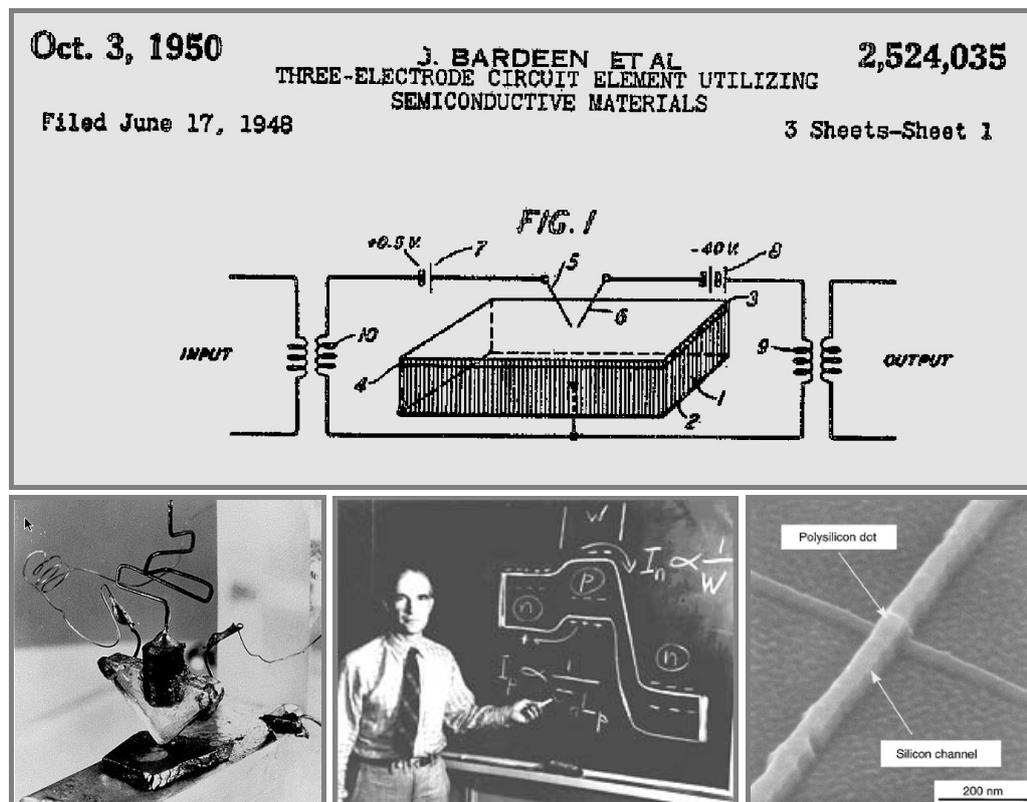


Dioden und Transistoren

baulemente.tex KB 20061118



Für analoge und digitale Anwendungen der modernen Elektronik sind Dioden und Transistoren die wichtigsten Basisbauelemente – die kleinsten Einheiten. Dioden zur Gleichrichtung, als Spannungsreferenzen, zur Lichterzeugung, als Detektoren von Licht oder Teilchen, Transistoren als Verstärker oder Schalter bestimmen unsere Technik – manchmal als Einzelbauelemente für spezielle Anwendungen, meist aber hoch oder höchst integriert in ICs.

Ziel dieses Teilprojekts des Elektronik-Praktikums ist es, typische Eigenschaften und einfache Anwendungen von Dioden und Transistoren kennen zu lernen. Experimente und Simulationen sollen sich dabei komplementär ergänzen.

Inhaltsverzeichnis

1 Grundlagen: Halbleiterphysik	3
2 Grundlagen: Dioden	3
2.1 Abrupter p-n-Übergang im Gleichgewicht	3
2.2 Spannung am p-n-Übergang	4
2.3 Strom am p-n-Übergang	5
2.4 Diodentypen	6
3 Grundlagen: Transistoren	11
3.1 Bipolare Transistoren	11
3.2 Feldeffekt-Transistoren (FETs)	12
3.3 Sperrschicht-FETs (JFETs)	13
3.4 MOSFETs	13
3.5 Schaltsymbole	14
3.6 Bauformen	14
3.7 Miniaturisierung	16
3.8 Kennlinien	16
3.9 Vier-Quadranten-Kennlinienfeld	17
3.10 Transistor als Vierpol	19
3.11 Arbeitspunkt	21
3.12 Grenzwerte	22
3.13 Grundsaltungen	23
3.14 Kennlinien von Feldeffekttransistoren	24
Literatur und Links	28

Die Titelseite zeigt einen Ausschnitt aus dem ersten für den Transistor erteilten Patent (oben), den Nachbau des ersten Transistor im Museum der Bell Laboratories (links), William Shockley, einen der Erfinder, beim Erklären der Bandstruktur eines Transistors (Mitte) und eine Rastertunnelaufnahme des ersten in Silizium realisierten Ein-Elektronen-Transistors (rechts).

1 Grundlagen: Halbleiterphysik

Überlegen Sie, ob sie die folgenden Begriffe aus der Halbleiterphysik physikalisch richtig einordnen können. Falls nicht — ein Lehrbuch zur Halbleiterphysik hilft weiter.

- Bändermodell im Orts- und k-Raum,
- Valenzband, Leitungsband, Bandlücke E_g ,
- Elektronen – Defektelektronen/Löcher,
- Eigenleitung – Störstellenleitung ('intrinsic', 'extrinsic'),
- Akzeptoren, Donatoren,
- p-Typ-, n-Typ-Halbleiter,
- effektive Masse,
- Zustandsdichte, Fermiverteilung, Ladungsträgerdichte,
- Fermienergie E_F ,
- Ladungsträgerbeweglichkeit, Leitfähigkeit,
- Majoritätsladungsträger, Minoritätsladungsträger,
- Ladungsträgeranreicherung, -verarmung,
- bewegliche und ortsfeste Ladungen, Raumladung,
- p-n-Übergang,
- Erzeugung und Rekombination von Ladungsträgern.

2 Grundlagen: Dioden

Ladungen, Potenziale und Felder in einer Diode kann man sich an einem Idealisierung veranschaulichen, dem abrupten p-n-Übergang. Gekennzeichnet ist diese Idealisierung durch eine eindimensionale Geometrie und konstante Dotierungen im p- und n-Bereich.

2.1 Abrupter p-n-Übergang im Gleichgewicht

Im idealisierten Modell des p-n-Übergangs wechselt in der Mitte eines Halbleiterstabs die Dotierung nach Art einer Stufenfunktion von p-Typ (Akzeptoren) nach n-Typ (Donatoren). Direkt am Übergang ändert sich zunächst auch die Dichte der (beweglichen) Ladungsträger abrupt, in der Skizze Abbildung 1 (a) sind die Energieverhältnisse – Bandkanten und Fermienergien – in einem eindimensionalen Bandschema im Ortsraum dargestellt. Der starke Gradient führt zu einem Diffusionsstrom. Dadurch wird der Dichtegradient abgebaut, gleichzeitig ein Gegenfeld

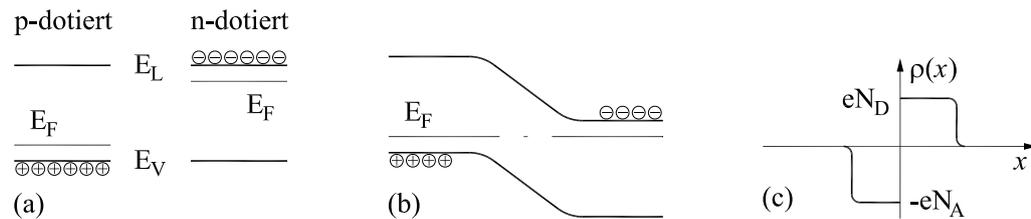


Abbildung 1: Idealisierter p-n-Übergang. (a): Eindimensionale Bandstruktur (Ortsraum) für p- und n-Dotierung (E_V : Valenzbandkante, E_L : Leitungsbandkante, E_F : Fermienergie). (b): Ladungsträgerverarmung und Bandverbiegung am Übergang nach Kontakt. (c): Ladungsdichte ρ am Übergang (N_A und N_D : Dotierungskonzentrationen).

durch die raumfesten Ladungen aufgebaut. Im stationären Zustand kompensieren sich Diffusions- und Feldstrom, die Fermi-Energie im Gesamtsystem ist konstant. Im Übergangsbereich bildet sich eine ladungsträgerverarmte Zone, die Raumladungszone - das Bandschema Abbildung 1 (b) veranschaulicht die resultierende 'Bandverbiegung'. Für eine idealisierte Berechnung nimmt man an, dass die Raumladungszone völlig frei von beweglichen Ladungsträgern ist. Die Ladungsdichte im Übergangsbereich hat dann etwa den in Abbildung 1 (c) dargestellten Verlauf (Akzeptor- und Donorkonzentrationen N_A und N_D).

Zur Feld- und Potenzialberechnung wird die (eindimensionale) Poisson-Gleichung ein- bzw. zweimal integriert. Die zu beiden Seiten des Übergangs konstante Ladungsdichte führt zu jeweils linearem Feldverlauf (Dreieck, Maximum an der Grenze zwischen p- und n-Gebiet) und parabolischem Potenzialverlauf. Neben- und Randbedingungen: Das E-Feld außerhalb des Übergangsbereichs ist Null, die gesamte Potenzialdifferenz ist gleich der ursprünglichen Differenz der beiden Fermi-Energien auf der p- und n-Seite.

2.2 Spannung am p-n-Übergang

Liegt eine äußere Spannung am p-n-Übergang, ändern sich die Randbedingungen bei der Integration der Poisson-Gleichung. Spannung in Durchlassrichtung verringert die Potenzialstufe am Übergang, die Raumladungszonen auf beiden Seiten des Übergangs werden schmaler, das Feld im Übergangsbereich wird kleiner. Bei einer Spannung in Sperrrichtung ist es umgekehrt, die Potenzialstufe vergrößert sich, die Raumladungszonen werden breiter, das Feld im Übergangsbereich wird größer. Abbildung 2 veranschaulicht dies.

Durch die Änderung der Breite der Raumladungszonen ändert sich auch die Kapazität des p-n-Übergangs; man kann den p-n-Übergang als Plattenkondensator auffassen, dessen 'Plattenabstand' durch die Raumladungszonen definiert ist. Der Effekt wird ausgenutzt, um spannungsgesteuerte Kapazitäten zu bauen (Kapazitätsdioden). Detektordioden (Photodioden) werden schneller (höheres Feld, geringere Kapazität), wenn man eine Sperrspannung geeigneter Größe anlegt.

