

# 1. Die p-n-Diode

Wir betrachten eine scharfe Grenzfläche zwischen n- und p-dotiertem Silizium. In der Praxis sind die Dotierungsprofile nicht so abrupt, dieser einfache Fall lässt sich aber leichter berechnen. Das p-Material mit einer Dotierungsdichte von  $N_A$  (Akzeptoratomen pro Volumeneinheit) befindet sich links, das n-Material mit einer Donatordichte  $N_D$  rechts. Dies ist in Figure 1 skizziert.

Wegen der ungleichen Konzentration der beweglichen Ladungsträger *diffundieren* Elektronen vom n-Material ins p-Material ( $\leftarrow$ ) und Löcher vom p- ins n-Material ( $\rightarrow$ ). Wegen der negativen Ladungen der Elektronen tragen *beide* Ladungsträger zu einem *Strom nach rechts* bei. Die abwandernden Elektronen und Löcher hinterlassen Bereiche ohne bewegliche Ladungsträger, die *Verarmungszone*, in der sich raumfeste, *geladene* Atomrümpfe befinden. Man spricht daher auch von einer *Raumladungszone* oder von der *Sperrschicht*. Der n-Bereich ist positiv geladen, der p-Bereich negativ. Dies führt zu einem elektrischen *Feld*, das in dieser Anordnung von rechts nach links zeigt, also negativ ist. Löcher *driften* in diesem Feld nach links, Elektronen nach rechts. Beide Ladungsträger tragen so zu einem *Feldstrom* bei, der von rechts nach links gerichtet ist. Im stationären Fall sind die Beträge dieser beiden Ströme gleich. Dies nutzen wir zur Berechnung der Diffusionsspannung aus, die sich durch den beschriebenen Ausgleichsvorgang ausbildet. Zur Berechnung legen wir den Nullpunkt der nach rechts gerichteten x-Achse auf die Grenzfläche. Die Raumladungszone im n-Bereich endet bei  $x_n > 0$ , die im p-Bereich bei  $x_p < 0$ .

Diffusionsstrom  $\rightarrow$

Feldstrom  $\leftarrow$

## 1.1 Berechnung der Diffusionsspannung

Wir beschreiben die Anordnung durch das ortsabhängige elektrische Feld  $E(x)$  und durch die ortsabhängigen Dichten der Ladungsträger. Wir betrachten hier nur die Löcherdichte  $p(x)$ , die Rechnung für die Elektronendichte  $n(x)$  verläuft analog.

Der Feldstrom ist proportional zum Feld, zur Dichte der Ladungsträger, zu deren Beweglichkeit  $\mu$  (von der wir vereinfachend annehmen, daß sie nicht vom Feld abhängt) und der Ladung  $q$ :

$$j_{Feld}(x) = q \mu p(x) E(x) \quad (1)$$

Feldstrom

Der Diffusionsstrom hängt *nicht* vom Feld ab. Ladungsträger wandern von Stellen hoher Konzentration in Richtung niedriger Konzentration, also *entgegen* dem Gradienten der Ladungsträgerkonzentration, daher das Minuszeichen. Die Stärke der Diffusion wird durch die Diffusionskonstante  $D$  beschrieben:

$$j_{Diff}(x) = -q D \frac{dp(x)}{dx} \quad (2)$$

Diffusionsstrom

Im stationären Gleichgewicht (es fließt kein Strom durch die Diode) müssen sich der Feldstrom und der entgegengerichtete Diffusionsstrom aufheben, also

Gleichgewicht

$$\begin{aligned} j_{Feld}(x) = q \mu p(x) E(x) &= q D \frac{dp(x)}{dx} = -j_{Diff}(x) \\ \frac{\mu}{D} E(x) &= \frac{1}{p(x)} \frac{dp(x)}{dx} \end{aligned} \quad (3)$$

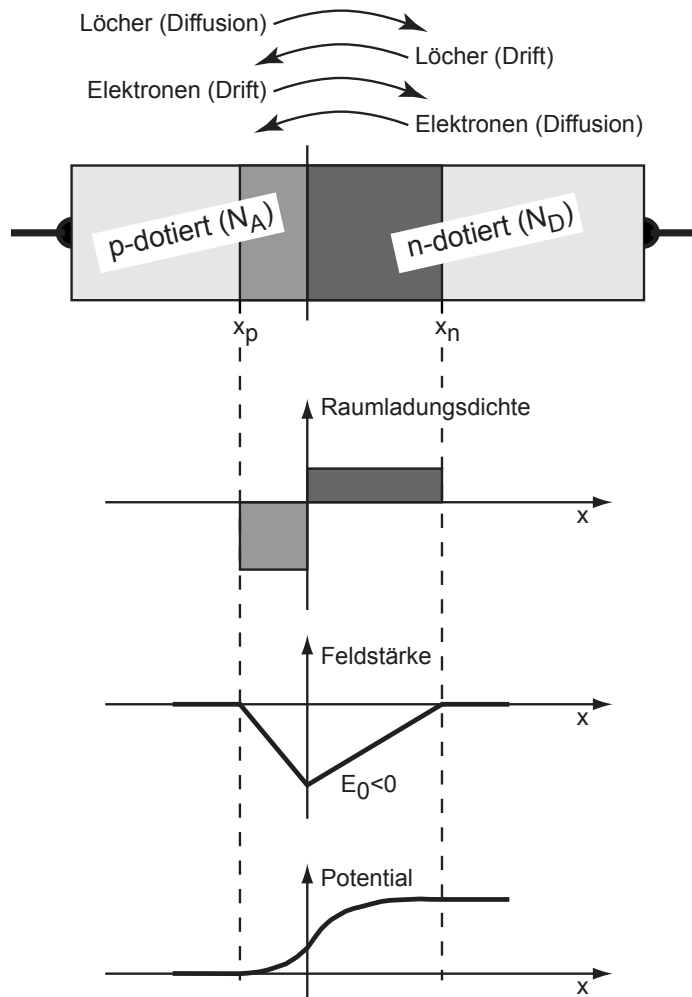


Abbildung 1: p-n-Übergang, Raumladungszone, Feld und Potential

$$-\frac{q}{kT} \frac{dV(x)}{dx} = \frac{1}{p(x)} \frac{dp(x)}{dx} \quad (4)$$

$$-\frac{q}{kT} dV(x) = \frac{dp(x)}{p(x)}. \quad (5)$$

Im Schritt von (3)  $\rightarrow$  (4) wurde das Feld  $E(x)$  durch den negativen Gradienten des Potentials ausgedrückt und die (nichttriviale!) Einstein-Beziehung  $\frac{D}{\mu} = \frac{kT}{q}$  benutzt. In (5) wurde zu einem Differential übergegangen, das direkt integriert werden kann:

Einstein-Beziehung

$$\begin{aligned} -\frac{q}{kT} \int_{V_p}^{V_n} dV(x) &= \int_{p_p}^{p_n} \frac{dp(x)}{p(x)} \\ -\frac{q}{kT} (V_n - V_p) &= \ln \left( \frac{p_n}{p_p} \right) \\ V_n - V_p &= \frac{kT}{q} \ln \left( \frac{p_p}{p_n} \right) \end{aligned} \quad (6)$$

