

# Halbleiterbauelemente

**4. Semester**

**2011**

**Prof. Andreas Schenk**

Copyright ©

# Kurze Geschichte des Transistors

## Karl Ferdinand Braun (1850–1918)

- 1874: Entdeckung von Leitfähigkeit und Gleichrichter-Effekt in Metall-Sulfid-Kristallen
- 14. November 1876 in Leipzig: Demonstration des Gleichrichter-Effekts eines Metall-Halbleiter-Kontakts (Blei-Sulfid) vor grossem Publikum



## Meilensteine der Physik

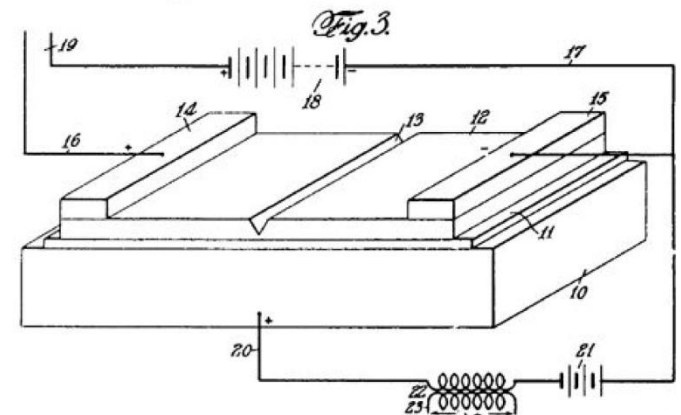
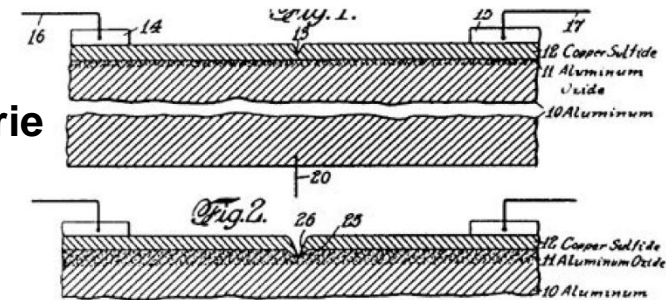
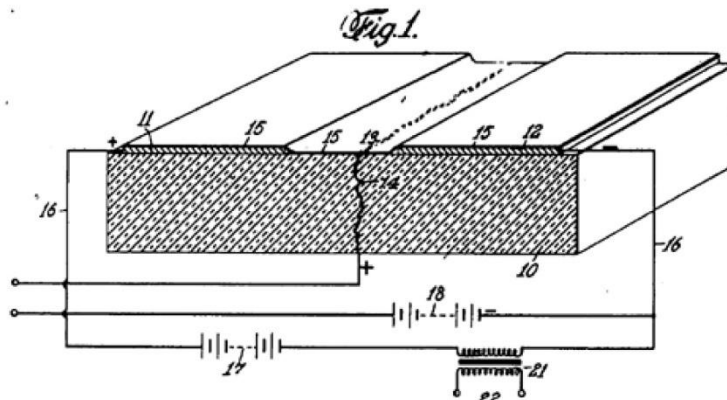
- 1831: Michael Faraday entdeckt die elektromagnetische Induktion
- 1864: James Maxwell publiziert seine Theorie des elektromagnetischen Feldes
- 1888: Heinrich Hertz entdeckt die Radio-Wellen
- 1897: Entdeckung des Elektrons durch Joseph J. Thomson
- Albert Einstein und das Wunderjahr 1905
- Alte (Niels Bohr, 1913) und neue Quantentheorie (Werner Heisenberg, ..., 1927)
- Begründung der Festkörperphysik in den 20er und 30er Jahren

# Julius Edgar Lilienfeld (1881-1963)



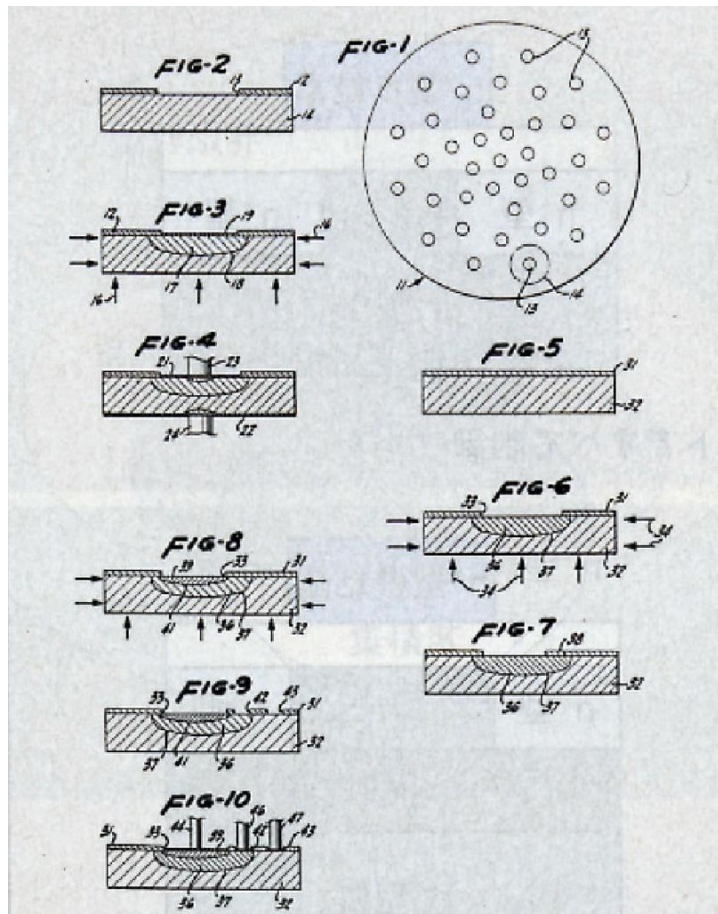
- Multitalent. 60 US-Patente, 15 deutsche Patente
- Erfinder des Prinzips des Feldeffekt-Transistors
- 8. Oktober 1926: Metall-Halbleiter (MES)FET
- 28. März 1928: depletion mode MOSFET
- Selbst wenn Lilienfeld versucht hätte, diese FETs zu fertigen, hätten sie nicht funktioniert wegen der zu schlechten HL-Materialeigenschaften
- Seine Ideen eines Festkörper-Verstärkers hatten daher keinen praktischen Wert für die Radio-Industrie der 20er und 30er Jahre

Jan. 28, 1930. J. E. LILIENFELD 1,745,175  
 METHOD AND APPARATUS FOR CONTROLLING ELECTRIC CURRENTS  
 Filed Oct. 8, 1926



# Oskar Ernst Heil (1908–1994)

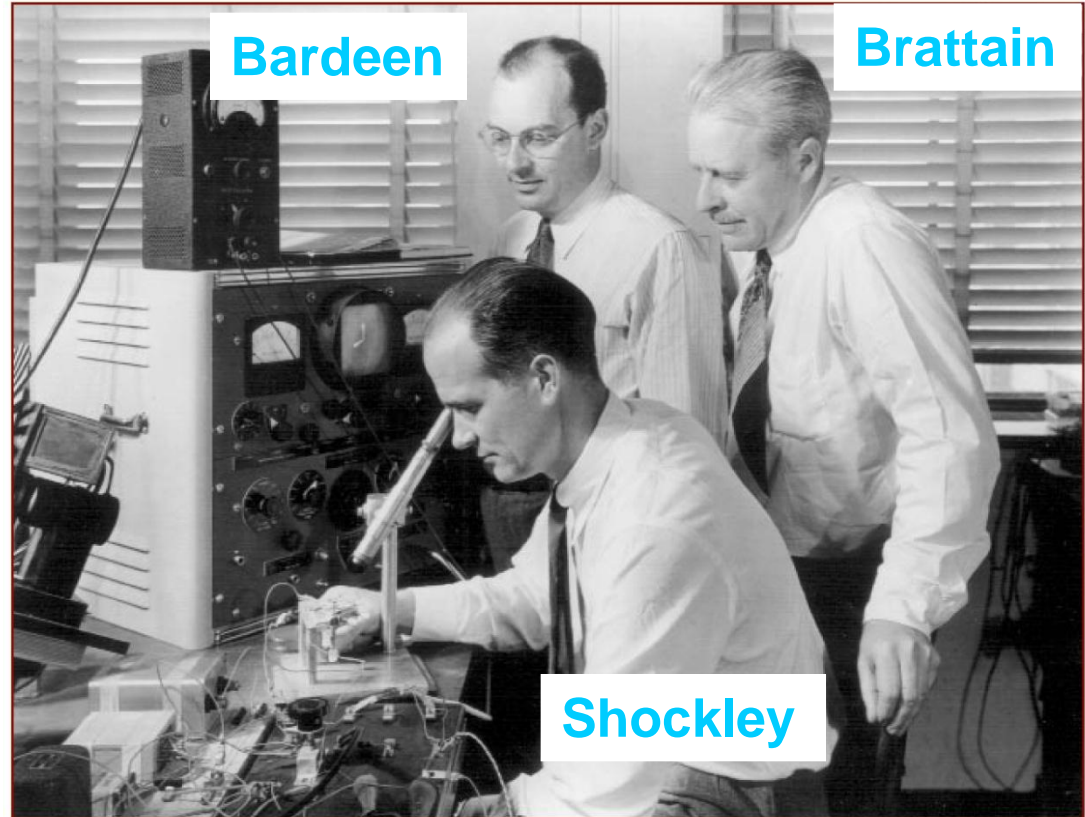
- 2. März 1934: Erfindung der kapazitiven Stromsteuerung in Feldeffekt-Transistoren (British Patent No. 439457)

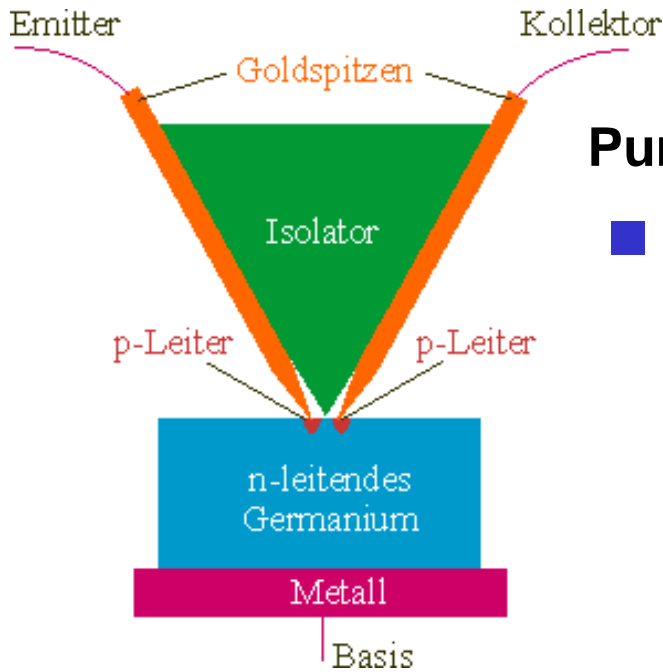


- Es sollte bis 1960 dauern, bevor der erste MOSFET experimentell realisiert werden konnte.
- Die extrem hohe Dichte von (geladenen) Defekten an der Halbleiter-Isolator-Grenzfläche verhinderte das Eindringen des elektrischen Feldes und damit den Feldeffekt.

# Die Erfindung des Bipolartransistors

- Dezember 1947: John Bardeen, Walter Brattain, William Shockley
- 1936 war bei Bell Laboratories durch Mervin Kelly eine Forschergruppe gegründet worden. Ziel: Untersuchung von Halbleiter-Eigenschaften, um mögliche Anwendungen zu verstehen.
- Neustart 1946 mit Fokus auf Germanium und Silizium





## Punkt-Kontakt-Transistor (Spitzen transistor)

- Goldfolie über Plastik-Keil wurde an der Spitze durchtrennt und auf Germanium-Oberfläche gepresst (Formierung)

- Unter den Spitzen (Abstand etwa 20 Mikrometer) bilden sich p-leitende Gebiete
- Während Brattain und Bardeen an Oberflächen-Effekte glauben, legt Shockley eine Theorie basierend auf Minoritätsladungsträgern vor

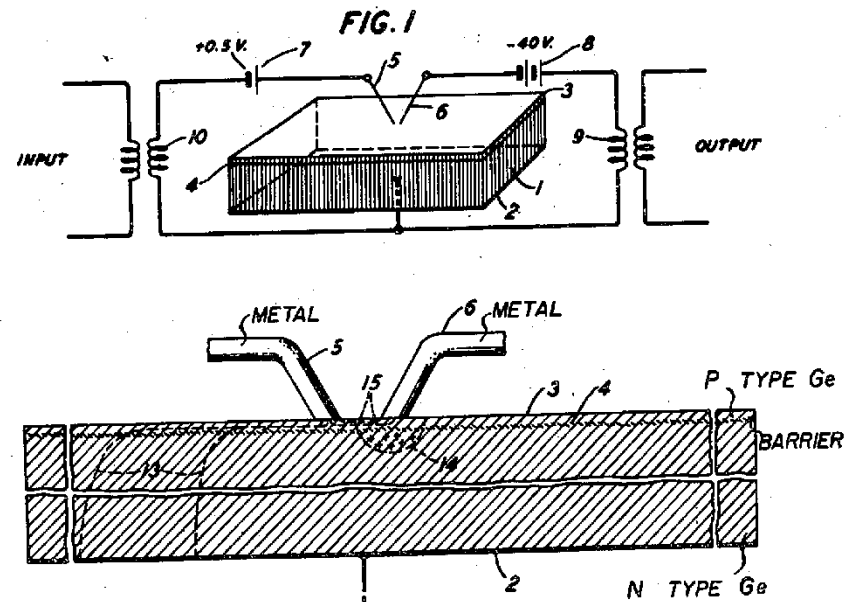
Oct. 3, 1950

J. BARDEEN ET AL  
THREE-ELECTRODE CIRCUIT ELEMENT UTILIZING  
SEMICONDUCTIVE MATERIALS

2,524,035

Filed June 17, 1948

3 Sheets-Sheet 1



# Wo kommt eigentlich der Name *Transistor* her?

BELL TELEPHONE LABORATORIES  
INCORPORATED

~~B. T. L. CONFIDENTIAL~~

## COVER SHEET FOR TECHNICAL MEMORANDA

SUBJECT: Terminology for Semiconductor Triodes - Committee Recommendations - Case 38139-8

### COPIES TO:

- 1 - Dept. 1000 File  
2 - R. Boyl - Case File  
3 - R. K. Potter  
4 - J. R. Wilson  
5 - G. W. Gilman  
6 - J. W. McRae  
7 - H. S. Black  
8 - H. C. Hart  
9 - R. C. Mathes  
10 - C. B. Feldman  
11 - W. E. Kock-R. L. Wallace  
12 - J. A. Becker-J. N. Shive  
13 - W. Shockley  
14 - J. H. Scaff-W. G. Pfann  
15 - J. A. Bardeen  
16 - W. H. Brattain  
17 - A. C. Norwine-D. M. Chapin  
18 - A. J. Rack-S. E. Michaels  
19 - F. Gray

MM-48-130-10  
DATE May 28, 1948  
AUTHOR L. A. Neacham  
C. O. Mallinckrodt  
H. L. Barney

Surface States -  
Terminology

### ABSTRACT

- 20 - J. R. Pierce  
21 - J. G. Kreer  
22 - J. O. Edson  
23 - M. E. Mohr  
24 - L. A. Neacham  
25 - C. O. Mallinckrodt  
26 - H. L. Barney-E. Dickten

### ABSTRACT

Recommendations are made for an equivalent circuit representation, and terminology relating to semiconductor triodes.

~~B. T. L. CONFIDENTIAL~~

Terminology for Semiconductor Triodes - Committee Recommendations - Case 38139-8

MM-48-130-10

May 28, 1948

## BALLOT

Designate by the numbers 1, 2 and 3, the order of your preference for the names listed below:

- \_\_\_ Semiconductor Triode  
\_\_\_ Surface States Triode  
\_\_\_ Crystal Triode  
\_\_\_ Solid Triode  
\_\_\_ Iotatron  
\_\_\_ Transistor  
\_\_\_ \_\_\_\_\_ (Other suggestion)

Comments: \_\_\_\_\_

Signed \_\_\_\_\_

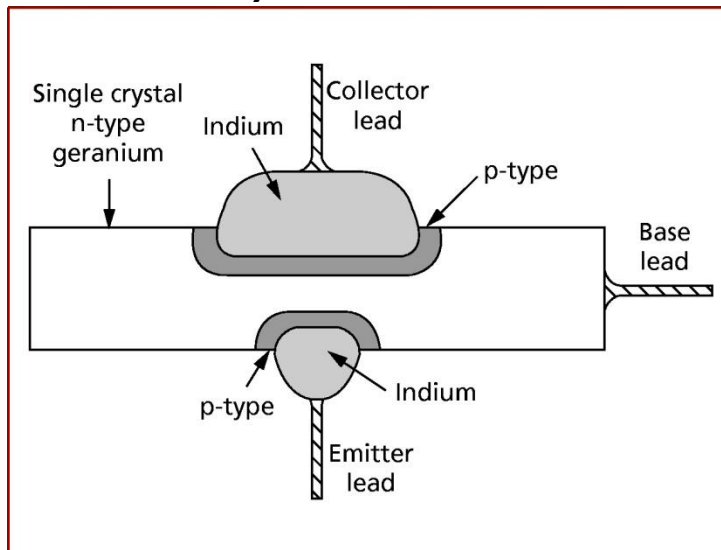
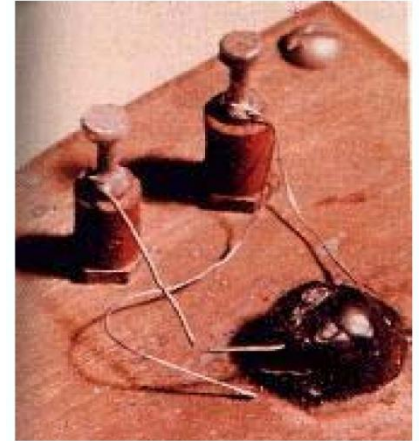
Please return this ballot to Miss G. R. Callender in 1A-323 at Murray Hill.

*transconductance or transfer + varistor*

transconductance = Gegenwärtigkeitwert  
varistor = voltage dependent resistor

# Junction-type Bipolartransistor

- 1948: John Shive (Bell Labs) benutzt eine 0.1mm dünne Germanium-Scheibe mit seitlichen Kontakten aus Phosphor-Bronze und erhält beträchtliche Stromverstärkung
- Dieses Experiment beweist den Volumen-Effekt und bestätigt Shockleys Theorie des pn-Übergangs
- Mit der Verbesserung des kontrollierten Dotierens während des Kristallwachstums konnte 1950 der erste gewachsene pnp-Bipolartransistor von Shockley, Teal und Sparks (Bell Labs) demonstriert werden (Bilder rechts)

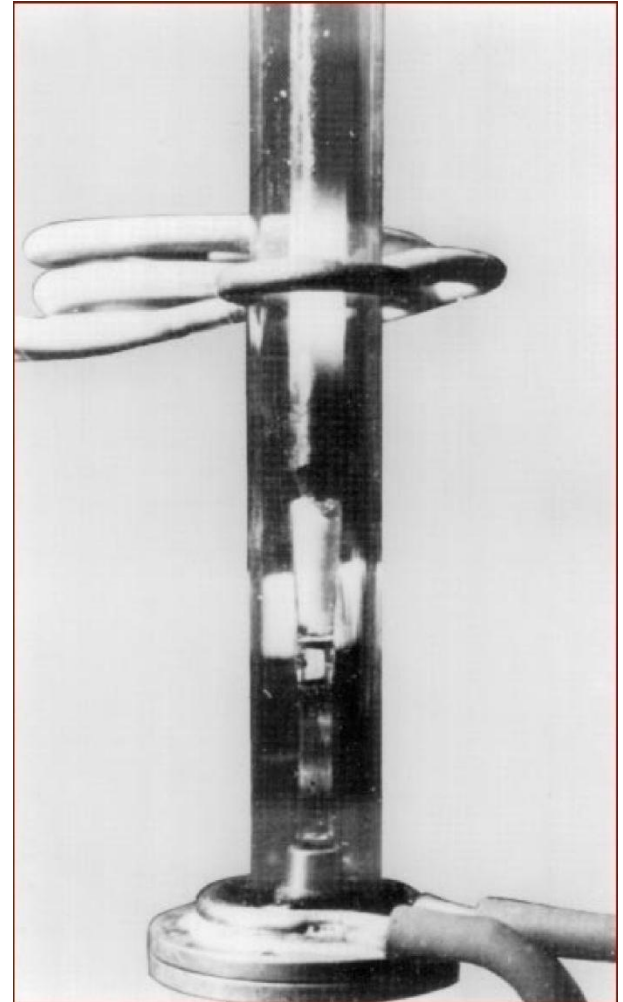


- Auch der Junction-type Bipolartransistor war für eine Massenfertigung nicht geeignet
- 1952: John Saby (GE) produziert den ersten Legierungstransistor (Bild links)
- Gute Ausbeute, reproduzierbare Kennlinien; wird Stütze der Industrie für einige Jahre



# Silizium ersetzt Germanium

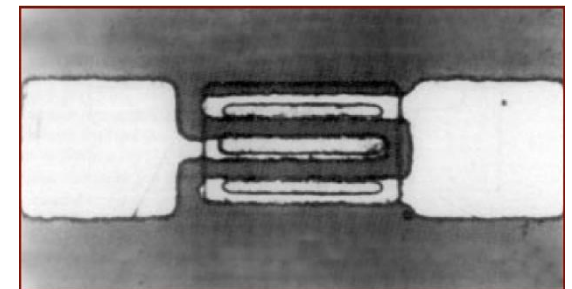
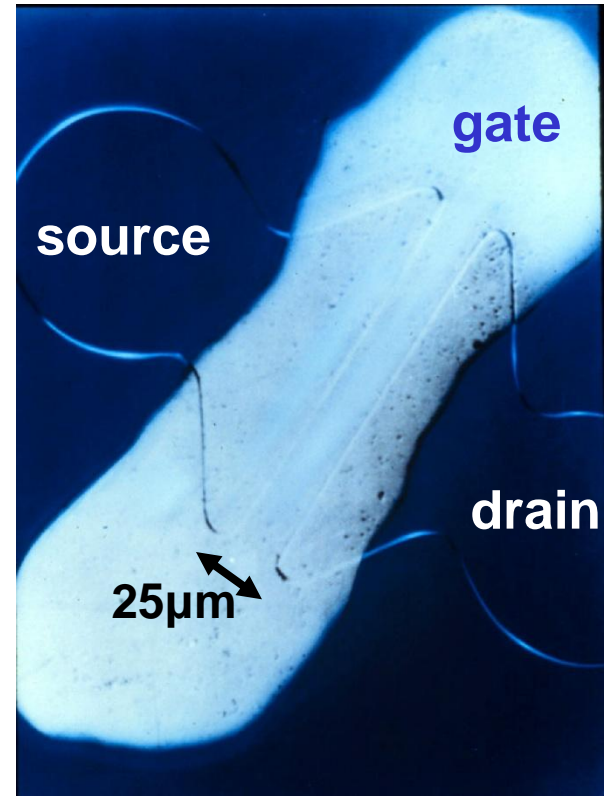
- 1823 entdeckt Jöns Jakob Berzelius das Silizium
- 1886 entdeckt Clemens Winkler das Germanium, das einige Jahre vorher von Mendelejew anhand seines PSE als “Ekasilizium” vorausgesagt worden war
- Für die Halbleiterelektronik erlangte Silizium erst Bedeutung, als man es in einer Reinheit von 99.9999999% herstellen konnte (ein Fremdatom pro eine Milliarde Si-Atome)
- 1952 erfindet Bill Pfann (Bell Labs) das Float-Zone-Verfahren (zuerst für Germanium)
- 1953 wendet Henry Theuerer (Bell Labs) das Verfahren auf Silizium an (senkrechte Anordnung, Bild)



# Der Weg zum MOS Feldeffekt-Transistor

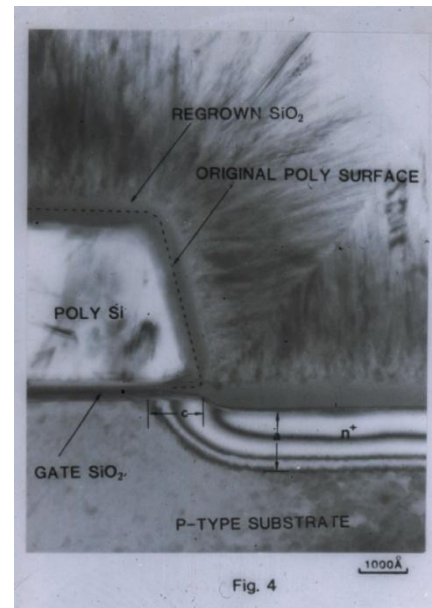
- 1955: C. J. Frosch und L. Derick entdecken die Maskierungseigenschaften von  $\text{SiO}_2$  (Oxid verhindert die Eindiffusion von Dotanden)
- 1959 zeigen Brattain und Shive, dass das  $\text{SiO}_2$  die Dichte der Grenzflächenzustände so stark reduziert, dass der Feldeffekt beobachtbar wird
- 1960 demonstrieren D. Kahng und M. M. Atalla (Bell Labs) den ersten funktionierenden MOSFET (Bild) mit 25  $\mu\text{m}$  Kanallänge
- Ein grosses Problem war die Zuverlässigkeit. Insbesondere Natrium an der  $\text{Si-SiO}_2$ -Grenzfläche hatte verheerenden Effekt
- Den Durchbruch bringt eine Erfindung von Jean A. Hoerni (Fairchild Semiconductors) von 1958: die  $\text{SiO}_2$ -Maske wird nicht entfernt, sondern dient als Schutzschild gegen Kontamination

erster planarer Bipolartransistor,  
Hoerni (1960)

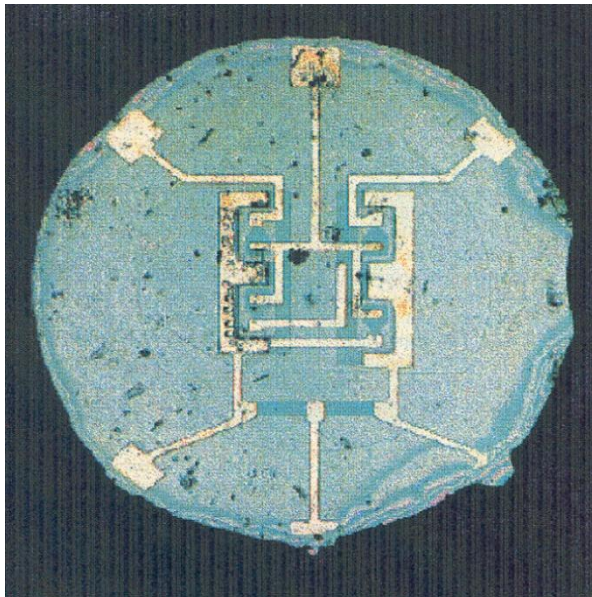
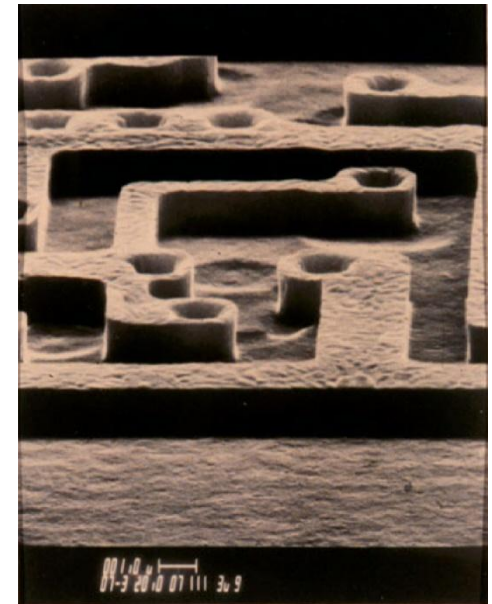


- 1963 zeigen Gordon E. Moore und Robert N. Noyce, dass Aluminium gute Kontakte auf n- und p-type Silizium macht
- Letzte Probleme mit den Oxiden werden bis etwa 1966 gelöst
- Erst mit Beginn der 70er Jahre finden planare MOSFETs massenhafte Verbreitung

MOSFET 1981



1µm nMOSFET 1984

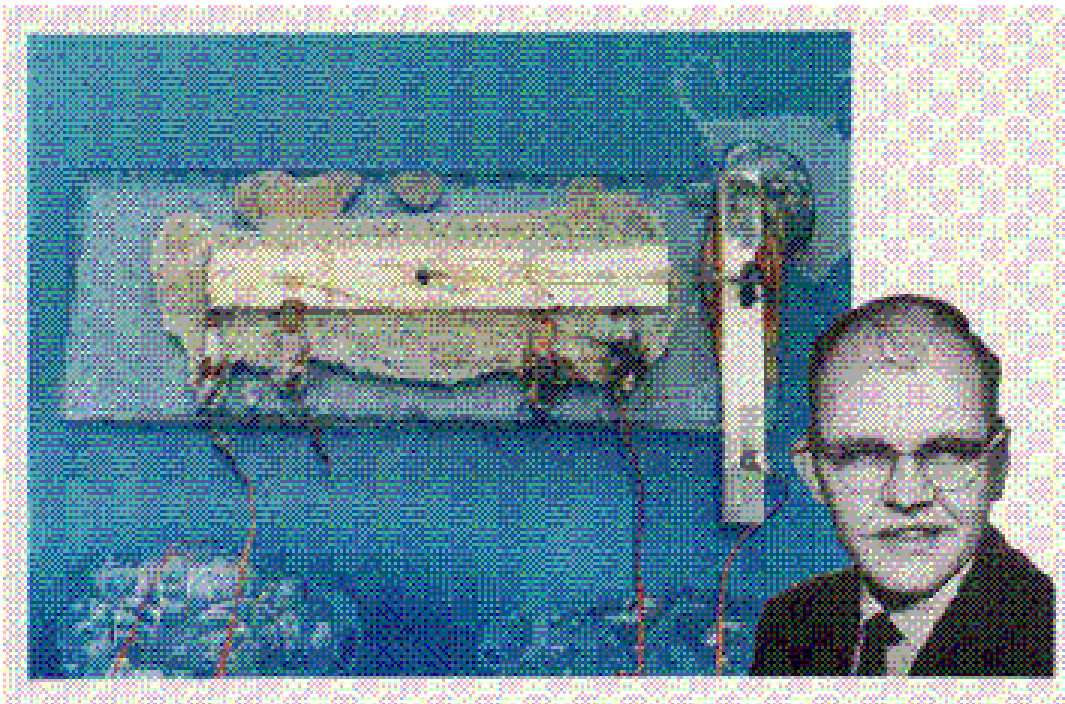


1961 Fairchild RTL IC (Bipolar Flip-Flop)

- Der erste vollständige Planarprozess wird von Jean A. Hoerni (Fairchild) 1960 beschrieben
- Dies führt 1961 zum ersten Fairchild IC mit einer Metallisierungsebene

# Der erste IC

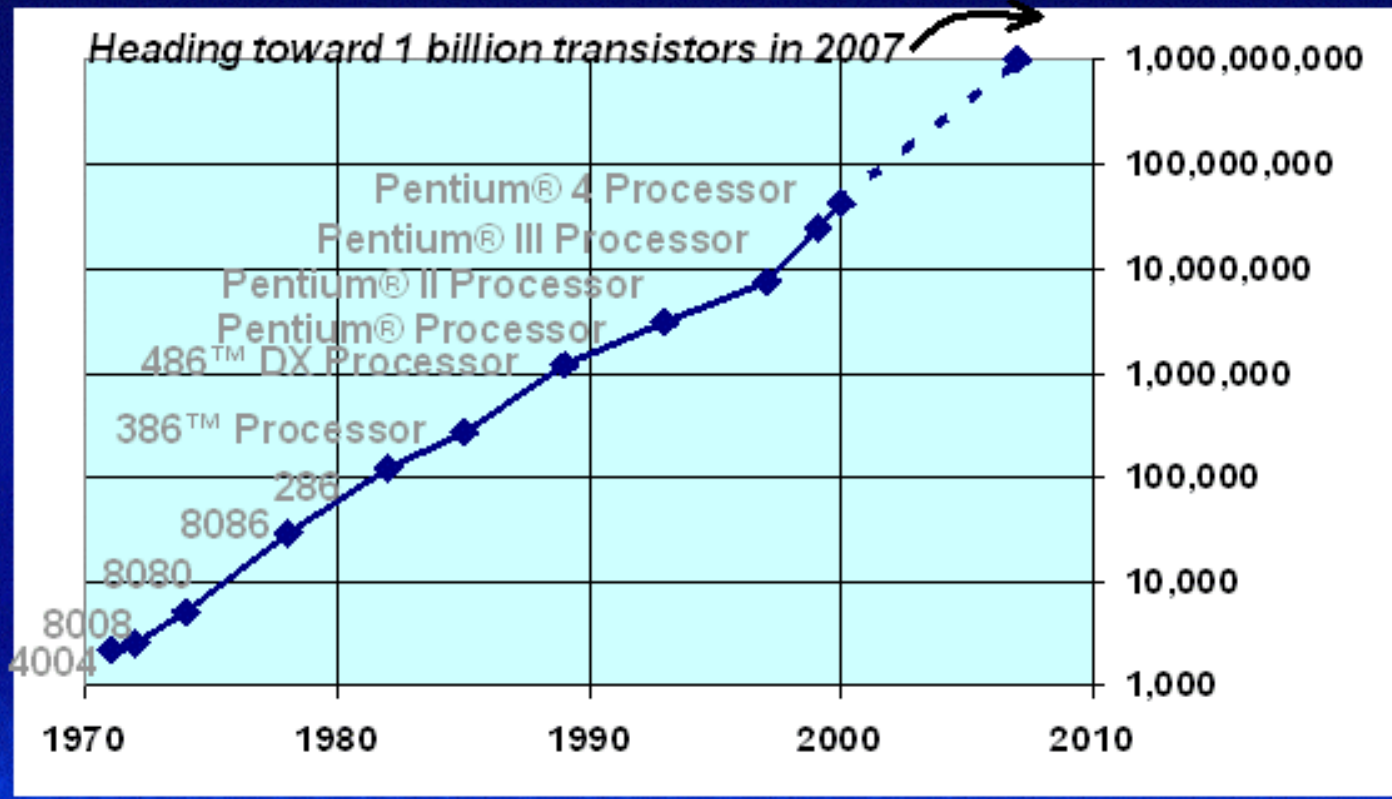
- 12. September 1958: Jack Kilby (Texas Instruments)
- Vorher wurden die Transistoren diskret verdrahtet, was zu einer grossen Zahl von Verbindungen führte
- Kilby hatte die Idee, die Silizium-Oberfläche selektiv zu dotieren und so Regionen im monolithischen Material zu erzeugen, die als Transistoren, Dioden, Widerstände und Kondensatoren fungieren



- Der erste IC war aus Germanium, weil Kilby kein geeignetes Stück Silizium auftreiben konnte und enthielt 1 Transistor, 1 Kondensator und 1 Widerstand

# Moore's Law Continues

## Transistors doubling every 2 years toward the billion-transistor microprocessor



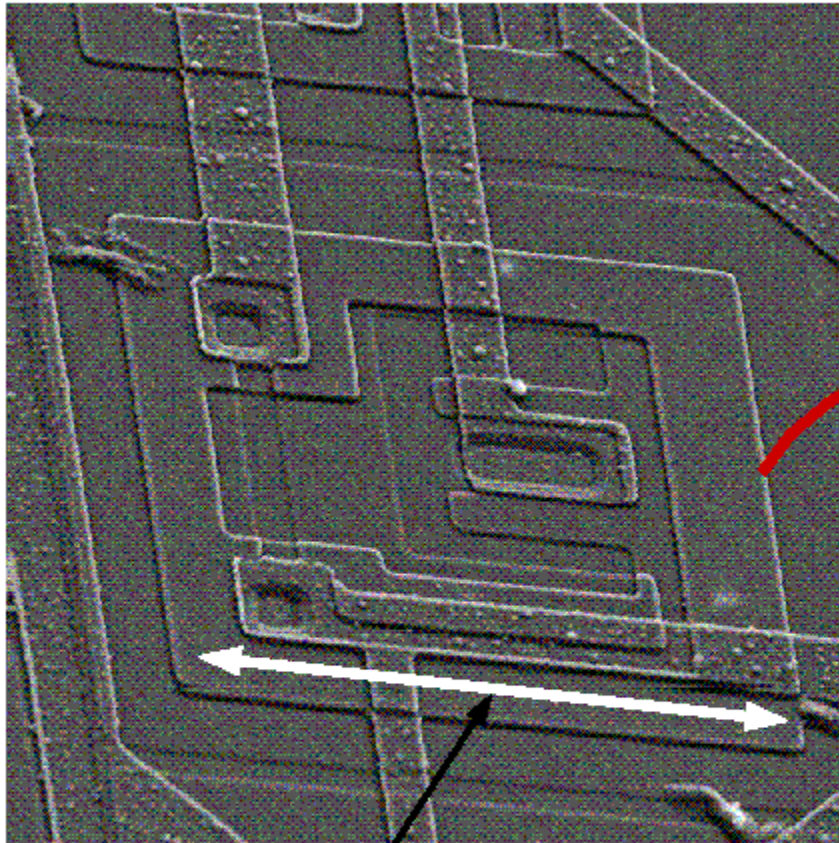
4



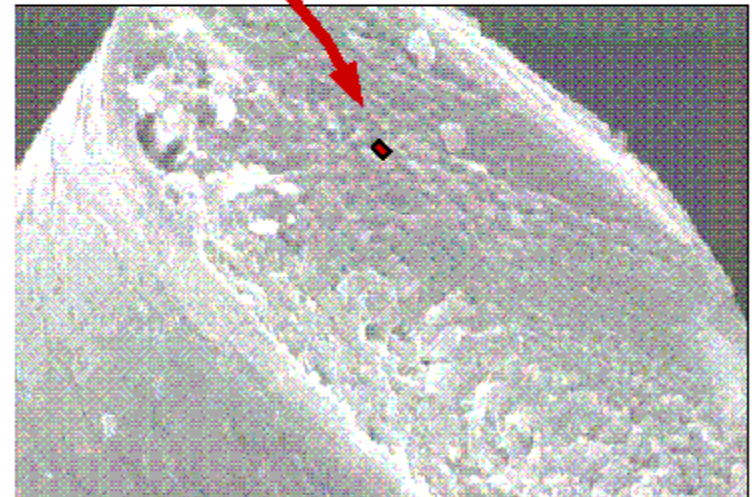
# Anfang 80er Jahre

Zum Vergleich:

Querschnitt eines  
menschlichen Haares

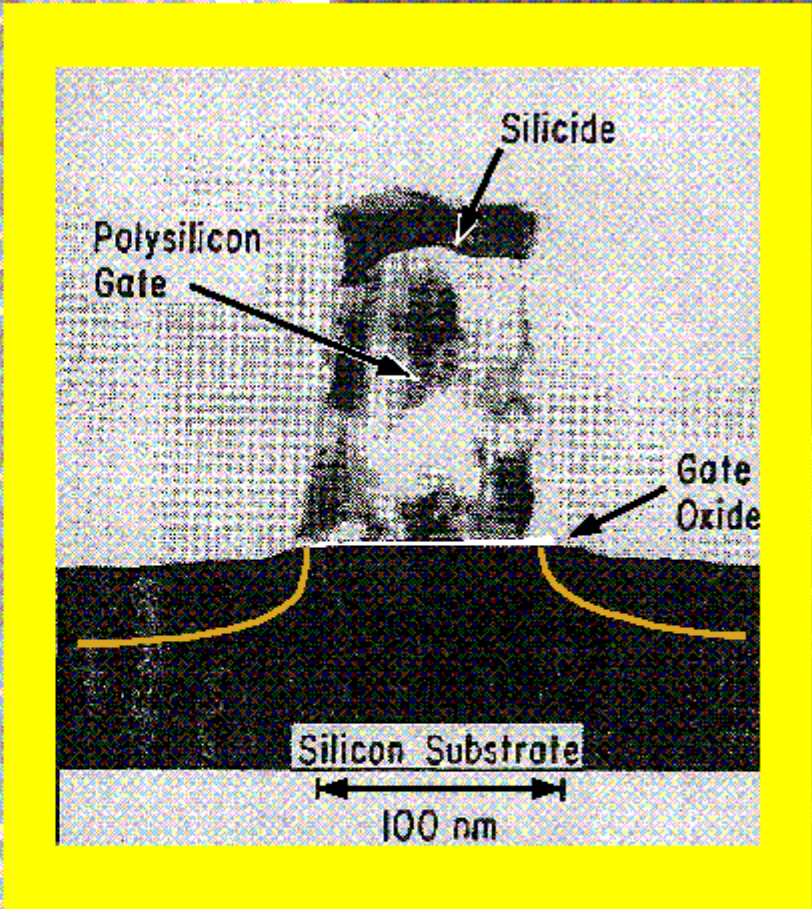


$\sim 2 \mu\text{m}$



$10 \mu\text{m}$

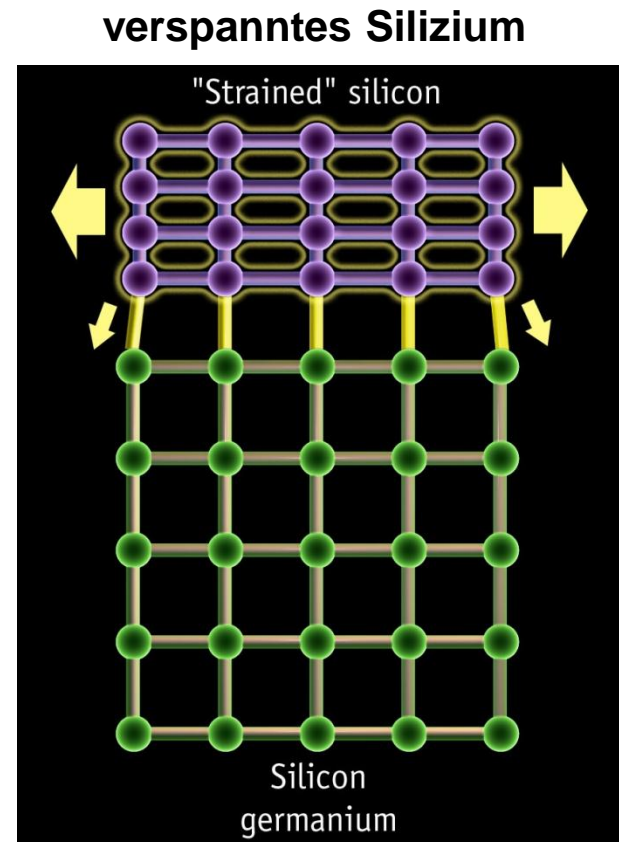
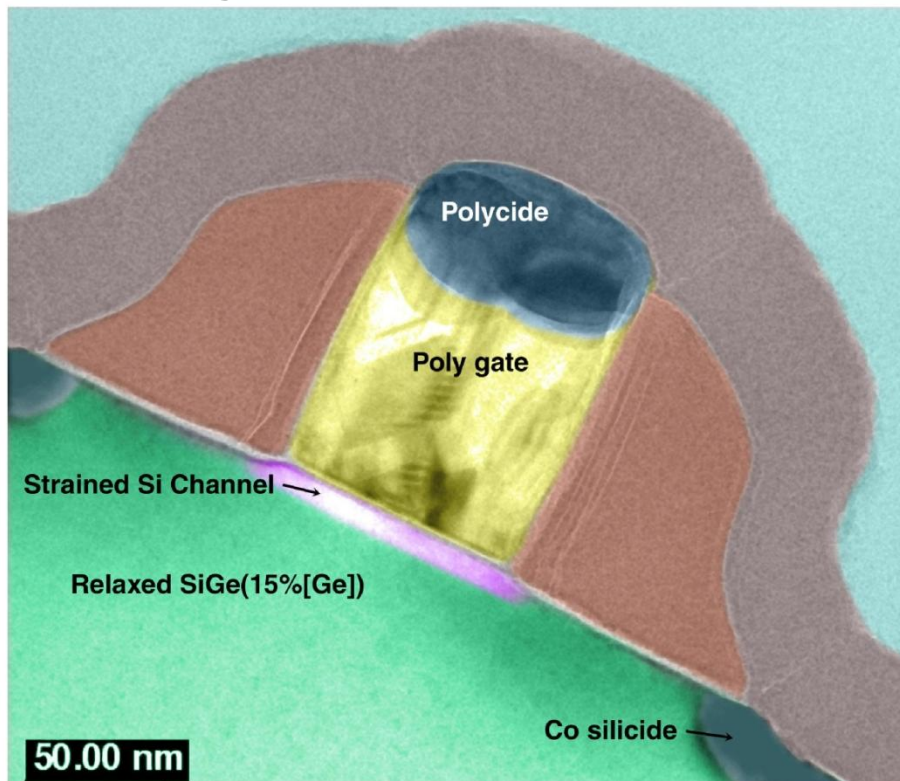
10  $\mu\text{m}$    
1  $\mu\text{m}$    
0.1  $\mu\text{m}$  



# Neue Materialien

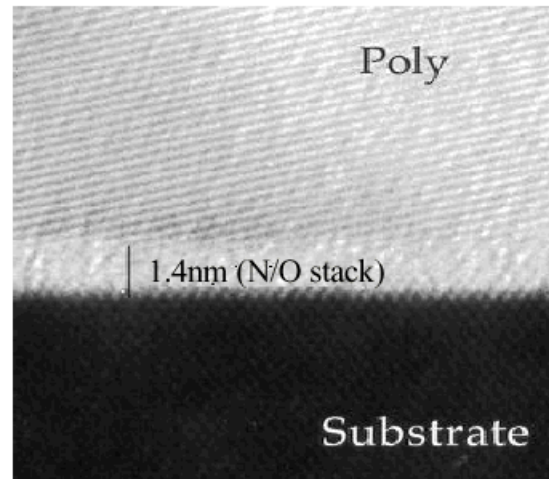
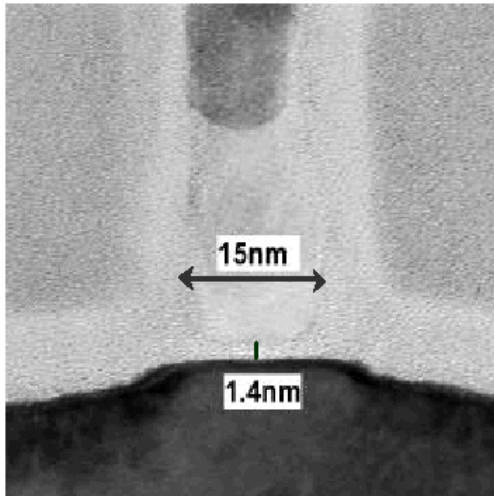
- 2001 veröffentlicht IBM einen nMOSFET mit einem Kanal aus verspanntem Silizium, der zu einer Strom-Verbesserung von 35% führt
- daneben: silizidierte Kontakte, nitridiertes Gate-Oxid

IBM 2001: nMOSFET-Kanal von 67nm  
Länge aus verspanntem Silizium

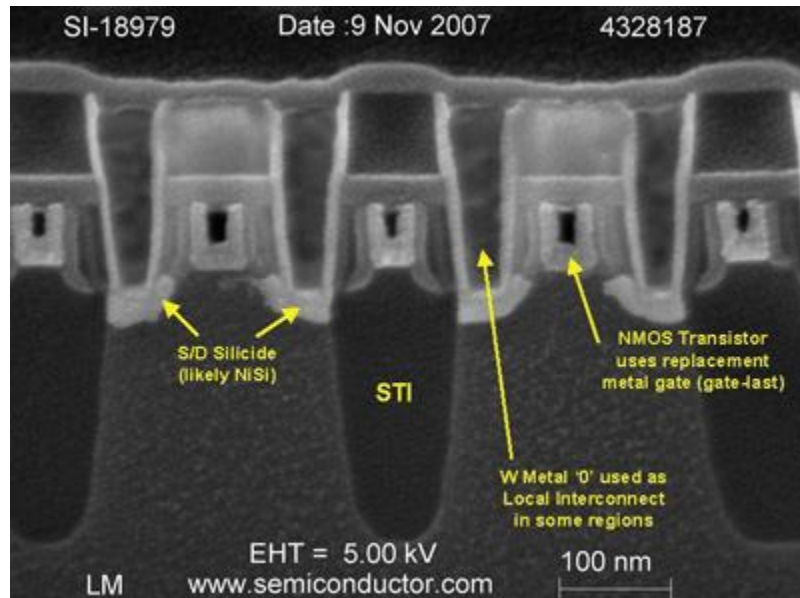




# In der Massenfertigung...



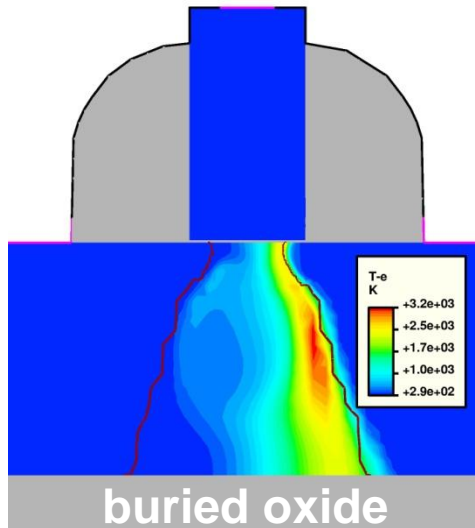
- AMD Dresden 2007
- Bulk-Technologie
- 15nm Gate-Länge
- 1.4nm Gate NO-Stack
- Globaler Strain durch Nitrid Cap-Layer



- INTEL 2008
- 45nm Bulk-Technologie
- High-K Oxid mit  $\text{HfO}_2$ -Schicht
- Metall-Gates
- Silizidierte Source und Drain
- Globaler Strain durch Nitrid Cap-Layer

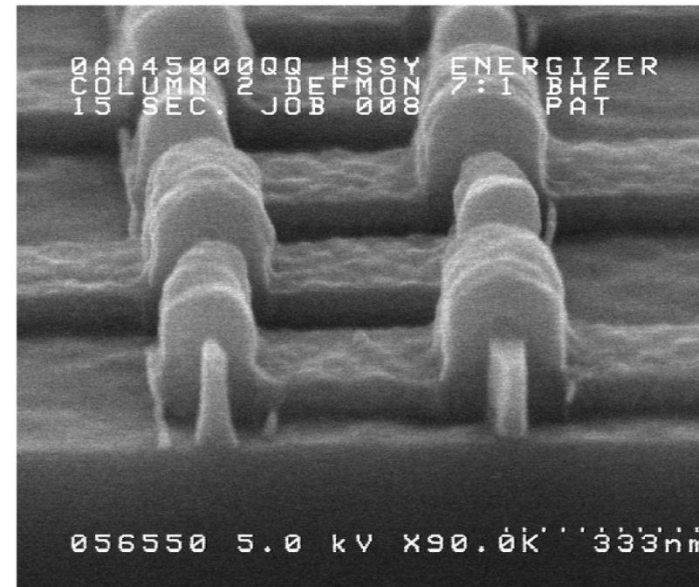
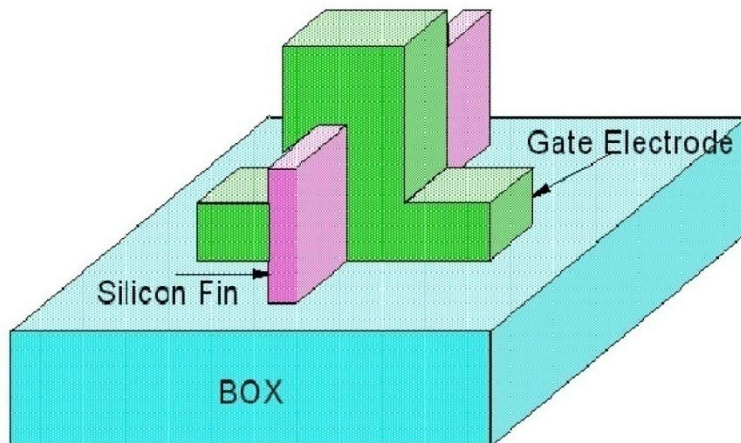
INTEL Penryn, 45nm, picture from [www.eetimes.com](http://www.eetimes.com)

# Die nähere Zukunft



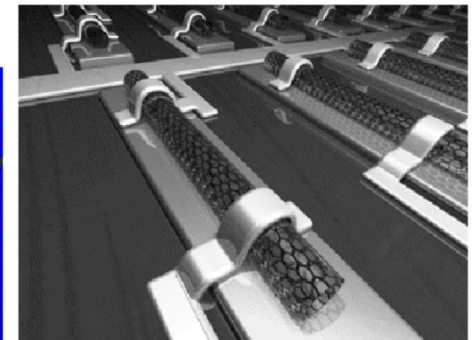
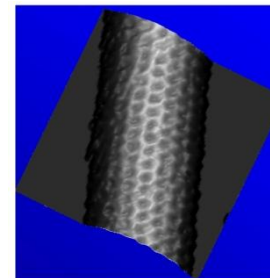
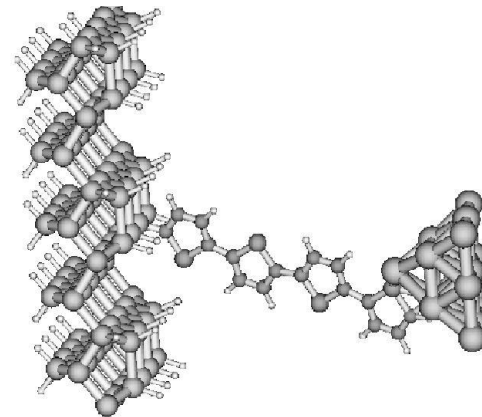
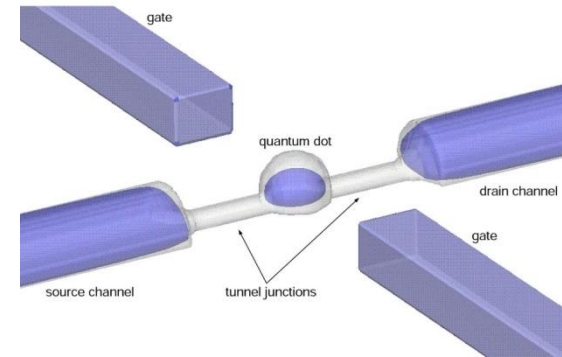
- SOI (Silicon-On-Insulator) setzt sich immer mehr durch
- Entscheidende Verbesserung bringen Double-Gate-Transistoren (kleinerer Off-Strom, grösserer On-Strom)
- DG-Transistoren können als FinFETs realisiert werden

## FinFET Structure

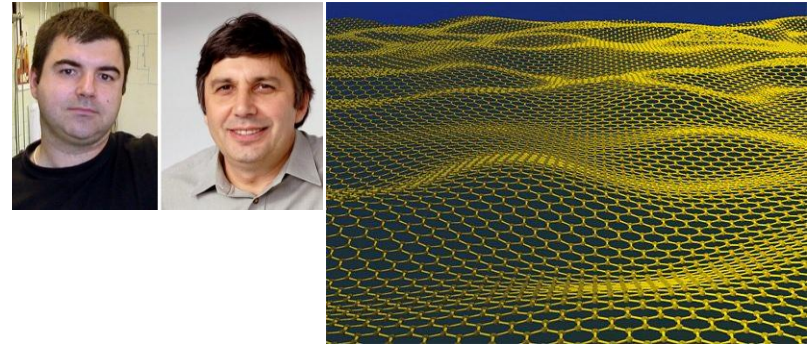
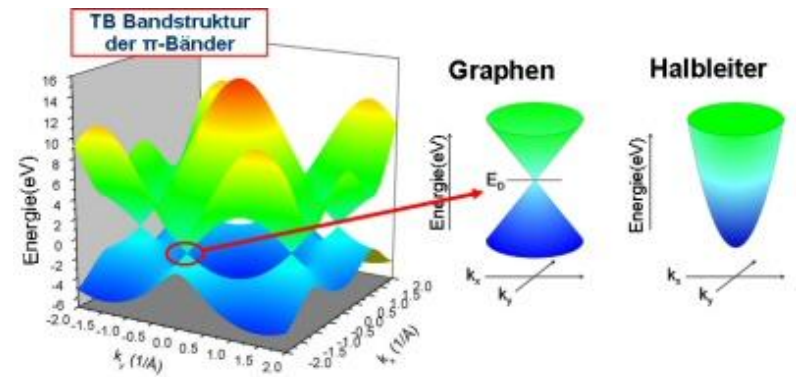


# Die ferne(re) Zukunft (“post-CMOS”)

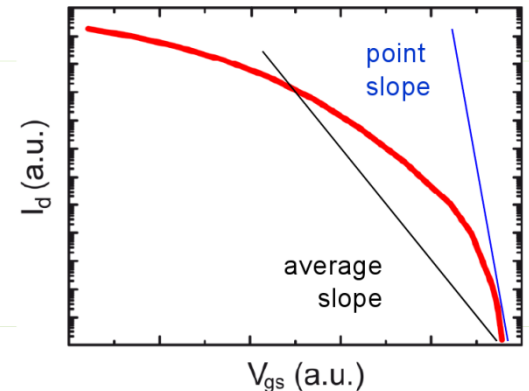
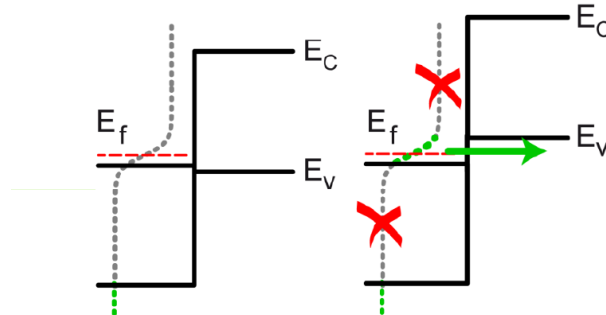
- Halbleiter-basierte Nanoelektronik?
- z.B. Einzel-Elektronen-Transistoren, Einzel-Elektronen-Speicher
- Toleranz-Probleme, ...
  
- Molekulare Nanoelektronik?
- z.B. Oligo-thiophen, C60
- Gigantischer Widerstand, ...
  
- Kohlenstoff-Nanoröhren (Carbon Nanotubes (CNTs))?
- Sehr hohe theoretische Beweglichkeit, aber ...



- Graphene (Nobelpreis 2010)
- Monoschicht Graphit, sehr hohe theoretische Beweglichkeit, aber kein Gap
- Für FET: Nano-ribbons oder Multilayer, ...



- Tunnel-FET
- Wunsch: steile Anstiege und extrem kleine Off-Ströme (“green transistor”)





# 1. Die Kristallstruktur der Festkörper

## 1.1 Halbleitermaterialien (N1.1)

## 1.2 Typen von Festkörpern (N1.2)

## 1.3 Raumgitter (N1.3)

- Elementarzelle
- Gitterarten
- Kristallebenen und Millersche Indizes
- Reziprokes Gitter\*
- Diamantstruktur

## 1.4 Atomare Bindung (N1.4)

## 1.5 Störstellen und Verunreinigungen (N1.5)

# 1.1 Halbleitmaterialien

Table 1.1 | A portion of the periodic table

III	IV	V
B	C	
Al	Si	P
Ga	Ge	As
In		Sb

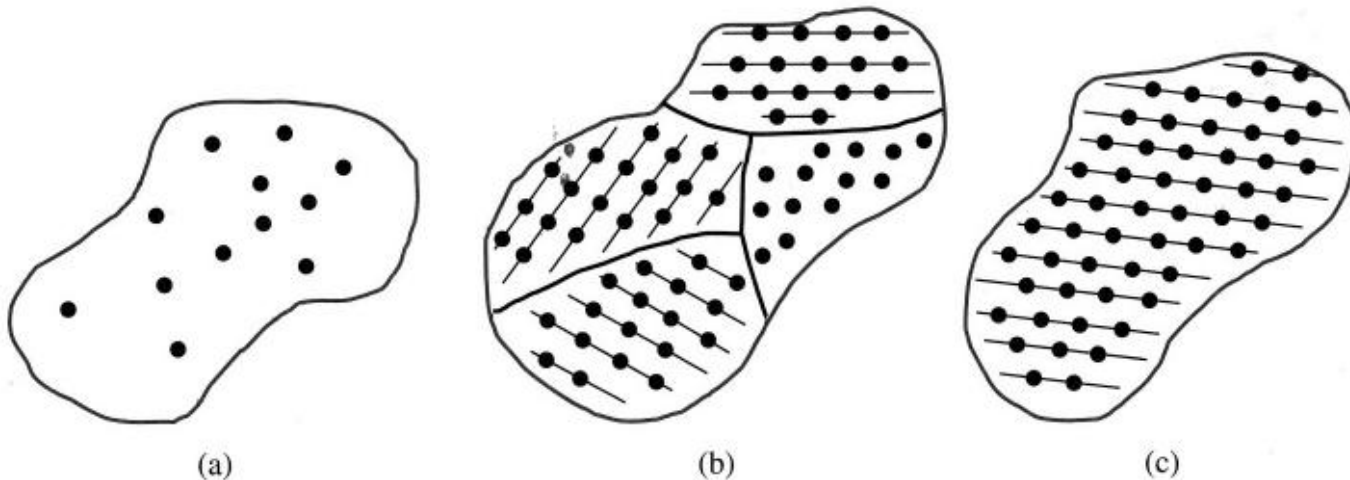
Table 1.2 | A list of some semiconductor materials

Elemental semiconductors	
Si	Silicon
Ge	Germanium
Compound semiconductors	
AlP	Aluminum phosphide
AlAs	Aluminum arsenide
GaP	Gallium phosphide
GaAs	Gallium arsenide
InP	Indium phosphide

- Elementhalbleiter (C, Si, Ge)
- Verbindungshalbleiter (binär, ternär, quarternär)
- III-V, II-VI, IV-VI

# 1.2 Typen von Festkörpern

- **Festkörper:** Materie im festen Zustand. Festkörper können nach dem Ordnungszustand ihrer Strukturbausteine (Atome, Ionen, Moleküle) klassifiziert werden:
- **Einkristallin:** Aufbau in Form eines einzigen Kristalls (Fernordnung)
- **Polykristallin:** Aufbau aus einer Vielzahl von kleinen Einkristallen (Kristallite), deren Orientierung statistisch verteilt ist
- **Amorphes Material:** Festkörper ohne Fernordnung der Strukturbausteine (nur Nahordnung), wie Legierungen, Gläser,...



Schematische Darstellung dreier allgemeiner Kristalltypen: (a) amorph, (b) polykristallin, (c) einkristallin.



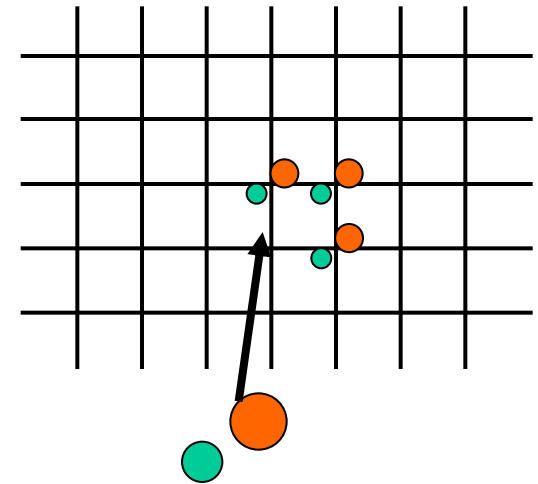
# 1.3 Raumgitter

Im Gegensatz zum strukturlosen Zustand nennen wir einen festen Körper mit *dreidimensionaler periodischer Anordnung* der Atome oder Ionen einen **Kristall**.

**Kristallgitter**: Die periodische Anordnung der Atome, Ionen, Moleküle.

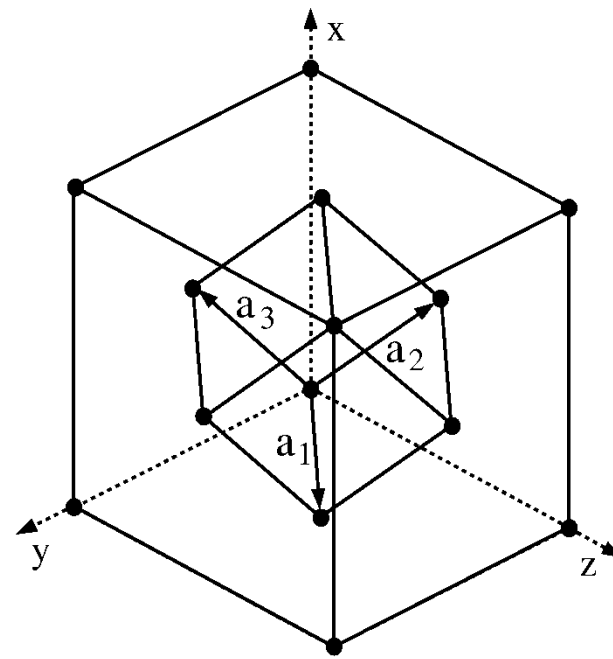
**Raumgitter**: Die *mathematische* Abstraktion des Kristallgitters, wobei die Art der Atome/ Moleküle an den Gitterpunkten vernachlässigt ist.

**Basis**: Gruppe von Atomen oder Molekülen, die jedem Gitterpunkt zugeordnet ist.



**Raumgitter + Basis = Kristallstruktur**

# Elementarzelle



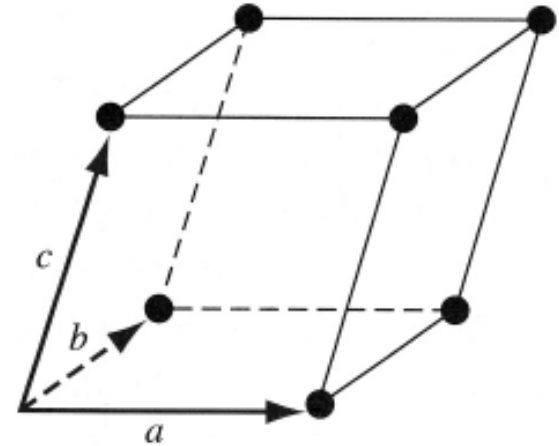
**Translation. Räumliche Periodizität. Elementarzelle.** Das periodische Raumgitter kann man sich durch räumliche Aneinanderreihung (Translation) von *Elementarzellen* aufgebaut denken. Die Elementarzelle ist der von drei nicht-koplanaren Vektoren  $\mathbf{a}_1$ ,  $\mathbf{a}_2$ ,  $\mathbf{a}_3$ , den Grund- oder Basisvektoren, aufgespannte Raum mit dem Volumen  $V = [ \mathbf{a}_1 \cdot (\mathbf{a}_2 \times \mathbf{a}_3) ]$ .

*Translation* heisst, dass das Gitter bei einer Verschiebung von  $\mathbf{r}$  nach  $\mathbf{r}'$  auf sich selbst abgebildet wird. Dazu muss gelten:

$$\mathbf{r}' = \mathbf{r} + n_1 \mathbf{a}_1 + n_2 \mathbf{a}_2 + n_3 \mathbf{a}_3 = \mathbf{r} + \mathbf{T} \text{ mit ganzzahligen } n_i .$$

Das Gitter kann durch Translation von Elementarzellen aufgebaut werden.

# Primitive Elementarzelle



Beispiel einer primitiven Elementarzelle

**Primitive Elementarzelle.** Diejenige Elementarzelle, die bei gegebener Gitterstruktur das kleinstmögliche Volumen hat.

Die primitive Elementarzelle enthält nur einen Gitterpunkt.

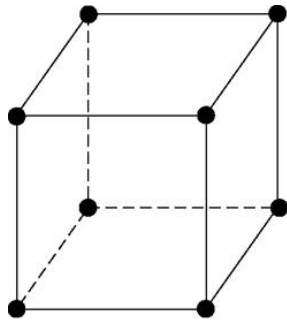
Die im obigen Bild gezeigten Gittervektoren  $\mathbf{a}$ ,  $\mathbf{b}$ , und  $\mathbf{c}$  spannen eine primitive Elementarzelle auf. Die acht in den Ecken des Parallelepipeds befindlichen Gitterpunkte werden jeweils von den acht angrenzenden Zellen geteilt. ( $8 \times 1/8 = 1$ ).

Die kleinsten Abstände  $|\mathbf{a}|, |\mathbf{b}|, |\mathbf{c}|$  heissen Gitterkonstanten  $a_0, b_0, c_0$ .

# Gitterarten

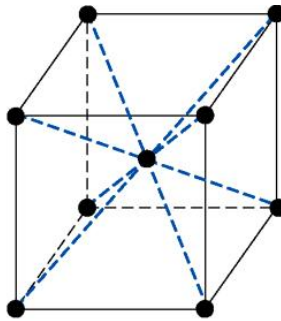
**Primitives Gitter.** Alle Gitterpunkte liegen auf Eckpunkten der Elementarzelle.

**Translationsgitter.** In der primitiven Einheitszelle befinden sich die Gitterpunkte in den Ecken des Parallelepipeds. Weitere Punkte können sich in den Flächen oder im Volumen befinden.



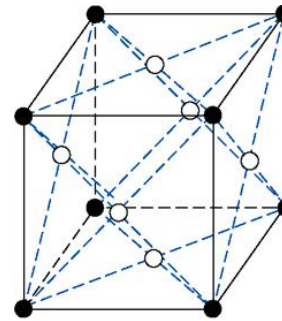
(a)

primitiv



(b)

raumzentriert  
englisch: *bcc*



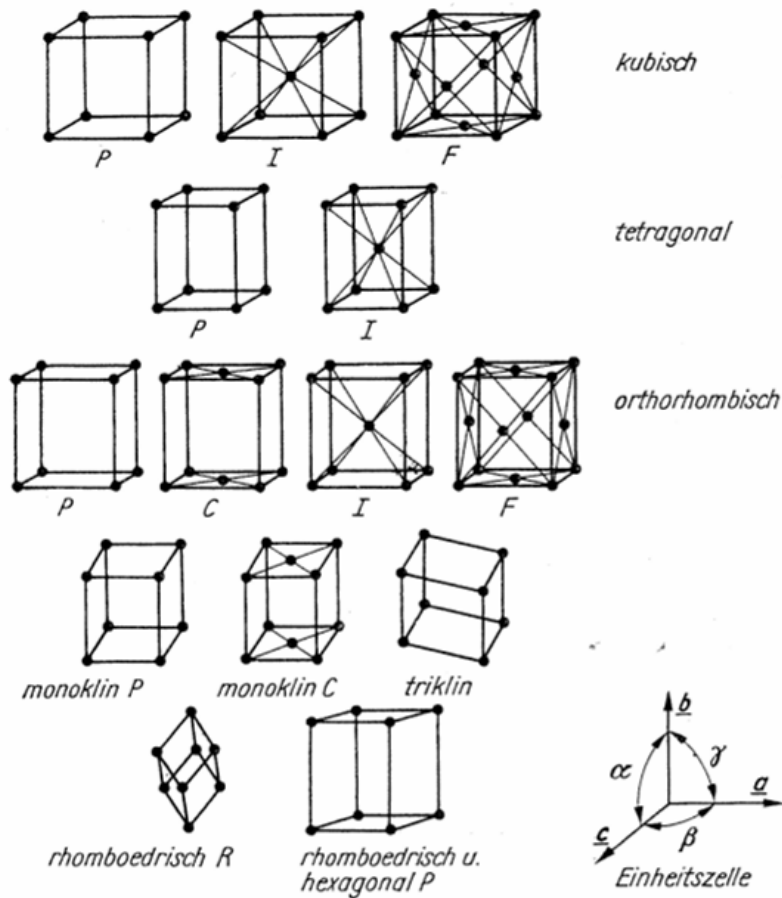
(c)

flächenzentriert  
englisch: *fcc*

**Kristallklassen.** Durch sorgfältige Symmetriebetrachtung findet man sieben verschiedene Systeme von Basisvektoren, denen sieben Kristallklassen entsprechen.

Die Gesamtheit der möglichen 14 primitiven, flächen- oder raumzentrierten Gitter heisst Translations- oder **Bravais-Gitter**. Man unterteilt sie durch Angabe der speziellen Achsen  $\mathbf{a}$  und Winkel  $\alpha, \beta, \gamma$ .

# Kristallklassen



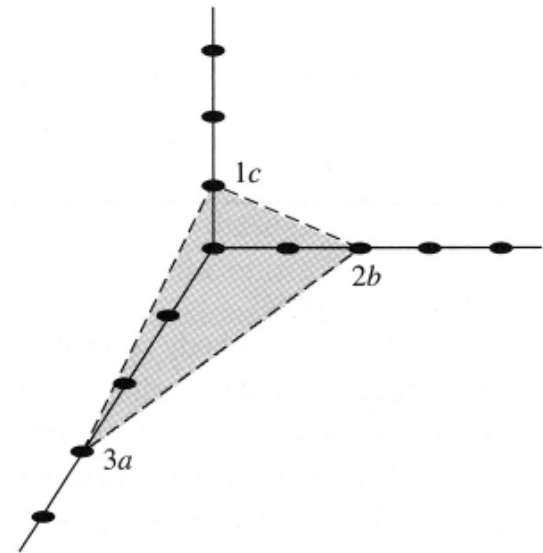
	Anzahl der Gitter im Symbol	Gittersymbol	Beträge der Einheitsvektoren	Achsenwinkel
triklin	1	<i>P</i>	<i>a, b, c</i> beliebig	beliebig $90^\circ$
monoklin	2	<i>P, C</i>	<i>a, b, c</i> beliebig	$\alpha = \gamma = 90^\circ \neq \beta$
rhombisch	4	<i>P, C, I, F</i>	<i>a, b, c</i> beliebig	$\alpha = \beta = \gamma = 90^\circ$
tetragonal	2	<i>P, I</i>	$a = b, c$ beliebig	$\alpha = \beta = \gamma = 90^\circ$
rhomboedrisch (trigonal)	1	<i>R</i>	$a = b = c$	$\alpha = \beta = \gamma < 120^\circ, 90^\circ$
hexagonal	1	<i>P</i>	$a = b, c$ beliebig	$\alpha = \beta = \gamma = 90^\circ$
kubisch	3	<i>P</i> oder <i>sc</i> <i>I</i> oder <i>bcc</i> <i>F</i> oder <i>fcc</i>	$a = b = c$	$\alpha = \beta = \gamma = 90^\circ$

**1.3.3 Kristallebenen und Millersche Indizes.** Eine Kristallebene ist durch drei nicht auf einer Geraden liegende Punkte gekennzeichnet. Dazu wählt man Punkte auf den Kristallachsen. Als zweckmässig haben sich **Millersche Indizes** erwiesen: Um sie zu bestimmen, drückt man

1. die Schnittpunkte der Ebene mit den Achsen  $\mathbf{a}_1$ ,  $\mathbf{a}_2$ ,  $\mathbf{a}_3$  in Einheiten der Gitterkonstanten aus und bildet
2. die Kehrwerte der Zahlen. Dann multipliziert man mit dem kleinsten gemeinsamen Nenner und schreibt das Ganze in Klammern  $(hkl)$ .

Ferner gilt:

- Schneidet eine Ebene eine Achse auf der negativen Seite des Ursprungs, so bekommt der zugehörige Index einen oberen Querstrich.
- Liegt ein Schnittpunkt im Unendlichen, so ist der zugehörige Index 0.
- Symmetrische Gleichwertigkeit  $\{hkl\}$



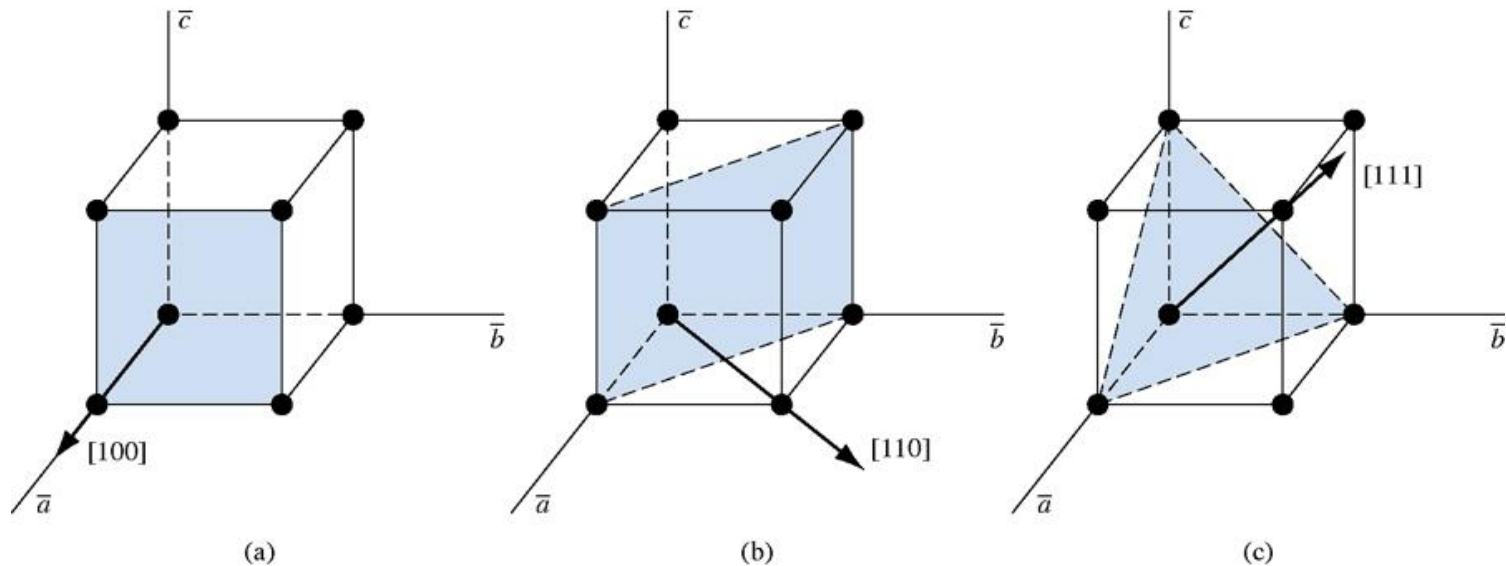
Beispiel: Eine Ebene hat die Schnittpunkte 3,2,1. Dann lauten die Kehrwerte  $1/3$ ,  $1/2$ , 1, und die Millerschen Indizes sind  $(2,3,6)$ .

**Richtungsindizes.** Die Indizes einer *Richtung* im Kristall drücken wir durch diejenige Gruppe kleinster ganzer Zahlen aus, die das gleiche Verhältnis untereinander haben wie die Komponenten des Vektors in der betreffenden Richtung, bezogen auf die Kristallachse.

Wir schreiben diese Indizes in eckigen Klammern:  $[hkl]$ .

Im kubischen Kristall steht die Richtung  $[hkl]$  stets senkrecht zur Ebene  $(hkl)$  gleicher Indizes.

Beispiel: Im kubischen Kristall weist die x- Achse in die  $[100]$ -Richtung und die y-Achse in die  $[010]$ - Richtung.



3 Kristallebenen und Kristallrichtungen: (a)  $(100)$ -Ebene und  $[100]$ -Richtung, (b)  $(110)$ -Ebene und  $[110]$ -Richtung, (c)  $(111)$ -Ebene und  $[111]$ -Richtung.

Allgemein gilt:

$[100]$  beschreibt die spezielle Kristallrichtung 100

$\langle 100 \rangle$  beschreibt alle äquivalenten 100-Richtungen  
(Rotation der Indizes)

$(100)$  beschreibt die spezielle 100-Ebene

$\{100\}$  beschreibt alle äquivalenten 100-Ebenen



# Das reziproke Gitter

**Definition.** Zu jedem Gitterpunkt

$$\mathbf{R} = l_1 \mathbf{a}_1 + l_2 \mathbf{a}_2 + l_3 \mathbf{a}_3$$

mit den Basisvektoren  $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3$  lässt sich ein *reziproker Gittervektor*  $\mathbf{K}$

$$\mathbf{K} = h_1 \mathbf{b}_1 + h_2 \mathbf{b}_2 + h_3 \mathbf{b}_3$$

so angeben, dass

$$\mathbf{R} \cdot \mathbf{K} = 2\pi n \quad (n \text{ ganz}) .$$

Dies trifft zu für

$$\mathbf{b}_1 = 2\pi \frac{\mathbf{a}_2 \times \mathbf{a}_3}{\mathbf{a}_1 \cdot (\mathbf{a}_2 \times \mathbf{a}_3)} \quad \mathbf{b}_2 = 2\pi \frac{\mathbf{a}_3 \times \mathbf{a}_1}{\mathbf{a}_1 \cdot (\mathbf{a}_2 \times \mathbf{a}_3)} \quad \mathbf{b}_3 = 2\pi \frac{\mathbf{a}_1 \times \mathbf{a}_2}{\mathbf{a}_1 \cdot (\mathbf{a}_2 \times \mathbf{a}_3)}$$

Die Vektoren  $\mathbf{b}$  sind die Translations- oder Basisvektoren des *reziproken*,  $\mathbf{a}$  jene des *gewöhnlichen* Gitters.  $\mathbf{b}$  hat die Dimension einer reziproken Länge.

# Das reziproke Gitter....

Aus den *Basisvektoren*  $\mathbf{b}$  lässt sich das unendliche periodische reziproke Gitter aufbauen; sein Raum hat die Dimension  $(\text{Länge})^{-3}$ . Die Vektoren  $\mathbf{b}$  spannen - analog zum direkten Gitter - die *Elementarzelle* des reziproken Gitters auf mit dem Volumen

$$V_{\text{rez}} = \mathbf{b}_1 \cdot (\mathbf{b}_2 \times \mathbf{b}_3) = \frac{(2\pi)^3}{\mathbf{a}_1 \cdot (\mathbf{a}_2 \times \mathbf{a}_3)} = \frac{(2\pi)^3}{V_{\text{gew}}}$$

# Feststellungen zum reziproken Gitter

1. Zu jeder Kristallstruktur gehören das Kristallgitter und das reziproke Gitter. Beide stehen durch die oben erwähnte Vektorrelation in Beziehung zueinander.
2. Das Kristallgitter ist ein Gitter im wirklichen Raum, das reziproke Gitter dagegen ist ein Gitter im sog. Fourierraum („k-Raum“). Es gilt

$$\exp(j\mathbf{K}\cdot\mathbf{R}) = \exp(j2\pi n) = 1$$

3. Der reziproke Gittervektor steht senkrecht auf der  $(h_1 h_2 h_3)$  Ebene
4. Der Abstand  $d(l_1 l_2 l_3)$  der Ebenen  $(l_1 l_2 l_3)$  des Kristallgitters ist:

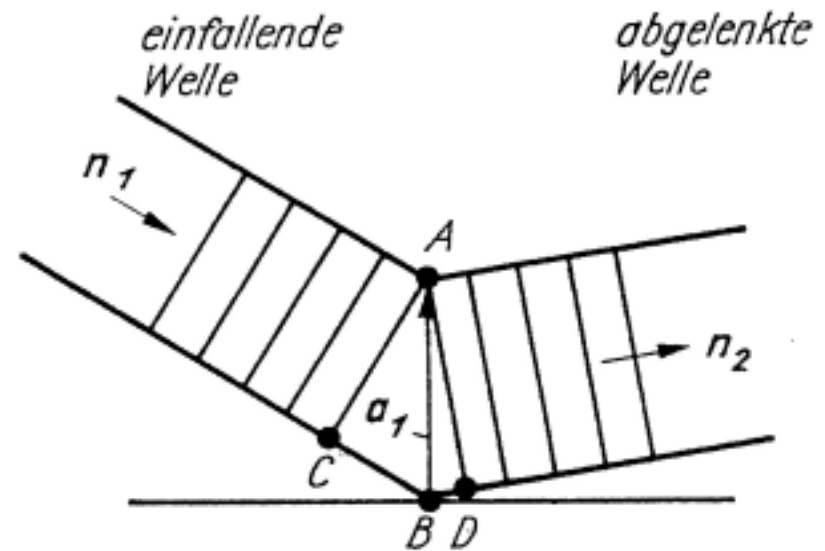
$$d(l_1 l_2 l_3) = \frac{2\pi}{|\mathbf{K}(l_1 l_2 l_3)|}$$

5. Lässt sich das wirkliche Kristallgitter durch ein Mikrofoto abbilden, so liefert das reziproke Gitter ein Beugungsbild. Das begründet seine Relevanz.

**Beugung einer Welle im Kristall.** Wir untersuchen die Beugung einer ebenen Welle (Wellenlänge  $\lambda$ ) an zwei im Abstand des Basisvektors  $\mathbf{a}_1$  befindlichen Atomen. Die einfallende Welle habe den Normaleinheitsvektor  $\mathbf{n}_1$ , die gestreute den Vektor  $\mathbf{n}_2$ .

An beiden Atomen A und B erfolgt eine Streuung, und infolge des Gangunterschieds längs beider Strecken kommt es zur teilweisen oder vollständigen Löschung der gestreuten Welle.

Die gestreute Welle hat die grösste Intensität, wenn der relative Gangunterschied  $CB+BD$  ein ganzes Vielfaches der Wellenlänge beträgt, also



$$CB + BD = -(\mathbf{a}_1 \cdot \mathbf{n}_1) + (\mathbf{a}_1 \cdot \mathbf{n}_2) = m_1 \lambda$$

Im dreidimensionalen Kristall (Basisvektoren  $\mathbf{a}_1$ ,  $\mathbf{a}_2$ ,  $\mathbf{a}_3$ ) lassen sich dann drei Gleichungen der Form

$$\begin{aligned} \mathbf{a}_1 \left( \frac{\mathbf{n}_2}{\lambda} - \frac{\mathbf{n}_1}{\lambda} \right) &= m_1 & \mathbf{a}_2 \left( \frac{\mathbf{n}_2}{\lambda} - \frac{\mathbf{n}_1}{\lambda} \right) &= m_2 \\ \mathbf{a}_3 \left( \frac{\mathbf{n}_2}{\lambda} - \frac{\mathbf{n}_1}{\lambda} \right) &= m_3 & m_1 \dots m_3 &\text{ ganz} \end{aligned}$$

aufschreiben (Laue-Bedingung).

Das entspricht völlig der Bedingung für den reziproken Gittervektor  $\mathbf{k} = \mathbf{K}/2\pi$ , so dass

$$\frac{\mathbf{n}_2}{\lambda} - \frac{\mathbf{n}_1}{\lambda} = \mathbf{k} = \frac{\mathbf{K}}{2\pi}$$

ein Vektor des reziproken Gitters sein muss.

Ordnet man der ein- und auslaufenden Welle die *Wellenvektoren*

$$\mathbf{K}_1 = \frac{2\pi}{\lambda} \mathbf{n}_1 \qquad \mathbf{K}_2 = \frac{2\pi}{\lambda} \mathbf{n}_2$$

zu, so lautet die Bedingung

$$\mathbf{K}_2 - \mathbf{K}_1 = \mathbf{K}$$

oder

$$\mathbf{K}_2 = \mathbf{K}_1 + \mathbf{K}$$

Der Wellenvektor der gestreuten Welle ist gleich dem ursprünglichen Wellenvektor zuzüglich des reziproken Gittervektors  $\mathbf{K}$ .

Ist  $\mathbf{K}_2 - \mathbf{K}_1 = \mathbf{K}$  gleich einem reziproken Gittervektor, so sind die Laueschen Gleichungen der Wellenbeugung erfüllt.

Man nennt diese Gleichung auch **Erhaltungssatz** für die Wellenvektoren.

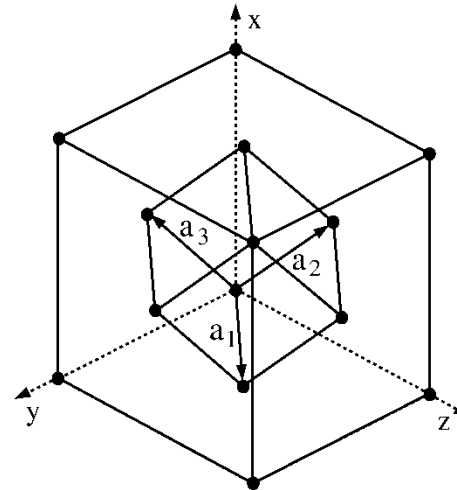
## Beispiel: Kubisch flächenzentriertes (fcc) direktes Gitter

Die Basisvektoren des fcc-Gitters mit Einheitsvektoren in Richtung der Würfelkanten lauten

$$\mathbf{a}_1 = \frac{a_0}{2} (\mathbf{e}_y + \mathbf{e}_z)$$

$$\mathbf{a}_2 = \frac{a_0}{2} (\mathbf{e}_x + \mathbf{e}_z)$$

$$\mathbf{a}_3 = \frac{a_0}{2} (\mathbf{e}_x + \mathbf{e}_y)$$



Über die Definition des reziproken Gitters ergeben sich dessen Basisvektoren zu

$$\mathbf{b}_1 = \frac{2\pi}{a_0} (-\mathbf{e}_x + \mathbf{e}_y + \mathbf{e}_z) \quad \mathbf{b}_2 = \frac{2\pi}{a_0} (\mathbf{e}_x - \mathbf{e}_y + \mathbf{e}_z) \quad \mathbf{b}_3 = \frac{2\pi}{a_0} (\mathbf{e}_x + \mathbf{e}_y - \mathbf{e}_z)$$

Der allgemeine reziproke Gittervektor  $\mathbf{K}$  für allgemeine Zahlen  $hkl$  wird dann

$$\mathbf{K} = \frac{2\pi}{a_0} [(-h + k + l)\mathbf{e}_x + (h - k + l)\mathbf{e}_y + (h + k - l)\mathbf{e}_z]$$

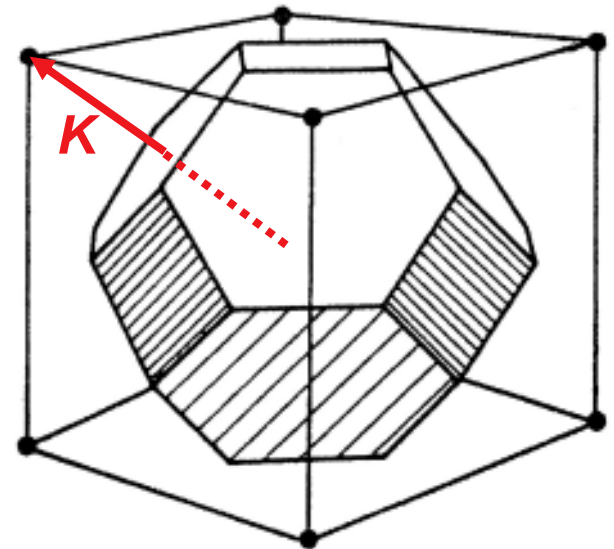
Mit den kleinsten, nichtverschwindenden acht Werten  $\mathbf{K} = \frac{2\pi}{a_0} [\pm \mathbf{e}_x \pm \mathbf{e}_y \pm \mathbf{e}_z]$

Das reziproke Gitter des fcc-Gitters ist also ein bcc-Gitter! Als primitive Elementarzelle wählt man eine *symmetrische* Zelle, die wie folgt konstruiert wird: Auf den Mitten der acht  $\mathbf{K}$ -Vektoren errichtet man senkrechte Ebenen. Sie bilden zunächst einen Oktaeder, der die primitive Elementarzelle begrenzt. Die Ecken des Oktaeders werden von Ebenen abgeschnitten, die senkrecht auf den Mitten der sechs reziproken Verschiebungen

$$\mathbf{v}_1 = \frac{2\pi}{a_0} (\pm 2\mathbf{e}_x) \quad \mathbf{v}_2 = \frac{2\pi}{a_0} (\pm 2\mathbf{e}_y) \quad \mathbf{v}_3 = \frac{2\pi}{a_0} (\pm 2\mathbf{e}_z)$$

liegen.

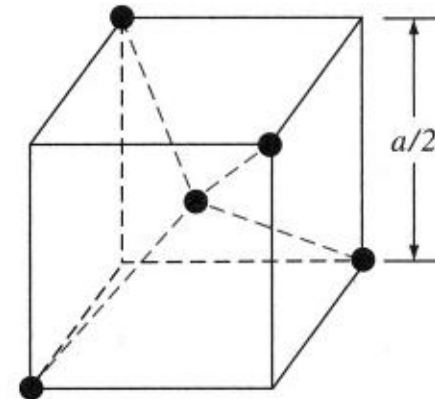
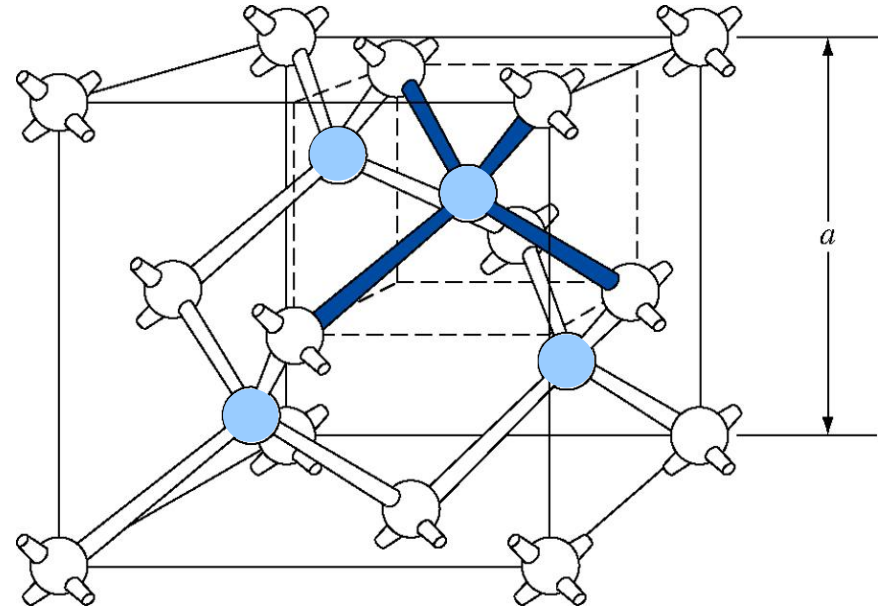
Die symmetrische Zelle hat also die Form eines gekappten Oktaeders. Sein Volumen ist die sog. **1. Brillouin-Zone**.





# 1.3.4 Die Diamantstruktur

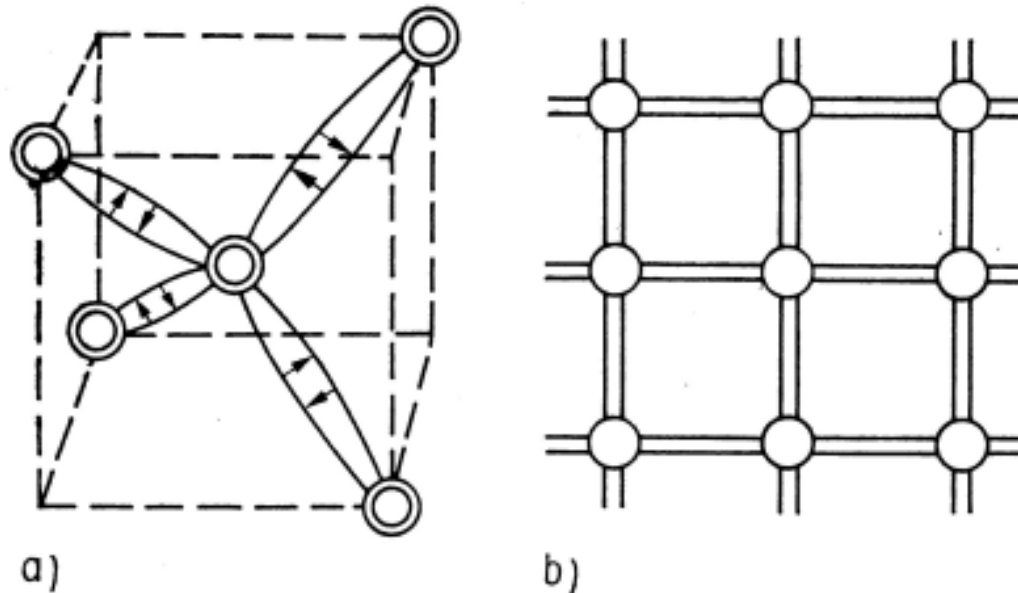
**Diamantstruktur.** Eine Reihe wichtiger Halbleitermaterialien (Si, Ge) kristallisiert in der Diamantstruktur. Das Raumgitter ist ein fcc-Gitter. Die *Basis* von Si besteht aus zwei Atomen. Diese Basis ist auf dem fcc-Gitter periodisch fortgesetzt. Deshalb besteht das Kristallgitter aus zwei ineinander verschachtelten fcc-Gittern, die entlang der Raumdiagonalen der Elementarzelle (Kantenlänge  $a$ ) um  $1/4$  ihrer Länge verschoben sind. In einer Elementarzelle befinden sich acht Atome auf den Würfecken, sechs auf den Flächenmitten und vier im Innern. Damit hat die kubische Zelle insgesamt  $8 \cdot 1/8 + 6 \cdot 1/2 + 4 = 8$  Atome und jedes von ihnen vier nächste und zwölf übernächste Nachbarn.



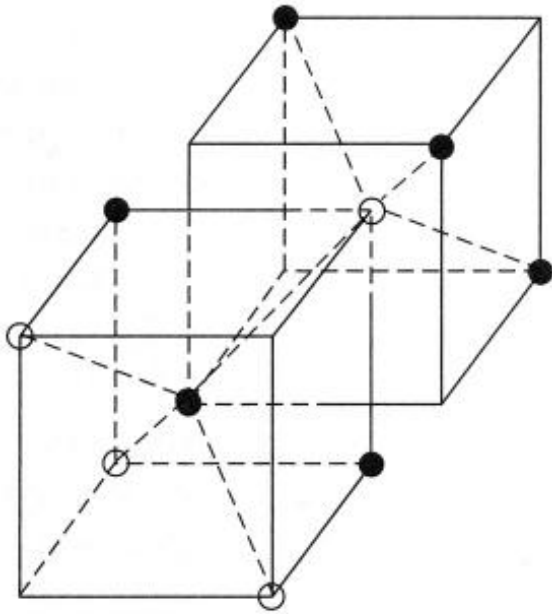
Die tetraedrische Konfiguration der nächsten Nachbarn im Diamantgitter.

# 1.4 Atomare Bindung

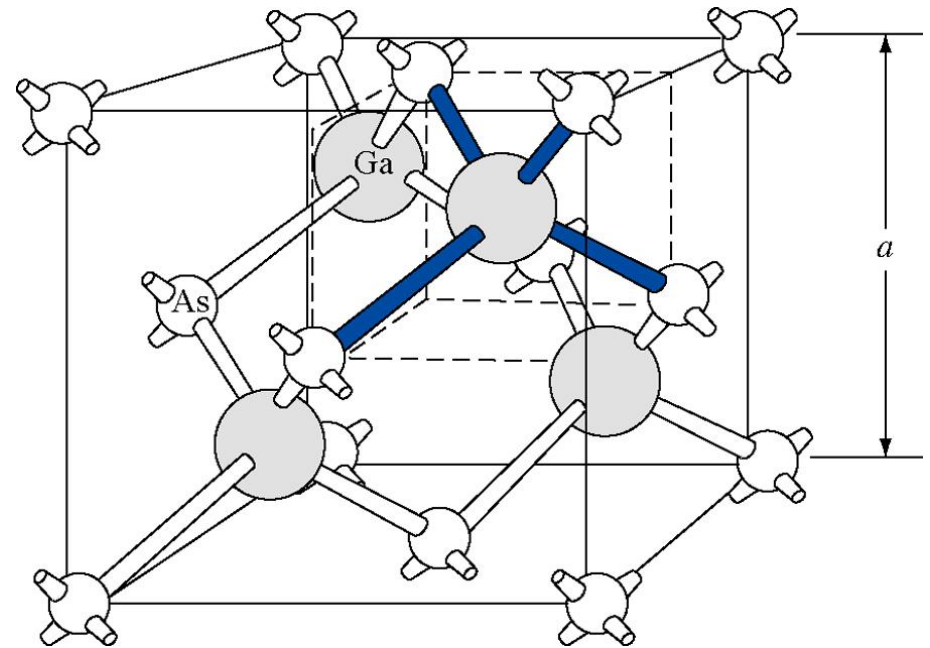
**Eigenleitung:** Die Elementhalbleiter Si und Ge sind chemisch vierwertig, so dass jedes Atom symmetrisch von vier Nachbaratomen in Form eines Tetraeders umgeben ist. Bei Si ist der Atomrumpf (Atomkern und äussere Schalen 1s, 2s, 2p) über Valenzelektronen (je zwei 3s und 3p) mit dem Nachbaratom verbunden.



(a) Silizium-Valenzelektronen, (b) kovalente Bindung im Siliziumkristall.



Die tetraedrische Konfiguration der nächsten Nachbarn in der Zinkblendestruktur.

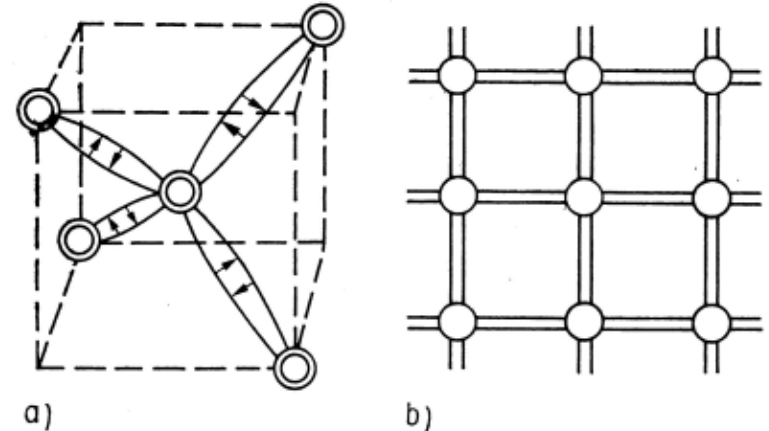


**Zinkblendestruktur.** In dieser Struktur kristallisiert eine Reihe binärer Halbleiter aus Verbindungen der IV-VI-, III-V- und II-VI-Gruppen des periodischen Systems (z.B. SiC, GaAs, GaP, InSb, CdS). Es besteht aus zwei Subgittern, z.B. bei GaAs aus einem Ga-Subgitter und einem As-Subgitter. Die Zinkblendestruktur ergibt sich aus der Diamantstruktur, wenn die Ga-Atome auf das eine, und die As-Atome auf das andere fcc-Gitter verteilt werden.

# Eigenleitung

Im absoluten Nullpunkt der Temperatur stehen infolge der abgesättigten kovalenten Bindung keine beweglichen Elektronen zur Verfügung. Damit gibt es unter der Wirkung eines äusseren Feldes auch keinen Stromtransport.

Durch zugeführte Energie geraten die Atome in Schwingungen und Valenzelektronen können sich als freie Ladungsträger im Gitter mit statistischer Geschwindigkeitsverteilung bewegen, bis sie wieder in einen leeren Gitterplatz zurückfallen.



(a) Wasserstoff-Valenzelektronen, (b) kovalente Bindung im Wasserstoffmolekül.

# Eigenleitung.....

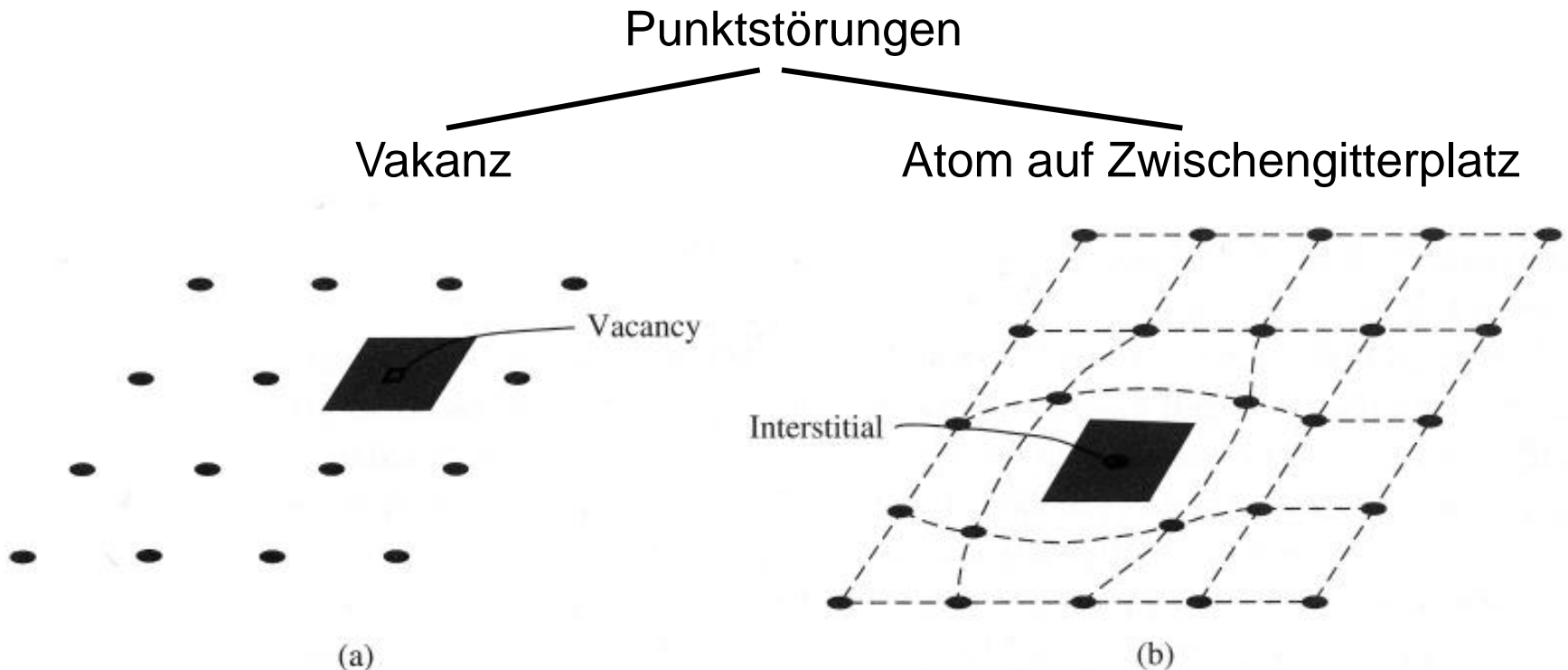
Durch statistische Auffüllung und Neuerzeugung wechseln auch die Gitterlücken ihren Platz wie Elektronen, nur in umgekehrter Richtung.

Das Verhalten des lückenbehafteten Bindungssystems lässt sich phänomenologisch durch ein System beschreiben, in dem die Lücken durch positive Ladungsträger - die sogenannten **Löcher** - repräsentiert sind, deren Zahl gleich der Anzahl der beweglichen Elektronen ist.

Elektronische Halbleiter haben also neben der Elektronenleitung eine Löcherleitung.

# 1.5 Störstellen und Verunreinigungen

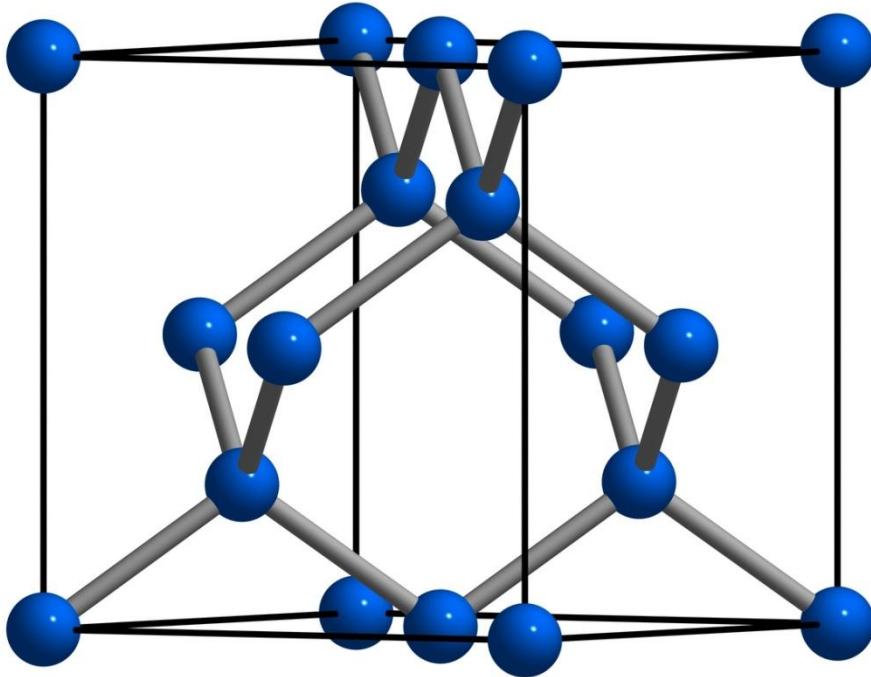
Der Realkristall weist immer Verunreinigungen und Störstellen auf. Dazu gehören Punktdefekte (Fremdatome auf Gitter- oder Zwischengitterplatz, Vakanzen, Self-Interstitials) und ausgedehnte Defekte (Versetzungen).



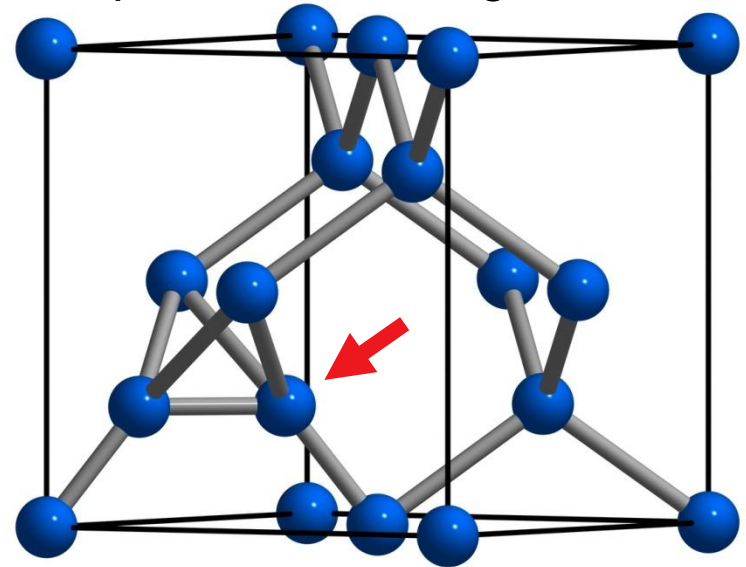
Zweidimensionale Darstellung eines einkristallinen Gitters mit Punktdefekten: (a) Vakanz, (b) Atom auf Zwischengitterplatz (Interstitial).

# Self-Interstitial

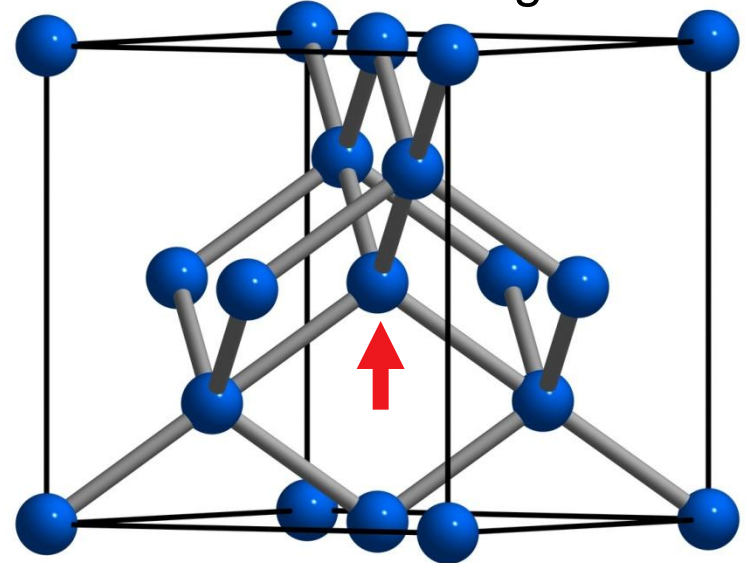
idealer Siliziumkristall



split  $\langle 110 \rangle$  Konfiguration



tetraedrische Konfiguration



# Störleitung

Im Gegensatz zu Metallen hängt die Leitfähigkeit in Halbleitern stark von den Gitterstörungen ( Abweichungen von der Gitterstruktur oder Dotieratome, die eingeführt wurden) ab. Diese Störungen können unter bestimmten Bedingungen durch Dissoziation freie Ladungsträger (Elektronen oder Löcher) in den Kristall abgeben. Man spricht von **Störleitung**.

Bei Si mit einer Atomkonzentration von  $5 \cdot 10^{22} \text{ cm}^{-3}$  würde bereits der Zusatz einer Störstelle auf  $10^{10}(!)$  Si Atome eine Verdopplung der freien Ladungsträgerdichte ergeben.

Damit wird klar, warum die Reinheitsanforderungen in der Halbleiterindustrie so herausfordernd sind.

Die Leitungseigenschaften der Halbleitermaterialien werden völlig von der Störstellenkonzentration bestimmt.





# 2. Einführung in die Quantenmechanik

2.1 Prinzipien der Quantenmechanik (N2.1)

2.2 Die Schrödinger-Gleichung (N2.2)

2.3 Anwendungen der Schrödinger-Gleichung (N2.3)

2.4 Anwendung der Wellentheorie auf Atome\* (N2.4)

# 2.1 Prinzipien der Quantenmechanik

**Energiequanten.** Der photoelektrische Effekt demonstriert eine Inkonsistenz zwischen den Aussagen der klassischen Physik und dem Experiment. Die Austrittsarbeit eines Elektrons ist von der Frequenz abhängig.

Planck hat 1900 postuliert, dass thermische Strahlung von einem heißen Körper gequantelt emittiert wird:

$$E = h\nu.$$

$$h = 6.625 \times 10^{-34} \text{ Js}$$

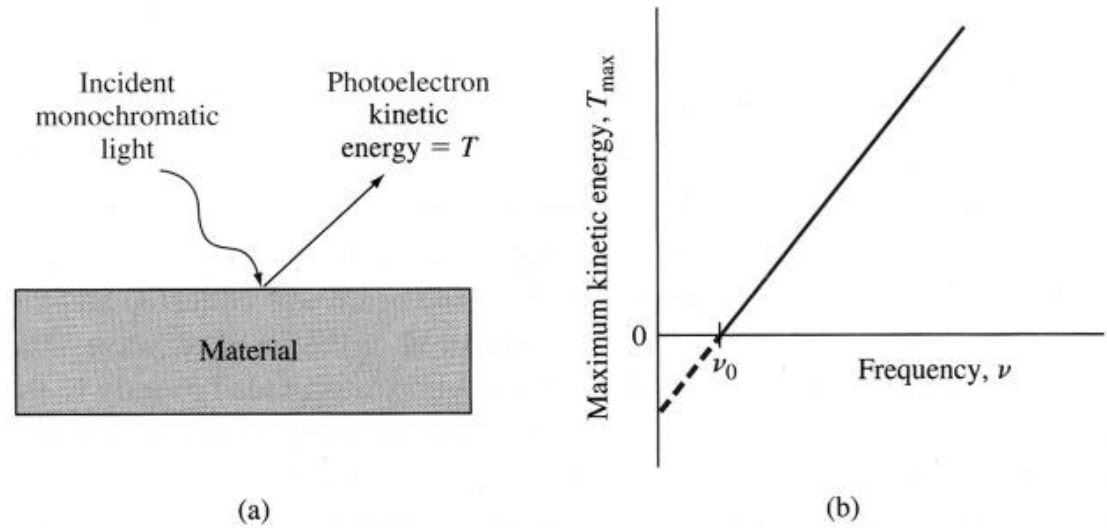
Einstein (1905) interpretiert den photoelektrischen Effekt als Resultat diskreter

**Energiepakete = Photonen.**

Austrittsarbeit = Energie, um ein Elektron vom Körper ins Vakuum zu bringen.

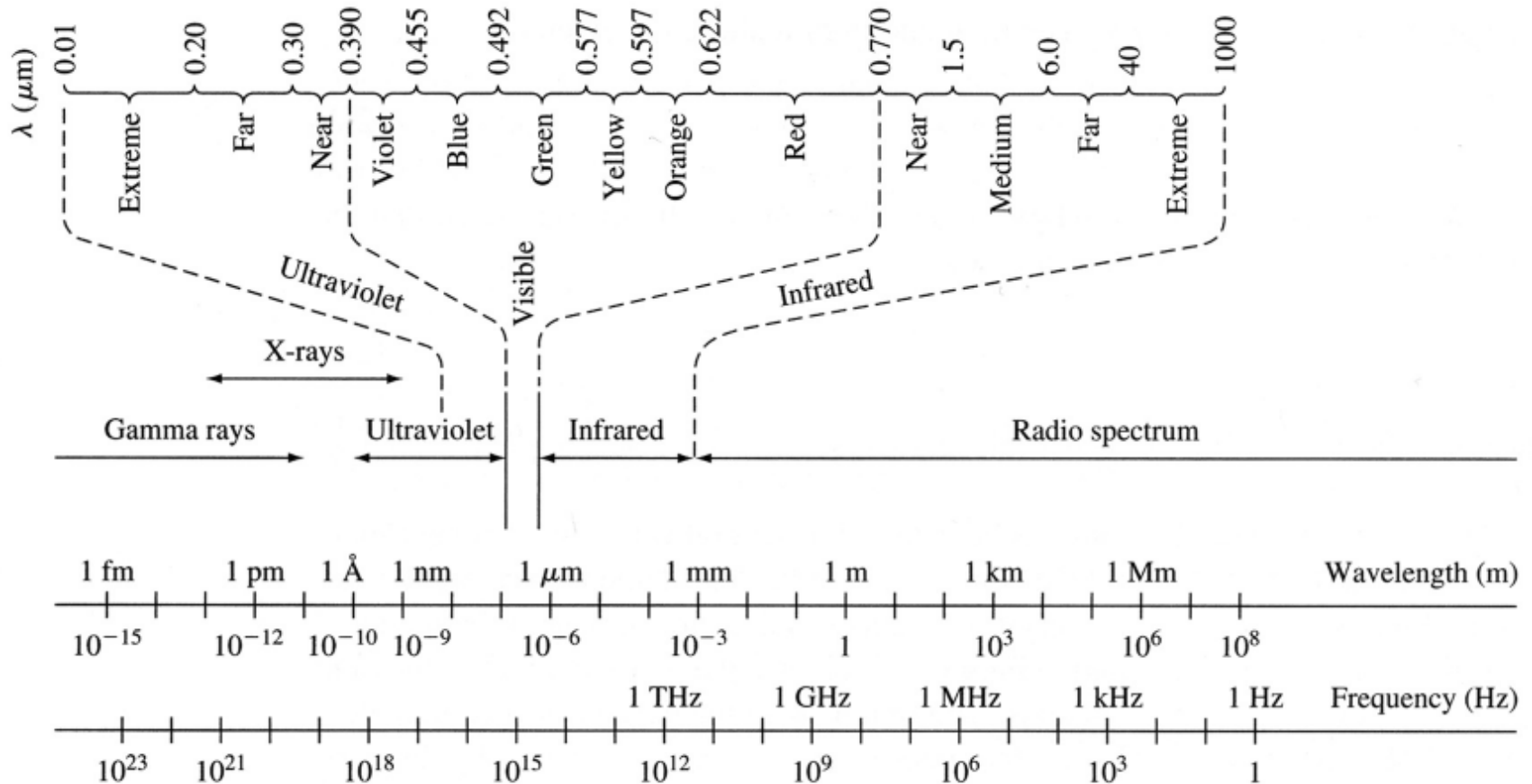
Beziehung zwischen Energie, Wellenlänge und Frequenz:  $E = h\nu = \frac{hc}{\lambda}$

⇒ **Elektromagnetische Wellen können sich wie Teilchen verhalten**



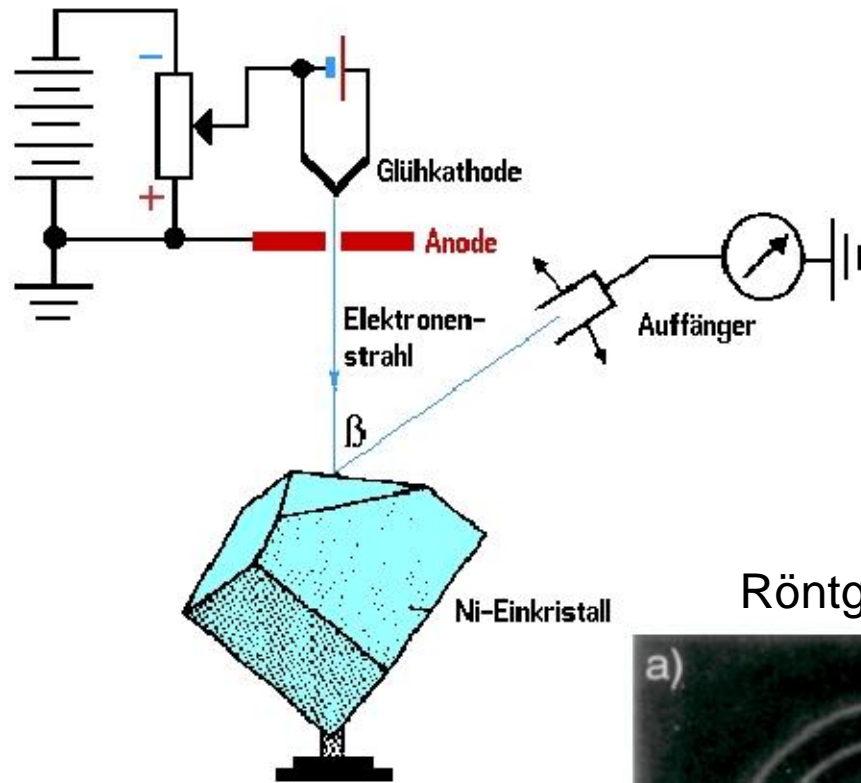
(a) Photoelektrischer Effekt, (b) maximale kinetische Energie als Funktion der Lichtfrequenz.

**Der Welle-Teilchen-Dualismus.** Louis de Broglie postuliert 1924, dass sich Teilchen wie Wellen verhalten können. Da der Impuls eines Photons  $p=h/\lambda$  ist, kann die Wellenlänge eines Teilchens analog dazu als  $\lambda=h/p$  definiert werden.



Das elektromagnetische Frequenzspektrum.

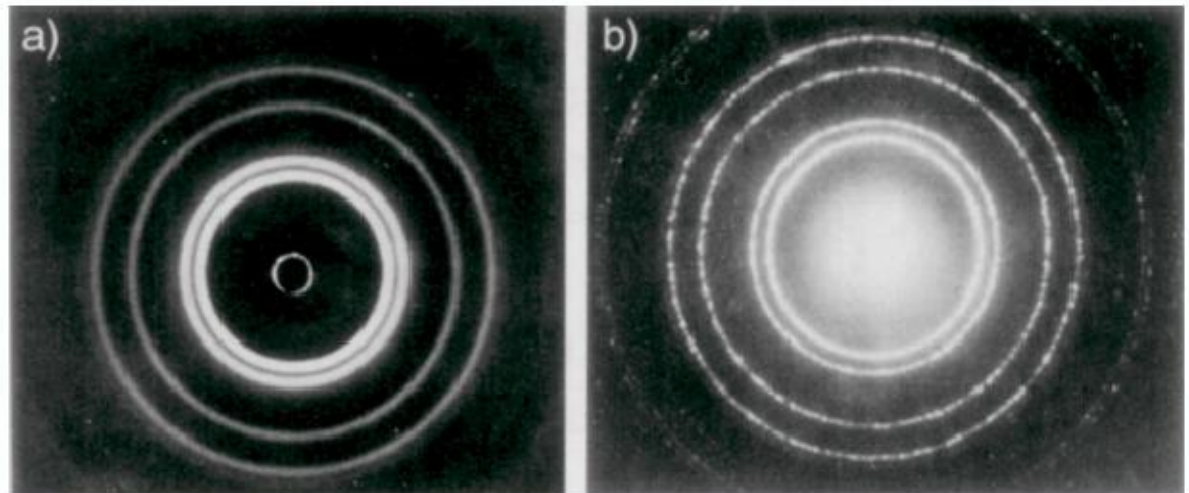
# Clinton Davisson und Lester Germer (Bell Labs) beweisen 1927 die Wellennatur von Elektronen.



- Elektronen wurden senkrecht auf die Oberfläche eines Nickel-Einkristalls geschossen und die Intensität der reflektierten Strahlen als Funktion des Streuwinkels gemessen.
- Es ergaben sich ausgeprägte Maxima, besonders für eine Beschleunigungsspannung von 54 V.

Röntgenstrahlen

Elektronenstrahlen



**Die Unschärfe-Relation.** Nach der Korpuskularvorstellung kann man den Zustand der Teilchen durch Angabe ihres Ortsvektors  $\mathbf{r}$  und ihres Impulses  $\mathbf{p}$  eindeutig festlegen.

Im *mikroskopischen* Bereich der Elementarteilchenabmessungen schränkt die *Heisenbergsche Unschärferelation* dies ein. Ihr zufolge sind gewisse Paare von Größen, z.B Ort  $\mathbf{r}$  und Impuls  $\mathbf{p}$  oder Energie  $E$  und Zeit  $t$ , nicht mehr gleichzeitig scharf messbar. Vielmehr gelten für das Produkt der Ungenauigkeiten die Relationen

$$\Delta x \Delta p_x \geq \hbar = h/2\pi \qquad \Delta E \Delta t \geq \hbar$$

Umso genauer der Ort bei einer Messung bestimmt wird, desto ungenauer ist der Impuls bekannt. Erfolgt bei einer Messung eine scharfe Ortsbestimmung, bleibt der Impuls völlig unbekannt.

Der quantenmechanische Zustand der Teilchen wird durch die Schrödinger-Gleichung beschrieben. Ihre Lösungen sind sogenannte *Wellenfunktionen*. Mit ihrer Hilfe können alle messbaren Größen eines Systems (Ort, Energie, Impuls, Drehimpuls, ...) berechnet werden.

## 2.2 Die Schrödinger-Gleichung

Die Wellenfunktion  $\Psi(\mathbf{r}, t)$  eines quantenmechanischen Systems, z.B. eines Elektrons, gehorcht der Gleichung

$$\frac{-\hbar^2}{2m} \Delta \Psi + V(\mathbf{r}, t) \Psi = j \hbar \frac{\partial \Psi}{\partial t}$$

Hängt die potentielle Energie  $V$  nicht von der Zeit ab, liegt der stationäre Fall vor. Mit dem Produktansatz  $\Psi(\mathbf{r}, t) = \psi(\mathbf{r}) \phi(t)$  wird die Gleichung dann *separierbar*

$$\frac{-\hbar^2}{2m} \frac{\Delta \psi}{\psi} + V(r) = E$$
$$j \hbar \frac{\partial \phi}{\partial t} \frac{1}{\phi} = E$$

bzw.  $\frac{\hbar^2}{2m} \Delta \psi + (E - V) \psi = 0$  *stationäre Schrödinger-Gleichung*

Zur Lösung werden gebraucht:

1. Der räumliche Verlauf von  $V(\mathbf{r})$
2. Aussagen über die Grösse des Gebietes, in dem sich das Elektron bewegt
3. Werte an den Rändern dieses Gebietes (Randwerte)

Für den zeitabhängigen Teil der Wellenfunktion  $\phi(t)$  ergibt sich die Lösung

$$\phi(t) = \exp\left(-j\frac{E}{\hbar} t\right)$$

Sie beschreibt einen zeitlich harmonischen Vorgang mit der (Kreis-)Frequenz

$$\omega = 2\pi f = \frac{E}{\hbar}$$

**Das kräftefreie Elektron:** Für  $V=0$  vereinfacht sich die stationäre Schrödinger-Gleichung zu

$$\left(\frac{\hbar^2}{2m}\Delta + E\right)\psi = 0 \quad \text{oder} \quad (\Delta + k^2)\psi = 0$$

mit dem Quadrat der Wellenzahl  $k^2 = k_x^2 + k_y^2 + k_z^2 = \frac{2mE}{\hbar^2}$  bzw.  $E = \frac{\hbar^2 k^2}{2m}$

Die beiden linear unabhängigen Lösungen lauten

$$\psi = \exp(\pm j\mathbf{k}\cdot\mathbf{r})$$



Die Wellenfunktion des kräftefreien Elektrons ist eine *ebene Welle*

$$\Psi = \psi(r)\phi(t) \sim \exp j(\mathbf{k} \cdot \mathbf{r} - \omega t)$$

die sich in Richtung des Wellen- oder Ausbreitungsvektors  $\mathbf{k}$  mit der (Materie-) Wellenlänge  $\lambda = 2\pi/k$  und der Frequenz  $\omega$  ausbreitet.