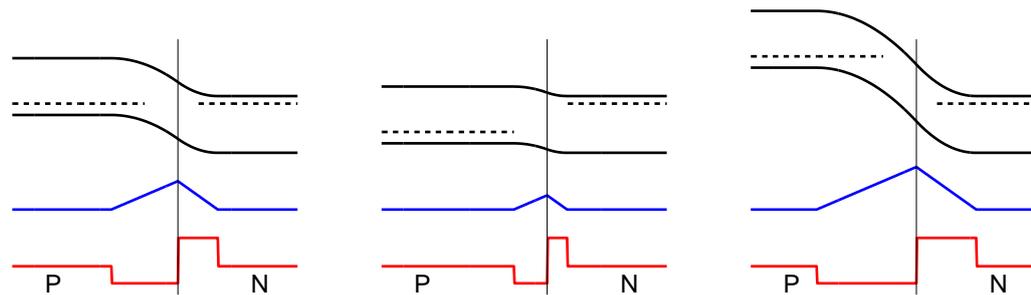


Abbildung 2: Spannung am p-n-Übergang: links ohne anliegende Spannung, Mitte: Spannung in Durchlassrichtung, rechts: Spannung in Sperr-Richtung. Die senkrechte dünne Linie ist die als scharf angenommene Grenze zwischen p- und n-Gebiet. Die untere Kurve stellt jeweils die stufenfunktionsartige Raumladungsdichte dar, die sich in einem solch einfachen Modell einstellt, die mittlere den dreieckigen Verlauf des Betrags des elektrischen Feldes, oben der parabolische Potenzialverlauf für die Elektronen, der Verlauf von Valenz- und Leitungsbandkante (Bandstruktur im Ortsraum) im Übergangsgebiet. Die Lage der Fermi-Energien ist gestrichelt eingezeichnet, die Differenz zwischen den Fermi-Energien auf p- und n-Seite entspricht der am p-n-Übergang anliegenden Spannung, $\Delta E_F = e \cdot U$. Deutlich werden die Unterschiede in der Breite der Raumladungszonen und im Maximalwert des elektrischen Feldes für die drei Fälle.

2.3 Strom am p-n-Übergang

Die Dichte der Minoritätsladungsträger am Rand des Übergangsgebiets ändert sich exponentiell mit der angelegten Spannung. Ist die p-Seite gegenüber der n-Seite negativ gepolt (Sperr-Richtung), nimmt die Dichte der Minoritätsladungsträger ab; im Übergangsgebiet thermisch generierte Ladungsträger (das System strebt eine Fermi-Verteilung an) führen zu einem kleinen, bei genügend hoher Spannung konstanten Sperrstrom. Bei entgegengesetzter Polung (Durchlassrichtung) steigt die Dichte der Minoritätsladungsträger exponentiell mit der Spannung an, der Durchlassstrom steigt proportional dazu an.

Die Abhängigkeit des Stroms I von der angelegten Spannung U wird recht gut durch die erstmals von Shockley [1] 1951 aufgestellte Formel

$$I = I_S \left(\exp \frac{eU}{kT} - 1 \right) \quad (1)$$

beschrieben. Der Vorfaktor I_S (Sättigungssperrstrom) wird durch Materialparameter, Geometrie der Diode und Temperatur bestimmt.

Bei hohen Sperrspannungen weicht der Kennlinienverlauf vom obigen Ausdruck ab: Durch die dann hohen Feldstärken im Übergangsgebiet werden die Ladungsträger dort derart beschleunigt, dass sie sekundäre Elektron-Loch-Paare anregen können (Lawineneffekt). Die Folge ist ein mit der Sperrspannung drastisch zunehmender Sperrstrom, der bei zu

großer thermischer Belastung schließlich zu einer Zerstörung des p-n-Übergangs führt.

Im Durchlassbereich ist der Kennlinienverlauf dadurch modifiziert, dass der ohmsche Widerstand von p- und n-Bereich im allgemeinen nicht vernachlässigt werden kann. Der Strom wird dadurch kleiner als im durch Gleichung 1 gegebenen Idealfall.

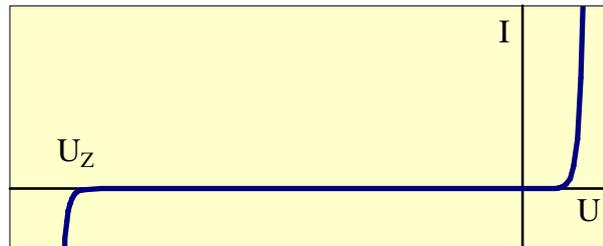


Abbildung 3: Strom-Spannungs-Kennlinie einer Diode (U_Z : Zener-, Durchbruchspannung).

Weiterführende Stichworte: Feldstrom, Diffusionsstrom, Quasiferminiveaus, Tunnel-, Zener-, Lawineneffekt.

2.4 Diodentypen

Dioden werden in der Elektronik und Elektrik sowie der physikalischen und elektrischen Messtechnik für die unterschiedlichsten Anwendungen eingesetzt. Dafür wurden viele Spezialbauformen entwickelt; einige davon sind im folgenden aufgeführt.

2.4.1 Schaltodiode, Universaldiode



Geringer Durchlasswiderstand, sehr kleiner Sperrstrom, Material: Silizium (überwiegend) oder Germanium (heute selten, praktisch ausgestorben), Größenordnung der Strombelastbarkeit: 100 mA, der Durchbruchspannung: 100 V, der Kapazität: einige pF, der Durchlassspannung: 0.7 V (Si) bzw. 0.3 V (Ge), preisgünstig.

2.4.2 Gleichrichterdiode



Ähnliche Eigenschaften wie bei der Schottkydiode, meist höhere Durchbruchspannung (1000 V), Strombelastbarkeit der Anwendung entsprechend (1 . . . 1000 A), die höhere Belastbarkeit wird unter anderem durch eine größere Fläche erreicht, dadurch auch deutlich höhere Kapazität.

2.4.3 Schottky-Diode



Kein p-n-Überhang, sondern ein Metall-Halbleiter-Übergang (Schottky-Kontakt). Zum Strom tragen nur Majoritätsladungsträger bei, dadurch schaltet die Schottky-Diode sehr schnell. Die Durchlassspannung ist etwa halb so groß wie bei Silizium-Dioden. Sie wird dort verwendet, wo es auf eine dieser beiden Eigenschaften ankommt: Hochfrequenztechnik, Gleichrichter mit geringen Verlusten – hohem Wirkungsgrad.

2.4.4 Zenerdiode



Diode mit gut definierter Durchbruchspannung, die so gebaut ist, dass sie dauerhaft im Bereich des steilen Anstiegs der Sperrkennlinie betrieben werden kann, ohne dadurch zerstört zu werden. Sie wird zur Spannungsstabilisierung und Spannungsbegrenzung verwendet (kleiner differentieller Widerstand im steilen Anstieg der Kennlinie). Typische Nennspannungen zwischen etwa 2 und 200 V.

2.4.5 Kapazitätsdiode



Die Kapazität einer Diode ist abhängig von der anliegenden Sperrspannung. Diodenbauformen, die auf diesen Effekt hin optimiert sind (Kapazitätsdioden, *Varaktor*-Dioden), werden als variable Kapazität in Schwingkreisen verwendet (spannungsgesteuerte Frequenzabstimmung in jeglicher Art von Empfängern).

2.4.6 Tunnel diode



Diode mit extrem hoch dotierten p- und n-Gebieten, damit sehr dünnem Übergangsgebiet. Durch die hohe Dotierung liegt die Fermi-Energie im Valenz- bzw. Leitungsband. Wird eine kleine Spannung in Durchlass- oder Sperr-Richtung an die Tunnel diode angelegt, verschieben sich die Bänder so, dass sich besetzte und unbesetzte Zustände auf gleichem Potenzial gegenüberstehen, ein Tunnelstrom fließt durch das dünne Übergangsgebiet. Dieser wächst in Sperr-Richtung mit zunehmender Sperrspannung stetig an, ebenso für kleine Spannungen in Durchlassrichtung. Dagegen stehen sich bei etwas höheren Spannungen in Durchlassrichtung besetzte und verbotene Zustände (Bandlücke) gegenüber. Dies führt in einem gewissen Spannungsbereich zu einer Abnahme des Stroms mit steigender Spannung (negativer differentieller Widerstand), bis bei weiterer Spannungserhöhung 'normales' Durchlassverhalten einsetzt.

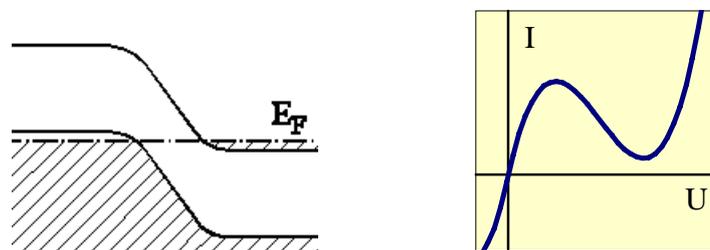


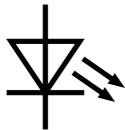
Abbildung 4: Links das idealisierte Bandschema im p-n-Übergangsgebiet einer Tunnel diode, gezeichnet ist der örtliche Verlauf der Valenzband- und Leitungsbandkanten. Elektronenzustände unterhalb der Fermienergie sind besetzt (schraffiert). Rechts die Strom-Spannungs-Kennlinie einer Tunnel diode, das lokale Strommaximum liegt bei einer Durchlassspannung von etwa 100 mV.

2.4.7 Backward-Diode



Tunneldiode, bei der die Fermienergie an der Valenz- bzw. Leitungsbandkante liegt. Dadurch erhält man die Sperrkennlinie einer Tunneldiode und (annähernd) die Durchlasskennlinie eines normalen p-n-Übergangs. Die Backward-Diode ist somit für kleine Spannungen in Sperrrichtung durchlässig, in Durchlassrichtung sperrend (Name!). Sie kann zur Gleichrichtung kleiner Spannungen verwendet werden.

2.4.8 Lumineszenzdiode, Laserdiode



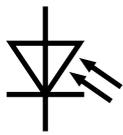
Bei allen Dioden ist im Durchlassbetrieb die Ladungsträgerdichte im Übergangsbereich gegenüber der Gleichgewichtssituation deutlich erhöht. Dieser Überschuss wird durch die Rekombination von Elektron-Loch-Paaren abgebaut. Wird die Energie des Elektron-Loch-Paars ($\approx E_g$) als Lichtquant frei, spricht man von strahlender Rekombination – Lumineszenz. Grundsätzlich lumineszieren alle Dioden, effizient allerdings nur bei einem Halbleitermaterial mit *direkter* Bandstruktur, bei Halbleitern mit *indirekter* Bandstruktur (Si, Ge) überwiegt die nichtstrahlende Rekombination (Phononenerzeugung).

Die Energie der Lichtquanten entspricht in etwa der Bandlücke (E_g) des verwendeten Halbleitermaterials. Gebräuchliche Materialien sind 'maßgeschneiderte' Mischungen aus III-V-Verbindungshalbleitern wie GaAs, GaAlAs, InGaAsP u. ä. für Infrarot, Rot, Orange, Gelb, Grün. Effiziente violette und blaue Lumineszenzdioden sind erst seit kurzem auf dem Markt (GaN, InGaN). In Kombination mit geeigneten lumineszierenden Substanzen (*Lumineszenz-Konverter*), die in ähnlicher Weise auch in Leuchtstoff- und Energiesparlampen verwendet werden, lassen sich damit 'weiße' Lumineszenzdioden herstellen. Wegen ihrer Zuverlässigkeit und ihres hohen Wirkungsgrads steht diesen ein immenser Boom bevor.