Dabei sind  $V_p$  und  $p_p$  das Potential bzw. die Löcherdichte weit links von der Grenzschicht im p-Bereich,  $V_n$  und  $p_n$  die Werte weit rechts im n-Bereich.  $V_{bi} := V_n - V_p$  ist die so genannte *Diffusionsspannung*, die sich zwischen den zwei Seiten des p-n-Übergangs einstellt. Die Bezeichnung  $V_{bi}$  kommt vom englischen Ausdruck 'built-in Potential'. Die Löcherdichte weit links im p-Bereich kommt fast ausschließlich von der p-Dotierung:

$$p_p = N_A. \quad (7)$$

Im rechten n-Bereich ist die Elektronendichte  $n_n = N_D$  bekannt. Sie kann mit Hilfe des (nichttrivialen!) Massenwirkungsgesetzes  $n_n p_n = n_i^2$  in die Löcherdichte umgerechnet werden:

$$p_n = \frac{n_i^2}{N_D} \quad (8)$$

Hier ist  $n_i$  die 'intrinsische' Elektronen- bzw. Löcherdichte von reinem Silizium von  $1.45 \times 10^{10} \text{ cm}^{-3}$  bei Raumtemperatur (300K). Mit (7) und (8) erhält man aus (6) den Wert der Diffusionsspannung als Funktion der Dotierungen

$$V_{bi} = \frac{kT}{q} \ln \left( \frac{N_A N_D}{n_i^2} \right) \quad (9)$$

$$\approx \left[ \log \left( \frac{N_A}{n_i} \right) + \log \left( \frac{N_D}{n_i} \right) \right] \times 60 \text{ mV}. \quad (10)$$

Der rechte Ausdruck gilt bei Raumtemperatur. Wie zu erwarten war ist dieser Ausdruck symmetrisch in  $N_D$  und  $N_A$ .  $V_{bi}$  ist immer positiv, die n-Seite des Übergangs liegt also auf positiverem Potential als die p-Seite.

## 1.2 Betrachtung im Bändermodell (mit Hilfe der Fermi-Levels)

Elektronen sind Fermionen. Zwei Fermionen können sich nicht im selben Quantenzustand befinden. Bringt man viele Fermionen in ein begrenztes Raumvolumen, so müssen die neu hinzukommenden immer höhere Energieniveaus einnehmen, da die Niveaus mit niedriger Energie schon besetzt sind. (Da Fermionen einen Spin von 1/2 haben, kann jedes Energieniveau mit zwei Teilchen mit entgegengesetzten Spins besetzt werden.) Die Wahrscheinlichkeit, daß ein Fermion in einem Ensemble die Energie  $E$  hat, wird durch die quantenmechanische *Fermi-Dirac-Verteilung* (Fig. 2) gegeben:

$$f(E) = \frac{1}{1 + e^{(E-E_F)/kT}} \quad (11)$$

Hierin ist  $k$  die Boltzmann-Konstante,  $T$  die absolute Temperatur (in Grad Kelvin) und  $E_F$  die *Fermi-Energie*. Im Grenzfall  $T \rightarrow 0K$  sind alle Niveaus bis zur Fermi-Energie besetzt, die Verteilung wird zu einer Kastenfunktion. Bei höheren Temperaturen schmilzt die 'Fermi-Kante' ab und es gibt Fermionen mit höheren Energien. Die Wahrscheinlichkeit, ein Fermion mit einer Energie von  $E_F$  anzutreffen, ist immer 1/2. Für  $E - E_F \gg kT$  (also für hohe Fermion-Energien) ist die Exponentialfunktion im Nenner wesentlich größer als die Eins, so daß diese vernachlässigt werden kann. Die Fermi-Verteilung kann dann durch eine Boltzmann-Verteilung angenähert werden:

$$f(E) \approx e^{-(E-E_F)/kT} \quad \text{with} \quad E \gg E_F + kT \quad (12)$$

Die Fermi-Verteilung bleibt auch gültig, wenn gewisse Energien nicht besetzt wer-

Diffusionsspannung

Massenwirkungsgesetz

Fermi-Verteilung

Fermi-Energie

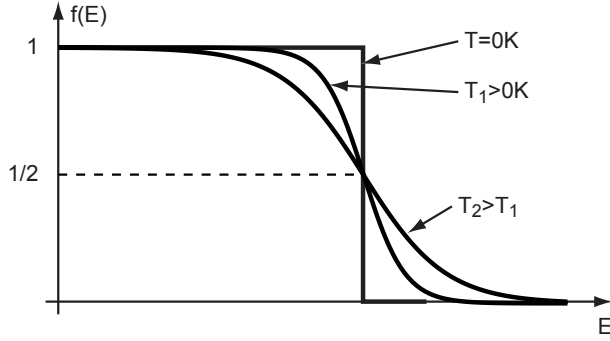


Abbildung 2: Fermi-Dirac-Verteilung für  $T_2 > T_1 > T_0 = 0K$

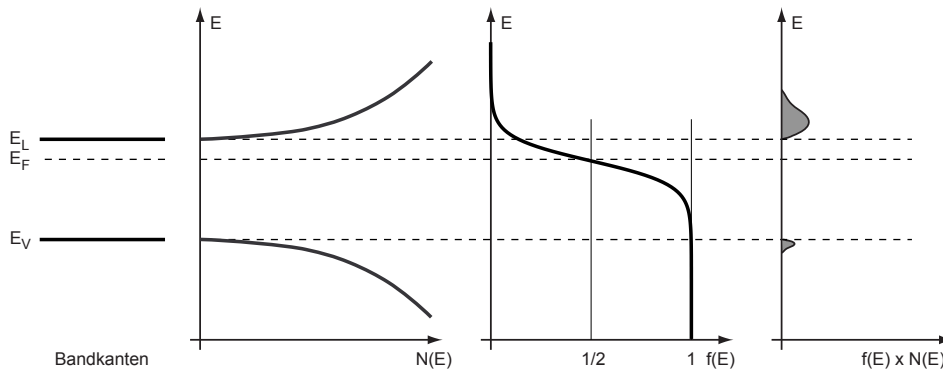


Abbildung 3: Fermiverteilung in der Bandlücke eines n-Halbleiters

den können, wie dies in der Bandlücke eines Halbleiters der Fall ist. Aus Fig. 3 wird ersichtlich, daß z.B. die Anzahl der Elektronen im Leitungsband mit der Fläche unterhalb der Fermiverteilung ab  $E = E_L$  zusammenhängt. Da in einem *intrinsischen Halbleiter* die Anzahl Elektronen im Leitungsband gleich der Anzahl Löcher im Valenzband ist, muß die Fermi-Energie in diesem Fall (etwa) in der Mitte der Bandlücke liegen (denn die Fermi-Verteilung ist punktsymmetrisch um  $(E_F, 1/2)$ ). Für einen intrinsischen Halbleiter gilt daher

intrinsischer Halbleiter

$$E_F \approx \frac{E_L + E_V}{2} \quad (13)$$

Später wird klar, warum das nicht ganz exakt gilt. In einem n-dotierten Halbleiter gibt es viel mehr Elektronen im Leitungsband als Löcher im Valenzband. Wir erwarten daher, daß das Fermi-Niveau höher liegt. Um den allgemeinen Fall zu berechnen, muß man beachten, daß aus quantenmechanischen Gründen im Kristallgitter die Dichte von besetzbaren Zuständen energieabhängig ist. Die *Zustandsdichte* ist im Leitungsband (ohne Herleitung)

Zustandsdichte

$$N(E) = \frac{(2m^*)^{3/2}}{2\pi^2} \left(\frac{2\pi}{h}\right)^3 (E - E_L)^{1/2} = C_1 (E - E_L)^{1/2}. \quad (14)$$

$m^*$  ist die *effektive Masse* der Ladungsträger, die sich durch quantenmechanische Effekte von der freien Masse der Ladungsträger unterscheiden kann und die i.A. für Elektronen und Löcher unterschiedlich ist. Die Anzahl Elektronen im Leitungsband ist also

effektive Masse

$$n = \int_{E_L}^{\infty} N(E) f(E) dE \quad (15)$$

$$\approx C_1 \int_{E_L}^{\infty} (E - E_L)^{1/2} e^{-(E-E_L)/kT} dE \quad (16)$$

$$= C_1 e^{-(E_L-E_F)/kT} \int_0^{\infty} y^{1/2} e^{-y/kT} dy$$

$$= C_1 (kT)^{3/2} e^{(E_F-E_L)/kT} \underbrace{\int_0^{\infty} x^{1/2} e^{-x} dx}_{\sqrt{\pi}/2} \quad (17)$$

$$= C_1 (kT)^{3/2} e^{(E_F-E_L)/kT} \cdot \frac{\sqrt{\pi}}{2}$$

$$= \underbrace{2 \left( \frac{2\pi m^* kT}{h^2} \right)^{3/2}}_{N_C} e^{(E_F-E_L)/kT}$$

$$n = N_C \cdot e^{(E_F-E_L)/kT} \quad (18)$$

Dabei haben wir von (15)→(16) die Boltzmann-Näherung (12) benutzt. Nach zweifacher Substitution ( $E - E_L = y$ ,  $y/kT = x$ ) kann das Integral in (17) gelöst werden. Aus (18) bekommt man nun die Lage des Fermi-Niveaus, wenn man annimmt, daß die Elektronen im Leitungsband alle von der Dotierung herrühren, also daß  $n = N_D$  ist.

