

Vorlesung 2

In dieser Vorlesung werden folgende Themen behandelt

Herstellung des MOSFETs

Funktionsweise des MOSFETs und die Herleitung der Stromgleichung

Einfluss der vertikalen E-Feld Komponente: Kanalladung als Funktion von Gate-Spannung

Einfluss der horizontalen E-Feld Komponente: Drain-Strom Strom als Funktion von Drain-Source Spannung

Sättigung des Stromes

Herstellung des MOSFETs

Abbildung 1 zeigt das 3D Bild eines N-Kanal Metal-Oxid-Semiconductor Feldeffekt-Transistors (N-MOSFETs) oder kurz NMOS Transistors.

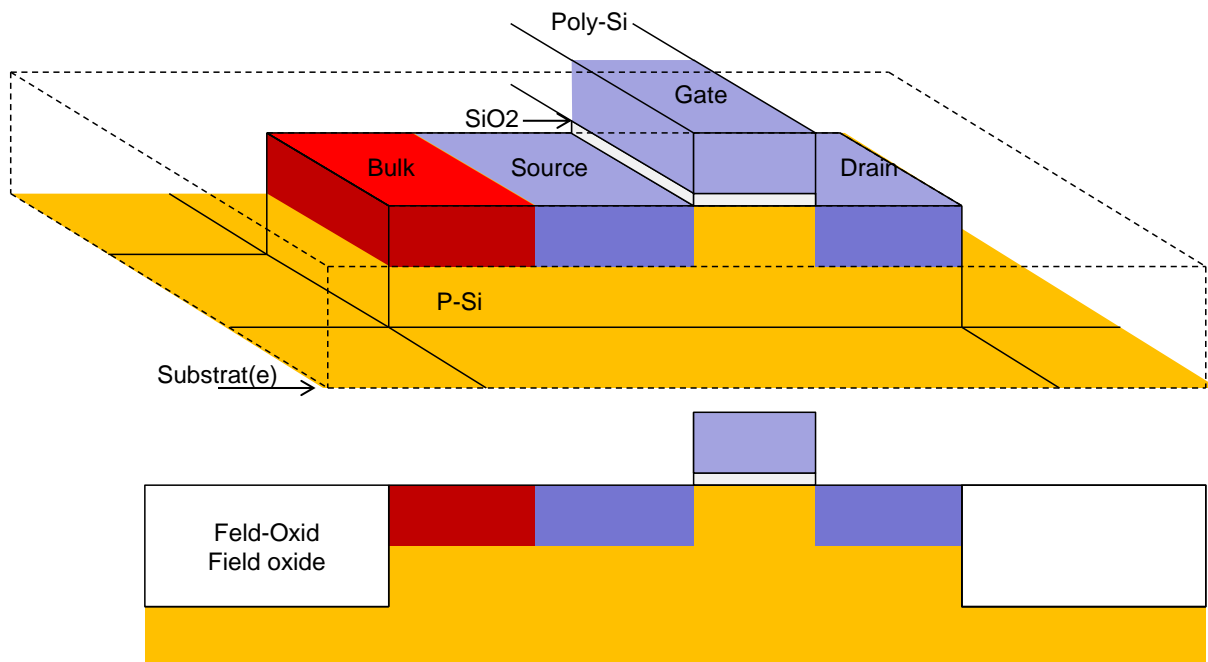


Abbildung 1: 3D-Darstellung eines NMOS Transistors

Ein MOSFET Transistor enthält vier Elektroden: Source, Drain, Gate und Substrat (Bulk). Die Source ist die Quelle für die freien Ladungsträger (NMOS: Elektronen, PMOS: Löcher) und der Drain sammelt sie. Mithilfe von Gate-Elektrode kann man Transistorstrom verändern. Source und Drain befinden sich im Substrat. Das Substrat kontaktiert man über „Bulkkontakt“. Im Fall vom NMOS, ist das Substrat P-dotiert. Source, Drain und Gate sind N-dotiert.

In diesem Kurs betrachten wir die Planartransistoren in einem Bulkprozess, wie sie z.B. in einem 65 nm Prozess realisiert werden. Neuere Chiptechnologien verwenden andere Transistorarten, wie z.B. FINFET.

Ein Transistor befindet sich in einer isolierenden Umgebung. Sie wird als eine Silizium-Dioxid- (SiO_2 -)Schicht realisiert – das *Feldoxid*.

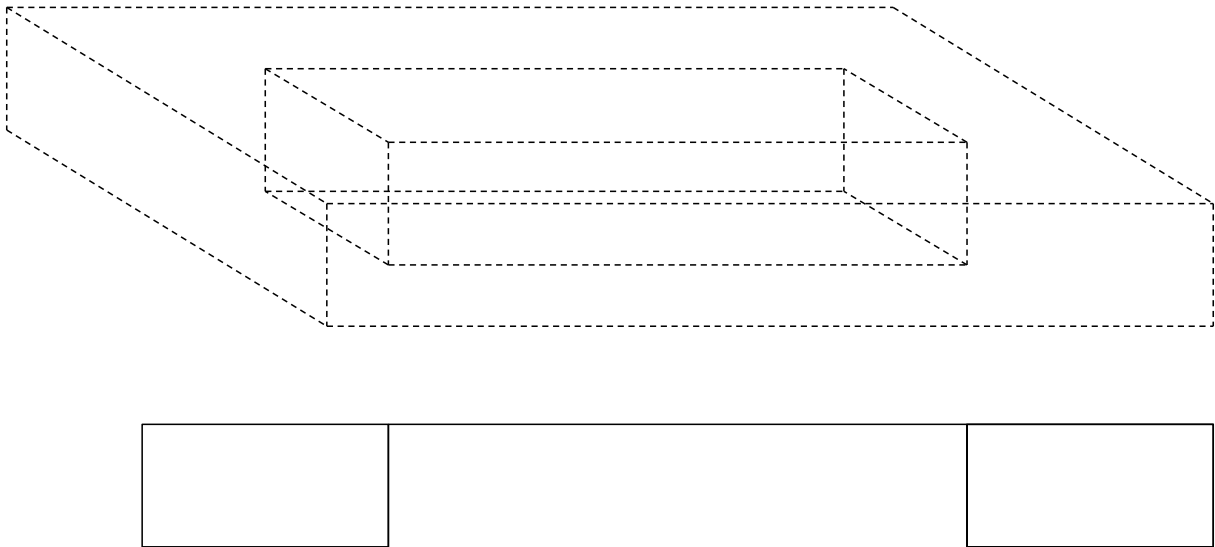


Abbildung 2: Transistor befindet sich im Feldoxid

Das Feldoxid soll die einzelnen Transistoren voneinander isolieren und die Entstehung von parasitären Transistoren verhindern.

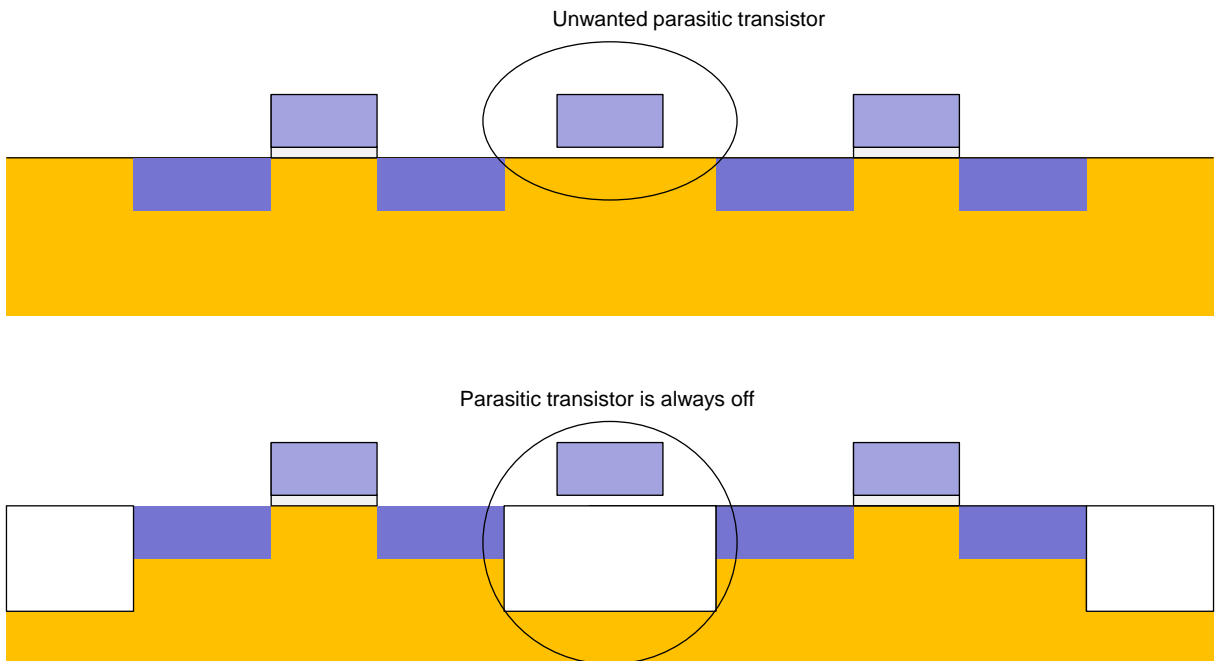


Abbildung 3: Isolierung von Transistoren

Weitere Abbildungen zeigen die Reihenfolge von Prozessschritten bei der Herstellung eines Transistors.

Eine Maskenlage (genannt Diffusionsmaske) definiert die aktiven Silizium-Bereiche wo sich später Transistor befinden wird.

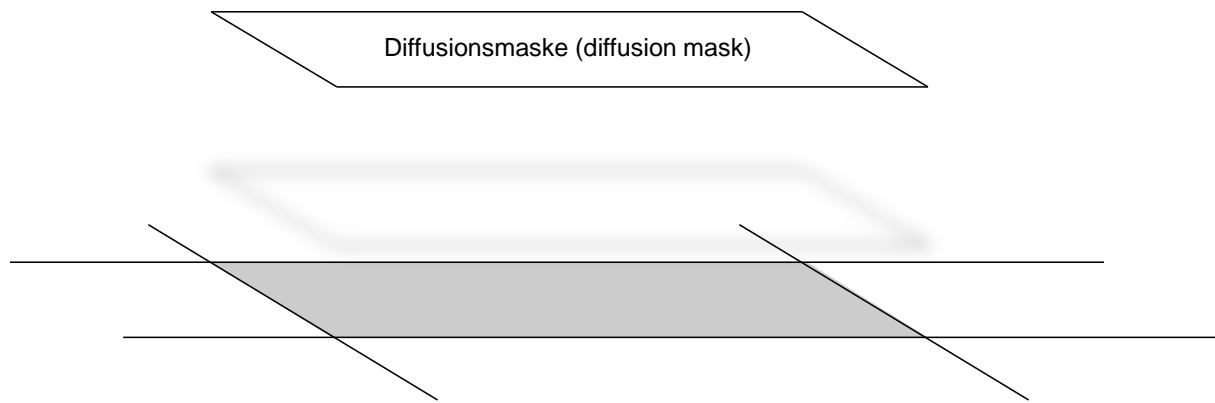


Abbildung 4: „Diffusionsmaske“ definiert die Stelle wo sich Transistor befinden wird

Der erste Schritt in MOSFET-Fertigung ist die Erzeugung von trenches für das Feldoxid.

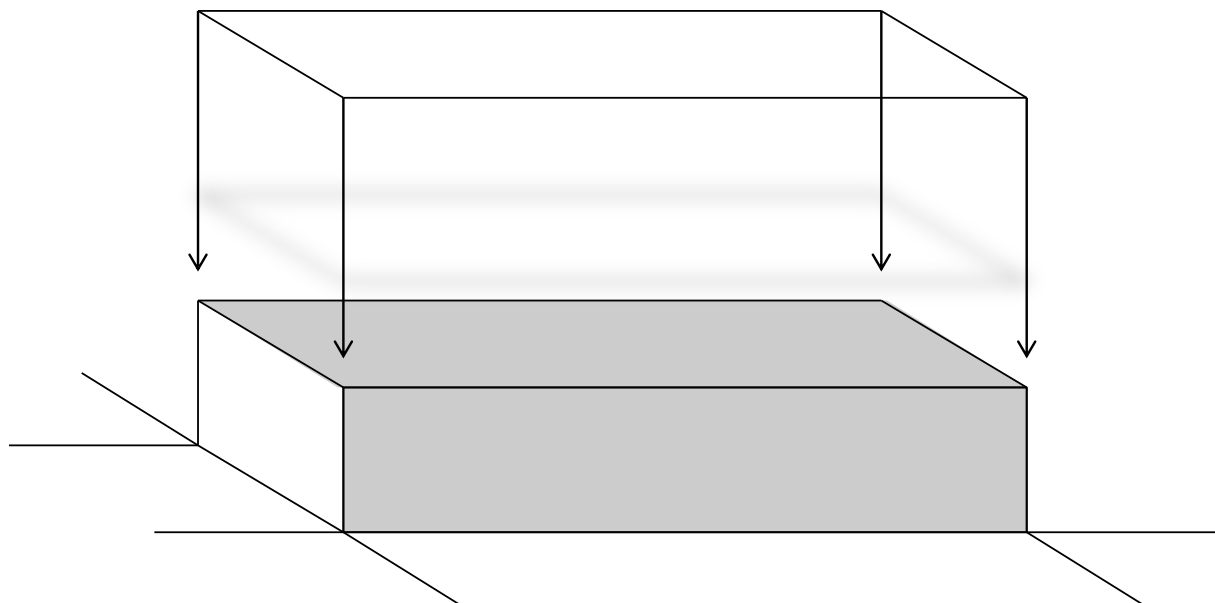


Abbildung 5: Erzeugung von trenches

Ätzverfahren (z.B. reaktives Ionenätzen) werden benutzt. Diese trenches (Gräben) werden mit Silizium-Dioxid gefüllt. Dafür wird CVD Verfahren (chemische Dampfabcheidung) verwendet. Der Wafer reagiert nicht mit den Gasen. Eine detailliertere Beschreibung finden Sie hier: <https://de.wikipedia.org/wiki/Grabenisolation>

Im nächsten Schritt werden die lokalen Substrate (Wannen) erzeugt. Für NMOS brauchen wir eine P-Wanne und für PMOS (P-Kanal MOSFET) eine N-Wanne.