Laserdioden sind spezielle Bauformen von Lumineszenzdioden. Durch geeignete Maßnahmen (Heteroübergänge, Strukturierung, Resonator) erreicht man hohe Ladungsträgerkonzentrationen in einem begrenzten Volumen (*carrier confinement*), Lichtführung in einem Wellenleiter (*light confinement*) und Rückkopplung des Lichts (Reflexion an Spaltflächen, 'distributed Bragg reflector' o. ä.), dadurch eine sehr niedrige Laserschwelle (Durchlasstrom, ab dem Lasertätigkeit einsetzt). Durch die mit dem vergleichsweise kleinen Betriebsstrom verbundene geringe thermische Belastung ist bei modernen Halbleiterlasern ein sehr zuverlässiger Betrieb möglich. Dies ist insbesondere bei den Anwendungen in der optischen Nachrichtentechnik ein wichtiger Aspekt.

Die Bandlückenenergie (E_g) im Halbleiter ist abhängig von äußeren Parametern wie Temperatur und Druck. Durch Änderung dieser Parameter lässt sich die Emissionwellenlänge von Laserdioden in gewissen Grenzen verschieben. Dies wird für Anwendungen in der optischen Spektroskopie ausgenutzt (Infrarotspektroskopie von Schadstoffmolekülen in der Umweltüberwachung mit Blei-Chalkogenid-Lasern).

2.4.9 Photodiode



Beleuchtet man einen p-n-Übergang mit Licht geeigneter Wellenlänge, werden Elektron-Loch-Paare erzeugt, die durch das Feld im Übergangsbereich getrennt werden, (Sperr-)Strom fließt. Photodioden sind für diesen Effekt – hohen Quantenwirkungsgrad – optimierte Dioden. Sie werden in der optischen Nachrichtentechnik, in vielen Bereichen der Elektronik (Potenzialtrennung, Lichtschranken etc.) und in vielen Bauformen und aus verschiedenen Materialien (unterschiedliche langwellige Grenze wegen der materialabhängigen Mindestanregungsenergie) als Detektoren in der physikalischen Meßtechnik (optische Spektroskopie) eingesetzt.

Speziell für eine gute Leistungsausbeute bei Beleuchtung mit Sonnenlicht optimiert sind Solarzellen. Auch hier ist ein hoher Quantenwirkungsgrad entscheidend, andere für die Anwendung von Photodioden wichtige Parameter wie Linearität und Geschwindigkeit sind allerdings nur von untergeordneter Bedeutung.

Lawinenphotodioden (avalanche photo diodes, APDs) sind Photodioden, die für eine kontrollierte Lawinenvervielfachung konstruiert sind. Sie werden mit einer genau festgelegten, hohen Sperrspannung betrieben, entweder in einem Bereich konstanter Ladungsträgervervielfachung (100 . . . 1000) oder – ähnlich einem Geiger-Müller-Zählrohr – im Durchbruchbereich. Für den sichtbaren und nahen infraroten Spektralbereich sind APDs derzeit die empfindlichsten Detektoren (Einzelphotonennachweis mit einer Quanteneffizienz von etwa 80 %).

Ähnlich wie Photodioden arbeiten die Halbleiterdetektoren in der kernphysikalischen Messtechnik. Zum spektroskopischen Nachweis von Alpha- und Beta-Teilchen sowie Röntgenquanten verwendet man p-i-n-Strukturen aus Silizium, für Gamma-Quanten solche aus Germanium (höhere Elektronendichte und damit höherer Absorptionskoeffizient). Sie werden in Sperr-Richtung mit hoher Sperrspannung betrieben, das dadurch über der i-Schicht aufgebaute E-Feld trennt die vom absorbierten Teilchen erzeugten Elektron-Loch-Paare, die abfließende Ladung ist in guter Näherung proportional zur absorbierten Energie.

2.4.10 P-I-N-Diode



Ein in den p-n-Übergang eingeschobener undotierter (*intrinsic*) Bereich kann dazu dienen, die Sperrspannung zu erhöhen oder das aktive Volumen von Detektordioden zu vergrößern. Im Durchlassbetrieb zeigen solche Strukturen eine ausgeprägte Stromabhängigkeit des Widerstands. Spezielle, darauf hin optimierte p-i-n-Dioden werden als Abschwächer oder Schalter in der Hochfrequenztechnik verwendet.

2.4.11 DIAC (Diode AC switch)



Eine symmetrische Struktur aus zwei gegeneinander gesetzten p-n-Übergängen mit gemeinsamer dünner Zwischenschicht. Ein p-n-Übergang

wird jeweils in Durchlass-, einer in Sperr-Richtung betrieben. Das Bauelement sperrt bis zu einer bestimmten Spannung, wird beim Überschreiten dieser Grenzspannung schlagartig niederohmig. Man verwendet DIACs, um Trigger-Impulse (Zünd-Impulse) für *Thyristoren* oder *TRIACs* zu erzeugen.

2.4.12 Gunn-Diode

Die Funktionsweise der Gunn-Diode beruht auf einem physikalisch interessanten Bandstruktur-Effekt: In Halbleitern, bei denen ein Leitungsbandminimum für Elektronen mit kleiner effektiver Masse und hoher Beweglichkeit energetisch knapp unterhalb eines Leitungsbandminimums für Elektronen mit großer effektiver Masse und kleiner Beweglichkeit liegt, treten bei hohen elektrischen Feldern Streueffekte auf, die die Leitfähigkeit verkleinern. Es werden schnelle Elektronen aus dem Minimum mit hoher Beweglichkeit in das mit kleinerer Beweglichkeit gestreut, dies führt zu einem Feldstärkebereich mit negativem differentiellen Widerstand.

Zu den Halbleitern, die diesen Gunn-Effekt zeigen, gehören GaAs und InP. Durch geeignete Konstruktion kann man zeitliche Feldstärke-, Elektronendichte- und damit Stromfluktuationen im GHz-Bereich erreichen. Gunn-Dioden werden als Mikrowellengeneratoren mit hohem Wirkungsgrad eingesetzt.

2.4.13 IMPATT-Diode

Ebenfalls zur Mikrowellenerzeugung dienen IMPATT-Strukturen (IMPact Avalanche and Transit Time), bei denen ein p-n-Übergang mit einer kurzen Laufstrecke kombiniert ist. Eine genügend große Sperrspannung veranlasst einen Lawinendurchbruch im Übergangsbereich, die Ladungslawine senkt die Feldstärke solange unter die Durchbruchfeldstärke, bis die Ladung abgeflossen ist. Dies wiederholt sich periodisch und verursacht Stromfluktuationen im GHz-Bereich.

Sowohl Gunn- als auch IMPATT-Dioden werden zur Frequenzfestlegung mit geeigneten Mikrowellenresonatoren kombiniert.

2.4.14 Temperatursensoren

Die Temperaturabhängigkeit der Durchlasskennlinie einer Diode (Shockley-Formel) lässt sich zur Temperaturmessung ausnutzen; *Bandgap-Referenz-Dioden* werden als Sensoren mit nahezu linearer Kennlinie im Temperaturbereich $-30 \dots 120^\circ\text{C}$ angeboten, für tiefe Temperaturen ($1 \dots 100\text{ K}$) geeignete Dioden weisen dagegen eine stark gekrümmte Kennlinie auf.

3 Grundlagen: Transistoren

Transistoren arbeiten je nach Bauart nach zwei völlig unterschiedlichen Grundprinzipien. *Bipolare* Transistoren sind Mehrschicht-Strukturen aus unterschiedlich dotierten Schichten, Dioden-Effekte spielen eine Rolle, durch Strom wird Strom gesteuert. *Unipolare* oder *Feldeffekt*-Transistoren dagegen können als Halbleiterwiderstände betrachtet werden, deren Leitfähigkeit durch ein von außen angelegtes Feld beeinflusst wird, durch Spannung wird der Widerstand verändert.

3.1 Bipolare Transistoren

Bipolare Transistoren bestehen aus einer Abfolge von drei abwechselnd dotierten Halbleiterschichten, N-P-N oder P-N-P, *Emitter*, *Basis* und *Kollektor*¹. Sie enthalten zwei diodenartige Übergänge, Basis-Emitter und Basis-Kollektor. Je nach Zonenabfolge spricht man von npn- oder pnp-Transistoren. Die überwiegende Zahl von (bipolaren) Transistoren sind heute npn-Transistoren aus Silizium. Andere Materialien sind nur noch für ganz wenige Spezialtransistoren relevant, so etwa GaAs oder InP wegen ihrer hohen Ladungsträgerbeweglichkeiten für extreme Hochfrequenzanwendungen.

Für den Transistoreffekt ist es notwendig, dass die beiden pn-Übergänge des Transistors räumlich direkt zusammenhängen. Das Funktionsprinzip kann man sich an Abbildung 5 für den Fall eines npn-Transistors klarmachen. Die Basis-Emitter-Diode wird in Durchlassrichtung betrieben,

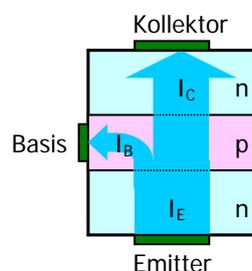


Abbildung 5: Schema eines npn-Transistors, eingezeichnet ist der Elektronenstrom vom Emitter zu Basis und Kollektor.

ben, es fließt ein Emitterstrom I_E , der wie bei jeder Diode exponentiell von der angelegten Spannung abhängt. Ein Teil der zur Basis fließenden Elektronen diffundiert zum Kollektor, eine daran angelegte, gegenüber dem Emitter positive Spannung führt zu einem Stromfluss zum Kollektoranschluss. Durch einen Basisstrom I_B kann man mithin einen Kollektorstrom I_C steuern.

Für das Verhältnis I_C/I_E hat man die Abkürzung α eingeführt, eine 'Stromverstärkung', die kleiner als 1 ist. Die interessantere Stromverstärkung B ergibt sich daraus zu

$$B = \frac{I_C}{I_B} = \frac{I_C}{I_E - I_C} = \frac{\alpha}{1 - \alpha} \quad (2)$$

¹ Die Begriffe *Emitter* und *Kollektor* entsprechen der jeweiligen Funktion, *Basis* kommt von der Bauform der ersten Transistoren (s. Titelseite).

Daneben definiert man eine differentielle Stromverstärkung β als

$$\beta = \frac{dI_C}{dI_B} \approx B \quad . \quad (3)$$

Die Stromverstärkungen B und β sollen möglichst groß sein, der Kollektorstrom soll durch einen möglichst kleinen Basisstrom gesteuert werden können. Das heißt, dass möglichst wenig vom gesamten Emittierstrom I_E für die Basis (oder andere Verluste) investiert werden soll, α soll möglichst nahe bei 1 liegen. Das kann man durch verschiedene Maßnahmen optimieren:

- Der Strom über die Basis-Emitter-Diode setzt sich aus Elektronen und Löchern zusammen. Es können aber nur die Majoritätsladungsträger des Emitters (Elektronen bei npn-, Löcher bei pnp-Transistoren) zum Kollektor diffundieren. Die Ladungsträgerinjektion vom Emitter muss daher die von der Basis weit überwiegen. Das erreicht man dadurch, dass man den Emitter erheblich höher dotiert als die Basis.
- Die Ladungsträger sollen in der Basis möglichst wenige Partner zur Rekombination finden, dies erreicht man durch eine niedrig dotierte Basisschicht.
- Die Ladungsträger sollen möglichst vollständig zum Kollektor gelangen, dies erreicht man durch eine dünne Basisschicht.
- Im Kollektor sollte nichts mehr verloren gehen, das erreicht man durch einen geeigneten geometrischen Aufbau mit großer Kollektorfläche (Abbildung 6).

