entstanden sind, handelt es sich um eine unipolare Erzeugung von Elektronen. Das Ganze funktioniert auch mit Löchern. Eine unipolare Erzeugung von Löchern werden durch solche Prozesse erfolgen:



Um diese Prozesse mathematisch zu erfassen, führen wir die Größen g_n und g_p ein, nämlich die Anzahl der pro Volumen und Zeit absorbierten Photonen.

$$\left(\frac{\partial n}{\partial t}\right)_{\text{unipolare Generation}} = g_n, \quad \left(\frac{\partial p}{\partial t}\right)_{\text{unipolare Generation}} = g_p. \quad (1.30)$$

Für die bipolare Generation gilt:



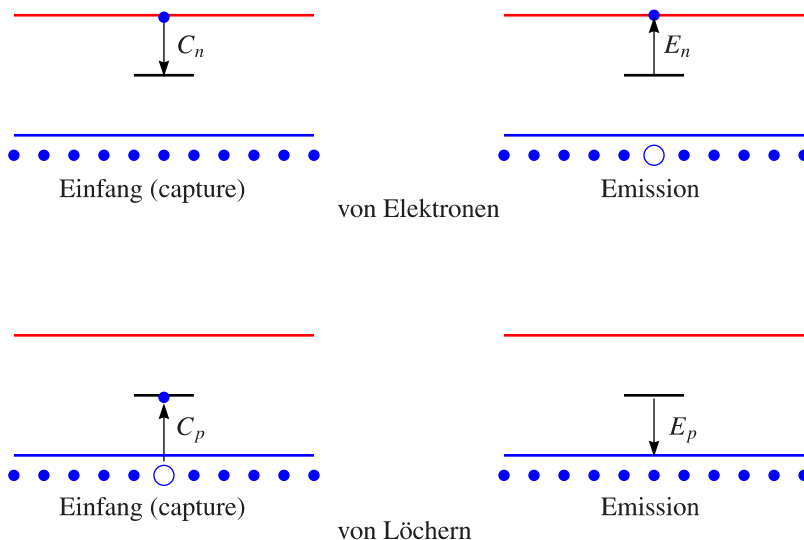
$$g_n = g_p = g = \left(\frac{\partial n}{\partial t}\right)_{\text{bipolare Generation}} = \left(\frac{\partial p}{\partial t}\right)_{\text{bipolare Generation}}. \quad (1.31)$$

Allgemein gilt dann weiter

$$g = \frac{\alpha I}{\hbar\omega}, \quad (1.32)$$

wobei I die Lichtintensität, α der Absorptionskoeffizient und $\hbar\omega$ die Energie der Photonen ist. Mit den typischen Werten $I = 1 \text{ W/cm}^2$, $\hbar\omega = 1 \text{ eV}$, $\alpha = 10^3 \text{ 1/cm}$ ergibt sich ein Wert der Größenordnung $g = 6 \cdot 10^{21} \text{ 1/(cm}^3 \cdot \text{s)}$. Man muss jedoch berücksichtigen, dass sich diese freien Ladungsträger auch wieder gegenseitig vernichten; die Zahl der Ladungsträger, die im Gleichgewicht übrig bleibt, ist um einiges kleiner.

Auf der anderen Seite gibt es natürlich auch Vernichtungsprozesse, die wir nun anschauen möchten. Bei der **unipolaren Rekombination** handelt es sich um den Einfang von freien Ladungsträgern an tiefen Zentren (Störstellen).



$$\left(\frac{\partial n}{\partial t}\right)_{\text{Einfang}} = -[C_n - E_n], \quad (1.33)$$

wobei auf der linken Seite die Netto-Einfangrate steht. C_n ist die Bruttoeinfangrate und E_n die Emissionsrate. Diese Größen können wir wie folgt schreiben als

$$C_n = c_n \cdot n \cdot (1 - f)N_t, \quad (1.34)$$

wobei n_f die Anzahl der tiefen Zentren, f die Besetzungswahrscheinlichkeit und c_n der Einfangkoeffizient ist. Analog gilt für die Emission

$$E_n = e \cdot f \cdot N_t, \quad (1.35)$$

mit dem Emissionskoeffizienten. Damit lässt sich die Netto-Einfangrate schreiben als

$$\left(\frac{\partial n}{\partial t} \right)_{\text{Einfang}} = -[c_n n(1-f)N_t - e_n f N_t]. \quad (1.36)$$

Betrachten wir nun den Spezialfall eines thermodynamischen Gleichgewichts. Dann sind alle makroskopischen Größen zeitunabhängig:

$$\left(\frac{\partial n}{\partial t} \right)_{\text{Einfang}} = 0. \quad (1.37)$$

Seien $n = n_0$, $p = p_0$ und $f = f_0$ die Gleichgewichtswerte, wobei

$$f_0 = \frac{1}{\beta_{t_n} \exp\left(\frac{E_F - E_g}{k_B T}\right) + 1}, \quad (1.38)$$

die Fermi-Verteilungsfunktion ist. β_{t_n} ist das Verhältnis von unbesetzten und besetzten Zuständen im statischen Gleichgewicht. Aus (1.36) folgt für das thermodynamische Gleichgewicht:

$$\boxed{e_n = n_t \cdot c_n, \quad n_t \equiv n_0 \frac{1-f_0}{f_0} = \beta_{t_n} N_L \exp\left(-\frac{E_g - E_F}{k_B T}\right)}. \quad (1.39)$$

Man spricht hierbei von einem „detaillierten Gleichgewicht“. n_t gibt bis auf β_{t_n} die Konzentration freier Elektronen für den Fall an, dass E_F genau im tiefsten Niveau liegt. Da im Allgemeinen E_F viel näher am Leitungsband liegt, gilt $n_t \ll n_0$. Im **Nichtgleichgewicht** weicht f von f_0 ab. Sowohl die Besetzung im Band als auch die der tiefsten Störstellen weicht somit von f_0 ab. Es sind somit nur Gleichungen für die Änderung von f , jedoch nicht f selbst angebar:

$$\frac{\partial n}{\partial t} + N_t \frac{\partial f}{\partial t} = 0, \quad (1.40)$$

was aus der Ladungsträgererhaltung folgt. Das Ganze gilt, falls es nur einen Austausch zwischen Leitungsband und tiefer Störstelle gibt. Anmerkung: Ab $t = t_0$ erfolgt keine weitere Erzeugung freier Ladungsträger aus tiefen Zentren, sondern nur noch Vernichtung durch Einfang.

$$\frac{\partial n}{\partial t} = \left(\frac{\partial n}{\partial t} \right)_{\text{Einfang}}, \quad (1.41)$$

für $t > t_0$. (1.36), (1.40) und (1.41) ist ein System nichtlinearer Differentialgleichungen, das exakt lösbar ist. Wir wollen hier jedoch eine lineare Näherung durchführen. Wir betrachten also eine kleine Abweichung von den Gleichgewichtswerten:

$$n = n_0 + \Delta n, \quad |\Delta n| \ll n_0, \quad f = f_0 + \Delta f, \quad |\Delta f| \ll f. \quad (1.42)$$

Die Rechnung schenken wir uns und kommen direkt zum Ergebnis:

$$\boxed{\Delta n(t) = \Delta n(t_0) \exp\left(-\frac{t-t_0}{\tau_n}\right)}, \quad (1.43)$$

die Abweichung vom Gleichgewicht zerfällt also exponentiell. Bei hoher Konzentration N_t und vollständig unbesetzten Zentren ($f_0 = 0$) gilt:

$$\tau_n \mapsto (\tau_n)_0 = \frac{1}{c_n N_t}, \quad (1.44)$$

was als Einfangszeit bezeichnet wird. $(\tau_n)_0$ ist die Zeit, in welcher die Gesamtheit von N_t unbesetzten tiefen Niveaus pro Kubikzentimeter ein Nichtgleichgewichtselektron aus dem Leitungsband einfangen kann.

1.2 Bipolare Rekombination

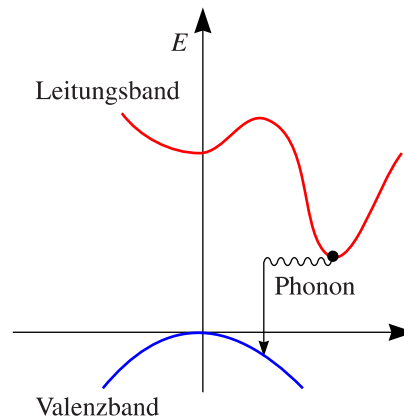
Betrachten wir zunächst die strahlende Rekombination, wobei ein Elektron-Loch-Paar unter Aussendung eines Photons rekombiniert. Man muss dabei unterscheiden zwischen:

- monomolekulare Rekombination (beispielsweise ein Exziton, bei dem das Elektron nur mit einem bestimmten Loch rekombinieren kann):

$$\frac{dn}{dt} \sim n, \quad (1.45)$$

was zu einem exponentiellen Zerfallsgesetz führt.

- bimolekulare Rekombination (beispielsweise Band-Band-Rekombination bei einem indirekten Halbleiter):



$$\frac{dn}{dt} \sim n \cdot p. \quad (1.46)$$

Hat man einen optisch gepaarten Prozess, gilt $n = p$, also

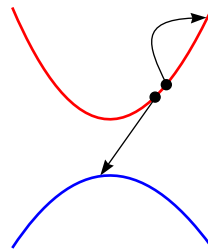
$$\frac{dn}{dt} \sim n^2, \quad (1.47)$$

und das Zerfallsgesetz wird nicht mehr exponentiell sein. Das ist meist nur in der Halbleiteroptik wichtig, da die stimulierte Emission Grundlage für Halbleiter-LASER ist.

Einfangmechanismen bei tiefen Störstellen sind die Multi-Phonon-Emission, Photon-Emission (Photolumineszenz, strahlender Übergang) und Auger-Prozesse (Anregung anderer Elektronen, bei der die freiwerdende Energie benutzt wird, um einen anderen Ladungsträger anzuregen.)

1.3 Nichtstrahlende Rekombination

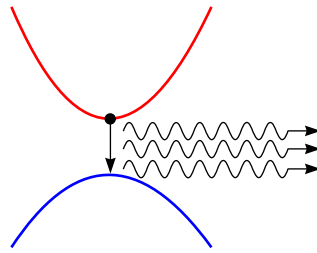
Nichtstrahlende Rekombination kann man dann bekommen, wenn zum ersten Auger-Rekombination vorliegt, also wenn ein weiteres Elektron oder auch Loch angeregt wird.



Da es sich um einen Dreikörperprozess handelt, ist die Abhängigkeit

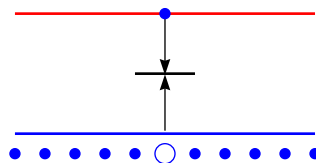
$$\frac{dn}{dt} \sim n^3. \quad (1.48)$$

Die Auger-Rekombination wird also vor allem bei hohen Ladungsträgerkonzentrationen wichtig. Bei einem intrinsischen Halbleiter findet der Prozess nicht statt, weil es keine freien Elektronen gibt. Anhand der zeitlichen Abhängigkeit von $n(t)$ lässt sich experimentell also bestimmen, welche Art von Rekombination vorliegt. Der zweite wichtige Prozess im Zusammenhang mit der nichtstrahlenden Rekombination ist die Multi-Phonon-Emission.



Für einen solchen Prozess sind jedoch sehr viele Phononen möglich und damit ist dieser unwahrscheinlich. Effizienter verläuft dagegen die **Schockley-Read-Hall-Rekombination**. Das ist der dominierende Mechanismus in Bauelementen. Hierbei benutzt man ein tiefes Zentrum als „Zwischenstation“; die Störstelle fängt sowohl Elektronen als auch Löcher an. Dann sitzen diese auf derselben Störstelle und können rekombinieren. Danach ist die tiefe Störstelle wieder frei und der Prozess kann sich wiederholen. Damit das Ganze möglich wird, muss ein tiefes Zentrum sowohl Elektronen als auch Löcher einfangen können. Die Rekombination über ein tiefes Zentrum ist um Größenordnungen wahrscheinlicher als nichtstrahlende Band-Band-Rekombination. Das tiefe Zentrum wirkt als

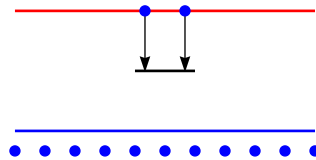
i.) Rekombinationszentrum



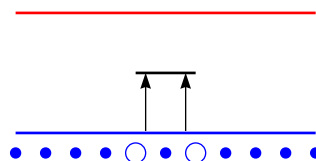
Dann muss die Einfangwahrscheinlichkeit für ein Elektron aus dem Leitungsband viel größer als der Einfang eines Elektrons aus dem Valenzband: $C_n \ll E_p$. Andererseits muss auch $C_p \gg E_n$ sein, so dass ein Loch effektiv eingefangen wird.

ii.) Elektronen-Reinfangzentrum:

Hier gilt $C_n \gg E_p$ und $C_p \ll E_n$, so dass effektiv Elektronen, aber keine Löcher eingefangen werden.

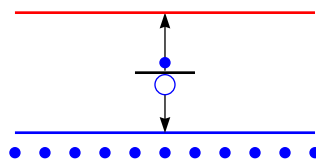


iii.) Loch-Einfangzentrum:



Es gilt $C_p \ll E_n$ und $C_n \ll E_p$; dann werden effektiv Löcher, aber keine Elektronen eingefangen.

iv.) Generationszentrum:



Mit $C_p \ll E_n$ und $C_n \ll E_p$ wirkt das Generationszentrum entgegengesetzt zum Rekombinationszentrum; erzeugt als Elektron-Loch-Paare anstelle diese zu vernichten.

Bei Fall (1), der Schockley-Read-Hall-Rekombination ist die treibende Kraft der Rekombination die Abweichung des Konzentrationsproduktes $n \cdot p$ vom Gleichgewichtswert $n_0 \cdot p_0$. Die Rekombinationslebensdauer ist durch die Minoritätsladungsträger (Minoritäten) bestimmt. Im intrinsischen Falle wird diese Lebensdauer gerade maximal. Die Dotierung eines Halbleiters reduziert die Rekombinationslebensdauer.

1.4 Räumliche Inhomogenitäten und Transportphänomene

Wir suchen die Gleichungen, durch welche man solche Phänomene beschreiben kann: man bezeichnet diese als Transportgleichungen. Ströme werden getrieben durch Gradienten im Ortsraum. Wir gehen zunächst von der Boltzmann-Transportgleichung in Relaxationszeitznäherung aus:

$$\frac{\delta f}{\tau} + \dot{\mathbf{k}} \nabla_{\mathbf{k}} f_0 + \dot{\mathbf{r}} \nabla_{\mathbf{r}} f = 0. \quad (1.49)$$

In $\nabla_{\mathbf{r}} f$ tauchen Terme der Form $\nabla \zeta$ (Gradienten im chemischen Potential) und auch $\nabla_{\mathbf{r}} T$ (Gradienten in der Temperatur) auf. Die Kraft durch elektrische, magnetische Felder und Kompositionsänderungen (durch Zusammenführen verschiedener Halbleiter, $\text{Al}_x \text{Ge}_{1-x} \text{As}_s$) ist gegeben durch

$$\dot{\mathbf{k}} = \frac{1}{\hbar} \left\{ -\nabla_r (E_g - e\varphi) - \frac{e}{\hbar} (\nabla_{\mathbf{k}} \mathbf{E} \times \mathbf{B}) \right\}. \quad (1.50)$$

Einsetzen in die Boltzmann-Gleichung (1.49) liefert die elektrische Stromdichte

$$\mathbf{j} = -\frac{e}{\hbar} \int \nabla_{\mathbf{k}} E(\mathbf{k}) f(\mathbf{r}, \mathbf{k}) D(\mathbf{k}) d^3 k, \quad (1.51)$$

und die mittlere Energiestromdichte (Gewichtung der Teilchenstromdichte mit der Energie)

$$\boldsymbol{\omega} = \frac{1}{\hbar} \int (\nabla_{\mathbf{k}} E(\mathbf{k})) E(\mathbf{k}) f(\mathbf{r}, \mathbf{k}) D(\mathbf{k}) d^3 k. \quad (1.52)$$

Nur Terme mit δf liefern Ströme, also

$$\mathbf{j} = N_{11} \nabla \left(\frac{\eta}{e} \right) + N_{12} \left(-\frac{\nabla T}{T} \right), \quad (1.53)$$

und den Wärmestrom

$$\boldsymbol{\omega} = N_{21} \nabla \left(\frac{\eta}{e} \right) + N_{22} \left(-\frac{\nabla T}{T} \right). \quad (1.54)$$

Elektrische Ströme lassen sich somit auch die Temperaturgradienten treiben (Thermoelement), genauso wie sich ein Wärmestrom durch einen Potentialgradienten treiben lässt. Der elektrische Strom ist also mit dem Wärmestrom verknüpft, auch weil die Elektronen sowohl Ladung tragen als auch für die Wärmeleitung im Metall verantwortlich sind. Im Allgemeinen sind die Koeffizienten N_{ik} Tensoren für anisotrope Materialien, da dann effektive Masse und Beweglichkeit der Elektronen von der Richtung abhängen. Anisotropie kann ebenso durch ein Magnetfeld \mathbf{B} verursacht werden, welche an der Probe anliegt. Für isotrope Halbleiter sind die N_{ik} skalare (ohne Magnetfeld). Weiterhin gelten die **Onsager-Beziehungen**:

$$N_{ik}(\mathbf{B}) = N_{ki}^\dagger(-\mathbf{B}). \quad (1.55)$$

Um experimentell zugängliche Größen zu erhalten, formt man (1.49) um (ohne Magnetfeld):

$$\boxed{\nabla \left(\frac{\eta}{e} \right) = \frac{\mathbf{j}}{\sigma} + \varepsilon \nabla T}, \quad (1.56)$$

und der Wärmestromdichte

$$\boxed{\boldsymbol{\omega}_q = \Pi \mathbf{j} - \chi \nabla T}, \quad (1.57)$$

mit der spezifischen elektrischen Leitfähigkeit $\sigma = N_{11}$. $\varepsilon = N_{12}/(TN_{11})$ heißt Seebeck-Koeffizient (Thermospannung). $\kappa = N_{11}N_{22} - N_{12}N_{21}$ ist die spezifische Wärmeleitfähigkeit. $\Pi = N_{21}/N_{11} = \varepsilon T$ nennt man Peltier-Koeffizient. Beim Peltier-Element erzeugt man einen Temperaturgradient durch einen elektrischen Strom.

