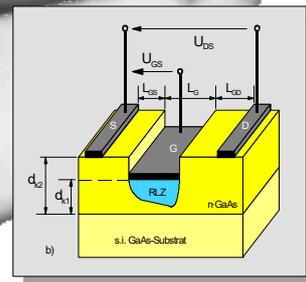
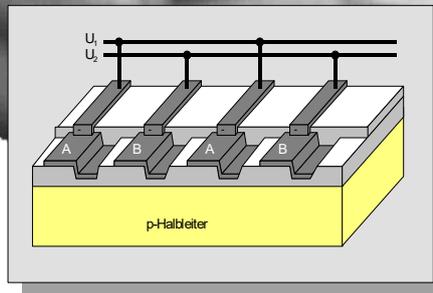
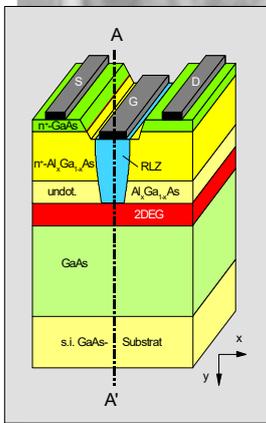


Elektronische Bauelemente



Universität Duisburg-Essen

Campus Duisburg

Fachbereich Ingenieurwissenschaften

Elektrotechnik und Informationstechnik

Bauelemente der Höchstfrequenzelektronik / Halbleitertechnologie

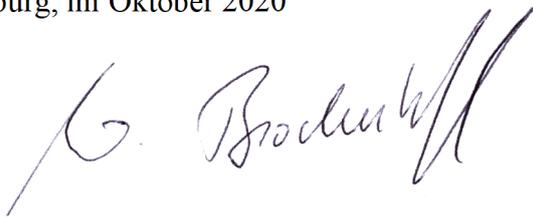
Vorwort

Das vorliegende Skript zur Vorlesung "*Elektronische Bauelemente*" behandelt die Grundlagen elektronischer Bauelemente wie MOS-Kondensatoren, Feldeffekttransistoren und bipolarer Bauelemente.

Es stellt eine Zusammenfassung des im Rahmen der Vorlesung vorgestellten Stoffes dar. Da das Skript aber nicht unbedingt den gesamten, für die Klausur relevanten Inhalt der Vorlesung abdeckt, soll und kann es auch die Vorlesung nicht ersetzen.

Darüber hinaus wird vor dem unkritischen Gebrauch des vorliegenden Manuskripts gewarnt. Zwar wurde versucht, für alle Berechnungen einheitliche Formelzeichen und eine einheitliche Vorzeichenregelung zu verwenden. Mögliche Fehler sind jedoch nicht auszuschließen.

Duisburg, im Oktober 2020

A handwritten signature in black ink, appearing to read 'G. Brockerhoff', with a stylized flourish at the end.

Dr.-Ing. Wolfgang Brockerhoff

Inhalt

0 Festkörperelektronik - eine kurze Zusammenfassung	0-2
0.1 Atommodell und Bandstruktur	0-2
0.1.1 Das Bohr'sche Atommodell	0-2
0.1.2 Das Festkörpermodell und die Bandstruktur des Festkörpers	0-4
0.2 Die elektronischen Eigenschaften des Festkörpers	0-5
0.2.1 Isolatoren, Metalle und Halbleiter	0-5
0.2.1.1 Isolatoren	0-7
0.2.1.2 Halbleiter	0-7
0.2.1.3 Metalle	0-8
0.2.2 Die Konzentration von Elektronen und Löchern in Festkörpern	0-10
0.2.2.1.1 Zustandsdichte $g(W)$	0-11
0.2.2.1.2 Die Fermi-Energie und die Fermi-Dirac-Verteilungsfunktion $f(W)$	0-11
0.2.2.2 Besetzungsstatistik in Halbleitern; Eigen- und Störstellenhalbleiter	0-12
0.2.2.2.1 Eigenhalbleiter (intrinsischer Halbleiter)	0-12
0.2.2.2.1.1 Elektronenkonzentration im Leitungsband:	0-12
0.2.2.2.1.2 Löcherkonzentration im Valenzband:	0-13
0.2.2.2.1.3 Die Eigenleitungskonzentration n_i :	0-14
0.2.2.2.2 Störstellenhalbleiter	0-15
0.2.2.2.2.1 Dotierung von Halbleitern; Donatoren, Akzeptoren, tiefe Störstellen	0-15
0.2.2.2.3 Ladungsträgerstatistik im Störstellenhalbleiter	0-18
0.2.3 Ladungstransport in Festkörpern	0-21
0.2.3.1 Die Ladungsträgerbeweglichkeit und die elektrische Leitfähigkeit	0-22
0.2.3.2 Ströme infolge eines elektrischen Feldes: Feldströme	0-26
0.2.3.3 Ströme infolge eines Konzentrationsgradienten: Diffusionsströme	0-27
0.3 Halbleiter-Festkörper-Grenzflächen	0-29
0.3.1 Das Metall-Isolator-Halbleiter-System, MIS- und MOS-Kondensatoren	0-29
0.3.1.1 Bändermodell und Ladungsverteilung	0-29
0.3.1.1.1 Die ideale MIS-Struktur	0-29
0.3.1.1.2 Die nicht ideale MIS-Struktur	0-32
0.3.1.1.3 Die Kapazitäts-Spannungs-Kennlinie der MIS-Diode	0-34
0.3.2 Metall-Halbleiter-Grenzflächen; Schottky-Dioden und Ohm-Kontakte	0-36
0.3.2.1 Das Bändermodell des Metall-Halbleiter-Kontaktes	0-37
0.3.2.2 Die Kapazitäts-Spannungskennlinie	0-38
0.3.2.3 Die Strom-Spannungs-Kennlinie des Metall-Halbleiter-Kontaktes	0-40
0.3.2.4 Ohm-Kontakte	0-42
0.4 Halbleiter-Halbleiter-Grenzflächen	0-44
0.4.1 Der pn-Übergang zwischen Halbleitern gleichen Materials	0-44
0.4.1.1 Stromloser pn-Übergang; Bänderdiagramm und Ladungsverteilung	0-44
0.4.1.2 Die Strom-Spannungscharakteristik des pn-Übergangs	0-49
0.4.1.3 Die Sperrschichtkapazität des pn-Übergangs	0-55
1 Der MOS-Kondensator	1-2
1.1 Physikalische Grundlagen	1-2
1.2 Anwendungen des MOS-Kondensators	1-5
1.2.1 Ladungsgekoppelte Bauelemente CCD	1-5
1.2.1.1 Prinzip der Signalspeicherung in CCD-Sensoren	1-5
1.2.1.2 Der Signaltransport	1-6

1.2.1.2.1	Drei-Phasen-CCD	1-8
1.2.1.2.2	Zwei-Phasen-CCD	1-10
1.2.1.3	Prinzipieller Aufbau der Fotodetektoren in CCD-Zellen	1-10
1.2.1.4	CCD mit Fotodiode	1-12
1.2.1.5	CCD mit Foto-Gate	1-13
1.2.1.6	Sensortypen	1-14
1.2.1.6.1	Interline-Transfer-Sensor	1-14
1.2.1.7	Frame-Transfer-Sensor	1-15
1.2.1.8	Full-Frame-Sensor	1-16
1.2.2	CCD-Farbkameras	1-16
1.2.2.1	Drei-Chip Kamera	1-16
1.2.2.2	Ein-Chip Kamera	1-17
1.2.2.3	Weitere Bauformen	1-18
1.2.3	Geschaltete Kapazitäten	1-19
1.2.3.1	Wirkungsweise	1-19
1.2.3.2	Anwendungsbeispiele	1-20
1.2.3.2.1	Filter	1-20
1.2.3.2.2	Integrator mit Operationsverstärkern	1-22
1.2.3.3	Realisierung in MOS-Technik	1-23
2	Feldeffekttransistoren	2-2
2.1	MOS-Feldeffekttransistoren	2-4
2.1.1	Aufbau und Wirkungsweise	2-4
2.1.2	Berechnung des Drain-Stroms	2-6
2.1.2.1	Der MOSFET als steuerbarer Widerstand	2-9
2.1.2.2	Der MOSFET im Sättigungsgebiet	2-10
2.1.3	Die Kennlinienfelder des MOSFET	2-15
2.1.3.1	Die Kennlinienfelder der Source-Schaltung	2-16
2.1.3.1.1	Kennlinien der p-Kanal-Anreicherungstypen (Enhancement-FET, E-FET)	2-17
2.1.3.1.2	Kennlinien der n- und p-Kanal-Verarmungstypen (Depletion-FET, D-FET)	2-19
2.1.4	Temperaturabhängigkeit der Kennlinien	2-20
2.1.5	Einfluss der Substratvorspannung	2-20
2.1.6	Arbeitspunkteinstellung und -stabilisierung	2-22
2.1.6.1	n-Kanal-Anreicherungs-MOSFET (E-MOSFET)	2-22
2.1.6.2	n-Kanal-Verarmungs-MOSFET (D-MOSFET)	2-23
2.1.7	Komplementäre MOS-Technik (CMOS)	2-24
2.2	Sperrschicht-Feldeffekttransistoren	2-28
2.2.1	Aufbau und Wirkungsweise	2-29
2.2.2	Kennlinienberechnung	2-32
2.3	Der Heterostruktur-Feldeffekttransistor (HFET)	2-37
2.3.1	Der Heteroübergang	2-37
2.3.2	Der Ladungstransport im 2DEG	2-39
2.3.3	Der Aufbau eines Heterostruktur-Feldeffekttransistors	2-40
2.3.4	Berechnung der Schichtkonzentration	2-42
2.3.4.1	Der Gleichgewichtsfall	2-43
2.3.4.2	Der Einfluss des Schottky-Kontaktes auf die Schichtkonzentration	2-47
2.3.4.3	Gleichstromverhalten	2-48
2.4	Das Elektrische Ersatzschaltbild	2-49
2.4.1	Das Gleichstrom- (DC-) verhalten des realen Feldeffekttransistors	2-50

2.4.2	Das Wechselspannungs- (AC-) Verhalten	2-53
2.4.2.1	Funktionsprinzip	2-53
2.4.2.2	Allgemeine Beschreibung des Kleinsignalverhaltens	2-54
2.4.2.3	Quasistationäres Kleinsignalverhalten von Feldeffekttransistoren	2-56
3	Dioden	3-2
3.1	pn-Dioden	3-2
3.1.1	Die Grundgleichungen des pn-Übergangs	3-2
3.1.2	Der pn-Übergang ohne Vorspannung	3-4
3.1.3	Der stromdurchflossene pn-Übergang	3-7
3.1.4	Die reale Diode	3-15
3.1.5	Der Temperatureinfluss	3-17
3.1.5.1	Einfluss der Temperatur auf den Sperrstrom I_o	3-17
3.1.5.2	Einfluss der Temperatur auf U_D	3-18
3.1.5.3	Einfluss der Temperatur auf die gesamte Kennlinie	3-19
3.1.6	Das Elektrische Ersatzschaltbild der pn-Diode	3-21
3.1.6.1	Das Gleichstrom- (DC-) Verhalten	3-21
3.1.6.1.1	Das quasistationäre Kleinsignalverhalten von pn-Dioden	3-22
3.2	Spezielle Dioden	3-24
3.2.1.1	Varaktor-Diode (Kapazitätsdiode)	3-24
3.2.2	Zenerdiode	3-28
3.2.2.1	Tunnel-(Zener-)Effekt	3-29
3.2.2.2	Lawineneffekt	3-32
3.2.3	Tunnel-Diode	3-35
3.2.4	Backward-Diode	3-36
3.2.5	pin-Diode	3-37
3.2.6	Schottky-Dioden	3-39
3.3	Kennzeichnung	3-40
3.3.1	Europäische Kennzeichnung (Pro Electron)	3-40
3.3.2	JEDEC	3-41
4	Bipolartransistoren	4-2
4.1	Funktionsweise	4-2
4.2	Aufbau und Technologie der Bipolartransistoren	4-4
4.3	Die Gleichstromeigenschaften der Bipolartransistoren	4-8
4.3.1	Das Ebers-Moll Modell	4-8
4.3.2	Die Kennlinienfelder des Bipolartransistors	4-15
4.3.2.1	Die Basis-Schaltung	4-15
4.3.2.1.1	Die Eingangskennlinie (EKL bzw. EKF)	4-15
4.3.2.1.2	Das Ausgangskennlinienfeld (AKF)	4-17
4.3.2.1.2.1	Das AKF mit U_{EB} als Steuerparameter ($I_C = f(U_{CB}, U_{EB})$)	4-18
4.3.2.1.2.2	Das AKF mit I_E als Steuerparameter ($I_C = f(U_{CB}, I_E)$)	4-20
4.3.2.2	Die Emitterschaltung	4-23
4.3.2.2.1	Die Eingangskennlinie $I_B = f(U_{BE})$	4-23
4.3.2.2.2	Das Ausgangskennlinienfeld (AKF)	4-24
4.3.2.2.2.1	Das AKF mit U_{BE} als Steuerparameter ($I_C = f(U_{CE}, U_{BE})$)	4-24

4.3.2.2.2	AKF mit I_B als Steuerparameter ($I_C = f(U_{CE}, I_B)$)	4-26
4.3.3	Das Transportmodell	4-33
4.3.4	Der Gummel-Plot	4-36
4.3.5	Der Early-Effekt	4-36
4.4	Das quasistationäre Kleinsignalverhalten des Bipolartransistors in Emitter-Schaltung	4-38
4.4.1	Der Eingangswiderstand	4-38
4.4.2	Die Stromverstärkung	4-40
4.4.3	Die Steilheit	4-41
4.4.4	Der Ausgangswiderstand	4-42
4.4.5	Das quasistationäre Kleinsignal-Ersatzschaltbild	4-43
4.5	Der Schaltvorgang im Bipolartransistor	4-43
4.5.1	Sperrbereich	4-44
4.5.2	Aktiver Bereich	4-45
4.5.3	Sättigungsbereich	4-45
4.6	Grenzwerte	4-48
4.6.1	Durchbruchverhalten	4-48
4.6.1.1	Durchbruch 1. Art	4-48
4.6.1.1.1	Basis-Emitter Sperrspannung	4-48
4.6.1.1.2	Kollektor-Basis Sperrspannung	4-48
4.6.1.1.3	Kollektor-Emitter Sperrspannung	4-48
4.6.1.2	Durchbruchverhalten 2. Art	4-48
4.6.2	Grenzströme	4-49
4.6.3	Sperrströme	4-49
4.6.4	Maximale Verlustleistung	4-49
4.7	Der Heterostrukturbipolartransistor (HBT)	4-53
5	Weitere Bipolare Bauelemente	5-2
5.1	Thyristoren	5-2
5.1.1	Aufbau und Wirkungsweise	5-2
5.1.2	Technologie von Thyristoren	5-8
5.1.2.1	Thyristor mit Mesastruktur	5-8
5.1.2.2	Thyristor mit Isolationsdiffusion und geätzten Isolationsgräben	5-9
5.2	Sonderformen	5-10
5.2.1	Rückwärts leitender Thyristor (RCT)	5-10
5.2.2	DIAC	5-11
5.2.3	TRIAC	5-12
5.2.4	Abschaltbare Thyristoren (GTO)	5-12
5.3	Permeable Base Transistor (PBT)	5-14
5.4	Insulated-Gate Bipolar Transistor (IGBT)	5-15
6	Literatur	1
7	Stichwortverzeichnis	1

Kapitel 0

Festkörperelektronik - eine kurze Zusammenfassung

Inhalt

0 Festkörperelektronik - eine kurze Zusammenfassung	0-2
0.1 Atommodell und Bandstruktur	0-2
0.1.1 Das Bohr'sche Atommodell	0-2
0.1.2 Das Festkörpermodell und die Bandstruktur des Festkörpers	0-4
0.2 Die elektronischen Eigenschaften des Festkörpers	0-5
0.2.1 Isolatoren, Metalle und Halbleiter	0-5
0.2.1.1 Isolatoren	0-7
0.2.1.2 Halbleiter	0-7
0.2.1.3 Metalle	0-8
0.2.2 Die Konzentration von Elektronen und Löchern in Festkörpern	0-10
0.2.2.1 Zustandsdichte $g(W)$	0-11
0.2.2.1.2 Die Fermi-Energie und die Fermi-Dirac-Verteilungsfunktion $f(W)$	0-11
0.2.2.2 Besetzungsstatistik in Halbleitern; Eigen- und Störstellenhalbleiter	0-12
0.2.2.2.1 Eigenhalbleiter (intrinsischer Halbleiter)	0-12
0.2.2.2.1.1 Elektronenkonzentration im Leitungsband:	0-12
0.2.2.2.1.2 Löcherkonzentration im Valenzband:	0-13
0.2.2.2.1.3 Die Eigenleitungskonzentration n_i :	0-14
0.2.2.2.2 Störstellenhalbleiter	0-15
0.2.2.2.2.1 Dotierung von Halbleitern; Donatoren, Akzeptoren, tiefe Störstellen	0-15
0.2.2.2.2.3 Ladungsträgerstatistik im Störstellenhalbleiter	0-18
0.2.3 Ladungstransport in Festkörpern	0-21
0.2.3.1 Die Ladungsträgerbeweglichkeit und die elektrische Leitfähigkeit	0-22
0.2.3.2 Ströme infolge eines elektrischen Feldes: Feldströme	0-26
0.2.3.3 Ströme infolge eines Konzentrationsgradienten: Diffusionsströme	0-27
0.3 Halbleiter-Festkörper-Grenzflächen	0-29
0.3.1 Das Metall-Isolator-Halbleiter-System, MIS- und MOS-Kondensatoren	0-29
0.3.1.1 Bändermodell und Ladungsverteilung	0-29
0.3.1.1.1 Die ideale MIS-Struktur	0-29
0.3.1.1.2 Die nicht ideale MIS-Struktur	0-32
0.3.1.1.3 Die Kapazitäts-Spannungs-Kennlinie der MIS-Diode	0-34
0.3.2 Metall-Halbleiter-Grenzflächen; Schottky-Dioden und Ohm-Kontakte	0-36
0.3.2.1 Das Bändermodell des Metall-Halbleiter-Kontaktes	0-37
0.3.2.2 Die Kapazitäts-Spannungskennlinie	0-38
0.3.2.3 Die Strom-Spannungs-Kennlinie des Metall-Halbleiter-Kontaktes	0-40
0.3.2.4 Ohm-Kontakte	0-42
0.4 Halbleiter-Halbleiter-Grenzflächen	0-44
0.4.1 Der pn-Übergang zwischen Halbleitern gleichen Materials	0-44
0.4.1.1 Stromloser pn-Übergang; Bänderdiagramm und Ladungsverteilung	0-44
0.4.1.2 Die Strom-Spannungscharakteristik des pn-Übergangs	0-49
0.4.1.3 Die Sperrschichtkapazität des pn-Übergangs	0-55

0 Festkörperelektronik - eine kurze Zusammenfassung

0.1 Atommodell und Bandstruktur

0.1.1 Das Bohr'sche Atommodell

Die elektrische Leitfähigkeit eines Stoffes ist in erster Linie von der Anzahl der freien Ladungsträger abhängig. Der elektrische Strom entsteht durch den Transport von freien Elektronen, deren "Wanderung" die Struktur des leitenden Stoffes nicht beeinflusst. Die Atome des Kristalls bleiben fest an ihren Gitterplätzen gebunden.

Jedes Atom besitzt entsprechend seiner im Kern enthaltenen Protonenzahl eine positive Ladung, die durch die um den Kern kreisenden negativen Ladungen, die Elektronen, kompensiert wird. Die Protonenzahl ist gleich der Elektronenzahl. Für das chemische Verhalten des jeweiligen Elements sind die sogenannten Valenzelektronen - das sind die den Kern in den größten Abständen umkreisenden Elektronen - verantwortlich. Sie stellen die Bindung zu den benachbarten Atomen im Kristall her. Diese Valenzelektronen sind durch "äußere Störungen", wie z.B. mechanische Einwirkungen, vor allem aber elektrische Felder, am ehesten zu beeinflussen.

Jedes Elektron besitzt eine Energie, die mit dem Abstand vom Kern zunimmt.

Für das Einzelatom kann man nach der Quantentheorie von *Planck* den Elektronen **Energieschalen** zuordnen.

Das einfachste atomphysikalische System ist das Wasserstoffatom:

Es besteht aus **einem negativ geladenen Elektron** und dem einfach **positiv geladenen Atomkern**. Für die spätere **Beschreibung des Festkörpers** kann man sich diesen aus Einzelbausteinen ähnlich dem Wasserstoffatom zusammengesetzt denken.

Die Kraft, die zwischen zwei Ladungen auftritt, lässt sich berechnen entsprechend:

$$F = \frac{Q_1 Q_2}{4\pi \varepsilon_0 \varepsilon_r r^2} \quad (0.1)$$

wobei

Q_1 bzw. Q_2 die Ladungen des Kerns (H+) bzw. des Elektrons sind,

r der Abstand der Ladungen und

ε_0 bzw. ε_r , die dielektrischen Eigenschaften der Umgebung beschreibt.

Für das Wasserstoffatom gilt: $Q_1 = +q$, $Q_2 = -q$, $\varepsilon_r = 1$ und somit für die (Anziehungs-)Kraft, die zwischen beiden Ladungen auftritt, in Abhängigkeit vom Abstand r :

$$F(r) = -\frac{q^2}{4\pi \varepsilon_0 r^2} \quad (0.2)$$

Mit der Beziehung

$$F = -\frac{dV}{dr} \quad (0.3)$$

lässt sich das **energetische Potenzial** des Atomkerns bestimmen:

$$V(r) = -\frac{q^2}{4\pi \varepsilon_0 r} \quad (0.4)$$

Dabei bedeuten negative Werte, dass es sich um Bindungsenergie handelt. Es hat im Grenzfall $r \rightarrow \infty$ den Wert $V = 0$.

Diese Verhältnisse lassen sich anschaulich durch einen sogenannten **Potenzialtopf** mit endlich hohen Wänden beschreiben, in dem sich die Elektronen befinden (Abb. 0.1).

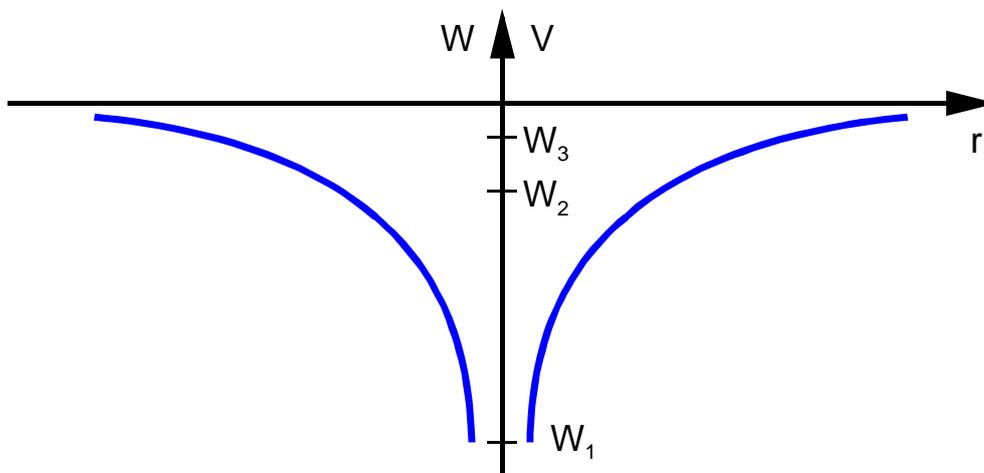


Abb. 0.1 Coulomb-Potenzial des H-Atoms

Ein größerer Abstand des Elektrons vom Atomkern ist mit einer höheren Energie des Elektrons gegenüber diesem Atomkern gleichbedeutend.

Dabei sind für das Elektron nach Planck aber nur diskrete Energiewerte (-niveaus), die Elektronenschalen, zulässig.

Das bedeutet: Ein größerer Abstand des Elektrons zum Atomkern, z.B. auf einer höheren Elektronenschale, entspricht einer höheren Energie dieses Elektrons gegenüber dem Atomkern.

Mit wachsendem Abstand wird andererseits die Kraft zwischen Atomkern und Elektron, also die Bindungsenergie, gemäß Gl. (0.1) kleiner.

Auf Basis festkörperphysikalischer Grundlagen lassen sich die **zulässigen** – diskreten – Energiewerte (Eigenwerte) berechnen. Die Berechnung selbst ist sehr aufwändig. Sie liefert unter anderem das sogenannte **Eigenwertspektrum** des H-Atoms:

$$W_n = -\frac{q^2}{8\pi \varepsilon_0 r_0} \cdot \frac{1}{n^2} \quad n = 1, 2, 3, \dots \text{ (Quantenzahlen)} \quad (0.5)$$

mit

$$r_o = \frac{4\pi \varepsilon_o \hbar^2}{m q^2} = 5,29 \cdot 10^{-9} \text{ cm} \quad (0.6)$$

dem **Bohr'schen Atomradius**, benannt nach dem Physiker *Niels Bohr*. n sind dabei die **Hauptquantenzahlen** und beschreiben die Nummerierung der Energiewerte, die gleichzeitig der Bindungsenergie des Elektrons zum Atomkern entsprechen.

Anschaulich beschrieben bedeutet das:

Aus dem Gleichgewicht zwischen Coulombkraft und Zentrifugalkraft ergeben sich dann für diese Bahnen die erlaubten Energieniveaus.

Die diskreten Energiewerte W_n rücken mit wachsenden n näher zusammen ($\Delta W_n \rightarrow 0$). Dies ist anschaulich auch verständlich, denn, wie oben beschrieben, nimmt mit wachsender Energie W_n des Elektrons die Bindung an den Kern aufgrund des Verlaufs von $V(r)$ ab.

Mit $W_n \rightarrow 0$ ist der Fall des **freien Elektrons** realisiert, das nun jede beliebige Energie aufnehmen kann. Das negative Vorzeichen der W_n spiegelt den Bindungscharakter im Gegensatz zum freien Elektron wider.

In diesem Zusammenhang ist der Begriff der **Ionisierungsenergie** W_{ion} wichtig, die aufgebracht werden muss, um das Elektron völlig vom H^+ Kern zu trennen, d.h. zu einem ungebundenen freien Elektron zu machen:

$$W_{ion} = W_{n \rightarrow \infty} - W_{n=1} = \frac{q^2}{8\pi \varepsilon_o r_o} = 13,6 \text{ eV} \quad (0.7)$$

Auf der Basis dieses Modells kann der Aufbau der Elemente mit steigender Protonenzahl im Kern und entsprechend steigender Elektronenzahl in der Atomhülle (ein vollständiges Atom ist elektrisch neutral!) beschrieben werden, wenn man zusätzlich das **Pauli-Prinzip** beachtet.

Das Pauli-Prinzip besagt, dass in einem atomaren System ein irgendwie ausgewähltes Teilchenpaar nicht exakt den gleichen Zustand einnehmen darf.

0.1.2 Das Festkörpermodell und die Bandstruktur des Festkörpers

Betrachtet man zunächst ein Molekül aus zwei Atomen, so ergeben sich aufgrund der gegenseitigen Krafteinwirkung der Atome Doppelniveaus (Abb. 0.2), d.h. die eigentlich auf gleichem Energieniveau W_1 liegenden Elektronen haben nicht exakt die gleiche Energie.

Je näher die beiden Atome zusammengebracht werden, desto stärker ist die gegenseitige Wechselwirkung und desto größer die Energieaufspaltung ΔW .

In einem Festkörper sind nun außerordentlich viele Atome mit jeweils mehr als nur einem Elektron auf engem Raum zusammengebracht ($N \approx 10^{22} \text{ cm}^{-3}$). Entsprechend viele Energieniveaus müssen zur Erfüllung des Pauli-Prinzips entstehen, und diese liegen derart dicht beieinander, dass sie wegen der Gültigkeit der Unschärferelationen (s. Vorlesung Festkörperelektronik) nicht unterscheidbar sind.

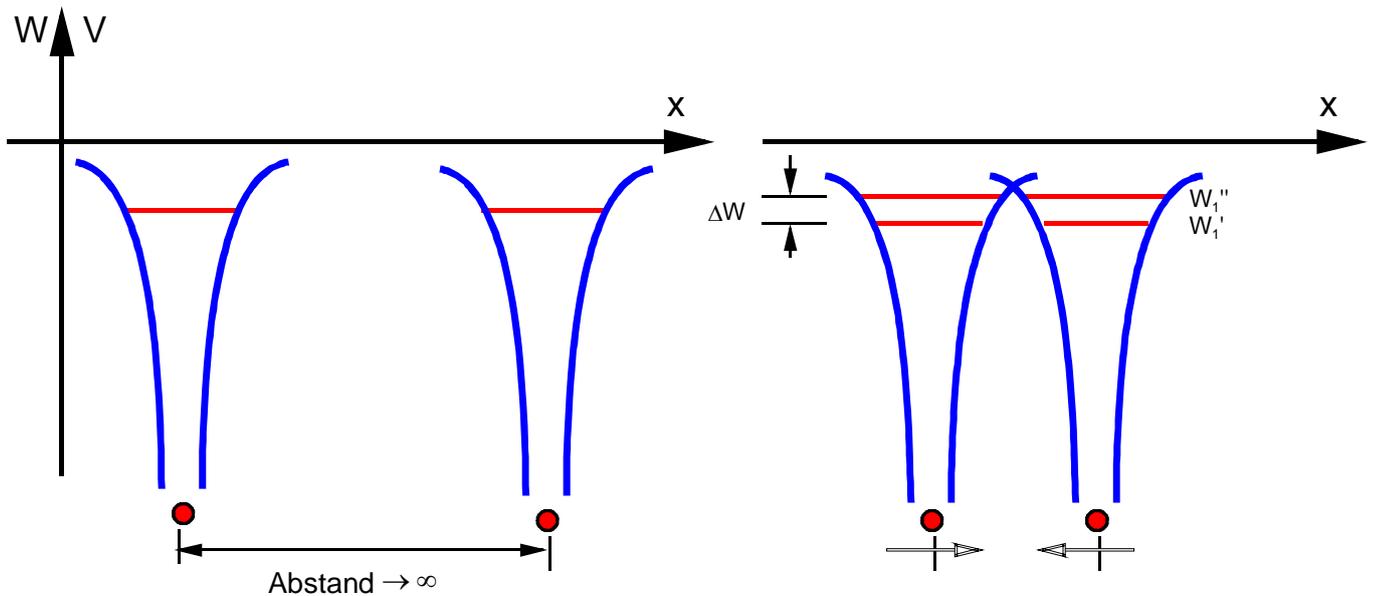


Abb. 0.2 Aufspaltung zweier identischer Energieniveaus aufgrund des Pauli Prinzips

Im Gegensatz zum Potenzialtopfproblem eines einzelnen Wasserstoffatoms, bei dem nur diskrete Energiewerte erlaubt waren, sind jetzt ganze Energiebereiche, die **Energiebänder**, erlaubt. Auf diesem Sachverhalt beruhen die Begriffe "**Bandstruktur**" oder "**Bändertheorie**" des Festkörpers (Abb. 0.3).

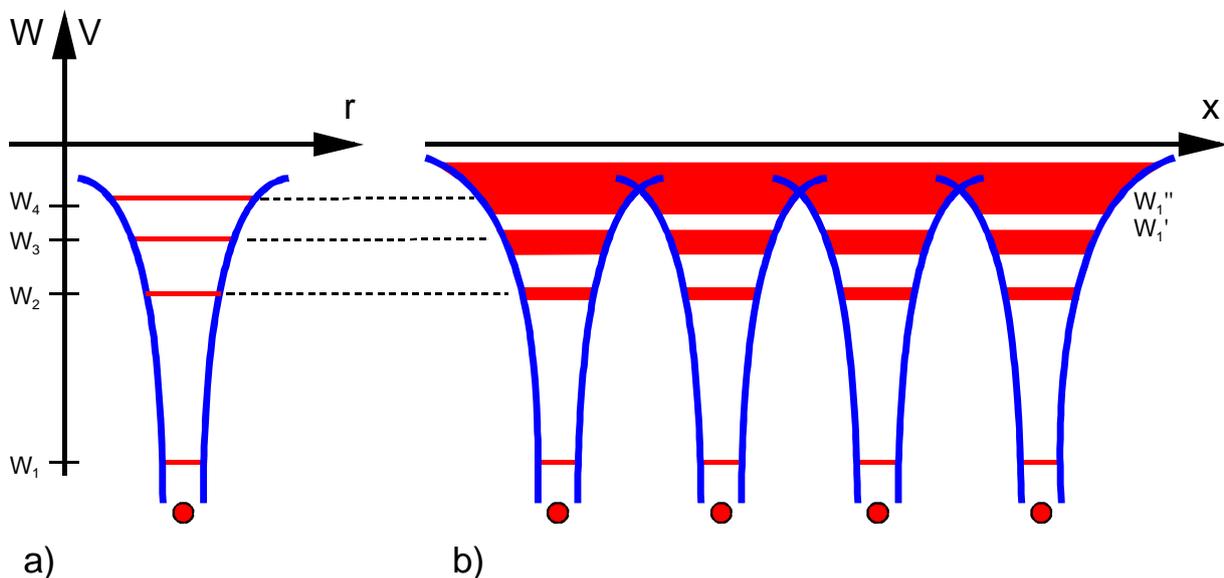


Abb. 0.3 Zum Zusammenhang zwischen den Energieniveaus des Einzelatoms (a) und den Energiebändern des Festkörpers (b) (eindimensional)

0.2 Die elektronischen Eigenschaften des Festkörpers

0.2.1 Isolatoren, Metalle und Halbleiter

Die **gegenseitige chemische Bindung der Festkörperatome** bewirken die Elektronen im **obersten** besetzten Band, weil gerade sie ja die Wechselwirkung mit den Nachbaratomen bewerkstelligen.

Man nennt sie **Valenzelektronen** und das dazugehörige Band **Valenzband**.

Die Besetzung des Valenzbandes und des nächst höheren Leitungsbandes entscheidet über die elektrische Leitfähigkeit des Festkörpers.

Die Energieaufspaltung in Bänder hängt, wie in Abb.0.2 gezeigt, vom Atomabstand a ab.

Je stärker die Einzelatome samt deren Elektronen in Wechselwirkung treten, desto weiter spalten die aus den diskreten Niveaus entstehenden Bänder auf, bis sie sich schließlich überlappen. In Abb. 0.4 sind qualitativ die Bandverhältnisse für die drei hinsichtlich der elektrischen Leitfähigkeit zu unterscheidenden Festkörper

- Isolator,
- Halbleiter und
- Metall

eingezeichnet: Isolator und Halbleiter unterscheiden sich nur quantitativ durch die Breite des verbotenen Energiebandes, der Energielücke W_g , die das Valenzband vom Leitungsband trennt. Bei Metallen liegen die Verhältnisse etwas komplizierter.

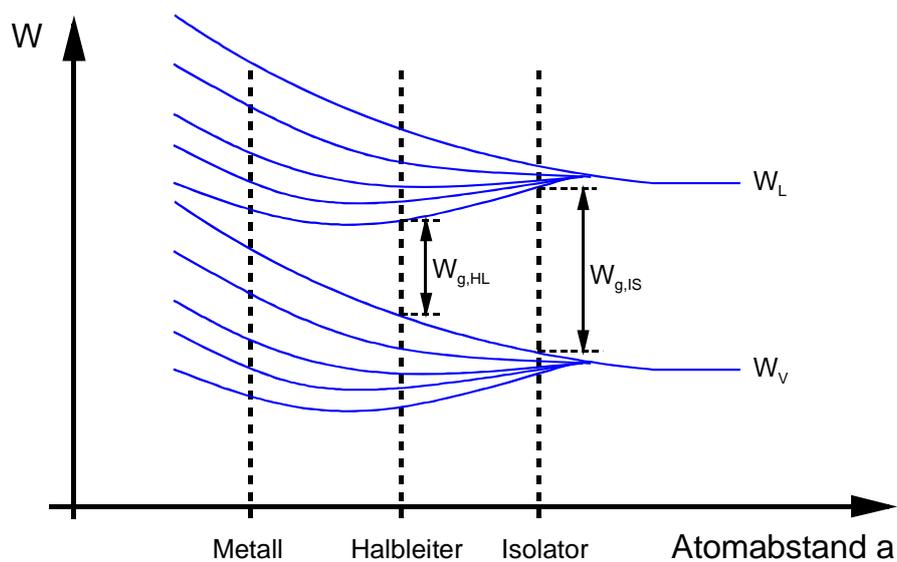


Abb. 0.4 Aufspaltung zweier Energieniveaus zu Bändern

0.2.1.1 Isolatoren

In **Isolatoren** ist das **Valenzband vollständig mit Elektronen** besetzt und das nächst höhere Band, das Leitungsband, ist völlig leer (Abb. 0.5).

Dies gilt zumindest bei $T = 0\text{K}$. Für die Bewegung eines Elektrons, d.h. für den Ladungstransport und somit den Stromfluss, muss dieses Elektron kinetische Energie aufnehmen, also einen anderen energetischen Zustand einnehmen. Dies ist aber im Valenzband nicht möglich, da alle verfügbaren Plätze bereits mit Elektronen besetzt sind. Die nächst höher liegenden erlaubten Energiewerte sind erst im Leitungsband anzutreffen, das aber durch die sehr große Energielücke von mehreren Elektronenvolt (eV) vom Valenzband getrennt ist. Derart hohe kinetische Energien kann das Elektron üblicherweise, z. B. aus einem elektrischen Feld, nicht aufnehmen.

Das bedeutet, dass im Isolator kein Strom fließt und somit der Name gerechtfertigt ist. Selbst wenn eine endliche Temperatur $T > 0\text{K}$ zugelassen wird, besitzen laut Statistik nur ganz vereinzelt Elektronen eine derart hohe thermische Energie, dass sie die Energielücke W_g überspringen können.

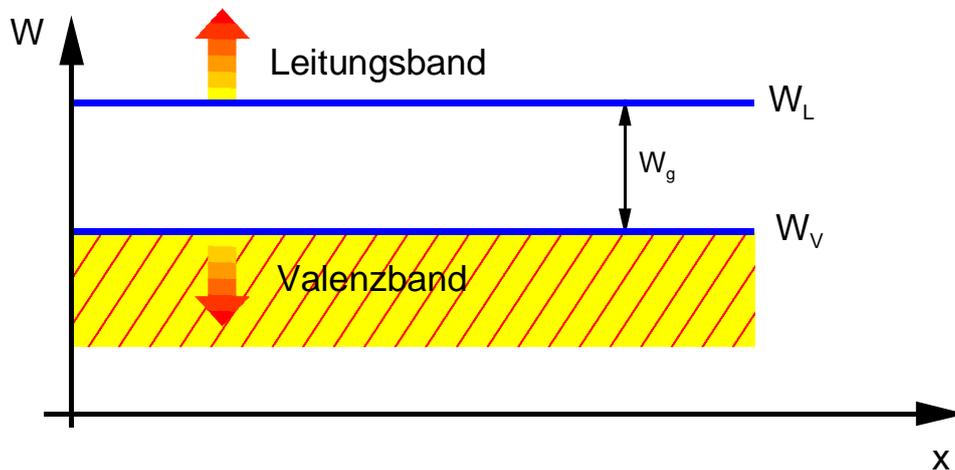


Abb. 0.5 Bändermodell des Isolators bei $T = 0\text{K}$ (gilt auch für intrinsische Halbleiter)

0.2.1.2 Halbleiter

Bei reinen Halbleitern liegen prinzipiell die gleichen Verhältnisse wie beim Isolator gemäss Abb. 0.5 vor, nur ist nun die Energielücke W_g geringer, etwa 1eV.

Hier ist bei endlichen Temperaturen $T > 0\text{K}$ die Wahrscheinlichkeit erheblich höher, dass einige Elektronen ausreichend thermische Energie zur Überwindung der Lücke besitzen.

Für die thermische Energie gilt:

$$W = k \cdot T \quad k: \text{ Boltzmann-Konstante } (k = 8,62 \cdot 10^{-5} \text{ eV} / \text{K})$$

und mit $T = 300\text{K}$ (Raumtemperatur) folgt:

$$W(300\text{K}) \approx 26\text{meV}$$

Die aufgrund der thermischen Energie angeregten Elektronen befinden sich dann im Leitungsband mit beliebig vielen freien Plätzen, die für die Elektronen nach Aufnahme kinetischer Energie zur Verfügung stehen. Damit ist Elektronenbewegung und folglich Stromtransport möglich. Die

Leitfähigkeit ist somit deutlich höher als bei Isolatoren, aber immer noch außerordentlich gering im Vergleich zu Metallen (siehe unten).

Man kann allerdings durch gezieltes Einbringen von Fremdatomen in das Gitter des reinen Halbleiters die elektrische Leitfähigkeit ganz erheblich erhöhen. Durch dieses sogenannte Dotieren kann gezielt sowohl Elektronen- als auch Löcherleitfähigkeit erzeugt werden.

Typische Halbleiter sind **Silizium** (Si) und **Germanium** (Ge) aus der 4. Hauptgruppe des Periodischen Systems. Es gibt aber neben diesen Elementhalbleitern auch **Verbindungshalbleiter**, die sich aus Atomen der 3. und 5. Gruppe (**III-V Halbleiter**) oder auch aus der 2. und 6. Gruppe (**II-VI Halbleiter**) zusammensetzen. **Galliumarsenid** (GaAs) ist ein wichtiger Vertreter der III-V Halbleiter und, da er aus zwei Komponenten besteht, ein **binärer Halbleiter**. Es gibt auch ternäre (z. B. Ga Al As) und quaternäre (z. B. In Ga As P), die sich aus drei bzw. vier Komponenten zusammensetzen, und deren Eigenschaften sich in weiten Grenzen über den Mischungsanteil der einzelnen Komponenten gezielt verändern lässt.

0.2.1.3 Metalle

Die Metalle unterscheiden sich grundsätzlich von Halbleitern und Isolatoren dadurch, dass besetzte und freie Zustände nicht durch eine Energielücke voneinander getrennt sind. Im Fall der monovalenten Metalle ist das Valenzband nur etwa zur Hälfte gefüllt (vgl. Abb. 0.6).

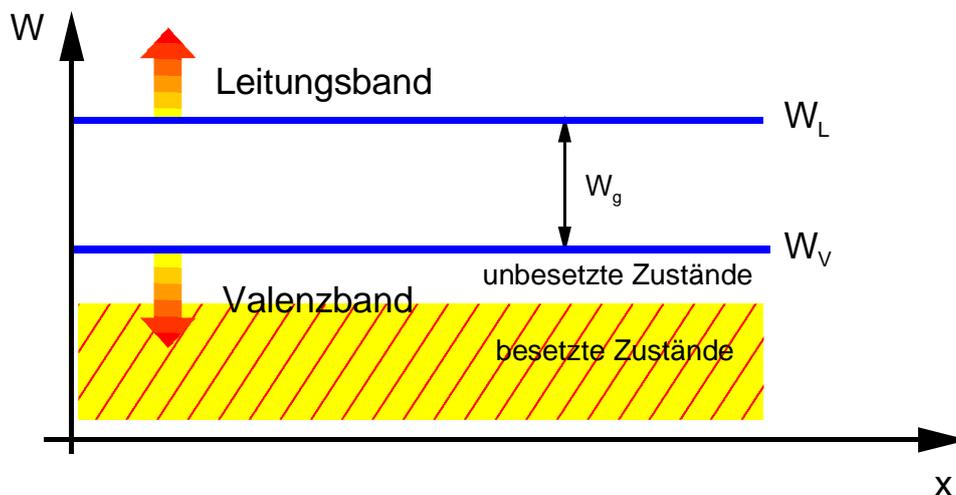


Abb. 0.6 Bändermodell für ein monovalentes Metall

Freie Plätze für Elektronen, die eine beliebige (kleine) Bewegungsenergie aufgenommen haben, finden also noch im Valenzband ihren neuen Platz, d.h. Stromtransport kann stattfinden.

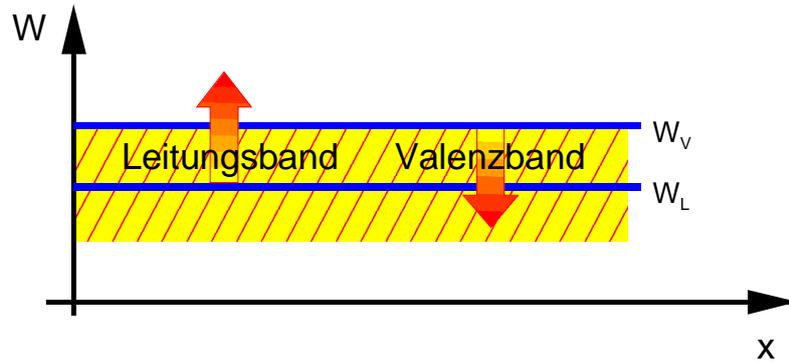


Abb. 0.7 Bändermodell eines Metalls mit überlappenden Bändern

Dies gilt sogar für alle im Valenzband befindlichen Elektronen, so dass im Vergleich zu den außerordentlich wenigen Elektronen im reinen Halbleiter, die die Energielücke überspringen können, eine um Größenordnungen höhere Elektronenkonzentration am Stromtransport teilnehmen kann, und daher auch eine entsprechend höhere elektrische Leitfähigkeit resultiert. Es gibt darüber hinaus einen anderen Typ Metall, bei dem das Leitungs- und das Valenzband überlappen (vgl. Abb. 0.4 und Abb. 0.7).

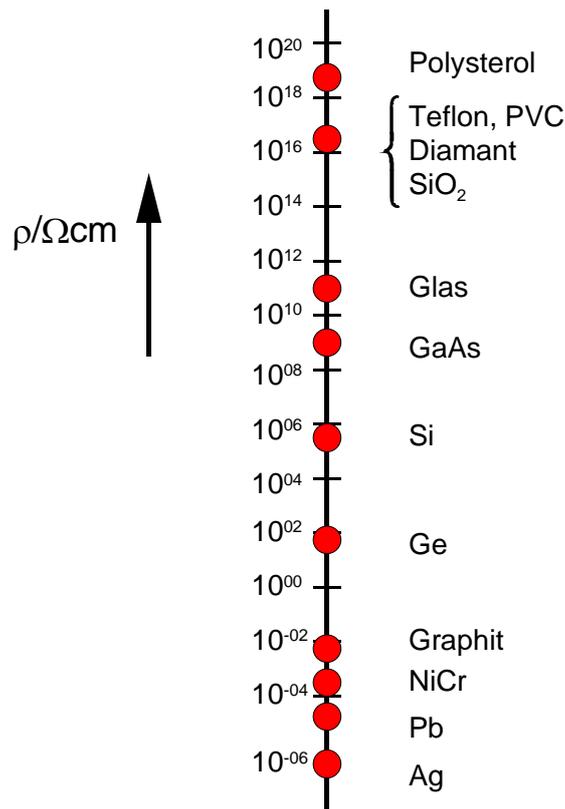
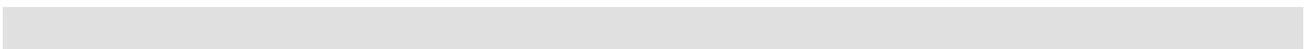


Abb. 0.8 Der spezifische Widerstand einiger Festkörper



Ein Vergleich der drei Typen von Festkörpern Isolator, Halbleiter und Metall liefert also:

Isolatoren und **Halbleiter** unterscheiden sich hinsichtlich der elektronischen Leitfähigkeit nicht qualitativ, sondern nur quantitativ in Form der Breite der Energielücke. Für beide ist die Zahl der Elektronen im Leitungsband bzw. die Zahl der freien Plätze im Valenzband maßgebend für die elektrische Leitfähigkeit. Im Vergleich zur Gesamtzahl der besetzbaren Zustände im Leitungsband bzw. der besetzten Zustände im Valenzband sind diese verschwindend gering. Daraus resultiert die niedrige Leitfähigkeit.

Bei den **Metallen** tragen nahezu alle Valenzelektronen zur Leitfähigkeit bei, weil im Valenzband etwa ebenso viele freie wie besetzte Plätze und etwa so viele Valenzelektronen wie Atome vorhanden sind. Abb. 0.8 zeigt im logarithmischen Maßstab einige Beispiele des spezifischen Widerstandes (das ist der Kehrwert der Leitfähigkeit) von Metallen bis zu Isolatoren

0.2.2 Die Konzentration von Elektronen und Löchern in Festkörpern

Die **Ladungsträgerkonzentration** n in Abhängigkeit von der Energie W lässt sich aus den Größen

- Zustandsdichte $g(W)$,
- Fermi-Energie W_F und
- Fermi-Dirac-Verteilungsfunktion $f(W)$

entsprechend dem Zusammenhang

$$n(W) = g(W) \cdot f(W) \quad (0.8)$$

bestimmen.

Dabei gibt die

- **Zustandsdichte** $g(W)$ die Dichte der möglichen besetzbaren Zustände pro Energieintervall und die
- **Verteilungsfunktion** $f(W)$ die Verteilung der Elektronen auf diese Zustände in Abhängigkeit von der Energie an. Das Produkt liefert die Verteilung der Elektronen über die Energiebänder.

Die **Gesamtzahl bzw. die Konzentration von Elektronen** in einem Band gewinnt man durch Integration über das Band

$$n = \int_{\text{Band}} n(W) dW = \int_{\text{Band}} g(W) f(W) dW \quad (0.9)$$

0.2.2.1.1 Zustandsdichte $g(W)$

Zur Berechnung der Zustandsdichte beschränkt man sich auf die Bereiche in unmittelbarer Nähe der Bandkanten, weil beim Halbleiter gerade die Elektronen an der Unterkante des Leitungsbandes bzw. an der Oberkante des Valenzbandes die Leitfähigkeit tragen.

Für die **Unterkante des Leitungsbandes** gilt:

$$g_L(W) = \frac{\sqrt{2}}{\pi^2 \hbar^3} (m^*)^{3/2} \cdot \sqrt{W - W_L}; \quad W > W_L \quad (0.10)$$

und für die Oberkante des Valenzbandes:

$$g_V(W) = \frac{\sqrt{2}}{\pi^2 \hbar^3} (m^*)^{3/2} \cdot \sqrt{W_V - W}; \quad W < W_V \quad (0.11)$$

mit der Energie der Leitungsbandkante W_L bzw. der Valenzbandkante W_V .

0.2.2.1.2 Die Fermi-Energie und die Fermi-Dirac-Verteilungsfunktion $f(W)$

Für $T = 0\text{K}$ bildet die **Fermi-Energie** W_F eine scharfe Grenze zwischen **besetzten** und **unbesetzten Zuständen**:

- alle Zustände unterhalb von W_F sind besetzt, d.h. die Verteilungsfunktion hat den Wert $f(W) = 1$;
- alle Zustände oberhalb von W_F sind unbesetzt, d.h. $f(W) = 0$.

Wird aber die Temperatur erhöht, so kann auf einige Elektronen thermische Energie übertragen werden. Als Folge davon nehmen einige Elektronen Energiewerte $W > W_F$ an und hinterlassen bei $W < W_F$ freie, unbesetzte Plätze.

Die Wahrscheinlichkeit, besetzte Plätze oberhalb von W_F und unbesetzte Plätze unterhalb von W_F zu finden, wächst also mit der Temperatur.

Diese Wahrscheinlichkeit wird durch die **Fermi-Dirac-Verteilungsfunktion** $f(W)$ beschrieben. Nach der Quantenstatistik gilt (ohne Herleitung):

$$f(W) = \frac{1}{e^{(W-W_F)/kT} + 1} \quad (0.12)$$

Diese Funktion liefert auch eine allgemeine, d.h. auch für alle Temperaturen gültige Definition der Fermi-Energie:

$$f(W_F) = \frac{1}{2} \quad (0.13)$$

0.2.2.2 Besetzungsstatistik in Halbleitern; Eigen- und Störstellenhalbleiter

Im folgenden werden nur noch die für die Herstellung elektronischer Bauelemente und Schaltungen wichtigen Halbleiter betrachtet. Man unterscheidet

- **Eigenhalbleiter**, das sind solche ohne Störstellen und Verunreinigungen und
- **Störstellenhalbleiter**, die meist gezielt mit Fremdatomen "dotiert" sind.

0.2.2.2.1 Eigenhalbleiter (intrinsischer Halbleiter)

Wie bereits bekannt, besitzen Halbleiter eine im Vergleich zu Isolatoren kleinere Lücke W_g zwischen Leitungs- und Valenzband, so dass bei Raumtemperatur (300K) schon einige Elektronen durch Zufuhr thermischer Energie aus dem Valenz- ins Leitungsband angehoben werden. Dann liegen freie Elektronen im Leitungsband und unbesetzte Plätze im Valenzband vor, die man als Löcher, also als Ladungsträger mit positiver Ladung interpretieren kann. Da Elektronen und Löcher im Eigenhalbleiter, wie soeben beschrieben, immer als Paar erzeugt werden, ist deren Konzentration n (Elektronen) und p (Löcher) gleich:

$$n = p = n_i \quad (0.14)$$

n_i heißt **intrinsische** oder **Eigenleitungskonzentration**.

Die soeben beschriebenen Verhältnisse erfordern, dass die Verteilungsfunktion im Valenzband Werte < 1 und im Leitungsband Werte > 0 annimmt, somit muss die Fermi-Energie in der Bandlücke zwischen W_L und W_V liegen.

Dabei interessieren jetzt die Konzentrationen n , p , n_i .

0.2.2.2.1.1 Elektronenkonzentration im Leitungsband:

Für die Elektronenkonzentration im Leitungsband gilt:

$$n = \int_{LB} n(W) dW = \int_{LB} g_L(W) \cdot f(W) dW$$

Für die Verteilungsfunktion $f(W)$ kann, da die Fermi-Energie W_{Fi} im intrinsischen Halbleiter um mehrere kT unterhalb der Leitungsbandkante liegt, folgende Näherung (**Boltzmann-Näherung**) benutzt werden:

$$f(W) = \frac{1}{e^{(W-W_{Fi})/kT} + 1} \approx e^{-(W-W_{Fi})/kT} \quad \text{für } W_L - W_{Fi} \gg kT \quad (0.15)$$

Die Integration ist von der Unterkante bis zur Oberkante des Leitungsbandes durchzuführen.

Daraus resultiert:

$$n = N_L \cdot e^{-(W_L - W_{Fi})/kT} \quad [n] = \text{cm}^{-3} \quad (0.16)$$

mit

$$N_L = \frac{1}{\sqrt{2} \hbar^3} \left(\frac{m_n^* kT}{\pi} \right)^{3/2} \quad [N_L] = \text{cm}^{-3} \quad (0.17)$$

N_L wird als **effektive Zustandsdichte des Leitungsbandes** bezeichnet. Sie gibt die Konzentration der Elektronen im Leitungsband an, die vorliegt, wenn die Fermi-Energie mit der Leitungsbandkante zusammenfällt.

0.2.2.2.1.2 Löcherkonzentration im Valenzband:

Analog gilt:

$$p = \int_{VB} p(W) dW = \int_{VB} g_V(W) \cdot (1 - f(W)) dW \quad (0.18)$$

Statt $f(W)$ tritt hier $(1 - f(W))$ auf, da die Löcher fehlende Elektronen an der Oberkante des Valenzbandes darstellen: Sind keine Löcher vorhanden, sind also alle Plätze im Valenzband mit Elektronen besetzt, so ist $f(W) = 1$ und $(1 - f(W)) = 0$.

	m_n^*/m_o	m_p^*/m_o	W_g/eV	n_i/cm^{-3}	N_L/cm^{-3}	N_V/cm^{-3}	ϵ_r
Halbleiter:							
Ge	0,55	0,37	0,66	$2,4 \cdot 10^{13}$	$1,0 \cdot 10^{19}$	$6,0 \cdot 10^{18}$	16,0
Si	0,98	0,49	1,1	$1,4 \cdot 10^{10}$	$3,0 \cdot 10^{19}$	$1,5 \cdot 10^{19}$	11,9
GaAs	0,067	0,45	1,42	$2,5 \cdot 10^6$	$4,7 \cdot 10^{17}$	$7,0 \cdot 10^{18}$	13,1
InAs	0,022	0,41	0,35	$2,4 \cdot 10^{15}$	$2,3 \cdot 10^{17}$	$1,8 \cdot 10^{19}$	12,5
InP	0,078	0,56	1,35	$6,6 \cdot 10^6$	$5,4 \cdot 10^{17}$	$2,9 \cdot 10^{18}$	12,4
InGaAs	0,045	0,43	0,76	$1,7 \cdot 10^{12}$	$6,8 \cdot 10^{17}$	$2,0 \cdot 10^{19}$	12,8
Isolator:							
Diamant (C)	----	0,25	5,47	≈ 0	$1,8 \cdot 10^{18}$	$2,5 \cdot 10^{18}$	

Tab. 0.1 Daten einiger Festkörper bei $T = 300\text{K}$

Mit ähnlichen Umformungen und Substitutionen wie für die Elektronenkonzentration folgt:

$$p = N_V \cdot e^{-(W_{Fi} - W_V)/kT} \quad [p] = \text{cm}^{-3} \quad (0.19)$$

mit

$$N_V = \frac{1}{\sqrt{2} \hbar^3} \left(\frac{m_p^* kT}{\pi} \right)^{3/2} \quad [N_V] = \text{cm}^{-3} \quad (0.20)$$

der **effektiven Zustandsdichte des Valenzbandes**.

0.2.2.2.1.3 Die Eigenleitungskonzentration n_i :

Nach Gl. (0.14) erhält man in Anlehnung an ähnliche Sachverhalte in der Chemie das **Massenwirkungsgesetz** für Elektronen und Löcher im Halbleiter:

$$\sqrt{n \cdot p} = n_i \text{ bzw. } n \cdot p = n_i^2 \tag{0.21}$$

und:

$$n_i = A \cdot T^{3/2} \cdot e^{-W_g/2kT} \tag{0.22}$$

mit

$$A = \frac{1}{\sqrt{2} \hbar^3} \left[(m_n^* m_p^*)^{1/2} \cdot \left(\frac{k}{\pi} \right) \right]^{3/2}$$

Die Ladungsträgerkonzentrationen $n = p = n_i$ hängen gemäss Gl. (0.22) von der Temperatur und der Bandlücke ab.

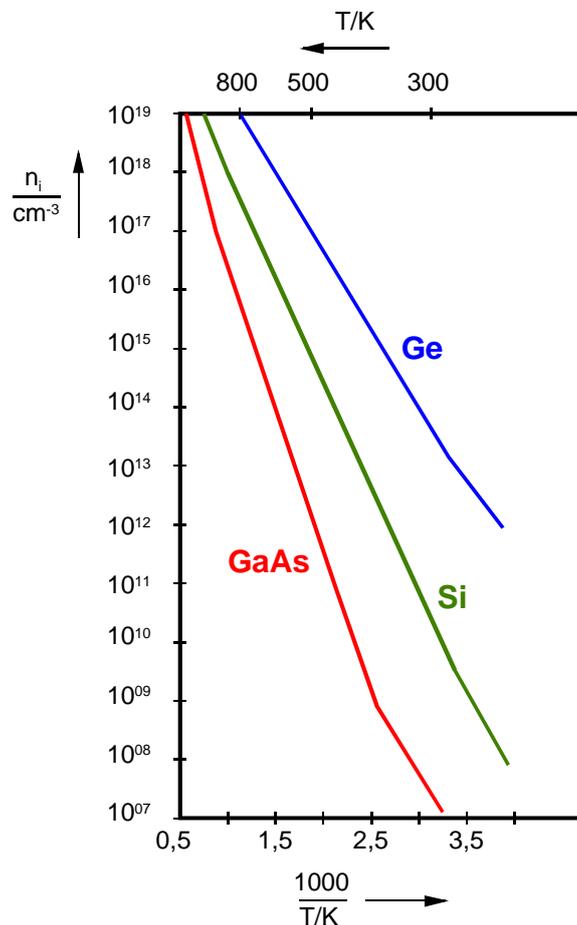


Abb. 0.9 Temperaturabhängigkeit der Eigenleitungskonzentration

Man erkennt, dass sowohl mit steigender Temperatur als auch mit fallender Energielücke W_g die Ladungsträgerkonzentration exponentiell ansteigt.

Beispielsweise für Raumtemperatur unterscheiden sich daher die n_i verschiedener Halbleiter erheblich (Tab. 0.1). Die Temperaturabhängigkeit dieser Halbleiter ist in Abb. 0.9 dargestellt.

Ein Halbleiter, für den die Fermi-Energie $W_{Fi} > W_L$ (bzw. $W_F > W_L$) ist, heißt **entarteter Halbleiter**. Man beachte, dass die Eigenschaft "**Entartung**" temperaturabhängig ist!

0.2.2.2.2 Störstellenhalbleiter

Der bisher behandelte Eigenhalbleiter stellt eine Idealisierung dar, die in reiner Form in der Natur nicht vorkommt und auch durch eine noch so ausgereifte und aufwendige Technologie nicht realisiert werden kann. Selbst wenn man Verunreinigungen durch Fremdatome völlig unterdrücken könnte (was nicht der Fall ist), ist aus thermodynamischen Gründen die exakte Einhaltung der periodischen Gitterstruktur nicht möglich. In den meisten technisch interessanten Anwendungen ist aber gerade ein gestörter Halbleiter erwünscht, der durch gezielte Zugabe von Fremdatomen (Dotierung) erzeugt werden kann. Beispielsweise können so **n-Typ** ($n \gg p$) oder **p-Typ** ($p \gg n$) Halbleiter hergestellt werden.

0.2.2.2.2.1 Dotierung von Halbleitern; Donatoren, Akzeptoren, tiefe Störstellen

Bei der **Dotierung** von Halbleitern mit Fremdatomen werden Gitteratome durch Fremdatome ersetzt. Am Beispiel des Siliziums sollen die zwei wichtigen zu unterscheidenden Fälle verdeutlicht werden. Silizium ist vierwertig, im idealen ungestörten Kristall ist jedes Siliziumatom durch sogenannte kovalente Bindungen, bestehend aus jeweils zwei Elektronen, an vier benachbarte Atome gebunden, die in den Ecken eines Tetraeders liegen. Schematisch kann man diesen Sachverhalt in einer ebenen Zeichnung nach Abb. 0.10 bzw. 0.11 darstellen, in der die Atome in einem Rechteckgitter sitzen. Die Anzahl der Bindungen bleibt korrekt, nicht aber ihre dreidimensionale Ausrichtung. Da die Elektronen die Bindung zwischen den Siliziumatomen bewirken, liegen sie hier fest und stehen für einen Leitungsvorgang nicht zur Verfügung. In den Abbildungen 0.10 bzw. 0.11 sind jedoch bereits die Dotierstoffatome berücksichtigt, und zwar Arsen (As) in Abb. 0.10 und Bor in Abb. 0.11. Zunächst sei der Fall "Arsen in Silizium" behandelt.

Donatoren

Arsen ist fünfwertig, d. h. es hat fünf Bindungs-(Valenz-)elektronen in seiner äußersten Schale und somit eines mehr als zur Absättigung der Bindungen im Siliziumgitter notwendig ist. Dieses Elektron ist daher nicht in eine Bindung eingebaut, somit weitgehend frei beweglich, so dass es zur elektrischen Leitfähigkeit beitragen kann: Das Arsenatom stellt im Siliziumkristall ein quasi freies Elektron zur Verfügung, es wirkt als Donator (Spender).

Im **Bänderdiagramm** wird das Donatorverhalten folgendermaßen beschrieben: Zu den Elektronen, die das Valenzband füllen (vier pro Atom), wird das überzählige des Arsenatoms hinzugefügt. Wenn dieses vollkommen frei wäre, würde es das niedrigste Niveau des Leitungsbandes besetzen. Tatsächlich aber wird es von der zusätzlichen positiven Ladung des Arsenatomrumpfes schwach angezogen. Dies erzeugt neue Zustände (Abb. 0.12a), die knapp unterhalb der Leitungsbandunterkante liegen, da bereits eine geringe Energie ausreicht, das Elektron vom Arsenatomrumpf zu lösen. Diese Energie heißt **Ionisierungsenergie des Donators** ΔW_D .

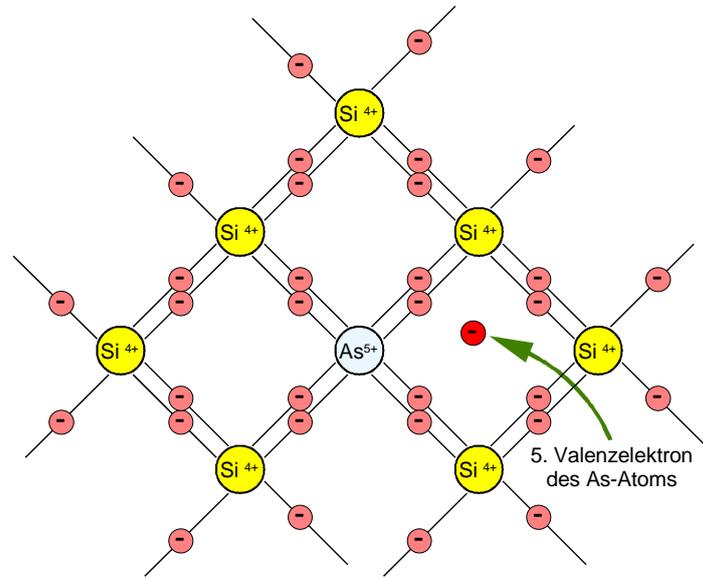


Abb. 0.10 Arsen in Silizium: Donatorverhalten

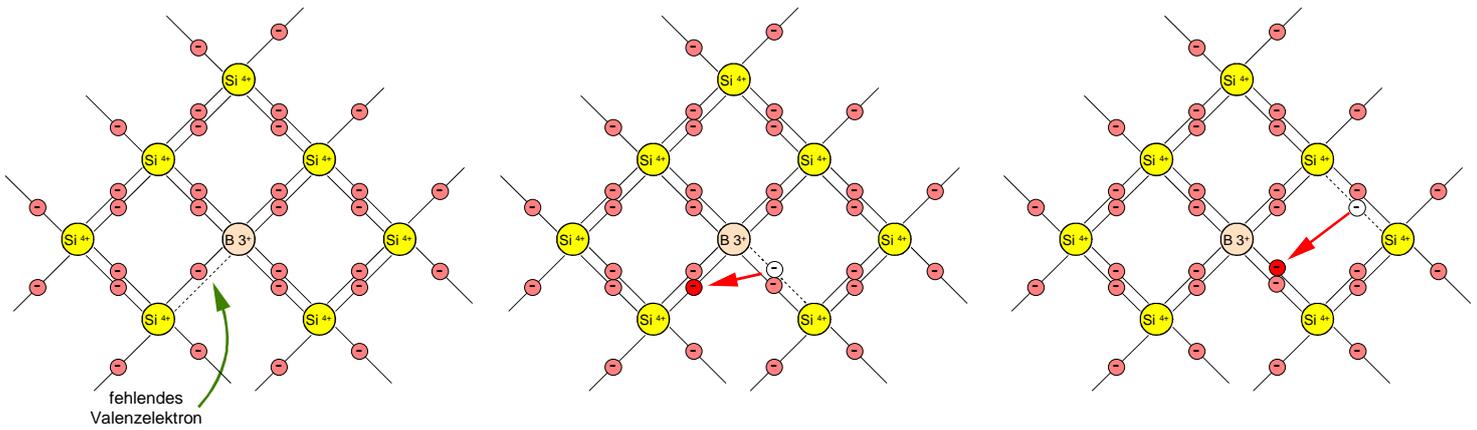


Abb. 0.11 Bor in Silizium: Akzeptorverhalten

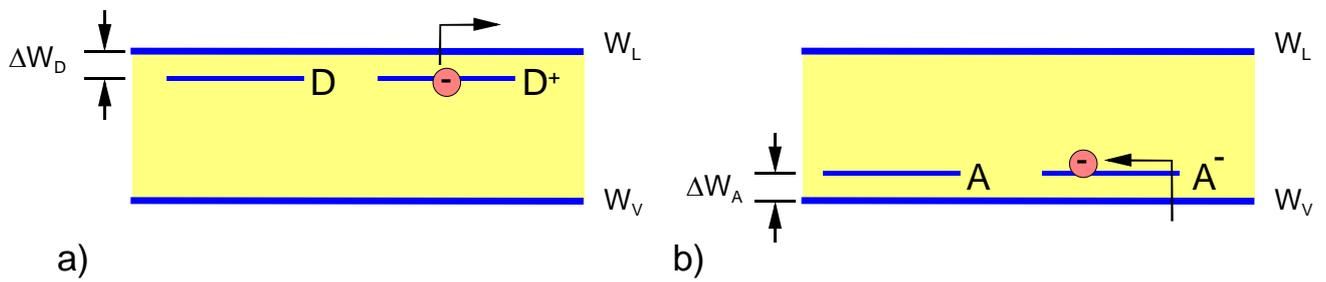


Abb. 0.12 a) Anordnung eines Donators im Bänderdiagramm
 b) Anordnung eines Akzeptors im Bänderdiagramm

Akzeptoren

Im Falle der Abb. 0.11 wird in das Siliziumgitter ein dreiwertiges Atom (Bor) eingebaut. Im Vergleich zum Silizium fehlt ein Valenzelektron, so dass eine Bindung nur mit einem Elektron statt mit zwei besetzt ist. In diese Lücke, die einem Loch im Valenzband entspricht, können besonders leicht Elektronen eingefangen werden, die beispielsweise aus Nachbarbindungen stammen. Das dreiwertige Atom nimmt also ein zusätzliches Elektron auf, es wirkt als Akzeptor. Der Leitungsmechanismus, getragen von Löchern im Valenzband, ist in Abb. 0.11 angedeutet. Dieses Verhalten ist in Abb. 0.12b im Bänderdiagramm gezeigt. Die geringe Energie, die benötigt wird, ein Elektron aus einer benachbarten Bindung, d. h. aus dem Valenzband, in die "Bindungslücke" zu heben, heißt **Ionisierungsenergie des Akzeptors** ΔW_A .

Je nach Dotierung mit Donatoren oder Akzeptoren werden mehr Elektronen bzw. Löcher als im intrinsischen Halbleiter vorliegen, da die Ionisierungsenergie der Störstellenatome sehr viel geringer ist als die Bandlückenenergie.

Im Fall der

Donatordotierung spricht man von **n-Typ Halbleitern**, die eine erhöhte Elektronenkonzentration besitzen, und im Falle der

Akzeptordotierung von **p-Typ Halbleitern** mit erhöhter Löcherkonzentration.

Nach wie vor gilt jedoch das **Massenwirkungsgesetz**

$$n \cdot p = n_i^2$$

Damit sinkt im n-Typ Halbleiter die Löcherkonzentration tief unter die des intrinsischen Halbleiters und im p-Typ Halbleiter die Elektronenkonzentration weit unter die des intrinsischen Halbleiters.

Die Ionisierungsenergien ΔW_D und ΔW_A können in Anlehnung an die Ionisierungsenergie des H-Atoms abgeschätzt werden (vgl. Gl.(0.5))

$$W_{ion}(H - Atom) = \frac{q^2}{8\pi \varepsilon_o r_o} = 13,6eV$$

mit

$$r_o = \frac{4\pi \varepsilon_o \hbar^2}{m q^2}$$

Im Falle des dotierten Halbleiters hat man statt der Elektronenmasse die effektive Masse der Elektronen bzw. Löcher sowie statt ε_o die Größe zu verwenden, entsprechend der Permittivität des Halbleiters:

$$\Delta W_D = \frac{m_n^* q^4}{2(4\pi \varepsilon_o \varepsilon_r \hbar)^2} \quad \text{und} \quad \Delta W_A = \frac{m_p^* q^4}{2(4\pi \varepsilon_o \varepsilon_r \hbar)^2} \quad (0.23)$$

Für die verschiedenen Halbleiter sind die Ionisierungsenergien in Tab. 0.2 angegeben. Ebenfalls angegeben sind die "Bohr'schen Radien" für das H-Atom ähnliche System "Ladungsträger-Störstellenatom". Die Ausdehnung erstreckt sich über viele hundert Gitterzellen.

Für **Verbindungshalbleiter** wie GaAs sind die Verhältnisse meist komplizierter als für Elementhalbleiter (z. B. Si): Ga ist dreiwertig (3 Valenzelektronen) und As ist fünfwertig (5 Valenzelektronen), so dass zusammen wieder, wie im Fall des Si, acht Elektronen für die Bindung mit den vier nächsten Nachbarelektronen zur Verfügung stehen.

Halbleiter	ΔW_D /meV	r_o^* / r_o	ΔW_A /meV	r_o^* / r_o
Si	94	12	47	24
Ge	29	29	20	43
GaAs	5,3	196	36	29
InAs	1,9	568	36	30
InP	8	159	57	22
InGaAs	3,7	284	36	30

Tab. 0.2 Donator- und Akzeptorionisierungsenergien verschiedener Halbleiter

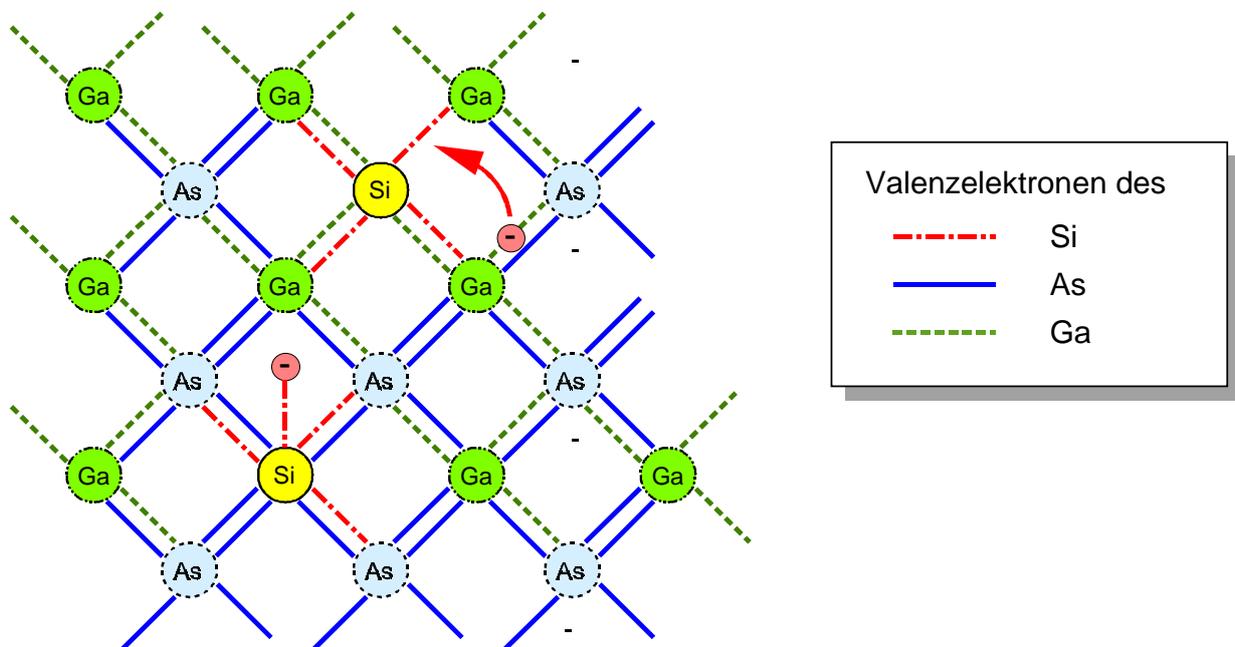


Abb. 0.13 Amphoteres Verhalten von Si in GaAs

Jedes Ga-Atom ist von vier nächsten As-Atomen, jedes As-Atom von vier nächsten Ga-Atomen umgeben. Jetzt kommt es aber darauf an, ob beispielsweise ein vierwertiges Fremdatom (z. B. Si) auf einem Ga- oder auf einem As-Platz eingebaut wird: Auf dem Ga-Platz liefert es ein überzähliges Elektron und wirkt als Donator, auf einem As-Platz fehlt ein Elektron und es wirkt daher als Akzeptor (Abb. 0.13). Dieses Verhalten wird als **amphot** bezeichnet.

0.2.2.2.3 Ladungsträgerstatistik im Störstellenhalbleiter

Die Ladungsträgerkonzentrationen (n bzw. p) im intrinsischen Halbleiter sind, wie gezeigt, durch die drei Größen Zustandsdichte $g(W)$, Fermi-Dirac-Verteilungsfunktion $f(W)$ sowie Fermi-Energie W_F bestimmt. Dies gilt auch für den Störstellenhalbleiter (extrinsischer Halbleiter), allerdings sind die Unterschiede zum Eigenhalbleiter zu beachten:

Es existieren Zustände in der Energielücke und es gilt i.A.:

$$n \neq p.$$

Nach wie vor gelten aber die Gl. (2.16) bzw. (2.19) für die Elektronen- bzw. Löcherkonzentrationen:

$$n = N_L \cdot e^{-(W_L - W_F)/kT} \quad (0.24)$$

$$p = N_V \cdot e^{-(W_F - W_V)/kT} \quad (0.25)$$

Nur unterscheidet sich jetzt die Fermi-Energie von der im Eigenhalbleiter:

Im **n-Typ** Halbleiter liegt die Fermi-Energie näher an der Leitungsbandunterkante und

im **p-Typ** Halbleiter näher an der Valenzbandoberkante.

Mit den Gl. (2.24) und (2.25) ist auch die Gültigkeit des **Massenwirkungsgesetzes** gegeben:

$$n \cdot p = n_i^2$$

Da jedes Donator- und Akzeptoratom für sich neutral ist, gilt auch die **Ladungsneutralität** für dotierte Halbleiter, auch für eine simultane p- und n-Dotierung, solange sich der Halbleiter im **thermodynamischen Gleichgewicht** befindet:

$$n + N_A^- = p + N_D^+ \quad (0.26)$$

Temperaturabhängigkeit der Konzentration

Eine analytische Auswertung lässt sich nur für spezielle Voraussetzungen durchführen. Dies soll hier beispielhaft für einen n-dotierten Halbleiter getan werden.

Dazu wird angenommen, dass

1. nur ein Donortyp vorliegt, charakterisiert durch die Parameter ΔW_D und N_D
2. zunächst die Eigenleitungskonzentration n_i vernachlässigt wird.

1. Schwache Dotierung und hohe Temperatur

Genauer soll gelten:

- a) $N_D \ll N_L$ und
- b) $kT \gg \Delta W_D$

Dann folgt:

$$n = N_D \quad (0.27)$$

Die Elektronenkonzentration ist identisch mit der Donatorkonzentration, denn alle Donatoren haben ihr Elektron abgegeben und sind somit ionisiert.

Man sagt, die Störstellen sind erschöpft (**Erschöpfungsbereich**). Dies ist auch einleuchtend, weil wegen a) $N_D \ll N_L$ für die Donatorelektronen genügend freie Plätze im Leitungsband vorhanden sind und wegen b) $kT \gg \Delta W_D$ die thermische Energie sehr viel höher ist, als zur Ionisierung der Donatoren nötig ist.

2. Tiefe Temperatur

Das heißt, es gilt $kT \ll \Delta W_D$. Dafür ergibt sich:

$$n = \sqrt{N_D N_L} e^{-\Delta W_D / 2kT} \quad (0.28)$$

Mit steigender Temperatur nimmt die Elektronenkonzentration also stark zu. Es befindet sich eine "Elektronen-Reserve" auf den noch nicht ionisierten Donatoren, daher wird dieser Temperaturbereich **Reservebereich** genannt. Bei der Temperatur T_{\min} (Abb. 0.14) geht dieser in den **Erschöpfungsbereich** mit $n = N_D$ über.

Wird die Temperatur weiter bis T_{\max} gesteigert, werden Elektron-Loch-Paare durch Band-Band-Übergänge über die Bandlücke hinweg erzeugt, wie bereits für den Fall des Eigenhalbleiters besprochen. In diesem Temperaturbereich überwiegt die Eigenleitung, dementsprechend heißt er **Eigenleitungsbereich**. Die Elektronenkonzentration gehorcht hier der Gl. (0.22):

$$n_i = \sqrt{N_L N_V} e^{-W_g / 2kT}$$

Die Abb. 0.14 gibt schematisch die Temperaturabhängigkeit eines n-dotierten Halbleiters von niedrigsten bis zu höchsten Temperaturen wieder. Analoge Überlegungen lassen sich für einen p-dotierten Halbleiter durchführen.

Bauelemente aus dotierten Halbleitern werden im Erschöpfungsbereich betrieben, also zwischen T_{\min} und T_{\max} .

Denn Elemente, die einen pn-Übergang enthalten, verlieren oberhalb von T_{\max} ihre Funktion, weil ein Eigenhalbleiter per Definition weder n- noch p-leitend, sondern gleich viele Elektronen wie Löcher besitzt. Für Temperaturen unterhalb T_{\min} "frieren die Ladungsträger aus", die Ladungsträgerkonzentration hat keinen festgelegten Wert mehr.

Bezieht man sich auf eine n-Dotierung von $N_D = 10^{17} \text{ cm}^{-3}$, dann ergibt sich für Silizium ein Wert für $T_{\max} \approx 510^\circ \text{ C}$ für GaAs wegen des höheren Bandabstandes ein deutlich höherer Wert von $T_{\max} \approx 810^\circ \text{ C}$. Dabei ist zu beachten, dass diese Temperaturen z. B. bei Bipolartransistoren am Ort des pn-Übergangs nicht überschritten werden dürfen, so dass die Gehäusetemperatur noch deutlich niedriger liegen muss. Ferner wird klar, dass für elektronische Bauelemente, die in einer Hochtemperaturumgebung (Automobil, Energietechniksteuerungen usw.) arbeiten müssen, vorteilhaft Halbleiter mit größerer Bandlücke verwendet werden.

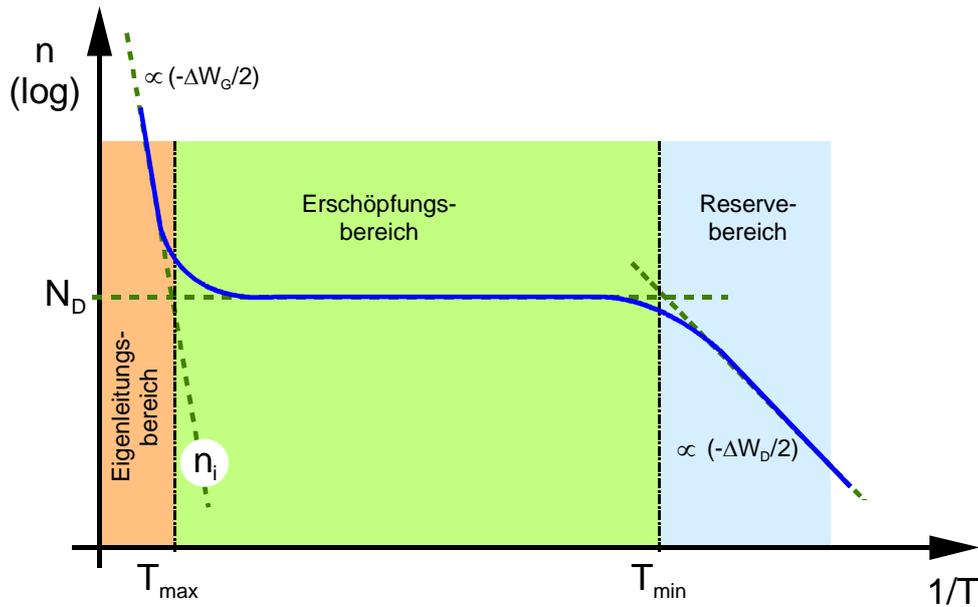


Abb. 0.14 Die temperaturabhängige Elektronenkonzentration in einem n-dotierten Halbleiter

Wenn man gemäss der oben durchgeführten Überlegungen im n-Halbleiter die Elektronenkonzentration berechnet hat, so ist die Bestimmung der dann vorliegenden Löcherkonzentration sehr einfach durch Anwendung des Massenwirkungsgesetzes durchzuführen. Es gilt:

$$p = \frac{n_i^2}{n} = \frac{n_i^2}{N_D} \quad \text{Erschöpfungsbereich; } p \ll n_i \ll n \quad (0.29)$$

Wegen der um viele Größenordnungen unterschiedlichen Konzentrationen nennt man die Elektronen im n-Halbleiter **Majoritätsträger** und die Löcher die **Minoritätsträger**.

0.2.3 Ladungstransport in Festkörpern

Für die elektrische Leitfähigkeit von und die Stromdichte in Festkörpern ist die **Ladungsträgerkonzentration** von großer Bedeutung. Darüber hinaus spielt aber auch die **Geschwindigkeit** eine Rolle, mit der sich die Ladungsträger unter Einwirkung einer äußeren Kraft bewegen. Sieht man im Festkörper von einem Ionentransport ab, **so wird die Leitfähigkeit von Elektronen und Löchern getragen.**

Allgemein gilt dann für die **Stromdichte**

$$\vec{J} = -q n \vec{v}_n + q p \vec{v}_p \quad n, p: \quad \text{Ladungsträgerkonzentration} \quad (0.30)$$

$q, -q:$ Elementarladung für Elektronen bzw. Löcher
 $\vec{v}_n, \vec{v}_p:$ Elektronen- bzw. Löchergeschwindigkeit

Wird die äußere Kraft durch ein elektrisches Feld hervorgerufen, so ist wegen des Vorzeichenunterschiedes der Ladung \vec{v}_n antiparallel und \vec{v}_p parallel zum elektrischen Feld gerichtet.

Somit addieren sich die beiden Stromdichtebeiträge.

Elektronen und Löcher bewegen sich aber bei Raumtemperatur auch ohne eine von außen wirkende Kraft mit erheblichen Geschwindigkeiten, da sie thermische Energie besitzen. Die mittlere thermische Energie jedes Ladungsträgers beträgt:

$$W_{th} = \frac{3}{2} kT = \frac{1}{2} m^* v_{th}^2 = \frac{1}{2} m^* \left(\frac{\hbar k}{m^*} \right)^2$$

oder, umgestellt nach der thermischen Geschwindigkeit:

$$v_{th} = \sqrt{\frac{3kT}{m^*}} \quad (0.31)$$

Für gängige Halbleiter beträgt diese Geschwindigkeit bei 300K immerhin etwa $v_{th} \approx 10^7 \text{ cm/s}$. Für unterschiedliche Elektronen sind aber auch die Geschwindigkeitsrichtungen völlig unkorreliert und unterschiedlich.

0.2.3.1 Die Ladungsträgerbeweglichkeit und die elektrische Leitfähigkeit

Die Kraft, die ein geladenes Teilchen im elektrischen Feld \vec{E} erfährt, ist

$$\vec{F}_n = -q \cdot \vec{E} \quad (0.32)$$

Dabei wird ein Elektron mit der Ladung $-q$ betrachtet. Die Folge ist eine Beschleunigung des Elektrons nach dem mechanischen Kraftgesetz:

$$\vec{F} = m_n^* \cdot \vec{a} \quad \vec{a}: \quad \text{Beschleunigung} \quad (0.33)$$

Den Proportionalitätsfaktor (ohne Vorzeichen) zwischen der resultierenden Geschwindigkeit \vec{v}_D und der diese verursachenden Feldstärke \vec{E} nennt man sehr anschaulich **Beweglichkeit** der Elektronen μ_n im Festkörper: je beweglicher das Elektron ist, desto höher wird bei vorgegebener Feldstärke \vec{E} die Driftgeschwindigkeit v_D sein.

Mit (s. Vorlesung "Festkörperelektronik")

$$\begin{aligned} \vec{v}_{Dn} &= -\mu_n \cdot \vec{E} \\ \vec{v}_{Dp} &= \mu_p \cdot \vec{E} \end{aligned} \quad (0.34)$$

gilt:

$$\vec{J} = q(n \cdot \mu_n + p \cdot \mu_p) \cdot \vec{E} \quad (0.35)$$

Damit ist die allgemeingültige Definition der Ladungsträgerbeweglichkeit gegeben

$$\mu = \frac{d|\vec{v}_D|}{d|\vec{E}|} \quad (0.36)$$

Dieser Ausdruck wird genauer **Driftbeweglichkeit** genannt.

Der Ausdruck (0.33) liefert ferner die elektrische **Leitfähigkeit** als Proportionalitätsfaktor zwischen Stromdichte und elektrischer Feldstärke

$$\vec{J} = \kappa \cdot \vec{E} = (\kappa_n + \kappa_p) \cdot \vec{E} \quad (0.37)$$

Diese Gleichung gibt erneut die Aufteilung in Elektronen- und Löcheranteil wieder. Der Kehrwert von κ wird als **spezifischer Widerstand** ρ bezeichnet:

$$\rho = \frac{1}{\kappa} \quad (0.38)$$

Temperaturabhängigkeit

Mit steigender Temperatur schwingen die Gitteratome immer stärker um ihre Ruhelage und stören somit die strenge Periodizität der Gitteranordnung. Diese Gitterschwingungen werden Phononen genannt, die man ähnlich den Photonen und Elektronen wie Teilchen behandeln kann. Die Störung der Elektronen- (Löcher-) bewegung (Phononenstreuung) kann man in diesem Bild als Zusammenstoß mit Phononen interpretieren, der umso häufiger auftritt, je höher die Temperatur ist.

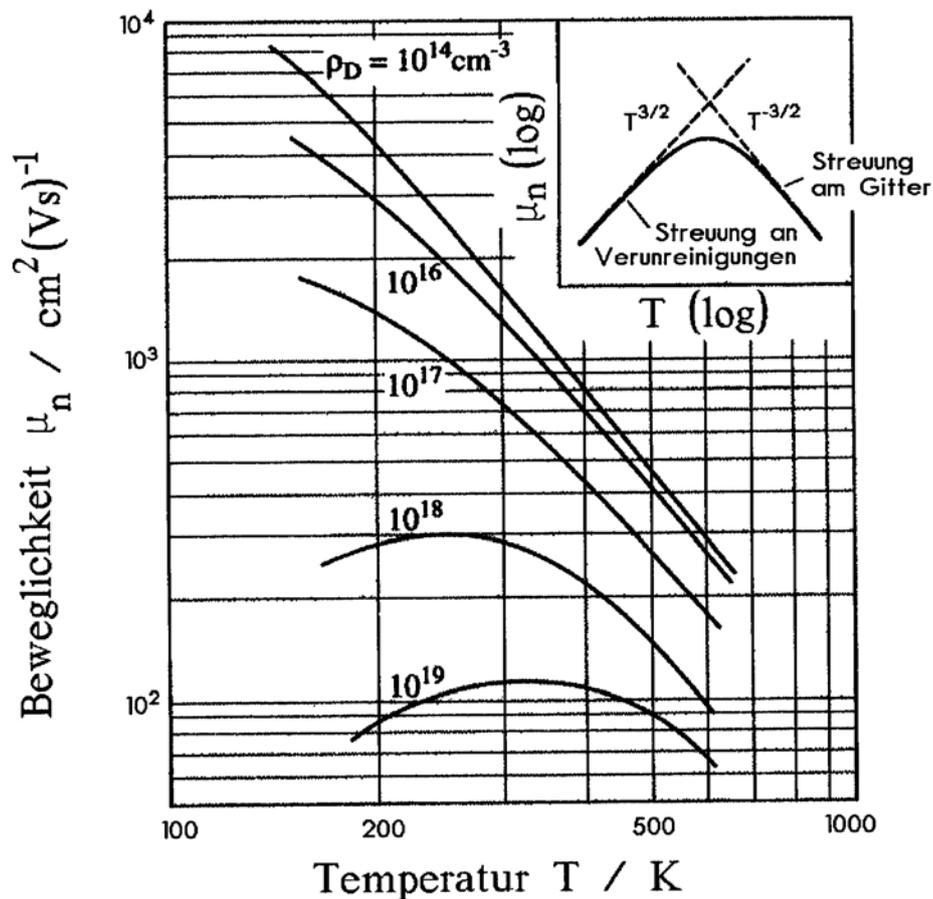


Abb. 0.15 Abhängigkeit der Elektronendriftbeweglichkeit von der Temperatur und der Dotierung (n-Silizium) nach H.Schaumburg: "Halbleiter", B.G.Teubner-Verlag, 1991, S.145

Die Beweglichkeit und die Driftgeschwindigkeit wird also mit steigender Temperatur abnehmen.

Man findet:

$$\mu_{phon}(T) \propto T^{-3/2} \tag{0.39}$$

mit Schwankungen des Exponenten um den Wert $-(3/2)$. Abb. 0.15 zeigt diesen Verlauf für sehr schwache Dotierungen.

Dotierungsabhängigkeit

Oben wurde bereits dargelegt, dass die Störung der Periodizität des Gitters z.B. durch eine Dotierung zu häufigeren Zusammenstößen und somit zur Beeinträchtigung der Beweglichkeit führt.

Eingebaute Störungen, insbesondere in Form von geladenen Donatoren oder Akzeptoren, erniedrigen dann erst recht die Beweglichkeit. Durch die Ladung ist ihr Wirkungsbereich besonders groß.

Mit steigender Störstellenkonzentration wird daher die Beweglichkeit abnehmen. Die Streuung der Elektronen an ionisierten (geladenen) Störstellen (**Coulombstreuung**) hat ebenfalls eine charakteristische Temperaturabhängigkeit:

$$\mu_{ion} \propto T^{3/2} \tag{0.40}$$

Die Beweglichkeit steigt also mit zunehmender Temperatur an (Abb. 0.15, höhere Störstellenkonzentrationen). Dies ist leicht einzusehen, da die thermische Ladungsträrgeschwindigkeit ansteigt und damit die Wechselwirkung mit den Störstellen abnimmt. Je nach Dotierung durchläuft daher die Beweglichkeit mit der Temperatur ein Maximum. Entscheidend ist hier die Summe von Donator- und Akzeptorkonzentrationen, da beide Anteile die Beweglichkeit reduzieren. Für die Ladungsträgerkonzentration ist dagegen die Differenz von Donator- und Akzeptorkonzentration entscheidend, z.B.

$$n = N_D^+ - N_A^-, \text{ wenn } N_D^+ > N_A^-$$

Abhängigkeit von der elektrischen Feldstärke

Bei kleinen Feldstärken steigt die Driftgeschwindigkeit linear mit der Feldstärke gemäß Gl. (0.34).

Übertragen auf Stromdichte und Spannung ist dies der Bereich, in dem das **Ohm'sche Gesetz** gilt:

$$J = \kappa \cdot E \quad \text{bzw.} \quad I = \frac{U}{R} \tag{Ohm'sches Gesetz} \tag{0.41}$$

mit

$E = \frac{U}{L}$	$L:$	Kontaktabstand
$R = \frac{1}{\kappa} \cdot \frac{L}{A} = \rho \cdot \frac{L}{A}$	$A:$	Querschnitt

und

$$I = J \cdot A$$

Mit steigender Feldstärke und somit steigender Geschwindigkeit nimmt die kinetische Energie der Elektronen zu. Dies führt dazu, dass die Ladungsträgerbeweglichkeit fällt.

Ferner werden die Zeitabstände zwischen den Zusammenstößen bei steigender Geschwindigkeit (aufgrund steigender Feldstärke) immer kürzer, so dass dies schließlich bei Halbleitern wie Silizium, zu einer Sättigung der Geschwindigkeit mit der Feldstärke führt (vgl. Abb. 0.16).

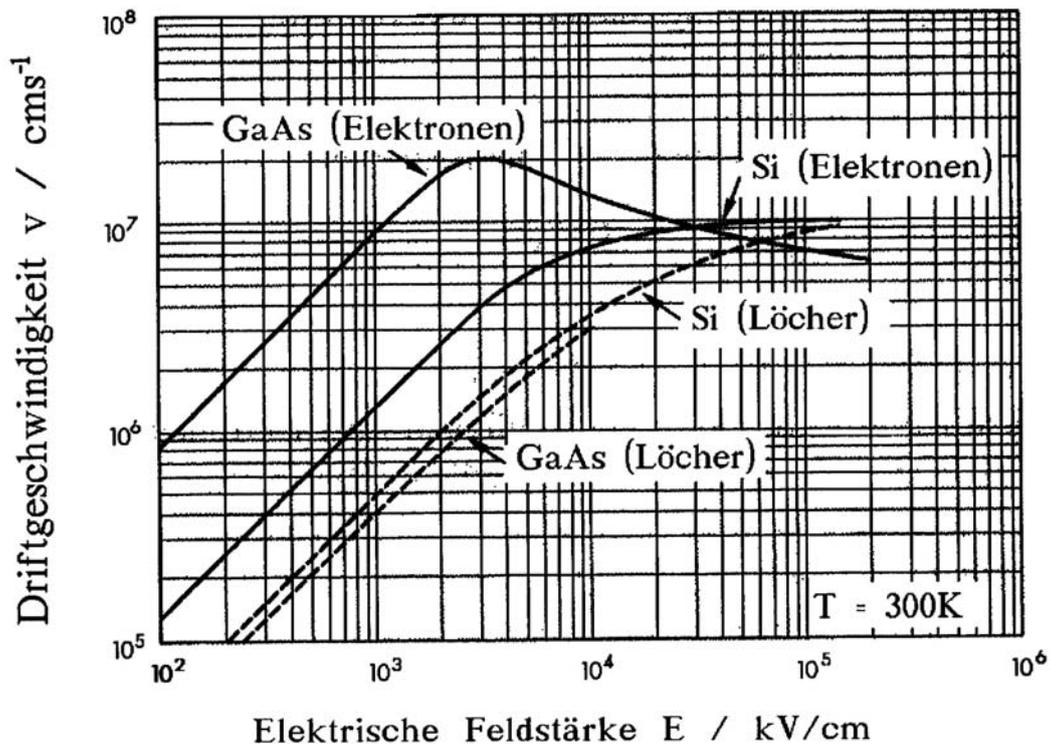


Abb. 0.16 Feldstärkeabhängigkeit von Ladungsträgern in unterschiedlichen Halbleitern [17]

Eine Geschwindigkeitssättigung ist aber gleichbedeutend mit einer verschwindenden Beweglichkeit (Gl. (0.36)).

Werkstoffe wie **GaAs** weisen die Besonderheit auf, dass die Elektronendriftgeschwindigkeit nach Durchlaufen eines Maximums sogar wieder fällt, obwohl die Feldstärke ansteigt. Dieser Sachverhalt ist gleichbedeutend mit einer negativen Beweglichkeit und damit auch mit einem negativem Widerstand. Bauelemente aus solchen Materialien (z.B. Gunn-Elemente) können zur Verstärkung und Schwingungserzeugung benutzt werden.

0.2.3.2 Ströme infolge eines elektrischen Feldes: Feldströme

Grundsätzlich wurde bereits der Zusammenhang zwischen Stromdichte und elektrischer Feldstärke (Gl. (0.37)) angegeben:

$$\vec{J}_{Feld} = \kappa \cdot \vec{E} = (\kappa_n + \kappa_p) \cdot \vec{E}$$

Dabei wurde angenommen, dass die Leitfähigkeit κ , und damit auch die Ladungsträgerkonzentrationen und die Beweglichkeiten, **ortsunabhängig** waren. Bei räumlich konstanter Feldstärke ist dann auch die Stromdichte **ortsunabhängig**.

Im folgenden soll einen Zusammenhang der für den Feldstrom relevanten Größen mit dem Bänderdiagramm hergestellt werden, das sich bei der Beschreibung der elektronischen Eigenschaften von Festkörpern bereits ausgezeichnet bewährt hat.

Für den Ladungsträgertransport ist die Geschwindigkeit der Elektronen und Löcher und damit deren kinetische Energie entscheidend, die für Elektronen ihren Bezugspunkt an der Leitungsbandunterkante hat.

Andererseits besteht zwischen der Energie W einer Ladung in einem elektrischen Feld \vec{E} , diesem Feld und dem dazugehörigen elektrostatischen Potenzial φ , die Beziehungen

$$W = -q \varphi + const. \quad (0.42)$$

$$grad \varphi = -\vec{E} \quad (0.43)$$

Gemäss der Wahl des Nullpunktes der kinetischen Energie wird aus diesen Gleichungen

$$W_L = -q \varphi \quad (0.44)$$

$$grad W_L = -q grad \varphi = q \vec{E} \quad (0.45)$$

Wegen $W_g = const.$ für einen bestimmten Halbleiter gilt auch

$$grad W_L = grad (W_V + W_g) = grad W_V \quad (0.46)$$

Das bedeutet, dass bei Anlegen eines E-Feldes die Bandkanten W_L und W_V nicht mehr waagrecht, sondern geneigt, und somit W_L und W_V ortsabhängig sind. Entsprechend der Voraussetzung sollen jedoch die Ladungsträgerkonzentrationen ortsunabhängig sein, d.h. nach Gl. (0.24) und (0.25) muss $(W_L - W_F)$ bzw. $(W_F - W_V)$ konstant sein. Somit ergibt sich die Abb. 0.17.

Der Potenzialunterschied zwischen linkem und rechtem Ende des betrachteten Halbleiterkörpers ist identisch mit der angelegten Spannung, die das E-Feld verursacht.

Es gilt also (eindimensional):

$$-U = \varphi(x=L) - \varphi(x=0) = \int_{\varphi(x=0)}^{\varphi(x=L)} d\varphi = -\frac{1}{q} \int_{W_L(x=0)}^{W_L(x=L)} dW_L \quad (0.47)$$

Die Neigung der Bandkanten ist offensichtlich gleichbedeutend mit einem Potenzialunterschied und dem Vorliegen eines elektrischen Feldes. Dieses Ergebnis ist wichtig für die Verhältnisse beim pn-Übergang, aber auch beim Metall-Halbleiterkontakt und an der Halbleiteroberfläche.

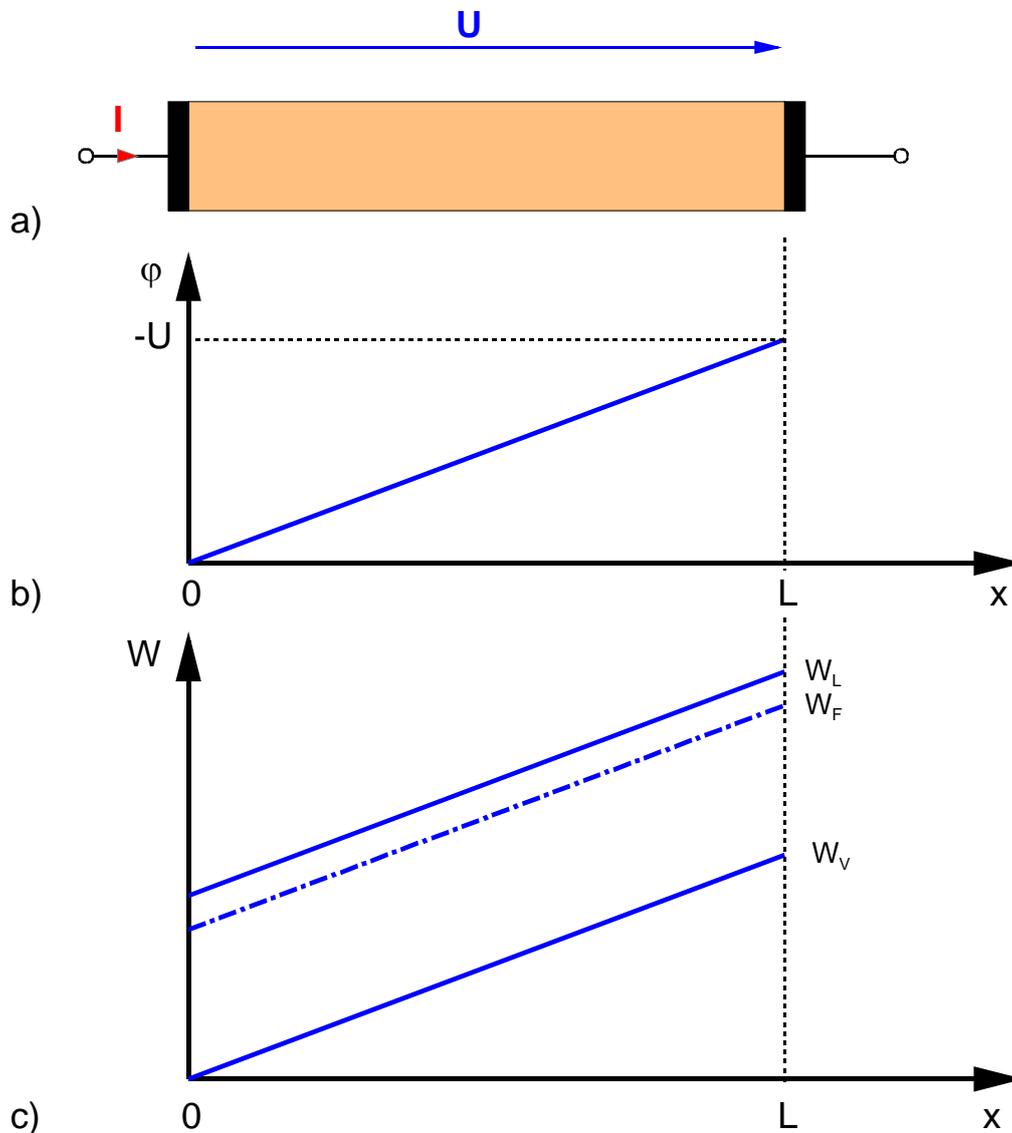


Abb. 0.17 Verlauf der Bandkanten und des Potentials für einen Halbleiter im E-Feld

0.2.3.3 Ströme infolge eines Konzentrationsgradienten: Diffusionsströme

Von Gasen weiß man, dass sie die Tendenz haben, den ganzen verfügbaren Raum gleichmäßig auszufüllen. Anders ausgedrückt besagt dies, dass Konzentrationsunterschiede der frei beweglichen Gasteilchen ausgeglichen werden. Dieses Verhalten gilt ganz allgemein für bewegliche Teilchen, auch für Elektronen und Löcher im Festkörper. Der Mechanismus des Ausgleichs von Konzentrationsunterschieden durch die Bewegung von Teilchen heißt Diffusion. Für die Teilchenstromdichte gilt:

$$\vec{s} = -D \text{grad } n(x) \quad (0.48)$$

Sie ist umso größer, je steiler das Konzentrationsprofil ist, und ist in Richtung kleinerer Konzentrationen gerichtet. D wird **Diffusionskoeffizient** genannt.

Er ist von der Temperatur abhängig, da die Energiequelle für die Diffusionsbewegung die thermische Energie ist, die für die thermische Geschwindigkeit verantwortlich ist (vgl. Gl. (0.31)). Nur für kleine Konzentrationen ist er näherungsweise konzentrationsunabhängig.

Sind die Teilchen elektrisch geladen, so ist mit dem Teilchendiffusionsstrom ein elektrischer Strom verbunden. Für Elektronen bzw. Löcher findet man

$$\vec{J}_{n,Diff} = -q \cdot \vec{s}_n = q D_n \text{ grad } n(x) \quad (0.49)$$

$$\vec{J}_{p,Diff} = q \cdot \vec{s}_p = -q D_p \text{ grad } p(x) \quad (0.50)$$

Ohne eine detaillierte Behandlung kann bereits an dieser Stelle die wichtige **Einstein Beziehung** zwischen Beweglichkeit und Diffusionskonstante hergeleitet werden, da sie auf ganz allgemein gültigen Prinzipien beruht:

Zunächst gilt nach dem thermodynamischen "*Prinzip des detaillierten Gleichgewichtes*", dass im Gleichgewicht zu jedem einzelnen Mikroprozess ein Gegenprozess existiert. Im Fall des pn-Übergangs ist der zum Diffusionsprozess der Elektronen existierende Gegenprozess der Elektronenfeldstrom. Das treibende Feld entsteht aus dem Diffusionsladungsträgertransport, der Ladungen verschiebt und somit einen Potenzialunterschied aufbaut. Wenn von außen keine Spannung anliegt, herrscht Gleichgewicht, und außen kann auch kein Strom fließen; sonst hätte man ein *Perpetuum Mobile* gebaut. Es gilt also:

$$\begin{aligned} \vec{J}_n = 0 &= \vec{J}_{n,Diff} + \vec{J}_{n,Feld} \\ &= q \left(\mu_n \vec{E}(x) + D_n \text{ grad } n(x) \right) \end{aligned} \quad (0.51)$$

Mit

$$n(x) = N_L \cdot e^{-(W_L(x)-W_F)/kT}$$

gewinnt man aus (0.51)

$$0 = q n E(x) \left(\mu_n - D_n \frac{q}{kT} \right)$$

Ohne Herleitung wurde dabei benutzt, dass die Fermi-Energie W_F im Gleichgewicht konstant im ganzen betrachteten System, also ortsunabhängig ist. Feldstärke und Elektronenkonzentration sind aber ortsabhängig anzusetzen.

Diese Gleichung hat als allgemeine Lösung die **Einsteinsche Beziehung**:

$$\mu_n = D_n \frac{q}{kT} \quad \text{bzw. für die Löcher} \quad \mu_p = D_p \frac{q}{kT} \quad (0.52)$$

0.3 Halbleiter-Festkörper-Grenzflächen

Für reale Halbleitergebilde spielen insbesondere die **Grenzflächen** eine entscheidende Rolle. Beispielsweise werden die Eigenschaften von speziellen Halbleitern erst durch metallische Kontakte, d.h. über **Metall-Halbleiter-Grenzflächen** verfügbar. Die Computertechnologie basiert zum großen Teil auf Eigenschaften der Halbleiter-Isolator-Grenzflächen, und bipolare Bauelemente wie die klassischen Bipolar-Transistoren, aber auch Komponenten der Höchstfrequenztechnik und der optischen Nachrichtentechnik beziehen ihre funktionsbestimmenden Merkmale aus **Halbleiter-Halbleiter-Grenzflächen**, an denen zunehmend nicht allein der Leitfähigkeitstyp (p bzw. n), sondern auch die Halbleitersubstanz eine Veränderung erfährt (z. B. **Heterostruktur-Transistoren**, **Heterostruktur-Laser** und **-Detektoren**).

0.3.1 Das Metall-Isolator-Halbleiter-System, MIS- und MOS-Kondensatoren

Dieses System ist das in der Mikroelektronik mit Abstand am weitesten verbreitete System überhaupt. In Form der **CMOS** und insbesondere der Speichertechnologie hat es die moderne Rechnertechnik erst möglich gemacht.

Das Metall-Isolator-Halbleiter (MIS)-System enthält bereits zwei Festkörpergrenzflächen, von denen die Isolator-Halbleitergrenzfläche die wichtigste ist. Das Metall dient lediglich der isolatorseitigen Kontaktierung der Anordnung. Für den Zweck des chemischen und mechanischen Schutzes der Halbleiteroberfläche (Passivierung) fehlt es sogar. Die technisch wichtigste Kombination ist die des Silizium/SiO₂ mit einer SiO₂-Schichtdicke im Bereich von nur wenigen Nanometern. Es wird also das natürliche Oxid des Halbleitermaterials (MOS: Metall-Oxid-Semiconductor) als Isolator verwendet, und das hat sowohl technologische als auch Qualitätsvorteile gegenüber anderen MIS-Kombinationen.

0.3.1.1 Bändermodell und Ladungsverteilung

Grundsätzlich können an den Grenzfläche zwischen Isolator und Halbleiter sogenannte Grenzflächenzustände auftreten. Zusätzlich können jetzt noch Zustände und Ladungen im Isolator selbst vorliegen. Ihre Wirkung überlagert sich additiv mit den Grenzflächenzuständen .

0.3.1.1.1 Die ideale MIS-Struktur

Zunächst werde die ideale MIS-Struktur ohne Grenzflächen- und Isolatorladungen betrachtet, die folgendermaßen charakterisiert ist:

1. Die Austrittsarbeiten für Metall und Halbleiter sind identisch
2. Grenzflächen- sowie Isolatorladungen liegen nicht vor.
3. Der Isolator ist ideal und lässt somit keinen Stromfluss zu.

Legt man keine Spannung zwischen Metall und Halbleiter an, so liegt **thermodynamisches Gleichgewicht** vor und die Fermi-Energie W_F verläuft in der gesamten Anordnung waagrecht. Sie dient als Referenzlinie für die weiteren Betrachtungen. Für diese ideale MIS-Struktur verlaufen die Bandkanten des Halbleiters und des Isolators horizontal, d.h. W_L , W_V und $W_{L,IS} = const.$ Abb 0.18 zeigt dieses Verhalten für den Fall eines p-dotierten Halbleiters (p-HL).

Da die Austrittsarbeiten aus thermodynamischen Gründen auf W_F bezogen werden, sind sie für n- und p-Typ Halbleiter unterschiedlich. Daher müssen in beiden Fällen unterschiedliche Metalle zur Realisierung der idealen Struktur verwendet werden.