$$E_{F,n} = E_L + kT \ln \frac{N_D}{N_{C,n}} < E_L \quad (19)$$

$$E_{F,p} = E_V - kT \ln \frac{N_A}{N_{C,p}} > E_V \quad (20)$$

Equation 20 ist der entsprechende Ausdruck für das Fermi-Niveau in einem p-dotierten Halbleiter. Die beiden Konstanten  $N_{C,n}$  und  $N_{C,p}$  müssen unterschieden werden, denn sie enthalten die unterschiedlichen effektiven Massen von Elektronen und Löchern. Sie sind i.A. größer als  $N_D$  oder  $N_A$ . Für einen intrinsischen Halbleiter gilt für Elektronen und Löcher

$$E_{F,i} = E_L + kT \ln \frac{n_i}{N_{C,n}} \quad (21)$$

$$E_{F,i} = E_V - kT \ln \frac{n_i}{N_{C,p}}. \quad (22)$$

Wenn  $N_{C,n} = N_{C,p}$  wäre würde hieraus (13) folgen. Die Abweichung davon ist in Silizium klein. Für den intrinsischen Halbleiter folgt aus den beiden obigen Gleichungen

$$E_V - E_L = kT \ln \frac{n_i}{N_{C,n}} + kT \ln \frac{n_i}{N_{C,p}}. \quad (23)$$

Bringt man p- und n-dotiertes Material in Kontakt, so muß das Fermi-Niveau auf beiden Seiten gleich sein, denn die zugehörige Fermi-Verteilung gilt universell. Da  $E_F$  im n-Bereich kurz unter der Leitungsbandkante liegt, werden die Energieniveaus

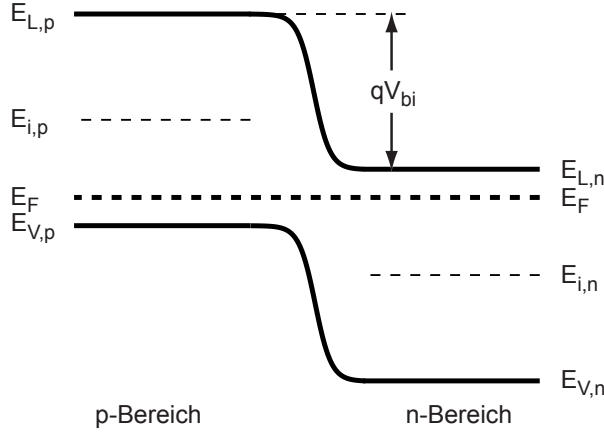


Abbildung 4: Lage der Bänder und des Fermi-Niveaus beim p-n-Übergang

dort abgesenkt, im p-Bereich werden sie entsprechend angehoben. Dies ist in Fig. 4 skizziert. Der sich ausbildende Potentialunterschied ergibt sich aus der Abbildung zu

$$\begin{aligned} qV_{bi} &= E_{F,p} - E_{F,n} \\ &= E_V - E_L - kT \ln \frac{N_A}{N_{C,p}} - kT \ln \frac{N_D}{N_{C,n}} \end{aligned} \quad (24)$$

$$= kT \ln \frac{n_i}{N_{C,n}} + kT \ln \frac{n_i}{N_{C,p}} - kT \ln \frac{N_A}{N_{C,p}} - kT \ln \frac{N_D}{N_{C,n}} \quad (25)$$

$$qV_{bi} = kT \ln \frac{n_i n_i}{N_A N_D} = -kT \ln \frac{N_A N_D}{n_i n_i} \quad (26)$$

Im Schritt von (24)  $\rightarrow$  (25) wurde (23) eingesetzt und die Logarithmen dann zusammengefasst. Berücksichtigt man, daß im Bänderdiagramm das Potential der negativ geladenen Elektronen dargestellt wird, so entspricht (26) genau dem vorherigen Ergebnis (9).

### 1.3 Berechnung der Dicke der p-n-Sperrschicht

Die Raumladungszone endet in Wirklichkeit nicht abrupt, sondern innerhalb einer dünnen Übergangszone. Diese ist jedoch so dünn, daß die Annahme des abrupten Übergangs sehr gut gerechtfertigt ist. Die (positive) Gesamtladung im n-Bereich der Raumladungszone (mit der Dicke  $x_n$ ) beträgt bei einer Diode der Fläche  $A$

$$Q_n = Ax_n q N_D > 0. \quad (27)$$

Für den p-Bereich gilt wegen  $x_p < 0$  vorzeichenrichtig

$$Q_p = Ax_p q N_A < 0. \quad (28)$$

Die beiden Raumladungszonen entstehen durch das Wegwandern von Elektronen bzw. Löchern. Da diese wegen Ladungserhaltung nicht verloren gehen können, muß

Ladungserhaltung

$Q_p = -Q_n$  sein, also

$$\begin{aligned} Q_p &= Ax_p q N_A = -Ax_n q N_D = -Q_n \\ x_p N_A &= -x_n N_D. \end{aligned} \quad (29)$$

Durch die Raumladungszone n entsteht ein elektrisches Feld, das an der Grenzfläche seinen höchsten Wert  $E_{max}$  hat. Dieser kann mit Hilfe des Gauß'schen Satzes berechnet werden. Dazu betrachten wir eine Dose der Fläche  $A$ , die z.B. weit links im (nicht verarmten) p-Bereich beginnt und an der Grenzschicht endet. Das Integral über die austretenden Felder enthält nur den Beitrag  $A \cdot E_{max}$  der rechten Dosenfläche (da links das Feld Null ist und am Rand der Dose die Feldlinien parallel zur Oberfläche verlaufen). Die eingeschlossene Ladung ist durch (28) gegeben. Der Gauß'sche Satz liefert

$$A \cdot E_{max} = \frac{Ax_p q N_A}{\epsilon} \Rightarrow E_{max} = \frac{q}{\epsilon} x_p N_A < 0. \quad (30)$$

Hierbei ist  $\epsilon = \epsilon_{Si} \epsilon_0$ . Die Feldstärke ist negativ, die Feldlinien zeigen also, wie in der Einleitung bereits vermutet, nach links. Wie in Fig. 1 angedeutet ist, fällt das Feld also von  $x_p$  an von Null linear auf  $E_{max}$  ab und steigt dann wieder linear an, bis es bei  $x_n$  wieder Null ist. Der *Potentialunterschied* zwischen der linken und der rechten Seite lässt sich nun durch Integration des Feldes berechnen. Das Integral ist einfach die Fläche unter der dreieckigen  $E(x)$ -Kurve, also

$$\Delta V = V_n - V_p = - \int_{x_p}^{x_n} E(x) dx = -\frac{1}{2} (x_n - x_p) E_{max} > 0. \quad (31)$$