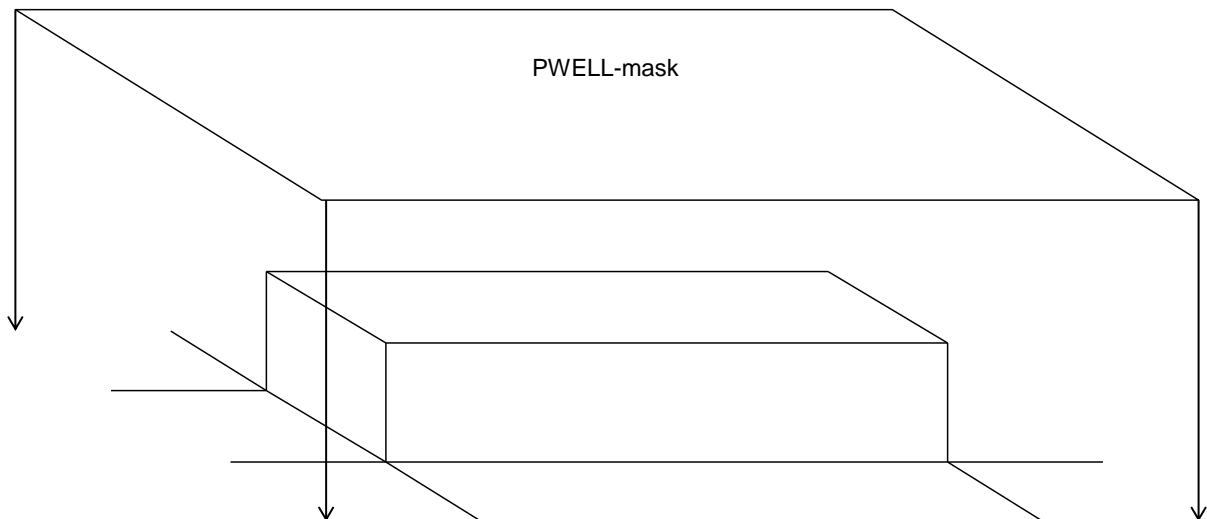


Abbildung 6: Ein NMOS befindet sich in einem lokalen P-Typ Substrat (P-Wanne)

Die Wannen werden mithilfe von Ionenimplantern erzeugt.

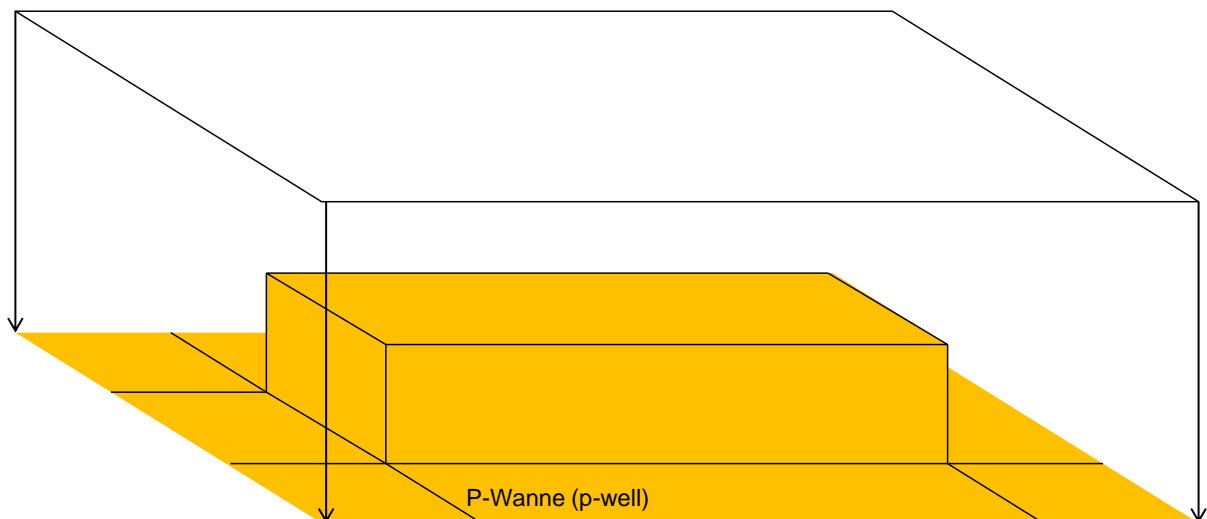


Abbildung 7: P-Wanne

N-Siliziumbereiche können mit Phosphor oder Arsen dotiert werden, P-Bereiche mit Bor. Die Fremdatome (Ionen) dringen in Siliziumsubstrat durch und haben eine relativ konstante Reichweite, die von kinetischer Energie der Ionen abhängig ist. Die Dotierungsdichte hat deswegen ein gut definiertes Maximum. Bei der Implantation entstehen in Abhängigkeit von der Masse der implantierten Ionen und der Implantationsdosis Strahlenschäden im Kristallgitter des Halbleiters. Daher muss das Substrat nach einem Implantationsschritt ausgeheilt werden. Das erreicht man durch hohe Temperatur. <https://de.wikipedia.org/wiki/Ionenimplantation>

Dünnes Gateoxid wird auf dem aktiven Bereich durch thermische Oxidation erzeugt. Neuere Technologien verwenden kein SiO_2 sondern ein Dielektrikum mit größerer relativen Permittivität z.B. HfO_2 .

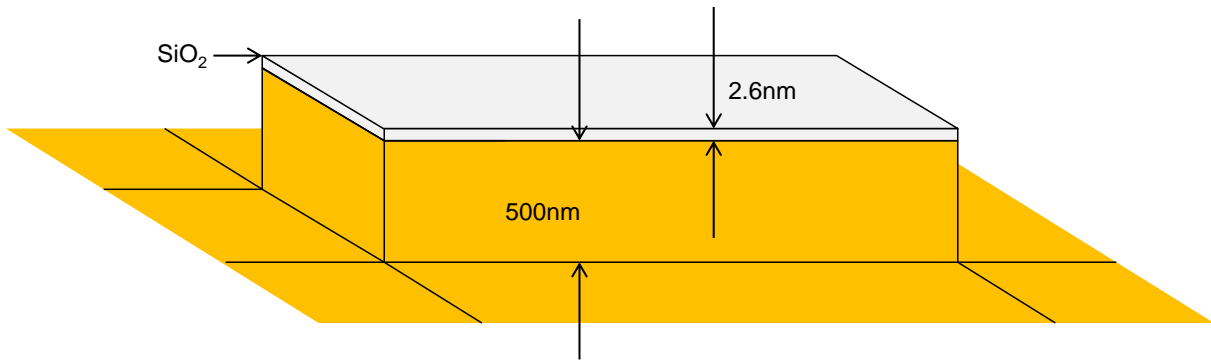


Abbildung 8: Thermische Oxidation

Eine dünne und hochwertige Oxidlage ist für gute elektrische Eigenschaften der Transistoren entscheidend. Die Oxidkapazität beträgt etwa $13 \text{ fF}/\mu\text{m}^2$. Oxidkapazität bestimmt die Schwellenspannung und die Transkonduktanz des Transistors. Typische Dicke vom Gateoxid beträgt 2.6 nm (65 nm Technologie). Das entspricht nur etwa 5 Atomlagen, da die Gitterkonstante von SiO₂ etwa 0,5 nm ist.

In einem weiteren Schritt wird die Gate-Elektrode erzeugt. Das Gate besteht aus polykristallinem Silizium, dass mittels LPCVD (Niederdruck-Dampfabscheidung) erzeugt wird.

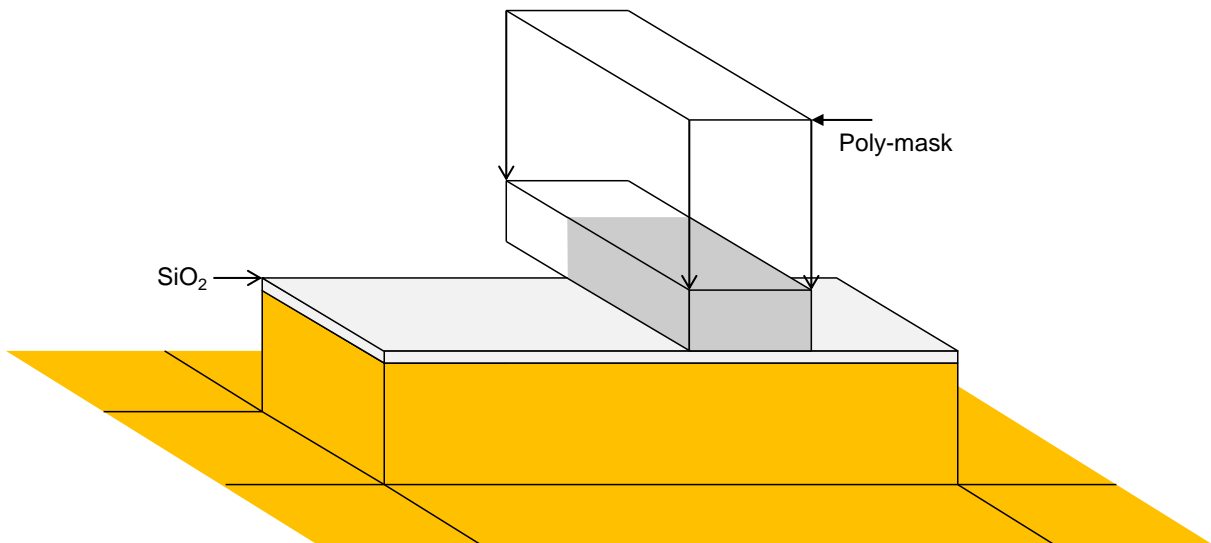


Abbildung 9: Polysilizium Maske definiert die Gate-Elektrode und die Polysilizium-Leitung

Die Polysilizium Strukturen werden durch Ätzen (nasschemisch, Plasma- und Ionenätzen) strukturiert. Fotolack wird verwendet um die Bereiche die nicht geätzt werden sollen zu schützen - Lithographie. Fotolack wird mit UV Licht strukturiert (belichtet und dann entwickelt), die Schablone für diese Strukturierung nennen wir Maske.

Die Polysilizium-Maske bestimmt die Stellen wo Polysilizium nach bleibt. Das dünne Oxid bleibt nur im Überlappbereich von Diffusions- und Polysilizium-Maske.

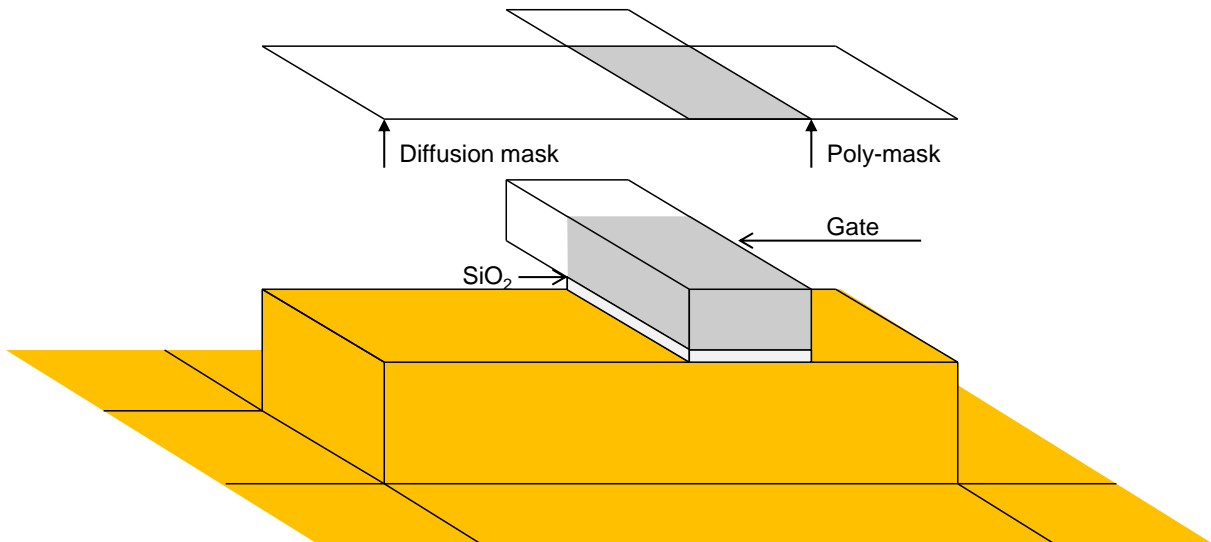


Abbildung 10: Gate-Oxid

Warum wird Silizium für die Gate-Elektrode verwendet und nicht Metall (z.B. Aluminium)?
Dafür gibt es drei Gründe:

1. Gate-Elektrode aus Polysilizium kann dotiert werden – dadurch kann die Schwellenspannung eingestellt werden.
2. Polysilizium-Gate kann als Maske für nachfolgende Dotierung von Source und Drain verwendet werden – der Prozess ist „selbstjustierend“.
3. Silizium hat deutlich höhere Schmelztemperatur als Aluminium, deswegen können die nachfolgenden Schritte bei höheren Temperaturen erfolgen – z.B. die Dotierung von Source und Drain durch Diffusion.