1.5 Drift von Ladungsträgern

Sei $\nabla T = 0$, $n = n_0$, $p = p_0$ (homogen). Bei konstanten Ladungsträgerdichten hat man kein chemisches Potential und das elektrochemische Potential ist durch das äußere elektrische Potential gegeben:

$$\nabla \left(\frac{\eta}{e} \right) = -\nabla \varphi, \quad (1.58)$$

ζ ist also konstant. Dann gilt

$$\mathbf{j} = -\sigma \nabla \varphi = \sigma \mathbf{E}, \quad (1.59)$$

und dabei handelt es sich um das Ohmsche Gesetz mit dem Driftstrom \mathbf{j} . Grundsätzlich hat man Beiträge sowohl von den Elektronen als auch von den Löchern:

$$\mathbf{j} = \mathbf{j}_n + \mathbf{j}_p = (\sigma_n + \sigma_p) \mathbf{E}. \quad (1.60)$$

Kommen wir kurz nochmal auf das **Drude-Modell** zurück. Für Elektronen gilt mit ihrer effektiven Masse m^* nach der Newtonschen Mechanik:

$$m^* \frac{d\langle \mathbf{v} \rangle}{dt} = m^* \left. \frac{d\langle \mathbf{v} \rangle}{dt} \right|_{\text{Feld}} + m^* \left. \frac{d\langle \mathbf{v} \rangle}{dt} \right|_{\text{Stöße}}, \quad \left. \frac{d\langle \mathbf{v} \rangle}{dt} \right|_{\text{Feld}} = -e \mathbf{E}, \quad m^* \left. \frac{d\langle \mathbf{v} \rangle}{dt} \right|_{\text{Stöße}} = -\frac{m^*}{\tau} \langle \mathbf{v} \rangle, \quad (1.61)$$

wobei τ die Impulsrelaxationszeit ($\tau(\mathbf{k})$) ist. Im stationären Zustand gilt

$$\frac{d\langle \mathbf{v} \rangle}{dt} = 0 \Rightarrow \langle \mathbf{v}_e \rangle = \mathbf{v}_D = -\frac{e\tau_e}{m_e^*} \mathbf{E} = -\mu_n \mathbf{E}, \quad (1.62)$$

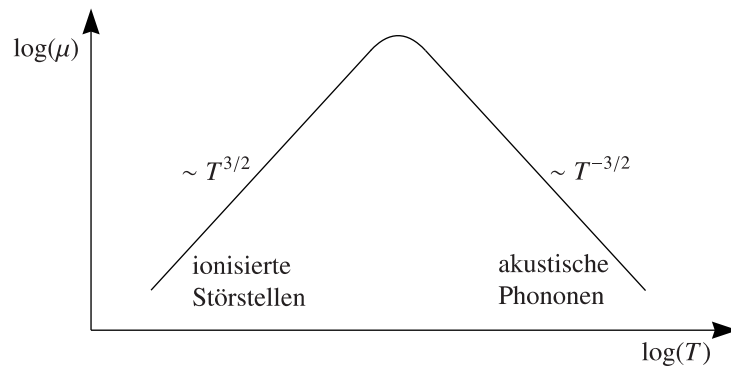
mit der Beweglichkeit μ_n für Elektronen. Für Löcher ist analog die Beweglichkeit μ_h einzusetzen. Dann gilt:

$$\mathbf{j}_h = -en\mathbf{v}_{D,e} = en\mu_n \mathbf{E} \Rightarrow \sigma_n = en\mu_n = \frac{e^2 n \tau_e}{m_e^*}, \quad (1.63)$$

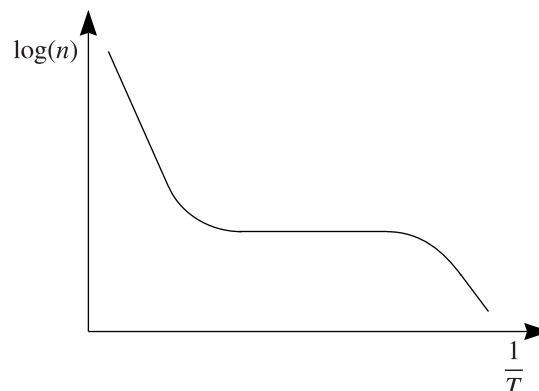
$$\mathbf{j}_p = ep\mathbf{v}_{D,h} = ep\mu_p \mathbf{E} \Rightarrow \sigma_p = ep\mu_p = \frac{e^2 p \tau_h}{m_h^*}. \quad (1.64)$$

Es stellt sich ein Gleichgewicht zwischen der äußeren beschleunigenden Kraft und der Reibung ein. Beispiel: Für Galliumarsenid ist m_e^* klein im Leitungsband, weshalb dieses Halbleitermaterial sehr geeignet in der Hochgeschwindigkeitselektronik ist.

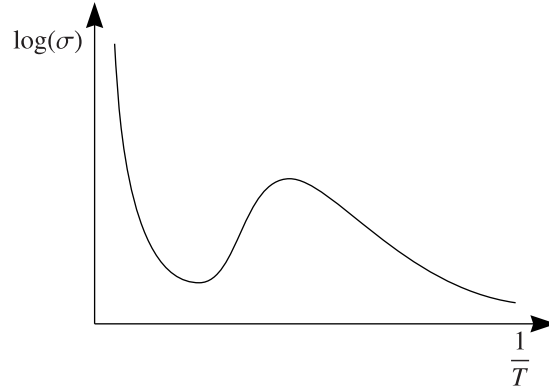
Aus der Festkörperphysik wissen wir, dass σ (als Funktionen von $n(T)$ und $p(T)$) und ebenso μ temperaturabhängig ist.



Die Streuung an ionisierten Störstellen ist nicht der einzige Prozess, welcher die Beweglichkeit verändert. Bei hohen Temperaturen ist ebenso die Streuung an Phononen, also quantisierten Gitterschwingungen wichtig. Dies führt bei Raumtemperatur wieder zu einem Abfall der Beweglichkeit. Mit dem Verhalten der Ladungsträgerkonzentration



ergibt sich schließlich die Temperaturabhängigkeit der Leitfähigkeit:



Das Bild gilt in der Form nur, wenn man keine Entartung hat, wenn also die Dotierung nicht zu stark ist.

1.5.1 Hochfeldeffekte

Die Beziehung $\mathbf{v}_D = \pm\mu\mathbf{E}$ ist nur eine gute Näherung für kleine Felder \mathbf{E} . Hier schlägt sich also die Tatsache nieder, dass μ unabhängig vom Feld ist und $v_D \ll v_{\text{thermisch}}$, $v_D \ll v_{\text{Schall}}$, $v_D \ll v_F$, wobei v_F die Geschwindigkeit der Elektronen auf der Oberfläche der Fermi-Kugel ist. Für hohe Felder sind die letzten genannten Bedingungen nicht mehr erfüllt, was zu einer Nichtlinearität und Sättigung der Drift führt. Außerdem können die Elektronen so eine hohe Energie haben, dass sie Valenzelektronen ins Leitungsband anheben (Stoßionisation). Sättigung der Driftgeschwindigkeit: Im thermodynamischen Gleichgewicht können Elektronen ohne weiteres Phononen absorbieren oder emittieren. Im klassischen Grenzfall gilt für beide eine Boltzmannverteilung mit $T_e = T_{\text{Gitter}}$. Bei anliegendem Feld nehmen die Elektronen Energie auf und geben sie durch Emission von Phononen wieder ans Gitter ab. Bei hohen Feldern spielen die optischen Phononen eine größere Rolle als die akustischen, weil die optischen im Allgemeinen eine größere Energie haben. Dadurch heizt sich sowohl das Elektronensystem als auch das Gitter auf. Für $v_D \ll v_{\text{thermisch}}$ gilt $T_e = T_{\text{Gitter}}$. Für sehr hohe Felder werden die Elektronen jedoch so stark beschleunigt, dass sie ihre Energie nicht mehr schnell genug an das Gitter abgeben können. (Die Zeit geht proportional zum Kehrwert der Schwingungsdauer eines Phonons.) Das führt dazu, dass die Elektronen eine Fermi-/Boltzmannverteilung mit $T_e > T_{\text{Gitter}}$ („heiße Elektronen“) haben. Falls $\mu\mathbf{E} \approx \mathbf{v}_{\text{Schall}}$, gilt

$$T_e \approx T_{\text{Gitter}} \left[1 + \frac{3\pi}{32} \left(\frac{\mu_0 E}{v_{\text{Schall}}} \right)^2 \right], \quad (1.65)$$

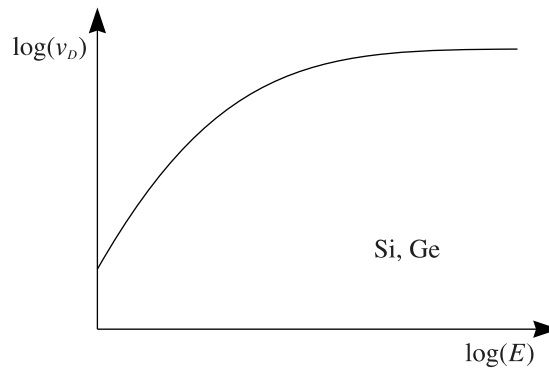
und

$$v_D \approx \mu_0 E \left[1 - \frac{3\pi}{64} \left(\frac{\mu_0 E}{v_{\text{Schall}}} \right)^2 \right], \quad (1.66)$$

wobei μ_0 die Beweglichkeit bei kleinen Feldern ist. Dann tritt somit ein Sättigungseffekt in der Driftgeschwindigkeit auf; diese nimmt nicht weiter zu, sondern flacht ab. Bei höchsten Feldern tritt eine Sättigung der Form

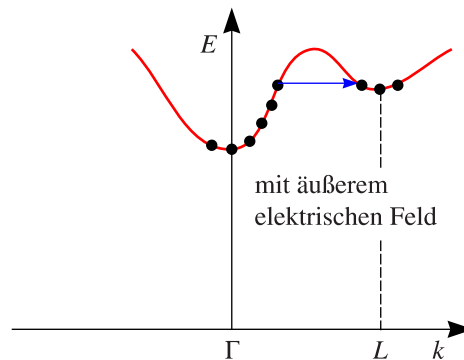
$$v_{\text{Sat}} = \sqrt{\frac{8\hbar\omega_{\text{lo}}}{3\pi m}} \approx 10^7 \frac{\text{cm}}{\text{s}}, \quad (1.67)$$

auf. Hier ist $\hbar\omega_{\text{lo}}$ die Energie eines longitudinalen optischen Phonons. (Das liegt daran, dass das elektrische Feld solche Phononen anregt; da die positiven Ionen in die eine Richtung und die negativen in die entgegengesetzte Richtung ausgelenkt werden und diese gegenphasig schwingen. Optische Phononen regt man vor allem optisch an. Für einen Kristall gibt es $3N$ Phononen, von denen 3 akustisch sind, weil sie eine Schwingung in jeder der drei Dimensionen ausbreiten kann.)



1.5.2 Negativer differentieller Widerstand

Dabei handelt es sich um Hochfeldeffekte in GaAs beeinflusst durch die Vielval-Bandstruktur.



Für genügend hohe elektrische Felder ist ein Transfer in Seitentäler (hier der L-Punkt) möglich. Am L-Punkt ist die Bandstruktur jedoch flacher als beim Γ -Punkt und damit ist die effektive Masse größer: $m_L^* \gg m_\Gamma^*$. (Typische Zahlenwerte sind $m_L^* \approx 1,2m_0 \approx m_0$ und $m_\Gamma^* \approx 0,068m_0$.) Eine hohe effektive Masse führt jedoch dann auch zu einer hohen Zustandsdichte im L-Minimum, da die Zustandsdichte proportional zu $(m^*)^{\frac{3}{2}}$ ist. Die Beweglichkeit dieser Elektronen ist jedoch klein, da $\mu \sim 1/m^*$. Dies führt zu einem negativen differentiellen Widerstand (negative differential resistance, NDR) bei hohem Feld, also $dj/dE < 0$. (Der Strom nimmt also ab bei Erhöhung des elektrischen Feldes.) Unter bestimmten Bedingungen gibt es dann keinen Stromfluss mehr durch den Halbleiter, sondern es entstehen Strominstabilitäten und Oszillationen. Diese Effekte werden im Gamm-Oszillator ausgenutzt zur Erzeugung von Mikrowellen.

1.5.3 Stoßionisation

Bei Feldern oberhalb eines Schwellenwerts haben die Ladungsträger genügend Energie, um ein Elektron-Loch-Paar anzuregen. Die Generationsrate ist dann $G = \alpha_n n v_D^e = \alpha_p p v_D^p$. Die Parameter α_n bzw. α_p ist die sogenannte Ionisationsrate; diese ist feldabhängig. Der Effekt führt zur Erhöhung der freien Ladungsträgerdichten und beispielsweise zum Lawinendurchbruch bei in Sperrichtung gepoltem p-n-Übergang, welcher zur Zerstörung des Bauelements führen kann.

Kapitel 2

Die Differentialgleichungen der inneren Elektronik

Dabei handelt es sich um ein System von Basisgleichungen der Elektronik. Um dieses aufzustellen, müssen wir zusätzlich die räumlichen Inhomogenitäten der Ladungsträgerdichte in die Betrachtung mit einbeziehen. Dann treten neben den Driftströmen auch Diffusionsströme auf wegen der vorliegenden Konzentrationsgradienten. Dann bilden sich Raumladungen (analog zum p-n-Übergang) an bestimmten Stellen im Bauelement aus, also hängen die Konzentrationen und Felder vom Ort ab: $n = n(\mathbf{r}, p = p(\mathbf{r}), \mathbf{E} = \mathbf{E}(\mathbf{r}) \neq 0$. Dann ist die Stromdichte gegeben durch

$$\mathbf{J}_n = \sigma_n \mathbf{E}(\mathbf{r}) + e D_n \nabla n(\mathbf{r}), \quad (2.1)$$

für die Elektronen und

$$\mathbf{J}_p = \sigma_p \mathbf{E}(\mathbf{r}) - e D_p \nabla p(\mathbf{r}), \quad (2.2)$$

für die Löcher. Bei den ersten Summanden handelt es sich um die Feldströme, bei den zweiten um die Diffusionsströme. Im thermodynamischen Gleichgewicht herrscht die Neutralitätsbedingung $n + n_A^- = p + n_D^+$. Diese gilt jetzt aber aufgrund der sich bildenden Raumladungen nicht mehr lokal, sondern nur noch für den gesamten Halbleiter. Es herrscht somit **keine** lokale Neutralität mehr. Die lokale Ladungsdichte $\rho(\mathbf{r})$ ist

$$\rho(\mathbf{r}) = -e[n(\mathbf{r}) - p(\mathbf{r}) - n_D^+(\mathbf{r}) + n_A^-(\mathbf{r})]. \quad (2.3)$$

Die Ladungsträgerneutralität für den gesamten Halbleiter kann man dann schreiben als

$$\int_{\text{Halbleiter}} d^3r \rho(\mathbf{r}) = 0. \quad (2.4)$$

Mit der Raumladung ist ein elektrisches Feld verknüpft: $\mathbf{E}(\mathbf{r}) = -\nabla\varphi(\mathbf{r})$ mit dem elektrostatischen Potential, das sich aus der **Poisson-Gleichung** ergibt:

$$\Delta\varphi(\mathbf{r}) = \nabla \cdot (\nabla\varphi(\mathbf{r})) = -\frac{1}{\varepsilon\varepsilon_0} \rho(\mathbf{r}), \quad (2.5)$$

wobei wir hier angenommen haben, dass ε nicht vom Ort abhängt. Ansonsten müsste man dies bei der Differentiation auf der linken Seite berücksichtigen. Hängt ε vom Ort ab, führt dies dazu, dass sich Oberflächenladungen ansammeln. Falls zusätzlich eine zeitliche Inhomogenität $n(\mathbf{r}, t)$, $p(\mathbf{r}, t)$ und $\mathbf{E}(\mathbf{r}, t)$ auftritt, hängen die Ströme selbst zusätzlich von der Zeit ab: $\mathbf{j}_{n,p}(\mathbf{r}, t)$, $\varphi(\mathbf{r}, t)$. Wie sich die Ladungsträgerdichten mit der Zeit ändern, beschreibt man durch die Kontinuitätsgleichung

$$\frac{\partial n(\mathbf{r}, t)}{\partial t} = \frac{1}{e} \nabla \cdot \mathbf{j}_n(\mathbf{r}, t) + (G_n - R_n), \quad (2.6)$$

in der prinzipiell die Teilchenzahlerhaltung der Elektronen steckt. Zusätzlich müssen Generations- oder Rekombinationsprozesse (Shockley-Reed-Hall-Rekombination, Einfang durch Störstellen) berücksichtigt werden, was durch den zweiten und dritten Term geschieht. Eine analoge Gleichung gilt für die Löcher:

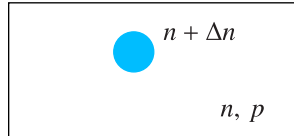
$$\frac{\partial p(\mathbf{r}, t)}{\partial t} = -\frac{1}{e} \nabla \cdot \mathbf{j}_p(\mathbf{r}, t) + (G_p - R_p). \quad (2.7)$$

Die Stromgleichungen, die Kontinuitätsgleichungen und die Poissongleichung bilden ein System **gekoppelter** Differentialgleichungen. Dabei handelt es sich um die Differentialgleichungen der inneren Elektronik. Zusätzlich gilt die Einstein-Relation

$$D_{n,p} = \frac{k_B T}{e} \mu_{n,p}. \tag{2.8}$$

2.1 Die dielektrische Relaxationszeit und die Debye-Länge

Die dielektrische Relaxationszeit τ_{rel} ist eine Zeitkonstante, auf der sich ausgehend von einem Nichtgleichgewichtszustand das Gleichgewicht wieder einstellt. Betrachten wir einen Halbleiter, in dem sich eine Raumladung ausgebildet hat:



Die Raumladung wird durch Diffusion und Abstoßung der Elektronen (Feldströme) abgebaut. Wir nehmen an, dass die Rekombinationslebensdauer τ_{rek} unendlich groß ist. (Später wird sich herausstellen, dass diese im Bereich von 10^{-9} s liegt und viel größer als τ_{rel} ist.) Vernachlässigen wir zusätzlich die Diffusion, ergibt sich eine obere Grenze für τ_{rel} .

Sei $\Delta n \ll n_0$ eine kleine Abweichung der Ladungsträgerkonzentration im Gleichgewicht. Dann gilt die Kontinuitätsgleichung in der Form

$$\frac{\partial \varrho}{\partial t} = -\nabla \cdot \mathbf{j} = -\nabla \cdot (\sigma \mathbf{E}). \tag{2.9}$$

Wir vernachlässigen hierbei außerdem die Abhängigkeit der Leitfähigkeit von der Ladungsträgerkonzentration und setzen $\sigma \approx \sigma_0$:

$$\frac{\partial \varrho}{\partial t} \approx -\sigma_0 \nabla \cdot \mathbf{E} = -\frac{\sigma_0}{\varepsilon \varepsilon_0} \varrho(x). \tag{2.10}$$

Diese Differentialgleichung wird gelöst durch

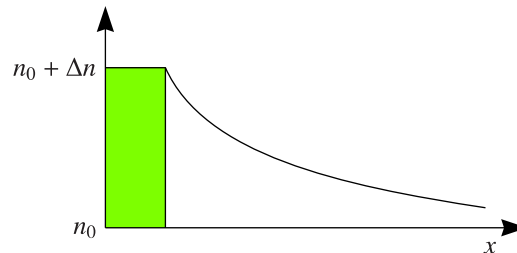
$$\varrho(t) \sim \exp\left(-\frac{t}{\tau_{\text{rel}}}\right), \quad \tau_{\text{rel}} = \frac{\varepsilon \varepsilon_0}{\sigma_0}, \tag{2.11}$$

mit der dielektrischen Relaxationszeit τ_{rel} . Beispielsweise besitzt Germanium mit $\varepsilon = 16$, $n_0 = 10^{16} \text{ cm}^{-3}$ eine Relaxationszeit $\tau_{\text{rel}} = 2,5 \cdot 10^{-13} \text{ s} \ll \tau_{\text{rek}}$ im Einklang mit obiger Annahme. Wir betrachten den Abbau des Nichtgleichgewichtswerts beispielsweise der Elektronenwolke im p-Halbleiter, wo $\sigma_p \ll \sigma_n$ gilt:

$$\frac{\sigma_p}{\sigma_n} = \frac{\mu_p p e}{\mu_n n e} = \frac{\mu_p p}{\mu_n n}. \tag{2.12}$$

Die Elektronenwolke wird dadurch kompensiert, dass von außen Löcher in die Wolke fließen, anstelle dass Elektronen nach außen fließen. Im Allgemeinen zerfließt eine Majoritätswolke und eine Minoritätswolke wird durch Zufluss von Majoritäten neutralisiert.

Als nächstes möchten wir untersuchen, über welche Strecke hinweg die Störung räumlich abklingt.



Mit

$$\mathbf{j} = \sigma \mathbf{E} + e D_n \nabla (\Delta n(\mathbf{r})), \tag{2.13}$$

erhalten wir

$$0 = \frac{\partial \varrho}{\partial t} = -\nabla \cdot \mathbf{j} = -\sigma_0 \nabla \cdot \mathbf{E} - eD_n \nabla \cdot (\nabla \Delta n(\mathbf{r})), \quad (2.14)$$

mit dem Feldstrom als ersten und dem Diffusionsstrom als zweiten Beitrag. Unter Verwendung von (2.10) gilt dann:

$$\boxed{-\frac{\varrho}{\tau_{\text{rel}}} + D_n \Delta \varrho = 0.} \quad (2.15)$$

Die Lösung im eindimensionalen Falle ist

$$\varrho(x) \sim \exp\left(-\frac{x}{L_D}\right), \quad L_D = \sqrt{D\tau_{\text{rel}}} = \sqrt{\frac{\varepsilon\varepsilon_0 k_B T}{e^2 n_0}}, \quad (2.16)$$

wobei L_D die sogenannte Debye-Abschirmlänge ist. Sie gibt Auskunft darüber, über welche Länge das Feld, welches durch die Ladungsträgerverteilung gemacht wird, abgeschirmt wird. Beispielsweise gilt für Germanium ($n_0 = 10^{16} \text{ cm}^{-3}$, $\varepsilon = 16$) $L_D = 50 \text{ nm}$ und ist somit sehr begrenzt.