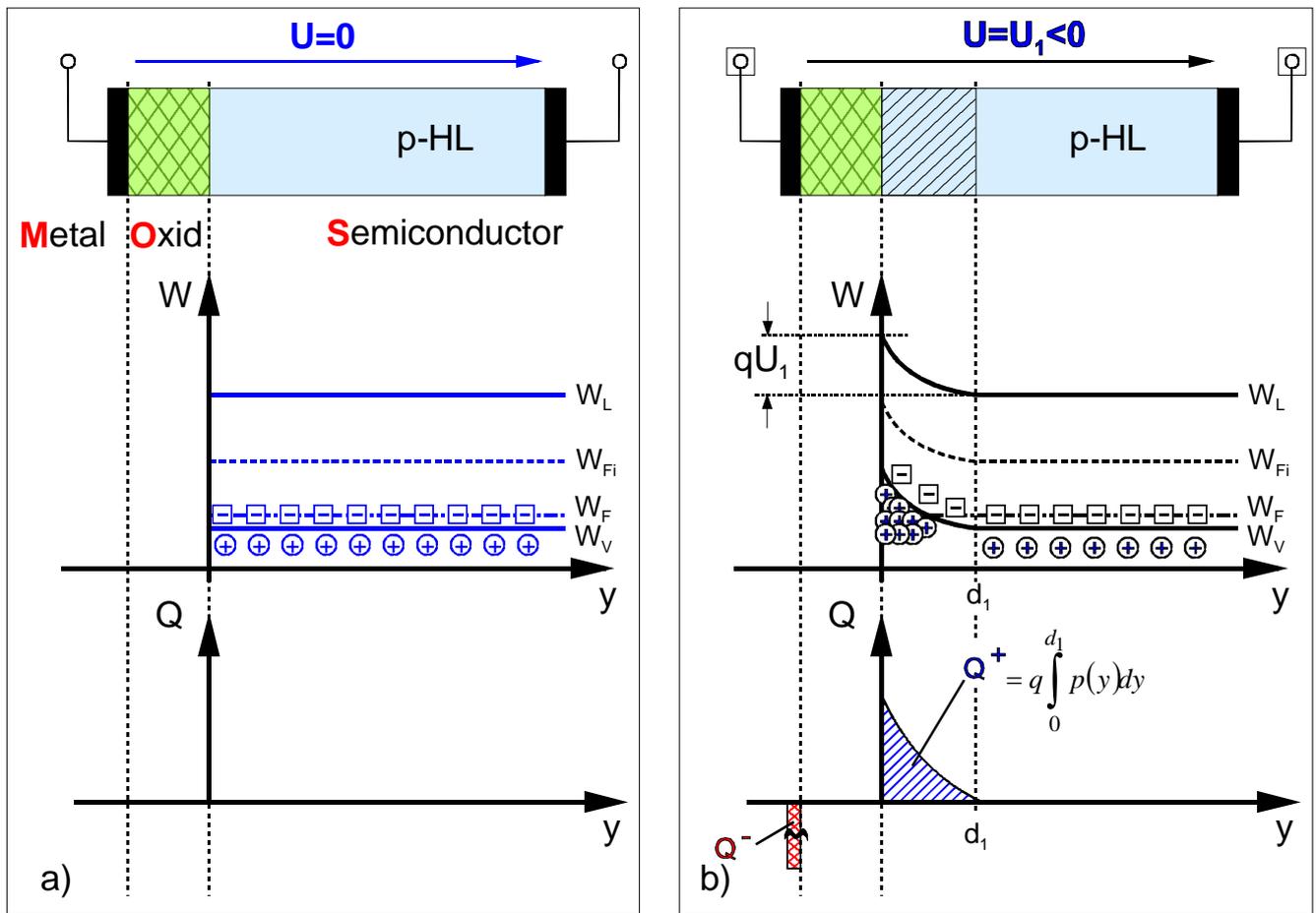


Abb. 0.18 Aufbau und Bändermodell einer idealen MIS-Struktur a) $U = 0$ (Flachbandfall), b) Anreicherung ($U < 0$)

Wird eine Spannung an die Struktur angelegt, so wird ein Feld erzeugt, welches zum Auftreten von Ladungen entsprechend der Poisson-Gleichung führt, die im Falle der idealen MIS-Struktur (MIS-Diode) wegen der Voraussetzung 2 **nur auf der Metallplatte und als Raumladung im Halbleiter** vorliegen können.

Unter dem Einfluss des elektrischen Feldes müssen aber auch die Bandkanten des Halbleiters geneigt sein. In Abb. 0.18 bzw. Abb. 0.19 sind Bänderdiagramm und Ladungsverteilung für einen p-Halbleiter mit $U < 0$ bzw. $U > 0$ dargestellt.

Die angelegte Spannung $U < 0$ an der Metallelektrode führt zu einer negativen Aufladung und einer Anhebung von W_F gegenüber dem Halbleiter. Ladungsneutralität und Poisson-Gleichung fordern eine positive Ladung im p-Typ Halbleiter, die durch eine Anreicherung von Löchern an der Grenzfläche über den durch die Dotierung vorgegebenen Wert hinaus bereitgestellt wird, sowie die damit einhergehende nach oben gerichtete Bandverbiegung im Halbleiter über die Länge d_1 . Wegen der extrem hohen Ladungsträgerdichte im Metall ist die Ausdehnung hier verschwindend gering und die Raumladung entartet zu einer Flächenladung.

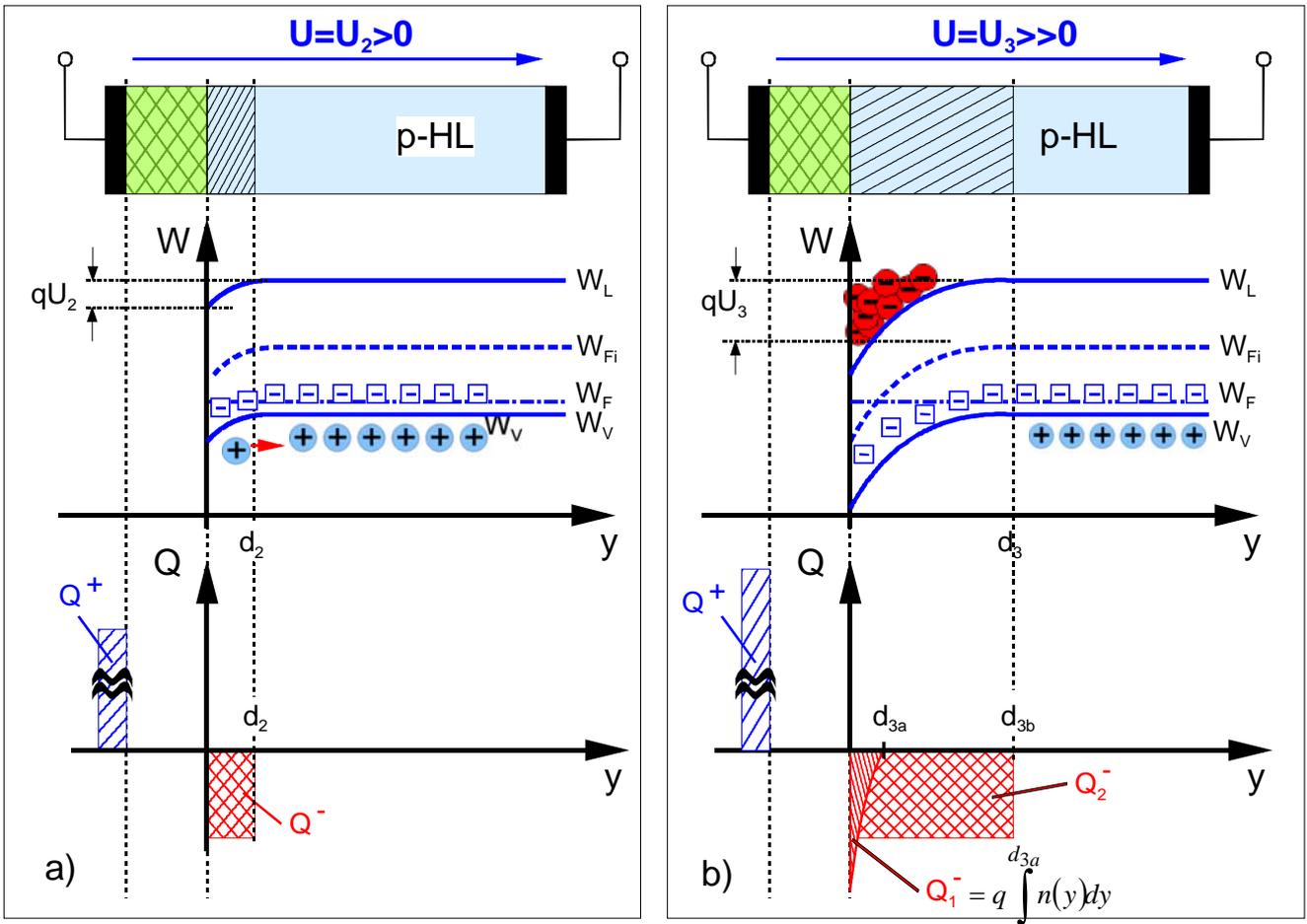


Abb. 0.19 Aufbau und Bändermodell einer **idealen** MIS-Struktur a) Verarmung ($U > 0$), b) Inversion ($U \gg 0$)

Für $U > 0$ (Abb. 0.19a) wird die Metallplatte positiv aufgeladen, W_F gegenüber dem Halbleiter abgesenkt und somit die Bänder im Halbleiter nach unten verbogen. Die Folge ist eine Verarmung an Löchern an der Grenzfläche Isolator-Halbleiter, die die notwendige negative Gegenladung aus ionisierten negativ geladenen Akzeptoren bereitstellt. Die Ausdehnung der verarmten Halbleiter-Raumladungszone ist d_2 .

Wird die positive Spannung weiter vergrößert (Abb. 0.19b), bis an der Grenzfläche $W_{Fi} > W_F$ eintritt, kommt es zur **Inversion**: Die weiter erhöhte positive Ladung auf der Metallelektrode wird nicht mehr ausschließlich durch negativ geladene ionisierte Akzeptoren bereitgestellt, sondern auch durch negativ geladene Elektronen unmittelbar an der Grenzfläche. Dies wird möglich, weil die Ladungsträgerkonzentrationen vom Abstand der jeweiligen Bandkante zum Fermi-Niveau bestimmt werden. Für den Fall:

$$W_{Fi} < W_F$$

folgt trotz p-Dotierung

$$n(x=0) = N_L \cdot e^{-(W_L - W_F)/kT} > N_V \cdot e^{-(W_F - W_V)/kT} = p(x=0)$$

Die durch diese **Inversion** erzeugte **Grenzflächenleitfähigkeit**, die offensichtlich durch die angelegte Spannung gesteuert werden kann, wird in **Feldeffekttransistoren** ausgenutzt: Durch Anlegen eines Feldes können Ströme nahezu leistungslos, eben durch einen reinen Feldeffekt, gesteuert werden.

Eine analoge Behandlung kann für die MIS-Struktur mit einem n-Halbleiter erfolgen, soll hier aber nicht durchgeführt werden.

In allen Fällen gilt für die ideale MIS-Struktur:

$$W_{F,HL} - W_{F,M} = -q \cdot U \quad (0.53)$$

Das heißt aber nicht, dass auch die Bandverbiegung im Halbleiter identisch ist mit $-qU$. Vielmehr fällt auch ein Teil der Spannung und damit der Bandverbiegung über dem Isolator ab, im Falle der Anreicherung und der Inversion sogar der deutlich größere.

0.3.1.1.2 Die nicht ideale MIS-Struktur

Die **nicht ideale MIS-Struktur** unterscheidet sich von der idealen durch

- unterschiedliche Austrittsarbeiten von Metall und Halbleiter
- Grenzflächenladungen
- Ladungen im Isolator

In allen Fällen tritt im Gegensatz zur idealen MIS-Diode bereits bei $U = 0$ eine Bandverbiegung im Halbleiter auf, die in der Folge Influenzladungen im Halbleiter oder auch auf der Metallelektrode verursacht.

Je nach technologischer Realisierung spielen diese zusätzlichen Beiträge eine unterschiedlich große Rolle. Abb. 0.20 stellt den Fall zusätzlicher Ladungen im Isolator für $U = 0$ dar.

Die Fälle **Anreicherung**, **Verarmung** und **Inversion** können ebenso gut auch durch Grenzflächen- und Oxidladungen je nach Vorzeichen und Konzentration realisiert werden. Die charakteristische Größe, die in Summe die Einflüsse der unterschiedlichen Austrittsarbeiten und Grenz- sowie Isolatorladungen beschreibt, ist die **Flachbandspannung** U_{FB} .

Die **Flachbandspannung** U_{FB} ist die Spannung, die man an die reale MIS-Struktur anlegen muss, um das Bänderdiagramm der idealen MIS-Struktur für $U = 0$, eben den flachen Bandverlauf, zu erzeugen.

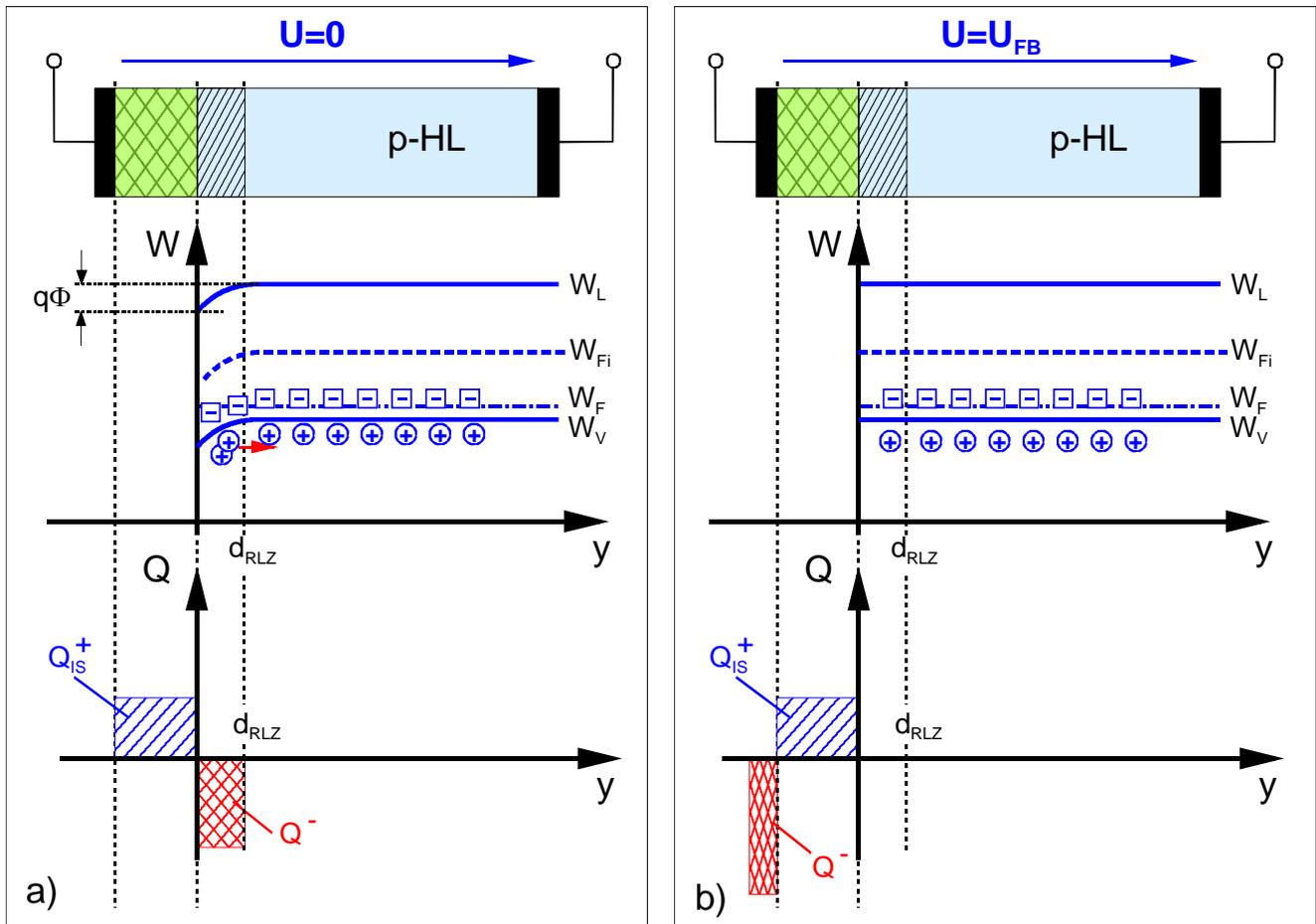


Abb. 0.20 MIS-Struktur, Bändermodell und Ladungsverteilung für einen p-Halbleiter im Fall positiver Isolatorladungen für a) $U = 0$ und b) $U = U_{FB}$

Da für $U = U_{FB}$ im Halbleiter keine Bandverbiegung, somit kein Potenzialabfall und gemäss der Poisson-Gleichung keine Ladung auftritt, beschränkt sich die Ladungsverteilung auf die Metallplatte und den Isolator bzw. dessen Grenzfläche.

Wenn der Unterschied zur idealen MIS-Struktur allein aufgrund der Isolatorladungen Q_{IS} besteht, folgt für die Flachbandspannung:

$$U_{FB} = -\frac{Q_{IS}}{c_{IS}} \quad (0.54)$$

mit

$$c_{IS} = \frac{\epsilon_0 \epsilon_r, IS}{d} \quad c_{IS}: \text{Isolatorkapazität } \underline{\text{pro Flächeneinheit}} \text{ (!)} \quad (0.55)$$

Reale MIS-Strukturen verhalten sich bei angelegter Spannung wie ideale MIS-Strukturen, wenn man eine Verschiebung der Bezugsspannung von $U = 0$ zu $U = U_{FB}$ vornimmt und die zusätzlichen Ladungen gemäss Abb. 0.20 berücksichtigt. Bei der folgenden Behandlung der Kapazitäts-Spannungskennlinien von MIS-Strukturen wird dies deutlich werden.

0.3.1.1.3 Die Kapazitäts-Spannungs-Kennlinie der MIS-Diode

Die MIS-Struktur mit p-Halbleiter stellt eine typische Kondensatoranordnung dar, bei der zwei elektrisch leitfähige Bereiche, Metallplatte und Halbleiter, durch ein Dielektrikum, Isolator und/bzw. Raumladungszone, getrennt sind. Abb. 0.21 zeigt den Querschnitt und die beiden in Serie geschalteten Kapazitätsanteile c_{IS} (Isolator-Kapazität) und c_R (Raumladungskapazität im Halbleiter).

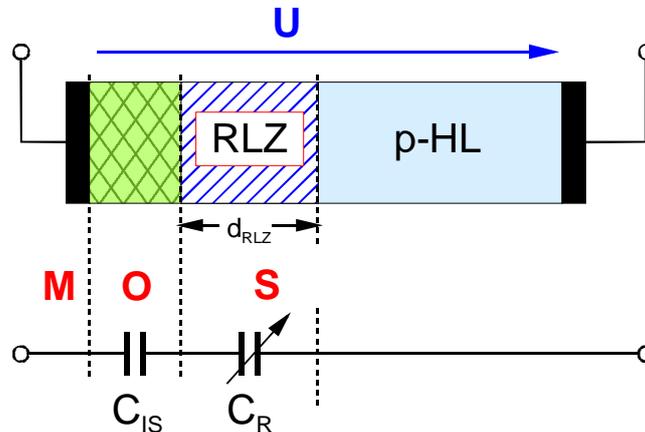


Abb. 0.21 Anteile der Isolator- (c_{IS}) und der Raumladungskapazität (c_R) an der gesamten MIS-Kapazität

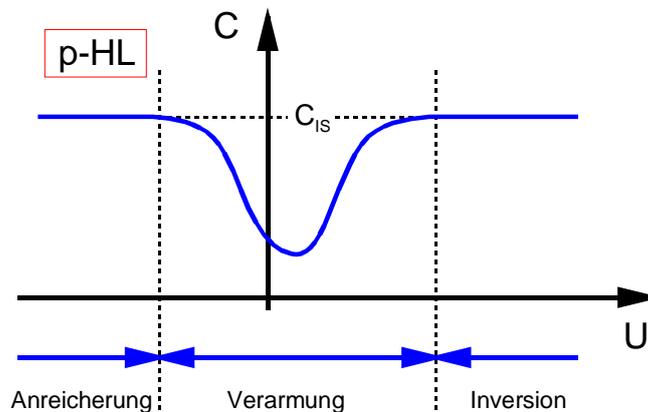


Abb. 0.22 Spannungsabhängigkeit der MIS-Kapazität

Die Isolatorkapazität c_{IS} entspricht der Kapazität eines Plattenkondensators und ist mit

$$C_{IS} = \frac{\epsilon_0 \epsilon_r}{d_{IS}} \cdot A = c_{IS} \cdot A \tag{0.56}$$

gegeben. Die Raumladungskapazität c_R ist im Gegensatz zu c_{IS} spannungsabhängig, da sich die Dicke d_{RLZ} der Raumladungszone, die für c_R als Dielektrikum wirkt, mit der angelegten Spannung ändert. Im Fall der Verarmung gilt für die **Raumladung in der Raumladungszone:**

$$Q_R = q \cdot N_A^- \cdot d_{RLZ} \quad Q_R: \text{Ladung pro Flächeneinheit} \tag{0.57}$$

wenn eine konstante Dotierung des p-Halbleiters vorausgesetzt wird. Allgemein gilt für die Kapazität (pro Flächeneinheit)

$$c_R = \frac{dQ_R}{dU} \tag{0.58}$$

und daher:

$$c_R = q N_A^- \frac{d(d_{RLZ})}{dU} \quad (0.59)$$

Für die Abhängigkeit der Dicke der Raumladungszone von der Spannung gilt im Fall eines p-Halbleiters:

$$d_{RLZ} = \sqrt{\frac{2\varepsilon_o\varepsilon_{r,HL}}{q N_A^-} (U_{FB} - U)} \quad (0.60)$$

und dementsprechend für die Kapazität aufgrund dieser Raumladungszone:

$$c_R = \sqrt{\frac{\varepsilon_o\varepsilon_{r,HL} q N_A^-}{2 (U_{FB} - U)}} \quad \text{für } U < U_{FB} \quad (0.61)$$

Die Spannungsabhängigkeit der Gesamtkapazität der MIS-Struktur, gegeben durch die Serienschaltung von c_R und c_{IS} , folgt entsprechend:

$$c_{MIS}(U) = \frac{c_{IS} \cdot c_R(U)}{c_{IS} + c_R(U)} \quad \text{für } U < U_{FB} \quad (0.62)$$

und ist in Abb. 0.22 dargestellt. Man erkennt, dass c_{MIS} nur im Bereich der Verarmung der Halbleitergrenzfläche spannungsabhängig ist: Verschwindet die Verarmungszone zusammen, so ist c_R im Vergleich zu c_{IS} sehr groß, der Halbleiter ist bis zur Grenzfläche zum Isolator leitfähig, und als Dielektrikum bleibt nur der Isolator übrig, dessen Dicke spannungsunabhängig ist.

Bei weiterer Erhöhung der angelegten Spannung werden Elektronen an der Grenzfläche angereichert. Sie erhöhen zwar die Leitfähigkeit an der Grenzfläche, ändern aber nichts an der Dicke des Dielektrikums. Erniedrigt man, ausgehend von der Verarmung, die angelegte Spannung, so werden mit Auftreten der Inversion wieder Ladungsträger (jetzt Minoritäten, hier Löcher) an der Grenzfläche erzeugt. Die halbleiterseitige "Kondensatorplatte" verschiebt sich somit aus dem Halbleiterinneren wieder an die Grenzfläche, so dass erneut nur die Isolatorkapazität wirksam ist.

Die Kapazität c_{MIS} ist nicht nur von der angelegten Spannung, sondern auch noch von der Betriebsfrequenz abhängig. Sowohl die Grenzflächenladungen Q_{ss} als auch die Ladungen im Isolator Q_{IS} können spannungsabhängig sein, so dass bei Wechselspannung die entsprechenden Zustände umgeladen werden können.

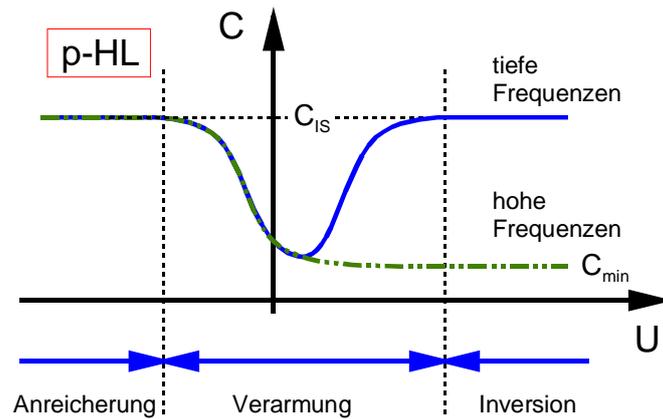


Abb. 0.23 Spannungsabhängigkeit der Gesamtkapazität einer MIS-Struktur und Einfluss der Messfrequenz im Inversionsbereich

Für die Umladung ist aber im Mittel eine Umladezeit τ_s nötig. Für Frequenzen $f > 1/\tau_s$ werden daher nur ein Teil der Zustände umgeladen und gemäss $c = dQ/dU$ ist c_{MIS} für hohe Frequenzen kleiner als für niedrige.

Besonders ausgeprägt ist die Frequenzabhängigkeit im Bereich der Inversion (Abb. 0.23). Für hohe Frequenzen bleibt die Gesamtkapazität bei dem Verarmungswert C_{min} . Dies ist dadurch bedingt, dass sich bei Inversion Minoritäten an der Grenzfläche, also Elektronen in einem p-Halbleiter, sammeln. Diese können nur mittels thermischer Generation von Elektron-Loch-Paaren durch Aufbringen der Bandlückenenergie W_g erzeugt werden, und dies ist für 300K ein extrem unwahrscheinlicher, d.h. langsamer Prozess. Je nach Bandabstand und Reinheit des Materials liegen die Zeitkonstanten für diesen Prozess bei $\tau_{inv} = 10^{-6} s \dots 1s$. Für $f \gg 1/\tau_{inv}$ gilt gemäss $dQ_{inv}/dU \approx 0$ also wieder $C \approx C_{min}$.

Im Gegensatz dazu handelt es sich bei der Verarmung am Rande der Raumladungszone lediglich um eine Bewegung von Majoritäten, die reichlich vorhanden sind und nicht erst erzeugt werden müssen.

0.3.2 Metall-Halbleiter-Grenzflächen; Schottky-Dioden und Ohm-Kontakte

Erst über metallische Kontakte, d.h. durch Metall-Halbleiter-Grenzflächen, werden die Vorgänge im Innern von Halbleitern nach außen verfügbar. Bauelemente und erst recht Schaltungen sind also ohne Metall-Halbleiter-Kontakte nicht denkbar, so dass der Grenzfläche zwischen diesen beiden Festkörpern eine für die Technik zentrale Bedeutung zukommt. Grundsätzlich unterscheidet man zwischen **sperrfreien (Ohm-)** und **sperrenden (Schottky-) Kontakten**. Der erste hat eine lineare Strom-Spannungskennlinie und dient der reinen, niederohmigen Kontaktierung, der zweite besitzt eine stark nicht lineare Diodencharakteristik und kann bereits Träger einer Bauelementfunktion (Diode, Transistor) sein.

0.3.2.1 Das Bändermodell des Metall-Halbleiter-Kontaktes

Das Bändermodell des Metall-Halbleiter-Überganges liefert den Zugang zum elektrischen Verhalten der Anordnung. Denn es bestimmt die Ladungsträgerkonzentration, den Feld- und den Potenzialverlauf im Halbleiter an der Grenzfläche zum Metall. In Abb. 0.24 sind die Bandverläufe idealer Metall- und n-Halbleiteroberflächen getrennt gezeigt, wobei als Bezugspunkt W_o die Energie gewählt wird, die ein Elektron nach Aufbringen der Austrittsarbeit mit $W_{kin} = 0$ besitzt. Da beide Festkörper noch getrennte Systeme darstellen, liegen die Fermi-Energien auf unterschiedlichen Niveaus, weil sie i. a. unterschiedliche Austrittsarbeiten aufweisen. Werden beide idealen Festkörper zusammengefügt (Abb. 0.25) unter der Voraussetzung, dass keine Grenzflächenladungen auftreten (ideales Verhalten), so bilden sie ein gemeinsames System, das sich ohne von außen angelegte Spannung (oder andere äußere Einflüsse) im thermodynamischen Gleichgewicht befindet. Daher hat die Fermi-Energie W_F im ganzen System den gleichen Wert!

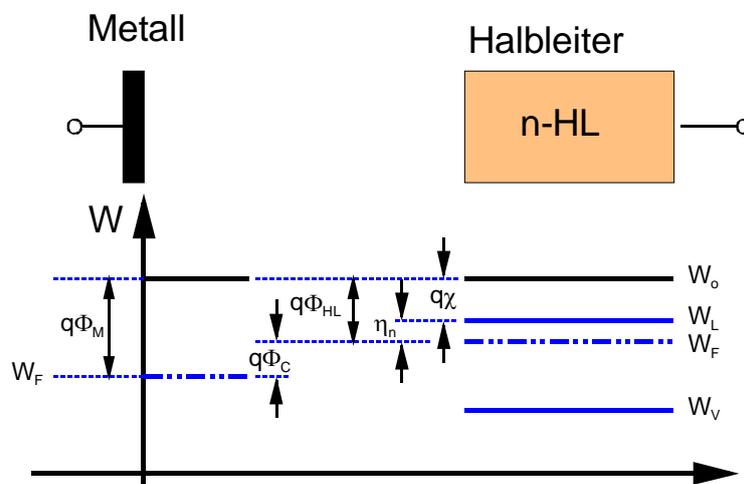


Abb. 0.24 Zum Entstehen eines Metall-Halbleiter-Kontaktes: getrennte Systeme

Im gewählten Fall $\Phi_M > \Phi_{HL}$ ist es für Elektronen einfacher, aus dem Halbleiter ins Metall zu wechseln als umgekehrt aus dem Metall in den Halbleiter. Es wird sich daher ein Gleichgewicht einstellen derart, dass an der Grenzfläche im Halbleiter eine an Elektronen verarmte und durch die zurückbleibenden Donatorrümpfe positiv geladene Zone ausbildet, deren Ladung durch eine entsprechende Elektronenladung "metallseitig" an der Grenzfläche kompensiert wird. Die Ausdehnung dieser zusätzlichen Elektronenanreicherung ist im Metall wegen der um Größenordnungen höheren Elektronenkonzentration extrem klein. Man kann sich anschaulich vorstellen, dass dies gerade die Elektronen sind, die aus dem Halbleiter ins Metall gewechselt sind.

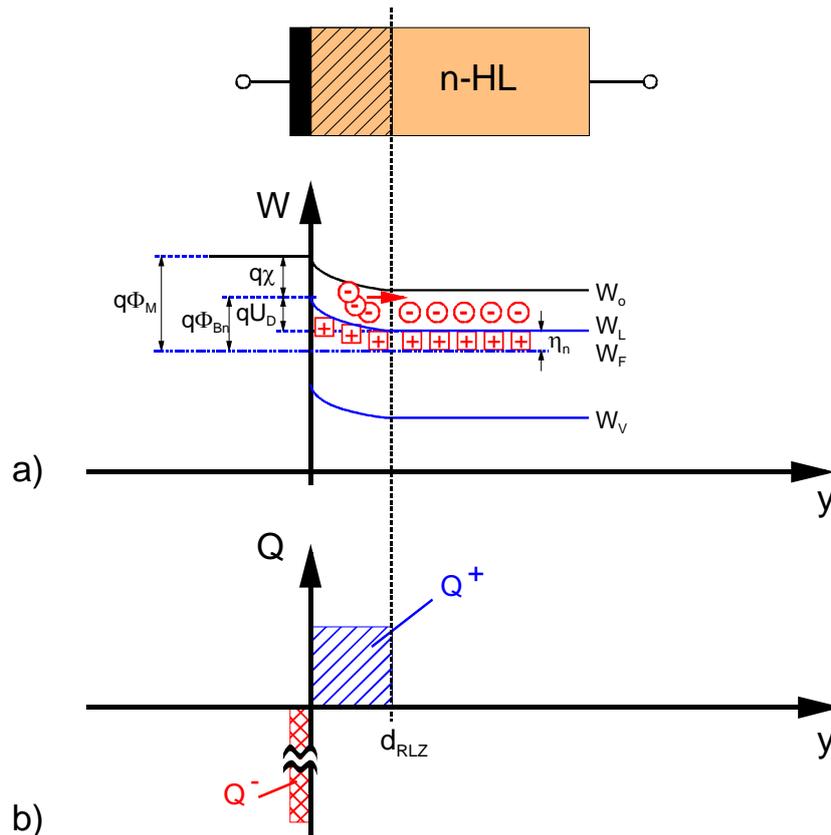


Abb. 0.25 Zum Entstehen des Metall-Halbleiter-Kontaktes: vereinte Systeme ($U = 0$)

Am Übergang Metall \rightarrow Halbleiter bleibt quasi eine Barriere, die durch die **Diffusionsspannung** U_D beschrieben werden kann:

Sie wird nach **W. Schottky Schottky-Barriere** bzw. Barrierenhöhe genannt. Die Barrierenhöhe wird durch das jeweilige System Metall \rightarrow Halbleiter bestimmt.

0.3.2.2 Die Kapazitäts-Spannungskennlinie

Der Metall-Halbleiterkontakt stellt, ähnlich wie die MIS-Struktur, eine Kondensatoranordnung dar, bei der die an freien Ladungsträgern verarmte Raumladungszone das Dielektrikum darstellt. Wie in der MIS-Struktur heißt die entsprechende Kapazität c_R , die Kapazität c_{IS} fehlt jedoch. Völlig analog gelten wieder

$$c_R = \frac{dQ_R}{dU} \quad Q_r : \quad \text{Ladung pro Flächeneinheit !} \quad (0.63)$$

$$Q_R = q N_D^+ d_{RLZ} \quad (0.64)$$

Für $U = 0$ stellt sich als Dicke d_{RLZ0} der Raumladungszone ein:

$$d_{RLZ0} = \sqrt{\frac{2\varepsilon_o\varepsilon_{r,HL}}{q N_D^+} U_D} \quad \text{mit } U_D > 0 \quad (0.65)$$

Dabei entspricht U_D der Flachbandspannung U_{FB} der realen MIS-Struktur. Durch eine außen angelegte Spannung kann die Dicke der Raumladungszone verändert werden:

$$d_{RLZ}(U) = \sqrt{\frac{2\varepsilon_o\varepsilon_{r,HL}}{q N_D^+}} (U_D - U) \tag{0.66}$$

Für $U = U_D$ ist die Raumladungszone völlig abgebaut, d.h. $d_{RLZ} = 0$. Mit den Gln. (0.63) - (0.66) folgt:

$$c_R = \sqrt{\frac{\varepsilon_o\varepsilon_{r,HL} q N_D^+}{2 (U_D - U)}} \text{ (Kapazität pro Flächeneinheit) für } U \leq U_D \tag{0.67}$$

Der Fall $U > U_D$ ist für die Kapazitätsberechnung uninteressant, da kein Dielektrikum mehr vorhanden ist. Für die Strom-Spannungs-Kennlinie wird dieser Fall wieder aufgegriffen. In Abb. 0.26 sind die Bandverläufe für unterschiedliche Spannungen gezeigt.

Die Gln. (0.66) zeigt, dass mit steigender negativer Spannung die Ausdehnung der Raumladungszone mit $\sqrt{U_D - U}$ zunimmt, und die Kapazität mit dem Kehrwert $1/\sqrt{U_D - U}$ abnimmt: Das Bild des Plattenkondensators mit dem Plattenabstand d_{RLZ} ist auch für die Schottky-Diode ein funktionsfähiges Modell. Ferner ist c_R von der Dotierung des Halbleiters abhängig.

Auf der Spannungsabhängigkeit der Schottky-Diode beruht sogar das (neben dem Hall-Effekt) wichtigste Verfahren zur Bestimmung der Dotierstoffkonzentration. Aus Gl. (0.67) folgt

$$\frac{1}{c_R^2} = \frac{2}{q \varepsilon_o \varepsilon_{r,HL} N_D^+} (U_D - U) = \frac{K}{N_D^+} \cdot (U_D - U) \text{ für } U \leq U_D \tag{0.68}$$

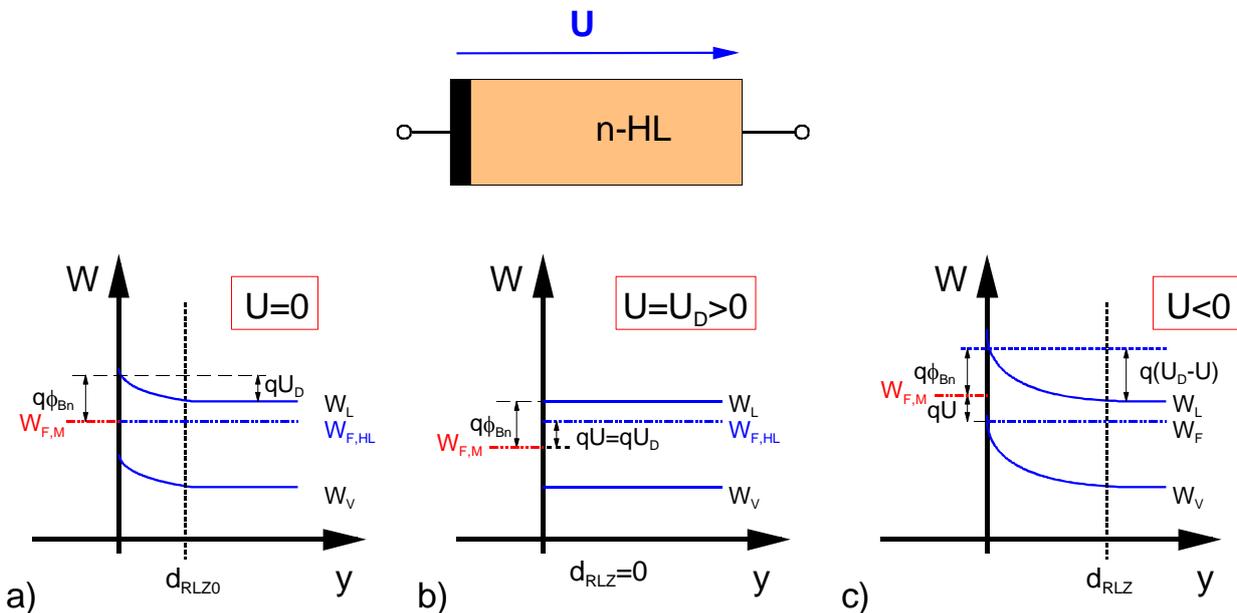


Abb. 0.26 Verlauf der Leitungsbandkante für unterschiedliche Spannungen

Für eine bestimmte Halbleitersubstanz ist K eine Konstante. Aus einer grafischen Darstellung $1/c_R^2 = f(U)$ folgt zunächst aus der Steigung die unbekannte Dotierstoffkonzentration N_D , vorausgesetzt, dass alle Donatoren ionisiert sind ($N_D = N_D^+$). Ferner ist aus dem Achsenabschnitt die Diffusionsspannung bestimmbar. Abb. 0.27 zeigt das Diagramm für zwei unterschiedliche Konzentrationen N_{D1} , N_{D2} . Man beachte, dass sich auch zwei verschiedene U_D ergeben, da die Dotierung auch die Größe $\eta = W_L - W_F$ verändert.

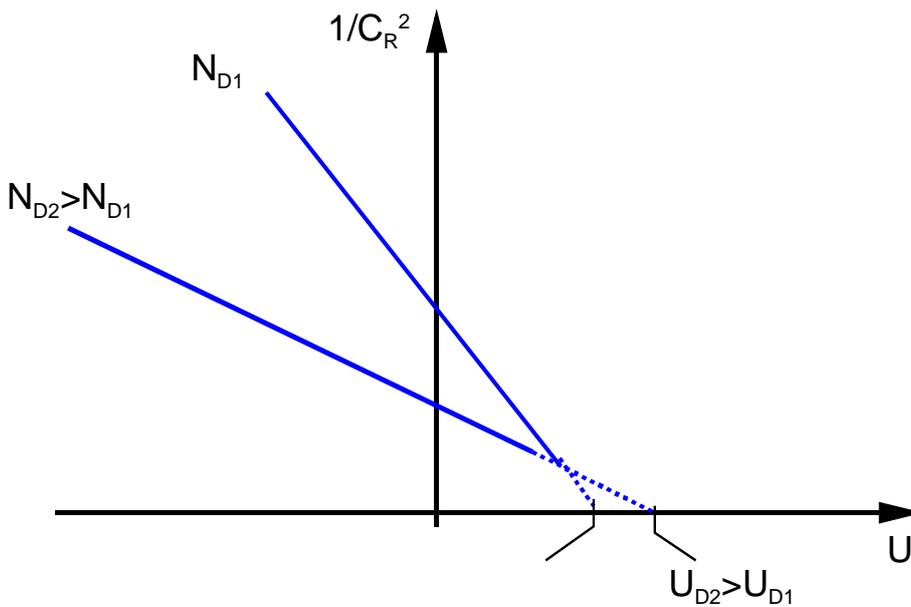


Abb. 0.27 Spannungsabhängigkeit der Raumladungskapazität für zwei verschiedene Dotierungen

Neben dieser messtechnischen Anwendung wird die Schottky-Diode als **Varactor** (variable reactor) meist in der Hochfrequenztechnik als einstellbare Kapazität eingesetzt. Jeder Radioempfänger enthält diese Abstimmeelemente in Schwingkreisen.

0.3.2.3 Die Strom-Spannungs-Kennlinie des Metall-Halbleiter-Kontaktes

Der Stromfluss durch die Metall-Halbleitergrenzfläche muss von Elektronen getragen werden, die die Barriere an der Grenzfläche ($q\Phi_{Bn}$ oder qU_D) überwinden können.

Es ergibt sich eine Diodencharakteristik mit Durchlass- und Sperrbereich.

Der Durchlassstrom I_{DU} steigt für $U > \text{einige } kT$ exponentiell an, so dass man näherungsweise schreiben darf:

$$I_{Du} = I_o \cdot e^{qU/kT} \quad \text{für } U > 3 \frac{kT}{q} \quad (0.69)$$

Das hier entwickelte Modell gilt nur bis zum Flachbandfall, d.h. bis etwa $U \approx 0,6V$ für Silizium. Darüber hinaus fällt die zusätzlich angelegte Spannung im Wesentlichen in den Bahngebieten in größerer Entfernung vom Kontakt selbst ab.

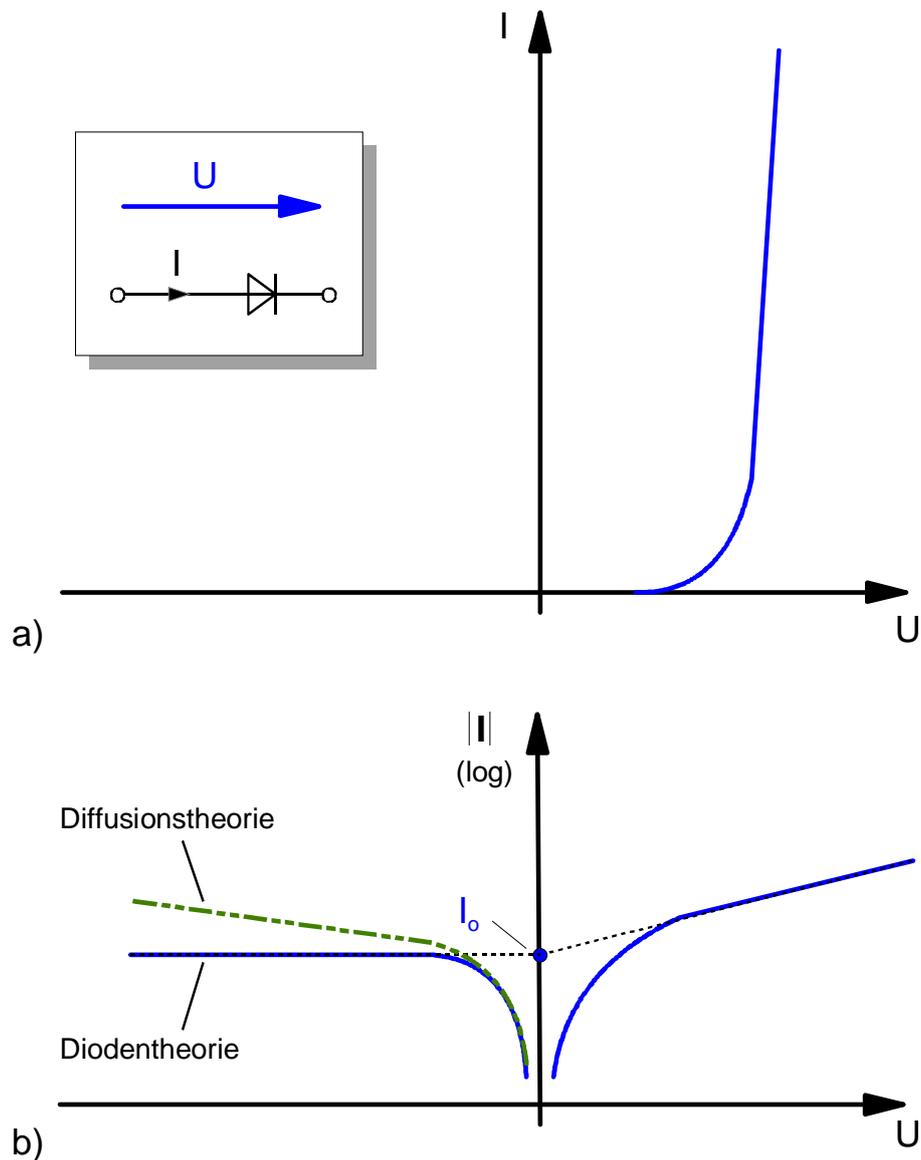


Abb. 0.28 Strom-Spannungskennlinien einer Schottky-Diode
a) linearer Maßstab, b) logarithmischer Maßstab

Für sehr kleine Spannungen kann man die exp-Funktion entwickeln:

$$\frac{I}{I_0} = e^{qU/kT} - 1 \approx \left(1 + \frac{U}{U_{th}}\right) - 1 = \frac{U}{U_{th}} \quad \text{mit} \quad U_{th} = \frac{kT}{q} \quad (0.70)$$

Das Verhalten ist linear, so dass sich die Schottky-Diode bei kleinen Spannungen wie ein ohm'scher Widerstand verhält.

0.3.2.4 Ohm-Kontakte

Ohm-Kontakte sind solche mit einer ohm'schen, d.h. linearen Strom-Spannungs-Kennlinie über den ganzen Betriebsbereich. Nach der bisherigen Behandlung existieren derartige Kontakte nicht, da der Metall-Halbleiter-Kontakt immer auch einen Sperrbereich der Kennlinie aufweist. Für einen sperrfreien Kontakt muss zusätzlich zur bisherigen Behandlung der quantenmechanische Tunneleffekt berücksichtigt werden. Dieser tritt aber nur bei sehr dünnen Barrieren in Erscheinung, wie sie bei einfachem Aufbringen von Metallen auf Halbleiteroberflächen nicht realisierbar sind. Die Breite der Potenzialbarriere hängt analog zu Gl. (0.66) von der Dotierung ab: Wenn diese im Bereich des Kontaktes stark erhöht wird, so wird die Barriere sehr schmal und der Tunnelstrom überwiegt den vorher berechneten Strom bei weitem. Dabei ist das Vorzeichen der angelegten Spannung unerheblich, weil die Elektronen im Metall die gleiche Barrierendicke wie die im Halbleiter sehen. Ferner ist auch die Barrierenhöhe für beide etwa gleich, wie aus Abb. 0.29 hervorgeht.

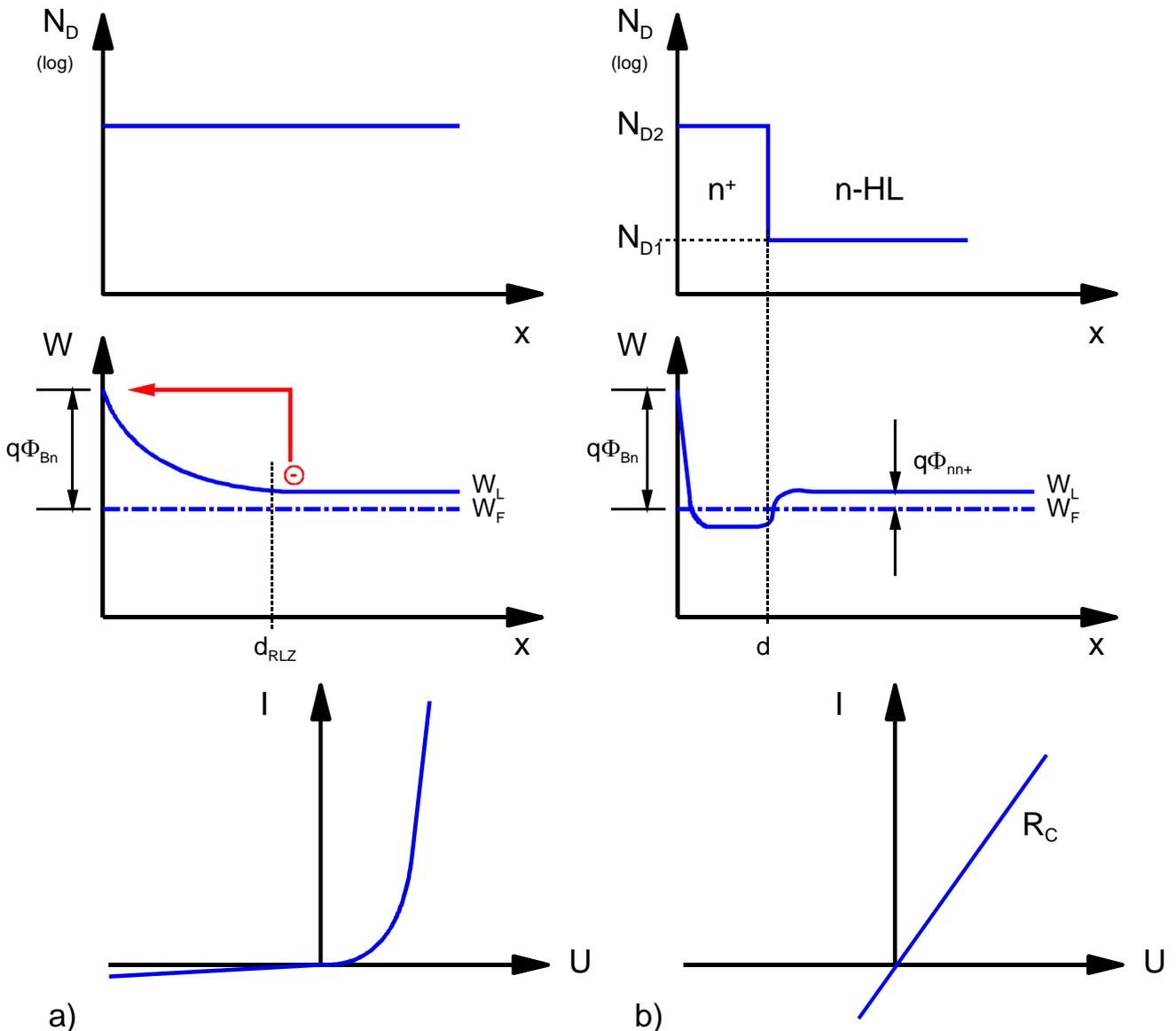


Abb. 0.29 Bändermodell und Kennlinie eines Schottky- (a) und Ohm-Kontaktes (b)

Hier ist angenommen, dass in einer Dicke d der Halbleiteroberfläche die Dotierung von N_{D1} auf N_{D2} so stark erhöht wurde, dass der Halbleiter entartet ist und folglich das Fermi-Niveau oberhalb der Leitungsbandkante liegt.

In Abb. 0.29 ist neben der eigentlichen Schottky-Barriere eine weitere mit $q\Phi_{mn}^+$ gekennzeichnete Barriere erkennbar. Bei geringer Dotierung N_{D1} ist sogar diese und nicht die eigentliche Schottky-Barriere bestimmend für den Kontaktwiderstand R_C . Die Herstellung von Ohm-Kontakten ist in Abb. 0.30 skizziert.

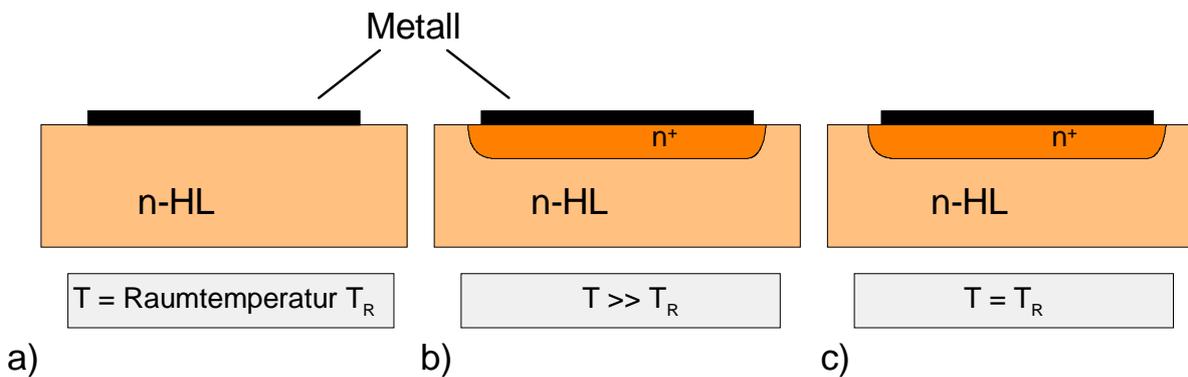


Abb. 0.30 Herstellung eines Ohm-Kontaktes durch Legieren

Zunächst wird im Hochvakuum das Kontaktmetall zusammen mit dem n-Dotierstoff aufgedampft. Dies wird heute meist in mindestens zwei Schritten durchgeführt, bei der die zuerst aufgedampfte Schicht einen besonders hohen Dotierstoffanteil enthält. Sodann wird die ganze Probe kurzzeitig so stark erhitzt, dass der n-Dotierstoff aus der Aufdampfschicht kontrolliert in die Halbleiteroberfläche eindiffundiert und hier eine dünne n⁺-Schicht erzeugt. Nach dem Abkühlen auf Raumtemperatur wird dieser Zustand eingefroren, und damit liegen die Verhältnisse nach Abb. 0.30 vor.

0.4 Halbleiter-Halbleiter-Grenzflächen

Man unterscheidet zwischen Homoübergängen mit gleichem Halbleitermaterial, aber unterschiedlicher Dotierung, und Heteroübergängen, bei denen an der Grenzfläche das Halbleitermaterial wechselt, also beispielsweise auch der Bandabstand W_g .

0.4.1 Der pn-Übergang zwischen Halbleitern gleichen Materials

Der pn-Übergang in ein und demselben Halbleitermaterial ist das Standardbeispiel für Halbleiter-Homoübergänge. Er bildet das Kernstück vieler Bauelemente wie pn-Diode, Bipolartransistor, Thyristor, Photodiode, Solarzelle und, mit Einschränkung, auch LED und Laserdiode.

0.4.1.1 Stromloser pn-Übergang; Bänderdiagramm und Ladungsverteilung

Ströme werden durch ein

elektrisches Feld und bzw. oder durch

Ladungsträgerdiffusion.

angetrieben. Elektrische Felder korrespondieren mit Bandverbiegungen, Ladungsträgerdiffusion mit Konzentrationsgradienten. Ohne Verfälschung der grundsätzlichen Verhältnisse werden zur Vereinfachung folgende **Voraussetzungen** gemacht:

1. Abruptes Dotierstoffprofil:

Die Dotierstoffkonzentrationen N_D bzw. N_A sind im n- bzw. p-Gebiet konstant und fallen an der Grenzfläche abrupt auf Null (vgl. Abb. 0.31b), d.h.

$$N_D = \text{const.} \quad \text{für } x > l_g \quad (0.71)$$

$$N_D = 0 \quad \text{sonst}$$

$$N_A = \text{const.} \quad \text{für } x < l_g \quad (0.72)$$

$$N_A = 0 \quad \text{sonst}$$

2. Alle Dotierstoffatome sind ionisiert:

$$N_A^- = N_A \quad (0.73)$$

$$N_D^+ = N_D \quad (0.74)$$

3. Andere Störstellen, insbesondere Grenzflächenzustände, existieren nicht.

Weit entfernt von der Grenzfläche gelten (vgl. Abb.0.31b) ferner:

$$p_{po} = N_A^- = N_A \quad (0.75)$$

$$n_{no} = N_D^+ = N_D \quad (0.76)$$

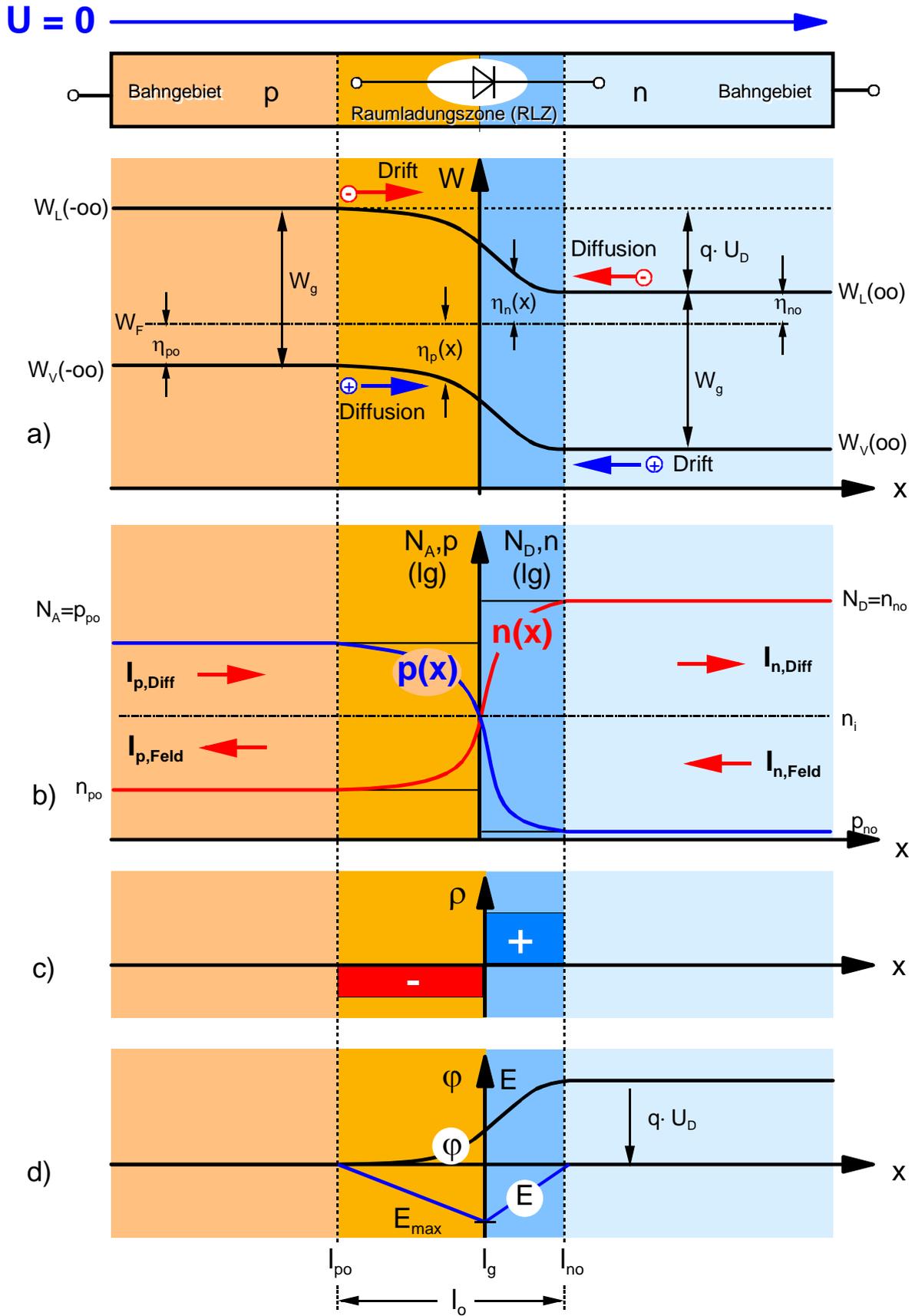


Abb. 0.31 a) Bändermodell eines abrupten pn-Übergangs
 b) Dotierstoff- und Ladungsträgerkonzentration
 c) Verlauf der Raumladung
 d) Potenzial- und Feldverlauf

Die Indizes n bzw. p in Gl. (0.75), (0.76) geben das n bzw. p-Gebiet an, der Index 'o' kennzeichnet den Gleichgewichtszustand ($U = 0, I = 0$). Für diesen Gleichgewichtszustand (thermodynamisches Gleichgewicht) gelten die "Massenwirkungsgesetze", die die Minoritätenkonzentrationen festlegen:

$$n_{po} = \frac{n_i^2}{p_{po}} = \frac{n_i^2}{N_A}; \quad \text{Minoritäten (Elektronen) im p-Gebiet} \quad (0.77)$$

$$p_{no} = \frac{n_i^2}{n_{no}} = \frac{n_i^2}{N_D}; \quad \text{Minoritäten (Löcher) im n-Gebiet} \quad (0.78)$$

Als Bezugsgröße wurde in Abb. 0.31 $W_F = const.$ für den stromlosen Fall (thermodynamisches Gleichgewicht) genutzt.

Die Gl. (0.77), (0.78) stellen in den von der Grenzfläche weit entfernten Bereichen die lokale **Ladungsneutralität** sicher, da hier die Dotierstoffladung (N_D^+ bzw. N_A^-) jeweils von der Ladung freier Ladungsträger (n bzw. p) kompensiert wird.

Im Bereich der Grenzfläche werden die freien Ladungsträger aufgrund ihrer Beweglichkeit nicht dem abrupten Profil der Dotierstoffatome folgen, weil die Diffusion die Elektronen vom n- ins p-Gebiet und die Löcher vom p- ins n-Gebiet treibt. Hier findet daher keine Kompensation von freien (n, p) und ortsfesten (N_D^+, N_A^-) Ladungen statt, und somit ist die lokale Ladungsneutralität nicht mehr gegeben: Im n-Gebiet sorgen die Donatorrümpfe (N_D^+) für eine positive, im p-Gebiet die Akzeptorrümpfe (N_A^-) für eine negative **Raumladung**. Gemäss der Poisson-Gleichung entsteht ein elektrisches Feld und eine Bandverbiegung (Abb.0.31a), die der Diffusionsbewegung der Elektronen und Löcher entgegenwirken und zu einem **Gleichgewichtszustand** führen.

Der Potenzialunterschied U_D (**Diffusionsspannung**) bzw. der Energieunterschied $q \cdot U_D$ zwischen n- und p-Seite und damit die Barriere, die sowohl Elektronen als auch Löcher für den Strom über den pn-Übergang überwinden müssen, ist aus der Abbildung direkt ablesbar:

$$q U_D = W_L(x = -\infty) - W_L(x = \infty) \quad (0.79)$$

und lässt sich angeben (s. Skript "Festkörperelektronik") mit:

$$U_D = U_{th} \ln \frac{N_A N_D}{n_i^2} \quad \text{mit} \quad U_{th} = \frac{kT}{q} \quad (0.80)$$

D.h., diese Diffusionsspannung $U_D > 0$ entspricht der Spannung, die theoretisch von außen anzulegen ist, um den Flachbandfall zu erzeugen. Praktisch ist dies beim realen pn-Übergang nicht möglich, da dies durch den Spannungsabfall in den Bereichen außerhalb des eigentlichen Überganges (Bahngebiete) verhindert wird.

Aus Gl. (0.80) ist die Abhängigkeit der Diffusionsspannung von der Dotierung (N_D, N_A) und den intrinsischen Halbleiterdaten (n_i und damit N_L, N_V, W_g) ersichtlich. Da diese Größen als Argumente der (ln)-Funktion auftreten, ist deren Einfluss auf U_D nur schwach. Für eine ansteigende Dotierung N_D und N_A strebt die Fermi-Energie W_F im n-Bereich gegen W_L und im p-Bereich gegen W_V , so dass sich U_D dem Wert W_g / q nähert. Für die Dotierungen $N_D = N_A = 1 \cdot 10^{17} \text{ cm}^{-3}$ wird für Silizium $U_D \approx 0,8V$ und für GaAs $U_D \approx 1,3V$. Wegen der geringen Zustandsdichte im Leitungs- und Valenzband des GaAs liegt speziell dieser U_D -Wert nahe dem Wert, der der Bandlückenenergie entspricht: $W_g / q = 1,42V$.

Aus dem ortsabhängigen Verlauf von n und p können die Diffusionsströme und aus dem Bandverlauf (und damit dem Potenzialverlauf) die Feldströme mit Hilfe die **Poisson-Gleichung**

$$\Delta\varphi = -\operatorname{div} \vec{E} = -\frac{\rho}{\varepsilon_0 \varepsilon_r} \quad \text{mit}$$

$$\rho = q \left[N_D^+ + p(x) - N_A^- - n(x) \right] \quad (0.81)$$

berechnet werden. Da die Konzentration der freien Ladungsträger n , p innerhalb der Raumladungszone am pn-Übergang stark reduziert ist, können sie in einem ersten Ansatz gegenüber den ortsfesten Ladungen (N_D^+ , N_A^-) vernachlässigt werden.

Wenn mit diesem vereinfachenden Ansatz, $n = p = 0$ innerhalb der Raumladungszone, der Bandverlauf $W(x)$ in zulässiger Näherung berechnet worden ist, können im nachhinein die Konzentrationen gemäss (Gl.(0.24) und (0.25)) berechnet werden.

Somit gilt vereinfachend (vgl. Abb. 0.31c) und mit den Voraussetzungen (1)-(3):

$$\rho = q N_D^+ \quad \text{für } l_g \leq x \leq l_{no} \quad (\text{n-Seite}) \quad (0.82)$$

$$\rho = -q N_A^- \quad \text{für } l_{po} \leq x \leq l_g \quad (\text{p-Seite}) \quad (0.83)$$

Die Poissongleichung kann nun für die n- bzw. p-Seite getrennt integriert werden, und an der Grenzfläche ($x = l_g$) müssen beide Lösungen einander angepasst werden.

Unter Voraussetzung der Ladungsneutralität:

$$q N_A (l_g - l_{po}) = q N_D (l_{no} - l_g) \quad (0.84)$$

kann die n- bzw. p-seitige Ausdehnungen der Raumladungszone bestimmt werden:

$$l_g - l_{po} = \sqrt{\frac{2\varepsilon_0 \varepsilon_r N_D}{q N_A (N_A + N_D)}} \cdot \sqrt{U_D} \quad (0.85)$$

$$l_{no} - l_g = \sqrt{\frac{2\varepsilon_0 \varepsilon_r N_A}{q N_D (N_A + N_D)}} \cdot \sqrt{U_D} \quad (0.86)$$

und somit ist die Gesamtausdehnung der Raumladungszone

$$l_o = l_{no} - l_{po} = (l_{no} - l_g) + (l_g - l_{po}) = \sqrt{\frac{2\varepsilon_0 \varepsilon_r}{q} \left(\frac{1}{N_A} + \frac{1}{N_D} \right)} \cdot \sqrt{U_D} \quad (0.87)$$

(ausführliche Berechnung siehe Skript "Festkörperelektronik")

Die Ausdehnung der Raumladungszone ist für $U = 0$ (stromloser Fall) schwach abhängig von intrinsischen Halbleitergrößen (ϵ_r , n_i über U_D), wesentlich stärker aber von den Dotierungskonzentrationen N_A und N_D . Gl. (0.86) zeigt insbesondere, dass für $N_D \neq N_A$ die Ausdehnung in den schwächer dotierten Bereich größer ist als in den hoch dotierten. Durch stark unterschiedliche Dotierung kann man erreichen, dass die Raumladungszone fast völlig auf der n- oder p-Seite liegt:

$$N_A \gg N_D \quad (\text{p}^+\text{n-Übergang}) \quad l_{no} - l_g \gg l_g - l_{po}$$

$$N_D \gg N_A \quad (\text{pn}^+\text{-Übergang}) \quad l_g - l_{po} \gg l_{no} - l_g$$

Abb. 0.31d gibt den Verlauf des elektrischen Feldes im pn-Übergang wieder, das bei $(x = l_g)$ sein Maximum erreicht. Dieses lässt sich berechnen zu

$$E_{\max} = \frac{1}{q} \frac{dW_L}{dx} (x = l_g) = -\frac{q N_A}{\epsilon_o \epsilon_r} (l_g - l_{po}) = -\frac{q N_D}{\epsilon_o \epsilon_r} (l_{no} - l_g)$$

Aus dem ortsabhängigen Verlauf der Leitungs-, und wegen $W_L - W_V = W_g = \text{const.}$, auch der der Valenzbandkante können die ortsabhängigen Konzentrationsverläufe der freien Ladungsträger $n(x)$ und $p(x)$ bestimmt werden:

$$n_n(x) = N_D \cdot \exp\left[-\frac{1}{2}\left(\frac{l_{no} - x}{L_{Dn}}\right)^2\right] \quad \text{für } l_g \leq x \leq l_{no} \quad (0.88)$$

sowie

$$p_p(x) = N_A \cdot \exp\left[-\frac{1}{2}\left(\frac{x - l_{po}}{L_{Dp}}\right)^2\right] \quad \text{für } l_{po} \leq x \leq l_g \quad (0.89)$$

Die Minoritätsträgerkonzentration gewinnt man aus der Bedingung:

$$n_n \cdot p_n = n_i^2 \quad \text{bzw.} \quad n_p \cdot p_p = n_i^2$$

Somit ist

$$n_p(x) = \frac{n_i^2}{p_p(x)} = N_D \cdot e^{-U_D/U_{th}} \cdot \exp\left[\frac{1}{2}\left(\frac{x - l_{po}}{L_{Dp}}\right)^2\right] \quad \text{für } l_{po} \leq x \leq l_g \quad (0.90)$$

$$p_n(x) = \frac{n_i^2}{n_n(x)} = N_A \cdot e^{-U_D/U_{th}} \cdot \exp\left[\frac{1}{2}\left(\frac{l_{no} - x}{L_{Dn}}\right)^2\right] \quad \text{für } l_g \leq x \leq l_{no} \quad (0.91)$$

Unter Berücksichtigung von:

$$E = - \operatorname{grad} \varphi = \frac{1}{q} \operatorname{grad} W \quad (0.92)$$

ergeben sich die Stromdichteanteile für Elektronen mit

$$J_{n,Feld} = -q n \mu E \quad (0.93)$$

bzw.

$$J_{n,Diff} = q D_n \operatorname{grad} n \quad (0.94)$$

und analoger Ausdrücke für die Löcher lassen sich die Feld- und Diffusionsströme berechnen. Im spannungslosen Fall (thermodynamisches Gleichgewicht) muss der Gesamtstrom gleich Null sein und es muss gelten:

$$J_{ges} = 0 = J_{Feld} + J_{Diff} = J_{n,Feld} + J_{n,Diff} + J_{p,Feld} + J_{p,Diff}$$

Genauer gilt aufgrund des Prinzips des detaillierten Gleichgewichts sogar

$$J_n = J_{n,Feld} + J_{n,Diff} = 0 \quad (0.95)$$

$$J_p = J_{p,Feld} + J_{p,Diff} = 0 \quad \text{bei } U = 0 \quad (0.96)$$

weil jeder Einzelprozess (Elektronen und Löcher getrennt) für sich im Gleichgewicht sein muss.

0.4.1.2 Die Strom-Spannungscharakteristik des pn-Übergangs

Durch eine außen angelegte Spannung wird das durch die Gl. (0.95), (0.96) beschriebene Gleichgewicht gestört. Es zeigt sich aber, dass diese Störung in den meisten Fällen nur sehr gering ist, obwohl sie den Stromfluss durch den pn-Übergang trägt. Das liegt daran, dass die Feld- bzw. Diffusionsströme für sich genommen aufgrund der hohen Konzentrationen, Konzentrationsgradienten sowie elektrischen Felder entsprechend groß sind. Im Vergleich zu den Einzelkomponenten Feld- oder Diffusionsstrom ist aber der Gesamtstrom als Differenz beider Größen sehr klein. Grundsätzlich liegen daher die gleichen Verhältnisse wie im stromlosen Fall vor, und es gilt für nicht zu große, außen anliegende Spannungen (im Vergleich zu U_D):

$$J_n = J_{n,Feld} + J_{n,Diff} \approx 0 \quad (0.97)$$

$$J_p = J_{p,Feld} + J_{p,Diff} \approx 0 \quad \text{bei } U \neq 0, |U| \ll U_D \quad (0.98)$$

Zur Herleitung der Strom-Spannungscharakteristik wird angenommen, dass die gesamte außen angelegte Spannung über der Raumladungszone abfällt. Nach der obigen Darstellung ist dann in den Gleichungen für den stromlosen Fall lediglich die Größe U_D durch $(U_D - U)$ zu ersetzen.

Insbesondere gilt für die Ausdehnung der Raumladungszone

$$l_g - l_p = \sqrt{\frac{2\varepsilon_o\varepsilon_r N_D}{q N_A (N_A + N_D)}} \cdot \sqrt{U_D - U} \quad (0.99)$$

$$l_n - l_g = \sqrt{\frac{2\varepsilon_o\varepsilon_r N_A}{q N_D (N_A + N_D)}} \cdot \sqrt{U_D - U} \quad (0.100)$$

$$l = l_n - l_p = (l_n - l_g) + (l_g - l_p) = \sqrt{\frac{2\varepsilon_o\varepsilon_r}{q} \left(\frac{1}{N_A} + \frac{1}{N_D} \right)} \cdot \sqrt{U_D - U} \quad (0.101)$$

Je nach Vorzeichen von U wird die Ausdehnung der Raumladungszone vergrößert oder verringert.

Mit der Verschiebung der Raumladungszonengrenzen verschieben sich auch die Majoritätenkonzentrationen, die aber nach wie vor an den (jetzt neuen) Grenzen nahezu ihre Gleichgewichtskonzentration n_{no} bzw. p_{po} annehmen.

Dagegen erreichen die Minoritäten ihre Gleichgewichtskonzentrationen nicht an den Raumladungszonengrenzen, sondern erst tiefer in den Bahngebieten (vgl. Abb. 0.32 und 0.33b).

An den Grenzen $x = l_p$ bzw. $x = l_n$ gilt daher

$$n_p(x = l_p) = N_D \cdot e^{-U_D/U_{th}} \cdot e^{U/U_{th}} = n_{po} \cdot e^{U/U_{th}} \quad (0.102)$$

$$p_n(x = l_n) = N_A \cdot e^{-U_D/U_{th}} \cdot e^{U/U_{th}} = p_{no} \cdot e^{U/U_{th}} \quad (0.103)$$

Die Minoritätenkonzentrationen sind also an den **Grenzen der Raumladungszonen** gegenüber dem stromlosen Fall je nach Vorzeichen von U um den Faktor $\exp(U/U_{th})$ angehoben oder abgesenkt (vgl. Abb. 0.32 und 0.33).

Wegen $n_{no} \gg p_n(l_n)$ und $p_{po} \gg n_p(l_p)$ bleiben jedoch, wie bereits erwähnt, die Majoritätenkonzentrationen nahezu unverändert. Die zugehörigen Bänderdiagramme sind ebenfalls angegeben

Die außen angelegte Spannung U tritt (mit der Elementarladung q multipliziert) als Unterschied der Fermi-Energie im n- bzw. p-Gebiet auf (vgl. Schottky-Kontakt).

Durch die Spannung U wird die Bandverbiegung und damit auch das elektrische Feld erhöht (Sperrichtung) oder erniedrigt (Durchlassrichtung).

Für die ortsabhängige Stromdichte ergibt sich:

$$J_{pn}(x) = \frac{q \mu_p U_{th}}{L_p} p_{no} \left(e^{U/U_{th}} - 1 \right) \frac{\cosh\left(\frac{l_1 - x}{L_p}\right)}{\sinh\left(\frac{l_1 - l_n}{L_p}\right)} \quad (0.104)$$

bzw.

$$J_{np}(x) = \frac{q \mu_n U_{th}}{L_p} n_{po} \left(e^{U/U_{th}} - 1 \right) \frac{\cosh\left(\frac{x-l_2}{L_n}\right)}{\sinh\left(\frac{l_p-l_2}{L_n}\right)} \quad (0.105)$$

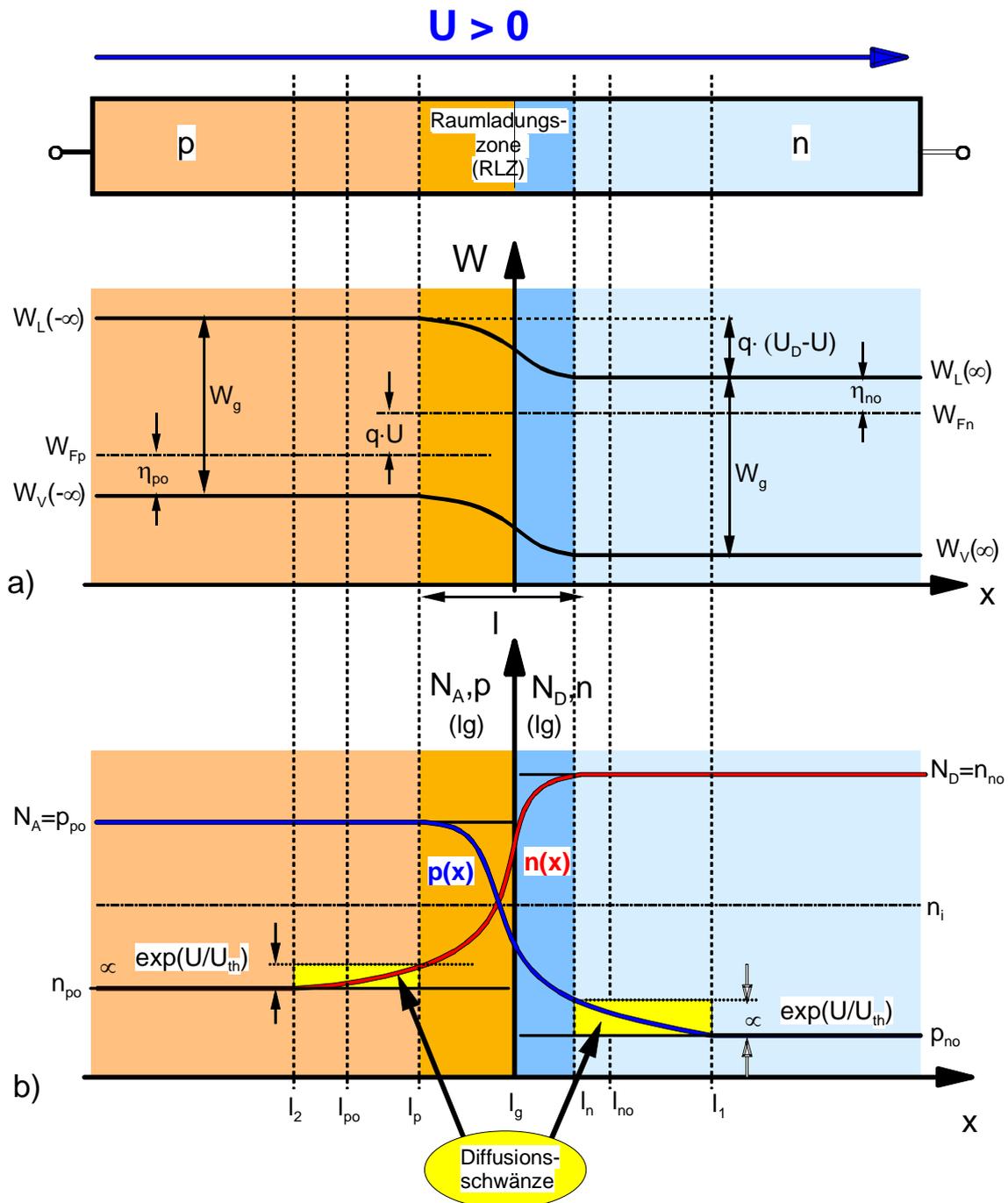


Abb. 0.32 pn-Übergang in Flussrichtung ($U > 0$); Ladungsträgerverteilung und Bändermodell

Die Ausläufer der Minoritäten in die Bahngebiete hinein werden **Diffusionsschwänze** genannt. Durch folgende Näherungen und darauf aufbauenden Überlegungen kann man zeigen, dass der Strom durch den pn-Übergang von den gerade berechneten Diffusionsstromanteilen der Minoritäten im Bereich der Diffusionsschwänze getragen wird:

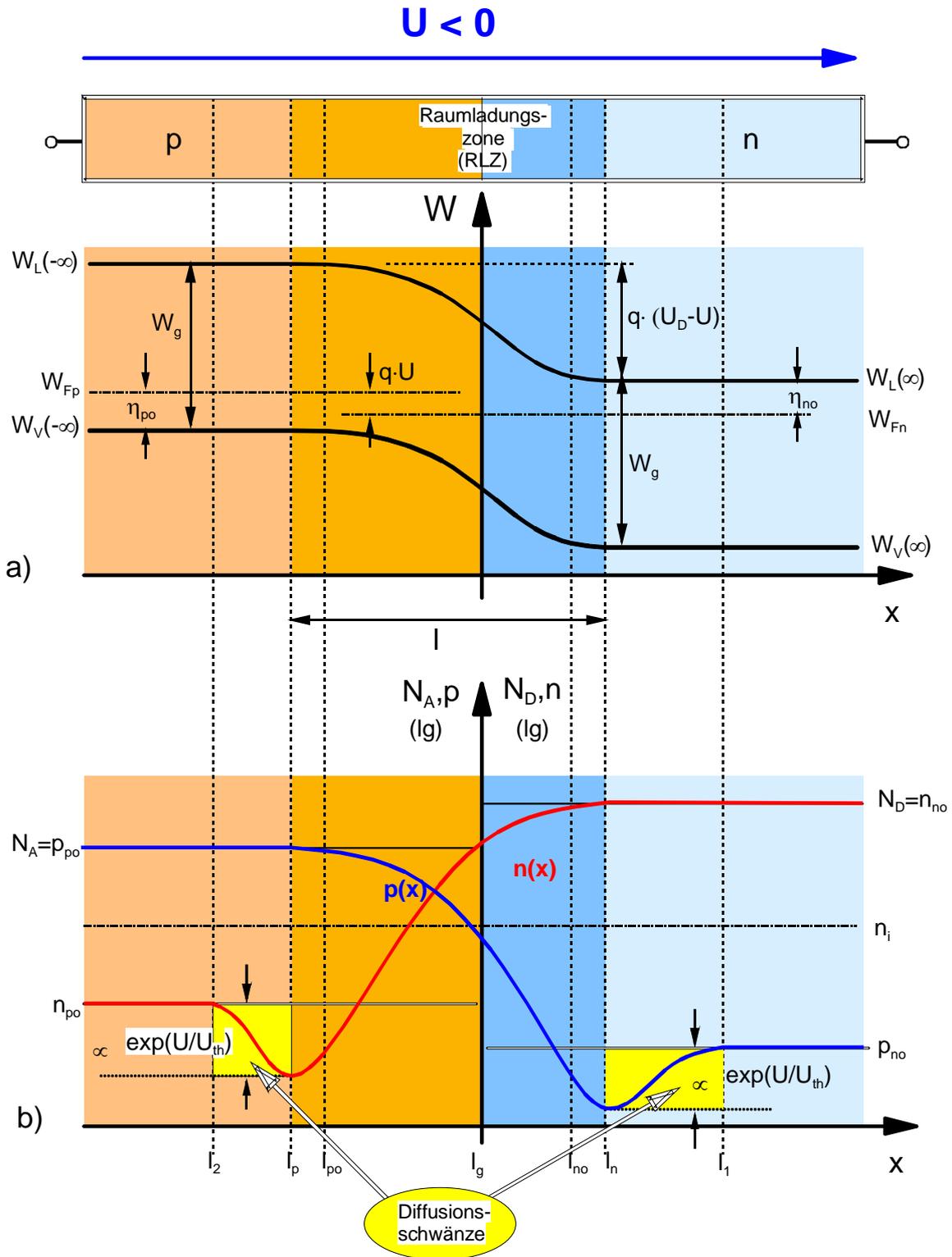


Abb. 0.33 pn-Übergang in Sperrrichtung ($U < 0$); Ladungsträgerverteilung und Bändermodell

1. Keine Rekombination in der Raumladungszone (Shockley-Näherung)

Dies ist gleichbedeutend mit der Bedingung

$$L_{n,p} \gg l_n - l_p \quad (0.106)$$

2. Kein Spannungsabfall in den Bahngebieten,

d.h. insbesondere, dass an den Grenzen der Raumladungszone und in den Bahngebieten keine Feldströme fließen. Real müssen aber Felder in den Bahngebieten existieren, da hier Diffusionsströme mangels Diffusionsgradienten nicht auftreten und somit die Kontinuitätsgleichungen verletzt wären.

Mit diesen Annahmen sind die Gradienten der Minoritätenkonzentrationen innerhalb der Raumladungszone etwa konstant:

$$\frac{dp_n(x)}{dx} \approx const.; \quad \frac{dn_p(x)}{dx} \approx const. \quad \text{für } l_p \leq x \leq l_n \quad (0.107)$$

Diese Konstanz überträgt sich auf die Diffusionsströme:

$$J_{pn}(l_n) \approx J_{pn}(l_g) \quad (0.108)$$

$$J_{np}(l_p) \approx J_{np}(l_g) \quad (0.109)$$

Da an den Grenzen der Raumladungszone die Feldströme verschwinden, die Diffusionsströme also die einzigverbleibenden Anteile darstellen, kann mit Anwendung der Kontinuität des Stromes die Gesamtstromdichte aus dem Diffusionslöcherstrom an der Stelle $x = l_n$ und dem Diffusionselektronenstrom an der Stelle $x = l_p$ berechnet werden. Der Gesamtstrom ergibt sich aus der Summe der Ströme bei $x = l_g$.

Es folgt die Stromdichte-Spannungskennlinie des pn-Überganges:

$$J = J_o \cdot (e^{U/U_{th}} - 1) \quad (0.110)$$

mit

$$J_o = q U_{th} n_i^2 \left[\frac{\mu_p}{L_p N_D} \coth \left(\frac{l_1 - l_n}{L_p} \right) + \frac{\mu_n}{L_n N_A} \coth \left(\frac{l_p - l_2}{L_n} \right) \right] \quad (0.111)$$

J_o enthält dann die technologisch einstellbaren Dotierstoffkonzentrationen.

Die Form der Gl. (0.110) ist wie bereits bei der Schottky-Diode die einer Diode mit Durchlass- und Sperrrichtung:

$$J = J_o \cdot (e^{U/U_{th}} - 1) \quad \text{für } U > 0 \quad (0.112)$$

bzw.

$$J = J_o \cdot e^{U/U_{th}} \quad \text{für } U \gg U_{th} \quad \text{und} \quad \text{Durchlassrichtung} \quad (0.113)$$

$$J = -J_o \quad \text{für } U < 0 \quad \text{Sperrrichtung} \quad (0.114)$$

Die Kennlinien haben die gleiche Form wie die der Schottky-Diode.

Für den Einsatz in Transistoren und Dioden ist eine Fallunterscheidung wichtig, die eine Vereinfachung und Spezialisierung des Ausdrucks für J_o liefert.

1. Langbasisdiode

In diesem Fall können auch die Minoritäten in der pn-Struktur ihren Gleichgewichtswert annehmen. Dann wird vereinfachend angenommen:

$$l_1 - l_n \gg L_p \tag{0.115}$$

$$l_p - l_2 \gg L_n \tag{0.116}$$

Wegen $\coth x \approx 1$ für $x \gg 1$ folgt für J_o der vereinfachte Ausdruck (aus Gl. (0.111)):

$$J_{o,l} = q U_{th} n_i^2 \left[\frac{\mu_p}{L_p N_D} + \frac{\mu_n}{L_n N_A} \right] \tag{0.117}$$

2. Kurzbasisdiode

In diesem Fall gilt umgekehrt:

$$l_1 - l_n \ll L_p \tag{0.118}$$

$$l_p - l_2 \ll L_n \tag{0.119}$$

Mit $\coth x \approx 1/x$ für $|x| \ll 1$ gilt nun

$$J_{o,k} = q U_{th} n_i^2 \left[\frac{\mu_p}{(l_1 - l_n) N_D} + \frac{\mu_n}{(l_p - l_2) N_A} \right] \tag{0.120}$$

Die jeweils kürzere der beiden Längen L_n , L_p bzw. $(l_p - l_2)$, $(l_1 - l_n)$ bestimmt also in beiden Fällen den Sättigungsstrom J_o .

Eine weitere Vereinfachung folgt für stark unsymmetrische pn-Übergänge mit entweder $N_A \gg N_D$ (p⁺n-Übergang) bzw. $N_A \ll N_D$ (n⁺p-Übergang) (Tab. 0.3).

	p ⁺ n ($N_A \gg N_D$)	n ⁺ p ($N_A \ll N_D$)
Langbasisdiode	$J_{o,l} = q U_{th} n_i^2 \frac{\mu_p}{L_p N_D}$ (0.121)	$J_{o,l} = q U_{th} n_i^2 \frac{\mu_n}{L_n N_A}$ (0.122)
Kurzbasisdiode	$J_{o,k} = q U_{th} n_i^2 \frac{\mu_p}{(l_1 - l_n) N_D}$ (0.123)	$= q U_{th} n_i^2 \frac{\mu_n}{(l_p - l_2) N_A}$ (0.124)

Tab.0.3 Vereinfachungen für stark unsymmetrische pn-Übergänge

Der Fall der Kurzbasisdiode gewinnt für den **Bipolartransistor** tragende Bedeutung. Hier liegen zwei pn-Übergänge so dicht benachbart, dass die Minoritätenkonzentrationen des einen pn-Übergangs wesentlich von den Diffusionsausläufern des anderen beeinflusst werden.

Die Sperrströme J_o sind über die Größe n_i^2 stark temperaturabhängig. In den Durchlassströmen macht sich zusätzlich noch die Temperaturabhängigkeit über U_{th} mit

$$J = J_o \cdot (e^{U/U_{th}} - 1)$$

bemerkbar.

0.4.1.3 Die Sperrschichtkapazität des pn-Übergangs

Ähnlich wie der Schottky-Kontakt besitzt der pn-Übergang eine Raumladungs- oder Sperrschichtkapazität, bei der die Raumladungszone von freien Ladungsträgern verarmt ist und das Dielektrikum bildet.

Unter Verwendung von Gl. (0.101) gilt:

$$c = \frac{\epsilon_o \epsilon_r}{l} = \frac{dQ}{dU} = \sqrt{\frac{q \epsilon_o \epsilon_r N_A N_D}{2(N_A + N_D)}} \cdot \frac{1}{\sqrt{U_D - U}} \quad (0.125)$$

Mit steigender Sperrspannung wächst die Ausdehnung der Raumladungszone, somit fällt die Kapazität. In Flussrichtung ist zusätzlich zur **Sperrschichtkapazität** noch die **Diffusionskapazität** zu berücksichtigen.

Die Diffusionskapazität ist dadurch bedingt, dass Spannungsänderungen nicht nur die Ausdehnung der Raumladungszone verändern, sondern auch die Minoritätenkonzentrationen an deren Grenzen. Der damit verbundene Transportprozess ist die Diffusion, die aber nicht momentan, sondern zeitversetzt der angelegten Spannung folgt. Dieses Verhalten lässt sich als Kapazität beschreiben.

Stichwortverzeichnis:Akzeptoren 0-17

amphoter 0-18

Anreicherung 0-32

Austrittsarbeit 0-29, 0-37

Bändertheorie 0-5

Bandstruktur 0-5

Besetzungsstatistik 0-12

Beweglichkeit 0-22

binärer Halbleiter 0-8

Bindungsenergie 0-3

Bipolartransistor 0-54

Bohr'scher Radius 0-17

Bohr'scher Radius 0-4

CMOS 0-29

Coulombkraft 0-4

Coulomb-Potenzial 0-3

Coulombstreuung 0-24

Detektoren 0-29

Diffusionskapazität 0-55

Diffusionskoeffizient 0-28

Diffusionsspannung 0-38

Diffusionsstrom 0-27

Diodencharakteristik 0-36

Donator 0-15

Dotierung 0-15

Dotierungsabhängigkeit 0-24

Driftbeweglichkeit 0-22

Driftgeschwindigkeit 0-22

Eigenhalbleiter 0-12

Eigenleitungsbereich 0-20

Eigenleitungskonzentration 0-12

Eigenwerte 0-3

Eigenwertspektrum 0-3

Einstein Beziehung 0-28

elektrisches Feld 0-22

Energieband 0-5

Energieschalen 0-2

Energieschalen 0-2

Entartung 0-14

Erschöpfungsbereich 0-19, 0-20

Feldeffekttransistor 0-32

Feldstrom 0-26

Fermi-Dirac-Verteilungsfunktion 0-11

Fermi-Energie 0-11

Festkörpermodell 0-5Flachbandspannung 0-32

Galliumarsenid (GaAs) 0-8

Germanium (Ge) 0-8

Geschwindigkeit

thermische Geschwindigkeit 0-22

Geschwindigkeitssättigung 0-25

Grenzflächen 0-29

Grenzflächenladung 0-32

Halbleiter-Halbleiter-Grenzflächen 0-29, 0-44

Metall-Halbleiter-Grenzfläche 0-29

Grenzflächenladung 0-32

Halbleiter 0-7

entarteter Halbleiter 0-14

Halbleiter-Halbleiter-Grenzflächen 0-44Hauptquantenzahlen 0-4

Heterostruktur-Laser 0-29

Heterostruktur-Transistoren 0-29

Homoübergang 0-44

III-V Halbleiter 0-8

intrinsische Konzentration 0-12

Inversion 0-31, 0-32

Ionisierungsenergie 0-4, 0-15**Isolator** 0-7

Isolatorkapazität 0-33

Kapazitäts-Spannungs-Kennlinie 0-34

Kontakt

Ohm-Kontakt 0-36, 0-42

Schottky-Kontakt 0-36

Kontaktwiderstand 0-43

Konzentration

extrinsische 0-15

intrinsische 0-12

Konzentrationsgradient 0-27

kovalente Bindungen 0-15

Kurzbasisdiode 0-54

Ladungsneutralität 0-19

Ladungsträgerbeweglichkeit 0-22

Ladungsträgerkonzentration 0-10

Ladungstransport 0-21

Langbasisdiode 0-54

Legieren 0-43

Leitfähigkeit 0-23

Leitungsband 0-6

effektive Zustandsdichte 0-13

Leitungsbandkante 0-11

Majoritätsladungsträger 0-21

Massenwirkungsgesetz 0-17, 0-19, 0-46

Metall 0-8**Metall-Halbleiter-Kontakt** 0-37

Metall-Halbleiter-Übergang 0-37

Minoritäten 0-46

Minoritätsladungsträger 0-21

MIS-Kapazität 0-34

MIS-Struktur 0-29

MOS 0-29
n-Halbleiter 0-15
Ohm-Kontakt 0-36, 0-42
Ohm'sches Gesetz 0-24
Pauli-Prinzip 0-4
p-Halbleiter 0-15
Phonon 0-23
Phononenstreuung 0-23
Planck 0-2
Potenzialbarriere 0-42
Potenzialtopf 0-3
Prinzip des detaillierten Gleichgewichts 0-49
Quantentheorie 0-2
Raumladung 0-34
Raumladungskapazität 0-34
Raumladungszone 0-39, 0-47
Reservebereich 0-20
Schottky 0-38
Schottky-Kontakt 0-36
Silizium (Si) 0-8
Sperrschichtkapazität 0-55
spezifischer Widerstand 0-9, 0-23
Störstellenhalbleiter 0-12, 0-15
Stromdichte 0-21
thermische Energie 0-7
thermische Geschwindigkeit 0-22
thermodynamisches Gleichgewicht 0-19
Valenzband 0-5
 effektive Zustandsdichte 0-13
Valenzbandkante 0-11
Valenzelektronen 0-2, 0-5
Verarmung 0-32
Verbindungshalbleiter 0-8, 0-18
Verteilungsfunktion 0-10
Widerstand
 spezifischer Widerstand 0-9
Zustandsdichte 0-10
 effektive Zustandsdichte des
 Leitungsbandes 0-13
 effektive Zustandsdichte des Valenzbandes
 0-13

Kapitel 1

Der MOS-Kondensator

Inhalt

1 Der MOS-Kondensator	1-2
1.1 Physikalische Grundlagen	1-2
1.2 Anwendungen des MOS-Kondensators	1-5
1.2.1 Ladungsgekoppelte Bauelemente CCD	1-5
1.2.1.1 Prinzip der Signalspeicherung in CCD-Sensoren	1-5
1.2.1.2 Der Signaltransport	1-6
1.2.1.2.1 Drei-Phasen-CCD	1-8
1.2.1.2.2 Zwei-Phasen-CCD	1-10
1.2.1.3 Prinzipieller Aufbau der Fotodetektoren in CCD-Zellen	1-10
1.2.1.4 CCD mit Fotodiode	1-12
1.2.1.5 CCD mit Foto-Gate	1-13
1.2.1.6 Sensortypen	1-14
1.2.1.6.1 Interline-Transfer-Sensor	1-14
1.2.1.7 Frame-Transfer-Sensor	1-15
1.2.1.8 Full-Frame-Sensor	1-16
1.2.2 CCD-Farbkameras	1-16
1.2.2.1 Drei-Chip Kamera	1-16
1.2.2.2 Ein-Chip Kamera	1-17
1.2.2.3 Weitere Bauformen	1-18
1.2.3 Geschaltete Kapazitäten	1-19
1.2.3.1 Wirkungsweise	1-19
1.2.3.2 Anwendungsbeispiele	1-20
1.2.3.2.1 Filter	1-20
1.2.3.2.2 Integrator mit Operationsverstärkern	1-22
1.2.3.3 Realisierung in MOS-Technik	1-23

1 Der MOS-Kondensator

1.1 Physikalische Grundlagen

Abb. 1.1a, b zeigt die Struktur und die CV-Kennlinie eines MOS-Kondensators auf p-leitendem Substrat. Für die Flachbandspannung gelte $U_F < 0$, z.B. aufgrund positiver Oxidladungen.

1. Bei der Spannung $U = U_F$ liegt der Flachbandfall vor.
2. Bei Anlegen einer Spannung $U < U_F$ tritt eine Anreicherung von Majoritätsladungsträgern (hier: Löcher) an der Grenzfläche *Oxid-Halbleiter* auf. Als wirksame Kapazität verbleibt die Oxidkapazität C_{ox} (Abb. 1.1c).
3. Bei Spannungen $U > U_F$ muss auf der Halbleiterseite eine negative Ladung auftreten. Für den Fall, dass die Spannung U nur wenig größer als U_F ist, besteht diese zunächst aus den ionisierten Akzeptoren. Die oberflächennahe Raumladungszone der Dicke d_2 ist an beweglichen Ladungsträgern verarmt => Verarmung (Abb. 1.1d).
4. Wird die Spannung $U \gg U_F$, so sind zwei Fälle zu unterscheiden:
 - Wird die Spannung $U \gg U_F$ sprungförmig angelegt, so wird die erforderliche negative Ladung zunächst dadurch erzeugt, dass Löcher aus der Oberflächennähe verdrängt werden. Es bleiben die negativ geladenen Akzeptoren übrig (Abb. 1.2c). Die Löcher bewegen sich relativ schnell aus der Raumladungszone hinaus. Dieser Zustand ist jedoch ein Nicht-Gleichgewichtszustand.
 - Im stationären Gleichgewicht gilt überall:

$$n \cdot p = n_i^2 \quad (1.1)$$

mit

$$p = N_V \cdot e^{-(W_F - W_V)/kT} = N_V \cdot e^{-\eta_p/kT} \quad (1.2)$$

Da η_p zur Oberfläche hin stark zunimmt, nimmt p ab und n zu.

Wenn

$$W_F \geq W_{Fi} = W_V + \frac{1}{2}W_g + \frac{3}{4}kT \cdot \ln\left(\frac{m_p^*}{m_n^*}\right) \quad (1.3)$$

ist $n \geq p$ => Inversion (Abb. 1.2d).

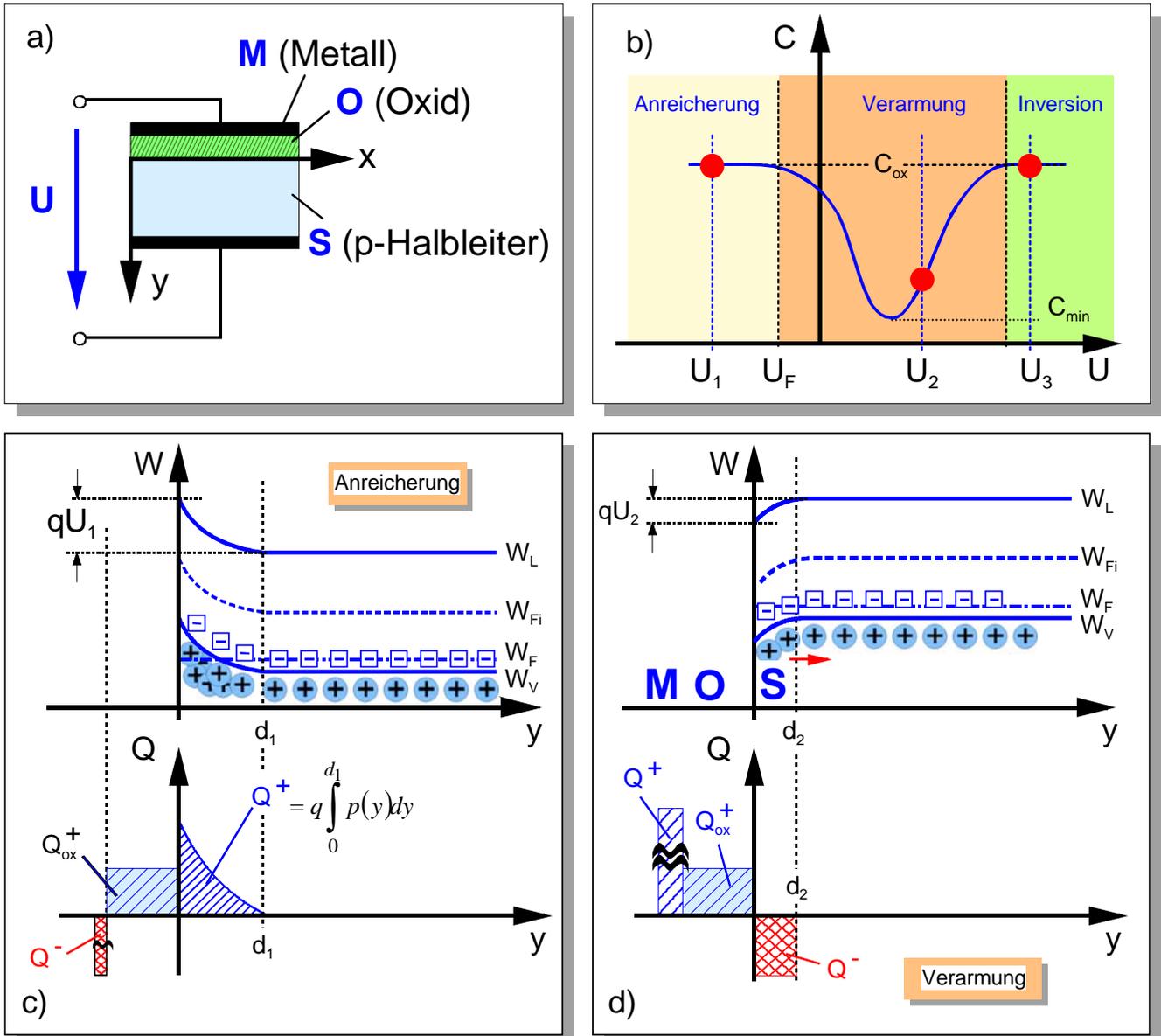


Abb. 1.1 a) MOS-Struktur
 b) Kapazitäts-Spannungs Kennlinie
 c) Anreicherung
 d) Verarmung

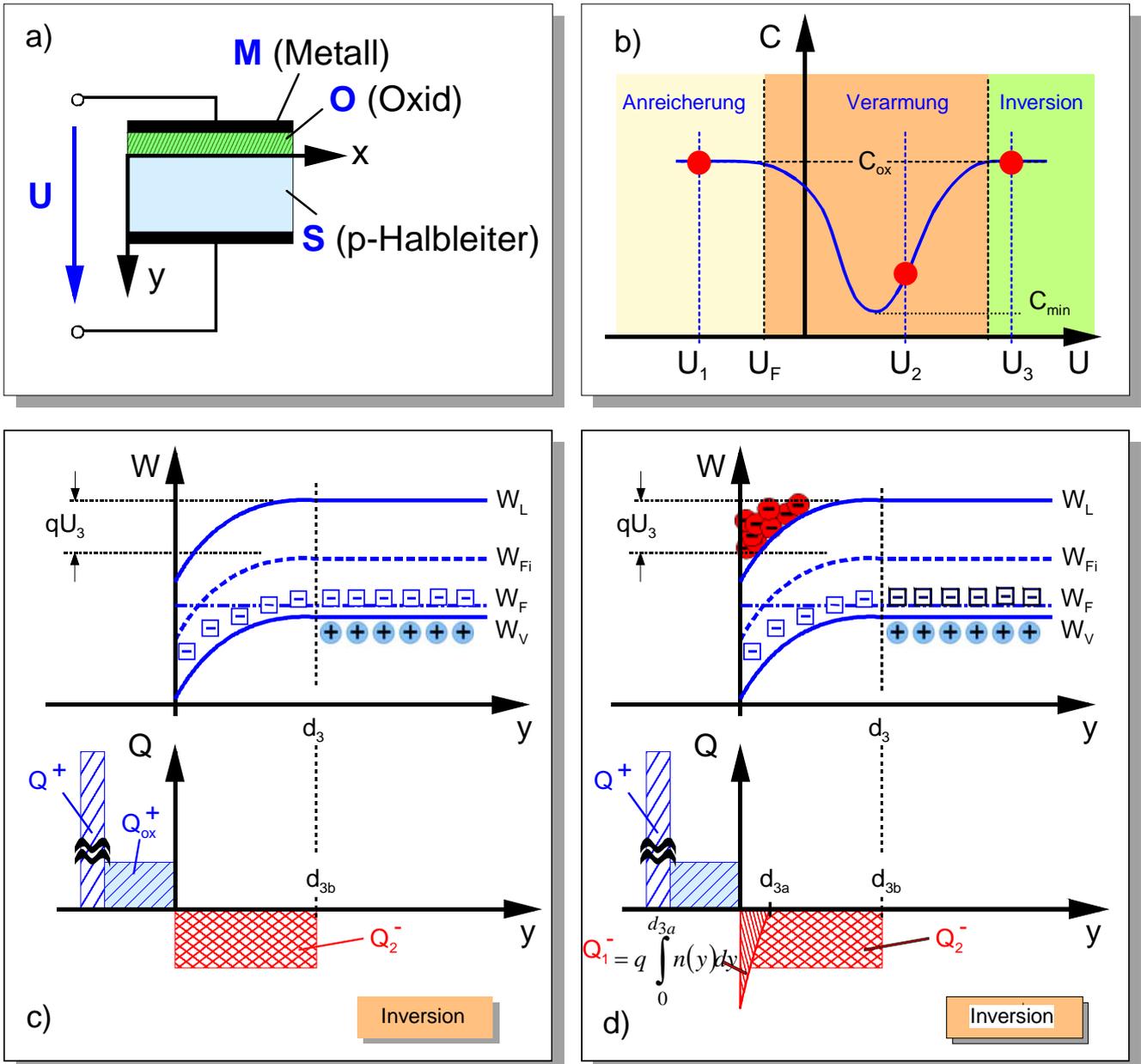


Abb. 1.2 a) MOS-Struktur
 b) Kapazitäts-Spannungs Kennlinie
 c) Inversion (unmittelbar nach dem Umschalten)
 d) Inversion (stationärer Zustand)

Die Elektronen in der Inversionsschicht resultieren z.B. aus der thermischen Generation von Elektron-Loch-Paaren. Die dafür charakteristische Zeitkonstante beträgt in Silizium-Systemen:

$$\tau_{th} \approx 1 \dots 100s$$

Aufgrund der Potentialverhältnisse an der Grenzschicht werden die Elektronen am Oxid-Halbleiter-Übergang (Interface) angereichert, während gleichzeitig eine Verarmung an Löchern stattfindet.

Darüber hinaus können zusätzliche Elektronen durch elektrische oder optische Injektion aufgrund eines äußeren Signals diesen Potentialtopf füllen.

1.2 Anwendungen des MOS-Kondensators

1.2.1 Ladungsgekoppelte Bauelemente CCD

Durch Zusammenschaltung mehrerer MOS-Kondensatoren lassen sich "Ladungsgekoppelte Bauelemente" (engl.: CCD: Charge Coupled Device) realisieren.

CCD-Arrays wurden Ende der 1960er Jahre von den Forschern W.S. Boyle und G.E. Smith in den Bell-Laboratories in den USA entwickelt. Der erste CCD-Chip mit 96×1 Pixel (*picture element*) wurde 1972 vorgestellt.

Ursprünglich wurden mit CCD-Chips Verzögerungsleitungen, Schieberegister und Filter realisiert. Heute werden diese CCD-Sensoren fast ausschließlich als "Image Sensor" in Bilderfassungssystemen eingesetzt, so u.a. in

- Fotoapparaten und Videokameras:
(Eine spezielle Anwendung liegt aufgrund der wesentlich höheren Empfindlichkeit gegenüber fotografischen Schichten (bis etwa 40.000 ASA!) in der Astrofotografie)
- Scannern
- Systemen zur schnellen und berührungslosen Vermessung von Objekten

1.2.1.1 Prinzip der Signalspeicherung in CCD-Sensoren

Innerhalb von Zeitspannen Δt , die klein gegenüber der Zeitkonstanten für die thermische Generation sind, d.h. $\Delta t \ll \tau_{th}$, kann der MOS-Kondensator als Ladungsspeicher eingesetzt werden.

Diese Ladung wird durch Erzeugung von Elektronen-Loch Paaren durch Einstrahlen von Licht erzeugt (Abb. 1.3 und Skript zur *Festkörperelektronik* bzw. Kap. 0).

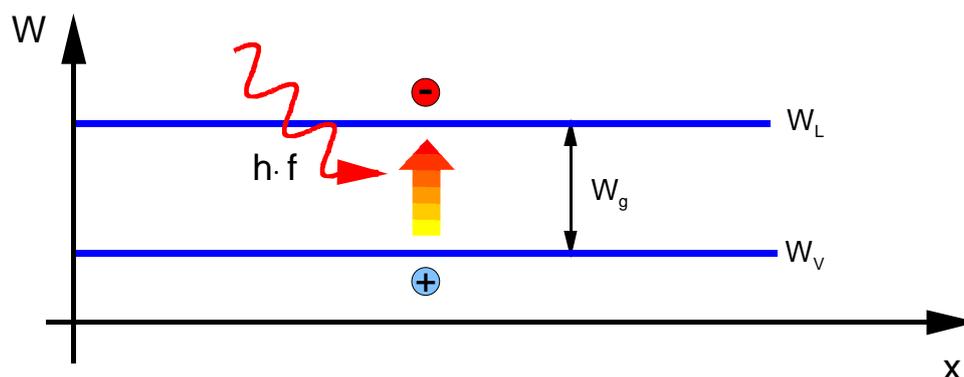


Abb. 1.3 Fotogeneration in einem intrinsischen kristallinen Halbleiter

Die Größe der gespeicherten Ladung hängt von der Dauer des Signals ab, das diese Ladung erzeugt, bzw. Intensität und Wellenlänge des eingestrahlt Lichts.

Die CCD-Zelle kann also als Potentialtopf betrachtet werden, dessen Füllstand linear mit der Signalladung ansteigt.

Dieser *Ladungspool* kann nur eine bestimmte maximale Anzahl von Ladungen speichern, wodurch letztendlich z.B. die Dynamik einer Kamera, in der diese CCD-chips immer häufiger eingesetzt werden, bestimmt wird. Einzelne Zellen sind primär voneinander isoliert, können einander aber unter bestimmten Umständen ungewollt beeinflussen. Das führt zu dem Problem, das man "blooming" nennt.

1.2.1.2 Der Signaltransport

Die Ladungen aus den einzelnen Zellen werden meist über Schieberegister ausgelesen. Im Schieberegister wird der Inhalt einer Zelle einer benachbarten übergeben. Dies kann durch verschiedene Schaltungen erreicht werden, die die Barrieren zwischen den Zellen auf- oder abbauen.

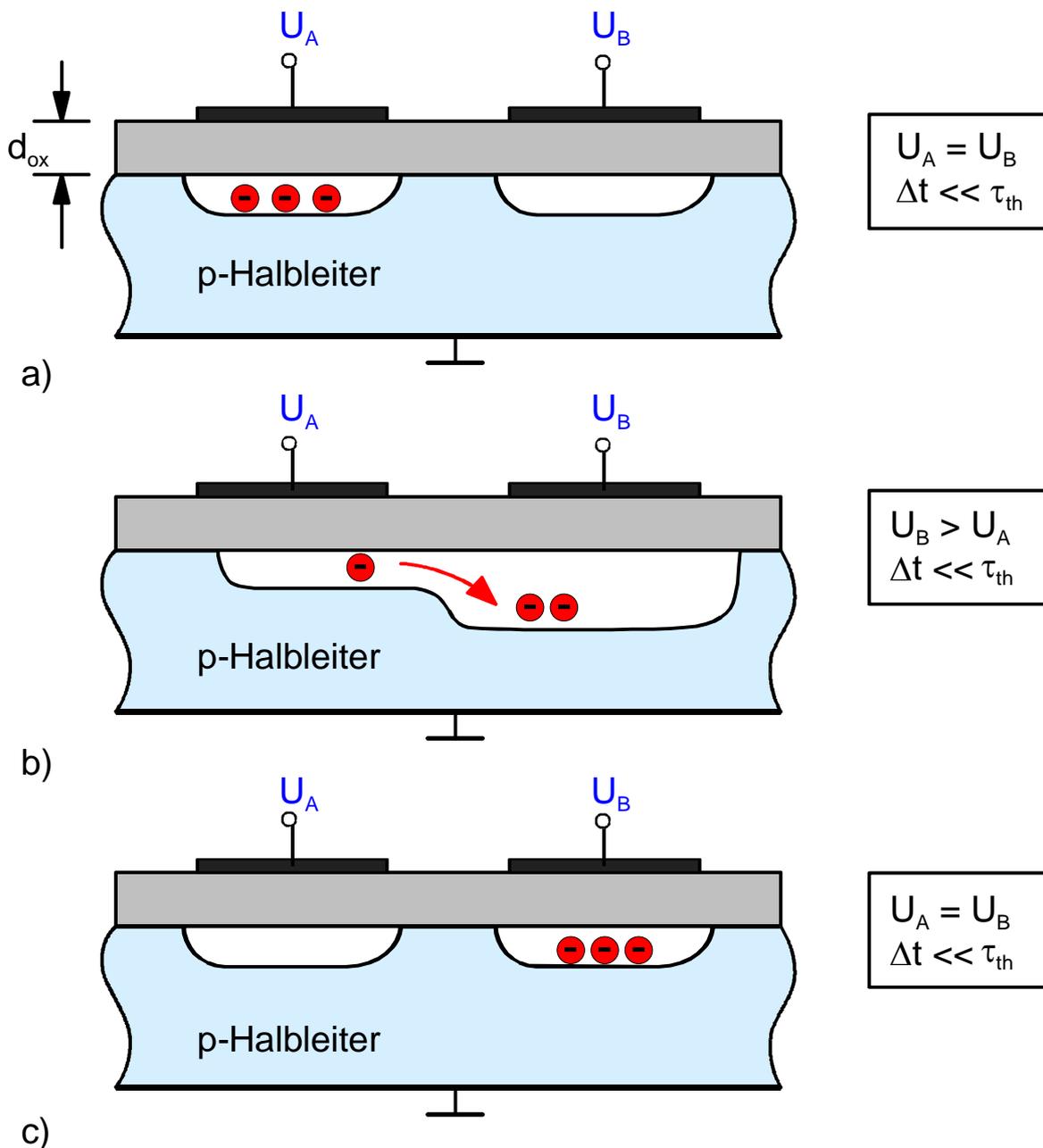


Abb. 1.4 Ladungsverschiebung zwischen MOS-Kondensatoren (Inversion)

Bei geeigneter räumlicher Anordnung der MOS-Kondensatoren kann die Ladung von einem Kondensator in einen anderen übertragen werden (Abb. 1.4):

- In Abb. 1.4a sind beide Kondensatoren so vorgespannt, dass Inversion auftreten kann. Jedoch nur in Kondensator 1 befinden sich Minoritätsladungsträger (hier: Elektronen).
- Wird $U_B > U_A$ (Abb. 1.4b), so dass sich die Inversionszonen überlappen, fließt Ladung von Kondensator 1 zu Kondensator 2, da dort energetisch tiefer liegende Plätze besetzt werden können.
- Wird danach $U_A = U_B$ (Abb. 1.4.c), so ist der Anfangszustand wieder hergestellt, die Ladung ist jedoch nach Kondensator 2 übertragen worden.
- Der Vorgang muss sich innerhalb einer Zeit $\Delta t \ll \tau_{th}$ abspielen.