Die Spannung ist also rechts positiver, wie es aufgrund der positiven Raumladung im n-Bereich ja auch sein muß. Nun kann man in (31) z.B.  $x_n$  mit Hilfe von (29) durch  $x_p$  ausdrücken. Nach Einsetzen von (30) erhält man

$$\Delta V = \frac{1}{2} \left( \frac{x_p N_A}{N_D} + x_p \right) \frac{x_p q N_A}{\epsilon} = \frac{q}{2\epsilon} \frac{(N_A + N_D) N_A}{N_D} x_p^2$$

oder

$$|x_p| = \sqrt{\frac{2\epsilon}{q} \frac{N_D}{(N_A + N_D) N_A} \Delta V} \xrightarrow{N_D \rightarrow \infty} \sqrt{\frac{2\epsilon}{q} \frac{\Delta V}{N_A}} \quad (32)$$

$$|x_n| = \sqrt{\frac{2\epsilon}{q} \frac{N_A}{(N_A + N_D) N_D} \Delta V} \xrightarrow{N_A \rightarrow \infty} \sqrt{\frac{2\epsilon}{q} \frac{\Delta V}{N_D}}. \quad (33)$$

Die Verarmungszonen sind um so dünner, je stärker dotiert die Bereiche sind (das ist ja bereits aus (29) klar). Die Gesamtdicke  $x_d$  der Sperrschicht kann z.B. aus (31) mit (30) und (32) berechnet werden. Sie ist

$$x_d = |x_p| + |x_n| = \sqrt{\frac{2\epsilon}{q} \frac{N_A + N_D}{N_A N_D} \Delta V}. \quad (34)$$

Die Spannung  $\Delta V$  setzt sich aus der Diffusionsspannung  $V_{bi}$  und einer von außen angelegten externen Spannung  $V_{ext}$  zusammen:

$$\Delta V = V_{bi} - V_{ext}. \quad (35)$$

Dabei haben wir das Vorzeichen so gewählt, daß  $\Delta V$  für  $V_{bi} = V_{ext}$  Null wird, der positive Pol der externen Spannung also an der p-Seite, der *Anode*, anliegt. Mit einer externen Spannung kann also die Dicke der Sperrschicht verändert werden.

Gauß'scher Satz

## 1.4 Die Kapazität der Sperrschicht

Wir berechnen hier nur die Kapazität einer in Sperrrichtung gepolten Diode, bei der sich eine von beweglichen Ladungsträgern freie Raumladungszone ausgebildet hat und in der daher (außer einem kleinen Leckstrom) kein Strom fließt. Im VLSI Design sind Dioden fast immer in Sperrrichtung geschaltet und bilden eigentlich unerwünschte 'parasitäre' Elemente, die insbesondere durch ihre Kapazität das Verhalten der Schaltung (meist negativ) beeinflussen.

Die Diode stellt einen Plattenkondensator mit dem Plattenabstand  $x_d$  gebildet dar, der mit Silizium mit der Dielektrizitätskonstante  $\epsilon_{Si}$  gefüllt ist. Die Kapazität ist daher

$$C_j = \epsilon \cdot \frac{A}{x_d} \quad (36)$$

$$\begin{aligned} &= A \sqrt{\frac{q\epsilon}{2} \frac{N_A N_D}{N_A + N_D} \frac{1}{V_{bi} - V_{ext}}} \\ &= A C_{j0} \left(1 - \frac{V_{ext}}{V_{bi}}\right)^{-1/2} \end{aligned} \quad (37)$$

Der Index von  $C_j$  zeigt an, daß es sich um die 'junction'-Kapazität handelt. Dieser Ausdruck wurde für einen abrupten Übergang vom p- ins n-dotierte Gebiet hergeleitet. In der Praxis ist der Verlauf fließend (je nach Herstellungsverfahren) wodurch insbesondere der Exponent in (37) nicht mehr 1/2, sondern kleiner ist.

In SPICE wird die Kapazität einer implantierten Diode daher meist mit folgender Formel berechnet:

$$C = A \times \mathbf{CJ} \times \left(1 - \frac{V}{\mathbf{PHA}}\right)^{-\mathbf{MJ}} + P \times \mathbf{CJSW} \times \left(1 - \frac{V}{\mathbf{PHP}}\right)^{-\mathbf{MJSW}} \quad (38)$$

Hierbei sind  $A$  und  $P$  die aus dem Design extrahierte Fläche (area) und Umfang (periphery) der Diode,  $V$  ist die angelegte Spannung. Die SPICE Modellparameter **CJ** und **CJSW** sind Konstanten für Fläche und 'sidewall'. **PHA** und **PHP** sind die eingebauten (Diffusions-)Potentiale und **MJ** und **MJSW** die Exponenten. Fläche und Peripherie haben unterschiedliche Parameter, da die Dotierungsverhältnisse und die Struktur der Grenzschichten sehr verschieden sein können. Bei kleinen Dioden sind beide Beträge oft vergleichbar!

## 1.5 Zahlenwerte

Wir betrachten einen abrupten p-n-Übergang mit einer (schwachen) n-Dotierung von  $N_D = 10^{16} \text{ cm}^{-3}$  und einer (starken) p-Dotierung von  $N_A = 10^{19} \text{ cm}^{-3}$ . Die Diffusionsspannung ist nach (10)

$$\begin{aligned} V_{bi} &= \left[ \log \left( \frac{10^{19}}{1.45 \times 10^{10}} \right) + \log \left( \frac{10^{16}}{1.45 \times 10^{10}} \right) \right] \times 60 \text{ mV} \\ &= [8.84 + 5.84] \times 60 \text{ mV} = 0.874 \text{ V.} \end{aligned} \quad (39)$$

Die Verarmungszone erstreckt sich fast nur in den schwächer dotierten n-Bereich. (Nach (29) ist der p-dotierte Bereich  $10^3$  mal dünner!). Equation 33 liefert mit



$\epsilon = \epsilon_{Si}\epsilon_0 = 11.9 \times 8.854 \times 10^{-12} \text{ F/m} = 1.05 \times 10^{-10} \text{ F/m}$  ohne äußere Spannung, also für  $\Delta V = V_{bi}$

$$\begin{aligned} x_n(0 \text{ V}) &\approx \sqrt{\frac{2\epsilon V_{bi}}{q N_D}} \\ &= \sqrt{\frac{2 \times 1.05 \times 10^{-10} \text{ F/m} \times 0.874 \text{ V}}{1.6 \times 10^{-19} \text{ C} \times 10^{22} \text{ m}^{-3}}} \\ &= 0.34 \mu\text{m}. \end{aligned} \quad (40)$$

Die Dotierungsdichte muß dabei von der gebräuchlichen Einheit  $\text{cm}^{-3}$  auf  $\text{m}^{-3}$  umgerechnet werden! Bei einer extern angelegten Sperrspannung von  $V_{ext} = -2 \text{ V}$  wächst die Dicke der Sperrschicht auf

$$x_n(-2 \text{ V}) \approx x_n(0 \text{ V}) \times \sqrt{\frac{V_{bi} - V_{ext}}{V_{bi}}} = 0.61 \mu\text{m}. \quad (41)$$

Die maximale Feldstärke (an der Grenzfläche) beträgt nach (30)

$$|E_{max}| = \frac{q}{\epsilon} x_p N_A = \frac{q}{\epsilon} x_n N_D = 5.2 \times 10^6 \text{ V/m} \quad (42)$$

Das liegt noch deutlich unter der Durchbruchfeldstärke von Silizium von  $3 \times 10^7 \text{ V/m}$ . Die Flächenkapazität einer  $2 \times 2 \mu\text{m}^2$  großen Diode beträgt nach (37)

$$C_j \approx \frac{\epsilon A}{x_d} = 1.2 \times 10^{-15} \text{ F} = 1.2 \text{ fF} \quad (43)$$

Der Rand dieser Diode ist  $4 \times 2 \mu\text{m}$  lang und hat mit seiner Tiefe von  $0.34 \mu\text{m}$  somit eine zusätzliche Fläche von  $2.7 \mu\text{m}^2$ . Bei gleichen Dotierungsprofilen erhöht der Rand also die Kapazität um zusätzlich 70%.