Energie  $E$  und Wellenzahl  $k$  hängen quadratisch zusammen:

$$E = \frac{\hbar^2 k^2}{2m}$$

Das nennt man auch *parabolische Dispersionsrelation*.

Durch Vergleich mit der klassischen kinetischen Energie  $E = p^2/2m$  erkennt man nochmal, dass in der Quantenmechanik der Zusammenhang

$$p = \hbar k$$

zwischen Impuls und Wellenzahl gilt.

# Quantenmechanische Fragestellungen

1. Welche physikalische Bedeutung hat die Wellenfunktion?
2. Wie berechnet man bestimmte physikalische Grössen (Energie, Impuls, Drehimpuls,...) mithilfe der Wellenfunktion?

Die Wellenfunktion ist komplex, sie selbst kann daher keine physikalische Grösse repräsentieren. Jedoch, gemäss der “Kopenhagener Schule” (Niels Bohr, 1926) gilt:

1. Die “Aufenthaltswahrscheinlichkeit”, d.h. die Wahrscheinlichkeit  $w(\mathbf{r},t)$ , ein Elektron am Ort  $\mathbf{r}$  zur Zeit  $t$  anzutreffen, beträgt,

$$w(\mathbf{r},t) = \Psi(\mathbf{r},t) \Psi^*(\mathbf{r},t) = |\Psi(\mathbf{r},t)|^2 \quad \text{mit} \quad \int \Psi^* \Psi dV = 1$$

2. Der Mittelwert der kinetischen Energie  $E_{\text{kin}}$  des Elektrons beträgt

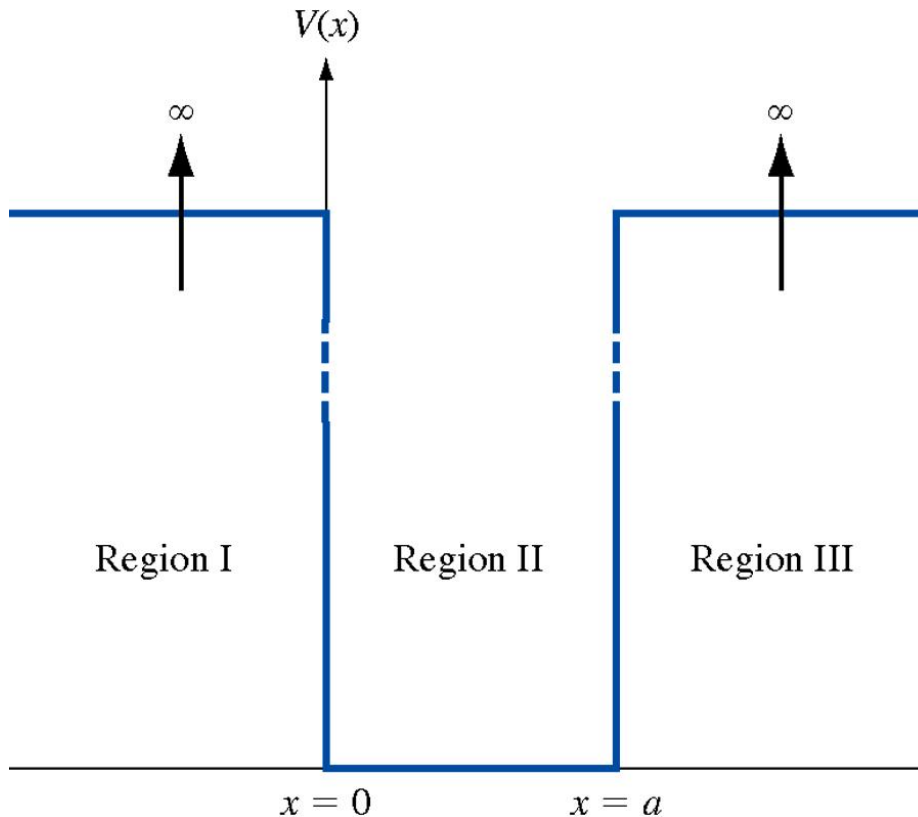
$$\bar{E} = -\int \Psi^* \frac{\hbar^2}{2m} \Delta \Psi dV$$

# 2.3 Anwendungen der Schrödinger-Gleichung

## Der Potentialtopf

$$V(x) = \infty \quad \text{für } x \leq 0, x \geq a$$

$$V(x) = 0 \quad \text{im Innern}$$



Potentialtopf mit unendlich hohen Wänden.

$$\psi(x) = A e^{jkx} + B e^{-jkx} \quad \text{im Innern}$$

$$\psi(0) = \psi(a) = 0$$

(die Elektronenwelle kann in eine unendlich hohe Potentialwand nicht eindringen!)

Ausnutzung der beiden Randbedingungen ergibt die Gleichungen

$$A + B = 0, \quad A e^{jka} + B e^{-jka} = 0$$

oder zusammen:

$$\sin(ka) = 0$$

$\sin(ka) = 0$  ist die *Quantisierungsbedingung* für die Wellenzahl  $k$ . Nur diskrete Werte

$$k_n = \frac{n\pi}{a} \quad \text{mit } n = 1, 2, \dots$$

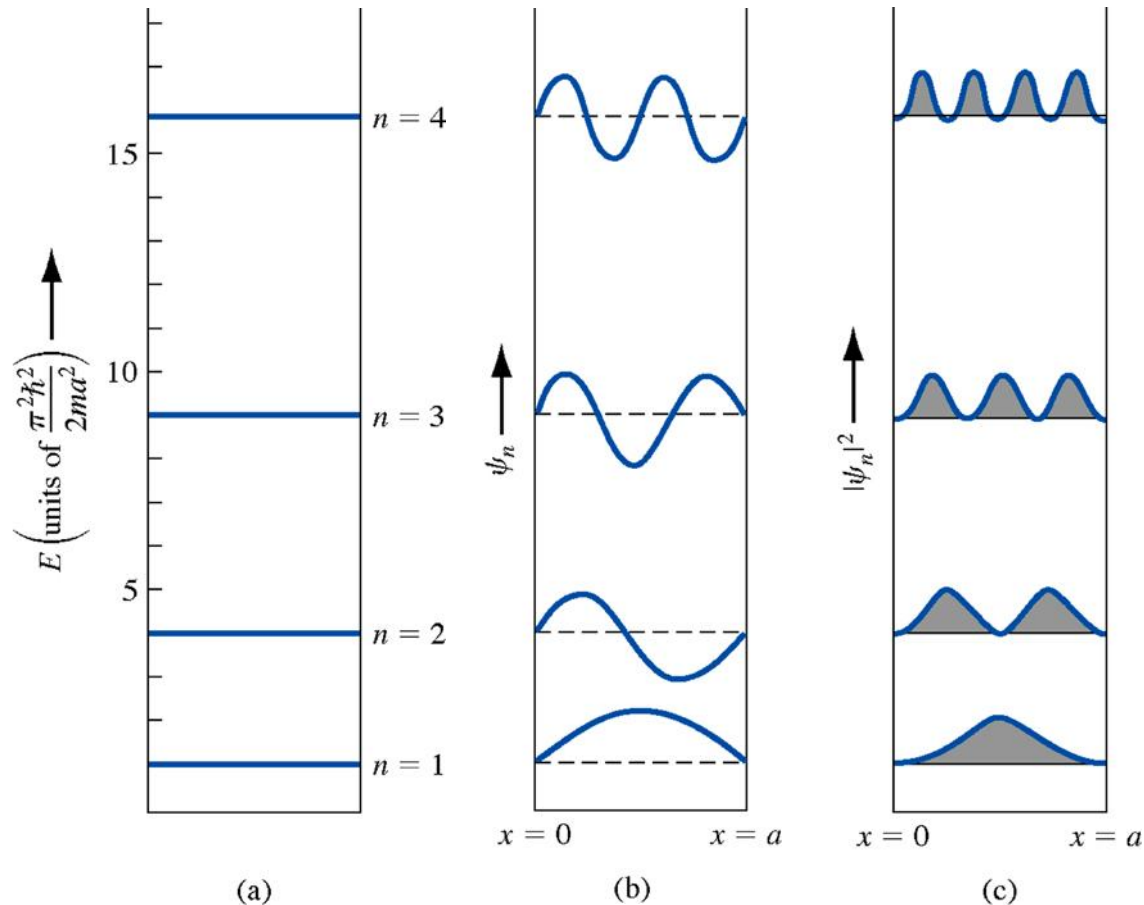
erfüllen diese Bedingung. Daher ist auch die Energie quantisiert:

$$E_n = \frac{\hbar^2 k_n^2}{2m} = \frac{\hbar^2 \pi^2}{2ma^2} n^2$$

Die Wellenfunktionen sind *stehende Wellen*:

$$\psi_n(x) = \sqrt{\frac{2}{a}} \sin\left(\frac{n\pi}{a} x\right)$$

Wellenlänge:  $\lambda_n = \frac{2\pi}{k_n} = \frac{2a}{n}$

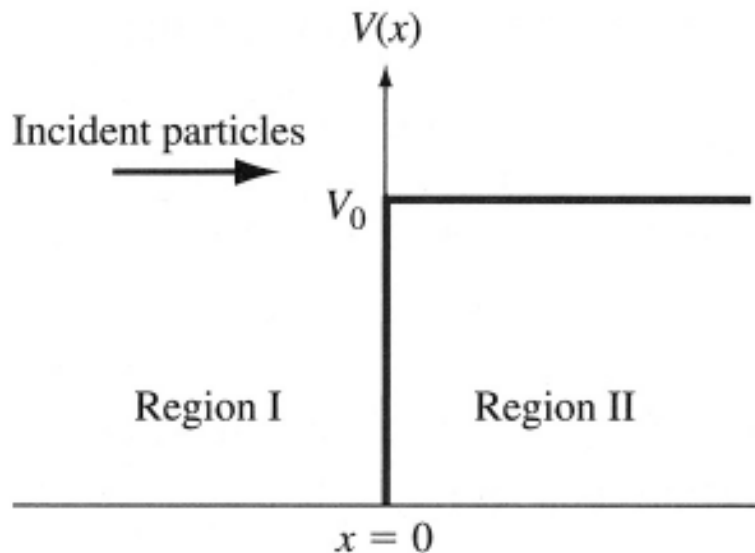


a) Die vier niedrigsten Energieniveaus, b) zugehörige Wellenfunktionen, c) zugehörige Aufenthaltswahrscheinlichkeiten.

# Anwendungen der Schrödinger-Gleichung: Die Potentialstufe

$$\psi_{II}(x) = \begin{cases} A_2 e^{-k_2 x} & E < V_0 \\ B_2 e^{jk_2 x} & E > V_0 \end{cases}$$

$$k_2 = \sqrt{2m |V_0 - E|} / \hbar$$



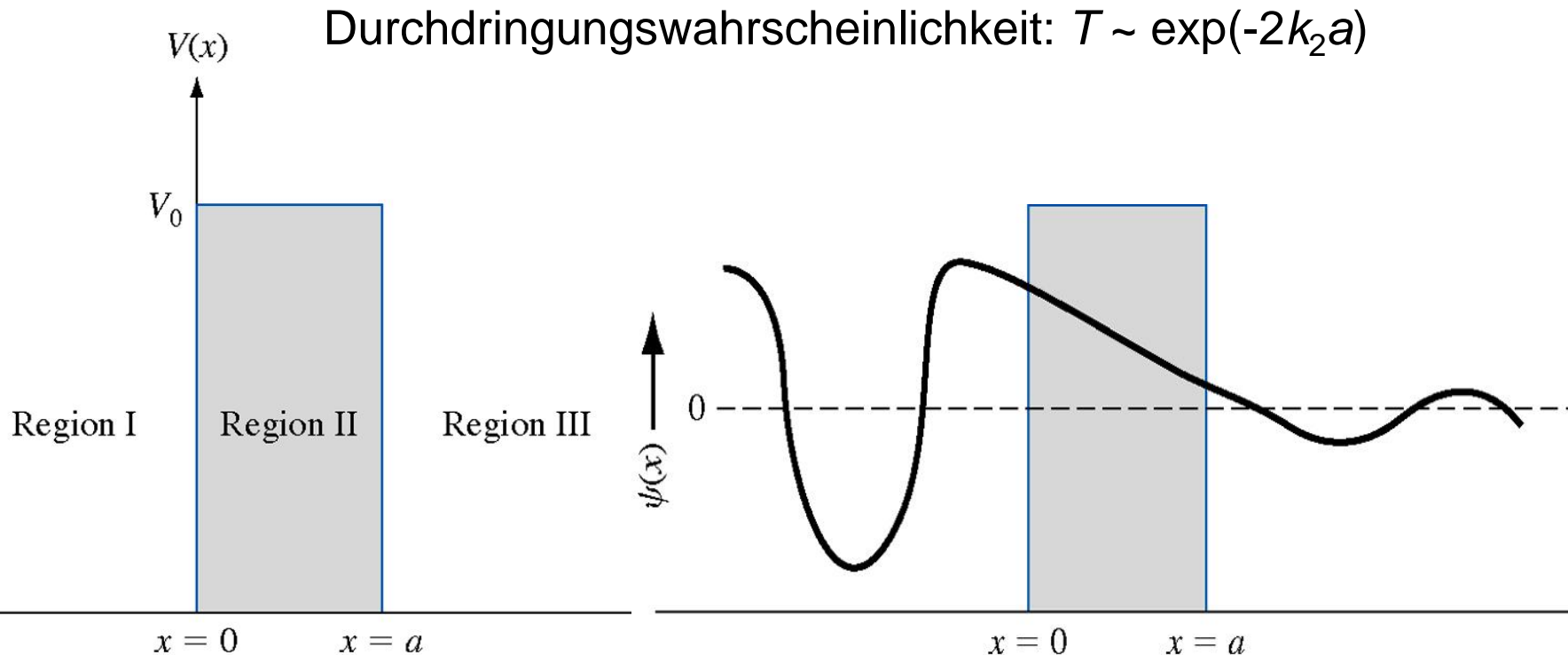
Funktionsverlauf der Potentialstufe.

Die Quantenmechanik erlaubt eine endliche Aufenthaltswahrscheinlichkeit eines Teilchens im Bereich II, der bei Energien  $E < V_0$  klassisch verboten ist.

Einfallende Wellen mit  $E < V_0$  werden vollständig reflektiert und dringen dabei gleichzeitig in den Bereich II ein.

Einfallende Wellen mit  $E > V_0$  werden teilweise reflektiert und teilweise über die Potentialstufe hinweg in den Bereich II hinein transmittiert.

# Anwendungen der Schrödinger-Gleichung: Die Potentialbarriere



Die Quantenmechanik erlaubt eine endliche Aufenthaltswahrscheinlichkeit eines Teilches im (klassisch verbotenen) Bereich II. Einfallende Wellen werden teilweise reflektiert, sie können aber den Bereich II durchdringen, und damit die Barriere überwinden (Tunneleffekt).

# 2.4 Die Quantenmechanik von Atomen

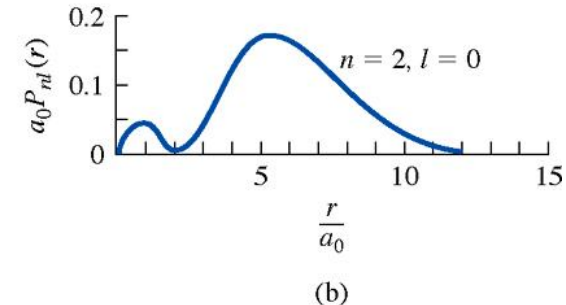
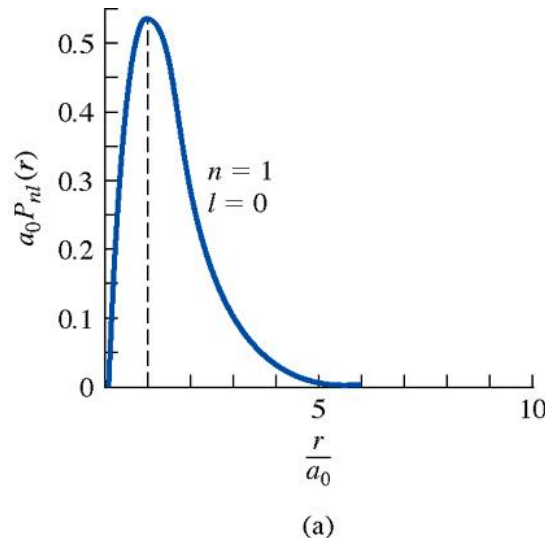
## Das Einelektronenatom

Die Schrödingergleichung kann für das Einelektronenatom exakt gelöst werden: Vier Quantenzahlen beschreiben den Zustand eines Elektrons:

1. die Hauptquantenzahl  $n = 1, 2, 3, \dots$
2. die Nebenquantenzahl  $l = n-1, n-2, \dots, 0$
3. die magnetische Quantenzahl  $m = -l, -l+1, \dots, 0, \dots, +l$
4. der Elektronenspin  $s = \pm 1/2$

$$E_n = \frac{-m_0 e^4}{(4\pi\epsilon_0)^2 2\hbar^2 n^2}$$

negative quantisierte  
Energien => *gebundene*  
Zustände



Radiale Aufenthaltswahrscheinlichkeit: a) im Grundzustand, b) im 1. angeregten Zustand.

# Das Periodensystem der Elemente

Table 2.1 | Initial portion of the periodic table

Element	Notation	$n$	$l$	$m$	$s$
Hydrogen	$1s^1$	1	0	0	$+\frac{1}{2}$ or $-\frac{1}{2}$
Helium	$1s^2$	1	0	0	$+\frac{1}{2}$ and $-\frac{1}{2}$
Lithium	$1s^2 2s^1$	2	0	0	$+\frac{1}{2}$ or $-\frac{1}{2}$
Beryllium	$1s^2 2s^2$	2	0	0	$+\frac{1}{2}$ and $-\frac{1}{2}$
Boron	$1s^2 2s^2 2p^1$	2	1	}	$m = 0, -1, +1$ $s = +\frac{1}{2}, -\frac{1}{2}$
Carbon	$1s^2 2s^2 2p^2$	2	1		
Nitrogen	$1s^2 2s^2 2p^3$	2	1		
Oxygen	$1s^2 2s^2 2p^4$	2	1		
Fluorine	$1s^2 2s^2 2p^5$	2	1		
Neon	$1s^2 2s^2 2p^6$	2	1		





# 3. Quantentheorie des Festkörpers

3.1 Erlaubte Energiebänder und verbotene Zonen (N3.1)

3.2 Elektrische Leitung in Festkörpern (N3.2)

3.3 Verallgemeinerung auf den dreidimensionalen Fall (N3.3)

3.4 Die Zustandsdichte (N3.4)

3.5 Statistische Mechanik (N3.5)

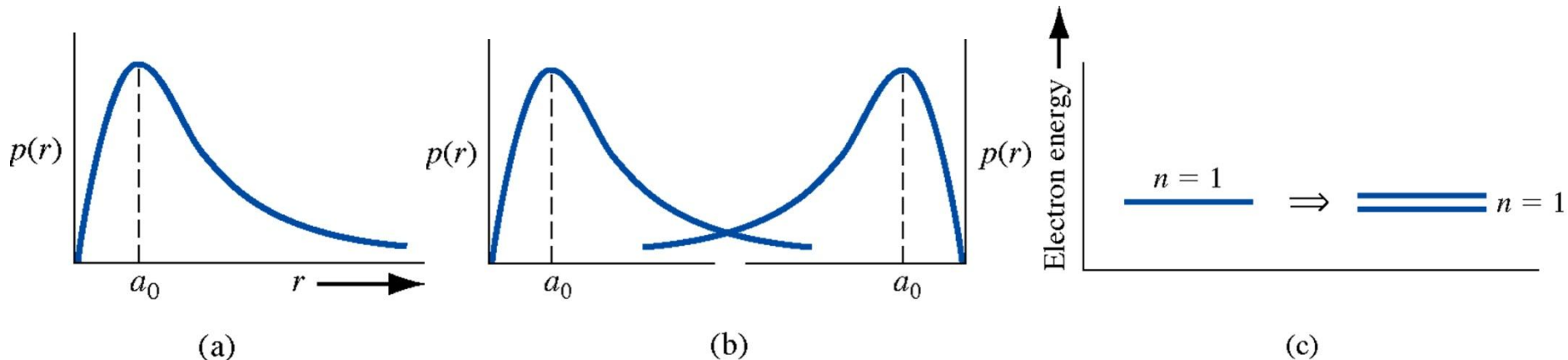
# 3.1 Erlaubte Energiebänder und verbotene Zonen

## Die Entstehung von Energiebändern

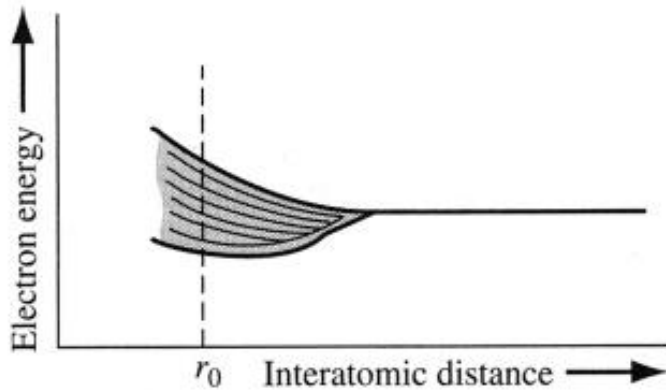
Im Kristall überlappen die Elektronen-Wellenfunktionen benachbarter Atome. Resultat ist eine Wechselwirkung zwischen diesen Elektronen.

Diese Wechselwirkung führt zur Aufspaltung der diskreten atomaren Energieniveaus.

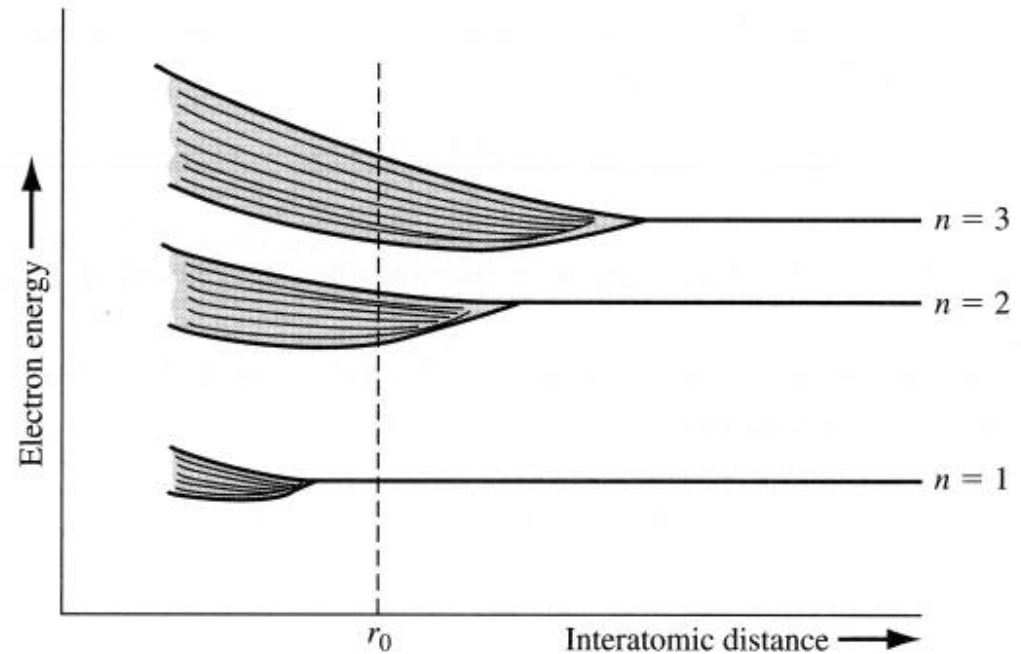
Beispiel: Grundzustand im Wasserstoffatom. Die Aufspaltung entspricht der Wechselwirkungsenergie der beiden Elektronen im  $H_2$ -Molekül.



Ein regulär angeordnetes Ensemble von Atomen führt zu einem Quasikontinuum von Energiebändern. Diese Bänder können auch überlappen.



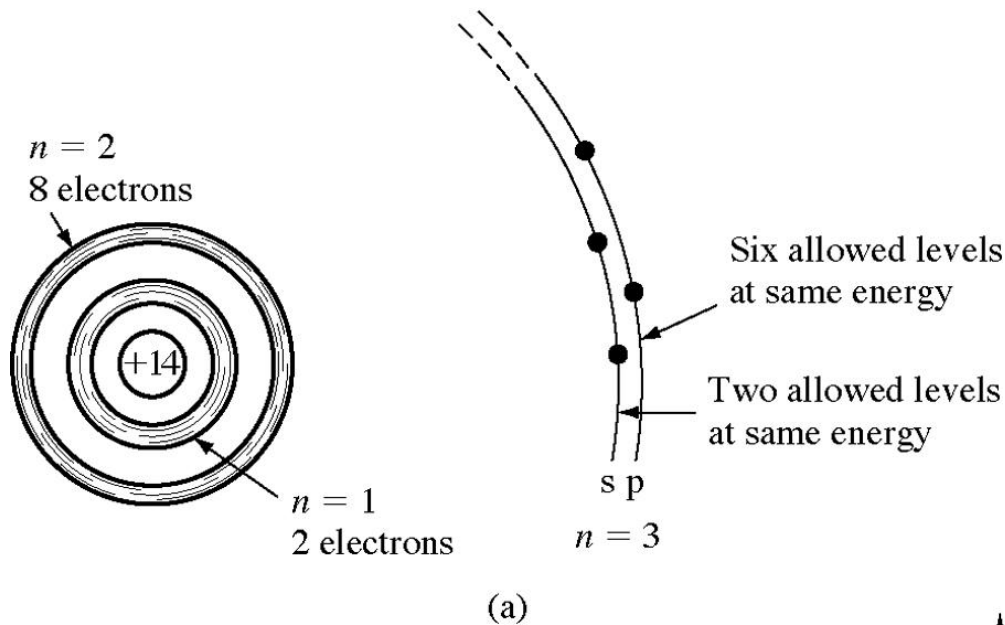
Aufspaltung eines Energieniveaus in ein Band erlaubter Energien.



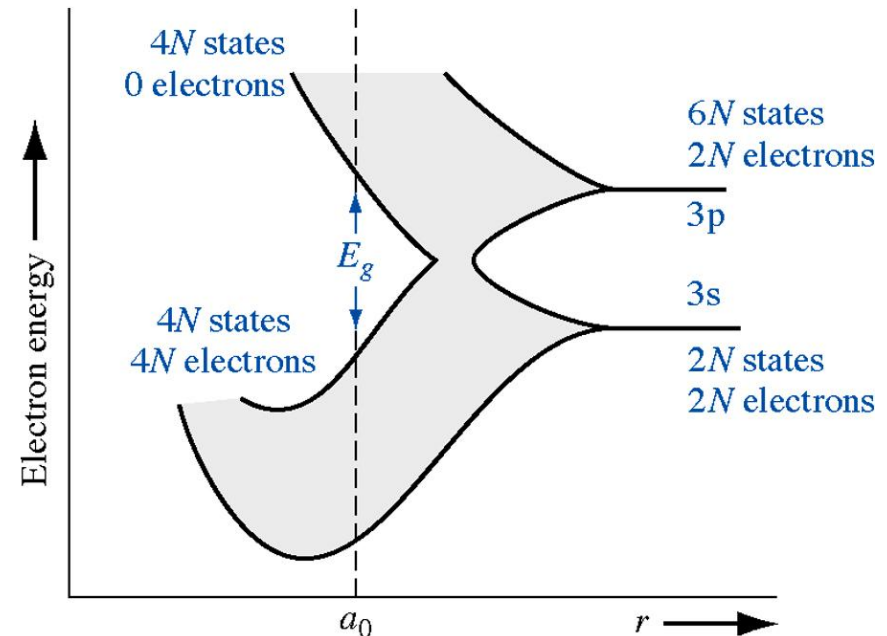
Schematische Darstellung der Aufspaltung von 3 Energieniveaus in 3 Bänder erlaubter Energien.

Was bedeutet *Quasikontinuum*? Nehmen an, dass sich  $10^{19}$  Einelektronenatome so zum Kristall anordnen, dass eine Bandweite von 1 eV resultiert. Dann sind die Abstände zwischen den Energieniveaus nur  $10^{-19}$  eV!

# Vom Silizium-Atom zum Silizium-Gitter



- nur Valenzelektronen der 3. Schale sind an chemischen Reaktionen beteiligt
- $2 \times 3s$  ( $n=3, l=0$ ) und  $2 \times 3p$  ( $n=3, l=1$ )
- Wechselwirkung im Kristall führt zu  $4 sp^3$ -Hybridorbitalen



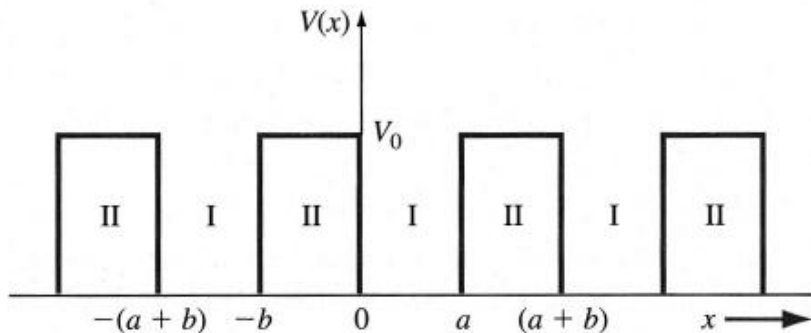
- die 8 Quantenzustände ergeben 4 Valenzbänder und 4 Leitungsbänder
- Bei  $T=0$  K sind alle Valenzbänder besetzt und alle Leitungsbänder leer
- Leitungs- und Valenzbänder sind durch die Energielücke  $E_g$  ("gap") voneinander getrennt

# Das Kronig-Penney Modell\*

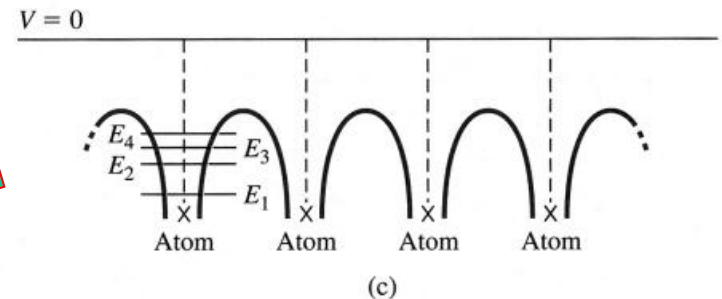
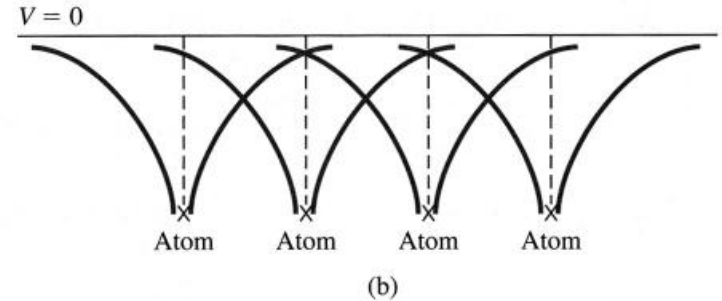
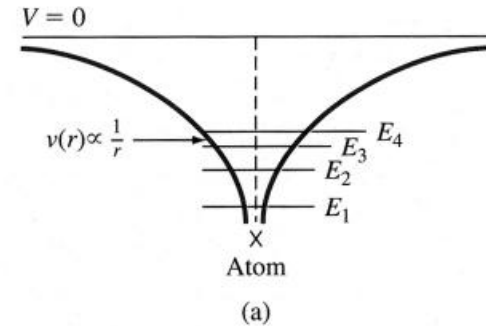
Im Festkörper sind die Atomrümpfe gitterperiodisch angeordnet, und die Valenzelektronen sind an diesen Atomrümpfen gebunden. Deshalb zeichnet sich ein Kristall durch einen periodischen Verlauf der potentiellen Energie aus.

Um das Verhalten der Elektronen in diesem gitterperiodischen Potential zu verstehen, kann man wie folgt vorgehen:

1. Betrachtung eines einfachen Modells
2. Verallgemeinerung



Das eindimensionale periodische Potential des Kronig-Penney-Modells.



(a) Verlauf des Potentials im Fall eines einzelnen isolierten Atoms, (b) überlappende Potentiale benachbarter Atome, (c) resultierendes Potential eines eindimensionalen Einkristalls.

Der Ortsverlauf der potentiellen Energie lässt sich in erster Näherung durch ein periodisches Kastenpotential simulieren. In ihm sind die Atome durch periodisch angebrachte Potentialwälle der Höhe  $V_0$  und der Breite  $b$  getrennt.

Die Frage lautet: Mit welcher Wahrscheinlichkeit kann ein Elektron aus dem Potentialtopf 1 in den Potentialtopf 3 oder 5 gelangen?

Ausgangspunkt ist die eindimensionale Schrödinger-Gleichung:

$$\frac{d^2\psi}{dx^2} + \frac{2mE}{\hbar^2} \psi = 0 \quad (0 \leq x \leq a)$$
$$\frac{d^2\psi}{dx^2} + \frac{2m}{\hbar^2} (E - V_0) \psi = 0 \quad (-b \leq x \leq 0)$$

Nach dem **Theorem von Bloch-Floquet** ist eine Lösung  $\psi_k$  stets von der Form

$$\psi_k(x) \equiv u_k(x) \exp jk_x x$$

wobei  $u_k$  bezüglich  $a_0 = a + b$  periodisch ist. Der Elektronenzustand wird durch eine (räumliche) Welle mit der Materiewellenlänge  $\lambda = 2\pi/k_x$  beschrieben. Ihre Amplitude ist gitterperiodisch moduliert.

Das Elektron im Kristall gehört damit nicht mehr zu einem individuellen Atom, sondern tritt als Wellenerscheinung im Kristall auf (Blochwelle, Blochfunktion).

**Lösung.** Die DGL ist linear und von 2. Ordnung. Mit dem Ansatz der Blochfunktion folgt

$$\frac{d^2 u_k}{dx^2} + 2jk_x \frac{du_k}{dx} - \left( k_x^2 - \alpha^2 + \frac{2mV}{\hbar^2} u_k(x) = 0 \right)$$

$$\alpha = \sqrt{\frac{2mE}{\hbar^2}} \equiv k_{0,x}$$

für ein Elektron mit der Energie  $0 < E < V_0$ . Das Potential  $V$  ist in den beiden Bereichen entweder 0 oder  $V_0$ :

*Bereich 1:*  $V = 0$  ( $0 \leq x \leq a$ )

$$\frac{d^2 u_k}{dx^2} + 2jk_x \frac{du_k}{dx} - (k_x^2 - \alpha^2) u_k(x) = 0;$$

*Bereich 2:*  $V = V_0$  ( $-b \leq x \leq 0$ )

$$\frac{d^2 u_k}{dx^2} + 2jk_x \frac{du_k}{dx} - (k_x^2 - \beta^2) u_k(x) = 0$$

$$\beta = \sqrt{\frac{2m(E - V_0)}{\hbar^2}} = j\gamma$$

Durch den Blochansatz kann man beide Gleichungen leicht lösen

$$u_{k1}(x) = A \exp j(\alpha - k_x)x + B \exp -j(\alpha + k_x)x \quad (0 \leq x \leq a),$$

$$u_{k2}(x) = C \exp j(\beta - k_x)x + D \exp -j(\beta + k_x)x \quad (-b \leq x \leq 0)$$

Zum Schluss müssen die Konstanten  $A, B, C, D$  als Funktion von  $\alpha$  und  $\beta$  (bzw.  $\gamma$ ), d.h. als Funktion der Energie bestimmt werden:



Die Konstanten  $A, B, C, D$  werden durch die Randwerte festgelegt:

1. Stetigkeit der Funktionen bei  $x = 0$  und  $x = -b, x = a$

$$u_{k_1}(0) = u_{k_2}(0) \quad u_{k_1}(a) = u_{k_2}(-b) \quad (\text{gitterperiodische Randbedingungen})$$

2. Stetigkeit der ersten Ableitungen

$$u'_{k_1}(0) = u'_{k_2}(0) \quad u'_{k_1}(a) = u'_{k_2}(-b)$$

Dies führt auf folgende Bedingungen:

$$A + B = C + D$$

$$j(\alpha - k_x)A - j(\alpha + k_x)B = j(\beta - k_x)C - j(\beta + k_x)D$$

$$A \exp j(\alpha - k_x)a + B \exp(-j(\alpha + k_x)a) = C \exp -j(\beta - k_x)b + D \exp j(\beta + k_x)b$$

$$j(\alpha - k_x)A \exp j(\alpha - k_x)a - j(\alpha + k_x)B \exp(-j(\alpha + k_x)a) =$$

$$j(\beta - k_x)C \exp j(\beta - k_x)b - j(\beta + k_x)D \exp j(\beta + k_x)b$$

Nichttriviale Lösungen existieren bei verschwindender Koeffizienten-Determinante: Eigenwertgleichung

$$-\frac{\alpha^2 + \beta^2}{2\alpha\beta} \sin\alpha a \sin\beta b + \cos\alpha a \cos\beta b = \cos k_x(a + b)$$

Im Bereich  $0 < E < V_0$  ist  $\beta = j\gamma$  rein imaginär und es gilt  $\frac{\gamma^2 - \alpha^2}{2\alpha\gamma} \sinh\gamma b \sin\alpha a + \cosh\gamma b \cos\alpha a = \cos k_x(a + b)$

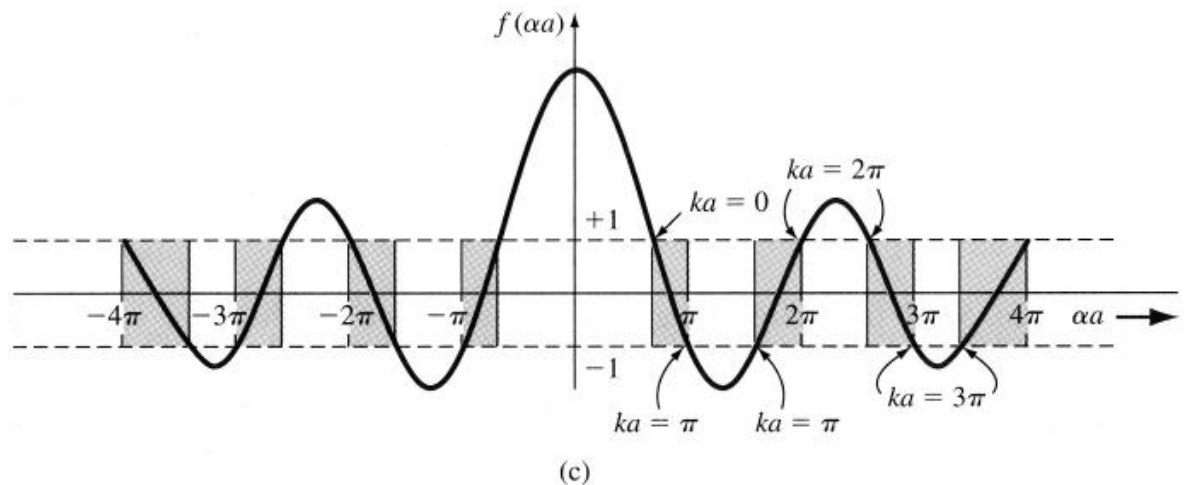
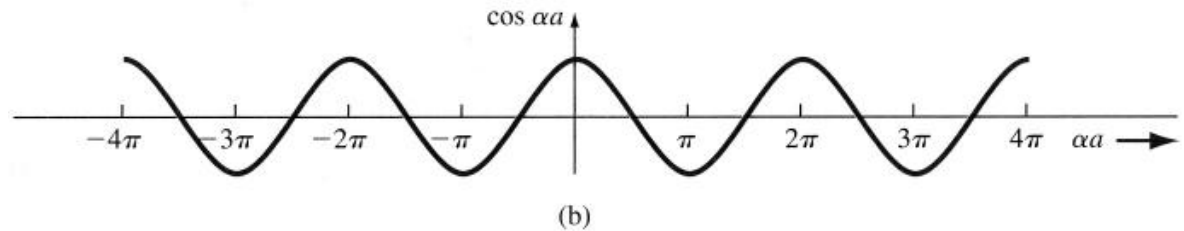
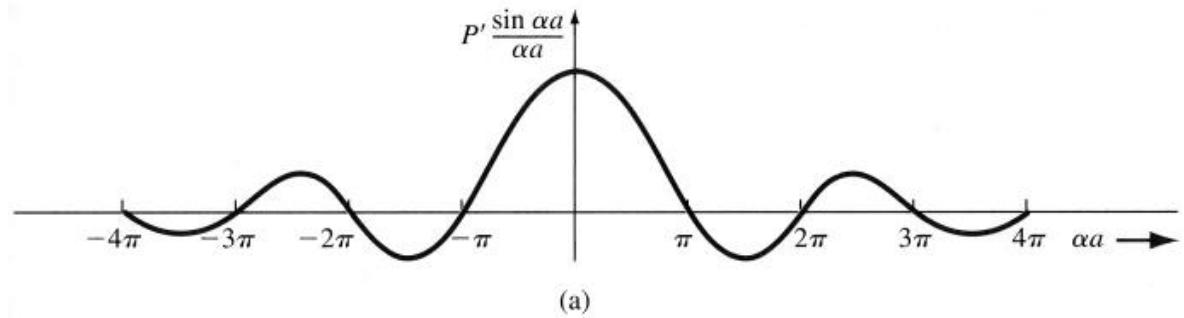
**Vereinfachung.** Wir vereinfachen jetzt das Potentialmodell, indem wir das Kastenpotential durch ein Deltapotential gleicher Fläche, aber verschwindender Breite  $b$  ersetzen. Dann bleibt  $\lim_{b \rightarrow 0} V_0 b = \text{const}$  bei  $b$  gegen Null ( $\gamma$  gegen unendlich) und es gilt

$$P' = \lim_{\substack{b \rightarrow 0 \\ \gamma \rightarrow \infty}} \left( \gamma^2 \frac{ab}{2} \right) = \frac{mV_0 ab}{\hbar^2}$$

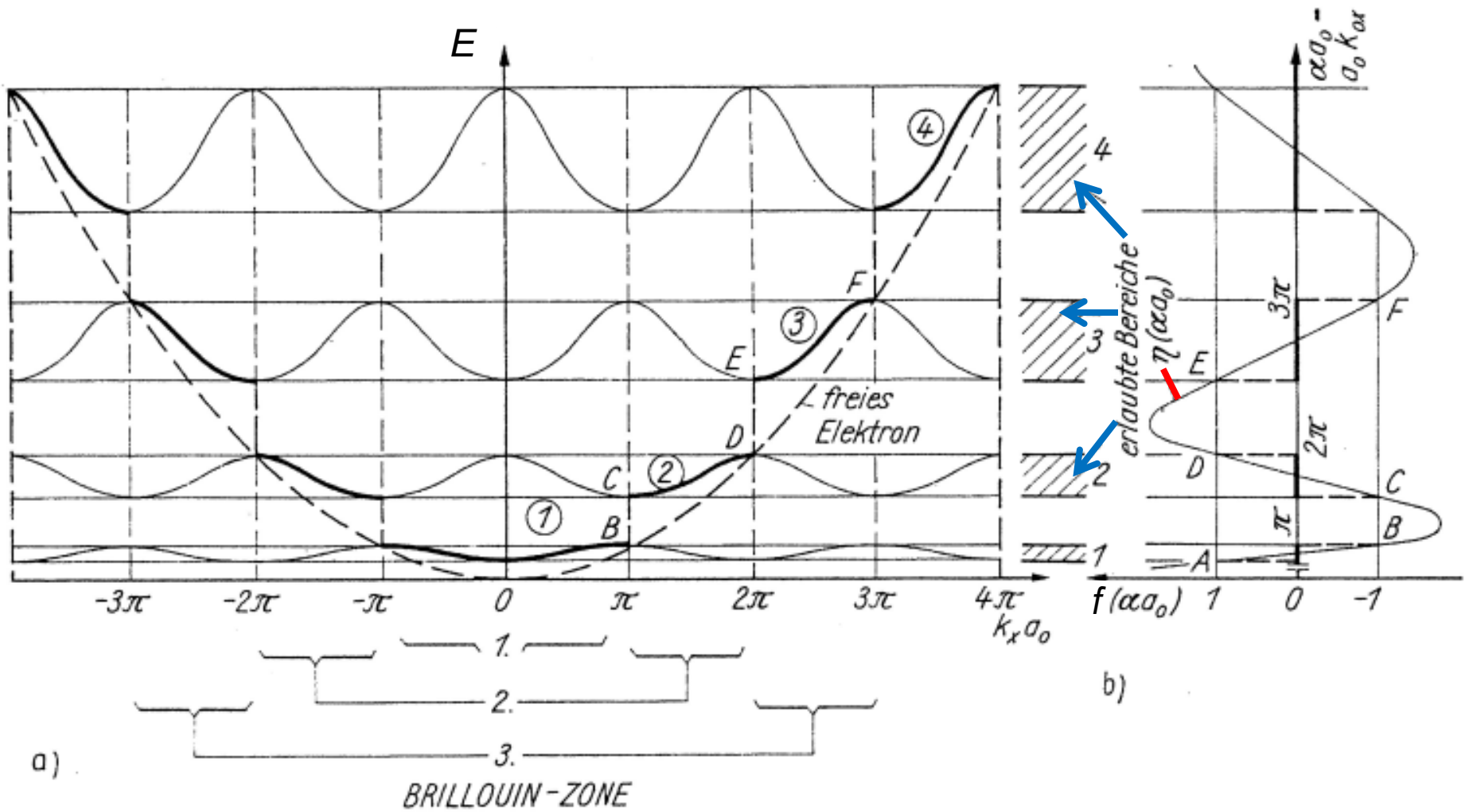
d.h.

$$\cos ak_x = P' \frac{\sin \alpha a}{\alpha a} + \cos \alpha a \equiv f(\alpha a)$$

Die hervorgehobenen Bereiche zeigen die erlaubten Werte von  $\alpha a \sim \sqrt{E}$ , die zu reellen Werten von  $k_x$  gehören. Lässt man  $k_x$  variieren, kann man daraus die Bandstruktur  $E(k_x)$  des Kronig-Penney Modells extrahieren.



(a) Erster Term von  $f(\alpha a)$ , (b) zweiter Term von  $f(\alpha a)$ , (c) gesamte Funktion  $f(\alpha a)$ . Die schattierten Bereiche zeigen die erlaubten Werte  $\alpha a$ , die reellen Werten von  $k$  entsprechen.



(a) Energie als Funktion der Wellenzahl, (b) Zusammenhang von  $f(\alpha a_0)$  mit  $\alpha a_0$ .

**Ergebnis:** Im Gegensatz zum Verhalten des Elektrons im unbegrenzt konstanten Potential mit allen möglichen  $k$ -Werten sind im unbegrenzt periodischen Potential des Kronig-Penney Modells nur noch Ausbreitungsvektoren nach folgender Beziehung erlaubt:

$$\cos(a k_x) = P' \frac{\sin(\alpha a)}{\alpha a} + \cos(\alpha a) \equiv f(\alpha a) .$$

$$\alpha = \sqrt{\frac{2mE}{\hbar^2}}$$

$$P' = \frac{mV_0ab}{\hbar^2}$$

Die linke Seite dieser Gleichung liegt immer zwischen -1 und 1, und gleiches muss für die rechte gelten. Innerhalb dieses Gebietes gibt es zu jedem Wert  $k_x a$  mehrere Werte von  $\alpha a$  bzw. der Energie  $E$ .

Da  $\alpha$  der Wurzel aus der Elektronenenergie proportional ist, kann man prinzipiell  $E$  über  $ak_x$  auftragen.

Damit sind die möglichen Energieeigenwerte  $E$  eingeschränkt. Es treten erlaubte und verbotene Energiebänder abwechselnd auf. Ferner hängt die Energie  $E$  periodisch von der Wellenzahl  $k_x$  mit der Periode  $2\pi/a$  ab.

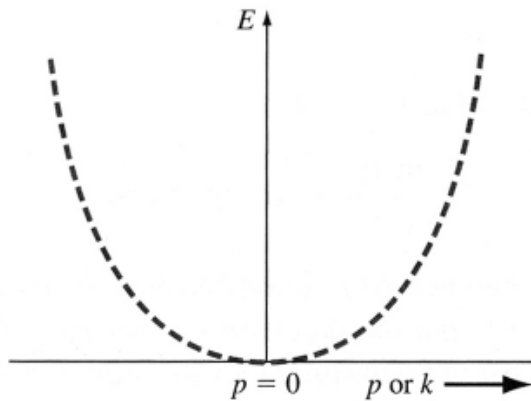
# Darstellungen des Energiebandverlaufes

**Ausgebreitetes Energiebandschema:** Die Wellenzahl variiert über dem gesamten Bereich. Die Energielücken reissen an den Rändern der Brillouinzonen (BZ) auf.

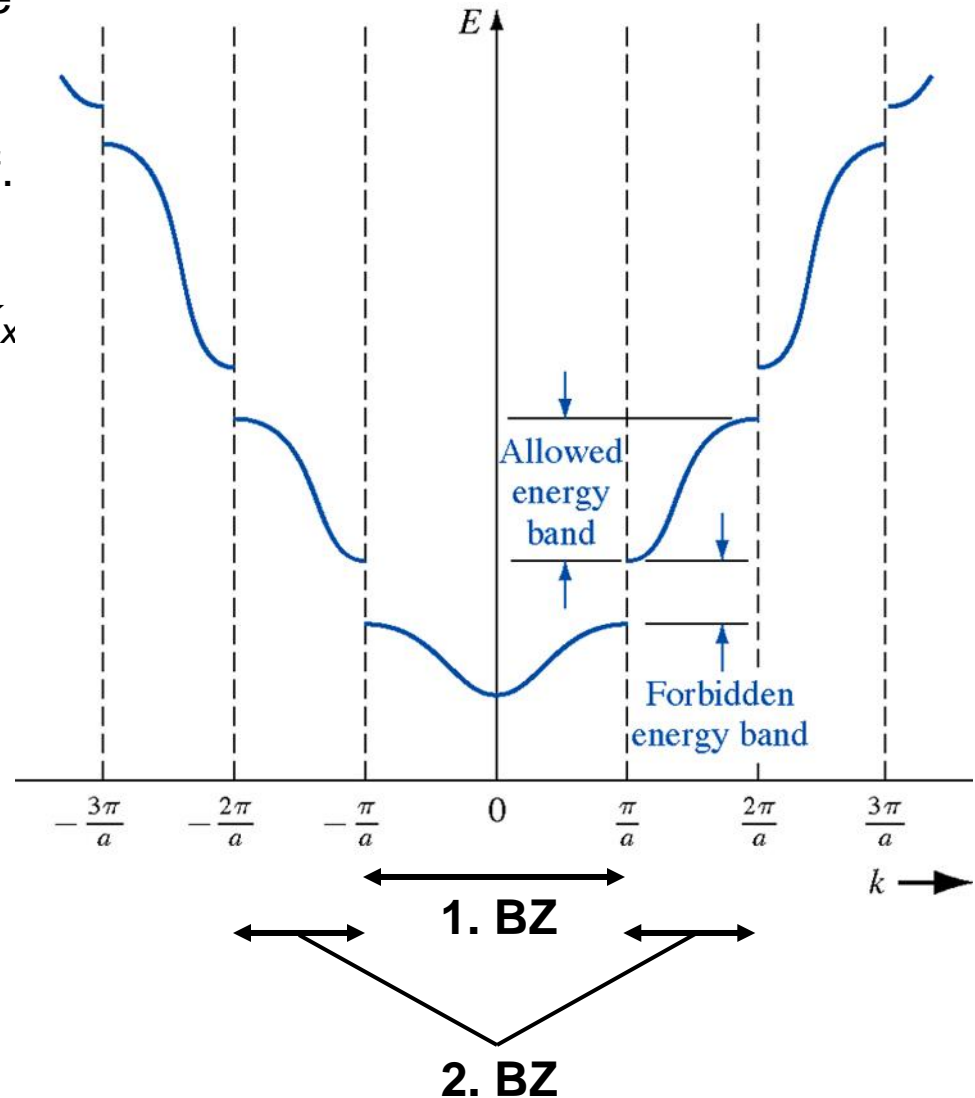
Lässt man die Potentialbarrieren verschwinden ( $V_0 \rightarrow 0$ ), ergibt sich  $\alpha = k_x$  oder

$$E(k_x) = \frac{\hbar^2 k_x^2}{2m}$$

d.h. die  $E(k)$ -Relation des kräftefreien Elektrons

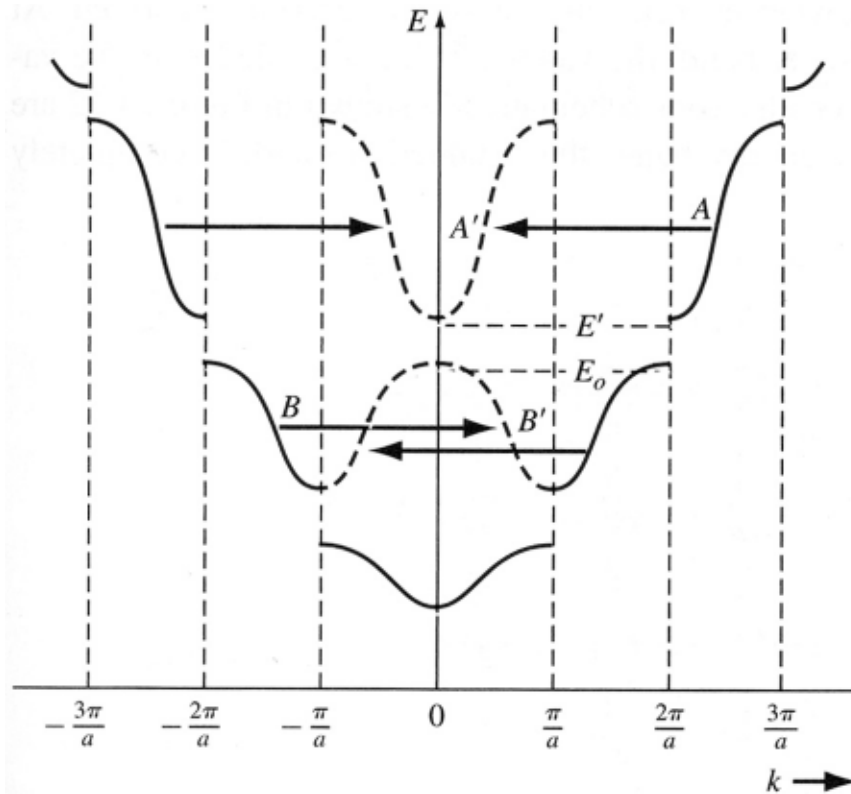


Die parabolische  $E(k)$ -Relation des freien Elektrons.

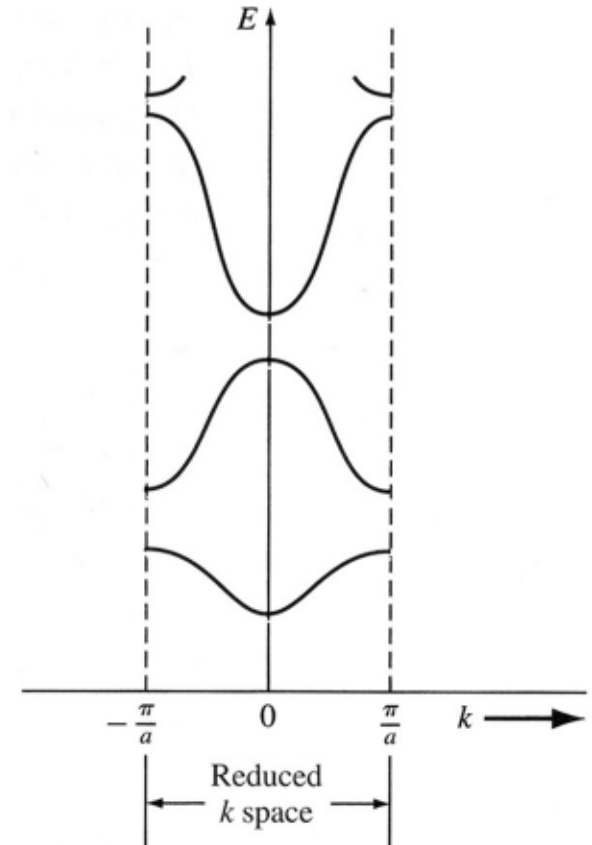


# Vom ausgebreiteten zum reduzierten Energiebandschema

**Reduziertes Energiebandschema:** Man verschiebt die Teile von  $E(k)$  aus höheren Brillouinzonen in die 1.BZ zurück ("Falten"). Verschiebung um ganzzahlige Vielfache von  $2\pi$ . Dies ist möglich, weil  $\cos(ka) = \cos(ka + 2n\pi)$ .



$E(k)$ -Diagramm mit Darstellung der Verschiebung von verschiedenen Abschnitten der erlaubten Energiebänder um  $2\pi$ .



$E(k)$ -Diagramm im reduzierten Zonenschema.

# Feststellungen

1. Im unendlich ausgedehnten Kristall mit gitterperiodischem Potential  $V(x)$  hat die Schrödinger-Gleichung Bloch-Funktionen als Lösung:

$$\psi_k(x) \equiv u_k(x) \exp jk_x x$$

2. Für die Energie  $E$  gibt es abwechselnd erlaubte und verbotene Bereiche  $E(k_x)$ .
3. Innerhalb eines erlaubten Bandes gibt es im Falle des unbegrenzt ausgedehnten Kristalls unendlich viele erlaubte Energiewerte, im Falle des begrenzten Kristalls jedoch nur endlich viele.
4. In erlaubten Bändern ist  $k_x$  reell. Dort liegt eine ungedämpfte Elektronenwelle mit dem Ausbreitungsvektor  $k_x$  und gitterperiodisch modulierter Amplitude vor. In verbotenen Bändern ist  $k_x$  dagegen imaginär und entspricht einem exponentiell abklingenden Wellenzug.
5. Die Breite eines erlaubten Energiebandes (= Bandweite) wächst mit steigender Elektronenenergie  $E$ .
6. Mit wachsendem  $P'$  verbreitert sich die verbotene Zone. Da  $P'$  ein Mass für die Höhe des Potentialwalls ist, bedeutet ein grosses  $P'$  starke Bindung des Elektrons an das Atom. Für  $P' \rightarrow \infty$  reduzieren sich die Energiebänder auf das Energiespektrum des Einzelatoms.

7. Für  $P' = 0$  (Keine Potentialwand, kräftefreies Elektron) gilt für alle Werte  $k_x$

$$k_x = k_{0x} = \alpha = \sqrt{\frac{2mE}{\hbar^2}} \quad \text{und damit} \quad E = \frac{\hbar^2 k^2}{2m}$$

In diesem Grenzfall gibt es keine verbotenen Energiebereiche.

8. Obere Bandränder treten auf für

$$ka = \pi n \quad n = \pm 1, \pm 2, \dots$$

untere Bandränder entsprechend für

$$ka = \pi(n-1) \quad n = \pm 1, \pm 2, \dots$$

So gibt es bei laufendem  $n$  unendlich viele Bandkanten. Die Zahl  $n$  hat demnach die Bedeutung einer Bandnummer (oder eines *Bandindex*):  $n = 1$  erstes erlaubtes Energieband, usw.

9. Wegen  $\cos(ak_x) = \cos(-ak_x)$  läuft die  $E(k_x)$ -Kurve symmetrisch zur  $E$ -Achse. Damit nimmt die erste Zone den Bereich  $-\pi/a < k_x < \pi/a$ , die zweite den Bereich  $-2\pi/a < k_x < -\pi/a$  und  $\pi/a < k_x < 2\pi/a$  an. Wir nennen diese Bereiche erste und zweite Brillouin-Zone, usw.
10. Physikalisch relevant ist nur die 1. Brillouin-Zone, da wegen  $\cos(ka) = \cos(ka+2n\pi)$  der Ausbreitungsvektor  $k' = k + 2n\pi/a$  physikalisch äquivalent zu  $k$  ist.



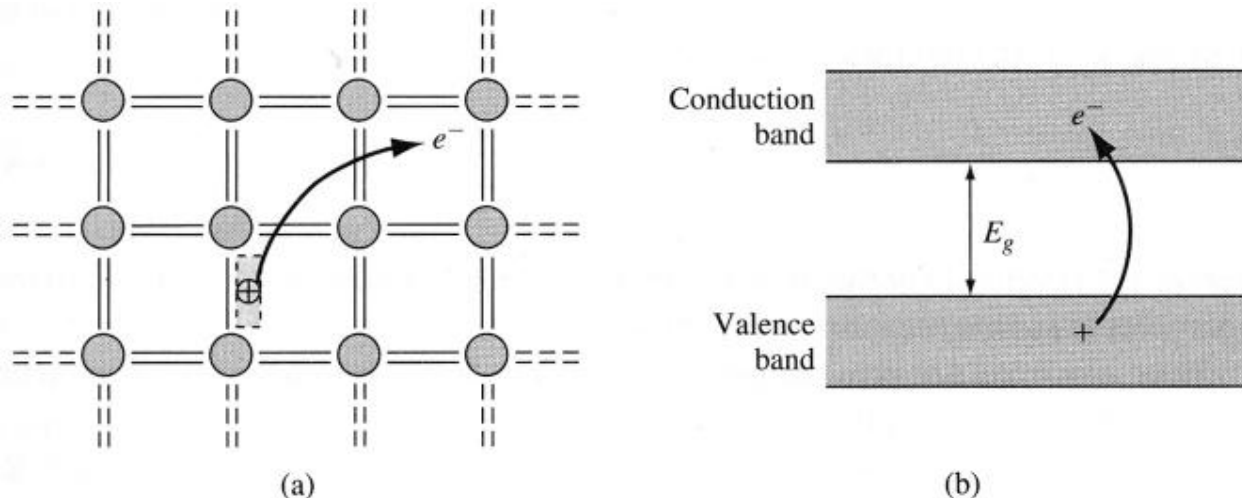
# 3.2 Elektrische Leitung im Festkörper

## Kovalente Bindung und Energieband

Bei  $T = 0$  K sind alle Valenzelektronen an der kovalenten Bindung beteiligt und füllen das Valenzband vollständig.

Bei  $T > 0$  K bekommen einige Elektronen genügend thermische Energie, um in einen angeregten Zustand im Leitungsband übergehen zu können.

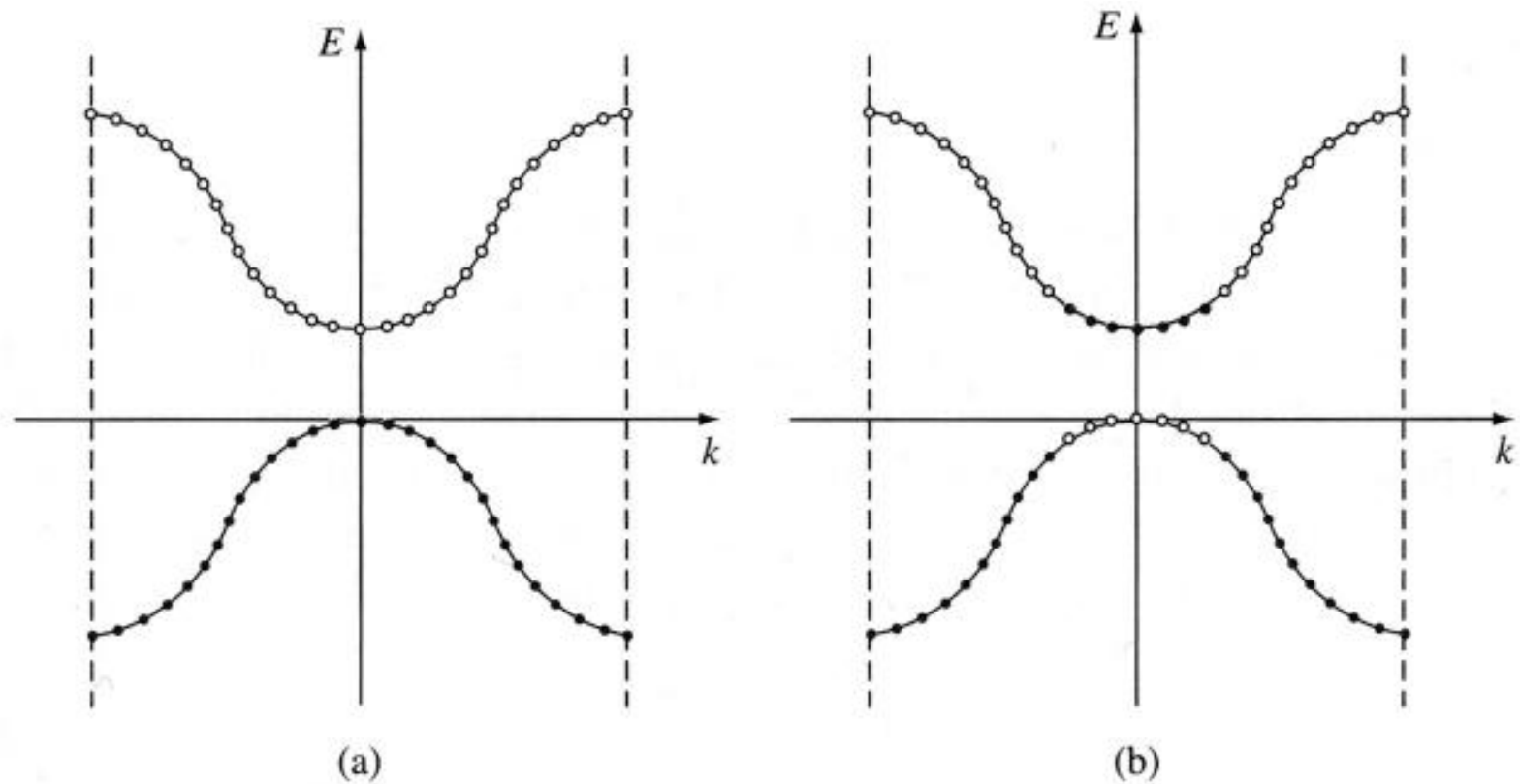
Sie hinterlassen positiv geladene Defektelektronen = Löcher im Valenzband



(a) Zweidimensionale Darstellung des Aufbrechens einer kovalenten Bindung, (b) Generation eines Elektrons im Leitungsband und eines Lochs im Valenzband infolge des Aufbrechens einer kovalenten Bindung.

# Besetzung von $E_C$ und $E_V$ : Einfluss der Temperatur

Im  $E(k)$ -Banddiagramm kann man dies so veranschaulichen:



$E(k)$ -Diagramm des Leitungs- und Valenzbandes eines Halbleiters bei (a)  $T = 0$  K und (b)  $T > 0$  K.

# Der Driftstrom

Elektrischer Strom wird durch die Bewegung von Ladungen getragen. Bewegen sich  $N$  positive Ladungsträger in einem Würfel der Kantenlänge  $d$  (cm) mit einer mittleren Driftgeschwindigkeit  $v_d$ , so ergibt sich eine Driftstromdichte (A/cm<sup>2</sup>) von

$$J = q n v_d \quad \text{mit } n = N / d^3$$

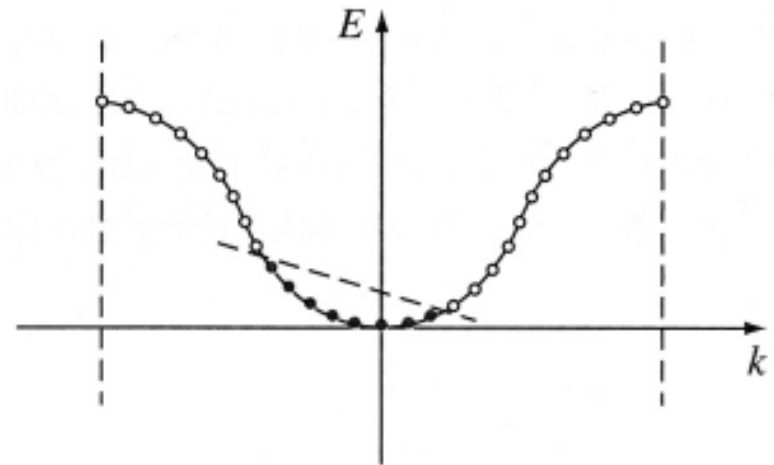
Dieser Strom setzt sich aus den individuellen Bewegungen der einzelnen Ladungsträger zusammen

$$J = \frac{q}{d^3} \sum_{i=1}^N v_i$$

In einem externen Feld kann ein geladenes Teilchen Energie gewinnen:

$$dE = F dx = F v dt$$

$F$  ist die wirkende Kraft.



Asymmetrische Elektronenverteilung im  $E(k)$ -Diagramm als Folge einer angelegten äusseren Kraft.

# Leitfähigkeit der Elektronen im Halbleiter

In einem Halbleiter mit der Elektronendichte  $n$  und der Löcherdichte  $p$  wirke eine äussere Feldstärke  $\mathbf{E}$ . Sie übt auf Elektronen die Kraft

$$\mathbf{F}_n = -e \mathbf{E}$$

aus. Die mittlere Driftgeschwindigkeit  $\mathbf{v}_{dn}$  ist proportional zur Feldstärke  $\mathbf{E}$  (für hinreichend kleine Felder):

$$\mathbf{v}_{dn} = -\mu_n \mathbf{E}.$$

Die Proportionalitätskonstante heisst *Beweglichkeit*  $\mu$ . Damit ergibt sich für die Feldstromdichte der Elektronen

$$\mathbf{J}_n = -en \mathbf{v}_{dn} = -\mu_n n \mathbf{F}_n = en \mu_n \mathbf{E}.$$

Teilchenstromdichte  $\mathbf{J}_n/-e$  und Feldstärke  $\mathbf{E}$  sind entgegengesetzt, (elektrische) Feldstromdichte und Feldstärke also gleichgerichtet.

# Die effektive Masse des Elektrons

Auf ein Elektron wirken neben der äusseren Kraft  $F_{\text{ext}}$  (angelegtes Feld) auch innere Kräfte durch die Wirkung der Atomrümpfe und aller anderer Elektronen

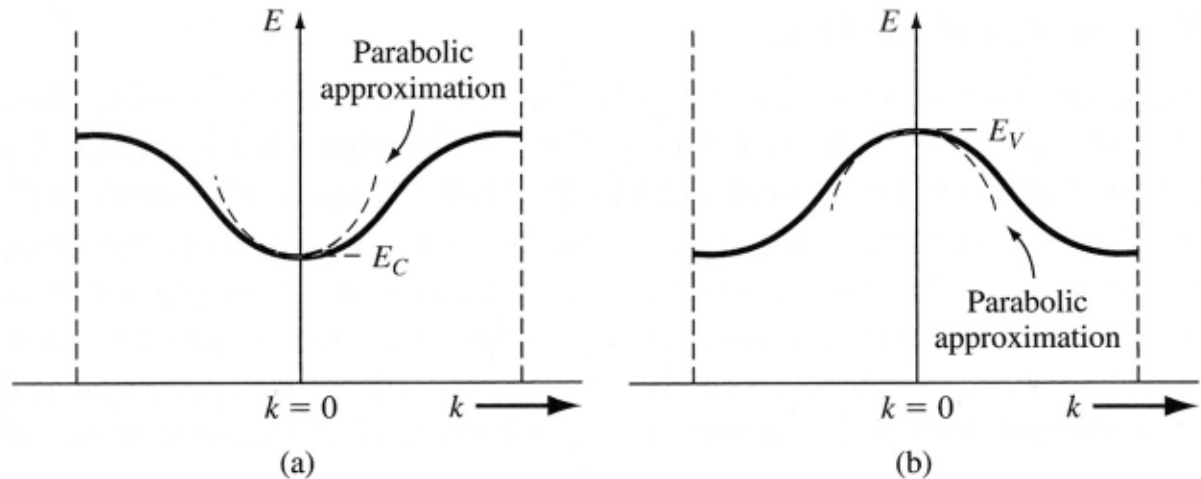
$$F_{\text{total}} = F_{\text{ext}} + F_{\text{int}} = m_0 a \quad (a = \text{Beschleunigung}).$$

Die komplizierte Wirkung von  $F_{\text{int}}$  kann man dadurch erfassen, dass man die Beschleunigung nur in Relation zu  $F_{\text{ext}}$  setzt:

$$F_{\text{ext}} = m^* a$$

Der Parameter  $m^*$  heisst **effektive Masse** und beinhaltet den Effekt aller inneren Kräfte.

Im Bandschema  $E(k)$  ist  $1/m^*$  proportional zur Krümmung des Bandes bei  $k = 0$ .



Parabolische Näherung der Dispersion (a) des Leitungsbandes und (b) des Valenzbandes für kleine  $k$ -Werte.

## Die effektive Masse des Elektrons ...

Für das kräftefreie Elektron gilt  $E(k) = \hbar^2 k^2 / (2m_0)$  und daher

$$\frac{1}{\hbar^2} \frac{d^2 E}{dk^2} = \frac{1}{m_0}$$

d.h.  $\hbar^2/m_0$  ist die konstante Krümmung ( $m_0 =$  Ruhemasse).

In Halbleitern wird die Energie in der Umgebung eines Bandminimums oft in parabolischer Form approximiert:

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

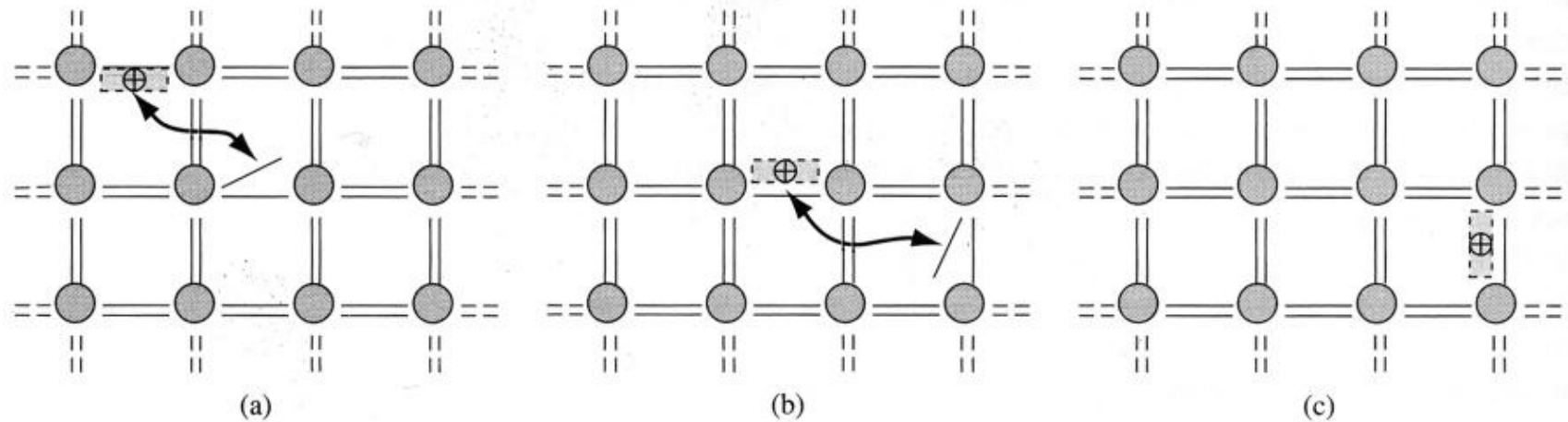
Die Krümmung ist dabei im allgemeinen verschieden von der der  $E(k)$ -Relation des kräftefreien Elektrons. Die Approximation ist in der Regel nur für kleine  $k$ -Werte (d.h. Energiewerte) gut.

# Das Lochkonzept

Der Elektronenstrom in einem fast vollen Valenzband entspricht dem Fluss positiv geladener Teilchen von einer Leerstelle zur nächsten:

$$\mathbf{J}_n = -e \sum_{i(\text{filled})} \mathbf{v}_i = -e \sum_{i(\text{total})} \mathbf{v}_i + e \sum_{i(\text{empty})} \mathbf{v}_i \stackrel{!}{=} e \sum_{i(\text{empty})} \mathbf{v}_i \quad (\alpha^3 \text{ wurde } 1 \text{ gesetzt})$$

Ein volles Valenzband liefert keinen Strom, da zu jedem  $v$  ein  $-v$  existiert ( $E(k)$  ist symmetrisch) und ein äusseres Feld keine Umverteilung der Elektronen bewirken kann, da ja alle Zustände besetzt sind.



Veranschaulichung der Löcherbewegung im Halbleiter.

# Elektrische Leitfähigkeit

Eine Kraft  $\mathbf{F}_p = e\mathbf{E}$  auf Löcher verursacht eine Löcherstromdichte

$$\mathbf{J}_p = \mu_p p \mathbf{F}_p = ep \mathbf{v}_{dp} = ep \mu_p \mathbf{E}$$

Die gesamte Feldstromdichte  $\mathbf{J}$  beträgt

$$\mathbf{J} = e(\mu_n n + \mu_p p)\mathbf{E} = \sigma \mathbf{E}$$

Dies ist das Ohmsche Gesetz des Halbleiters.

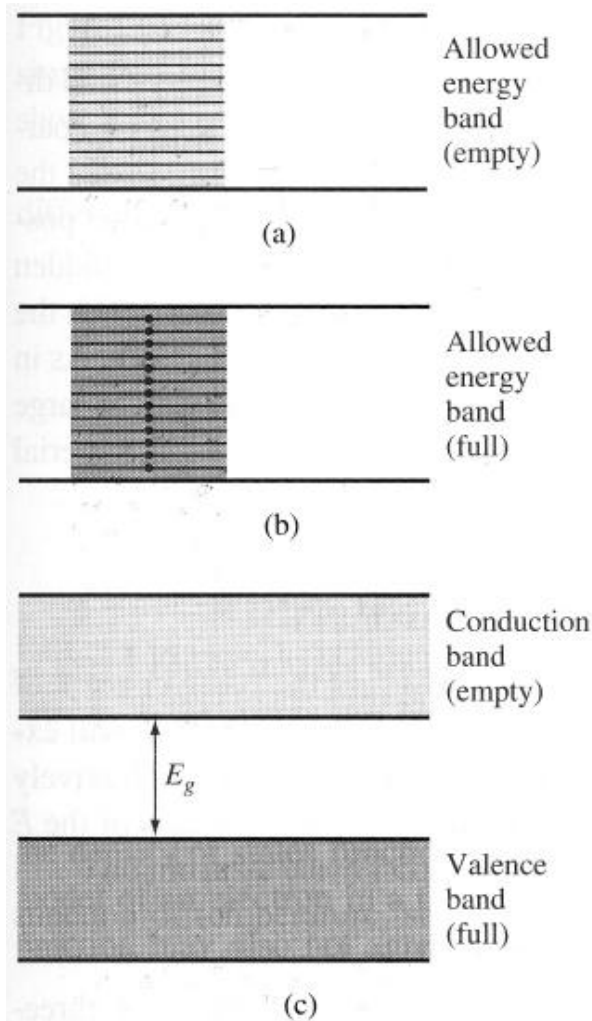
1. Zur Leitfähigkeit  $\sigma$  tragen jeweils die Produkte  $\mu_n n$  und  $\mu_p p$  bei.
2. Die Beweglichkeit ist für Elektronen und Löcher unterschiedlich.
3. Wegen der Abhängigkeit  $\mu(E)$  gilt das Ohmsche Gesetz nur für kleine Feldstärken.



# Metalle, Isolatoren, Halbleiter

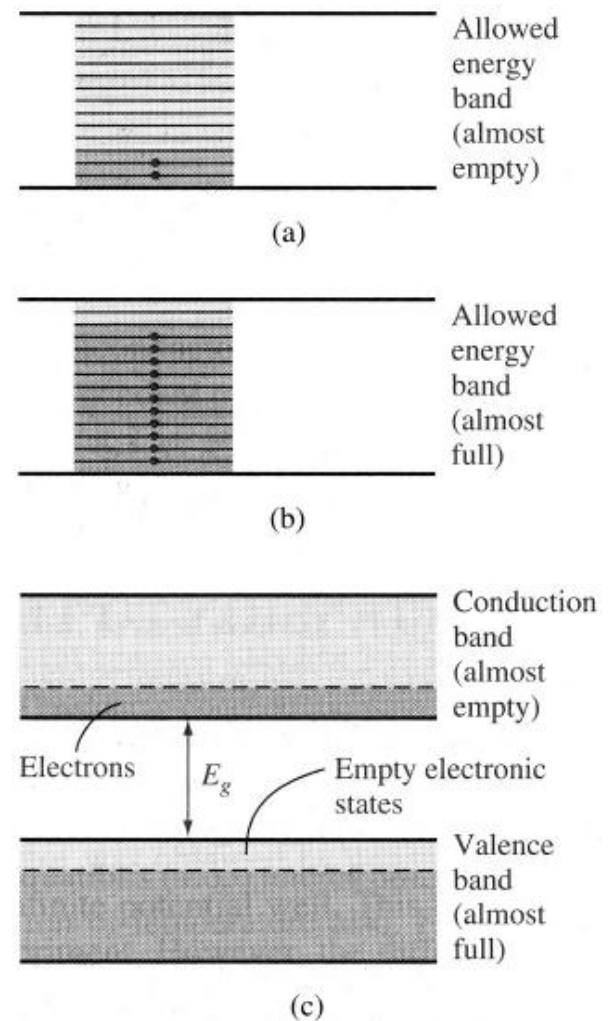
- Jeder Kristall hat seine eigene Bandstruktur  $E(\mathbf{k})$ .
- Die Bandstruktur bestimmt die elektrische Leitfähigkeit und damit die IV-Charakteristiken.
- Typische Unterschiede existieren zwischen Isolatoren, Halbleitern und Metallen.
- Isolator: Breite Energielücke (z.B.  $\text{SiO}_2$ :  $E_g \sim 9 \text{ eV}$ ); praktisch leeres Leitungsband bei 300 K.
- Halbleiter: Energielücke von Grössenordnung 1 eV; schwach gefülltes Leitungsband bei 300 K.
- Metall: Überlappung von Valenz- und Leitungsband; starke Bandfüllung, hohe Leitfähigkeit.

# Isolator



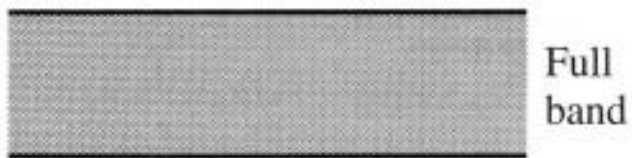
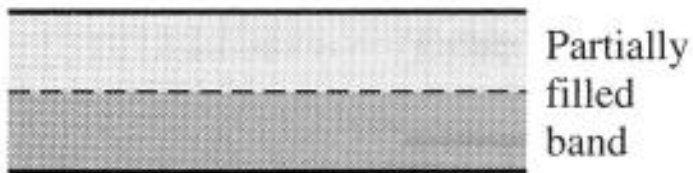
(a) Leeres Band, (b) vollständig gefülltes Band, (c) Energielücke (gap) zwischen den erlaubten Bändern.

# Halbleiter

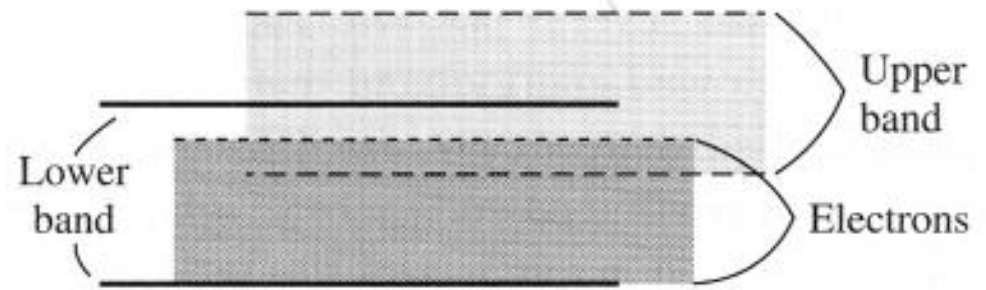


(a) Fast leeres Band, (b) fast vollständig gefülltes Band, (c) Energielücke (gap) zwischen den erlaubten Bändern.

# Metall



(a)

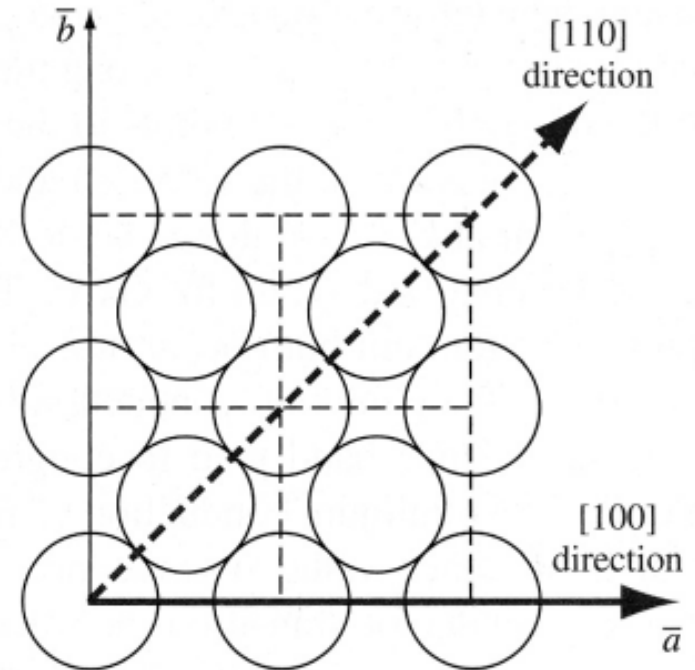


(b)

Zwei Möglichkeiten im Metall: (a) ein teilweise gefülltes Band, (b) Überlappung von erlaubten Bändern.

# 3.3 Verallgemeinerung auf den 3D-Fall

- Das Potential, das ein Elektron “spürt”, hängt von der Richtung im Kristall ab, da der Abstand zwischen den Atomen richtungsabhängig ist.
- Analogie Kronig-Penney Modell: Die Parameter  $a$  und  $b$  (Potentialtöpfe und Barrieren) hängen von der Richtung ab.
- Verschieden sind deshalb auch die Grenzen der 1. Brillouin-Zone und die effektiven Massen.



Die (100)-Ebene eines kubisch flächenzentrierten Kristalls und die [100]- und [110]-Richtungen.

# Bandverlauf und Iso-Energieflächen

Der Verlauf von  $E(k)$  lässt sich in zweierlei Weise darstellen:

1. Als Bandstruktur in der Darstellung der Energie als Funktion der Wellenzahl für bestimmte Richtungen im  $k$ -Raum.
2. Als Iso-Energiefläche, in der alle Ausbreitungsvektoren zusammengefasst werden, die zur gleichen Energie gehören.

Da nur ein teilweise gefülltes Band einen (gerichteten) Strom ergibt, interessiert uns praktisch nur der Bandverlauf des untersten **Leitungsbandes** und des obersten **Valenzbandes**.

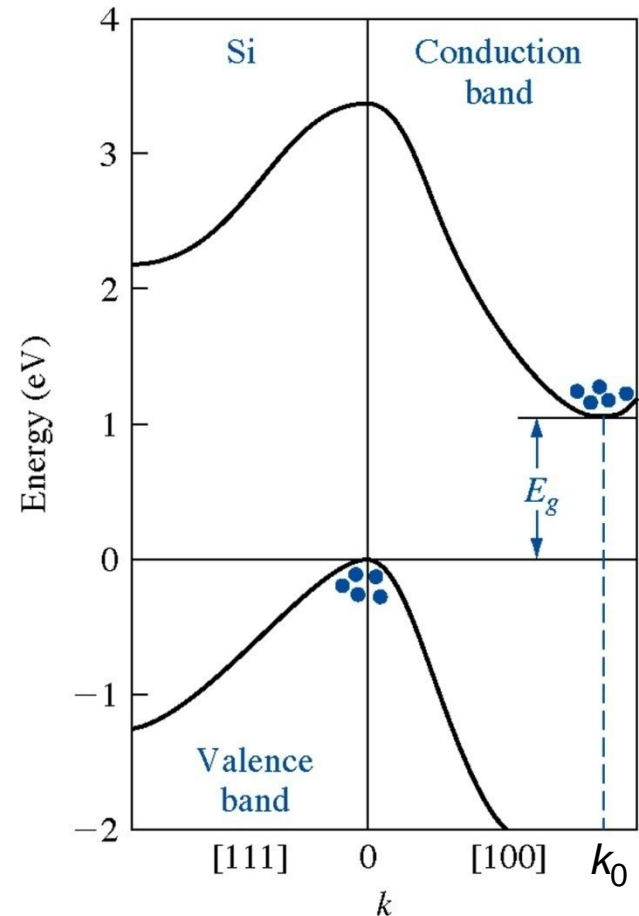
Aufgrund der energetischen Verteilungsfunktion halten sich Elektronen überwiegend an der unteren Kante des Leitungsbandes  $E_C$ , Löcher hingegen an der oberen Kante  $E_V$  des Valenzbandes auf.

# Bandstruktur von Silizium

Silizium ist ein *indirekter* Halbleiter.

**Leitungsband:** Es existieren sechs absolute Minima an der Stelle  $k_0$  in den [100] Richtungen (nämlich in  $[\pm 100]$   $[0\pm 10]$   $[00\pm 1]$ ). Die Iso-Energieflächen sind Rotationsellipsoide. Ihre verlängerten Achsen zeigen in die jeweilige  $\langle 100 \rangle$ -Richtung.

**Valenzband:** Es existieren drei Zweige des Valenzbandes, die alle ihr Minimum bei  $k = 0$  haben. Für jeden Zweig lässt sich eine Energiefläche  $E(\mathbf{k}) = \text{const}$  zeichnen, deren Krümmung mit der Richtung des Wellenvektors  $\mathbf{k}$  variiert.

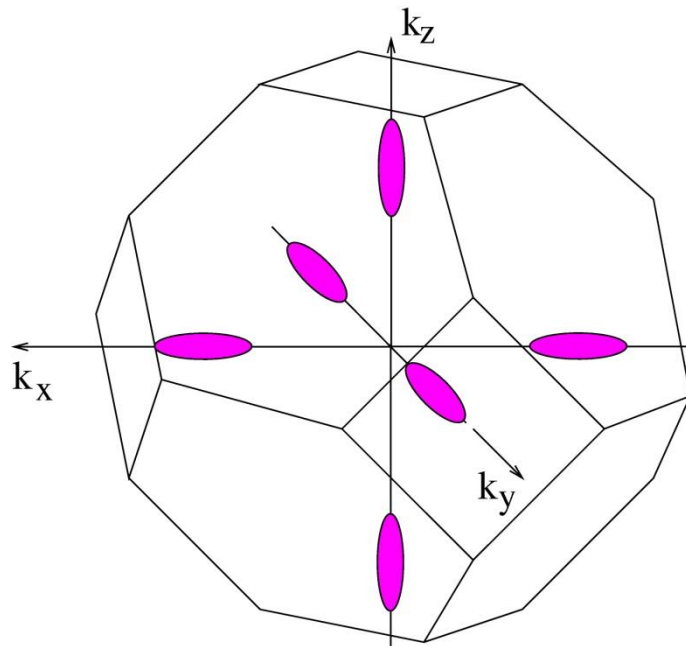


(b)

b) Reduzierte Darstellung für die [100]- und [111]-Richtungen (qualitativ).

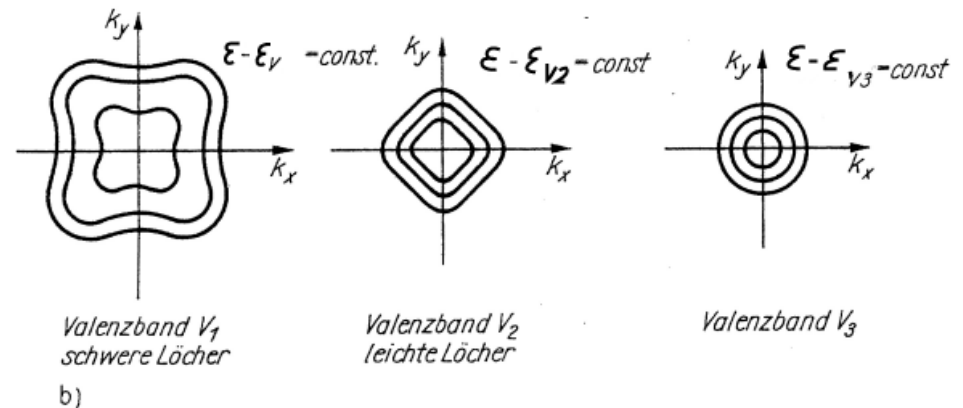
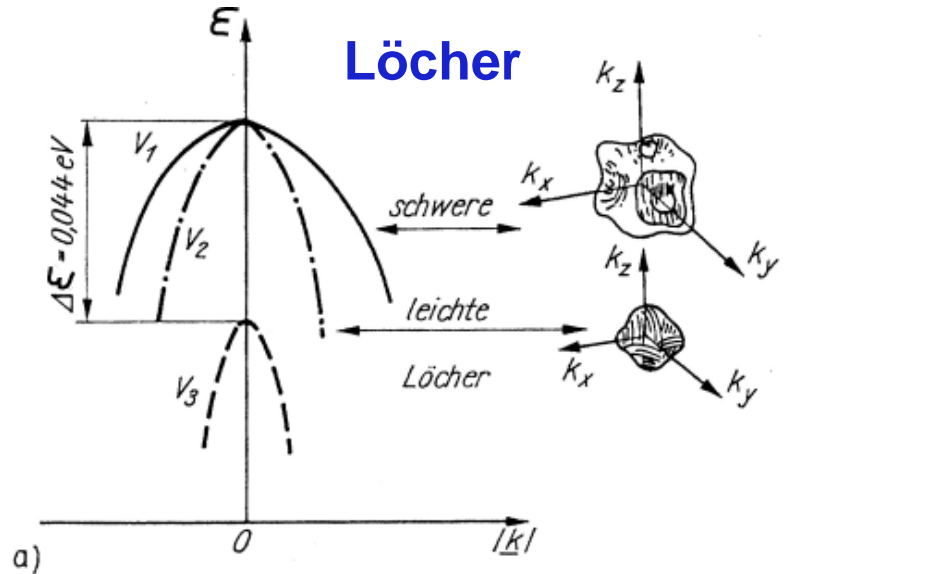
# Iso-Energieflächen von Silizium

## Elektronen



Die Iso-Energieflächen sind Rotationsellipsoide. Die schwache Krümmung längs der Achsen entspricht der grossen effektiven Masse  $m_{\text{long}} = 0.92m_0$ , die starke Krümmung senkrecht zur Achse der kleinen effektiven Masse  $m_{\text{trans}} = 0.19m_0$ .

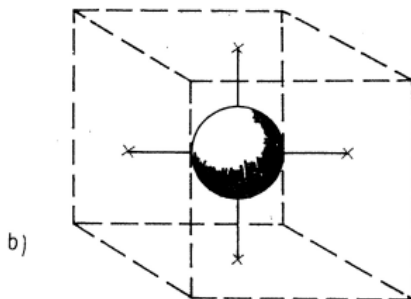
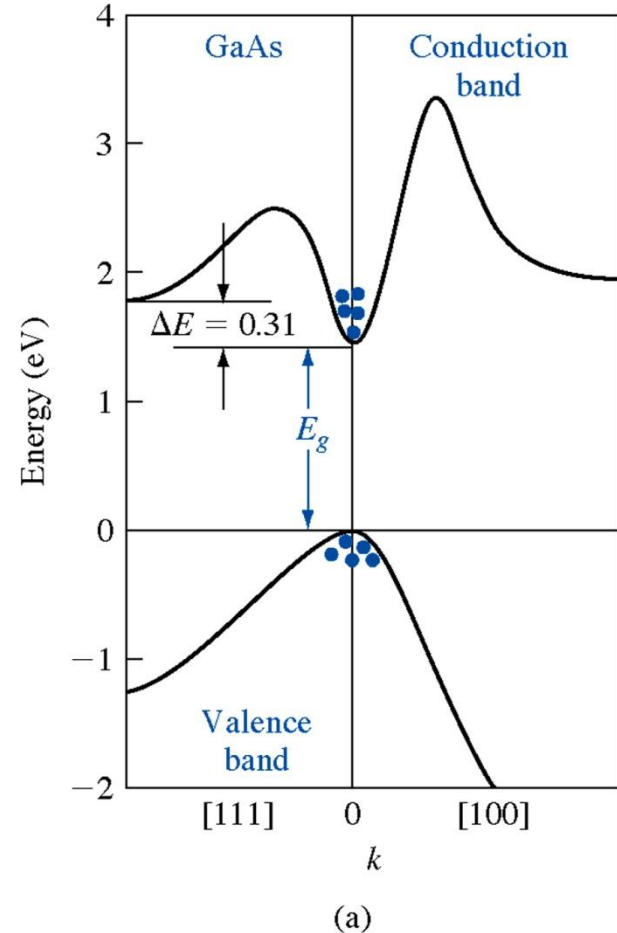
## Löcher



a) Valenzband und räumliche Darstellung der Energieflächen. b) Schnitte der Energieflächen  $E(\mathbf{k}) = \text{const}$  mit der Ebene  $k_z = 0$  für die drei Valenzbänder.

# Bandstruktur von GaAs

Dieses Material ist ein *direkter* Halbleiter. Es weist eine Besonderheit im Leitungsband auf: neben dem Hauptminimum existiert noch ein Nebenminimum entlang der [111]-Richtung. Seine Unterkante liegt um  $\Delta E = 0.31$  eV über dem absoluten Leitungsbandminimum bei  $\mathbf{k} = 0$ . Im Nebenminimum ist die effektive Masse der Elektronen grösser. Es gibt "leichte" ( $m_n = 0.07m_0$ ) und "schwere" ( $m_n = 0.4m_0$ ) Elektronen.



- a) Reduzierte Darstellung für die [100]- und [111]-Richtungen (qualitativ).
- b) Iso-Energiefläche  $E(\mathbf{k}) = \text{const}$  bei isotroper effektiver Masse.



# 3.4 Die Zustandsdichte

## Ableitung

- Teilchen im dreidimensionalen Kasten der Kantenlänge  $a$

$$\frac{2mE}{\hbar^2} = k^2 = k_x^2 + k_y^2 + k_z^2 = (n_x^2 + n_y^2 + n_z^2) \left( \frac{\pi^2}{a^2} \right)$$

- Abstand zweier Zustände

$$k_{x+1} - k_x = (n_x + 1) \left( \frac{\pi}{a} \right) - n_x \left( \frac{\pi}{a} \right) = \frac{\pi}{a}$$

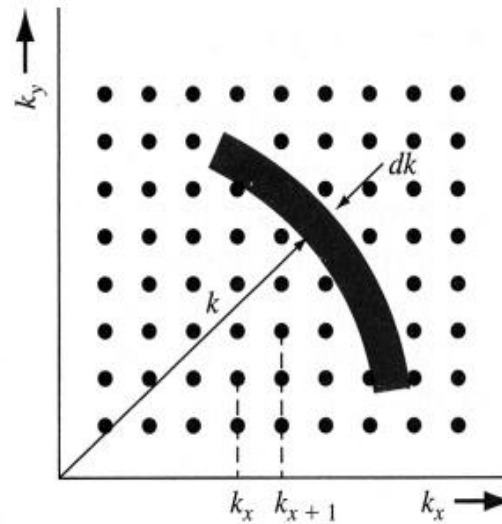
- Differentielle Zustandsdichte im  $k$ -Raum

$$g(k)dk = 2 \frac{1}{8} \frac{4\pi k^2}{\left( \frac{\pi}{a} \right)^3} dk$$

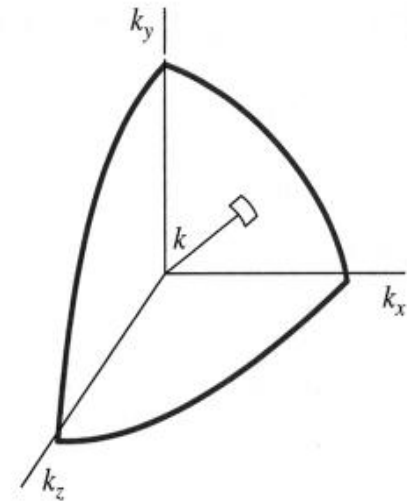
- Zustandsdichte auf Energieskala

$$g(E) = \frac{4\pi(2m)^{3/2}}{h^3} \sqrt{E}$$

Doppeldichte; Masseinheit  $1/(\text{eV cm}^3)$



(a)



(b)

(a) Zweidimensionale Anordnung von erlaubten Quantenzuständen im  $k$ -Raum. (b) Das positive Achteil des kugelförmigen  $k$ -Raums.

■ Annahme: parabolisches Band

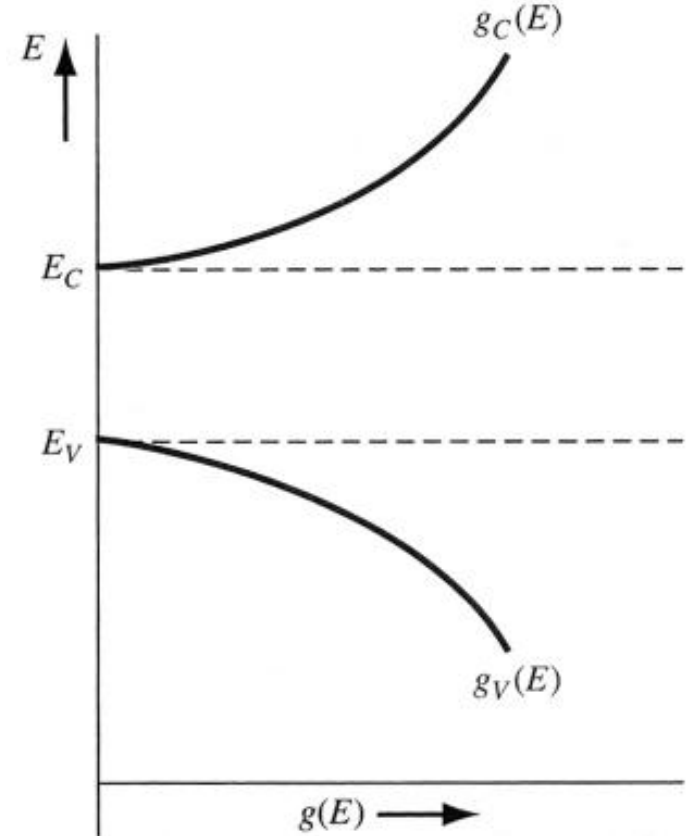
$$E = E_C + \frac{\hbar^2 k^2}{2m_n^*} \Rightarrow E - E_C = \frac{\hbar^2 k^2}{2m_n^*}$$

■ Zustandsdichte im Leitungsband

$$g_C(E) = \frac{4\pi(2m_n^*)^{3/2}}{h^3} \sqrt{E - E_C} \text{ für } E \geq E_C$$

■ Zustandsdichte im Valenzband

$$g_V(E) = \frac{4\pi(2m_p^*)^{3/2}}{h^3} \sqrt{E_V - E} \text{ für } E \leq E_V$$



Die Dichte der Zustände im Leitungs- und Valenzband als Funktion der Energie.

**Effektive Zustandsdichten.** Für die spätere Darstellung der Elektronen- und Löcherdichten definiert man zweckmässigerweise sog. effektive Zustandsdichten  $N_C$  und  $N_V$  mit derselben Masseinheit ( $\text{cm}^{-3}$ ).

*Isotropes parabolisches Band*

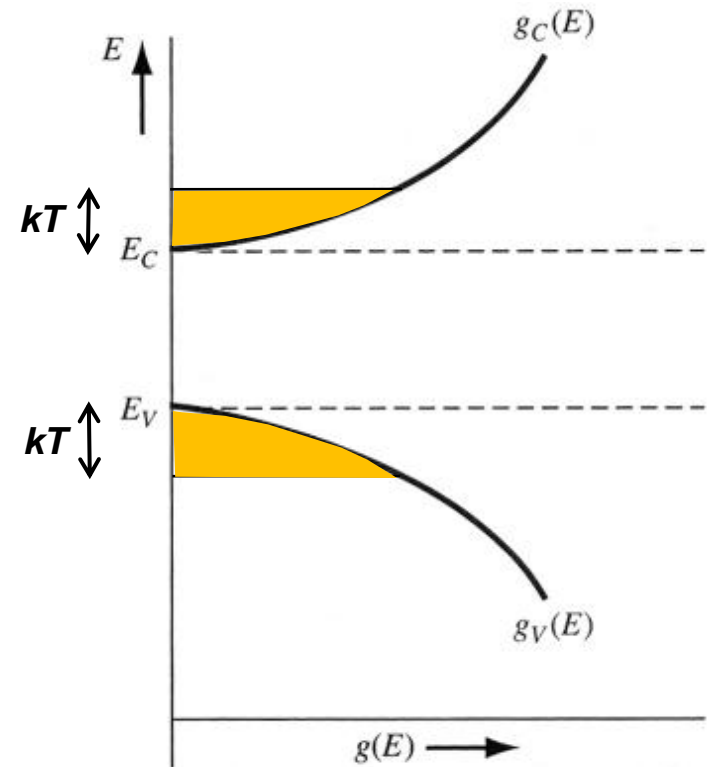
$$N_C = 2 \left( \frac{2\pi m_n kT}{h^2} \right)^{3/2} = \frac{\sqrt{\pi}}{(2\pi)^2} \left( \frac{2m_n kT}{\hbar^2} \right)^{3/2} = 2.4 \times 10^{19} \left( \frac{m_n}{m_0} \frac{T}{T_0} \right)^{3/2} \text{ cm}^{-3}$$

$$N_V = 2 \left( \frac{2\pi m_p kT}{h^2} \right)^{3/2} = 2.4 \times 10^{19} \left( \frac{m_p}{m_0} \frac{T}{T_0} \right)^{3/2} \text{ cm}^{-3}$$

$$g_C(E) dE = N_C \frac{2}{\sqrt{\pi}} \left( \frac{E - E_C}{kT} \right)^{1/2} d\left( \frac{E}{kT} \right),$$

$$g_V(E) dE = N_V \frac{2}{\sqrt{\pi}} \left( \frac{E_V - E}{kT} \right)^{1/2} d\left( \frac{E}{kT} \right)$$

Die effektive Zustandsdichte  $N_{C,V}$  beschreibt damit anschaulich die Dichte der besetzbaren Zustände innerhalb des Energieintervalls  $\Delta E = kT$  von der Bandkante aus gemessen. Sie liegt in der Grössenordnung von  $10^{19} \text{ cm}^{-3}$ .



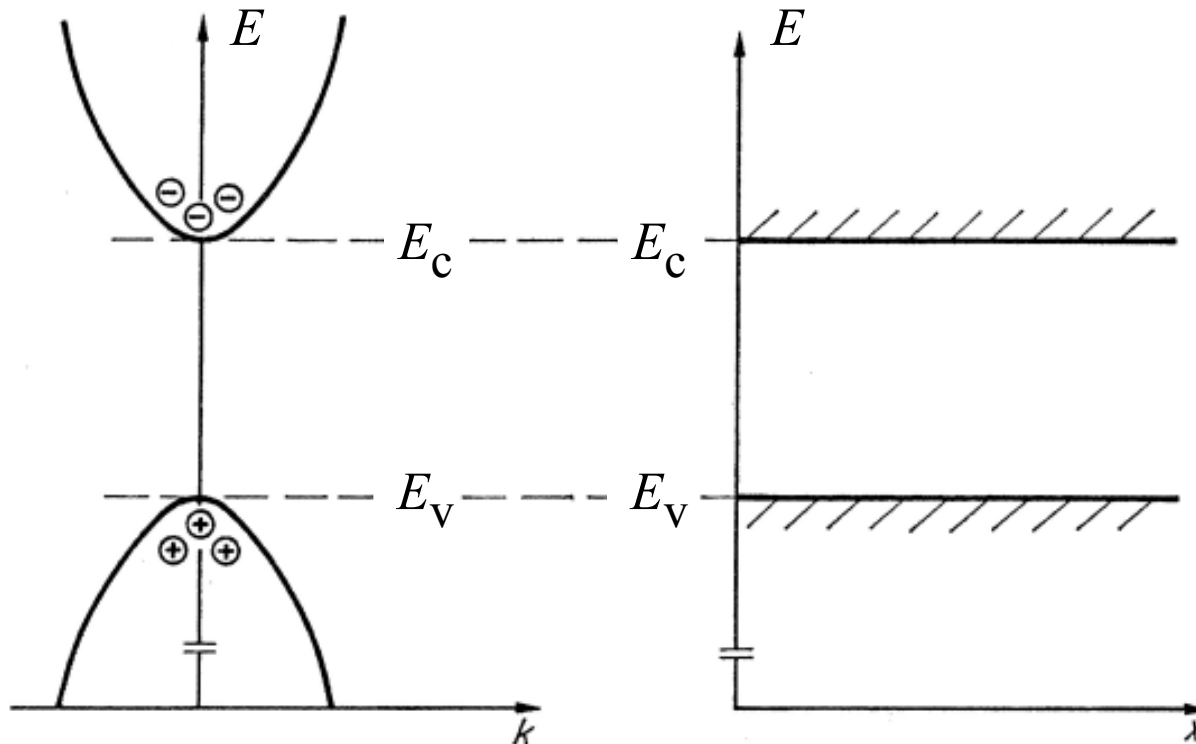
# Bandstruktur und Bandkantendiagramm

Die Energie der Elektronen als Funktion der Wellenzahl  $k$  entlang einer Richtung im  $k$ -Raum.

Die Energie der Bandkanten als Funktion des Ortes  $x$  entlang einer Richtung im Halbleiter oder Bauelement.

Bandstruktur  $E_{c,v}(k)$

Bandkantendiagramm  $E_{c,v}(x)$



# 3.5 Statistische Mechanik

## Die Fermi-Dirac Verteilung

$$n(E) = g(E) f_F(E)$$

Zahl der Elektronen  
pro Energieintervall  
 $[E, E+dE]$

Zustandsdichte im  
Energieintervall  
 $[E, E+dE]$

Wahrscheinlichkeit, dass  
ein Quantenzustand bei  
der Energie  $E$  mit einem  
Elektron *besetzt* ist

### 3 Arten von Verteilungen:

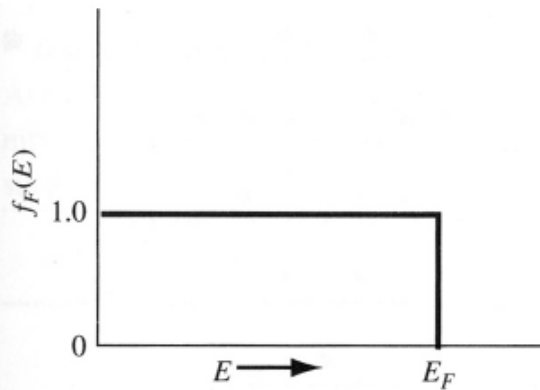
- Unterscheidbar  
⇒ Maxwell-Boltzmann Statistik
- Ununterscheidbar, kein Pauli-Prinzip  
⇒ Bose-Einstein Statistik
- Ununterscheidbar, Pauli-Prinzip  
⇒ Fermi-Dirac Statistik

# Die Fermi-Dirac Verteilung

$$f_F(E) = \frac{1}{1 + \exp\left(\frac{E - E_F}{kT}\right)}$$

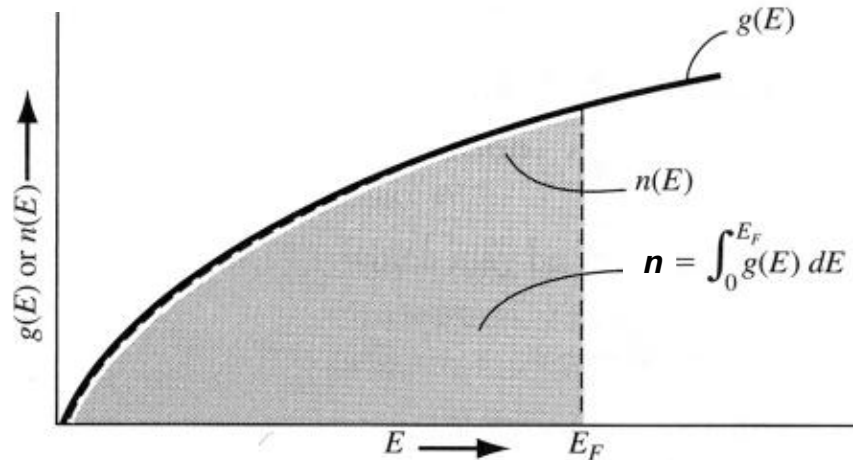
- Wahrscheinlichkeit, dass ein Zustand mit der Energie  $E$  **besetzt** ist.
- Gilt für das thermodynamische Gleichgewicht.
- Pauli-Prinzip: Jeder Quantenzustand kann nur mit einem Elektron (Fermion) besetzt werden. Daher muss die mittlere Besetzungszahl zwischen Null und Eins liegen.
- Auftreten der Temperatur: Das Elektron ist an ein Wärmebad der Temperatur  $T$  angekoppelt (ungehinderter Energieaustausch).
- Auftreten der Fermi-Energie  $E_F$ : Das Elektron befindet sich in einem Teilchenreservoir, das durch das chemische Potential  $\mu_C = E_F$  charakterisiert ist.
- Das chemische Potential ist diejenige (chemische) Energie, die nötig ist, um einem Elektronensystem ein weiteres Elektron hinzuzufügen.

# Die Fermi-Energie



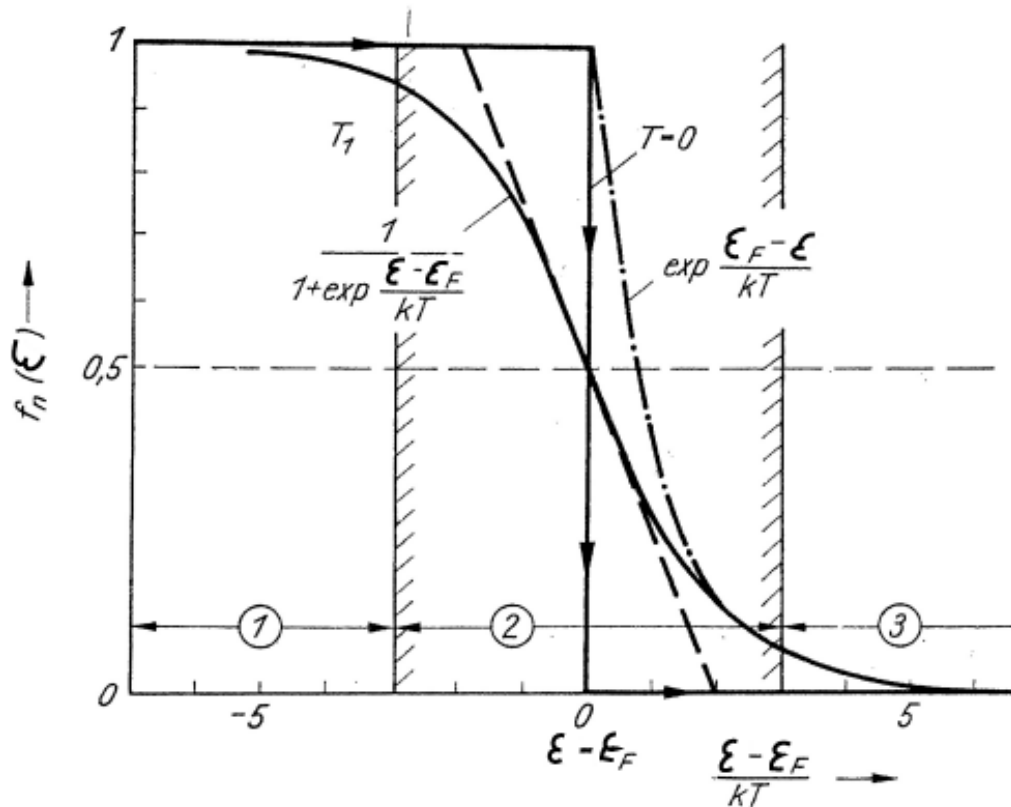
Die Fermi-Verteilungsfunktion bei  $T = 0$  K.

- Bei  $T = 0$  K sind alle Zustände bis zur Fermi-Energie  $E_F$  besetzt, darüber leer. Die Fermi-Dirac Verteilung entartet zur Stufenfunktion.
- Die Wahrscheinlichkeit für die Besetzung eines Zustands mit  $E \equiv E_F$  ist  $\frac{1}{2}$  bei jeder Temperatur.
- $E_F$  ist ein Parameter des statistischen Ensembles und muss nicht mit einer erlaubten Energie  $E$  zusammenfallen.



Die Dichte der Quantenzustände und die Dichte der Elektronen in einem System mit kontinuierlicher Energie bei  $T = 0$  K.

## Fermi-Dirac Verteilungsfunktion



- Gilt  $E > E_F$ , und  $|E - E_F| \gg kT$ , so kann man die "1" im Nenner vernachlässigen und  $f_F(E)$  geht in die Maxwell-Boltzmann Verteilung über.
- Die Elektronen sind dann natürlich immer noch ununterscheidbar, aber ihre Dichte ist so klein, dass die Verteilungsfunktion in die von unterscheidbaren Teilchen übergeht.

Im Bereich 3 stimmt die Maxwell-Boltzmann-Verteilung sehr gut mit der Fermi-Dirac-Verteilung überein. Im Bereich 2 wird sie als Näherung zunehmend falsch.



# Eigenschaften wichtiger Halbleiter

Halbleiter	Bandstruktur <sup>1)</sup>	Bandabstand		Beweglichkeit bei 300 K <sup>2)</sup>		Effektive Masse <sup>3)4)</sup>		Dielektrizitätskonstante	
		eV	eV	Elektronen	Löcher	$\frac{m^*}{m_0}$	$\frac{m^*}{m_0}$		
Elementgruppe IV		300 K	0 K	Elektronen	Löcher	Elektronen	Löcher		
C (Diamant)	Si	5,47	5,51	1800	1600	0,2	0,25	5,5	
Ge	Ge	0,67	0,69	3900	1900	$m_l^* = 1,58$ $m_t^* = 0,082$	$m_l^* = 0,04$ $m_t^* = 0,3$	16	
Si	Si	1,12	1,16	1500	600	$m_l^* = 0,98$ $m_t^* = 0,19$	$m_l^* = 0,16$ $m_t^* = 0,5$	11,8	
IV-IV	SiC	Si	3	3,1	400	50	0,6	1	10
III-V	AlSb	GaAs	1,63	1,75	200	420	0,3	0,4	11
	GaSb	GaAs	0,67	0,8	4000	1400	0,047	0,5	15
	GaAs	GaAs	1,43	1,52	8500	400	0,068	0,5	10,9
	GaP	Si	2,24	2,40	110	75	0,5	0,5	10
	InSb	GaAs	0,16	0,26	78000	750	0,013	0,6	17
	InAs	GaAs	0,33	0,46	33000	460	0,02	0,41	14,5
	InP	GaAs	1,29	1,34	4600	150	0,07	0,4	14
II-VI	CdS	GaAs	2,42	2,56	300	50	0,17	0,6	10
	CdSe	GaAs	1,7	1,85	800		0,13	0,45	10
	ZnO	GaAs	3,2		200		0,27		
	ZnS	GaAs	3,6	3,7	165		1,1		8
IV-VI	PbS		0,41	0,39	600	700	0,66	0,5	17
	PbTe	GaAs	0,32	0,24	6000	4000	0,22	0,29	30

1) Ge, Si, GaAs bedeutet diesen Materialien ähnliche Bandstruktur.

2) Driftbeweglichkeit

3)  $m_t^*$ ,  $m_l^*$  effektive transversale bzw. longitudinale Masse

4) Werte bezogen auf die Masse  $m$  des Elektrons.

# Eigenschaften von Si, Ge, und GaAs bei T = 300 K

Eigenschaften	Ge	Si	GaAs	
Effektive Zustandsdichte im Leitband $N_C$ [ $\text{cm}^{-3}$ ]	$1,04 \cdot 10^{19}$	$2,8 \cdot 10^{19}$	$4,7 \cdot 10^{17}$	
Effektive Zustandsdichte im Valenzband $N_V$ [ $\text{cm}^{-3}$ ]	$6,1 \cdot 10^{18}$	$1,02 \cdot 10^{19}$	$7,0 \cdot 10^{18}$	
Elektronenaffinität [eV]	4,0	4,05	4,07	
Bandbreite bei 300 K [eV]	0,67	1,12	1,43	
Austrittsarbeit [eV]	4,4	4,8	4,7	
Dielektrizitätskonstante	16	11,8	10,9	
Durchbruchfeldstärke [V/cm]	$10^5$	$3 \cdot 10^5$	$4 \cdot 10^5$	
Intrinsidichte [ $\text{cm}^{-3}$ ]	$2,33 \cdot 10^{13}$	$1,6 \cdot 10^{10}$	$1,3 \cdot 10^6$	
Beweglichkeit [ $\text{cm}^2/\text{Vs}$ ]				
$\mu_n$	3900	1500	8500	
$\mu_p$	1900	600	400	
Atome/ $\text{cm}^3$	$4,42 \cdot 10^{22}$	$5,0 \cdot 10^{22}$	$2,21 \cdot 10^{22}$	
Atomgewicht	72,6	28,08	144,63	
Gitterstruktur	Diamant	Diamant	Zinkblende	
Effektive Masse $m^*/m_0$				
Elektronen	$m_1^* = 1,58$	$m_t^* = 0,082$	$m_1^* = 0,98$ $m_t^* = 0,19$	0,068
Löcher	$m_1^* = 0,04$	$m_h^* = 0,3$	$m_1^* = 0,16$ $m_h^* = 0,5$	0,12, 0,5
Gitterkonstante [Å]	5,65	5,43	5,65	
Linearer thermischer Ausdehnungskoeffizient $\Delta L/\Delta T$ [ $^\circ\text{C}^{-1}$ ]	$5,8 \cdot 10^{-6}$	$2,6 \cdot 10^{-6}$	$5,9 \cdot 10^{-6}$	
Spezifische Wärme [Joule/g $^\circ\text{C}$ ]	0,31	0,7	0,35	
Thermische Leitfähigkeit [W/cm $^\circ\text{C}$ ]	0,64	1,45	0,46	
Schmelzpunkt [ $^\circ\text{C}$ ]	937	1420	1235	
Dampfdruck [Torr]	$10^{-3}$ bei $127^\circ\text{C}$ $10^{-8}$ bei $810^\circ\text{C}$	$10^{-3}$ bei $1600^\circ\text{C}$ $10^{-8}$ bei $930^\circ\text{C}$	1 bei $1050^\circ\text{C}$ 100 bei $1220^\circ\text{C}$	







# 4. Der Halbleiter im Gleichgewicht

4.1 Ladungsträger in Halbleitern (N4.1)

4.2 Dotieratome und ihre Energieniveaus (N4.2)

4.3 Der extrinsische Halbleiter (N4.3)

4.4 Statistik von Donatoren und Akzeptoren (N4.4)

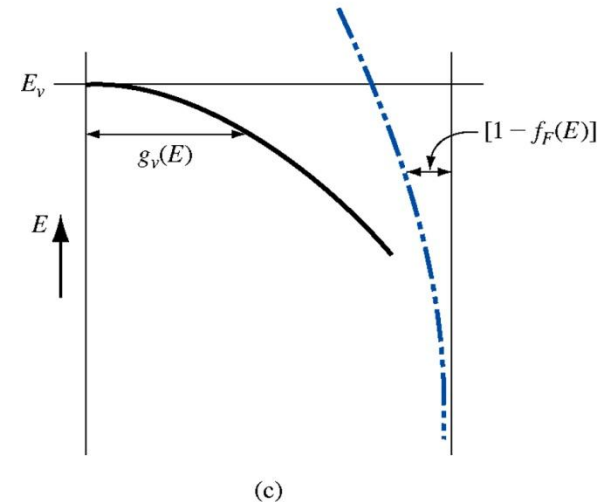
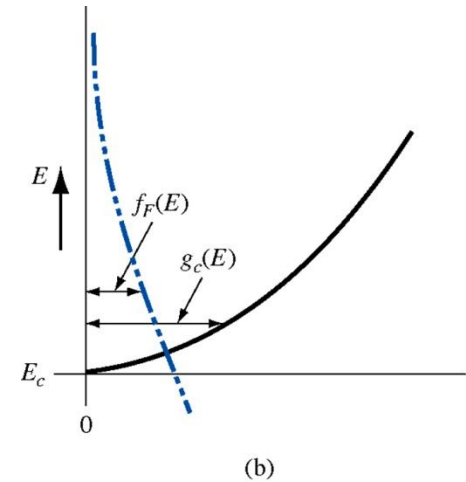
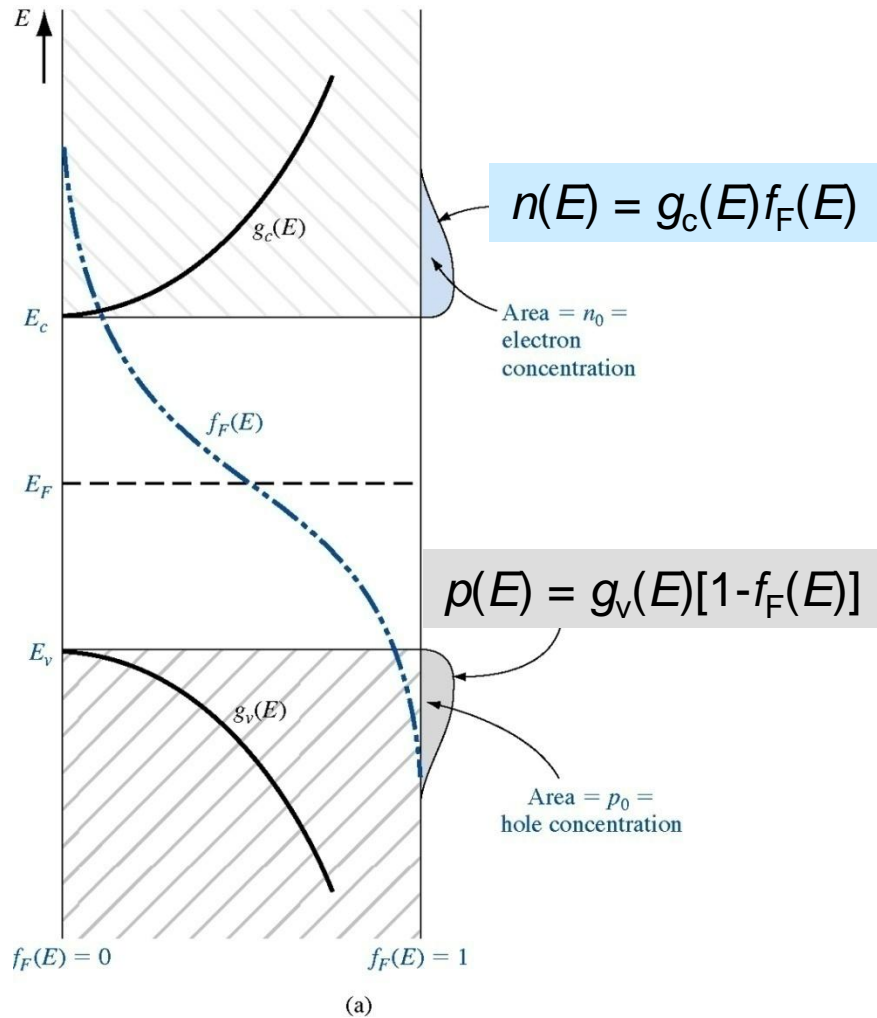
4.5 Ladungsneutralität (N4.5)

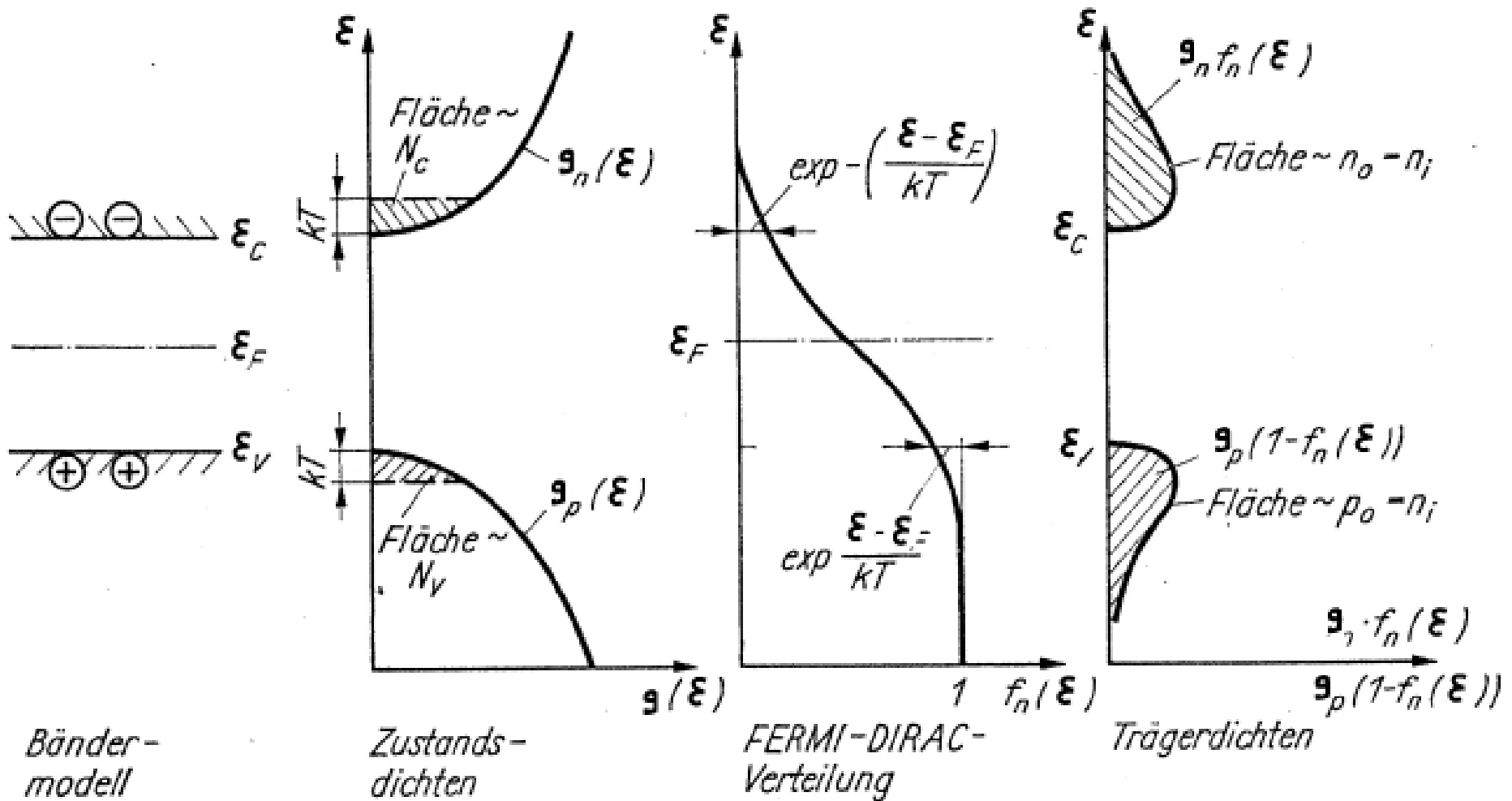
4.6 Lage des Fermi-niveaus (N4.6)

# 4.1 Ladungsträger in Halbleitern

Die Trägerdichte im Leitungs- und im Valenzband  
(undotierter Halbleiter)

hier:  
 $m_n^* = m_p^*$   
 $E_F$  midgap  
 da  $n_0 = p_0$







**Elektronendichte im Leitungsband:** Die gesamte Gleichgewichtsdichte  $n_0$  im Leitungsband lässt sich bei Kenntnis der Zustandsdichte  $g_c(E)$  und der Besetzungsfunktion  $f_n(E)$  als Integral über alle erlaubten Energiewerte angeben:

$$n_0 = \int_{E_C}^{\infty} g_c(E) f_n(E) dE$$

Mit den effektiven Zustandsdichten des isotropen parabolischen Bandes kann man das umschreiben

$$\frac{n_0}{N_C} = \frac{2}{\sqrt{\pi}} \int_0^{\infty} \frac{\sqrt{\eta_n} d\eta_n}{\exp(\eta_n - y_n) + 1} = F_{1/2}(y_n)$$

$$\eta_n = \frac{E - E_C}{kT} \quad y_n = \frac{E_F - E_C}{kT}$$

Dabei ist  $F_{1/2}$  das *Fermi-* Integral. **Löcherdichte im Valenzband:** Die Löcherdichte  $p_0$  im Valenzband lässt sich durch Integration über die nicht von Elektronen, sondern von Löchern besetzten Zustände bestimmen:

$$p_0 = \int_{-\infty}^{E_V} g_v(E) f_p(E) dE$$

Mit den effektiven Zustandsdichten kann man das wiederum wie oben umschreiben

$$\frac{p_0}{N_V} = \frac{2}{\sqrt{\pi}} \int_0^{\infty} \frac{\sqrt{\eta_p} d\eta_p}{\exp(\eta_p - y_p) + 1} = F_{1/2}(y_p)$$

$$\eta_p = \frac{E_V - E}{kT} \quad y_p = \frac{E_V - E_F}{kT}$$

# Diskussion

1. Die untere Integrationsgrenze für die Energie der Elektronen ist die Leitungsbandkante. Die obere Integrationsgrenze für die Energie der Löcher ist die Valenzbandkante.
2. Die Integrationsgrenzen „ $\infty$ “ sind aus mathematischer Bequemlichkeit gewählt. Dies ist möglich, weil die Verteilungsfunktion für grosse Energien  $E$  exponentiell gegen Null geht. Daher trägt der Integrand  $n(E)$  bzw.  $p(E)$  dort nichts mehr bei.
3. Die Zustandsdichte  $g(E)$  hängt bei isotropen und anisotropen *parabolischen* Bändern wurzelförmig von der Energie ab. Elektronen:  $\sim (E - E_c)^{1/2}$ , Löcher:  $\sim (E_v - E)^{1/2}$ .
4. Die dimensionsbehafteten Vorfaktoren sind die *effektiven Zustandsdichten* des jeweiligen Bandes. (Erinnerung: Integration über die thermische Energie  $kT$ )

**Table 4.1** | Effective density of states function and effective mass values

	$N_c$ (cm <sup>-3</sup> )	$N_v$ (cm <sup>-3</sup> )	$m_n^*/m_0$	$m_p^*/m_0$
Silicon	$2.8 \times 10^{19}$	$1.04 \times 10^{19}$	1.08	0.56
Gallium arsenide	$4.7 \times 10^{17}$	$7.0 \times 10^{18}$	0.067	0.48
Germanium	$1.04 \times 10^{19}$	$6.0 \times 10^{18}$	0.55	0.37

# Der Grenzfall der Nichtentartung

Hier gilt:  $y$  ist negativ und betragsmässig gross, d.h.  $E_C - E_F \gg kT$  bzw.  $E_F - E_V \gg kT$ . Die Fermienergie liegt mehrere  $kT$  unterhalb des Leitungsbandes bzw. oberhalb des Valenzbandes, also im Inneren der verbotenen Zone. Die Besetzungsfunktion  $f(E)$  lässt sich durch die Maxwell-Boltzmann-Verteilung annähern

$$f_{\text{MB}}(E) = \exp\left(-\frac{E - E_F}{kT}\right)$$

Das Fermi-Integral geht dann in die einfache e-Funktion über

$$F_{1/2}(y) \rightarrow \exp(y)$$

Damit betragen die Elektronen- und Löcherdichte

$$\frac{n_0}{N_C} = \frac{2}{\sqrt{\pi}} \int_0^\infty \frac{\sqrt{\eta_n} d\eta_n}{\exp(\eta_n - y_n)} = \exp y_n = \exp\left[-\frac{E_C - E_F}{kT}\right] \quad \frac{p_0}{N_V} = \exp\left[-\frac{E_F - E_V}{kT}\right]$$

Solche Halbleiter nennt man **nichtentartet**, und es gilt

$$F_{1/2}(y_n) \ll 1 \quad F_{1/2}(y_p) \ll 1 \\ n_0 \ll N_C \approx 10^{19} \text{ cm}^{-3} \quad p_0 \ll N_V \approx 10^{19} \text{ cm}^{-3}$$

# Der Grenzfall der Entartung

Hier gilt:  $y > 0$  und  $E_F - E_C \gg kT$  bzw.  $E_V - E_F \gg kT$ . Die Fermienergie liegt im Leitungsband (n-Halbleiter) oder im Valenzband (p-Halbleiter). Liegt das Fermi-niveau weiter als  $3kT$  im jeweiligen Band, dann wird die Besetzungswahrscheinlichkeit bereits innerhalb des erlaubten Bandes  $\approx 1$ , was charakteristisch ist für Metalle. Hochdotierte Halbleiter (= Halbleiter mit entarteten Trägerdichten) zeigen daher metallisches Verhalten.