Um die Ladungen aus den Pixeln gerichtet zu transportieren, werden 2-, 3- oder 4-Phasen Schieberegister aufgebaut.

Das Prinzip des Ladungstransports durch das gesamte Array ist in Abb. 1.5 dargestellt.

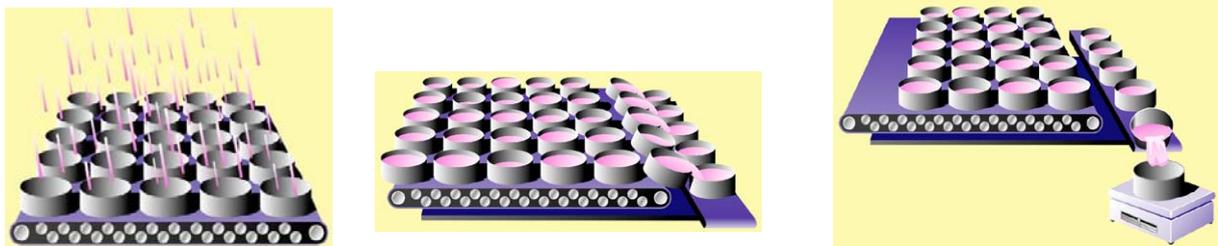


Abb. 1.5 Prinzip des Ladungstransports

Mittels transferierter bzw. gespeicherter Ladung kann über eine in Sperrrichtung betriebene pn-Diode Ladung detektiert werden (Abb. 1.6).

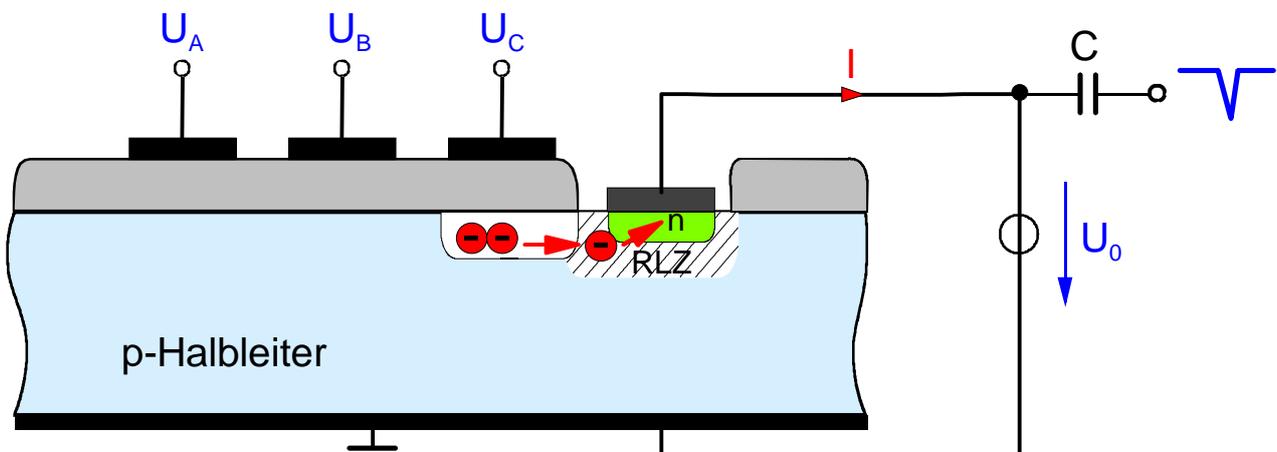


Abb. 1.6 Ladungsdetektion durch Absaugen der Ladung in einer n-dotierten Senke (Drain)

1.2.1.2.1 Drei-Phasen-CCD

Im Fall der Drei-Phasen-CCD besteht eine einzelne Speicherzelle aus 3 MOS-Kondensatoren, von denen jeder mit einer eigenen Spannung angesteuert wird (Abb. 1.7).

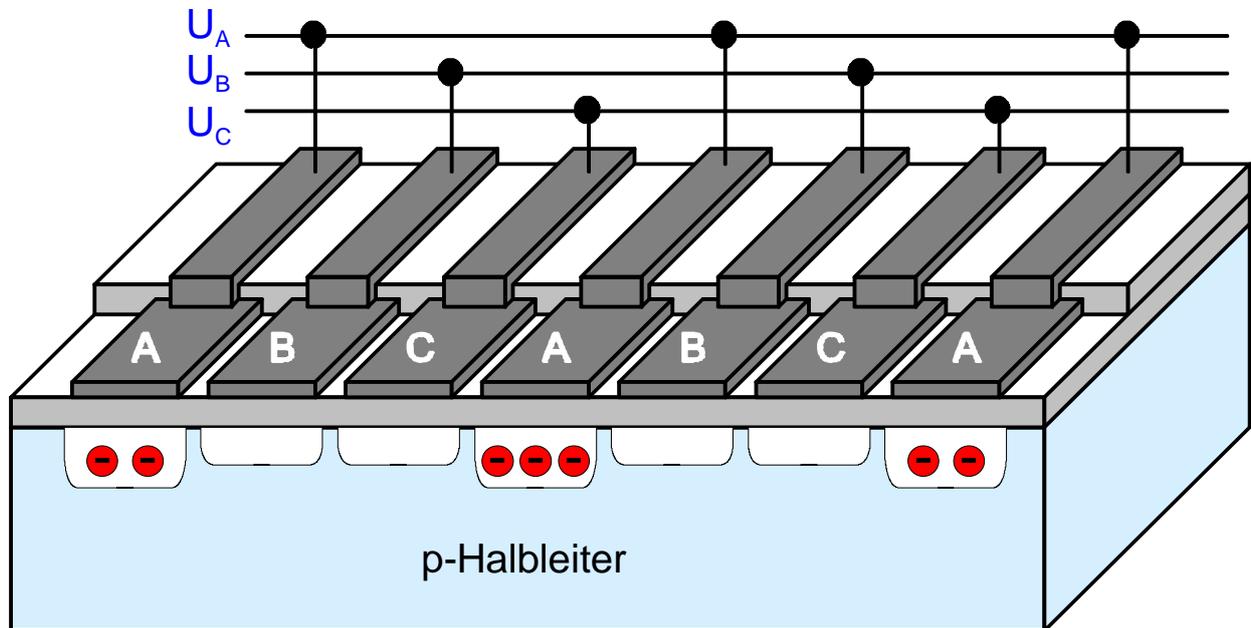


Abb. 1.7 Drei-Phasen-CCD

- Im Ausgangszustand (Abb. 1.8a) erzeugt die Spannung U_A Inversion, die Spannungen U_B und U_C nicht. Dadurch kann Ladung in dem jeweils ersten Kondensatoren einer Zelle gespeichert werden.
- Anschließend wird die Spannung U_B gegenüber U_A soweit erhöht, dass jeweils der Kondensator B jeder Zelle stärker in Richtung Inversion vorgespannt ist als der Kondensator A (Abb. 1.8b). Da U_C unverändert bleibt, ist gewährleistet, dass Ladungen nur von links nach rechts fließen können (auf Kondensator C kann daher nicht verzichtet werden!).
- Im nächsten Schritt (Abb. 1.8c) wird $U_A = U_C$ (keine Inversion!), aber $U_B > U_A = U_C$ (Inversion, jedoch weniger stark als vorher). Die Ladung ist damit vom Kondensator A zum Kondensator B übertragen worden.

Der Zyklus $a \rightarrow b \rightarrow c$ muss noch zweimal wiederholt werden, bis der Ausgangszustand:

- Ladung im Kondensator A
 - Kondensatoren B, C nicht in Inversion
- wieder hergestellt ist.

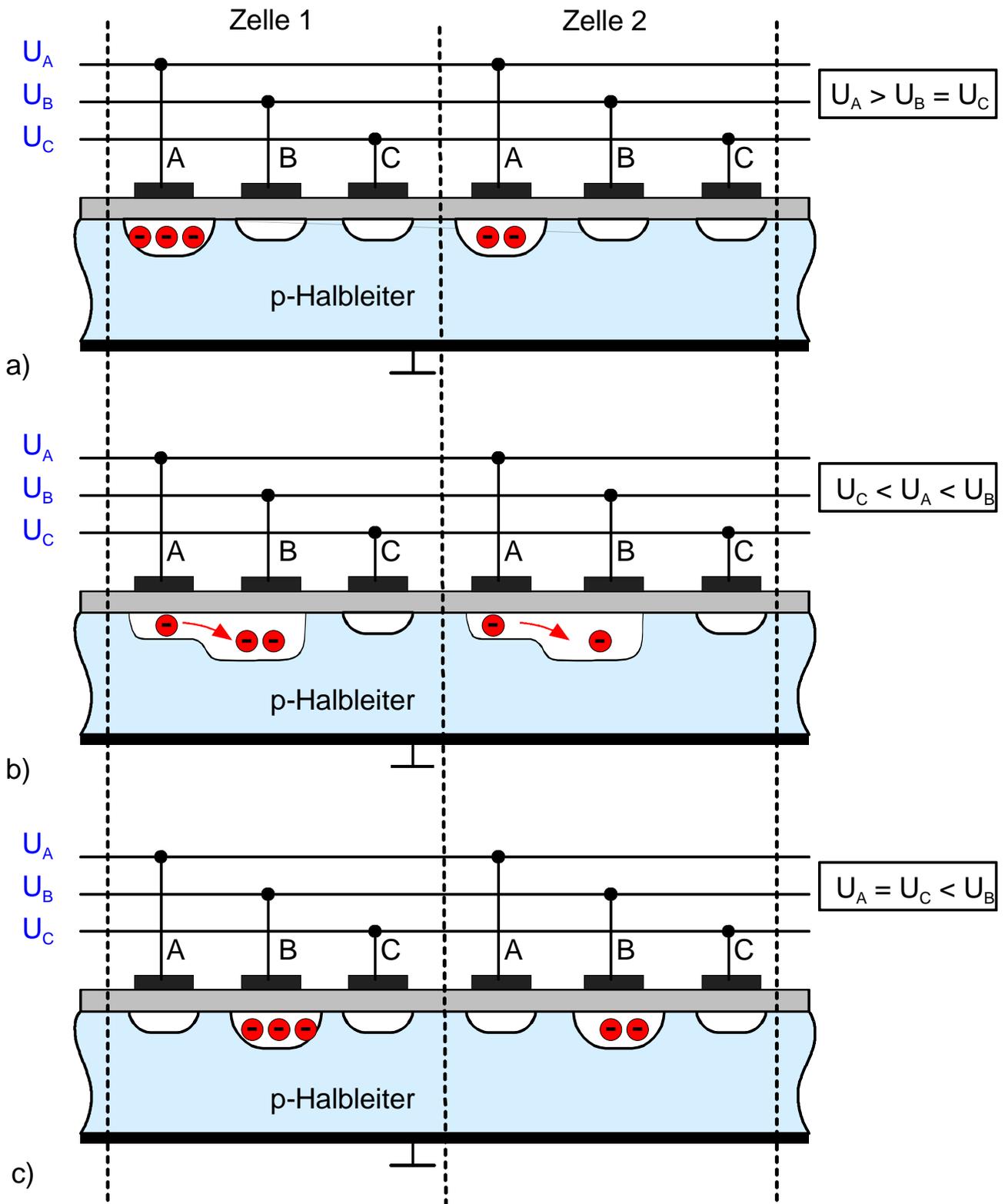


Abb. 1.8 3-Phasen-Ladungsgekoppeltes Bauelement

1.2.1.2.2 Zwei-Phasen-CCD

Damit ein gerichteter Transport der Ladungen mit nur 2 Taktphasen möglich ist, muss die CCD-Zelle unsymmetrisch, wie in Abb. 1.9 dargestellt, aufgebaut sein. Es werden dabei stets zwei direkt nebeneinander liegende Kondensatoren gleichzeitig angesteuert. Da einer von beiden Kontakten jeweils näher am Oxid-Halbleiter-Interface liegt, ist der darunter liegende Bereich weiter in Inversion getrieben als der gleichzeitig angesteuerte, andere Kondensator. Dadurch bedingt liegen dort für die Elektronen energetisch günstigere Zustände vor und es kommt zu einer Ansammlung der Ladungen in diesem Potentialtopf.

Bei Erhöhung des Potentials an den Kontakten B und gleichzeitiger Reduzierung des Potentials an den Zellen A wird damit ein Ladungstransfer von A nach B möglich.

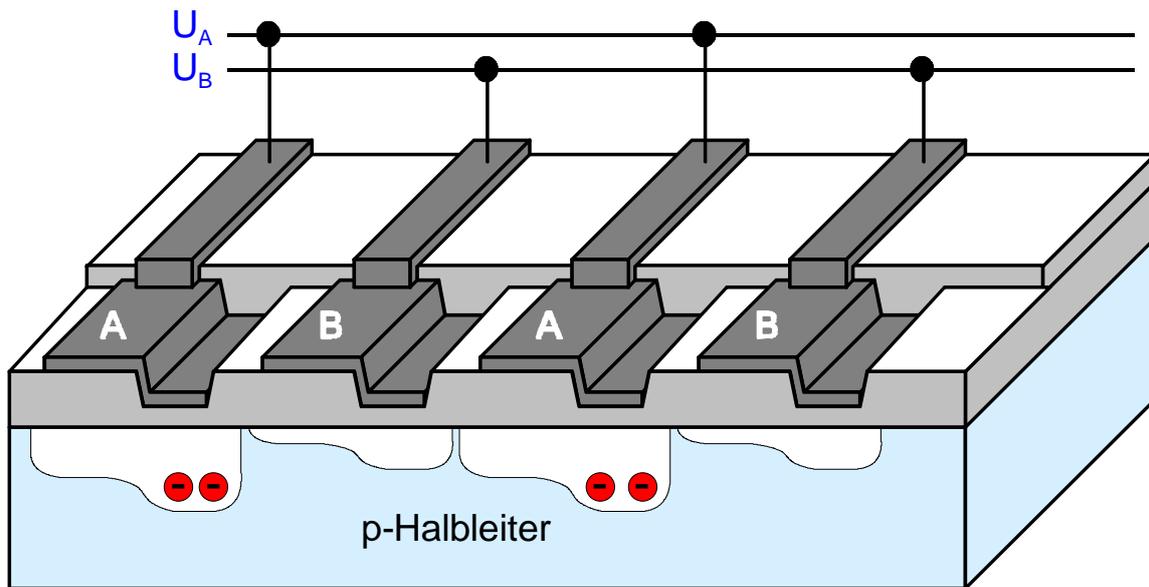


Abb. 1.9 Aufbau des Zwei-Phasen CCD

1.2.1.3 Prinzipieller Aufbau der Fotodetektoren in CCD-Zellen

Ein CCD-Sensor besteht aus einem geometrisch sehr exakten Raster der oben beschriebenen lichtempfindlichen Zellen. Bei der Bildbearbeitung entspricht eine Zelle genau einem Pixel.

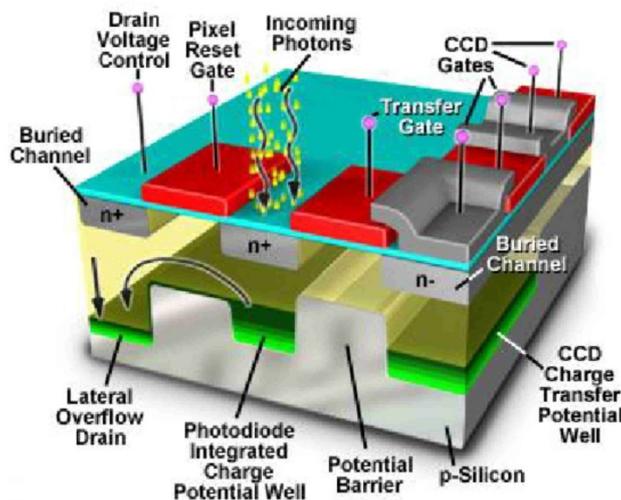


Abb. 1.10 Prinzipieller (möglicher) Aufbau eines CCD-Sensors

Abb. 1.11 zeigt einen typischen, handelsüblichen CCD-Chip.

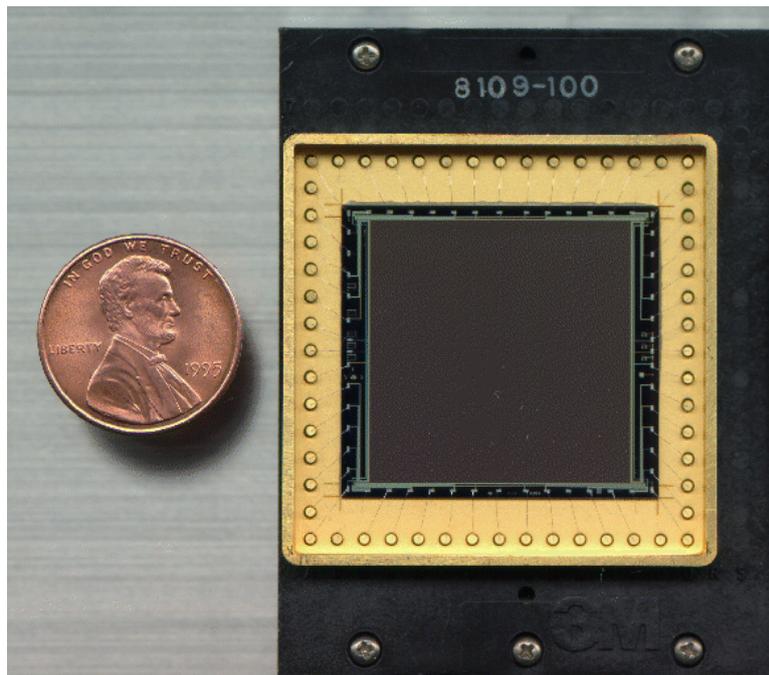


Abb. 1.11 Handelsübliches CCD-Array

Bei Lichteinfall auf eine dieser Zellen wird, wie beschrieben, eine Ladung aufgebaut, die umso größer ist, je länger die Zelle belichtet wird. Durch diese Form der Integration bedingt, liegt die Bildinformation auf dem Sensor nicht in digitaler, sondern analoger Form vor.

Die einzelnen Zellen berühren sich nicht direkt, sondern sind je nach Sensortyp, voneinander durch Stege oder Potentialwälle getrennt. So wird einerseits verhindert, dass die Ladungsträger von der einen Zelle in die andere überlaufen, zum anderen sind diese Stege auch für das Auslesen des Zelleninhalts von Bedeutung. Folglich füllen je nach Sensortyp die lichtempfindlichen Zellen nicht den ganzen Sektor, sondern ein Teil der Sensorfläche wird als Transport- bzw. Sperrfläche genutzt.

Nach der Belichtung der einzelnen Zellen werden die Ladungen ausgelesen. Dieser Vorgang funktioniert bei allen Sensoren nach dem sogenannten "Eimerkettenprinzip". (s. Kap. 1.2.1.2)

Fotoempfindliche CCD-Pixel können grundsätzlich in zwei verschiedenen Bauformen ausgeführt werden:

- CCDs unter Verwendung von n^+p -Fotodioden
- CCDs mit Foto-Gate, d.h. die MOS-Kapazität dient direkt zur Sammlung der Ladungsträger

Bei der Entwicklung von CCD-Chips muss weiterhin beachtet werden, dass Oberflächenzustände, die einen vorübergehenden Einfang der Signalladung bewirken, äußerst schädlich sind, da sie eine Reduzierung des Transportwirkungsgrades mit sich bringen.

Deshalb wird der Halbleiter unterhalb der Elektroden inhomogen dotiert. Die Potentialmulde zur Aufnahme der erzeugten Ladungen befindet sich dann nicht mehr an der Grenzfläche zwischen Halbleiter und Isolation, sondern etwas darunter. Diese werden als "*buried-channel-CCD*" bezeichnet.

1.2.1.4 CCD mit Fotodiode

In Abb. 1.12 sind der Aufbau und die Taktphasen für ein CCD mit Fotodiode schematisch dargestellt. Ein Pixel besteht aus einer n^+p -Fotodiode, einem Transfer-Gate und einem Transport-Gate. In dieser Bauform stellt die Fotodiode das fotoempfindliche Bauteil dar, wohingegen das eigentliche CCD, das sich aus mehreren Transport-Gates in jedem Pixel zusammensetzt, die fotogenerierten Ladungen zum Ausgang transportiert. Im dargestellten Querschnitt verläuft dessen Schieberichtung senkrecht zur "Papier"-Ebene. Das Transfer-Gate dient als Verbindungsglied zwischen Fotodiode und Transport-Gate.

Der Betrieb des Pixels erfolgt unabhängig von der Zahl der Taktphasen des CCD in zwei Schritten:

Zum Beginn der Belichtungsphase befindet sich die Fotodiode in Sperrrichtung, wobei die n^+ -Seite mit keiner Spannungsquelle verbunden ist (*floating node*). Die Fotodiode arbeitet somit im integrierenden Betriebsmodus.

In der Belichtungsphase werden durch Lichteinstrahlung Elektronen und Löcher erzeugt, die durch das elektrische Feld in der Raumladungszone voneinander getrennt und gespeichert werden. Während der Belichtung liegt das Transfer-Gate auf einem anderen niedrigen Potential φ_s . Solange das Potential der n^+ -Seite der Fotodiode noch oberhalb von φ_s liegt, existiert für die gespeicherten fotogenerierten Elektronen eine Potentialbarriere.

In der Austaktphase wird die Spannung am Transfer-Gate erhöht und damit die Potentialbarriere für die Elektronen verringert. Dies bewirkt, dass ein Teil der fotogenerierten Ladung über das Transfer-Gate abfließen kann. Das Transport-Gate liegt auf einem höheren positiven Potential als das Transfer-Gate, so dass die Ladung weitgehend vollständig dorthin überwechselt. Gleichzeitig wird damit das Transfer-Gate auf ein definiertes Potential zurückgesetzt, das einem ungefüllten Potentialtopf entspricht, und mit ihm auch die Fotodiode.

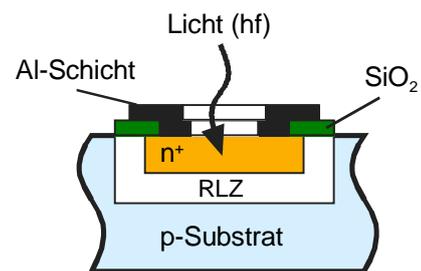
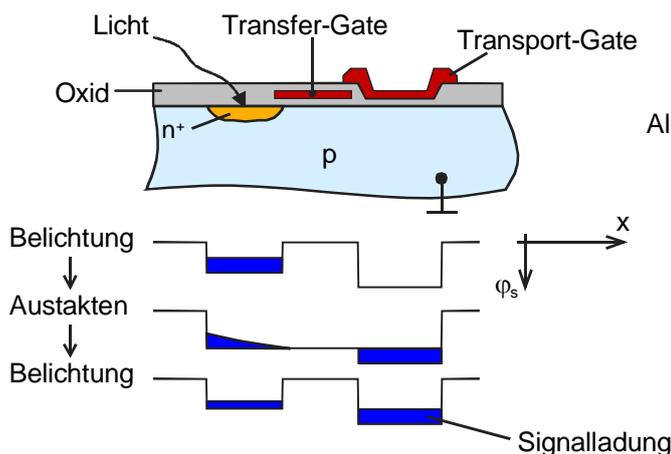


Abb. 1.12a Schematischer Aufbau eines CCD mit Fotodiode und örtlicher Verlauf des Oberflächenpotentials in verschiedenen Phasen der Bildaufnahme

Abb. 1.12b Querschnitt einer Fotodiode

Nach dem Auslesen liegt somit die n^+ -Seite auf einem positiven Potential, so dass die Fotodiode in Sperrrichtung vorgespannt ist. Dies ist genau der Ausgangszustand, wie er für die Belichtungsphase notwendig ist.

Das Transport-Gate, unter dem sich nach dem Auslesen der Fotodiode die Signalladung befindet, ist mit den anderen Gates im CCD-Element gekoppelt und wird so angesteuert, dass die Ladungspakete in Richtung des Ausgangs verschoben werden. Die hier dargestellte Anordnung von Fotodiode, Transfer-Gate und Transport-Gate entspricht dem Interline-Transfer-CCD (Kap. 1.2.1.6.1)

Ein Nachteil des CCD mit Fotodiode liegt darin, dass nur das Gebiet der n^+p Diode fotoelektrisch aktiv und somit der sogenannte *Fillfaktor* des Bauelements entsprechend gering ist.

1.2.1.5 CCD mit Foto-Gate

Beim CCD mit Foto-Gate wird der gesamte Raum unterhalb des Gates zum Sammeln der fotogenerierten Ladungsträger genutzt. Die Trennung zwischen fotoempfindlichen und ladungsverschiebenden Elementen entfällt somit. Abb. 1.13 zeigt eine Realisierung mit vier Gates pro Pixel (je zwei befinden sich auf gleichem Potential), bei der zum Ladungsträgertransport vier verschiedene Oberflächenpotentiale erzeugt werden (s. Zwei Phasen-CCD).

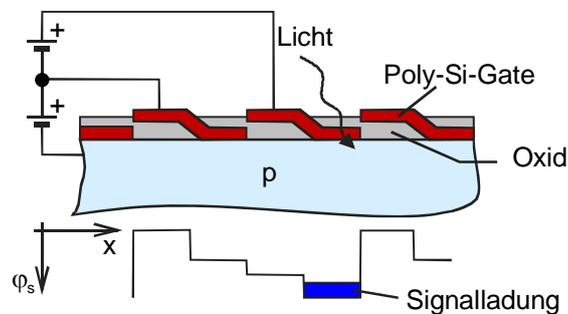


Abb. 1.13 Schematischer Aufbau eines CCD mit Foto-Gate und örtlicher Verlauf des Oberflächenpotentials für den statischen Fall (konstante Gate-Spannungen)

1.2.1.6 Sensortypen

1.2.1.6.1 Interline-Transfer-Sensor

Die Interline-Transfer-Sensoren bestehen aus einem streifenförmigen Muster von lichtempfindlichen Sensorflächen (Belichtungsbereich) und lichtundurchlässigen Schieberegistern.

CCD-Aufbau

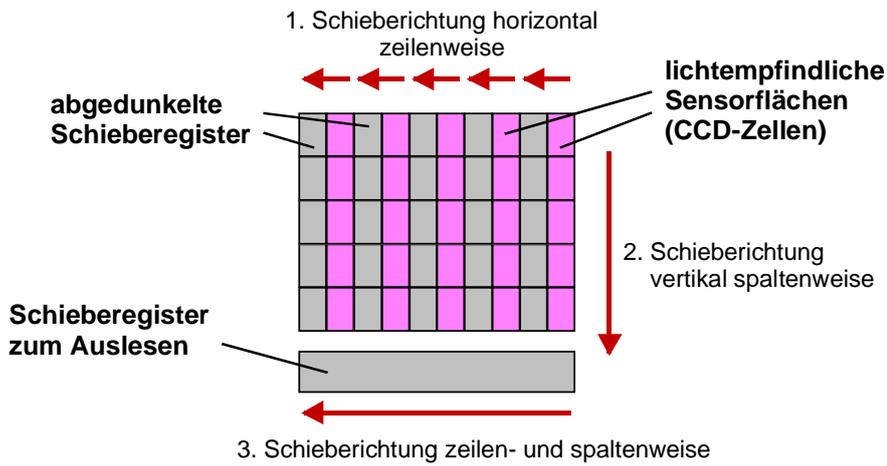


Abb. 1.14 prinzipieller Aufbau des CCD

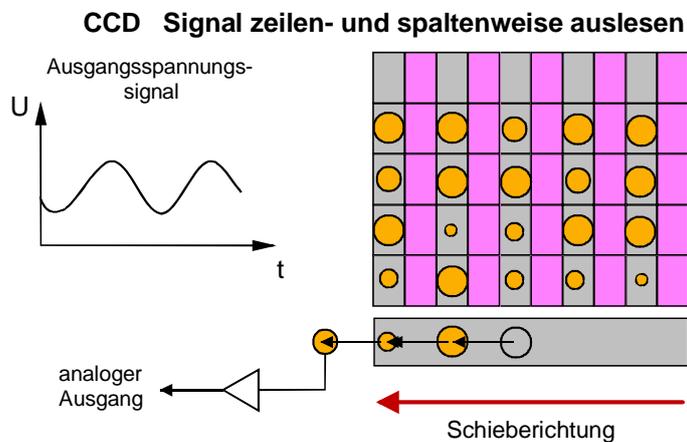
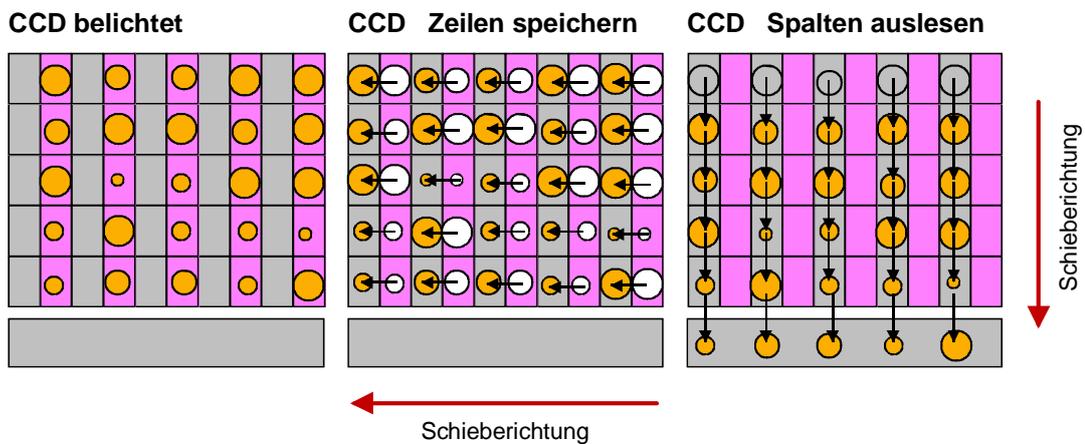


Abb. 1.15 Auslesen der Information

Die in den lichtempfindlichen Bereichen aufgenommenen Ladungen werden parallel in die angrenzenden Schieberegister übernommen. Daraufhin werden die Ladungen in den vertikalen Schieberegistern zeilenweise in das horizontale Ausleseregister geschoben. Von dort werden sie seriell an einem Verstärker in Spannungssignale umgewandelt und ausgelesen.

1.2.1.7 Frame-Transfer-Sensor

Bei Frame-Transfer-Sensoren sind Sensorfläche und Speicherbereich nicht streifenförmig, sondern in 2 Blöcke getrennt angeordnet. Die lichtempfindliche Sensorfläche dient dabei gleichzeitig als Schieberegister. So können die Ladungspakete auf ein Mal nach der Belichtungsphase vom Sensorbereich in den dunklen Speicherbereich verschoben und von dort in gleicher Weise wie beim Interline-Sensor ausgelesen werden. Die jeweils unterste Bildzeile wird ins Ausleseregister übernommen und von dort seriell ausgelesen. Frame-Transfer-Sensoren haben eine bessere Lichtquantenausbeute, sind aber in der Herstellung wesentlich teurer als Interline-Sensoren, da die Gesamtsensorfläche fast doppelt so groß ist.

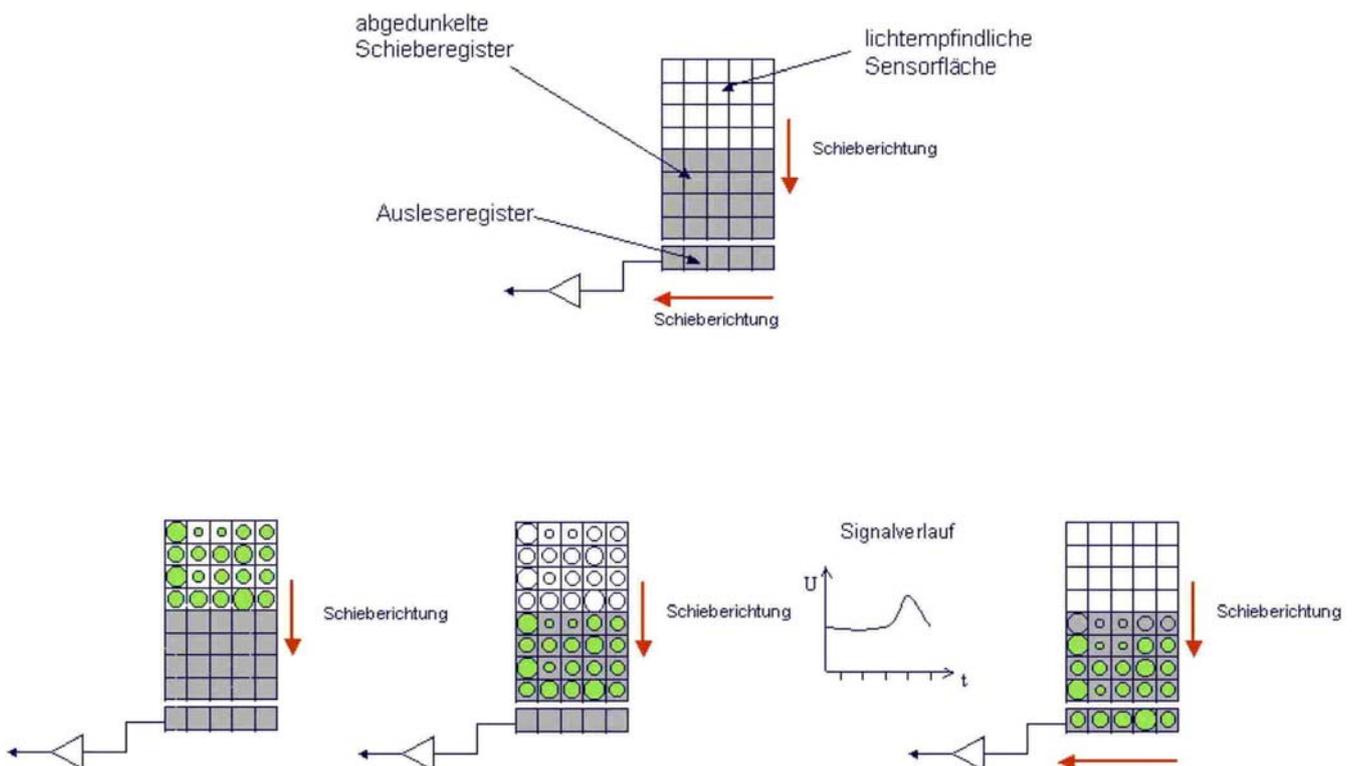


Abb. 1.16 Auslesevorgang beim Frame-Transfer-Sensor

1.2.1.8 Full-Frame-Sensor

Im Gegensatz zu Interline- und Frame-Transfer-CCDs besitzen Full-Frame-CCDs keinen eigenen Speicherbereich. Dafür ist die komplette Sensorfläche lichtempfindlich. Weil die Verschlusszeit von diesem Sensortyp nicht selbst gesteuert werden kann (während des Auslesens würden bei weiterer Belichtung sich noch mehr Ladungen ansammeln und das Bild würde total verschmieren) ist ein mechanischer Verschluss nötig. Der große Vorteil von Full-Frame-CCDs liegt darin, dass die Sensorfläche optimal für den Lichteinfall genutzt wird. Dadurch ist es möglich, Sensoren mit einer hohen Auflösung zu vergleichsweise günstigen Preisen herzustellen.

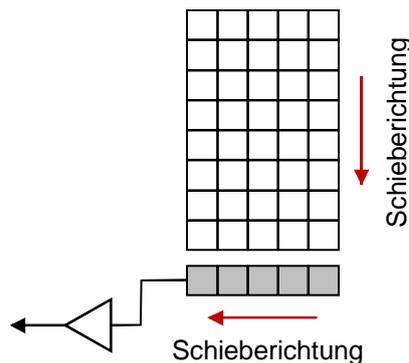


Abb. 1.17 Prinzip des Full-Frame-CCD

1.2.2 CCD-Farbkameras

Um mit einem CCD Element eine Farbkamera zu realisieren, sind optische Systeme erforderlich, mit denen das einfallende Licht in die drei Grundfarben Rot, Grün und Blau zerlegt wird, bevor es auf das CCD fällt.

1.2.2.1 Drei-Chip Kamera

Bei der Drei-Chip-Kamera werden drei CCDs verwendet, von denen je eines für die Detektion von rotem, grünem und blauem Licht zuständig ist. Das Licht wird zu diesem Zweck entweder mit einem Prisma in seine spektralen Bestandteile zerlegt oder fällt durch einen Strahlteiler und drei Farbfilter auf die CCDs. Diese Lösung ist aufwendig und teuer und verlangt neben den drei Chips auch eine exakte Justierung der Chips untereinander, deren Genauigkeit im Bereich der Pixelgröße, also in der Regel bei wenigen μm liegen muss.

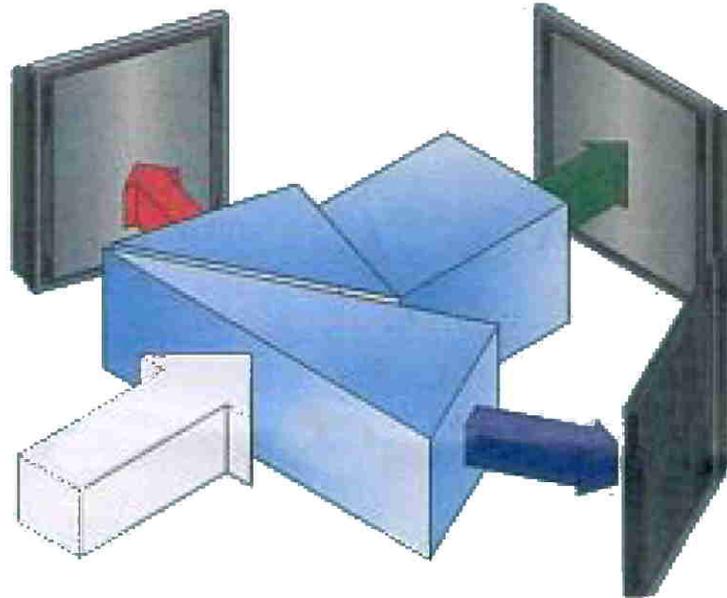


Abb. 1.18 Prinzip der Drei-Chip-Kamera: Zerlegung des einfallenden Lichtes in die drei Farbanteile

1.2.2.2 Ein-Chip Kamera

Für die Massenproduktion sind nur Ein-Chip-Kameras geeignet, bei denen die Farbaufteilung durch Mosaikfilter geschieht, die auf das CCD aufgebracht werden. Dadurch werden rot-, grün- und blauempfindliche Pixel definiert.

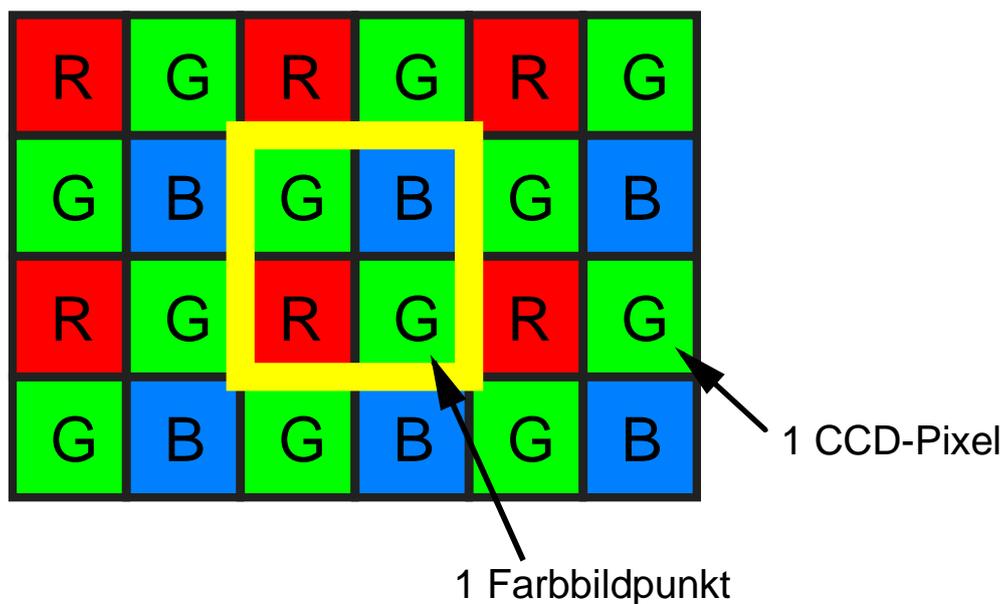


Abb. 1.19 Mosaikfilter für eine Ein-Chip-CCD-Videokamera

Abb. 1.19 zeigt eine mögliche Filteranordnung, bei der ein Farbbildpunkt aus je zwei grünempfindlichen, einem rot- und einem blauempfindlichen CCD-Pixel besteht. Die Anzahl der grünempfindlichen Pixel ist doppelt so groß, um die spektrale Empfindlichkeit der CCD-Kamera auf das menschliche Auge anzupassen, das bei etwa 550 nm, also im Bereich des grünen Lichts, am empfindlichsten ist.

1.2.2.3 Weitere Bauformen

Um von den Mosaikformen gängiger Bildaufnehmer wegzukommen, wurden bereits Systeme für Digitalkameras entwickelt, die aufgrund der unterschiedlichen Eindringtiefen der Photonen die Farbinformation in drei Schritten auslesen können.

Im Vergleich zum vorhergehenden Beispiel kann diese Architektur die Farbkanäle parallel statt sequentiell erfassen.

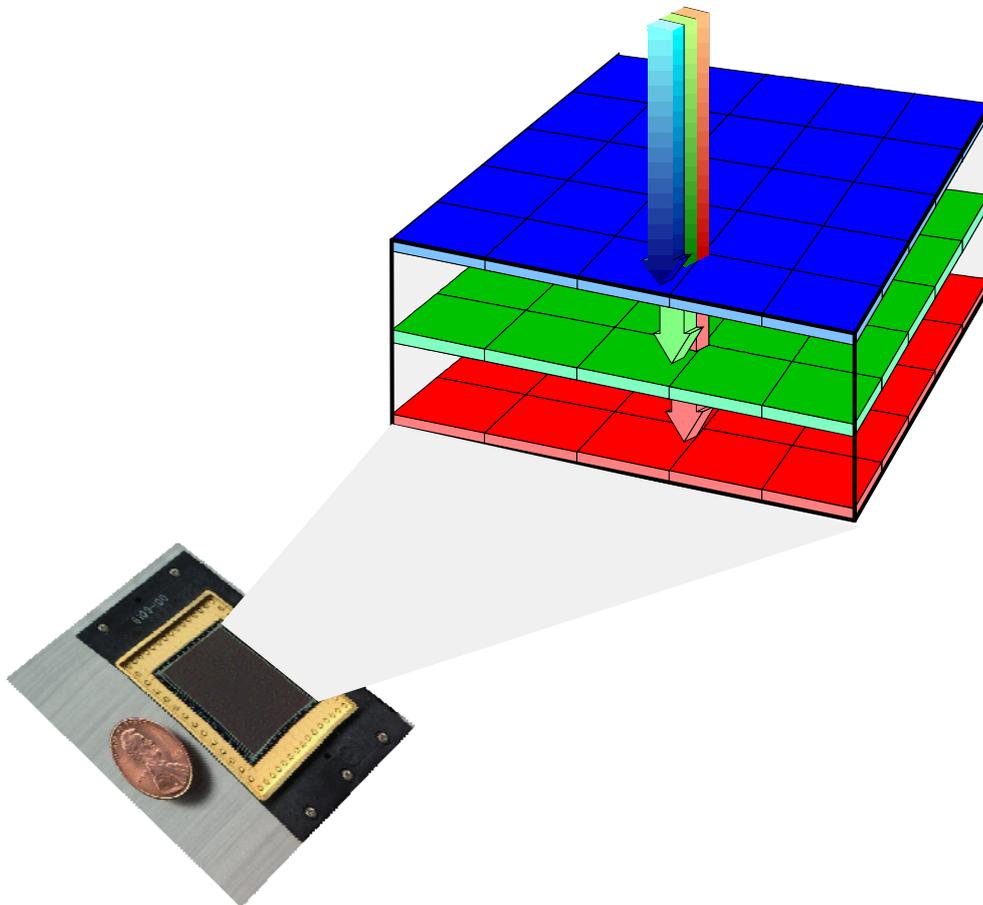


Abb. 1.20 Chip zur parallelen Erfassung der Farbinformation

Das Konzept unterscheidet sich von gängigen CCD-Sensoren und nähert sich dem Prinzip der herkömmlichen Fotografie. Das Licht dringt je nach Wellenlänge unterschiedlich tief in die Halbleiterelemente ein, d.h. längere Wellenlängen werden erst später absorbiert als kürzere. Zudem hat dieser Chip eine dreifache Lichtempfindlichkeit und es ist keine kosten- und zeitraubende Interpolation erforderlich.

Darüber hinaus befindet sich eine Vielzahl von Modifikationen und neuen Sensortypen derzeit im Entwicklungsstadium, so dass an dieser Stelle nur ein kurzer Überblick über CCD-Chips, jedoch keine umfassende Darstellung sämtlicher Typen und Verfahren gegeben werden kann.

1.2.3 Geschaltete Kapazitäten

Das Prinzip der geschalteten Kapazitäten (*switched capacitors*) ist schon sehr lange bekannt. In Verbindung mit MOS-IC's ist es jedoch besonders interessant geworden [1].

Bei den geschalteten Kapazitäten wird die Spannungsabhängigkeit der MOS-Kapazität explizit nicht ausgenutzt. Sinnvollerweise wird der Kondensator nur im **Anreicherungsmodus** betrieben.

1.2.3.1 Wirkungsweise

Ein Widerstand R kann durch eine geschaltete Kapazität (Abb. 1.20) ersetzt werden.

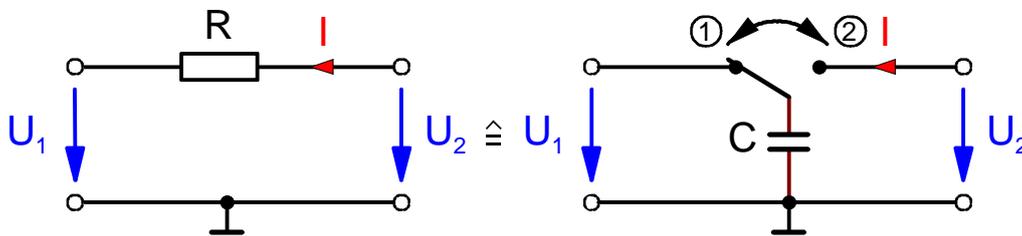


Abb. 1.21 Methode der "geschalteten Kapazität" (*switched capacitor*)

Prinzip:

In Schalterposition 1 wird C auf die Spannung U_1 aufgeladen.

In Schalterposition 2 wird C auf die Spannung U_2 umgeladen.

Dazu ist eine Änderung der Ladung vom Betrag

$$\Delta Q = C \cdot (U_2 - U_1) \quad (1.4)$$

erforderlich. Wird der Schalter S mit der Frequenz f_c hin- und hergeschaltet, fließt im Mittel ein Strom

$$I = \Delta Q \cdot f_c = C \cdot f_c \cdot (U_2 - U_1) . \quad (1.5)$$

Der Vergleich mit dem Strom durch das Widerstandsnetzwerk

$$I = \frac{U_2 - U_1}{R} \quad (1.6)$$

zeigt, dass die "switched-capacitor"-Anordnung einen äquivalenten Widerstand

$$R_{\text{äq}} = \frac{1}{C \cdot f_c} \quad (1.7)$$

aufweist. Voraussetzung ist dabei die Bedingung

$$f_c \gg f_{\text{max}} \quad (1.8)$$

mit f_{max} als größte zu verarbeitende Frequenz.

1.2.3.2 Anwendungsbeispiele

1.2.3.2.1 Filter

Ein Tiefpassfilter 1.Ordnung mit dem in Abb. 1.22 gezeigten Frequenzgang lässt sich entsprechend Abb. 1.23a aus einem Widerstand und einem Kondensator realisieren.

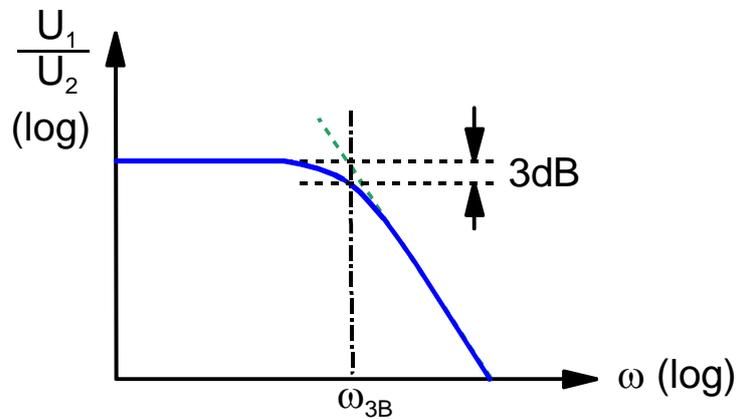


Abb. 1.22 Frequenzgang eines Tiefpassfilters 1.Ordnung

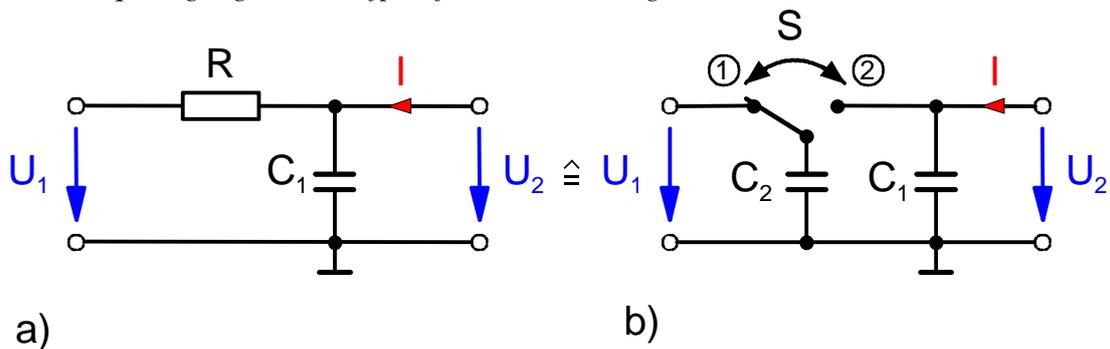


Abb. 1.23 Realisierung des Tiefpasses 1.Ordnung mit dem Frequenzgang aus Abb. 1.22
 a) konventionelle Lösung
 b) Lösung mit geschalteter Kapazität

Die 3dB-Grenzfrequenz des Filters ist bestimmt durch

$$\omega_{3dB} = \frac{1}{R C_1} \tag{1.9}$$

Der Widerstand R kann durch eine geschaltete Kapazität ersetzt werden (Abb. 1.23b).

Mit den Gln. (1.6) und (1.9) folgt

$$R = \frac{1}{C_2 f_c} \quad \text{und} \quad \omega_{3dB} = f_c \frac{C_2}{C_1}; \quad f_c \gg \omega_{3dB} \tag{1.10}$$

Vorteile der Lösung mit geschalteten Kapazitäten:

1. ω_{3dB} ist über f_c einstellbar
2. Je größer der gewünschte Wert des Widerstandes R , desto **kleiner** ist die für C_2 benötigte Fläche in einer MOS-Schaltung (Gl. (1.9)).

Zahlenbeispiel:

Filter im Audiobereich:

$$\omega_{3dB} = 10^4 \text{ Hz} \Rightarrow R = 10 \text{ M}\Omega, C_1 = 10 \text{ pF}$$

10 M Ω sind in MOS-Schaltungen praktisch nicht realisierbar.

Mit $f_c = 10^5 \text{ Hz} \gg \omega_{3dB}$ folgt

$$C_2 = \frac{1}{R f_c} = \frac{1}{10 \text{ M}\Omega \cdot 10^5 \text{ Hz}} = 1 \text{ pF}$$

1pF ist in MOS-Technik leicht realisierbar.

3. Fertigungstoleranzen gehen nicht ein, solange sie die Werte C_1, C_2 **in gleichem Maße** beeinflussen

Beispiel:

Variation der Oxiddicke um 10%

Dann ist $C'_1 = 1,1 \cdot C_1$

und $C'_2 = 1,1 \cdot C_2$

damit also $\frac{C'_2}{C'_1} = \frac{C_2}{C_1}$

Somit bleibt z.B. in diesem Fall die Grenzfrequenz trotz Fertigungstoleranzen konstant.

1.2.3.2 Integrator mit Operationsverstärkern

Abb. 1.24a zeigt einen Integrator, bestehend aus einem Operationsverstärker, einem Eingangswiderstand R_1 und einem Rückkopplungskondensator C_2 .

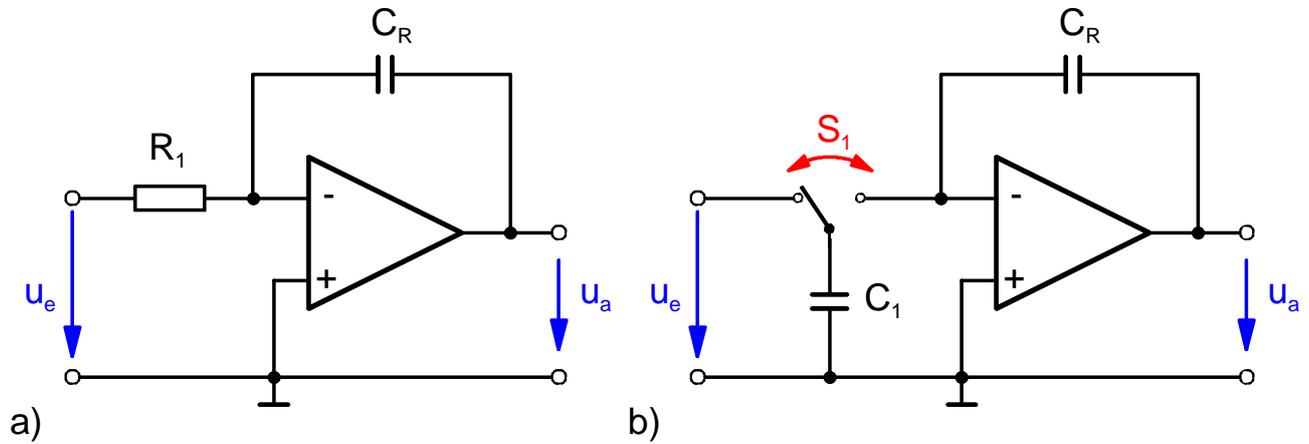


Abb. 1.24 Integrator mit Operationsverstärker
 a) konventionelle Lösung
 b) Lösung mit geschalteter Kapazität

Die Ausgangsspannung ist

$$U_a(t) = - \int_0^t K \cdot u_e dt + u_a(0) \quad \text{mit} \quad K = \frac{1}{R_1 C_2} \quad (1.11)$$

Der Ersatz des Widerstandes R_1 durch eine geschaltete Kapazität C_1 ergibt (Abb. 1.24b):

$$K = f_c \cdot \frac{C_1}{C_2} \quad (1.12)$$

Durch Variation der Frequenz f_c lässt sich dieser Faktor entsprechend einstellen.

1.2.3.3 Realisierung in MOS-Technik

Der Kondensator C wird in MOS-Technik genauso wie für CCD realisiert. Der Schalter S kann durch zwei MOSFET ersetzt werden (Abb. 1.25a), die gegenphasig geschaltet werden. Als Beispiel seien selbstleitende p-Kanal MOSFET verwendet. Die erforderlichen Spannungsverläufe (die von einem Taktgenerator zu liefern sind) an G_1 und G_2 zeigt Abb. 1.25b. Der MOSFET-Schalter arbeitet jedoch nicht verlustfrei.

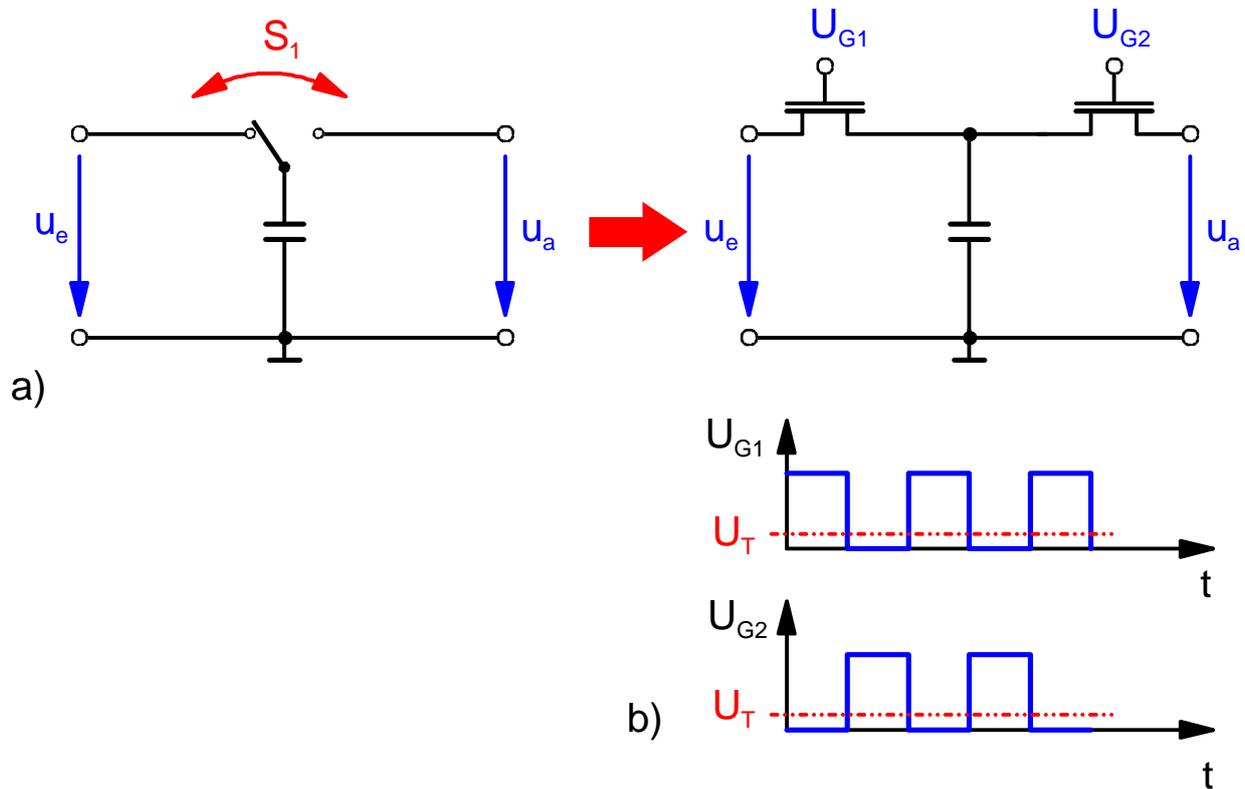


Abb. 1.25 a) Realisierung des "switched capacitors" in MOS-Technik (hier z.B. mit selbstleitenden MOSFET)
b) zugehöriges Taktdiagramm

Stichwortverzeichnis:

Anreicherung 1-2

CCD

 Drei-Phasen-CCD 1-8

 Fotodetektoren-Aufbau 1-10

 Fotodiode 1-12

 Foto-Gate 1-13

 Sensortypen 1-14

 Speicherungsart 1-5

 Signaltransport 1-6

 Zwei-Phasen-CCD 1-10

CCD-Farbkameras 1-16

 Drei-Chip Kamera 1-16

 Ein-Chip Kamera 1-17

Flachbandfall 1-2

Fotogeneration 1-5

Frame-Transfer-Sensor 1-15

Full-Frame-Sensor 1-16

Geschaltete Kapazitäten

 Filter 1-20

 Integrator mit Operationsverstärker 1-22

 Realisierung in MOS-Technik 1-23

 Wirkungsweise 1-19

Interline-Transfer-Sensor 1-14

Inversion 1-2

Kapazitäten

 Geschaltete Kapazitäten 1-19

MOS

 Kondensator 1-2

MOS-Kondensator 1-2

 Anreicherung 1-2

 Anwendungen 1-5

 CCD 1-5

 CV-Kennlinie 1-3

 Flachbandfall 1-2

 Inversion 1-2

MOS-Struktur 1-3

 Verarmung 1-2

Sensortypen

 Frame-Transfer-Sensor 1-15

 Full-Frame-Sensor 1-16

 Interline-Transfer-Sensor 1-14

Verarmung 1-2

Kapitel 2

Feldeffekttransistoren

Inhalt

2	Feldeffekttransistoren	2-2
2.1	MOS-Feldeffekttransistoren	2-4
2.1.1	Aufbau und Wirkungsweise.....	2-4
2.1.2	Berechnung des Drain-Stroms	2-6
2.1.2.1	Der MOSFET als steuerbarer Widerstand	2-9
2.1.2.2	Der MOSFET im Sättigungsgebiet	2-10
2.1.3	Die Kennlinienfelder des MOSFET.....	2-15
2.1.3.1	Die Kennlinienfelder der Source-Schaltung	2-16
2.1.3.1.1	Kennlinien der p-Kanal-Anreicherungstypen (Enhancement-FET, E-FET)	2-17
2.1.3.1.2	Kennlinien der n- und p-Kanal-Verarmungstypen (Depletion-FET, D-FET)	2-19
2.1.4	Temperaturabhängigkeit der Kennlinien	2-20
2.1.5	Einfluss der Substratvorspannung.....	2-20
2.1.6	Arbeitspunkteinstellung und -stabilisierung	2-22
2.1.6.1	n-Kanal-Anreicherungs-MOSFET (E-MOSFET)	2-22
2.1.6.2	n-Kanal-Verarmungs-MOSFET (D-MOSFET)	2-23
2.1.7	Komplementäre MOS-Technik (CMOS).....	2-24
2.2	Sperrschicht-Feldeffekttransistoren	2-28
2.2.1	Aufbau und Wirkungsweise.....	2-29
2.2.2	Kennlinienberechnung	2-32
2.3	Der Heterostruktur-Feldeffekttransistor (HFET)	2-37
2.3.1	Der Heteroübergang.....	2-37
2.3.2	Der Ladungstransport im 2DEG	2-39
2.3.3	Der Aufbau eines Heterostruktur-Feldeffekttransistors.....	2-40
2.3.4	Berechnung der Schichtkonzentration	2-42
2.3.4.1	Der Gleichgewichtsfall	2-43
2.3.4.2	Der Einfluss des Schottky-Kontaktes auf die Schichtkonzentration	2-47
2.3.4.3	Gleichstromverhalten	2-48
2.4	Das Elektrische Ersatzschaltbild	2-49
2.4.1	Das Gleichstrom- (DC-) verhalten des realen Feldeffekttransistors.....	2-50
2.4.2	Das Wechsellspannungs- (AC-) verhalten.....	2-53
2.4.2.1	Funktionsprinzip	2-53
2.4.2.2	Allgemeine Beschreibung des Kleinsignalverhaltens	2-54
2.4.2.3	Quasistationäres Kleinsignalverhalten von Feldeffekttransistoren	2-56

2 Feldeffekttransistoren

Allgemein können Halbleiterbauelemente unipolare oder bipolare Bauelemente sein.

Bipolare Bauelemente:

Majoritäts- und Minoritätsladungsträger sind bestimmen das Bauelementverhalten.

Steuermechanismus: **Injektion** und **Rekombination**.

Unipolare Bauelemente:

Vorwiegend Majoritätsladungsträger bestimmen das Bauelementverhalten.

=> **Feldeffekttransistoren**

Das Prinzip des Feldeffekttransistors zeigt Abb. 2.1.

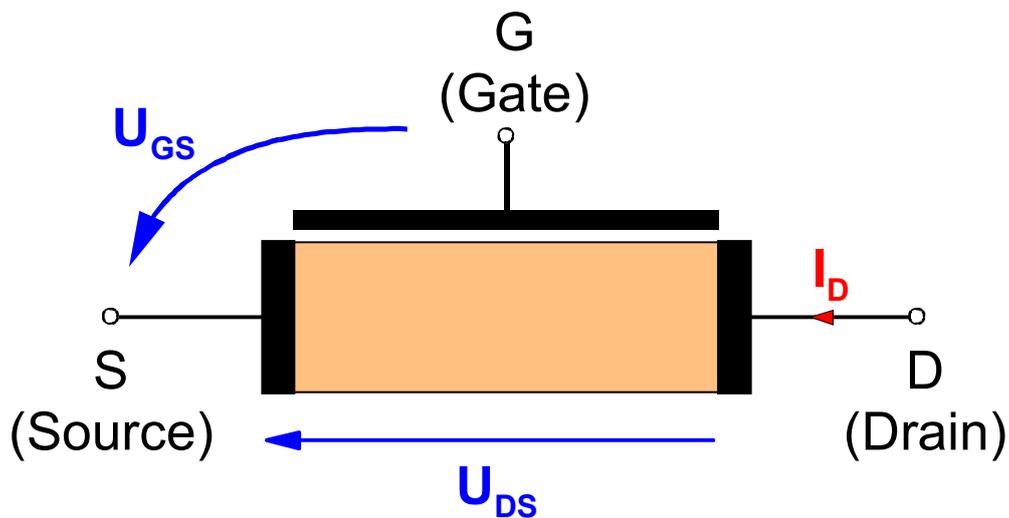


Abb. 2.1 Prinzip des Feldeffekttransistors:

Steuerung des Stromes $I_D = f(U_{DS})$ in einem Halbleiter durch eine Steuerelektrode (Gate)

Der von Majoritäten getragene Strom I_D zwischen den Anschlüssen S (Source) und D (Drain) wird über eine dritte Elektrode G (Gate) gesteuert.

Dabei kann "Steuerung" bedeuten:

- Schalten des Gleichstromes I_D → Anwendung in **Digital**schaltungen
- Verstärken eines (Wechsel-)Signals → Anwendung in **Analog**schaltungen

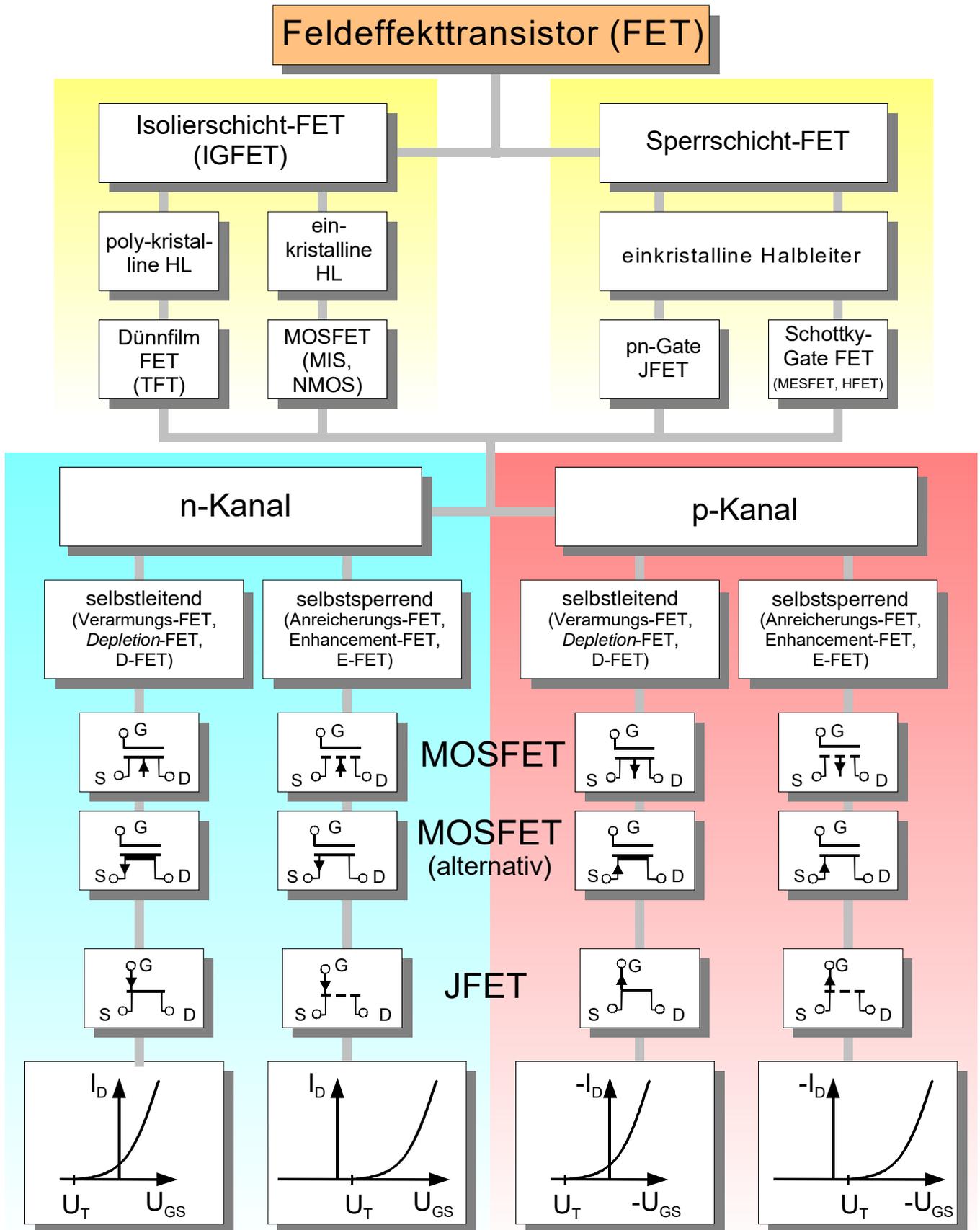


Abb. 2.2 Übersicht über die möglichen Typen von Feldeffekttransistoren

Steuerungsmöglichkeiten der Ladungsträger im Kanal:

1. Die Ladung $n \cdot q$ wird durch Influenzwirkung geändert. Dazu wird der MOS-Kondensator (s. Kap.1) im **Anreicherungs-** oder **Inversionsmodus** eingesetzt. Dies ist der eigentliche **Feldeffekt** => **MOS-Feldeffekt-Transistor (MOSFET)**.
2. Der Querschnitt A des stromführenden Kanals wird durch eine Raumladungszone eingeengt, deren Ausdehnung über das Gate beeinflusst wird. Die Raumladungszone kann von einem **pn-Übergang** oder von einem **Schottky-Kontakt** (Metall-Halbleiter-Kontakt) herrühren. Auch diese Steuermöglichkeit wird Feldeffekt genannt.

=> **Sperrschichtfeldeffekttransistoren** (s. Kap. 2.2)

- FET mit pn-Übergang als Gate => **Junction-FET (JFET)**
- FET mit Schottky-Kontakt als Gate => **Metal-Semiconductor-FET (MESFET)**

Aufgrund dieser beiden Steuermöglichkeiten und unter Verwendung der Kenntnisse aus Kapitel 1 über MOS-Kondensatoren folgt der "Stammbaum" der Feldeffekttransistoren (FET) (Abb. 2.2).

2.1 MOS-Feldeffekttransistoren

Kommerziell erhältliche MOSFETs und integrierte MOS-Schaltungen basieren ausschließlich auf dem Halbleitermaterial *Silizium*. Als Isolierschicht wird überwiegend Siliziumdioxid verwendet. SiO_2 (Quarz) ist ein leicht herstellbarer, chemisch sehr resistenter Werkstoff. Andere Isolatoren können auch Si_3N_4 , Al_2O_3 , sowie neuerdings TaN, ZrO_2 oder HfO_2 sein.

Aufbau, Wirkungsweise und Kennlinienberechnung werden am Beispiel eines n-Kanal-Anreicherungs-MOSFET erläutert.

2.1.1 Aufbau und Wirkungsweise

Der Aufbau eines n-Kanal-MOSFET für Anreicherungsbetrieb ist in Abb. 2.3 gezeigt:

Ein niedrig p-dotiertes **Substrat** (Silizium) dient als mechanischer Träger des MOSFET. Die Substratoberfläche ist mit einem SiO_2 -Film (Dicke ≈ 100 nm) überzogen. Durch Öffnungen im SiO_2 werden hoch n-dotierte (n^+) Bereiche diffundiert oder implantiert, die **Source** (S) und **Drain** (D) bilden. Die Anschlüsse werden über aufgedampfte und einlegierte **Metallkontakte** hergestellt. Eine Metallschicht auf dem Isolator zwischen S und D bildet die **Gate-Elektrode**.

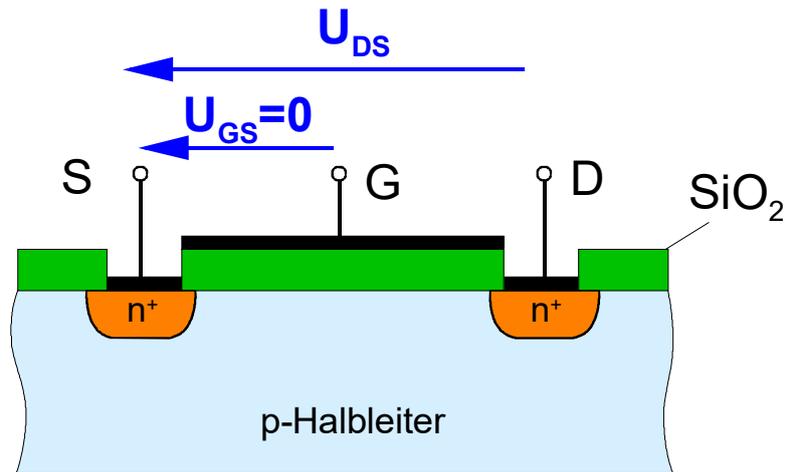


Abb. 2.3 Querschnitt durch einen MOSFET bei $U_{GS} = 0$

Die **Wirkungsweise** des n-MOSFET aus Abb. 2.3 ist folgendermaßen zu verstehen:

Bei Gate-Spannungen $U_{GS} = 0$ ist kein n-leitender Kanal vorhanden. Mindestens einer der beiden n⁺p-Übergänge S oder D ist stets gesperrt. Es fließt praktisch kein Strom I_D . Bei realen Bauelementen tritt jedoch ein kleiner Sperrstrom auf.

Bei positiver Gate-Spannung $U_{GS} > 0$ kann negative Ladung im Kanal influenziert werden. Wenn U_{GS} groß genug ist, entsteht **Inversion** und damit ein n-leitender Kanal zwischen S und D. Die Größe der Inversionsladung und damit die Größe des Stromes I_D wachsen mit zunehmender Gate-Spannung an.

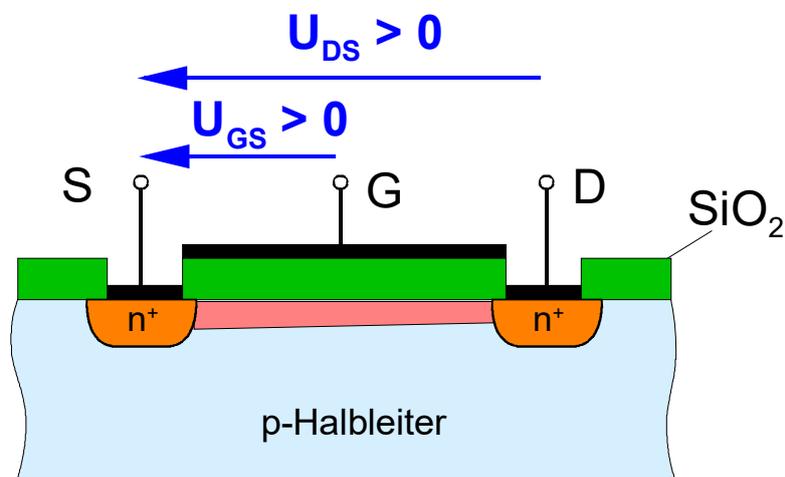


Abb. 2.4 Querschnitt durch einen MOSFET bei $U_{GS} > 0$

Die **Kanalladung** ist nicht gleichmäßig über die Kanallänge L_G verteilt. Sie hängt vom Spannungsabfall über dem Oxid ab. Da das Potenzial φ im Kanal nicht konstant ist, sondern vom Wert $\varphi = 0$ am Source auf den Wert $\varphi = U_{DS}$ am Drain ansteigt, nimmt der Spannungsabfall über dem Oxid vom Source zum Drain hin ab. Wenn $U_{DS} > U_{GS}$ wird, wird in Teilen des Kanals überhaupt keine Inversion mehr erzeugt. Dieser Effekt führt letztlich zur Sättigung der $I_D(U_{DS})$ -Kennlinien.

2.1.2 Berechnung des Drain-Stroms

Bei den folgenden Berechnungen wird ein **p-dotiertes Substrat** vorausgesetzt. Dementsprechend liegt ein **n-leitender Kanal** vor. Für einen p-Kanal MOSFET lassen sich die Berechnungen jedoch analog durchführen.

Für die weiteren Berechnungen werden folgende **Annahmen** gemacht:

1. $E_{y,ox} \gg E_x$

Die Feldstärke im Oxid senkrecht zum Stromfluß im Kanal $E_{y,ox}$ ist dem Betrage nach wesentlich größer als die durch U_{DS} erzeugte Feldstärke E_x im Kanal parallel zum Stromfluss. Das bedeutet, dass die Feldstärke im Oxid praktisch senkrecht zum Kanal und die Feldstärke im Kanal parallel zum Kanal ist, so dass eindimensional gerechnet werden kann.

Diese Voraussetzung wird auch als "*SHOCKLEY'sche gradual channel approximation*" bezeichnet [2].

2. $\mu = konst.$

Die Beweglichkeit der Ladungsträger im Kanal ist konstant und vom Feld E_x unabhängig.

Für den Strom I_D im Kanal gilt generell:

$$I_D = q \cdot n \cdot \left| \vec{v} \right| \cdot A \quad (2.1)$$

n : Konzentration der im Kanal zwischen S und D frei beweglichen Ladungsträger (hier: Elektronen) ($[n] = \text{cm}^{-3}$)

$\left| \vec{v} \right|$: Geschwindigkeit, mit der sich die Ladungsträger bewegen ($[v] = \text{cm/s}$)

A : stromführender Querschnitt des Kanals ($[A] = \text{cm}^2$)

Die Geschwindigkeit \vec{v} wird durch die Feldstärke im Kanal bestimmt. Im einfachsten Fall gilt:

$$\vec{v} = \mu \cdot \vec{E}$$

bzw. im eindimensionalen Fall:

$$v = -\mu \cdot E_x \quad \mu: \text{Beweglichkeit der Ladungsträger} \quad (2.2)$$

Dabei ist das negative Vorzeichen durch die negative Ladung der Elektronen begründet.

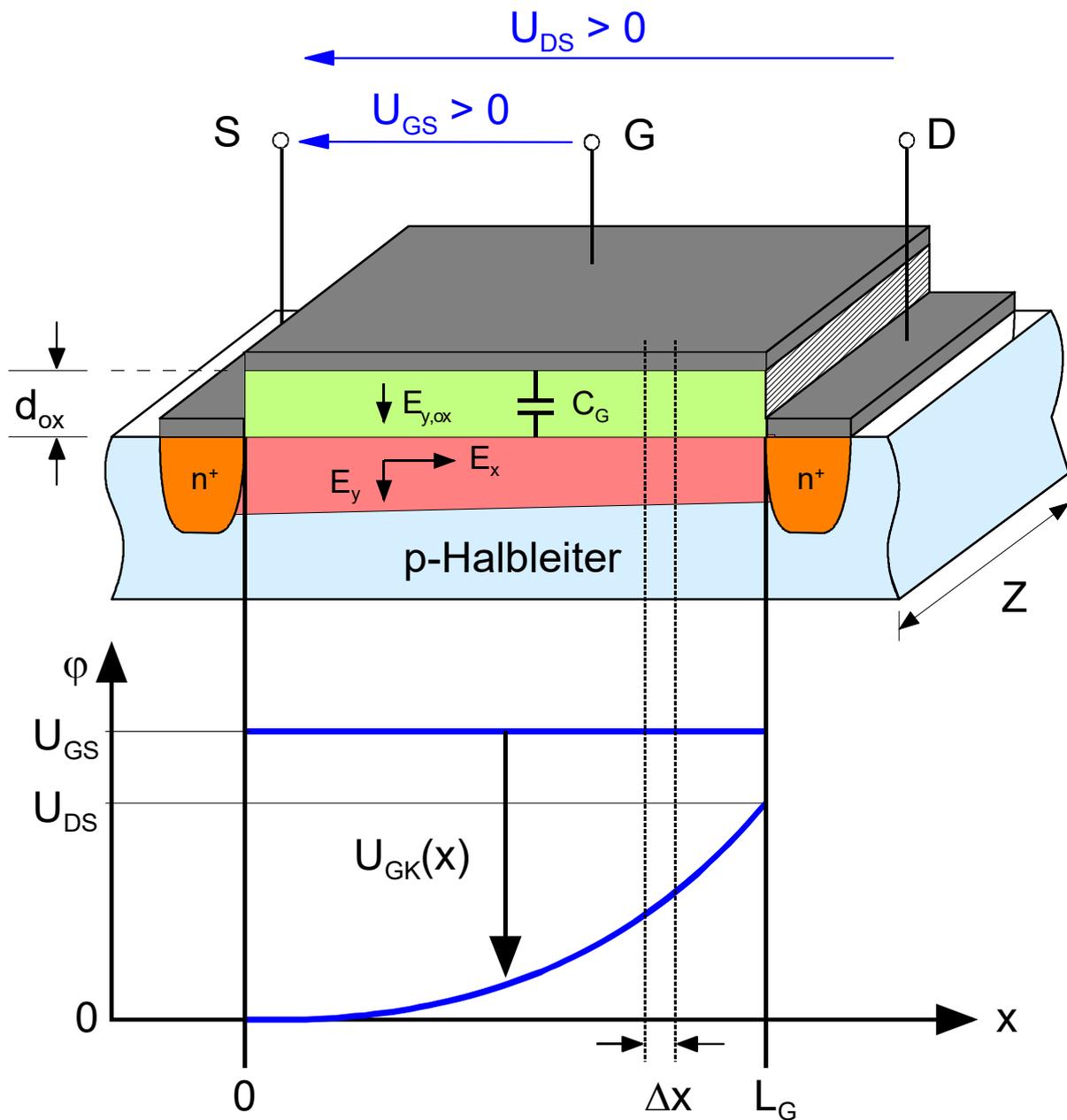


Abb. 2.5 Potenzialverteilung im Kanal des MOSFET

Mit :

$$E_x = -\frac{d\varphi}{dx} \quad \varphi : \text{Potenzial im Kanal} \quad (2.3)$$

folgt:

$$v = \mu \cdot \frac{d\varphi}{dx} \quad (2.4)$$

Die Konzentration der frei beweglichen Ladungsträger (n) und die Querschnittsfläche des Kanals (A) lässt sich über die Kapazität des MOSFET ermitteln:

Für die Gesamtkapazität des Gate-Kontaktes (Oxidkapazität) C_G gilt:

$$C_G = \frac{\varepsilon_o \cdot \varepsilon_r \cdot L_G \cdot Z}{d_{ox}} \quad Z: \text{ Gate-Weite} \quad (2.5)$$

Die an der Stelle x im Kanal in einem Volumenelement

$$\Delta V = A \cdot \Delta x \quad (2.6)$$

vorhandene bewegliche Ladung $q \cdot n(x)$ entspricht aufgrund des Ladungsgleichgewichtes an der Grenze zwischen Oxid und Halbleiter der in der entsprechenden Teilkapazität

$$\Delta C_G = C_G \cdot \frac{\Delta x}{L_G} \quad (2.7)$$

gespeicherten Ladung.

Damit folgt:

$$q \cdot n(x) = \frac{\Delta C_G}{\Delta V} \cdot U_{GK}(x) = \frac{\Delta C_G}{\Delta x \cdot A(x)} \cdot U_{GK}(x) = \frac{C_G}{L_G \cdot A(x)} \cdot U_{GK}(x) \quad (2.8)$$

Der Spannungsabfall $U_{GK}(x)$ über dem Oxid ergibt sich entsprechend Abb. 2.5 aus einem Maschenumlauf mit:

$$U_{GK}(x) = U_{GS} - \varphi(x) \quad (2.9)$$

mit

$$\varphi(x=0) = 0 \quad (2.10)$$

und

$$\varphi(x=L_G) = U_{DS} \quad (2.11)$$

Mit (2.9) folgt damit:

$$q \cdot n(x) = \frac{C_G}{L_G \cdot A(x)} \cdot (U_{GS} - \varphi(x)) \quad (2.12)$$

Ein Teil der Spannung U_{GS} wird jedoch benötigt, um

- ortsfeste Ladungen Q_{ox} im Oxid und an der Grenzfläche Oxid-Halbleiter umzuladen;
- Differenzen in den Austrittsarbeiten *Metall-Oxid* und *Oxid-Halbleiter* zu kompensieren und somit den Kanal durch Inversion im Substratmaterial erst zu erzeugen.

Diese Spannung wird als **Schwellesspannung** U_T (engl.: *threshold voltage*) bezeichnet. Das bedeutet aber auch, dass für Spannungen $U_{GS} < U_T$ kein Kanal existiert.

Damit folgt für die frei beweglichen Ladungen im Kanal:

$$q \cdot n(x) = \frac{C_G}{L_G \cdot A(x)} (U_{GS} - U_T - \varphi(x)) \quad (2.13)$$

Für den Drain-Strom I_D folgt mit Gl. (2.1), Gl. (2.4) und Gl. (2.13) an jeder Stelle x :

$$I_D(x) = \frac{\mu \cdot C_G}{L_G} [U_{GS} - U_T - \varphi(x)] \frac{d\varphi}{dx} \quad (2.14)$$

Integration in den Grenzen von $x=0$ bis $x=L_G$ mit den Randbedingungen $\varphi(x=0) = 0$ und $\varphi(x=L_G) = U_{DS}$ (Gl. (2.10), (2.11)) liefert:

$$\int_0^{L_G} I_D(x) dx = \frac{\mu \cdot C_G}{L_G} \int_0^{U_{DS}} [U_{GS} - U_T - \varphi(x)] d\varphi \quad (2.15)$$

Da aufgrund der Stromkontinuität an jeder Stelle im Kanal der gleiche Strom fließen muss und somit I_D nicht von x abhängt, folgt:

$$I_D = \frac{\mu \cdot C_G}{L_G^2} \left[(U_{GS} - U_T) \cdot U_{DS} - \frac{1}{2} U_{DS}^2 \right] \quad (2.16a)$$

oder

$$I_D = 2K \left[(U_{GS} - U_T) U_{DS} - \frac{1}{2} U_{DS}^2 \right] \quad (2.16b)$$

mit

$$2K = \frac{\mu \cdot C_G}{L_G^2} = \frac{\varepsilon_o \cdot \varepsilon_{r,ox} \cdot \mu}{d_{ox}} \cdot \frac{Z}{L_G} \quad (2.17)$$

Dabei wird K als "*Entwurfsparameter*" und $\frac{Z}{L_G}$ als "*aspect ratio*" bezeichnet.

2.1.2.1 Der MOSFET als steuerbarer Widerstand

Zunächst wird der Fall:

$$U_{DS} \ll (U_{GS} - U_T)$$

betrachtet. Damit gilt auch:

$$U_{DS}^2 \ll (U_{GS} - U_T) \cdot U_{DS}$$

so dass aus Gl. (2.16b) folgt:

$$I_D = 2K(U_{GS} - U_T) \cdot U_{DS} \quad \text{für } U_{DS} \ll (U_{GS} - U_T) \quad (2.18)$$

Damit ergibt sich eine lineare Änderung von I_D mit U_{DS} entsprechend:

$$I_D = \frac{1}{R} U_{DS}$$

mit

$$R = \frac{1}{2K(U_{GS} - U_T)} \tag{2.19}$$

Für sehr kleine U_{DS} -Werte verhält sich der MOSFET also wie ein ohmscher Widerstand R , dessen Größe entsprechend Gl. (2.19) durch U_{GS} gesteuert werden kann.

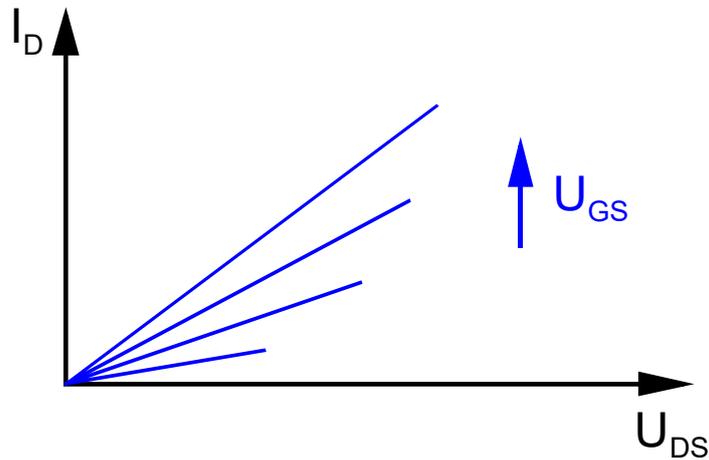


Abb. 2.6 Abhängigkeit des Drain-Stromes I_D von U_{DS} und U_{GS} für den Fall $U_{DS} \ll (U_{GS} - U_T)$

2.1.2.2 Der MOSFET im Sättigungsgebiet

Die Kennlinien realer MOSFET folgen Gl. (2.16) nur bis zu einem Maximalwert abhängig von der jeweiligen Gate-Source-Spannung U_{GS} (Abb. 2.7). Danach bleibt der Drain-Strom näherungsweise unabhängig von U_{DS} bzw. steigt nur noch schwach mit U_{DS} an.

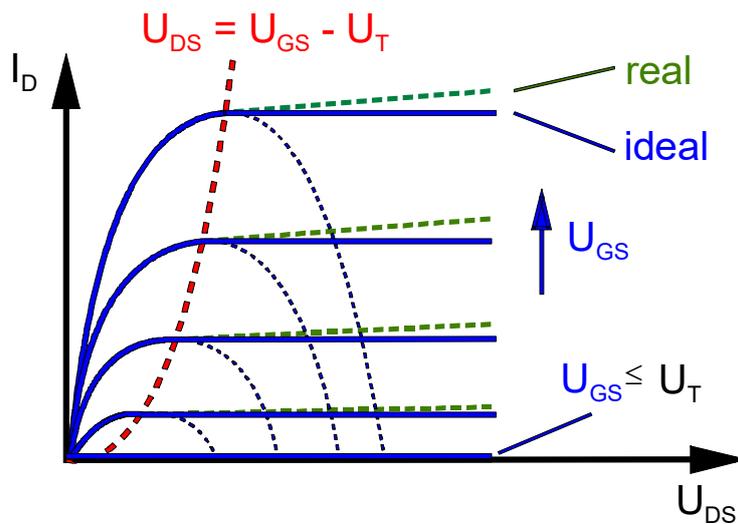


Abb. 2.7 Abhängigkeit des Drain-Stromes I_D von U_{DS} und U_{GS} entsprechend der vollständigen Formel nach Gl. (2.16)

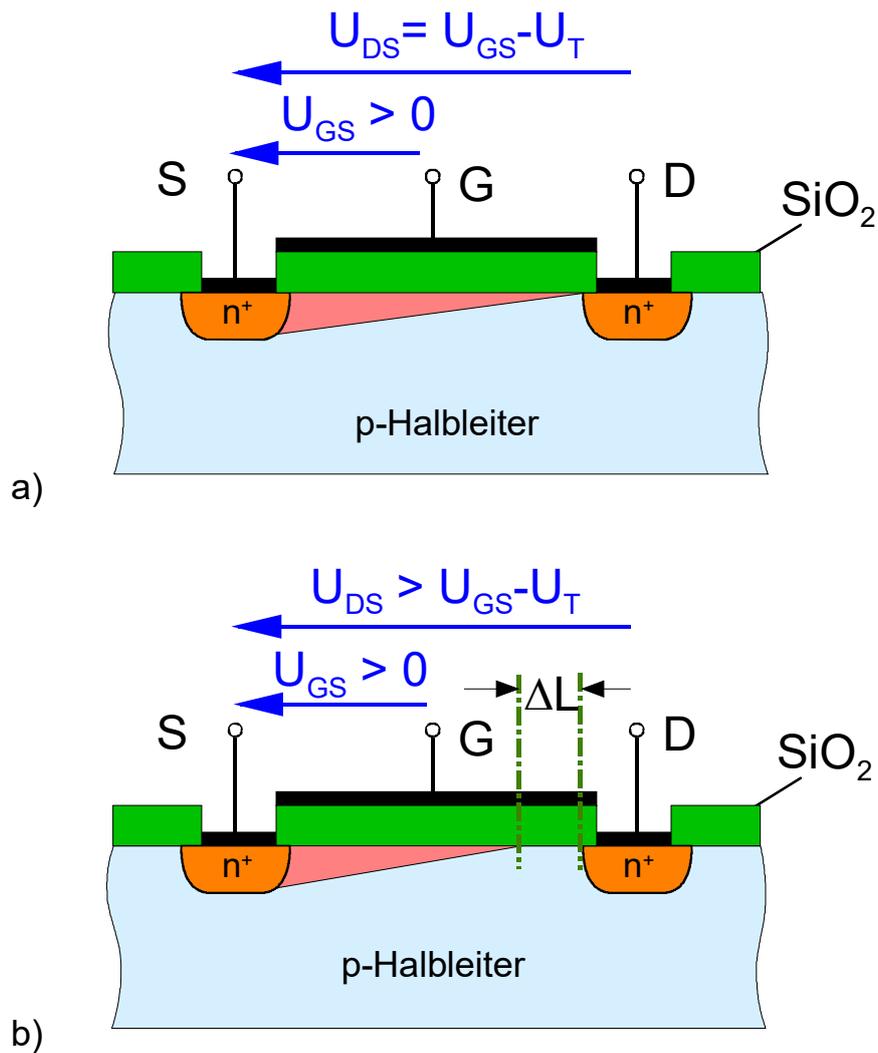


Abb. 2.8 Querschnitt durch den MOSFET für

a) $U_{DS} = U_{GS} - U_T$ und

b) $U_{DS} > U_{GS} - U_T$

Die Spannung $U_{DS,sat}$, bei der der maximale Sättigungsstrom $I_{D,sat}$ erreicht wird, folgt aus:

$$\frac{dI_D}{dU_{DS}} \stackrel{!}{=} 0 = 2K[(U_{GS} - U_T) - U_{DS}]$$

oder

$$U_{DS,sat} = U_{GS} - U_T \quad (2.20)$$

Einsetzen in Gl. (2.16b) ergibt:

$$I_{D,sat} = K \cdot (U_{GS} - U_T)^2 \quad (2.21)$$

Für $U_{DS} > U_{DS,sat} = U_{GS} - U_T$ gilt Gl. (2.16) nicht mehr!

Wenn:

$$U_{DS} = U_{GS} - U_T$$

ist, dann reicht der Potenzialabfall über dem Oxid an der Stelle $x = L_G$, d.h. am Drain-Kontakt, nicht mehr aus, um bewegliche Ladungsträger zu influenzieren. Der leitende Kanal verschwindet damit an der Stelle $x = L_G$ und es liegt **Kanalabschnürung (pinch-off)** vor (Abb. 2.8a)

Wenn

$$U_{DS} > U_{GS} - U_T$$

wird, so kehrt sich das Vorzeichen über dem Oxid um und statt Inversion entsteht nun Verarmung. Dies führt zur **Kanalverkürzung**.

Diese Verarmungszone ist eine hochohmige Raumladungszone, in der eine hohe Feldstärke E_x auftritt. Sie ist ähnlich der Raumladungszone in einer gesperrten Basis-Kollektor-Diode eines Bipolartransistors. So wie die vom Emitter in die Basis injizierten Ladungsträger die Kollektor-Raumladungszone durchqueren, durchqueren die vom Kanal "injizierten" Ladungsträger die hochohmige Zone im MOSFET, so dass der Strom I_D nicht absinkt, wenn $U_{DS} > U_{GS} - U_T$ gilt.

Da der Kanal mit wachsendem U_{DS} verkürzt wird, sinkt sein Widerstand, so dass I_D für Spannungen U_{DS} oberhalb von $U_{DS,sat}$ noch anwächst.

Die Ausdehnung ΔL der Raumladungszone ist [13]:

$$\Delta L \propto \sqrt{\Delta U_{DS}}$$

mit dem Spannungsabfall über ΔL

$$\Delta U_{DS} = U_{DS} - U_{DS,sat} \quad (2.22)$$

Bei kleinen Gate-Längen und den damit verbundenen hohen Feldstärken im Kanal ist die Driftgeschwindigkeit v der Ladungsträger nicht mehr proportional zum elektrischen Feld, d.h. die Beweglichkeit μ ist nicht mehr konstant. Die Geschwindigkeit wächst mit zunehmendem elektrischen Feld immer langsamer an bis sie schließlich von diesem unabhängig wird und einen Sättigungswert erreicht (s. Abb. 2.9). Das heißt, es liegt **Geschwindigkeitssättigung** vor.

Die Feldstärke E_C , oberhalb der $v = v_{\max} = \text{konst.}$ ist, beträgt für Silizium etwa

$$E_C \approx 10^4 \text{ V / cm}$$

mit

$$v_{\max} \approx 1 \cdot 10^7 \text{ cm / s}$$

Derart hohe Feldstärken werden bei MOSFET mit kurzen Kanälen leicht erreicht.

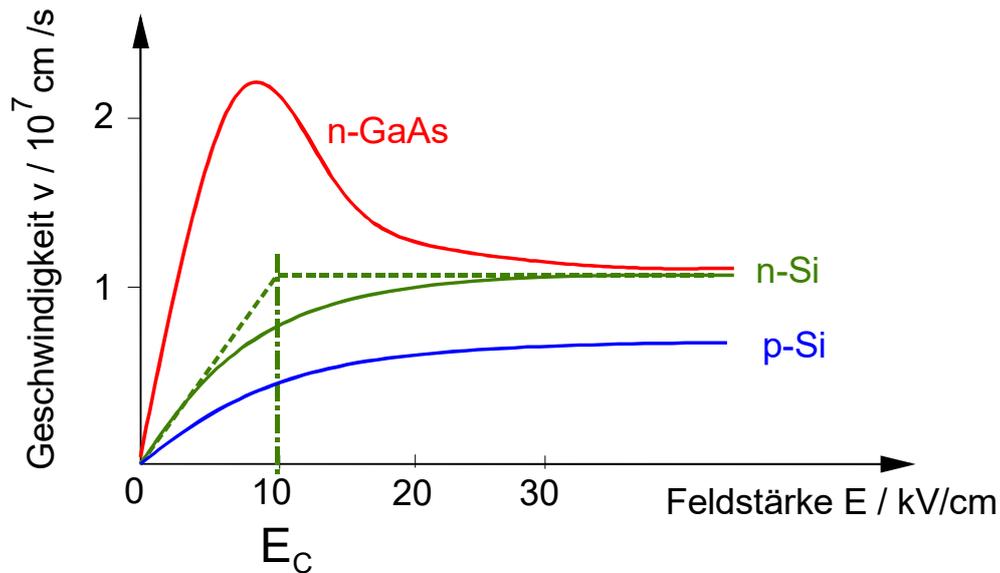


Abb. 2.9 Abhängigkeit der Geschwindigkeit von der auftretenden Feldstärke

Beispiel:

Die Feldstärke im Kanal eines MOSFET lässt sich für den einfachsten Fall entsprechend:

$$E_x \approx \frac{U_{DS}}{L_G}$$

abschätzen. D.h. für einen MOSFET mit einer Gate-Länge $L_G = 1\mu\text{m}$ beträgt die Feldstärke bei einer Drain-Source-Spannung $U_{DS} = 1\text{V}$ im einfachsten Fall:

$$E_x \approx \frac{U_{DS}}{L_G} = \frac{1\text{V}}{1\mu\text{m}} = 10^4 \text{V/cm}$$

Damit ist für diesen Fall bereits Sättigung im Kanal erreicht und für den Strom folgt dann:

$$I_{D,sat} = q \cdot n \cdot v_{\text{max}} \cdot A \quad (2.23a)$$

wobei $I_{D,sat}$ unabhängig von U_{DS} ist.

Die Spannung $U_{DS,sat}$, bei der $v = v_{\text{max}} = \text{konst.}$ erreicht wird, kann aus der Beziehung:

$$U_{DS,sat} \approx E_C \cdot L_G \quad (2.23b)$$

abgeschätzt werden. Sie muss mit $U_{DS,sat}$ aus Gl. (2.20) verglichen werden. Die kleinere der beiden Spannungen bestimmt $I_{D,sat}$. Je nachdem, ob zuerst Kanalabschnürung oder Driftgeschwindigkeits-sättigung erreicht wird, gilt Gl. (2.21) oder Gl. (2.23) für den Strom im Sättigungsbereich.

Die im einzelnen behandelten Fälle sind in Abb. 2.10 dargestellt.

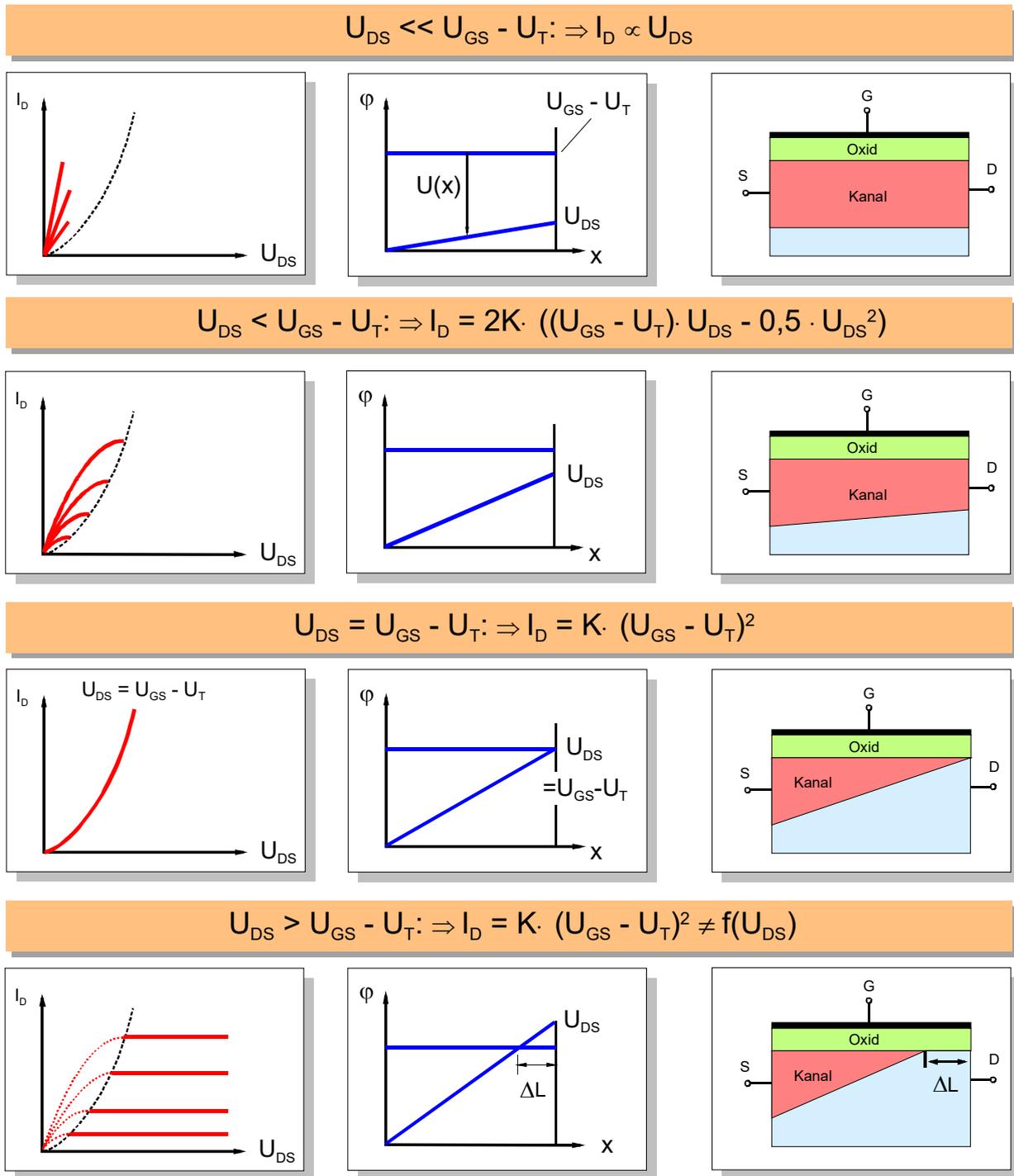


Abb. 2.10 Kennlinie, Potenzialverlauf und Querschnitt durch den MOSFET für die verschiedenen Spannungsbereiche

2.1.3 Die Kennlinienfelder des MOSFET

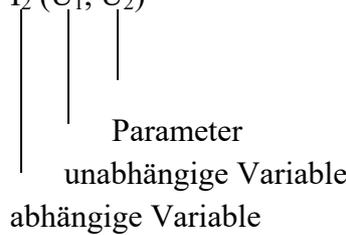
Verstärkerbauelemente werden üblicherweise als Vierpole (Zweitor) behandelt. Dazu muss beim MOSFET einer der drei Anschlüsse für Eingang und Ausgang gemeinsam sein.



Abb. 2.11 Allgemeiner Vierpol

Folgende Kennlinienfelder sind von Interesse:

- 1. Eingangskennlinienfeld $I_1(U_1, U_2)$
- 2. Ausgangskennlinienfeld $I_2(U_2, U_1)$
- 3. Übertragungskennlinienfeld $I_2(U_1, U_2)$



Für den MOSFET sind folgende Grundschaltungen denkbar (Abb. 2.12):

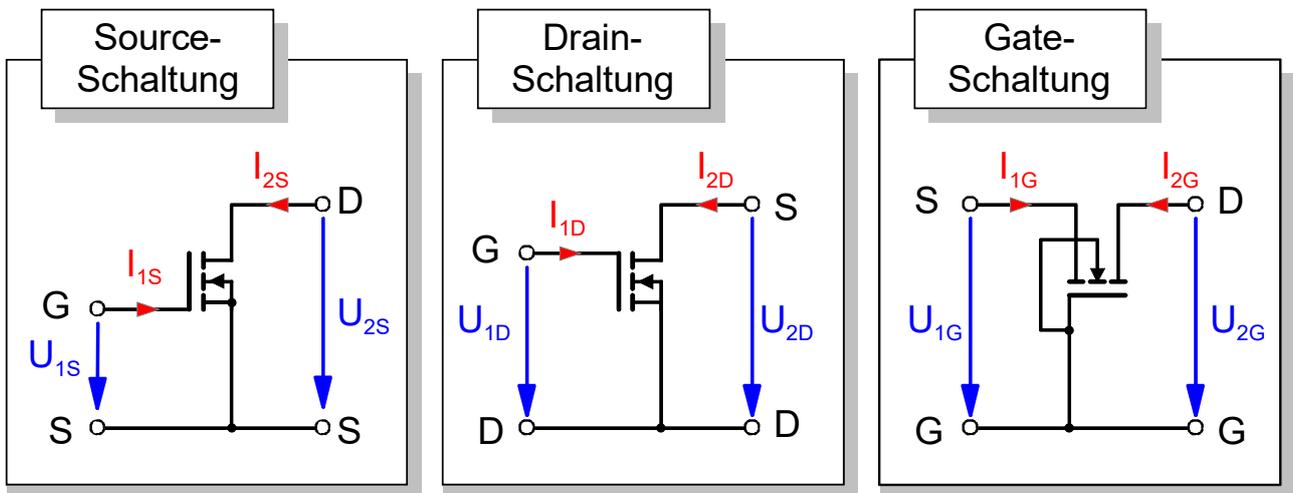


Abb. 2.12 Grundschaltungen des MOSFET (n-Kanal-Anreicherung)

Wegen der Symmetrie des MOSFET sind Source- und Drain-Schaltung vom Prinzip her identisch. Der Substratanschluss wird in der Praxis meist mit dem Source-Kontakt verbunden. Die pn-Diode am Drain-Kontakt wird damit in Sperrichtung vorgespannt. Damit wird der FET unsymmetrisch und Source- und Drain-Kontakt sind nicht mehr ohne weiteres zu vertauschen.

Gegenüber den beiden anderen Grundschaltungen weist die Gate-Schaltung bedingt durch den Kanalleitwert eine hohe Rückwirkung auf. Daher wird sie selten eingesetzt.

2.1.3.1 Die Kennlinienfelder der Source-Schaltung

Für die Source-Schaltung als häufigste Grundschaltung ergeben sich folgende Kennlinienfelder:

1. **Eingangskennlinienfeld (EKF):**

Im Idealfall ist $I_{1S} = I_G = 0$, so dass es kein Eingangskennlinienfeld gibt.

2. **Ausgangskennlinienfeld (AKF)**

$$I_{2S} = I_{2S}(U_{2S}, U_{1S}) = I_D(U_{DS}, U_{GS}) \quad [\text{Gl. (2.15), (2.21), (2.23) und Abb. 2.10}]$$

3. **Übertragungskennlinie(nfeld) (ÜKF):**

$$I_{2S} = I_{2S}(U_{1S}, U_{2S}) = I_D(U_{GS}, U_{DS})$$

Da Verstärker im Allgemeinen im Sättigungsbereich, also bei $U_{DS} \geq U_{GS} - U_T$ betrieben werden, gilt

$$I_D = I_{D,sat} = K(U_{GS} - U_T)^2$$

sofern $U_{GS} > U_T$ ist.

Man erhält das Übertragungskennlinienfeld des n-Kanal-Anreicherungs-MOSFET in Abb. 2.13.

Im Sättigungsgebiet, also für $U_{DS} > U_{GS} - U_T$, existiert (im Idealfall) nur eine Kennlinie.

Für

$$U_{DS} \leq U_{GS} - U_T$$

nimmt der Strom I_D bei konstanter Gate-Spannung U_{GS} ab, so dass sich ein entsprechendes Kennlinienfeld, wie in Abb. 2.13 dargestellt, ergibt.

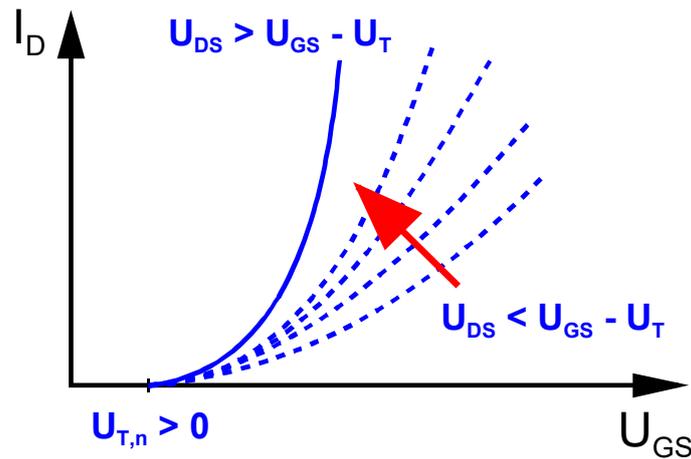


Abb. 2.13 Übertragungskennlinie(nfeld) des n-Kanal-Anreicherungs-MOSFET

Die Steigung der Übertragungskennlinie wird als Steilheit S bzw. g_m (engl.: *transconductance*) bezeichnet und ist ein Maß für die so genannte Kleinsignalverstärkung des Bauelements (s. Kap. 2.4.2).