Nach der Herstellung von Gate folgt die Dotierung von Source-, Drain- und Substratkontakten des Transistors. Hier werden zwei Verfahren benutzt – Diffusion und Ionenimplantation

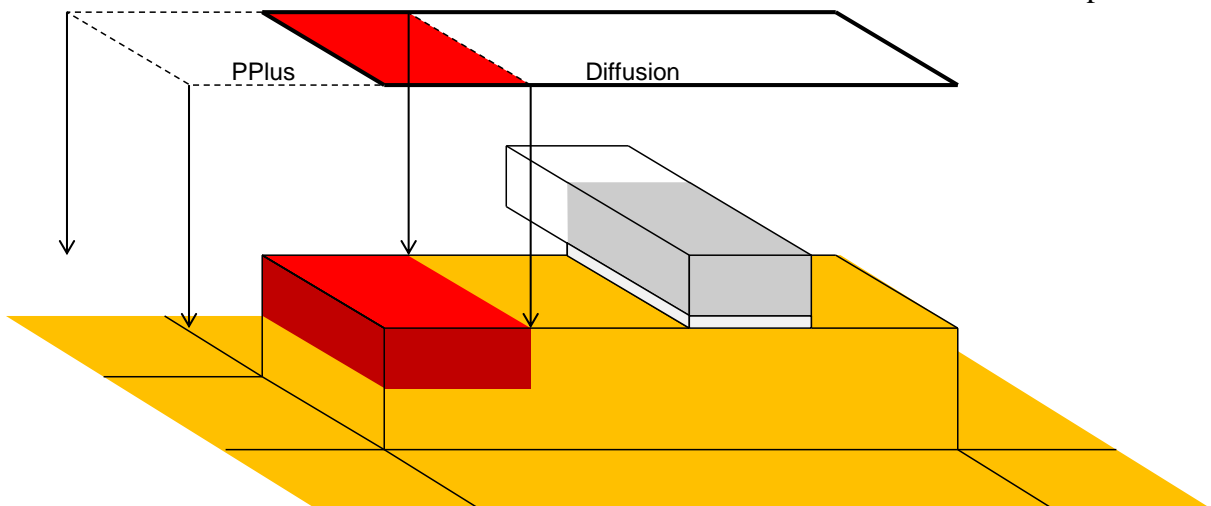


Abbildung 11: P-Puls Maske. Überlapp von Diffusions- und P-Plus Masken wird P-dotiert

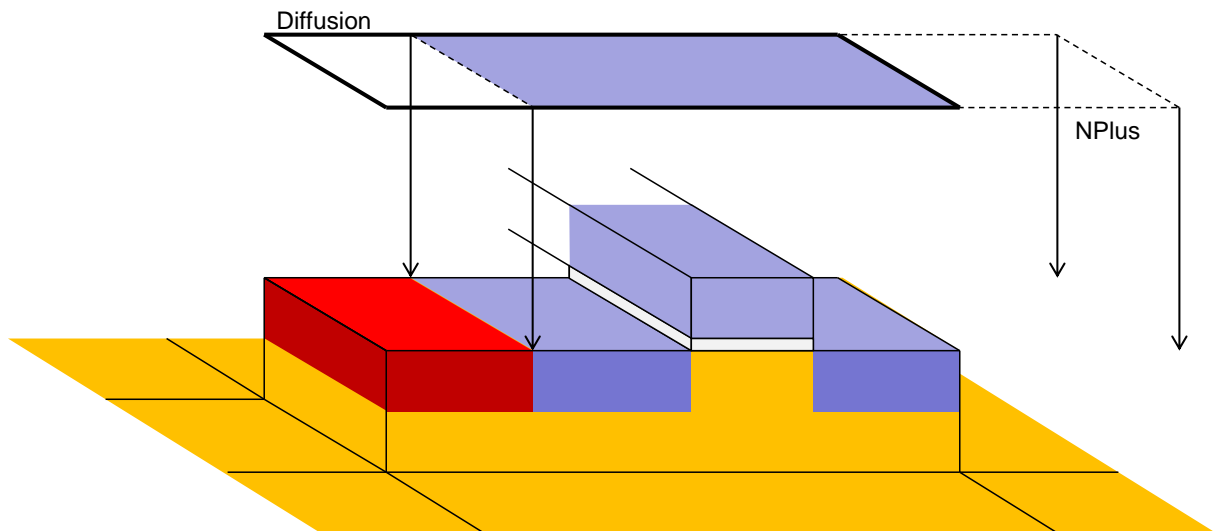


Abbildung 12: N-Puls Maske -> Überlapp von Diffusions- und N-Plus Masken wird N-dotiert (auch Poly-Gate)

Eine Maske („P/N-Plus Maske“), die Gate-Elektrode und das Feldoxid dienen, alle zusammen, als Masken für die Dotierung.

Damit ist die Herstellung von Grundstrukturen des Transistors abgeschlossen. Eine detailliertere Darstellung findet man hier: <https://www.halbleiter.org/chipfertigung/>

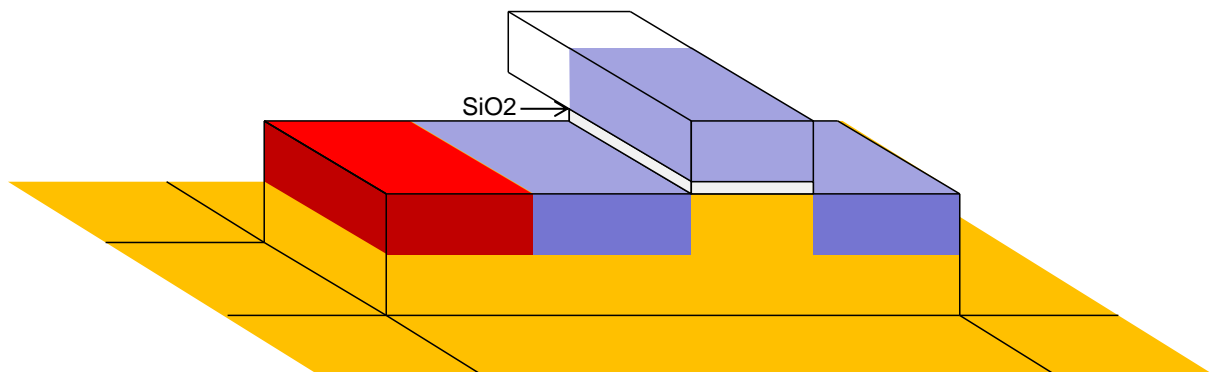


Abbildung 13: Vollständiger Transistor

Funktionsweise des MOSFETs

Ein MOSFET Transistor enthält vier Elektroden: Source, Drain, Gate und Substrat (Bulk).

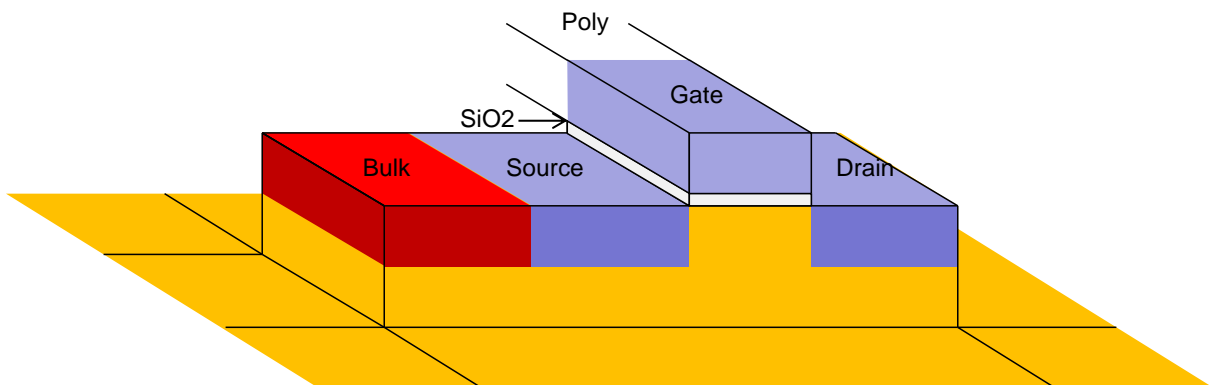


Abbildung 14: Vier Elektroden: Source, Drain, Gate und Substrat (Bulk)

Die Source ist die Quelle für die freien Ladungsträger (NMOS: Elektronen, PMOS: Löcher) und der Drain sammelt sie. Mithilfe von Gate-Elektrode kann man Transistorstrom verändern.

Source und Drain befinden sich im Substrat. Das Substrat hat einen eigenen Kontakt – „Bulkkontakt“.

Einen PMOS bekommt man indem man alle Dotierungen austauscht (N -> P, P -> N). Ein PMOS befindet sich in einem N-Typ Substrat. In Wirklichkeit befinden sich PMOS- und NMOS-Transistoren auf demselben Wafer. Die Transistoren befinden sich üblicherweise in den lokalen Substraten - „Wannen“.

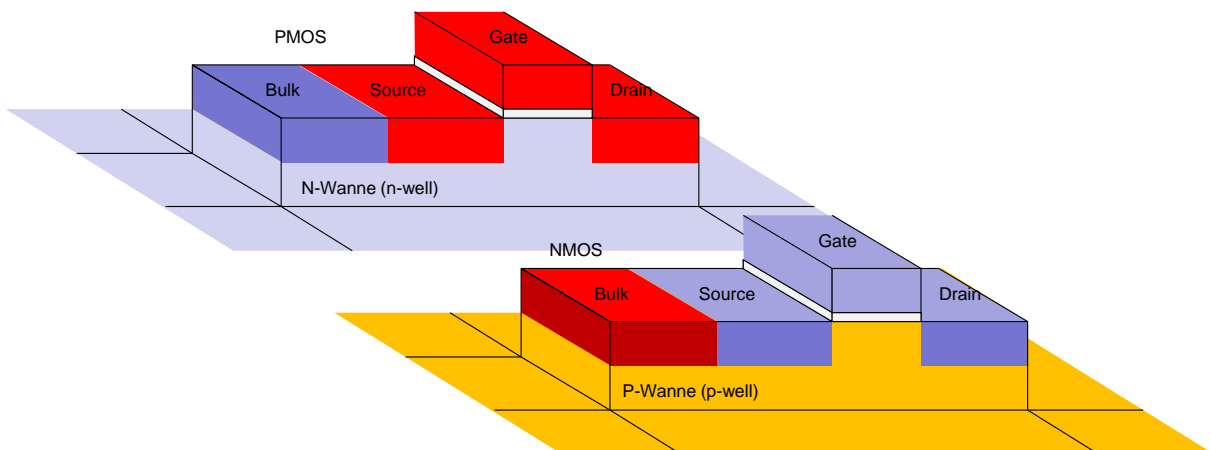


Abbildung 15: NMOS und PMOS

In diesem Kurs werden wir die vereinfachten Transistorsymbole (die Schalter-Symbole) verwenden, mit oder ohne Substratelektrode. Diese Symbole sind symmetrisch, wie die Transistorstruktur selbst.

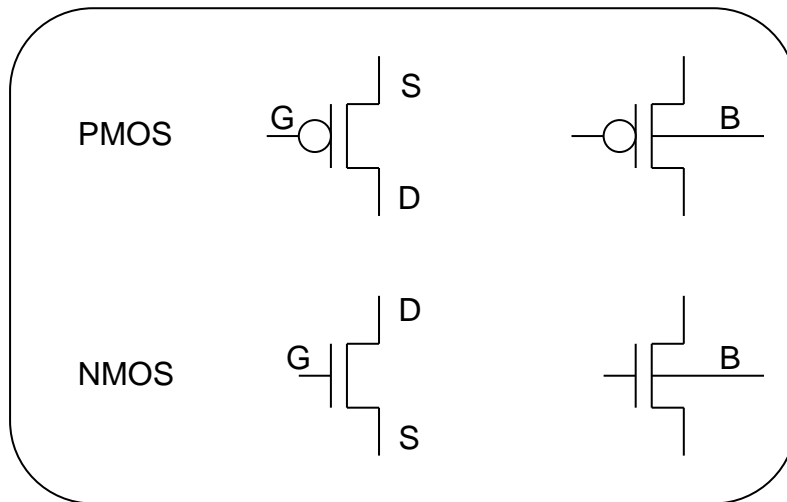


Abbildung 16: Vereinfachte Transistor-Symbole

Wie erkennen wir Source und Drain?

Im Fall vom NMOS ist die Source die Elektrode auf niedrigerem Potential. Im Fall vom PMOS, ist die Source auf höherem Potential.

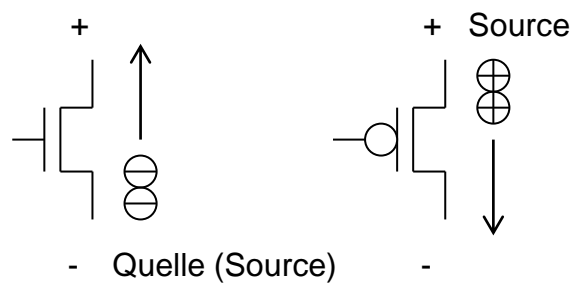


Abbildung 17: Links: NMOS-Source, rechts PMOS-Source

Alternativ kann man die asymmetrischen Symbole mit dem Pfeil benutzen. Wenn die Substratelektrode im Symbol fehlt, ist sie entweder an Source oder an eine feste Spannung angeschlossen.

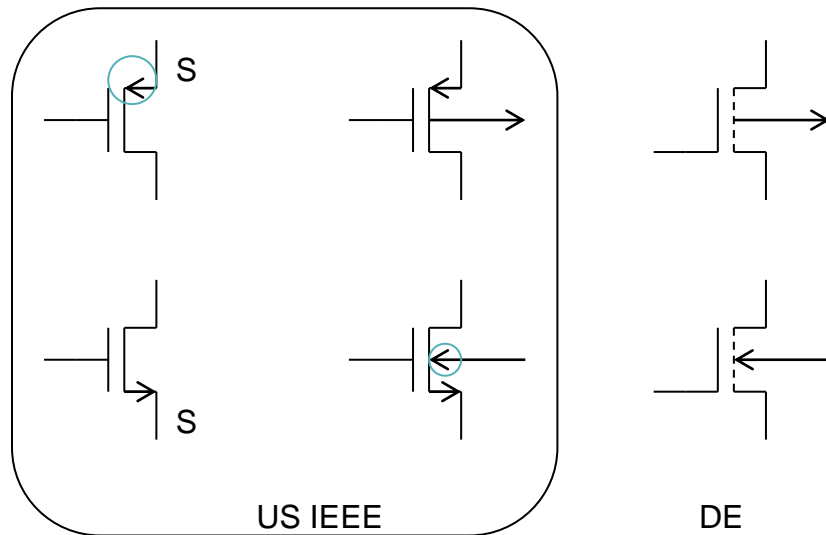


Abbildung 18: Symbole mit Pfeil

Wie kann man sich die Pfeilrichtung merken: Der Pfeil beim Substratkontakt zeigt vom P- zu N-Bereich. Im Fall von NMOS bedeutet es: Vom P-Substrat zum N-Kanal. Es ist ähnlich wie bei einer PN-Diode. Ihr pfeilförmiges Symbol zeigt die Stromrichtung, wenn die Diode in Durchlassrichtung gepolt ist. Der Strom fließt dann von P- zum N-Teil.

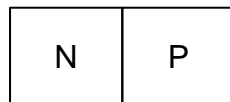
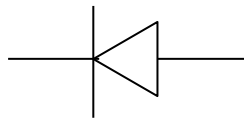


Abbildung 19: Der Pfeil im Symbol einer Diode zeigt von P zu N

Die Funktionsweise vom MOSFET wurde in der Vorlesung „Elektronische Schaltungen“ erklärt – wir werden hier das Wichtigste zusammenfassen und einige Spezialeigenschaften von kleinen MOSFETs beschrieben.

Betrachten wir einen NMOS. In der Struktur haben wir zwei PN Dioden: Source/Substrat und Drain/Substrat. Das Substrat-Potential muss so gewählt werden, dass beide Dioden in Sperrrichtung gepolt werden. Sonst funktioniert ein MOSFET nicht richtig. Also, im Fall von NMOS muss das Substrat auf niedrigerem Potential als Source und Drain liegen. In solch einem Zustand fließt kein Strom zwischen dem Drain und der Source, falls die Gate-Source Spannung null ist.