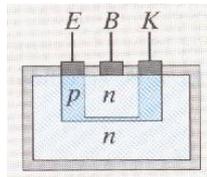


Abbildung 6: Geometrie eines npn-Transistors (das Bild ist ein Zitat aus Gerthsen Physik, 18. und 19. Auflage; überlegen Sie, was daran falsch ist).

Für die Entwicklung der (bipolaren) Transistortechnologie war es ein Glücksfall, dass die genannten Punkte miteinander kompatibel sind, das heißt, dass sie gleichzeitig optimiert werden können. Gebräuchliche Transistortypen liegen heute bei α -Werten zwischen 0.98 und 0.998, das heißt bei Stromverstärkungen B und β zwischen 50 und 500.

Hochdotierter Emitter und dünne Basis verursachen allerdings auch eine relativ geringe Sperrspannungsfestigkeit der Basis-Emitter-Diode, die Durchbruchspannung liegt meist nicht höher als etwa 5 V. Dies muss bei der Schaltungskonzeption berücksichtigt werden, wenn die Möglichkeit besteht, dass ein Transistor in diesem Bereich betrieben wird (zusätzliche Schutzdiode).

3.2 Feldeffekt-Transistoren (FETs)

Die Ladungsträgerdichte und damit die Leitfähigkeit in Halbleitern wird durch den energetischen Abstand der Fermi-Energie vom Valenz- bzw.

Leitungsband bestimmt. Feldeffekttransistoren sind Halbleiterwiderstände, in denen die Fermi-Energie durch ein von außen angelegtes elektrisches Feld verändert werden kann (durch die im Vergleich zu Metallen geringe Ladungsträgerdichte bewirkt ein äußeres Feld eine räumlich exponentiell abfallende Potenzialverschiebung an der Grenzfläche – Debye-Länge). *Source* und *Drain* sind die beiden Widerstandsanschlüsse, *Gate* der Anschluss an die Steuerelektrode. Durch die Gate-Spannung lässt sich der Widerstand zwischen *Source* und *Drain* in einem weiten Bereich verändern.

3.3 Sperrschicht-FETs (JFETs)

Bei Sperrschicht-Feldeffekttransistoren (Junction FET, JFET) wird der Widerstand durch die Veränderung der Breite der Raumladungszone eines p-n-Übergangs variiert. Ohne Gate-Spannung sind JFETs leitend, die Leitfähigkeit wird durch die – in Sperr-Richtung angelegte – Gate-Spannung verringert (selbstleitende FETs). Das Funktionsprinzip veranschaulicht Abbildung 7.

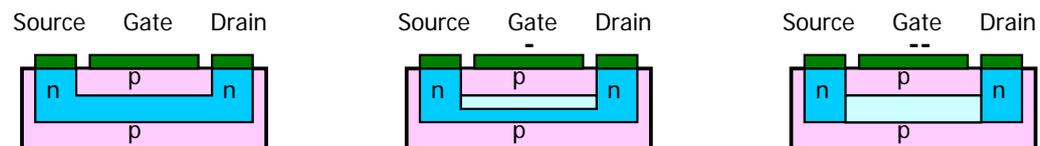


Abbildung 7: Idealisiertes Funktionsprinzip eines Sperrschicht-FETs. Links spannungslos: Leitfähiger Kanal, Mitte mit kleiner negativer Gatespannung: breitere Raumladungszone des p-n-Übergangs – geringere Leitfähigkeit, rechts mit großer Gatespannung: die Raumladungszone schnürt den Kanal ab – keine Leitfähigkeit. Der Einfluss der Drain-Spannung auf die Potenzialverhältnisse ist hier vernachlässigt (\Rightarrow trapezartige Form der Raumladungszone).

Während bei bipolaren Transistoren der p-n-Übergang an der Basis in Durchlassrichtung betrieben wird, ist bei JFETs der p-n-Übergang des Gates immer in Sperr-Richtung. Die Eingänge von Schaltungen mit JFETs können dadurch wesentlich hochohmiger sein als bei Bipolar-Transistoren.

3.4 MOSFETs

Bei MOSFETs (Metal Oxide Semiconductor FET)² ist das Gate durch eine dünne Isolatorschicht vom Kanal isoliert – kondensatorähnlich. Beim selbstsperrenden Typ (Anreicherungstyp, Enhancement Mode MOSFET) ist im spannungslosen Zustand kein leitender Kanal vorhanden, er wird durch eine Gate-Spannung und die damit verbundene Potenzialverschiebung induziert. Das Funktionsprinzip veranschaulicht Abbildung 8.

² Üblich sind auch die allgemeineren Bezeichnungen IGFET (Insulated Gate FET) und MISFET (Metal Insulator Semiconductor FET), MOSFET ist jedoch die am häufigsten verwendete (Google-Resultate: IGFET – 1050, MISFET – 609, MOSFET – 191000).

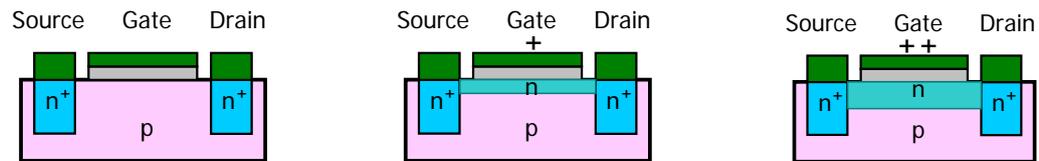


Abbildung 8: Funktionsprinzip eines selbstsperrenden MOSFETs. Links spannungslos: keine Verbindung zwischen Source und Drain, Mitte und rechts: kleinere und größere positive Gatespannung – durch die Potenzialverbiegung in der Randschicht wird n-Leitfähigkeit induziert. Der Einfluss der Drain-Spannung ist weggelassen (\Rightarrow trapezartige Form des Kanals).

Beim selbstleitenden Typ (Verarmungstyp, Depletion Mode MOSFET) verbindet im spannungslosen Zustand ein dünner Kanal in der Randschicht Source und Drain. Er ist vom gleichen Leitfähigkeitstyp wie diese Anschlüsse (im Beispiel in Abbildung 8 n-Typ). Durch die Gate-Spannung (im Beispiel negativ) wird dessen Leitfähigkeit vermindert. Das Funktionsprinzip hat somit gewisse Ähnlichkeit zum JFET.

Selbstsperrende FETs haben gegenüber selbstleitenden den Vorteil, dass die Gate-Spannung sowohl im gesperrten wie auch im leitenden Zustand zwischen Source- und Drain-Spannung liegen kann, man kann mithin geringere Betriebsspannungen verwenden. Dies ist insbesondere bei extrem hoch integrierten Chips wichtig, um die Verlustleistung gering zu halten.

PMOS, NMOS, CMOS sind gebräuchliche Abkürzungen für den Typ des MOSFETs: p-Kanal, n-Kanal und – bei integrierten Schaltungen – Complementary MOS. NMOS-Transistoren werden häufiger verwendet als PMOS wegen der höheren Beweglichkeit des relevanten Ladungsträgertyps (Elektronen im Leitungsband). Die CMOS-Technologie kombiniert auf Logik-ICs PMOS- und NMOS-Transistoren. Man spart damit weitgehend die Arbeitswiderstände an den Drain-Anschlüssen ein, die durch die komplementären MOS-Transistoren ersetzt werden. Das vermindert die Verlustleistung und erhöht außerdem die Schaltgeschwindigkeit.

3.5 Schaltsymbole

Die Schaltsymbole für Bipolar- und Feldeffekt-Transistoren sind in den Abbildungen 9 und 10 zusammengefasst. Früher war es üblich, um die Symbole Kreise zu zeichnen (ein Relikt aus der Zeit der Elektronenröhren), Sie finden das noch bei älteren Schaltbildern und/oder Leuten.

3.6 Bauformen

Praktisch alle Transistoren werden heute in *Planartechnik* hergestellt. Dabei wird eine große Siliziumscheibe (Durchmesser bis zu 30 cm) auf einer Seite bearbeitet: mit Hilfe von Photolacktechniken werden zweidimensionale – *planare* – Strukturen realisiert, verschiedene Dotierungs-

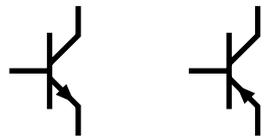


Abbildung 9: Schaltsymbole für bipolare Transistoren: links npn-, rechts pnp-Transistor. Der Pfeil gibt die Durchlassrichtung der Basis-Emitter-Diode an. Das Symbol ist der Bauform der ersten Transistoren nachempfunden (vgl. Titelseite): Die Basis als Plättchen (linker Anschluss), der Ladungsträger emittierende Emitter (Pfeil, unterer Anschluss), der dazu dicht benachbarte Kollektor (oberer Anschluss).

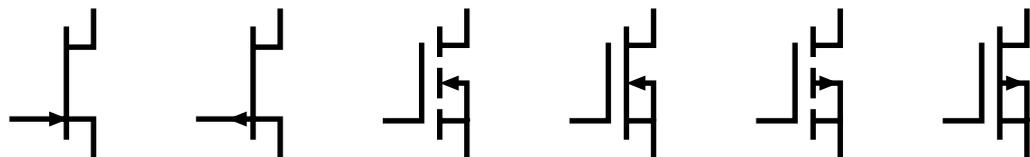


Abbildung 10: Schaltsymbole für Feldeffekt-Transistoren, von links nach rechts: n-Kanal-Sperrschicht-FET, p-Kanal-Sperrschicht-FET, selbstsperrender n-Kanal-MOSFET, selbstleitender n-Kanal-MOSFET, selbstsperrender p-Kanal-MOSFET, selbstleitender p-Kanal-MOSFET. Links jeweils der Gate-Anschluss, unten Source, oben Drain, Mitte Bulk (meist mit dem Source-Anschluss verbunden, es sind aber auch Bauformen mit getrenntem Bulk-Anschluss üblich). Die Symbole widerspiegeln die Funktion: Pfeil als Durchlassrichtung der jeweiligen Diode, isoliertes Gate bei den MOSFETs, unterbrochene Verbindung bei den selbstsperrenden Typen.

substanzen werden in mehreren Schritten aufgedampft und eindiffundiert, abschließend wird die Oberfläche durch Oxidation zu SiO_2 passiviert, im letzten Prozessschritt werden Metallkontakte und -verbindungen aufgedampft. Auf diese Weise werden sehr viele gleichartige Einzelbauteile (Chips) gleichzeitig – dadurch sehr kostengünstig – hergestellt. Am Ende des Fertigungsprozesses wird die große Scheibe in die Einzelchips zerteilt, diese werden dann einzeln getestet und konfektioniert. Abbildung 11 zeigt einen auf diese Weise hergestellten Bipolar-Transistor.

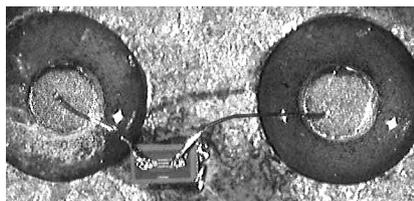


Abbildung 11: Bipolar-Transistor für Kleinsignal-Anwendungen (maximale Verlustleistung etwa 500 mW). Auf dem Chip ist die fingerartige Struktur der Emitter- und Basis-Kontakte zu erkennen. Diese sind durch dünne Golddrähte mit den beiden Anschlusspins verbunden (realer Abstand zwischen etwa 1.5 mm). Der Kollektor wird über die Chip-Unterseite kontaktiert.