2.2 Die Diffusionslänge

Wir betrachten eine **neutrale** Dichteabweichung, deren Kompensation vollständig durch Diffusion und nicht durch elektrische Felder getrieben wird. Sei $n = n_0 + \Delta n$ und $p = p_0 + \Delta p$, wobei $\Delta n = \Delta p$ ist. Die Kontinuitätsgleichung lautet für diesen Fall

$$\frac{\partial n}{\partial t} = \frac{\partial \Delta n}{\partial t} = \frac{1}{e} \nabla_{\mathbf{r}} \cdot \mathbf{j}_n + G - \frac{\Delta n}{\tau_{\text{rek}}}, \quad (2.17)$$

für Elektronen und

$$\frac{\partial p}{\partial t} = \frac{\partial \Delta p}{\partial t} = -\frac{1}{e} \nabla_{\mathbf{r}} \cdot \mathbf{j}_p + G - \frac{\Delta p}{\tau_{\text{rek}}}, \quad (2.18)$$

für Löcher, wobei jeweils die letzten beiden Terme die Rekombination (beispielsweise Shockley-Reed-Hall-Rekombination) berücksichtigen. Weiterhin gilt

$$\mathbf{j}_n = \sigma_n \mathbf{E} + eD_n \nabla n, \quad (2.19)$$

und

$$\mathbf{j}_p = \sigma_p \mathbf{E} - eD_p \nabla p. \quad (2.20)$$

\mathbf{E} ist ein internes elektrisches Feld durch die Raumladung, das bei unterschiedlich schneller Diffusion von Elektronen und Löchern hervorgerufen wird; man spricht dabei von **ambipolarer Diffusion**. Da wir dieses Feld nicht kennen, eliminieren wir es in (2.19) und (2.20), also

$$\frac{\mathbf{j}_n - eD_n \nabla n}{\sigma_n} = \frac{\mathbf{j}_p + eD_p \nabla p}{\sigma_p}, \quad (2.21)$$

mit $\sigma_n \sim n\mu_n \sim nD_n$ und $\sigma_p \sim pD_p$ gilt:

$$\frac{\mathbf{j}_n}{nD_n} - \frac{\mathbf{j}_p}{pD_p} = e \left(\frac{\nabla n}{n} + \frac{\nabla p}{p} \right). \quad (2.22)$$

Da das Ganze ambipolar erfolgt, gilt $\mathbf{j} = \mathbf{j}_n + \mathbf{j}_p = 0$ und hieraus folgt $\mathbf{j}_p = -\mathbf{j}_n$. Weiterhin ist $\nabla n = \nabla p$ und somit:

$$\frac{\mathbf{j}_n(nD_n + pD_p)}{nD_n \cdot pD_p} = e \frac{\nabla n(n+p)}{np}. \quad (2.23)$$

Also gilt

$$\mathbf{j}_n = eD \nabla n, \quad D = \frac{(n+p)D_n D_p}{nD_n + pD_p}, \quad (2.24)$$

mit dem Diffusionskoeffizienten D für ambipolare Diffusion (für einen n-Halbleiter: D_p , p-Halbleiter: D_n). Die Diffusion einer neutralen Dichteabweichung wird durch Diffusion von Minoritäten bestimmt.

Kapitel 3

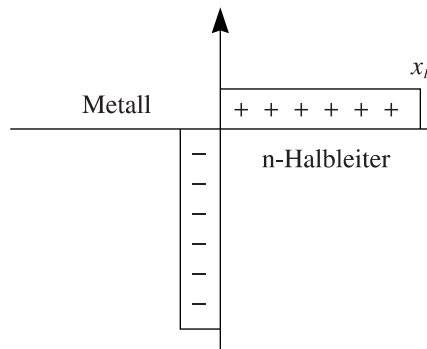
Metall-Halbleiter-Übergänge

3.1 Schottky-Kontakt und ohmscher Kontakt

Vor dem Einstellen des thermodynamischen Gleichgewichts: Typisch ist $E_F^M < E_F^{\text{HL}}$ (n-typ) und $E_F^M > E_F^{\text{HL}}$ (p-typ). Es bildet sich eine Schottky-Barriere ϕ_B , die gegeben ist durch:

$$\phi_B \approx \begin{cases} \phi - \chi \\ (E_0 - E_V) - \phi = \chi + E_g - \phi \end{cases} \quad (3.1)$$

für hochdotierte Halbleiter. Zur Einstellung des thermodynamischen Gleichgewichts diffundieren die Elektronen vom n-Halbleiter oder analog die Löcher vom p-Halbleiter ins Metall. Dann entstehen zwei Raumladungszonen an der Grenzfläche.



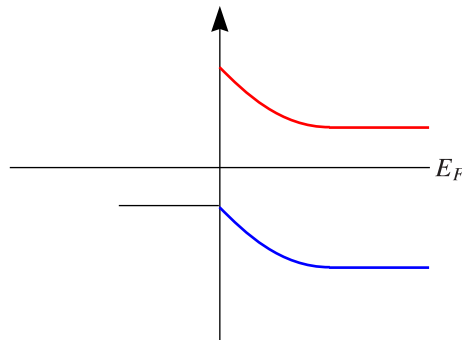
Wie breit die Abschirmmägen sind, lässt sich abschätzen. Im Metall ist die Randschicht extrem dünn. Im Grenzfall hoher Dichte gehen wir von der Debye-Länge über in die Thomas-Fermi-Länge:

$$l_D = \sqrt{\frac{\epsilon\epsilon_0 k_B T}{e^2 n_D}} \mapsto l_{\text{TF}} = \sqrt{\frac{2\epsilon\epsilon_0 k_B T}{3e^2 n_D}} \quad (3.2)$$

Setzt man beispielsweise $n_D \approx 10^{22} \text{ cm}^{-3}$, so ist $l_{\text{TF}} \approx 0,05 \text{ nm}$.

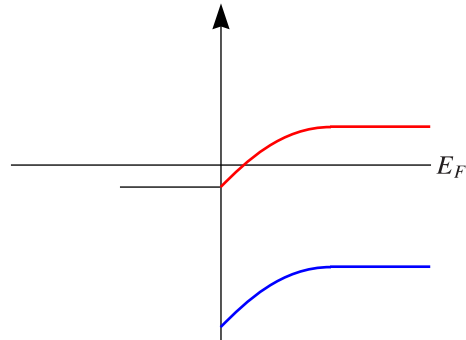
i.) $E_F^{\text{HL}} > E_F^M$:

Da sich die Elektronen im Halbleiter sammeln, biegt sich das Band nach oben.



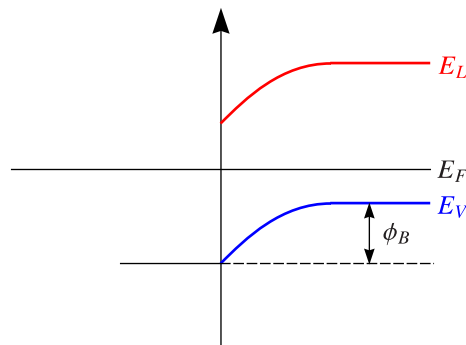
Beispiele: Si, GaAs usw. (fast alle Halbleiter). Hier müssen die Elektronen durch die Barriere tunneln und es handelt sich um einen Schottky-Kontakt. Ist die Tunnelbarriere sehr klein (bei hoher Dotierung), spricht man auch von einem ohmschen Kontakt. Bei kleiner Spannung fließt sofort ein Strom und der Kontakt verhält sich wie ein ohmscher Widerstand.

ii.) $E_F^{\text{HL}} < E_F^{\text{M}}$:



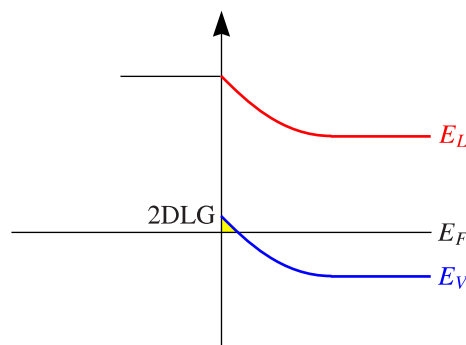
Für die wenigsten Halbleiter ist das der Fall; die prominentesten Beispiele sind InAs und InN. Es handelt sich um einen ohmschen Kontakt, weil das Kontaktpotential negativ ist.

iii.) $E_F^{\text{HL}} < E_F^{\text{M}}$:

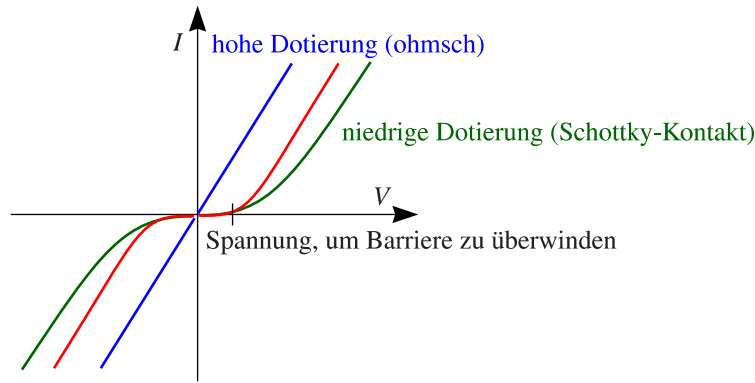


Hier stellt sich wieder ein Schottky-Kontakt ein; für die Löcher gibt es eine Barriere.

iv.) $E_F^{\text{HL}} > E_F^{\text{M}}$:



Dieser Kontakt ist wieder ohmsch, da es für die Löcher keine Barriere gibt.



Der Potentialverlauf ist ähnlich wie beim p-n-Übergang. Das gilt nur in der Schottky-Näherung, wenn also $eU_k \gg k_B T$. Die Randbedingungen zur Lösung der Poisson-Gleichung sind $\varrho(x)|_{x=0} = 0$ und $\varphi(x)|_{x=x_B=\infty} = U_k$. Die Lösung lautet dann:

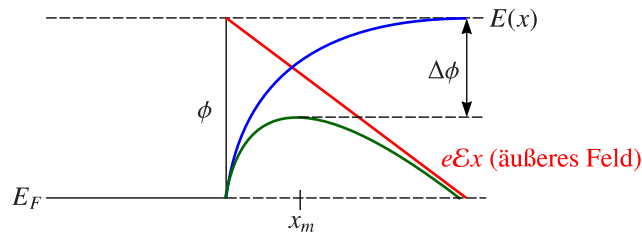
$$\varphi(x) = \begin{cases} 0 & \text{für } x \leq 0 \\ -e/(2\varepsilon\varepsilon_0)n_D(x^2 - 2xx_B) & \text{für } 0 < x < x_B \\ U_k & \text{für } x \geq x_B \end{cases}, \quad x_B = \sqrt{\frac{2\varepsilon\varepsilon_0}{en_D}U_k}. \quad (3.3)$$

3.2 Höhe der Schottky-Barriere

Notwendig zur Berechnung des Stromflusses im idealen Fall wäre $\phi_B = \phi - \chi$ (Mott-Beziehung). Diese Beziehung ist jedoch (fast) nie erfüllt, aufgrund der

- Bildladungen, die zum sogenannten Schottky-Effekt führen
 - und wegen des Fermineau-Pinning durch Grenzflächenzustände.
- i.) Schottky-Effekt: Betrachte den Metall-Vakuum-Übergang. Wenn das Elektron im Metall austritt, entsteht eine positive Bildladung im Metall im gleichen Abstand x von der Oberfläche. Die potentielle Energie ist gegeben durch:

$$E(x) = \int_{-\infty}^x F dx = \int_{-\infty}^x \frac{-e^2}{4\pi\varepsilon_0(2x)^2} dx = \frac{e^2}{16\pi\varepsilon_0 x}. \quad (3.4)$$



$$E(x) = \frac{e^2}{16\pi\varepsilon_0 x} + e\mathcal{E}x. \quad (3.5)$$

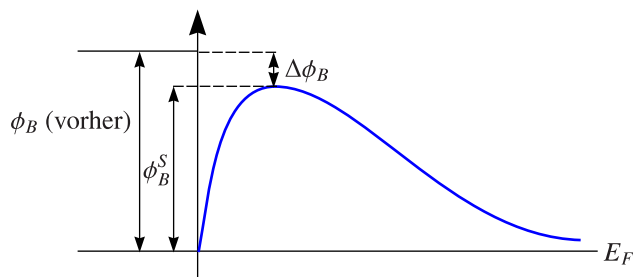
Damit wird die Barriere kleiner. Deren Maximum liegt bei

$$x_m = \frac{\sqrt{e}}{\sqrt{16\pi\varepsilon\varepsilon_0\mathcal{E}}}, \quad (3.6)$$

und es gilt

$$\Delta\phi = 2\mathcal{E}x_m = \sqrt{\frac{e\mathcal{E}}{4\pi\varepsilon\varepsilon_0}}, \quad (3.7)$$

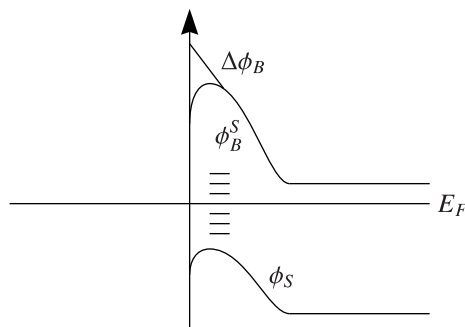
mit der Dielektrizitätskonstante ε des Halbleiters.



Der Effekt ist klein. Beispielsweise gilt für $\epsilon = 12$ und $\mathcal{E} = 10^5$ V/cm: $\Delta\phi_B = 3,5$ meV.

ii.) Fermi-niveau-Pinning

An der Oberfläche des Halbleiters befinden sich viele Defekte, ungesättigte Bindungen (dangling bonds) etc. Lokalisiert an der Oberfläche gibt es somit viele Zustände, beispielsweise in der Bandlücke. Falls die Dichte sehr hoch ist, ist das Fermi-niveau gepinnt in diesen Zuständen ϕ_S (E_F bezüglich E_V).

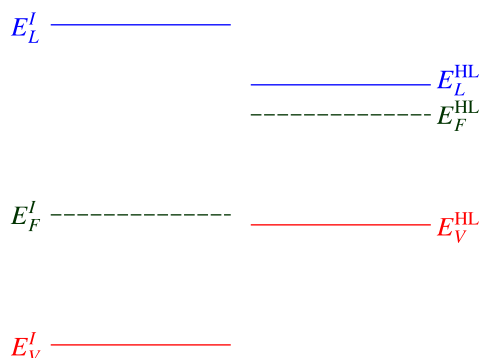


$$\phi_B^S = E_g - \phi_S - \Delta\phi, \quad \phi_S = \frac{1}{3}E_g. \tag{3.8}$$

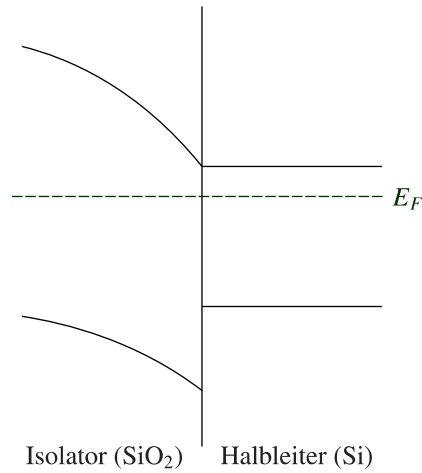
Die letzte Gleichung bezeichnet man als Bardeen-Beziehung.

3.3 Isolator-Halbleiter-Übergang

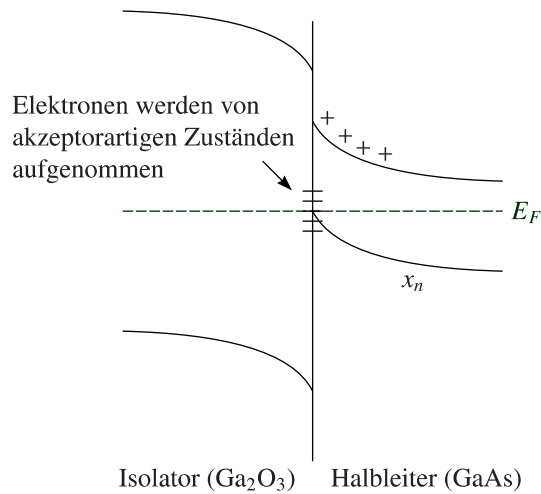
Das Ganze ist wichtig für die Passivierung von Halbleiter-Oberflächen und auch für MOSFETs. Ein Isolator (beispielsweise SiO_2) ist ein Halbleiter mit großer Bandlücke. Die typische Situation ist SiO_2/Si (p-Halbleiter aufgrund von Sauerstofffehlstellen, fehlende Oxidation/n-Halbleiter). Da die Bandlücke sehr groß ist, sind auch die Aktivierungsenergien für Akzeptoren und Donatoren sehr groß. Damit ist $n_A^I \ll n_D^{\text{HL}}$. Die Elektronen fließen also vom Halbleiter in den Isolator. Da es im Isolator nicht zu viele Löcher gibt, die es zu kompensieren gilt, fließen auch nicht viele Elektronen. Es entsteht im Isolator ein Raumladungsgebiet, das sehr viel größer als das im Halbleiter ist. Unter Umständen ist diese viel größer als die Dicke des Isolators.



Die Lösung der Poissongleichung in Schottky-Näherung führt darauf, dass die Diffusionsspannung fast vollständig im Isolator abfällt. Damit gilt $E_F \approx E_F^{\text{HL}}$.



Das elektrische Feld im Halbleiter ist quasi gleich Null. Das Ferminiveau wird nicht gepinnt. Unter Berücksichtigung der Grenzflächenzustände ergibt sich ein Ferminiveau-Pinning. Das ist beispielsweise der Fall für Isolator auf GaAs.



Wir möchten das Potential in Schottky-Näherung berechnen:

$$\varphi(x) = \begin{cases} -e/(2\varepsilon\varepsilon_0)n_D(x - x_n)^2 & \text{für } 0 < x < x_n \\ 0 & \text{für } x_n < x \end{cases}, \quad x_n = \sqrt{\frac{2\varepsilon\varepsilon_0}{en_D}U^S}, \quad (3.9)$$

mit der Grenzflächenspannung $U^S = E_F(x = \infty) - E_F(x = 0)$ vor dem Gleichgewicht. Die Grenzflächenladungsdichte ist

$$n_S = n_D \cdot x_n = \sqrt{\frac{2\varepsilon\varepsilon_0}{e}n_D|U^S|}. \quad (3.10)$$

Wir interessieren uns nun noch für die Feldstärke an der Grenzfläche:

$$|\mathcal{E}_S| = |-\nabla\varphi| = -\frac{e}{\varepsilon\varepsilon_0}n_S. \quad (3.11)$$

Typische Größen sind $n_D = 10^{16} \text{ cm}^{-3}$ und $\varepsilon = 10$. Mit $e|U^S| = 1 \text{ eV}$ ergibt sich $n_S = 3 \cdot 10^{11} \text{ cm}^{-2}$. für das elektrische Feld gilt $\mathcal{E}_S = 6 \cdot 10^4 \text{ V/cm}$. Ist das Oxid minderwertig, bricht es durch bei diesen hohen Feldstärken und es entsteht ein Lochstrom.

Kapitel 4

Halbleiterübergänge im Nichtgleichgewicht / Bauelemente

Das Grundprinzip der neuesten Halbleiter-Bauelemente ist die Störung des Gleichgewichts an einem Halbleiter-Übergang. Störungen sind:

- Spannung (Schalter, Verstärker, Gleichrichter, Speicher)
- Temperatur(gefälle) (Sensoren)
- Einstrahlung von Licht (Sensoren, Solarzelle, LED, LASER)
- Druck (Sensoren)
- Anlegen eines Magnetfeldes (Sensoren, Speicher)

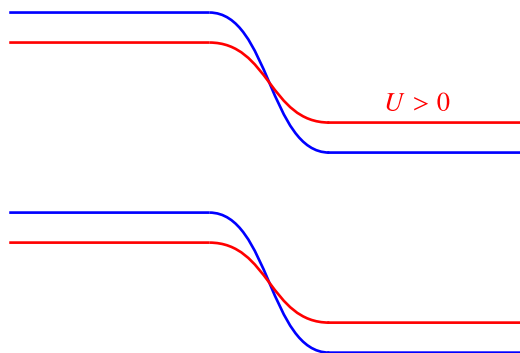
4.1 Der pn-Übergang bei angelegter Spannung

Nun ist das chemische Potential für Elektronen und Löcher getrennt zu betrachten.