Für grosse  $y$ -Werte (starke Entartung) entartet die Fermi-Dirac-Verteilung zur Stufenfunktion und das Fermi-Integral wird

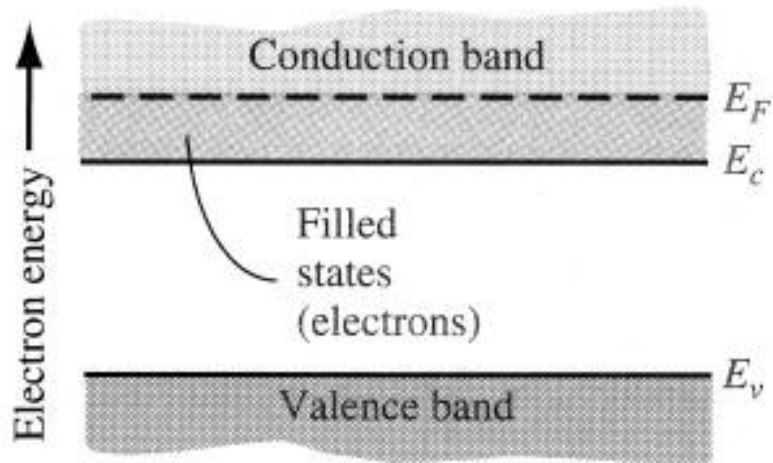
$$F_{1/2}(y_n) \rightarrow \frac{2}{\sqrt{\pi}} \int_0^{y_n} \sqrt{\eta_n} d\eta_n = \frac{2}{\sqrt{\pi}} \frac{2}{3} y_n^{3/2}$$

Die Elektronen- bzw. Löcherkonzentrationen ergeben sich zu

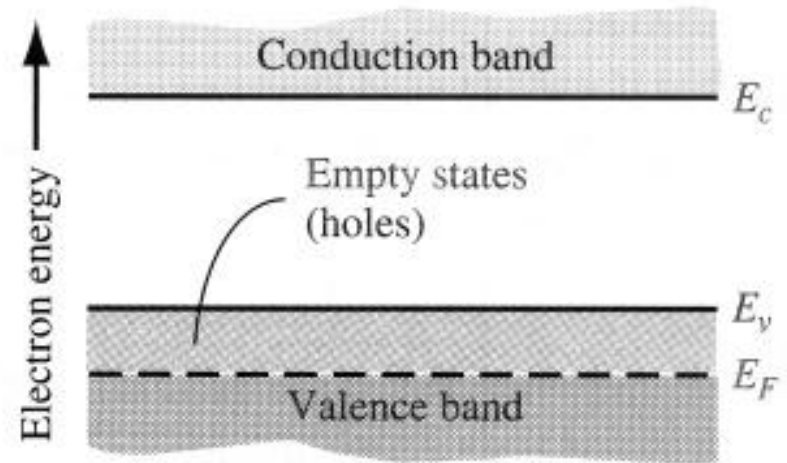
$$\frac{n_0}{N_C} = \frac{2}{\sqrt{\pi}} \int_0^{y_n} \sqrt{\eta_n} d\eta_n = \frac{4}{3\sqrt{\pi}} \left( \frac{E_F - E_C}{kT} \right)^{3/2} \quad \frac{p_0}{N_V} = \frac{4}{3\sqrt{\pi}} \left( \frac{E_V - E_F}{kT} \right)^{3/2}$$

$$n_0 \gg N_C \approx 10^{19} \text{cm}^{-3} \quad p_0 \gg N_V \approx 10^{19} \text{cm}^{-3}$$

## Bandkantendiagramm für entartete Halbleiter

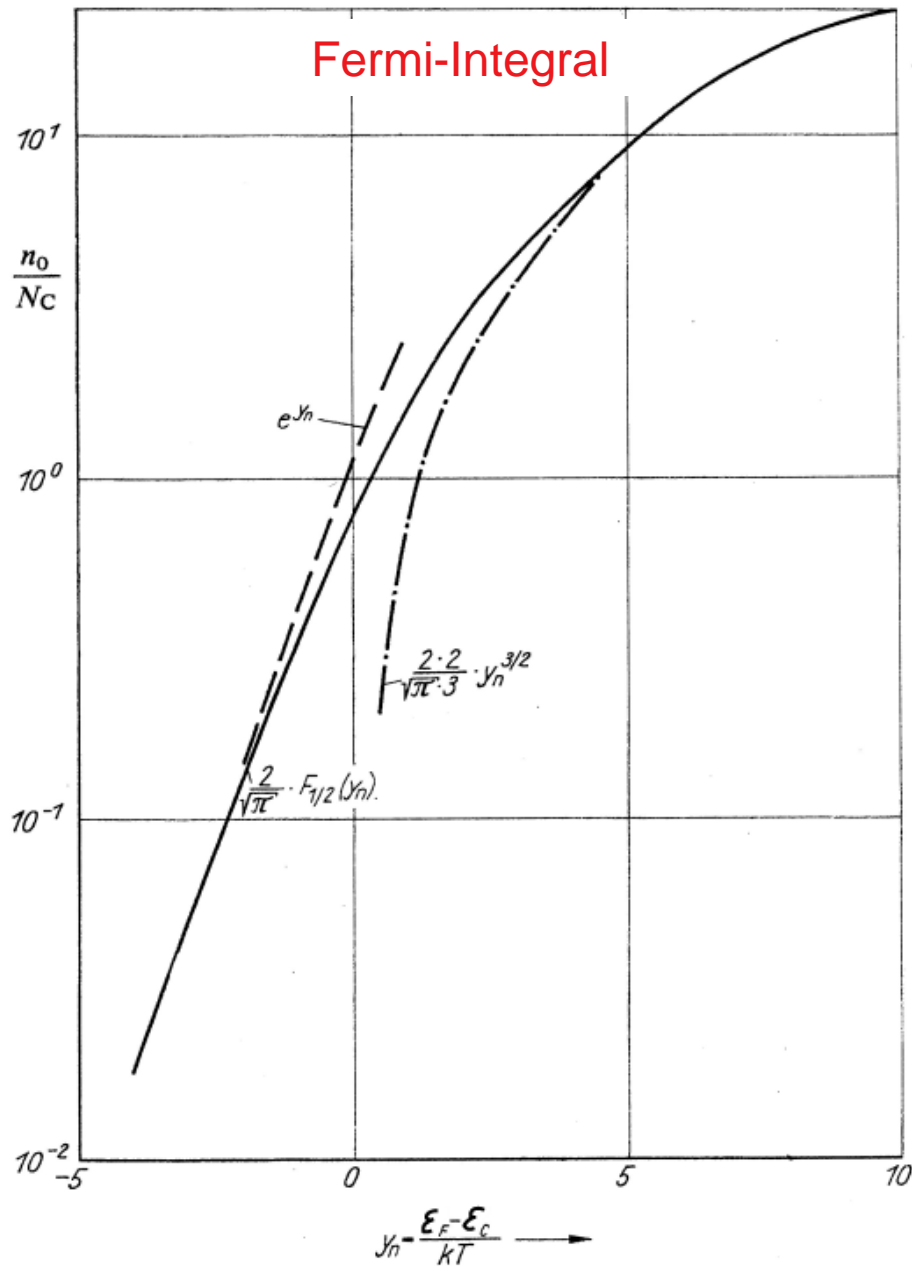


(a)



(b)

Vereinfachtes Bandkantendiagramm für entartete Halbleiter: (a) n-Typ, (b) p-Typ.



Fermi-Integral  $F_{1/2}(y_n)$  und Näherungen für negative und positive  $y$ -Werte. Gestrichelt: Nichtentartungsnäherung, punkt-gestrichelt: Entartungsnäherung.

# Feststellungen

1. Bei nichtentarteten Halbleitern überschreitet die Dichte beweglicher Ladungsträger (also Elektronen oder Löcher) die effektive Zustandsdichte bei Raumtemperatur ( $\approx 10^{19} \text{ cm}^{-3}$ ) nicht. Das Fermi-niveau liegt innerhalb der verbotenen Zone. Oft arbeiten Halbleiterbauelemente mit Trägerdichten im Bereich der Nichtentartung.
2. Nichtentartung erfordert einen Abstand von mindestens  $3kT$  zwischen Fermi-niveau und Bandkante. Ist dies für ein Band nicht der Fall, so nennt man den Halbleiter „in einem Band entartet“.
3. Entartete Halbleiter mit Trägerkonzentrationen  $n, p > N_C, N_V$  (d.h.  $n, p \approx 10^{19} \dots 10^{21} \text{ cm}^{-3}$ ) haben metallähnliche Eigenschaften. Die maximal mögliche Dotierung ist durch die Löslichkeit der Dotanden bzw. deren Aktivierung gegeben.
4. Sowohl die Elektronen- als auch die Löcherdichte hängen von der Lage des Fermi-niveaus ab. Zur eindeutigen Festlegung der Fermi-energie braucht man eine zusätzliche Beziehung zwischen beiden Trägerdichten - die *Neutralitätsgleichung*.

# Die intrinsische Dichte

Bisher wurden die Besetzungen des Leitungs- und des Valenzbandes im Halbleiter isoliert betrachtet.

Im Halbleiter verteilen sich Elektronen und Löcher nicht unabhängig voneinander auf die Bänder, sondern nach Massgabe eines dynamischen Reaktionsgleichgewichtes.

**Das Massenwirkungsgesetz des Halbleiters:** Im allgemeinen Fall bekommen wir für das Produkt der Dichten

$$n_0 p_0 = N_C N_V F_{1/2}(y_n) F_{1/2}(y_p)$$

In nichtentarteten Halbleitern wird daraus das sogenannte *Massenwirkungsgesetz*

$$n_0 p_0 = N_C N_V \exp\left(-\frac{E_C - E_V}{kT}\right) = N_C N_V \exp\left(-\frac{E_g}{kT}\right) = n_i^2$$

Es ist unabhängig von der Fermienergie (also unabhängig von der Dotierung), aber stark abhängig von den Eigenschaften des betreffenden Halbleiters (Bandstruktur-Parameter  $N_C$ ,  $N_V$  und Energielücke  $E_g$ ).

Das Produkt  $n_0 p_0$  ist somit eine Halbleiterkonstante und wird gleich dem Quadrat der *intrinsischen Dichte*  $n_i$  gesetzt.

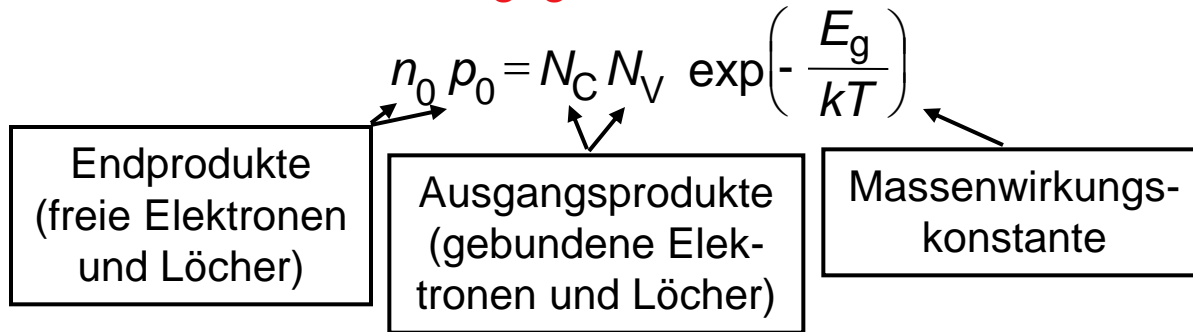


**Eigenhalbleiter:** Hier stimmen die Elektronen- und Löcherdichten überein. Elektron und Loch können immer nur paarweise generiert werden bzw. rekombinieren.

$$n_0 = p_0 = n_i$$

$$n_i = (N_C N_V)^{1/2} \exp\left(-\frac{E_g}{2kT}\right)$$

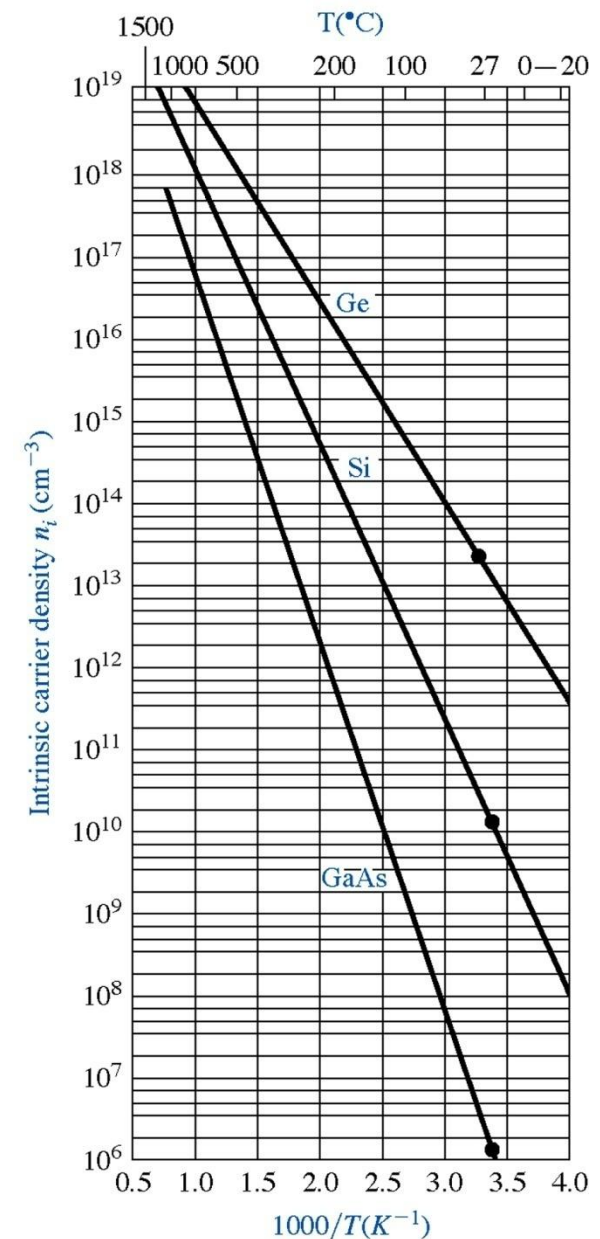
**Warum „Massenwirkungsgesetz“?**



Kann als chemische Reaktion interpretiert werden, die thermisch aktiviert ist und die Schwellenenergie  $E_g$  benötigt.

**Table 4.2** | Commonly accepted values of  $n_i$  at  $T = 300$  K

Silicon	$n_i = 1.0 \times 10^{10} \text{ cm}^{-3}$
Gallium arsenide	$n_i = 1.8 \times 10^6 \text{ cm}^{-3}$
Germanium	$n_i = 2.4 \times 10^{13} \text{ cm}^{-3}$



Intrinsische Dichte als Funktion der Temperatur.

# Die Lage des intrinsischen Fermienergielevels

Im intrinsischen Halbleiter lautet die Bedingung für Ladungsneutralität:  $n_0 = p_0 = n_i$

Setzt man die Gleichgewichtsdichten für den Fall der Nichtentartung ein, folgt

$$N_C \exp\left(-\frac{E_C - E_F}{kT}\right) = N_V \exp\left(-\frac{E_F - E_V}{kT}\right)$$

Auflösung nach der Fermienergie ergibt

$$E_F \equiv E_{Fi} = \frac{E_C + E_V}{2} + \frac{kT}{2} \ln \frac{N_V}{N_C}$$

Setzt man die Ausdrücke für die effektiven Zustandsdichten  $N_C$  und  $N_V$  ein, folgt

$$E_{Fi} = E_{\text{midgap}} + \frac{3kT}{4} \ln \frac{m_p^*}{m_n^*}$$

Die Fermienergie des Eigenhalbleiters  $E_F = E_{Fi}$  (auch Intrinsic-Niveau genannt) liegt bei gleichen effektiven Massen der Elektronen und Löcher in der Mitte der verbotenen Zone.

Für Ge und Si ist das etwa erfüllt. Beispielsweise beträgt der letzte Term für Si nur etwa 1meV.

**Die Trägerdichten:** Mit Bezug auf das Intrinsic-Niveau  $E_{Fi}$  ergeben sich folgende nützliche Beziehungen

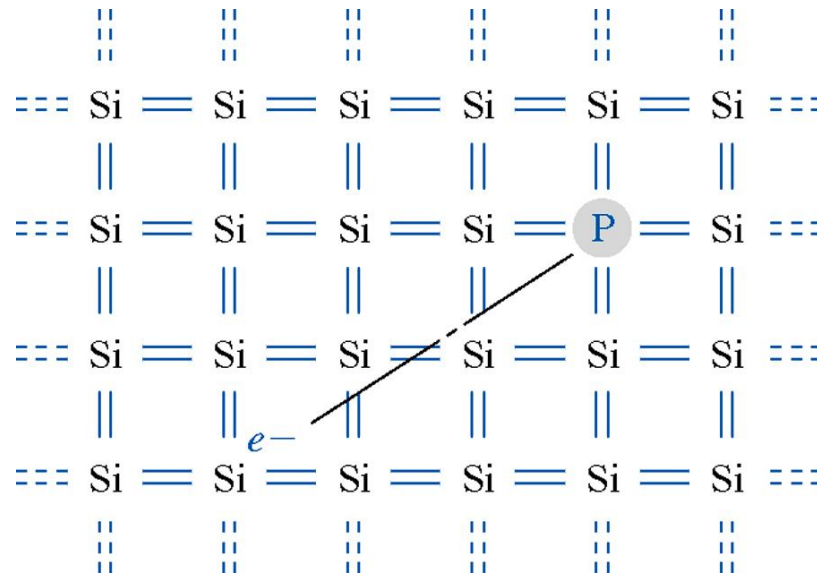
$$\begin{aligned}n_i &= N_C \exp\left(-\frac{E_C - E_{Fi}}{kT}\right) \\ &= N_V \exp\left(-\frac{E_{Fi} - E_V}{kT}\right)\end{aligned}$$

mit denen die Trägerdichten  $n, p$  eines allgemeinen Halbleiters unter Umgehung der effektiven Zustandsdichten geschrieben werden können:

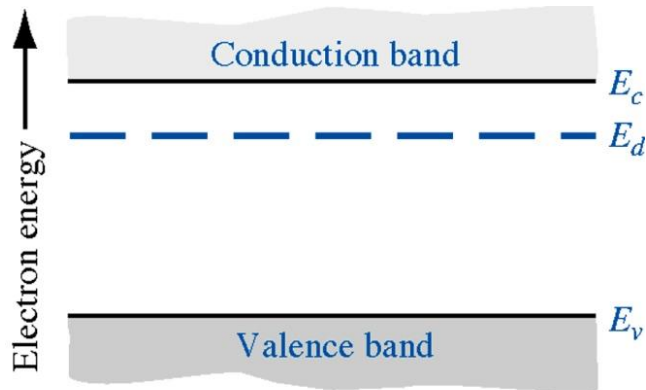
$$\begin{aligned}n_0 &= n_i \exp\left(\frac{E_F - E_{Fi}}{kT}\right) = N_C \exp\left(-\frac{E_C - E_F}{kT}\right) \\ p_0 &= n_i \exp\left(\frac{E_{Fi} - E_F}{kT}\right) = N_V \exp\left(-\frac{E_F - E_V}{kT}\right)\end{aligned}$$

# 4.2 Dotieratome und ihre Energieniveaus

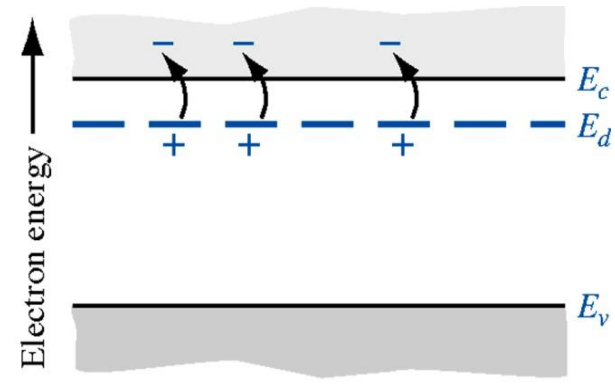
## Donatoren



Phosphor-Dotieratom auf Si-Gitterplatz.

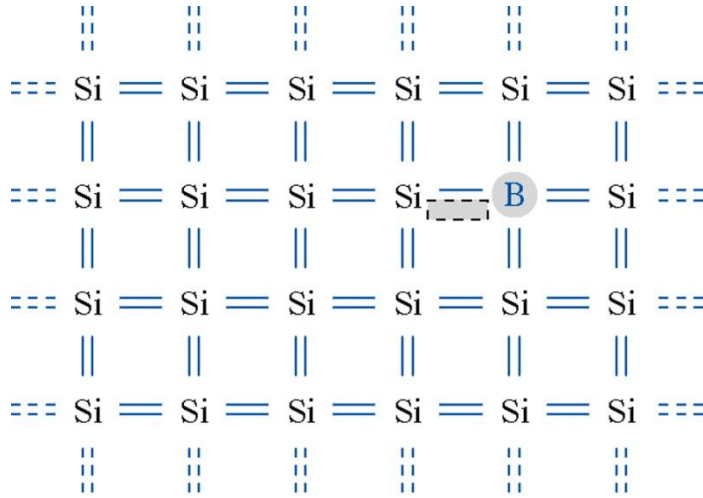


Donator-Energieniveaus im Bandkantendiagramm.

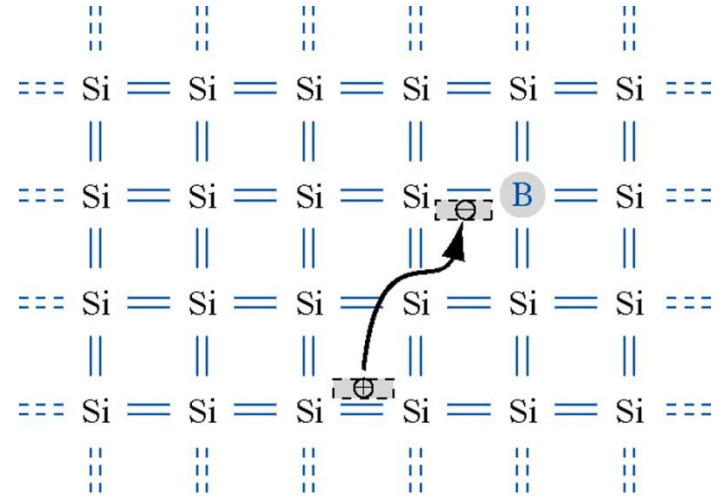


Ionisation der Donatoren.

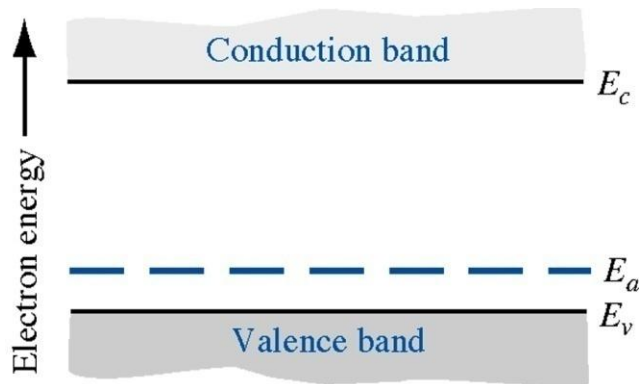
# Akzeptoren



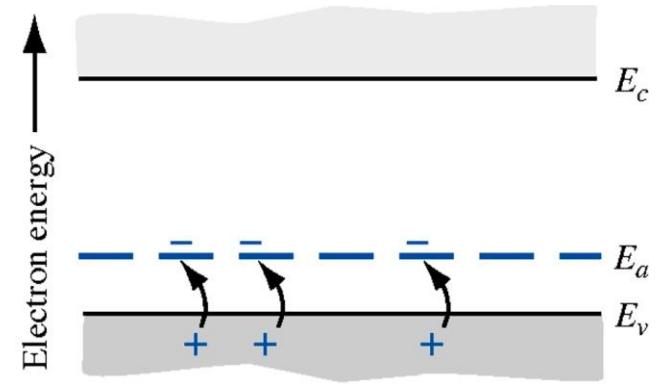
Bor-Dotieratom auf Si-Gitterplatz.



Die Ionisation des Boratoms resultiert in einem freien Loch.



Akzeptor-Energieniveaus im Bandkantendiagramm.



Ionisation der Akzeptoren.

# Aktivierungsenergie. Wasserstoffmodell der Störstelle. Flache und tiefe Störstellen

Im isolierten Atom sind die Hüllenelektronen stark gebunden. Nach dem Bohrschen Wasserstoffmodell beträgt der  $n$ -te Bahnradius

$$r_n = \frac{\varepsilon_0 h^2}{\pi m_0 q^2} n^2 = r_0 n^2$$

und die Energie

$$E_n = -\frac{1}{2} \frac{m_0 q^4}{(4\pi\varepsilon_0 \hbar n)^2}$$

Die Ionisierungsenergie ( $E_1$ ) beträgt 13.6 eV (Rydberg-Energie).

Im Kristallgitter des Halbleiters kreist das Hüllenelektron der Störstelle nicht im Vakuum, sondern in einem durch das Gitter und seine Elektronen polarisierten Medium.

Damit wird die Coulombkraft auf das Elektron um den Faktor  $\epsilon_r$  geschwächt, und die Elektronenmasse muss durch eine effektive Masse  $m^*$  ersetzt werden. Für den Bohr-Radius ergibt sich

$$r_n = \epsilon_r \frac{m_0}{m^*} r_0 n^2$$

und für die Bindungsenergie eines Donator-Elektrons

$$E_D = \frac{1}{\epsilon_r^2} \frac{m_n^*}{m_0} E_n$$

Der Radius wird um den Faktor  $\epsilon_r m_0 / m^*$  grösser, die Aktivierungsenergie um den Faktor  $\epsilon_r^2 m_0 / m^*$  kleiner.

Für Si ( $\epsilon_r = 11.7$ ,  $m^*/m_0 \sim 0.26$ ) wird der Bahnradius um etwa 45x grösser ( $\approx 24 \text{ \AA}$ ), und die Aktivierungsenergie um einen Faktor 526 kleiner ( $\approx 26 \text{ meV}$ ). Allgemein:

Donatoren: Ge:  $E_d \sim 0.006 \text{ eV}$    Si:  $E_d \sim 0.04 \text{ eV}$    GaAs:  $E_d \sim 0.007 \text{ eV}$

Akzeptoren: Ge:  $E_a \sim 0.006 \text{ eV}$    Si:  $E_a \sim 0.05 \text{ eV}$    GaAs:  $E_a \sim 0.050 \text{ eV}$

Die beträchtliche Zunahme des Bahnradius rechtfertigt die Verwendung makroskopischer Dielektrizitätskonstanten. Das locker gebundene Elektron hält vom Atomrumpf einen so grossen Abstand, dass seine Bewegung annähernd als die eines Elektrons um einen wasserstoffähnlichen Rumpf angesehen werden kann.

Als flach bezeichnet man eine Störstelle, wenn ihr Energieniveau nahe d.h.  $\approx (2 - 3) \cdot kT$  an der Leitungs- oder Valenzbandkante liegt. Von tiefen Störstellen spricht man dagegen, wenn ihre energetische Lage weiter von einer der beiden Bandkanten entfernt ist.

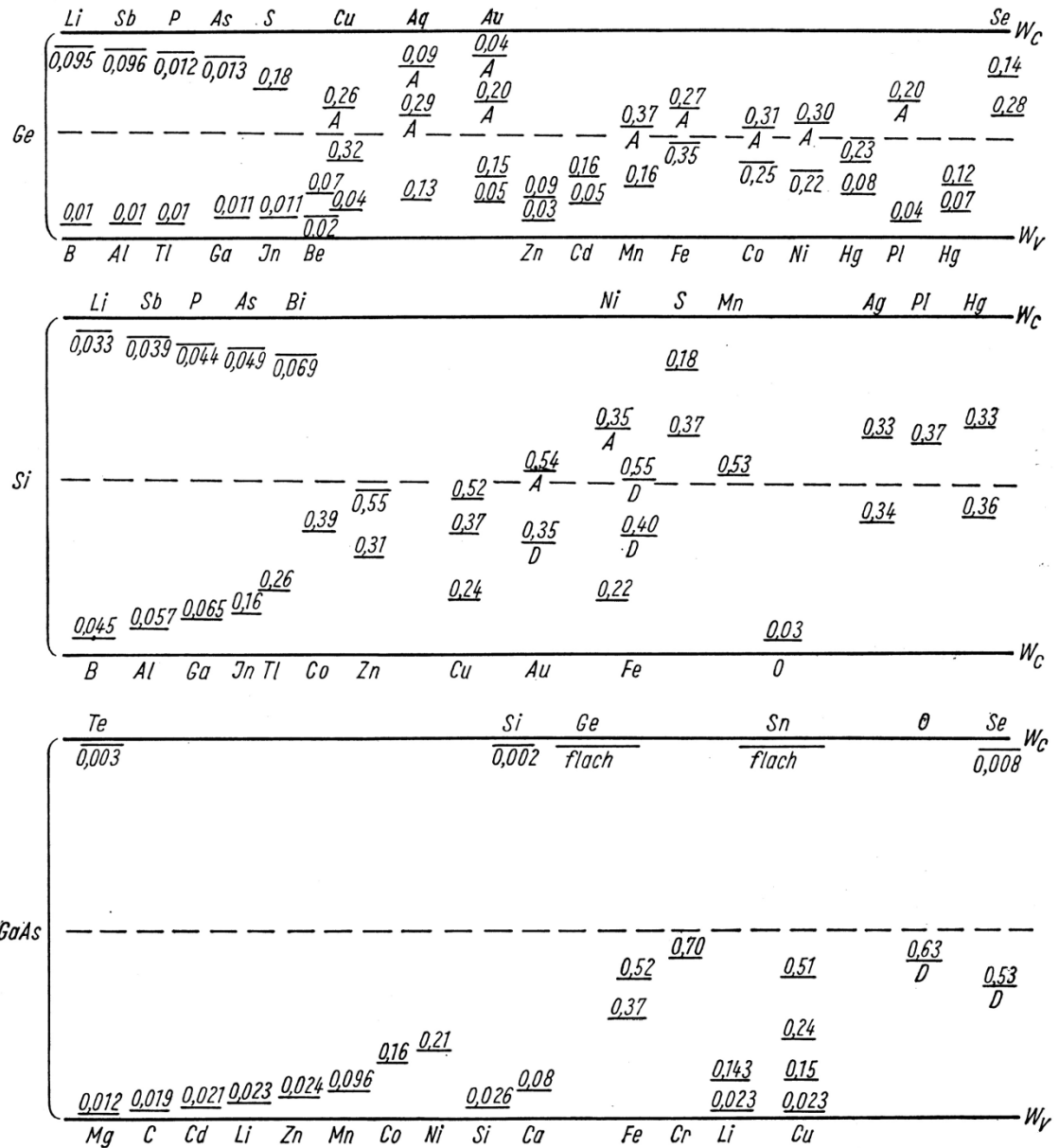
Zu den flachen Störstellen zählen normalerweise Elemente aus der 3. und 5. Hauptgruppe. Schwermetalle (Cd, Mn, Ni, Fe, Cu) wirken gewöhnlich als tiefe Störstellen.



# Wirkung von Störstellen

1. Elektrisch aktivierte Störstellen beeinflussen die Leitfähigkeit des Kristalls durch Abgabe von Elektronen oder Löchern.
2. Bei sehr hohen Störstellenkonzentrationen ( $> 5 \cdot 10^{19} \text{ cm}^{-3}$  in Si) bekommt der Halbleiter metallische Eigenschaften. Man spricht dann von Entartung.
3. Störstellen liefern einen Beitrag zur Streuung. Man unterscheidet Streuung an geladenen, ungeladenen und gruppierten Störstellen.
4. Störstellen beeinflussen nicht nur die elektrischen, sondern auch andere (chemische, optische, magnetische) Halbleitereigenschaften.
5. Liegen örtlich verschiedene Störstellendichten vor, so können Störstellen bei sehr hohen Temperaturen ( $> 800^\circ\text{C}$ ) durch Diffusion in Richtung des negativen Dichtegradienten wandern. Auf dieser Störstellendiffusion beruhen die wichtigsten Herstellungsmethoden für Festkörperbauelemente.
6. Tiefe Störstellen fungieren als sogenannte Rekombinationszentren und erhöhen die Rekombinationsrate.

# Energetische Lage von Störstellen in Si, Ge, und GaAs



# Ionisationsenergien von Dotanden in Si, Ge und GaAs

**Table 4.3** | Impurity ionization energies in silicon and germanium

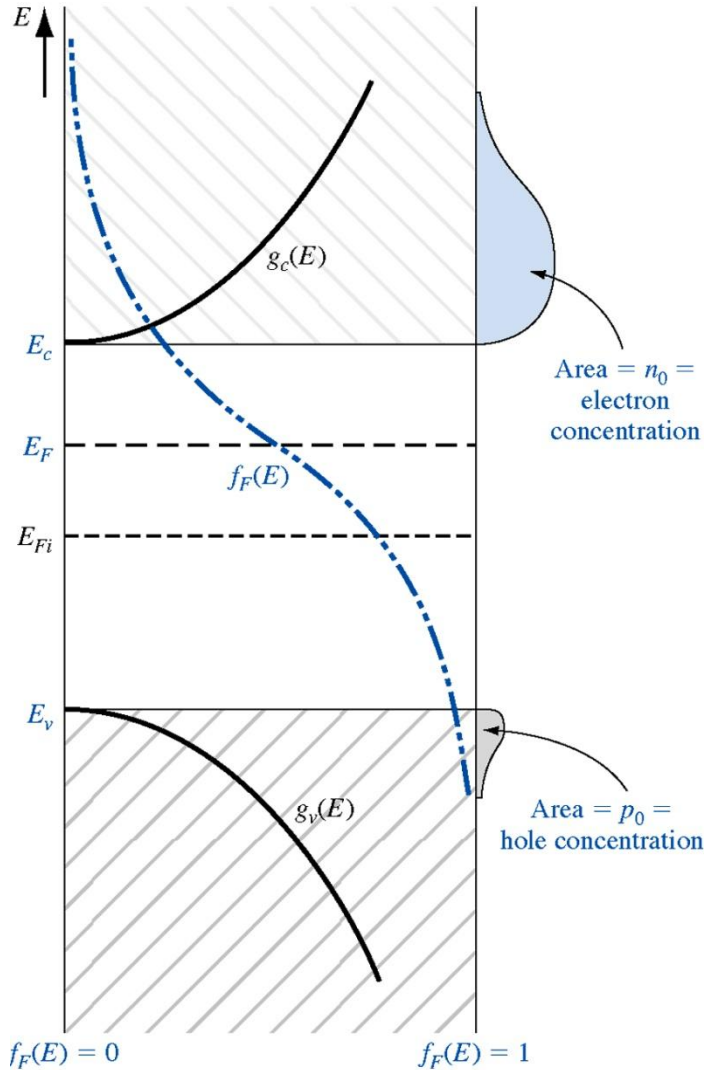
Impurity	Ionization energy (eV)	
	Si	Ge
<i>Donors</i>		
Phosphorus	0.045	0.012
Arsenic	0.05	0.0127
<i>Acceptors</i>		
Boron	0.045	0.0104
Aluminum	0.06	0.0102

**Table 4.4** | Impurity ionization energies in gallium arsenide

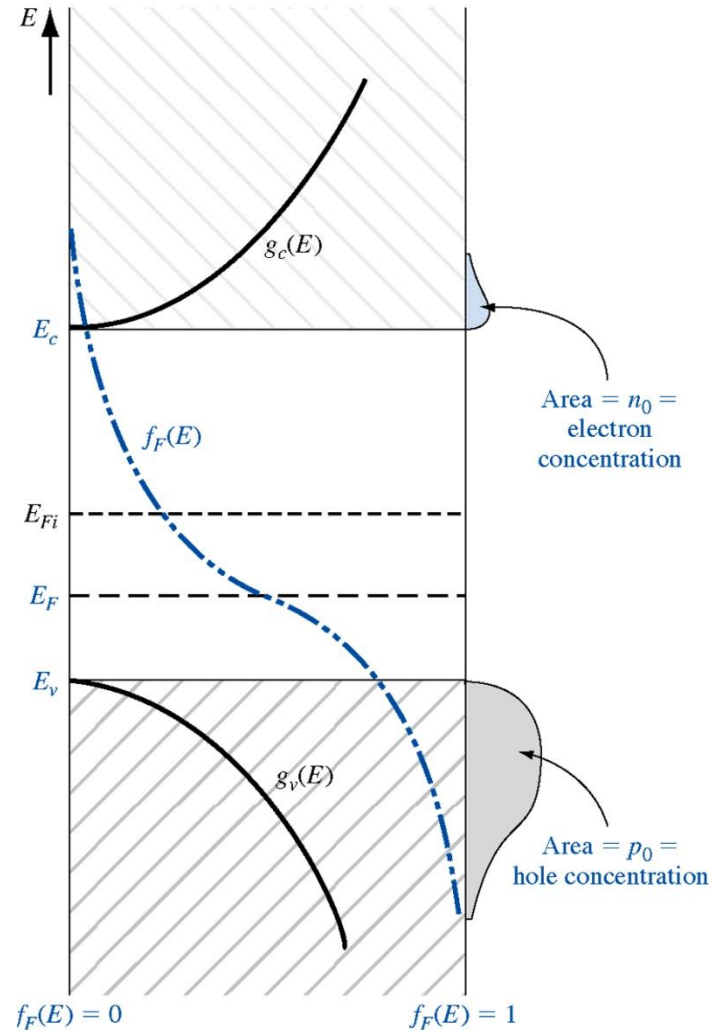
Impurity	Ionization energy (eV)
<i>Donors</i>	
Selenium	0.0059
Tellurium	0.0058
Silicon	0.0058
Germanium	0.0061
<i>Acceptors</i>	
Beryllium	0.028
Zinc	0.0307
Cadmium	0.0347
Silicon	0.0345
Germanium	0.0404

# 4.3 Der extrinsische Halbleiter

## n-Halbleiter



## p-Halbleiter



## Die Konstanz der Produktes

$$n_0 p_0 = N_C N_V \exp\left(-\frac{E_C - E_V}{kT}\right) = N_C N_V \exp\left(-\frac{E_g}{kT}\right) = n_i^2$$

schliesst nicht aus, dass sich  $n_0$  und  $p_0$  im allgemeinen Fall unterscheiden. Im Gegenteil: gerade durch (stark) unterschiedliche Dichten  $n_0$  und  $p_0$  kann die Leitfähigkeit einseitig sehr stark erhöht werden. So stellen den Hauptbeitrag entweder Löcher ( $p_0 \gg n_0$ , man spricht von  $p$ -Leitung), oder Elektronen ( $n_0 \gg p_0$ , man spricht von  $n$ -Leitung).

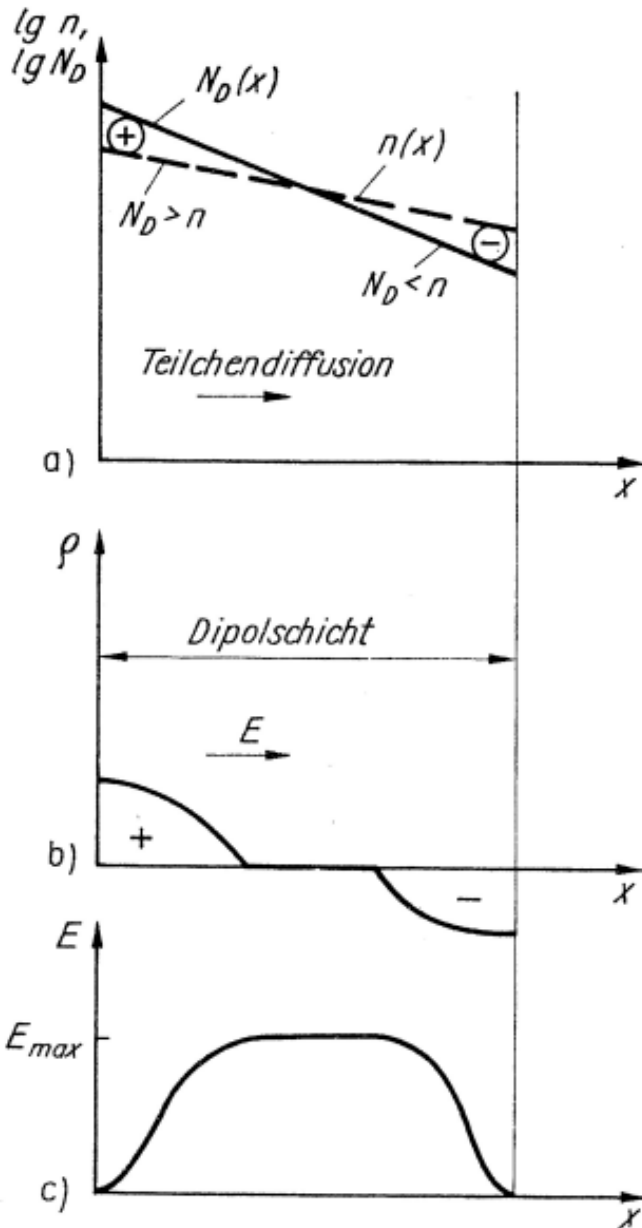
**Räumliche Verteilung.** Die Wirkung der Störstellen auf die elektronischen Eigenschaften wird nicht nur durch ihre Anzahl bestimmt, sondern in starkem Masse auch durch ihre Verteilung. Man unterscheidet:

1. *Homogene Verteilung* (homogene Dotierung).

Die Störstellen sind im Halbleiter gleichmässig verteilt. Dann herrscht im thermodynamischen Gleichgewicht stets örtliche Neutralität.

2. *Inhomogene Verteilung* (inhomogene Dotierung). Die Störstellen sind im Halbleiter inhomogen verteilt.

Im allgemeinen stellen sich bereits im thermodynamischen Gleichgewicht durch Wechselspiel mit beweglichen Trägern (Feld- und Diffusionsbewegung) innere Raumladungsfelder ein. Ein Mass der Inhomogenität ist der Störstellengradient, d.h. die räumliche Änderung der Störstellenverteilung.



Der Halbleiter ist rechts schwächer dotiert als links. Die Elektronenkonzentration links ist bestrebt, sich durch Diffusion von Elektronen in Richtung des Dichtegefälles abzubauen. Es kommt schliesslich zu einer Zunahme der Elektronendichte rechts. Da nun Elektronen- und Störstellendichte nicht mehr übereinstimmen, ist die Neutralität gebietsweise verletzt; es entsteht eine innere Raumladung und, mit ihr verknüpft, ein inneres Feld. Dieses Feld bremst die Elektronendiffusion, im Gleichgewicht kompensieren sich Diffusions- und Feldstrom gegenseitig.

Eine besonders starke Inhomogenität stellt der (abrupte) Wechsel der Störstellenkonzentration oder sogar des Typs (Donator, Akzeptor) dar, wie er etwa an Grenzflächen oder bei  $pn$ -Übergängen auftritt.

- a) Konzentrationsverlauf der ortsfesten Störstellen  $N_D(x)$  und Elektronen im Gleichgewicht zwischen Trägerdichtenausgleich durch Diffusion und entgegengerichteter Feldkraft
- b) stationär entstandene Raumladungsdichte als Ursache der Feldstärke
- c) Verlauf der Feldstärke

# 4.4 Statistik von Donatoren und Akzeptoren

**Störstellenbesetzung:** Einfach besetzbare Störstellen haben nur ein diskretes Energieniveau ( $E_d$ ,  $E_a$ , allgemein  $E_t$ ). Die Formel für die Zustandsdichte lautet:  $g_t(E) = N_t \delta(E - E_t)$ , wobei  $N_t$  die Dichte ( $\text{cm}^{-3}$ ) der jeweiligen Störstellensorte ist ( $t = d, a$ ).

Die Wahrscheinlichkeit für die Besetzung eines solchen Niveaus unterscheidet sich etwas von der Fermi-Dirac-Verteilung. Ein Donatorniveau könnte gemäss Pauli-Prinzip von einem Elektron „spin-down“ *und* einem Elektron „spin-up“ besetzt werden (verschiedene Quanten-Zustände!). Nachdem das Niveau von einem Elektron besetzt wurde, kann jedoch das (andere) zweite Elektron nicht mehr gebunden werden, weil die starke Coulomb-Abstossung dies verhindert. Diese starke Wechselwirkung ist eine Folge der starken Lokalisation der Wellenfunktion (Erinnerung: der Bohrsche Radius beträgt nur etwa 4 Gitterkonstanten). Als Folge tritt ein zusätzlicher Faktor  $\frac{1}{2}$  vor der e-Funktion im Nenner auf, der von den beiden möglichen Spineinstellungen des Donator-Elektrons herrührt

$$f_t(E_d) = \frac{1}{1 + \frac{1}{2} \exp \frac{E_d - E_F}{kT}} \quad n_d = \frac{N_d}{1 + \frac{1}{2} \exp \frac{E_d - E_F}{kT}} = N_d - N_d^+$$

Beim Akzeptor gibt es wegen der komplexen Valenzbandstruktur des Si sogar 4 mögliche Spinzustände. Allgemein führt man einen Entartungsfaktor  $g$  ein. Im Falle des einfachen Akzeptors in Silizium wird also  $g = 4$ .

$$f_t(E_a) = \frac{1}{1 + \frac{1}{g} \exp \frac{E_F - E_a}{kT}} \quad p_a = \frac{N_a}{1 + \frac{1}{g} \exp \frac{E_F - E_a}{kT}} = N_a - N_a^-$$



Unter der Annahme tiefer Temperaturen bzw.  $(E_d - E_F) \gg kT$ , ist die Dichte der Elektronen auf Donatorplätzen

$$n_d \approx \frac{N_d}{\frac{1}{g_d} \exp\left(\frac{E_d - E_F}{kT}\right)} = 2N_d \exp\left(-\frac{E_d - E_F}{kT}\right)$$

Für  $(E_d - E_F) \gg kT$  gilt die Boltzmann-Näherung auch für die Elektronen im Leitungsband,

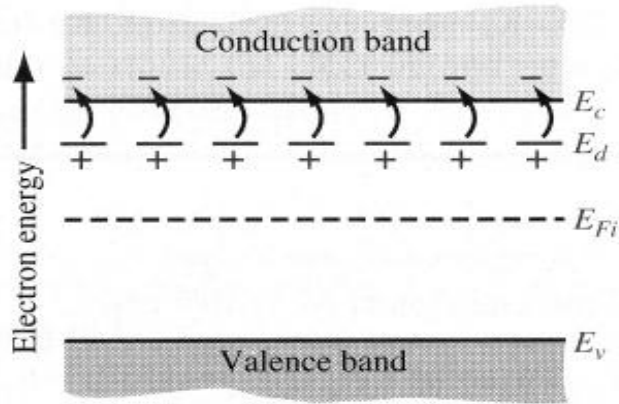
$$n_0 = N_C \exp\left(-\frac{E_C - E_F}{kT}\right)$$

Und man erhält für die relative Zahl der Elektronen auf Donatorplätzen

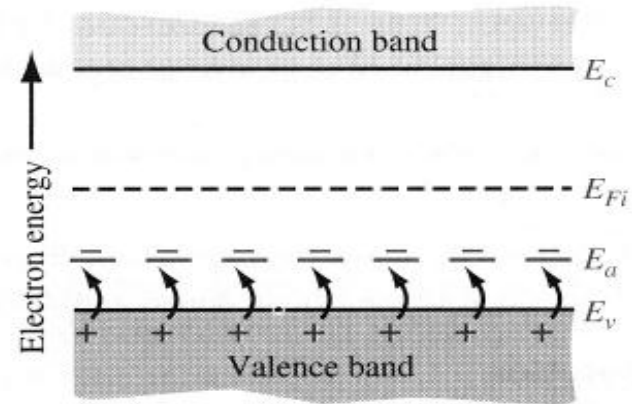
$$\frac{n_d}{n_d + n_0} \approx \frac{1}{1 + \frac{N_C}{2N_d} \exp\left(-\frac{E_C - E_d}{kT}\right)}$$

$(E_C - E_d)$  ist die Ionisierungsenergie des Donatoratoms.

# Ionisierung und Ausfrieren

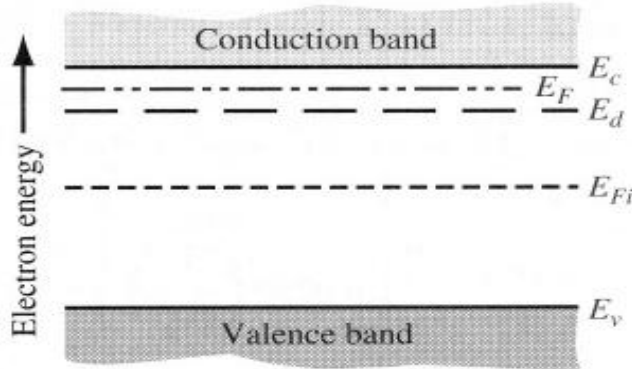


(a)

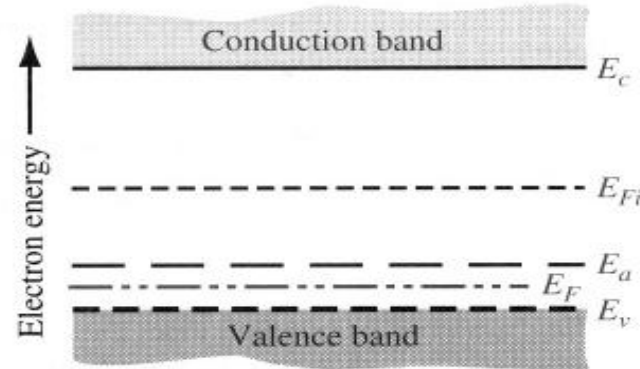


(b)

Bandkantendiagramm mit Veranschaulichung der vollständigen Ionisation:  
 (a) Donator-Zustände, (b) Akzeptor-Zustände.



(a)



(b)

Bandkantendiagramm für  $T = 0$  K: (a) n-Typ-Halbleiter, (b) p-Typ-Halbleiter.

# 4.5 Ladungsneutralität

**Allgemeine Bestimmung der Fermienergie.** Ein homogenes Halbleitergebiet ist elektrisch stets neutral. Es muss also die Neutralitätsbedingung gelten:

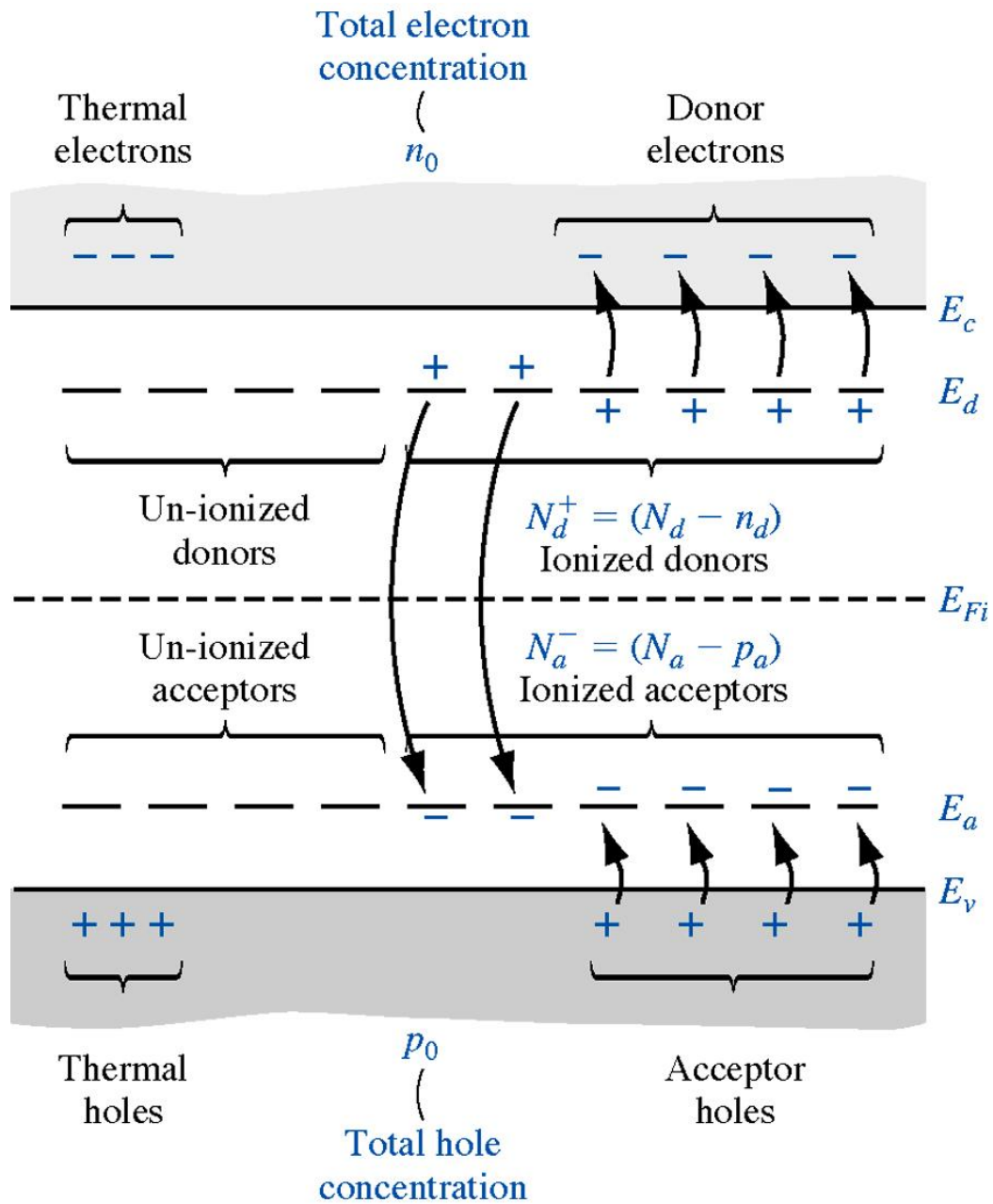
$$\sum (Q_{\text{beweglich}} + Q_{\text{fest}}) = 0$$

Sie umfasst alle im Halbleiter vorhandenen beweglichen und ortsfesten (ionisierten) Ladungen  $Q$ . Bei einfach ionisierten Störstellen (Donatoren  $N_d^+$ , Akzeptoren  $N_a^-$ ) folgt daraus

$$n_0 + N_a^- = p_0 + N_d^+$$

als Bestimmungsgleichung der Fermienergie.

Zur Bestimmung der Fermienergie müssen die Störstellenkonzentrationen, ihre Aktivierungsenergien, die effektiven Zustandsdichten, Temperatur und Bandabstand gegeben sein.



# Majoritäts- und Minoritätsträger

Bei neutralen, im thermodynamischen Gleichgewicht befindlichen Halbleitern bestimmen Neutralitätsbedingung und Massenwirkungsgesetz bei gegebenen Störstellendichten die Konzentrationen der Elektronen und Löcher eindeutig:

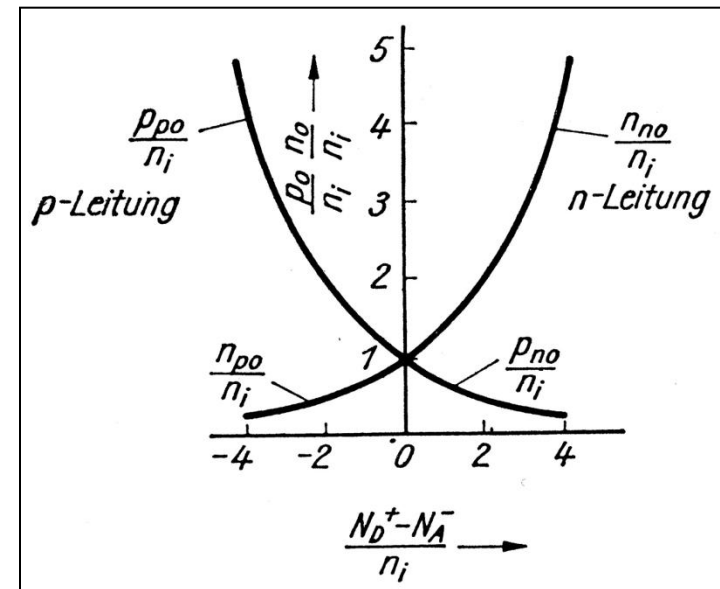
$$\left. \begin{matrix} p_0 \\ n_0 \end{matrix} \right\} = \mp \frac{1}{2} (N_d^+ - N_a^-) \pm \sqrt{\frac{1}{4} (N_d^+ - N_a^-)^2 + n_i^2}$$

Damit dominieren für  $(N_d - N_a) > 0$  stets Elektronen gegenüber den Löchern, für den Fall  $(N_d - N_a) < 0$  stets Löcher gegenüber den Elektronen.

Man nennt daher die in der Mehrzahl vorhandenen Träger **Majoritätsträger**, die in der Minderzahl vorhandenen **Minoritätsträger**. Ausserdem gilt:

$$p\text{-Leitung: } p_{p0} - n_{p0} = N_a^- - N_d^+$$

$$n\text{-Leitung: } n_{n0} - p_{n0} = N_d^+ - N_a^-$$



# Majoritäts- und Minoritätsträger: Folgerungen

1. Der Leitungstyp wird durch die im Überschuss vorhandenen aktivierten Störstellen bestimmt.
2. Falls ausgeprägte Störleitung vorliegt, d.h.  $|N_a - N_d^+| \gg n_i$  gilt, stimmt die Majoritätsdichte im thermodynamischen Gleichgewicht mit der ionisierten Nettostörstellendichte überein.  $p$ -Leitung:  $p_p \approx N_a$ ,  $n$ -Leitung:  $n_n \approx N_d^+$
3. Die Minoritätsträgerdichte stellt sich im thermodynamischen Gleichgewicht stets nach dem Massenwirkungsgesetz ein:

$$n_{n0} p_{p0} = n_i^2 = n_{p0} p_{n0}$$

Im Fall ausgeprägter Störleitung gilt somit für  $n$ - und  $p$ -Leitung

$$n_{p0} = \frac{n_i^2}{p_{p0}} \quad p_{n0} = \frac{n_i^2}{n_{n0}}$$

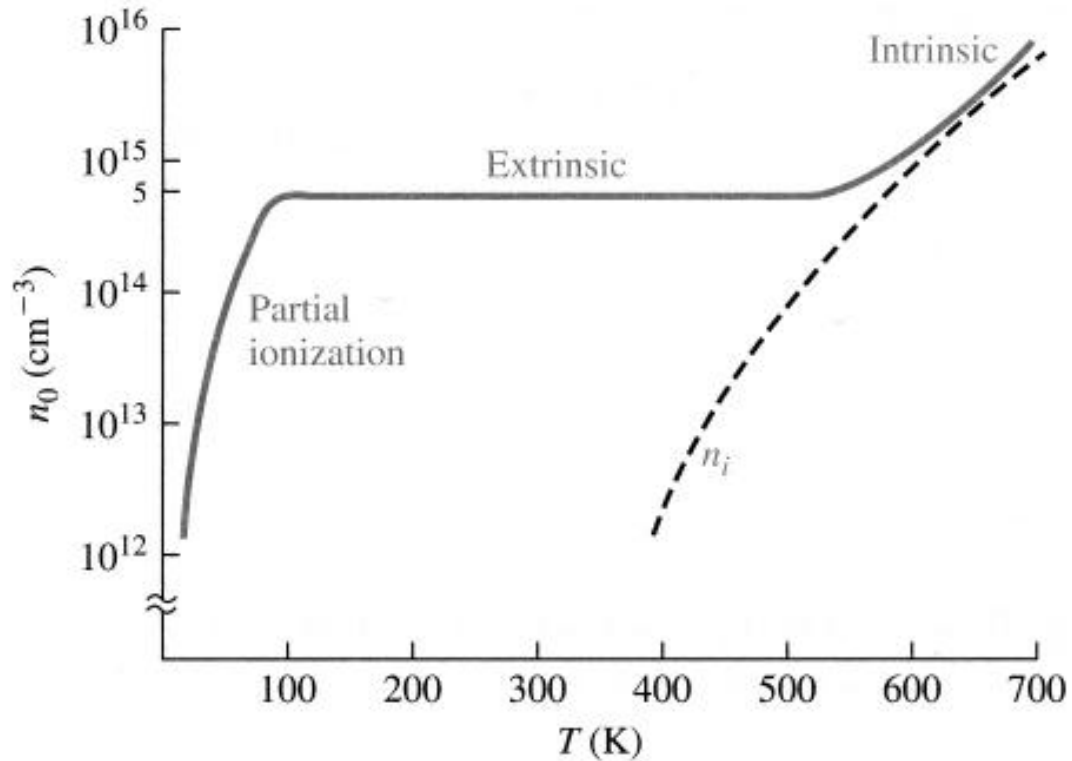
4. Eine Anhebung der Majoritätsträgerkonzentration bedingt eine Senkung der Minoritätsträgerkonzentration.
5. Da die intrinsische Dichte stark von der Temperatur abhängt, überträgt sich dies voll auf die Minoritäten.

6. Minoritätsträger bestimmen zwar nicht den Leitungstyp, aber in entscheidendem Masse die elektronischen Eigenschaften einer Reihe wichtiger Halbleiterbauelemente (Bipolar-Transistoren, Dioden).

Daher heissen Halbleiterbauelemente, bei denen vorzugsweise die Minoritätsträger das Verhalten bestimmen, *bipolare* Bauelemente. Die gleichzeitige Anwesenheit von Majoritätsträgern ist dabei sekundär.

Bestimmen dagegen nur Majoritätsträger, d.h. eine Trägersorte, das Verhalten des Bauelementes, so spricht man von *unipolaren* Bauelementen. Der wichtigste Vertreter ist der Feldeffekt-Transistor.

# Temperaturabhängigkeit der Ladungsträgerdichte



Elektronenkonzentration als Funktion der Temperatur. Man unterscheidet deutlich 3 Regionen: teilweise Ionisierung, extrinsischer Bereich, intrinsischer Bereich.



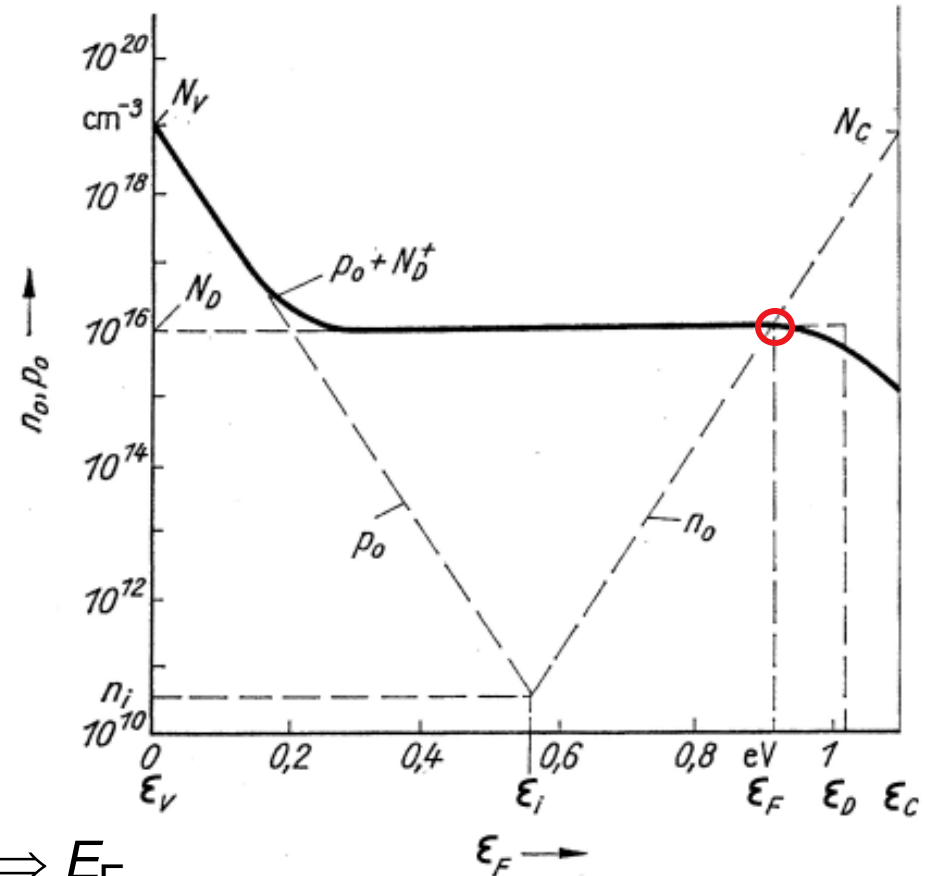
# 4.6 Lage des Ferminiveaus

Man kann die Lage des Ferminiveaus im homogenen Halbleiter durch graphische Lösung der Neutralitätsbedingung erhalten (hier für einen nicht-entarteten n-Halbleiter)

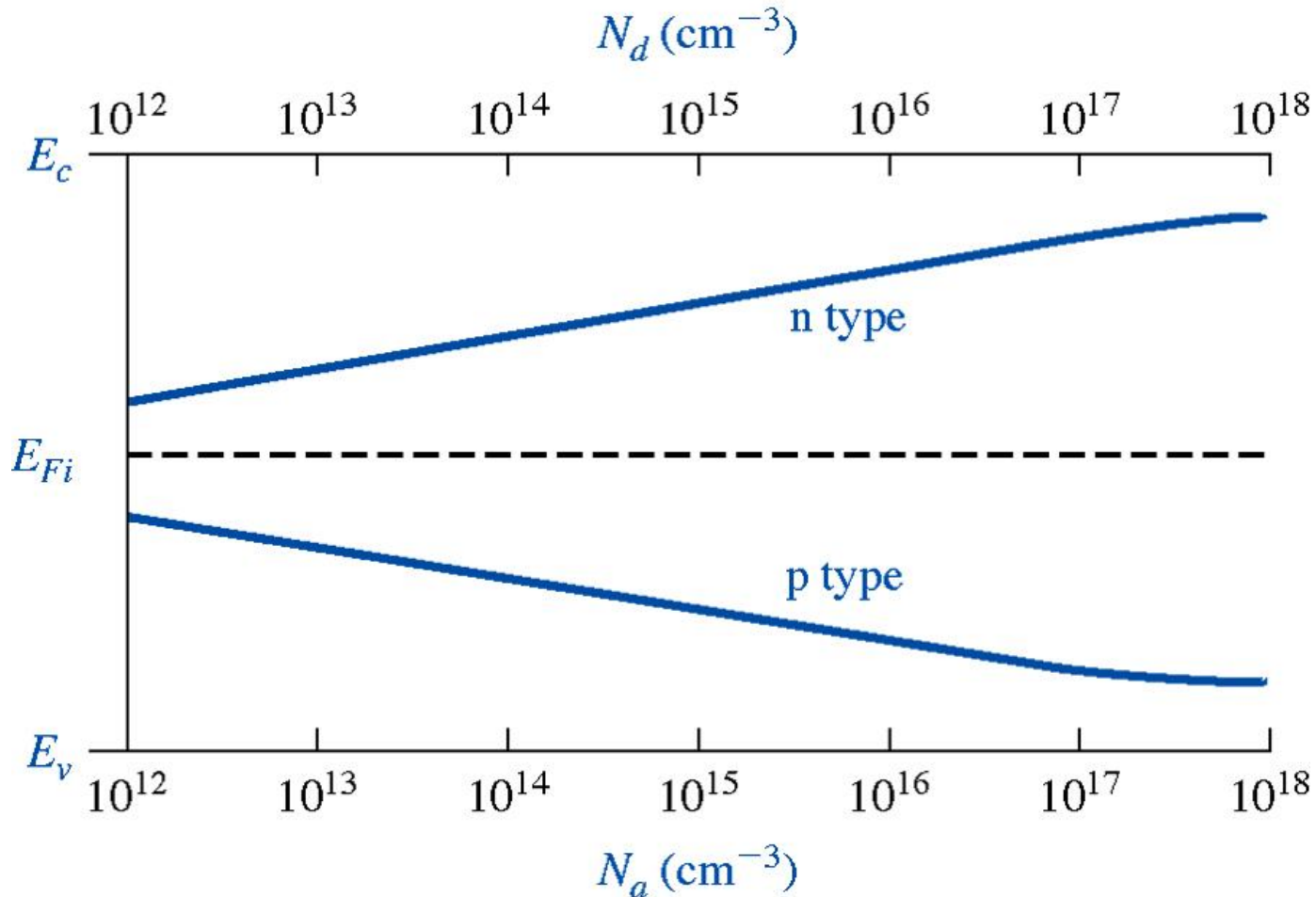
$$N_C \exp\left(-\frac{E_C - E_F}{kT}\right) + \frac{N_a}{1 + g_a \exp\left(\frac{E_A - E_F}{kT}\right)} =$$

$$= N_V \exp\left(-\frac{E_F - E_V}{kT}\right) + \frac{N_d}{1 + g_d \exp\left(\frac{E_F - E_d}{kT}\right)}$$

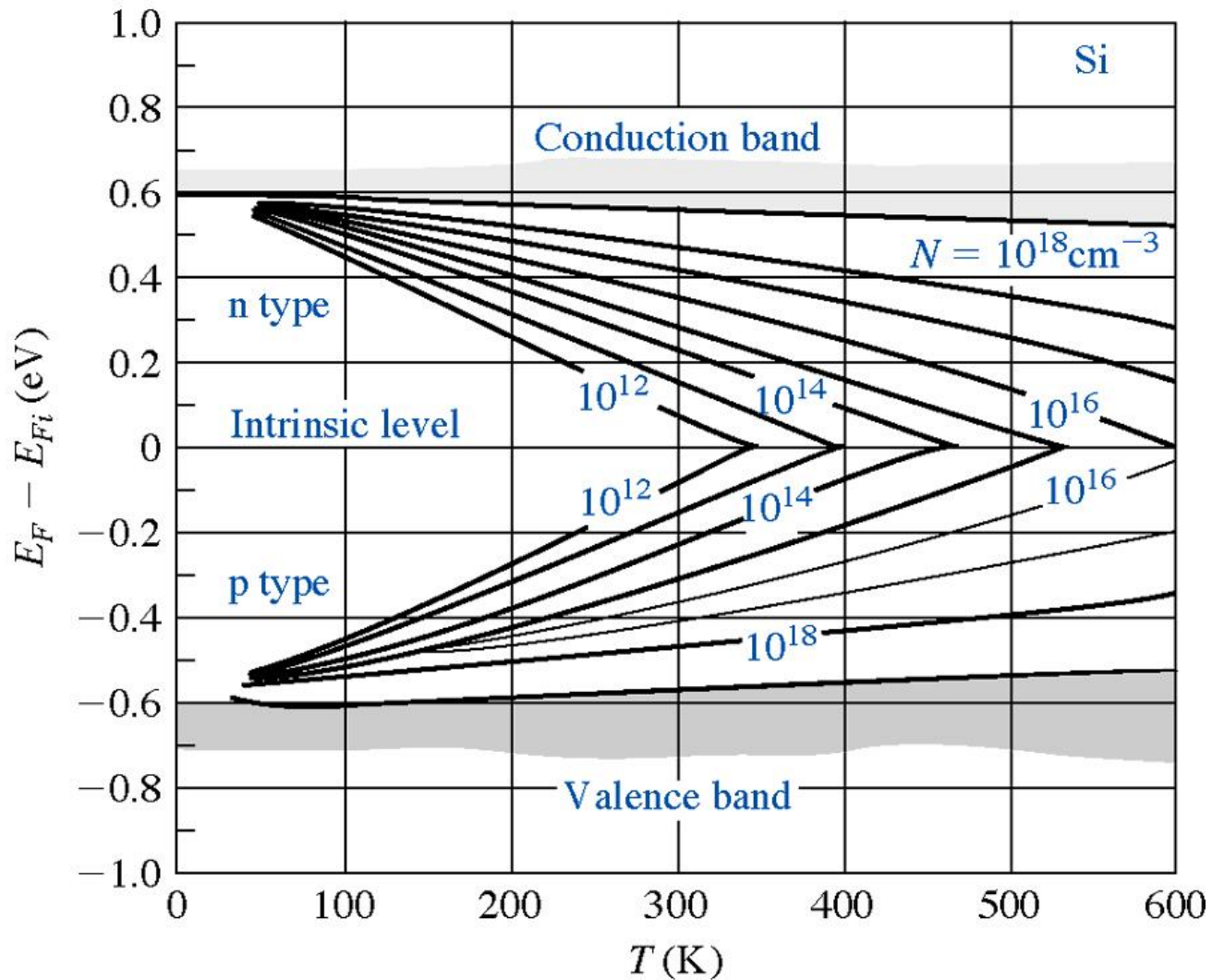
1.  $N_C, N_V, n_i, E_i$  sind bekannt
2. Zeichne  $n_0, p_0$  für  $E_V < E_F < E_C$
3. Zeichne  $p_0 + N_d^+$
4. Schnittpunkt von  $p_0 + N_d^+$  mit  $n_0 \Rightarrow E_F$



# Verlauf von $E_F$ als Funktion der Dotierung



# Verlauf von $E_F$ als Funktion der Temperatur







# 5. Transportphänomene

- 5.1 Einfluss des elektrischen Feldes: Drift (N5.1)
- 5.2 Einfluss von Dichtegradienten: Diffusion (N5.2)
- 5.3 Gradienten der Dotierkonzentration (N5.3)
- 5.4 Generation und Rekombination (N6.1)
- 5.5 Quasi-Ferminiveaus (N6.4)

# 5.1 Drift

Elektrischer Strom wird durch die Bewegung von Ladungen getragen. Haben Löcher mit einer Dichte  $p$  eine mittlere Driftgeschwindigkeit  $v_{dp}$ , so ergibt sich eine Driftstromdichte (siehe Folie 87)

$$J_{dp} = ep v_{dp} = ep \mu_p \mathbf{E}.$$

Masseinheit:

$$\frac{\text{As}}{\text{cm}^3} \frac{\text{cm}}{\text{s}} = \frac{\text{A}}{\text{cm}^2}$$

Analog für Elektronen mit der mittleren Driftgeschwindigkeit  $v_{dn}$  (siehe Folie 83):

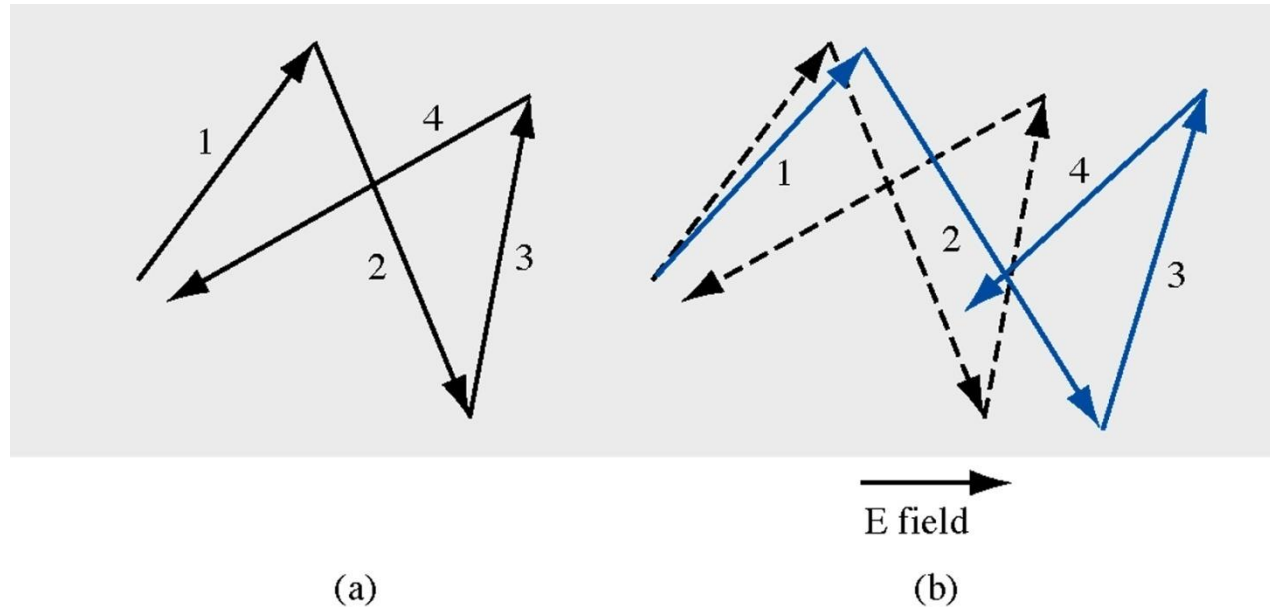
$$J_{dn} = -en v_{dn} = en \mu_n \mathbf{E}.$$

Die Proportionalitäts“konstante“ zwischen Driftgeschwindigkeit und Feldstärke ist die *Beweglichkeit*  $\mu$ .

# Die Beweglichkeit der Ladungsträger im Halbleiter

Elektronen- und Löcherbeweglichkeit gehören zu den wichtigsten Kenngrößen eines Halbleiters.

$$\mu = \frac{e\tau_c}{m^*}$$

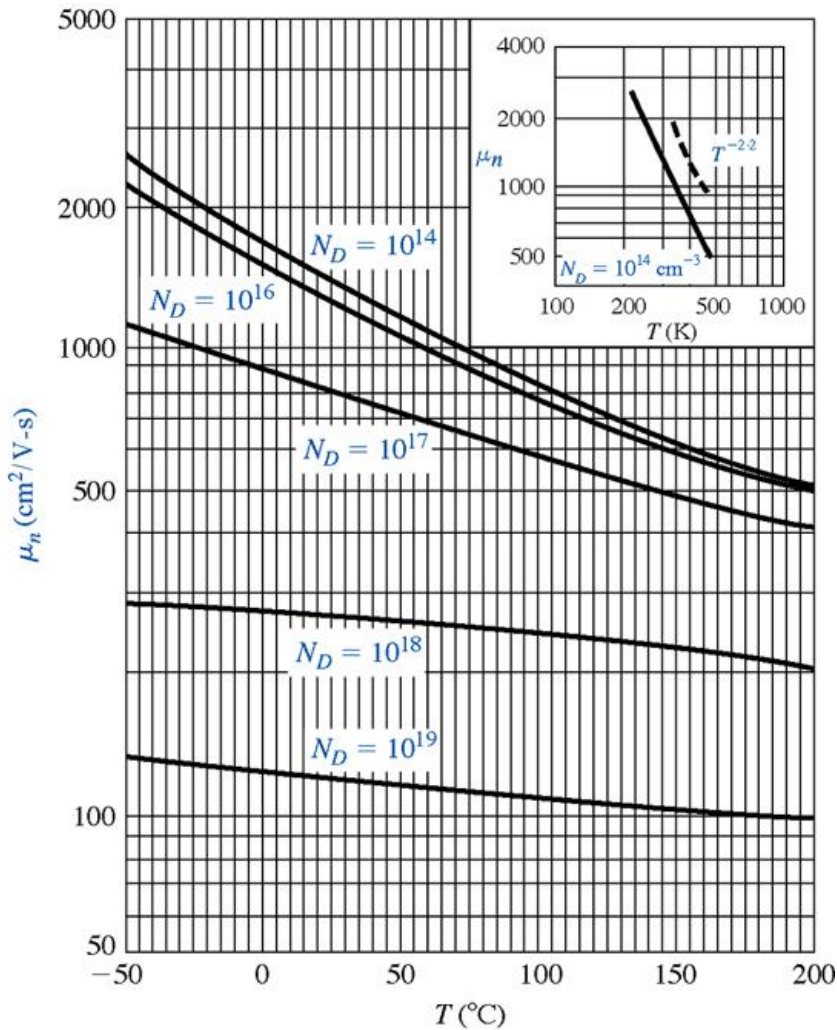


Stochastisches Verhalten eines Loches im Halbleiter a) ohne Feld, b) mit Feld.

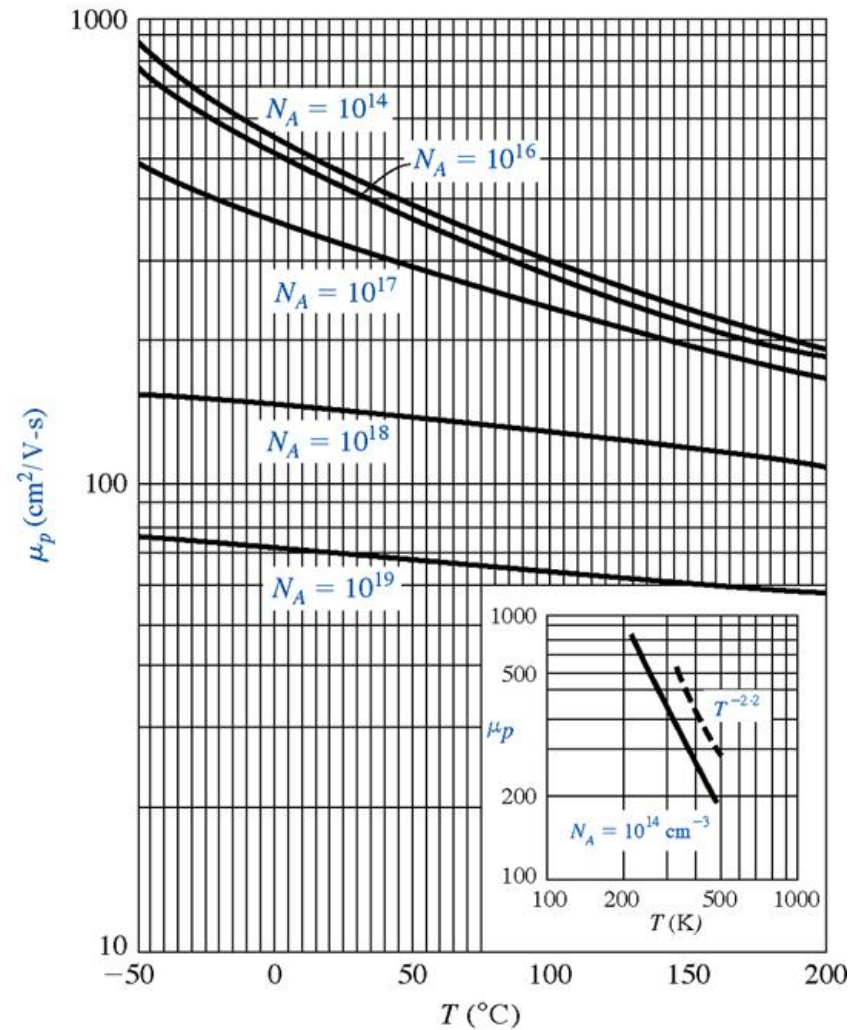
Die Beweglichkeit beschreibt den Effekt aller mikroskopischer Streuprozesse der Elektronen und Löcher im Halbleiter. Deshalb hängt  $\mu$  stark von Temperatur, Dotierung, Defektdichte, etc. ab.



# Beweglichkeit in Silizium



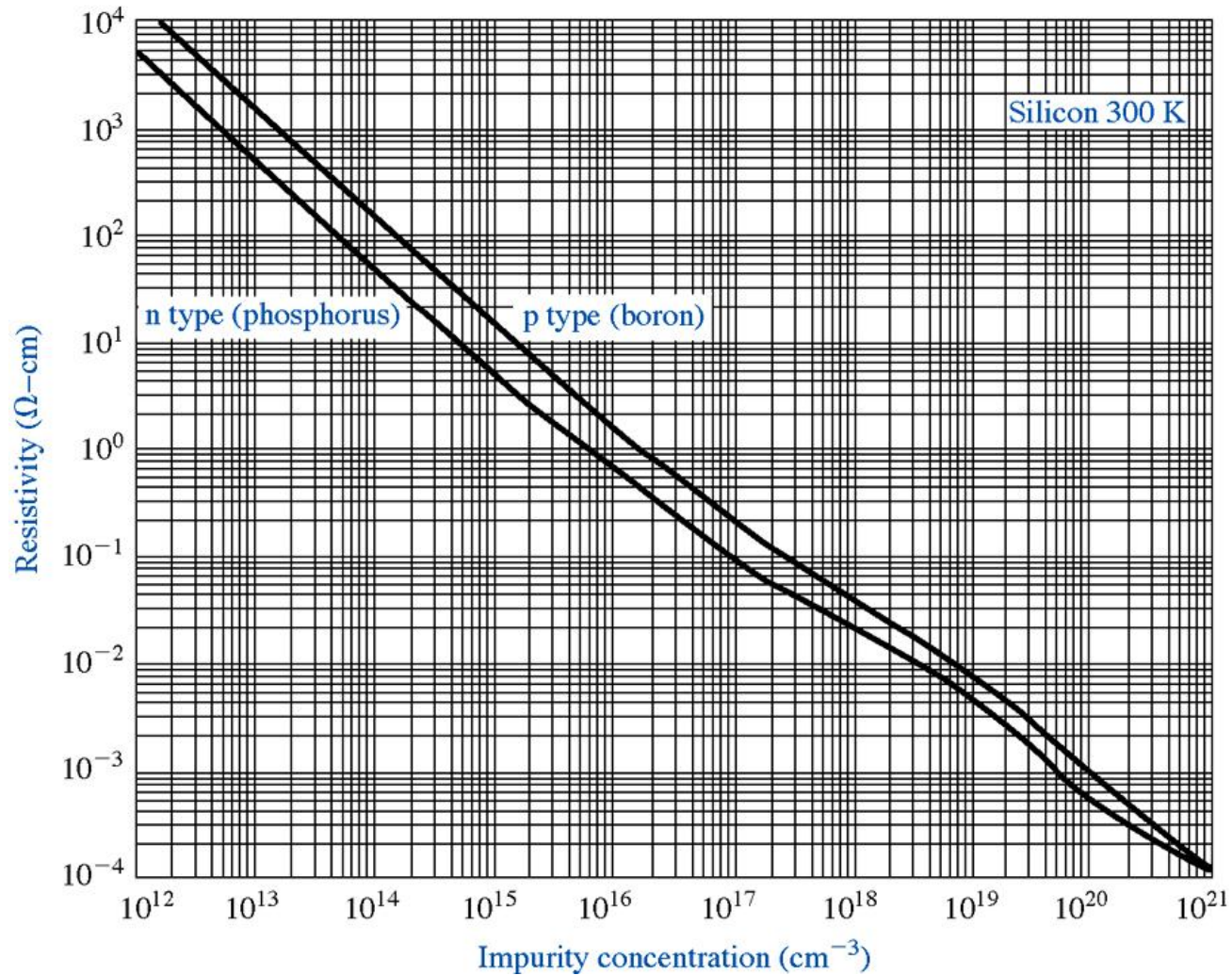
(a)



(b)

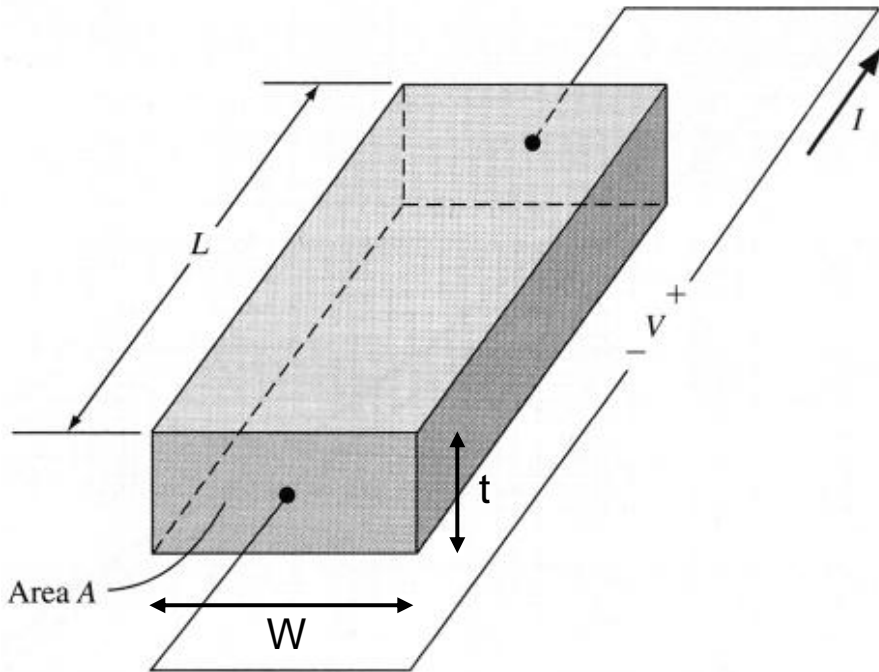
a) Elektronen- und b) Löcherbeweglichkeit in Silizium als Funktion der Temperatur für verschiedene Dotierkonzentrationen.

# Spezifischer Widerstand $\rho=1/\sigma$



Spezifischer Widerstand von Silizium als Funktion der Dotierung.

# Das Ohmsche Gesetz im Halbleiter



Ein Halbleiterblock mit angelegter Spannung leitet einen Strom  $J = I/A$ . Das elektrische Feld ist  $E = V/L$ . Daraus folgt

$$\frac{I}{A} = \sigma \left( \frac{V}{L} \right) = \frac{1}{\rho} \left( \frac{V}{L} \right) \Rightarrow \rho = \frac{E}{J} [\Omega \text{cm}]$$

bzw.

$$V = \left( \frac{L}{\sigma A} \right) I = \left( \frac{\rho L}{A} \right) I = IR \quad \text{“}U=RI\text{”}$$

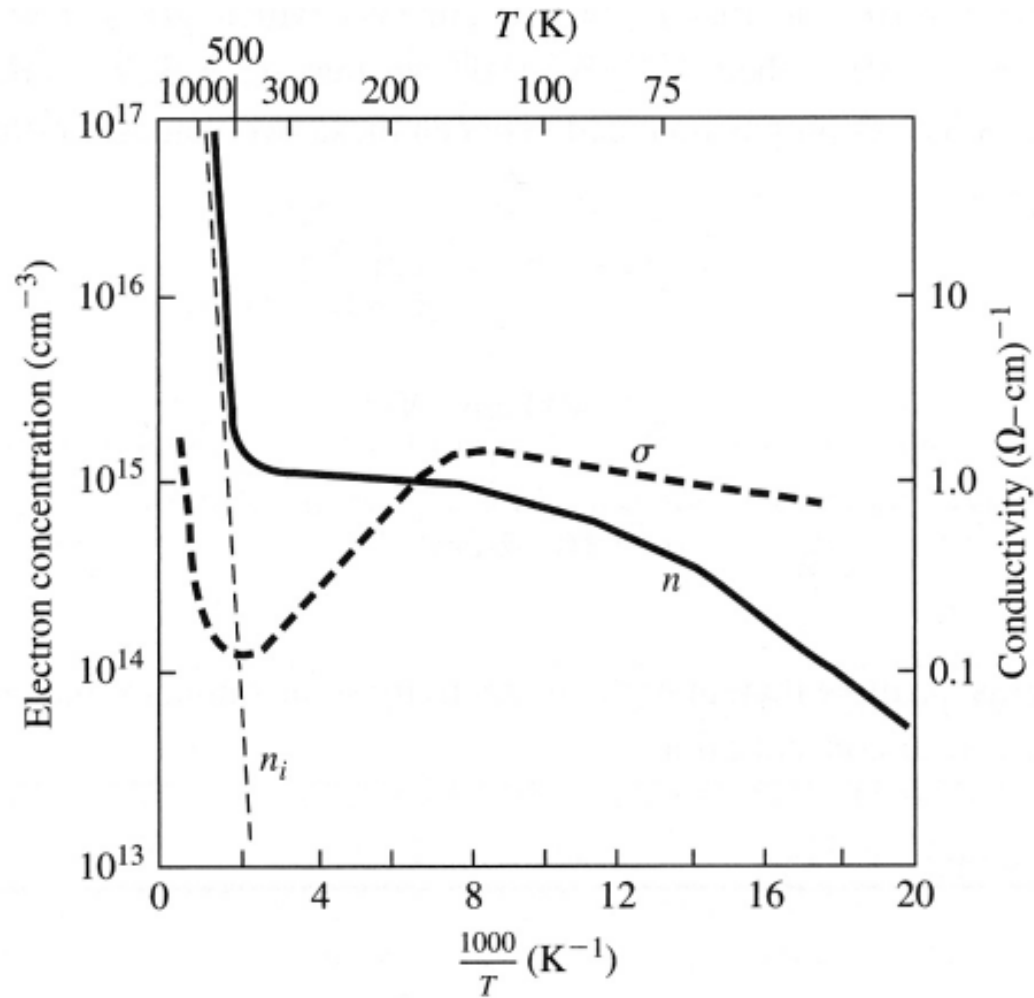
Kombination ergibt:

$$\rho = R \frac{A}{L} = R \frac{Wt}{L}$$

Definition des Schichtwiderstandes:

$$\rho_s \equiv \frac{\rho}{t} = R \frac{W}{L} \quad \left[ \frac{\Omega}{\square} \right]$$

# Die Temperaturabhängigkeit von $\sigma$



Elektronenkonzentration und Leitfähigkeit in Silizium als Funktion der inversen Temperatur (nach Sze).

# Sättigung der Driftgeschwindigkeit

Die Geschwindigkeit eines Ladungsträgers setzt sich aus der thermischen und der Driftgeschwindigkeit zusammen.

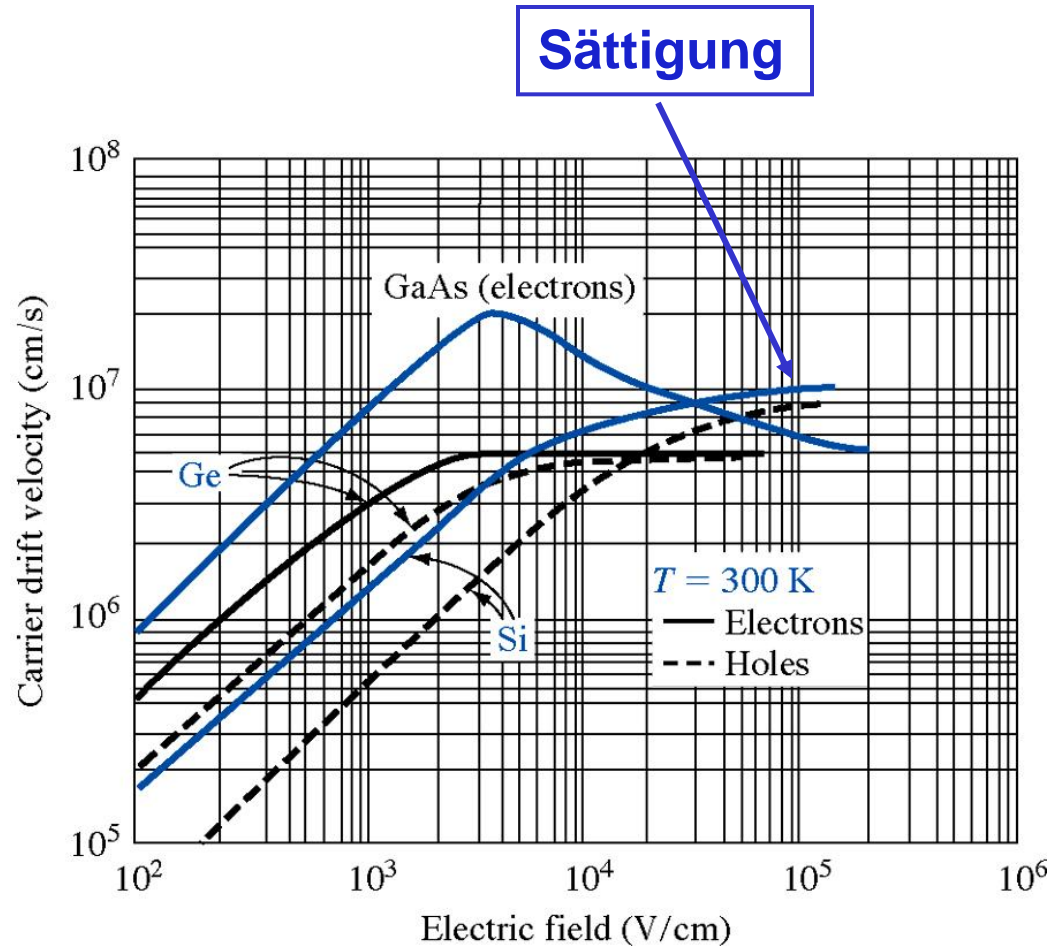
Die Annahme einer vom Feld unabhängigen Beweglichkeit bedeutet, dass die Driftgeschwindigkeit linear mit dem Feld ansteigt.

Für  $T = 300\text{K}$  ist der thermische Teil

$$\frac{1}{2}mv_{\text{th}}^2 = \frac{3}{2}kT = \frac{3}{2}0.0259\text{ eV} = 0.0389\text{ eV}$$

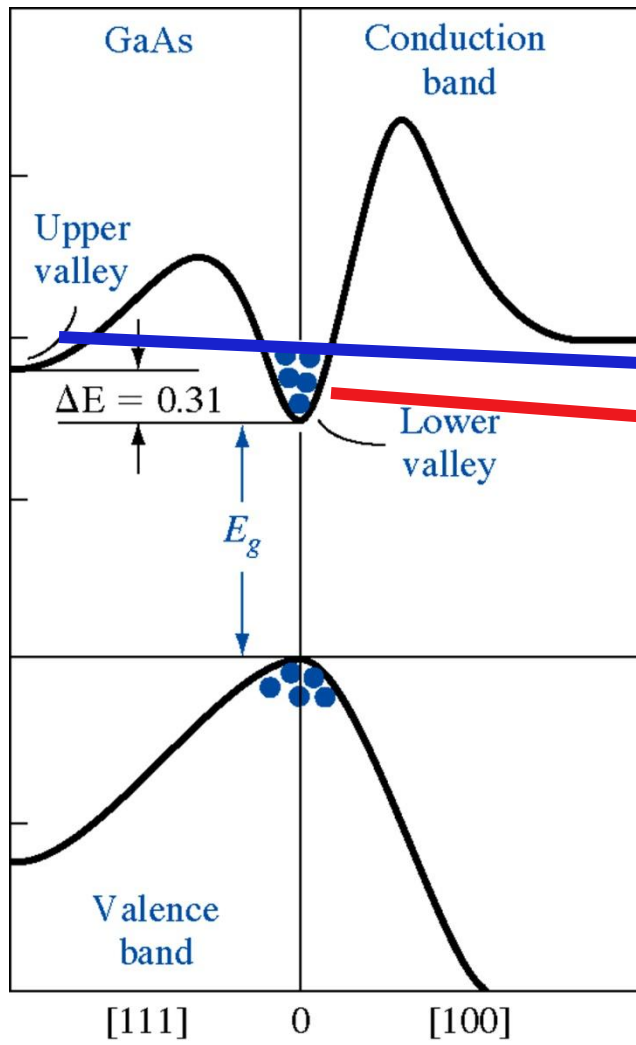
und  $v_{\text{th}} \approx 10^7\text{ cm/s}$



Für  $E > 10^3\text{ V/cm}$  wird die Beweglichkeit eine starke Funktion des elektrischen Feldes!

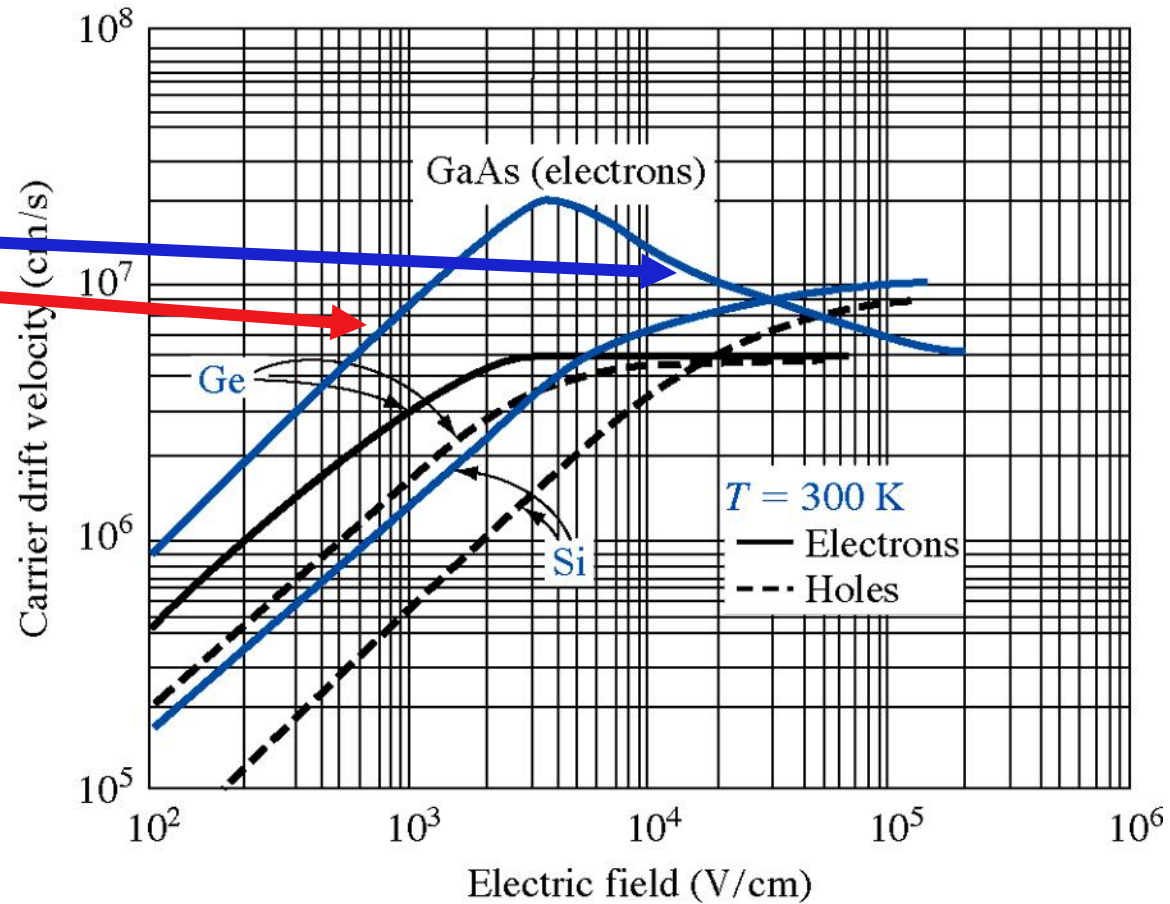


Driftgeschwindigkeit als Funktion der Feldstärke für reines Silizium, Germanium und GaAs.

# Negativ-differentielle Beweglichkeit in GaAs

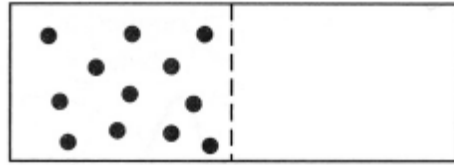


  $m_{dn,lower} = 0.067 m_0$   
  $m_{dn,upper} = 0.55 m_0$



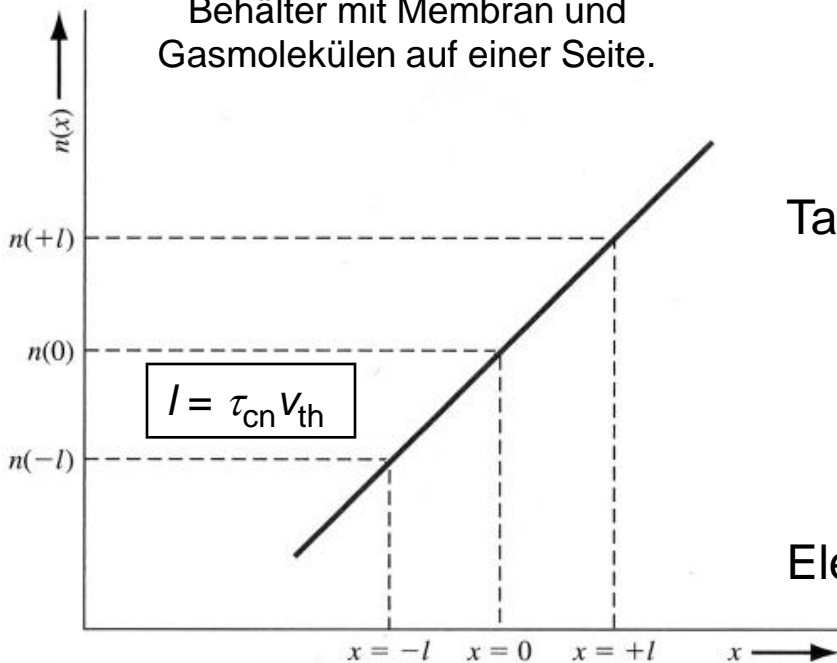
# 5.2 Diffusion

In einem Diffusionsprozess fließen Teilchen von einem Ort hoher Konzentration zu einem Ort niedrigerer Konzentration. Sind diese Teilchen geladen, so entsteht dadurch ein (elektrischer) Diffusionsstrom.



$x = 0$

Behälter mit Membran und Gasmolekülen auf einer Seite.



Verlauf der Elektronenkonzentration.

Bei konstanter Temperatur ist  $v_{th}$  konstant. Sei  $l =$  mittlere freie Weglänge. Teilchenfluss  $F_n$  bei  $x = 0$ :

$$F_n(0) = \frac{1}{2} n(-l) v_{th} - \frac{1}{2} n(+l) v_{th}$$

$$= \frac{1}{2} v_{th} [n(-l) - n(+l)]$$

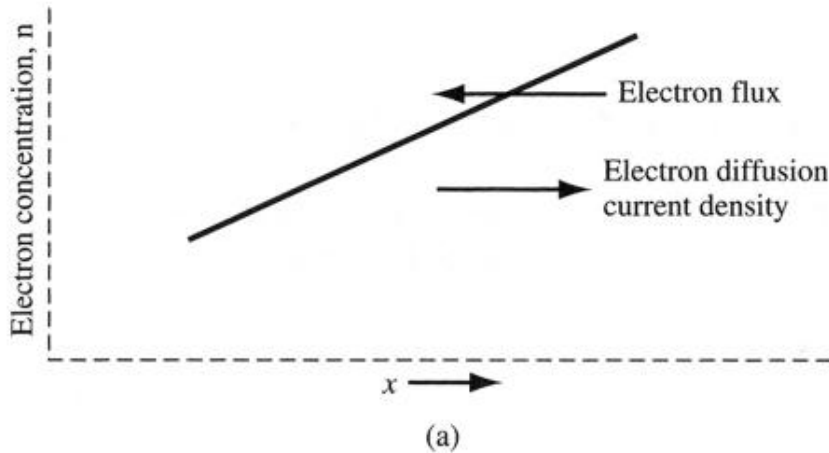
Taylor-Entwicklung:

$$= \frac{1}{2} v_{th} \left\{ \left[ n(0) - l \frac{dn}{dx} \right] - \left[ n(0) + l \frac{dn}{dx} \right] \right\}$$

$$F_n = - v_{th} l \frac{dn}{dx}$$

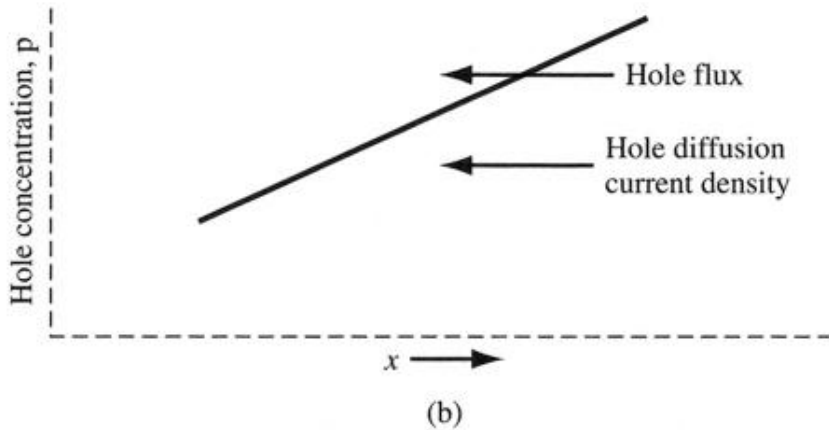
Elektronen-Diffusionsstrom:

$$J_{nx|diff} = - e F_n = + e v_{th} l \frac{dn}{dx} = e D_n \frac{dn}{dx}$$



Elektronen-Diffusionsstrom:

$$J_{nx|\text{diff}} = -eF_n = +ev_{\text{th},n}l_n \frac{dn}{dx} = eD_n \frac{dn}{dx}$$



Löcher-Diffusionsstrom:

$$J_{px|\text{diff}} = eF_p = -ev_{\text{th},p}l_p \frac{dp}{dx} = -eD_p \frac{dp}{dx}$$

(a) Diffusion von Elektronen aufgrund eines Dichtegradienten, (b) Diffusion von Löchern aufgrund eines Dichtegradienten.



# Der totale Strom

Der gesamte im Halbleiter fließende Strom setzt sich aus den Drift- und Diffusionskomponenten von Elektronen und Löchern zusammen:

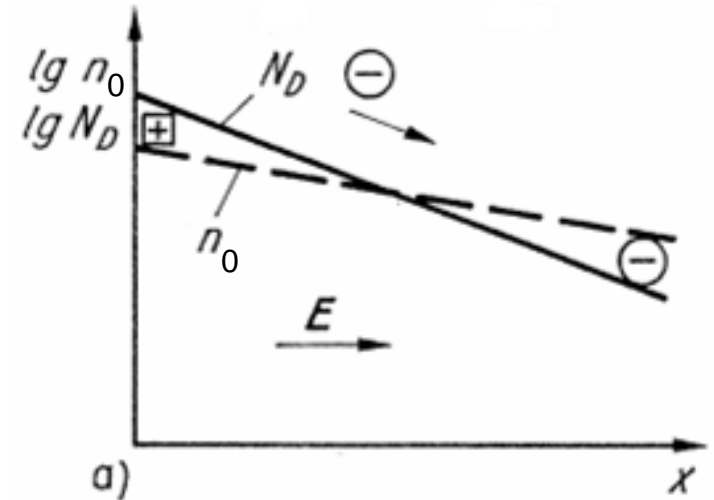
$$\mathbf{J} = en\mu_n\mathbf{E} + ep\mu_p\mathbf{E} + eD_n\nabla n - eD_p\nabla p$$

Dieses (einfachste) Modell des Ladungstransports nennt man ***Drift-Diffusions-Modell.***

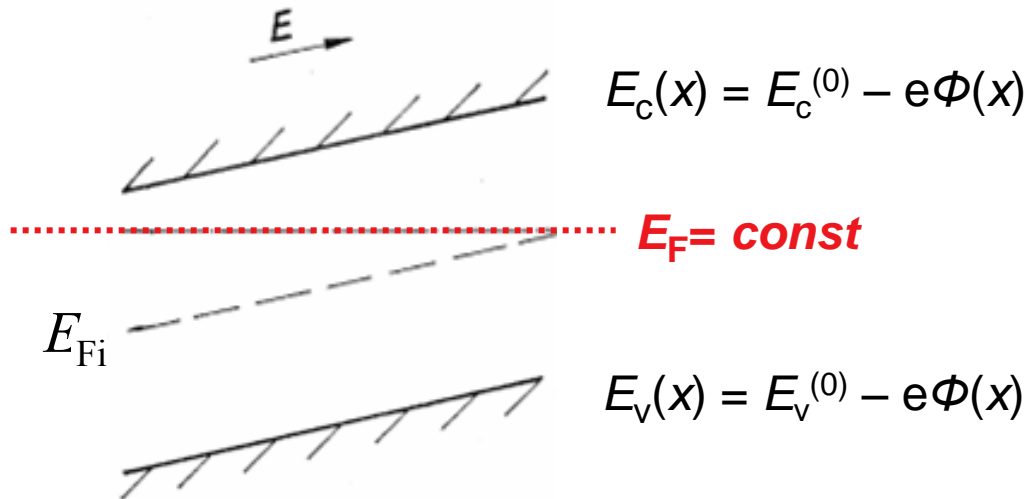
# 5.3 Gradienten der Dotierkonzentration

Ortsabhängige Dotierung führt so lange zu einer Elektronendiffusion nach rechts, bis die entstehende Raumladung ein Gegenfeld aufgebaut hat und den Diffusionsstrom kompensiert.

Dann verschwindet der Gesamtstrom, und das Fermi-Niveau verläuft waagrecht. Die Bandkanten sind geneigt.



a)

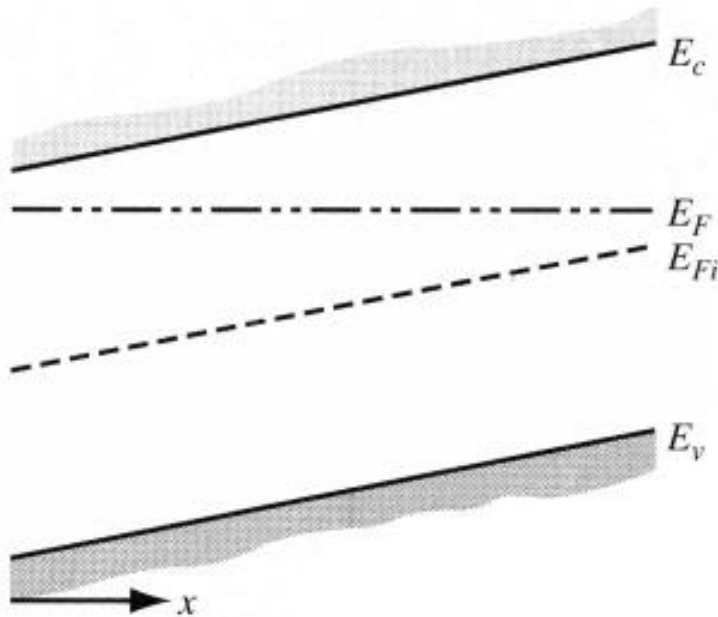


Die Feldstärke  $\mathbf{E}$  in x-Richtung ( $E_x$ ) ergibt sich als negativer Gradient des Potentials  $\Phi(x)$ :

$$E_x = -\frac{d\Phi(x)}{dx} = \frac{1}{e} \frac{dE_C(x)}{dx}$$

# Das eingebaute Feld

Halbleiter mit inhomogener Donatordotierung  $N_d(x)$



$$E_x = - \frac{d\Phi(x)}{dx} = \frac{1}{e} \frac{dE_C(x)}{dx} = \frac{1}{e} \frac{dE_{Fi}}{dx}$$

Bei Quasi-Neutralität ( $n_0 \approx N_d$ ) folgt

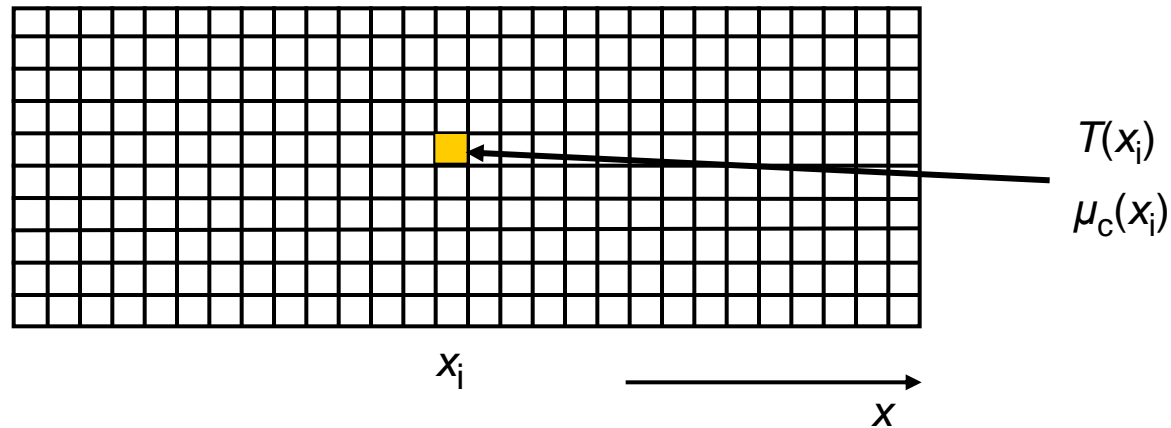
$$n_0(x) \approx N_d(x) = n_i \exp\left(\frac{E_F - E_{Fi}(x)}{kT}\right)$$

und das eingebaute Feld folgt aus der Ableitung der Dichte

$$E_x = - \frac{kT}{e} \frac{d \ln N_d}{dx}$$

Bandkantendiagramm eines Halbleiters mit inhomogener Donator-Konzentration im thermodynamischen Gleichgewicht.

# Physikalische Erklärung der ortsabhängigen Bandkanten (1)



Zur Beschreibung inhomogener Halbleiter denkt man sich das Bauelement in “physikalisch kleine” Volumina aufgeteilt. Diese müssen gross genug sein, um die darin befindlichen Ladungsträger noch als thermodynamisches System ansehen zu können, andererseits aber auch klein genug, um physikalische Grössen wie Dichte, Temperatur, chemisches Potential u.s.w. als kontinuierliche Funktionen des Ortes über das gesamte Bauelement hinweg beschreiben zu können. Annahme: in jedem dieser Volumina herrsche thermodynamisches Gleichgewicht! Dann kann man dort die Gleichgewichts-Verteilungsfunktion (Fermi oder Boltzmann) benutzen und den Teilchen in jeder Zelle eine Temperatur  $T(x_i)$  und ein chemisches Potential  $\mu_c(x_i)$  zuordnen. In verschiedenen Zellen des inhomogenen Halbleiters können diese Parameter verschiedene Werte haben. Die Teilchen-Temperatur  $T$  wird in dieser Vorlesung generell als konstant angenommen. Dotierungsgradienten führen aber in jedem Falle zu einem räumlich variierenden chemischen Potential  $\mu_c(x)$ .

# Physikalische Erklärung der ortsabhängigen Bandkanten (2)

Gleichgewichts-Elektronendichte im *homogenen* Halbleiter:

$$n_0 = N_C \exp\left(-\frac{E_C^{(0)} - \mu_C}{kT}\right)$$

$$= N_C \exp\left(-\frac{E_C^{(0)} - E_F}{kT}\right)$$

Leitungsbandkante des homogenen Halbleiters

Chemisches Potential  $\equiv$  "Ferminiveau"

Gleichgewichts-Elektronendichte im *inhomogenen* Halbleiter:

$$n_0(x) = N_C \exp\left(-\frac{E_C^{(0)} - \mu_C(x)}{kT}\right)$$

ortsabhängiges chemisches Potential wegen inhomogener Dotierung

Im Exponenten wird eine "Null" addiert:  $0 = -e\Phi(x) + e\Phi(x)$

$$n_0(x) = N_C \exp\left(-\frac{\overbrace{E_C^{(0)} - e\Phi(x)}^{E_C(x)} - \overbrace{[\mu_C(x) - e\Phi(x)]}^{E_F}}{kT}}{kT}\right)$$

$E_C(x)$  ist die ortsabhängige Bandkante des inhomogenen Halbleiters

$E_F$  ist das konstante *elektro-chemische* Potential  $\equiv$  "Ferminiveau"

# Physikalische Erklärung der ortsabhängigen Bandkanten (3)

$$n_0(x) = N_C \exp\left(-\frac{E_C(x) - E_F}{kT}\right)$$

Die ortsabhängige Bandkante  $E_C(x) = E_C^{(0)} - e\Phi(x)$  trägt die Ortsabhängigkeit des elektrostatischen Potentials  $\Phi(x)$ , das von der Raumladung im inhomogenen Halbleiter erzeugt wird.

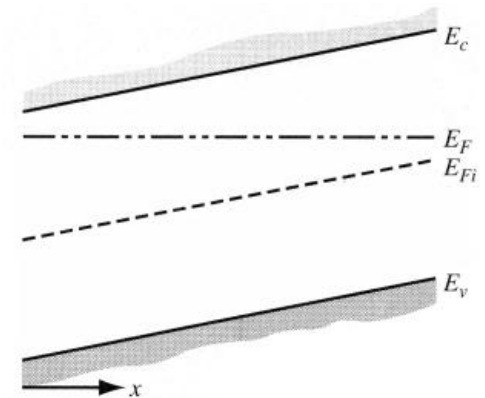
Das elektro-chemische Potential  $E_F = \mu_c(x) - e\Phi(x)$  ist die Summe aus chemischem und elektrostatischem Potential. Sie ist im Gleichgewicht *konstant* (ortsunabhängig,  $E_F = \text{const}$ ). Chemisches und elektrostatisches Potential verlaufen dann also spiegelbildlich und kompensieren einander zur Konstante "Ferminiveau". Im Nichtgleichgewicht (z.B. angelegte Spannung) wird  $E_F$  zur ortsabhängigen Funktion  $E_F(x)$  im obigen Sinne ( $E_F$  variiert von Zelle zu Zelle).