Die Gehäusebauformen und Gehäusegrößen von Transistoren werden im wesentlichen durch die abzuführende Wärmeleistung bestimmt. Eine Übersicht über gängige Bauformen finden Sie in Büchern zur Halbleiterelektronik (z. B. in [2]). Als Ergänzung dazu hier zwei Bauformen, die die Bandbreite der Anwendungen in etwa erahnen lassen. Abbildung 12 zeigt einen Kleinsignal-Transistor für die aktuelle SMD³-Technik und daneben einen Transistor für Hochleistungsanwendungen (z. B. in Motorsteuerungen).

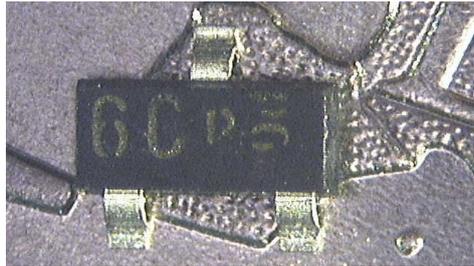


Abbildung 12: Baugrößen unterschiedlicher Transistoren. Links ein Kleinsignal-Transistor für Oberflächenmontage (Gehäuselänge 2.9 mm), zum Vergleich die Iberische Halbinsel auf dem 1-Euro-Stück. Rechts ein Hochleistungstransistor (maximale Kollektor-Emitter-Spannung 1200 V, maximaler Kollektorstrom 200 A).

3.7 Miniaturisierung

Wegen ihres einfachen Aufbaus und ihrer geringen Verlustleistung lassen sich MOS-, insbesondere CMOS-Strukturen sehr klein bauen und äußerst dicht packen. Die mit dieser Miniaturisierung stetig zunehmende Packungsdichte ist die Basis für die immense Zunahme der Leistungsfähigkeit moderner VLSI⁴-Chips. Abbildung 13 zeigt den Forschungsstand etwa 1999. Drei Jahre später (Frühjahr 2002) lagen die Strukturgrößen in den modernsten kommerziell hergestellten Chips bei etwa 60 bis 100 nm. Experimentiert wird mit Strukturdimensionen, die deutlich darunter liegen (Abbildung 14).

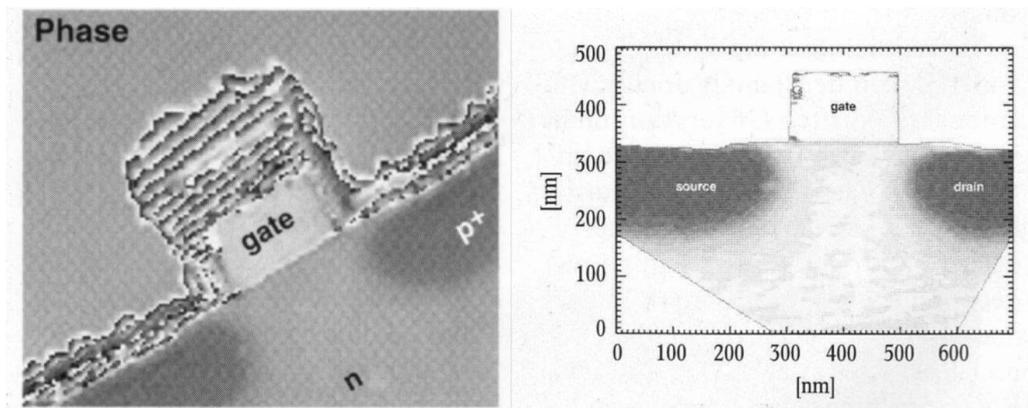
Daneben experimentieren verschiedene Forschungsgruppen mit Einzelelektronen-Transistoren, die zwar noch nicht ganz so klein zu bauen sind, aber mit der kleinstmöglichen Ladung, einem einzelnen Elektron, zwischen leitendem und nichtleitendem Zustand umgeschaltet werden können. Eine Realisierung, die bei Zimmertemperatur funktioniert, ein leicht geknicktes Kohlenstoff-Nano-Röhrchen, zeigt Abbildung 15.

3.8 Kennlinien

Die verschiedenen Spannungen und Ströme am Transistor sind alle mehr oder weniger stark jeweils von allen anderen abhängig. So ist der Kollektorstrom in erster Linie eine Funktion des Basisstroms, aber auch

³ Surface Mounted Device.

⁴ Very Large Scale Integration



Phasenbild eines „buried channel PMOS-Transistors“ mit einer Kanallänge von 180 nm, aufgenommen mit Elektronenholographie (links). Aus dem Phasenbild ergibt sich das zweidimensionale elektrostatische Potential (rechts). Deutlich zu erkennen sind die p-leitenden Bereiche, die die source- und drain-Kontakte definieren, sowie der n-leitende Kanal dazwischen mit dem aufgebrauchten Kontakt aus Polysilizium. Im Potentialbild entspricht der dunkelste Grauwert einer Spannung von $-0,50$ V, der hellste einer Spannung von $+0,45$ V. Die Schwarz-Weiß-Streifen im Phasenbild sind auf Phasendifferenzen größer als 2π zurückzuführen. (Quelle: W. D. Rau, IHP)

Abbildung 13: Experimentelle PMOS-Struktur. Quelle: Physikalische Blätter Band 55 (1999), Heft 4, Seite 16. Fehler im Text?

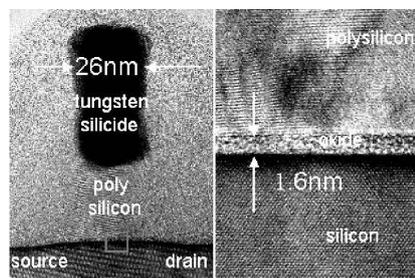


Abbildung 14: ‘A high-resolution transmission electron micrograph through a cross-section of a nominally 25nm gate length MOSFET. The red inset shows the nominally 1.4nm thick gate oxide’. Quelle: G. Timp et al., University of Illinois, 2001.

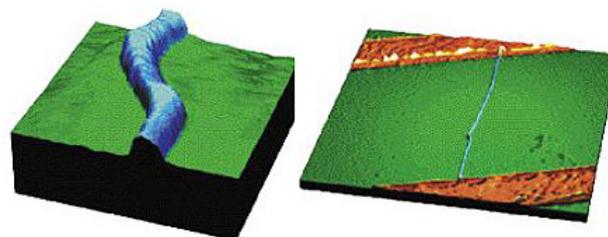


Abbildung 15: Einzel-elektronen-Transistor aus einem Kohlenstoff-Nano-Röhrchen. Quelle: Cees Dekker und Henk Postma, Universität Delft.

der Kollektor-Emitter-Spannung. Indirekt natürlich auch eine Funktion der Basis-Emitter-Spannung, da davon der Basisstrom abhängt. Die Spannungen und Ströme für einen bipolaren Transistor im Emitter-schaltung (Emitter als Bezugspotenzial) sind in Abbildung 16 definiert.

3.9 Vier-Quadranten-Kennlinienfeld

Einen Überblick über die Eigenschaften eines Transistors erhält man, wenn man den funktionalen Zusammenhang zwischen verschiedenen

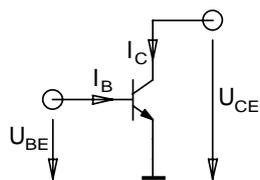


Abbildung 16: Spannungen und Ströme am bipolaren Transistor.

Größen graphisch – als *Kennlinien* – aufträgt. Dabei wird eine Größe in Abhängigkeit von einer zweiten aufgetragen, die beiden anderen bleiben dabei konstant oder werden sinngemäß mit verändert.

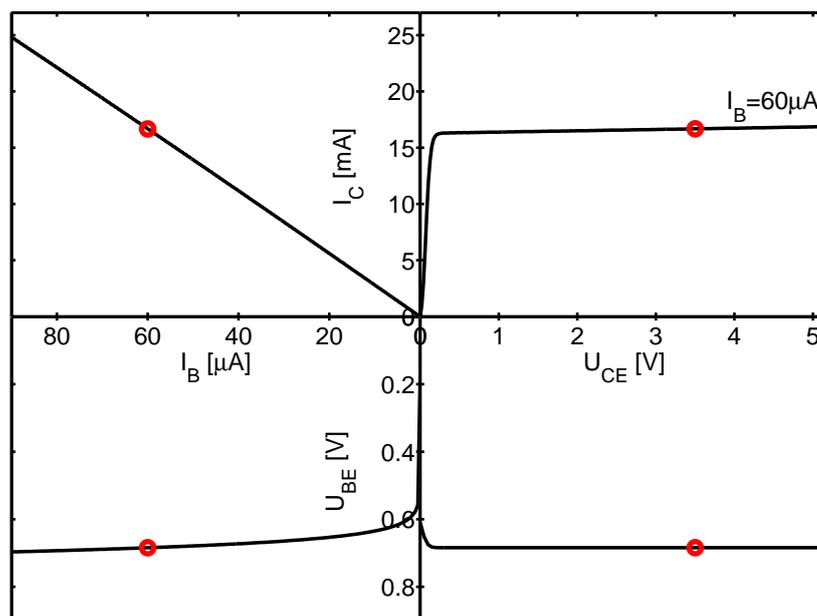


Abbildung 17: Vier-Quadranten-Kennlinienfeld eines Transistors. Die Punkte auf jeder Kennlinie definieren den jeweils gleichen Betriebszustand: $I_B = 60 \mu\text{A}$, $U_{CE} = 3.5 \text{ V}$ sowie die zugehörigen I_C und U_{BE} .

Einen guten Gesamteindruck kann man sich mit dem Vier-Quadranten-Kennlinienfeld verschaffen, ein Beispiel zeigt Abbildung 17. Darin sind die folgenden Kennlinien zusammengestellt:

Eingangskennlinie: Abhängigkeit des Basisstroms von der Basis-Emitter-Spannung, $I_B = I_B(U_{BE})$ bei $U_{CE} = \text{const.}$, links unten.

Ausgangskennlinie: Abhängigkeit des Kollektorstroms von der Kollektor-Emitter-Spannung, $I_C = I_C(U_{CE})$ bei $I_B = \text{const.}$, rechts oben.

Stromverstärkungskennlinie, Übertragungskennlinie: Abhängigkeit des Kollektorstroms vom Basisstrom, $I_C = I_C(I_B)$ bei $U_{CE} = \text{const.}$, links oben.

Rückwirkungskennlinie: Abhängigkeit der Basis-Emitter-Spannung von der Kollektor-Emitter-Spannung, $U_{BE} = U_{BE}(U_{CE})$ bei $I_B = \text{const.}$, rechts unten.

Darüber hinaus kann man noch die Steuerkennlinie definieren, die Ab-

hängigkeit des Kollektorstroms von der Basis-Emitter-Spannung.

Aus Gründen der Übersichtlichkeit ist in den Quadranten von Abbildung 17 nur jeweils eine einzelne Kennlinie gezeichnet, eine Variation der jeweils konstant gehaltenen Größe führt zu Kennlinienscharen mit mehr oder weniger ausgeprägter Verteilung der Kennlinien.

