

Springer-Lehrbuch

Th. Tille • D. Schmitt-Landsiedel

Mikroelektronik

Halbleiterbauelemente und deren Anwendung in
elektronischen Schaltungen

Mit 243 Abbildungen



Springer

Dr.-Ing. Thomas Tille

Technische Universität München
Arcisstr. 21
80333 München

BMW AG
Knorrstr. 147
80788 München

Prof. Dr. rer. nat. Doris Schmitt-Landsiedel

Technische Universität München
Arcisstr. 21
80333 München

ISBN 3-540-20422-9 Springer-Verlag Berlin Heidelberg New York

Bibliografische Information der Deutschen Bibliothek

Die Deutsche Bibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie;
detaillierte bibliografische Daten sind im Internet über <http://dnb.ddb.de> abrufbar.

Dieses Werk ist urheberrechtlich geschützt. Die dadurch begründeten Rechte, insbesondere die der Übersetzung, des Nachdrucks, des Vortrags, der Entnahme von Abbildungen und Tabellen, der Funksendung, der Mikroverfilmung oder Vervielfältigung auf anderen Wegen und der Speicherung in Datenverarbeitungsanlagen, bleiben, auch bei nur auszugsweiser Verwertung, vorbehalten. Eine Vervielfältigung dieses Werkes oder von Teilen dieses Werkes ist auch im Einzelfall nur in den Grenzen der gesetzlichen Bestimmungen des Urheberrechtsgesetzes der Bundesrepublik Deutschland vom 9. September 1965 in der jeweils geltenden Fassung zulässig. Sie ist grundsätzlich vergütungspflichtig. Zu widerhandlungen unterliegen den Strafbestimmungen des Urheberrechtsgesetzes.

Springer ist ein Unternehmen von Springer Science+Business Media

springer.de

© Springer-Verlag Berlin Heidelberg 2005
Printed in Germany

Die Wiedergabe von Gebrauchsnamen, Handelsnamen, Warenbezeichnungen usw. in diesem Buch berechtigt auch ohne besondere Kennzeichnung nicht zu der Annahme, dass solche Namen im Sinne der Warenzeichen- und Markenschutz-Gesetzgebung als frei zu betrachten wären und daher von jedermann benutzt werden dürften. Sollte in diesem Werk direkt oder indirekt auf Gesetze, Vorschriften oder Richtlinien (z. B. DIN, VDI, VDE) Bezug genommen oder aus ihnen zitiert worden sein, so kann der Verlag keine Gewähr für die Richtigkeit, Vollständigkeit oder Aktualität übernehmen. Es empfiehlt sich, gegebenenfalls für die eigenen Arbeiten die vollständigen Vorschriften oder Richtlinien in der jeweils gültigen Fassung hinzuzuziehen.

Satz: Gelieferte Daten des Autors

Einbandgestaltung: medionet AG, Berlin

Gedruckt auf säurefreiem Papier

7/3020/M - 5 4 3 2 1 0

Vorwort

Die Mikroelektronik stellt mit ihren innovativen Entwicklungen und breiten Anwendungsmöglichkeiten einen wichtigen Wirtschaftsfaktor in der heutigen Zeit dar, dessen Bedeutung mehr und mehr zunimmt. Auf Grund der wachsenden Komplexität in der Mikroelektronik ist die Kenntnis der Wirkungsweise von Halbleiterbauelementen und deren Nutzung in elektronischen Schaltungen von großer Bedeutung.

Dieses Buch vermittelt die wichtigsten Themengebiete der Mikroelektronik in prägnanter Form. Es gestattet dem Leser, das Verständnis für den Aufbau und die Funktionsweise elektronischer Halbleiterbauelemente der Mikroelektronik zu erlangen und den Einsatz in schaltungstechnischen Anwendungen kennen zu lernen. Dabei steht die Verknüpfung physikalischer Halbleiter-Effekte mit der Funktionsweise elektronischer Bauelemente und integrierter Schaltungen im Vordergrund. Durch die Vielfalt und Beschreibungsweise der behandelten elektronischen Halbleiterbauelemente und deren Anwendungsmöglichkeiten wird ein umfassendes Bild des Stoffgebietes geliefert.

Der Inhalt des Buches erstreckt sich über die Gebiete Halbleiter-Grundlagen, Halbleiterdioden, Bipolar- und Feldeffekttransistoren sowie Leistungsbauelemente, Halbleitersensoren und Integrierte Schaltungen. Die Halbleiter-Grundlagen umfassen u.a. das Bindungs- und Bändermodell, Relaxationsprozesse, den Ladungsträgertransport und weitere Halbleitereffekte. Die Gebiete Halbleiterdioden, Bipolar- und Feldeffekttransistoren beinhalteten jeweils den Aufbau, die Wirkungsweise, die Strom-Spannungs-Charakteristika und technologische Realisierung dieser elektronischen Bauelemente sowie die zugehörigen Netzwerkmodelle, das Schaltverhalten und die wesentlichen Grundschaltungen. Des Weiteren werden Leistungsbauelemente an Hand der wichtigsten Vertreter hinsichtlich ihrer Eigenschaften und Anwendbarkeit beschrieben. Aus dem Gebiet der Halbleitersensoren werden Temperatur-, Magneto-, Opto- und Chemosensoren sowie Sensoren für mechanische Größen betrachtet. Im Kapitel Integrierte Schaltungen wird auf die statische und dynamische CMOS-Logik sowie auf CMOS-Speicherschaltungen und integrierte Logikbausteine eingegangen.

Das Buch eignet sich als vorlesungsbegleitendes Lehrbuch an Universitäten und Fachhochschulen, zum Selbststudium und als hilfreiches Kompendium im Beruf. Es richtet sich sowohl an Studierende der Ingenieur- und Naturwissenschaften im Hauptstudium, als auch an Professionals in Wissenschaft und Wirtschaft.

Inhaltsverzeichnis

Formelzeichen und Naturkonstanten	XIII
1 Halbleiter-Grundlagen.....	1
1.1 Halbleitermaterialien.....	1
1.2 Bindungsmodell.....	3
1.2.1 Gitterstruktur	3
1.2.2 Eigenleitung	4
1.2.3 Störstellenleitung	7
1.3 Bändermodell.....	8
1.3.1 Aufbau des Bändermodells.....	9
1.3.2 Bänderstruktur im Energie-Impuls-Raum	11
1.3.3 Bänderbesetzung.....	11
1.3.3.1 Zustandsdichte	12
1.3.3.2 Besetzungswahrscheinlichkeit	12
1.3.3.3 Ladungsträgerverteilung.....	13
1.3.4 Entartete Halbleiter	15
1.4 Ladungsträgertransport.....	16
1.4.1 Driftstromdichte.....	16
1.4.2 Diffusionsstromdichte	18
1.4.3 Konvektionsstromdichte.....	19
1.5 Halbleiter im Nichtgleichgewicht	20
1.5.1 Majoritätsträgerinjektion	20
1.5.2 Minoritätsträgerinjektion	21
1.5.3 Minoritäts- und Majoritätsträgerinjektion	22
1.5.4 Minoritätsträgerdiffusion	24
1.5.5 Ladungsträgerbilanz.....	26
1.6 Weitere Halbleitereffekte	27
1.6.1 Lawineneffekt	27
1.6.2 Thermoeffekt	28
1.6.2.1 Thermospannung im homogenen Halbleiter	28
1.6.2.1 Thermospannung zwischen verschiedenen Materialien.....	29
1.6.3 Halleffekt	29
2 Halbleiterdioden.....	32
2.1 pn-Übergang	32
2.1.1 Abrupter pn-Übergang im thermodynamischen Gleichgewicht.....	32

2.1.2 Abrupter pn-Übergang unter Vorspannung	35
2.1.2.1 Raumladungszone des pn-Übergangs.....	36
2.1.2.2 pn-Übergang in Flusspolung	38
2.1.2.3 pn-Übergang in Sperrpolung	41
2.2 Diodenkennlinien.....	42
2.2.1 Kennlinie der idealen Diode	42
2.2.2 Kennlinie der realen Diode.....	43
2.2.3 Diodenkennlinie mit 2-Geraden-Näherung	43
2.3 Durchbruchsmechanismen	44
2.2.1 Lawineneffekt.....	44
2.2.2 Zenereffekt	46
2.2.3 Thermischer Durchbruch.....	46
2.4 Technologische Realisierung.....	47
2.5 Netzwerkmodelle der pn-Diode	48
2.5.1 Sperrsicht- und Diffusionskapazität	48
2.5.2 Großsignalersatzschaltbild	50
2.5.3 Kleinsignalersatzschaltbild.....	50
2.6 Schaltverhalten der pn-Diode	51
2.6.1 Schalten ohmscher Lasten.....	51
2.6.2 Schalten induktiver Lasten	53
2.7 Diodentypen	54
2.7.1 pin-Diode	54
2.7.2 Z-Diode	57
2.7.3 Tunneldiode.....	58
2.7.4 Schottky-Diode.....	59
2.7.5 Kapazitätsdiode	61
2.7.6 Fotodiode	62
2.7.7 Solarzelle.....	64
2.7.8 Lumineszenzdiode	67
2.8 Dioden-Grundschaltungen.....	71
2.8.1 Einweggleichrichter	71
2.8.2 Zweiwegbrückenschaltung.....	72
2.8.3 Spannungsvervielfacher	73
2.8.4 Frequenzmodulation	74
2.8.5 Spannungsbegrenzung	75
3 Bipolartransistoren	77
3.1 Aufbau und Funktionsweise des Bipolartransistors	77
3.1.1 Aufbau des Bipolartransistors.....	77
3.1.2 Bipolartransistor im thermodynamischen Gleichgewicht.....	78
3.1.3 Bipolartransistor in Vorwärtsbetrieb.....	79
3.1.4 Beschaltungsarten	82
3.2 Strom-Spannungs-Charakteristik des Bipolartransistors	84
3.2.1 Transistorkennlinie	84
3.2.2 Basisweitenmodulation	85
3.3 Durchbruchsmechanismen	86

3.3.1 Lawinendurchbruch	86
3.3.2 Durchbruch zweiter Art (thermischer Durchbruch)	87
3.3.3 Punch-Through-Effekt	87
3.4 Technologische Realisierung	88
3.5 Netzwerkmodelle des Bipolartransistors	89
3.5.1 Sperrsicht- und Diffusionskapazität	89
3.5.2 Großsignalersatzschaltbild	90
3.5.3 Kleinsignalersatzschaltbild	91
3.5.4 Grenzfrequenzen	93
3.6 Schaltverhalten des Bipolartransistors	94
3.7 Bipolartransistor-Grundschaltungen	97
3.7.1 Darlington-Schaltung	97
3.7.2 Stromquelle und Stromspiegel	99
3.7.3 Differenzverstärker	102
4 Feldeffekttransistoren	105
4.1 MOS-Struktur	105
4.1.1 MOS-Struktur im thermodynamischen Gleichgewicht	105
4.1.2 MOS-Struktur im Flachbandfall	106
4.1.3 MOS-Struktur in Akkumulation	107
4.1.4 MOS-Struktur in Verarmung	108
4.1.5 MOS-Struktur in Inversion	109
4.1.6 Kapazität der MOS-Struktur	111
4.1.7 Komplementäre MOS-Strukturen	113
4.2 Aufbau und Funktionsweise des MOS-Feldeffekttransistors	115
4.2.1 Aufbau des MOSFETs	115
4.2.2 Funktionsweise des MOSFETs	115
4.3 Strom-Spannungs-Charakteristik des MOSFETs	117
4.3.1 Theorie der Ladungssteuerung	117
4.3.2 Einsatzspannung des MOSFETs	120
4.3.3 Substratsteuerung des MOSFETs	121
4.3.4 Unterschwellstrom des MOSFETs	122
4.3.5 Thermisches Verhalten des MOSFETs	122
4.4 Geometrieabhängigkeit von MOSFET-Parametern	123
4.5 Durchbruchsmechanismen und Degradation	124
4.6 Technologische Realisierung	126
4.7 Netzwerkmodelle des MOSFETs	127
4.7.1 MOSFET-Kapazitäten	127
4.7.2 Großsignalersatzschaltbild	128
4.7.3 Kleinsignalersatzschaltbild	129
4.8 MOSFET-Grundschaltungen	130
4.8.1 Einfache Verstärkerschaltungen	130
4.8.1.1 Source-Schaltung	130
4.8.1.2 Drain- und Gate-Schaltung	132
4.8.2 CMOS-Inverter	132
4.8.2.1 Statisches Übertragungsverhalten	132

4.8.2.2 Schaltverhalten des CMOS-Inverters	134
4.8.3 Transmissiongate	136
4.8.4 Stromspiegel.....	137
4.8.5 Differenzverstärker	140
4.8.6 Operationsverstärker	141
4.9 Vergleich von Bipolartransistor und MOSFET	144
4.9.1 Technologischer und parametrischer Vergleich.....	144
4.9.2 Einsatzkriterien von Bipolar- und Feldeffekttransistoren.....	145
5 Leistungsbauelemente	146
5.1 psn-Leistungsdiode	146
5.2 Leistungs bipolartransistor	148
5.3 Leistungs-MOSFET	150
5.4 Thyristor.....	154
5.4.1 Aufbau und Wirkungsweise des Thyristors.....	154
5.4.2 Thyristorkennlinie.....	156
5.4.3 Zündverhalten des Thyristors	157
5.4.4 Schaltverhalten des Thyristors.....	159
5.4.5 Weitere Thyristortypen	161
5.4.5.1 Fotothyristor	161
5.4.5.2 GTO-Thyristor.....	161
5.4.6 Anwendung des Thyristors.....	162
5.5 IGBT	164
5.5.1 Aufbau und Wirkungsweise des IGBTs	164
5.5.2 Kennlinie des IGBTs.....	166
5.5.3 Schaltverhalten des IGBTs	166
5.5.4 Weitere IGBT-Typen	168
5.5.4.1 Punch-Through-IGBT	168
5.5.4.2 Trench-IGBT	169
5.5.5 Anwendung des IGBTs	170
5.6 Latch-up-Effekt	170
6 Halbleitersensoren.....	172
6.1 Temperatursensoren.....	172
6.1.1 Grenzflächen-Temperatursensoren	172
6.1.2 Spreading-Resistance Si-Temperatursensoren.....	174
6.1.3 Thermopiles.....	177
6.2 Magnetosensoren	178
6.3 Optische Sensoren.....	182
6.4 Chemosensoren.....	187
6.4.1 Metalloxid-Gassensoren.....	187
6.4.2 Chemosensitiver Feldeffekttransistor	190
6.5 Sensoren für mechanische Größen.....	192
6.5.1 Piezoresistive Kraft- und Drucksensoren	193
6.5.1.1 Halbleiter-Dehnungsmessstreifen	193
6.5.1.2 Piezoresistiver Drucksensor.....	195

6.5.2 Kapazitiver Drucksensor	196
7 Integrierte Schaltungen	197
7.1 CMOS-Grundschaltungen	197
7.1.1 Statische CMOS-Logik	197
7.1.1.1 Realisierung logischer Verknüpfungen	197
7.1.1.2 Tristate-Treiber	201
7.1.1.3 Addierer	202
7.1.2 Dynamische CMOS-Logik	204
7.1.2.1 Dynamisches Latch	204
7.1.2.2 Dynamisches Schieberegister	204
7.2 Speicher	205
7.2.1 Nichtflüchtige Speicher	205
7.2.1.1 ROM	206
7.2.1.2 PROM	207
7.2.1.3 EPROM	207
7.2.1.4 EEPROM	209
7.2.1.5 Flash-EPROM	211
7.2.1.6 FRAM	211
7.2.2 Flüchtige Speicher	213
7.2.2.1 SRAM	213
7.2.2.2 DRAM	214
7.3 Integrierte Logik-Bausteine	216
7.3.1 PLD	216
7.3.1.1 PAL und GAL	217
7.3.1.2 CPLD	218
7.3.1.3 FPGA	218
7.3.2 ASIC	220
7.3.2.1 Gate-Array	220
7.3.2.2 Standardzellen-IC	220
7.3.2.3 Voll-kundenspezifischer IC	221
Literaturverzeichnis	222
Sachverzeichnis	227

Formelzeichen und Naturkonstanten

Größe	Bedeutung	Einheit
A	Fläche	m^2
A_D	Differenzverstärkung	—
A_G	Gleichtaktverstärkung	—
A_G	Gatefläche	m^2
A_{IB}	Stromverstärkung in Basisschaltung	—
A_{IC}	Stromverstärkung in Kollektorschaltung	—
A_{IE}	Stromverstärkung in Emitterschaltung	—
A_{pn}	Fläche einer pn-Diode	m^2
A_{V0}	Leerlaufverstärkung	—
A_{VB}	Spannungsverstärkung in Basisschaltung	—
A_{VC}	Spannungsverstärkung in Kollektorschaltung	—
A_{VE}	Spannungsverstärkung in Emitterschaltung	—
B	Magnetische Induktion	$\text{V}\cdot\text{s}\cdot\text{m}^{-2}$
c	Lichtgeschwindigkeit $c = 2,998 \cdot 10^8 \text{ m}\cdot\text{s}^{-1}$	$\text{m}\cdot\text{s}^{-1}$
C_B	Basiskapazität	F
C_D	Depletionkapazität	$\text{F}\cdot\text{m}^{-2}$
C_{DB}	Drain-Bulk-Sperrschichtkapazität	$\text{F}\cdot\text{m}^{-2}$
C_{Diff}	Diffusionskapazität	$\text{F}\cdot\text{m}^{-2}$
C_{GB}	Gate-Bulk-Kapazität	$\text{F}\cdot\text{m}^{-2}$
$C_{GB,min}$	Minimale Gate-Bulk-Kapazität	$\text{F}\cdot\text{m}^{-2}$
C_{GD}	Gate-Drain-Kapazität	$\text{F}\cdot\text{m}^{-2}$
C_{GS}	Gate-Source-Kapazität	$\text{F}\cdot\text{m}^{-2}$
C_L	Lastkapazität	F
C_{ox}	Gateoxidkapazität	$\text{F}\cdot\text{m}^{-2}$
C_S	Sperrschichtkapazität	$\text{F}\cdot\text{m}^{-2}$
C_{SB}	Source-Bulk-Sperrschichtkapazität	$\text{F}\cdot\text{m}^{-2}$
d	Durchmesser	m
D_n	Diffusionskonstante der Elektronen	$\text{m}^{-2}\cdot\text{s}^{-1}$
D_p	Diffusionskonstante der Löcher	$\text{m}^{-2}\cdot\text{s}^{-1}$
e	Elementarladung $e = 1,602 \cdot 10^{-19} \text{ C}$	C
E	Elektrische Feldstärke	$\text{V}\cdot\text{m}^{-1}$
E	Energie	eV
E_λ	Spektrale Bestrahlungsstärke	$\text{W}\cdot\text{m}^{-2}$
E_\perp	Vertikale Feldstärke	$\text{V}\cdot\text{m}^{-1}$

Größe	Bedeutung	Einheit
E_{\parallel}	Laterale Feldstärke	V·m ⁻¹
E_A	Ionisierungsenergie für Akzeptoren, Akzeptorniveau	eV
E_C	Energieniveau der unteren Leitungsbandkante	eV
$E_{C,min}$	Leitungsbandminimum im E - k -Raum	eV
E_D	Ionisierungsenergie für Donatoren, Donatorniveau	eV
E_F	Ferminiveau	eV
E_G	Bandabstand	eV
E_{G0}	Bandabstand (zu $T = 0$ K extrapoliert)	eV
E_i	Eigenleitungs niveau	eV
E_M	Maximale Feldstärke	V·m ⁻¹
E_{Ph}	Photonenenergie	eV
E_{Sat}	Sättigungsfeldstärke	V·m ⁻¹
E_V	Energieniveau der oberen Valenzbandkante	eV
$E_{V,min}$	Valenzbandmaximum im E - k -Raum	eV
f	Frequenz	Hz
F	Füllfaktor	—
f_{α}	α -Grenzfrequenz	Hz
f_{β}	β -Grenzfrequenz	Hz
f_0	Resonanzfrequenz	Hz
f_G	Grenzfrequenz	Hz
f_{in}	Eingangssignalfrequenz	Hz
F_L	Lorentzkraft	N
f_T	Transitfrequenz	Hz
G	Gleichtaktunterdrückung	—
G	Generationsrate	m ⁻³ ·s ⁻¹
g	Zusätzliche Generationsrate	m ⁻³ ·s ⁻¹
g_0	Kleinsignalleitwert für $\omega \rightarrow 0$	Ω^{-1}
g_{BE}	Eingangsleitwert zwischen Basis und Emitter	Ω^{-1}
g_{CB}	Eingangsleitwert zwischen Kollektor und Basis	Ω^{-1}
g_{DS}	Ausgangsleitwert eines MOSFETs	Ω^{-1}
g_m	Steilheit	Ω^{-1}
G_T	Temperaturleitwert	W·K ⁻¹
G_{th}	Thermische Generationsrate	m ⁻³ ·s ⁻¹
\hbar	Plancksches Wirkungsquantum	J·s
	$\hbar = h/2\pi = 1,054 \cdot 10^{34}$ J·s	
I	Elektrische Stromstärke	A
i	Gesamtkonvektionsstromdichte	A·m ⁻²
I_{λ}	Relative spektrale Intensitätsverteilung	—
I_{2G}	Diodenstrom bei 2-Geraden-Näherung	A
I_A	Anodenstrom	A
I_B	Basisstrom	A
I_{BF}	Kippstrom	A
I_C	Kollektorstrom	A

Größe	Bedeutung	Einheit
I_D	Drainstrom	A
I_D	Strom am pn-Übergang nach Diffusionsmodell	A
I_{DSat}	Drain-Sättigungsstrom	A
I_{DSub}	Unterschwellstrom	A
I_E	Emitterstrom	A
I_{ES}	Transfersättigungsstrom	A
I_F	Strom in Flusspolung	A
I_G	Gatestrom	A
I_H	Haltestrom	A
I_{in}	Eingangsstrom	A
I_K	Kurzschlussstrom	A
I_L	Fotostrom	A
i_n	Konvektionsstromdichte der Elektronen	$\text{A} \cdot \text{m}^{-2}$
$i_{n,Diff}$	Diffusionsstromdichte der Elektronen	$\text{A} \cdot \text{m}^{-2}$
$i_{n,Drift}$	Driftstromdichte der Elektronen	$\text{A} \cdot \text{m}^{-2}$
I_{out}	Ausgangsstrom	A
i_p	Konvektionsstromdichte der Löcher	$\text{A} \cdot \text{m}^{-2}$
$i_{p,Diff}$	Diffusionsstromdichte der Löcher	$\text{A} \cdot \text{m}^{-2}$
$i_{p,Drift}$	Driftstromdichte der Löcher	$\text{A} \cdot \text{m}^{-2}$
I_R	Strom am pn-Übergang nach Rekombinations-/Generationsmodell	A
I_{RM}	Maximaler Sperrstrom	A
I_S	Sperrstrom	A
I_{SD}	Sperrsättigungsstrom nach Diffusionsmodell	A
I_{SR}	Sperrsättigungsstrom nach Rekombinations-/Generationsmodell	A
I_T	Transferstrom	A
I_V	Talwert des Diodenstroms einer Tunneldiode	A
k	k-Faktor von Dehnungsmessstreifen	—
k	Boltzmann-Konstante $k = 1,380 \cdot 10^{-23} \text{ J} \cdot \text{K}^{-1}$	$\text{J} \cdot \text{K}^{-1}$
k	Wellenzahl	m^{-1}
K_λ	Absolute Augenempfindlichkeit	$\text{lm} \cdot \text{W}^{-1}$
k_{B0}	Leerlaufempfindlichkeit bei Halleffekt	$\text{m}^2 \cdot \text{A}^{-1} \cdot \text{s}^{-1}$
k_f	Technologieabhängiger Fittingparameter	—
k_r	Technologieabhängiger Fittingparameter	—
L	Induktivität	H
L	Kanallänge eines MOSFETs	m
L_n	Diffusionslänge der Elektronen im p-Typ-Halbleiter	m
L_p	Diffusionslänge der Löcher im n-Typ-Halbleiter	m
ℓ_{BE}	Ausdehnung der Raumladungszone zwischen Basis und Emitter eines Bipolartransistors	m
ℓ_{CB}	Ausdehnung der Raumladungszone zwischen Kollektor und Basis eines Bipolartransistors	m
ℓ_D	Länge der Raumladungszone	m

Größe	Bedeutung	Einheit
ℓ_n	Ausdehnung der Raumladungszone im n-Typ-Halbleiter	m
ℓ_p	Ausdehnung der Raumladungszone im p-Typ-Halbleiter	m
m_n	Effektive Masse der Elektronen	kg
m_p	Effektive Masse der Löcher	kg
N	Nettodotierungsdichte	m^{-3}
n	Ladungsträgerdichte der Elektronen	m^{-3}
n'	Überschussträgerdichte der Elektronen	m^{-3}
n_0	Ladungsträgerdichte der Elektronen bei thermodynamischem Gleichgewicht	m^{-3}
N_A	Akzeptordotierungsdichte	m^{-3}
N_A^-	Dichte ionisierter Akzeptoren	m^{-3}
N_C	Effektive Zustandsdichte der Leitungsbandkante	m^{-3}
N_D	Donatordotierungsdichte	m^{-3}
N_D^+	Dichte ionisierter Donatoren	m^{-3}
n_E	Ladungsträgerverteilung der Elektronen	$J^1 \cdot m^{-3}$
n_i	Eigenleitungsträgerdichte (Si) $n_i = 1,5 \cdot 10^{10} \text{ cm}^{-3}$	m^{-3}
N_n	Zustandsdichte der Elektronen	m^{-3}
n_{n0}	Majoritätsträgerdichte im n-Typ-Halbleiter bei thermodynamischem Gleichgewicht	m^{-3}
N_p	Zustandsdichte der Löcher	m^{-3}
n_{p0}	Minoritätsträgerdichte im p-Typ-Halbleiter bei thermodynamischem Gleichgewicht	m^{-3}
n_S	Elektronendichte an der Grenzfläche Halbleiter/Oxid	m^{-3}
N_V	Effektive Zustandsdichte der Valenzbandkante	m^{-3}
p	Druck	Pa
p	Ladungsträgerdichte der Löcher	m^{-3}
p'	Überschussträgerdichte der Löcher	m^{-3}
p_0	Ladungsträgerdichte der Löcher bei thermodynamischem Gleichgewicht	m^{-3}
p_E	Ladungsträgerverteilung der Löcher	$J^1 \cdot m^{-3}$
P_{el}	Elektrische Leistung	W
P_L	Lichtleistung	W
p_{n0}	Minoritätsträgerdichte im n-Typ-Halbleiter bei thermodynamischem Gleichgewicht	m^{-3}
p_{p0}	Majoritätsträgerdichte im p-Typ-Halbleiter bei thermodynamischem Gleichgewicht	m^{-3}
p_S	Löcherdichte an der Grenzfläche Halbleiter/Oxid	m^{-3}
P_V	Verlustleistung	W
Q	Ladung	C
Q_{Akk}	Flächenladungsdichte in Akkumulation	$C \cdot m^{-2}$
Q_B	Basisladung	C
Q_D	Flächenladungsdichte in Depletion	$C \cdot m^{-2}$

Größe	Bedeutung	Einheit
Q_G	Gateladung	C
Q_I	Flächenladungsdichte in Inversion	$\text{C} \cdot \text{m}^{-2}$
Q_{SS}	Oberflächenladungsdichte	$\text{C} \cdot \text{m}^{-2}$
Q_T	Tail-Ladung	C
R	Elektrischer Widerstand	Ω
R	Rekombinationsrate	$\text{m}^{-3} \cdot \text{s}^{-1}$
r	Rekombinationskoeffizient	$\text{m}^3 \cdot \text{s}^{-1}$
R_B	Bahnwiderstand	Ω
r_B	Kleinsignalbahnwiderstand	Ω
r_{CE}	Kleinsignalausgangswiderstand (Kollektor-Emitter-Widerstand) eines Bipolartransistors	Ω
r_{DS}	Kleinsignalausgangswiderstand (Drain-Source-Widerstand) eines MOSFETs	Ω
R_{HE}	Heizwiderstand	Ω
R_{Hn}	Hallkonstante für n-Typ-Halbleiter	$\text{m}^3 \cdot \text{C}^{-1}$
R_{Hp}	Hallkonstante für p-Typ-Halbleiter	$\text{m}^3 \cdot \text{C}^{-1}$
R_i	Innenwiderstand	Ω
r_i	Kleinsignalinnennwiderstand	Ω
r_{in}	Eingangswiderstand	Ω
R_L	Lastwiderstand	Ω
R_{netto}	Nettorekombinationsrate	$\text{m}^{-3} \cdot \text{s}^{-1}$
R_{on}	ON-Widerstand	Ω
r_{out}	Ausgangswiderstand	Ω
R_P	Parallelwiderstand	Ω
R_{rad}	Strahlende Rekombinationsrate	$\text{m}^{-3} \cdot \text{s}^{-1}$
R_S	Sensorwiderstand	Ω
R_{th}	Thermischer Widerstand	$\text{K} \cdot \text{W}^{-1}$
S_λ	Spektrale Empfindlichkeit	$\text{A} \cdot \text{W}^{-1}$
t	Zeit	s
T	Temperatur	$\text{K}, \text{ }^\circ\text{C}$
T	Periodendauer	s
T_D	Taktverzögerung	s
t_{don}	Einschaltverzögerung	s
T_E	Eigenerwärmung	K
t_f	Abfallzeit	s
t_{fr}	Vorwärtserholzeit	s
T_G	Maximale Speicherzeit	s
t_{gd}	Zündverzugszeit des Thyristors	s
t_{gt}	Zündzeit des Thyristors	s
t_{ox}	Dicke des Gateoxids	m
t_q	Freiwerdezeit des Thyristors	s
t_r	Anstiegszeit	s
t_{rr}	Rückwärtserholzeit	s

Größe	Bedeutung	Einheit
t_s	Speicherzeit	s
V	Elektrische Spannung	V
V_ϑ	Temperaturspannung	V
V_0	Leerlaufspannung	V
V_{AK}	Anoden-Katoden-Spannung	V
V_B	Bulkspannung	V
V_{BE}	Basis-Emitter-Spannung	V
V_{BF}	Kippspannung	V
V_{Br}	Durchbruchspannung	V
V_{brss}	Brummspannung	V
V_{BS}	Bulk-Source-Spannung	V
V_{CB}	Kollektor-Basis-Spannung	V
V_{CC}, V_{DD}, V_B	Betriebsspannung	V
V_{CE}	Kollektor-Emitter-Spannung	V
V_{CESat}	Kollektor-Emitter-Sättigungsspannung	V
V_D	Diffusionsspannung	V
V_{DC}	Gleichspannung	V
V_{DS}	Drain-Source-Spannung	V
V_{DSat}	Drain-Source-Sättigungsspannung	V
V_{Early}	Early-Spannung	V
V_F	Flussspannung	V
V_{FB}	Flachbandspannung	V
V_{FP}	Einschaltüberspannung	V
V_G	Gatespannung	V
V_G	Gleichspannungsanteil	V
V_{GB}	Gate-Bulk-Spannung	V
V_{GD}	Gate-Drain-Spannung	V
V_{GE}	Gate-Emitter-Spannung	V
V_{GS}	Gate-Source-Spannung	V
V_H	Haltespannung des Thyristors	V
V_H	Hallspannung	V
V_{HE}	Heizspannung	V
V_{in}	Eingangsspannung	V
v_k	Mittlere Geschwindigkeit von Ladungsträgern	$\text{m}\cdot\text{s}^{-1}$
V_{Kanal}	Kanalspannung	V
v_n	Driftgeschwindigkeit der Elektronen	$\text{m}\cdot\text{s}^{-1}$
v_p	Driftgeschwindigkeit der Löcher	$\text{m}\cdot\text{s}^{-1}$
V_{off}	Offset-Spannung	V
V_{out}	Ausgangsspannung	V
V_R	Sperrspannung	V
V_{RM}	Maximale Sperrspannung	V
V_{RRM}	Rückschlagspannung	V
V_S	Schwellspannung	V
v_{Sat}	Geschwindigkeitssättigung	$\text{m}\cdot\text{s}^{-1}$

Größe	Bedeutung	Einheit
V_{SS}	Massepotential	V
V_T	Einsatzspannung MOSFET	V
V_{Tn}	Einsatzspannung n-Kanal-MOSFET	V
V_{Tn0}	Einsatzspannung n-Kanal-MOSFET bei $V_{BS} = 0$	V
V_{Tp}	Einsatzspannung p-Kanal-MOSFET	V
V_{Tp0}	Einsatzspannung p-Kanal-MOSFET bei $V_{BS} = 0$	V
V_Z	Zenerspannung	V
w	Basisweite	m
W	Kanalweite eines MOSFETs	m
W	Besetzungswahrscheinlichkeit	–
w_0	Neutrale Basisweite	m
x, y, z	Ortskoordinaten	m
X_G	Gaskonzentration	ppm
x_J	Tiefe des Source- und Draingebiets von MOSFETs	m
x_m	Ort des metallurgischen Übergangs	m
α	Differentieller Thermokoeffizient	$V \cdot K^{-1}$
α	Kleinsignalstromverstärkung in Basisschaltung	–
α_0	Kleinsignalstromverstärkung in Basisschaltung für $\omega \rightarrow 0$	–
α_A, α_B	Thermokoeffizienten der Materialien A und B	$V \cdot K^{-1}$
α_{AB}	Seebeck-Koeffizient	$V \cdot K^{-1}$
α_n	Ionisationskoeffizient für Elektronen	m^{-1}
α_p	Ionisationskoeffizient für Löcher	m^{-1}
β	Kleinsignalstromverstärkung in Emitterschaltung	–
β_0	Kleinsignalstromverstärkung in Emitterschaltung für $\omega \rightarrow 0$	–
χ_{HL}	Elektronenaffinität im Halbleiter	V
δ	Hallwinkel	°
δ	Basistransportfaktor	–
ϵ	Dielektrizitätskonstante $\epsilon = \epsilon_0 \cdot \epsilon_r$	$F \cdot m^{-1}$
ϵ_0	Vakuum-Dielektrizitätskonstante	$F \cdot m^{-1}$
	$\epsilon_0 = 8,854 \cdot 10^{-12} F \cdot m^{-1}$	
ϵ_{ox}	Dielektrizitätszahl von SiO_2 $\epsilon_{ox} = 3,9$	–
ϵ_r	Dielektrizitätszahl	–
ϵ_{Si}	Dielektrizitätszahl von Silizium $\epsilon_{Si} = 12$	–
ϕ	Elektrisches Potential	V
ϕ_F	Fermipotential	V
Φ_{Ph}	Photonenstrom	s^{-1}
ϕ_s	Oberflächenpotential	V
γ	Body-Faktor	$V^{1/2}$
γ_n	Body-Faktor für n-Kanal-MOSFET	$V^{1/2}$
γ_p	Body-Faktor für p-Kanal-MOSFET	$V^{1/2}$
η_Q	Quantenwirkungsgrad	–

Größe	Bedeutung	Einheit
η_P	Programmiereffizient	—
η_{rad}	Quantenausbeute	—
φ	Taktsignalspannung	V
φ_R	Phasenreserve	°
φ_S	Phasenverschiebung	°
κ	Absorptionsgrad	m^{-1}
λ	Kanallängenmodulationsfaktor	—
λ	Lichtwellenlänge	m
μ_{eff}	Effektive Beweglichkeit von Ladungsträgern	$m^2 \cdot V^{-1} \cdot s^{-1}$
μ_H	Hallbeweglichkeit	$m^2 \cdot V^{-1} \cdot s^{-1}$
μ_n	Beweglichkeit der Elektronen	$m^2 \cdot V^{-1} \cdot s^{-1}$
μ_p	Beweglichkeit der Löcher	$m^2 \cdot V^{-1} \cdot s^{-1}$
θ	Raumwinkel	sr
ρ	Raumladungsdichte	$C \cdot m^{-3}$
ρ	Spezifischer Widerstand	$\Omega \cdot cm$
ρ_n	Ladungsdichte der Elektronen	$C \cdot m^{-3}$
ρ_p	Ladungsdichte der Löcher	$C \cdot m^{-3}$
σ	Elektrische Leitfähigkeit	$\Omega^{-1} \cdot cm^{-1}$
σ_T	Thermische Leitfähigkeit	$W \cdot m^{-1} \cdot K^{-1}$
τ_B	Basistransitzeit	s
τ_D	Dielektrische Relaxationszeit	s
τ_d	Verzögerungszeit	s
τ_{df}	Verzögerungszeit der fallenden Flanke	s
τ_{dr}	Verzögerungszeit der steigenden Flanke	s
τ_n	Minoritätsträgerlebensdauer (der Elektronen) im p-Typ-Halbleiter	s
τ_p	Minoritätsträgerlebensdauer (der Löcher) im n-Typ-Halbleiter	s
τ_{RC}	RC-Zeitkonstante	s
τ_S	Effektive Ladungsträgerlebensdauer	s
τ_{SC}	Zeitkonstante von Schaltern	s
τ_T	Thermische Zeitkonstante	s
ω	Kreisfrequenz	Hz

1 Halbleiter-Grundlagen

Als *Halbleiter* bezeichnet man Materialien, deren spezifischer Widerstand zwischen dem der Leiter (Metalle) und dem der Isolatoren liegt. Der Ladungstransport im Halbleiter stellt die Grundlage für die Funktionsweise der *Halbleiterbauelemente* dar. In diesem Kapitel werden die Eigenschaften von Halbleitern an Hand des *Bindungsmodells* und *Bändermodells* beschrieben. Die Mechanismen des Ladungsträgertransports und das Verhalten des Halbleiters bei Störung des thermodynamischen Gleichgewichts sowie spezielle Halbleitereffekte werden dargestellt. Diese Betrachtungen bilden die Grundlage für das Verständnis der in den nachfolgenden Kapiteln beschriebenen Halbleiterbauelementen und deren schaltungstechnischer Anwendung.

1.1 Halbleitermaterialien

Halbleitermaterialien zeichnen sich insbesondere durch den Wert des spezifischen Widerstandes ρ bzw. der elektrischen Leitfähigkeit σ aus ($\rho = 1/\sigma$), der bei Raumtemperatur ($T = 300\text{ K}$) zwischen dem der Leiter und dem der Isolatoren liegt ($10^{-4}\Omega\cdot\text{cm} < \rho < 10^9\Omega\cdot\text{cm}$). In Abb. 1.1 ist der spezifische Widerstand ρ verschiedener Materialien und deren Zuordnung zu Leitern, Halbleitern und Isolatoren dargestellt. Die aufgeführten Materialien sind typische Stoffe, die zur Herstellung von Halbleiterbauelementen verwendet werden.

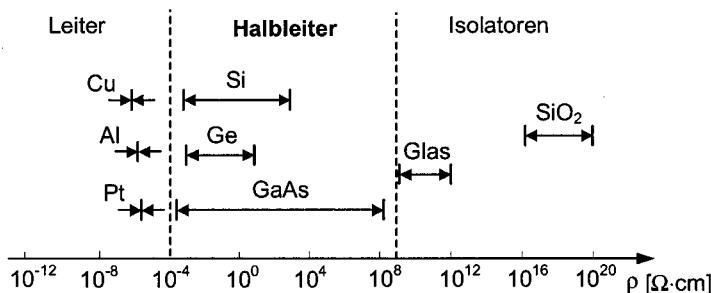


Abb. 1.1. Spezifischer Widerstand ρ von typischen Leitern, Halbleitern und Isolatoren bei Raumtemperatur

Im Gegensatz zu Leitern (Metallen), deren spezifischer Widerstand bei Temperaturabnahme absinkt, nimmt bei reinen Halbleitern der spezifische Widerstand

bei Temperaturverringerung zu. Bei einer Temperatur von $T = 0\text{ K}$ verhalten sich Halbleiter dann hinsichtlich ihrer elektrischen Leitfähigkeit bzw. ihres spezifischen Widerstandes wie Isolatoren.

Die elektrische Leitfähigkeit von Halbleitern kann durch *Dotierung* mit Störstellenatomen variiert und bis in den Bereich der Leiter gesteigert werden. Abhängig von der Dotierungsart (vgl. Abschn. 1.2) existieren im Halbleiter eine sog. *Elektronen-Leitfähigkeit* und eine *Löcher-Leitfähigkeit*, die zusammen die elektrische Leitfähigkeit ergeben. Halbleiter, in denen die Elektronen-Leitfähigkeit dominiert, werden als sog. *n-Typ-Halbleiter* bezeichnet; die Elektronen sind hier die *Majoritätsträger*. Im n-Typ-Halbleiter existiert aber auch eine Löcher-Leitfähigkeit, wobei die Löcher als *Minoritätsträger* bezeichnet werden. Beim *p-Typ-Halbleiter* gilt, dass Löcher die Majoritätsträger und Elektronen die Minoritätsträger sind. Durch Kombination von n-Typ- und p-Typ-Halbleiterschichten können elektronische Bauelemente realisiert werden, bei denen die Wechselwirkungen zwischen Elektronen und Löchern ausgenutzt werden.

Neben der Abhangigkeit der elektrischen Leitfahigkeit von der Dotierung und der Temperatur weisen Halbleiter eine starke Abhangigkeit von elektrischen und magnetischen Feldern sowie von optischen und mechanischen Einflussen auf. Die aus diesen Abhangigkeiten resultierenden Halbleitereffekte werden in entsprechenden Halbleiterbauelementen ausgenutzt, deren wesentliche Ausfuhrungen in diesem Buch beschrieben werden.

Die Halbleitermaterialien werden in sog. *Elementhalbleiter* und *Verbindungs- halbleiter* unterteilt. Elementhalbleiter wie Silizium (Si) und Germanium (Ge) sind aus einem der Elemente der IV. Hauptgruppe des Periodensystems der Elemente (vgl. Abb. 1.2) aufgebaut und weisen eine sog. *Diamantstruktur* auf (vgl. Abschn. 1.2).

I	II	III	IV	V	VI	VII	VIII
	B 5 10,82						Elementhalbleiter
	Al 13 26,97	Si 14 28,06	P 15 31,02	S 16 32,06			
Zn 30 65,38	Ga 31 69,72	Ge 32 72,64	As 33 74,91	Se 34 78,96			
Cd 48 112,41	In 49 114,76	Sn 50 118,71	Sb 51 121,76	Te 52 127,60			

Abb. 1.2. Auszug aus dem Periodensystem der Elemente mit typischen Stoffen für Elementhalbleiter und Komponenten für Verbindungshalbleiter

Im Gegensatz zu den Elementhalbleitern bestehen Verbindungshalbleiter aus mindestens zwei Elementen. Sie werden entsprechend der Anzahl der beteiligten Komponenten als *binäre*, *ternäre* oder *quaternäre Halbleiter* bezeichnet. Typische binäre Halbleiter sind in der Form $A^{IV}B^{IV}$ (z. B. SiGe), $A^{III}B^V$ (z. B. GaAs) oder $A^{II}B^{VI}$ (z. B. CdSe) aufgebaut, wobei die Elemente A und B der jeweils entsprechenden Hauptgruppe des Periodensystems zu entnehmen sind. Ternäre Halbleiter bestehen aus drei Komponenten (z. B. $\text{Ga}_{1-x}\text{Al}_x\text{As}$) und quaternäre Halbleiter aus

vier Elementen (z. B. $In_xGa_{1-x}As_yP_{1-y}$). Die elektrische Leitfähigkeit von Verbindungshalbleitern kann durch Stöchiometriestörungen oder gezielte Dotierung mit weiteren Komponenten beeinflusst werden. Wird beispielsweise der binäre Verbindungshalbleiter GaAs mit einem Element der II. Hauptgruppe (z. B. Zn) dotiert, so entsteht ein p-Typ-Halbleiter. Wird GaAs hingegen mit einem Element der VI. Hauptgruppe (z. B. Se) dotiert, so entsteht ein n-Typ-Halbleiter.

Auf die technologische Herstellung von Halbleiterstrukturen wird in den Folgekapiteln bei der Betrachtung elektronischer Bauelemente näher eingegangen (vgl. Abschn. 2.4, 3.4, 4.6).

1.2 Bindungsmodell

Das Bindungsmodell von Halbleitern beschreibt die grundlegenden Mechanismen, die am Ladungsträgertransport im Halbleiter beteiligt sind. Ausgehend von der Gitterstruktur von Halbleitern wird die im reinen Halbleiter auftretende *Eigenleitung* und die durch Dotierung erzeugte *Störstellenleitung* beschrieben. Die Betrachtungen werden hauptsächlich am Halbleitermaterial Silizium durchgeführt, da Silizium in der Mikroelektronik derzeit die größte Bedeutung besitzt.

1.2.1 Gitterstruktur

Halbleitermaterialien (vgl. Abschn. 1.1) werden zumeist in *einkristalliner* Form verwendet, die sich durch eine regelmäßige, räumlich wiederkehrende Atomanordnung auszeichnet. Die sich räumlich wiederholenden Grundstrukturen werden *Elementarzellen* genannt. Die chemische Bindung der Atome untereinander wird dabei insbesondere durch die Anzahl der Valenzelektronen der den Atomkern umgebenden Elektronenhülle festgelegt. Ein Silizium-Atom besitzt vier Valenzelektronen. Die Valenzen sind so ausgerichtet, dass sie zu vier flächendiagonal gegenüberliegenden Eckpunkten eines Würfels zeigen, in dessen Zentrum sich das Atom befindet [1.1]. Das daraus resultierende Diamantgitter kennzeichnet die Gitterstruktur von Silizium, deren Elementarzelle ein Würfel ist. Die Gitterstruktur von Silizium ist in Abb. 1.3a schematisch dargestellt.

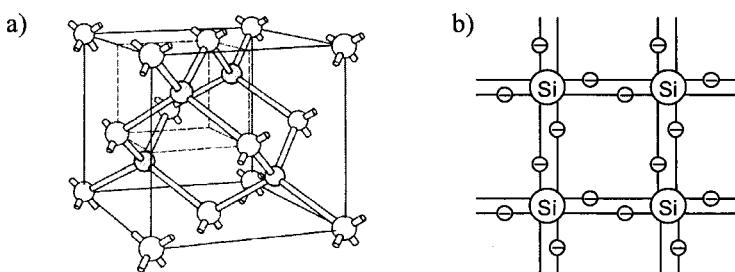


Abb. 1.3. Gitterstruktur von Silizium (Diamantgitter) a) Räumliche Darstellung nach [1.1] und b) Ebenes Schema

Die Atome im Kristallgitter sind so angeordnet, dass jedes Atom ein Valenzelektron mit jedem der vier Nachbaratome teilt. Zur Vereinfachung der Darstellung verwendet man das in Abb. 1.3b abgebildete ebene Schema der Gitterstruktur.

Die Wechselwirkung zwischen den Atomen im Kristallgitter wird charakterisiert durch abstoßende Kräfte zwischen den Atomkernen und anziehende Kräfte durch die Wechselwirkung der Valenzelektronen mit den Atomkernen. Die Bindungen im Kristallgitter ergeben einen optimalen Abstand der Atome, der durch den Zustand minimaler Energie charakterisiert wird. Die Bindungsenergie E im Kristallgitter in Abhängigkeit vom Abstand x der Atomkerne ist in Abb. 1.4 dargestellt.

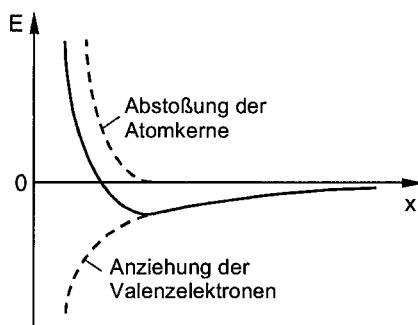


Abb. 1.4. Bindungsenergie der Atome im Kristallgitter eines Halbleiters in Abhängigkeit vom Atomabstand

Auf Grund des aus der Bindungsenergie resultierenden, stabilen Abstandes der Atome im Kristallgitter und durch die gerichteten Valenzen entsteht unter geeigneten Bedingungen bei der Halbleitermaterialherstellung ein regelmäßiger Aufbau der Kristallstruktur.

1.2.2 Eigenleitung

Die Eigenleitung tritt in vollkommen reinen Halbleitern auf. Diese Halbleiter werden *eigenleitende* oder *intrinsische* Halbleiter genannt. Durch thermische Gitterschwingungen brechen Bindungen des Kristallgitters auf und es entstehen *Elektron-Loch-Paare*. Dabei werden Valenzelektronen zu freien Elektronen (*Leitungselektronen*), die im Kristallgitter einen positiv geladenen Atomrumpf (*Loch*) hinterlassen. Wird dieses Loch durch ein Valenzelektron einer Nachbarbindung besetzt, so verschiebt sich das Loch an den neu entstandenen Platz der Nachbarbindung. Unter Einwirkung eines elektrischen Feldes ist also ein Stromfluss sowohl von Leitungselektronen als auch von Löchern möglich. Dieser als Eigenleitung bezeichnete Mechanismus beschreibt somit die Bewegung von Ladungsträgern im Kristall, wobei die *Ladungsträgerdichte* n der Elektronen und die Ladungsträgerdichte p der Löcher im Halbleitervolumen gleich ist ($n = p$). Das

Prinzip der Eigenleitung ist in Abb. 1.6 im ebenen Schema der Silizium-Gitterstruktur veranschaulicht.

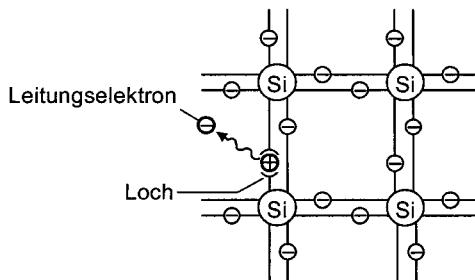


Abb. 1.5. Prinzip der Eigenleitung eines Halbleiters im ebenen Schema der Silizium-Gitterstruktur

Die zur Bildung eines Elektron-Loch-Paars erforderliche Mindestenergie ist vom Halbleitermaterial und der Temperatur abhängig. Die Mindestenergie E_G beträgt bei $T = 300\text{ K}$ für Silizium $E_G = 1,1\text{ eV}$, für Germanium $E_G = 0,7\text{ eV}$ und für GaAs $E_G = 1,4\text{ eV}$.

Führt man dem Halbleiter Energie zu (z. B. thermisch oder optisch), so werden durch diese Energie Elektron-Loch-Paare generiert. Dieser Mechanismus wird *Generation* genannt. Der Generation von Elektron-Loch-Paaren wirkt nach Erzeugung einer bestimmten Ladungsträgeranzahl ein gegenläufiger Prozess entgegen. Dieser Prozess wird *Rekombination* genannt. Bei der Rekombination wird ein Leitungselektron von einem Loch eingefangen, wobei das Loch und das Elektron rekombinieren und Energie (z. B. thermisch oder optisch) freigesetzt wird. Bei konstanter Temperatur werden ständig Elektron-Loch-Paare generiert, wobei bereits generierte Ladungsträger gleichzeitig wieder rekombinieren. Bei konstanter Temperatur und ohne weitere Störungen durch z. B. elektrische oder magnetische Felder bildet sich dann ein Gleichgewicht aus, das als *thermodynamisches Gleichgewicht* bezeichnet wird. Die Ladungsträgerdichte n_0 der Elektronen und die Ladungsträgerdichte p_0 der Löcher im Halbleiter bei thermodynamischem Gleichgewicht ergeben sich zu:

$$n_0 = p_0 = n_i(T). \quad (1.1)$$

Dabei wird von einem *homogenen* Halbleiter ausgegangen, d.h. im Halbleiter existiert eine örtliche Gleichverteilung der Ladungsträger. Der Index Null der Ladungsträgerdichten n_0 und p_0 kennzeichnet den Fall des thermodynamischen Gleichgewichts. Die Größe n_i wird als *Eigenleitungsträgerdichte* oder intrinsische Ladungsträgerdichte bezeichnet. Sie beträgt bei $T = 300\text{ K}$ für Silizium $n_i \approx 1,5 \cdot 10^{10}\text{ cm}^{-3}$, für Germanium $n_i \approx 2,5 \cdot 10^{13}\text{ cm}^{-3}$ und für GaAs $n_i \approx 1,8 \cdot 10^6\text{ cm}^{-3}$. Die Eigenleitungsträgerdichte ist stark von der Temperatur abhängig. Diese Temperaturabhängigkeit ist in Abb. 1.6 für verschiedene Halbleitermaterialien abgebildet.

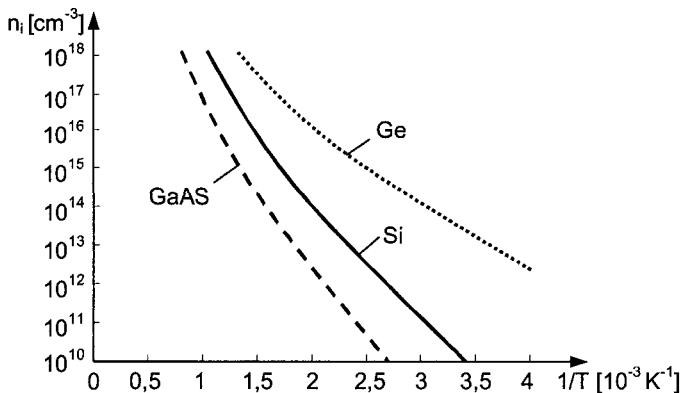


Abb. 1.6. Temperaturabhängigkeit der Eigenleitungsträgerdichte n_i für Ge, Si und GaAs

Die *Generationsrate* G definiert die Anzahl der erzeugten Elektron-Loch-Paare pro Zeit- und Volumeneinheit. Die thermische Generation ist in folgender Weise zum Bandabstand E_G und zum Produkt aus Temperatur T und Boltzmann-Konstante k proportional:

$$G(E_G, T) \sim \exp\left(-\frac{E_G}{k \cdot T}\right). \quad (1.2)$$

Die Rekombinationsrate R ist ebenfalls von der Temperatur abhängig und wird mit Hilfe der Ladungsträgerdichten und des *Rekombinationskoeffizienten* r wie folgt bestimmt:

$$R(T, n, p) = r(T) \cdot n \cdot p. \quad (1.3)$$

Der Rekombinationskoeffizient ist eine temperaturabhängige Materialkonstante, die das Produkt aus dem Wirkungsquerschnitt der Elektron-Loch-Wechselwirkung und der mittleren thermischen Geschwindigkeit der Ladungsträger darstellt [1.2].

Eine ungebrochene Bindung zerfällt bei Zuführen der Energie E_G in ein Leitungselektron und ein Loch. Umgekehrt gilt, dass ein Elektron-Loch-Paar bei Abgabe der Energie E_G in eine ungebrochene Bindung übergeht. Im thermodynamischen Gleichgewicht ist die Generationsrate gleich der Rekombinationsrate ($G = R = r(T) \cdot n_0 \cdot p_0$) und es gilt:

$$n_0 \cdot p_0 = \frac{G(T)}{r(T)} = n_i^2(T). \quad (1.4)$$

Gl. (1.4) gilt nicht nur im Fall der Eigenleitung, sondern allgemein im thermodynamischen Gleichgewicht. Aus Gl. (1.4) resultiert weiterhin, dass die Erhöhung einer Ladungsträgerdichte (z. B. n_0) zur Verringerung der anderen Ladungsträgerdichte (p_0) führt. Dieser Aspekt wird in Abschn. 1.2.3 bei der Betrachtung des Störstellenhalbleiters genutzt.

1.2.3 Störstellenleitung

Neben der thermischen oder optischen Generation von Elektron-Loch-Paaren können Ladungsträger auch durch Störstellen im Halbleiter erzeugt werden (Störstellenleitung). Halbleiter, in denen überwiegend Störstellenleitung auftritt, werden als Störstellenhalbleiter bezeichnet. Die Störstellen werden durch sog. *Dotierung* in den Halbleiter eingebracht. Bei Dotierung mit einem Element der V. Hauptgruppe des Periodensystems wird die Elektronenkonzentration im Halbleiter erhöht (n-Typ-Halbleiter). Durch Dotierung des Halbleiters mit einem Element der III. Hauptgruppe des Periodensystems wird die Löcherkonzentration erhöht (p-Typ-Halbleiter). Die Störstellen des n-Typ-Halbleiters werden *Donatoren*, die des p-Typ-Halbleiters werden *Akzeptoren* genannt. Die Störstellenleitung ist in Abb. 1.7a für einen n-Typ-Halbleiter und in Abb. 1.7b für einen p-Typ-Halbleiter am ebenen Schema der Silizium-Gitterstruktur veranschaulicht.

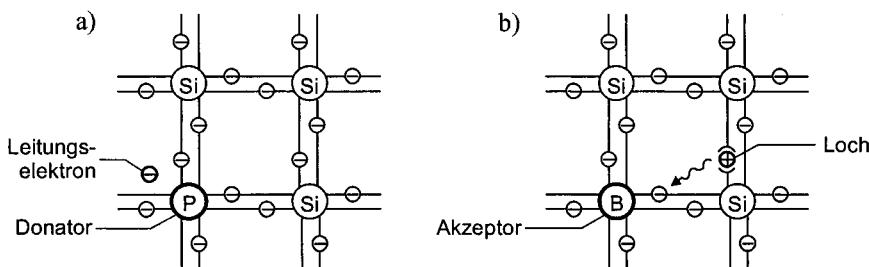


Abb. 1.7. Störstellenleitung in der ebenen Silizium-Gitterstruktur a) n-Typ-Halbleiter am Beispiel Phosphor (P)-Dotierung und b) p-Typ-Halbleiter am Beispiel Bohr (B)-Dotierung

Aus Abb. 1.7a wird ersichtlich, dass eine fünfwertige Störstelle mit fünf Valenzelektronen nur vier Valenzelektronen zur Bindung im Kristallgitter benötigt. Das überschüssige fünfte Valenzelektron kann durch Zufuhr einer relativ geringen Ionisierungsenergie ΔE_D (z. B. $\Delta E_D \approx 0,05$ eV) freigesetzt werden. Dabei entsteht aus einem neutralen Donator eine feste positive Ladung (ionisierter Donator N_D^+) und ein Leitungselektron. Umgekehrt entsteht aus einem ionisierten Donator und einem Leitungselektron unter Abgabe der Energie ΔE_D ein neutraler Donator. Bei Raumtemperatur $T = 300\text{K}$ und hoher *Donatordotierungsdichte* N_D gegenüber der Eigenleitungsträgerdichte n_i gilt für die Majoritätsträgerdichte n_{n0} im n-Typ-Halbleiter:

$$n_{n0} \approx N_D^+ \approx N_D. \quad (1.5)$$

Die Minoritätsträgerdichte p_{n0} im n-Typ-Halbleiter kann mit Hilfe der Eigenleitungsträgerdichte n_i und Gl. (1.4) wie folgt bestimmt werden:

$$p_{n0} \approx \frac{n_i^2}{N_D}. \quad (1.6)$$

Beim p-Typ-Halbleiter (vgl. Abb. 1.7b) wird durch eine dreiwertige Störstelle die Bindung nicht gesättigt, da ein Valenzelektron fehlt. Die somit existierende Leerstelle kann durch ein Valenzelektron der Nachbarbindung besetzt werden. Dabei erfolgt die Ionisation eines neutralen Akzeptors unter Zufuhr der Energie ΔE_A , wobei ein ionisierter Akzeptor N_A^- und ein Loch entsteht. Umgekehrt entsteht aus einem ionisierten Akzeptor und einem Loch unter Abgabe der Energie ΔE_A ein neutraler Akzeptor. Bei Raumtemperatur $T = 300\text{ K}$ und hoher *Akzeptordotierungsdichte* $N_A \gg n_i$ gilt für die Majoritätsträgerdichte p_{p0} im p-Typ-Halbleiter:

$$p_{p0} \approx N_A^- \approx N_A. \quad (1.7)$$

Für die Minoritätsträgerdichte n_{p0} eines n-Typ-Halbleiters im thermodynamischen Gleichgewicht folgt gemäß Gl. (1.4):

$$n_{n0} \approx \frac{n_i^2}{N_A}. \quad (1.8)$$

Existieren auf Grund der Dotierung sowohl Akzeptoren als auch Donatoren im Halbleiter in gleicher Konzentration, so kompensieren sich ihre Einflüsse. Der Leitfähigkeitsotyp (n- oder p-leitend) wird durch die höhere Dotierungsdichte festgelegt. Die Ladungsträgerdichte wird durch die sog. *Nettodotierung* $N = |N_D - N_A|$ bestimmt.

Bei vollständiger Ionisation der Dotierstoffatome ($N_D^+ = N_D$ und $N_A^- = N_A$) gilt im elektrisch neutralen, homogen dotierten Halbleiter bei Raumtemperatur die *Neutralitätsbedingung*:

$$\boxed{n_0 - p_0 = N_D - N_A} \quad \text{und} \quad \boxed{n_0 \cdot p_0 = n_i^2}. \quad (1.9)$$

Im dotierten Halbleiter ist die Minoritätsträgerdichte stets kleiner als die Eigenleitungsdichte n_i .

1.3 Bändermodell

Das Bändermodell ermöglicht die detaillierte Beschreibung der Energiezustände von Ladungsträgern im Halbleiter, die durch quantenmechanische Wechselwirkungen im Kristall und den sog. *Welle-Teilchen-Dualismus* der Elektronen bestimmt sind. Elektronen können in Atomen nur bestimmte *Energieniveaus* besetzen. Auf Grund der hohen Anzahl der im Festkörper in Wechselwirkung stehenden Atome spalten sich diese Energieniveaus auf und werden in kontinuierlichen *Energiebändern* zusammengefasst. Die Energiebänder sind durch sog. *verbotene Zonen* getrennt, die nicht von Elektronen besetzt werden können. Für die Leitungsvorgänge im Festkörper sind das *Leitungsband* und das darunter liegende *Valenzband* von Bedeutung. Die Energiebänder werden im Bändermodell zusammengefasst, wobei die besetzbaren Energieniveaus über dem Ort dargestellt werden.

1.3.1 Aufbau des Bändermodells

Das Bändermodell wird durch das Leitungsband mit seiner unteren Energiebandkante E_C und das Valenzband mit seiner oberen Energiebandkante E_V charakterisiert. Dabei beschreibt das Leitungsband die kinetische Energie der Elektronen und das Valenzband die kinetische Energie der Löcher über dem Ort x (im homogenen Halbleiter ohne externe Felder besteht im Kristall keine Abhängigkeit vom Ort). Das Bändermodell eines eigenleitenden Halbleiters ist in Abb. 1.8 dargestellt.

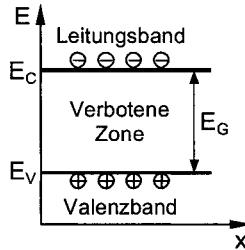


Abb. 1.8. Bändermodell eines eigenleitenden Halbleiters mit Lokalisierung von Leitungselektronen (Θ) und Löchern (\oplus) mit Kennzeichnung der oberen Valenzbandkante E_V , der unteren Leitungsbandkante E_C und des Bandabstandes E_G (verbotene Zone)

Der Abstand zwischen der unteren Leitungsbandkante E_C und der oberen Valenzbandkante E_V wird als Bandabstand E_G bezeichnet und entspricht der in Abschn. 1.2.2 beschriebenen Mindestenergie zur Bildung eines Elektron-Loch-Paares. Der Energiebereich des Bandabstandes (verbotene Zone) kann nicht durch freie Ladungsträger besetzt werden. Für den Bandabstand gilt:

$$E_G = E_C - E_V. \quad (1.10)$$

Bei eigenleitenden Halbleitern (vgl. Abschn. 1.2.2) ist bei einer Temperatur von $T = 0\text{ K}$ das Valenzband mit Elektronen vollständig besetzt und das Leitungsband ist unbesetzt. Bei höheren Temperaturen ($T > 0\text{ K}$) werden Elektronen durch thermische oder optische Generation energetisch aus dem Valenzband in das Leitungsband angehoben. Durch die Bildung von Elektron-Loch-Paaren (vgl. Abschn. 1.2.2) existieren dann die gleiche Anzahl von Elektronen im Leitungsband und Löchern im Valenzband. Bei der Rekombination fallen Elektronen dann energetisch in das Valenzband zurück, wobei thermische oder optische Energie abgegeben wird. Der Abstand eines Zustands im Leitungsband von der Bandkante E_C wird durch die kinetische Energie des Leitungselektrons bestimmt. Im Valenzband entspricht die Differenz zwischen dem Energieniveau E_V und einem nicht besetzten Elektronenzustand der kinetischen Energie des entsprechenden Loches.

Bei Störstellenhalbleitern (vgl. Abschn. 1.2.3) existieren in der verbotenen Zone diskrete Energiezustände, die bei n-Typ-Halbleitern als *Donatorniveau* und beim p-Typ-Halbleiter als *Akzeptorniveau* bezeichnet werden. In Abb. 1.9 sind die resultierenden Bändermodelle eines n-Typ- und p-Typ-Halbleiters dargestellt.

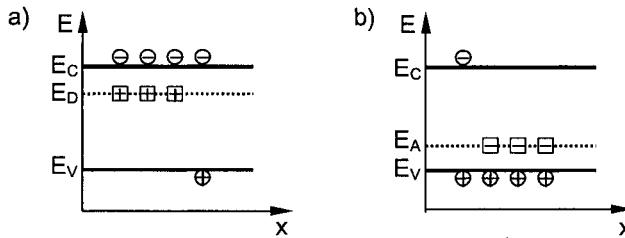


Abb. 1.9. Bändermodelle für Störstellenleitung mit Lokalisierung von Leitungselektronen (\ominus) und Löchern (\oplus) sowie ortsfesten ionisierten Donatoren (■) und Akzeptoren (□) für a) n-Typ-Halbleiter mit Donatorniveau E_D und b) p-Typ-Halbleiter mit Akzeptorniveau E_A

Bei n-Typ-Halbleitern befindet sich das Donatorniveau E_D nahe der unteren Leitungsbandkante und es genügt eine Energie von $\Delta E_D = E_C - E_D \approx 0,05$ eV, um Donatoren zu ionisieren (vgl. Abschn. 1.2.3). Dabei werden Elektronen in das Leitungsband abgegeben und gleichzeitig die ortsfesten Donatorstörstellen positiv aufgeladen. Dadurch entsteht die sog. *n-Leitung* ($n_0 > p_0$). Bei Raumtemperatur sind fast alle Donatoren im Störstellenhalbleiter ionisiert (*vollständige Ionisation*).

Bei p-Typ-Halbleitern befindet sich das Akzeptorniveau E_A nahe der oberen Valenzbandkante. Zur Ionisation von Akzeptoren genügt eine Energie von $\Delta E_A = E_A - E_V \approx 0,05$ eV. Bei der Ionisierung werden Elektronen energetisch aus dem Valenzband auf das Akzeptorniveau angehoben, wobei im Valenzband Löcher entstehen und die Akzeptorstörstellen negativ geladen werden. Daraus ergibt sich die sog. *p-Leitung* ($p_0 > n_0$). Bei Raumtemperatur sind auch nahezu alle Akzeptoren ionisiert.

Die Generations- und Rekombinationsmechanismen sind in Abb. 1.10a für einen eigenleitenden Halbleiter und in Abb. 1.10b für einen n-Typ-Störstellenhalbleiter veranschaulicht. Dabei gilt das Prinzip des detaillierten Gleichgewichts, d. h. es treten gegenläufige Prozesse auf (Generation G und Rekombination R), die sich bei thermodynamischem Gleichgewicht kompensieren.

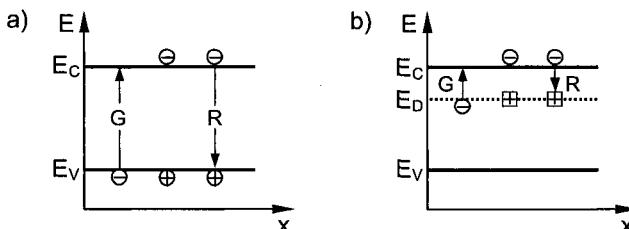


Abb. 1.10. Bändermodell zur Veranschaulichung von Generations- (G) und Rekombinations- (R) Prozessen im a) eigenleitenden Halbleiter und b) n-Typ-Halbleiter

Bei Anlegen eines elektrischen Feldes am Halbleiter werden durch die ortsabhängige potentielle Energie (entsprechend des Potentialverlaufes) die Bandkanten des Bändermodells geneigt; es entsteht eine sog. *Bandverbiegung*. Die Bandverbiegung kennzeichnet die vom Ort x der Halbleitergeometrie abhängigen Energiebandverläufe.

1.3.2 Bänderstruktur im Energie-Impuls-Raum

Zur Beschreibung spezieller Halbleitereigenschaften ist es notwendig, das Wellenverhalten der Elektronen zu berücksichtigen. Dies kann durch Darstellung der Bänderstruktur im Energie-Impuls (E - k)-Raum erfolgen. Dabei werden die Elektronenzustände in einem Band abhängig von der Wellenzahl k (Kristallimpuls) dargestellt. Abhängig von der Lage der Energiemaxima und -minima der Bänder unterscheidet man *direkte Halbleiter* und *indirekte Halbleiter*. Beim direkten Halbleiter liegen das Valenzbandmaximum $E_{V,max}(k)$ und das Leitungsbandminimum $E_{C,min}(k)$ bei demselben k -Wert. Bei indirekten Halbleitern ist dies nicht der Fall. In Abb. 1.11 sind die Bänderstrukturen im E - k -Raum eines direkten und eines indirekten Halbleiters dargestellt.

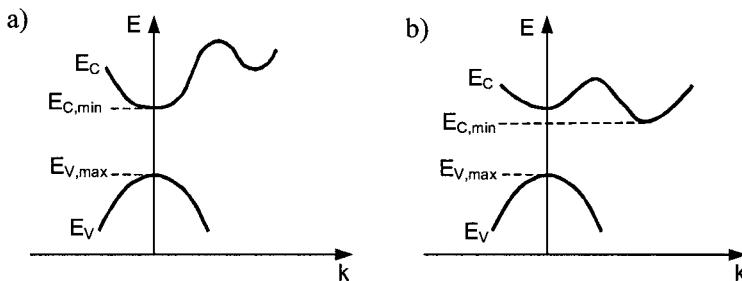


Abb. 1.11. Bänderstruktur im Energie-Impuls (E - k)-Raum a) eines direkten Halbleiters und b) eines indirekten Halbleiters

Die Bandverläufe der Bänderstruktur im E - k -Raum können in der Nähe der Extrema durch Parabeln angenähert werden. Die Bandverläufe werden hier durch die *effektiven Massen* der Ladungsträger bestimmt. Die effektive Masse m_n der Elektronen im Leitungsband und die effektive Masse m_p der Löcher im Valenzband können mit Hilfe der Wellenzahl k , des Planckschen Wirkungsquantums \hbar und der Energieniveaus E_C bzw. E_V wie folgt bestimmt werden:

$$m_n = \hbar^2 \cdot \left(\frac{d^2 E_C}{dk^2} \right)^{-1} \quad \text{und} \quad m_p = -\hbar^2 \cdot \left(\frac{d^2 E_V}{dk^2} \right)^{-1}. \quad (1.11)$$

Die Bänderstruktur im E - k -Raum wird beispielsweise für die Charakterisierung der Eigenschaften von lichtemittierenden Bauelementen (vgl. Abschn. 2.7.8) verwendet.

1.3.3 Bänderbesetzung

Die Bänderbesetzung eines Halbleiters gibt an, mit welcher Anzahl und Dichte die Energiebänder mit Ladungsträgern besetzt sind. Zur Charakterisierung der Bänderbesetzung werden insbesondere die Zustandsdichte, die Besetzungswahrscheinlichkeit und die Ladungsträgerverteilung verwendet.

1.3.3.1 Zustandsdichte

Die Ladungsträgeranzahl bzw. Ladungsträgerdichte in einem Energieband wird von der Anzahl der besetzbaren Energieplätze bestimmt, die von der sog. *Zustandsdichte* N_n für Elektronen im Leitungsband und der Zustandsdichte N_p für Löcher im Valenzband abhängt. Der Verlauf dieser Zustandsdichten ist in Abb. 1.12 veranschaulicht.

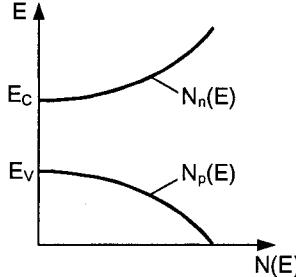


Abb. 1.12. Zustandsdichte N_n für Elektronen im Leitungsband und Zustandsdichte N_p für Löcher im Valenzband eines eigenleitenden Halbleiters

Der energieabhängige Verlauf der Zustandsdichte $N_n(E)$ für Elektronen und $N_p(E)$ für Löcher berechnet sich mit Hilfe der effektiven Masse m_n und m_p der Elektronen und Löcher (vgl. Abschn. 1.3.2) näherungsweise zu [1.2]:

$$N_n(E)dE \approx \frac{2\pi \cdot (2 \cdot m_n)^{3/2}}{h^3} \cdot \sqrt{E - E_C} dE, \quad (1.12)$$

$$N_p(E)dE \approx \frac{2\pi \cdot (2 \cdot m_p)^{3/2}}{h^3} \cdot \sqrt{E_V - E} dE. \quad (1.13)$$

Aus den Zustandsdichten in Gl. (1.12) und Gl. (1.13) können die *effektiven Zustandsdichten* N_C und N_V bestimmt werden, die die räumliche Zustandsdichte des Leitungs- und Valenzbandes nahe der Bandkanten beschreiben. Die effektiven Zustandsdichten berechnen sich wie folgt:

$$N_C = 2 \cdot \left(\frac{2\pi \cdot m_n \cdot kT}{h^2} \right)^{3/2} \quad \text{und} \quad N_V = 2 \cdot \left(\frac{2\pi \cdot m_p \cdot kT}{h^2} \right)^{3/2}. \quad (1.14)$$

Für Silizium bei Raumtemperatur betragen die effektiven Zustandsdichten $N_C = 2,92 \cdot 10^{19} \text{ cm}^{-3}$ und $N_V = 1,15 \cdot 10^{19} \text{ cm}^{-3}$. Die Zustandsdichten werden zur Bestimmung der Ladungsträgerverteilung im Halbleiter (vgl. Abschn. 1.3.3.3) verwendet.

1.3.3.2 Besetzungswahrscheinlichkeit

Die Wahrscheinlichkeit, mit der bestimmte quantenmechanische Energiezustände im Halbleiter besetzt sind, wird durch die sog. *Besetzungswahrscheinlichkeit* $W(E)$

angegeben. Dabei beschreibt $W(E)$ die Wahrscheinlichkeit für das Antreffen eines Elektrons in einem Zustand der Energie E . Die Funktion $1-W(E)$ hingegen bezeichnet die Wahrscheinlichkeit für das Antreffen eines Loches in einem Energiezustand E .

Elektronen und Löcher sind sog. *Fermiteilchen*, die der *Fermiverteilung* unterliegen. Die Grundlage für die Fermiverteilung besteht darin, dass ein entsprechend der quantenmechanischen Unschärferelation definierter Energiezustand nur von einem Fermiteilchen besetzt werden kann, d. h. $0 \leq W(E) \leq 1$. Die Fermiverteilung wird durch die Temperatur und das sog. *Ferminiveau* E_F bestimmt. Das Ferminiveau kennzeichnet die Energie, bei der die Besetzungswahrscheinlichkeit $W(E)$ den Wert $1/2$ besitzt. Bei eigenleitenden Halbleitern sind das Leitungs- und Valenzband mit der gleichen Anzahl von Elektronen und Löchern besetzt, sodass das Ferminiveau etwa in der Bandmitte (bei einem Wert E_i) liegt. Das Ferminiveau eines eigenleitenden Halbleiters berechnet sich dann mit Hilfe der Energieniveaus E_C und E_V sowie der effektiven Zustandsdichten N_C und N_V wie folgt:

$$E_F = E_i = \frac{E_C + E_V}{2} + \frac{kT}{2} \cdot \ln\left(\frac{N_V}{N_C}\right) \approx \frac{E_C + E_V}{2}. \quad (1.14)$$

Das Energieniveau E_i wird auch als *Eigenleitungs-Ferminiveau* bezeichnet und dient als Bezugspotential für bestimmte Betrachtungen.

Bei n-Typ-Halbleitern liegt das Ferminiveau stets oberhalb, bei p-Typ-Halbleitern stets unterhalb der Bandmitte bzw. des Eigenleitungs-Ferminiveaus E_i . Das Ferminiveau $E_{F,n}$ eines n-Typ-Halbleiters und $E_{F,p}$ eines p-Typ-Halbleiters berechnet sich dann zu:

$$E_{F,n} = E_C - kT \cdot \ln\left(\frac{N_C}{N_D^+}\right) \quad \text{und} \quad E_{F,p} = E_V - kT \cdot \ln\left(\frac{N_V}{N_A^-}\right). \quad (1.15)$$

Die Lage des Ferminiveaus und der Verlauf der Besetzungswahrscheinlichkeit $W(E)$ bei Fermiverteilung ist in Abb. 1.13a für einen eigenleitenden Halbleiter und in Abb. 1.13b für einen n-Typ-Halbleiter dargestellt.

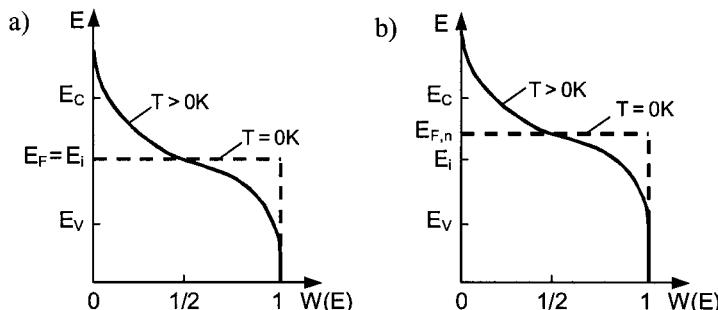


Abb. 1.13. Besetzungswahrscheinlichkeit $W(E)$ bei Fermiverteilung für die Temperaturen $T = 0\text{K}$ und $T > 0\text{K}$ und Lage des Ferminiveaus E_F eines a) eigenleitenden Halbleiters b) n-Typ-Halbleiters

Die Besetzungswahrscheinlichkeit $W(E)$ für Elektronen berechnet sich mit Hilfe der Ferminiveaus E_F und der Boltzmann-Konstante k wie folgt:

$$W(E) = \left(1 + \exp\left(\frac{E - E_F}{kT}\right) \right)^{-1}. \quad (1.16)$$

Bei $|E - E_F| \gg kT$ gilt für die Besetzungswahrscheinlichkeit $W(E)$ nach der sog. *Boltzmann-Näherung* folgende Vereinfachung:

$$W(E) \approx \exp\left(-\frac{E - E_F}{kT}\right) \text{ für } E \gg E_F. \quad (1.17)$$

Für die Besetzungswahrscheinlichkeit $1 - W(E)$ für Löcher folgt entsprechend der Boltzmann-Näherung:

$$1 - W(E) \approx \exp\left(-\frac{E_F - E}{kT}\right) \text{ für } E \ll E_F. \quad (1.18)$$

Liegt die Energie E bei einer Temperatur von $T = 0$ K unterhalb des Ferminiveaus ($E < E_F$), so ist die Besetzungswahrscheinlichkeit für Elektronen $W(E) = 1$. Liegt die Energie E bei $T = 0$ K oberhalb des Ferminiveaus ($E > E_F$), so ist die Besetzungswahrscheinlichkeit für Elektronen $W(E) = 0$. Bei Temperaturen $T > 0$ K beträgt die Besetzungswahrscheinlichkeit $W(E_F) = 1/2$ (vgl. Abb. 1.13).

1.3.3.3 Ladungsträgerverteilung

Multipliziert man die Zustandsdichte mit der Besetzungswahrscheinlichkeit, so erhält man die Ladungsträgerverteilungen $n_E(E)$ und $p_E(E)$ für Elektronen und Löcher im Halbleiter:

$$n_E(E) = N_n(E) \cdot W(E) \quad \text{und} \quad p_E(E) = N_p(E) \cdot (1 - W(E)). \quad (1.19)$$

Die Ladungsträgerdichten $n_0(E)$ und $p_0(E)$ des Leitungs- bzw. Valenzbandes bei thermodynamischem Gleichgewicht werden durch Integration der Ladungsträgerverteilungen über der Energie bestimmt:

$$n_0 = \int_{E_C}^{+\infty} n_E(E) dE \quad \text{und} \quad p_0 = \int_0^{E_V} p_E(E) dE. \quad (1.20)$$

Unter Verwendung der Boltzmann-Näherung (vgl. Abschn. 1.3.3.2) erhält man für die Ladungsträgerdichten:

$$n_0 \approx N_C \cdot \exp\left(-\frac{E_C - E_F}{kT}\right) \quad \text{und} \quad p_0 \approx N_V \cdot \exp\left(-\frac{E_F - E_V}{kT}\right). \quad (1.21)$$

In Gl. (1.21) bezeichnen N_C und N_V die effektiven Zustandsdichten für Elektronen im Leitungsband bzw. für Löcher im Valenzband (vgl. Abschn. 1.3.3.1). Die

Ladungsträgerverteilung kann also durch monoenergetische Zustände der Dichten N_C und N_V an den Bandkanten E_C bzw. E_V angenähert werden. Die Ladungsträgerverteilungen $n_E(E)$ und $p_E(E)$ eines eigenleitenden Halbleiters sind in Abb. 1.14b dargestellt. In Abb. 1.14c sind die zugehörigen Ladungsträgerdichten n_0 und p_0 veranschaulicht.

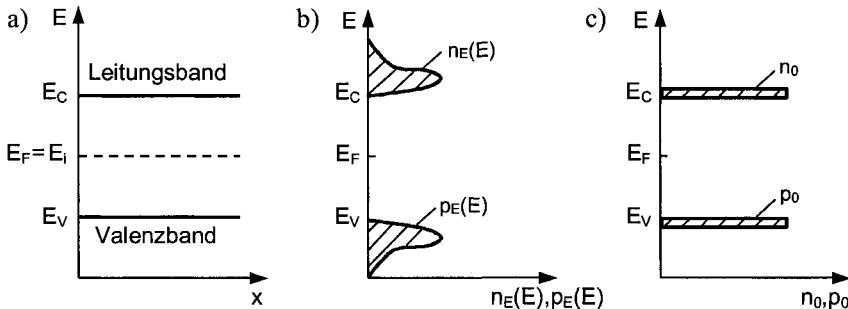


Abb. 1.14. Eigenleitender Halbleiter a) Bändermodell b) Ladungsträgerverteilungen $n_E(E)$ für Elektronen und $p_E(E)$ für Löcher c) Ladungsträgerdichten n_0 und p_0 bei Boltzmann-Näherung

Die Ladungsträgerverteilungen und Ladungsträgerdichten bei Boltzmann-Näherung bei thermodynamischem Gleichgewicht sind in Abb. 1.15 für einen n-Typ-Halbleiter dargestellt. Das Ferminiveau liegt nahe der Leitungsbandkante; im Leitungsband befinden sich mehr Elektronen als Löcher im Valenzband.

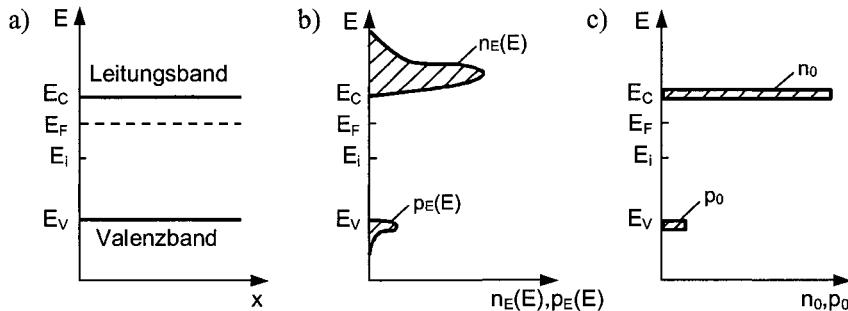


Abb. 1.15. n-Typ-Halbleiter a) Bändermodell b) Ladungsträgerverteilungen $n_E(E)$ für Elektronen und $p_E(E)$ für Löcher c) Ladungsträgerdichten n_0 und p_0 bei Boltzmann-Näherung

1.3.4 Entartete Halbleiter

Bei Störstellenhalbleitern mit sehr hoher Dotierung bilden sich zusätzliche Störbänder aus, die bei n-Typ-Halbleitern als Donatorband und bei p-Typ-Halbleitern als Akzeptorband bezeichnet werden. Diese Störbänder überlappen mit dem Leitungs- bzw. Valenzband, wodurch eine Verschiebung der Bandkante und eine

Verringerung des Bandabstandes erfolgt. Dadurch befindet sich das Ferminiveau E_F innerhalb des Leitungs- oder Valenzbandes (vgl. Abb. 1.16a) und die Leitfähigkeit des Halbleiters besitzt metallähnliches Verhalten. Die Zustandsdichte $N_n(E)$ für Elektronen im Leitungsband und $N_p(E)$ für Löcher im Valenzband eines entarteten n-Typ-Halbleiters sind in Abb. 1.16b dargestellt.

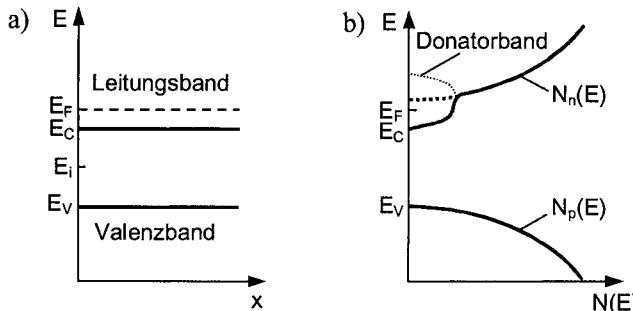


Abb. 1.16. Entarteter n-Typ-Halbleiter a) Bändermodell und b) Zustandsdichte $N_n(E)$ für Elektronen im Leitungsband und Zustandsdichte $N_p(E)$ für Löcher im Valenzband

Entartete Halbleiter finden insbesondere Anwendung als ohmsche Metall-Halbleiter-Kontakte.

1.4 Ladungsträgertransport

Der Ladungsträgertransport im Halbleiter wird einerseits durch elektrische oder magnetische Felder hervorgerufen, die sog. *Driftströme* bewirken; andererseits entstehen durch inhomogene Ladungsträgerverteilungen und thermische Energie sog. *Diffusionsströme*. Die Summe aus Drift- und Diffusionsströmen wird als *Konvektionsstromdichte* bezeichnet. Die am Ladungsträgertransport des Halbleiters beteiligten Stromkomponenten werden nachfolgend näher beschrieben.

1.4.1 Driftstromdichte

Driftströme im Halbleiter werden durch elektrische und magnetische Felder hervorgerufen. Im Folgenden wird zunächst von einem homogen dotierten Halbleiter ohne Einfluss eines magnetischen Feldes ($B = 0$) bei einer Temperatur von $T > 0\text{K}$ ausgegangen. Ein im Halbleiter auftretendes elektrisches Feld beschleunigt die beweglichen Ladungsträger so, dass sich auf Grund ihrer Polarität die Elektronen antiparallel und die Löcher parallel zur Feldrichtung bewegen. Die Energie der beschleunigten Ladungsträger wird durch Streuung an Störstellen und an der Gitterstruktur teilweise wieder abgegeben. Bei der Streuung werden die Ladungsträger von ihrer ursprünglichen Bahn abgelenkt. Als Maß für die gerichtete Bewegung der Ladungsträger dient die *Driftgeschwindigkeit*. Die von der elektrischen Feld-

stärke E abhängige Driftgeschwindigkeit v_n für Elektronen und v_p für Löcher berechnen sich wie folgt:

$$v_n = -\mu_n \cdot E \quad \text{und} \quad v_p = \mu_p \cdot E . \quad (1.22)$$

In Gl. (1.22) bezeichnet μ_n die *Beweglichkeit* der Elektronen und μ_p die Beweglichkeit der Löcher im Halbleiter. Da Elektronen gegenüber Löchern eine geringere Wechselwirkung mit der Gitterstruktur aufweisen, gilt $\mu_n > \mu_p$. Für Silizium bei Raumtemperatur beträgt $\mu_n = 1350 \text{ cm}^2 \text{V}^{-1} \text{s}^{-1}$ und $\mu_p = 480 \text{ cm}^2 \text{V}^{-1} \text{s}^{-1}$. Die Ladungsträgerbeweglichkeit in Halbleitern ist erheblich höher als in Metallen. In Abb. 1.17 sind die Driftgeschwindigkeiten v_n für Elektronen und v_p für Löcher eines Silizium-Halbleiters in Abhängigkeit von der elektrischen Feldstärke E dargestellt. Bei hoher Feldstärke tritt hier eine sog. *Geschwindigkeitssättigung* bei $v_{Sat} \approx 10^7 \text{ cm/s}$ auf. Die Feldstärkeabhängigkeit der Driftgeschwindigkeit ist durch die energieabhängigen Streumechanismen im Halbleiter begründet.

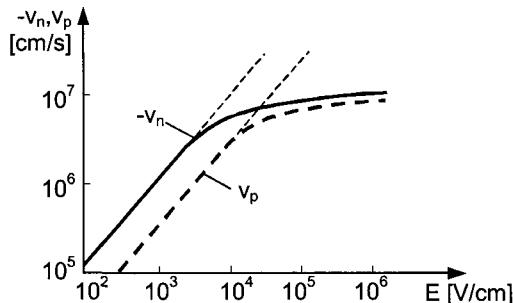


Abb. 1.17. Driftgeschwindigkeit v_n für Elektronen und v_p für Löcher eines Silizium-Halbleiters in Abhängigkeit von der elektrischen Feldstärke E ; Auftreten von Geschwindigkeitssättigung bei $v_{Sat} \approx 10^7 \text{ cm/s}$

Aus der Driftgeschwindigkeit der Ladungsträger im Halbleiter gemäß Gl. 1.22 resultiert die Driftstromdichte $i_{n,Drift}$ der Elektronen und die Driftstromdichte $i_{p,Drift}$ der Löcher:

$$i_{n,Drift} = -e \cdot n \cdot v_n = e \cdot n \cdot \mu_n \cdot E , \quad (1.23)$$

$$i_{p,Drift} = e \cdot p \cdot v_p = e \cdot p \cdot \mu_p \cdot E . \quad (1.24)$$

Die Summe der Driftstromdichten für Elektronen und Löcher ergibt die Driftstromdichte i_{Drift} des Halbleiters:

$$i_{Drift} = e \cdot (n \cdot \mu_n + p \cdot \mu_p) \cdot E . \quad (1.25)$$

Aus der Driftstromdichte kann mit Hilfe des lokalen ohmschen Gesetzes ($i = \sigma \cdot E$) die spezifische Leitfähigkeit σ bzw. der spezifische Widerstand ρ des Halbleiters ermittelt werden:

$$\boxed{\sigma = \frac{I}{\rho} = e \cdot (n \cdot \mu_n + p \cdot \mu_p).} \quad (1.26)$$

Die spezifische Leitfähigkeit beschreibt die Bewegung der Ladungsträger im Halbleiter. Dabei existiert neben der Dotierungsdichteabhängigkeit eine starke Temperaturabhängigkeit der spezifischen Leitfähigkeit σ , die in Abb. 1.18 beispielhaft für n-Typ-Silizium-Halbleiter mit unterschiedlichen Dotierungsdichten N_D dargestellt ist.

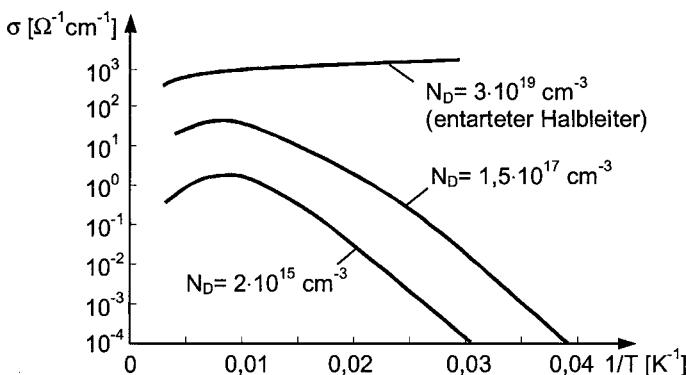


Abb. 1.18. Typische Verläufe der Temperaturabhängigkeit der spezifischen Leitfähigkeit σ für n-Typ-Silizium-Halbleiter mit unterschiedlichen Donatordotierungsdichten N_D

Bei niedrigen Temperaturen nimmt die spezifische Leitfähigkeit σ mit der Temperatur T zu, was durch die zunehmende Ionisierung der Dotierstoffatome begründet ist. Nachdem alle Dotierstoffatome in ionisierter Form vorliegen, bestimmt die durch thermische Gitterstreuung bedingte Abnahme der Ladungsträgerbeweglichkeit μ die Temperaturabhängigkeit. Dadurch nimmt die spezifische Leitfähigkeit ab. Wird die Temperatur weiter erhöht, so gelangt man in den Bereich der Eigenleitung ($n_i \geq N$), in dem die spezifische Leitfähigkeit wieder stark zunimmt. Die Temperaturabhängigkeit der spezifischen Leitfähigkeit von Halbleitern wird z. B. in Temperatursensoren zur Temperaturmessung ausgenutzt (vgl. Abschn. 6.1.2).

1.4.2 Diffusionsstromdichte

Diffusionsströme entstehen bei einem Ladungsträgerausgleich durch thermische Bewegungen in Halbleitergebieten mit örtlich unterschiedlichen (inhomogenen) Ladungsträgerdichten. Dabei entsteht ohne Wirken eines externen elektrischen Feldes ($E = 0$) ein Teilchenstrom, der vom Gebiet hoher Ladungsträgerdichte in das Gebiet niedrigerer Ladungsträgerdichte gerichtet ist. Ein eindimensionales Modell zur Visualisierung des Diffusionsstroms ist in Abb. 1.19 am Beispiel eines n-Typ-Halbleiters dargestellt.

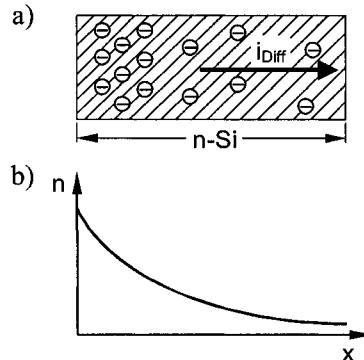


Abb. 1.19. Diffusionsmechanismus im n-Typ-Halbleiter a) Entstehen eines Diffusionsstroms i_{Diff} durch Ladungsträgerdichtegefälle b) Verlauf der Ladungsträgerdichte $n(x)$

Die entstehende Diffusionsstromdichte ist proportional zum Gradienten der Ladungsträgerdichte. Bei konstanter Temperatur folgt für die Diffusionsstromdichte $i_{n,Diff}$ der Elektronen und der Diffusionsstromdichte $i_{p,Diff}$ der Löcher folgender Zusammenhang:

$$i_{n,Diff} = e \cdot D_n \cdot \frac{\partial n}{\partial x} \quad \text{und} \quad i_{p,Diff} = -e \cdot D_p \cdot \frac{\partial p}{\partial x}. \quad (1.27)$$

Die in Gl. (1.27) auftretenden temperaturabhängigen *Diffusionskonstanten* D_n und D_p für Elektronen und Löcher berechnen sich mit Hilfe entsprechenden Ladungsträgerbeweglichkeiten μ_n und μ_p zu:

$$D_n = \frac{kT}{e} \cdot \mu_n \quad \text{und} \quad D_p = \frac{kT}{e} \cdot \mu_p. \quad (1.28)$$

Die Beziehungen in Gl. (1.28) werden *Einstein-Beziehungen* genannt. Der darin enthaltene Term aus Boltzmann-Konstante k , Temperatur T und Elementarladung e wird als *Temperaturspannung* V_ϑ bezeichnet:

$$V_\vartheta = \frac{kT}{e}. \quad (1.29)$$

Die Temperaturspannung beträgt bei Raumtemperatur $V_\vartheta \approx 26\text{mV}$.