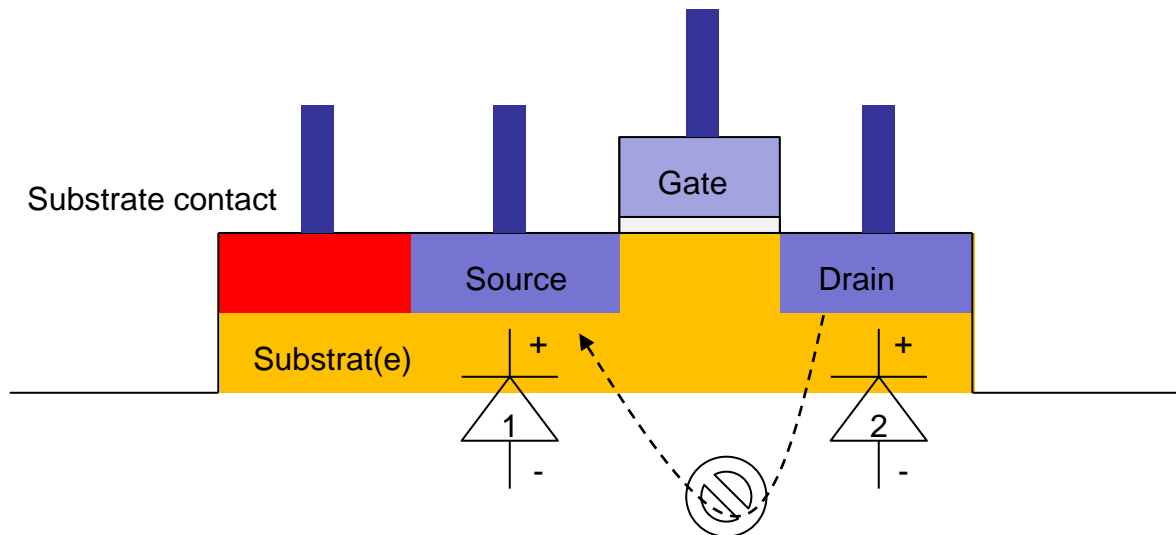


Abbildung 20: Zwei PN-Dioden (Source-Substrat und Drain-Substrat). Die Dioden müssen in Sperrichtung gepolt werden

Kontaktspannungen

Zwischen P- und N-Silizium und zwischen Silizium und Metallelektroden entstehen Kontaktspannungen.

Um die Analyse zu vereinfachen, werden wir annehmen, dass sich das Metall gleich wie das N-dotierte Silizium im Source/Drain/Gate verhält. (Das Metall – P-Substrat Übergang hat in unserer Annahme genau die gleiche Kontaktspannung hat wie der N-Source – P-Substrat Übergang). Wir werden ebenfalls annehmen, dass die N-Source, N-Drain und N-Gate gleiche Dotierungsdichte wie P-Substrat (Kanalbereich) haben. Das ist in Wirklichkeit nicht ganz richtig, aber es ändert die Ergebnisse der Analyse wenig. In unserem Fall heben sich die Kontaktpotentiale auf, und in Wirklichkeit nur teilweise.

Unter dieser Annahme gebe es zwischen den N-dotierten Bereichen – den Source, Drain und Gate (Gate ist, wie Source und Drain, n-dotiert) und dem Metall keine Kontaktspannung.

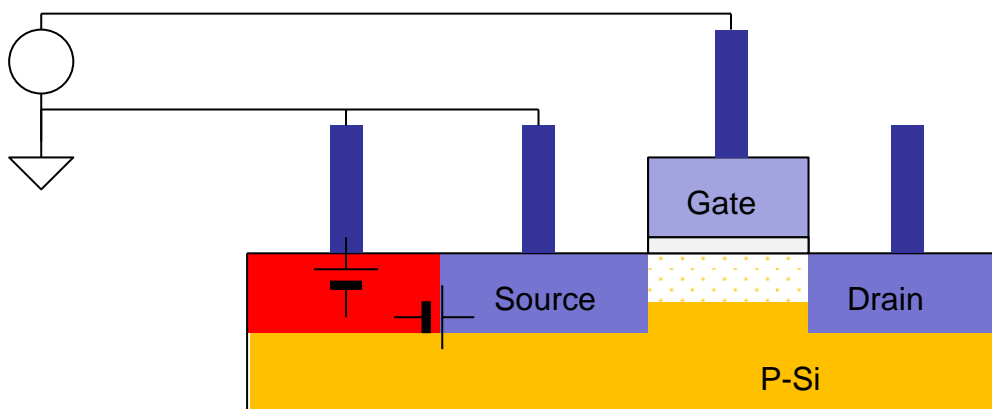


Abbildung 21: Zwischen Silizium und Metallelektroden entstehen Kontaktspannungen

Warum entsteht die Kontaktspannung?

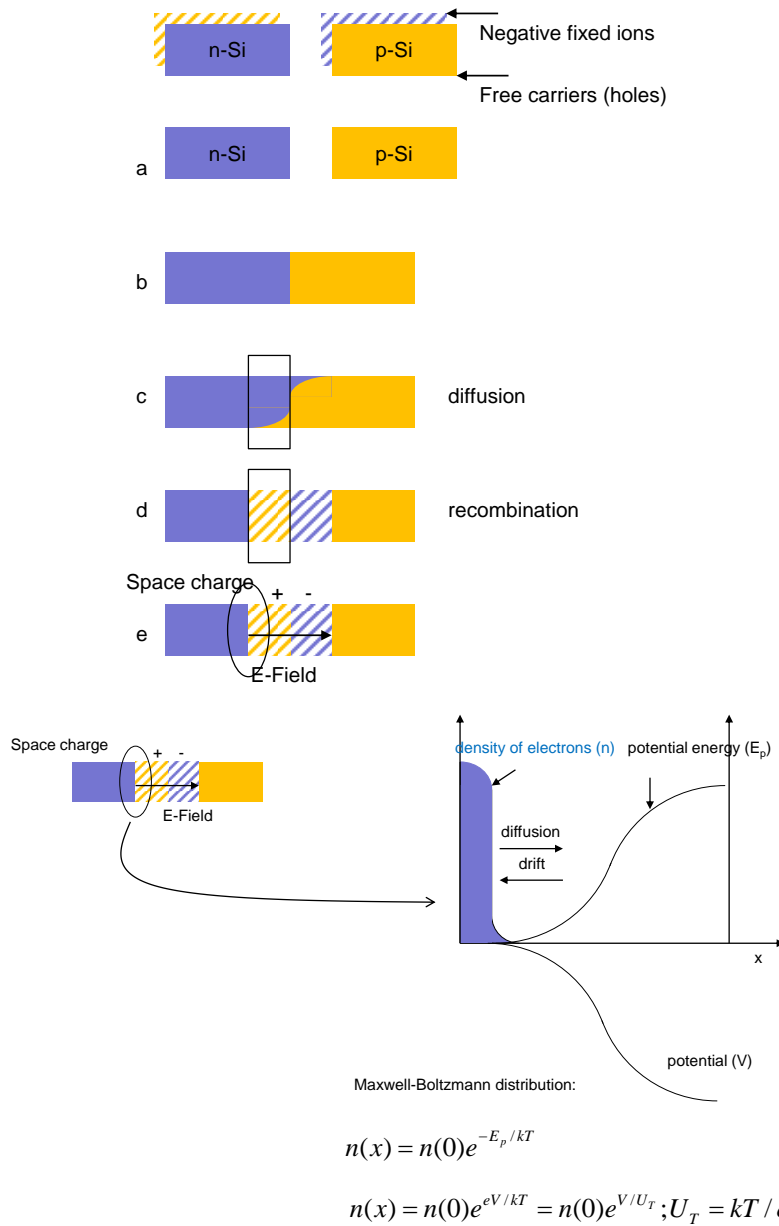


Abbildung 22: Entstehung einer Kontaktspannung

Vereinfachte Erklärung (Abbildung 22): Betrachten wir zuerst die Situation aus der Sicht der Elektronen. Das gleiche gilt für Löcher. Im N-Silizium und im Metall ist die Dichte der Elektronen höher als im P-Silizium. Wenn wir N-Silizium und Metall mit P-Silizium „verbinden“ (b), entsteht ein Diffusionsstrom von Elektronen in Richtung zum P-Silizium (c). Elektronen und Löcher im P-Silizium rekombinieren. Die Ladung von negativen Akzeptorionen ist dann nicht mehr durch die positive Ladung der Löcher kompensiert und entsteht eine negativ geladene Zone (d). Die Löcher diffundieren in N-Silizium und rekombinieren dort mit den Elektronen. Die Ladung von positiven Donatorionen ist nicht mehr durch die negative Ladung der Elektronen kompensiert und entsteht eine positiv geladene Zone.

Es entsteht das E-Feld und die Kontaktspannung (e). Das E-Feld verursacht einen Driftstrom der den Diffusionsstrom kompensiert – es entsteht Gleichgewichtszustand. Wenn sich der Drift- und der Diffusionsstrom kompensieren, haben wir Gleichgewichtszustand. Die Ladungsträgerdichte ist innerhalb eines Energiebandes durch Maxwell-Boltzmann Formel beschrieben. Die Elektronen sind eigentlich Fermi-Dirac-verteilt, diese Verteilung kann durch Maxwell-Boltzmann-Verteilung innerhalb eines Energiebandes approximiert werden. Die Elektronendichte im N-Silizium n_n ist etwa gleich die die Dichte von Donatoratomen N_d Die Elektronendichte im P-Silizium n_p ist durch folgende Gleichung gegeben $n_p = n_i / N_a$.

Mithilfe von Maxwell-Boltzmann Formel kann man die Kontaktspannung als Funktion von n_n und n_p berechnen:

$$V = U_T \ln\left(\frac{n_n}{n_p}\right) = U_T \ln\left(\frac{N_a N_d}{n_i^2}\right); n_i \sim e^{\frac{-E_g}{kT}}$$

Eigenleitungsichte n_i steigt mit Temperatur. Deswegen wird die Kontaktspannung kleiner wenn die Temperatur steigt. Die Kontaktspannungen modellieren wir mit Spannungsquellen.

Eigenleitungsichte $n_i = 10^{10}/\text{cm}^3$ auf 300 K, Silizium-Atomdichte: $n_{\text{si}} \sim 5 \times 10^{22}/\text{cm}^3$.

Tunneleffekt-Kontakt

Ein Metall-Silizium Kontakt ist normalerweise eine Schottky-Diode. Der Strom kann nur in eine Richtung fließen – und das nur wenn die externe Spannung die Potentialbarriere verringert.

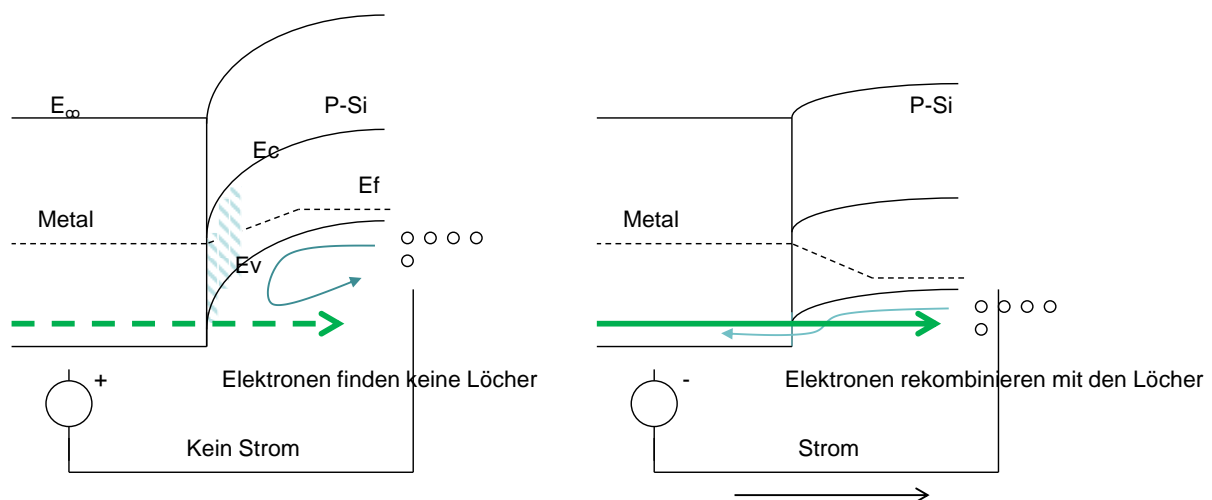


Abbildung 23: Schottky-Diode

Es wird ein Trick verwendet um einen normalen Kontakt herzustellen. Wenn das Silizium hochdotiert ist, ist die Potentialbarriere im Halbleiter sehr schmal und die Ladungsträger können durch die Barriere in beide Richtungen „tunneln“. (Quantenmechanischer

Tunneleffekt) Der Metall-Silizium-Kontakt ist in dem Fall in beide Richtungen leitend – es ist ein „Ohmscher Kontakt“.

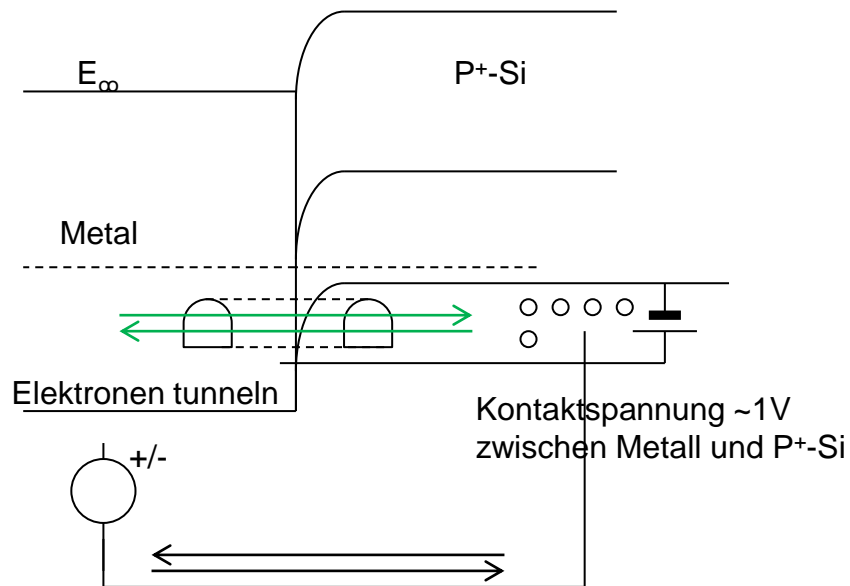


Abbildung 24: Ohmscher Kontakt

Der Bulkkontakt eines MOSFETS muss als Ohmscher Kontakt realisiert werden. Deswegen wird das Silizium-Teil des Bulk-Kontakts zusätzlich P⁺-dotiert und danach ein Metallkontakt gemacht. Auch ein Ohmscher Kontakt hat eine Kontaktspannung.