Die vier Kennlinien beschreiben das Verhalten eines Transistors einigermaßen vollständig, sind aber zum Teil nicht sehr aussagekräftig. Die Stromverstärkungskennlinie ist ziemlich gut linear, die Rückwirkungskennlinie ist nicht besonders wichtig für die praktische Anwendung von Transistoren. Man beschränkt sich daher meist auf die beiden restlichen, die Eingangskennlinie und das Ausgangskennlinienfeld (Abbildung 18). Die Eingangskennlinie kann für praktische Anwendun-

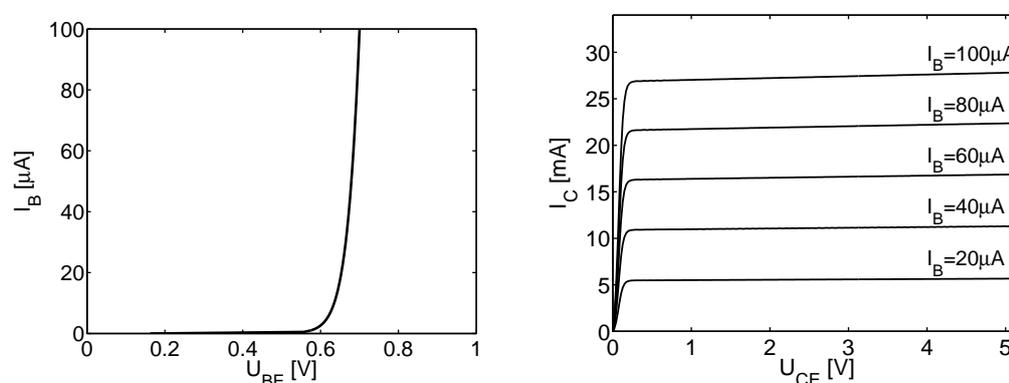


Abbildung 18: Eingangskennlinie (links) und Ausgangskennlinienfeld (rechts) eines bipolaren Transistors (BC 337-25).

gen als unabhängig von der Kollektor-Emitter-Spannung U_{CE} angesehen werden. Sie entspricht in guter Näherung einer Diodenkennlinie (vgl. Anleitung *Dioden*). Das Ausgangskennlinienfeld wird in der Regel über den Basisstrom I_B parametrisiert (in selteneren Fällen über die Basis-Emitter-Spannung U_{BE}). Die Ausgangskennlinien steigen bei kleinen Spannungen steil an, dieser *Sättigungsbereich* und die zugehörigen Spannungen $U_{CE,sat}$ sind für Schaltanwendungen wichtig. Für größere Spannungen verlaufen die Kennlinien nahezu horizontal (Bereich annähernd konstanter Stromverstärkung). Verlängert man die leicht ansteigenden Kurven nach links, schneiden alle die Spannungsachse etwa in einem Punkt, der so genannten *Early-Spannung*⁵ (-30... -150 V).

3.10 Transistor als Vierpol

Formalisiert kann man den Transistor – ein Bauelement mit drei Anschlüssen – als *Vierpol* beschreiben. Die Korrespondenz zwischen den Transistor- und Vierpolgrößen zeigt Abbildung 19.

Diese formale Beschreibung ist für die Netzwerktheorie und die Simulation wichtig. Der Transistor ist dann ganz allgemein durch eine Matrix

⁵ Benannt nach James M. Early, Bell Laboratories, der diesen Effekt erstmals erklärte.

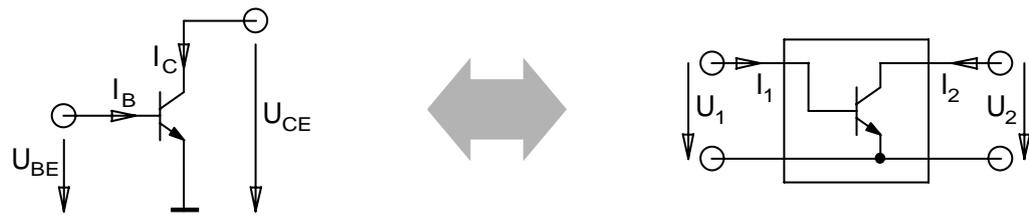


Abbildung 19: Spannungen und Ströme am Transistor (links) und korrespondierende Vierpolbeschreibung (rechts) für die Emitterschaltung.

definiert, die zwei der Variablen U_1 , I_1 , U_2 , I_2 auf die beiden anderen abbildet. Grundsätzlich sind diese Beschreibungen zueinander äquivalent, jede der möglichen Matrizen lässt sich in jede andere umrechnen.

Y-Matrix, Y-Parameter: Denkt man an das Knotenpotenzialverfahren bei der Netzwerksimulation, so liegt eine Beschreibung durch eine Leitwert- oder Y-Matrix nahe (überlegen Sie, warum). Mit einer derartigen Matrix werden die Ströme aus den Spannungen berechnet, die einzelnen Matrixelemente y_{ik} sind Leitwerte, reziproke Widerstände:

$$\begin{pmatrix} I_1 \\ I_2 \end{pmatrix} = \begin{pmatrix} y_{11} & y_{12} \\ y_{21} & y_{22} \end{pmatrix} \cdot \begin{pmatrix} U_1 \\ U_2 \end{pmatrix} \quad (4)$$

Die Beschreibung durch eine Y-Matrix ist für die Netzwerkberechnung recht nützlich, hat aber den Nachteil, dass die darin auftretenden Y-Parameter nicht sehr anschaulich sind und die physikalische Funktion – zumindest bei bipolaren Transistoren – nur sehr indirekt beschreiben.

H-Matrix, H-Parameter: Sehr viel häufiger verwendet wird daher die Hybrid- oder H-Matrix, die durch eine zunächst recht willkürlich erscheinende Abbildung definiert ist:

$$\begin{pmatrix} U_1 \\ I_2 \end{pmatrix} = \begin{pmatrix} h_{11} & h_{12} \\ h_{21} & h_{22} \end{pmatrix} \cdot \begin{pmatrix} I_1 \\ U_2 \end{pmatrix} \quad (5)$$

Erst bei genauerer Betrachtung wird die Korrespondenz mit dem Vier-Quadranten-Kennlinienfeld der Abbildung 17 augenfällig. Die H-Parameter lassen sich mithin vollständig aus den Kennlinien dort ableiten. Überlegen Sie, welcher H-Parameter der Stromverstärkung β entspricht.

Großsignal, Kleinsignal: Die Kennlinien in Abbildung 17 sind hochgradig nichtlinear, die Parameter (Matrixelemente) hängen daher von den Größen selbst ab, zwischen denen sie vermitteln. Bei so genannten Großsignalanwendungen muss man das berücksichtigen, die H- oder Y-Parameter sind dann nicht konstant, sondern hängen von den anliegenden Spannungen ab. Anders bei so genannten Kleinsignalanwendungen. Dort legt man einen *Arbeitspunkt* fest, um den die Spannungen und Ströme nur geringfügig verändert werden. Die H- oder Y-Matrizen vermitteln dann den Zusammenhang zwischen kleinen Spannungs- und Stromänderungen, am Arbeitspunkt werden die Kennlinien linearisiert, die H-Parameter sind durch die Steigungen der Kennlinien im Arbeitspunkt definiert und damit konstant. Aus Abbildung 17 ist zu erkennen,

dass die Steigungen der Kennlinien über relativ große Bereiche einigermaßen konstant sind. Das heißt, dass die Angabe konstanter Parameter (wie man sie in Datenblättern für Transistoren findet) eine sehr gute Näherung für Kleinsignalanwendungen ist.

S-Matrix, S-Parameter: Bei den bisher betrachteten Vierpol-Matrizen nimmt man im allgemeinen an, dass die Matrixelemente reell sind und nicht von den Eigenschaften der äußeren Quellen und Senken abhängen. Das ist eine gute Näherung, solange Störungen wie die Zeitverzögerung von Signalen, die Fehlanpassung von Impedanzen oder Signalreflexionen keine oder allenfalls eine untergeordnete Rolle spielen. Bei Hochfrequenz-Anwendungen ist das meist nicht mehr der Fall, Kapazitäten, Induktivitäten, der geometrische Aufbau beeinflussen das Verhalten von Bauelementen. Man beschreibt Transistoren in diesem Bereich durch die so genannten Streu- oder Scattering-Parameter, kurz S-Parameter bzw. die S-Matrix. Statt der Spannungen und Ströme an Ein- und Ausgang verwendet man einfallende und auslaufende Wellen (d. h. deren Amplituden) als Variablen. Zwischen den einfallenden Wellen a_1 und a_2 an den beiden Ports 1 und 2 und den auslaufenden Wellen b_1 und b_2 besteht dann der Zusammenhang

$$\begin{pmatrix} b_1 \\ b_2 \end{pmatrix} = \begin{pmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{pmatrix} \cdot \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} \quad (6)$$

S_{11} ist darin der Reflexionskoeffizient am Eingang, S_{21} der Transmissionskoeffizient am Eingang (= Verstärkung) usw. Zur sinnvollen Beschreibung werden diese Parameter für problemangepasste Quell- und Lastwiderstände angegeben (50 Ω , 60 Ω o. ä.). Näheres zur Beschreibung der Hochfrequenz-Eigenschaften durch S-Parameter finden Sie z. B. in [3, 4, 5].

3.11 Arbeitspunkt

Einen Lastwiderstand R im Kollektorkreis des Transistors (Abbildung 20) kann man im Ausgangskennlinienfeld dadurch berücksichtigen, dass man eine Widerstandsgerade einzeichnet, die durch

$$I_R = \frac{U - U_{CE}}{R} \quad (7)$$

beschrieben wird. Darin ist U die (feste) Betriebsspannung.

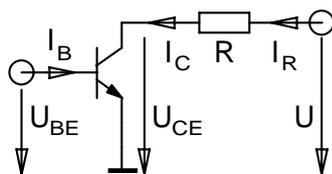


Abbildung 20: Transistor in Emitterschaltung mit Lastwiderstand R im Kollektorkreis.

Die möglichen Betriebszustände des Transistors sind durch

$$I_R = \frac{U - U_{CE}}{R} = I_C \quad (8)$$

definiert, das heißt durch Schnittpunkte der Widerstandsgeraden mit einer der Ausgangskennlinien. Da diese beliebig dicht liegen (der Übersichtlichkeit wegen werden immer nur ein paar davon gezeichnet), gibt es unendlich viele *Betriebszustände*, die durch Basisstrom bzw. Basis-Emitter-Spannung eingestellt werden. Von der spezifischen Anwendung hängt es ab, in welchem Bereich man den Transistor betreibt. Bei Analoganwendungen (z. B. Wechselspannungsverstärker mit einem einzelnen Transistor) sorgt man meist dafür, dass der Ruhebetriebszustand – der *Arbeitspunkt* – etwa in der Mitte des verfügbaren Bereichs liegt, dann ist der *Aussteuerbereich* groß, man erreicht gute Linearität (*A-Betrieb*). Nachteil ist der mit dem *Ruhestrom* verbundene Leistungsverbrauch. Geringeren Verbrauch erzielt man durch Arbeitspunkte mit geringerem oder ganz ohne Ruhestrom (*AB-Betrieb*, *B-Betrieb*, *C-Betrieb*). Die Anwendung muss allerdings dafür geeignet sein (*Gegentakt*-Schaltung, Hochfrequenz-Anwendungen). Bei Digitalanwendungen (Ein-Aus-Schalter) wechselt man zwischen stromlosem und maximal stromführendem Zustand des Transistors. Abbildung 21 gibt einen Überblick über die diskutierten Arbeitspunkte im Kennlinienfeld.