Alternativ kann man die Dichte durch die intrinsische Dichte  $n_i$  und das Intrinsic-Niveau  $E_{Fi}$  ausdrücken. Dazu addiert man in obigem Exponenten  $0 = E_{Fi}^{(0)} - E_{Fi}^{(0)}$

$$E_C^{(0)} - e\Phi(x) - E_F = E_C^{(0)} - E_{Fi}^{(0)} + \underbrace{E_{Fi}^{(0)} - e\Phi(x) - E_F}_{E_{Fi}(x)}$$

und erhält

$$(N5.37) \quad n_0(x) = n_i \exp\left(\frac{E_F - E_{Fi}(x)}{kT}\right) \quad E_{Fi}(x) \text{ verläuft also parallel zu den Bandkanten!}$$



# Die Einstein-Relation

Im stromlosen Fall und bei Quasi-Neutralität  $n(x) \approx N_d(x)$  ist

$$J_n = 0 = e \mu_n N_d(x) E_x + e D_n \frac{dN_d(x)}{dx}$$

Das Elektronengas sei nicht-entartet. Dann gilt für die Dichte im Gleichgewicht:

$$n(x) \approx N_d(x) = N_C \exp\left(-\frac{E_C(x) - E_F}{kT}\right) = N_C \exp\left(-\frac{E_C^{(0)} - e\Phi(x) - E_F}{kT}\right)$$

Bildet man die Ableitung

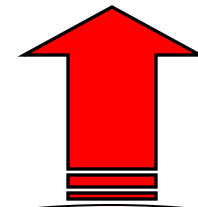
$$\frac{dN_d(x)}{dx} = N_d(x) \frac{e}{kT} \frac{d\Phi(x)}{dx} = N_d(x) \frac{e}{kT} (-E_x)$$

und setzt oben ein, ergibt sich

$$\frac{D_n}{\mu_n} = \frac{kT}{e}$$

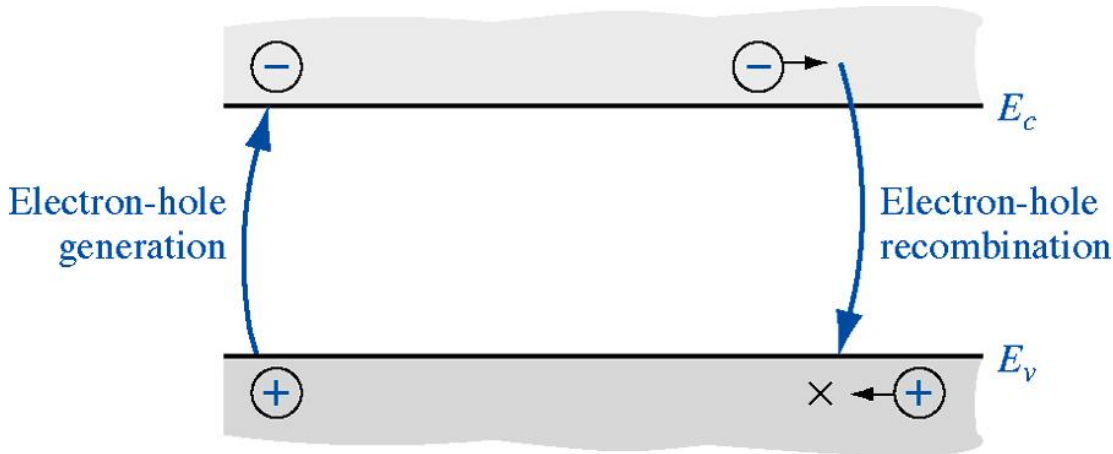
Analog für Löcher. Den Zusammenhang zwischen Beweglichkeit und Diffusionskoeffizient nennt man

$$\frac{D_n}{\mu_n} = \frac{D_p}{\mu_p} = \frac{kT}{e}$$



Einstein-Relation.

# 5.4 Rekombination und Generation



Thermische Generation und Rekombination [ $\#/cm^3s$ ] halten sich im Gleichgewicht die Waage:

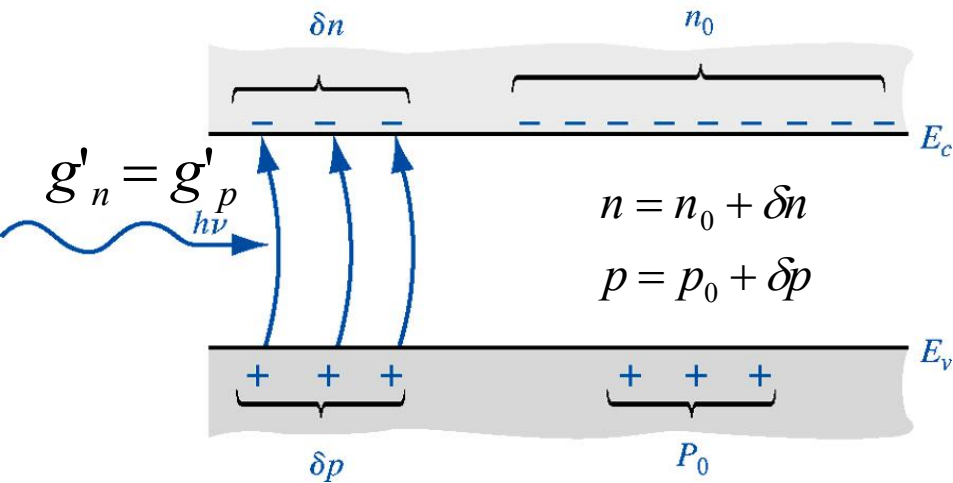
$$G_{n0} = G_{p0} = R_{n0} = R_{p0}$$

Table 6.1 | Relevant notation used in Chapter 6

Symbol	Definition
$n_0, p_0$	Thermal equilibrium electron and hole concentrations (independent of time and also usually position).
$n, p$	Total electron and hole concentrations (may be functions of time and/or position).
$\delta n = n - n_0$ $\delta p = p - p_0$	Excess electron and hole concentrations (may be functions of time and/or position).
$g'_n, g'_p$	Excess electron and hole generation rates.
$R'_n, R'_p$	Excess electron and hole recombination rates.
$\tau_{n0}, \tau_{p0}$	Excess minority carrier electron and hole lifetimes.



# Rekombination und Generation von Überschuss-Ladungsträgern



Die Änderung der Überschussladung muss der Abweichung vom Gleichgewicht proportional sein:

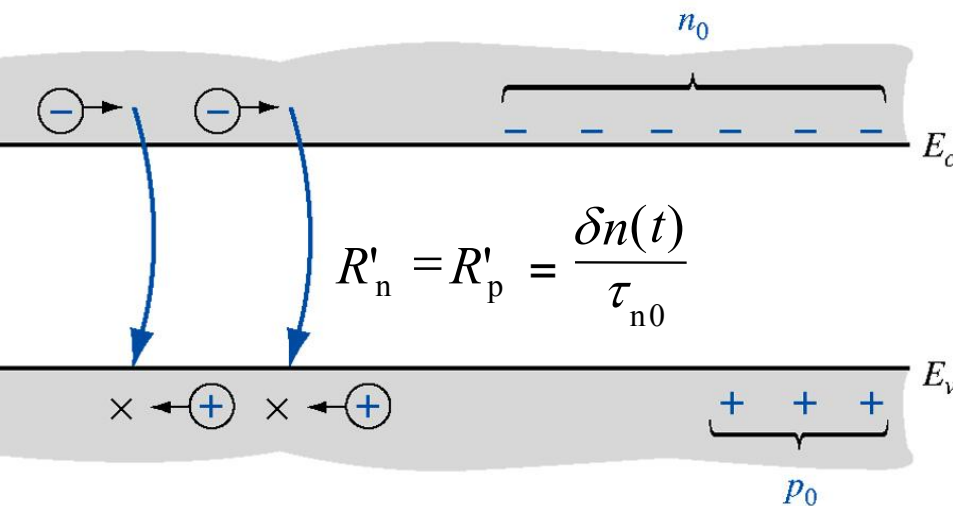
$$\begin{aligned} \frac{dn(t)}{dt} &= \alpha_r [n_i^2 - n(t)p(t)] \\ &= -\alpha_r \delta n(t) [(n_0 + p_0) + \delta n(t)] \end{aligned}$$

Für  $p$ -Material und Niedrig-Injektion gilt:

$$p_0 \gg n_0 \quad \text{und} \quad \delta n(t) \ll p_0$$

$$\frac{d(\delta n(t))}{dt} = -\alpha_r p_0 \delta n(t)$$

$$\delta n(t) = \delta n(0) e^{-\alpha_r p_0 t} = \delta n(0) e^{-t/\tau_{n0}}$$



Definition der Lebensdauer von Überschussminoritäten:

$$\tau_{n0} = \frac{1}{\alpha_r p_0}$$

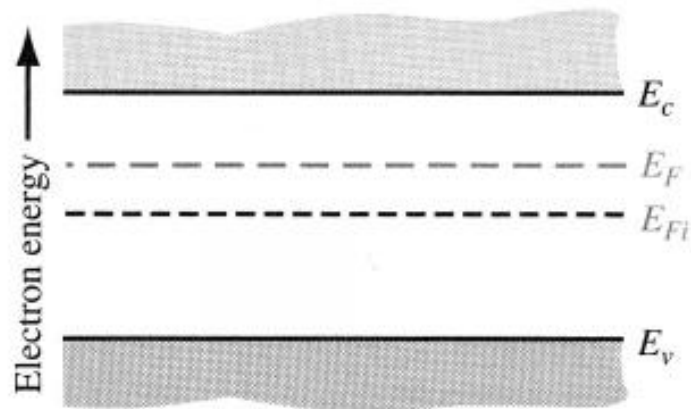
# 5.5 Quasi-Ferminiveaus

Gleichgewichts-Dichten der Ladungsträger:

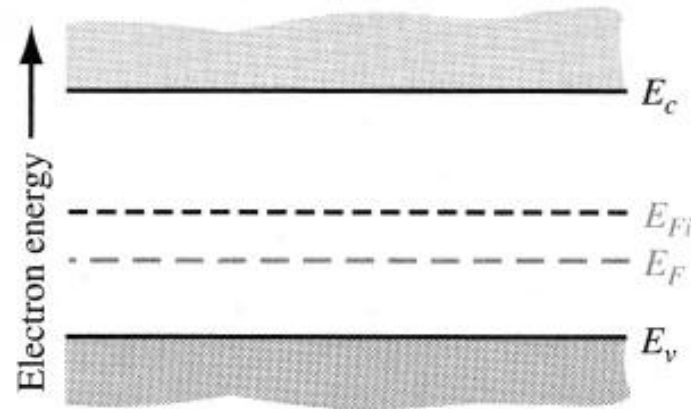
Die Dichten hängen exponentiell von der Fermienergie ab.

$$n_0 = n_i \exp\left(\frac{E_F - E_{Fi}}{kT}\right)$$

$$p_0 = n_i \exp\left(\frac{E_{Fi} - E_F}{kT}\right)$$



(a)



(b)

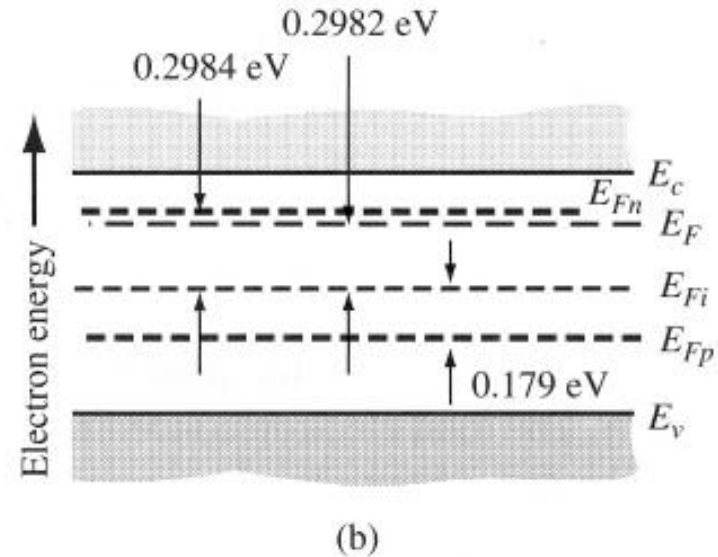
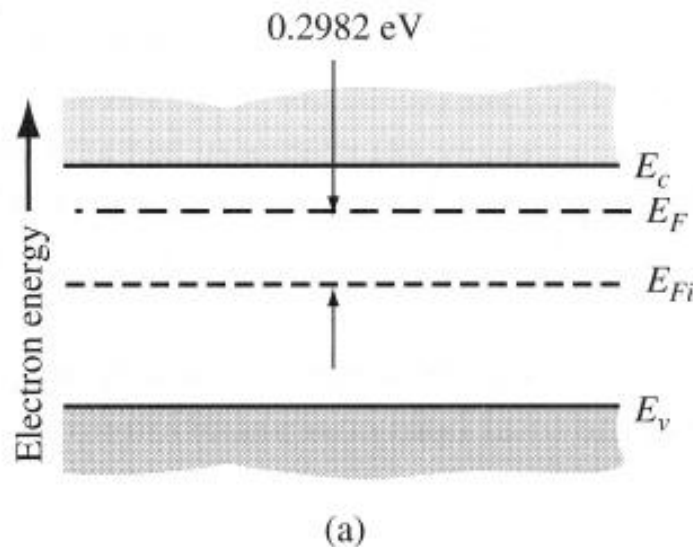
Bandkantendiagramm im thermodynamischen Gleichgewicht für (a) einen n-Typ-Halbleiter, (b) einen p-Typ-Halbleiter.

Wie sieht das im Nichtgleichgewicht aus?

Unter der Annahme, dass die Verteilungen der Elektronen und Löcher im Halbleiter voneinander unabhängig sind, kann man separate Fermi-niveaus auch für den Nichtgleichgewichtsfall definieren:

$$n = n_i \exp\left(\frac{E_{Fn} - E_{Fi}}{kT}\right)$$

$$p = n_i \exp\left(\frac{E_{Fi} - E_{Fp}}{kT}\right)$$



- (a) Bandkantendiagramm im thermodynamischen Gleichgewicht für  $N_D = 10^{15} \text{ cm}^{-3}$  und  $n_i = 10^{10} \text{ cm}^{-3}$ .  
 (b) Quasi-Fermi-Niveaus für Elektronen und Löcher, wenn Überschussladungen der Dichte  $10^{13} \text{ cm}^{-3}$  existieren.





# 6. Der pn-Übergang

6.1 Die Struktur des pn-Übergangs (N7.1)

6.2 Der Fall  $V = 0$  (N7.2)

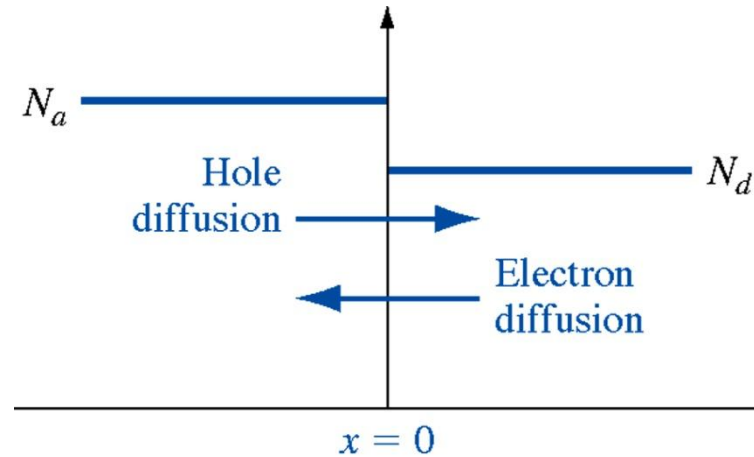
6.3 Der Fall des gesperrten pn-Übergangs (N7.3)

# 6.1 Die Struktur des pn-Übergangs



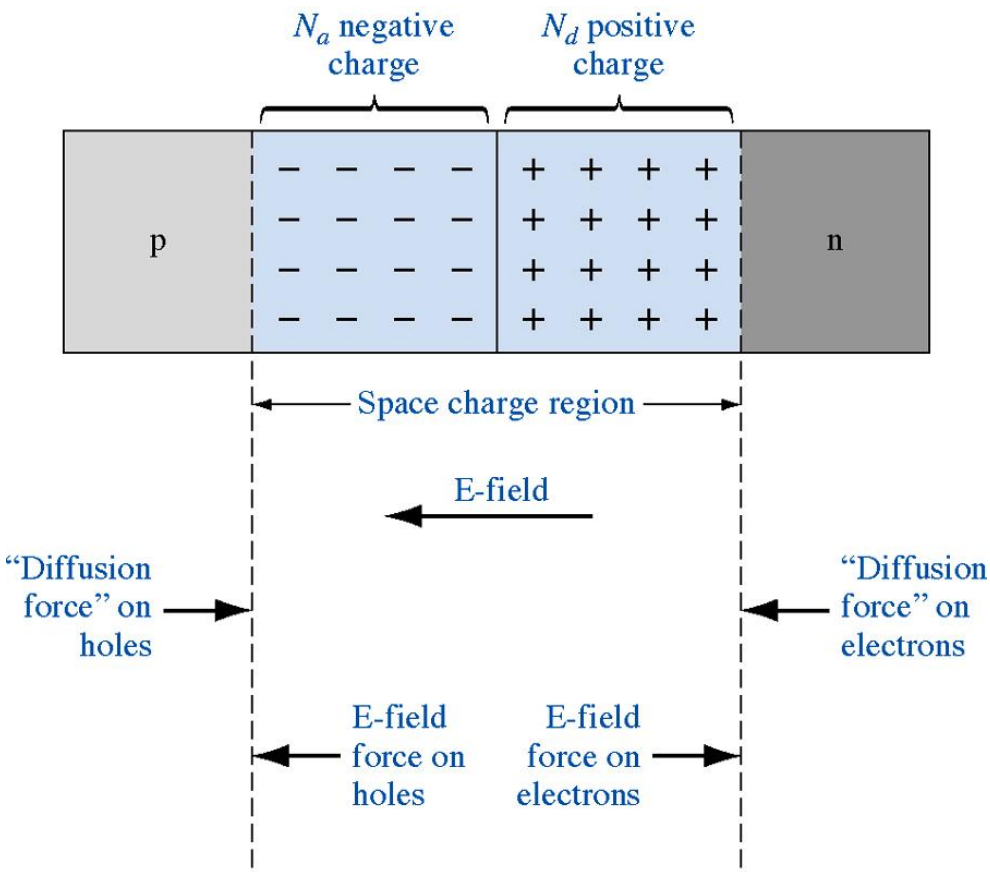
Metallurgical  
junction

(a)



(b)

- pn-Übergänge entstehen, wenn ein p- und ein n-leitendes Halbleitergebiet aneinanderstossen (a).
- Die Konzentrationsgradienten treiben einen Löcher-Diffusionsstrom aus dem p- ins n-Gebiet und eine Elektronen-Diffusionsstrom aus dem n- ins p-Gebiet (b).
- Als Minoritäts-Ladungsträger rekombinieren die Löcher im n-Gebiet und die Elektronen im p-Gebiet. Es entsteht ein sogenanntes *Verarmungsgebiet* (depletion region) und gleichzeitig eine *Raumladungsschicht* (space charge region) wegen der ionisierten Dotanden.



- Das elektrische Feld der Raumladungszone übt eine Gegenkraft auf die diffundierenden Träger aus.
- Im dynamischen Gleichgewicht halten sich Diffusion und Drift die Waage. Die Raumladungsschicht bleibt endlich dick und wird von den *neutralen Bahngebieten* begrenzt.
- Der Verlauf des elektrostatischen Potentials wird durch die *Poisson-Gleichung* bestimmt. In 1D:

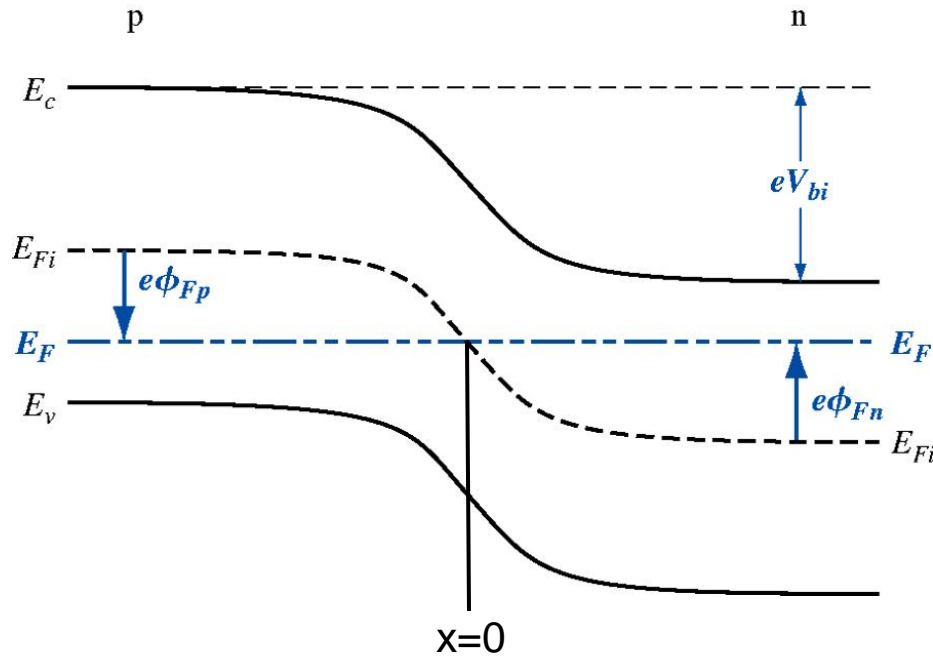
$$-\frac{d^2\Phi}{dx^2} = \frac{dE}{dx} = \frac{e}{\epsilon} [N_d^+(x) - N_a^-(x) + p(x) - n(x)]$$

**Ableitung:**

Satz von Gauss:  $\nabla \cdot E = \frac{\rho}{\epsilon}$  (integral:  $\oiint \mathbf{D} \cdot d\mathbf{A} = Q$ ) und  $-\nabla\Phi = E$



## 6.2 Der Fall $V = 0$



Bandkanten-Diagramm eines pn-Übergangs im thermodynamischen Gleichgewicht

Ohne äussere Spannung ( $V=0$ ) fliesst kein Netto-Elektronenstrom:

$$J_n = 0 = e\mu_n n(x) E_x + eD_n \frac{dn(x)}{dx}$$

Benutzt man die Einstein-Relation (Boltzmann-Statistik)

$$\frac{D_n}{\mu_n} = \frac{kT}{e} \stackrel{\text{def}}{=} V_t$$

wird daraus

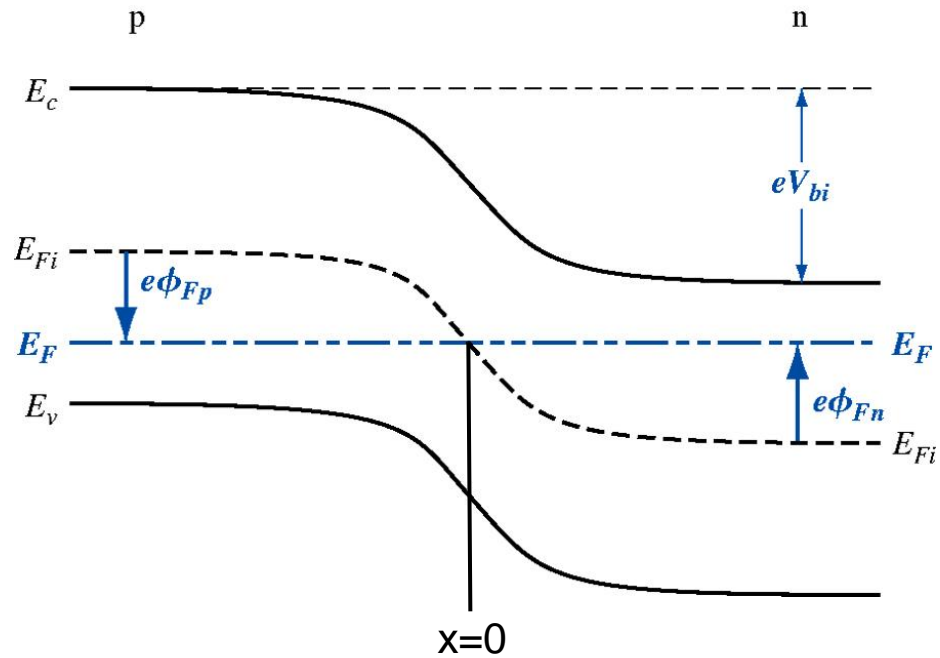
$$\frac{1}{n(x)} \frac{dn(x)}{dx} = \frac{1}{V_t} \frac{d\Phi(x)}{dx}$$

Integration von  $x=0$  bis  $x$  ergibt  $\ln(n) \Big|_{x=0}^x = \frac{\Phi(x) - \Phi(x=0)}{V_t}$

und damit die ortsabhängige Dichte als Funktion des Potentials

$$n(x) = n_i \exp\left(\frac{\Phi(x)}{V_t}\right) \quad \text{weil } n(x=0) = n_i \quad \text{und} \quad \Phi(x=0) = 0$$

# Das eingebaute Potential (built-in potential)



Bandkanten-Diagramm eines pn-Übergangs im thermodynamischen Gleichgewicht

$$n(x) = n_i \exp\left(\frac{\Phi(x)}{V_t}\right)$$

$$p(x) = n_i \exp\left(-\frac{\Phi(x)}{V_t}\right)$$

In den Bahngebieten gilt:

$$n(x \rightarrow \infty) = N_d$$

$$p(x \rightarrow -\infty) = N_a$$

Bildet man das Produkt, folgt

$$n(\infty)p(-\infty) = N_d N_a = n_i^2 \exp\left(\frac{\Phi(\infty) - \Phi(-\infty)}{V_t}\right) = n_i^2 \exp\left(\frac{V_{bi}}{V_t}\right)$$

Auflösung nach dem *eingebauten Potential*:

$$V_{bi} = V_t \ln \frac{N_d N_a}{n_i^2}$$

## Das eingebaute Potential (2)

$$V_{bi} = V_t \ln \frac{N_d N_a}{n_i^2}$$

kann umgeschrieben werden, indem die Elektronendichten in den Bahngelieten benutzt werden

$$n_{n0} = N_d \qquad n_{p0} = \frac{n_i^2}{N_a}$$

$$V_{bi} = V_t \ln \frac{n_{n0}}{n_{p0}}$$

Achtung! Von hier ab Wechsel in der Bezeichnung:

$$N_d = N_d^+ - N_a^- \quad \text{im n-Gebiet}$$

$$N_a = N_a^- - N_d^+ \quad \text{im p-Gebiet}$$

# Elektrisches Feld

Vereinfachende Annahmen:

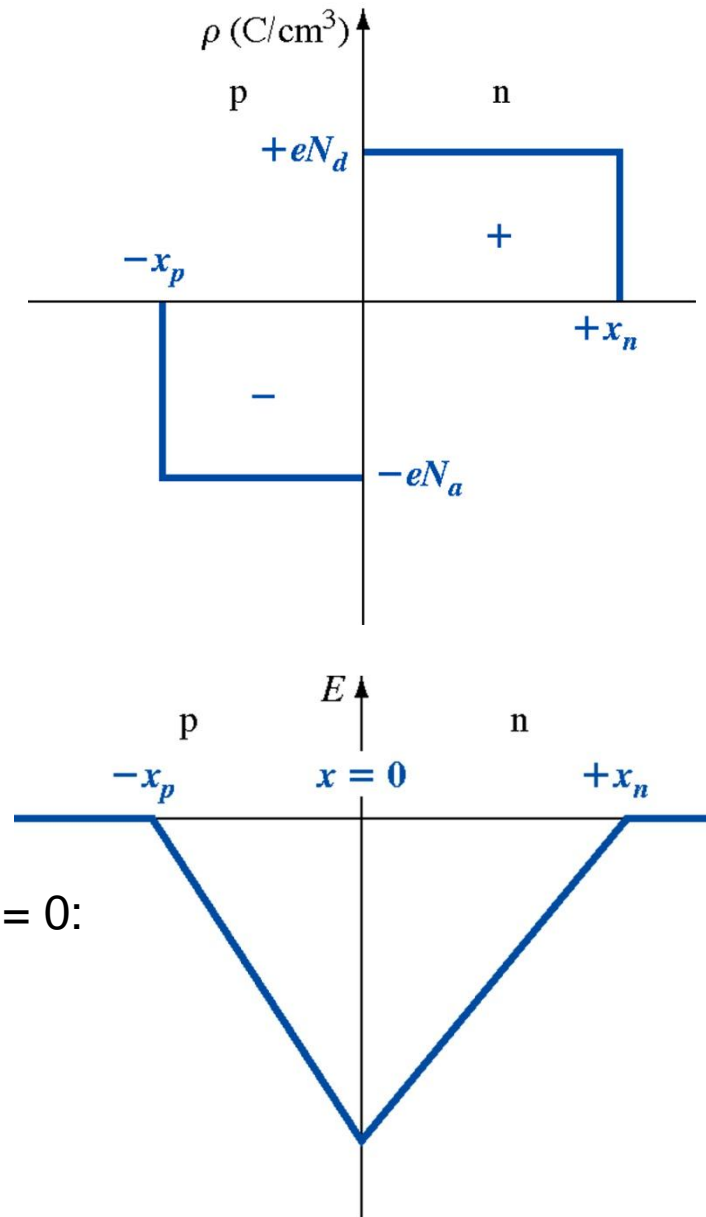
- konstante n- und p-Dotierung,
- abrupter Übergang von Verarmungszone zu neutralem Bahnggebiet,
- keine freien Ladungsträger in den Verarmungs-zonen (Schottky-Näherung, depletion approximation)

Damit wird die rechte Seite der Poisson-Gleichung konstant in den Verarmungsgebieten und Null sonst:

$$-\frac{d^2\Phi}{dx^2} = \frac{dE}{dx} = \begin{cases} \frac{e}{\varepsilon} N_d & 0 \leq x \leq x_n \\ -\frac{e}{\varepsilon} N_a & -x_p \leq x < 0 \end{cases}$$

Einmal integrieren mit Randbedingung  $E(-x_p) = E(x_n) = 0$ :

$$E(x) = \begin{cases} \frac{e}{\varepsilon} N_d (x - x_n) & 0 \leq x \leq x_n \\ -\frac{e}{\varepsilon} N_a (x + x_p) & -x_p \leq x < 0 \end{cases}$$



# Elektrostatistisches Potential

Nochmaliges integrieren liefert das Potential

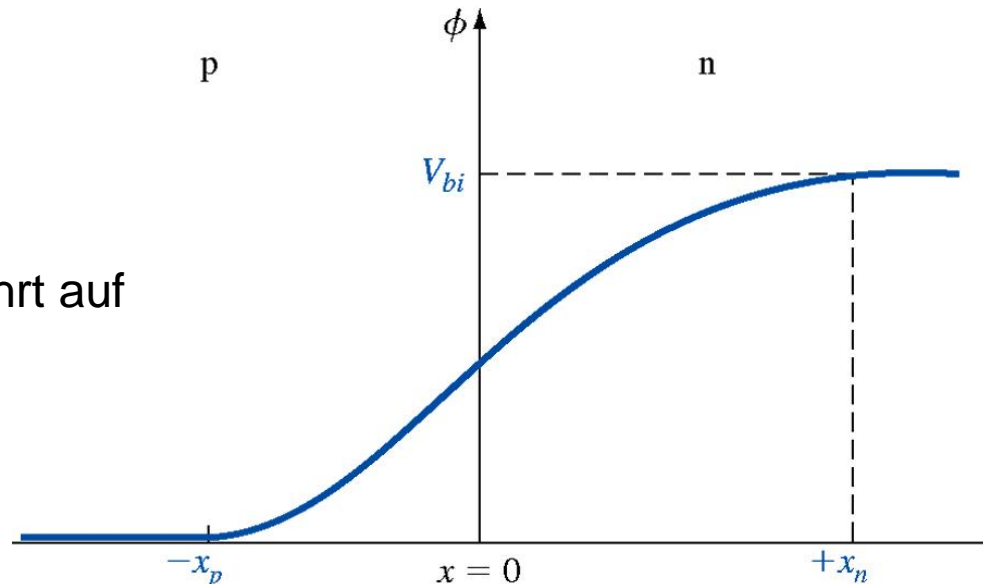
$$\Phi(x) = - \int_{-x_p}^x E(x') dx' = \frac{e}{2\epsilon} N_a (x + x_p)^2 \quad -x_p \leq x < 0$$

$$\Phi(x) = - \int_{x_n}^x E(x') dx' + C = - \frac{e}{2\epsilon} N_d (x - x_n)^2 + C \quad 0 \leq x \leq x_n$$

Die Integrationskonstante  $C$  erhält man aus  $\Phi(x_n) = V_{bi}$  ( $\Rightarrow C = V_{bi}$ )

Die Stetigkeit des Potentials bei  $x = 0$  führt auf

$$V_{bi} = \frac{e}{2\epsilon} (N_a x_p^2 + N_d x_n^2)$$



# Weite der Raumladungsschicht

$$V_{bi} = \frac{e}{2\epsilon} (N_a x_p^2 + N_d x_n^2)$$

Die Stetigkeit des  $E$ -Feldes bei  $x = 0$  ergibt andererseits

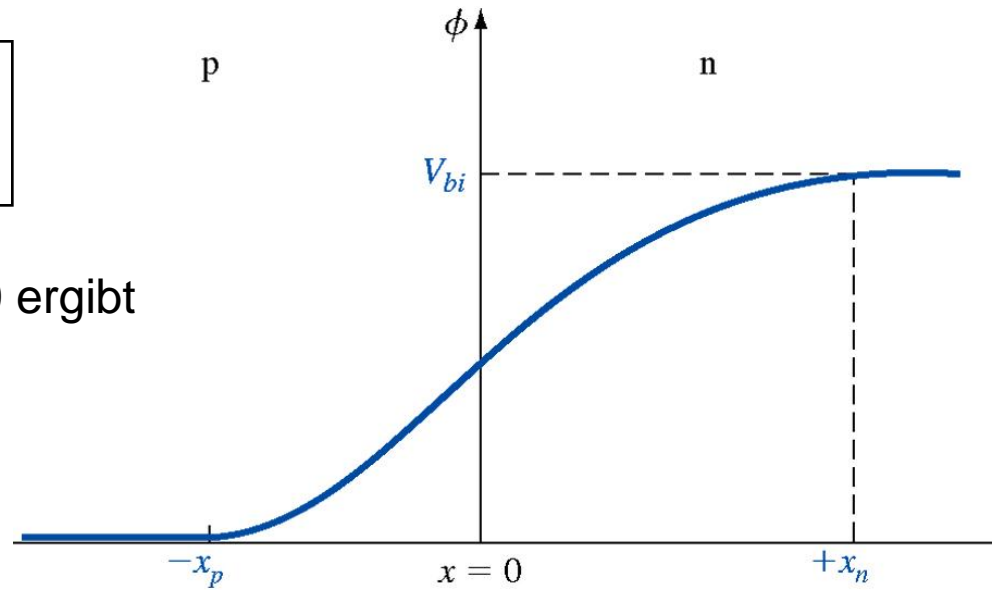
$$N_d x_n = N_a x_p$$

Einsetzen von  $x_p = x_n N_d / N_a$  in obige Gleichung und Auflösung nach  $x_n$ :

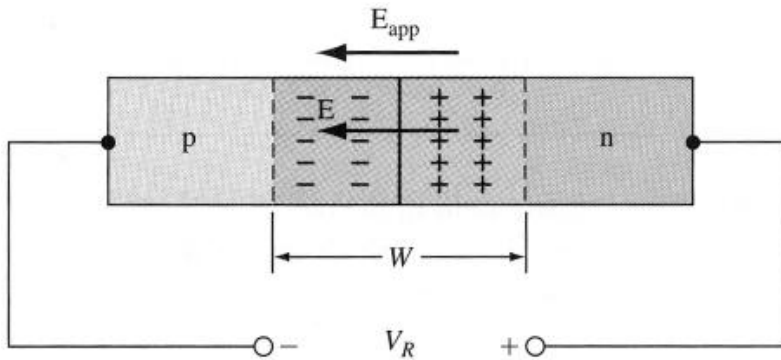
$$x_n = \left\{ \frac{2\epsilon V_{bi}}{e} \frac{N_a}{N_d(N_a + N_d)} \right\}^{1/2} \quad x_p = \left\{ \frac{2\epsilon V_{bi}}{e} \frac{N_d}{N_a(N_a + N_d)} \right\}^{1/2}$$

Die *Raumladungszonenweite*  $W = x_n + x_p$  wird damit

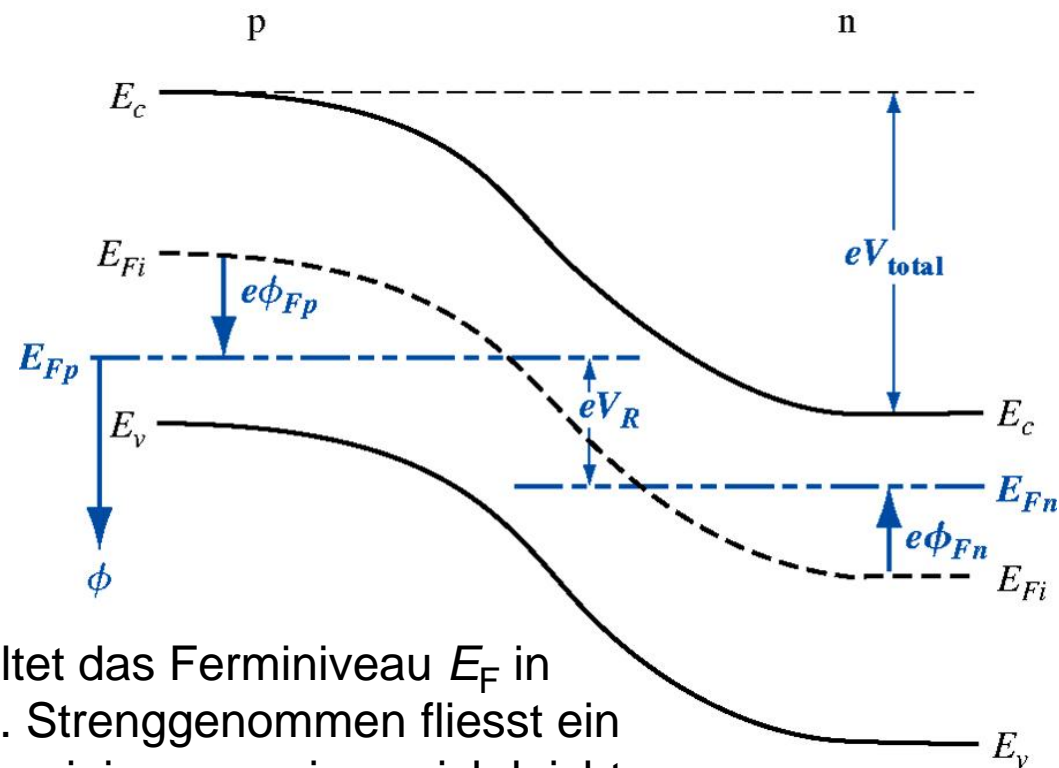
$$W = \left\{ \frac{2\epsilon V_{bi}}{e} \frac{N_a + N_d}{N_a N_d} \right\}^{1/2}$$



# 6.3 Der Fall des gesperrten pn-Übergangs



pn-Übergang bei angelegter Sperrspannung. Das Feld, das von der Raumladung induziert wird, hat dieselbe Richtung wie das Feld, welches von  $V_R$  induziert wird.



Legt man eine Spannung  $V_R$  an, spaltet das Fermi-niveau  $E_F$  in Quasi-Fermi-niveaus  $E_{Fn}$  und  $E_{Fp}$  auf. Strenggenommen fließt ein kleiner Sperrstrom, und die Quasi-Fermi-niveaus neigen sich leicht. Davon ist hier abgesehen.

Merke: *Spannung = Differenz von elektrochemischen Potentialen*

$$V_{\text{total}} = V_{\text{bi}} + V_R$$

$$W = \left\{ \frac{2\epsilon(V_{\text{bi}} + V_R)}{e} \frac{N_a + N_d}{N_a N_d} \right\}^{1/2}$$

# Die Sperrschichtkapazität

Definition:

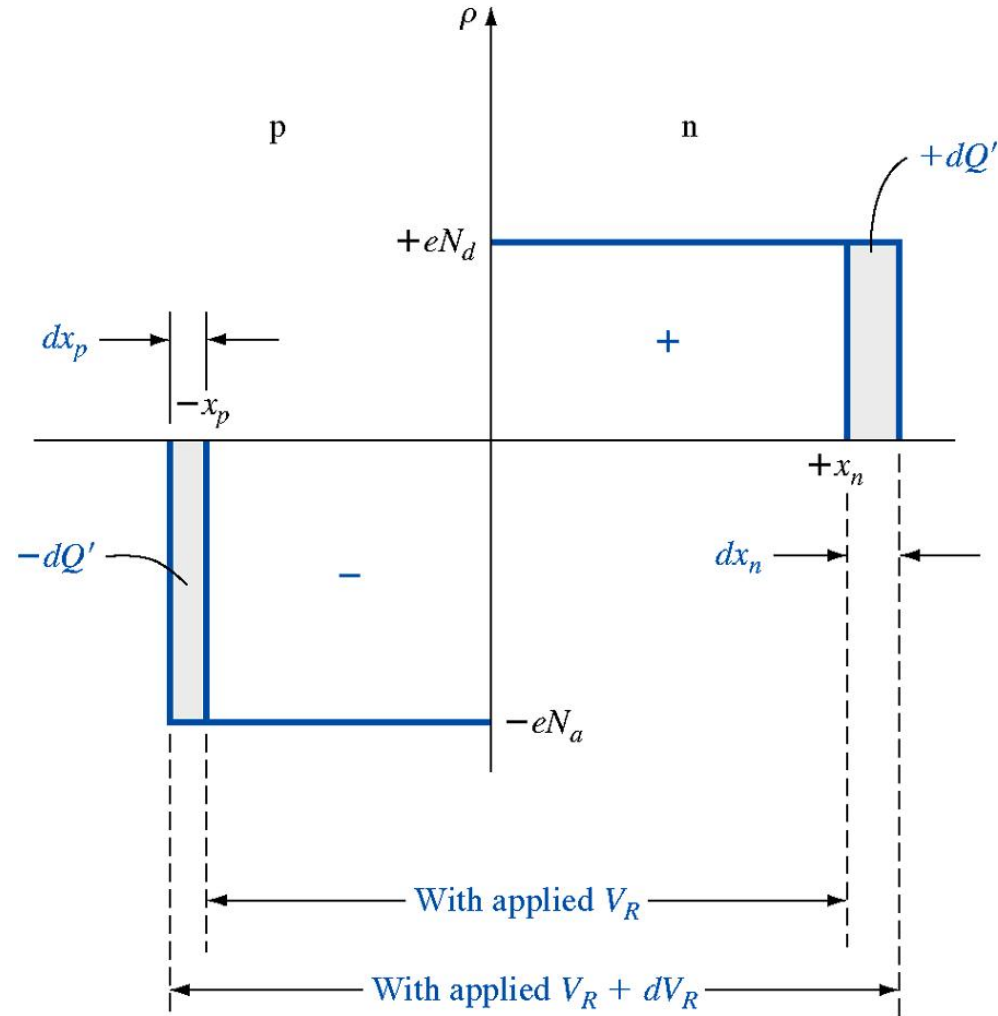
$$C' = \frac{dQ'}{dV_R}$$

$$dQ' = e N_d dx_n$$

$$x_n = \left\{ \frac{2 \varepsilon (V_{bi} + V_R)}{e} \frac{N_a}{N_d (N_a + N_d)} \right\}^{1/2}$$

$$C' = e N_d \frac{dx_n}{dV_R} =$$

$$= \left\{ \frac{e \varepsilon}{2(V_{bi} + V_R)} \frac{N_a N_d}{(N_a + N_d)} \right\}^{1/2}$$







# 7. Die pn-Diode

**7.1 Stromfluss im pn-Übergang (N8.1)**

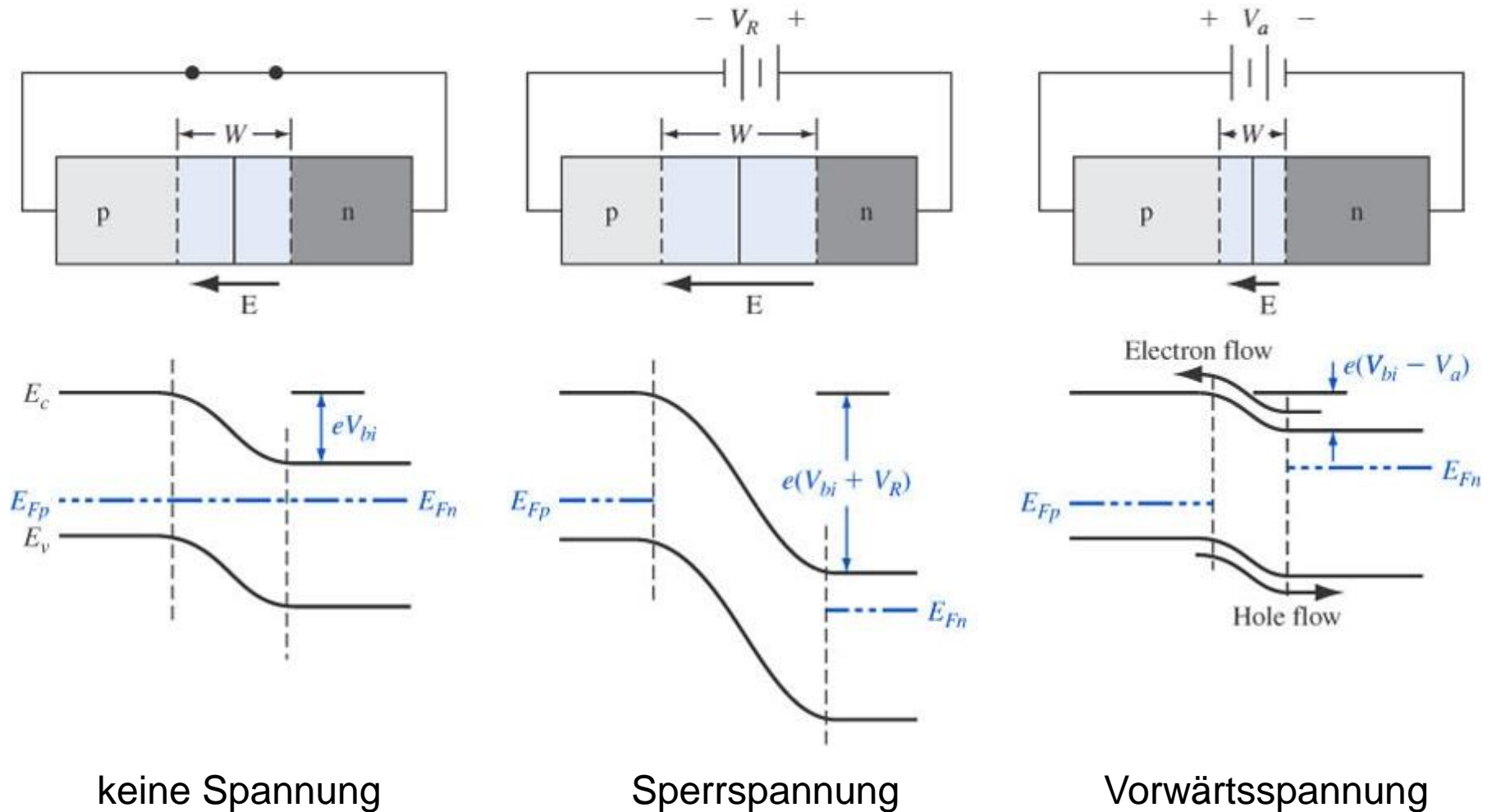
**7.2 Kleinsignal-Verhalten der pn-Diode (N8.2)**

**7.3 Generations-Rekombinationsströme (N8.3)**

**7.4 Durchbrüche von Dioden (N8.4)**

# 7.1 Stromfluss im pn-Übergang

## Bandkanten-Diagramme

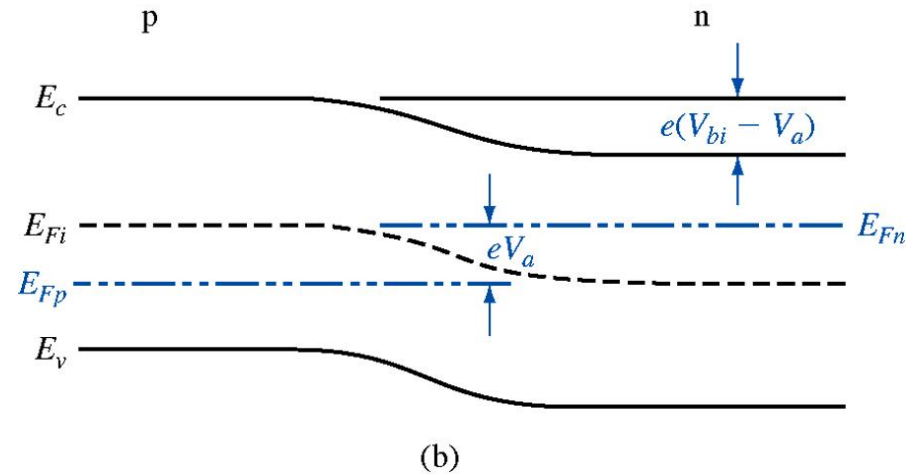
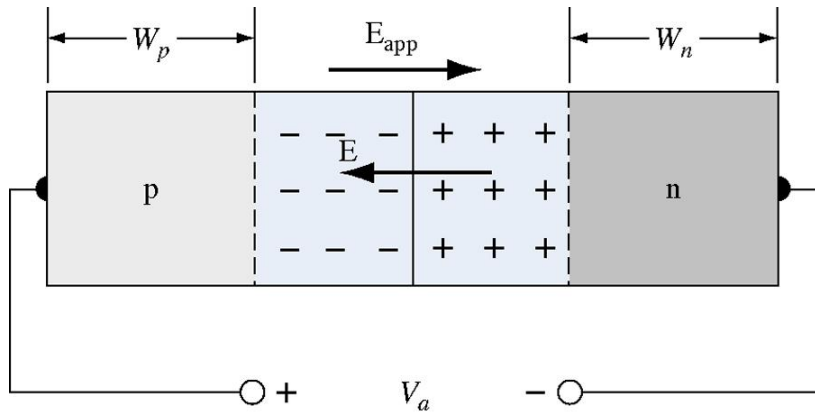


# Ideale Strom-Spannungs-Charakteristik

Annahmen:

- Abrupter pn-Übergang in Schottky-Näherung (keine Träger in der Raumladungszone).
- Maxwell-Boltzmann-Statistik.
- Niedrig-Injektion.
- Totaler Strom ist konstant quer durch die gesamte pn-Struktur.
- Die individuellen Ströme der Elektronen und Löcher sind kontinuierliche Funktionen des Ortes.
- Die individuellen Ströme der Elektronen und Löcher sind konstant quer durch die Raumladungszone (keine Rekombination in diesem Gebiet).

# Minoritätsladungsträgerdichten an den Rändern der Raumladungszone



$$V_{bi} = V_t \ln \frac{n_{n0}}{n_{p0}}$$

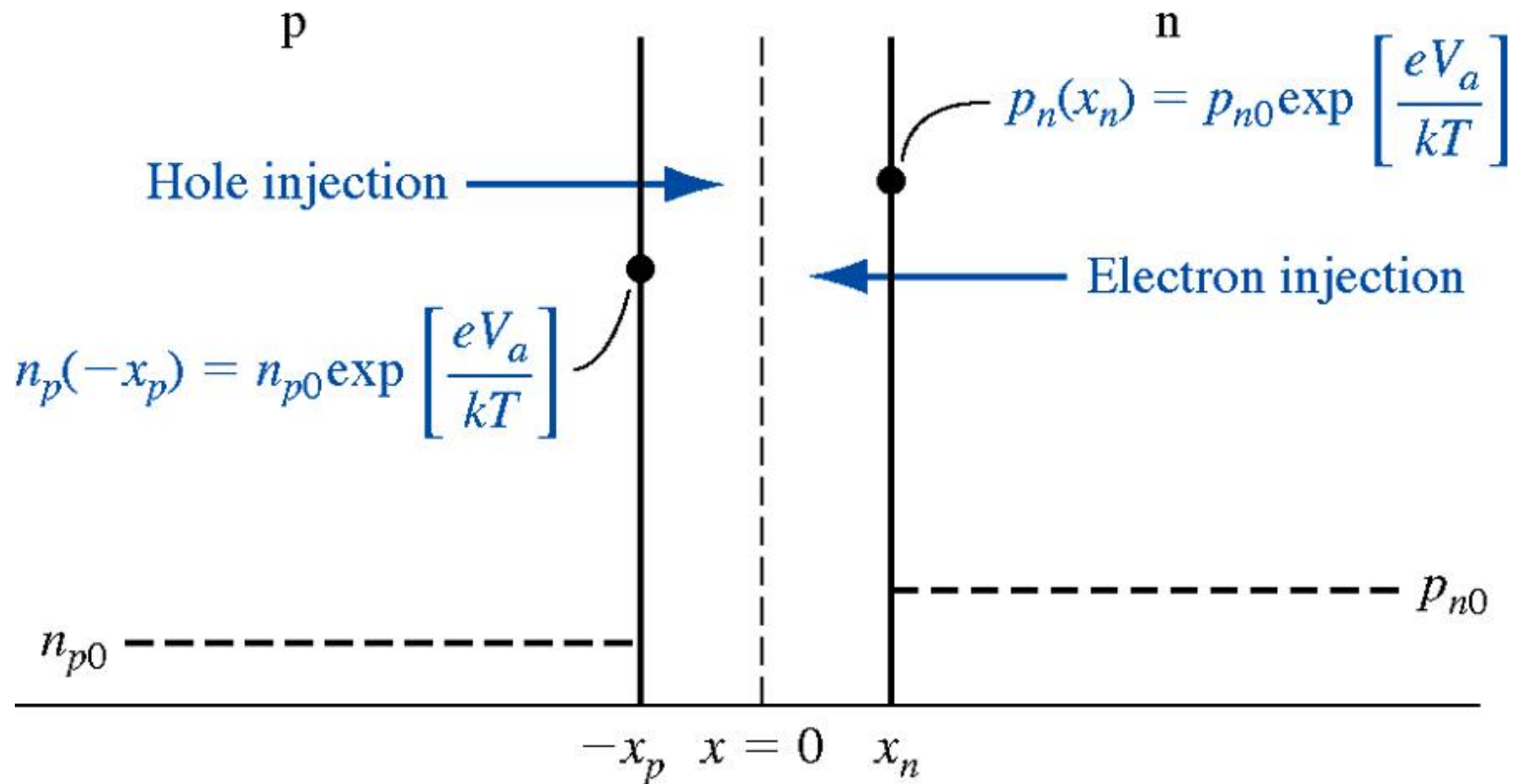


$$n_{p0} = n_{n0} \exp\left[-\frac{V_{bi}}{V_t}\right]$$

$$n_p = n_{n0} \exp\left[-\frac{e(V_{bi} - V_a)}{kT}\right]$$

$$= n_{p0} \exp\left[\frac{eV_a}{kT}\right]$$

- Die ins p-Gebiet injizierten Elektronen und die ins n-Gebiet injizierten Löcher sind dort Minoritäts- und Überschuss-Ladungsträger. Sie diffundieren und rekombinieren in den Bahngebieten.
- Daher gelten die Formeln nur am jeweiligen Rand der Raumladungszone.



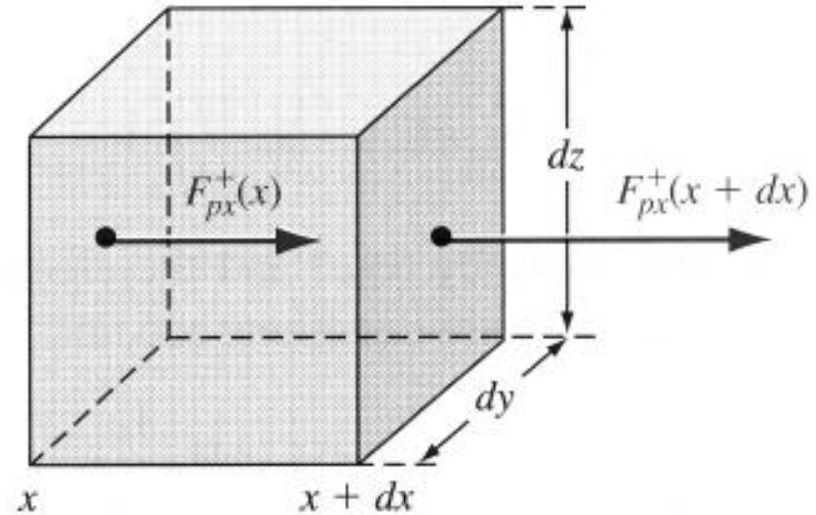
# Die Kontinuitätsgleichung

- Zur Ableitung der Stromdichten in der Diode brauchen wir die *Kontinuitätsgleichung*.

$p dx dy dz$  ist die Zahl der Löcher im rechts dargestellten infinitesimalen Würfel. Diese Zahl kann sich während des Zeitintervalls  $dt$  auf zweierlei Art ändern:

1.) Bei  $x+dx$  fließen während  $dt$  weniger (oder mehr) Löcher aus dem Würfel raus als bei  $x$  reinfließen. Der Fluss sei  $F_{px}^+$  (=Zahl/( $cm^2s$ )). Die zeitliche Änderung der Löcherzahl aufgrund des Fluss-Unterschiedes ist

$$\begin{aligned} \frac{\partial}{\partial t} p dx dy dz &= [F_{px}^+(x) - F_{px}^+(x+dx)] dy dz \\ &= - \frac{\partial F_{px}^+}{\partial x} dx dy dz \end{aligned}$$



Differentielles Volumenelement und x-Komponente des Teilchenflusses der Löcher.

2.) Während  $dt$  können im Würfel auch Löcher verschwinden (=Rekombination) bzw. neu entstehen (=Generation). Die zeitliche Änderung der Löcherzahl aufgrund der Netto-Generation ist

$$\frac{\partial}{\partial t} p dx dy dz = \left[ g_p - \frac{\delta p}{\tau_{pt}} \right] dx dy dz$$

Generations- und Rekombinationsraten wurden in Kap. 5.1 der Vorlesung eingeführt. Die Grösse  $\tau_{pt}$  bezeichnet die Lebensdauer der Löcher. Die totale zeitliche Änderung der Löcherzahl ergibt sich aus der Summe der Prozesse 1) und 2):

$$\frac{\partial p}{\partial t} = -\frac{\partial F_{px}^+}{\partial x} + g_p - \frac{\delta p}{\tau_{pt}}$$

Die Stromdichten unterscheiden sich von den Flüssen durch die jeweilige Ladung:

$$j_{px} = eF_{px}^+ \qquad j_{nx} = -eF_{nx}^+$$

$$\frac{\partial p}{\partial t} = -\frac{1}{e} \frac{\partial j_{px}}{\partial x} + g_p - \frac{\delta p}{\tau_{pt}} \qquad \frac{\partial n}{\partial t} = +\frac{1}{e} \frac{\partial j_{nx}}{\partial x} + g_n - \frac{\delta n}{\tau_{nt}}$$

Die zeitliche Änderung einer Ladung in einem Volumenelement ist gleich der räumlichen Zu- oder Abnahme einer Stromdichte *plus* der im Element (netto) generierten Ladung.



# Die Diffusionsgleichung

Die Löcher-Stromdichte (Drift + Diffusion)  $j_{px} = ep\mu_p E_x - eD_p \frac{\partial p}{\partial x}$  wird in die Kontinuitätsgleichung eingesetzt

$$\frac{\partial p}{\partial t} = -\frac{1}{e} \frac{\partial j_{px}}{\partial x} + g_p - \frac{\delta p}{\tau_{pt}}$$

Man erhält

$$\frac{\partial p}{\partial t} = -\mu_p \frac{\partial}{\partial x} (p E_x) + D_p \frac{\partial^2 p}{\partial x^2} + g_p - \frac{\delta p}{\tau_{pt}} \quad \text{zeitabhängige Drift-Diffusionsgleichung für Löcher}$$

Annahmen:

- Stationärer Fall.
- Elektrisches Feld in den Bahngebieten wird wegen seiner Kleinheit vernachlässigt.
- Überschuss-Löcher *rekombinieren* (keine Generation).

$$\cancel{\frac{\partial p}{\partial t}} = -\cancel{\mu_p} \cancel{\frac{\partial}{\partial x}} (\cancel{p} E_x) + D_p \frac{\partial^2 p}{\partial x^2} + \cancel{g_p} - \cancel{\frac{\delta p}{\tau_{pt}}}$$

Dann verbleibt für die Überschuss-Löcherdichte im n-Bahnggebiet die einfache Differentialgleichung

$$\frac{\partial^2 \delta p}{\partial x^2} = \frac{\delta p}{D_p \tau_{p0}} = \frac{\delta p}{L_p^2} \quad x > x_n$$

$L_p = (D_p \tau_{p0})^{1/2}$  ist die *Diffusionslänge* der Löcher. (Beachte  $\tau_{pt} \rightarrow \tau_{p0}$  wegen Stationarität.) Eine entsprechende Gleichung folgt für die Überschuss-Elektronendichte im p-Bahnggebiet

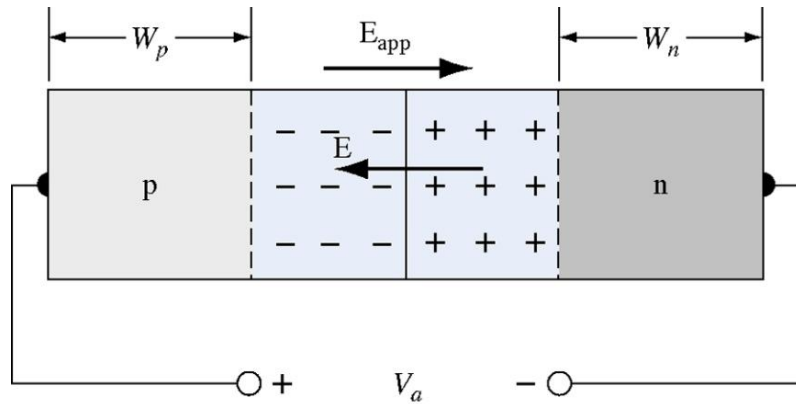
$$\frac{\partial^2 \delta n}{\partial x^2} = \frac{\delta n}{D_n \tau_{n0}} = \frac{\delta n}{L_n^2} \quad x < -x_p$$

Die Überschussdichten klingen in den Bahnggebieten exponentiell auf der Längenskala der Diffusionslänge ab!

$$\delta p_n(x) = p_n(x) - p_{n0} = A \exp\left(\frac{x}{L_p}\right) + B \exp\left(-\frac{x}{L_p}\right) \quad x > x_n$$

$$\delta n_p(x) = n_p(x) - n_{p0} = C \exp\left(\frac{x}{L_n}\right) + D \exp\left(-\frac{x}{L_n}\right) \quad x < -x_p$$

# Die "lange" pn-Diode



Die Länge der Bahngebiete  $W_n$  und  $W_p$  sei gross gegen die Diffusionslängen! Dann sind die Überschusssichten bereits vor den Kontakten auf Null abgeklungen:

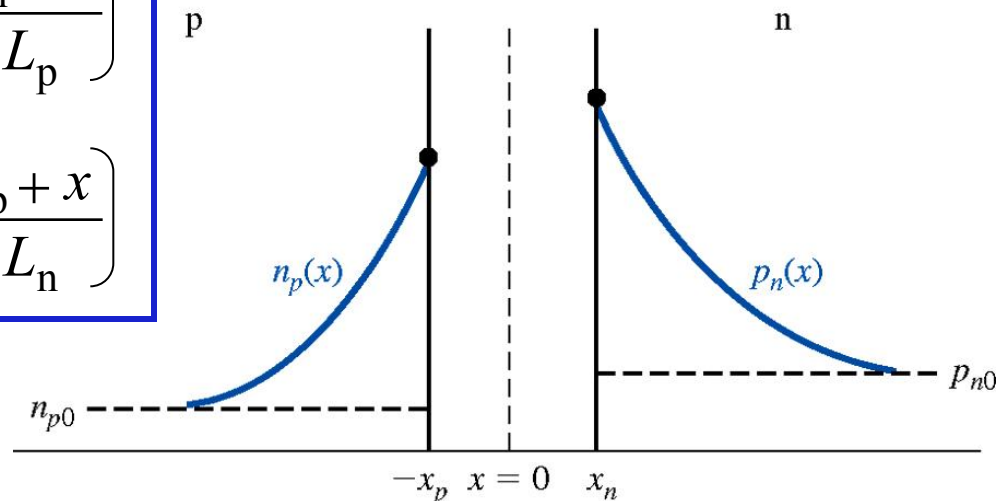
$$p_n(x \rightarrow \infty) = p_{n0} \quad \text{und} \quad n_p(x \rightarrow -\infty) = n_{p0}$$

Die Randwerte der Minoritäts-Ladungsträgerdichten an den Rändern der RLZ sind uns bereits bekannt:

$$\begin{cases} p_n(x_n) = p_{n0} \exp\left(\frac{eV_a}{kT}\right) \\ n_p(-x_p) = n_{p0} \exp\left(\frac{eV_a}{kT}\right) \end{cases}$$

$$\delta p_n(x) = p_{n0} \left\{ \exp\left(\frac{eV_a}{kT}\right) - 1 \right\} \exp\left(\frac{x_n - x}{L_p}\right)$$

$$\delta n_p(x) = n_{p0} \left\{ \exp\left(\frac{eV_a}{kT}\right) - 1 \right\} \exp\left(\frac{x_p + x}{L_n}\right)$$



# Die ideale Strom-Spannungs-Charakteristik der "langen" Diode

Die Ortsabhängigkeit der Überschuss-Minoritätsladungsträgerdichten in den Bahngebieten kann benutzt werden, um die Diffusionsströme zu berechnen. Man beachte, dass Driftströme wegen des vernachlässigten  $E$ -Feldes nicht auftreten.

$$j_p(x_n) = -eD_p \left. \frac{\partial p_n}{\partial x} \right|_{x=x_n} \quad j_n(-x_p) = eD_n \left. \frac{\partial n_p}{\partial x} \right|_{x=-x_p}$$

Wir nehmen an, dass die Dotierung in den Bahngebieten konstant ist. Dann ist die Ableitung der Minoritätsladungsträgerdichten gleich der Ableitung der Überschuss-Minoritätsladungsträgerdichten. Ausführung der Ableitungen ergibt

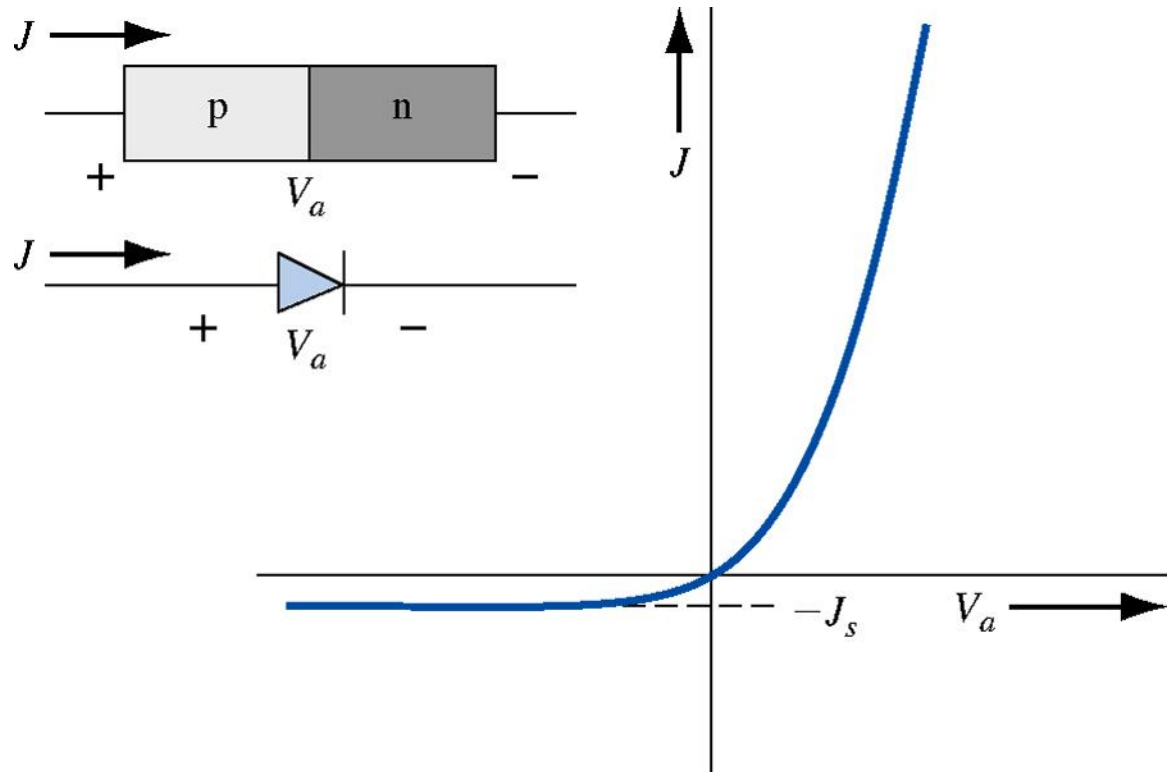
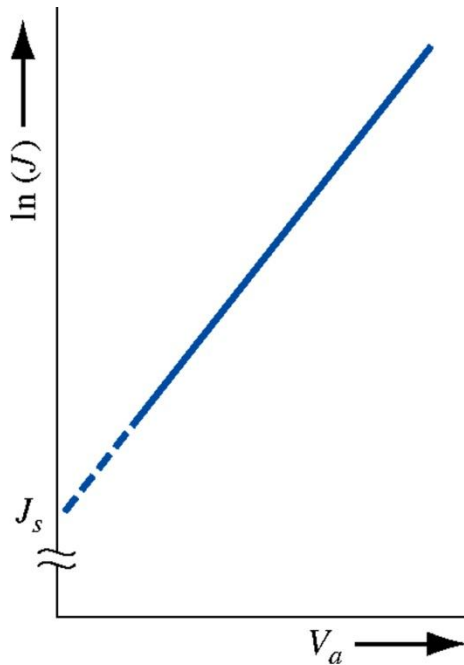
$$j_p(x_n) = \frac{eD_p p_{n0}}{L_p} \left\{ \exp\left(\frac{eV_a}{kT}\right) - 1 \right\} \quad j_n(-x_p) = \frac{eD_n n_{p0}}{L_n} \left\{ \exp\left(\frac{eV_a}{kT}\right) - 1 \right\}$$

Nach Voraussetzung sind die individuellen Ströme in der gesamten RLZ konstant. Die Summe ergibt daher den totalen (Gesamt-)strom in der RLZ. Da der totale Strom in der ganzen Diode konstant ist, liefert die Summe den totalen Diodenstrom. (Die Konstanz des Gesamtstroms folgt aus der Summe der Kontinuitätsgleichungen und der Annahme der Stationarität.)

$$j = j_p(x_n) + j_n(-x_p) = \left( \frac{eD_p p_{n0}}{L_p} + \frac{eD_n n_{p0}}{L_n} \right) \left\{ \exp\left( \frac{eV_a}{kT} \right) - 1 \right\}$$

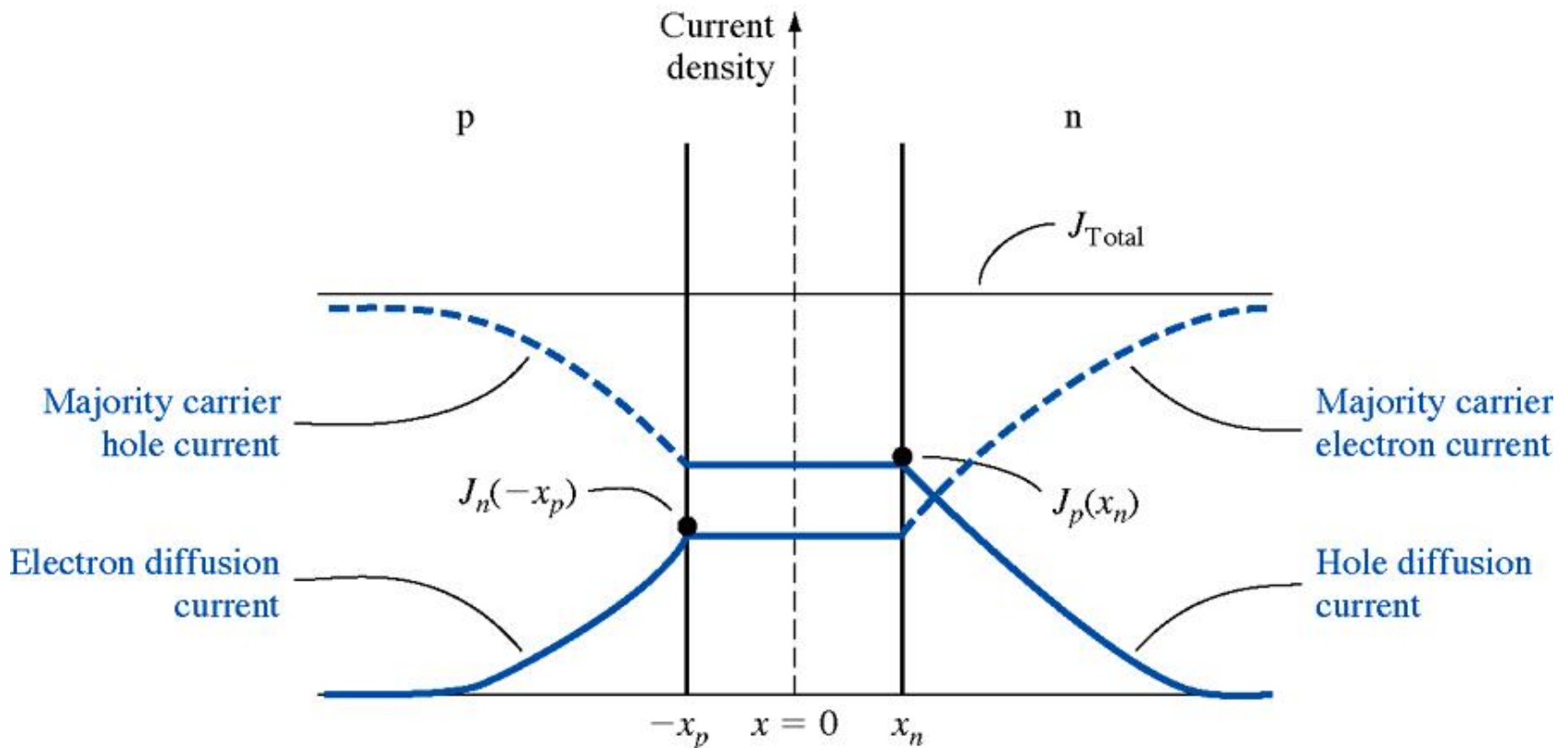
$$j = j_S \left\{ \exp\left( \frac{eV_a}{kT} \right) - 1 \right\}$$

$j_S$  ist der Sättigungs-Sperrstrom



$$j_p(x) = \frac{eD_p p_{n0}}{L_p} \left\{ \exp\left(\frac{eV_a}{kT}\right) - 1 \right\} \exp\left(\frac{x_n - x}{L_p}\right)$$

$$j_n(x) = \frac{eD_n n_{p0}}{L_n} \left\{ \exp\left(\frac{eV_a}{kT}\right) - 1 \right\} \exp\left(\frac{x_p + x}{L_n}\right)$$



# Die "dünne" pn-Diode

Die Länge des Bahngebietes  $W_n$  sei klein gegen die Diffusionslänge der Löcher  $L_p$ ! Dadurch kann die vormalige Randbedingung für  $x \rightarrow \infty$  nicht mehr angewandt werden.

Man nimmt an, dass am Metallkontakt bei  $x = x_n + W_n$  die Löcher unendlich schnell rekombinieren (Ohmscher Kontakt), so dass dort die Dichte instantan auf den Gleichgewichtswert  $p_{n0}$  zurückgeht:

$$p_n(x_n + W_n) = p_{n0}$$

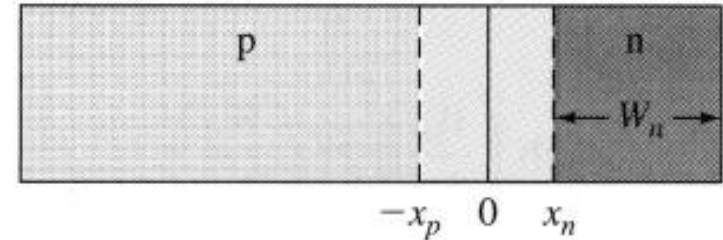
Die zweite Randbedingung bleibt wie gehabt:  $p_n(x_n) = p_{n0} \exp\left(\frac{eV_a}{kT}\right)$

Die Konstanten  $A$  und  $B$  der allgemeinen Lösung

$$\delta p_n(x) = p_n(x) - p_{n0} = A \exp\left(\frac{x}{L_p}\right) + B \exp\left(-\frac{x}{L_p}\right)$$

können nun eindeutig bestimmt werden (der exp. ansteigende Teil wird wichtig!):

$$\delta p_n(x) = p_{n0} \left\{ \exp\left(\frac{eV_a}{kT}\right) - 1 \right\} \frac{\sinh[(x_n + W_n - x)/L_p]}{\sinh[W_n/L_p]}$$



Geometrie der dünnen pn-Diode.

Falls  $W_n \ll L_p$ , kann man den  $\sinh(*)$  linear entwickeln und es bleibt:

$$\delta p_n(x) = p_{n0} \left\{ \exp\left(\frac{eV_a}{kT}\right) - 1 \right\} \frac{(x_n + W_n - x)}{W_n}$$

Die Löcherdichte fällt linear von  $x_n$  nach  $x_n + W_n$  ab. Der Diffusionsstrom

$$j_p(x) = - e D_p \frac{\partial \delta p_n}{\partial x}$$

in der “kurzen” Diode wird daher

$$j_p(x) = \frac{e D_p p_{n0}}{W_n} \left\{ \exp\left(\frac{eV_a}{kT}\right) - 1 \right\}$$

**Diskussion:** 1.) Die Löcherstromdichte in der “kurzen” Diode wird konstant. Das Bahngebiet ist so kurz, dass keine Rekombination stattfindet. 2.) Wegen der Voraussetzung  $W_n \ll L_p$  ist der Diffusionsstrom der “kurzen” Diode grösser als der der “langen” Diode.



# 7.2 Kleinsignal-Verhalten der pn-Diode

## Der Diffusionswiderstand

Der ideale Diodenstrom ist  $I_D = I_S \left\{ \exp\left(\frac{eV_a}{kT}\right) - 1 \right\}$

Die Diode sei durch die Spannung  $V_0$  in Flussrichtung vorgespannt. Dadurch fließt ein dc-Strom  $I_{DQ}$ .

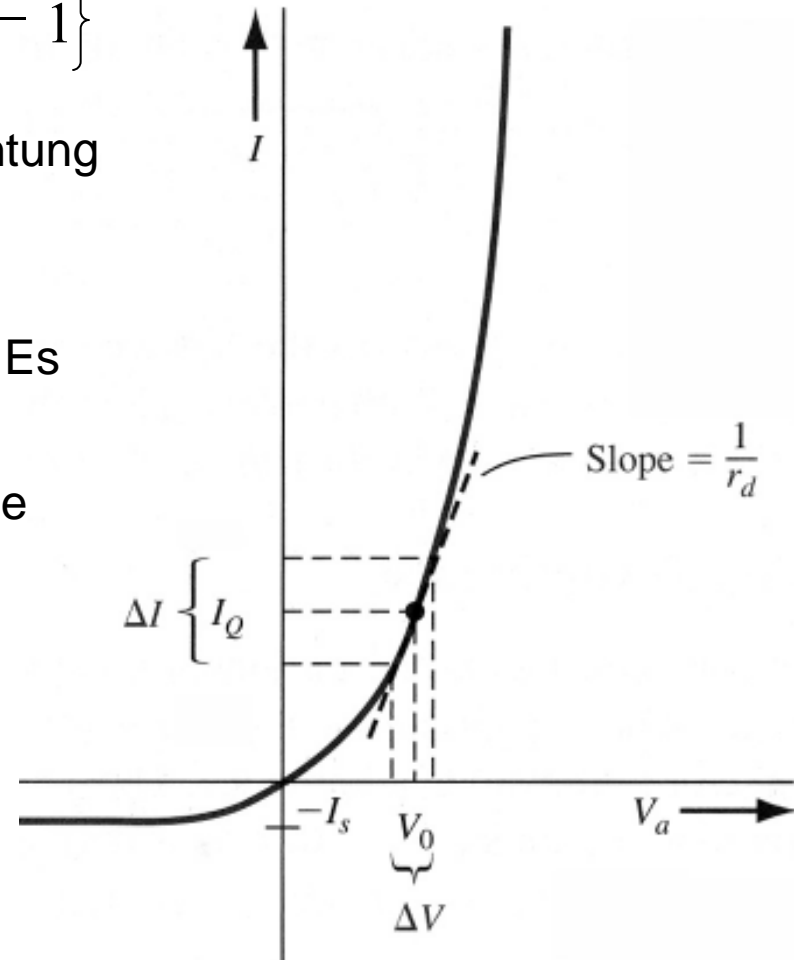
Ein kleines, niedrig-frequentes, sinusförmiges Wechsignal wird der dc-Spannung überlagert. Es erzeugt einen kleinen ac-Strom.

Der Wechselstrom-Leitwert ist die Tangente an die Kennlinie:

$$g_d = \left. \frac{dI_D}{dV_a} \right|_{V_a = V_0}$$

Der Diffusionswiderstand ist der Kehrwert davon

$$r_d = \left. \frac{dV_a}{dI_D} \right|_{I_D = I_{DQ}}$$



Die Gleichspannung sei gross gegen die Thermospannung:  $V_0 \gg kT/e$ . Dann kann man die "1" in der Formel für den Diodenstrom vernachlässigen und der Wechselstrom-Leitwert wird

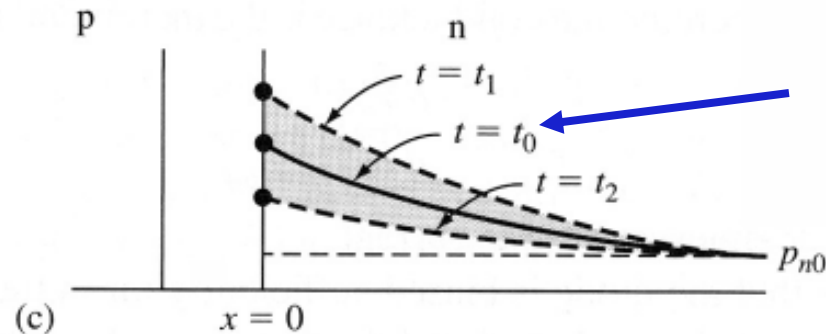
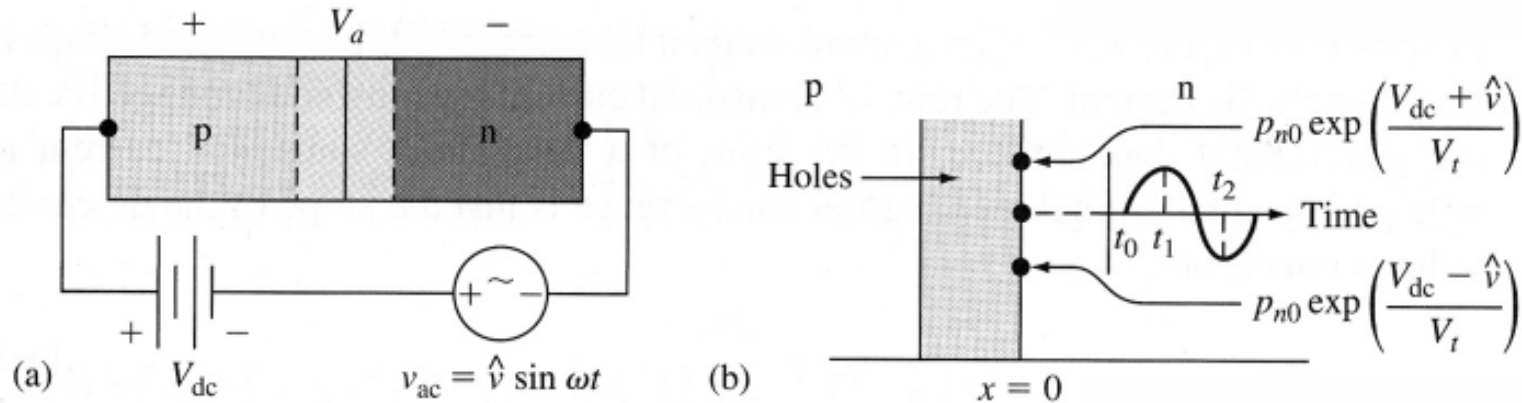
$$g_d = \left. \frac{dI_D}{dV_a} \right|_{V_a = V_0} = \left( \frac{e}{kT} \right) I_S \exp\left( \frac{eV_0}{kT} \right) \approx \frac{I_{DQ}}{V_t}$$

Der Diffusionswiderstand ist durch das Verhältnis aus Thermospannung und dc-Strom gegeben:

$$r_d = \frac{V_t}{I_{DQ}}$$

# Die Kleinsignal-Admittanz

Die Frequenz des Kleinsignals sei klein gegen die inversen Lebensdauern der Minoritäten. Das führt zum steady-state, d.h. die Minoritätsdichten können der Wechselfspannung ohne Verzögerung folgen.



(a) Vorwärts gepolter pn-Übergang ( $V_{dc}$ ) mit Wechselfspannung  $v_{ac}$ , (b) zeitliches Verhalten der Löcherkonzentration am Rand der Raumladungszone, (c) räumlicher Verlauf der Löcherkonzentration in der n-Region zu 3 verschiedenen Zeiten.

Der Phasor  $\hat{V}_1$  der Wechselspannung  $v_1(t) = \hat{V}_1 \exp(j\omega t)$  sei klein gegen die Thermospannung.

Die dc-Diffusionsströme folgen nach Voraussetzung dem Wechselsignal,

$$I_{p,n} = \hat{I}_{p,n} \exp(j\omega t)$$

Die Admittanz ist definiert als 
$$Y = \frac{\hat{I}_p + \hat{I}_n}{\hat{V}_1} = \frac{1}{V_t} \left[ I_{p0} (1 + j\omega\tau_{p0})^{1/2} + I_{n0} (1 + j\omega\tau_{n0})^{1/2} \right]$$

Letzteres wird im *Neamen* auf S. 289 – 293 vorgerechnet.

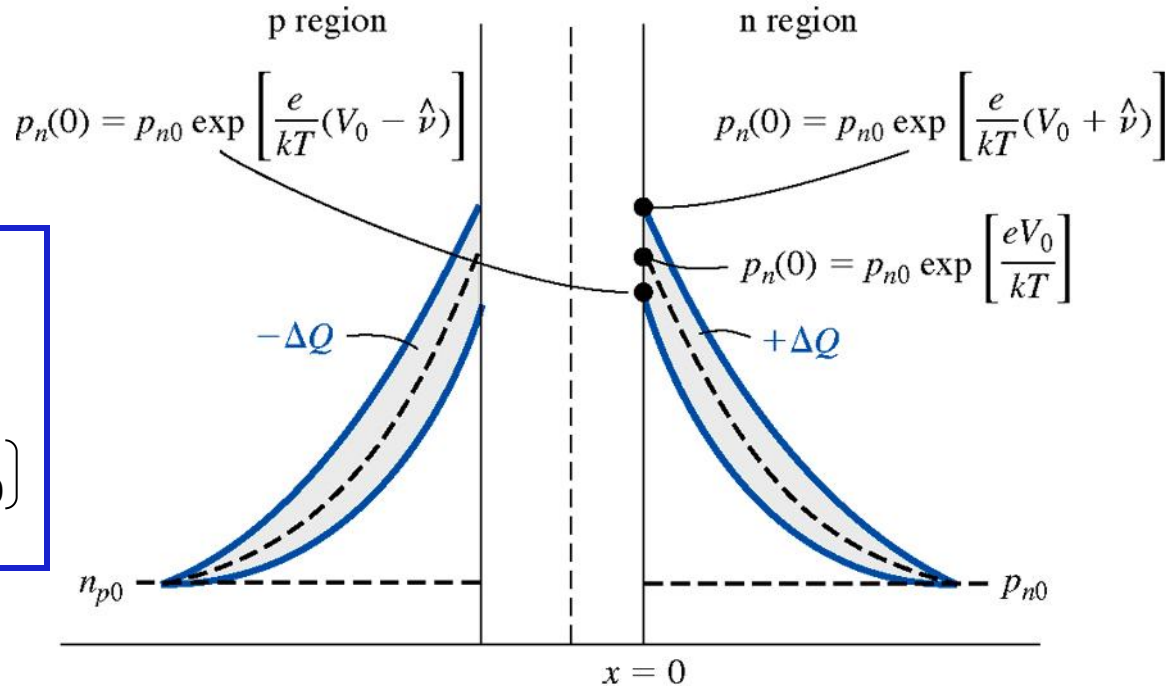
Andererseits

$$Y = g_d + j\omega C_d$$

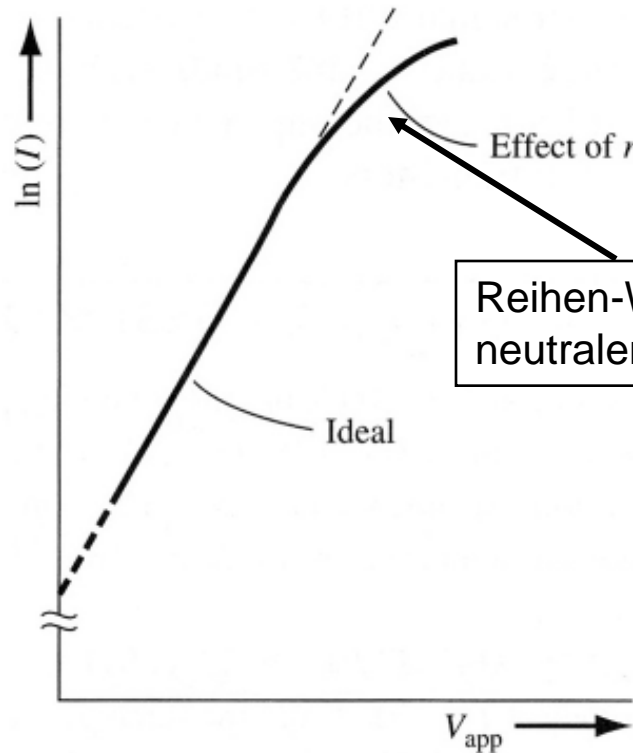
Man erhält mit  $\omega_{p,n0} \ll 1$

$$g_d = \frac{I_{DQ}}{V_t}$$

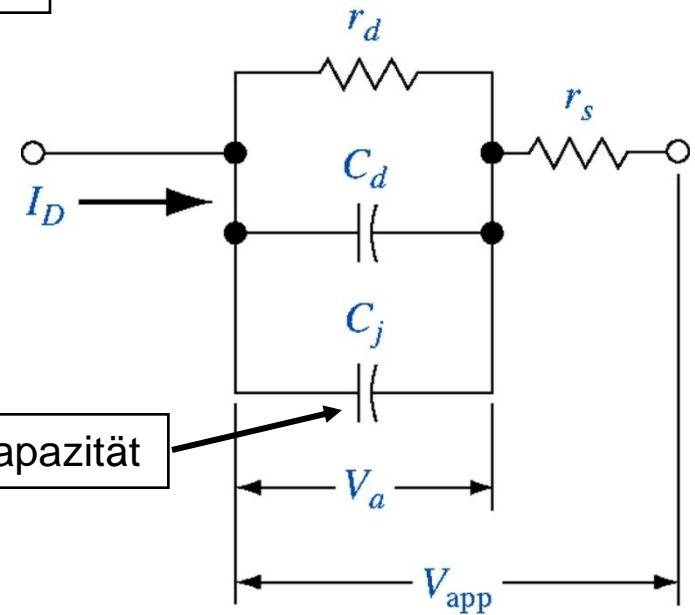
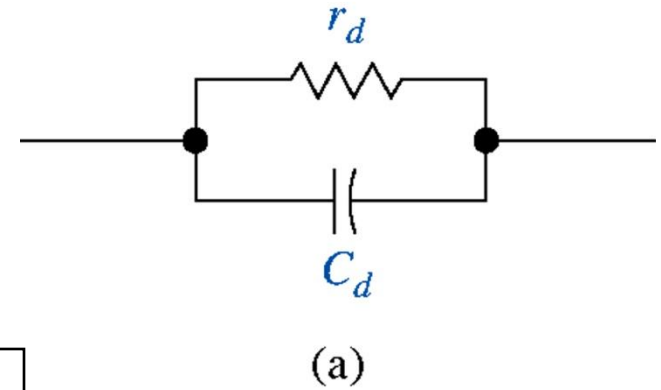
$$C_d = \frac{1}{2V_t} (I_{p0} \tau_{p0} + I_{n0} \tau_{n0})$$



# Ersatzschaltbilder

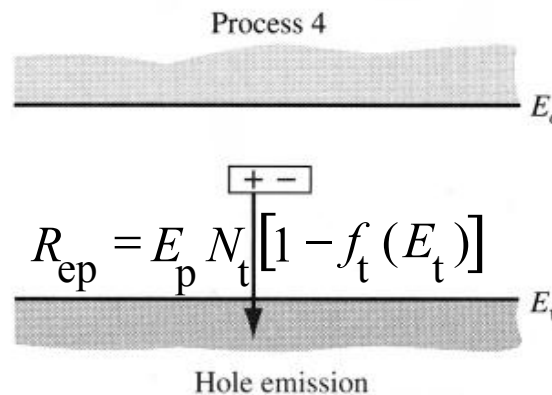
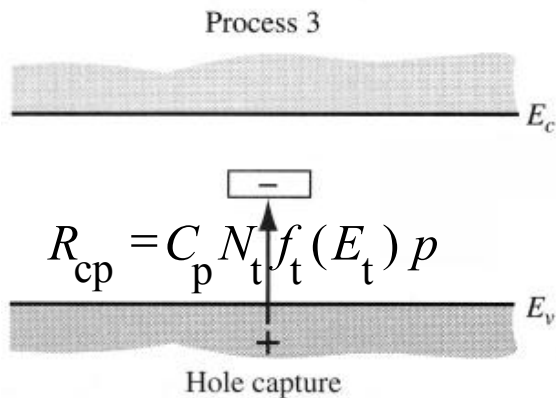
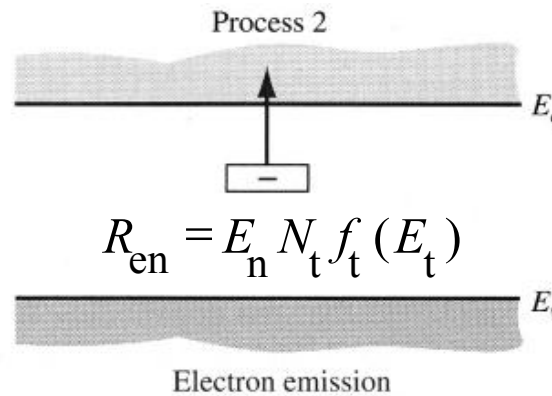
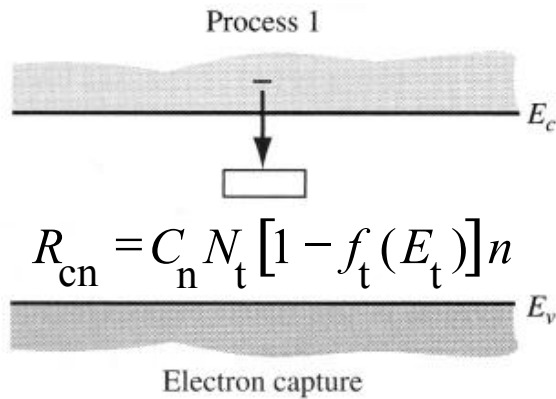


Die  $I/V$ -Kurve des vorwärts gepolten pn-Übergangs knickt wegen des Reihenwiderstands ab.



# 7.3 Generations-Rekombinationsströme

## Shockley-Read-Hall-Rekombination



1. Elektroneneinfang
2. Elektronenemission
3. Locheinfang
4. Lochemission

Dichte der Rekombinationszentren ("traps")

$$N_t \text{ [cm}^{-3}\text{]}$$

Einfangkoeffizient

$$C_{n,p} \text{ [cm}^3\text{/s]}$$

Emissionsrate

$$E_{n,p} \text{ [1/s]}$$

Die 4 grundlegenden Einfang- und Emissionsprozesse an einer tiefen Störstelle.

Die wesentlichen Schritte der Ableitung der SRH-Formel (sh. *Neamen* S. 219-222):

1.) Besetzungswahrscheinlichkeit der tiefen Störstelle *im Gleichgewicht*:

$$f_t^{(0)}(E_t) = \frac{1}{1 + \frac{1}{g} \exp \frac{E_t - E_F}{kT}}$$

2.) Im Gleichgewicht muss  $R_{en} = R_{cn}$  und  $R_{ep} = R_{cp}$  gelten. Damit kann man die Emissionsraten  $E_{n,p}$  durch die Einfangkoeffizienten  $C_{n,p}$  ausdrücken.

3.) Im stationären Zustand muss die Netto-Rekombinationsrate für Elektronen  $R_n = R_{cn} - R_{en}$  gleich der für Löcher  $R_p = R_{cp} - R_{ep}$  sein. Daraus kann man die Besetzungswahrscheinlichkeit der tiefen Störstelle im Nichtgleichgewicht  $f_t(E_t)$  gewinnen.

4.) Setzt man diese wieder in  $R_n$  ein, folgt die SRH-Formel

$$R_n = R_p = R = \frac{C_n C_p N_t (np - n_i^2)}{C_n (n + n') + C_p (p + p')} = \frac{(np - n_i^2)}{\tau_{p0} (n + n') + \tau_{n0} (p + p')}$$

Die Minoritätsladungsträger-Lebensdauern sind definiert als

$$\tau_{p0} = (C_p N_t)^{-1} \quad \tau_{n0} = (C_n N_t)^{-1}$$

# Generationsstrom in der gesperrten pn-Diode

Bedeutung der Dichten  $n'$  und  $p'$ :

Offenbar ist  $n'p' = n_0p_0 = n_i^2$

Liegt das tiefe Niveau  $E_t$  ("trap") in der Mitte der Energielücke, haben  $n'$  und  $p'$  die Größenordnung von  $n_i$

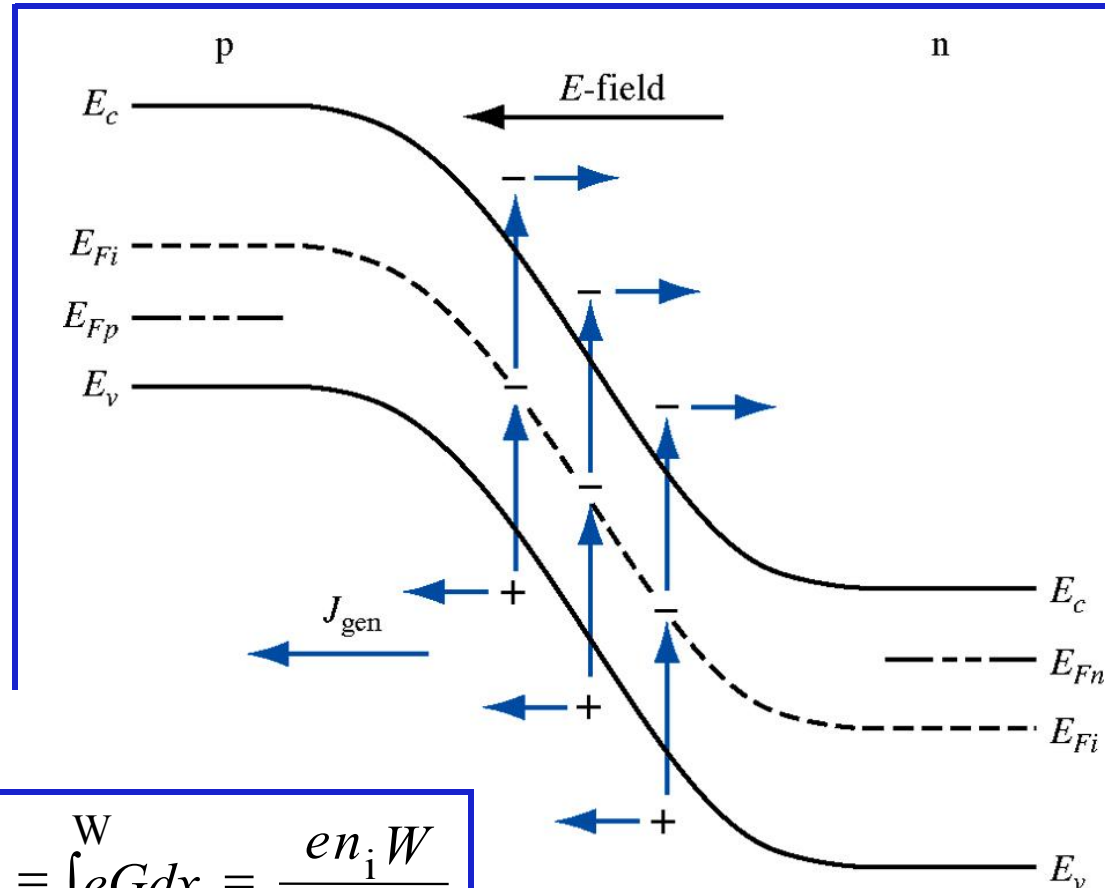
Im gesperrten pn-Übergang ist  $np \ll n_i^2$  sowie  $p \ll n_i$  und  $n \ll n_i$ . Daher reduziert sich die SRH-Formel auf

$$R = \frac{-n_i^2}{\tau_{p0}n' + \tau_{n0}p'} \approx \frac{-n_i}{\tau_{p0} + \tau_{n0}}$$

$R = -G$  ist eine Generationsrate. Der Generationsstrom wird

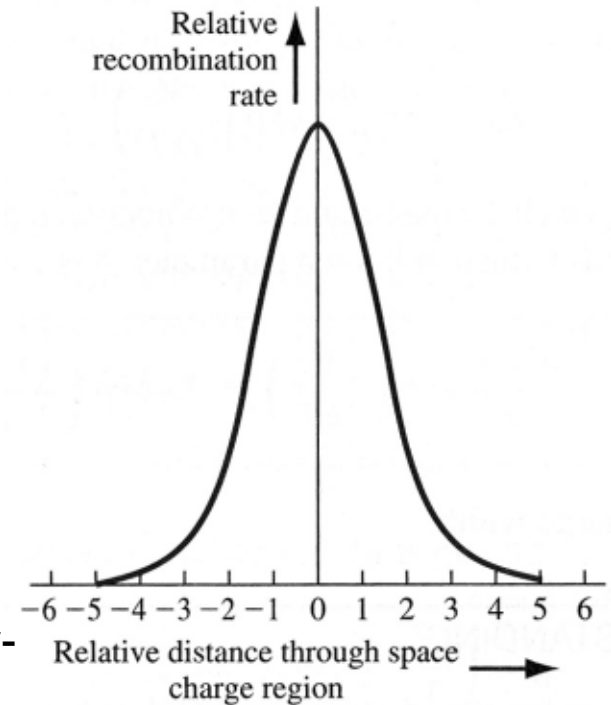
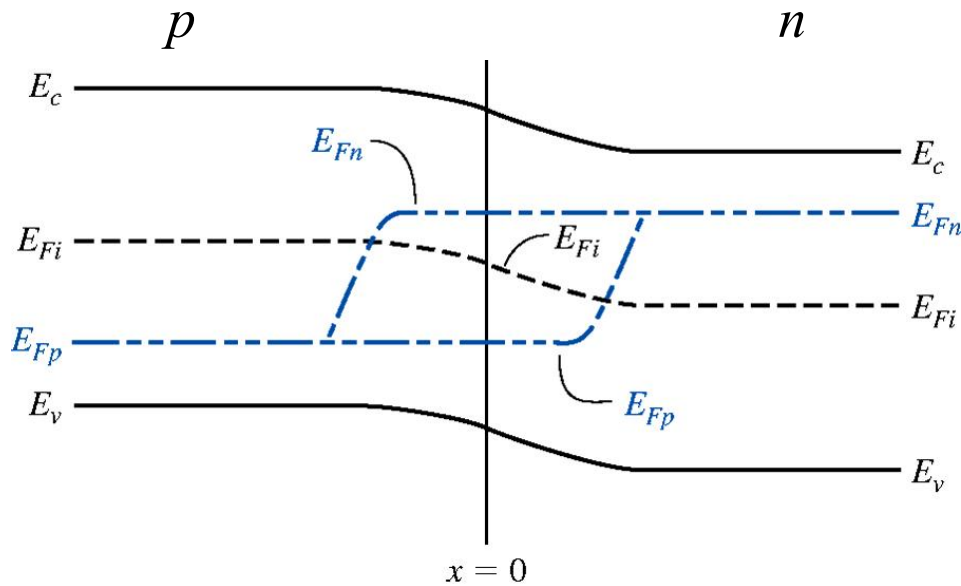
$$j_{\text{gen}} = \int_0^W eG dx = \frac{en_i W}{\tau_{p0} + \tau_{n0}}$$

$$n' = n_0 \frac{1 - f_t^{(0)}}{f_t^{(0)}} \quad p' = p_0 \frac{f_t^{(0)}}{1 - f_t^{(0)}}$$





# Rekombinationsstrom in der vorwärtsgepolten pn-Diode



Im vorwärtsgepolten pn-Übergang ist  $np \gg n_i^2$  sowie  $p \gg n_i$  und  $n \gg n_i$ . Die SRH-Rate hat ihr Maximum am metallurgischen Übergang (bei  $x = 0$ ). Dort kann man annehmen, dass  $n = p$  gilt (zumindest für eine symmetrische Diode). Die max. SRH-Rate wird also

$$R_{\max} = \frac{np}{\tau_{p0}n + \tau_{n0}p} \approx \frac{n}{\tau_{p0} + \tau_{n0}}$$

Die Dichte  $n$  ist bei  $x = 0$ :

$$n(0) = n_i \exp\left(\frac{E_{Fn} - E_{Fi}}{kT}\right) = n_i \exp\left(\frac{eV_a}{2kT}\right)$$

Relative Stärke der Rekombinationsrate in der Raumladungszone eines gesperrten pn-Übergangs.

Das Integral über die SRH-Rate wird ersetzt durch das Produkt aus  $R_{\max}$  und einer Länge  $x'$  ( $x'$  ist ein Teil der RLZ-Weite).

Dann folgt für den Rekombinationsstrom

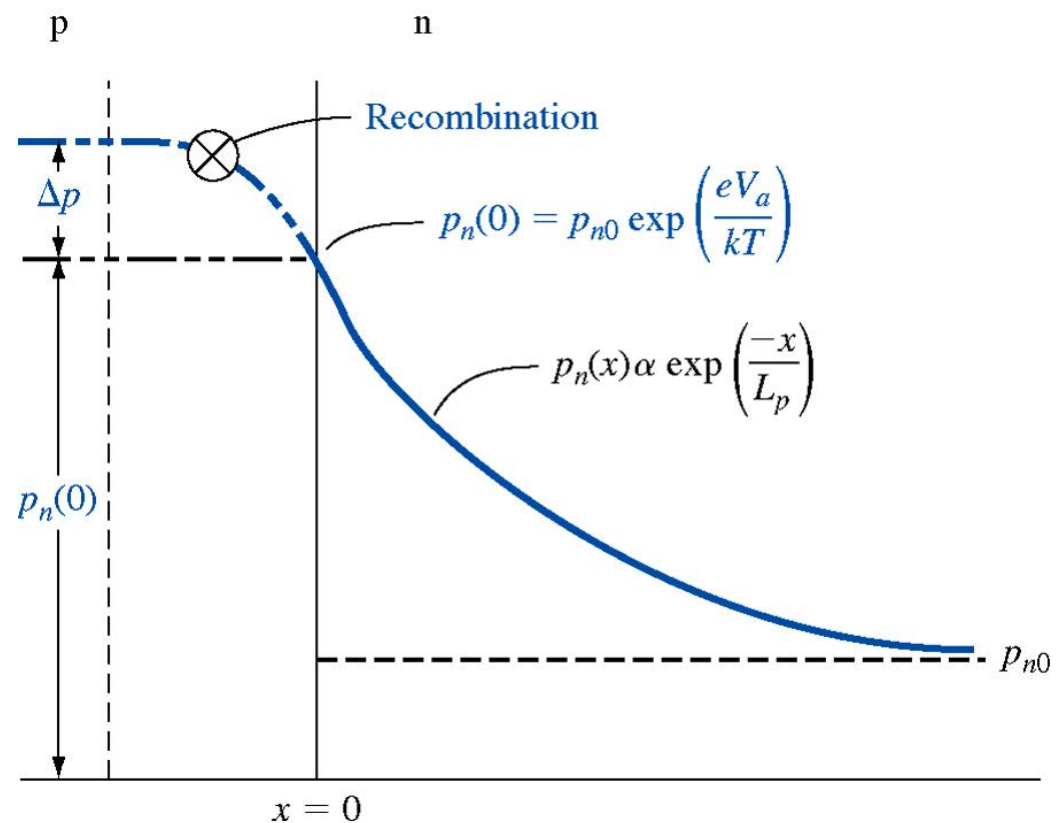
$$j_{\text{rec}} = \int_0^W eR dx$$

$$= \frac{en_i x'}{\tau_{p0} + \tau_{n0}} \exp\left(\frac{eV_a}{2kT}\right)$$

Allgemein schreibt man

$$j_{\text{rec}} = j_{r0} \exp\left(\frac{eV_a}{2kT}\right)$$

Löcher, die in der RLZ durch Rekombination verloren gehen, müssen zusätzlich vom p-Gebiet geliefert werden.

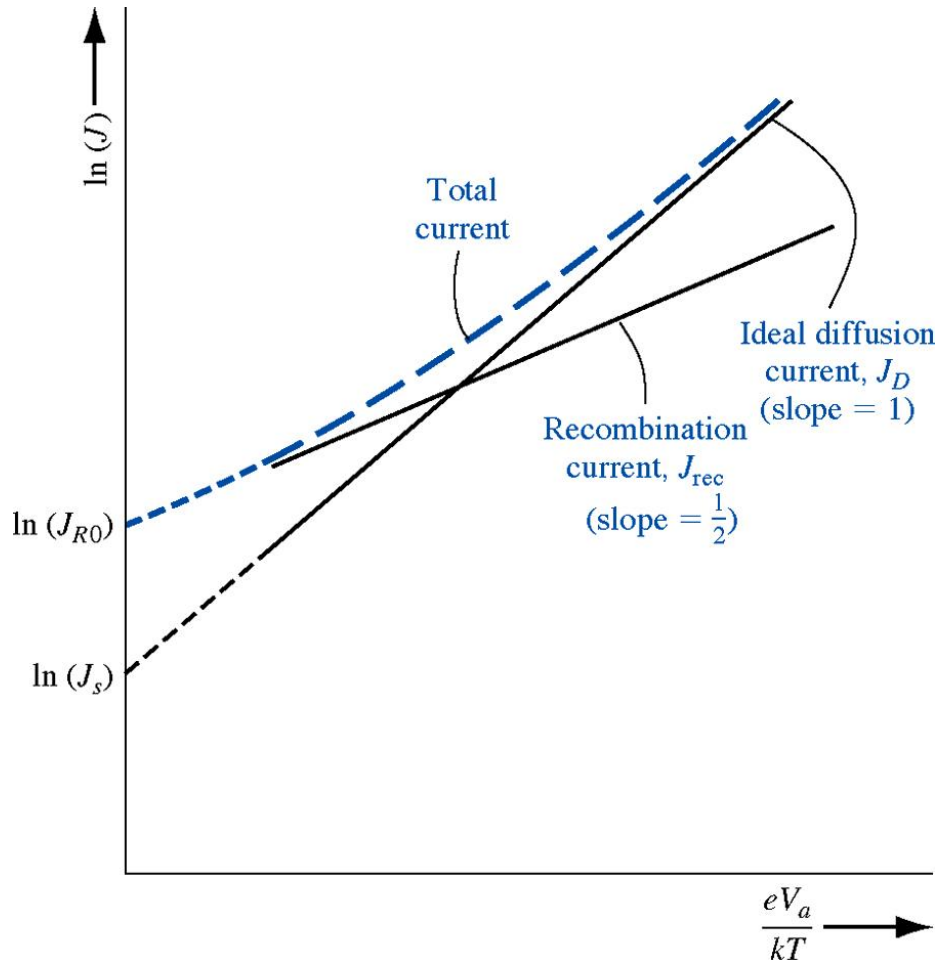


Die totale Stromdichte in der vorwärtsgepolten pn-Diode ist die Summe aus Rekombinations- und idealer Diffusionsstromdichte:

$$j = j_{\text{rec}} + j_D = j_{r0} \exp\left(\frac{eV_a}{2kT}\right) + j_S \exp\left(\frac{eV_a}{kT}\right)$$

# Gesamtstrom in der vorwärtsgepolten pn-Diode

$$j = j_{\text{rec}} + j_{\text{D}} = j_{\text{r0}} \exp\left(\frac{eV_{\text{a}}}{2kT}\right) + j_{\text{S}} \exp\left(\frac{eV_{\text{a}}}{kT}\right)$$

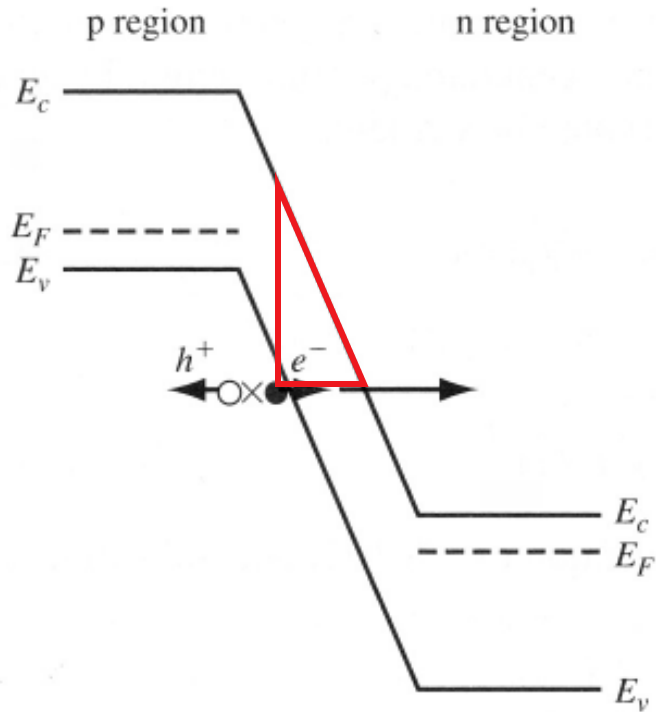


In der Praxis wird die Summe aus Rekombinations- und idealer Diffusionsstromdichte zu *einer* Diodenstrom-Formel zusammengefasst:

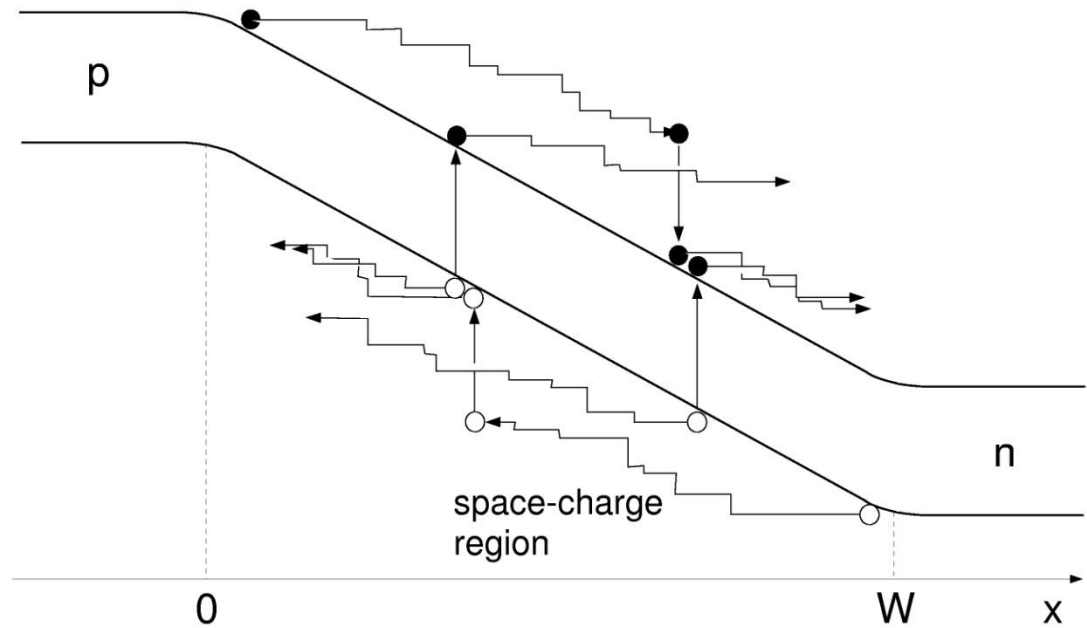
$$j = j_{\text{S}} \exp\left(\frac{eV_{\text{a}}}{nkT}\right)$$

Der Faktor  $n$  heisst *Idealitätsfaktor*. Er variiert im Bereich moderater Vorwärtsspannungen zwischen 2 und 1.

# 7.4 Durchbrüche von Dioden



(a)



(b)

(a) "Zener"-Durchbruch (Interband-Tunneln) im gesperrten pn-Übergang. (b) Avalanche-Durchbruch im gesperrten pn-Übergang.

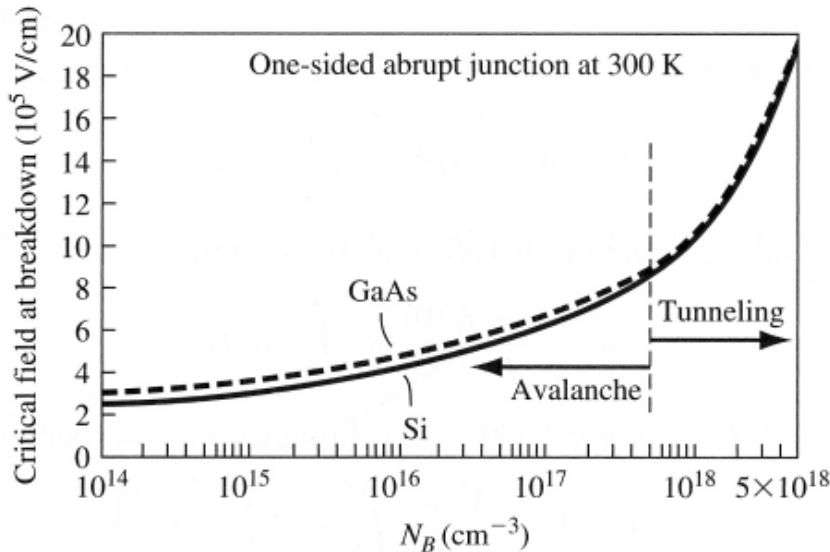
# Der Lawinendurchbruch

Stossionisationsrate:  $G_{II} = \alpha_n j_n(x) + \alpha_p j_p(x)$

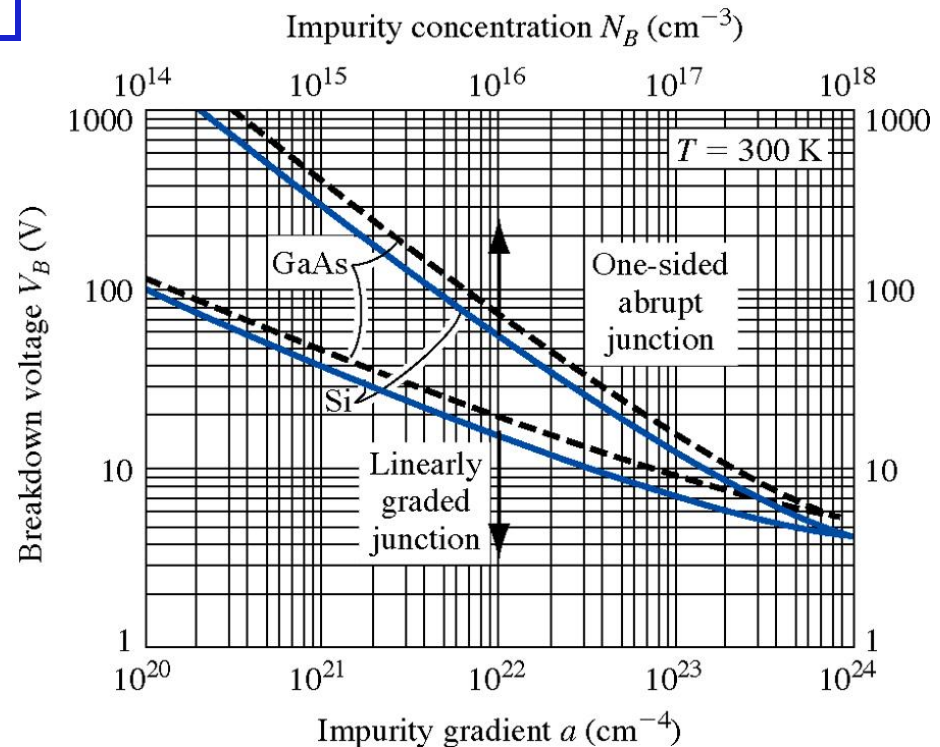
Die  $\alpha_n, \alpha_p$  heissen Stossionisationskoeffizienten [1/cm] und sind anschaulich der reziproke mittlere Weg zwischen zwei Stößen, die zur Ionisation führen. Sie hängen exponentiell vom elektrischen Feld ab:  $\alpha_{n,p} \sim \exp(-E_{crit}/E)$ .

Mit  $\alpha_n = \alpha_p = \alpha = \text{const}$  lautet die Bedingung für Lawinendurchbruch

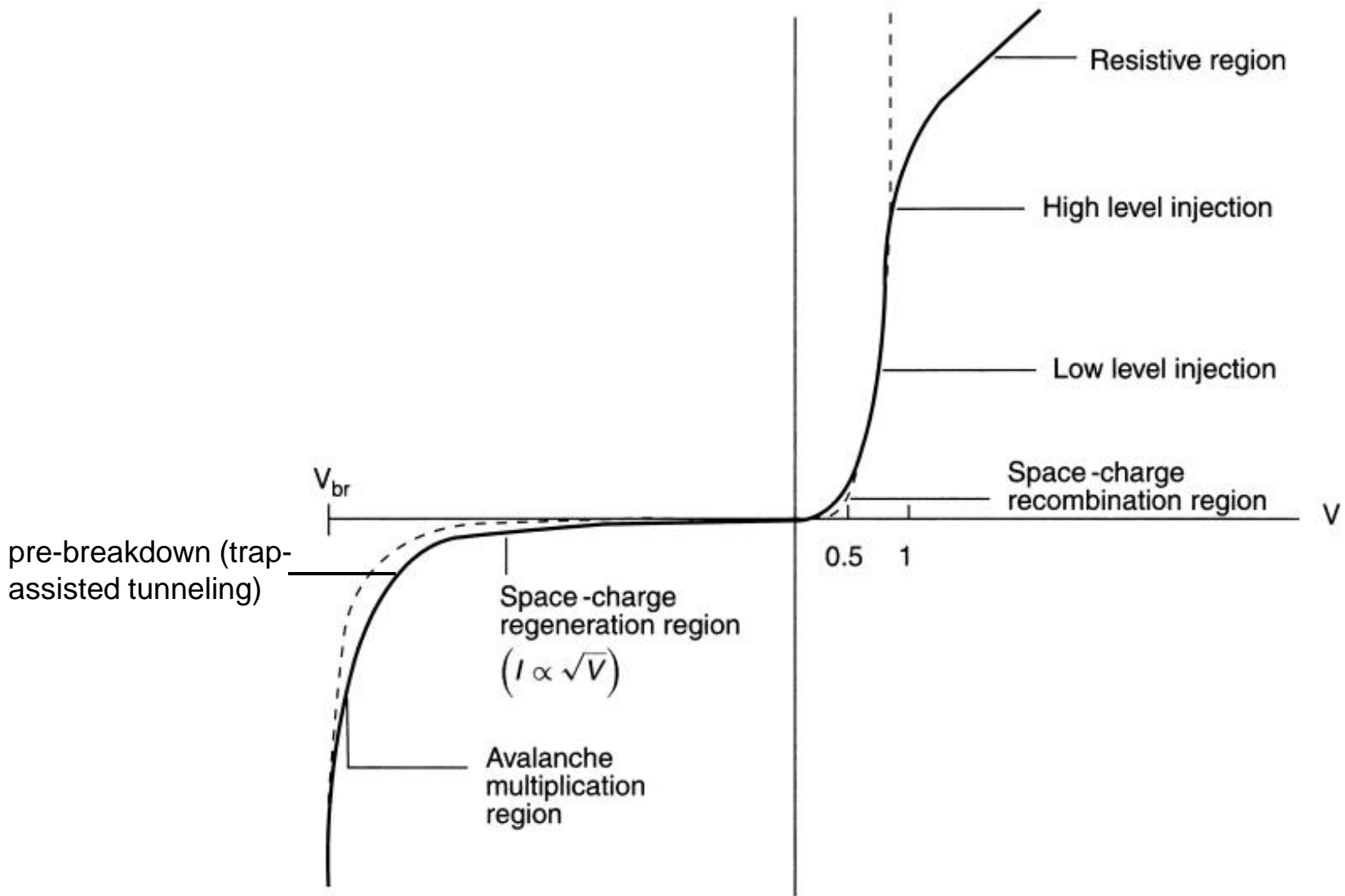
$$\alpha W = 1$$



Kritische Feldstärke beim Durchbruch in einem einseitigen pn-Übergang als Funktion der Dotierungskonzentration (nach Sze).