Folgende Abbildung gibt die Zusammenfassung der wichtigen Formeln für Halbleitertechnik.

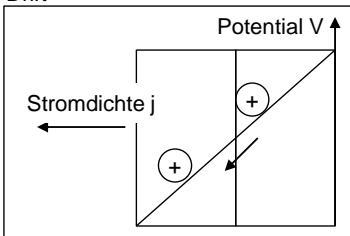
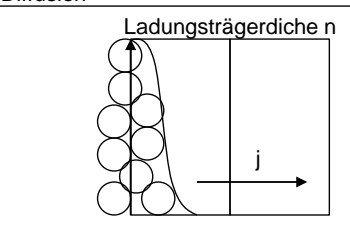
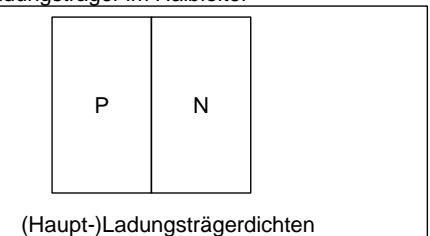
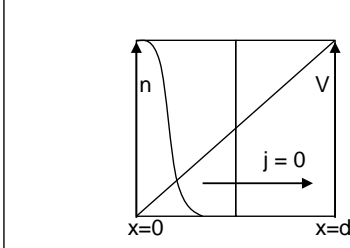
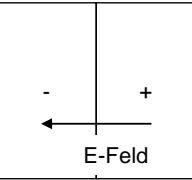
<p>Drift</p>  <p>Driftstrom $j = q n\mu E$</p>	<p>Diffusion</p>  <p>Diffusionsstrom $j = -qD \frac{dn}{dx}$</p>	<p>Ladungsträger im Halbleiter</p>  <p>(Haupt-)Ladungsträgerdichten $p_p = N_a$ $n_n = N_d$ $np = n_i^2$ Ni: Eigenleitungsichte Minoritätsladungsträger $p_n = n_i^2 / N_a$ $n_p = n_i^2 / N_d$</p>
<p>Gleichgewichtszustand</p>  <p>Kontaktspannung $U_T \ln \frac{n(0)}{n(d)} = \pm V$ $n(x) = n(0)e^{\pm V/U_T}$ Maxwell-Boltzmann Formel $D/\mu = kT/e = U_t$ Einsteingleichung Ut - Thermische Spannung</p>		<p>E-Feld, Spannung</p> <p>Gaußsches Gesetz $E = -\frac{dV}{dx}$ $\frac{dE}{dx} = \frac{Q}{\epsilon}$ Dicke der Verarmungszone  $d = \sqrt{\frac{2\epsilon V}{eN_a}}$</p>

Abbildung 25: V ist Potential, E ist E-Feld, n ist Ladungsträgerdichte, q ist die Ladung des Ladungsträgers (mit Vorzeichen), Q ist die Ladungsdichte der Ladungsträger: ($Q = qn$); e ist die elementare Ladung ($+1.6 \cdot 10^{-19} \text{ C}$), μ ist Beweglichkeit der Ladungsträger, D ist die Diffusionskonstante, ϵ ist elektrische Leitfähigkeit (Permittivität) ($\epsilon = \epsilon_r \epsilon_0$), N_a und N_d sind die Dichten von Akzeptoren und Donatoren.

Im E-feld entsteht ein Driftstrom (Abbildung 25 Drift)

Wenn die Ladungsträgerdichte inhomogen ist, entsteht ein Diffusionsstrom (Abbildung 25 Diffusion).

Wenn sich der Drift- und der Diffusionsstrom kompensieren, haben wir Gleichgewichtszustand. Die Ladungsträgerdichte ist innerhalb eines Energiebandes durch Maxwell-Boltzmann Formel beschrieben (Abbildung 25 Gleichgewichtszustand). Die Elektronen sind eigentlich Fermi-Dirac-verteilt, diese Verteilung kann durch Maxwell-Boltzmann-Verteilung innerhalb eines Energiebandes approximiert werden.

Das Gaußsche Gesetz beschreibt wie Ladung Q E-Feld erzeugt. E-feld ist die Änderung (Gradient) des Potentials. (Abbildung 25 E-Feld, Spannung)

Potentialverlauf im Substrat einer MOSFET Struktur

Abbildung 26 zeigt die Potentiale innerhalb des MOSFETs wenn wir Source, Drain und Gate-Elektroden auf 0 V legen. Wir machen hier die Annahme, dass nur die E-Feldkomponente in vertikaler z -Richtung die Ladungsträgerdichten beeinflusst. Das E-Feld in horizontaler Richtung ist für den Strom verantwortlich.

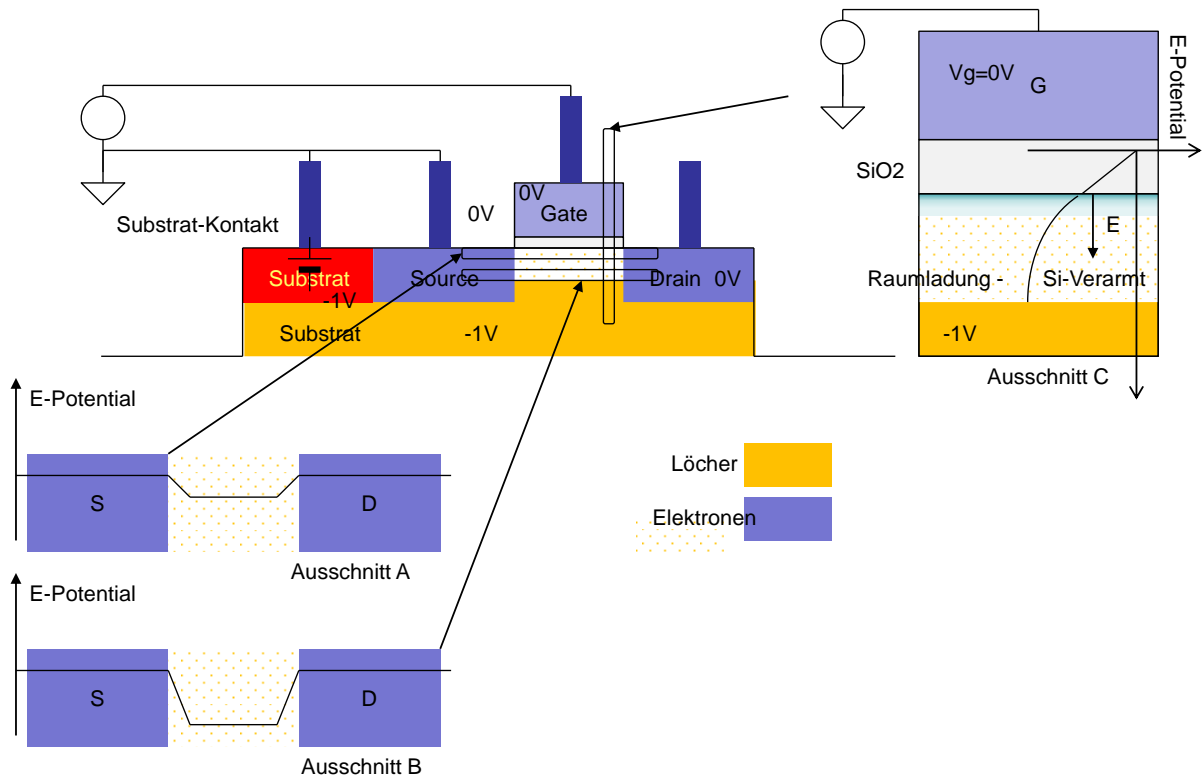


Abbildung 26: Potentiale in verschiedenen Bereichen der MOS Struktur

Wir definieren das Potential am Source-Metallkontakt als das Referenzpotential 0V. Da wir die Annahme machen dass es keine Kontaktspannung zwischen dem Metallkontakt und n-Silizium gibt, ist das Potential im n-Silizium Source auch 0 V. Der Drain (Metall und Silizium) sind ebenfalls auf 0 V. Der Bulkkontakt (Metall) ist zwar 0 V, wegen der Kontaktspannung ist das Substratpotential niedriger. Die Kontaktspannung ist gleich wie die zwischen n- und p-Silizium (Source und Substrat). Diese Spannung ist durch die Formel

$$V_{\text{cont}} = U_T \ln\left(\frac{N_a N_d}{n_i^2}\right) \quad (1)$$

gegeben. N_a und N_d sind die Dichten von Akzeptoren in Substrat (Kanalbereich) und Donatoren im Source, n_i ist die Eigenleitungsichte, etwa 10^{10} cm^{-3} auf 300 K. Thermische Spannung $U_T = kT/e \sim 26 \text{ mV}$ auf 300 K. Wir werden folgende Werte annehmen: $N_a \sim 10^{18} \text{ cm}^{-3} = N_d$. Wenn wir diese Werte einsetzen bekommen wir:

$$V_{\text{cont}} = 0.958 \text{ mV} \sim 1V.$$

Das P-Substratpotential ist also etwa -1 V.

Wir haben den Potentialverlauf in drei Ausschnitten A, B, C (zwei horizontale Ausschnitte A und B und ein vertikaler C) in Abbildung 26 gezeichnet.

Dort wo die elektrischen Potentiale hoher sind (und die Potentialenergie für Elektronen niedriger ist) (in Source und Drain) sammeln sich freie Elektronen. Elektronen sammeln sich in den Bereichen mit höherem elektrischen Potential da sie dorthin driften. Die Löcher werden aus diesen Bereichen verdrängt. Die Elektronendichte (und Löcherdichte) kann in einem Energieband im Gleichgewichtszustand durch Maxwell-Boltzmann Formel approximiert werden. Betrachten wir zuerst die Löcher im Substrat. Die Veränderung des Potentials im Substrat führt dazu dass die Löcherdichte sehr schnell von N_a auf $\ll N_a$ fällt. Es entsteht eine Verarmungszone (Raumladungszone) mit einem scharfen Übergang im Substrat, wo die negative Ladung der Akzeptor-Ionen lokal nicht kompensiert ist (Abbildung 27). Die negative Gesamtladung der Raumladungszone muss der positiven Ladung in der Gate-Elektrode gleich sein (Elektroneutralität).

Wenn wir die Raumladungszone im P-Bereich (Ausschnitt C) betrachten, erwarten wir, dass die Elektronendichte am größten nahe Si-SiO₂ Grenze ist, da dort das elektrische Potential am größten ist. Wenn wir eine positive Spannung an Gate anlegen (Abbildung 27), werden die Löcher aus dem Substrat noch weiter verdrängt. Die Verarmungszone (Sperrschicht) erweitert sich. Es bleibt die unkomensierte negative Ladung von Akzeptor-Ionen in der Verarmungszone. Die Elektronendichte an Si-SiO₂ Grenze steigt.

Wir nennen den Effekt wenn sich in einem P-Siliziumbereich N-Ladungsträger ansammeln „Inversion“.

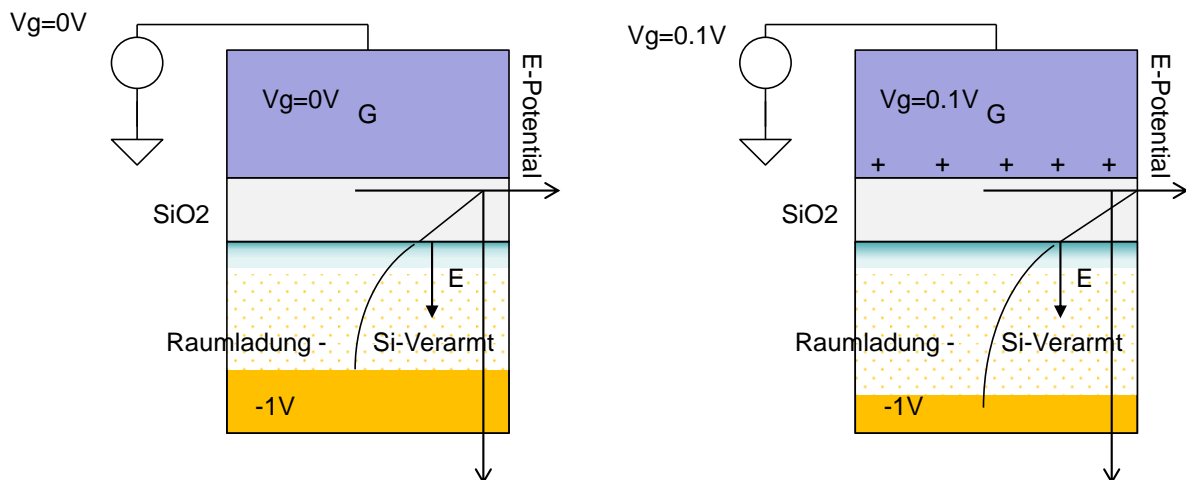


Abbildung 27: Wenn wir eine positive Spannung an Gate anlegen, werden die Löcher aus dem Substrat noch weiter verdrängt

Einfluss der vertikalen E-Feld Komponente: Kanalladung als Funktion von Gate-Spannung

Wir werden im nachfolgenden Text den Transistorstrom als Funktion von V_g (bzw. V_{gs}) herleiten. Um den Strom zu berechnen, sollen wir zuerst die Elektronendichte bestimmen. Die

Elektronendichte hängt vom Potential an der Si-SiO₂ Grenze ab. Im ersten Schritt rechnen wir das Potential an der Si-SiO₂ Grenze V_{SiO_2} . Dieses Potential ist durch die Formel beschrieben:

$$V_{\text{Si,SiO}_2} = V_{\text{sub}} + V_{\text{dep}} = V_s - V_{\text{cont}} + V_{\text{dep}} \approx -1\text{V} + V_{\text{dep}} \quad (2)$$

V_{sub} ist das Potential im Substrat (p-Silizium), V_s ist das Source-Potential ($V_s = 0\text{V}$), V_{dep} ist die Potentialänderung innerhalb der Verarmungszone. Um V_{dep} als Funktion von V_g zu berechnen, werden wir die MOS Struktur mit einer Ersatzschaltung modellieren.