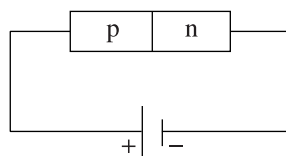
4.1.1 Auswirkungen auf das elektrische Potential

Die äußere Spannung U verhält sich additiv zur Diffusionsspannung U_D :

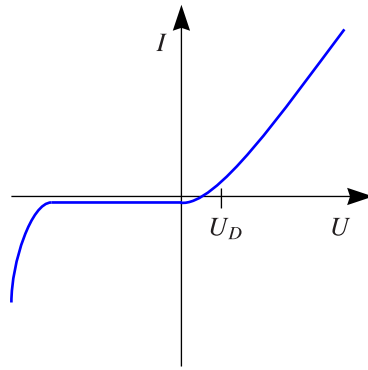
$$\varphi(\infty) - \varphi(-\infty) = U_D - U. \quad (4.1)$$



U ist positiv, falls bei positiven Ladungsträgern die potentielle Energie bei $x = \infty$ relativ zu $x = -\infty$ erniedrigt wird.



Wir berechnen $\varphi(x)$ analog zum pn-Übergang im Gleichgewicht, falls die Schottky-Näherung $\sqrt{U_D - U} \gg \sqrt{k_B T}$ gilt. Diese gilt für $U < 0$ immer und für $U > 0$ nur für $U < U_D$.



Früher hatten wir

$$\varphi(x_n) = \frac{n_A}{n_A + n_D}(U_D - U), \quad \varphi(x_p) = -\frac{n_D}{n_A + n_D}(U_D - U), \quad (4.2)$$

berechnet und daraus ergibt sich die Breite der Raumladungszone:

$$w = \sqrt{\frac{2\varepsilon\varepsilon_0}{e} \frac{n_A + n_D}{n_A n_D} (U_D - U)}. \quad (4.3)$$

Für $U = -100 \text{ V}$ ($U_D = 0,7 \text{ V}$) und $n_D = n_A = 10^{16} \text{ cm}^{-3}$ mit $\varepsilon = 50$ ergibt sich $w \approx 4 \mu\text{m}$.

4.1.2 Mechanismus des Stromtransports

Bei angelegter Spannung gibt es einen Übergang vom Löcherstrom im p-Bahngebiet zu einem Elektronenstrom im n-Bahngebiet. Die Generation und Rekombination ist nötig für den Übergang. Für die Rekombination müssen Nichtgleichgewichtsladungsträger vorhanden sein.

- i.) $U > 0$: Beispielsweise ins p-Gebiet werden Elektronen aus dem n-Gebiet emittiert, da die potentielle Energie angehoben wurde. Wir injizieren also Minoritätsladungsträger. Analoges geschieht in der umgekehrten Richtung. Die injizierte Minoritäten rekombinieren mit vorhandenen Majoritäten. Die freiwerdenden Zustände werden durch Majoritätenstrom aus dem p-Bahngebiet aufgefüllt. Die Größe des Stroms hängt von der Schnelligkeit der Injektion, der Rekombination und dem Serienwiderstand im Bahngebiet ab.
- ii.) $U < 0$: Die einzige Möglichkeit, einen Strom zu erhalten, ist die Extraktion von Minoritätsladungsträgern. Dies funktioniert über die Generation von Elektron-Loch-Paaren, was jedoch (viel) Energie benötigt. Das führt dazu, dass der Strom für $U < 0$ sehr klein ist verglichen mit dem Strom für $U > 0$.

Wir zeigen nun, dass der Gesamtstrom durch Rekombination plus Injektion gegeben ist. Die Gesamtstromdichte ist gegeben durch $j(x) = j_n(x) + j_p(x)$. Die Kontinuitätsgleichung für Teilströme im stationären Zustand lautet:

$$\frac{dj_n}{dx} = eR(x), \quad \frac{dj_p}{dx} = -eR(x). \quad (4.4)$$

Einschub: Kontinuitätsgleichung

Mittels

$$\frac{\partial n}{\partial t} = -R_n + G_n + \frac{1}{e} \nabla \cdot \mathbf{j}_n, \quad (4.5)$$

und der Tatsache, dass keine Generation stattfindet, also $G_n = 0$ ist und dass wir einen stationären Zustand betrachten, $\partial n / \partial t = 0$ folgt die obige Gleichung.

Der Gesamtstrom $\mathbf{j}(x)$ ist räumlich konstant. Betrachte eine bestimmte Position x_1 im p-Gebiet und x_2 im n-Gebiet. Damit muss also

$$j = j_n(x_1) + j_p(x_1) = j_n(x_2) + j_p(x_2), \quad (4.6)$$

sein und mittels Integration erhalten wir:

$$j_n(x_2) = j_n(x_1) + e \int_{x_1}^{x_2} R(x) dx. \quad (4.7)$$

Damit folgt

$$j = j_n(x_1) + j_p(x_2) + e \int_{x_1}^{x_2} R(x) dx, \quad (4.8)$$

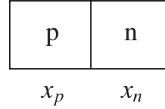
mit den Minoritätsströmen $j_n(x_1)$ und $j_p(x_2)$.

i.) Falls $x_1 = -\infty$ und $x_2 = \infty$ gilt:

$$j = j_n(-\infty) + j_p(\infty) + e \int_{x_1}^{x_2} R(x) dx = e \int_{x_1}^{x_2} R(x) dx, \quad (4.9)$$

wegen $j_n(-\infty) = j_p(\infty) = 0$. Damit ist der Gesamtstrom im Wesentlichen ein Rekombinationsstrom.

ii.) Falls $x_1 \geq x_p$ und $x_2 \leq x_n$:



$$j = j_n(x_p) + j_p(x_n) + e \int_{x_1}^{x_2} R(x) dx = j_n(x_p) + j_p(x_n), \quad (4.10)$$

der Gesamtstrom ist im Wesentlichen ein Injektionsstrom.

4.1.3 Verlauf des chemischen Potentials

Wir betrachten speziell das chemische Potential der injizierten N Gleichgewichtsminoritäten und beschreiben das Ganze durch Annahme eines separaten lokalen Gleichgewichts für Elektronen und Löcher. (Also seien Elektronen und Löcher untereinander nicht im Gleichgewicht.) Es gilt $\zeta^n(x) \neq \zeta^p(x)$ und:

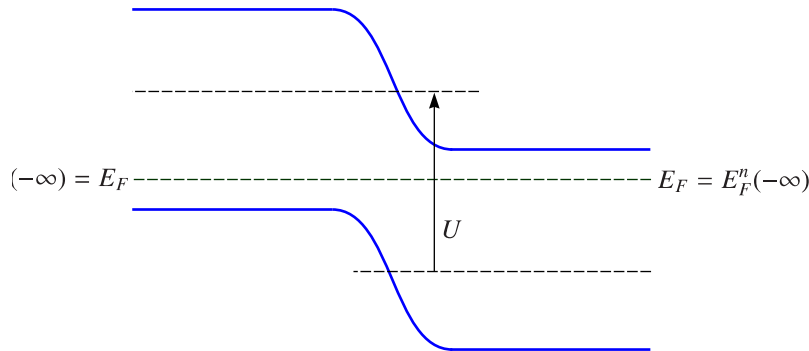
$$n(x) = N_C \exp\left(\frac{\zeta^n(x) - E_g}{k_B T}\right), \quad p(x) = N_V \exp\left(-\frac{\zeta^p(x)}{k_B T}\right). \quad (4.11)$$

Definition über das intrinsische elektrochemische Potential E_F^i :

$$n(x) = n_i(x) \exp\left(\frac{E_F^n(x) - E_F^i(x)}{k_B T}\right), \quad p(x) = n_i(x) \exp\left(\frac{E_F^i(x) - E_F^p(x)}{k_B T}\right). \quad (4.12)$$

Im Gleichgewicht gilt $E_F^p = E_F^n = E_F$ und somit $n \cdot p = n_i^2$. Im Nichtgleichgewicht ist jedoch

$$n \cdot p = n_i^2(x) \exp\left(-\frac{E_F^p(x) - E_F^n(x)}{k_B T}\right) = n_i(x)^2 \exp\left(-\frac{eU}{k_B T}\right). \quad (4.13)$$



Für $U > 0$ gilt also $n \cdot p < n_i^2$ und für $U < 0$ ist $n \cdot p > n_i^2$. Für die Ladungsträgerdichten im Bahngebiet gilt:

i.) Majoritäten: Die Injektion ändert die Dichte der Majoritäten wenig im Vergleich zur Dichte der Minoritäten. Es ist $p(x) = p_{p0}$ für $x < x_p$ und $n(x) = n_{n0}$ für $x > x_n$.

ii.) Minoritäten: Im stationären Fall folgt aus der Kontinuitätsgleichung:

$$0 = \frac{\partial n}{\partial t} = \frac{1}{e} \nabla \cdot \mathbf{j}_n - R_n. \quad (4.14)$$

(Aus S.Sze: Physics of Semiconductor Devices) Und entsprechend:

$$0 = \frac{\partial p}{\partial t} = -\frac{1}{e} \nabla \cdot \mathbf{j}_p - R_p. \quad (4.15)$$

Der Strom setzt sich zusammen aus einem Strom, der durch ein äußeres elektrisches Feld \mathcal{E} kommt, und einem Diffusionsstrom der aus dem Konzentrationsunterschied der Ladungsträger herrührt:

$$\mathbf{j} = en\mu_n\mathcal{E} + eD_n\nabla n. \quad (4.16)$$

Da die Ladungsträgerkonzentration n und das elektrische Feld \mathcal{E} vom Ort abhängen, folgt weiter:

$$\nabla \cdot \mathbf{j} = e\mu_n\mathcal{E} \cdot (\nabla n) + en\mu_n\nabla \cdot \mathcal{E} + eD_n\nabla \cdot (\nabla n). \quad (4.17)$$

Der erste Term auf der rechten Seite ist sehr klein im Bahngebiet, weil dort die Ladungsträgerkonzentration im Wesentlichen konstant ist. Ist ebenso der zweite Term klein (kleine Spannung). Das Ganze führt uns zu den sogenannten Diffusions/Rekombinationsgleichungen:

$$D_n \frac{\partial^2 n_p}{\partial x^2} - \frac{n_p - n_{p0}}{\tau_n} = 0, \quad D_p \frac{\partial^2 p_n}{\partial x^2} - \frac{p_n - p_{n0}}{\tau_p} = 0. \quad (4.18)$$

Randbedingungen:

- 1.) $n_p(-\infty) = n_{p0}$, $p_n(\infty) = p_{n0}$ (Gleichgewichtsdichten)
- 2.) Rand der Verarmungszone: Die gesamte Spannung fällt an der Verarmungszone ab, sofern wir Serienwiderstände ignorieren:

$$\varphi(x_n) - \varphi(x_p) = e(U_D - U), \quad (4.19)$$

wobei zusätzlich die Diffusionsspannung U_D berücksichtigt wird, die sich durch Ausgleich der Majoritätsladungsdichten bereits aufgebaut hat. Weiterhin gilt:

$$n(x) = n_{n0} \exp\left(\frac{e(\varphi(x) - \varphi(\infty))}{k_B T}\right), \quad p(x) = p_{p0} \exp\left(-\frac{e(\varphi(x) - \varphi(-\infty))}{k_B T}\right). \quad (4.20)$$

Mit $\varphi(\infty) = \varphi(x_n)$ und $\varphi(-\infty) = \varphi(x_p)$ ergibt sich weiter

$$n(x_p) = n_{n0} \exp\left(-\frac{e(U_D - U)}{k_B T}\right) = n_{p0} \exp\left(\frac{eU}{k_B T}\right), \quad p(x_n) = p_{p0} \exp\left(-\frac{e(U_D - U)}{k_B T}\right) = p_{n0} \exp\left(\frac{eU}{k_B T}\right). \quad (4.21)$$

Hier haben wir die Näherung $n(x_p) = n(-\infty)|_{U=0}$ verwendet und damit

$$n_{n0} \exp\left(-\frac{eU_D}{k_B T}\right) = n_{p0}. \quad (4.22)$$

Für $U = 0$ gilt also

$$p(x_n) = p_{n0}, \quad n(x_p) = n_{p0}, \quad (4.23)$$

im Gleichgewicht. Für $U > 0$ sind die Minoritäten deutlich höher als im Gleichgewicht wegen der Injektion. Beispiel: Mit $U = 0,25$ V und $T = 300$ K gilt

$$\exp\left(\frac{eU}{k_B T}\right) \approx \exp(10) \approx 2 \cdot 10^4. \quad (4.24)$$

Lösung der Diffusions/Rekombinationsgleichungen führt auf die Minoritätsdichten im Bahngebiet

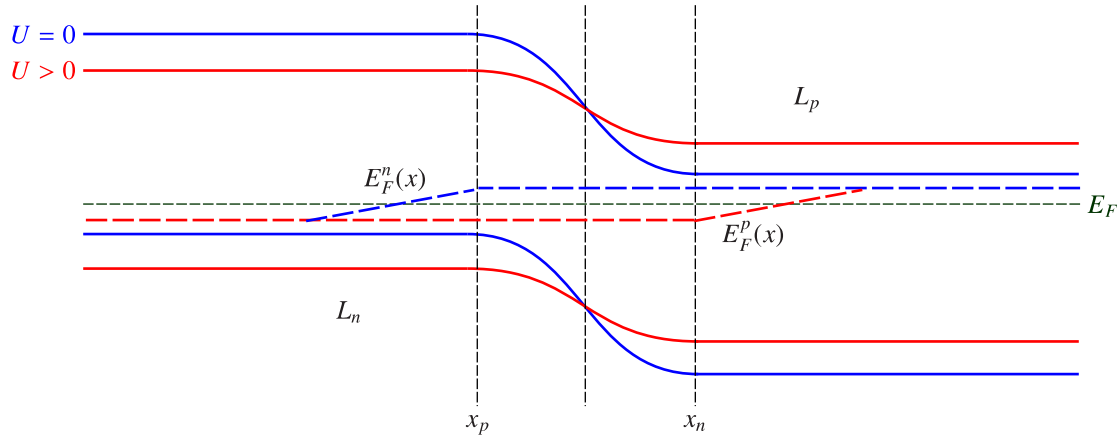
$$n(x) = n_{p0} + [n(x_p) - n_{p0}] \exp\left(\frac{x - x_p}{L_n}\right), \quad x < x_p, \quad (4.25a)$$

$$p(x) = p_{n0} + [p(x_n) - p_{n0}] \exp\left(-\frac{x - x_n}{L_p}\right), \quad x > x_n, \quad (4.25b)$$

wobei L_n und L_p die Diffusionslängen sind.

4.1.4 Verlauf des Quasi-Fermi-Niveaus

Für $U > 0$ sind die Minoritätsdichten auf einer Längenskala L_p , L_n höher als im Gleichgewicht. Dann liegen E_F^p und E_F^n näher am jeweiligen Band. $E_F^{p,n}$ verlaufen linear. Für $U < 0$ liegt der umgekehrte Fall vor.



4.1.5 Diodenkennlinie

Die Berechnung der Kennlinie $I(U)$ ist über zwei Wege möglich:

- i.) über die Formel für die Rekombinationsstromdichte
- ii.) über die Formel für die Injektionsstromdichte.

Die Stromdichte ist gegeben durch

$$j = j_n(x_p) + j_p(x_n) + e \int_{x_p}^{x_n} R(x) dx, \quad (4.26)$$

wobei der Rekombinationsstrom (dritte Term) in der Raumladungszone vernachlässigbar klein ist. $j_n(x_p)$ und $j_p(x_n)$ folgt aus $j = \sigma E \pm eD_{n,p}d(n,p)/dx$. Da φ im Bahngebiet konstant ist, gilt hier auch $\mathcal{E} = 0$, also

$$j_n(x_p) = eD_n \left. \frac{dn}{dx} \right|_{x=x_p} = e \frac{D_n}{L_n} [n(x_p) - n_{p0}], \quad (4.27)$$

und analog

$$j_p(x_n) = -eD_p \left. \frac{dp}{dx} \right|_{x=x_n} = e \frac{D_p}{L_p} [p(x_n) - p_{n0}], \quad (4.28)$$

mit

$$n(x_p) = n_{p0} \exp\left(\frac{eU}{k_B T}\right), \quad p(x_n) = p_{n0} \exp\left(\frac{eU}{k_B T}\right). \quad (4.29)$$

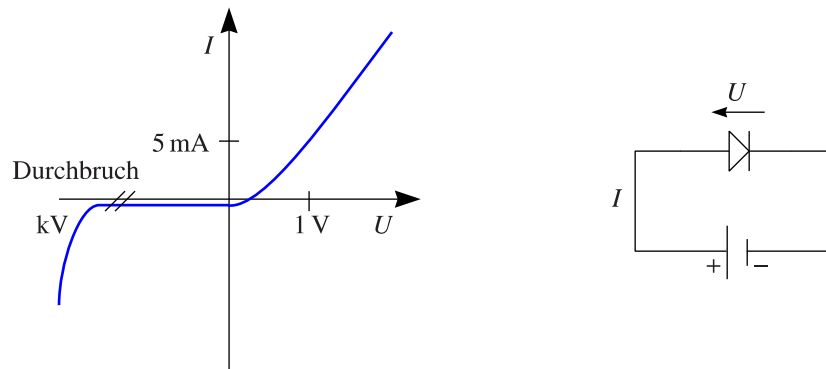
Damit ergibt sich die Schockley-Gleichung

$$j = j_s \left\{ \exp\left(\frac{eU}{k_B T}\right) - 1 \right\}, \quad j_s = e \left(n_{p0} \frac{D_n}{L_n} + p_{n0} \frac{D_p}{L_p} \right), \quad (4.30)$$

als ideale Diodenkennlinie. j_s nennt man auch Sättigungsstromdichte. Wir berechnen j_s in Silizium mit den Zahlenwerten $L_{n,p} = 10 \mu\text{m}$, $n_A = n_D = 10^{16} \text{cm}^{-3}$. Mit $n_i = 10^{10} \text{cm}^{-3}$ ergibt sich

$$n_{p0} = p_{n0} = \frac{n_i^2}{n_A} = \frac{n_i^2}{n_D} = 10^4 \text{cm}^{-3}, \quad (4.31)$$

und somit $j_s = 10^{-10} \text{A/cm}^2 \approx 0$. Der pn-Übergang sperrt also für $U \leq 0$.

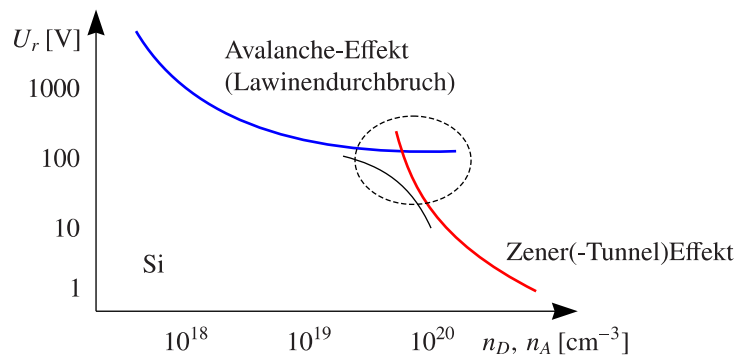


Von der Idealform der Kennlinie gibt es Abweichungen durch folgende Effekte:

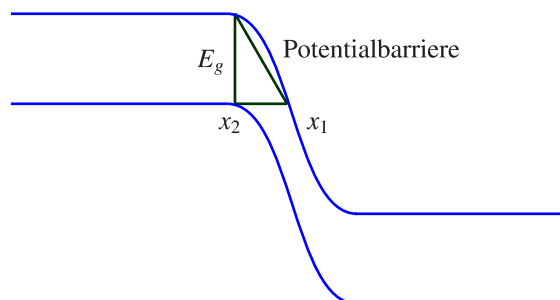
- i.) Oberflächenleckströme (in Sperrichtung)
- ii.) Generation von Elektron-Loch-Paaren in Verarmungszonen (Sperrichtung), Rekombination von Elektron-Loch-Paaren in Verarmungszone (Durchlassrichtung)
- iii.) Tunnelprozesse zwischen Zuständen in der Bandlücke
- iv.) hohe Injektion: zusätzlicher Beitrag
- v.) Widerstände im Bahngebiet sowie am Kontakt
- vi.) Durchbruchverhalten

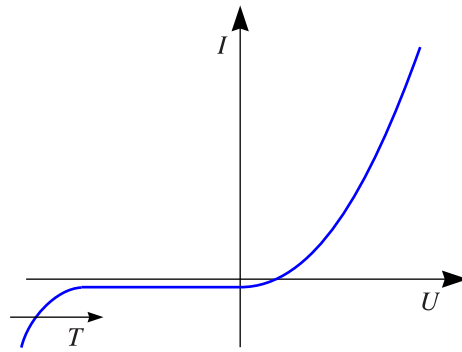
4.1.6 Durchbruchverhalten

Durchbruch bedeutet ein schneller Anstieg des Stromes bei Überschreiten eines Schwellwertes $U_{R/B}$ in der Sperrspannung. Dies kann die Diode durch lokales Schmelzen zerstören. Der Durchbruch kann jedoch durch einen äußeren Widerstand vermieden werden. Im Wesentlichen gibt es in diesem Zusammenhang zwei Effekte, die auftreten:



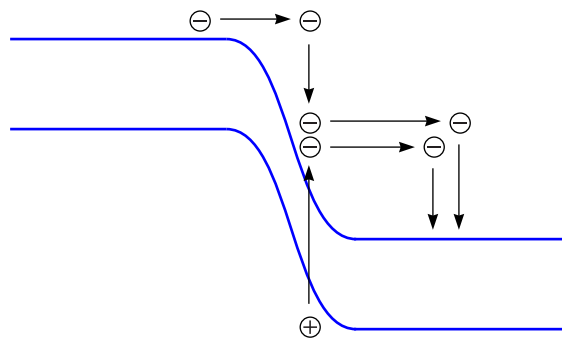
- i.) Zener-Effekt: Dabei kann ein Elektron vom Valenzband ins Leitungsband tunneln (dreieckige Tunnelbarriere mit Breite $W_T = x_2 - x_1$ und Höhe: E_g).