# Die reale Diodenkennlinie











# 8. Metall-Halbleiter- und Halbleiter-Halbleiter-Übergänge

**8.1 Die Schottky-Diode (N9.1)**

**8.2 Der Ohmsche Kontakt (N9.2)**

**8.3 Hetero-Übergänge (N9.3)**

# 8.1 Die Schottky-Diode

## Bandkanten-Diagramme

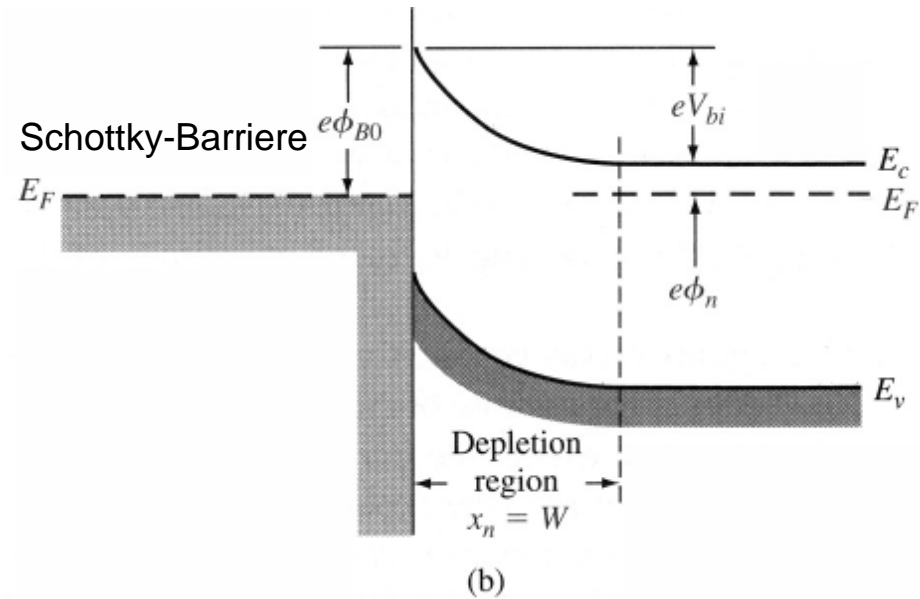
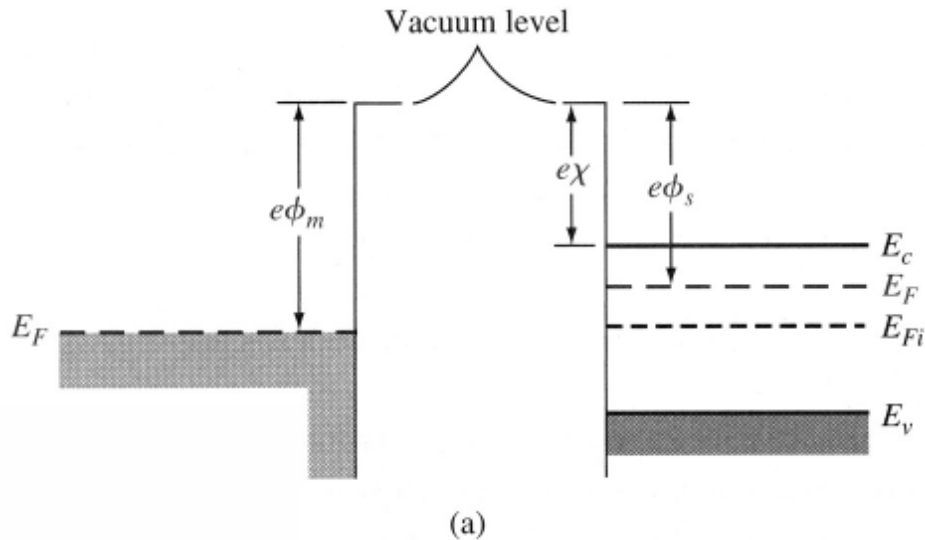


Table 9.1 | Work functions of some elements

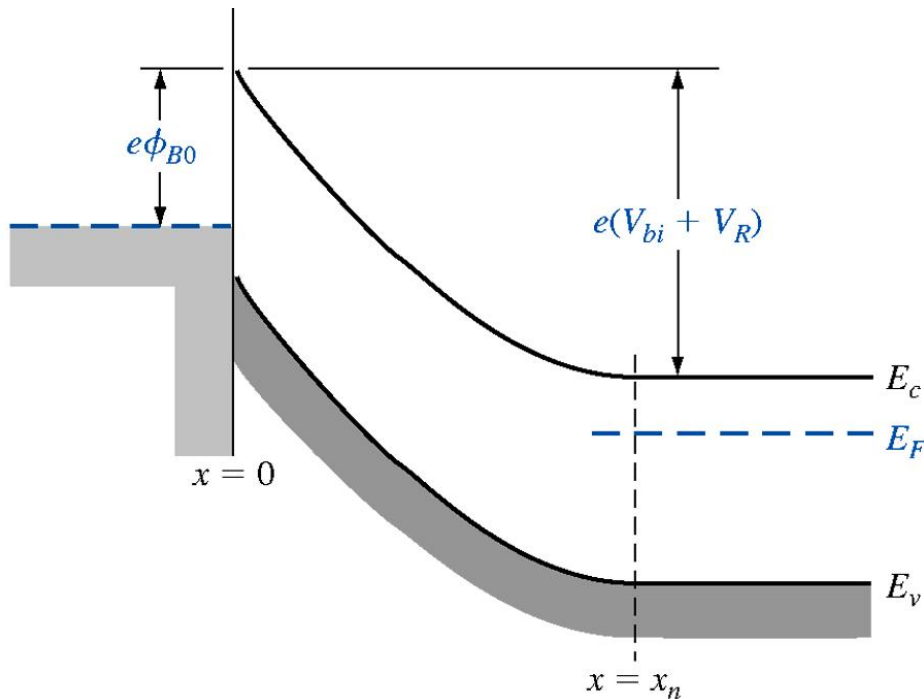
Element	Work function, $\phi_m$
Ag, silver	4.26
Al, aluminum	4.28
Au, gold	5.1
Cr, chromium	4.5
Mo, molybdenum	4.6
Ni, nickel	5.15
Pd, palladium	5.12
Pt, platinum	5.65
Ti, titanium	4.33
W, tungsten	4.55

$$\Phi_{B0} = \Phi_m - \chi$$

$$V_{bi} = \Phi_{B0} - \Phi_n$$

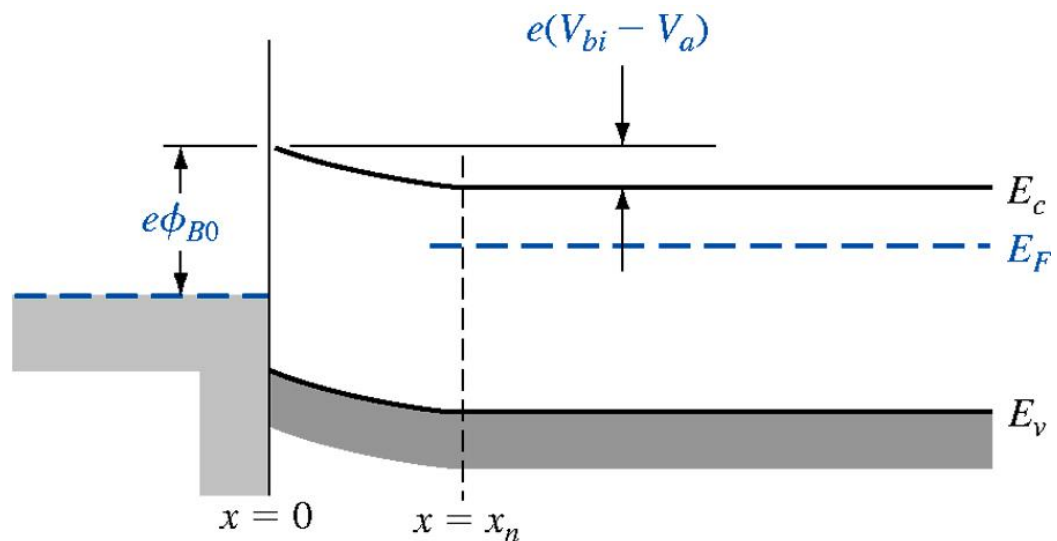
Table 9.2 | Electron affinity of some semiconductors

Element	Electron affinity, $\chi$
Ge, germanium	4.13
Si, silicon	4.01
GaAs, gallium arsenide	4.07
AlAs, aluminum arsenide	3.5



(a)

Legt man eine Vorwärtsspannung  $V_a$  an, fließen Elektronen vom HL ins Metall. Die Potentialbarriere wird verkleinert:  $V_{bi} - V_a$ . Der Strom steigt exponentiell an.



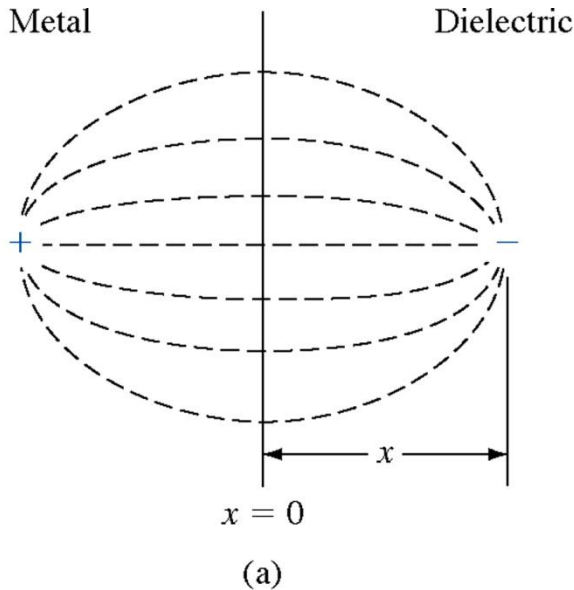
(b)

RLZ-Weite wie im einseitigen pn-Übergang mit  $N_a \rightarrow \infty$

$$W = \left\{ \frac{2\varepsilon(V_{bi} + V_R)}{eN_d} \right\}^{1/2}$$

Legt man eine Sperrspannung  $V_R$  an, fließt ein Elektronenstrom vom Metall in den HL. Die Potentialbarriere bleibt immer gleich der Schottky-Barriere.

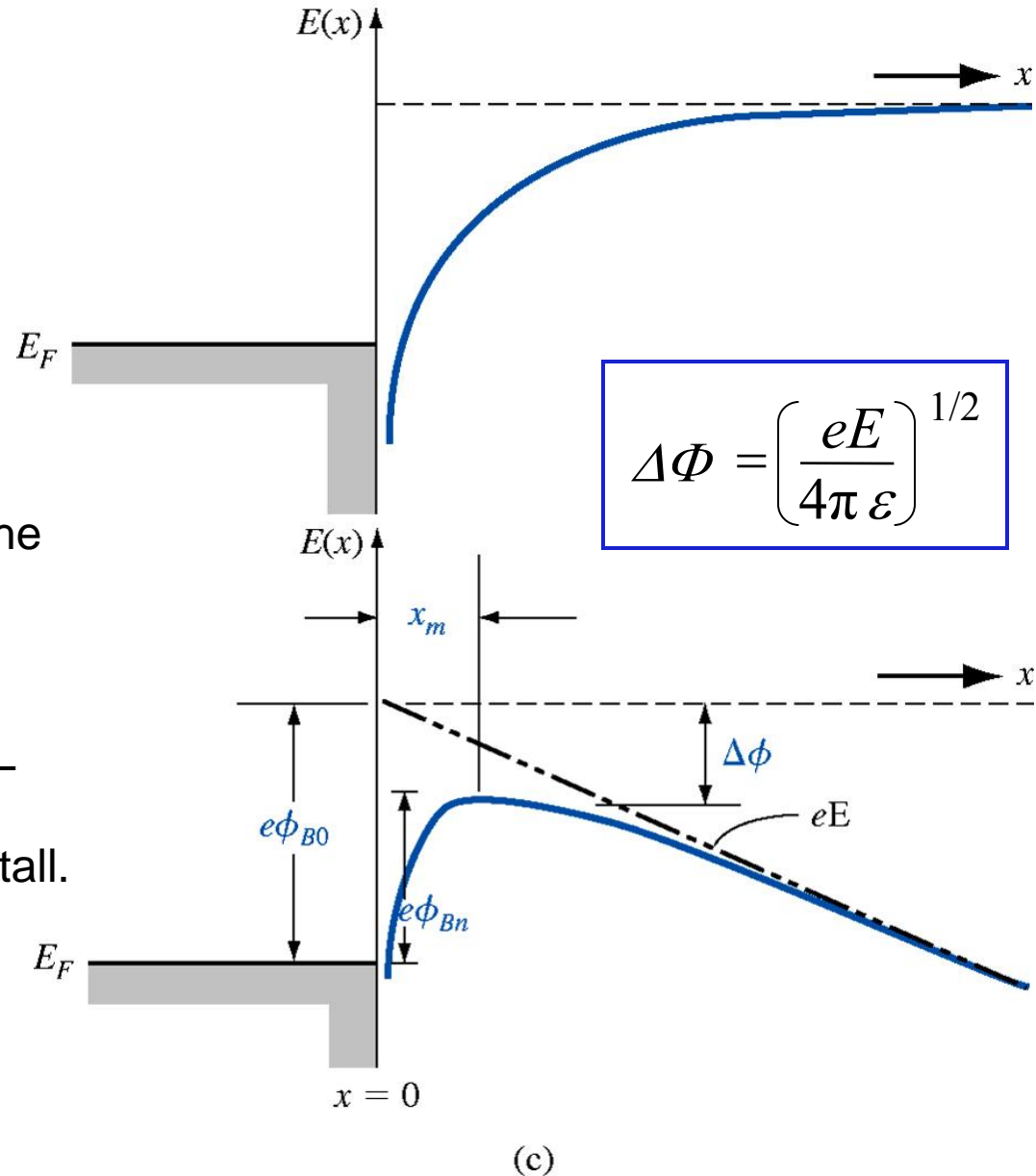
# Schottky-Effekt



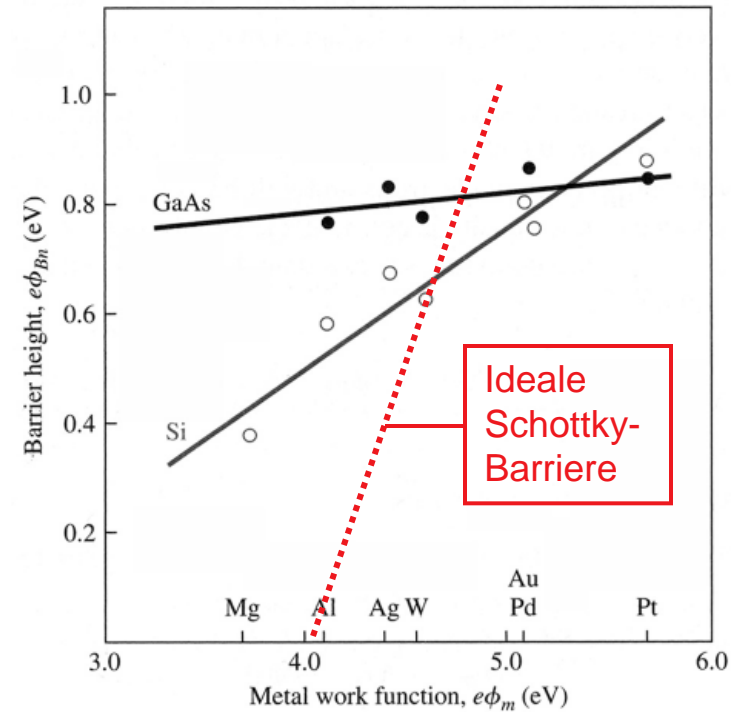
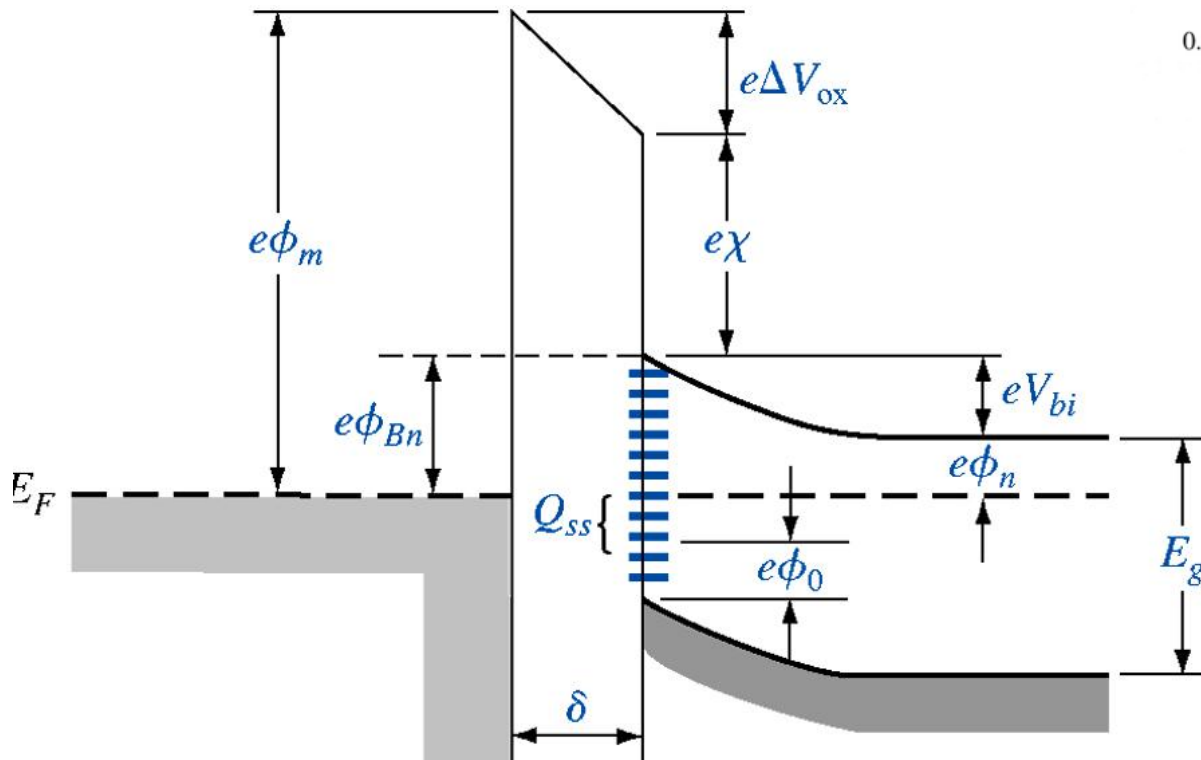
**Bildkraft-Effekt:** An einer Metallfläche müssen die Feldlinien senkrecht einmünden (sonst könnte man Ladungen parallel zur Grenzfläche verschieben). Jedes Elektron im HL erzeugt daher eine symmetrisch liegende positive Bildladung im Metall. Dieses Loch zieht das Elektron mit der Kraft

$$F = \frac{e^2}{16\pi \epsilon x^2}$$

an.



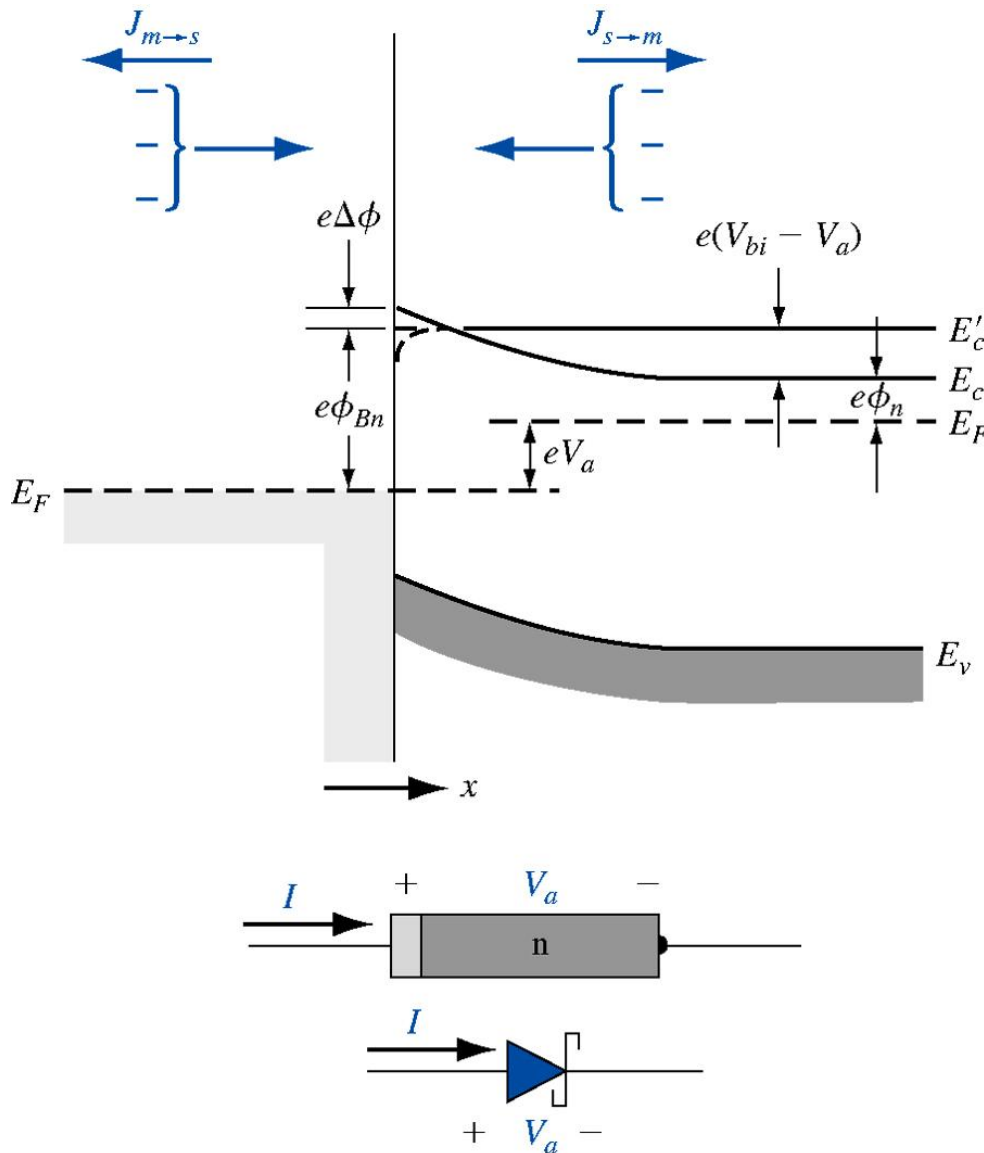
Das Schottky-Modell des Metall-HL-Kontakts ist eine Idealisierung. In Wirklichkeit bestimmen Grenzflächenzustände die Höhe der Barriere. Das Fermi-niveau wird durch diese Zustände "gepinnt" (Bardeen-Modell).



Gemessene Barrierenhöhen als Funktion der Metall-Austrittsarbeit für GaAs und Si (nach *Crowley and Sze*).

Realistisches Modell des Metall-HL-Kontakts mit ultra-dünner Isolatorschicht und Grenzflächenzuständen  $Q_{ss}$ .

# Strom-Spannungs-Charakteristik



Der Strom in der Schottky-Diode wird durch Majoritätsladungen getragen. Im Falle eines n-HLs werden Elektronen thermisch über die Barriere hinweg emittiert (thermionische Emission). Daher bestimmt die Schottky-Barriere den Sperrstrom.

Die Gesamtstromstärke ist

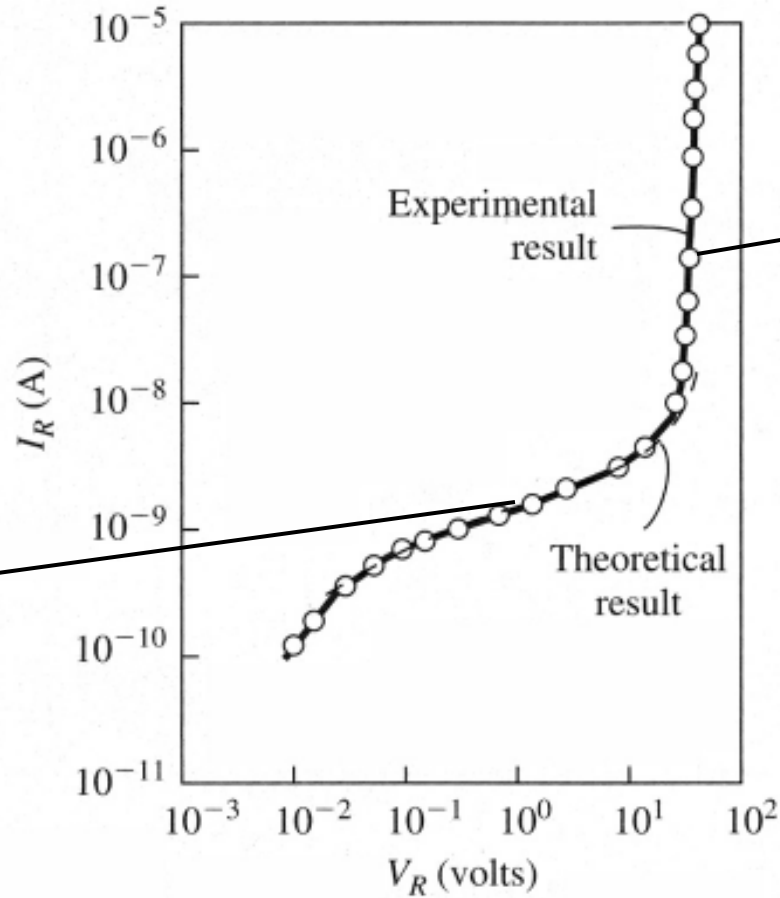
$$j = j_{sT} \left\{ \exp\left(\frac{eV_a}{kT}\right) - 1 \right\}$$

mit

$$j_{sT} = A^* T^2 \exp\left(-\frac{e(\Phi_{B0} - \Delta\Phi)}{kT}\right)$$

und der *Richardson-Konstante*

$$A^* = \frac{4\pi e m_n^* k^2}{h^3}$$

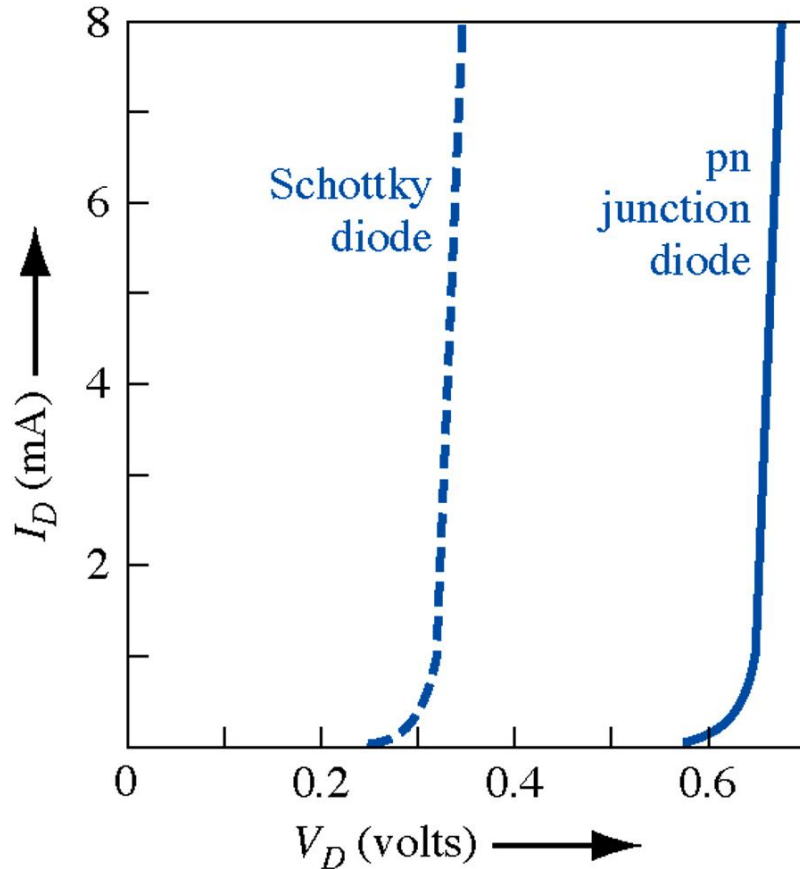


Bildkrafteffekt führt zu Anstieg des Sperrstroms wegen wachsender Barrierenabsenkung  $\Delta\Phi$

Gemessene und berechnete Sperrströme einer PtSi-Si-Diode.



# Vergleich zwischen Schottky-Diode und pn-Diode



Die Verhältnisse hängen stark ab (i) vom Metall-HL-System (Schottky-Barriere, eff. Masse) und (ii) vom Diffusionspotential der pn-Diode (Dotierung).

Schottky-Diode:

$$j_{sT} = A^* T^2 \exp\left[-\frac{e(\Phi_{B0} - \Delta\Phi)}{kT}\right]$$

(thermionische Emission von Majoritätsladungen)

pn-Diode:

$$j_S = \left[ \frac{e D_p p_{n0}}{L_p} + \frac{e D_n n_{p0}}{L_n} \right]$$

(Diffusion von Minoritätsladungsträgern)

Der Generationsstrom in der gesperrten Schottky-Diode ist vernachlässigbar.

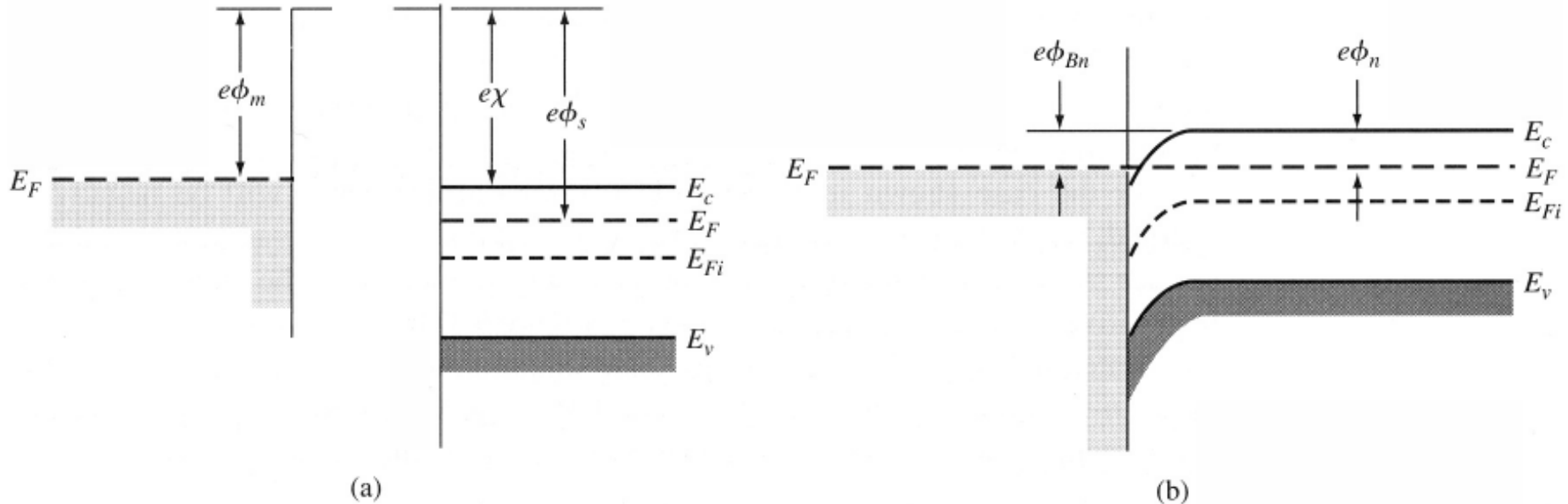
Die Einsatzspannung der pn-Diode ist i.a. grösser als die der Schottky-Diode.

**Schaltverhalten:** Schottky-Dioden haben praktisch keine Diffusionskapazität und keine Speicherzeit für die Minoritäten.

→ Schaltzeiten im Pikosekundenbereich

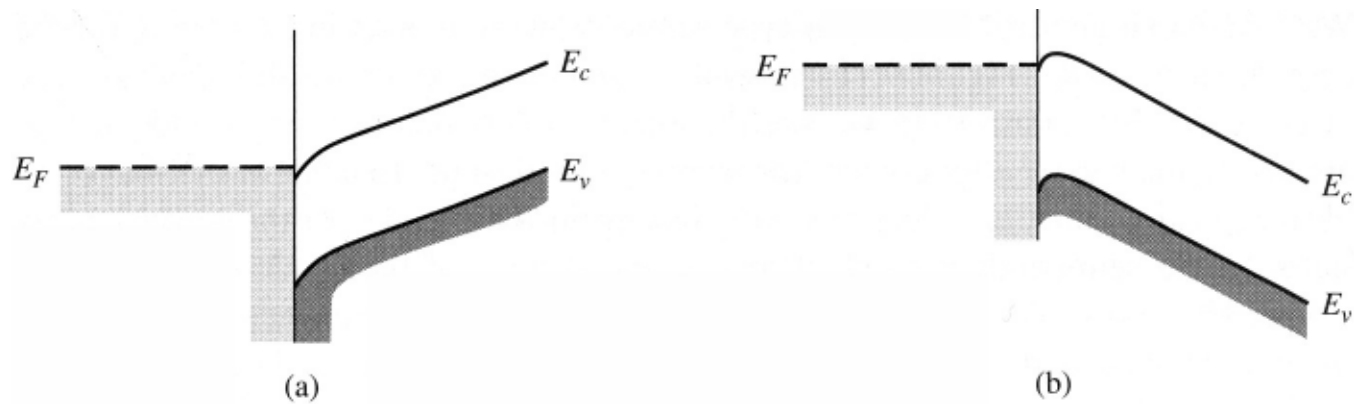
# 8.2 Der Ohmsche Kontakt

## Bandkanten-Diagramme (Fall n-HL und $\Phi_m < \Phi_s$ )

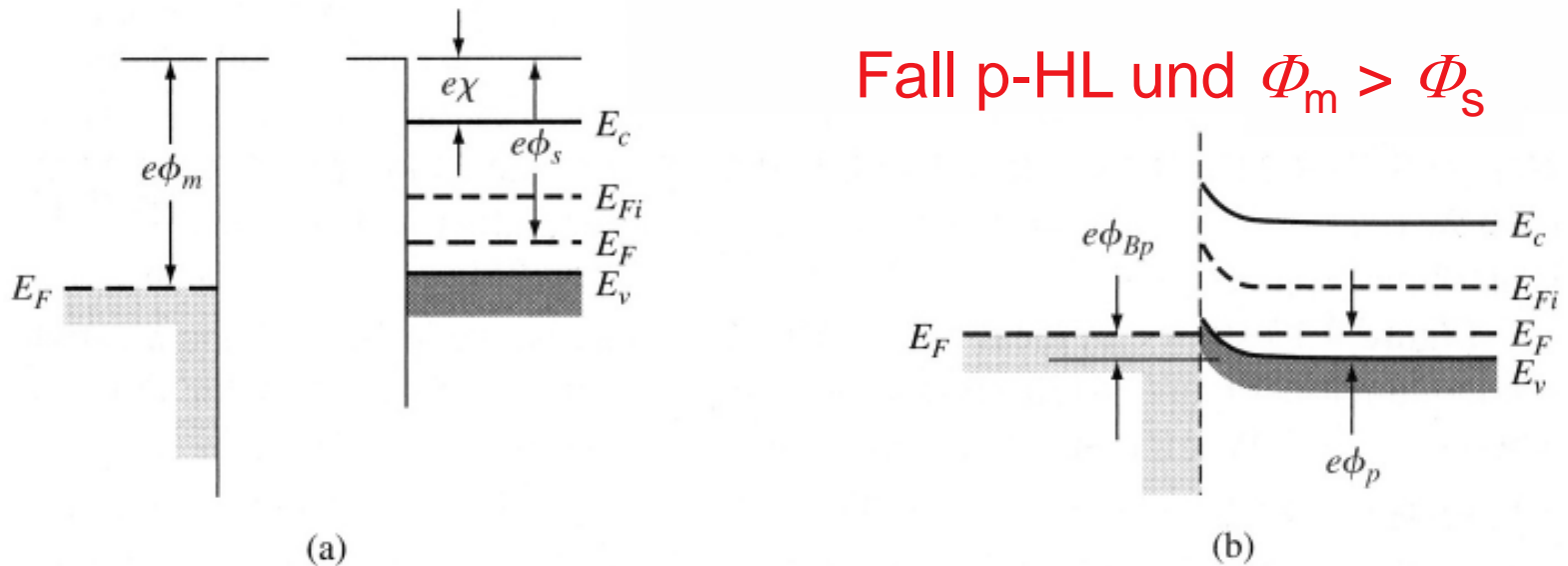


Ideale Bandkendiagramme eines Metall-n-HL-Übergangs für den Fall  $\Phi_m < \Phi_s$  (a) vor der Kontaktierung und (b) nach der Kontaktierung.

- Elektronen fließen vom Metall in den n-HL → dünne Anreicherungsschicht im HL
- Der Elektronen-Übergang HL → Metall enthält überhaupt keine Barriere
- Für den Übergang Metall → HL ist die Potentialbarriere gleich  $e\Phi_n$ , die durch hohe Dotierung klein gemacht werden kann.



Ideale Bandkendiagramme eines Ohmschen Metall-n-HL-Übergangs (a) mit positiver Spannung am Metall und (b) mit positiver Spannung am Halbleiter.



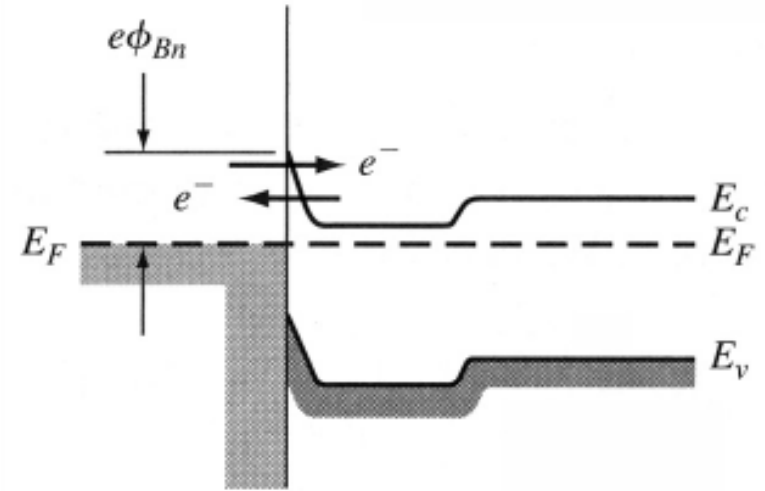
Fall p-HL und  $\Phi_m > \Phi_s$

Ideale Bandkendiagramme eines Metall-p-HL-Übergangs für den Fall  $\Phi_m > \Phi_s$  (a) vor der Kontaktierung und (b) nach der Kontaktierung.

# Tunnel-Barriere

Fall n-HL und  $\Phi_m > \Phi_s$   
 und Fall p-HL,  $\Phi_m < \Phi_s$

- Bei zu geringer HL-Dotierung entstehen in diesen Fällen gleichrichtende Kontakte (breite Verarmungszonen).
- Ohmsche Kontakte werden durch sehr hohe Dotierung des HLs im Kontaktgebiet erreicht.
- Trotz der hohen Schottky-Barriere können die Ladungsträger diese dann leicht durchtunneln, da die Barriere extrem dünn wird (im Ångstrom-Bereich).

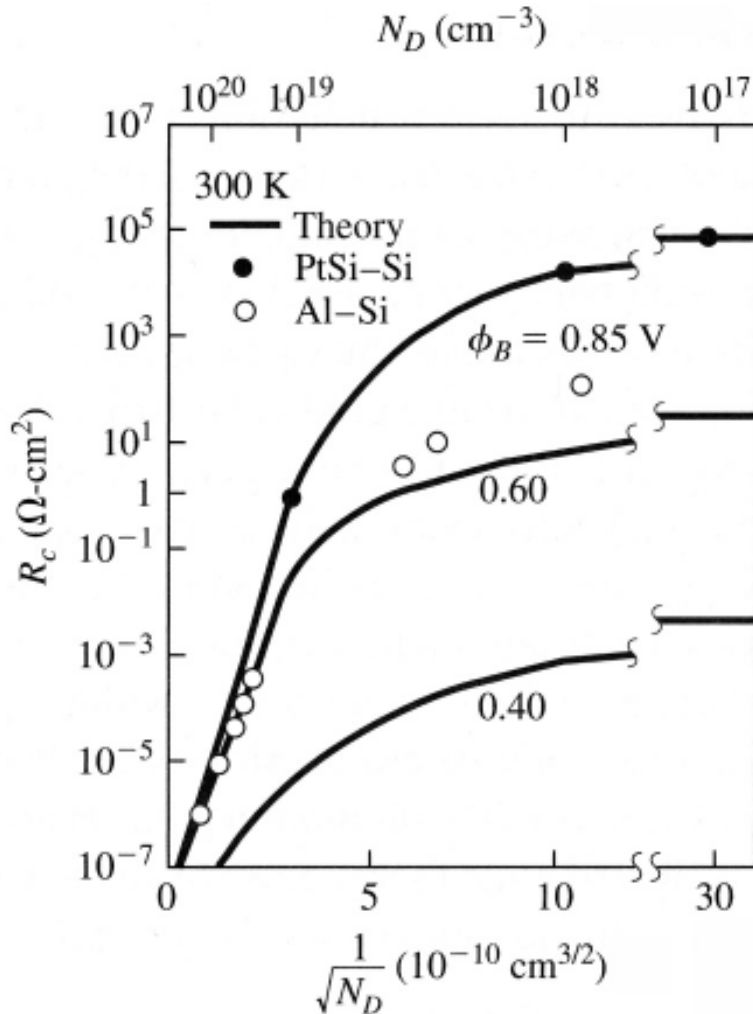


Bandkantendiagramm eines stark dotierten Metall-n-HL-Übergangs.

$$j_t \sim \exp\left[-\frac{e\Phi_{Bn}}{E_{00}}\right]$$

$$E_{00} = \frac{e\hbar}{2} \left(\frac{N_d}{\epsilon m_n^*}\right)^{1/2}$$

# Spezifischer Kontakt-Widerstand



Theoretischer und gemessener spezifischer Kontakt-Widerstand als Funktion der Dotierung (nach Sze).

$$R_c = \left. \left( \frac{\partial j}{\partial V} \right)^{-1} \right|_{V=0}$$

Schottky-Kontakt:

$$R_c = \frac{kT}{e} \frac{\exp\left(\frac{e\Phi_{Bn}}{kT}\right)}{A^*T^2}$$

(niedrige und moderate Dotierung, thermionische Emission)

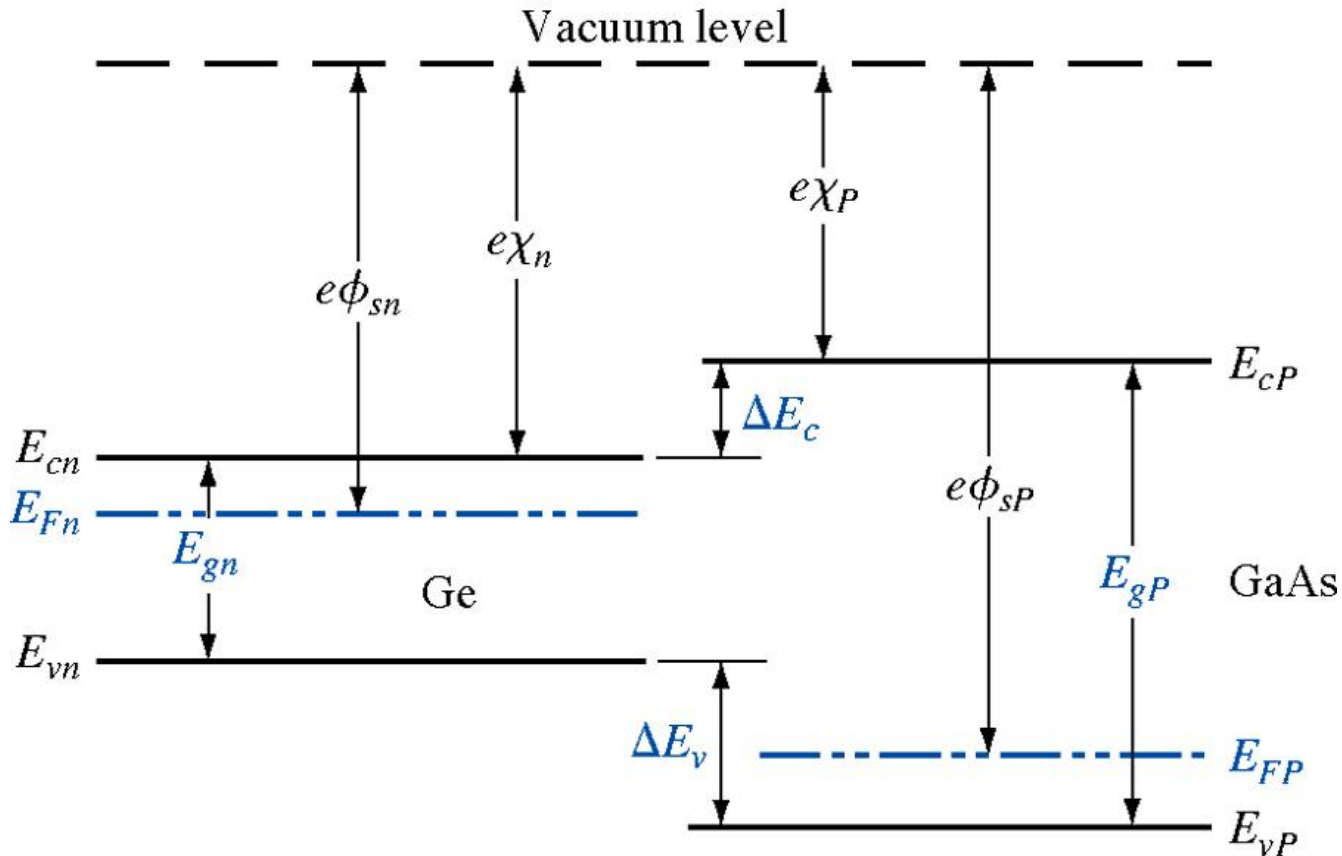
Tunnel-Barriere:

$$R_c \sim \exp\left(\frac{2\Phi_{Bn}}{\hbar} \left(\frac{\epsilon m_n^*}{N_d}\right)^{1/2}\right)$$

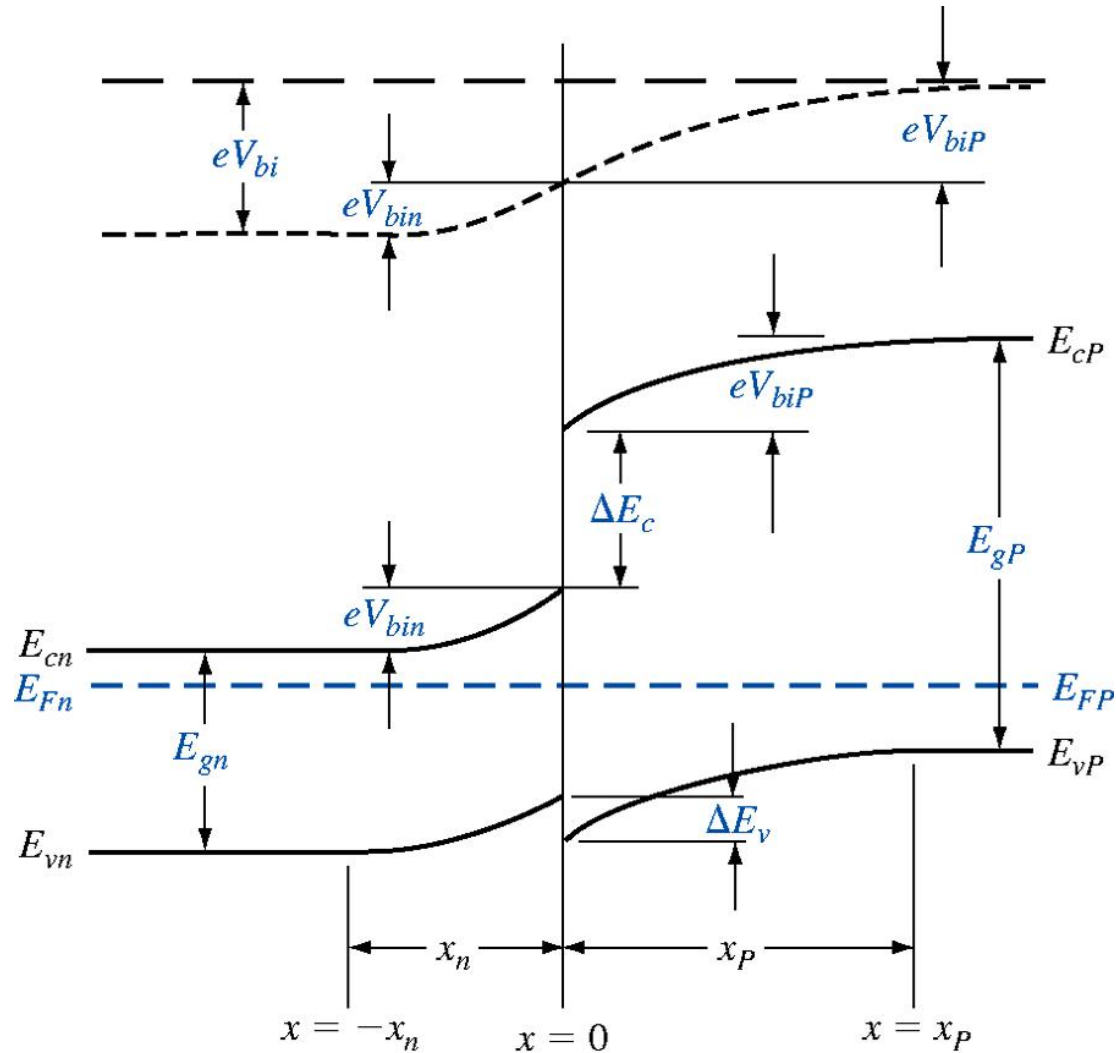
(sehr hohe Dotierung, Tunnelstrom)

# 8.3 Hetero-Übergänge

## Bandkanten-Diagramme

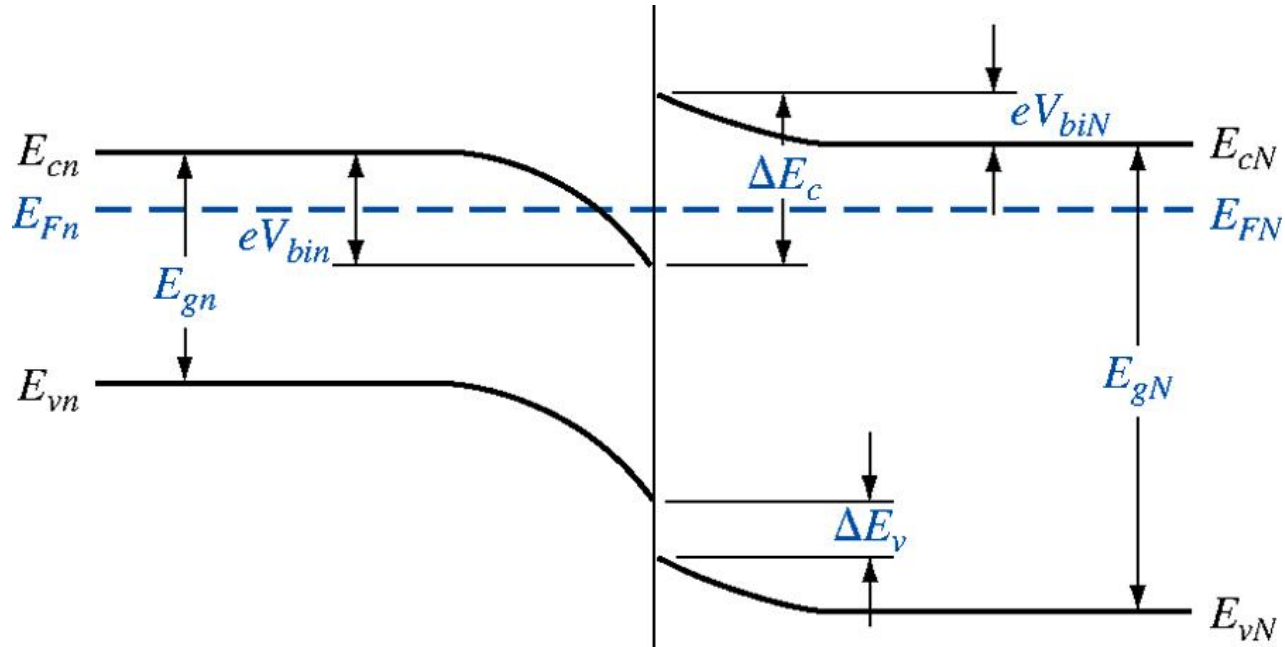


Isolierte n- und P-Materialien mit Vakuum-Niveau als Referenzenergie. Nach dem Kontaktieren entsteht ein anisotyper nP-Übergang.



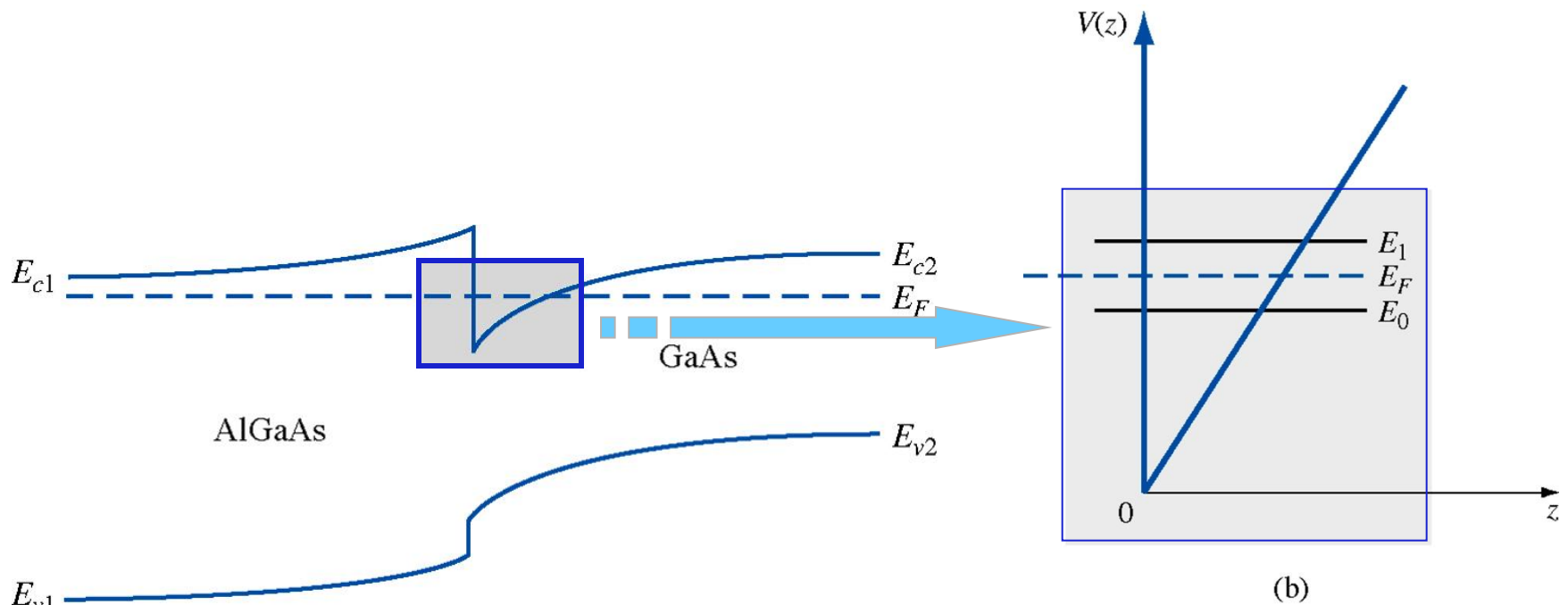
Idealer nP-Hetero-Übergang im thermodynamischen Gleichgewicht. Die Verbiegung des Vakuum-Niveaus entspricht dem ortsabhängigen elektrostatischen Potential.

## 2-dimensionales Elektronengas

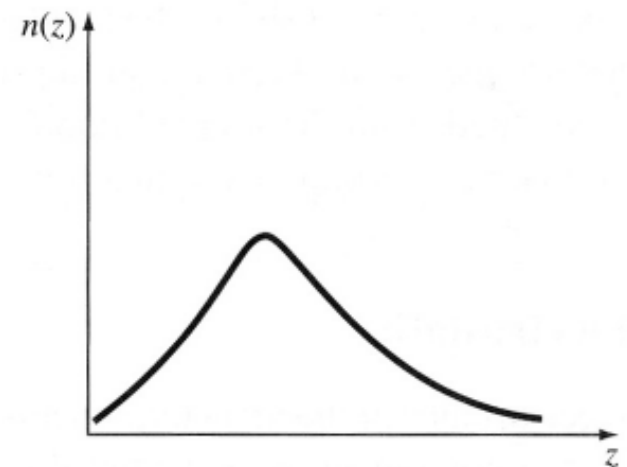


nN-Hetero-Übergang (GaAs-AlGaAs) im thermodynamischen Gleichgewicht. Im Potentialtrichter sammeln sich Elektronen und bilden ein zweidimensionales Gas.

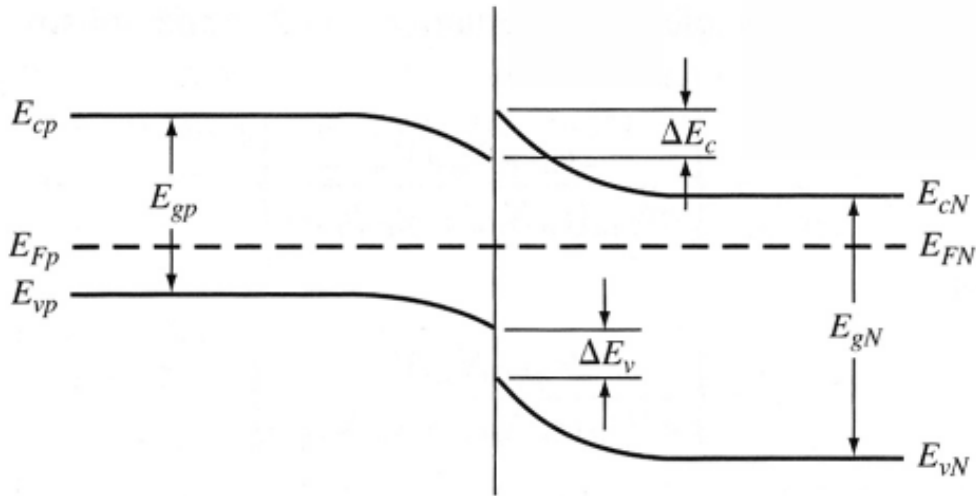




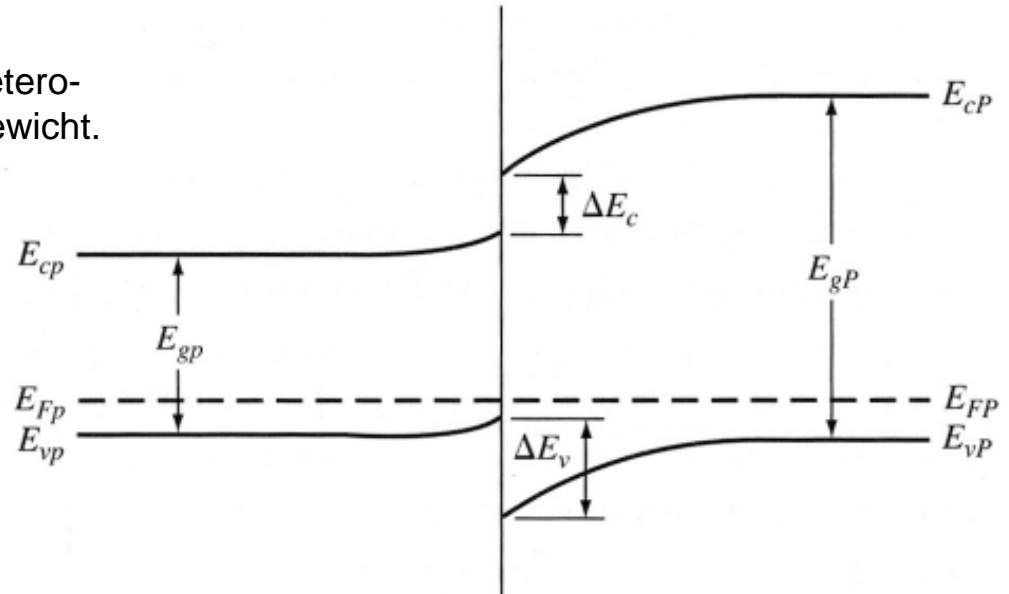
nN-Hetero-Übergang (GaAs-AlGaAs) im thermodynamischen Gleichgewicht. Die Energien im 2DEG sind quantisiert. In der Nähe der Grenzfläche ist der Potentialtrichter dreiecksförmig. Wird die GaAs-Seite nur schwach oder gar nicht dotiert, steigt die Beweglichkeit im Vergleich zum Homo-Material mit derselben Dichte stark an (keine Störstellenstreuung).



Elektronendichte im dreiecksförmigen Potentialtopf.



Ideales Bandkantendiagramm eines Np-Hetero-Übergangs im thermodynamischen Gleichgewicht.



Ideales Bandkantendiagramm eines pP-Hetero-Übergangs im thermodynamischen Gleichgewicht.







# 9. Der Bipolar-Transistor (1)

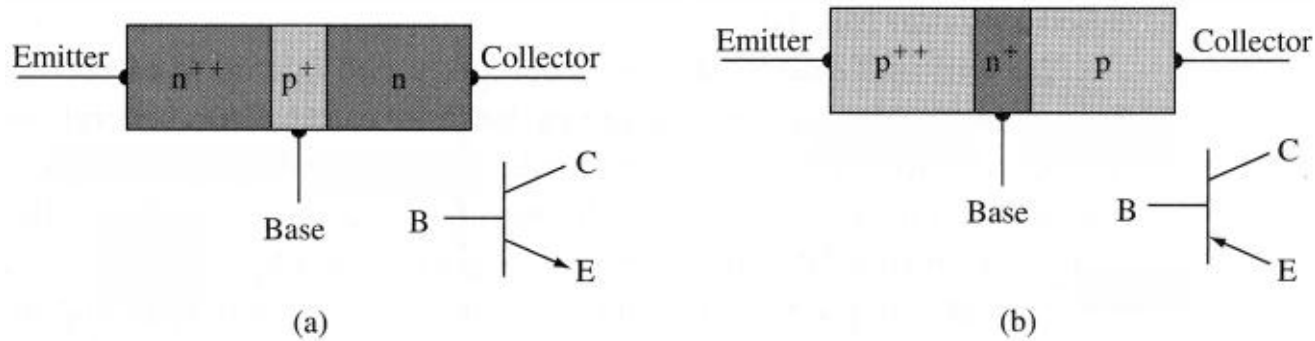
**9.1 Wirkungsweise des Bipolar-Transistors (N10.1)**

**9.2 Minoritätsladungsträger-Verteilungen (N10.2)**

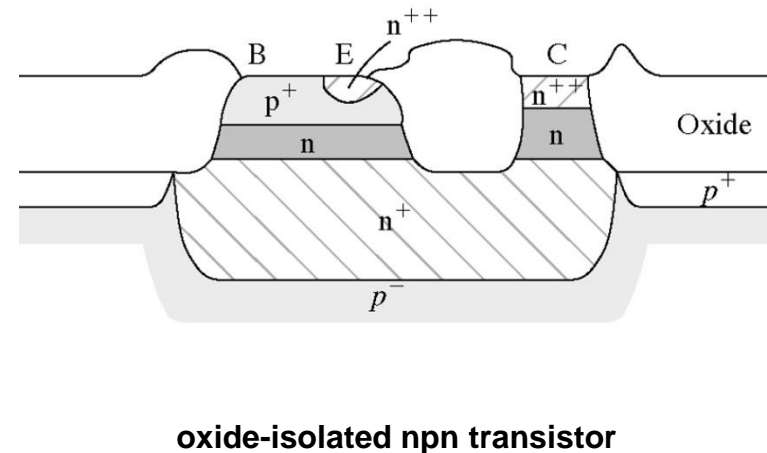
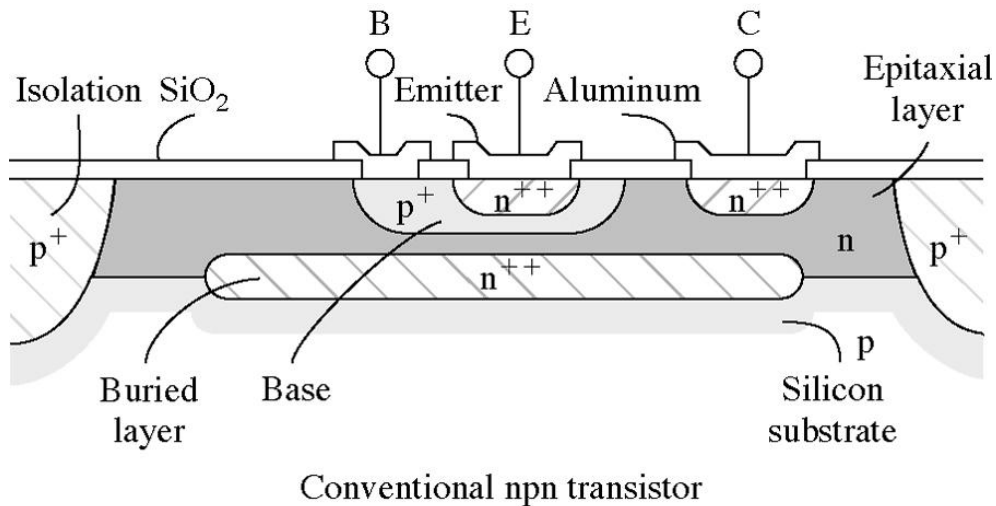
**9.3 Stromverstärkung in Basisschaltung (N10.3)**

# 9.1 Wirkungsweise des Bipolar-Transistors

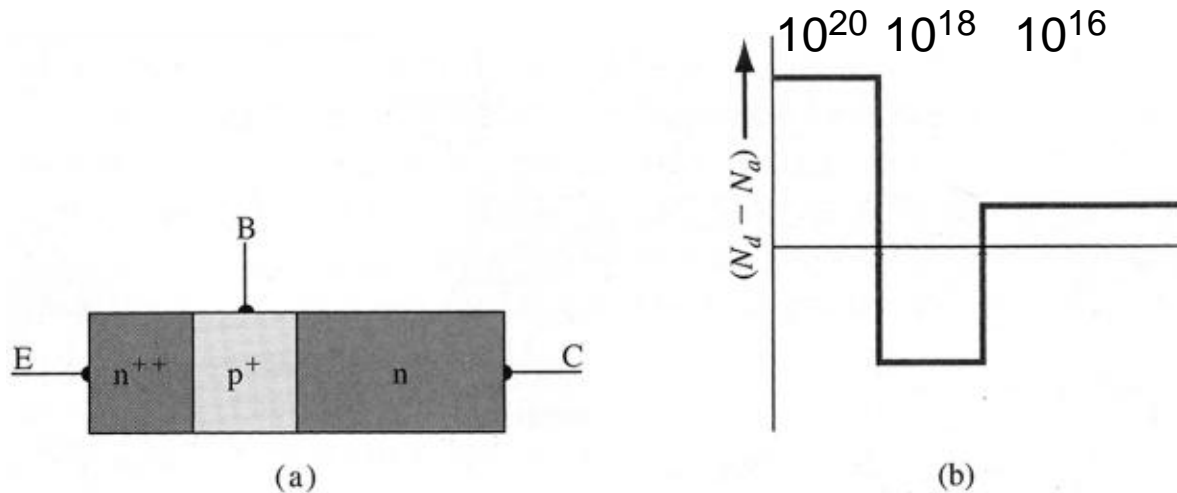
## Die Transistor-Struktur



Vereinfachtes Blockdiagramm und Schaltkreissymbole eines (a) npn- und (b) pnp-Bipolar-Transistors.



# Vereinfachte Struktur



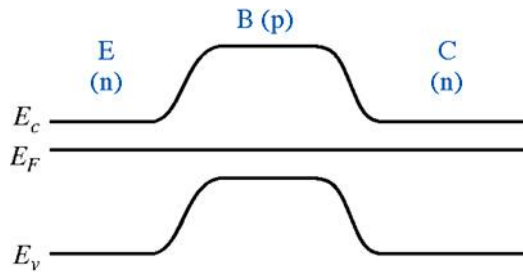
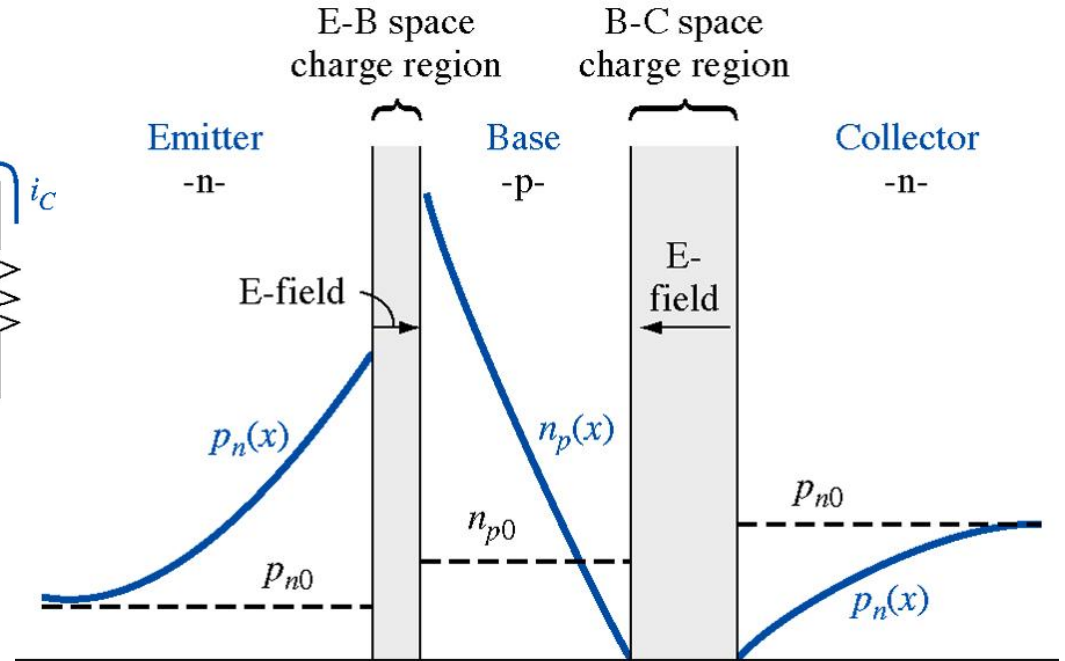
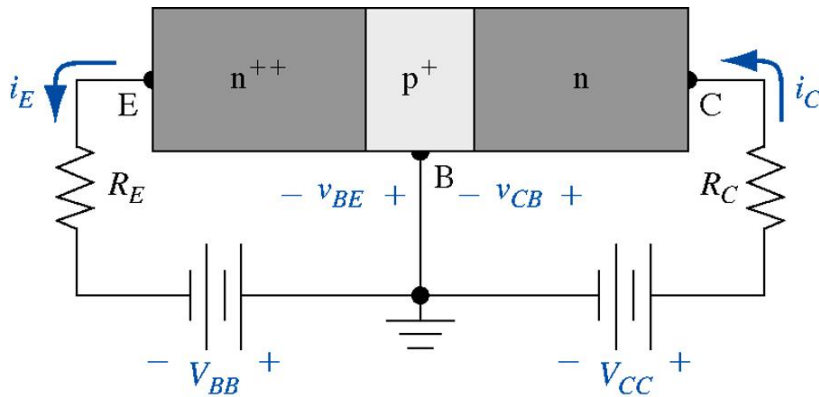
Idealisiertes Dotierprofil eines homogen dotierten npn-Bipolar-Transistors.

- Der Bipolar-Transistor ist eine spannungskontrollierte Stromquelle.
- Seine Funktionalität beruht auf der Wechselwirkung von zwei (eng) benachbarten pn-Übergängen.
- Die Dotierungen in Emitter/Basis/Kollektor sind etwa  $10^{20}/10^{18}/10^{16}$  cm<sup>-3</sup>.
- Die Basisweite muss klein sein, um die Wechselwirkung der beiden pn-Übergänge zu ermöglichen.

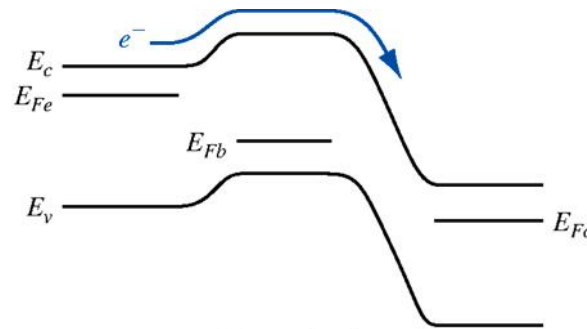


# Das grundlegende Funktionsprinzip

## Vorwärts-aktiver Modus

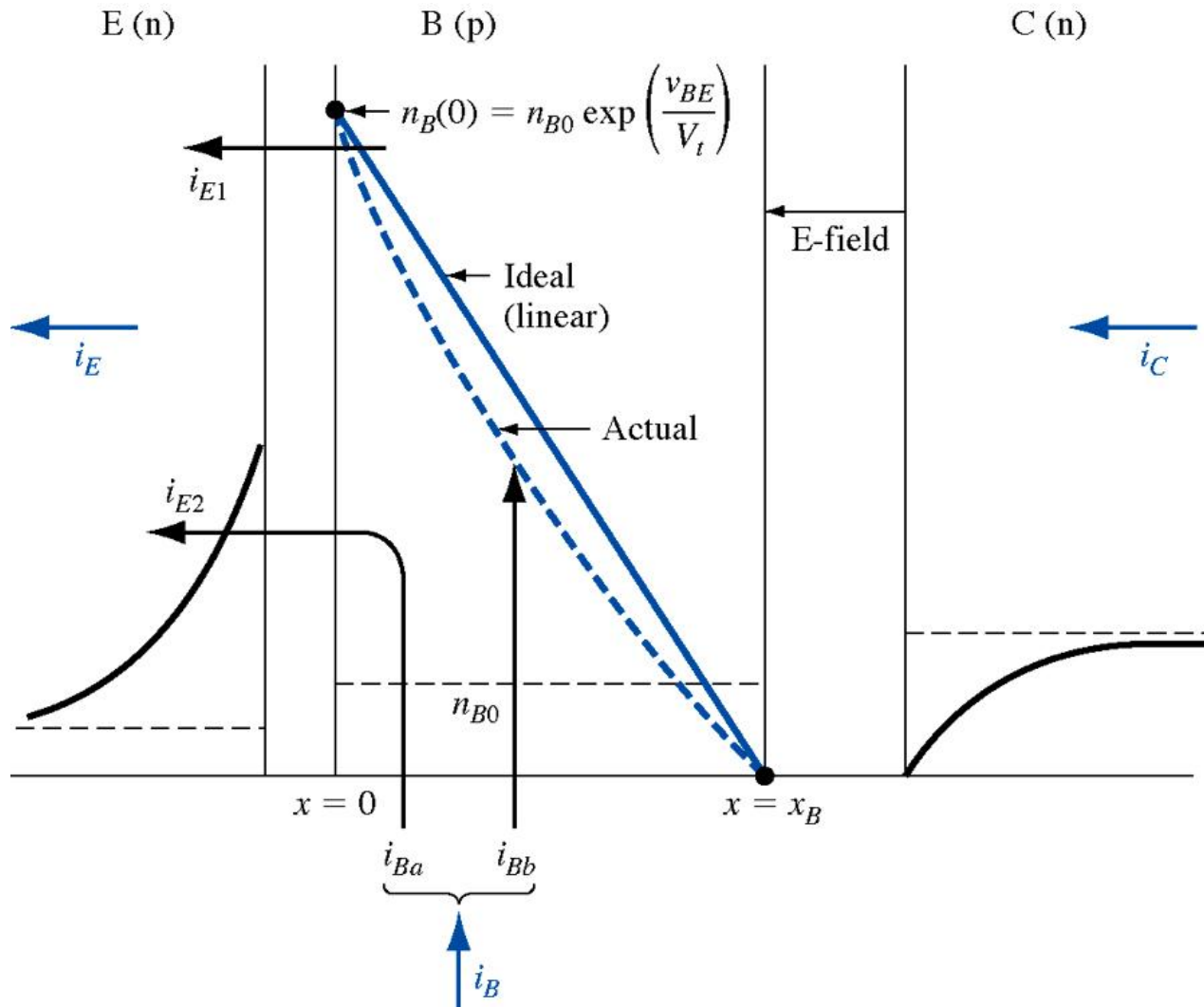


Zero bias



Forward active

# Gleichstromverhalten (vereinfacht)



## Kollektor-Strom

$$\begin{aligned}
 i_C &= eA_{BE}D_n \frac{dn}{dx} \\
 &= -eA_{BE}D_n n_B(0)/x_B \\
 &= I_S \exp(V_{BE}/V_t)
 \end{aligned}$$

## Emitter-Strom

$$\begin{aligned}
 i_E &= i_{E1} + i_{E2} = i_C + i_{E2} = \\
 &= I_{SE} \exp(V_{BE}/V_t)
 \end{aligned}$$

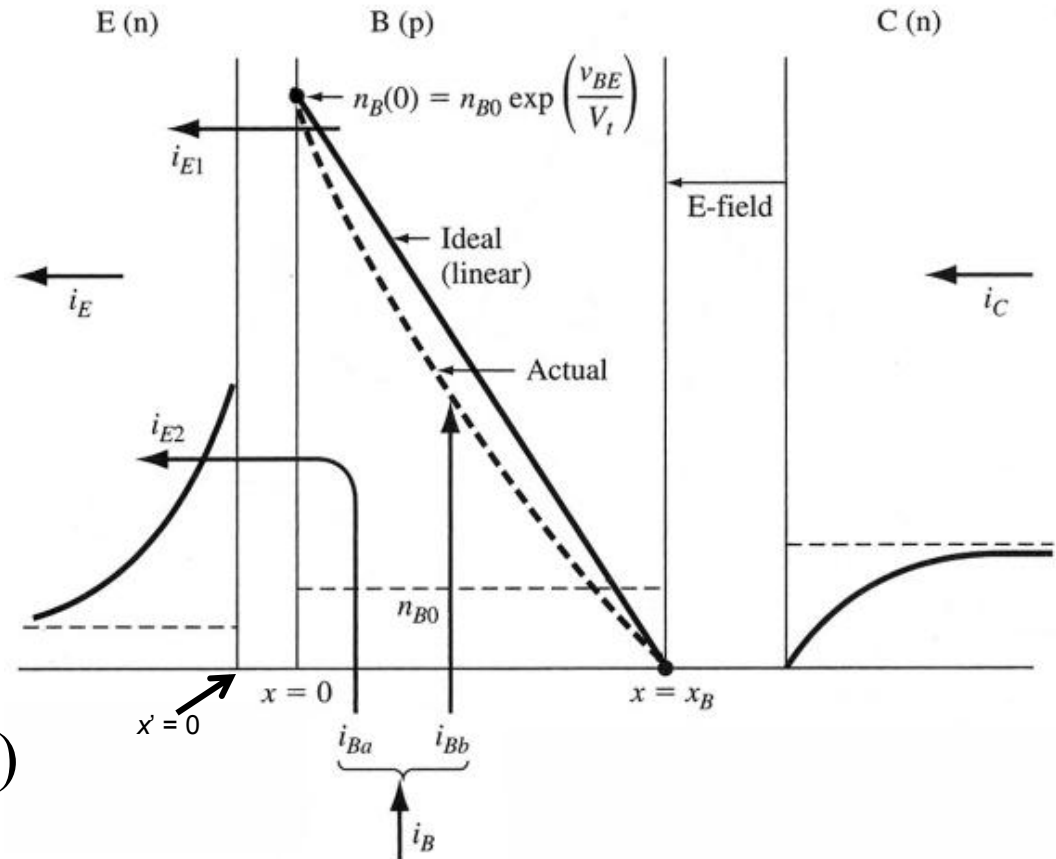
$$i_{E2} = I_{S2} \exp(V_{BE}/V_t)$$

## Basis-Strom

$$i_B = i_{Ba} + i_{Bb} = I_{SB} \exp(V_{BE}/V_t)$$

$$\frac{i_C}{i_E} \equiv \alpha \quad \text{und} \quad \frac{i_C}{i_B} \equiv \beta$$

$$\begin{aligned}
 n_B(x=0) &= n_{B0} \exp(V_{BE}/V_t) \\
 p_E(x=0) &= p_{E0} \exp(V_{BE}/V_t)
 \end{aligned}$$



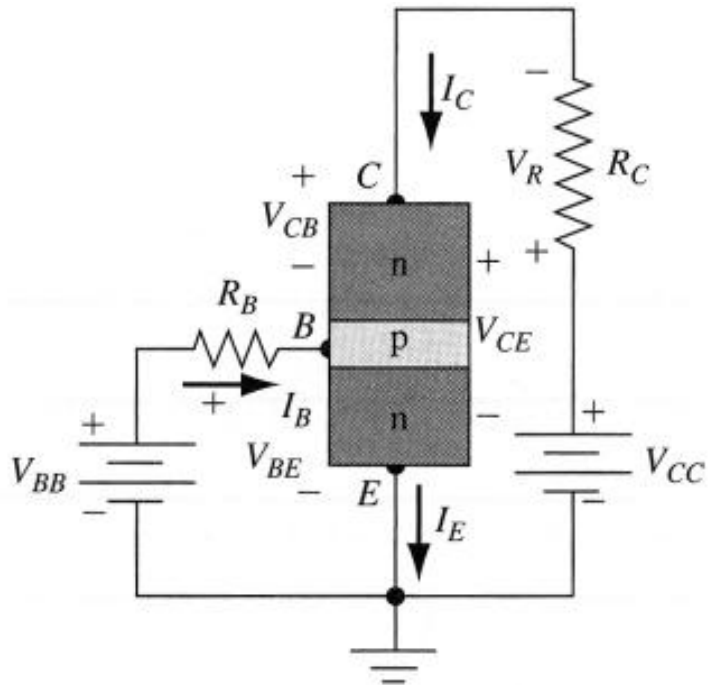
Minoritätsträger-Verteilung und grundlegende Ströme in einem vorwärts gepolten npn-Bipolar-Transistor.

# Operationsmodi

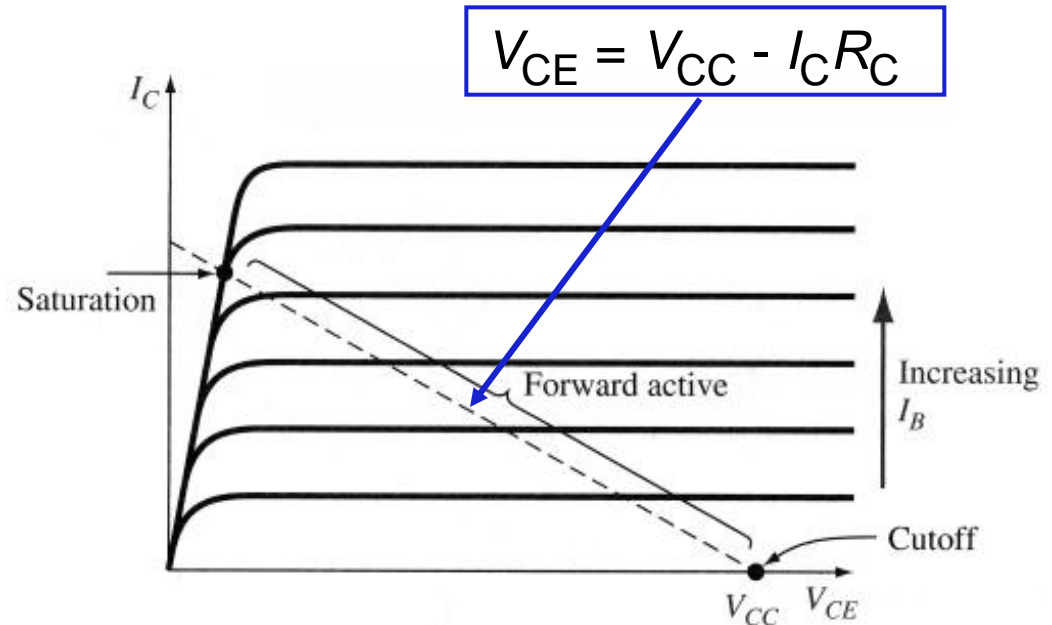
- Kirchhoffsches Gesetz im Kollektor-Emitter-Kreis:

$$V_{CC} = I_C R_C + V_{CB} + V_{BE} = V_R + V_{CE}$$

- $V_{CB} > 0$ : gesperrter Kollektor-Basis-Übergang  $\rightarrow$  vorwärts-aktiver Modus
- wachsende  $V_{BE} > 0 \rightarrow V_R$  steigt an, bis  $V_{CB} < 0$  wird; dann "Sättigung"

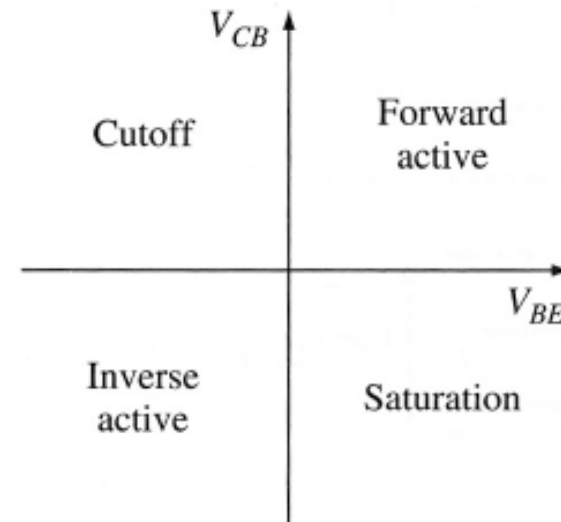
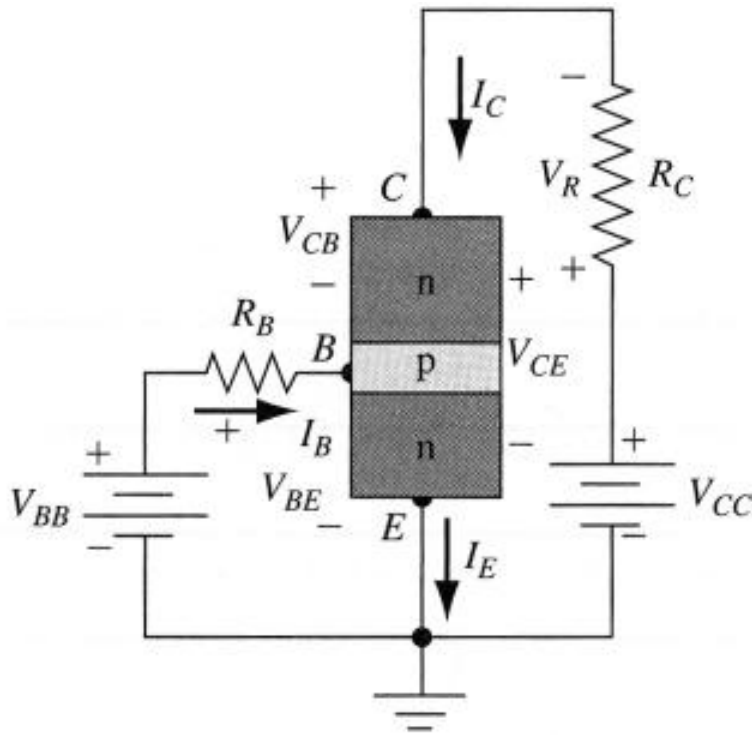


Der npn-Bipolar-Transistor in Emitterschaltung ("common emitter").



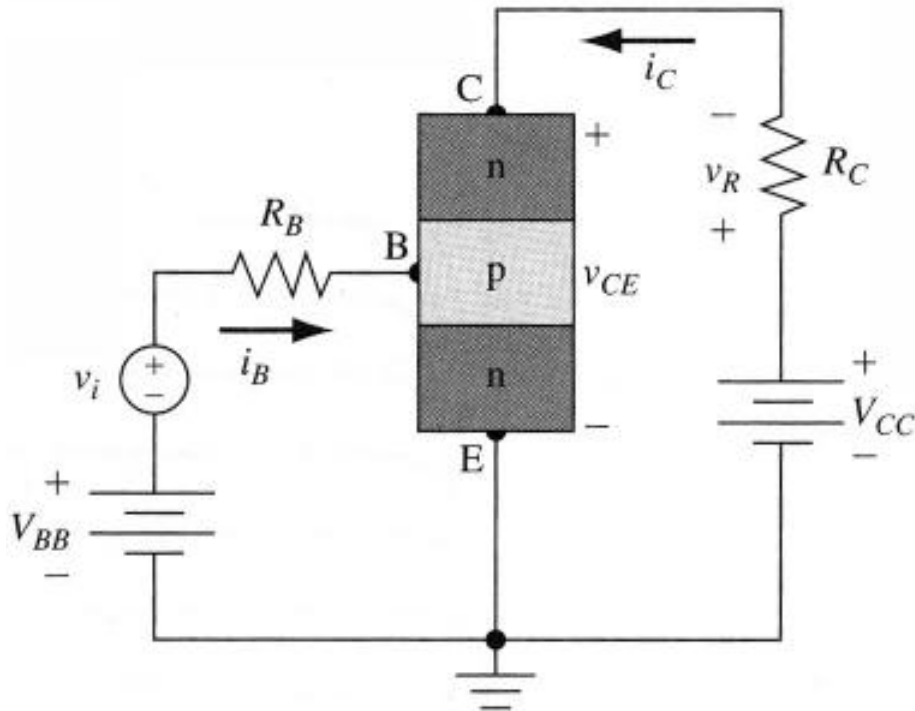
Strom-Spannungs-Charakteristiken eines npn-Bipolar-Transistor in Emitterschaltung ("common emitter") mit Load Line.

- Wenn  $V_{BE} \leq 0$ , dann keine Injektion von Elektronen aus dem Emitter in die Basis  
→ “cutoff”
- Wenn  $V_{BE} \leq 0$  und gleichzeitig  $V_{BC} \geq 0$  → Rollen von Emitter und Kollektor vertauscht → rückwärts-aktiver Modus

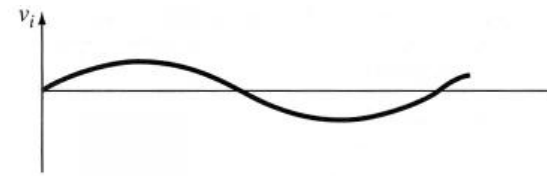


Spannungsbedingungen für die vier Operationsmodi des Bipolar-Transistors.

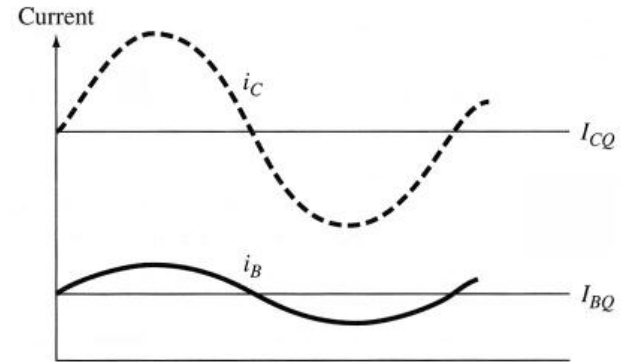
# Verstärkung im Bipolar-Transistor



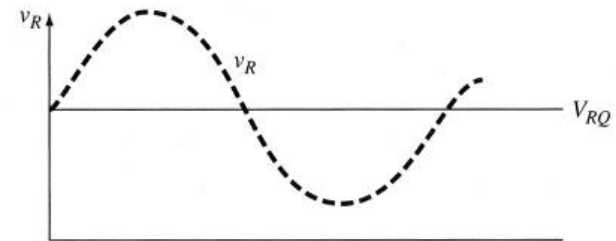
Der npn-Bipolar-Transistor in Emitterschaltung ("common emitter") mit zeitabhängiger Signalspannung  $v_i$  im Basis-Emitter-Kreis.



(a)



(b)

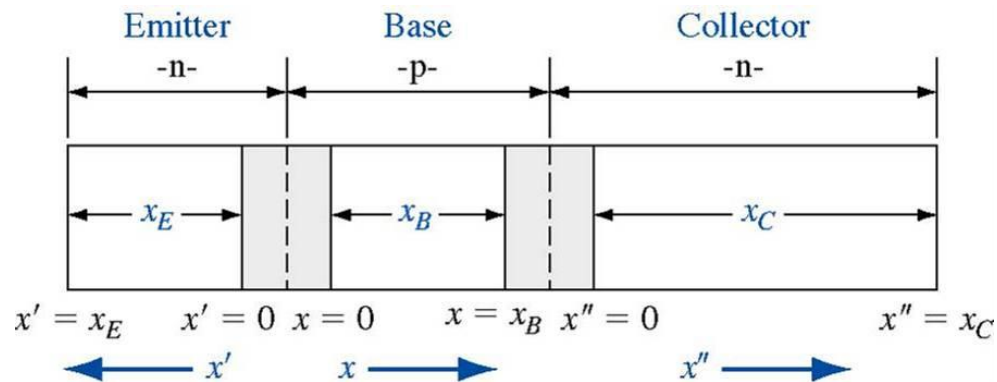


(c)

Ströme und Spannungen im linken Schaltkreis: (a) Sinusförmiges Eingangssignal, (b) den dc-Werten überlagerte sinusförmige Basis- und Kollektorströme, (c) dem dc-Wert überlagerter sinusförmiger Spannungsabfall über dem Lastwiderstand  $R_C$ .

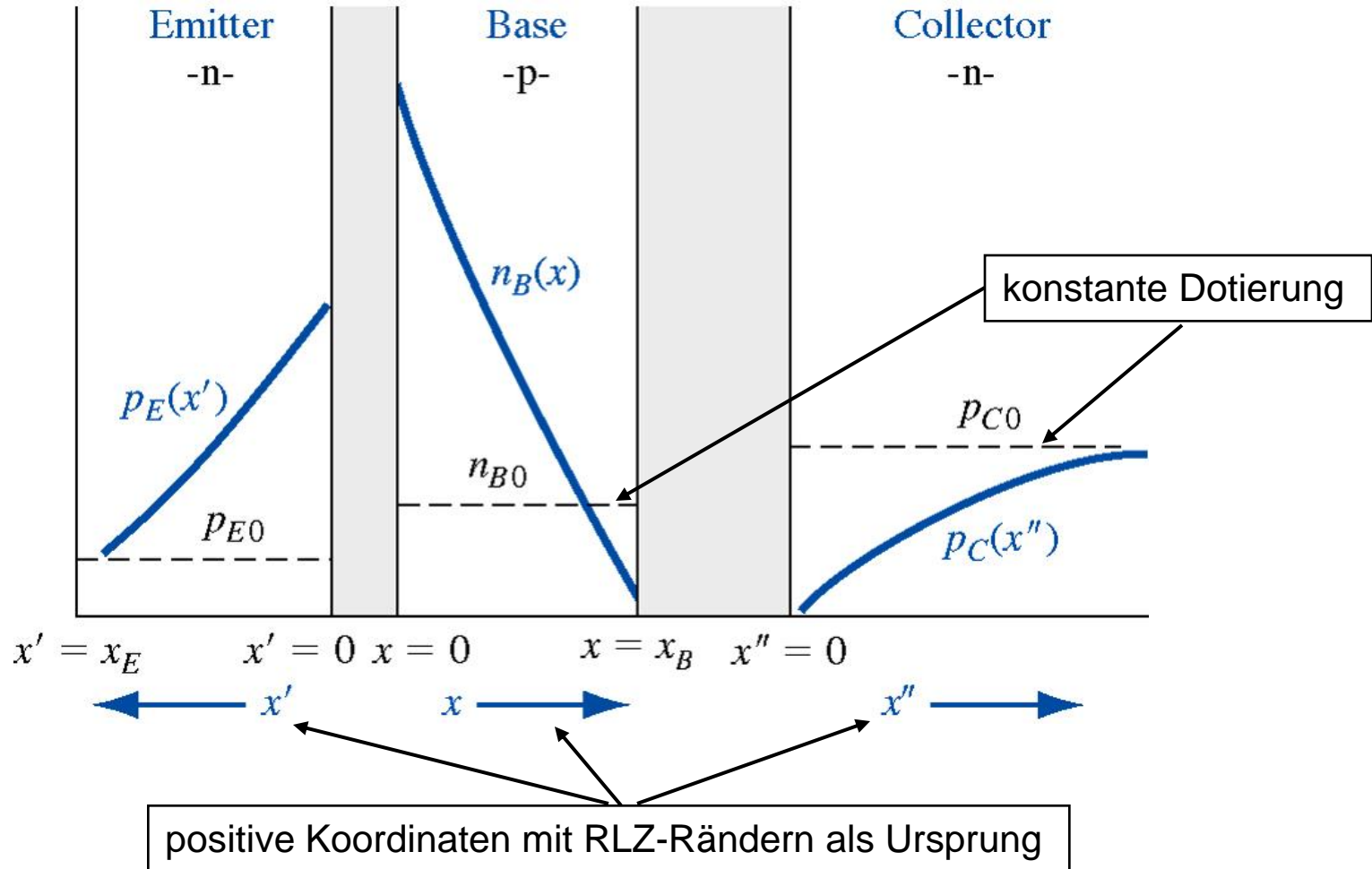
**Table 10.1** | Notation used in the analysis of the bipolar transistor

Notation	Definition
<b>For both the npn and pnp transistors</b>	
$N_E, N_B, N_C$	Doping concentrations in the emitter, base, and collector
$x_E, x_B, x_C$	Widths of neutral emitter, base, and collector regions
$D_E, D_B, D_C$	Minority carrier diffusion coefficients in emitter, base, and collector regions
$L_E, L_B, L_C$	Minority carrier diffusion lengths in emitter, base, and collector regions
$\tau_{E0}, \tau_{B0}, \tau_{C0}$	Minority carrier lifetimes in emitter, base, and collector regions
<b>For the npn</b>	
$p_{E0}, n_{B0}, p_{C0}$	Thermal equilibrium minority carrier hole, electron, and hole concentrations in the emitter, base, and collector
$p_E(x'), n_B(x), p_C(x'')$	Total minority carrier hole, electron, and hole concentrations in the emitter, base, and collector
$\delta p_E(x'), \delta n_B(x), \delta p_C(x'')$	Excess minority carrier hole, electron, and hole concentrations in the emitter, base, and collector
<b>For the pnp</b>	
$n_{E0}, p_{B0}, n_{C0}$	Thermal equilibrium minority carrier electron, hole, and electron concentrations in the emitter, base, and collector
$n_E(x'), p_B(x), n_C(x'')$	Total minority carrier electron, hole, and electron concentrations in the emitter, base, and collector
$\delta n_E(x'), \delta p_B(x), \delta n_C(x'')$	Excess minority carrier electron, hole, and electron concentrations in the emitter, base, and collector



# 9.2 Minoritätsladungsträger-Verteilungen

## Vorwärts-aktiver Modus





# Basis-Region

Diffusionsgleichung für Elektronen in der Basis:

$$\frac{\partial^2 \delta n_B}{\partial x^2} = \frac{\delta n_B}{D_B \tau_{B0}} = \frac{\delta n_B}{L_B^2}$$

Allgemeine Lösung:

$$\delta n_B(x) = n_B(x) - n_{B0} = A \exp\left[\frac{x}{L_B}\right] + B \exp\left[-\frac{x}{L_B}\right]$$

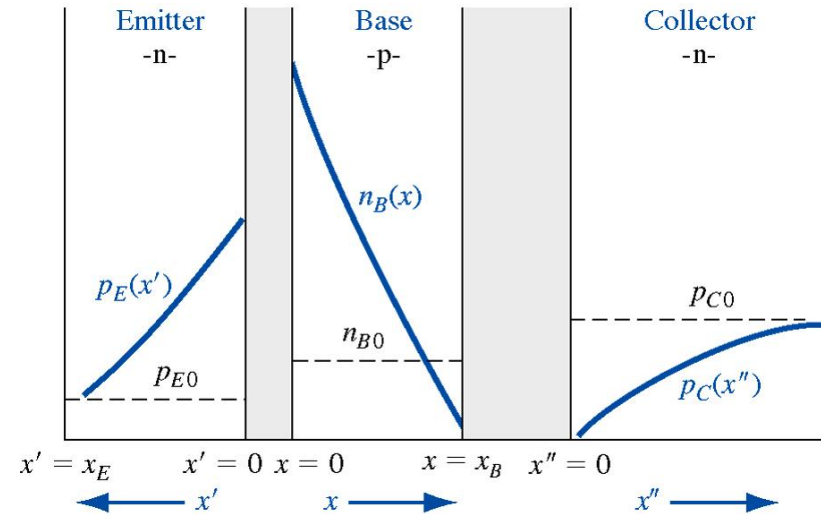
Zwei Randbedingungen:

$$\delta n_B(0) = n_B(0) - n_{B0} = n_{B0} \left\{ \exp\left[\frac{eV_{BE}}{kT}\right] - 1 \right\}$$

$$\delta n_B(x_B) = n_B(x_B) - n_{B0} = 0 - n_{B0} = -n_{B0}$$

Spezielle Lösung:

$$\delta n_B(x) = n_{B0} \frac{\left\{ \exp\left[\frac{eV_{BE}}{kT}\right] - 1 \right\} \sinh[(x_B - x)/L_B] - \sinh[x/L_B]}{\sinh[x_B/L_B]}$$



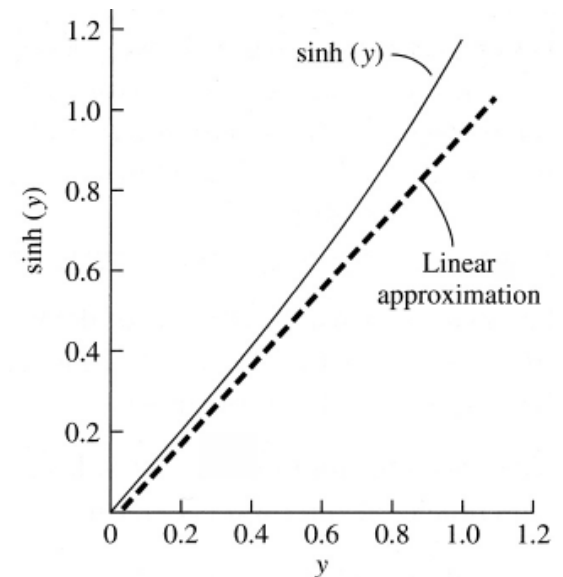
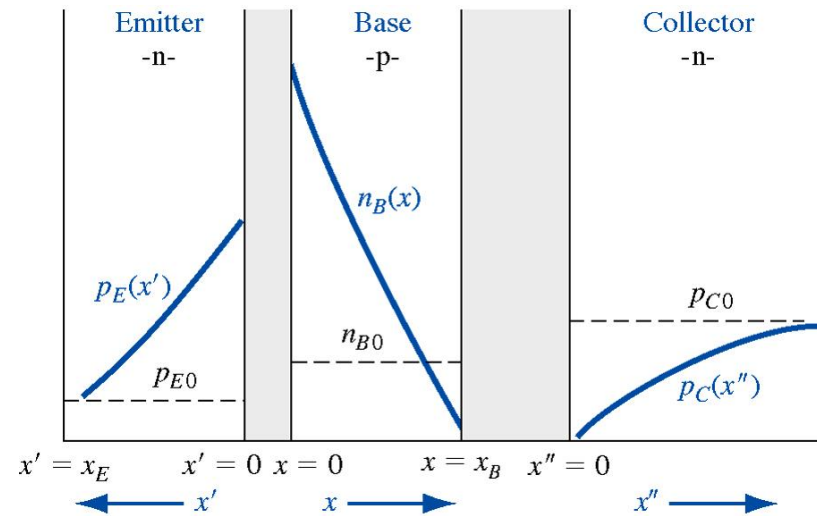
# Basis-Region (2)

- Im allgemeinen ist  $x_B \ll L_B$  (sonst bekommt man keinen guten Transistor)

Entwickelt man die  $\sinh(*)$  linear, so folgt:

$$\delta n_B(x) = n_{B0} \frac{\left\{ \exp\left(\frac{eV_{BE}}{kT}\right) - 1 \right\} (x_B - x) - x}{x_B}$$

Die Elektronen-Überschusskonzentration in der Basis ist näherungsweise eine lineare Funktion des Ortes.



Der Sinus-Hyperbolicus und seine lineare Approximation.

# Emitter-Region

- $x_E$  "endlich", d.h.  $x_E \gg L_E$  gilt i.a. nicht!
- Die Oberflächenrekombinationsgeschwindigkeit bei  $x' = x_E$  sei unendlich  $\rightarrow p_E(x_E) = p_{E0}$  (idealer Ohmscher Kontakt).

Diffusionsgleichung für Löcher im Emmitter:

$$\frac{\partial^2 \delta p_E}{\partial x'^2} = \frac{\delta p_E}{D_E \tau_{E0}} = \frac{\delta p_E}{L_E^2}$$

Allgemeine Lösung:  $\delta p_E(x') = p_E(x') - p_{E0} = C \exp\left(\frac{x'}{L_E}\right) + D \exp\left(-\frac{x'}{L_E}\right)$

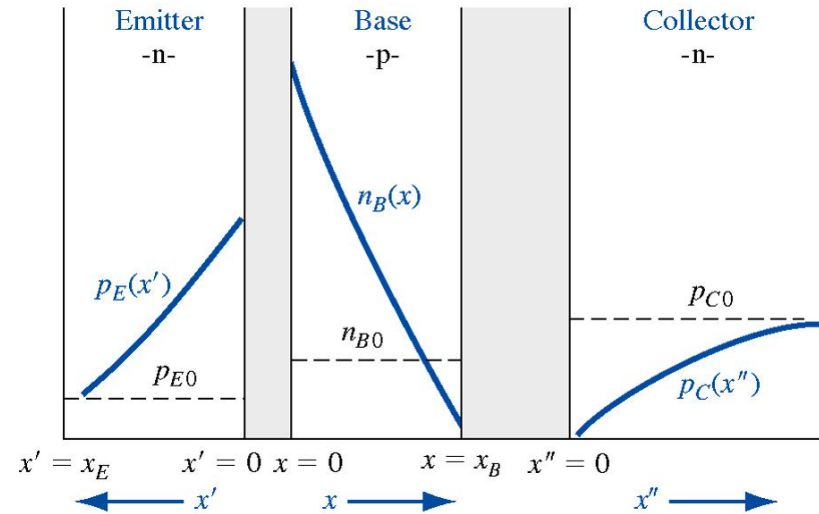
Zwei Randbedingungen:

$$\delta p_E(0) = p_E(0) - p_{E0} = p_{E0} \left\{ \exp\left(\frac{eV_{BE}}{kT}\right) - 1 \right\}$$

$$\delta p_E(x_E) = 0$$

Spezielle Lösung:

$$\delta p_E(x') = p_{E0} \frac{\left\{ \exp\left(\frac{eV_{BE}}{kT}\right) - 1 \right\} \sinh[(x_E - x')/L_E]}{\sinh[x_E/L_E]}$$



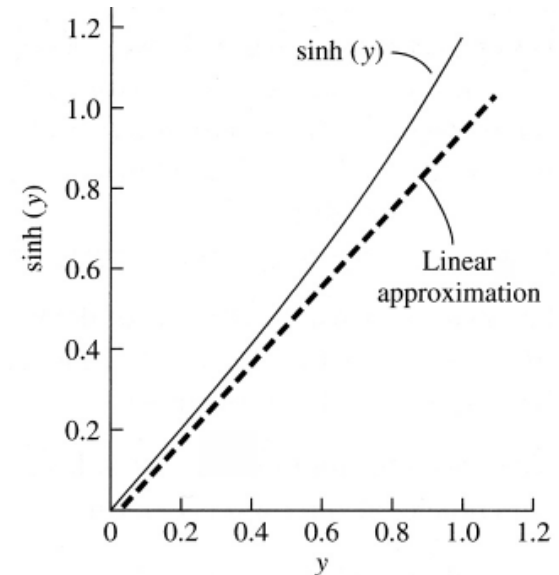
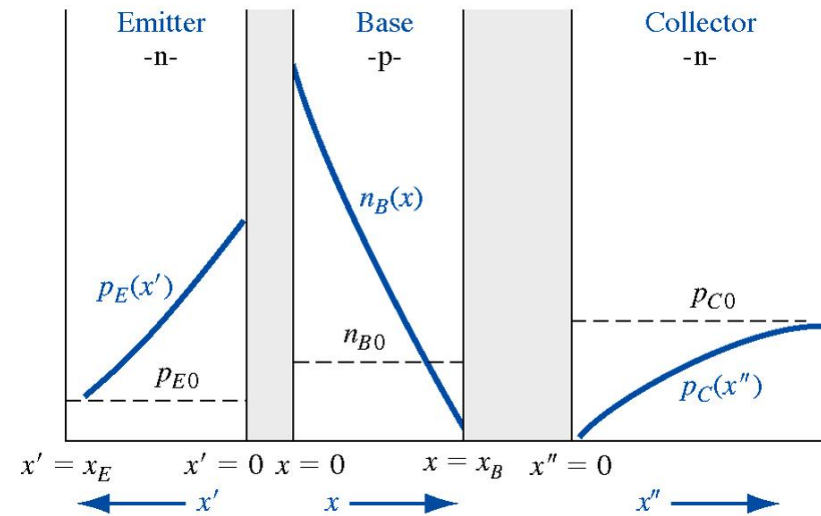
# Emitter-Region (2)

- Oft gilt  $x_E \ll L_E$ .

Entwickelt man die  $\sinh(*)$  linear, so folgt:

$$\delta p_E(x') = p_{E0} \frac{\left\{ \exp\left(\frac{eV_{BE}}{kT}\right) - 1 \right\} (x_E - x')}{x_E}$$

Die Löcher-Überschusskonzentration im Emitter ist dann näherungsweise eine lineare Funktion des Ortes.



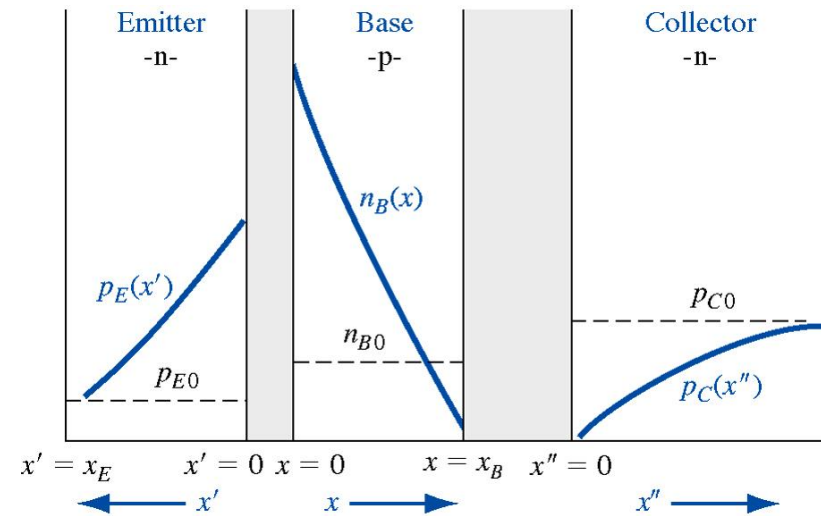
Der Sinus-Hyperbolicus und seine lineare Approximation.

# Kollektor-Region

- Annahme:  $x_C \gg L_C$

Diffusionsgleichung für Löcher im Kollektor:

$$\frac{\partial^2 \delta p_C}{\partial x''^2} = \frac{\delta p_C}{D_C \tau_{C0}} = \frac{\delta p_C}{L_C^2}$$



Allgemeine Lösung:  $\delta p_C(x'') = p_C(x'') - p_{C0} = G \exp\left[\frac{x''}{L_C}\right] + H \exp\left[-\frac{x''}{L_C}\right]$

Zwei Randbedingungen:

$$\delta p_C(0) = p_C(0) - p_{C0} = 0 - p_{C0} = -p_{C0}$$

$$\delta p_C(x_C) = 0$$

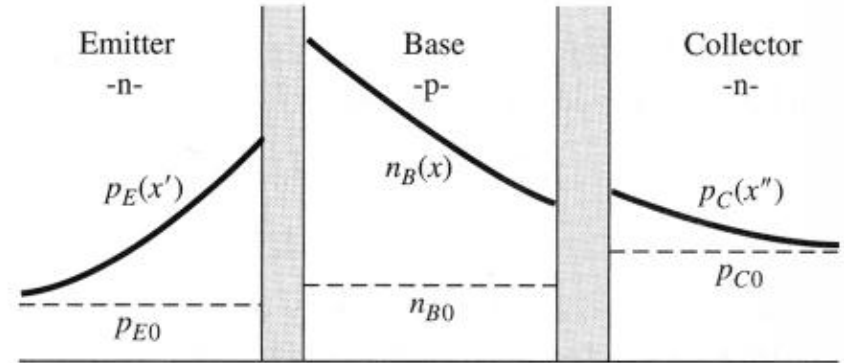
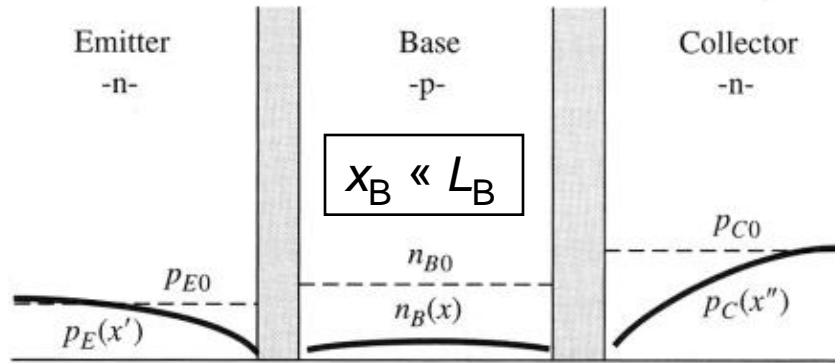
Spezielle Lösung:

$$\delta p_C(x'') = -p_{C0} \exp\left[-\frac{x''}{L_C}\right] \quad \text{wie beim gesperrten pn-Übergang}$$

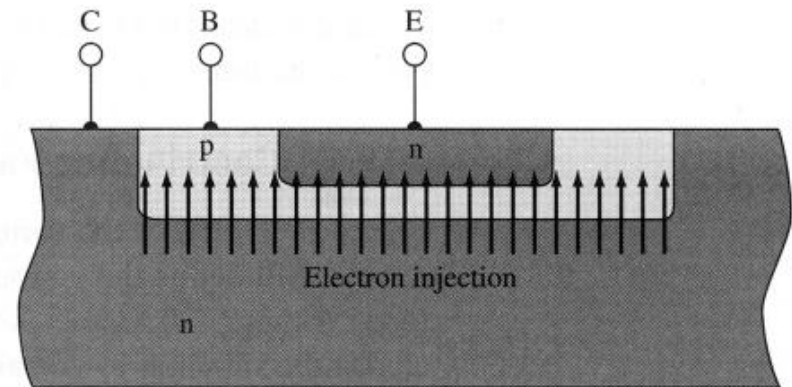
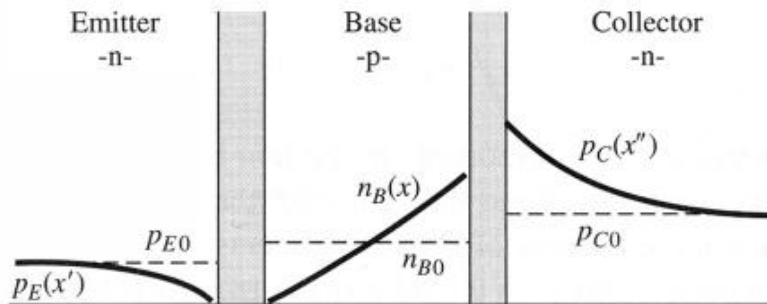
# Minoritätsladungsträger-Dichteprofile in anderen Operationsmodi

**cutoff**

**saturation**



**inverse-active**



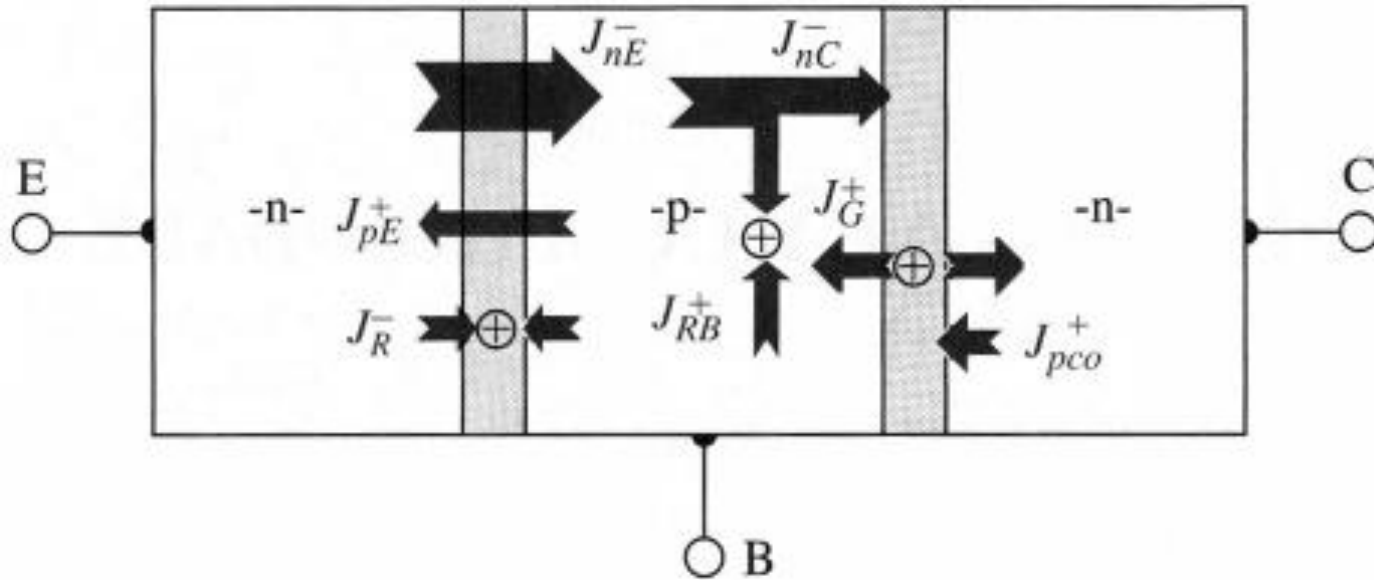
(a)

(b)

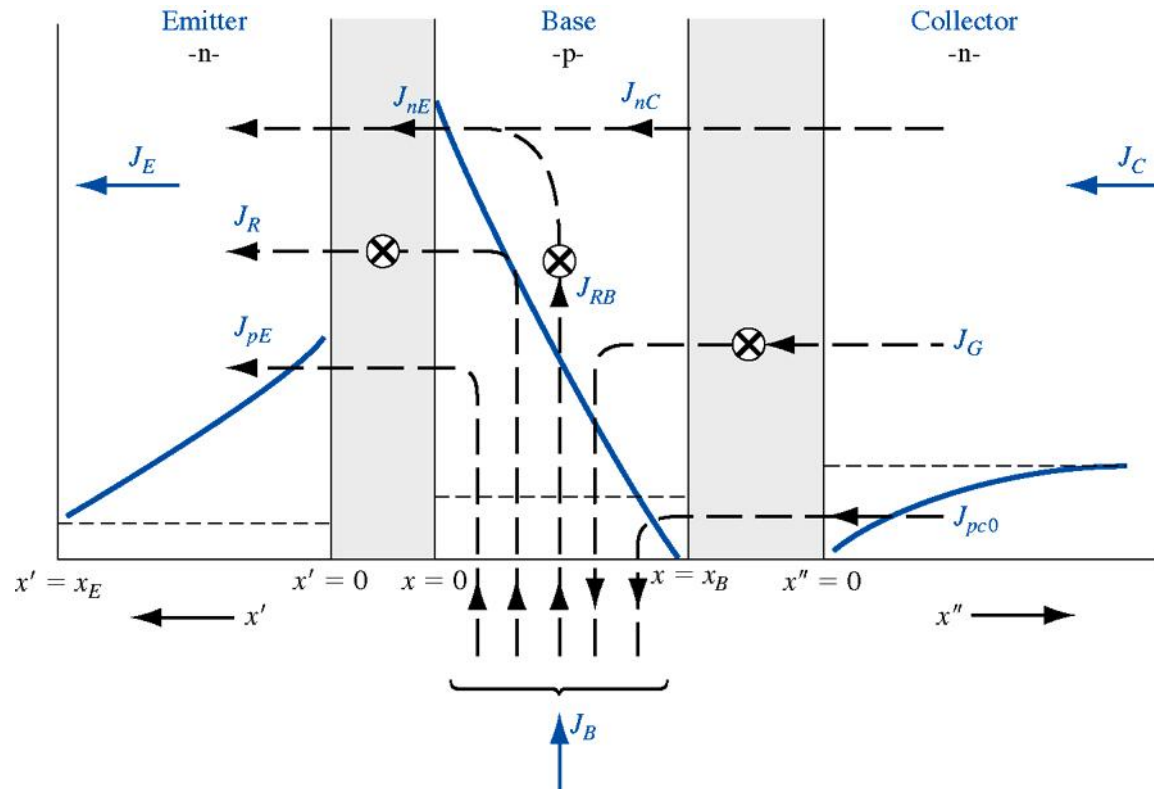
- (a) Minoritätsladungsträgerverteilung in einem npn-Bipolar-Transistor, der im invers-aktiven Modus arbeitet.  
 (b) Im Querschnitt ist die Injektion und das Aufsammeln der Elektronen im invers-aktiven Modus dargestellt.

# 9.3 Stromverstärkung in Basisschaltung

## Strom-Beiträge im Bipolar-Transistor

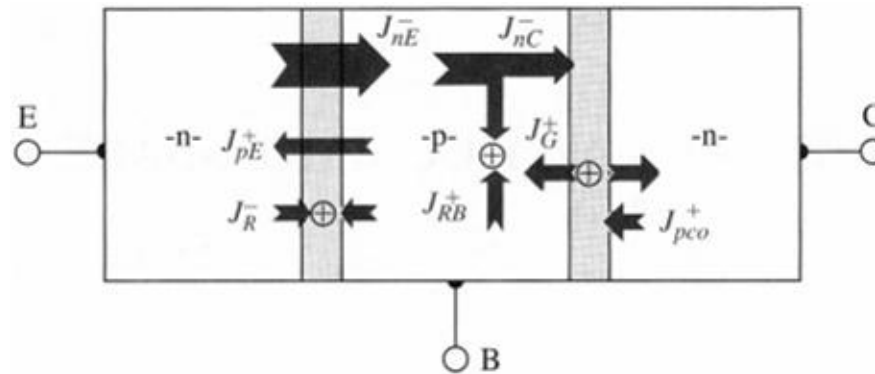


Teilchenstromdichten (oder Fluss-Komponenten) in einem npn-Bipolar-Transistor, der im vorwärts-aktiven Modus arbeitet.



- $J_{nE}$ : Diffusion von Minoritätselektronen in der Basis bei  $x = 0$ .
- $J_{nC}$ : Diffusion von Minoritätselektronen in der Basis bei  $x = x_B$ ;
- $J_{RB}$ : Differenz zwischen  $J_{nE}$  und  $J_{nC}$  wegen Rekombination von Minoritätselektronen mit Majoritätslöchern in der Basis.  $J_{RB}$  = Strom in die Basis, der die rekombinierten Löcher ersetzt.
- $J_{pE}$ : Diffusion von Minoritätslöchern im Emitter bei  $x' = 0$ .
- $J_R$ : Rekombination von Ladungsträgern im vorwärtsgepolten Basis-Emitter-Übergang.
- $J_{pc0}$ : Diffusion von Minoritätslöchern im Kollektor bei  $x'' = 0$ .
- $J_G$ : Generation von Ladungsträgern im gesperrten Basis-Kollektor-Übergang.





- Stromverstärkung in Basisschaltung:  $\alpha_0 = I_C/I_E$
- aktive Querschnittsfläche sei für Kollektor und Emitter gleich  $\rightarrow$

$$\alpha_0 = \frac{j_C}{j_E} = \frac{j_{nC} + j_G + j_{pC0}}{j_{nE} + j_R + j_{pE}}$$

- Kleinsignal-Stromverstärkung in Basisschaltung:  $\alpha = \partial j_C / \partial j_E$

$$\alpha = \frac{\partial j_C}{\partial j_E} = \frac{j_{nC}}{j_{nE} + j_R + j_{pE}}$$

- die Sperrströme  $j_G$  und  $j_{pC0}$  des gesperrten Basis-Kollektor-Übergangs sind keine Funktionen des Emitterstroms.

- Kleinsignal-Stromverstärkung in Basisschaltung wird umgeschrieben:

$$\alpha = \frac{\partial j_C}{\partial j_E} = \frac{j_{nC}}{j_{nE} + j_R + j_{pE}} = \left( \frac{j_{nE}}{j_{nE} + j_{pE}} \right) \left( \frac{j_{nC}}{j_{nE}} \right) \left( \frac{j_{nE} + j_{pE}}{j_{nE} + j_R + j_{pE}} \right)$$

$$\alpha = \gamma \alpha_T \delta$$

$$\gamma = \frac{j_{nE}}{j_{nE} + j_{pE}} \quad \text{Emitter-Injektionseffizienzfaktor (Emitter-Effizienz)}$$

$$\alpha_T = \frac{j_{nC}}{j_{nE}} \quad \text{Basis-Transportfaktor}$$

$$\delta = \frac{j_{nE} + j_{pE}}{j_{nE} + j_R + j_{pE}} \quad \text{Rekombinationsfaktor}$$

## Emitter-Effizienz

$$\gamma = \frac{j_{nE}}{j_{nE} + j_{pE}}$$

- Diffusionsströme an den Rändern der E-B-RLZ

$$j_{nE} = -eD_B \left. \frac{\partial(\delta n_B(x))}{\partial x} \right|_{x=0}$$

$$j_{pE} = -eD_E \left. \frac{\partial(\delta p_E(x'))}{\partial x'} \right|_{x'=0}$$

- Benutzen die linearen Näherungen, d.h. setzen voraus, dass  $x_B \ll L_B$  und  $x_E \ll L_E$ .
- Die E-B-Spannung sei gross gegen  $kT$ .
- Die Überschuss-Minoritätsdichten an den Rändern der E-B-RLZ vereinfachen sich zu

$$\delta n_B(x) = \frac{n_{B0}}{x_B} \left\{ \exp\left(\frac{eV_{BE}}{kT}\right) (x_B - x) - x \right\} \quad \delta p_E(x') = \frac{p_{E0}}{x_E} \left\{ \exp\left(\frac{eV_{BE}}{kT}\right) (x_E - x') \right\}$$

- Emitter-Effizienz:

$$\gamma = \frac{1}{1 + \frac{D_E}{D_B} \frac{p_{E0}}{n_{B0}} \frac{x_B}{x_E}} = \frac{1}{1 + \frac{D_E}{D_B} \frac{N_B}{N_E} \frac{x_B}{x_E}}$$

- Für  $\gamma \approx 1$  muss  $N_E \gg N_B$  sein!