1.4.3 Konvektionsstromdichte

Die Konvektionsstromdichte setzt sich aus den Driftstromdichten (vgl. Abschn. 1.4.1) und den Diffusionsstromdichten (vgl. 1.4.2) zusammen. Dabei treten die Drift- und Diffusionsvorgänge im Halbleiter unabhängig voneinander auf. Die Konvektionsstromdichte i_n für Elektronen und die Konvektionsstromdichte i_p für Löcher berechnen sich dann wie folgt:

$$i_n = i_{n,Drift} + i_{n,Diff} = e \cdot n \cdot \mu_n \cdot E + e \cdot D_n \cdot \frac{\partial n}{\partial x}, \quad (1.30)$$

$$i_p = i_{p,Drift} + i_{p,Diff} = e \cdot p \cdot \mu_p \cdot E - e \cdot D_p \cdot \frac{\partial p}{\partial x}. \quad (1.31)$$

Die Gesamtkonvektionsstromdichte i im Halbleiter setzt sich aus den Konvektionsstromdichten i_n und i_p wie folgt zusammen:

$$i = \sum \rho_k \cdot v_k = i_n + i_p. \quad (1.32)$$

In Gl. (1.32) kennzeichnet ρ_k die Ladungsdichte der Ladungsträger des entsprechenden Leitungstyps ($\rho_n = -e \cdot n$ oder $\rho_p = e \cdot p$). Die Größe v_k bezeichnet die zugehörigen mittleren Driftgeschwindigkeiten v_n und v_p der Elektronen bzw. Löcher (vgl. Abschn. 1.4.1).

1.5 Halbleiter im Nichtgleichgewicht

Die Störung des thermodynamischen Gleichgewichts bewirkt eine Veränderung der Ladungsträgerdichten, sodass ein Überschuss oder Mangel an beweglichen Ladungsträgern entsteht. Im Folgenden werden Störungen betrachtet, die durch Majoritätsträger- oder Minoritätsträgerinjektion oder durch gleichzeitige Injektion von Majoritäts- und Minoritätsträgern erfolgen. Tritt die Störung nicht im gesamten Volumen zugleich auf, dann wird sie von der Ladungsträgerdiffusion begleitet.

Die Störung des thermodynamischen Gleichgewichts wird durch die sog. *Überschussträgerdichten* n' und p' beschrieben:

$$n' = n - n_0 \quad \text{und} \quad p' = p - p_0. \quad (1.33)$$

Die folgenden Betrachtungen erfolgen unter der Annahme, dass ein homogen dotierter Halbleiter vorliegt, die Temperatur konstant ist und außer der jeweils betrachteten keine weitere Störung auf den Halbleiter einwirkt. Für den zeitlichen Verlauf der Ladungsträgerdichten und Überschussträgerdichten gilt:

$$\frac{\partial n}{\partial t} = \frac{\partial n'}{\partial t} \quad \text{und} \quad \frac{\partial p}{\partial t} = \frac{\partial p'}{\partial t}. \quad (1.34)$$

Die Rückkehr des sich wieder selbst überlassenen Systems in das thermodynamische Gleichgewicht wird als *Relaxation* bezeichnet.

1.5.1 Majoritätsträgerinjektion

Bei der Majoritätsträgerinjektion wird das thermodynamische Gleichgewicht dadurch gestört, dass in einen bestimmten Bereich des Halbleiters Majoritätsträger eingebracht werden. Der Injektionsvorgang und der resultierende Verlauf der Überschussträgerdichte $n'(t)$ ist in Abb. 1.20 für einen n-Typ-Halbleiter dargestellt.

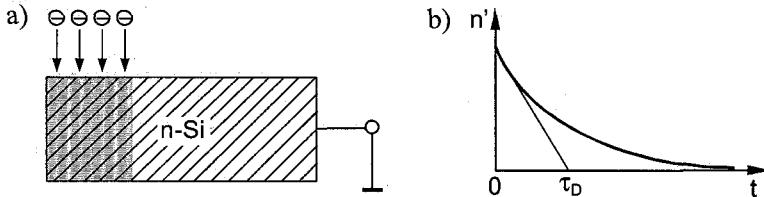


Abb. 1.20. a) Injektion von Majoritätsträgern in einen n-Typ-Halbleiter b) zeitlicher Verlauf der Überschussträgerdichte $n'(t)$ nach Beendigung der Majoritätsträgerinjektion (dielektrische Relaxation) mit Kennzeichnung der dielektrischen Relaxationszeit τ_D

Findet die Majoritätsträgerinjektion bei $t = 0$ impulsartig statt, so stellt sich der in Abb. 1.20b dargestellte zeitliche Verlauf der Überschussdichte n' ein ($p' = 0$). Durch die Störung der Neutralität ($n' \neq p'$) wird eine Majoritätsträgerverschiebung (Driftstrom) und damit ein elektrisches Feld im Halbleiter hervorgerufen. Der Driftstrom fließt so lange, bis nach Entfall der Störung alle injizierten Majoritätssträger abgeflossen sind und das thermodynamische Gleichgewicht wieder hergestellt ist (dielektrische Relaxation). Der Abklingvorgang $n'(t)$ wird durch eine Exponentialfunktion gemäß Gl. 1.35 beschrieben. Die Abklingzeitkonstante, die zum Erreichen der Neutralität ($n = p$) benötigt wird, bezeichnet man als *dielektrische Relaxationszeit* τ_D . Der Verlauf der Überschussträgerdichte $n'(t)$ und die dielektrische Relaxationszeit τ_D berechnen sich wie folgt:

$$n'(t) = n'(0) \cdot \exp\left(-\frac{t}{\tau_D}\right) \text{ und } \tau_D = \frac{\epsilon}{\sigma} \text{ mit } \epsilon = \epsilon_r \cdot \epsilon_0 . \quad (1.35)$$

Die dielektrische Relaxationszeit τ_D ist in Abb. 1.20b veranschaulicht; sie beträgt für Silizium bei Raumtemperatur $\tau_D \approx 10^{-12}$ s.

1.5.2 Minoritätsträgerinjektion

Die Störung des thermodynamischen Gleichgewichts eines Halbleiters durch Injektion von Minoritätsträgern in einen bestimmten Bereich des Halbleiters und der zeitliche Verlauf der resultierenden Überschusssichten n' und p' ist in Abb. 1.21 für einen p-Typ-Halbleiter dargestellt.

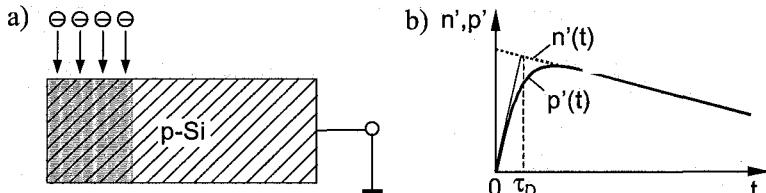


Abb. 1.21. a) Injektion von Minoritätsträgern in einen p-Typ-Halbleiter b) zeitlicher Verlauf der Überschussträgerdichten $n'(t)$ und $p'(t)$ nach Beendigung der Minoritätsträgerinjektion (dielektrische Relaxation) mit Kennzeichnung der dielektrischen Relaxationszeit τ_D

Die durch Injektion von Minoritätsträgern auftretende Überschussträgerdichte führt analog zur Majoritätsträgerinjektion (vgl. Abschn. 1.5.1) auf Grund der gestörten Neutralität ($n' \neq p'$) zu einem elektrischen Feld und einem Driftstrom im Halbleiter. Innerhalb der dielektrischen Relaxationszeit τ_D tritt im Halbleiter wieder Neutralität ein, aber es gilt $n' = p' \neq 0$. Im Gegensatz zur Majoritätsträgerinjektion werden bei der Minoritätsträgerinjektion die injizierten Ladungsträger durch zufließende Majoritätsträger kompensiert.

1.5.3 Minoritäts- und Majoritätsträgerinjektion

Die Störung des thermodynamischen Gleichgewichts durch gleichzeitige Injektion von Minoritäts- und Majoritätsträgern kann beispielsweise durch Bestrahlung eines Halbleiters mit Licht hinreichend kurzer Wellenlänge (vgl. Gl. (1.36)) erfolgen. Dabei erfolgt die Bildung von Elektron-Loch-Paaren durch optische Generation. Der Effekt der optischen Generation von Ladungsträgern wird beispielsweise in der Fotodiode (vgl. Abschn. 2.7.6) angewendet. Die optische Generation von Minoritäts- und Majoritätsträgern im homogenen Halbleiter, die durch Licht der Leistung P_L hervorgerufen wird, ist in Abb. 1.22 schematisch dargestellt.

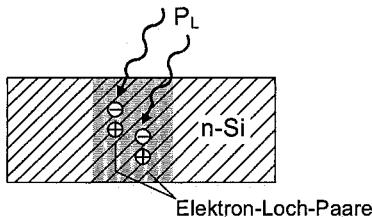


Abb. 1.22. Prinzip der Minoritäts- und Majoritätsträgerinjektion durch optische Generation von Elektron-Loch-Paaren im Halbleiter bei Bestrahlung mit der Lichtleistung P_L

Die Elektron-Loch-Paarbildung findet statt, wenn die Energie eines Photons größer ist als der Bandabstand. Die Bedingung für die Lichtwellenlänge λ wird mit Hilfe des Planckschen Wirkungsquantums h , der Lichtgeschwindigkeit c und des Bandabstandes E_G des Halbleiters wie folgt bestimmt:

$$\lambda < h \cdot \frac{c}{E_G}. \quad (1.36)$$

Durch die paarweise Generation von Elektronen und Löchern bleibt der Halbleiter neutral ($n' = p'$); der Abbau der Überschussladung erfolgt im homogenen Halbleiter nur durch erhöhte Rekombination. Während der Bestrahlung ($t < 0$) des Halbleiters stellt sich ein stationärer Zustand mit den Trägerdichten n_s und p_s ein. Im Folgenden wird der Ablauf des Relaxationsprozesses nach dem Abschalten der Bestrahlung ($t \geq 0$) betrachtet. Für die Überschussträgerdichten bei konstanter Temperatur und konstanter Dotierungsdichte gilt: $n' = n - n_0$ und $p' = p - p_0$, wobei auf Grund der Neutralität im Halbleiter $n' = p'$ gilt. Die Differenz aus der Rekom-

binationssrate R und der Generationsrate G (vgl. Abschn. 1.2.2) wird als Nettorekombinationsrate R_{netto} bezeichnet:

$$R_{netto} = R - G . \quad (1.37)$$

Die Generationsrate G wird zum einen durch die thermische Generation G_{th} von Ladungsträgern bestimmt, die von der Temperatur T und vom Bandabstand E_G des Halbleiters abhängt; zum anderen geht in die Generationsrate G die *zusätzliche Generationsrate* g ein, die aus der Bestrahlung des Halbleiters mit Licht resultiert. Für die Generationsrate ergibt sich dann:

$$G = G_{th}(E_g, T) + g \text{ mit } G_{th} = r \cdot n_0 \cdot p_0 . \quad (1.38)$$

Die Rekombinationsrate R wird über die Anzahl der Ladungsträger n und p bestimmt ($R = r \cdot n \cdot p$), wobei r den Rekombinationskoeffizienten bezeichnet (vgl. Abschn. 1.2.2). Demnach erhält man gemäß Gl. (1.37) für die Nettorekombinationsrate R_{netto} unter Verwendung der Überschussdichten n' und p' folgende Beziehung:

$$R_{netto} = r(n_0 p' + p_0 n' + n' p') - g . \quad (1.39)$$

Der zeitliche Verlauf der Trägerdichten bei Minoritäts- und Majoritätsträgerinjektion ergibt sich zu:

$$\frac{\partial n}{\partial t} = \frac{\partial n'}{\partial t} = \frac{\partial p}{\partial t} = \frac{\partial p'}{\partial t} = -R_{netto} . \quad (1.40)$$

Für den Fall der sog. *schwachen Injektion* von Minoritäts- und Majoritätsträgern ergibt sich für einen n-Typ-Halbleiter $p' \ll n_0$ und $n' \ll n_0$, woraus folgt, dass $n \approx n_0$. Für einen p-Typ-Halbleiter gilt entsprechend $n' \ll p_0$ und $p' \ll p_0$, woraus $p \approx p_0$ resultiert. Für die Nettorekombinationsrate $R_{n,netto}$ eines n-Typ-Halbleiters und die Nettorekombinationsrate $R_{p,netto}$ eines p-Typ-Halbleiters bei schwacher Injektion gilt:

$$R_{n,netto} = \frac{p'}{\tau_p} - g \quad \text{und} \quad R_{p,netto} = \frac{n'}{\tau_n} - g . \quad (1.41)$$

In Gl. (1.41) bezeichnet τ_p die Minoritätsträgerlebensdauer (der Löcher) im n-Typ-Halbleiter und τ_n die Minoritätsträgerlebensdauer (der Elektronen) im p-Typ-Halbleiter, die aus der Anzahl der Ladungsträger n_0 bzw. p_0 bei thermodynamischem Gleichgewicht bestimmt werden:

$$\tau_p = \frac{1}{r \cdot n_0} \quad \text{und} \quad \tau_n = \frac{1}{r \cdot p_0} . \quad (1.42)$$

Für einen n-Typ-Halbleiter ergibt sich der zeitliche Verlauf der Überschussminoritätsträgerdichte $\partial p' / \partial t$ wie folgt:

$$\frac{\partial p'}{\partial t} = -r \cdot n_0 p' + g . \quad (1.43)$$

Nach dem in Abb. 1.23a gezeigten Verlauf ist für die Zeit $t < 0$, in der die Bestrahlung des Halbleiters mit Licht erfolgt, die zusätzliche Generationsrate $g \neq 0$ und die Überschussträgerdichten p' und n' sind konstant. Damit ist $\partial p'/\partial t = 0$ und es folgt für die Überschussminoritätsträgerdichte p' :

$$p' = \frac{g}{r \cdot n_0} = g \cdot \tau_p = p'_s . \quad (1.44)$$

Nach Abschalten der Bestrahlung ($t \geq 0$) besitzt die zusätzliche Generationsrate g den Wert Null und es findet Relaxation (vgl. Abschn. 1.6.1) statt. Dabei beträgt das zeitliche Verhalten der Minoritätsträgerüberschussdichte $\partial p'/\partial t = -r \cdot n_0 \cdot p'$ und die Minoritätsträgerüberschussdichte $p'(t)$ klingt exponentiell mit folgender Funktion ab:

$$p'(t) = p'_s \cdot \exp\left(-\frac{t}{\tau_p}\right) . \quad (1.45)$$

Der zeitliche Verlauf der durch Bestrahlung eines n-Typ-Halbleiters hervorgerufenen zusätzlichen Generationsrate g und das resultierende zeitliche Verhalten der Überschussdichten n' und p' ist in Abb. 1.23 dargestellt.

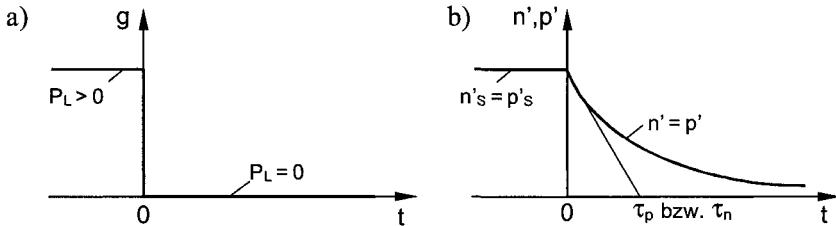


Abb. 1.23. a) Zeitlicher Verlauf der zusätzlichen Generationsrate g bei Bestrahlung eines Halbleiters mit Licht der Leistung P_L bei $t < 0$ und b) resultierendes zeitliches Verhalten der Überschussdichten n' und p' bei Relaxation ($t \geq 0$)

1.5.4 Minoritätsträgerdiffusion

Ist die Ladungsträgerdichte nicht im gesamten Halbleitervolumen gleichermaßen gestört, dann erfolgt die Diffusion von Ladungsträgern. Im Folgenden wird eine stationäre, schwache Minoritätsträgerinjektion am Rand eines homogen dotierten Halbleiters betrachtet, sodass gilt: $\partial p'/\partial t = \partial p/\partial t$. Unter der Annahme, dass die dielektrische Relaxation viel schneller erfolgt als die auf Diffusion und Rekombination beruhenden Relaxationsvorgänge, ist der Halbleiter neutral: $p'(x) = n'(x)$. Der Verlauf der Minoritätsträgerdichte und Diffusionsstromdichte bei stationärer Minoritätsträgerinjektion ist in Abb. 1.24 dargestellt.

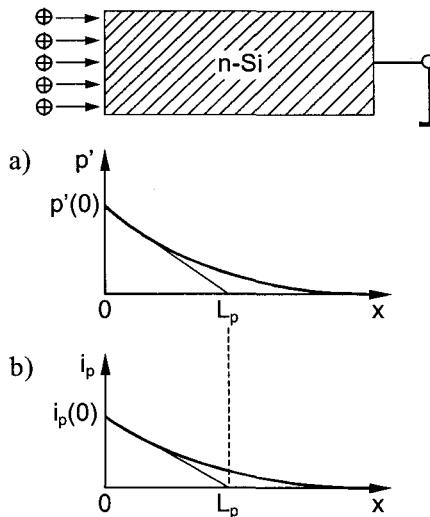


Abb. 1.24. Minoritätsträgerdiffusion durch stationäre, schwache Minoritätsträgerinjektion am Rand eines n-Typ-Halbleiters a) örtlicher Verlauf der Überschussminoritätsträgerdichte $p'(x)$ und b) örtlicher Verlauf der Diffusionsstromdichte $i_p(x)$ der Minoritätsträger

Auf Grund des Konzentrationsunterschieds der Minoritätsträgerdichte ergibt sich eine Löcherdiffusionsstromdichte $i_p(x)$ vom Rand des n-Typ-Halbleiters ins Volumen unter Verwendung der Diffusionskonstante D_p der Löcher von:

$$i_p(x) = -e \cdot D_p \cdot \frac{\partial p'}{\partial x}. \quad (1.46)$$

Dadurch ist nahe des Injektionsorts die Minoritätsträgerdichte erhöht. Dies führt gemäß Abschn. 1.5.3 zu einer Nettorekombinationsrate $R_{\text{netto}}(x) = p'(x)/\tau_p$. Im stationären Fall gilt mit der sog. *Kontinuitätsgleichung* (vgl. Abschn. 1.5.5):

$$\frac{1}{e} \cdot \frac{\partial i_p}{\partial x} + \frac{p'(x)}{\tau_p} = 0. \quad (1.47)$$

Für die Überschussminoritätsträgerdichte $p'(x)$ folgt dann ein exponentielles räumliches Abklingen gemäß:

$$p'(x) = p'(0) \cdot \exp\left(-\frac{x}{L_p}\right). \quad (1.48)$$

In Gl. (1.48) bezeichnet L_p die Diffusionslänge der Löcher, die in Abb. 1.24 veranschaulicht ist und sich mit Hilfe der Diffusionskonstanten D_p gemäß der Einstein-Beziehungen (vgl. Abschn. 1.4.2, Gl. (1.27)) und der Ladungsträgerlebensdauer τ_p berechnet. Für die Diffusionslänge L_p der Löcher und analog für die Diffusionslänge L_n der Elektronen gilt:

$$L_p = \sqrt{D_p \cdot \tau_p} \quad \text{und} \quad L_n = \sqrt{D_n \cdot \tau_n}. \quad (1.49)$$

Die Diffusionslänge gibt die innerhalb der Minoritätsträgerlebensdauer durch Diffusion im Mittel zurückgelegte Strecke der Minoritätsträger an. Sie stellt die kennzeichnende Größe für das räumliche Abklingen der Überschussminoritätssträgerdichte, die in einen homogenen Halbleiter injiziert wird, dar. Für die Diffusionsstromdichte $i_p(x)$ resultiert dann aus Gl. (1.46) die folgende Beziehung:

$$i_p(x) = i_p(0) \cdot \exp\left(-\frac{x}{L_p}\right) \quad \text{mit} \quad i_p(0) = -e \cdot D_p \cdot \left(-\frac{p'(0)}{L_p}\right). \quad (1.50)$$

Die Ladung der Überschussminoritätsträgerdichte $p'(x)$ wird durch Zufluss von Elektronen neutralisiert, wobei sich im Fall der schwachen Injektion die Majoritätsträgerdichte nicht merklich ändert. Es gilt somit im gesamten Halbleitervolumen $p'(x) = n'(x)$ und $n(x) = n_0(x) + n'(x) \approx n_0$. Die Gesamtstromdichte in x-Richtung ist konstant: $i = i_p(x) + i_n(x) = i_p(0)$.

1.5.5 Ladungsträgerbilanz

In einem betrachteten Volumenelement kann sich die Ladungsträgerdichte der Elektronen und Löcher nur ändern, wenn Ladungsträger zu- oder abfließen oder ein Ungleichgewicht zwischen Generation und Rekombination auftritt. Die zeitliche Änderung der Ladungsträgerdichte $\partial n/\partial t$ der Elektronen und $\partial p/\partial t$ der Löcher wird in den Kontinuitätsgleichungen für Elektronen und Löcher beschrieben. Bei eindimensionaler Betrachtung gilt:

$$\frac{\partial n}{\partial t} = \frac{1}{e} \cdot \frac{\partial i_n}{\partial x} - (R - G) \quad \text{und} \quad \frac{\partial p}{\partial t} = -\frac{1}{e} \cdot \frac{\partial i_p}{\partial x} - (R - G). \quad (1.51)$$

Bei schwacher Injektion erhält man unter Berücksichtigung der Nettorekombinationsrate R_{netto} gemäß Abschn. 1.5.3, Gl. (1.40) die Kontinuitätsgleichungen eines n-Typ-Halbleiters für die zeitliche Änderung der Überschussdichte $\partial n'/\partial t$ der Elektronen und die zeitliche Änderung der Überschussdichte $\partial p'/\partial t$ der Löcher:

$$\frac{\partial n'}{\partial t} = \frac{1}{e} \cdot \frac{\partial i_n}{\partial x} - \left(\frac{p'}{\tau_p} - g\right) \quad \text{und} \quad \frac{\partial p'}{\partial t} = -\frac{1}{e} \cdot \frac{\partial i_p}{\partial x} - \left(\frac{p'}{\tau_p} - g\right). \quad (1.52)$$

Analog gilt für die Kontinuitätsgleichungen eines p-Typ-Halbleiters für die zeitliche Änderung der Überschussdichte $\partial n'/\partial t$ der Elektronen und die zeitliche Änderung der Überschussdichte $\partial p'/\partial t$ der Löcher:

$$\frac{\partial n'}{\partial t} = \frac{1}{e} \cdot \frac{\partial i_n}{\partial x} - \left(\frac{n'}{\tau_n} - g\right) \quad \text{und} \quad \frac{\partial p'}{\partial t} = -\frac{1}{e} \cdot \frac{\partial i_p}{\partial x} - \left(\frac{n'}{\tau_n} - g\right). \quad (1.53)$$

Als weitere Beziehung zur eindimensionalen Beschreibung des Halbleiters im Nichtgleichgewicht dient die *Poisson-Gleichung*:

$$\Delta\phi = \frac{\partial^2\phi}{\partial x^2} = -\frac{\rho}{\epsilon} = \frac{e \cdot (p - n + N_D^+ - N_A^-)}{\epsilon_0 \cdot \epsilon_r}. \quad (1.54)$$

Die Poisson-Gleichung ermöglicht, das elektrische Potential ϕ bzw. die elektrische Feldstärke $E = -\partial\phi/\partial x$ aus den Ladungsträgerverteilungen bzw. der Ladungsdichte ρ zu berechnen. Aus der Poisson-Gleichung, den Kontinuitätsgleichungen und der Gleichung für die Drift- und Diffusionsstromdichte kann das stationäre und dynamische Verhalten von elektronischen Bauelementen bestimmt werden.

1.6 Weitere Halbleitereffekte

Im Halbleiter treten neben den beschriebenen Hauptmechanismen verschiedene zusätzliche Effekte auf, die in speziellen elektronischen Bauelementen genutzt werden. Im Folgenden werden der sog. *Lawineneffekt*, *Thermoeffekt* und *Halleffekt* näher betrachtet. Weitere Halbleitereffekte, die insbesondere an Grenzflächen zwischen unterschiedlichen Halbleitern auftreten, werden bauelementespezifisch in den nachfolgenden Kapiteln beschrieben.

1.6.1 Lawineneffekt

Bei einem hohen elektrischen Feld am Halbleiter (typischerweise $E > 3 \cdot 10^5$ V/cm) erhalten die freien Ladungsträger eine so hohe kinetische Energie, dass die beim Stoß mit dem Kristallgitter abgegebene Energie zum Aufbrechen einer Gitterbindung führt (Stoßionisation). Dadurch entsteht ein Elektron-Loch-Paar, welches durch das hohe elektrische Feld ebenfalls beschleunigt wird und durch die beim Stoß mit dem Kristallgitter freigesetzte Energie erneut Elektron-Loch-Paare erzeugt. Dadurch entsteht ein lawinenartiges Anwachsen der Ladungsträger im Halbleiter. Das Prinzip des Lawineneffekts ist in Abb. 1.25 veranschaulicht.

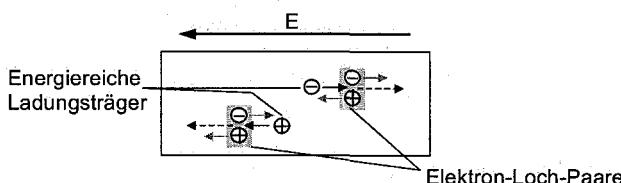


Abb. 1.25. Lawineneffekt im Halbleiter durch Stoßionisation energiereicher Ladungsträger bei hoher elektrischer Feldstärke E

Die Generationsrate G bei Auftreten des Lawineneffekts kann mit den Ladungsträgerdichten n und p sowie den Driftgeschwindigkeiten v_n und v_p der Ladungsträger (vgl. Abschn. 1.4.1) berechnet werden. Unter Berücksichtigung der Driftströme $i_{n,Drift}$ und $i_{p,Drift}$ (vgl. Abschn. 1.4.1, Gl. (1.23) und Gl. 1.24)) erhält man für die Generationsrate nach [1.2]:

$$G = \alpha_n \cdot v_n \cdot n + \alpha_p \cdot v_p \cdot p = \frac{1}{e} \cdot (\alpha_n \cdot i_{n,Drift} + \alpha_p \cdot i_{p,Drift}). \quad (1.55)$$

In Gl. (1.55) bezeichnen α_n und α_p die sog. *Ionisationskoeffizienten*. Durch das lawinenartige Anwachsen der Ladungsträger resultiert ein starker Stromanstieg im Halbleiter. Der Lawineneffekt tritt in elektronischen Bauelementen insbesondere als Durchbruchsmechanismus (vgl. Abschn. 2.2.1 und 3.3.1) auf.

1.6.2 Thermoefekt

1.6.2.1 Thermospannung im homogenen Halbleiter

Das Auftreten einer Thermospannung im homogenen Halbleiter wird *Thermodiffusion* genannt und ist in Abb. 1.26 für einen n-Typ- und p-Typ-Halbleiter bei einer Temperaturdifferenz $\Delta T = T_2 - T_1$ für $T_1 < T_2$ veranschaulicht.

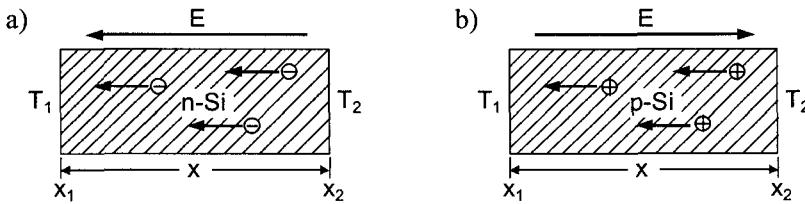


Abb. 1.26. Thermodiffusion bei einer Temperaturdifferenz $\Delta T = T_2 - T_1$ für $T_1 < T_2$ am a) n-Typ-Halbleiter und b) p-Typ-Halbleiter

Die Thermodiffusion kennzeichnet das elektrische Verhalten eines Halbleiters bei Auftreten einer Temperaturdifferenz zwischen zwei Halbleiterorten x_1 und x_2 unterschiedlicher Temperatur. Bei Erhöhung der Temperatur nimmt die mittlere Geschwindigkeit v_k der Ladungsträger zu. Dadurch erfolgt eine Diffusion der Ladungsträger vom Ort x_2 hoher Temperatur zum Ort x_1 niedriger Temperatur. Die dadurch im Halbleiter auftretenden Diffusionsströme $i_{n,Diff}$ und $i_{p,Diff}$ berechnen sich gemäß Abschn. 1.4.2, Gl. (1.27). Die resultierende elektrische Feldstärke

$$E(x) = \alpha \cdot \frac{\partial T}{\partial x} \quad (1.56)$$

führt zu einer Spannungsdifferenz ΔV_ϑ am Halbleiter, die *Thermospannung* genannt wird:

$$\Delta V_\vartheta = \alpha \cdot \Delta T. \quad (1.57)$$

Die material- und temperaturabhängige Größe α bezeichnet den differentiellen Thermokoeffizienten, der typische Werte in der Größenordnung von $100 \mu\text{V/K}$ besitzt. Der differentielle Thermokoeffizient eines n-Typ-Halbleiters ist negativ ($\alpha < 0$), der eines p-Typ-Halbleiters ist positiv ($\alpha > 0$).

1.6.2.2 Thermospannung zwischen verschiedenen Materialien

Eine Temperaturdifferenz $\Delta T = T_2 - T_1$ zwischen den Übergängen von zwei verschiedenen Materialien (z. B. zwischen Halbleiter und Metall oder zwischen zwei unterschiedlichen Halbleitern) ruft ebenfalls eine Thermospannung hervor. Das Auftreten einer Thermospannung zwischen einem n-Typ-Halbleiter und einer Metallkontakteierung ist in Abb. 1.27 dargestellt.

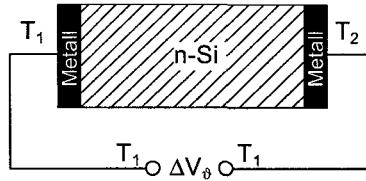


Abb. 1.27. Auftreten einer Thermospannung zwischen zwei verschiedenen Materialien (n-Typ-Halbleiter und Metallisierung) bei einer Temperaturdifferenz zwischen den beiden Materialübergängen in Höhe von $\Delta T = T_2 - T_1$

Die zwischen unterschiedlichen Materialien bei einer Temperaturdifferenz ΔT auftretende Thermospannung ΔV_θ berechnet sich zu:

$$\Delta V_\theta = \alpha_{AB} \cdot \Delta T \text{ mit } \alpha_{AB} = \alpha_A + \alpha_B. \quad (1.58)$$

Dabei bezeichnet α_{AB} den sog. *Seebeck-Koeffizienten*, der sich summarisch aus den differentiellen Thermokoeffizienten α_A und α_B (vgl. Abschn. 1.6.2.1) der unterschiedlichen Materialien (hier n-Typ-Halbleiter und Metall) zusammensetzt. Bei Halbleitern treten besonders große Seebeck-Koeffizienten auf.

Der Thermoeffekt zwischen verschiedenen Materialien wird beispielsweise in Temperatursensoren (vgl. Abschn. 6.1.3) genutzt.

1.6.3 Halleffekt

Der Halleffekt beschreibt den Einfluss eines magnetischen Feldes auf die Ladungsträgerbewegung im Halbleiter. Ein transversales Magnetfeld mit der magnetischen Induktion B ruft eine auf bewegte Ladungsträger wirkende Lorentzkraft F_L hervor, die sich wie folgt berechnet:

$$F_L = \pm e \cdot (v_k \times B) \text{ mit } v_k = \pm \mu \cdot E. \quad (1.59)$$

Dabei bezeichnet v_k die Driftgeschwindigkeit v_n oder v_p der Elektronen bzw. Löcher (vgl. Abschn. 1.4.1). Im thermodynamischen Gleichgewicht ist die Lorentzkraft ebenso wie die Driftgeschwindigkeit im Mittelwert gleich Null.

Wird der Halbleiter von einem Strom I_0 durchflossen und wirkt senkrecht dazu ein Magnetfeld, so werden die Ladungsträger durch die Lorentzkraft senkrecht zum Magnetfeld und zur Stromrichtung ausgelenkt. Dadurch entsteht eine Aufla-

dung der Halbleiterprobe gemäß Abb. 1.28. Dies bewirkt die Entstehung eines elektrischen Feldes mit der Feldstärke E , das auf die Ladungsträger eine Kraft entgegen der Lorentzkraft ausübt. Es stellt sich im unbelastetem Fall ein stationärer Zustand ein, in dem kein Strom in senkrechter Richtung fließt. Dieser Leerlaufzustand ist in Abb. 1.28 für einen n-Typ- und p-Typ-Halbleiter dargestellt.

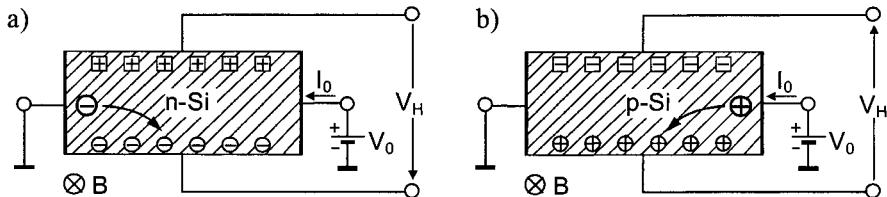


Abb. 1.28. Halleffekt bei Leerlauf a) eines n-Typ-Halbleiters ($V_H > 0$) und b) eines p-Typ-Halbleiters ($V_H < 0$)

Abhängig von der Stärke des magnetischen Feldes wird am Halbleiter eine sog. *Hallspannung* V_H erzeugt, die senkrecht zur extern angelegten Spannung V_0 wirkt. Die Hallspannung kann mit Hilfe der *Leerlaufempfindlichkeit* k_{B0} , die von der Geometrie des Halbleiters abhängt, wie folgt bestimmt werden:

$$V_H = k_{B0} \cdot I_0 \cdot B . \quad (1.60)$$

Bei einem n-Typ-Halbleiter ist die Hallspannung in der Anordnung gemäß Abb. 1.28 positiv ($V_H > 0$) und bei einem p-Typ-Halbleiter negativ ($V_H < 0$).

Liegt am n-Typ-Halbleiter kein magnetisches Feld an ($B = 0$), so berechnet sich die elektrische Feldstärke aus dem Verhältnis von Driftgeschwindigkeit v_n und Beweglichkeit μ_n der Elektronen zu $E = -v_n / \mu_n$. Wird der Halbleiter einem magnetischen Feld ausgesetzt ($B \neq 0$), so resultiert daraus folgende Beziehung zwischen magnetischem und elektrischem Feld:

$$E + r \cdot (v_n \times B) = -\frac{v_n}{\mu_n} . \quad (1.61)$$

Die Größe r stellt einen Korrekturfaktor dar, der von der Temperatur und der Dotierungsdichte des Halbleiters abhängig ist und Werte im Bereich $1 \geq r \geq 2$ annnehmen kann.

Die im Halbleiter auftretende Driftstromdichte $i_{n,Drift}$ der Elektronen und die resultierende spezifische Leitfähigkeit σ des Halbleiters können gemäß Abschn. 1.4.1, Gl. (1.23) und Gl. (1.26) bestimmt werden. Die elektrische Feldstärke E setzt sich dann mit Hilfe dieser Größen aus einem parallel wirkenden Anteil $E_{||}$ und einem senkrecht wirkenden Anteil E_{\perp} zusammen:

$$E = \underbrace{i_{n,Drift} / \sigma_n}_{E_{||}} - \underbrace{R_{Hn} \cdot i_{n,Drift} \times B}_{E_{\perp}} \quad \text{mit} \quad R_{Hn} = -\frac{r}{n \cdot e} . \quad (1.62)$$

Analog gilt für einen p-Typ-Halbleiter unter Berücksichtigung der Driftstromdichte $i_{p,Drift}$ der Löcher gemäß Gl. (1.24) folgender Zusammenhang:

$$E = \underbrace{i_{p,Drift} / \sigma_p}_{E_{\parallel}} - \underbrace{R_{Hp} \cdot i_{p,Drift} \times B}_{E_{\perp}} \quad \text{mit} \quad R_{Hp} = \frac{r}{p \cdot e}. \quad (1.63)$$

In Gl. (1.62) und Gl. (1.63) bezeichnen R_{Hn} und R_{Hp} die sog. *Hallkonstante* für einen n-Typ- bzw. p-Typ-Halbleiter. Das Produkt aus der Hallkonstanten und der spezifischen Leitfähigkeit wird als *Hallbeweglichkeit* μ_H bezeichnet:

$$\mu_H = |R_H| \cdot \sigma = r \cdot \mu \quad (1.64)$$

Zwischen dem Feldstärke- und Driftstromdichtevektor tritt der sog. *Hallwinkel* δ auf, der sich über die *Hallbeweglichkeit* μ_H wie folgt berechnet:

$$\tan |\delta| = \frac{|E_{\perp}|}{|E_{\parallel}|} = r \cdot \mu \cdot |B| = \mu_H \cdot |B|. \quad (1.65)$$

Zur Erzielung eines möglichst großen Hallwinkels δ und damit einer hohen Hallspannung V_H werden Halbleitermaterialien mit hoher Hallbeweglichkeit μ_H (z. B. InSb, InAs, GaAs) genutzt.

Der Halleffekt findet beispielsweise in Magnetosensoren (vgl. Abschn. 6.2) Anwendung. Weiterhin ermöglicht der Halleffekt auf Grund der für n-Typ- und p-Typ-Halbleiter unterschiedlichen Polarität der Hallspannung eine experimentelle Bestimmung des Leitungstyps von Halbleitern.

2 Halbleiterdioden

Wird ein p-dotierter mit einem n-dotierten Halbleiter metallurgisch in Kontakt gebracht, so entsteht ein sog. *pn-Übergang*. Das auf dem Mechanismus eines pn-Überganges basierende elektronische Bauelement nennt man *Halbleiterdiode*. Ausgehend von den Eigenschaften des pn-Überganges werden in diesem Kapitel der Aufbau und die Funktionsweise von Halbleiterdioden betrachtet. Verschiedene Typen von Dioden und deren Einsatzkriterien werden beschrieben und an Hand ausgewählter Schaltungen im praktischen Anwendungsfall dargestellt.

2.1 pn-Übergang

2.1.1 Abrupter pn-Übergang im thermodynamischen Gleichgewicht

Der sog. *abrupte pn-Übergang* beschreibt den metallurgischen Übergang von zwei entgegengesetzt dotierten Halbleitern mit den Dotierungsdichten N_A und N_D , wobei die Dotierung jeweils im gesamten Volumen konstant ist. In Abb. 2.1 ist der schematische Aufbau eines abrupten pn-Überganges mit einem beispielhaften Dotierungsprofil von $N_D = 2 \cdot N_A$ dargestellt.

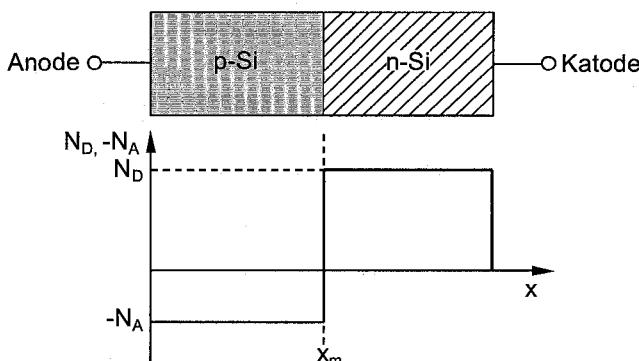


Abb. 2.1. Schematischer Aufbau eines abrupten pn-Überganges mit beispielhaftem Dotierungsprofil $N_D = 2 \cdot N_A$

Der Anschluss des p-Si-Gebietes wird als *Anode* bezeichnet, der Anschluss des n-Si-Gebietes wird *Katode* genannt. Die folgenden Betrachtungen erfolgen für das

thermodynamische Gleichgewicht, d. h. es liegt keine externe Spannung zwischen Anode und Katode an.

Auf Grund des Konzentrationsgefälles der freien Ladungsträger am pn-Übergang treten Diffusionsströme auf, die zu einer Verarmung der jeweiligen Ladungsträgerdichten in der Nähe des metallurgischen Überganges x_m führen. Eine sog. Raumladungszone (RLZ) entsteht (vgl. Abschn. 2.1.2.1). Daraus resultiert der in Abb. 2.2a dargestellte Verlauf der Elektronen- und Löcherdichte am pn-Übergang. In der Raumladungszone wird die Raumladung der festen Ladungsträger N_A bzw. N_D^+ nicht von der Raumladung der freien Ladungsträger kompensiert, so dass der in Abb. 2.2b dargestellte ortsabhängige Verlauf der Raumladungsdichte ρ entsteht.

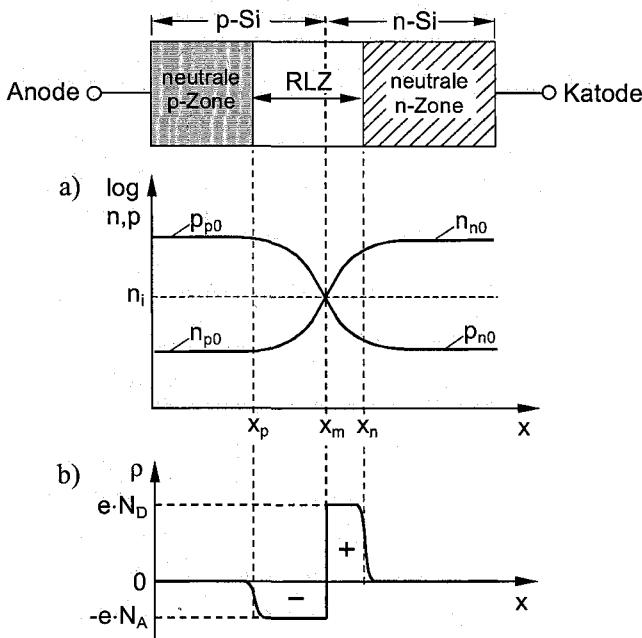


Abb. 2.2. Verläufe am abrupten pn-Übergang bei thermodynamischem Gleichgewicht
a) Ladungsträgerdichten der freien Ladungsträger b) Raumladungsdichte

Die außerhalb ($x < x_p$ bzw. $x > x_n$) der Raumladungszone befindlichen neutralen Halbleiterzonen besitzen näherungsweise konstante Ladungsträgerdichten. Für die Ladungsträgerdichten n_{p0} und p_{p0} der neutralen p-Zone gilt im thermodynamischen Gleichgewicht:

$$n_{p0} = \frac{n_i^2}{N_A} \quad \text{und} \quad p_{p0} = N_A. \quad (2.1)$$

Die konstanten Ladungsträgerdichten n_{n0} und p_{n0} der neutralen n-Zone sind entsprechend:

$$n_{n0} = N_D \text{ und } p_{n0} = \frac{n_i^2}{N_D}. \quad (2.2)$$

Das auf Grund der Raumladungszone entstehende, ortsabhängige elektrische Feld ist im Feldverlauf in Abb. 2.3a dargestellt. Durch dieses elektrische Feld werden Driftströme der freien Ladungsträger verursacht, die den Diffusionsströmen entgegenwirken. Es stellt sich daher ein Gleichgewichtszustand ein. Die maximale Feldstärke E_M tritt dabei am metallurgischen Übergang x_m zwischen n- und p-dotiertem Halbleiter auf und berechnet sich gemäß Abschn. 2.1.2.1, Gl. (2.7).

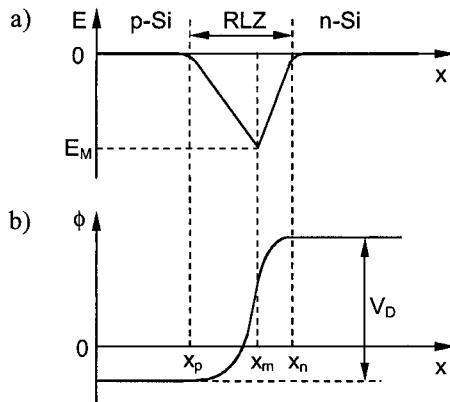


Abb. 2.3. a) Verlauf des elektrischen Feldes und b) Potentialverlauf am pn-Übergang bei thermodynamischem Gleichgewicht

Wird die elektrische Feldstärke E über x integriert, so erhält man den in Abb. 2.3b dargestellten Verlauf des elektrischen Potentials $\phi(x)$. Dieser Verlauf macht das Auftreten eines Potentialunterschieds über der Raumladungszone sichtbar. Dieser Potentialunterschied wird *Diffusionsspannung* V_D genannt. Die Diffusionsspannung V_D kann durch Integration des elektrischen Feldes $E(x)$ in den Grenzen x_p und x_n der Raumladungszone bestimmt werden und ist von der Akzeptor- und Donatordotierungsdichte N_A bzw. N_D sowie von der Temperaturspannung V_ϑ abhängig. Die Diffusionsspannung berechnet sich wie folgt:

$$V_D = - \int_{x_p}^{x_n} E(x) dx = V_\vartheta \cdot \ln \left(\frac{N_A \cdot N_D}{n_i^2} \right) \quad \text{mit } V_\vartheta = \frac{kT}{e}. \quad (2.3)$$

Die Diffusionsspannung stellt die Kontaktspannung zwischen den unterschiedlich dotierten Halbleiterzonen dar. Auch beim Übergang zwischen verschiedenen stark dotierten Halbleitern gleichen Leitungstypen treten Diffusionsspannungen auf.

Zwischen Anoden- bzw. Katodenkontakt und der jeweiligen Halbleiterzone entstehen ebenfalls Kontaktspannungen, deren Differenz gleich der Diffusionsspannung ist. Mit der Anzahl n der Kontaktspannungen resultiert die Summe der Kontaktspannungen in einem geschlossenen Kreis zu:

$$\sum_{j=1}^n V_{Dj} = 0. \quad (2.4)$$

Aus Gl. (2.4) resultiert, dass die Diffusionsspannung des pn-Überganges im thermodynamischen Gleichgewicht nicht an den Kontakten abgreifbar ist.

Der in Abb. 2.3b dargestellte Potentialverlauf $\phi(x)$ bestimmt auch die Energieverläufe des Bändermodells für den pn-Übergang (Abb. 2.4). Das Maß für die auftretende Bandverbiegung ist die Diffusionsspannung V_D gemäß Gl. (2.3). Das Ferminiveau $E_F(x)$ der im Kontakt befindlichen Halbleiter besitzt im thermodynamischen Gleichgewicht an jedem Ort den gleichen Wert. Die Diffusionsspannung kann daher auch aus der Differenz der Ferminiveaus der nicht im Kontakt stehenden n- und p-Halbleiterbereiche bezüglich der Bandkanten bestimmt werden.

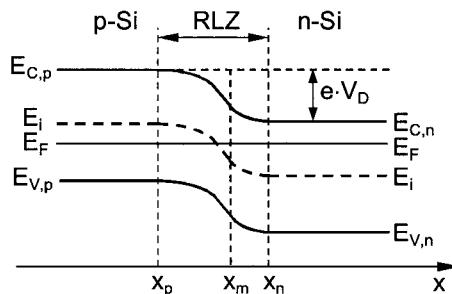


Abb. 2.4. Bändermodell eines abrupten pn-Überganges bei thermodynamischem Gleichgewicht

2.1.2 Abrupter pn-Übergang unter Vorspannung

Liegt eine Spannung $V = 0$ am pn-Übergang an (thermodynamisches Gleichgewicht), so fällt über der Raumladungszone die Diffusionsspannung V_D ab. Der Driftstrom ist dann gleich dem Diffusionsstrom und es fließt netto kein Elektronen- und Löcherstrom. Wird der pn-Übergang nun mit Hilfe einer externen Spannung $V \neq 0$ gemäß Abb. 2.5 vorgespannt, so tritt über der Raumladungszone ein Spannungsabfall von $(V_D - V)$ auf und die Länge ℓ der Raumladungszone nimmt zu für $V < 0$ ($\Delta\ell > 0$) bzw. nimmt ab für $V > 0$ ($\Delta\ell < 0$).

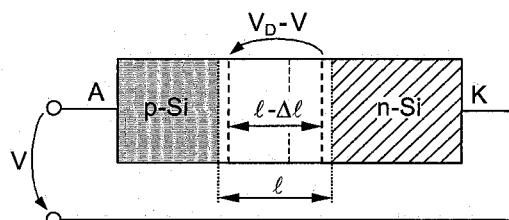


Abb. 2.5. Abrupter pn-Übergang unter Vorspannung (hier beispielhaft für $V > 0$)

Im Folgenden werden ausgehend von der Berechnung der Raumladungszone des pn-Überganges die Betriebsfälle für externe Spannungen von $V > 0$ und $V < 0$ näher beschrieben.

2.1.2.1 Raumladungszone des pn-Überganges

Nachfolgend soll die Ausdehnung der Raumladungszone in Abhängigkeit von der extern angelegten Spannung V berechnet werden. Dazu wird von ohmschen Kontakten und vernachlässigbaren Spannungsabfällen in den neutralen Zonen des pn-Überganges ausgegangen. Da die Raumladung der freien Ladungsträger gegenüber der Raumladung der ionisierten Dotieratome vernachlässigt werden kann (*Depletion-Näherung*), resultiert eine Gesamtraumladungsdichte in der Raumladungszone von:

$$\rho_{RLZ} = e \cdot (N_D - N_A). \quad (2.5)$$

Der Potentialverlauf $\phi(x)$ am abrupten pn-Übergang ist bei dieser Raumladungsverteilung parabelförmig, der Feldverlauf $E(x)$ ist linear (vgl. Abb. 2.6). Die auf diesen Annahmen basierende Berechnung der Raumladungszone nennt man daher auch *Schottkysche Parabelnäherung*.

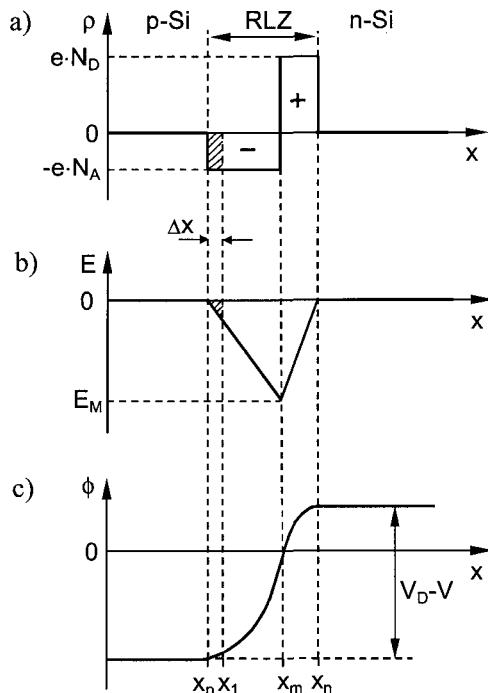


Abb. 2.6. Verläufe am pn-Übergang nach der Schottkyschen Parabelnäherung a) Raumladungsdichte b) elektrisches Feld c) elektrisches Potential

In den neutralen Zonen gilt für die Raumladung $\rho = 0$ und auch das elektrische Feld wird zu Null. Innerhalb der Raumladungszone im p-Gebiet gilt für die elektrische Feldstärke E im Abstand Δx vom Rand:

$$E(x) = -\frac{e \cdot N_A}{\epsilon} \cdot \Delta x. \quad (2.6)$$

Die maximale Feldstärke E_M tritt am metallurgischen Übergang zwischen n- und p-Gebiet auf:

$$E_M = -\frac{e \cdot N_A}{\epsilon} \cdot \ell_p. \quad (2.7)$$

In Gl. (2.7) gibt $\ell_p = x_m - x_p$ die Ausdehnung der Raumladungszone im p-Gebiet bzw. die Grenze zur neutralen p-Zone an. Da die Gesamtladung der Raumladungszone Null ist, gilt die Beziehung:

$$N_A \cdot \ell_p = N_D \cdot \ell_n. \quad (2.8)$$

Aus dieser Verhältnisgleichung resultiert für unterschiedlich stark dotierte Zonen, dass sich die Raumladungszone weiter in die niedriger dotierte Zone ausdehnt. Für die Ausdehnung der Raumladungszone in das p- und n-Gebiet folgt dann aus Gl. (2.7) und Gl. (2.8):

$$\ell_p = -\frac{\epsilon}{e \cdot N_A} \cdot E_M \quad \text{und} \quad \ell_n = -\frac{\epsilon}{e \cdot N_D} \cdot E_M. \quad (2.9)$$

Die Potentialdifferenz am pn-Übergang ($V_D - V$) entspricht der Fläche unter dem Feldstärkeverlauf $-E(x)$ gemäß Abb. 2.6b und es gilt:

$$V_D - V = -\frac{1}{2} \cdot E_M \cdot (\ell_p + \ell_n). \quad (2.10)$$

Die maximale Feldstärke E_M in Abhängigkeit von der externen Spannung V am pn-Übergang berechnet sich aus Gln. (2.9) und Gl. (2.10):

$$E_M = -\sqrt{\frac{2 \cdot e}{\epsilon} \cdot (V_D - V) \cdot \frac{1}{\frac{1}{N_A} + \frac{1}{N_D}}}. \quad (2.11)$$

Die spannungsabhängige Ausdehnung der Raumladungszone kann dann für $V < V_D$ wie folgt bestimmt werden:

$$\ell = \sqrt{\frac{2 \cdot \epsilon}{e} \cdot (V_D - V) \cdot \left(\frac{1}{N_A} + \frac{1}{N_D} \right)}, \quad (2.12)$$

$$\text{mit } \ell_p = \ell \cdot \frac{N_D}{N_D + N_A}, \quad \ell_n = \ell \cdot \frac{N_A}{N_D + N_A}. \quad (2.13)$$

Steigt die externe Spannung V am pn-Übergang an, so verringert sich die Ausdehnung ℓ der Raumladungszone und die Feldstärke E_M nimmt ab. Im Folgenden werden die Fälle einer positiven und negativen externen Spannung näher beschrieben.

2.1.2.2 pn-Übergang in Flusspolung

Wird eine externe Spannung $V > 0$ an den pn-Übergang angelegt, so bewirkt die nun verringerte Spannung ($V_D - V$) über der Raumladungszone einen Anstieg des Diffusionsstroms, durch den Elektronen in die p-Zone und Löcher in die n-Zone injiziert werden. Da eine hohe Anzahl von Ladungsträgern zur Injektion zur Verfügung steht, kommt es zu einem mit der Spannung schnell ansteigenden Stromfluss. In den an die Raumladungszone angrenzenden Diffusionsgebieten ist die Minoritätsträgerdichte erhöht. Dieser Fall wird *Flusspolung* genannt. In Abb. 2.7 sind die Minoritätsträgerdichteverteilung und die resultierenden Stromkomponenten bei Flusspolung dargestellt.

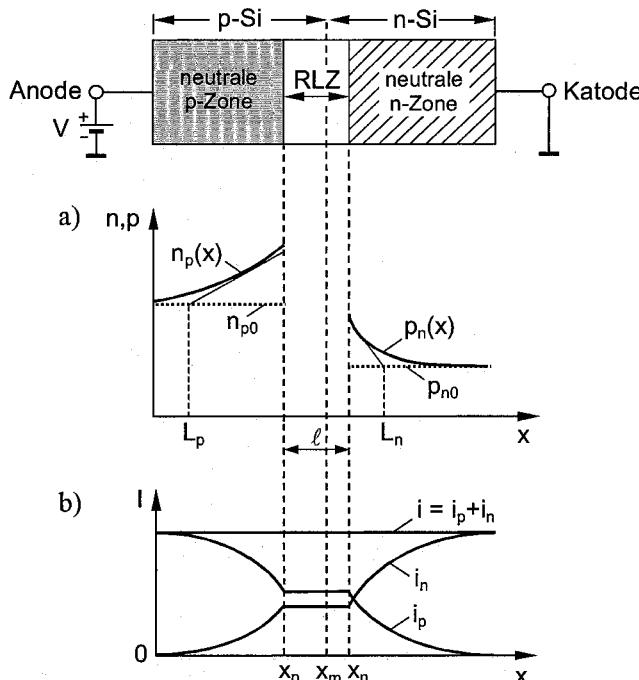


Abb. 2.7. Abrupter pn-Übergang in Flusspolung ($V > 0$)
a) Ladungsträgerdichteverteilung
b) Stromkomponenten i_n , i_p und Gesamtkonvektionsstromdichte i (Diffusionsnäherung)

Bei Spannungen $V \neq 0$ wird die Minoritätsträgerdichte am Rand der Raumladungszone im Wesentlichen durch die Majoritätsträgerdiffusion in der Raumladungszone bestimmt. Im Folgenden wird angenommen, dass sich am Rand der

neutralen Zone die Minoritätsträger in einem Quasigleichgewicht mit den Majoritätsträgern im Bereich jenseits der Raumladungszone befinden. Die Löcher am Rand des neutralen n-Gebietes sind dann im Quasigleichgewicht mit den Löchern im p-Gebiet:

$$\left. \frac{n_p}{n_n} \right|_{Rand} = \exp\left(-\frac{(V_D - V)}{V_\vartheta}\right). \quad (2.14)$$

Analog gilt, dass sich die Elektronen am Rand des neutralen p-Gebietes im Quasigleichgewicht mit den Elektronen im n-Gebiet befinden:

$$\left. \frac{p_n}{p_p} \right|_{Rand} = \exp\left(-\frac{(V_D - V)}{V_\vartheta}\right). \quad (2.15)$$

Im Abstand von mehr als einer Diffusionslänge (vgl. Abschn. 1.4.3) in die neutrale Zone hinein sind Minoritätsträger und Majoritätsträger örtlich im Quasigleichgewicht ($n \cdot p = n_i^2$). Für geringe Flusspolung ($V_D - V \gg V_\vartheta$) gilt die Näherung der schwachen Injektion ($n_n = n_{n0}$ und $p = p_{p0}$). Für die Elektronendichte n_p am Rand der Raumladungszone bei schwacher Injektion folgt:

$$\boxed{n_p|_{Rand} = n_{p0} \cdot \exp\left(\frac{V}{V_\vartheta}\right)}. \quad (2.16)$$

Für die Löcherdichte p_n am Rand der Raumladungszone bei schwacher Injektion gilt analog:

$$\boxed{p_n|_{Rand} = p_{n0} \cdot \exp\left(\frac{V}{V_\vartheta}\right)}. \quad (2.17)$$

Die Ermittlung des Stroms, der durch den pn-Übergang fließt, erfolgt unter der Annahme, dass keine Rekombinations- und Generationsvorgänge in der Raumladungszone auftreten (nach Diffusionsmodell). Wie in Abb. 2.7b dargestellt, besitzen die Stromkomponenten i_n und i_p dann innerhalb der Raumladungszone einen konstanten Wert. Zur Bestimmung der Gesamtkonvektionsstromdichte i kann der in die p-Zone injizierte Minoritätsträgerstrom i_n und der in die n-Zone injizierte Majoritätsträgerstrom i_p am Rand der Raumladungszone wie folgt berechnet werden:

$$i_n = e \cdot D_n \cdot \left. \frac{dn'}{dx} \right|_{x=x_p} \quad \text{und} \quad i_p = -e \cdot D_p \cdot \left. \frac{dp'}{dx} \right|_{x=x_n}. \quad (2.18)$$

Außerhalb der Raumladungszone klingen die Stromdichten exponentiell mit der jeweiligen Diffusionslänge L_n und L_p ab (vgl. Abb. 2.7a). In Gl. (2.18) bezeichnen $n'(x)$ die Überschussträgerdichte der Elektronen und $p'(x)$ die Überschussträgerdichte der Löcher, die sich gemäß Gl. (2.16) und Gl. (2.17) berechnen:

$$n_p'(x) = n_p(x) - n_{p0} = n_{p0} \cdot \left(\exp\left(\frac{V}{V_\vartheta}\right) - 1 \right), \quad (2.19)$$

$$p_n'(x) = p_n(x) - p_{n0} = p_{n0} \cdot \left(\exp\left(\frac{V}{V_\vartheta}\right) - 1 \right). \quad (2.20)$$

Die Summe der in Gl. (2.18) bestimmten Diffusionsstromdichten ergibt die Gesamtkonvektionsstromdichte i des pn-Überganges:

$$i = i_n + i_p = -e \cdot \left(D_n \cdot \left. \frac{dn'}{dx} \right|_{x=x_p} - D_p \cdot \left. \frac{dp'}{dx} \right|_{x=x_n} \right). \quad (2.21)$$

Damit erhält man dann für den Gesamtstrom I_D des idealen pn-Überganges bzw. der idealen pn-Diode die Beziehung:

$$\boxed{I_D = I_{SD} \cdot \left(\exp\left(\frac{V}{V_\vartheta}\right) - 1 \right)}. \quad (2.22)$$

Dabei bezeichnet I_{SD} den Sperrsättigungsstrom nach dem Diffusionsmodell. Mit der Diodenfläche A_{pn} berechnet sich der Sperrsättigungsstrom wie folgt:

$$I_{SD} = A_{pn} \cdot e \cdot \left(p_{n0} \cdot \frac{L_p}{\tau_p} + n_{p0} \cdot \frac{L_n}{\tau_n} \right). \quad (2.23)$$

Die aus Gl. (2.22) und Gl. (2.23) resultierende Strom-Spannungs-Kennlinie wird in Abschn. 2.2.1 näher beschrieben.

Berücksichtigt man zusätzlich die Rekombinations- und Generationsvorgänge in der Raumladungszone, so erhält man den Strom I der realen pn-Diode. Dieser Strom setzt sich aus dem Stromanteil I_D nach dem Diffusionsmodell gemäß Gl. (2.22) und einem zusätzlichen Stromanteil I_R nach dem Rekombinations-/Generationsmodell zusammen:

$$\boxed{I = I_D + I_R = I_{SD} \cdot \left(\exp\left(\frac{V}{V_\vartheta}\right) - 1 \right) + I_{SR} \cdot \left(\exp\left(\frac{V}{2V_\vartheta}\right) - 1 \right)}. \quad (2.24)$$

Dabei ergibt sich der Sperrsättigungsstrom I_{SR} als Folge der Generation in der Raumladungszone nach dem Rekombinations-/Generationsmodell zu:

$$I_{SR} = A_{pn} \cdot e \cdot \ell(V) \cdot \frac{n_i}{\tau_e} \quad \text{mit} \quad \left. \tau_e \right|_{E_T=E_i} = \tau_p + \tau_n. \quad (2.25)$$

Demnach hängt der Sperrsättigungsstrom I_{SR} von der Diodenfläche A_{pn} , der spannungsabhängigen Ausdehnung der Raumladungszone $\ell(V)$ und der Generationsrate n_i/τ_e ab. Die Größe τ_e stellt die charakteristische Zeitkonstante für den Generationsmechanismus dar und berechnet sich für ein Zwischenenergieniveau

E_T , das die gleiche Energie wie das Eigenleitungsniveau E_i besitzt, aus der Summe der Ladungsträgerlebensdauern τ_p und τ_n . Die resultierenden Strom-Spannungs-Verläufe werden in Abschn. 2.2 näher betrachtet.

2.1.2.3 pn-Übergang in Sperrpolung

Legt man eine externe Spannung $V < 0$ an den pn-Übergang an, so wird die Injektion von Elektronen aus der n-Zone bzw. von Löchern aus der p-Zone erschwert. Der Diffusionsstrom wird daher schon bei geringen Sperrspannungen unterdrückt; es verbleibt der Driftstrom. Dabei werden die Minoritätsträger der jeweiligen Zone durch das elektrische Feld über den metallurgischen pn-Übergang gezogen, was zum Fließen eines geringen *Sperrstroms* führt. Dieser Betriebsfall wird *Sperrpolung* genannt.

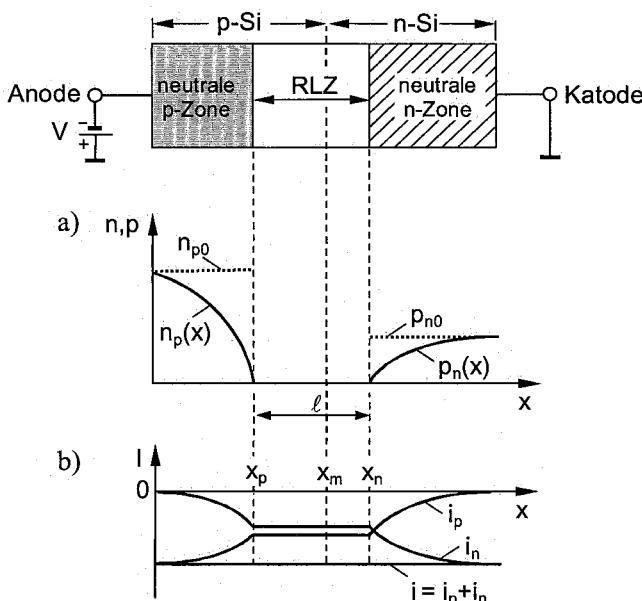


Abb. 2.8. Abrupter pn-Übergang in Sperrpolung ($V < 0$) a) Minoritätsträgerdichteverteilung in den Diffusionsgebieten b) Stromkomponenten i_n , i_p und Gesamtstrom i

Für die Ladungsträgerdichten n_p und p_n an den Rändern der Raumladungszone gelten die in Abschn. 2.1.2.2 beschriebenen Abhängigkeiten:

$$n_p|_{\text{Rand}} = n_{p0} \cdot \exp\left(\frac{V}{V_\vartheta}\right) \quad \text{und} \quad p_n|_{\text{Rand}} = p_{n0} \cdot \exp\left(\frac{V}{V_\vartheta}\right). \quad (2.26)$$

Daraus folgt, dass die Minoritätsträgerdichte am Rand der Raumladungszone für $V >> V_\vartheta$ auf nahezu Null absinkt (vgl. Abb. 2.8a). Der am pn-Übergang bei Sperrpolung auftretende Sperrsättigungsstrom I_{SD} nach dem Diffusionsmodell

wurde bereits in Abschn. 2.1.2.2, Gl. (2.23) berechnet. Der zusätzlich auftretende Sperrsättigungsstrom I_{SR} nach dem Rekombinations-/Generationsmodell kann gemäß Abschn. 2.1.2.2, Gl. (2.25) bestimmt werden. Eine Veranschaulichung der Sperrstromverläufe nach dem Diffusionsmodell und dem Rekombinations-/Generationsmodell erfolgt in Abschn. 2.2.

2.2 Diodenkennlinien

Die Kennlinien von pn-Dioden resultieren aus den am pn-Übergang bestimmten Stromverhältnissen für Fluss- und Sperrpolung gemäß Abschn. 2.1.2.2 und Abschn. 2.1.2.3. Nachfolgend werden die Kennlinien der idealen und realen pn-Diode näher betrachtet.

2.2.1 Kennlinie der idealen Diode

Die Darstellung des Strom-Spannungs-Verlaufes der idealen pn-Diode erfolgt gemäß der Berechnungen für den Diodenstrom und den Sperrsättigungsstrom nach dem Diffusionsmodell, d. h. Rekombinations- und Generationsvorgänge in der Raumladungszone werden nicht berücksichtigt. Dadurch wird der Diodenstrom durch die Minoritätsträgerdiffusion in den neutralen Zonen begrenzt. Die Kennlinie einer idealen Diode resultiert aus Gl. (2.22) und ist schematisch in Abb. 2.9a dargestellt.

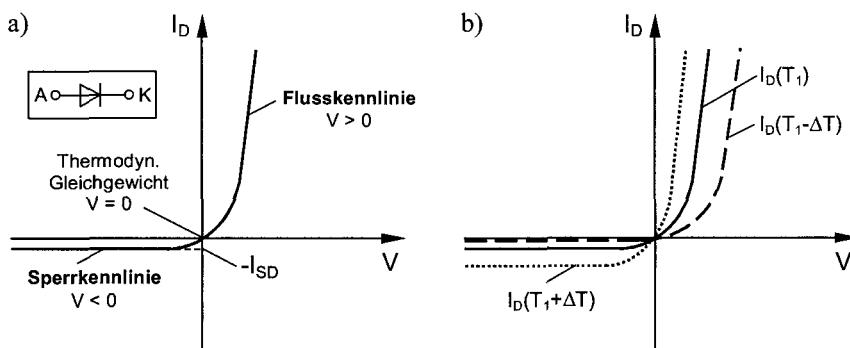


Abb. 2.9. a) Schematische Strom-Spannungs-Kennlinie einer idealen pn-Diode nach dem Diffusionsmodell und Schaltsymbol der pn-Diode b) Temperaturabhängigkeit der idealen Diodenkennlinie

Bei Flusspolung ($V > 0$) steigt der Diodenstrom schon bei geringen Flussspannungen stark an, was gemäß Gl. (2.16) auf den exponentiellen Zusammenhang zwischen der externen Spannung V und der Minoritätsträgerdichte zurückzuführen ist. Bei Sperrpolung hingegen setzt bereits bei geringer Sperrspannung Sättigung ein, sodass $I_D \approx -I_{SD}$ gilt.

Der Sperrsättigungsstrom I_{SD} weist auf Grund der Proportionalität der Minoritätsträgerdichten n_{p0} und p_{n0} zur Eigenleitungsträgerdichte $n_i^2(T)$ eine starke Temperaturabhängigkeit auf (exponentielle Abhängigkeit von V_ϑ). Der Diodenstrom I_D besitzt nach Gl. (2.22) ebenfalls eine starke Temperaturabhängigkeit. Das resultierende Temperaturverhalten der idealen Diodenkennlinie ist beispielhaft in Abb. 2.9b für die Temperaturen $T_1 + \Delta T$ und $T_1 - \Delta T$ veranschaulicht.

2.2.2 Kennlinie der realen Diode

Die Strom-Spannungs-Kennlinie der realen Diode berücksichtigt insbesondere den Stromanteil I_R nach dem Rekombinations-/Generationsmodell nach Gl. (2.24). Die Kennlinie einer realen Diode ist in Abb. 2.10 schematisch dargestellt.

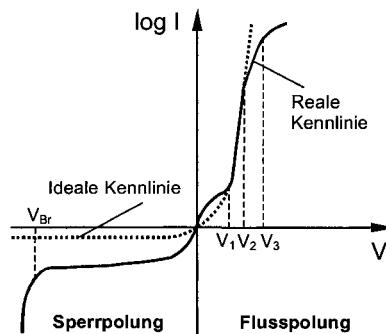


Abb. 2.10. Schematische Strom-Spannungs-Kennlinie einer realen pn-Diode

In Flusspolung ist bei geringen Spannungen $0 < V \leq V_1$ der Einfluss der Rekombination von Ladungsträgern in der Raumladungszone nach dem Rekombinations-/Generationsmodell sichtbar. Bei Spannungen $V_1 < V \leq V_2$, wobei V_2 unterhalb der Diffusionsspannung V_D liegt, verhält sich die reale pn-Diode analog dem Diffusionsmodell (vgl. Abschn. 2.2.1). Bei höheren Spannungen $V_2 < V \leq V_3$ wird der Einfluss starker Injektion ($I_R \sim \exp(V/2V_\vartheta)$) sichtbar. Bei hohen Spannungen $V > V_3$ überwiegt dann jeweils der Einfluss des endlichen Widerstandes der neutralen Zonen, was zu einem Abknicken der Kennlinie führt.

In Sperrpolung ist der Sperrsättigungsstrom I_{SD} nach dem Diffusionsmodell bei Silizium-pn-Dioden so gering, dass der Sperrsättigungsstrom I_{SR} nach dem Rekombinations-/Generationsmodell dominiert (vgl. Abschn. 2.1.2.2). Weiterhin treten ab einer bestimmten Sperrspannung $|V| > V_{Br}$ sog. *Durchbruchsmechanismen* auf, die ein starkes Ansteigen des Sperrstroms bewirken (vgl. Abschn. 2.3).

2.2.3 Diodenkennlinie mit 2-Geraden-Näherung

Insbesondere für Gleichrichteranwendungen (vgl. Abschn. 2.8.1) kann die Diodenkennlinie durch eine 2-Geraden-Näherung vereinfacht werden. Dabei wird die

Kennlinie in Flusspolung oberhalb der sog. *Schwellspannung* durch eine Gerade angenähert. Die Schwellspannung V_S stellt einen wichtigen Parameter zur Spezifikation einer pn-Diode dar. Sie gibt die Spannung an, bei der ein merklicher Stromfluss durch die Diode erfolgt. Für $V \geq V_S$ berechnet sich der Diodenstrom I_{2G} bei Berücksichtigung des sog. *Bahnwiderstandes* R_B der neutralen Zonen, der in Serie zur idealen pn-Diode wirkt, über die Beziehung:

$$I_{2G} = \frac{V - V_S}{R_B}. \quad (2.27)$$

In Abb. 2.11 ist der schematische Verlauf einer Diodenkennlinie nach der 2-Geraden-Näherung dargestellt.

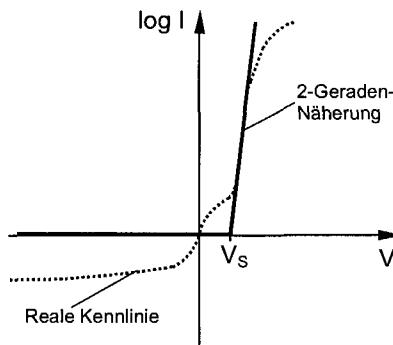


Abb. 2.11. Diodenkennlinie mit 2-Geraden-Näherung

Unterhalb der Schwellspannung ($V < V_S$) gilt nach der 2-Geraden-Näherung, dass kein Sperrstrom fließt ($I_{SD} = I_{SR} = 0$). Diese Näherung gilt für $|V| < V_{Br}$.

2.3 Durchbruchmechanismen

Bei Sperrpolung treten verschiedene Durchbruchmechanismen am pn-Übergang auf, die einen starken Stromanstieg unterhalb einer bestimmten Sperrspannung bewirken. Durchbrucheffekte beeinflussen somit die Sperrfestigkeit der pn-Diode. Im Folgenden werden die wichtigsten Durchbruchmechanismen, der *Lawineneffekt*, der *Zenereffekt* und der *thermische Durchbruch* näher beschrieben.

2.2.1 Lawineneffekt

Einzelne Ladungsträger nehmen im starken elektrischen Feld der Raumladungszone hinreichend Energie auf, damit bei einem Stoß mit dem Kristallgitter (vgl. Kap. 1) ein Elektron-Loch-Paar gebildet werden kann (Stoßionisation). Die so entstehenden freien Ladungsträger können nach Beschleunigung im elektrischen Feld

weitere Paarbildungen verursachen, wodurch oberhalb einer kritischen maximalen Feldstärke E_M ein lawinenartiges Anwachsen von freien Ladungsträgern entsteht. Der Durchbruchseffekt ist in Abb. 2.12 veranschaulicht.

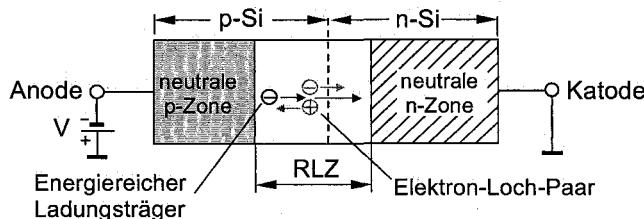


Abb. 2.12. pn-Übergang bei Lawinendurchbruch, Erzeugung von Elektron-Loch-Paaren durch energiereiche Ladungsträger

Bei Temperaturerhöhung verringert sich die mittlere freie Weglänge der Teilchen. Das zum Auslösen des Lawinendurchbruchs notwendige elektrische Feld muss dementsprechend größer sein. Daraus resultiert ein *positiver Temperaturkoeffizient* der Durchbruchsspannung als Folge des Lawineneffekts.

Der Lawinendurchbruch ist bei Strombegrenzung reversibel. In Silizium-pn-Dioden ist er der am häufigsten auftretende Durchbruchsmechanismus.

Bei zunehmender Dotierung (und gleicher Spannung am pn-Übergang) steigt die maximale Feldstärke E_M in der Raumladungszone an. Die Durchbruchsspannung V_{Br} infolge des Lawinendurchbruchs nimmt somit bei steigender Dotierung ab. In Abb. 2.13 ist die Durchbruchsspannung in Abhängigkeit von der Dotierungsdichte für verschiedene Materialien dargestellt.

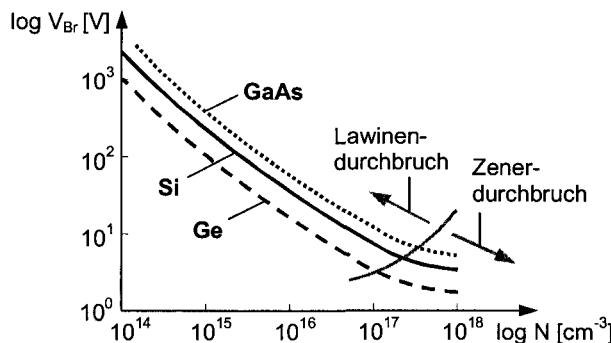


Abb. 2.13. Schematischer Verlauf der Durchbruchsspannung V_{Br} einer pn-Diode als Funktion der Dotierungsdichte N für die Halbleitermaterialien Si, Ge und GaAs bei $T = 300K$ mit Einsatzbereich für Lawinen- und Zenerdurchbruch

In Abb. 2.13 wird sichtbar, dass bei Dotierungsdichten unterhalb $N \approx 10^{18} cm^{-3}$ und Durchbruchsspannungen von $V_{Br} > 10V$ der Lawinendurchbruch dominiert. Bei höheren Dotierungsdichten und geringeren Durchbruchsspannungen wirkt dann

ein weiterer Durchbruchsmechanismus, der Zenerdurchbruch genannt wird und in Abschn. 2.2.2 näher beschrieben wird.

2.2.2 Zenereffekt

Bei hoher Dotierungsdichte (vgl. Abb. 2.13) treten bereits bei geringen Sperrspannungen hohe Feldstärken am pn-Übergang auf. Dementsprechend existieren Elektronenzustände gleicher Energie im Valenzband des p-Gebietes und im Leitungsband des n-Gebietes, die sich in geringem örtlichen Abstand befinden. Dadurch können Elektronen des Valenzbandes aus der p-Zone durch die Bandlücke in unbesetzte Zustände des Leitungsbands der n-Zone tunneln. Dieser Tunnelmechanismus ist in Abb. 2.14 veranschaulicht.

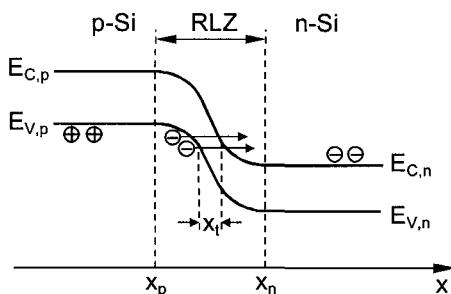


Abb. 2.14. Bänderschema des pn-Überganges bei Erzeugung freier Ladungsträger durch den Zenereffekt

Die aus dem Valenzband des pn-Überganges ins Leitungsband getunnelten Elektronen und die im Valenzband verbleibenden Löcher stehen jeweils als zusätzliche Ladungsträger zur Verfügung, die den Sperrstrom der pn-Diode erhöhen. Die Tunnelwahrscheinlichkeit nimmt exponentiell mit abnehmender Breite der Tunnelbarriere x_t zu. Der Durchbruch tritt ein, wenn der Tunnelstrom der pn-Diode größer als der thermisch bedingte Sperrstrom ist. Der Zenerdurchbruch ist bei Strombegrenzung reversibel.

Mit steigender Temperatur verringert sich die Energie E_G der Bandlücke, und das zum Einsetzen des Zenerdurchbruchs erforderliche elektrische Feld ist geringer, was zu einem *negativen Temperaturkoeffizienten* der Durchbruchsspannung als Folge des Zenereffekts führt. Dioden, die im Übergang von Lawinen- zu Zenerdurchbruch betrieben werden, besitzen einen besonders geringen Temperaturkoeffizienten der Durchbruchsspannung.

2.2.3 Thermischer Durchbruch

Der Sperrstrom I_S führt zu einer Verlustleistung $P_V = I_S \cdot V$, die in der pn-Diode umgesetzt wird. Die damit verbundene Eigenerwärmung $T(P_V)$ kann auf Grund

der starken Temperaturabhängigkeit des Sperrstroms zum irreversiblen thermischen Durchbruch führen. Das Einsetzen des thermischen Durchbruchs erfolgt unter der Bedingung:

$$R_{th} \cdot V \cdot \frac{dI_R}{dT} \geq I, \quad (2.28)$$

wobei R_{th} den thermischen Widerstand zwischen dem pn-Übergang und einer Wärmesenke darstellt. Der thermische Durchbruch tritt insbesondere bei pn-Dioden mit geringem Bandabstand und hohem thermischen Widerstand auf.

2.4 Technologische Realisierung

Die technologische Realisierung von pn-Dioden in sog. *Silizium-Planartechnik* erfolgt an Hand der in Abb. 2.15 beispielhaft dargestellten Technologieschritte.

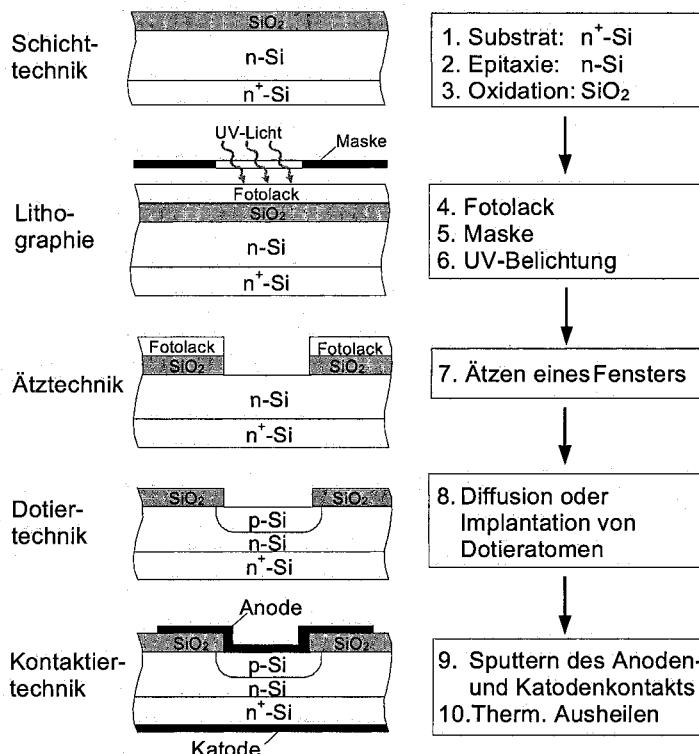


Abb. 2.15. Typische Technologieschritte zur Herstellung einer Silizium-pn-Diode

Ausgangspunkt der Herstellung einer pn-Diode ist beispielweise ein n⁺-Silizium-Substrat. Auf diesem Substrat wird eine n-Si-Schicht aufgewachsen und an-

schließend durch thermische Oxidation bei z. B. $T = 1100^\circ\text{C}$ eine Isolatorschicht aus SiO_2 erzeugt. Im nächsten Prozessschritt (Lithographie) wird ein sog. *Fotolack* aufgetragen und durch eine Maske mit UV-Licht bestrahlt. Durch das bestrahlte Gebiet wird nun ein Fenster in die Isolatorschicht geätzt und der restliche Fotolack entfernt. Durch Diffusion oder Ionenimplantation erfolgt dann die Dotierung des p-Gebietes der pn-Diode. Im weiteren wird die Metallisierung des Rück- und Frontseitenkontakts durch Aufdampfen oder sog. *Sputtern* (Katodenerstörung) aufgebracht.

2.5 Netzwerkmodelle der pn-Diode

Für den schaltungstechnischen Einsatz von pn-Dioden sind die dynamischen Eigenschaften von besonderem Interesse. Diese Eigenschaften können in Netzwerkmodellen bzw. sog. *Ersatzschaltungen* beschrieben werden und dienen insbesondere zur Schaltungssimulation. Ausgehend von der Betrachtung der wichtigsten Elemente dieser Netzwerkmodelle wird nachfolgend das Großsignal- und Kleinsignalersatzschaltbild erörtert.

2.5.1 Sperrsicht- und Diffusionskapazität

Eine Spannungsänderung ΔV am pn-Übergang bewirkt eine Ladungsänderung ΔQ in der Raumladungszone, woraus eine Längenänderung $\Delta \ell$ der Raumladungszone ℓ resultiert (vgl. Abschn., 2.1.2, Abb. 2.5). Das Verhältnis aus der Ladungsänderung ΔQ zur Spannungsänderung ΔV wird Sperrsichtkapazität C_s genannt. Sie wird durch Änderung der Majoritätsträgerkonzentration verursacht; das dynamische Verhalten ist durch dielektrische Relaxation bestimmt. Die Sperrsichtkapazität eines pn-Überganges mit der Fläche A wird wie folgt berechnet:

$$C_s(V) = \frac{\Delta Q}{\Delta V} = \frac{\epsilon_0 \epsilon_r \cdot A}{\ell(V)}. \quad (2.29)$$

Die Sperrsichtkapazität ist nichtlinear von der Spannung V am pn-Übergang abhängig. Wird die pn-Diode in Sperrpolung betrieben, so wird die Gesamtkapazität maßgeblich durch die Sperrsichtkapazität bestimmt (vgl. Abb. 2.17).

In Flusspolung erfolgt zusätzlich eine Minoritätsträgerspeicherung in der sog. *Diffusionskapazität*. Die Minoritätsträgerspeicherung erfolgt durch die spannungsabhängige Erhöhung der Minoritätsträgerdichten in den neutralen Zonen nahe der Raumladungszone (vgl. Abb. 2.16) über den Gleichgewichtswert. In Abb. 2.16 veranschaulichen die schraffierten Bereiche die Erhöhung der Diffusionsladung auf Grund einer Spannungsänderung ΔV . Bei Änderung der Minoritätsträgerdichten ΔQ_n und ΔQ_p tritt eine entgegengesetzte, gleich große Änderung der Majoritätsträgerdichten auf. Dadurch herrscht in den Diffusionsgebieten Ladungsneutralität vor.

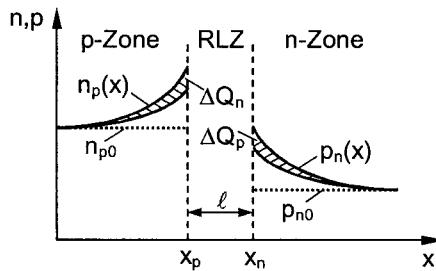


Abb. 2.16. Entstehen der Diffusionskapazität: Ladungsspeicherung in den neutralen Zonen einer pn-Diode

Aus den Ladungsänderungen ΔQ_n und ΔQ_p resultiert die Diffusionskapazität C_{Diff} , die sich wie folgt berechnet:

$$C_{Diff} = \frac{1}{2} \cdot \frac{\partial I}{\partial V} \cdot \frac{p_{n0} L_p + n_{p0} L_n}{p_{n0} L_p / \tau_p + n_{p0} L_n / \tau_n}. \quad (2.30)$$

In Gl. (2.30) bezeichnen p_{n0} und n_{p0} die Minoritätsträgerdichten im thermodynamischen Gleichgewicht. Die Größe L_p ist die Diffusionslänge der Löcher in der n-Zone; L_n beschreibt Diffusionslänge der Elektronen in der p-Zone. Die Zeiten τ_n und τ_p bezeichnen die Minoritätsträgerlebensdauer von Elektronen in der p-Zone bzw. von Löchern in der n-Zone. Für geringe Frequenzen ($\omega \rightarrow 0$) gilt für die Diffusionskapazität C_{Diff} einer p^+n -Diode, bei der die Diffusionsladung im n-Gebiet überwiegt, die folgende Beziehung:

$$C_{Diff} = \frac{I_D \cdot \tau_p}{2 \cdot V_\vartheta}. \quad (2.31)$$

Die Gesamtkapazität der pn-Diode stellt eine Parallelschaltung der Sperrschichtkapazität C_S und der Diffusionskapazität C_{Diff} dar. Der Verlauf der spannungsabhängigen Gesamtkapazität einer pn-Diode ist in Abb. 2.17 dargestellt. Wird die pn-Diode mit hinreichend hoher Flussspannung betrieben, so wird die Gesamtkapazität maßgeblich durch die Diffusionskapazität bestimmt.

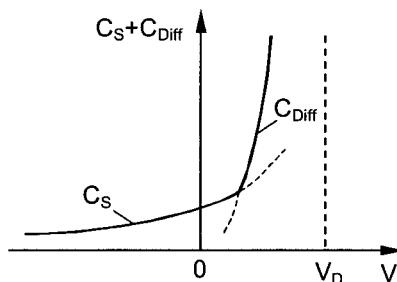


Abb. 2.17. Gesamtkapazität einer pn-Diode mit den Kapazitätsanteilen C_S und C_{Diff}

2.5.2 Großsignalersatzschaltbild

Das Großsignalverhalten einer pn-Diode wird im sog. *Großsignalersatzschaltbild* modelliert. Es dient zur Schaltungssimulation und Parametrisierung von pn-Dioden. Es enthält die in Abschn. 2.5.1 hergeleiteten nichtlinearen Sperrsicht- und Diffusionskapazitäten C_S und C_{Diff} , den Großsignalbahnwiderstand R_B und spannungsgesteuerte Stromquellen. Diese Stromquellen beschreiben das nichtlineare Verhalten der Dioden-Kennlinie in Fluss- und Sperrpolung. Das Großsignalersatzschaltbild einer pn-Diode ist in Abb. 2.18 dargestellt.

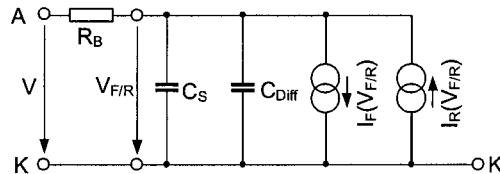


Abb. 2.18. Großsignalersatzschaltbild einer pn-Diode

Die externe Spannung V an der pn-Diode teilt sich in einen Spannungsabfall über dem Bahnwiderstand R_B und einen Spannungsabfall $V_{F/R}$ über der Sperrsicht auf. Diese Spannung steuert die Stromquelle $I_F(V_{F/R})$ für Flusspolung und die Stromquelle $I_R(V_{F/R})$ für Sperrpolung der pn-Diode.

2.5.3 Kleinsignalersatzschaltbild

Das sog. *Kleinsignalverhalten* der pn-Diode beschreibt den Zusammenhang zwischen kleinen Spannungsänderungen ΔV an der pn-Diode bei gegebenem Arbeitspunkt (AP) und den daraus resultierenden Stromänderungen ΔI . Dieser Zusammenhang wird im sog. *Kleinsignalersatzschaltbild* (Abb. 2.19a) dargestellt und mit entsprechenden Kleinsignalparametern beschrieben. Das Kleinsignalersatzschaltbild enthält die in Abschn. 2.5.1 hergeleiteten nichtlinearen Sperrsicht- und Diffusionskapazitäten C_S und C_{Diff} , den Kleinsignalbahnwiderstand r_B und den Kleinsignalleitwert g_0 .

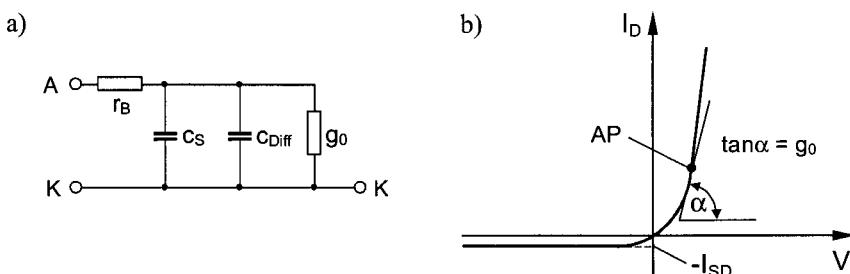


Abb. 2.19. a) Kleinsignalersatzschaltbild einer pn-Diode b) Diodenkennlinie nach dem Diffusionsmodell mit Arbeitspunkt (AP) und Kleinsignalleitwert g_0

Zur Bestimmung des Kleinsignalverhaltens wird an die pn-Diode eine Gleichspannung V angelegt und mit einer Kleinsignalspannungsänderung ΔV der Kreisfrequenz ω überlagert. Man erhält das Kleinsignalmodell aus dem Großsignalmodell durch Linearisierung der Diodenkennlinie im Arbeitspunkt (AP) [2.2]. Die Steigung der Diodenkennlinie in Abb. 2.19b ergibt dann den Kleinsignalleitwert g_0 , der sich nach dem Diffusionsmodell (vgl. Abschn. 2.2.1) für geringe Frequenzen ($\omega \rightarrow 0$) wie folgt berechnet:

$$g_0 = \frac{\partial I}{\partial V} \approx \left. \frac{I_D}{V_\vartheta} \right|_{I_D \gg I_{SD}} . \quad (2.32)$$

Der Kleinsignalbahnwiderstand r_B entspricht näherungsweise dem Großsignalbahnwiderstand R_B .

2.6 Schaltverhalten der pn-Diode

Die dynamischen Eigenschaften der pn-Diode können durch das Schaltverhalten charakterisiert werden. Bei den nachfolgenden Betrachtungen wird in das Schalten ohmscher und induktiver Lasten unterschieden.

2.6.1 Schalten ohmscher Lasten

Zur Bestimmung des Schaltverhaltens einer pn-Diode D bei ohmscher Last R_L wird die in Abb. 2.20 dargestellte Schaltungsanordnung verwendet.

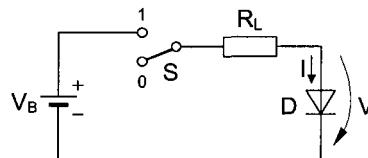


Abb. 2.20. Anordnung zur Bestimmung des Schaltverhaltens einer pn-Diode bei ohmscher Last

Vor dem Einschalten ($t \leq t_1$, $S = 0$) gilt, dass die Diodenspannung V und der Diodenstrom I jeweils Null sind und die Minoritätsträgerverteilung $p_n(x)$ außerhalb der Raumladungszone den Wert p_{n0} (Gleichgewichtsträgerdichte) besitzt. Die folgenden Betrachtungen gelten für eine p⁺n-Diode unter der Annahme, dass die Betriebsspannung viel größer als die Flussspannung ist ($V_B \gg V_F$). Nach dem Einschalten ($t > t_1$, $S = 1$) ergibt sich die in Abb. 2.21a dargestellte Minoritätsträgerverteilung $p_n(x)$ und Minoritätsträgerüberschussdichte $p'_n(x)$ in der n-Zone der pn-Diode und der in Abb. 2.21b abgebildete zeitliche Verlauf des Diodenstroms I und der Diodenspannung V . Dabei bezeichnet $x = 0$ den Rand der Raumladungszone im n-Gebiet.