Wenn sich V_{dep} erhöht, erhöht sich auch die negative Ladung in der Verarmungszone Q_{dep} - die Verarmungszone verhält sich also wie eine Kapazität C_{dep} . Da sich in der Gate-Elektrode die gleiche aber positive Ladungsmenge wie Q_{dep} sammelt, bilden die Kapazitäten der Verarmungszone C_{dep} und die Gate-Kapazität C_{ox} eine Reihenschaltung, Abbildung 28. C_{ox} ist als $Q_{\text{gate}}/V_{\text{ox}}$ definiert (V_{ox} ist die Spannung im Gate-Oxid). Es ist:

$$C_{\text{ox}} = \epsilon_0 \epsilon_{\text{SiO}_2} \frac{A}{t_{\text{ox}}} \sim 8.854 \cdot 10^{-12} \frac{\text{As}}{\text{Vm}} \times 3.9 \times \frac{A}{t_{\text{ox}}} \quad (3)$$

A ist die Gate-Fläche und t_{ox} die Gate-Dicke, z.B. für eine 65 nm Technologie $t_{\text{ox}} = 2.6\text{ nm}$.

C_{dep} kann als $Q_{\text{dep}}/V_{\text{dep}}$ (normale Kapazität) oder als $dQ_{\text{dep}}/dV_{\text{dep}}$ (dynamische Kapazität) definiert werden. In beiden Fällen hängt C_{dep} von V_{dep} ab.

Wenn wir C_{dep} als dynamische Kapazität definieren, gilt

$$C_{\text{dep}} \equiv \frac{dQ_{\text{dep}}}{dV_{\text{dep}}} = \epsilon_0 \epsilon_{\text{Si}} \frac{A}{t_{\text{dep}}} \quad (4)$$

Q_{dep} ist die Ladung in der Verarmungszone, V_{dep} ist die Potentialänderung innerhalb der Verarmungszone und t_{dep} die Dicke der verarmten Zone, $\epsilon_{\text{Si}} \sim 12$. Interessanterweise ist die Formel gleich wie die für Plattenkondensator (das beweisen wir später).

(Es ist für meiste Berechnungen sinnvoll C_{dep} als dynamische Kapazität zu definieren, da wir normalerweise Änderungen dV_{dep} rechnen.)

Da sich t_{dep} mit Zunahme von V_{dep} verändert, ist C_{dep} nicht eindeutig. Es ist nützlich die dynamische C_{dep} Kapazität für die Maximale $t_{\text{dep,max}}$ zu definieren. Diese $t_{\text{dep,max}}$ bekommen wir für $V_{\text{dep}} = V_{\text{cont}}$:

$$C_{\text{dep,min}} = \epsilon_0 \epsilon_{\text{Si}} \frac{A}{t_{\text{dep,max}}} \quad (5)$$

Die Dotierung und die Oxiddicke werden üblicherweise so gewählt, dass C_{ox} etwa $4 \times$ größer als $C_{\text{dep,min}}$ ist.

Wir definieren ebenfalls einen Faktor (genannt „slope factor“) als:

$$n = (C_{\text{dep,min}} + C_{\text{ox}})/C_{\text{ox}} \sim 1.25 \quad (6)$$

C_{ox} und C_{dep} bilden einen Kapazitiven Spannungsteiler.

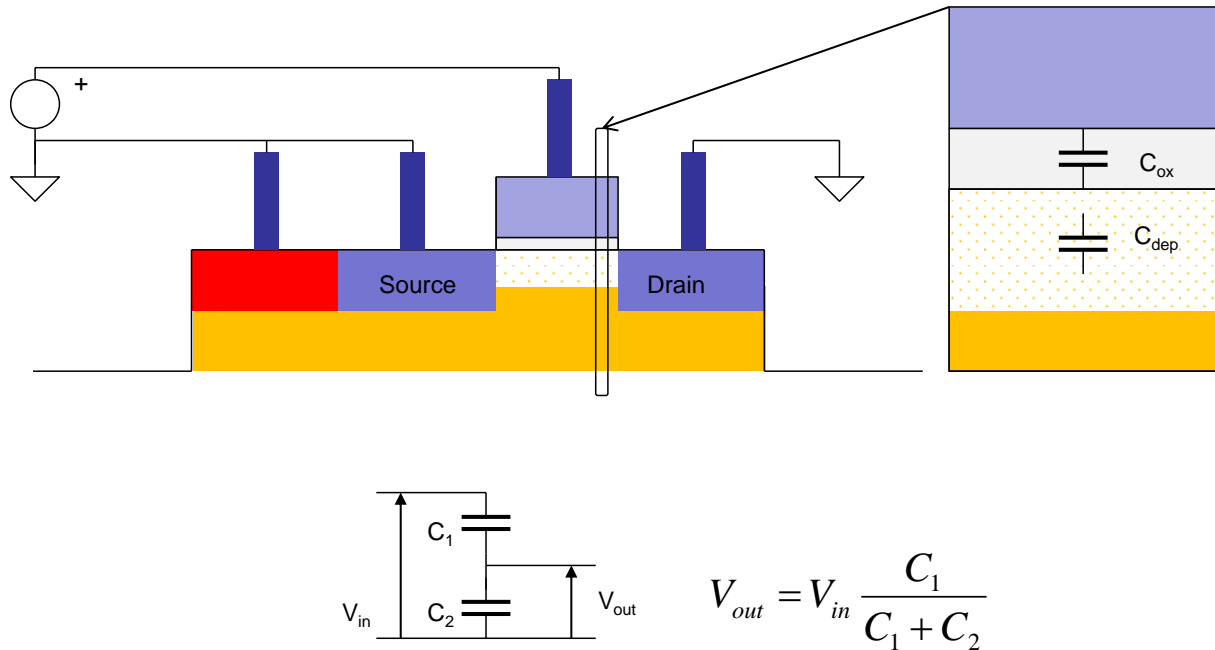


Abbildung 28: Die Kapazitäten in der MOS Struktur bilden einen Spannungsteiler

Wir haben zwei Kapazitäten 1) die Oxidkapazität C_{ox} – sie ist durch die Oxiddicke bestimmt und 2) die Kapazität der Verarmungszone C_{dep}

Berechnen wir nun das Potential an Si-SiO₂ Fläche – $V_{\text{Si,SiO}_2}$.

Wir können die Formel für Spannungsteiler verwenden. Daraus folgt:

$$V_{\text{Si,SiO}_2} = V_{\text{sub}} + (V_{\text{g}} - V_{\text{sub}}) \times \frac{C_{\text{ox}}}{C_{\text{ox}} + C_{\text{dep}}} \quad (7).$$

Wir behandeln hier C_{dep} als eine konstante Kapazität. Das ist eine gute Approximation wenn wir folgendes annehmen.

$$C_{\text{dep}} = 2 C_{\text{dep,min}} \quad (8)$$

weil näherungsweise $Q_{\text{dep}} = 2 C_{\text{dep,min}} V_{\text{dep}}$ gilt.

Beweis: (optional)

Die Verarmungszonentiefe kann wie folgend gerechnet werden.

Berechnen wir das E-Feld in z-Richtung. Z-Koordinate ist null am unteren Rand der Verarmungszone und zeigt nach oben. Gaußsches Gesetz:

$$\frac{dE_z}{dz} = -\frac{eN_a}{\epsilon_0 \epsilon_{\text{Si}}} \Rightarrow E_z = -\frac{eN_a}{\epsilon_0 \epsilon_{\text{Si}}} z \quad (9)$$

Potentialgleichung:

$$-\frac{dV_z}{dz} = E_z \Rightarrow V_z = \frac{eN_a}{\epsilon_0 \epsilon_{Si}} \frac{z^2}{2} \quad (10)$$

Daraus folgt:

$$t_{dep} = \sqrt{\frac{2\epsilon_0 \epsilon_{Si} V_{dep}}{eN_a}} \quad (11)$$

Die Ladung in Verarmungszone ist (Absolutwert):

$$Q_{dep} = AeN_a t_{dep} = A\sqrt{eN_a 2\epsilon_0 \epsilon_{Si} V_{dep}} \quad (12)$$

Die dynamische Kapazität der Verarmungszone ist:

$$C_{dep} \equiv \frac{dQ_{dep}}{dV_{dep}} = A\sqrt{\frac{eN_a \epsilon_0 \epsilon_{Si}}{2V_{dep}}} \quad (13)$$

Es gilt also, wegen (13) und (11):

$$C_{dep} = A\frac{\epsilon_0 \epsilon_{Si}}{t_{dep}} \quad (14)$$

Es gilt auch:

$$Q_{dep} = A\sqrt{eN_a 2\epsilon_0 \epsilon_{Si} V_{dep}} = V_{dep} A\sqrt{\frac{2\epsilon_0 \epsilon_{Si} eN_a}{V_{dep}}} = 2C_{dep} V_{dep} \quad (15)$$

Es gilt etwa: $C_{ox} = 4 \times C_{dep, min}$ (5) und $C_{ox} = 2 \times C_{dep}$ (8).

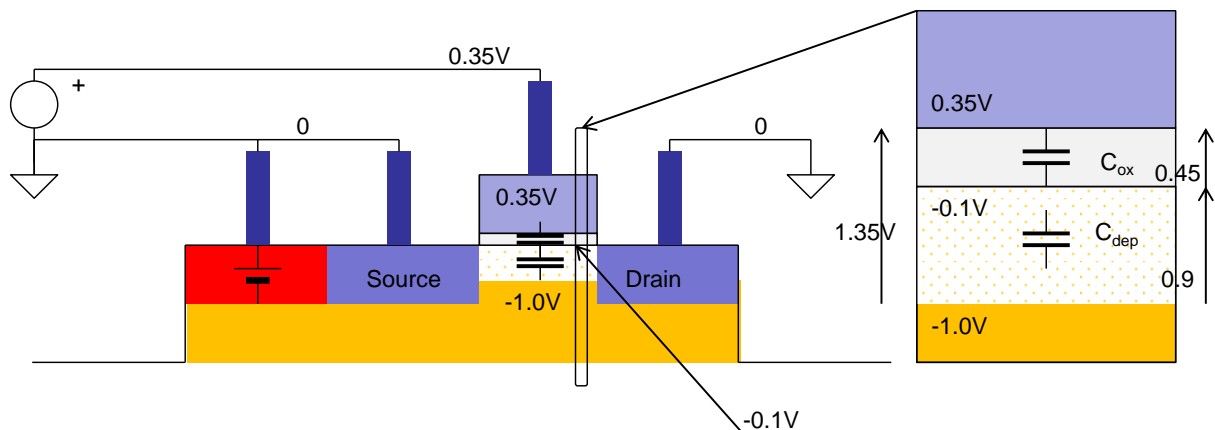


Abbildung 29: Potentiale für $V_g = 0.35 V$

Beispiel: Wir setzen $V_g = 0.35 V$ (Abbildung 29). Dieser Wert ist interessant da er knapp unterhalb der Schwellenspannung liegt. In der Verarmungszone steigt das Potential von $-1.0V$ auf etwa $\sim -0.1V$ (um $0.9V$). Im Oxid zwischen dem Substrat und der Gate Elektrode steigt das

Potential um noch etwa 0.45V auf 0.35V. Wenn das Gate auf 0.35V liegt, ist das Potential auf der Silizium-Oberfläche -0.1V. Das zeigt auch das folgende Bild.

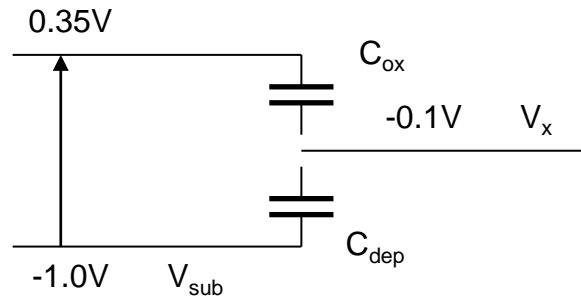


Abbildung 30: Spannungsteiler

Ein positives Potential ist eine Barriere für die positive Ladung (Löcher) und ein negatives Potential ist die Barriere für Elektronen. Daraus folgt: ein Elektron kann schwer von Source zum Drain, durch die tiefere P-Silizium Schichten gelangen, da dort die Potentialbarriere groß ist. (Abbildung 31)

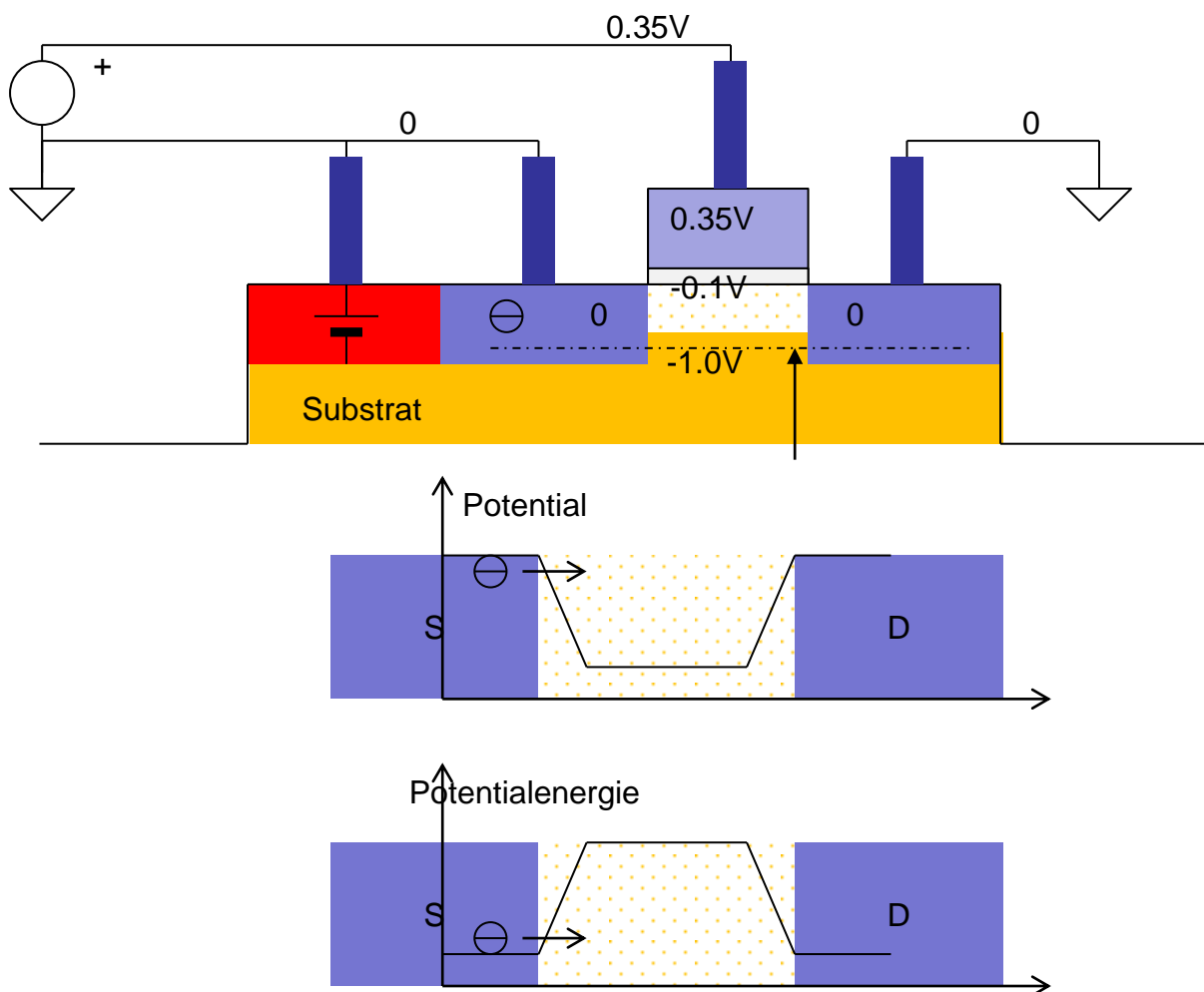


Abbildung 31: Potentialbarrieren

Im Unterschied zu den tieferen Substratschichten, ist auf der Substratoberfläche das Potential nur 0.1V niedriger wie in Source und Drain. Die Barriere ist kleiner und ein kleiner Diffusionsstrom kann entstehen (Abbildung 32). Darauf werden wir in der Vorlesung 3 detaillierter eingehen.