Es gilt:

$$S = \left. \frac{\partial I_D}{\partial U_{GS}} \right|_{U_{DS,AP}} = S = g_m \quad (2.24a)$$

Für den Fall der Sättigung ergibt sich dann mit

$$I_D = I_{D,sat} = K(U_{GS} - U_T)^2$$

$$S = \left. \frac{\partial I_D}{\partial U_{GS}} \right|_{U_{DS}=konst.} = 2K(U_{GS} - U_T) \quad (2.24b)$$

2.1.3.1.1 Kennlinien der p-Kanal-Anreicherungstypen (Enhancement-FET, E-FET)

p-Kanal-Anreicherungstypen können in völliger Analogie zu n-Kanal-Anreicherungstypen berechnet werden, wenn die Elektronen durch Löcher ersetzt werden und die Vorzeichen der anzulegenden Spannung vertauscht werden. Damit ergeben sich folgende Kennlinienfelder:

- a) das **Ausgangskennlinienfeld** (Abb. 2.14) und
- b) das **Übertragungskennlinienfeld** (Abb. 2.15)

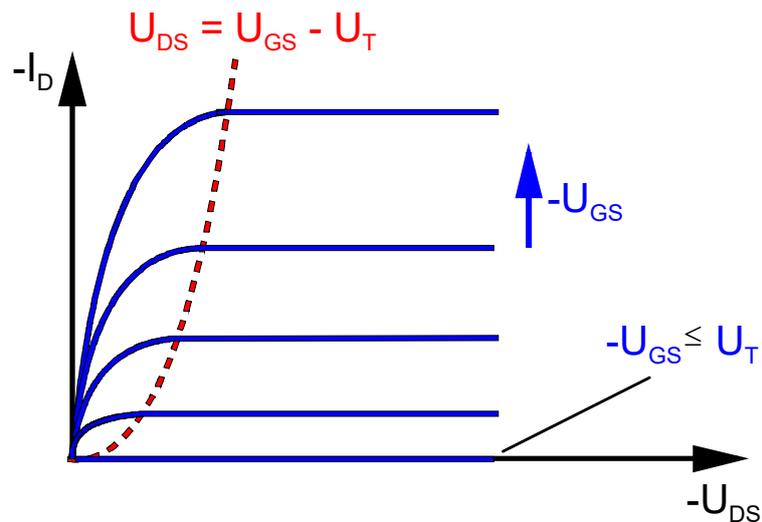


Abb. 2.14 Ausgangskennlinienfeld eines p-Kanal-Anreicherungs-FET

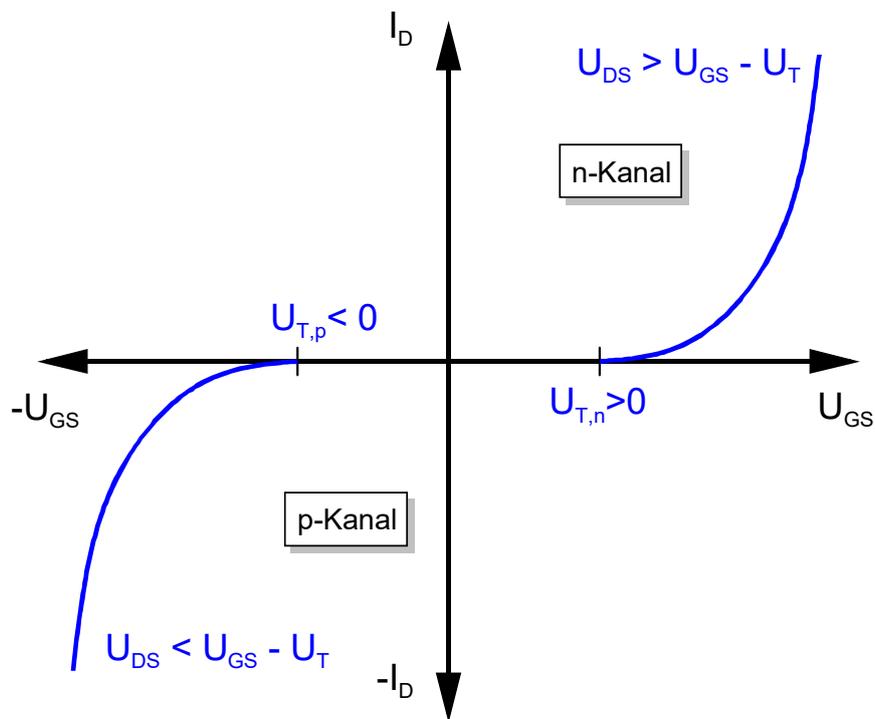


Abb. 2.15 Vergleich der Übertragungskennlinienfelder von p- und n-Kanal-Anreicherungs-MOSFET

Aus Abb. 2.15 ist zu erkennen, dass

- beim **n-Kanal** Anreicherungs-FET (E-FET): $U_T = U_{T,n} > 0$ und
- beim **p-Kanal** Anreicherungs-FET (E-FET): $U_T = U_{T,p} < 0$ ist.

2.1.3.1.2 Kennlinien der n- und p-Kanal-Verarmungstypen (Depletion-FET, D-FET)

Durch eine Dotierung des Substrates in einer dünnen Schicht an der Oberfläche kann ein Kanal erzeugt werden. Dadurch wird ein Stromfluß bereits bei einer äußeren Gate-Spannung $U_{GS} = 0$ möglich. Je nach Vorzeichen der Gate-Spannung und der Kanaldotierung kann die bewegliche Ladung im Kanal erhöht oder erniedrigt werden.

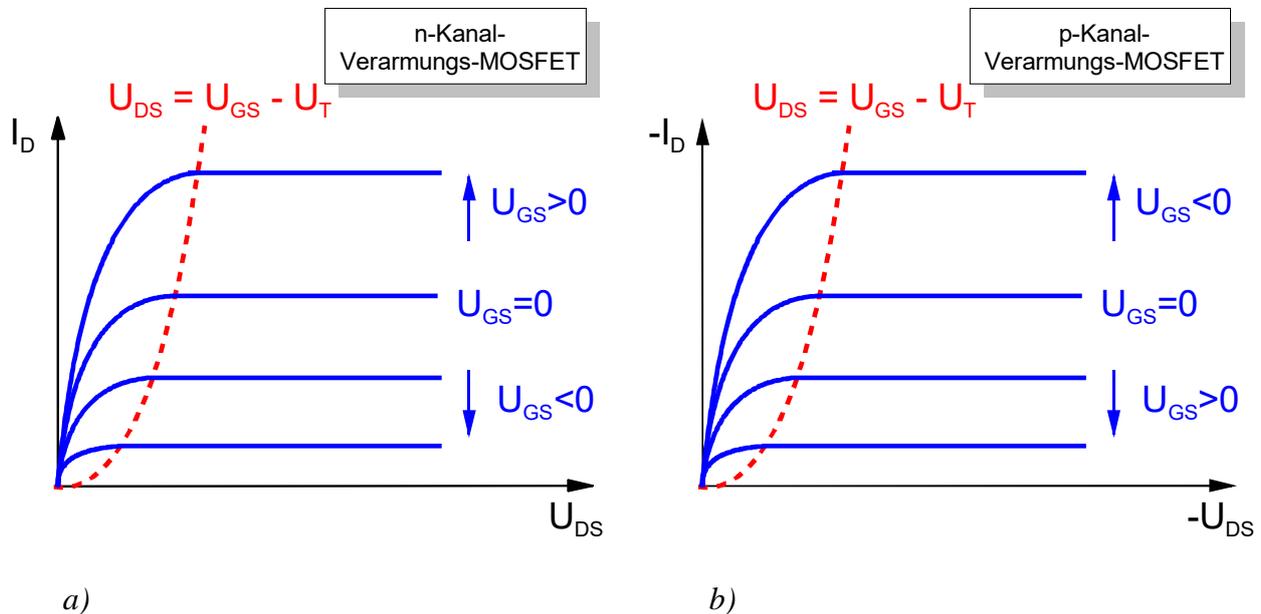


Abb. 2.16 Ausgangskennlinienfelder von a) n-Kanal- und b) p-Kanal-Verarmungs-MOSFET

Damit ergeben sich:

- das **Ausgangskennlinienfeld** (Abb. 2.16) und
- das **Übertragungskennlinienfeld** (Abb. 2.17)

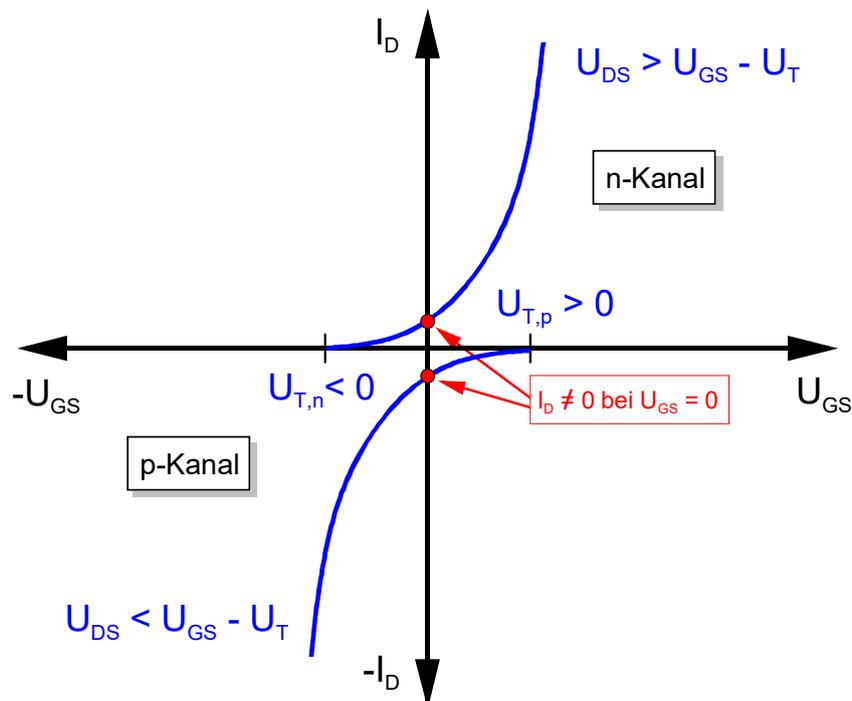


Abb. 2.17 Übertragungskennlinien von n-Kanal- und p-Kanal-Verarmungs-MOSFET

Abb. 2.16 lässt erkennen, dass bei entsprechend gewählter Gate-Spannung der Strom I_D unterbunden werden kann. Im Gegensatz zum Anreicherungs-FET muss in diesem Fall eine zusätzliche **Abschnürspannung** oder **Schwellesspannung** U_T angelegt werden, um den Kanal abzuschneiden.

Aus Abb. 2.17 ist zu erkennen, dass

- beim **n-Kanal Verarmungs-FET (D-FET)**: $U_T = U_{T,n} < 0$ und
- beim **p-Kanal Verarmungs-FET (D-FET)**: $U_T = U_{T,p} > 0$ ist.

Anmerkung: Die Schwellesspannung U_T wird oftmals auch mit U_P (*pinch-off-voltage*) bezeichnet. Um die Bezeichnungen jedoch für MESFET (vgl. Kap.3.2) und MOSFET zu vereinheitlichen, wird für die weitere Berechnung die **außen anzulegende Spannung**, die notwendig ist, den Kanal abzuschneiden, mit U_T und die dabei **über dem Kanal abfallende Spannung** mit U_P bezeichnet.

2.1.4 Temperaturabhängigkeit der Kennlinien

In Gl. (2.21):

$$I_{D,sat} = \frac{\mu C_G}{2 L_G^2} (U_{GS} - U_T)^2$$

ist die temperaturempfindlichste Größe die Beweglichkeit. Für sie gilt nach [13]:

$$\mu \sim T^{-m}$$

mit $m = 1,5 \dots 2,5$ je nach Halbleiter und im Bereich der Raumtemperatur.

Für die Temperaturabhängigkeit folgt:

$$\frac{d\mu}{dT} \sim -m \cdot T^{-m-1} < 0$$

Damit folgt für den Temperaturkoeffizienten des Stromes I_D :

$$\frac{dI_{D,sat}}{dT} = \frac{1}{2} \frac{C_G}{L_G^2} (U_{GS} - U_T)^2 \cdot \frac{d\mu}{dT} < 0 \quad (2.25)$$

D.h. mit steigender Temperatur nimmt der Strom ab. Im Gegensatz dazu nimmt zum Beispiel beim Bipolartransistor der Strom mit steigender Temperatur zu, was zur Selbstzerstörung führen kann. Diese Gefahr besteht beim FET nicht.

2.1.5 Einfluss der Substratvorspannung

Um sicherzustellen, dass die pn-Übergänge an Source und Drain auf jeden Fall gesperrt sind und somit kein Strom über das Substrat abfließen kann, liegt der Substratanschluß im allgemeinen

- im Fall des **n-Kanal-FET** auf dem **niedrigsten** und
- im Fall des **p-Kanal-FET** auf dem **höchsten**

in der Schaltung auftretenden Potenzial.

Ist der Substratanschluß, im Fall von diskreten Bauelementen, aus dem Gehäuse herausgeführt, kann das Substrat auch auf eine gegenüber Source negativere (n-Kanal) oder positivere (p-Kanal) Spannung gelegt werden. Dadurch wird jedoch der Drain-Strom beeinflusst: Zwischen den n^+ -dotierten Bereichen an Source, Drain sowie dem n-leitenden, durch Inversion hervorgerufenen Kanal einerseits und dem p^- -dotierten Substrat andererseits baut sich eine Raumladungszone auf (Abb. 2.18). Je negativer im Fall des n-Kanal-FET die Substratvorspannung U_{BS} wird, um so breiter wird diese Raumladungszone, und umso schmaler wird der Kanal selbst. Dadurch wird also die Dicke des Kanals moduliert und die Schwellenspannung U_T verschoben.

Die Folge ist, dass bei gleicher Gate-Source-Spannung U_{GS} die Kanalabschnürung schon bei kleineren Drain-Source-Spannungen U_{DS} auftritt und daher ein kleinerer Drain-Sättigungsstrom $I_{D,sat}$ fließt. Der gleiche Effekt tritt im Fall des p-Kanal-FET bei positiver Substrat-Source-Spannung U_{BS} auf. Dadurch wird die Dicke des Kanals moduliert und die Schwellenspannung U_T verschoben.

Prinzipiell besteht also auch die Möglichkeit, den Drain-Strom sowohl mit der Gate-Source-Spannung U_{GS} als auch mit der Substrat-Source-Spannung U_{BS} zu steuern. Die Steuerwirkung über das Substrat ist jedoch relativ gering.

Statt Gl. (2.16) erhält man:

$$I_D = 2K \left[(U_{GS} - U_T) U_{DS} - \frac{1}{2} U_{DS}^2 - \frac{1}{2} U_{BS}^2 \right] \quad (2.26)$$

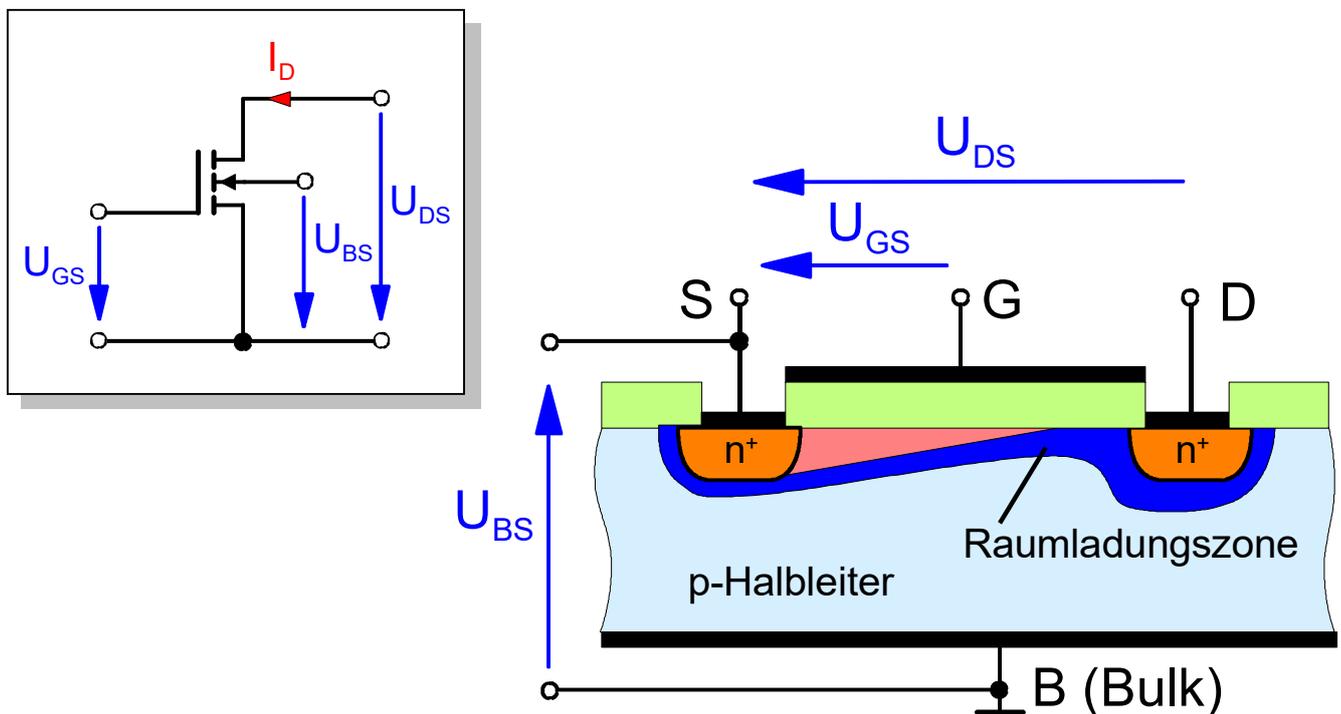


Abb. 2.18 Einfluß einer Substratvorspannung

2.1.6 Arbeitspunkteinstellung und -stabilisierung

Einstellung und Stabilisierung des Arbeitspunktes werden am Beispiel eines n-Kanal-Anreicherungs- und eines n-Kanal-Verarmungs-MOSFET behandelt.

2.1.6.1 n-Kanal-Anreicherungs-MOSFET (E-MOSFET)

Prinzipiell gibt es viele verschiedene Möglichkeiten der Arbeitspunkteinstellung. Hier seien drei aufgezeigt:

- U_{GS} und U_{DS} können getrennt über zwei separate Spannungsversorgungen eingestellt werden (Abb. 2.19b). Da der Gate-Strom I_G im Eingangszweig des MOSFET vernachlässigbar klein ist, entspricht die Gate-Source-Spannung U_{GS} der Spannung am Eingang U_{GG} .

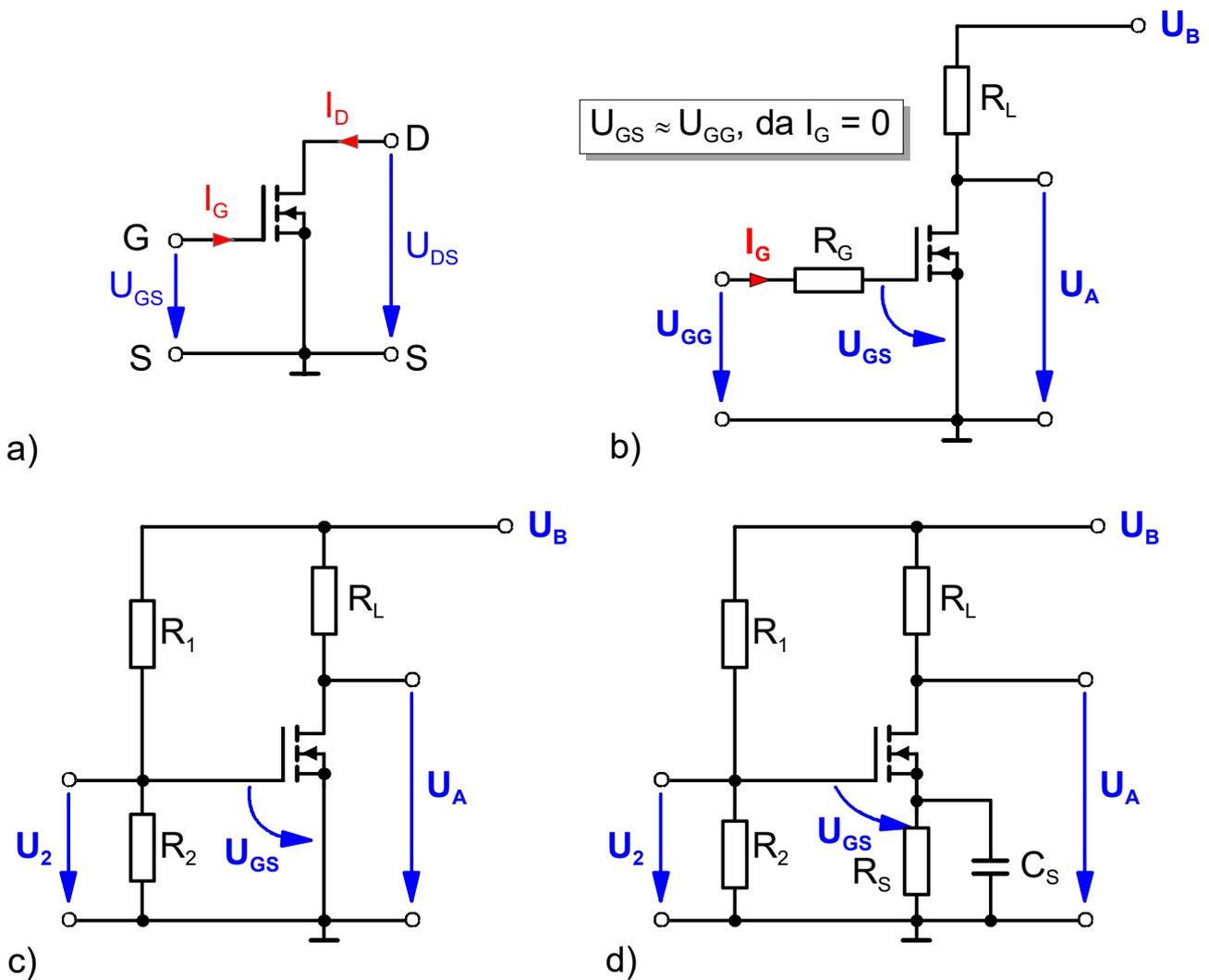


Abb. 2.19 Arbeitspunkteinstellung und -stabilisierung von n-Kanal-Anreicherungs-MOSFET
 a) n-Kanal-MOSFET in Source-Schaltung
 b) Arbeitspunkteinstellung mittels zweier Spannungsversorgungen (U_B und U_{GG})
 c) Arbeitspunkteinstellung mit nur einer Spannungsversorgung
 d) zusätzliche Stabilisierung durch Gegenkopplung

Diese Lösung wird aber insbesondere für integrierte Schaltungen nicht eingesetzt, da für jede Spannungsversorgung eigene Zuleitungen erforderlich sind, die Platz und Leistung verbrauchen und parasitäre Kapazitäten, ggf. auch Induktivitäten zur Folge haben.

- Da U_{GS} und U_{DS} gleiches Vorzeichen haben, können beide über eine Spannungsversorgung auch mittels Spannungsteiler eingestellt werden (Abb. 2.19c).

Es gilt:

$$U_2 = U_{GS} = \frac{R_2}{R_1 + R_2} U_B \quad (2.27)$$

- Der Arbeitspunkt kann gegen Schwankungen infolge von Temperaturänderungen, Alterung u.ä. durch eine zusätzliche Gegenkopplung stabilisiert werden. Eine der häufig verwendeten Gegenkopplungen nutzt einen zusätzlichen Widerstand R_S im Source-Zweig (Abb. 2.19d). Für die Steuerspannung U_{GS} folgt:

$$U_{GS} = U_2 - I_D \cdot R_S \quad (2.28)$$

mit U_2 nach Gl. (2.27). Nimmt der Strom I_D z.B. aufgrund eines Temperaturanstiegs ab, nimmt U_{GS} zu und kompensiert die Stromabnahme.

Da die Gegenkopplung nur **gleichstrommäßig** wirksam werden soll, muss R_S durch einen Block-Kondensator C_S **wechselstrommäßig** überbrückt werden.

2.1.6.2 n-Kanal-Verarmungs-MOSFET (D-MOSFET)

Wenn der n-Kanal-Verarmungs-MOSFET im Anreicherungsbetrieb ($U_{GS} > 0$) arbeiten soll, haben U_{GS} und U_{DS} gleiches Vorzeichen. Dann kann ebenfalls die Schaltung nach Abb. 2.19d benutzt werden.

Meistens wird der selbstleitende MOSFET jedoch im Verarmungsmodus ($U_{GS} < 0$) betrieben. Dann haben U_{GS} und U_{DS} entgegengesetzte Vorzeichen.

Lässt man die Spannungsversorgung mit Hilfe von zwei Spannungsversorgungen (Abb. 2.19b) außer Acht, so bleibt zur Einstellung von U_{GS} die Möglichkeit nach Abb. 2.20b:

Wird angenommen, dass der DC-Eingangsstrom $I_E = 0$ ist, folgt aus einem Spannungsumlauf im Eingangskreis:

$$U_{GS} + I_D \cdot R_S - I_G \cdot R_G = 0$$

Mit $I_G = 0$ (idealer MOSFET), folgt damit

$$U_{GS} = -I_D \cdot R_S \quad (2.29)$$

Die gegenkoppelnde Wirkung von R_S bleibt erhalten.

Dieser Arbeitspunkt könnte auch mit $R_G = 0$ eingestellt werden. In diesem Fall würde allerdings ein ggfs. am Eingang angelegtes Wechselsignal kurzgeschlossen.

Aus diesem Grund sollte R_G möglichst hochohmig gewählt werden.

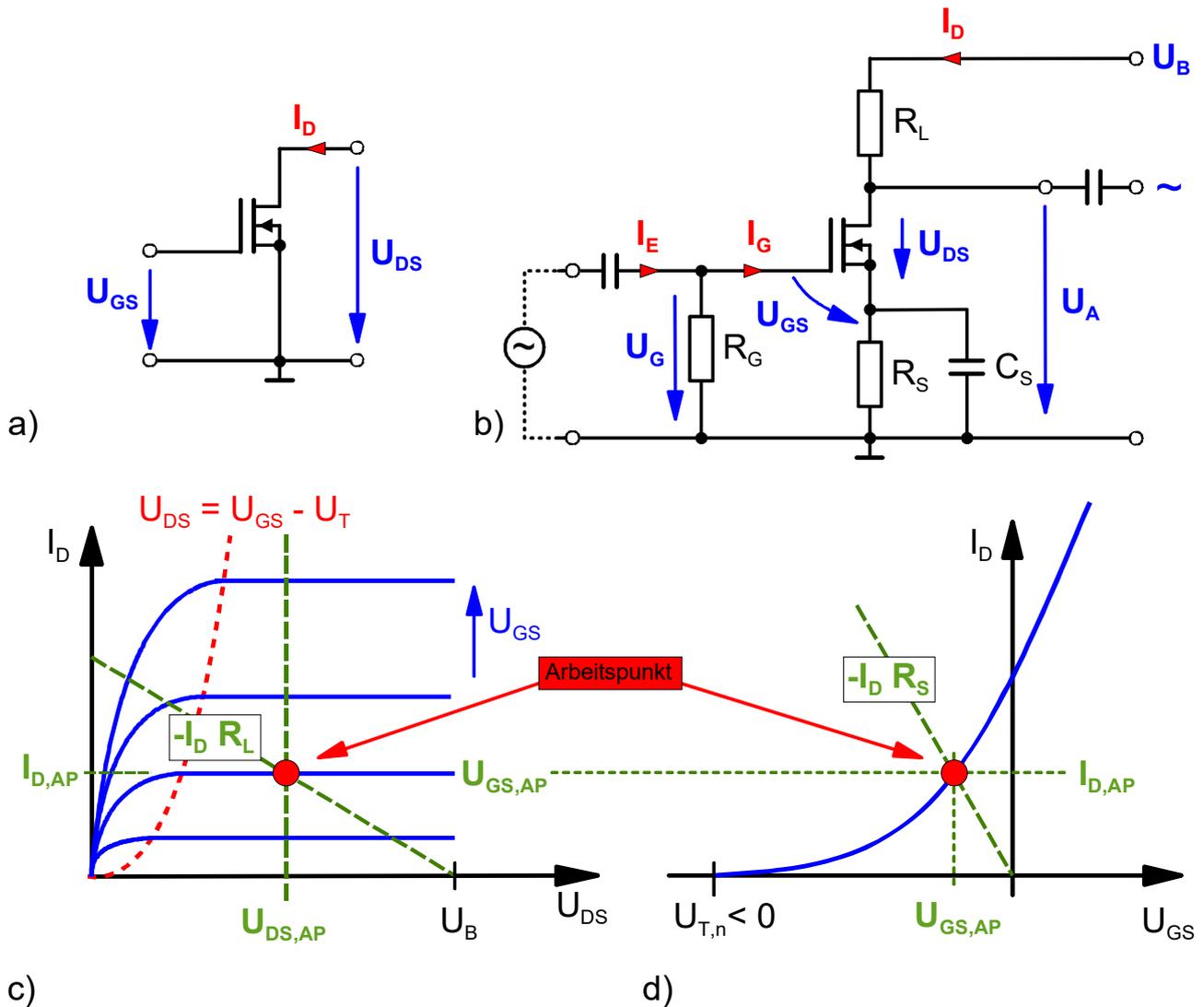


Abb. 2.20 Arbeitspunkteinstellung von n-Kanal-Verarmungs-MOSFET (a,b)

Arbeitspunkt im Ausgangskennlinienfeld (c) und in der Übertragungskennlinie (d)

2.1.7 Komplementäre MOS-Technik (CMOS)

In der Praxis werden FET sowohl im analogen Bereich, z.B. als Verstärker, als auch in der Digitaltechnik als elektronische Schalter eingesetzt.

Da für Prozessoren, Speicher u.ä. Anwendungen eine große Anzahl von Transistoren zum Einsatz kommt, spielt die dabei umgesetzte Leistung eine große Rolle. Insbesondere, da die dabei entstehende Wärme abgeführt werden muss, was vor allem bei hoher Integrationsdichte ein großes Problem darstellt.

Eine klassische, einfache Schaltstufe kann mit Hilfe eines FET und eines Lastwiderstands realisiert werden (Abb. 2.21):

Liegt am Eingang eine niedrige Spannung $U_A = U_A^L < U_T$ an, so ist der FET gesperrt und am Ausgang fällt eine hohe Spannung $U_Q = U_Q^H \approx U_B$ ab.

Wird die Spannung am Eingang erhöht, schaltet der FET durch, die Ausgangsspannung sinkt auf $U_Q = U_Q^L$. Die Schaltung zeigt also invertierendes Verhalten und wird daher auch als Inverter bezeichnet. Es gilt also:

$$U_A = U_A^L < U_T \quad \Rightarrow U_Q = U_Q^H \approx U_B$$

$$U_A = U_A^H \gg U_T \quad \Rightarrow U_Q = U_Q^L$$

Im durchgeschalteten Zustand "ON" fließt ein ggfs. recht großer Strom und dementsprechend wird in diesem stationären Fall Leistung umgesetzt.

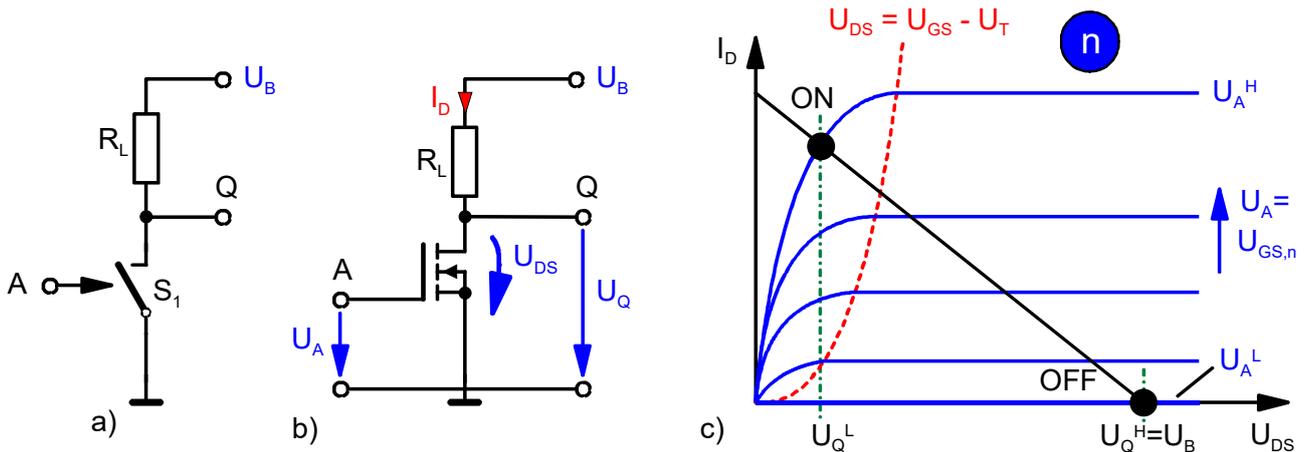


Abb. 2.21 a) einfache Schaltstufe
 b) Realisierung mit n-Kanal MOSFET
 c) stationäre Zustände (ON/OFF) im Ausgangskennlinienfeld (AKF)

Um den Leistungsverbrauch deutlich zu reduzieren, werden in der Praxis bei hoher Integrationsdichte komplementäre Schaltstufen verwendet, durch die erreicht wird, dass in beiden stationären Zuständen ("ON" oder "OFF") kein Strom fließt, somit auch keine Leistung umgesetzt und damit Wärme erzeugt wird (Abb. 2.22). Strom fließt nur im Umschaltmoment.

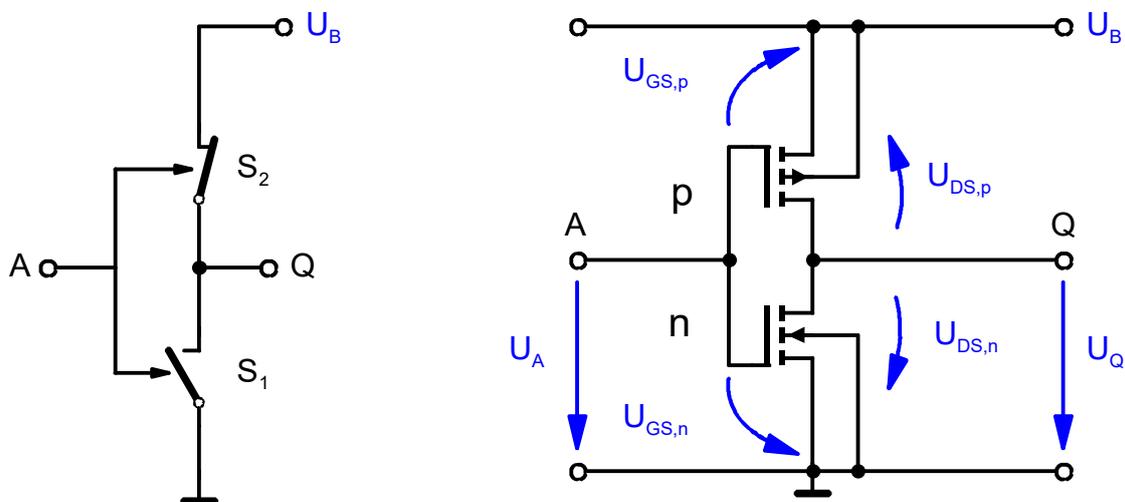


Abb. 2.22 a) Komplementäre Schaltstufe
 b) Realisierung mit komplementären MOSFET

Diese Schaltstufe wird mit Hilfe eines n- und eines p-Kanal MOSFET, beide selbstsperrend, realisiert (*Complementary Metal Oxide Semiconductor* (CMOS)).

Beide Gate-Elektroden werden gleichzeitig durch das Eingangssignal U_A angesteuert. Dadurch bedingt, sperrt immer einer der FET, der andere ist leitend, d.h. beide FET werden gegenphasig angesteuert.

Mit Hilfe der Übertragungskennlinien $|I_D| = f(U_{GS})$ wird die Wirkungsweise des CMOS-Inverters verständlich:

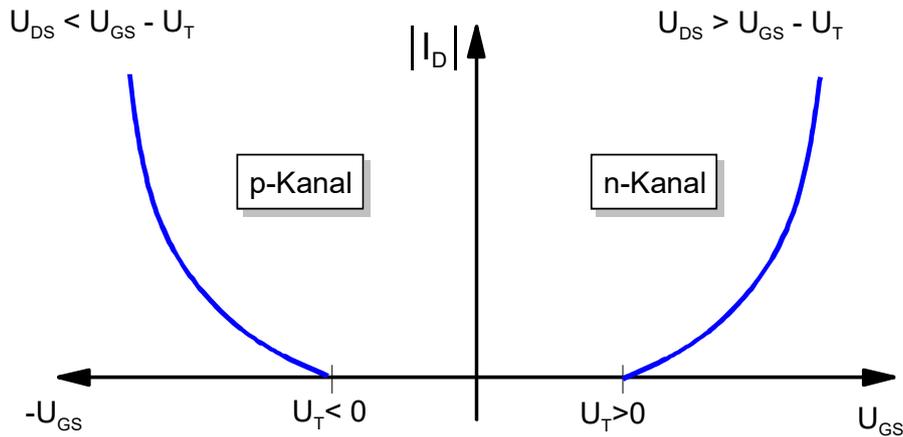


Abb. 2.23 Übertragungskennlinien $I_D = f(U_{GS})$ der n- und p-Kanal MOSFET

Es gilt:

$$U_{GS,n} = U_A \quad U_{GS,p} = U_A - U_B$$

$$U_{DS,n} = U_Q \quad U_{DS,p} = U_Q - U_B$$

Eine niedrige Spannung ($U_A < U_{T,n}$) am Eingang A des Inverters entspricht dabei der logischen "0". Sie sorgt dafür, dass der n-Kanal FET sperrt und nur die p-Kanal-Komponente leitet und somit quasi die Versorgungsspannung mit dem Ausgang Q verbunden ist (AP1 in Abb. 2.24):

$$U_A = U_A^L < U_{T,n} \quad \Rightarrow U_Q = U_Q^H \approx U_B$$

Die logische "1" entspricht einer höheren positiven Eingangsspannung und bewirkt, dass nur die n-Kanal-Komponente leitet und somit der n-Kanal MOSFET durchschaltet (AP6 in Abb. 2.24). Damit gilt:

$$U_A = U_A^H \geq U_B + U_{T,p} \Rightarrow U_Q = U_Q^L \approx 0$$

da für

$$U_{GS,p} = U_A - U_B > U_{T,p}$$

der p-Kanal FET sperrt.

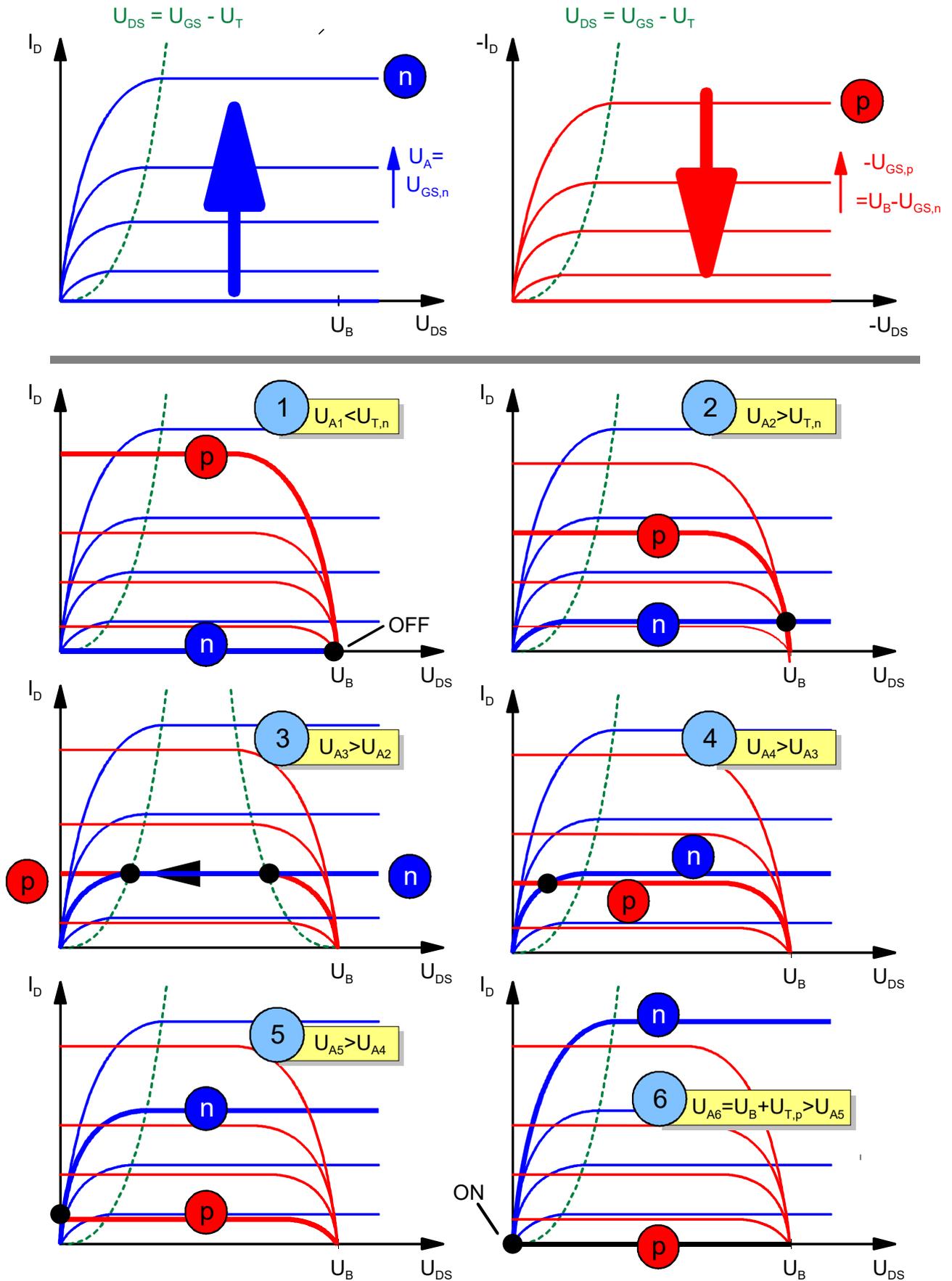


Abb. 2.24 Umschaltvorgang der MOSFET und Konstruktion der jeweiligen Arbeitspunkte ($K_n \neq K_p$)

Abb. 2.23 zeigt den prinzipiellen Aufbau einer CMOS-Schaltstufe. Dabei ist darauf zu achten, dass die pn-Übergänge an den ohmschen Kontakten Drain und Source auf jeden Fall gesperrt sind, um Leckströme zu vermeiden. Dies wird dadurch erreicht, dass im Fall des n-Kanal-FET der Substratanschluss (bzw. die n-"Wanne") auf das niedrigste in der Schaltung auftretende Potential (hier: 0V, Masseanschluss) und im Fall des p-Kanal-FET der Substratanschluss (bzw. die p-"Wanne") auf das höchste in der Schaltung auftretende Potential (hier: U_B) gelegt wird.

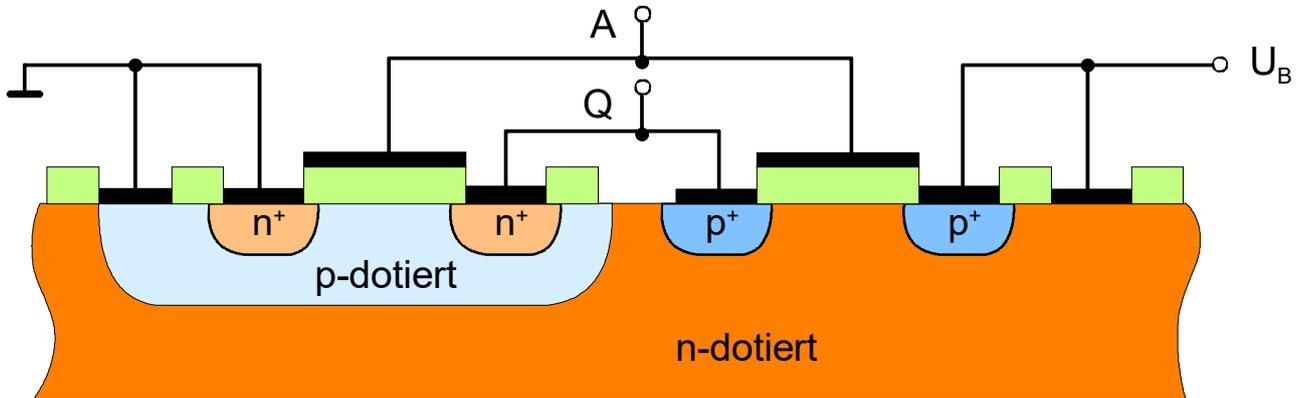


Abb. 2.25 prinzipieller Aufbau einer CMOS-Schaltstufe

CMOS-Eingänge sind empfindlich gegenüber statischen Aufladungen und Überspannungen, weshalb nach Möglichkeit vor CMOS-Eingänge Schutzschaltungen z.B. mit Dioden gesetzt werden.

2.2 Sperrschicht-Feldeffekttransistoren

Bei Sperrschicht-Feldeffekttransistoren erfolgt die Steuerung des Stromes anhand der Modulation der Breite eines leitfähigen Kanals mittels Steuerung einer Raumladungszone, hervorgerufen durch einen:

- a) **pn-Übergang** (engl. *junction*) => **J**unction-**F**ET (**JFET**) (Abb. 2.26a) oder
- b) **Metall-Halbleiter-(Schottky-)Kontakt** => **M**etall-**S**emiconductor-**F**ET (**MESFET**) (Abb. 2.26b) oder **H**eterostruktur-**F**ET (**HFET**) (Kap. 3.3)

Als Kanalmaterial für JFET und MESFET kommen Silizium, vor allem aber III/V-Halbleiter wie GaAs, InGaAs o.a. zum Einsatz. Diese werden insbesondere für Anwendungen bei höchsten Frequenzen (>1GHz) bzw. höchsten Bitraten (>1Gbit/s) bevorzugt eingesetzt, da die III/V-Halbleiter eine deutlich höhere Ladungsträgerbeweglichkeit als Si aufweisen. (vgl. Abb. 2.9). Darüber hinaus sind die III/V-Halbleiter aufgrund ihrer Bandstruktur und damit ihrer Eigenschaften für den Aufbau optoelektronischer Schaltungen hervorragend geeignet.

2.2.1 Aufbau und Wirkungsweise

Als **Substrat** dient ein hochohmiger Halbleiter:

Si: $\rho \approx 10^3 \Omega\text{cm}$;

GaAs, InP: $\rho \approx 10^7\text{-}10^8 \Omega\text{cm}$ (s.i.: semi-isolierend).

Der Kanal wird durch Ionenimplantation (oder seltener durch Diffusion) in das Substrat hinein erzeugt. Die implantierten oder diffundierten Atome dotieren das Substrat in einer dünnen oberflächennahen Schicht.

Es wird im folgenden von einer **n-dotierten** Kanalschicht ausgegangen, d.h. von **n-Kanal FET**.

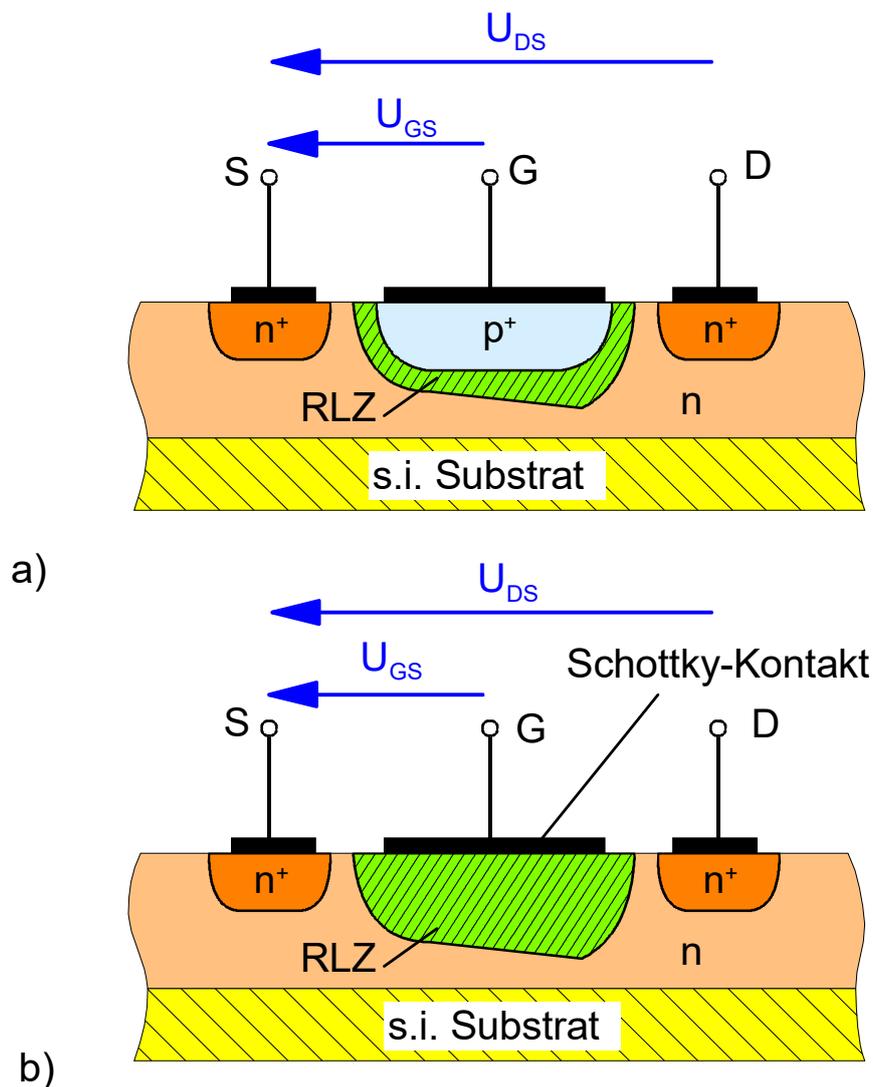


Abb. 2.26 Prinzipieller Aufbau eines

a) Sperrschicht-Feldeffekttransistors (JFET)

b) Feldeffekttransistors mit Schottky-Kontakt als Steuerelektrode (MESFET)

Source und **Drain** werden durch Aufdampfen und Einlegieren von Metallschichten erzeugt, die Dotierstoffe enthalten (Ohm'sche Kontakte).

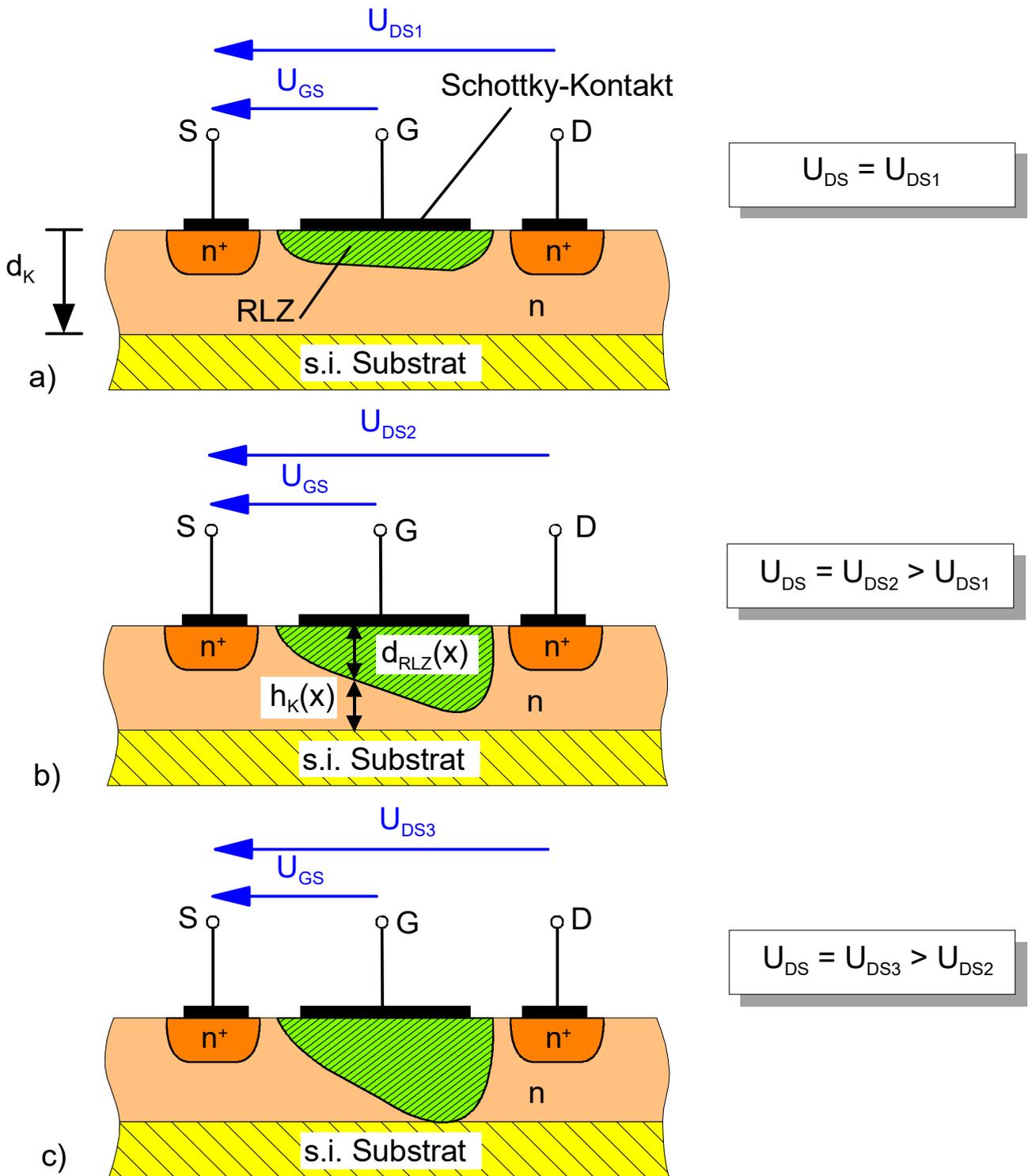


Abb. 2.27 Querschnitt durch einen n-Kanal D-MESFET bei verschiedenen Spannungen U_{DS} und einer festen Gate-Spannung $U_{GS} = \text{konst.}$

Das **Gate** wird

- im Fall des **JFET** durch zusätzliche Diffusion oder Implantation eines p-Gebietes und anschließendem Aufdampfen eines Ohmschen Kontaktes (Abb. 2.26a) oder
- im Fall des **MESFET** durch Aufdampfen einer Metallschicht (sperrender Kontakt, Schottky-Kontakt, Abb. 2.26b)

hergestellt.

Die Funktionsweise eines ist MESFET ist aus Abb. 2.27 zu erkennen:

Aufgrund der inneren Spannung U_D des Schottky-Kontaktes (Gate) (Diffusionspannung, engl.: *built-in voltage*) bildet sich unter dem Kontakt eine Raumladungszone (RLZ) aus. Wird zusätzlich eine Gate-Source Spannung $U_{GS} < 0$ angelegt, so nimmt die Ausdehnung der Raumladungszone d_{RLZ} zu und damit die Höhe des Kanals

$$h_K = d_K - d_{RLZ}$$

ab. Bei Anlegen einer Spannung $U_{DS} > 0$ fließt dann ein Strom I_D durch den Kanal, der von der Steuerspannung U_{GS} abhängt.

Wird U_{GS} konstant gehalten und U_{DS} erhöht, so dehnt sich die Raumladungszone zum Drain-Kontakt hin immer weiter in die leitende Kanalschicht aus bis schließlich die Ausdehnung der Raumladungszone fast der Kanaldicke entspricht (Abb. 2.27b). In dem verbleibenden Kanal herrschen hohe Stromdichten und Feldstärken, d.h. U_{DS} fällt überwiegend über diesem Bereich ab und I_D bleibt damit weitgehend konstant. Es kommt damit wieder - wie im Fall des MOSFET - zur **Sättigung** des Stromes.

Für größere Spannungen U_{DS} dehnt sich dieser Bereich weiter zum Source-Kontakt hin aus (Abb. 2.27c).

Die Kennlinien des MESFET sind daher qualitativ ähnlich denen des MOSFET:

- kleine Spannungen U_{DS} : → steuerbarer Widerstand
- große Spannungen U_{DS} : → Sättigung des Drain-Stromes

Man unterscheidet auch hier wieder:

- **selbstsperrender MESFET**

(Anreicherungs-MESFET / engl.: *enhancement-MESFET*, E-MESFET)

d.h. die leitende Kanalschicht unter dem Gate-Kontakt ist so dünn, dass bei $U_{GS} = 0$ und $U_{DS} = 0$ die Diffusionspannung des Schottky-Kontaktes ausreicht, die leitende Schicht zu verarmen. Damit ist dort **kein** leitender Kanal vorhanden und somit kann auch kein Strom für $U_{DS} > 0$ fließen.

Um Stromfluss zu ermöglichen, muss **zusätzlich** eine positive Spannung $U_{GS} > 0$ angelegt werden, damit die Ausdehnung der RLZ reduziert und der Kanal des MESFET aufgesteuert wird. Dabei muss jedoch $U_{GS} \leq U_D$ bleiben, wenn der Gate-Strom $I_G \approx 0$ bleiben soll.

- **selbstleitender MESFET**

(Verarmungs-MESFET / engl.: *depletion-MESFET*, D-MESFET)

d.h. bei $U_{GS} = 0$ und $U_{DS} = 0$ ist bereits ein leitender Kanal unter dem Gate-Kontakt vorhanden, und es muss **zusätzlich** eine negative Spannung $U_{GS} < 0$ angelegt werden, um den Kanal dort abzuschneiden.

2.2.2 Kennlinienberechnung

Der Kanal wird in Abschnitte der Länge Δx unterteilt, von denen jeder den Widerstand $dR(x)$ hat. $dR(x)$ hängt von der Dicke des verbleibenden Kanals unterhalb der Raumladungszone und den geometrischen Abmessungen ab. Die Dicke der Raumladungszone wiederum wird durch die räumlich konstante Dotierung und den Spannungsabfall $U_{GK}(x)$ über der Raumladungszone bestimmt.

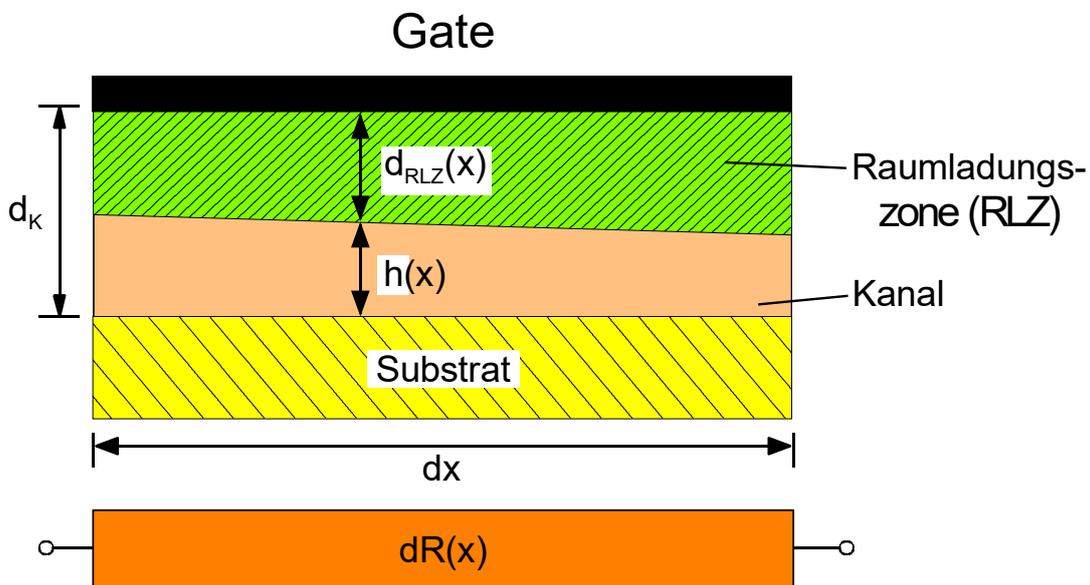


Abb. 2.28 Ausschnitt aus dem Kanal eines MESFET

Für die Berechnung des Drain-Stromes gilt als Voraussetzung wieder die Shockley'sche Näherung (siehe Kap. 3.1.2.) und $\mu = konst.$

Für das Widerstandselement $dR(x)$ folgt:

$$dR(x) = \frac{\rho}{h(x) \cdot Z} dx \quad h(x) : \text{ Dicke des Kanals}$$

Z : Gate-Weite

mit dem spezifischen Widerstand

$$\rho = \frac{1}{n \cdot q \cdot \mu}$$

bzw. mit $n = N_D$

$$dR = \frac{1}{qN_D \mu Z h(x)} dx \quad (2.30)$$

Dabei gilt für die Dicke des leitenden Kanals $h(x)$:

$$h(x) = d_K - d_{RLZ}(x)$$

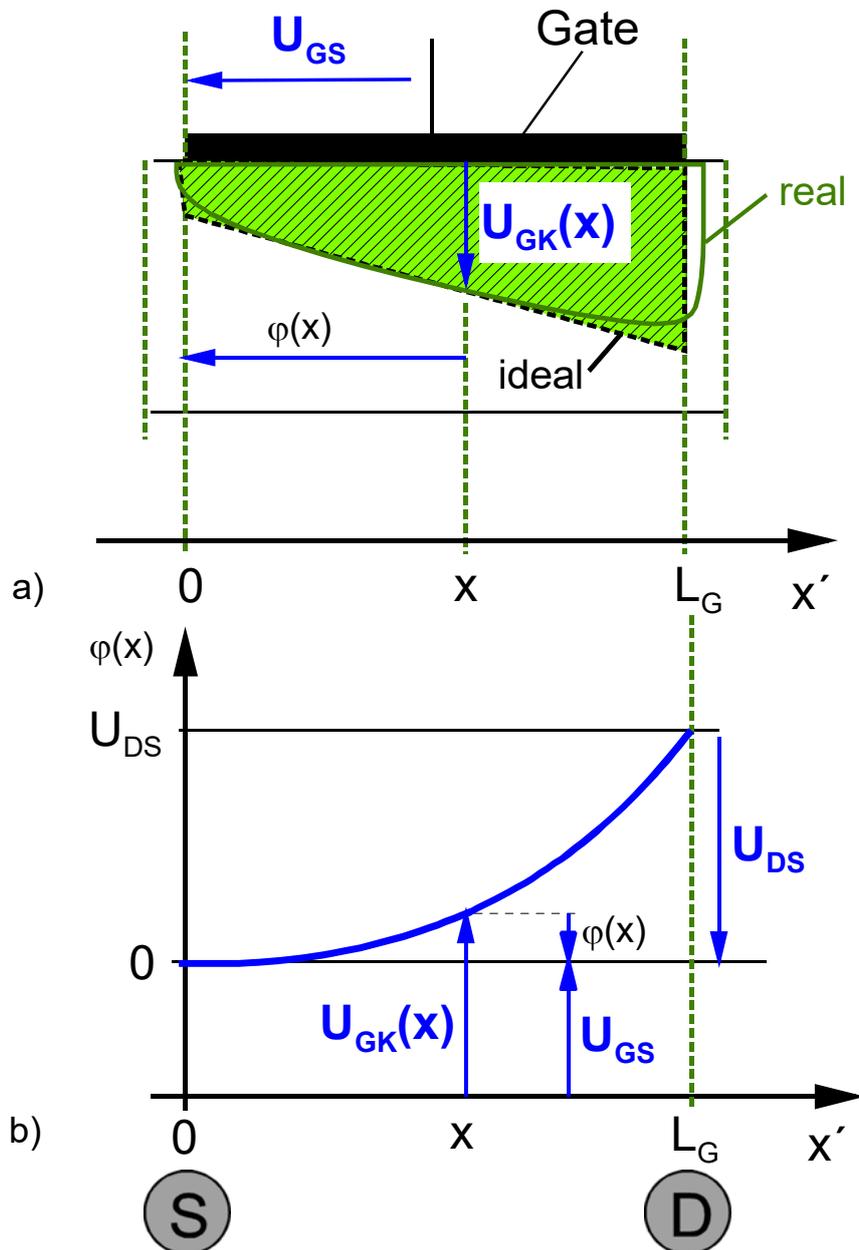


Abb. 2.29 Potenzialverteilung in einem Sperrschicht-FET bzw. MESFET

Die Dicke der Raumladungszone $d_{RLZ}(x)$ lässt sich berechnen mit (s. Vorlesung Festkörperelektronik):

$$d_{RLZ}(x) = \sqrt{\frac{2\epsilon_0\epsilon_r}{qN_D}} \cdot \sqrt{-U_{GK}(x) + U_D} \quad \text{mit } U_D > 0$$

bzw.:

$$d_{RLZ}(x) = B \cdot \sqrt{-U_{GK}(x) + U_D} \quad \text{mit} \quad B = \sqrt{\frac{2\varepsilon_0\varepsilon_r}{qN_D}} \quad (2.31)$$

Dabei bezeichnet $U_{GK}(x)$ die Spannung zwischen dem Gate und dem Kanal an der Stelle x :

$$-U_{GK}(x) = -U_{GS} + \varphi(x) \quad (2.32)$$

Das Potenzial $\varphi(x)$ im Kanal an der Stelle x erhält man aus:

$$\varphi(x) = \int_{R(0)}^{R(x)} I_D dR(\xi) = I_D \int_{R(0)}^{R(x)} dR(\xi) \quad (2.33)$$

Einsetzen in Gl. (2.32) ergibt:

$$-U_{GK}(x) = -U_{GS} + I_D \int_{R(0)}^{R(x)} dR(\xi) \quad (2.34)$$

und mit Gl. (2.30), (2.31) und $h(x) = d_K - d_{RLZ}(x)$ folgt:

$$-U_{GK}(x) = -U_{GS} + I_D \int_0^x \frac{1}{qN_D \mu Z h(\xi)} d\xi$$

bzw.:

$$-U_{GK}(x) = -U_{GS} + \frac{I_D}{qN_D \mu Z} \int_0^x \frac{1}{(d_K - B \cdot \sqrt{-U_{GK}(\xi) + U_D})} d\xi \quad (2.35)$$

Gl. (2.35) ist eine Integralgleichung für $U_{GK}(x)$. Differenzieren und Trennen der Variablen dU_{GK} , dx ergibt:

$$\frac{d(-U_{GK})}{dx} = \frac{I_D}{qN_D \mu Z} \cdot \frac{1}{d_K - B \cdot \sqrt{-U_{GK}(x) + U_D}}$$

$$d(-U_{GK}(x)) \left[d_K - B \cdot \sqrt{-U_{GK}(x) + U_D} \right] = \frac{I_D}{qN_D \mu Z} dx \quad (2.36)$$

Dabei sind die Randbedingungen entsprechend Gl. (2.32) für die Integration:

$$\begin{aligned} -U_{GK}(x=0) &= -U_{GS} \\ -U_{GK}(x=L_G) &= -U_{GS} + U_{DS} \end{aligned} \quad \text{vgl. Abb.(2.29b)}$$

Damit folgt aus Gl. (2.36):

$$d_K \int_{-U_{GS}}^{-U_{GS}+U_{DS}} d(-U_{GK}) - B \int_{-U_{GS}}^{-U_{GS}+U_{DS}} \sqrt{-U_{GK} + U_D} d(-U_{GK}) = \frac{I_D}{qn\mu Z} \int_0^{L_G} dx \quad (2.37)$$

mit

$$y = -U_{GK}$$

$$a = 1$$

$$b = +U_D$$

$$n = \frac{1}{2}$$

Das zweite Integral hat die Form:

$$\int_1^2 (ay + b)^n dy = \frac{1}{a(n+1)} (ay + b)^{n+1} \Big|_1^2$$

Damit ergibt sich:

$$d_K U_{DS} - \frac{2}{3} B (-U_{GK} + U_D)^{3/2} \Big|_{-U_{GS}}^{-U_{GS}+U_{DS}} = \frac{I_D L_G}{qN_D \mu Z} \quad (2.38)$$

Nach Einsetzen der Grenzen und Auflösen nach I_D folgt:

$$I_D = G_{K0} \left\{ U_{DS} - \frac{2}{3} \frac{B}{d_K} \left[(-U_{GS} + U_D + U_{DS})^{3/2} - (-U_{GS} + U_D)^{3/2} \right] \right\} \quad (2.39)$$

mit

$$G_{K0} = \frac{qN_D \mu d_K Z}{L_G} = \frac{\kappa \cdot A}{L_G}$$

κ = Leitfähigkeit des Kanals

A = Querschnitt der leitenden Schicht unter dem Gate-Kontakt (2.40)

1. Anlaufbereich:

Für den Betrieb des D-MESFET bei

$$|U_{DS}| \ll |-U_{GS} + U_D|$$

folgt damit:

$$I_D = G_{K0} \cdot U_{DS} \quad (2.41)$$

2. Sättigung des Stromes I_D :

Die Sättigung des Stromes I_D tritt ein, wenn sich an der Stelle $x = L_G$ (Drain) die Raumladungszone über die gesamte Kanaldicke erstreckt (Abb. 2.29):

Dann ist:

$$d_{RLZ}(x = L_G) = d_K \quad (2.42)$$

Die dazu erforderliche Drainspannung sei die Sättigungsspannung $U_{DS,sat}$.

Die Spannung, die dabei über der Raumladungszone abfällt und die notwendig ist, den Kanal zu verarmen, wird als Pinch-Off-Spannung U_P bezeichnet (s. Anmerkung Kap. 3.1.4.2).

Dabei gilt entsprechend Gl. (2.31) am Ende des Kanals bei $x = L_G$:

$$d_{RLZ} = d_K = B \sqrt{-U_{GK0} + U_D} = B \cdot \sqrt{U_P}$$

d.h.:

$$U_P = \frac{qN_D d_K^2}{2\varepsilon_0 \varepsilon_r} \quad \text{bzw.:} \quad U_P = \left(\frac{d_K}{B} \right)^2 \quad (2.43)$$

Aufgrund eines Spannungsumlaufes (Abb. 2.29b) ergibt sich für die Sättigungsspannung:

$$U_{DS,sat} = U_{GS} - U_{GK0} = U_{GS} - U_D + U_P = U_{GS} - U_T \quad (2.44)$$

Aufgrund von Gl. (2.39) und Gl. (2.44)

$$I_{D,sat} = G_{K0} \left\{ U_{DS,sat} - \frac{2}{3} \frac{B}{d_K} \left[(-U_{GS} + U_D + U_{DS,sat})^{3/2} - (-U_{GS} + U_D)^{3/2} \right] \right\} \quad (2.45)$$

folgt mit

$$U_P = \left(\frac{d_K}{B} \right)^2$$

$$I_{D,sat} = \frac{G_{K0} \cdot U_P}{3} \left\{ 3 \frac{U_{GS} - U_D + U_P}{U_P} - 2 \left(\frac{-U_{GS} + U_D + U_{GS} - U_D + U_P}{U_P} \right)^{3/2} + 2 \left(\frac{-U_{GS} + U_D}{U_P} \right)^{3/2} \right\} \quad (2.46)$$

und mit

$$I_{K0} = \frac{G_{K0} U_P}{3} \quad (2.47)$$

$$I_{D,sat} = I_{K0} \left\{ 1 - 3 \left(\frac{-U_{GS} + U_D}{U_P} \right) + 2 \left(\frac{-U_{GS} + U_D}{U_P} \right)^{3/2} \right\} \quad (2.48)$$

2.3 Der Heterostruktur-Feldeffekttransistor (HFET)

Neben den bereits behandelten Transistor-Typen MOSFET und MESFET gewinnt der sogenannte Heterostruktur-Feldeffekttransistor (HFET) immer mehr an Bedeutung. Er besteht aus mindestens zwei Halbleitern mit unterschiedlichem Bandabstand.

2.3.1 Der Heteroübergang

Das derzeit am meisten eingesetzte Heterosystem ist das $\text{Al}_x\text{Ga}_{1-x}\text{As}/\text{GaAs}$ -System. Dabei bezeichnet der Index x den prozentualen Al-Gehalt, d.h. die Anzahl der Al-Atome, die bei der Herstellung des Halbleitermaterials statt der Ga-Atome eingebaut werden.

Durch den Al-Gehalt lässt sich der Bandabstand der AlGaAs-Schicht einstellen.

Dabei gilt näherungsweise:

$$W_{g,\text{AlGaAs}} \approx W_{g,\text{GaAs}} + 1,247 \cdot x \quad \text{eV} \quad \text{für } x < 0,45$$

mit

$$W_{g,\text{GaAs}} = 1,42 \text{ eV} \quad (2.49)$$

Die AlGaAs-Schicht sei stark n-dotiert (Dotierstoffschicht) während die GaAs-Schicht undotiert bzw. leicht p-dotiert (Buffer) ist. Bringt man zwei Halbleiter mit unterschiedlichem Bandabstand (Abb. 2.25a) zusammen, so führt dies im Bändermodell am Grenzübergang zu einer Unstetigkeitsstelle. Im Fall des AlGaAs/GaAs-Systems entsteht dadurch an der Grenzschicht auf der AlGaAs-Seite eine Bandaufwölbung und somit eine Verarmungszone. Auf der GaAs-Seite bildet sich zwischen dem dotierten AlGaAs und dem undotierten GaAs ein **Potenzialtopf** aus. Damit werden im undotierten Material besetzbare Energiezustände unterhalb des Fermi-niveaus geschaffen, die sogenannten **Subbänder**, die durch die freien Elektronen aus dem **dotierten** Material (AlGaAs) besetzt werden können (Abb. 2.30b). Dies entspricht in etwa der Inversionsrandschicht des MOSFET, in der der Ladungstransport stattfindet. Da der Bereich, in dem sich die Ladungsträger aufhalten können, nur sehr schmal ($\approx 10\text{nm}$) ist, spricht man vom **zweidimensionalen Elektronengas (2DEG)**.

Zur Beschreibung des Bänderschemas am Heteroübergang existieren mehrere unterschiedliche Modelle. Ein Modell, mit dessen Hilfe sich die Leitungsbanddiskontinuität ΔW_L berechnen lässt, ist das **ANDERSON-Modell**.

Dabei lässt sich der Leitungsbandsprung aus den unterschiedlich großen Elektronenaffinitäten X_1 und X_2 berechnen. Das Vakuum-Niveau wird dabei als Bezugspunkt gewählt (Abb. 2.30). $q \cdot X_1$ und $q \cdot X_2$ werden als die Energien definiert, die notwendig sind, Elektronen von der Leitungsbandunterkante bis ins Vakuum-Niveau anzuheben.

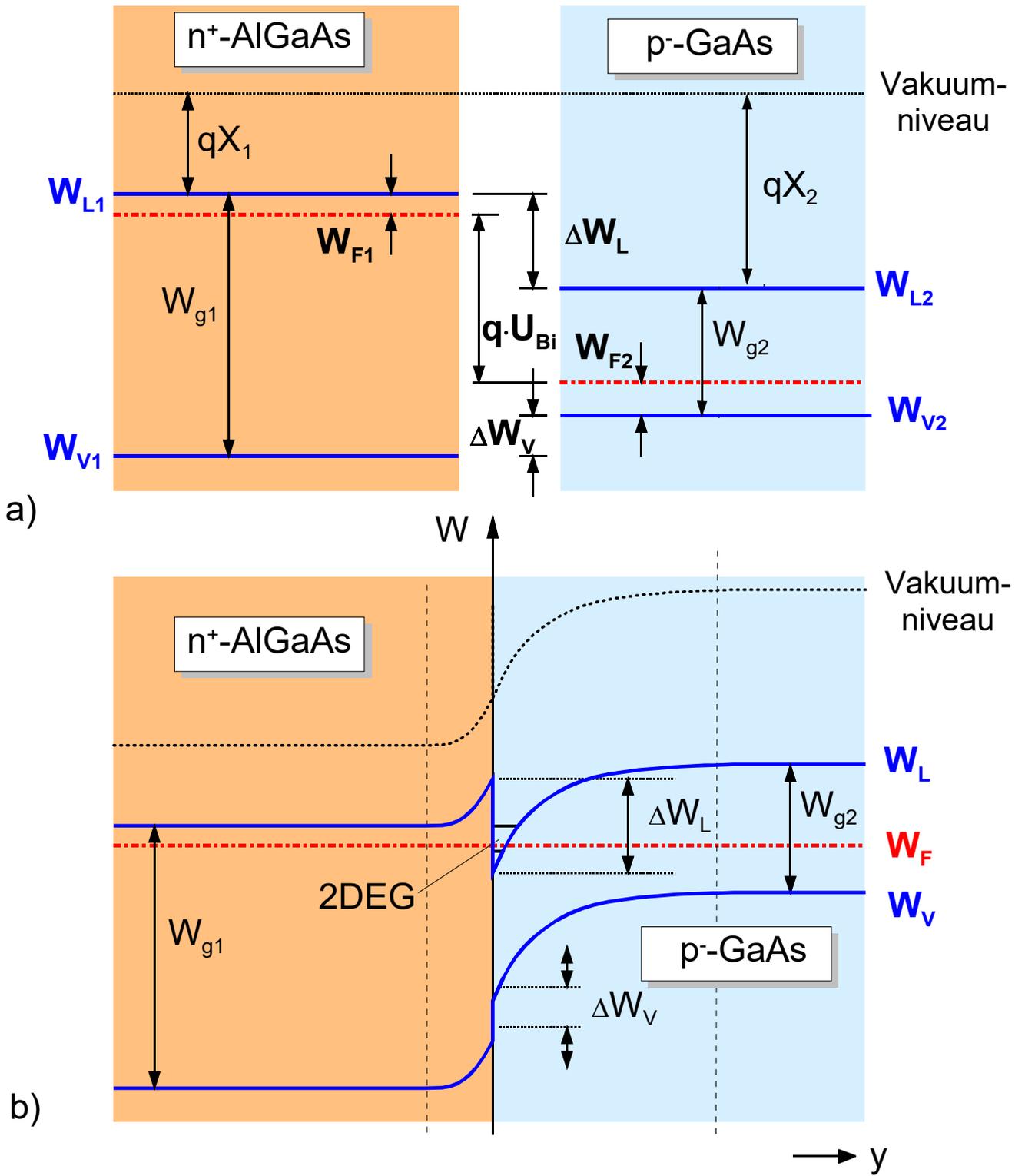


Abb. 2.30 Bänderdiagramm für zwei verschiedene Halbleiter: n -AlGaAs und p -GaAs.

(Das Vakuum-Niveau wird als Referenz gewählt)

a) die Halbleiter sind getrennt

b) die Halbleiter sind verbunden

Danach gilt:

$$\Delta W_L = q(X_2 - X_1) \quad (2.50)$$

Für die Valenzband-Diskontinuität ΔW_V ergibt sich damit:

$$\Delta W_V = \Delta W_g - \Delta W_L \quad (2.51)$$

mit

$$\Delta W_g = W_{g,AlGaAs} - W_{g,GaAs}$$

Die Elektronenaffinität X und damit auch der Bandsprung ΔW_L hängen vom Al-Gehalt x ab. Nach dem ANDERSON-Modell ergibt sich näherungsweise:

$$\Delta W_L \approx 0,85 \cdot \Delta W_g \quad (2.52)$$

und mit Gl. (2.49)

$$\Delta W_L \approx 1,06eV \cdot x \quad (2.53)$$

(In der Literatur sind auch von dem ANDERSON-Modell leicht abweichende Zahlenwerte für das Verhältnis von Leitungsbandsprung ΔW_L und der Differenz der Bandabstände ΔW_g bekannt.)

Damit ergibt sich bei einem Heteroübergang $Al_{0,3}Ga_{0,7}As/GaAs$ für den Leitungsbandsprung ΔW_L :

$$\Delta W_L \approx 300meV \quad (2.54)$$

2.3.2 Der Ladungstransport im 2DEG

An der Heterogrenzfläche im undotierten GaAs bildet sich, wie bereits gezeigt wurde, ein Potenzialtopf aus. In diesem Potenzialtopf existieren Energiezustände, die sogenannten Subbänder, die von den Elektronen aus der AlGaAs-Schicht besetzt werden können.

Da der Ladungstransport damit im undotierten Material (hier: GaAs) stattfindet, ist auch die Beweglichkeit der freien Elektronen im 2DEG bereits bei Raumtemperatur ($T=300K$) wesentlich höher als im dotierten Material.

Zahlenbeispiel:

$$\mu_{GaAs} (N_D \approx 4 \cdot 10^{17} cm^{-3}) \approx 2500 cm^2/Vs$$

$$\mu_{2DEG} = \mu_{GaAs} (N_D \approx 1 \cdot 10^{14} cm^{-3}) \approx 8000 cm^2/Vs$$

Die Beweglichkeit im Halbleiter wird durch die im Material auftretenden Streuprozesse bestimmt. Diese hängen ihrerseits von der Temperatur ab. Bei Raumtemperatur wird das Ladungsträgerverhalten vor allem durch die **Phononenstreuung**, d.h. durch die Wechselwirkung der Elektronen mit dem Gitter, beeinträchtigt (s. Vorlesung *Festkörperelektronik*).

Bei Abnahme der Temperatur geht der Einfluß der Gitterstreuung aufgrund der geringeren thermischen Energie zurück und die Beweglichkeit steigt je nach Höhe der Dotierstoffkonzentration an. Bei niedrigen Temperaturen macht sich jedoch im dotierten Material zunehmend die **Coulombstreuung** bemerkbar. Das bedeutet, das durch die ionisierten Donatorrümpfe hervorgerufene elektrische Feld reduziert die Beweglichkeit der Ladungsträger drastisch. Aus

diesem Grund unterscheidet sich die Beweglichkeit im dotierten Material bei tiefen Temperaturen kaum von Werten bei Raumtemperatur.

In den Heterostrukturen jedoch sind die ionisierten Donatorrümpfe im AlGaAs von den freien Elektronen im 2DEG (im GaAs) räumlich voneinander getrennt. Damit lässt sich der Einfluß der Coulombstreuung drastisch verringern. Dadurch bedingt steigt die Ladungsträgerbeweglichkeit im 2DEG bei Reduktion der Temperatur stark an (Abb. 2.31).

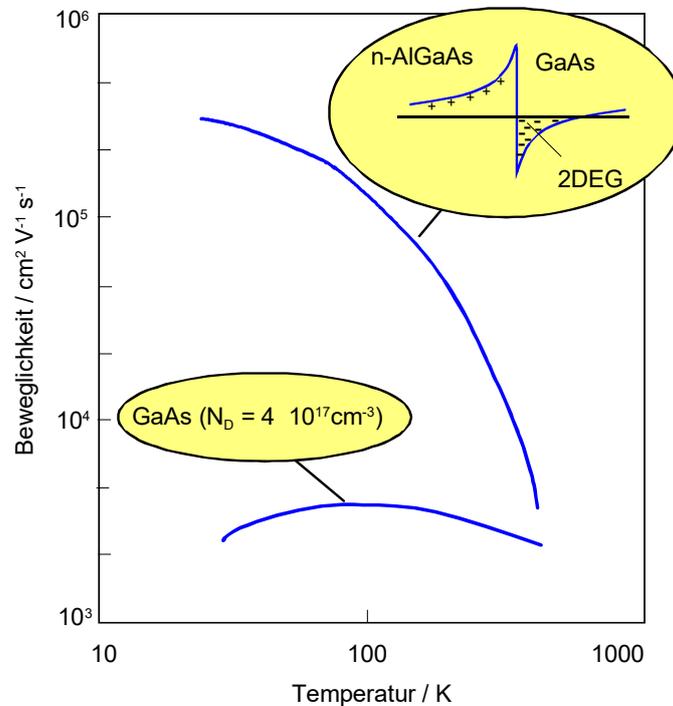


Abb. 2.31 Temperaturabhängigkeit der Ladungsträgerbeweglichkeit im dotierten GaAs und im 2DEG

2.3.3 Der Aufbau eines Heterostruktur-Feldeffekttransistors

Im zweidimensionalen Elektronengas lässt sich eine hohe Ladungsträgerkonzentration bei gleichzeitig hoher Ladungsträgerbeweglichkeit realisieren. Dies ist mit konventionellen dotierten Strukturen (z.B. MESFET) nicht möglich, da im dotierten Material die Beweglichkeit mit wachsender Dotierung abnimmt.

Werden auf der Halbleiteroberfläche zwei ohmsche Kontakte (Source und Drain) aufgebracht und einlegiert, fließt bei Anlegen einer Spannung zwischen diesen beiden Kontakten ein Strom durch das 2DEG. Wird zusätzlich in der Mitte ein sperrender Schottky-Kontakt (Gate) aufgebracht, so bildet sich auch in diesem Fall eine Raumladungszone (RLZ) unterhalb des Kontaktes aus. Über die Vorspannung an dieser Elektrode lässt sich die Ausdehnung der RLZ und auch die Ladungsträgerkonzentration und somit der Strom durch das 2DEG steuern. Abb. 2.32 zeigt den kompletten Aufbau eines Heterostruktur-Feldeffekttransistors (HFET) mit zugehörigem Bänderdiagramm.

Zwischen der undotierten GaAs-Schicht und dem dotierten AlGaAs wird in der Regel noch eine zusätzliche Zwischenschicht, der sogenannte spacer, eingebaut. Dieser sorgt dafür, dass die räumliche Trennung zwischen den ionisierten Donatorrümpfen im AlGaAs und den freien Ladungsträgern im Kanal noch größer wird. Dadurch lässt sich erreichen, dass gerade bei tiefen Temperaturen die Beweglichkeit noch stärker ansteigt. Je dicker diese Zwischenschicht allerdings ist, umso geringer wird die Ladungsträgerkonzentration im 2DEG. Für die Bauelementoptimierung ist daher ein vernünftiger Kompromiss zu finden. Für praktische Anwendungen liegt die Dicke der Zwischenschicht bei etwa $d_{sp} \approx 2 - 5 \text{ nm}$.

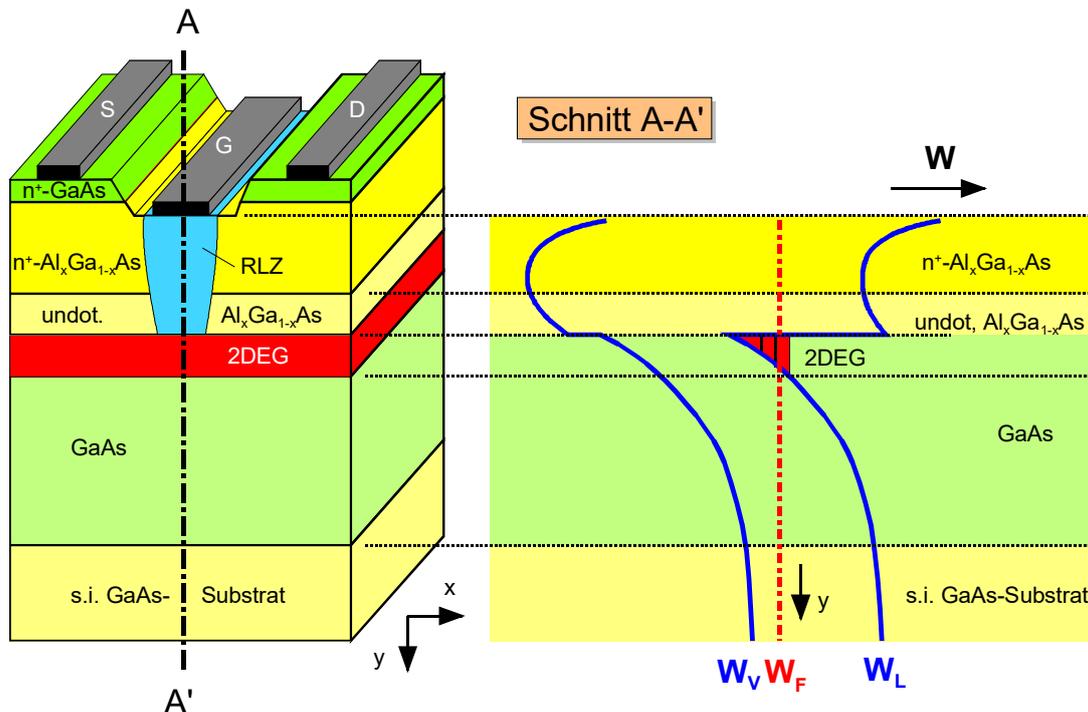


Abb. 2.32 Querschnitt durch einen AlGaAs/GaAs-Heterostruktur-FET und zugehöriges Bänderdiagramm (unterhalb des Gate-Kontakts)

Die **Vorteile** des HFET gegenüber konventionellen dotierten MESFET-Strukturen sind:

- höhere Verstärkungen aufgrund des besseren Transportverhaltens;
- geringeres Rauschen, da der Ladungstransport im undotierten Material stattfindet;
- Verbesserung der elektrischen Eigenschaften durch Abkühlen des Bauelements.

Der **Nachteil** liegt dabei in der aufwendigeren Technologie (z.B. Materialwachstum).

Die Heterostrukturen können mit Hilfe verschiedener Epitaxieverfahren (Epitaxie: Wachstum) hergestellt werden. Die derzeit am häufigsten eingesetzten Verfahren sind:

- Die Metallorganische Gasphasenepitaxie (MOVPE: Metall-Organic-Vapor-Phase-Epitaxy)
- Die Molekularstrahlepitaxie (MBE: Molecular-Beam-Epitaxy)

Abb. 2.33 zeigt die Aufnahme eines HFETs inkl. der Kontaktpads zur Kontaktierung des Bauelements.

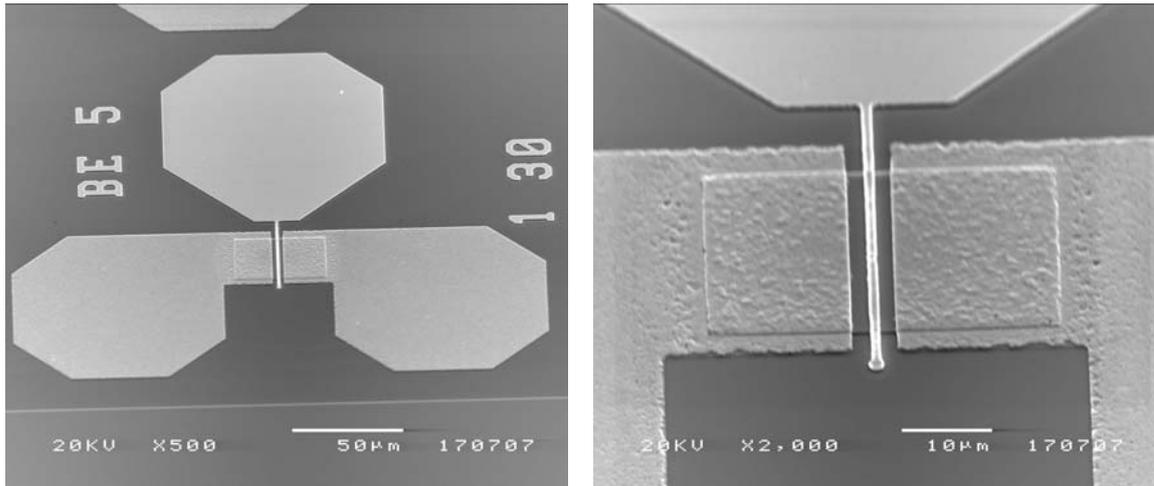


Abb. 2.33 a) im Fachgebiet Halbleitertechnik/Halbleitertechnologie der Universität Duisburg-Essen hergestellter InP HFET inkl. Kontaktpads von oben, b) Detailaufnahme des Gate-Bereichs

2.3.4 Berechnung der Schichtkonzentration

Da die räumliche Ausdehnung des Kanals (2DEG) sehr gering ist, spricht man nicht mehr von einer Volumenkonzentration, sondern von der **Schichtkonzentration** n_s .

Z.B. entspricht die Schichtkonzentration im MESFET-Kanal:

$$n_s(\text{MESFET}) = N_D \cdot d_K \quad [n_s] = \text{cm}^{-2}$$

Die Schichtkonzentration im 2DEG lässt sich mit Hilfe des sogenannten **charge control model** berechnen.

Dabei werden folgende Annahmen gemacht:

- Der Potenzialtopf an der Grenzfläche ist dreieckförmig, d.h., die Feldstärke im Potenzialtopf senkrecht zur Grenzfläche ist konstant.
- Es werden nur zwei quantisierte Zustände im Potenzialtopf (Subbänder) berücksichtigt.
- Die Konzentration der Ladungsträger im 2DEG entspricht im Gleichgewichtsfall der Raumladung aus der verarmten AlGaAs-Schicht an der Grenzfläche (Ladungsgleichgewicht).
- Die Fermi-Energie am Heteroübergang ist stetig.
- Es existieren keine Grenzflächenladungen am Heteroübergang.

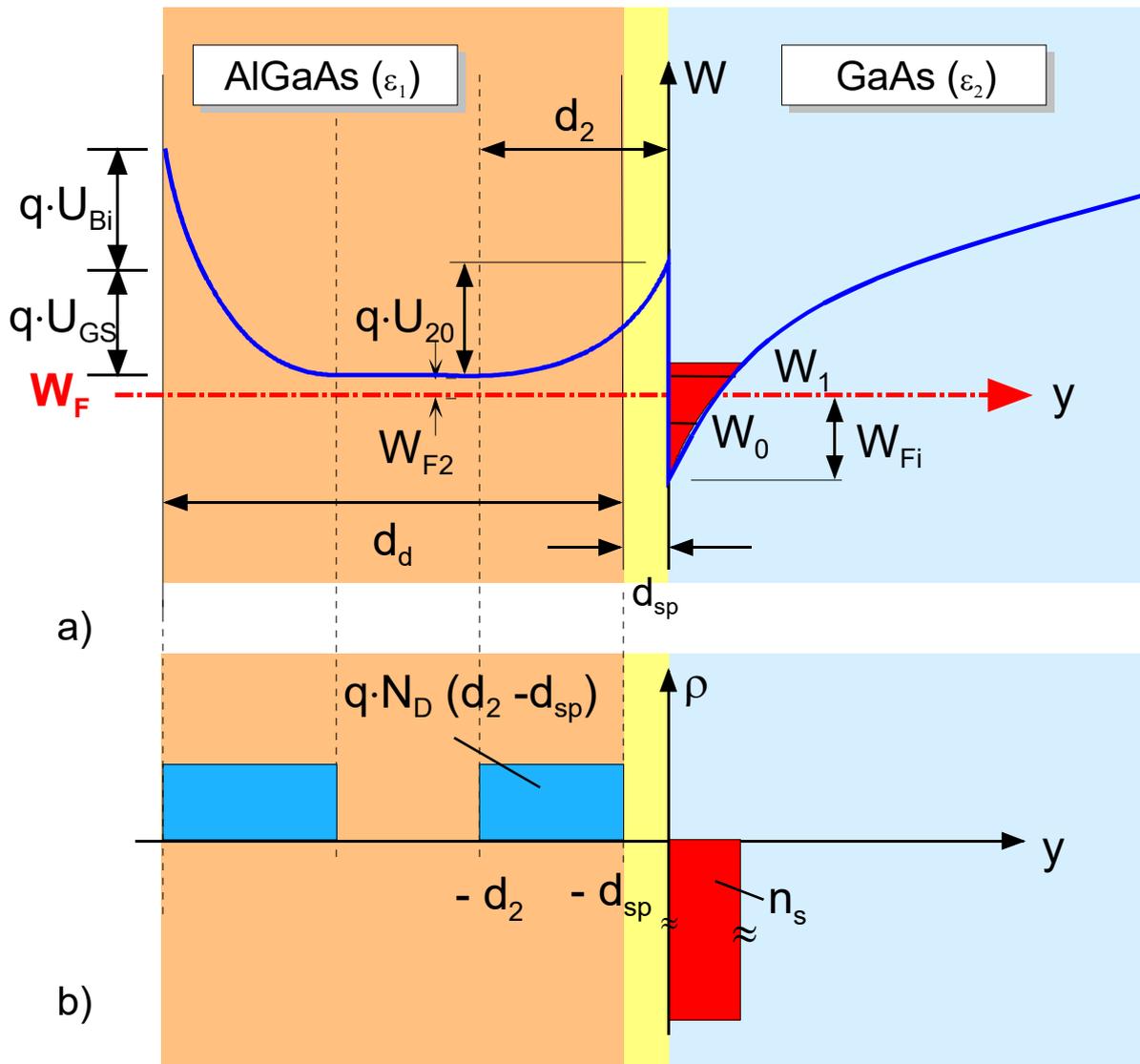


Abb. 2.34 a) Banddiagramm des Heteroüberganges unter Berücksichtigung des Schottky-Kontaktes an der Oberfläche und
b) Ladungsverteilung [3]

2.3.4.1 Der Gleichgewichtsfall

Zunächst sei vorausgesetzt, dass die Raumladungszone aufgrund des Schottky-Kontaktes an der Oberfläche noch nicht den Bandverlauf am Heteroübergang beeinflusst. Damit kann diese zunächst vernachlässigt werden. Es wird daher zunächst nur der Heteroübergang an sich betrachtet.

Wird vorausgesetzt, dass keine Grenzflächenladungen auftreten, muss die Normalkomponente der Verschiebungsdichte an der Grenzfläche stetig sein.

Das heißt, es muss gelten:

$$\varepsilon_1 E_{10} = \varepsilon_2 E_{20} \quad (2.55)$$

E_{10} und E_{20} sind die Feldstärken direkt am Heteroübergang im AlGaAs bzw. im GaAs senkrecht zur Grenzfläche.

Werden nur zwei diskrete Energieniveaus W_0 und W_1 im 2DEG (s. Abb. 2.34) betrachtet, so lässt sich die Schichtkonzentration der Ladungsträger im 2DEG angeben mit:

$$n_s = \frac{DkT}{q} \ln \left[\left(1 + \exp \left(\frac{W_{Fi} - W_0}{kT} \right) \right) \left(1 + \exp \left(\frac{W_{Fi} - W_1}{kT} \right) \right) \right] \quad (2.56)$$

mit

$$D = \frac{qm^*}{\pi \hbar^2} \quad \text{D: Zustandsdichte eines einzelnen Energieniveaus im 2DEG}$$

$$= 3,24 \cdot 10^{13} \text{ cm}^{-2} \text{ V}^{-1}$$

W_{Fi} bezeichnet den Abstand zwischen Fermi-Niveau und Leitungsbandkante im Potenzialtopf.

Die Lage der Energieniveaus W_0 und W_1 hängt von der Höhe der Schichtkonzentration ab mit:

$$W_0 = \gamma_0 n_s^{2/3} \approx 1,16 \cdot 10^{-9} \text{ eV cm}^{4/3} \cdot n_s^{2/3} \quad (2.57)$$

$$W_1 = \gamma_1 n_s^{2/3} \approx 1,48 \cdot 10^{-9} \text{ eV cm}^{4/3} \cdot n_s^{2/3}$$

Die Spannung U_{20} , die über der Raumladungszone in der AlGaAs-Schicht am Heteroübergang abfällt, lässt sich aus der Poisson-Gleichung bestimmen.

Im AlGaAs gilt:

$$\Delta \varphi = -\frac{\rho}{\varepsilon_1} = \frac{qN_D}{\varepsilon_1} \quad (2.58)$$

An dieser Stelle soll nur der eindimensionale Fall betrachtet werden. Damit folgt:

$$\frac{dE}{dy} = \frac{qN_D}{\varepsilon_1}$$

bzw.

$$E_y = \frac{qN_D}{\varepsilon_1} y \quad (2.59)$$

und nach einer weiteren Integration in den Grenzen zwischen der Zwischenschicht ($-d_{sp}$) und dem Ende der Raumladungszone ($-d_2$)

$$U_{20} = \int_{-d_{sp}}^{-d_2} \frac{qN_D}{\varepsilon_1} \cdot y dy$$

bzw.

$$U_{20} = \frac{qN_D}{2\varepsilon_1} (d_2^2 - d_{sp}^2) \quad (2.60)$$

(Bei dieser Berechnung wurde allerdings der Spannungsabfall über der undotierten Zwischenschicht der Dicke d_{sp} vernachlässigt.)

Da Ladungsgleichgewicht vorausgesetzt wird (s. Annahmen), gilt:

$$qn_s = qN_D (d_2 - d_{sp}) \quad (2.61)$$

Aus Gl. (2.60) ergibt sich für die Ausdehnung der Raumladungszone d_2 :

$$d_2 = \sqrt{\frac{2U_{20}\varepsilon_1}{qN_D} + d_{sp}^2} \quad (2.62)$$

Einsetzen in Gl. (2.61) führt zu:

$$qn_s = qN_D \left(\sqrt{\frac{2U_{20}\varepsilon_1}{qN_D} + d_{sp}^2} - d_{sp} \right) \quad (2.63)$$

und damit zu der Schichtkonzentration:

$$n_s = \frac{1}{q} \sqrt{2\varepsilon_1 qN_D U_{20} + q^2 N_D^2 d_{sp}^2} - N_D d_{sp} \quad (2.64)$$

Die Spannung U_{20} lässt sich aus einem Energieumlauf (Abb. 2.34) ermitteln:

$$qU_{20} = \Delta W_L - W_{F2} - W_{Fi} \quad (2.65)$$

und damit ergibt sich:

$$n_s = \frac{1}{q} \sqrt{2\varepsilon_1 qN_D (\Delta W_L - W_{F2} - W_{Fi}) + q^2 N_D^2 d_{sp}^2} - N_D d_{sp} \quad (2.66)$$

Mit Gl. (2.56) und Gl. (2.66) existieren zwei Bestimmungsgleichungen für die Schichtkonzentration n_s . Die Energie W_{Fi} , die wiederum von n_s abhängt, ist in beiden Gleichungen die unbekannte Größe. Durch z.B. ein iteratives Lösungsverfahren lässt sich damit der Zusammenhang zwischen n_s und W_{Fi} ermitteln (Abb. 2.34)

Der Zusammenhang zwischen W_{Fi} und n_s kann aber auch in weiten Bereichen für n_s linear angenähert werden mit:

$$W_{Fi} = W_{F0}(T) + \alpha \cdot n_s$$

mit

$$\begin{aligned} \alpha &\approx 0,125 \cdot 10^{-12} \text{ eV cm}^2 \\ W_{F0}(300\text{K}) &\approx 0 \\ W_{F0}(77\text{K}) &\approx 0,025 \text{ eV} \end{aligned} \quad (2.67)$$

Wie Gl. (2.66) zeigt, hängt die Schichtkonzentration vor allem von der Dotierstoffkonzentration in der AlGaAs-Schicht und der Spacerdicke ab (Abb. 2.36).

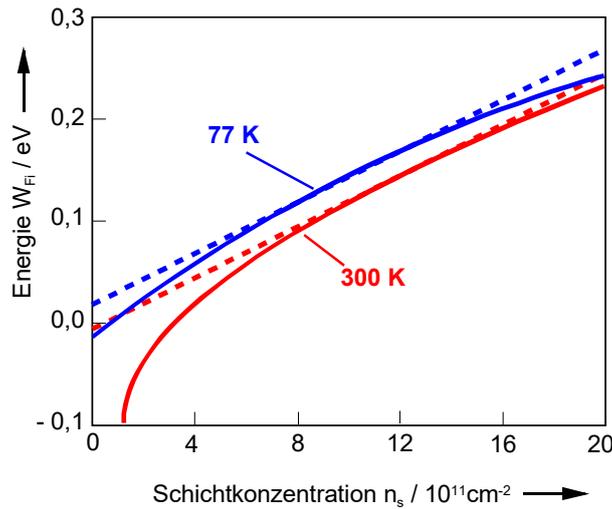


Abb. 2.35 Die Energie W_{Fi} in Abhängigkeit von der Schichtkonzentration mit der Temperatur T als Parameter [3]

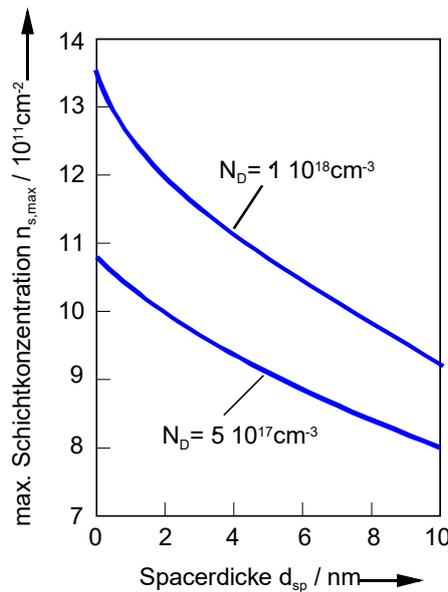


Abb. 2.36 Maximale Schichtkonzentration n_s als Funktion der Spacerdicke d_{sp} mit der Dotierstoffkonzentration in der AlGaAs-Schicht als Parameter [3]

Dabei wird deutlich:

- je höher die Dotierstoffkonzentration, desto größer wird die Schichtkonzentration
- je größer die Spacerdicke d_{sp} wird, desto kleiner wird die Schichtkonzentration

Wie bereits erwähnt besitzt der Kanal eine gewisse, wenn auch nur sehr kleine Ausdehnung. Dadurch bedingt, liegt der Ladungsschwerpunkt der Ladungen im 2DEG nicht direkt an der Heterogrenzfläche. Diese Verschiebung wird in den einfachen Modellen durch die Größe Δd mit

$$\Delta d \approx \frac{\epsilon_1 \alpha}{q^2}$$

berücksichtigt. Damit ändert sich Gl. (2.66) zu:

$$n_s = \frac{1}{q} \sqrt{2\varepsilon_1 q N_D U_{20} + q^2 N_D^2 (d_{sp} + \Delta d)^2} - q N_D (d_{sp} + \Delta d) \quad (2.68)$$

2.3.4.2 Der Einfluss des Schottky-Kontaktes auf die Schichtkonzentration

Wird die AlGaAs-Schichtdicke reduziert oder die Spannung U_{GS} (s. Abb. 2.34) derart verkleinert, dass sich die Raumladungszone aufgrund des Schottky-Kontaktes weiter in die AlGaAs-Schicht hinein ausdehnt, so überlappen sich ab einem bestimmten Arbeitspunkt die Verarmungsschichten aufgrund des Schottky-Kontaktes und aufgrund des Heteroübergangs.

Die Spannung, bei der dies geschieht, ist

$$U_{T1} = U_{Bi} - \frac{q N_D}{2\varepsilon_1} (d_d + d_{sp} - d_2)^2 \quad (2.69)$$

Für $U_{GS} > U_{T1}$ beeinflussen sich die beiden Raumladungszonen nicht gegenseitig. Die Schichtkonzentration im 2DEG ist maximal und unabhängig von der außen angelegten Spannung.

Für $U_{GS} < U_{T1}$ ist die AlGaAs-Schicht vollkommen verarmt und die Gate-Spannung steuert auch den Bandverlauf am Heteroübergang (Abb. 2.37). Damit verringert sich die Konzentration der Ladungsträger im Kanal.

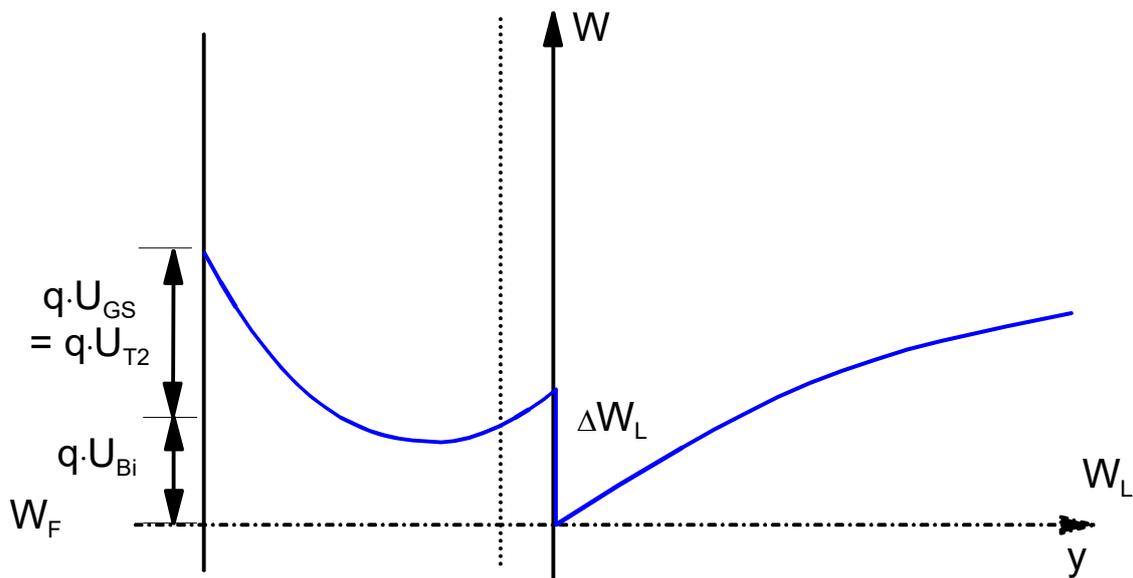


Abb. 2.37 Bandverlauf am Heteroübergang unter Berücksichtigung der Gate-Spannung für $U_{GS} < U_{T1}$

Wird die Spannung U_{GS} weiter vergrößert bis der Potentialtopf über das Fermi-Niveau angehoben wird, so existieren keine besetzten Zustände mehr an der Grenzfläche und der Kanal verschwindet. Die Abschnürspannung U_{T2} , bei der dieser Fall eintritt, ergibt sich aus einem Umlauf entsprechend Abb. 2.38:

$$U_{T2} = U_{Bi} - U_{P2} - \frac{1}{q} \Delta W_L \tag{2.70}$$

wobei U_{P2} die Spannung ist, die über der verarmten AlGaAs-Schicht abfällt:

$$U_{P2} = \frac{qN_D d_2^2}{2\epsilon_1} \tag{2.71}$$

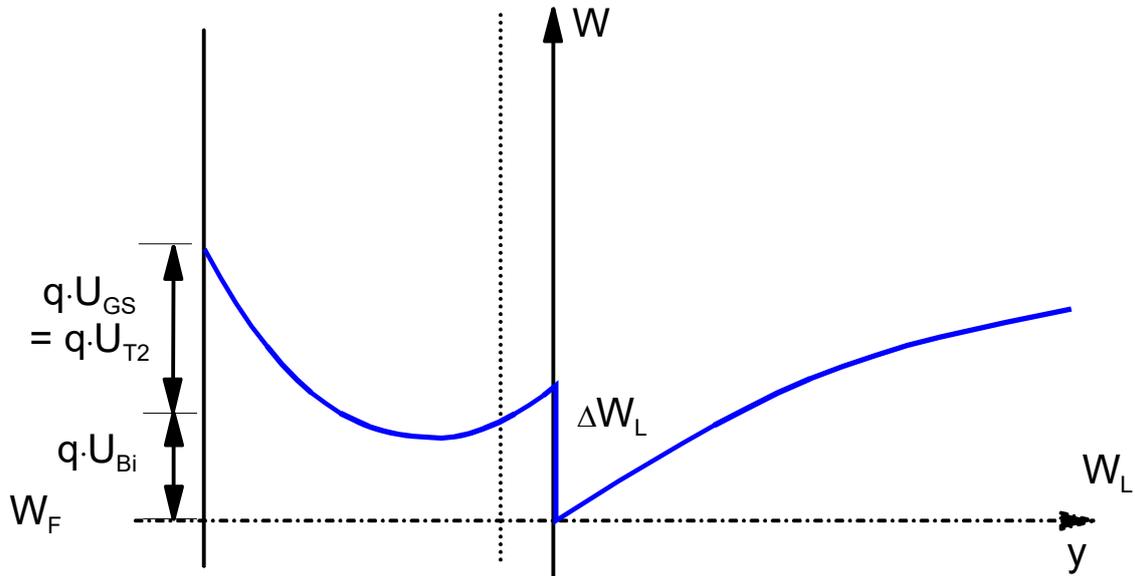


Abb. 2.38 Bandverlauf am Heteroübergang für den Abschneirfall $U_{GS} = U_{T2}$

Für den Spannungsbereich, in dem das 2DEG durch die Vorspannung am Schottky-Kontakt gesteuert wird, d.h. für $U_{T1} > U_{GS} > U_{T2}$, gilt:

$$n_s = \frac{\epsilon_0 \epsilon_r}{q(d_d + d_{sp} + \Delta d)} (U_{GS} - U_{T2}) \tag{2.72}$$

$d = d_d + d_{sp} + \Delta d$ ist dabei der Abstand des Ladungsschwerpunktes des 2DEG von der Oberfläche. Damit gibt Gl. (2.72) vereinfacht die induzierte Ladung in einem Plattenkondensator der Dicke d an, dessen beide Elektroden einerseits durch das 2DEG und andererseits durch den Schottky-Kontakt gebildet werden.

2.3.4.3 Gleichstromverhalten

Der Strom durch den HFET lässt sich analog zu den Herleitungen für den MOSFET (Kap.3.1.2) berechnen.

Für den Strom I_D gilt:

$$I_D = q n_s(x) v(x) Z \tag{2.73}$$

Die Schichtkonzentration n_s im Kanal hängt jetzt nicht nur von der Gate-Spannung U_{GS} , sondern von dem jeweiligen Potenzial $U_{GK}(x)$ im Kanal ab (s. Abb. 2.5). Das heißt, die Steuerspannung

U_{GS} in Gl. (2.73) muss durch die jeweilige Gate-Kanal-Spannung an der Stelle x im Kanal ersetzt werden.

Mit

$$v = \mu \cdot E = \mu \cdot \frac{dU(x)}{dx} \quad (2.74)$$

ergibt sich:

$$I_D = \frac{Z\mu\epsilon_1}{d_d + d_{sp} + \Delta_d} (U_{GS} - U(x) - U_{T2}) \frac{dU(x)}{dx} \quad (2.75)$$

Analog zu der Herleitung in Kap.3.1.1 lässt sich damit der Strom I_D angeben:

$$I_D = 2K \left[(U_{GS} - U_{T2}) U_{DS} - \frac{1}{2} U_{DS}^2 \right]$$

mit

$$2K = \frac{Z\mu\epsilon_1}{(d_d + d_{sp} + \Delta_d)L} \quad (2.76)$$

Die weiteren Berechnungen verlaufen analog zu den Herleitungen für den MOSFET (s.Kap.3).

2.4 Das Elektrische Ersatzschaltbild

Zur Optimierung der Bauelemente, vor allem aber für die Schaltungssimulation, versucht man das elektrische Verhalten z.B. der Feldeffekttransistoren durch Modelle bzw. elektrische Ersatzschaltbilder zu beschreiben. Dabei unterscheidet man zwischen

- **Gleichspannungs-** (DC-) und
- **Wechselspannungs-** (AC-) Verhalten.

Bei Letzterem differenziert man weiter zwischen **Kleinsignal-** und **Großsignal-** Verhalten. Im Rahmen dieser Vorlesung wird jedoch nur auf das Gleichspannungs- und das Kleinsignalverhalten sowie entsprechende Ersatzschaltbilder eingegangen.

Hinweis: Soweit dies möglich ist, werden im Folgenden

Gleichstrom- (DC-) Größen mit **großen Buchstaben** und **großen Indizes** (U_{DS}, U_{GS}, I_D),

Wechsel- (AC-) Größen mit **kleinen Buchstaben** und **kleinen Indizes** (u_{ds}, u_{gs}, i_d) bezeichnet

2.4.1 Das Gleichstrom- (DC-) Verhalten des realen Feldeffekttransistors

Bei realen Bauelementen unterscheidet man zwischen

- **intrinsischem** und
- **extrinsischem** Bereich.

Der **intrinsische** Bereich beschreibt das ideale Bauelement, wie es in den vorangegangenen Kapiteln behandelt wurde.