## 2. Der Sperrschicht-Feldeffekttransistor (JFET)

In einem Sperrschicht-Feldeffekttransistor (Junction-FET, JFET) wird der Querschnitt eines leitenden Kanals durch eine von der Seite eindringende Raumladungszone verändert. Die Raumladungszone ist die Sperrschicht einer p-n-Diode. Sie wird über eine an die Gate-Elektrode angelegte Sperrspannung verändert. In Fig. 5 ist der Aufbau eines n-Kanal JFET schematisch dargestellt. Ein Stück schwach n-dotiertes

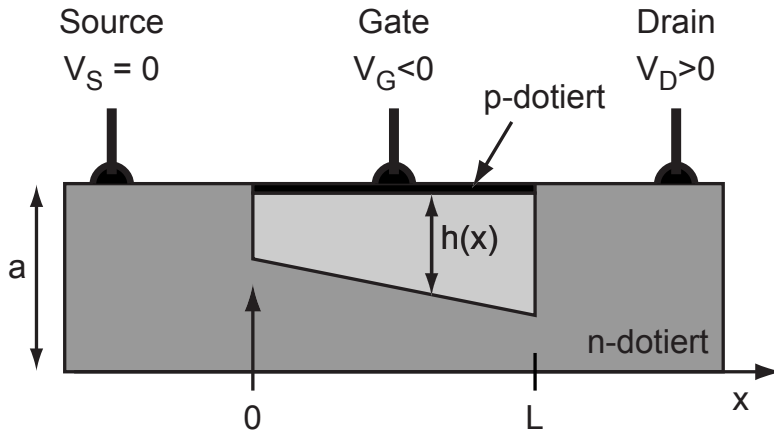


Abbildung 5: Schematischer Aufbau eines JFET

Silizium der Dicke  $a$  ist an *Source* und *Drain* kontaktiert, so daß zwischen diesen Elektroden bei einer angelegten Spannung ein Strom durch den *Kanal* fließen kann. Wir vereinbaren, daß die Source auf Massepotential liegt ( $V_S = 0\text{ V}$ ). An der Oberfläche des Siliziums ist eine dünne, stark p-dotierte Schicht der Länge  $L$  eingebaut, die mit der *Gate* Elektrode verbunden ist. An der Grenzfläche zum Kanal bildet sich eine ladungsträgerfreie und daher isolierende Verarmungszone aus deren Dicke vom Potentialunterschied zwischen dem Gate und dem Potential im Kanal abhängt. Dieses ist Null an der Source-Seite und steigt auf  $V_D$  an der Drain-Seite an, so daß die Raumladungszone dort dicker ist (s. Fig. 5). Mit einer negativen Spannung am Gate (bei positive Spannungen würde die Gate-Kanal-Diode leiten!) wächst eine Verarmungszone in den n-Bereich und verkleinert dort den Querschnitt des Kanals, so daß der Drainstrom abnimmt. Ab einer bestimmten Spannung (der 'Pinch-Off'-Spannung  $V_p$ ) versperrt die Raumladungszone den gesamten Kanal und der Stromfluß wird vollständig unterbrochen. Dies ist in der Übertragungskennlinie in Fig. 6 schematisch dargestellt. Bei einer Gatespannung  $V_G > V_p$  hängt die Dicke der Raumladungszone an der Drain-Seite auch von der Drainspannung ab. Für  $V_D = V_G - V_p$  schnürt der Kanal lokal ab und eine weitere Erhöhung von  $V_D$  führt nicht zu einem höheren Strom. Dies ist in der Ausgangskennlinie in Fig. 6 dargestellt. Den Bereich (fast) konstanten Drainstroms oberhalb von  $V_D = V_G - V_p$  nennt man den *Sättigungsbereich*, den Bereich in der Nähe von  $V_D = 0$ , wo der Drainstrom fast linear mit der Drainspannung zusammenhängt, den *linearen Bereich*.

Source, Drain  
Kanal

Gate

Sättigungsbereich  
Linearer Bereich

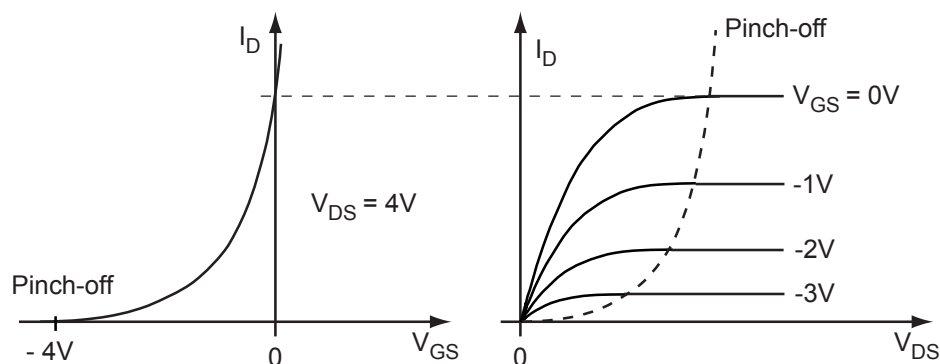


Abbildung 6: Übertragungskennlinie (links) und Ausgangskennlinie (rechts) eines n-Kanal JFET

## 2.1 Berechnung der Kennlinie

Zur einfachen Berechnung betrachten wir nur den Bereich unter der Dotierung, wir vernachlässigen also den (nahezu konstanten) Zuleitungswiderstand von den Drain/Source-Kontakten bis zum Kanal. Die untere Begrenzung des Kanals wird in der praktischen Realisierung des JFET durch eine weitere, großflächig p-dotierte Schicht erreicht, an der sich eine weitere Raumladungszone ausbildet. Die Ergebnisse dieses Abschnitts ändern sich dadurch ein wenig.

Das Potential im Kanal  $V(x)$  ist Null bei der Source (bei  $x = 0$ ) und steigt auf  $V_D$  bei  $x = L$  an. Es ist also nicht konstant. Die an der Gate-Kanal-Diode anliegende Spannung ist somit  $V_{ext} = |V_G| + V(x)$ , wobei  $|V_G|$  hier der Betrag der negativen Gate-Spannung sei. Die Raumladungszone hat daher nach (33) an der Position  $x$  längs des Kanals eine Dicke von

$$h(x) = \left[ \frac{2\epsilon}{qN_D} (|V_G| + V(x) + V_{bi}) \right]^{1/2} \quad (44)$$

Die Spannungen gehen mit positivem Vorzeichen ein, weil das Gate p-dotiert ist und die Diode in Sperrichtung polarisiert ist. Zur einfacheren Schreibweise absorbieren wir das zur Gatespannung additive Diffusionspotential in  $V_G'$ :

$$V_G' := |V_G| + V_{bi} > 0 \quad (45)$$

Da das Kanalpotential von links nach rechts ansteigt wird die Verarmungszone nach rechts immer dicker, wie dies in Fig. 5 angedeutet ist. Die Dicken an Source und Drain sind insbesondere

$$h_S = h(0) = \left[ \frac{2\epsilon}{qN_D} V_G' \right]^{1/2} \quad (46)$$

$$h_D = h(L) = \left[ \frac{2\epsilon}{qN_D} (V_G' + V_D) \right]^{1/2} \quad (47)$$

Kanalpotential

Bei genügend hoher Spannung füllt die Verarmungszone zunächst an der Drain-Seite die gesamte Dicke  $a$  des Kanals aus. Die ist der Fall, wenn

$$\begin{aligned} h_D &= \left[ \frac{2\epsilon}{qN_D} (V_{G'} + V_D) \right]^{1/2} = a \\ \Rightarrow V_p &:= V_{G'} + V_D = \frac{qa^2 N_D}{2\epsilon} \end{aligned} \quad (48)$$

Hier haben wir die häufig benutzte *Abschnürspannung* (engl. pinch-off-voltage)  $V_p$  eingeführt. Unter Benutzung von  $V_p$  können wir die Dicken der Verarmungszone bei Source (46) und Drain (47) einfacher schreiben:

Abschnürspannung

$$\frac{h_S}{a} = \sqrt{\frac{V_{G'}}{V_p}} \quad \text{und} \quad \frac{h_D}{a} = \sqrt{\frac{V_{G'} + V_D}{V_p}} \quad (49)$$

Die horizontal fließende Stromdichte an einer Stelle  $x$  im Kanal ist (vgl. (1))

$$j(x) = q N_D \mu E_x(x) = q N_D \mu \frac{dV(x)}{dx} \quad (50)$$

Dabei ist  $\mu = \mu_e$  die Mobilität der im Kanal fließenden Elektronen. Der gesamte Strom ergibt sich durch Multiplikation mit der Querschnittsfläche  $W(a - h(x))$ . Hier ist  $W$  die *Breite* (engl. width) des Gates. Der Gesamtstrom muß (Ladungserhaltung!) an jeder Stelle des Kanals gleich sein, er ist also unabhängig von  $x$ :

Gatebreite

$$I_D = W(a - h(x)) q N_D \mu \frac{dV(x)}{dx} \quad (51)$$

Das Differential  $dV(x)$  erhält man, indem man in (44)  $dh(x)/dx$  bildet und die Kettenregel verwendet:

$$dV(x) = \frac{qN_D}{\epsilon} h dh \quad (52)$$

Einsetzen von (52) in (51) liefert

$$I_D dx = \frac{Wq^2 N_D^2 \mu}{\epsilon} (a - h) h dh$$

Integration beider Seiten von der Source bis zur Drain liefert

$$\begin{aligned} \int_0^L I_D dx &= \frac{Wq^2 N_D^2 \mu}{\epsilon} \int_{h_S}^{h_D} (a - h) h dh \\ LI_D &= \frac{Wq^2 N_D^2 \mu}{\epsilon} \left[ \frac{ah^2}{2} - \frac{h^3}{3} \right]_{h_S}^{h_D} \\ I_D &= \frac{W}{L} \frac{q^2 N_D^2 \mu a^3}{2\epsilon} \left[ \frac{h_D^2}{a^2} - \frac{2}{3} \frac{h_D^3}{a^3} - \frac{h_S^2}{a^2} + \frac{2}{3} \frac{h_S^3}{a^3} \right] \end{aligned}$$

Nach Einsetzen der Ausdrücke aus (49) ergibt sich der Drainstrom im JFET im linearen Bereich zu

$$I_D = \frac{W}{L} I_0 \left[ \frac{V_D}{V_p} + \frac{2}{3} \left( \frac{V_{G'}}{V_p} \right)^{3/2} - \frac{2}{3} \left( \frac{V_D + V_{G'}}{V_p} \right)^{3/2} \right]. \quad (53)$$

## 2.2 Interpretation des Ergebnisses

Wir wollen Ausdruck diese Ergebnisse nun analysieren:

- Der Strom ist proportional zur  $W/L$ , also dem Verhältnis von Breite und Länge des Gates. Die Abhängigkeit von  $W$  ist trivial (Vergrößerung von  $W$  entspricht die Parallelschaltung mehrerer Transistoren), die Abhängigkeit von  $L$  weniger. Der gleiche Zusammenhang wird uns auch beim Feldeffekttransistor begegnen. Die geschickte Wahl dieser Parameter beim Entwurf einer Schaltung ist von zentraler Bedeutung.
- Der Strom hängt sowohl von  $V_G$  wie von  $V_D$  ab, da das Zusammenspiel dieser Spannungen insbesondere am drainseitigen Ende den Querschnitt des Kanals festlegt.
- Für  $V_D = 0$  V darf natürlich kein Strom fließen. Dies wird von (53) erfüllt, weil sich die beiden rechten Terme wegheben.
- Für sehr kleine Drainspannungen, genauer für  $V_D \ll V_{G'}$ , kann man den dritten Term entwickeln:

$$\begin{aligned} \left(\frac{V_D + V_{G'}}{V_p}\right)^{3/2} &= \left[\frac{V_{G'}}{V_p} \left(1 + \frac{V_D}{V_{G'}}\right)\right]^{3/2} \\ &\approx \left(\frac{V_{G'}}{V_p}\right)^{3/2} \left(1 + \frac{3}{2} \frac{V_D}{V_{G'}}\right). \end{aligned}$$

Damit erhält man

$$\begin{aligned} I_D &= \frac{W}{L} I_0 \left[ \frac{V_D}{V_p} - \left(\frac{V_{G'}}{V_p}\right)^{3/2} \frac{V_D}{V_{G'}} \right] \\ &= \frac{W}{L} \frac{I_0}{V_p} \left[ 1 - \sqrt{\frac{V_{G'}}{V_p}} \right] V_D \end{aligned} \quad (54)$$

Der Drainstrom ist also proportional zur Drainspannung. Der Transistor verhält sich wie ein Widerstand. Die Kennlinie im  $I_D(V_D)$  ist in der Nähe des Ursprungs linear. Man nennt diesen Arbeitsbereich den *linearen* oder *ohm'schen Bereich*. Im Englischen spricht man von 'linear' oder 'ohmic region' oder von der *'triode region'*.

linearer Bereich

- Nach (48) schnürt der Kanal an der Drainseite ab, wenn  $V_D = V_p - V_{G'}$ . Mit diesen Werten für die Drainspannung wird aus (53)

$$I_D = \frac{W}{L} I_0 \left[ \frac{1}{3} + \frac{2}{3} \left(\frac{V_{G'}}{V_p}\right)^{3/2} - \frac{V_{G'}}{V_p} \right] \quad (55)$$

Diese Gleichung gibt den Sättigungsstrom als Funktion der Gatespannung an. Wie wir es erwarten fällt der Strom bei steigendem  $V_{G'}$  ab. Für  $V_{G'} = V_p$  wird er Null, der Kanal ist dann selbst für  $V_D = 0$  undurchlässig, der Transistor ist vollständig abgeschaltet. (Die Ableitung von (55) mit  $x := V_{G'}/V_p$  ist  $\sqrt{x} - 1$ , die Steigung an der Stelle vollständiger Abschaltung ist also Null.)

- Der Drainstrom steigt mit der Drainspannung an, bis der Kanal am drainseitigen Ende abschnürt, denn dort ist die Spannungsdifferenz zwischen der Drain und dem (negativen) Gate am höchsten. Bei einer weiteren Erhöhung der Drainspannung ändert sich im Kanal praktisch nichts, die zusätzliche Spannung fällt über dem abgeschnürten Bereich ab. Der Strom steigt daher nicht weiter an, der Transistor ist in *Sättigung* (engl.: saturation).

Sättigung

### 3. Der MOSFET

Das heute wichtigste Bauelement im VLSI Design ist der MOSFET. 'FET' steht für Feld-Effekt-Transistor, was andeutet, daß der Stromfluß im Kanal über ein extern generiertes elektrisches Feld beeinflusst wird (also stromlos). Das 'MOS' steht für 'Metall-Oxid-Semiconductor', was die Gate-Struktur beschreibt. Inzwischen besteht das Gate nicht mehr aus Metall, sondern aus polykristallinem Silizium, so daß man allgemeiner auch oft IGFET für 'Isolated Gate FET' sagt.