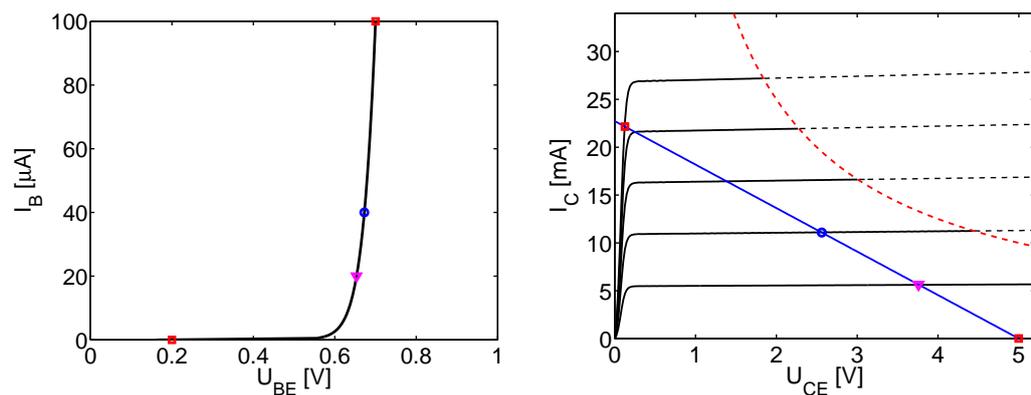


Abbildung 21: Arbeitspunkte im Kennlinienfeld des Transistors: \circ – Symmetrischer Arbeitspunkt (Verstärker in A-Betrieb), ∇ – Unsymmetrischer Arbeitspunkt (AB-Betrieb), \square – Stromloser und maximal stromführender Zustand (Transistor als Schalter). — : Widerstandsgerade ($R = 220 \Omega$), - - - : Hyperbel maximaler Verlustleistung (hier willkürlich mit 50 mW angenommen).

3.12 Grenzwerte

Der zulässige Betriebsbereich (SOA – Safe Operating Area) eines Transistors ist durch mehrere Grenzen definiert. Im wesentlichen sind dies der maximale Kollektorstrom $I_{C,\text{max}}$, die maximale Verlustleistung $P_{\text{tot,max}}$ und die maximale Kollektor-Emitter-Spannung $U_{CE,\text{max}}$ ⁶. Diese Grenzwerte, die im Datenblatt des Transistors vermerkt sind (*Maximum Ratings*, *Limiting Values*), dürfen im Betrieb auf keinen Fall überschrit-

⁶ Eine weitere Grenze ist durch den *Durchbruch 2. Art* gegeben. \Rightarrow Stichwort ‘Durchbruch 2. Art’ oder ‘Second Breakdown’ in der Literatur oder in Suchmaschinen.

ten werden⁷. Abbildung 22 zeigt als Beispiel diese Grenzwerte für den Transistor BC 337 (npn-Transistor im Praktikum) – $I_{C,max} = 500 \text{ mA}$, $P_{tot,max} = 650 \text{ mW}$, $U_{CE,max} = 45 \text{ V}$.

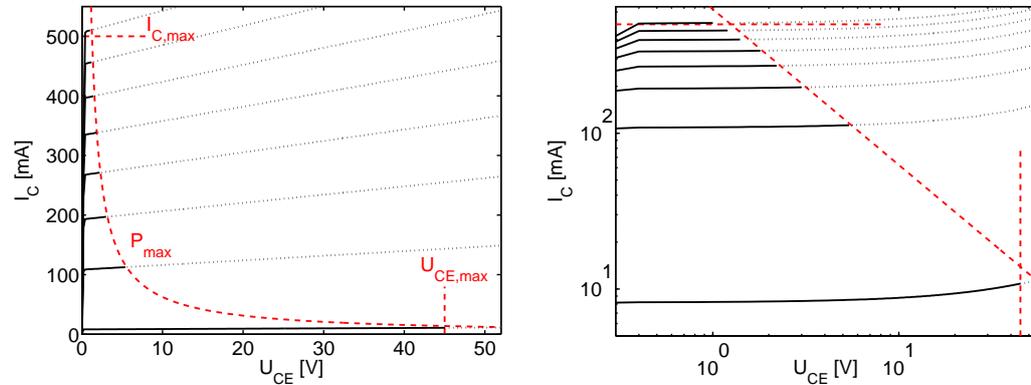


Abbildung 22: Grenzwerte im Ausgangskennlinienfeld, links lineare, rechts logarithmische Skalen.

3.13 Grundschaltungen

Bisher hatten wir angenommen, dass der Emitter das Bezugspotenzial darstellt (Emitterschaltung). Vom Transistorprinzip her gesehen ist dies nahe liegend. Man kann aber auch einen der beiden anderen Anschlüsse als Bezugspotenzial (Masse) festlegen, kommt dann zur Basis- und Kollektorschaltung. Abbildung 23 zeigt diese drei Grundschaltungen.

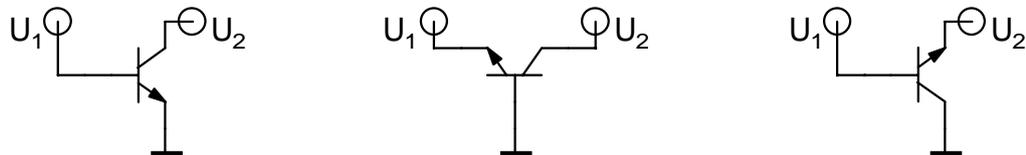


Abbildung 23: Transistorgrundschaltungen, links Emitterschaltung, Mitte Basisschaltung, rechts Kollektorschaltung (Emitterfolger).

Die drei Grundschaltungen unterscheiden sich in den Eingangs-, Ausgangs- und Verstärkungscharakteristiken deutlich, typische Werte sind in Tabelle 1 zusammengestellt.

Bei der Emitterschaltung ist die Ausgangsspannung gegenüber der Eingangsspannung um π in der Phasenlage verschoben. Das führt zu einer bei höheren Frequenzen sehr störenden Spannungsgegenkopplung durch die Basis-Kollektor-Kapazität (Miller-Effekt).

Wie die Emitterschaltung lassen sich auch die anderen Grundschaltungen formalisiert als Vierpole beschreiben mit Parametern, die in entsprechender Weise linearisierbar sind. Falls Verwechslungsmöglichkei-

⁷ Kurzzeitige Überschreitungen von maximalem Kollektorstrom und maximaler Verlustleistung sind in gewissen Grenzen zulässig, oft ist in Datenblättern ein zweiter Wert für den maximalen Kollektorstrom für Impulsbetrieb angegeben.

	Emitterschaltung	Basisschaltung	Kollektorschaltung
Eingangswiderstand	100 Ω ... 10 k Ω	10 ... 100 Ω	10 ... 100 k Ω
Ausgangswiderstand	1 ... 10 k Ω	10 ... 100 k Ω	10 ... 100 Ω
Spannungsverstärkung	20 ... 100	100 ... 1000	1
Stromverstärkung	10 ... 500	<1	10 ... 500
Phasendrehung	π	0	0
Anwendung	Verstärker Schalter	HF-Verstärker	Impedanzwandler

Tabelle 1: Charakteristika der drei Transistorgrundschaltungen.

ten bestehen, sollte man die Schaltungsart bei den Parametern vermerken (h_{21e} , oft auch als h_{fe} in Datenblättern, für die Stromverstärkung in Emitterschaltung).

3.14 Kennlinien von Feldeffekttransistoren

Das Gate bei Feldeffekttransistoren ist entweder eine isolierte Halbleiterschicht oder eine Seite eines in Sperr-Richtung gepolten p-n-Übergangs. Zumindest bei niedrigen Frequenzen fließt daher kein (MOSFET) oder allenfalls ein minimaler Gate-Strom (Diodensperrstrom beim JFET). Die Steuerung des Drain-Stroms erfolgt nahezu leistungslos, der Eingangswiderstand ist praktisch unendlich. Eine Eingangskennlinie $I_G = I_G(U_{GS})$ (Bezeichnungen der Spannungen und Ströme am Feldeffekttransistor sind in Abbildung 24 definiert) wie bei bipolaren Transistoren macht daher bei Feldeffekttransistoren wenig Sinn.

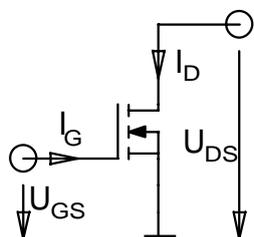


Abbildung 24: Spannungen und Ströme am Feldeffekttransistor.

Kennlinienbereiche: Man kann beim FET drei Arbeitsbereiche definieren, die sich in den Kennlinien widerspiegeln und die sich deutlich voneinander unterscheiden. Im *Sperrbereich* ist kein leitender Kanal vorhanden, die Gate-Source-Spannung U_{GS} ist kleiner⁸ als die Schwellenspannung (Threshold-Spannung U_{th}). Bei kleiner Drain-Source-Span-

⁸ Die Spannungen, Kleiner-Größer-Relationen usw. gelten für N-Kanal-FETs, für P-Kanal-FETs sind sie sinngemäß umzukehren.

nung U_{DS} ist der Drain-Strom annähernd linear von U_{DS} abhängig, die Steigung (Leitfähigkeit des Kanals) hängt von U_{GS} ab. Der FET entspricht in diesem Bereich einem (spannungsgesteuerten) ohmschen Widerstand, man spricht vom *Ohmschen Bereich*. Bei größerem U_{DS} wird der Kanal in der Nähe des Drain-Anschlusses abgeschnürt, wenn die dort wirkende Gate-Spannung, d. h. die Differenz zwischen Gate- und Drain-Spannung, kleiner wird als U_{th} . Der Drain-Strom steigt mit weiter zunehmendem U_{DS} nur noch wenig an, der *Abschnürungsbereich*⁹ ist erreicht.