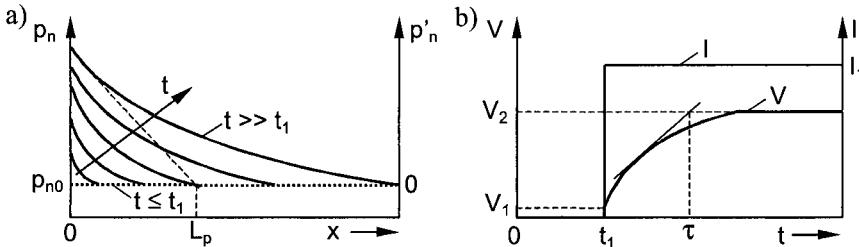


Abb. 2.21. a) Minoritätsträgerverteilung $p_n(x)$ und Minoritätsträgerüberschussdichte $p'_n(x)$ in der n-Zone und b) Verlauf von Diodenstrom und -spannung während des Einschaltvorganges einer pn-Diode

Nach dem Einschalten ($t > t_1$) gilt $I \approx I_1 = V_B/R_L$. Für den Diodenstrom in der n-Zone der pn-Diode gilt die Beziehung:

$$I_1 \sim \left. \frac{\partial p_n}{\partial x} \right|_{x=0}. \quad (2.33)$$

Für die Minoritätsträgerüberschussdichte gilt am Anfang des Einschaltvorganges (Zone mit $p'_n > 0$ ist klein im Vergleich zur Diffusionslänge L_p):

$$\int_0^{\infty} p'_n(x) dx = \frac{1}{e} \cdot \int_{t_1}^{\tau} I(t) dt. \quad (2.34)$$

Der in Gl. (2.34) beschriebene Zuwachs der Minoritätsträgerladung wird während des Einschaltvorganges durch die mit Zunahme der Minoritätsträgerdichte im Diffusionsgebiet stärker werdende Rekombination abgeschwächt und schließlich kompensiert. Es stellt sich ein stationärer Zustand ein. Die Spannung an der Diode bei $t = t_1$ berechnet sich aus dem Produkt des Diodenstroms I_1 und des Bahnwiderstandes R_B zu $V_1(t_1) = I_1 \cdot R_B$. Bei $t > t_1$ steigt die Spannung V an der Diode mit der Zeit bis auf den Wert V_2 an. Die Diodenspannung V ist dabei entsprechend Gl. (2.17) über folgende Beziehung mit der Minoritätsträgerladungsdichte verknüpft:

$$V = \frac{kT}{e} \cdot \ln \left(\frac{p_n(x=0)}{p_{n0}} \right). \quad (2.35)$$

Beim Ausschaltvorgang (Schalter S : 1 → 0) und beim Umschalten in den Sperrbereich werden analog zum Einschaltvorgang die Verläufe von Strom und Spannung an der Diode wesentlich durch die im n-Diffusionsgebiet gespeicherten Überschussladungsträger p' beeinflusst. Wird vom eingeschalteten Zustand in den Sperrbetrieb umgeschaltet, so fließt zunächst ein relativ hoher negativer Strom $I_{RM} = V_R / R_B$. Dadurch werden die Überschussladungsträger abgebaut und der Diodenstrom I beginnt zu sinken bis die pn-Diode schließlich sperrt. Das Ausschaltverhalten einer pn-Diode ist in Abb. 2.22 dargestellt.

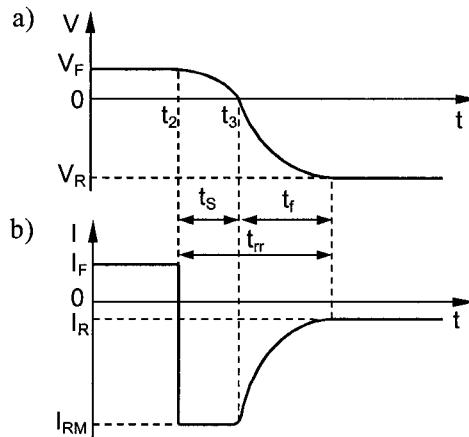


Abb. 2.22. Ausschaltverhalten einer pn-Diode bei Wechsel von Fluss- in Sperrpolung mit Kennzeichnung der Speicherzeit t_s , Abfallzeit t_f und Rückwärtserholzeit t_{rr} a) zeitlicher Verlauf der Diodenspannung und b) zeitlicher Verlauf des Diodenstroms

Direkt nach dem Umschalten ($t_2 < t \leq t_3$) von Fluss- in Sperrpolung sind die Bahngebiete der pn-Diode noch mit Ladungsträgern überschwemmt. Während dieser Zeit fließt ein näherungsweise konstanter negativer Sperrstrom I_{RM} ; die pn-Diode verbleibt jedoch durch die in den Bahngebieten gespeicherte Minoritätsträgerladung noch im Durchlassbetrieb. Dieses von der Diffusionskapazität (vgl. Abschn. 2.5.1) bestimmte Verhalten wird Sperrverzögerung genannt und ist durch die Speicherzeit t_s charakterisiert.

Bei $t > t_3$ gelangt die Diode in den Sperrbetrieb. Die Diodenspannung erreicht dann die negative Sperrspannung V_R , wobei der Sperrstrom auf den Wert $|I_R|$ absinkt. Die Zeit, in der sich die Diodenspannung und der Diodenstrom z. B. auf 90% von V_R bzw. $|I_R|$ annähern, wird als Abfallzeit t_f bezeichnet. Dieses zeitliche Verhalten wird durch das Umladen der Sperrsichtkapazität C_S der pn-Diode (vgl. Abschn. 2.5.1) über dem Lastwiderstand R_L bestimmt. Die Summe aus Speicher- und Abfallzeit wird Rückwärtserholzeit t_{rr} genannt.

2.6.2 Schalten induktiver Lasten

Zur Bestimmung des Schaltverhaltens einer pn-Diode bei induktiver Last (externe Induktivität oder induktives Verhalten der Diode bei hoher Strombelastung) wird die in Sperrpolung betriebene Diode zunächst mit einer sehr hohen Stromsteilheit dI/dt bzw. sog. *Kommutierungssteilheit* beaufschlagt. Diese Kommutierungssteilheit bewirkt einen Anstieg der Diodenspannung V auf den Wert V_{FP} (Einschaltüberspannung), die nach der Vorwärtserholzeit t_f auf den Wert der Flussspannung V_F absinkt. Hierbei kann eine hohe Verlustleistung P_V (vgl. Abschn. 2.2.3) an der pn-Diode auftreten. Das Umschaltverhalten der pn-Diode bei Stromkommunikation ist in Abb. 2.22a veranschaulicht.

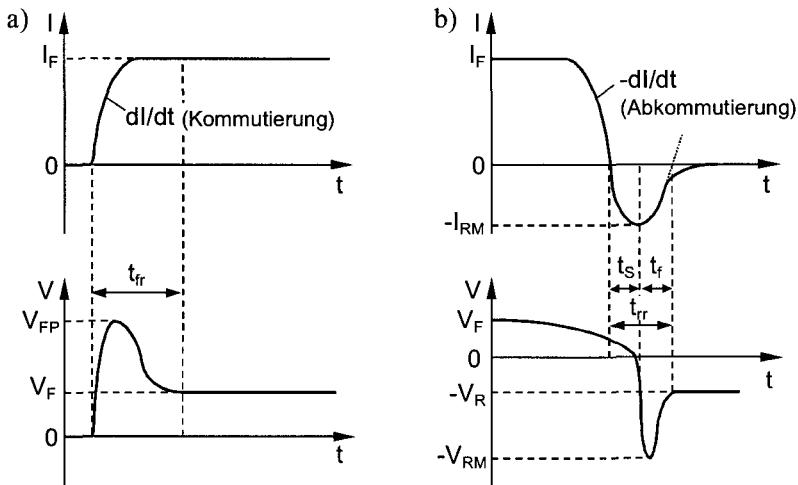


Abb. 2.23. Schaltverhalten der pn-Diode bei induktiver Last a) Umschaltverhalten bei Stromkommutierung und b) Abschaltverhalten bei Abkommunierung

Beim Schalten der pn-Diode in Sperrpolung begrenzt die Abkommunierungssteilheit $-dI/dt$ den Abschaltvorgang so, dass nach Nulldurchgang des Diodenstroms I die pn-Diode noch für eine bestimmte Zeit (Speicherzeit t_S) leitend bleibt. Nach der Speicherzeit geht die Diode in den Sperrzustand über und erreicht die maximale Sperrspannung V_{RM} (vgl. Abb. 2.23b). Nach der Abfallzeit t_f steigt die Sperrspannung auf den Wert V_R an. Die Summe aus der Speicherzeit t_S und der Abfallzeit t_f charakterisiert dabei die Rückwärtserholzeit t_{rr} der pn-Diode.

Dioden, die speziell auf das Schaltverhalten optimiert sind, werden Schaltdioden genannt. Weitere Diodentypen werden in Abschn. 2.7 näher beschrieben.

2.7 Diodentypen

2.7.1 pin-Diode

Durch Separation des p- und n-Gebietes einer pn-Diode durch eine sehr schwach dotierte bzw. intrinsische Halbleiterschicht (i-Zone) entsteht die sog. *pin-Diode*. Die pin-Diode ist durch eine hohe Ladungsträgerlebensdauer τ in der Raumladungszone und eine lange Rückwärtserholzeit t_{rr} (vgl. Abschn. 2.6.1) charakterisiert. Im folgenden wird von einer schmalen i-Zone der Struktur ausgegangen (Weite der i-Zone $w_i \ll L_{Di}$), wodurch vorteilhafte Hochfrequenzeigenschaften der pin-Diode resultieren. Pin-Dioden mit einer sehr breiten i-Zone ($w_i \gg L_{Di}$) werden in Kap. 5 in der Anwendung als Leistungsbauelement gesondert behandelt. Der schematische Aufbau einer pin-Diode und der ortsabhängige Verlauf der Raumladungsdichte $\rho(x)$ und der Feldstärke $E(x)$ ist in Abb. 2.24 dargestellt.

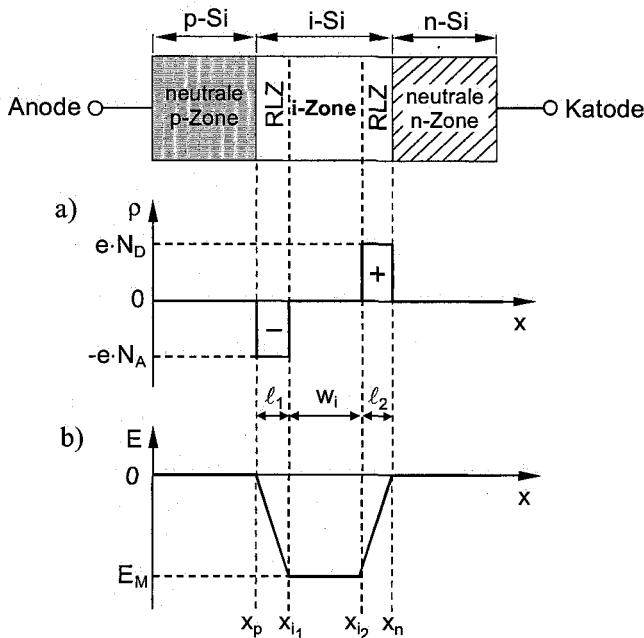


Abb. 2.24. Schematischer Aufbau einer pin-Diode mit ortsabhängigem Verlauf a) der Raumladungsdichte und b) des elektrischen Feldes

Zwischen p- und i-Zone sowie n- und i-Zone bildet sich jeweils eine Raumladungszone aus ($x_p \leq \ell_1 \leq x_{i1}$ und $x_{i2} \leq \ell_2 \leq x_n$). Im Bereich $x_{i1} < w_i < x_{i2}$ der i-Zone ist die Raumladungsdichte $\rho(x) = 0$ und die Feldstärke E ist konstant ($E(x) = E_M$). Die Durchbruchspannung der pin-Diode ergibt sich dann näherungsweise aus dem Produkt aus maximaler Feldstärke E_M und der Weite w_i der i-Zone ($V_{Br} \approx E_M \cdot w_i$). Auf Grund der zwei Raumladungszenosen treten im Bänderschema der pin-Diode zwei Potentialbarrieren auf, die jeweils durch die entsprechende Diffusionsspannung V_D (vgl. Abschn. 2.1.1) charakterisiert sind. Das Bänderdiagramm einer pin-Diode im thermodynamischen Gleichgewicht ist in Abb. 2.25 dargestellt.

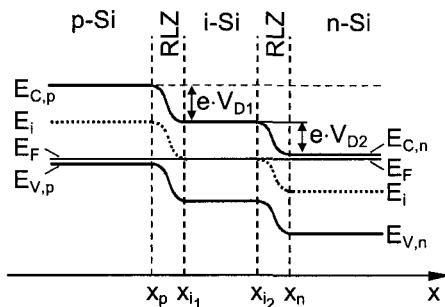


Abb. 2.25. Bänderdiagramm einer pin-Diode im thermodynamischen Gleichgewicht

Bei Sperrpolung besitzt die pin-Diode der Fläche A aufgrund der gegenüber w_i geringen Änderung der Raumladungszonenweiten eine nahezu spannungsunabhängige Sperrschnittskapazität C_S (vgl. Abschn. 2.5.1):

$$C_S = \frac{\epsilon_0 \cdot \epsilon_r \cdot A}{w_i}. \quad (2.36)$$

Darüber hinaus hängt der Bahnwiderstand R_B der pin-Diode im wesentlichen von der Geometrie und Dotierungsdichte der n- und p-Zone ab, sodass sich für die Grenzfrequenz f_G der pin-Diode folgende Beziehung ergibt:

$$f_G = \frac{1}{C_S \cdot R_B} = \frac{w_i}{\epsilon_0 \cdot \epsilon_r \cdot A \cdot R_B}. \quad (2.37)$$

Bei Flusspolung der pin-Diode erfolgt die Injektion von Ladungsträgern von den hoch dotierten Randzonen in die i-Zone, sodass die i-Zone mit Ladungsträgern überschwemmt wird. Die Kennliniengleichung einer pin-Diode mit symmetrischem Aufbau lautet dann:

$$I = 2I_S \cdot \left(\exp\left(\frac{V}{V_\vartheta}\right) - 1 \right) + 2I_{S_i} \cdot \left(\exp\left(\frac{V}{V_\vartheta}\right) - 1 \right), \quad (2.38)$$

wobei die Sperrsättigungsströme I_S und I_{S_i} von der Geometrie der drei Zonen abhängen und sich gemäß Gl. (2.23) berechnen. Der Widerstand R der pin-Diode bei Flusspolung ist vom Diodenstrom I und der Weite w_i der i-Zone abhängig und berechnet sich über die Ladungsträgerlebensdauer τ und Ladungsträgerbeweglichkeiten μ_n und μ_p wie folgt:

$$R = \frac{w_i^2}{I \cdot \tau \cdot (\mu_n + \mu_p)}. \quad (2.39)$$

Ein typischer Widerstandsverlauf einer pin-Diode als Funktion des Diodenstroms bei Flusspolung ($V > 0$) und das Schaltsymbol der pin-Diode sind in Abb. 2.25 dargestellt.

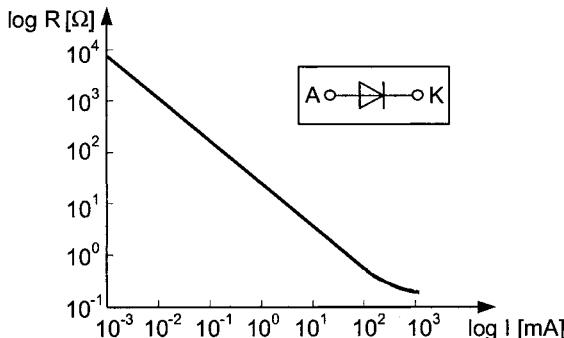


Abb. 2.26. Typischer Widerstandsverlauf einer pin-Diode als Funktion des Diodenstroms bei Flusspolung ($V > 0$), Schaltsymbol der pin-Diode

Da der Widerstand der pin-Diode über mehrere Größenordnungen umgekehrt proportional zum Diodenstrom ist, kann die pin-Diode als regelbarer Widerstand verwendet werden.

Die Anwendung der pin-Diode mit schmaler i-Zone erfolgt insbesondere als Hochfrequenz-Schalter. Hierbei erfolgt bei großer Aussteuerung ein Umschalten von Sperrpolung in Flusspolung und umgekehrt. Bei Flusspolung ist die pin-Diode dabei sehr niederohmig und bei hoher Sperrspannung ist die Sperrsichtkapazität konstant, wodurch Grenzfrequenzen von über 100 GHz erreicht werden. Verwendet man in der pin-Diode anstatt einer schmalen eine sehr breite i-Zone, so resultiert daraus eine sehr hohe Durchbruchspannung, womit die pin-Diode als Leistungsbauelement Anwendung findet (vgl. Abschn. 5.1).

2.7.2 Z-Diode

Z-Dioden (auch Zenerdioden genannt) sind relativ hoch dotierte pn-Dioden, die sich bei Flusspolung und Sperrpolung unterhalb der sog. *Zenerspannung* wie herkömmliche pn-Dioden verhalten. Z-Dioden sind insbesondere für den Betrieb in Sperrpolung ausgelegt und weisen eine definierte Durchbruchscharakteristik auf. Bei Spannungen unterhalb der Zenerspannung V_Z erfolgt ein abrupter Sperrstromanstieg, der entweder auf dem Zenereffekt (vgl. Abschn. 2.2.2) oder auf dem Lawineneffekt (vgl. Abschn. 2.2.1) basiert. Die Strom-Spannungs-Kennlinie nach dem Diffusionsmodell und das Schaltsymbol einer Z-Diode sind in Abb. 2.27 dargestellt.

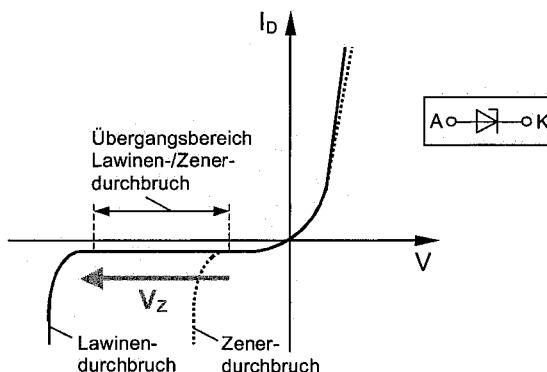


Abb. 2.27. Schaltsymbol und schematische Diodenkennlinie einer Z-Diode nach dem Diffusionsmodell mit Kennzeichnung der auftretenden Durchbruchseffekte

Bei geringer Zenerspannung (z. B. $2,5V \leq V_Z \leq 5V$) erfolgt der Durchbruch durch den Zenereffekt, wobei die Durchbruchspannung einen negativen Temperaturkoeffizienten hat. Bei hoher Zenerspannung (z. B. $V_Z \geq 15V$) erfolgt der Durchbruch durch den Lawineneffekt, wodurch die Durchbruchspannung einen positiven Temperaturkoeffizienten besitzt.

Im Übergangsbereich (z. B. $5V < V_Z < 15V$) ist sowohl der Zenereffekt als auch der Lawineneffekt am Durchbruch beteiligt (vgl. Abb. 2.27). Eine weitere wesentliche Größe zur Charakterisierung von Z-Dioden ist der differentielle Widerstand (Zenerwiderstand) $r_Z = \partial V_Z / \partial I$.

Die Z-Diode findet insbesondere Anwendung zur Spannungsbegrenzung (vgl. Abschn. 2.8.5) und zur Spannungsstabilisierung.

2.7.3 Tunneldiode

Tunneldioden (auch Esaki-Dioden genannt) sind sehr hoch dotierte Dioden mit sehr dünner Sperrsicht und zeichnen sich durch einen Kennlinienbereich mit negativem differentiellem Widerstand aus (vgl. Abb. 2.29). Elektronen aus Zuständen nahe der Leitungsbandkante „durchtunnen“ die Sperrsicht bei Flusspolung, obwohl die angelegte Spannung V unterhalb der Schwellspannung V_S liegt.

Durch die sehr hohe Dotierungsdichte der Halbleitergebiete (Entartung, vgl. Abschn. 1.3.4) liegt das Fermi niveau E_F sowohl im Leitungsband des n-Halbleiters als auch im Valenzband des p-Halbleiters der Tunneldiode; es tritt teilweise eine Überlappung der Energiebänder auf. Dadurch stehen sich mit Elektronen besetzte und unbesetzte Energieniveaus gegenüber. Durch die geringe Sperrsichtbreite und die hohe elektrische Feldstärke am pn-Übergang kommt es an den Stellen der Bänderüberlappung zum „Durchtunnen“ von Ladungsträgern und einem merklichen Stromfluss schon bei geringen Spannungen. Dieser Mechanismus ist in Abb. 2.28 im Bändermodell der Tunneldiode im Betrieb bei geringer Flusspolung dargestellt.

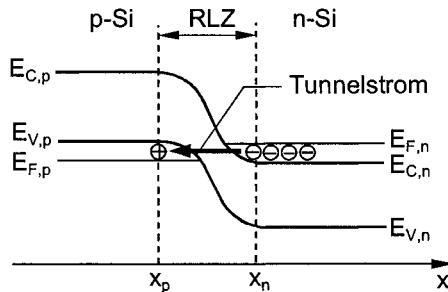


Abb. 2.28. Bändermodell der Tunneldiode bei geringer Flussspannung ($V < 0$)

Nach Erreichen des Peakstroms I_P , der bei der Spannung V_P durch die Tunneldiode fließt, beginnt der Diodenstrom I auf den Talwert I_V abzusinken. Das Absinken erfolgt für Diodenspannungen im Bereich $V_P < V \leq V_V$ im Maß der Abnahme der Überlappung der Zustände von Leitungsbandelektronen und gegenüberliegenden unbesetzten Energiezuständen. Bei Diodenspannungen von $V > V_V$ ist schließlich die Bandüberlappung aufgehoben; es findet kein Tunnelvorgang mehr statt. Die Tunneldiode verhält sich bei Spannungen $V > V_V$ wie eine herkömmliche pn-Diode (vgl. Abschn. 2.2). Im Spannungsbereich $V_P < V \leq V_V$ besitzt die Tunneldi-

ode damit einen negativen differentiellen Widerstand $-dV/dI$. Wird die Tunneldiode in Sperrrichtung betrieben, so befindet sich die Diode im Zenerdurchbruch (vgl. Abschn. 2.2.2). Die Diodenkennlinie und das Schaltsymbol einer Tunneldiode ist in Abb. 2.29 dargestellt.

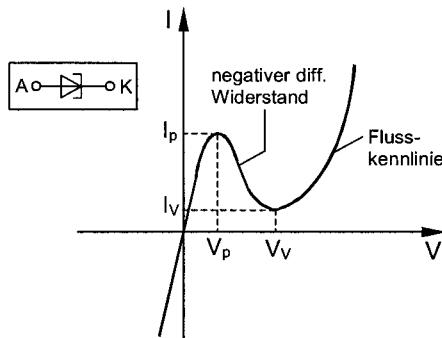


Abb. 2.29. Schaltsymbol und Diodenkennlinie einer Tunneldiode mit Kennzeichnung des Bereiches des negativen differentiellen Widerstandes

Die Kennlinie der Tunneldiode kann näherungsweise durch folgende Beziehung beschrieben werden:

$$I = I_S \cdot \exp\left(\frac{V}{V_\vartheta}\right) + I_P \cdot \frac{V}{V_P} \cdot \exp\left(1 - \frac{V}{V_p}\right) + I_V \cdot \exp(\kappa \cdot (V - V_V)). \quad (2.40)$$

In Gl. (2.40) ist κ ein Fitting-Parameter zur Anpassung an eine gemessene Diodenkennlinie und I_S beschreibt den Sperrsättigungsstrom der Tunneldiode (vgl. Abschn. 2.1.2.2).

Tunneldioden finden insbesondere Anwendung in HF-Oszillatoren, in denen sie bei Nutzung des Kennlinienbereichs negativer Steigung als „negativer differentieller Widerstand“ verwendet werden.

2.7.4 Schottky-Diode

Schottky-Dioden sind Dioden mit Metall-Halbleiter-Übergang, die als Gleichrichter mit niedriger Schwellspannung oder als niederohmiger Kontakt (Schottky-Kontakt) verwendet werden. Die Eigenschaften des Metall-Halbleiter-Überganges sind insbesondere von der Elektronenaustrittsarbeit $e \cdot \phi_{HL}$ bzw. der Elektronenaffinität χ_{HL} des Halbleiters abhängig.

Wird ein Metall mit einem Halbleiter (z. B. n-Typ-Halbleiter) gemäß Abb. 2.30b in Kontakt gebracht ($x_{Me} = x_{HL} = x_m$), so diffundieren wegen der kleineren Austrittsarbeit ($e \cdot \phi_{HL} < e \cdot \phi_{Me}$) Elektronen vom Halbleiter in das Metall. Dadurch entsteht eine Verarmungszone der Breite ℓ_D , die aus einer positiven Raumladung (ortsfeste Donatoren) im Halbleiter und einer Elektronenladung sehr geringer

Ausdehnung im Metall besteht. Die positive Raumladung im Halbleiter stellt eine Barriere für den Elektronenfluss vom Halbleiter zum Metall dar. In Abb. 2.30 sind die Bänderschemata einer Schottky-Diode vor und nach der Kontaktierung von Metall und Halbleiter dargestellt. Als Bezugsenergie dient hier die Vakuumenergie E_0 , die ein Elektron zum Verlassen eines Festkörpers mindestens benötigt.

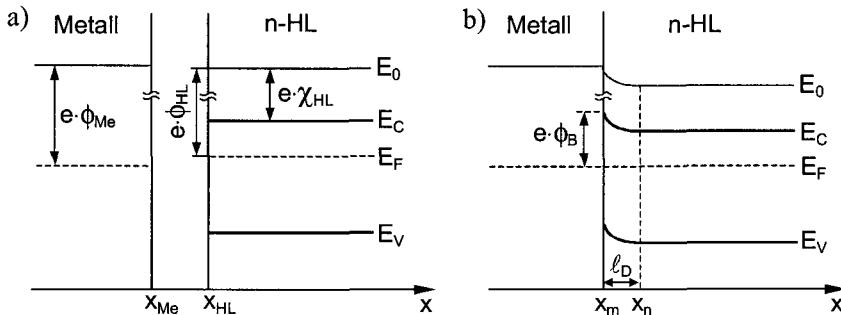


Abb. 2.30. Bänderschema einer Schottky-Diode a) vor der Kontaktierung von Metall und Halbleiter b) nach der Kontaktierung von Metall und Halbleiter

Der Ladungsträgertransport erfolgt bei der Schottky-Diode (im Gegensatz zur pn-Diode) nahezu ausschließlich durch Majoritätsträger. Durch die Ladungsneutralität existiert eine negative Grenzflächenladung im Metall, wodurch eine Potentialbarriere $e \cdot \phi_M$ entsteht. Bei negativer Spannung am Halbleiter wird die Potentialbarriere zwischen Halbleiter und Metall abgebaut; es kann ein Elektronenfluss zum Metall erfolgen (Schottky-Diode ist leitend). Wird eine positive Spannung am Halbleiter angelegt, so bleibt die Potentialbarriere $e \cdot \phi_B$ zwischen Halbleiter und Metall bestehen und die Schottky-Diode sperrt.

Für die Strom-Spannungs-Kennlinie einer Schottky-Diode ergibt sich die gleiche exponentielle Abhängigkeit wie bei der Kennlinie einer pn-Diode (vgl. Abschn. 2.1.2.2, Gl. 2.22). Der Unterschied zum Verhalten einer pn-Diode besteht im erheblich höheren Sperrsättigungsstrom I_S und seiner Temperaturabhängigkeit. Schottky-Dioden besitzen darüber hinaus eine geringere Schwellspannung V_S und eine geringere Durchbruchspannung V_B , als pn-Dioden.

Mit Schottky-Dioden sind sehr kurze Schaltzeiten ($t < 0,1\text{ns}$) realisierbar. Dadurch eignen sich Schottky-Dioden insbesondere als Schaltdioden in Höchstfrequenzschaltungen und schnellen Logikschaltungen.

Bei der Anwendung der Schottky-Diode als Schottky-Kontakt (zwischen einem Metall und einem Halbleiter) lässt sich ein sehr geringer Kontaktwiderstand realisieren, wenn der Ladungsträgertransport nicht über die Potentialbarriere erfolgt, sondern mittels Tunnelprozess durch die Potentialbarriere hindurch. Dazu muss die Potentialbarriere hinreichend schmal sein, was durch Einfügen einer sehr hoch dotierten (entarteten) Halbleiterzone zwischen Metall und Halbleiter realisiert wird. In Abb. 2.31 ist der Tunnelmechanismus am Schottky-Kontakt an Hand eines Bänderschemas veranschaulicht und das Schaltsymbol der Schottky-Diode dargestellt.

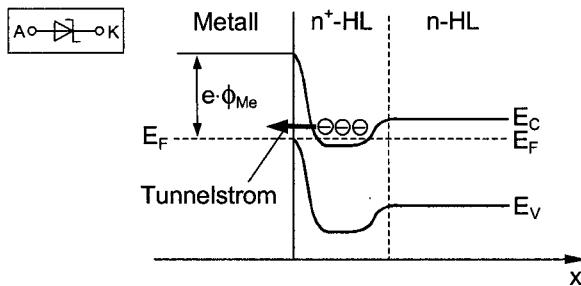


Abb. 2.31. Schaltsymbol und Bänderdiagramm einer Schottky-Diode mit hoch dotierter Halbleiterzone (hier n⁺-Zone) bei ohmschem Kontakt mit Kennzeichnung des auftretenden Tunnelstroms

2.7.5 Kapazitätsdiode

Kapazitätsdioden bzw. sog. *Sperrsichtvaraktoren* sind pn-Dioden, deren spannungsabhängige Sperrsichtkapazität als charakteristische Größe genutzt wird. Gemäß Abschn. 2.5.1 bewirkt eine Spannungsänderung am pn-Übergang eine Ladungsänderung in der Raumladungszone, woraus eine Längenänderung der Raumladungszone resultiert. Die Sperrsichtkapazität C_s , die das Verhältnis aus der Ladungsänderung ΔQ und der Spannungsänderung ΔV darstellt, ändert sich bei Variation der Sperrspannung V_R .

Die Kapazitäts-Spannungs-Charakteristik wird metallurgisch über die Wahl des Dotierprofils des pn-Überganges festgelegt. In Abb. 2.32 ist eine typische Kapazitäts-Sperrspannungs-Kennlinie und das Schaltsymbol einer Kapazitätsdiode dargestellt.

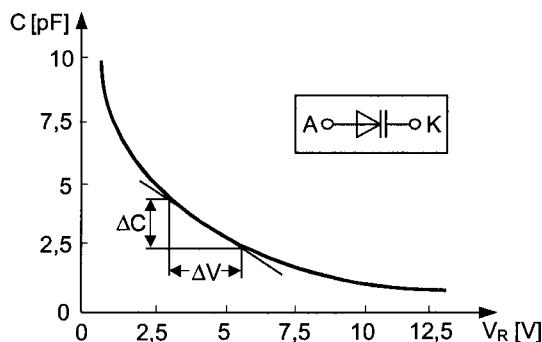


Abb. 2.32. Typische Kapazitäts-Sperrspannungs-Kennlinie und Schaltsymbol einer Kapazitätsdiode

Die wesentlichen Kenngrößen einer Kapazitätsdiode sind das Verhältnis aus maximaler und minimaler Kapazität C_{max}/C_{min} , die Steilheit $\partial C/\partial V$ der Kapazitäts-

Spannungs-Charakteristik und die Grenzfrequenz. Die Grenzfrequenz f_G berechnet sich mit Hilfe des Bahnwiderstandes R_B und der Sperrsichtkapazität C_S wie folgt:

$$f_G = \frac{I}{2\pi \cdot R_B \cdot C_S}. \quad (2.41)$$

Um sehr hohe Grenzfrequenzen zu erreichen, wird neben Silizium auch GaAs als Halbleitermaterial verwendet, da dieses eine erheblich höhere Ladungsträgerbeweglichkeit besitzt und sich somit der Bahnwiderstand R_B der Kapazitätsdiode reduziert.

Kapazitätsdioden werden insbesondere in Schwingkreisen eingesetzt, in denen die Resonanzfrequenz mit Hilfe einer Abstimmspannung geändert wird und eine Frequenzmodulation erfolgt (vgl. Abschn. 2.8.4).

2.7.6 Fotodiode

Fotodioden ermöglichen die Umwandlung von Lichtleistung P_L (Photonenstrahlung) in elektrische Leistung. Der Sperrstrom einer pn-Diode wird durch die Ladungsträgergeneration in der Raumladungszone und den angrenzenden Gebieten der neutralen Zonen bestimmt.

Wie in herkömmlichen pn-Dioden tritt in Fotodioden eine thermische Generation von Elektron-Loch-Paaren auf, wodurch ein Sperrstrom bzw. sog. *Dunkelstrom* fließt. Darüber hinaus werden Elektron-Loch-Paare optisch generiert. Die Absorption von Photonen bewirkt dabei die Anhebung von Elektronen aus dem Valenzband in das Leitungsband und das Fließen eines Fotostroms I_L . Die optische und thermische Generation von Ladungsträgern in einer pn-Fotodiode ist in Abb. 2.33 veranschaulicht.

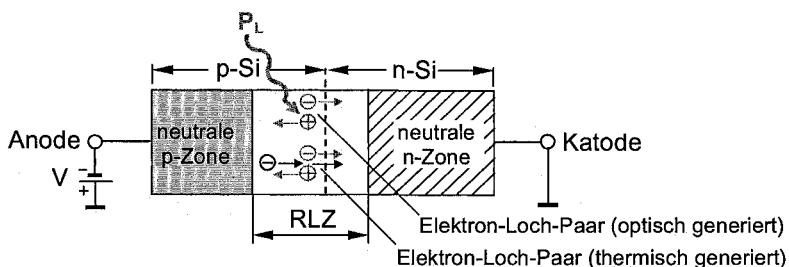


Abb. 2.33. Prinzip der optischen und thermischen Generation von Ladungsträgern und Stromfluss in einer pn-Fotodiode bei Sperrpolung ($V < 0$) und Einstrahlung der Lichtleistung P_L

Nach dem Diffusionsmodell (vgl. Abschn. 2.2.1) resultiert für die Diodenkennlinie das in Abb. 2.34 dargestellte Verhalten einer Fotodiode im unbestrahlten ($P_L = 0$) und bestrahlten ($P_L \neq 0$) Zustand.

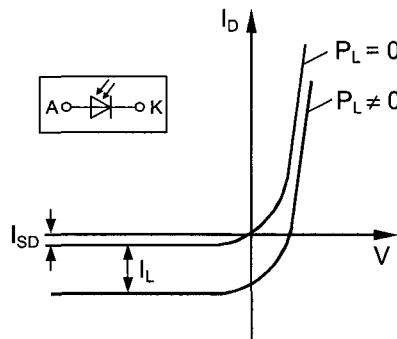


Abb. 2.34. Schaltsymbol und Diodenkennlinie einer Photodiode nach dem Diffusionsmodell im unbestrahlten ($P_L = 0$) und bestrahlten ($P_L \neq 0$) Zustand mit Dunkelstrom I_{SD} und Photostrom I_L

Die Kennliniengleichung der Photodiode unter Einbeziehung des Photostroms I_L folgt aus Gl. (2.22):

$$I_D = I_{SD} \cdot \left(\exp\left(\frac{V}{V_\vartheta}\right) - 1 \right) - I_L, \quad (2.42)$$

wobei sich der Photostrom I_L berechnet mit:

$$I_L = \frac{P_L}{h \cdot f} \cdot e \cdot n_Q(\lambda) = S_\lambda \cdot P_L. \quad (2.43)$$

In Gl. (2.43) gibt das Verhältnis von P_L zu $h \cdot f$ die Anzahl der absorbierten Photonen an. Die Größe $n_Q(\lambda)$ beschreibt den wellenlängenabhängigen externen Quantenwirkungsgrad und S_λ charakterisiert die spektrale Empfindlichkeit der Photodiode. Der externe Quantenwirkungsgrad n_Q bezeichnet das Verhältnis zwischen den erzeugten Elektron-Loch-Paaren, die zum Nutzstrom beitragen, und der Zahl der auftreffenden Photonen. Ein typischer Verlauf des Quantenwirkungsgrades $n_Q(\lambda)$ einer Si-Photodiode ist in Abb. 2.35 dargestellt.

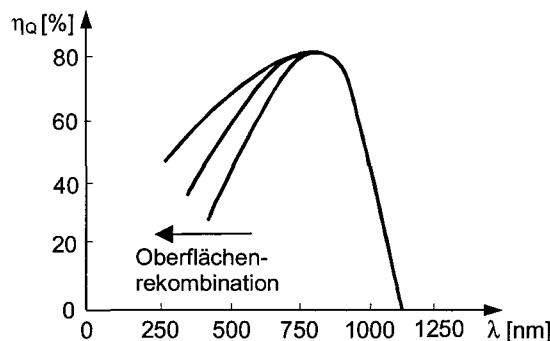


Abb. 2.35. Typischer Verlauf des externen Quantenwirkungsgrades η_Q einer Si-Photodiode

Der Quantenwirkungsgrad wird im wesentlichen durch zwei Effekte beeinflusst: Mit steigender Wellenlänge erfolgt eine abrupte Abnahme des Quantenwirkungsgrades, da die Photonenenergie E_{Ph} ab einem bestimmten Wert nicht mehr ausreicht, Elektron-Loch-Paare zu generieren. Der Grenzwert, bis zu dem Elektron-Loch-Paare gebildet werden können, lautet:

$$E_{Ph} = h \cdot f = h \cdot \frac{c}{\lambda} > E_G . \quad (2.44)$$

Aus Gl. (2.44) folgt, dass die Grenzwellenlänge λ zur Elektron-Loch-Paarbildung in Silizium ca. 1,1 µm beträgt. Der zweite Effekt zur Beeinflussung des Quantenwirkungsgrades bei abnehmender Wellenlänge stellt die zunehmende Absorption von Photonen nahe der Halbleiteroberfläche dar. Dort existiert eine hohe Anzahl von Rekombinationszentren, an denen die Ladungsträger nichtstrahlend rekombinieren (vgl. Abschn. 2.7.8). Der externe Quantenwirkungsgrad n_Q und die spektrale Empfindlichkeit S_λ von Silizium- und Germanium-Fotodioden mit Gegenüberstellung zur sog. *relativen Augenempfindlichkeit* ist in Abb. 2.36 dargestellt.

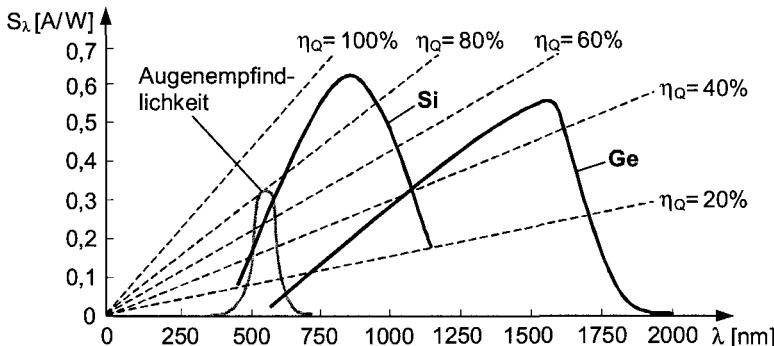


Abb. 2.36. Spektrale Empfindlichkeiten S_λ von Si- und Ge-Fotodioden und externer Quantenwirkungsgrad n_Q (bezogen auf $E_{G, Si}$) und relative Augenempfindlichkeit

Die relative Augenempfindlichkeit stellt den spektralen Hellempfindlichkeitsgrad des photometrischen Normalbeobachters dar und dient zur Beschreibung der Wirkung von Licht.

Fotodioden finden beispielsweise Anwendung zur Lichtintensitätsmessung und in der optischen Nachrichtenübertragung. Vorteilhaft ist dabei der hohe Dynamikbereich und die geringe Ansprechzeit der Fotodiode. Betreibt man eine Fotodiode nahe am Lawinendurchbruch (vgl. Abschn. 2.2.1), so wird der Fotostrom zusätzlich durch Lawinenmultiplikation verstärkt (Avalanche- oder Lawinen-Fotodiode).

2.7.7 Solarzelle

Die Solarzelle ist eine spezielle pn-Diode, die hinsichtlich der Umwandlung von (Sonnen-) Licht in elektrische Energie optimiert ist. Wie bei der Fotodiode (vgl.

Abschn. 2.7.6) werden bei Einstrahlung der Lichtleistung P_L Photonen im Halbleiter absorbiert. Die Energie dieser Photonen dient dazu, Elektronen aus dem Valenzband in das Leitungsband anzuheben, wodurch Elektron-Loch-Paare gebildet werden und das Fließen eines Fotostroms I_L resultiert. Die Funktionsweise einer Solarzelle ist in Abb. 2.37 veranschaulicht.

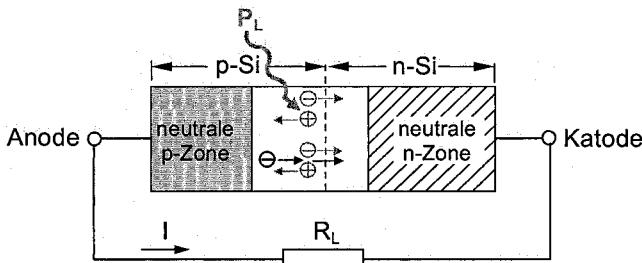


Abb. 2.37. Schaltprinzip einer pn-Diode im Betrieb als Solarzelle

Die Strom-Spannungs-Kennlinie einer Solarzelle (Abb. 2.38b) kann aus dem Ersatzschaltbild in Abb. 2.38a bestimmt werden. Hierbei muss das Wirken des parasitären Parallelwiderstandes R_P und des Bahnwiderstandes R_B berücksichtigt werden.

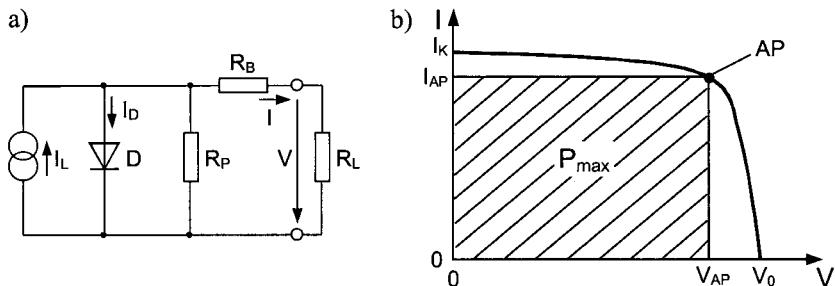


Abb. 2.38. a) Ersatzschaltung und b) Strom-Spannungs-Kennlinie einer Solarzelle mit Arbeitspunkt (AP) und maximaler Leistung P_{max}

Der von einer Solarzelle an einen elektrischen Verbraucher R_L abgegebene Strom I kann wie folgt berechnet werden:

$$I = I_L - I_S \cdot \left[\exp\left(\frac{V - I \cdot R_B}{V_\theta}\right) - 1 \right] - \frac{V - I \cdot R_B}{R_P}. \quad (2.45)$$

Der Kurzschlussstrom I_K einer Solarzelle entspricht dem Fotostrom I_L ($I_K = I_L$) und verhält sich proportional zur Lichtleistung P_L . Die Leerlaufspannung V_0 einer Solarzelle stellt eine Funktion des Bandabstandes E_G , der Temperatur und der Lichtleistung P_L dar. Die Leerlaufspannung V_0 tritt auf, wenn der Gesamtstrom durch die Solarzelle Null ist und berechnet sich ohne Berücksichtigung von R_p zu:

$$V_0 = \frac{kT}{e} \cdot \ln \left(\frac{I_L(P_L)}{I_S(E_G, T)} + 1 \right). \quad (2.46)$$

Eine steigende Lichtleistung hat demnach eine schwach zunehmende Leerlaufspannung V_0 der Solarzelle zur Folge. Wird durch die Materialauswahl ein größerer Bandabstand E_G festgelegt, so resultiert ein geringerer Sperrstrom I_S und damit eine höhere Leerlaufspannung der Solarzelle. Bei einer Temperaturerhöhung steigt der Sperrstrom und die Leerlaufspannung sinkt. Da beim Betrieb einer Solarzelle eine hohe Leerlaufspannung erwünscht ist, eignen sich die Materialien Silizium und GaAs hinsichtlich ihres größeren Bandabstandes deutlich besser als Germanium. Die Verwendung von Materialien mit noch höherem Bandabstand als GaAs ist nicht sinnvoll, da die Absorptionskante dann im UV-Bereich liegt und somit das Sonnenpektrum nicht sinnvoll genutzt werden kann. Ein Vergleich der spektralen Empfindlichkeit einer Si-Solarzelle und der spektralen Bestrahlungsstärke E_λ der Sonne an der Erdoberfläche ist in Abb. 2.39 dargestellt.

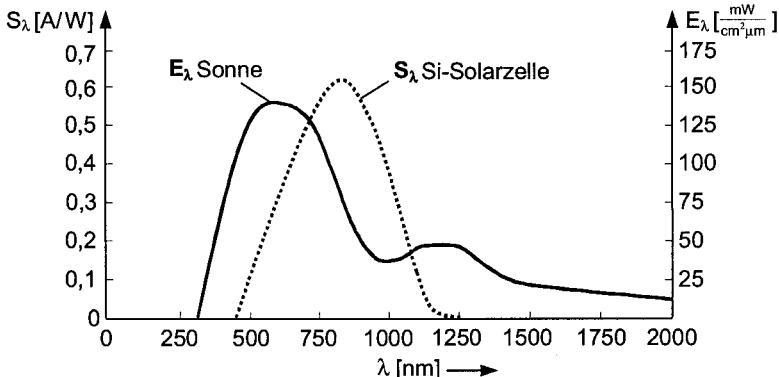


Abb. 2.39. Vergleich der spektralen Empfindlichkeit S_λ einer Si-Solarzelle und der spektralen Bestrahlungsstärke E_λ der Sonne an der Erdoberfläche

Der Wirkungsgrad η einer Solarzelle wird über das Verhältnis der maximalen Leistungsabgabe P_{max} zur einfallenden Lichtleistung P_L bestimmt. Unter Einbeziehung des sog. *Füllfaktors* F , der das Verhältnis von P_{max} zu $I_K \cdot V_0$ angibt und vom Lastwiderstand R_L abhängt, erhält man für den Wirkungsgrad einer Solarzelle folgenden Ausdruck:

$$\eta = \frac{P_{max}}{P_L} = F \cdot \frac{V_0 \cdot I_K}{P_L}. \quad (2.47)$$

Die technologische Optimierung des Wirkungsgrades von Solarzellen erfolgt unter anderem durch die Wahl des Bandabstandes bzw. des Halbleitermaterials, durch Reduktion von Reflexionsverlusten an der Solarzelloberfläche und Verringerung des Bahnwiderstandes R_B . Der schematische Aufbau einer Silizium-Solarzelle mit Kontaktfingerstruktur ist in Abb. 2.40 im Querschnitt dargestellt.

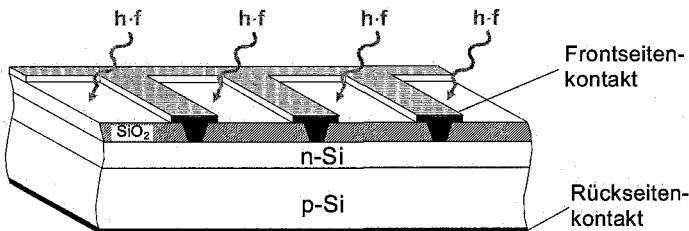


Abb. 2.40. Schematischer Aufbau einer Silizium-Solarzelle im Querschnitt mit Frontseitenkontakt in Fingerstruktur

Zur Erhöhung des Kurzschlussstroms oder der Leerlaufspannung ist es möglich, mehrere Solarzellen parallel oder seriell in Solarzellenarrays zu verschalten. Werden beispielsweise zwei identische Solarzellen parallel geschaltet und beide Solarzellen gleich stark bestrahlt, so verdoppelt sich bei konstanter gehaltener Spannung der an den Verbraucher abgegebene Gesamtstrom.

Solarzellen finden beispielsweise Anwendung zur Energieerzeugung für portable Elektronikapplikationen und zur Energieeinspeisung in das Stromnetz.

2.7.8 Lumineszenzdiode

Lumineszenzdioden (LEDs) sind pn-Dioden, die in Flusspolung elektrische Leistung in Lichtleistung umwandeln. Diese Lichtleistung entsteht durch die strahlende Rekombination von Elektron-Loch-Paaren in der Raumladungszone der Lumineszenzdiode. Der Wirkmechanismus ist in Abb. 2.41 veranschaulicht.

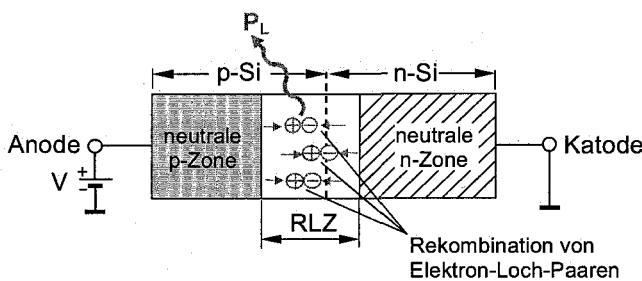


Abb. 2.41. Erzeugung der Lichtleistung P_L durch Rekombination von Elektron-Loch-Paaren in der Raumladungszone einer Lumineszenzdiode bei Flusspolung ($V > 0$)

Die Rekombination von Elektron-Loch-Paaren erfolgt entweder durch einen direkten oder einen indirekten Übergang von Elektronen aus dem Leitungs- in das Valenzband des Halbleiters. Insbesondere bei indirektem Übergang treten neben strahlenden auch nichtstrahlende Rekombinationsprozesse auf. Abb. 2.42 zeigt die Bänderdiagramme für strahlende Rekombination bei direktem und indirektem Bandübergang im Energie-Impuls ($E-k$)-Raum.

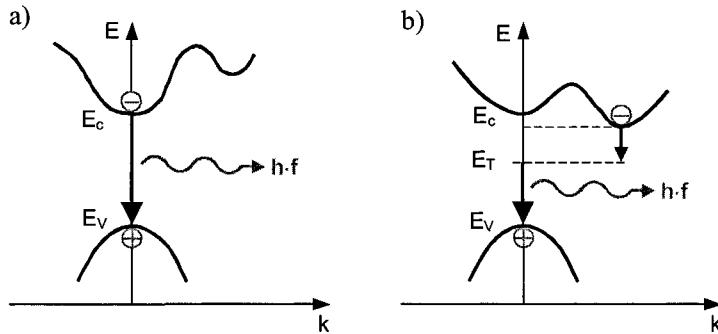


Abb. 2.42. Bänderschema im Energie-Impuls (E - k)-Raum für strahlende Rekombination bei a) direktem und b) indirektem Übergang

Das Verhältnis der strahlenden Rekombinationsrate R_{rad} zur Gesamtrekombinationsrate R bezeichnet die Quantenausbeute η_{rad} der Lumineszensdiode:

$$\eta_{rad} = \frac{R_{rad}}{R}. \quad (2.48)$$

Das Eigenabsorptionsverhalten des Halbleiters und Reflexionserscheinungen an der Halbleiteroberfläche bewirken eine Beeinträchtigung der Transmissionseigenschaften der Lumineszenzdiode. Insbesondere der hohe Brechungsindex von Halbleitern verringert den Grenzwinkel der Totalreflexion, sodass nur ein geringer Teil der erzeugten Photonen den Halbleiter verlassen kann. Das Verhältnis des tatsächlich von der Lumineszenzdiode ausgesandten Photonenstroms Φ_{ph} zum aufgewendeten Elektronenstrom I wird als externer Quantenwirkungsgrad η_Q bezeichnet:

$$\eta_Q = \frac{\Phi_{ph}}{I / e} = \frac{e \cdot P_L}{I \cdot h \cdot f}. \quad (2.49)$$

Der Photonenstrom Φ_{ph} ergibt sich aus dem Verhältnis der Lichtleistung P_L zur Photonenergie $h \cdot f$. Als Maß für die Effektivität der Lichtleistungserzeugung von Lumineszenzdioden dient der Leistungswirkungsgrad η_P , der das Verhältnis von Lichtleistung P_L zur elektrischen Leistung P_{el} darstellt und sich proportional zum externen Quantenwirkungsgrad η_Q verhält:

$$\eta_P = \frac{P_L}{P_{el}} = \frac{h \cdot f \cdot \Phi_{ph}}{I \cdot V} = \eta_Q \cdot \frac{h \cdot f}{eV}. \quad (2.50)$$

Zur Spezifikation von Lumineszenzdioden dient weiterhin die Helligkeit, die als physiologische Größe aus der Augenempfindlichkeit hervorgeht. Zur Bestimmung der Helligkeit wird der sog. *Leuchtwirkungsgrad* definiert, der die in Abb. 2.43 dargestellte Überlappung der relativen spektralen Intensitätsverteilung I_λ der Lumineszenzdiode mit der absoluten Augenempfindlichkeit K_λ beschreibt.

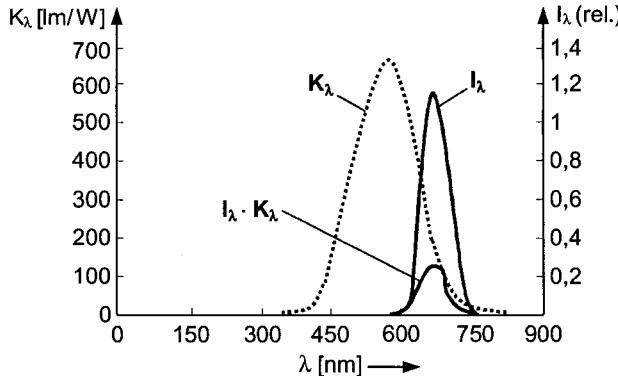


Abb. 2.43. Beispielhafte relative Intensitätsverteilung I_λ einer Lumineszenzdiode und absolute Augenempfindlichkeit K_λ zur Bestimmung des Leuchtwirkungsgrades

Der Leuchtwirkungsgrad η_L berechnet sich dann als Quotient aus dem Integral von $I_\lambda \cdot K_\lambda$ über dem Spektralbereich und dem Integral von I_λ [2.4]:

$$\eta_L = \frac{\int_0^{\infty} I_\lambda \cdot K_\lambda d\lambda}{\left(\int_0^{\infty} I_\lambda d\lambda \right)^{-1}}. \quad (2.51)$$

Die Helligkeit einer Lumineszenzdiode wird über die Leuchtdichte L_v beschrieben, die sich bei Berücksichtigung des Raumwinkels θ (Richtungsabhängigkeit der Lichtstrahlung) und der Emissionsfläche A wie folgt berechnet:

$$L_v = \frac{\eta_L \cdot P_L}{A \cdot \theta}. \quad (2.52)$$

Die Diodenkennlinie der Lumineszenzdiode verhält sich wie die Kennlinie einer pn-Diode (vgl. Abschn. 2.2). Die Farbe des emittierten Lichts von Lumineszenzdioden ist vom verwendeten Halbleitermaterial bzw. dessen Bandabstand abhängig. In der nachfolgenden Tabelle sind beispielhaft verschiedene Halbleitermaterialien mit zugehörigerem emittierter Wellenlänge und Farberscheinung dargestellt.

Material	ZnS	GaN	GaP	GaAs _x P _{1-x}	GaAs	InP
λ [nm]	340	440	550	580 - 800	870	925
Farbe	ultraviolett	blau	grün	gelb - rot	infrarot	infrarot

Durch die Verwendung von Mischkristallen kann der Bandabstand und somit die emittierte Wellenlänge der Lumineszenzdiode variiert werden. Wird beispielsweise im Material GaAs_xP_{1-x} ein Teil der As-Atome des GaAs-Kristalls durch P-Atome ersetzt, kann Licht im Wellenlängenbereich von 580nm bis 800nm erzeugt werden.

Für den Einsatz in der optischen Nachrichtentechnik finden vor allem Heterostrukturen (vgl. Abb. 2.45) Anwendung, in denen insbesondere quaternäre Verbindungshalbleiter (vgl. Kap. 1) verwendet werden. Diese können zum Beispiel in der Form $\text{In}_{1-x}\text{Ga}_x\text{As}_y\text{P}_{1-y}$ vorliegen und liefern abhängig von der Zusammensetzung Wellenlängen im Bereich von 920nm bis 1650nm.

Der schematische Aufbau einer einfachen Lumineszenzdiode mit diffundiertem pn-Übergang ist in Abb. 2.44 im Querschnitt dargestellt.

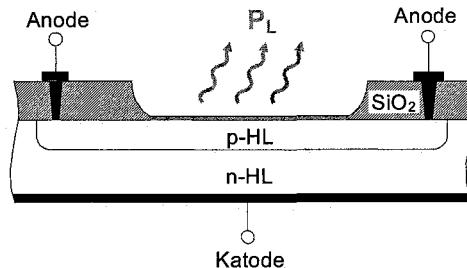


Abb. 2.44. Schematischer Aufbau einer Lumineszenzdiode mit diffundiertem pn-Übergang im Querschnitt

Die Erhöhung der Leuchtdichte von Lumineszenzdioden bei geringer Emissionsfläche kann durch Einkopplung des erzeugten Lichts in einen Lichtleiter erfolgen. In Abb. 2.45 ist der Aufbau einer Lumineszenzdiode mit Lichtleiter am Beispiel einer Doppelheterostruktur (LED mit Heteroübergängen aus Mischkristallen zur Erhöhung des Wirkungsgrades) im Querschnitt dargestellt.

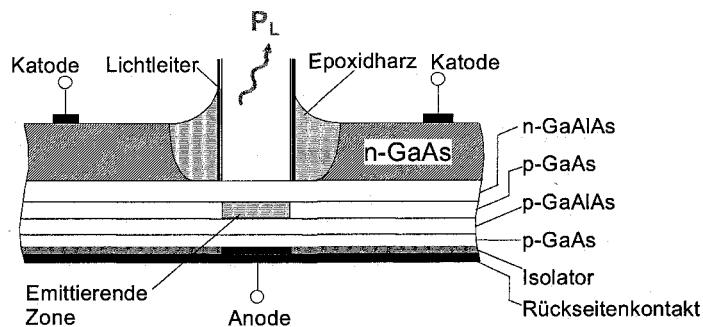


Abb. 2.45. Schematischer Aufbau einer Doppelheterostruktur-Lumineszenzdiode im Querschnitt

Bei zunehmender Betriebsdauer der Lumineszenzdiode nimmt bei konstantem Strom der externe Quantenwirkungsgrad η_O und damit die Leuchtdichte L_V ab (Degradation). Dieser Effekt ist in der Entstehung von Kristalldefekten begründet, die als zusätzliche nichtstrahlende Rekombinationszentren wirken. Die Abnahme der Lichtstärke erfolgt näherungsweise exponentiell mit der Zeit. Als Lebensdauer der Lumineszenzdiode wird die Zeit definiert, bei der die Leuchtdichte um 50%

gegenüber dem Nominalwert abgesunken ist. Die Lebensdauer ist stark von der Temperatur T und vom Flussstrom I abhängig.

Typische Anwendungsgebiete von Lumineszenzdioden sind die Anzeigetechnik (Symbole und symbolische Displays) und die optische Signalübertragung (Lichtleitertechnik und Optokoppler). Auch das optische Zünden von sog. *Fotothyristoren* (vgl. Kap. 5) kann mit Lumineszenzdioden erfolgen.

2.8 Dioden-Grundschatungen

2.8.1 Einweggleichrichter

Das grundlegende Prinzip zur Gleichrichtung einer Wechselspannung V_{in} in eine Gleichspannung V_{out} stellt der sog. *Einweggleichrichter* dar. Dabei wird ein Kondensator C über eine pn-Diode D aufgeladen, wodurch über einem Lastwiderstand R_L , der parallel zum Kondensator wirkt, eine pulsierende Gleichspannung abfällt. Das Schaltungsprinzip eines Einweggleichrichters mit ohmscher Last R_L ist in Abb. 2.46 dargestellt.

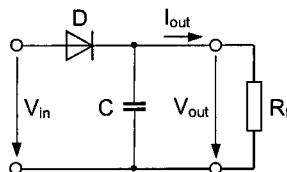


Abb. 2.46. Einweggleichrichter mit ohmscher Last R_L

Bei unbelastetem Ausgang ($R_L = \infty$, $V_{out} = V_{out0}$) wird der Kondensator C während der positiven Halbwelle der Eingangsspannung V_{in} auf den Scheitelwert $V_{out0} = \sqrt{2} \cdot V_{in,eff}$ abzüglich der Schwellspannung V_S (vgl. Abschn. 2.2.3) aufgeladen. Die maximale Sperrspannung V_R der Diode tritt auf, wenn die Eingangsspannung V_{in} ihren negativen Scheitelwert erreicht ($V_R \approx 2 \cdot \sqrt{2} \cdot V_{in,eff}$). Bei belastetem Ausgang ($V_{out} = V_{outL}$) entlädt die ohmsche Last R_L den Kondensator C , solange die pn-Diode D sperrt. Sobald die Leerlaufspannung der Wechselspannungsquelle um V_S größer als die Ausgangsspannung V_{outL} ist, wird der Kondensator wieder nachgeladen. Die erreichbare Ausgangsspannung hängt vom Innenwiderstand R_i der Wechselspannungsquelle ab und berechnet sich bei Verwendung eines hinreichend großen Kondensators wie folgt:

$$V_{outL} = V_{out0} \cdot \left(1 - \sqrt{\frac{R_i}{R_L}} \right) \quad \text{mit} \quad V_{out0} = \sqrt{2} \cdot V_{in,eff} - V_S . \quad (2.53)$$

Als Maß für die Qualität der Gleichrichtung dient die sog. *Brummspannung*, die die Differenz zwischen maximaler und minimaler Ausgangsspannungsamplitude

angibt (vgl. Abb. 2.47). Die Brummspannung V_{brss} berechnet sich abhängig von der Schaltungsdimensionierung und der Eingangsspannungs frequenz f_{in} zu:

$$V_{brss} = \frac{I_{out}}{C \cdot f_{in}} \left(1 - \sqrt{\frac{R_i}{R_L}} \right). \quad (2.54)$$

Der Eingangs- und Ausgangsspannungsverlauf bei Einweggleichrichtung ist in Abb. 2.47 schematisch dargestellt.

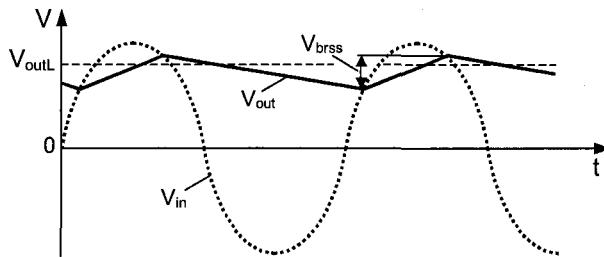


Abb. 2.47. Eingangs- und Ausgangsspannungsverlauf bei Einweggleichrichtung

Auf Grund des ungünstigen Verhältnisses von Nachlade- zu Entladzeit des Kondensators C sinkt die Ausgangsspannung V_{out} schon bei geringer Ausgangslast R_L stark ab. Die Einweggleichrichtung ist demnach nur bei geringen Ausgangsströmen sinnvoll einsetzbar.

2.8.2 Zweiwegbrückenschaltung

Eine Verbesserung des Verhältnisses von Nachlade- zu Entladzeit gegenüber des in Abschn. 2.8.1 betrachteten Einweggleichrichters erfolgt durch Aufladen eines Kondensators C sowohl während der positiven als auch während der negativen Halbwelle der Eingangsspannung V_{in} . Dies wird durch die sog. *Zweiwegbrückenschaltung* (auch Graetzbrücke genannt) erreicht, deren Schaltungsprinzip in Abb. 2.48 dargestellt ist.

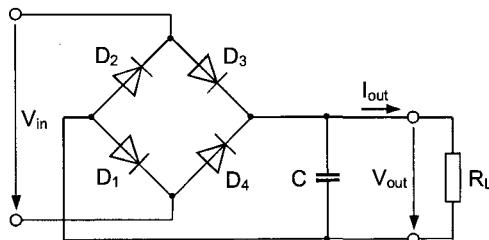


Abb. 2.48. Gleichrichter in Zweiwegbrückenschaltung (Graetzbrücke)

Bei positiver Halbwelle der Eingangsspannung V_{in} sind die Dioden D_1 und D_3 leitend, bei negativer Halbwelle leiten D_2 und D_4 . Am Ausgang stellt sich dann eine pulsierende Gleichspannung V_{out} ein. Die mittlere Last-Ausgangsspannung V_{outL} , die über dem Lastwiderstand R_L abfällt, berechnet sich bei Verwendung eines hinreichend großen Kondensators wie folgt:

$$V_{outL} = V_{out0} \left(1 - \sqrt{\frac{R_i}{2R_L}} \right) \quad \text{mit} \quad V_{out0} = \sqrt{2} \cdot V_{in,eff} - 2V_S . \quad (2.55)$$

In Gl. (2.55) bezeichnet R_i den Innenwiderstand der Wechselspannungsquelle und V_S die Schwellspannung der pn-Diode. Die Brummspannung V_{brss} als Maß für die Qualität der Gleichrichtung nimmt dann für die Zweiwegbrückenschaltung folgenden Wert an:

$$V_{brss} = \frac{I_{out}}{2C \cdot f_{in}} \left(1 - \sqrt{\frac{R_i}{2R_L}} \right) . \quad (2.56)$$

Demnach besitzt die Zweiwegbrückenschaltung eine erheblich geringere Brummspannung als der Einweggleichrichter und gestattet einen höheren Ausgangsstrom I_{out} .

2.8.3 Spannungsvervielfacher

Mit Hilfe der sog. *Delon-Schaltung* erfolgt eine Verdopplung der Eingangswechselspannung durch antiserielle Verschaltung von zwei Einweggleichrichtern (vgl. Abschn. 2.8.1). Dazu sind zwei pn-Dioden D_1 und D_2 so gepolt, dass ein Kondensator C_1 während der positiven Halbwelle und ein weiterer Kondensator C_2 während der negativen Halbwelle der Eingangsspannung V_{in} aufgeladen werden. Beide Kondensatoren sind in Reihe geschaltet, was bei gleicher Kondensatordimensionierung ($C_1 = C_2$) zu einer Verdopplung des Eingangsspannungsscheitelwertes über einer ohmschen Last R_L führt. Das Schaltungsprinzip der Delon-Schaltung ist in Abb. 2.49 dargestellt.

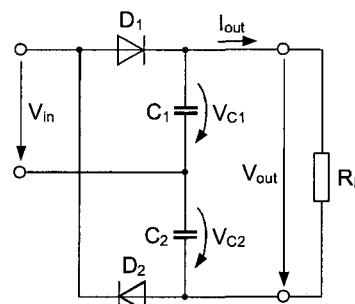


Abb. 2.49. Schaltungsprinzip zur Spannungsverdopplung mit Delon-Schaltung

Eine weitere Möglichkeit zur Spannungsverdopplung stellt die sog. *Villard-Schaltung* dar. Während der negativen Halbwelle der Eingangswechselspannung V_{in} wird ein Kondensator C_1 über einer pn-Diode D_1 auf die Amplitude des Eingangsspannungsscheitelwerts aufgeladen. Während der positiven Halbwelle sperrt die Diode D_1 und über eine zweite Diode D_2 wird ein Kondensator C_2 auf die Summe aus dem Eingangsspannungsscheitelwert und der Kondensatorspannung V_{C1} aufgeladen, wobei näherungsweise eine Verdopplung des Eingangsspannungsscheitelwertes resultiert.

Das Schaltungsprinzip der Villard-Schaltung ist in Abb. 2.50a dargestellt. Es bietet den Vorteil der Kaskadierbarkeit. In Abb. 2.50b eine entsprechende kaskadierte Villard-Schaltung zur Spannungsverviervfachung abgebildet.

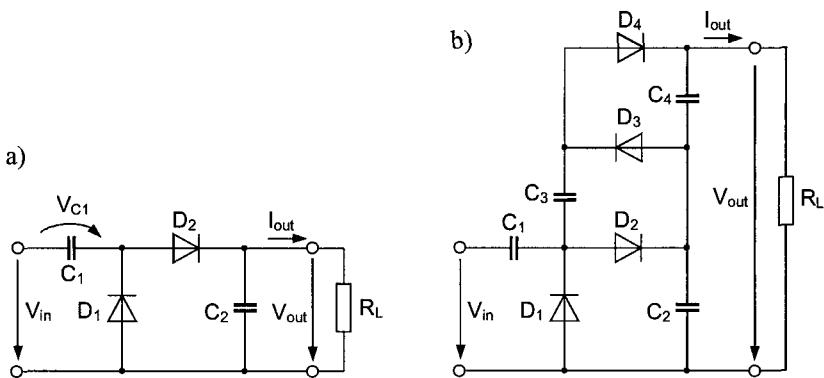


Abb. 2.50. Schaltungsprinzipien der Spannungsvervielfachung a) Spannungsverdopplung mit Villard-Schaltung und b) Spannungsverviersfachung mit Kaskadenschaltung

Bei der kaskadierten Villard-Schaltung wird während der negativen Halbwelle der Eingangsspannung V_{in} der Kondensator C_3 über die pn-Diode D_3 auf den doppelten Eingangsspannungsscheitelwert aufgeladen. Während der positiven Halbwelle fällt dann über der ohmschen Last R_L näherungsweise der vierfache Scheitelwert der Eingangswechselspannung ab, die sich aus den verdoppelten Spannungen über den Kapazitäten C_2 und C_4 zusammensetzt. Zur weiteren Erhöhung der Ausgangsspannung können weitere Stufen zur Kaskadenschaltung hinzugefügt werden.

2.8.4 Frequenzmodulation

Die Steuerung der Resonanzfrequenz eines Schwingkreises ermöglicht z. B. die Modulation einer hochfrequenten Trägerfrequenz mit einem niederfrequenten Signal (Frequenzmodulation). Die Variation der Resonanzfrequenz wird dabei mit Hilfe einer Kapazitätsdiode (vgl. Abschn. 2.7.5) realisiert, deren Arbeitspunkt durch die niederfrequente Spannung V_{var} verändert wird. Die Prinzipschaltung eines Parallelschwingkreises mit steuerbarer Resonanzfrequenz ist in Abb. 2.51 dargestellt.

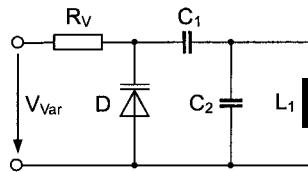


Abb. 2.51. Frequenzmodulation im Parallelschwingkreis durch eine Kapazitätsdiode

Die Kapazität C_1 und der Widerstand R_V bewirken eine Entkopplung der Spannungsquelle V_{var} von dem aus L_1 und C_2 bestehenden Parallelschwingkreis. Die von der niederfrequenten Spannung V_{var} abhängigen Resonanzfrequenz f_0 des Schwingkreises berechnet sich dann wie folgt:

$$f_0(V_{var}) = \frac{1}{2\pi \cdot \sqrt{L_1 \cdot (C_2 + C_S(V_{var}))}}. \quad (2.57)$$

In Gl. (2.59) bezeichnet C_S die Sperrsichtkapazität (vgl. Abschn. 2.5.1) der Kapazitätsdiode D , wobei C_S kleiner als die Entkoppelkapazität C_1 ist ($C_S \ll C_1$). Der Bereich, in dem die Frequenz des Schwingkreises moduliert werden kann, ist abhängig vom Verhältnis der minimalen zur maximalen Sperrsichtkapazität der Kapazitätsdiode.

2.8.5 Spannungsbegrenzung

Eine Spannungsbegrenzung mit Hilfe von Z-Dioden (vgl. Abschn. 2.7.2) wird mit der in Abb. 2.52a dargestellten Schaltung erreicht, wobei zwei Z-Dioden D_1 und D_2 antiparallel verschaltet sind.

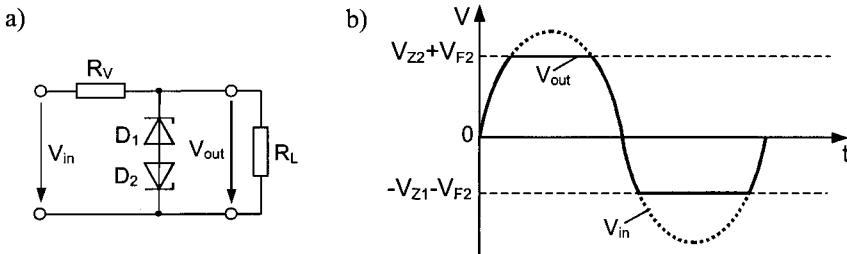


Abb. 2.52. a) Spannungsbegrenzungsschaltung mit Zenerdioden und b) beispielhafter Einangs- und Ausgangsspannungsverlauf für $D_1 \neq D_2$

Im dargestellten Beispiel wird eine Eingangswechselspannung V_{in} begrenzt, wobei die in Abb. 2.52b dargestellte Ausgangsspannung V_{out} über der ohmschen Last R_L resultiert. Unabhängig von der Polarität der Eingangsspannung V_{in} ist jeweils eine der Dioden in Flusspolung und die andere in Sperrpolung. Bei Überschreiten der jeweiligen Zenerspannung V_{Z1} bzw. V_{Z2} wird der Schaltungszweig

der antiseriell geschalteten Dioden niederohmig. Durch die damit verbundene Stromzunahme vergrößert sich der Spannungsabfall über dem Vorwiderstand R_V und begrenzt somit den Spannungsabfall über dem Lastwiderstand R_L in den Grenzen $(-V_{Z1} - V_{F1}) \leq V_{out} \leq (V_{Z2} + V_{F2})$, wobei V_{F1} und V_{F2} die jeweiligen Flussspannungen der Dioden D_1 und D_2 sind. Auf Grund der dabei umgesetzten Leistung muss bei der Dimensionierung der Schaltung die Verlustleistung der verwendeten Dioden berücksichtigt werden.

Das Schaltungsprinzip in Abb. 2.52a wird insbesondere als Überspannungsschutzschaltung (Schutz vor Spannungsspitzen) verwendet. Auf Grund der Temperaturabhängigkeit von Z-Dioden ist die Nutzung dieser Schaltung als Spannungsreferenz nur eingeschränkt geeignet.

3 Bipolartransistoren

Der sog. *Bipolartransistor* besteht aus zwei pn-Übergängen, die antiseriell auf einem Substrat angeordnet sind und ein gemeinsames Mittengebiet nutzen. Ausgehend vom Aufbau und der Funktionsweise von Bipolartransistoren werden in diesem Kapitel die Strom-Spannungs-Charakteristika, Durchbruchsmechanismen und technologische Realisierung dieser Bauelemente betrachtet. Neben dem Schaltverhalten und Netzwerkmodellen wird weiterhin der Einsatz von Bipolartransistoren in ausgewählten Schaltungsbeispielen beschrieben.

3.1 Aufbau und Funktionsweise des Bipolartransistors

Bipolartransistoren bestehen aus zwei benachbarten pn-Übergängen, die über ein gemeinsames Mittengebiet verfügen, das als Basis (B) bezeichnet wird. Die außenliegenden Gebiete werden als Emitter (E) und Kollektor (C) bezeichnet. Man unterscheidet beim Bipolartransistor Strukturen der Schichtfolge npn und pnp. Der schematische Aufbau eines npn- und pnp-Bipolartransistors und die zugehörigen Schaltsymbole sind in Abb. 3.1 dargestellt.

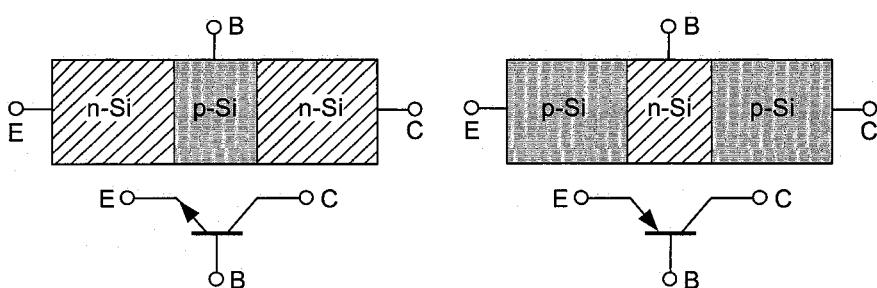


Abb. 3.1. Schematischer Aufbau und Schaltsymbol a) eines npn-Bipolartransistors und b) eines pnp-Bipolartransistors

Auf Grund der geringen Basisweite gegenüber der Diffusionslänge findet eine wechselseitige Beeinflussung der pn-Übergänge statt. Die Basis dient als Steuerelektrode, über die der sog. *Transferstrom* zwischen Emitter und Kollektor gesteuert werden kann. In den weiteren Betrachtungen wird, soweit nicht anders

vermerkt, von einem npn-Bipolartransistor ausgegangen. In Abb. 3.2 ist der beispielhafte Querschnitt eines npn-Bipolartransistors dargestellt, der in Silizium-Planartechnik (vgl. Abschn. 3.4) hergestellt ist.

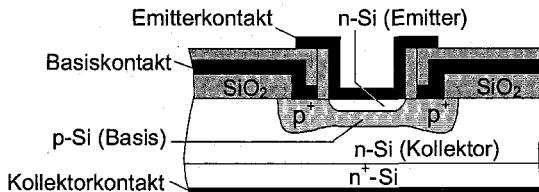


Abb. 3.2. Beispielhafter Querschnitt eines npn-Bipolartransistors in Silizium-Planartechnik

3.1.2 Bipolartransistor im thermodynamischen Gleichgewicht

Der Bipolartransistor befindet sich im thermodynamischen Gleichgewicht, wenn keine externe Spannung anliegt (Basis-Emitter-Spannung $V_{BE} = 0$, Kollektor-Basis-Spannung $V_{CB} = 0$). Dann bilden sich die in Abb. 3.3a dargestellten Raumladungszonen ℓ_{BE} und ℓ_{CB} zwischen Emitter und Basis sowie zwischen Kollektor und Basis aus ($\ell_{BE} = x_{p1} - x_{n1}$ und $\ell_{CB} = x_{n2} - x_{p2}$). Die Ausdehnung der Raumladungszonen ist vom Dotierungsprofil abhängig, welches in Abb. 3.3b beispielhaft dargestellt ist.

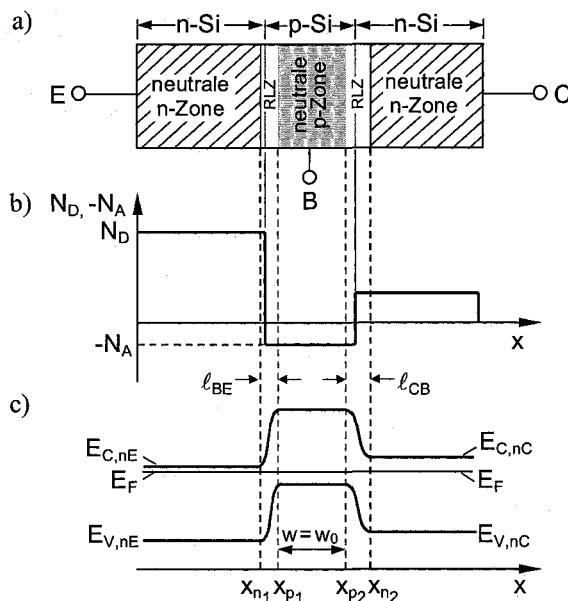


Abb. 3.3. a) Raumladungszonen im npn-Bipolartransistor bei thermodynamischem Gleichgewicht ($V_{BE} = V_{CB} = 0$) b) beispielhaftes Dotierungsprofil und c) Bänderschema

Die sich im thermodynamischen Gleichgewicht einstellenden Bandverläufe sind im Bänderdiagramm Abb. 3.3c dargestellt. Dabei liegt das Ferminiveau E_F des Emitter-, Basis- und Kollektorgebietes jeweils auf dem gleichen Energieniveau. Es findet kein Stromfluss statt. Im Bereich $x_{p1} \leq x \leq x_{p2}$ ist die Basis neutral (bei homogen dotierter Basiszone), d.h. die Raumladungsdichte besitzt den Wert $\rho(x) = 0$. Die Länge dieses Bereiches wird neutrale Basisweite w genannt. Die neutrale Basisweite im thermodynamischen Gleichgewicht ($V_{BE} = V_{CB} = 0$) wird mit w_0 bezeichnet.

3.1.3 Bipolartransistor in Vorwärtsbetrieb

Befindet sich der pn-Übergang zwischen Emitter und Basis in Flusspolung (Basis-Emitter-Spannung $V_{BE} > 0$), so werden vom Emitter Majoritätsträger (Elektronen) in die Basis injiziert, die dann vom pn-Übergang wegdiffundieren. Aufgrund der geringen Weite der Basis diffundieren diese injizierten Ladungsträger zum in Sperrpolung betriebenen pn-Übergang zwischen Basis und Kollektor (Kollektor-Basis-Spannung $V_{CB} < 0$). Hier werden sie als Teil des Sperrstroms zum Kollektor hin abgesaugt. Dieser Transferstrom I_T , der näherungsweise dem Kollektorstrom I_C entspricht, wird sehr stark von der Basis-Emitter-Spannung V_{BE} beeinflusst. Die Abhängigkeit dieses Stroms von der Kollektor-Basis-Spannung V_{CB} ist dagegen gering. Das Ansteuerungsprinzip des Bipolartransistors und das Bänderdiagramm für diesen sog. *Vorwärtsbetrieb* ($V_{BE} > 0$, $V_{CB} < 0$) ist in Abb. 3.4 dargestellt.

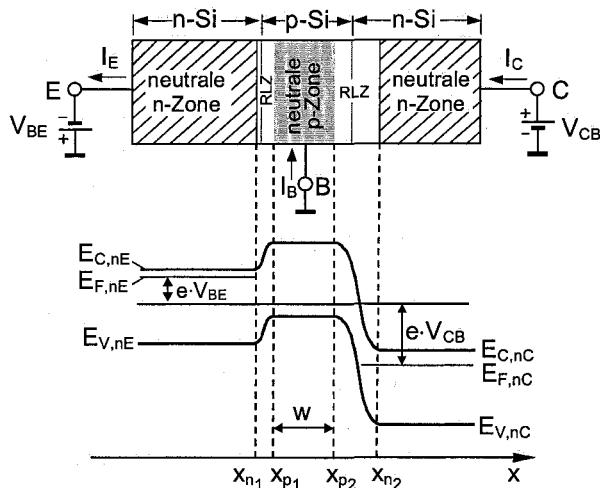


Abb. 3.4. Ansteuerungsprinzip und Bänderschema eines npn-Bipolartransistors in Vorwärtsbetrieb ($V_{BE} > 0$, $V_{CB} < 0$)

Die Minoritätsträgerverteilung eines Bipolartransistors im Vorwärtsbetrieb ist in Abb. 3.5 verdeutlicht. Dabei werden die Ladungsträgerverteilungen für die

neutrale Basis und die neutralen Zonen des Emitter- und Kollektorgebietes betrachtet, aus denen sich nachfolgend die Ströme des Bipolartransistors ermitteln lassen.

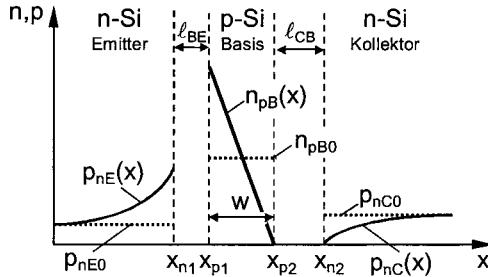


Abb. 3.5. Minoritätsträgerverteilung eines npn-Bipolartransistors in Vorwärtsbetrieb ($V_{BE} > 0, V_{CB} < 0$)

Die Minoritätsträgerdichten am emitterseitigen Rand der neutralen Basis werden durch das dort vorherrschende Potential bestimmt und sind somit von der Basis-Emitter-Spannung V_{BE} abhängig. Für die Minoritätsträgerdichte n_{pB} in der neutralen Basis am emitterseitigen Rand $x = x_{p1}$ und für die Minoritätsträgerdichte p_{nE} im Emittergebiet am basisseitigen Rand $x = x_{n1}$ gilt entsprechend Gl. 2.14 bzw. Gl. 2.15:

$$n_{pB}\Big|_{x=x_{p1}} = n_{pB0} \cdot \exp\left(\frac{V_{BE}}{V_\vartheta}\right) \quad \text{und} \quad p_{nE}\Big|_{x=x_{n1}} = n_{nE0} \cdot \exp\left(\frac{V_{BE}}{V_\vartheta}\right). \quad (3.1)$$

Am kollektorseitigen Rand der neutralen Basis ist die Minoritätsträgerdichte n_{pB} sehr gering, da der pn-Übergang zwischen Basis und Kollektor in Sperrrichtung gepolt ist ($V_{CB} < 0$). Die Minoritätsträgerdichte n_{pB} in der neutralen Basis am kollektorseitigen Rand $x = x_{p2}$ und die Minoritätsträgerdichte p_{nC} im Kollektorgebiet am basisseitigen Rand $x = x_{n2}$ sind jeweils von der Kollektor-Basis-Spannung V_{CB} abhängig und werden wie folgt bestimmt:

$$n_{pB}\Big|_{x=x_{p2}} = n_{p0} \cdot \exp\left(\frac{V_{CB}}{V_\vartheta}\right) \quad \text{und} \quad p_{nC}\Big|_{x=x_{n2}} = n_{n0} \cdot \exp\left(\frac{V_{CB}}{V_\vartheta}\right). \quad (3.2)$$

Da der pn-Übergang zwischen Basis und Kollektor gesperrt ist, sind beide Werte sehr gering und in der linearen Darstellung in Abb. 3.5 näherungsweise Null.

Innerhalb der neutralen Basis existiert gemäß Gl. (3.1) und Gl. (3.2) ein starkes Minoritätsträgerkonzentrationsgefälle. Auf Grund der erheblich größeren Diffusionslänge der Elektronen gegenüber der neutralen Basisweite ($L_n \gg w$) ergibt sich für den in Abb. 3.5 dargestellten Verlauf der Minoritätsträgerdichte $n_{pB}(w)$ innerhalb der neutralen Basis näherungsweise die folgende Beziehung:

$$n_{pB}(w) \approx n_{pB}(x = x_{p1}) - \frac{(x - x_{p1})}{w} \cdot n_{pB}(x = x_{p1}). \quad (3.3)$$

Dieses Minoritätsträgerkonzentrationsgefälle in der neutralen Basis stellt die Ursache für die Ladungsträgerdiffusion vom Emittergebiet in das Kollektorgebiet dar. Der dadurch resultierende Diffusionsstrom bzw. Transferstrom I_T , der näherungsweise dem Kollektorstrom I_C entspricht, berechnet sich unter Berücksichtigung der Emitterfläche A und der Diffusionskonstante D_n gemäß Gl. (3.4). Dabei wird von einer konstanten Dotierungsdichte in den Bahngebieten ausgegangen und die pn-Übergänge werden als abrupt angesehen.

$$I_T \approx I_C \approx \frac{e \cdot A \cdot D_n \cdot n_{pB} \Big|_{x=x_{pI}}}{w} = \frac{e \cdot A \cdot D_n \cdot n_i^2}{w \cdot N_A} \cdot \exp\left(\frac{V_{BE}}{V_\vartheta}\right) \quad (3.4)$$

Wegen der exponentiellen Abhängigkeit genügt eine kleine Änderung der Basis-Emitter-Spannung V_{BE} , um einen großen Kollektorstrom I_C zu steuern oder zu schalten. Sowohl die Rekombination in der Basis als auch die Generation und Rekombination in den Raumladungszonen wurden in Gl. (3.1) - (3.4) vernachlässigt. Weiterhin wurden die Bahnwiderstände nicht berücksichtigt. Die Näherung eines konstanten Gradienten der Minoritätsträgerkonzentration in Gl. (3.3) gilt für den Fall, dass der Basisstrom gegenüber dem Transferstrom I_T vernachlässigbar ist.

Die vom Emitter zum Kollektor diffundierenden Ladungsträger haben eine begrenzte Aufenthaltsdauer in der Basiszone. Die mittlere Laufzeit wird *Basistransitzzeit* τ_B genannt. Sie ergibt sich aus der neutralen Basisweite w_0 und der Diffusionskonstanten D_n wie folgt:

$$\tau_B = \frac{w_0^2}{2 \cdot D_n}. \quad (3.5)$$

Für eine sichere Funktionsweise des Bipolartransistors muss die Basistransitzzeit sehr viel geringer sein als die Minoritätsträgerlebensdauer in der Basis ($\tau_B \ll \tau_n$).

Neben der Ladungsträgerdiffusion vom Emitter zum Kollektor treten weitere Mechanismen auf, die am Ladungsträgertransport des Bipolartransistors beteiligt sind. Einige Ladungsträger rekombinieren in der Basis mit Majoritätsträgern, was zu einem Rekombinationsverlust führt und durch einen zusätzlichen Basisstromanteil I_{BB} ausgeglichen werden muss. Weiterhin werden Majoritätsträger aus der Basis in den Emitter injiziert, woraus ein Stromanteil I_{BE} resultiert. In Abb. 3.6 sind die Stromanteile, die am Ladungsträgertransport des Bipolartransistors beteiligt sind dargestellt.

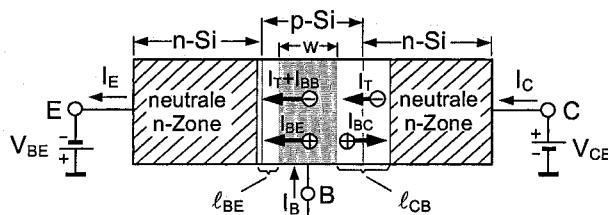


Abb. 3.6. npn-Bipolartransistor mit Stromkomponenten, die neben dem Transferstrom I_T am Ladungsträgertransport beteiligt sind

Als weiterer Effekt tritt auch bei gesperrtem pn-Übergang zwischen Kollektor und Basis ein Sperrstrom I_{BC} auf. Dadurch steigt der Kollektorstrom I_C an, ohne dass dies vom Emitter beeinflusst werden kann.

Der Basisstrom I_B des Bipolartransistors setzt sich aus den beschriebenen Stromanteilen wie folgt zusammen:

$$I_B = I_{BE} + I_{BB} + I_{BC}. \quad (3.6)$$

Im Vorwärtsbetrieb des Bipolartransistors ist der Stromanteil I_{BC} zumeist sehr gering und bei hinreichend geringer Basisweite überwiegt meist der Stromanteil I_{BE} gegenüber I_{BB} . Der Basisstrom I_B ist demnach nur noch von dem zum Emitter fließenden Löcherstrom I_{BE} abhängig ($I_B \approx I_{BE}$). Der Strom I_{BE} ist der Vorwärtstrom des pn-Überganges zwischen Emitter und Basis, der jedoch geringer als der Vorwärtstrom eines herkömmlichen pn-Überganges ist, was durch den Abfluss des Transferstroms I_T zum Kollektor begründet ist. Der Strom I_{BE} ist in gleicher Weise wie der Kollektorstrom I_C von der Basis-Emitter-Spannung V_{BE} abhängig ($I_{BE} \sim \exp(V_{BE}/V_\theta)$).

3.1.4 Beschaltungsarten

Der Bipolartransistor kann in drei wesentlichen Beschaltungsarten betrieben werden. Die bedeutendste Beschaltungsart ist die sog. *Emitterschaltung* (Abb. 3.7a). Neben der Emitterschaltung existieren weiterhin die *Basisschaltung* (Abb. 3.7b) und die *Kollektorschaltung* (Abb. 3.7c). Wird der Bipolartransistor beschaltet, dann nutzen der Eingangs- und Ausgangskreis der Schaltung jeweils einen Anschluss des Transistors gemeinsam. Dieser Anschluss wird während des Steuervorganges auf konstantem Potential gehalten und gibt der zugehörigen Schaltung den Namen.

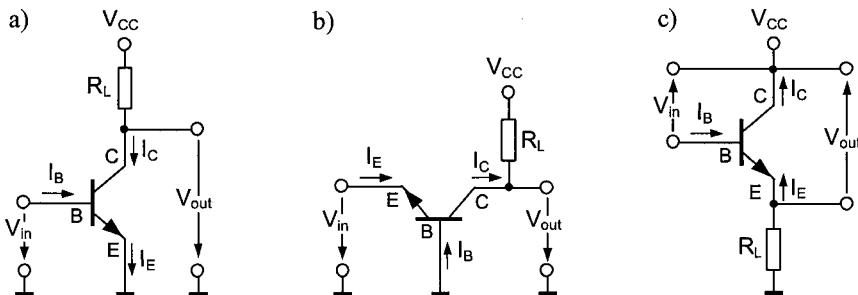


Abb. 3.7. Beschaltungsarten des Bipolartransistors a) Emitterschaltung b) Basisschaltung c) Kollektorschaltung

Bei der Emitterschaltung (Abb. 3.7a) liegt der Emitter auf konstantem Potential (hier Massepotential) und der Kollektorstrom I_C wird über den Basisstrom I_B gesteuert. Die Emitterschaltung zeichnet sich durch eine sehr hohe Spannungsver-

stärkung A_{VE} aus, die über das Verhältnis von Ausgangsspannungsänderung ΔV_{out} und Eingangsspannungsänderung ΔV_{in} bestimmt wird. Aus Gl. (3.4) folgt:

$$A_{VE} = \frac{\Delta V_{out}}{\Delta V_{in}} = -\frac{I_C \cdot R_L}{V_\theta}. \quad (3.7)$$

Die Spannungsverstärkung ist somit proportional zum Lastwiderstand R_L . Die Stromverstärkung A_{IE} einer Emitterschaltung ist ebenfalls sehr hoch und berechnet sich wie folgt:

$$A_{IE} = \frac{I_E}{I_B} \approx \frac{I_C}{I_B}. \quad (3.8)$$

Bei der Basisschaltung (Abb. 3.7b) liegt die Basis des Bipolartransistors auf konstantem Potential (z. B. Massepotential) und der Kollektorstrom I_C wird durch den eingeprägten Emitterstrom I_E bestimmt. Somit findet keine Stromverstärkung A_{IB} statt und es gilt:

$$A_{IB} = \frac{I_C}{I_E} = \frac{A_{IE} - 1}{A_{IE}} \approx 1. \quad (3.9)$$

Die Spannungsverstärkung A_{VB} eines Bipolartransistors in Basisschaltung ist dagegen ebenso hoch wie bei der Emitterschaltung und kann über das Verhältnis von Ausgangsspannungsänderung ΔV_{out} und Eingangsspannungsänderung ΔV_{in} ermittelt werden:

$$A_{VB} = \frac{\Delta V_{out}}{\Delta V_{in}} = -\frac{I_C \cdot R_L}{V_\theta}. \quad (3.10)$$

Bei der Kollektorschaltung (Abb. 3.7c), die auch Emitterfolger genannt wird, liegt der Kollektor auf Betriebsspannungspotential V_{CC} und der Emitterstrom I_E wird über den Basisstrom I_B gesteuert. Die Stromverstärkung A_{IC} einer Kollektorschaltung entspricht daher näherungsweise der Stromverstärkung in Emitterschaltung:

$$A_{IC} = \frac{I_E}{I_B} \approx A_{IE}. \quad (3.11)$$

Die Basis-Emitter-Spannung ist im Vorwärtsbetrieb etwa gleich der Schleusenspannung ($V_{BE} \approx 0,65V$). Das Emitterpotential folgt daher der Eingangsspannung (Emitterfolger) und es tritt näherungsweise keine Spannungsverstärkung auf:

$$A_{VC} = \frac{\Delta V_{out}}{\Delta V_{in}} \approx 1. \quad (3.12)$$

Die dargestellten Schaltungsarten des Bipolartransistors dienen im folgenden der Bestimmung weiterer Transistormerkmale, die schaltungsabhängig sind. Die Emitterschaltung besitzt auf Grund ihrer hohen Spannungs- und Stromverstärkung schaltungstechnisch die größte Bedeutung. Ausgewählte Grundschaltungen mit Bipolartransistoren werden in Abschn. 3.7 behandelt.