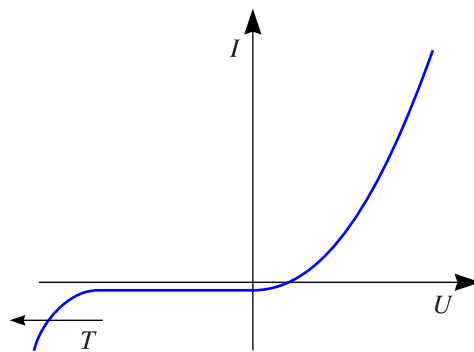


Die Bandlücke E_g wird mit steigender Temperatur T kleiner. Damit findet schon bei kleinerer Spannung ein Durchbruch statt.

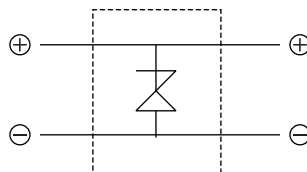
- ii.) Avalanche-Effekt: Hier reicht die kinetische Energie eines Elektrons zur Erzeugung eines Elektron-Loch-Paares aus (Stoßionisation).



Zuerst findet eine Multiplikation von freien Ladungsträgern (Elektronen und Löchern) statt und dann ein unkontrollierter Durchbruch: $U_r \sim 1/n$. Die kritische Durchbruchfeldstärke ist beispielsweise in Silizium $3 \cdot 10^7$ V/m. U_r steigt mit der Temperatur, weil dann mehr Phononen angeregt werden, mit denen das Elektron wechselwirkt. Damit wird die freie Weglänge des Elektrons kleiner und dieses gibt bei einem Stoß mit einem Phonon Energie ab.



Bei geeigneter Wahl der Dotierung tragen Zener-Effekt und Avalanche-Effekt gleichermaßen zur U_r bei und damit ist U_r temperaturunabhängig. Die Anwendung ist eine Zener-Diode mit einer extrem stabilen Spannungsbegrenzung (Spannungsreferenz).



4.1.7 Wechselstromverhalten der Diode

Eine große Wechselspannung (bei kleinen Frequenzen): Gleichrichtung; hier: kleine Wechselspannung

- i.) Sei $W = W(U)$ die Dicke der Raumladungszone. Die Ladungen (Majoritäten) müssen zu oder abtransportiert werden: Kondensatorwirkung (Schottky-Kapazität). Für die Ladung Q , die man zu- oder abführen muss, gilt $Q \sim W(U) \sim \sqrt{U_D - U}$. Weiterhin gilt dann $\partial Q / \partial U \sim 1 / \sqrt{U_D - U}$. Die Kapazität kann dann berechnet werden zu

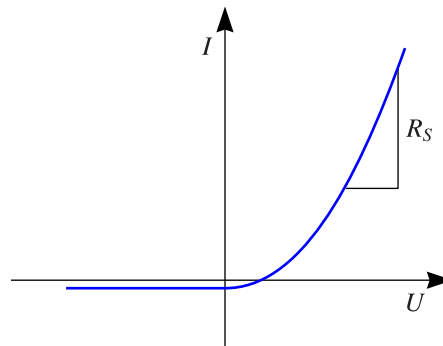
$$C_R = A \sqrt{\frac{\varepsilon \varepsilon_0 n_A n_D}{2e(n_A + n_D)}} \frac{1}{\sqrt{U_D - U}}, \quad (4.32)$$

für $W < 1/t$, wobei t die Laufzeit in der Sperrschicht ist. Das Ganze ist also maßgeblich für kleine Frequenzen in der Sperrichtung.

- ii.) Die Minoritäten im Diffusionsgebiet müssen durch Rekombination abgebaut bzw. injiziert werden, wenn $U > U_D$. Die endliche Rekombinationszeit führt zu einem kapazitiven Verhalten, das man als Schockley-Kapazität bezeichnet. Die Anzahl der injizierten Ladungsträger $n(x_p)$ und $p(x_n)$ ist abhängig von U : $\sim \exp(eU / (k_B T))$. Damit gilt

$$C_D \sim \exp\left(\frac{eU}{k_B T}\right) [L_n n_p + L_p p_n]. \quad (4.33)$$

C_D ist maßgeblich für kleine Frequenzen in Flußrichtung.



Die $I(U)$ -Kurve kann nun geschrieben werden als

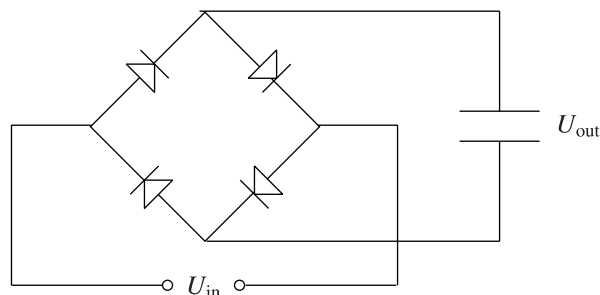
$$I = I_0 \left\{ \exp\left(\frac{eU - R_S I}{nk_B T}\right) - 1 \right\}, \quad (4.34)$$

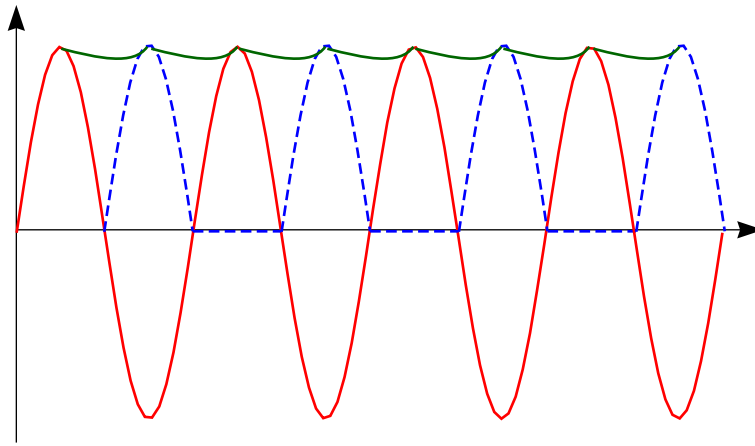
wobei n der Nichtidealitätsfaktor (non-ideality factor) ist und Abweichungen der Diode vom idealen Verhalten bei kleinen Spannungen beschreibt. Für sehr gute Silizium-Dioden gilt $n \geq 1,05$. Der Summand $-R_S I$ führt zur Strombegrenzung; gefittet wird dann die Kurve $U(I)$.

4.2 Anwendungen der Diode

Eine Diode ist ein „two-terminal device“.

- a.) Gleichrichter: Nach der Diodenkennlinie besitzt diese einen kleinen Widerstand in Durchlassrichtung und einen hohen Widerstand in Sperrichtung. Mit einer Diode lässt sich somit eine Wechselspannung gleichrichten.



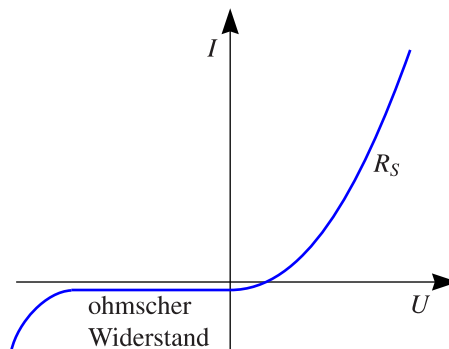


Die Frequenz hängt von der Lebensdauer der injizierten Minoritäten ab. Typische Schaltzeiten liegen bei 50 bis 500 ns. Schnellere Schaltzeiten erreicht man durch

- i.) Verwenden eines direkten Halbleiters oder
- ii.) durch Einbringen von Rekombinationszentren.

b.) Varistor (variable resistor): Dabei handelt es sich um einen nicht-ohmschen Widerstand. Man betreibt die Diode in Durchlassrichtung:

$$R_F = \frac{U_F}{I_F} = \frac{U_F}{I_S} \left\{ \exp\left(\frac{eU_F}{k_B T}\right) - 1 \right\}^{-1} . \quad (4.35)$$



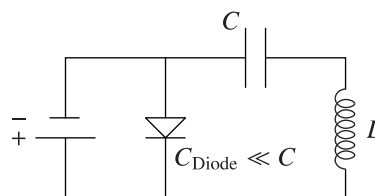
Das Ganze wendet man an in der symmetrischen Teilspannungsbegrenzung.

c.) Varactor (variable reactor): Kapazitätsdiode

Hier wird die Diode in der Sperrichtung betrieben. Kleinsignal-Kapazität der Diode:

$$C(U) = \frac{dQ}{dU} = \frac{\varepsilon \varepsilon_0}{W(U)} . \quad (4.36)$$

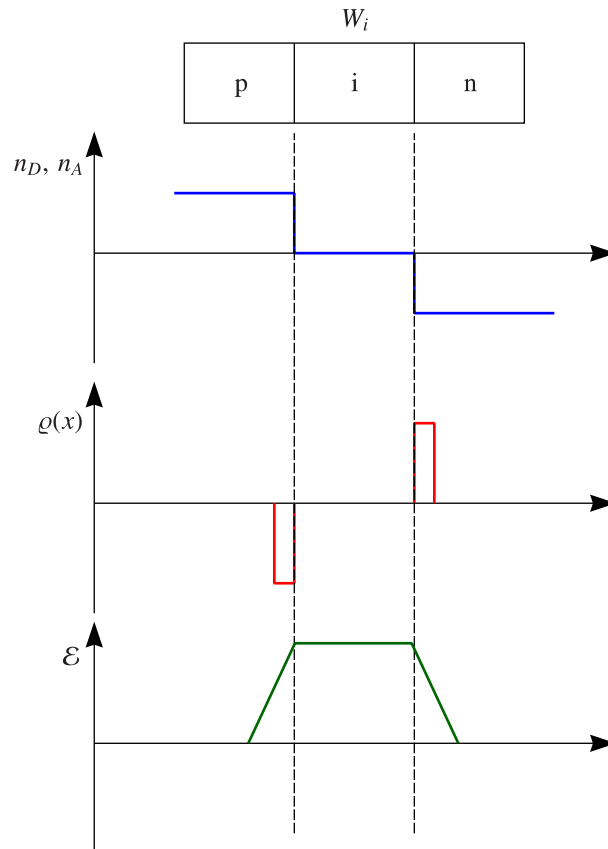
Feinabstimmung eines Schwingkreises:



Die maximale Frequenz ist gegeben durch $\tau = C(U)R_{\text{Bahngebiet}}$, wobei $C(U)$ die Schottky-Kapazität ist. (Also spielen die Majoritäten die entscheidende Rolle.)

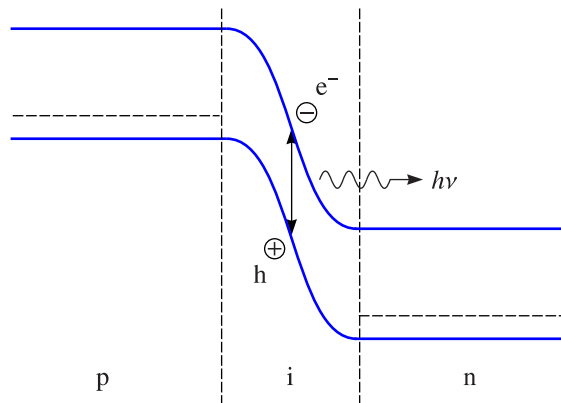
d.) p-i-n-Diode:

Es wird eine intrinsische Schicht zwischen n- und p-Gebiet eingeschoben.



Eine hohe Spannung fällt im intrinsischen Gebiet ab. Das führt zu einer hohen Grenzfrequenz und zwar $\tau = W_i/(2v_s)$, wobei v_s die Sättigungsgeschwindigkeit im i-Gebiet ist. Das Ganze wird im Mikrowellenbereich (Consumer-Bereich) angewendet. Man kann außerdem Schalter und Modulatoren bauen (IMPATT-Diode).

- e.) Photodiode: Die p-i-n-Diode wird als schneller Photodetektor angewendet. Das Anlegen einer kleinen Sperrspannung führt zu einem großen Verarmungsgebiet (typischerweise der Breite von einigen $10 \mu\text{m}$).



Das Feld trennt Ladungsträger und es fließt damit ein äußerer Strom $I = I_S + I_p$. Beispielsweise folgt für $U = 10 \text{ V}$ und $W = 50 \mu\text{m}$ ein elektrisches Feld $\mathcal{E} = 2 \cdot 10^5 \text{ V/m}$. Die Transitzeit von Ladungsträgern im i-Gebiet ist somit $2,5 \text{ ns}$. Die Anforderungen an das Bauelement sind derart, dass es zum einen schnell ist. Also benötigen wir ein dünnes Verarmungsgebiet. Andererseits möchten wir auch eine hohe Quanteneffizienz, um viele Photonen zu absorbieren, was wiederum ein breites Verarmungsgebiet erfordert. Wir definieren die Quanteneffizienz durch

$$\eta = \frac{I_{\text{Photo}/e}}{I_{\text{Licht}/h\nu}} \quad (4.37)$$

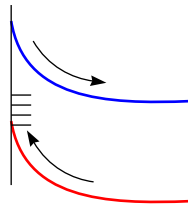
Quanteneffizienz:

- Cut-Off bei großer Wellenlänge λ : Absorptionskante

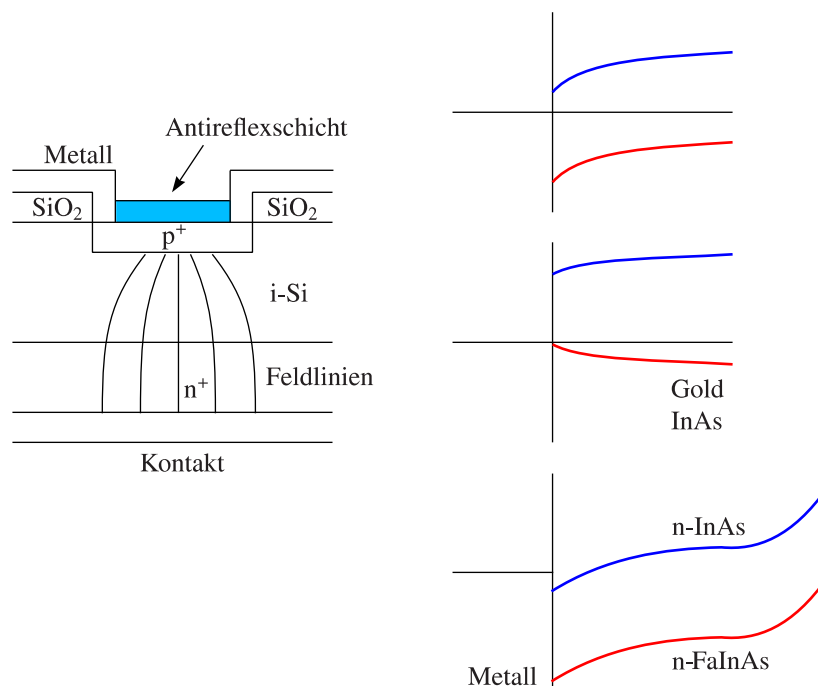
- Cut-Off bei kleiner Wellenlänge λ : Elektron-Loch-Paare

nahe der Oberfläche generiert + Oberflächenrekombination Die Lebensdauer ist besonders da klein, wo viele Zustände sind und die befinden sich an der Oberfläche.

Photodioden sind so gebaut, dass sie bei einer bestimmten Wellenlänge sehr hohe Quanteneffizienzen haben.



Der Vorteil der p-i-n-Diode gegenüber der Solarzelle ist deren kleine Fläche und die effektive Detektion einer Wellenlänge. Bauteilausführung:



Die Antireflexschicht dient dazu, dass möglich sehr viel Licht eingegangen wird (beispielsweise durch Mehrfachreflexion). Mittels der Isolationsschicht trennt man die Metallkontakte von den anderen Bereichen. Das Feld wird generiert durch einen Schottky-Kontakt (Metall + Halbleiter). Also benötigt man einen Halbleiter, der eine Schottky-Barriere hat. (Bei Gold und InAs gibt es keine große Verarmungszone, weshalb die Elektron-Loch-Paare sich nicht bewegen.)

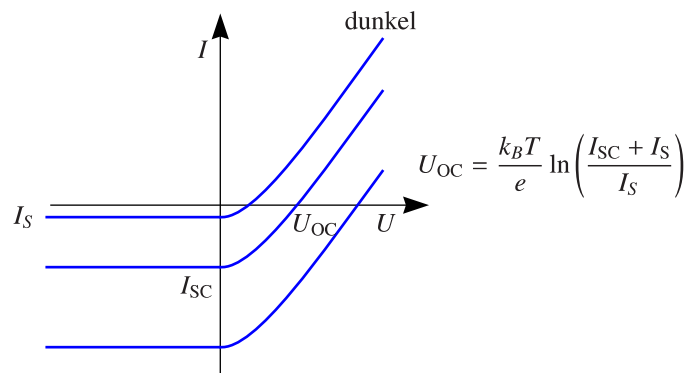
Die Avalanche-Photodiode wird mit hoher Sperrspannung kurz vor dem Durchbruch betrieben. Die Verstärkung durch Avalanche-Multiplikation liefert uns bei Lichtpulsen ein sehr gutes Signal-Rausch-Verhältnis (S/N-Verhältnis). Solche Detektoren verwendet man vor allem bei Kernforschungsexperimenten.

f.) Solarzelle:

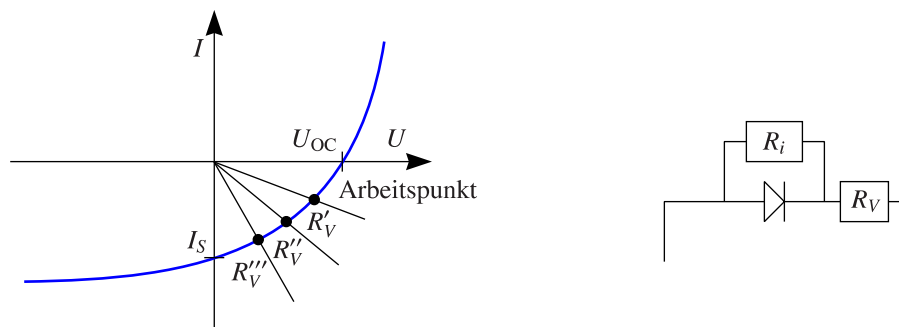
Die Solarzelle wird so gebaut, dass ein möglichst großer intrinsischer Bereich vorliegt. Im pn-Übergang ohne Vorspannung werden optisch generierte Elektron-Loch-Paare getrennt. Dadurch entsteht ein Kurzschlussstrom I_{SC} , welcher proportional zur Intensität des Lichtes ist. Die Kennlinie setzt sich aus dem Dunkelstrom und dem Kurzschlußstrom zusammen:

$$I = I_S \left\{ \exp \left(\frac{eU}{k_B T} \right) - 1 \right\} - I_{SC} . \tag{4.38}$$

Es ergibt sich bei offenen Klemmen eine Leerlaufspannung U_{OC} (open circuit). Die ist größer als Null, weil die Elektron-Loch-Paare die Raumladungszone verringern.