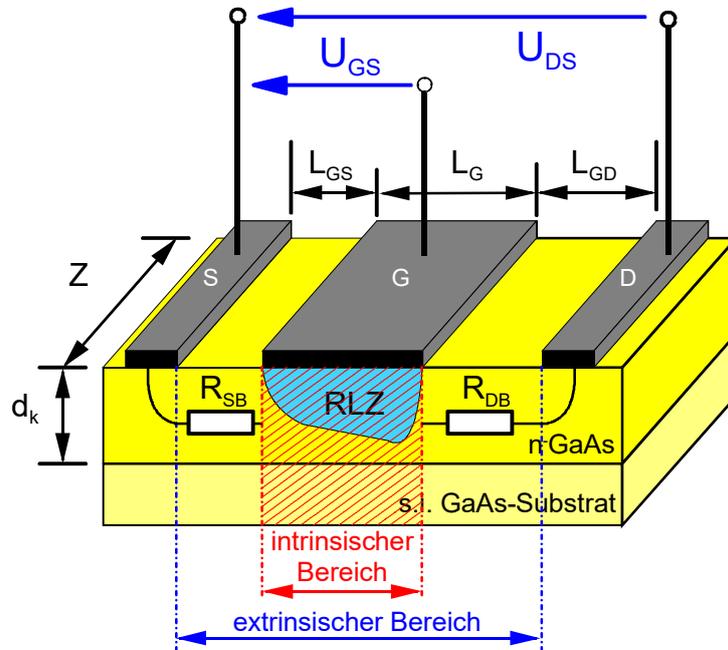


Abb. 2.39 Querschnitt durch einen "realen" MESFET mit Unterscheidung zwischen intrinsischem und extrinsischem Bereich

Der **extrinsische** Bereich umfasst zusätzlich die parasitären Größen, die bei realen Bauelementen berücksichtigt werden müssen. Diese sind oftmals durch die Technologie bzw. den Aufbau bedingt und lassen sich in der Realität nie ganz vermeiden.

Abb. 2.39 zeigt den prinzipiellen Aufbau eines realen MESFET. Aufgrund der Tatsache, dass die ohmschen Source- und Drain-Kontakte nicht direkt am eigentlichen FET (intrinsischer Bereich) angeschlossen werden können, muss der Einfluss der Zuleitungen über das Halbleitermaterial berücksichtigt werden. Diese können durch einen Source- und einen Drain-seitigen Bahnwiderstand R_{SB} bzw. R_{DB} beschrieben werden. Damit ergibt sich das in Abb. 2.40 dargestellte elektrische Ersatzschaltbild.

Für die Berechnung des Stroms I_D , wie in den vorangegangenen Kapiteln für die verschiedenen Bauelemente beschrieben, müssen die **intern wirksamen**, intrinsischen Spannungen $U_{DS,i}$ und $U_{GS,i}$ verwendet werden.

Sie können aufgrund der durch den Drain-Strom I_D verursachten Spannungsabfälle über den Bahnwiderständen stark von den **extern angelegten**, extrinsischen Spannungen $U_{DS,e}$ und $U_{GS,e}$ abweichen.

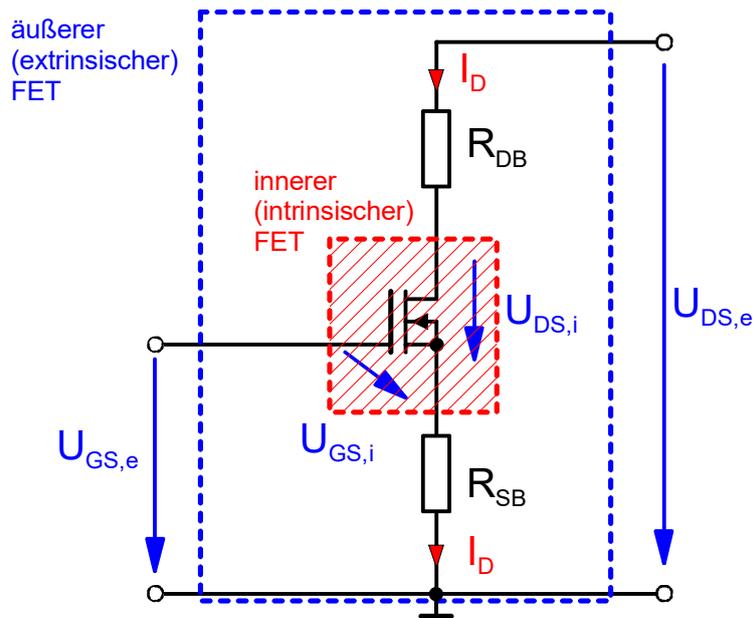


Abb. 2.40 Elektrisches Ersatzschaltbild des realen MESFET unter Berücksichtigung der durch die Zuleitungen bedingten Bahnwiderstände

Für den Zusammenhang zwischen extrinsischen und intrinsischen Spannungen gilt:

$$U_{DS,e} = I_D \cdot R_{DB} + U_{DS,i} + I_D \cdot R_{SB}$$

bzw.

$$U_{DS,e} = U_{DS,i} + I_D \cdot (R_{DB} + R_{SB}) \quad (2.77)$$

Damit folgt für die intrinsische Drain-Source-Spannung $U_{DS,i}$, die für die Berechnung des Stroms zugrunde gelegt werden muss:

$$U_{DS,i} = U_{DS,e} - I_D \cdot (R_{DB} + R_{SB}) \quad (2.78)$$

Reicht die innere Spannung $U_{DS,i}$ aus, den **intrinsischen** FET in Sättigung zu betreiben, so hat der zusätzliche Spannungsabfall über den Bahnwiderständen keine Auswirkung auf die Berechnung des Stroms, da - im Idealfall - in Sättigung $I_D \neq f(U_{DS,i})$ gilt.

Ganz anders verhält es sich dagegen in Bezug auf die innere Spannung $U_{GS,i}$:

Es gilt:

$$U_{GS,e} = U_{GS,i} + I_D \cdot R_{SB} \quad (2.79)$$

bzw.

$$U_{GS,i} = U_{GS,e} - I_D \cdot R_{SB} \quad (2.80)$$

Der zusätzliche Spannungsabfall über R_{SB} reduziert die intern wirksame Spannung $U_{GS,i}$, schnürt damit den Kanal zusätzlich ab und der resultierende Strom nimmt ab. Damit hat auch hier der Source-seitige Widerstand eine gegenkoppelnde Wirkung.

Um diesen gegenkoppelnden Einfluss möglichst gering zu halten, versucht man einerseits den Abstand zwischen Gate- und Source-Kontakt so klein wie möglich zu wählen. Aus technologischen Gründen lässt sich dieser jedoch nicht ganz vermeiden.

Daher wird die leitende Schicht dicker als für die Abschnürspannung notwendig hergestellt. Anschließend wird nur im Gate-Bereich die Schicht auf das notwendige Maß heruntergeätzt. Man spricht dabei von einem "versenkten Gate" oder "*recessed gate*" (s. Abb. 2.41).

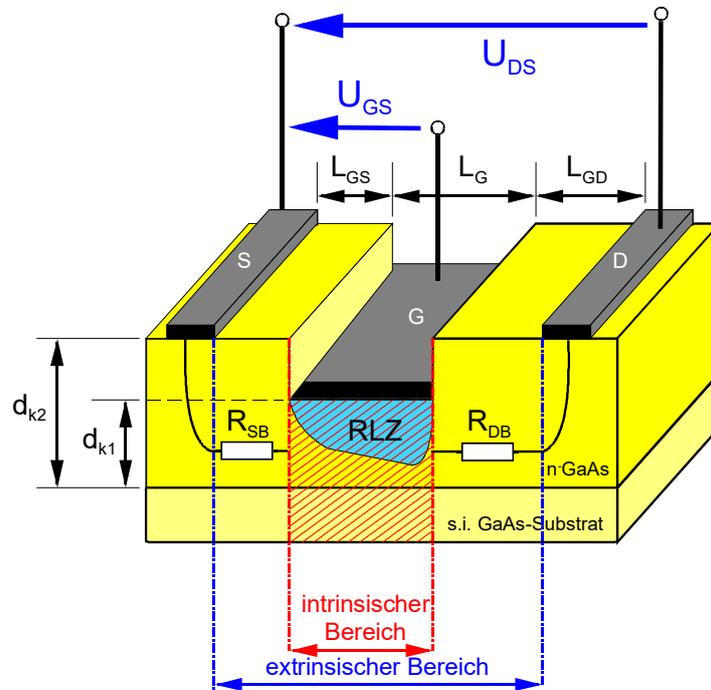


Abb. 2.41 Querschnitt durch einen MESFET mit "versenktem Gate" (*recessed gate*)

Abb. 2.42 zeigt eine Rasterelektronenmikroskopaufnahme (REM) des Gate-Bereichs eines HFETs mit versenktem Gate-Kontakt.

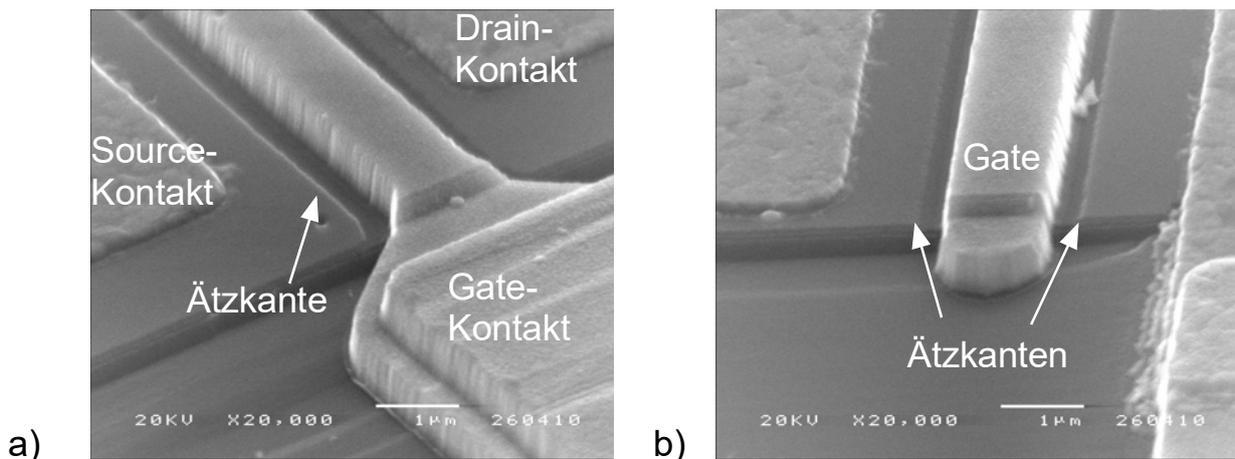


Abb. 2.42 Rasterelektronenmikroskop- (REM-)aufnahmen des Gate-Bereichs eines HFETs mit versenktem Gate-Kontakt a) vom Kontakt-Pad aus und b) vom Ende des Gate-Streifens aus gesehen

2.4.2 Das Wechselspannungs- (AC-) verhalten

2.4.2.1 Funktionsprinzip

Transistoren werden häufig in Verstärkern eingesetzt. Abb. 2.43 zeigt als Beispiel eine einfache Verstärkerschaltung mit einem n-Kanal Verarmungs-MOSFET (*depletion-type*, D-MOSFET) in Source-Schaltung.

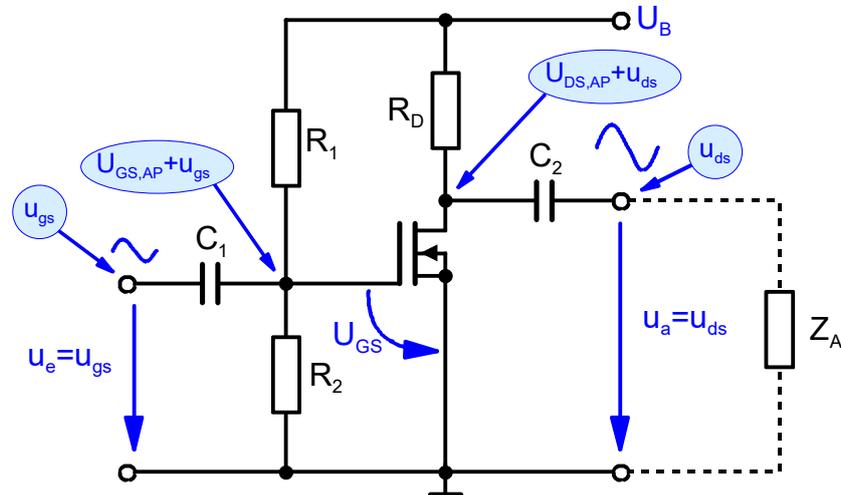


Abb. 2.43 Prinzipschaltung einer Verstärkerstufe (Inverter) mit einem n-Kanal MOSFET in Source-Schaltung

Der Arbeitspunkt (AP) des Transistors wird, wie in Kap. 2.1.6 beschrieben, zunächst durch die äußere Beschaltung und damit durch die Widerstände R_1 , R_2 und R_D festgelegt.

Kann der Gate-Strom vernachlässigt werden, d. h. ist $I_G = 0$, so wird die Gate-Source Spannung U_{GS} durch den (in diesem Fall) unbelasteten Spannungsteiler R_1 und R_2 eingestellt mit:

$$U_{GS} = U_B \cdot \frac{R_2}{R_1 + R_2} \quad (2.81a)$$

R_D bestimmt entsprechend

$$U_{DS} = U_B - I_D \cdot R_D \quad (2.81b)$$

die Arbeitsgerade (AG) (Abb. 2.44). Der resultierende Arbeitspunkt (AP) ergibt sich dann aus dem Schnittpunkt der Arbeitsgeraden mit dem zu U_{GS} gehörigen Kennlinienast.

Wird zusätzlich an den Eingang eine Wechselspannung u_e angelegt, so wird der Transistor um den Arbeitspunkt herum angesteuert: Eine positive Aussteuerung führt zu einem Anstieg des Drain-Stroms und somit zu einem größeren Spannungsabfall am Lastwiderstand R_D . Somit nimmt die Ausgangsspannung U_{DS} ab. Entsprechendes gilt bei negativer Aussteuerung. Ein- und Ausgangsspannung haben demnach bei dieser Verstärkerschaltung entgegengesetzte Polarität. Man spricht daher in diesem Fall von einem Inverter.

Da nur Wechselspannungen verstärkt werden sollen, sind zur Vermeidung ungewünschter DC-Koppeleffekte zwischen einzelnen Stufen zusätzliche Koppelkondensatoren C_1 und C_2 notwendig, die die Wechselspannung ein- und auskoppeln.

Diese Koppelkondensatoren sind in der Regel so groß ausgelegt, dass sie für den betrachteten Frequenzbereich als wechselfspannungsmäßige Kurzschlüsse anzunehmen sind und deren Einfluss daher meist bei den AC-Betrachtungen vernachlässigt werden kann.

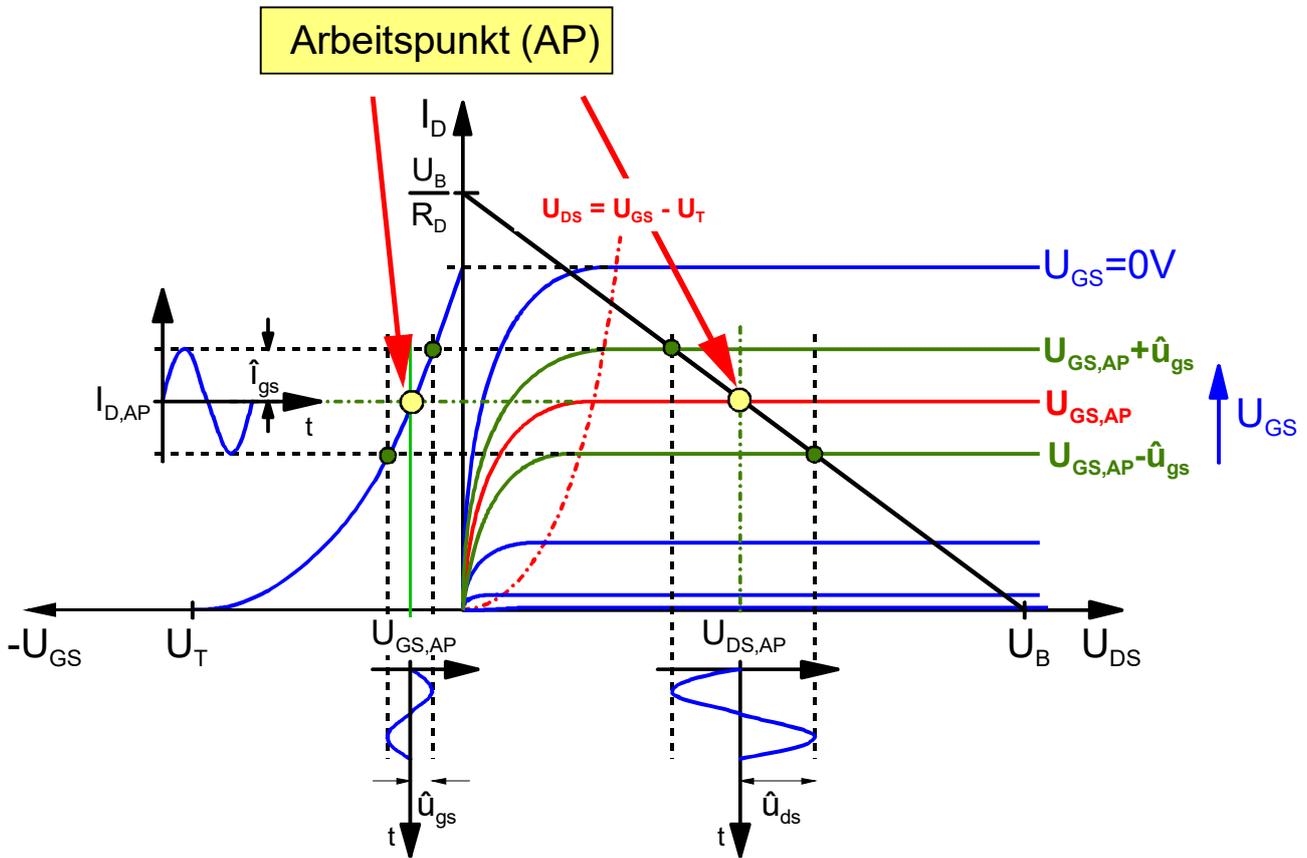


Abb. 2.44 Veranschaulichung des Verstärkungsvorgangs mittels der Kennlinienfelder

2.4.2.2 Allgemeine Beschreibung des Kleinsignalverhaltens

In der dargestellten Betriebsart wird dem gleichstrommäßig eingestellten Arbeitspunkt AP das zu verstärkende Signal in Form einer zeitlich variablen Größe (Spannung, Strom) überlagert (Abb. 2.44).

Bei beliebig großer Amplitude des Signals werden die im allgemeinen **nichtlinearen** Kennlinien der Transistoren angesteuert. Eine vollständige Berechnung des Bauelementverhaltens ist aufwändig und meist nur numerisch möglich. Die nichtlinearen Kennlinien führen außerdem zu einer **Verzerrung** des verstärkten Signals.

Ist die Amplitude des Signals "klein", so lässt sich das Verhalten **linearisieren**. Der Transistor kann durch eine Ersatzschaltung aus linearen Bauelementen dargestellt und durch Vierpolgleichungen beschrieben werden. Der Transistor verhält sich dann in diesem Arbeitspunkt wie ein **linearer aktiver Vierpol**.

Sind die Signalfrequenzen so niedrig, dass Kapazitäten und ggf. Induktivitäten sowie Ladungsträgerlaufzeiten im Transistor das Verhalten nicht beeinflussen, spricht man von **quasistationärem Verhalten** und dieses **Kleinsignalverhalten** kann im Arbeitspunkt durch Tangenten an die Kennlinie näherungsweise charakterisiert werden. Diese Näherung ist umso besser, je kleiner die Signalamplitude ist. (Abb. 2.45).

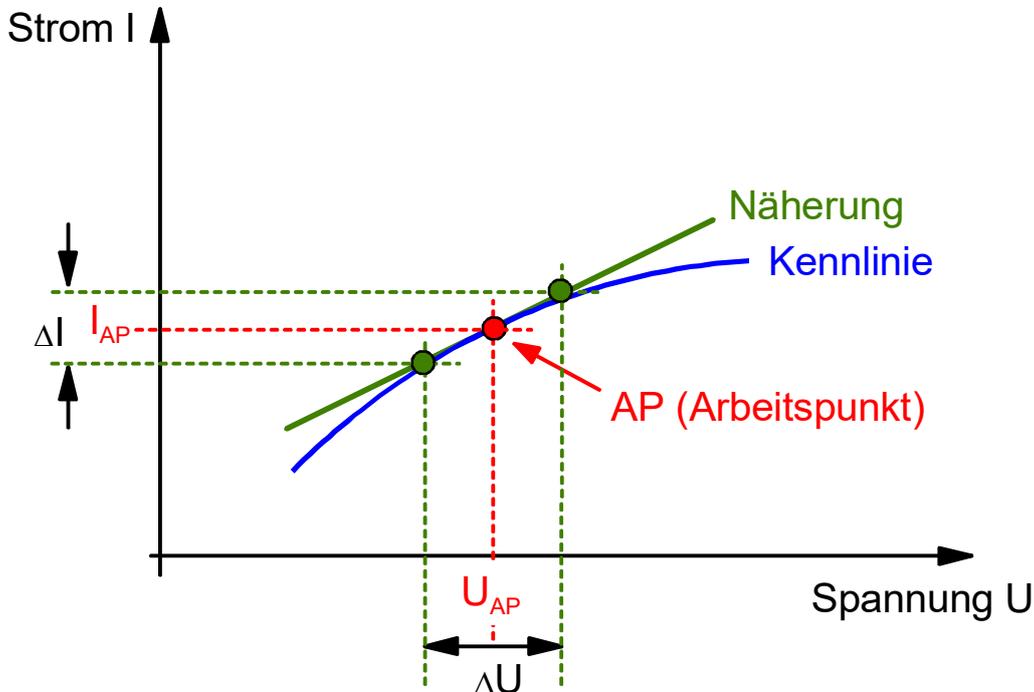


Abb. 2.45 Kennlinie und Näherung für Kleinsignalbetrieb

Im Fall des Feldeffekttransistors ist der Strom I_D eine Funktion von zwei variablen Spannungen U_{DS} und U_{GS} .

Mit Hilfe der Taylor-Reihenentwicklung lässt sich die Änderung eines Funktionswertes, der von zwei Variablen abhängt, bei Linearisierung darstellen durch:

$$f(x, y) = f(x_0 + \Delta x, y_0 + \Delta y) = f(x_0, y_0) + \left[\frac{\partial f}{\partial x} \Big|_{y_0} \cdot \Delta x + \frac{\partial f}{\partial y} \Big|_{x_0} \cdot \Delta y \right] \quad (2.87)$$

Damit lässt sich die Aussteuerung des Transistors um den Arbeitspunkt $f(x_0, y_0)$ durch arbeitspunkt**abhängige**, aber im Arbeitspunkt selbst konstante Parameter beschreiben. Da sowohl das DC- als auch das AC-Verhalten der Transistoren von mehreren Größen (ein- und ausgangsseitige Ströme und Spannungen) abhängt, benötigt man zur vollständigen Charakterisierung des Kleinsignalverhaltens entsprechend viele Parameter.

Der Transistor, der sich in einem festen Arbeitspunkt wie ein linearer, aktiver Vierpol verhält, kann durch **Vierpolparameter** beschrieben werden.

Für Transistoren hat sich die Darstellung des Vierpols durch

- **Leitwertparameter (y-Parameter)**
- **Hybrid-Parameter (h-Parameter)** oder
- **Streu-Parameter (s-Parameter)**

als zweckmäßig erwiesen.

Die **Leitwertform** wird meist bei Schaltungen mit Feldeffekttransistoren herangezogen. Ferner ist die Leitwertform bei Untersuchungen im Bereich hoher Frequenzen vorteilhaft.

Die **Hybridform** der Vierpolgleichungen wird hauptsächlich bei der Beschreibung von Schaltungen mit Bipolartransistoren angewandt.

Die Darstellung durch **s-Parameter** wird vor allem zur Charakterisierung des Hochfrequenzverhaltens von Transistoren eingesetzt.

Die verschiedenen Vierpolparameter werden in der Vorlesung "Elektronische Schaltungen" ausführlich behandelt.

Dabei ist zu beachten, dass alle Parameter **signalunabhängig** sind, der Wert der Parameter jedoch vom jeweiligen Arbeitspunkt und ggfs. der Frequenz abhängt!

2.4.2.3 Quasistationäres Kleinsignalverhalten von Feldeffekttransistoren

Für die Ausgangskennlinie des (MOS)FET in Source-Schaltungen gilt entsprechend Kap. 2.1.3.1:

$$I_D = I_D(U_{DS}, U_{GS})$$

Nimmt man nur eine sehr kleine Aussteuerung um den Arbeitspunkt herum an, so lässt sich die Änderung des Stroms ohne großen Fehler linearisieren. Man spricht in diesem Fall von Kleinsignalaussteuerung bzw. Kleinsignalverhalten:

Mit Gl. (2.87) folgt für den Stromwert in der Nähe des Arbeitspunktes:

$$I_D(U_{DS}, U_{GS}) = I_{D,AP}(U_{DS,AP}, U_{GS,AP}) + \left. \frac{\partial I_D}{\partial U_{GS}} \right|_{U_{DS,AP}=\text{konst.}} \cdot \Delta U_{GS} + \left. \frac{\partial I_D}{\partial U_{DS}} \right|_{U_{GS,AP}=\text{konst.}} \cdot \Delta U_{DS}$$

bzw. für die Änderung des Stroms:

$$\Delta I_D = \left. \frac{\partial I_D}{\partial U_{GS}} \right|_{U_{DS,AP}=\text{konst.}} \cdot \Delta U_{GS} + \left. \frac{\partial I_D}{\partial U_{DS}} \right|_{U_{GS,AP}=\text{konst.}} \cdot \Delta U_{DS}$$

oder abgekürzt:

$$\Delta I_D = a_1 \cdot \Delta U_{GS} + a_2 \cdot \Delta U_{DS}$$

mit

$$a_1 = \left. \frac{\partial I_D}{\partial U_{GS}} \right|_{U_{DS,AP}} = \text{konst.}$$

$$a_2 = \left. \frac{\partial I_D}{\partial U_{DS}} \right|_{U_{GS,AP}} = \text{konst.} \quad (2.88)$$

Bei sinusförmiger Aussteuerung entspricht die Änderung des Stromes ΔI_D der Amplitude des Wechselstromes \hat{i}_d und die Änderungen der Spannungen ΔU_{GS} bzw. ΔU_{DS} den Amplituden \hat{u}_{gs} bzw. \hat{u}_{ds} der Wechselspannungen.

Damit kann endgültig geschrieben werden:

$$i_d = a_1 \cdot u_{gs} + a_2 \cdot u_{ds} \quad (2.89)$$

Aus Gleichung (2.88) ist erkennbar, dass a_1 und a_2 die Dimension eines Leitwertes haben.

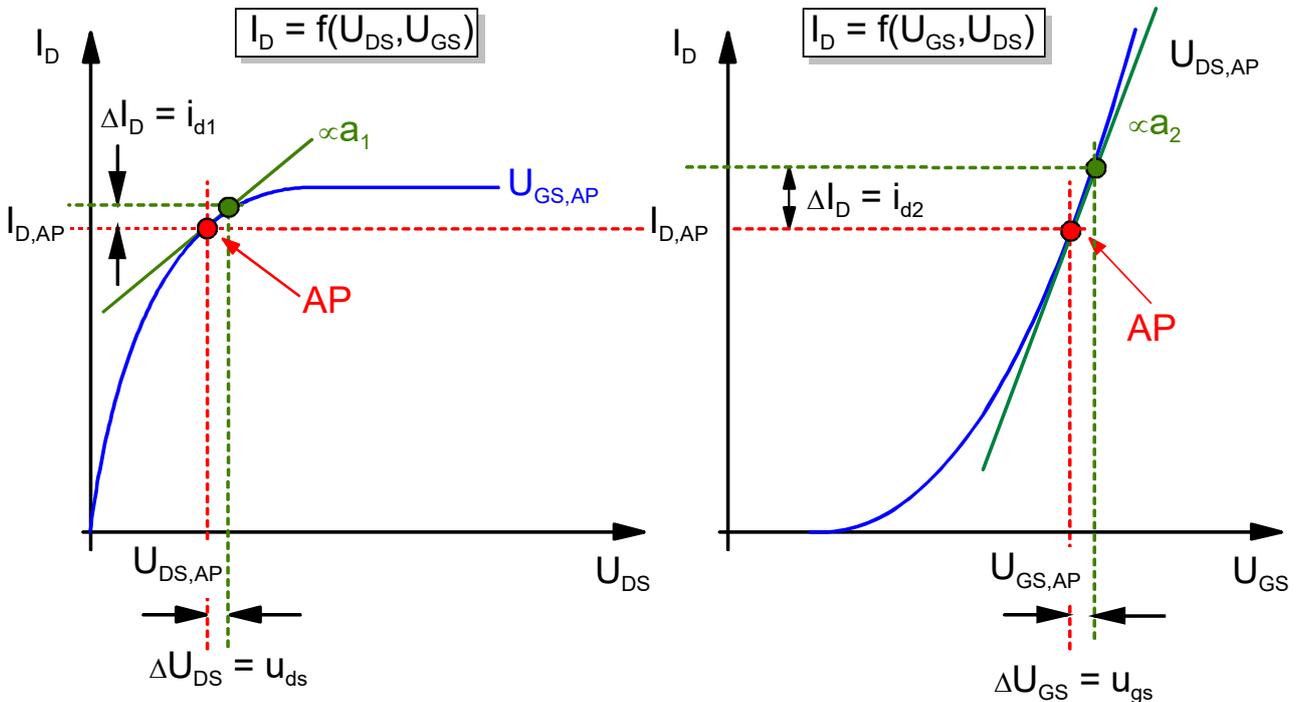


Abb. 2.46 Ausgangskennlinie $I_D = f(U_{DS})$ für $U_{GS} = \text{konst.}$

Abb. 2.47 Übertragungskennlinie $I_D = f(U_{GS})$ für $U_{DS} = \text{konst.}$

a_1 beschreibt die Änderung des Drain-Stroms mit der **Gate-Source**-Spannung, während die Drain-Source-Spannung $U_{DS,AP}$ konstant gehalten wird (Abb. 2.47), d.h. dass keine Signalspannung u_{ds} anliegt:

$$U_{DS} = U_{DS,AP} = \text{konst.} \Rightarrow u_{ds} = 0$$

und damit gilt:

$$a_1 = \left. \frac{\partial I_D}{\partial U_{GS}} \right|_{U_{DS,AP}} \approx \left. \frac{i_{d1}}{u_{gs}} \right|_{u_{ds}=0} = S = g_m \quad (2.90)$$

Die Steigung der Übertragungskennlinie im Arbeitspunkt wird als **Steilheit** S bezeichnet.

Anmerkung: In der Fachliteratur findet man diese Größe auch oft unter der Bezeichnung g_m (engl.: *transconductance*)

a_2 beschreibt die Änderung des Drain-Stroms mit der **Drain-Source**-Spannung, während die Gate-Source-Spannung $U_{GS,AP}$ konstant gehalten wird, d.h. keine Signalspannung u_{gs} anliegt:

$$U_{GS} = U_{GS,AP} = \text{konst.} \Rightarrow u_{gs} = 0$$

und damit gilt:

$$a_2 = \left. \frac{\partial I_D}{\partial U_{DS}} \right|_{U_{GS,AP}} \approx \left. \frac{i_{d2}}{u_{ds}} \right|_{u_{gs}=0} = g_{ds} = \frac{1}{r_{ds}} \quad (2.91)$$

In Abb. 2.46 bedeutet dies die Bewegung auf einer Kennlinie mit $U_{GS} = konst.$

Die Steigung der Ausgangskennlinie im Arbeitspunkt wird als **Ausgangsleitwert** g_{ds} bezeichnet.

Der gesamte Signalstrom ergibt sich zu:

$$i_d = i_{d1} + i_{d2}$$

$$i_d = a_1 \cdot u_{gs} + a_2 \cdot u_{ds}$$

und somit zu:

$$i_d = S \cdot u_{gs} + g_{ds} \cdot u_{ds} \quad (2.92)$$

Damit kann das Kleinsignalersatzschaltbild des FET ausgangsseitig angegeben werden.

Im einfachsten Fall gilt eingangsseitig $I_G = 0$ und wegen der quasistationären Näherung dann auch $i_g = 0$. Damit bildet der Eingangskreis einen Leerlauf.

Damit ergibt sich das Kleinsignalersatzschaltbild für den intrinsischen FET nach Abb. 2.48.

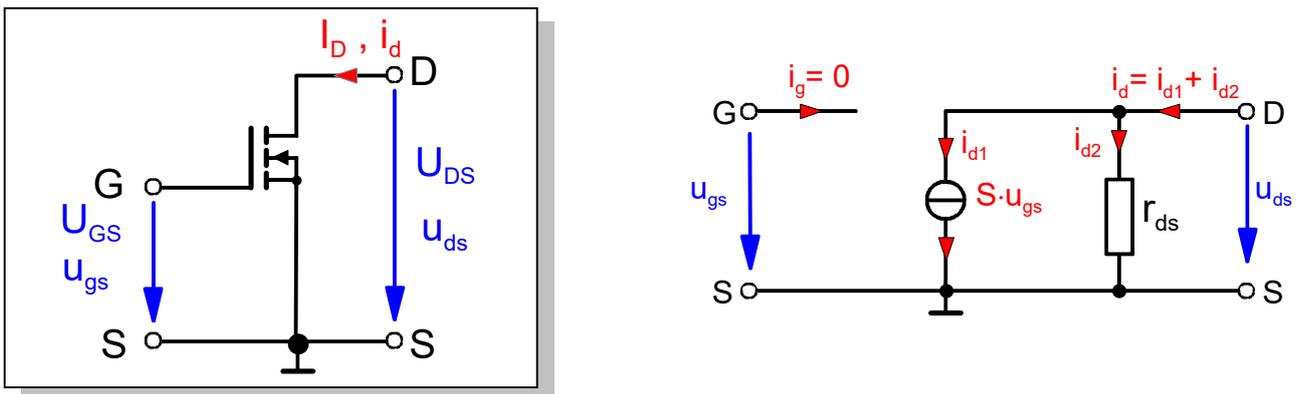


Abb. 2.48 a) (MOS)FET in Source-Schaltung

b) Kleinsignalersatzschaltbild des (MOS)FET in Source-Schaltung (quasistationäre Näherung)

In Sättigung gilt für den idealen FET $r_{ds} \rightarrow \infty$ bzw. $g_{ds} \rightarrow 0$, d.h. $i_{d2} < i_{d1}$, so dass das einfachste Ersatzschaltbild in Abb. 2.49 folgt.

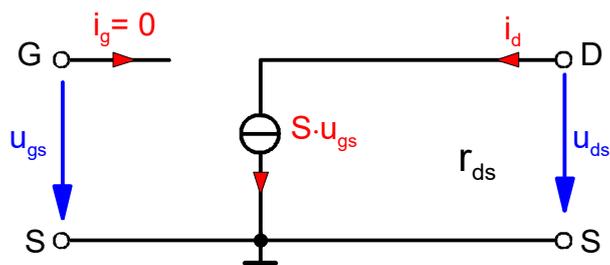


Abb. 2.49 Einfachstes Ersatzschaltbild eines (MOS)FET

Mit Hilfe dieses Kleinsignal-Ersatzschaltbilds lässt sich dann relativ einfach z.B. die Spannungsverstärkung einer Schaltung bei sinusförmigem Eingangssignal als Verhältnis der Wechselspannungsamplitude am Ausgang u_a zur Wechselspannungsamplitude am Eingang u_e berechnen:

$$v_u = \frac{u_a}{u_e} \quad (2.93)$$

Beispiel:

Untersucht werden soll die Verstärkerschaltung in Abb. 2.43.

Wie bereits erwähnt, sind die Koppelkondensatoren in der Regel so groß ausgelegt, dass sie für den betrachteten Frequenzbereich als **wechselspannungsmäßige Kurzschlüsse** anzunehmen sind und deren Einfluss in dem betrachteten Frequenzbereich vernachlässigt werden kann.

Darüber hinaus kann die **konstante DC-Spannungsquelle** für die **wechselspannungsmäßige Betrachtung** (AC) ebenfalls durch einen Kurzschluss gegen Masse ersetzt werden.

Der FET wird im Arbeitspunkt durch das AC-Ersatzschaltbild in Abb. 2.49 ersetzt. Die restlichen Bauelemente werden ergänzt. Damit ergibt sich das Kleinsignalersatzschaltbild der Schaltung in Abb. 2.50.

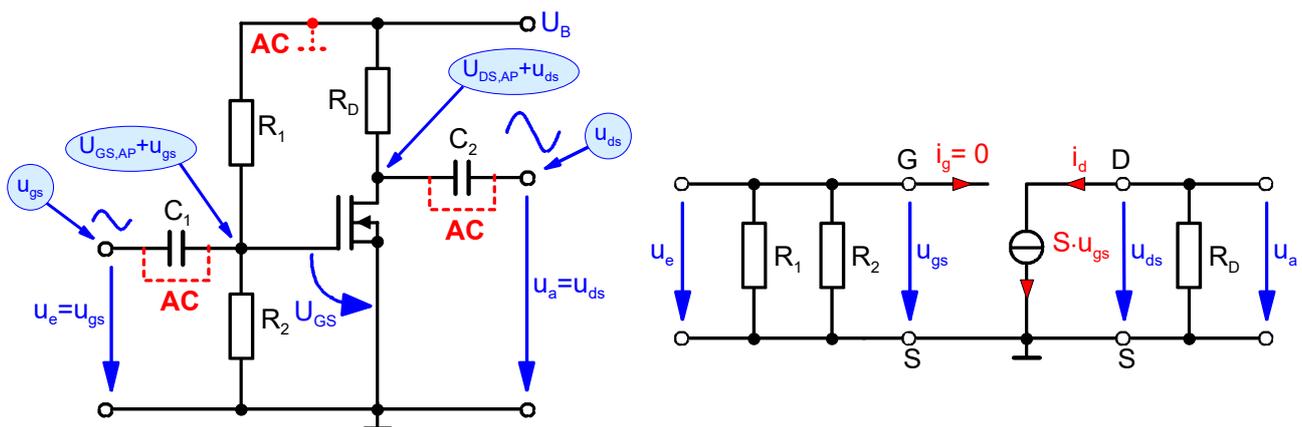


Abb. 2.50 Verstärkerschaltung und zugehöriges Kleinsignalersatzschaltbild

Betrachtet man den Ausgangskreis, ergibt sich die Ausgangsspannung u_a zu:

$$u_a = -S \cdot u_{gs} \cdot R_D \quad (2.94)$$

Mit

$$u_{gs} = u_e$$

folgt

$$u_a = -S \cdot R_D \cdot u_e$$

und damit

$$v_u = \frac{u_a}{u_e} = -S \cdot R_D \quad (2.95)$$

Stichwortverzeichnis

- AC-Verhalten 2-49
- ANDERSON-Modell 2-37**
- aspect ratio* 2-9
- Ausgangsleitwert 2-58
- Bahnwiderstand 2-50
- bipolare Bauelemente 2-2
- CMOS 2-24
- Coulombstreuung 2-39**
- DC-Verhalten 2-49
- Entwurfparameter 2-9
- Ersatzschaltbild 2-49
- extrinsischer Bereich 2-50
- extrinsischer FET 2-50
- Feldeffekttransistor (FET)
 - Ausgangsleitwert 2-58
 - Gleichstromverhalten 2-2
 - Kleinsignalersatzschaltbild in Source-Schaltung (Näherung) 2-58
 - Prinzip 2-2
 - quasistationäres Kleinsignalverhalten 2-56
 - Spannungsverstärkung 2-59
 - Steilheit 2-57
 - transconductance 2-17, 2-57
 - Typen-Übersicht 2-3
- Geschwindigkeitssättigung 2-12
- Gleichspannungsverhalten 2-49
- Heterostruktur-Feldeffekttransistor (HFET) 2-37
- Heteroübergang 2-37
- HFET 2-28, 2-37**
 - Aufbau 2-40
 - Bänderdiagramm 2-38
 - Gleichgewichtsfall 2-43
 - Gleichstromverhalten 2-48
 - Heteroübergang 2-37
 - Ladungstransport im 2DEG 2-39
 - Querschnitt 2-41
 - Schichtkonzentrations-Berechnung 2-42
 - Schottky-Kontakt 2-47
- Hybrid-Form 2-56
- Hybrid-Parameter 2-55
- intrinsischer Bereich 2-50
- intrinsischer FET 2-50
- Inverter
 - Prinzipschaltung 2-53
- JFET 2-4, **2-28**
 - Aufbau 2-29
- Junction-FET 2-28**
- Kanalabschnürung 2-12
- Kanalverkürzung 2-12
- Kennlinienfeld 2-15
 - Ausgangskennlinienfeld 2-15
 - Eingangskennlinienfeld 2-15
 - Übertragungskennlinienfeld 2-15
- Kleinsignalverhalten 2-55
 - allgemein 2-54
 - Kennlinie 2-55
- Komplementäre MOS-Technik 2-24
- Leitwert-Form 2-56
- Leitwert-Parameter 2-55
- MESFET 2-4, **2-28**
 - Anlaufbereich 2-35
 - Aufbau 2-29
 - Kennlinienberechnung 2-32
 - Potenzialverteilung 2-33
 - Sättigung 2-36
 - selbstleitender MESFET 2-32
 - selbstperrender MESFET 2-31
- Metallorganische Gasphasenepitaxie 2-41
- Metall-Semiconductor-FET 2-28**
- Molekularstrahlepitaxie 2-41
- MOS
 - MOS-Feldeffekttransistor (MOSFET) 2-4
- MOSFET 2-4
 - Anreicherungstypen 2-17
 - Arbeitspunkteinstellung 2-22
 - Aufbau und Wirkungsweise 2-4
 - Ausgangskennlinienfeld 2-15
 - Drain-Strom-Berechnung 2-6
 - Eingangskennlinienfeld 2-15
 - Grundsaltungen 2-15
 - Kennlinien-Temperaturabhängigkeit 2-20
 - n-Kanal-Verarmungs-MOSFET 2-23
 - Potenzialverteilung im Kanal 2-7
 - Sättigungsgebiet 2-10
 - Source-Schaltung 2-16
 - Spannungsbereichs-Darstellung 2-14
 - steuerbarer Widerstand 2-9
 - Substratvorspannung-Einfluß 2-20
 - Übertragungskennlinienfeld 2-15
 - Verarmungstypen 2-19
- parasitäre Größen 2-50
- Phononenstreuung 2-39**
- pinch-off 2-12
- Potenzialtopf 2-37
- quasistationäres Verhalten 2-55
- recessed gate 2-52
- Schwellenspannung 2-8

- Source-Schaltung 2-16
- Spannungsverstärkung 2-59
- Sperrschicht-Feldeffekttransistoren 2-28
 - Aufbau und Wirkungsweise 2-29
- Steilheit 2-17, 2-57
- Streu-Parameter 2-55
- Streu-Parameter-Form 2-56
- Subbänder 2-37**
- Taylor-Reihe 2-55
- transconductance 2-17, 2-57
- unipolare Bauelemente 2-2
- versenktes Gate 2-52
- Verstärkerstufe
 - Prinzipschaltung 2-53
- Vierpol
 - Vierpolparameter 2-55
- Wechselspannungsverhalten 2-49, 2-53
 - Funktionsprinzip 2-53
 - Kleinsignalverhalten 2-54
- zweidimensionales Elektronengas 2-37**

Kapitel 3

Dioden

Inhalt

3 Dioden	3-2
3.1 pn-Dioden	3-2
3.1.1 Die Grundgleichungen des pn-Übergangs	3-2
3.1.2 Der pn-Übergang ohne Vorspannung	3-4
3.1.3 Der stromdurchflossene pn-Übergang	3-7
3.1.4 Die reale Diode	3-14
3.1.5 Der Temperatureinfluss	3-16
3.1.5.1 Einfluss der Temperatur auf den Sperrstrom I_o	3-16
3.1.5.2 Einfluss der Temperatur auf U_D	3-17
3.1.5.3 Einfluss der Temperatur auf die gesamte Kennlinie	3-18
3.1.6 Das Elektrische Ersatzschaltbild der pn-Diode	3-20
3.1.6.1 Das Gleichstrom- (DC-)Verhalten	3-20
3.1.6.1.1 Das quasistationäre Kleinsignalverhalten von pn-Dioden	3-21
3.2 Spezielle Dioden	3-23
3.2.1.1 Varaktor-Diode (Kapazitätsdiode)	3-23
3.2.2 Zenerdiode	3-27
3.2.2.1 Tunnel-(Zener-)Effekt	3-28
3.2.2.2 Lawineneffekt	3-31
3.2.3 Tunnel-Diode	3-34
3.2.4 Backward-Diode	3-35
3.2.5 pin-Diode	3-36
3.2.6 Schottky-Dioden	3-38
3.3 Kennzeichnung	3-39
3.3.1 Europäische Kennzeichnung (Pro Electron)	3-39
3.3.2 JEDEC	3-40

3 Dioden

3.1 pn-Dioden

3.1.1 Die Grundgleichungen des pn-Übergangs

In der Vorlesung "Festkörperelektronik" sowie in Kap.0 sind die Grundgleichungen des pn-Übergangs ausführlich hergeleitet worden. Aus diesem Grund werden an dieser Stelle nur die wichtigsten Fakten noch einmal zusammengefasst.

Abb. 3.1 zeigt einen pn-Übergang mit zugehörigem Bändermodell, Ladungsträger- und Dotierstoffkonzentrationen sowie Raumladungs- und Feldverlauf.

Für die Berechnung der Strom-Spannungskennlinie $I = f(U)$ wird vorausgesetzt, dass der pn-Übergang abrupt ist, und dass alle Donator- und Akzeptoratomer ionisiert sind, d.h. es gilt:

	p-Gebiet	n-Gebiet	
Dotierstoffkonzentration	$N_D = 0$ $N_A^- = N_A = konst.$	$N_A = 0$ $N_D^+ = N_D = konst.$	
Bahngebiete:			
Majoritäten	$p_{po} = N_A^- = N_A$	$n_{no} = N_D^+ = N_D$	(3.1)
Minoritäten	$n_{po} = \frac{n_i^2}{p_{po}} = \frac{n_i^2}{N_A}$	$p_{no} = \frac{n_i^2}{n_{no}} = \frac{n_i^2}{N_D}$	(3.2)

Tab. 3.1 Konzentrationen im n- und p-Bereich im stationären Gleichgewichtsfall

In dem Übergangsbereich nimmt die Löcherkonzentration von p_{po} auf p_{no} bzw. die Elektronenkonzentration n_{no} auf n_{po} ab. Die Folge davon sind **Raumladungen** in dem Übergangsbereich. Diese führen in dieser Raumladungszone (RLZ) zu:

1. **Diffusionsströme** infolge eines **Konzentrationsgradienten:**

$$J_{n,Diff} = qD_n grad n = qD_n \frac{dn}{dx} = qD_n n' \tag{3.3}$$

$$J_{p,Diff} = -qD_p grad p = -qD_p \frac{dp}{dx} = -qD_p p' \tag{3.4}$$

mit den Diffusionskonstanten

$$D_n = \frac{kT}{q} \mu_n = U_{th} \mu_n \quad \text{bzw.} \quad D_p = \frac{kT}{q} \mu_p = U_{th} \mu_p \tag{3.5}$$

μ_n bzw. μ_p : Beweglichkeit

J_n bzw. J_p : Stromdichte

$U_{th} = \frac{kT}{q}$: Temperaturspannung

Hinweis: Dabei gilt stets: $q = +1,6 \cdot 10^{-19} \text{ As}$, also $q > 0$!

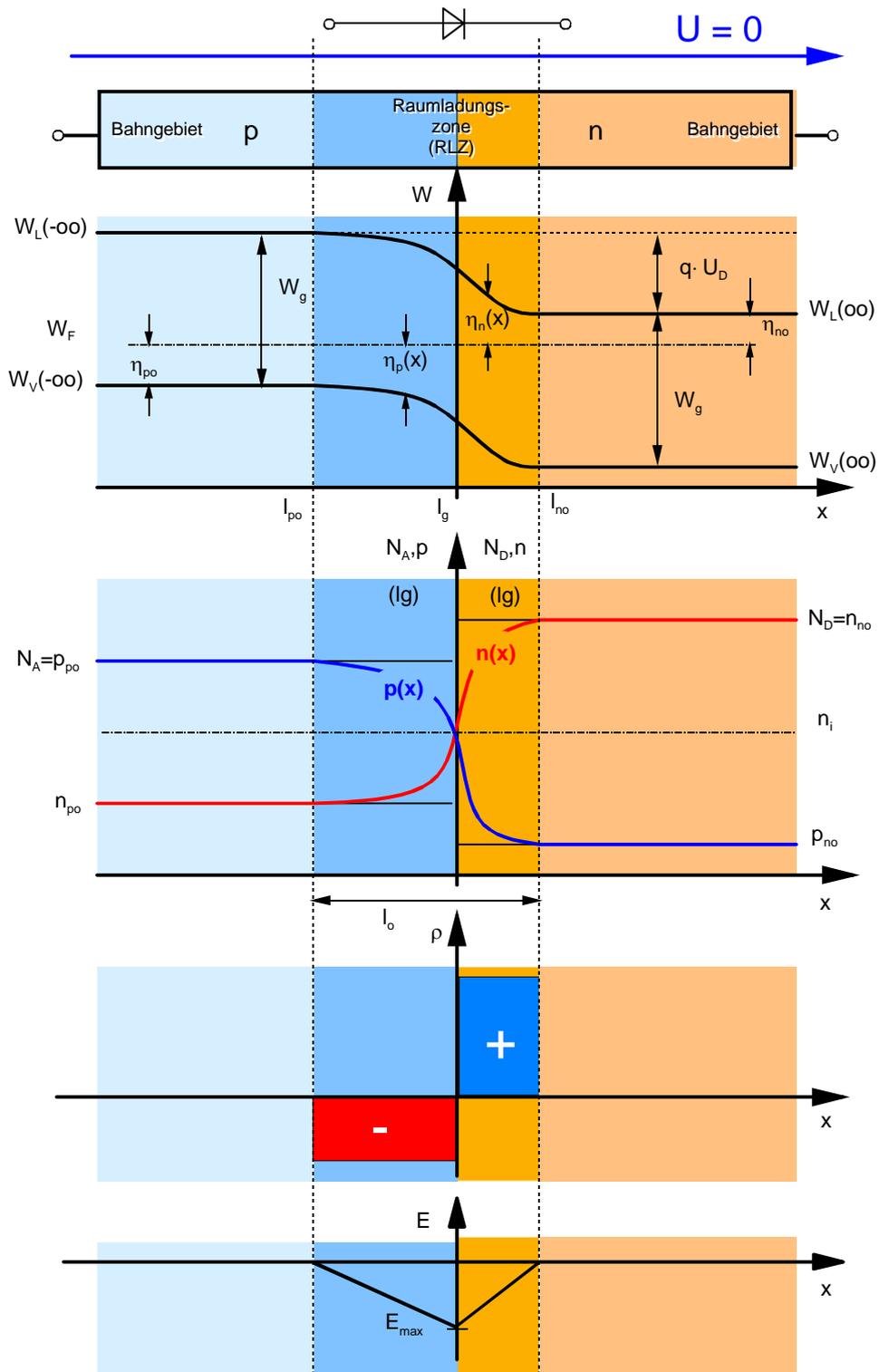


Abb. 3.1 Dotierstoffkonzentration, Ladungsträgerkonzentrationen und Bändermodell eines abrupten pn-Übergangs

und

2. **Feldströme** infolge **elektrischer Felder**

$$J_{n,Feld} = q \cdot \mu_n \cdot n \cdot E = -q \cdot \mu_n \cdot n \cdot \text{grad } \varphi \tag{3.6}$$

$$J_{p,Feld} = -q \cdot \mu_p \cdot p \cdot E = q \cdot \mu_p \cdot p \cdot \text{grad } \varphi \tag{3.7}$$

3.1.2 Der pn-Übergang ohne Vorspannung

Es gilt das Prinzip des detaillierten Gleichgewichts:

$$\sum J_n = 0 \quad \text{d.h. } J_{n,Diff} + J_{n,Feld} = 0 \quad (3.8)$$

und

$$\sum J_p = 0 \quad \text{d.h. } J_{p,Diff} + J_{p,Feld} = 0 \quad (3.9)$$

Die Feldstärke ergibt sich aus dem Bandverlauf am pn-Übergang (Abb. 3.1) mit:

$$qE = \text{grad} W_L = \frac{dW_L}{dx} = W_L'$$

Daraus lassen sich die Konzentrationen berechnen mit

$$n(x) = n_{no} \exp\left[-\frac{W_L(x) - W_L(\infty)}{qU_{th}}\right] \quad (3.10)$$

und:

$$p(x) = p_{po} \exp\left[-\frac{W_V(-\infty) - W_V(x)}{qU_{th}}\right] \quad (3.11)$$

Mit der Definition:

$$W_L(-\infty) - W_L(\infty) = qU_D$$

folgt die **Diffusionsspannung**

$$U_D = U_{th} \ln \frac{N_D N_A}{n_i^2} \quad (3.12)$$

Für Ge, Si und GaAs gilt $U_D = 0,2V \dots 1,3V$.

Hinweis: Häufig wird die Diffusionsspannung in der Literatur negativ angegeben. In diesem Fall ist das Vorzeichen in den zugehörigen Gleichungen entsprechend berücksichtigt.

Mit Hilfe der **Poisson-Gleichung** lassen sich der Bandverlauf und die Raumladung ρ berechnen:

$$\rho = q \left[N_D^+ + p - N_A^- - n \right] \quad (3.13)$$

Da die Konzentration der beweglichen Ladungen $n(x)$ und $p(x)$ schnell abnehmen (vgl. Vorlesung *Festkörperelektronik*), werden nur die Raumladung aufgrund der ortsfesten Donatoren und Akzeptoren berücksichtigt.

In Tab. 3.1 sind die Berechnung des Bandverlaufs und der Konzentrationen zusammengefasst.

	p-Gebiet	n-Gebiet
Bereich der RLZ	$l_{po} \leq x \leq l_g$	$l_g \leq x \leq l_{no}$
Raumladung (Näherung statt Gl. (3.13))	$\rho = -qN_A$	$\rho = qN_D$
Poisson-Gleichung:	$\frac{d^2\varphi}{dx^2} = -\frac{\rho}{\varepsilon_o\varepsilon_r} = \frac{qN_A^-}{\varepsilon_o\varepsilon_r} = -\frac{1}{q} \frac{d^2W_L}{dx^2}$	$\frac{d^2\varphi}{dx^2} = -\frac{\rho}{\varepsilon_o\varepsilon_r} = -\frac{qN_D^+}{\varepsilon_o\varepsilon_r} = -\frac{1}{q} \frac{d^2W_L}{dx^2}$ (3.14)
1. Integration	$\frac{dW_{L,p}}{dx} = -\frac{q^2 N_A}{\varepsilon_o\varepsilon_r} x + const.$	$\frac{dW_{L,n}}{dx} = \frac{q^2 N_D}{\varepsilon_o\varepsilon_r} x + const.$
Randbedingungen	$\left. \frac{dW_{L,p}}{dx} \right _{x=l_{po}} = 0$ $\Rightarrow \frac{dW_{L,p}}{dx} = -\frac{q^2 N_A}{\varepsilon_o\varepsilon_r} (x - l_{po})$	$\left. \frac{dW_{L,n}}{dx} \right _{x=l_{no}} = 0$ (3.15) $\Rightarrow \frac{dW_{L,n}}{dx} = -\frac{q^2 N_D}{\varepsilon_o\varepsilon_r} (l_{no} - x)$
2. Integration	$W_{L,p}(x) = -\frac{q^2 N_A}{\varepsilon_o\varepsilon_r} \int (x - l_{po}) dx$ $= W_L(l_{po}) - \frac{q^2 N_A}{2\varepsilon_o\varepsilon_r} (x - l_{po})^2$	$W_{L,n}(x) = -\frac{q^2 N_D}{\varepsilon_o\varepsilon_r} \int (l_{no} - x) dx$ $= W_L(l_{no}) + \frac{q^2 N_D}{2\varepsilon_o\varepsilon_r} (l_{no} - x)^2$ (3.16)

Tab. 3.2 Berechnung des Bandverlaufs

Die Größen l_{no} und l_{po} werden durch Anpassung der Lösungen für die p- und die n-Seite an der Stelle $x = l_g$ bestimmt, unter der Annahme, dass

- keine Grenzflächenladungen ($Div D = 0$)
- keine Dipolschichten

auftreten. An der Grenze muss **Stetigkeit**:

$$W_{L,p}(x = l_g) = W_{L,n}(x = l_g) \quad (3.17)$$

sowie **stetige Differenzierbarkeit** mit

$$\left. \frac{dW_{L,p}}{dx} \right|_{x=l_g} = \left. \frac{dW_{L,n}}{dx} \right|_{x=l_g} \quad (3.18)$$

und **Ladungsgleichgewicht**

$$N_A \cdot (l_g - l_{po}) = N_D \cdot (l_{no} - l_g) \quad (3.19)$$

bestehen. Damit lassen sich die Grenzen der Raumladungszone bestimmen mit

$$l_{no} - l_g = \sqrt{\frac{2\varepsilon_o\varepsilon_r N_A}{qN_D(N_D + N_A)}} \cdot \sqrt{U_D} \quad (3.20)$$

$$l_g - l_{po} = \sqrt{\frac{2\varepsilon_o\varepsilon_r N_D}{qN_A(N_D + N_A)}} \cdot \sqrt{U_D} \quad (3.21)$$

Daraus ergibt sich die Ausdehnung der Raumladungszone:

$$l_o = l_{no} - l_{po} \quad (3.22)$$

Damit sind

- die Ausdehnung der Raumladungszone $l_g - l_{po}$ im p-Gebiet
- die Ausdehnung der Raumladungszone $l_{no} - l_g$ im n-Gebiet
- die gesamte Ausdehnung der Raumladungszone $l_{no} - l_{po}$ und
- der Verlauf des Leitungsbandes $W_L(x)$

berechenbar, wenn N_D , N_A und die Materialgrößen ε_r bekannt sind. Der Verlauf des Valenzbandes lässt sich entsprechend über

$$W_V(x) = W_L(x) - W_g$$

berechnen.

Unter Verwendung der **Debye-Länge** L_D mit

$$L_D = \sqrt{\frac{\varepsilon_o\varepsilon_r U_{th}}{qN_D}} \quad (3.23)$$

ergeben sich mit Gl. (3.10) und (3.11) die Konzentrationen von Elektronen und Löchern als freie Ladungsträger in der Raumladungszone zu:

	RLZ im p-Gebiet	RLZ im n-Gebiet
	$l_{po} \leq x \leq l_g$	$l_g \leq x \leq l_{no}$
Löcher-konzentration	$p(x) = p_{po} \exp\left[-\frac{1}{2}\left(\frac{x-l_{po}}{L_{Dp}}\right)^2\right]$ (3.24)	$p(x) = \frac{n_i^2}{N_D} \cdot \exp\left[\frac{1}{2}\left(\frac{l_{no}-x}{L_{Dn}}\right)^2\right]$ (3.25)
Elektronen-konzentration	$n(x) = \frac{n_i^2}{N_A} \cdot \exp\left[\frac{1}{2}\left(\frac{x-l_{po}}{L_{Dp}}\right)^2\right]$ (3.26)	$n(x) = n_{no} \exp\left[-\frac{1}{2}\left(\frac{l_{no}-x}{L_{Dn}}\right)^2\right]$ (3.27)

Tab. 3.3 Konzentrationen der Minoritäten in der Raumladungszone

3.1.3 Der stromdurchflossene pn-Übergang

Durch Anlegen einer Spannung wird das Gleichgewicht zwischen Feld- und Diffusionsdrift der Ladungsträger gestört. Wird eine positive Spannung (Abb. 3.2) angelegt, werden die Löcher in dem p-Gebiet und die Elektronen im n-Gebiet in Richtung des pn-Übergangs gedrückt, gleichzeitig die Minoritätsladungsträger (Elektronen im p-Gebiet und Löcher im n-Gebiet) vom pn-Übergang weggezogen. Es ergeben sich neue Dichteverteilungen $n(x)$ und $p(x)$, die aber jetzt nicht mehr zu einem statischen Gleichgewicht führen, sondern nur durch einen dauernden Stromfluss aufrecht erhalten werden können, dessen Richtung der angelegten Spannung U entspricht.

Wird dagegen eine negative Spannung (Abb. 3.3) angelegt, werden die Löcher in dem p-Gebiet und die Elektronen im n-Gebiet vom pn-Übergang weggezogen. Die Raumladungszone dehnt sich aus.

Zur Berechnung ist in den Gleichungen zur Berechnung der Raumladungszone (Gl. (3.20),(3.21)) jetzt

$$U_D \text{ durch } U_D - U$$

zu ersetzen. Dadurch bedingt verändern sich die Grenzen der Raumladungszonen und es ist dann

$$l_n, l_p \text{ statt } l_{no}, l_{po}$$

zu verwenden.

An den Grenzen der Raumladungszone gilt nun nicht mehr der Gleichgewichtsfall aus Gl. (3.2), sondern die Minoritätenkonzentrationen an den Grenzen der Raumladungszone sind nunmehr von der angelegten Spannung abhängig und es gilt:

$$n_p(l_p) = \frac{n_i^2}{N_A} \exp\left(\frac{U}{U_{th}}\right) = n_{po} \cdot e^{U/U_{th}} \quad (3.28)$$

bzw.:

$$p_n(l_n) = \frac{n_i^2}{N_D} \exp\left(\frac{U}{U_{th}}\right) = p_{no} \cdot e^{U/U_{th}} \quad (3.29)$$

Dieses Ergebnis sagt folgendes aus:

1. Die Konzentration der Minoritätsladungsträger n_p und p_n an den Sperrschichträndern l_p und l_n wird bei Anlegen einer Spannung U um den Faktor $\exp(U/U_{th})$ gegenüber den Gleichgewichtswerten n_{po} und p_{no} verändert.
2. Die Minoritätsladungsträgerkonzentrationen n_p und p_n klingen innerhalb der Bahngebiete auf die Gleichgewichtswerte n_{po} und p_{no} ab.

Diese abklingenden Konzentrationen heißen **Diffusionsschwänze**

3. Innerhalb der Diffusionsschwänze herrscht Ladungsneutralität.
4. Die Stromergiebigkeit eines pn-Übergangs (bzw. der Widerstand eines pn-Übergangs) hängt **nur** von dem Strom ab, der in den Konzentrationsgradienten der Diffusionsschwänze transportiert werden kann.

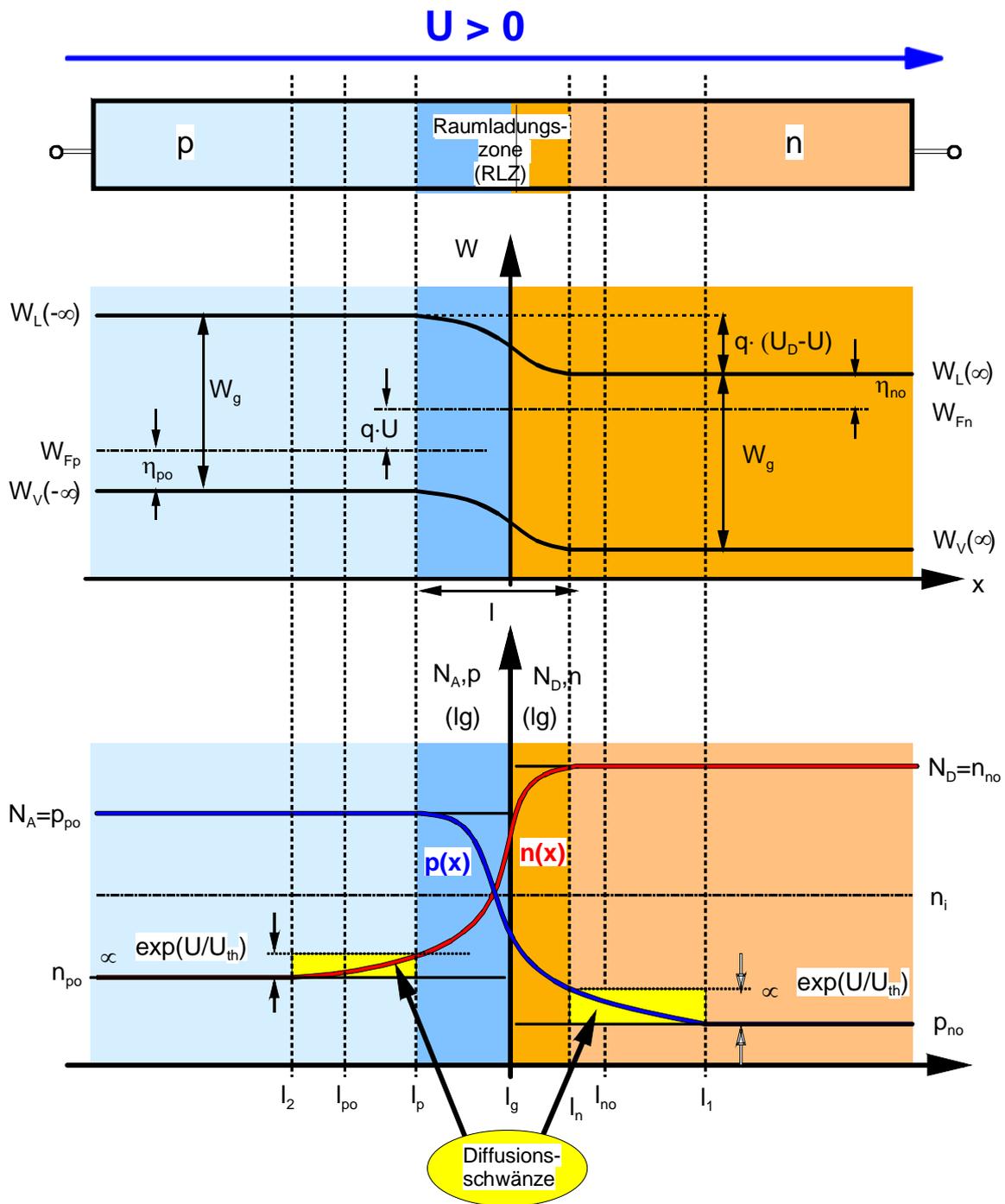


Abb. 3.2 Ladungsträgerverteilung und Bändermodell eines in Flussrichtung gepolten, stromdurchflossenen pn-Übergangs ($U > 0$)

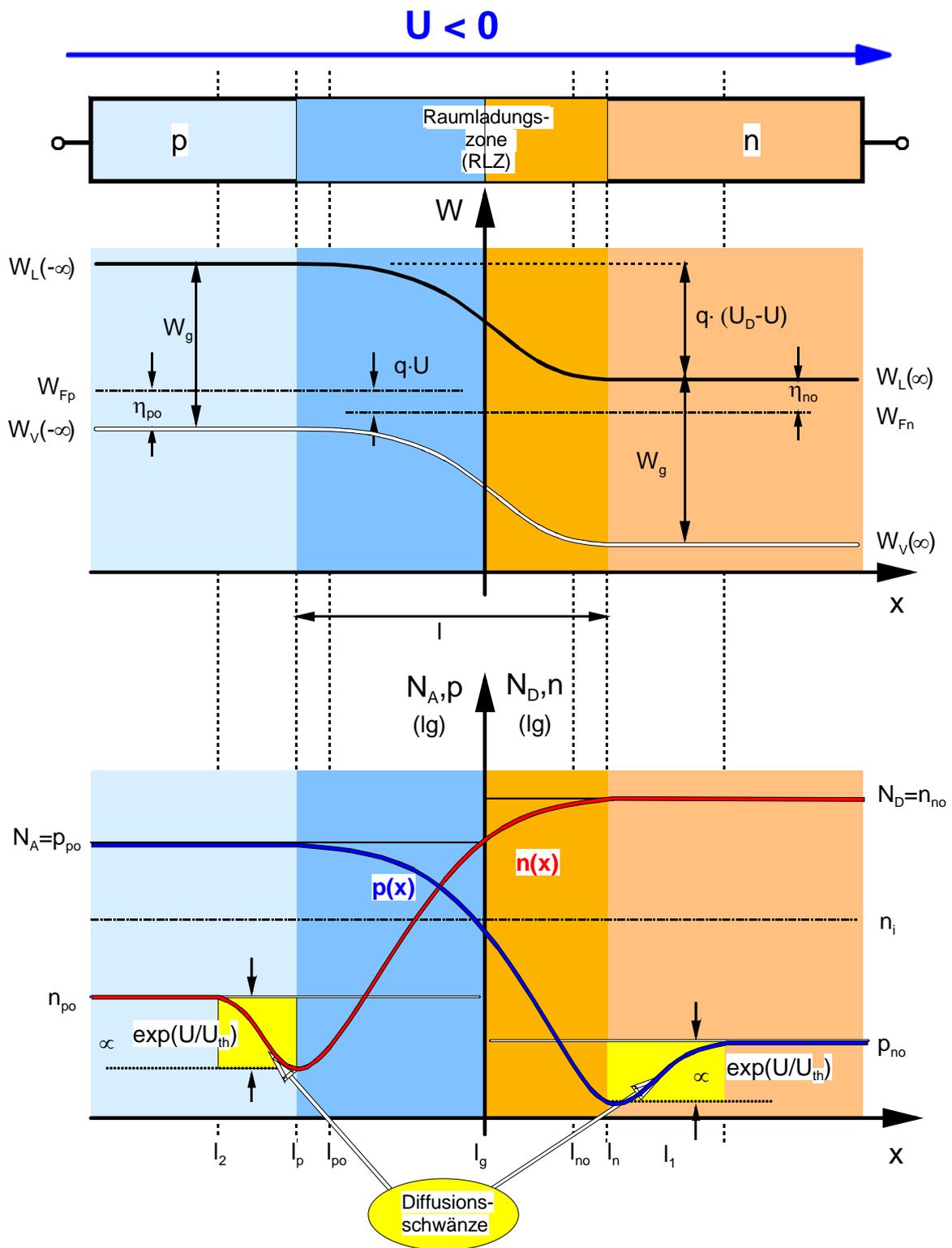


Abb. 3.3 Ladungsträgerverteilung und Bändermodell eines in Sperrrichtung gepolten pn-Übergangs ($U < 0$)

Berechnung der Stromdichteanteile der Diffusionsschwänze:

Der Stromdichteanteil der Löcher im **n-Bereich** ergibt sich entsprechend Gl. (3.4) zu:

$$J_{pn}(x) = -q\mu_p U_{th} p_n'(x) \quad \text{für } x \geq l_n \quad (3.30)$$

$J_{pn}(x)$ und $p_n'(x)$ lassen sich unter Verwendung der **Kontinuitätsgleichung**:

$$\frac{\partial p(x)}{\partial t} = \dot{p}(x) = -\frac{1}{q} \frac{\partial J_{pn}(x)}{\partial x} - r_{th} + G \quad \text{für } x \geq l_n \quad (3.31)$$

r_{th} : Rekombinationsrate

G : Generationsrate

berechnen. Für den Gleichstromfall gilt:

$$\frac{\partial p(x)}{\partial t} = 0$$

Außer der thermischen Generation, die in der **Rekombinationsrate** r_{th} enthalten ist, soll keine zusätzliche Generation vorhanden sein, das heißt:

$$G = 0$$

Für die Rekombinationsrate gilt entsprechend der Vorlesung *Festkörperelektronik*:

$$r_{th} = \frac{p_n(x) - p_{no}}{\tau_p} \quad (3.32)$$

Dabei bezeichnet τ_p die **Lebensdauer** der Löcher im n-dotierten Bereich. Damit ergibt sich für die Löcher aus Gl. (3.30) eingesetzt in (3.31):

$$p_n''(x) = \frac{1}{L_p^2} (p_n(x) - p_{no}) \quad \text{für } x \geq l_n \quad (3.33)$$

und analog dazu für die Elektronen:

$$n_p''(x) = \frac{1}{L_n^2} (n_p(x) - n_{po}) \quad \text{für } x \leq l_p \quad (3.34)$$

wobei L_p die **Diffusionslänge** der Löcher im n-Gebiet und L_n die **Diffusionslänge** der Elektronen im p-Gebiet bezeichnen mit:

$$\tau_p \mu_p U_{th} = \tau_p D_p = L_p^2 \quad (3.35)$$

bzw.

$$\tau_n \mu_n U_{th} = \tau_n D_n = L_n^2 \quad (3.36)$$

Die allgemeine Lösung der Differentialgleichung (3.33) lautet:

$$p_n(x) - p_{no} = A e^{x/L_p} + B e^{-x/L_p} \quad (3.37)$$

a) **Allgemeine Lösung:**

Die Randbedingungen zur Lösung der Differentialgleichung sind gegeben durch:

$$1. \quad p_n(l_n) = p_{no} \cdot e^{U/U_{th}} \quad \text{bzw.} \quad n_p(l_p) = n_{po} \cdot e^{U/U_{th}} \quad (3.38)$$

und

$$2. \quad p_n(l_1) = p_{no} \quad \text{bzw.} \quad n_p(l_2) = n_{po} \quad (3.39)$$

Einsetzen der Randbedingungen und Berechnung der Konstanten A,B liefert:

$$p_n(x) - p_{no} = p_{no} \left(e^{U/U_{th}} - 1 \right) \frac{\sinh\left(\frac{l_1 - x}{L_p}\right)}{\sinh\left(\frac{l_1 - l_n}{L_p}\right)} \quad \text{für } x \geq l_n \quad (3.40)$$

und mit Gl. (3.30):

$$J_{pn} = \frac{q\mu_p U_{th}}{L_p} p_{no} \left(e^{U/U_{th}} - 1 \right) \frac{\cosh\left(\frac{l_1 - x}{L_p}\right)}{\sinh\left(\frac{l_1 - l_n}{L_p}\right)} \quad \text{für } x \geq l_n \quad (3.41)$$

Analog gilt für den Stromdichteanteil der Elektronen im p-Gebiet:

$$J_{np} = \frac{q\mu_n U_{th}}{L_n} n_{po} \left(e^{U/U_{th}} - 1 \right) \frac{\cosh\left(\frac{x - l_2}{L_n}\right)}{\sinh\left(\frac{l_p - l_2}{L_n}\right)} \quad \text{für } x \leq l_p \quad (3.42)$$

Die Gesamtstromdichte J wird unter folgenden Annahmen berechnet:

1. In den Bahngebieten tritt kein Spannungsabfall über dem Bereich der Diffusionsschwänze auf, d.h. in den Bahngebieten gilt: $J_{Feld} = 0$. Damit wird der Strom nur durch den Diffusionsstromanteil bestimmt:

$$J = J_{Diff}$$

mit $J_{pn,Diff}$ im n-Gebiet und $J_{np,Diff}$ im p-Gebiet.

2. In der Raumladungszone tritt nur eine schwache Rekombination auf (*Shockley'sche Näherung*), d.h.:

$$J_{pn}(x) \approx \text{konst.} \quad \text{bzw.}$$

$$J_{np}(x) \approx \text{konst.}$$

in der Raumladungszone.

Dies bedeutet auch, dass die Diffusionslängen L_n, L_p der Minoritäten groß gegen die Ausdehnung der Raumladungszone sein müssen.

$$L_n, L_p \gg l_n - l_p \tag{3.43}$$

Die Diffusionsstromanteile $J_{pn}(x)$ und $J_{np}(x)$ müssen an einer geeigneten Stelle addiert werden, z.B. an der Stelle $x = l_g$.

Aufgrund der oben genannten Näherungen gilt dort:

$$J_{pn}(x = l_g) = J_{pn}(x = l_n) \tag{3.44}$$

und

$$J_{np}(x = l_g) = J_{np}(x = l_p) \tag{3.45}$$

Diese gemachten Annahmen sind in Abb. 3.4 zusammenfassend dargestellt.

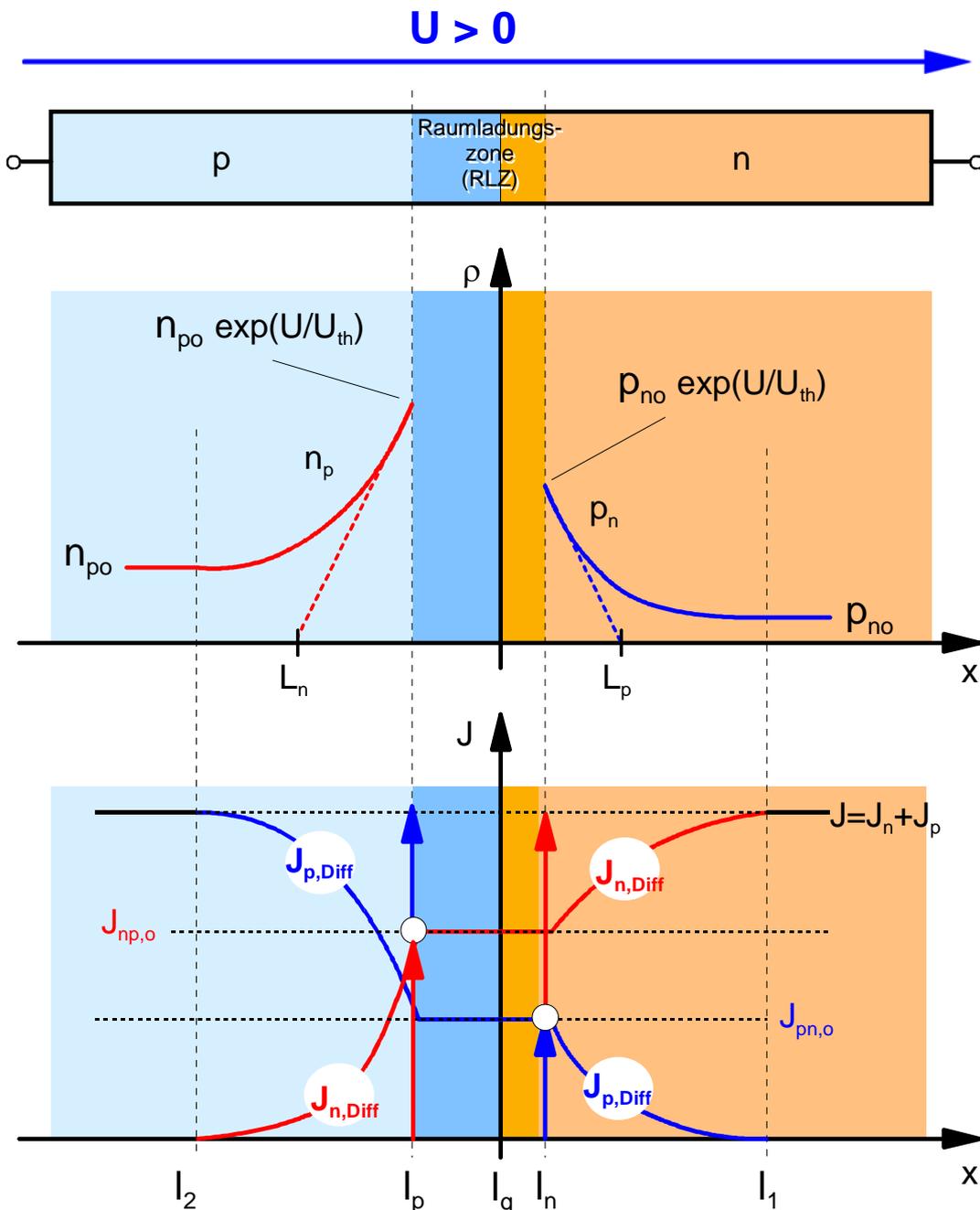


Abb. 3.4 Näherungen zur Berechnung der gesamten Stromdichte

Einsetzen von Gl. (3.41) und Gl. (3.42) in Gl. (3.45) ergibt:

$$J = J_o \left(e^{U/U_{th}} - 1 \right) \quad (3.46)$$

mit

$$J_o = qU_{th} \left(\frac{\mu_p}{L_p} p_{no} \coth\left(\frac{l_1 - l_n}{L_p}\right) + \frac{\mu_n}{L_n} n_{po} \coth\left(\frac{l_p - l_2}{L_n}\right) \right) \quad (3.47)$$

Der Ausdruck (3.47) kann für Spezialfälle weiter vereinfacht werden:

b) Langbasisdiode

Die Ausdehnung der Bahngebiete ist groß gegenüber der Diffusionslänge der Minoritäten:

$$l_1 - l_n \gg L_p \quad (3.48)$$

$$l_p - l_2 \gg L_n \quad (3.49)$$

Wegen

$$\coth(a) = 1 \quad \text{für } a \gg 1$$

folgt:

$$J_o^L = qU_{th} \left(\frac{\mu_p}{L_p} p_{no} + \frac{\mu_n}{L_n} n_{po} \right) \quad (3.50)$$

bzw. mit Gl. (3.5)

$$J_o^L = q \cdot \left(\frac{D_p}{L_p} p_{no} + \frac{D_n}{L_n} n_{po} \right) \quad (3.51)$$

c) Kurzbasisdiode

$$l_1 - l_n \ll L_p \quad (3.52)$$

$$l_p - l_2 \ll L_n \quad (3.53)$$

Wegen

$$\coth(a) = \frac{1}{a} \quad \text{für } a \ll 1$$

gilt:

$$J_o^K = qU_{th} \left(\frac{\mu_p}{l_1 - l_n} p_{no} + \frac{\mu_n}{l_p - l_2} n_{po} \right) = q \cdot \left(\frac{D_p}{l_1 - l_n} p_{no} + \frac{D_n}{l_p - l_2} n_{po} \right) \quad (3.54)$$

Das heißt, in die Näherungen (3.51) bzw. (3.54) sind jeweils die kürzeren der beiden Längen $l_1 - l_n$, L_p bzw. $l_p - l_2$, L_n einzusetzen.

1. Durchlassbereich:

Für $U \gg U_{th}$ folgt näherungsweise

$$I = I_o e^{U/U_{th}} \tag{3.55}$$

Anmerkung: Bereits für $U/U_{th} = 4$, d.h. für $U \approx 100mV$ bei $T = 300K$ beträgt der Fehler der Gl. (3.54) gegenüber der Gl. (3.47) weniger als 2%.

2. Sperrbereich:

Für $U < 0$ und $|U| \gg |U_{th}|$ gilt:

$$I = -I_o \tag{3.56}$$

3.1.4 Die reale Diode

Bislang wurde die Stromcharakteristik der idealen Diode berechnet. Dabei wurden parasitäre Widerstände aufgrund der Bahngebiete sowie Kontaktwiderstände vernachlässigt. Diese können je jedoch in der Praxis einen erheblichen Einfluss haben. Der zusätzliche Spannungsabfall über dem parasitären Widerstand verschiebt die Kennlinie und sorgt für eine Scherung der Kennlinie wie in Abb. 3.5 dargestellt. Letztendlich verhindern diese parasitären Widerstände, dass der idealer Weise angenommene Flachbandfall bei $U = U_D$ tatsächlich eintritt und der Strom somit unendlich stark ansteigen würde.

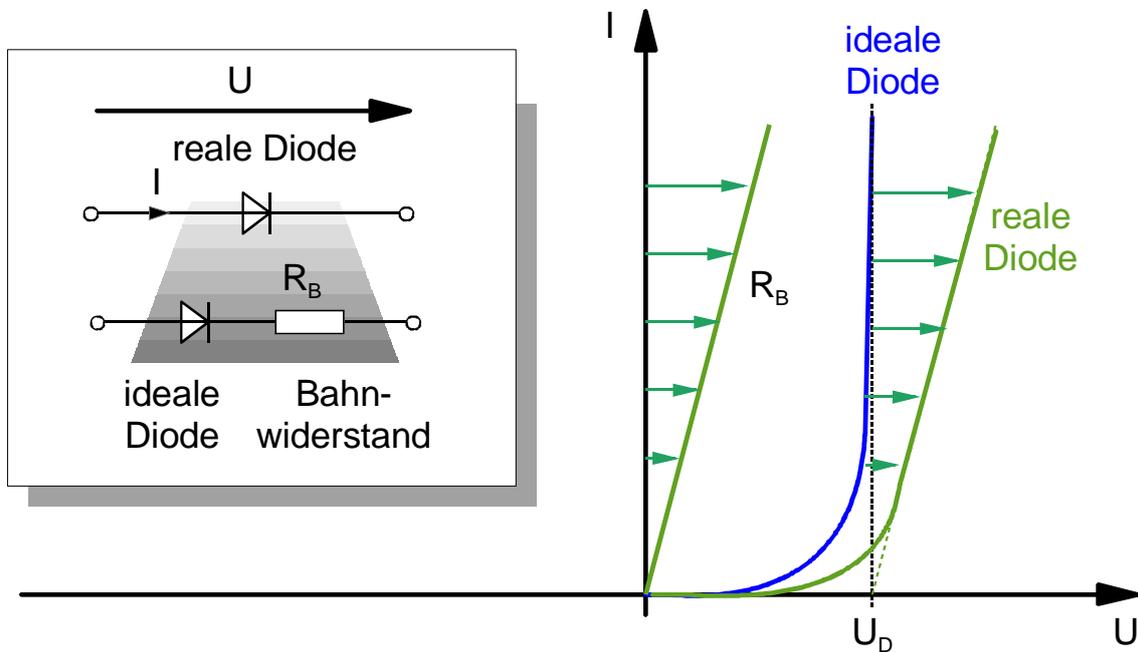


Abb. 3.5 Scherung der Kennlinie aufgrund parasitärer Bahn- und Kontaktwiderstände

Darüber hinaus ist die Kennlinie stark materialabhängig. Nach Gl. (3.12) gilt für die Diffusionsspannung

$$U_D = U_{th} \ln \frac{N_D N_A}{n_i^2}$$

Zahlenbeispiel:

Material	n_i / cm^{-3}	N_D / cm^{-3}	N_A / cm^{-3}	U_D / V
Si	$1,45 \cdot 10^{10}$	$1 \cdot 10^{15}$	$1 \cdot 10^{15}$	0,56
		$1 \cdot 10^{16}$	$1 \cdot 10^{16}$	0,67
		$1 \cdot 10^{15}$	$1 \cdot 10^{18}$	0,73
		$1 \cdot 10^{18}$	$1 \cdot 10^{18}$	0,90
Ge	$2,4 \cdot 10^{13}$	$1 \cdot 10^{15}$	$1 \cdot 10^{15}$	0,19
		$1 \cdot 10^{16}$	$1 \cdot 10^{16}$	0,30
		$1 \cdot 10^{15}$	$1 \cdot 10^{18}$	0,36
		$1 \cdot 10^{18}$	$1 \cdot 10^{18}$	0,53
GaAs	$1,8 \cdot 10^6$	$1 \cdot 10^{15}$	$1 \cdot 10^{15}$	1,00
		$1 \cdot 10^{16}$	$1 \cdot 10^{16}$	1,12
		$1 \cdot 10^{15}$	$1 \cdot 10^{18}$	1,18
		$1 \cdot 10^{18}$	$1 \cdot 10^{18}$	1,35

Tab. 3.4 Diffusionsspannungen für verschiedene Materialien und Dotierungen

Man erkennt die relativ schwache Abhängigkeit der Diffusionsspannung von den gewählten Dotierungen. Den größeren Einfluss hat das jeweilige Halbleitermaterial.

Abb. 3.6. zeigt den prinzipiellen Unterschied zwischen pn-Dioden auf Silizium und Germanium.

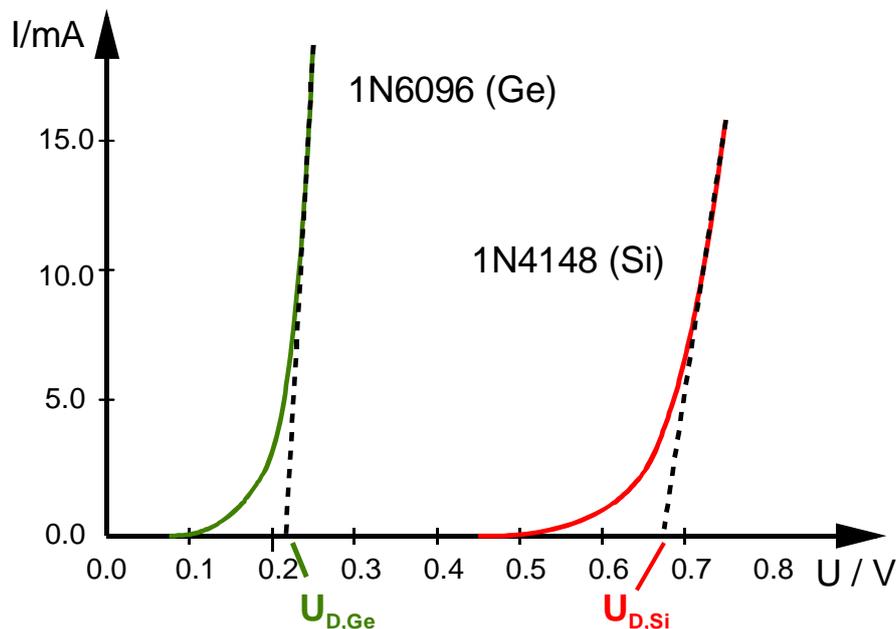


Abb. 3.6 Unterschied zweier kommerzieller Si-(1N4148) und Ge-(1N6096) Dioden

3.1.5 Der Temperatureinfluss

Wenn die Temperatur zunimmt, können weitere Ladungen durch Band-Band-Übergänge erzeugt werden. Eine Temperaturänderung hat sowohl Einfluss auf den Sperrstrom I_o als auch die Diffusionsspannung U_D

3.1.5.1 Einfluss der Temperatur auf den Sperrstrom I_o

Nach Gl. (3.51) gilt für eine Langbasisdiode:

$$J_o^L = \left(\frac{D_p}{L_p} p_{no} + \frac{D_n}{L_n} n_{po} \right)$$

Mit

$$p_{no} = \frac{n_i^2}{N_D} \quad \text{bzw.} \quad n_{po} = \frac{n_i^2}{N_A}$$

folgt:

$$J_o^L = n_i^2 \left(\frac{D_p}{N_D L_p} + \frac{D_n}{N_A L_n} \right) \propto n_i^2 \quad (3.57)$$

Für die Eigenleitungskonzentration gilt:

$$n_i(T) = A \cdot T^{3/2} e^{-W_g/2KT} \quad (3.58)$$

und damit für das Verhältnis der Konzentrationen bei zwei verschiedenen Temperaturen:

$$\frac{n_i(T_1)}{n_i(T_0)} = \frac{T_1^{3/2}}{T_0^{3/2}} e^{-\frac{W_g}{2KT_1} + \frac{W_g}{2KT_0}}$$

bzw.

$$n_i(T_1) = n_i(T_0) \cdot \left(\frac{T_1}{T_0} \right)^{3/2} e^{-\frac{W_g}{2K} \left(\frac{1}{T_1} - \frac{1}{T_0} \right)} \quad (3.59)$$

Bei relativ geringen Temperaturänderungen gilt:

$$\frac{T_1}{T_0} = \frac{T_0 + \Delta T}{T_0} \approx 1$$

und damit:

$$n_i(T_1) \approx n_i(T_0) e^{\frac{W_g}{2K} \left(\frac{1}{T_0} - \frac{1}{T_1} \right)} = n_i(T_0) e^{\frac{W_g}{2K} \cdot \frac{\Delta T}{T_0 T_1}} \quad (3.60)$$

Für die Stromdichte mit

$$J_o \propto n_i^2$$

folgt:

$$J_o(T_1) = J_o(T_0) \cdot e^{\frac{W_g}{KT_0 T_1} \cdot \Delta T} = J_o(T_0) \cdot e^{\lambda_T \cdot \Delta T} \quad (3.61)$$

mit

$$\lambda_T = \frac{Wg}{KT_0T_1} \quad (3.62)$$

Für kleine Temperaturänderungen mit

$$T_0 \cdot T_1 \approx T_0^2$$

lässt sich der Ansatz vereinfachen zu:

$$\lambda_T = \frac{Wg}{K \cdot T_0^2} \quad (3.63)$$

und damit:

$$\frac{J_o(T_1)}{J_o(T_0)} = e^{\lambda_T \cdot \Delta T} \quad (3.64)$$

bzw.

$$I_o(T_1) = I_o(T_0) \cdot e^{\lambda_T \cdot \Delta T} \quad (3.65)$$

Zahlenbeispiel für $T_0 = 300K$ und $\Delta T = 10K$

Material	Wg / eV	λ_T / K^{-1}	$I_o(T_1) / I_o(T_0)$
Si	1,12	0,14	4,1
Ge	0,66	0,085	2,3
GaAs	1,42	0,18	6,04

Tab. 3.5 Temperaturkoeffizient für verschiedene Materialien

Aus der Gegenüberstellung folgt:

$$\lambda_T(\text{GaAs}) > \lambda_T(\text{Si}) > \lambda_T(\text{Ge})$$

Mit zunehmender Temperatur nehmen also die Minoritätenkonzentration und damit der Sperrstrom zu. Da die Minoritätenkonzentrationen exponentiell mit der Temperatur ansteigen, nimmt auch der Sperrstrom **exponentiell** mit der Temperatur zu.

3.1.5.2 Einfluss der Temperatur auf U_D

Nach Gl. (3.12) gilt für die Diffusionsspannung:

$$U_D = U_{th} \cdot \ln \left(\frac{N_A N_D}{n_i^2} \right) \quad (3.66)$$

Mit Gl. (3.61) und (3.64) folgt damit

$$U_D(T) = U_{th} \cdot \left[\ln \frac{N_A N_D}{n_i^2(T_0) e^{\lambda_T(T-T_0)}} \right] = U_{th} \cdot \left[\ln \frac{N_A N_D}{n_i^2(T_0)} - \ln \left(e^{\lambda_T(T-T_0)} \right) \right] \quad (3.67)$$

$$U_D(T) = U_{th} \cdot \left[\ln \frac{N_A N_D}{n_i^2(T_0)} \right] - U_{th} \cdot \lambda_T (T - T_0) \quad (3.68)$$

Unter der Annahme relativ kleiner Temperaturänderungen, so dass näherungsweise gilt:

$$U_{th} \approx \text{konst.}$$

folgt

$$\frac{dU_D}{dT} \approx -U_{th} \cdot \lambda_T = -d_T \quad \text{mit} \quad d_T : \text{Temperaturdurchgriff} \quad (3.69)$$

3.1.5.3 Einfluss der Temperatur auf die gesamte Kennlinie

Für die Strom-Spannungskennlinie gilt

$$I \approx I_0 e^{U/U_{th}} \quad , (U \gg U_{th})$$

Mit Gl. (3.67) folgt damit bei Temperaturänderung

$$I(T) = I_0(T_0) e^{\lambda_T(T-T_0)} e^{U/U_{th}} \quad (3.70)$$

und dementsprechend für die Änderung:

$$\frac{dI}{dT} = I_0(T_0) \cdot \left(\lambda_T + \frac{1}{U_{th}} \cdot \frac{dU}{dT} \right) e^{[\lambda_T(T-T_0)+U/U_{th}]} = I(T) \cdot \left(\lambda_T + \frac{1}{U_{th}} \cdot \frac{dU}{dT} \right) \quad (3.71)$$

Für den Fall gleicher Ströme bei zwei verschiedenen Temperaturen, d.h.

$$\frac{dI}{dT} = 0$$

ergibt sich die damit verbundene Spannungsänderung aus

$$\lambda_T + \frac{1}{U_{th}} \cdot \frac{dU}{dT} = 0$$

zu:

$$\frac{dU}{dT} = -\lambda_T \cdot U_{th} = -d_T \quad \text{bei} \quad I = \text{Konst.} \quad (3.72)$$

Im linearen Maßstab verschiebt sich damit die Kennlinie im Durchlassbereich praktisch parallel zur Stromachse, wobei die Verschiebung ΔU bei konstantem Strom I proportional zur Temperaturänderung ΔT ist, solange ΔT klein ist.

Damit gilt also für diesen Fall::

$$\begin{aligned} U(T) &= U(T_0) - d_T (T - T_0) \\ \Delta U &= U(T_0) - U(T) = -d_T \Delta T \end{aligned} \quad (3.73)$$

Zahlenbeispiel:

Für $T = 300\text{K}$ und $\Delta T = 10\text{K}$

Material	$d_T / \text{mV/K}$	$\Delta U (\Delta T = 10\text{K})$
----------	---------------------	------------------------------------

Si	3,6	36 mV
Ge	2,2	22 mV
GaAs	4,7	47 mV

Tab. 3.6 Temperaturdurchgriff für verschiedene Materialien

Für die Steigung der Kennlinie im logarithmischen Maßstab folgt aus

$$I \approx I_o e^{U/U_{th}} \quad , (U \gg U_{th})$$

mit

$$\ln \frac{I}{I_o} = U / U_{th}$$

$$\frac{d\left(\ln \frac{I}{I_o}\right)}{dU} = \frac{1}{U_{th}} = \frac{q}{KT} \propto \frac{1}{T} \quad (3.74)$$

In Durchlassrichtung dominiert der Exponentialfaktor $\exp(U/U_{th})$ die Temperaturabhängigkeit. Im einfach-logarithmischen Maßstab ist die Steigung der Geraden I/I_o proportional zu $1/U_{th} = q/kT$. Sie nimmt mit steigender Temperatur ab (Abb. 3.8).

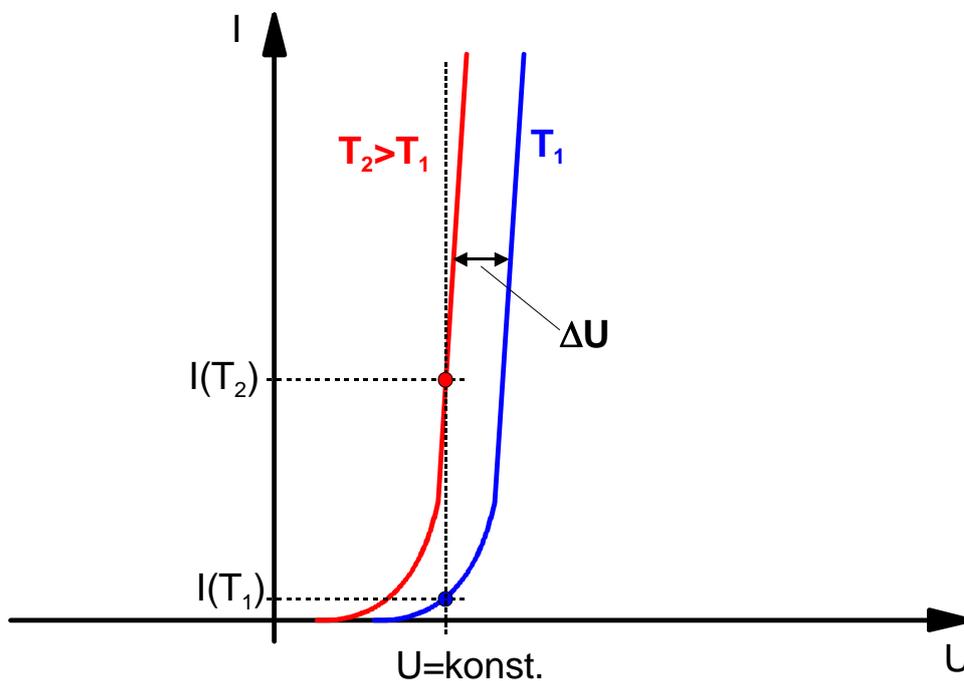


Abb. 3.7 Temperaturabhängigkeit der Diodenkennlinie im linearen Maßstab

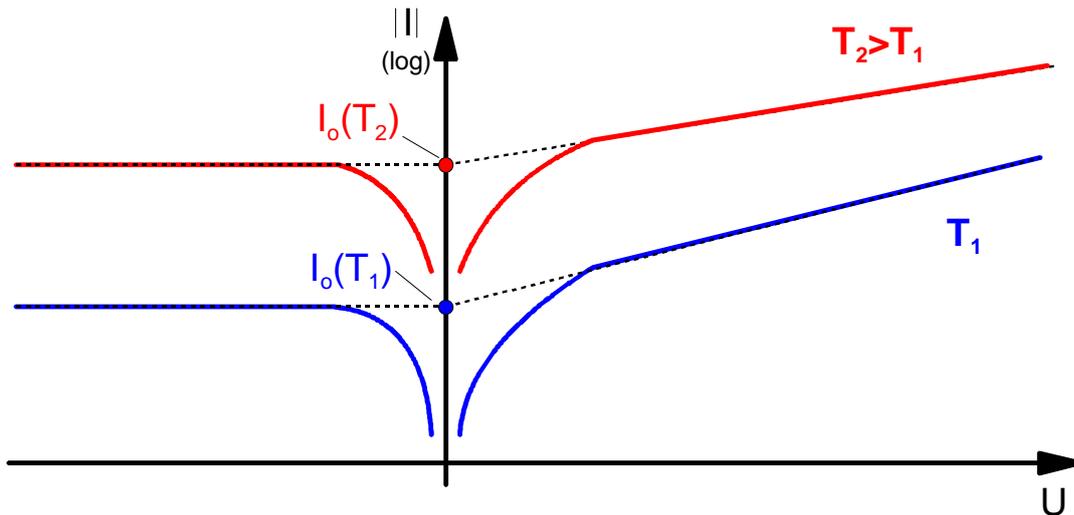


Abb. 3.8 Temperaturabhängigkeit der Diodenkennlinie im einfach logarithmischen Maßstab

3.1.6 Das Elektrische Ersatzschaltbild der pn-Diode

Wie bereits in Kap. 3.4 beschrieben, muss man auch hier bei der Beschreibung des Bauelements durch ein passendes Modell zwischen dem DC- und AC-Verhalten unterscheiden.

3.1.6.1 Das Gleichstrom- (DC-)Verhalten

Nach Gl. (3.46) gilt:

$$J = J_o \left(e^{\bar{U}/U_{th}} - 1 \right)$$

wobei für die Berechnung der Stromdichten J bzw. J_o jeweils nach Lang- und Kurzbasisdiode unterschieden wurde.

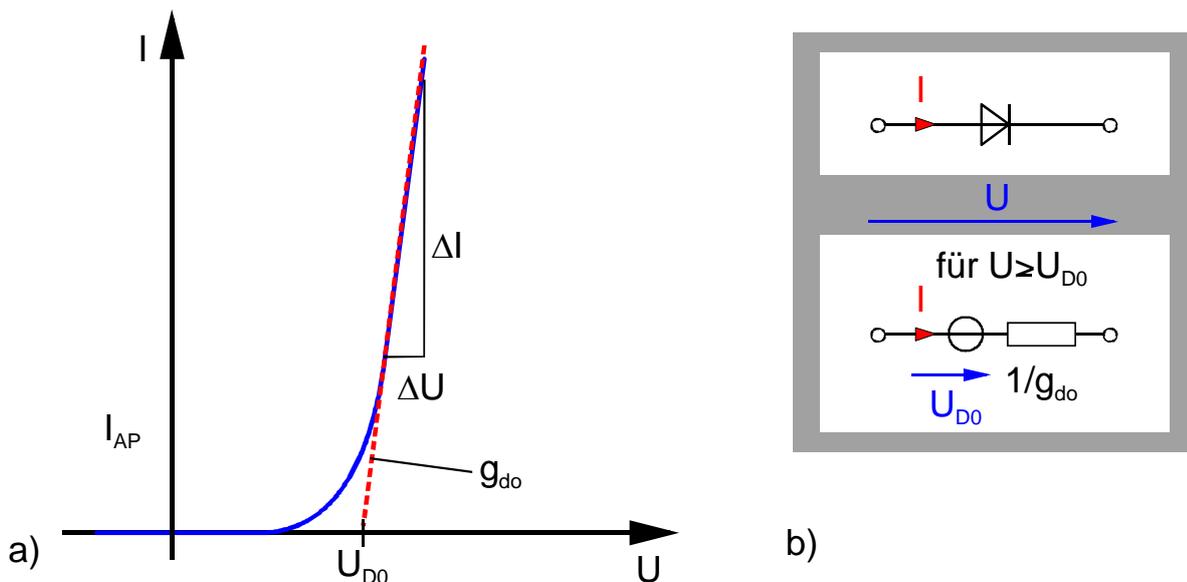


Abb. 3.9 a) DC-Kennlinie der pn-Diode und b) einfaches Ersatzschaltbild für $U \geq U_{D0}$

Der Strom hängt exponentiell von der angelegten Spannung ab. Für relativ große Ströme ändert sich die Steigung der Kennlinie nur geringfügig und das Verhalten der pn-Diode lässt sich in diesem

Bereich linearisieren (Abb. 3.9).

Für Spannungen $U \geq U_{D0}$ gilt für die an der Diode abfallende Spannung U näherungsweise:

$$U \approx U_{D0} + I \cdot \frac{1}{g_{do}} \quad (3.75)$$

Der Leitwert g_{do} ergibt sich aus der Tangente an die Kennlinie für große Ströme (s. Abb. 3.9):

$$g_{do} = \frac{\Delta I}{\Delta U} = \frac{1}{r_{do}}$$

und U_{D0} folgt bei Extrapolation der Tangente aus dem Schnittpunkt mit der Spannungsachse.

3.1.6.1.1 Das quasistationäre Kleinsignalverhalten von pn-Dioden

An dieser Stelle soll nur auf das quasistationäre Kleinsignalverhalten der pn-Dioden, also das Wechselspannungsverhalten bei sehr niedrigen Frequenzen eingegangen werden. Das Verhalten bei höheren Frequenzen wird in der Vorlesung "Elektronische Schaltungen" behandelt.

Wie bereits in Kap. 3.4.2 für die Beschreibung des Kleinsignalverhaltens von Feldeffekttransistoren wird auch in diesem Fall der Gleichspannung im Arbeitspunkt eine Wechselspannung überlagert.

Dabei werden folgende Annahmen gemacht:

Für die Gesamtspannung gilt:

$$U = \bar{U} + \tilde{u}$$

wobei \tilde{u} eine zeitlich periodische Funktion darstellt mit

$$\tilde{u} = \hat{u} \cdot e^{j\omega t}$$

und für die Amplitude gilt:

$$\hat{u} \ll U_{th} = \frac{KT}{q} \quad (3.76)$$

Weiterhin seien Laufzeiten im Bauelement vernachlässigbar, d.h. die Ladungsträgerkonzentrationen $p_n(l_n), n_p(l_p)$ an den Grenzen der Raumladungszonen stellen sich momentan ein.

Im vorangegangenen Kapitel wurde die Kennlinie im Fall größerer Ströme linearisiert und nur durch eine Spannungsquelle und einen **konstanten** differentiellen Leitwert beschrieben.

Das Verhalten bei kleinen Strömen wird damit nicht richtig beschrieben. Wird in diesem Bereich einem Arbeitspunkt eine Wechselspannung überlagert, so lässt sich das Verhalten durch einen Arbeitspunkt abhängigen Leitwert G_d (s. Abb. 3.10) beschreiben:

$$g_d = \left. \frac{dI}{dU} \right|_{AP} \neq konst.$$

Wie aus Abb. 3.9 zu erkennen ist, kann dieser Leitwert bei kleinen Strömen stark von dem Wert G_{d0} aus Kap. 4.2.1.6.1 abweichen.

Aus

$$I = J \cdot A = A \cdot J_o \left(e^{\bar{U}/U_{th}} - 1 \right) = I_o \left(e^{\bar{U}/U_{th}} - 1 \right)$$

mit A als Querschnittsfläche folgt:

$$I(U) = I_o \left(e^{\bar{U}/U_{th}} - 1 \right) \approx I_o \cdot e^{\bar{U}/U_{th}} \quad \text{für} \quad \bar{U} \gg U_{th}$$

und damit

$$g_d = \frac{dI}{d\bar{U}} = \frac{1}{U_{th}} I_o \cdot e^{\bar{U}/U_{th}} = \frac{I}{U_{th}} = f(I) = f(AP) \tag{3.77}$$

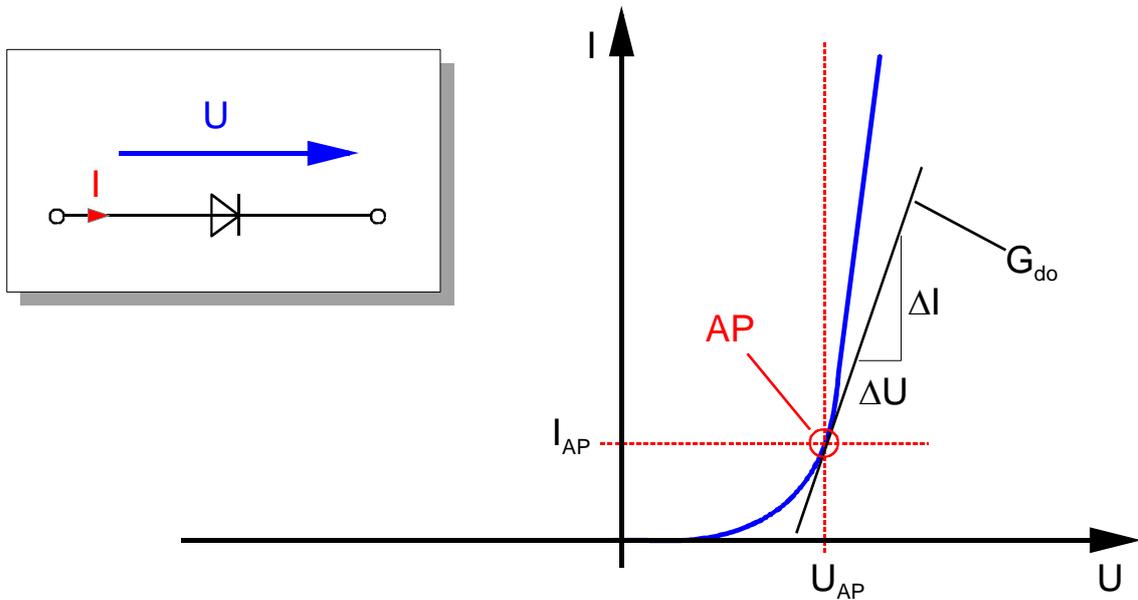


Abb. 3.10 Strom-Spannungs-Kennlinie einer pn-Diode und Definition des differentiellen (Diffusions-)Leitwertes im Arbeitspunkt (AP)

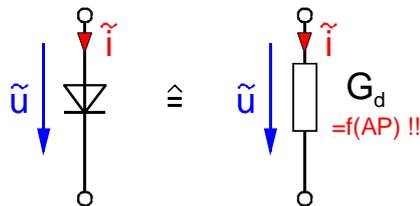


Abb. 3.11 einfachstes **Kleinsignal**ersatzschaltbild einer pn-Diode bei Kleinsignalaussteuerung im Arbeitspunkt (AP). Es beschreibt nur die Änderung des Stromes!

3.2 Spezielle Dioden

Durch Variation der Störstellenkonzentration, des Dotierstoffprofils beim pn-Übergang und darüber hinaus durch den Aufbau mehrschichtiger Halbleiteranordnungen erhält man Diodenbauelemente mit besonderen Eigenschaften, die meist für spezielle Anwendungen eingesetzt werden. Dazu gehören z.B. die **Varaktor-Diode**, die **Zener-Diode** und die **Tunneldiode**.

3.2.1.1 Varaktor-Diode (Kapazitätsdiode)

Kapazitätsdioden sind in Sperrrichtung gepolte Dioden, bei denen die Änderung der Sperrschichtkapazität C_S in Abhängigkeit von der Sperrspannung ausgenutzt wird. Silizium als Basismaterial hat gegenüber Germanium den Vorteil des kleineren Sperrstromes.

Man kann einen in Sperrrichtung vorgespannten pn-Übergang als Plattenkondensator auffassen, dessen Dielektrikum die Raumladungszone darstellt (Abb. 3.12).

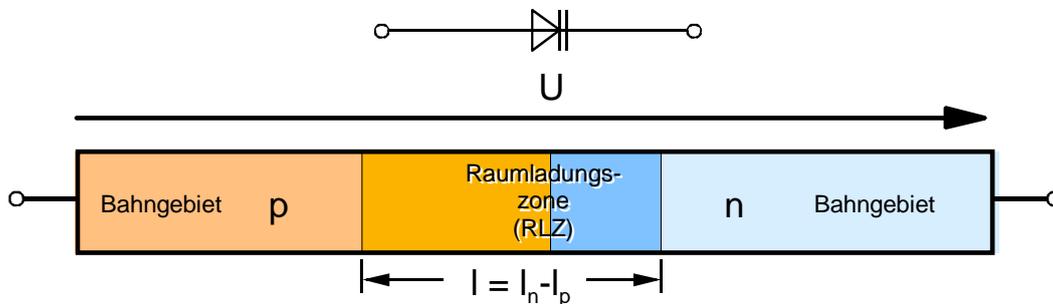


Abb. 3.12 pn-Übergang als Varaktordiode

Die Ausdehnung der Raumladungszone ist durch die Dotierung auf der n- und p-Seite (N_D , N_A) sowie durch die angelegte Spannung und durch die Materialkonstante ϵ_r bestimmt.

Ausgehend von der Kapazität eines Plattenkondensators mit

$$C = \frac{\epsilon_0 \epsilon_r \cdot A}{l} \quad A: \text{Fläche}$$

lässt sich nach Gl. (0.125) die Kapazität der RLZ für einen abrupten pn-Übergang angeben mit

$$C_S = \frac{dQ}{dU} = \frac{\epsilon_0 \epsilon_r \cdot A}{l} = \sqrt{\frac{q \epsilon_0 \epsilon_r N_A N_D}{2(N_A + N_D)}} \cdot \frac{1}{\sqrt{U_D - U}} \cdot A \quad (3.78)$$

bzw.

$$C_S = C_{S0} \cdot \frac{1}{(1 - (U/U_D))^{1/2}} \quad \text{mit} \quad C_{S0} = \sqrt{\frac{q \epsilon_0 \epsilon_r N_A N_D}{2U_D (N_A + N_D)}} \cdot A$$

Gl. (3.78) gilt nur unter der Voraussetzung, dass die Dotierung am pn-Übergang sprunghaft von n- auf p-Dotierung wechselt (abrupter Übergang, Abb. 3.13 a). Allgemein gilt für die Spannungsabhängigkeit der Sperrschichtkapazität

$$C_S = C_{S0} \cdot \frac{1}{(1 - (U/U_D))^m} \quad (3.79)$$

Liegt ein linearer pn-Übergang (Abb. 3.13b) vor, beträgt der Wert $m=0,33$. Bei Dioden mit hyperabrupten Übergang (Abb. 3.13c) ist $m>0,5$. Die Änderung der Sperrschichtkapazität bei gleich großer Änderung der Sperrspannung ist umso stärker, je größer der Exponent m ist.