3.2 Strom-Spannungs-Charakteristik des Bipolartransistors

3.2.1 Transistorkennlinien

Das *Ausgangskennlinienfeld* (Abb. 3.8a) des Bipolartransistors in Emitterschaltung kennzeichnet die Abhängigkeit des Kollektorstroms I_C von der Kollektor-Emitter-Spannung V_{CE} und macht die Arbeitsbereiche des Bipolartransistors mit V_{BE} als Parameter sichtbar. Die Arbeitsgerade im Ausgangskennlinienfeld wird durch die Betriebsspannung V_{CC} und den Kollektorstrom I_C festgelegt, der durch das Verhältnis von Betriebsspannung zum Lastwiderstand R_L bestimmt wird. Die *Eingangskennlinie* des Bipolartransistors (Abb. 3.8b) charakterisiert den Zusammenhang zwischen Basisstrom I_B und Basis-Emitter-Spannung V_{BE} . Sie macht sichtbar, dass eine geringe Änderung der Basis-Emitterspannung ΔV_{BE} eine relativ große Änderung des Basisstroms ΔI_B und damit eine große Änderung des Kollektorstroms ΔI_C in der Ausgangskennlinie (Abb. 3.8a) bewirkt.

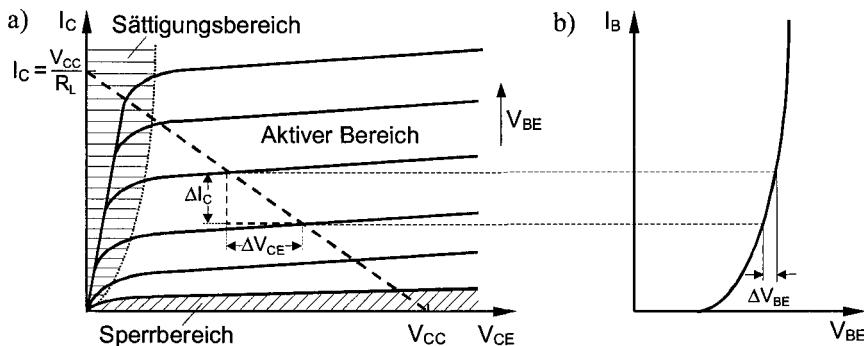


Abb. 3.8. a) Ausgangskennlinienfeld und **b)** Eingangskennlinie eines npn-Bipolartransistors in Emitterschaltung

Im *Sperrbereich* des Ausgangskennlinienfeldes des Bipolartransistors sind beide pn-Übergänge gesperrt, wodurch beidseitig keine Injektion von Ladungsträgern erfolgen kann; der Bipolartransistor im Schalterbetrieb ist im geöffneten Zustand.

Im *aktiven Bereich* (Vorwärtsbetrieb) befindet sich der pn-Übergang zwischen Basis und Emitter in Flusspolung und der pn-Übergang zwischen Basis und Emitter in Sperrpolung. Dadurch wird der Bipolartransistor im normalen Verstärkungsbereich betrieben. Unter der Bedingung $V_{BE} >> V_\vartheta$ und Vernachlässigung der Bahnwiderstände erhält man für den Verlauf des Kollektorstroms I_C die Beziehung:

$$I_C = I_{ES} \cdot \exp\left(\frac{V_{BE}}{V_\vartheta}\right) \quad \text{mit} \quad I_{ES} = \frac{e \cdot A \cdot D_n \cdot n_{pB0}}{w_0}. \quad (3.13)$$

In Gl. (3.13) bezeichnet I_{ES} den Transfersättigungsstrom, der von der Diffusionskonstanten D_n der Elektronen, der Emitterfläche A und der neutralen Basisweite w_0 bei thermodynamischem Gleichgewicht (vgl. Abschn. 3.1.2) abhängt. Der Grund für die Zunahme des Kollektorstroms I_C bei steigender Kollektor-Emitter-Spannung V_{CE} liegt in der Erhöhung des Gradienten der Minoritätsträgerdichte in der Basis des Bipolartransistors. Dieser Effekt wird in Abschn. 3.2.2 näher beschrieben.

Im *Sättigungsbereich* befinden sich beide pn-Übergänge des Bipolartransistors in Flusspolung, wodurch ein unvollständiges Absaugen bzw. eine Rückinjektion von Ladungsträgern stattfindet; der Bipolartransistor im Schalterbetrieb befindet sich im geschlossenen Zustand. Dabei verbleibt eine Restspannung am Bipolartransistor, die als Kollektor-Emitter-Sättigungsspannung V_{CESat} bezeichnet wird. Sie berechnet sich beim idealen Transistor aus dem Spannungsabfall über den an tisierell geschalteten, in Flusspolung betriebenen Dioden zwischen Basis und Emitter sowie zwischen Basis und Kollektor ($V_{CESat} = V_{BE} - V_{CB}$). Beim realen Transistor besitzen die Bahnwiderstände R_C und R_E des Kollektorbereiches und Emitterbereiches (vgl. Abschn. 3.5.2) einen erheblichen Einfluss auf den Kollektorstrom I_C , sodass sich für die Kollektor-Emitter-Sättigungsspannung V_{CESat} des Bipolartransistors folgender Zusammenhang ergibt:

$$V_{CESat} = (V_{BE} - V_{CB}) + I_C \cdot R_C + I_E \cdot R_E. \quad (3.14)$$

Der in der Eingangskennlinie Abb. 3.8b dargestellte Basisstrom I_B bezeichnet die Differenz zwischen Kollektorstrom I_C und Emitterstrom I_E . Für die Kennliniengleichung gilt unter Berücksichtigung des Sättigungsstroms I_{ES} und der Stromverstärkung A_{IE} in Emitterschaltung (vgl. Abschn. 3.1.4) für $V_{BE} \gg V_\vartheta$ folgender genauerer Zusammenhang [3.6]:

$$I_B = I_C - I_E \approx \frac{I_{ES}}{A_{IE}} \cdot \exp\left(\frac{V_{BE}}{V_\vartheta}\right). \quad (3.15)$$

3.2.2 Basisweitenmodulation

Eine Erhöhung der Kollektor-Basis-Spannung V_{CE} bewirkt eine Ausdehnung der Raumladungszone des pn-Überganges zwischen Kollektor und Basis, wodurch sich die neutrale Basisweite w des Bipolartransistors verringert. Bei konstanter Minoritätsträgerdichte n_{pB} (x_{pl}) am emitterseitigen Rand der Raumladungszone nimmt der Gradient der Minoritätsträgerdichte n_{pB} in der Basis zu, was zu einem Anstieg des Kollektorstroms I_C führt. Dieser Effekt wird *Basisweitenmodulation* oder *Early-Effekt* genannt. Die Minoritätsträgerverteilung $n_{pB}(w)$ in der neutralen Basis für zwei unterschiedliche Kollektor-Emitter-Spannungen $V_{CE} > V'_CE$ ist in Abb. 3.9a dargestellt. Der Effekt der Basisweitenmodulation wird in den Ausgangskennlinien des Bipolartransistors durch die nicht horizontal verlaufenden Kennlinien im aktiven Bereich sichtbar. Ein beispielhafter Verlauf der Ausgangskennlinie eines npn-Bipolartransistors in Emitterschaltung bei konstantem Basisstrom I_B mit und ohne Basisweitenmodulation ist in Abb. 3.9b dargestellt.

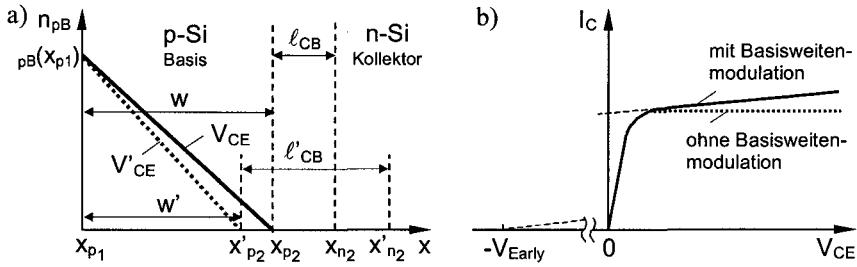


Abb. 3.9. a) Minoritätsträgerverteilung in der neutralen Basis für Kollektor-Emitter-Spannungen $V_{CE} > V'_CE$ und b) Ausgangskennlinie eines npn-Bipolartransistors in Emitterschaltung bei konstantem Basisstrom I_B mit und ohne Basisweitenmodulation, Extrapolation der Early-Spannung V_{Early}

Die Berechnung der von der Kollektor-Emitter-Spannung $V_{CE} > 0$ abhängigen neutralen Basisweite w unter Berücksichtigung der neutralen Basisweite w_0 bei thermodynamischem Gleichgewicht erfolgt näherungsweise über die Beziehung:

$$w \approx w_0 - \sqrt{\frac{2\epsilon_0 \cdot \epsilon_r \cdot |V_D - V_{CE}|}{e \cdot N_A}}. \quad (3.16)$$

Als Parameter für die Basisweitenmodulation dient die sog. *Early-Spannung* V_{Early} , die durch den Abszissenschnittpunkt bei $V_{CE} = -V_{Early}$ der zu $I_C = 0$ extrapolierten Ausgangskennlinien näherungsweise bestimmt wird (vgl. Abb. 3.9b). Durch die Wahl der Dotierungskonzentrationen $N_C \ll N_B$ kann die Basisweitenmodulation gering gehalten werden.

3.3 Durchbruchsmechanismen

Entsprechend der Durchbruchsmechanismen von pn-Dioden (vgl. Abschn. 2.3) treten auch am Bipolartransistor Durchbrüche auf, die einen starken Stromanstieg bewirken. Im Folgenden wird neben dem Lawinendurchbruch der sog. *Durchbruch zweiter Art* und der sog. *Punch-Through-Effekt* beschrieben.

3.3.1 Lawinendurchbruch

Analog zur Wirkungsweise des Lawineneffektes an pn-Dioden (vgl. Abschn. 2.2.1) wirkt am Bipolartransistor der reversible Lawinendurchbruch an der Basis-Kollektor-Diode des Bipolartransistors. Zu unterscheiden ist der Betrieb bei offenem Emitter ($I_E = 0$) oder in Emitterschaltung bei offener Basis ($I_B = 0$). Der erste Fall entspricht dem Lawinendurchbruch der pn-Diode. Im zweiten Fall fließt der Sperrstrom durch die Basis-Emitter-Diode. Dabei gelangen Elektronen aus dem Emitter in die Basis und der Sperrstrom wird durch die anwachsende Minoritätsträgerdichte erhöht. Damit tritt auch der Lawinendurchbruch in diesem Fall früher

auf. Erfolgt schaltungstechnisch keine Strombegrenzung, so kann aus dem Lawinendurchbruch eine Zerstörung des Bipolartransistors resultieren.

3.3.2 Durchbruch zweiter Art (thermischer Durchbruch)

Ebenso wie in der pn-Diode kann durch die mit dem Sperrstrom verbundene Erwärmung ein thermischer Durchbruch des Bipolartransistors erfolgen. Besonders bei offener Basis und Vorwärtsbetrieb wird dieser Effekt durch das Auftreten von inhomogenen Temperaturverteilungen begünstigt. Dabei konzentrieren sich hohe Stromdichten auf einen sehr geringen Kristallbereich, wobei die Sperrsichttemperatur stark ansteigt (Hot Spot) und die Kollektor-Emitter-Spannung V_{CE} lokal zusammenbricht. Findet keine Strombegrenzung durch die äußere Beschaltung des Transistors statt, so steigt der Strom stark an und der Bipolartransistor wird zerstört. Dieser irreversible thermische Durchbruch wird als Durchbruch zweiter Art bezeichnet. Er begrenzt die maximal zulässige Verlustleistung P_{max} des Bipolartransistors im aktiven Bereich, wie beispielhaft im Ausgangskennlinienfeld eines npn-Bipolartransistors in Emitterschaltung in Abb. 3.10 dargestellt ist.

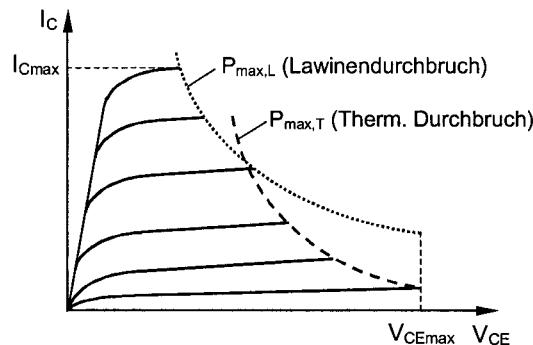


Abb. 3.10. Ausgangskennlinienfeld eines npn-Bipolartransistors mit Kennzeichnung der maximalen Verlustleistung bei Lawinendurchbruch $P_{max,L}$ und Reduktion der maximalen Verlustleistung durch thermischen Durchbruch $P_{max,T}$

Die durch eine lokale Leistungsdichte-Zunahme bedingte Sperrsichttemperaturerhöhung muss durch Wärmeleitung zu einer Wärmesenke ausgeglichen werden. Je größer der spezifische Widerstand der Basis ist, um so wahrscheinlicher ist das Auftreten eines Durchbruchs zweiter Art.

3.3.3 Punch-Through-Effekt

Der Punch-Through-Effekt bezeichnet die Berührung der Raumladungszonen zwischen Basis und Kollektor sowie zwischen Basis und Emitter bei geringer metallurgischer Basisweite. Dadurch werden Emitter und Kollektor kurzgeschlossen und der Bipolartransistor kann nicht mehr über die Basis gesteuert werden. Durch thermische Emission von Ladungsträgern [3.3] kommt es zu einem starken An-

stieg des Kollektorstroms I_C . Durch die Wahl einer hinreichend großen Basisweite des Bipolartransistors kann erreicht werden, dass der Lawineneffekt zeitlich vor dem Punch-Through-Effekt einsetzt.

3.4 Technologische Realisierung

Die Prozessschritte zur technologischen Realisierung eines npn-Bipolartransistors in Silizium-Planartechnik mit sog. *selbstjustierender Emitter-Strukturierung* sind in Abb. 3.11 veranschaulicht.

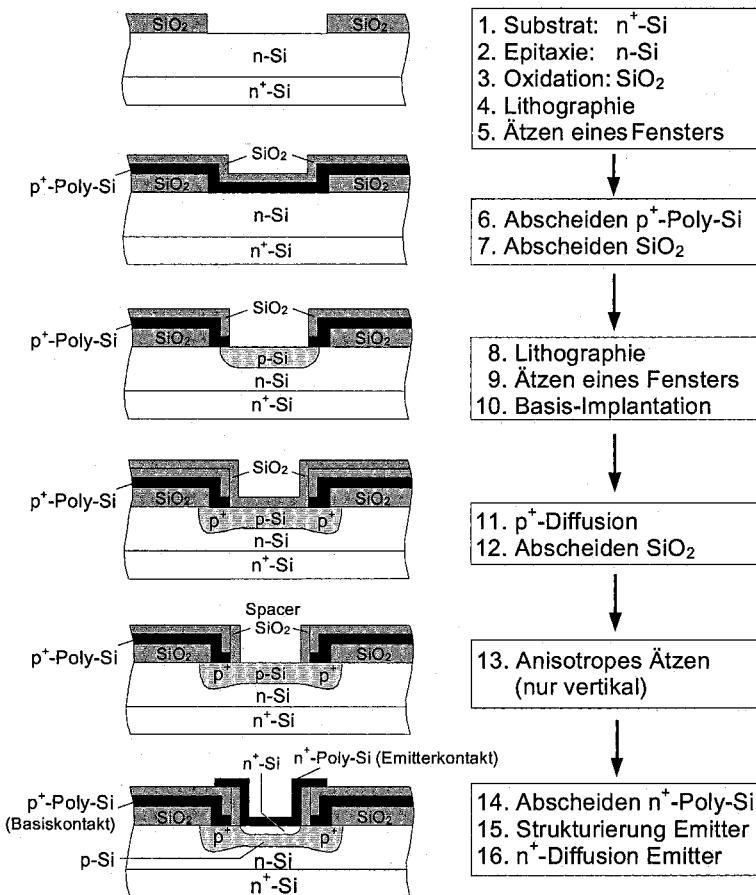


Abb. 3.11. Typische Technologieschritte zur Herstellung eines npn-Bipolartransistors (Beispiel für Silizium-Planarprozess mit selbstjustierender Emitter-Strukturierung)

Die Herstellung eines npn-Bipolartransistors erfolgt beispielsweise ausgehend von einem n^+ -Silizium-Substrat, auf welches durch Epitaxie eine n-Si-Schicht auf-

gewachsen wird. Durch anschließende thermische Oxidation wird eine SiO_2 -Isolatorschicht erzeugt. Im nächsten Prozessschritt (Lithographie) wird Fotolack aufgetragen und durch eine Maske mit UV-Licht bestrahlt (vgl. Abschn. 2.4). Durch das bestrahlte Gebiet wird ein Fenster in die Isolatorschicht geätzt und der restliche Fotolack entfernt. Anschließend wird ganzflächig eine Doppelschicht aus p^+ -Poly-Silizium (Basiskontaktgebiet) und Oxid (SiO_2) abgeschieden. Nach dem folgenden Lithographieschritt wird ein Fenster in diese Doppelschicht geätzt, durch welches die Basis-Implantation (p -Si) vorgenommen wird. Nach der Ausdiffusion der Basiskontaktgebiete (p^+ -Si) wird an der Oberfläche der Struktur SiO_2 abgeschieden. Durch vertikales anisotropes Ätzen werden sog. *Spacer* erzeugt, die eine Isolation des Basiskontaktgebietes (p^+ -Poly-Si) vom anschließend erzeugten n^+ -Emitter-Diffusionsgebiet und des Emitter-Kontaktgebietes ermöglichen.

3.5 Netzwerkmodelle des Bipolartransistors

3.5.1 Sperrsicht- und Diffusionskapazität

Die spannungsabhängigen *Sperrsichtkapazitäten* zwischen Basis und Emitter sowie zwischen Kollektor und Basis entstehen durch Abtransport und Zuliefern von Majoritätsträgern am Rande der neutralen Zonen bei Änderung der Raumladungszonenweite. Die Sperrsichtkapazitäten $C_{S_{BE}}$ und $C_{S_{CB}}$ berechnen sich über die zugehörigen Raumladungszonen ℓ_{BE} und ℓ_{CB} (vgl. Abschn. 2.1.2.1) zu:

$$C_{S_{BE}} = \frac{A \cdot \epsilon_r \epsilon_0}{\ell_{BE}} \quad \text{und} \quad C_{S_{CB}} = \frac{A \cdot \epsilon_r \epsilon_0}{\ell_{CB}}. \quad (3.17)$$

Weiterhin tritt eine spannungsabhängige Sperrsichtkapazität C_{SC} am pn-Übergang zwischen Kollektor und Substrat des Bipolartransistors auf.

Durch Speicherung von Diffusionsladungen in der Basis des Bipolartransistors entsteht die Diffusionskapazität (Basiskapazität) C_B . Die Speicherung der Diffusionsladung bei Änderung der Basis-Emitter-Spannung V_{BE} um ΔV_{BE} ist in Abb. 3.12 veranschaulicht.

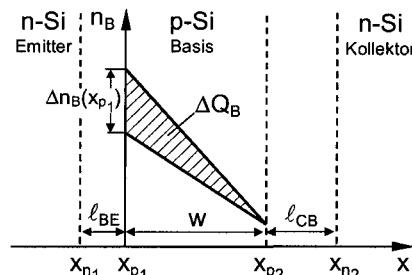


Abb. 3.12. Speicherung von Diffusionsladungen in der Basis eines npn-Bipolartransistors (Entstehung der Diffusionskapazität)

Die sog. *Basisladung* Q_B des Bipolartransistors berechnet sich in Abhängigkeit von der Minoritätsträgerdichte n_{B0} in der Basis bei thermodynamischem Gleichgewicht, der Basis-Emitter-Spannung V_{BE} , der Emitterfläche A und der neutralen Basisweite w wie folgt:

$$|Q_B| = \frac{1}{2} \cdot e \cdot A \cdot n_{B0} \cdot \exp\left(\frac{V_{BE}}{V_\vartheta}\right) \cdot w. \quad (3.18)$$

Aus Gl. (3.18) folgt für die Basiskapazität C_B durch partielle Ableitung der Basisladung Q_B nach der Basis-Emitter-Spannung V_{BE} bei konstanter Kollektor-Emitter-Spannung V_{CE} die nachfolgende Beziehung:

$$C_B = \left| \frac{\partial Q_B}{\partial V_{BE}} \right| = \frac{1}{2} \cdot \frac{e \cdot A \cdot n_{B0}}{V_\vartheta} \cdot \exp\left(\frac{V_{BE}}{V_\vartheta}\right) \cdot w = \frac{I_C}{2 \cdot V_\vartheta} \cdot \frac{w^2}{D_{nB}}. \quad (3.19)$$

Gemäß Gl. (3.19) kann die Basiskapazität C_B somit über den Kollektorstrom I_C , die Temperaturspannung V_ϑ , die neutrale Basisweite w und die Diffusionskonstante D_{nB} der Elektronen in der Basis berechnet werden.

Die Diffusionsladung der Basiskapazität C_B teilt sich in einen Emitter- und einen Kollektorladungsanteil auf. Der Ladungsanteil im Emitter entsteht durch Löcherinjektion bei $V_{BE} > 0$ und der Ladungsanteil im Kollektor wird durch Löcherinjektion bei $V_{CB} > 0$ erzeugt [3.3]. Die resultierenden Diffusionskapazitätsanteile $C_{B_{BE}}$ und $C_{B_{CB}}$ sind mit der zugehörigen Basis-Emitter-Diode bzw. Basis-Kollektor-Diode verknüpft.

3.5.2 Großsignalersatzschaltbild

Das Großsignalverhalten des Bipolartransistors kann durch das Großsignalersatzschaltbild in Abb. 3.13 beschrieben werden. Es enthält eine gesteuerte Stromquelle, eine Basis-Kollektor- und Basis-Emitter-Diode sowie Diffusions- und Sperrschichtkapazitäten.

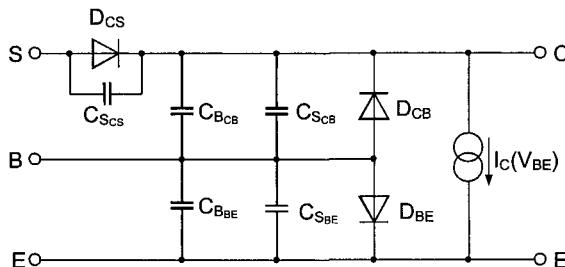


Abb. 3.13. Großsignalersatzschaltbild eines npn-Bipolartransistors

Die spannungsgesteuerte Stromquelle $I_C(V_{BE})$ beschreibt den Steuereffekt des Transfer- bzw. Kollektorstroms in Abhängigkeit von der Basis-Emitter-Spannung.

Die Basis-Emitter-Diode D_{BE} und Basis-Kollektor-Diode D_{BC} kennzeichnen das Durchlass- und Sperrverhalten der pn-Übergänge des Bipolartransistors und beinhalten damit das nichtideale Verhalten, das durch Leckströme entsteht. Parallel zur Basis-Emitter- und Basis-Kollektor-Diode liegen jeweils die zugehörigen Sperrsichtkapazitäten $C_{S_{BE}}$ und $C_{S_{CB}}$ sowie die Diffusionskapazitätsanteile $C_{B_{BE}}$ und $C_{B_{CB}}$ (vgl. Abschn. 3.5.1).

Die Diode D_{CS} beschreibt das Verhalten des pn-Überganges zwischen Kollektor und Substrat. Parallel zur Diode D_{CS} liegt die zugehörige Sperrsichtkapazität $C_{S_{CS}}$.

3.5.3 Kleinsignalersatzschaltbild

Das Kleinsignalverhalten des Bipolartransistors beschreibt analog zur Kleinsignalbetrachtung der pn-Diode (vgl. Abschn. 2.5.3) den Zusammenhang zwischen kleinen Spannungsänderungen ΔV_{BE} und ΔV_{CB} und den resultierenden Stromänderungen ΔI_C und ΔI_E . Dabei wird die Steuerwirkung der Basis-Emitter-Spannung durch die sog. Steilheit g_m und der Effekt der Basisweitenmodulation durch den Ausgangsleitwert g_o beschrieben. Das Kleinsignalersatzschaltbild eines Bipolartransistors (ohne Berücksichtigung des Bahnwiderstands- und Substrateinflusses) ist in Abb. 3.14 dargestellt.

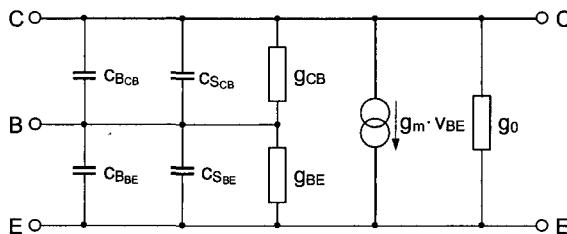


Abb. 3.14. Kleinsignalersatzschaltbild eines npn-Bipolartransistors (ohne Berücksichtigung des Bahnwiderstands- und Substrateinflusses)

Die in Abschn. 3.5.2 im Großsignalersatzschaltbild (Abb. 3.13) dargestellten Dioden D_{BE} und D_{CB} werden im Kleinsignalersatzschaltbild durch die Eingangsleitwerte g_{BE} und g_{CB} ersetzt:

$$g_{BE} = \frac{dI_{BE}}{dV_{BE}} \quad \text{und} \quad g_{CB} = \frac{dI_{BC}}{dV_{CB}}. \quad (3.20)$$

Der Leitwert g_{CB} ist bei Vorwärtsbetrieb des Bipolartransistors vernachlässigbar klein. Er muss nur bei Flusspolung des Basis-Kollektor-pn-Überganges berücksichtigt werden. Parallel zu den Eingangsleitwerten g_{BE} und g_{CB} liegen jeweils die zugehörigen Sperrsicht- und Diffusionskapazitäten (vgl. Abschn. 3.5.1).

Als wesentliche Größe zur Beschreibung des Kleinsignalverhaltens des Bipolartransistors dient die Steilheit g_m (auch Gain), die das Verhältnis der Kollektormenge

stromänderung ΔI_C zur Änderung der Basis-Emitter-Spannung ΔV_{BE} bei konstanter Kollektor-Emitter-Spannung V_{CE} angibt. Die spannungsgesteuerte Stromquelle $I_C(V_{BE})$ des Großsignalersatzschaltbildes (Abb. 3.13) wird somit durch eine lineare Stromquelle ($g_m \cdot v_{BE}$) ersetzt. Die Steilheit g_m eines npn-Bipolartransistors mit der Fläche A wird mit Hilfe der Minoritätsträgerdichte n_{B0} in der Basis bei thermodynamischem Gleichgewicht, der Diffusionskonstanten D_{nB} von Elektronen in der Basis und der neutralen Basisweite w wie folgt berechnet:

$$g_m = \frac{\partial I_C}{\partial V_{BE}} \Big|_{V_{CE}=\text{const.}} = \frac{e \cdot A \cdot D_{nB} \cdot n_{B0}}{w} \cdot \exp\left(\frac{V_{BE}}{V_\vartheta}\right) \cdot \frac{1}{V_\vartheta} = \frac{I_C}{V_\vartheta}. \quad (3.21)$$

Parallel zur linearen Stromquelle ($g_m \cdot v_{BE}$) wirkt der Ausgangsleitwert g_o , der den reziproken Ausgangswiderstand r_{CE} des Bipolartransistors und damit die Steigung der Ausgangskennlinie im Arbeitspunkt kennzeichnet:

$$g_o = \frac{\partial I_C}{\partial V_{CE}} \Big|_{V_{BE}=\text{const.}} = \frac{1}{r_{CE}} = \frac{I_C}{V_{Early}}. \quad (3.22)$$

Die Berechnung des Ausgangsleitwertes erfolgt mit Hilfe der Early-Spannung V_{Early} (vgl. Abschn. 3.2.2).

Insbesondere für höhere Frequenzen müssen zusätzlich die arbeitspunktabhängigen Bahnwiderstände von Emitter, Basis und Kollektor im Kleinsignalersatzschaltbild berücksichtigt werden, die jeweils in Serie zum entsprechenden Transistorkontakt geschaltet werden.

Neben den Elementen des Kleinsignalersatzschaltbildes sind die Verstärkeigenschaften von besonderer Bedeutung. Diese sind von der verwendeten Beschaltungsart (vgl. Abschn. 3.1.4) abhängig. Die Kleinsignalstromverstärkung β in Emitterschaltung und die Kleinsignalstromverstärkung α in Basisschaltung berechnen sich über die Beziehungen:

$$\beta = \frac{\Delta I_C}{\Delta I_B} \text{ und } \alpha = \frac{\Delta I_C}{\Delta I_E}. \quad (3.23)$$

Für niedrige Frequenzen ($\omega \rightarrow 0$) werden die Kleinsignalstromverstärkungen bei Emitterschaltung mit β_0 und bei Basisschaltung mit α_0 bezeichnet, die wie folgt bestimmt werden:

$$\beta_0 = \frac{\Delta I_C}{\Delta I_B} \Big|_{\omega \rightarrow 0} = \frac{1}{\delta} \quad \text{und} \quad \alpha_0 = -\frac{\Delta I_C}{\Delta I_E} \Big|_{\omega \rightarrow 0} = \frac{1}{1 + \delta}. \quad (3.24)$$

In Gl. (3.24) beschreibt δ den sog. *Basistransportfaktor*, der sich über das Verhältnis aus Basisstromänderung zur Kollektorstromänderung bei niedrigen Frequenzen ergibt:

$$\delta = \frac{\Delta I_B}{\Delta I_C} \Big|_{\omega \rightarrow 0} = \frac{D_{pE} \cdot w \cdot p_{E0}}{D_{nB} \cdot L_{pE} \cdot n_{B0}} + \frac{1}{2} \cdot \left(\frac{w}{L_{nB}} \right)^2. \quad (3.25)$$

Die Spannungsverstärkungen A_{VE} in Emitterschaltung und A_{VB} in Basisschaltung berechnen sich analog Gl. (3.7) und Gl. (3.10) unter Verwendung der Steilheit g_m und des Lastwiderstandes R_L (vgl. Abb. 3.7) wie folgt:

$$A_{VE} = \frac{\Delta V_{CE}}{\Delta V_{BE}} \approx -g_m \cdot R_L \quad \text{und} \quad A_{VB} = \frac{\Delta V_{CB}}{\Delta V_{BE}} \approx g_m \cdot R_L. \quad (3.26)$$

Die Frequenzabhängigkeit der Verstärkung des Bipolartransistors in Emitter- und Basisschaltung wird in Abschn. 3.5.4 näher beschrieben.

3.5.4 Grenzfrequenzen

Die Kleinsignalstromverstärkung eines Bipolartransistors ist frequenzabhängig. Zur Spezifikation dieser Frequenzabhängigkeit werden insbesondere die α -Grenzfrequenz f_α , die β -Grenzfrequenz f_β und die Transitfrequenz f_T verwendet. Die Frequenzabhängigkeit der Stromverstärkung eines Bipolartransistors und Kennzeichnung der Grenzfrequenzen ist in Abb. 3.15 schematisch dargestellt.

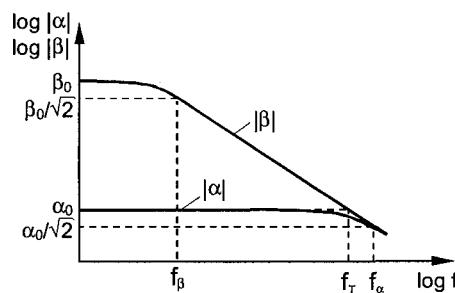


Abb. 3.15. Frequenzabhängigkeit der Stromverstärkung eines Bipolartransistors in Basis- und Emitterschaltung mit Kennzeichnung der Grenzfrequenzen f_α , f_β und f_T

Die α -Grenzfrequenz bezeichnet die Frequenz, bei der die Kleinsignalstromverstärkung $|\alpha|$ des Bipolartransistors in Basisschaltung auf den Wert $\alpha_0/\sqrt{2}$ abgesunken ist. Dabei bezeichnet α_0 die Kleinsignalstromverstärkung für niedrige Frequenzen (vgl. Abschn. 3.5.3). Für die α -Grenzfrequenz f_α folgt [3.6]:

$$f_\alpha = \frac{C_{B_{BE}}}{2\pi \cdot g_m}. \quad (3.27)$$

Hierbei bezeichnet g_m die Steilheit gemäß Gl. (3.21) und $C_{B_{BE}}$ den Diffusionskapazitätsanteil der Basiskapazität (vgl. Abschn. 3.5.1).

Die Frequenzabhängigkeit der Kleinsignalstromverstärkung $|\beta|$ eines Bipolartransistors in Emitterschaltung wird durch die β -Grenzfrequenz gekennzeichnet, bei der die Stromverstärkung um den Wert $\beta_0/\sqrt{2}$ abgesunken ist. Gemäß Abschn. 3.5.3 bezeichnet β_0 die Kleinsignalstromverstärkung für niedrige Frequenzen. Für die β -Grenzfrequenz f_β folgt [3.2]:

$$f_\beta = (1 - \alpha_0) \cdot f_\alpha \approx \frac{f_\alpha}{\beta_0} \quad \text{mit} \quad \beta_0 = \left. \frac{\Delta I_C}{\Delta I_B} \right|_{\omega \rightarrow 0} = \frac{1}{\delta}. \quad (3.28)$$

Die Transitfrequenz f_T kennzeichnet die Frequenz, bei der die Kleinsignalstromverstärkung des Bipolartransistors in Emitterschaltung auf den Wert $\beta = 1$ abgesunken ist. Die Transitfrequenz berechnet sich mit Hilfe der β -Grenzfrequenz wie folgt:

$$f_T = \beta \cdot f_\beta = \frac{I}{2 \cdot \tau_B}. \quad (3.29)$$

Für die Charakterisierung des Hochfrequenzverhaltens von Bipolartransistoren ist die Transitfrequenz eine maßgebliche Größe. Sie wird gemäß Gl. (3.29) durch die Basistransitzeit τ_B (vgl. Abschn. 3.1.3) bestimmt. Für sog. *Hochfrequenztransistoren* mit sehr geringer Basisweite gehen neben der Basistransitzeit auch die Sperrsicht- und Diffusionskapazitäten (vgl. Abschn. 3.5.1) und die Bahnwiderstände in die Bestimmung der Grenzfrequenzen ein.

Neben der Frequenzabhängigkeit der Kleinsignalstromverstärkung ist das Frequenzverhalten der Kleinsignalspannungsverstärkung (vgl. Abschn. 3.5.3) von Bedeutung. Auf diese Frequenzabhängigkeit wird in Abschn. 3.7 schaltungsspezifisch näher eingegangen.

3.6 Schaltverhalten des Bipolartransistors

Die dynamischen Eigenschaften eines Bipolartransistors werden überwiegend durch die Änderung der gespeicherten Ladung bestimmt. Die Minoritätsträgerverteilung $n_B(x)$ in der Basis ist für einen Einschaltvorgang bei Stromsteuerung in Abb. 3.16a und für einen Ausschaltvorgang bei Spannungssteuerung in Abb. 3.16b dargestellt.

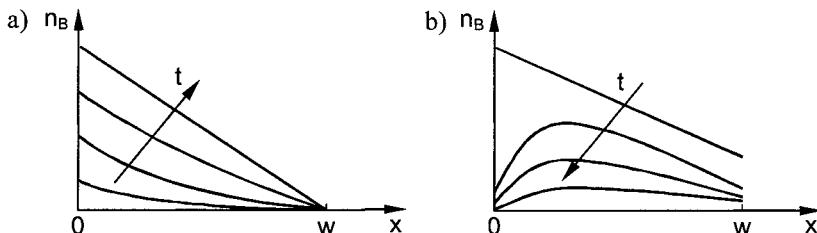


Abb. 3.16. Minoritätsträgerverteilung in der Basis eines npn-Bipolartransistors a) für Einschaltvorgang bei Stromsteuerung und b) für Ausschaltvorgang bei Spannungssteuerung

Das Einschaltverhalten eines Bipolartransistors in Basisschaltung kann durch den zeitlichen Verlauf des Kollektorstroms I_C bei Einprägung eines Emitterstromsprunges I_E charakterisiert werden. Dieses Verhalten ist in Abb. 3.17 veranschaulicht.

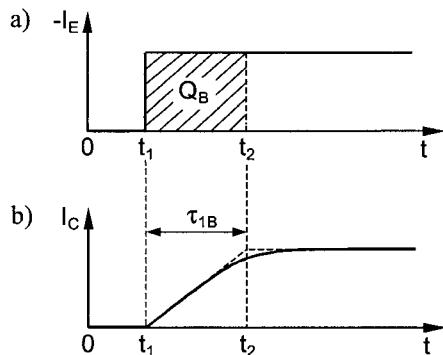


Abb. 3.17. Einschaltvorgang eines npn-Bipolartransistors in Basisschaltung a) zeitlicher Verlauf des eingeprägten Emittorstromsprunges I_E und b) resultierendes Zeitverhalten des Kollektorstroms I_C

Nach dem Einschalten zum Zeitpunkt $t = t_1$ steigt die Minoritätsträgerkonzentration $n_B(x)$ in der Basis des Bipolartransistors solange an, bis auf Grund der ebenfalls stärker werdenden Rekombination ein stationärer Zustand der Minoritätsträgerkonzentration erreicht wird (vgl. Abb. 3.16a). Bei Eintreten dieses Zustandes zum Zeitpunkt $t = t_2$ ist in der Basis die Basisladung Q_B (Minoritätsträgerladung) gespeichert, die aus der Basiskapazität C_B resultiert (vgl. Abschn. 3.5.1). Die Schaltzeit τ_{IB} des Einschaltvorganges eines Bipolartransistors in Basisschaltung stellt dann näherungsweise das Verhältnis aus der Basiskapazität zur Steilheit g_m (vgl. Abschn. 3.5.3) dar:

$$\tau_{IB} \approx \frac{C_B}{g_m} = \frac{w^2}{2 \cdot D_{nB}} . \quad (3.30)$$

Zur Realisierung kurzer Schaltzeiten müssen Bipolartransistoren für Hochfrequenzanwendungen eine möglichst geringe neutrale Basisweite w aufweisen.

Das Einschaltverhalten eines Bipolartransistors in Emitterschaltung kann bei Einprägung eines Basisstromsprunges I_B durch den resultierenden zeitlichen Verlauf des Kollektorstroms I_C ermittelt werden (Abb. 3.18). Dabei muss die Basisladung Q_B durch den Basisstrom I_B aufgebracht werden. Im Vergleich zum Emittorstrom I_E in Basisschaltung ist der Basisstrom I_B in Emitterschaltung um den Faktor β_0 (vgl. Abschn. 3.5.3) kleiner. Die volle Verstärkung erfolgt erst, nachdem die Basisladung Q_B aufgebaut wurde. Für die Schaltzeit τ_{IE} eines Bipolartransistors in Emitterschaltung gilt daher:

$$\tau_{IE} = \tau_{IB} \cdot \beta_0 \approx \frac{w^2}{2 \cdot D_{nB}} \cdot \frac{\Delta I_C}{\Delta I_B} . \quad (3.31)$$

Zur Verkürzung der Schaltzeit kann die Basis des Bipolartransistors über eine Parallelschaltung aus einem Widerstand R_I und einer Kapazität C angesteuert werden (vgl. Abb. 3.18). Dazu wird die RC-Zeitkonstante τ_{RC} näherungsweise gleich der Schaltzeit τ_{IE} gewählt ($\tau_{RC} = R_I \cdot C \approx \tau_{IE}$).

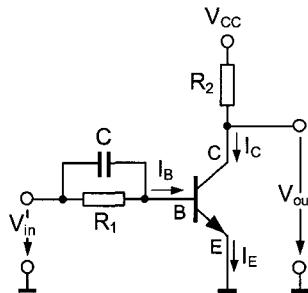


Abb. 3.18. Schaltungsprinzip eines npn-Bipolartransistors in Emitterschaltung mit Ansteuerung der Basis über eine Parallelschaltung aus R_1 und C zur Verkürzung der Transistor schaltzeit

Das Einschaltverhalten eines Bipolartransistors in Emitterschaltung mit Ansteuerung der Basis über eine Kapazität C ist in Abb. 3.19 dargestellt. Dabei wird sichtbar, dass die Basisladung Q_B schneller aufgebaut wird, wodurch die Schaltzeit auf τ'_{IE} verkürzt wird.

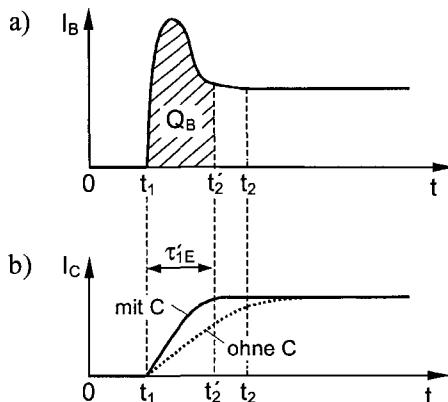


Abb. 3.19. Einschaltvorgang eines npn-Bipolartransistors in Emitterschaltung mit Ansteuerung der Basis über eine Kapazität C a) eingeprägter Basisstromsprung I_B und b) resultierendes Zeitverhalten des Kollektorstroms I_C

Der Ausschaltvorgang eines Bipolartransistors erfolgt über die Abnahme der Minoritätsträgerkonzentration in der Basis (vgl. Abb. 3.15b). Die Basisladung Q_B wird dabei abgebaut, was durch Abfließen der Ladung über den Emitter oder durch Rekombination erfolgt. Wird zum Zeitpunkt $t = t_3$ der Bipolartransistor durch einen Spannungssprung abgeschaltet, so ergibt sich das in Abb. 3.20 dargestellte zeitliche Verhalten des Kollektor- und Emitterstroms.

Die Zeitkonstanten t_s und t_f charakterisieren die Speicher- und Abfallzeit. In der Speicherzeit t_s fließt zunächst noch ein näherungsweise konstanter Emittersperrstrom I_R durch den pn-Übergang zwischen Basis und Emitter des Bipolartransistors.

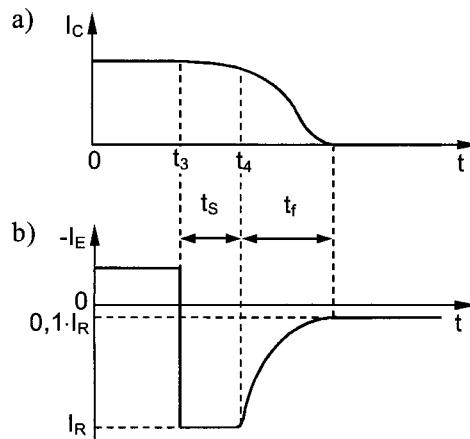


Abb. 3.20. Ausschaltvorgang eines npn-Bipolartransistors a) zeitlicher Verlauf des Kollektorstroms I_C und b) zeitlicher Verlauf des Emitterstroms I_E

Ab dem Zeitpunkt $t = t_4$ beginnen der Kollektorstrom I_C und Emitterstrom I_E abzusinken. Dabei gibt die Abfallzeit t_f an, zu welchem Zeitpunkt der Emittersperrstrom I_R auf 10% seines Endwertes abgesunken ist.

Das Schaltverhalten ist insbesondere bei der Auslegung von Bipolartransistoren als sog. *Schalttransistoren* von Bedeutung. Das Schalten von großen Lasten mit sog. *Leistungsbipolartransistoren* wird in Kap. 5 näher betrachtet.

3.7 Bipolartransistor-Grundschatungen

In den vorangegangenen Abschnitten wurden bereits die Beschaltungsarten des Bipolartransistors (Emitter-, Basis- und Kollektorschaltung) beschrieben und die entsprechenden schaltungsspezifischen Eigenschaften des Bipolartransistors hinsichtlich Verstärkungs- und Schaltverhalten hergeleitet. Nachfolgend werden weitere Grundschatungen mit Bipolartransistoren vorgestellt, deren wesentliche Eigenschaften charakterisiert und Einsatzgebiete beschrieben werden.

3.7.1 Darlington-Schaltung

Die sog. *Darlington-Schaltung* wird verwendet, wenn die Stromverstärkung eines einzelnen Bipolartransistors nicht ausreicht. Sie ermöglicht die Steuerung eines hohen Stroms im Ausgangskreis durch einen geringen Strom im Eingangskreis. Dazu werden mehrere Bipolartransistoren gemäß Abb. 3.21 so verschaltet, dass der Emitter des Transistors T_1 mit der Basis des Transistors T_2 verbunden ist und die Kollektoren beider Transistoren kurzgeschlossen sind. Die Darlington-Schaltung verhält sich an ihren äußeren Kontakten (B' , C' , E') wie ein Einzeltransistor jedoch mit höherer Stromverstärkung.

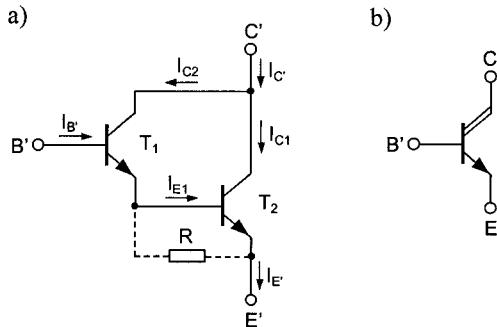


Abb. 3.21. Darlington-Schaltung a) Schaltungsprinzip und b) Schaltsymbol

Die Kleinsignalstromverstärkung β' einer Darlington-Schaltung ist annähernd gleich dem Produkt der Kleinsignalstromverstärkungen β_1 und β_2 der Einzeltransistoren T_1 und T_2 , da der Emitterstrom beider Transistoren gleich ist:

$$\beta' \approx \beta_1 \cdot \beta_2 . \quad (3.32)$$

Die Steilheit g'_m (vgl. Abschn. 3.5.3) einer Darlington-Schaltung berechnet sich auf Grund der doppelten Basis-Emitter-Spannung zu:

$$g'_m = \frac{\partial I_{C'}}{\partial V_{B'E'}} \Big|_{V_{C'E'} = \text{const.}} = \frac{I_{C'}}{2 \cdot V_\vartheta} . \quad (3.33)$$

Der Eingangswiderstand r_{in} der Darlington-Schaltung wird über die Eingangswiderstände der Transistoren T_1 und T_2 bestimmt. Sind sie für beide Transistoren gleich ($r_{BE} = r_{BE1} = r_{BE2}$), so erhält man für den Eingangswiderstand r_{in} :

$$r_{in} = r_{B'E'} = 2 \cdot r_{BE} = \frac{\beta'}{g_m} . \quad (3.34)$$

Der Ausgangswiderstand r_{out} der Darlington-Schaltung berechnet sich mit Hilfe des Ausgangswiderstandes r_{CE_2} von Transistor T_2 zu [3.5]:

$$r_{out} = r_{C'E'} = \frac{2}{3} \cdot r_{CE_2} \quad \text{mit} \quad r_{CE_2} = \frac{V_{Early_2}}{I'_C} . \quad (3.35)$$

Dabei wird der Ausgangswiderstand r_{CE_2} des Transistors T_2 über die Early-Spannung V_{Early_2} (vgl. Abschn. 3.2.2) festgelegt. Durch Ansteuerung des Transistors T_2 über den Emitter von Transistor T_1 fließt der Strom I_{E1} nur in die Basis von Transistor T_2 . Daher kann beim Abschalten kein Abbau der Basisladung Q_{B1} des Transistors T_1 durch den Rückwärtsstrom des pn-Überganges zwischen Basis und Emitter erfolgen, wodurch die Schaltgeschwindigkeit der Darlington-Schaltung reduziert wird. Durch Verwendung eines zusätzlichen Basis-Emitter-Widerstandes

R (vgl. Abb. 3.21) zwischen Basis und Emitter des Transistors T_2 kann dieser Effekt verringert werden.

Neben dem in Abb. 3.21 dargestellten Schaltungsprinzip ist auch eine komplementäre Anordnung von zwei Bipolartransistoren möglich (z. B. $T_1 = \text{pnp}$, $T_2 = \text{npn}$). In diesem Fall berechnet sich die Kleinsignalstromverstärkung β' und die Steilheit g_m' analog Gl. (3.32) und Gl. (3.33). Für den Eingangs- und Ausgangswiderstand r_{in} und r_{out} der komplementären Darlington-Schaltung resultieren folgende Beziehungen:

$$r_{in} = r_{B'E'} = r_{BE_1} = \beta' \cdot g_{m_1} \quad \text{und} \quad r_{out} = r_{C'E'} = \frac{1}{2} \cdot r_{CE_2}. \quad (3.36)$$

Die komplementäre Darlington-Schaltung besitzt damit einen geringeren Eingangs- und Ausgangswiderstand als die Darlington-Schaltung mit Bipolartransistoren gleichen Typs.

3.7.2 Stromquelle und Stromspiegel

Durch Verwendung eines Bipolartransistors als sog. *Konstantstromquelle* kann ein hoher differentieller Innenwiderstand (Ausgangswiderstand r_{out}) erzielt werden, der durch das Verhalten der Ausgangskennlinie des Bipolartransistors (vgl. Abschn. 3.2.1) begründet ist. Eine weitere Erhöhung des Innenwiderstandes kann durch Gegenkopplung erreicht werden. Das Schaltungsprinzip einer Konstantstromquelle, welches auf einer stromgegengekoppelten Emitterschaltung basiert, ist in Abb. 3.22 dargestellt. Dabei wird in eine ohmsche Last R_L der Strom I_{out} eingeprägt, der nahezu unabhängig vom Spannungsabfall V_{out} an der Last R_L ist.

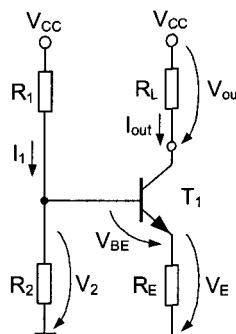


Abb. 3.22. Schaltungsprinzip einer Konstantstromquelle mit einem npn-Bipolartransistor in stromgegengekoppelter Emitterschaltung

Der Ausgangsstrom I_{out} der Konstantstromquelle berechnet sich dann mit den in Abb. 3.22 dargestellten Spannungskomponenten wie folgt:

$$I_{out} = \frac{V_E}{R_E} = \frac{V_2 - V_{BE}}{R_E} \quad \text{mit} \quad V_2 = \frac{V_{CC} \cdot R_2}{R_1 + R_2}. \quad (3.37)$$

Der Wert des Konstantstroms I_{out} kann bei gegebener Basis-Emitter-Spannung V_{BE} durch das Widerstandsverhältnis von R_1 zu R_2 eingestellt werden.

Der differentielle Ausgangswiderstand r_{out} der Schaltung wird über das Verhältnis der Ausgangsspannungsänderung dV_{out} zur Ausgangsstromänderung dI_{out} bestimmt und kann mit Hilfe der Kleinsignalsstromverstärkung β in Emitterschaltung ermittelt werden:

$$r_{out} = -\frac{dV_{out}}{dI_{out}} = r_{CE} \left[1 + \frac{\beta \cdot R_E}{(R_1 // R_2) + r_{BE} + R_E} \right] \quad \text{mit} \quad r_{CE} \approx \frac{V_{Early}}{I_{out}}. \quad (3.38)$$

Bei der Dimensionierung der Konstantstromquelle gibt man zunächst den Spannungsabfall V_E am Widerstand R_E vor. Je größer dieser Spannungsabfall gewählt wird, desto höher wird der Ausgangswiderstand r_{out} . Gleichzeitig verringert sich bei gegebener Betriebsspannung V_{CC} der maximale Spannungsabfall V_{out} am Lastwiderstand R_L . Der Spannungsteiler aus R_1 und R_2 wird möglichst niederohmig ausgelegt, um den Ausgangswiderstand r_{out} nicht merklich zu erhöhen.

Kompensiert man die Temperaturabhängigkeit des Basispotentials der Konstantstromquelle, so erhält man den sog. *Stromspiegel*. Diese Temperaturkompen-sation wird durch Verwendung eines weiteren Bipolartransistors T_2 mit kurzgeschlossenem Basis- und Kollektoranschluss erreicht. Für die Spannungen des als Diode betriebenen Transistors T_2 gilt dann $V_{CE} = V_{BE} > V_{CE,Sat}$. Das Schaltungs-prinzip eines Stromspiegels mit zwei npn-Bipolartransistoren T_1 und T_2 ist in Abb. 3.23 dargestellt.

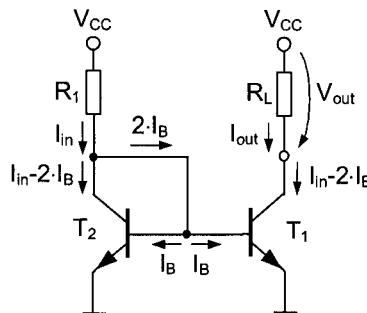


Abb. 3.23. Schaltungsprinzip eines Stromspiegels mit zwei identischen npn-Bipolartransistoren und Darstellung der auftretenden Stromkonstellation

Sind beide Bipolartransistoren T_1 und T_2 identisch, so besitzen sie gemäß ihrer Beschaltung nach Abb. 3.23 den gleichen Kollektorstrom ($I_{C1} = I_{C2} = I_{in} - 2 \cdot I_B$). Die Gleichheit der Kollektorströme ist dabei nahezu unabhängig von der ohmschen Last R_L , solange der Transistor T_2 im aktiven Bereich betrieben wird. Somit wird der Strom durch den Transistor T_2 in Transistor T_1 „gespiegelt“. Für den Ausgangsstrom I_{out} folgt dann, dass dieser näherungsweise gleich dem Eingangsstrom I_{in} ist:

$$I_{out} = \frac{V_{CC} - V_{BE}}{R_E} = I_{in} - 2 \cdot I_B \approx I_{in}. \quad (3.39)$$

Der differentielle Ausgangswiderstand r_{out} des Stromspiegels wird durch das Verhältnis von Ausgangsspannungsänderung dV_{out} zur Ausgangsstromänderung dI_{out} gekennzeichnet. Der Ausgangswiderstand des Stromspiegels wird dann durch den Ausgangswiderstand r_{CE1} des Bipolartransistors T_1 bestimmt (vgl. Abschn. 3.5.3), der mit Hilfe der Early-Spannung V_{Early1} berechnet werden kann:

$$r_{out} = -\frac{dV_{out}}{dI_{out}} = r_{CE1} = \frac{V_{Early1}}{I_{out}}. \quad (3.40)$$

Zur Erhöhung des Ausgangswiderstandes r_{out} des Stromspiegels können jeweils Widerstände zwischen Masse und Emitter von T_1 und T_2 eingesetzt werden. Durch Parallelschaltung von weiteren Transistoren lassen sich Vielfache des Eingangsstroms erzeugen.

Eine weitere Möglichkeit, einen hoch präzisen Stromspiegel mit sehr hohem Ausgangswiderstand zu realisieren, stellt der sog. *Wilson-Stromspiegel* dar. Bei dieser Schaltung wird der Ausgangsstrom I_{out} über den aus den npn-Bipolartransistoren T_1 und T_2 bestehenden Stromspiegel rückgekoppelt. Das Schaltungsprinzip eines Wilson-Stromspiegels und die auftretende Stromkonstellation ist in Abb. 3.24 dargestellt.

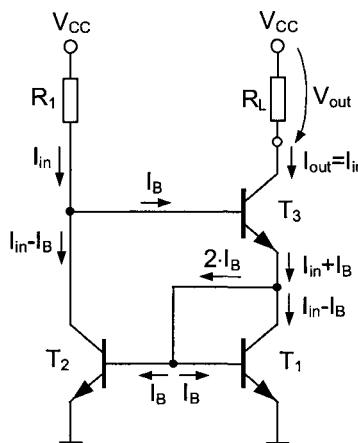


Abb. 3.24. Schaltungsprinzip eines Wilson-Stromspiegels mit npn-Bipolartransistoren und Darstellung der auftretenden Stromkonstellation

Der Ausgangsstrom I_{out} des Wilson-Stromspiegels ist gleich dem Eingangsstrom I_{in} . Der differentielle Ausgangswiderstand r_{out} des Wilson-Stromspiegels kann bei Verwendung von identischen Transistoren näherungsweise mit Hilfe der Kleinsignalstromverstärkung β in Emitterschaltung und des Ausgangswiderstandes r_{CE3} des Bipolartransistors T_3 berechnet werden:

$$r_{out} = -\frac{dV_{out}}{dI_{out}} \approx \frac{r_{CE3} \cdot \beta}{2} \text{ mit } r_{CE3} = \frac{V_{Early3}}{I_{out}}. \quad (3.41)$$

3.7.3 Differenzverstärker

Der sog. *Differenzverstärker* dient zur Verstärkung der Differenz von zwei Eingangsspannungen. Differenzverstärker weisen eine geringe Abhängigkeit der Ausgangsspannung gegenüber Parameterstreuungen und Temperaturschwankungen auf. Bei der Differenzverstärkerschaltung liegen die Eingangsspannungen an den Basisanschlüssen von zwei Bipolartransistoren T_1 und T_2 , deren Emitter verbunden sind, wobei die Summe der Emitterströme I_{E1} und I_{E2} mit einer Konstantstromquelle konstant gehalten wird ($I_E = I_{E1} + I_{E2}$). Das Schaltungsprinzip des Differenzverstärkers ist in Abb. 3.25 dargestellt.

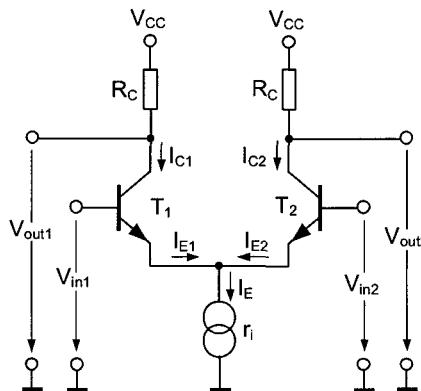


Abb. 3.25. Schaltungsprinzip eines Differenzverstärkers mit npn-Bipolartransistoren und Stromquelle

Sind die Eingangsspannungen V_{in1} und V_{in2} des Differenzverstärkers gleich, so ergibt sich für die Emitterströme der Bipolartransistoren $I_{E1} = I_{E2} = \frac{1}{2} \cdot I_E$. Daraus resultiert für die Kollektorströme die Beziehung $I_{C1} = I_{C2} \approx \frac{1}{2} \cdot I_E$. Besitzt beispielsweise die Eingangsspannung V_{in1} einen höheren Wert als die Eingangsspannung V_{in2} , so steigt der Kollektorstrom I_{C1} an und der Kollektorstrom I_{C2} sinkt im gleichen Maße ab. Die Summe der Kollektorströme bleibt dabei gleich ($I_{C1} + I_{C2} \approx I_E$). Das Verhältnis der Kollektorströme berechnet sich dann für eine Eingangsspannungsdifferenz ΔV_{in} wie folgt:

$$\frac{I_{C1}}{I_{C2}} = \exp\left(\frac{\Delta V_{in}}{V_\theta}\right) \quad \text{mit } \Delta V_{in} = V_{in1} - V_{in2}. \quad (3.42)$$

Aus der Eingangsspannungsdifferenz ΔV_{in} resultiert die Ausgangsspannungsdifferenz ΔV_{out} , die über die Kleinsignalstromverstärkung α in Basisschaltung, den Kollektorstrom I_E und den Wert des Kollektorwiderstandes R_C bestimmt werden kann:

$$\Delta V_{out} = V_{out1} - V_{out2} = -\alpha \cdot I_E \cdot R_C \cdot \tanh\left(\frac{\Delta V_{in}}{2 \cdot V_\theta}\right). \quad (3.43)$$

Der Verlauf der Kollektorströme I_{C1} und I_{C2} sowie der Ausgangsspannungsdifferenz $\Delta V_{out} = V_{out1} - V_{out2}$ ist in Abb. 3.26 in Abhängigkeit von der Eingangsspannungsdifferenz ΔV_{in} dargestellt [3.1].

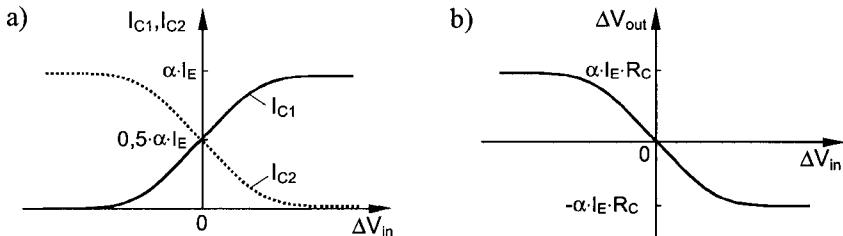


Abb. 3.26. a) Verlauf der Kollektorströme und b) Verlauf der Ausgangsspannungsdifferenz ΔV_{out} eines Differenzverstärkers in Abhängigkeit von der Eingangsspannungsdifferenz ΔV_{in}

Als wichtiges Merkmal dient die *Differenzverstärkung* A_D , die das Verhältnis der jeweiligen Ausgangsspannungsänderung dV_{out1} bzw. dV_{out2} zur Eingangsspannungsänderung dV_{in} angibt und sich wie folgt berechnet:

$$A_D = \frac{dV_{out1}}{dV_{in}} = -\frac{dV_{out2}}{dV_{in}} = -\frac{1}{2} \cdot g_m \cdot \left(\frac{R_C \cdot r_{CE}}{R_C + r_{CE}} \right). \quad (3.44)$$

Dabei bezeichnet g_m die Steilheit der Bipolartransistoren. Der Ausgangswiderstand r_{out} eines Differenzverstärkers wird aus der Parallelschaltung des Kollektowiderstandes R_C und des Ausgangswiderstandes r_{CE} des Bipolartransistors bestimmt:

$$r_{out} = \frac{R_C \cdot r_{CE}}{R_C + r_{CE}} \text{ mit } r_{CE} = \frac{V_{Early}}{I_C}. \quad (3.45)$$

Die sog. *Gleichaktverstärkung* A_G eines Differenzverstärkers berechnet sich bei Berücksichtigung des Innenwiderstandes r_i der Stromquelle wie folgt:

$$A_G = \frac{dV_{out1}}{dV_G} = \frac{dV_{out2}}{dV_G} = -\frac{1}{2} \cdot \frac{R_C}{r_i} \text{ mit } V_G = \frac{1}{2} \cdot (V_{in1} + V_{in2}). \quad (3.46)$$

Als weitere charakteristische Größe des Differenzverstärkers dient die *Gleichaktunterdrückung* G , die das Verhältnis von Differenzverstärkung A_D und Gleichaktverstärkung A_G darstellt und näherungsweise dem Produkt aus der Steilheit g_m und dem Innenwiderstand r_i der Konstantstromquelle entspricht:

$$G = \frac{A_D}{A_G} \approx g_m \cdot r_i. \quad (3.47)$$

Ein typischer Verlauf des Frequenzganges der Differenzverstärkung A_D und der Gleichaktunterdrückung G eines Differenzverstärkers ist in Abb. 3.27 dargestellt [3.5].

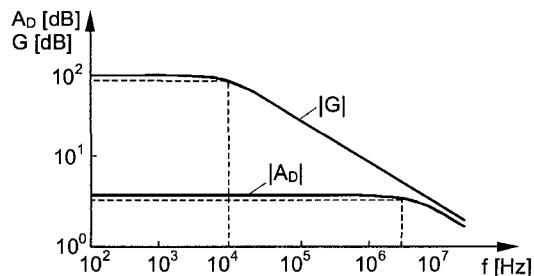


Abb. 3.27. Typischer Verlauf des Frequenzganges der Differenzverstärkung A_D und Gleichtaktunterdrückung G eines Differenzverstärkers

Der Differenzverstärker findet insbesondere Anwendung als Eingangsstufe von sog. *Operationsverstärkern*, die in Kap. 4 anhand des Feldeffekttransistors näher beschrieben werden.

4 Feldeffekttransistoren

Die Funktion von Feldeffekttransistoren beruht auf dem Effekt, dass die Spannung über einer isolierenden Schicht die Art und Dichte der beweglichen Ladungsträger an der Grenzfläche eines Halbleiters und damit die Leitfähigkeit steuert. Dieser Effekt wird *Feldeffekt* genannt. Zur Bestimmung der Spannungsabhängigkeit des Feldeffektes wird das Bändermodell einer *MOS* (Metall-Oxid-Halbleiter)-Struktur betrachtet und die primären Parameter hergeleitet. Aufbauend auf den Eigenschaften der MOS-Struktur wird die Funktionsweise des MOS-Feldeffekttransistors (*MOSFET*) beschrieben und die wichtigsten Merkmale, Modelle und die technologische Realisierung dargestellt. Die schaltungstechnische Anwendung des MOSFETs wird anhand ausgewählter Grundschaltungen erläutert.

4.1 MOS-Struktur

Die Funktionalität des MOS-Feldeffekttransistors basiert auf den Eigenschaften der MOS-Struktur, deren prinzipieller Aufbau in Abb. 4.1 dargestellt ist.



Abb. 0.1. Prinzipieller Aufbau einer MOS-Struktur

Übliche Materialkompositionen einer MOS-Struktur sind Metall oder n⁺⁺- bzw. p⁺⁺-Poly-Silizium als Metallisierung, SiO₂ als Oxid und p- oder n-Silizium als Halbleitersubstrat. Der Metallisierungskontakt wird als *Gate* bezeichnet; das Halbleitersubstrat wird *Bulk* genannt. Die zugehörigen Spannungen sind die *Gatespannung* V_G und *Bulkspannung* V_B . In den folgenden Betrachtungen wird, soweit nicht anders vermerkt, von p-Silizium-Halbleitern mit n⁺⁺-Silizium-Gate ausgegangen.

4.1.1 MOS-Struktur im thermodynamischen Gleichgewicht

Auch ohne externe Spannung am Gate ($V_G = 0$) tritt innerhalb der MOS-Struktur eine inhomogene Ladungsverteilung auf. Diese wird im nachfolgenden Bändermodell durch die auftretende Bandverbiegung sichtbar.

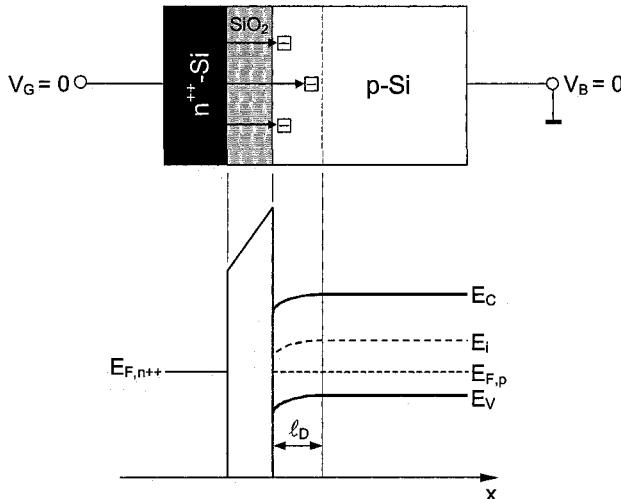


Abb. 4.2. Prinzip und Bändermodell einer MOS-Struktur bei thermodynamischem Gleichgewicht ($V_G = 0$)

Die Ursache für die inhomogene Ladungsverteilung ist das auftretende Kontaktpotential durch unterschiedliche Austrittsarbeiten des Gate- und Bulk-Materials. Die Kontaktspannung setzt sich aus einer Spannung über dem Oxid und einer Spannung über der Raumladungszone zusammen, die durch eine Bandverbiegung sichtbar wird. Die Raumladungszone der Weite ℓ_D enthält negativ geladene Akzeptoren, deren Ladung durch Löcher nicht mehr kompensiert wird. Das Gate lädt sich positiv auf und es entsteht eine Potentialdifferenz über dem Oxid.

4.1.2 MOS-Struktur im Flachbandfall

Um einen völlig flachen Verlauf der Bandkanten im Bändermodell zu erhalten, muss eine negative Spannung an das Gate der MOS-Struktur angelegt werden. Die Spannung, bei der sich ein gerader Verlauf der Bandkanten einstellt, wird *Flachbandspannung* genannt. Die Flachbandspannung V_{FB} stellt bei Verwendung von Metall als Gatematerial die Differenz des Ferminiveaus $E_{F,p}$ im p-Typ-Halbleiter und des Ferminiveaus $E_{F,Me}$ im Metall bezogen auf das Vakuumniveau E_0 dar und berechnet sich wie folgt:

$$V_{FB} = \frac{1}{e} \cdot (E_{F,p} - E_{F,Me}). \quad (4.1)$$

Für eine MOS-Struktur mit n⁺⁺-Silizium-Gate erhält man für die Flachbandspannung:

$$V_{FB} = -\frac{1}{e} \cdot (E_C - E_i) - \phi_F < 0 \quad \text{mit} \quad \phi_F = \frac{1}{e} \cdot (E_i - E_F) = V_\vartheta \cdot \ln \frac{N_A}{n_i}. \quad (4.2)$$

Dabei bezeichnet E_i das Eigenleitungs-Ferminiveau und E_C die Energie der unteren Leitungsbandkante sowie ϕ_F das sog. *Fermipotential*. Über das Fermipotential wird hier die Donatordotierungsdichte N_A mit der Eigenleitungsdichte n_i in Beziehung gebracht. Für einen n-Typ-Halbleiter gilt analog $\phi_F = -V_D \cdot \ln(N_D/n_i)$.

In Abb. 4.3 ist das Bändermodell einer MOS-Struktur für den Flachbandfall mit den charakteristischen Größen dargestellt.

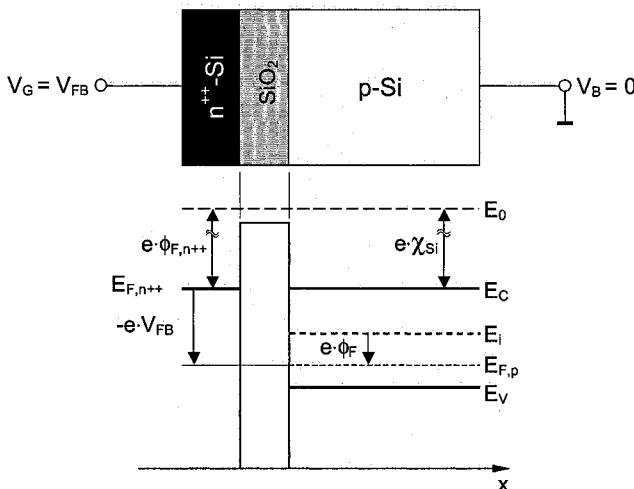


Abb. 4.3. Prinzip und Bändermodell einer MOS-Struktur im Flachbandfall ($V_G = V_{FB} \neq 0$)

4.1.3 MOS-Struktur in Akkumulation

Im *Akkumulationsfall* ($V_G < V_{FB}$) lagern sich positive Ladungsträger (Löcher) im p-Silizium-Substrat unter dem Oxid an. Da diese Ladungsträger frei beweglich sind, wird eine Flächenladung an der Halbleiteroberfläche erzeugt (Majoritätsträgeranreicherung). Die Flächenladungsdichte in Akkumulation kann mit Hilfe der Gateoxidkapazität C_{ox} berechnet werden:

$$C_{ox} = \frac{\epsilon_0 \cdot \epsilon_{ox}}{t_{ox}}, \quad (4.3)$$

wobei t_{ox} die Dicke der Oxidschicht kennzeichnet. Für die Flächenladungsdichte $Q_{Akk,p}$ des p-Typ-Halbleiters bzw. $Q_{Akk,n}$ des n-Typ-Halbleiters resultiert dann:

$$Q_{Akk,p} = C_{ox} \cdot (V_{FB} - V_G), \quad (4.4)$$

$$Q_{Akk,n} = -C_{ox} \cdot (V_{FB} - V_G). \quad (4.5)$$

Dabei bezeichnet V_{FB} die Flachbandspannung und V_G die Gatespannung der MOS-Struktur. Der Bänderverlauf für eine MOS-Struktur in Akkumulation ist in Abb. 4.4 dargestellt.

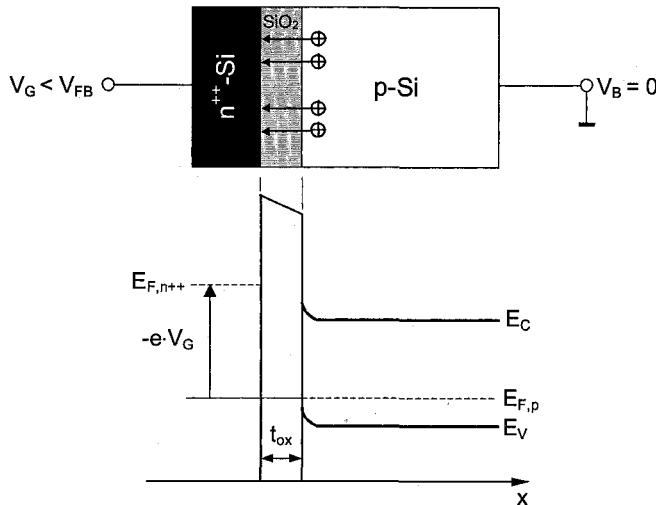


Abb. 4.4. Prinzip und Bändermodell einer MOS-Struktur in Akkumulation ($V_G < V_{FB}$)

4.1.4 MOS-Struktur in Verarmung

Im *Verarmungs-* bzw. *Depletionfall* ($V_{FB} < V_G \leq V_T$) bewirkt die nun positive Gateladung eine Majoritätsträgerverarmung an der Substratoberfläche. Die verbleibenden negativen, fest gebundenen Ionen bilden eine räumlich verteilte Ladung und es prägt sich eine Sperrschiicht (Depletion-Schicht) aus. Die Flächenladungsdichte Q_D im Depletionfall berechnet sich aus der Akzeptordotierungsdichte N_A und der Länge der Raumladungszone ℓ_D wie folgt:

$$Q_D = -e \cdot N_A \cdot \ell_D . \quad (4.6)$$

Die Potentialdifferenz V_{ox} über dem Oxid (vgl. Abb. 4.5) kann dann mit Hilfe der Gateoxidkapazität C_{ox} gemäß Gl. 4.3 ermittelt werden:

$$V_{ox} = \frac{-Q_D}{C_{ox}} . \quad (4.7)$$

Die Potentialdifferenz über der Raumladungszone entspricht bei einer Bulkspannung von $V_B = 0$ dem Grenzflächenpotential ϕ_S mit:

$$\phi_S = \frac{e \cdot N_A \cdot \ell_D^2}{2 \cdot \epsilon_0 \cdot \epsilon_{Si}} . \quad (4.8)$$

Das Bändermodell einer MOS-Struktur im Depletionfall mit dem zugehörigen Potentialverlauf $\phi(x)$ ist in Abb. 4.5 dargestellt.

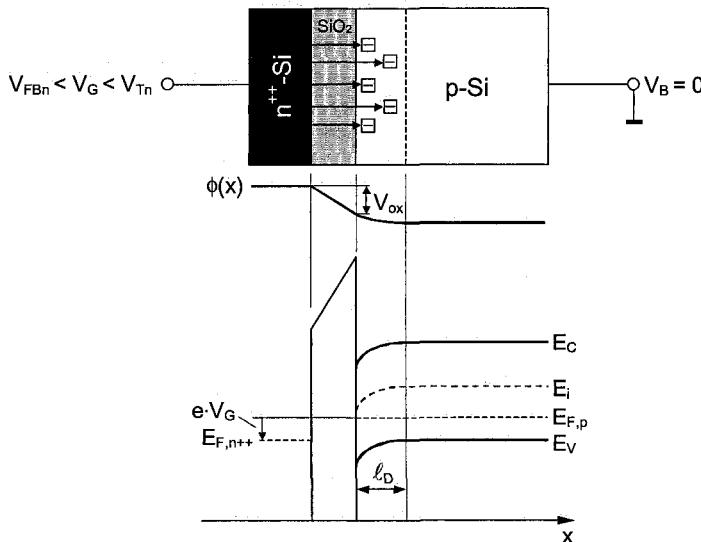


Abb. 4.5. Prinzip und Bändermodell einer MOS-Struktur in Depletion ($V_{FB} < V_G \leq V_T$)

Die Bandkanten im Depletionfall sind nahe der Grenzfläche Oxid/Halbleiter nach unten gebogen, so dass $E_i(x=0) - E_F < e \cdot \phi_F$ ist. Für die Ladungsträgerdichten p_S und n_S an der Grenzfläche Oxid/Halbleiter gilt dann $p_S < p_{p0}$ und $n_S > n_{p0}$.

4.1.5 MOS-Struktur in Inversion

Im *Inversionsfall* ($V_G \geq V_T$) ist die Gatespannung an der MOS-Struktur so hoch, dass $E_i(x=0) = E_F$ an der Grenzfläche gilt. Dann nimmt dort die Ladungsträgerdichte der Elektronen den Wert der Ladungsträgerdichte der Löcher an ($n_S = p_S = n_i$). Bei $E_i(x=0) < E_F$ stellt sich $n_S > n_i > p_S$ ein; man spricht dann von *Inversion*. Ist die Elektronendichte n_S an der Grenzfläche klein gegenüber der Akzeptordotierungsdichte ($n_S < N_A$), so hat n_S nur einen sehr geringen Einfluss auf den Potentialverlauf. Dieser Zustand wird als *schwache Inversion* bezeichnet und es gilt:

$$E_F - E_i(x=0) < e \cdot \phi_F . \quad (4.9)$$

Die Bandverbiegung am Inversionspunkt berechnet sich dann zu:

$$e \cdot \phi_S = -(E_i(x=0) - E_i(x > \ell_D)) = e \cdot \phi_F . \quad (4.10)$$

An der Grenzfläche Oxid/Halbleiter liegt bei Erreichen der sog. *Einsatzspannung* ($V_G = V_T$) das Eigenleitungsniveau um $e \cdot \phi_F$ unterhalb des Ferminiveaus E_F . Dann ist die Elektronendichte n_S an der Grenzfläche etwa gleich der Löcherdichte im Volumen ($n_S \approx p_{p0}$) und es folgt:

$$E_F - E_i(x=0, V_G = V_T) = E_i(x > \ell_D) - E_F = e \cdot \phi_F . \quad (4.11)$$

Für das Grenzflächenpotential ϕ_S bzw. die resultierende Bandverbiegung bei $V_G = V_T$ folgt dann:

$$\phi_S = -\frac{1}{e} \cdot (E_i(x=0) - E_i(x > \ell_D)) = 2 \cdot \phi_F. \quad (4.12)$$

Man spricht hier vom Einsatz der *starken Inversion*. Das Bändermodell einer MOS-Struktur in starker Inversion ist in Abb. 4.6 dargestellt.

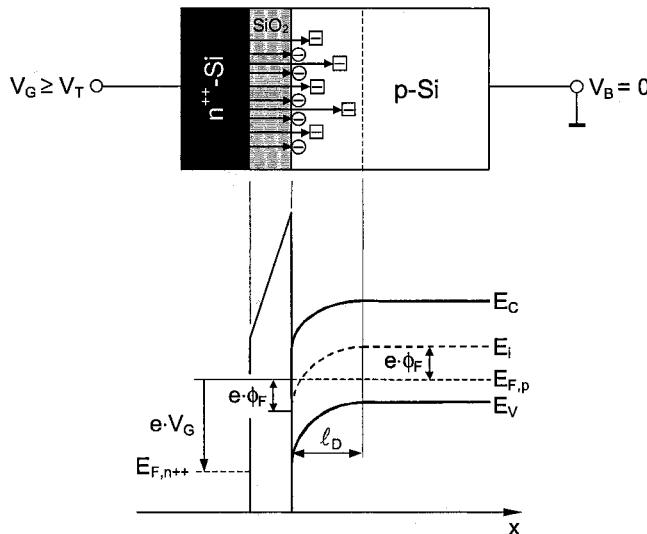


Abb. 4.6. Prinzip und Bändermodell einer MOS-Struktur in starker Inversion ($V_G \geq V_T$)

Für $V_G > V_T$ nimmt die Elektronendichte weiter zu; die Weite der Raumladungszone wächst jedoch fast nicht weiter an. Die Flächenladungsdichte Q_I in Inversion wird dann mit Hilfe der Gateoxidkapazität C_{ox} wie folgt bestimmt werden:

$$Q_I \approx -C_{ox} \cdot (V_G - V_T). \quad (4.13)$$

Die Einsatzspannung V_T für starke Inversion setzt sich aus der Flachbandspannung V_{FB} , der Potentialdifferenz in der Raumladungszone $2 \cdot \phi_F$, der Potentialdifferenz über dem Oxid $V_{ox} = -Q_D/C_{ox}$ und einem Beitrag fester Ladungen zusammen:

$$V_T = V_{FB} + 2\phi_F + \frac{1}{C_{ox}} \cdot \sqrt{2 \cdot \epsilon_0 \cdot \epsilon_{Si} \cdot e \cdot N_A \cdot 2\phi_F} - \frac{1}{C_{ox}} \cdot Q_{SS}. \quad (4.14)$$

Feste Ladungen an der Grenzfläche zwischen Oxid und Halbleiter bewirken bei der Einsatzspannung V_T einen zusätzlichen Spannungsanteil $\Delta V_T(Q_{SS})$, der von der Oberflächenladungsdichte Q_{SS} und der Gateoxidkapazität C_{ox} abhängig ist:

$$\Delta V_T(Q_{SS}) = -\frac{1}{C_{ox}} \cdot Q_{SS}. \quad (4.15)$$

Der Einfluss fester Ladungen in einer MOS-Struktur bei starker Inversion ist in Abb. 4.7 veranschaulicht.

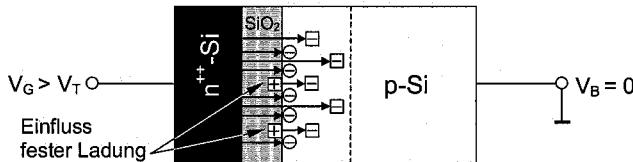


Abb. 4.7. MOS-Struktur in starker Inversion mit Einfluss fester Ladungen (hier: \blacksquare)

Die Flächenladungsdichte Q_D in der Raumladungszone berechnet sich wie folgt:

$$Q_D = -\sqrt{2 \cdot \epsilon_0 \cdot \epsilon_{Si} \cdot e \cdot N_A \cdot \phi_S} = -e \cdot N_A \cdot \ell_D. \quad (4.16)$$

Die Einsatzspannung kann vor allem durch die Wahl des Gatematerials, die Oxiddicke, die Substratdotierung, feste Ladungen nahe der Grenzfläche oder zusätzliche Dotierung im Bereich der Raumladungszone (*Einsatzspannungsimplantation*) technologisch beeinflusst werden.

4.1.6 Kapazität der MOS-Struktur

Im Akkumulationsfall ($V_G < V_{FB}$) wird die Gate-Bulk-Kapazität C_{GB} einer MOS-Struktur mit der Fläche A_G im wesentlichen durch die Flächenladung an der Halbleiteroberfläche gebildet, die dem Wert der Gateoxidkapazität C_{ox} entspricht:

$$C_{GB} \approx C_{ox} = \frac{\epsilon_0 \cdot \epsilon_{ox} \cdot A_G}{t_{ox}}. \quad (4.17)$$

Im Depletionfall ($V_{FB} < V_G \leq V_T$) entsteht durch die negativen, fest gebundenen Ionen eine Sperrschiicht, die eine sog. *Depletionkapazität* C_D besitzt. Diese liegt in Serie zur Gateoxidkapazität C_{ox} , woraus für die Gate-Bulk-Kapazität C_{GB} folgt:

$$C_{GB} \approx \frac{C_{ox} \cdot C_D}{C_{ox} + C_D} \quad \text{mit} \quad C_D = \frac{\epsilon_0 \cdot \epsilon_{Si}}{\ell_D}. \quad (4.18)$$

Bei Erhöhung der Gatespannung V_G nimmt die Weite der Raumladungszone ℓ_D (vgl. Kap. 2) zu, wodurch die Depletionkapazität C_D und damit die Gate-Bulk-Kapazität C_{GB} abnimmt. Das Wirken der Gateoxid- und Depletionkapazität ist schematisch in Abb. 4.8 dargestellt.

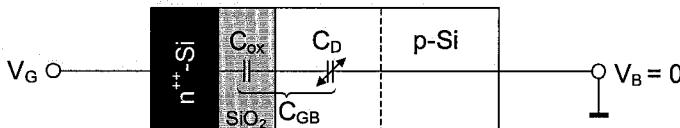


Abb. 4.8. Kapazitätsmodell einer MOS-Struktur in Depletion ($V_{FB} < V_G \leq V_T$)

Bei schwacher Inversion ($\phi_S \approx \phi_F$) und niedrigen Frequenzen ($f \ll 1/\tau_p$) erfolgt ein Teil der Ladungsänderung in der Inversionsschicht. Dies entspricht einer zusätzlichen frequenzabhängigen Kapazität parallel zur Depletionkapazität C_D . Die minimale Gate-Bulk-Kapazität $C_{GB,min}$ der MOS-Struktur tritt dann zwischen dem Inversionspunkt ($\phi_S = \phi_F$) und dem Einsatz starker Inversion ($\phi_S = 2 \cdot \phi_F$) auf.

Bei starker Inversion ($\phi_S \geq 2 \cdot \phi_F$) und niedrigen Frequenzen ($f \ll 1/\tau_p$) folgt die Inversionsladung der Spannungsänderung entsprechend der stationären Abhängigkeit durch Nettorekombination bzw. -generation. Die Inversionsladung verhindert dabei das Wirken der Depletionkapazität C_D , wobei sich der Wert der Gate-Bulk-Kapazität dem Wert der Gateoxidkapazität annähert:

$$C_{GB} \approx C_{ox} = \frac{\epsilon_0 \cdot \epsilon_{ox} \cdot A_G}{t_{ox}}. \quad (4.19)$$

Bei einer Kleinsignalspannung mit hoher Frequenz ($f \gg 1/\tau_p$) folgt die Inversionsladung zwar der Spannung im Arbeitspunkt, aber nicht der Wechselspannung. Die Ladungsänderung durch das hochfrequente Signal erfolgt am substratseitigen Rand der Raumladungszone durch die Änderung der Majoritätsträgerdichte, d. h. durch Änderung der Weite der Raumladungszone ℓ_D . Die Raumladungszone dehnt sich bei Änderung des Arbeitspunktes in starker Inversion ($V_G > V_T$) kaum mehr aus. Die Gate-Bulk-Kapazität nähert sich dann der minimalen Kapazität $C'_{GB,min}$ an:

$$C_{GB} \approx C'_{GB,min} \sim \ell_D(V_G = V_T) = \sqrt{\frac{2\epsilon \cdot 2\phi_F}{e \cdot N_A}}. \quad (4.20)$$

Die resultierende Kapazitäts-Spannungs-Charakteristik einer MOS-Struktur ist in Abb. 4.9 für die entsprechenden Betriebsbereiche dargestellt.

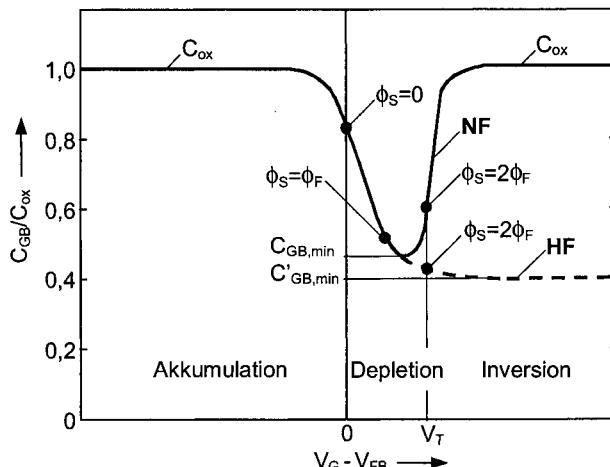


Abb. 4.9. Kapazitäts-Spannungs-Charakteristik einer MOS-Struktur bei Akkumulation, Depletion und Inversion für Niederfrequenz (NF)- und Hochfrequenz (HF)-Kleinsignalverhalten

Die in Abb. 4.9 dargestellte Kapazitäts-Spannungs-Charakteristik enthält die auf die Gatoxidkapazität normierte Gate-Bulk-Kapazität C_{GB}/C_{ox} in Abhängigkeit von der auf die Flachbandspannung bezogenen Gate-Bulk-Spannung $V_{GB} - V_{FB}$. Somit tritt bei einer Spannung von $V_{GB} - V_{FB} = 0$ an der MOS-Struktur ein Grenzflächenpotential von $\phi_S = 0$ auf, da hier die Gate-Bulk-Spannung den Wert der Flachbandspannung besitzt und somit der Flachbandfall (vgl. Abschn. 4.1.2) vorliegt.

Für die Kapazitäts-Spannungs-Charakteristik des MOS-Feldeffekttransistors (vgl. Abschn. 4.7.1) gilt der in Abb. 4.9 dargestellte Verlauf für Niederfrequenzverhalten.

4.1.7 Komplementäre MOS-Strukturen

Um Löcherleitung anstelle Elektronenleitung im Inversionsfall ($V_G \geq V_T$) der MOS-Struktur zu erhalten, wird anstatt eines p-Typ-Halbleiters ein n-Typ-Halbleiter als Substrat verwendet. Die Schichtfolge und das Bänderdiagramm einer MOS-Struktur mit n⁺⁺-Silizium-Gate und n-Silizium-Substrat bei thermodynamischem Gleichgewicht ($V_G = 0$) sind in Abb. 4.10 dargestellt.

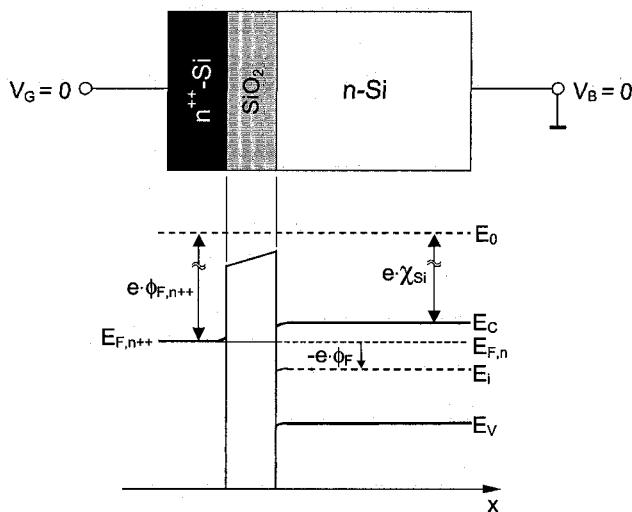


Abb. 4.10. Prinzip und Bändermodell einer MOS-Struktur mit n⁺⁺-Silizium-Gate und n-Silizium-Substrat bei thermodynamischem Gleichgewicht ($V_G = 0$)

In modernen Herstellungsprozessen wird für p-Kanal-MOSFETs häufig p⁺⁺-Poly-Silizium als Gatematerial eingesetzt. In der folgenden Übersicht (Abb. 4.11) sind daher die Bänderschemata für MOS-Strukturen mit den Materialkompositionen n⁺⁺-Poly-Silizium/SiO₂/p-Silizium und p⁺⁺-Poly-Silizium/SiO₂/n-Silizium für die Betriebsfälle Akkumulation, Depletion und starke Inversion gegenübergestellt.

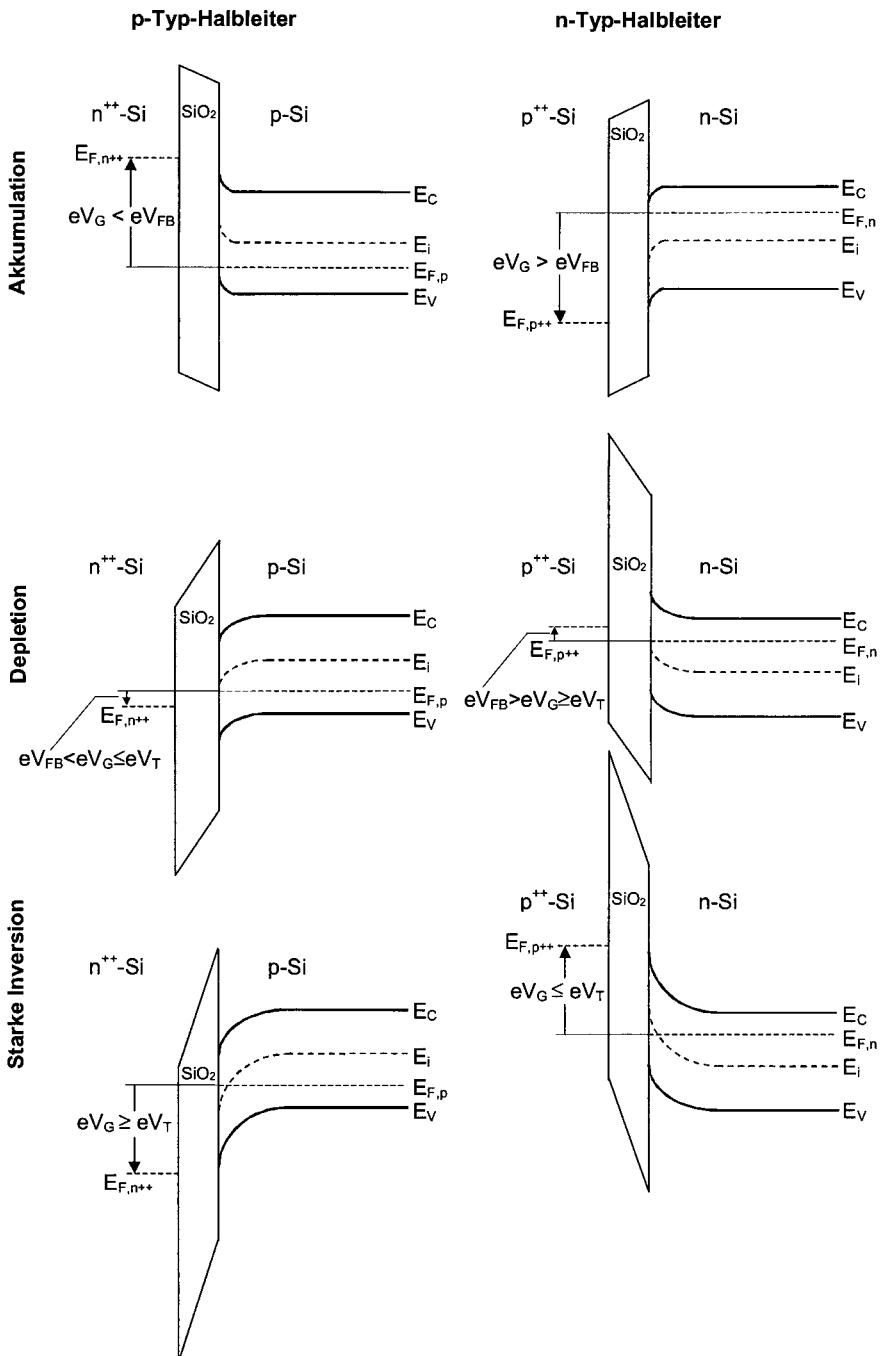


Abb. 4.11. Bändermodelle für MOS-Strukturen mit p-Silizium- und n-Silizium-Substrat und n⁺⁺-Poly-Silizium- bzw. p⁺⁺-Poly-Silizium-Gate (schematische Darstellung)

4.2 Aufbau und Funktionsweise des MOS-Feldeffekttransistors

4.2.1 Aufbau des MOSFETs

Der MOSFET besteht aus einer MOS-Struktur (vgl. Abschn. 4.1), die um zwei hoch dotierte Gebiete im Substrat erweitert ist. Diese Gebiete werden als *Source* und *Drain* bezeichnet. Dabei werden das Drain- und Sourcegebiet jeweils von den gegenüberliegenden Kanten des Gateoxids überlappt. Verwendet man eine Struktur mit p-Substrat und n⁺-dotiertem Drain- und Sourcegebiet, so spricht man von einem *n-Kanal-MOSFET*. Ist das Substrat hingegen n-dotiert und das Drain- und Sourcegebiet p⁺-dotiert, so erhält man einen *p-Kanal-MOSFET*.