Die Solarzelle soll als Stromquelle dienen; ein angeschlossener Lastwiderstand verbraucht Leistung. Man versucht, den Arbeitspunkt so zu wählen, dass das Rechteck maximiert wird. Die Fläche des Rechtecks gibt nämlich die maximale Leistung $P_{\text{mpp}} = I_{\text{mp}} \cdot U_{\text{mp}}$ an.



Man definiert nun den sogenannten Füllfaktor durch

$$\text{FF} = \frac{I_{\text{mp}} \cdot U_{\text{mp}} = P_{\text{mpp}}}{I_{\text{SC}} \cdot U_{\text{OC}}} . \quad (4.39)$$

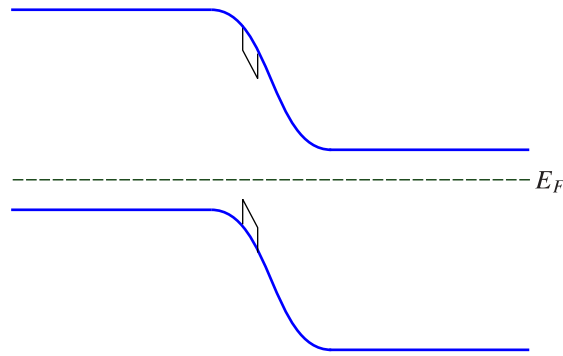
Heutzutage erreicht man Füllfaktoren von 0,75 bis 0,85 für kristallines Silizium und 0,7 für amorphes Silizium. Amorphes Silizium ist zwar billiger als kristallines; das letztere besitzt jedoch aufgrund weniger Defekte einen besseren Füllfaktor. Möchte man verschiedene Wellenlängen absorbieren, verwendet man Halbleiter-Heterostrukturen. Dann sinkt der Füllfaktor auf etwa 0,6, was jedoch durch die Tatsache ausgeglichen wird, dass man Licht verschiedener Wellenlängen absorbieren kann. Kommen wir schließlich noch zum Wirkungsgrad:

$$\eta = \frac{P_{\text{mpp}}}{A \cdot P_{\text{opt}}} , \quad (4.40)$$

wobei A die Fläche und P_{opt} die Bestrahlungsstärke ist. Für η sind Werte zwischen 30% und 85% möglich. Anforderungen an Solarzellen sind:

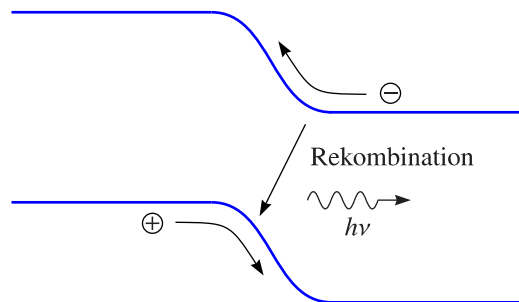
- maximale Leistungsabgabe
- großflächig
- Schnelligkeit spielt keine Rolle
- möglichst effizient für das ganze Sonnenspektrum
- Dicke der Verarmungszone groß

Für die Effizienz benötigt man idealerweise eine Bandlücke von 1,2 eV bis 1,8 eV (Si oder GaAs). Man benötigt eine dicke Verarmungszone, so dass möglichst viel Licht absorbiert wird. Andererseits sollte sie jedoch auch klein genug sein, um nichtstrahlende Rekombination zu verhindern. Man verwendet deshalb eine nicht allzu dicke Verarmungszone und verlängert den Lichtweg durch Mehrfachreflexion. Eine andere Möglichkeit stellen plasmonische Strukturen dar. Dazu bringt man Gold-Nanopartikel mit einem Durchmesser von etwa 100 nm auf die Oberfläche. In diesen Goldpartikeln können Plasmonen angeregt werden, welche zerfallen und Licht emittieren. Das Licht wird damit auf einen sehr kleinen Bereich fokussiert.

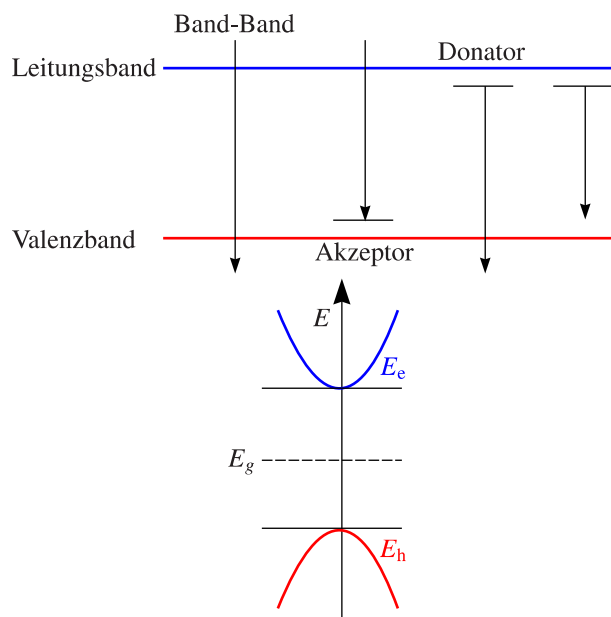


g.) Leuchtdioden:

Man benutzt einen pn-Übergang, welcher in Vorwärtsrichtung gepolt ist. In der Verarmungszone findet strahlende Rekombination statt:



Dabei treten folgende Rekombinationsprozesse, auf, wobei Band-Band-Übergänge bei Zimmertemperatur besonders wichtig sind:



Lumineszenzintensität (spontane Emission beim Band-Band-Übergang):

$$I(\hbar\omega) \sim D_e(E_e)f_e(E_e) \cdot D_h(E_h)f_h(E_h) \cdot |\mathcal{M}|^2, \quad \hbar\omega = E_g + E_e + E_h, \quad (4.41)$$

wobei $D(E)$ die Zustandsdichte, $f(E)$ die Fermi-Dirac-Verteilungsfunktion und \mathcal{M} das Übergangsmatrixelement ist. Das Matrixelement \mathcal{M} gibt an, ob ein entsprechender Übergang überhaupt erst erlaubt ist (anhand von Gruppentheorie und Symmetrieüberlegungen), unabhängig von der Bandlücke. Es kann beispielsweise der Fall sein, dass die Elektronen bestimmter Bänder nur mit Licht wechselwirken, sofern dieses senkrecht oder parallel zu bestimmten Vektoren polarisiert ist. (Bei Kristallen mit der Wurtzit-Struktur gibt es Probleme mit polarisiertem Licht.)

Materialien, die man zum Bau von LEDs verwendet, sind:

- $\text{In}_x\text{Ga}_{1-x}\text{As}_y\text{P}_{1-y}/\text{InP}$: 1100 - 2100 nm; 1,3 μm , 1,55 μm (Glasfaser)
- $\text{In}_x\text{Ga}_{1-x}\text{As}/\text{GaAs}$: 880 - 1100 nm
- $\text{Al}_x\text{Ga}_{1-x}\text{AsGaAs}$: 780 - 890 nm
- $(\text{Al}_x\text{Ga}_{1-x})_{0,5}\text{In}_{0,5}\text{P}/\text{GaAs}$: 610 - 690 nm (rot/gelb)
- $\text{In}_x\text{Ga}_{1-x}\text{N}/\text{GaN}$: 520 nm (grün), 460 nm (blau)

g.) Diodenlaser: LED + Resonator

- DVD, optische Speicher
- Drucker, Kopierer
- Datenübertragung, Glasfaser
- Laserpointer

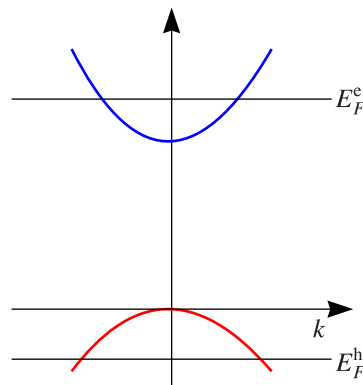
Damit der Laser funktioniert, muss die sogenannte Laserbedingung erfüllt sein. Die Gesamtemissionsrate ist gegeben durch

$$|\mathcal{M}|^2 D(\hbar\omega) [f_e f_h (1 + N) - N(1 - f_h)(1 - f_e)], \quad (4.42)$$

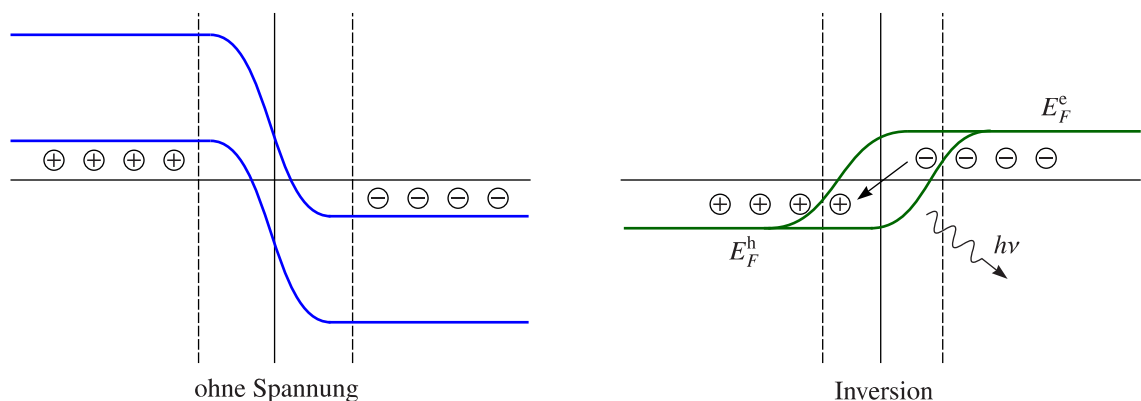
wobei der erste Summand die Emission und der zweite die Absorption beschreibt. Innerhalb des ersten Summanden steht $f_e f_h$ für die spontane Emission und $f_e f_h N$ für die stimulierte, wobei N die Dichte der Photonen ist. Umformen ergibt:

$$|\mathcal{M}|^2 D(\hbar\omega) [f_e f_h + N(f_e + f_h - 1)]. \quad (4.43)$$

Für $f_e + f_h > 1$ liegt Besetzungsinversion und damit optische Verstärkung (gain) vor.



Es muss entartete Besetzung vorliegen, also $\mu = E_F^e + E_F^h > E_g$. Wir bauen eine p-i-n-Diode, die in beiden Bereichen sehr hoch dotiert ist. Dann liegt die Fermienergie sehr hoch. Die Ladungsträger können jedoch noch nicht rekombinieren, da sie räumlich getrennt sind. Deshalb legen wir eine äußere Spannung $U = U_D > 0$ an, so dass die Bänder abgeflacht werden und sich infolgedessen die Fermienergie verbiegen. Dann liegen in der früheren Verarmungszone sowohl Elektronen als auch Löcher vor, die leicht rekombinieren können.

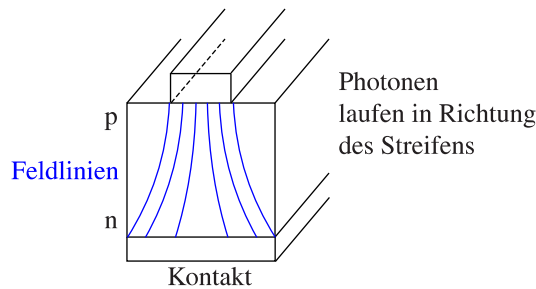


Aufgrund der hohen Dotierung gibt es auch eine große Anzahl von Störstellen, an denen nichtstrahlende Rekombination stattfinden kann. Deshalb ist die hohe Dotierung nicht die beste Lösung. Um stimulierte Emission zu erzeugen, wird außerdem ein Resonator benötigt. Einen Resonator kann man bauen, indem man

- den Kristall entlang der Hauptspaltebenen zerbricht oder
- man die Spaltebenen auch noch verspiegelt.

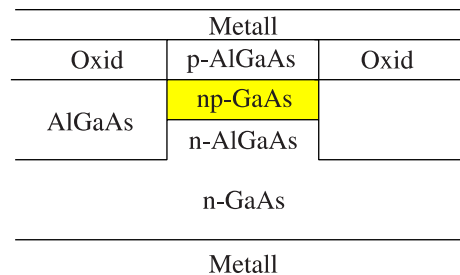
Eine zusätzliche Wellenlängenselektion kann man durch ein Gitter im externen Resonator durchführen. Verbesserung der Lasereffizienz erreicht man durch

- Bau eines Doppelheterostrukturlasers (Einschluss der Ladungsträger und der Lichtwelle)
- Gain Guiding:



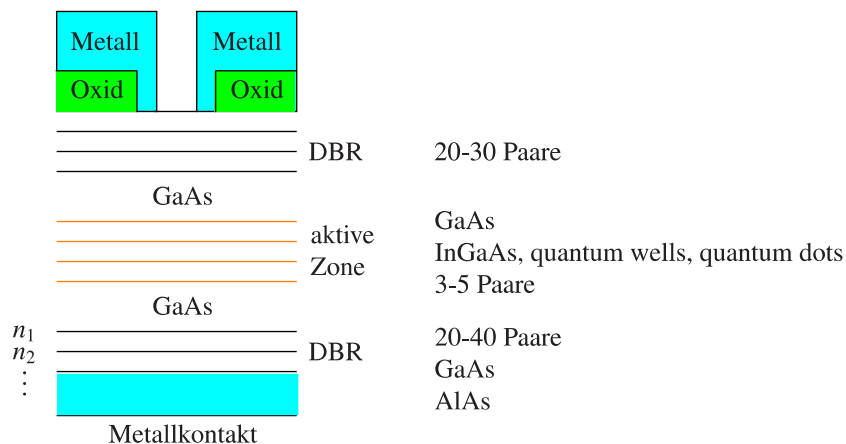
Beim Gain Guiding werden die zentralen Bereiche des Laserprofils verstärkt und die äußeren Bereiche abgeschwächt. Damit beschränkt man die Dicke des Laserstrahls und verhindert, dass dieser auffächert (ähnlich dem Einschluss der Lichtwelle).

- Index Guiding:

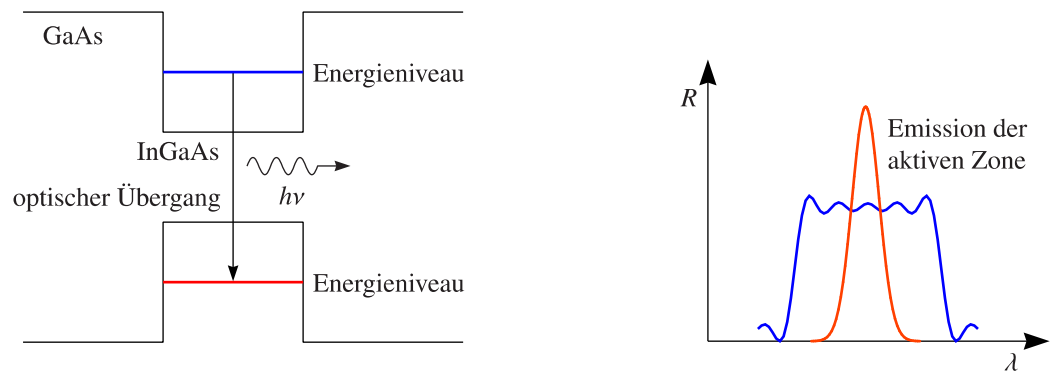


Der aktive Bereich ist mit Material hoher Bandlücke umgeben (kleiner Brechungsindex). In dieses Material kann der Laserstrahl somit nicht eindringen, womit die Struktur als Wellenleiter dienen kann. Auch so verhindert man eine Ausweitung des Laserstrahls.

- Vertical Cavity Surface Emitting Laser (VCSEL):



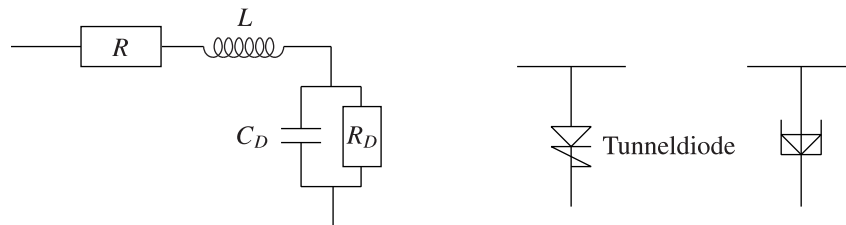
Die aktive Zone des VCSEL besteht aus drei bis fünf Paaren von GaAs- und InGaAs-Schichten, in die Quantentöpfe eingebettet sind. In diesen Quantentöpfen wird das Laserlicht erzeugt. Die aktive Zone ist auf beiden Seiten von sogenannten DBR-Spiegeln (Distributed Bragg-Reflector) umgeben. Diese bestehen aus Schichten mit abwechselnd niedriger und hoher Brechungszahl n_1 und n_2 . Die Dicke der Schichten liegt bei einem Viertel der Wellenlänge des Laserlichts, womit die DBR-Spiegel als Resonator funktionieren.



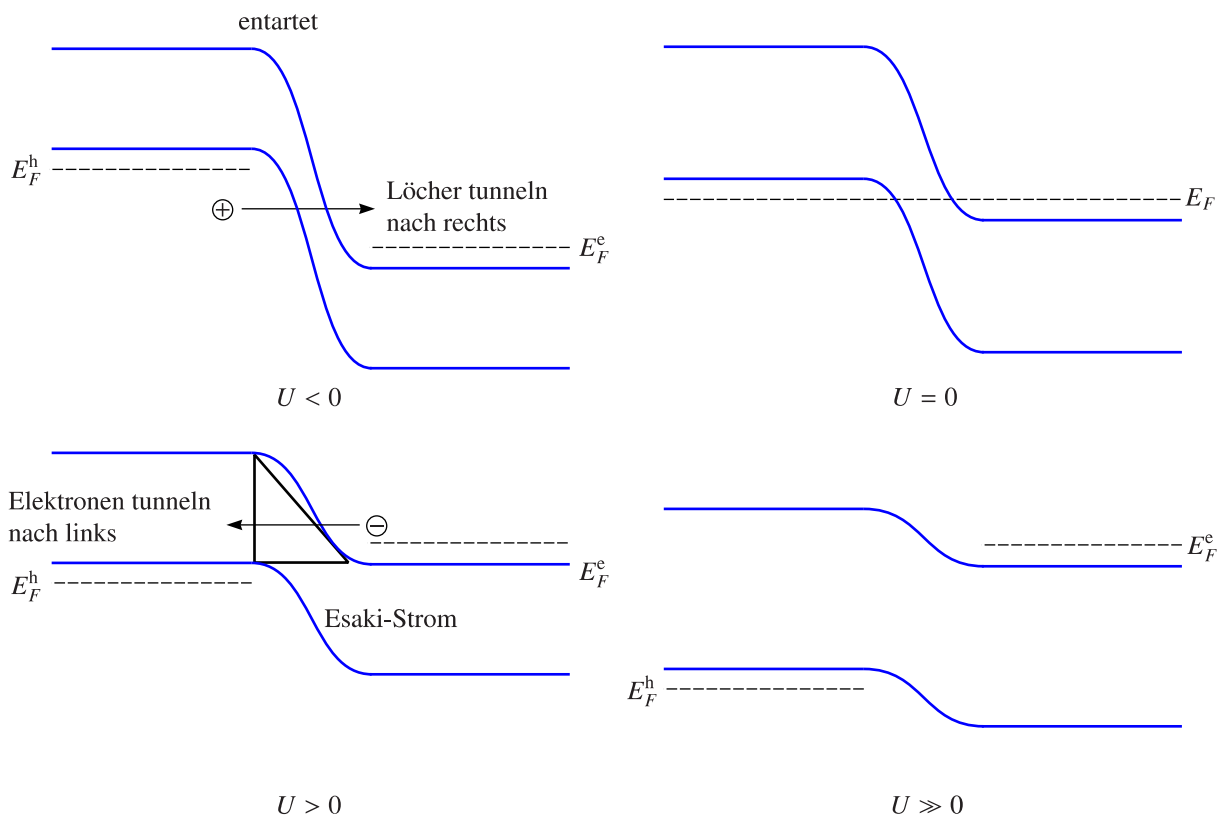
In den GaAs- und InGaAs-Schichten sind Quantentöpfe eingelagert. Liegt deren Breite in der Größenordnung der de-Broglie-Wellenlänge der Elektronen, so gibt es in einem Quantentopf nur ein einziger Quantenzustand und das Laserlicht wird durch optische Übergänge zwischen den Zuständen verschiedener Quantentöpfe erzeugt.