Ohmscher Bereich: Beim Feldeffekttransistor wird der Kanal sowohl durch die Gate-Spannung wie auch durch die Drain-Source-Spannung beeinflusst. Bei kleinen Drain-Source-Spannungen ist der Einfluss beider auf die Leitfähigkeit des Kanals jeweils annähernd linear. Zunehmende Gate-Spannung erhöht die Leitfähigkeit, zunehmende Drain-Source-Spannung verkleinert sie. Die durch die Drain-Spannung verursachte Änderung der Feldverhältnisse im Kanal wächst linear von Null am Source-Anschluss bis zu einem Maximalwert am Drain-Anschluss, die Wirkung der Drain-Source-Spannung ist mithin etwa halb so groß wie die der Gate-Spannung. Insgesamt ergibt sich dadurch für den Drain-Strom im Bereich kleiner Drain-Source-Spannungen⁸

$$I_D = KU_{DS}(U_{GS} - U_{th} - U_{DS}/2) \quad . \quad (9)$$

Abschnürungsbereich: Sobald U_{DS} die Größe von $U_{GS} - U_{th}$ ($= U_{DS,ab}$) erreicht hat, fängt der Kanal an, sich abzuschnüren (*pinch off*), der Stromfluss wird durch Injektion aus dem Restkanal aufrecht erhalten. Der Strom nimmt dann mit steigender Spannung nur noch wenig zu. Diese Zunahme ist einigermaßen linear – bedingt durch die Änderung der Kanalgeometrie mit steigender Drain-Spannung (Stichwort: *Kanallängenmodulation*). Setzt man $U_{DS,ab}$ in Gleichung 9 ein, wird der Strom im Abschnürungsbereich

$$I_D = \frac{K}{2}(U_{GS} - U_{th})^2 \quad (10)$$

bzw. unter Berücksichtigung des linearen Anstiegs mit steigendem U_{DS}

$$I_D = \frac{K}{2}(U_{GS} - U_{th})^2(1 + \lambda U_{DS}) \quad . \quad (11)$$

Wie beim bipolaren Transistor (Kapitel 3.9) kann der lineare Anstieg auch durch eine Early-Spannung $U_A = 1/\lambda$ beschrieben werden.

Übertragungskennlinie: Gleichung 10 bzw. Gleichung 11 mit konstantem U_{DS} beschreiben die Übertragungskennlinie $I_D = I_D(U_{GS})$. Reale Kennlinien weichen natürlich etwas von diesem idealisierten theoretischen Verlauf ab. Die Übertragungskennlinie hängt nur sehr wenig von

⁹ Da der Drain-Strom nicht mehr ansteigt, wird dieser Bereich oft auch als Sättigungsbereich bezeichnet. *Sättigung* bezeichnet hier allerdings etwas völlig anderes als beim bipolaren Transistor.

der Drain-Spannung ab, diese Abhängigkeit braucht meist nicht berücksichtigt zu werden. Je nach Typ des Feldeffekttransistors liegt der relevante Bereich bei $U_{GS} < 0$ (n-Kanal JFET, selbstsperrender p-Kanal MOSFET), $U_{GS} \approx 0$ (selbstleitende MOSFETs) oder $U_{GS} > 0$ (p-Kanal JFET, selbstsperrender n-Kanal MOSFET). Abbildung 25 zeigt typische Verläufe.

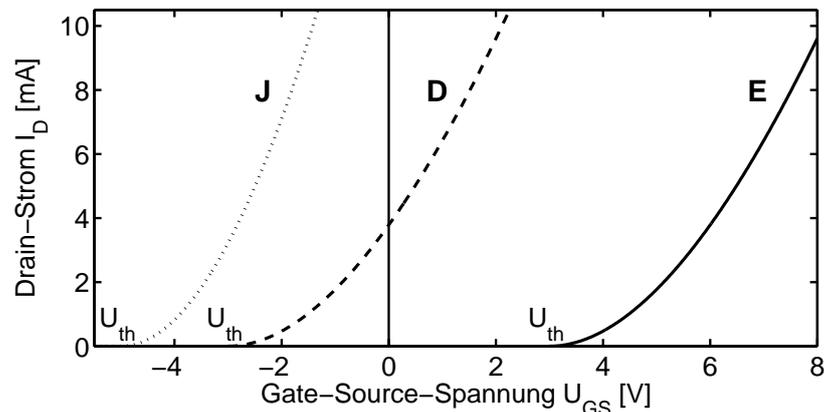


Abbildung 25: Typische Übertragungskennlinien $I_D = I_D(U_{GS})$ für verschiedene n-Kanal-Feldeffekttransistoren. J – Sperrschicht-FET (Junction FET), D – selbstleitender MOSFET (Depletion Type, Verarmungs-Typ), E – selbstsperrender MOSFET (Enhancement Type, Anreicherungs-Typ). Für p-Kanal-FETs müssten die Vorzeichen von U_{GS} und I_D umgekehrt werden.

Ausgangskennlinie: Eine ähnliche Rolle wie bei den bipolaren Transistoren spielt das Ausgangskennlinienfeld $I_D = I_D(U_{DS})$, Parameter ist dort die Gate-Spannung. Die Kennlinien werden für kleine U_{DS} näherungsweise durch Gleichung 9, für größere durch Gleichung 11 beschrieben. Die Kennlinienfelder für die im Praktikum verwendeten FETs sind in den Abbildungen 26 und 27 zusammengestellt.

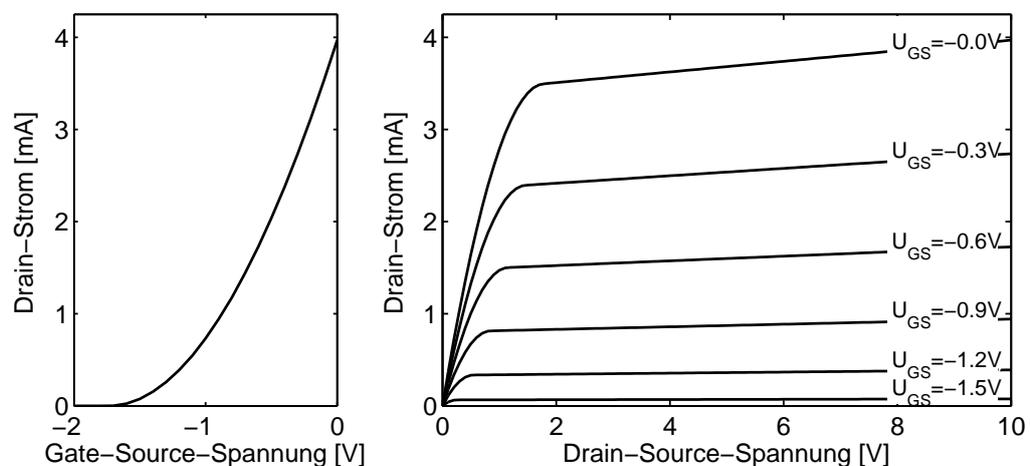


Abbildung 26: Übertragungskennlinie und Ausgangskennlinienfeld für den im Praktikum verwendeten JFET BF 245A.

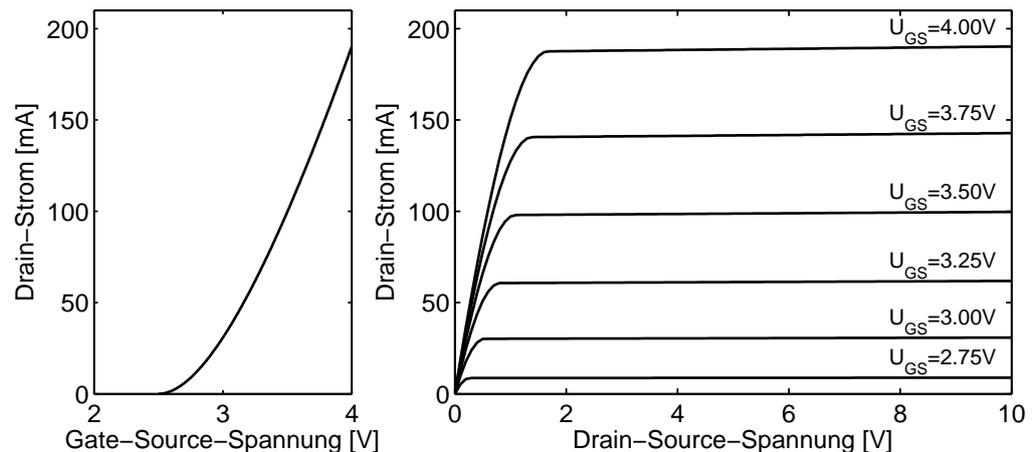


Abbildung 27: Übertragungskennlinie und Ausgangskennlinienfeld für den im Praktikum verwendeten MOSFET BS 170.

Arbeitspunkt: Zur Bestimmung eines geeigneten Arbeitspunktes wird das Ausgangskennlinienfeld durch die Widerstandsgerade für den verwendeten Lastwiderstand ergänzt. Die notwendige Gate-Spannung kann dann graphisch ermittelt werden.

Grenzwerte: Der Arbeitsbereich wird durch entsprechende Grenzwerte wie beim bipolaren Transistor eingeschränkt.

Grundschaltungen: Wie beim bipolaren Transistor kann man jeden der drei Anschlüsse als gemeinsamen Potenzialbezugspunkt (Masse) definieren (Abbildung 28). Man spricht von Source-, Gate- und Drain-Schaltung (Common Source, Common Gate, Common Drain). Die Unterschiede in

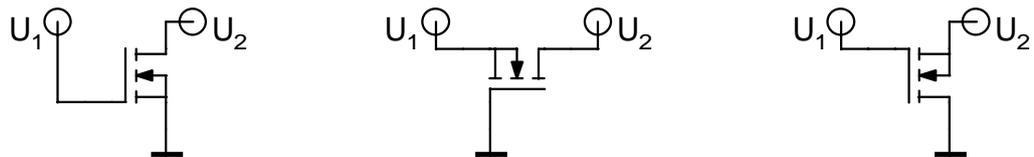


Abbildung 28: Grundschaltungen bei Feldeffekttransistoren, links Source-Schaltung, Mitte Gate-Schaltung, rechts Drain-Schaltung (Source-Folger).

den Charakteristika sind ähnlich wie bei den bipolaren Transistoren. Die Definition einer Stromverstärkung macht nur bei der Gate-Schaltung Sinn (dort etwa 1), bei den anderen Grundschaltungen wäre sie formal unendlich. Wegen der Phasendrehung von π zwischen Aus- und Eingangsspannung der Source-Schaltung begrenzt auch dort die Spannungsgegenkopplung durch die Drain-Gate-Kapazität die Hochfrequenzeigenschaften (Miller-Effekt).

Vierpoldarstellung: Eine Beschreibung als Vierpol wird etwas einfacher als beim bipolaren Transistor, da der Gate-Strom fest auf Null gesetzt werden kann, d. h. nicht weiter berücksichtigt werden muss. Bei höhe-

ren Frequenzen ist diese Vereinfachung allerdings nicht mehr möglich, da dann die Auf- bzw. Entladung der Gate-Kapazität eine Rolle spielt. Für die Beschreibung des Hochfrequenz-Verhaltens spielen wiederum die S-Parameter (vgl. entsprechenden Abschnitt in 3.10) eine wichtige Rolle.

Literatur und Links

- [1] F. S. Goucher, G. L. Pearson, M. Spark, G. K. Teal, W. Shockley. *Theory and Experiment for a Germanium p-n Junction*. Phys. Rev. **81**, 637 (1951).
- [2] Ulrich Tietze, Christian Schenk. *Halbleiterschaltungstechnik, n-te Auflage*. Springer, 199x.
- [3] <http://www.janson-soft.de/links.html>.
- [4] <http://www.sss-mag.com/spara.html>.
- [5] <http://www.sss-mag.com/smith.html>.
- [6] *Datenblätter von Dioden:*
<http://www.physik.uos.de/kbetzler/ep/proj/dsdiodes.shtml>.
- [7] *Halbleiterhersteller und mehr:*
<http://www.aufzu.de/semi/>.
- [8] *Applets zu Halbleiterbauelementen:*
<http://jas.eng.buffalo.edu/index.html>.
- [9] *Datenblätter von Transistoren:*
<http://www.physik.uos.de/kbetzler/ep/proj/dstrans.shtml>.