In Abb. 4.12a ist der schematische Aufbau eines n-Kanal-MOSFETs im Querschnitt dargestellt. Die Schaltsymbole eines n- und p-Kanal-MOSFETs sind in Abb. 4.12b bzw. Abb. 4.12c abgebildet. Dabei werden die vereinfachten Schaltsymbole ohne Darstellung des Bulkanschlusses genutzt, wenn Bulk und Source auf dem gleichen Potential liegen.

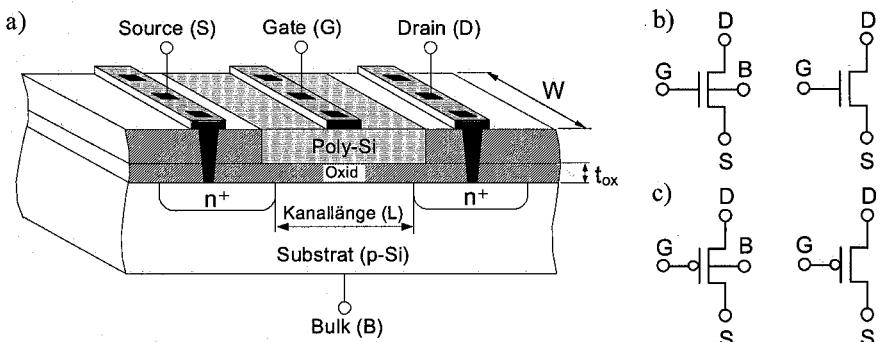


Abb. 4.12. a) Schematischer Aufbau eines n-Kanal-MOSFETs im Querschnitt b) Schaltsymbole eines n-Kanal-MOSFETs und c) Schaltsymbole eines p-Kanal-MOSFETs

Ist der Kanal des MOSFETs bei einer Gate-Source-Spannung $V_{GS} = 0$ leitend, so bezeichnet man diesen MOSFET als *Verarmungstyp* (Depletion-Typ). Wird der Kanal bei einer Spannung $V_{GS} \neq 0$ leitend, so spricht man vom *Anreicherungstyp* (Enhancement-Typ). In den weiteren Betrachtungen wird der Begriff MOSFET, soweit nicht anders vermerkt, für den MOSFET vom Anreicherungstyp verwendet.

4.2.2 Funktionsweise des MOSFETs

Liegt am Gate (Steuereingang) des MOSFETs keine Spannung an, so kann zwischen Source und Drain kein Strom fließen, da Source und Drain gegenüber dem Substrat in Sperrrichtung gepolt sind. Der Kanalbereich ist für diesen Fall an be-

weglichen Ladungsträgern verarmt. Wird eine hinreichend hohe Steuerspannung $V_{GS} > V_T$ an das Gate angelegt, so entsteht im Halbleiter an der Grenze zum Oxid eine Inversionsschicht aus Elektronen, die einen Stromfluss zwischen Source und Drain ermöglicht. Diese Inversionsschicht wird als *Inversionskanal* bezeichnet. Dabei ist die Größe des sog. *Drainstroms* I_D zwischen Drain und Source von der Höhe der Gate-Source-Spannung V_{GS} abhängig. In Abb. 4.13 ist schematisch der Querschnitt eines n-Kanal-MOSFETs mit ausgebildetem Inversionskanal bei einer Drain-Source-Spannung von $V_{DS} = 0$ (thermodynamisches Gleichgewicht im Inversionskanal) und das zugehörige Bändermodell der MOS-Struktur dargestellt.

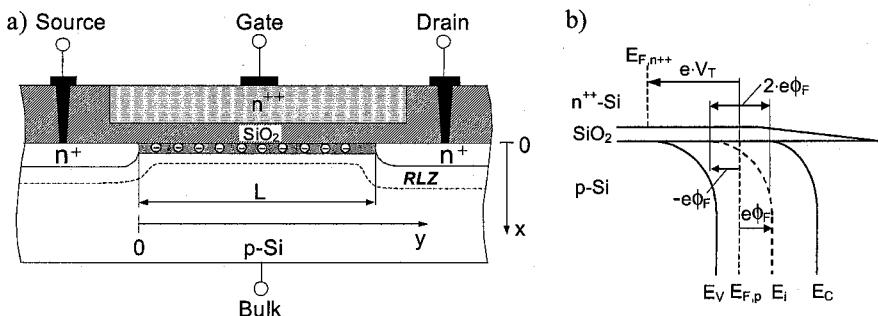


Abb. 4.13. a) Schematischer Querschnitt eines n-Kanal-MOSFETs mit ausgebildetem Inversionskanal und b) zugehöriges Bändermodell der MOS-Struktur im Inversionskanalbereich ($0 \leq y \leq L$)

Der Stromfluss im Inversionskanal wird im Folgenden für drei Fälle betrachtet. Im ersten Fall ist $V_{GS} > V_T$ und $V_{DS} > 0$, wobei V_{DS} einen geringen Wert besitzt. Für diesen Fall gilt dann $I_D \sim V_{DS} \cdot (V_{GS} - V_T)$ und der MOSFET arbeitet als spannungsgesteuerter Widerstand. Abb. 4.14 zeigt den entsprechenden Querschnitt mit ausgebildeter Raumladungszone und Inversionskanal.

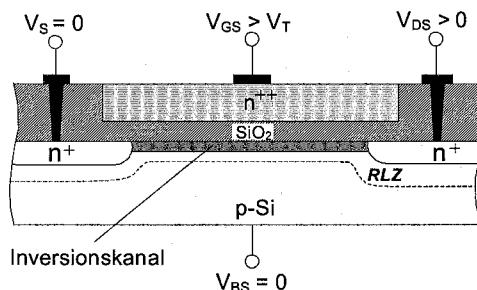


Abb. 4.14. Querschnitt eines n-Kanal-MOSFETs mit Raumladungszone und Inversionskanal mit $V_{GS} > V_T$ und $V_{DS} > 0$; MOSFET arbeitet als spannungsgesteuerter Widerstand

Im zweiten Fall ist die Drain-Source-Spannung V_{DS} gleich der Sättigungsspannung V_{DSat} ($V_{DS} = V_{DSat} = V_{GS} - V_T$). Der Drainstrom ist dann gleich dem Sätti-

gungsstrom ($I_D = I_{DSat}$) und es erfolgt ein Abschnüren (*Pinch-off*) des Inversionskanals am Drain. Der entstehende *Pinch-off-Punkt* ist in Abb. 4.15 veranschaulicht.

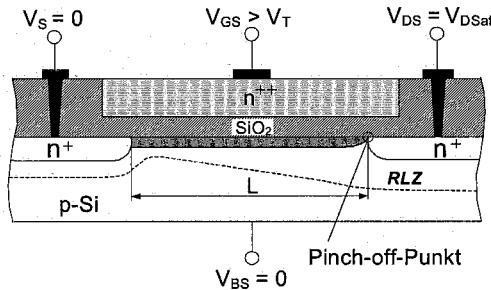


Abb. 4.15. Querschnitt eines n-Kanal-MOSFETs bei Kanalabschnürung (Pinch-off)

Im dritten Fall gilt, dass V_{DS} größer als die *Drain-Source-Sättigungsspannung* V_{DSat} ist, was ein Wandern des Pinch-off-Punktes in Richtung Source bewirkt. Damit erfolgt eine Kanalverkürzung um ΔL gemäß Abb. 4.16. Mit $\Delta L \ll L$ behält der Drainstrom I_D näherungsweise den Wert des *Drain-Sättigungsstroms* I_{DSat} .

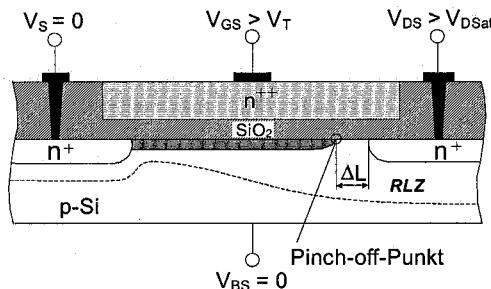


Abb. 4.16. Querschnitt eines n-Kanal-MOSFETs in Sättigung (Pinch-off) mit Kanallängenverkürzung ΔL bei $V_{DS} > V_{DSat}$

4.3 Strom-Spannungs-Charakteristik des MOSFETs

4.3.1 Theorie der Ladungssteuerung

Die Ladungssteuerung des MOSFETs wird anhand der *Parabelnäherung* (Gradual Channel Approximation) betrachtet. Dabei wird von der Annahme ausgegangen, dass die Sourcespannung den Wert $V_S = 0$ besitzt, keine Spannung am Substrat anliegt ($V_{BS} = 0$) und dass für die Gate-Source-Spannung $V_{GS} > V_T$ gilt. Als Randbedingungen für die Betrachtungen wird eine *Kanalspannung* V_{Kanal} vorgegeben, für die gilt: $V_{Kanal}(y = 0) = V_S = 0$ und $V_{Kanal}(y = L) = V_{DS}$ (vgl. Abb. 4.17).

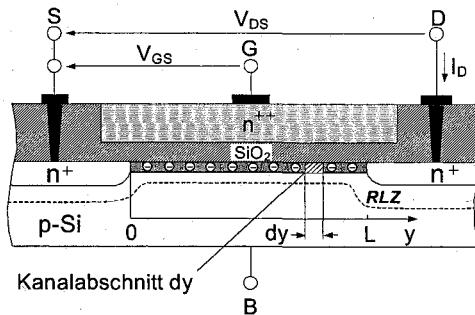


Abb. 4.17. Veranschaulichung der Kanalgeometrien eines n-Kanal-MOSFETs zur Bestimmung des Drainstroms

Zur Bestimmung des Drainstroms werden zwei Betriebsbereiche unterschieden. Für Drain-Source-Spannungen $V_{DS} \leq V_{GS} - V_{Th}$ arbeitet der MOSFET als linear spannungsgesteuerter Widerstand. Diesen Betriebsbereich nennt man *Triodenbereich* oder aktiven Bereich. Bei Drain-Source-Spannungen $V_{DS} \geq V_{GS} - V_{Th} > 0$ erfolgt eine Abschnürung des Inversionskanals (vgl. Abschn. 4.2.2). Diesen Betriebsbereich bezeichnet man als *Sättigungsbereich*.

Für den Triodenbereich ($V_{DS} \leq V_{GS} - V_{Th}$) erfolgt die Bestimmung des Drainstroms über die Flächenladungsdichte $Q_I(y)$ der beweglichen Ladung im Kanal:

$$Q_I(y) = -C_{ox} \cdot (V_{GS} - V_{Th} - V_{Kanal}(y)). \quad (4.21)$$

Des Weiteren wird zur Berechnung des Drainstroms die mittlere Geschwindigkeit $\bar{v}_n(y)$ der Elektronen im Kanalabschnitt dy herangezogen:

$$\bar{v}_n = \mu_n \cdot \frac{dV_{Kanal}(y)}{dy}. \quad (4.22)$$

Der Drainstrom I_D wird dann aus dem Produkt aus der Kanalweite W , der Flächenladungsdichte $Q_I(y)$ und der mittleren Geschwindigkeit der Elektronen $\bar{v}_n(y)$ wie folgt berechnet:

$$I_D = -W \cdot Q_I(y) \cdot \bar{v}_n(y). \quad (4.23)$$

Durch Integration dieser Gleichung über die Kanallänge L kann der rechte Term in eine Integration der Flächenladungsdichte $Q_I(V)$ über die Drain-Source-Spannung V_{DS} umgeformt werden:

$$\int_0^L I_D dy = -W \cdot \mu_n \int_0^{V_{DS}} Q_I(V) dV,$$

$$I_D \cdot L = W \cdot \mu_n \cdot C_{ox} \cdot \left((V_{GS} - V_{Th}) \cdot V_{DS} - \frac{1}{2} \cdot V_{DS}^2 \right). \quad (4.24)$$

Damit erhält man für die Kennliniengleichung bzw. für den Drainstrom I_D im Triodenbereich:

$$I_D = \mu_n \cdot C_{ox} \cdot \frac{W}{L} \cdot \left((V_{GS} - V_{Tn}) \cdot V_{DS} - \frac{1}{2} \cdot V_{DS}^2 \right). \quad (4.25)$$

An der Gültigkeitsgrenze $V_{DS} = V_{GS} - V_{Tn}$ tritt Stromsättigung auf, d. h. die Flächenladung $Q_f(L)$ geht gegen Null und die Spannungsänderung dV/dy bzw. die Spannung V geht bei Parabelnäherung gegen unendlich. Jenseits dieser Gültigkeitsgrenze arbeitet der MOSFET im Sättigungsbereich ($V_{DS} \geq V_{GS} - V_{Tn} > 0$); es gilt (Langkanalnäherung):

$$\frac{\partial I_D}{\partial V_{DS}} = 0. \quad (4.26)$$

Die Kennliniengleichung für den Sättigungsbereich lautet dann:

$$I_{DSat} = \frac{\mu_n \cdot C_{ox} \cdot W}{2} \cdot (V_{GS} - V_{Tn})^2. \quad (4.27)$$

Das *Ausgangskennlinienfeld* des MOSFETs (Abb. 4.18a) veranschaulicht die Lage des Trioden- und Sättigungsbereiches mit V_{GS} als Parameter. Die *Übertragungskennlinie* in Abb. 4.18b macht die Steigung dI_D/dV_{GS} sichtbar, die die Steuerwirkung der Gate-Source-Spannung auf den Drainstrom beschreibt.

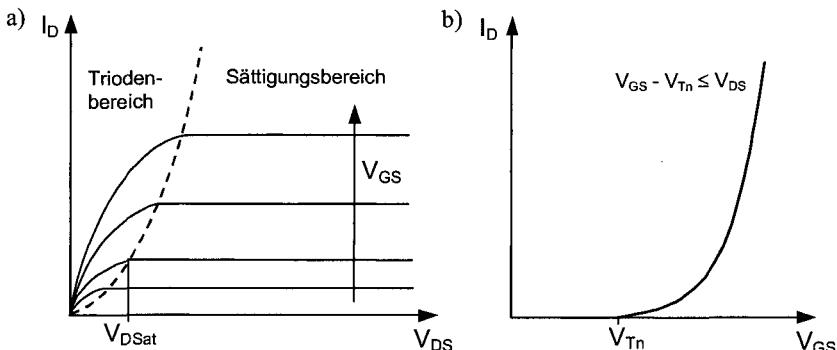


Abb. 4.18. a) Ausgangskennlinienfeld und b) Übertragungskennlinie eines n-Kanal-MOSFETs

Für einen p-Kanal-MOSFET berechnet sich der Drainstrom I_D für den Triodenbereich ($V_{GS} - V_{Tp} \leq V_{DS} < 0$) wie folgt:

$$I_D = -\mu_p \cdot C_{ox} \cdot \frac{W}{L} \cdot \left((V_{GS} - V_{Tp}) \cdot V_{DS} - \frac{1}{2} \cdot V_{DS}^2 \right). \quad (4.28)$$

Für den Sättigungsbereich des p-Kanal-MOSFETs ($0 > V_{GS} - V_{Tp} \geq V_{DS}$) erhält man die folgende Kennliniengleichung:

$$I_{DSat} = -\frac{\mu_p \cdot C_{ox}}{2} \cdot \frac{W}{L} \cdot (V_{GS} - V_{Tp})^2. \quad (4.29)$$

Typische Verhältnisse der Ladungsträgerbeweglichkeiten und Einsatzspannungen sind $\mu_p \approx \frac{1}{2} \cdot \mu_n$ und $V_{Tp} \approx -V_{Th}$.

Ein Vergleich der Ausgangs- und Übertragungskennlinien von n-Kanal- und p-Kanal-MOSFETs vom Anreicherungs- und Verarmungstyp sind in Abb. 4.19 dargestellt.

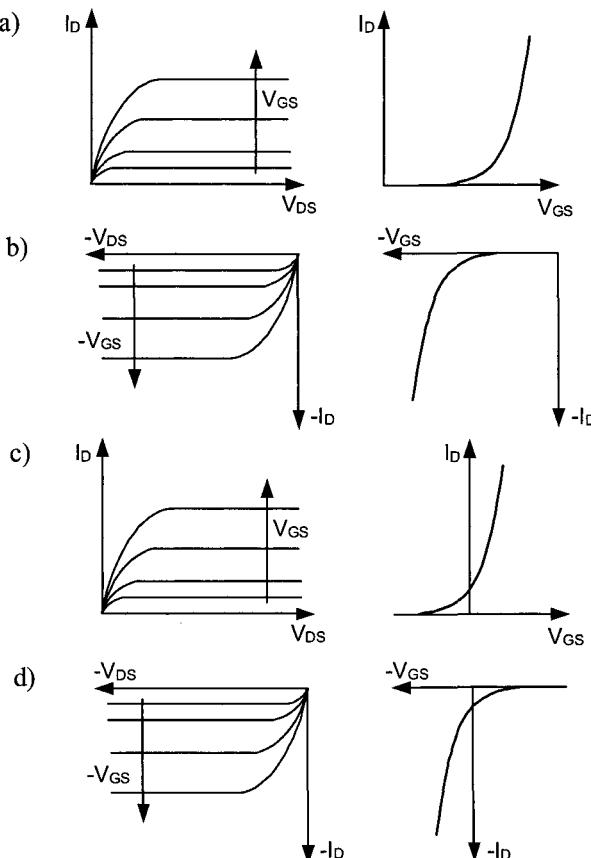


Abb. 4.19. Vergleich der Ausgangs- und Übertragungskennlinien für eine a) n-Kanal-MOSFET vom Anreicherungstyp b) p-Kanal-MOSFET vom Anreicherungstyp c) n-Kanal-MOSFET vom Verarmungstyp und d) p-Kanal-MOSFET vom Verarmungstyp

4.3.2 Einsatzspannung des MOSFETs

Analog zur Einsatzspannung V_T der MOS-Struktur bei starker Inversion (vgl. Abschn. 4.1.5) kann man die Einsatzspannung V_{Th0} eines n-Kanal-MOSFETs mit

Hilfe der Flachbandspannung V_{FB} , dem Fermipotential ϕ_F und einem Beitrag fester Ladungen Q_{SS} sowie der Gateoxidkapazität C_{ox} bestimmen:

$$V_{Tn0} = V_{FB} + 2\phi_F + \gamma_n \cdot \sqrt{2\phi_F - \frac{Q_{SS}}{C_{ox}}} \quad \text{mit } \gamma_n = \frac{\sqrt{2e \cdot N_A \cdot \epsilon_0 \cdot \epsilon_{Si}}}{C_{ox}}. \quad (4.30)$$

Dabei bezeichnet γ den *Body-Faktor*, der den Substrateinfluss des MOSFETs beschreibt (vgl. Abschn. 4.3.3). Für einen p-Kanal-MOSFET ergibt sich analog zum n-Kanal-MOSFET eine Einsatzspannung V_{Tp0} von:

$$V_{Tp0} = V_{FB} + 2\phi_F - \gamma_p \cdot \sqrt{2\phi_F - \frac{Q_{SS}}{C_{ox}}} \quad \text{mit } \gamma_p = \frac{\sqrt{2e \cdot N_D \cdot \epsilon_0 \cdot \epsilon_{Si}}}{C_{ox}}. \quad (4.31)$$

Die angegebenen Einsatzspannungen V_{Tn0} und V_{Tp0} des MOSFETs gelten für eine Substratspannung $V_{BS} = 0$.

4.3.3 Substratsteuerung des MOSFETs

Beim Wirken einer Substratspannung ($V_{BS} \neq 0$) entsteht eine Einsatzspannungsverschiebung des MOSFETs und damit auch eine Änderung des Drainstroms I_D . Man spricht von der Substratsteuerung oder vom *Body-Effekt*. Die Spannung über der Raumladungszone beträgt dann ($2\cdot\phi_F - V_{BS}$). Die Raumladungszone dehnt sich aus, was zu einer betragsmäßigen Erhöhung der Depletionladung Q_D führt. Die Einsatzspannung V_T des Feldeffekttransistors wird damit größer. Der Einfluss der Substratspannung auf den Bandverlauf in der Raumladungszone eines n-Kanal-MOSFETs ist in Abb. 4.20 im Bänderschema dargestellt.

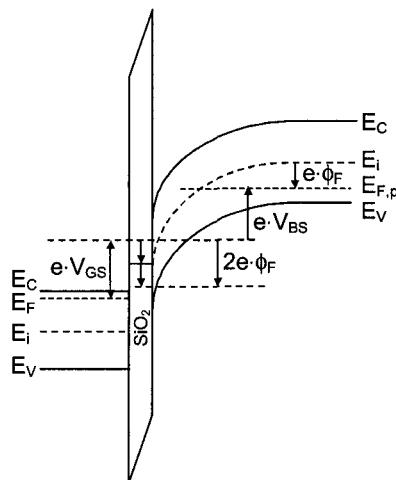


Abb. 4.20. Bänderschema eines n-Kanal-MOSFETs bei Einfluss einer Substratspannung $V_{BS} \neq 0$ (schematische Darstellung)

Die Einsatzspannung $V_{Tn}(V_{BS})$ bei Substratspannungseinfluss berechnet sich für einen n-Kanal-MOSFET mit Hilfe des Body-Faktors γ_n wie folgt:

$$V_{Tn}(V_{BS}) = V_{Tn0} + \gamma_n \cdot \left(\sqrt{2\phi_F - V_{BS}} - \sqrt{2\phi_F} \right). \quad (4.32)$$

Für einen p-Kanal-MOSFET folgt für die substratspannungsabhängige Einsatzspannung $V_{Tp}(V_{BS})$ analog:

$$V_{Tp}(V_{BS}) = V_{Tp0} - \gamma_p \cdot \left(\sqrt{-2\phi_F + V_{BS}} - \sqrt{-2\phi_F} \right). \quad (4.33)$$

Werden die substratspannungsabhängigen Einsatzspannungen in Gl. (4.25) und Gl. (4.26) bzw. in Gl. (4.28) und Gl. (4.29) eingesetzt, so erhält man die zugehörigen substratspannungsabhängigen Kanalströme $I_D(V_{BS})$ für den Trioden- und Sättigungsbereich.

4.3.4 Unterschwellstrom des MOSFETs

Für Bandverbiegungen im Potentialbereich $\phi_F \leq \phi_S < 2 \cdot \phi_F$ (schwache Inversion) besteht ein gering leitender Inversionskanal an der Grenzfläche zwischen Oxid und Halbleiter. Für die an dieser Grenzfläche vorherrschende Oberflächenladungsdichte n_s gilt:

$$n_s = n_{p0} \cdot \exp\left(\frac{\phi_S}{V_\vartheta}\right). \quad (4.34)$$

Daraus resultiert das Fließen eines *Unterschwellstroms* I_{DSub} auch bei Gate-Source-Spannungen unterhalb der Einsatzspannung ($V_{GS} < V_T$):

$$I_{DSub} \approx I_0 \cdot \exp\left(\frac{V_{GS} - V_T}{V_\vartheta}\right). \quad (4.35)$$

Der Unterschwellstrom nimmt bei kurzen Kanallängen zu (vgl. Abschn. 4.4) und wird zunehmend abhängig von der Drain-Source-Spannung V_{DS} .

4.3.5 Thermisches Verhalten des MOSFETs

Das thermische Verhalten eines MOSFETs wird vor allem durch die Temperaturabhängigkeit der Ladungsträgerbeweglichkeit bestimmt. Bei einer Temperaturerhöhung nimmt diese Beweglichkeit ab, wodurch der Drainstrom betragsmäßig ebenfalls abnimmt. Weiterhin kommt die Temperaturabhängigkeit der Einsatzspannung (insbesondere bei geringer Betriebsspannung) zum Tragen. Im wesentlichen wird die Temperaturabhängigkeit der Einsatzspannung durch die Temperaturabhängigkeit des Fermipotentials ϕ_F bestimmt:

$$|\phi_F| = \frac{kT}{e} \cdot \ln\left(\frac{N}{n_i(T)}\right) \sim \ln\left(\frac{1}{T^{3/2}}\right). \quad (4.36)$$

Bei einer Temperaturerhöhung sinkt $|\phi_F|$ ab, wodurch die Einsatzspannung $|V_T|$ des MOSFETs abnimmt. Dies hat zur Folge, dass der Drainstrom I_D betragsmäßig größer wird. Weiterhin steigt der Unterschwellstrom (vgl. Abschn. 4.3.4) exponentiell mit der Temperatur an.

4.4 Geometrieabhängigkeit von MOSFET-Parametern

Durch Verringerung der Geometrie des MOSFETs (W , L , t_{ox}) aus Gründen der Miniaturisierung integrierter elektronischer Bauelemente ergeben sich verschiedene Effekte, die Einfluss auf die Betriebsparameter des MOSFETs haben.

Die sog. *Kanallängenmodulation* beruht darauf, dass sich bei Drain-Source-Spannungen $V_{DS} > V_{DSat}$ der Pinch-off-Punkt (vgl. Abschn. 4.2.2) in Richtung Sourcegebiet bewegt. Dadurch wird die effektive Kanallänge des MOSFETs beeinflusst ($L - \Delta L$), was sich in einer Änderung des Drainstroms I_D widerspiegelt. Die effektive Kanallänge wird durch den *Kanallängenmodulationsfaktor* λ ausgedrückt und es gilt:

$$L - \Delta L = L \cdot (1 - \lambda \cdot V_{DS}) \text{ mit } \lambda \cdot V_{DS} \ll 1. \quad (4.37)$$

Für den Drainstrom I_D eines n-Kanal-MOSFETs unter Berücksichtigung der Kanallängenmodulation folgt dann:

$$I_D = \frac{\mu_n \cdot C_{ox}}{2} \cdot \frac{W}{L \cdot (1 - \lambda \cdot V_{DS})} \cdot (V_{GS} - V_T)^2. \quad (4.38)$$

Es resultiert eine Abweichung von der Ausgangskennlinie nach Parabelnäherung, die sich in einem Anstieg der Kennlinie im Sättigungsbereich äußert.

Einen weiteren Effekt stellt die von der Gate-Source-Spannung abhängige *Beweglichkeitsreduktion* $\mu_s(V_{GS})$ der Ladungsträger im Inversionskanal des MOSFETs dar. Diese Reduktion erfolgt auf Grund steigender Grenzflächenstreuungen bei hohen Feldern z. B. durch Grenzflächenphononen, Grenzflächenrauhigkeit oder Coulomb-Streuung an Störstellen. Die Übertragungskennlinie eines MOSFETs bei Berücksichtigung des Einflusses von $\mu_s(V_{GS})$ ist in Abb. 4.21 dargestellt.

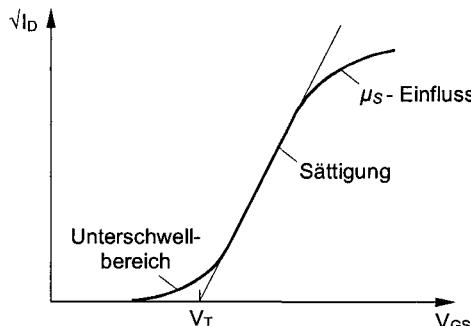


Abb. 4.21. Übertragungskennlinie eines n-Kanal-MOSFETs mit Einfluss von $\mu_s(V_{GS})$

Weiterhin kann bei Geometrieverkleinerung eine *Geschwindigkeitssättigung* bei hohem lateralem Feld auftreten. Dies kann bei kurzem Kanal und hoher Drain-Source-Spannung schon vor Erreichen der Pinch-off-Bedingung ($V_{DS} > V_{GS} - V_T$) erfolgen. Die Geschwindigkeitssättigung kann durch eine Verringerung der effektiven Beweglichkeit μ_{eff} der Ladungsträger im Inversionskanal beschrieben werden. Sie hängt von der Sättigungsfeldstärke E_{sat} , der Kanallänge L und der Drain-Source-Spannung V_{DS} sowie von der Beweglichkeitsreduktion μ_s ab:

$$\mu_{eff} = \frac{\mu_s}{1 + \frac{V_{DS}}{E_{sat} \cdot L}}. \quad (4.39)$$

Außerdem ist die Abhängigkeit der Einsatzspannung von der Kanallänge zu betrachten. Von einem kurzen Kanal spricht man, wenn die Kanallänge in der Größenordnung der Raumladungszonenweite ℓ_D oder der Tiefe des Drain- bzw. Sourcegebietes x_J liegt (vgl. Abb. 4.22).

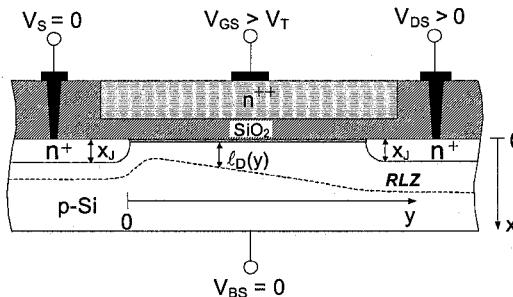


Abb. 4.22. Kanalgeometrie eines n-Kanal-MOSFETs mit Einfluss der Raumladungszonen von Drain und Source

Die Ladungen in der Verarmungszone in der Nähe des Drain- und Sourcegebietes werden von den Drain- und Sourcekontakten und dem Gate beeinflusst. In einem typischen Kurzkanaltransistor bilden diese gemeinsam kontrollierten Ladungen einen relativ großen Anteil der Gesamtladung in der Verarmungszone unter dem Gate. Es verringert sich die Einsatzspannung des MOSFETs. Eine weitere Reduktion der Einsatzspannung erfolgt durch die Ausdehnung der vom Drain gesteuerten Verarmungszone in der Nähe des Draingegebietes bei hohen Drain-Source-Spannungen V_{DS} . Darüber hinaus wächst der Unterschwellstrom des MOSFETs (vgl. Abschn. 4.3.4) bei kurzen Kanallängen an und wird zunehmend abhängig von der Drain-Source-Spannung.

4.5 Durchbruchsmechanismen und Degradation

Um bei kurzer Kanallänge die Steuerwirkung der Gateelektrode aufrecht zu erhalten, kann die Oxiddicke t_{ox} des MOSFETs reduziert werden. Dies führt zu hohen

vertikalen Feldstärken E_{\perp} . Zusätzlich tritt am Ende des Draingegebietes im Pinch-off-Bereich eine hohe laterale Feldstärke $E_{\parallel}(\Delta L)$ auf. Die entsprechenden Feldstärkekomponenten sind in Abb. 4.23 veranschaulicht.

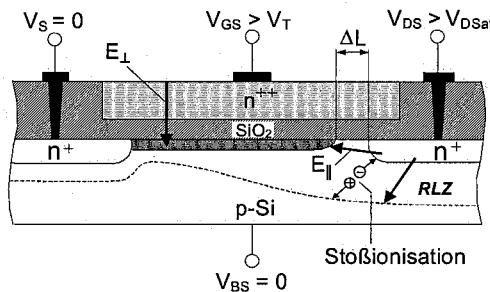


Abb. 4.23. Feldstärkekomponenten E_{\perp} und E_{\parallel} am Ende des Draingegebietes eines n-Kanal-MOSFETs

Die auftretenden, hohen Feldstärken können zwei Arten von Durchbruchmechanismen hervorrufen, den *Lawinen-Durchbruch* und den *Gate-Durchbruch*. Der Lawinendurchbruch (beim MOSFET auch *Draindurchbruch* genannt) tritt nahe des Draingegebietes im Pinch-off-Bereich auf, wenn die Spannung V_{DK} zwischen Draingegebiet und Kanal-Ende den Wert

$$V_{DK} = V_{DS} - (V_{GS} - V_T) \quad (4.38)$$

annimmt und die kritische Feldstärke ($E_M \approx 3 \cdot 10^5 \text{ V/cm}$) für den Lawineneffekt überschritten wird. Zugleich tritt ein hoher Substratstrom auf.

Der Gate-Durchbruch des Gateoxids erfolgt bei hoher Gate-Source-Spannung V_{GS} oder hoher Gate-Drain-Spannung V_{GD} . Die Durchbruchfeldstärke beträgt etwa $E_{Br} \approx 10^7 \text{ V/cm}$. Sie kann im Anschlussbereich von integrierten Schaltungen auch durch elektrostatische Entladung (ESD) hervorgerufen werden. Schaltungstechnisch kann diesem Effekt durch Schutzschaltungen vorgebeugt werden. Die in Abb. 4.24 beispielhaft dargestellte Schutzschaltung zwischen Anschluss-Pad und Gatekontakt eines MOSFETs T_1 erfolgt durch Beschaltung mit den Schutzdiode D_1 und D_2 sowie einem Vorwiderstand R .

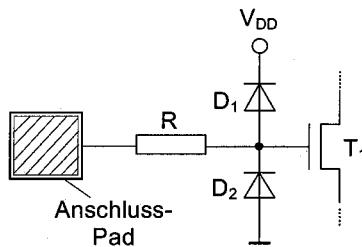


Abb. 4.24. Schutzschaltung zur Vermeidung eines ESD-bedingten Gatedurchbruchs eines MOSFETs

Neben den genannten Durchbruchsmechanismen kann bei kurzem Kanal und hoher Drain-Source-Spannung eine Berührung der Raumladungszonen des pn-Überganges zwischen Drain und Substrat sowie zwischen Source und Substrat auftreten (Punch-Through-Effekt, vgl. Abschn. 3.3.3). In diesem Fall wird die Abhängigkeit des Drainstroms von der Gate-Source-Spannung aufgehoben.

Durch die hohe Beschleunigung der Ladungsträger im starken elektrischen Feld des Pinch-off-Bereiches (vgl. Abb. 4.23) entstehen sog. *heiße Ladungsträger*, die die Barriere zwischen Halbleiter und Oxid überwinden können. Es resultiert das Fließen eines Gatestroms und durch Eindringen von Ladungsträgern in Störstellen im Oxid kann eine Drift der Einsatzspannung V_T des MOSFETs erfolgen. Dieser Mechanismus wird *Degradation* genannt.

4.6 Technologische Realisierung

In Abb. 4.25 sind die typischen Prozessschritte zur technologischen Realisierung eines n-Kanal-MOSFETs dargestellt.

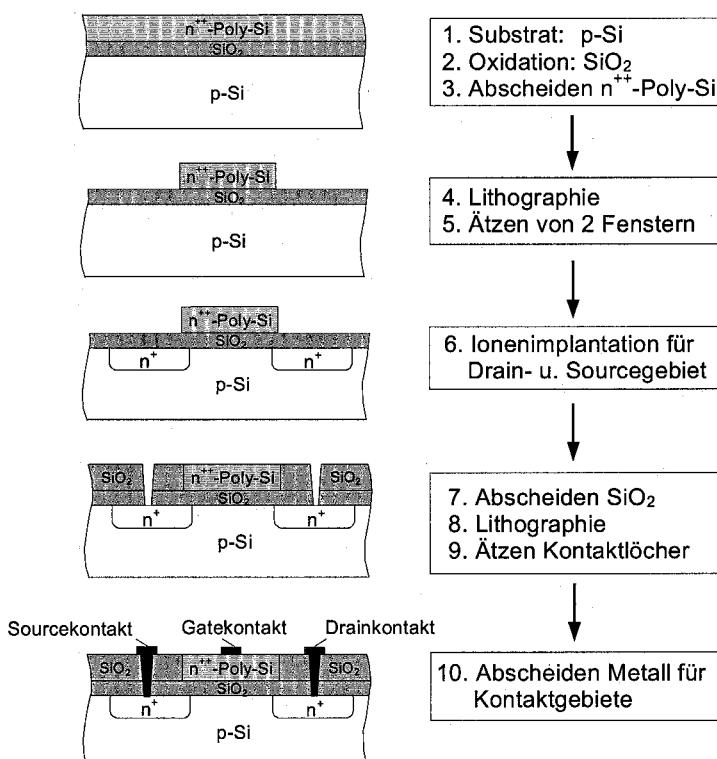


Abb. 4.25. Grundlegende Prozessschritte zur technologischen Realisierung eines n-Kanal-MOSFETs

Der Ausgangspunkt zur Herstellung eines n-Kanal-Feldeffekttransistors ist ein gering dotiertes p-Silizium-Substrat, auf dem durch thermische Oxidation eine SiO₂-Isolatorschicht aufgebracht wird. Anschließend erfolgt das Abscheiden einer Schicht aus n⁺⁺-Poly-Silizium, aus der später das Gate des Feldeffekttransistors entsteht. Im nächsten Prozessschritt (Lithographie) wird Fotolack aufgetragen und durch eine Maske mit UV-Licht bestrahlt (vgl. Abschn. 2.4). Durch das bestrahlte Gebiet werden zwei Fenster in die n⁺⁺-Poly-Silizium-Schicht geätzt und der restliche Fotolack entfernt. Durch Ionenimplantation erfolgt dann die Dotierung des Source- und Draingegebietes (n⁺-Si); an der Oberfläche der Struktur wird anschließend eine SiO₂-Schicht abgeschieden. Nach einem weiteren Lithographieschritt werden die Kontaktlöcher für Drain und Source geätzt und im nächsten Prozessschritt die Metallisierung des Source-, Drain-, und Gategebietes durch Metallabscheidung (z. B. Aluminium) durchgeführt. Neben den hier beschriebenen grundlegenden Technologieschritten werden weitere Schritte z. B. zur Einstellung der Einsatzspannung (vgl. Abschn. 4.1.5 und Abschn. 4.3.2) durchgeführt.

Die Herstellung eines p-Kanal-MOSFETs erfolgt analog, jedoch wird die MOSFET-Struktur in eine sog. *n-Wanne* (n-Si-Bulkgebiet im p-Si-Substrat) eingebettet. Man spricht hierbei von einem *n-Wannen-Prozess*.

4.7 Netzwerkmodelle des MOSFETs

4.7.1 MOSFET-Kapazitäten

Ausgehend von den Kapazitäten an der MOS-Struktur (vgl. Abschn. 4.1.6) können unter zusätzlicher Berücksichtigung des Source- und Draingegebietes die Kapazitäten des MOSFETs bestimmt werden. In Abb. 4.26 sind die wesentlichen Kapazitäten eines MOSFETs für den Inversionsfall im Querschnitt dargestellt.

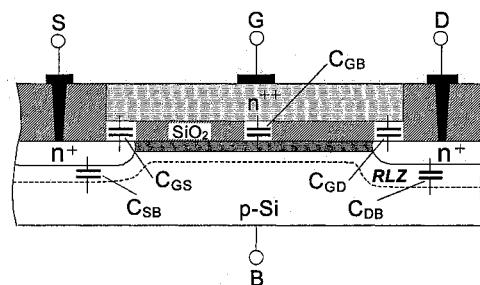


Abb. 4.26. Kapazitäten eines n-Kanal-MOSFETs in Inversion im Querschnitt

Die *Oxidkapazitäten* C_{GB} , C_{GS} und C_{GD} resultieren zum einen aus der Wechselwirkung zwischen Gatespannung und Kanalladung (spannungsabhängig) und zum anderen aus der Überlappung der Gateelektrode mit dem Source- und Draingegebiet (nahezu spannungsunabhängig). Die spannungsabhängigen Kapazitätsanteile der Oxidkapazitäten entstehen durch unterschiedliche Ladungsverteilungen in den Be-

triebsbereichen Akkumulation, Depletion und Inversion (vgl. Abschn. 4.1.6). Weiterhin bilden sich durch die pn-Übergänge zwischen Source und Bulk sowie zwischen Drain und Bulk die spannungsabhängigen *Sperrsichtkapazitäten* C_{SB} und C_{DB} aus. Die aus der Überlappung von Schichten resultierenden Kapazitätsanteile der Oxidkapazitäten, sowie die Sperrsichtkapazitäten C_{SB} und C_{DB} nennt man *extrinsische* Kapazitäten. Die auf den Kanalbereich des MOSFETs bezogenen Kapazitätsanteile werden als *intrinsische* Kapazitäten bezeichnet. Die Kapazitäts-Spannungs-Charakteristik eines MOSFETs ist in Abschn. 4.1.6, Abb. 4.9 (Verlauf für NF-Verhalten) dargestellt.

Die Kapazitäten eines p-Kanal-MOSFETs werden analog ermittelt. Jedoch ist für die Bestimmung der Spannungsabhängigkeit der Sperrsichtkapazitäten von einer umgekehrten Polarität der pn-Übergänge auszugehen. Weiterhin wirkt bei Verwendung eines n-Wannen-Prozesses (vgl. Abschn. 4.6) eine zusätzliche extrinsische Kapazität zwischen Substrat- und Bulkgebiet (*Wannenkapazität*), die ebenfalls spannungsabhängig ist.

4.7.2 Großsignalersatzschaltbild

Mit den in Abschn. 4.7.1 genannten Kapazitäten kann das Großsignalersatzschaltbild des MOSFETs in Abb. 4.27 aufgestellt werden.

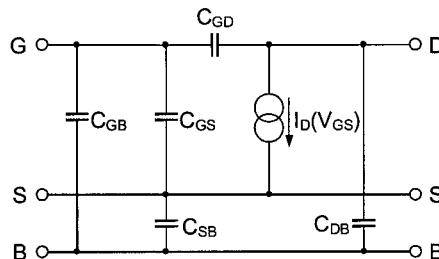


Abb. 4.27. Großsignalersatzschaltbild eines n-Kanal-MOSFETs

Neben den Oxid- und Sperrsichtkapazitäten enthält das Großsignalersatzschaltbild eine spannungsgesteuerte Stromquelle, die den Steuereffekt des Drainstroms I_D in Abhängigkeit von der Gate-Source-Spannung V_{GS} darstellt. Für den Drainstrom gelten zusammenfassend folgende Abhängigkeiten (vgl. Abschn. 4.3):

$$V_{GS} < V_{Tn} : \quad I_D \approx I_0 \cdot \exp\left(\frac{V_{GS} - V_{Tn}}{V_\vartheta}\right) \quad (4.41)$$

$$V_{GS} - V_{Tn} > V_{DS} : \quad I_D = \mu_n \cdot C_{ox} \cdot \frac{W}{L} \cdot \left((V_{GS} - V_{Tn}) \cdot V_{DS} - \frac{1}{2} \cdot V_{DS}^2 \right) \quad (4.42)$$

$$V_{DS} \geq V_{GS} - V_{Tn} > 0: \quad I_D = \frac{\mu_n \cdot C_{ox}}{2} \cdot \frac{W}{L \cdot (1 - \lambda \cdot V_{DS})} \cdot (V_{GS} - V_{Tn})^2 \quad (4.43)$$

4.7.3 Kleinsignalersatzschaltbild

Im Kleinsignalersatzschaltbild wird die Steuerwirkung der Gate-Source-Spannung V_{GS} durch den *Ausgangsleitwert* g_{DS} und die *Steilheit* g_m des MOSFETs beschrieben. Das Kleinsignalersatzschaltbild eines n-Kanal-MOSFETs (ohne Substratsteuereinfluss: $V_{BS} = 0$) ist in Abb. 4.28 dargestellt.

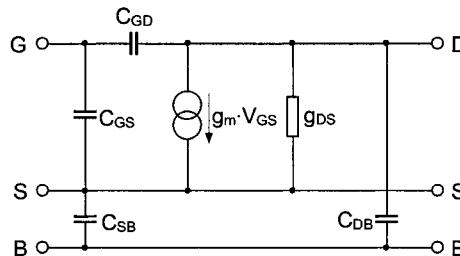


Abb. 4.28. Kleinsignalersatzschaltbild eines n-Kanal-MOSFETs

Die Steilheit g_m kennzeichnet die Steigung der Übertragungskennlinie des MOSFETs (vgl. Abschn. 4.3.1) und gibt damit die Steuerwirkung der Gate-Source-Spannung V_{GS} an. Die Steilheit wird über die partielle Ableitung des Drainstroms I_D nach der Gate-Source-Spannung V_{GS} bei konstanter Drain-Source-Spannung V_{DS} und konstanter Substratspannung V_{BS} berechnet.

$$g_m = \left. \frac{\partial I_D}{\partial V_{GS}} \right|_{V_{DS}=\text{konst.}, V_{BS}=\text{konst.}} \quad (4.44)$$

Wird der n-Kanal-MOSFET im Triodenbereich ($V_{DS} \leq V_{GS} - V_{Th}$) betrieben, so gilt für die Steilheit g_m :

$$g_m = \mu_n \cdot C_{ox} \cdot \frac{W}{L} \cdot V_{DS} \quad (4.45)$$

Für den Sättigungsbereich ($V_{DS} \geq V_{GS} - V_{Th} > 0$) berechnet sich die Steilheit hingegen wie folgt:

$$g_m = \mu_n \cdot C_{ox} \cdot \frac{W}{L} \cdot (V_{GS} - V_{Th}) \cdot (1 + \lambda \cdot V_{DS}). \quad (4.46)$$

Der Ausgangsleitwert (auch Kleinsignaladmittanz) g_{DS} wird über die Änderung des Drainstroms I_D nach der Drain-Source-Spannung V_{DS} bei konstanter Gate-Source-Spannung V_{GS} und konstanter Substratspannung V_{BS} berechnet. Der Ausgangsleitwert entspricht dem Reziproken des Drain-Source-Widerstandes r_{DS} :

$$g_{DS} = \left. \frac{\partial I_D}{\partial V_{DS}} \right|_{V_{GS}=\text{konst.}, V_{BS}=\text{konst.}} = \frac{I}{r_{DS}}. \quad (4.47)$$

Abhängig davon, ob der MOSFET im Trioden- oder Sättigungsbereich betrieben wird, ergeben sich für den Ausgangsleitwert die folgenden Beziehungen:

$$V_{DS} \leq V_{GS} - V_{Th}: \quad g_{DS} = \mu_n \cdot C_{ox} \cdot \frac{W}{L} \cdot ((V_{GS} - V_{Th}) - V_{DS}), \quad (4.48)$$

$$V_{DS} \geq V_{GS} - V_{Th} > 0: \quad g_{DS} \approx \lambda \cdot \frac{\mu_n \cdot C_{ox}}{2} \cdot \frac{W}{L} (V_{GS} - V_{Th})^2 = \lambda \cdot I_{DSat}. \quad (4.49)$$

Bei Vernachlässigung der Kanallängenmodulation (Parabelnäherung) erhält man für den Ausgangsleitwert im Sättigungsbereich $g_{DS} = 0$.

Um die entsprechenden Kleinsignalparameter für einen p-Kanal-MOSFET zu bestimmen, ist die Elektronenbeweglichkeit μ_n in Gl. (4.45) und Gl. (4.46) sowie in Gl. (4.48) und Gl. (4.49) durch die Löcherbeweglichkeit μ_p zu ersetzen.

4.8 MOSFET-Grundschaltungen

4.8.1 Einfache Verstärkerschaltungen

Analog zu den grundlegenden Verstärkerschaltungen des Bipolartransistors (vgl. Abschn. 3.6) unterscheidet man beim MOSFET die Source-, Drain- und Gate-Schaltung. Die Source-Schaltung wird hier auf Grund ihrer schaltungstechnischen Bedeutung ausführlicher beschrieben.

4.8.1.1 Source-Schaltung

Die Source-Schaltung zeichnet sich durch eine hohe Spannungsverstärkung und einen hohen Eingangswiderstand aus. In Abb. 4.29 ist die Source-Schaltung eines n-Kanal-MOSFETs T_1 mit ohmscher Last R_L und kapazitiver Last C_L . Die Eingangsspannung V_{in} ist z. B. eine Wechselspannung, die mit einem Gleichspannungsanteil V_{DC} beaufschlagt wird.

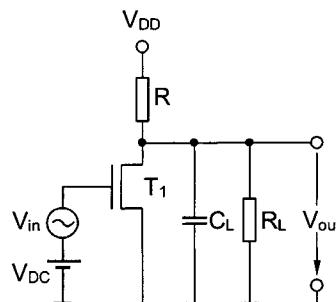


Abb. 4.29. Schaltungsprinzip eines n-Kanal-MOSFETs in Source-Schaltung mit ohmscher und kapazitiver Last

Aus dem Schaltungsprinzip in Abb. 4.29 resultiert das vereinfachte Kleinsignalersatzschaltbild, das in Abb. 4.30 dargestellt ist.

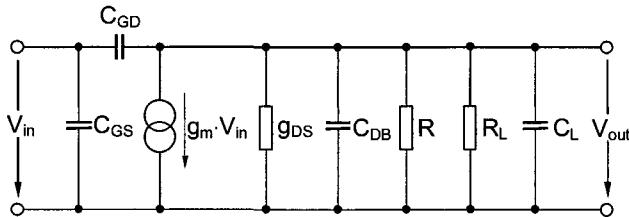


Abb. 4.30. Vereinfachtes Kleinsignalersatzschaltbild eines n-Kanal-MOSFETs in Source-Schaltung mit ohmscher und kapazitiver Last

Die Spannungsverstärkung A_V eines n-Kanal-MOSFETs in Source-Schaltung ohne Substratsteuerung ($V_{BS} = 0$) berechnet sich aus dem Verhältnis der Ausgangsspannungsänderung ΔV_{out} zur Eingangsspannungsänderung ΔV_{in} :

$$A_V = \frac{\Delta V_{out}}{\Delta V_{in}} = \frac{-g_m \cdot (R // R_L) + j\omega C_{GD} \cdot (R // R_L)}{1 + g_{DS} \cdot (R // R_L) + j\omega(R // R_L) \cdot (C_L + C_{DB} + C_{GD})}. \quad (4.50)$$

Mit der Näherung $C_{GD} \approx 0$ (Pinch-off, vgl. Abschn. 4.2.2) und unter Vernachlässigung des Ausgangswiderstandes r_{DS} des MOSFETs ($r_{DS} \gg R_L \parallel 1/j\omega C_L$) erhält man vereinfachend für die Spannungsverstärkung A_V der Source-Schaltung folgenden Zusammenhang:

$$A_V \approx \frac{-g_m(R // R_L)}{1 + j\omega(R // R_L) \cdot (C_L + C_{DB})}. \quad (4.51)$$

Für niedrige Frequenzen ($\omega \rightarrow 0$) ergibt sich für die Spannungsverstärkung A_{VSO} vereinfachend:

$$A_{V0} = -g_m \cdot (R // R_L). \quad (4.52)$$

Die Transitfrequenz f_T der Source-Schaltung hängt von der Steilheit g_m , der Lastkapazität C_L und der Drain-Bulk-Kapazität C_{DB} ab. Sie gibt die Frequenz an, bei der die Spannungsverstärkung den Wert Eins annimmt ($|A_V(f_T)| = 1$) und berechnet sich wie folgt:

$$f_T = \frac{g_m}{2\pi \cdot (C_L + C_{DB})}. \quad (4.53)$$

Zur Bestimmung der Stromverstärkung der Source-Schaltung ist die Berechnung der Eingangskapazität C_{in} notwendig, die aus der Gate-Source- und der Gate-Drain-Kapazität sowie der Spannungsverstärkung resultiert:

$$C_{in} = C_{GS} + C_{GD} \cdot (1 - A_v). \quad (4.54)$$

Damit erhält man für die Stromverstärkung A_I der Source-Schaltung:

$$A_I = \frac{g_m}{j\omega C_{in}}. \quad (4.55)$$

Der Term $C_{DG} \cdot (1 - A_V)$ in Gl. (4.54) wird auch als *Miller-Kapazität* bezeichnet. Durch die Miller-Kapazität wird die Ausgangsspannung V_{out} auf den Eingang der Source-Schaltung zurückgekoppelt, wodurch sich bei höheren Frequenzen der Eingangswiderstand r_{in} verringert. Bei niedrigen Frequenzen ($\omega \rightarrow 0$) besitzt die Source-Schaltung einen sehr hohen Eingangswiderstand, der sich wie folgt berechnet:

$$r_{in} = j\omega C_{in}. \quad (4.56)$$

4.8.1.2 Drain- und Gate-Schaltung

Die Prinzipschaltung eines n-Kanal-MOSFETs in Drain- und Gate-Schaltung ist in Abb. 4.31 dargestellt.

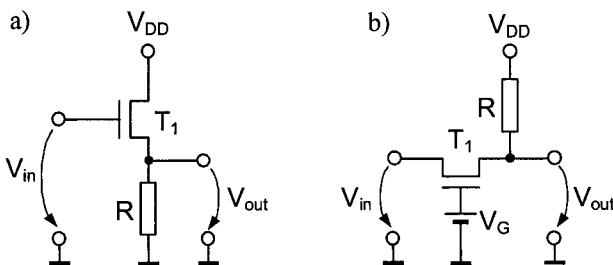


Abb. 4.31. n-Kanal-MOSFET in a) Drain-Schaltung und b) Gate-Schaltung

Die Drain-Schaltung besitzt eine Spannungsverstärkung von $A_V \approx 1$ und einen sehr hohen Eingangswiderstand r_{in} bei niedrigem Ausgangswiderstand r_{out} . Man nutzt die Drain-Schaltung daher als *Impedanzwandler*.

Die Gate-Schaltung besitzt die gleiche Spannungsverstärkung wie die Source-Schaltung, jedoch eine Stromverstärkung von $A_I = 1$ und eine geringe Spannungsrückwirkung vom Ausgang auf den Eingang. Auf Grund des geringen Eingangswiderstandes kommt die Gate-Schaltung im wesentlichen bei Hochfrequenz-Anwendungen und als *Kaskodestufe* nach einer Source-Schaltung (vgl. Abschn. 4.8.4) zum Einsatz.

4.8.2 CMOS-Inverter

4.8.2.1 Statisches Übertragungsverhalten

Der *CMOS* (Complementary MOS)-Inverter nutzt die komplementären Eigenschaften des n-Kanal- und p-Kanal-MOSFETs, um ein zum Eingangspegel (Low

oder High, vgl. Abschn. 7.1) invertiertes Ausgangssignal V_{out} zu erzeugen. In Abb. 4.32 sind das Schaltungsprinzip und das Schaltsymbol eines CMOS-Inverters dargestellt.

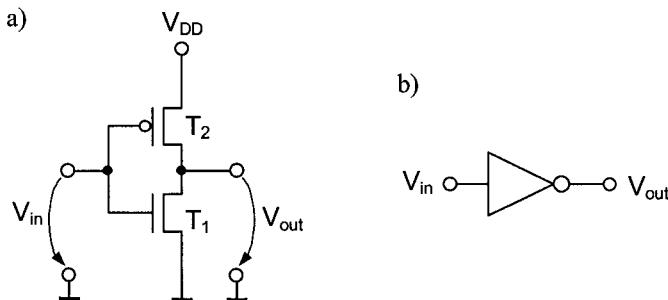


Abb. 4.32. CMOS-Inverter a) Prinzipschaltung und b) Schaltsymbol

Die Schaltschwelle des CMOS-Inverters liegt bei $V_{DD}/2$, wenn folgende Beziehungen gelten:

$$\mu_n \cdot C_{ox} \cdot \frac{W_n}{L_n} = \mu_p \cdot C_{ox} \cdot \frac{W_p}{L_p} \quad \text{und} \quad |V_{Tn}| = |V_{Tp}|. \quad (4.57)$$

Für ein symmetrisches Verhalten der Anstiegs- und Abfallzeit der Flanken des Ausgangssignals V_{out} werden die Weiten der Transistoren auf $W_p \approx 2 \cdot W_n$ dimensioniert, da die Beweglichkeiten der Ladungsträger näherungsweise $\mu_n \approx 2 \cdot \mu_p$ betragen.

Die Übertragungskennlinie und Arbeitsbereiche eines unbelasteten CMOS-Inverters sind in Abb. 4.33 dargestellt.

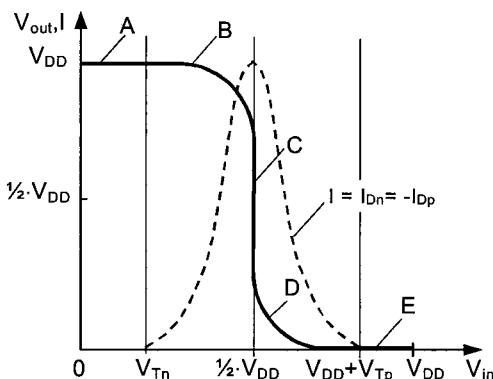


Abb. 4.33. Übertragungskennlinie und Arbeitsbereiche eines CMOS-Inverters

Die Wirkungsweise des CMOS-Inverters lässt sich anhand der Übertragungskennlinie Abb. 4.33 für die Bereiche A bis E wie folgt beschreiben: Für Spannung-

gen $0 \leq V_{in} \leq V_{Tn}$ (Bereich A) arbeitet der n-Kanal-MOSFET T_1 im Unterschwellbereich (vgl. Abschn. 4.3.4) und der p-Kanal-MOSFET T_2 im Triodenbereich. Die Ausgangsspannung besitzt im Bereich A den Wert der Betriebsspannung ($V_{out} = V_{DD}$).

Im Bereich B, d. h. für Eingangsspannungen im Bereich $V_{Tn} \leq V_{in} \leq \frac{1}{2} \cdot V_{DD}$, befindet sich T_1 in Sättigung und T_2 im Triodenbereich. Die Ausgangsspannung ist dann abhängig von der Eingangsspannung.

$$V_{out} = f(V_{in}), \frac{V_{DD}}{2} + |V_{Tp}| < V_{out} < V_{DD} \quad (4.58)$$

Bei einer Eingangsspannung von $V_{in} = \frac{1}{2} \cdot V_{DD}$ befinden sich beide Transistoren in Sättigung (Bereich C) und die Ausgangsspannung ist bei Parabelnäherung unabhängig von der Eingangsspannung:

$$V_{out} \neq f(V_{in}), \frac{V_{DD}}{2} - V_{Tn} < V_{out} < \frac{V_{DD}}{2} + |V_{Tp}|. \quad (4.59)$$

Der Bereich D der Übertragungskennlinie des CMOS-Inverters gilt für Eingangsspannungen im Bereich $\frac{1}{2} \cdot V_{DD} < V_{in} < V_{DD} - |V_{Tp}|$. Dann befindet sich der MOSFET T_1 im Triodenbereich und MOSFET T_2 arbeitet in Sättigung. Die Ausgangsspannung ist dann abhängig von der Eingangsspannung:

$$V_{out} = f(V_{in}), 0 < V_{out} < \frac{V_{DD}}{2} - V_{Tn}. \quad (4.60)$$

Im Bereich E ($V_{DD} - |V_{Tp}| \leq V_{in} \leq V_{DD}$) befindet sich der MOSFET T_1 im Triodenbereich und MOSFET T_2 im Unterschwellbereich. Die Ausgangsspannung besitzt dann den Wert $V_{out} = 0$.

4.8.2.2 Schaltverhalten des CMOS-Inverters

Das Schaltverhalten des CMOS-Inverters kann anhand der in Abb. 4.34 dargestellten Schaltungsanordnung aus mehreren Invertern mit kapazitiven Lasten C_L beschrieben werden.

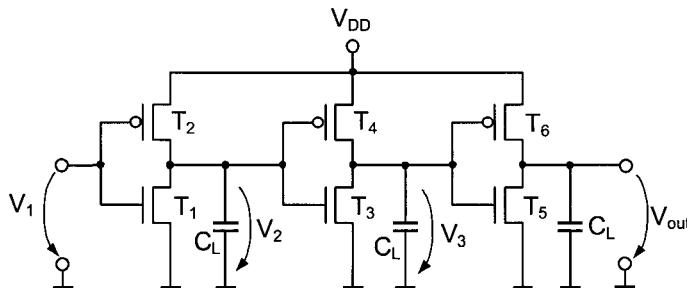


Abb. 4.34. Schaltungsanordnung zur Bestimmung des Schaltverhaltens von CMOS-Invertern mit kapazitiven Lasten C_L

Während der Schaltvorgänge treten Verluste z. B. durch das Umladen der kapazitiven Last C_L weiterer CMOS-Inverter auf. Die mittlere Verlustleistung P_V ist proportional zur Schaltfrequenz f und berechnet sich wie folgt:

$$\boxed{P_V = \frac{I}{2} \cdot f \cdot C_L \cdot V_{DD}^2} \quad (4.61)$$

Das Schaltverhalten kann durch Anlegen eines Spannungssprungs am Eingang des Inverters bestimmt werden. Charakteristische Größe für das Schaltverhalten ist die Verzögerungszeit τ_d , die die Zeitdifferenz zwischen dem 50%-Wert des Eingangssignals und dem 50%-Wert des Ausgangssignals des Inverters darstellt. Bei einem positiven Eingangsspannungssprung V_I am Invertereingang wird die Lastkapazität C_L zu Beginn mit dem maximalen Drain-Sättigungsstrom $I_{DSat,n}$ des n-Kanal-MOSFETs aufgeladen. Die Verzögerungszeit $\tau_{df,Sprung}$ der fallenden Flanke des Ausgangssignals berechnet sich dann wie folgt:

$$\tau_{df,Sprung} \approx \frac{C_L}{2} \cdot \frac{V_{DD}}{I_{DSat,n}} \text{ mit } I_{DSat,n} = \frac{\mu_n \cdot C_{ox}}{2} \cdot \frac{W}{L} \cdot (V_{DD} - V_{Tn})^2. \quad (4.62)$$

Bei einem negativen Spannungssprung V_I am Invertereingang wird die Lastkapazität C_L zu Beginn mit dem maximalen Drain-Sättigungsstrom $I_{DSat,n}$ des p-Kanal-MOSFETs aufgeladen. Die resultierende Verzögerungszeit $\tau_{dr,Sprung}$ der nun steigenden Flanke des Ausgangssignals berechnet sich über die folgende Beziehung:

$$\tau_{dr,Sprung} \approx \frac{C_L}{2} \cdot \frac{V_{DD}}{|I_{DSat,p}|} \text{ mit } I_{DSat,p} = -\frac{\mu_p \cdot C_{ox}}{2} \cdot \frac{W}{L} \cdot (V_{DD} + V_{Tp})^2. \quad (4.63)$$

Bei einer Spannungsrampe V_2 am Invertereingang (vgl. Abb. 4.34, zweite Inverterstufe) liegt an den Gates der MOSFETs nicht sofort die Betriebsspannung oder Masse an, sodass die Lastkapazität C_L nicht mit dem maximalen Drain-Sättigungsstrom umgeladen werden kann. Daraus resultiert eine größere Verzögerungszeit als bei einem Eingangsspannungssprung. Die Verzögerungszeit τ_{dr} der steigenden Flanke und die Verzögerungszeit τ_{df} der fallenden Flanke des Ausgangssignals können näherungsweise mit Hilfe der technologieabhängigen *Fitting-Parameter* k_r und k_f bestimmt werden:

$$\tau_{df} \approx k_f \cdot C_L \cdot \frac{V_{DD}}{I_{DSat,n}} \text{ und } \tau_{dr} \approx k_r \cdot C_L \cdot \frac{V_{DD}}{|I_{DSat,p}|}. \quad (4.64)$$

Die Fitting-Parameter besitzen beispielsweise für eine $0,25\mu\text{m}$ -Technologie einen Wert von $k_f = k_r \approx 0,75$ [6].

Die Verzögerungszeiten für einen Spannungssprung V_I und eine Spannungsrampe V_2 am Invertereingang sind in Abb. 4.35 veranschaulicht. Dabei stellt die Spannungsrampe die durch einen Spannungssprung V_I erzeugte Ausgangsspannung V_2 der ersten Inverterstufe dar, die als Eingangsspannung für die zweite In-

verstärkungsstufe dient (vgl. Abb. 4.34). Die Ausgangsspannung der zweiten Inverterstufe besitzt dann den Verlauf V_3 .

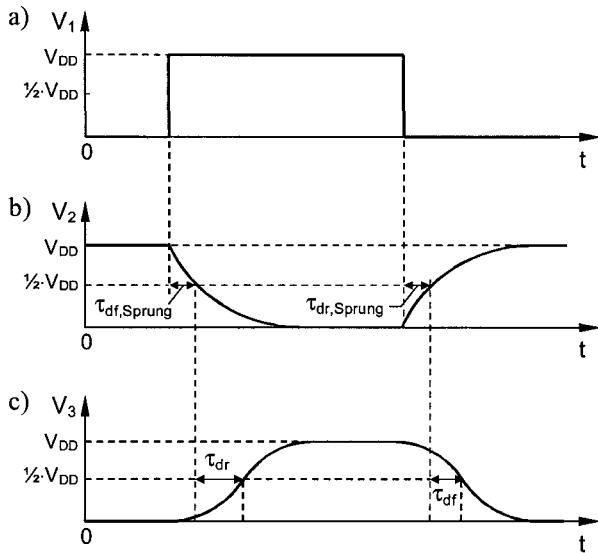


Abb. 4.35. Schaltverhalten eines CMOS-Inverters gemäß Schaltung in Abb. 4.34 a) Eingangsspannungssprung V_1 b) Ausgangsspannung V_2 der ersten Inverterstufe c) Ausgangsspannung V_3 der zweiten Inverterstufe

4.8.3 Transmissiongate

Durch die Parallelschaltung eines n-Kanal-MOSFETs T_1 und eines p-Kanal-MOSFETs T_2 , deren Gates mit invertierten Taktsignalen φ_1 und φ_2 der Frequenz $f = 1/T$ angesteuert werden, nennt man *Transmissiongate* oder Transfergate. Die Prinzipschaltung und das Takschema eines Transmissiongates sind in Abb. 4.36 dargestellt.

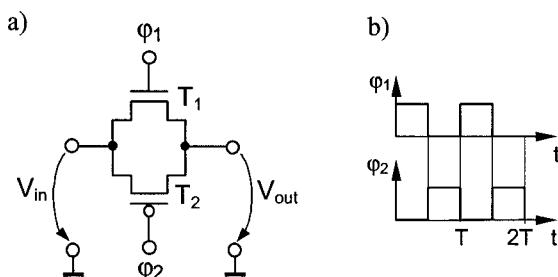


Abb. 4.36. Transmissiongate a) Prinzipschaltung und b) Takschema

Das Transmissiongate dient als bidirektonaler Schalter. Sind die Transistoren T_1 und T_2 gesperrt ($\varphi_1 < V_{Th}$, $\varphi_2 > V_{Tp}$), so ist das Transmissiongate hochohmig (geöffneter Schalterzustand). Ist einer der MOSFETs leitend, so wird das Transmissiongate niederohmig (geschlossener Schalterzustand).

Der Kanalwiderstand der MOSFETs im Transmissiongate stellt einen Serienwiderstand in der Signalstrecke zwischen Eingang und Ausgang dar. Um diesen Widerstand möglichst gering zu halten, ist die Wahl minimaler Transistorlängen bei der Dimensionierung des Transmissiongates vorteilhaft. Der Widerstand R_{on} bei geschlossenem Schalterzustand des Transmissiongates kann dann über die Drainströme $I_{DSat,n}$ und $I_{DSat,p}$ der in Sättigung befindlichen MOSFETs bestimmt werden und berechnet sich bei $V_{GSn} = V_{DD}$ und $V_{GSp} = 0$ wie folgt:

$$R_{on} = \frac{V_{DD}}{I_{DSat,n} - I_{DSat,p}} = \frac{V_{DD}}{\frac{C_{ox} \cdot W}{2} \cdot \frac{L}{\mu_n \cdot (V_{GSn} - V_{Th})^2 + \mu_p \cdot (V_{GSp} - V_{Tp})^2}}. \quad (4.65)$$

Der Widerstand R_{on} des Transmissiongates ist beispielsweise für die Bestimmung der Zeitkonstanten τ_{SC} zur Umladung einer kapazitiven Last C_L notwendig, die sich wie folgt berechnet:

$$\tau_{SC} = R_{on} \cdot C_L \quad (4.66)$$

Um eine vollständige Auf- bzw. Entladung der Lastkapazität C_L zu ermöglichen, ist $\tau_{SC} \ll T/2$ zu wählen.

4.8.4 Stromspiegel

Analog zu den Stromquellen- und Stromspiegelschaltungen mit Bipolartransistoren (vgl. Abschn. 3.7.2) können entsprechende Schaltungen mit MOSFETs aufgebaut werden. Nachfolgend ist die Prinzipschaltung des Stromspiegels mit n-Kanal-MOSFETs (Abb. 4.37a) sowie mit p-Kanal-MOSFETs (Abb. 4.37b) dargestellt.

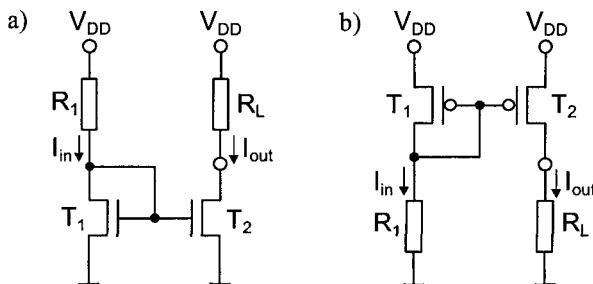


Abb. 4.37. Stromspiegel mit a) n-Kanal-MOSFETs und b) p-Kanal-MOSFETs

In Abb. 4.37a wird der n-Kanal-MOSFET T_1 durch den Kurzschluss von Drain und Gate ($V_{GS} = V_{DS}$) in Sättigung betrieben. Wenn die Drain-Source-Spannung

V_{DS} des n-Kanal-MOSFETs T_2 größer als die Drain-Sättigungsspannung ist ($V_{DS} > V_{DSat}$), dann arbeitet T_2 ebenfalls in Sättigung. Sind beide Transistoren T_1 und T_2 gleich dimensioniert, so besitzen sie nach der Beschaltung in Abb. 4.37a den gleichen Drainstrom unabhängig von der Last R_L . Damit gilt $I_{in} = I_{out}$, wenn der Einfluss des endlichen Ausgangsleitwertes g_{DS} der Transistoren vernachlässigt wird. Der gleiche Sachverhalt gilt auch für den Stromspiegel mit p-Kanal-MOSFETs in Abb. 4.37b. Das Spiegelverhältnis des Eingangsstroms I_{out} zum Ausgangsstrom I_{in} des Stromspiegels kann über das Verhältnis der Transistorflächen von T_1 und T_2 eingestellt werden.

Der differentielle Ausgangswiderstand r_{out} des Stromspiegels ist gleich dem Ausgangswiderstand r_{DS2} des MOSFETs T_2 und kann mit Hilfe des Kanallängenmodulationsfaktors λ (vgl. Abschn. 4.4) und des Drainstroms I_D von MOSFET T_2 bzw. des Ausgangsstroms I_{out} des Stromspiegels näherungsweise wie folgt berechnet werden:

$$r_{out} = r_{DS2} \approx \frac{1}{I_{out} \cdot \lambda}. \quad (4.67)$$

Ein weitaus höherer Ausgangswiderstand kann durch Verwendung einer *Kaskodestufe* erreicht werden. In Abb. 4.38 ist die Prinzipschaltung eines Stromspiegels mit Kaskodestufe dargestellt. Durch die endliche Steilheit der Ausgangskennlinie von MOSFETs im aktiven Bereich ist der Ausgangsstrom des aus den n-Kanal-MOSFETs T_1 und T_2 bestehenden Stromspiegels noch gering vom Drainpotential des Ausgangs abhängig. Durch die Kaskodestufe, die aus den n-Kanal-MOSFETs T_3 und T_4 besteht, kann dieser Effekt kompensiert werden.

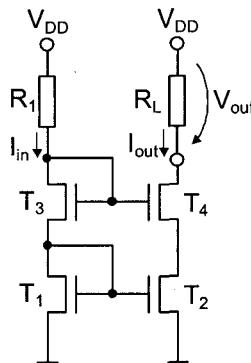


Abb. 4.38. Stromspiegel mit Kaskodestufe (bestehend aus n-Kanal-MOSFETs)

Der differentielle Ausgangswiderstand r_{out} des Stromspiegels mit Kaskodestufe berechnet sich näherungsweise aus dem Ausgangswiderstand r_{DS4} und der Steilheit g_{m4} des MOSFETs T_4 sowie des Ausgangswiderstandes r_{DS2} des MOSFETs T_2 wie folgt:

$$r_{out} \approx r_{DS2} + r_{DS4} + r_{DS2} \cdot r_{DS4} \cdot g_{m4}. \quad (4.68)$$

Der Nachteil des Stromspiegels mit Kaskodestufe liegt im reduzierten Aussteuerbereich der Ausgangsspannung, d. h. wenn sich die MOSFETs T_2 und T_4 in Sättigung befinden, wird die minimale Ausgangsspannung $V_{out,min}$ durch die Sättigungsspannung V_{DSat} der beiden MOSFETs T_2 und T_4 sowie durch die Einsatzspannung V_{Tn} bestimmt:

$$V_{out,min} = V_{Tn} + 2 \cdot V_{DSat}. \quad (4.69)$$

Durch Verwendung von großen Transistorweiten W kann die Sättigungsspannung V_{DSat} und damit die minimale Ausgangsspannung $V_{out,min}$ des Stromspiegels verringert werden.

Eine weitere Möglichkeit, einen hoch präzisen Stromspiegel mit sehr hohem Ausgangswiderstand zu realisieren, stellt der im Zusammenhang mit Bipolartransistoren (vgl. Abschn. 3.7.2) beschriebene Wilson-Stromspiegel dar. Das Schaltungsprinzip eines Wilson-Stromspiegels mit n-Kanal-MOSFETs ist in Abb. 4.39a dargestellt.

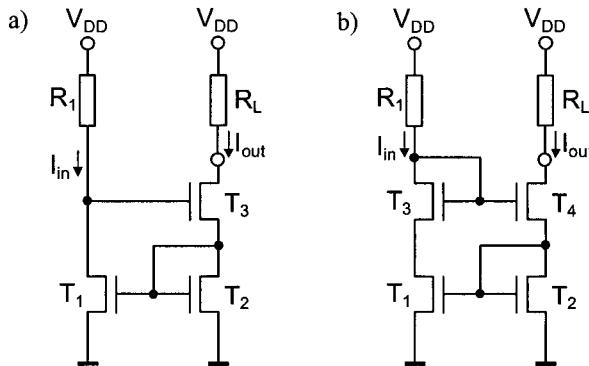


Abb. 4.39. Wilson-Stromspiegel mit n-Kanal-MOSFETs a) Grundschaltung und b) erweiterte Schaltung

Der differentielle Ausgangswiderstand r_{out} des Wilson-Stromspiegels in Abb. 4.39a berechnet sich mit Hilfe des Ausgangswiderstandes r_{DS3} und der Steilheit g_{m3} des MOSFETs T_3 sowie des Ausgangswiderstandes r_{DS1} des MOSFETs T_1 näherungsweise wie folgt:

$$r_{out} \approx r_{DS3} \cdot (g_{m3} \cdot r_{DS1} + 2) \quad (4.70)$$

In Schaltung Abb. 4.39a besitzt die Drain-Source-Spannung von MOSFET T_1 einen höheren Wert als die von MOSFET T_2 , sodass bei höheren Einsatzspannungen unterschiedliche Drainströme auf Grund des endlichen Ausgangswiderstandes dieser Transistoren resultieren. Um dies zu verhindern, kann die Schaltung in Abb. 4.39a um einen weiteren MOSFET gemäß Abb. 4.39b erweitert werden. Dadurch gilt für die Drain-Source-Spannungen der MOSFETs T_1 und T_2 : $V_{DS1} = V_{DS2}$.

Auch beim Wilson-Stromspiegel zeigt sich analog zum Stromspiegel mit Kasko-destufe ein verringelter Aussteuerbereich der Ausgangsspannung.

4.8.5 Differenzverstärker

Der Differenzverstärker besteht aus zwei gegengekoppelten Source-Schaltungen (vgl. 4.8.1.1) und dient der Verstärkung der Differenz von zwei Eingangssignalen V_{in1} und V_{in2} , die gemäß Abb. 4.40 jeweils am Gate der MOSFETs T_1 und T_2 anliegen. Dabei ist die Abhängigkeit der Ausgangsspannung des Differenzverstärkers gegenüber Temperaturschwankungen und Parameterstreuungen nur gering. Die generelle Funktionsweise gestaltet sich anlog zum Differenzverstärker mit Bipolartransistoren (vgl. Abschn. 3.7.3).

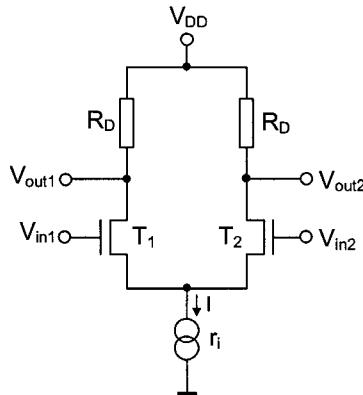


Abb. 4.40. Prinzipschaltung eines Differenzverstärkers mit n-Kanal-MOSFETs

Für die Differenzverstärkung der in Abb. 4.40 dargestellten Schaltung, die das Verhältnis der jeweiligen Ausgangsspannungsänderung dV_{out1} bzw. dV_{out2} zur Eingangsspannungsänderung $dV_{in} = d(V_{in1} - V_{in2})$ angibt, gilt dann:

$$A_D = \frac{dV_{out1}}{dV_{in}} = -\frac{dV_{out2}}{dV_{in}} = -\frac{1}{2} \cdot g_m \cdot \left(\frac{R_D \cdot r_{DS}}{R_D + r_{DS}} \right). \quad (4.71)$$

In Gl. (4.71) gibt g_m die Steilheit und r_{DS} den Ausgangswiderstand der identischen MOSFETs T_1 und T_2 an. Der Ausgangswiderstand r_{out} des Differenzverstärkers beträgt:

$$r_{out} \approx \frac{R_D \cdot r_{DS}}{R_D + r_{DS}} \text{ mit } r_{DS} \approx \frac{1}{I_D \cdot \lambda}. \quad (4.72)$$

Die Gleichtaktunterdrückung G berechnet sich aus dem Verhältnis der Differenzverstärkung A_D und der Gleichtaktverstärkung A_G mit Hilfe des Innenwiderstandes r_i der Stromquelle wie folgt:

$$G = \frac{A_D}{A_G} \approx g_m \cdot r_i. \quad (4.73)$$

Da der dynamische Ausgangswiderstand eines MOSFETs erheblich höher ist als die statische Last eines Widerstandes, werden die in Abb. 4.40 verwendeten Widerstände R_D durch eine aktive Last aus den p-Kanal-MOSFETs T_3 und T_4 gemäß Abb. 4.41 ersetzt. Dadurch kann eine höhere Differenzverstärkung A_D erzielt werden. Somit besitzt ein aus MOS-Feldeffekttransistoren aufgebauter Differenzverstärker mit aktiver Last gegenüber einem mit Bipolartransistoren realisierten Differenzverstärker (vgl. Abschn. 3.7.4) eine erheblich höhere Gleichtaktunterdrückung.

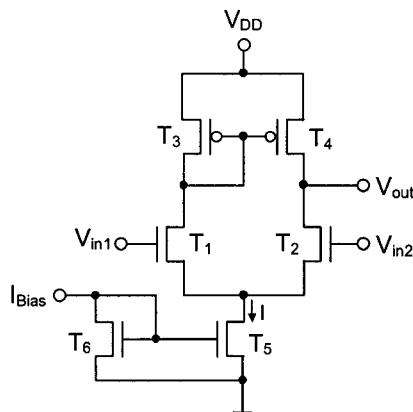


Abb. 4.41. Schaltungsprinzip eines Differenzverstärkers mit aktiver Last (T_3 und T_4) und Stromspiegel (T_5 und T_6) als Konstantstromquelle

In der in Abb. 4.41 dargestellten Schaltung wird nur ein Ausgang V_{out} genutzt. Es wurde weiterhin die Konstantstromquelle durch einen Stromspiegel aus den n-Kanal-MOSFETs T_5 und T_6 (vgl. Abb. 4.37a) ersetzt, der durch einen Bias-Strom I_{Bias} gespeist wird. Der Bias-Strom kann beispielsweise über einen Widerstand gegen Betriebsspannung V_{DD} bereitgestellt werden. Der Strom I am Source-Knoten der Eingangstransistoren T_1 und T_2 wird damit durch eine Spiegelung des Bias-Stroms erzeugt und über das Spiegelverhältnis des Stromspiegels z. B. auf $I_{Bias}/I = 10/1$ eingestellt (vgl. Abschn. 4.8.4).

4.8.6 Operationsverstärker

Der Operationsverstärker zeichnet sich in erster Linie durch eine hohe Differenzspannungsverstärkung A_D , einen hohen Eingangswiderstand r_{in} und einen geringen Ausgangswiderstand r_{out} aus. Weiterhin besitzt ein Operationsverstärker eine hohe Nullpunktstabilität und einen definierten Frequenzgang der Spannungsverstärkung (vgl. Abb. 4.43).

Zum Aufbau eines Operationsverstärkers kann der in Abschn. 4.8.5 beschriebene Differenzverstärker mit aktiver Last (vgl. Abb. 4.41) als Eingangsstufe verwendet werden. Um eine möglichst hohe Verstärkung zu erzielen, wird diese Eingangsstufe um eine Ausgangsverstärkerstufe erweitert. In Abb. 4.42 ist ein einfacher zweistufiger Operationsverstärker dargestellt, dessen Ausgangsverstärkerstufe aus den komplementären MOSFETs T_7 und T_8 besteht.

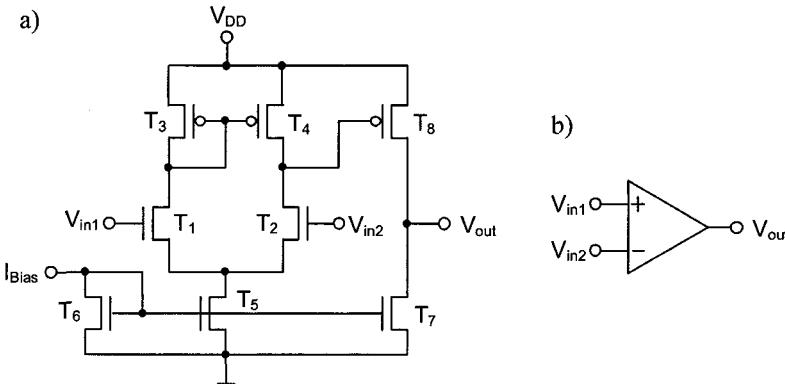


Abb. 4.42. a) Schaltungsprinzip eines einfachen zweistufigen Operationsverstärkers mit MOSFETs und b) Schaltsymbol des Operationsverstärkers

Der Eingang V_{in1} (+) des Operationsverstärkers wird als *nicht-invertierender* Eingang bezeichnet; der Eingang V_{in2} (-) wird *invertierender* Eingang genannt. Bei niedrigen Frequenzen befindet sich die Ausgangsspannung V_{out} in Phase mit der Eingangsspannungsdifferenz $\Delta V_{in} = V_{in1} - V_{in2}$.

Die (Leerlauf-) Differenzverstärkung A_D des Operationsverstärkers, die auch als *Open-Loop-Verstärkung* bezeichnet wird, berechnet sich für die in Abb. 4.42 dargestellte Schaltung bei Kleinsignalaussteuerung unter Verwendung der Steilheit g_m und des Ausgangswiderstandes r_{DS} der MOSFETs wie folgt:

$$A_D = \frac{dV_{out}}{d(V_{in1} - V_{in2})} = -\frac{g_{m1} \cdot g_{m8}}{\left(\frac{1}{r_{DS2}} + \frac{1}{r_{DS4}} \right) \cdot \left(\frac{1}{r_{DS7}} + \frac{1}{r_{DS8}} \right)} \quad \text{mit } r_{DS} \approx \frac{1}{I_D \cdot \lambda}. \quad (4.74)$$

In Abb. 4.43 ist der schematische Verlauf der Ausgangsspannung eines Operationsverstärkers in Abhängigkeit von der Eingangsspannungsdifferenz dargestellt. Dabei ist die Ausgangsspannung V_{out} im Bereich $V_{out,min} < V_{out} < V_{out,max}$ näherungsweise linear von der Eingangsspannungsdifferenz ΔV_{in} abhängig. Außerhalb dieses Spannungsbereichs ist der Operationsverstärker übersteuert und es findet kein weiterer Anstieg der Ausgangsspannung statt.

Beim idealen Operationsverstärker existiert keine Nullpunkt drift, sodass die Übertragungskennlinie den Nullpunkt schneidet. Beim realen Operationsverstärker muss zum Erreichen des Nullpunkt durchgangs eine *Offset-Spannung* V_{off} an die Eingänge des Operationsverstärkers angelegt werden (vgl. Gl. (4.76)).

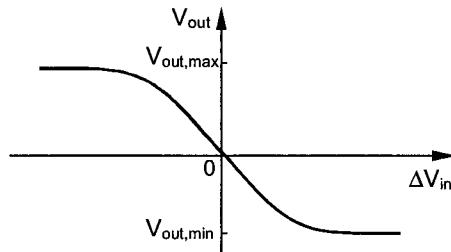


Abb. 4.43. Übertragungskennlinie eines Operationsverstärkers

Liegt an beiden Eingängen eines realen Operationsverstärkers die gleiche Spannung V_G an, so kann man die Gleichtaktverstärkung A_G und die Gleichtaktunterdrückung G wie folgt berechnen:

$$A_G = \frac{dV_{out}}{dV_G} \text{ und } G = \frac{A_D}{A_G}. \quad (4.75)$$

Im linearen Bereich der Übertragungskennlinie gilt dann für die Ausgangsspannung V_{out} unter Berücksichtigung der Offset-Spannung V_{off} :

$$V_{out} = A_D \cdot \left((\Delta V_{in} - V_{off}) + \frac{V_G}{G} \right). \quad (4.76)$$

Zur weiteren Spezifikation eines Operationsverstärkers dient die Darstellung des Frequenzganges der Differenzverstärkung nach Betrag und Phase. In Abb. 4.44 ist ein typischer Verlauf des Frequenzganges der Differenzverstärkung A_D und die zugehörige Phasenverschiebung φ_s dargestellt.

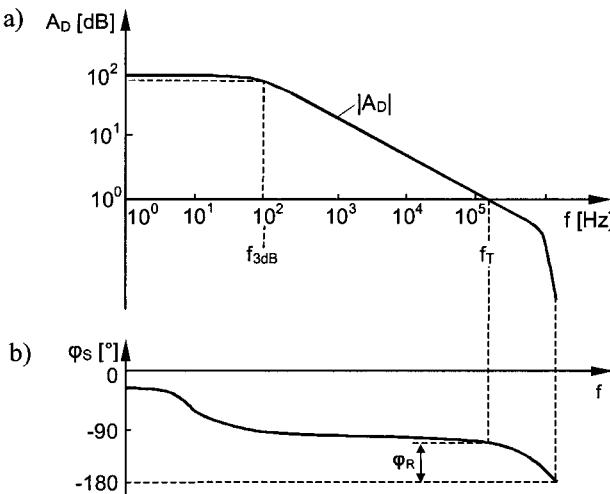


Abb. 4.44. a) Typischer Verlauf des Frequenzganges der Differenzverstärkung A_D eines Operationsverstärkers und b) zugehörige Phasenverschiebung φ_s mit Phasenreserve φ_R

In Abb. 4.44a bezeichnet f_{3dB} die Frequenz, bei der die Differenzverstärkung A_D um 3dB abgesunken ist. Die Differenzverstärkung sinkt dann ab, bis sie bei der Transitfrequenz f_T den Wert $|A_D| = 1$ annimmt. Die in Abb. 4.44b bei der Transitfrequenz auftretende Phasenverschiebung φ_S charakterisiert die sog. *Phasenreserve* φ_R , die die Differenz zu $\varphi_S = -180^\circ$ darstellt. Die Phasenreserve stellt ein Stabilitätskriterium der Schaltung dar, wobei Stabilität dann vorherrscht, wenn $\varphi_R > 45^\circ \dots 60^\circ$ gilt. Da sich der Operationsverstärker auf Grund seines mehrstufigen Aufbaus und seiner parasitären Kapazitäten wie ein Tiefpass höherer Ordnung verhält, müssen entsprechende schaltungstechnische Maßnahmen zur Frequenzgang-Korrektur durchgeführt werden. Da Operationsverstärker zumeist in einem Gegenkopplungsnetzwerk (Rückkopplung der Ausgangsspannung auf den Eingang) betrieben werden, neigen diese Schaltungsanordnungen zum Schwingen. Die Phasenreserve φ_R charakterisiert dann die Robustheit der Schaltung gegenüber diesen Schwingneigungen.

4.9 Vergleich von Bipolartransistor und MOSFET

4.9.1 Technologischer und parametrischer Vergleich

In Abb. 4.45 sind die wesentlichen technologischen und parametrischen Unterschiede zwischen einem npn-Bipolartransistor und einem n-Kanal-Feldeffekttransistor zusammengefasst.

	Bipolartransistor	Feldeffekttransistor	
Schnittbild			
Eingangs-gleichstrom	$I_B \sim \exp\left(\frac{V_{BE}}{V_\vartheta}\right)$	$I_G = 0$	
Übertragungs-verhalten	exponentiell	quadratisch	
Steilheit	$g_m = \frac{I_C}{V_\vartheta}$	$>$	$g_m \approx \mu \cdot C_{ox} \cdot \frac{W}{L} \cdot (V_{GS} - V_T)$
Stromfluss-richtung	vertikal	lateral	

Abb. 4.45. Technologische und parametrische Unterschiede zwischen npn-Bipolartransistor und n-Kanal-MOSFET

Der Eingangsleitwert von MOSFETs ist bei niedrigen Frequenzen vernachlässigbar klein; im Gegensatz zum Bipolartransistor benötigt der MOSFET keinen Steuerstrom ($I_G = 0$). Das Übertragungsverhalten des Bipolartransistors weist ein exponentielles Verhalten auf, womit die Abhängigkeit des Kollektorstroms I_C von der Steuerspannung V_{BE} stark ausgeprägt ist. Hingegen besitzt der MOSFET nur eine quadratische Abhängigkeit des Drainstroms I_D von der Steuerspannung V_{GS} im Sättigungsbereich. Demnach ist die Steilheit g_m des Bipolartransistors zumeist größer als die des Feldeffekttransistors. Der Stromfluss beim Bipolartransistor erfolgt in vertikaler Richtung (vgl. Abb. 4.45). Im Gegensatz erfolgt der Stromfluss beim MOSFET lateral im Inversionskanal.

4.9.2 Einsatzkriterien von Bipolar- und Feldeffekttransistoren

Bipolartransistoren werden insbesondere für große Verstärkungen im hohen Frequenzbereich z. B. für Oszillatoren und Frequenzteiler im oberen GHz-Bereich eingesetzt. Weiterhin sind Bipolartransistoren rauscharm, besitzen geringe Parameterstreuungen und bieten eine hohe Stromergiebigkeit pro Fläche. Die Temperaturabhängigkeit von Bipolartransistoren ist zwar stark, aber wohldefiniert. Gegenüber elektrostatischer Aufladung (ESD) sind sie robust. Demnach werden Bipolartransistoren vor allem in Hochgeschwindigkeitsschaltungen, hochpräzisen Analogschaltungen und für Spannungsreferenzen eingesetzt.

Feldeffekttransistoren, die in CMOS-Technologie hergestellt werden, eignen sich auf Grund der realisierbaren hohen Integrationsdichte und des sehr geringen Ruhestroms insbesondere für Logikgatter und Speicherschaltungen (vgl. Kap. 7). Auf Grund des hohen Eingangswiderstandes bei geringen Frequenzen werden Feldeffekttransistoren auch in Analogschaltungen eingesetzt. Da Feldeffekttransistoren im Betrieb als Schalter keinen Gleichstrom am Steuereingang führen und kein Spannungsabfall bei kapazitiver Last auftritt, eignen sich diese sowohl für Speicherschaltungen, als auch für bestimmte Digital- und Analogschaltungen. Auf Grund regenerierender Logikpegel z. B. des aus MOSFETs bestehenden CMOS-Inverters (vgl. Abschn. 4.8.2) ist die CMOS-Technologie robust gegen Parameterstreuungen. Somit sind Feldeffekttransistoren insbesondere für hochintegrierte Speicher, Logik-Schaltungen und sog. *Mixed-Signal-Schaltungen* (z. B. Analog-Digital-Wandler) geeignet.

Die vorteilhaften Eigenschaften der Bipolar- und CMOS-Technologie werden in der sog. *BiCMOS-Technologie* vereint. Hierbei werden Bipolartransistoren und MOSFETs auf einem Substrat integriert. Diese Technologie benötigt jedoch einen höheren Prozessaufwand als die Bipolar- und CMOS-Technologie. Bauelemente, die in BiCMOS-Technologie hergestellt werden, finden vor allem als schnelle Treiberschaltungen sowie als schnelle und hochpräzise Mixed-Signal-Schaltungen Anwendung.

5 Leistungsbauelemente

Aufbauend auf die in Kap. 2 - Kap. 4 behandelten Eigenschaften von Halbleiterdioden, Bipolar- und Feldeffekttransistoren werden in diesem Kapitel Bauelemente beschrieben, die speziell für hohe Schaltleistungen konzipiert sind. Die Anforderungen an diese sog. *Leistungsbauelemente* liegen in einer hohen Sperrfestigkeit und guten Eigenschaften im Flusspolung bei einfacher Ansteuerbarkeit, hoher Schaltgeschwindigkeit und Überlastfähigkeit. Nachfolgend werden der Aufbau und die Wirkungsweise von Leistungsbauelementen betrachtet, die entsprechenden Einsatzbereiche aufgezeigt sowie die Nutzung in ausgewählten Schaltungsbeispielen dargestellt.

5.1 psn-Leistungsdiode

Die wesentlichen Anforderungen an Leistungsdioden sind eine hohe Schaltleistung und geringe Sperrverluste, d. h. ein geringer Sperrsättigungsstrom. Die Schaltleistung entspricht dabei dem Produkt aus Sperrspannung und Flussstrom ($V_R \cdot I_F$).

Als Kriterien für die Auslegung der Sperrfähigkeit einer Leistungsdiode werden die maximale Sperrspannung und die maximale Feldstärke herangezogen. Die maximale Feldstärke E_M einer herkömmlichen pn-Diode (vgl. Abschn. 2.1.2.1) wird bereits bei relativ geringen Sperrspannungen erreicht. Wird die maximale Feldstärke überschritten, so kommt es zu einem Durchbruch der pn-Diode (vgl. Abschn. 2.3). Zum Erreichen einer hohen Durchbruchspannung V_{Br} und damit einer hohen maximalen Feldstärke E_M ist eine niedrige Dotierung der Leistungsdiode im Bereich der Raumladungszone notwendig.