4.3 Spezielle Bauelemente zur Mikrowellenerzeugung

Mikrowellen kann man nicht mit gewöhnlichen Schwingkreisen erzeugen, weil diese eine zu hohe Frequenz besitzen. Man muss deshalb auf andere Techniken zurückgreifen, nämlich auf die Verwendung von Dioden:

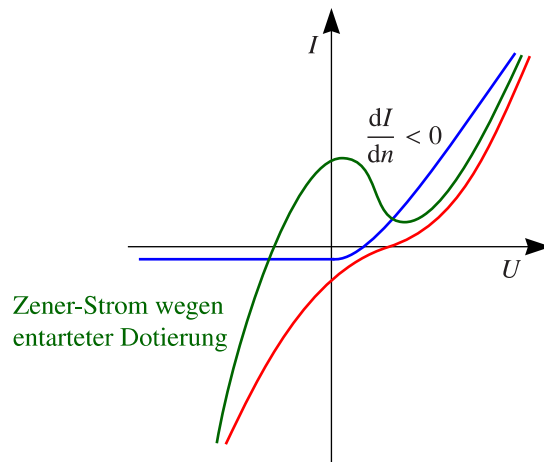


Wir dotieren das Ganze entartet, um den Zenerdurchbruch auf die Sprünge zu helfen.



Ist die angelegte Spannung < 0 , so können Löcher von links nach rechts tunneln. Ist dagegen $U > 0$, so tunneln Elektronen von rechts nach links und zwar durch eine näherungsweise dreieckige Potentialbarriere. Dies ist möglich, wenn Leitungsband im n-dotierten und Valenzband im p-dotierten Bereich auf derselben Höhe

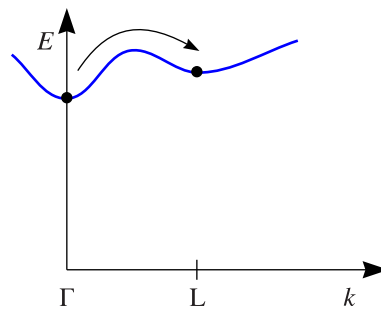
liegen. (Ladungsträger fangen in den Bereichen an zu tunneln, wo das Quasiferminiveau höher liegt, sofern die Barriere nicht zu hoch oder zu breit ist.) Der dabei entstehende Esaki-Strom führt zum charakteristischen Verlauf $dI/dU < 0$ und somit zu einer Phasenverschiebung, womit sich das Bauteil wie ein Kondensator verhält. Damit liegt dann ein Schwingkreis vor. Im Falle von $U \gg 0$ ist kein Tunneln möglich, höchstens indem zusätzlich Phononen erzeugt werden.



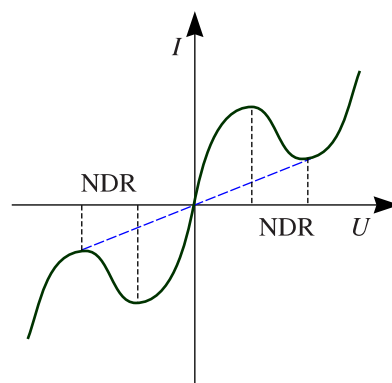
Überall, wo man Mikrowellen oder noch höherfrequente Strahlung erzeugen möchte, benutzt man solche negative differential resistances (NDR).

Es gibt noch ähnliche Bauelemente, wie beispielsweise:

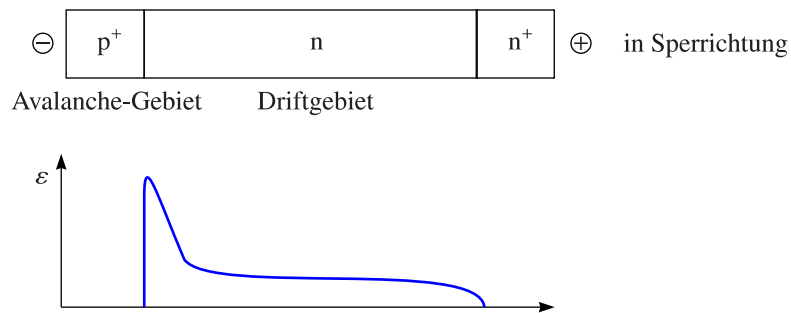
- i.) Gamm-Element: Man tunnelt vom Γ -Punkt zum L-Punkt im \mathbf{k} -Raum.



Bei hohem Feld ist ein Transfer von Elektronen ins L-Tal möglich:



- ii.) IMPATT-Diode (IMPact-Avalanche-Transit-Time-Diode):



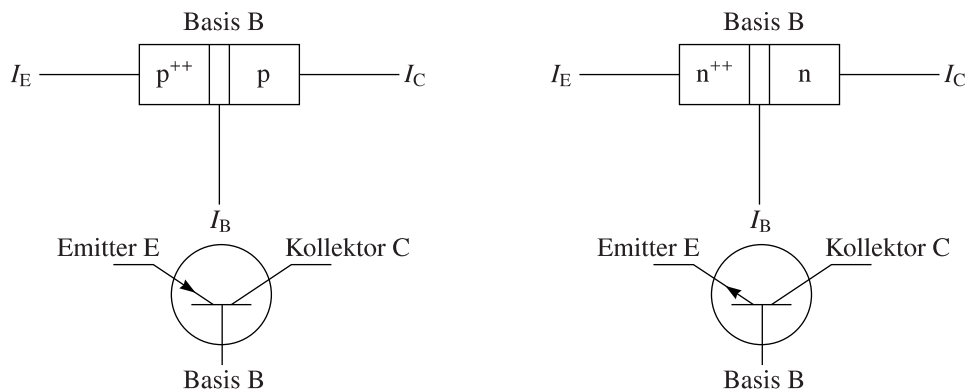
Eingebaut in einen Schwingkreis steigt die Spannung (Feld), bis der Avalanche-Durchbruch stattfindet. Dann werden Elektronen vom p^+ ins n -Gebiet injiziert. Diese driften dann zum n^+ -Gebiet. Löcher werden schnell am p^+ -Kontakt abgesaugt. Driftende Elektronen im n -Gebiet bewirken einen lokalen Anstieg des Feldes. Somit wird das Feld im Avalanche-Gebiet reduziert, bis es unter die kritische Durchbruchfeldstärke sinkt. Dann stoppt die Injektion. Der Strom fließt, bis alle Elektronen das Driftgebiet verlassen haben. Es gibt zwei Verzögerungszeiten:

- 1.) Aufbau des Avalanche-Durchbruchs
- 2.) Drift- und Transitzeit

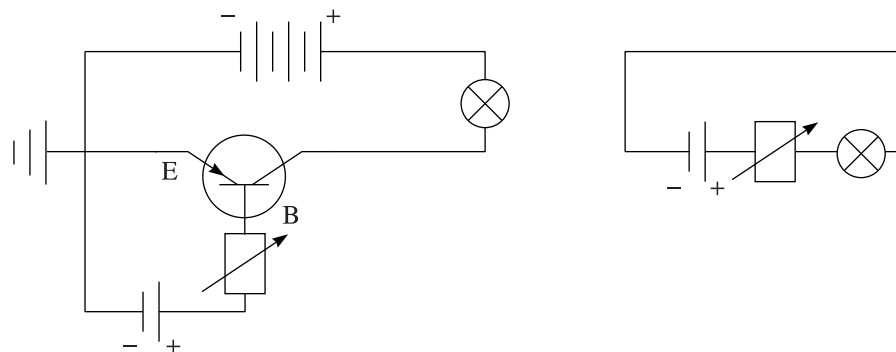
4.4 Der bipolare Transistor

4.4.1 Funktionsweise

Der bipolare Transistor wurde im Jahre 1947 erfunden. Es handelt sich dabei im Prinzip um zwei gegeneinander geschaltete pn-Dioden mit sehr dünner Basis. Dabei gibt es zwei Möglichkeiten, nämlich pnp oder npn-Transistoren.



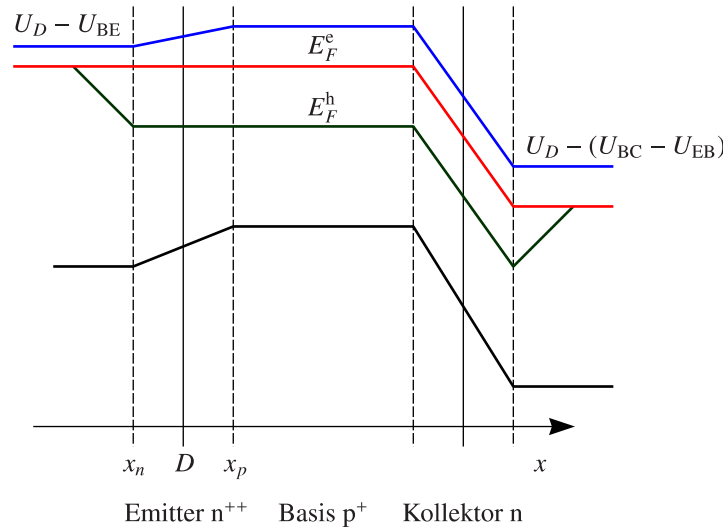
Typische Emitterschaltung (Basis-Emitterschaltung in Durchlassrichtung):



Kollektor und Basiskreis ist gesperrt. Die Potentiale in den Gebieten sind gegeben durch

$$\varphi_E < \varphi_B < \varphi_C \Rightarrow \varphi_B - \varphi_E = U_{BE} > 0, \quad \varphi_C - \varphi_E = U_{CE} > U_{BE} > 0. \quad (4.44)$$

Emitter injiziert Elektronen in Basis, die größtenteils vom Kollektor abgezogen werden. Damit besteht der Emitterstrom aus zwei Teilströmen: $I_E = I_B + I_C$.



Der Emitterstrom ist der Strom durch den ersten np-Übergang:

- Injektionsstrom von Elektronen $j_n(x_p)$
- Injektionsstrom von Löchern $j_p(x_n)$
- Rekombinationsstrom im Raumladungsgebiet ($x \approx 0$): Dieser ist vernachlässigbar, da die Diode in Durchlassrichtung geschaltet ist.

Mit der Emitterfläche F ergibt sich:

$$I_E = j_n(x_p) \cdot F + j_p(x_n) \cdot F, \quad j_n(x_p) = -eD_n \left. \frac{dn}{dx} \right|_{x=x_p}, \quad j_p(x_n) = eD_p \left. \frac{dp}{dx} \right|_{x=x_n}. \quad (4.45)$$

$j_p(x_n)$ berechnen wir wie beim pn-Übergang. (Wir nehmen dazu an, dass der Emitter unendlich groß ist.)

$$j_p(x_n) = ep_n \frac{D_p}{L_p} \left\{ \exp\left(\frac{eU_{EB}}{k_B T}\right) - 1 \right\}. \quad (4.46)$$

Die Breite der Basis B ist klein: $b < L_n$, wobei L_n die Diffusionslänge ist. Die Elektronen kommen als im Kollektor an, bevor sie in der Basis rekombinieren können. Nur ein Teil des Elektronenstroms rekombiniert in der Basis. Der Rest wird vom Kollektor abgesaugt. Der Gleichgewichtswert der Minoritäten in der Basis n'_{p0} wird nicht erst für $x \rightarrow \infty$, sondern $x = b$ eingenommen. Damit ergibt sich eine Randbedingung $\Delta n(x = b) = 0$. Aus den Diffusions-Rekombinationsgleichungen folgt

$$\Delta n(x) = [n(x_p) - n_{p0}] \frac{\sinh\left(\frac{b-x}{L_n}\right)}{\sinh\left(\frac{b-x_p}{L_n}\right)}. \quad (4.47)$$

und damit weiter

$$j_n(x_p) = en_p \frac{D_n}{L_n} \left\{ \exp\left(\frac{eU_{BE}}{k_B T}\right) - 1 \right\} \coth\left(\frac{b-x_p}{L_n}\right). \quad (4.48)$$

4.4.2 Verstärkungsverhalten des Transistors

Dieses beruht auf der Kleinheit des Basistroms. Also

$$j_{n,r} \cdot F + j_p(x_n) \cdot F, \quad (4.49)$$

wobei $j_{n,r}$ der Rekombinationsstrom und $j_p(x_n)$ die Löcherinjektion im Emitter ist.

- 1.) Da die Dicke der Basis b viel kleiner als die Rekombinationslänge L_n ist, kann der Rekombinationsstrom vernachlässigt werden.
- 2.) Die Basis ist geringer dotiert als der Emitter. Es gilt $j_p(x_n) \sim p_n$ und $n_n \gg p_p$, wobei n_n die Emitterdotierung und p_p die Basisdotierung ist. Das Ganze folgt aus dem Massenwirkungsgesetz (Übung).

Wähle also die Randbedingungen für das Potential so, dass das Potential der Basis φ_B um $d\varphi_B$ geändert wird, während φ_E und φ_C konstant bleiben. Hieraus folgt, dass sich U_{BE} um $dU_{BE} = d\varphi_B$ ändert, während U_{CE} konstant bleibt. Wir möchten I_C in Abhängigkeit von I_B bestimmen.

$$\left. \frac{dI_B}{dU_{BE}} \right|_{U_{CE}=\text{const.}} = \frac{\cosh\left(\frac{b-x_p}{L_n}\right) - 1}{\cosh\left(\frac{b-x_p}{L_n}\right)} \cdot \left. \frac{dI_E}{dU_{BE}} \right|_{U_{CE}=\text{const.}} \quad (4.50)$$

Weiterhin

$$\left. \frac{dI_C}{dU_{BE}} \right|_{U_{CE}=\text{const.}} = \frac{1}{\cosh\left(\frac{b-x_p}{L_n}\right)} \cdot \left. \frac{dI_E}{dU_{BE}} \right|_{U_{CE}=\text{const.}} \quad (4.51)$$

Die Stromverstärkung ist das Verhältnis der Stromänderung im Kollektor-Kreis zur Stromänderung im Basis-Kreis:

$$\beta = \left. \frac{dI_C/dU_{BE}}{dI_B/dU_{BE}} \right|_{U_{CE}=\text{const.}} \approx \frac{1}{\cosh\left(\frac{b}{L_n}\right) - 1}, \quad (4.52)$$

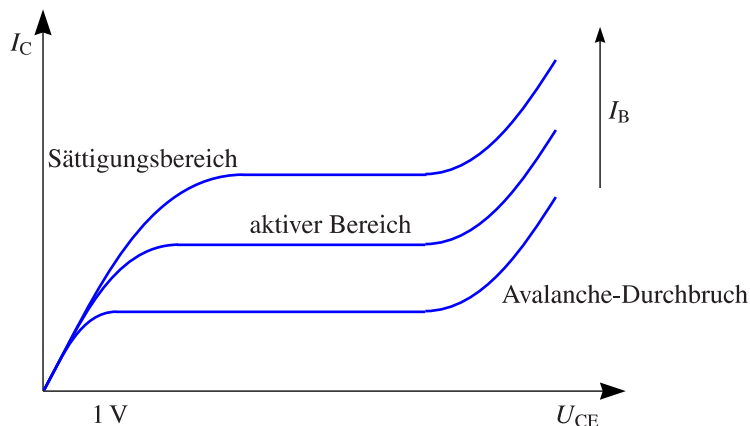
wobei x_p vernachlässigbar ist. Falls also $b \ll L_n$, wird β sehr groß. Eine kleine Änderung im Basisstrom führt somit zu einer großen Änderung im Kollektorstrom. Interessanter ist die Leistungsverstärkung. Dazu fügen wir einen Lastwiderstand R_L in den Kollektorkreis ein. Falls R_L gleich groß wie der Transistorwiderstand (gegeben durch den Widerstand des gesperrten pn-Übergangs) ist, gilt:

$$\frac{dP_L}{dP_B} = \beta \frac{U_{CE}}{U_{BE}} > \beta, \quad (4.53)$$

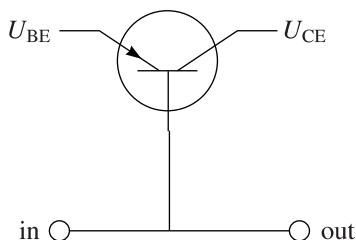
da $U_{CE} > U_{BE}$. Das Ganze ist so zu verstehen, dass eine kleine Steuerleistung eine große Nutzleistung regelt. Der bipolare Transistor besitzt einen kleinen Eingangswiderstand und einen großen Ausgangswiderstand (Einsatz beispielsweise im Handy). Bei einem unipolaren Bauelement (MOSFET) ist dies genau umgekehrt.

4.4.3 Das Kennlinienfeld

Wichtig ist die Ausgangskennlinie der Emitterschaltung, also $I_C(U_{CE})$ bei konstantem I_B . Es gilt weiterhin $U_{CE} = U_{CB} + U_{BE}$.

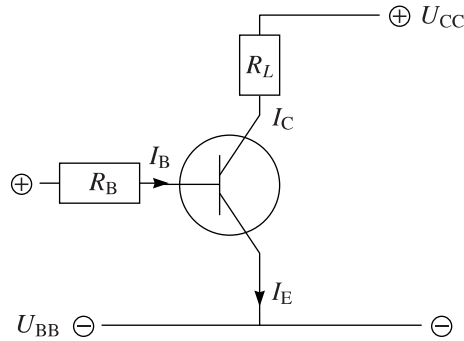


Falls $U_{CB} < 0$, sind beide pn-Übergänge in Durchlassrichtung und dann verschwindet I_C . Die Eingangskennlinien $I_B(U_{BE})$ ist die einer normalen Diode.



4.4.4 Betrieb des Transistors

Wir möchten nun die Widerstandskennlinie und den Arbeitspunkt konstruieren. Die Spannungen sind U_{BB} am Eingang und U_{CC} am Ausgang.



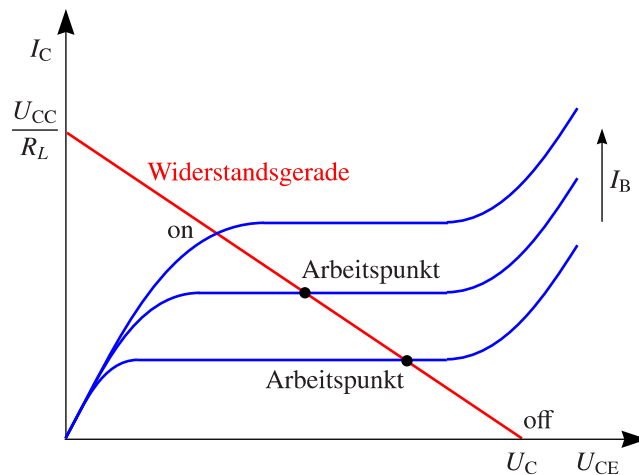
Aus $U_{EB} = I_B R_B + U_{BE}$ folgt

$$I_B = \frac{U_{BB} - U_{BE}}{R_B}, \tag{4.54}$$

wobei $U_{BB} \gg U_{BE}$, weil der Eingangswiderstand klein ist (Diode in Durchlassrichtung). Weiterhin ist $U_C = U_{CE} + I_C R_L$, also ist

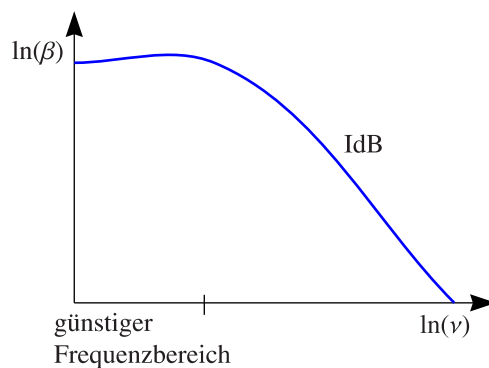
$$I_C = -\frac{1}{R_L} \cdot U_{CE} + \frac{U_{CC}}{R_L}, \tag{4.55}$$

die Widerstandsgerade.