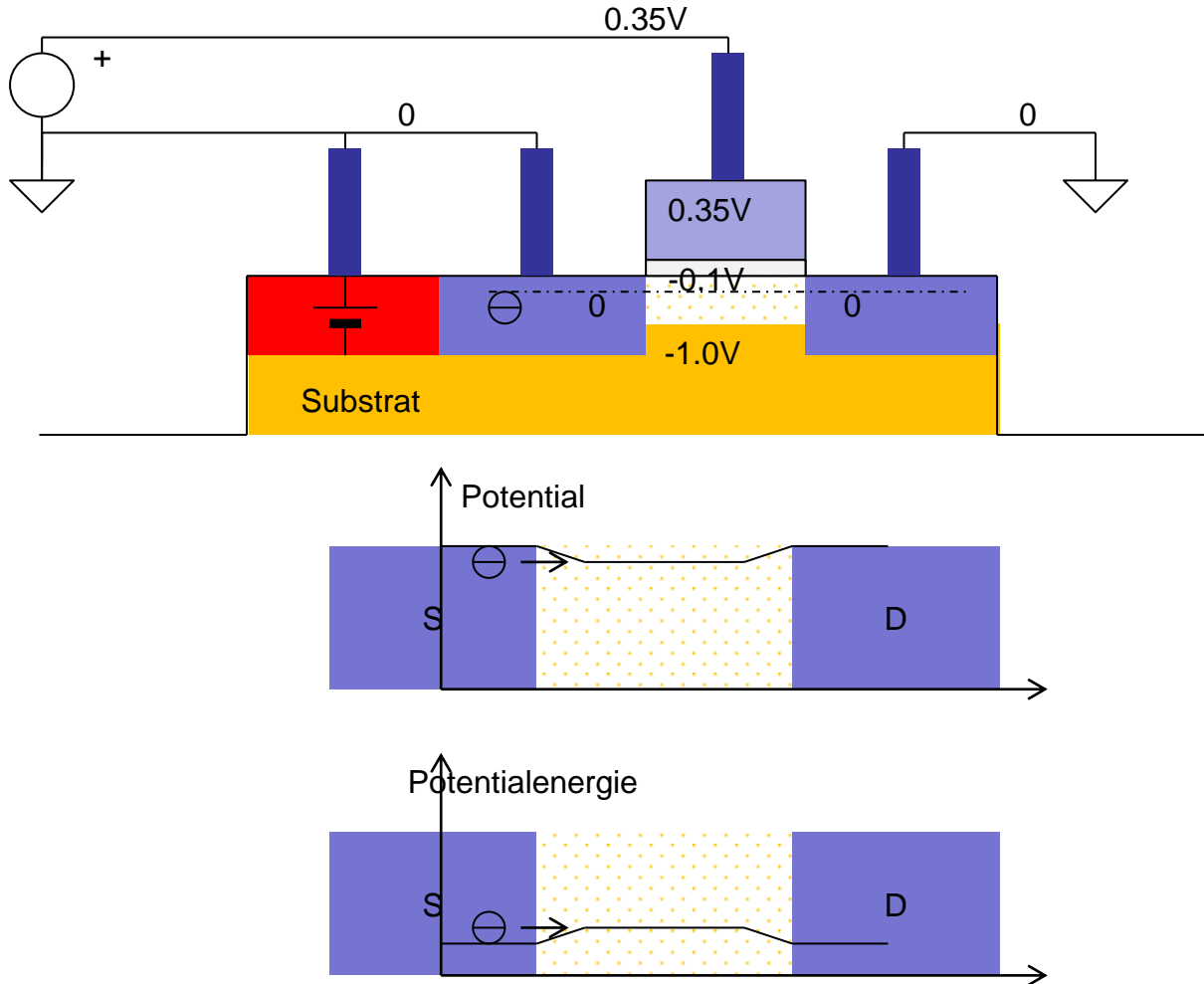


Abbildung 32: Kleinere Potentialbarriere nahe SiO₂

Erhöhen wir jetzt das Gate-Potential auf 0.5V:

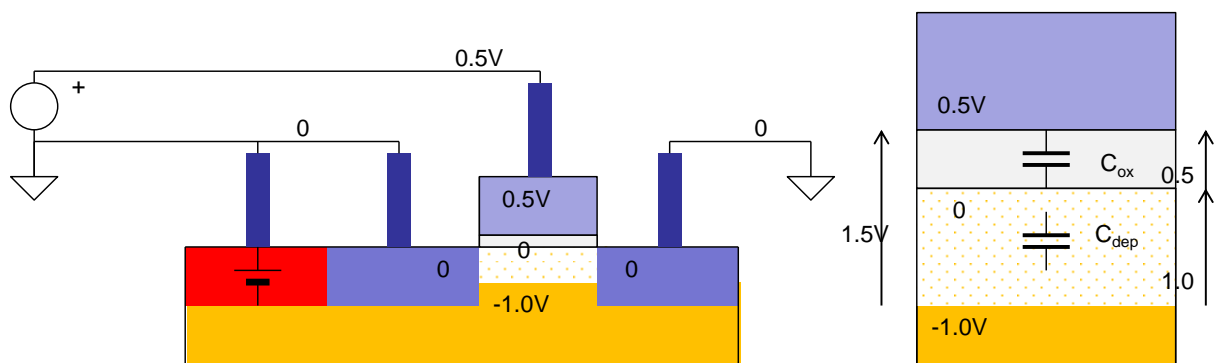


Abbildung 33: Schwellenspannung

Das Potential an der Substratoberfläche ($V_{Si,SiO_2} \equiv V_x$) ist nun 0V, also genau gleich wie in Source und Drain.

Wir definieren die Schwellenspannung V_{th} als die Gate-Source Spannung für welche $V_{Si,SiO_2} = V_s = V_d$ ist.

Aus der Formel für Spannungsteiler gilt:

$$V_{th} = -\frac{C_{dep}}{C_{ox}} \times V_{sub} = \frac{C_{dep}}{C_{ox}} \times V_{cont} = \frac{2C_{dep,min}}{C_{ox}} \times V_{cont} \quad (16)$$

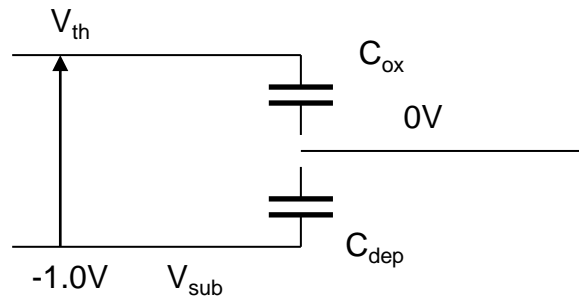


Abbildung 34: Spannungsteiler

Beachten wir dass die Schwellenspannung kleiner wird, wenn C_{ox} erhöht wird.

Berechnen wir V_{th} nur als Funktion von V_{cont} . (optional)

Es ist (siehe Gleichung 13):

$$C_{dep,min} = A \sqrt{\frac{eN_a \epsilon_0 \epsilon_{Si}}{2V_{cont}}} \quad (17)$$

Deswegen ist es:

$$V_{th} = \frac{2C_{dep,min}}{C_{ox}} \times V_{cont} = \frac{A \sqrt{2eN_a \epsilon_0 \epsilon_{Si} V_{cont}}}{C_{ox}} \quad (18)$$

Die Kontaktspannung ist mit (1) beschrieben. Da wir $N_d = N_a$ angenommen haben, gilt:

$$V_{cont} = 2U_T \ln\left(\frac{N_a}{n_i}\right) \quad (19)$$

Formel (18) wird normalerweise in Literatur hergeleitet.

Kleinere Schwellen sind vom Vorteil wenn die Versorgungsspannung niedrig ist. Eine niedrige Versorgungsspannung führt zu einem geringeren Leistungsverbrauch. Deswegen versucht man C_{ox} zu maximieren, bzw. die Dicke des Oxids möglichst klein zu machen.

Mit Temperaturanstieg sinkt die Schwellenspannung (18), da das Kontaktpotential auch sinkt (19). Es ist die Folge des Anstiegs von Eigenleitendichte in Silizium n_i (und daraus folgend der Elektronendichte in P-Silizium).

Wenn wir das Gate-Potential (Gate-Source Spannung) über die Schwellenspannung erhöhen, sollte das Potential an der Substratoberfläche über 0 V steigen. Das würde aber bedeuten, dass viele Elektronen aus der Source und Drain in die Regionen unterhalb des Oxids fließen, da dort für sie ein Potentialminimum ist (Abbildung 35).

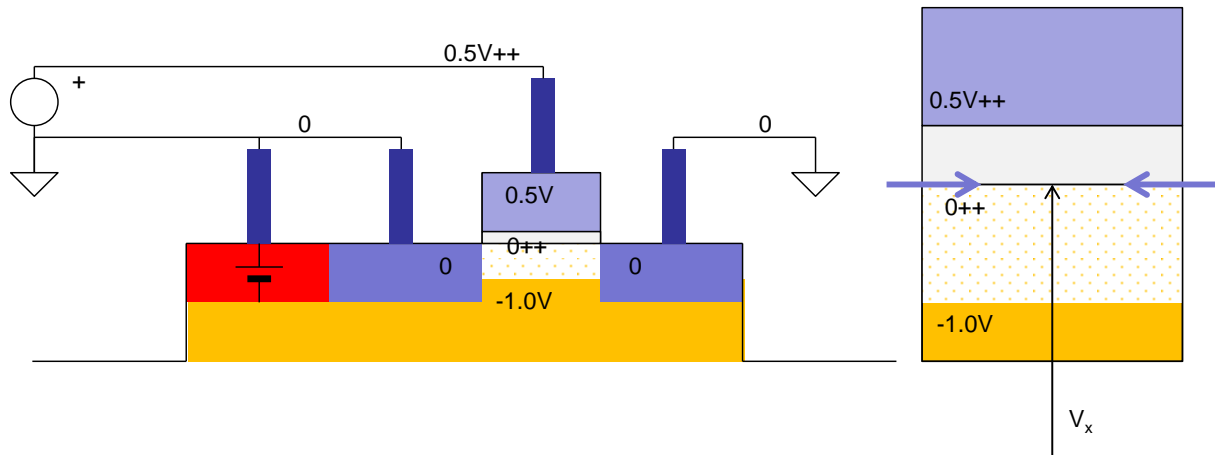


Abbildung 35: Potentiale in der MOS Struktur für $V_g > V_{th}$

In Wirklichkeit sammeln sich die Elektronen in Silizium unterhalb Gate und bilden einen leitenden Kanal. Die Elektronen im Kanal schließen Source, Drain und die Substratoberfläche kurz und halten auf diese Weise, durch ihre eigene Ladung, das Kanal-Potential auf dem Niveau von Source und Drain. Der Kanal und die Source/Drain sind daher kurzgeschlossen. (Abbildung 36). Wir nennen solch einen Arbeitsbereich „starke Inversion“.

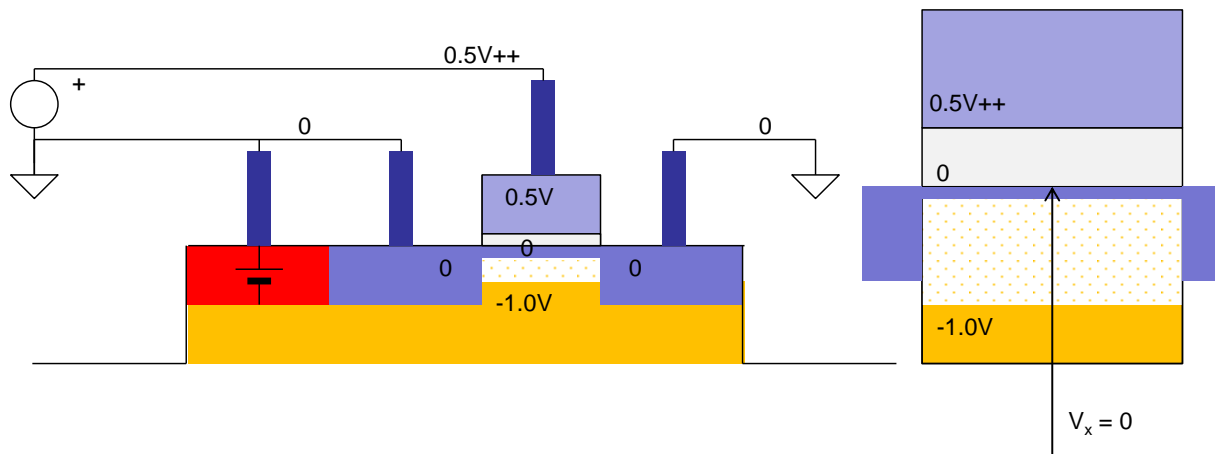


Abbildung 36: Elektronen aus der Source und dem Drain bilden einen leitenden Kanal. Source, Drain und die Substratoberfläche werden kurzgeschlossen $\rightarrow V_x = 0$

Berechnen wir nun die Kanalladung.