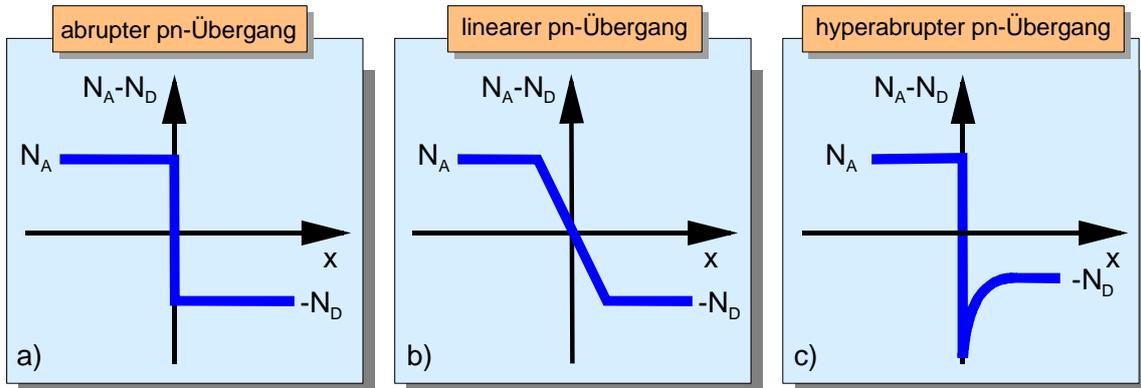


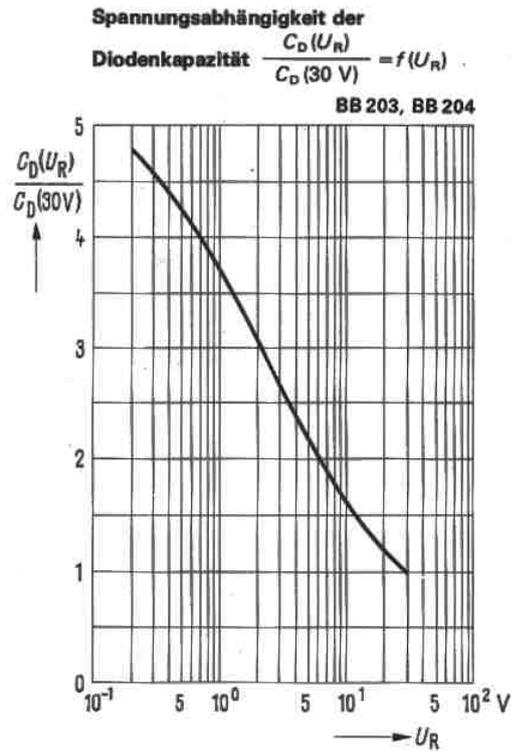
Abb. 3.13 verschiedene Typen von pn-Übergängen
 a) abrupter pn-Übergang
 b) linearer pn-Übergang
 c) hyperabrupter pn-Übergang

Um eine möglichst große Spannungsabhängigkeit der Kapazität zu erhalten, werden besondere Störstellenprofile gezüchtet, bei denen die Störstellenkonzentration nach beiden Seiten ins Halbleiterinnere abnimmt. Dioden, bei denen die Spannungsabhängigkeit der Sperrschichtkapazität ausgenutzt wird, werden als **Varaktordioden** oder **Varicaps** bezeichnet. Man kann Kapazitäten von einigen 100pF realisieren. Der Variationsbereich beträgt etwa 1:2 bis 1:13.

In Abb. 3.14 sind die typischen Kennlinien zweier kommerzieller Kapazitätsdioden dargestellt und in Tab. 3.1 die wichtigsten Kenndaten dieser Dioden angegeben.

statische Kenndaten		BB 203	BB 204	
Durchbruchspannung ($I_R = -I = 10\mu A$)	U_{Br}	> 32	> 32	V
Sperrstrom ($U_R = -U = 30V; T = 25^\circ C$)	I_R	< 20	< 20	nA
Sperrstrom ($U_R = -U = 30V; T = 60^\circ C$)	I_R	< 0,2	< 0,2	μA
dynamische Kenndaten				
Kapazität ($U_R = -U = 3V; T = 25^\circ C, f = 1MHz$)	C_S bzw. C_D	27 ...31	34 ...39	pF
Kapazität ($U_R = -U = 30V; T = 25^\circ C, f = 1MHz$)	C_S bzw. C_D	11	14	pF
Kapazitätsverhältnis	$\frac{C_S(3V)}{C_S(30V)}$ bzw. $\frac{C_D(3V)}{C_D(30V)}$	2,55...2,8	2,4...2,8	

Tab. 3.7 Kenndaten der Kapazitätsdioden in Abb. 3.14



**BB 203
BB 204**

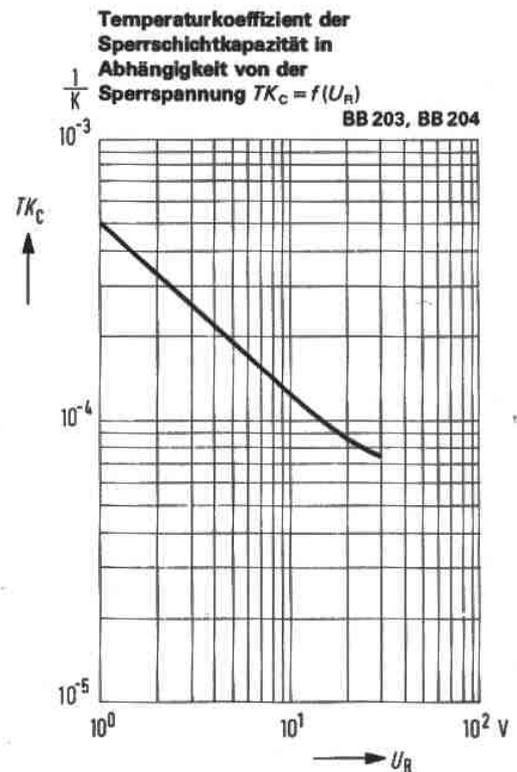
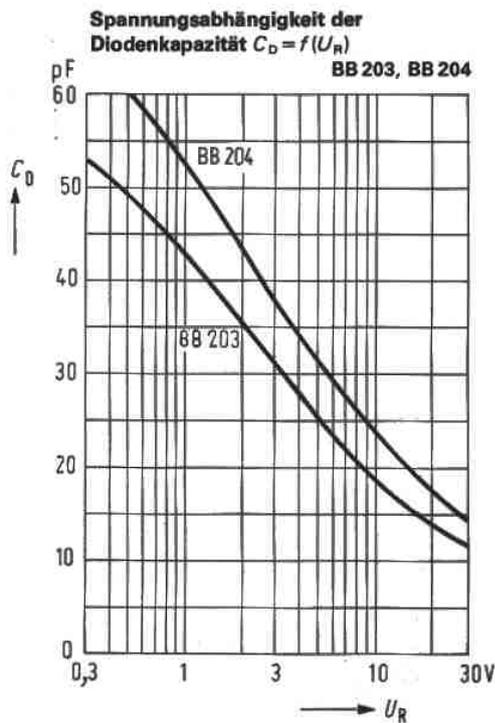
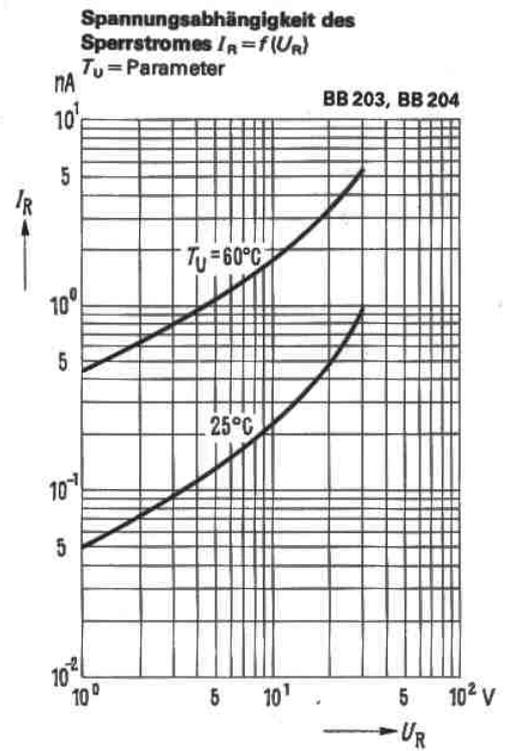


Abb. 3.14 Kennlinien einer Kapazitäts-Diode

Anwendung finden die Kapazitätsdioden vor allem in automatischen Abstimmungen für UKW-Oszillatoren, in Frequenzmodulatoren sowie als Ersatz für die früher häufig eingesetzten Drehkondensatoren.

Abb. 3.15 zeigt die Rasterelektronenmikroskop-(REM-)Aufnahme einer Varaktordiode, die im Fachgebiet Halbleitertechnik/Halbleitertechnologie der Universität Duisburg-Essen auf InP hergestellt wurde.

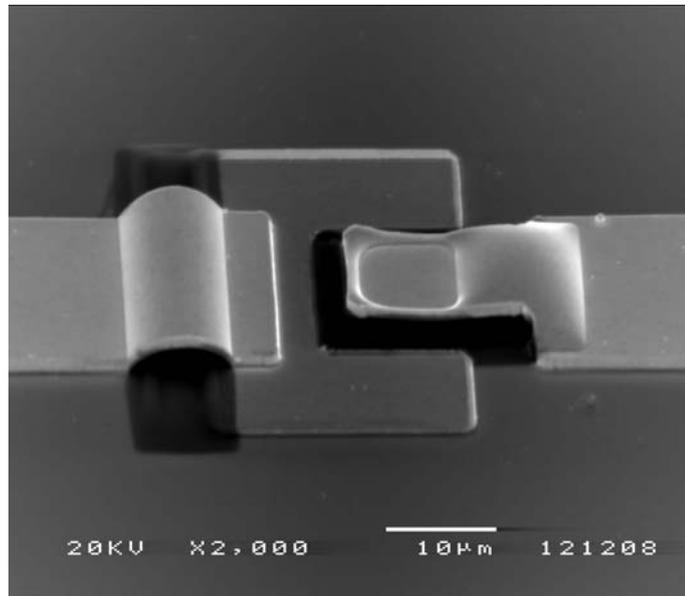


Abb. 3.15 im Fachgebiet Halbleitertechnik/Halbleitertechnologie der Universität Duisburg-Essen hergestellte Varaktordiode auf InP

Anwendungsbeispiel:

Abb. 3.16 zeigt eine von vielen Möglichkeiten der Abstimmung eines Schwingkreises mit Hilfe einer Kapazitätsdiode. Die Kapazitätsdiode D mit der spannungsabhängigen Kapazität C_S ist über den in Serie geschalteten Kondensator C_2 parallel zum Schwingkreis L und C_1 geschaltet. Über den Widerstand R wird der Diode D die Gleichspannung U_B zugeführt, wodurch der Arbeitspunkt und damit die Kapazität der Diode C_S eingestellt wird.

Die Kapazität C_2 soll verhindern, dass die Gleichspannung U_B am Ausgang anliegt, soll aber das Frequenzverhalten nicht beeinflussen. Somit sollte sie deutlich größer als C_S sein, z.B. $C_2 > 5 \cdot C_{S,\max}$. Der Widerstand R muss groß gewählt werden, so dass eine nennenswerte Bedämpfung des Schwingkreises durch die Steuerspannung U_B nicht auftritt, d.h. er muss beträchtlich größer sein als der Resonanzwiderstand des Kreises (z.B. $R = 100\text{k}\Omega$).

Eine ausgangsseitige Wechselfspannung führt jedoch ebenfalls zu einer Kapazitätsänderung der Diode, die von der Amplitude der Wechselfspannung abhängt. D.h. neben der Grundschwingung treten störende Oberwellen auf, so dass die Anwendung von Kapazitätsdioden in dieser einfachen Anordnung auf möglichst kleine Wechselfspannungsamplituden begrenzt ist.

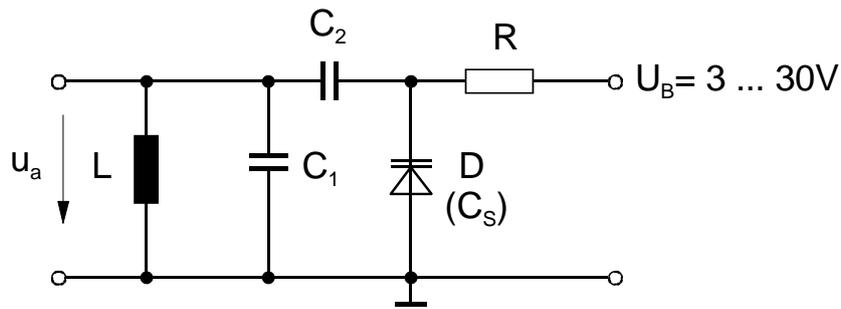


Abb. 3.16 Schwingkreisabstimmung mit Kapazitätsdiode

Die Resonanzfrequenz des Schwingkreises ergibt sich zu

$$f_o = \frac{1}{2\pi \cdot \sqrt{(C_1 + C_S) \cdot L}} \quad (3.79)$$

3.2.2 Zenerdiode

Zenerdioden (Z-Diode) sind Halbleiterdioden mit pn-Übergängen, die im Durchbruchbereich betrieben werden, während sie sich in Durchlassrichtung wie eine konventionelle pn-Diode verhalten. Durch entsprechende Herstellungstechnologien wird erreicht, dass der Übergang von sehr kleinem zu sehr großem Sperrstrom sehr steil bei einer Sperrspannung U_{Z0} geschieht. In Abb. 3.17 ist die prinzipielle Strom-Spannungskennlinie einer Z-Diode gezeigt.

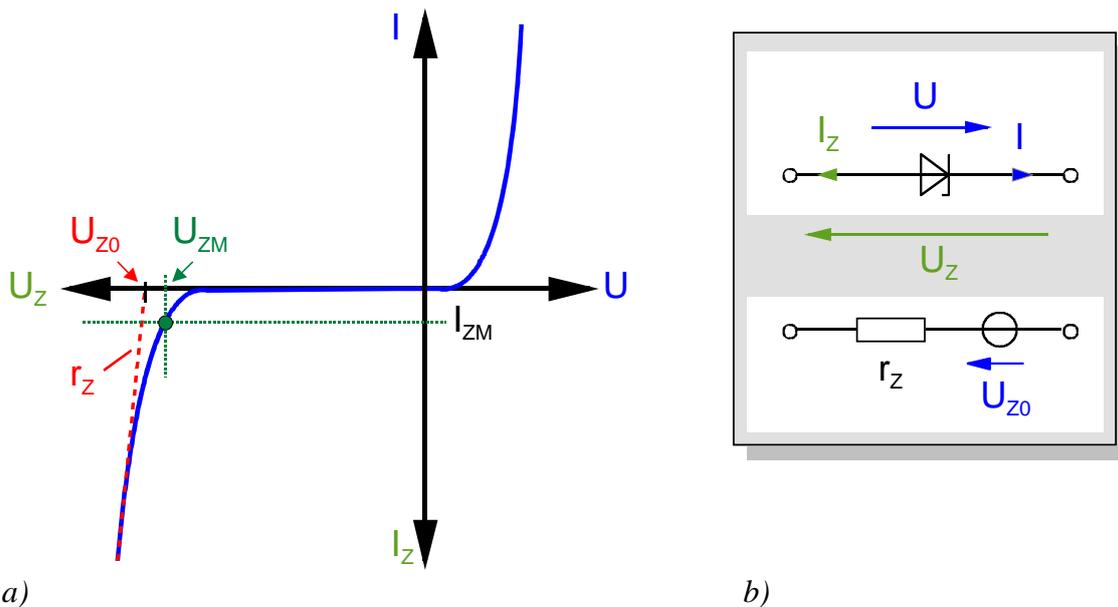


Abb. 3.17 a) Kennlinie der Z-Diode und
b) einfaches DC-Ersatzschaltbild

In Sperrrichtung, jenseits des Durchbruchs, lässt sich die Z-Diode durch ein einfaches DC-Ersatzschaltbild bestehend aus einer Spannungsquelle U_{Z0} und einem Widerstand r_Z , der dem differentiellen Widerstand der Kennlinie im Durchbruchbereich entspricht, beschreiben (Abb. 3.17 b).

Der Durchbruch, der zu einem sprunghaften Anstieg des Stromes in Sperrrichtung führt, hat zwei wesentliche Ursachen:

- Tunnel- (Zener-)Effekt
- Lawineneffekt-Effekt

die im folgenden erklärt werden.

3.2.2.1 Tunnel-(Zener-)Effekt

Steigt die Sperrspannung, so erhöht sich die Feldstärke an der Sperrschicht. Bei genügend hohen Sperrspannungen können Elektronen aus besetzten Zuständen des Valenzbandes in unbesetzte Zustände des Leitungsbandes tunneln (Abb. 3.18). Die Tunnelwahrscheinlichkeit hängt dabei exponentiell von der zu durchtunnelnden Breite a und der Höhe der Potentialbarriere ab.

Mit zunehmender Dotierung der p- und n-Gebiete nimmt die Breite der Raumladungszone l ab und damit die Tunnelwahrscheinlichkeit zu. Der Zener-Effekt ist daher überwiegend in Dioden mit hochdotierten Bereichen zu erwarten. Die Höhe der Barriere entspricht in etwa dem Bandabstand W_g , da die Elektronen aus dem Valenzband (p-Gebiet) ins Leitungsband (n-Gebiet) angehoben werden müssen. Mit zunehmender Temperatur nimmt die Tunnelwahrscheinlichkeit zu, die Sperrspannung U_{Z0} (auch Zener-Spannung genannt), bei der der Tunnelstrom merklich zunimmt, sinkt infolgedessen ab.

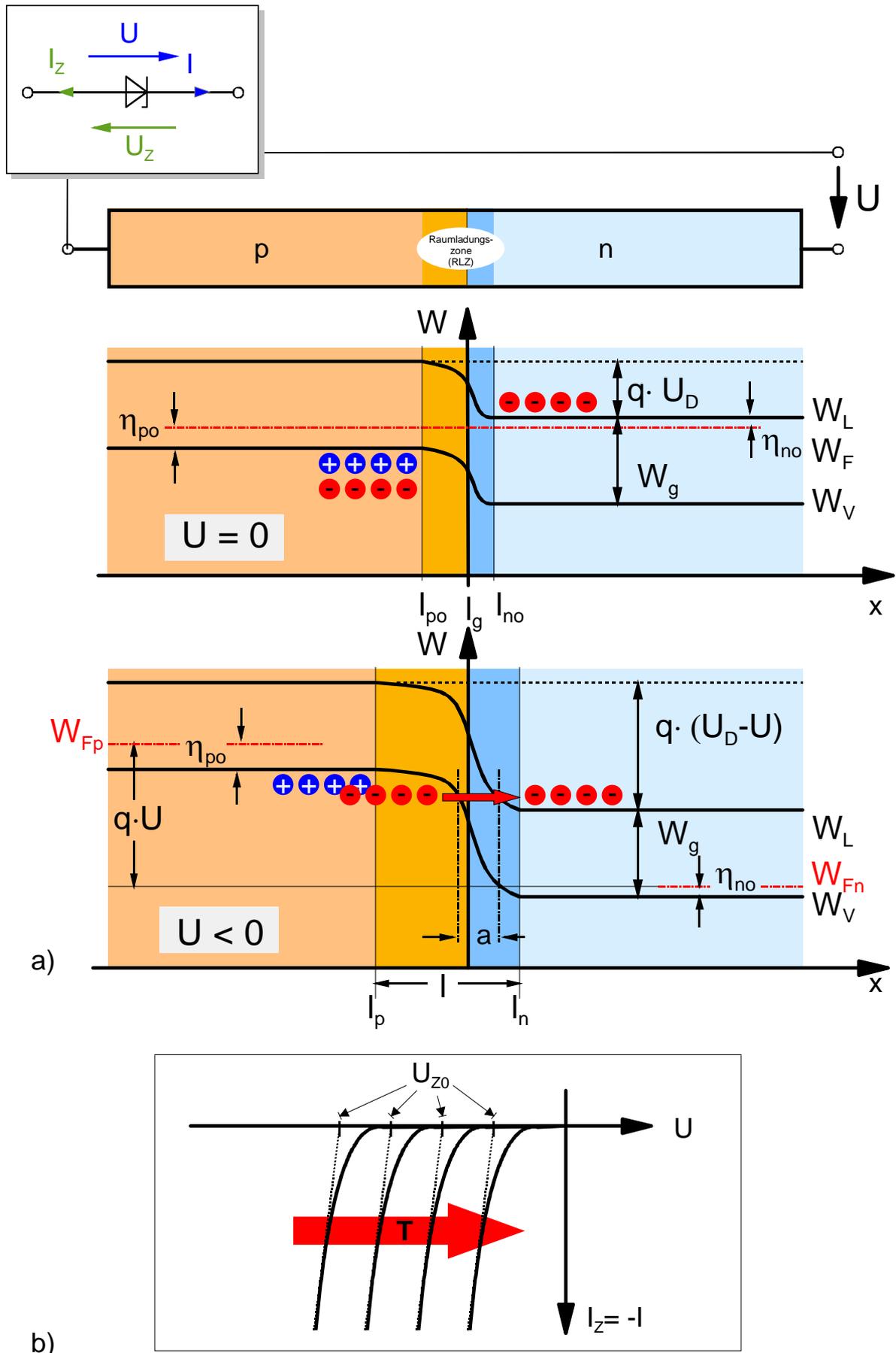


Abb. 3.18 a) Der Tunnel- (Zener-)Effekt bei hohen Sperrspannungen und
 b) Temperaturabhängigkeit der Zener-Spannung

3.2.2.2 Lawineneffekt

Sind die Dotierungen relativ gering, so dehnt sich Raumladungszone weit in die Bahngebiete aus und die Feldstärke am Übergang nimmt deutlich ab.

In diesem Fall können sich die Ladungsträger eine gewisse Strecke λ bewegen, ohne mit Gitterbausteinen oder Fremdatomen zusammenzustoßen. Dabei können sie kinetische Energie aus dem elektrischen Feld aufnehmen ((1) in Abb. 3.19). Ist die aufgenommene Energie $W_{kin} \geq W_g$, so kann bei Wechselwirkung des Ladungsträgers mit einem Gitteratom die kinetische Energie auf ein Elektron in einer besetzten Bindung übertragen werden, so dass dieses Elektron in das Leitungsband angehoben wird ((2) in Abb. 3.19).

Nach dem Zusammenstoß sind zwei Elektronen und ein Loch vorhanden, die erneut beschleunigt werden und die Lawinenbildung auslösen können ((3) in Abb. 3.19).

Mit steigender Temperatur nimmt die mittlere freie Weglänge der Elektronen zwischen zwei Zusammenstößen ab. Damit längs der kürzeren Strecke wieder die notwendige kinetische Energie $W_{kin} \geq W_g$ erreicht wird, müssen die Feldstärke und damit auch die notwendige Sperrspannung erhöht werden. Die Zener-Spannung U_{Z0} nimmt also mit steigender Temperatur zu.

Vergleicht man den entgegengesetzten Temperaturgang der Durchbruchspannungen bei Zener- und bei Lawineneffekt, so muss es offensichtlich eine Dotierung geben, bei der sich beide Effekte kompensieren, so dass die Zener-Spannung temperaturunabhängig wird. In Silizium-Dioden, deren p-Gebiet beispielsweise sehr hoch dotiert ist ($N_A \gg N_D$), beträgt diese Dotierung $N_D = 4 \cdot 10^{17} \text{ cm}^{-3}$. Die zugehörige Zener-Spannung beträgt etwa $U = -U_{Z0} \approx -6\text{V}$ bei $T = 300\text{K}$.

Abb. 3.20. zeigt die Sperrcharakteristik verschiedener Zener-Dioden.

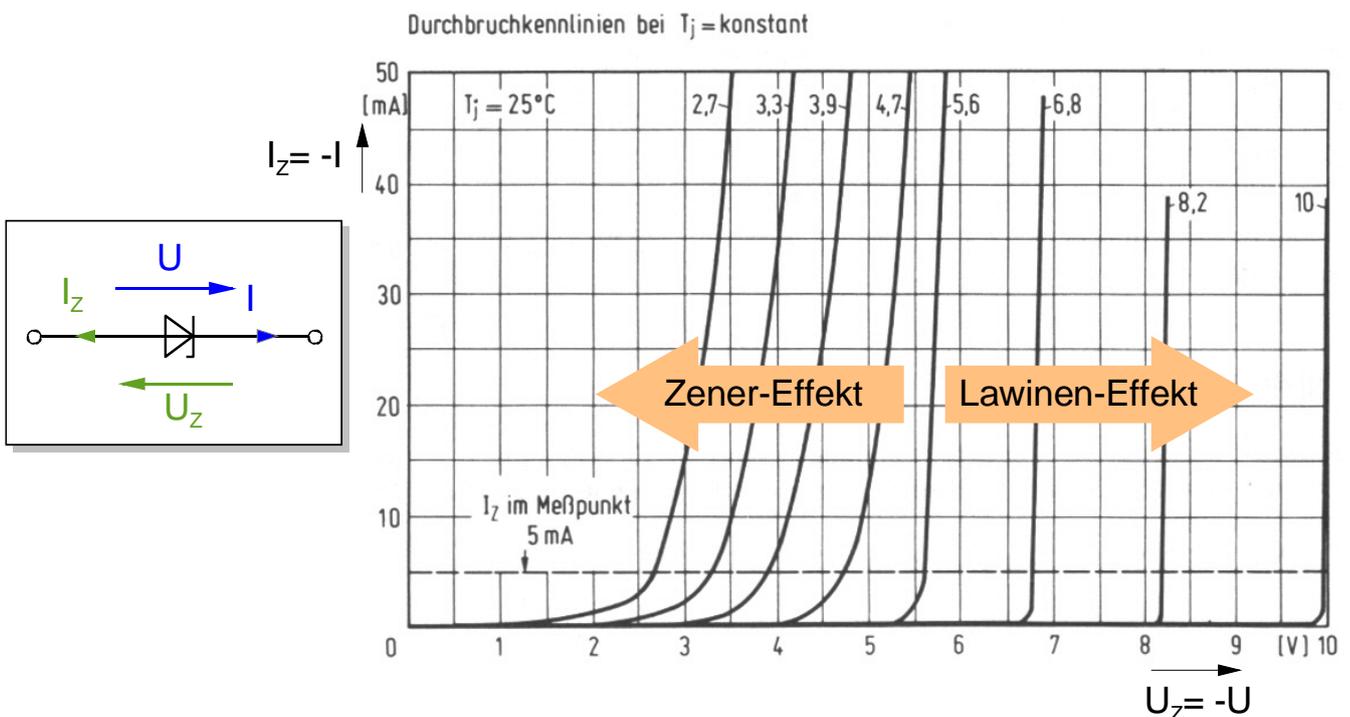


Abb. 3.20 Strom-Spannungs-Kennlinie diverser Zenerdioden in Sperrrichtung

Man erkennt das unterschiedliche Durchbruchverhalten der Z-Dioden:

In Dioden mit relativ hohen Dotierungen tritt der Zener-Effekt bereits bei kleinen Spannungen auf und ist somit in Dioden mit kleiner Zener-Spannung für den Durchbruch "verantwortlich". Der Durchbruch erfolgt allerdings nicht so abrupt wie in Z-Dioden mit relativ niedrigen Dotierungen, in denen der Lawineneffekt für den sprunghaften Anstieg des Sperrstromes sorgt.

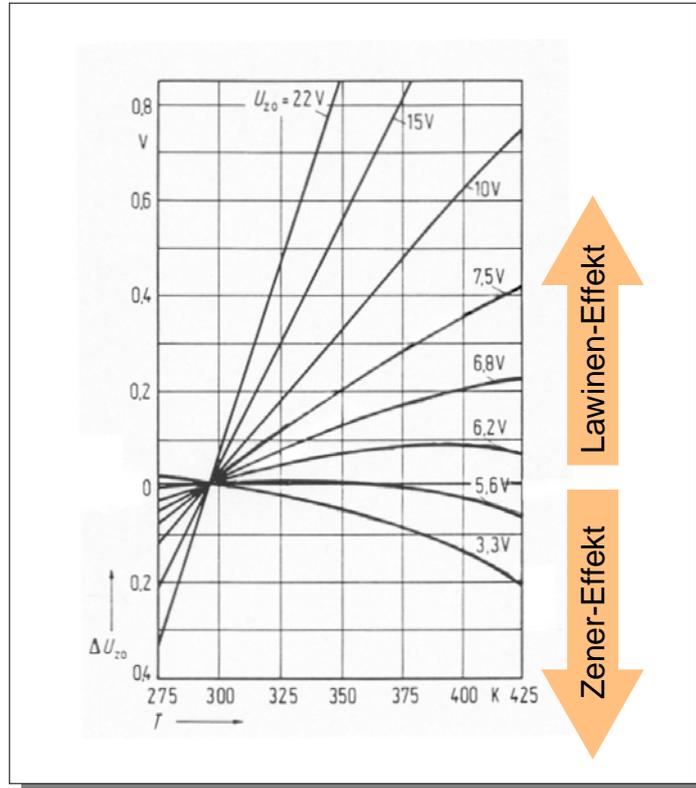


Abb. 3.21 Temperaturabhängigkeit der Zener-Spannung [17]

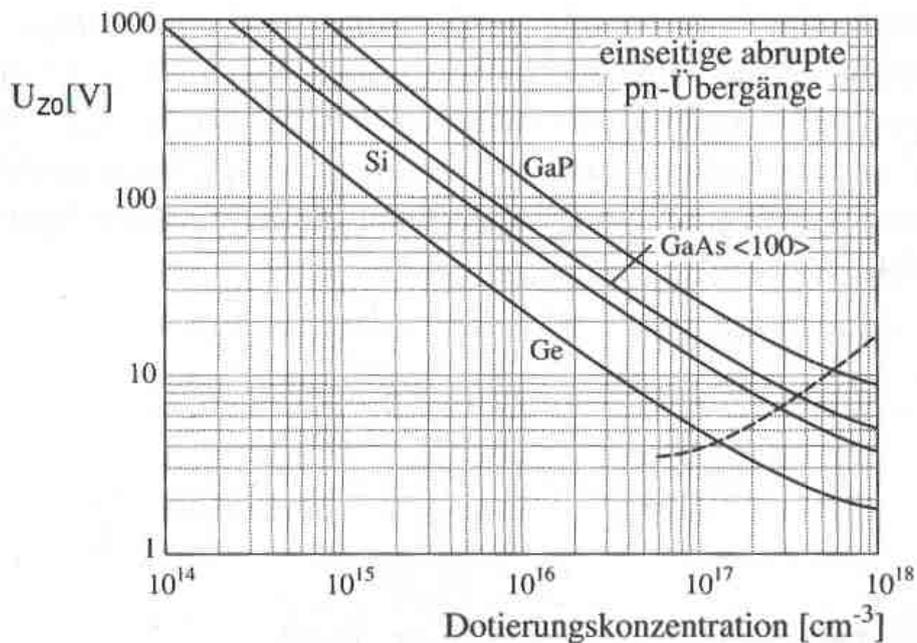


Abb. 3.22 Durchbruchspannung in Abhängigkeit von der Dotierung für verschiedene Halbleitermaterialien [19]; die gestrichelte Linie gibt die Grenze zwischen Lawinen- und Zenerdurchbruch an

Abb. 3.21 zeigt die Temperaturabhängigkeit der Zener-Spannung. Man erkennt das oben beschriebene unterschiedliche Verhalten bei kleinen und großen Werten.

Es gilt:

$$\Delta U_{Z0} = \alpha_T \cdot U_{Z0} \cdot \Delta T \quad (3.80)$$

mit

$$\alpha_T < 0 \quad \text{für} \quad U_{Z0} \leq 6V \text{ (Zenerdurchbruch)}$$

$$\alpha_T > 0 \quad \text{für} \quad U_{Z0} \geq 6V \text{ (Lawinendurchbruch)}$$

Die Durchbruchspannung hängt neben der Dotierung auch vom eingesetzten Halbleiter ab, wie Abb. 3.23 zeigt.

Anwendungsbeispiel:

Aus der Kennlinie ist die spannungsstabilisierende Wirkung der Zenerdiode abzulesen. In Sperrichtung betrieben wird dieses Bauelement daher vor allem für Schaltungen zur Spannungsstabilisierung eingesetzt.

Abb. 3.23 zeigt eine einfache Schaltung zur Stabilisierung der Ausgangsspannung mit Hilfe einer Z-Diode. Bei Schwankungen der Eingangsspannung U_1 sowie des Ausgangsstromes I_L soll die Ausgangsspannung U_L möglichst konstant bleiben.

Dies funktioniert umso besser, je kleiner der differentielle Widerstand r_Z ist, d.h. je steiler die Kennlinie im Durchbruchbereich verläuft.

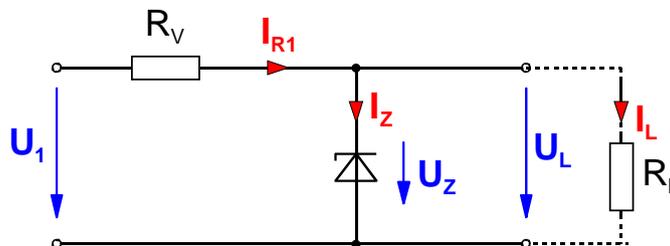


Abb. 3.23 Einfache Spannungsstabilisierung mit Z-Diode

3.2.3 Tunnel-Diode

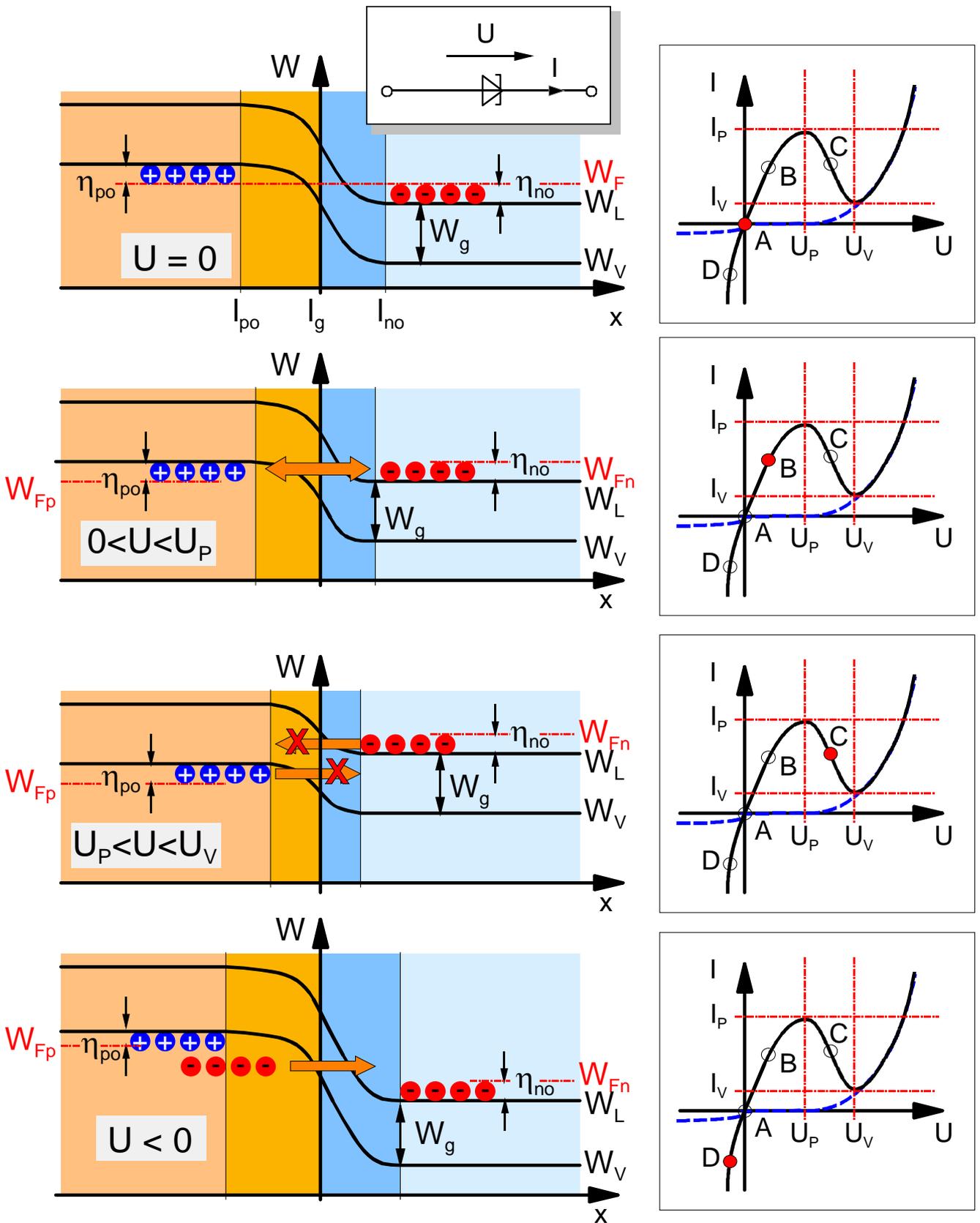


Abb. 3.24 Tunnel-Effekt

Bei sehr hohen Dotierungen ($N_D > N_L$ bzw. $N_A > N_V$) werden sowohl Sperrstrom als auch Durchlassstrom durch den Tunnelprozess bestimmt. Wenn die Dotierungskonzentrationen höher als die effektiven Zustandsdichten sind, ist der Halbleiter **entartet** und die Fermi-Energie liegt innerhalb der jeweiligen Bänder (Abb. 3.24a, $U = 0$).

Bei kleinen Flussspannungen (Abb. 3.24b, $0 < U < U_P$) können Elektronen aus dem Leitungsband in freie Plätze im Valenzband tunneln: der Strom steigt mit wachsender Spannung stark an. In einer nicht bis zur Entartung dotierten Diode können Elektronen nur die Barriere überwinden, nicht jedoch tunneln. Es gilt die gestrichelt gezeichnete Kennlinie.

Mit zunehmender Spannung geht die Tunnelwahrscheinlichkeit wieder zurück (Abb. 3.24c, $U_P < U < U_V$), weil die Elektronen des Leitungsbandes der n-Seite sich der verbotenen Zone der p-Seite "gegenübersehen", in die hinein sie nicht tunneln können, weil es keine besetzbaren Zustände gibt. Entsprechendes gilt für die Löcher im p-Gebiet. Der Tunnelstrom sinkt daher wieder ab.

Bei noch höheren Spannungen ($U > U_V$) setzt die Bewegung der Ladungsträger über die Barriere hinweg ein: die bekannte Dioden-Kennlinie wird erreicht .

In Sperrrichtung (Abb. 3.24d, $U < 0$) können wegen der Entartung beider Seiten bereits bei sehr kleinen Spannungen sehr viele Elektronen aus dem Valenzband der p-Seite in das Leitungsband der n-Seite tunneln, wie es im Fall des Zener-Effekts bereits beschrieben wurde (Kap. 4.1.4.2.1)

Für Spannungen $U_P < U < U_V$ weist die Diodenkennlinie einen Bereich negativen differentiellen Widerstandes (*NDR: Negative Differential Resistance*)

$$r = \frac{dU}{dI} < 0 \quad (3.81)$$

auf, d.h. trotz zunehmender Spannung nimmt der Strom ab. Aufgrund dieses negativen, differentiellen Widerstandes kann die Tunneldiode zur Entdämpfung von Schwingkreisen und damit zur Schwingungserzeugung (Oszillatoren) eingesetzt werden. Da der quantenmechanische Tunnelprozess sehr schnell ist, können Schwingungen sehr hoher Frequenz, jedoch nur relativ kleiner Amplitude erzeugt werden.

3.2.4 Backward-Diode

Das Verhältnis I_P / I_V ist umso ausgeprägter, je stärker die Entartung auf beiden Seiten ist. Bei geringer oder gerade noch nicht erreichter Entartung fehlt der Tunnelstrom in Durchlassrichtung ganz, während er in Sperrrichtung auch bei kleinen Spannungen noch vorhanden ist. Diese Diode leitet in Sperrrichtung besser als in Flussrichtung und wird, da sie in Sperrrichtung betrieben wird, **Backward-** oder **Back-Diode** genannt. Sie eignen sich zum Gleichrichten kleiner Hochfrequenzspannungen, die in konventionellen pn-Dioden mit deutlich höherer Flussspannung unterdrückt oder verfälscht würden.

Die Bedeutung dieser Dioden ist allerdings rückläufig, da kleine HF-Spannungen mittlerweile wesentlich besser verstärkt werden können bzw. zur Gleichrichtung Schottky-Dioden mit besonders kleiner Schwellenspannung (Zero-Bias-Detektordioden) zur Verfügung stehen.

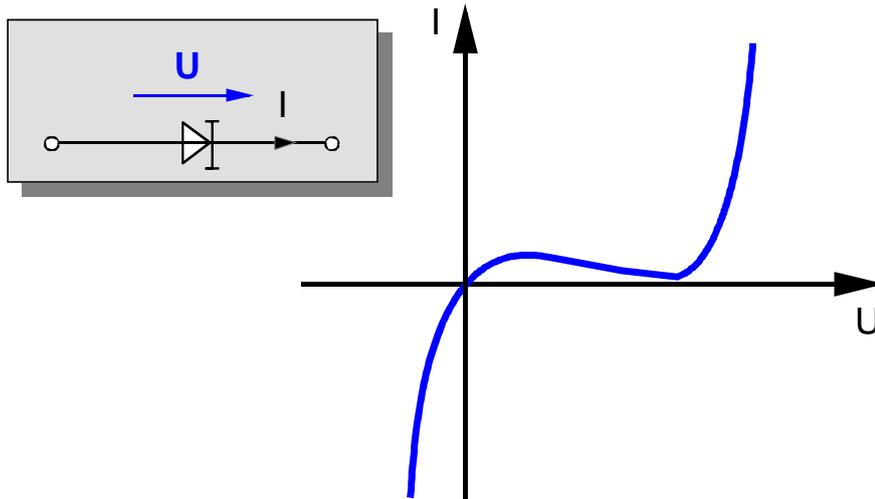


Abb. 3.25 Kennlinie der Backward-Diode

3.2.5 pin-Diode

Für Leistungsanwendungen werden Dioden mit einerseits guter Leitfähigkeit in Durchlassrichtung, andererseits hohen Sperrspannungen benötigt. Dies lässt sich durch Einbringen einer zusätzlichen intrinsischen bzw. nur schwach dotierten Schicht zwischen die hoch p- und n-dotierten Gebiete erreichen (Abb. 3.26a).

Diese Struktur ergibt eine gute Leitfähigkeit in Durchlassrichtung, da der mittlere Bereich von den beiden stark dotierten Randzonen mit Ladungsträgern überschwemmt wird. Dieser Effekt wird als Trägerinjektion bezeichnet.

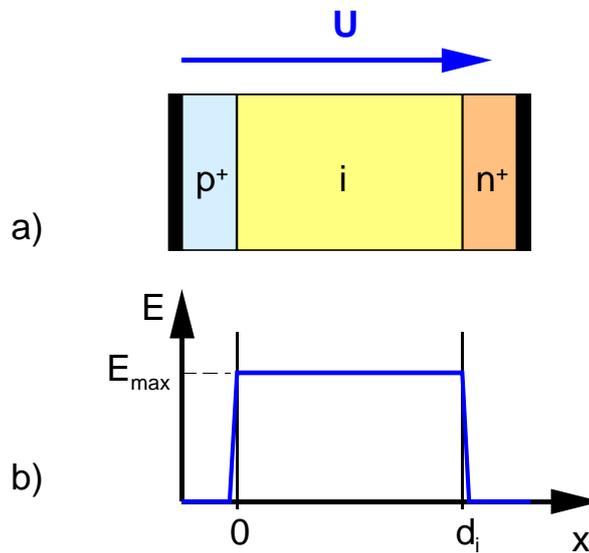


Abb. 3.26 p^+in^+ -Diode
 a) prinzipieller Querschnitt
 b) Feldstärkeverlauf

Gleichzeitig erhält man wegen des großen spezifischen Widerstands der i-Zone eine hohe Sperrspannung. Der größte Teil der angelegten Sperrspannung fällt über dieser eigenleitenden bzw. schwach dotierten Zwischenschicht ab, und es ergibt sich der in Abb. 3.26b dargestellte prinzipielle Feldstärkeverlauf.

Über die Dicke und die schwache Dotierung der mittleren Schicht lässt sich die maximal mögliche Sperrspannung einstellen. Nach [20] ergibt sich die Durchbruchspannung U_{Br} zu:

$$U_{Br} = E_{krit} \cdot d_i - \frac{q \cdot N_i \cdot d_i^2}{2\epsilon_0 \epsilon_r} \quad (3.82)$$

Dabei ist E_{krit} die kritische Durchbruchfeldstärke. Bei Silizium gilt $E_{krit} \approx 3 \cdot 10^5 \text{ V/cm}$. N_i ist die Dotierung der schwach dotierten mittleren Zone, d_i die Dicke der mittleren Schicht.

Darüber hinaus kann die pin-Diode zur Realisierung optoelektronischer Bauelemente wie z.B. Fotodetektoren (Fotodiode) eingesetzt werden. In diesem Fall wird die Diode in Sperrrichtung betrieben. Durch Einstrahlen von Licht werden Elektronen-Loch Paare erzeugt und der Sperrstrom steigt somit bei Lichteinfall an. Bzgl. weiterer Details sei an dieser Stelle auf die Veranstaltungen und Literatur aus dem Bereich der Optoelektronik verwiesen.

In Abb. 3.27 ist eine im Fachgebiet Halbleitertechnik/Halbleitertechnologie der Universität Duisburg-Essen hergestellte pin-Diode auf InP dargestellt.

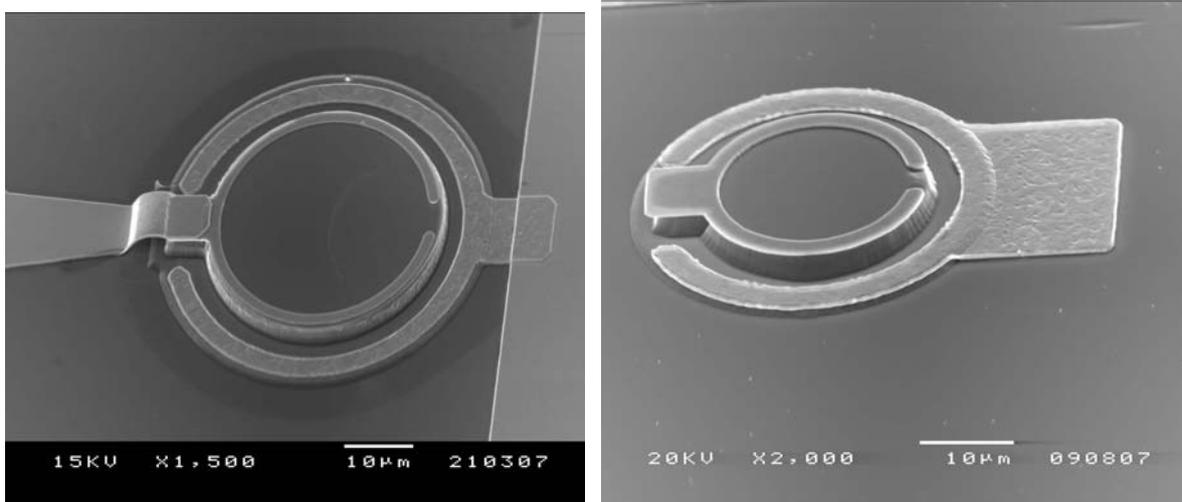


Abb. 3.27 im Fachgebiet Halbleitertechnik/Halbleitertechnologie der Universität Duisburg-Essen hergestellte pin-Diode auf InP

3.2.6 Schottky-Dioden

Die Schottky-Diode wurde bereits in Kap. 3 in Zusammenhang mit dem MESFET behandelt. Bei dieser Art von FET wird sie zur Modulation des leitenden Kanals eingesetzt. Schottky-Dioden werden aber auch als diskrete Bauelemente eingesetzt.

Da in diesem Bauelement nur die Majoritätsladungsträger zum Stromfluss beitragen und es sich somit nicht um ein bipolares Bauelement handelt, sei sie auch hier nur der Vollständigkeit halber erwähnt.

Wird der Metall-Halbleiterübergang in Sperrrichtung vorgespannt, bildet sich eine isolierende Sperrschicht.

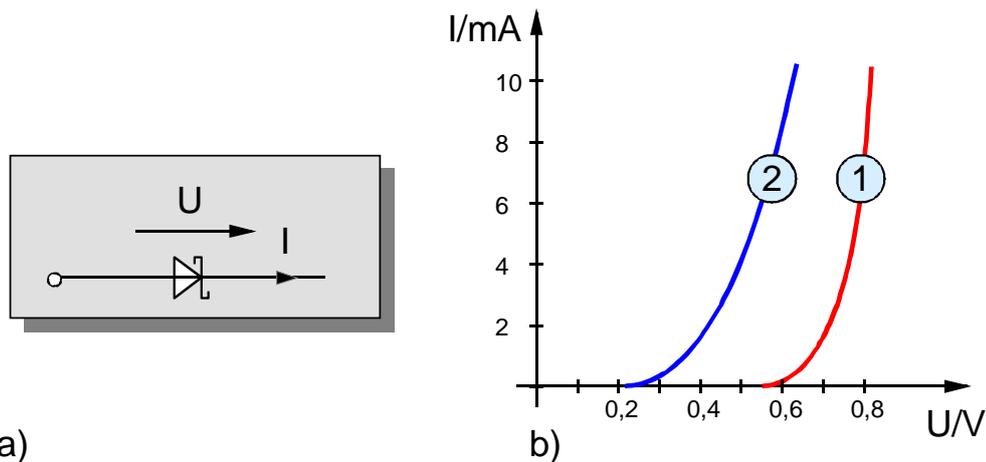


Abb. 3.28 a) Schaltbild der Schottky-Diode

b) Durchlasskennlinien von Signal-Dioden: konventionelle pn-Diode (1) und Schottky-Diode (2)

Da nur Majoritätsladungsträger eines Typs (n oder p) bei Änderung der anliegenden Spannung verschoben werden müssen, sind Schottky-Dioden besonders für HF-Anwendungen gut geeignet. Abhängig von ihrem Aufbau und der Kontaktierung eignen sich diese Schottky-Dioden bis zu Frequenzen von 40GHz. Wegen der hohen Schaltgeschwindigkeit heißen sie auch *Hot-Carrier-Dioden*.

Die Durchlassspannung einer Schottky-Diode ist bei kleinen Strömen (0,1mA bis 1mA) wesentlich kleiner (0,2 V bis 0,4V) als bei Sperrschichtdioden (0,6 bis 0,7V) .

Die hohen Arbeitsfrequenzen erfordern ein System mit kleiner Kapazität, die sich u.a. durch eine kleine aktive Fläche des Chips erreichen lässt. Dadurch steigt der Durchlasswiderstand der Diode und die Kennlinie ist weniger steil als bei einer konventionellen pn-Diode (Abb. 3.28).

3.3 Kennzeichnung

Zur Kennzeichnung der unterschiedlichen Dioden sowie auch Transistoren und deren Einsatzbereiche existieren zwei verschiedene Normen:

- die europäische Kennzeichnung (Pro Electron)
- die JEDEC-Norm

3.3.1 Europäische Kennzeichnung (Pro Electron)

Die Dioden-Bezeichnung nach *Pro Electron* besteht aus zwei bis drei Buchstaben und einer darauffolgenden 2stelligen Zahlenfolge (z.B. BAW 75).

Der **erste Buchstabe** der Typenbezeichnung gibt immer das Halbleitermaterial an:

1. Buchstabe	Halbleitermaterial
A	Germanium
B	Silizium
C	Gallium-Arsenid o.ä. (Bandabstand $W_g > 1,3eV$)
D	Indium-Antimonid o.ä. (Bandabstand $W_g < 0,6eV$)
R	Material für Opto-Elemente (z.B. Kadmium-Sulfid)

Tab. 3.8 Kennzeichnung der Dioden (1.Buchstabe) nach Pro Electron-Norm

Der zweite Buchstabe kennzeichnet Art und Funktion:

2. Buchstabe	Funktion
A	Diode
B	Kapazitätsdiode
E	Tunneldiode
G	Mikrowellendiode
H	Magnetfelddiode
X	Vervielfacherdiode
Y	Leistungsdiode
Z	Z-Diode u.ä.

Tab. 3.9 Kennzeichnung der Dioden (2.Buchstabe) nach Pro Electron-Norm

Weitere Informationen findet man auf dem jeweiligen Datenblatt der betreffenden Diode.

Diese Ziffernfolge kann in Klartext oder als Farbcode auf dem Diodengehäuse aufgedruckt sein. Der Farbcode lässt sich anhand der folgenden Tabelle aufschlüsseln und so den Diodentyp ermitteln.

Den ersten Ring erkennt man an der größeren Breite. Er ist auch die Kathode der Diode.

Farbe	1. Ring	2. Ring	3. Ring	4. Ring
schwarz		X	0	0
braun	AA		1	1
rot	BA		2	2
orange		S	3	3
gelb		T	4	4
grün		V	5	5
blau		W	6	6
violett			7	7
grau		Y	8	8
weiß		Z	9	9

Tab. 3.10 Farbcodierung von Dioden) nach Pro Electron-Norm

3.3.2 JEDEC

Alle Typenbezeichnungen nach der *Jedec*-Norm (*JEDEC: Joint Electron Device Engineering Council*) beginnen mit 1N.... Danach folgt eine meist vierstellige Nummer (z.B. 1N4148), die nur anhand des Datenblattes zu weiteren Informationen führt.

Diese Ziffernfolge kann in Klartext oder als Farbcode auf dem Diodengehäuse aufgedruckt sein.

Der Farbcode lässt sich anhand der folgenden Tabelle aufschlüsseln und so den Diodentyp ermitteln. Den ersten Ring erkennt man an der größeren Breite. Er ist auch die Kathode der Diode.

Farbe	1. Ring	2. Ring	3. Ring	4. Ring
schwarz	0	0	0	0
braun	1	1	1	1
rot	2	2	2	2
orange	3	3	3	3
gelb	4	4	4	4
grün	5	5	5	5
blau	6	6	6	6
violett	7	7	7	7
grau	8	8	8	8
weiß	9	9	9	9

Tab. 3.11 Farbcodierung von Dioden) nach JEDEC-Norm

Stichwortverzeichnis:

- Backward-Diode 3-36
- Debye-Länge 3-6
- Diffusionskonstanten 3-2
- Diffusionslänge 3-10
- Diffusionsschwänze 3-7
 - Stromdichteanteile 3-10
- Diffusionsspannung 3-4
- Diffusionsstrom 3-2
- Diode 3-2
 - Backward-Diode 3-36
 - Kennzeichnung 3-40
 - Kurzbasisdiode 3-13
 - Langbasisdiode 3-13
 - pin-Diode 3-37
 - reale Diode 3-14
 - Schottky-Diode 3-39
 - Spezielle Dioden 3-24
 - Temperatureinfluss 3-16
 - Temperatureinfluss auf
 - Diffusionsspannung 3-18
 - Temperatureinfluss auf gesamte Kennlinie 3-19
 - Temperatureinfluss auf Sperrstrom 3-16
 - Tunnel-Diode 3-35
 - Varaktor-Diode (Kapazitätsdiode) 3-24
 - Zenerdiode 3-28
- Feldstrom 3-3
- Kontinuitätsgleichung 3-10
- Kurzbasisdiode 3-13
- Langbasisdiode 3-13
- Lawineneffekt 3-32
- Lebensdauer 3-10
- pin-Diode 3-37
- pn-Diode
 - Kleinsignalverhalten 3-22
- pn-Übergang
 - abrupter Übergang 3-25
 - Grundgleichungen 3-2
 - Hyperabrupter Übergang 3-25
 - linearer Übergang 3-25
 - stromdurchflossen 3-7
- Raumladung 3-2
- Rekombinationsrate 3-10
- Schottky-Dioden 3-39
- Temperaturspannung 3-2
- Tunnel-(Zener-)Effekt 3-29
- Tunnel-Diode 3-35
- Varaktor-Diode 3-24
- Zenerdiode 3-28

Kapitel 4

Bipolartransistoren

Inhalt:

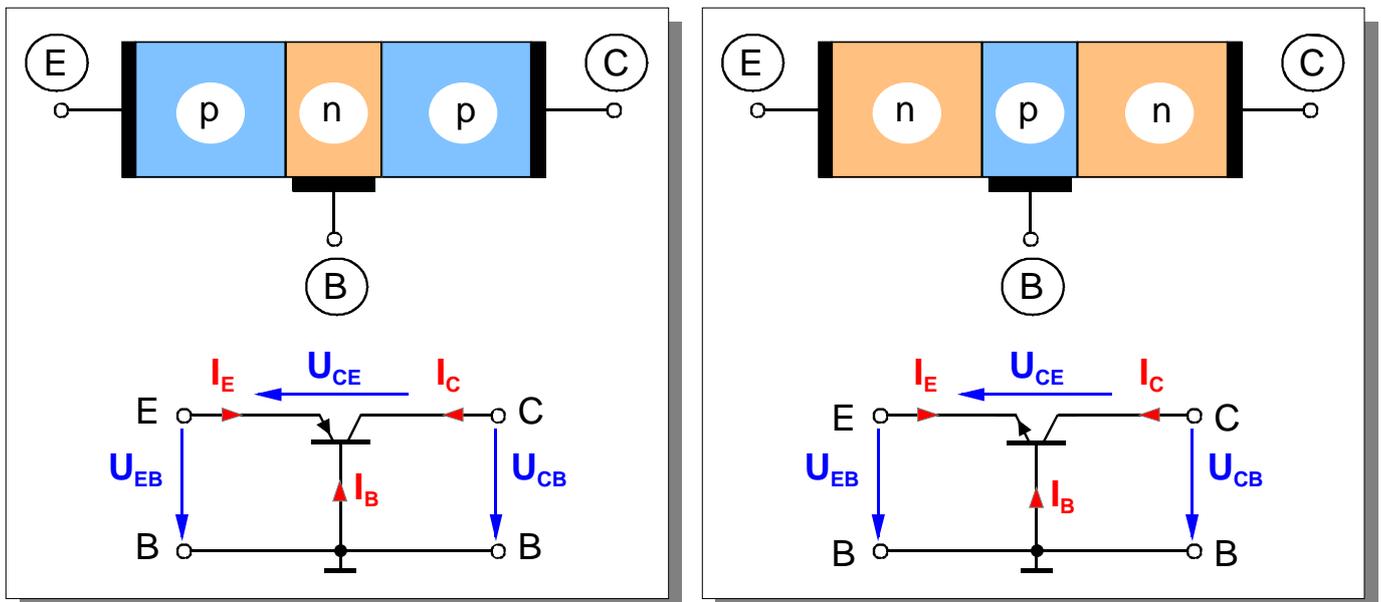
4	Bipolartransistoren	4-2
4.1	Funktionsweise	4-2
4.2	Aufbau und Technologie der Bipolartransistoren	4-4
4.3	Die Gleichstromeigenschaften der Bipolartransistoren	4-8
4.3.1	Das Ebers-Moll Modell	4-8
4.3.2	Die Kennlinienfelder des Bipolartransistors	4-15
4.3.2.1	Die Basis-Schaltung	4-15
4.3.2.1.1	Die Eingangskennlinie (EKL bzw. EKF)	4-15
4.3.2.1.2	Das Ausgangskennlinienfeld (AKF)	4-17
4.3.2.1.2.1	Das AKF mit U_{EB} als Steuerparameter ($I_C = f(U_{CB}, U_{EB})$)	4-18
4.3.2.1.2.2	Das AKF mit I_E als Steuerparameter ($I_C = f(U_{CB}, I_E)$)	4-20
4.3.2.2	Die Emitterschaltung	4-23
4.3.2.2.1	Die Eingangskennlinie $I_B = f(U_{BE})$	4-23
4.3.2.2.2	Das Ausgangskennlinienfeld (AKF)	4-24
4.3.2.2.2.1	Das AKF mit U_{BE} als Steuerparameter ($I_C = f(U_{CE}, U_{BE})$)	4-24
4.3.2.2.2.2	AKF mit I_B als Steuerparameter ($I_C = f(U_{CE}, I_B)$)	4-26
4.3.3	Das Transportmodell	4-32
4.3.4	Der Gummel-Plot	4-35
4.3.5	Der Early-Effekt	4-35
4.4	Das quasistationäre Kleinsignalverhalten des Bipolartransistors in Emitter-Schaltung	4-37
4.4.1	Der Eingangswiderstand	4-37
4.4.2	Die Stromverstärkung	4-39
4.4.3	Die Steilheit	4-40
4.4.4	Der Ausgangswiderstand	4-41
4.4.5	Das quasistationäre Kleinsignal-Ersatzschaltbild	4-42
4.5	Der Schaltvorgang im Bipolartransistor	4-42
4.5.1	Sperrbereich	4-43
4.5.2	Aktiver Bereich	4-44
4.5.3	Sättigungsbereich	4-44
4.6	Grenzwerte	4-47
4.6.1	Durchbruchverhalten	4-47
4.6.1.1	Durchbruch 1. Art	4-47
4.6.1.1.1	Basis-Emitter Sperrspannung	4-47
4.6.1.1.2	Kollektor-Basis Sperrspannung	4-47
4.6.1.1.3	Kollektor-Emitter Sperrspannung	4-47
4.6.1.2	Durchbruchverhalten 2. Art	4-47
4.6.2	Grenzströme	4-48
4.6.3	Sperrströme	4-48
4.6.4	Maximale Verlustleistung	4-48
4.7	Der Heterostrukturbipolartransistor (HBT)	4-52

4 Bipolartransistoren

4.1 Funktionsweise

Der Bipolartransistor (BJT: engl: *Bipolar Junction Transistor*) besteht aus der Reihenschaltung zweier nahe benachbarter pn-Übergänge, wobei die mittlere Dotierungszone, die **Basis B**, beiden Übergängen als gemeinsamer Bereich dient. Die beiden äußeren Bereiche werden mit **Emitter E** und **Kollektor C** (engl.: Collector) bezeichnet. Emitter und Kollektor weisen zwar den gleichen Typ von Majoritätsladungsträgern auf, sie sind jedoch unterschiedlich stark dotiert, so dass sich bei Vertauschen der Anschlüsse jeweils andere Verhältnisse ergeben. Je nach Dotierung lassen sich pnp- und npn-Transistoren herstellen (Abb. 4.1), wobei letzterer im Folgenden weiter behandelt werden soll.

Bei der Betrachtung der Wirkungsweise eines pnp-Transistors sind unter Beibehaltung der Strom- und Spannungsbezugsrichtungen gegenüber den Ergebnissen für den npn-Transistor negative Vorzeichen einzusetzen. Außerdem sind, durch die entgegengesetzte Dotierung der Bereiche bedingt, die Ladungsträgerarten, die Elektronen und Löcher, zu vertauschen.



a)

b)

Abb. 4.1 a) pnp- b) npn-Transistor mit den jeweiligen Schaltsymbolen und Spannungsbezugs Pfeilen

Funktionsweise des npn-Transistors:

Wird zwischen Kollektor und Basis eine Spannung $U_{CB} > 0$ angelegt, wird dieser pn-Übergang in Sperrrichtung betrieben. Wird keine zusätzliche Spannung zwischen Basis und Emitter angelegt, fließt nur ein geringer Sperrstrom, der in Abb. 4.2, welche die einzelnen Stromanteile im Transistor dargestellt, mit I_{CB0} bezeichnet ist. Da dieser Strom bei offenem Emitter zwischen Kollektor und Basis fließt, wird er **Kollektor-Basis Reststrom** genannt.

Legt man nun an den Emitter eine Spannung $U_{BE} = -U_{EB} > 0$ an, so wird der pn-Übergang zwischen diesen beiden Elektroden in Durchlassrichtung betrieben. Bereits kleine Spannungswerte von U_{BE} verursachen einen großen Elektronenstrom vom Emitter zur Basis; in Gegenrichtung

fließt aber nur ein geringer Löcherstrom I_{E+} von der Basis zum Emmitter, da die Dotierung der Basis geringer als die des Emitters ist.

Die aus dem Emmitter heraustretenden Elektronen können nur zu einem geringen Teil mit den Löchern der Basis rekombinieren, da die Basiszone sehr dünn und gegenüber dem Emmitter geringer dotiert ist.

Der größte Teil der Elektronen diffundiert durch die Basis hindurch und gelangt in das Sperrgebiet zwischen Basis und Kollektor, das an Elektronen verarmt ist. Durch die positive Kollektor-Basis-Spannung U_{CB} werden dann die Elektronen in den Kollektor gezogen. Es ist daher im Ausgangskreis ein Elektronenstrom feststellbar.

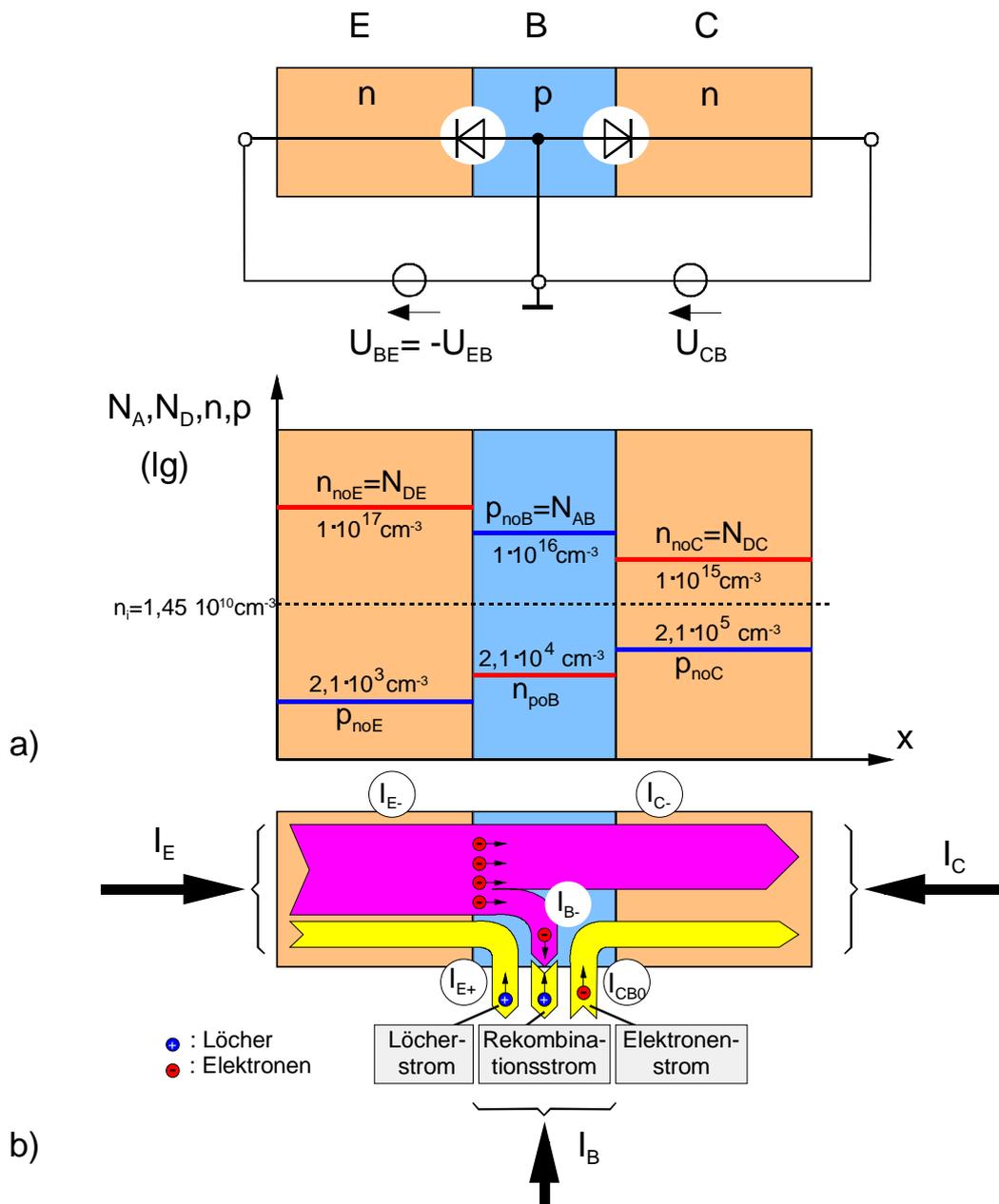


Abb. 4.2 Zur Funktionsweise des Transistors:

- a) typische Dotierungsverhältnisse im Silizium-Bipolartransistor
- b) Stromanteile im Bipolartransistor

Auch die Elektronen, die in der Basis mit Löchern rekombinieren, erzeugen einen Strom im Eingangskreis, der aber weitaus geringer ist.

Entscheidende Voraussetzung dafür, dass die vom Emitter ausgesandten Elektronen in ihrer Mehrzahl zum Kollektor und nicht zur Basiselektrode wandern, ist die Tatsache, dass die Basiszone sehr dünn (1-50 μm) und wesentlich geringer mit Akzeptoren als die Emitterschicht mit Donatoren dotiert ist.

Im Gegensatz zu den Feldeffekttransistoren als unipolare Bauelemente, deren Verhalten durch die Majoritätsladungsträger bestimmt wird, bestimmen in bipolaren Transistoren die Minoritätsladungsträger maßgeblich das Verhalten.

4.2 Aufbau und Technologie der Bipolartransistoren

Im Laufe der nunmehr über 40jährigen Geschichte des Bipolartransistors (er wurde 1948 von *Shockley, Bardeen* und *Brattain* in den Bell Laboratorien erfunden) wurden große Anstrengungen unternommen, um npn- bzw. pnp-Strukturen mit möglichst guten elektrischen Eigenschaften (hohe Verstärkung, gute Hochfrequenzeigenschaften) zu realisieren. Es entstanden so eine Vielzahl von Bauformen und Technologien [20].

Über fast 2 Jahrzehnte hatte der in Abb. 4.3 gezeigte Legierungstransistor die Palette der Transistorbauformen entscheidend mitbestimmt. Bei diesem Bauelement wird eine pnp-Struktur dadurch erzeugt, dass in ein n-leitendes Halbleiterplättchen (ca. 0,1 mm dick) von beiden Seiten Indiumpillen einlegiert werden. Da Indium ein dreiwertiges Element ist, wirkt es als Akzeptor, und nach der Rekristallisation beim Legierungsvorgang reichert sich Indium in den Rekristallisationszonen an, wodurch p-leitende Gebiete entstehen, wie es in Abb. 4.4 gezeigt ist..

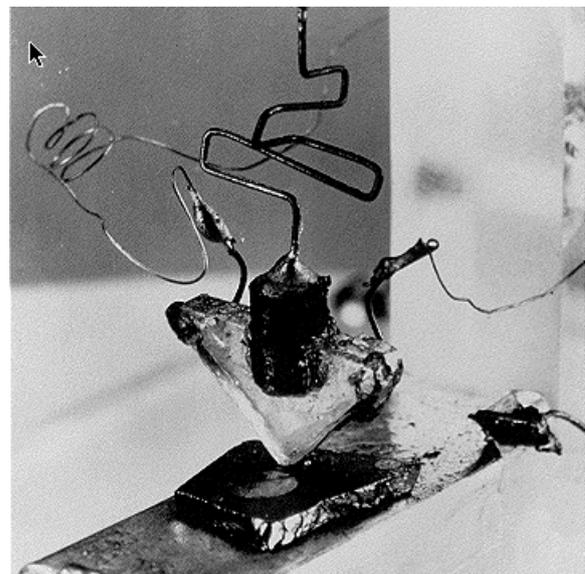
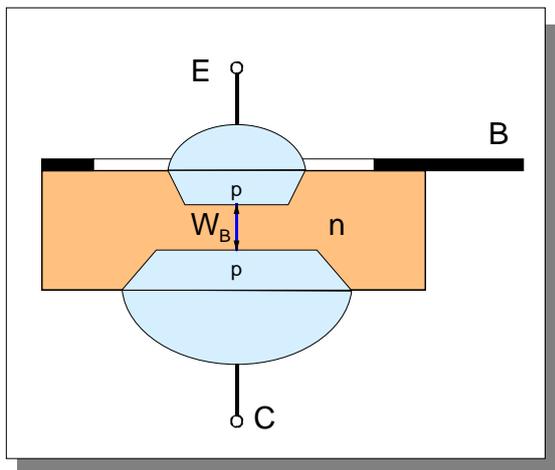


Abb. 4.3 a) Legierungstransistor (prinzipieller Querschnitt)
b) die erste technologische Realisierung des Bipolartransistors von Shockley, Bardeen und Brattain (1948)

Die kleinere Indiumpille dient als Emitteranschluss, die größere als Kollektoranschluss, und das

n-Halbleiterplättchen wird mit dem Basiskontakt versehen. Diese Bauform hat zwei entscheidende Nachteile: Zunächst ist diese Herstellungsweise für die Massenfertigung äußerst schlecht geeignet, da jeder Transistor durch Aufsetzen der Pillen einzeln behandelt werden muss. Weiterhin sind die erzielbaren elektrischen Eigenschaften nur mittelmäßig, da infolge des Legierungsprozesses die Basisweite W_B nicht wesentlich dünner als $10\ \mu\text{m}$ gemacht werden kann. Damit sind aber weder gute Hochfrequenzeigenschaften noch ein gutes Verstärkerverhalten zu erzielen.

Etwa 1960 wurde daher auf Siliziumbasis die Epitaxie-Planar-Technologie eingeführt und der Epitaxie-Planar-Transistor nach Abb. 4.4 realisiert. Hier wird zunächst auf ein hochdotiertes n^+ -Substrat (Dicke 0,1 bis 0,2 mm) eine dünne n-leitende einkristalline Siliziumschicht (Dicke etwa 5 bis $10\ \mu\text{m}$) durch Epitaxie aufgebracht. Das ist die eigentlich aktive Schicht, in der der Transistor realisiert wird. Danach erfolgt von der Oberfläche her die **Eindiffusion** 3-wertiger Atome als Akzeptoren (z.B. Bor), und es entsteht ein p-Gebiet (die Basis). Danach erfolgt in dieses p-Gebiet durch nochmalige Eindiffusion von 5-wertigen Atomen (Donatoren, z.B. Phosphor oder Arsen) mit noch höherer Konzentration die Erzeugung eines hochdotierten n^+ -Gebietes als Emitter.

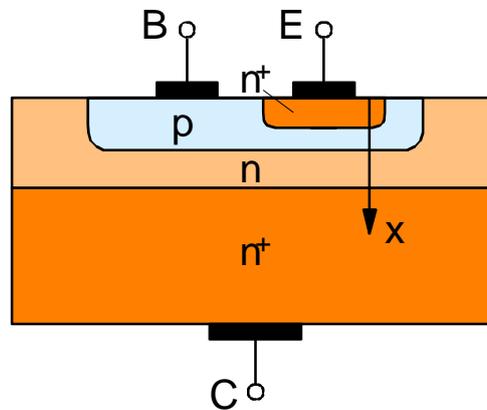


Abb. 4.4 Epitaxie-Planar-Transistor

In Abb. 4.5 ist das Störstellenprofil eines typischen Epitaxie-Planar-Transistors, wie er durch diese Doppeldiffusion erzeugt werden kann, gezeigt. Man erkennt, dass sich hier sehr dünne Basisgebiete (in Abb. 4.6 $W_B \approx 1,5\ \mu\text{m}$) herstellen lassen. Man sieht ebenfalls, dass der Emitter zur Erzielung eines hohen Stromverstärkungsfaktors sehr viel stärker dotiert ist als die Basis.

Eine neuere Version der Dotierung ist die **Ionenimplantation**. Hier lassen sich noch feinere Strukturen erzeugen. Der Technologieaufwand ist allerdings größer, so dass man dies nur dort anwendet, wo sich damit echte Fortschritte erzielen lassen. Die Epitaxie-Planar-Technik hat sich als Grundgedanke bis heute bei der Herstellung von Transistoren und integrierten Schaltkreisen erhalten. Das hat seine Ursache in der ausgezeichneten Eignung für die Massenfertigung (alle Transistorstrukturen können auf einer Halbleiterscheibe, Wafer genannt, gleichzeitig bearbeitet werden) und in der Erzielung sehr guter elektrischer Eigenschaften.

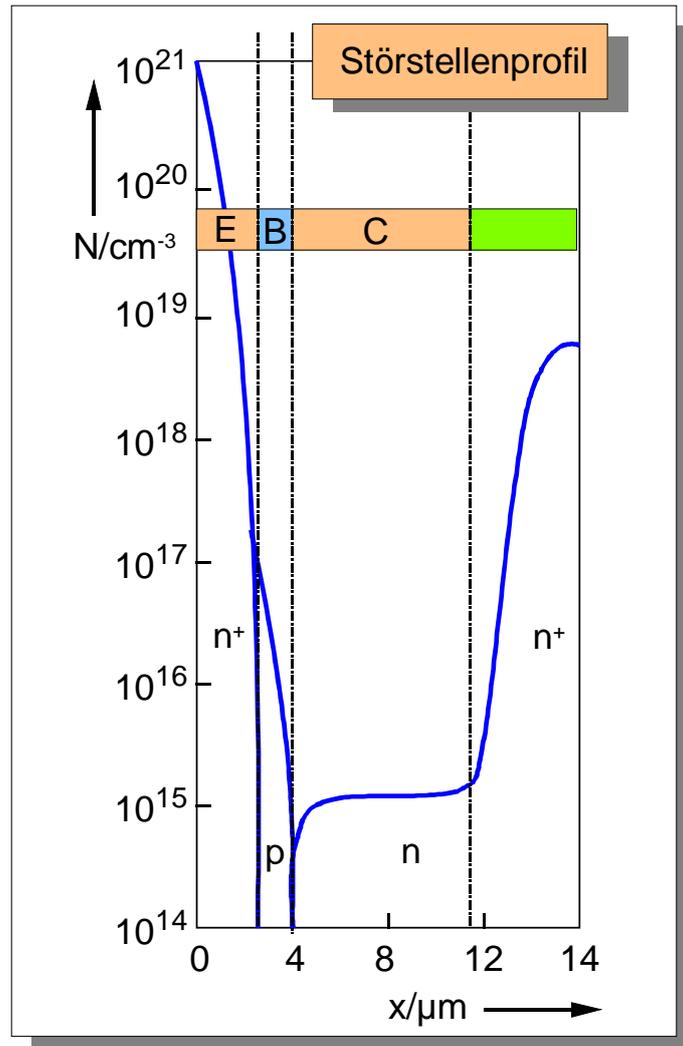


Abb. 4.5 Störstellenprofil eines Epitaxie-Planar-Transistor

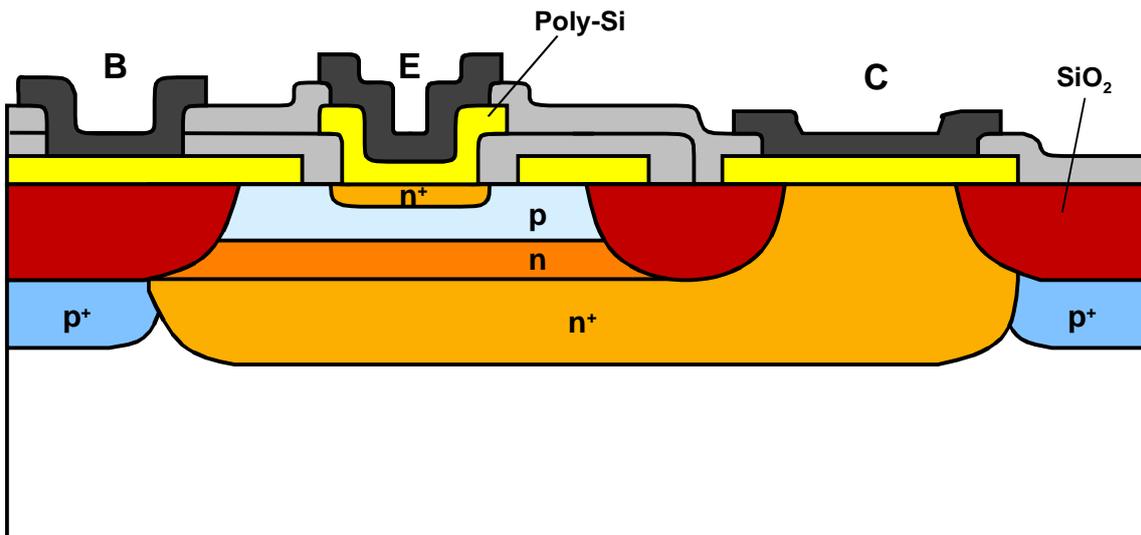


Abb. 4.6 ISOPLANAR-Transistor mit Polysiliziumemitter- und Basiskontakten (selbstjustierend)

Trotzdem hat es besonders in integrierten Schaltkreisen Weiterentwicklungen gegeben, die die Elementarstruktur in Abb. 4.4 nur noch in den Wesenszügen erkennen lassen. In integrierten Schaltkreisen ist besonders die seitliche Isolierung der Bauelemente gegeneinander zu realisieren. Das wird heute mit SiO_2 -Gräben (ISOPLANAR-Verfahren, TRENCH-Isolation o.a. ausgeführt. In Abb. 4.6 ist der Querschnitt einer solchen modernen Bipolartransistorstruktur für integrierte Schaltkreise gezeigt, Man erkennt zunächst die npn-Struktur, die den Bipolartransistor bildet, und die SiO_2 -Schichten zur seitlichen Isolation. Der Kollektoranschluss muss in integrierten Schaltkreisen auch von oben (von der Halbleiteroberfläche her) erfolgen. Dies geschieht, wie in Abb. 4.6 gezeigt, durch den Anschluss an ein sehr hochdotiertes n-Gebiet (n^+). Für Transistoren mit guten Hochfrequenz-Eigenschaften für Einsatz bei Frequenzen $f > 1 \text{ GHz}$ muss die Breite des Emitters W_E sehr klein (möglichst $W_E \leq 1 \mu\text{m}$) sein. Da der Emitter aber auch noch kontaktiert werden muss, wurden zur Lösung dieses schwierigen Strukturierungsprozesses selbstjustierende Strukturen mit Hilfe dotierter polykristalliner Siliziumschichten angewendet.

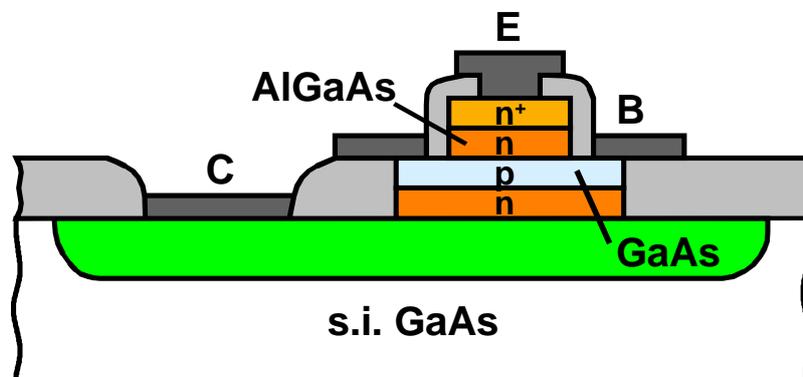


Abb. 4.7 AlGaAs-GaAs Heterostruktur-Bipolartransistor (HBT)

Eine besondere Möglichkeit, den Stromverstärkungsfaktor groß zu machen, ohne den Emitter sehr viel stärker als die Basis dotieren zu müssen, bieten die Heterostruktur-Bipolartransistoren. In Abb. 4.7 ist als Beispiel ein AlGaAs-GaAs-Heterostruktur-Bipolartransistor (AlGaAs/GaAs-HBT) prinzipiell dargestellt. Durch die verschiedenen Bandabstände von AlGaAs und GaAs haben diese auch unterschiedliche Eigenleitungsdichten, so dass damit die Injektionsverhältnisse I_n/I_p gesteuert werden können. Darauf wird noch zu einem späteren Zeitpunkt eingegangen.

4.3 Die Gleichstromeigenschaften der Bipolartransistoren

4.3.1 Das Ebers-Moll Modell

Zur Berechnung der DC-Eigenschaften der Bipolartransistoren wird entsprechend der Berechnung, die bereits von einzelnen pn-Übergängen her bekannt ist (Kap. 3.1), verfahren. Abb. 4.8 zeigt den prinzipiellen Bandverlauf sowie den Konzentrationsverlauf der Ladungsträger für den Fall, dass die Basis-Emitter-Diode in Flussrichtung und die Basis-Kollektor-Diode in Sperrichtung vorgespannt ist. Wie bereits im Fall der einfachen pn-Diode ist auch hier das Verhalten der Minoritäten entscheidend.

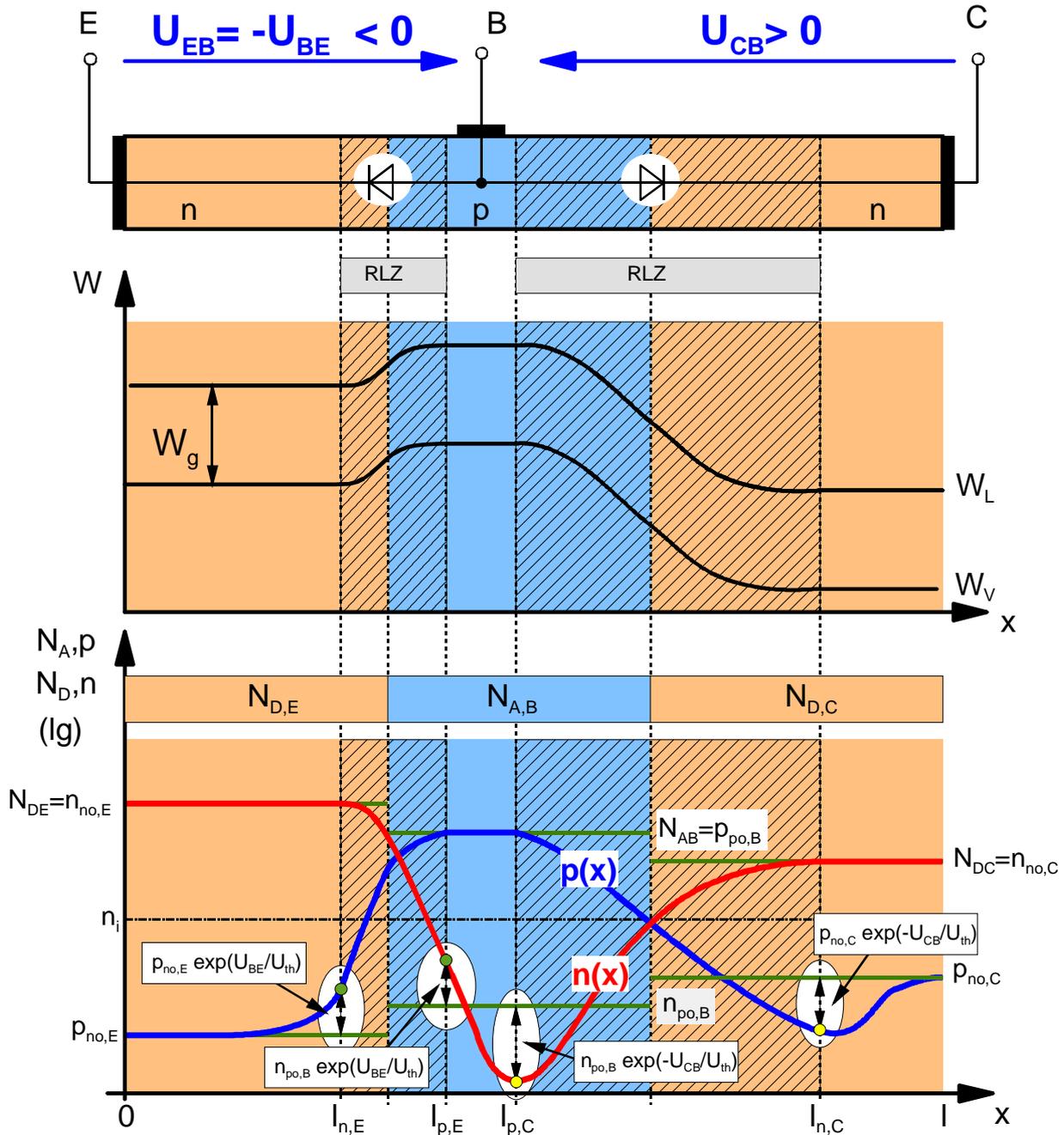


Abb. 4.8 Querschnitt und Ladungsverteilung in einem vorgespannten npn-Transistor mit $U_{EB} = -U_{BE} < 0$ und $U_{CB} > 0$

Es gilt daher für die Elektronenkonzentration in der p-dotierten Basis am emitterseitigen Rand:

$$n_p(x=l_{pE}) = n_{po,B} e^{-U_{EB}/U_{th}} \quad (4.1)$$

Am kollektorseitigen Rand der p-dotierten Basis gilt:

$$n_p(x=l_{pC}) = n_{po,B} e^{-U_{CB}/U_{th}} \quad (4.2)$$

Als weitere Randbedingungen sind zu berücksichtigen:

$$p_{nE}(0) = p_{no,E} = \frac{n_i^2}{N_{DE}} \quad (4.3)$$

und

$$p_{nC}(x=l) = p_{no,C} = \frac{n_i^2}{N_{DC}} \quad (4.4)$$

In Abb. 4.9 ist die Ladungsträgerkonzentration in der Basis und an den Rändern der Raumladungszonen im Emitter und Kollektor prinzipiell dargestellt.

Aufgrund der in Durchlassrichtung gepolten Basis-Emitter-Diode und dementsprechend hohen Konzentration der Elektronen an der Emitterseite sowie der in Sperrrichtung gepolten Basis-Kollektor-Diode mit entsprechend niedriger Konzentration auf der Kollektorseite bildet sich ein Konzentrationsgradient dn/dx aus. Bedingt durch Diffusion führt dieser Konzentrationsgradient letztendlich zu dem Kollektorstrom I_C . Es gilt:

$$\frac{dn}{dx} \propto I_C$$

Für diesen Diffusionsvorgang ist allerdings notwendig, dass die Basisweite W_B klein gegenüber der Diffusionslänge L_n der Minoritäten (hier: Elektronen) in der Basis ist. Abb. 4.10 zeigt den Konzentrationsverlauf für verschiedene Werte von W_B/L_n .

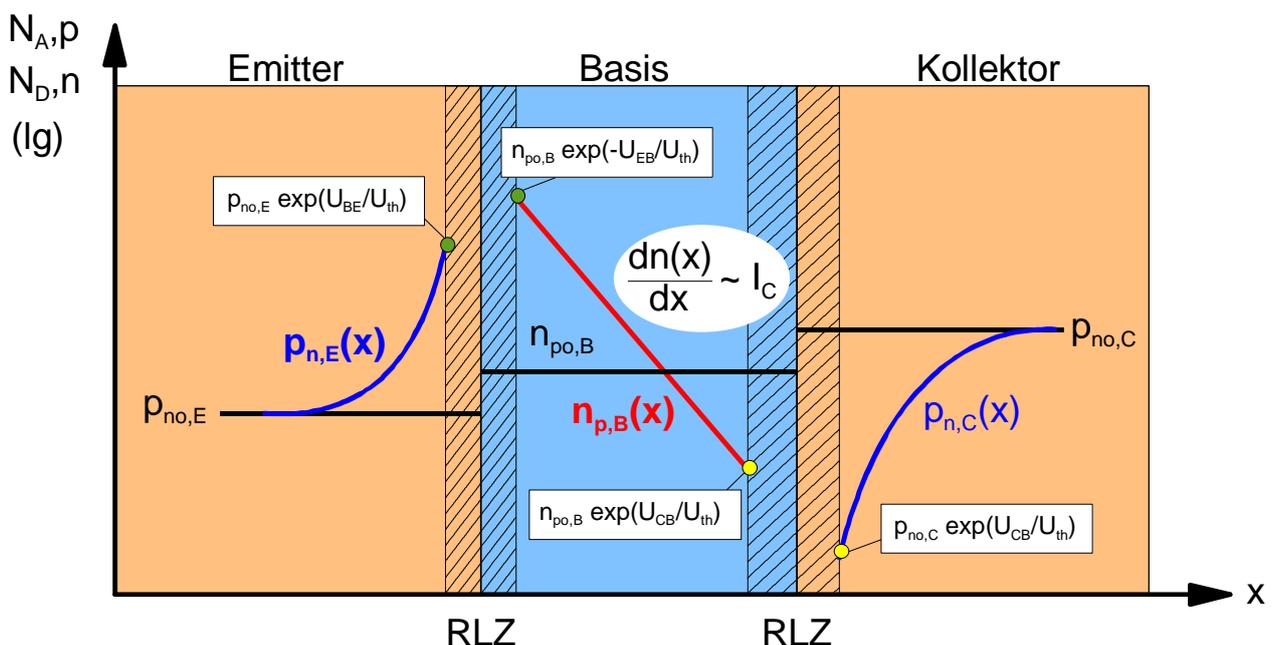


Abb. 4.9 Prinzipieller Verlauf der Minoritäten in Basis-Nähe

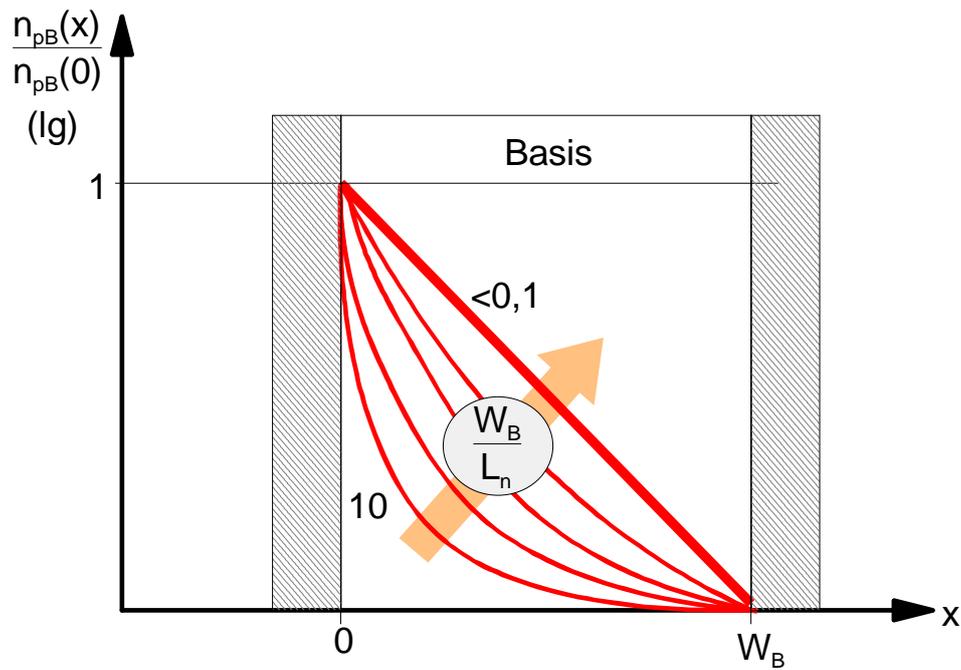


Abb. 4.10 Prinzipieller Verlauf der Minoritäten in der Basis bei verschiedenen Verhältnissen der Basisweite W_B zu Diffusionslänge L_n

Mit

$$I_{np,B}(l_{pE}) = J_{np,B}(l_{pE}) \cdot A_E \tag{4.5}$$

und

$$I_{pn,E}(l_{nE}) = J_{pn,E}(l_{pE}) \cdot A_E \tag{4.6}$$

und unter Annahme derselben Voraussetzungen wie zur Berechnung des stromdurchflossenen pn-Übergangs (Kap. 3.1.3, Abb. 3.4) mit

$$J_{pn}(x) \approx \text{konst. in der Raumladungszone} \tag{4.7}$$

$$J_{np}(x) \approx \text{konst. in der Raumladungszone}$$

erhält man den Gesamtstrom:

$$I_E = I_{np,B}(x = l_{pE}) + I_{pn,E}(x = l_{nE}) \tag{4.8}$$

1. Emittorstrom:

Unter der Voraussetzung, dass für die Basisweite W_B gilt:

$$W_B = l_{pC} - l_{pE} \ll L_{nB} \quad , \quad L_{nB} \approx 10 \mu\text{m}$$

ergibt sich für den Emittorstrom:

$$I_E = -I_{EE} \left(e^{-U_{EB}/U_{th}} - 1 \right) + I_{EC} \left(e^{-U_{CB}/U_{th}} - 1 \right) \quad (4.9)$$

mit

$$I_{EE} = qU_{th}A_E \left(\frac{\mu_n n_{p0B}}{W_B} + \frac{\mu_p p_{n0E}}{l_{nE}} \right) \quad (4.10)$$

bzw.

$$I_{EE} = qU_{th}A_E n_i^2 \left(\frac{\mu_n}{W_B N_{AB}} + \frac{\mu_p}{l_{nE} N_{DE}} \right) \quad (4.11)$$

und

$$I_{EC} = qU_{th}A_E \frac{\mu_n n_{p0B}}{W_B} = qU_{th}A_E \frac{\mu_n n_i^2}{W_B N_{AB}} \quad (4.12)$$

2. Kollektorstrom:

Analog gilt für den Kollektorstrom unter Voraussetzung des Bezugspfeilsystems in Abb. 4.1

$$I_C = I_{np,B}(l_{pC}) + I_{pn,C}(l_{nC})$$

$$I_C = +I_{CE} \left(e^{-U_{EB}/U_{th}} - 1 \right) - I_{CC} \left(e^{-U_{CB}/U_{th}} - 1 \right) \quad (4.13)$$

mit

$$I_{CE} = qU_{th}A_C \frac{\mu_n n_{p0B}}{W_B} = qU_{th}A_C \frac{\mu_n n_i^2}{W_B N_{AB}} \quad (4.14)$$

und

$$I_{CC} = qU_{th}A_C n_i^2 \left(\frac{\mu_n}{W_B N_{AB}} + \frac{\mu_p}{l_{nC} N_{DC}} \right) \quad (4.15)$$

Für den Fall, dass

$$A_E = A_C$$

gilt, ergibt sich

$$I_{EC} = I_{CE}$$

Dabei wurde vorausgesetzt, dass:

$$l - l_{nC} \gg L_{pC}$$

ist, d.h. die Berechnung wurde für den Fall einer Langbasisdiode durchgeführt. Die Gleichungen Gl. (4.9) und Gl. (4.13) werden als **Ebers-Moll Gleichungen** bezeichnet.

Zusammenfassung:

Nach Gl. (4.9), (4.13) sowie Gln. (4.11) - (4.14) gilt:

$$I_E = -I_{EE} \left(e^{-U_{EB}/U_{th}} - 1 \right) + I_{EC} \left(e^{-U_{CB}/U_{th}} - 1 \right) \quad (4.9)$$

$$I_C = +I_{CE} \left(e^{-U_{EB}/U_{th}} - 1 \right) - I_{CC} \left(e^{-U_{CB}/U_{th}} - 1 \right) \quad (4.13)$$

mit:

$$I_{EE} = qU_{th}A_E n_i^2 \left(\frac{\mu_n}{W_B N_{AB}} + \frac{\mu_p}{l_{nE} N_{DE}} \right) \quad (4.11)$$

$$I_{EC} = qU_{th}A_E \frac{\mu_n n_{p0B}}{W_B} = qU_{th}A_E \frac{\mu_n n_i^2}{W_B N_{AB}} \quad (4.12)$$

$$I_{CE} = qU_{th}A_C \frac{\mu_n n_{p0B}}{W_B} = qU_{th}A_C \frac{\mu_n n_i^2}{W_B N_{AB}} \quad (4.14)$$

$$I_{CC} = qU_{th}A_C n_i^2 \left(\frac{\mu_n}{W_B N_{AB}} + \frac{\mu_p}{l_{nC} N_{DC}} \right) \quad (4.15)$$

Damit ergibt sich das Ebers-Moll Großsignal-Modell in Abb. 4.11.

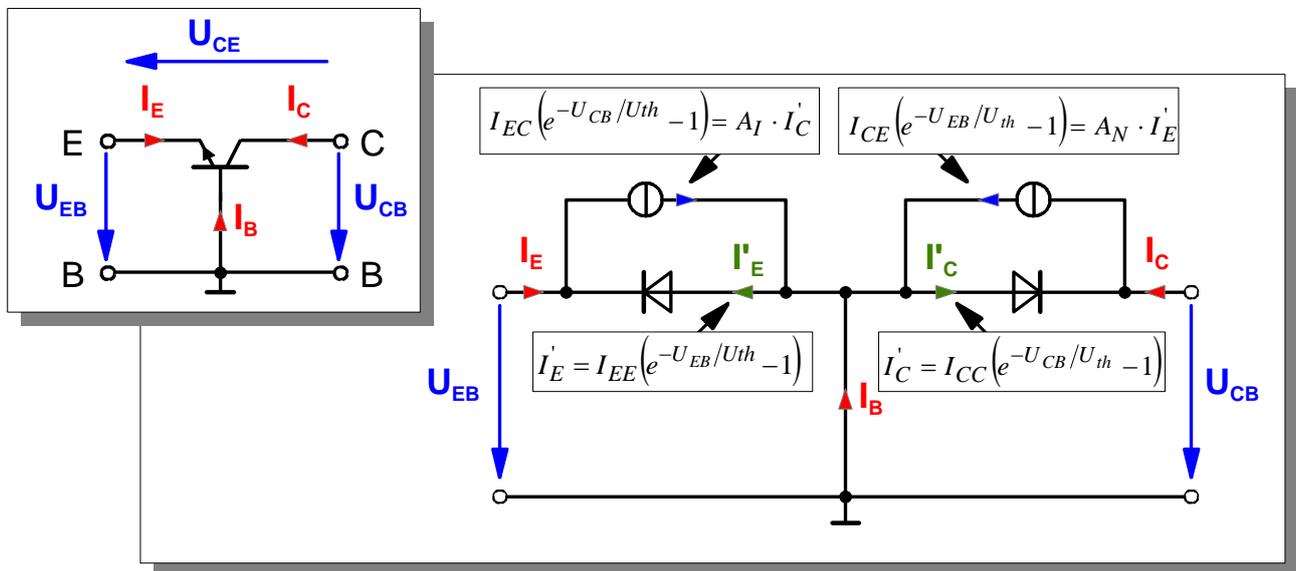


Abb. 4.11 Großsignal-Ersatzschaltbild des Bipolartransistors nach Ebers-Moll

Analyse des Modells:

Die Ströme I_E und I_C setzen sich jeweils aus zwei Anteilen zusammen:

Zum einen fließt der Strom aufgrund des pn-Übergangs I'_E bzw. I'_C entsprechend der über der Diode abfallenden Spannung, wobei der Strom durch diese Dioden dem bekannten exponentiellen Zusammenhang (Gl. (3.46)) folgt.

Parallel zu diesen Dioden liegt jeweils eine Stromquelle, die durch den Strom der jeweils anderen Diode gesteuert wird. Der Zusammenhang wird durch die Proportionalitätsfaktoren A_N und A_I beschrieben.

Damit gilt:

$$I_E = -I_E' + A_I \cdot I_C' \quad (4.16)$$

$$I_C = +A_N \cdot I_E' - I_C' \quad (4.17)$$

mit

$$I_E' = I_{EE} \left(e^{-U_{EB}/U_{th}} - 1 \right) \quad (4.18)$$

$$I_C' = I_{CC} \left(e^{-U_{CB}/U_{th}} - 1 \right) \quad (4.19)$$

bzw.

$$I_E = -I_{EE} \left(e^{-U_{EB}/U_{th}} - 1 \right) + A_I \cdot I_{CC} \left(e^{-U_{CB}/U_{th}} - 1 \right) \quad (4.20)$$

$$I_C = A_N \cdot I_{EE} \left(e^{-U_{EB}/U_{th}} - 1 \right) - I_{CC} \left(e^{-U_{CB}/U_{th}} - 1 \right) \quad (4.21)$$

Dabei bedeuten:

A_N : Stromverstärkung der Basis-Schaltung in Vorwärtsrichtung (N: Normal)

A_I : Stromverstärkung der Basis-Schaltung in Rückwärtsrichtung (I: Invers)

Aus einem Koeffizientenvergleich folgt:

$$I_{EC} = A_I \cdot I_{CC} \quad \Rightarrow \quad I_{CC} = \frac{I_{EC}}{A_I} \quad \text{bzw.} \quad A_I = \frac{I_{EC}}{I_{CC}} \quad (4.22)$$

$$I_{CE} = A_N \cdot I_{EE} \quad \Rightarrow \quad I_{EE} = \frac{I_{CE}}{A_N} \quad \text{bzw.} \quad A_N = \frac{I_{CE}}{I_{EE}} \quad (4.23)$$

Aufgrund der Kontenregel mit:

$$I_B + I_C + I_E = 0$$

ergibt sich der Basisstrom zu:

$$I_B = -I_C - I_E$$

$$I_B = +I_E' - A_I \cdot I_C' - A_N \cdot I_E' + I_C'$$

$$I_B = (1 - A_N) \cdot I_E' + (1 - A_I) \cdot I_C' \quad (4.24)$$

bzw.

$$I_B = (1 - A_N) \cdot I_{EE} \cdot \left(e^{-U_{EB}/U_{th}} - 1 \right) + (1 - A_I) \cdot I_{CC} \cdot \left(e^{-U_{CB}/U_{th}} - 1 \right) \quad (4.25)$$

Der Bipolartransistor wird sehr häufig unter den Bedingungen:

- die Basis-Emitter-Diode ist leitend ($-U_{EB} / U_{th} \gg 0$)
- die Basis-Kollektor-Diode ist gesperrt ($U_{CB} / U_{th} \gg 0$)

betrieben. Daher wird der Betrieb in diesem Fall auch als "**Normalbetrieb**" des Bipolartransistors bezeichnet.

In diesem Fall gilt:

$$I_C' = I_{CC} \left(e^{-U_{CB}/U_{th}} - 1 \right) = -I_{CC}$$

Wird I_{CC} in erster Näherung vernachlässigt, so folgt $I_C' = 0$. Und damit

$$I_E = -I_E' + A_I \cdot I_C' = -I_E'$$

Mit

$$I_E' = I_{EE} \left(e^{-U_{EB}/U_{th}} - 1 \right)$$

folgt:

$$I_E = -I_{EE} \left(e^{-U_{EB}/U_{th}} - 1 \right)$$

und

$$I_C = A_N \cdot I_E'$$

Dementsprechend ergibt sich das reduzierte Ersatzschaltbild in Abb. 4.12 mit einer idealen Diode zwischen Basis und Emitter sowie einer gesteuerte Stromquelle im Kollektorzweig.

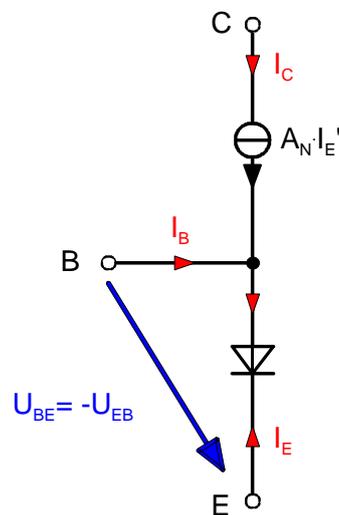


Abb. 4.12 Reduziertes Ebers-Moll Modell für den npn-Transistor im Normalbetrieb

4.3.2 Die Kennlinienfelder des Bipolartransistors

Entsprechend dem MOSFET sind auch im Fall des Bipolartransistors verschiedene Grundschaltungen möglich:

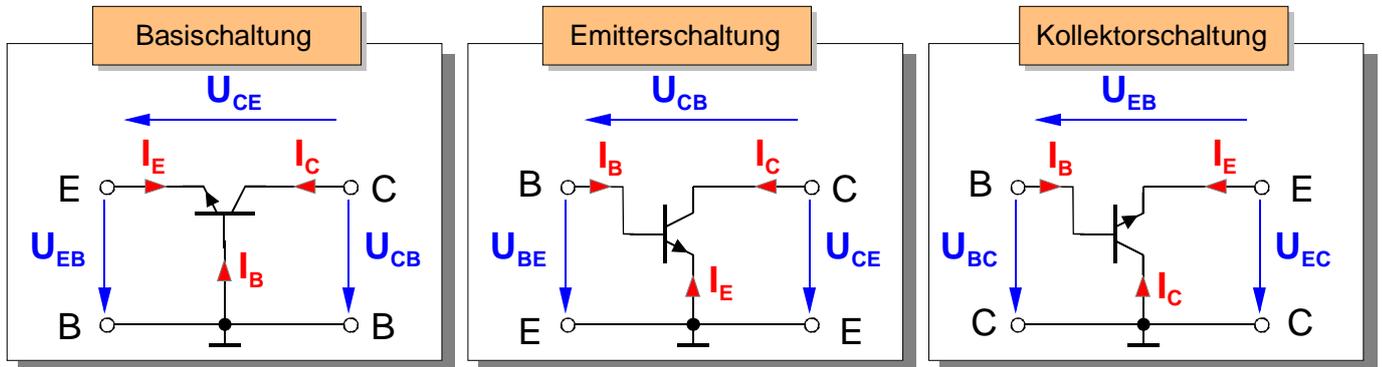


Abb. 4.13 Grundschaltungen des Bipolartransistors

4.3.2.1 Die Basis-Schaltung

Betrachtet man den Emittor als Eingang, den Kollektor als Ausgang und den Basisanschluss als den gemeinsamen Kontakt für Ein- und Ausgang, so liegt die **Basischaltung** (Abb. 4.13a) vor.

Zur Beschreibung des Transistorverhaltens unterscheidet man verschiedene Kennlinien bzw. Kennlinienfelder:

- **Eingangskennlinie** (EKL bzw. EKF)

$$I_E = f(U_{EB}, U_{CB})$$
- **Ausgangskennlinienfeld** (AKF)

$$I_C = f(U_{CB}, U_{EB})$$
- **Übertragungskennlinie** (ÜKL bzw. ÜKF)

$$I_C = f(U_{EB}, U_{CB}) \text{ bzw. } I_C = f(I_E, U_{CB})$$

4.3.2.1.1 Die Eingangskennlinie (EKL bzw. EKF)

Die Eingangskennlinie ist gegeben durch

$$I_E = f(U_{EB})$$

Es gelten die bekannten Ebers-Moll Gleichungen mit:

$$I_E = -I_{EE} \left(e^{-U_{EB}/U_{th}} - 1 \right) + I_{EC} \left(e^{-U_{CB}/U_{th}} - 1 \right) \quad (4.9)$$

$$I_C = +I_{CE} \left(e^{-U_{EB}/U_{th}} - 1 \right) - I_{CC} \left(e^{-U_{CB}/U_{th}} - 1 \right) \quad (4.13)$$

bzw.

$$I_E = -I_E' + A_I \cdot I_C' \quad (4.16)$$

$$I_C = +A_N \cdot I_E' - I_C' \quad (4.17)$$

Für die Diskussion werden 3 Fälle betrachtet:

- a) $U_{CB} = 0$
- b) U_{CB} offen
- c) $U_{CB} / U_{th} \gg 0$

zu a) $U_{CB} = 0$

In diesem Fall vereinfacht sich Gl. (4.9) zu

$$I_E = -I_{EE} \left(e^{-U_{EB}/U_{th}} - 1 \right) \quad (4.26)$$

Es ergibt sich eine Diodenkennlinie.

zu b) Leerlauf auf der Kollektorseite, d.h. es gilt $I_C = 0$ und somit folgt

$$I_C = 0$$

Mit Gl. (4.17) folgt:

$$I_C' = A_N \cdot I_E'$$

Eingesetzt in Gl. (4.16) ergibt sich:

$$\begin{aligned} I_E &= -I_E' + A_I \cdot I_C' = -I_E' + A_I \cdot A_N \cdot I_E' \\ \Rightarrow I_E &= I_E' (A_I \cdot A_N - 1) \end{aligned}$$

Mit Gl.(4.22) und (4.23)

$$A_I = \frac{I_{EC}}{I_{CC}} \quad \text{und} \quad A_N = \frac{I_{CE}}{I_{EE}}$$

gilt:

$$I_E = \left(\frac{I_{EC} \cdot I_{CE}}{I_{CC} \cdot I_{EE}} - 1 \right) \cdot I_E'$$

bzw. mit. Gl. (4.18)

$$I_E = \left(\frac{I_{EC} I_{CE}}{I_{CC}} - I_{EE} \right) \cdot \left(e^{-U_{EB}/U_{th}} - 1 \right) \quad (4.27)$$

Es ergibt sich wiederum eine Diodenkennlinie. Dabei ist der Betrag des Stromes $|I_E|$ bei gleicher Spannung U_{EB} im Fall b) kleiner als im Fall a). Dies führt zu einer Verschiebung der Kennlinie hin zu größeren Werten von $|U_{EB}|$ (s. Abb. 4.14).

zu c) $U_{CB} / U_{th} \gg 0$, d.h. die BC-Diode ist gesperrt:

Damit folgt

$$I_E = -I_{EE} \left(e^{-U_{EB}/U_{th}} - 1 \right) - I_{EC} \quad (4.28)$$

Auch in diesem Fall ergibt sich wiederum eine Diodenkennlinie. Der Strom I_E ist dabei nahezu so groß wie unter a) berechnet (s. Gl. (4.19)).

Da der Bipolartransistor sehr häufig unter diesen Bedingungen:

- die Basis-Emitter-Diode ist leitend ($-U_{EB} / U_{th} \gg 0$)
- die Basis-Kollektor-Diode ist gesperrt ($U_{CB} / U_{th} \gg 0$)

betrieben wird, wird der Betrieb in diesem Fall auch als "**Normalbetrieb**" des Bipolartransistors bezeichnet.