MOS

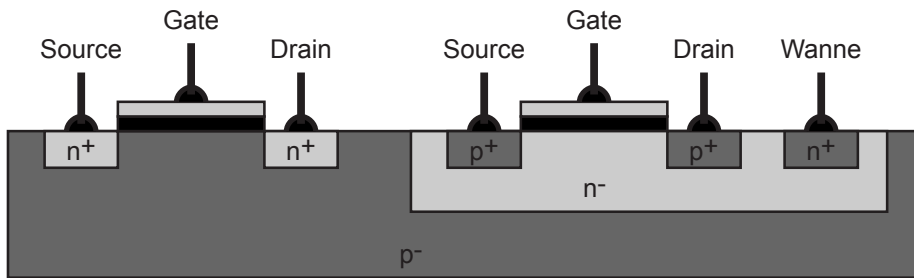


Abbildung 7: Schematischer Aufbau eines n-Kanal MOSFET (links) und eines p-Kanal MOSFET (rechts)

Der Aufbau von MOSFETs ist in Fig. 7 schematisch dargestellt für den Fall, daß das Grundmaterial, das Substrat schwach p-dotiert ist. Der n-Kanal Transistor ('NMOS') besteht aus  $n^+$ -dotierten Drain- und Source-Gebieten, zwischen denen sich der Kanal ausbilden soll. Die Oberfläche zwischen Source und Drain ist durch ein dünnes Oxid (meist  $SiO_2$ ) vom Gate isoliert ('IGFET'), das meist aus dotiertem Polysilizium besteht. Das Gatematerial spielt in unserer weiteren Betrachtung keine Rolle (in der Praxis schon!). Der rechts dargestellte PMOS befindet sich in einem schwach  $n^-$ -dotierten Gebiet, der *n-Wanne* (englisch 'n-well'), die über einen eigenen  $n^+$ -Kontakt angeschlossen ist. Drain und Source sind hier  $p^+$ -Gebiete.

Substrat  
NMOS

PMOS  
Wanne

#### 3.1 Akkumulation, Verarmung, Inversion

Wir betrachten zunächst eine MOS Struktur (Metall - Isolator - Halbleiter) auf p-dotiertem Silizium, wie das beim NMOS der Fall ist. Je nach der Spannung des Gates unterscheidet man drei Bereiche

1. *Akkumulation*: Wenn die Gatespannung negativ ist werden die im Substrat vorhandenen positiv geladenen Löcher unter das Gate gezogen. Löcher sind frei verfügbar, so daß sich sehr leicht viele Löcher (je nach Gatespannung) 'akkumulieren' lassen. Unter dem Gate befindet sich also eine leitende Schicht. Betrachtet man die Struktur als Plattenkondensator, so ist die Kapazität jetzt einfach durch  $C_0 = \epsilon_0 \epsilon_{SiO_2} A / t_{OX}$  gegeben. Da Löcher jederzeit zu- und abgeführt werden können, ist dieser Wert bei hohen und niedrigen Frequenzen

Akkumulation

gültig.

- |   |   |
|---|---|
| <p>2. <i>Verarmung</i>: Bei positiver Gatespannung werden die Löcher von der Grenzschicht weggedrückt (man kann auch sagen Elektronen werden angezogen und besetzen die freien Bindungsstellen der Akzeptor-Atome). Es bildet sich eine <i>Raumladungszone</i> (Verarmungszone) aus, deren Dicke mit steigender Spannung ansteigt. Direkt unter dem Gate sind also <i>keine</i> beweglichen Ladungsträger mehr vorhanden. Die Kapazität sinkt mit steigendem Plattenabstand.</p>  | <p>Verarmung</p> <p>Raumladungszone</p>                           |
| <p>3. <i>Inversion</i>: Bei stark positiver Spannung werden weitere Elektronen unter das Gate gezogen. Sie bilden eine dünne, negativ geladene Schicht mit <i>Minoritätsträgern</i>. Inversion ist erreicht, wenn die Elektronendichte größer ist als die Akzeptordichte. Ab dann sind 'viele' bewegliche Ladungsträger vorhanden. Die zugehörige Gatespannung nennt man die <i>Schwellenspannung</i> <math>V_T</math> ('Threshold Voltage'). Die Kapazität steigt wieder auf den Wert des Plattenkondensators mit dünnem Oxid. Bei Spannungsänderungen am Gate müssen Elektronen zu- und abfließen. Da diese aus dem p-Material nicht einfach beschafft werden können, gilt die hohe Kapazität nur bei <i>niedrigen Frequenzen</i>. Bei hohen Frequenzen bleibt die Kapazität klein (s. Fig. 8).</p> | <p>Inversion</p> <p>Minoritätsträger</p> <p>Schwellenspannung</p> |

Der Transistor wird meist im Bereich *starker Inversion* ('strong inversion') betrieben, also wenn  $V_G > V_T$ . Ein Betrieb in *schwacher Inversion* ('weak inversion') ist jedoch auch möglich und wird im Low-Power-Design zunehmend eingesetzt. Man spricht dann auch von 'Subthreshold Operation'.

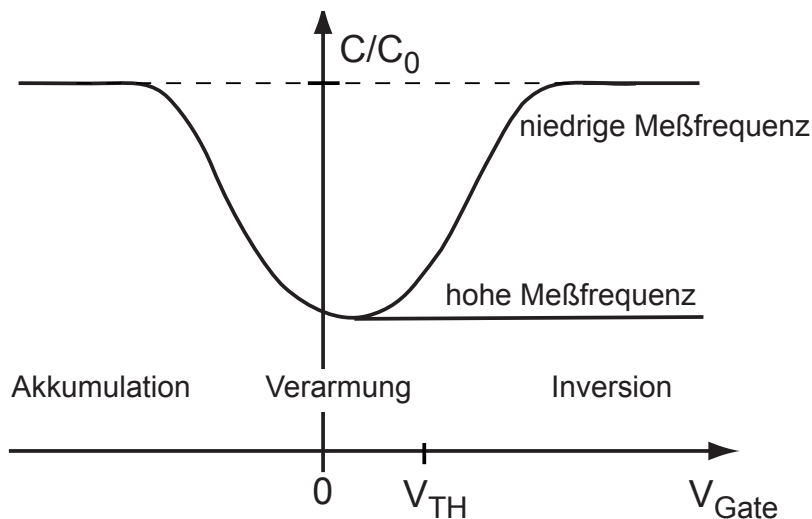


Abbildung 8: Kapazität einer NMOS Struktur als Funktion der angelegten Spannung



### 3.2 Berechnung der NMOS Kennlinie in starker Inversion

Sobald die Gatespannung über der Schwellenspannung liegt bildet sich eine Schicht aus beweglichen Ladungsträgern unter dem Gate. Die Anzahl Elektronen pro Fläche hängt von der Gatespannung und dem Potential im Kanal ab. Für die Herleitung setzen wir wieder das Sourcepotential auf Null:  $V_S = 0 \text{ V}$ . Im Abstand  $x$  von der Source (die Geometrie ist wie in Fig. 5) ist die Flächenladung (Einheit:  $\text{Cm}^{-2}$ )

$$Q(x) = C_{ox}(V_G - V_T - V(x)) \quad (56)$$

Wenn das Potential im Kanal positiver ist als an der Source, so muß also nach (56) eine höhere Gatespannung angelegt werden um den Kanal auszubilden. Die Ladung im Kanal besteht aus beweglichen Minoritäts-Ladungsträgern (hier Elektronen), die den Kanalwiderstand wesentlich beeinflussen. Ein Stück Kanal der Länge  $dx$  hat den Widerstand

$$dR(x) = \frac{dx}{\mu W Q(x)} \quad (57)$$

In diese Gleichung geht auch die Beweglichkeit (Mobilität)  $\mu$  der Ladungsträger und die Breite  $W$  des Kanals ein. Bei einem Kanalstrom von  $I_D$  ist der Spannungsabfall über diesem Stück Kanal

$$\begin{aligned} dV(x) &= I_D dR(x) = \frac{I_D dx}{\mu W Q(x)} \\ \text{oder} \quad I_D dx &= \mu W Q(x) dV(x) \end{aligned} \quad (58)$$

Diesen Ausdruck kann man von 0 bis zum Ende Kanals bei  $L$  integrieren:

$$\begin{aligned} \int_0^L I_D dx &= \mu W \int_0^{V_D} Q(x) dV(x) \\ LI_D &= \mu W C_{ox} \int_0^{V_D} (V_G - V_T - V) dV \\ I_D &= \mu C_{ox} \frac{W}{L} \left[ (V_G - V_T) V \Big|_0^{V_D} - \frac{1}{2} V^2 \Big|_0^{V_D} \right] \\ I_D &= \frac{K W}{2 L} \left[ 2 (V_{GS} - V_T) V_{DS} - V_{DS}^2 \right] \end{aligned} \quad (59)$$

Diese wichtige Gleichung beschreibt den Zusammenhang zwischen Drainstrom und Gate- und Drain-Spannung. In der letzten Zeile wurde der wichtige *Transkonduktanzparameter*  $K := \mu C_{ox}$  eingeführt, der für n-Kanal Transistoren etwa  $K_n = 100 \mu\text{A V}^{-2}$  und, wegen der geringeren Mobilität der Löcher, für p-Kanal Transistoren etwa  $K_p = 40 \mu\text{A V}^{-2}$ . In Technologien mit kurzer Gatelänge wird meist auch das Gateoxid dünner und  $K$  steigt daher.