# Basis-Transportfaktor

$$\alpha_T = \frac{j_{nC}}{j_{nE}}$$

- Diffusionsströme an den Rändern der Basis

$$j_{nC} = -eD_B \left. \frac{\partial(\delta n_B(x))}{\partial x} \right|_{x=x_B}$$
$$j_{nE} = -eD_B \left. \frac{\partial(\delta n_B(x))}{\partial x} \right|_{x=0}$$

- Die E-B-Spannung sei gross gegen  $kT$ .
- Würde man hier die lineare Näherung benutzen, dann ergäbe sich sofort  $\alpha_T = 1$ .

$$\delta n_B(x) = n_{B0} \frac{\exp\left(\frac{eV_{BE}}{kT}\right) \sinh[(x_B - x)/L_B] - \sinh[x/L_B]}{\sinh[x_B/L_B]}$$

- Basis-Transportfaktor:

$$\alpha_T \approx \frac{1}{\cosh[x_B/L_B]} \approx \frac{1}{1 + \frac{1}{2}(x_B/L_B)^2}$$

# Rekombinationsfaktor

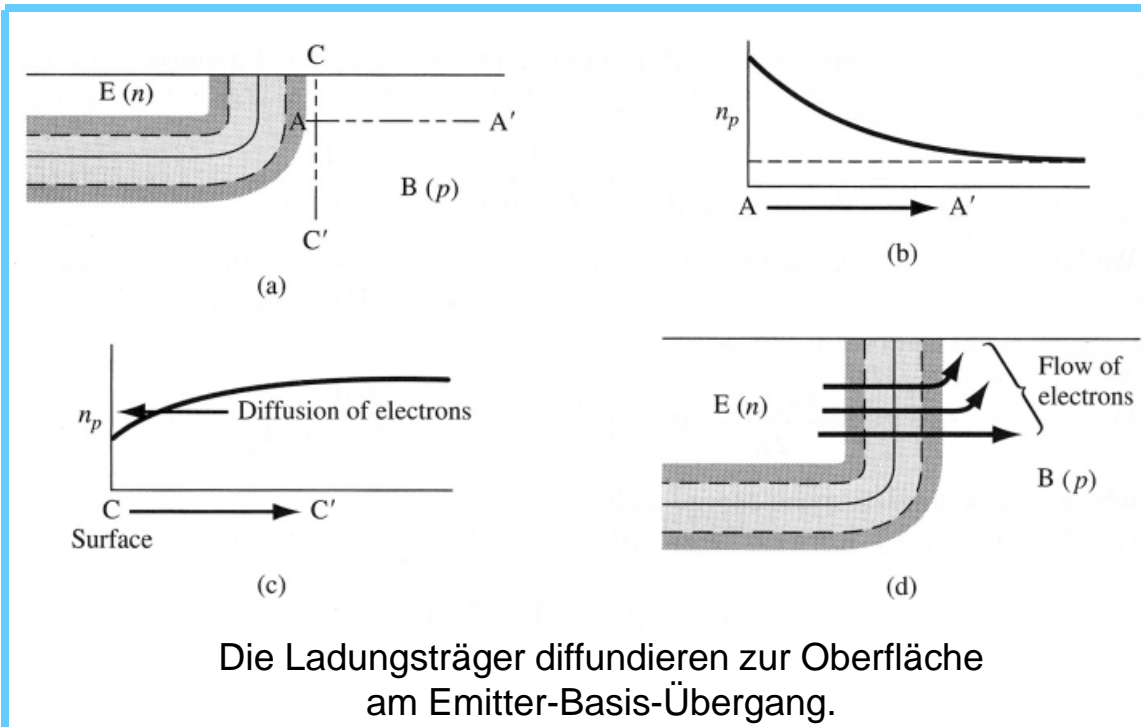
$$\delta = \frac{j_{nE} + j_{pE}}{j_{nE} + j_R + j_{pE}} \approx \frac{j_{nE}}{j_{nE} + j_R} = \frac{1}{1 + j_R/j_{nE}}$$

- Rekombinationsstrom und Diffusionsstrom sind gegeben durch (siehe vorwärts-gepolte Diode)

$$j_R = j_{r0} \exp\left(\frac{eV_{BE}}{2kT}\right) \quad j_{nE} = j_{s0} \exp\left(\frac{eV_{BE}}{kT}\right)$$

- Rekombinationsfaktor:

$$\delta = \frac{1}{1 + \frac{j_{r0}}{j_{s0}} \exp\left(\frac{-eV_{BE}}{2kT}\right)}$$



# 10. Der Bipolar-Transistor (2)

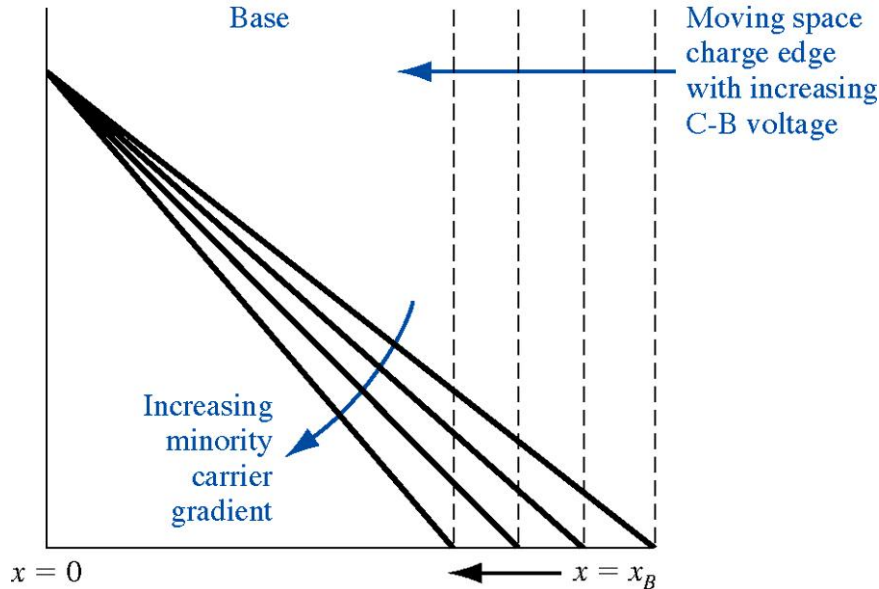
**10.1 Abweichungen vom idealen Verhalten (N10.4)**

**10.2 Frequenz-Verhalten (N10.6)**

**10.3 Moderne Bipolar-Transistoren\* (N10.8)**

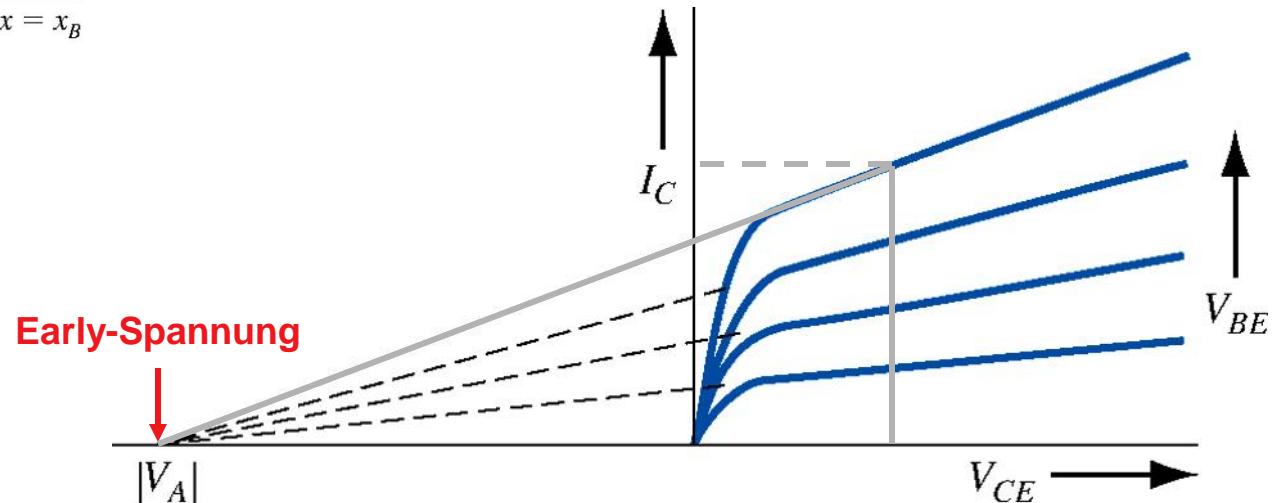
# 10.1 Abweichung vom idealen Verhalten

## Basisweiten-Modulation (Early-Effekt)



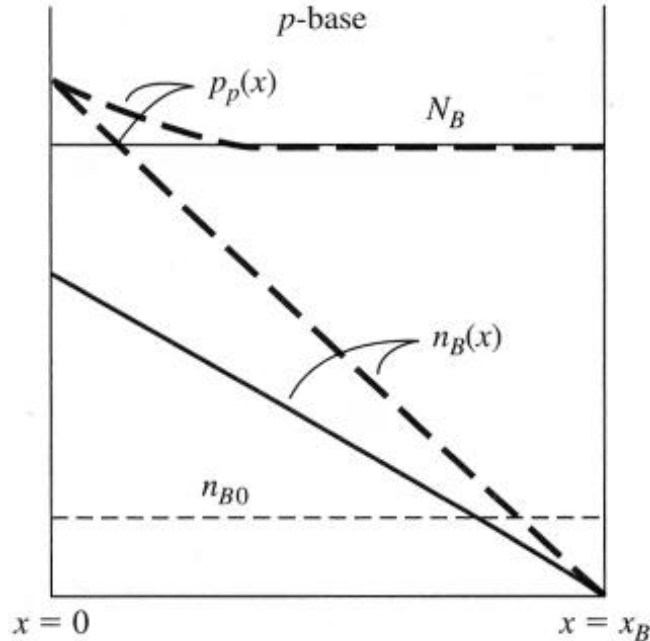
- RLZ-Weite des B-C-Übergangs weitet sich mit steigender Kollektor-Spannung.
- Dadurch wächst der Gradient der Minoritätslöcherdichte in der Basis → der Diffusionsstrom steigt an.
- Der Ausgangsleitwert, der idealerweise Null ist, wird endlich.

$$g_0 = \frac{\partial I_C}{\partial V_{CE}} = \frac{I_C}{V_{CE} + V_A}$$

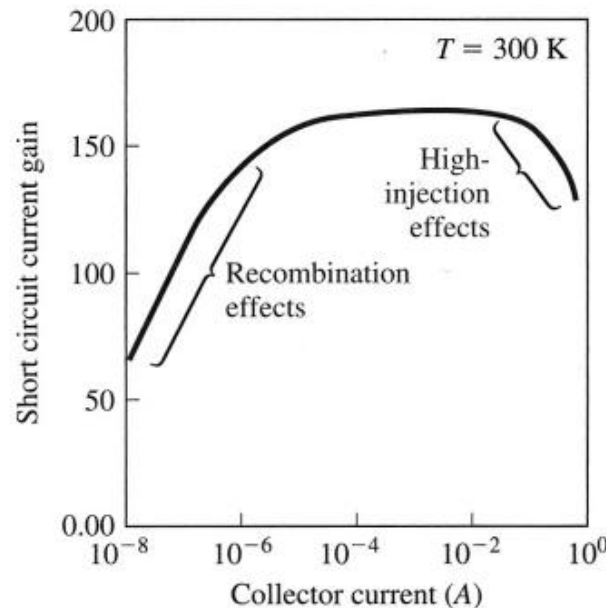


# Hochinjektion in der Basis

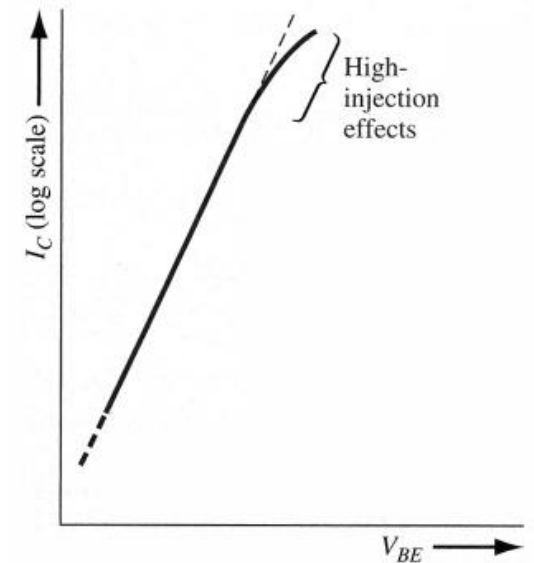
- Injizierte Löcherdichte übersteigt die Majoritätsdichte in der Basis = Hochinjektion
- 1.) Löcher-Rückinjektion aus der Basis in den Emitter steigt stark an  $\rightarrow J_{pE}$  steigt und verringert die Emitter-Effizienz
- 2.)  $n_p(0) \sim \exp[eV_{BE}/(2kT)]$ , da  $n_p(0) \sim p_p(0) \rightarrow j_C$  flacht ab



Konzentration der Minoritäts- und Majoritätsladungsträger in der Basis unter Niedriginjektion (durchgezogene Kurven) und unter Hochinjektion (gestrichelte Kurven).



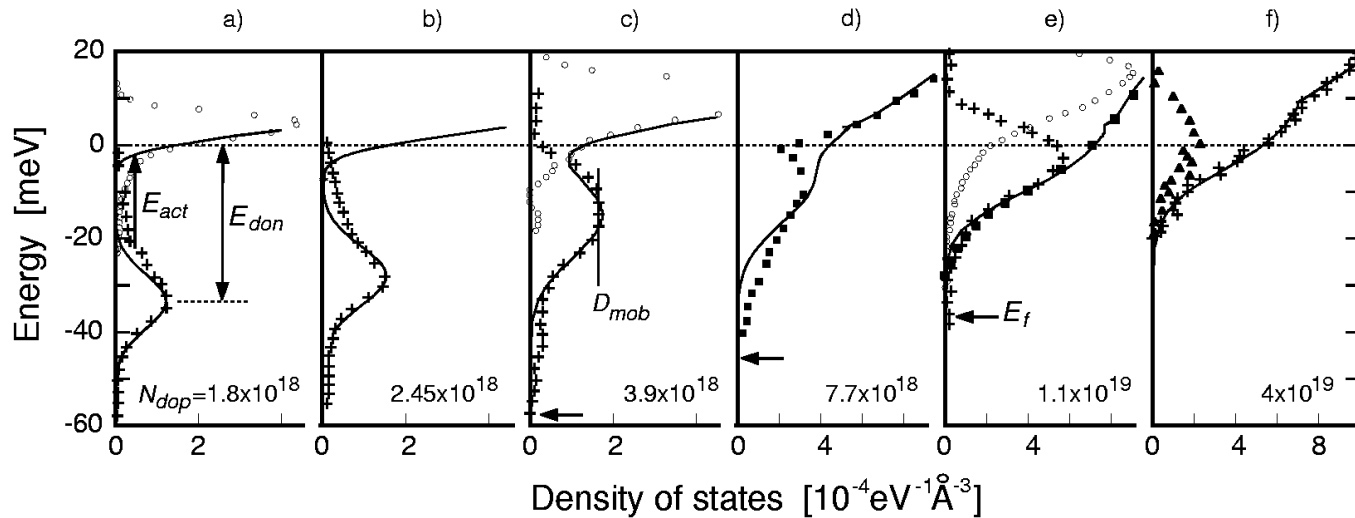
Gain als Funktion des Kollektorstroms in Emitterschaltung.



Hochinjektionseffekt im Kollektorstrom als Funktion der Basis-Emitter-Spannung.



# Band Gap Narrowing im Emitter



Band tailing: das Donator-niveau verbreitert sich zu einem schmalen Band, das mit dem Leitungsband verschmilzt. Es verbleibt ein Zustandsdichte-Ausläufer in der Energielücke.

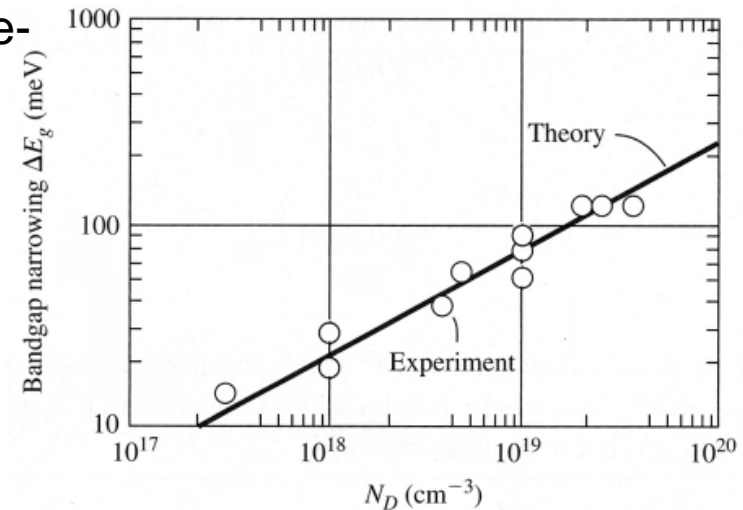
■ Bei hohen Dotierungen wird das Elektronengas “real” (starke e-e-WW). Die Bandkanten verschieben sich, und die Energielücke schrumpft.

■ Die Konsequenz ist eine höhere intrinsische Dichte im Emitter!

$$n_{iE} = (N_C N_V)^{1/2} \exp\left(-\frac{E_{g0} - \Delta E_g}{2kT}\right) = n_i \exp\left(\frac{\Delta E_g}{2kT}\right)$$

■ Die Stromverstärkung in Emitterschaltung wird kleiner

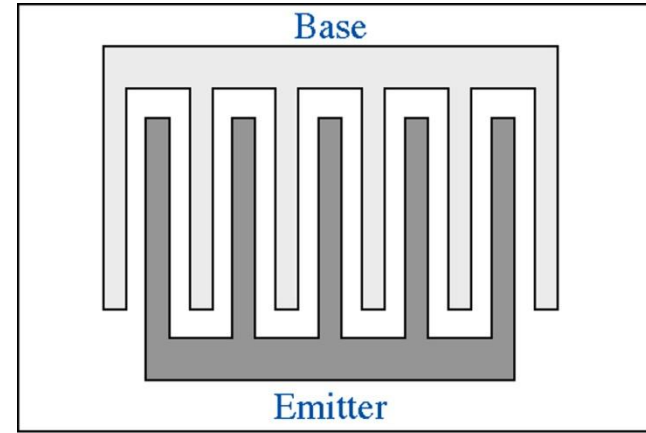
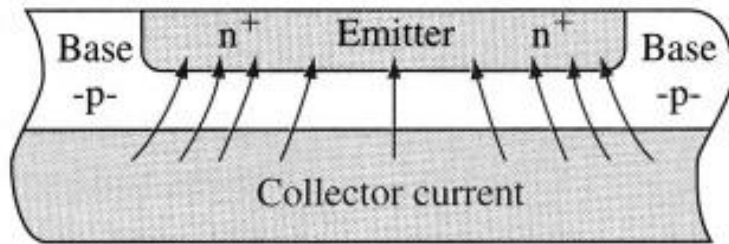
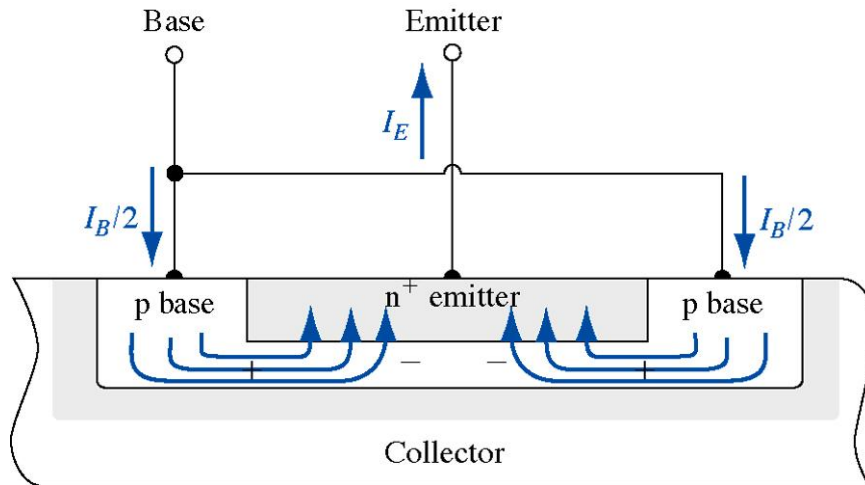
$$\beta \sim \frac{n_{B0}}{p_{E0}} \sim \exp\left(\frac{-\Delta E_g}{2kT}\right)$$



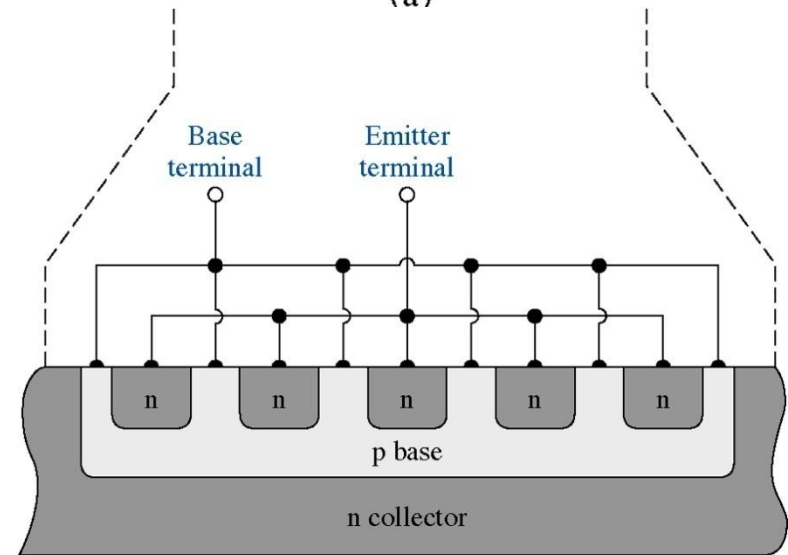
Bandgap-Narrowing als Funktion der Dotierung in Si (nach Sze).

# Current Crowding

- dünne Basis → lateraler Widerstand → Potentialabfall von den Kanten zur Mitte → current crowding im Kantenbereich



(a)

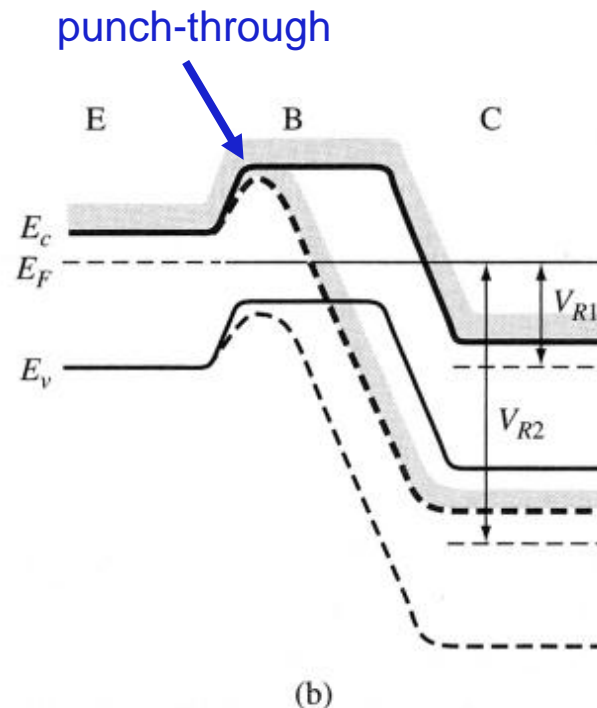
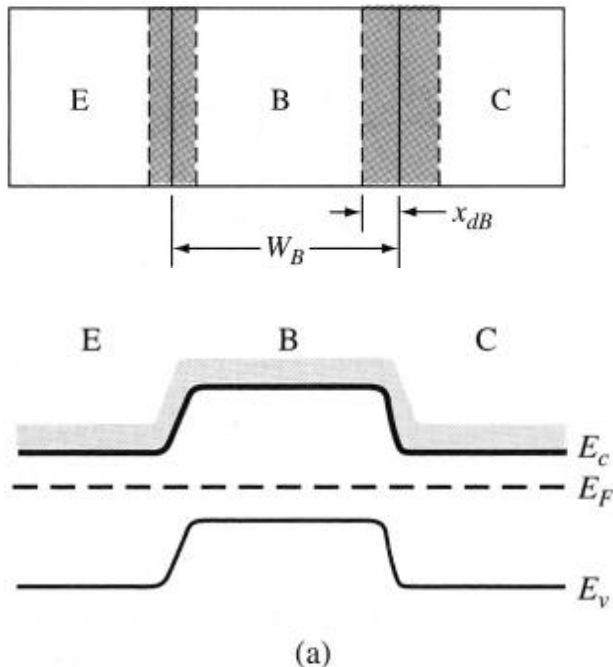


(b)

# Durchbruch-Spannung (punch-through)

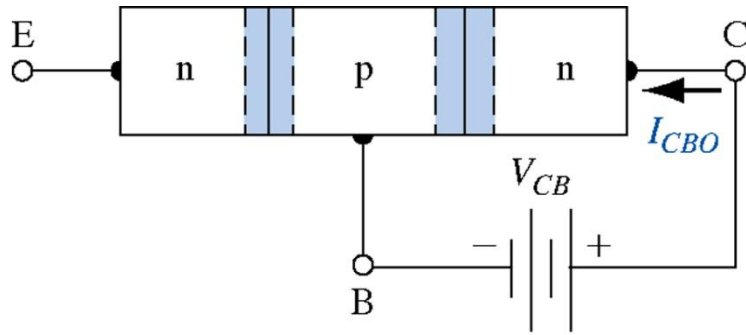
$$W_B = \left\{ \frac{2 \varepsilon (V_{bi} + V_{pt})}{e} \frac{N_C}{N_B (N_B + N_C)} \right\}^{1/2}$$

$$V_{pt} = \frac{e W_B^2}{2 \varepsilon} \frac{N_B (N_B + N_C)}{N_C}$$



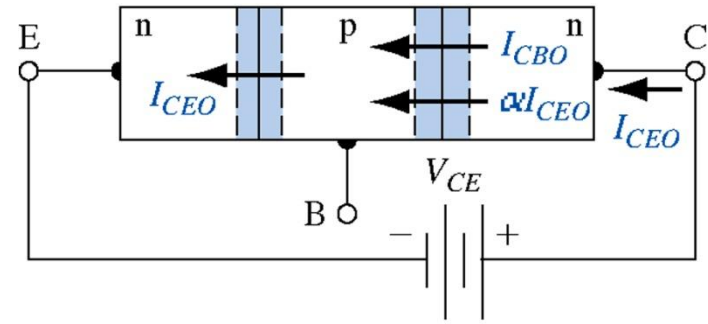
Bandkendiagramm eines npn-Bipolar-Transistors (a) im thermodynamischen Gleichgewicht und (b) unter Sperrspannung am B-C-Übergang vor dem punch-through ( $V_{R1}$ ) und nach dem punch-through ( $V_{R2}$ ).

# Durchbruch-Spannung (Avalanche)



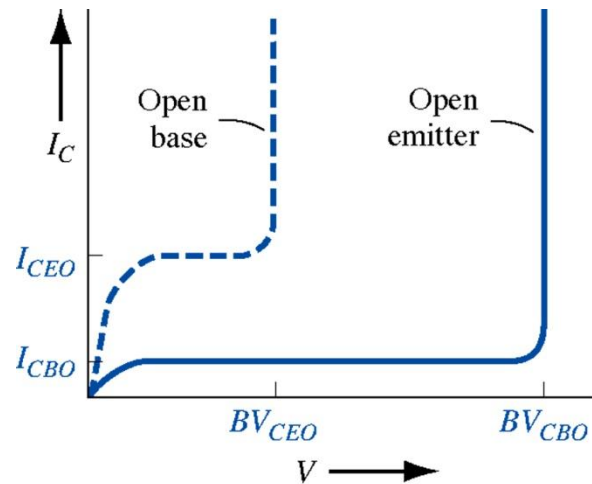
(a)

open-emitter configuration



(b)

open-base configuration



# 10.2 Frequenz-Verhalten

## Transit-Zeiten (Zeitverzögerungsfaktoren)

Totale Emitter-Kollektor Transit-Zeit:

$$\tau_{ec} = \tau_e + \tau_b + \tau_d + \tau_c$$

Kapazitäts-Aufladungszeit  
des EB-Übergangs

Basis-  
Transit-Zeit

Kollektor-RLZ-  
Transit-Zeit

Kollektor-Kapazitäts-  
Aufladungszeit

$$\tau_e = r_e' (C_{je} + C_p)$$

Emitter-  
Diffusions-  
widerstand

alle parasi-  
tären Kapa-  
zitäten

BE-  
Diffusions-  
kapazität

$$\tau_b = x_B^2 / (2D_n)$$

Neamen  
Seite 423

$$\tau_d = x_{dc} / v_s$$

BC-RLZ-  
Weite

Sättigungs-  
geschwin-  
digkeit

$$\tau_c = r_c (C_\mu + C_s)$$

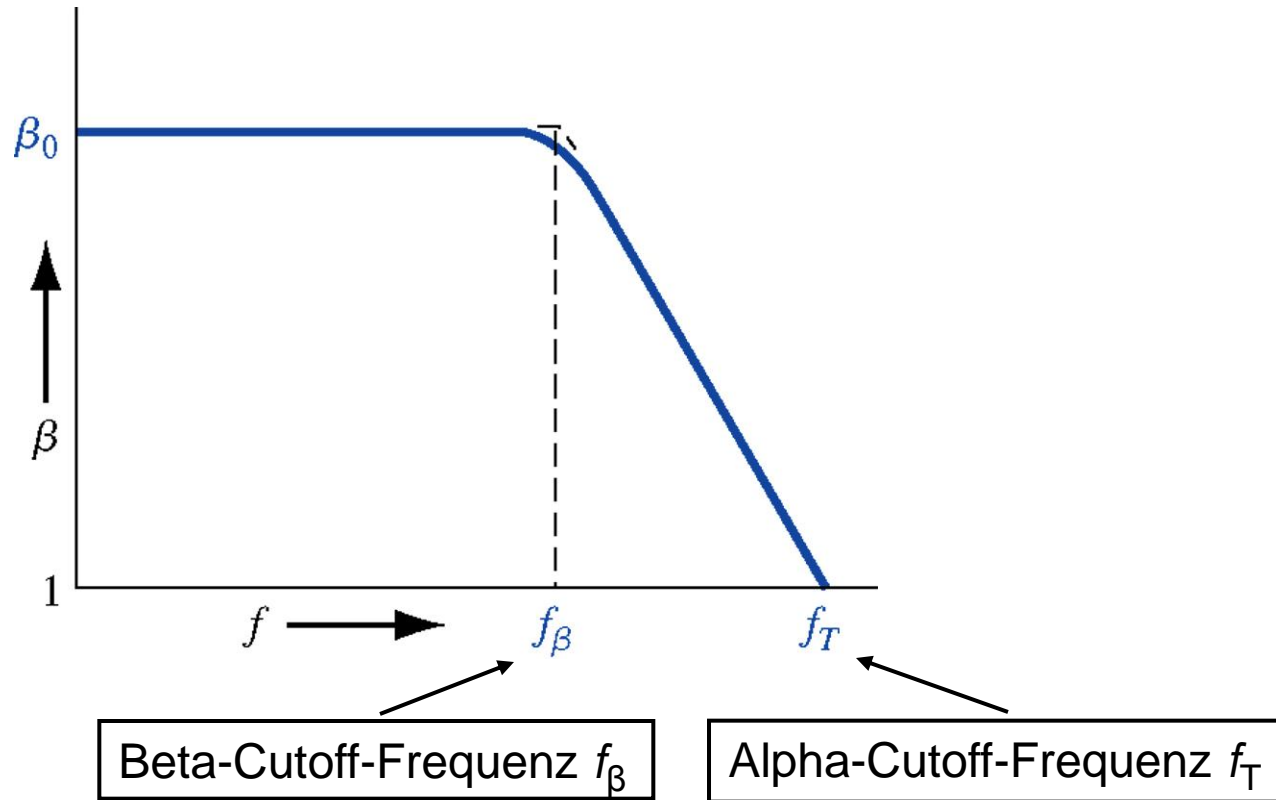
Kollektor-  
Reihen-  
widerstand

Kollektor-  
Substrat-  
Kapazität

BC-  
Diffusions-  
kapazität

$$r_e' = V_t / I_E$$

# Cutoff-Frequenzen

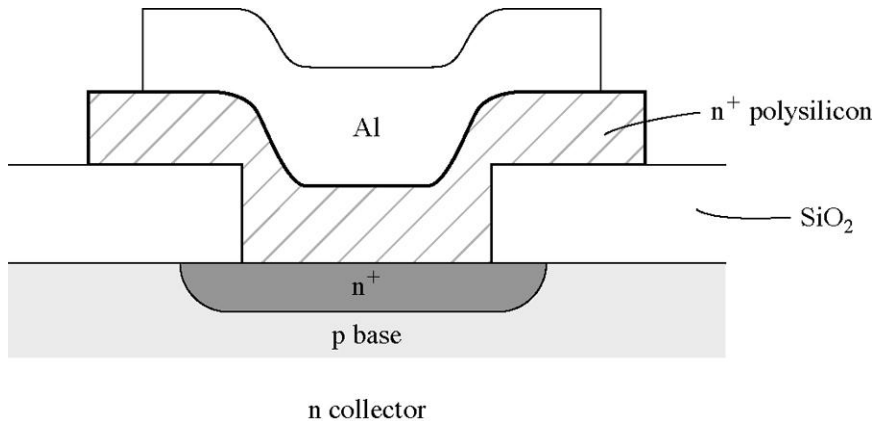


$$f_\beta = \frac{f_T}{\beta_0}$$

$$f_T = \frac{1}{2\pi\tau_{ec}}$$

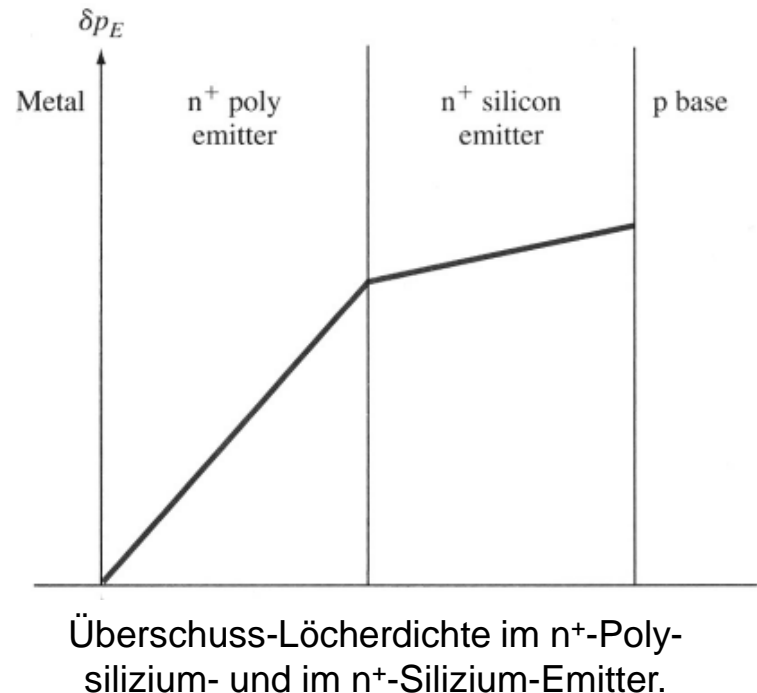
# 10.3 Moderne Bipolar-Transistoren\*

## Polysilizium-Emitter BJT



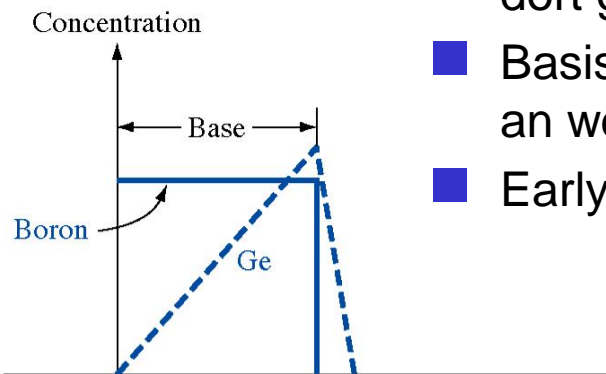
- Polysilizium hat kleine Beweglichkeit → kleine Diffusionslänge.
- Kontinuität der Stromdichte an der Grenze zwischen kristallinem Emitter und Poly-Emitter → Reduktion des Gradienten an der RLZ-Grenze des B-E-Übergangs.
- Reduktion der Löcher-Rückinjektion → bessere Emitter-Effizienz.

- Man benützt dünne Emitter → kleine  $r_e$  und  $\tau_e$ .
- Dadurch steigt der Gradient der Löcher-Minoritätsdichte → Degradation der Emitter-Effizienz.
- Lösung: Polysilizium Emitter.

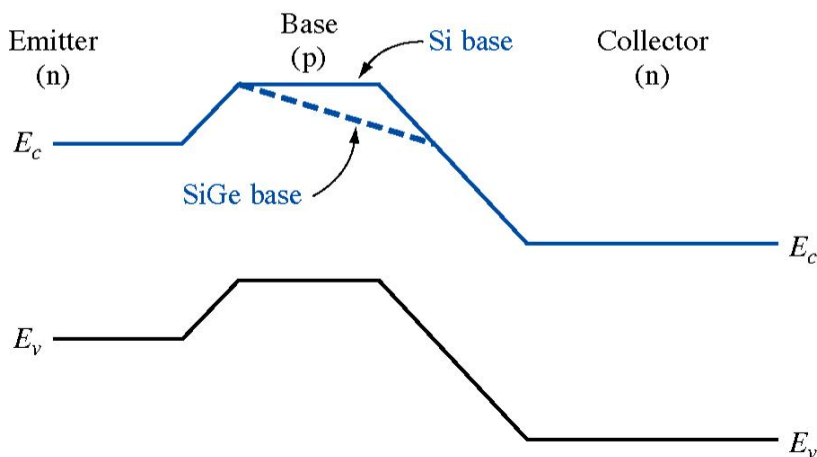


# Bipolar-Transistor mit SiGe-Basis

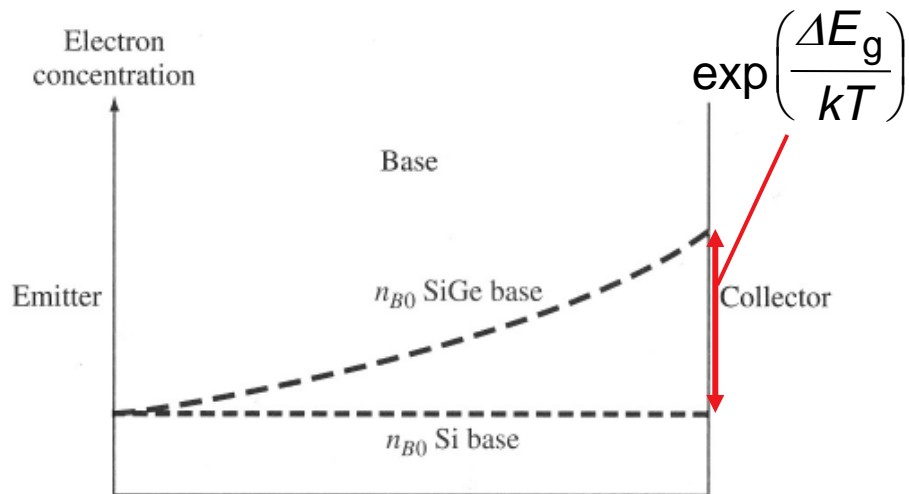
- Energielücke von Germanium  $\sim 0.67$  eV.
- Basis enthält Ge-Profil mit Maximum am B-C-Übergang, dort grösste Reduktion vom Gap.
- Basisstrom bleibt unverändert, nur Kollektorstrom steigt an wegen grösster Gleichgewichts-Löcherdichte bei  $x_B$ .
- Early-Spannung steigt, Basis-Transitzeit wird kleiner.



(a)



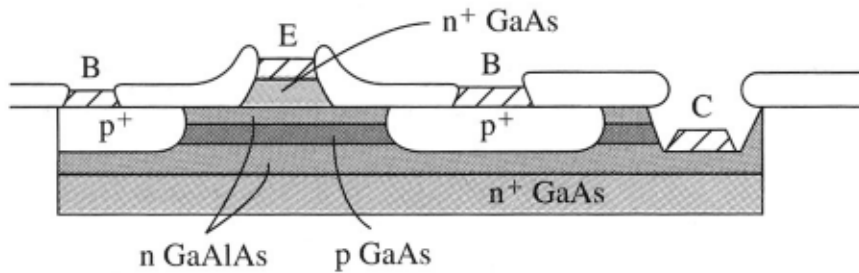
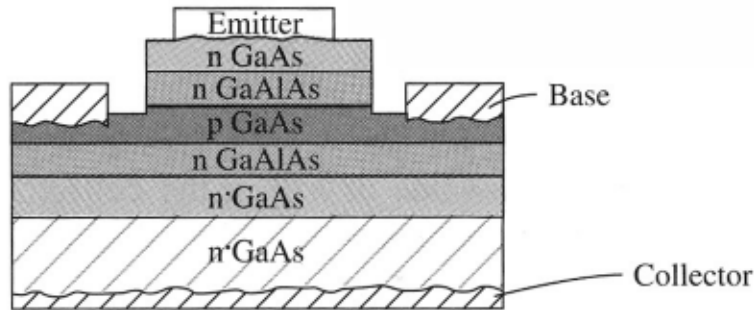
(b)



Minoritäts-Elektronenkonzentration in der Basis im thermodynamischen Gleichgewicht. Unterschied zwischen Si- und SiGe-Basis.

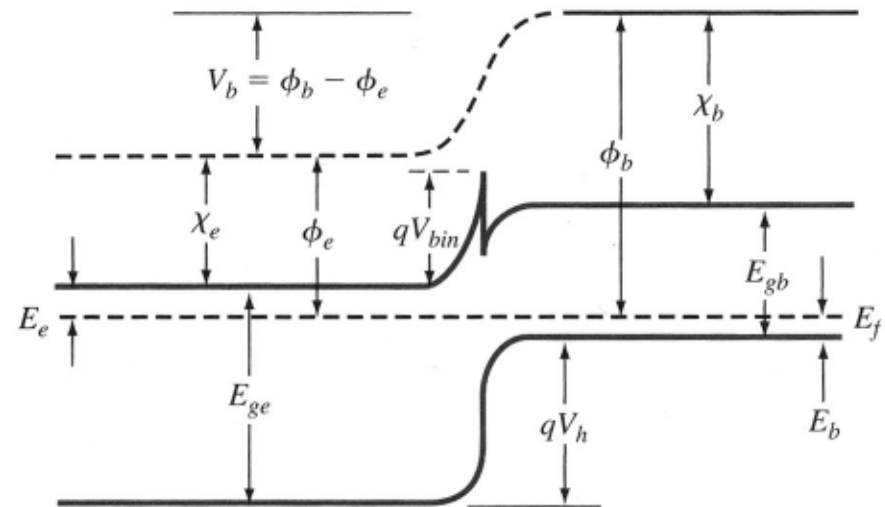


# Bipolar-Transistoren mit Hetero-Übergängen

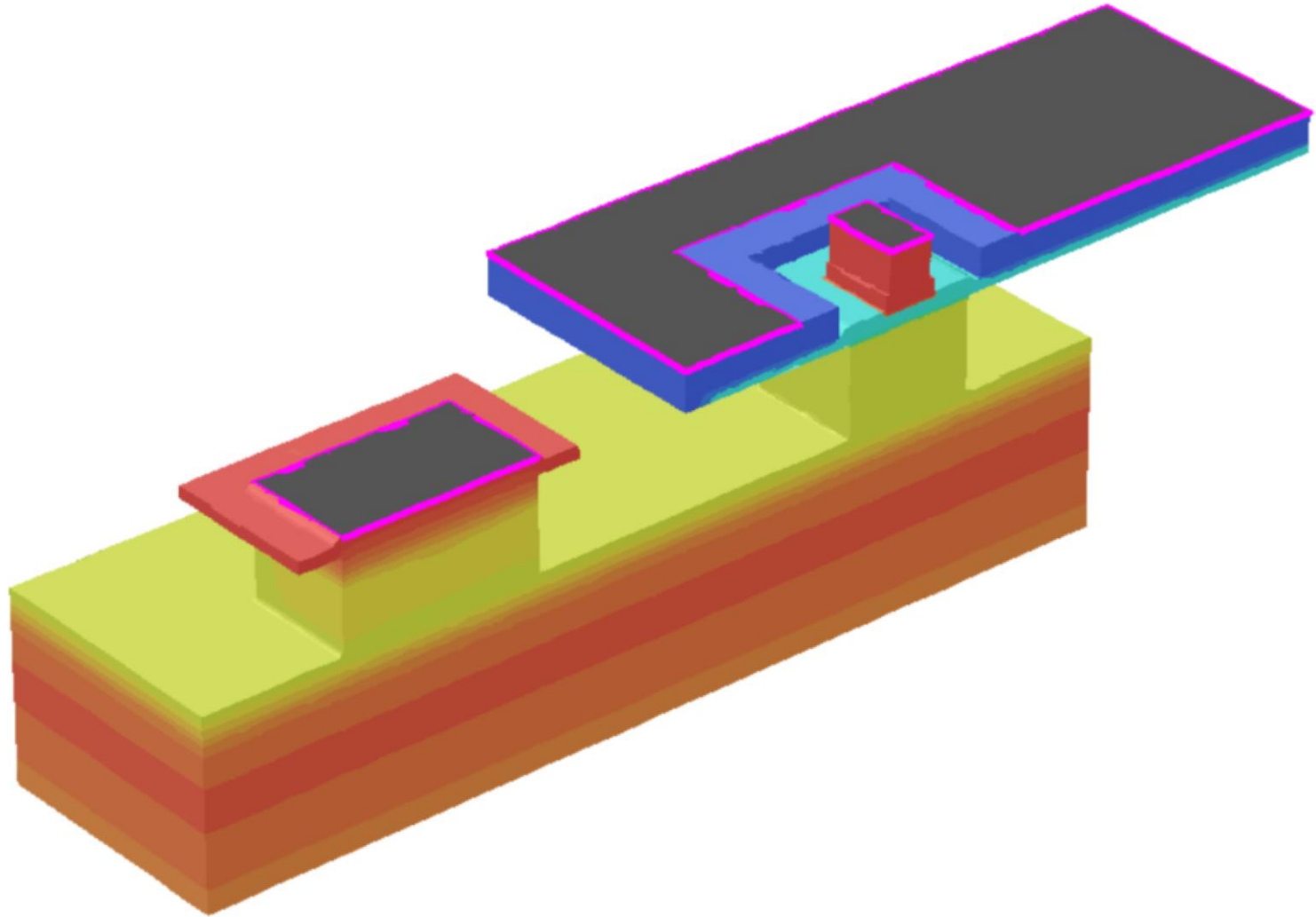


- Man benützt im Emitter ein Material mit grösserer Energielücke → kleinere Löcher-Minoritätsdichte → bessere Emitter-Effizienz.

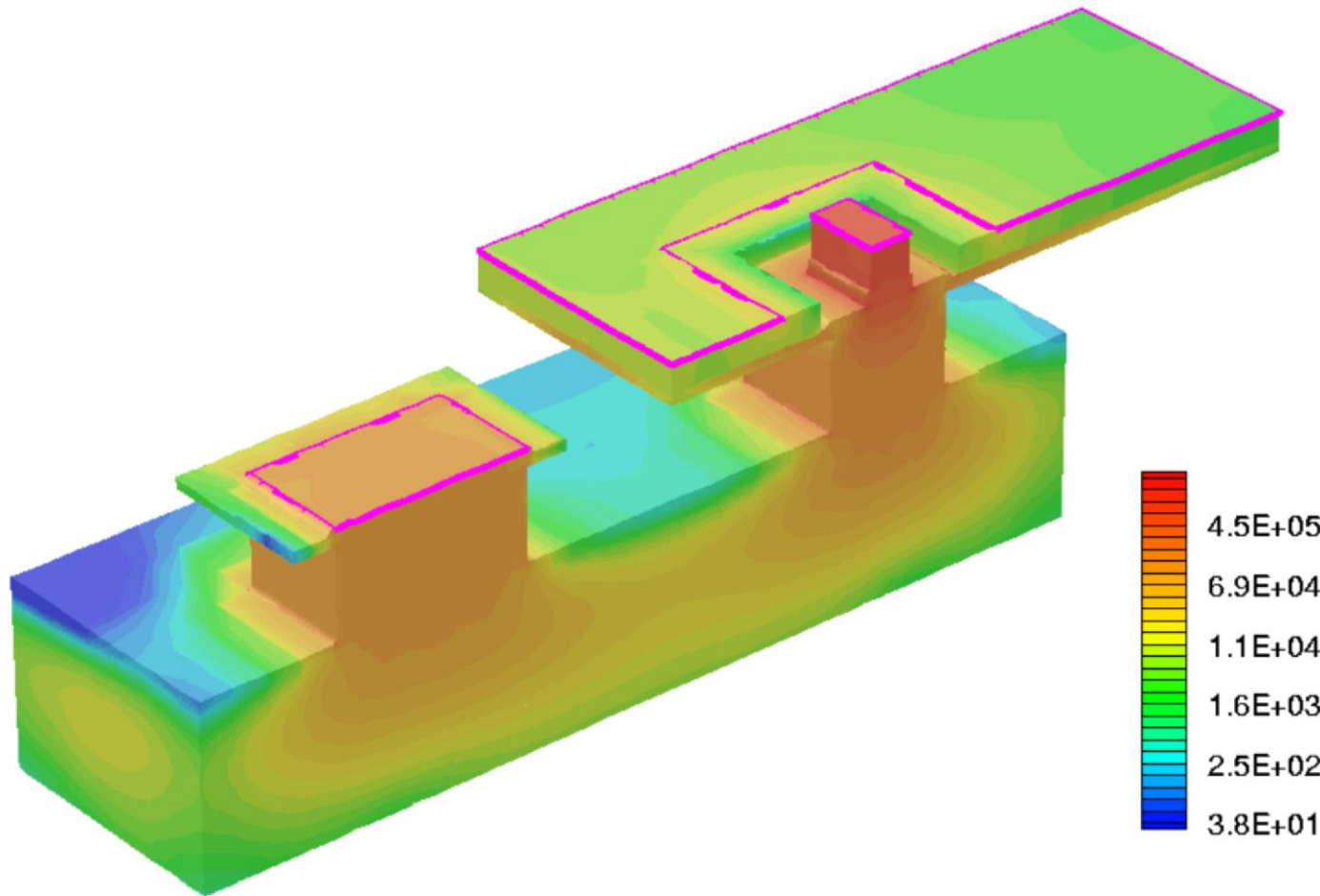
- Die höhere Potentialbarriere  $V_h$  reduziert die Dichte der injizierten Löcher um  $\exp(\Delta E_g/kT)$ .
- Emitter muss nicht mehr so hoch dotiert werden, weniger BGN.
- Kleinere Dotierung → kleinere B-E-Kapazität → schnellerer Transistor



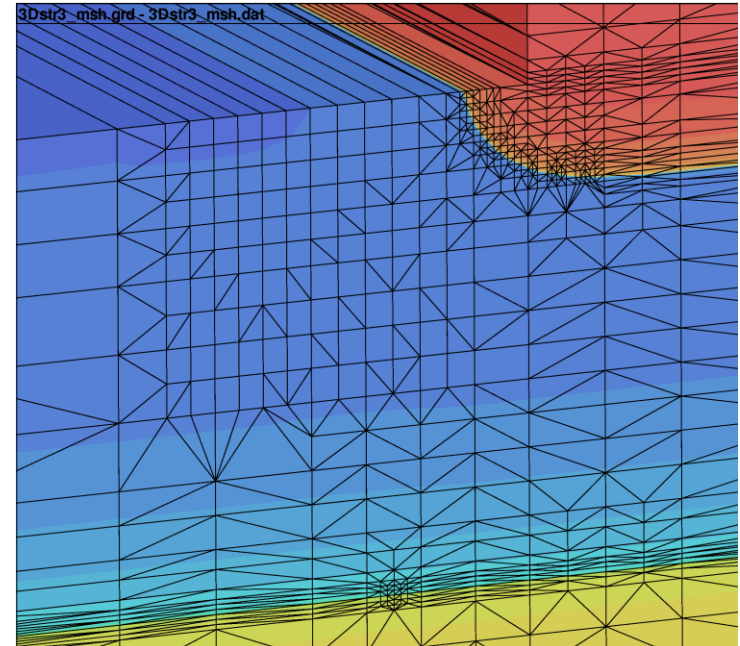
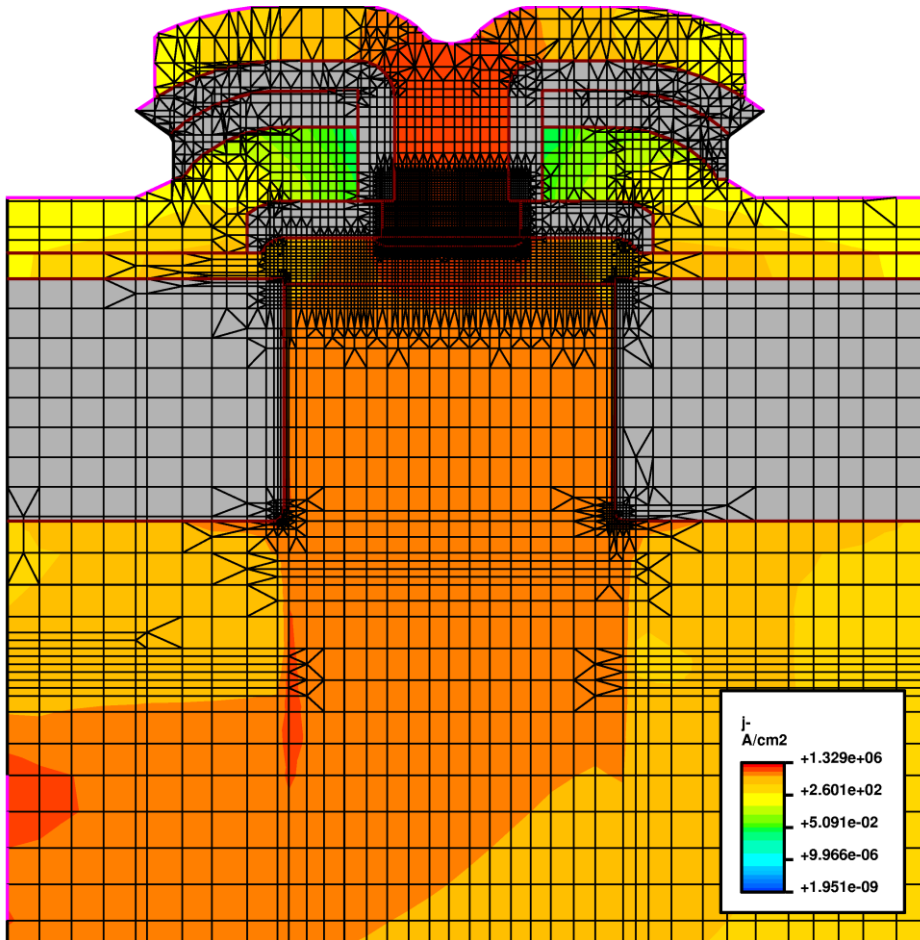
# Moderner Silizium-Bipolar-Transistor (Dotierung)



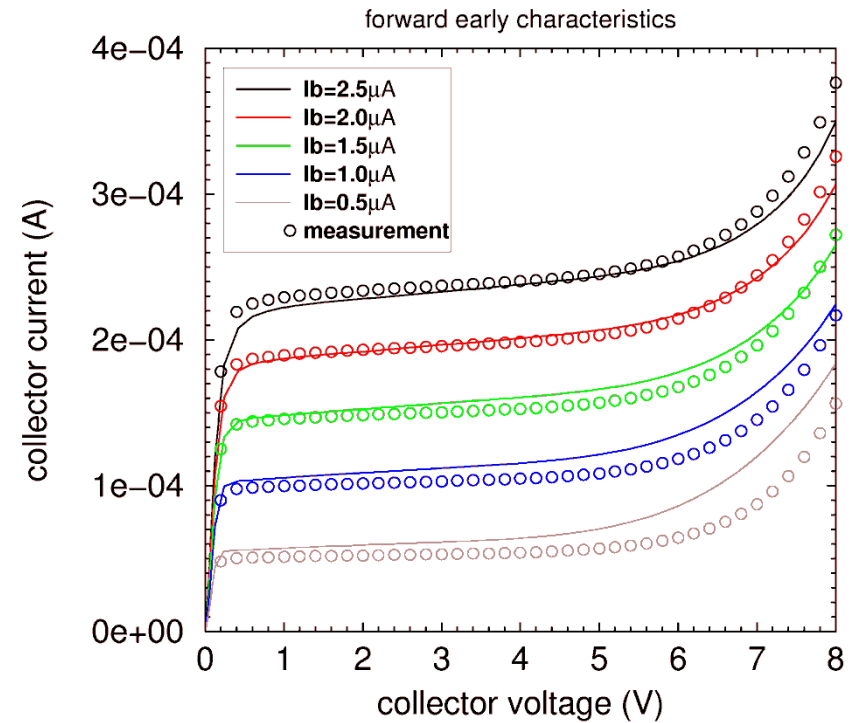
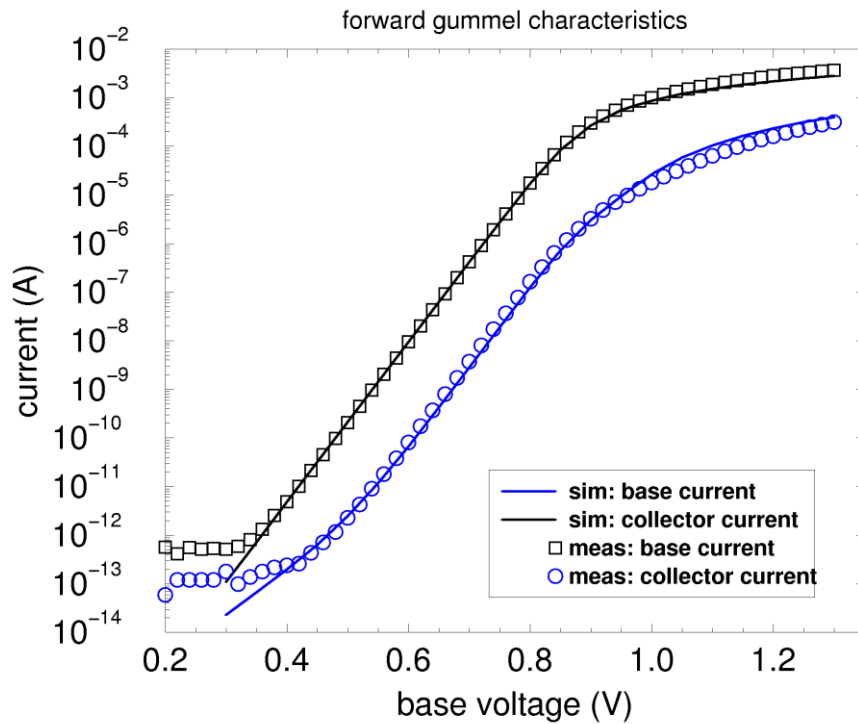
# Stromdichte-Verteilung @ $V_{BE} = 1.3V$ und $V_{CE} = 2V$



# Simulationsgitter (2D und 3D)



# Simulation von Gummel- und Early-Kennlinien



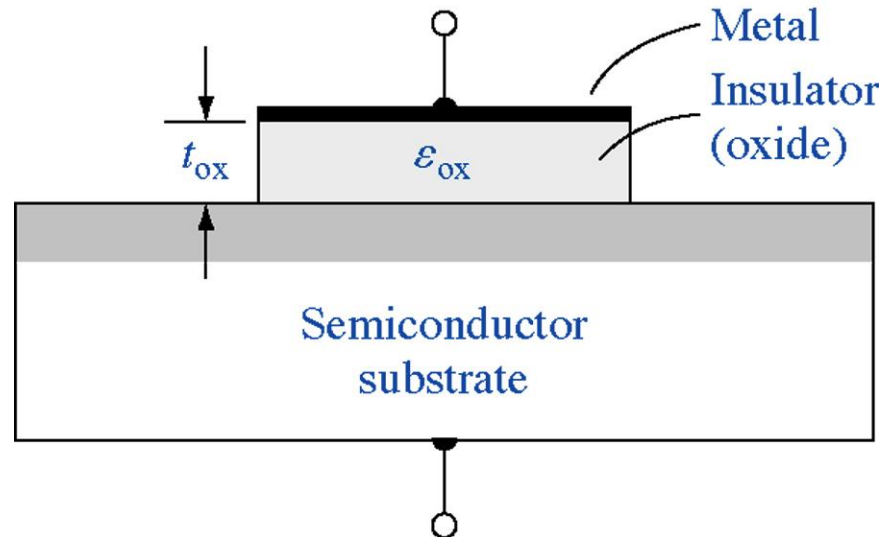
# 11. Die MOS-Diode

**11.1 Grundprinzipien der MOS-Struktur (N11.1)**

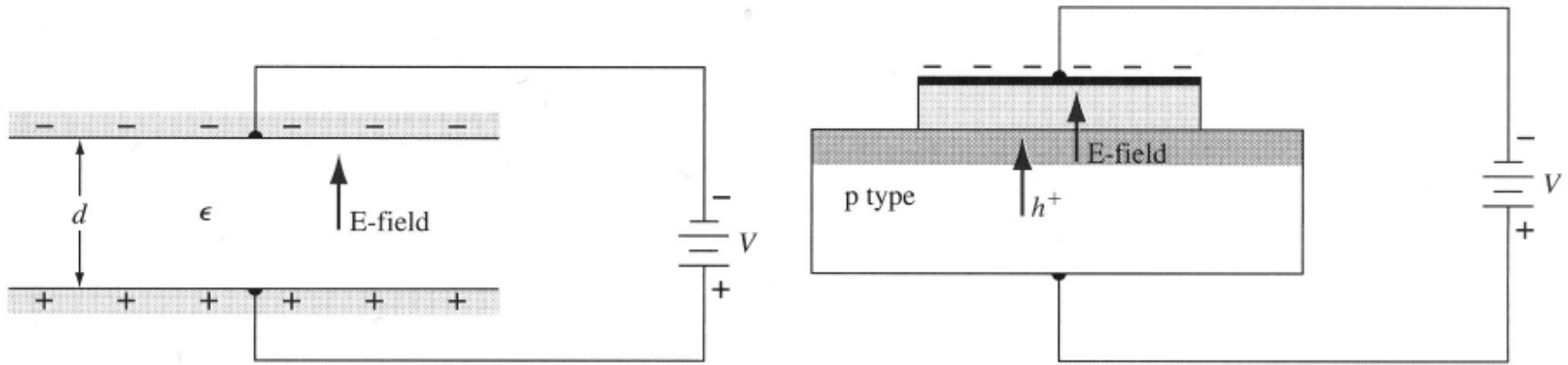
**11.2 CV-Kennlinien (N11.2)**

# 11.1 Grundprinzipien der MOS-Struktur

- Eine MOS Struktur wird gebildet, indem eine dünne  $\text{SiO}_2$ -Schicht auf einer Si Oberfläche (= Substrat) aufgewachsen wird. Ein Kontakt (= Gate - typischerweise aus hochdotiertem polykristallinen Si oder ein Metall) auf dem Oxid führt zu einer Kondensator-Struktur.
- Das Grundprinzip der Wirkungsweise des MOS-Kondensators beruht darauf, dass beim Anlegen einer Spannung mit richtiger Polarität am Gate im Silizium bewegliche Ladungsträger induziert werden. Man nennt den Bereich der induzierten Ladung *Kanal*.
- Eine MOSFET-Struktur entsteht, wenn man zu beiden Seiten des Kanals Kontaktbereiche definiert.
- Ein MOSFET ist im Gegensatz zum Bipolar-Transistor von Majoritätsladungsträgern dominiert.

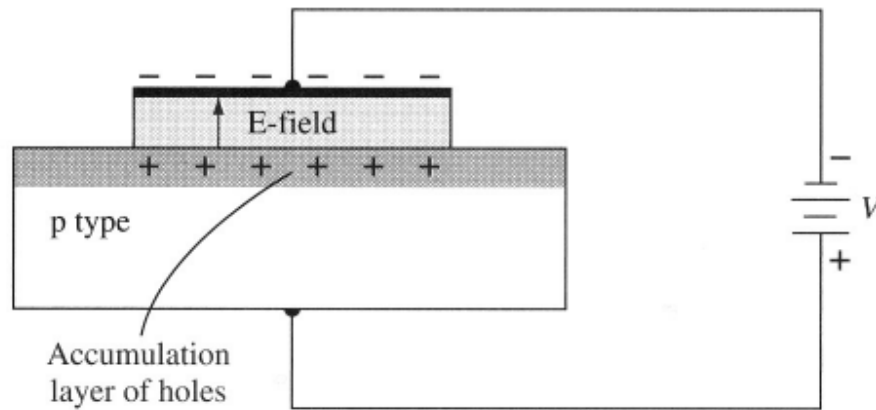


# Plattenkondensator-Modell



(a)

(b)



(c)

$$C' = \frac{\epsilon_{ox}}{d}$$

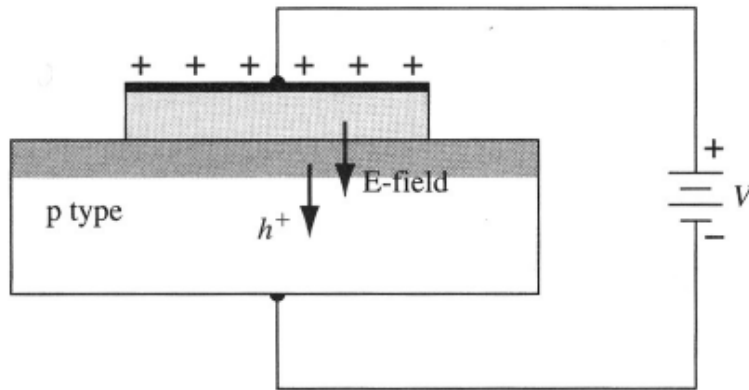
$$Q' = C'V$$

$$E = \frac{V}{d} = \frac{Q'}{\epsilon_{ox}}$$

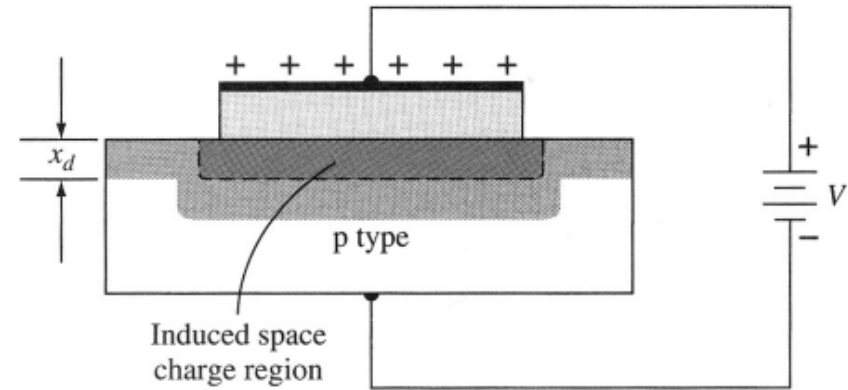
(a) Plattenkondensator-Modell mit elektrischem Feld und Ladungen, (b) MOS-Kondensator unter negativer Gatespannung,  $E$ -Feld und Ladungsfluss, (c) MOS-Kapazität mit Akkumulationsschicht von Löchern.



# Bandkanten-Diagramme

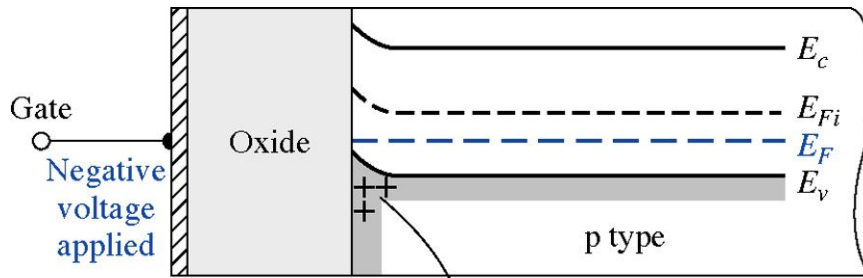


(a)

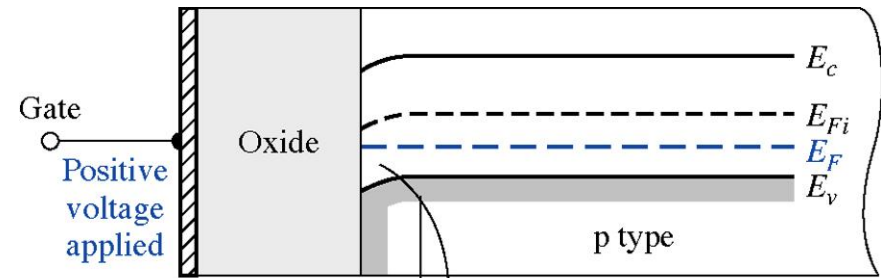


(b)

MOS-Kondensator unter moderater positiver Gatespannung, (a)  $E$ -Feld und Ladungsfluss, (b) induzierte Raumladungszone.



Accumulation layer of holes

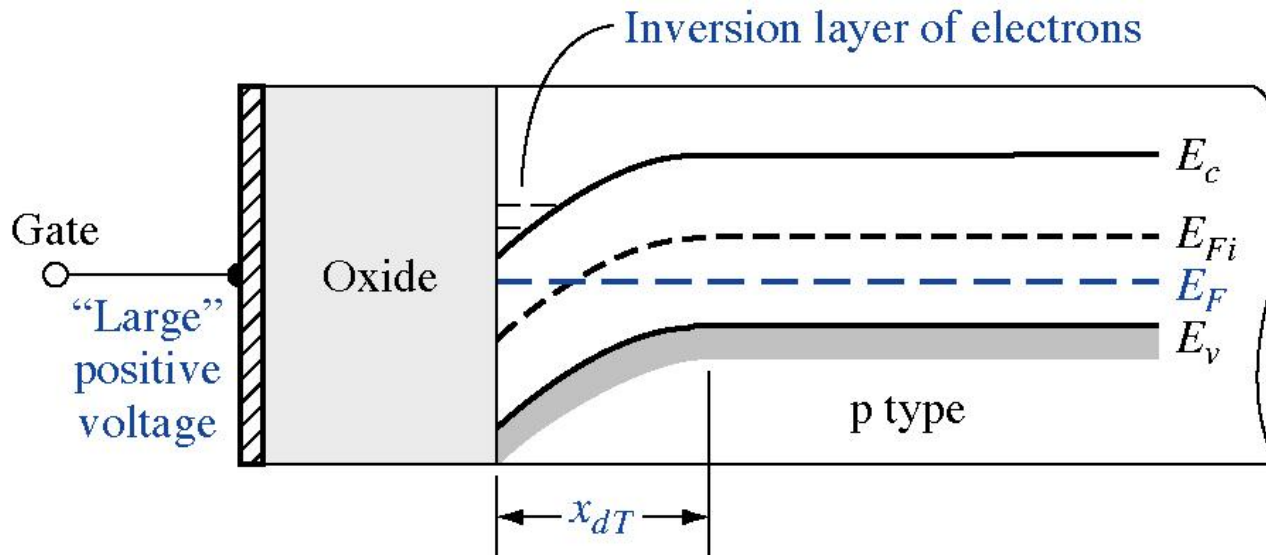


Induced space charge region

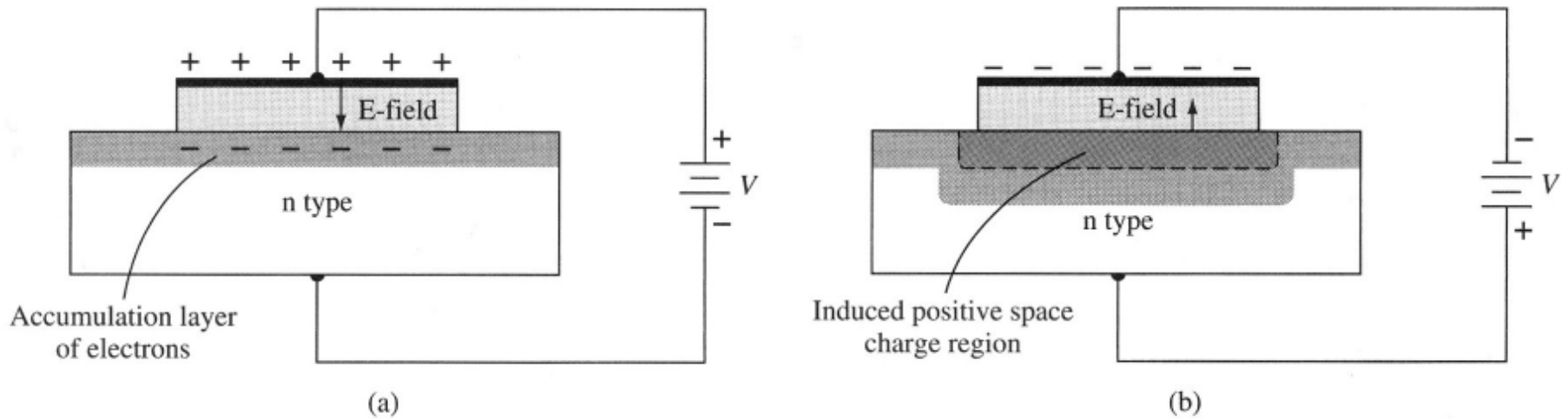
Bandkanten-Verlauf im MOS-Kondensator mit p-Substrat. Links: negative Gatespannung, rechts: positive Gatespannung.

# Der n-Kanal MOS-Kondensator in Inversion

Legt man eine genügend grosse positive Spannung an das Gate an, so verbiegt sich das intrinsische Fermi-niveau unter das extrinsische Fermi-niveau. Die Oberfläche des p-Halbleiters wird n-leitend (= Inversion). Die sehr dünne Elektronenschicht wird *Inversionsschicht* genannt.

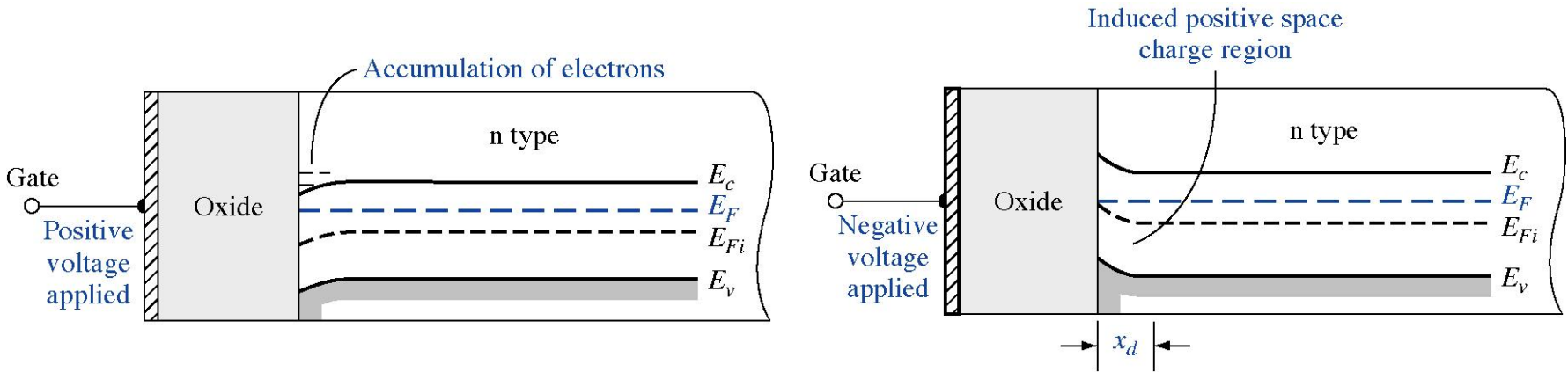


# Der p-Kanal MOS-Kondensator

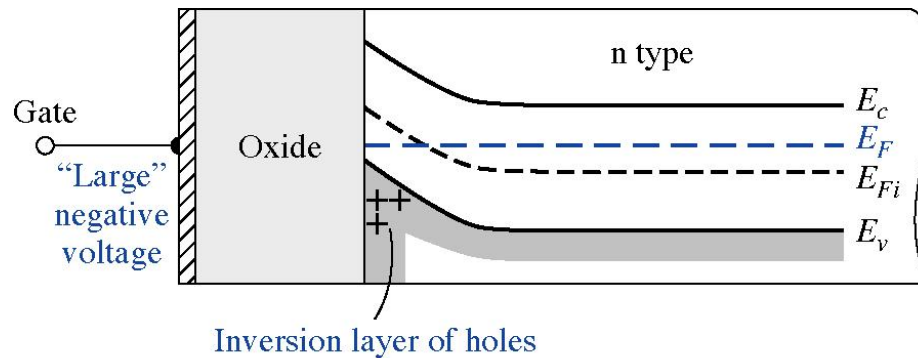


MOS-Kondensator mit n-Substrat unter (a) positiver Gatespannung, (b) moderater negativer Gatespannung.

# Der p-Kanal MOS-Kondensator

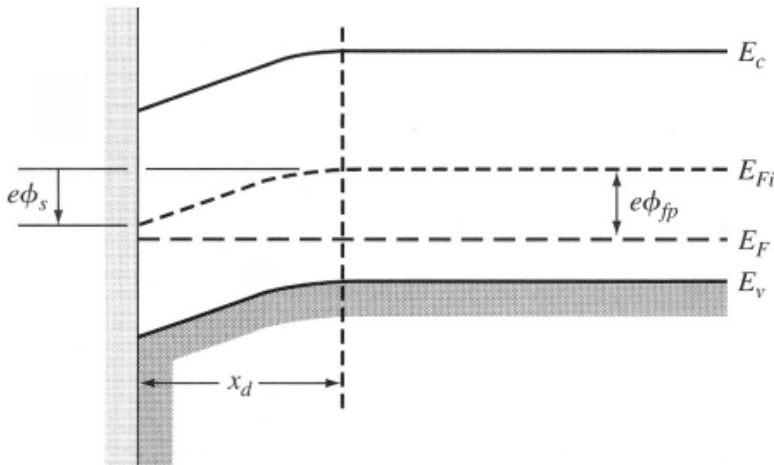


Bandkanten-Verlauf im MOS-Kondensator mit n-Substrat. Links: positive Gatespannung, rechts: negative Gatespannung.

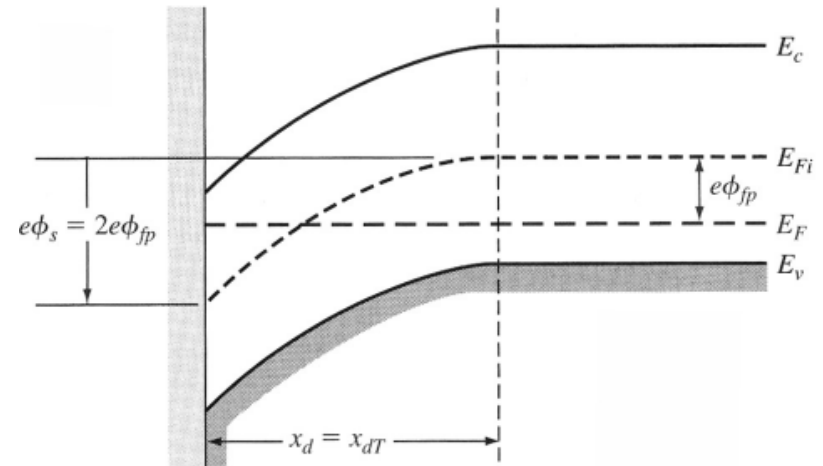


Legt man eine genügend grosse negative Spannung an das Gate an, so verbiegt sich das intrinsische Fermi-niveau über das extrinsische Fermi-niveau. Die Oberfläche des n-Halbleiters wird p-leitend (= Inversion). Die sehr dünne Löcherschicht wird **Inversionsschicht** genannt.

# Die Weite der Verarmungsschicht



Bandkantediagramm im p-Typ-Halbleiter.  
 $\Phi_s$  ist das Oberflächenpotential.



Bandkantediagramm im p-Typ-Halbleiter  
 bei Anlegen der Schwellspannung

$$\Phi_{fp} = V_t \ln \left( \frac{N_a}{n_i} \right)$$

$$x_{dT} = \left( \frac{4 \epsilon_s \Phi_{fp}}{e N_a} \right)^{1/2}$$

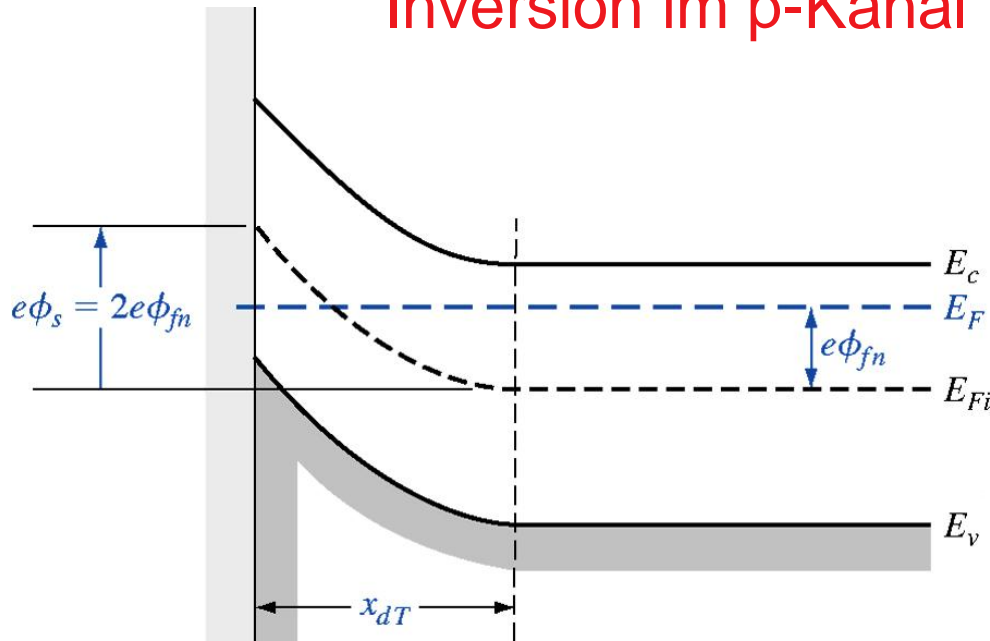
Das Potential  $\Phi_s$  heisst *Oberflächenpotential* (Spannungsdifferenz zwischen intrinsischem Fermienergielevel im Substrat und an der Oberfläche). Es entspricht der gesamten Bandverbiegung im HL.

$$x_d = \left( \frac{2 \epsilon_s \Phi_s}{e N_a} \right)^{1/2}$$

Die Weite der Verarmungsschicht entspricht der RLZ-Weite des einseitigen pn-Übergangs mit eingebautem Potential  $\Phi_s$ .

Für  $\Phi_s = 2\Phi_{fp}$  ist die Elektronendichte im Kanal so gross wie die Löcherdichte im Bulk. Die entsprechende Gatespannung heisst *Schwellspannung*.

# Inversion im p-Kanal MOS-Kondensator

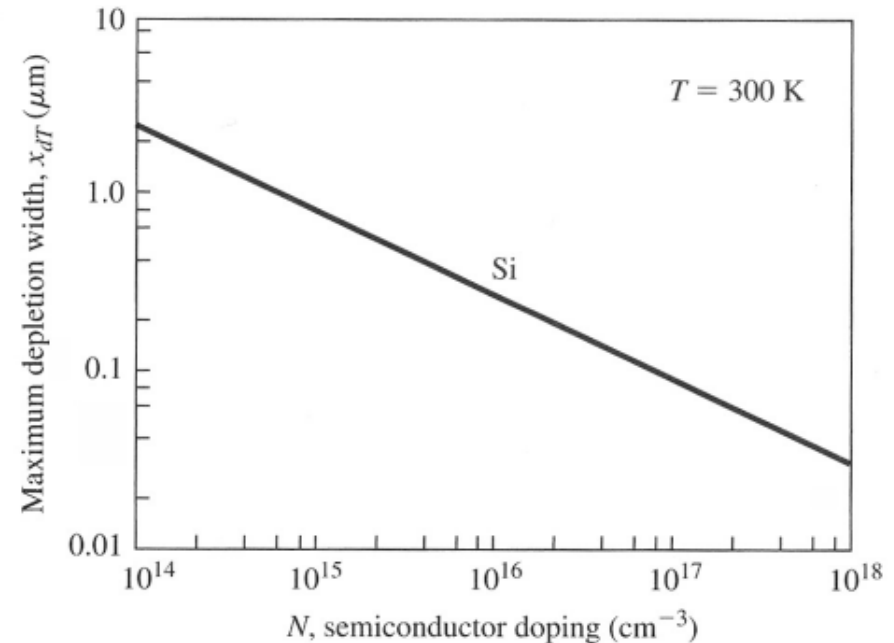


$$\Phi_{fn} = V_t \ln \left[ \frac{N_d}{n_i} \right]$$

Für  $\Phi_s = 2\Phi_{fn}$  ist die Löcherdichte im Kanal so gross wie die Elektronendichte im Bulk.

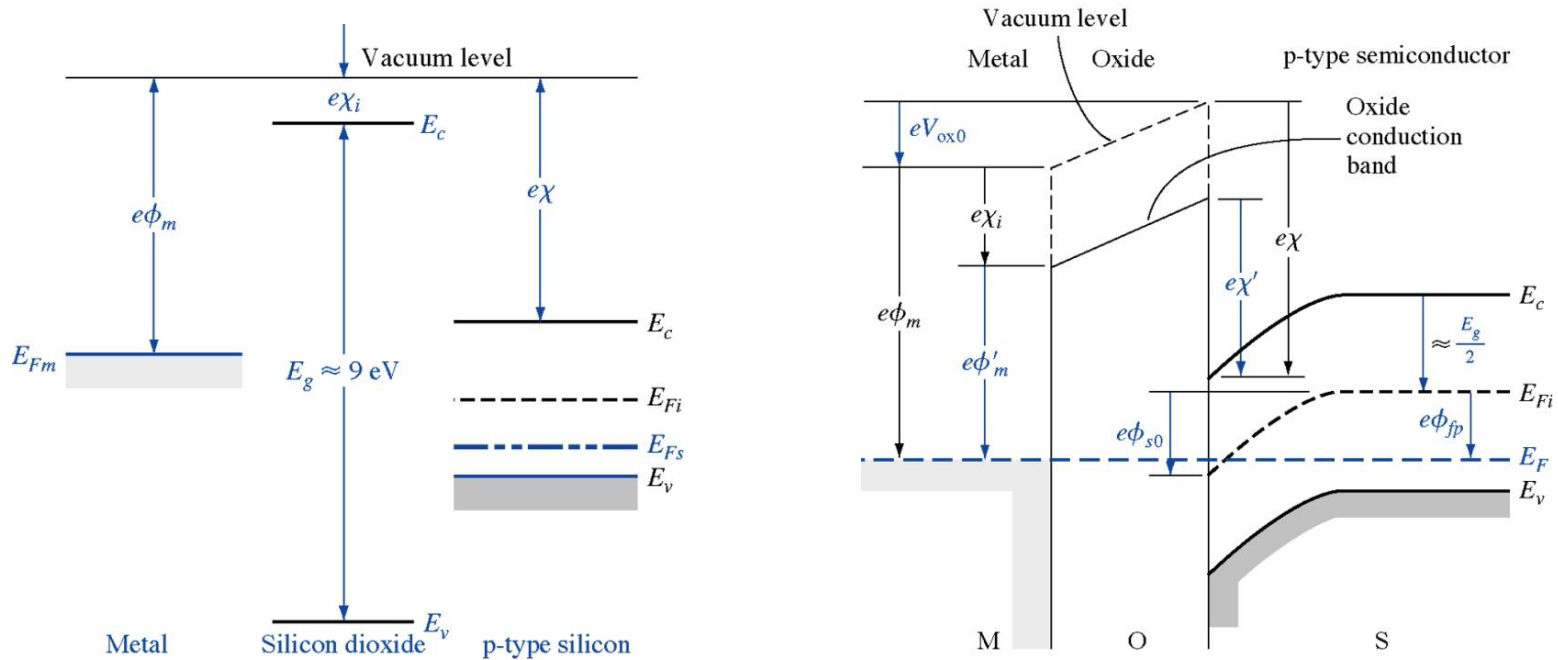
$$x_{dT} = \left[ \frac{4 \epsilon_s \Phi_{fn}}{e N_d} \right]^{1/2}$$

Nach Einsetzen der Inversion ändert sich die Weite der Verarmungsschicht kaum noch. Die Bänder verbiegen sich nur noch leicht, aber die Inversionsdichte steigt exponentiell ( $\rightarrow$  Abschirmeffekt für das eindringende Feld).



Maximale Ausdehnung der Raumladungszone als Funktion der Dotierung im Halbleiter.





$$e\Phi'_m + eV_{ox0} = e\chi' + \frac{E_g}{2} - e\Phi_{s0} + e\Phi_{fp}$$

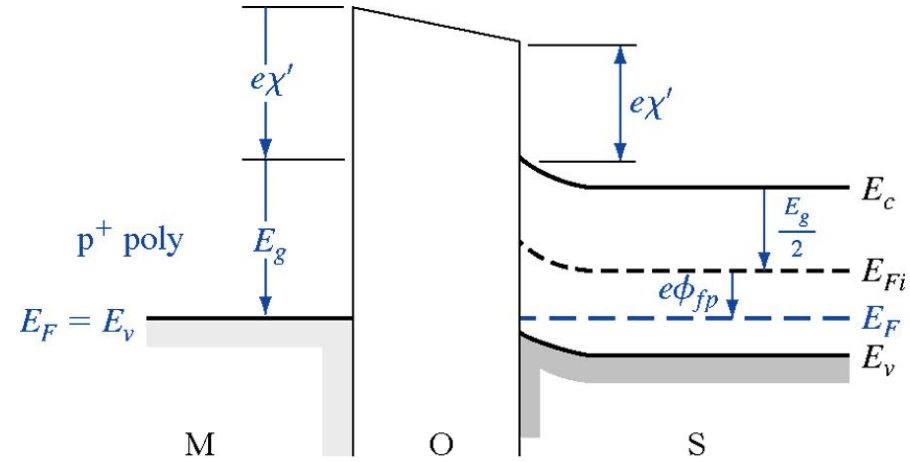
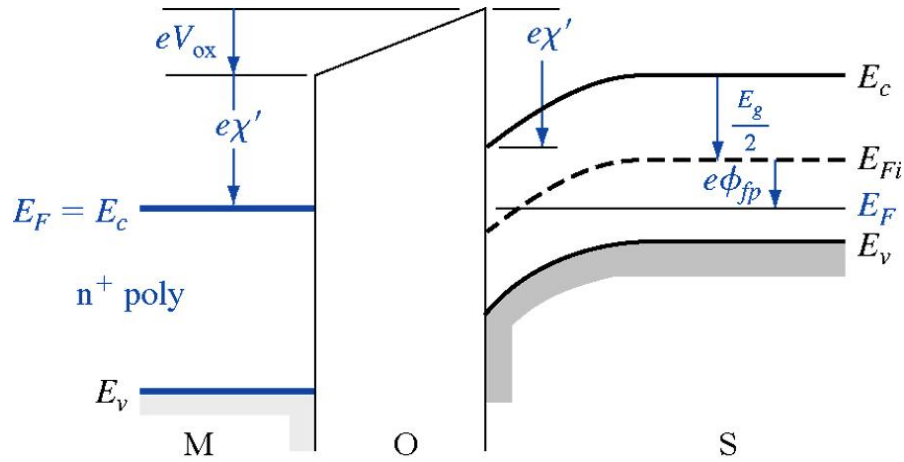
$$\Phi_{s0} + V_{ox0} = -\left\{ \Phi'_m - \left[ \chi' + \frac{E_g}{2e} + \Phi_{fp} \right] \right\}$$

$$\Phi_{ms} = \Phi'_m - \left[ \chi' + \frac{E_g}{2e} + \Phi_{fp} \right]$$

Differenz der Austrittsarbeiten



# n- oder p-Polysilizium als Gate-Material

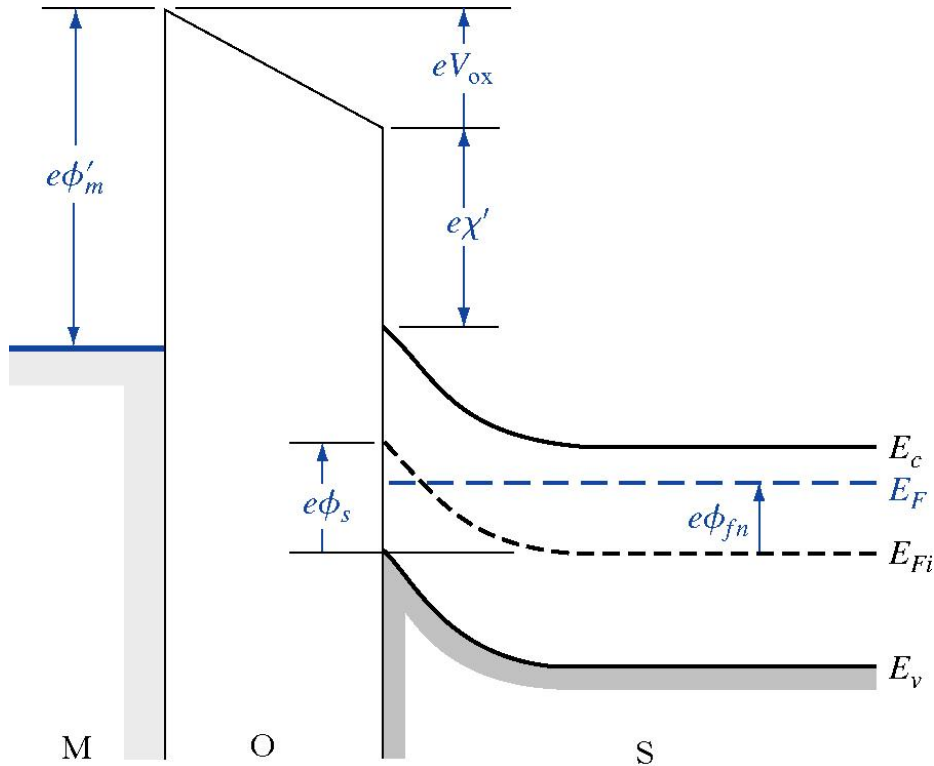


Bandkanten-Verlauf im MOS-Kondensator mit p-Substrat. Links: n<sup>+</sup>-Polysilizium-Gate, rechts: p<sup>+</sup>-Polysilizium-Gate. Das Zusammenfallen der Fermi-niveaus im Polysilizium mit den Bandkanten ist eine Näherung.

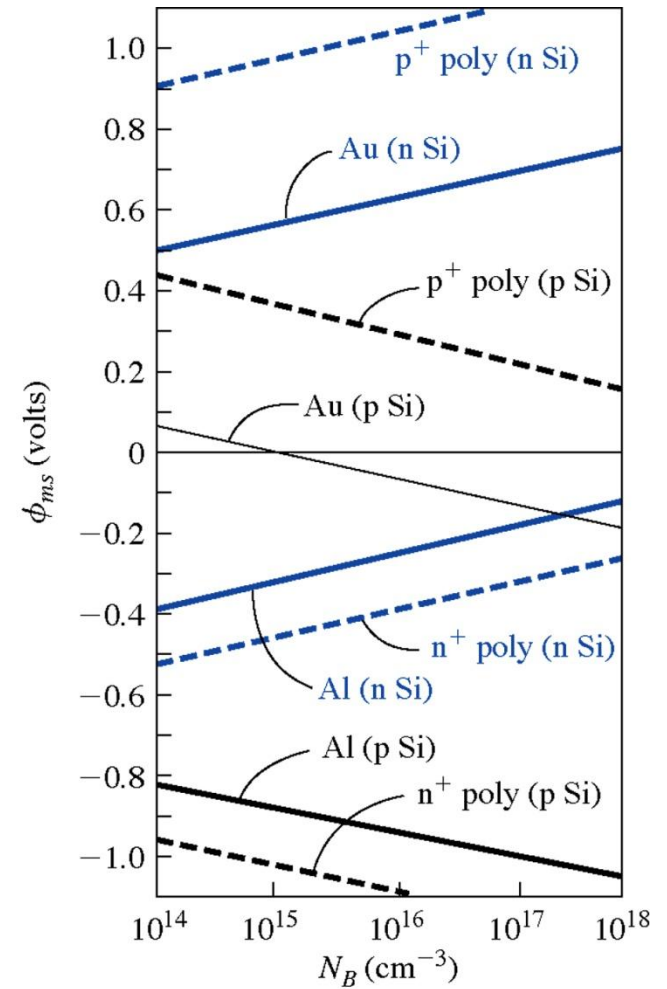
$$\begin{aligned}\Phi_{ms} &= \chi' - \left( \chi' + \frac{E_g}{2e} + \Phi_{fp} \right) \\ &= - \left( \frac{E_g}{2e} + \Phi_{fp} \right)\end{aligned}$$

$$\begin{aligned}\Phi_{ms} &= \chi' + \frac{E_g}{e} - \left( \chi' + \frac{E_g}{2e} + \Phi_{fp} \right) \\ &= \frac{E_g}{2e} - \Phi_{fp}\end{aligned}$$

# Austrittsarbeit im p-Kanal MOS-Kondensator



$$\Phi_{ms} = \Phi'_m - \left[ \chi' + \frac{E_g}{2e} - \Phi_{fn} \right]$$

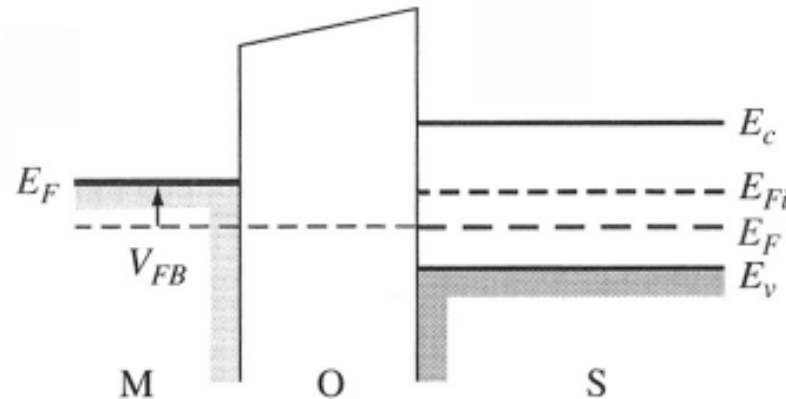


# Die Flachbandspannung

Diejenige Gatespannung, bei der es keine Bandverbiegung und damit auch keine Raumladung im Halbleiter gibt. Der Spannungsabfall über dem Oxid ist nicht notwendigerweise Null wegen der Differenz der Austrittsarbeiten und wegen diverser Ladungen:

- Ladungen im Innern des Oxids ( $Q_{ox}$ ),
- fixe Ladungen an der Si-SiO<sub>2</sub>-Grenzfläche ( $Q_{ss}$ )
- umladbare ungesättigte Bindungen an der Si-SiO<sub>2</sub>-Grenzfläche ( $Q_{it}$ )

Bei modernen MOS-Technologien kann man  $Q_{ox}$  und  $Q_{it}$  vernachlässigen.



Bandkantendiagramm eines MOS-Kondensators unter Flachband-Bedingungen.

## Die Flachbandspannung (2)

Fassen alle Ladungen zu einer äquivalenten Flächenladung  $Q'_{ss}$  zusammen. Die Nettoladung im Halbleiter ist Null (Flachband-Bedingung). Im Gate existiert dann eine Ladung  $Q'_m$ , die wegen globaler Ladungsneutralität gerade  $Q'_{ss}$  kompensiert:

$$0 = Q'_m + Q'_{ss}$$

Ohne Gatespannung gilt allgemein

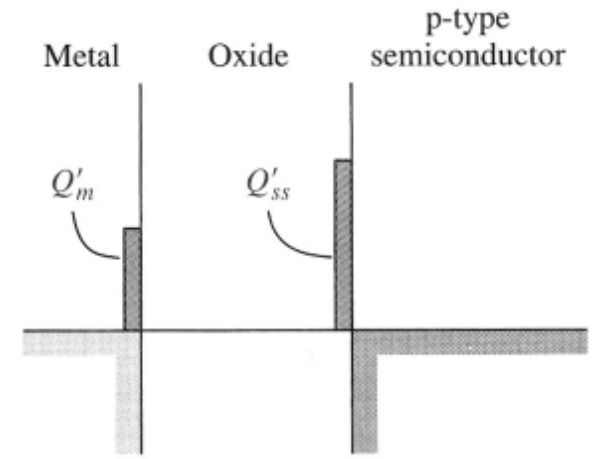
$$\Phi_{s0} + V_{ox0} = -\Phi_{ms}$$

d.h. die Summe aus Bandverbiegung im HL und Spannungsabfall über dem Oxid ist gerade der negativen Austrittsarbeitsdifferenz. Eine angelegte Gatespannung führt zu

$$\begin{aligned} V_G &= \Delta V_{ox} + \Delta \Phi_s = V_{ox} - V_{ox0} + (\Phi_s - \Phi_{s0}) \\ &= V_{ox} + \Phi_s + \Phi_{ms} \end{aligned}$$

Wegen  $V_{ox} = Q'_m / C_{ox} = -Q'_{ss} / C_{ox}$  und der Flachbandbedingung  $\Phi_s = 0$  folgt

$$V_{FB} = \Phi_{ms} - \frac{Q'_{ss}}{C_{ox}}$$



Ladungsverteilung in einem MOS-Kondensator unter Flachband-Bedingungen.

# Die Schwellspannung

Die Spannung, die man an das Gate anlegen muss, um den Inversionspunkt  $\Phi_s = 2\Phi_{fp}$  (beim n-Kanal MOS-Kondensator) zu erhalten, heisst **Schwellspannung**. Bezeichnung:  $V_{TN}$ .

Am Inversionspunkt hat die Ausdehnung der Raumladungsschicht ihr Maximum erreicht. Unter Vernachlässigung der Inversionsladung führt die globale Ladungserhaltung auf

$$Q'_{mT} + Q'_{ss} = |Q'_{SD}(\max)| = eN_a x_{dT} = \left[4\epsilon_s e N_a \Phi_{fp}\right]^{1/2}$$

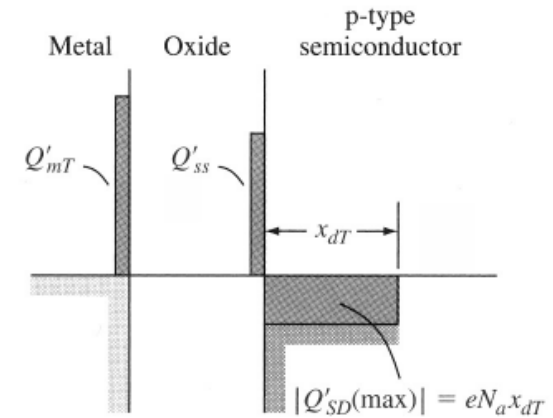
Aus der allgemeinen Spannungsbilanz  $V_G = V_{ox} + \Phi_s + \Phi_{ms}$  folgt nach Definition die Schwellspannung

$$V_{TN} = V_{oxT} + 2\Phi_{fp} + \Phi_{ms}$$

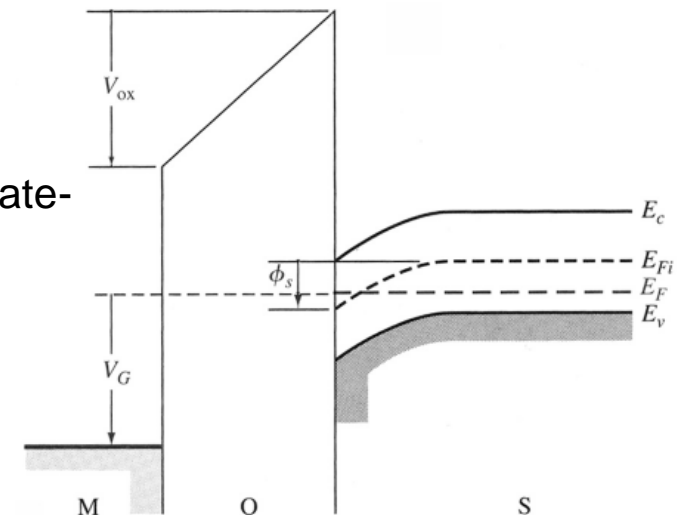
Den Spannungsabfall über dem Oxid kann man mittels der Gateladung am Inversionspunkt ausdrücken:  $V_{oxT} = Q'_{mT}/C_{ox}$ .

$$V_{TN} = \left[|Q'_{SD}(\max)| - Q'_{ss}\right]/C_{ox} + 2\Phi_{fp} + \Phi_{ms}$$

$$V_{TN} = \frac{|Q'_{SD}(\max)|}{C_{ox}} + V_{FB} + 2\Phi_{fp}$$

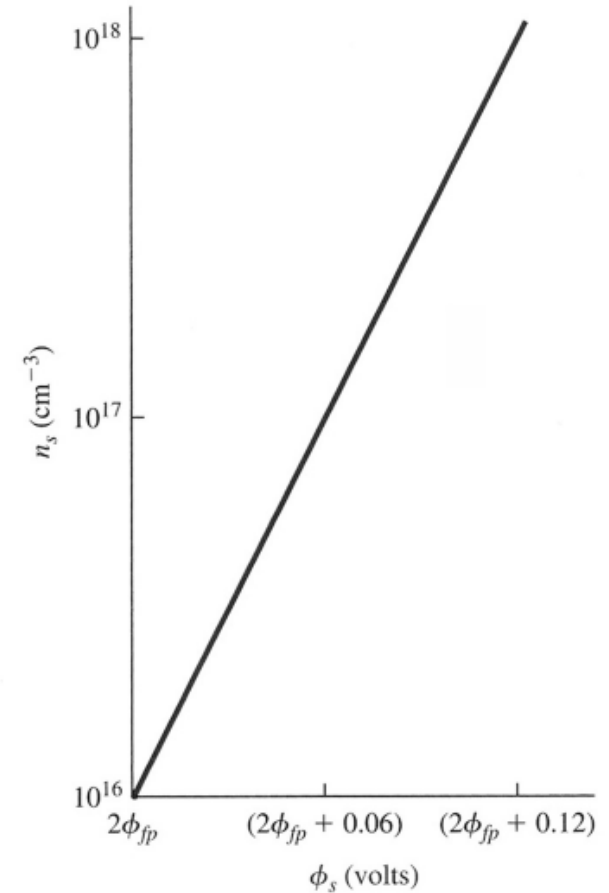
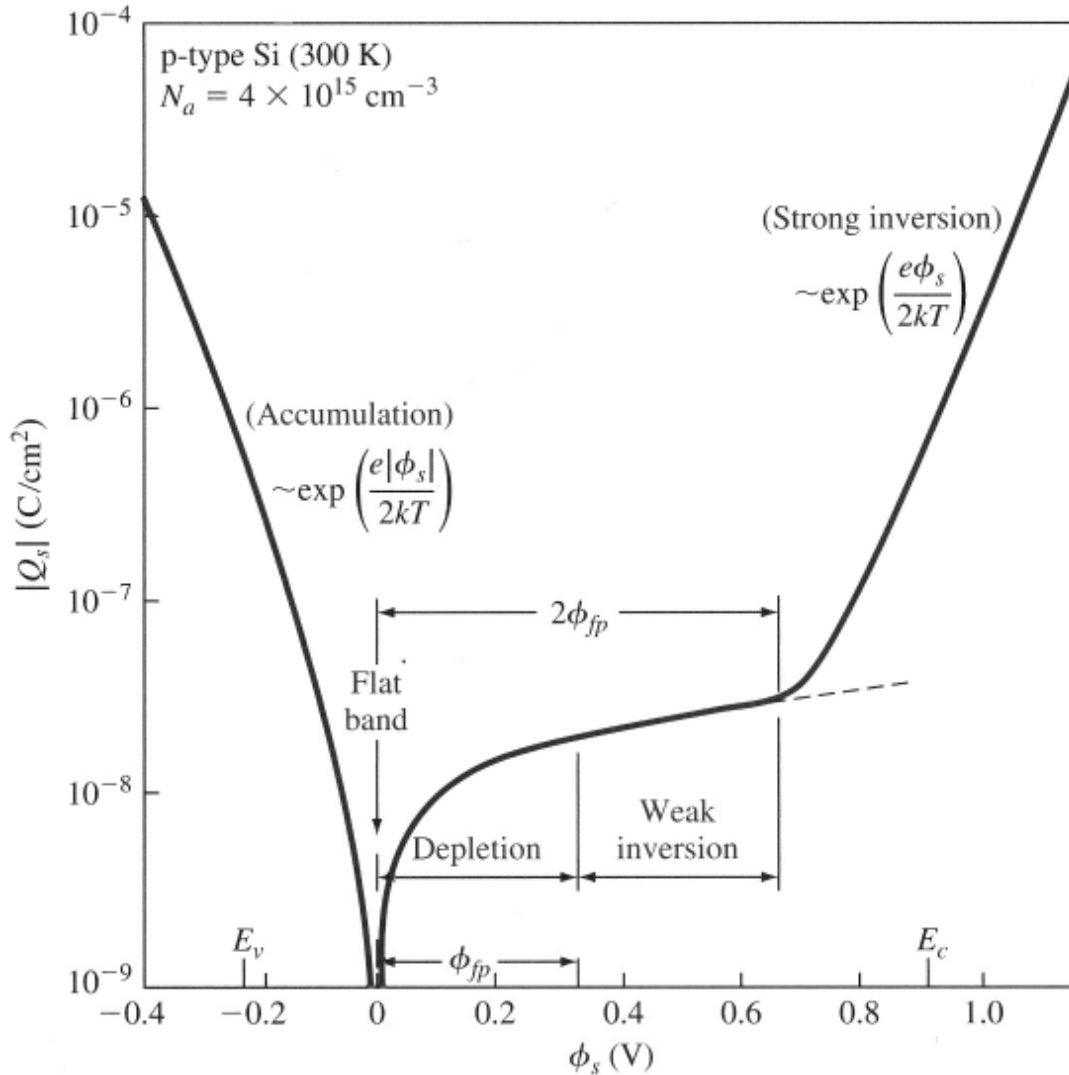


Ladungsverteilung in einem MOS-Kondensator mit p-Substrat bei Anlegen der Schwellspannung.



Bandkantendiagramm einer MOS-Struktur unter positiver Gatespannung.

# Die Ladungsverteilung im MOS-Kondensator

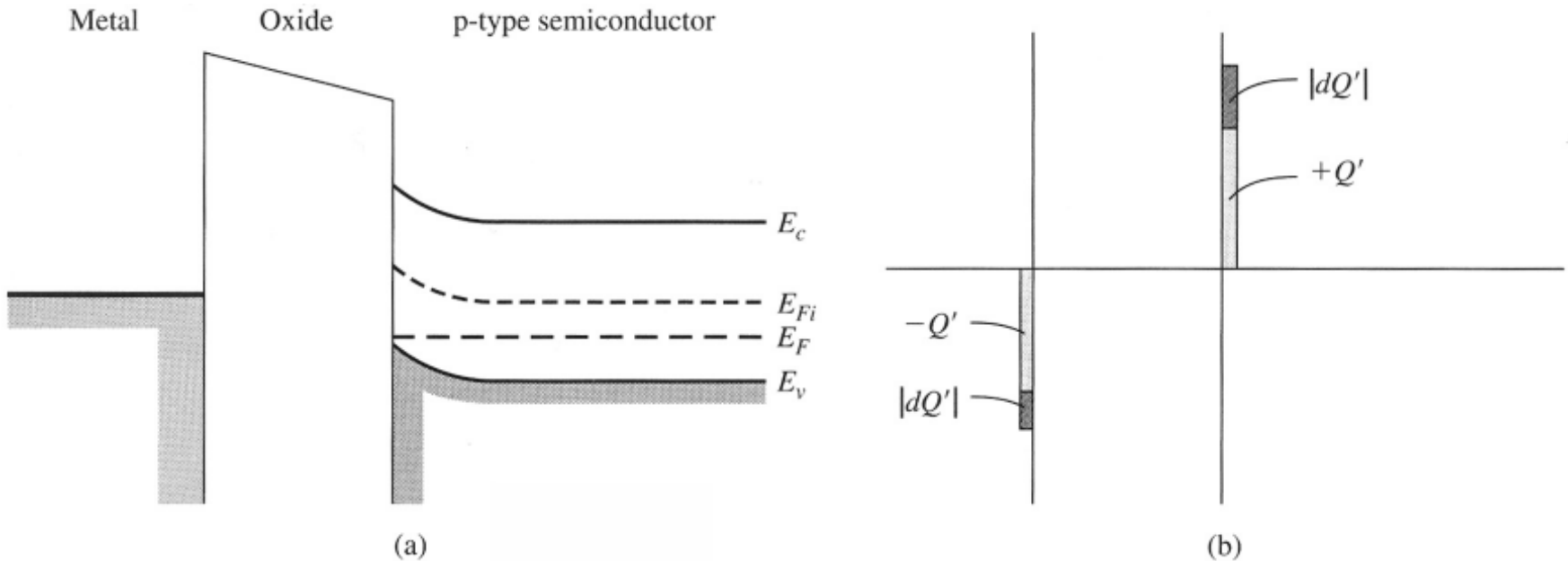


Inversionsladungsdichte der Elektronen als Funktion des Oberflächenpotentials.

Änderung der Oberflächenladungsdichte (Akkumulations- und Inversionsladung) als Funktion des Oberflächenpotentials (nach Sze).

# 11.2 CV-Kennlinien

## Der Idealfall. Akkumulation

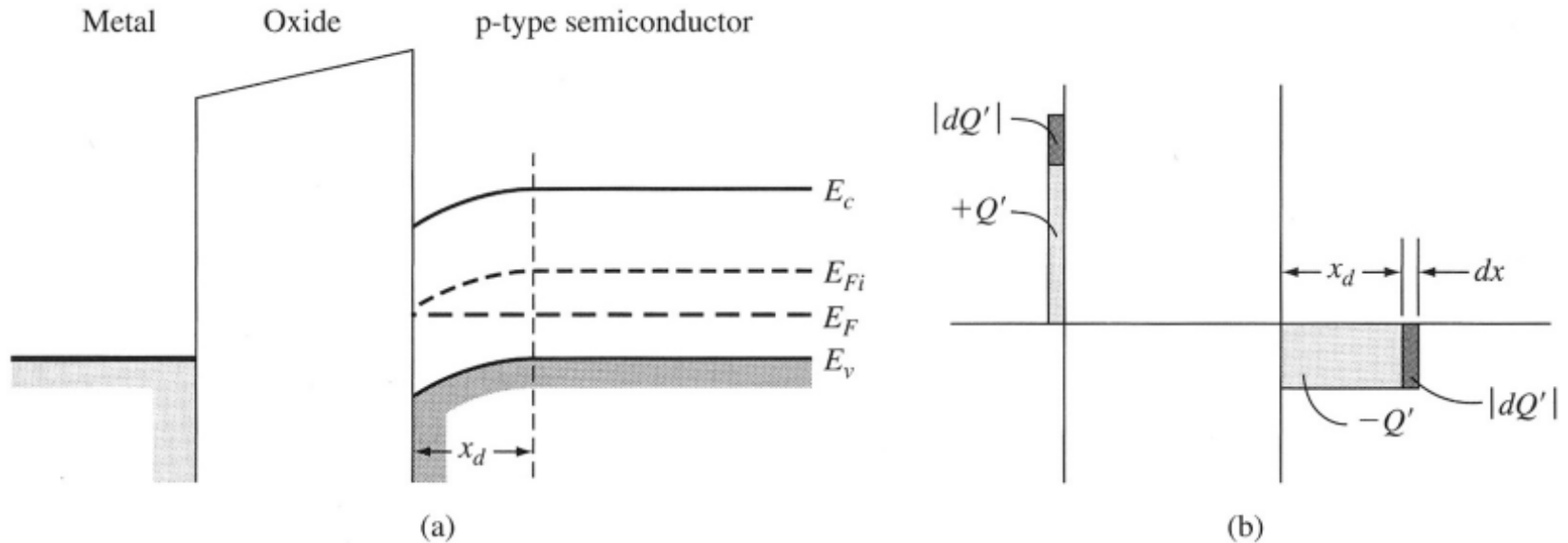


(a) Bandkantendiagramm eines MOS-Kondensators unter Akkumulationsbedingungen. (b) Differentielle Ladungsverteilung unter Akkumulationsbedingungen bei differentieller Änderung der Gatespannung.

**Akkumulation**

$$C'(\text{acc}) = C_{\text{ox}} = \frac{\epsilon_{\text{ox}}}{t_{\text{ox}}}$$

# Der Idealfall. Verarmung



(a) Bandkantediagramm eines MOS-Kondensators unter Verarmungsbedingungen. (b) Differentielle Ladungsverteilung unter Verarmungsbedingungen bei differentieller Änderung der Gatespannung.

## Verarmung

$$\frac{1}{C'(\text{depl})} = \frac{1}{C_{\text{ox}}} + \frac{1}{C'_{\text{SD}}}$$

$$C'(\text{depl}) = \frac{C_{\text{ox}} C'_{\text{SD}}}{C_{\text{ox}} + C'_{\text{SD}}}$$

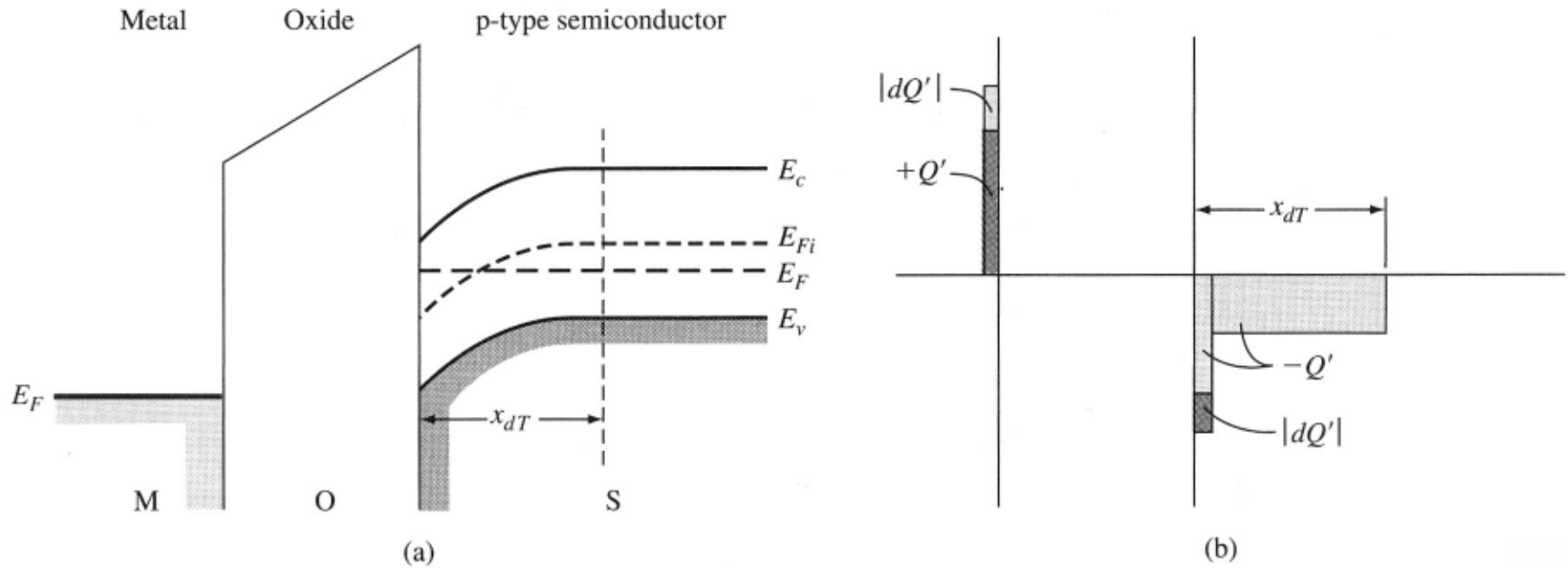
$$C'(\text{depl}) = \frac{\epsilon_{\text{ox}}}{t_{\text{ox}} + \frac{\epsilon_{\text{ox}}}{\epsilon_s} x_d}$$

Minimum der Kapazität bei max. Weite der Verarmungsschicht:

$$C'_{\text{min}} = \frac{\epsilon_{\text{ox}}}{t_{\text{ox}} + \frac{\epsilon_{\text{ox}}}{\epsilon_s} x_{dT}}$$



# Der Idealfall. Inversion

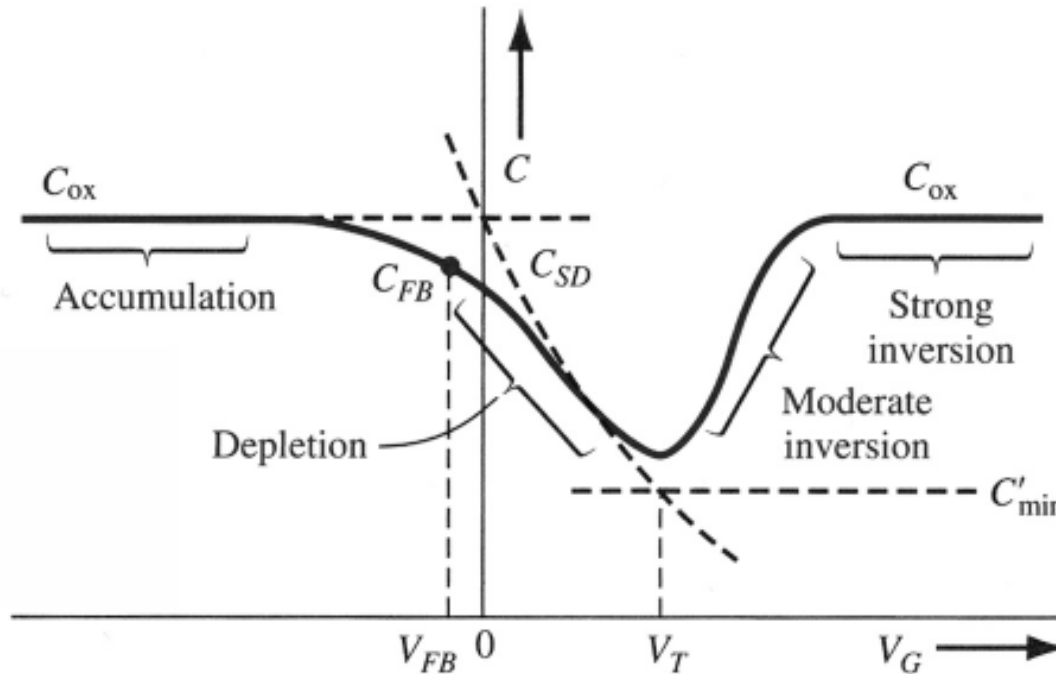


(a) Bandkantendiagramm eines MOS-Kondensators unter Inversionsbedingungen. (b) Differentielle Ladungsverteilung unter Inversionsbedingungen bei niedrigfrequenter differentieller Änderung der Gatespannung.

## Inversion

$$C'(\text{inv}) = C_{\text{ox}} = \frac{\epsilon_{\text{ox}}}{t_{\text{ox}}}$$

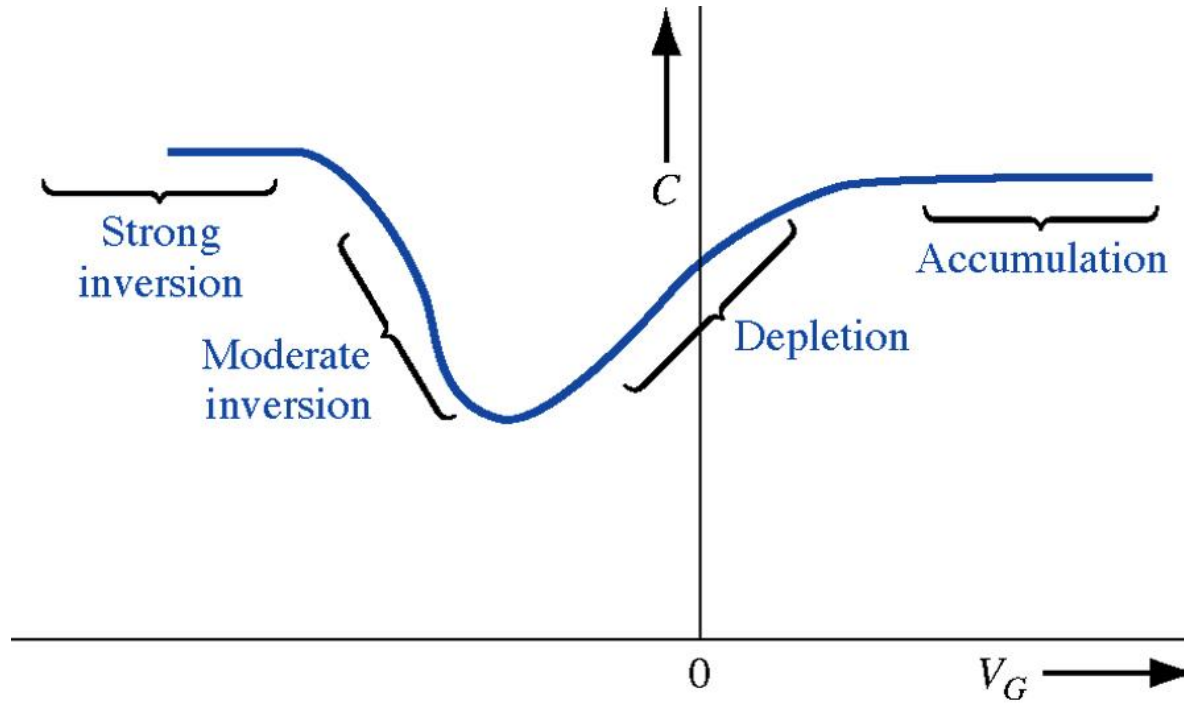
# Der Idealfall. Niedrigfrequenz-CV-Kurve



Ideale Niedrigfrequenz-Kapazität eines MOS-Kondensators mit p-Substrat als Funktion der Gatespannung. Die individuellen Kapazitätsbeiträge sind ebenfalls dargestellt.

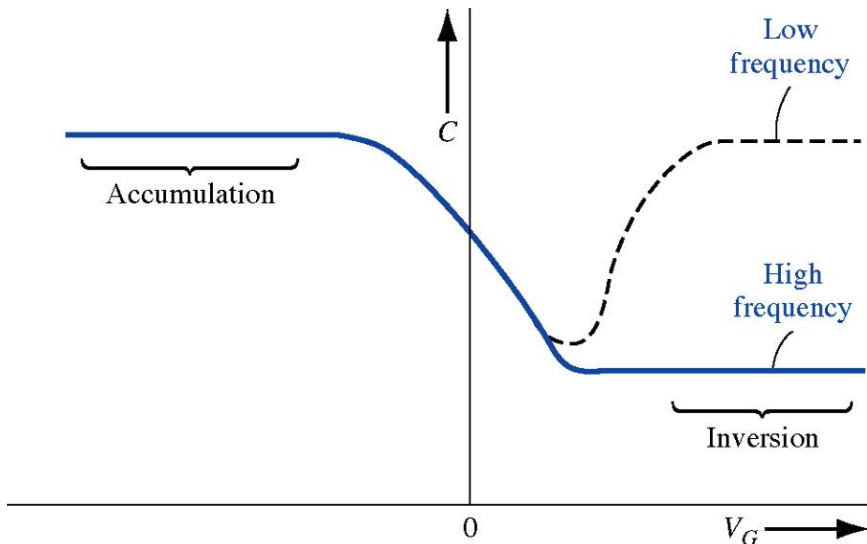
$$C'_{FB} = \frac{\epsilon_{ox}}{t_{ox} + \frac{\epsilon_{ox}}{\epsilon_s} \left( \frac{\epsilon_s V_t}{e N_a} \right)^{1/2}}$$

# CV-Kennlinie des p-Kanal MOS-Kondensators



Niedrigfrequenz-CV-Kurve des MOS-Kondensators mit n-Substrat. Akkumulation tritt für positive Gatespannungen auf, Inversion für negative. Die CV-Kurve ist die an der C-Achse gespiegelte Charakteristik des Kondensators mit p-Substrat.

# Frequenz-Effekte



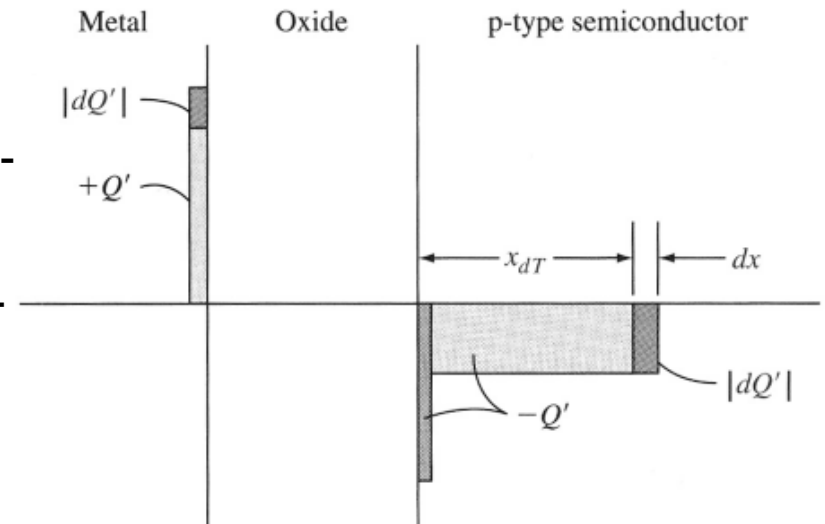
Welche Elektronen tragen zur differentiellen Änderung der Inversionsladung aufgrund der differentiellen Änderung der Kondensatorspannung bei?

- 1.) diffundierende Minoritäten aus dem p-Substrat
- 2.) thermische Generation von e-h-Paaren in der RLZ

Die Elektronenkonzentration in der Inversionsschicht kann sich nicht instantan ändern.

Bei Frequenzen  $> 1$  MHz kann die Inversionsladung dem Wechselsignal nicht mehr folgen.

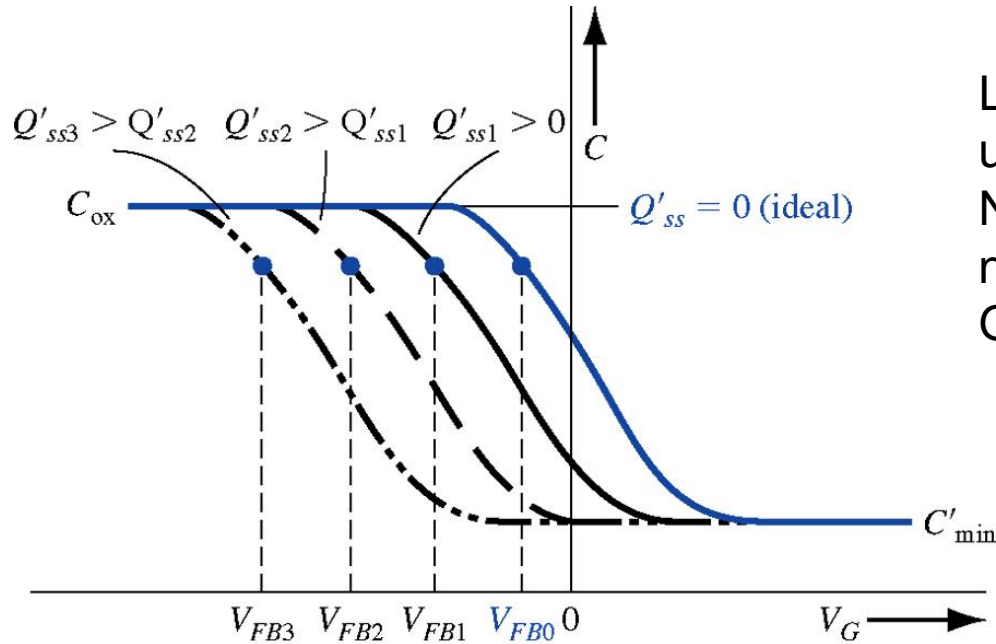
Die differentielle Ladungsänderung tritt dann nur im Gate und am Rand der RLZ auf. Man hat den Fall der minimalen Kapazität  $C'_{\min}$ .