Das Ziel nach einer geringen Verlustleistung bei Flusspolung ($V > 0$) kann durch einen geringen Spannungsabfall V über dem pn-Übergang bei hohem Flussstrom I_F erreicht werden. Dazu ist eine hohe Dotierung beider Halbleiterzonen erforderlich (entgegen einer zur Optimierung der Sperrfähigkeit notwendigen niedrigen Dotierung). Diese gegenläufigen Anforderungen an die Dotierungshöhe werden durch Separation des p- und n-Gebietes der pn-Diode durch eine sehr schwach dotierte Halbleiterschicht (s-Zone) analog zur pin-Diode (vgl. Abschn. 2.7.1) erfüllt; man spricht dann bei der für Leistungsanwendungen optimierten Struktur von einer *psn-Leistungsdiode*.

Der schematische Aufbau einer $p^+ s_n^+$ -Leistungsdiode ist in Abb. 5.1a im Querschnitt dargestellt. Ein beispielhaftes Dotierungsprofil der Struktur ist in Abb. 5.1b ersichtlich.

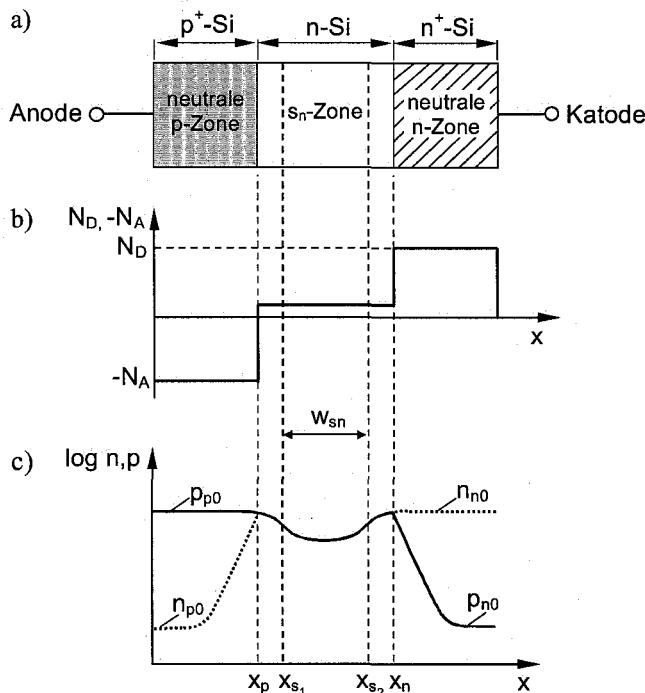


Abb. 5.1. a) Schematischer Aufbau einer p^+Si-n^+ -Leistungsdiode b) beispielhaftes Dotierungsprofil und c) LadungsträgerdichteVerteilung bei Flusspolung ($V > 0$)

Wird die p^+Si-n^+ -Leistungsdiode in Flusspolung ($V > 0$) betrieben, so wird die schwach dotierte s_n -Zone von beiden benachbarten stark dotierten Zonen mit freien Ladungsträgern überschwemmt und die Diode wird leitend. Die hierbei auftretende Verteilung der Ladungsträgerdichten ist in Abb. 5.1c dargestellt. Auf Grund der starken Injektion von Ladungsträgern gilt in der schwach dotierten s_n -Zone für die Minoritätsträgerdichte $p \gg N_D$. Dadurch folgt für die Ladungsträgerverteilung in der s_n -Zone $p(x) \approx n(x)$. Der Gesamtstrom bzw. die Kennliniengleichung der p^+Si-n^+ -Leistungsdiode wird dann gemäß Abschn. 2.7.1, Gl. (2.38) bestimmt. Dabei wird die psn-Struktur so dimensioniert, dass einerseits der Spannungsabfall über der s_n -Zone nicht zu groß wird, andererseits aber der Rekombinationsstrom in der s_n -Zone dominiert. Für die neutrale Weite w_{sn} der s_n -Zone (vgl. Abb. 5.1) gilt dann in bezug auf die Diffusionslänge: $w_{sn} \ll L_{Dsn}$. Man erhält für den Diodenstrom I näherungsweise:

$$I \approx 2 \cdot I_{Ssn} \cdot \left(\exp\left(\frac{V}{2 \cdot V_\vartheta}\right) - 1 \right). \quad (5.1)$$

In Gl. (5.1) bezeichnet I_{Ssn} den Sperrsättigungsstrom der s_n -Zone, der sich näherungsweise mit Hilfe der Ladungsträgerlebensdauer τ_{sn} in der s_n -Zone und der neutralen Weite w_{sn} der s_n -Zone wie folgt berechnet:

$$I_{SSn} \approx \frac{e \cdot A \cdot n_i \cdot w_{sn}}{2 \cdot \tau_{sn}}. \quad (5.2)$$

Befindet sich die $p^+s_nn^+$ -Leistungsdiode in Sperrpolung ($V < 0$), so dehnt sich die Raumladungszone vor allem in die schwach dotierte s_n -Zone aus. Demnach kann näherungsweise die gesamte Sperrspannung als Feldstärkeintegral über der s_n -Zone dargestellt werden. Die Feldstärkeverläufe $E(x)$ einer $p^+s_nn^+$ -Leistungsdiode bei verschiedenen Sperrspannungen $V < 0$ sind in Abb. 5.2 dargestellt.

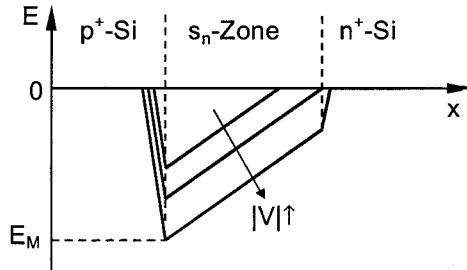


Abb. 5.2. Feldstärkeverläufe einer $p^+s_nn^+$ -Leistungsdiode für verschiedene Sperrspannungen $V < 0$

Um eine optimale Sperrfähigkeit der $p^+s_nn^+$ -Leistungsdiode zu erreichen, sollte ein möglichst konstanter Feldstärkeverlauf in der s_n -Zone erzielt werden. Dies kann durch eine besonders niedrige Dotierung der s_n -Zone erreicht werden. Erreicht die Raumladungszone das Ende der schwach dotierten s_n -Zone (Punch-Through-Effekt, vgl. Abschn. 3.3.3), dann steigt die maximale Feldstärke E_M rasch mit der Spannung V an und es kommt zum Lawinendurchbruch (vgl. Abschn. 2.2.1).

Mit $p^+s_nn^+$ -Leistungsdioden können typische Sperrspannungen bis zu 10 kV erreicht werden.

5.2 Leistungsbipolartransistor

Die wesentlichen Anforderungen an einen Leistungsbipolartransistor sind ein hoher Strom bei geringem Durchlasswiderstand, eine hohe Sperrspannung (hohe Kollektor-Emitter-Spannung V_{CE} im blockierten Zustand), ein gutes Hochstromverhalten der Stromverstärkung und eine hohe Sicherheit gegen einen Durchbruch zweiter Art (vgl. Abschn. 3.3.2). Um eine hohe Stromverstärkung zu realisieren, wird beim Bipolartransistor (vgl. Kap. 3) eine geringe Basisweite gewählt; dies führt jedoch zu einer niedrigen Durchbruchsspannung auf Grund des Punch-Through-Effekts (vgl. Abschn. 3.3.3). Zur Vermeidung des Punch-Through-Effekts muss die Basisweite des Leistungsbipolartransistors hinreichend groß gewählt werden, was jedoch zu einer relativ geringen Stromverstärkung führt. Um eine geringe Basisweitenmodulation (vgl. Abschn. 3.2.2) zu erzielen, muss die

Kollektordotierung niedriger sein als die Basisdotierung. Dies führt allerdings zu einem hohen Kollektorbahnwiderstand. Um diesen Nachteil möglichst gering zu halten, kann das hoch dotierte Kollektoranschlussgebiet durch eine sehr schwach dotierte Halbleiterschicht (s_n -Zone) von der Basis separiert werden. Der resultierende schematische Aufbau eines Leistungsbipolartransistors mit beispielhaftem Dotierungsprofil und zugehörigem Feldstärkeverlauf $E(x)$ ist in Abb. 5.3 dargestellt.

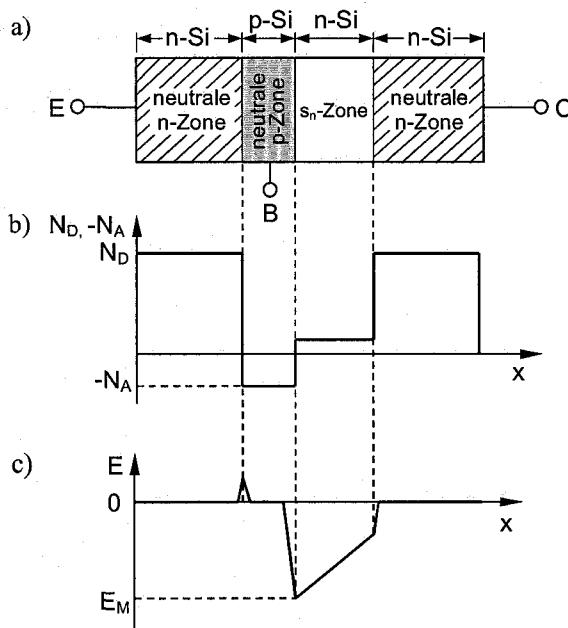


Abb. 5.3. a) Schematischer Aufbau eines Leistungsbipolartransistors b) beispielhaftes Dotierungsprofil und c) Feldstärkeverlauf

Durch die in Abb. 5.3a dargestellte s_n -Zone kann die maximale Feldstärke E_M verringert werden (im Blockierfall gilt $V_{CE} \approx V_{CB}$). Da sich die Raumladungszone im wesentlichen in der gering dotierten s_n -Zone ausbildet, kann die Basisweite schmal gehalten werden, ohne den Punch-Through-Effekt hervorzurufen. Jedoch wird durch die s_n -Zone ein zusätzlicher Bahnwiderstand verursacht, der das Verhalten des Leistungsbipolartransistors im Durchlassbetrieb verschlechtert. Auf Grund der Ladungsträgerspeicherung in der s_n -Zone werden außerdem die dynamischen Eigenschaften beeinträchtigt.

Gegenüber einem herkömmlichen Bipolartransistor tritt beim Leistungsbipolartransistor bei hohen Kollektorströmen eine Verringerung der Stromverstärkung β auf. Ein typischer Verlauf der Stromverstärkung eines Leistungsbipolartransistors in Abhängigkeit vom Kollektorstrom I_C ist in Abb. 5.4 bei konstanter Kollektor-Emitter-Spannung V_{CE} für unterschiedliche Temperaturen $T_1 < T_2$ aufgezeigt.

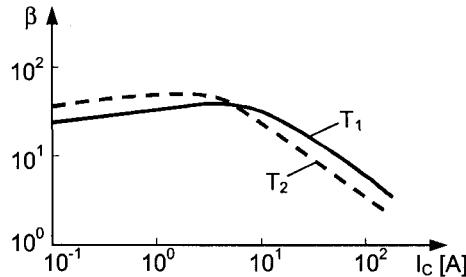


Abb. 5.4. Beispielhafter Verlauf der Stromverstärkung β eines Leistungsbipolartransistors als Funktion des Kollektorstroms I_C bei konstanter Kollektor-Emitter-Spannung für unterschiedliche Temperaturen $T_1 < T_2$

Die Verringerung der Stromverstärkung β bei höheren Kollektorströmen wird insbesondere durch die sog. *Basisaufweitung* hervorgerufen, die vor allem durch einen Abbau der Raumladungszone zwischen Basis und Kollektor im Sättigungsfall erfolgt. Auf Grund der geringen Stromverstärkung bei hohen Kollektorströmen ist eine hohe Steuerleistung notwendig. Mit Hilfe einer Darlington-Schaltung (vgl. Abschn. 3.7.1) kann die Stromverstärkung von Leistungsbipolartransistoren erhöht werden.

Durch eine großflächige Ausführung des Leistungsbipolartransistors und Montage in einem Gehäuse mit geringem thermischen Widerstand kann ein Durchbruch zweiter Art (vgl. Abschn. 3.3.2) vermieden werden. Durch die Wahl einer geeigneten Schichtfolge (z. B. $n^+ps_{n,n}$ -Struktur) wird eine hohe Durchbruchsspannung $V_{CE} = V_{Br}$ zwischen Kollektor und Basis erzielt (vgl. Abschn. 3.3.1).

Typische Grenzwerte eines Leistungsbipolartransistors sind z. B. eine Sperrspannung von $V_{CE} \approx 1000$ V bei einem maximalen Kollektorstrom von $I_C \approx 50$ A. Bei geringeren Sperrspannungen (z. B. $V_{CE} \approx 300$ V) werden maximale Kollektorströme von bis zu $I_C \approx 500$ A erreicht. Die Schaltgeschwindigkeit von Leistungsbipolartransistoren ist relativ gering, da die große Basisweite und die hohe Basisladung zu einer langen Schaltzeit führen (vgl. Abschn. 3.6.). Typische Schaltfrequenzen liegen bei 5 kHz.

Das Anwendungsgebiet von Leistungsbipolartransistoren liegt vor allem im unteren bis mittleren Leistungsbereich. Sie werden z. B. in Analogverstärkern und Motorsteuerungen eingesetzt.

5.3 Leistungs-MOSFET

Zur Erhöhung der Stromergiebigkeit eines MOSFETs (vgl. Kap. 4) werden mehrere Einzeltransistoren parallel geschaltet, wodurch der s.g. *Leistungs-MOSFET* entsteht. Zur flächenoptimalen Anordnung der parallel geschalteten Einzeltransistoren werden MOSFETs verwendet, deren Drainanschlüsse sich auf der Unterseite des Substrats befinden. MOSFETs dieses Strukturaufbaus werden durch Doppelimplantation der Kanalstruktur hergestellt und als *DMOSFET* bezeichnet. Der

schematische Aufbau eines n-Kanal-DMOSFET-Elements ist in Abb. 5.5a im Querschnitt dargestellt.

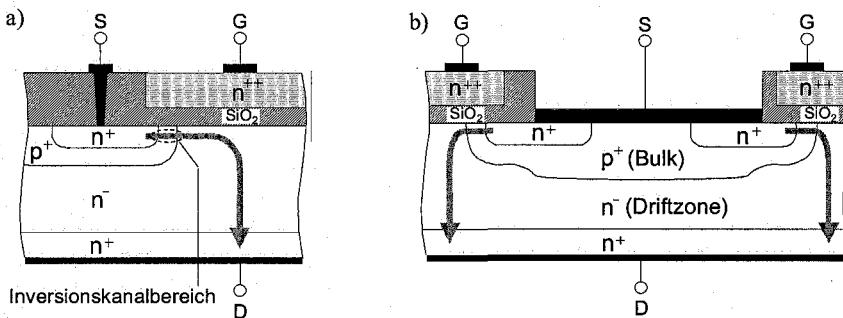


Abb. 5.5. Schematischer Aufbau (Querschnitt) a) eines Elements eines n-Kanal-DMOSFETs und b) von zwei parallel geschalteten Elementen eines DMOSFETs

Der Leistungs-MOSFET entsteht durch die integrierte Parallelschaltung von mehreren hundert bis tausend Einzelstrukturen, wobei die Verbindung der n⁺-Source-Gebiete mit dem p⁺-Bulk-Gebiet durch die Source-Metallisierung realisiert wird. Der schematische Aufbau eines Leistungs-MOSFETs, der hier beispielhaft aus zwei n-Kanal-DMOSFETs besteht, ist in Abb. 5.5b im Querschnitt dargestellt.

Der Leistungs-MOSFET befindet sich im leitenden Zustand, wenn für $V_{GS} > V_T$ der Inversionskanal zwischen n⁺-Source-Gebiet und s. g. n⁻-Driftzone gemäß Abb. 5.5 aufgebaut wird. Auf Grund der Doppelimplantation der Kanalstruktur sind sehr kurze Kanallängen realisierbar, wodurch eine hohe Sättigungsstromdichte erreicht wird. Durch niedrige Dotierung und große Breite der Driftzone kann eine hohe Sperrspannung erreicht werden, was jedoch mit einem hohen On-Widerstand R_{on} im Durchlassbetrieb verbunden ist. Mit zunehmender Sperrspannung wird deshalb die Stromergiebigkeit des Leistungs-MOSFETs geringer. In Abb. 5.6 sind typische Verläufe des On-Widerstandes eines Leistungs-MOSFETs in Abhängigkeit vom Drainstrom I_D für verschiedene Gate-Source-Spannungen $V_{GS1} < V_{GS2} < V_{GS3}$ dargestellt.

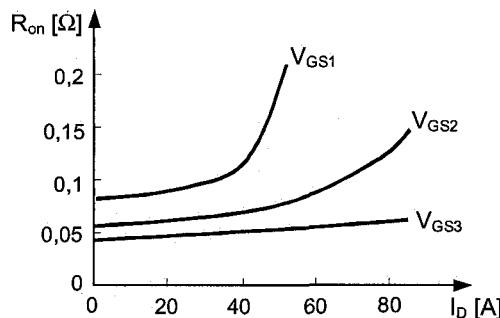


Abb. 5.6. Verlauf des On-Widerstandes R_{on} eines Leistungs-MOSFETs für verschiedene Gate-Source-Spannungen $V_{GS1} < V_{GS2} < V_{GS3}$

Im Vergleich zum Leistungsbipolartransistor (vgl. Abschn. 5.2) weist der Leistungs-MOSFET bei näherungsweise gleicher Transistorfläche einen höheren On-Widerstand auf.

Zwischen Source- und Draingebiet des Leistungs-MOSFETs befindet sich ein parasitärer npn-Bipolartransistor, der bei Schaltvorgängen leitend werden kann und einen Kurzschluss zwischen Source- und Draingebiet des Leistungs-MOSFETs bewirkt. Diesem Vorgang kann durch eine hohe Dotierung des Bulkgebietes entgegengewirkt werden. Das vereinfachte Ersatzschaltbild eines Leistungs-MOSFETs mit parasitärem npn-Bipolartransistor T_2 und Bahnwiderstand R_B des Bulkgebietes ist in Abb. 5.7a dargestellt.

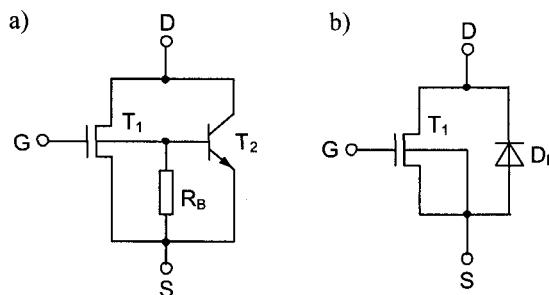


Abb. 5.7. Vereinfachtes Ersatzschaltbild eines n-Kanal-Leistungs-MOSFETs a) mit parasitärem npn-Bipolartransistor T_2 und b) mit Inversdiode D_F

Bei Rückwärtsbetrieb ($V_{DS} < 0$) weist der Leistungs-MOSFET keine Sperrfähigkeit auf, da sich die Inversdiode D_F des pn-Überganges zwischen Basis und Kollektor des parasitären npn-Bipolartransistors bei $V_{DS} < 0$ in Flusspolung befindet (vgl. Abb. 5.6b). Beim Schalten induktiver Lasten kann die Inversdiode die Funktion einer sog. *Freilaufdiode* übernehmen. Eine Freilaufdiode verhindert durch Kurzschluss die Induktion einer hohen Spannung beim Abschalten des Stromflusses durch die induktive Last (vgl. Abschn. 5.5.5).

Die Übertragungskennlinie und das Ausgangskennlinienfeld eines Leistungs-MOSFETs verhalten sich im wesentlichen wie die Verläufe eines herkömmlichen MOSFETs (vgl. Abschn. 4.3). Die Einsatzspannung von Leistungs-MOSFETs ist auf Grund erwünschter Störsicherheit höher als bei herkömmlichen MOSFETs; $|V_T|$ liegt etwa bei 2 V bis 5V. Folgende Betriebszustände des selbstsperrenden n-Kanal- und p-Kanal-Leistungs-MOSFETs werden unterschieden:

Leistungs-MOSFET-Typ	Durchlassbetrieb	Blockierbetrieb	Rückwärtsbetrieb
n-Kanal	$V_{GS} > V_{Tn}, V_{DS} \geq 0$	$V_{GS} < V_{Tn}, V_{DS} \geq 0$	$V_{DS} < 0$
p-Kanal	$V_{GS} < V_{Tp}, V_{DS} \leq 0$	$V_{GS} > V_{Tp}, V_{DS} \leq 0$	$V_{DS} > 0$

Wesentliche Unterschiede zwischen Leistungsbipolartransistor und Leistungs-MOSFET liegen im Schaltverhalten. Im Gegensatz zum Bipolartransistor fließen beim Leistungs-MOSFET nur für kurze Zeit Steuerströme. Das Schaltverhalten er-

folgt meist mit konstantem Gatestrom. Die kann durch einen Spannungssprung mit einem Vorwiderstand am Gate charakterisiert werden. Während des Einschaltvorganges resultiert dann ein näherungsweise konstanter Gatestrom I_G , mit dem die Gatekapazität umgeladen wird. Bis zum Erreichen der Einsatzspannung V_T fließt noch kein Drainstrom I_D . Wird der Leistungs-MOSFET leitend ($V_{GS} > V_{Th}$), so wird die Drain-Source-Spannung V_{DS} von der zuvor anliegenden Sperrspannung auf die niedrige Durchlassspannung entladen. Somit wird die Gate-Drain-Kapazität C_{GD} durch den Miller-Effekt (vgl. Abschn. 4.8.1.1) erhöht. Im Schaltpunkt genügt eine geringe Änderung der Gate-Source-Spannung V_{GS} , um eine sehr starke Änderung der Drain-Source-Spannung zu bewirken. Demnach ist eine hohe Ladung notwendig, um die Miller-Kapazität in diesem Arbeitspunkt umzuladen. Die Gate-Source-Spannung bleibt dabei nahezu konstant; man spricht vom sog. *Miller-Plateau*. Der bei Vorherrschen des Miller-Plateaus fließende Gatestrom I_G berechnet sich näherungsweise mit Hilfe der Gate-Drain-Kapazität C_{GD} und der zeitlichen Änderung der Drain-Source-Spannung V_{DS} wie folgt:

$$I_G \approx C_{GD} \cdot \frac{dV_{DS}}{dt}. \quad (5.3)$$

Der Wert der Miller-Kapazität ist am Ende des Einschaltvorganges (Übergang vom Sättigungs- in den Triodenbereich) am größten. Erst wenn der zum Umladen der Millerkapazität notwendige Strom kleiner als der Gatestrom I_G ist, steigt die Gate-Source-Spannung weiter an. Das Schaltverhalten eines Leistungs-MOSFETs ist in Abb. 5.8 mit den auftretenden Schaltzeiten dargestellt.

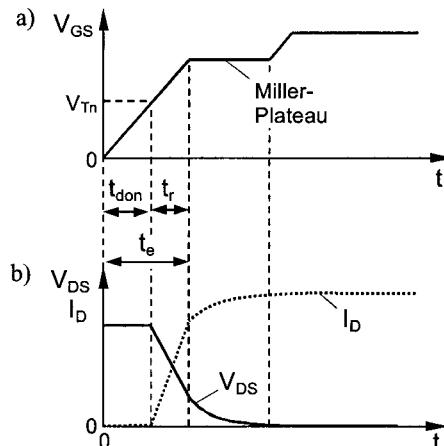


Abb. 5.8. Schaltverhalten eines n-Kanal-Leistungs-MOSFETs bei ohmscher Last und mit eingeprägtem Gatestrom a) Zeitverhalten der Gate-Source-Spannung V_{GS} mit Miller-Plateau und b) Zeitverhalten der Drain-Source-Spannung V_{DS} und des Drainstroms I_D

In Abb. 5.8 kennzeichnet t_{don} die Einschaltverzögerung des Leistungs-MOSFETs; sie charakterisiert die Zeit, die die Gate-Source-Spannung V_{GS} bis zum Erreichen der Einsatzspannung V_T benötigt. Der Anstieg des Drainstroms I_D bzw. der

Abfall der Drain-Source-Spannung V_{DS} wird durch die Anstiegszeit t_r beschrieben. Die Summe aus Einschaltverzögerung und Anstiegszeit wird als Einschaltzeit t_e bezeichnet ($t_e = t_{don} + t_r$). Das Ausschalten eines Leistungs-MOSFETs erfolgt im wesentlichen umgekehrt zum Einschalten.

Leistungs-MOSFETs besitzen eine hohe thermische Stabilität, d. h. auf Grund des negativen Temperaturkoeffizienten des Drainstroms führt ein Kurzschluss zwischen Drain und Betriebsspannung V_{DD} nicht unmittelbar zu einem thermischen Durchbruch. Erst nach einer Kurzschlussdauer von ca. $10\mu\text{s}$ erfolgt ein thermischer Durchbruch. Diese Zeit gestattet es, mit Hilfe einer externen Begrenzungsschaltung diesem Zustand vorzubeugen.

Leistungs-MOSFETs sind typischerweise für Blockierspannungen bis 1000 V und Drainströme bis 30 A ausgelegt; sie können bei Schaltfrequenzen bis zu einigen 100 kHz betrieben werden. Leistungs-MOSFETs werden insbesondere im unteren bis mittleren Leistungsbereich z. B. in Schaltnetzteilen, Motorsteuerungen, Gleichspannungswandlern und Breitbandverstärkern eingesetzt. Darüber hinaus lassen sich Leistungs-MOSFETs direkt über die Ausgänge von integrierten CMOS-Gattern (vgl. Kap. 7) ansteuern.

5.4 Thyristor

5.4.1 Aufbau und Wirkungsweise des Thyristors

Der Thyristor ist ein gut steuerbarer Leistungsgleichrichter in Vierschichtstruktur der Schichtfolge $\text{ps}_n\text{s}_p\text{n}$ mit den Anschlüssen Anode, Gate (Steuereingang) und Katode (Emitter). Die äußere p- und n-Zone des Thyristors ist jeweils hochdotiert; die s_n - und s_p -Zonen sind jeweils schwach dotiert. Die Breite der s_p -Zone ist relativ schmal ausgeprägt und liegt in der Größenordnung der Diffusionslänge L_n (vgl. Abschn. 1.5.4) der Minoritätsträger. Der schematische Aufbau eines Thyristors mit beispielhaftem Dotierungsprofil ist in Abb. 5.9 dargestellt.

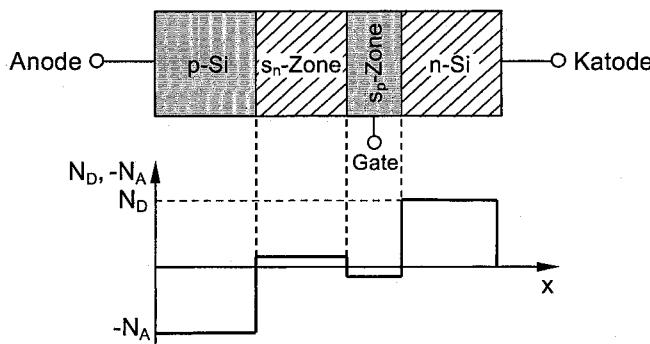


Abb. 5.9. Schematischer Aufbau eines Thyristors in $\text{ps}_n\text{s}_p\text{n}$ -Schichtfolge mit beispielhaftem Dotierungsprofil

Im Gegensatz zur herkömmlichen Gleichrichterdiode (vgl. Kap. 2), die für eine der beiden Halbwellen eines sinusförmigen Stroms stets durchlässig ist, sperrt der Thyristor zunächst beide Halbwellen. Erst durch einen Zündstromimpuls $I_G > 0$ am Gate kann der Thyristor während der positiven Halbwelle (Anoden-Katoden-Spannung $V_{AK} > 0$) eingeschaltet werden. Dabei bleibt der Thyristor auch nach Wegfall des Zündstromimpulses (Gatestrom $I_G = 0$) im leitenden Zustand. Durch die Wahl des entsprechenden Zündzeitpunktes kann die Dauer des Stromflusses und damit der Stromverlauf bzw. der mittlere Gleichstrom durch den Thyristor gesteuert werden. Das Abschalten des Thyristors erfolgt durch Umpolung der Anoden-Katoden-Spannung ($V_{AK} < 0$) bei Nulldurchgang der Wechselspannung (vgl. Abschn. 5.4.4).

Im thermodynamischen Gleichgewicht (externe Ströme und Spannungen sind gleich Null) bildet sich an den drei pn-Übergängen der Thyristorstruktur jeweils eine Raumladungszone aus, deren Diffusionsspannungen V_D (vgl. Abschn. 2.1) mit den Kontaktspannungen in der Summe den Wert Null ergeben (vgl. Abschn. 2.1.1, Gl. (2.4)).

Man unterscheidet drei Betriebsarten des Thyristors: den Blockier-, Durchlass- und Sperrbetrieb. Der Blockierbetrieb (Anoden-Katoden-Spannung $V_{AK} > 0$) liegt vor, wenn sich die pn-Übergänge zwischen p- und s_n-Zone sowie zwischen n- und s_p-Zone in Flusspolung befinden. Der pn-Übergang zwischen s_n- und s_p-Zone ist dann sperrgepolt, wodurch an diesem pn-Übergang der überwiegende Spannungsabfall auf Grund der stark ausgedehnten Raumladungszone auftritt. Dadurch ist der Stromfluss zwischen Anode und Katode sehr gering (Thyristor ist ausgeschaltet). Die Raumladungszone und der resultierende Potentialverlauf eines Thyristors im Blockierbetrieb sind in Abb. 5.10 (bei Vernachlässigung der Diffusionsspannungen) dargestellt.

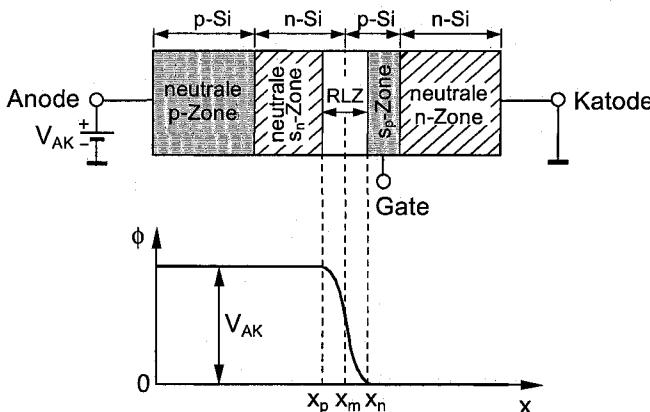


Abb. 5.10. Raumladungszone und Potentialverlauf eines Thyristors im Blockierbetrieb (Diffusionsspannungen vernachlässigt)

Wird die Anoden-Katoden-Spannung V_{AK} so weit erhöht, dass der pn-Übergang zwischen s_n- und s_p-Zone die Sperrwirkung verliert, so erfolgt das sog. *Zünden* des

Thyristors. Dabei kippt der Thyristor ab einer Spannung $V_{AK} \geq V_{BF}$ (*Kippspannung*) in den Durchlassbetrieb. Im Durchlassbetrieb des Thyristors befinden sich alle drei pn-Übergänge in Flussbetrieb, sodass aus der hochdotierten n- und p-Zone Ladungsträger in die niedrig dotierte s_n - bzw. s_p -Zone injiziert werden. Dadurch erfolgt ein starker Anstieg des Anodenstroms I_A des Thyristors.

Im Sperrbetrieb ($V_{AK} \leq 0$) verhält sich der Thyristor analog einer pn-Diode. Dabei befinden sich die pn-Übergänge zwischen p- und s_n -Zone sowie zwischen n- und s_p -Zone in Sperrpolung. Der pn-Übergang zwischen s_n - und s_p -Zone ist dann flussgepolt. Die bei Sperrbetrieb auftretenden Ausdehnungen der Raumladungszonen und der Potentialverlauf des Thyristors ist in Abb. 5.11 bei Vernachlässigung der Diffusionsspannungen dargestellt.

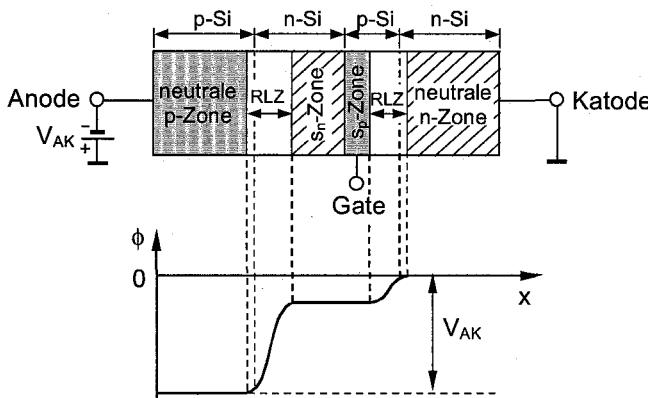


Abb. 5.11. Raumladungszone und Potentialverlauf eines Thyristors in Sperrbetrieb (Diffusionsspannungen vernachlässigt)

Die zum Zünden des Thyristors erforderliche Kippspannung V_{BF} kann deutlich verringert werden, wenn ein zusätzlicher Steuerstrom (Zündstrom) I_G am Gate anliegt. Dieser Strom erhöht die Minoritätsträgerdichte in der s_p -Zone, sodass die Raumladungszone am pn-Übergang zwischen s_n - und s_p -Zone nahezu unabhängig von V_{AK} ihre Sperrwirkung verliert (vgl. Abschn. 5.4.3).

5.4.2 Thyristorkennlinie

Die Abhängigkeit des Anodenstroms I_A von der Anoden-Katoden-Spannung V_{AK} für die in Abschn. 5.4.1 beschriebenen Betriebsarten wird in der sog. *Hauptstromkennlinie* des Thyristors dargestellt. In Abb. 5.12 ist ein schematischer Verlauf der Hauptstromkennlinie und das Schaltsymbol des Thyristors abgebildet.

Die Blockierkennlinie ($0 < V_{AK} < V_{BF}$) kennzeichnet das Verhalten des Thyristors bis zum Einsetzen des Zündvorganges (Vorwärtsdurchbruch ohne Gatestrom). Das Zünden erfolgt bei der Kippspannung $V_{AK} = V_{BF}$ bei einem Kippstrom $I_A = I_{BF}$ und es gilt in diesem Arbeitspunkt, dass der Anstieg dV_{AK}/dI_A den Wert Null besitzt.

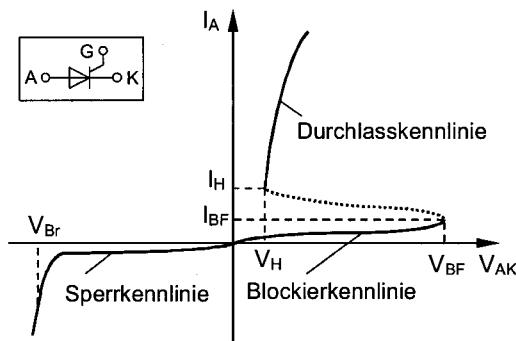


Abb. 5.12. Schaltsymbol und Hauptstromkennlinie eines Thyristors für Blockier-, Durchlass- und Sperrbetrieb

Nach Überschreiten der Kippspannung bzw. des Kippstroms findet ein Ansteigen des Anodenstroms I_A statt; der Thyristors befindet sich im Durchlassbetrieb. Die entsprechende Durchlasskennlinie verhält sich näherungsweise wie die einer pn-Diode (vgl. Abschn. 2.2.2). Der Durchlassbetrieb des Thyristors bleibt so lange erhalten, bis die sog. Haltespannung V_H bzw. der Haltestrom I_H unterschritten wird (vgl. Abb. 5.12).

Die Sperrkennlinie ($V_{AK} \leq 0$) des Thyristors verhält sich analog der Sperrkennlinie einer pn-Diode (vgl. Abschn. 2.2.2). Als charakteristische Größe dient die Durchbruchspannung V_{Br} , ab der Durchbruchsmechanismen (vgl. Abschn. 2.3) auftreten und ein starkes Ansteigen des Sperrstroms des Thyristors bewirken.

5.4.3 Zündverhalten des Thyristors

Zur Veranschaulichung des Zündvorganges kann der Thyristor durch eine Ersatzschaltung aus zwei komplementären Bipolartransistoren (vgl. Kap. 3) T_1 und T_2 beschrieben werden, die in Abb. 5.13a dargestellt ist. Dabei wird über einen Schalter S der Zündstrom I_G pulsförmig am Gate des Thyristors bereitgestellt. Die Grundschaltung für die Ansteuerung eines Thyristors bei Verwendung einer ohmschen Last R_L ist in Abb. 5.13b veranschaulicht.

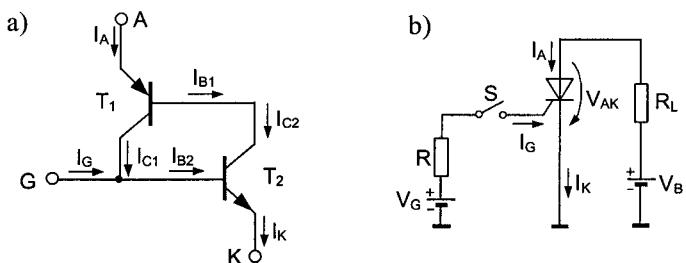


Abb. 5.13. a) Ersatzschaltung des Thyristors aus zwei komplementären Bipolartransistoren T_1 (pnp) und T_2 (npn) zur Veranschaulichung des Zündvorganges b) Grundschatzung zur Ansteuerung eines Thyristors

Die Zündbedingung eines Thyristors ergibt sich aus den Stromverstärkungen α_1 und α_2 (vgl. Abschn. 3.5.3) der in Abb. 5.13a dargestellten Bipolartransistoren T_1 und T_2 wie folgt:

$$\alpha_1 + \alpha_2 \geq 1. \quad (5.4)$$

Erfolgt das Zünden des Thyristors durch das Überschreiten der Kippsspannung V_{BF} (vgl. Abschn. 5.4.1) ohne einen zusätzlichen Steuerstrom am Gate ($I_G = 0$), so spricht man vom sog. *Überkopfzünden*. Hierbei wird die Spannungsabhängigkeit der Stromverstärkungen α_1 und α_2 ausgenutzt. Bei ansteigender Anoden-Katoden-Spannung V_{AK} verringern sich die Basisweiten der Transistoren T_1 und T_2 auf Grund der Basisweitenmodulation (vgl. Abschn. 3.2.2) und der Einfluss des Lawineneffektes (vgl. Abschn. 3.3.1) nimmt zu, wodurch die Stromverstärkungen bei der Kippsspannung $V_{AK} \geq V_{BF}$ den zum Zünden des Thyristors erforderlichen Wert überschreiten.

Die zum Zünden des Thyristors notwendige Kippsspannung V_{BF} kann deutlich verringert werden, wenn ein zusätzlicher Zündstrom I_G am Gate anliegt (vgl. 5.4.1), wodurch eine Erhöhung der Minoritätsträgerdichte in der s_p -Zone erfolgt. Dabei wird die Stromabhängigkeit der Stromverstärkungen α_1 und α_2 ausgenutzt. Eine Erhöhung des Anodenstroms I_A hat eine Erhöhung der Stromverstärkungen zur Folge, wodurch bei einem Kippstrom $I_A \geq I_{BF}$ die Stromverstärkungen den zum Zünden des Thyristors erforderlichen Wert ($\alpha_1 + \alpha_2 = 1$) überschreiten. Die Abhängigkeit des Anodenstroms I_A von der Stromverstärkung α_1 des pnp-Transistors T_1 und der Stromverstärkung α_2 des npn-Transistors T_2 sowie der Summe von α_1 und α_2 ist beispielhaft in Abb. 5.14 dargestellt.

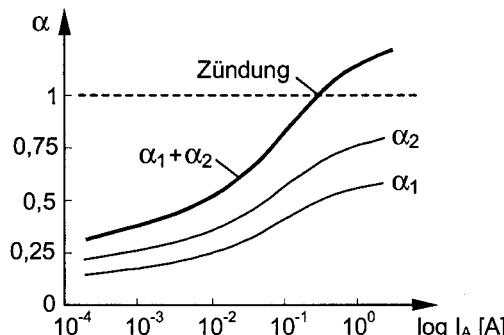


Abb. 5.14. Typischer Verlauf der Stromverstärkungen α_1 und α_2 der komplementären Bipolartransistoren des Thyristors in Abhängigkeit vom Anodenstrom I_A

Wird der Anodenstrom I_A durch einen Zündimpuls am Gate auf einen Wert oberhalb des Haltestroms I_H (vgl. Abb. 5.12) erhöht, so wird der Einschaltvorgang des Thyristors auch bei Wegfall des Zündstroms I_G fortgesetzt. Die Hauptstromkennlinie eines Thyristors für verschiedene Zündströme I_G ist in Abb. 5.15 veranschaulicht.

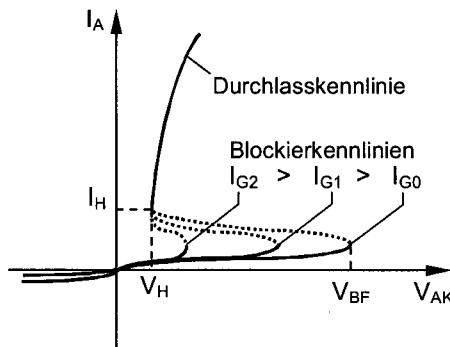


Abb. 5.15. Hauptstromkennlinie eines Thyristors für verschiedene Steuerströme I_G

Neben der Beschleunigung des Zündvorganges durch einen Zündstrom kann auch eine Temperaturerhöhung oder Bestrahlung des Thyristors mit Licht zur Erhöhung der Minoritätsträgerdichte in der s_p -Zone führen. Bei Temperaturerhöhung wächst der Sperrstrom des Thyristors und damit auch der Anodenstrom I_A exponentiell mit der Temperatur an. Dadurch wird bei höherer Temperatur die Zündbedingung Gl. (5.4) schneller erfüllt.

Bei Bestrahlung der s_p -Zone des Thyristors mit Licht werden gemäß der Mechanismen einer Photodiode (vgl. Abschn. 2.7.6) in dieser Zone zusätzlich Elektron-Loch-Paare optisch generiert, wobei die Minoritätsträgerdichte ansteigt und das Zünden des Thyristors ausgelöst wird (Fotothyristor, vgl. Abschn. 5.4.5.1). Weiterhin kann ein Zünden des Thyristors durch schnelle Spannungsänderungen dV_{AK}/dt hervorgerufen werden, was zumeist als parasitärer Effekt z. B. bei Spannungsspitzen auftritt (vgl. Abschn. 5.6).

5.4.4 Schaltverhalten des Thyristors

Beim Einschalten des Thyristors über einen Zündstrom I_G gemäß Abb. 5.11b müssen sich zunächst die Ladungsträgerverteilungen in der s_n - und s_p -Zone auf die sich neu ergebenen stationären Werte einstellen. Die daraus resultierende Einschaltzeit (*Zündzeit*) t_{gt} setzt sich aus der *Zündverzugszeit* t_{gd} und die durch laterale Ausbreitung des Zündvorganges über die Katodenfläche bedingte *Durchschaltzeit* t_{gr} zusammen:

$$t_{gt} = t_{gd} + t_{gr}. \quad (5.5)$$

Dabei gibt die Zündverzugszeit t_{gd} die Zeit an, die zwischen dem Beginn des Zündimpulses und dem Zeitpunkt vergeht, bei dem die Anoden-Katoden-Spannung V_{AK} auf 90% ihres Wertes abgesunken ist. Die Durchschaltzeit t_{gr} charakterisiert die Zeit, in der die Spannung V_{AK} von 90% auf 10% ihres Wertes abgesunken ist. Das Einschaltverhalten eines Thyristors bei Anlegen eines Zündstroms I_G ab dem Zeitpunkt $t = 0$ ist in Abb. 5.16 dargestellt.

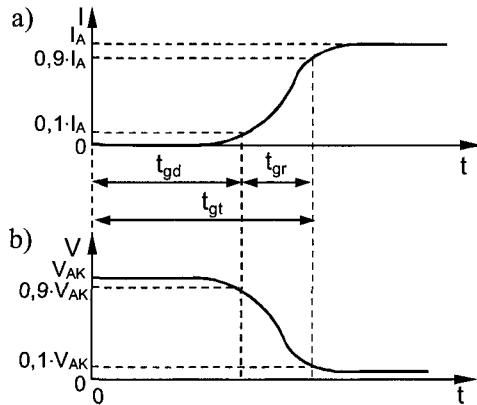


Abb. 5.16. Einschaltverhalten eines Thyristors a) zeitlicher Verlauf des Anodenstroms I_A und b) zeitlicher Verlauf der Anoden-Katoden-Spannung V_{AK} bei Anlegen eines Zündstroms ab dem Zeitpunkt $t = 0$

Zum Ausschalten eines Thyristors muss der Anodenstrom unter den Haltestrom ($I_A < I_H$) verringert werden. Dazu wird entweder der Laststromkreis (vgl. Abb. 5.11b) geöffnet, wobei der Strom über einen anderen Strompfad (z. B. über einen gezündeten Hilfsthryistor) abkommutiert werden muss oder es wird die Anoden-Katoden-Spannung V_{AK} umgepolzt. Für die nachfolgenden Betrachtungen wird der Ausschaltvorgang an Hand der in Abb. 5.17a dargestellten Schaltung mit induktiver Last und Ansteuerung mit einer Rechteckspannung V_{in} beschrieben.

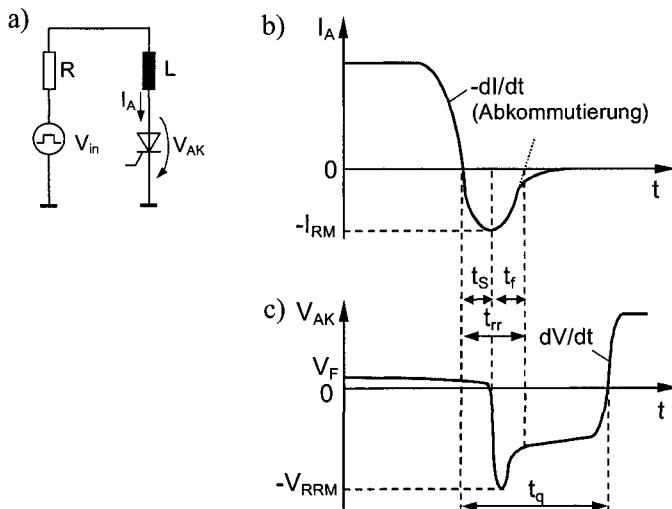


Abb. 5.17. Ausschaltverhalten eines Thyristors bei Abkommutierung a) Messschaltung b) zeitlicher Verlauf des Anodenstroms I_A und b) zeitlicher Verlauf der Anoden-Katoden-Spannung V_{AK}

Beim Ausschaltvorgang sinkt der Anodenstrom I_A des Thyristors näherungsweise linear mit der Abkommunierungssteilheit $-dI/dt$. Die Schaltzeit wird zum einen durch die Rückwärtserholzeit t_{rr} bestimmt, die sich aus der Speicherzeit t_S und der Abfallzeit t_f zusammensetzt. Zum anderen wird die Schaltzeit durch die sog. *Freiwerdezeit* t_q bestimmt, die zur Wiederherstellung der Blockierfähigkeit des Thyristors notwendig ist. Der beim Ausschaltvorgang eines Thyristors auftretende zeitliche Verlauf des Anodenstroms I_A und der Anoden-Katoden-Spannung V_{AK} mit den zugehörigen Schaltzeiten ist in Abb. 5.17b und Abb. 5.17c veranschaulicht. Während des Ausschaltvorganges muss die in der s_p -Zone gespeicherte Ladung durch einen Strom mit dem Maximalwert $-I_{RM}$ abgebaut werden. Bei induktiver Last wird durch den Strom $-I_{RM}$ eine sog. *Rückschlagsspannung* $-V_{RRM}$ hervorgerufen, die schaltungstechnisch begrenzt werden muss.

5.4.5 Weitere Thyristortypen

5.4.5.1 Fotothyristor

Die Zündung eines Thyristors (vgl. Abschn. 5.4.3) kann ebenfalls durch eine Bestrahlung der s_p -Zone mit Licht erfolgen; man spricht hierbei von einem *Fotothyristor*. Dabei werden in dieser Zone zusätzlich Elektron-Loch-Paare optisch generiert, wobei die Minoritätsträgerdichte ansteigt und das Zünden des Thyristors ausgelöst wird. Als Lichtquelle dienen z. B. Lumineszenzdioden (vgl. Abschn. 2.7.8), die Photonen mit genügend hoher Energie emittieren, um Elektron-Loch-Paare zu generieren. Die Lichtquelle kann sich dabei direkt im Thyristorgehäuse befinden oder extern angeordnet sein; bei externer Anordnung wird das Licht über einen Lichtleiter eingekoppelt. Der Vorteil der optischen Zündung liegt in der galvanischen Trennung des Steuerkreises vom Lastkreis.

5.4.5.2 GTO-Thyristor

Der *GTO-Thyristor* (Gate-Turn-Off Thyristor) wird wie ein herkömmlicher Thyristor über einen positiven Zündstromimpuls I_G gezündet. Das Ausschalten erfolgt jedoch mit einem negativen Strompuls am Gate. Der schematische Aufbau und das Schaltsymbol eines GTO-Thyristors ist in Abb. 5.18 dargestellt.

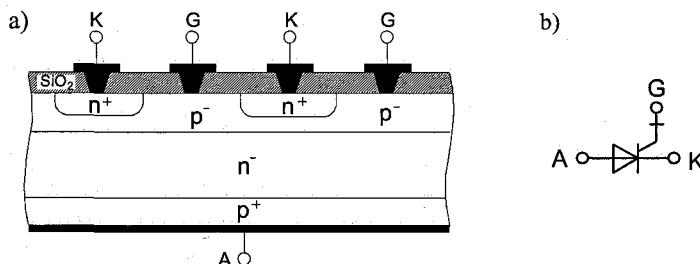


Abb. 5.18. GTO-Thyristor a) schematischer Aufbau (Querschnitt) mit alternierenden Gate- und Katodenanschlüssen b) Schaltsymbol

Beim Ausschalten des GTO-Thyristors über einen negativen Stromimpuls am Gate werden die Überschussladungsträger aus der s_p -Zone abgesaugt. Um beim Ausschalten die von der Basis des pnp-Bipolartransistors T_1 (vgl. Ersatzschaltbild Abb. 5.13) injizierten Löcher sicher abfließen zu lassen, muss der laterale Gatewiderstand des Thyristors reduziert werden. Dazu werden streifenförmig ausgeprägte, alternierende Gate- und Katodenanschlüsse verwendet. In Abb. 5.19 ist das Ausschaltverhalten eines GTO-Thyristors dargestellt. Dabei bleibt nach Einprägen des negativen Gatestromimpulses I_G für die Speicherzeit t_s der Anodenstrom I_A noch erhalten. Während der sich anschließenden Abfallzeit t_f beginnt der Anodenstrom durch das Absaugen von Löchern aus der p^- -Zone zu sinken und die Anoden-Katoden-Spannung V_{AK} steigt stark an.

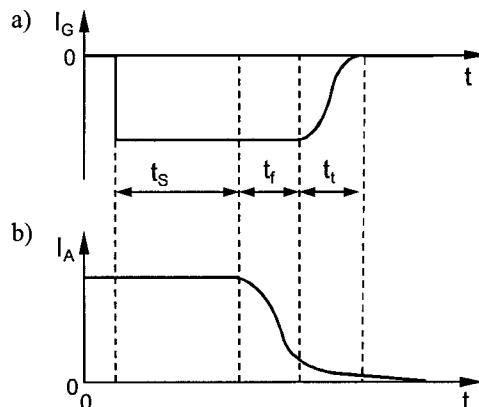


Abb. 5.19. Ausschaltverhalten eines GTO-Thyristors bei negativem Gatestromimpuls
a) Zeitverhalten des Gatestroms I_G und b) Zeitverhalten des Anodenstroms I_A

Nach der Abfallzeit t_f befindet sich der pnp-Bipolartransistor T_1 (vgl. Ersatzschaltung Abb. 5.13) noch im aktiven Bereich. Der resultierende Reststrom fließt während der Abklingzeit t_t über die Gateelektrode ab, da der npn-Bipolartransistor T_2 der Ersatzschaltung bereits gesperrt ist. Da die sog. *Ausschaltstromverstärkung* $-I_A/I_G$ des GTO-Thyristors gering ist, wird zum Ausschalten ein relativ hoher negativer Gatestrom I_G benötigt ($|I_G| \approx 1/3 I_A$).

Mit GTO-Thyristoren werden Sperrspannungen bis ca. 7000 V und Dauerströme bis ca. 3000 A erreicht. Die Schaltfrequenzen liegen dabei im Bereich von 1 kHz bis 10 kHz.

5.4.6 Anwendung des Thyristors

Thyristoren werden insbesondere im mittleren bis hohen Leistungsbereich eingesetzt. Mit ihnen können Sperrspannungen bis ca. 10 kV und Dauerströme bis ca. 4000 A realisiert werden. Kleinere Thyristoren werden vor allem als elektronische Schalter (z. B. Helligkeitsregler) und Drehzahlregler für Drehstrommotoren einge-

setzt. In Abb. 5.20 ist der Einsatz eines Thyristors in einer sog. *Phasenanschnittsteuerung* (Impulssteuerung) dargestellt.

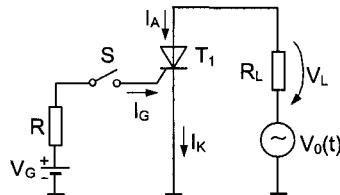


Abb. 5.20. Phasenanschnittsteuerung (Impulssteuerung) mit Thyristor

Dabei befindet sich der Thyristor T_1 in Serie zur ohmschen Last R_L . Bei der Phasenanschnittsteuerung findet dann während jeder positiven Halbwelle der Wechselspannung $V_0(t)$ ein Stromfluss I_A durch den Lastwiderstand R_L statt. Dabei fließt der Strom jeweils vom Zündzeitpunkt bis zum Nulldurchgang der Wechselspannung. Bei einer negativen Halbwelle der Wechselspannung sperrt der Thyristor ($I_A = 0$). Der aus dem Zeitverlauf des Gatestroms I_G resultierende Spannungsverlauf $V_L(t)$ über dem Lastwiderstand R_L ist in Abb. 5.21 dargestellt.

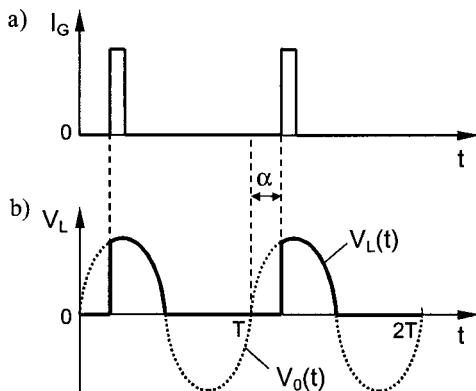


Abb. 5.21. a) Zeitverhalten des Gatestroms I_G und b) Zeitverhalten der Spannung über der ohmschen Last R_L einer Phasenanschnittsteuerung

Der Winkel zwischen dem Nulldurchgang der Wechselspannung $V_0(t)$ und dem Zünden der Thyristoren wird als *Phasenanschnittwinkel* α bezeichnet. Durch Verändern des Phasenanschnittwinkels kann der Mittelwert der Wechselspannung am Lastwiderstand R_L variiert werden. Die Phasenanschnittsteuerung ermöglicht somit eine stufenlose Leistungssteuerung bei geringer Reaktionszeit. Durch antiparallele Verschaltung von zwei Thyristoren kann auch die negative Halbwelle der Wechselspannung $V_0(t)$ genutzt werden.

Ein weiteres Anwendungsgebiet für Thyristoren stellt die Drehstrombrückenschaltung zur Ansteuerung von Gleichstrommotoren dar. Diese Schaltung wird in Zusammenhang mit der Anwendung des IGBTs in Abschn. 5.5.5 näher betrachtet.

5.5 IGBT

5.5.1 Aufbau und Wirkungsweise des IGBTs

Der *IGBT* (Insulated-Gate Bipolar-Transistor) weist einen ähnlichen Aufbau wie der Leistungs-MOSFET auf (vgl. Abschn. 5.3), besitzt aber anstelle einer n^+ -Silizium-Schicht eine p^+ -Silizium-Schicht unterhalb der n^- -Driftzone. Dadurch benötigt der IGBT nur eine einfache Ansteuerung analog zum MOSFET, weist aber die hohe Stromergiebigkeit eines Bipolartransistors auf. Der Steueranschluss des IGBTs wird analog zum MOSFET als Gate (G), die Ausgangselektroden analog zum Bipolartransistor mit Emitter (E) und Kollektor (C) bezeichnet. Der schematische Aufbau eines IGBT-Elements ist in Abb. 5.22a im Querschnitt dargestellt; in Abb. 5.22b ist die Parallelschaltung von zwei Elementen zur Erhöhung der Stromergiebigkeit abgebildet.

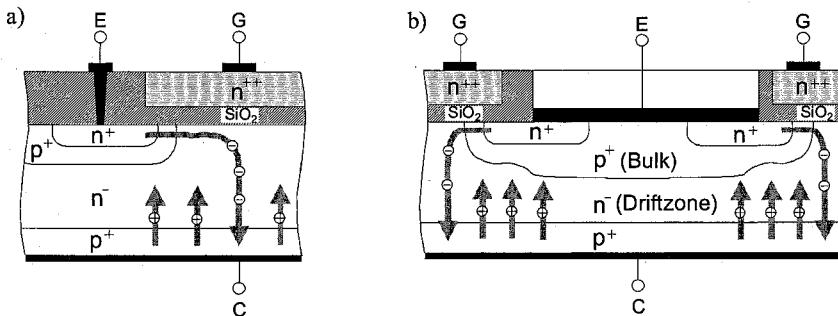


Abb. 5.22. Schematischer Aufbau (Querschnitt) a) eines IGBT-Elements und b) Parallelschaltung von zwei Elementen zur Erhöhung der Stromergiebigkeit

In der n^- -Driftzone des IGBTs (vgl. Abb. 5.22) erfolgt ein bipolarer Stromtransport, da aus der darunter liegenden p^+ -Silizium-Schicht Löcher in die Driftzone emittiert werden können. Die Löcherinjektion aus dem p^+n^- -Übergang erhöht analog zur psn-Leistungsdiode (vgl. Abschn. 5.1) die Ladungsträgerdichte von Elektronen und Löchern (starke Injektion). Dadurch wird der On-Widerstand R_{on} des IGBTs im Durchlassbetrieb gegenüber einem sonst identisch aufgebauten Leistungs-MOSFET (vgl. Abschn. 5.3) erheblich reduziert. Der p^+n^- -Übergang des IGBTs stellt gleichzeitig die Emitter-Basis-Strecke eines pnp-Bipolartransistors dar, dessen Kollektor der Substratanschluss der internen MOSFET-Struktur des IGBTs ist. Fließt durch den pnp-Bipolartransistor ein Kollektorstrom, so trägt dieser ebenfalls zum Durchlassstrom bei. Dadurch kann der On-Widerstand des IGBTs weiter verringert werden.

Die Wirkungsweise des IGBTs im Durchlassbetrieb ($V_{CE} \geq 0$) kann anhand des Ersatzschaltbildes in Abb. 5.23a veranschaulicht werden. Demnach kann der IGBT durch die Verschaltung eines n-Kanal-MOSFETs T_1 und eines pnp-Bipolartransistors T_2 beschrieben werden. Wie der Leistungs-MOSFET enthält auch der

IGBT einen parasitären lateralen npn-Bipolartransistor T_3 , der sich parallel zum MOSFET T_1 befindet (vgl. Abschn. 5.3, Abb. 5.7). Als weiteres parasitäres Element wirkt der Bahnwiderstand R_B des p⁺-Bulkgebietes zwischen n⁺-Diffusionsgebiet (Emitter) und n⁻-Driftzone (Basis von T_2).

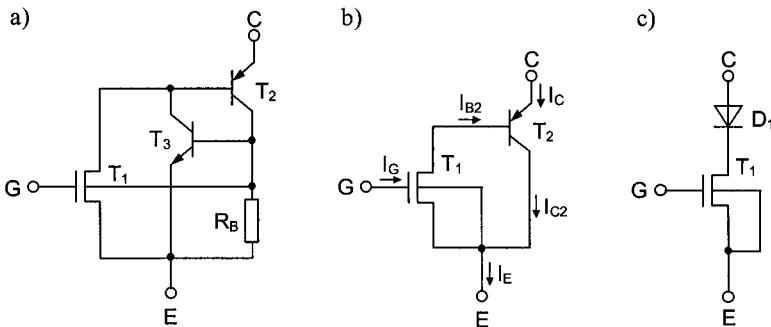


Abb. 5.23. a) Ersatzschaltbild eines IGBTs im Durchlassbetrieb ($V_{CE} \geq 0$) b) vereinfachtes Ersatzschaltbild im Durchlassbereich für $R_B \rightarrow 0$ c) vereinfachtes Ersatzschaltbild im Schwellenbereich ($0 \leq V_{CE} < V_S$)

Wenn der Bahnwiderstand R_B gering ist ($R_B \rightarrow 0$), resultiert daraus ein Kurzschluss zwischen Basis und Emitter des parasitären npn-Bipolartransistors T_3 , so dass dieser inaktiv bleibt und im Ersatzschaltbild vernachlässigt werden kann. Das resultierende vereinfachte Ersatzschaltbild ist in Abb. 5.23b dargestellt. Der enthaltene Stromanteil I_{B2} entsteht durch den Elektronenfluss im Kanal des n-Kanal-MOSFETs T_1 , wobei ein Teil der Elektronen in der n⁻-Driftzone rekombiniert. Der Stromanteil I_{B2} dient gleichzeitig als Basisstrom des pnp-Bipolartransistors T_2 , dessen Kollektorstrom I_{C2} sich aus der Stromverstärkung β_2 und dem Stromanteil I_{B2} wie folgt zusammensetzt:

$$I_{C2} = \beta_2 \cdot I_{B2}. \quad (5.6)$$

Da die Basisweite von T_2 groß ist, ist die Stromverstärkung $\beta_2 < 1$. Der Kollektorstrom I_C des IGBTs berechnet sich dann mit Hilfe der Stromanteile I_{B2} und I_{C2} wie folgt:

$$I_C = I_{C2} + I_{B2} \quad (5.7)$$

Wird der IGBT im Schwellenbereich ($0 \leq V_{CE} < V_S$) betrieben, so ist der pn-Übergang zwischen Basis und Kollektor des pnp-Bipolartransistors T_2 nur schwach leitend und die Stromverstärkung β_2 ist sehr gering. Dadurch kann die Basis-Kollektor-Diode des pnp-Bipolartransistors T_2 vernachlässigt werden; es ergibt sich das in Abb. 5.23c dargestellte vereinfachte Ersatzschaltbild aus der Serienschaltung des n-Kanal-MOSFETs T_1 und der Diode D_1 . Wird die Schleusenspannung V_S der Diode D_1 überschritten, gelangt man in den aktiven Bereich des IGBTs (vgl. Abschn. 5.5.2).

Bei Sperrpolung des IGBTs ($V_{DS} < 0$) ist die Diode zwischen Kollektor und n-Driftzone gesperrt. Somit ist der IGBT in Rückwärtspolung sperrfähig und besitzt keine interne Rückwärtsdiode wie der Leistungs-MOSFET.

5.5.2 Kennlinie des IGBTs

Die Abhängigkeit des Kollektorstroms I_C von der Kollektor-Emitter-Spannung V_{CE} wird im Ausgangskennlinienfeld des IGBTs dargestellt. Im Gegensatz zum Leistungs-MOSFET zeigt der IGBT im Schwellenbereich ein diodentypisches Verhalten, d. h. erst ab einer bestimmten Schwellspannung $V_{CE} \geq V_S$ nimmt der Kollektorstrom I_C merklich mit der Kollektor-Emitter-Spannung V_{CE} zu. In Abb. 5.24 ist ein typischer Verlauf des Ausgangskennlinienfeldes und das Schaltsymbol des IGBTs abgebildet.

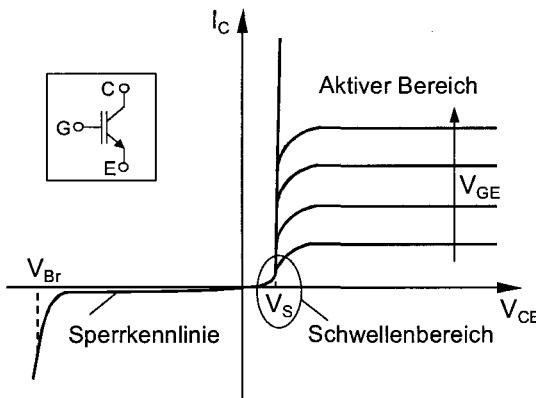


Abb. 5.24. Schaltsymbol und Ausgangskennlinienfeld eines IGBTs für verschiedene Gate-Emitter-Spannungen V_{GE}

Bei geringer Kollektor-Emitter-Spannung V_{CE} wird gemäß der in Abb. 5.23c dargestellten Serienschaltung von MOSFET T_I und p⁺n-Diode D_I der Kollektorstrom I_C des IGBTs durch die Diode begrenzt. Bei höherer Kollektor-Emitter-Spannung $V_{CE} \geq V_S$ geht die Diodenkennlinie in die Ausgangskennlinie des Transistors über. Der Kollektorstrom I_C des IGBTs berechnet sich dann gemäß Gl. (5.7).

In Sperrpolung ($V_{DS} < 0$) verhält sich der IGBT wie eine herkömmliche pn-Diode (vgl. Abschn. 2.2.2). Ist die Sperrspannung größer als die Durchbruchspannung $|V_{CE}| > V_{Br}$, so findet ein starker Anstieg des Sperrstroms statt.

5.5.3 Schaltverhalten des IGBTs

Das Einschaltverhalten des IGBTs verhält sich nahezu analog zum Einschaltverhalten des Leistungs-MOSFETs (vgl. Abschn. 5.3). In Abb. 5.25 wird zum Zeit-

punkt $t = 0$ wird ein positiver Spannungssprung V_{GE} am Gate des IGBTs erzeugt, wobei mit Hilfe einer Spannungsteuerung am Gate das Auftreten eines Miller-Plateaus (vgl. Abschn. 5.3) vermieden wird. Nach Erreichen der Einsatzspannung $V_{GE} \geq V_{Tn}$ steigt der Kollektorstrom I_C des IGBTs stark an. Auf Grund der hohen, nahezu konstanten Kollektor-Emitter-Spannung V_{CE} befindet sich der IGBT hierbei im Sättigungsbereich. Nach der Anstiegszeit t_r erreicht der Strom I_{B2} seinen Endwert und die Kollektor-Emitter-Spannung V_{CE} sinkt auf einen niedrigen Wert V_{CESat} ab. Der Kollektorstrom I_{C2} durch den pnp-Bipolartransistor steigt erst an, nachdem der Strom I_{B2} fließt. Das zeitliche Verhalten der Kollektor-Emitter-Spannung V_{CE} und des Kollektorstroms I_C beim Einschaltverhalten eines IGBTs bei ohmscher Last und eingeprägter Spannungsrampe am Gate ist in Abb. 5.25 dargestellt.

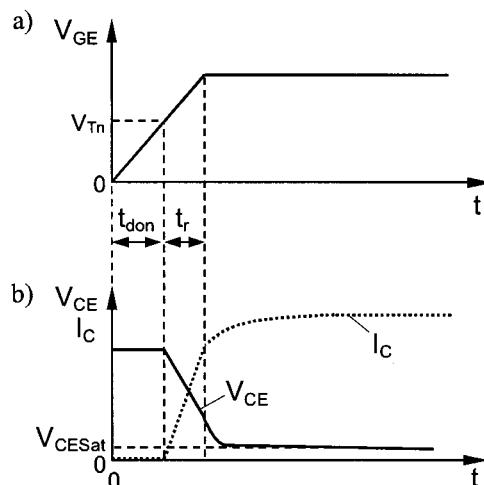


Abb. 5.25. Einschaltverhalten eines IGBTs bei ohmscher Last und eingeprägter Spannungsrampe am Gate a) Zeitverhalten der Gate-Emitter-Spannung V_{GE} und b) Zeitverhalten der Kollektor-Emitter-Spannung V_{CE} und des Kollektorstroms I_C

Beim Ausschalten des IGBTs wird die Gate-Emitter-Spannung V_{GE} auf den Wert $V_{GE} = 0$ reduziert. Nach Unterschreiten der Einsatzspannung ($V_{GE} < V_{Tn}$) beginnt der Kollektorstrom I_C mit ähnlicher Geschwindigkeit wie beim Leistungs-MOSFET zu sinken. Anders als beim Leistungs-MOSFET sind aber beim IGBT sowohl Elektronen als auch Löcher als Diffusionsladung der p⁺n-Diode in der n⁻Driftzone gespeichert. Bei abruptem Absinken der Gate-Emitter-Spannung V_{GE} beim Ausschaltvorgang können die Elektronen nicht über den Inversionskanal abfließen. Der Löcherstrom I_{C2} (vgl. Abb. 5.23b) und damit der Kollektorstrom I_C des IGBTs fließt dann so lange weiter, bis die Diffusionsladung (sog. *Tail-Ladung* Q_T) durch Rekombination in der n⁻-Driftzone abgebaut ist. Die hierbei auftretende Zeitkonstante τ_T kann über die Stromverstärkung β_2 des internen pnp-Bipolartransistors (vgl. Abb. 5.23b) und seine Basistransitzeit τ_{B2} (vgl. Abschn. 3.1.3) wie folgt bestimmt werden:

$$\tau_T = \beta_2 \cdot \tau_{B2} \quad \text{mit} \quad \tau_{B2} = \frac{w_0^2}{2 \cdot D_p}. \quad (5.8)$$

Das Ausschaltverhalten eines IGBTs ist anhand des Zeitverhaltens der Gate-Emitter-Spannung V_{GE} , der Kollektor-Emitter-Spannung V_{CE} und des Kollektorstroms I_C mit Auftreten der Tail-Ladung Q_T mit der Zeitkonstanten τ_T in Abb. 5.26 dargestellt.

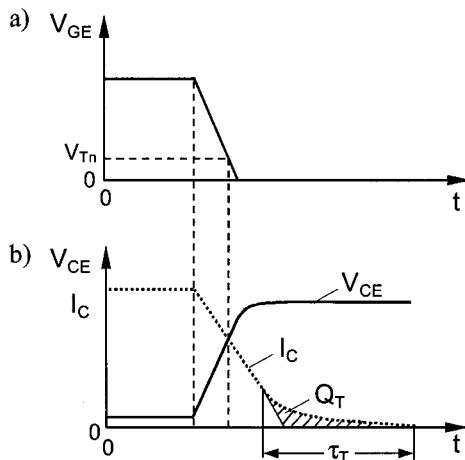


Abb. 5.26. Ausschaltverhalten eines IGBTs a) Zeitverhalten der Gate-Emitter-Spannung V_{GE} und b) Zeitverhalten der Kollektor-Emitter-Spannung V_{CE} und des Kollektorstroms I_C mit Auftreten der Tail-Ladung Q_T

Das langsame Abklingen des Löcherstroms über den durch die Tail-Ladung Q_T hervorgerufenen sog. *Tail-Strom* erfolgt in einem Zeitabschnitt, in dem bereits eine relativ hohe Sperrspannung am IGBT anliegt. Dadurch tritt eine erhöhte Verlustleistung während des Ausschaltvorganges auf. Durch den Tail-Strom wird darüber hinaus die maximale Schaltfrequenz des IGBTs reduziert (vgl. Abschn. 5.5.5).

5.5.4 Weitere IGBT-Typen

5.5.4.1 Punch-Through-IGBT

Beim *Punch-Through-IGBT (PT-IGBT)* wird die n^- -Driftzone durch eine n^+ -Silizium-Schicht von der darunter liegenden p^+ -Silizium-Schicht separiert. Auf Grund des dadurch geänderten Feldstärkeverlaufs kann die Driftzone schmäler und niedriger dotiert ausgelegt werden. Durch diese Maßnahme wird der Durchlasswiderstand R_{on} reduziert. In Abb. 5.27 ist der schematische Aufbau eines PT-IGBTs im Querschnitt dargestellt.

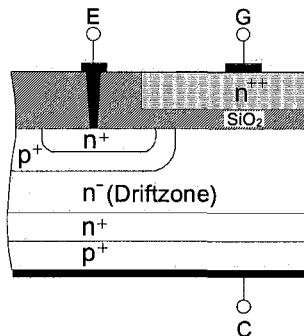


Abb. 5.27. Schematischer Aufbau eines Punch-Through-IGBTs im Querschnitt

Bei hoher Sperrspannung tritt hier der Punch-Through-Effekt (vgl. Abschn. 3.3.3) auf, d. h. die Sperrsicht ist über die n⁻-Driftzone bis zur n⁺-Silizium-Schicht ausgedehnt. Der Feldstärkeverlauf ist dann nahezu rechteckig, weil das elektrische Feld in der darunter liegenden n⁺-Silizium-Schicht schnell mit geringer örtlicher Ausdehnung abnimmt. Der Punch-Through-IGBT besitzt einen geringen Durchlasswiderstand, ist jedoch nicht für so hohe Sperrspannungen geeignet wie der herkömmliche IGBT.

5.5.4.2 Trench-IGBT

Beim *Trench-IGBT* wird im Gegensatz zum herkömmlichen IGBT eine senkrecht positionierte Gate-Elektrode verwendet (Grabenstruktur). Im eingeschalteten Zustand ($V_{GE} \geq V_{Th}$) bilden sich jeweils neben dem Gate im p-Bereich senkrechte Inversionskanäle aus. Der schematische Aufbau eines Trench-IGBTs mit ausgebildeten Inversionskanälen ist in Abb. 5.28 im Querschnitt dargestellt.

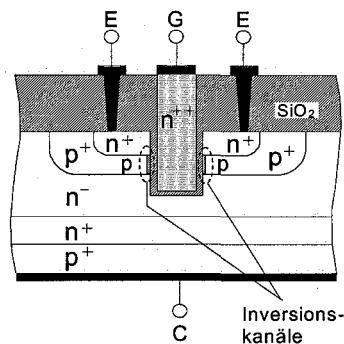


Abb. 5.28. Schematischer Aufbau (Querschnitt) eines Trench-IGBTs mit ausgebildeten Inversionskanälen

Die Vorteile des Trench-IGBTs sind ein niedriger Durchlasswiderstand R_{on} , ein geringer Flächenbedarf und ein günstiger Verlauf der elektrischen Feldstärke.

5.5.5 Anwendung des IGBTs

Der IGBT wird insbesondere im geringen bis mittleren Leistungsbereich bei hohen Blockierspannungen eingesetzt. Es werden Blockierspannungen bis $V_{CE} \approx 3000$ V und Dauerströme bis $I_C \approx 100$ A erreicht. Zur Erhöhung der Stromerhöhung werden mehrere IGBTs parallel geschaltet und zusammen mit Freilaufdioden (vgl. Abschn. 5.3, Abb. 5.7) in einem Modul integriert. Da der Tail-Strom (vgl. Abschn. 5.5.3) die maximale Schaltfrequenz des IGBTs gegenüber dem Leistungs-MOSFET reduziert, werden Schaltfrequenzen bis maximal ca. 100 kHz erreicht.

Ein wichtiges Anwendungsgebiet des IGBTs ist die getaktete Ansteuerung induktiver Lasten z. B. in Spannungswandlern, Frequenzkonvertern oder Motorsteuerungen. Dabei wird beim Ein- bzw. Ausschalten des IGBTs der Laststrom zwischen der Freilaufdiode und dem IGBT kommutiert.

In Abb. 5.28 ist beispielhaft eine Einphasen-Brückenschaltung mit IGBTs zur Ansteuerung eines Gleichstrommotors M_1 dargestellt.

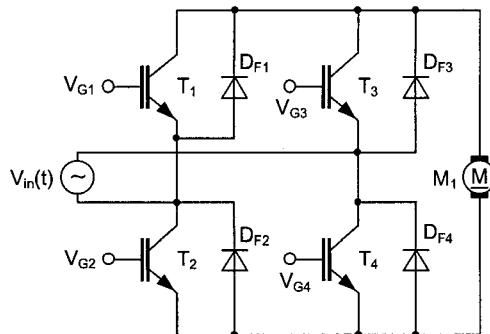


Abb. 5.29. Prinzip einer Einphasen-Brückenschaltung mit IGBTs zur Ansteuerung eines Gleichstrommotors

Der Gleichstrom im Lastkreis entsteht durch abschnittsweises Durchschalten der Wechselspannungsquelle auf die Gleichspannungsseite, dass durch geeignete Ansteuerung der IGBTs T_1 bis T_4 über die zugehörigen Gatespannungen V_{G1} bis V_{G4} realisiert wird. Die Freilaufdioden D_{F1} bis D_{F4} nehmen beim Ausschalten der IGBTs den Strom auf, der durch die induktive Last nahezu konstant gehalten wird. Sie müssen so dimensioniert werden, dass sie in Flussrichtung den durch die Induktivität fließenden Strom, der vor dem Abschalten geflossenen ist, aufnehmen können.

5.6 Latch-up-Effekt

Der *Latch-up-Effekt* kennzeichnet das unerwünschte Zünden parasitärer Thyristorstrukturen z. B. im IGBT oder in CMOS-Schaltungen. Der IGBT (vgl.

Abschn. 5.5) besitzt in vertikaler Richtung eine $n^+p^+n^-p^+$ -Schichtfolge, die eine Thyristorstruktur (vgl. Abschn. 5.4) darstellt, welche unter bestimmten Randbedingungen zünden kann. Dabei stellt der n^+p^+ -Übergang die Steuerstrecke zwischen Kollektor und p-Bulkgebiet dar. Wird der IGBT in starker Sättigung betrieben ($V_{CE} \gg V_{CESat}$), so werden am kollektorseitigen Ende des Inversionskanals auf Grund des einsetzenden Lawinendurchbruchs (vgl. Abschn. 2.2.1) Elektron-Loch-Paare generiert. Daraus resultiert ein Löcherfluss über das p^+ -Gebiet zum Emitter, wodurch der parasitäre Thyristor gezündet werden kann. Der Latch-up-Effekt beim IGBT tritt insbesondere beim Ausschaltvorgang des IGBTs auf, kann aber auch bei schnellem Anstieg der Sperrspannung durch ein dV/dt -Zünden der parasitären Thyristorsstruktur erfolgen (vgl. Abschn. 5.4.3). Dem Latch-up-Effekt im IGBT kann z. B. durch die Wahl eines niedrigen Bahnwiderstandes R_b des Bulkgebietes entgegengewirkt werden. Deshalb werden die p-Zonen außerhalb des Kanalgebietes p^+ -dotiert (vgl. Abb. 5.28).

Der Latch-up-Effekt in CMOS-Schaltungen kann an Hand der parasitären n^+pnp^+ -Thyristorstruktur eines CMOS-Inverters (vgl. Abschn. 4.8.2) beschrieben werden, wobei der p-Kanal-MOSFET in einer n-Wanne im p-Substrat eingebettet ist. Wird diese Thyristorstruktur leitend, so fließt ein hoher Strom zwischen Betriebsspannung und Masse des Inverters, wodurch eine thermische Überlastung und meist die Zerstörung des Inverters resultiert. Der entstehende Strompfad zwischen Betriebsspannung V_{DD} und Masse V_{SS} ist im Schnittbild eines CMOS-Inverters in Abb. 5.30 veranschaulicht.

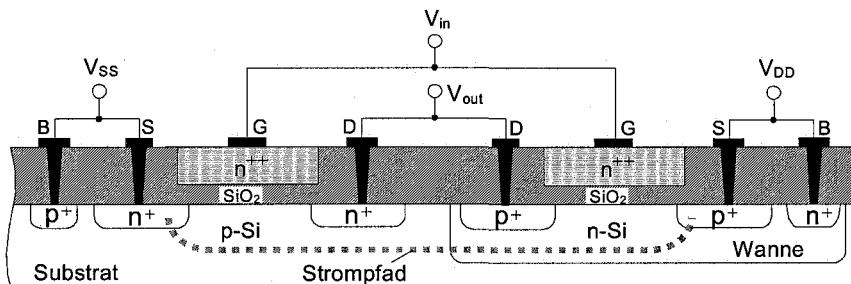


Abb. 5.30. Querschnitt eines CMOS-Inverters in n-Wannen-Technologie mit Strompfad einer parasitären Thyristorstruktur

Zündursachen der parasitären Thyristorstruktur sind vor allem das Über- oder Unterschwingen der Ausgangsspannung V_{out} des CMOS-Inverter, wodurch Ladungsträger durch den dann kurzzeitig leitenden pn-Übergang am Drain injiziert werden. Weiterhin führen hohe Störspannungsspitzen auf der Versorgungsspannung zum Zünden der Thyristorstruktur, falls die Kippspannung V_{BF} überschritten wird (vgl. Abschn. 5.4.3). Maßnahmen gegen den Latch-up-Effekt in CMOS-Schaltungen sind insbesondere die Wahl eines ausreichenden Abstandes des n^+ -bzw. p^+ -Draingegebietes von der Wannengrenze oder die Verwendung von in Sperrrichtung gepolten pn-Übergängen, die die injizierenden Gebiete ringförmig umgeben (sog. *Guard-Ringe*).

6 Halbleitersensoren

Sensoren ermöglichen die Umwandlung von physikalischen, chemischen oder biologischen Größen in ein äquivalentes elektrisches Messsignal. Somit stellen Sensoren die Schnittstelle zwischen Umgebung und elektronischer Schaltung dar. Die von einem Sensorelement erzeugten elektrischen Signale können über eine Messschaltung detektiert, verstärkt und in ein normiertes analoges Messsignal umgewandelt werden. Das Sensorelement und die Messschaltungen bilden zumeist eine abgeschlossene Einheit. Durch die Umwandlung des normierten analogen Messsignals in ein Digitalsignal mit Hilfe eines Analog-Digital-Wandlers kann eine digitale Weiterverarbeitung des Sensorsignals erfolgen.

In diesem Kapitel werden die wesentlichen halbleiterbasierten Sensorprinzipien an Hand der dominierenden Halbleitereffekte beschrieben; Messschaltungen zur Sensorsignal konditionierung und ausgewählte Anwendungsbeispiele werden dargestellt. Neben Halbleitersensoren zur Temperatur- und Magnetfeldmessung werden Halbleitersensoren zur Detektion von optischen, chemischen und mechanischen Größen behandelt.

6.1 Temperatursensoren

Zur Temperaturmessung können verschiedene Halbleiter-Effekte, wie die Temperaturabhängigkeit der Ladungsträgerbeweglichkeit oder der Seebeck-Effekt (vgl. Kap. 1) ausgenutzt werden. Die wesentlichen Vertreter der auf diesen Effekten basierenden Sensoren sind *Grenzflächen-Temperatursensoren*, *Spreading-Resistance Temperatursensoren* und *Thermopiles*, die in den folgenden Abschnitten hinsichtlich Aufbau, Funktionsweise und Anwendung beschrieben werden.

6.1.1 Grenzflächen-Temperatursensoren

Grenzflächentemperatursensoren nutzen die Temperaturabhängigkeit des Ladungsträgertransports durch in Flussrichtung betriebene pn-Übergänge von Dioden oder Transistoren. Typischer Weise findet ein als Diode geschalteter Bipolartransistor (vgl. Kap. 3) Anwendung, bei dem der Kollektor und die Basis kurzgeschlossen sind (Abb. 6.1a). Die Temperaturabhängigkeit der Basis-Emitter-Spannung $V_{BE}(T)$ wird dabei zur Temperaturmessung genutzt. Eine typische Spannungs-Temperatur-Kennlinie eines Grenzflächen-Temperatursensors ist in Abb. 6.1b dargestellt.

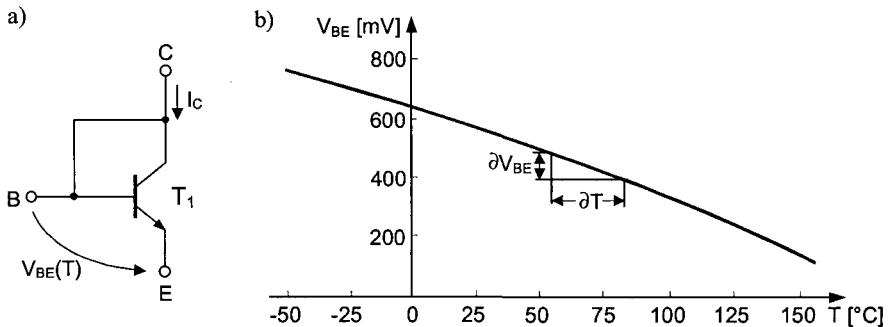


Abb. 6.1. Bipolartransistor als Grenzflächen-Temperatursensor a) Schaltungsprinzip und b) typische Spannungs-Temperatur-Kennlinie

Aus Gl. (3.13), Abschn. 3.2.1 kann die temperaturabhängige Basis-Emitter-Spannung $V_{BE}(T)$ des als Temperatursensor betriebenen Bipolartransistors bei eingeprägtem Kollektorstrom I_C näherungsweise wie folgt berechnet werden:

$$V_{BE}(T) \approx \frac{kT}{e} \cdot \ln\left(\frac{I_C}{I_S(T)}\right), \quad (6.1)$$

wobei $I_S(T)$ den temperaturabhängigen Sperrsättigungsstrom des pn-Überganges gemäß Gl. (2.27), Abschn. 2.1.3.3 beschreibt. Die Temperaturabhängigkeit von I_S ist dabei durch die Proportionalität der Minoritätsträgerdichten p_{n0} und n_{p0} zur temperaturabhängigen Eigenleitungsträgerdichte $n_i^2(T)$ begründet.

Trotz des nichtlinearen Zusammenhangs von V_{BE} und der Temperatur besitzt die Basis-Emitter-Spannung bei konstantem Kollektorstrom insbesondere im Temperaturbereich $-50^\circ\text{C} \leq T \leq +150^\circ\text{C}$ eine relativ hohe Linearität. Der Temperaturgradient $\partial V_{BE}/\partial T$ bei konstantem Kollektorstrom I_C ist bei Verwendung eines npn-Transistors negativ und berechnet sich nach [6.2] mit Hilfe des zu $T = 0$ extrapolierten Bandabstandes E_{G0} wie folgt:

$$\left. \frac{\partial V_{BE}}{\partial T} \right|_{I_C=\text{konst.}} = \frac{\frac{E_{G0}}{e} - V_{BE}}{T}. \quad (6.2)$$

Der Temperaturgradient ist in Abb. 6.1b veranschaulicht. Er ist stark von Exemplarstreuungen abhängig, die sich im wesentlichen durch technologiespezifische Abweichungen auf den Sperrsättigungsstrom $I_S(T)$ auswirken. Daraus resultieren für unterschiedliche Transistoren bei gleicher Temperatur und gleichem Kollektorstrom unterschiedliche Basis-Emitter-Spannungen.

Die in Abb. 6.2 dargestellte Schaltung dient dazu, eine Spannung proportional zur Temperatur zu erzeugen, die neben der Temperatur nur vom Flächenverhältnis der Transistoren T_1 und T_2 abhängig ist. Diese Schaltung wird insbesondere zur Temperaturmessung in integrierten Schaltungen eingesetzt.

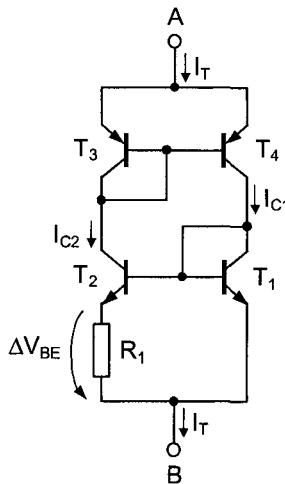


Abb. 6.2. Schaltungsprinzip zur Temperaturmessung in integrierten Schaltungen auf Basis eines Grenzflächen-Temperatursensors (Bipolartransistor als Diode)

In Schaltung Abb. 6.2 ermöglichen die Transistoren T_1 und T_2 die Temperaturmessung. Der aus den flächengleichen pnp-Transistoren T_3 und T_4 bestehende Stromspiegel (vgl. Abschn. 3.7.2) bewirkt die Aufteilung des Gesamtstroms I_T in zwei gleichgroße Teilströme I_{C1} und I_{C2} , die als Kollektorströme der Transistoren T_1 und T_2 dienen. Die Flächen A_1 und A_2 der Transistoren T_1 und T_2 sind im Verhältnis von $A_1/A_2 > 1$ (z. B. $A_1/A_2 = 10$) dimensioniert, woraus für die temperaturabhängige Basis-Emitter-Spannungsdifferenz $\Delta V_{BE}(T)$ folgt:

$$\Delta V_{BE}(T) = \frac{kT}{e} \cdot \ln\left(\frac{I_{C2}}{I_S(T) \cdot A_2}\right) - \frac{kT}{e} \cdot \ln\left(\frac{I_{C1}}{I_S(T) \cdot A_1}\right) = \frac{kT}{e} \cdot \ln\left(\frac{A_1}{A_2}\right). \quad (6.3)$$

Aus dieser über dem Widerstand R_1 abfallenden, temperaturproportionalen Spannungsdifferenz resultiert der zur absoluten Temperatur proportionale Strom I_{C2} . Durch die symmetrischen Eigenschaften des Stromspiegels ist demnach auch der Gesamtstrom I_T zur absoluten Temperatur proportional.

6.1.2 Spreading-Resistance Si-Temperatursensoren

Spreading-Resistance Si-Temperatursensoren arbeiten nach dem Prinzip des radia- len Ausbreitungswiderstandes, welches die Ladungsträgerbeweglichkeit im Halbleiter zur Temperaturmessung ausnutzt. Bei Einschnürung der Stromlinien an einem Kontaktloch wird der temperaturabhängige ohmsche Widerstand $R_T(T)$ des Temperatursensors bei Vernachlässigung des Kontaktwiderstandes zwischen Metallisierung und Halbleiter durch den lokalen spezifischen Widerstand $\rho(T)$ bestimmt. Der Aufbau eines planaren Spreading-Resistance Si-Temperatursensors

mit schematischer Darstellung des am Kontaktloch eingeschnürten Stromverlaufes ist in Abb. 6.3 dargestellt.

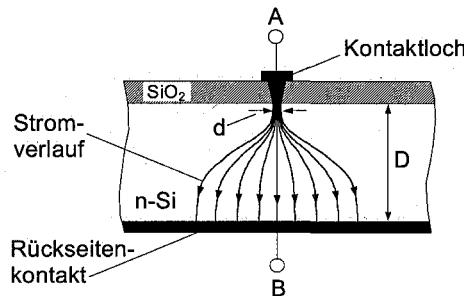


Abb. 6.3. Aufbau und Stromverlauf eines planaren Spreading-Resistance Si-Temperatursensors

Der temperaturabhängige ohmsche Widerstand $R_T(T)$ eines Spreading-Resistance Si-Temperatursensors kann über die Beziehung

$$R_T(T) = \frac{\rho(T)}{\pi \cdot d} \quad (6.4)$$

bestimmt werden [6.2], wobei $\rho(T)$ den dotierungsdichteabhängigen spezifischen Widerstand und d den Durchmesser des Kontaktloches bezeichnet. Gl. (6.4) gilt für die Annahme, dass der Kontaktlochdurchmesser viel kleiner als die Dicke D des Substrats ist. Die Temperaturabhängigkeit des spezifischen Widerstandes eines Spreading-Resistance Si-Temperatursensors mittlerer Dotierungsdichte ist in Abb. 6.4 dargestellt.

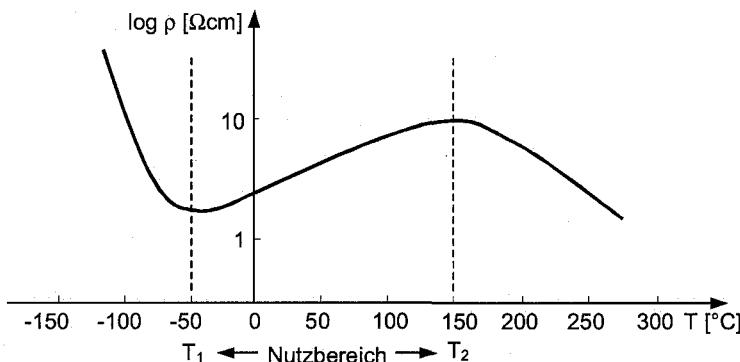


Abb. 6.4. Typischer Verlauf des temperaturabhängigen spezifischen Widerstandes eines Spreading-Resistance Si-Temperatursensors mittlerer Dotierungsdichte

Bei Temperaturen $T < T_1$ sind die Dotieratome des Halbleiters (hier n-Halbleiter) nur unvollständig ionisiert ($n = N_D^+ \neq N_D$). Dadurch steigt die Ladungsträ-

gerdichte n exponentiell mit der Temperatur an. Bis zum Eintritt der vollständigen Ionisation der Dotieratome nimmt der spezifische Widerstand des Halbleiters ab. Bei Temperaturen $T_1 \leq T \leq T_2$ (Nutzbereich) sind alle Dotieratome vollständig ionisiert und die Ladungsträgerbeweglichkeit μ dominiert die Temperaturabhängigkeit des Widerstandes. Durch thermische Gitterstreuung sinkt die Ladungsträgerbeweglichkeit ab, woraus ein Anstieg des spezifischen Widerstandes bei Temperaturerhöhung resultiert. Bei Temperaturen $T > T_2$ herrscht im Halbleiter Eigenleitung vor. Durch thermische Generation von Elektron-Loch-Paaren nimmt die Ladungsträgerdichte bei Temperaturerhöhung zu, was ein Absinken des spezifischen Widerstands bewirkt.

Die praktische Realisierung eines Spreading-Resistance Si-Temperatursensors erfolgt zumeist durch symmetrische Anordnung von zwei antiseriell geschalteten Strukturen (Abb. 6.5). Durch diese Anordnung wird eine Polungshängigkeit infolge der unterschiedlichen Wirkung der Kontaktflächen von Vorder- und Rückseitenkontakt vermieden.

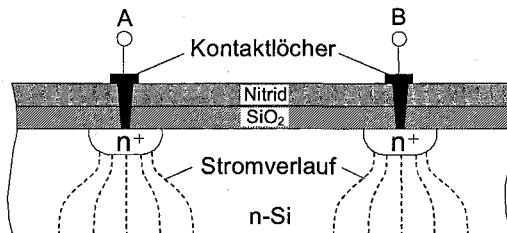


Abb. 6.5. Symmetrische Realisierungsform eines Spreading-Resistance Si-Temperatursensors nach [6.2]

Die Auswirkungen von Fertigungsstreuungen durch Schwankungen der Dotierungsdichte oder geometrische Abweichungen können durch Anpassung der Dotierungsdichte der n^+ -Kontakt-Diffusionsgebiete [6.3] verringert werden. Weiterhin ist eine Kompensation der Auswirkungen von Fertigungsstreuungen durch Verwendung zusätzlicher Abgleichwiderstände möglich.

An den Kontaktlöchern der Anordnung ergeben sich hohe Stromdichten. Um eine starke Eigenerwärmung ΔT_E des Sensors zu vermeiden, sind nur geringe Messströme zulässig. Die Eigenerwärmung berechnet sich über die Leistungsaufnahme P des Sensors und den Temperaturleitwert G_T :

$$\Delta T_E = \frac{P}{G_T}. \quad (6.5)$$

Da die Kennlinie des Spreading-Resistance Si-Temperatursensors auch im Nutzbereich ein nichtlineares Verhalten aufweist (vgl. Abb. 6.4), ist eine Linearisierung beispielsweise durch Verwendung eines Parallelwiderstandes zum Sensor-element möglich. Eine Linearisierungsschaltung mit Sensoransteuerung durch einen konstanten Steuerstrom I_M , der über eine Konstantstromquelle gemäß Abschn. 4.8.4 bereitgestellt werden kann, ist in Abb. 6.6 dargestellt.