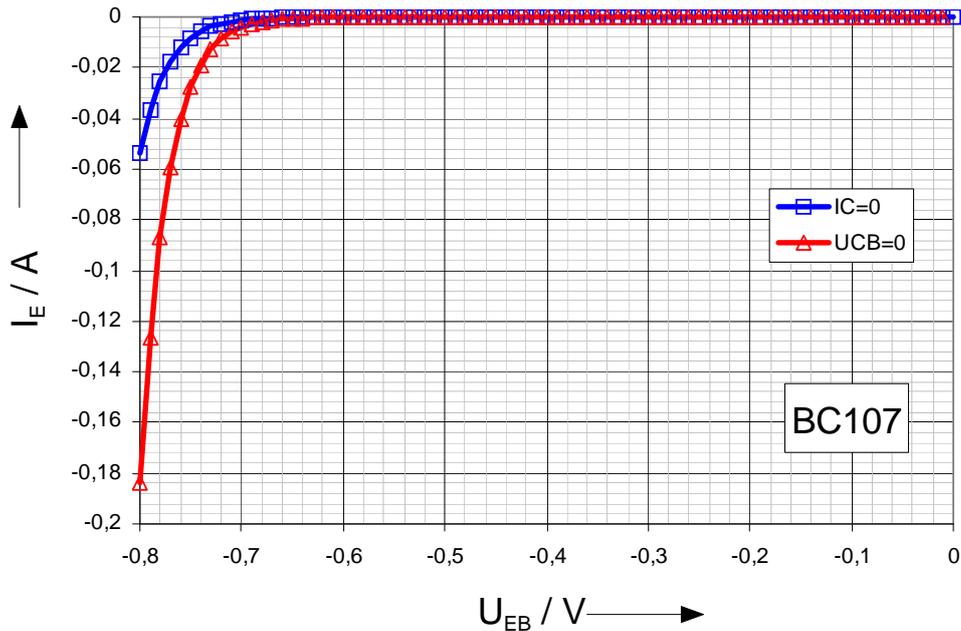


Abb. 4.14 Eingangskennlinie für $U_{CB} = 0$ und Leerlauf auf Kollektorseite ($I_C = 0$)

4.3.2.1.2 Das Ausgangskennlinienfeld (AKF)

Da im Gegensatz zum Feldeffekttransistor das Verhalten des Bipolartransistors eingangseitig durch zwei Parameter, U_{EB} und I_E , gesteuert werden kann, lassen sich auch zwei Ausgangskennlinienfelder angeben:

- $I_C = f(U_{CB}, U_{EB})$ mit U_{EB} als Parameter
- $I_C = f(U_{CB}, I_E)$ mit I_E als Parameter

Anmerkung: Die im Folgenden dargestellten Kennlinien sind auf Basis des Kleinsignaltransistors BC107 mit Hilfe des Simulationsprogramms SPICE ermittelt worden.

4.3.2.1.2.1 Das AKF mit U_{EB} als Steuerparameter ($I_C = f(U_{CB}, U_{EB})$)

Auch hier werden für die Diskussion wieder drei verschiedene Fälle betrachtet:

- a) $U_{CB} > 0$ bzw. $U_{CB}/U_{th} \gg 1$
- b) $U_{CB} = 0$
- c) $U_{CB} < 0$

zu a) $U_{CB} > 0$ bzw. $U_{CB}/U_{th} \gg 1$ (Normalbetrieb)

Nach Gl. (4.13) gilt:

$$I_C = +I_{CE} \left(e^{-U_{EB}/U_{th}} - 1 \right) - I_{CC} \left(e^{-U_{CB}/U_{th}} - 1 \right)$$

Für $U_{CB} > 0$ bzw. $U_{CB}/U_{th} \gg 1$ ist die BC-Diode gesperrt. Damit folgt für den ausgangsseitigen Strom I_C :

$$I_C = I_{CE} \left(e^{-U_{EB}/U_{th}} - 1 \right) + I_{CC} \quad (4.29)$$

Somit ist der Ausgangsstrom nicht mehr von der Spannung U_{CB} abhängig.

$$I_C \neq f(U_{CB}).$$

Für kleine Spannungen U_{EB} unterhalb der Flussspannung ist der Strom $I_C = -I_{CE} + I_{CC} \approx 0$. Für größere Spannungen steigt I_C exponentiell mit $-U_{EB}$ an.

zu b) Bei $U_{CB} = 0$ folgt:

$$I_C = I_{CE} \left(e^{-U_{EB}/U_{th}} - 1 \right), \quad (4.30)$$

und entspricht damit nahezu dem Strom aus Gl. (4.29) bei Normalbetrieb des Transistors.

zu c) Wird die BC-Diode mit $U_{CB} < 0$ in Durchlassrichtung vorgespannt, muss die vollständige Ebers-Moll Gleichung herangezogen werden und es ergibt sich:

$$I_C = I_{CE} \left(e^{-U_{EB}/U_{th}} - 1 \right) - I_{CC} \left(e^{-U_{CB}/U_{th}} - 1 \right), \quad (4.31)$$

Der zweite Term in Gl. (4.29) ($-I_{CC} \left(e^{-U_{CB}/U_{th}} - 1 \right)$) kompensiert zunehmend den 1. Term und führt zu einer Abnahme des Stroms I_C im 2. Quadranten des Kennlinienfeldes (Abb. 4.16, 4.17).

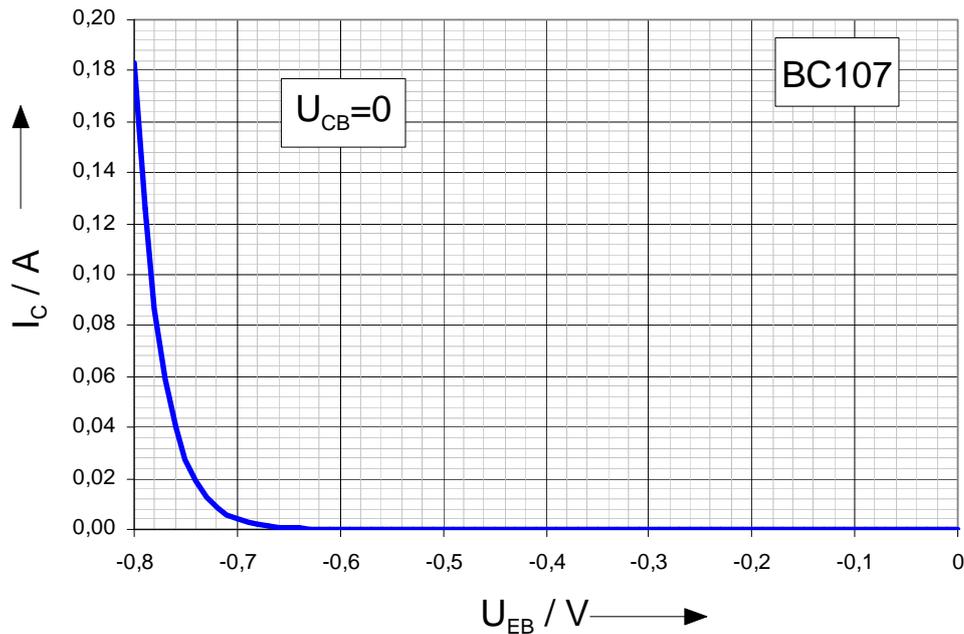


Abb. 4.15 Kollektorstrom des BC 107 in Abhängigkeit von der Spannung U_{EB} bei $U_{CB} = 0$

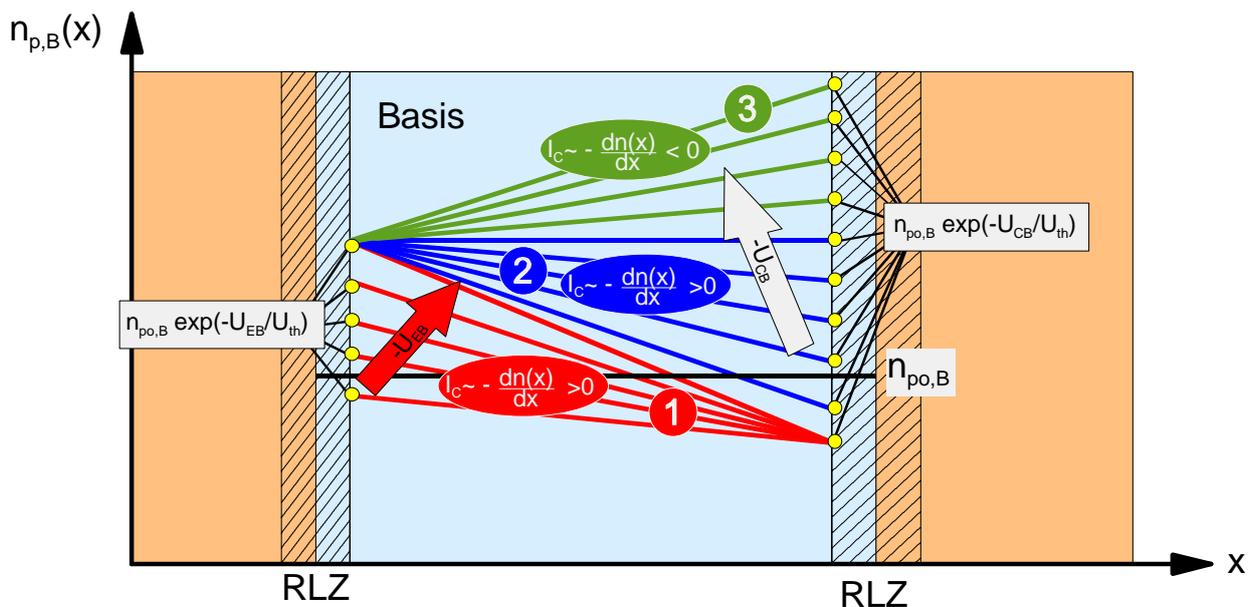


Abb. 4.16 Prinzipieller Verlauf der Minoritätenkonzentration in der Basis für verschiedene Arbeitspunkte

In Abb. 4.16 ist das prinzipielle Verhalten schematisch dargestellt:

1. Fall: Die BE-Diode ist leitend, die BC-Diode gesperrt. Bei konstanter Spannung $-U_{CB}$ wird die Spannung $-U_{EB}$ vergrößert, der Konzentrationsgradient dn/dx steigt an und somit nimmt der Kollektorstrom I_C zu (1. Quadrant in Abb. 4.17).
2. Fall: Bei konstanter Spannung $-U_{EB}$ wird die Spannung $-U_{CB}$ erhöht. Die BC-Diode wird leitend, der Konzentrationsgradient nimmt ab und der Strom I_C sinkt. Dies ist gleichbedeutend mit dem oben beschriebenen Fall, dass der zweite Term in Gl. (4.106) den ersten Term kompensiert, bis der Strom auf 0 abgenommen hat (2. Quadrant in Abb. 4.17).
3. Fall: Wird dann $-U_{CB}$ noch weiter vergrößert, so wird der Konzentrationsgradient sogar negativ und damit der Strom I_C ebenfalls (3. Quadrant in Abb. 4.17).

Der Abstand der Kennlinienäste nimmt bei linearem Anstieg des Steuerparameters $-U_{EB}$ entsprechend Gl. (4.29)-(4.31) exponentiell zu.

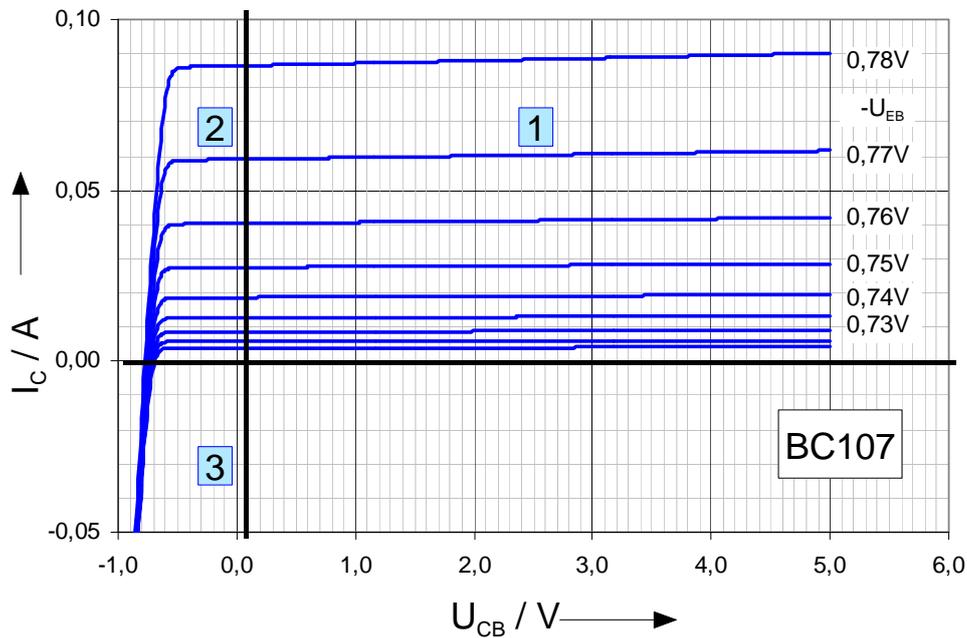


Abb. 4.17 Ausgangskennlinienfeld $I_C = f(U_{CB}, U_{EB})$ des BC 107

4.3.2.1.2 Das AKF mit I_E als Steuerparameter ($I_C = f(U_{CB}, I_E)$)

In den Ebers-Moll Gleichungen ist der steuernde Parameter jeweils die Spannung U_{EB} . Um den Strom als Steuerparameter zu erhalten, muss die Spannung U_{EB} eliminiert werden.

Gl.(4.16):

$$I_E = -I_E' + A_I \cdot I_C'$$

$$\Rightarrow I_E' = -I_E + A_I \cdot I_C'$$

eingesetzt in Gl. (4.17)

$$I_C = A_N \cdot I_E' - I_C'$$

folgt:

$$I_C = A_N \cdot (-I_E + A_I \cdot I_C') - I_C'$$

$$\Rightarrow I_C = -I_E \cdot A_N + I_C' \cdot (A_N \cdot A_I - 1)$$

und damit

$$I_C = -I_E \cdot \frac{I_{CE}}{I_{EE}} + I_C' \cdot \left(\frac{I_{CE}}{I_{EE}} \cdot \frac{I_{EC}}{I_{CC}} - 1 \right)$$

bzw.

$$I_C = -\frac{I_{CE}}{I_{EE}} \cdot I_E + \left(\frac{I_{CE} I_{EC}}{I_{EE}} - I_{CC} \right) \cdot (e^{-U_{CB}/U_{th}} - 1)$$

Der Strom setzt sich damit aus einem von I_E und einem von U_{CB} abhängigen Anteil zusammen.

$$I_C = A \cdot I_E + I_{CB0} \cdot (e^{-U_{CB}/U_{th}} - 1) \quad (4.33)$$

mit

$$A = -\frac{I_{CE}}{I_{EE}} = -A_N < 0 \quad (!) \quad (4.34)$$

und

$$I_{CB0} = \frac{I_{CE} I_{EC}}{I_{EE}} - I_{CC} \quad (4.35)$$

A wird als **Stromverstärkung der Basisschaltung** bezeichnet und es gilt $|A| \approx 1$ wie auch die Übertragungskennlinie des BC107 in Abb. 4.18 zeigt. Das **negative Vorzeichen** ergibt sich auf Grund des verwendeten Bezugspfeilsystems (Abb. 4.1).

Auch hier werden wieder drei verschiedene Fälle betrachtet:

- $U_{CB} / U_{th} \gg 1$ (Normalbetrieb)
- $U_{CB} = 0$
- $U_{CB} < 0$

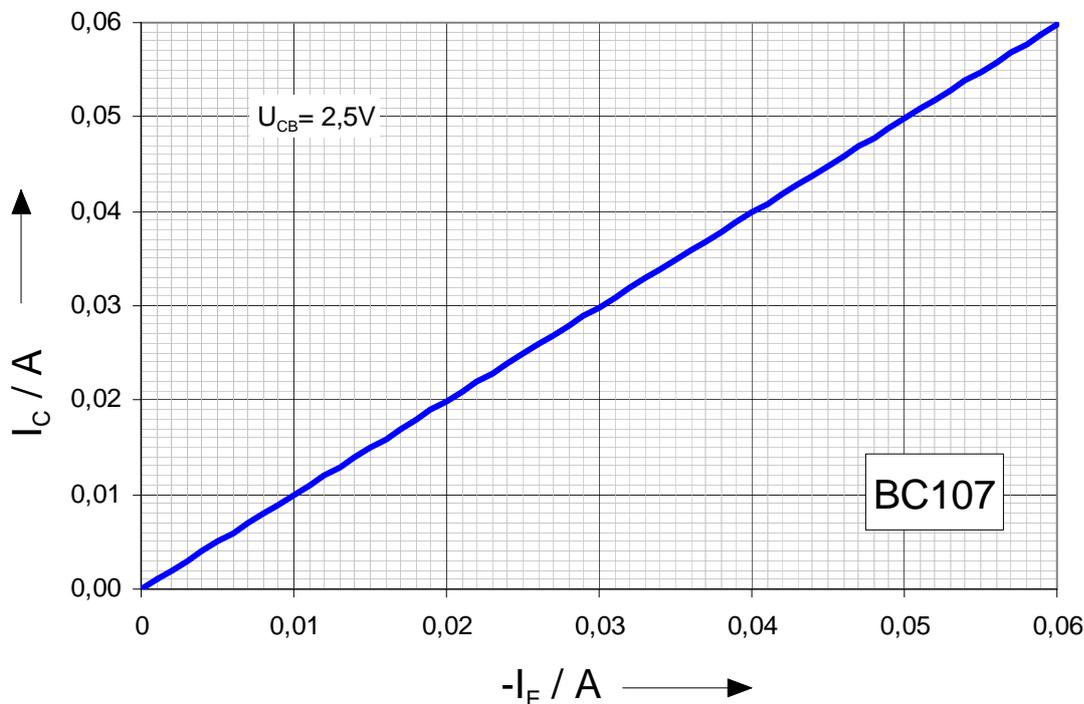


Abb. 4.18 Kennlinie $I_C = f(I_E)$ des BC 107 für $U_{CB} = 2V$ (Normalbetrieb)

zu a) $U_{CB} / U_{th} \gg 1$ (Normalbetrieb)

Damit ergibt sich der Strom (Abb. 4.18) aus Gl. (4.33):

$$I_C = A \cdot I_E - I_{CB0} \approx -I_E \tag{4.36}$$

zu b) Bei $U_{CB} = 0$ folgt:

$$I_C = A \cdot I_E \tag{4.37}$$

Damit ergibt sich nahezu der gleiche Strom wie unter a) (Gl.(4.36)).

zu c) Wird die BC-Diode in Durchlassrichtung vorgespannt mit $U_{CB} < 0$, muss die vollständige Gl. (4.33) herangezogen werden:

$$I_C = I_E \cdot A + I_{CB0} \cdot (e^{-U_{CB} / U_{th}} - 1) \tag{4.38}$$

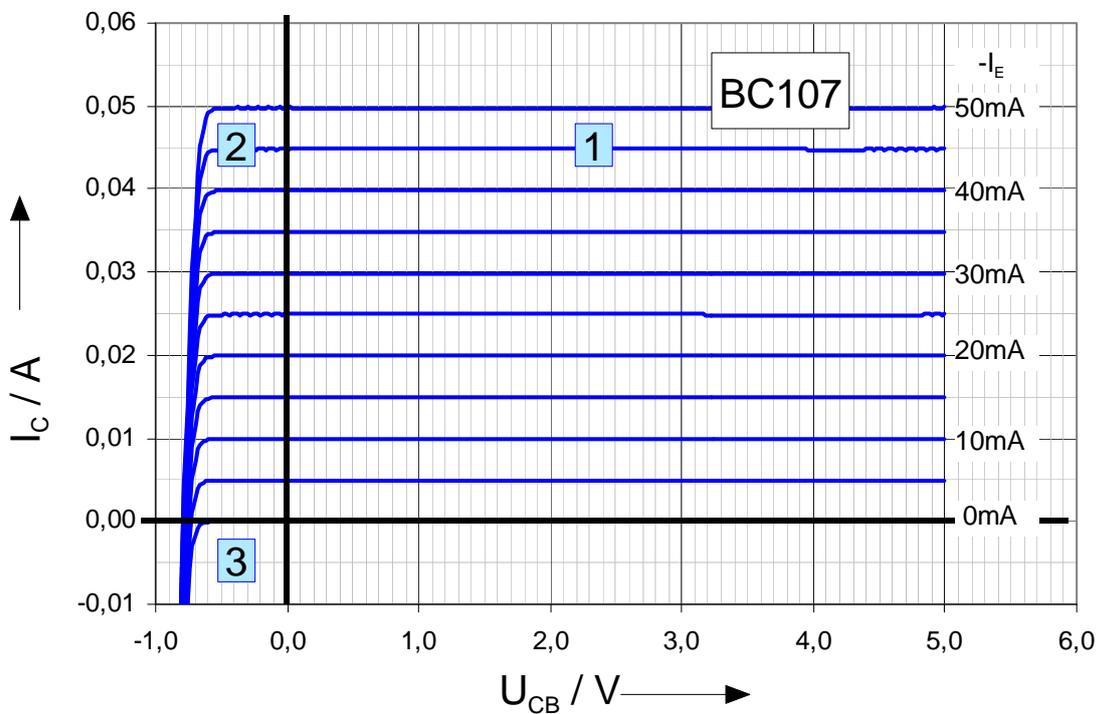


Abb. 4.19 Ausgangskennlinienfeld $I_C = f(U_{CB}, I_E)$ des BC 107

Dies führt dann, wie oben bereits dargestellt, bei $U_{CB} < 0$ zu einer Abnahme des Stromes I_C bis hin zu negativen Werten.

Damit ergibt sich das in Abb. 4.19 dargestellte AKF. Aufgrund der konstanten Stromverstärkung $|A| \approx 1$ bleibt der Abstand der Kennlinienäste konstant.

4.3.2.2 Die Emitterschaltung

Werden die Basis als Eingang und der Kollektor als Ausgang sowie der Emitter als gemeinsamer Anschluss für Ein- und Ausgang beschaltet, so spricht man von der Emitterschaltung.

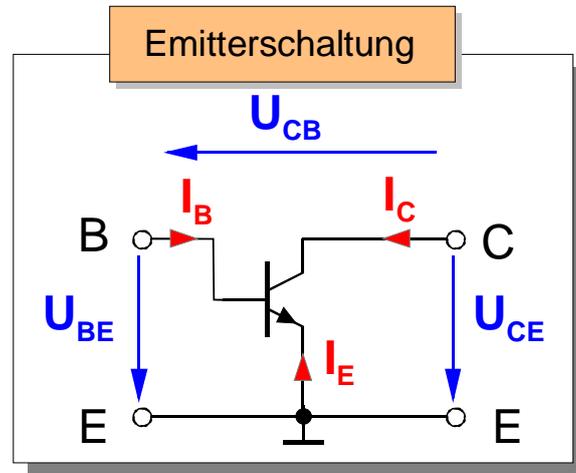


Abb. 4.20 Emitterschaltung

4.3.2.2.1 Die Eingangskennlinie $I_B = f(U_{BE})$

Aus den Ebers-Moll-Gleichungen folgte (Gl.(4.24))

$$I_B = (1 - A_N) \cdot I_E + (1 - A_I) \cdot I_C$$

bzw. mit Gl. (4.20) und (4.21)

$$\begin{aligned} I_B &= (1 - A_N) \cdot I_{EE} \cdot \left(e^{-U_{EB}/U_{th}} - 1 \right) + (1 - A_I) \cdot I_{CC} \cdot \left(e^{-U_{CB}/U_{th}} - 1 \right) \\ \Rightarrow I_B &= (I_{EE} - I_{CE}) \cdot \left(e^{-U_{EB}/U_{th}} - 1 \right) + (I_{CC} - I_{EC}) \cdot \left(e^{-U_{CB}/U_{th}} - 1 \right) \end{aligned} \quad (4.39)$$

Für den **Normalbetrieb** des Transistors mit $U_{CB}/U_{th} \gg 1$ und mit

$$U_{EB} = -U_{BE} \quad (4.40)$$

ergibt sich:

$$I_B = (I_{EE} - I_{CE}) \cdot \left(e^{U_{BE}/U_{th}} - 1 \right) - (I_{CC} - I_{EC}) \quad (4.41)$$

Damit ergibt sich, wie zu erwarten ist, als Eingangskennlinie eine Diodenkennlinie. In Abb. 4.21 ist die Kennlinie des BC 107 für zwei verschiedene Temperaturen dargestellt. Eine Temperaturerhöhung führt zu einer Verschiebung der Kennlinie und damit bei konstanter Eingangsspannung U_{BE} zu einer Erhöhung des Eingangsstroms I_B (vgl. Kap. 3.1.5.3). Daran erkennt man bereits das stark temperaturabhängige Verhalten des Bipolartransistors.

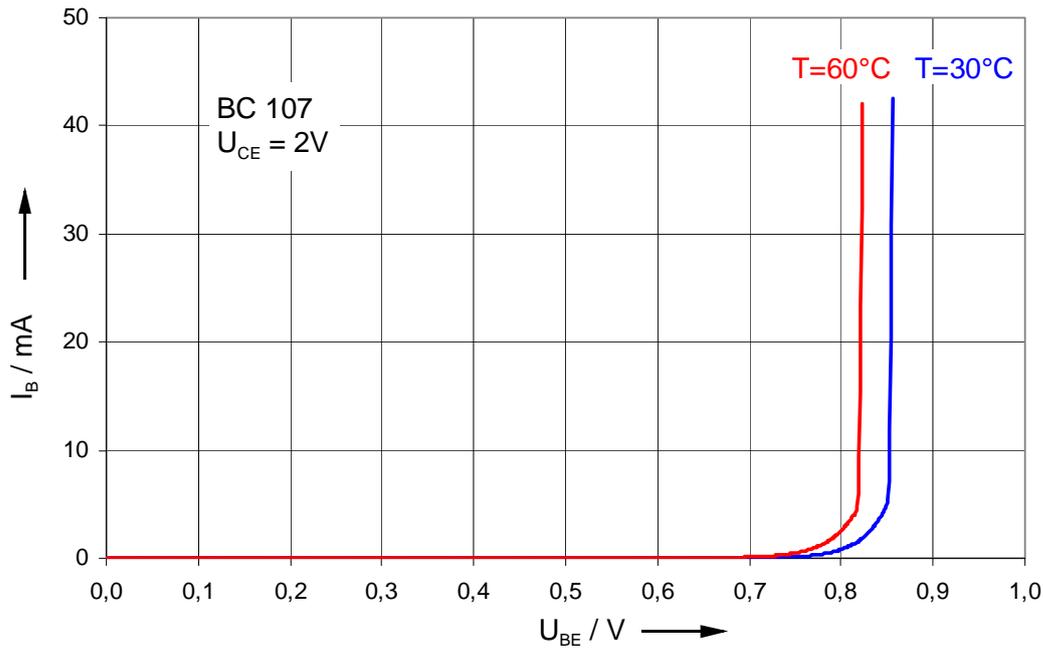


Abb. 4.21 Eingangskennlinie $I_B = f(U_{BE})$ des BC 107 in Emitterschaltung für $U_{CE} = 2V$ und zwei verschiedene Temperaturen

4.3.2.2.2 Das Ausgangskennlinienfeld (AKF)

Auch in diesem Fall existieren zwei verschiedene Darstellungen.

4.3.2.2.2.1 Das AKF mit U_{BE} als Steuerparameter ($I_C = f(U_{CE}, U_{BE})$)

Legt man wiederum die Ebers-Moll-Gleichungen zugrunde, so folgt mit Gl. (4.13):

$$I_C = I_{CE} \cdot \left(e^{U_{BE}/U_{th}} - 1 \right) - I_{CC} \cdot \left(e^{-U_{CB}/U_{th}} - 1 \right)$$

Auch hier werden wieder die drei bekannten Fälle betrachtet:

- a) $U_{CB} / U_{th} \gg 1$ (Normalbetrieb)
- b) $U_{CB} = 0$
- c) $U_{CB} < 0$

a) Im Normalbetrieb des Transistors ist die BC-Diode mit $U_{CB} / U_{th} \gg 1$ wiederum gesperrt.

Damit folgt:

$$I_C = I_{CE} \cdot \left(e^{U_{BE}/U_{th}} - 1 \right) + I_{CC} \tag{4.42}$$

bzw.

$$I_C = I_{CE} \cdot e^{U_{BE}/U_{th}} - (I_{CE} - I_{CC}) \quad (4.43)$$

D.h., der Abstand der Kennlinienäste steigt **exponentiell** mit linear wachsendem Parameter U_{BE} und der Strom ist unabhängig von der Ausgangsspannung U_{CE} .

b) Bei $U_{CB} = 0$ ergibt sich mit:

$$I_C = I_{CE} \cdot \left(e^{U_{BE}/U_{th}} - 1 \right) \quad (4.44)$$

wieder eine Diodenkennlinie.

zu c) Wird die BC-Diode mit $U_{CB} < 0$ in Durchlassrichtung vorgespannt, so müssen wiederum die vollständigen Ebers-Moll Gleichungen (Gl. (4.18) bzw. (4.19)) herangezogen werden:

Mit

$$U_{CB} = U_{CE} - U_{BE}$$

folgt dann für den Kollektorstrom:

$$I_C = I_{CE} \cdot \left(e^{U_{BE}/U_{th}} - 1 \right) - I_{CC} \cdot \left(e^{-U_{CE}/U_{th}} \cdot e^{U_{BE}/U_{th}} - 1 \right) \quad (4.45)$$

Mit

$$U_{CB} < 0$$

ergibt sich auch

$$U_{CE} < U_{BE}$$

Für kleine Spannungen U_{CE} kompensiert wieder der 2. Term in Gl. (4.45) den 1. Term und führt somit zu einer Reduzierung des Stroms I_C bis sich letztendlich bei $U_{CE} = 0$

$$I_C = \left(e^{U_{BE}/U_{th}} - 1 \right) \cdot (I_{CE} - I_{CC}) \approx 0 \quad (4.46)$$

ergibt.

Die Ergebnisse spiegeln sich im AKF in Abb. 4.22 wider.

Man unterscheidet dabei drei Bereiche:

- **Sättigungsbereich** (für kleine Spannungen U_{CE} mit $U_{CB} < 0$) (Gl. (4.45))
- **aktiver Bereich (Normalbetrieb)**, für große Spannungen U_{CE} mit $U_{CB} > 0$) (Gl.(4.43))
- **Sperrbereich** (für Eingangsspannungen unterhalb der Flussspannung, so dass die BE-Diode gesperrt ist)

WICHTIG: Beim **Bipolartransistor** liegt das Sättigungsgebiet im Bereich **niedriger Ausgangsspannungen** U_{CE} , während das Sättigungsgebiet beim **Feldeffekttransistor (FET)** im Bereich **hoher Ausgangsspannungen** U_{DS} liegt!

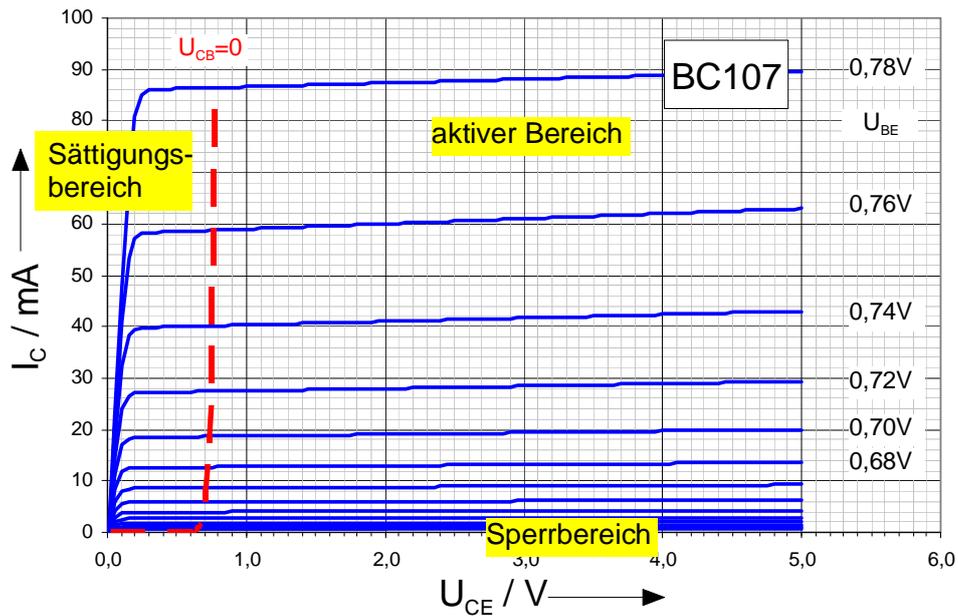


Abb. 4.22 Ausgangskennlinienfeld $I_C = f(U_{CE}, U_{BE})$ des BC 107 für $U_{CE} > 0$

4.3.2.2.2 AKF mit I_B als Steuerparameter ($I_C = f(U_{CE}, I_B)$)

In den Ebers-Moll Gleichungen ist der steuernde Parameter jeweils die Spannung U_{EB} . Um den Strom als Steuerparameter zu erhalten, muss die Spannung U_{EB} eliminiert werden.

Aus den Ebers-Moll-Gleichungen folgte (Gl.(4.25)):

$$I_B = (1 - A_N) \cdot I_{EE} \cdot (e^{-U_{EB}/U_{th}} - 1) + (1 - A_I) \cdot I_{CC} \cdot (e^{-U_{CB}/U_{th}} - 1)$$

Und damit

$$(e^{-U_{EB}/U_{th}} - 1) = I_B \frac{1}{(1 - A_N) \cdot I_{EE}} - \frac{(1 - A_I) \cdot I_{CC}}{(1 - A_N) \cdot I_{EE}} \cdot (e^{-U_{CB}/U_{th}} - 1) \quad (4.47)$$

Mit Gl. (4.21):

$$I_C = A_N \cdot I_{EE} (e^{-U_{EB}/U_{th}} - 1) - I_{CC} (e^{-U_{CB}/U_{th}} - 1)$$

folgt dann:

$$I_C = A_N \cdot I_{EE} \left(I_B \cdot \frac{1}{(1 - A_N) \cdot I_{EE}} - \frac{1 - A_I}{1 - A_N} \cdot \frac{I_{CC}}{I_{EE}} \cdot (e^{-U_{CB}/U_{th}} - 1) \right) - I_{CC} (e^{-U_{CB}/U_{th}} - 1)$$

$$I_C = \frac{A_N}{1 - A_N} \cdot I_B - I_{CC} \cdot \left(A_N \cdot \frac{1 - A_I}{1 - A_N} + 1 \right) \cdot (e^{-U_{CB}/U_{th}} - 1) \quad (4.48)$$

Der Kollektorstrom setzt sich damit aus einem von I_B und einem von U_{CB} abhängigen Anteil zusammen. Dementsprechend lässt sich schreiben:

$$I_C = B \cdot I_B - I_{CE0} \cdot (e^{-U_{CB}/U_{th}} - 1) \quad (4.49)$$

mit (Gln. (4.34), (4.22), (4.23)):

$$B = \frac{A_N}{1 - A_N} = \frac{-A}{1 + A} = \frac{I_{CE}}{I_{EE} - I_{CE}} \gg 1 \quad (A < 0!!) \quad (4.50)$$

und

$$I_{CE0} = I_{CC} \cdot \left(A_N \cdot \frac{1 - A_I}{1 - A_N} + 1 \right) = \frac{I_{CC} I_{EE} - I_{CE} I_{EC}}{I_{EE} - I_{CE}} = I_{CC} - \frac{I_{CE} I_{EC}}{I_{EE} - I_{CE}} \quad (4.51)$$

I_{CE0} ist der Reststrom bei $I_B = 0$.

B wird als **Stromverstärkung der Emitterschaltung** bezeichnet.

Dabei gilt:

$$B \gg 1$$

Auch hier werden wieder die drei bekannten Fälle betrachtet:

- a) $U_{CB} / U_{th} \gg 1$ (Normalbetrieb)
- b) $U_{CB} = 0$
- c) $U_{CB} < 0$

a) Für $U_{CB} / U_{th} \gg 1$ ist die BC-Diode wiederum gesperrt und es gilt entsprechend Gl.(4.49):

$$I_C = B \cdot I_B + I_{CE0} \quad (4.52)$$

d.h., im Normalbetrieb (aktiver Bereich) steigt der Ausgangsstrom I_C entsprechend der Stromverstärkung B **linear** mit I_B an.

Dementsprechend sind die Kennlinienäste bei linear steigendem Parameter I_B unter der Annahme nahezu konstanter Stromverstärkung B äquidistant.

b) Bei $U_{CB} = 0$ ergibt sich mit Gl.(4.49)

$$I_C = B \cdot I_B \quad (4.53)$$

Der Ausgangsstrom I_C steigt ebenfalls entsprechend der Stromverstärkung B **linear** mit I_B an.

c) Wird die BC-Diode mit $U_{CB} < 0$ in Durchlassrichtung vorgespannt, so muss die vollständige Gl. (4.49) herangezogen werden:

$$I_C = B \cdot I_B - I_{CE0} \cdot \left(e^{-U_{CB}/U_{th}} - 1 \right) \tag{4.54}$$

Mit

$$U_{CB} = U_{CE} - U_{BE}$$

ergibt sich der Kollektorstrom zu:

$$I_C = B \cdot I_B - I_{CE0} \cdot \left(e^{-U_{CE}/U_{th}} \cdot e^{U_{BE}/U_{th}} - 1 \right)$$

Mit

$$U_{CB} < 0$$

ergibt sich auch

$$U_{CE} < U_{BE}$$

Für kleine Spannungen U_{CE} kompensiert wieder der 2. Term in Gl. (4.54) den 1. Term und führt somit zu einer Reduzierung des Stroms I_C bis sich letztendlich bei $U_{CE} = 0$

$$I_C \approx 0$$

ergibt.

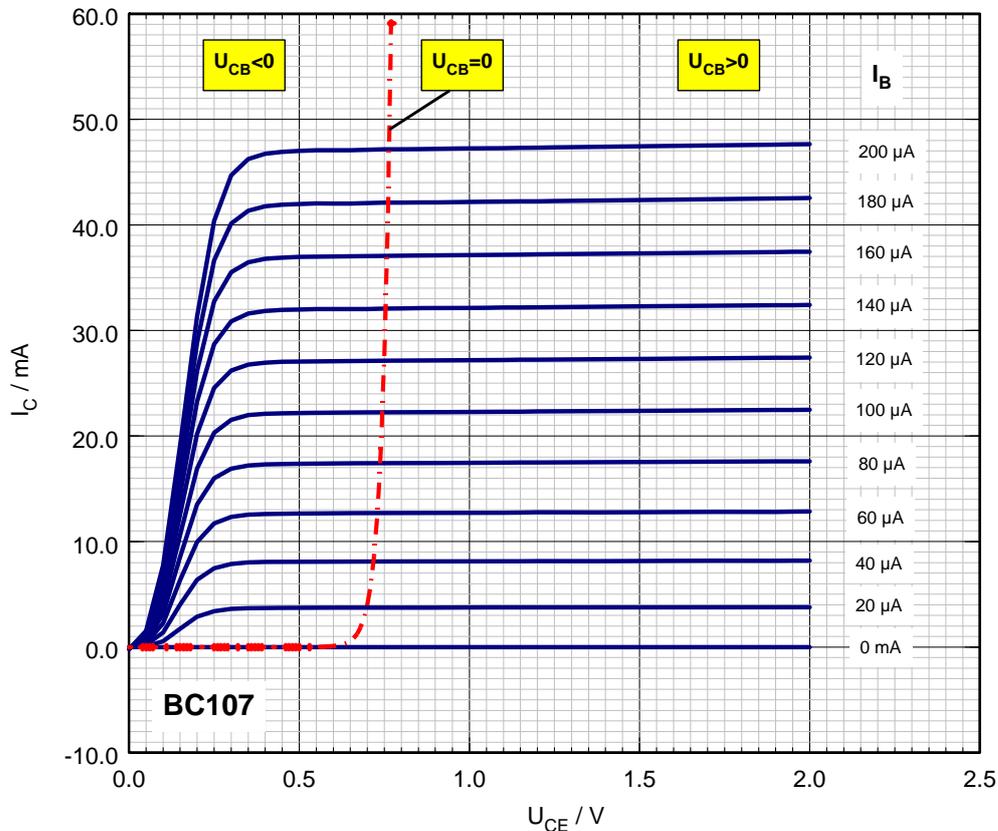


Abb. 4.23 Ausgangskennlinienfeld $I_C = f(U_{CE}, I_B)$ des BC107 für $U_{CE} > 0$

Abb. 4.23 zeigt das Ausgangskennlinienfeld des BC 107 mit dem Steuerstrom I_B als Parameter.

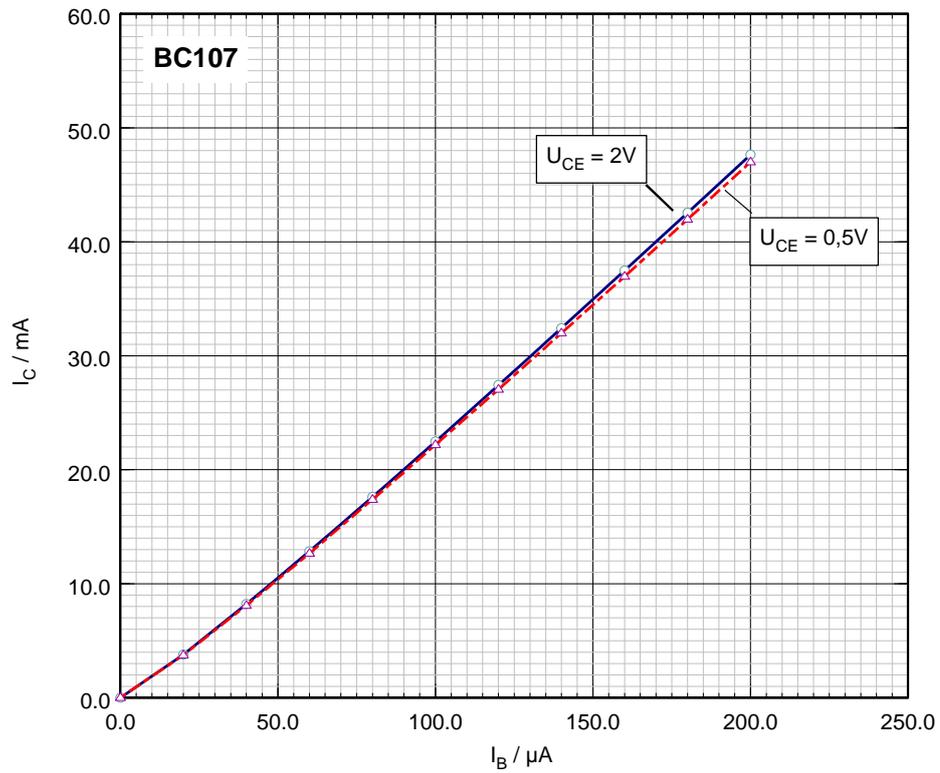


Abb. 4.24 Übertragungskennlinie $I_C = f(I_B)$ des BC 107

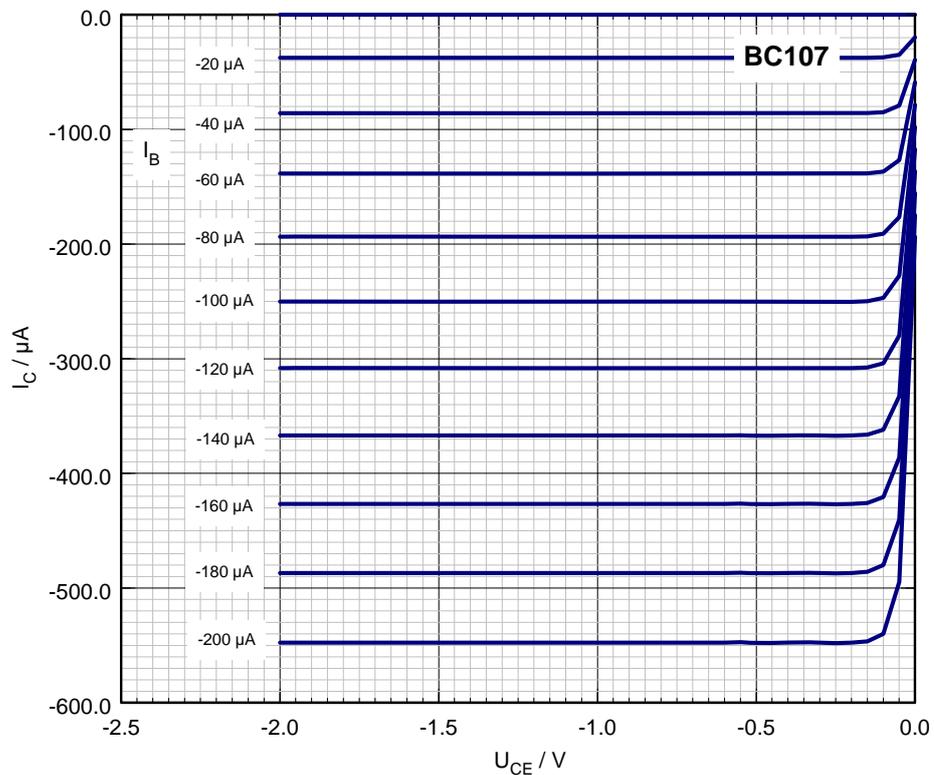


Abb. 4.25 Ausgangskennlinienfeld $I_C = f(U_{CE}, I_B)$ des BC107 für $U_{CE} < 0$

Abb. 4.24 zeigt die Übertragungskennlinie des Bipolartransistors BC107. Man erkennt, dass die ideale Annahme der linearen Abhängigkeit in weiten Bereichen erfüllt ist.

Auch für Spannungen $U_{CE} < 0$ existiert ein Ausgangskennlinienfeld (Abb. 4.25), wenn auch die Rückwärtsstromverstärkung deutlich geringer ist.

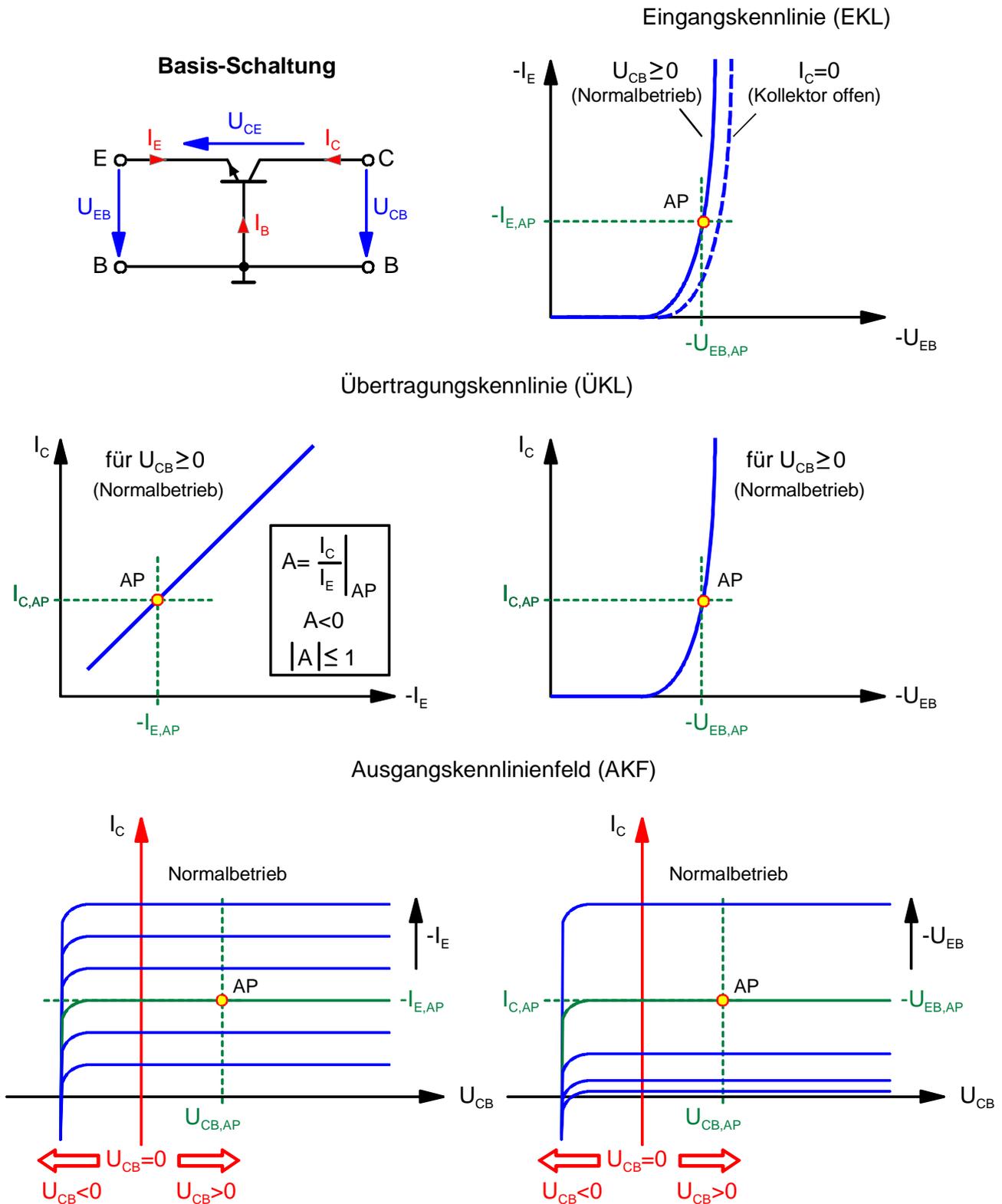
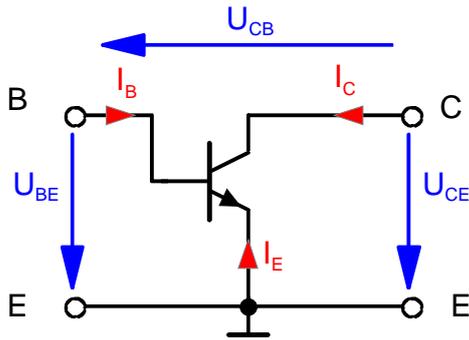
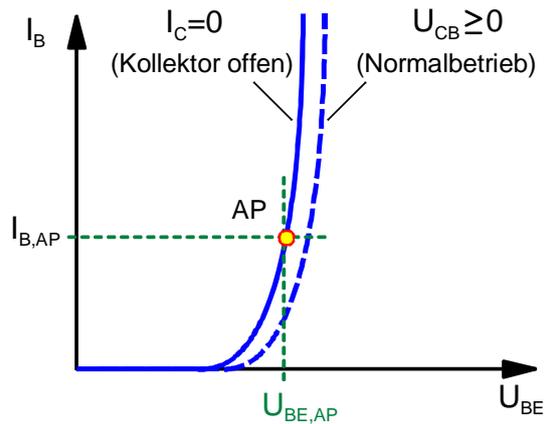


Abb. 4.25b Prinzipielle Kennlinien der Basis-Schaltung

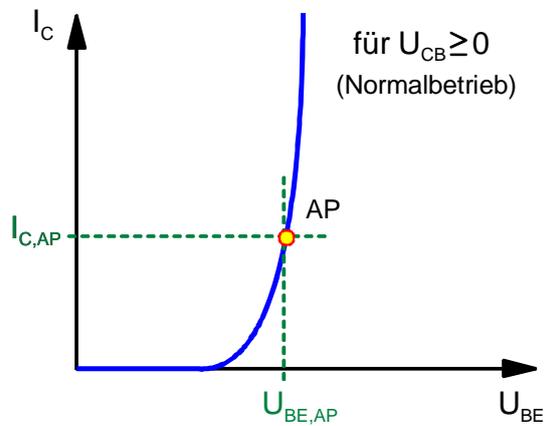
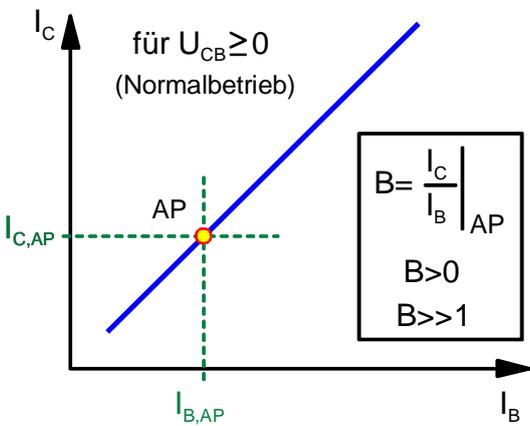
Emitter-Schaltung



Eingangskennlinie (EKL)



Übertragungskennlinie (ÜKL)



Ausgangskennlinienfeld (AKF)

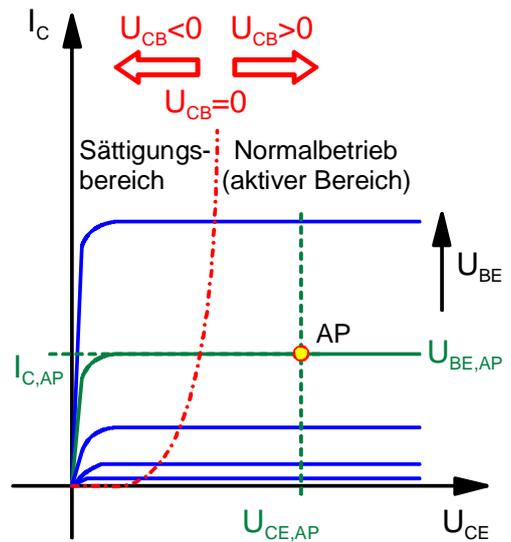
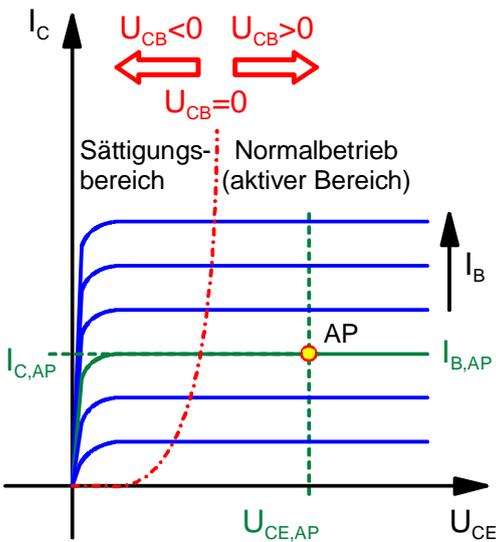


Abb. 4.25c Prinzipielle Kennlinien der Emitter-Schaltung

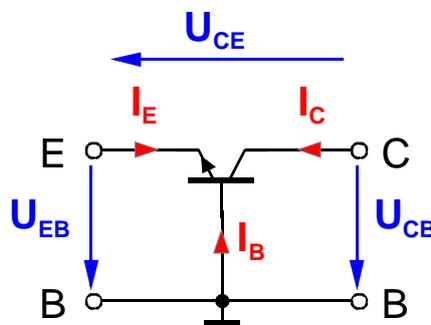
Für eine bessere Übersicht sind in Abb. 4.15.b und 4.25c die prinzipiellen Kennlinien der Basis- und der Emitter-Schaltung gegenübergestellt.

4.3.3 Das Transportmodell

In der Literatur findet man neben dem Ebers-Moll Modell häufig das so genannte **Transportmodell**. Es erleichtert in vielen Fällen die Berechnung bzw. Modellierung des Bipolartransistors und soll daher hier auch vorgestellt werden.

Die beiden Quellen im Ebers-Moll Modell (Abb. 4.11) lassen sich zu einer Quelle zusammenfassen, so dass aus der Betrachtung eines *Injektionsmodells* (Ladungsträger werden wie bei einer Diode in die Basis injiziert) ein *Transportmodell* wird (die Basisströme errechnen sich aus dem Transferstrom zwischen Emitter und Kollektor) (Abb. 4.26). Diese Umwandlung ist zunächst rein mathematischer Natur, führt schließlich aber zu einer einfacheren Umsetzungsmöglichkeit in so genannten Kompaktmodellen, die häufig für die Simulation von Transistoren verwendet werden.

Der Vorteil liegt darin, dass nur noch eine gesteuerte Stromquelle verwendet wird. Um die Gültigkeit des Modells zu bewahren, müssen dafür aber auch die Ströme durch die jeweiligen Dioden entsprechend modifiziert werden.



Ebers-Moll Modell (EMM)

Transportmodell (TM)

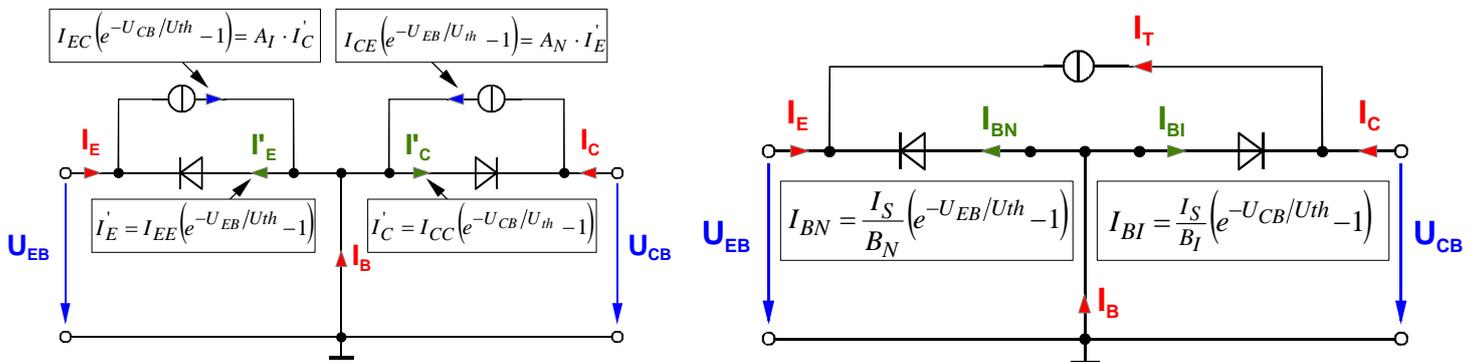


Abb. 4.26 Das Transport-Modell im Vergleich zum vorher behandelten Ebers-Moll Modell

Die Größen des Transport-Modells (TM) lassen sich aus den Parametern des Ebers-Moll Modells (EMM) herleiten:

Nach Gl.(4.25) gilt:

$$I_B = (1 - A_N) \cdot I_{EE} \cdot \left(e^{-U_{EB}/U_{th}} - 1 \right) + (1 - A_I) \cdot I_{CC} \cdot \left(e^{-U_{CB}/U_{th}} - 1 \right) \quad (4.55)$$

Unter Verwendung von Gln. (4.22) und (4.23) mit

$$I_{EC} = A_I \cdot I_{CC} \quad \Rightarrow \quad I_{CC} = \frac{I_{EC}}{A_I} \quad \text{bzw.} \quad A_I = \frac{I_{EC}}{I_{CC}}$$

$$I_{CE} = A_N \cdot I_{EE} \quad \Rightarrow \quad I_{EE} = \frac{I_{CE}}{A_N} \quad \text{bzw.} \quad A_N = \frac{I_{CE}}{I_{EE}}$$

folgt:

$$I_B = \frac{1 - A_N}{A_N} \cdot I_{CE} \cdot \left(e^{-U_{EB}/U_{th}} - 1 \right) + \frac{1 - A_I}{A_I} \cdot I_{EC} \cdot \left(e^{-U_{CB}/U_{th}} - 1 \right) \quad (4.56)$$

Setzt man voraus, dass gilt

$$A_E = A_C$$

und somit

$$I_{CE} = I_{EC} = I_S \quad (4.57)$$

gilt, ergibt sich der Basisstrom zu:

$$I_B = \frac{1 - A_N}{A_N} \cdot I_S \cdot \left(e^{-U_{EB}/U_{th}} - 1 \right) + \frac{1 - A_I}{A_I} \cdot I_S \cdot \left(e^{-U_{CB}/U_{th}} - 1 \right) \quad (4.58)$$

und lässt sich damit durch zwei Diodenströme beschreiben:

$$I_B = \frac{1}{B_N} \cdot I_S \cdot \left(e^{-U_{EB}/U_{th}} - 1 \right) + \frac{1}{B_I} \cdot I_S \cdot \left(e^{-U_{CB}/U_{th}} - 1 \right) \quad (4.59)$$

bzw.

$$I_B = I_{BN} + I_{BI} \quad (4.60)$$

mit

$$B_N = \frac{A_N}{1 - A_N} \quad A_N = \frac{B_N}{B_N + 1} \quad (4.61)$$

$$B_I = \frac{A_I}{1 - A_I} \quad A_I = \frac{B_I}{B_I + 1} \quad (4.62)$$

Dabei bedeuten:

B_N : Stromverstärkung der Emitter-Schaltung in Vorwärtsrichtung (N: Normal)

B_I : Stromverstärkung der Emitter -Schaltung in Rückwärtsrichtung (I: Invers)

und

$$I_{BN} = \frac{I_S}{B_N} \cdot \left(e^{-U_{EB}/U_{th}} - 1 \right) \quad (4.63)$$

$$I_{BI} = \frac{I_S}{B_I} \cdot \left(e^{-U_{CB}/U_{th}} - 1 \right) \quad (4.64)$$

Aus dem Ersatzschaltbild (Abb. 4.26) folgt:

$$I_E = -I_{BN} - I_T \quad (4.65)$$

$$\Rightarrow I_T = -I_{BN} - I_E \quad (4.66)$$

Mit Gl. (4.9) ergibt sich:

$$\begin{aligned} I_T &= -I_{BN} + I_{EE} \left(e^{-U_{EB}/U_{th}} - 1 \right) - I_{EC} \left(e^{-U_{CB}/U_{th}} - 1 \right) \\ \Rightarrow I_T &= -I_{BN} + \frac{I_{EC}}{A_N} \left(e^{-U_{EB}/U_{th}} - 1 \right) - I_{EC} \left(e^{-U_{CB}/U_{th}} - 1 \right) \end{aligned}$$

und mit Gl. (4.57), (4.63) und (4.64) lässt sich schreiben:

$$\begin{aligned} I_T &= -I_{BN} + \frac{1}{A_N} \cdot I_S \cdot \left(e^{-U_{EB}/U_{th}} - 1 \right) - I_S \left(e^{-U_{CB}/U_{th}} - 1 \right) \\ \Rightarrow I_T &= -I_{BN} + \frac{1}{A_N} \cdot B_N I_{BN} - B_I I_{BI} \\ \Rightarrow I_T &= -I_{BN} + \frac{B_N + 1}{B_N} \cdot B_N I_{BN} - B_I I_{BI} \\ \Rightarrow I_T &= -I_{BN} + (B_N + 1) \cdot I_{BN} - B_I I_{BI} \\ \Rightarrow I_T &= B_N \cdot I_{BN} - B_I I_{BI} \end{aligned} \quad (4.67)$$

Einsetzen in Gl. (4.65) ergibt:

$$I_E = -I_{BN} - B_N \cdot I_{BN} + B_I I_{BI} = -(1 + B_N) \cdot I_{BN} + B_I I_{BI}$$

und mit Gl. (4.63) und (4.64) folgt:

$$\begin{aligned} I_E &= -(1 + B_N) \cdot \frac{I_S}{B_N} \cdot \left(e^{-U_{EB}/U_{th}} - 1 \right) + B_I \frac{I_S}{B_I} \cdot \left(e^{-U_{CB}/U_{th}} - 1 \right) \\ I_E &= I_S \left[- \left(1 + \frac{1}{B_N} \right) \cdot e^{-U_{EB}/U_{th}} + \frac{1}{B_N} + 1 + e^{-U_{CB}/U_{th}} - 1 \right] \\ \Rightarrow I_E &= I_S \left[- \left(1 + \frac{1}{B_N} \right) \cdot e^{-U_{EB}/U_{th}} + \frac{1}{B_N} + e^{-U_{CB}/U_{th}} \right] \end{aligned} \quad (4.68)$$

Analog dazu mit

$$I_C = -I_{BI} + I_T$$

folgt

$$\begin{aligned} I_C &= -I_{BI} + B_N \cdot I_{BN} - B_I I_{BI} = -(1 + B_I) \cdot I_{BI} + B_N \cdot I_{BN} \\ \Rightarrow I_C &= -\frac{I_S}{B_I} \cdot \left(e^{-U_{CB}/U_{th}} - 1 \right) \cdot (B_I + 1) + B_N \cdot \frac{I_S}{B_N} \cdot \left(e^{-U_{EB}/U_{th}} - 1 \right) \end{aligned}$$

$$\begin{aligned} \Rightarrow I_C &= I_S \cdot \left[-\frac{B_I + 1}{B_I} \cdot (e^{-U_{CB}/U_{th}} - 1) + (e^{-U_{EB}/U_{th}} - 1) \right] \\ \Rightarrow I_C &= I_S \cdot \left[-\frac{B_I + 1}{B_I} \cdot e^{-U_{CB}/U_{th}} + \frac{1}{B_I} + 1 + e^{-U_{EB}/U_{th}} - 1 \right] \\ \Rightarrow I_C &= I_S \cdot \left[\frac{1}{B_I} + e^{-U_{EB}/U_{th}} - \left(1 + \frac{1}{B_I} \right) \cdot e^{-U_{CB}/U_{th}} \right] \end{aligned} \quad (4.69)$$

4.3.4 Der Gummel-Plot

Um für die Charakterisierung der Bipolartransistoren sowohl sehr kleine als auch sehr große Ströme übersichtlich darstellen zu können, verwendet man den so genannten **Gummel-Plot**.

Dabei werden Ein- und Ausgangsstrom in Abhängigkeit von der Basis-Emitter Spannung U_{BE} in einer halblogarithmischen Darstellung aufgetragen (Abb. 4.27). Dabei wird in die Kollektor-Basis Spannung U_{CB} konstant (meist $U_{CB} = 0$) gehalten. In diesem Fall werden die Ströme an der Sättigungsgrenze des Bipolartransistors aufgenommen (vgl. Abb. 4.23).

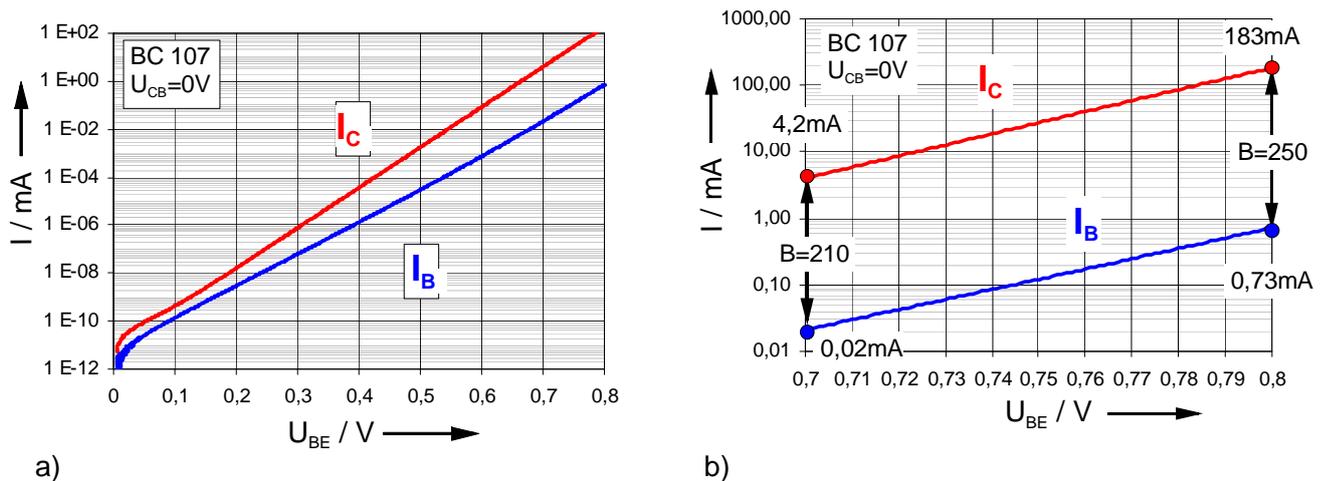


Abb. 4.27 Gummel-Plot a) für $0 \leq U_{BE} \leq 0,8V$ und b) für $0,7V \leq U_{BE} \leq 0,8V$ bei $U_{CB} = 0V$

Aus dem Gummel-Plot kann man auch direkt die DC-Stromverstärkung des Transistors $B = I_C / I_B$ im entsprechenden Arbeitspunkt ablesen. Betrachtet man den gesamten Spannungsbereich, erkennt man die starke Variation der Stromverstärkung. Beschränkt man sich jedoch auf den eigentlichen Arbeitsbereich des Transistors z. B. für $0,7V \leq U_{BE} \leq 0,8V$ (Abb. 4.27b), so bleibt die Stromverstärkung relativ konstant.

4.3.5 Der Early-Effekt

Für den aktiven Bereich (Normalbetrieb) gilt entsprechend Gl. (4.44):

$$I_C = B \cdot I_B + I_{CE0} \quad (4.70)$$

Danach ist in diesem Betrieb der Kollektorstrom unabhängig von der Kollektorspannung U_{CE} . Diese Annahme stimmt aber in vielen Fällen nicht mit den Kennlinien realer Transistoren überein.

Vielmehr zeigt die Praxis einen Anstieg des Kollektorstroms auch im aktiven Betrieb mit wachsender Kollektorspannung U_{CE} , wie in Abb. 4.28 dargestellt ist.

Die Ursache für diesen Anstieg ist der so genannte **Early-Effekt**: Mit wachsender Sperrspannung am Kollektor-Basis-Übergang weitet sich die Sperrschicht dieses Überganges immer weiter in das Basisgebiet aus, und dadurch wird die elektronisch wirksame Basisweite W_B kleiner. Wenn aber die elektronisch wirksame Basisweite kleiner wird und andererseits die Emitter-Basis-Spannung und damit die Randkonzentration der Minoritätsträger am Emitterring der Basis entsprechend

$$n_{pB0} \cdot \exp(U_{BE} / U_{th})$$

konstant bleibt, wird der Ladungsträgerdichtegradient und damit der Strom größer.

Im Extremfall kann sich sogar die Sperrschicht des Basis-Kollektor-Überganges bis zur Sperrschicht des Emitter-Basis-Überganges ausdehnen. Dann erfolgt der sogenannte *Punchthrough* (Durchgreifen), was zu einem durchbruchartigen, steilen Stromanstieg führt [20].

Eine Modellierung dieses Verhaltens ist durch folgende lineare empirische Approximation möglich:

$$I_C = B \cdot I_B + I_{CE0} \left(1 + \frac{U_{CE}}{U_{Early}} \right) \tag{4.71}$$

U_{Early} , ist hierbei ein empirisch aus dem Schnittpunkt der Kennlinien gemäß Abb. 4.28 zu ermittelnder Parameter.

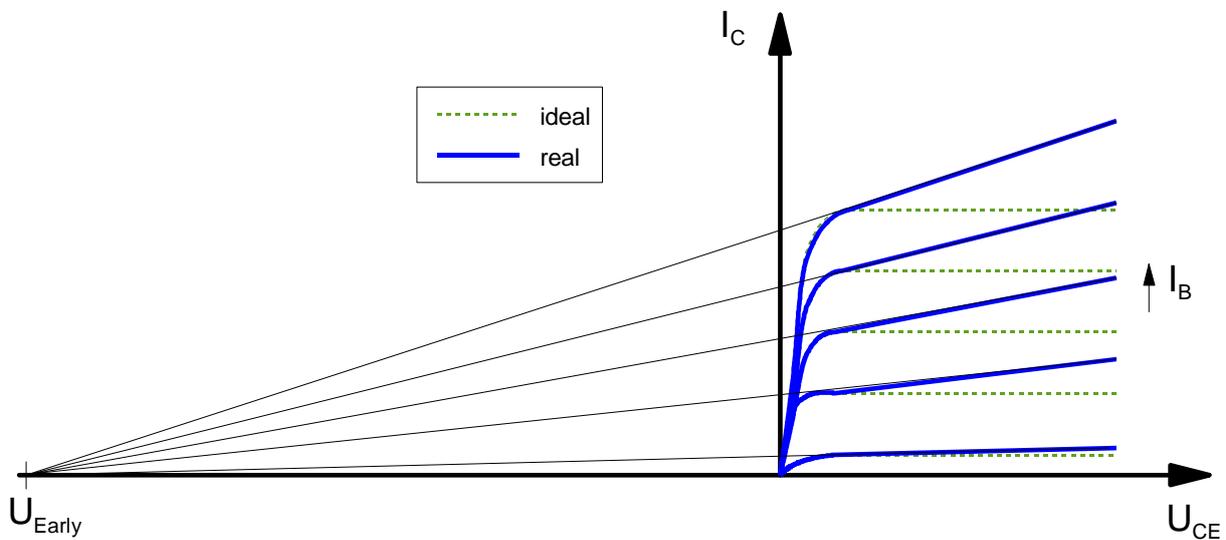


Abb. 4.28 Ausgangskennlinienfeld für den idealen und den realen Fall mit Bestimmung der Early-Spannung U_{Early}

4.4 Das quasistationäre Kleinsignalverhalten des Bipolartransistors in Emitter-Schaltung

Wie bereits in Zusammenhang mit der Behandlung der Feldeffekttransistoren erwähnt, müssen zur praktischen Dimensionierung einer Schaltung die elektrischen Eigenschaften des Transistors bekannt sein, damit man sie durch eine geeignete Beschaltung optimal dem Anwendungszweck anpassen kann.

Der Bipolartransistor ist wie der FET als Vierpol bzw. Zweitor aufzufassen. Analog zu Kap. 4.3 unterscheidet man auch bei der Kleinsignalbeschreibung nach

- Emitter-
- Basis-
- Kollektorschaltung

je nachdem, welcher Anschluss als gemeinsamer Bezugspunkt für Ein- und Ausgang gewählt wird.

An dieser Stelle wird nur auf das quasistationäre Kleinsignalverhalten der Emitter-Schaltung eingegangen, das sich wieder aus den DC-Kennlinien ableiten lässt (s. Kap. 3.1.6.1.1). Es handelt sich damit auch in diesem Fall wieder um differentielle Größen!

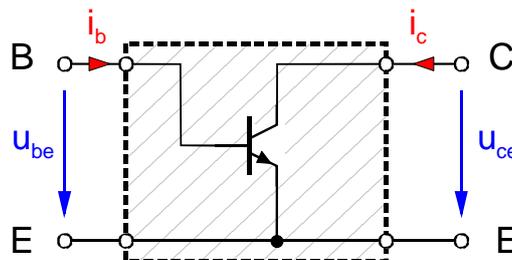


Abb. 4.29 Der Transistor in Emitter-Schaltung als Vierpol bzw. Zweitor

4.4.1 Der Eingangswiderstand

Im Fall des Feldeffekttransistors wurde davon ausgegangen, dass kein Strom in den Gate-Kontakt fließt, und somit wurde von einem unendlich hohen Eingangswiderstand ausgegangen (Kap. 2.4.2.3).

Für den Betrieb eines Bipolartransistors ist dagegen ein Steuerstrom absolut notwendig. Dementsprechend ergibt sich aus der Eingangskennlinie (Abb. 4.21) auch ein differentieller Eingangswiderstand.

Unter Verwendung von Gl. (4.35)

$$I_B = (I_{EE} - I_{CE}) \cdot \left(e^{U_{BE}/U_{th}} - 1 \right) - (I_{CC} - I_{EC})$$

lässt sich bei Vernachlässigung der Sperrströme die Eingangskennlinie $I_B = f(U_{BE})$ darstellen durch:

$$I_B = I_{B0} \cdot \left(e^{U_{BE}/U_{th}} - 1 \right) \quad (4.50)$$

Für einen konventionellen Kleinsignal npn-Silizium-Transistor gilt $I_{B0} \approx 0,1nA$.

Für $U_{BE} \gg U_{th}$ (z.B. für $U_{BE} > 100mV$ bei $T = 300K$ und $U_{th} \approx 25mV$), also in Durchlassrichtung der Basis-Emitter Diode, folgt:

$$I_B \approx I_{B0} \cdot e^{U_{BE}/U_{th}} \tag{4.51}$$

und analog zur Berechnung des Leitwerts der pn-Diode (Kap. 3.1.6.1.1) folgt für den Eingangswiderstand des Bipolartransistors in Emitter-Schaltung r_{be} :

$$\frac{1}{r_{be}} = \frac{i_b}{u_{be}} = \left. \frac{dI_B}{dU_{BE}} \right|_{AP} \tag{4.52}$$

$$\Rightarrow \frac{1}{r_{be}} = \frac{1}{U_{th}} \cdot I_B \tag{4.53}$$

bzw.

$$r_{be} = \frac{U_{th}}{I_B} = f(I_B) \tag{4.54}$$

Der Eingangswiderstand ist auch in diesem Fall abhängig vom jeweiligen Arbeitspunkt! Je größer der Basis-Strom ist, desto steiler wird die Kennlinie und desto kleiner ist der differentielle Eingangswiderstand.

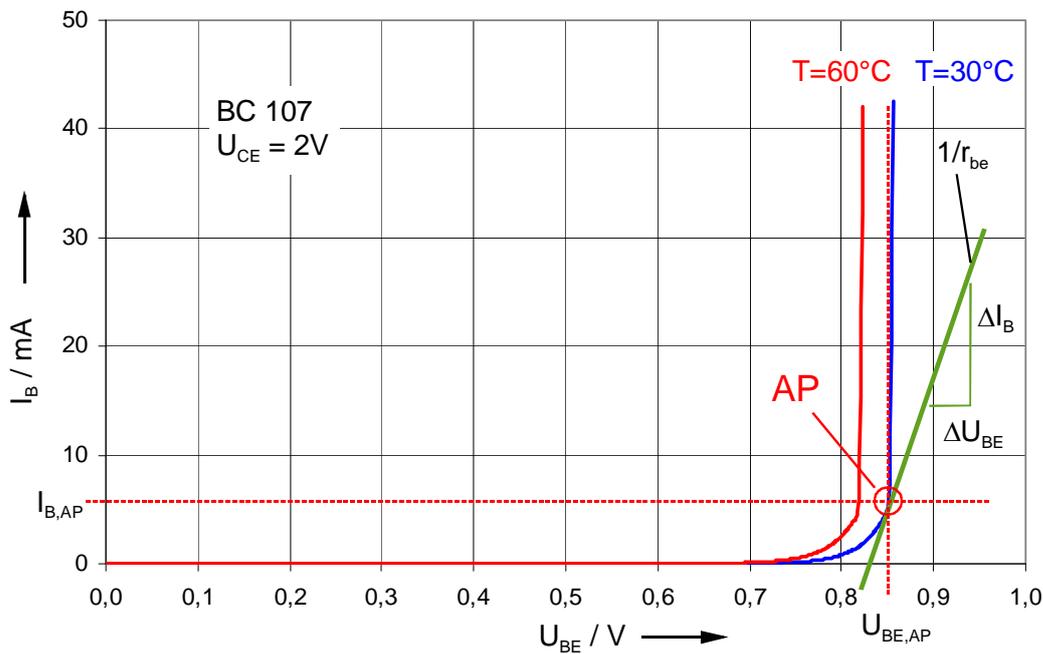


Abb. 4.30 Eingangskennlinie $I_B = f(U_{BE})$ und Bestimmung des Eingangswiderstands r_{be}

Dabei ist zu berücksichtigen, dass sowohl die Kennlinie $I_B = f(U_{BE})$ temperaturabhängig ist und sich mit steigender Temperatur nach links verschiebt (s. Kap. 3.1.5.2 und Abb. 4.30) als auch $U_{th} = kT/q$ mit der Temperatur ansteigt.

4.4.2 Die Stromverstärkung

Der Ausgangsstrom, hier der Kollektorstrom I_C bzw. kleinsignalmäßig i_c , hängt sowohl von den Ein- als auch den Ausgangsgrößen ab.

Der Zusammenhang zwischen Aus- und Eingangsstrom lässt sich aus der **Übertragungskennlinie** des Transistors $I_C = f(I_B)$ (Abb. 4.31) ermitteln und definiert die Stromverstärkung.

Im Fall der Emitter-Schaltung ergibt sich für die Kleinsignal-Stromverstärkung:

$$\beta = \left. \frac{i_c}{i_b} \right|_{AP} = \left. \frac{dI_C}{dI_B} \right|_{AP} \quad (4.55)$$

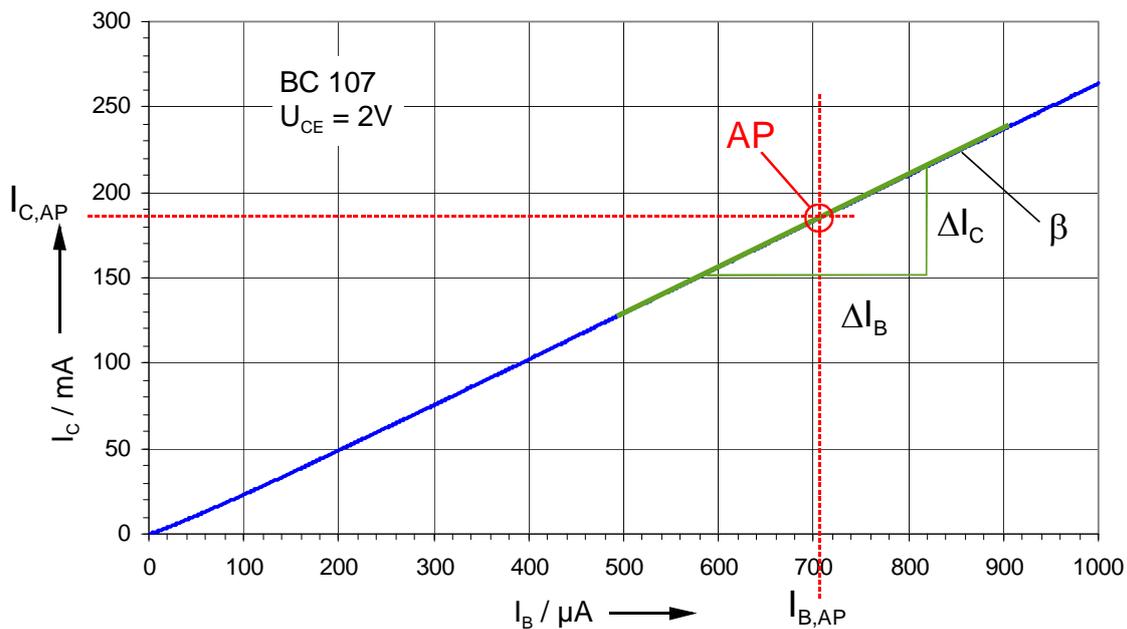


Abb. 4.31 Übertragungskennlinie $I_C = f(I_B)$ und Bestimmung der Kleinsignalstromverstärkung β

Wie Abb. 4.31 zeigt, ist die Übertragungskennlinie nahezu linear, d.h. die quasistationäre Kleinsignal-Stromverstärkung β entspricht näherungsweise der DC-Stromverstärkung B aus Gl. (4.42):

$$\beta = \left. \frac{i_c}{i_b} \right|_{AP} \approx \left. \frac{I_C}{I_B} \right|_{AP} = B \quad (4.56)$$

Größere Abweichungen treten erst bei höheren Strömen und höheren Frequenzen auf (s. Vorlesung "Elektronische Schaltungen")

4.4.3 Die Steilheit

Der Kollektorstrom hängt sowohl vom Basisstrom als auch von der Basis-Emitter Spannung ab.

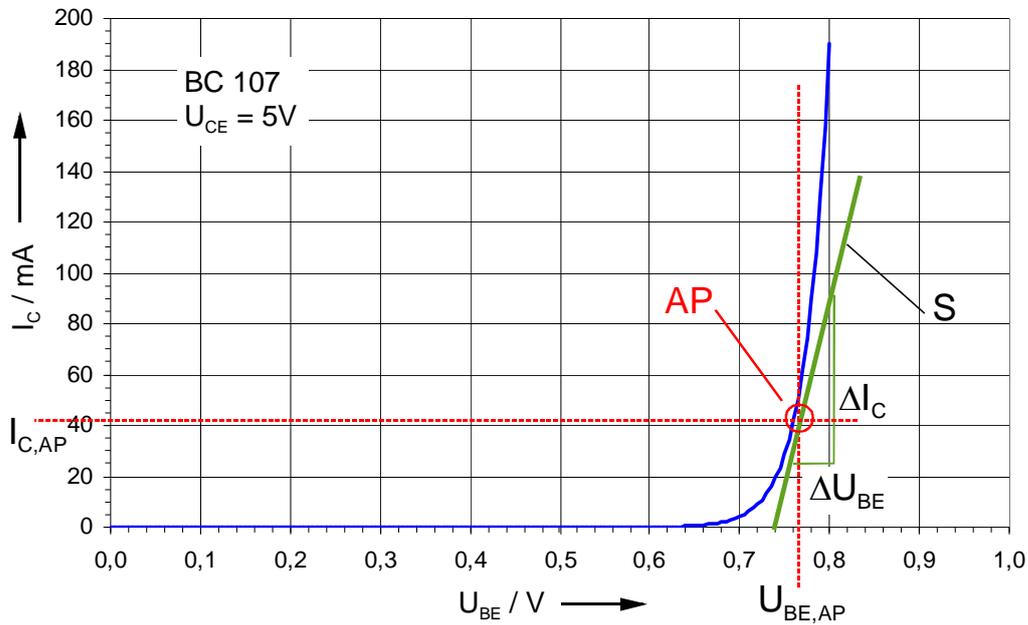


Abb. 4.32 Übertragungskennlinie $I_C = f(U_{BE})$ und Bestimmung der Steilheit Da für die Stromverstärkung $B \approx \text{konst.}$ gilt, ergibt sich der Eingangskennlinie entsprechend ein exponentieller Verlauf (Abb. 4.32):

$$I_C = B \cdot I_B = B \cdot I_{B0} \cdot \left(e^{\frac{U_{BE}}{U_{th}}} - 1 \right) \approx B \cdot I_{B0} \cdot e^{\frac{U_{BE}}{U_{th}}} \quad \text{für } U_{BE} \gg U_{th} \quad (4.57)$$

Die Steilheit lässt sich wieder aus der Steigung der Kennlinie ermitteln:

$$S = \left. \frac{i_c}{u_{be}} \right|_{AP} = \left. \frac{dI_C}{dU_{BE}} \right|_{AP} \quad (4.58)$$

$$\Rightarrow S = \frac{dI_C}{dU_{BE}} = \frac{B}{U_{th}} \cdot I_{B0} \cdot e^{\frac{U_{BE}}{U_{th}}} = \frac{B}{U_{th}} \cdot I_B \quad (4.59)$$

Mit Gl. (4.52):

$$u_{be} = i_b \cdot r_{be}$$

und Gl. (4.56):

$$i_c = \beta \cdot i_b$$

in Gl. (4.68) eingesetzt, folgt:

$$S = \frac{\beta}{r_{be}} \quad (4.60)$$

D.h., sind zwei der Größen gegeben, lässt sich die dritte direkt daraus berechnen.

4.4.4 Der Ausgangswiderstand

Je nach Arbeitspunkt hängt der Kollektorstrom nicht nur von den Eingangsgrößen I_B und U_{BE} , sondern auch von der Ausgangsspannung U_{CE} ab.

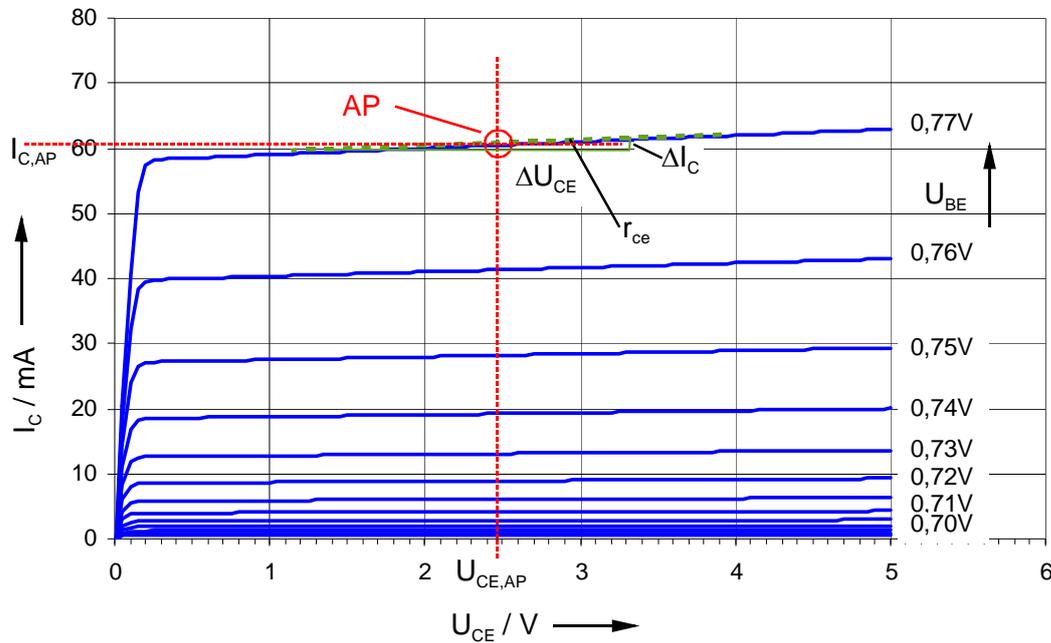


Abb. 4.33 Ausgangskennlinienfeld $I_C = f(U_{CE}, U_{BE})$ und Bestimmung des Ausgangswiderstands r_{ce}

Die Steigung der Ausgangskennlinie bei konstantem Eingangsparameter (I_B oder U_{BE}) wird, wie in Kap. 2.4.2.3 bereits dargestellt, durch den Ausgangswiderstand beschrieben:

$$\frac{1}{r_{ce}} = g_{ce} = \left. \frac{i_c}{u_{ce}} \right|_{AP} = \left. \frac{dI_C}{dU_{CE}} \right|_{AP} \quad (4.61)$$

Als Verstärker wird der Transistor meist im aktiven Bereich betrieben. Da sich der Ausgangsstrom dort nur wenig ändert, ist der Ausgangswiderstand relativ hoch. Im Idealfall gilt $r_{ce} \rightarrow \infty$ bzw. $g_{ce} \rightarrow 0$. Abhängig vom in Kap. 4.6 vorgestellten Early-Effekt kann aber auch dort der Strom I_C mit wachsender Kollektor-Emitter Spannung U_{CE} ansteigen.

4.4.5 Das quasistationäre Kleinsignal-Ersatzschaltbild

Die gesamte Kollektorstromänderungen setzt sich also wieder aus 2 Anteilen zusammen:

$$i_c = i_{c1} + i_{c2} \tag{4.62}$$

mit

$$i_{c1} = S \cdot u_{be} \tag{4.63}$$

$$i_{c2} = \frac{u_{ce}}{r_{ce}} \tag{4.64}$$

Damit ergibt sich das in Abb. 4.34 dargestellte Kleinsignalerersatzschaltbild.

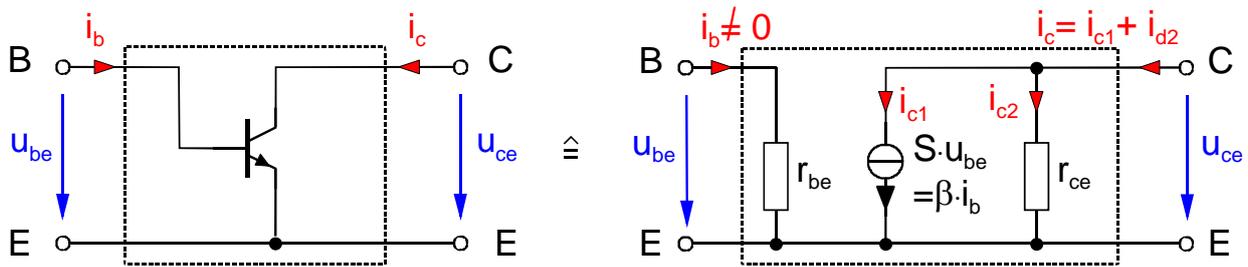


Abb. 4.34 Quasistationäres Kleinsignalerersatzschaltbild für den Bipolartransistor in Emitter-Schaltung

4.5 Der Schaltvorgang im Bipolartransistor

Im Folgenden wird der Schaltvorgang im Bipolartransistor an Hand einer einfachen Transistorschaltung (Abb. 4.35a) untersucht. Abb. 4.35b zeigt die Anordnung der pn-Dioden im Bipolartransistor mit den entsprechenden Spannungsbezugspfeilen.

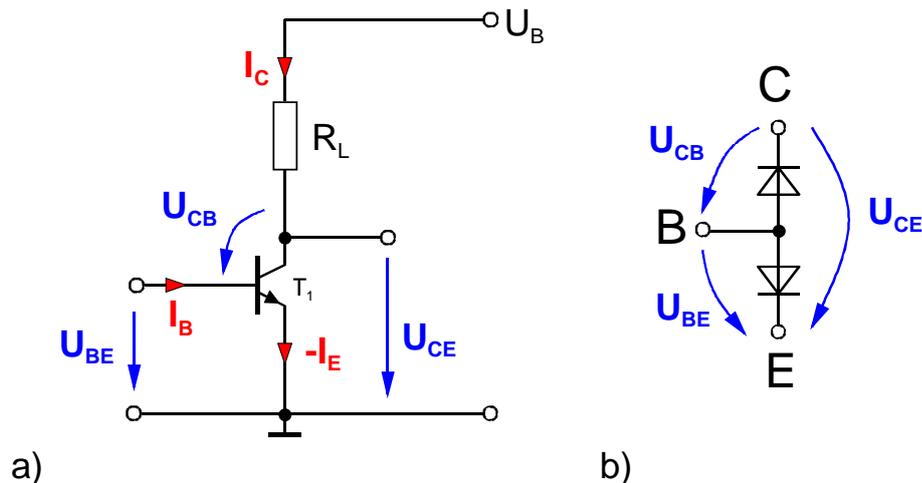


Abb. 4.35 a) Der Bipolartransistor als Inverter

b) stark vereinfachtes Großsignalersatzschaltbild des Bipolartransistors mit den entsprechenden Spannungsbezugspfeilen; die gesteuerten Stromquellen (Ebers-Moll-Modell) sind dabei weggelassen worden.

Der jeweilige Arbeitspunkt des Transistors ergibt sich aus dem Schnittpunkt von Kennlinie und

Arbeitsgerade gemäß:

$$U_B = U_{CE} + I_C \cdot R_L \quad (4.48)$$

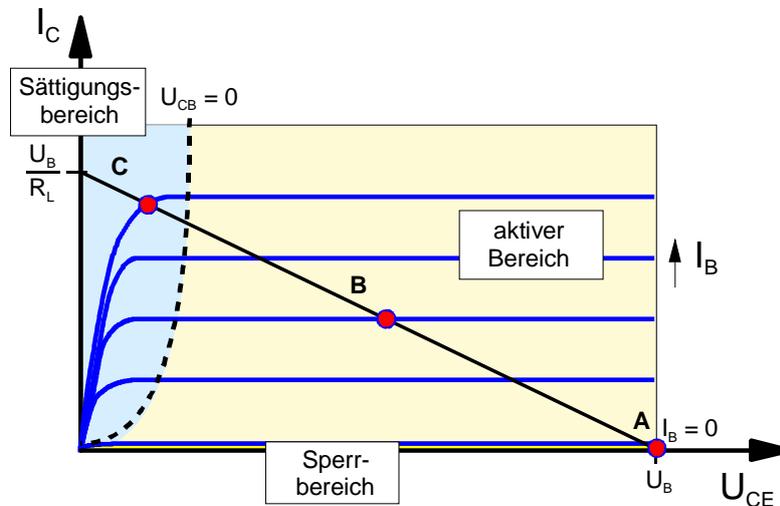


Abb. 4.36 Ausgangskennlinienfeld und die bekannten Arbeitsbereiche eines Bipolartransistors

Wie bereits beschrieben, unterscheidet man prinzipiell drei Arbeitsbereiche:

- Sperrbereich
- aktiver Bereich
- Sättigungsbereich

4.5.1 Sperrbereich

Ist der Transistor ausgeschaltet, d.h. $I_B = 0$, dann gilt auch $I_C \approx 0$. Es fließen nur noch Sperr- bzw. Restströme. Der Arbeitspunkt liegt im Sperrbereich. Damit fällt keine Spannung über R_L ab und am Ausgang der Schaltung liegt $U_{CE} = U_B = U_{CE}^H$ ("H":="High"). In diesem Arbeitsbereich (Punkt A in Abb. 4.36) sind Kollektor-Basis- und Emitter-Basis-Diode gesperrt ($U_{BE} \leq 0$, $U_{CB} > 0$). Die Minoritätsverteilung in der Basis entspricht qualitativ der Kurve in Abb. 4.37. Der Transistor als Schalter ist offen (gesperrt, AUS). Er bildet einen sehr großen Widerstand R_{sperr} .

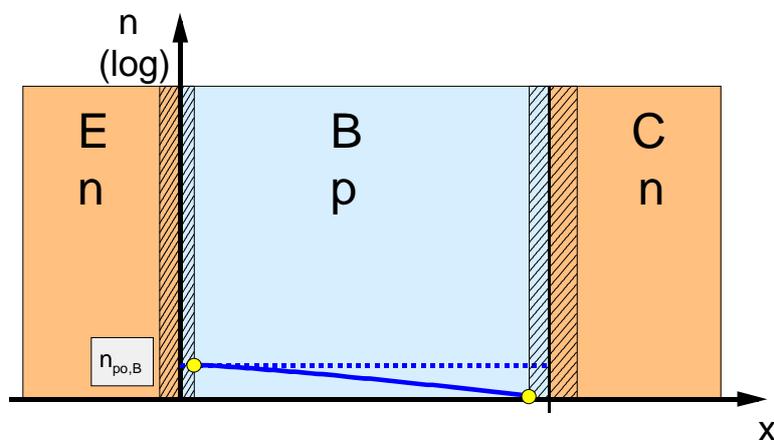


Abb. 4.37 Qualitative Darstellung der Verteilung der Minoritätsladungsträger in der Basis für einen Arbeitspunkt im Sperrbereich (Punkt A in Abb. 4.35)

4.5.2 Aktiver Bereich

Wird die Basis-Emitter-Spannung U_{BE} bzw. der Basisstrom I_B erhöht, so wird die Basis-Emitter-Diode in Flussrichtung durchgesteuert, während die Kollektor-Basis-Diode noch in Sperrrichtung betrieben wird. Der Arbeitspunkt verläuft entlang der Arbeitsgeraden nach Gl. (4.48) im aktiven Bereich (Punkt **B** in Abb. 4.36).

Der Spannung U_{BE} entsprechend steigt die Konzentration der Minoritäten in der **Basis an der Grenze zum Emittor** an, während diese an der Grenze zum Kollektor weiterhin sehr gering bleibt. Damit liegt die in Abb. 4.38 dargestellte, prinzipielle Minoritätsträgerkonzentration in der Basis vor.

Da der Kollektorstrom I_C vom Konzentrationsgradienten entsprechend:

$$I_C \propto \frac{dn}{dx} \tag{4.65}$$

abhängt, und dieser mit wachsender Basis-Emitter-Spannung U_{BE} bzw. ansteigendem Strom I_B ansteigt (Abb. 4.36), wächst auch der Kollektorstrom an, wie aus Abb. 4.36 hervorgeht.

Die Fläche unter dem Verlauf von $n(x)$ entspricht dem steuernden Basisstrom I_B und damit auch der in der Basis gespeicherten Ladung Q_B .

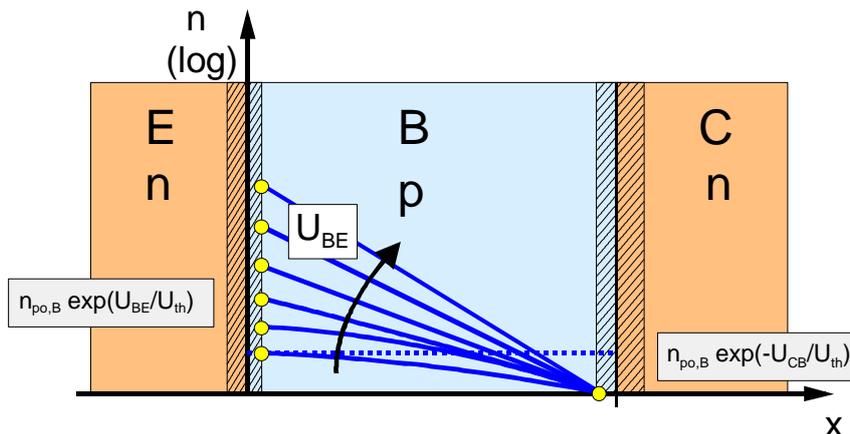


Abb. 4.38 Qualitative Verteilung der Minoritätsträgerkonzentration in der Basis für einen Arbeitspunkt im aktiven Bereich (Punkt B in Abb. 4.36)

4.5.3 Sättigungsbereich

Wird der Strom $I_B > 0$ noch weiter erhöht, verschiebt sich der Arbeitspunkt entlang der Arbeitsgeraden bis in den Sättigungsbereich (Punkt **C** in Abb. 4.36).

Der Transistor als Schalter ist geschlossen (durchgeschaltet, EIN) und bildet nur noch einen geringen Widerstand, den Durchlasswiderstand. Am Ausgang liegt nur eine relativ kleine Spannung U_{CE}^L ("L":="Low").

Die Grenze zwischen Sättigungs- und aktivem Bereich ist gegeben durch die Kurve $I_C(U_{CE})$ bei $U_{CB} = 0$ (s. Abb. 4.36). Im Sättigungsbereich sind nun beide Dioden in Flussrichtung geschaltet ($U_{BE} > 0, U_{CB} < 0$). Da die Spannung U_{CE} in diesem Arbeitspunkt nahezu konstant bleibt, steigt mit der Spannung U_{BE} die Konzentration der Minoritäten zur Basis hin, gleichzeitig aber auch zum Kollektor hin an. Die Minoritätsträgerverteilung entspricht qualitativ der Kurve in Abb. 4.39.

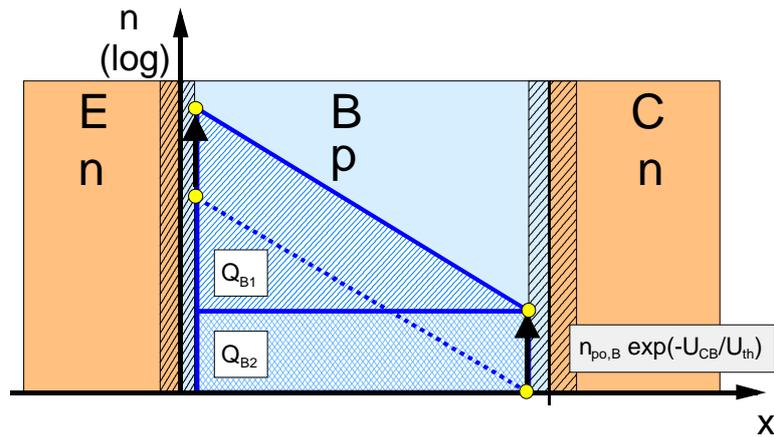


Abb. 4.39 Qualitative Verteilung der Minoritätsträger in der Basis für einen Arbeitspunkt im Sättigungsbereich (Punkt C in Abb. 4.36)

Der Konzentrationsgradient dn/dx nimmt nicht mehr weiter zu, d.h. der Kollektorstrom I_C steigt trotz Erhöhung des Basisstroms I_B bzw. der Basis-Emitter-Spannung U_{BE} nicht mehr weiter an.

Dagegen wächst die Fläche unterhalb der Kurve $n(x)$ und damit auch die in der Basis gespeicherte Ladung $Q_{B1} + Q_{B2}$, d.h. sowohl die Ladung, die dem Basisstrom entspricht, der benötigt wird, den Transistor durchzuschalten (Q_{B1}) als auch überschüssige Ladung Q_{B2} , die nichts zum Aufbau des Konzentrationsgradienten und somit zum Kollektorstrom beiträgt.

Da in diesem Fall die Basis mit Minoritäten gesättigt ist, spricht man bei Arbeitspunkten, die in diesem Bereich des Ausgangskennlinienfelds liegen, vom **Sättigungsbereich**. Der Wert von U_{CE} fällt in diesem Arbeitsbereich bis zu einer Kollektor-Emitter-Restspannung $U_{CE,rest}$ ab (Abb. 4.40).

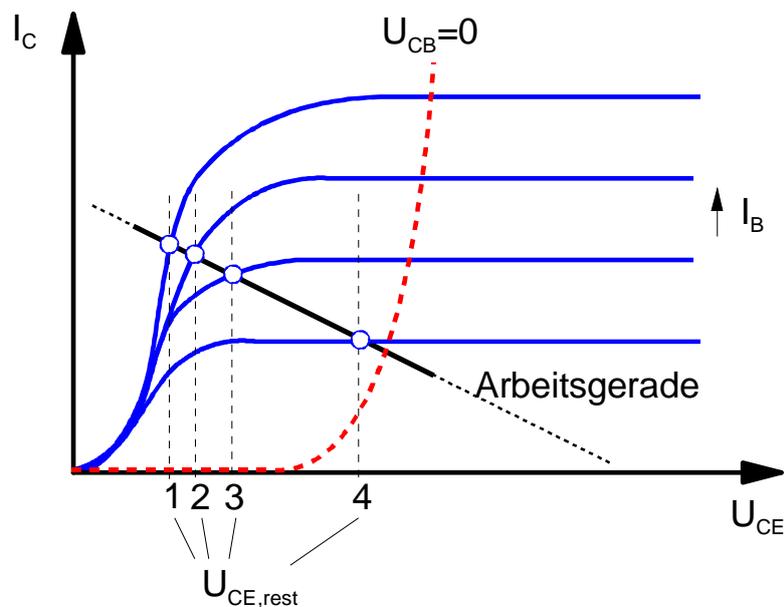


Abb. 4.40 Abhängigkeit der Restspannung $U_{CE,rest}$ vom Arbeitspunkt im Sättigungsbereich

Wird der Transistor als Schalter eingesetzt, so sollte die Kollektor-Emitter-Restspannung $U_{CE,rest}$ möglichst klein sein. Dazu wird der Transistor übersteuert, d.h. der Steuerstrom I_B wird auf ein Mehrfaches des erforderlichen Wertes festgelegt:

$$I_{BÜ} = \frac{I_C}{B^*} \gg \frac{I_C}{B} \Rightarrow B^* \ll B \quad (4.66)$$

Eine Übersteuerung ist allerdings nur bis zu einem bestimmten Maße sinnvoll, da der Grad der Übersteuerung auch die Schaltzeit des Transistors, und hier insbesondere die Ausschaltzeit, beeinflusst. Dies liegt daran, dass im Sättigungsbereich überschüssige Ladung (Q_{B2} in Abb. 4.39) gespeichert wird, die erst abgebaut werden muss, bevor der Transistor ausgeschaltet werden kann. Üblich ist eine Übersteuerung derart, dass ein Kollektor-Basis-Stromverhältnis von

$$B^* = \frac{I_C}{I_B} \approx 10 \quad (4.67)$$

eingestellt wird.

Bei vorgegebenem U_B und R_L wird die Kollektor-Emitter-Spannung U_{CE} über dem Basisstrom I_B aufgetragen. Dabei kann an einigen markanten Punkten das Verhältnis von I_C zu I_B eingetragen werden (s. Abb. 4.41).

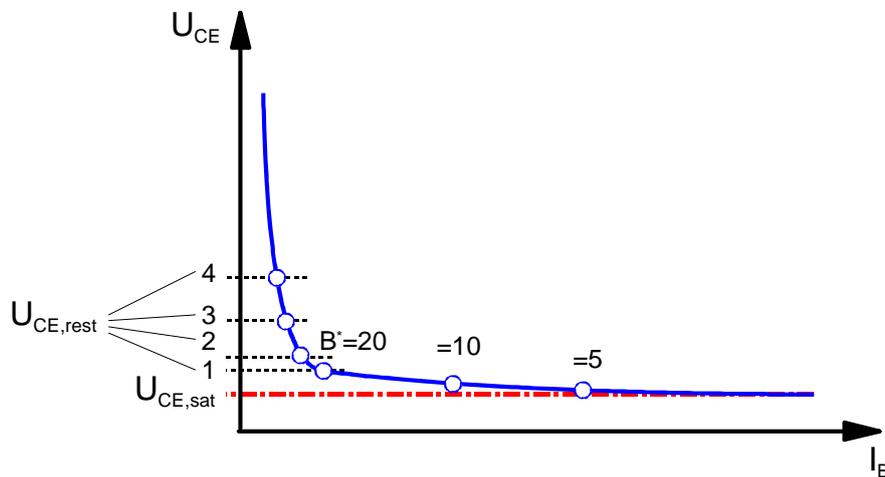


Abb. 4.41 Abhängigkeit der Kollektor-Emitter-Spannung $U_{CE,rest}$ vom Basisstrom

Aus dem Kurvenverlauf ist deutlich erkennbar, dass mit dem Stromverhältnis $B^* \approx 10$ der Bereich der Sättigung erreicht wird und die Spannung U_{CE} fast auf die minimal mögliche Sättigungsspannung $U_{CE,sat}$ abgefallen ist. Eine weitere Erhöhung des Basisstromes hat nur noch eine unwesentliche Verringerung der Restspannung zur Folge. Der erreichbare Durchlasswiderstand für $B^* = 10$ hängt vom Transistortyp und von der Höhe des Kollektorstromes $I_{C,max}$ ab. Für Kleinsignaltransistoren liegen die Werte für Silizium bei einigen Ohm ($\approx 10 \text{ Ohm}$), für Germanium geringfügig tiefer ($\approx 5 \text{ Ohm}$). Für Leistungstransistoren werden noch geringere Widerstandswerte erreicht (Silizium ca. 1 Ohm , Germanium ca. $0,1 \text{ Ohm}$); dies gilt jedoch nicht für Hochspannungstransistoren. Aufgrund der invertierenden Funktion der Schaltung:

$$U_{BE} < U_F \text{ bzw. } I_B = 0 \Rightarrow I_C = 0 \text{ und } U_{CE} = U_B = U_{CE}^H \text{ bzw.}$$

$$U_{BE} \gg U_F \text{ bzw. } I_B \gg 0 \Rightarrow I_C \gg 0 \text{ und } U_{CE} = U_{CE,rest} = U_{CE}^L$$

wird diese Schaltung auch **Inverter** genannt.

4.6 Grenzwerte

Bei Aufbau und Dimensionierung von Schaltungen ist es absolut wichtig, die maximal zulässigen Grenzwerte eines Bauelements zu kennen und zu beachten. Aus diesem Grund soll an dieser Stelle kurz auf die wichtigsten Grenzdaten eingegangen werden.

4.6.1 Durchbruchverhalten

4.6.1.1 Durchbruch 1. Art

Als Dreischichtbauelement besteht der Bipolartransistor aus 2 pn-Dioden. Im aktiven Bereich (Normalbetrieb) ist die Basis-Kollektor Diode gesperrt, die Basis-Emitter Diode leitend. Im Betrieb als Schalter kommt es jedoch auch vor, dass die Basis-Emitter Diode gesperrt ist. Bei der Dimensionierung von Transistorschaltungen müssen daher die maximal zulässigen Sperrspannungen bekannt sein.

Der in den Dioden auftretende Durchbruch bei Überschreiten der Sperrspannungen auftretende Zener- oder Lawinendurchbruch (s. Kap. 3.2.2) wird beim Transistor als Durchbruch 1. Art bezeichnet.

4.6.1.1.1 Basis-Emitter Sperrspannung

Die Sperrspannung U_{EB0} der Basis-Emitter Diode bei offenem Kollektor ($I_C = 0$) ist wegen der hohen Dotierung des Emitters relativ klein und liegt im Bereich $5V < U_{EB0} < 10V$. Als Durchbruchmechanismus wirkt der Zener-Effekt, d.h. eine Überschreitung der Sperrspannung führt nicht unbedingt zur Zerstörung des Transistors, solange die maximal zulässige Verlustleistung der Diode nicht überschritten wird.

4.6.1.1.2 Kollektor-Basis Sperrspannung

Die Sperrspannung U_{CB0} der Kollektor- Basis Diode bei offenem Emitter ($I_E = 0$) ist wegen der verhältnismäßig niedrigen Dotierung des Kollektors relativ hoch. Als Durchbruchmechanismus ist der Lawineneffekt wirksam. Die maximale Sperrspannung kann bei Hochspannungstransistoren bis zu 1000V betragen.

4.6.1.1.3 Kollektor-Emitter Sperrspannung

Die wichtigste Sperrspannung des Transistors ist die Kollektor-Emitter Sperrspannung U_{CE0} bei offener Basis ($I_B = 0$). Sie bestimmt die höchste zulässige Kollektor-Emitter Spannung U_{CE} des Transistors in Emitterschaltung.

4.6.1.2 Durchbruchverhalten 2. Art

Besonders bei Leistungstransistoren, die mit großem Kollektorstrom und hoher Kollektor-Emitter Spannung betrieben werden, beobachtet man, dass diese gelegentlich zerstört werden, auch wenn ihre maximal zulässige Verlustleistung nicht überschritten wird. Hierbei handelt es sich um einen speziellen thermischen Durchbruch. Um ihn von dem schon bekannte Zener- oder Lawinendurchbruch, dem Durchbruch 1. Art zu unterscheiden, spricht man in diesem Fall vom Durchbruch 2. Art.

Dabei tritt durch inhomogene Stromverteilung eine lokale Übertemperatur auf, die zu einem lokalen Schmelzen und damit zur Zerstörung des Transistors führt. Es findet zunächst ein normaler Durchbruch statt, in dessen Verlauf eine Einschnürung der Strompfade auftritt. Anschließend kommt es dann zum Durchbruch 2. Art.

4.6.2 Grenzströme

Bei den Grenzströmen wird zwischen maximalen Dauerströmen und maximalen Spitzenwerten unterschieden. Die maximalen Spitzenwerte gelten für den gepulsten Betrieb mit vorgegebener Pulsdauer und Wiederholrate. Sie sind um den Faktor 1,2 - 2 größer als die Dauerströme.

4.6.3 Sperrströme

Für die Emitter- und die Kollektor-Basis Diode sind im Datenblatt neben den Durchbruchspannungen auch die Sperrströme I_{EB0} und I_{CB0} angegeben, die bei einer Spannung unterhalb der jeweiligen Durchbruchspannung gemessen werden.

4.6.4 Maximale Verlustleistung

Eine besonders wichtige limitierende Größe ist die maximal zulässige Verlustleistung. Die Verlustleistung entspricht der im Transistor in Wärme umgesetzten Leistung:

$$P_V = U_{CE} \cdot I_C + U_{BE} \cdot I_B \approx U_{CE} \cdot I_C \quad (4.68)$$

Sie entsteht im Wesentlichen in der Sperrschicht der Basis-Kollektor Diode. Die Temperatur der Sperrschicht erhöht sich auf einen Wert, der von dem Temperaturgefälle zwischen Sperrschicht und Transistorgehäuse und der thermischen Leitfähigkeit bestimmt wird.