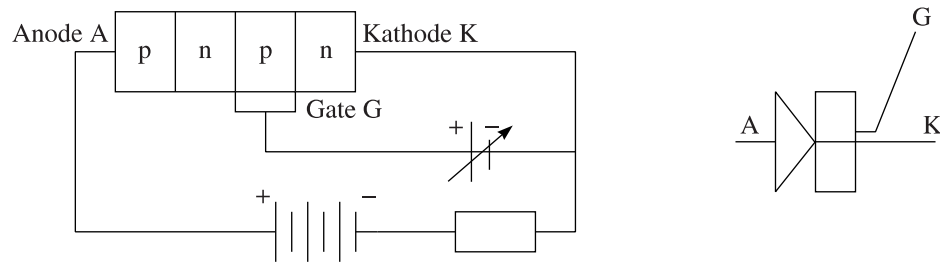


- Betrieb an den Punkten „on“ und „off“: Schalterbetrieb
- Betrieb an den Arbeitspunkten: Verstärkerbetrieb, da kleinen Änderungen in I_B große Änderungen in I_C hervorrufen

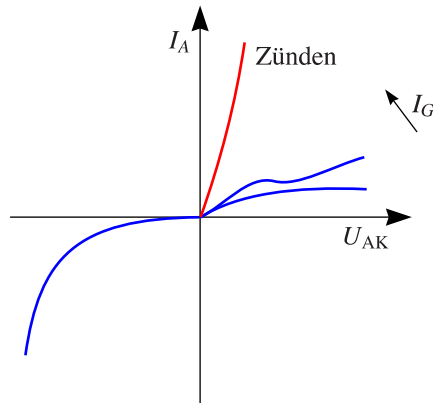
Die Grenzfrequenz hängt im Wesentlichen von der Transitzeit durch die Basis ab, also $\tau_t = b^2/(2D)$. Diese ist dann $\nu_{\text{Grenz}} = D/(\pi b^2)$. Diese Frequenz kann man zusätzlich durch ein intrinsisches Feld erhöhen. Die Grenzfrequenz liegt bei ungefähr 10 GHz.



4.5 Der Thyristor



Dabei handelt es sich um zwei stark gekoppelte Transistoren (anders als beim OPamp, wo man zwei hintereinander geschaltete Transistoren hat).



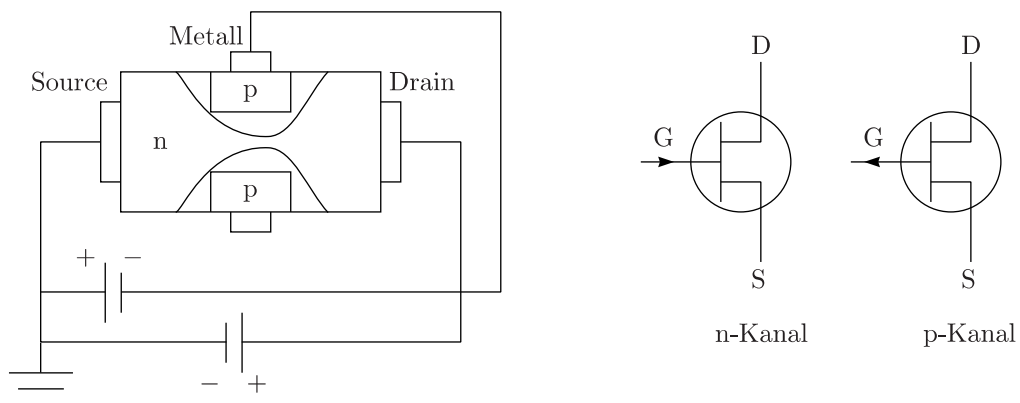
Anwendungen:

- Steuern, Schalter
- Gleichrichter für hohe Leistungen (kA bei kV), bei niedrigen Frequenzen (50 Hz)

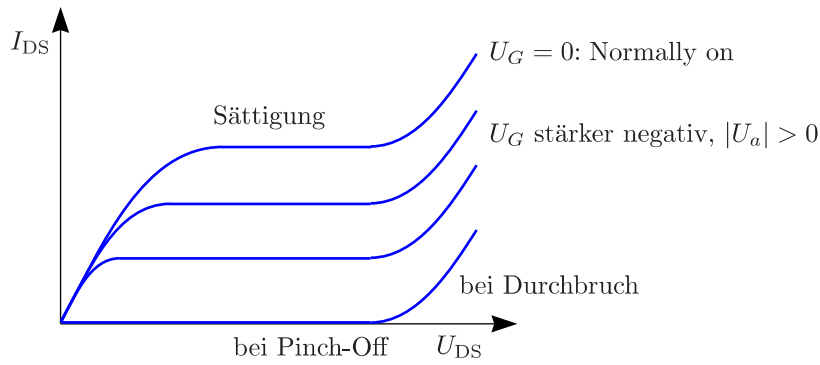
4.6 Unipolare Bauelemente

In unipolaren Bauelementen kommt vorwiegend ein Ladungsträgertyp vor.

4.6.1 Junction Field Effect Transistor (JFET)



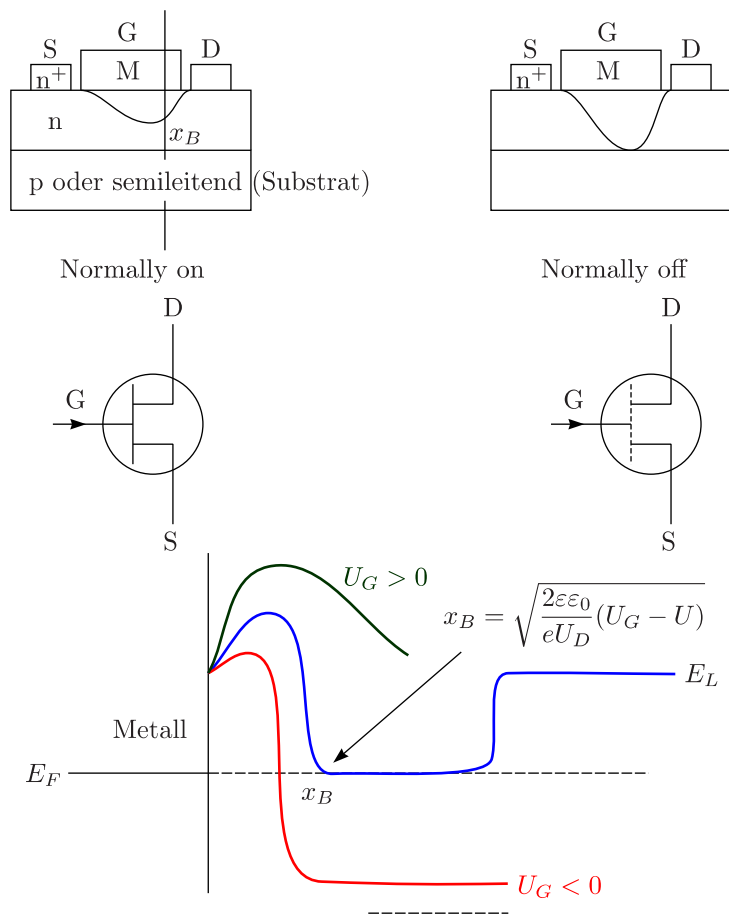
Die Verarmungszonen lassen einen Kanal zwischen Source und Drain. Die Spannung am Gate regelt die Breite des Kanals und damit den Widerstand. Wenn die beiden Verarmungszonen sich anfangen zu berühren, befinden wir uns im Pinch-Off-Bereich. Oft baut man n-Kanal-Bauelemente, da Löcher in p-Halbleitern eine geringere Mobilität haben.



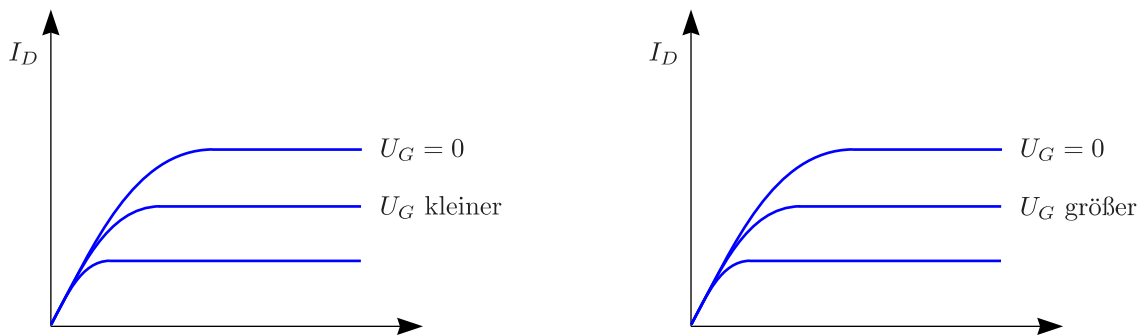
Die Vorteile des Bauelements sind:

- Es gibt keine Überschüsse von Minoritäten, die abgebremst werden müssen. Damit ist das Bauelement schnell.
- Es gibt nur eine Sorte von Ladungsträgern und damit kann die maximale Mobilität ausgenutzt werden.
- hohe Eingangsimpedanz
- Es regelt sich selbst, weil die Stromstärke abnimmt, je heißer das Bauelement wird. (Es gilt $dI/dT < 0$.) Damit ist der JFET temperaturstabil.
- „Normally on“ bzw. „Normally off“ sind realisierbar.

4.6.2 Metal Semiconductor Field Effect Transistor (MESFET)



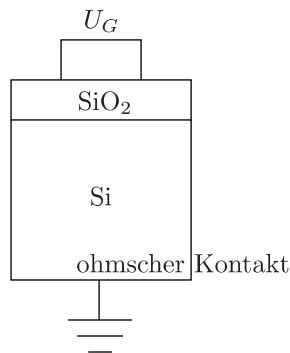
Berührt die Verarmungszone den semiisolierten Bereich, ist der Stromfluss stark verringert. Die Verarmungszone selbst besitzt eine asymmetrische Form, da die Gate-Spannung in Bezug auf Source anliegt. Über die Gate-Spannung lässt sich die Breite der Verarmungszone variieren; es besteht eine direkte Flexibilität über den Metall-Halbleiter-Übergang. Heutzutage setzt man eher MESFETs als JFETs ein, da letztere keine planare Geometrie haben und deshalb aufwändiger herzustellen sind.



Der Nachteil von MESFETs ist jedoch, dass Leckströme zum Gate hin auftreten können. Dies lässt sich jedoch verhindern, indem man eine Oxidschicht zwischen Source und Gate aufbringt. Diese Barriere verhindert Leckströme.

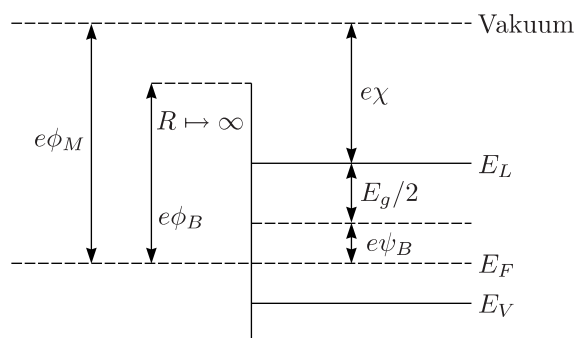
4.6.3 Metal Insulator Semiconductor Diode (MIS-Diode)

Eine alternative Bezeichnungsweise ist „Metal Oxide Semiconductor Diode“ (MOS-Diode). Bei der Herstellung verwendet man amorphes Silizium, da sich dieses gut oxidieren lässt und man sich keine Sorgen über Gitterfehler machen muss.

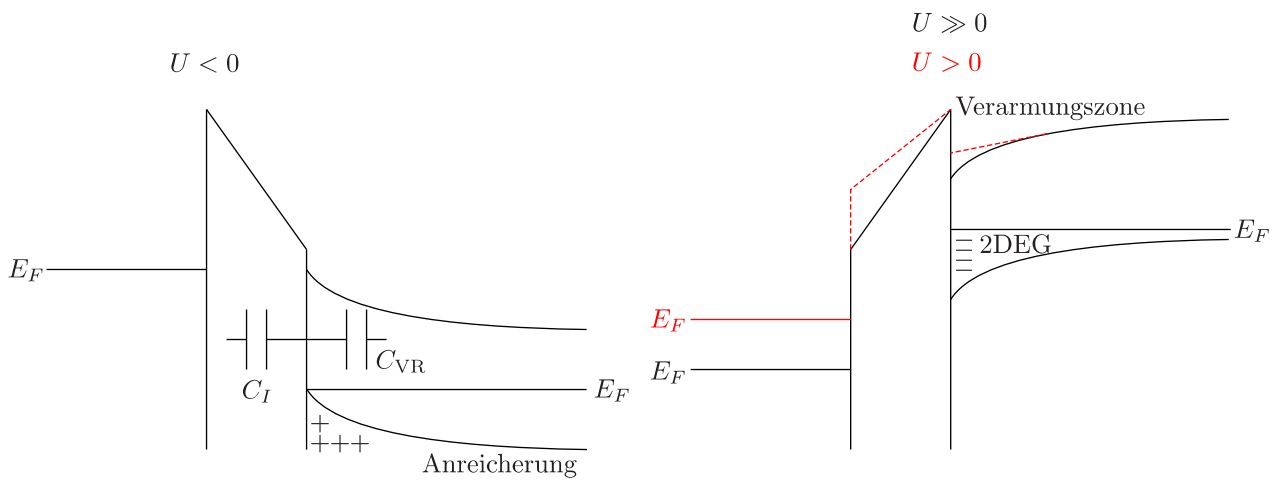


Wir betrachten eine ideale Diode; die Flachband-Bedingung lautet

$$e\phi_M - \left(e\chi + \frac{E_g}{2} + e\psi_B \right) = 0. \quad (4.56)$$



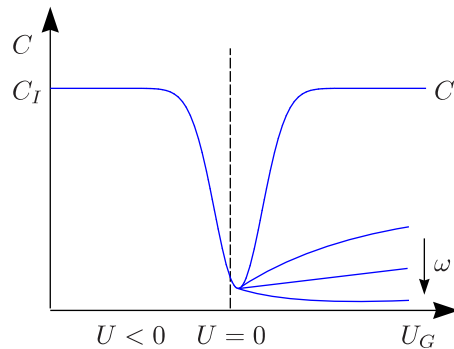
Für $U > 0$ werden Löcher von der Grenzfläche abgestoßen.



Mit einer CV-Messung bestimmt man:

$$C = \frac{C_I \cdot C_{VR}(V)}{C_I + C_{VR}(V)}, \quad C_I = \frac{\varepsilon \varepsilon_0 A}{d}. \quad (4.57)$$

Die Kapazität vom Isolator ist C_I und die Kapazität der Verarmungszone durch C_{VR} gegeben.



Bei hohen Frequenzen bildet sich eine große Verarmungszone und es lassen sich keine Minoritäten ansammeln. Damit kann kein zweidimensionales Elektronengas (2-DEG) entstehen. Auf diese Weise lässt sich untersuchen, bei welchen Frequenzen man das Bauelement betreiben kann.

4.6.4 Metal Oxid Semiconductor Field Effect Transistor (MOSFET)

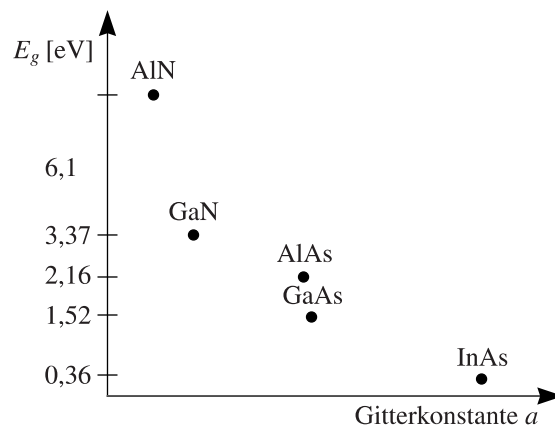
Beim MOSFET fließt zwischen Source und Gate kein großer Strom wegen der eingebrachten Oxidschicht. Weil das Gate keinen Strom durchlässt, besitzt das MOSFET somit einen großen Eingangswiderstand und jedoch kleine Ausgangswiderstände (im Gegensatz zum bipolaren Transistor). Geschaltet wird über die Spannung und nicht über den Strom. Eine Oxidschicht wird erzeugt, indem man das Oxid in den n-Bereich hineindiffundieren lässt. Für n- und p-Bereiche wachsen Oxide fast gleich schnell.

Kapitel 5

Halbleitertechnologie

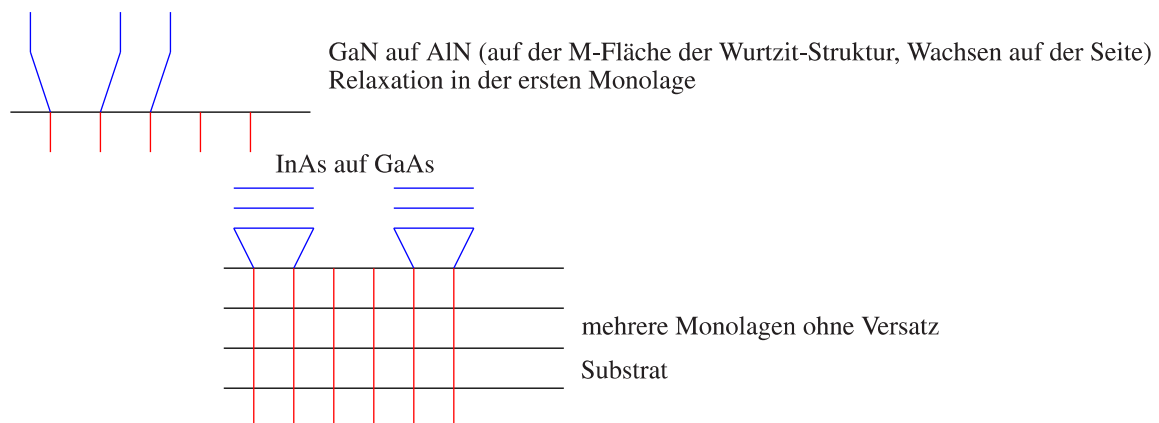
- Czochralski-Verfahren: Ziehen von Halbleitern aus einer Schmelze
- Epitaxie (Aufwachsen einkristalliner Schichten)

Man sollte Materialien mit ähnlicher Gitterkonstanten aufeinander wachsen lassen. Ansonsten entstehen Versetze im dreidimensionalen Gitter.

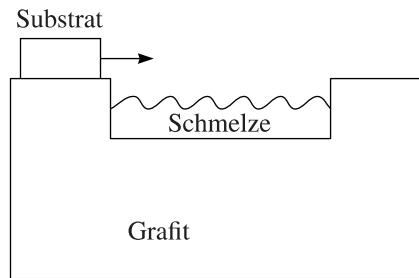


Verspannte Systeme entstehen für $a > b$, sofern a auf b .

- Vollmer-Weber (3D-Schichten): Die Adhäsion auf der neuen Schicht ist viel höher als auf der Oberfläche des Substrats, weshalb sich vor allem auf den neuen Schichten weitere Atome sammeln werden. Dies führt zu hohen Inseln an den Stellen, wo sich bereits Atome angesammelt haben.
- Frank-van-der-Merve-Wachstum (2D-Schichten): Hier wächst jede Monolage für sich komplett, weil die Adhäsion auf der Monolage der Adhäsion auf der Oberfläche des Substrats entspricht. Damit entstehen gleichmäßig immer neue Lagen (beispielsweise AlAs auf GaAs).
- Stranski-Krastanov-Wachstum: Hier ist die Adhäsion auf der ersten Monolage viel höher als auf der Oberfläche des Substrats. Es entsteht eine komplette Monolage und danach findet das Wachstum einzelner Inseln statt ($1 - x$ Monolagen auf 2D-Schicht, ab $x + 1$ Monolagen auf 3D-Schicht).



a.) Liquid phase epitaxy



b.) Molecular beam epitaxy

Das Verfahren ist ziemlich teuer und wird vor allem eingesetzt, wenn Halbleiter ohne Fehlstellen, an denen nichtstrahlende Rekombination stattfinden kann, hergestellt werden sollen (beispielsweise zur Verwendung in Satelliten).

c.) Channel Vapor Deposition (CVD), Metal Organic CVD (MOCVD)

Im Gegensatz zur Molecular beam epitaxy ist diese Methode nicht so sauber, weil man bei hohen Temperaturen arbeitet. Daher wird sie vor allem dann verwendet, wenn es um die möglichst billige Herstellung von Massenware wie beispielsweise LEDs geht.



Das Ganze ist für Nitride einfacher, weil man da mit gasförmigem Ammoniak NH_3 arbeitet, anstelle des giftigen AsH_3 , was im ersten Weltkrieg als Kampfgas eingesetzt wurde.

Damit sind wir in der Lage, InGaAs-Quantenpunkte in einer GaAs-Matrix herzustellen. Ober- und unterhalb fügen wir zusätzlich DBRs (Distributed Bragg-Reflectors) hinzu und können auf diese Weise einen Laser herstellen. Was man dann noch braucht, sind Kontakte.

5.1 Dotierungen

- Legieren
- Diffundieren
- Implantieren