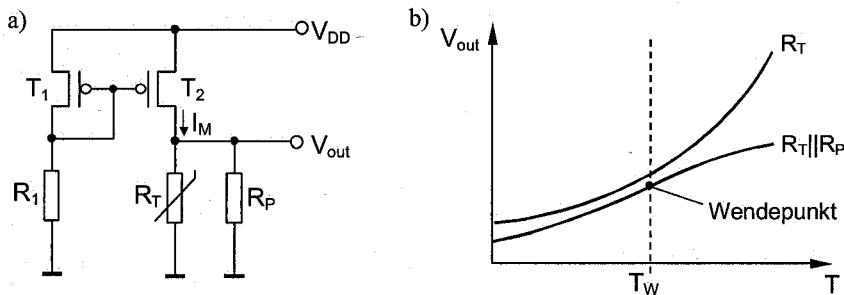


Abb. 6.6. a) Messschaltung zur Ansteuerung und Linearisierung eines Spreading-Resistance Si-Temperatursensors R_T b) Verlauf der Ausgangsspannung ohne und mit Parallelwiderstand R_P

Der Verlauf der Ausgangsspannung $V_{out}(T)$ der Schaltung mit linearisiertem Si-Temperatursensor ist in Abb. 6.6b dargestellt und berechnet sich wie folgt:

$$V_{out}(T) = I_M \cdot \frac{R_T(T) \cdot R_P}{R_T(T) + R_P} \quad (6.6)$$

Der Parallelwiderstand R_P wird so dimensioniert, dass die resultierende Kennlinie einen Wendepunkt bei $T = T_W$ gemäß Abb. 6.6b besitzt [6.3]. Diese Linearisierung ist für eine Kalibrierung bei analoger Signalverarbeitung von Vorteil, geht aber mit einer betragsmäßigen Verringerung des Gradienten der Kennlinie und damit einem Empfindlichkeitsverlust des Sensors einher.

6.1.3 Thermopiles

Der in Abschn. 1.6.2 beschriebene Thermoeffekt wird zur Temperaturmessung in sog. *Thermopiles* genutzt. Durch Absorption von Wärmestrahlung bzw. Infrarot (IR)-Strahlung wird eine Thermospannung zwischen zwei unterschiedlichen Materialien hervorgerufen und ausgewertet. Dazu wird eine größere Anzahl von Thermoelementen in Serie geschaltet, um eine hohe Empfindlichkeit zu erzeugen. Der schematische Aufbau eines Thermopiles ist in Abb. 6.7 im Querschnitt dargestellt.

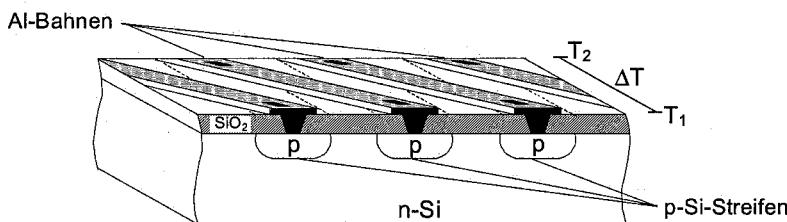


Abb. 6.7. Schematischer Aufbau eines Thermopiles (Serienschaltung von Thermoelementen) mit Temperaturdifferenz ΔT zwischen vorderer und hinterer Kontaktlochreihe

Die einzelnen Thermoelemente bestehen hier jeweils aus einem p-Silizium-Diffusionsstreifen und einer Aluminiumbahn, die über ein Kontaktloch verbunden sind. Die Temperaturdifferenz $\Delta T = T_2 - T_1$ zwischen der vorderen und hinteren Kontaktlochreihe wird dabei ausgewertet. Die Temperaturspannung des Thermopiles setzt sich aus den Thermospannungen ΔV_ϑ der einzelnen Thermoelemente zusammen:

$$\Delta V_\vartheta = \alpha_{AB} \cdot \Delta T \text{ mit } \alpha_{AB} = \alpha_A - \alpha_B, \quad (6.7)$$

wobei α_{AB} den Seebeck-Koeffizienten beschreibt, der sich aus den differentiellen Thermokoeffizienten α_A und α_B der verwendeten Materialien ergibt. Die Empfindlichkeit S eines aus Thermoelementen der Anzahl N bestehenden Thermopiles berechnet sich nach [6.1] bei sinusförmigem Strahlungsfluss mit der Frequenz ω wie folgt:

$$S = N \cdot \frac{\kappa \cdot \alpha_{AB}}{G_T \cdot \sqrt{1 + \omega^2 \cdot \tau_T}}. \quad (6.6)$$

In Gl. (6.6) bezeichnet τ_T die thermische Zeitkonstante, G_T den Temperaturleitwert und κ den Absorptionsgrad, mit dem das Sensorelement den Strahlungsfluss absorbiert. Zur Erreichung einer hohen Empfindlichkeit S sollten die verwendeten Materialien differentielle Thermokoeffizienten α_A und α_B mit möglichst hohem unterschiedlichen Wert aufweisen. Weiterhin führt ein geringer spezifischer Widerstand $\rho(T)$ und eine geringe thermische Leitfähigkeit σ_T des Materials zu einer hohen Empfindlichkeit des Sensorelements.

Durch entsprechende Dotierung des Halbleitermaterials können positive und negative Thermokräfte sowie unterschiedliche Volumenwiderstände realisiert werden [6.1]. Bei höheren Volumenwiderständen nimmt der Seebeck-Koeffizient α_{AB} zu, aber auch das thermische Rauschen, welches dem Nutzsignal überlagert ist.

6.2 Magnetosensoren

Magnetosensoren (auch *Hallgeneratoren* genannt) basieren auf dem in Abschn. 1.6.3 beschriebenen Halleffekt, der den Einfluss eines magnetischen Feldes auf die Ladungsträgerbewegung im Halbleiter ausnutzt. Abhängig von der Stärke und Richtung des magnetischen Feldes wird an einem stromdurchflossenen Halbleiter eine Hallspannung V_H erzeugt. Diese Hallspannung kann messtechnisch ausgewertet und weiterverarbeitet werden. In Abb. 6.8 sind typische Bauformen in Rechteckform und symmetrischer Kreuzform sowie das Symbol des Hallgenerators dargestellt [6.2].

Für die Funktion des Hallgenerators sind im wesentlichen der steuerseitige Innenwiderstand R_{AB} zwischen Kontakt A und B sowie der hallseitige Innenwiderstand R_{CD} zwischen Kontakt C und D von Bedeutung.

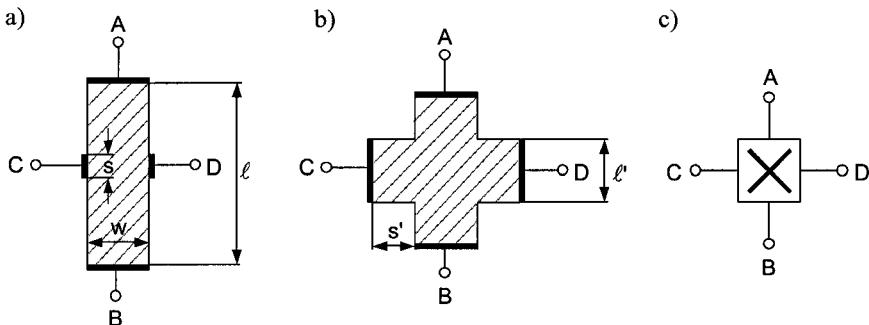


Abb. 6.8. Bauformen eines Hallgenerators a) in Rechteckform b) in symmetrischer Kreuzform und c) Symbol des Hallgenerators

Die Geometrie des rechteckigen Hallgenerators (Abb. 6.8a) wird zumeist in einem Längen-Breiten-Verhältnis von $\ell/w \geq 2$ ausgeführt, sodass der hallseitige Innenwiderstand kleiner als der steuerseitige Innenwiderstand ist ($R_{CD} < R_{AB}$). Die Länge der Hall-Elektrode s ist dabei möglichst klein auszulegen. Der symmetrisch kreuzförmige Hallgenerator (Abb. 6.8b) wird meist mit dem Verhältnis $s'/\ell \geq 0,4$ dimensioniert. Da beim symmetrisch kreuzförmigen Hallgenerator $R_{CD} = R_{AB}$ gilt, ist eine Festlegung von Hall- und Steueranschlüssen nicht notwendig.

Der von der magnetischen Induktion B abhängige hallseitige Innenwiderstand $R_{CD}(B)$ des Hallgenerators dient als charakteristische Größe. Ein typischer Verlauf des normierten Innenwiderstandes R_{CD}/R_0 ist in Abb. 6.9 für verschiedene Halbleitermaterialien dargestellt, wobei $R_0 = R_{CD}(B = 0)$.

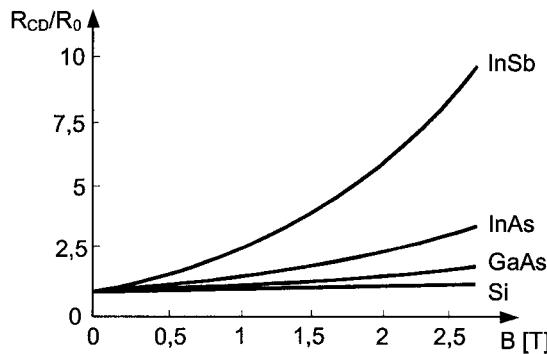


Abb. 6.9. Typische Kennlinien von Hallgeneratoren aus verschiedenen Halbleitermaterialien

Die materialabhängigen Widerstandsverläufe resultieren aus dem unterschiedlich starken Einfluss der Ladungsträgerbeweglichkeiten der verschiedenen Materialien bzw. der Hallbeweglichkeit μ_H gemäß Gl. (1.64), wobei eine möglichst große Hallbeweglichkeit angestrebt wird. Weiterhin ist die Leistungsdichte ρ_p des Hall-

generators (vgl. Abschn. 1.6.3) zu minimieren, um die Eigenerwärmung des Sensors möglichst gering zu halten.

Die Temperaturabhängigkeit des Innenwiderstandes bzw. der Hallspannung des Hallgenerators wird im wesentlichen durch den Bandabstand E_G des verwendeten Halbleitermaterials bestimmt (großer Bandabstand bedeutet einen geringen Temperaturgradienten) und kann durch die Dotierungsdichte in bestimmten Grenzen beeinflusst werden. Der Widerstand der Materialien InSb und InAs hat jeweils einen negativen Temperaturkoeffizienten, da auf Grund der Eigenleitung eine Ladungsträgerzunahme bei Temperaturerhöhung erfolgt. GaAs und Si besitzen hingegen einen positiven Temperaturkoeffizienten, der auf eine Abnahme der Elektronenbeweglichkeit μ_n bei steigender Temperatur zurückzuführen ist. Den geringsten Temperaturkoeffizienten der betrachteten Materialien besitzt GaAs mit $\alpha = 5 \cdot 10^{-4} \text{ K}^{-1}$ [6.2].

Die Ansteuerung von Hallgeneratoren erfolgt zumeist mit einem konstanten Steuerstrom I_M . Dieser kann über eine Konstantstromquelle gemäß Abschn. 3.7.2 oder Abschn. 4.8.4 bereitgestellt werden. Die Ausgangsspannung des Hallgenerators V_H kann mit Hilfe eines Operationsverstärkers OP_1 (vgl. Abschn. 4.8.6) verstärkt werden. Eine entsprechende Ansteuer- und Auswerteschaltung ist in Abb. 6.10 dargestellt. Bei Verwendung von z. B. Silizium als Hallgeneratormaterial kann die Ansteuer- und Auswerteschaltung auf dem gleichen Substrat integriert werden.

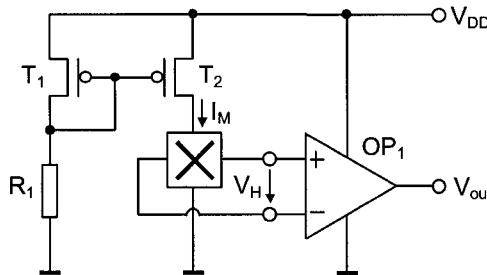


Abb. 6.10. Hallgenerator mit Ansteuer- und Auswerteschaltung

Schaltungstechnisch kann weiterhin eine Kompensation der Temperaturabhängigkeit der Hallspannung V_H vorgenommen werden. Diese Temperaturkompensation kann abhängig vom verwendeten Hallgeneratormaterial durch Bauelemente (z. B. Widerstände) mit komplementären Temperaturkoeffizienten im Steuer- oder Hallkreis erfolgen.

Am realen Hallgenerator tritt auch bei einem magnetischen Feld mit $B = 0$ eine sog. *ohmsche Nullspannung* auf, die durch technologische Toleranzen hervorgerufen wird. Bedingt durch diese geometrischen Abweichungen liegen die Äquipotentiallinien des Hallgenerators bei $B = 0$ nicht exakt senkrecht zu den Hall-elektroden (vgl. Abschn. 1.6.3). Der ohmschen Nullspannung kann durch eine Offsetkompensation am Operationsverstärker OP_1 entgegengewirkt werden. Weiterhin tritt auf Grund der Leitungsführung zu den Hall-Elektroden eine sog. *induk-*

tive Nullspannung auf, die eine nichtlineare magnetische Induktion zur Folge hat. Dadurch wird auch bei einem Steuerstrom von $I_M = 0$ eine Spannung am Hallgenerator erzeugt.

Hallgeneratoren finden im wesentlichen Anwendung als kontaktloser Schalter, bei der berührungslosen Positionserkennung und bei der kontaktlosen Strommessung. Abb. 6.11a zeigt die Anwendung des Hallgenerators als kontaktlosen Schalter. Dazu wird die Ansteuer- und Auswerteschaltung (Abb. 6.10) um einen sog. *Komparator* und eine Ausgangsstufe erweitert.

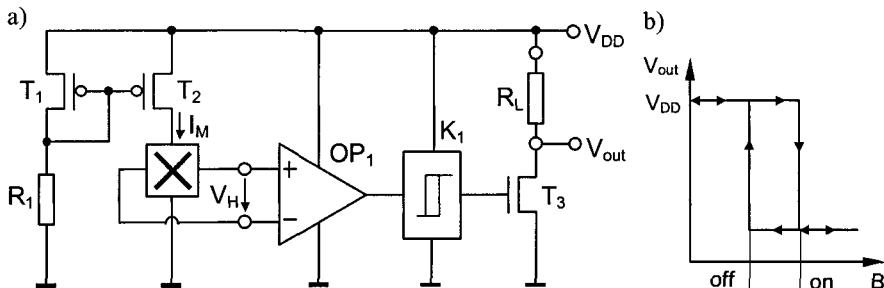


Abb. 6.11. Hallgenerator als kontaktloser Schalter a) Schaltung b) Ausgangskennlinie mit Schalterzuständen

Der Komparator K_1 schaltet ab einem vorgegebenen Schwellwert (z. B. $V_{DD}/2$) seinen Ausgang von V_{SS} auf V_{DD} . Liegt die Ausgangsspannung des Operationsverstärkers OP_1 unterhalb dieses Schwellwertes, so besitzt das Ausgangssignal des Komparators den Wert V_{SS} . Demnach sperrt der MOSFET T_3 und die Ausgangsspannung V_{out} der Schaltung besitzt näherungsweise den Wert V_{DD} (Schalterzustand „off“). Steigt die Ausgangsspannung des Operationsverstärkers bedingt durch die Zunahme der Hallspannung V_H des Hallgenerators soweit an, dass der Schwellwert des Komparators überschritten wird, so wird der MOSFET T_3 leitend und die Ausgangsspannung der Schaltung wird auf V_{SS} gezogen (Schalterzustand „on“). Da der reale Komparator eine Schalthysterese aufweist, ergibt sich die in Abb. 6.11b dargestellte Kennlinie des kontaktlosen Schalters mit den entsprechenden Schalterzuständen.

Bei der berührungslosen Positionserkennung wird die Position des zu messenden Objektes durch einen Permanentmagneten markiert und die Relativbewegung zwischen Hallgenerator und Permanentmagnet ausgewertet. Diese Relativbewegung kann auch durch eine Winkeländerung zwischen Hallgenerator und Magnet hervorgerufen werden, was zur berührungslosen Winkelmessung genutzt werden kann. Eine Drehzahlmessung kann dadurch realisiert werden, dass z. B. in die Magnetflussstrecke zwischen einem Permanentmagneten und einem Weicheisenrotationskörper ein Hallsensor im Abstand d positioniert wird (vgl. Abb. 6.12a). Ist der Abstand d zwischen Hallsensor und Weicheisenkörper konstant und wird die Hallspannung V_H über die in Abb. 6.10 dargestellte Schaltung ausgewertet, so erhält man den in Abb. 6.12b dargestellten beispielhaften Ausgangsspannungsverlauf V_{out} .

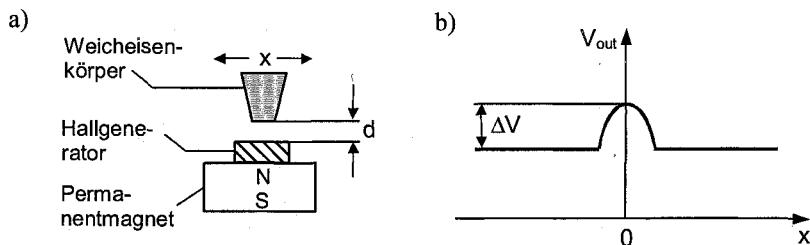


Abb. 6.12. Anwendung eines Hallgenerators zur berührungslosen Positionserkennung
a) Schematischer Aufbau mit Hallgenerator, Permanentmagnet und Weicheisenkörper
b) beispielhafter Verlauf der Ausgangsspannung V_{out} bei Verwendung von Schaltung Abb. 6.10 (Abstand d = konstant)

Bei der kontaktlosen Strommessung wird die durch einen Strom in einer Leitung induzierte magnetische Feldstärke ausgewertet, die mit der zu messenden Stromstärke korreliert. Dazu wird der Hallgenerator im Luftspalt eines Eisenkerns positioniert und mit Hilfe der in Abb. 6.10 dargestellten Schaltung angesteuert und ausgewertet. Die resultierende Ausgangsspannung V_{out} ist proportional zur Stromstärke im Leiter. Durch Verwendung der kontaktlosen Strommessung wird eine Auf trennung der Leitung gegenüber der amperometrischen Strommessung vermieden.

6.3 Optische Sensoren

Optische Sensoren ermöglichen die Umwandlung von Lichtleistung in elektrische Leistung. In Kap. 2 und Kap. 3 wurden bereits die optischen Bauelemente Fotodiode, Solarzelle und Fototransistor behandelt. Ein weiteres optisches Bauelement ist das *CCD-Element* (Charge Coupled Device). Das CCD-Element ermöglicht die Messung der Intensitätsverteilung von Strahlung durch kurzzeitige Ladungsspeicherung mit anschließender Ladungsverschiebung sowie das Erzeugen eines seriellen Ausgangssignals für den Bildinhalt.

Das CCD-Element besteht aus einer Anzahl von MOS-Strukturen (vgl. Abschn. 4.1), die jeweils Ladungen speichern, welche durch optische Generation von zusätzlichen Elektron-Loch-Paaren in der Raumladungszone entstehen. Die Speicherung dieser Ladung erfolgt in der Raumladungszone, die hier als *Potentialmulde* bezeichnet wird. Die Potentialmulde mit der potentialbezogenen Muldentiefe ϕ_s ist entsprechend der optisch generierten Ladung mehr oder weniger stark gefüllt. Das Gate ist als transparentes Fotogate ausgeführt.

Eine MOS-Struktur als grundlegender Bestandteil eines CCD-Elements mit symbolisierter Potentialmulde ist in Abb. 6.13a dargestellt. Dabei entspricht die räumliche Ausdehnung der Potentialmulde der Weite ℓ_D der Raumladungszone (vgl. Abschn. 4.1). In Abb. 6.13b und Abb. 6.13c sind die zugehörigen Bänderdiagramme für den Fall einer leerer und einer gefüllten Potentialmulde mit den entsprechenden charakteristischen Größen veranschaulicht.

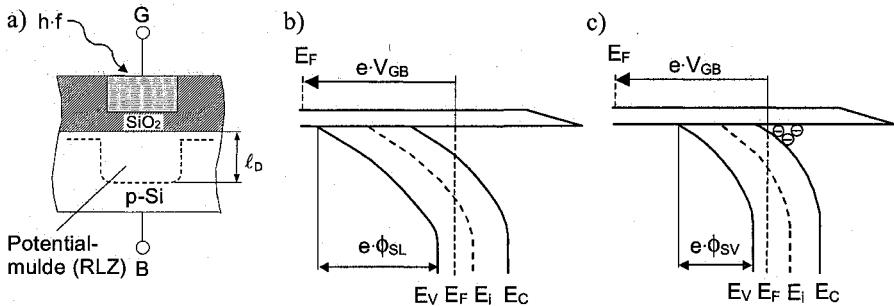


Abb. 6.13. a) MOS-Struktur als Bestandteil eines CCD-Elements mit symbolisierter Potentialmulde und Bänderdiagramme b) bei leerer Potentialmulde und c) bei gefüllter Potentialmulde

Durch einen positiven Spannungssprung am Gate erfolgt eine abrupte Ausdehnung der Raumladungszone der MOS-Struktur und damit eine starke Bandverbiegung (Abb. 6.13b), ohne jedoch eine Inversionsschicht auszubilden (Deep-Depletion). Das dabei entstehende, hohe elektrische Feld bewirkt eine Ladungstrennung der durch optische Generation in der Raumladungszone erzeugten Elektron-Loch-Paare. Die resultierenden freien Elektronen häufen sich an der Grenzfläche zwischen Oxid und Halbleiter an und erzeugen eine negative Grenzflächenladung, wodurch die Bandverbiegung verringert wird (vgl. Abb. 6.13c). Die Größe der Grenzflächenladung stellt den Füllgrad der Potentialmulde dar.

Die Berechnung des Potentialverlaufs für den Füllgrad der Potentialmulde erfolgt gemäß der Spannungsbilanz an der MOS-Struktur (vgl. Abschn. 4.1):

$$V_{GB} - V_{FB} = \phi_S + V_{ox}. \quad (6.7)$$

Es gilt weiterhin die Neutralitätsbedingung:

$$Q_G + Q_D + Q_n = 0, \quad (6.8)$$

wobei Q_G die Gateladung pro Fläche bezeichnet. Q_n ist die Flächenladungsdichte der Elektronen in der Inversionsschicht, die der Speicherladung in der Potentialmulde entspricht. Die Ladungsdichte Q_D der ionisierten Akzeptoren pro Fläche in der Verarmungsschicht (Depletion-Ladung) wird gemäß Gl. (4.16), Abschn. 4.1.5 bestimmt:

$$Q_D = -e \cdot N_A \cdot \ell_D = -\sqrt{2 \cdot \epsilon_0 \cdot \epsilon_{Si} \cdot e \cdot N_A \cdot \phi_S}. \quad (6.9)$$

Analog Abschn. 4.1.4, Gl. (4.7) erhält man für die über dem Oxid abfallende Spannung V_{ox} :

$$V_{ox} = \frac{Q_G}{C_{ox}} = \frac{t_{ox}}{\epsilon_0 \cdot \epsilon_{ox}} \cdot Q_G = \frac{t_{ox}}{\epsilon_0 \cdot \epsilon_{ox}} \cdot (-Q_D - Q_n). \quad (6.10)$$

Zwischen der Gate-Bulk-Spannung V_{GB} und der Muldentiefe ϕ_{SL} einer leeren Potentialmulde ($Q_n = 0$) ergibt sich dann nach Gl. (6.7) unter Einbeziehung von Gl. (6.9) und Gl. (6.10) folgender Zusammenhang:

$$V_{GB} - V_{FB} = \phi_{SL} + \frac{t_{ox}}{\epsilon_0 \cdot \epsilon_{ox}} \cdot \sqrt{2 \cdot \epsilon_0 \cdot \epsilon_{Si} \cdot e \cdot N_A \cdot \phi_{SL}} . \quad (6.11)$$

Für eine gefüllte Potentialmulde mit der Muldentiefe ϕ_{SV} und der in der Mulde durch optische Generation gespeicherten Ladung Q_n folgt für die Gate-Bulk-Spannung V_{GB} :

$$V_{GB} - V_{FB} = \phi_{SV} + \frac{t_{ox}}{\epsilon_0 \cdot \epsilon_{ox}} \cdot \left(\sqrt{2 \cdot \epsilon_0 \cdot \epsilon_{Si} \cdot e \cdot N_A \cdot \phi_{SV}} + |Q_n| \right) . \quad (6.12)$$

Abb. 6.14 veranschaulicht die Beziehung zwischen Oberflächenpotential ϕ_S und Gate-Bulk-Spannung V_{GB} für die Fälle $Q_n = 0$ und $Q_n \neq 0$, wobei die Muldentiefen ϕ_{SL} (leere Mulde) und ϕ_{SV} (gefüllte Mulde) ersichtlich werden.

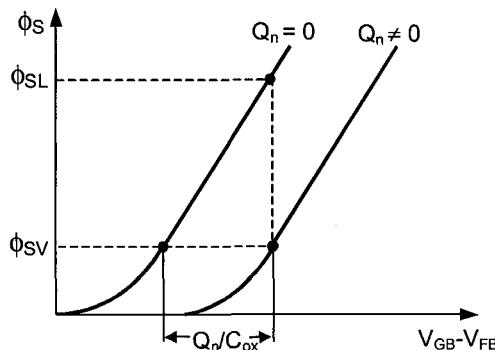


Abb. 6.14. Potentialmuldentiefe der MOS-Struktur in Abhängigkeit von der Gate-Bulk-Spannung bei $Q_n = 0$ und $Q_n \neq 0$

Die maximal in einer Potentialmulde speicherbare Ladung $Q_{n\max}$ kann näherungsweise über das Produkt aus Gate-Bulk-Spannung V_{GB} und Oxidkapazität C_{ox} berechnet werden [6.7]:

$$Q_{n\max} \approx V_{GB} \cdot C_{ox} = \frac{V_{GB} \cdot \epsilon_0 \cdot \epsilon_{Si}}{t_{ox}} . \quad (6.13)$$

Durch das Verändern der Gate-Bulk-Spannung V_{GB} von jeweils benachbarten MOS-Strukturen können Minoritätsträgerladungen von einer MOS-Struktur zur anderen transportiert werden. Die Verkettung mehrerer MOS-Strukturen zum CCD-Element ist in Abb. 6.15a dargestellt. Um einen möglichst störungsfreien Ladungstransport zu ermöglichen, werden die Elektroden beispielsweise mit einem Dreiphasentakt (Abb. 6.15c) angesteuert.

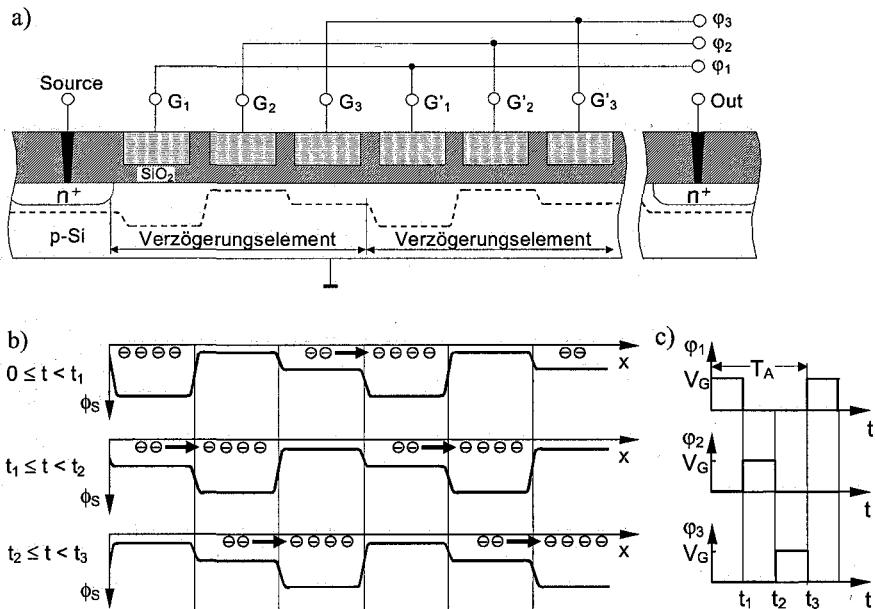


Abb. 6.15. a) CCD-Element (Querschnitt) mit symbolisierten Potentialmulden b) Potentialverläufe und Ladungstransport in drei Phasen c) Dreiphasen-Taktschema

Der Ladungstransport des CCD-Elements erfolgt im wesentlichen durch Diffusion aufgrund des vorherrschenden Konzentrationsgefälles. Es sind aber auch Driftmechanismen und Streufelder am Ladungstransport beteiligt. Die zu transportierende Ladung wird über die Gates des CCD-Elements gesteuert und nach dem sog. *Schieberegisterprinzip* sequentiell am Ausgang *Out* ausgelesen. Jeweils drei MOS-Strukturen bilden in dem in Abb. 6.15 dargestellten Fall ein Verzögerungselement. Durch geeignete Taktspannungen an den Gates der MOS-Strukturen (vgl. Abb. 6.15c) und durch die Wahl der Struktur- und Prozessparameter muss sichergestellt werden, dass sich jeweils unter benachbarten MOS-Strukturen unterschiedlich tiefe Potentialmulden ausbilden, damit die Ladung von Stufe zu Stufe transportiert werden kann.

Während der ersten Phase ($0 \leq t < t_1$) ruht die Ladung Q_n unter den mit der Takteleitung ϕ_1 auf V_G geschalteten MOS-Struktur-Grenzschichten ($\phi_2 = \phi_3 = 0$). Während der zweiten Phase ($t_1 \leq t < t_2$) liegt die Takteleitung ϕ_2 auf V_G , sodass sich unter den Grenzflächen der mit ϕ_2 getakteten MOS-Strukturen ebenfalls Potentialmulden ausbilden, welche die Ladung Q_n von den mit ϕ_1 getakteten MOS-Strukturen übernehmen. In der dritten Phase ($t_2 \leq t < t_3$) besitzt die Takteleitung ϕ_3 das Potential V_G ($\phi_1 = \phi_2 = 0$) und die mit ϕ_3 getakteten MOS-Strukturen bilden Potentialmulden aus. An diese Potentialmulden werden jeweils die gespeicherten Ladungen der mit ϕ_2 getakteten MOS-Strukturen abgegeben. Damit ist ein Zyklus in der Taktzeit T_A abgeschlossen und die Ladung Q_n befindet sich unter der mit ϕ_1 getakteten MOS-Struktur des nächsten Verzögerungselements. Die Ladungsver-

schiebung wird fortgesetzt, bis das Signal bzw. die Ladung am Ausgang des CCD-Elements anliegt. Bei geeigneter Taktansteuerung können die Ladungspakete wahlweise in beiden Richtungen transportiert werden.

Bei jedem Ladungstransfer zwischen zwei Mulden kommt es zu einem Ladungsverlust. Je mehr Zeit für den Übertragungsvorgang zur Verfügung steht, um so mehr Ladung kann übertragen werden. Die in einem CCD-Element wirkenden parasitären Einflüsse sind der *Dunkelstrom*, das Rauschen und externe Störeinflüsse. Der Dunkelstrom ist ein Sperrstrom, der ohne Lichteinwirkung durch thermische Generation und Diffusion in die Potentialmulde gelangt. Die Zwischenspeicherung der Ladung in einer Potentialmulde kann solange erfolgen, bis die Eigenerzeugung von Elektronen durch thermische Generation in der Mulde merklich wird. Diese durch Generation begrenzte maximale Speicherzeit T_G des Einzelelements berechnet sich über die effektive Ladungsträgerlebensdauer τ_S in der Mulde wie folgt:

$$T_G = \frac{Q_G \cdot \tau_S}{e \cdot n_i \cdot \ell_D}. \quad (6.14)$$

Durch den Dunkelstrom wird somit die maximale Gesamtspeicherzeit eines CCD-Elements begrenzt. Bei einem CCD-Element mit N Verzögerungselementen muss die Taktfrequenz entsprechend größer sein als das Verhältnis aus der Anzahl der Verzögerungselemente und der maximalen Speicherzeit T_G :

$$f_A > \frac{N}{T_G}. \quad (6.15)$$

Die Ursachen für Rausch- und Störsignaleinflüsse in CCD-Elementen sind das Rauschen der Eingangsstufe, Transferrauschen beim Ladungstransport, Rauschen der Ausgangsstufe und Störungen der Taktspannungen. Störungen auf den Taktleitungen überlagern sich direkt mit dem Oberflächenpotential ϕ_S und verfälschen damit die Signalinformation. Zusätzlich tritt ein sog. *Pixel-Rauschen* auf, das durch technologische Abweichungen zwischen den einzelnen Gates hervorgerufen wird. Diese deterministischen Abweichungen können durch eine Fehlerkorrektur kompensiert werden.

Neben der dargestellten Form des CCD-Elements mit Oberflächenkanal existiert eine Realisierungsform mit vergrabenem Kanal (BCCD), die einem Berühren der gespeicherten Ladung mit der Oxidgrenzfläche und damit einer Ladungsverfälschung durch Wirken von Grenzflächenladungen entgegenwirkt.

Zur Auswertung des Ausgangssignals muss eine sehr empfindliche Ladungsbewertung erfolgen. Demnach bietet sich die Integration der Auswerteschaltung mit den CCD-Elementen auf dem gleichen Chip an, um eine möglichst geringe Eingangskapazität der Auswerteschaltung zu realisieren, sowie Rauscheinflüsse zwischen Ausgangselektrode des CCD-Elements und Eingang der Auswerteschaltung zu minimieren.

Das Hauptanwendungsgebiet der CCD-Technik liegt im Bildwandlerbereich. Die CCD-Technik ermöglicht die Umwandlung eines Bildes in ein elektrisches Ladungsbild sowie das serielle Auslesen des Bildinhaltes ohne großen techni-

schen Aufwand. Dazu werden die CCD-Elemente in zweidimensionalen Arrays angeordnet. Das gespeicherte Bild kann dann über einen Videoausgang zeilenweise ausgelesen werden.

6.4 Chemosensoren

Die hier beschriebenen Chemosensoren ermöglichen die Bestimmung der Konzentration X_G eines Stoffes durch Wandlung der Messgröße in ein elektrisches Signal. Dabei können verschiedene chemisch-physikalische Effekte genutzt werden. Im Folgenden werden nur die auf Halbleitern basierenden Verfahren beschrieben. Die Änderung der Leitfähigkeit spezieller Halbleiter infolge der Adsorption oxidierender oder reduzierender Gase wird in sog. *Metalloxid-Gassensoren* angewendet. Des weiteren kann die Beeinflussung des Feldeffekts (vgl. Kap. 4) durch die gassensitiven Eigenschaften bestimmter Halbleiter in chemosensitiven Feldeffekttransistoren (*CHEMFETs*) genutzt werden.

6.4.1 Metalloxid-Gassensoren

Das Wirkprinzip von Metalloxid-Gassensoren basiert auf der Leitfähigkeitsänderung eines Metalloxid-Halbleiters unter Einfluss von reduzierenden oder oxidierenden Gasen. Da die Temperatur eine wichtige Einflussgröße auf das gassensitive Verhalten des Metalloxidhalbleiters darstellt, ist der Aufbau von Metalloxid-Gassensoren zumeist durch eine Sensorwirkschicht mit zusätzlichem Heizelement (Heizwendel oder Heizschicht) gekennzeichnet. Die klassische Herstellung eines Metalloxid-Gassensors erfolgt in Sintertechnik (Pillen- oder Röhrchenform) oder Dickschichttechnik. Modernere Verfahren verwenden ein mikrostrukturiertes Silizium-Substrat, auf das ein Metalloxid z. B. durch Sputtern aufgebracht wird. Auf Grund der geringeren zu temperierenden Fläche und des geringeren Abstandes zwischen Heizschicht und Wirkchichtoberfläche ist die notwendige Heizleistung von mikrostrukturierten Sensoren geringer als bei klassischen Sensorprinzipien. Der Aufbau eines Metalloxid-Gassensors in mikrostrukturierter Siliziumtechnologie ist in Abb. 6.16 dargestellt.

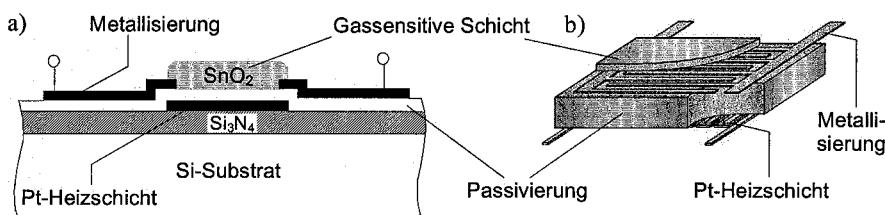


Abb. 6.16. Schematischer Aufbau eines mikrostrukturierten Metalloxid-Gassensors a) im Querschnitt b) Metallisierung als Interdigitalstruktur und Heizschicht als Platin-Mäanderstruktur

Als gassensitive Materialien finden vor allem die Metalloxide SnO_2 , ZnO , WO_3 und Fe_2O_3 Anwendung. Die gassensitive Schicht wird durch eine Metallisierung, die beispielsweise als Interdigitalstruktur ausgelegt ist, kontaktiert (vgl. Abb. 6.16b). Die z. B. als Platin-Mäanderstruktur ausgeprägte Heizschicht wird durch eine Passivierung von der gassensitiven Schicht elektrisch isoliert.

Reduzierende Gase (z. B. CO , H_2 , H_2S , C_xH_y) bewirken eine Leitfähigkeitserhöhung; oxidierbare Gase (z. B. O_2 , NO_2) erzeugen eine Verringerung der Leitfähigkeit des Metalloxid-Halbleiters. An der Halbleiteroberfläche (hier z. B. Zinn-dioxid) wird Sauerstoff aus der Umgebungsluft adsorbiert. Die Adsorptionsplätze an der Halbleiteroberfläche liegen energetisch unterhalb des Ferminiveaus des n-Typ-Halbleiters, wodurch Elektronen des Halbleiters vom Sauerstoff aufgenommen werden (negativ geladener Sauerstoff entsteht). Durch die Abgabe von Elektronen vom Halbleiter entsteht eine Verarmung an Majoritätsträgern an der Halbleiteroberfläche (Depletion), was durch eine Bandverbiegung sichtbar wird (vgl. Abb. 6.17b). Durch die Majoritätsträgerverarmung verringert sich die Leitfähigkeit des Metalloxid-Halbleiters [6.9].

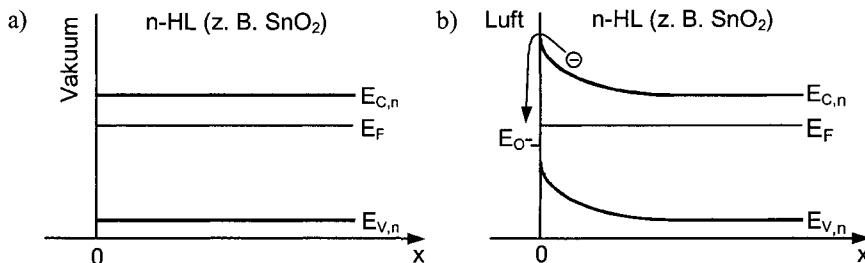


Abb. 6.17. Bänderschemata für a) Übergang zwischen Vakuum und Metalloxid-Halbleiter und b) Übergang zwischen Luft und Metalloxid-Halbleiter

Führt man der Halbleiteroberfläche nun Moleküle eines reduzierenden Gases (z. B. CO) zu, so reagieren diese mit dem adsorbierten, negativ geladenen Sauerstoff zu ebenfalls negativ geladenem CO_2 . Die dadurch auftretenden Mechanismen sind im Bänderschema in Abb. 6.18 veranschaulicht.

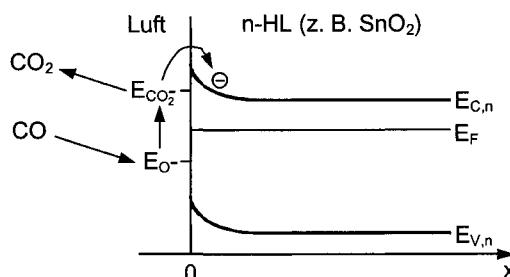


Abb. 6.18. Bänderschema für den Übergang zwischen reduzierendem Gas und Metalloxid-Halbleiter am Beispiel CO-Adsorption / CO_2 -Desorption

Das negativ geladene CO_2 nimmt eine energetisch höhere Lage als das Fermi-niveau ein, wodurch Elektronen an den Metalloxid-Halbleiter abgegeben werden. Daraus resultiert eine Leitfähigkeitserhöhung. Das nun ungeladene CO_2 desorbiert von der Halbleiteroberfläche und der Adsorptionsplatz kann durch neuen Sauerstoff wieder besetzt werden. Die für die Leitfähigkeitsänderung zur Verfügung stehende Zeit wird demnach durch die Desorption von CO_2 bis zur Neubesetzung des Adsorptionsplatzes durch Sauerstoff bestimmt [6.9].

Die Stärke der Leitfähigkeitsänderung $\Delta\sigma_s$ stellt das Maß für die Konzentrationsänderung ΔX_G des adsorbierten Gases dar und wird über den Sensorwiderstand R_s des Metalloxid-Gassensors ausgewertet. Da auf Grund der Herstellungs-technologie eine Reproduzierbarkeit der Sensoreigenschaften nur bedingt gegeben ist, wird mit Metalloxid-Gassensoren in erster Linie die relative Änderung des konzentrationsabhängigen Sensorwiderstandes $\Delta R_s/R_0$ ausgewertet. R_0 gibt dabei den Sensorwiderstand bei $X_G = 0$ an. Ein typischer Verlauf des normierten Sensorwiderstandes R_s/R_0 ist in Abb. 6.19a beispielhaft für die Gase CO, CH_4 und C_4H_{10} dargestellt.

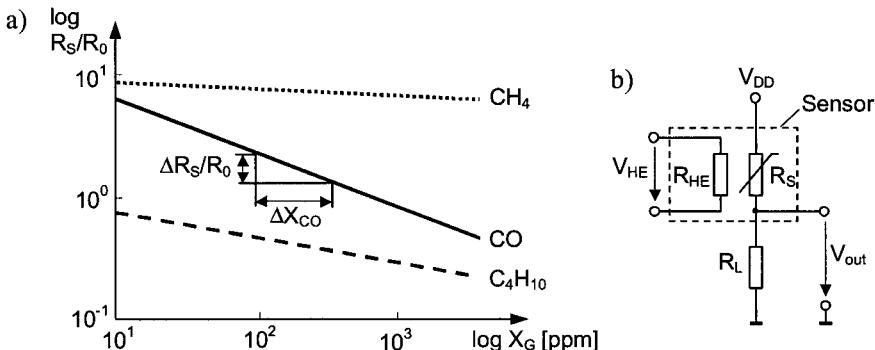


Abb. 6.19. a) Typische Widerstands-Konzentrations-Kennlinie eines Metalloxid-Gassensors für verschiedene Gase b) einfache Ansteuer- und Auswertschaltung eines Metalloxid-Gassensors

Der Gradient des normierten Sensorwiderstandes über der Konzentration ist in Abb. 6.19a beispielhaft für CO hervorgehoben und charakterisiert die Empfindlichkeit des Sensors auf die entsprechende Gaskonzentration (hier X_{CO}). Die Auswertung der Änderung des Sensorwiderstandes R_s kann über die in Abb. 6.19b dargestellte Schaltung in eine äquivalente Spannung V_{out} umgewandelt werden. Dabei wird der Gassensor mit einer konstanten Betriebsspannung V_{DD} versorgt. Der Heizwiderstand R_{HE} des Gassensors wird mit einer Heizspannung V_{HE} betrieben. In Reihe zum Sensorwiderstand $R_s(X_G)$ befindet sich ein Lastwiderstand R_L , über den die folgende Ausgangsspannung abfällt:

$$V_{\text{out}} = \frac{V_{\text{DD}}}{R_s(X_G)/R_L + 1} \quad (6.16)$$

Die Betriebstemperaturen von Metalloxid-Gassensoren, deren Wirkungsweise auf den geschilderten Oberflächeneffekten basiert, liegen typischerweise im Bereich von 200°C bis 600°C. Bei höheren Temperaturen dominieren Volumeneffekte im Metalloxidhalbleiter, die durch Diffusion von Sauerstoffatomen bzw. Sauerstoffionen gekennzeichnet sind. Dieser Diffusionsmechanismus beruht auf der Wechselwirkung zwischen der Sauerstoffkonzentration an der Sensoroberfläche und der Anzahl von Sauerstofffehlstellen im Halbleiter. Die resultierende Leitfähigkeitsänderung stellt dann ein Maß für die effektive Dotierungsänderung im Halbleiter dar.

Als Eigenschaften von Metalloxid-Gassensoren werden zumeist eine hohe Selektivität auf das zu detektierende Gas, eine geringe oder wohldefinierte Querempfindlichkeit auf andere Gase und eine hohe Empfindlichkeit bzw. geringe Nachweisgrenze angestrebt. Des Weiteren soll eine konstante Empfindlichkeit bei hoher Lebensdauer ohne Auftreten von Kontaminationseffekten sowie eine möglichst geringe Feuchte- und Temperaturabhängigkeit des Sensorsignals erreicht werden.

Metalloxid-Gassensoren können für bestimmte Gase (z. B. NO₂) Konzentrationen bis in den unteren ppb-Bereich detektiert werden. Die Temperaturabhängigkeit des Sensorsignals im Nutztemperaturbereich ist auf Grund der relativ hohen Betriebstemperatur nur gering. Die gassensitiven Eigenschaften von Metalloxid-Gassensoren lassen sich im wesentlichen durch die Wahl des Metalloxids, die Dotierungsdichte des Materials, die Oberflächenbeschaffenheit bzw. das Herstellverfahren sowie durch die Betriebstemperatur beeinflussen. Weiterhin ist es möglich, durch Einsatz von Katalysatoren oder Verringerung der Schichtdicke (Dünnschichttechnik) die Empfindlichkeit des Metalloxid-Gassensors zu erhöhen.

Metalloxid-Gassensoren finden z. B. im Bereich der Luftgüte- und Prozessüberwachung Anwendung.

6.4.2 Chemosensitiver Feldeffekttransistor

Die gassensitiven Eigenschaften spezieller Halbleiter können auch zur Beeinflussung des in Kap. 4 beschriebenen Feldeffekts genutzt werden. Diese Bauelemente werden chemosensitive Feldeffekttransistoren (CHEMFETs) genannt. Man unterscheidet hierbei Feldeffekttransistoren zur Detektion von Gasmolekülen (GASFETs) und zur Detektion von Ionen (ISFETs).

Die Funktionsweise des CHEMFETs basiert auf den Eigenschaften der MOS-Struktur (vgl. Abschn. 4.1), deren Gate-Metallisierung jedoch durch gassensitives Material wie Palladium ersetzt wird (vgl. Abb. 6.20).

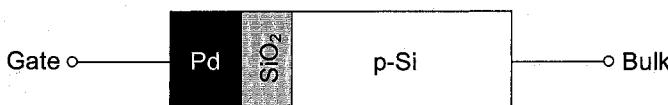


Abb. 6.20. Schematischer Aufbau einer MOS-Struktur bei Verwendung von Palladium als gassensitives Gatematerial

Die Pd-Schicht lässt z. B. Wasserstoff relativ ungehindert zur Grenzschicht Pd/SiO_2 hindurchtreten. An dieser Grenzschicht bilden sich dann Wasserstoffdipole, welche die Austrittsarbeit des Gatematerials beeinflussen. Dadurch verändert sich die Flachbandspannung V_{FB} und damit die Kapazitäts-Spannungs-Charakteristik der MOS-Struktur (vgl. Abschn. 4.1.6), wobei die auftretende Spannungsänderung ΔV_G als Maß für die Gaskonzentration X_G dient. Der schematische Kapazitäts-Spannungs-Verlauf und die zugehörige konzentrationsabhängige Spannungsänderung für das Beispiel Wasserstoff ist in Abb. 6.21 dargestellt.

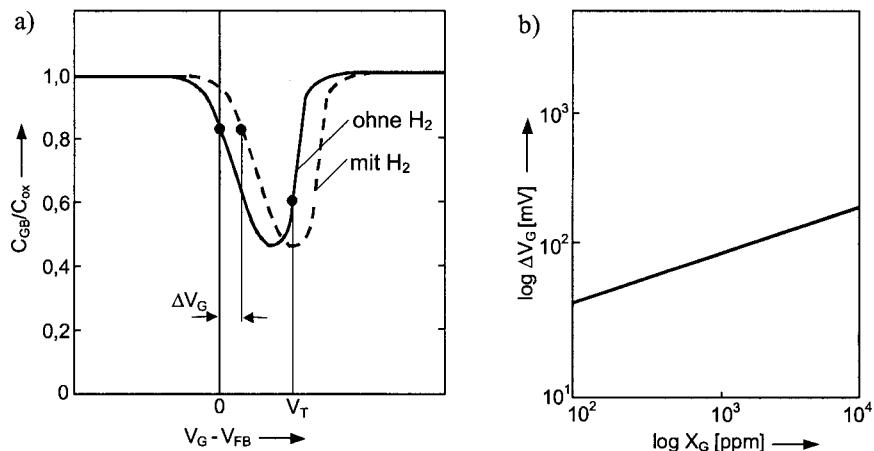


Abb. 6.21. a) Kapazitäts-Spannungs-Charakteristik einer gassensitiven MOS-Struktur bei Beaufschlagung mit/ohne H_2 für Niederfrequenzkleinsignalverhalten b) Spannungs-Konzentrations-Kennlinie für H_2

Bei Erweiterung der MOS-Struktur (Abb. 6.20) zu einem Feldeffekttransistor erhält man den gassensitiven Feldeffekttransistor (GASFET). Hier wird die Änderung der Einsatzspannung V_T um ΔV_G ausgewertet (vgl. Abb. 6.21a). Dabei stellt der Drainstrom I_D das Maß für die Konzentration X_G dar. Der Querschnitt eines GASFETs ist in Abb. 6.22 dargestellt.

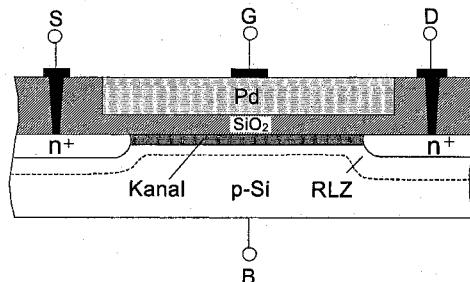


Abb. 6.22. Prinzipieller Aufbau eines gassensitiven Feldeffekttransistors (GASFET) mit Palladium-Gate im Querschnitt

Beim ionensensitiven Feldeffekttransistor wird die Gatemetallisierung durch eine als Gateelektrode wirkende Flüssigkeit ersetzt, in der die Ionenkonzentration bestimmt wird. Das Potential der Flüssigkeit wird dabei über eine Referenzelektrode festgelegt. Der schematische Aufbau eines ISFETs ist in Abb. 6.23 im Querschnitt dargestellt.

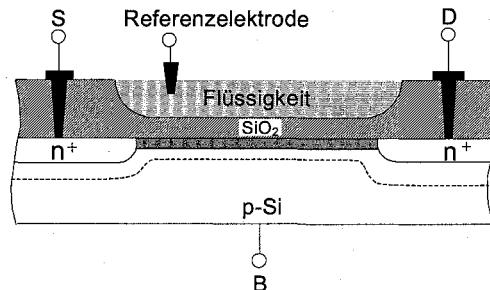


Abb. 6.23. Schematischer Aufbau eines ionensensitiven Feldeffekttransistors (ISFET) im Querschnitt

Abhängig von der Ionenkonzentration in der Flüssigkeit ändert sich die Einsatzspannung V_T und damit der Drainstrom I_D des ISFETs. Der ISFET kann zur Bestimmung des pH-Wertes von Flüssigkeiten verwendet werden. Ein typischer Verlauf der äquivalenten Gatespannungsänderung in Abhängigkeit vom pH-Wert einer Flüssigkeit ist in Abb. 6.24 dargestellt [6.5].

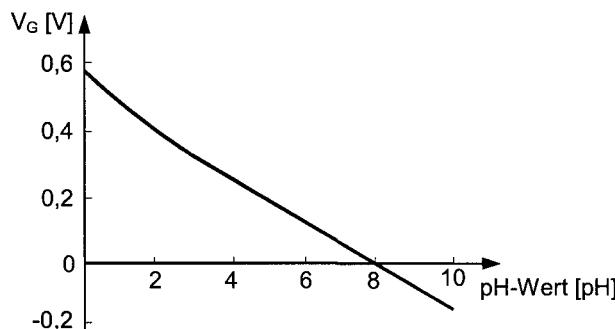


Abb. 6.24. Typischer Verlauf der äquivalenten Gatespannungsänderung eines ISFETs in Abhängigkeit vom pH-Wert einer Flüssigkeit

6.5 Sensoren für mechanische Größen

In diesem Abschnitt werden die wesentlichen Prinzipien zur Messung von mechanischen Größen wie Dehnung, Weg, Torsion, Kraft, Beschleunigung und Druck beschrieben. Ein wesentlicher Halbleitereffekt, der zur Bestimmung dieser Größen

genutzt werden kann, ist der sog. *piezoresistive Effekt*. Unter Einfluss mechanischer Spannungen erfolgt eine Atomverschiebung im Halbleiterkristall. Daraus resultiert eine Änderung der Gitterkonstanten sowie eine Strukturänderung des Leitungs- und Valenzbandes, was in einer Änderung des elektrischen Widerstands des Halbleiters sichtbar wird. Neben dem piezoresistiven Effekt kann auch der kapazitive Effekt zur Messung mechanischer Größen genutzt werden, der die Umwandlung einer mechanischen Größe in eine äquivalente Kapazitätsänderung beschreibt. Im Folgenden werden Sensoren, die auf diesen Effekten basieren, näher betrachtet.

6.5.1 Piezoresistive Kraft- und Drucksensoren

6.5.1.1 Halbleiter-Dehnungsmessstreifen

Der piezoresistive Effekt wird in sog. *Dehnungsmessstreifen* (DMS) zur Erfassung mechanischer Größen wie Dehnung, Weg, Torsion und Kraft genutzt. Der schematische Aufbau eines Halbleiter-Dehnungsmessstreifens ist in Abb. 6.25 dargestellt. Dabei ist der DMS eine dünne Silizium-Schicht (z. B. Poly-Silizium), die auf einem Verformungskörper (hier p-Silizium-Substrat) elektrisch isoliert aufgebracht ist. Die verformungsabhängige Widerstandsänderung des DMS wird über zwei Messelektroden erfasst.

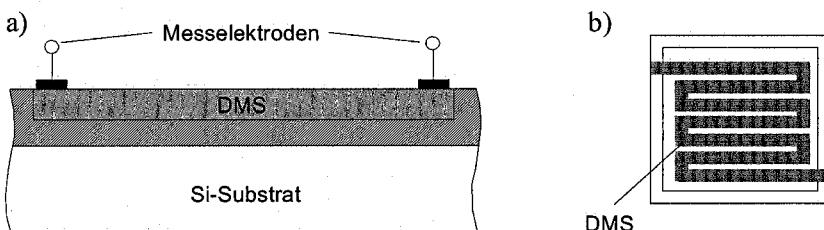


Abb. 6.25. Schematischer Aufbau eines Halbleiter-Dehnungsmessstreifens a) im Querschnitt b) Draufsicht

Der mechanische Spannungszustand wird über die Oberfläche des Verformungskörpers auf den DMS übertragen. Beim longitudinalen piezoresistiven Effekt erhält man für die relative Widerstandsänderung des DMS folgenden Zusammenhang:

$$\frac{\Delta R}{R} = k \cdot \frac{\Delta \ell}{\ell} . \quad (6.17)$$

Dabei gibt $\Delta \ell / \ell$ die relative Längenänderung bei Dehnung bzw. Stauchung des DMS an. Die Konstante k bezeichnet den sog. *k-Faktor*, der die Empfindlichkeit des DMS charakterisiert. Der *k-Faktor* ist kristallrichtungs- und dotierungsabhängig (p-Typ-Halbleiter: $k > 0$, n-Typ-Halbleiter: $k < 0$). Die relative Widerstands-

änderung in Abhängigkeit von der relativen Längenänderung für n- und p-Silizium ist in Abb. 6.26 schematisch dargestellt.

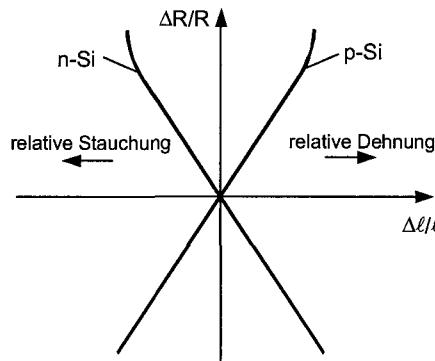


Abb. 6.26. Kennlinie eines Silizium-Dehnungsmessstreifen (relative Widerstandsänderung in Abhängigkeit von der relativen Längenänderung) für n- und p-Silizium

Die Auswertung der Widerstandsänderung des DMS erfolgt zumeist mit Hilfe einer Wheatstone-Brückenschaltung gemäß Abb. 6.27, in der beispielsweise vier Dehnungsmessstreifen verschaltet sind.

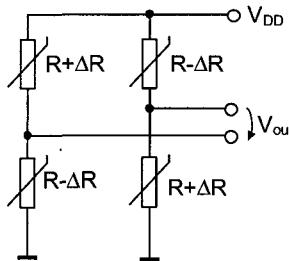


Abb. 6.27. Wheatstone-Brückenschaltung von Halbleiter-Dehnungsmessstreifen

Das Verhältnis der Ausgangsspannung V_{out} zur konstant gehaltenen Betriebsspannung V_{DD} ergibt die relative Widerstandänderung $\Delta R/R$ des piezoresistiven Sensors:

$$\frac{V_{out}}{V_{DD}} = \frac{R + \Delta R}{2R} - \frac{R - \Delta R}{2R} = \frac{\Delta R}{R}. \quad (6.18)$$

Durch die Verwendung einer Brückenschaltung wird die Empfindlichkeit des piezoresistiven Sensors erhöht, da sich die Widerstandänderungen der einzelnen Dehnungsmessstreifen addieren. Gleichsinnige Effekte, wie die Drift des Widerstandsabsolutwertes und der Temperaturkoeffizient des Widerstandes der DMS, werden hingegen kompensiert.

6.5.1.2 Piezoresistiver Drucksensor

Die in Abschn. 6.5.1.1 dargestellten piezoresistiven Effekte an Dehnungsmessstreifen können ebenfalls zur Druckmessung genutzt werden. Bei druckabhängiger Durchbiegung einer Silizium-Membran bilden sich Zonen aus, die gedehnt bzw. gestaucht werden. Hier ordnet man Dehnungsmessstreifen ($DMS_1 - DMS_4$) z. B. durch mikrostrukturierte Herstellung an. Der schematische Aufbau eines Halbleiter-Drucksensors zur Absolutdruckmessung ist in Abb. 6.28a dargestellt. Um eine Relativdruckmessung zur realisieren, ist in der Druckkammer ein Druck $p \neq 0$ zu erzeugen (Abb. 6.28b), wodurch eine zum Differenzdruck äquivalente Widerstandsänderung der DMS detektiert werden kann.

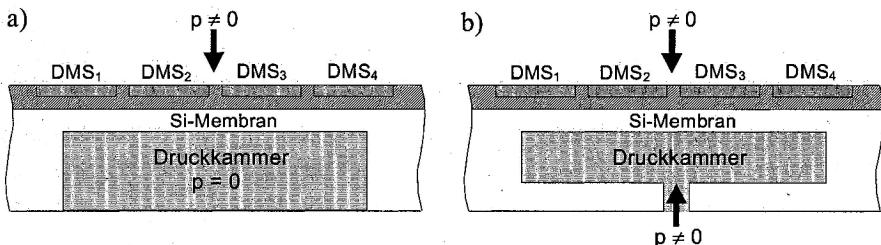


Abb. 6.28. Schematischer Aufbau eines piezoresistiven Drucksensors im Querschnitt a) zur Absolutdruckmessung b) zur Relativdruckmessung

Die Dehnungsmessstreifen werden gemäß Abb. 6.29a in einer Wheatstone-Brückenschaltung angeordnet. Wirkt ein Druck p auf diese Brückenschaltung, so ergibt sich die folgende Ausgangsspannung V_{out} :

$$V_{out} = S \cdot p \cdot V_{DD} \quad \text{mit} \quad S = \frac{\Delta R}{\Delta p \cdot R}, \quad (6.19)$$

wobei S die Empfindlichkeit des Drucksensors charakterisiert. Ein typischer Verlauf der Ausgangsspannung eines piezoresistiven Drucksensors ist in Abb. 6.29b für verschiedene Betriebstemperaturen dargestellt.

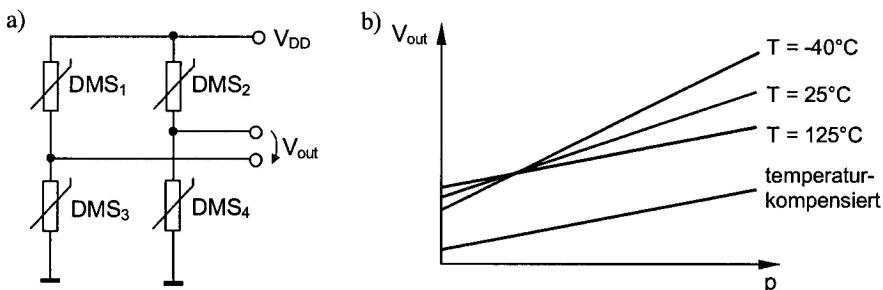


Abb. 6.29. a) Wheatstone-Brückenschaltung eines piezoresistiven Drucksensors b) Ausgangskennlinie des Drucksensors für verschiedene Temperaturen und temperaturkompensiert

Auf Grund der Temperaturabhängigkeit des spezifischen Widerstandes und des k-Faktors von Halbleiter-Dehnungsmessstreifen ist eine Temperaturkompensation der Drucksensoranordnung notwendig. Dies kann z. B. durch Verwendung eines Grenzflächen-Temperatursensors erfolgen (vgl. Abschn. 6.1.1). Eine weitere Möglichkeit besteht darin, die Temperaturabhängigkeit der DMS selbst zur Temperaturkompensation zu nutzen. Dazu wird die Brückenschaltung (Abb. 6.29a) mit einem konstanten Strom anstelle einer konstanten Spannung betrieben. Die Ausgangskennlinie eines temperaturkompensierten Drucksensors ist aus Abb. 6.29b ersichtlich.

Das Hauptanwendungsgebiet von piezoresistiven Drucksensoren ist die Durchfluss- und Mengenmessung von Gasen und Flüssigkeiten sowie die Beschleunigungsmessung in Sicherheitssystemen (Airbag-Anwendung).

6.5.2 Kapazitiver Drucksensor

Neben der piezoresistiven Druckmessung kann der Druck auch mit Hilfe eines kapazitiven Prinzips erfasst werden. Hierzu wird die Abstandsänderung von Kondensatorplatten bei Druckbelastung ausgewertet. Aus dieser Abstandsänderung resultiert eine Kapazitätsänderung, die messtechnisch detektiert werden kann. Der schematische Aufbau eines mikrostrukturierten kapazitiven Drucksensors ist in Abb. 6.30 im Querschnitt dargestellt [6.8].

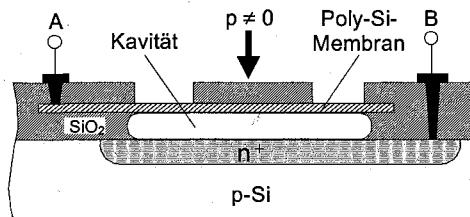


Abb. 6.30. Schematischer Aufbau eines mikrostrukturierten kapazitiven Drucksensors im Querschnitt

Hierbei befindet sich eine druckempfindliche Poly-Si-Membran über einer hermetisch verschlossenen Kavität. Bei Druckbeaufschlagung wird die Membran ausgelenkt, wodurch die zwischen Membran und n^+ -Diffusionsgebiet befindliche Kapazität geändert wird, die an den Elektroden A und B abgegriffen werden kann. Die Empfindlichkeit des kapazitiven Drucksensors wird in erster Linie durch die Dicke und Fläche der Membran bestimmt. Die schaltungstechnische Auswertung des Sensorsignals kann integriert erfolgen. Dazu wird eine Brückenschaltung von vier kapazitiven Drucksensoren genutzt, wobei zwei Sensoren druckempfindlich ausgelegt sind und zwei Sensoren als Referenz dienen. Die Wandlung des analogen Ausgangssignals in ein digitales Signal kann störungssarm beispielsweise durch einen sog. *Sigma-Delta-Wandler* auf dem gleichen Substrat erfolgen.

Kapazitive Drucksensoren finden Anwendung zur Detektion von Druckwellen, insbesondere als schneller Beschleunigungssensor in Airbag-Anwendungen.

7 Integrierte Schaltungen

Integrierte Schaltungen (*ICs*) ermöglichen die Realisierung komplexer elektronischer Schaltungen auf einem Halbleitersubstrat. Sie weisen eine hohe Packungsdichte bei geringer Chipfläche auf und sind für hohe Schaltgeschwindigkeiten geeignet. Die Herstellung von integrierten Schaltungen erfolgt zumeist in Planartechnik, wobei die Einzelprozessschritte der Schichttechnik, Lithographie, Ätztechnik und Dotiertechnik zugeordnet werden (vgl. Abschn. 2.4, 3.4, 4.6). Auf Grund des hohen Herstellungsaufwandes sind integrierte Schaltungen erst bei großen Stückzahlen preisgünstig.

In diesem Kapitel wird insbesondere die Anwendung von Feldeffekttransistoren (vgl. Kap. 4) in sog. *digitalen integrierten CMOS-Schaltungen* beschrieben. CMOS-Schaltungen besitzen eine hohe Störsicherheit, weisen im statischen Betrieb eine geringe Verlustleistung auf und sind hinsichtlich des Schaltungsentwurfs relativ einfach zu realisieren. Ausgehend von CMOS-Grundschatungen in statischer und dynamischer CMOS-Logik, die auch Schaltungselemente aus den in Kap. 4 beschriebenen MOSFET-Grundschatungen enthalten, werden hier die wichtigsten Vertreter von CMOS-Speicherschaltungen und integrierten Logikbausteinen anhand ihres Aufbaus, der Funktionsweise und der praktischen Anwendbarkeit behandelt.

7.1 CMOS-Grundschatungen

In CMOS-Schaltungen werden die komplementären Eigenschaften von n- und p-Kanal-MOSFETs (vgl. Kap. 4) zur Realisierung von nahezu ruhestromfreien Logikschatungen genutzt. Eine bedeutende CMOS-Grundschatzung stellt der CMOS-Inverter dar, der in Abschn. 4.8.2. hinsichtlich statischem Übertragungsverhalten und Schaltverhalten beschrieben wurde. Nachfolgend werden weitere CMOS-Schaltungen der statischen und dynamischen CMOS-Logik betrachtet.

7.1.1 Statische CMOS-Logik

7.1.1.1 Realisierung logischer Verknüpfungen

Logische Verknüpfungen können durch sog. *CMOS-Logikgatter* in CMOS-Technologie realisiert werden. Die prinzipielle Schaltungsstruktur eines statischen CMOS-Logikgatters ist in Abb. 7.1 dargestellt [7.4].

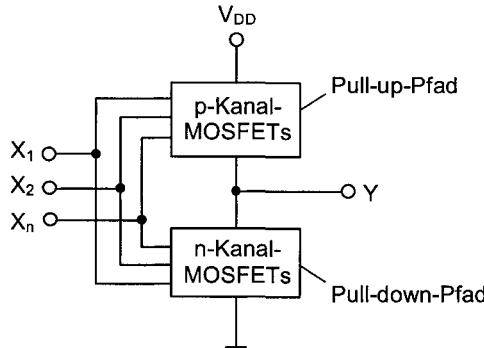


Abb. 7.1. Prinzipielle Schaltungsstruktur eines statischen CMOS-Logikgatters

Die Gatter der statischen CMOS-Logik bestehen prinzipiell aus einem Netzwerk von n-Kanal-MOSFETs im sog. *Pull-down-Pfad* (niederohmige Verbindung zu Masse) und einem Netzwerk von p-Kanal-MOSFETs im *Pull-up-Pfad* (niederohmige Verbindung zur Betriebsspannung V_{DD}). Die n digitalen Eingangsgrößen werden mit $X_1, X_2 \dots X_n$ und die Ausgangsgröße mit Y bezeichnet.

Eine NAND-Verknüpfung bzw. negierte AND-Funktion von digitalen Signalen (Low = 0, High = 1) erfolgt über eine Serienschaltung von n-Kanal-MOSFETs im Pull-down-Pfad und eine Parallelschaltung von p-Kanal-MOSFETs im Pull-up-Pfad. In Abb. 7.2a ist beispielhaft ein *NAND-Gatter* mit zwei Eingängen X_1, X_2 und einem Ausgang Y dargestellt. In Abb. 7.2b ist das zugehörige Schaltsymbol des NAND-Gatters abgebildet.

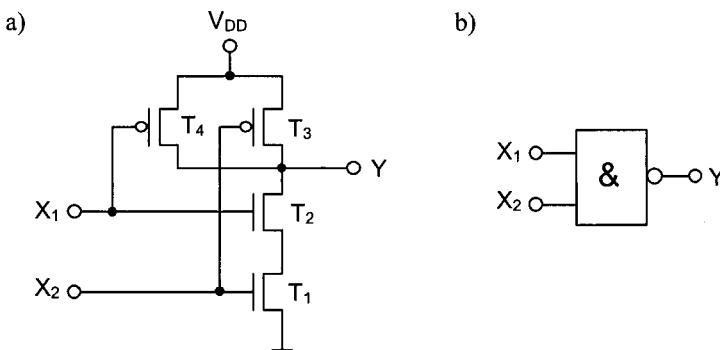


Abb. 7.2. NAND-Gatter in statischer CMOS-Logik a) Schaltungsprinzip mit zwei Eingängen und b) Schaltsymbol

Die logische Funktion von CMOS-Gattern wird in einer sog. *Wahrheitstabelle* abgebildet. In der nachfolgenden Tabelle sind die logische Verknüpfung einer NAND-Funktion und die Schaltzustände (ON = geschlossen, OFF = offen) der n-Kanal-MOSFETs T_1 und T_2 sowie der p-Kanal-MOSFETs T_3 und T_4 gemäß Abb. 7.2a dargestellt.

X_1	X_2	T_1	T_2	T_3	T_4	Y
0	0	OFF	OFF	ON	ON	1
0	1	ON	OFF	OFF	ON	1
1	0	OFF	ON	ON	OFF	1
1	1	ON	ON	OFF	OFF	0

Die Ausgangsgröße Y der NAND-Funktion ergibt sich dann aus den Eingangsgrößen X_1 und X_2 wie folgt:

$$Y = \overline{X_1 \wedge X_2} . \quad (7.1)$$

Eine NOR-Funktion (negierte OR-Funktion) in statischer CMOS-Logik wird über eine Serienschaltung von p-Kanal-MOSFETs im Pull-down-Pfad und eine Parallelschaltung von n-Kanal-MOSFETs im Pull-up-Pfad erreicht. Das Schaltungsprinzip und Schaltsymbol eines NOR-Gatters mit zwei Eingängen ist in Abb. 7.3 ersichtlich.

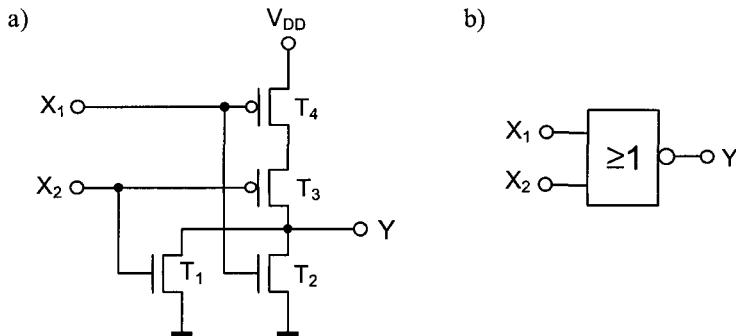


Abb. 7.3. NOR-Gatter in statischer CMOS-Logik a) Schaltungsprinzip mit zwei Eingängen und b) Schaltsymbol

In der nachfolgenden Tabelle sind die logische Verknüpfung und die Schaltzustände der n-Kanal-MOSFETs T_1 und T_2 sowie der p-Kanal-MOSFETs T_3 und T_4 für eine NOR-Verknüpfung gemäß Abb. 7.3a dargestellt.

X_1	X_2	T_1	T_2	T_3	T_4	Y
0	0	OFF	OFF	ON	ON	1
0	1	ON	OFF	OFF	ON	0
1	0	OFF	ON	ON	OFF	0
1	1	ON	ON	OFF	OFF	0

Die Ausgangsgröße Y der NOR-Funktion ergibt sich dann aus den Eingangsgrößen X_1 und X_2 wie folgt:

$$Y = \overline{X_1 \vee X_2} . \quad (7.2)$$

Eine Verknüpfung von mehr als drei Eingangsgrößen in NAND- oder NOR-Gattern bewirkt hohe Schaltzeiten durch die in Serie geschalteten MOSFETs.

Durch eine Zusammenschaltung mehrerer CMOS-Gatter mit verringriger Anzahl von Eingängen in sog. *Komplexgattern* kann einer Schaltzeiterhöhung entgegenwirken. In CMOS-Logikschaltungen können auch Transmissionsgates (vgl. Abschn. 4.8.3) verwendet werden. Als Beispiel eines CMOS-Gatters mit Transmissionsgates ist in Abb. 7.4 ein sog. *Multiplexer* dargestellt, der abhängig vom Zustand des Schaltsignals N das Eingangssignal X_1 oder X_2 auf den Ausgang Y durchschaltet.

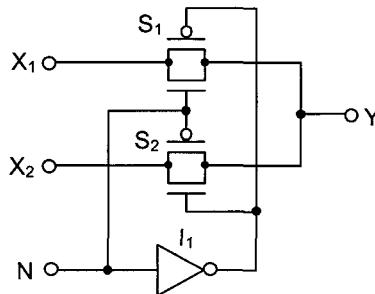


Abb. 7.4. Schaltungsprinzip eines Multiplexers in statischer CMOS-Logik

Als Schalter S_1 und S_2 kommen hier Transmissionsgates zum Einsatz, die über das Schaltsignal N und das durch einen CMOS-Inverter I_1 (vgl. Abschn. 4.8.2) invertierte Schaltsignal gegenphasig angesteuert werden. Die boolesche Ausgangsfunktion Y des Multiplexers kann dann wie folgt beschrieben werden:

$$Y = (X_1 \wedge N) \vee (X_2 \wedge \overline{N}). \quad (7.3)$$

Ein weiteres Beispiel für ein Komplexgatter ist das sog. *XOR-Gatter*, das mit Hilfe von Invertern und eines Transmissionsgates aufgebaut werden kann. Das Schaltungsprinzip und das Schaltsymbol eines XOR-Gatters in statischer CMOS-Logik sind in Abb. 7.5 dargestellt.

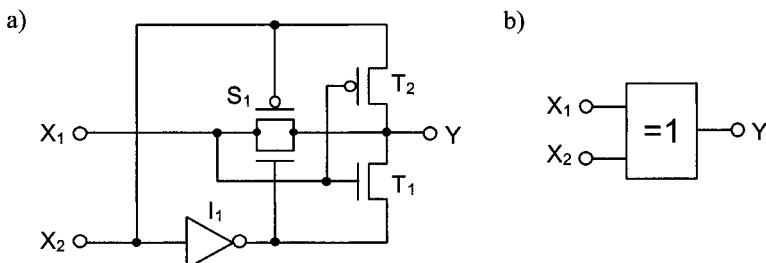


Abb. 7.5. XOR-Gatter in statischer CMOS-Logik a) Schaltungsprinzip und b) Schaltsymbol

Die Verknüpfung der Eingangssignale X_1 und X_2 zur Ausgangsfunktion Y des XOR-Gatters ergibt folgende boolesche Funktion:

$$Y = (\overline{X_1} \wedge X_2) \vee (X_1 \wedge \overline{X_2}). \quad (7.4)$$

Analog zum XOR-Gatter kann man ein *XNOR-Gatter* aufbauen, das sich durch die schaltungstechnische Anordnung des Inverters I_1 vom XOR-Gatter unterscheidet. Das Schaltungsprinzip und das Schaltsymbol eines XNOR-Gatters in statischer CMOS-Logik sind in Abb. 7.6 dargestellt sind.

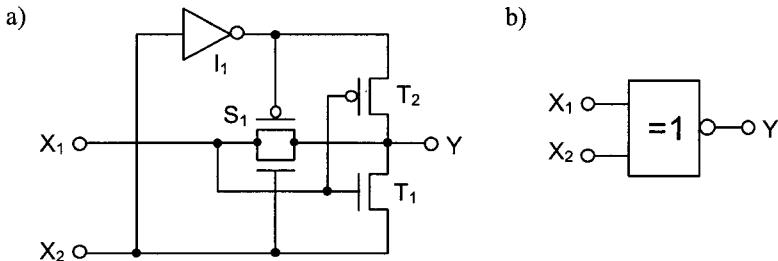


Abb. 7.6. XNOR-Gatter in statischer CMOS-Logik a) Schaltungsprinzip und b) Schaltsymbol

Mit einem XNOR-Gatter gemäß Abb. 7.6 wird folgende logische Ausgangsfunktion realisiert:

$$Y = \overline{(\overline{X_1} \wedge X_2) \vee (X_1 \wedge \overline{X_2})}. \quad (7.5)$$

7.1.1.2 Tristate-Treiber

Befinden sich mehrere Ausgangstreiber (Ansteuerung von Lasten) an einer Busleitung, so darf stets nur ein Ausgangstreiber auf die Leitung wirken; die restlichen Treiber müssen entkoppelt werden. Um diese Entkopplung zu realisieren, ist die Verwendung von Treibern mit einem zusätzlichen, hochohmigen Ausgangszustand Z notwendig. Der sog. *Tristate-Treiber* bietet diese Eigenschaft und besitzt somit die Ausgangszustände Low, High oder Z . Das Schaltungsprinzip eines Tristate-Treibers in statischer CMOS-Logik wird aus Abb. 7.7 ersichtlich.

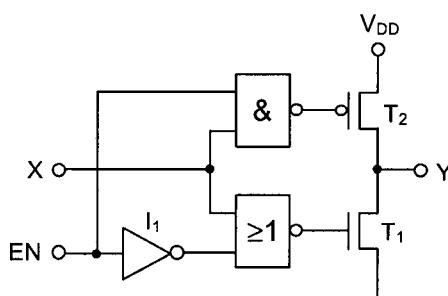


Abb. 7.7. Schaltungsprinzip eines Tristate-Treibers in statischer CMOS-Logik

Besitzt das Auswahlsignal EN des Tristate-Treibers den Zustand Low ($EN = 0$), so sind die MOSFETs T_1 und T_2 nicht durchgeschaltet und der Ausgang Y des Treibers ist hochohmig ($Y = Z$). Besitzt das Auswahlsignal hingegen den Zustand High ($EN = 1$), so verhält sich der Treiber wie ein herkömmlicher Inverter (vgl. Abschn. 4.8.2). Die Wahrheitstabelle des Tristate-Treibers mit den entsprechenden Schaltzuständen der MOSFETs T_1 und T_2 ist nachfolgend abgebildet.

X	EN	T_1	T_2	Y
0	0	OFF	OFF	Z
1	0	OFF	OFF	Z
0	1	ON	OFF	0
1	1	OFF	ON	1

7.1.1.3 Addierer

Ein *Addierer* in statischer CMOS-Logik ermöglicht die binäre Addition von Eingangszuständen. In Abb. 7.8 ist das Schaltungsprinzip eines sog. *1-Bit-Halbaddierers* dargestellt, der aus einem XOR-Gatter (vgl. Abb. 7.5) und einem AND-Gatter besteht. Der Halbaddierer verfügt über die Eingänge A und B , einen Ausgang Y und einen weiteren Ausgang C , der den Übertrag der Additionsfunktion darstellt.

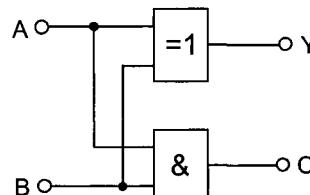


Abb. 7.8. Schaltungsprinzip eines 1-Bit-Halbaddierers in statischer CMOS-Logik

Die Zustände der Ausgänge Y und C eines 1-Bit-Halbaddierers gemäß Abb. 7.8 können in Abhängigkeit von den Eingangsgrößen folgender Wahrheitstabelle entnommen werden.

A	B	Y	C
0	0	0	0
0	1	1	0
1	0	1	0
1	1	0	1

Die logische Verknüpfung der Eingangsgrößen A und B zur Ausgangsgröße Y und dem Übertrag C des Halbaddierers erfolgt somit über die nachfolgende boolesche Funktion:

$$Y = (\overline{A} \wedge B) \vee (A \wedge \overline{B}) \quad \text{und} \quad C = A \wedge B. \quad (7.6)$$

Werden zwei 1-Bit-Halbaddierer schaltungstechnisch verbunden und die jeweiligen Überträge mit einem OR-Gatter verknüpft, so erhält man einen *1-Bit-Volladdierer* mit den Eingängen A_i und B_i und den Ausgängen Y_i und C_i (Übertrag). Die Eingangsgröße C_{i-1} bezeichnet den Übertrag, der beispielsweise von einem vorgeschalteten Addierer geliefert wird. Das Schaltungsprinzip eines 1-Bit-Volladdierers in statischer CMOS-Logik ist in Abb. 7.9 dargestellt.

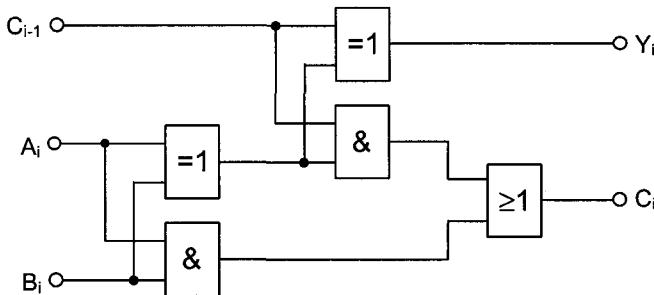


Abb. 7.9. Schaltungsprinzip eines aus zwei Halbaddierern aufgebauten 1-Bit-Volladdierers in statischer CMOS-Logik

Die Wahrheitstabelle eines 1-Bit-Volladdierers, der gemäß Abb. 7.9 aufgebaut ist, ist nachfolgend dargestellt.

A_i	B_i	C_{i-1}	Y_i	C_i
0	0	0	0	0
0	0	1	1	0
0	1	0	1	0
0	1	1	0	1
1	0	0	1	0
1	0	1	0	1
1	1	0	0	1
1	1	1	1	1

An den Ausgängen Y_i und C_i des 1-Bit-Volladdierers treten folgende boolesche Funktionen auf:

$$Y_i = ((\overline{A_i} \wedge B_i) \vee (A_i \wedge \overline{B_i})) \wedge C_{i-1} \vee ((\overline{A_i} \wedge B_i) \vee (A_i \wedge \overline{B_i})) \wedge \overline{C_{i-1}} \quad \text{und} \quad (7.7)$$

$$C_i = (A_i \wedge B_i) \vee (A_i \wedge C_{i-1}) \vee (B_i \wedge C_{i-1}).$$

Mit mehreren 1-Bit-Volladdierern können sog. *Mehrbit-Addierer* realisiert werden, um beispielsweise zwei 4-Bit-Zahlen zu addieren. Der Volladdierer dient auch als Grundzelle für den schaltungstechnischen Aufbau von CMOS-Subtrahierern und -Multiplizierern.

7.1.2 Dynamische CMOS-Logik

Bei der dynamischen CMOS-Logik erfolgt eine Speicherung von Ladungen als Signalgröße in der Gatekapazität von MOSFETs (vgl. Abschn. 4.7.1). Die dynamische CMOS-Logik erfordert im Gegensatz zur statischen CMOS-Logik einen getakteten Schaltungsbetrieb. Der Nachteil besteht in der Störanfälligkeit dieser Schaltungstechnik, die durch Ladungsverluste am Speicherknoten entsteht und z. B. durch Leckströme oder Einkopplungen hervorgerufen wird.

7.1.2.1 Dynamisches Latch

Das dynamische Latch besteht aus einem CMOS-Inverter (vgl. Abschn. 4.8.2) mit einem vorgeschalteten Transmissiongate (vgl. Abschn. 4.8.3). Das Schaltungsprinzip des dynamischen CMOS-Latches ist in Abb. 7.10a dargestellt. Die Ansteuerung des Transmissiongates erfolgt dabei mit einem Zwei-Phasen-Takt mit den zueinander invertierten Signalen φ_1 und φ_2 gemäß des Takschemas in Abb. 7.10b.

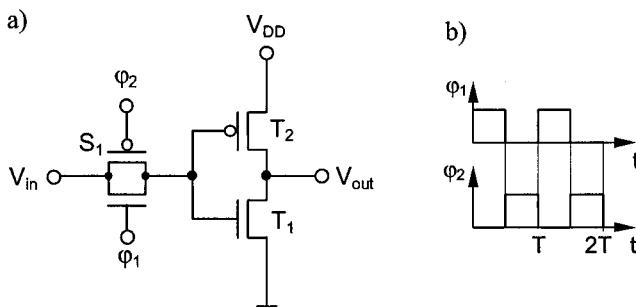


Abb. 7.10. Dynamisches CMOS-Latch a) Schaltungsprinzip und b) Takschema für Zwei-Phasen-Takt

Solange das Transmissiongate S_1 leitend ist ($\varphi_1 = 1$ und $\varphi_2 = 0$), wird der Inverteereingang des Latches mit der Eingangsspannung V_{in} aufgeladen. Wird das Transmissiongate hochohmig ($\varphi_1 = 0$ und $\varphi_2 = 1$), so wird die Information als Ladung in der Eingangskapazität des Inverters gespeichert. Die Taktfrequenz $f = 1/T$, mit der das Transmissiongate S_1 angesteuert wird, muss hoch genug sein, um das Auftreten von Leckströmen und einen damit verbundenen Ladungsverlust zu vermeiden.

7.1.2.2 Dynamisches Schieberegister

Die Kopplung mehrerer gegenphasig getakteter dynamischer CMOS-Latches wird als dynamisches Schieberegister bezeichnet. Das Schieberegister ermöglicht es, ein am Eingang V_{in} angelegtes Signal mit jedem Takt um eine Latch-Stufe weiterzuschieben. Nach Durchlauf durch alle Latch-Stufen liegt das unveränderte Eingangssignal am Ausgang V_{out} zeitlich verzögert an. Das Schaltungsprinzip eines dynamischen Schieberegisters, das beispielhaft aus vier Latch-Stufen aufgebaut

ist, ist in Abb. 7.11 dargestellt. Die Taktansteuerung kann mit Hilfe des in Abb. 7.10b gezeigten Takschemas erfolgen.

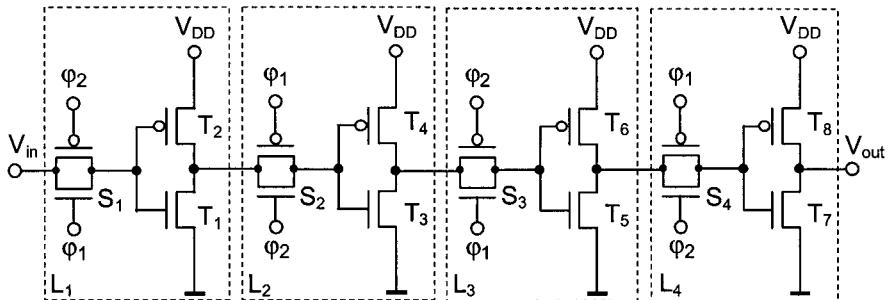


Abb. 7.11. Schaltungsprinzip eines dynamischen Schieberegisters

Besitzen die Taktsignale die Zustände $\varphi_1 = 0$ und $\varphi_2 = 1$, so sind die Transmissiongates S_2 und S_4 der dynamischen Latches L_2 und L_4 niederohmig. Alle weiteren Transmissiongates sind hochohmig. Dabei wird die als Ladung in der jeweiligen Eingangskapazität gespeicherte Information von Latch L_1 auf Latch L_2 bzw. von Latch L_3 auf Latch L_4 übertragen. Nehmen die Taktsignale die Zustände $\varphi_1 = 1$ und $\varphi_2 = 0$ an, so werden die Transmissiongates S_1 und S_3 der Latches L_1 und L_3 hochohmig und die übrigen Transmissiongates sind niederohmig. Dadurch wird die Information von Latch L_2 auf Latch L_3 übertragen.

Das in Abb. 7.11 dargestellte Schaltungsprinzip besitzt den Nachteil, dass es anfällig gegen Störungen (sog. *Hazards*) ist. Diese Störungen treten insbesondere bei Phasenversatz der Takte (sog. *Slew*) oder langsamem Taktflanken auf. Diesem Nachteil kann durch eine Ansteuerung mit einem 4-Phasen-Takt entgegengewirkt werden, sodass die Latches aufeinanderfolgender Stufen nicht gleichzeitig leitend werden können.

7.2 Speicher

Integrierte Speicherbausteine dienen der Speicherung von digitalen Informationen. Hierbei unterscheidet man zunächst die Klasse der nichtflüchtigen und flüchtigen Speicher.

7.2.1 Nichtflüchtige Speicher

Nichtflüchtige Speicher behalten ihre Speicherinformation auch nach Abschalten der Betriebsspannung bei. Sie werden unterschieden in festprogrammierte, anwenderprogrammierbare sowie programmierbare/elektrisch löschen Speicher. Die wesentlichen Vertreter dieser nichtflüchtigen Speicher sind in Abb. 7.12 zusammengefasst und werden nachfolgend näher beschrieben.

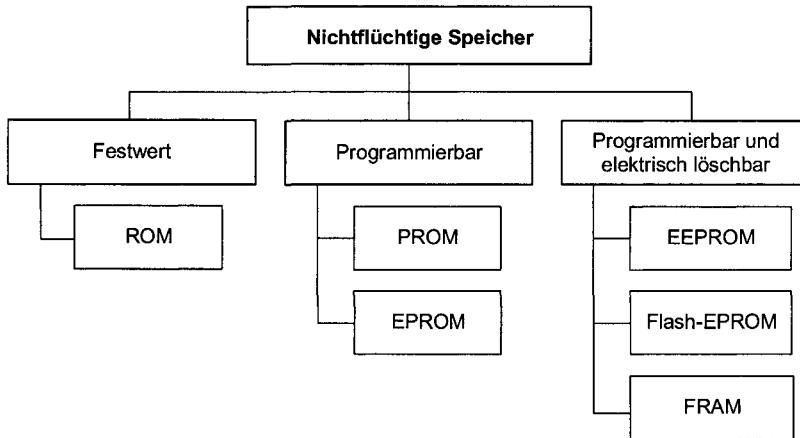


Abb. 7.12. Übersicht nichtflüchtiger Speicher

7.2.1.1 ROM

Der *ROM* (Read Only Memory) bezeichnet einen Festwertspeicher, der im Normalbetrieb nur ausgelesen und nicht beschrieben werden kann. Die Programmierung eines ROMs erfolgt während der technologischen Realisierung über eine spezielle Belichtungsmaske (Lithographie, vgl. Abschn. 2.4), die die Programm-informationen enthält. Dabei kann die Programmierung über eine sog. *Einsatzspannungsimplantation* erfolgen. Der CMOS-ROM ist in einer Matrixstruktur aus MOSFETs aufgebaut, wobei jeder MOSFET eine Speicherzelle mit einer Speichertiefe von z. B. 1 Bit (leitend oder nicht leitend) darstellt. Die MOSFETs sind jeweils mit einer sog. *Bitleitung* und einer sog. *Wortleitung* verbunden. Über die Wortleitung *WL* wird die entsprechende Speicherzelle ausgewählt und über die Bitleitung *BL* der zugehörige Speicherinhalt (Low oder High) ausgelesen. Das Schaltungsprinzip eines CMOS-ROMs mit n-Kanal-MOSFETs ist in Abb. 7.13 dargestellt.

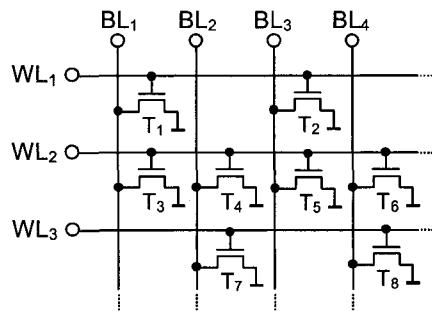


Abb. 7.13. Schaltungsprinzip eines CMOS-ROMs mit n-Kanal-MOSFETs zwischen Wortleitungen *WL* und Bitleitungen *BL*

Wird beispielsweise die Wortleitung WL_1 ausgewählt ($WL_1 = 1$), so wird die Einsatzspannung V_{Th} der n-Kanal-MOSFETs T_1 und T_2 überschritten, wodurch die Bitleitungen BL_1 und BL_3 auf Masse gezogen werden ($BL_1 = BL_3 = 0$). Die Bitleitungen BL_2 und BL_4 liegen hingegen auf Betriebsspannung ($BL_2 = BL_4 = 1$), da sie nicht über durchgeschaltete MOSFETs auf Masse liegen.

Die Herstellung flächeneffizienter ROMs kann durch Verwendung von mehrwertigen Speicherzellen (Speichertiefe > 1 Bit) erfolgen, die z. B. über unterschiedliche W/L-Verhältnisse der MOSFETs oder mehrere unterschiedliche Einsatzspannungen realisiert werden können.

7.2.1.2 PROM

Der *PROM* (Programmable ROM) bezeichnet einen anwenderprogrammierbaren Festwertspeicher. Hierbei kann die Verbindung zwischen Wort- und Bitleitungen technologisch durch dünne Metallstege (*Fuses*) realisiert werden, die bei Anlegen eines kurzen Programmierstromimpulses durchgeschmolzen werden. Ein weiteres Prinzip basiert darauf, dass zwischen zwei Metallbahnen eine dünne Isolatorschicht bei Anlegen einer Programmierspannung zerstört wird (*Antifuses*) und einen Kurzschluss zwischen den Leitungen verursacht. Das Schaltungsprinzip eines Fuse-basierten PROMs ist in Abb. 7.14 dargestellt.