Für kleine Drainspannungen (genauer für  $V_{DS} \ll V_{GS} - V_T$ ) kann man in (59) den quadratischen Teil vernachlässigen und bekommt so das Verhalten im *linearen Bereich*

$$I_D = K \frac{W}{L} (V_{GS} - V_T) V_{DS} \quad \text{fr} \quad V_{DS} \ll V_{GS} - V_T \quad (60)$$

linearer Bereich

Der Transistor verhält sich also wie ein gesteuerter Widerstand mit

Widerstand

$$R = \frac{V_{DS}}{I_D} = \frac{1}{K \frac{W}{L} (V_{GS} - V_T)} \quad (61)$$

Der Widerstand wird unendlich, wenn  $V_{GS}$  die Schwellenspannung erreicht. Gleichung 59 ist nur gültig, solange der Kanal noch nicht abgeschnürt ist. Dies passiert, wenn an der Drain keine Inversionsschicht mehr existiert, also nach (56) wenn

$$V_D = V_{DSat} = V_G - V_T \quad (62)$$

Setzt man das in (59) ein, so erhält man den Drainstrom in Sättigung

$$I_D = \frac{K W}{2 L} (V_{GS} - V_T)^2 \quad (63)$$

Dieser Strom fließt näherungsweise auch für höhere Drainspannungen  $V_D > V_{DSat}$ , da diese über der Abschnürregion abfallen und die Verhältnisse im Kanal sich nur wenig ändern. In Realität wächst die Abschnürregion mit steigender Drainspannung jedoch langsam an und der Kanal wird effektiv kürzer. Der Drainstrom steigt dadurch leicht an. Diese *Kanallängenmodulation* führt zu einem endlichen Ausgangswiderstand im *Sättigungsbereich*. Er ist um so höher ('besser'), je länger der Transistor ist (großes  $L$ ), da sich dann die Vergrößerung der Abschnürregion weniger auswirkt. Man parametrisiert die Kanallängenmodulation in meist guter Näherung durch einen multiplikativen Anteil:

$$I_D = \frac{\mu C_{ox} W}{2 L} (V_{GS} - V_T)^2 (1 + \lambda V_{DS}) \quad (64)$$

Die unterschiedlichen Betriebsbereiche eines NMOS sind in Fig. 9 dargestellt. Dort ist auch angedeutet, daß die Verlängerungen aller Sättigungskurven zu negativen Drainspannungen hin die x-Achse bei  $-1/\lambda$  schneiden. Man nennt diese Spannung in Anlehnung an den Bipolartransistor auch die 'Early-Spannung'.

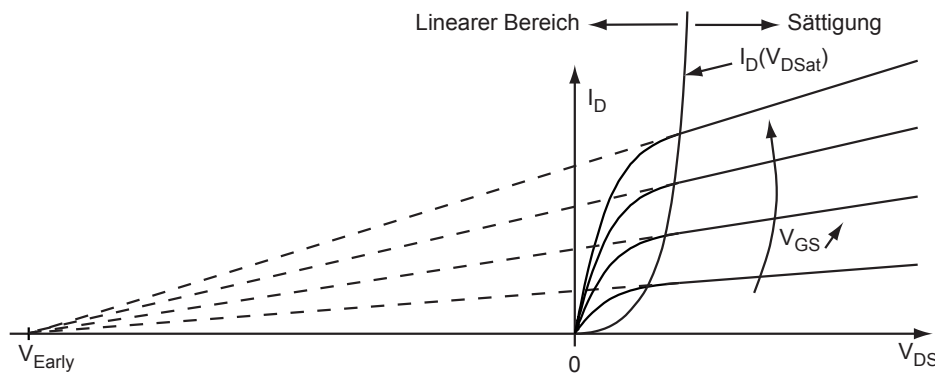


Abbildung 9: Idealisiertes Ausgangs-Kennlinienfeld eines n-Kanal MOSFET

Wir sind bisher davon ausgegangen, daß Source und Substrat des NMOS auf Masse liegen. In vielen Schaltungen ist das Substrat (der 'Bulk') jedoch negativer also die Source (genauer: das Substrat liegt auf Masse und die Source ist positiver.) Die negative Spannung am p-Substrat vergrößert die 'Verarmungszone' unterhalb des Kanals und der Strom wird etwas kleiner. Man kann dies durch einen JFET modellieren, der parallel zur Kanal liegt. Es gibt verschiedene Formeln, die diesen *Substrat-Effekt* (engl. 'body effect') beschreiben. Man kann ihn z.B. durch eine

Änderung der Schwellenspannung modellieren:

$$V_T = V_{T0} + \gamma \sqrt{2\Phi_F - V_{BS}} - \sqrt{2\Phi_F} \quad (65)$$

Hier ist  $V_{T0}$  die Schwellenspannung für  $V_{BS} = 0$  V, d.h. für  $V_S = V_G$ .  $\Phi_F$  ist der Abstand des Fermi-Niveaus des dotierten Halbleiters vom intrinsischen Fermi-Niveau des undotierten Halbleiters ( $\Phi_F$  hängt also nur von der Substratdotierung ab).  $\gamma = \sqrt{2N_A q \epsilon_{Si} / C_{ox}}$  ist der 'body factor', der die Sensitivität der Schwelle von der Substratspannung beschreibt. Die effektive Schwellenspannung steigt an, wenn das Substrat negativer wird. (Dies ist z.B. der Grund, weswegen die Verstärkung eines Source-Folgers oft wesentlich kleiner als Eins ist.)

### 3.3 Subthreshold Betrieb

Bisher sind wir davon ausgegangen, daß unterhalb der Schwellenspannung kein Drainstrom fließt. Im 'Subthreshold' Bereich ist die Inversion noch nicht stark ausgeprägt, man spricht daher auch von schwacher Inversion ('weak Inversion') im Gegensatz zur bisher behandelten starken Inversion ('strong inversion'). In diesem Bereich fließt ein kleiner Strom, der exponentiell von der Gatespannung abhängt. Der Ausdruck für den Drainstrom lautet

$$I_D = I_{D0} \frac{W}{L} e^{\frac{V_{GS}}{nV_{th}}} \left( 1 - e^{-\frac{V_{DS}}{V_{th}}} \right) e^{\frac{n-1}{n} \frac{V_{BS}}{V_{th}}} \quad (66)$$

Hier ist explizit auch die Abhängigkeit vom Substratpotential beschrieben.  $I_{D0}$  ist ein technologieabhängiger Parameter.  $V_{th} = kT/q$  ist die Temperaturspannung. Der Parameter  $n$  liegt in der Größenordnung von 1.5. Er bestimmt unter anderem wie steil die  $\ln[I_D(V_{GS})]$ -Kurve ist und wird daher 'subthreshold slope factor' genannt. Es ist bemerkenswert, daß in schwacher Inversion die Sättigungsspannung nicht von der Gatespannung abhängt und sehr niedrig ist. Sie beträgt nur wenige  $V_{th}$ , also  $< 200$  mV.