Die Temperatur der Sperrschicht darf einen materialabhängigen Grenzwert, bei Silizium $T \approx 175^\circ C$, nicht überschreiten. Aus Sicherheitsgründen wird in der Praxis mit einem Grenzwert von $T \approx 150^\circ C$ gerechnet. Die maximale Verlustleistung, bei der dieser Grenzwert erreicht wird, hängt vom Aufbau des Transistors und der Montage ab. Sie wird im Datenblatt mit P_{tot} bezeichnet.

Abb. 4.42 zeigt im Prinzip den gültigen Arbeitsbereich eines Bipolartransistors. Als Beispiel ist in Abb. 4.43 das Datenblatt des Kleinsignaltransistors BC 107 mit den zugehörigen Grenzwerten gegeben.

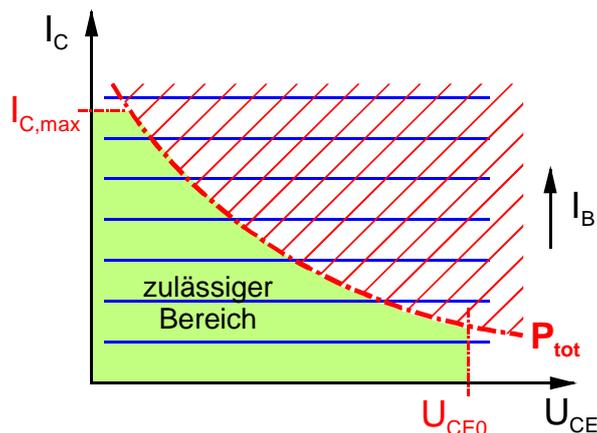


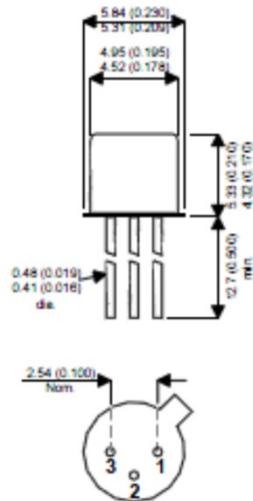
Abb. 4.42 Erlaubter Arbeitsbereich eines Bipolartransistors. Die Grenzen für den Durchbruch 2. Art sind nicht eingetragen



BC107
BC108
BC109

MECHANICAL DATA

Dimensions in mm (inches)



TO-18 METAL PACKAGE

Underside View

PIN 1 – Emitter PIN 2 – Base PIN 3 – Collector

GENERAL PURPOSE SMALL SIGNAL NPN BIPOLAR TRANSISTOR

FEATURES

- SILICON NPN
- HERMETICALLY SEALED TO18
- SCREENING OPTIONS AVAILABLE

ABSOLUTE MAXIMUM RATINGS ($T_A = 25^\circ\text{C}$ unless otherwise stated)

V_{CBO}	Collector – Base Continuous Voltage	BC017	50V
		BC108, BC109	30V
V_{CEO}	Collector – Emitter Continuous Voltage With Zero Base Current	BC107	45V
		BC108, BC109	20V
V_{CES}	Collector – Emitter Continuous Voltage With Base Shortcircuited to Emitter	BC107	50V
		BC108, BC109	30V
V_{EBO}	Emitter – Base Continuous Voltage Reverse Voltage	BC107	6V
		BC108, BC109	5V
I_C	Continuous Collector Current		100mA
I_{CM}	Peak Collector Current		200mA
P_{tot}	Power Dissipation @ $T_{amb} = 25^\circ\text{C}$		300mW
T_{amb}	Ambient Operating Temperature Range		-65 to +175°C
T_{stg}	Storage Temperature Range		-65 to +175°C



BC107
BC108
BC109

ELECTRICAL CHARACTERISTICS ($T_A = 25^\circ\text{C}$ unless otherwise stated)

Parameter	Test Conditions	Min.	Typ.	Max.	Unit
$I_{CBO(1)}$ Collector-Base Leakage Current	$V_{CB} = 45\text{V}$ BC107 $V_{CB} = 25\text{V}$ BC108, BC109			15 15	nA
$I_{CBO(1)}$ Collector-Emitter Leakage Current @ $T_{amb} = 125^\circ\text{C}$	$V_{CB} = 45\text{V}$ BC107 $V_{CB} = 25\text{V}$ BC108, BC109			4 4	μA
I_{EBO} Emitter Cut-off Current	$V_{EB} = 4\text{V}$ $I_C = 0$			1	μA
h_{21E} Static Forward Current Transfer Ratio	$V_{CE} = 5\text{V}$ $I_C = 2\text{mA}$ Group A BC107, BC108 Group B All Types Group C BC108, BC109	110 180 380		220 460 800	
	BC107 BC108 BC109	110 110 180		460 800 800	
V_{BE} Base – Emitter Breakdown	$V_{CE} = 5\text{V}$ $I_C = 2\text{mA}$			0.7	V
$V_{BE(sat)(1)}$ Base – Emitter Saturation Voltage	$I_B = 0.5\text{mA}$ $I_C = 10\text{mA}$			0.83	V
$V_{CE(sat)(1)}$ Collector – Emitter Saturation Voltage	$I_B = 0.5\text{mA}$ $I_C = 10\text{mA}$			0.25	V
f_T Transition Frequency	$V_{CE} = 5\text{V}$ $I_C = 10\text{mA}$ $f = 100\text{MHz}$	150			MHz
F Noise Factor	$V_{CE} = 5\text{V}$ $I_C = 0.2\text{mA}$ $R = 2\text{k}\Omega$ $f = 1\text{kHz}$ $\Delta F = 200\text{Hz}$ BC109 BC107, BC108			4 10	dB
h_{21e} Small Signal Forward Current Transfer Ratio	$V_{CE} = 5\text{V}$ $I_C = 2\text{mA}$ $f = 100\text{kHz}$ Group A BC107, BC108 Group B All Types Group C BC108, BC109	125 240 450		260 500 900	
	BC107 BC108 BC109	125 125 240		500 900 900	
h_{11e} Common Emitter Input Impedance	$V_{CE} = 5\text{V}$ $I_C = 2\text{mA}$ $f = 1\text{kHz}$ Group A BC107, BC108 Group B All Types Group C BC108, BC109	1.6 3.2 6.0		4.5 8.5 15	$\text{k}\Omega$
h_{22e} Common Emitter Output Admittance	$V_{CE} = 5\text{V}$ $I_C = 2\text{mA}$ $f = 1\text{kHz}$ Group A BC107, BC108 Group B All Types Group C BC108, BC109			30 60 110	μS
C_{22b} Common Base Output Capacitance	$V_{CB} = 10\text{V}$ $f = 1\text{MHz}$			6	pF
$R_{th(j-amb)}$ Thermal Resistance: Junction to Ambient				500	$^\circ\text{C/W}$

Semelab plc. Telephone +44(0)1455 556565. Fax +44(0)1455 552612. e-mail sales@semelab.co.uk
Website <http://www.semelab.co.uk>

3/99

Abb. 4.43b Datenblatt des BC 107 der Fa. SEMELAB

Tab. 4.1 gibt einen groben Überblick über die Klassifizierung der Bipolartransistoren mit typischen Werten.

	Einsatz für		U_{CE}	I_C	B	f_T	I_B
1	NF-Kleinsignal	nnp	1 ... 20 V	0,1 ... 10 mA	250 ^{*)}	300 MHz	0,4 μ A ... 40 μ A
2	NF-Kleinsignal	pnp	-1 ... -10 V	-0,1 ... -10 mA	200	150 MHz	0,5 μ A ... 50 μ A
3	mittlere Leistung	nnp	10 ... 40 V	5 ... 250 mA	120	130 MHz	40 μ A ... 2mA
4	mittlere Leistung	pnp	-10 ... -40 V	-5 ... -250 mA	80	90 MHz	60 μ A ... 3mA
5	große Leistung	nnp	10 ... 80 V	0,1 ... 5A	100	100 MHz	1mA ... 50mA
6	sehr große Leistung	nnp	10 ... 60 V	1 ... 15A	20 ... 40	1 MHz	25mA ... 370mA
7	Leistungsdarlington	nnp	3 ... 100 V	0,05 ... 10A	800	7 MHz	60 μ A ... 15mA
8	Hochspannung (Videostufen)	nnp	10 ... 400 V	3 ... 100 mA	30	100 MHz	100 μ A ... 4mA
9	Hochfrequenz	nnp	1 ... 25 V	0,1 ... 20 mA	~ 100	> 1GHz	>1 μ A ... 100 μ A

^{*)}: Für Kleinsignaltransistoren gilt meist $B=250$.

In Gruppen A,B und C unterteilt gilt näherungsweise:

für Gruppe A: $B= 150 \dots 300$

für Gruppe B: $B= 250 \dots 600$

für Gruppe C: $B= 450 \dots 900$

Tab. 4.1 Klassifizierung der Bipolartransistoren mit typischen Werten.

4.7 Der Heterostrukturbipolartransistor (HBT)

Nach Gl. (4.45) gilt für den Bipolartransistor bei $U_{CB} = 0$ (Sättigungsgrenze)

$$I_C = B \cdot I_B$$

mit

$$B = \frac{I_{CE}}{I_{EE} - I_{CE}} \quad (4.68)$$

wobei

$$I_{CE} = qU_{th}A_C \frac{\mu_n n_{p0B}}{W_B} = C \frac{\mu_n}{W_B N_{AB}} \quad (4.69)$$

und

$$I_{EE} = qU_{th}A_E n_i^2 \left(\frac{\mu_n}{W_B N_{AB}} + \frac{\mu_p}{l_{nE} N_{DE}} \right) = C \cdot \left(\frac{\mu_n}{W_B N_{AB}} + \frac{\mu_p}{l_{nE} N_{DE}} \right) \quad (4.70)$$

unter der Voraussetzung, dass $A_E = A_C$ ist.

Eingesetzt in Gl. (4.68) folgt damit:

$$B = \frac{l_{nE}}{W_B} \cdot \frac{\mu_n}{\mu_p} \cdot \frac{N_{DE}}{N_{AB}} \quad (4.71)$$

Ziel bei der Entwicklung und Optimierung von Transistoren ist in der Regel, eine möglichst große Stromverstärkung zu erreichen. Dazu bieten sich entsprechend Gl. (4.71) verschiedene Möglichkeiten an:

a) Basisweite W_B möglichst klein:

Damit lassen sich vor allem auch gute HF-Eigenschaften erzielen. Nachteil ist die damit geringere Durchbruchspannung des Bauelements. Es wird dadurch wesentlich empfindlicher in Bezug auf höhere Spannungen.

b) Verhältnis von μ_n / μ_p möglichst groß:

Die Beweglichkeiten sind eng verknüpft mit den Dotierstoffkonzentrationen in den einzelnen Bereichen (Abb. 0.15). Je höher die Dotierung wird, umso kleiner ist die Beweglichkeit. Relativ hohe Dotierungen sind aber für gute Eigenschaften der Bipolartransistoren notwendig (siehe auch c)), was zwangsläufig zu einer Reduzierung der Beweglichkeit führt.

Andererseits bewirkt eine niedrige Löcherbeweglichkeit hohe Bahnwiderstände, die wiederum die HF-Eigenschaften des Transistors negativ beeinflussen.

c) Verhältnis von N_{DE} / N_{AB} möglichst groß:

Eine hohe Emittierdotierung führt wiederum zu niedrigen Beweglichkeiten. Eine niedrige Basisdotierung hat in der ohnehin schon sehr dünnen Basisschicht bei gleichzeitig relativ niedriger Löcherbeweglichkeit einen hohen parasitären Zuleitungs-(Bahn-)widerstand zur Folge, was letztendlich wieder zu schlechten HF-Eigenschaften führt.

Für eine hohe Stromverstärkung $B = I_C / I_B$ ist ein großer Elektronenstrom I_n bei gleichzeitig

möglichst geringem Löcherstrom I_p notwendig.

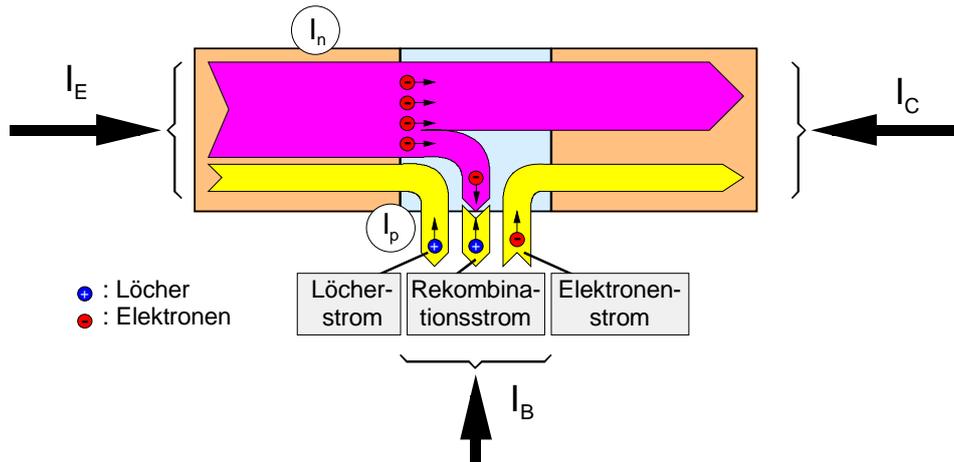


Abb. 4.44 Stromanteile im Bipolartransistor (entsprechend Abb. 4.2)

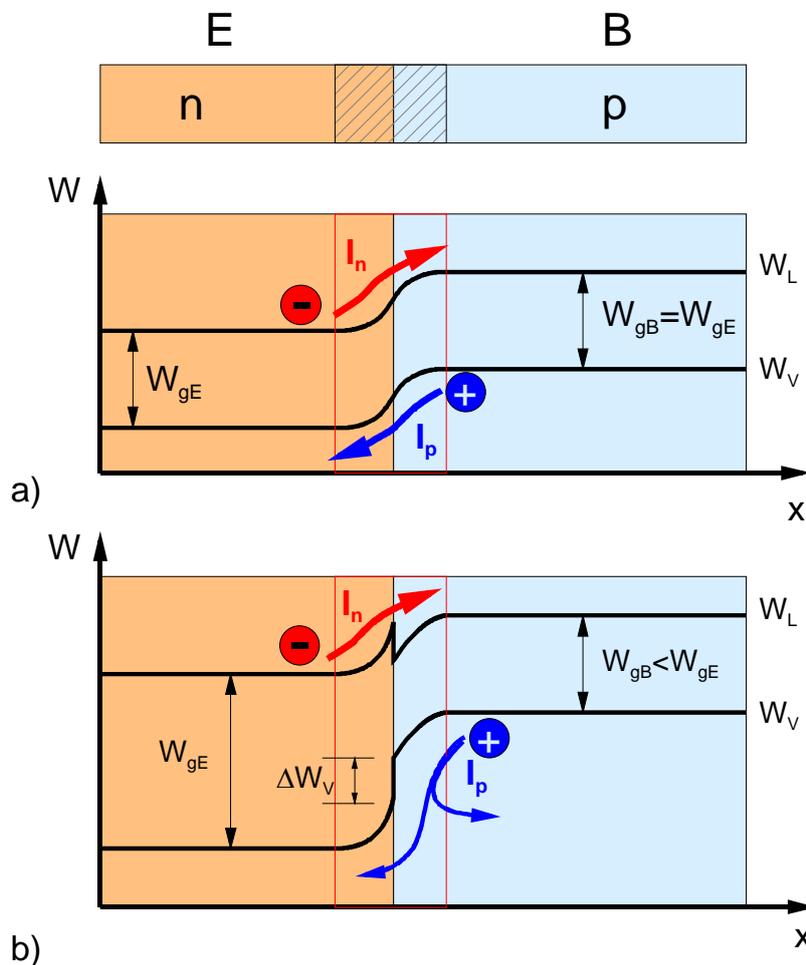


Abb. 4.45 a) homogener pn-Übergang (Emitter-Basis)
 b) heterogener pn-Übergang (Emitter-Basis)

Dieser Löcherstrom lässt sich dadurch reduzieren, dass statt eines homogenen Emitter-Basis-Übergangs mit $W_{gE} = W_{gB}$, bei dem in beiden Bereichen das gleiche Halbleitermaterial verwendet wird, für den Emitter ein Halbleitermaterial mit größerem Bandabstand $W_{gE} > W_{gB}$ verwendet wird (Abb. 4.45). Es ergibt sich ein Heteroübergang, wie er bereits aus Kap. 2.3 in Zusammenhang

mit dem Heterostruktur-FET (HFET) bekannt ist. Dementsprechend wird ein Bipolartransistor mit Heteroübergang zwischen Emitter und Basis als **Heterostruktur-Bipolartransistor** (HBT) bezeichnet. Während der Stromfluss beim HFET parallel zum Übergang (*interface*) stattfindet, fließt der Strom beim HBT jedoch senkrecht zum Heteroübergang.

Bei geeigneter Komposition macht sich die Barriere für Elektronen im Leitungsband kaum bemerkbar, während die Barriere im Valenzband (Abb. 4.45b) zu einer deutlichen Reduzierung des Löcherstroms zwischen Basis und Emitter führt.

Unter der Voraussetzung, dass:

$$\Delta W_V \approx \Delta W_g = W_{gE} - W_{gB} \tag{4.72}$$

ist, lässt sich nach [23] zeigen, dass gilt:

$$\left(\frac{I_n}{I_p} \right)_{heterogen} = \left(\frac{I_n}{I_p} \right)_{homogen} \cdot e^{\left(\frac{\Delta W_V}{kT} \right)} \tag{4.73}$$

Damit folgt für die Stromverstärkung

$$B \propto \frac{\mu_n}{\mu_p} \cdot \frac{N_{DE}}{N_{AB}} \cdot e^{\left(\frac{\Delta W_V}{kT} \right)} \tag{4.74}$$

Damit ist es jetzt möglich, die Basisdotierung N_{AB} deutlich zu erhöhen, ohne die Stromverstärkung zu reduzieren. Somit lässt sich gleichzeitig auch der Basis-Bahnwiderstand deutlich reduzieren und die HF-Eigenschaften des Transistors entsprechend verbessern. Darauf wird noch zu einem späteren Zeitpunkt weiter eingegangen.

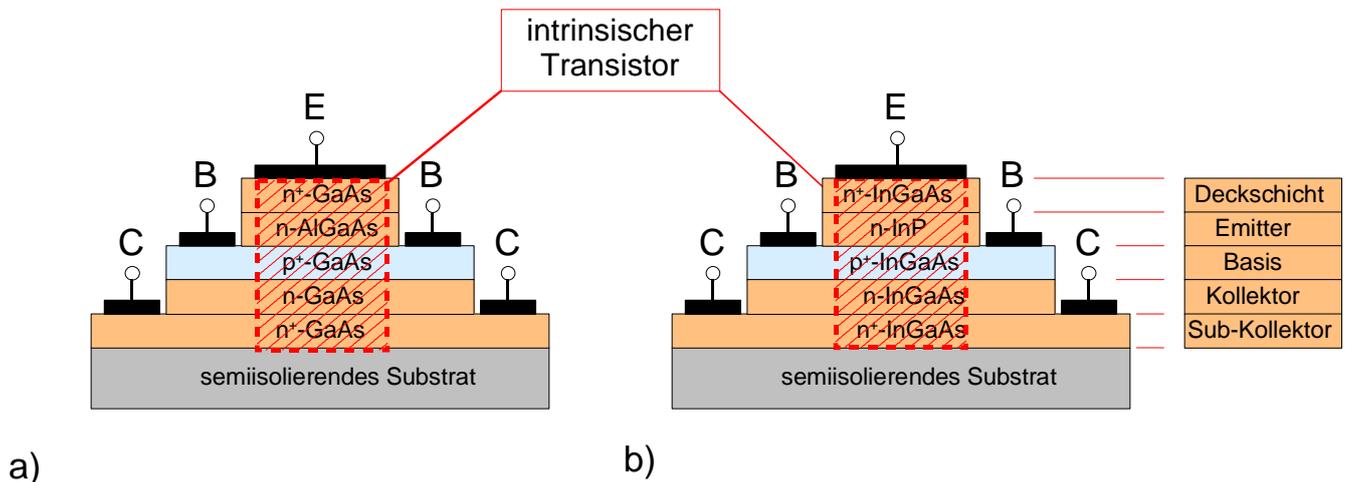


Abb. 4.46 Mögliche Materialsysteme für HBT

a) AlGaAs/GaAs-HBT und

b) InP/InGaAs-HBT

Als Materialsystem bietet sich das weite Spektrum der III/V-Halbleiter an, da sich die gewünschten Bandstrukturen über geeignete Verfahren zur Schichtherstellung bei gleichzeitiger Anpassung des Kristallgitters bereitstellen lassen und somit die Realisierung verspannungsfreier Schichten möglich ist. Während beim Bipolartransistor auf Siliziumsubstrat ein planares Design möglich ist (Kap. 4.2),

erfordert der Heterostrukturbipolartransistor, basierend auf den Materialsystemen der III. und V. Hauptgruppe des Periodensystems, eine sogenannte 3-Mesa-Struktur. Abb. 4.46 zeigt zwei HBT's im Querschnitt mit möglichen III/V-Kompositionen (AlGaAs/GaAs bzw. InP/InGaAs).

Diese Mesa-Technologie ist notwendig, da die einzelnen Schichten sukzessive auf dem Ausgangssubstrat (*Wafer* aus GaAs oder InP) mit Hilfe der Metallorganischen Gasphasenepitaxie (MOVPE) oder Molekularstrahlepitaxie (MBE) aufgewachsen werden. Die zusätzliche hoch dotierte Emitter- (Deckschicht) und Kollektorschicht (Sub-Kollektor) werden benötigt, um einen niederohmigen Zuleitungswiderstand (Bahnwiderstand) von den äußeren Kontakten zum intrinsischen Transistor (Abb. 4.46) sowie einen niederohmigen Metall-Halbleiter Übergang zu ermöglichen.

Abb. 4.47 zeigt Rasterelektronenmikroskop (REM) Aufnahmen von zwei realisierten HBT. Für weitere Details sei auf [23] und auf die Veranstaltungen "*Halbleitertechnologie 1*" bzw. "*Halbleitertechnologie 2*" verwiesen.

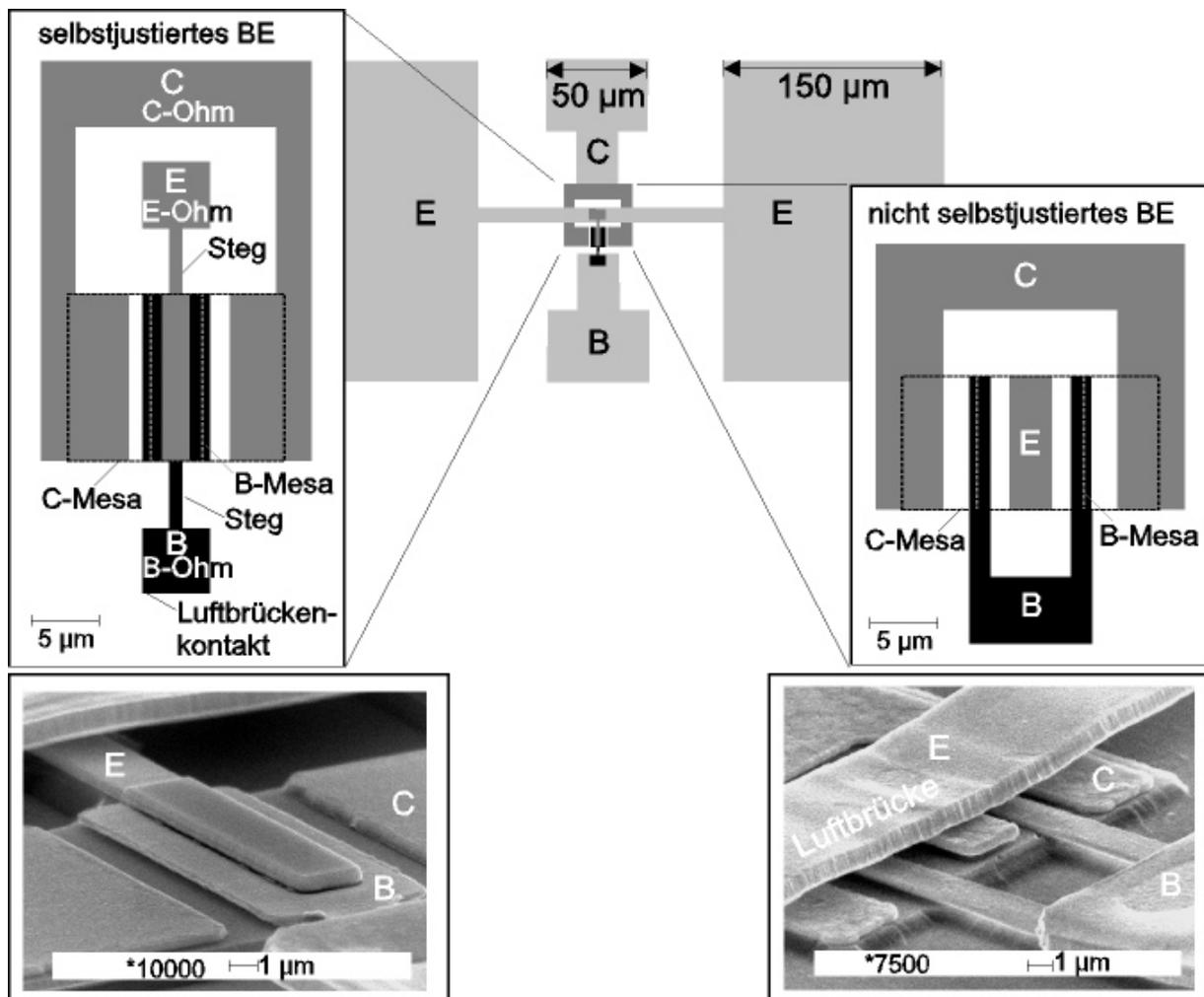


Abb. 4.47 Design und REM-Bilder eines links: selbstjustierten HBT (links, $A_E = 2 \cdot 10 \mu\text{m}^2$) und eines rechts: nicht-selbstjustierten HBT (rechts, $A_E = 3 \cdot 10 \mu\text{m}^2$) [23]

Stichwortverzeichnis:

- aktiver Bereich 4-27
- Ausgangswiderstand 4-41
- Basis-Emitter 4-47
- Basis-Schaltung 4-15
 - Ausgangskennlinienfeld 4-17
 - Eingangskennlinie 4-15
- Bipolartransistor 4-2
 - Basis 4-2
 - Basis-Emitter Sperrspannung 4-47
 - Basis-Schaltung 4-15
 - Durchbruch 1. Art 4-47
 - Durchbruchverhalten 4-47
 - Durchbruchverhalten 2. Art 4-47
 - Emitter 4-2
 - Emitter-Schaltung 4-23
 - Emitterstrom 4-11
 - Funktionsweise 4-2
 - Grenzströme 4-48
 - Kennlinienfelder 4-15
 - Kollektor 4-2
 - Kollektor-Basis Sperrspannung 4-47
 - Kollektor-Emitter Sperrspannung 4-47
 - Kollektorstrom 4-11
 - Maximale Verlustleistung 4-48
 - Normalbetrieb 4-14
 - nnp-Transistor 4-2
 - pnp-Transistor 4-2
 - Restspannung 4-45
 - Sättigungsbereich 4-44
 - Schaltvorgang 4-42
 - Sperrbereich 4-43
 - Sperrströme 4-48
 - Technologie 4-4
 - Übersteuerung 4-46
- DC Stromverstärkung 4-39
- differentielle Stromverstärkung 4-39
- Durchbruch 1. Art 4-47
- Durchbruchverhalten 4-47
- Durchbruchverhalten 4-47
- Early-Effekt 4-35
- Ebers-Moll Gleichungen 4-12
- Ebers-Moll-Gleichungen 4-8
- Emitter-Schaltung
 - Ausgangswiderstand 4-41
 - DC-Stromverstärkung 4-39
 - differentielle Stromverstärkung 4-39
 - Eingangswiderstand 4-37
 - Kleinsignalverhalten 4-37
 - quasistationäres Kleinsignal-
Ersatzschaltbild 4-42
- Emitter-Schaltung 4-23
 - Ausgangskennlinienfeld 4-24
 - Eingangskennlinie 4-23
- Emitter-Schaltung
 - Steilheit 4-40
- Epitaxie-Planar-Transistor 4-5
- Grenzströme 4-48
- Gummel-Plot 4-35
- Heterostruktur-Bipolartransistor (HBT) 4-7,
4-52
- Injektionsmodell 4-32
- Inverter 4-46
- Ionenimplantation 4-5
- ISOPLANAR-Transistor 4-6
- Kollektor-Basis 4-47
- Kollektor-Emitter 4-47
- Maximale 4-48
- Normalbetrieb 4-14, 4-17
- nnp-Transistor 4-2
- Punchthrough 4-36
- Sperrströme 4-48
- Steilheit 4-40
- Transportmodell 4-32

Kapitel 5

Weitere Bipolare Bauelemente

5 Weitere Bipolare Bauelemente	5-2
5.1 Thyristoren	5-2
5.1.1 Aufbau und Wirkungsweise	5-2
5.1.2 Technologie von Thyristoren	5-8
5.1.2.1 Thyristor mit MESAstruktur	5-8
5.1.2.2 Thyristor mit Isolationsdiffusion und geätzten Isolationsgräben	5-9
5.2 Sonderformen	5-10
5.2.1 Rückwärts leitender Thyristor (RCT)	5-10
5.2.2 DIAC	5-11
5.2.3 TRIAC	5-12
5.2.4 Abschaltbare Thyristoren (GTO)	5-12
5.3 Permeable Base Transistor (PBT)	5-14
5.4 Insulated-Gate Bipolar Transistor (IGBT)	5-15

5 Weitere Bipolare Bauelemente

5.1 Thyristoren

5.1.1 Aufbau und Wirkungsweise

Thyristoren oder SCR (*semiconductor controlled rectifier*) haben die Funktion von steuerbaren Gleichrichtern: sie ermöglichen den Stromfluss in Durchlassrichtung erst dann, wenn ein definierter Steuerimpuls an der Steuerelektrode, dem Gate, angelegt wird.

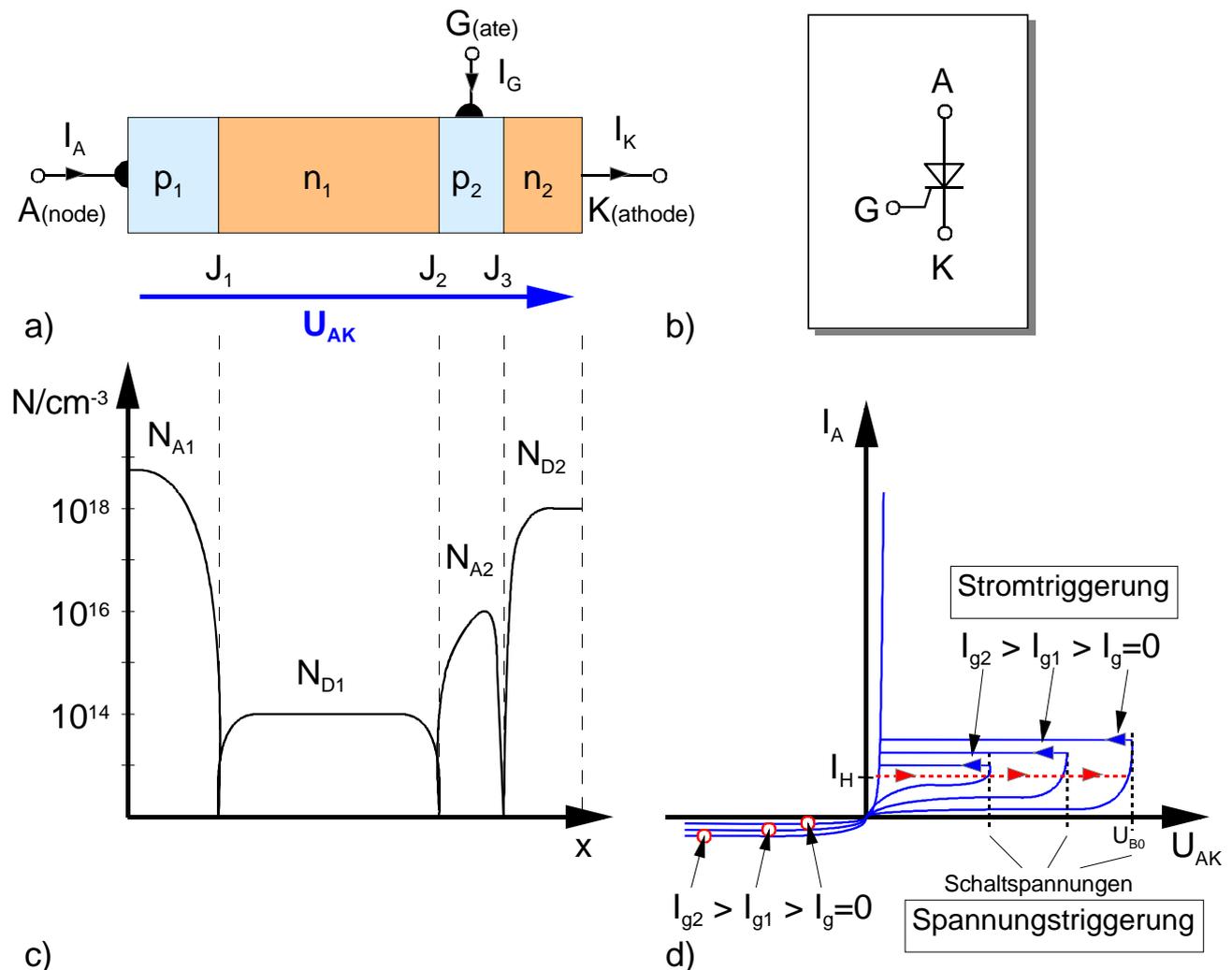


Abb. 5.1 a) Vierschichtstruktur des Thyristors

b) Schaltzeichen

c) Dotierstoffprofil

d) Kennlinie des Thyristors: wird der EIN-Zustand des Thyristors durch Überschreiten der Schaltspannung herbeigeführt, spricht man von **Spannungstriggerung**. Durch Einführen eines Gate-Stroms lässt sich die Schaltspannung reduzieren (**Stromtriggerung**) [17,20]

Diese Funktion wird durch ein Halbleiterbauelement mit vier hintereinandergeschalteten alternierend p- und n-dotierten Halbleiterschichten realisiert (Abb. 5.1).

Wird eine positive Spannung U_{AK} zwischen Anode und Kathode angelegt, sind die beiden äußeren

pn-Übergänge J_1 und J_3 in Flussrichtung, der innere Übergang J_2 in Sperrichtung gepolt. Es kann nur ein geringer Sperrstrom fließen, solange die angelegte Spannung nicht die Schaltspannung U_{B0} überschreitet. Wird die Spannung größer als die Schaltspannung, wechselt das Bauelement abrupt den Zustand, der Strom steigt steil an und die Spannung am Bauelement bricht auf ca. 0,7V zusammen (Abb. 5.1d). Wird der Strom unter den sogenannten Haltestrom reduziert, kehrt der Thyristor in den Sperrzustand zurück.

Dieses Verhalten lässt sich mit Hilfe der Betrachtung zweier kombinierter Bipolartransistoren (Abb. 5.2) erklären:

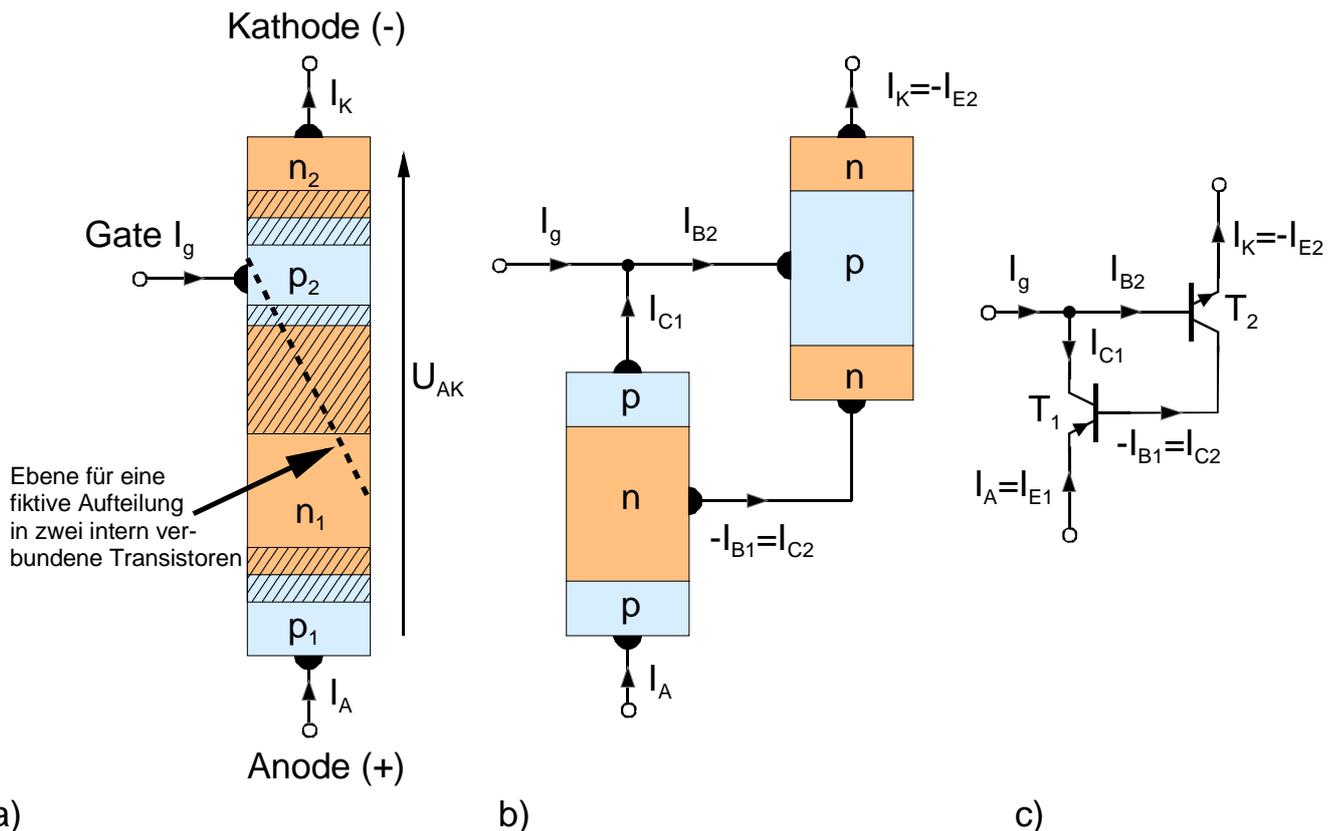


Abb. 5.2 zur Funktionsweise des Thyristors:

a) die Vierschichtstruktur des Thyristors

b,c) zur Analyse des elektrischen Verhaltens wird der Thyristor aus a) fiktiv in zwei Dreischicht-Bauelemente getrennt: ein npn- (T_1) und ein pnp-Transistor (T_2), wobei die Basis von T_1 und der Emitter von T_2 , sowie der Kollektor von T_1 und die Basis von T_2 intern miteinander verbunden sind.

Nach Gl. (4.36) ergibt sich für beide Transistoren im aktiven Bereich (Normalbetrieb):

$$I_{C1} = A_1 \cdot I_{E1} - I_{CB01} \quad (5.1)$$

bzw.

$$I_{C2} = A_2 \cdot I_{E2} - I_{CB02} \quad (5.2)$$

Für den Basisstrom des Transistors T_1 gilt:

$$I_{B1} = -(I_{E1} + I_{C1}) \quad (5.3)$$

und mit Gl. (5.1) sowie:

$$I_A = I_{E1}$$

ergibt sich:

$$I_{B1} = -(I_A + A_1 \cdot I_A - I_{CB01}) = -(1 + A_1) \cdot I_A + I_{CB01} \quad (5.4)$$

Für den Kollektorstrom von T_2 gilt:

$$I_{C2} = A_2 \cdot I_{E2} - I_{CB02} = -A_2 \cdot I_K - I_{CB02} \quad (5.5)$$

Aus

$$-I_{B1} = I_{C2} \quad (5.6)$$

folgt:

$$(1 + A_1) \cdot I_A - I_{CB01} = -A_2 \cdot I_K - I_{CB02} \quad (5.7)$$

und mit den "äußeren" Thyristor-Strömen:

$$I_K = I_A + I_g \quad (5.8)$$

ergibt sich:

$$(1 + A_1) \cdot I_A = -A_2 \cdot (I_A + I_g) + I_{CB01} - I_{CB02}$$

$$\Rightarrow I_A = \frac{-A_2 \cdot I_g + I_{CB01} - I_{CB02}}{(1 + (A_1 + A_2))} \quad (5.9)$$

Im Zähler von Gl. (5.9) stehen nur kleine Ströme, so dass der Thyristorstrom I_A ebenfalls klein ist, sofern

$$A_1 + A_2 \neq -1 \quad (5.10)$$

Diese Bedingung ist entscheidend für das Verhalten des Thyristors. Dabei nutzt man die Abhängigkeit der Stromverstärkung vom Kollektor aus: Bei sehr kleinen Kollektorströmen nimmt der Wert von B ab, damit nach Gl. (4.50) auch A .

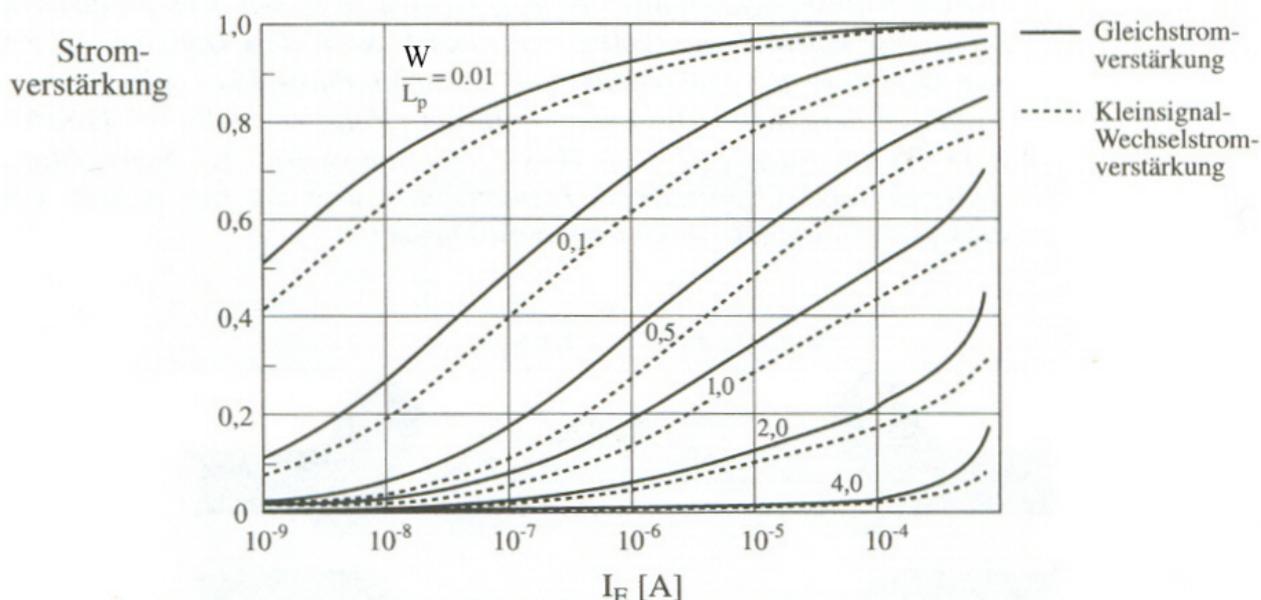


Abb. 5.3 Stromverstärkungen in einem Thyristor in Abhängigkeit vom Emitterstrom mit dem

Verhältnis von Basisweite W und Diffusionslänge L_p [17]

Im Allgemeinen wird davon ausgegangen, dass $A \approx 1$ und dementsprechend $B \gg 1$ ist. Dies gilt allerdings nur bei hinreichend kleinen Basisweiten und großen Kollektorströmen. Diese Bedingung ist in der Regel beim Transistor auch erfüllt.

Beim Thyristor können beide Randbedingungen *nicht* zutreffen, insbesondere kann bei dem beschriebenen Aufbau die Basisweite W_B groß im Vergleich zur Diffusionslänge der Minoritätsträger sein.

In Abb. 5.3 ist die Emitterstromabhängigkeit der A-Stromverstärkung für verschiedene Basisweiten W_B dargestellt.

Bei Anlegen einer äußeren Spannung an den Thyristor fließen zunächst nur sehr kleine Basis- und Sperrströme, unter dieser Voraussetzung gilt zunächst $|A_1 + A_2| < 1$. Wird jetzt der Thyristorstrom durch Vergrößerung der äußeren Spannung erhöht, dann steigen die Stromverstärkungen A_1 und A_2 schnell an, so dass Werte erreicht werden, bei denen schließlich gilt $|A_1 + A_2| = 1$, so dass der Nenner in Gl. (5.9) Null wird. Dieser Zustand kennzeichnet das Durchbrechen des Thyristors. Derselbe Effekt kann durch Einprägen eines Gate-Stromes erfolgen, auch dadurch werden die Stromverstärkungen A_1 und A_2 vergrößert, so dass der Thyristor durchschaltet.

Neben den beschriebenen Verfahren zur Durchschaltung (**Triggerung**) des Thyristors (**Strom- und Spannungstriggerung**) kann eine Vergrößerung der Stromverstärkungen A auch durch einen schnellen **Spannungsanstieg** (dieser verursacht einen kurzen dielektrischen Verschiebungsstrom, der den Emitterstrom vergrößert, **dU/dt -Triggerung**) oder durch lokale optische Bestrahlung (**optische Triggerung**) erreicht werden.

Eine **alternative Erklärung** der Funktionsweise kann mit Hilfe des Bänderdiagramms der Vierschichtdiode gegeben werden (Abb. 5.4). Ohne Vorspannung zeigt das Bänderdiagramm ein konstantes Fermi-Niveau (Abb. 5.4a). Für Spannungen in Vorwärtsrichtung bis zur Schaltspannung U_{B0} sind die beiden äußeren Sperrschichten, J_1 und J_3 , geringfügig in Vorwärtsrichtung gepolt und an der mittleren Sperrschicht J_2 liegt der größte Teil der angelegten Spannung in Rückwärtsrichtung an und das Bänderdiagramm nimmt die in (Abb. 5.4b) gezeigte Form an.

Mit Erreichen der Schaltspannung wird eine große Anzahl an Elektronen und Löchern in der Mitte des Bereiches mit hohem elektrischen Feld durch **Lawinenmechanismen** erzeugt. Die Ladungsträger werden durch das Feld separiert und in die benachbarten p- bzw. n-Zonen abgedrängt. Der plötzliche Ansturm der Ladungen kann nicht sofort durch Injektion in die in Durchlass gepolten Sperrschichten ausgeglichen werden, dadurch wird der n-dotierte Bereich negativ und der p-dotierte positiv geladen. Das Bänderdiagramm (Abb. 5.4c) zeigt, dass deswegen die Spannung an der inneren Sperrschicht reduziert wird und diese sogar in Durchlassrichtung gebracht werden kann.

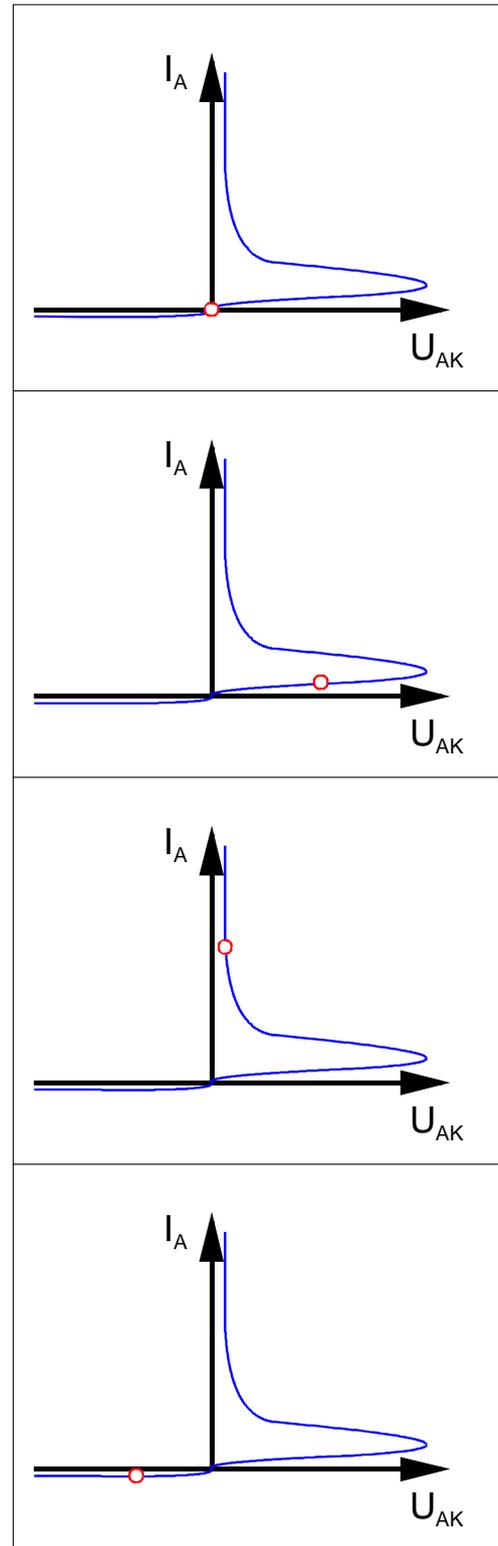
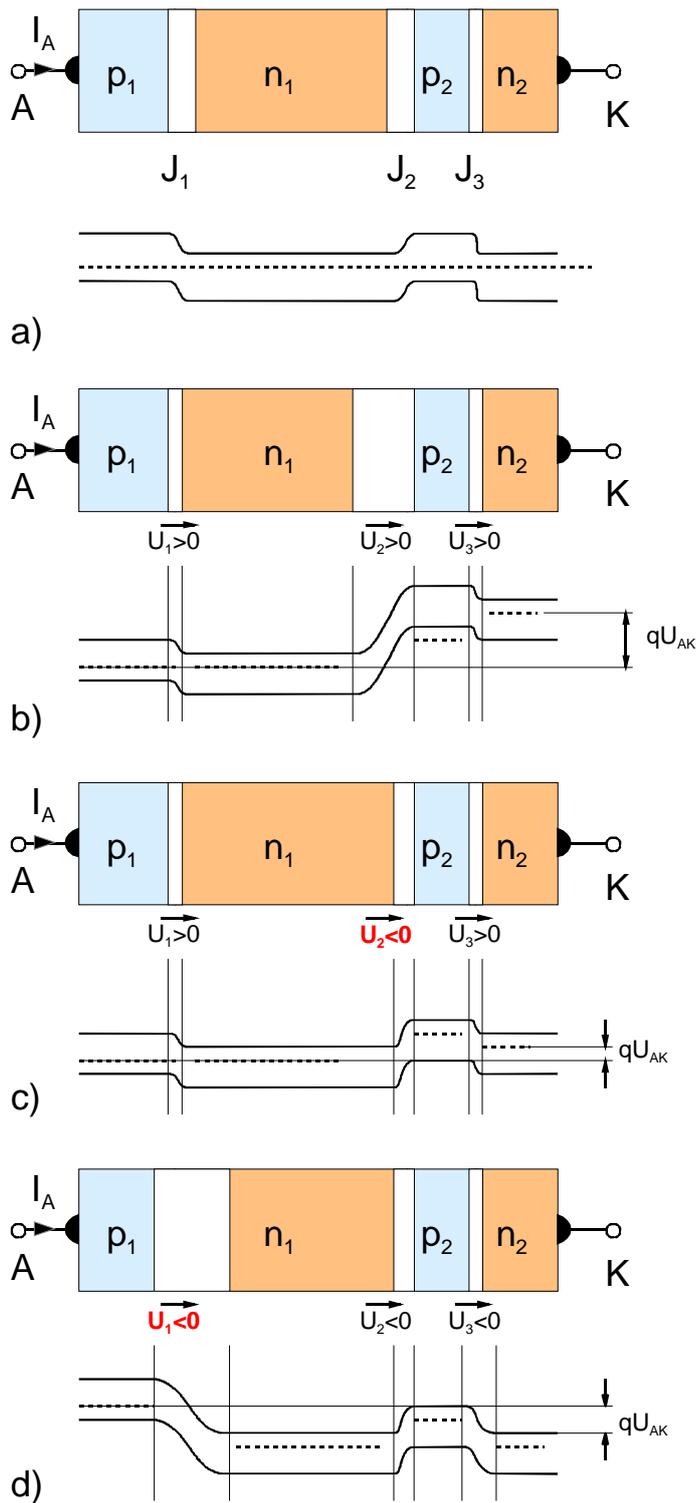


Abb. 5.4 Aufbau und Wirkungsweise von Thyristoren
 a) ohne äußere Spannung
 b) äußere Spannung in Flussrichtung, Thyristor ist noch gesperrt (AUS-Zustand)
 c) äußere Spannung in Flussrichtung, Thyristor ist durchgeschaltet (EIN-Zustand)
 d) äußere Spannung in Sperrrichtung

Durch diese Änderung werden die beiden außenliegenden Sperrschichten weit in Durchlass gebracht. Sie reagieren mit der Injektion großer Mengen von Ladungsträgern in Richtung zur mittleren Zone. Es stellt sich ein neuer stationärer Zustand ein, bei dem alle drei Sperrschichten in Durchlassrichtung gepolt sind, wobei allerdings der Spannungsabfall über der mittleren Sperrschicht gegenüber den anderen umgekehrt gerichtet ist.

Dabei wirken die beiden inneren Bereiche als Basis eines Bipolartransistors (Abb. 5.2c). Von den äußeren Sperrschichten injizierte Ladungsträger diffundieren durch diese Region und werden auf der anderen Seite als Minoritätsträger abgesaugt. Diese Doppeltransistorfunktion hält den Ladungsträgerfluss aufrecht, der durch Lawinenmultiplikation bei der Einschaltspannung begann. Wird der Strom soweit reduziert, dass der Anteil der Injektion durch die Vorwärtsspannung an der mittleren Sperrschicht den abgesaugten Anteil überwiegt, entsteht der in Abb. 5.4d dargestellte, instabile Zustand und die mittlere Sperrschicht gelangt rasch wieder in den Sperrbereich.

Wird der innere p-Bereich an einen dritten Anschluss geführt, lässt sich die Einschaltspannung extern kontrollieren. Spannt man die Gate-Elektrode gegenüber der n-dotierten Kathode positiv vor, wird diese Sperrschicht in Vorwärtsrichtung gepolt. Ein Teil des in die p-Region injizierten Stromes I_g wird von der inneren Sperrschicht aufgenommen und verstärkt den Leckstrom im abgeschalteten Zustand, dadurch wird ein früherer Übergang in den Durchlassbereich bewirkt (s. Kennlinie in Abb. 5.4d).

Mit dem Thyristor ist es also möglich, einen Strom mit Hilfe kleiner Leistungen zu steuern. Die Leistungsverstärkung beträgt etwa 10^4 bis 10^5 und die Betriebsfrequenz einige Hz bis zu 100 kHz (in Spezialfällen auch darüber). Hochleistungsthyristoren sind für Spannungen bis zu 5kV und 2kA ausgelegt. Von diesen Eigenschaften werden die Hauptanwendungsgebiete bestimmt [20]:

- Gleichrichteranlagen mit regelbarer Gleichspannung
- Stromrichterantriebe, insbesondere mit Stromrichtungsumkehr
- Helligkeitssteuerung von Beleuchtungseinrichtungen
- Drehzahlregelung von Drehstrommotoren
- Umrichter zur Frequenzumwandlung (z.B. von 50 Hz auf 16 2/3 Hz)
- Wechselrichter (z.B. für transportable Geräte zur induktiven Erhitzung bei etwa 10 kHz)
- Gleichstromsteuerung, elektronische Schalter- und Speicherelemente in der Digitaltechnik

5.1.2 Technologie von Thyristoren

Im Folgenden werden zwei Herstellungstechnologien von Thyristoren vorgestellt [17].

5.1.2.1 Thyristor mit Mesastruktur

In die beidseitig polierte n-Halbleiterscheibe werden aus einer Gallium/Siliziumquelle Akzeptoren eindiffundiert, so dass auf beiden Seiten p-leitende Bereiche entstehen (Abb. 5.5a). Die Diffusionstiefen ($60\mu\text{m}$) sind bei Leistungsbau-elementen außerordentlich groß, so dass sehr viel Dotierstoff eingebracht werden muss. Um die Diffusionszeiten nicht zu lang werden zu lassen, werden daher hohe Diffusionstemperaturen (1260°C) verwendet. Nach dem Dotierschritt erfolgt eine Oxidation, die Emitterfenster (n^+) werden durch Photolithographie und Oxidätzen hergestellt. Die Emitterdiffusion erfolgt ebenfalls bei 1260°C aus einer Phosphorquelle (PA). Dieser Diffusionsschritt hat gleichzeitig eine **Getterwirkung**: Unerwünschte Fremdatome haben die Eigenschaft, sich in stark phosphordotierten Gebieten anzusammeln [17].

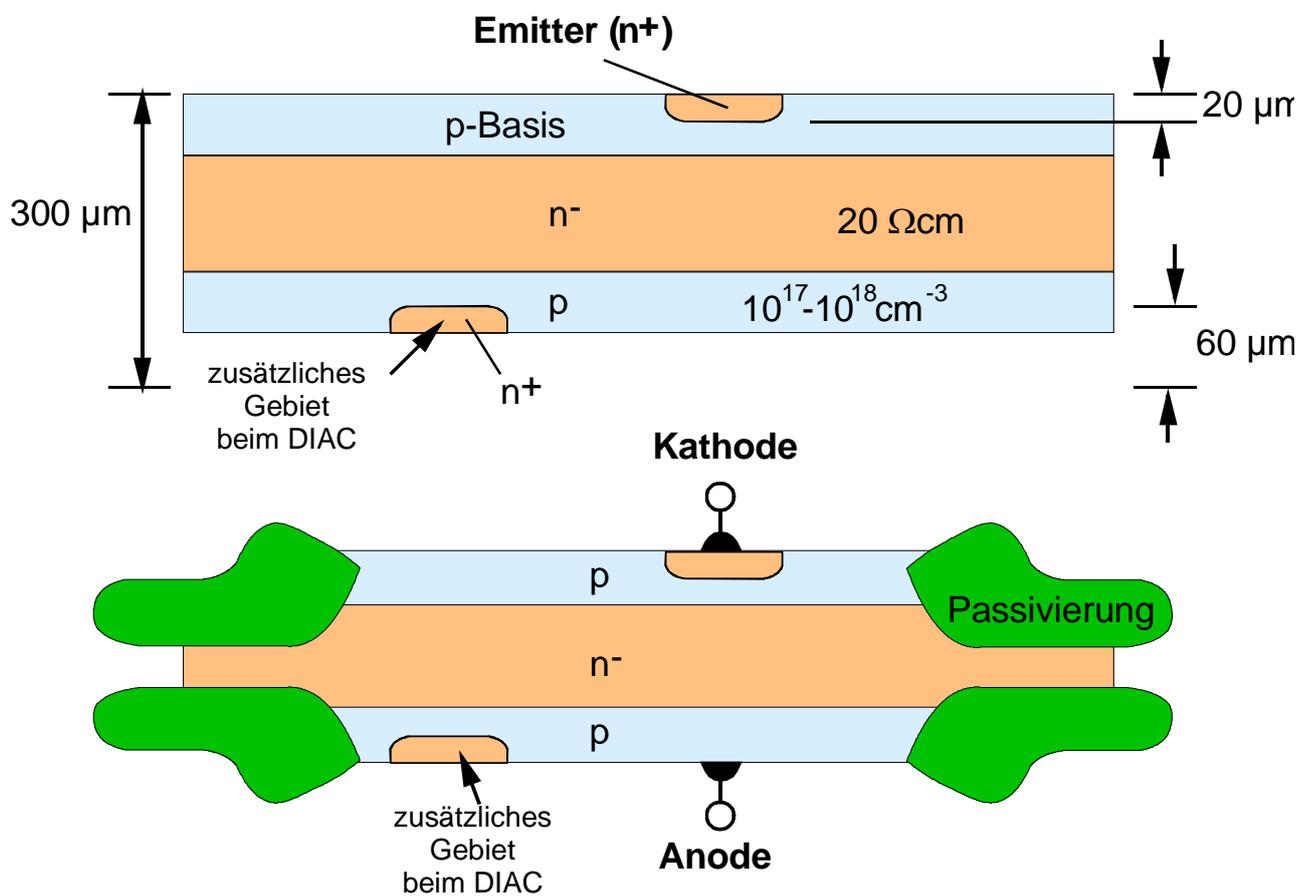


Abb. 5.5 Thyristor mit Mesastruktur [17]

Durch eine tiefe Mesaätzung werden die p-Gebiete benachbarter Thyristoren voneinander getrennt (Abb. 5.5b). Die entstandenen Mesagräben werden mit einer Oxid- oder Glaspassivierung aufgefüllt. Anschließend werden die Bauelemente durch Ritzen und Brechen der Scheibe vereinzelt. Ein Nachteil dieser Struktur ist, dass die n-Elektrode offen liegt zur Ritzkante. Dieses wird bei der Technologie in Abb. 5.6 vermieden.

5.1.2.2 Thyristor mit Isolationsdiffusion und geätzten Isolationsgräben

In die Ausgangsscheibe wird zunächst an den dafür vorgesehenen Stellen eine sehr tiefe p-Diffusion eingeführt, welche die Scheibe praktisch vollständig durchdringt. Nach der Basisdiffusion werden hierdurch Gebiete mit der Ausgangs-n-Dotierung voneinander isoliert. Der Dotierstoff für eine derart intensive Eindiffusion kann z.B. durch aufgedampfte Aluminiumschichten bereitgestellt werden. Die Diffusionszeiten können dabei mehr als 24h bei 1260°C betragen. Anschließend erfolgt wie in Abb. 5.5 Basisdiffusion auf beiden Seiten der Halbleiterscheibe, gefolgt von der n^+ -Emitterdiffusion. Die Abb. 5.5a zeigt zwei Ausführungsformen: Ohne (a1) und mit (a2) Anschluss der Gate-Elektrode [17]

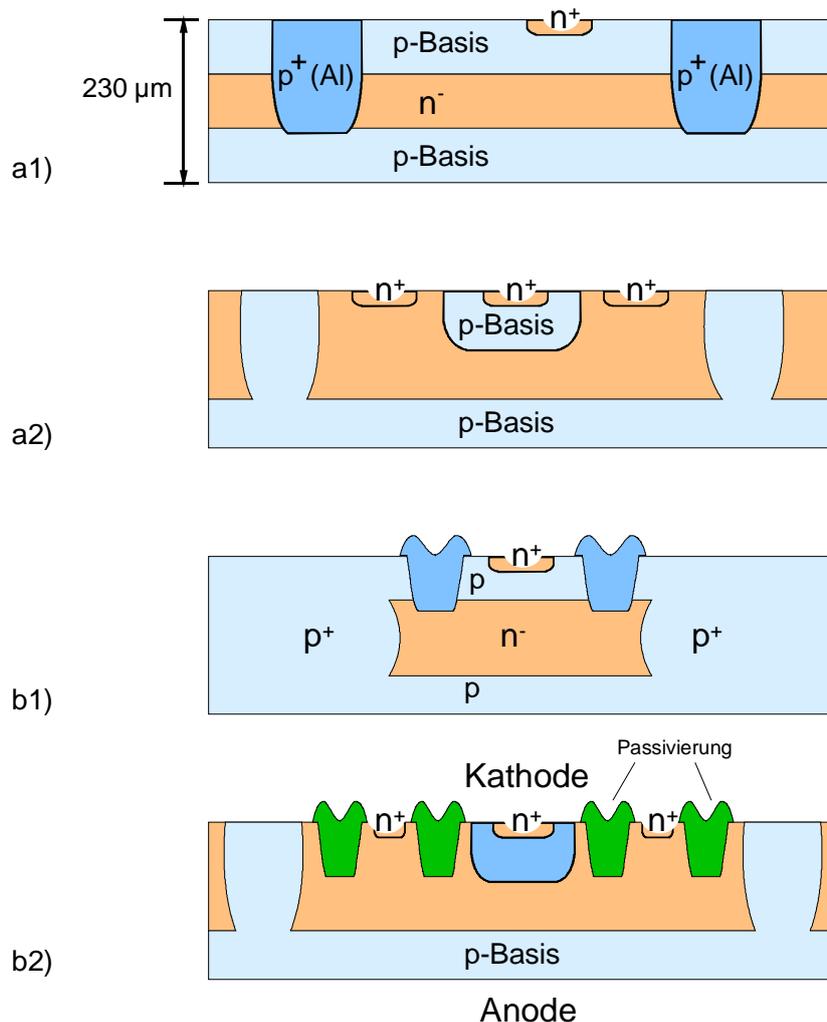


Abb. 5.6 Thyristor mit Isolationsdiffusion und geätzten Isolationsgräben [17]

Die nachfolgende Grabenätzung kann eine unterschiedliche Funktion haben: Bei der Struktur ohne Gate-Anschluss (b1) bewirkt sie eine Trennung (Isolation) der beiden p-Gebiete. Bei der anderen Struktur (b2) sind die p-Gebiete bereits getrennt (Abb. 5.6) durch die Ätzgräben werden aber gekrümmte Bereiche an den Dotierungsübergängen entfernt, die zu Feldstärkeüberhöhungen und damit zu einer geringeren Spannungsbelastbarkeit führen. Dieser Aspekt kann bei Hochspannungsbaulementen (Spannungsbelastbarkeit z.B. mehr als 1000 V) von großer Bedeutung sein. Die Auffüllung der Isolationsgräben mit einer Passivierungsschicht, sowie die Vereinzelung der Bauelemente erfolgt wie in Bild Abb. 5.5 [17].

5.2 Sonderformen

Durch Modifikation der Thyristorstruktur können die Thyristoreigenschaften erheblich verändert werden. Nachfolgend werden drei Bauformen vorgestellt.

5.2.1 Rückwärts leitender Thyristor (RCT)

Abb. 5.7 zeigt einen rückwärts leitenden Thyristor (RCT: *reverse conducting thyristor*): Durch die Anoden- und Kathodenmetallisierung werden jeweils die p- und n-leitenden Schichten auf der Kathoden- und Anodenseite des Thyristors kurzgeschlossen. Ist der entsprechende pn-Übergang in Flussrichtung gepolt, dann wirkt der Kurzschluss wie ein Parallelwiderstand, d.h. ein Großteil des Stroms fließt wegen des geringen Widerstands in Flussrichtung nach wie vor über den pn-Übergang, so dass das Funktionsprinzip des Thyristors weitgehend erhalten bleibt.

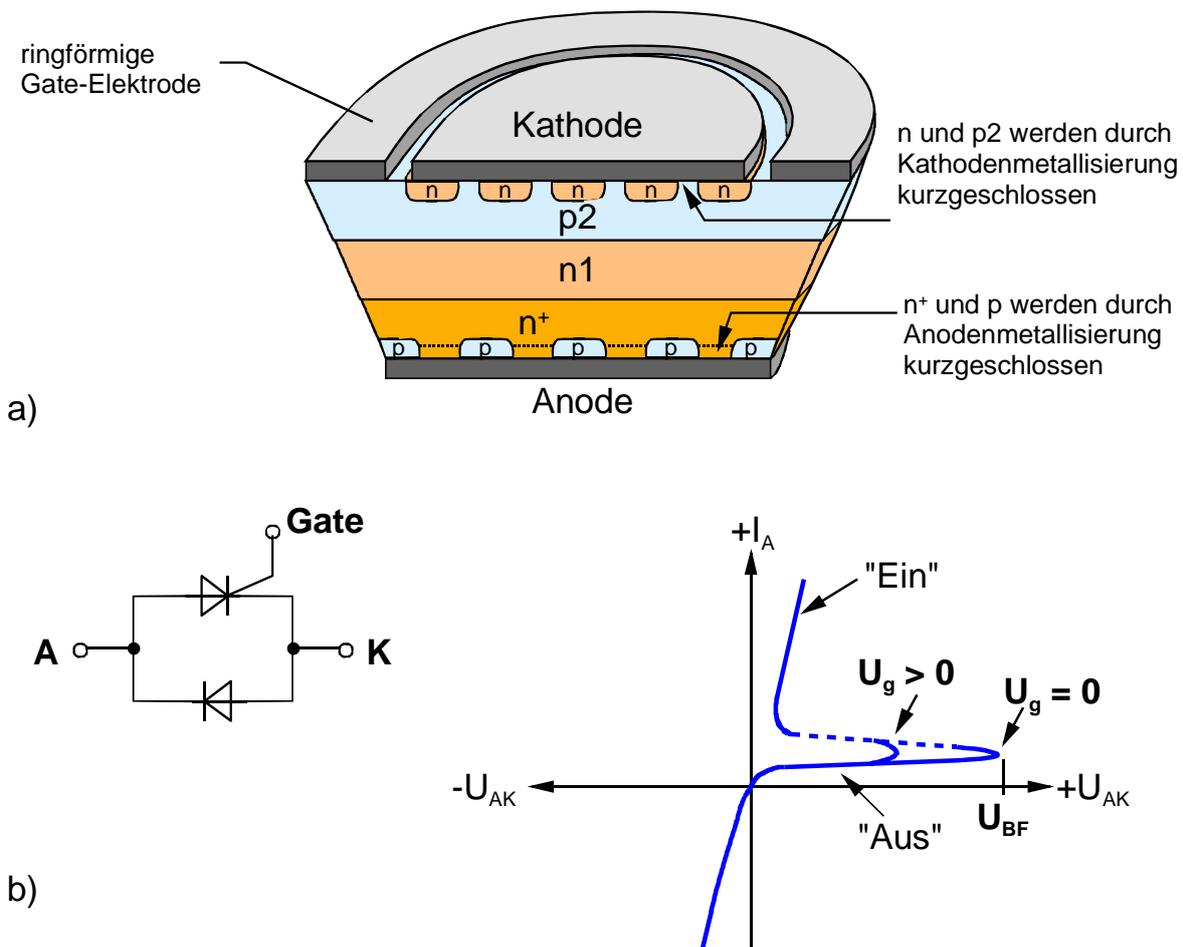


Abb. 5.7 Rückwärts leitender Thyristor (RCT) [17]

a) Aufbau des RCT

b) Kennlinie des RCT

Bei Anliegen von Sperrspannungen werden aber die nichtleitenden pn-Übergänge kurzgeschlossen (Abb. 5.4d), d.h. anstelle eines niedrigen Sperrstroms fließt jetzt der volle Strom über die Kurzschlussbrücken [17].

5.2.2 DIAC

Ein DIAC (*diode ac switch*) besteht aus zwei antiparallel geschalteten gate-losen Thyristoren (auch **Shockley-Dioden** genannt, Abb. 5.8). Jeweils einer von beiden befindet sich in Durchlassrichtung, so dass der Sperrbereich des anderen kurzgeschlossen wird.

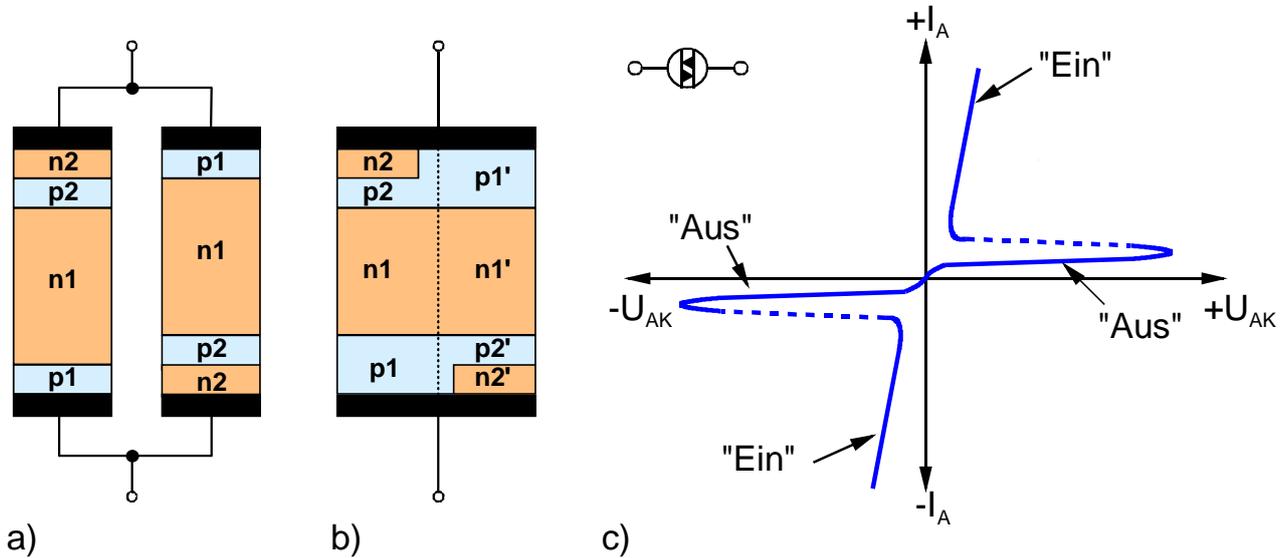


Abb. 5.8 DIAC

- a) Aufbau des DIAC
- b) integrierte Herstellung
- c) Kennlinie

Die beiden Thyristoren nach Abb. 5.8a lassen sich auf einem Halbleiterkristall integriert herstellen (Abb. 5.8b). Abb. 5.8c zeigt die Kennlinie des DIAC.

5.2.3 TRIAC

Ein DIAC mit zusätzlicher Steuermöglichkeit durch Gate-Anschlüsse, wobei diese zu einem einzigen integriert werden, wird als TRIAC (*triode ac switch*) bezeichnet (Abb. 5.9). Dabei kann die Zündung in beiden Schaltrichtungen getriggert werden.

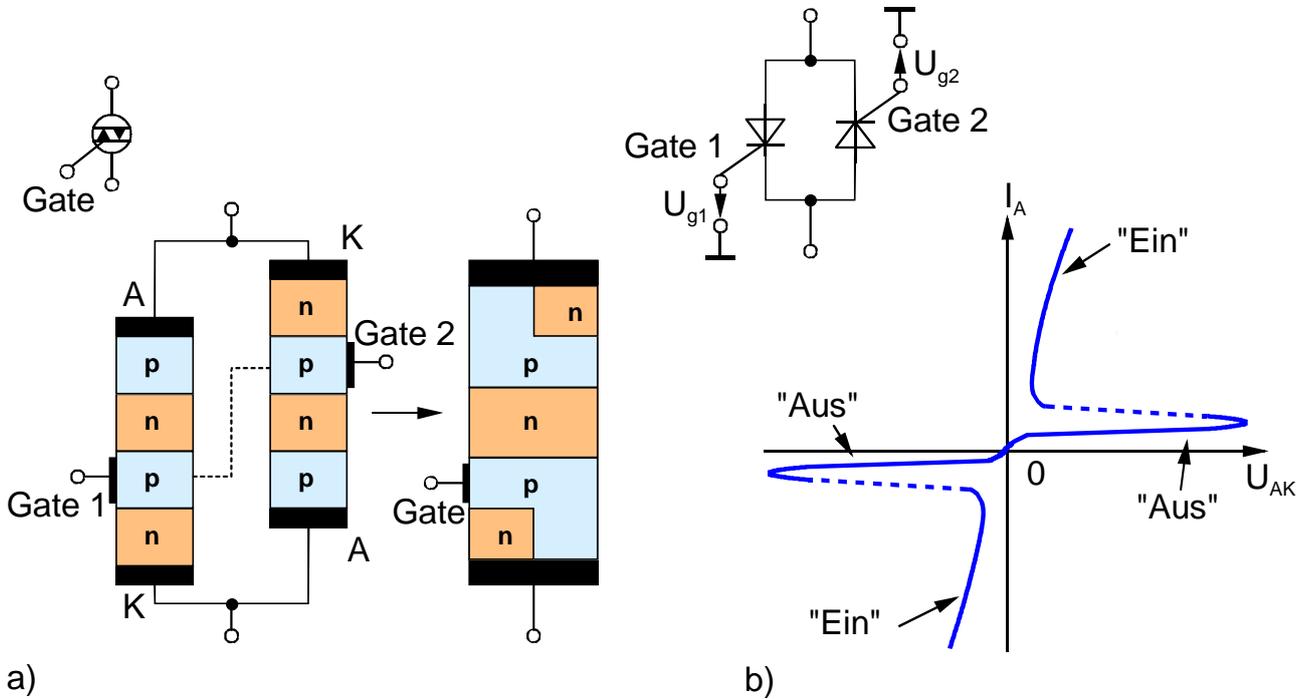


Abb. 5.9 TRIAC

- a) Integration der beiden Thyristoren in a) auf einem gemeinsamen Halbleiterkristall
- b) Kennlinie

Ein einzelner Thyristor kann bei Wechselspannung pro Zyklus nur einmal eingeschaltet werden, während der anderen Halbwelle bleibt er gesperrt. Mit dem TRIAC können nun beide Halbwellen ausgenutzt werden.

5.2.4 Abschaltbare Thyristoren (GTO)

Beim GTO (*gate turn-off thyristor*) kann ein Thyristor, der durch einen positiven Gate-Strom eingeschaltet worden ist (Abb. 5.2d), durch einen negativen wieder ausgeschaltet werden (Abb. 5.10). Es ist also nicht wie beim Thyristor erforderlich, dass zu einem Ausschalten der Strom durch den Thyristor erst auf einen so kleinen Wert reduziert werden muss, dass im Zwei-Transistor-Analogon (Kapitel 5.4.1) die Summe der Einzelstromverstärkungen $|A_1 + A_2|$ kleiner als 1 wird, oder dass der Thyristor ganz in den Sperrbereich geschaltet wird [17]. Mit diesem Element ist es also unnötig, bei Wechselspannungsanwendungen erst auf einen Rückgang des Wechselstromes unter den Haltestrom zu warten, damit das Bauelement abgeschaltet werden kann.

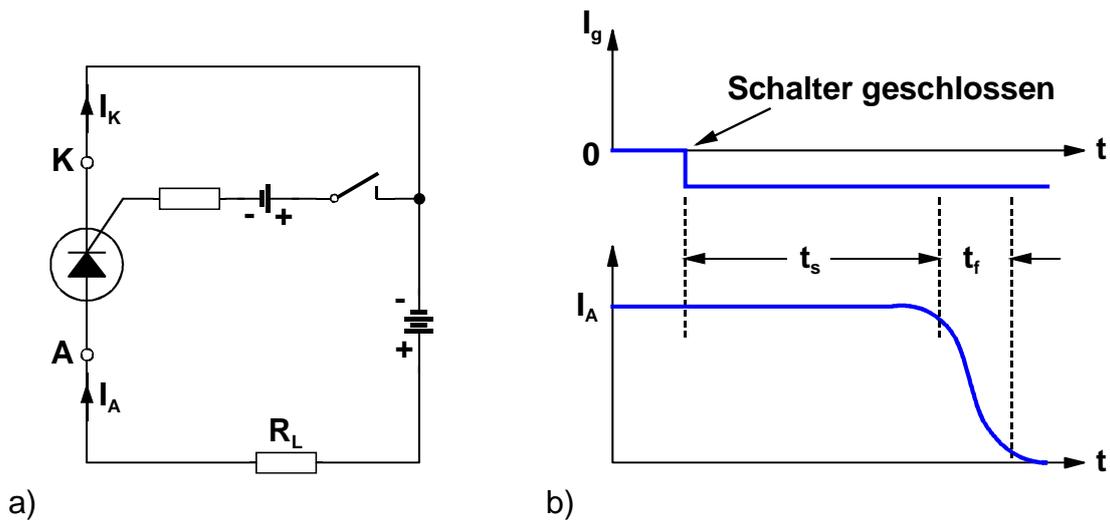


Abb. 5.10 a) Schaltung und
b) Abschaltverhalten eines GTO [17]

Die Funktion des GTOs wird mit Hilfe des Zwei-Transistor-Analogons erläutert: Für den Basisstrom des Transistors T_2 gilt bei Vernachlässigung des Sperrstroms I_{CB0} analog zu Gl. (5.3)-(5.8) und Abb. 5.2:

$$I_{B2} = -(I_{E2} + I_{C2}) = -(1 + A_2)I_{E2} = (1 + A_2)I_K$$

Dieser Basisstrom setzt sich nach dem Knotensatz zusammen aus den Beiträgen (s. Abb. 5.2)

$$I_{B2} = -A_1 \cdot I_A + I_g$$

Im eingeschalteten Zustand ist der Gate-Strom Null, d.h. es gilt

$$I_{B2} = -A_1 \cdot I_A = (1 + A_2) \cdot I_K$$

Bei Anlegen eines negativen Gate-Stroms wird ein Teil des Kollektorstroms von T_1 nicht in die Basis von T_2 geführt, sondern über die Gate-Elektrode aus dem Bauelement herausgeleitet. Dadurch wird der Thyristor abgeschaltet, es gilt

$$I_{B2} = -A_1 \cdot I_A - |I_g| < (1 + A_2) \cdot I_K \quad (5.11)$$

Mit der Knotenregel:

$$I_K = I_A + I_g$$

folgt

$$-|I_g| + I_A = I_K \quad (5.12)$$

Setzt man (5.12) ein in (5.11), dann ergibt sich nach Auflösung die Bedingung für den Gate-Strom:

$$|I_g| > \frac{(A_1 + A_2) + 1}{A_2} \cdot I_A \quad (5.13)$$

5.3 Permeable Base Transistor (PBT)

Schon sehr früh nach der Nutzung von Halbleitern im bipolaren Transistor wurden Vorschläge gemacht, die bei Vakuumröhren erprobten Steuerungsmechanismen durch Gitter auf den Festkörper zu übertragen. Dadurch könnte die Unempfindlichkeit von Festkörpern und ihr Potential zur Miniaturisierung genutzt werden. Man sprach damals vom Analogtransistor, ohne dass die technischen Mittel zu seiner Realisierung gegeben waren. Inzwischen ist die Halbleitertechnologie so weit fortgeschritten, dass man sowohl pn-Übergänge als auch Schottky-Kontakte in gitterähnlichen Strukturen einsetzen kann. Im ersten Fall spricht man vom **Static Induction Transistor (SIT)**, im zweiten Fall von **Permeable Base Transistor (PBT)**.

Abb. 5.11 zeigt den PBT in zwei Schnitten, die senkrecht zueinander orientiert sind.

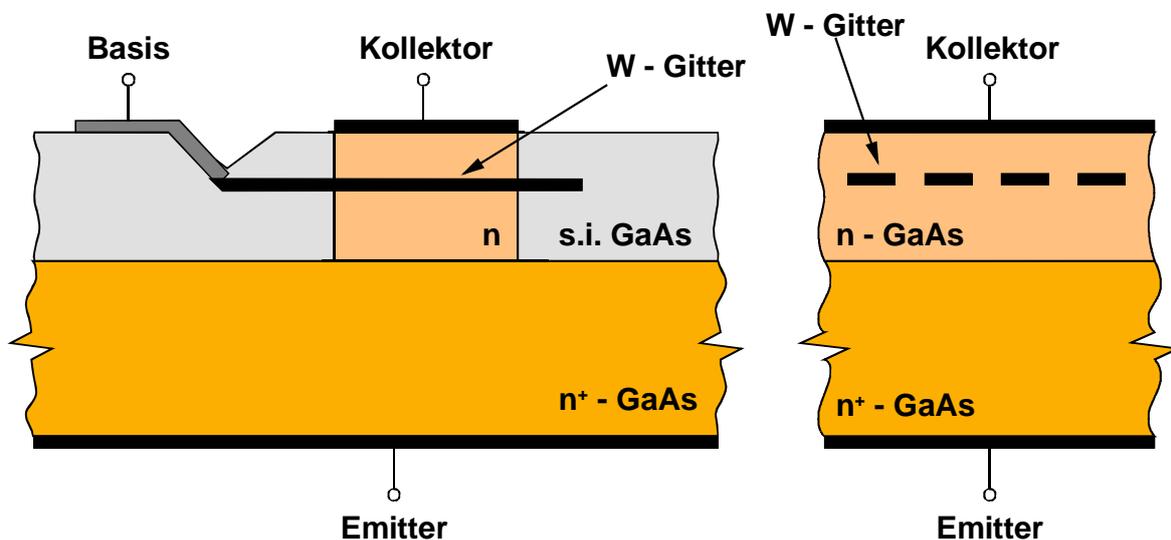


Abb. 5.11 Permeable Base Transistor (PBT): Schnitte in Richtungen, die senkrecht zueinander stehen [22]

Auf einem hoch n-dotierten GaAs-Substrat befindet sich ein dünner einkristalliner GaAs-Film mit einer Elektronenkonzentration im Bereich von 10^{16} cm^{-3} . In den Film ist ein hochfeines Wolfram-Gitter eingebaut, d.h. der Film ist in zwei Etappen vor und nach der Strukturierung des Gitters gewachsen. Die Periode des Gitters ist z. B. $0,4 \mu\text{m}$ mit einer Fingerbreite von $0,2 \mu\text{m}$ und einer Dicke von 20 nm . Die Bereiche um das Gitter herum sind durch Protonenbeschuss semiisierend gemacht, d.h. ihr spezifischer Widerstand liegt bei $10^6 \Omega\text{cm}$. Dadurch bleibt nur der innere Gitterbereich für den Elektronenfluss zwischen den Elektroden Emitter und Kollektor übrig. Der Kollektor ist gegenüber dem Emitter positiv vorgespannt.

Jeder Gitterfinger ist als Schottky-Kontakt von einer Verarmungszone umgeben. Da die Strukturierung des Gitters sehr fein ist, berühren sich die Verarmungszone benachbarter Finger. Daher wird der Stromfluss zwischen Emitter und Kollektor durch eine Potentialbarriere sehr stark eingeschränkt. Wird dagegen die Basis gegenüber dem Emitter positiv vorgespannt, so ziehen sich die Verarmungszone zurück, die Potentialbarriere zwischen den Gitterfingern werden abgebaut, und es kommt zu einem rasch anwachsenden Strom, der - wie bei der Vakuumröhre auch - schließlich durch Raumladungen begrenzt wird.

PBT's haben als Bauelemente, die auf Majoritätsträgereffekten beruhen, den Vorteil großer Schnelligkeit. Zudem weisen sie eine hohe Stromverstärkung auf, weil sie sperrschicht-kontrolliert sind. Grenzfrequenzen f_T , bei der die Stromverstärkung gerade Eins wird, nahe 40 GHz, wurden experimentell erreicht. Als maximale Oszillationsfrequenz f_{\max} wurde 100 GHz gemessen. Zum Vergleich: HEMTs mit einer Gate-Länge von 80 nm liegen etwa bei $f_{\max} = 270\text{GHz}$ [22].

5.4 Insulated-Gate Bipolar Transistor (IGBT)

Um die nahezu leistungslose Ansteuerbarkeit der MOSFETs mit der hohen Stromtragfähigkeit des bipolaren Transistors zu verbinden, ist der IGBT für die Leistungselektronik entwickelt worden (Abb. 5.12).

Im rechten oberen Gebiet befindet sich eine n^+ -Wanne, die man als Source S ansehen kann. Denn ihr gegenüber liegt - durch das p^+ -Kanalgebiet getrennt - der n^- -Drain-Bereich D . Über die Gate-Elektrode G , die durch den SiO_2 -Film vom Halbleiter isoliert ist, kann ein Inversionskanal influenziert werden, so dass ein Strom zwischen S und D fließen kann. Diese Struktur nennt man auch *insulated - gate FET*, ein Synonym zu MOSFET.

Abb. 5.12a enthält aber auch einen bipolaren Transistor, dessen Emmitter E die p^+ -Wanne, dessen Basis die n^-/n^+ -Epitaxieschicht und dessen Kollektor C das p^+ -Substrat darstellen. Bauelemente der beschriebenen Art nutzen dieselben Halbleiterbereiche gemeinsam. So ist die p^+ -Wanne sowohl Kanalbereich des MOSFET als auch Emmitter des p^+np^+ -Transistors. Ebenso dient der n^- -Film sowohl als Drain wie auch als Basis. In solchen Fällen spricht man daher auch von Superintegration.

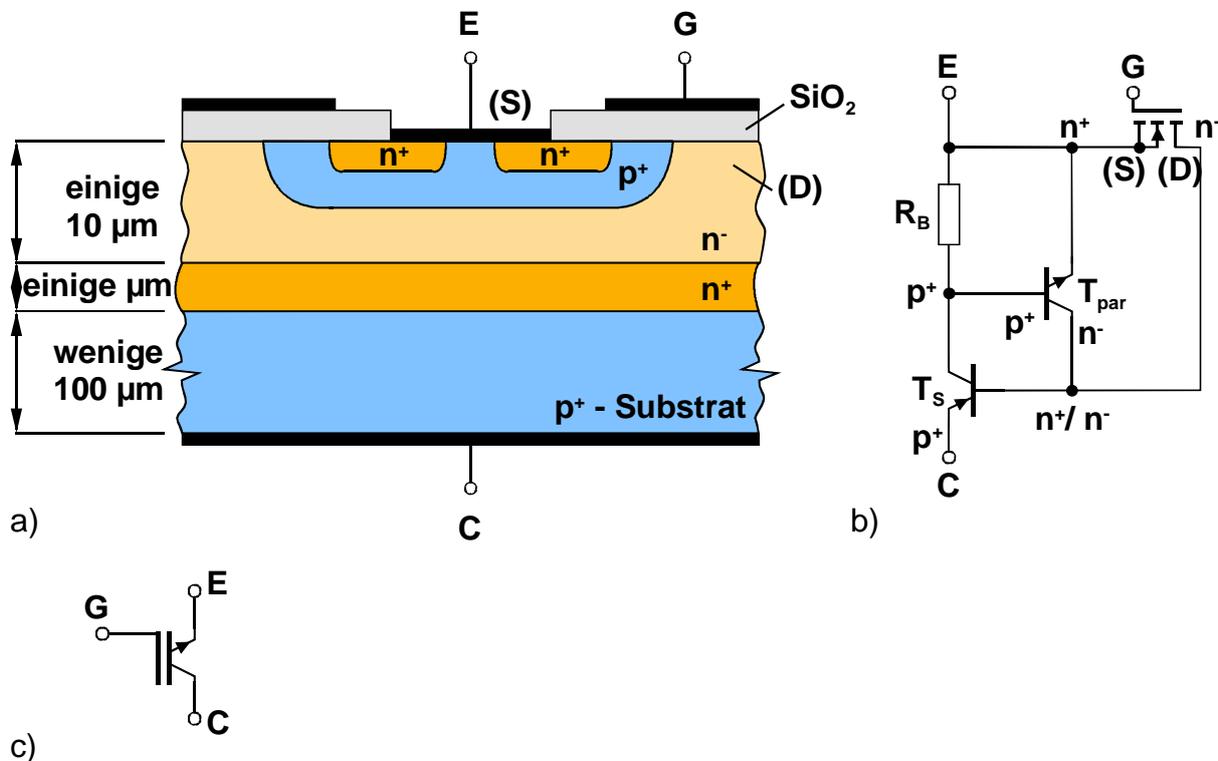


Abb. 5.12 a) Aufbau einer Zelle eines IGBT
 b) zugehöriges Ersatzschaltbild
 c) Schaltsymbol eines n-Kanal IGBT [22]

Im normalen Betrieb ist C positiv gegenüber E vorgespannt, so dass - wenn auch zwischen E und G keine Spannung liegt - der pn-Übergang zwischen p^+ -Wanne und n^- -Epitaxieschicht gesperrt ist. Der p^+np^+ -Transistor, der im Ersatzschaltbild Abb. 5.12b durch T_s symbolisiert ist, leitet daher nicht.

Wenn dagegen eine ausreichend hohe Spannung zwischen G und E liegt, baut sich im MOSFET ein Inversionskanal auf. Es können Elektronen von der n^+ -Wanne durch das p^+ -Gebiet entlang dem Inversionskanal in das n^- -Gebiet fließen. Sie werden zum Basisstrom für die p^+np^+ -Struktur, schalten also T_s ein. Als Konsequenz werden Löcher aus der p^+ -Wanne in die n^- -Basis injiziert. Daher kann nun der Schalttransistor T_s einen Strom tragen. Dieser Zustand dauert so lange an, wie der Inversionskanal über eine hinreichend hohe Gate-Spannung aufrechterhalten bleibt.

Abb. 5.12a ist spiegelsymmetrisch aufgebaut, d.h. dieselbe Diskussion wie eben lässt sich sinngemäß für die linke Hälfte führen. Darüber hinaus können mehrere Zellen der Art von Abb. 5.12a nebeneinander in einer Halbleiterscheibe integriert werden, so dass ein IGBT hohe Ströme tragen kann. Wenn man zudem noch mehrere Halbleiterchips hybrid integriert, kann man sehr leistungsfähige Module aufbauen.

Abb. 5.12b zeigt, dass der parasitäre Transistor T_{par} - gebildet aus n^+ -Source-Wanne, p^+ -Emitter-Wanne und n^- -Epitaxieschicht - überhaupt nicht berücksichtigt wurde. T_{par} und T_s bilden eine Thyristorstruktur. Es kommt darauf an, das Zünden des Thyristors ("Einrasten" oder *latch effect*) zu verhindern, weil sonst der IGBT nicht mehr abgeschaltet werden kann. Dies wird durch Verschlechterung der Stromverstärkung von T_{par} erreicht, indem z.B. die p^+ -Basiswanne sehr hoch dotiert und die Basis-Emitter-Strecke durch die Kontaktmetallisierung wirkungsvoll kurzgeschlossen wird. Es bleibt aber doch ein Bahnwiderstand, der durch R_B in Abb. 5.12b erfasst ist.

Wegen seines bipolaren Ausgangsteils T_s kann der IGBT sehr hohe Ströme tragen. Er ist in diesem Punkt mit Leistungstransistoren vergleichbar, den Leistungs-MOSFETs aber deutlich überlegen. Andererseits ist der IGBT wegen seines MOS-Eingangs rasch und mit geringen Verlusten schaltbar, so dass er in diesen Eigenschaften an Leistungs-MOSFETs heranreicht, aber besser als Leistungstransistoren ist. Kommerziell erhältliche IGBTs sind durch die folgenden typischen Werte charakterisiert:

Kollektorstrom:	10 ... 50 A
Schaltfrequenz:	10 ... 20, auch 100 kHz
maximale Kollektor Emitter-Sperrspannung:	500 ... 1200 V.

Die maximalen Ströme können bei Modulbauweise um nahezu eine Größenordnung höher als angegeben sein.

Stichwortverzeichnis:

Insulated-Gate Bipolar Transistor *5-15*
Permeable Base Transistor *5-14*
Thyristor *5-2*
 Aufbau *5-2*

Sonderformen *5-10*
Technologie *5-8*
Wirkungsweise *5-2*

Kapitel 6

Literaturverzeichnis

6. Literatur

- [1] Brodersen, Gray, Hodges
MOS Switched Capacitor Filters
Proc. IEEE, Vol.67, 1979, p.61-75)
- [2] W.Shockley
A Unipolar Field-Effect Transistor
Proceedings of the I.R.E., 1952, pp.1365-1376)
- [3] W.Brockerhoff
Feldeffekttransistoren aus III/V-Halbleitern: Simulatiuon mit stationären und nichtstationären Modellen
Dissertation, Universität Gesamthochschule Duisburg, 1989, Duisburger Mikroelektronik, Band 5, Verlagsbuchhandlung Nellissen-Wolff
- [4] W.Groß
Digitale Schaltungstechnik
Vieweg Verlag, Studium Technik, ISBN-3-528-03373-8, Braunschweig/Wiesbaden, 1994
- [5] K.-H. Rumpf, K.Pulvers
Elektronische Halbleiterbauelemente – Vom Transistor zur VLSI-Schaltung
Dr. Alfred Hüthig Verlag Heidelberg, ISBN 3-7785-1345-1, 1987
- [6] R.Köstner, A.Möschwitzer
Elektronische Schaltungen
Carl Hanser Verlag, München Wien, Studienbücher, ISBN 3-446-16588-6, 1993
- [7] K.Bystron, J.Borgmeyer
Grundlagen der Technischen Elektronik
Carl Hanser Verlag, München Wien, Studienbücher, ISBN 3-446-15869-3, 1990
- [8] D. A. Neamen
Electronic Circuit Analysis and Design
Irwin Book Team, ISBN 0-256-11919-8, 1996
- [9] A.S.Sedra, K.C.Smith
Microelectronic Circuits
Oxford University Press, 1991, ISBN 019-510369-6
- [10] R.S. Muller, T.I.Kamins
Device Electronics for Integrated Circuits
John Wiley & Sons, 1986, ISBN 0-471-88758-7
- [11] R.J.Baker, H.W.Li, D.E.Boyce
CMOS: Circuit Design, Layout, And Simulation
IEEE Press Series on Microelectronic Systems, IEEE Press, 1998, ISBN 0-7803-3416-7
- [12] H.Tholl
Bauelemente der Halbleiterelektronik
B.G.Teubner, Stuttgart, 1978, II, Teil 2, ISBN 3-519-06419-7

- [13] N. Weimann
Festkörperelektronik
Skript zur Vorlesung, Universität Duisburg-Essen
- [14] U.Tietze, Ch.- Schenk
Halbleiterschaltungstechnik
Springer-Verlag, Berlin
- [15] J. Borgmeyer
Grundlagen der Digitaltechnik
Hanser Lehrbuch, Carl Hanser Verlag München, ISBN 3-446-15624-0
- [16] M.Shur
GaAs Devices and Circuits
Plenum Press, Microdevices: Physics and Fabrication Technologies, New York 1987, ISBN 0-306-42192-5
- [17] H.Schaumburg
Halbleiter
B.G.Teubner Stuttgart
ISBN 3-519-06124-4
- [18] R.Kassing
Physikalische Grundlagen der elektronischen Halbleiterbauelemente
AULA-Verlag, Wiesbaden, Studien-Text
ISBN 3-89104-598-0
- [19] S.M.Sze
Physics of Semiconductor Devices
John Wiley & Sons
- [20] A.Möschwitzer
Grundlagen der Halbleiter-und Mikroelektronik
Band 1: Elektronische Halbleiterbauelemente
Hanser Verlag
ISBN 3-446-16456-1
- [21] M.Cooke
Halbleiter-Bauelemente
Hanser Verlag
ISBN 3-446-16316-6
- [22] A.Schalchetzki
Halbleiter-Elektronik
Teubner Studienbücher
ISBN 3-519-03070-5
- [23] D.Peters
Modellbildung und Simulation von Heterostruktur bipolartransistoren aus III/V Halbleitermaterialien
Dissertation, Universität Duisburg
Shaker-Verlag
ISBN 3-8265-3466-2

Kapitel 7

Stichwortverzeichnis

7. Stichwortverzeichnis

- AC-Verhalten 2-45
- aktiver Bereich 4-27
- Akzeptoren 0-17
- amphoter 0-18
- ANDERSON-Modell 2-33
- Anreicherung 0-32, 1-2
- aspect ratio 2-9
- Ausgangsleitwert 2-54
- Ausgangswiderstand 4-42
- Austrittsarbeit 0-29, 0-37
- Backward-Diode 3-36
- Bahnwiderstand 2-46
- Bändertheorie 0-5
- Bandstruktur 0-5
- Basis-Emitter 4-48
- Basis-Schaltung 4-15
 - Ausgangskennlinienfeld 4-17
 - Eingangskennlinie 4-15
- Besetzungsstatistik 0-12
- Beweglichkeit 0-22
- binärer Halbleiter 0-8
- Bindungsenergie 0-3
- bipolare Bauelemente 2-2
- Bipolartransistor 0-54, 4-2
 - Basis 4-2
 - Basis-Emitter Sperrspannung 4-48
 - Basis-Schaltung 4-15
 - Durchbruch 1. Art 4-48
 - Durchbruchverhalten 4-48
 - Durchbruchverhalten 2. Art 4-48
 - Emitter 4-2
 - Emitter-Schaltung 4-23
 - Emitterstrom 4-11
 - Funktionsweise 4-2
 - Grenzströme 4-49
 - Kennlinienfelder 4-15
 - Kollektor 4-2
 - Kollektor-Basis Sperrspannung 4-48
 - Kollektor-Emitter Sperrspannung 4-48
 - Kollektorstrom 4-11
 - Maximale Verlustleistung 4-49
 - Normalbetrieb 4-14
 - nnp-Transistor 4-2
 - pnp-Transistor 4-2
 - Restspannung 4-47
 - Sättigungsbereich 4-45
 - Schaltvorgang 4-43
 - Sperrbereich 4-44
 - Sperrströme 4-49
 - Technologie 4-4
 - Übersteuerung 4-47
- Bohr'scher Radius 0-17
- Bohr'scher Radius 0-4
- CCD
 - Drei-Phasen-CCD 1-8
 - Fotodetektoren-Aufbau 1-10
 - Fotodiode 1-12
 - Foto-Gate 1-13
 - Sensortypen 1-14
 - Speicherspeicherung 1-5
 - Signaltransport 1-6
 - Zwei-Phasen-CCD 1-10
- CCD-Farbkameras 1-16
 - Drei-Chip Kamera 1-16
 - Ein-Chip Kamera 1-17
- CMOS 0-29
- Coulombkraft 0-4
- Coulomb-Potenzial 0-3
- Coulombstreuung 0-24, 2-35
- DC Stromverstärkung 4-40
- DC-Verhalten 2-45
- Debye-Länge 3-6
- Detektoren 0-29
- differentielle Stromverstärkung 4-40
- Diffusionskapazität 0-55
- Diffusionskoeffizient 0-28
- Diffusionskonstanten 3-2
- Diffusionslänge 3-10
- Diffusionsschwänze 3-7
 - Stromdichteanteile 3-10
- Diffusionsspannung 0-38, 3-4
- Diffusionsstrom 0-27
- Diffusionsstrom 3-2
- Diode 3-2
 - Backward-Diode 3-36
 - Kennzeichnung 3-40
 - Kurzbasisdiode 3-13
 - Langbasisdiode 3-13
 - pin-Diode 3-37
 - reale Diode 3-15
 - Schottky-Diode 3-39
 - Spezielle Dioden 3-24
 - Temperatureinfluss 3-17
 - Temperatureinfluss auf
 - Diffusionsspannung 3-18
 - Temperatureinfluss auf gesamte Kennlinie 3-19
 - Temperatureinfluss auf Sperrstrom 3-17
 - Tunnel-Diode 3-35
 - Varaktor-Diode (Kapazitätsdiode) 3-24
 - Zenerdiode 3-28
- Diodencharakteristik 0-36

- Donator 0-15
- Dotierung 0-15
- Dotierungsabhängigkeit 0-24
- Driftbeweglichkeit 0-22
- Driftgeschwindigkeit 0-22
- Durchbruch 1. Art 4-48
- Durchbruchverhalten 4-48
- Durchbruchverhalten 4-48
- Early-Effekt 4-36
- Ebers-Moll Gleichungen 4-12
- Ebers-Moll-Gleichungen 4-8
- Eigenhalbleiter 0-12
- Eigenleitungsbereich 0-20
- Eigenleitungskonzentration 0-12
- Eigenwerte 0-3
- Eigenwertspektrum 0-3
- Einstein Beziehung 0-28
- elektrisches Feld 0-22
- Emitter-Schaltung
 - Ausgangswiderstand 4-42
 - DC-Stromverstärkung 4-40
 - differentielle Stromverstärkung 4-40
 - Eingangswiderstand 4-38
 - Kleinsignalverhalten 4-38
 - quasistationäres Kleinsignal-
Ersatzschaltbild 4-43
- Emitter-Schaltung 4-23
 - Ausgangskennlinienfeld 4-24
 - Eingangskennlinie 4-23
- Emitter-Schaltung
 - Steilheit 4-41
- Energieband 0-5
- Energielücke 0-7
- Energieschalen 0-2
- Entartung 0-14
- Entwurfsparameter 2-9
- Epitaxie-Planar-Transistor 4-5
- Ersatzschaltbild 2-45
- Erschöpfungsbereich 0-19, 0-20
- extrinsischer Bereich 2-46
- extrinsischer FET 2-46
- Feldeffekttransistor 0-32
- Feldeffekttransistor (FET)
 - Ausgangsleitwert 2-54
 - Gleichstromverhalten 2-2
 - Kleinsignalersatzschaltbild in Source-
Schaltung (Näherung) 2-54
 - Prinzip 2-2
 - quasistationäres Kleinsignalverhalten 2-52
 - Spannungsverstärkung 2-55
 - Steilheit 2-53
 - transconductance 2-17, 2-53
 - Typen-Übersicht 2-3
- Feldstrom 0-26, 3-3
- Fermi-Dirac-Verteilungsfunktion 0-11
- Fermi-Energie 0-11
- Festkörpermodell 0-5
- Flachbandfall 1-2
- Flachbandspannung 0-32
- Fotogeneration 1-5
- Frame-Transfer-Sensor 1-15
- Full-Frame-Sensor 1-16
- Galliumarsenid (GaAs) 0-8
- Germanium (Ge) 0-8
- Geschaltete Kapazitäten
 - Filter 1-20
 - Integrator mit Operationsverstärker 1-22
 - Realisierung in MOS-Technik 1-23
 - Wirkungsweise 1-19
- Geschwindigkeit
 - thermische Geschwindigkeit 0-22
- Geschwindigkeitssättigung 0-25, 2-12
- Gleichspannungsverhalten 2-45
- Grenzflächen 0-29
 - Grenzflächenladung 0-32
 - Halbleiter-Halbleiter-Grenzflächen 0-29,
0-44
 - Metall-Halbleiter-Grenzfläche 0-29
- Grenzflächenladung 0-32
- Grenzströme 4-49
- Gummel-Plot 4-36
- Halbleiter 0-7
 - entarteter Halbleiter 0-14
- Halbleiter-Halbleiter-Grenzflächen 0-44
- Hauptquantenzahlen 0-4
- Heterostruktur-Bipolartransistor (HBT) 4-7,
4-53
- Heterostruktur-Feldeffekttransistor (HFET) 2-33
- Heterostruktur-Laser 0-29
- Heterostruktur-Transistoren 0-29
- Heteroübergang 2-33
- HFET 2-24, 2-33
 - Aufbau 2-36
 - Bänderdiagramm 2-34
 - Gleichgewichtsfall 2-39
 - Gleichstromverhalten 2-44
 - Heteroübergang 2-33
 - Ladungstransport im 2DEG 2-35
 - Querschnitt 2-37
 - Schichtkonzentrations-Berechnung 2-38
 - Schottky-Kontakt 2-43
- Homöübergang 0-44
- Hybrid-Form 2-52
- Hybrid-Parameter 2-51
- III-V Halbleiter 0-8

- Injektionsmodell 4-33
- Insulated-Gate Bipolar Transistor 5-15
- Interline-Transfer-Sensor 1-14
- intrinsischer Bereich 2-46
- intrinsische Konzentration 0-12
- intrinsischer FET 2-46
- Inversion 0-31, 0-32, 1-2
- Inverter 4-48
 - Prinzipschaltung 2-49
- Ionenimplantation 4-5
- Ionisierungsenergie 0-4, 0-15
- Isolator 0-7
- Isolatorkapazität 0-33
- ISOPLANAR-Transistor 4-6
- JFET 2-4, 2-24
 - Aufbau 2-25
- Junction-FET 2-24
- Kanalabschnürung 2-12
- Kanalverkürzung 2-12
- Kapazitäten
 - Geschaltete Kapazitäten 1-19
- Kapazitäts-Spannungs-Kennlinie 0-34
- Kennlinienfeld 2-15
 - Ausgangskennlinienfeld 2-15
 - Eingangskennlinienfeld 2-15
 - Übertragungskennlinienfeld 2-15
- Kleinsignalverhalten 2-51
 - allgemein 2-50
 - Kennlinie 2-51
- Kollektor-Basis 4-48
- Kollektor-Emitter 4-48
- Kontakt
 - Ohm-Kontakt 0-36, 0-42
 - Schottky-Kontakt 0-36
- Kontaktwiderstand 0-43
- Kontinuitätsgleichung 3-10
- Konzentration
 - extrinsische 0-15
 - intrinsische 0-12
- Konzentrationsgradient 0-27
- kovalente Bindungen 0-15
- Kurzbasisdiode 0-54, 3-13
- Ladungsneutralität 0-19
- Ladungsträgerbeweglichkeit 0-22
- Ladungsträgerkonzentration 0-10
- Ladungstransport 0-21
- Langbasisdiode 0-54, 3-13
- Lawineneffekt 3-32
- Lebensdauer 3-10
- Legieren 0-43
- Leitfähigkeit 0-23
- Leitungsband 0-6
 - effektive Zustandsdichte 0-13
- Leitungsbandkante 0-11
- Leitwert-Form 2-52
- Leitwert-Parameter 2-51
- Majoritätsladungsträger 0-21
- Massenwirkungsgesetz 0-17, 0-19, 0-46
- Maximale 4-49
- MESFET 2-4, 2-24
 - Anlaufbereich 2-31
 - Aufbau 2-25
 - Kennlinienberechnung 2-28
 - Potenzialverteilung 2-29
 - Sättigung 2-32
 - selbstleitender MESFET 2-28
 - selbstperrender MESFET 2-27
- Metall 0-8
- Metall-Halbleiter-Kontakt 0-37
- Metall-Halbleiter-Übergang 0-37
- Metallorganische Gasphasenepitaxie 2-37
- Metall-Semiconductor-FET 2-24
- Minoritäten 0-46
- Minoritätsladungsträger 0-21
- MIS-Kapazität 0-34
- MIS-Struktur 0-29
- Molekularstrahlepitaxie 2-37
- MOS 0-29
 - Kondensator 1-2
 - MOS-Feldeffekttransistor (MOSFET) 2-4
- MOSFET 2-4
 - Anreicherungstypen 2-17
 - Arbeitspunkteinstellung 2-22
 - Aufbau und Wirkungsweise 2-4
 - Ausgangskennlinienfeld 2-15
 - Drain-Strom-Berechnung 2-6
 - Eingangskennlinienfeld 2-15
 - Grundsaltungen 2-15
 - Kennlinien-Temperaturabhängigkeit 2-20
 - n-Kanal-Verarmungs-MOSFET 2-23
 - Potenzialverteilung im Kanal 2-7
 - Sättigungsgebiet 2-10
 - Source-Schaltung 2-16
 - Spannungsbereichs-Darstellung 2-14
 - steuerbarer Widerstand 2-9
 - Substratvorspannung-Einfluß 2-20
 - Übertragungskennlinienfeld 2-15
 - Verarmungstypen 2-19
- MOS-Kondensator 1-2
 - Anreicherung 1-2
 - Anwendungen 1-5
 - CCD 1-5
 - CV-Kennlinie 1-3
 - Flachbandfall 1-2
 - Inversion 1-2
 - MOS-Struktur 1-3

- Verarmung 1-2
- n-Halbleiter 0-15
- Normalbetrieb 4-14, 4-17
- npn-Transistor 4-2
- Ohm-Kontakt 0-36, 0-42
- Ohm'sches Gesetz 0-24
- parasitäre Größen 2-46
- Pauli-Prinzip 0-4
- Permeable Base Transistor 5-14
- p-Halbleiter 0-15
- Phonon 0-23
- Phononenstreuung 0-23, 2-35
- pinch-off 2-12
- pin-Diode 3-37
- Planck 0-2
- pn-Diode
 - Kleinsignalverhalten 3-22
- pn-Übergang
 - abrupter Übergang 3-25
 - Grundgleichungen 3-2
 - Hyperabrupter Übergang 3-25
 - linearer Übergang 3-25
 - stromdurchflossen 3-7
- Potenzialbarriere 0-42
- Potenzialtopf 0-3, 2-33
- Prinzip des detaillierten Gleichgewichts 0-49
- Punchthrough 4-37
- Quantentheorie 0-2
- quasistationäres Verhalten 2-51
- Raumladung 0-34, 3-2
- Raumladungskapazität 0-34
- Raumladungszone 0-39, 0-47
- recessed gate 2-48
- Rekombinationsrate 3-10
- Reservebereich 0-20
- Schottky 0-38
- Schottky-Dioden 3-39
- Schottky-Kontakt 0-36
- Schwellenspannung 2-8
- Sensortypen
 - Frame-Transfer-Sensor 1-15
 - Full-Frame-Sensor 1-16
 - Interline-Transfer-Sensor 1-14
- Silizium (Si) 0-8
- Source-Schaltung 2-16
- Spannungsverstärkung 2-55
- Sperrschicht-Feldeffekttransistoren 2-24
 - Aufbau und Wirkungsweise 2-25
- Sperrschichtkapazität 0-55
- Sperrströme 4-49
- spezifischer Widerstand 0-9, 0-23
- Steilheit 2-17, 2-53
- Steilheit 4-41
- Störstellenhalbleiter 0-12, 0-15
- Streu-Parameter 2-51
- Streu-Parameter-Form 2-52
- Stromdichte 0-21
- Subbänder 2-33
- Taylor-Reihe 2-51
- Temperaturspannung 3-2
- thermische Energie 0-7
- thermische Geschwindigkeit 0-22
- thermodynamisches Gleichgewicht 0-19
- Thyristor 5-2
 - Aufbau 5-2
 - Sonderformen 5-10
 - Technologie 5-8
 - Wirkungsweise 5-2
- transconductance 2-17, 2-53
- Transportmodell 4-33
- Tunnel-(Zener-)Effekt 3-29
- Tunnel-Diode 3-35
- unipolare Bauelemente 2-2
- Valenzband 0-5
 - effektive Zustandsdichte 0-13
- Valenzbandkante 0-11
- Valenzelektronen 0-2, 0-5
- Varaktor-Diode 3-24
- Verarmung 0-32, 1-2
- Verbindungshalbleiter 0-8, 0-18
- versenktes Gate 2-48
- Verstärkerstufe
 - Prinzipschaltung 2-49
- Verteilungsfunktion 0-10
- Vierpol
 - Vierpolparameter 2-51
- Wechselspannungsverhalten 2-45, 2-49
 - Funktionsprinzip 2-49
 - Kleinsignalverhalten 2-50
- Widerstand
 - spezifischer Widerstand 0-9
- Zenerdiode 3-28
- Zustandsdichte 0-10
 - effektive Zustandsdichte des Leitungsbandes 0-13
 - effektive Zustandsdichte des Valenzbandes 0-13
- zweidimensionales Elektronengas 2-33