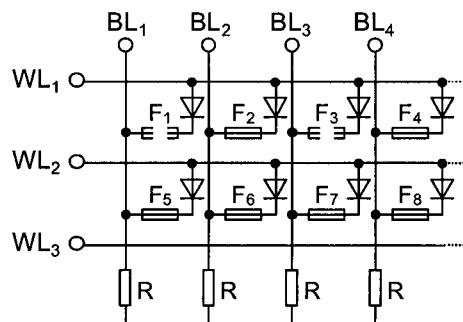


Abb. 7.14. Schaltungsprinzip eines Fuse-basierten PROMs

Die Bitleitungen BL_i werden jeweils über einen Widerstand R auf Masse gelegt. Im dargestellten Beispiel wurden die Fuses F_1 und F_3 durchgeschmolzen, sodass bei Auswahl der Wortleitung $WL_1 = 1$ die Bitleitungen BL_1 und BL_3 auf Masse liegen ($BL_1 = BL_3 = 0$). Die Bitleitungen BL_2 und BL_4 erhalten hingegen über die in Flussrichtung betriebenen Dioden das Potential der ausgewählten Wortleitung ($WL_1 = WL_2 = WL_4 = 1$), da die Fuses F_2 und F_4 leitend sind.

7.2.1.3 EPROM

Der *EPROM* (Erasable PROM) ist ein anwenderprogrammierbarer Speicher, der nicht elektrisch, sondern über Bestrahlung mit UV-Licht gelöscht werden kann. Die im EPROM verwendeten Speicherzellen sind MOSFETs, in deren Gateoxid

eine leitende Schicht aus Poly-Silizium eingebettet ist. Man spricht dann von einem sog. *Floating-Gate*. Der schematische Aufbau eines Floating-Gate-Speichertransistors ist in Abb. 7.15 im Querschnitt dargestellt.

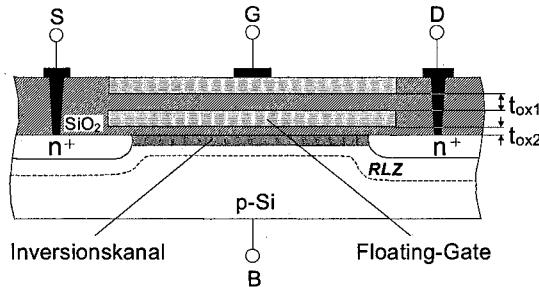


Abb. 7.15. Schematischer Aufbau eines Floating-Gate-Speichertransistors im Querschnitt

Die Programmierung eines Floating-Gate-Speichertransistors erfolgt durch einen Ladungstransfer vom Inversionskanal des MOSFETs zum Floating-Gate durch die untere Isolatorschicht, wobei für die Oxiddicken $t_{ox2} < t_{ox1}$ gilt. Dadurch resultiert eine Verschiebung der Einsatzspannung ΔV_T des MOSFETs. Das Schaltsymbol sowie die Übertragungskennlinie eines n-Kanal-Floating-Gate-Speichertransistors vor und nach dem Programmieren sind in Abb. 7.16 dargestellt.

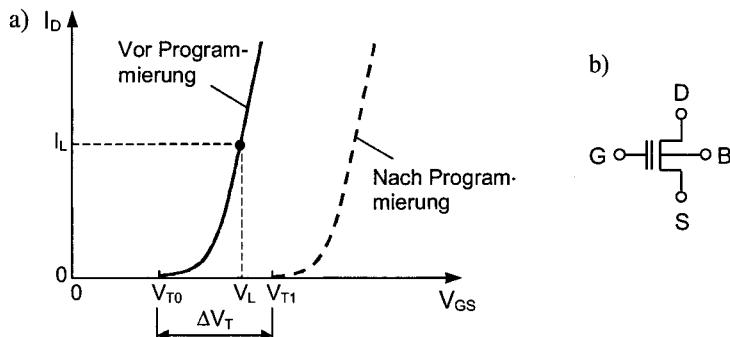


Abb. 7.16. a) Übertragungskennlinie eines n-Kanal-Floating-Gate-Speichertransistors vor und nach dem Programmervorgang b) Schalsymbol eines Floating-Gate-Speichertransistors

Die Verschiebung der Einsatzspannung ΔV_T berechnet sich aus der Ladung Q_F des Floating-Gates und der Oxidkapazität C_I zwischen Gate und Floating-Gate wie folgt:

$$\Delta V_T = -\frac{Q_F}{C_I} \quad \text{mit} \quad C_I = \frac{\epsilon_0 \cdot \epsilon_{ox} \cdot W \cdot L}{t_{ox1}}. \quad (7.8)$$

Die Speicherinformation des Floating-Gate-MOSFETs wird über eine Lese- spannung V_L ($V_{T0} < V_L < V_{T1}$) ausgelesen. Wurde der Speichertransistor program-

miert, so besitzt der resultierende Lesestrom den Wert $I_L \approx 0$ und im unprogrammierten Zustand $I_L > 0$ (vgl. Abb. 7.16).

Die Programmierung eines Floating-Gate-Speichertransistors erfolgt durch die sog. *Hot-Electron-Injektion*. Dabei erfolgt die Injektion von heißen Elektronen in das Floating-Gate durch Anlegen einer bestimmten Drain-Source-Spannung V_{DS} (Programmierspannung). Dadurch erhalten die Ladungsträger am drainseitigen Ende des Inversionskanals eine hohe kinetische Energie. Ein geringer Anteil dieser Ladungsträger mit einer Energie $E > 3,2$ eV kann dann die Potentialbarriere zwischen Inversionskanal und Floating-Gate überwinden und durch das Gateoxid auf das Floating-Gate wechseln. Auf Grund der Höhe des Drainstroms I_D , der viel größer ist als der Programmierstrom ($I_D \gg I_P$), ist ein hoher Energiebedarf beim Programmieren notwendig. Als Kenngröße für die Programmireffizienz dient dabei das Verhältnis von Programmierstrom zum Drainstrom $\eta_P = I_P/I_D$. Um einem Ladungsverlust während der Lebensdauer des EPROMs vorzubeugen, wird zumeist eine mehrfache *Überprogrammierung* der Information durchgeführt.

Der Aufbau des EPROMs erfolgt analog zum ROM in einer Matrixstruktur, wobei jeder Floating-Gate-Speichertransistor eine Speicherzelle darstellt. Das Auslesen des Speicherinhalts erfolgt analog zum ROM (vgl. Abschn. 7.2.1.1). Das Schaltungsprinzip eines EPROMs ist in Abb. 7.17 ersichtlich.

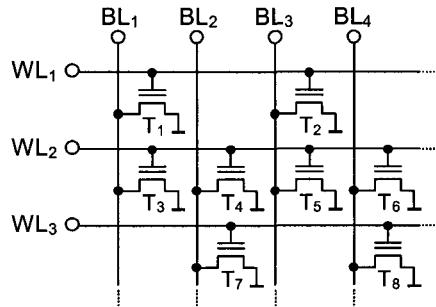


Abb. 7.17. Schaltungsprinzip eines EPROMs mit Floating-Gate-Speichertransistoren

Der Löschvorgang eines EPROMs, der aus Floating-Gate-Speichertransistoren aufgebaut ist und über Hot-Electron-Injektion programmiert wurde, erfolgt durch Einstrahlung von UV-Licht. Dadurch erhalten die im Floating-Gate gespeicherten Ladungsträger durch optische Anregung eine genügend hohe Energie ($E > 3,2$ eV), um wieder in den Inversionskanal oder in das Gate zu gelangen. Der Löschvorgang wird durch ein für UV-Licht durchlässiges Quarzglasfenster im Gehäuse des EPROMs vorgenommen und dauert einige Minuten.

7.2.1.4 EEPROM

Der *EEPROM* (Electrical Erasable PROM) ist im Gegensatz zum EPROM elektrisch lösbar. Als Speicherzelle dient ebenfalls ein Floating-Gate-Speichertransistor (vgl. Abschn. 7.2.1.3). Jedoch wird hier das Oxid zwischen Floating-Gate und Draingebiet so dünn ausgelegt (Dicke t_T), dass ein Tunnelfenster entsteht,

durch welches bei hoher Gate-Drain-Spannung V_{GD} ein merklicher Tunnelstrom fließen kann. Dadurch kann das Floating-Gate durch Tunnelströme geladen und auch entladen werden. Der Programmervorgang des Floating-Gate-Speichertransistors mit Tunnelfenster wird als *Fowler-Nordheim-Tunneln* bezeichnet; der Energieverbrauch ist geringer als bei der Hot-Electron-Injektion (vgl. Abschn. 7.2.1.3). Der schematische Aufbau eines Floating-Gate-Speichertransistors mit Tunnelfenster ist in Abb. 7.18 im Querschnitt dargestellt.

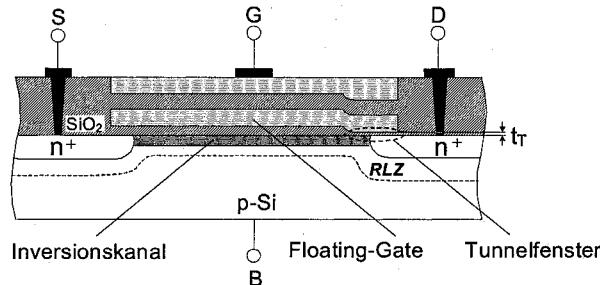


Abb. 7.18. Schematischer Aufbau eines Floating-Gate-Speichertransistors mit Tunnelfenster (Querschnitt)

Die Vorteile des EEPROMs bestehen darin, dass ein elektrisches Schreiben und Löschen möglich ist und beim Schreiben nur der Tunnelstrom (kein Drainstrom I_D) fließt. Nachteilig ist dagegen, dass die Ansteuerung durch einen zusätzlichen Auswahltransistor stattfinden muss, um ein unbeabsichtigtes Programmieren bzw. Löschen benachbarter Speicherzellen durch Leckströme zu verhindern. Die EEPROM-Speicherzelle besitzt daher eine relativ große Fläche. Das Schaltungsprinzip eines EEPROMs mit Floating-Gate-Speichertransistoren (mit Tunnelfenster) und Auswahltransistoren ist in Abb. 7.19 dargestellt.

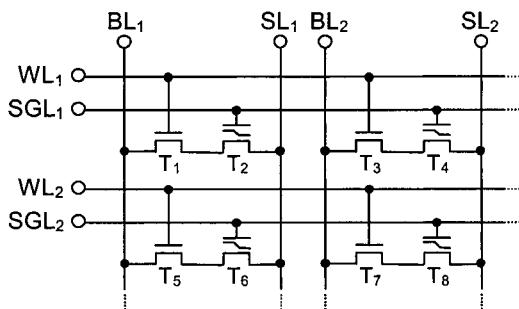


Abb. 7.19. Schaltungsprinzip eines EEPROMs mit Floating-Gate-Speichertransistoren (mit Tunnelfenster) und Auswahltransistoren

Wird beispielsweise die Wortleitung WL_1 ausgewählt ($WL_1 = 1$), so wird die Einsatzspannung V_T des Auswahltransistors T_1 überschritten, wodurch die auf $BL_1 = 1$ vorgeladene Bitleitung BL_1 abhängig vom Programmierzustand des Spei-

chertransistors T_2 (Floating-Gate geladen oder ungeladen) über die Sourceleitung SL_1 den Wert $BL_1 = 1$ behält oder auf $BL_1 = 0$ umgeladen wird. Dabei besitzt der Speichertransistor T_2 eine separate Steuergateleitung SGL_1 .

7.2.1.5 Flash-EPROM

Der sog. *Flash-EPROM* ist ein weiterer anwenderprogrammierbarer Speicher, der elektrisch löschbar ist. Hierbei kann jedoch nicht eine Einzelspeicherzelle, sondern nur ein ganzer Sektor gelöscht werden (z. B. alle Speicherzellen an einer Bitleitung). Zur Realisierung von Flash-EPROMs können verschiedene Speicherzellen angewendet werden. In sog. *CHE-Flash-Zellen* erfolgt die Programmierung durch Hot-Electron-Injektion (vgl. Abschn. 7.2.1.3) und das Löschen durch Fowler-Nordheim-Tunneln (vgl. Abschn. 7.2.1.4). In sog. *FN-Flash-Zellen* wird dagegen die Programmierung und das Löschen durch das Prinzip des Fowler-Nordheim-Tunnelns realisiert. In Abb. 7.20 ist das beispielhafte Schaltungsprinzip eines Flash-EPROMs in NAND-Architektur dargestellt.

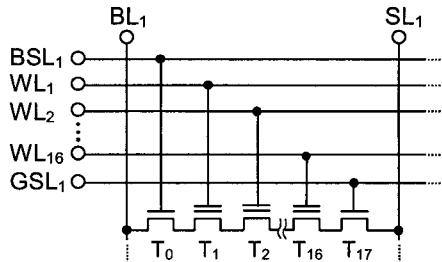


Abb. 7.20. Schaltungsprinzip eines Flash-EPROMs in NAND-Architektur

Der Aufbau einer Einheit besteht aus zwei Auswahltransistoren T_0 und T_{17} sowie z. B. 16 Floating-Gate-Speichertransistoren $T_1 - T_{16}$. Die in Serie geschalteten Floating-Gate-Speichertransistoren können nur seriell beschrieben und über die Bitleitung BL_1 ausgelesen werden. Der Vorteil dieses Prinzips liegt in der hohen Flächeneffizienz und der damit verbundenen preisgünstigen Realisierung.

7.2.1.6 FRAM

Der *FRAM* (Ferroelectric Random Access Memory) besitzt einen ähnlichen Aufbau wie ein DRAM (vgl. Abschn. 7.2.2.2), wobei die Speicherzelle aus einer Speicherkapazität und einem Auswahltransistor aufgebaut ist. Beim FRAM besitzt das Dielektrikum der Speicherkapazität jedoch ferroelektrische Eigenschaften und weist eine Dielektrizitätskonstante im Bereich $1000 \leq \epsilon \leq 1500$ auf. Als Materialien für das ferroelektrische Dielektrikum werden vor allem Blei-Zirkonium-Titanat (*PZT*) oder Strontium-Wismut-Tantalat (*SBT*) verwendet. In Abb. 7.21 ist das Schaltungsprinzip eines FRAMs dargestellt, wobei die Speicherzellen jeweils aus einem Auswahltransistor und einer ferroelektrischen Speicherkapazität bestehen (*IT/IC-Speicherzelle*).

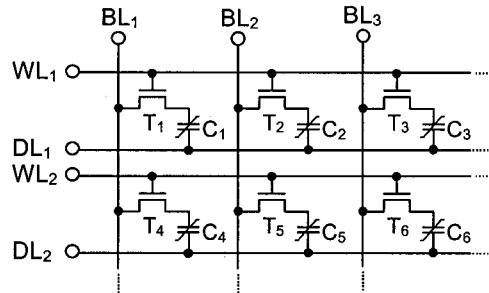


Abb. 7.21. Schaltungsprinzip eines FRAMs mit Speicherzellen, die jeweils aus einem Auswahltransistor und einer ferroelektrischen Speicherkapazität bestehen (1T/1C-Speicherzelle)

Wird z. B. die Wortleitung WL_1 des FRAMs ausgewählt ($WL_1 = 1$), so wird die Einsatzspannung V_{Th} des Auswahltransistors T_1 überschritten, wodurch über die Bitleitung BL_1 abhängig von der Polarisation der Speicherkapazität C_1 der Wert $BL_1 = 0$ oder $BL_1 = 1$ ausgegeben wird. Die Speicherinformation erhält die Kapazität durch eine Polarisation mit Spannungsimpulsen über die Datenleitung DL_1 . Der Wert der Polarisation bzw. die daraus resultierende Speicherinformation High oder Low kann dann über Spannungsimpulse auf der Bitleitung wieder ausgelesen werden. Die Grundlage dafür ist die spannungsabhängige Hysterese der dielektrischen Polarisation $P(V)$, deren typischer Verlauf mit den korrespondierenden Speicherzuständen in Abb. 7.22 dargestellt ist.

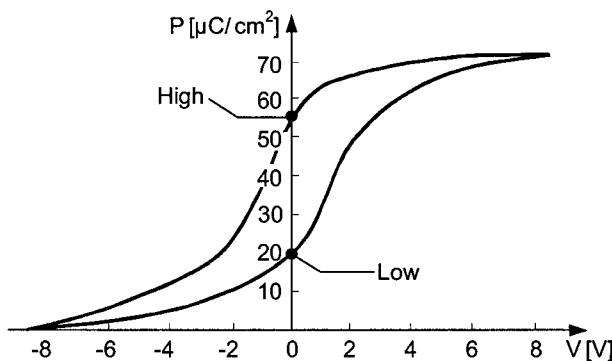


Abb. 7.22. Typischer Verlauf der spannungsabhängigen Hysterese der dielektrischen Polarisation P einer FRAM-Speicherzelle

Die Ladungsdifferenz zwischen dem Speicherzustand High und Low einer FRAM-Speicherzelle liegt bei $\Delta Q > 20 \mu\text{C}/\text{cm}^2$. Es erfolgt eine nichtflüchtige Speicherung im ferroelektrischen Medium, da eine permanente elektrische Polarisation des Dielektrikums vorliegt.

Das Auslesen einer FRAM-Speicherzelle erfolgt destruktiv, d. h. der Speicherinhalt kann beim Auslesen zerstört werden und muss nach dem Auslesen wieder in

die Speicherzelle geschrieben werden (*Restore*). Das Auslesen erfolgt durch einen negativen Spannungsimpuls an der ferroelektrischen Speicherkapazität. Für den Speicherinhalt High erfolgt ein Ladungsfluss über die Bitleitung nach dem aber das programmierte Low in der Zelle vorliegt (Zerstören des High-Speicherinhaltes). Es ist ein Restore der High-Information durch einen positiven Spannungsimpuls erforderlich. Für den Speicherinhalt Low findet kein Netto-Ladungsfluss über die Bitleitung statt, sodass der Low-Zustand der Zelle erhalten bleibt und ein Restore nicht erforderlich ist.

7.2.2 Flüchtige Speicher

Im Gegensatz zu nichtflüchtigen Speichern (vgl. Abschn. 7.2.1) verlieren flüchtige Speicher ihre Speicherinformation nach Abschalten der Betriebsspannung. Flüchtige Speicher sind sog. *Schreib-Lese-Speicher*, die als *RAM* (Random Access Memory) bezeichnet werden. Sie gestatten den wahlfreien Zugriff auf einen Speicherort, d. h. über die Auswahl einer Adresse erfolgt die Speicherung und das Auslesen von Informationen einer bestimmten Speicherzelle. Die wesentlichen Ausführungsformen des RAMs werden nachfolgend näher beschrieben.

7.2.2.1 SRAM

Im *SRAM* (Statischer RAM) wird die Information als Schaltzustand einer bistabilen Kippschaltung (*Flipflop*) gespeichert, die den Zustand High oder Low annehmen kann. Der Flipflop wird durch zwei rückgekoppelte CMOS-Inverter (vgl. Abschn. 4.8.2) realisiert, die durch zwei Auswahltransistoren mit den komplementären Bitleitungen verbunden sind (*6-Transistor-Zelle*). Die Adressierung erfolgt durch Auswahl von bestimmten Wortleitungen. Solange die Betriebsspannung anliegt, bleibt der Speicherinhalt erhalten (statisches Verhalten) und muss nicht wie beim dynamischen RAM (vgl. Abschn. 7.2.2.2) zyklisch aufgefrischt werden. Das Schaltungsprinzip eines SRAMs in CMOS-Technologie ist in Abb. 7.23 dargestellt.

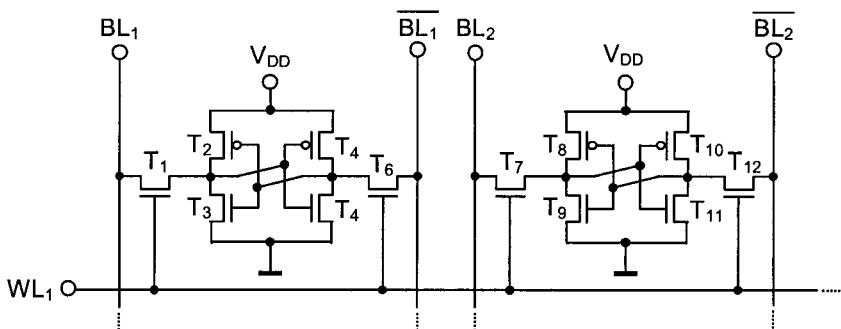


Abb. 7.23. Schaltungsprinzip eines SRAMs in CMOS-Technologie mit 6-Transistor-Speicherzellen

Die Verwendung komplementärer Bitleitungen BL und \overline{BL} verringert die Empfindlichkeit gegenüber Parameterstreuungen der MOSFETs. Beim Lesevorgang wird die Spannungsdifferenz zwischen den komplementären Bitleitungen mit Hilfe eines Ausleseverstärkers verstärkt, der durch einen Differenzverstärker gemäß Abschn. 4.8.5, Abb. 4.40 realisiert werden kann.

Als charakteristische Größe für das zeitliche Verhalten des Auslesevorgangs dient die sog. *Zugriffszeit*. Sie gibt an, in welcher Zeit nach Anlegen der Adressen die Information an den Datenausgängen des SRAMs zur Verfügung steht. Die Zugriffszeit kann verringert werden, wenn nicht die Spannungsdifferenz zwischen den komplementären Bitleitungen ausgewertet wird, sondern der durch die ausgewählte Speicherzelle fließende Strom bewertet wird. SRAMs besitzen eine sehr geringe Zugriffszeit, weisen aber einen relativ hohen Flächenverbrauch auf, so dass die Speichergröße begrenzt ist. Flächensparende SRAM-Zellen in Spezialprozessen können durch Ersetzen der p-Kanal-MOSFETs der Inverter (vgl. Abb. 7.23) durch Widerstände aus Poly-Silizium, die in der Ebene über den n-Kanal-MOSFETs angeordnet sind, erreicht werden (*4-Transistor-Zelle*). SRAMs werden vor allem als schnelle Zwischenspeicher (*Cache-Speicher*) in Rechnersystemen eingesetzt.

7.2.2.2 DRAM

Der *DRAM* (Dynamischer RAM) speichert die Information als Ladungsmenge in einer Speicherkapazität. Da die Ladung der Speicherkapazität zyklisch aufgefrischt werden muss, spricht man hier im Gegensatz zum statischen RAM von einer dynamischen Schaltung. Eine DRAM-Speicherzelle in CMOS-Technologie besteht somit aus einer Speicherkapazität und einem Auswahltransistor (*1-Transistor-Zelle*). Das Schaltungsprinzip eines DRAMs ist in Abb. 7.24 abgebildet.

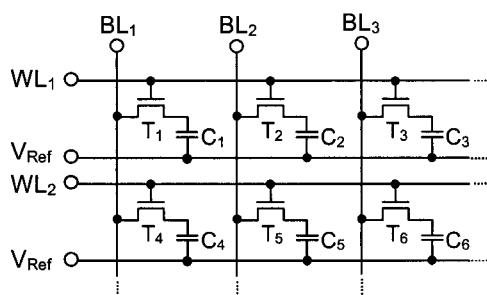


Abb. 7.24. Schaltungsprinzip eines DRAMs mit Speicherkapazität und Auswahltransistor (1-Transistor-Zelle)

Bei der Programmierung wird die Speicherkapazität z. B. C_1 , die zuvor durch die Wortleitung WL_1 ausgewählt wurde, über den Auswahltransistor T_1 auf das Potential der Bitleitung $BL_1 = 1$ oder $BL_1 = 0$ aufgeladen. An einem Anschluss der Speicherkapazität liegt dabei eine Referenzspannung (z. B. $V_{Ref} = V_{DD}/2$). Im Gegensatz zur Verwendung von Masse als Referenzspannung wird dadurch der

Spannungsabfall über dem Dielektrikum der Speicherkapazität auf $V_{DD}/2$ reduziert.

Das Auslesen des DRAMs erfolgt ebenfalls durch Selektion der entsprechenden Speicherzelle über eine Wortleitung (z. B. $WL_1 = 1$). Abhängig von der Ladung der Speicherkapazität C_1 wird die auf ca. $V_{DD}/2$ vorgeladene Bitleitung geringfügig aufgeladen oder entladen. In einem Leseverstärker wird diese Differenz auf die Ausgangsspannung High oder Low verstärkt. Anschließend wird die Information in die Speicherkapazität zurückgeschrieben (Restore). Da die Ladung der Speicherkapazität durch Leckströme verringert wird, kann es zum Informationsverlust der DRAM-Speicherzelle kommen. Um diesem Ladungsverlust entgegenzuwirken, werden die Speicherinhalte zyklisch aufgefrischt (Refresh). Dabei wird durch eine entsprechende Refresh-Logik der Speicherinhalt ausgelesen, bewertet und wieder auf die Speicherkapazität übertragen. Während der Refresh-Zeit, die einige Millisekunden beträgt, kann kein regulärer Zugriff auf den Inhalt der Speicherzelle erfolgen.

Zur Realisierung der Speicherkapazität kann in CMOS-Technologie das dünne Gateoxid der Dicke t_{ox} als Dielektrikum genutzt werden. Die Speicherkapazität $C = \epsilon_0 \cdot \epsilon_{ox} \cdot A / t_{ox}$ besitzt typische Werte im Bereich $25\text{ fF} \leq C \leq 50\text{ fF}$. Der Querschnitt einer planaren DRAM-Speicherzelle in CMOS-Technologie ist in Abb. 7.25 schematisch dargestellt.

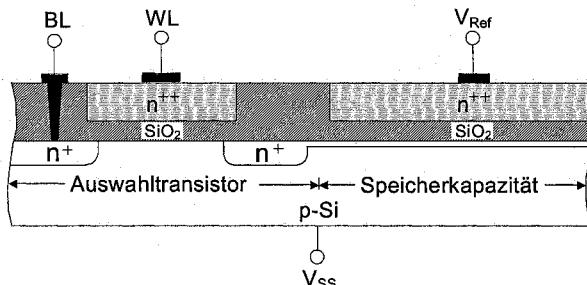


Abb. 7.25. Schematischer Aufbau (Querschnitt) einer planaren DRAM-Speicherzelle in CMOS-Technologie mit Speicherkapazität und Auswahltransistor

Um eine ausreichende Störsicherheit z. B. gegen Leckströme und Einkopplungen zu erreichen, muss eine hinreichende Größe der Speicherladung und damit der Speicherkapazität berücksichtigt werden. Dadurch kann die Speicherkapazität nicht mit der Technologie skaliert werden, was bei Verwendung einer planaren DRAM-Speicherzelle gemäß Abb. 7.25 eine Begrenzung der Integrationsdichte zur Folge hat. Dieser Begrenzung kann mit der Verwendung von dreidimensionalen Strukturen wie der sog. *Trench-Speicherkapazität* oder der sog. *Stack-Speicherkapazität* entgegengewirkt werden.

DRAMs besitzen auf Grund ihres relativ einfachen Aufbaus eine hohe Packungsdichte, wobei Speichergrößen von z. B. 1 GBit erreicht werden. Die Zugriffszeit von DRAMs ist hingegen höher als die von SRAMs (vgl. Abschn. 7.2.2.1). DRAMs werden vor allem als Hauptspeicher in Rechnersystemen eingesetzt.

7.3 Integrierte Logik-Bausteine

Integrierte Logik-Bausteine können in die Rubriken programmierbare Logikbausteine (*PLD*), kundenspezifische Logikbausteine (*ASIC*) und Mikroprozessoren eingeteilt werden. Sie unterscheiden sich in der Anzahl ihrer kundenspezifischen Fertigungsschritte und der realisierbaren Schaltungsfunktion. In Abb. 7.26 ist ein Überblick über die wesentlichen Logik-Bausteine dargestellt.

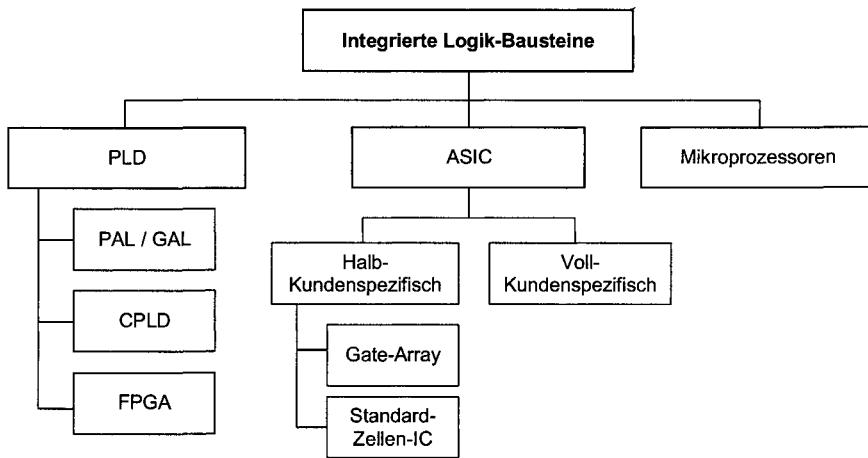


Abb. 7.26. Integrierte Logik-Bausteine im Überblick

Die Gruppe der PLDs verfügt über keine kundenspezifischen Fertigungsschritte und die Schaltungsfunktion wird durch den Anwender bestimmt. Im Gegensatz zum PLD werden bei ASICs einige oder alle Fertigungsschritte kundenspezifisch durchgeführt. Die Schaltungsfunktion wird jedoch auch durch den Anwender festgelegt. Bei nicht kundenspezifischen Logik-Bausteinen wird die Funktion durch ein Programm festgelegt, das z. B. in Mikroprozessoren über eine *Arithmetik-Logik-Einheit*, einen *Befehlsdekoder* und eine Ablaufsteuerung abgearbeitet wird. Im Folgenden werden die wesentlichen Vertreter von PLDs und ASICs näher beschrieben

7.3.1 PLD

Beim PLD (Programmable Logic Device) erfolgt die Programmierung der gewünschten Abhängigkeit der Ausgangsvariablen von den Eingangsvariablen durch einen Verknüpfungsalgorithmus. Ein PLD besteht aus einem Eingangsblock, der eine Verstärkung der Eingangssignale über Treiber und die Bereitstellung der negierten und nichtnegierten Eingangssignale realisiert. Weiterhin beinhaltet ein PLD ein AND-Array zur Realisierung der *konjunktiven* Terme und ein OR-Array zur Realisierung der *disjunktiven* Verknüpfungen. Die Verbindungen zwischen dem

Eingangsblock und AND-Array bzw. AND- und OR-Array sind je nach Ausführung festverdrahtet oder programmierbar.

7.3.1.1 PAL und GAL

Der *PAL* (Programmable Array Logic) besteht aus einem programmierbaren AND-Array und einem festverdrahteten OR-Array. Der Aufbau eines PALs ist in Abb. 7.27 dargestellt.

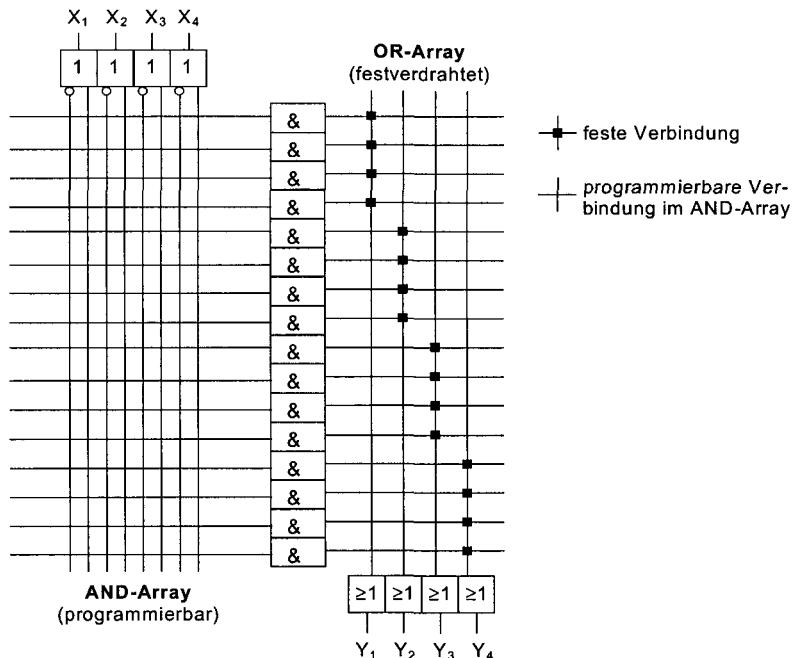


Abb. 7.27. Schaltungsprinzip eines PALs mit programmierbarem AND-Array und festverdrahtetem OR-Array

Die Programmierung eines PALs erfolgt analog zum PROM z. B. durch einen kurzen Programmierstromimpuls, bei dem die Verbindung zwischen den ausgewählten Zeilen und Spalten des AND-Arrays über ein Durchschmelzen von Fuses realisiert werden kann (vgl. Abschn. 7.2.1.2). Die logische Verknüpfung wird dann z. B. durch Eingabe der booleschen Gleichungen oder durch Eingabe einer Zustandstabelle über eine Programmiersoftware realisiert.

Der *GAL* (Generic Array Logic) besteht im Gegensatz zum PAL aus einem programmierbaren AND-Array und einem programmierbaren OR-Array. Der Vorteil des GALs gegenüber dem PAL besteht darin, dass gleiche Terme einer booleschen Gleichung nur einmal zu realisieren sind. Nachteilig stellt sich jedoch die aufwendigere Schaltungstopologie dar. Der Aufbau eines GALs ist in Abb. 7.28 dargestellt.

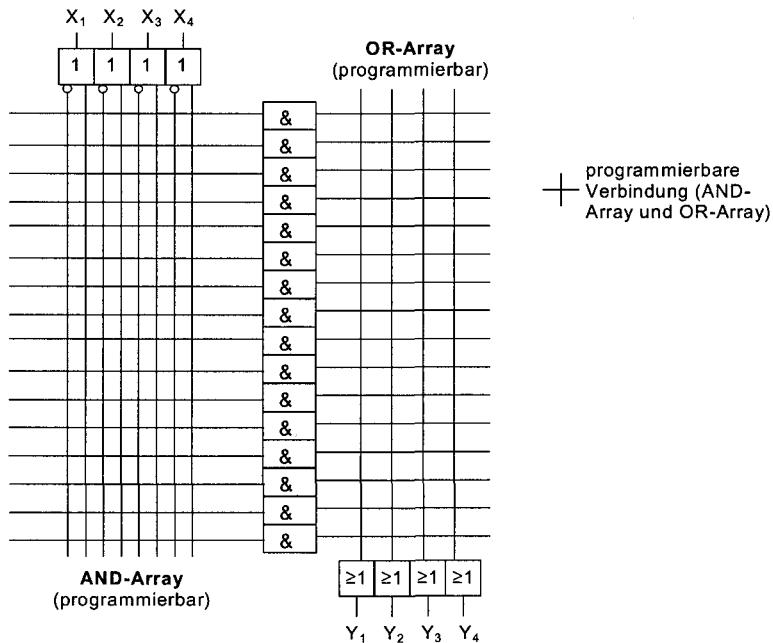


Abb. 7.28. Schaltungsprinzip eines GALs mit frei programmierbarem AND- und OR-Array

Der GAL bietet durch die frei programmierbaren AND- und OR-Arrays eine hohe Flexibilität bei der Programmierung logischer Verknüpfungen. Der Programmieraufwand entspricht dem des PALs.

7.3.1.2 CPLD

Ein *CPLD* (Complex PLD) ist eine Mehrfach-Array-Struktur mit PAL-ähnlichen Schaltungsblöcken (vgl. Abschn. 7.3.1.2). Die Verknüpfung der Funktionsblöcke erfolgt über eine komplexe Schaltmatrix, deren Einzelverknüpfungen über EEPROM- oder Flash-EPROM-Speicherzellen (vgl. Abschn. 7.2) realisiert werden. Die Verdrahtung verläuft zwischen den einzelnen Blöcken über Verdrahtungskanäle. Die Vorteile von CPLDs bestehen in der hohen Schaltgeschwindigkeit und in der Vorhersagbarkeit der Signallaufzeiten zwischen den Anschlüssen. Der Nachteil solcher CPLDs besteht darin, dass keine RAM- oder ROM-Komponenten implementierbar sind und dass das CPLD während des Betriebs nicht rekonfigurierbar ist.

7.3.1.3 FPGA

Beim *FPGA* (Field Programmable Gate Array) erfolgt eine Programmierung der Funktion einzelner Logikblöcke sowie die Programmierung horizontaler und vertikaler Verbindungen zwischen den Logikblöcken. Der schematische Aufbau eines FPGAs ist in Abb. 7.29 dargestellt.

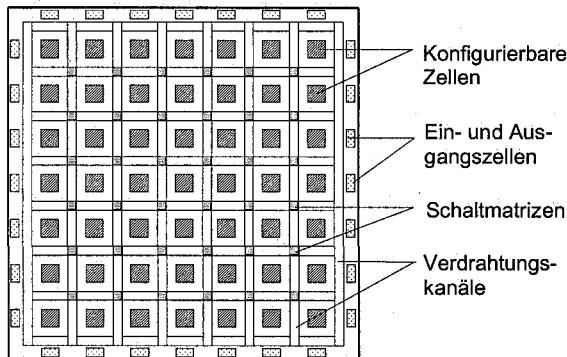


Abb. 7.29. Schematischer Aufbau eines FPGAs

Die Programmierung eines FPGAs kann sowohl nichtflüchtig als auch flüchtig erfolgen. Bei der nichtflüchtigen Programmierung werden bei Anlegen einer Programmierspannung die ausgewählten Antifuses dauerhaft leitend (vgl. Abschn. 7.2.1.2). Vorteile dieser Antifuse-basierten FPGAs sind die hohe Programmiersgeschwindigkeit, da nur ein geringer Teil der möglichen Verbindungen genutzt wird (ca. 2%). Weiterhin besitzen Antifuse-basierte FPGAs einen geringen Flächenbedarf und die Programmierinformation bleibt auch nach Abschalten der Betriebsspannung erhalten. Der Nachteil Antifuse-basierter FPGAs besteht darin, dass sie nur einmal programmierbar sind und die Signallaufzeit zwischen den Anschlüssen nicht exakt vorhersagbar ist.

Eine flüchtige Programmierung von FPGAs kann über SRAM-Zellen erfolgen (vgl. Abschn. 7.2.2.1). Diese Art von FPGAs enthält zumeist Logikelemente, die in Form von SRAMs aufgebaut sind und in Wertetabellen von booleschen Funktionen geladen werden können. Verschieden lange Leitungselemente zwischen den Logikelementen und einem Ein-/Ausgangsblock können dann durch sog. *Pass-Transistoren* verbunden werden. Dabei erfolgt die Gate-Ansteuerung dieser Pass-Transistoren jeweils durch eine SRAM-Speicherzelle. Eine typische Architektur eines SRAM-basierten FPGAs ist in Abb. 7.30 dargestellt.

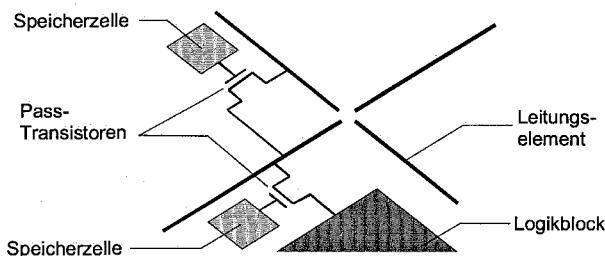


Abb. 7.30. Schematischer Aufbau eines SRAM-basierten FPGAs mit Pass-Transistoren

Im Gegensatz zum Antifuse-basierten FPGA ist ein SRAM-basierter FPGA beliebig oft programmierbar. Allerdings geht die Programmierinformation nach Ab-

schalten der Betriebsspannung verloren. Es wird dadurch ein Nachladen der Programmierinformation z. B. aus einem EPROM (vgl. Abschn. 7.2.1.3) erforderlich. SRAM-basierte FPGAs besitzen zudem einen hohen Flächenbedarf und weisen einen Geschwindigkeitsverlust auf Grund der Pass-Transistoren in den Signalpfaden auf.

7.3.2 ASIC

Der *ASIC* (Application Specific IC) bezeichnet eine integrierte Schaltung, die kundenspezifische Elemente enthält. Man unterscheidet hierbei halb-kundenspezifische und voll-kundenspezifische Ausführungen. Als halb-kundenspezifische Schaltungen kommen zumeist das Gate-Array und der Standardzellen-IC zum Einsatz.

7.3.2.1 Gate-Array

Das *Gate-Array* besteht aus einem vorgefertigten Feld von n- und p-Kanal-MOSFETs. Durch eine kundenspezifische Fertigung der Metallisierung werden die logischen Funktionen realisiert. Das Gate-Array ist vorgefertigt bis auf die Verbindungsleitungen (meist drei Metall-Lagen). Es besitzt eine kurze Fertigungszeit, da nur drei kundenspezifische Fertigungsschritte notwendig sind. Gate-Arrays haben schnellere Schaltzeiten als PLDs und sind flächeneffizienter.

7.3.2.2 Standardzellen-IC

Beim Standardzellen-IC werden vom Hersteller vorentwickelte Grundzellen kundenspezifisch platziert und verdrahtet. Dabei erfolgt eine Anordnung der Zellen in Reihen mit variablem Abstand (Verdrahtungskanäle). Der schematische Aufbau eines Standardzellen-ICs ist in Abb. 7.31 dargestellt.

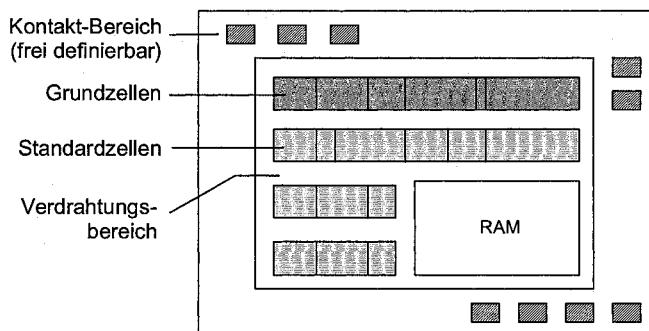


Abb. 7.31. Schematischer Aufbau eines Standardzellen-ICs

Gegenüber Gate-Arrays besitzen Standardzellen-ICs eine höhere Packungsdichte und eine höhere erzielbare Systemfrequenz. Nachteilig stellt sich die längere

re Fertigungszeit dar, da eine höhere Anzahl von kundenspezifischen Fertigungsschritten durchgeführt werden muss.

7.3.2.3 Voll-kundenspezifischer IC

Der voll-kundenspezifische IC wird vom Hersteller vollständig nach Spezifikation des Anwenders gefertigt. Die Elemente werden dabei aus Polygonzügen (Einzeltransistoren) aufgebaut. Der voll-kundenspezifische IC ist optimal auf das Anwenderproblem angepasst, besitzt jedoch eine hohe Entwicklungs- und Fertigungszeit und rentiert sich erst bei sehr hohen Stückzahlen (> 1 Mio.).

Literaturverzeichnis

Kapitel 1

- [1.1] Müller, R., „Grundlagen der Halbleiter-Elektronik“, Springer-Verlag, Berlin, Heidelberg, New York, 1984.
- [1.2] Paul, R., „Elektronische Halbleiterbauelemente“, B. G. Teubner, Stuttgart, Heidelberg, 1986.
- [1.3] Sze, S. M., „Physics of Semiconductor Devices“, J. Wiley, New York, 1982.
- [1.4] Seeger, K., „Semiconductor Devices“, Springer-Verlag, Berlin, 1985.
- [1.5] Grove, A. S., „Physics and Technology of Semiconductor Devices“, J. Wiley, New York, 1967.
- [1.6] Hess, K., „Theory of Semiconductor Devices“, IEEE Press, New York, 2000.

Kapitel 2

- [2.1] Müller, R., „Grundlagen der Halbleiter-Elektronik“, Springer-Verlag, Berlin, Heidelberg, New York, 1984.
- [2.2] Reisch, M., „Elektronische Bauelemente. Funktion, Grundschatungen, Modellierung mit SPICE“, Springer-Verlag, Berlin, Heidelberg, New York, 1998.
- [2.3] Paul, R., „Elektronische Halbleiterbauelemente“, B. G. Teubner, Stuttgart, Heidelberg, 1986.
- [2.4] Bleicher, M., „Halbleiter-Optoelektronik“, Hüthig Verlag, Heidelberg, 1986.
- [2.5] Sze, S. M., „Physics of Semiconductor Devices“, J. Wiley & Sons, New York, 1982.
- [2.6] Kessel, G., Hammerschmitt, J., Lange, E., „Signalverarbeitende Dioden“, Springer-Verlag, Berlin, Heidelberg, New York, 1982.

Kapitel 3

- [3.1] Gray, P. R., Meyer, R. G., „Analysis and Design of Analog Integrated Circuits“, John Wiley & Sons, New York, Chichester, 1993.
- [3.2] Müller, R., „Bauelemente der Halbleiter-Elektronik“, Springer-Verlag, Berlin, Heidelberg, New York, 1991.
- [3.3] Reisch, M., „Elektronische Bauelemente. Funktion, Grundschatungen, Modellierung mit SPICE“, Springer-Verlag, Berlin, Heidelberg, New York, 1998.

-
- [3.4] Paul, R., „Elektronische Halbleiterbauelemente“, B. G. Teubner, Stuttgart, Heidelberg, 1986.
 - [3.5] Tietze, U., Schenk, C., „Halbleiter-Schaltungstechnik“, Springer-Verlag, Berlin, Heidelberg, New York, 1993.
 - [3.6] Schrenk, H., „Bipolare Transistoren“, Springer-Verlag, Berlin, Heidelberg, New York, 1978.
 - [3.7] Gray, P. R., Hurst, P. J., Lewis, S. H., Meyer, R. G., „Analysis and Design of Analog Integrated Circuits“, John Wiley & Sons, New York, Chichester, 2001.
 - [3.8] Howe, R. T., Sodini, C. G., „Microelectronics. An Integrated Approach“, Prentice Hall, New Jersey, 1997.
 - [3.9] Sze, S. M., „Physics of Semiconductor Devices“, J. Wiley & Sons, New York, 1982.

Kapitel 4

- [4.1] Sze, S. M., „Physics of Semiconductor Devices“, J. Wiley & Sons, New York, 1982.
- [4.2] Arora, N., „MOSFET Models for VLSI Circuit Simulation. Theory and Practice“, Springer-Verlag, Wien, New York, 1993.
- [4.3] Paul, R., „MOS-Feldeffekttransistoren“, Springer-Verlag, Berlin, Heidelberg, New York 1994.
- [4.5] Müller, R., „Bauelemente der Halbleiter-Elektronik“, Springer-Verlag, Berlin, Heidelberg, New York, 1991.
- [4.6] Reisch, M., „Elektronische Bauelemente. Funktion, Grundschaltungen, Modellierung mit SPICE“, Springer-Verlag, Berlin, Heidelberg, New York, 1998.
- [4.7] Taur, Y., Ning, T. H., „Fundamentals of modern VLSI Devices“, Cambridge University Press, 1998.
- [4.8] Gray, P. R., Meyer, R. G., „Analysis and Design of Analog Integrated Circuits“, John Wiley & Sons, New York, Chichester, 1993.
- [4.9] Allen, E. P., Holberg, D. R., „CMOS Analog Circuit Design“, Oxford University Press, New York, Oxford, 1987.
- [4.10] Paul, R., „Elektronische Halbleiterbauelemente“, B. G. Teubner, Stuttgart, Heidelberg, 1986.
- [4.11] Razavi, B., „Design of Analog CMOS Integrated Circuits“, McGraw Hill, New York, 2001.
- [4.12] Gray, P. R., Hurst, P. J., Lewis, S. H., Meyer, R. G., „Analysis and Design of Analog Integrated Circuits“, John Wiley & Sons, New York, Chichester, 2001.
- [4.13] Howe, R. T., Sodini, C. G., „Microelectronics. An Integrated Approach“, Prentice Hall, New Jersey, 1997.
- [4.14] Widmann, D., Mader, H., Friedrich, H., „Technologie hochintegrierter Schaltungen“, Springer-Verlag, Berlin, Heidelberg, New York, 1996.
- [4.15] Tsividis, Y. P., „Operation and Modeling of the MOS Transistor“, McGraw Hill, New York, 1988.

- [4.16] Selberherr, S., „Analysis and Simulation of Semiconductor Devices“, Springer-Verlag, Wien, New-York, 1984.
- [4.17] Chen, W.-K., „The VLSI Handbook“, CRC Press & IEEE Press, New York, 2000.
- [4.18] Baker, R. J., Li, H. W., Boyce, P. E., „CMOS. Circuit Design, Layout, and Simulation“, IEEE Press, New York, 1998.
- [4.19] De Graaff, H. C., Klaassen, F. M., „Compact Transistor Modelling for Circuit Design“, Springer-Verlag, Wien, New York, 1990.

Kapitel 5

- [5.1] Schröder, D., „Elektrische Antriebe 3. Leistungselektronische Bauelemente“, Springer-Verlag, Berlin, Heidelberg, New York, 1996.
- [5.2] Jayant Baliga, B., „Modern Power Devices“, John Wiley & Sons, New York, Chichester, 1987.
- [5.3] Gerlach, W., „Thyristoren“, Springer-Verlag, Berlin, Heidelberg, New York, 1979.
- [5.4] Paul, R., „Elektronische Halbleiterbauelemente“, B. G. Teubner, Stuttgart, Heidelberg, 1986.
- [5.5] Müller, R., „Bauelemente der Halbleiter-Elektronik“, Springer-Verlag, Berlin, Heidelberg, New York, 1991.
- [5.6] Schrenk, H., „Bipolare Transistoren“, Springer-Verlag, Berlin, Heidelberg, New York, 1978.
- [5.7] Reisch, M., „Elektronische Bauelemente. Funktion, Grundschaltungen, Modellierung mit SPICE“, Springer-Verlag, Berlin, Heidelberg, New York, 1998.
- [5.8] Anke, D., „Leistungselektronik“, Oldenbourg Verlag, München, Wien, 2000.
- [5.9] Hagmann, G., „Leistungselektronik. Grundlagen und Anwendung“, Aula-Verlag, Wiesbaden, 1993.
- [5.10] Brosch., P. F., Landrath, J., Wehberg, J., „Leistungselektronik. Kompakte Grundlagen und Anwendung“, Vieweg Verlag, Braunschweig, Wiesbaden, 2000.
- [5.11] Sze, S. M., „Physics of Semiconductor Devices“, J. Wiley, New York, 1982.

Kapitel 6

- [6.1] Tränkler, H.-R., Obermeier, E., „Sensortechnik. Handbuch für Praxis und Wissenschaft“, Springer-Verlag, Berlin, Heidelberg, New York, 1998.
- [6.2] Heywang, W., „Sensorik“, Springer-Verlag, Berlin, Heidelberg, New York, 1988.
- [6.3] Schaumburg, H., „Sensoren“, B. G. Teubner, Stuttgart, 1992.
- [6.4] Reisch, M., „Elektronische Bauelemente. Funktion, Grundschaltungen, Modellierung mit SPICE“, Springer-Verlag, Berlin, Heidelberg, New York, 1998.
- [6.5] Müller, R., „Bauelemente der Halbleiter-Elektronik“, Springer-Verlag, Berlin, Heidelberg, New York, 1991.
- [6.6] Tietze, U., Schenk, C., „Halbleiter-Schaltungstechnik“, Springer-Verlag, Berlin, Heidelberg, New York, 1993.

- [6.7] Bleicher, M., „Halbleiter-Optoelektronik“, Hüthig Verlag, Heidelberg, 1986.
- [6.8] Infineon Technologies (Hrsg.), „Halbleiter. Technische Erläuterungen, Technologien und Kenndaten“, Publicis Corporate Publishing, Erlangen, 2004.
- [6.9] Tille, T., „Air-Quality-Sensorik in der Fahrzeugklimatisierung“, in Schlenz (Hrsg.), „PKW-Klimatisierung III“, Expert Verlag, Renningen, 2003.

Kapitel 7

- [7.1] Itoh, K., „VLSI Memory Chip Design“, Springer-Verlag, Berlin, Heidelberg, New York, 2001.
- [7.2] Cappelletti, P., Golla, C., Olivo, P., Zanoni, E., „Flash Memories“, Kluwer Academic Publishers, Boston, Dordrecht, London, 1999.
- [7.3] Scott, J. F., „Ferroelectric Memories“, Springer-Verlag, Berlin, Heidelberg, New York, 2000.
- [7.4] Reisch, M., „Elektronische Bauelemente. Funktion, Grundschaltungen, Modellierung mit SPICE“, Springer-Verlag, Berlin, Heidelberg, New York, 1998.
- [7.5] Hodges, D. A., Jackson, H. G., „Analysis and Design of Digital Integrated Circuits“, McGraw Hill, New York, 1988.
- [7.6] DeMassa, T. A., Ciccone, Z., „Digital Integrated Circuits“, John Wiley & Sons, New York, Chichester, 1996.
- [7.7] Bogart, T. F., „Introduction to Digital Circuits“, Glencoe, McGraw Hill, Westerville, 1992.
- [7.8] Vai, M. M., „VLSI Design“. CRC Press, Boca Raton, London, 2001.

Sachverzeichnis

- 1-Bit-Halbaddierer 202
1-Bit-Volladdierer 203
1T/1C-Speicherzelle 211
1-Transistor-Zelle 214
4-Phasen-Takt 205
4-Transistor-Zelle 214
6-Transistor-Zelle 213
- Abfallzeit 53, 96, 133, 161, 162
Abklingzeit 162
Abkommunitierungssteilheit 54, 161
Abschnürung 118
Absolutdruckmessung 195
Absorption 62
Absorptionsgrad 178
Addierer 202
Adresse 213
Adressierung 213
Adsorption 187
Adsorptionsplatz 188
Akkumulation 107, 111, 113, 128
Aktiver Bereich 84, 118
Akzeptor 8, 10
 ionisierter 8, 183
 neutraler 8
Akzeptordotierungsdichte 8, 109
Akzeptorniveau 9, 10
Analog-Digital-Wandler 145, 172
Analoverstärker 150
AND-Array 216, 217
Anode 32, 154
Anoden-Katoden-Spannung 155, 156, 158, 159, 160, 161, 162
Anodenstrom 156, 157, 158, 159, 160, 161, 162
Anreicherungstyp 115, 120
Anstiegszeit 133, 154
Antifuse 207, 219
Äquipotentiallinie 180
Arbeitsgerade 84
Arbeitspunkt 50, 156
- Arithmetik-Logik-Einheit 216
ASIC 220
Atomkern 3
Atomverschiebung 193
Ätztechnik 197
Augenempfindlichkeit, absolute 68
Augenempfindlichkeit, relative 64, 68
Ausbreitungswiderstand, radialer 174
Ausgangskennlinie 84, 138, 166
Ausgangskennlinienfeld 84, 119
Ausgangsleitwert 92, 129, 130, 138
Ausgangsverstärkerstufe 142
Ausgangswiderstand 92, 98, 99, 100, 101, 103, 138, 139, 140, 142
Ausgangswiderstand, differentieller 138
Ausschaltstromverstärkung 162
Ausschaltverhalten 52, 96, 154, 161, 162, 167
Aussteuerbereich 140
Austrittsarbeit 59
Auswahlsignal 202
Auswahltransistor 210, 211, 212, 213, 214, 215
Auswerteschaltung 180
- Bahngebiet 53
Bahnwiderstand 44, 50, 52, 56, 62, 65, 84, 149, 152, 165, 171
Bandabstand 9, 22, 23, 47, 65, 66, 69, 173, 180
Bändermodell 8, 9, 10, 35, 55, 58, 105, 106, 107, 116, 182, 188
 Akzeptorniveau 9
 Bänderbesetzung 11
 Bandverbiegung 10
 Donatorband 15
 Donatorniveau 9
 Energie-Impuls (E-k)-Raum 11
 Ferminiveau 13
 Leitungsband 8
 Störband 15

- Valenzband 8
Bänderüberlappung 58
Bandkante 10
Bandlücke 46
Bandverbiegung 10, 35, 105, 109, 110, 188
Basis 77
Basisaufweitung 150
Basis-Emitter-Spannung 78, 79, 80, 81, 82, 83, 84, 89, 90, 91, 92, 98, 100, 172, 173
Basis-Implantation 89
Basiskapazität 90, 93, 95
Basisladung 90, 95, 96
Basispotential 100
Basissschaltung 82, 92, 95
Basisstrom 82, 83, 95
Basisstromänderung 92
Basistransitzeit 81, 167
Basistransportfaktor 92
Basisweite 82, 88, 95, 149
Basisweitenmodulation 85, 148, 158
Basiszone 79
BCCD 186
Befehlsdekoder 216
Beschleunigungsmessung 196
Beschleunigungssensor 196
Besetzungswahrscheinlichkeit 12, 13, 14
Betriebsspannung 51, 100, 141, 154, 171, 189, 205, 213, 220
Beweglichkeit 17, 133
 Elektronen 17
 Löcher 17
Beweglichkeit, effektive 124
Beweglichkeitsreduktion 123, 124
Bewegung, thermische 18
Bias-Strom 141
BiCMOS-Technologie 145
Bidirektionaler Schalter 137
Bildwandler 186
Bindungsenergie 4
Bindungsmodell 3
Bipolartransistor 77, 172
Bistabile Kippschaltung 213
Bit 206
Bitleitung 206, 210, 211, 212, 213, 214, 215
Blockierbetrieb 155
Blockierfähigkeit 161
Blockerkennlinie 156
Blockierspannung 154, 170
Blockierzustand 148
Body-Faktor 121, 122
Boltzmann-Konstante 6, 14, 19
Boltzmann-Näherung 14
Boolesche Funktion 202
Boolesche Gleichung 217
Brechungsindex 68
Brückenschaltung 194, 196
Brummspannung 72, 73
Bulk 105
Bulkanschluss 115
Bulkspannung 105
Cache-Speicher 214
CCD-Element 182
CHE-Flash-Zelle 211
CHEMFET 190
Chemosensor 187
Chipfläche 197
CMOS 132, 197
CMOS-Gatter 198
CMOS-Inverter 171
CMOS-Logik 197, 198, 199, 200, 201, 202, 203, 204
 dynamische 204
 statische 200
CMOS-Logikgatter 197
CMOS-Technologie 145, 213
Coulomb-Streuung 123
CPLD 218
Darlington-Schaltung 97, 98, 150
Datenleitung 212
Deep-Depletion 183
Degradation 126
Dehnung, DMS 193
Dehnungsmessstreifen 193, 194, 195
Delon-Schaltung 73
Depletion 108, 109, 111, 113, 128, 188
Depletionkapazität 111, 112
Depletion-Ladung 183
Depletion-Näherung 36
Depletion-Schicht 108
Depletion-Typ 115
Desorption 189
Diamantstruktur 2
Dielektrikum 215
Dielektrizitätskonstante 211
Differenzdruck 195
Differenzspannungsverstärkung 141

-
- Differenzverstärker 102, 103, 140, 141, 214
 Differenzverstärkung 103, 140, 142, 143, 144
 Diffusion 24, 26
 Diffusionsgebiet 38, 52, 176, 196
 Diffusionskapazität 48, 49, 91, 93
 Diffusionskonstante 19, 25, 81, 85, 90
 Elektronen 19
 Löcher 19
 Diffusionsladung 48, 49, 90, 167
 Diffusionslänge 25, 26, 39, 49, 52, 77, 147, 154
 Elektronen 25
 Löcher 25
 Diffusionsmodell 41, 43, 62
 Diffusionsspannung 34, 35, 43, 55, 155
 Diffusionsstreifen 178
 Diffusionsstrom 18, 33, 38, 41, 81
 Diffusionsstromdichte 19, 26, 27, 40
 Elektronen 19
 Löcher 19
 Digitalsignal 172
 Diode 32, 172, 207
 Diodenfläche 40
 Diodenkennlinie 42, 43, 44, 50, 59, 62, 166
 Diodenspannung 51, 52, 53, 58
 Diodenstrom 42, 43, 52, 53, 54, 58, 147
 Disjunktiver Term 216
 Display 71
 DMOSFET 150
 DMS 193
 Donator 7, 10
 ionisierter 7
 neutraler 7
 ortsfester 59
 Donatorband 15
 Donatordotierungsdichte 34
 Donorniveau 9, 10
 Doppelheterostruktur 70
 Doppelimplantation 150, 151
 Dotieratom 175
 Dotierstoffatom 18
 Ionisierung 18
 Dotiertechnik 197
 Dotierung 2, 7, 15, 48, 148
 Nettodorierung 8
 Störstellenatome 2
 Dotierungsänderung 190
 Dotierungsdichte 8, 18, 30, 32, 45, 56, 58, 81, 175, 176, 180, 190
 Dotierungsprofil 32, 78, 146, 149, 154
 Drain 115
 Draingebiet 127
 Drainpotential 138
 Drain-Sättigungsspannung 138
 Drain-Sättigungsstrom 117, 135
 Drain-Schaltung 130
 Drain-Source-Sättigungsspannung 117
 Drain-Source-Spannung 116, 118, 122, 124, 126, 129, 137, 139, 153, 154, 209
 Drain-Source-Widerstand 129
 Drainstrom 116, 117, 118, 119, 121, 122, 123, 128, 138, 151, 153, 154, 191, 192, 209, 210
 DRAM 214
 Drehzahlregler 162
 Dreiphasentakt 184
 Driftgeschwindigkeit 17, 27, 29, 30
 Elektronen 17
 Löcher 17
 Driftstrom 16, 21, 22, 27, 34, 41
 Driftstromdichte 16, 17, 30, 31
 Driftzone 151
 Druckmessung 195, 196
 Drucksensor 195, 196
 Dunkelstrom 62, 186
 Dünnschichttechnik 190
 Durchbiegung 195
 Durchbruch zweiter Art 87, 150
 Durchbruchfeldstärke 125
 Durchbruchsmechanismen 44, 86, 124
 Durchbruchsspannung 45, 46, 55, 57, 60, 146, 148, 150, 157, 166
 Durchbruchsspannung 45
 Durchflussmessung 196
 Durchlassbetrieb 53, 149, 151, 152, 155, 156, 157, 164, 165
 Durchlassspannung 153
 Durchlasswiderstand 148, 168, 169
 Durchschaltzeit 159
 Dynamisches Latch 204
 Dynamisches Schieberegister 204
 Early-Effekt 85
 Early-Spannung 86, 101
 EEPROM 209, 218
 Effektive Masse 11, 12
 Eigenabsorption 68

- Eigenerwärmung 46, 176, 180
Eigenleitung 3, 4, 5, 6, 18, 176, 180
Eigenleitungsichte 8, 107
Eigenleitungs-Ferminiveau 13, 107
Eigenleitungsniveau 41, 109
Eigenleitungsträgerdichte 5, 6, 7, 173
Eingangskapazität 205
Eingangskennlinie 84, 85
Eingangsleitwert 91
Eingangspegel 132
Eingangsspannungsdifferenz 102, 142
Eingangswiderstand 98, 132, 141
Einphasen-Brückenschaltung 170
Einsatzspannung 109, 110, 111, 120,
 121, 122, 124, 126, 127, 139, 152,
 153, 167, 191, 192, 207, 208, 210,
 212
Einsatzspannungimplantation 111, 206
Einschaltüberspannung 53
Einschaltverhalten 51, 94, 153, 158,
 166
Einschaltverzögerung 153
Einschaltzeit 154, 159
Einstein-Beziehung 19
Einwegaufgleichrichter 71
Elektrisches Feld
 Feldrichtung 16
Elektronenaffinität 59
Elektronenaustrittsarbeit 59
Elektronenbeweglichkeit 130
Elektronendichte 109
Elektronenhülle 3
Elektronenkonzentration 7
Elektronenladung 59
Elektronenzustand 9
Elektron-Loch-Paar 4, 5, 7, 9, 22, 27,
 44, 62, 64, 67, 171, 176, 183
Elektrostatische Entladung 125
Elementarladung 19
Emissionsfläche 69
Emitter 77, 164
Emitterfläche 81, 90
Emitterschaltung 82, 83, 87, 92, 95, 100
Emitterstrom 83
Emitterstromsprung 94
Empfindlichkeit 177, 178, 189, 190,
 193, 194, 195, 196
Energie, kinetische 27
Energieband 8
Energie-Impuls (E - k)-Raum 67
Energieniveau 8, 9, 11, 13, 58, 79
Enhancement-Typ 115
Entkopplung 201
Entladezeit 72
Epitaxie 88
EPROM 207
ESD 125, 145
Exemplarstreuung 173
Extrinsische Kapazität 128
Feld, elektrisches 10, 18, 27, 30, 34, 44,
 46
Feld, magnetisches 29, 178
Feldeffekt 105
Feldeffekttransistor, chemosensitiver
 190
Feldeffekttransistor, gassensensitiver 191
Feldeffekttransistor, ionensensitiver 192
Feldeffekttransistor, MOS- 105
Feldstärke
 Durchbruch- 125
 kritische 125
 laterale 125
 vertikale 125
Feldstärke, elektrische 17, 27, 28, 30,
 34, 37, 54, 149, 168
Feldstärke, maximale 34, 37, 45, 55,
 146, 148, 149
Feldstärkeintegral 148
Feldstärkeverlauf 37
Ferminiveau 13, 14, 35, 58, 79, 106,
 109, 188, 189
 Eigenleitungs-Ferminiveau 13
Fermipotential 107, 122
Fermiteilchen 13
Fermiverteilung 13
Ferroelektrisches Dielektrikum 211
Festkörper 60
Festwertspeicher 206
Fitting-Parameter 59, 135
Flachbandspannung 106, 107, 110, 113,
 121, 191
Flächeneffizienz 211
Flächenladung 107, 119
Flächenladungsdichte 107, 108, 110,
 111, 118, 183
Flash-EPROM 211, 218
Flipflop 213
Floating-Gate 208, 209, 210, 211
Floating-Gate-Speichertransistor 208,
 209, 210

- Flusspolung 38, 42, 43, 44, 50, 56, 58, 79, 84, 91, 147, 155
 Flussspannung 42, 49, 51, 76
 Flussstrom 71, 146
 FN-Flash-Zelle 211
 Fotodiode 62, 182
 Fotogate 182
 Fotolack 48, 89, 127
 Fotostrom 63
 Fotothyristor 161
 Fototransistor 182
 Fowler-Nordheim-Tunneln 210, 211
 FPGA 218
 Antifuse-basierter 219
 SRAM-basierter 219
 FRAM 211
 Freilaufdiode 152, 170
 Freiwerdezeit 161
 Frequenz 49, 51, 92
 Frequenzgang 103, 141, 143
 Frequenzkonverter 170
 Frequenzmodulation 62, 74
 Frequenzteiler 145
 Fuse 207, 217
- GaAs 3, 66
 Eigenleitungsträgerdichte 5
 Mindestenergie 5
 Temperaturkoeffizient 180
 GaAs_xP_{1-x} 69
 GAL 217
 Galvanische Trennung 161
 Gas
 oxidierendes 187
 reduzierendes 187
 GASFET 191
 Gaskonzentration 189, 191
 Gate 105, 115, 154, 164
 Gate-Array 220
 Gate-Bulk-Kapazität 111, 112, 113
 Gate-Bulk-Kapazität, minimale 112
 Gate-Bulk-Spannung 113, 184
 Gate-Drain-Kapazität 153
 Gate-Durchbruch 125
 Gateelektrode 127, 192
 Gatekapazität 153, 204
 Gateladung 183
 Gatemetallisierung 192
 Gateoxid 115, 125, 215
 Gateoxidkapazität 107, 108, 110, 111, 112, 121
 Gate-Schaltung 130
 Gate-Source-Spannung 116, 117, 119, 122, 129, 151, 153
 Gatespannung 105, 109, 127
 Gatestrom 153, 155, 156, 162, 163
 Gatestromimpuls 162
 Gegenkopplung 99
 Generation 5, 9, 10, 22, 23, 26, 40, 62, 81
 optische 9, 22, 62, 184
 thermische 6, 9, 23, 62, 176, 186
 Generationsrate 6, 23, 27
 zusätzliche 23, 24
 Germanium 2
 Eigenleitungsträgerdichte 5
 Mindestenergie 5
 Geschwindigkeit
 mittlere thermische 6
 Geschwindigkeit, mittlere 28, 118
 Geschwindigkeitssättigung 17, 124
 Gitterbindung 27
 Gitterstreuung 176
 thermische 18
 Gitterstruktur 3, 5, 7, 17
 Streuung an Störstellen 16
 Wechselwirkung 17
 Gleichrichter 43, 59
 Gleichrichterdiode 155
 Gleichspannungswandler 154
 Gleichstrom 155
 Gleichtaktunterdrückung 103, 140, 141, 143
 Gleichtaktverstärkung 103, 143
 Grabenstruktur 169
 Grenzfläche 105, 109, 122, 183
 Grenzflächenladung 60, 183
 Grenzflächenpotential 108, 113
 Grenzflächenrauhigkeit 123
 Grenzflächenstreuungen 123
 Grenzflächen-Temperatursensor 172, 196
 Grenzfrequenz 56, 57, 62, 93, 94
 Grenzschicht 191
 Grenzwellenlänge 64
 Grenzwinkel 68
 Großsignalbahnwiderstand 50
 Großsignalersatzschaltbild 90, 128
 Großsignalersatzschaltbild, Diode 50
 Großsignalverhalten 50
 GTO-Thyristor 161
 Guard-Ringe 171

- Halbaddierer 202
Halbleiter 1
 Bandabstand 9, 22
 Bändermodell 8
 Besetzungswahrscheinlichkeit 13
 Bestrahlung mit Licht 22
 binärer 2
 Brechungsindex 68
 direkter 11
 Dotierung 2, 7
 eigenleitender 4, 9, 10
 einkristalliner 3
 Elektronen-Leitfähigkeit 2
 Elementarzelle 3
 Elementhalbleiter 2
 entarteter 16
 Halbleitermaterialien 1
 homogener 5, 22
 indirekter 11
 intrinsischer 4
 Ladungsträgertransport 16
 Löcher-Leitfähigkeit 2
 Majoritätsträger 2
 Minoritätsträger 2
 n-dotierter 32
 n-Typ-Halbleiter 2
 p-dotierter 32
 p-Typ-Halbleiter 2
 quaternärer 2
 reiner 3, 4
 Störstellenhalbleiter 7, 9, 15
 Streumechanismen 17
 ternärer 2
 Verbindungshalbleiter 2
Halbleitermaterialien 1
Halbleiteroberfläche 188
Halbleitersensor 172
Hallbeweglichkeit 31, 179
Halleffekt 29, 31, 178
Hallektrode 180
Hallgenerator 178, 179, 180, 181, 182
 rechteckiger 179
 symmetrisch kreuzförmiger 179
Hallkonstante 31
Hallsensor 181
Hallspannung 30, 178, 180, 181
Hallwinkel 31
Haltespannung 157
Haltestrom 157, 158, 160
Hauptspeicher 215
Hauptstromkennlinie 157, 159
Hazard 205
Heizelement 187
Heizleistung 187
Helligkeit 68, 69
Helligkeitsregler 162
HF-Oszillator 59
Hot-Electron-Injektion 209, 210, 211
Hysterese 212
IC 197
IGBT 164, 171
Impedanzwandler 132
Impulssteuerung 163
Induktion, magnetische 29, 181
Induktivität 53, 170
Infrarot-Strahlung 177
Injektion, schwache 23, 26, 39
Injektion, starke 43, 147
Innenwiderstand 71, 103, 140, 178, 179, 180
Integrierte Schaltung 197
Integrierter Logik-Baustein 216
Integrierter Speicherbaustein 205
Intensitätsverteilung 182
Interdigitalstruktur 188
Intrinsische Kapazität 128
Inversdiode 152
Inversion 109, 128
 schwache 109
 starke 110
Inversionskanal 116, 151, 167, 171, 208, 209
Inversionsladung 112
Inversionspunkt 109, 112
Inversionsschicht 112, 116, 183
Inverter 132, 197, 200, 202, 204, 213, 214
Invertereingang 204
Invertierender Eingang 142
Ionenimplantation 127
Ionenkonzentration 192
Ionisation 8, 10
 vollständige 8, 10, 176
Ionisationskoeffizient 28
IR-Strahlung 177
ISFET 192
Isolator 1
i-Zone 54
Kalibrierung 177
Kanalladung 127

- Kanallänge 118, 122, 123
 Kanallängenmodulation 123, 130
 Kanallängenmodulationsfaktor 123, 138
 Kanalspannung 117
 Kanalverkürzung 117
 Kanalwiderstand 137
 Kapazitätsdiode 61, 74, 75
 Kapazitäts-Spannungs-Charakteristik 62, 112, 113, 128, 191
 Kapazitiver Drucksensor 196
 Kaskadenschaltung 74
 Kaskodestufe 132, 138, 139
 Katode 32, 154
 Katodenfläche 159
 Katodenzerstäubung 48
 k-Faktor 193, 196
 Kinetische Energie 9
 Kippspannung 156, 157, 158, 171
 Kippstrom 156, 157, 158
 Kleinsignalersatzschaltbild 50, 91, 129, 131
 Kleinsignalleitwert 50
 Kleinsignalbahnwiderstand 50
 Kleinsignalstromverstärkung 92, 93, 98, 99, 101, 102
 Kleinsignalverhalten 50
 Kollektor 77, 164
 Kollektorbahnwiderstand 149
 Kollektor-Basis-Spannung 78, 79, 80
 Kollektor-Basis-Sperrspannung 79
 Kollektor-Emitter-Sättigungsspannung 85
 Kollektor-Emitter-Spannung 84, 90, 148, 149, 150, 166, 167, 168
 Kollektorschaltung 82, 83
 Kollektorstrom 79, 81, 82, 83, 84, 85, 88, 90, 95, 97, 100, 102, 149, 150, 164, 165, 166, 167, 173
 Kollektowiderstand 102
 Kommutierungssteilheit 53
 Komparator 181
 Komplexgatter 200
 Konjunktiver Term 216
 Konstantstromquelle 99, 100, 102, 103, 141, 176
 Kontaktloch 127, 174, 175, 178
 Kontaktlochdurchmesser 175
 Kontaktspannung 34, 106, 155
 Kontaktwiderstand 60
 Kontinuitätsgleichung 25, 26
 Konvektionsstromdichte 19, 20, 39
 Elektronen 19
 Gesamt- 20
 Löcher 19
 Konzentrationsänderung 189
 Kristaldefekt 70
 Kristallgitter 4, 7, 27, 44
 Kurzkanaltransistor 124
 Kurzschlussstrom 65, 67
 Ladungsänderung 48
 Ladungsdichte 20, 27, 183
 Ladungsneutralität 48
 Ladungsträger 23, 44, 105
 bewegliche 16
 Energiezustände 8
 freie 27
 Generation 5
 Rekombination 5
 Zustandsdichte 12
 Ladungsträgerbeweglichkeit 19, 56, 62, 120, 122, 172, 174, 176, 179
 Ladungsträgerbilanz 26
 Ladungsträgerdichte 4, 6, 8, 12, 14, 15, 18, 19, 20, 26, 27, 33, 41, 109, 147, 164, 176
 Elektronen 4
 intrinsische 5
 Löcher 4
 Ladungsträgerdiffusion 20, 81
 Ladungsträgerlebensdauer 41, 54, 56, 147, 186
 Ladungsträgertransport 16, 60, 172
 Ladungsträgerverteilung 11, 12, 14, 15, 27, 147, 159
 Ladungstransfer 186
 Ladungstransport 184, 185
 Ladungsverlust 204, 209
 Ladungsverschiebung 182
 Ladungsverteilung 105, 106
 Längenänderung 48, 193
 Langkanalnäherung 119
 Last
 aktive 141, 142
 induktive 152, 160
 ohmsche 157, 167
 Lastkapazität 131, 135
 Lastwiderstand 53, 84, 163, 189
 Latch 205
 Latch-up-Effekt 170
 Lawinendurchbruch 45, 64, 86, 125, 148

- Lawineneffekt 27, 44, 57, 86, 125, 158
Lawinen-Fotodiode 64
Lawinenmultiplikation 64
Lebensdauer
 Minoritätsträger- 23
Leckstrom 204, 215
LED 67
Leerlauf 30
Leerlaufempfindlichkeit 30
Leerlaufspannung 65, 66
Leistungsaufnahme 176
Leistungsbauelement 146
Leistungs bipolartransistor 148
Leistungsdichte 179
Leistungsdiode 146
Leistungsgleichrichter 154
Leistungs-MOSFET 150
Leiter 1
Leitfähigkeit 1, 105, 178, 187, 188
 elektrische 1
 Elektronen-Leitfähigkeit 2
 Löcher-Leitfähigkeit 2
 spezifische 17, 18, 30, 31
Leitfähigkeitsänderung 189
Leitungsband 9, 10, 15, 46, 58, 62
 unteren Leitungsbandkante 9
Leitungbandelektron 58
Leitungselektron 4, 7
Leitungstyp 20
Lesestrom 209
Leseverstärker 215
Leuchtdichte 70
Leuchtwirkungsgrad 68, 69
Lichtintensitätsmessung 64
Lichtleistung 62, 65, 66, 68, 182
Lichtleiter 161
Lichtwellenlänge 22
Lithographie 48, 89, 127, 197, 206
Loch 4
Löcherbeweglichkeit 130
Löcherdichte 109
Löcherdiffusionsstromdichte 25
Löcherinjektion 90, 164
Logikblock 218
Logikschaltungen 197, 200
Lorentzkraft 29
Löschtorgang 209
Luftgüteüberwachung 190
Lumineszenzdiode 67, 68, 69, 161

Mäanderstruktur 188

Magnetfeld 29
Magnitosensor 178
Majoritätsträger 2, 60, 79
Majoritätsträgeranreicherung 107
Majoritätsträgerdichte 7, 48, 112
Majoritätsträgerinjektion 20, 21, 22
Majoritätsträgerkonzentration 48
Majoritätsträgerverarmung 108
Majoritätsträgerverschiebung 21
Massepotential 82
Matrixstruktur 206
Mehrbit-Addierer 203
Mehrfach-Array-Struktur 218
Mengenmessung 196
Messsignal 172
Metall-Halbleiter-Kontakt 16
Metallisierung 48, 127, 174
Metalloxid 188
Metalloxid-Gassensor 187, 189
Metalloxidhalbleiter 187
Mikroprozessor 216
Mikrostrukturierte Siliziumtechnologie 187
Mikrostrukturierter kapazitiver Drucksensor 196
Miller-Effekt 153
Miller-Kapazität 132, 153
Miller-Plateau 153
Mindestenergie 5, 9
Minoritätsträger 2, 26
Minoritätsträgerdichte 7, 24, 25, 38, 41, 42, 43, 48, 49, 52, 80, 85, 90, 92, 147, 156, 158, 159, 161, 173
Minoritätsträgerdiffusion 42
Minoritätsträgerinjektion 20, 22, 24
Minoritätsträgerkonzentration 81, 95
Minoritätsträgerladung 52, 53
Minoritätsträgerladungsdichte 52
Minoritätsträgerlebensdauer 23, 26, 49
Minoritätsträgerüberschussdichte 24, 52
Minoritätsträgerverteilung 51, 79, 94
Mischkristalle 69
Mixed-Signal-Schaltung 145
Modulation 74
Molekül 188
MOSFET 115, 181, 198, 206
 n-Kanal- 115
 p-Kanal- 115
MOS-Struktur 105, 108, 115, 182, 183, 184, 185, 190, 191
Motorsteuerung 150

-
- Mulde 184
 Muldentiefe 182, 184
 Multiplexer 200
- n^{++} -Poly-Silizium 127
 n^{++} -Silizium-Gate 105
 Nachladezeit 72
 Nachweisgrenze 190
 NAND-Architektur 211
 NAND-Funktion 198
 NAND-Gatter 198
 n^- -Driftzone 151, 164, 165, 168
 Nettodotierung 8
 Nettorekombinationsrate 23, 25, 26
 n-Typ-Halbleiter 23
 p-Typ-Halbleiter 23
 Neutrale Basis 80
 Neutrale Basisweite 79, 81, 85, 90, 92
 Neutrale Weite 147
 Neutralität 21, 22
 Neutralitätsbedingung 183
 Nicht-invertierender Eingang 142
 n-Kanal-MOSFET 115
 n-Leitung 10
 NOR-Funktion 199
 NOR-Gatter 199
 npn-Bipolartransistor 77
 n-Typ-Halbleiter 13
 Nullpunkt drift 142
 Nullpunktstabilität 141
 Nullspannung
 induktive 181
 ohmsche 180
 n-Wanne 127
 n-Wannen-Prozess 127
- Oberflächeneffekt 190
 Oberflächenkanal 186
 Oberflächenladungsdichte 110, 122
 Oberflächenpotential 184, 186
 Offsetkompensation 180
 Offset-Spannung 142, 143
 Ohmsches Gesetz, lokales 17
 On-Widerstand 151, 152, 164
 Open-Loop-Verstärkung 142
 Operationsverstärker 141, 143, 180, 181
 Optischer Sensor 182
 Optokoppler 71
 OR-Array 216, 217, 218
 Oszillator 145
 Oxidation, thermische 48, 89, 127
- Oxiddicke 208
 Oxidkapazität 127, 208
 Oxidschicht 107
- p^{++} -Poly-Silizium 113
 $p^+ s_n n^+$ -Leistungsdiode 146
 Packungsdichte 197
 PAL 217
 Palladium 190
 Parabelnäherung 117, 123, 130
 Parallelschwingkreis 74, 75
 Parameterstreuungen 102, 214
 Parasitärer npn-Bipolartransistor 152, 165
 Pass-Transistor 219, 220
 Peakstrom 58
 Periodensystem 2
 Permanentmagnet 181
 Phase 142, 185
 Phasenanschnittsteuerung 163
 Phasenanschnittwinkel 163
 Phasenreserve 144
 Phasenversatz 205
 Phasenverschiebung 143
 Photon 22, 62, 68
 Energie 22
 Photonenenergie 64
 Photonenstrahlung 62
 Photonenstrom 68
 pH-Wert 192
 Piezoresistiven Sensor 194
 Piezoresistiver Drucksensor 195
 Piezoresistiver Effekt 193
 Pinch-off 117, 131
 Pinch-off-Bedingung 124
 Pinch-off-Punkt 117, 123
 pin-Diode 54
 Pixel-Rauschen 186
 p-Kanal-MOSFET 115
 Planartechnik 197
 Plancksches Wirkungsquantum 11, 22
 PLD 216
 p-Leitung 10
 pn-Diode 44
 pnp-Bipolartransistor 77
 pn-Übergang 32, 33, 34, 35, 36, 37, 38,
 39, 41, 42, 44, 45, 46, 47, 48, 58, 61,
 70, 77, 155, 156, 173
 abrupter 32
 Poisson-Gleichung 27
 Polarisation 212

- Polygonzug 221
Poly-Si-Membran 196
Positionserkennung 181
Potential, elektrisches 34
Potentialbarriere 60
Potentialdifferenz 106, 108, 110
Potentialmulde 182, 183, 184, 186
Potentialverlauf 35, 36, 108, 155
Programmieraufwand 218
Programmiereffizienz 209
Programmierspannung 209
Programmierstrom 209
Programmierstromimpuls 207, 217
Programmierung 206, 211, 218
PROM 207, 217
Prozessüberwachung 190
psn-Leistungsdiode 146
PT-IGBT 168
p-Typ-Halbleiter 13
Pull-down-Pfad 198
Pull-up-Pfad 198
Punch-Through-Effekt 87, 126, 148, 149
Punch-Through-IGBT 168
PZT 211
- Quantenausbeute 68
Quantenwirkungsgrad, externer 63, 68, 70
Quasigleichgewicht 39
Querempfindlichkeit 190
- RAM 213
 dynamischer 214
 statischer 213
Raumladung 37, 59, 60
Raumladungsdichte 54, 55, 79
Raumladungszone 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 48, 51, 54, 55, 61, 62, 67, 78, 81, 85, 89, 106, 108, 110, 112, 116, 121, 146, 148, 149, 155, 156, 182, 183
Raumladungszonenweite 124
Raumwinkel 69
Rauschen
 thermisches 178
RC-Zeitkonstante 95
Referenzspannung 214
Reflexionsverlust 66
Refresh 215
Refresh-Zeit 215
- Rekombination 5, 10, 22, 24, 26, 52, 81, 95, 96, 167
 nichtstrahlende 67
 strahlende 67
Rekombinations-/Generationsmodell 40
Rekombinationskoeffizient 6, 23
Rekombinationsrate 6, 23
Rekombinationsstrom 147
Rekombinationsverlust 81
Rekombinationszentrum 64
Relaxation 20, 22, 24, 48
 dielektrische 21
Relaxationszeit, dielektrische 21, 22
Resonanzfrequenz 62, 74, 75
Restore 213, 215
ROM 206
Rückinjektion 85
Rückkopplung 144
Rückschlagspannung 161
Rückwärtsbetrieb 152
Rückwärtsdiode 166
Rückwärtserholzeit 53, 54, 161
- Sättigung 42
Sättigungsbereich 118, 119, 129, 130, 134, 153
Sättigungsspannung 139
Sättigungsstrom 85, 117
Sättigungsstromdichte 151
Sauerstoff 188
SBT 211
Schaltdiode 54, 60
Schaltfrequenz 162, 170
Schaltgeschwindigkeit 98, 150, 197, 218
Schalthysterese 181
Schaltleistung 146
Schaltmatrix 218
Schaltpunkt 153
Schaltschwelle 133
Schaltverhalten, Bipolartransistor 94
Schaltverhalten, CMOS-Inverter 134
Schaltverhalten, Diode 51
Schaltverhalten, Leistungs-MOSFET 152
Schaltzeit 60, 95, 96, 150, 153, 161, 199
Schaltzustand 198
Scheitelwert 74
Schichtdicke 190
Schichttechnik 197

-
- Schieberegister 185, 204
Schleusenspannung 165
Schottky-Diode 59, 60
Schottky-Kontakt 59, 60
Schottkysche Parabelnäherung 36
Schreib-Lese-Speicher 213
Schutzdiode 125
Schwache Inversion 109, 112
Schwellenbereich 165
Schwellspannung 44, 58, 60, 71
Schwellwert 181
Schwingkreis 75
Schwingneigung 144
Seebeck-Koeffizient 29, 178
Selektivität 190
Sensor 172
Sensorelement 172
Sensoroberfläche 190
Sensorwiderstand 189
Serienwiderstand 137
Sigma-Delta-Wandler 196
Signallaufzeit 218
Silizium 2
 Beweglichkeit der Ladungsträger 17
 Diamantgitter 3
 effektive Zustandsdichte 12
 Eigenleitungsträgerdichte 5
 Gitterstruktur 3, 5, 7
 Mindestenergie 5
 Silizium-Atom 3
 Valenzelektron 3, 4
Silizium-Planartechnik 47, 78
Silizium-Substrat 47
Sintertechnik 187
SiO₂-Isolatorschicht 127
Slew 205
Solarzelle 64, 65, 66, 67, 182
Sonnenpektrum 66
Source 115
Sourcegebiet 127
Sourceleitung 211
Source-Schaltung 130
Sourcespannung 117
Spacer 89
Spannung
 Anoden-Katoden- 155
 Basis-Emitter- 78
 Betriebs- 100, 141
 Brumm- 72
 Bulk- 105
 Diffusions- 34
 Dioden- 51
 Drain-Source- 116
 Early- 86
 Einsatz- 109
 Flachband- 106
 Fluss- 42, 49
 Gate- 105
 Gate-Source- 116
 Hall- 30
 Kanal- 117
 Kipp- 156
 Kollektor-Basis- 78
 Kontakt- 34
 Offset- 142
 Schwell- 44
 Substrat- 121
 Temperaturspannung 19
 Thermo- 28, 29
 Zener- 57
 Spannungsbegrenzung 75
 Spannungsrampe 135
 Spannungsspitzen 159
 Spannungssprung 135
 Spannungssteuerung 94
 Spannungsteiler 100
 Spannungsverstärkung 83, 130, 131
 Spannungsvervielfacher 73
 Speicher 205
 anwenderprogrammierbarer 205
 festprogrammierter 205
 flüchtiger 213
 nichtflüchtiger 205
 programmierbarer/elektrisch
 löschbarer 205
 Speicherinhalt 206, 213
 Speicherkapazität 211, 212, 213, 214,
 215
 Speichertiefe 206
 Speichertransistor 211
 Speicherzeit 53, 54, 96, 161, 162, 186
 Speicherzelle 206, 207, 209, 210, 211,
 212, 213, 214, 215, 219
 Spektralbereich 69
 Spektrale Bestrahlungsstärke 66
 Spektrale Empfindlichkeit 63
 Spektrale Intensitätsverteilung 68
 Spektraler Hellempfindlichkeitsgrad 64
 Sperrbereich 52, 84
 Sperrbetrieb 53, 155, 156
 Sperrfähigkeit 146, 148
 Sperrpolung 41, 43, 50, 53, 56, 79, 148

- Sperrrichtung 115
Sperrsättigungsstrom 40, 41, 42, 43, 56, 146, 147, 173
Sperrsicht 50, 58, 108, 111
Sperrschichtkapazität 48, 49, 53, 56, 61, 62, 75, 89, 91, 128
Sperrschichttemperatur 87
Sperrspannung 53, 61, 71, 146, 148, 150, 151, 153, 162, 166, 168, 169, 171
Sperrspannung, maximale 54
Sperrstrom 41, 44, 46, 53, 62, 79, 82, 159, 166
Sperrverlust 146
Sperrverzögerung 53
Spiegelverhältnis 138, 141
Spreading-Resistance Si-
Temperatursensor 174, 176
Sputtern 48, 187
SRAM 213
Stabilität 144, 154
Stack-Speicherkapazität 215
Standardzellen-IC 220
Starke Inversion 110, 112, 113
Stauchung, DMS 193
Steilheit 91, 93, 95, 98, 103, 129, 131, 140, 142
Steuergateleitung 211
Steuerspannung 116
Steuerstrom 152, 176
Störanfälligkeit 204
Störsicherheit 152
Störstellenhalbleiter 6, 7, 15
Störstellenleitung 3, 7
Stoßionisation 27, 44
Strahlungsfluss 178
Streufeld 185
Strombegrenzung 46, 87
Stromergiebigkeit 150, 164
Stromfluss 79
Stromkommutierung 53
Strommessung, kontaktlose 182
Stromquelle 50, 90, 92, 99, 103, 128, 140
Stromsättigung 119
Stromspiegel 99, 100, 101, 137, 138, 141, 174
Stromsteilheit 53
Stromsteuerung 94
Stromverstärkung 83, 131, 148, 149, 150, 158, 165, 167
Substrat 117
Substrateinfluss 91, 121
Substratoberfläche 108
Substratspannung 121
Substratsteuerung 121, 131
s-Zone 146
Tail-Ladung 167, 168
Tail-Strom 168
Taktansteuerung 186
Taktflanke 205
Taktfrequenz 204
Takteitung 185
Taktschema 136, 204, 205
Taktsignal 136, 205
Teilchenstrom 18
Temperatur 187
Raumtemperatur 1, 7
Temperaturabhängigkeit 173
Temperaturbereich 173
Temperaturgradient 173
Temperaturkoeffizient 45, 46, 57, 154, 180, 194
Temperaturkompensation 100, 196
Temperaturleitwert 176, 178
Temperaturmessung 173, 174
Temperatursensor 172, 173, 177
Temperaturspannung 19, 178
Thermische Gitterschwingung 4
Thermischer Durchbruch 47, 87, 154
Thermodiffusion 28
Thermodynamisches Gleichgewicht 5, 6, 8, 10, 14, 23, 29, 33, 49, 55, 79, 85, 90, 106, 113, 116, 155
Störung 20, 21, 22
Thermoeffekt 28, 177
Thermoelement 177, 178
Thermokoeffizient 178
differentieller 28, 178
Thermokraft 178
Thermopile 177
Thermospannung 28, 29
Thyristor 154, 155, 171
Thyristorkennlinie 156
Tiefpass 144
Totalreflexion 68
Trägerfrequenz 74
Transfergate 136
Transferrauschen 186
Transfersättigungsstrom 85
Transferstrom 77, 81, 82

-
- Transistor 172
 Bipolartransistor 77
 Leistungsbipolartransistor 148
 Leistungs-MOSFET 151
 MOSFET 105
 Transistorfläche 138
 Transistorweite 139
 Transitfrequenz 94, 131, 144
 Transmission 68
 Transmissiongate 136, 137, 200, 204, 205
 Treiber 145, 216
 Trench-IGBT 169
 Trench-Speicherkapazität 215
 Triodenbereich 118, 119, 129, 134, 153
 Tristate-Treiber 201, 202
 Tunneldiode 58, 59
 Tunnelfenster 209, 210
 Tunnelmechanismus 46
 Tunnelstrom 210
 Tunnelwahrscheinlichkeit 46

 Überkopfzünden 158
 Überprogrammierung 209
 Überschussdichte 26
 Überschussladungsträger 52, 162
 Überschussminoritätsträgerdichte 23, 24, 25, 26
 Überschussträgerdichte 20, 22, 39
 Übertragungskennlinie 119, 120, 123, 133, 142, 143, 152

 Unterschwellbereich 134
 Unterschwellstrom 122, 123, 124
 UV-Licht 207, 209

 Vakuumenergie 60
 Vakuumniveau 106
 Valenzband 9, 10, 46, 62
 obere Valenzbandkante 9
 Valenzelektron 3, 7, 8
 Verarmung 108
 Verarmungsschicht 183
 Verarmungstyp 115, 120
 Verarmungszone 59, 124
 Verdrahtungskanal 218, 220
 Verformungskörper 193
 Verlustleistung 46, 53, 76, 135, 146
 Verstärkerschaltung 130
 Verstärkung 95, 140
 Verzögerungselement 185, 186

 Verzögerungszeit 135
 Videoausgang 187
 Villard-Schaltung 74
 kaskadierte 74
 Voll-kundenspezifischer IC 221
 Volumeneffekt 190
 Vorwärtsbetrieb 79, 84, 91
 Vorwärtserholzeit 53
 Vorwärtsstrom 82
 Vorwiderstand 76, 153

 Wahrheitstabellen 198, 202, 203
 Wannenkapazität 128
 Wärmesenke 87
 Wasserstoff 191
 Wasserstoffdipol 191
 Weglänge, mittlere freie 45
 Weicheisenrotationskörper 181
 Wellenlänge 22, 64, 69
 Wellenzahl 11
 Welle-Teilchen-Dualismus 8
 Wendepunkt 177
 Wheatstone-Brückenschaltung 194, 195
 Widerstand 1, 174
 Ausgangs- 92
 differentieller 58, 59
 Drain-Source- 129
 Durchlass- 137
 Eingangs- 132
 Innen- 71
 Kontakt- 60
 Last- 53, 66
 negativer differentieller 59
 ohmscher 174
 On- 151
 Parallel- 65, 176
 regelbarer 57
 Sensor- 189
 Serien- 137
 spannungsgesteuerter 116
 spezifischer 1, 17, 174, 175, 176, 196
 thermischer 47, 150
 Vor- 76
 Zener- 58
 Wilson-Stromspiegel 101, 139, 140
 Winkeländerung 181
 Winkelmessung 181
 Wirkungsgrad 66
 Wortleitung 206, 207, 210, 212, 214, 215

- XNOR-Gatter 201
XOR-Gatter 200, 202
- Z-Diode 57, 75
Zeitkonstante
 thermische 178
Zeitkonstante für Schalter 137
Zenerdurchbruch 46
Zenereffekt 46, 57
Zenerspannung 57, 75
Zenerwiderstand 58
Zinndioxid 188
- Zugriff, wahlfreier 213
Zugriffszeit 214
Zünden, Thyristor 155
Zündimpuls 158, 159
Zündstrom 156, 157, 158, 159
Zündstromimpuls 155, 161
Zündverzugszeit 159
Zündzeit 159
Zündzeitpunkt 155
Zustandsdichte 12, 14, 16
 effektive 12, 14
Zwei-Phasen-Takt 204
Zweiwegbrückenschaltung 72