Differentielle Ladungsverteilung unter Inversionsbedingungen bei einer hochfrequenten differentiellen Änderung der Gatespannung.

# Nichtideales Verhalten infolge Oxid- und Grenzflächenladungen

## feste positive Oxidladungen

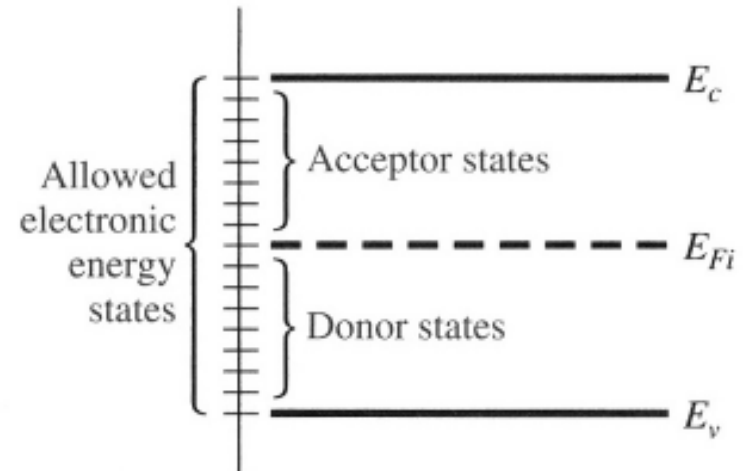


$$V_{FB} = \Phi_{ms} - \frac{Q'_{ss}}{C_{ox}}$$

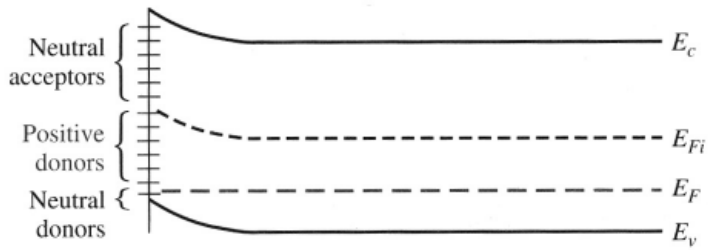
CV-Messungen dienen zur Diagnose von MOS-Bauelementen!

## Grenzflächenladungen

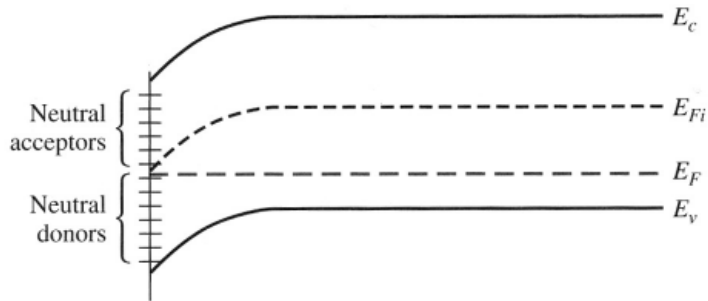
Ladungsaustausch zwischen Substrat und Grenzflächenzuständen → Nettoladung hängt von Lage des Fermi-niveaus ab, also von der angelegten Gatespannung.



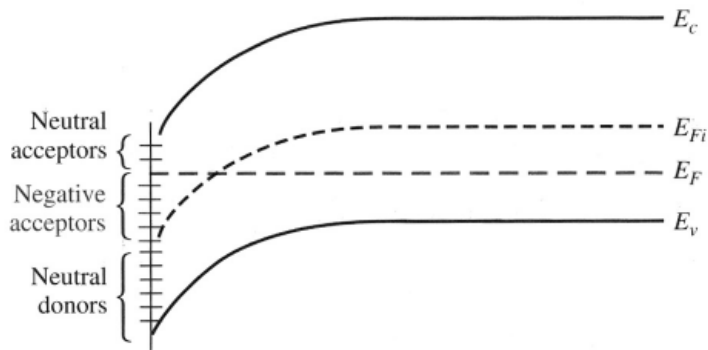
Grenzflächenzustände an der Oxid-Halbleiter-Grenzfläche.



(a)



(b)



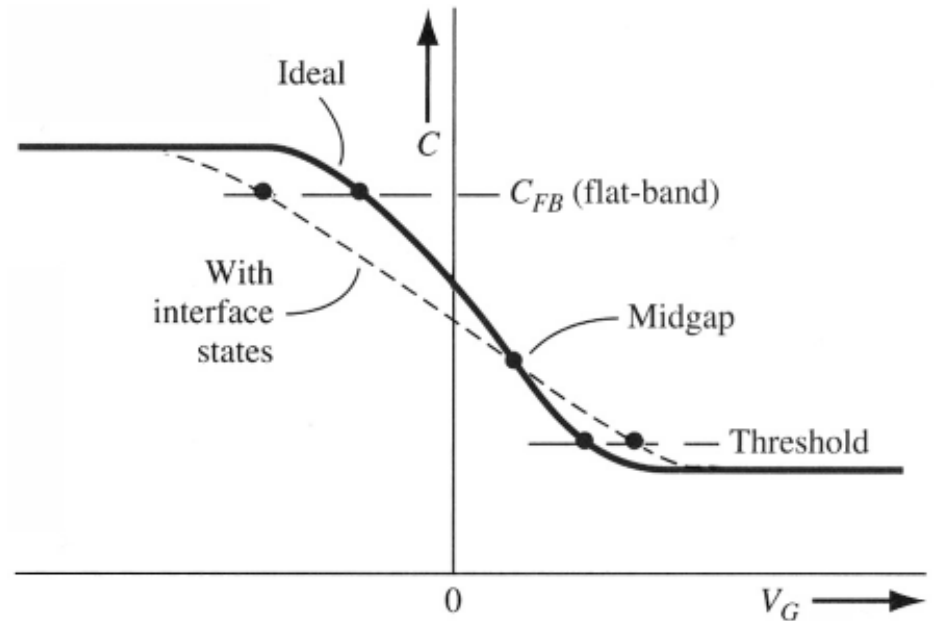
(c)

Bandkantendiagramm im p-Typ-Halbleiter mit eingefangenen Ladungen in den Grenzflächenzuständen. (a) Akkumulation, (b) Midgap-Bedingung, (c) Inversion.

Die Nettoladung in den Grenzflächenzuständen wechselt von positiv zu negativ, wenn man von Akkumulation zu Inversion geht.

Die CV-Kurve wird "ausgeschmiert".

Die Steilheit der Flanke ist eine Messung für die Dichte der Grenzflächenzustände.



Wirkung der Grenzflächenzustände auf die Hochfrequenz-CV-Kurve eines MOS-Kondensators.









# 12. Der MOSFET (1)

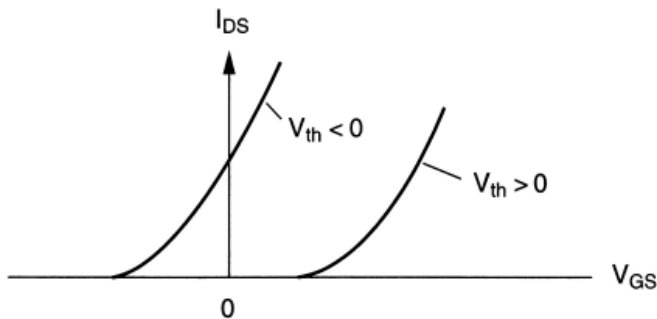
**12.1 Die Wirkungsweise des MOSFETs (N11.3)**

**12.2 Das Frequenz-Verhalten (N11.4)**

**12.3 CMOS-Technologie\* (N11.5)**

# 12.1 Die Wirkungsweise des MOSFETs

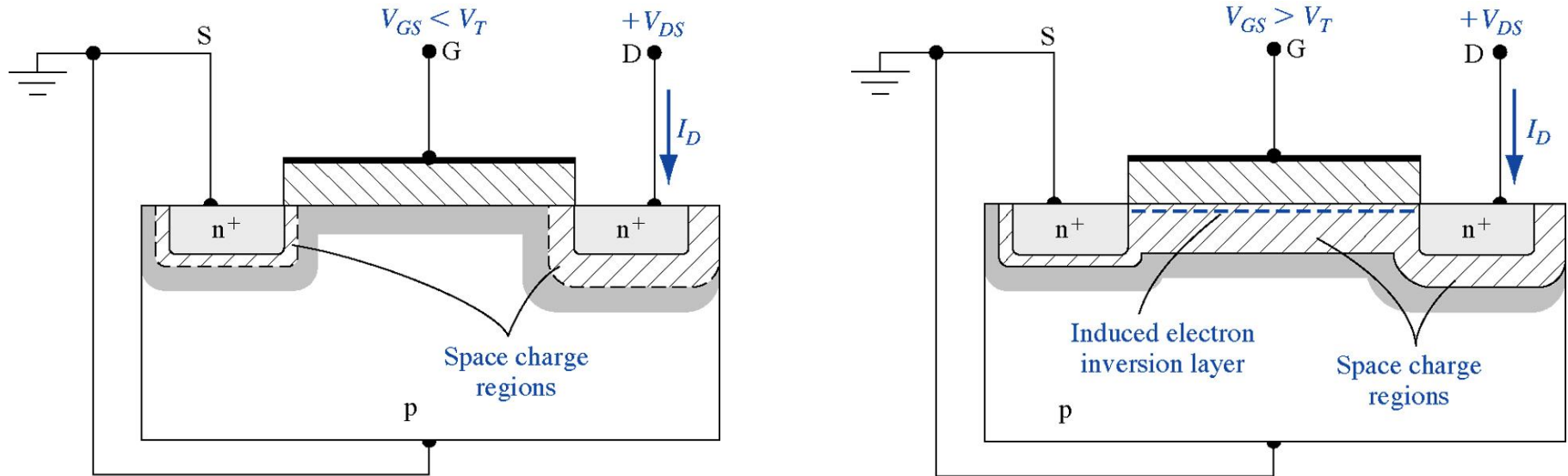
- 4 Arten: Je nach der Schwellspannung unterscheiden wir Anreicherungs- (Inversion, Enhancement) und Verarmungs- (Depletion)-Transistoren.



TYPE	CROSS SECTION	OUTPUT CHARACTERISTICS	TRANSFER CHARACTERISTICS
n-CHANNEL ENHANCEMENT (NORMALLY OFF)			
n-CHANNEL DEPLETION (NORMALLY ON)			
p-CHANNEL ENHANCEMENT (NORMALLY OFF)			
p-CHANNEL DEPLETION (NORMALLY ON)			

# Strom-Spannungs-Charakteristik (Grundprinzip)

Der MOSFET besteht aus einem MOS-Kondensator und Source- und Drain-Gebieten (die deutschen Ausdrücke "Quelle" und "Senke" werden nicht verwendet).



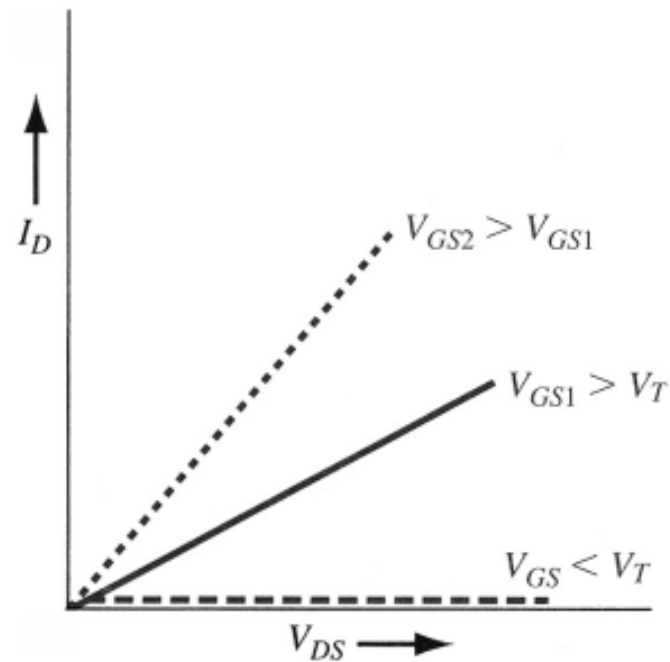
Der n-Kanal Anreicherungs-MOSFET mit  $V_{GS} < V_T$  (links) und  $V_{GS} > V_T$  (rechts).

Bei kleinen Drainspannungen  $V_{DS}$  verhält sich der MOSFET-Kanal wie ein Widerstand:

$$I_D = g_d V_{DS}$$

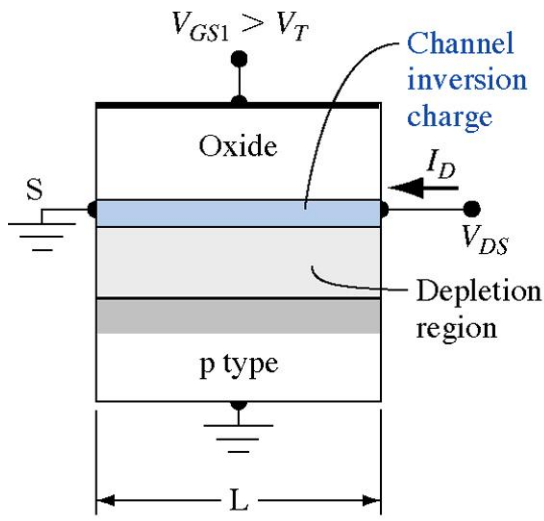
$g_d$  ist der Kanal-Leitwert:  $g_d = \frac{W}{L} \mu_n |Q'_n|$

$\mu_n$  = Kanal-Beweglichkeit  
 $|Q'_n|$  = Inversionsladung /cm<sup>2</sup>

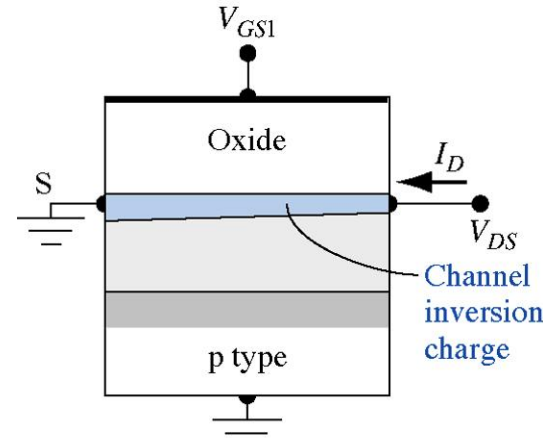
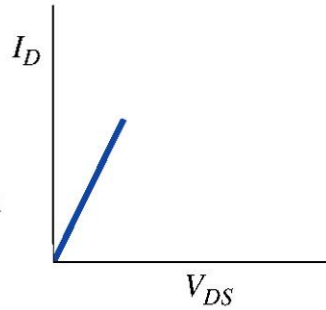


$I_D$ - $V_{DS}$ -Charakteristik für kleine  $V_{DS}$  bei 3 verschiedenen Werten von  $V_{GS}$ .

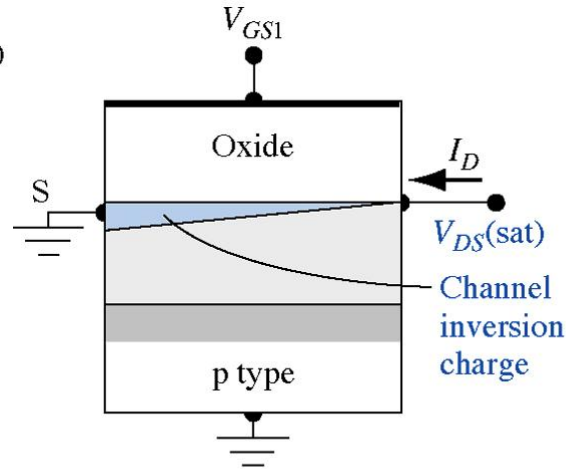
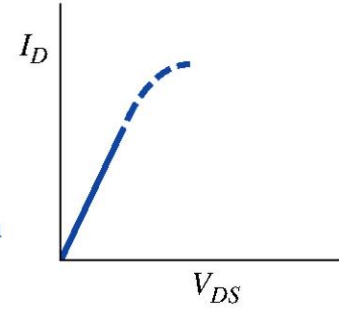
Die Ladung in der Inversionsschicht ist eine Funktion der Gatespannung. Deshalb ist die grundlegende Funktionsweise des MOSFETs die ***Modulation der Kanal-Leitfähigkeit durch die Gatespannung.***



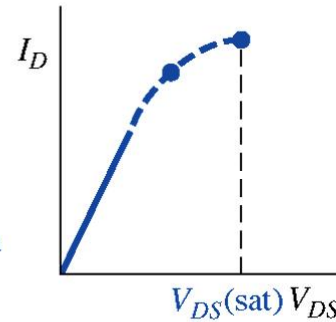
(a)



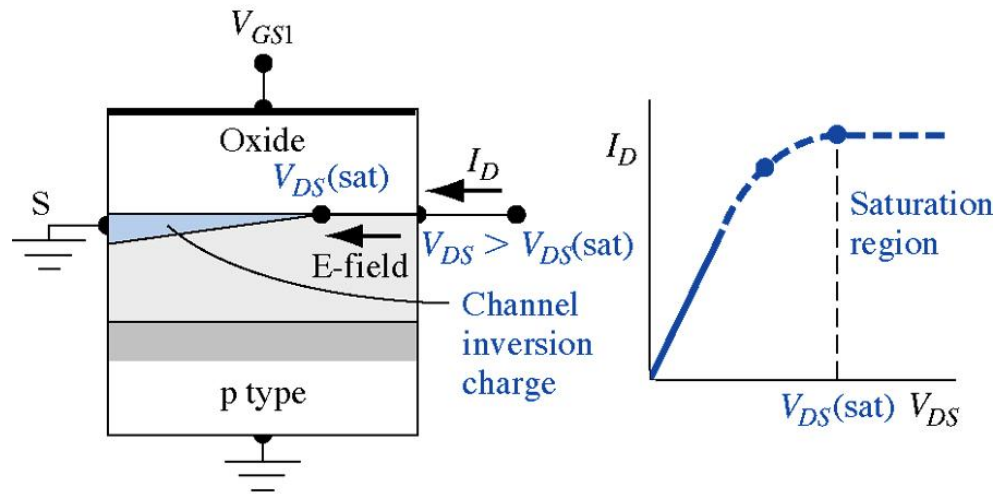
(b)



(c)



Drainstrom als Funktion der Drainspannung: Steigende Drainspannung  $\rightarrow$  kleinerer Spannungsabfall über dem Oxid am Drain  $\rightarrow$  schwächere Inversion am Drain  $\rightarrow$  Abflachung der Kennlinie. Wenn  $V_{DS} = V_{DS(sat)}$ , dann ist der Spannungsabfall am Drain gerade gleich  $V_T$   $\rightarrow$  Kennlinie wird konstant.



(d)

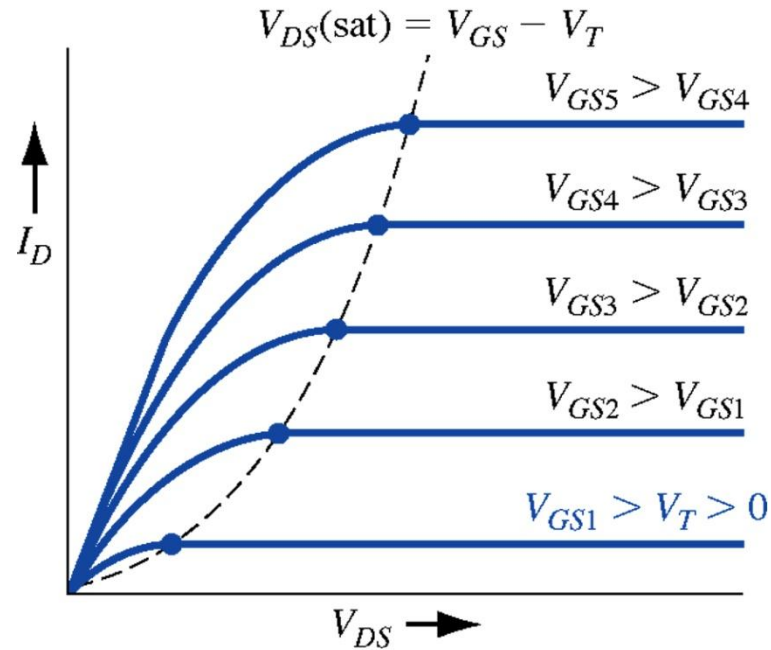
Wenn  $V_{DS} = V_{DS}(\text{sat})$ , dann ist der Spannungsabfall am Drain gerade gleich  $V_T$ , also

$$V_{GS} - V_{DS}(\text{sat}) = V_T$$

$$V_{DS}(\text{sat}) = V_{GS} - V_T$$

$V_{DS}(\text{sat})$  ist *die* Source-Drain-Spannung, die eine verschwindende Inversionsladung am Drain-Kontakt bewirkt.

Mit weiter wachsendem  $V_{DS}$  wandert der Abschnürpunkt in Richtung Source. Elektronen werden im gesperrten Substrat-Drain-Übergang abgesaugt. Vernachlässigt man die Reduktion der effektiven Kanallänge, bleibt der Sättigungsstrom konstant.

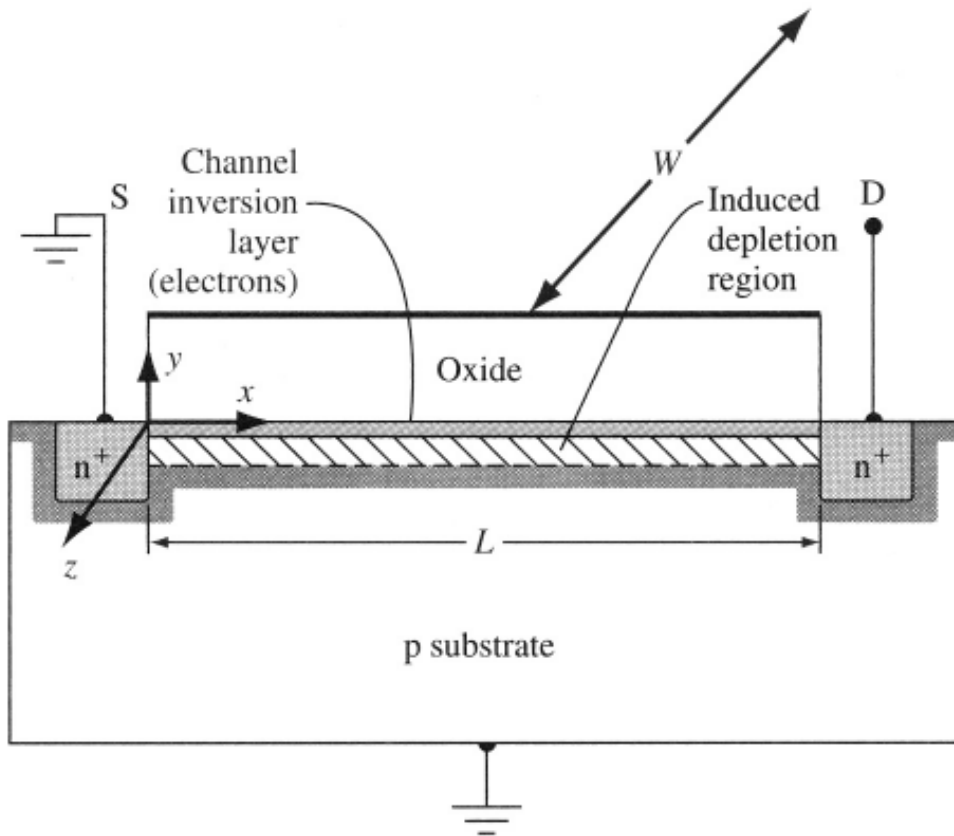


Mit steigender Gatespannung  $V_{GS}$  wächst der Sättigungsstrom  $I_D(\text{sat})$  und auch der Anstieg des Drainstroms im linearen Bereich.

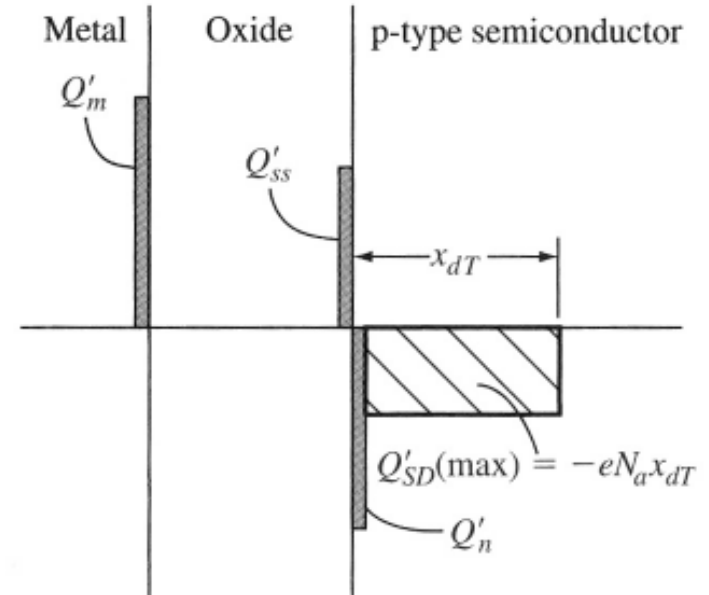
Ersteres, da  $V_{DS}(\text{sat}) = V_{GS} - V_T$ , letzteres, da  $I_D(\text{lin}) = \frac{W}{L} \mu_n |Q_n'| V_{DS}$



# Ableitung der Gleichstrom-Formeln



Geometrie des MOSFETs für die Ableitung der  $I_D$ - $V_{DS}$ -Relation.



Ladungsverteilung in einem n-Kanal Enhancement-Mode MOSFET für  $V_{GS} > V_T$ .

Nehmen an, dass die Bildung der Inversionsladung im Kanal durch das vertikale Gatefeld dominiert wird. Das elektrische Feld, das von der Drain-Source-Spannung hervorgerufen wird, sei um vieles kleiner. Man nennt diesen Ansatz *gradual-channel approximation*.

## Annahmen:

1. Der Ladungstransport erfolgt nur durch Drift (keine Diffusion).
2. Das elektrische Feld senkrecht zum Kanal dominiert: *gradual channel approximation*. ( $E_x$  ist näherungsweise konstant.)
3. Das Oxid ist ein idealer Isolator (kein Leckstrom durch das Gateoxid).
4. Alle Oxidladungen werden zu einer äquivalenten Flächenladung  $Q'_{ss}$  zusammengefasst.
5. Konstante Kanal-Beweglichkeit.
6. n-Kanal MOSFET.

Wir beginnen mit dem Ohmschen Gesetz  $J_x = \sigma E_x = e \mu_n n(y) E_x$

mit der Elektronenkonzentration  $n(y)$  am Ort  $x$  im Kanal. Der totale Strom wird durch Integration der Stromdichte in den  $y$ - $z$ -Richtungen erhalten:

$$I_x = \iint_{yz} J_x dydz$$

Definieren wir die *Inversionsladung* pro Fläche als

$$Q'_n = - \int en(y) dy$$

so bekommen wir für den Strom

$$I_x = -W\mu_n Q'_n E_x$$

mit der Transistorbreite  $W$  durch Integration über  $z$ .

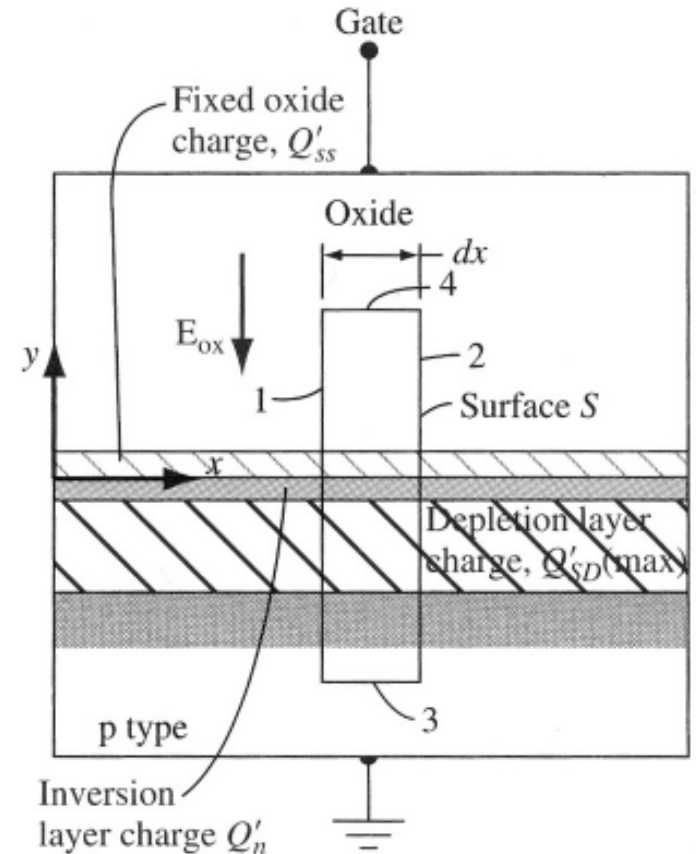
Globale Ladungsneutralität  $\rightarrow$

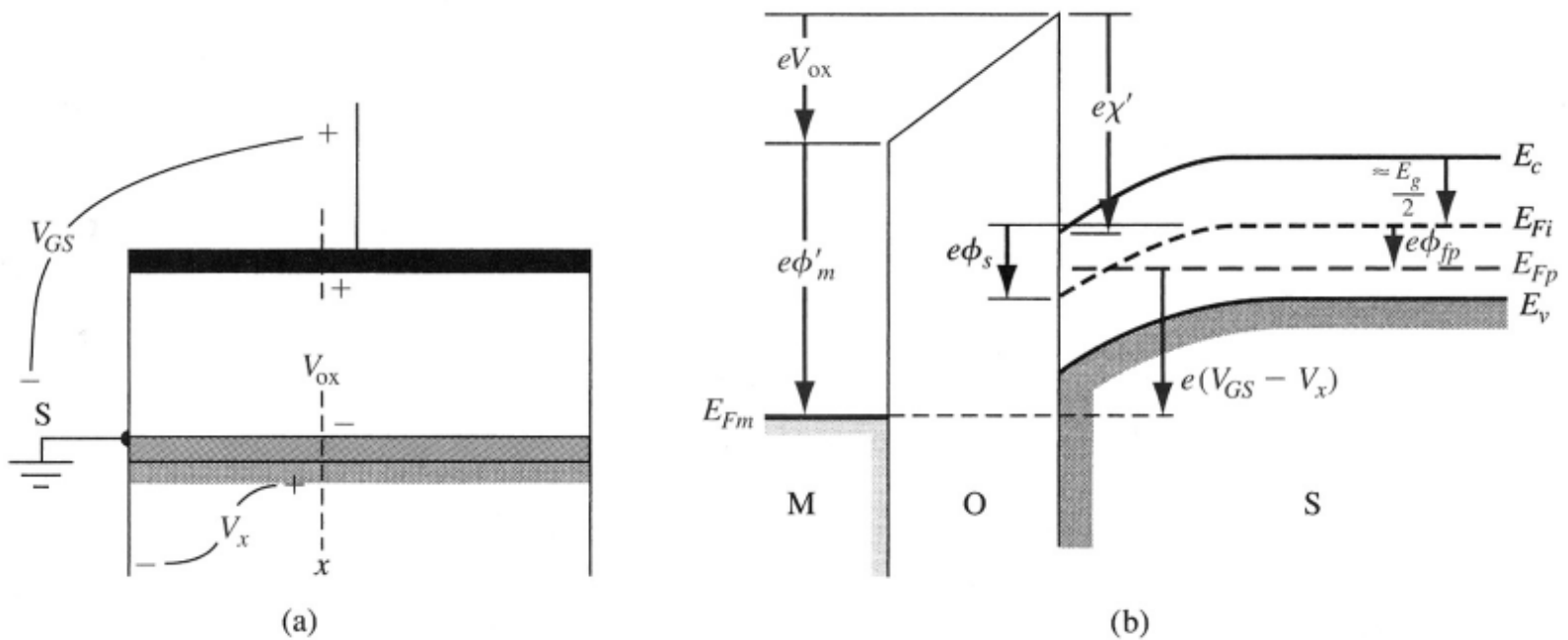
$$Q'_m + Q'_{SS} + Q'_n + Q'_{SD}(\max) = 0$$

wobei Inversionsladung und Substratladung negative Größen sind. Gauss'scher Satz über eine geschlossene Fläche im Bauelement:

$$\oint_S \epsilon E_n dS = Q_T = (Q'_{SS} + Q'_n + Q'_{SD}(\max))Wdx = -\epsilon_{ox} E_{ox} Wdx$$

(Front- und Rückfläche ergeben Null, da keine  $z$ -Komponente des  $E$ -Feldes; Beiträge von 1 und 2 kompensieren sich, da  $E_x \approx \text{const}$ ; 3 ergibt Null wegen neutralem  $p$ -Gebiet)





(a) Potentiale an einem Punkt  $x$  entlang des Kanals, (b) Bandkantenendiagramm quer durch die MOS-Struktur am Punkt  $x$ .

Die Differenz der Fermi-niveaus am Ort  $x$  im Kanal ist  $E_{Fp} - E_{Fm} = e(V_{GS} - V_x)$

Diese Spannungsdifferenz kann man ausdrücken durch

$$\begin{aligned}
 V_{GS} - V_x &= \Phi'_m + V_{ox} - \left[ \chi' + \frac{E_g}{2e} - \Phi_s + \Phi_{fp} \right] \\
 &= V_{ox} + \underbrace{\Phi'_m - \left[ \chi' + \frac{E_g}{2e} + \Phi_{fp} - \Phi_s \right]}_{\Phi_{ms}} \approx 2\Phi_{fp}
 \end{aligned}$$

Nach Einsetzen der Inversion bleibt das Oberflächenpotential näherungsweise auf diesem Wert!

$$V_{GS} - V_x = V_{ox} + \Phi_{ms} + 2\Phi_{fp}$$

Da  $E_{ox} = V_{ox}/t_{ox}$ , wird aus dieser Gleichung

$$\begin{aligned} -\varepsilon_{ox} E_{ox} &= -\frac{\varepsilon_{ox}}{t_{ox}} [V_{GS} - V_x - (\Phi_{ms} + 2\Phi_{fp})] \\ &= Q'_{ss} + Q'_n + Q'_{SD}(\max) \end{aligned}$$

Siehe zwei Folien zurück

Auflösung nach der Inversionsladung  $Q'_n$  ergibt

$$Q'_n = -C_{ox} \left[ V_{GS} - V_x - \underbrace{\left( \Phi_{ms} + 2\Phi_{fp} - \frac{Q'_{ss}}{C_{ox}} - \frac{Q'_{SD}(\max)}{C_{ox}} \right)}_{V_T} \right]$$

Die Inversionsladung  $Q'_n$  kann nun in  $I_x = W\mu_n Q'_n E_x$  eingesetzt werden:

$$I_x = -W\mu_n C_{ox} \frac{dV_x}{dx} (V_{GS} - V_x - V_T)$$

Integration über die gesamte Kanallänge:

$$\int_0^L I_x dx = -W\mu_n C_{ox} \int_{V_x(0)}^{V_x(L)} dV_x (V_{GS} - V_T - V_x)$$

Der Strom ist entlang des Kanals konstant  $I_x = -I_D$  und  $V_x(0) = 0$ . Integration ergibt

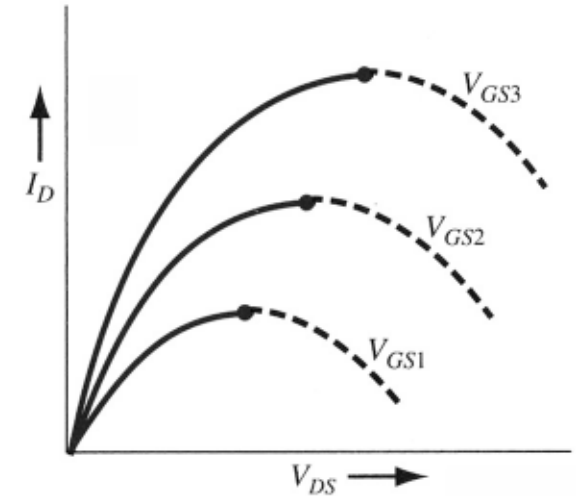
$$I_D = \frac{W\mu_n C_{ox}}{2L} [2(V_{GS} - V_T)V_{DS} - V_{DS}^2]$$

Diese Formel gilt für  $V_{GS} \geq V_T$  und  $0 \leq V_{DS} \leq V_{DS}(\text{sat})$ . Die Sättigungsspannung ist diejenige Spannung, bei der  $I_D$  maximal wird:

$$V_{DS}(\text{sat}) = V_{GS} - V_T$$

Für  $V_{DS} \geq V_{DS}(\text{sat})$  wird der ideale Drainstrom konstant und gleich dem Sättigungsstrom

$$I_D(\text{sat}) = \frac{W\mu_n C_{ox}}{2L} (V_{GS} - V_T)^2 = \frac{W\mu_n C_{ox}}{2L} V_{DS}^2(\text{sat})$$



Auswertung der untenstehenden  $I_D$ - $V_{DS}$ -Relation.

Im Falle des p-Kanal MOSFETs erhält man

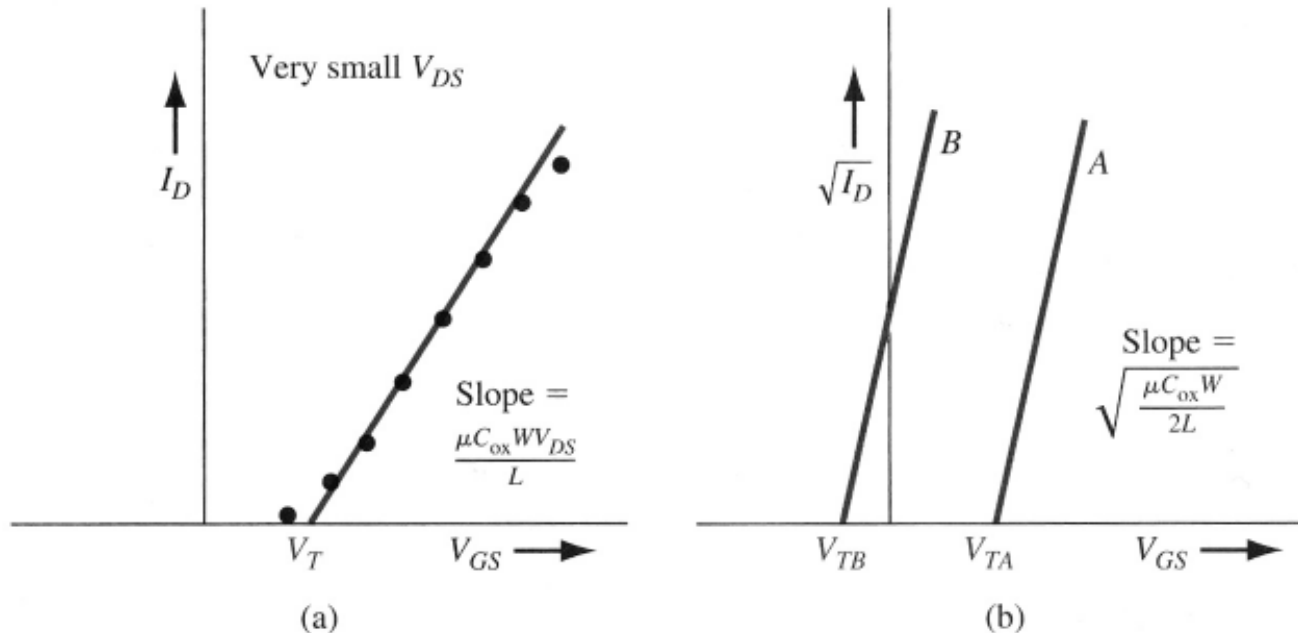
$$I_D = \frac{W\mu_p C_{ox}}{2L} [2(V_{GS} + V_T)V_{DS} - V_{DS}^2]$$

Man beachte, dass hier  $V_T < 0$ .

# Bestimmung der Schwellspannung

Sowohl der Bereich sehr kleiner Drainspannungen als auch der Sättigungsbereich können benutzt werden, um die Schwellspannung zu extrahieren. Der Anstieg von  $I_D(V_{GS})$  liefert die Kanal-Beweglichkeit.

$$I_D = \frac{W\mu_n C_{ox}}{L} (V_{GS} - V_T) V_{DS} \qquad \sqrt{I_D(\text{sat})} = \sqrt{\frac{W\mu_n C_{ox}}{2L}} (V_{GS} - V_T)$$



(a)  $I_D$ - $V_{GS}$  (für kleine  $V_{DS}$ ) bei einem Enhancement-Mode MOSFET. (b) Ideale  $I_D^{1/2}$ - $V_{GS}$ -Kurve im Sättigungsbereich für einen Enhancement-Mode (A) und Depletion-Mode (B) n-Kanal-MOSFET.

# Transconductance (Gegenwirkleitwert)

$$g_m = \frac{\partial I_D}{\partial V_{GS}}$$

Die Transconductance ist ein wichtiger technischer Parameter des MOSFETs. Im Nichtsättigungsbereich steigt  $g_m$  linear mit  $V_{DS}$  an:

$$g_{mL} = \frac{W\mu_n C_{ox}}{L} V_{DS}$$

Im Sättigungsbereich ist  $g_m$  eine lineare Funktion von  $V_{GS}$  und unabhängig von  $V_{DS}$ :

$$g_{ms} = \frac{\partial I_D(\text{sat})}{\partial V_{GS}} = \frac{W\mu_n C_{ox}}{L} (V_{GS} - V_T)$$

Bei der Angabe von Strömen und  $g_m$  dividiert man durch die Transistorbreite  $W$ . Ströme gibt man also in  $A/\mu\text{m}$  an,  $g_m$  in  $\text{mS}/\mu\text{m}$ .

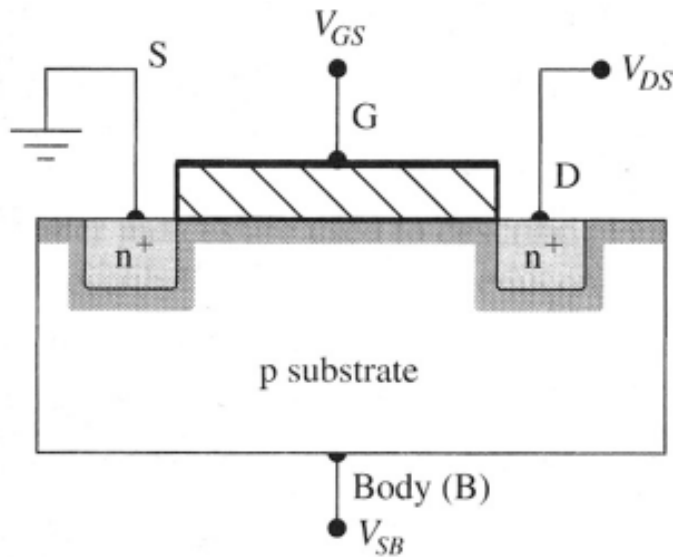


# Einfluss der Substratspannung

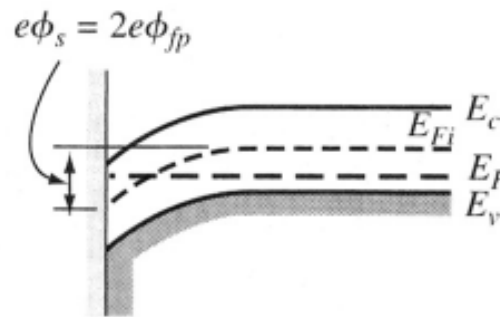
Substratspannung (Body, Bulk) kann unabhängig gewählt werden, aber nur  $V_{SB} \geq 0$  ist sinnvoll.

Die RLZ-Weite unter dem Gate wird grösser. Die Änderung der negativen Raumladung durch die Substratspannung nach Einsetzen der Inversion ist

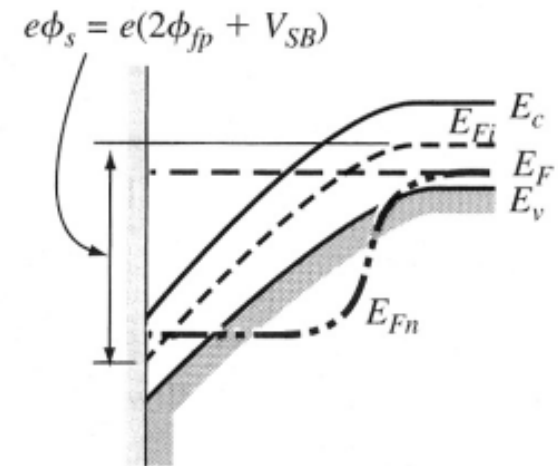
$$\Delta Q'_{SD} = -\sqrt{2 \epsilon_s e N_a} \left( \sqrt{2 \Phi_{fp} + V_{SB}} - \sqrt{2 \Phi_{fp}} \right)$$



(a)



(b)



(c)

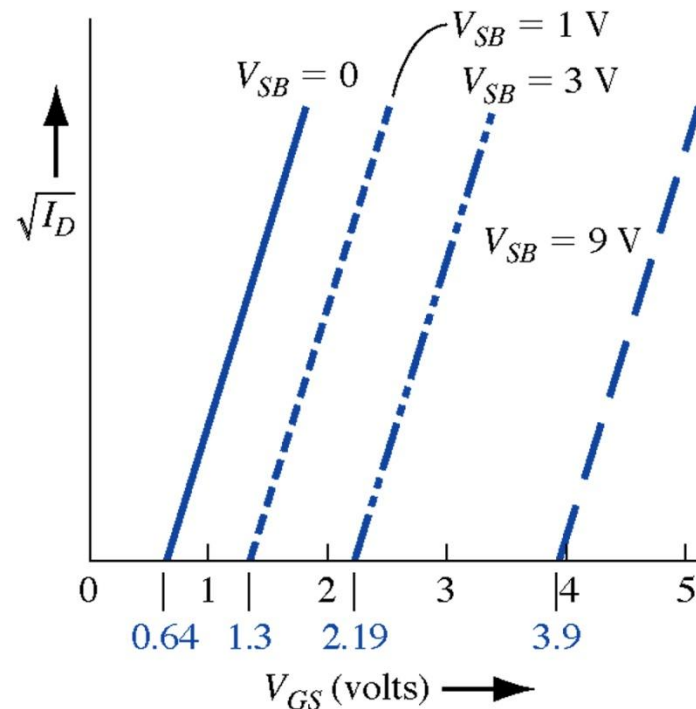
(a) Angelegte Spannungen an einem n-Kanal-MOSFET, (b) Bandkantendiagramm am Inversionspunkt, wenn  $V_{SB} = 0$  ist. (c) Bandkantendiagramm am Inversionspunkt, wenn  $V_{SB} > 0$  ist.

## Einfluss der Substratspannung (2)

Um die zusätzliche Raumladung zu kompensieren, muss eine grössere Spannung am Gate angelegt werden. Die (positive) Schwellspannungsänderung ergibt sich zu

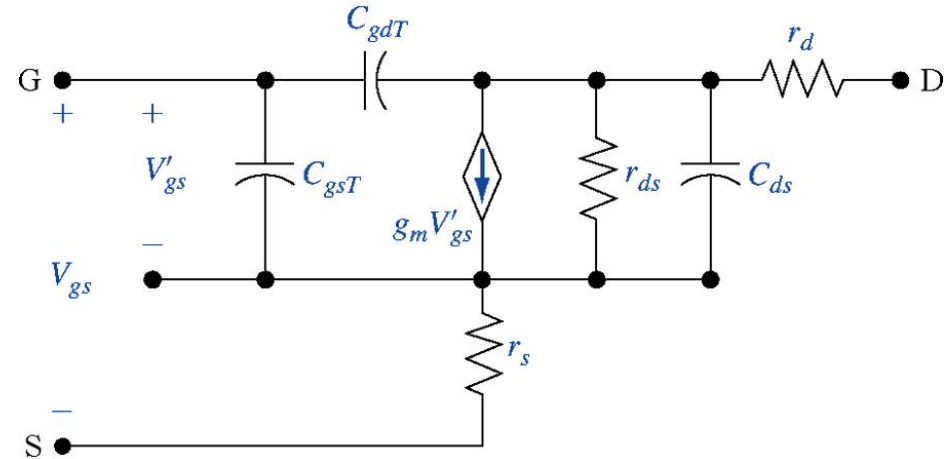
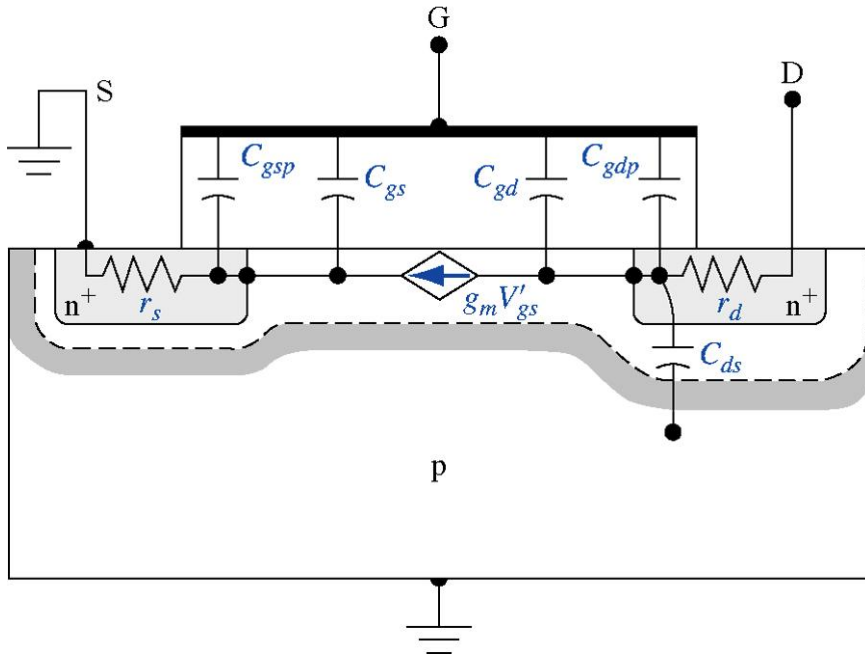
$$\Delta V_T = - \frac{\Delta Q'_{SD}}{C_{ox}} = \frac{\sqrt{2 \epsilon_s e N_a}}{C_{ox}} \left[ \sqrt{2 \Phi_{fp} + V_{SB}} - \sqrt{2 \Phi_{fp}} \right]$$

siehe MOS-Diode  
"Schwellspannung"

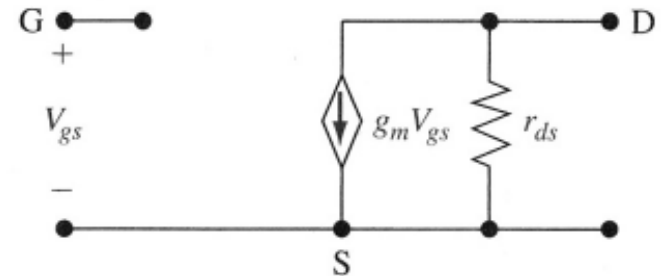


# 12.2 Das Frequenz-Verhalten

## Die Kleinsignal-Ersatzschaltung



- $C_{gs, gd}$  = Kapazitäten zwischen Gate- und Kanal-ladung (wie in MOS-Diode)
- $C_{gsp, gdp}$  = Überlapp-Kapazitäten zu den LDD-Gebieten an den Oxidkanten
- $I_d = g_m V'_{gs}$
- Der Widerstand  $r_{ds}$  kommt vom Anstieg der  $I_D(V_{DS})$ -Kurve (idealerweise ist  $r_{ds}$  unendlich).



Vereinfachtes Niedrigfrequenz-Ersatzschaltbild.

## Die Kleinsignal-Ersatzschaltung (2)

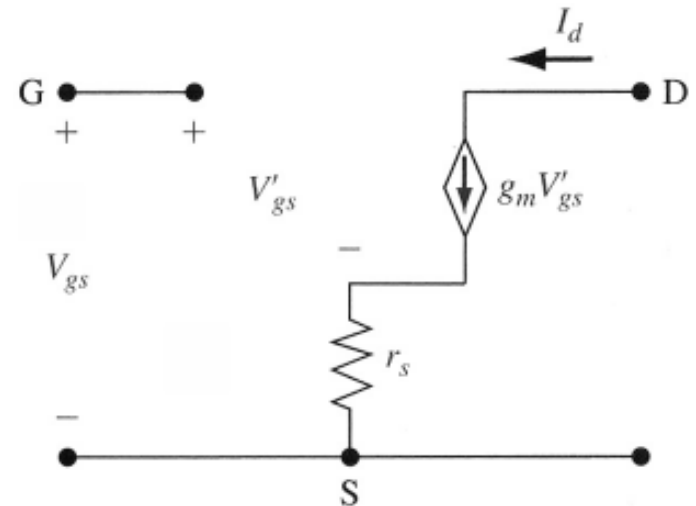
Der Source-Widerstand  $r_s$  kann erheblichen Einfluss haben. Vernachlässigt man stattdessen  $r_{ds}$ , so folgt

$$\begin{aligned}V_{gs} &= V'_{gs} + (g_m V'_{gs}) r_s \\ &= (1 + g_m r_s) V'_{gs}\end{aligned}$$

Damit ergibt sich für den Drainstrom

$$\begin{aligned}I_d &= g_m V'_{gs} \\ &= \left( \frac{g_m}{1 + g_m r_s} \right) V_{gs} \\ &= g'_m V_{gs}\end{aligned}$$

Der Source-Widerstand reduziert die effektive Transconductance und den Gain des Transistors.



Vereinfachtes Niedrigfrequenz-Ersatzschaltbild mit Source-Widerstand  $r_s$ .

# Transitzeit und Eingangsstrom

Intrinsische Verzögerung durch Transitzeit im Kanal:  $\tau_t = \frac{L}{v_{\text{sat}}}$

Beispiel:  $L = 0.1 \mu\text{m}$ ,  $v_{\text{sat}} = 10^7 \text{ cm/s}$   $\tau_t = 1 \text{ ps}$

Die Transitzeit ist also *nicht* der begrenzende Faktor für das Frequenz-Verhalten. Entscheidend ist die Gate- oder Kapazitäts-Ladezeit.

Strom am Eingang:

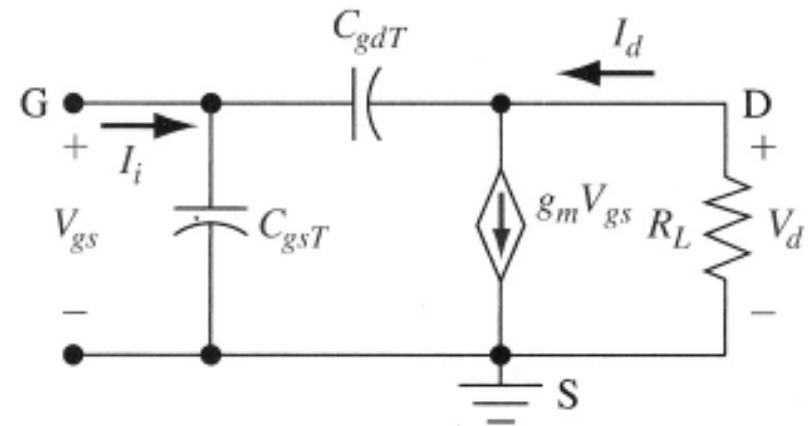
$$I_i = j\omega C_{\text{gsT}} V_{\text{gs}} + j\omega C_{\text{gdT}} (V_{\text{gs}} - V_d)$$

Strombilanz am Ausgang:

$$\frac{V_d}{R_L} + g_m V_{\text{gs}} + j\omega C_{\text{gdT}} (V_d - V_{\text{gs}}) = 0$$

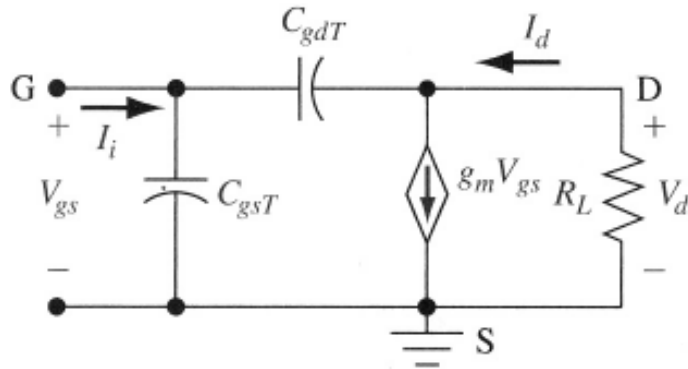
Eliminierung von  $V_d$ :

$$I_i = j\omega \left\{ C_{\text{gsT}} + C_{\text{gdT}} \left[ \frac{1 + g_m R_L}{1 + j\omega R_L C_{\text{gdT}}} \right] \right\} V_{\text{gs}}$$

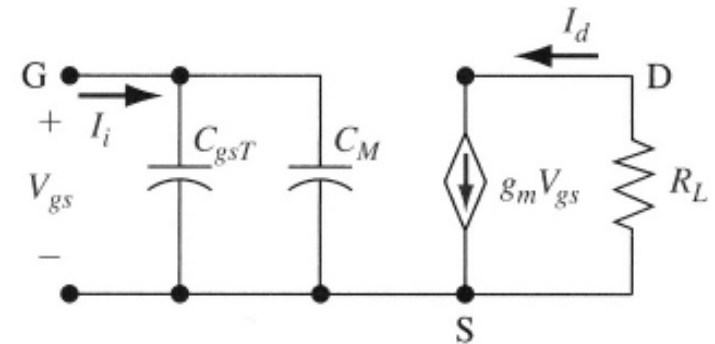


Vereinfachtes Hochfrequenz-Ersatzschaltbild.

## Die Miller-Kapazität



Vereinfachtes Hochfrequenz-Kleinsignal-Ersatzschaltbild für den Common-Source n-Kanal MOSFET.



Kleinsignal-Ersatzschaltbild mit Miller-Kapazität.

Normalerweise ist  $\omega R_L C_{gdT} \ll 1$ , daher kann der Term im Nenner vernachlässigt werden

$$\begin{aligned} I_i &= j\omega \left\{ C_{gsT} + C_{gdT} \left( 1 + g_m R_L \right) \right\} V_{gs} \\ &= j\omega \left[ C_{gsT} + C_M \right] V_{gs} \end{aligned}$$

$C_M$  ist die *Miller-Kapazität*:  $C_M = C_{gdT} \left( 1 + g_m R_L \right)$

Im Sättigungsbereich verschwindet  $C_{gd}$ , aber die Überlapp-Kapazität  $C_{gdp}$  wird konstant und, mit dem Gain multipliziert, ein wichtiger Beitrag zur Eingangsimpedanz.

# Die Grenzfrequenz

Die Grenzfrequenz  $f_T$  ist definiert als jene Frequenz, bei der die Stromverstärkung eins wird und damit der Betrag des Eingangsstromes gleich jenem des Ausgangsstromes:

$$\left. \begin{array}{l} I_i = j\omega (C_{gsT} + C_M) V_{gs} \\ I_d = g_m V_{gs} \end{array} \right\} \left| \frac{I_d}{I_i} \right| = \frac{g_m}{2\pi f (C_{gsT} + C_M)} \rightarrow 1$$

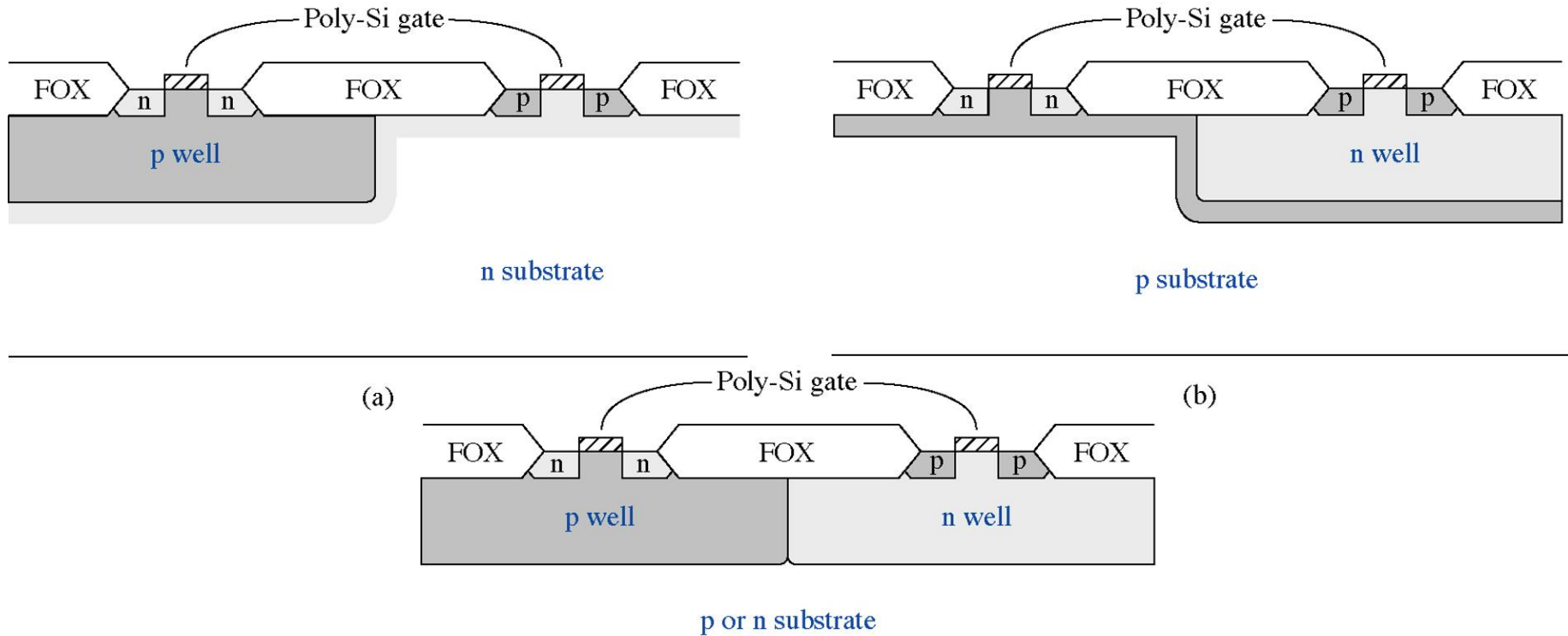
$$f_T = \frac{g_m}{2\pi (C_{gsT} + C_M)} = \frac{g_m}{2\pi C_G} \leftarrow \text{äquivalente Gate-Kapazität}$$

Für den idealen MOSFET in Sättigung kann man annehmen, dass  $C_{gsp} = C_{gdp} = 0$ ,  $C_{gd} \rightarrow 0$  (wegen pinch-off) und  $C_{gs} \rightarrow C_{ox} WL$ . Die Transconductance ist (Folie 319):

$$g_{ms} = W\mu_n C_{ox} (V_{gs} - V_T)/L$$

$$f_T = \frac{g_m}{2\pi C_G} = \frac{W\mu_n C_{ox} (V_{gs} - V_T)/L}{2\pi C_{ox} WL} = \frac{\mu_n (V_{gs} - V_T)}{2\pi L^2}$$

# 12.3 CMOS-Technologie

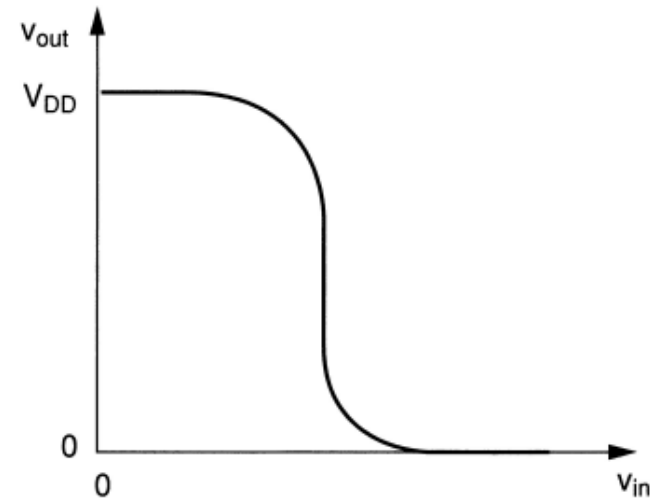
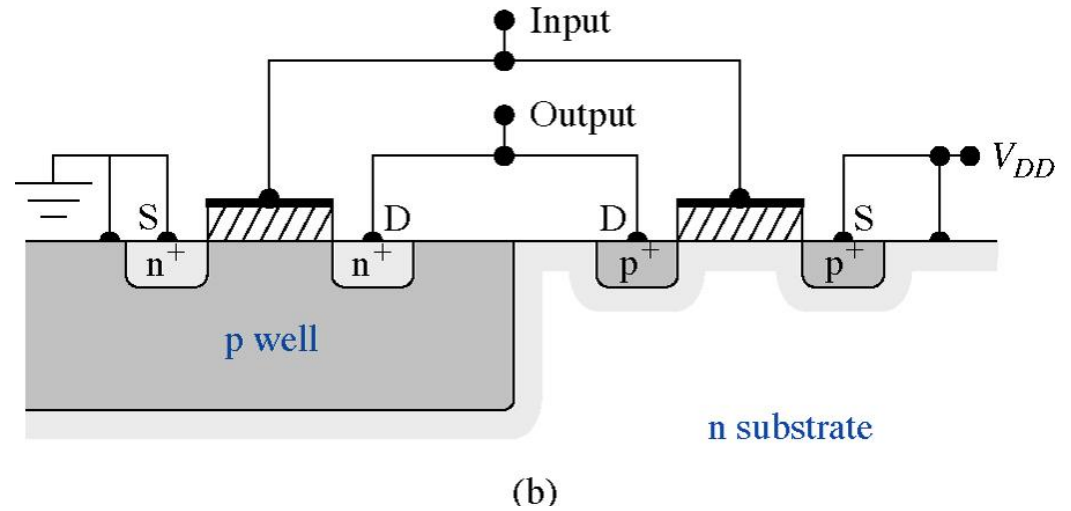
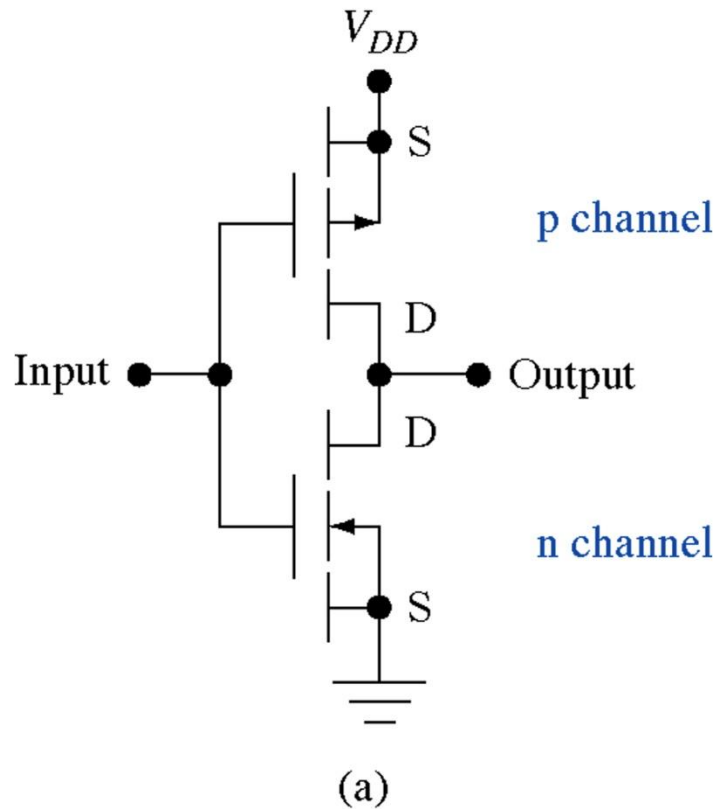


- Isolation zwischen n- und p-Kanal MOSFETs durch dickes “Feldoxid”.
- Well-Substrat- bzw. p-well-n-well-Übergänge müssen gesperrt sein.
- Schwellspannungsadjustierung durch Einstellung der Dosis bei Implantation.

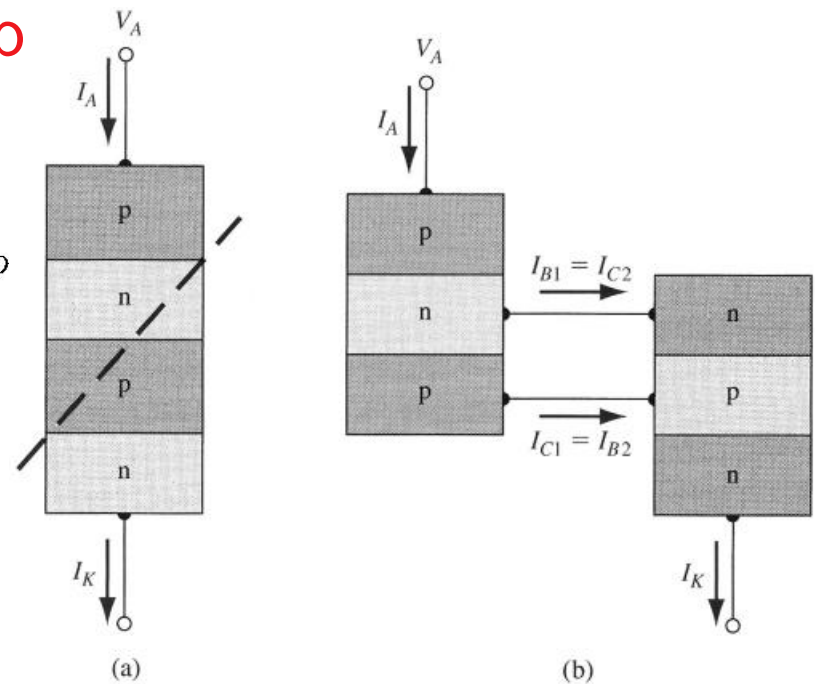
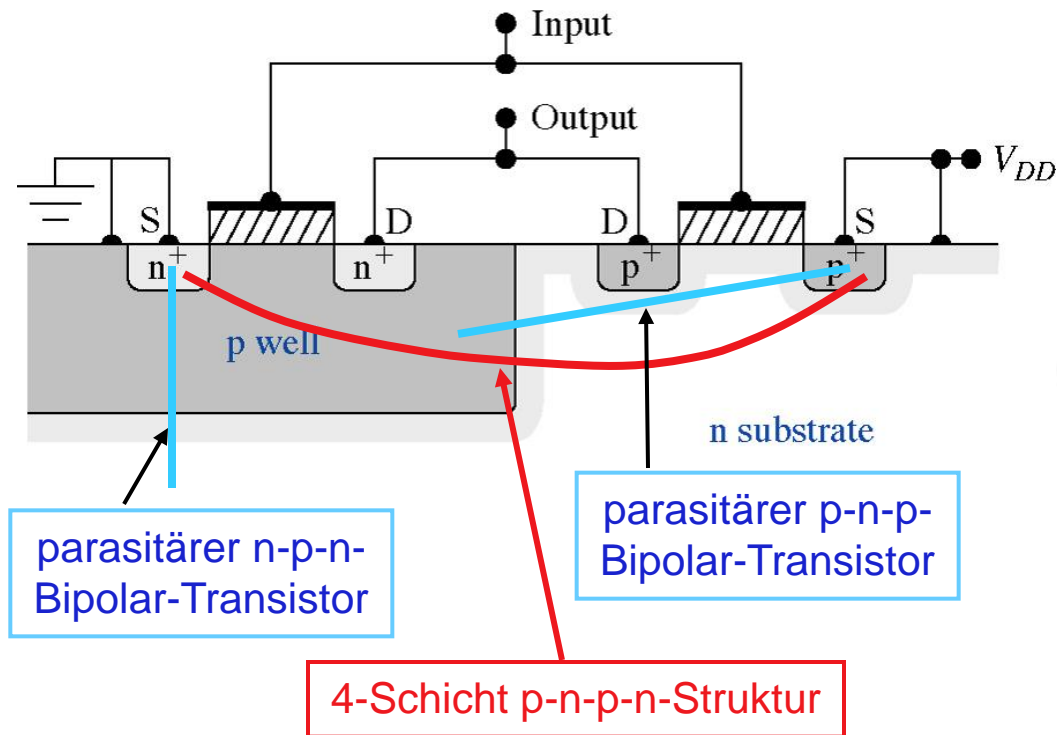


# Der CMOS-Inverter

Der CMOS-Inverter ist ein Grundbaustein der digitalen Logik-Schaltkreise.



# Latch-up



(a) Aufspaltung der ursprünglichen pnpn-Struktur, (b) 2-Transistor-Ersatzschaltkreis des 4-schichtigen pnpn-Bauelements.

- Im Normalzustand sind beide parasitären Bipolar-Transistoren im “cut off”.
- Kommt es allerdings zum Lawinen-Durchbruch im gesperrten p-well-n-Substrat-Übergang, können beide Bipolar-Transistoren in die Sättigung getrieben werden, da der grosse Kollektorstrom in die Basis des jeweiligen anderen Transistors fließt.
- Produkt der Stromverstärkungen beider Bipolar-Transistoren muss kleiner 1 gemacht werden, z.B. durch “Lebensdauer-Killer” (tiefe Störstellen).







# 13. Der MOSFET (2)

**13.1 Das nichtideale Verhalten des MOSFETs (N12.1)**

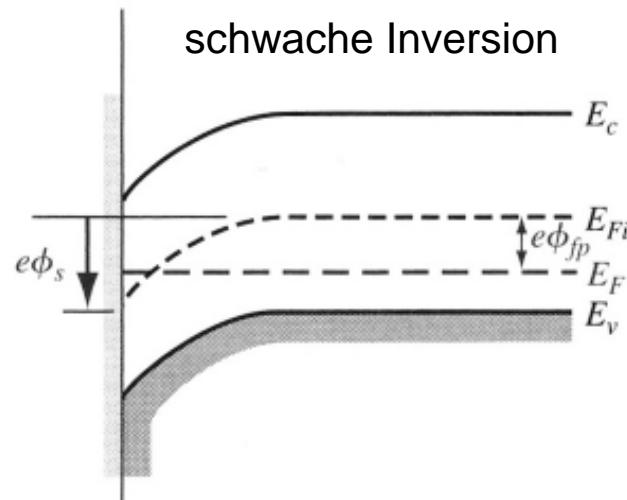
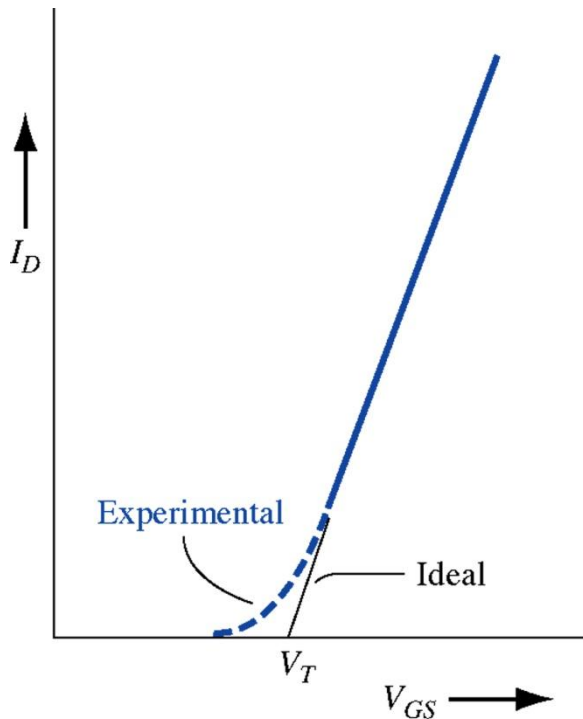
**13.2 Schwellspannungseffekte\* (N12.3)**

**13.3 Weitere wichtige Effekte\* (N12.4-5)**

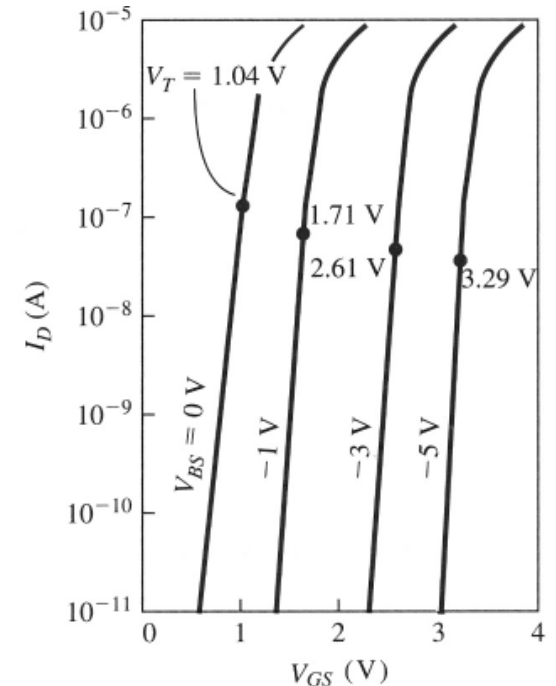
# 13.1 Das nichtideale Verhalten des MOSFETs

## Leitfähigkeit unterhalb der Schwellspannung (Subthreshold-Strom)

Der Strom, der im realen MOSFET bei Gatespannungen  $V_{GS} < V_T$  fließt, heisst Subthreshold-Strom. Er wird durch Diffusion (thermionische Emission) verursacht.

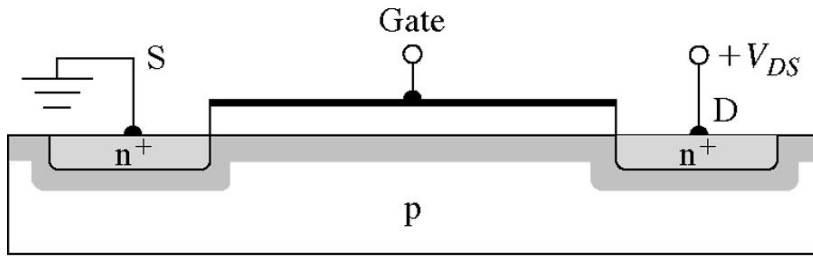


Bandkantendiagramm für  $\Phi_{fp} < \Phi_s < 2\Phi_{fp}$ .



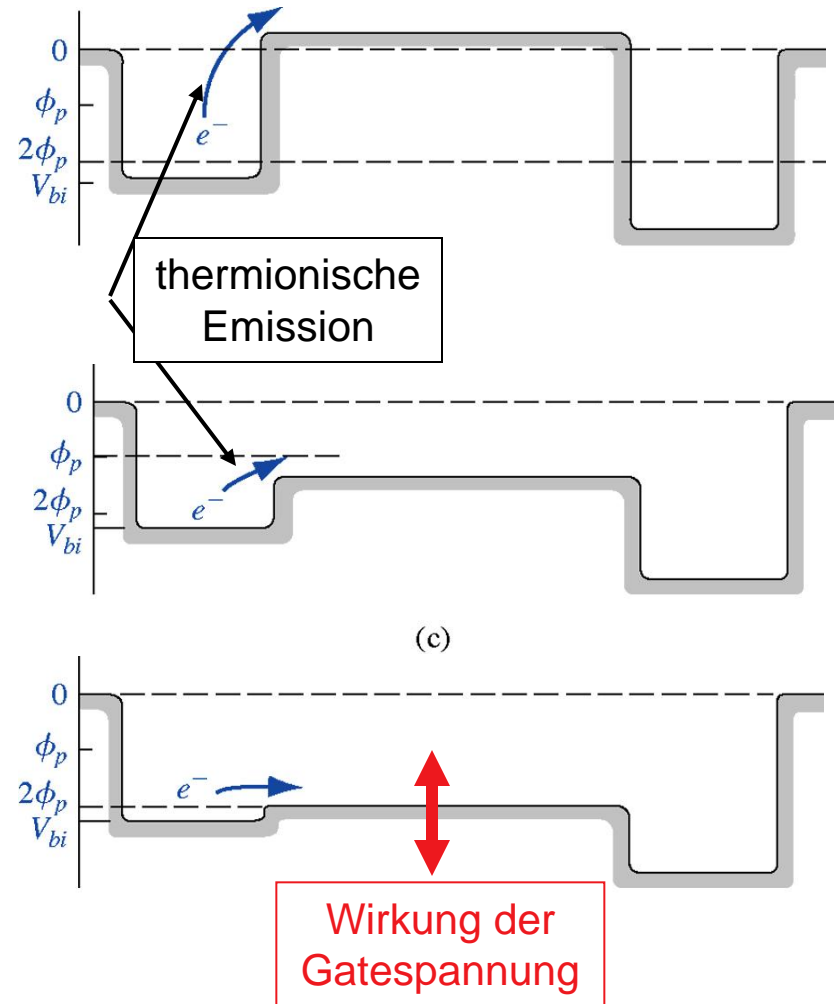
Subthreshold  $I_D$ - $V_{GS}$ -Kurven für verschiedene Werte der Substratspannung.  $V_T$  ist jeweils explizit angegeben (nach Schroder).

## Subthreshold-Strom (2)



$$I_{D(\text{sub})} \sim \exp\left(\frac{eV_{GS}}{kT}\right) \left\{ 1 - \exp\left(-\frac{eV_{DS}}{kT}\right) \right\}$$

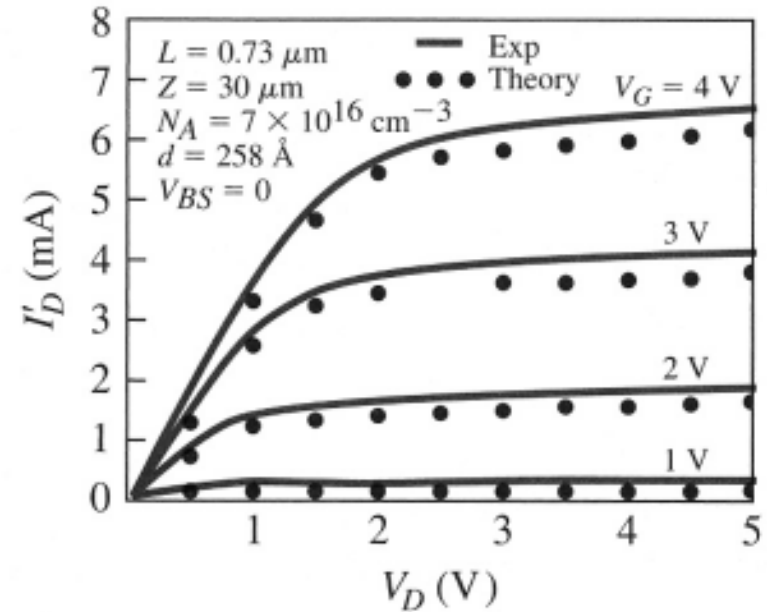
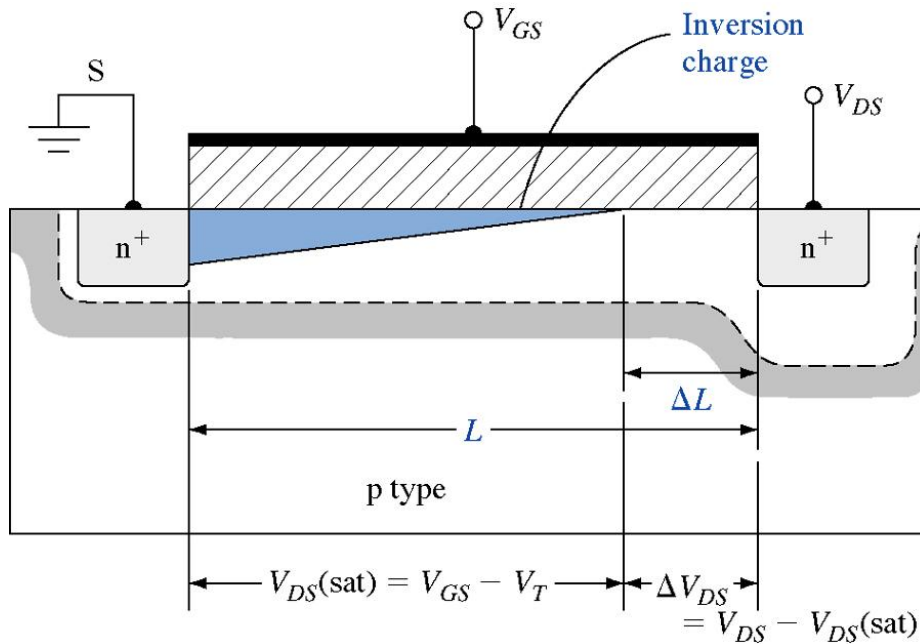
- Der Subthreshold-Strom ist (solange  $V_{GS} < V_T$ ) eine exponentielle Funktion der Gate-spannung. Elektronen können die Source-Drain-Barriere nur thermionisch überwinden.
- Idealerweise bewirkt eine Änderung der Gatespannung um 60 mV eine Änderung des Stroms um eine Größenordnung.
- Der Anstieg hängt von der Grenzflächenzustandsdichte ab.





# Die Kanallängen-Modulation

Im realen MOSFET wird bei Drainspannungen  $V_{DS} > V_{DS}(\text{sat})$  die Kanallänge kleiner.



Da der Drainstrom umgekehrt proportional zur Kanallänge ist, gilt

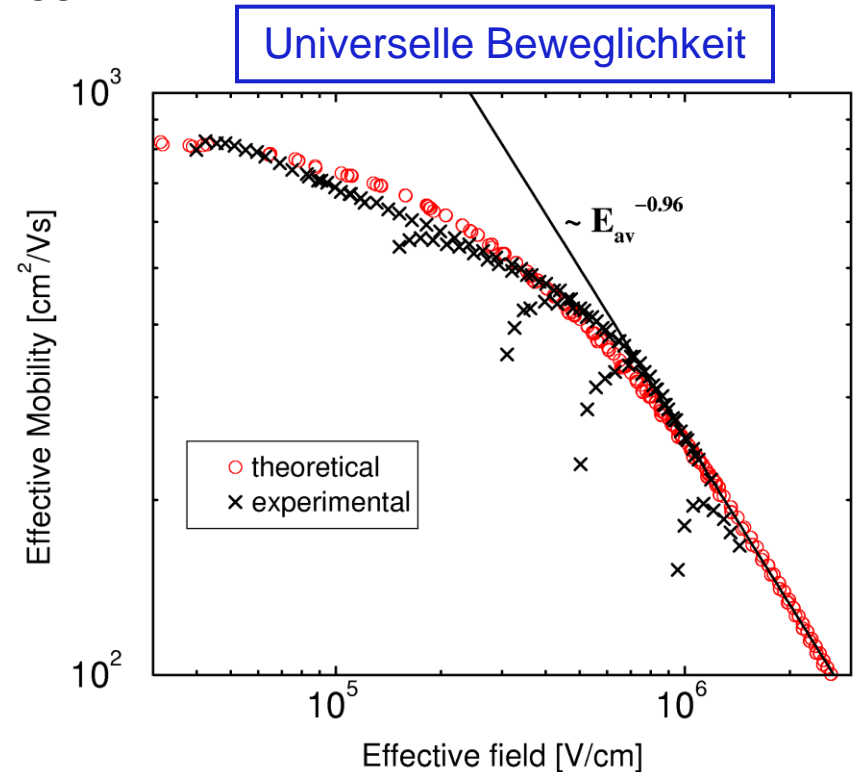
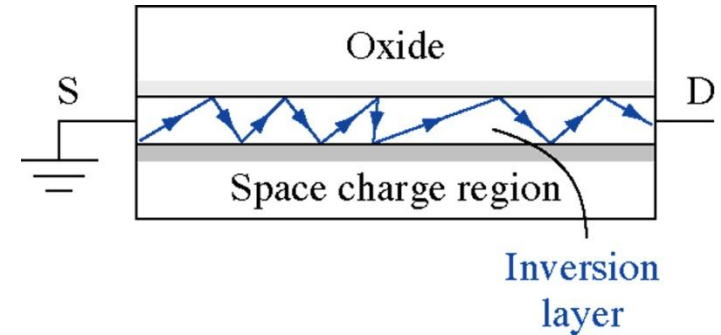
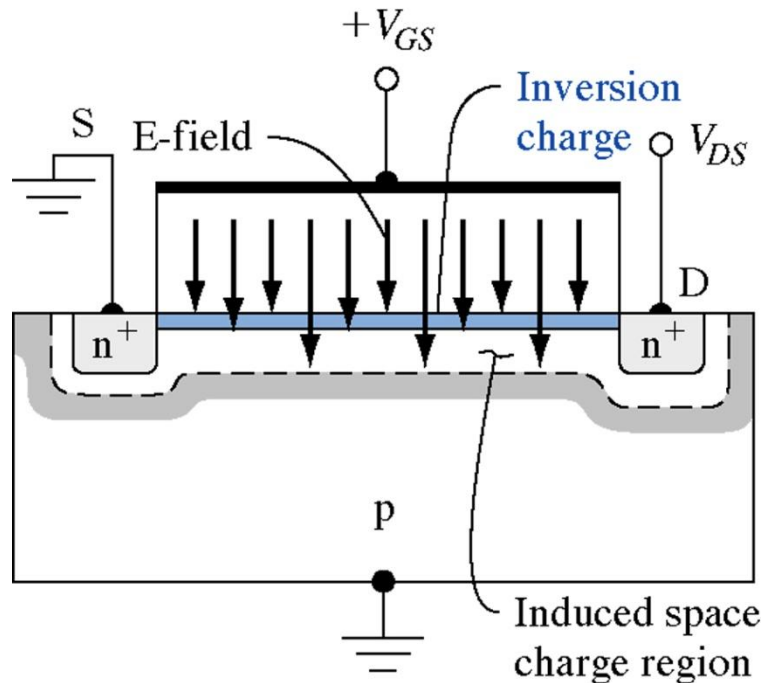
$$I'_D = \frac{L}{L - \Delta L} I_D$$

$$\Delta L = \sqrt{\frac{2 \epsilon_s}{e N_a}} \left[ \sqrt{\Phi_{\text{sat}} + V_{DS} - V_{DS}(\text{sat})} - \sqrt{\Phi_{\text{sat}}} \right]$$

# Die Reduktion der Kanal-Beweglichkeit

Die Beweglichkeit der Ladungsträger im Kanal ist eine starke Funktion der Gatespannung, weil

- 1) ihre Bewegung  $V_{GS}$ -abhängig eingeschränkt ist (2-dimensionales Trägergas)
- 2) Streuung an den Rauigkeiten der Si-SiO<sub>2</sub>-Grenzfläche stattfindet (je mehr, je grösser  $V_{GS}$ ).



# Die Sättigung der Driftgeschwindigkeit

Die Driftgeschwindigkeit sättigt ab  $E \approx 10^4$  V/cm. Da  $v_D = \mu E$ , muss die Beweglichkeit abnehmen. Grund:

Im stationären Zustand ex. Kräfte-Gleichgewicht aufgrund von Energiegewinn im longitudinalen elektrischen Feld und Energieverlust durch Emission von optischen Phononen (Generation von Wärme). Die Sättigungsdriftgeschwindigkeit beträgt für Elektronen in Bulk-Silizium etwa  $1 \times 10^7$  cm/s, im Kanal etwas weniger.

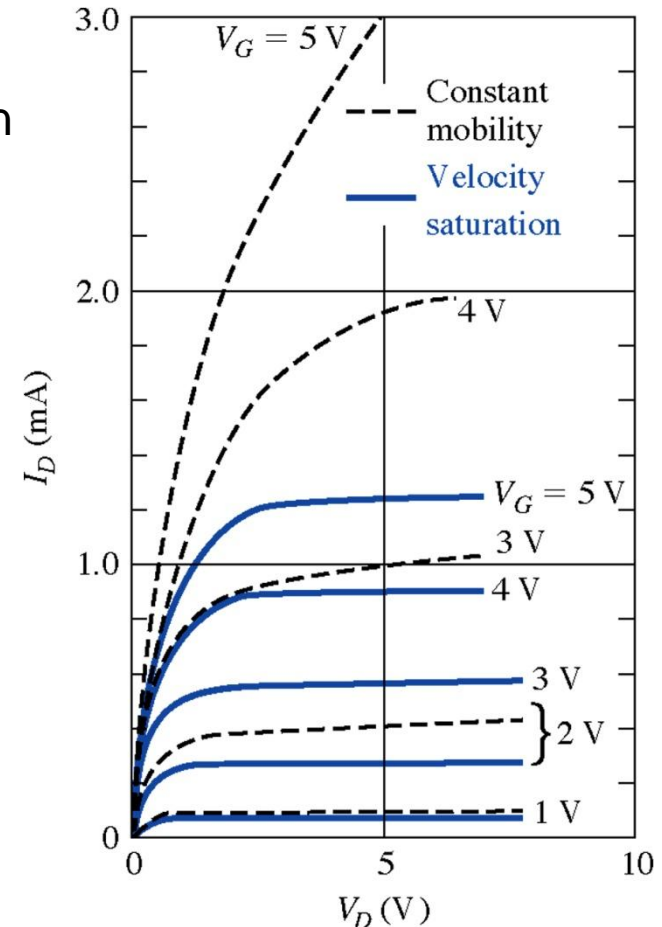
Beispiel: Kanallänge 60 nm,  $V_{DS} = 1.2$  V  $\rightarrow E = 2 \times 10^5$  V/cm.

Der Sättigungs-Drainstrom wird jetzt linear in  $V_{GS}$ :

$$I_D(\text{sat}) = WC_{\text{ox}} (V_{GS} - V_T) v_{\text{sat}}$$

Ein oft benutztes Fitmodell für die Beweglichkeit ist

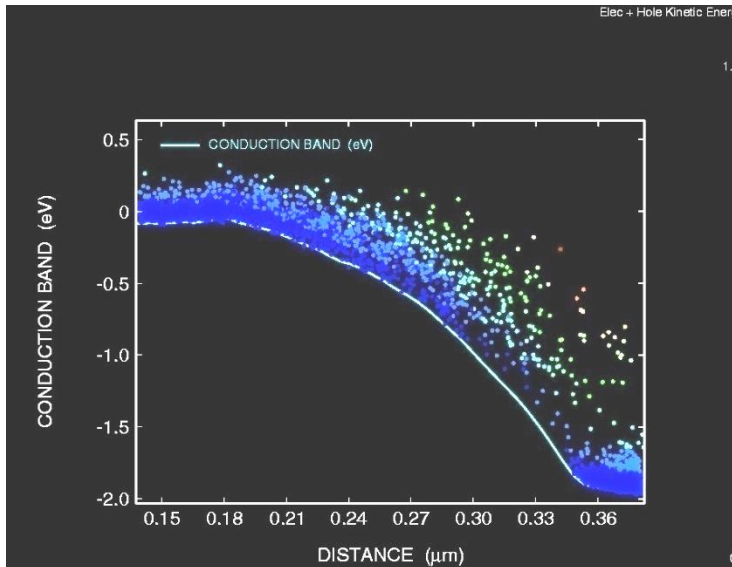
$$\mu = \frac{\mu_{\text{low}}}{\left\{ 1 + \left[ \frac{\mu_{\text{low}} E}{v_{\text{sat}}} \right]^\beta \right\}^{1/\beta}}$$



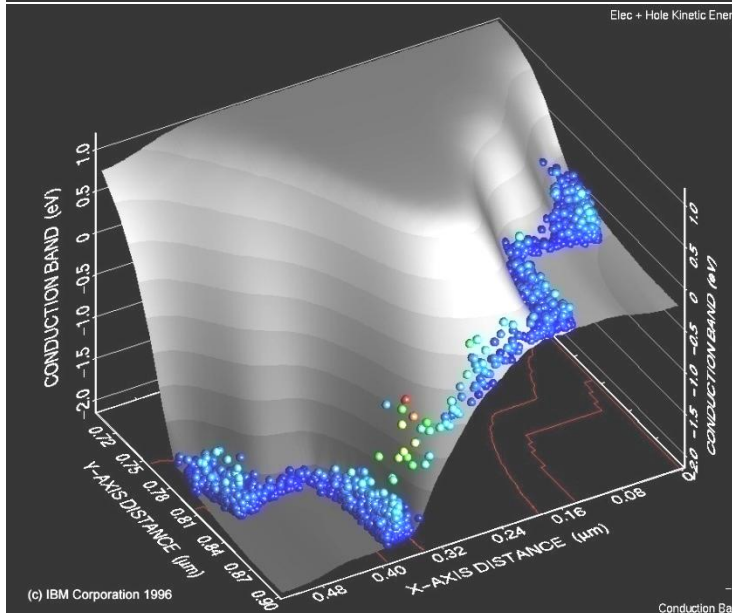
# Ballistischer Transport

Voraussetzung für Sättigung der Driftgeschwindigkeit: Kanallänge gross gegen mittlere freie Weglänge der Elektronen.

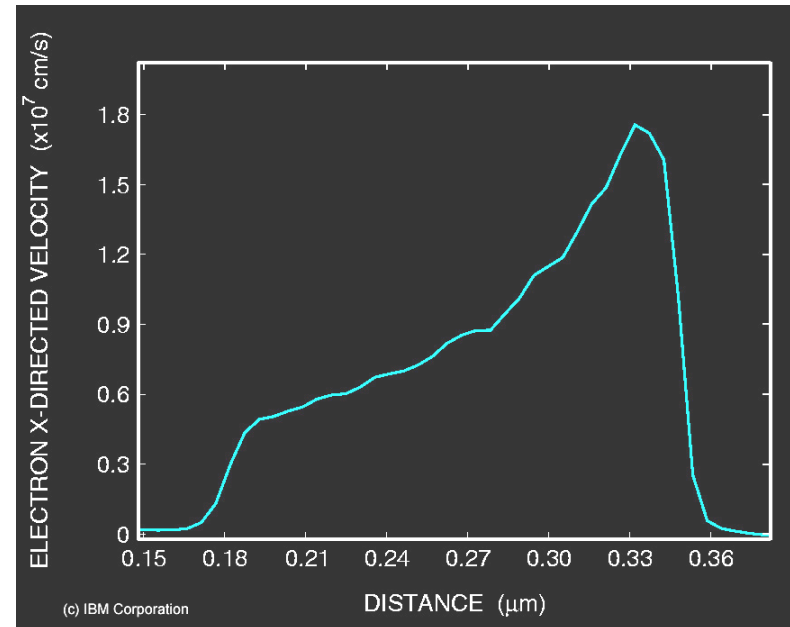
In Kurzkanal-MOSFETs nur noch wenige inelastische Streuungen → kein stationärer Zustand → velocity overshoot



(c) IBM Corporation 1996



(c) IBM Corporation 1996



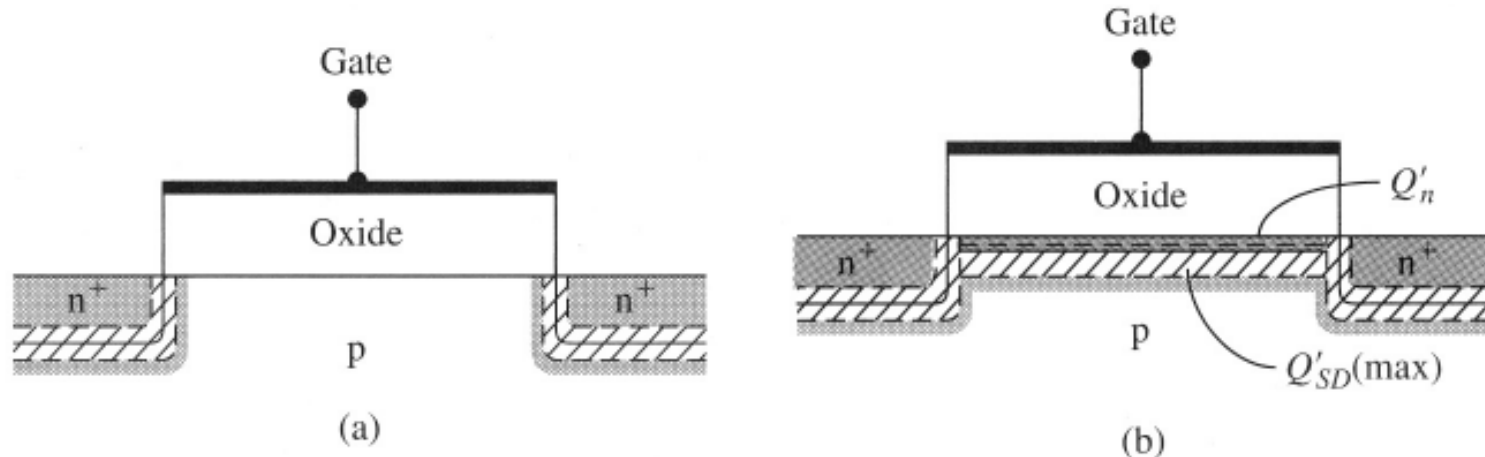
(c) IBM Corporation

“Velocity overshoot” am drainseitigen Ende des Kanals

# 13.2 Schwellspannungseffekte\*

## Kurzkanal-Effekte

In MOSFETs mit langem und weitem Kanal ist das Konzept der äquivalenten Flächenladungen sinnvoll, da weder Kanteneffekte noch der Einfluss der Source/Drain-Gebiete von Bedeutung sind. Auch kann man annehmen, dass die aktive Fläche im Substrat genauso gross ist, wie die Gatefläche.

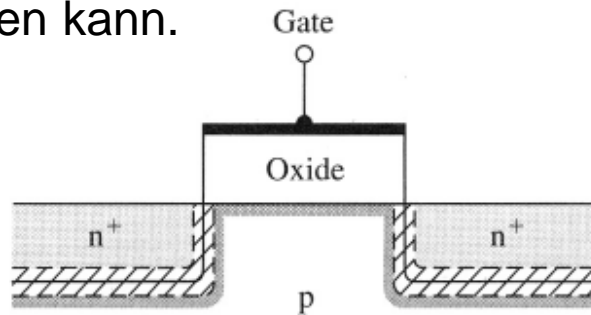


Querschnitt eines Langkanal-MOSFETs unter (a) Flachbandbedingungen und (b) Inversionsbedingungen.

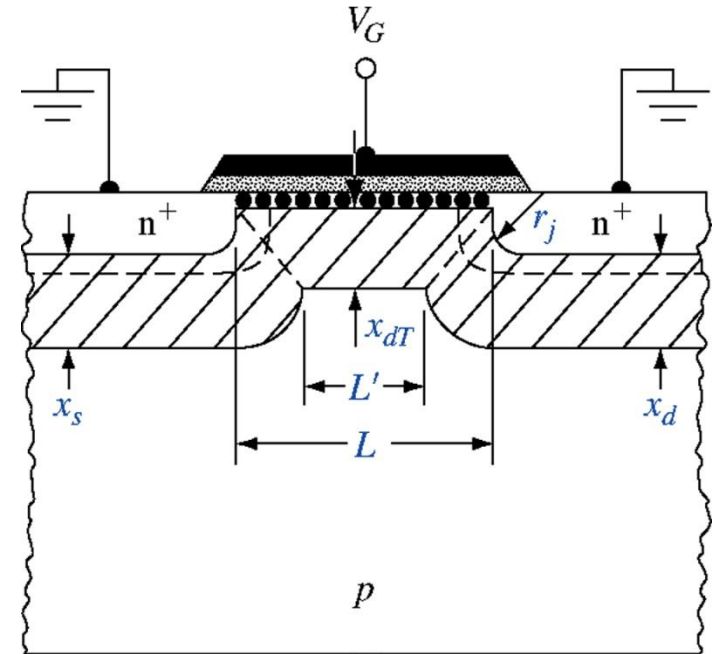
Die Gatespannung kontrolliert dann im wesentlichen die gesamte Raumladung, die in der Kanalregion induziert wird.

## Kurzkanal-Effekte (2)

Im Kurzkanal-MOSFET nimmt der gesperrte  $pn$ -Übergang zwischen Drain und Kanal immer mehr Platz in Anspruch, so dass immer weniger Bulk-Ladung vom Gate kontrolliert werden kann.



Querschnitt eines Kurzkanal-nMOSFETs unter Flachbandbedingungen.



Definition Schwellspannung

$$V_T = \Phi_{ms} + 2\Phi_{fp} - \frac{Q'_{ss}}{C_{ox}} + \frac{|Q'_{SD}(\max)|}{C_{ox}}$$

Annahme  $x_{dT} \approx x_d$ . Die totale Bulk-Ladung (in As) unter dem Gate im trapezförmigen Gebiet ist

$$|Q_B| = eN_a W x_{dT} \frac{L+L'}{2} \quad \text{und somit die mittlere Bulk-Ladung pro Fläche}$$

$$|Q'_B| = \frac{|Q_B|}{WL} = eN_a x_{dT} \frac{L+L'}{2L} = |Q'_{SD}(\max)|$$

## Kurzkanal-Effekte (3)

$$|Q'_B| = \frac{|Q_B|}{WL} = eN_a x_{dT} \frac{L+L'}{2L} = |Q'_{SD}(\max)|$$

Aufgrund der Geometrie kann  $L+L'/(2L)$  leicht berechnet werden (Satz des Pythagoras):

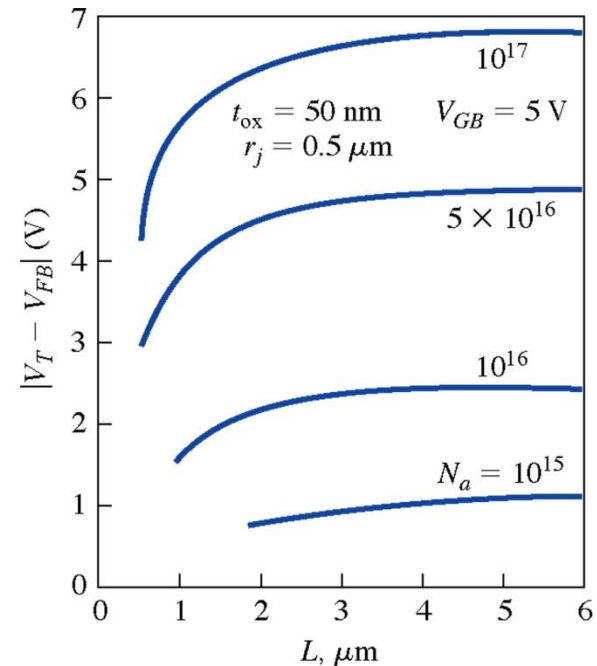
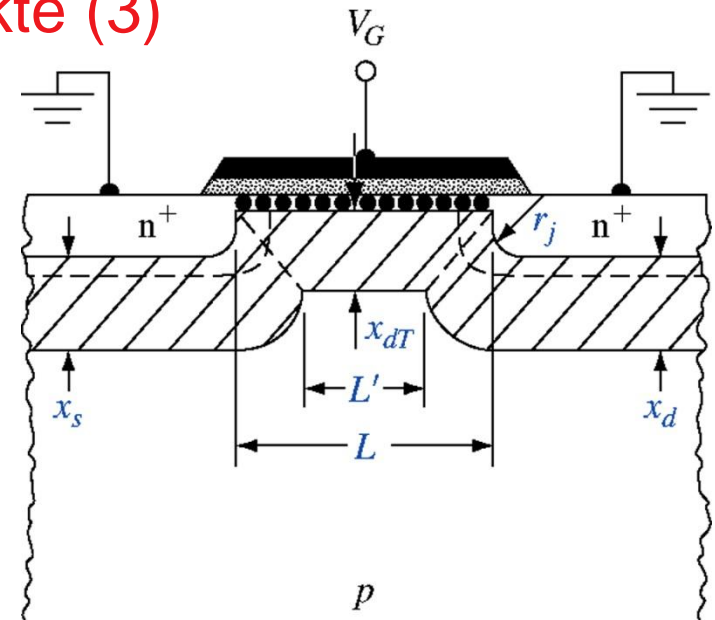
$$\frac{L+L'}{2L} = 1 - \frac{r_j}{L} \left( \sqrt{1 + \frac{2x_{dT}}{r_j}} - 1 \right)$$

$$|Q'_{SD}(\max)| = eN_a x_{dT} \left\{ 1 - \frac{r_j}{L} \left( \sqrt{1 + \frac{2x_{dT}}{r_j}} - 1 \right) \right\}$$

Man erhält eine *negative* Schwellspannungsverschiebung:

$$\Delta V_T = - \frac{eN_a x_{dT}}{C_{ox}} \frac{r_j}{L} \left( \sqrt{1 + \frac{2x_{dT}}{r_j}} - 1 \right)$$

Der Effekt wächst mit der Substratdotierung. Der Transistor verschiebt sich in Richtung Verarmungstyp. Die Source/Drain-Implantationen müssen möglichst steil sein, d.h. möglichst kleines  $r_j$ !

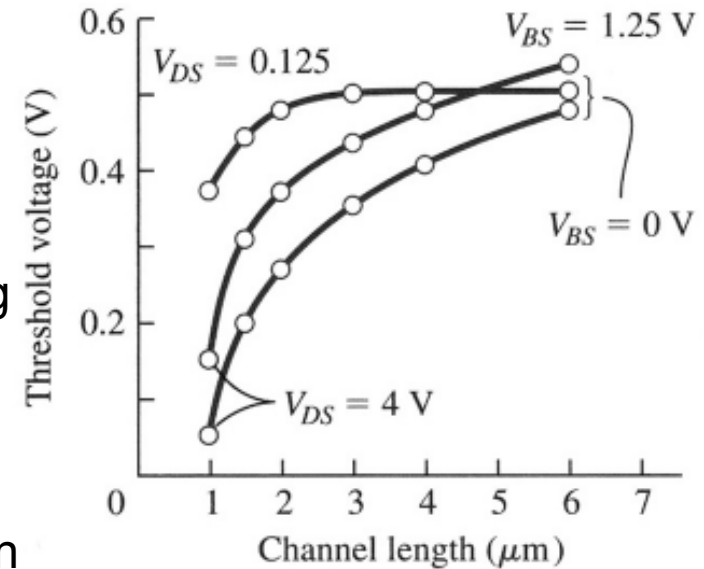


## Kurzkanal-Effekte (4)

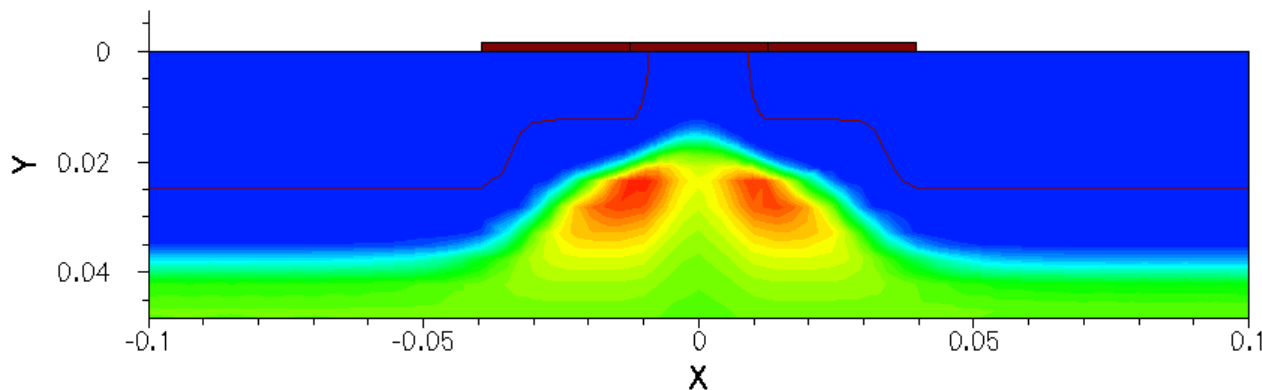
Die Erhöhung der Drainspannung bewirkt, dass  $L'$  noch kleiner wird und damit die Bulk-Ladung, die vom Gate kontrolliert werden kann.

*Die Schwellspannung wird eine Funktion der Drainspannung!* Der Drainkontakt beeinflusst die Steuerung des Transistors. Dieser wichtige Kurzkanal-Effekt heisst DIBL (Drain Induced Barrier Lowering).

Der DIBL ist ein 2D-Elektrostatik-Effekt. Um den Transistor möglichst "1D-like" zu machen, benutzt man in Bulk-MOSFETs standardmässig sogenannte HALO-Implantationen. Zu einer Alternative hat sich die SOI-Technologie entwickelt (Silicon On Insulator).

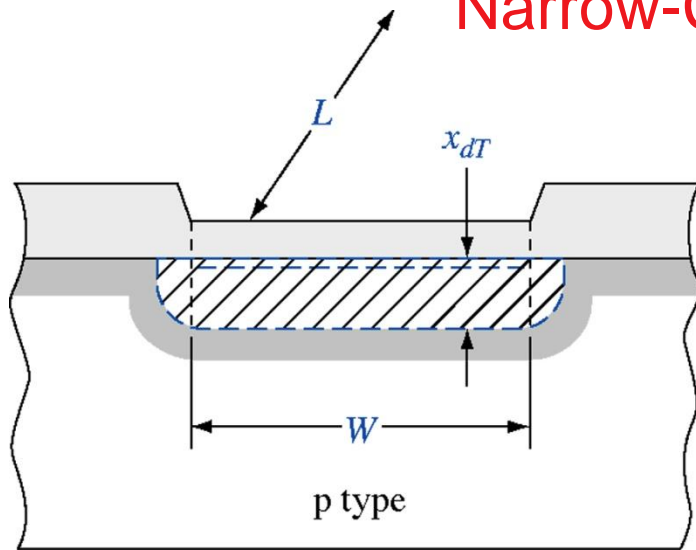


Schwelspannung als Funktion der Kanallänge für 2 Werte der Drain-Source- und Body-Source-Spannung.





# Narrow-Channel-Effekt



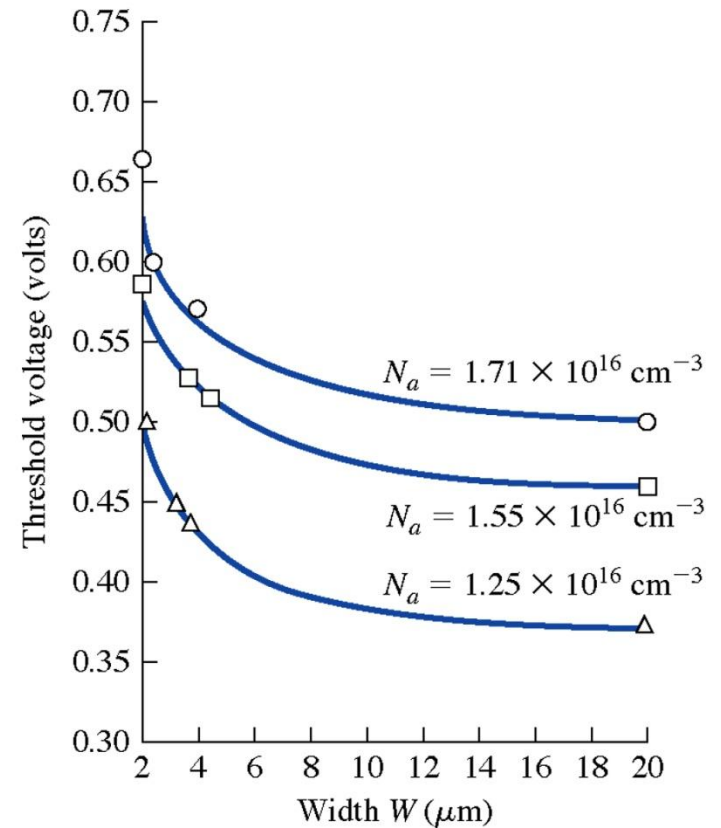
$$\Delta Q_B = eN_a L x_{dT} \chi x_{dT}$$

Der Fitparameter  $\chi$  hängt von der tatsächlichen lateralen Unterdiffusion ab.

$$|Q_B| = eN_a W L x_{dT} \left( 1 + \frac{\chi x_{dT}}{W} \right)$$

Man erhält eine *positive* Schwellspannungsverschiebung  $\sim 1/W$ :

$$\Delta V_T = \frac{eN_a x_{dT}}{C_{ox}} \left( \frac{\chi x_{dT}}{W} \right)$$



# 13.3 Weitere wichtige Effekte

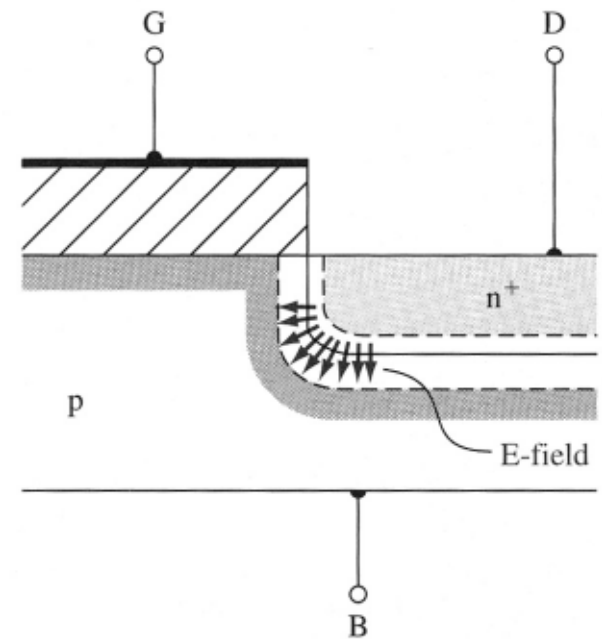
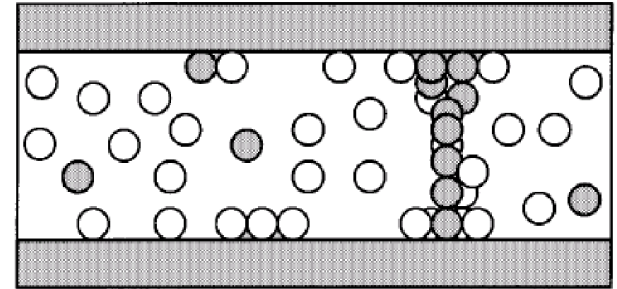
## Durchbruch-Spannung

### Oxid-Durchbruch

Bulk-Oxid bricht bei einer Feldstärke von etwa  $F_{ox} \approx 6 \times 10^6$  V/cm durch (Avalanche-Effekt). Ein 50nm-Oxid würde etwa 30 V aushalten, aber Defekte im Gateoxid verringern die Durchbruchspannung. In den heutigen ultradünnen Oxiden kommt es nicht mehr zum Avalanche-Effekt. Man beobachtet "wear-out", nm-grosse Löcher als Folge von Gatestrom-Filamentierung.

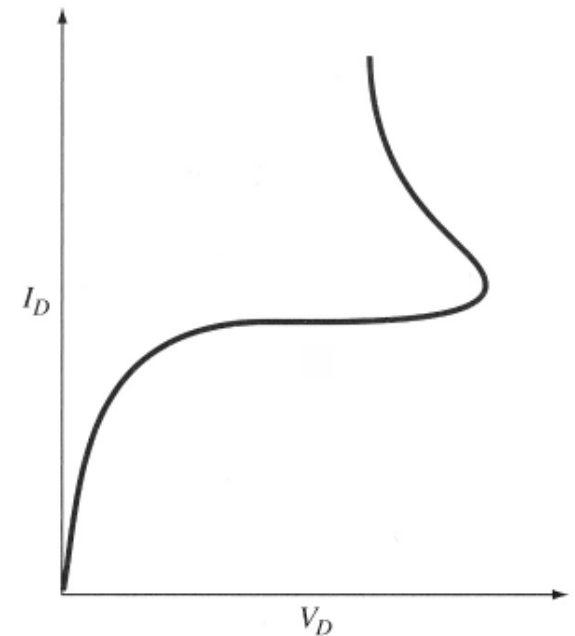
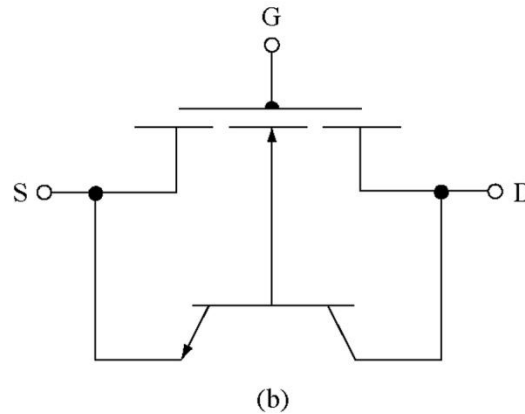
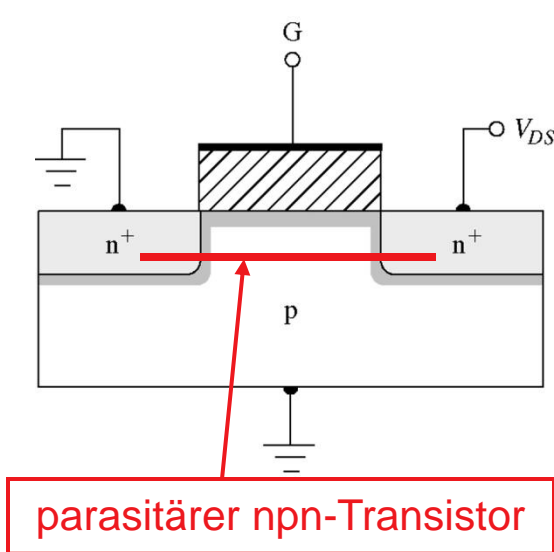
### Lawinen-Durchbruch

Wie in der pn-Diode kann es zum Durchbruch infolge Stossionisation kommen. Das elektrische Feld wird im gesperrten Drain-Substrat-Übergang im Bereich der grössten Krümmung der RLZ-Grenze konzentriert.

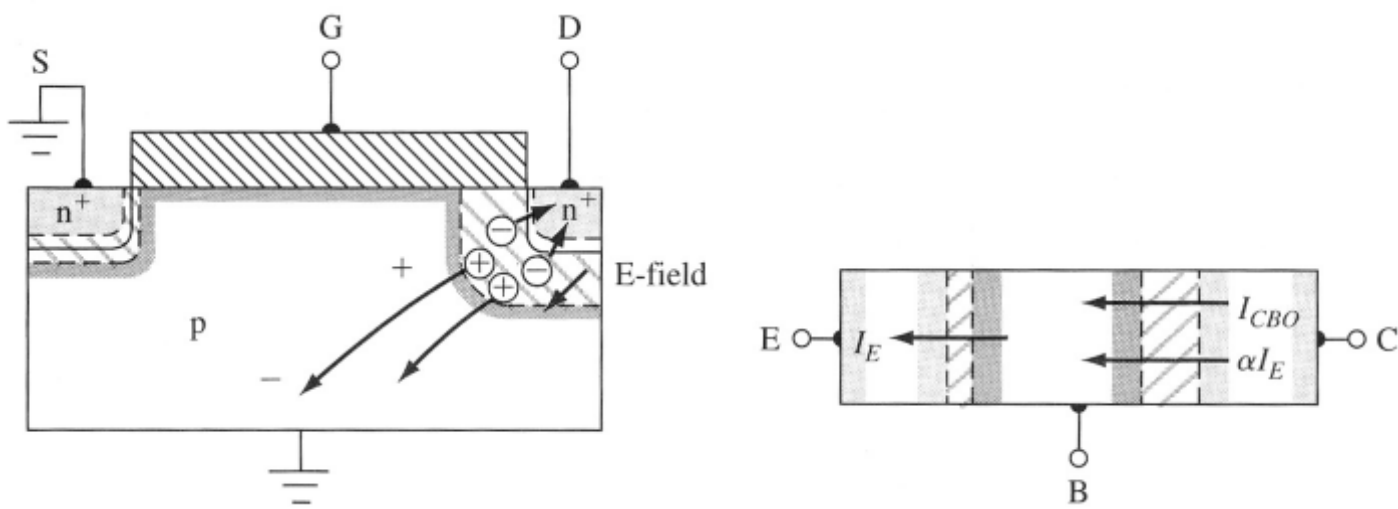


## Snapback-Durchbruch

Durchbruch als Folge der Wirkung eines parasitären Bipolar-Transistors.



$I_D$ - $V_D$ -Charakteristik mit Snapback-Durchbruch.



Dadurch werden Elektronen aus der  $n^+$ -Source emittiert, die teilweise zum Drain fließen. Dort verstärken sie die Stossionisationsrate (Rate  $\sim$  Stromdichte)  $\rightarrow$  Rückkopplungseffekt. Der Kollektorstrom ist

$$I_C = \alpha I_E + I_{CB0}$$

Wegen der offenen Basis (das Potential ist "floating") gilt  $I_C = I_E$ . Am Avalanche-Durchbruch wird der Kollektor-Sperrstrom verstärkt (Multiplikationsfaktor  $M$ ).

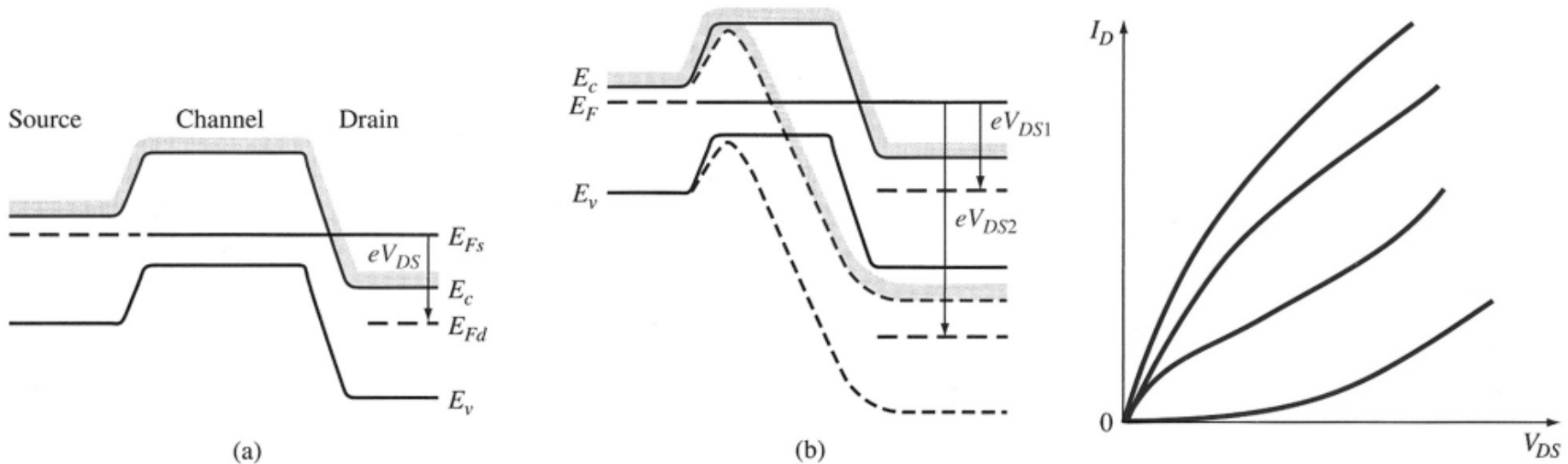
$$I_C = M(\alpha I_C + I_{CB0})$$

$$I_C = \frac{MI_{CB0}}{1 - \alpha M}$$

Die Stromverstärkung  $\alpha$  ist eine starke Funktion des Kollektorstroms (bei kleinem  $I_C$  klein wegen Rekombination in der BE-Diode). Mit wachsendem  $\alpha$  reicht ein kleineres  $M$  für die Durchbruchbedingung  $\rightarrow$  snapback-Effekt.

## Punch-Through

Bedingung, wo sich Source-Substrat- und Drain-Substrat-Übergang berühren. Bereits vorher steigt der Drainstrom stark an, da er exponentiell von der Höhe der Source-Drain-Potentialbarriere abhängt. Dies ist der bereits erwähnte DIBL-Effekt (Drain Induced Barrier Lowering).

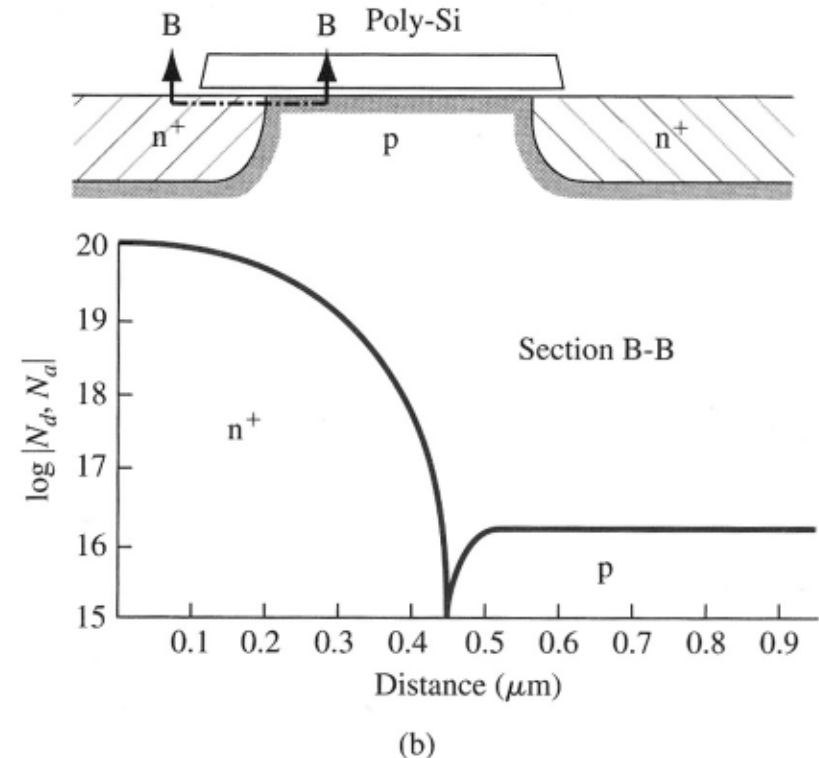
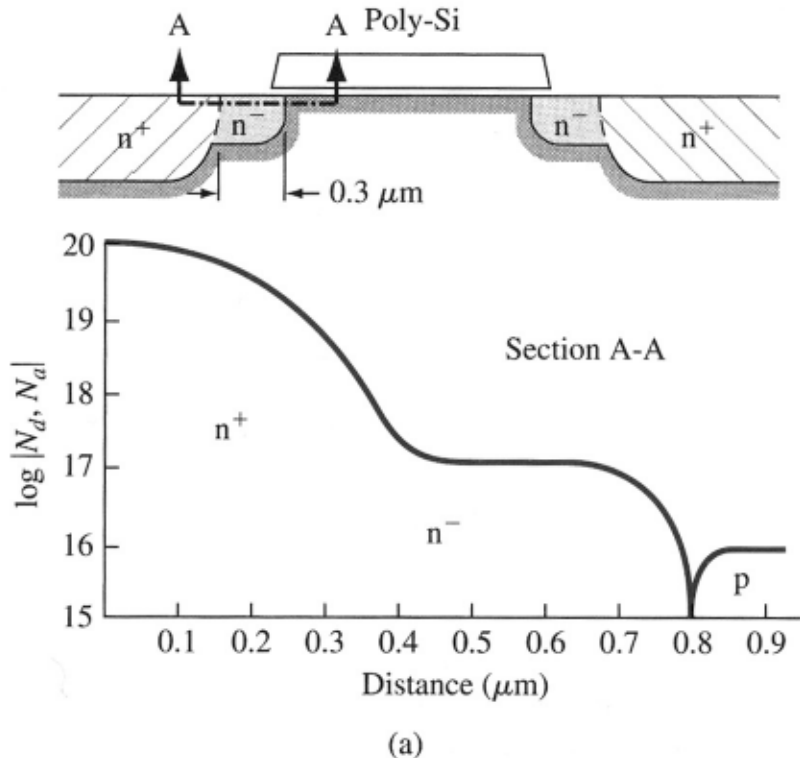


(a) Bandkantendiagramm eines Langkanal-MOSFETs entlang des Kanals, (b) dasselbe für einen Kurzkanal-MOSFET vor und nach dem punch-through.

Typische  $I_D$ - $V_{DS}$ -Kurve eines MOSFETs mit punch-through.

# Der LDD-Transistor

LDD steht für “Lightly Doped Drain”. Um die diskutierten Durchbruch-Mechanismen abzuschwächen, muss das elektrische Feld im Drain-Substrat-Übergang verkleinert werden. Man implantiert “Source/Drain-Extensions” mit geringerer Dosis.



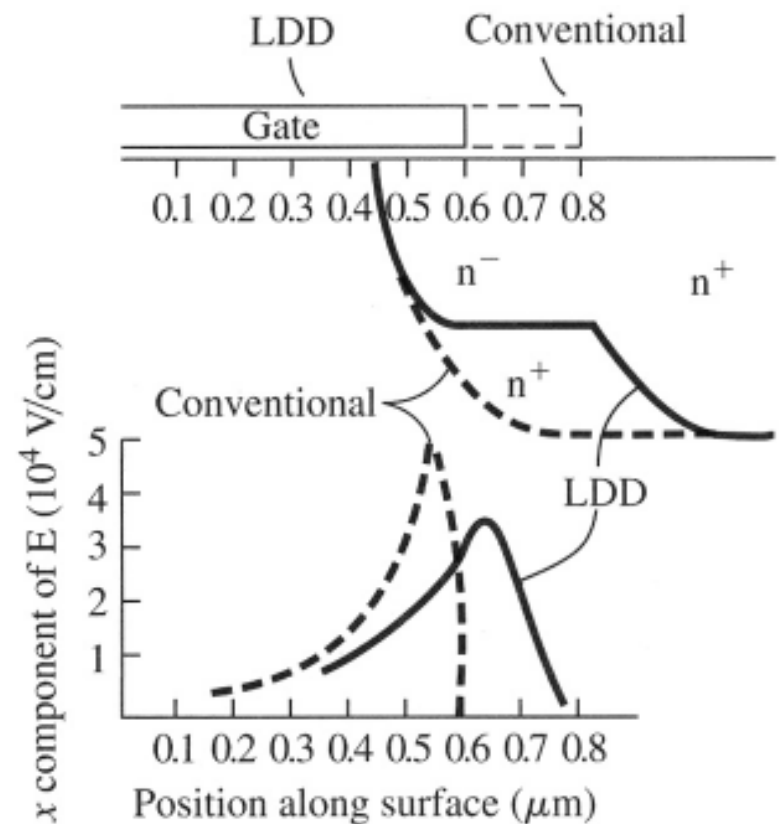
(a) Lightly Doped Drain (LDD) Struktur, (b) conventionelle Struktur (nach Ogura).

## Der LDD-Transistor (2)

Der Feldstärke-Peak am Drain wird kleiner und wandert mehr ins Draingebiet hinein. Dadurch wird auch die Degradation des MOSFETs durch heisse Elektronen verringert (siehe unten).

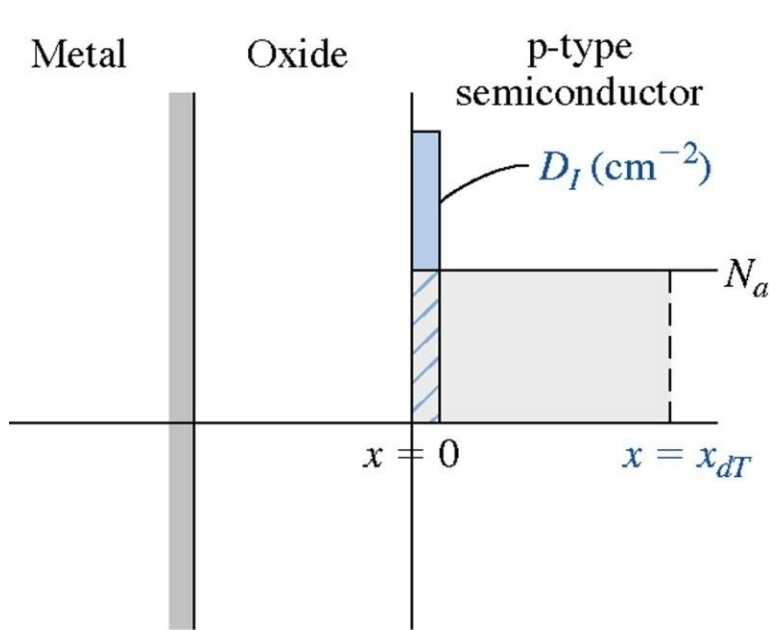
Der Nachteil ist neben der grösseren Fabrikationskomplexität die Zunahme des Drain-Widerstands.

Die symmetrische Behandlung von Drain und Source hat seinen Grund lediglich in einem einfacheren Prozessablauf; am Source bewirkt das LDD-Gebiet (oder "LDS") keine Verbesserung der MOSFET-Charakteristiken.

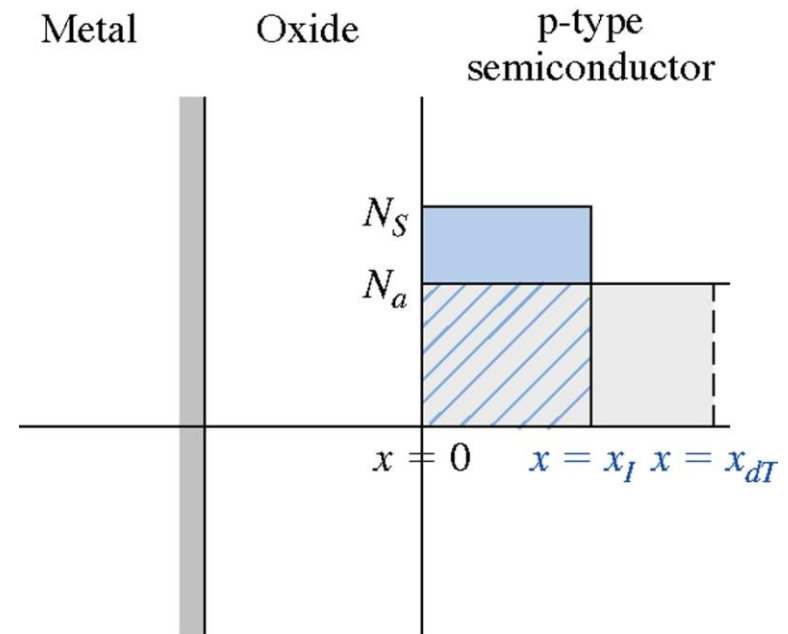


Feldstärke als Funktion des Abstands von der Si-SiO<sub>2</sub>-Grenzfläche (nach Ogura).

# Einstellung der Schwellspannung durch Ionenimplantation



$$\Delta V_T = \frac{eD_I}{C_{ox}}$$



$$\Delta V_T = \frac{ex_I}{C_{ox}} [N_S - N_a]$$

Faktoren, die die Schwellspannung beeinflussen: Feste Oxidladungen,  $\phi_{ms}$ , Oxid-  
dicke, Dotierung.

Schwellspannung wird durch gezielte, präzise Dotierung der Region nahe der Si-SiO<sub>2</sub>-  
Grenzfläche eingestellt. Implantierung von Akzeptoren verschiebt  $V_T$  nach rechts,  
Implantierung von Donatoren nach links.

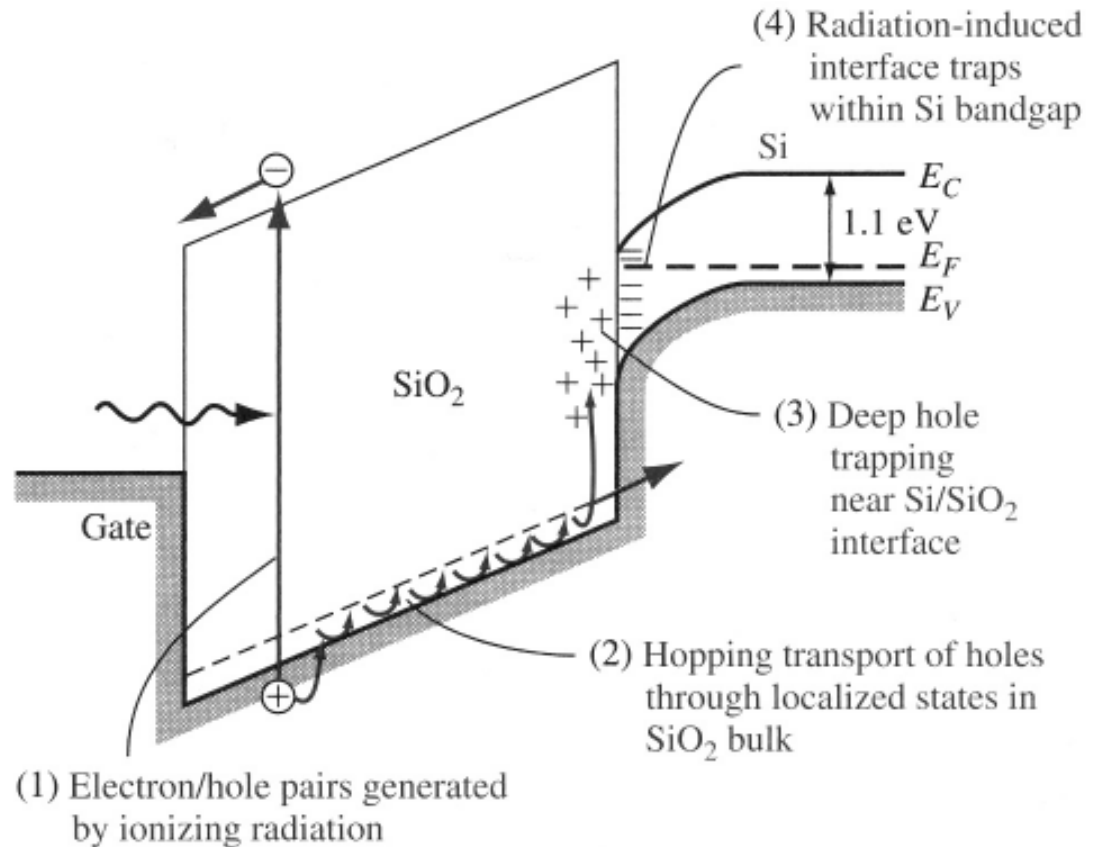


# Strahlungsinduzierte Oxidladungen

Ionisierende Strahlung kann Elektron-Loch-Paare im Oxid generieren. Zwei Arten: (i) kosmische Strahlung, (ii)  $\alpha$ -Teilchen aus Verpackung.

Beweglichkeit der Elektronen im Oxid ist relativ gross ( $\sim 20 \text{ cm}^2/\text{Vs}$ ). Elektronen fließen innerhalb von Pikosekunden aus dem Oxid ohne wesentliche Wirkung. Die Beweglichkeit der Löcher ist dagegen sehr klein ( $10^{-4} - 10^{-11} \text{ cm}^2/\text{Vs}$ , hopping transport).

Löcher werden teilweise in der Nähe der Grenzfläche eingefangen. Diese Ladung ist stabil auf einer Zeitskala von Stunden bis Jahren und bewirkt einen negativen  $V_T$ -shift.



Strahlungsinduzierte Ionisationsprozesse in einem MOS-Kondensator unter positiver Gatespannung (nach *Ma et al.*).

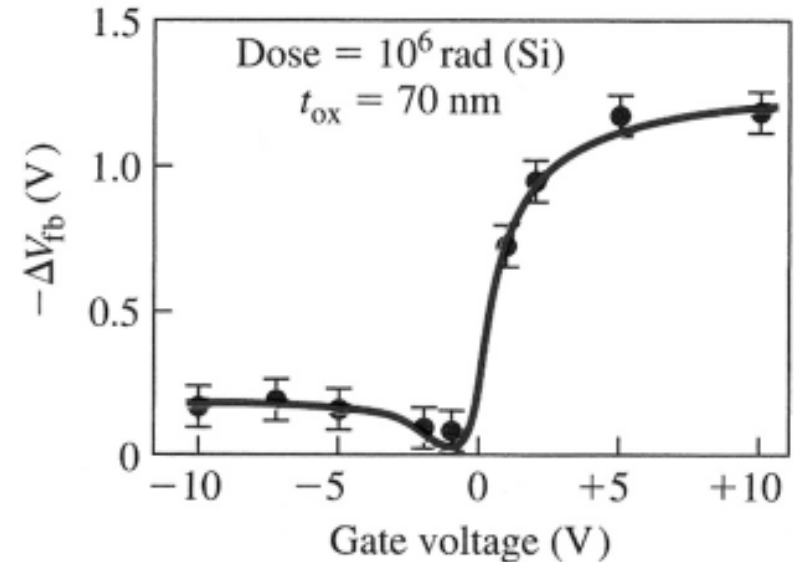
## Strahlungsinduzierte Oxidladungen (2)

Die Einfangzentren sind mit O-Vakanzen in der Si-reichen Schicht in der Nähe der Si-SiO<sub>2</sub>-Grenzfläche korreliert. Löcher werden dort in einer etwa 50Å dicken Schicht eingefangen.

Die Flächendichte der eingefangenen Löcher hängt von der Gatespannung ab. Bei kleinen Gatespannungen rekombinieren viele e-h-Paare, bevor Löcher eingefangen werden können. Bei grossen positiven Gatespannungen erreichen im wesentlichen alle Löcher die Grenzschicht.

Eine grosse negative Gatespannung zieht die Löcher zur Si-Gate-Grenzfläche, wo einige eingefangen werden. Deren Effekt auf  $V_T$  ist aber gering.

Als Konsequenz der Strahlung kann der MOSFET vom Anreicherungstyp zum Verarmungstyp driften und sich nicht mehr abschalten lassen.

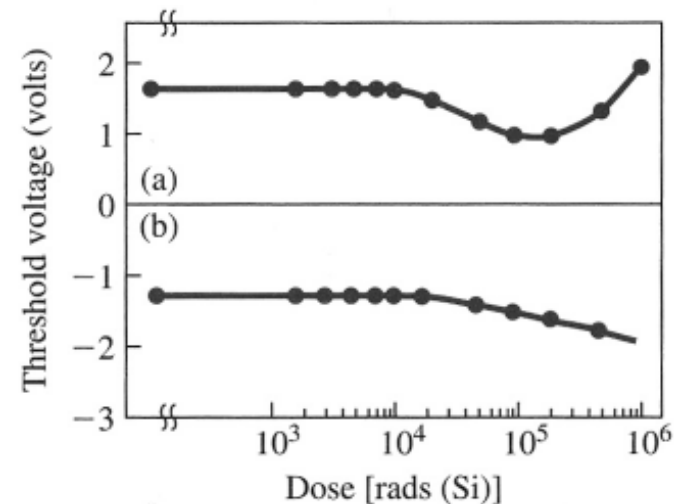
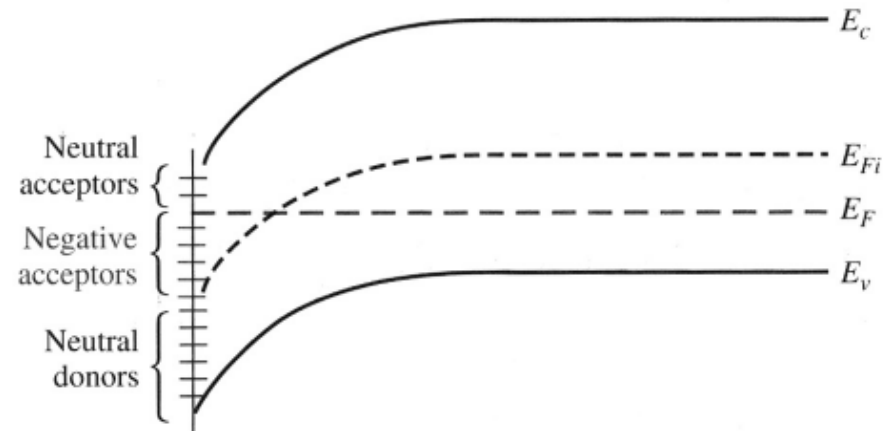


Strahlungsinduzierte Verschiebung der Flachbandspannung in einem MOS-Kondensator während der Bestrahlung als Funktion der Gatespannung (nach *Ma et al.*).

# Strahlungsinduzierte Grenzflächenzustände

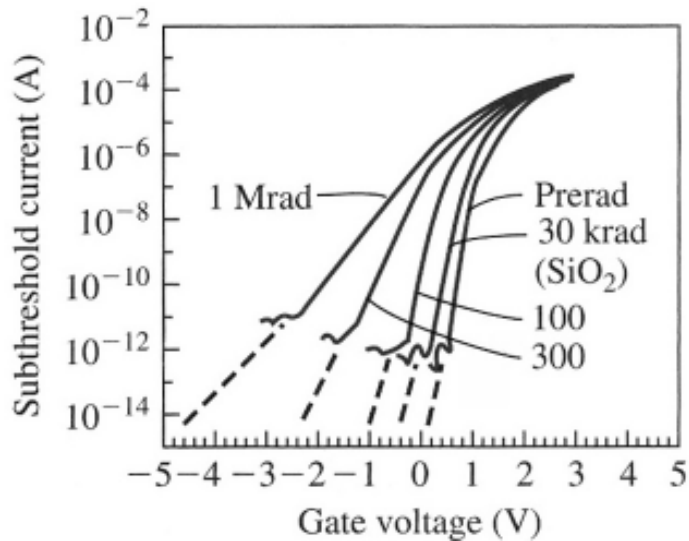
Am Inversionspunkt ist die Netto-Ladung in den Grenzflächenzuständen eines  $n$ -Kanal-MOSFETs negativ. Die damit verbundene Schwellspannungsverschiebung ist positiv!

Ionisierende Strahlung generiert weitere Grenzflächenzustände. Der  $V_T$ -shift als Funktion der Strahlungs-dosis zeigt zuerst den negativen Trend aufgrund der strahlungsinduzierten Oxidladungen. Der  $n$ -Kanal-MOSFET zeigt dann eine Umkehr wegen der Kompensation durch strahlungsinduzierte Grenzflächen-zustände.



Schwelspannung als Funktion der totalen Dosis der ionisierenden Strahlung für (a) einen  $n$ -Kanal-MOSFET und (b) einen  $p$ -Kanal-MOSFET (nach *Ma et al.*).

# Strahlungsinduzierte Grenzflächenzustände (2)

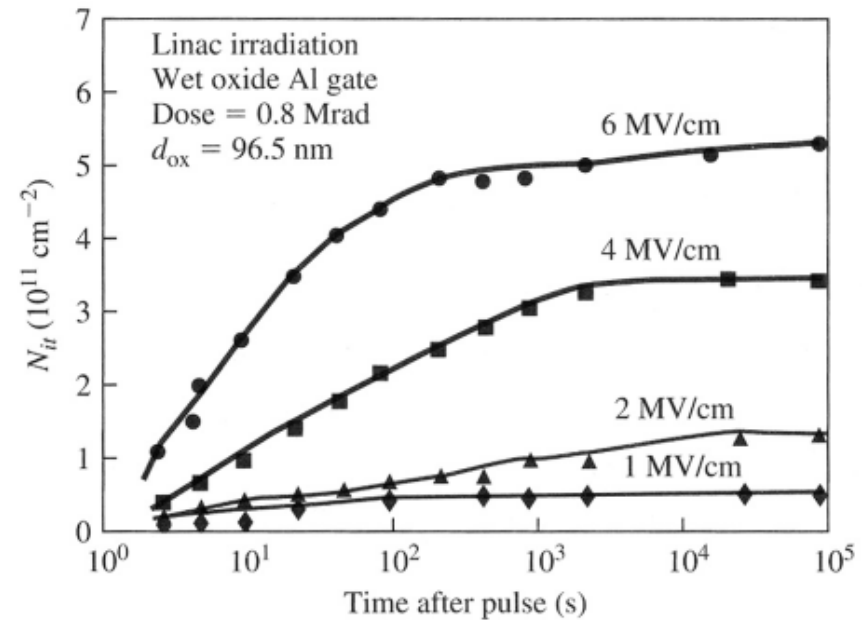


Subthreshold-Strom vor der Bestrahlung und für 4 verschiedene Strahlungs-Dosen.

Die Generationsrate dieser Zustände ist eine starke Funktion der Oxid-Feldstärke.

Die endgültige Dichte wird nach einer Zeit von 100 bis 10'000 Sekunden nach dem Strahlungspuls erreicht.

Der Anstieg der Eingangs-Kennlinie im Subthreshold-Bereich ("subthreshold swing") ist stark von der Dichte der Grenzflächenzustände abhängig.

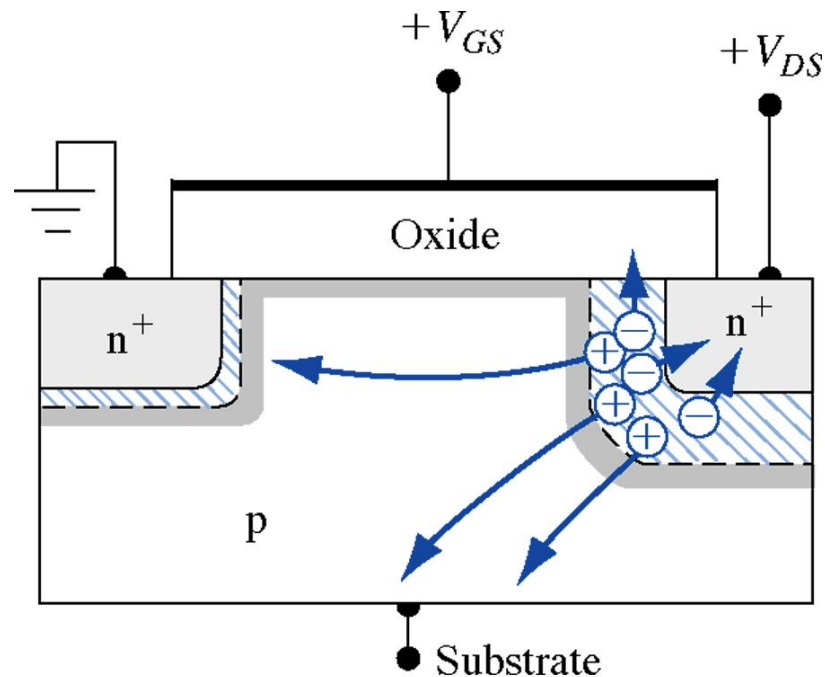


Zeitliche Änderung der strahlungsinduzierten Grenzflächenzustandsdichte nach einem Strahlungspuls für verschiedene Werte der Oxidfeldstärke (nach *Ma et al.*).

# Effekte heisser Elektronen

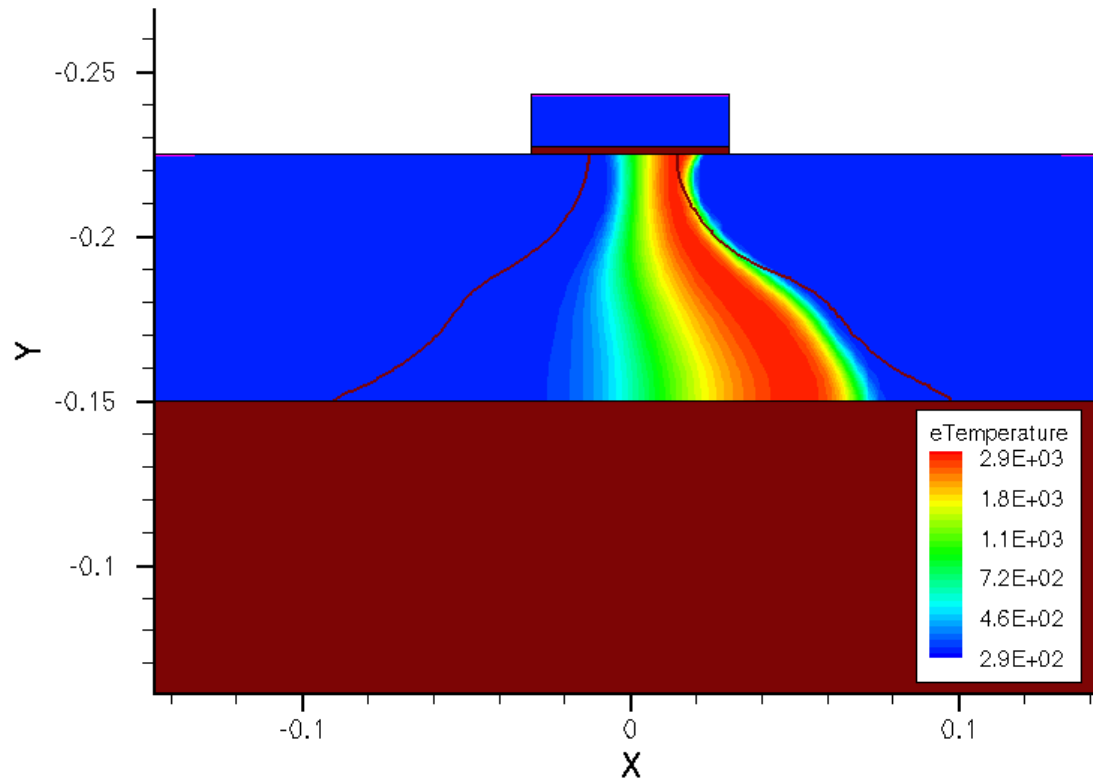
Im gesperrten Drain-Substrat-Übergang werden e-h-Paare durch Stossionisation erzeugt. Im  $n$ -Kanal-MOSFET fließen die Elektronen mehrheitlich zum Drainkontakt, die Löcher zum Substratkontakt.

Einige energiereiche Elektronen gelangen ins Oxid oder können die Oxid-Potentialbarriere sogar überwinden und einen Gatestrom produzieren. Negative Oxidladungen können sich akkumulieren und einen positiven lokalen  $V_T$ -shift bewirken.



## Effekte heisser Elektronen (2)

Eines der Hauptprobleme der heutigen CMOS-Technologie besteht darin, dass der Gatestrom der heissen Elektronen zusätzliche Grenzflächenzustände generiert. Der wahrscheinlichste Mechanismus ist das Aufbrechen der Si-H-Bindungen der passivierten "dangling bonds". Folge ist eine Langzeitverschiebung von  $V_T$  und die Degradation der Beweglichkeit der Kanal-Elektronen.











# 14. Optische Bauelemente

**14.1 Optische Absorption (N14.1)**

**14.2 Die Solarzelle (N14.2)**

**14.3 Der Photodetektor (N14.3)**

**14.4 Photo- und Elektrolumineszenz (N14.4)**

**14.5 Die lichtemittierende Diode (N14.5)**

**14.6 Die Laserdiode (N14.6)**

# 14.1 Optische Absorption

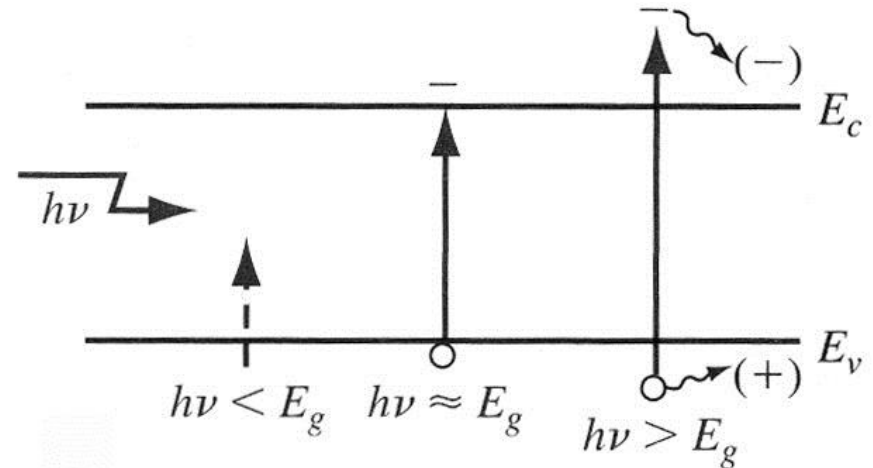
## Absorptionskoeffizient

Photonen, die Quanten des elektromagnetischen Feldes, können mit dem Halbleiterkristall wechselwirken. Neben Wechselwirkung mit Phononen, flachen Störstellen oder anderen Defekten, ist die technisch wichtigste die Wechselwirkung mit den Valenzelektronen, die zur Generation von e-h-Paaren führt.

$$\lambda = \frac{c}{\nu} = \frac{hc}{E} = \frac{1.24}{E} \mu\text{m}$$

Das Verhältnis aus Photonenenergie  $E=h\nu$  und Bandlücke  $E_g$  entscheidet darüber, ob das Photon absorbiert wird oder nicht.

Wenn  $h\nu > E_g$ , wird ein e-h-Paar generiert, wobei die überschüssige kinetische Energie in Wärme umgewandelt wird (Dissipation).



Optische Generation von Elektron-Loch-Paaren in einem Halbleiter.

## Absorptionskoeffizient (2)

Die Intensität des Photonenflusses bezeichnet man mit  $I_\nu(x)$  [eV/cm<sup>2</sup>s]. Treffen Photonen bei  $x$  auf, so ist die absorbierte Photonenenergie pro Zeit und Fläche zwischen  $x$  und  $x+dx$

$$\alpha I_\nu(x) dx$$

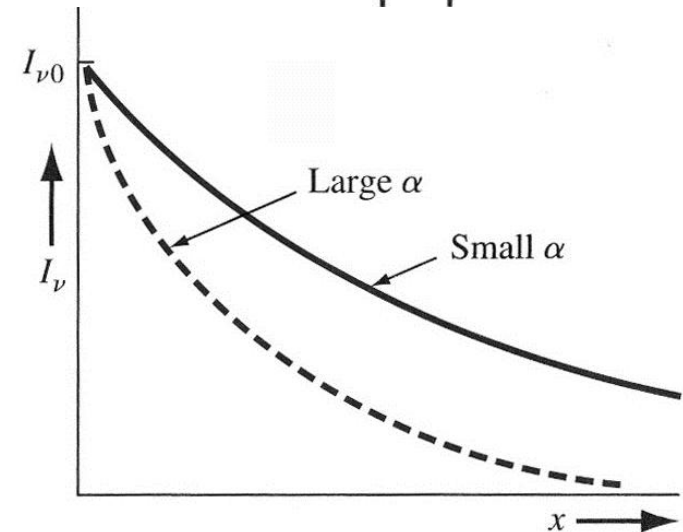
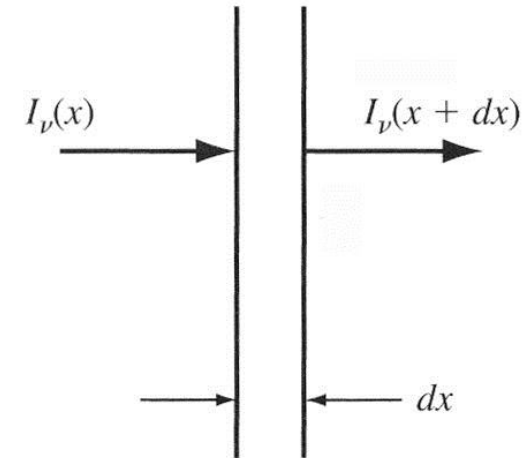
Der *Absorptionskoeffizient*  $\alpha$  ist die relative Zahl von Photonen, die pro Längeneinheit absorbiert werden [1/cm]

$$I_\nu(x+dx) - I_\nu(x) = \frac{dI_\nu(x)}{dx} dx = -\alpha I_\nu(x) dx$$

$$\frac{dI_\nu(x)}{dx} = -\alpha I_\nu(x)$$

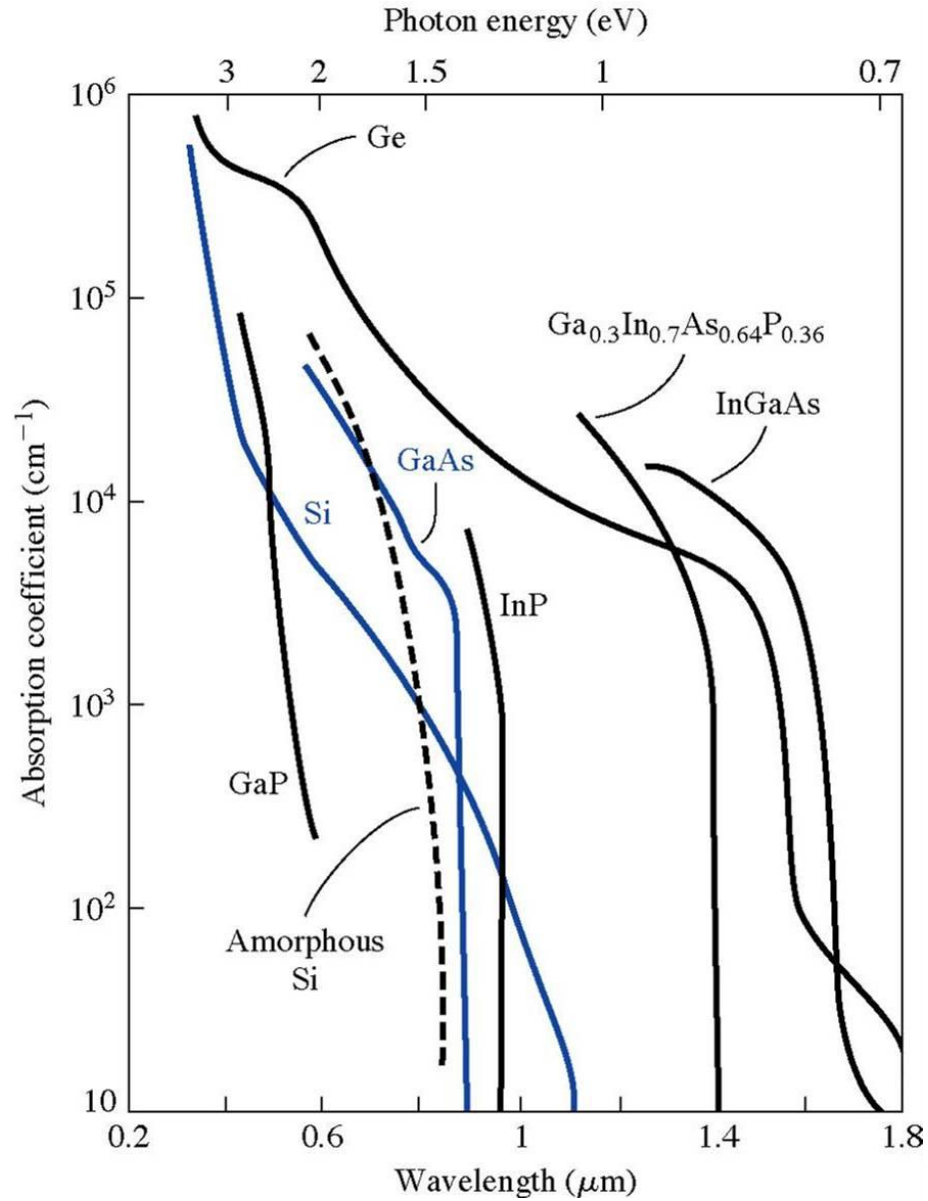
Mit der Anfangsbedingung  $I_\nu(0) = I_{\nu 0}$  lautet die Lösung

$$I_\nu(x) = I_{\nu 0} e^{-\alpha x}$$

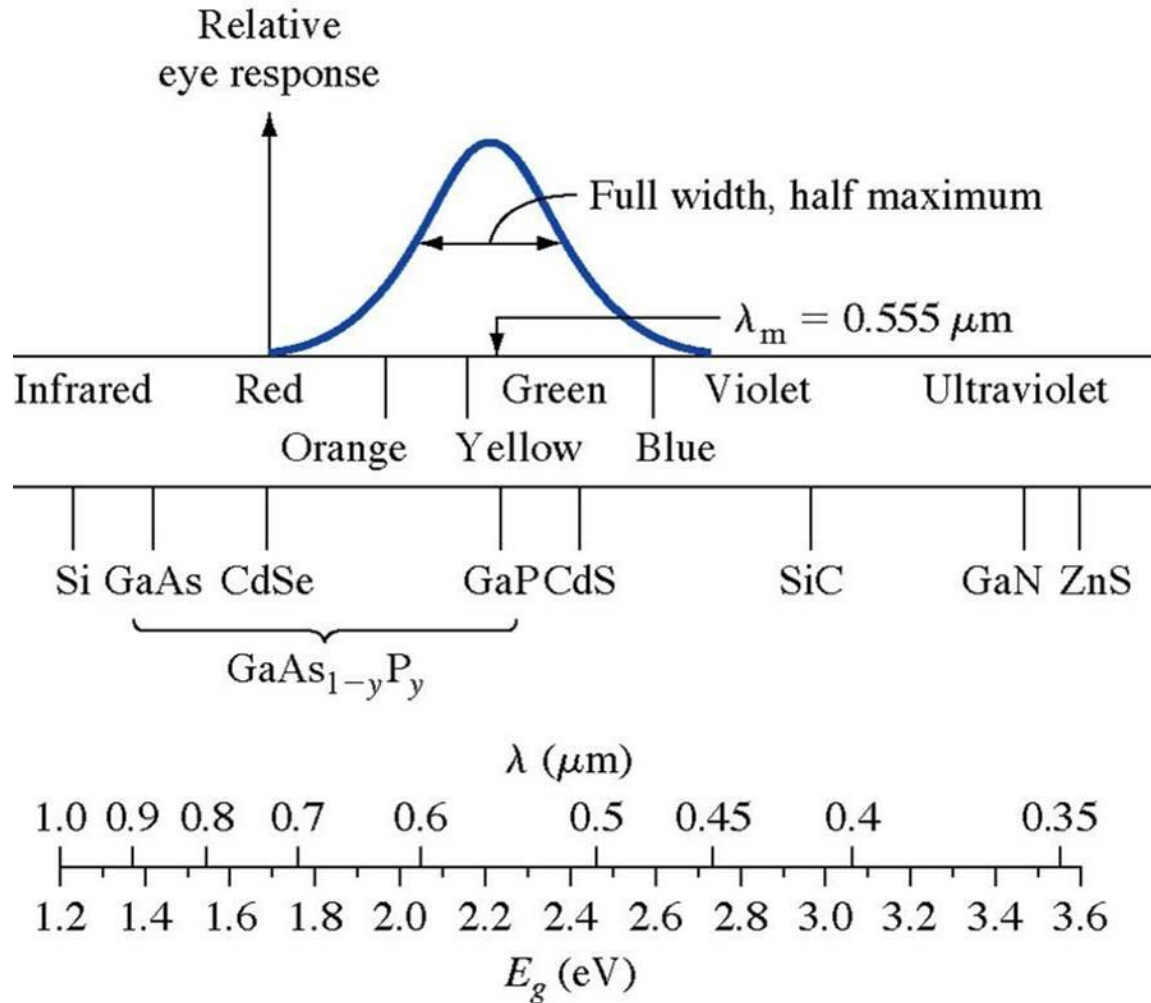


Räumlicher Verlauf der Lichtintensität für zwei Werte des Absorptionskoeffizienten.

# Absorptionskoeffizient (3)



# Spektrale Empfindlichkeit verschiedener Halbleiter



# Generation von Elektron-Loch-Paaren

Da die Intensität des Photonenflusses  $I_\nu(x)$  [eV/cm<sup>2</sup>s] ist, ist die *Absorptionsrate* gleich  $\alpha I_\nu(x)$  [eV/cm<sup>3</sup>s]. Dies ist die pro Volumen und Zeit absorbierte Energie (= Rate).

Nimmt man an, dass ein absorbiertes Photon der Energie  $h\nu$  ein e-h-Paar generiert, dann ist die entsprechende Generationsrate

$$g' = \frac{\alpha I_\nu(x)}{h\nu}$$

mit der Masseinheit [# / cm<sup>3</sup>s].

Das Verhältnis  $I_\nu(x)/h\nu$  ist der Photonenfluss (Teilchenfluss). Im allgemeinen ist die sogenannte *Quantenausbeute* kleiner als 1, d.h. nicht jedes Photon generiert ein e-h-Paar. Dann muss  $g'$  mit einem Effizienzfaktor multipliziert werden.

**Beispiel:** GaAs bei 300K. Die Intensität des Photonenflusses am Ort  $x$  sei  $I_\nu(x) = 0.05$  W/cm<sup>2</sup> bei einer Wellenlänge von 0.75  $\mu$ m, was typisch für Sonnenlicht ist. Der Absorptionskoeffizient von GaAs bei dieser Wellenlänge ist  $\alpha = 7 \cdot 10^3$ /cm. Die Photonenenergie ist  $E = h\nu = 1.24/0.75$  eV = 1.65 eV. Daraus erhält man  $g' = 0.05 \cdot 7 \cdot 10^3 / (1.65 \cdot 1.6 \cdot 10^{-19}) = 1.33 \cdot 10^{21} / (\text{cm}^3\text{s})$ .

Mit einer Minoritätsladungsträger-Lebensdauer von  $\tau = 10^{-7}$ s folgt daraus eine Überschussdichte von  $\delta n = g' \tau = 1.33 \cdot 10^{14}$  cm<sup>-3</sup>.

# 14.2 Die Solarzelle

## Der $pn$ -Übergang als Solarzelle

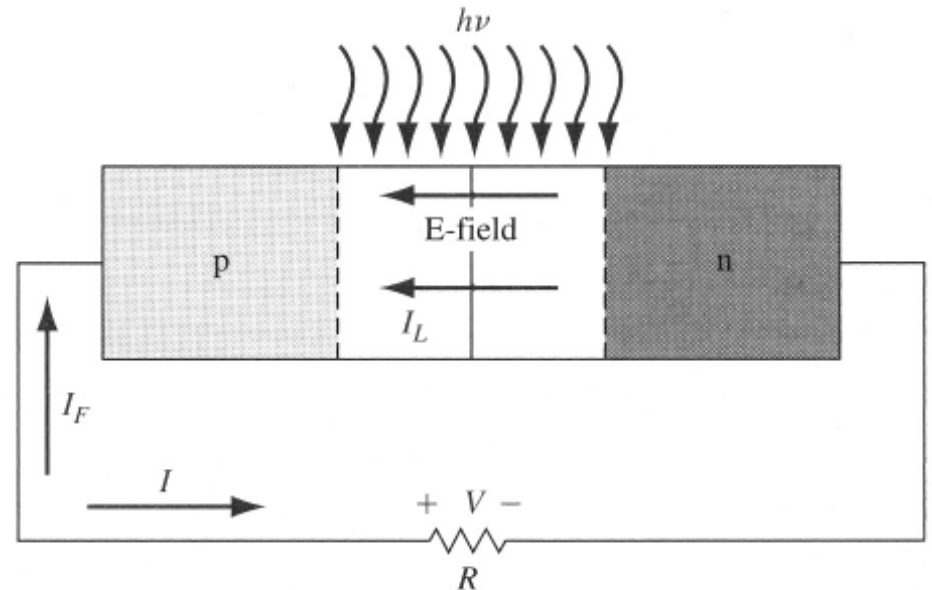
Eine Solarzelle ist eine  $pn$ -Diode ohne von aussen angelegte Spannung, die Lichtleistung in elektrische Leistung konvertiert und diese einer resistiven Last zuführt.

Die generierten e-h-Paare werden im eingebauten Potential getrennt. Es fliesst ein Photostrom  $I_L$  in Sperrrichtung, der einen Spannungsabfall über der Last erzeugt. Dieser Spannungsabfall polarisiert die Diode in Vorwärtsrichtung und treibt einen Vorwärtsstrom  $I_F$ .

Der Nettostrom ist

$$I = I_L - I_F = I_L - I_S \left\{ \exp \left( \frac{eV}{kT} \right) - 1 \right\}$$

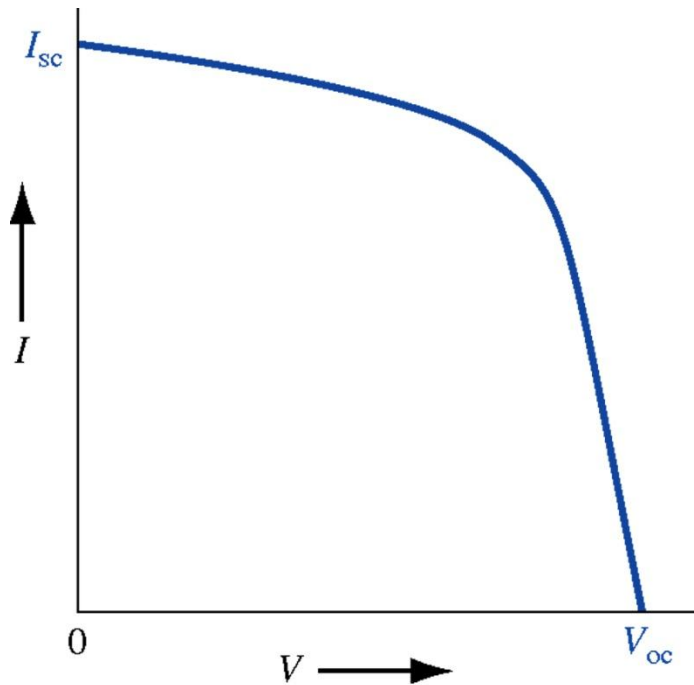
und bleibt in jedem Falle ein Strom in Sperrrichtung.



Solarzelle auf Basis eines  $pn$ -Übergangs mit resistiver Last.



## Der *pn*-Übergang als Solarzelle (2)



$V_{oc}$  für Silizium:

610 – 630 mV (Industrie)

≈ 700 mV (teure Zellen)

717 mV (Weltrekord)

Zwei Grenzfälle sind von Bedeutung:

1.) Kurzschluss:  $R = 0$ ,  $V = 0$

$$I = I_{sc} = I_L$$

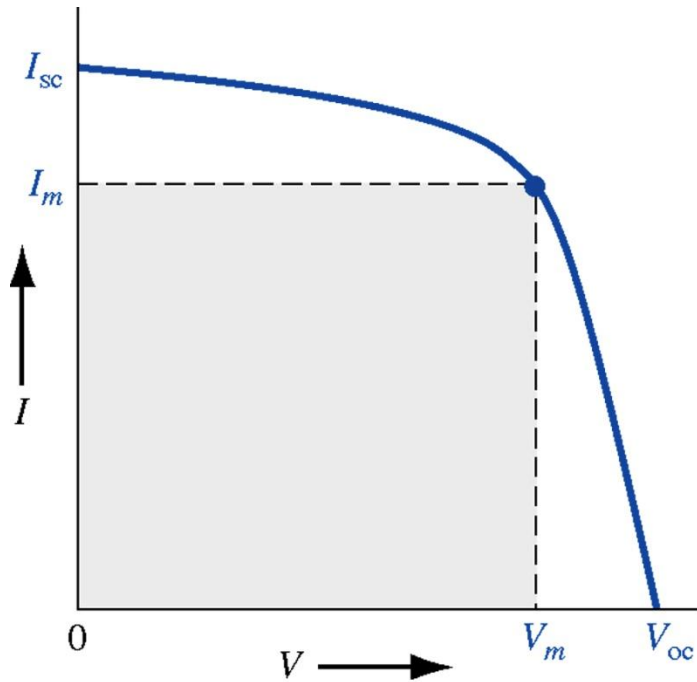
2.) Leerlauf:  $R = \infty$ ,  $I = 0$

$$I = 0 = I_L - I_S \left\{ \exp \left( \frac{eV_{oc}}{kT} \right) - 1 \right\}$$

woraus die Spannung am offenen Stromkreis folgt:

$$V_{oc} = V_t \ln \left( 1 + \frac{I_L}{I_S} \right)$$

# Die Leistung der Solarzelle



Die Leistung über der Last ist:

$$P = I V = I_L V - I_S \left\{ \exp \left( \frac{eV}{kT} \right) - 1 \right\} V$$

Die Bedingung für maximale Leistung erhält man durch Nullsetzen der ersten Ableitung:

$$\frac{dP}{dV} = 0 = I_L - I_S \left\{ \exp \left( \frac{V_m}{V_t} \right) - 1 \right\} - \frac{I_S V_m}{V_t} \exp \left( \frac{V_m}{V_t} \right)$$

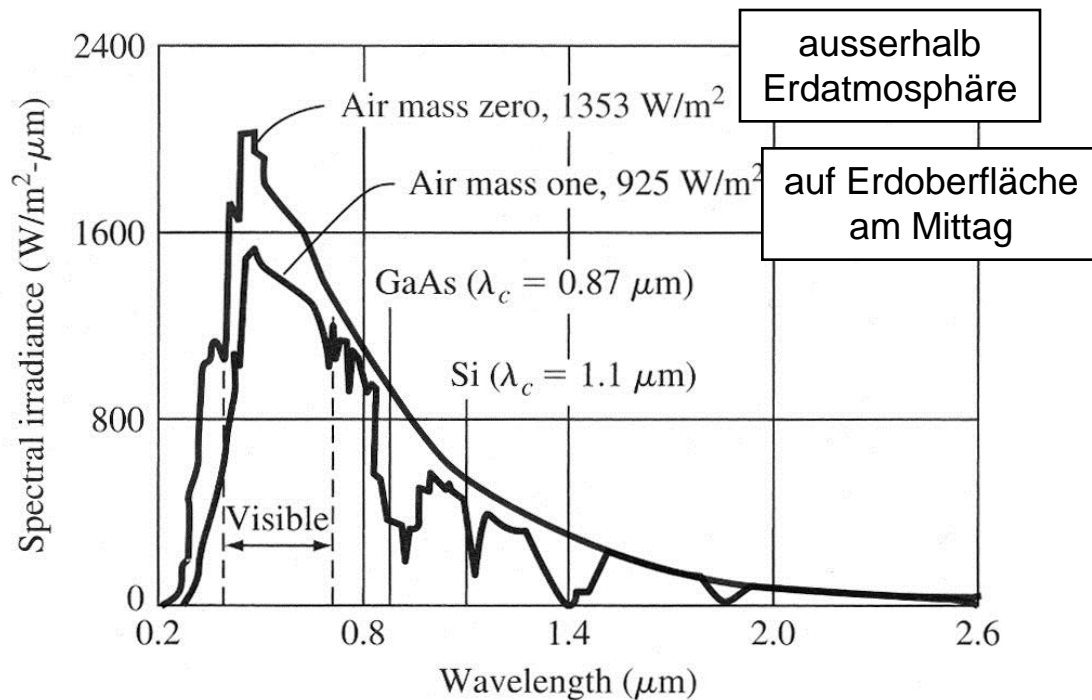
$$\left( 1 + \frac{V_m}{V_t} \right) \exp \left( \frac{V_m}{V_t} \right) = 1 + \frac{I_L}{I_S} = \exp \left( \frac{V_{oc}}{V_t} \right)$$

# Effizienz und Lichtbündelung

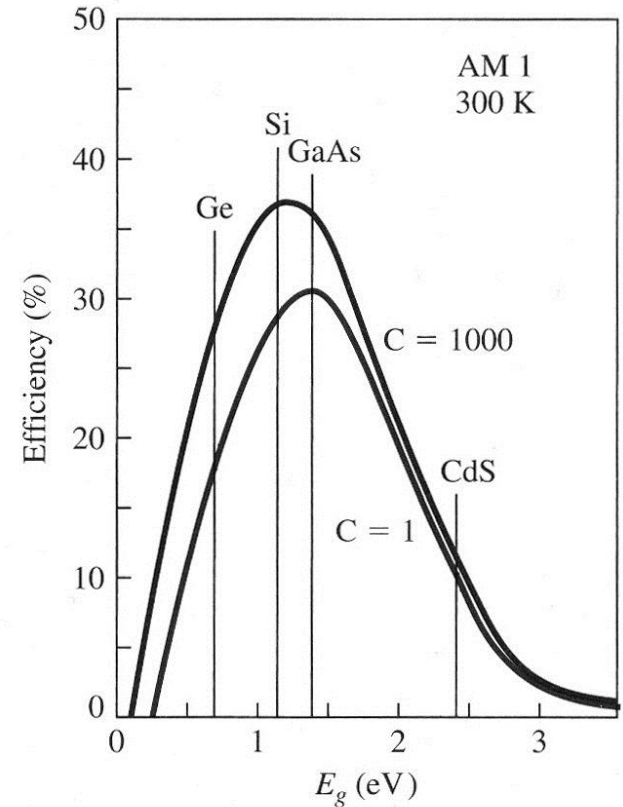
Die Effizienz der Umwandlung von Lichtenergie in elektrische Energie ist im Falle maximaler Leistung:

$$\eta = \frac{P_m}{P_{in}} * 100\% = \frac{I_m V_m}{P_{in}} * 100\%$$

Die Weltrekord-Solarzelle aus kristallinem Silizium hat eine Effizienz von 24.7%.

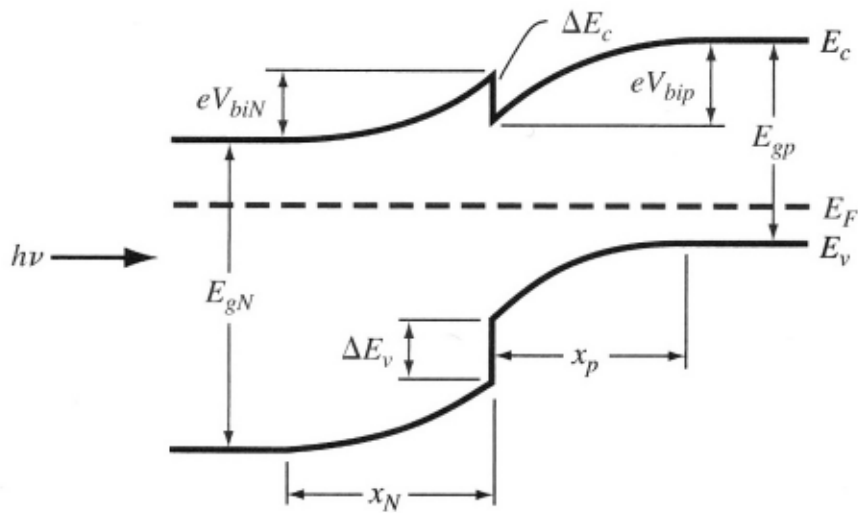


Strahlungsdichte der Sonne (nach Sze).

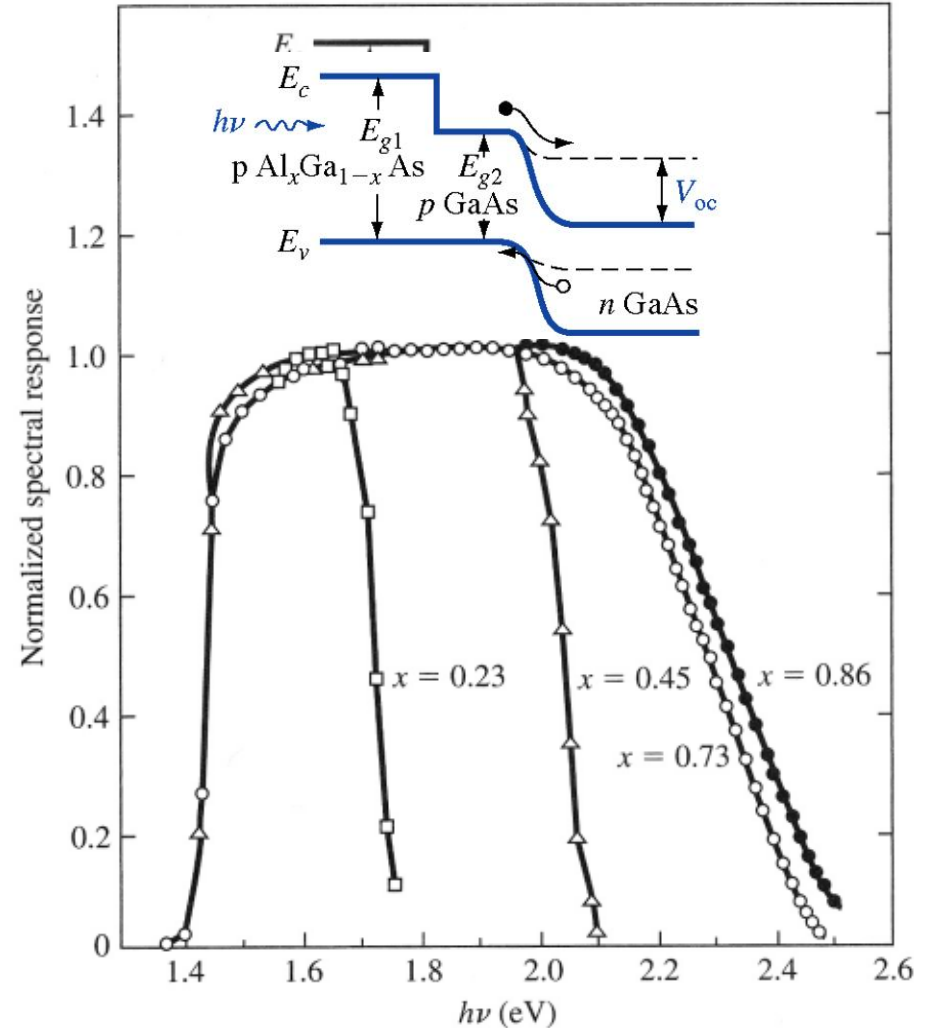


Ideale Effizienz einer Solarzelle bei  $T = 300 K$  für  $C = 1$  Sonne und  $C = 1000$  Sonnen als Funktion der Energielücke des Halbleiters (nach Sze).

# Hetero-Übergang als Solarzelle



Bandkantendiagramm eines pN-Hetero-Übergangs im thermodynamischen Gleichgewicht.

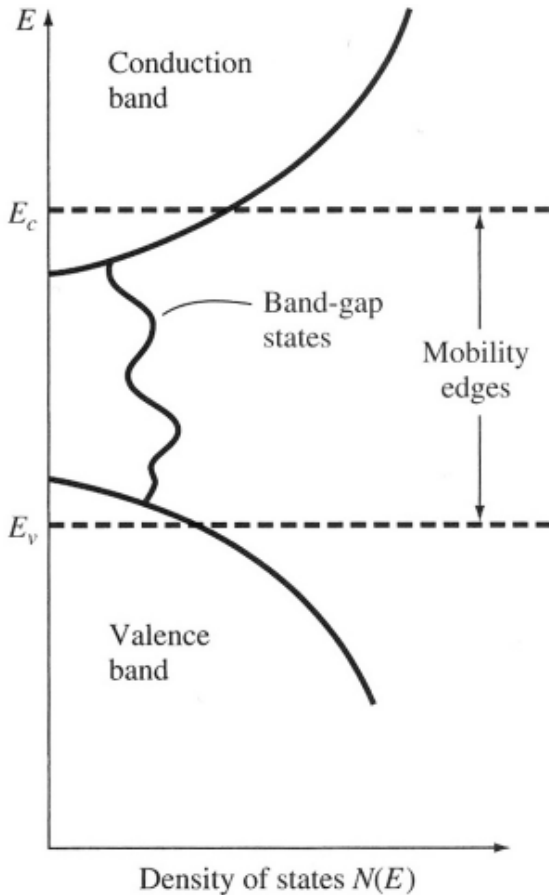


Normierte spektrale Empfindlichkeit verschiedener AlGaAs/GaAs-Solarzellen mit unterschiedlicher Zusammensetzung (nach Sze).

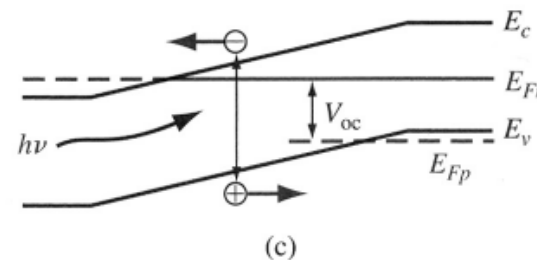
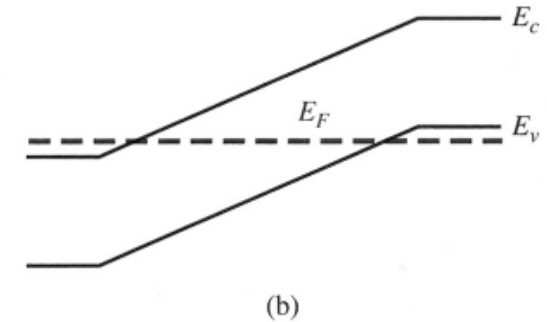
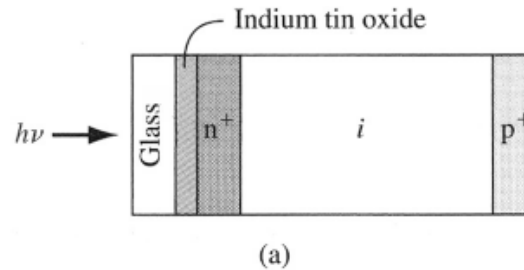
# Solarzellen mit amorphem Silizium

Solarzellen für die Massenanwendung müssen kostengünstig sein → Dünnschicht-Technologie mit amorphem Silizium.

Grosser Absorptionskoeffizient → 1  $\mu\text{m}$  Filmdicke reicht aus. Effizienz momentan > 10%.



Energetische Zustandsdichte von amorphem Silizium (nach Yang).



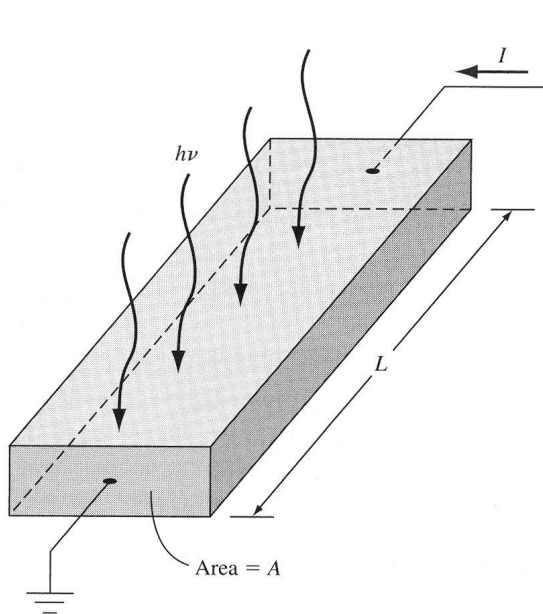
(a) Querschnitt, (b) Bandkantendiagramm im thermodynamischen Gleichgewicht, (c) Bandkantendiagramm einer amorphem Silizium PIN-Solarzelle unter Beleuchtung (nach Yang).

# 14.3 Der Photodetektor

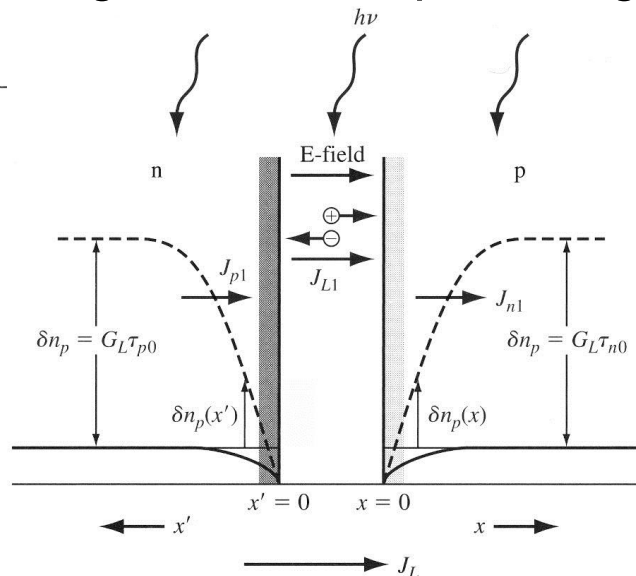
Es gibt verschiedene Typen von Bauelementen zur Umwandlung von elektro-magnetischer Strahlung in elektrische Signale.

Die Generation von Elektron-Loch-Paaren erhöht die lokale Leitfähigkeit des Halbleiters. In einem solchen Fall spricht man von einem **Photoleiter**.

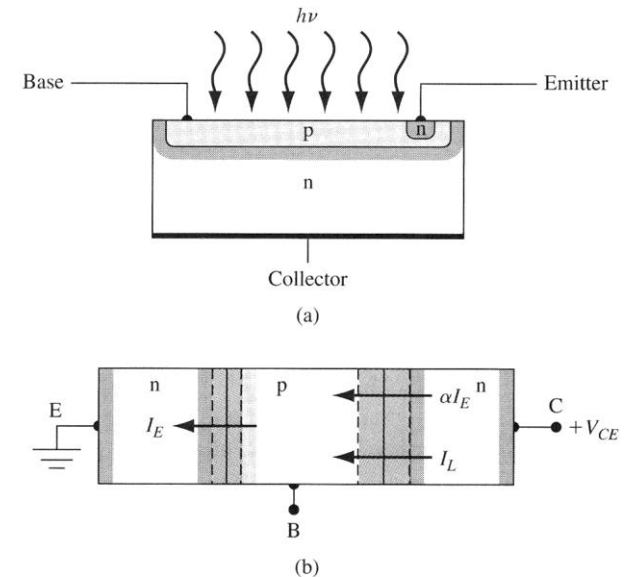
In einer **Photodiode** und im **Phototransistor** werden die Elektron-Loch-Paare in der Raumladungszone eines  $pn$ -Übergangs generiert.



Photoleiter schematisch.



Stationäre photoinduzierte Minoritätsträgerdichte und Photoströme in einer "langen" rückwärts-gepolten Diode.



(a) Photo-Bipolar-Transistor: (b) Blockdiagramm bei offener Basis.

# Der Photoleiter

Durch Bestrahlung wird die Leitfähigkeit eines Halbleitermaterials vergrößert:

$$\sigma = e[\mu_n(n_0 + \delta n) + \mu_p(p_0 + \delta p)]$$

Man spricht von *Photoleitung*:

$$\Delta\sigma = e\delta p(\mu_n + \mu_p)$$

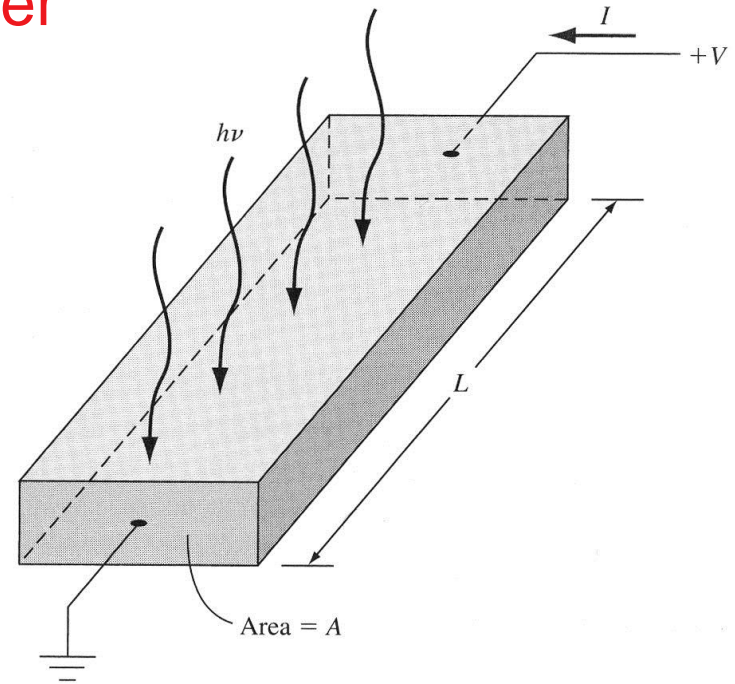
Im stationären Zustand ist die Überschussdichte proportional zur Generationsrate und der Lebensdauer:

$$\delta p = G_L \tau_p$$

Bei einer angelegten Spannung wird der Strom um den *Photostrom* vergrößert

$$J = (J_0 + J_L) = (\sigma_0 + \Delta\sigma) E$$

$$I_L = J_L A = \Delta\sigma A E = e G_L \tau_p (\mu_n + \mu_p) A E$$



Photoleiter schematisch.

*Verstärkungsfaktor*: Verhältnis der Rate, mit der die Ladung von den Kontakten aufgesammelt wird, zur Generationsrate im Innern

$$\Gamma_{ph} = \frac{I_L}{e G_L A L} = \frac{\tau_p}{t_n} \left( 1 + \frac{\mu_p}{\mu_n} \right) \quad \text{mit} \quad t_n = \frac{L}{\mu_n E}$$

# Die Photodiode

Eine Photodiode ist ein gesperrter  $pn$ -Übergang. Bei Beleuchtung werden überschüssige Ladungsträger durch das Feld aus der Raumladungszone entfernt. Der photo-generierte "prompte" Strom ist

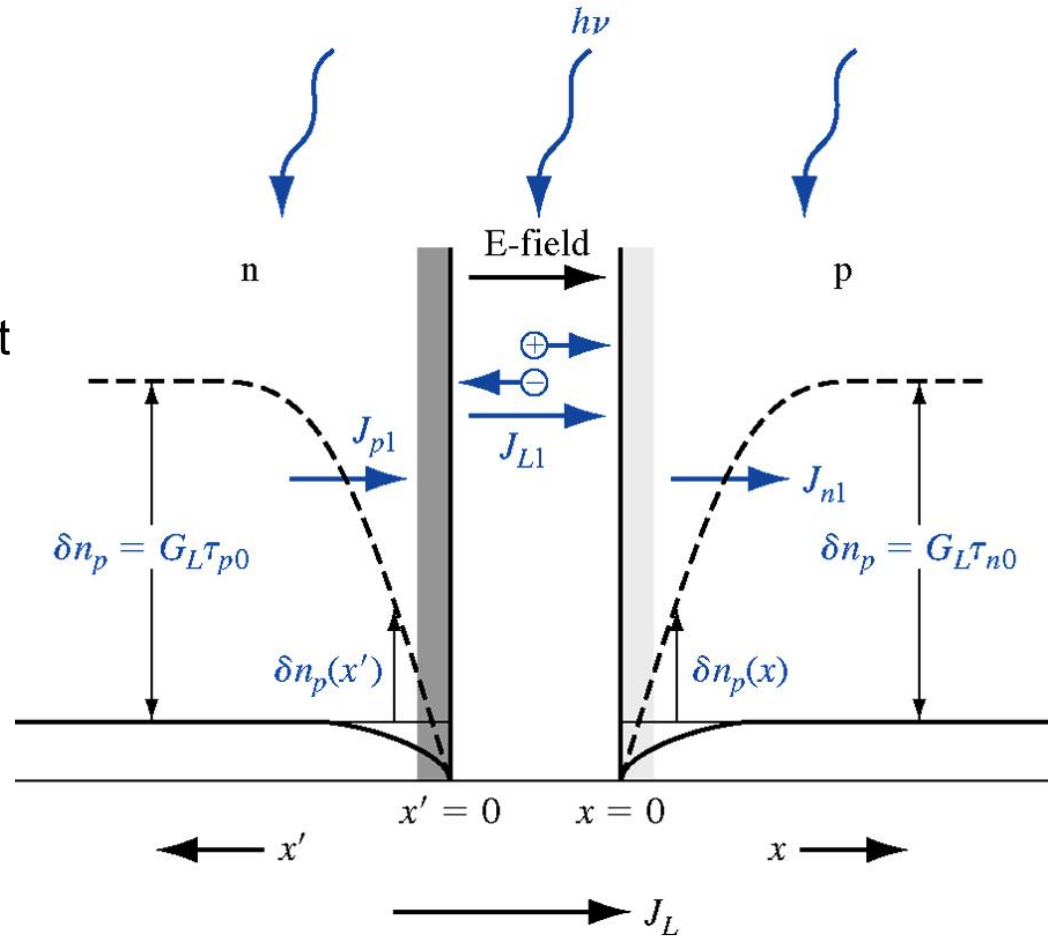
$$J_{L1} = e \int G_L dx = eG_L W$$

und die Verstärkung ist gleich eins.

Photogeneration passiert auch in den neutralen Bahngebieten. Minoritätsträger diffundieren in die RLZ, und führen zu einem "verzögerten" Photostrom.

Der totale Photostrom ist

$$J_L = eG_L W + eG_L L_n + eG_L L_p = eG_L [W + L_n + L_p]$$





# PIN-Photodiode

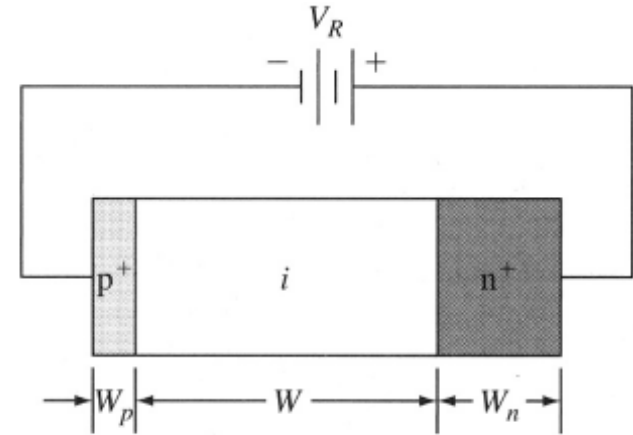
Für Anwendungen, bei denen es auf hohe Geschwindigkeiten ankommt, ist nur der photo-generierte “prompte” Strom von Interesse. Also muss die RLZ möglichst gross sein. Dies leisten *PIN-Dioden*.

Die RLZ erstreckt sich über die gesamte intrinsische Zone.

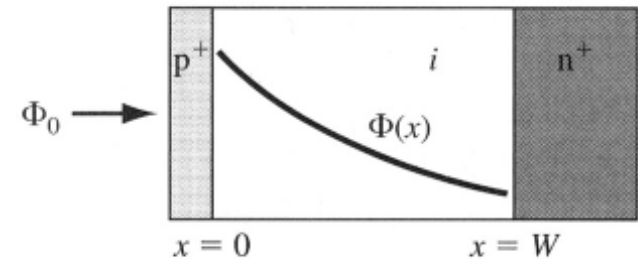
Der Photonenfluss  $\Phi_0$  treffe auf die  $p^+$ -Seite. Unter Vernachlässigung der Ausdehnung des  $p^+$ -Gebietes fällt der Photonenfluss in der intrinsischen Region wie  $\Phi_0 e^{-\alpha x}$  ab. Der generierte Photostrom ist dann

$$\begin{aligned} J_L &= e \int_0^W G_L dx = e \int_0^W \Phi_0 \alpha e^{-\alpha x} dx \\ &= e \Phi_0 \left[ 1 - e^{-\alpha W} \right] \end{aligned}$$

Dabei Annahme: keine Rekombination in der RLZ und Quantenausbeute = 1.



(a)



(b)

(a) Rückwärts gepolte PIN-Photodiode, (b) Geometrie mit inhomogener Lichtabsorption.

# Der Phototransistor

Bipolar-Transistor kann ebenfalls als Photodetektor genutzt werden. Dazu grosse Basis-Kollektor-Fläche und offene Basis. Im gesperrten Basis-Kollektor-Übergang generierte Löcher fließen ins  $p$ -Gebiet und laden die Basis positiv auf. Dies bewirkt Elektronen-Emission aus dem Emitter. Es gilt

$$I_C = \alpha I_E + I_L$$

Wegen der offenen Basis ist  $I_C = I_E$ . Daher

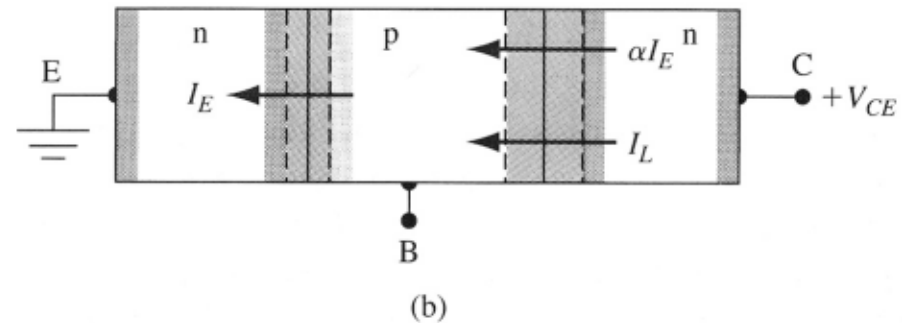
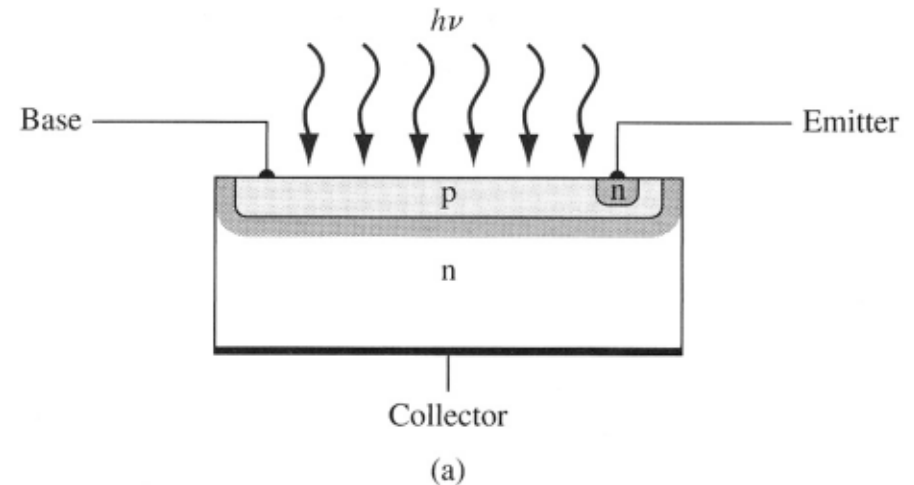
$$I_C = \alpha I_C + I_L$$

und

$$I_C = \frac{I_L}{1 - \alpha} = (1 + \beta) I_L$$

Die Verstärkung im Phototransistor ist  $1 + \beta$ .

Frequenzverhalten durch grosse B-C-Kapazität begrenzt!



(a) Photo-Bipolar-Transistor: (b) Blockdiagramm bei offener Basis.

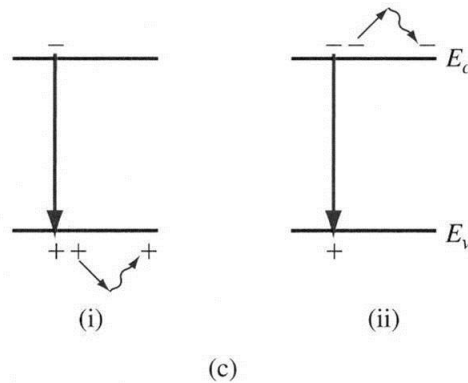
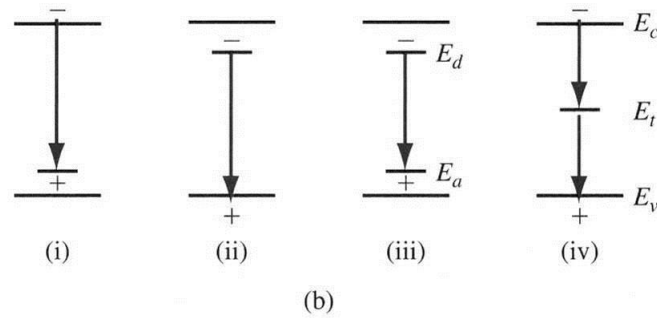
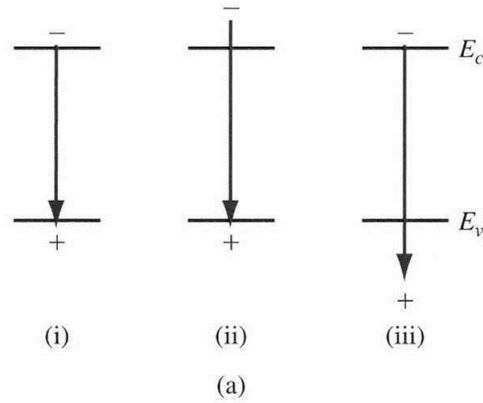
# 14.4 Photo- und Elektrolumineszenz

Die Eigenschaft, Licht zu emittieren, wird Lumineszenz genannt. Wenn Überschussladungen rekombinieren, werden Photonen generiert. Gelangen diese Photonen an die Halbleiter-Oberfläche, kann man von aussen Licht detektieren. Dazu dürfen sie nicht im Innern reabsorbiert werden.

In Abhängigkeit von der Art der Generation der Überschussladung unterscheidet man:

1. *Photolumineszenz*: Elektronen und Löcher werden durch Strahlung erzeugt.
2. *Elektro-(Injektions-)lumineszenz*: Elektronen und Löcher werden durch einen (Injektions)Strom infolge einer Spannung erzeugt.
3. *Thermolumineszenz*
4. *Kathodenlumineszenz*

# Grundlegende Übergänge in einem Halbleiter



# Das Emissionsspektrum

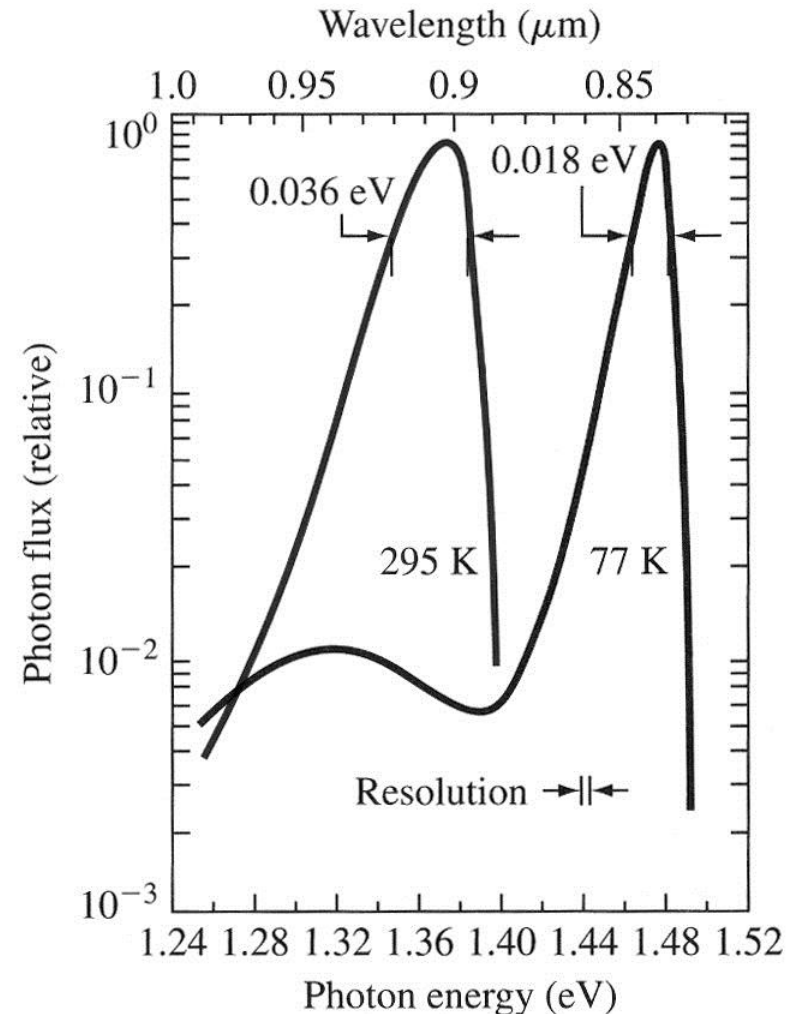
Rekombinationsprozesse können über einen breiten Energiebereich auftreten, je nachdem welche Energien die beteiligten Teilchen vor der Rekombination hatten.

Werden Elektron und Loch *nicht* durch eine externe Störung zur Rekombination angeregt, spricht man von **spontaner Emission**.

Die spontane Emissionsrate hat die Form

$$I(\nu) \sim \nu^2 \sqrt{h\nu - E_g} \exp\left(-\frac{h\nu - E_g}{kT}\right)$$

Die Bandbreite des Emissionsspektrums kann durch einen optischen Resonator (z.B. in Laserdiode) stark reduziert werden.



Emissionsspektrum einer GaAs-Diode bei  $T = 300 \text{ K}$  and  $T = 77 \text{ K}$  (nach Sze).

# Die Effizienz der Lumineszenz

Nicht alle Rekombinationsprozesse führen zur Generation von Strahlung. Die Quanteneffizienz  $\eta_q$  eines Materials ist das Verhältnis von strahlender zu totaler Rekombinationsrate

$$\eta_q = \frac{R_r}{R} = \frac{\tau_{nr}}{\tau_{nr} + \tau_r}$$

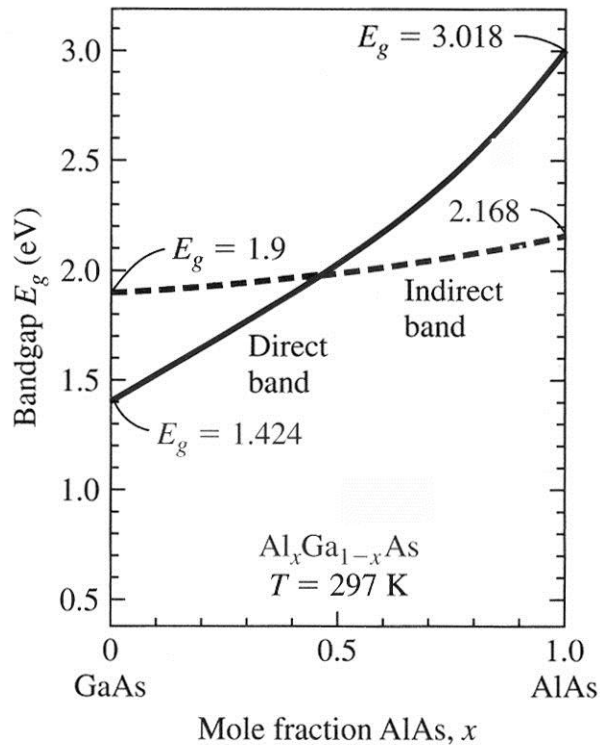
wobei  $R_r$  die strahlende Rekombinationsrate, und  $R$  die totale Rekombinationsrate der Überschussträger mit den entsprechenden Lebensdauern ist.

Ein Material hat dann eine hohe Effizienz in der Lumineszenz, wenn die nichtstrahlende Lebensdauer im Vergleich zur strahlenden gross ist.

Die Interband-Rekombinationsrate der Elektronen und Löcher hängt direkt von der Zahl der verfügbaren Elektronen und Löcher ab

$$R_r = Bnp$$

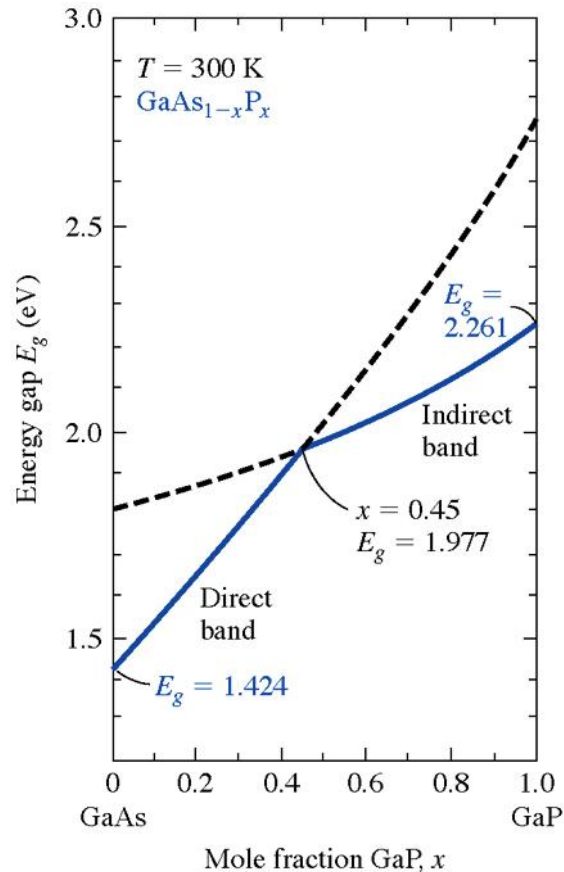
# Bandgap Engineering



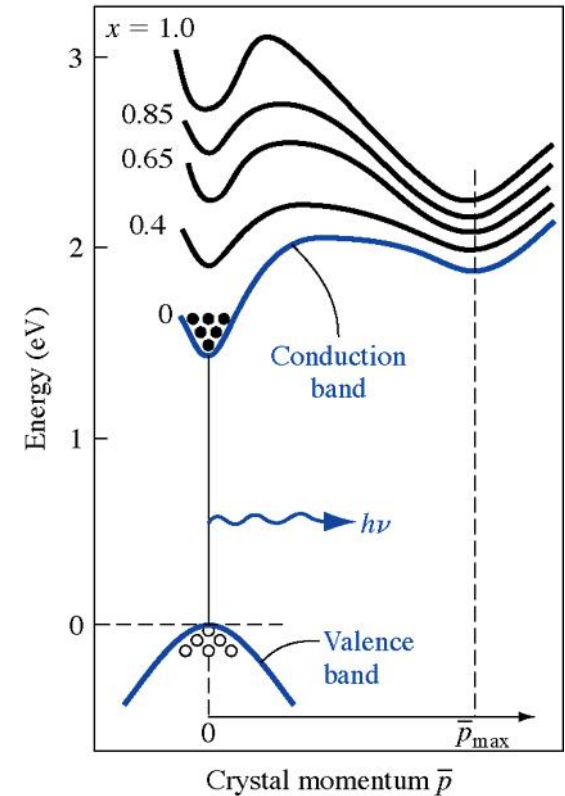
Energielücke von  $\text{Al}_x\text{Ga}_{1-x}\text{As}$  als Funktion des molaren Anteils  $x$ .

$$E_g = (1.424 + 1.247 x) \text{ eV}$$

$x < 0.35$



Energielücke von  $\text{GaAs}_{1-x}\text{P}_x$  als Funktion des molaren Anteils  $x$ .



$E(k)$ -Diagramm von  $\text{GaAs}_{1-x}\text{P}_x$  für verschiedene Werte des molaren Anteils  $x$ .

# 14.5 Die lichtemittierende Diode

In einer lichtemittierenden Diode (**LED**) wird durch Elektrolumineszenz Licht erzeugt. Die Diode wird dabei in Flussrichtung gepolt, wodurch Überschussladungsträger in die Raumladungszone injiziert werden. Hat der Halbleiter eine direkte Energielücke, so können Elektronen und Löcher unter Generation von Photonen rekombinieren.

Das Emissionsspektrum einer LED ist etwa 30 bis 40 nm breit, was schmal genug ist, um (im Sichtbaren) eine bestimmte Farbe zu beobachten.

Die Wellenlänge der emittierten Strahlung ist

$$\lambda = \frac{hc}{E_g} = \frac{1.24}{E_g} \mu m$$

Der LED Strom ist proportional zur Rekombinationsrate, womit auch der totale Photonenfluss direkt vom Diffusionsstrom in der Diode abhängt.



# Die interne Quanteneffizienz

Die IQE einer LED ist jener Anteil des Diodenstroms, welcher Lumineszenz produziert. Sie ist eine Funktion der Injektionseffizienz und des Anteils der strahlenden Rekombination an der totalen Rekombination.

Die Stromanteile einer Diode setzen sich aus dem Elektronen- und Löcherdiffusionsstrom und dem Rekombinationsstrom zusammen:

$$J = J_n + J_p + J_R$$

In einer GaAs LED wird die Lumineszenz durch Rekombination von Elektronen im p-Gebiet dominiert. Wir definieren die Injektionseffizienz

$$\gamma = \frac{J_n}{J_n + J_p + J_R}$$

In einer  $n^+p$  Diode kann  $\gamma$  nahezu eins werden, da der Löcherstrom klein ist und der Rekombinationsstrom durch eine grosse Vorwärtsspannung ebenfalls vernachlässigbar wird.

## Die interne Quanteneffizienz (2)

Überschusselektronen im p-Gebiet rekombinieren strahlend oder nicht-strahlend

$$R_r = \frac{\delta n}{\tau_r} \quad R_{nr} = \frac{\delta n}{\tau_{nr}} \quad R = R_r + R_{nr} = \frac{\delta n}{\tau}$$

Die Strahlungseffizienz ist durch den Anteil der strahlenden Rekombination definiert

$$\eta = \frac{R_r}{R_r + R_{nr}} = \frac{\tau}{\tau_r}$$

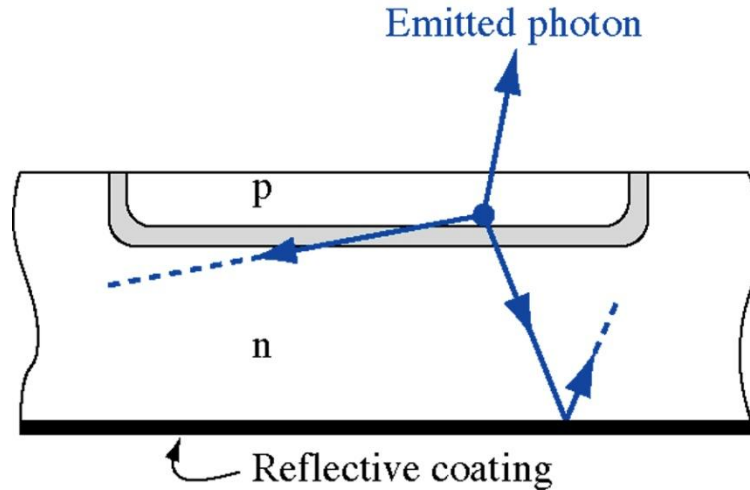
Die interne Strahlungseffizienz ist

$$\eta_i = \gamma \eta$$

# Die externe Quanteneffizienz

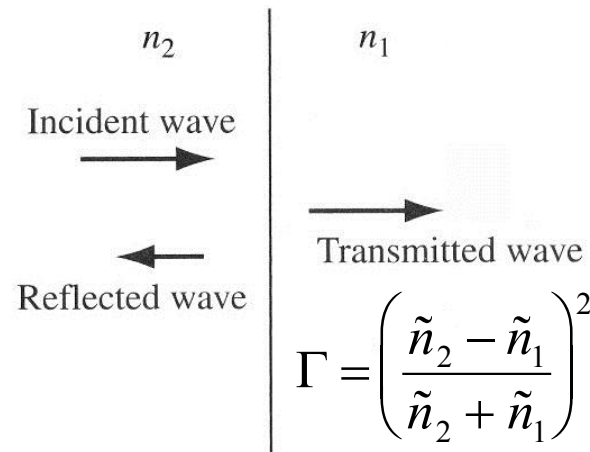
Die EQE ist jener Anteil an Photonen, welche effektiv aus dem Halbleiter emittiert werden. Bedingt durch interne Reabsorption, Fresnel-Verluste und totale interne Reflexion ist die EQE immer kleiner als die IQE.

Die EQE ist im Bereich weniger Prozente.

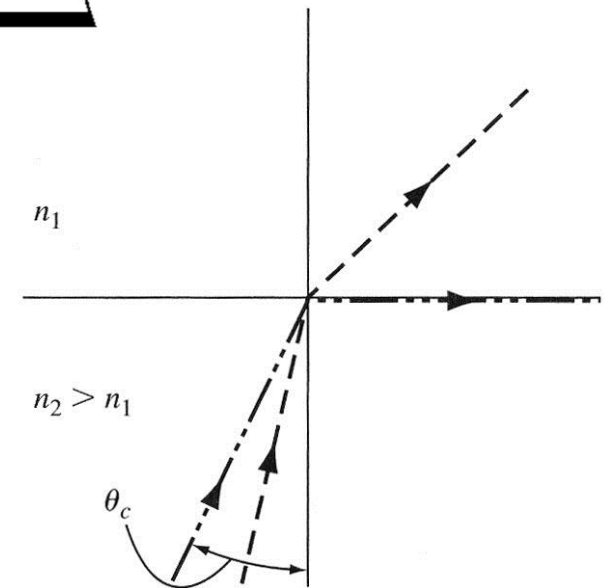


Snell:

$$\theta_c = \sin^{-1}\left(\frac{\tilde{n}_1}{\tilde{n}_2}\right)$$



Einfallende, reflektierte und transmittierte Photonen an einer dielektrischen Grenzfläche.



Lichtbrechung und totale interne Reflexion bei einem kritischen Winkel an einer dielektrischen Grenzfläche.

# Die weiße LED

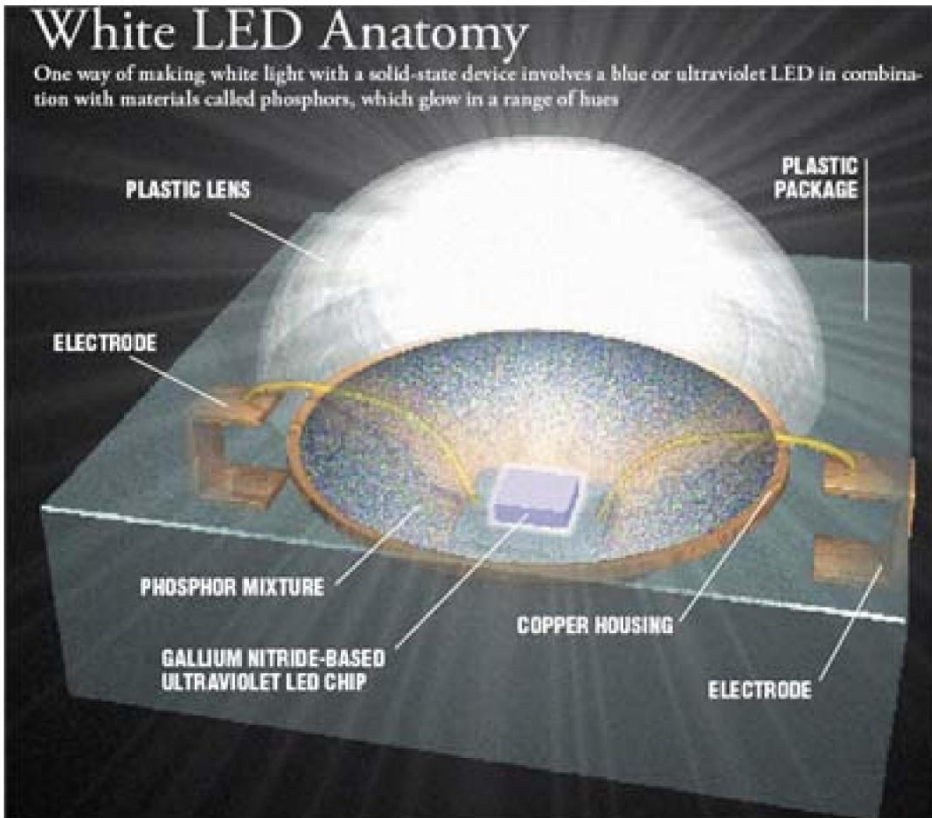
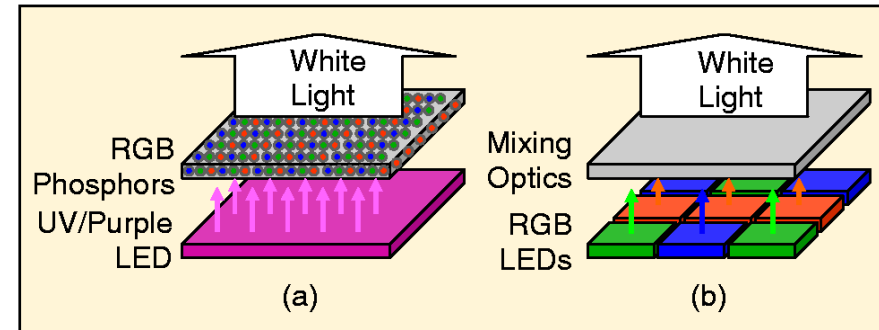
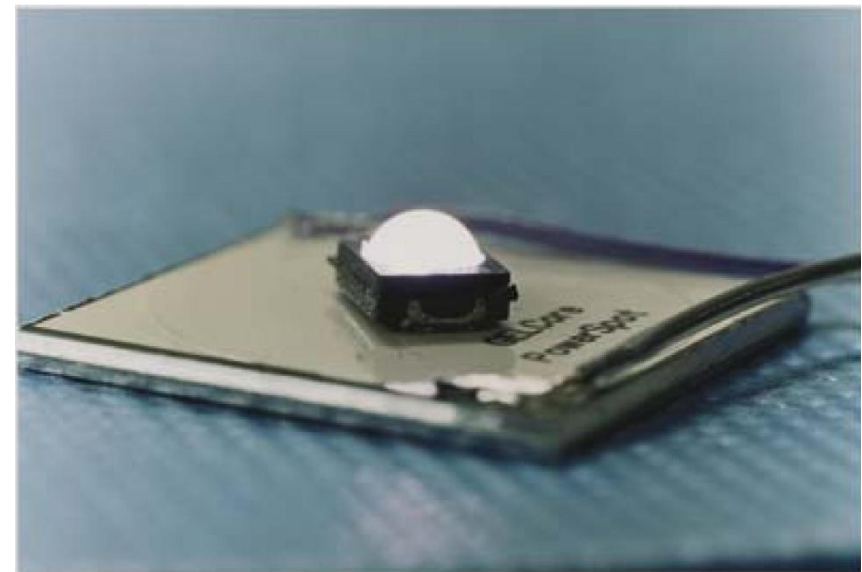


Figure 6. In this one-watt seven-millimeter device from GELcore, a one-millimeter chip emits ultraviolet light. The phosphor absorbs the ultra-violet light, emitting a broad spectrum of colors that mix to produce white illumination. A submount and copper housing remove heat.

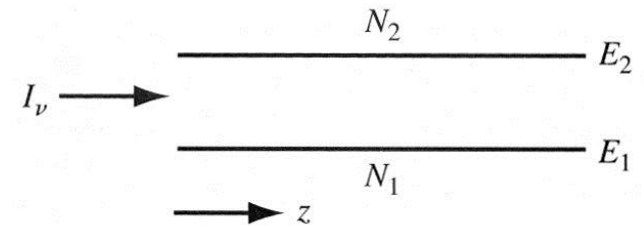
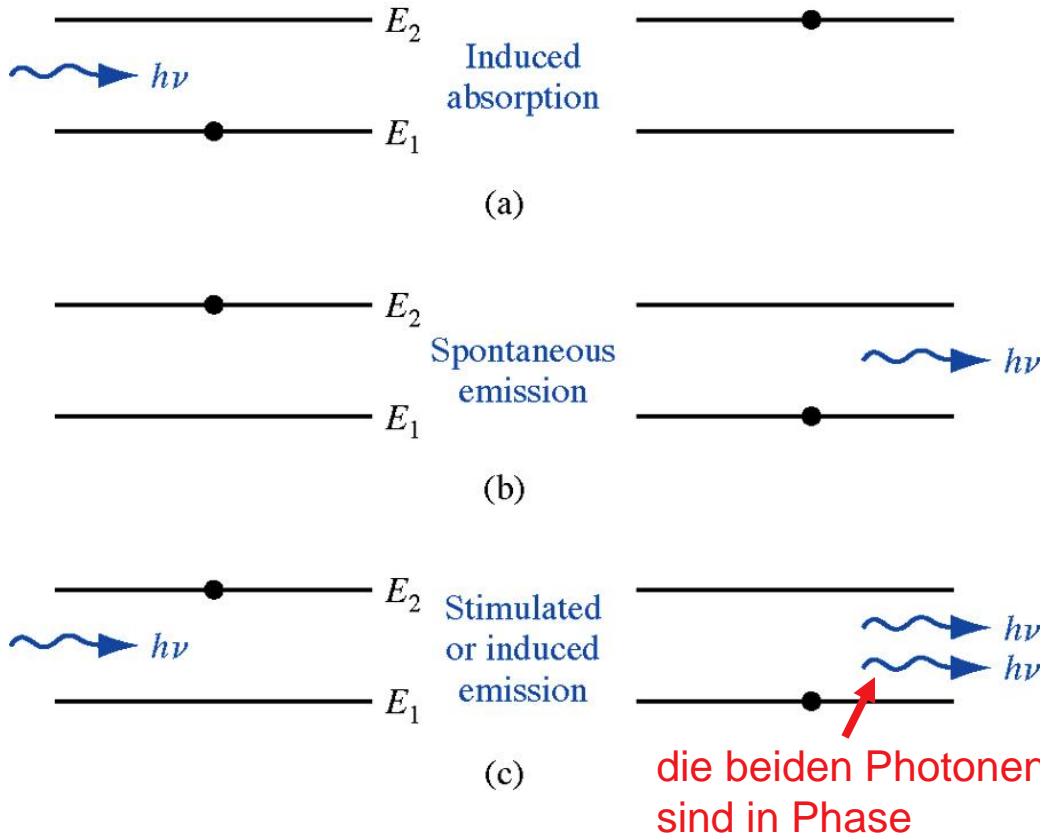


2. (a) Phosphor down-conversion and (b) color mixing approaches to SSL-LED lamps.



# 14.6 Die Laserdiode

## Stimulierte Emission und Besetzungsinversion



mit Boltzmann-Statistik gilt für die beiden Elektronendichten:

$$\frac{N_2}{N_1} = \exp\left[\frac{-(E_2 - E_1)}{kT}\right]$$

Laser-Betrieb wird erreicht, wenn  $N_2 > N_1 \rightarrow$  **Besetzungsinversion**

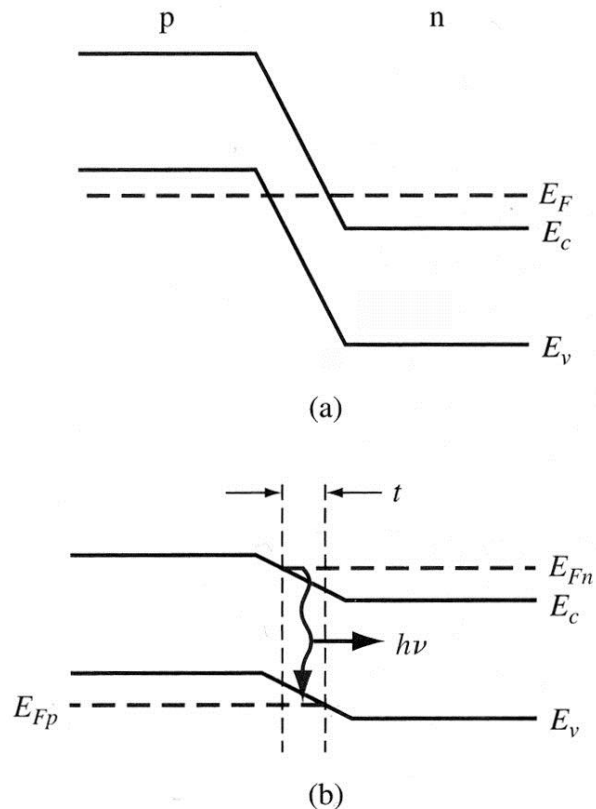
Die Änderung der Lichtintensität in  $z$ -Richtung ist

$$\frac{dI_\nu}{dz} = \gamma(\nu) I_\nu \sim (N_2 - N_1) I_\nu$$

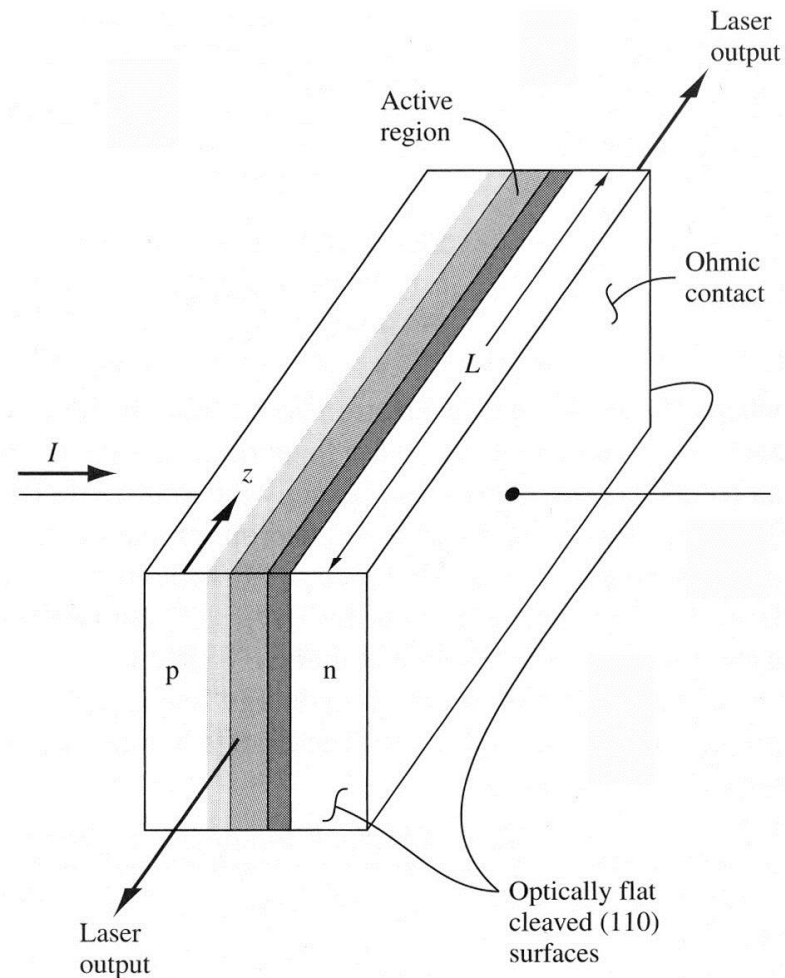
$$I_\nu = I_\nu(0) e^{\gamma(\nu)z}$$

# Stimulierte Emission und Besetzungsinversion (2)

$$\gamma(\nu) \sim \left\{ 1 - \exp\left[ \frac{h\nu - (E_{Fn} - E_{Fp})}{kT} \right] \right\}$$



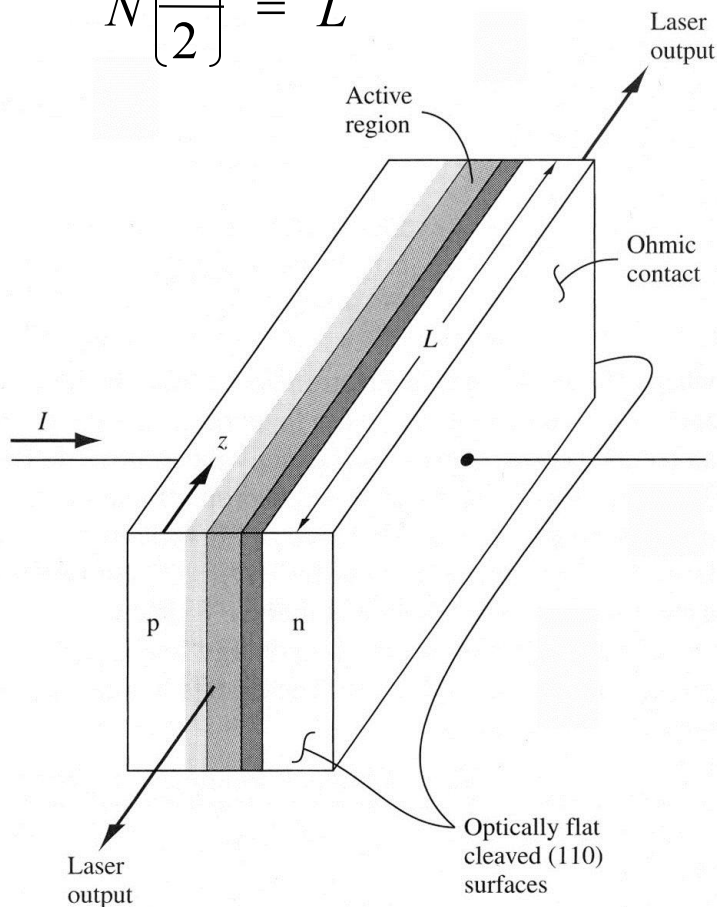
Hochdotierter (entarteter) pn-Übergang (a) ohne Spannung, (b) unter Vorwärtsspannung mit Photonen-Emission.



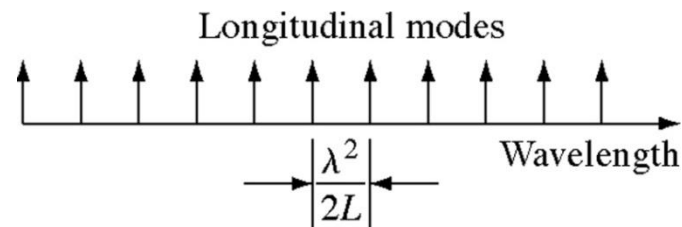
Laser-Diode mit Fabry-Perot-Hohlraum, gebildet von gespaltenen (100)-Flächen.

# Der optische Hohlraum

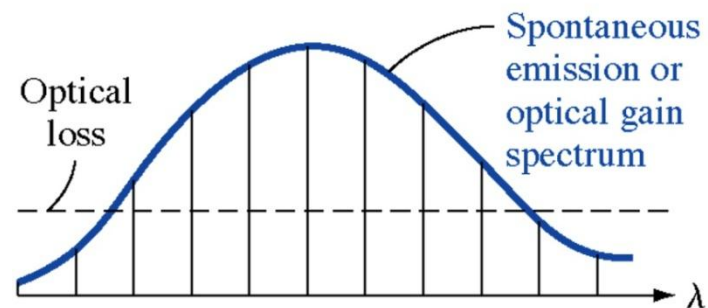
$$N \left\lfloor \frac{\lambda}{2} \right\rfloor = L$$



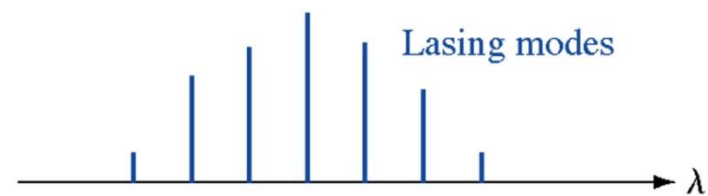
Laser-Diode mit Fabry-Perot-Hohlraum, gebildet von gespaltenen (100)-Flächen.



(a)



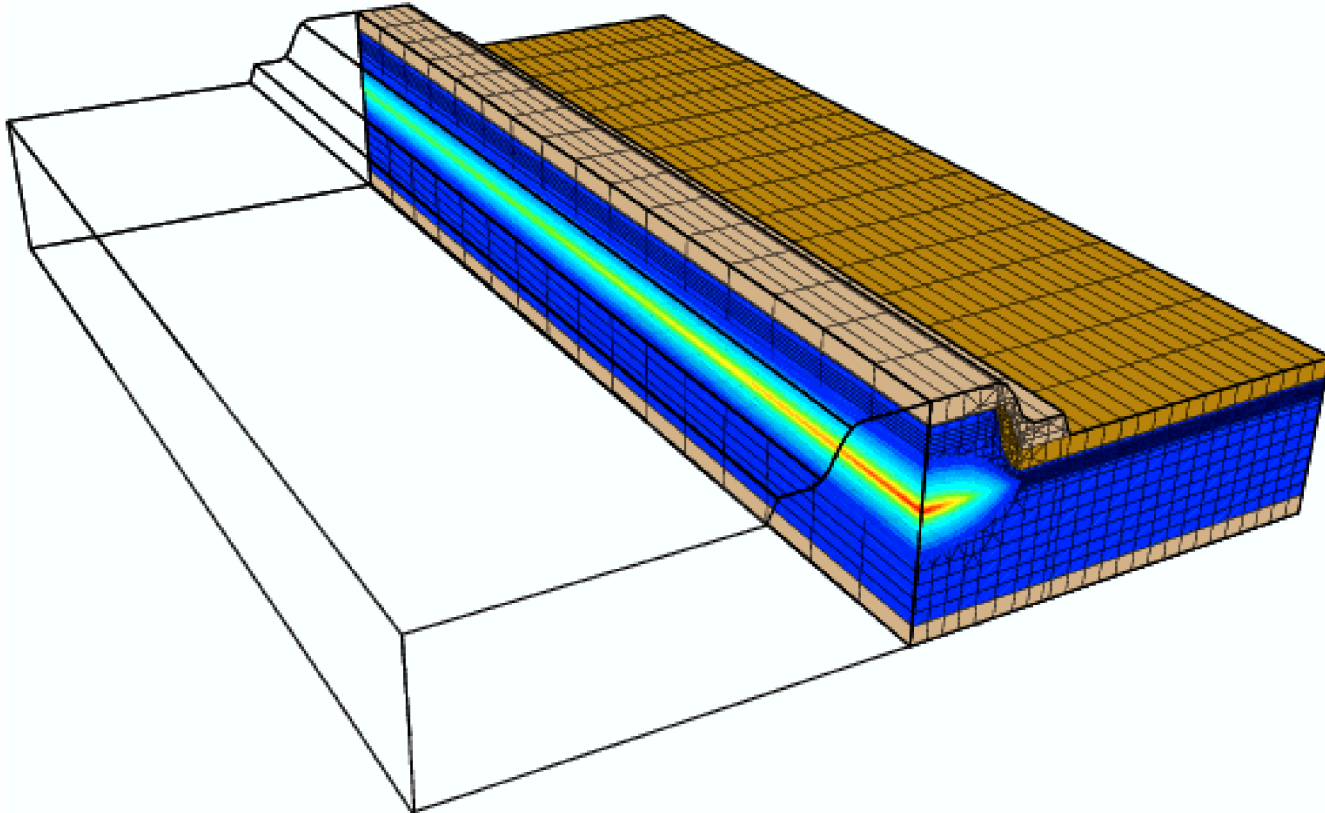
(b)



(c)

(a) Resonanz-Moden eines Hohlraums der Länge  $L$ , (b) Kurve der spontanen Emission, (c) tatsächliche Emissionsmoden einer Laserdiode.

# Edge-Emitting Laser





# Multisection Edge-Emitting Laser

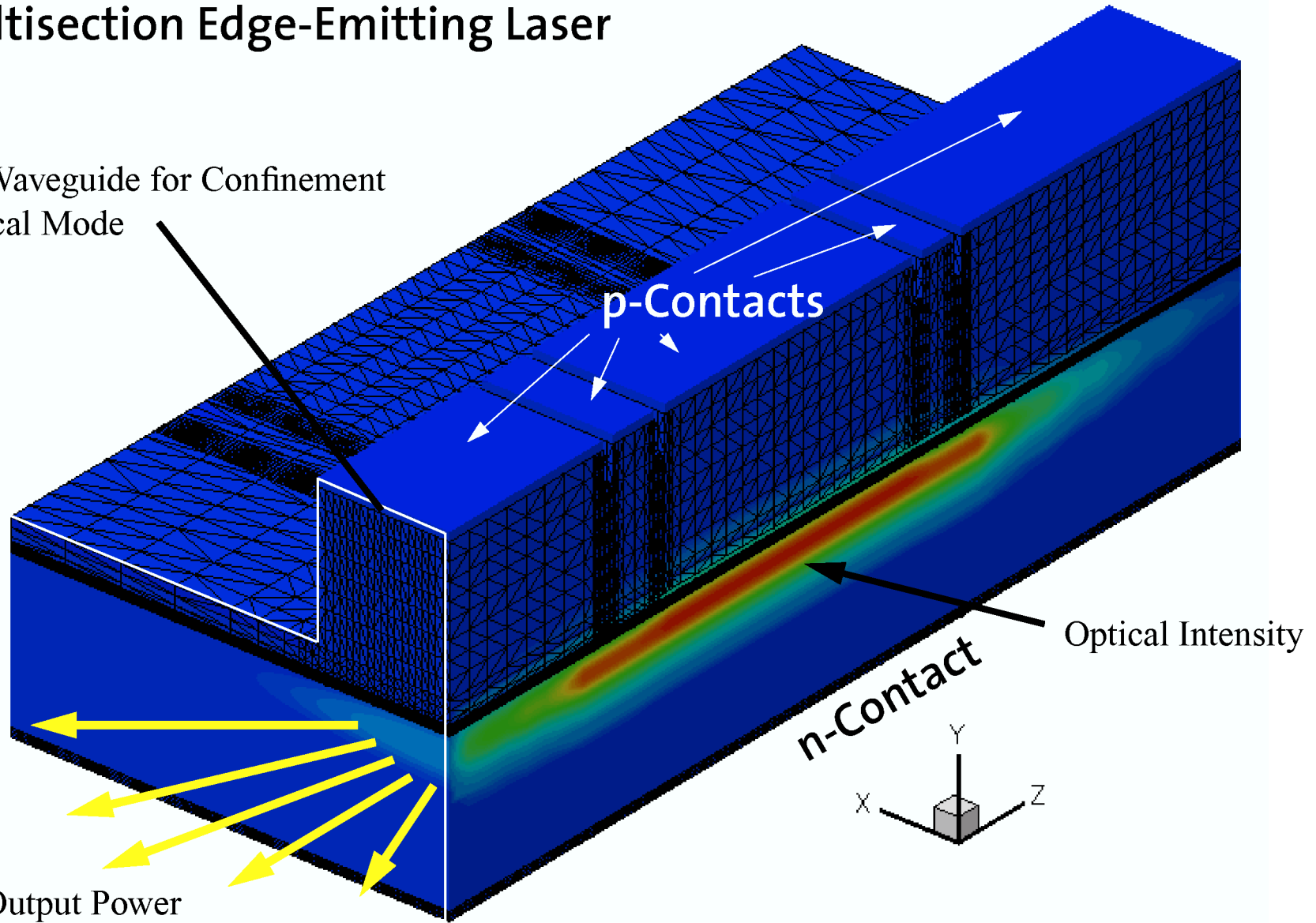
Ridge Waveguide for Confinement of Optical Mode

p-Contacts

Optical Intensity

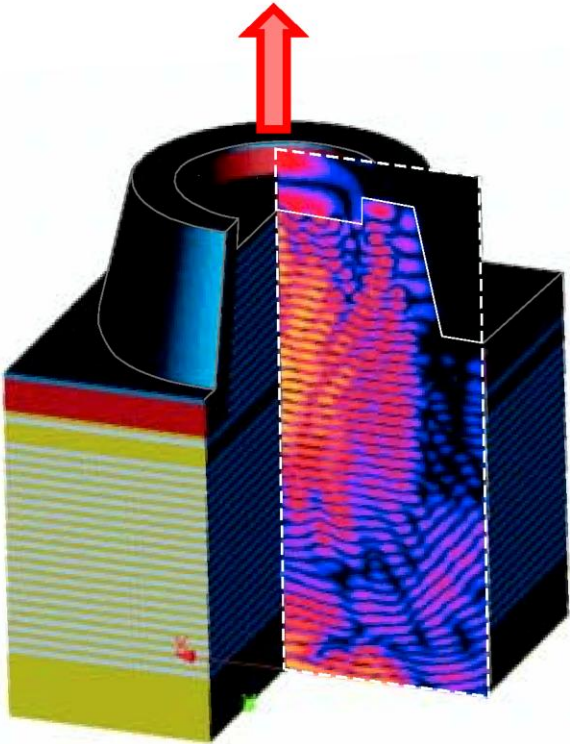
n-Contact

Laser Output Power (Edge-Emitting)

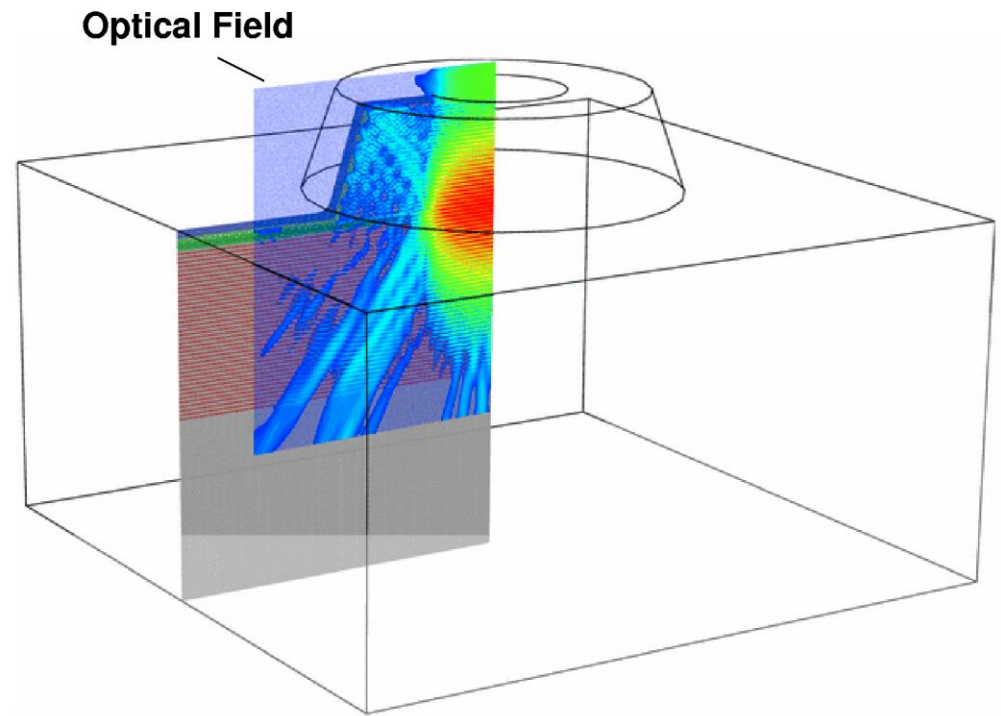


# Vertical-Cavity Surface-Emitting Laser (VCSEL)

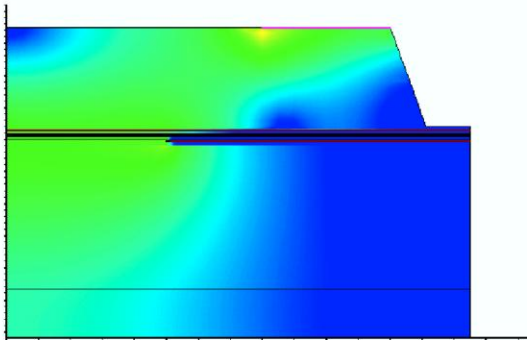
Light Output



## VCSEL Field Characteristics



Current Density



Temperature