Die untere Elektrode der Kapazität C_{ox} liegt an einem festen Potential. Die Spannung am C_{dep} ist konstant. Die Spannungsquelle am Gate „sieht“ also nur die Eingangskapazität C_{ox} : Wenn sich die Gate-Spannung um dV_g ändert, fließt die Ladung $C_{ox} dV_g$ durch die Spannungsquelle. Genau dieselbe Ladung bildet sich im Kanal.

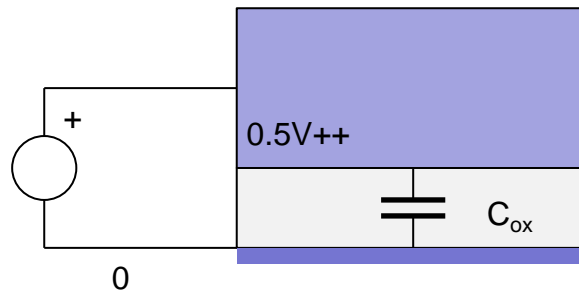


Abbildung 37: Ladung im Kanal

Für $V_g = V_{th} = 0.5 \text{ V}$ hatten wir noch keine Ladung im Kanal. Für $V_g > V_{th}$ gilt $dQ = C_{ox} dV_g$. Daraus folgt für die Ladung im Kanal:

$$Q = C_{ox}(V_g - V_{th})$$

Da $V_s = 0$ ist, können wir auch folgendes schreiben:

$$Q = C_{ox}(V_{gs} - V_{th}) \quad (20)$$

Zusammenfassung:

Wir definieren die Schwellenspannung $V_{th} \sim 0.5 \text{ V}$ als die Gate-Source Spannung für welche die Potentialen in Source und auf der Substratoberfläche (Si-SiO₂ Grenze) etwa gleich sind.

(Für $V_{gs} = V_{th}$ ist die Potentialbarriere null.)

Wenn die Gate-Source Spannung über die Schwelle steigt, sammeln sich die Elektronen im Kanal. **Wir reden dann von einer starken Inversion.** Für die Gate-Spannungen unter der Schwelle ist das Potential an der Substratoberfläche für die Kanalbildung nicht ausreichend.

Einfluss der horizontalen E-Feld Komponente: Drain-Strom Strom als Funktion von Drain-Source Spannung

Berechnen wir jetzt den Transistorstrom für kleine Spannungen V_{ds} .

Falls $V_{gs} > V_{th}$ entsteht ein Kanal aus Elektronen zwischen der Source und dem Drain, also eine Ohmsche Verbindung oder einen Widerstand. (Abbildung 38).

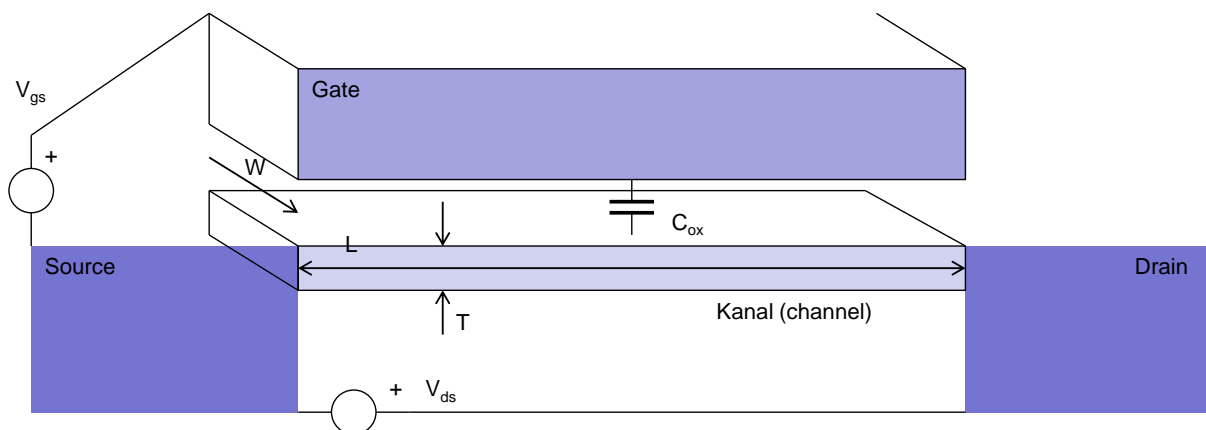


Abbildung 38: Ohmsche Verbindung zwischen der Source und dem Drain

Wenn wir eine kleine Spannung zwischen Drain und Source haben (V_{ds}), fließt ein Strom vom Drain in die Source (I_{ds}).

Der Kanal bildet einen Widerstand – der Strom ist durch die folgende Gleichung gegeben (Abbildung 38):

$$I_{ds} = A e \mu n E \quad (21),$$

e ist die elementare Ladung, μ Beweglichkeit, n Dichte der Elektronen im Kanal, E die horizontale E-Feld Komponente. Der Querschnitt A ist die Kanalbreite W multipliziert mit der Kanaltiefe t .

Daraus folgt:

$$I_{ds} = W t e \mu n E$$

Man kann daraus herleiten:

$$n t e = \frac{Q}{WL} = \frac{C_{ox} (V_{gs} - V_{th})}{WL}$$

Q ist die Gesamtladung im Kanal, L ist die Länge des Kanals. Wir haben das Ergebnis $Q = C_{ox} (V_{gs} - V_{th})$ benutzt.

Wir bekommen:

$$I_{ds} = \mu C'_{ox} W (V_{gs} - V_{th}) E$$

C'_{ox} ist die Kapazität pro Fläche (Kapazitätsbelag).

E-Feld ist in erster Näherung:

$$E = V_{ds}/L.$$

Deshalb gilt:

$$I_{ds} = \mu C'_{ox} \frac{W}{L} (V_{gs} - V_{th}) V_{ds} \quad (22).$$

Das ist die einfachste Gleichung für den Transistorstrom.

Sie gilt für kleine V_{ds} . Nur dann kann angenommen werden, dass die Ladung im Kanal gleichmäßig verteilt ist.

Wir sehen, dass der Strom vom Verhältnis W/L (genannt aspect ratio) abhängig ist. Das ist typisch für die MOSFETs. Im Gegensatz zu diesem Ergebnis, beeinflusst die Größe eines Bipolar-Transistors seinen Strom nicht.

Sättigung

Wie viel steigt der Strom wenn wir V_{ds} erhöhen?

Wir haben gesehen (20), dass die Ladung im Kanal durch die folgende Formel beschrieben ist wenn V_{ds} klein ist:

$$Q = C_{ox}(V_{gs} - V_{th})$$

Was passiert, wenn $V_{ds} \gg 0$ ist?

Wegen Symmetrie ist die Kanalladung nahe Source $C_{ox}(V_{gs} - V_{th})$ und nahe Drain $C_{ox}(V_{gd} - V_{th})$. (Das ist nicht ganz Korrekt da V_d ungleich 0 V ist. Wir werden darauf in Vorlesung 3 genauer eingehen.)

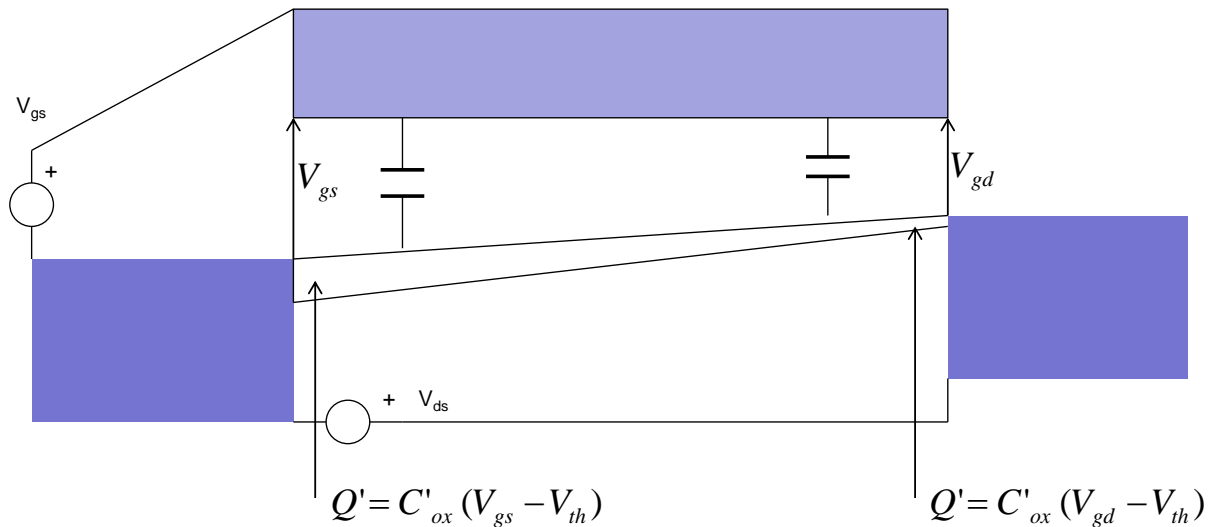


Abbildung 39: Strom als Funktion von V_{ds} . $V_{ds} > 0$

Also, für eine Drain Spannung $V_d = V_g - V_{th}$ haben wir keinen Kanal an der Drain-Seite mehr. Wir sagen, dass der Kanal abgeschnürt ist (pinched-off).

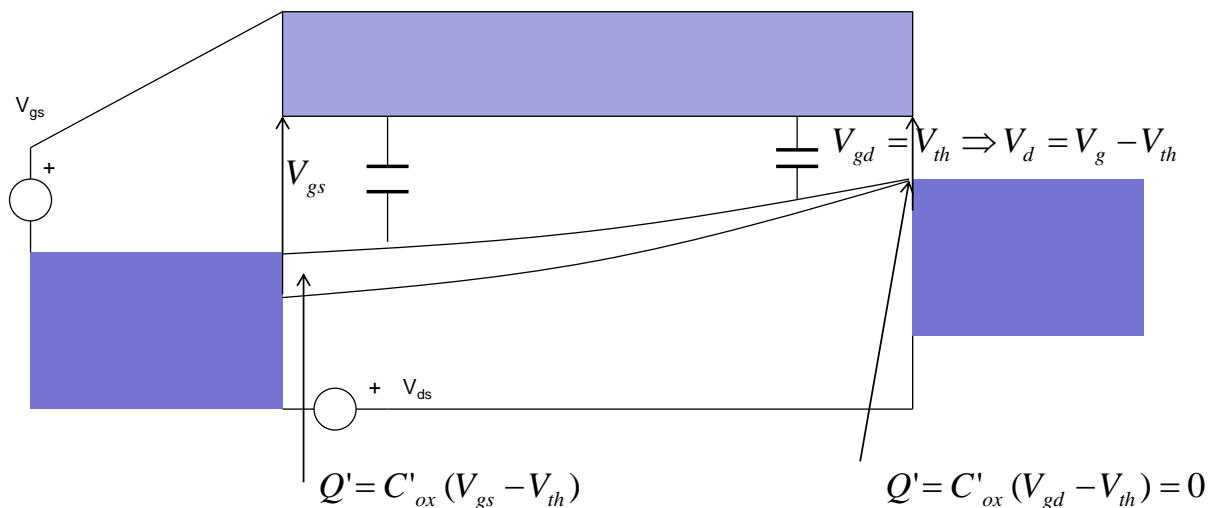


Abbildung 40: Strom als Funktion von V_{ds} . $V_{ds} = V_{gs} - V_{th}$

Ein weiterer Stromanstieg wird dadurch stark gebremst. Wir haben eine Stromsättigung. In erster Näherung steigt der Strom nicht mehr, wenn wir weiter V_{ds} erhöhen.

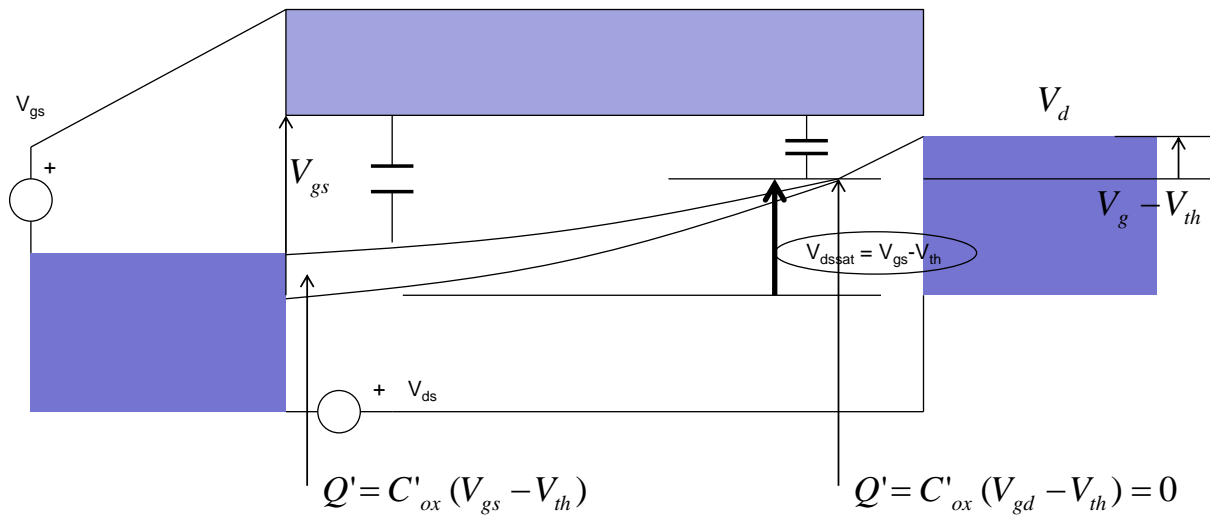


Abbildung 41: Strom als Funktion von V_{ds} . $V_{ds} > V_{gs} - V_{th}$

Die Bedingung für den Anfang von Sättigung: $V_d = V_g - V_{th}$ oder $V_{gd} = V_{th}$ kann man auch als $V_{gs} - V_{ds} = V_{th}$ (beide Seiten - V_s) oder als $V_{ds} = V_{gs} - V_{th}$ umschreiben.

Wir definieren die Sättigungsspannung als:

$$V_{dssat} = V_{gs} - V_{th} \quad (23)$$

Für höhere V_{ds} steigt der Strom nur ganz wenig weil sich L verkürzt.

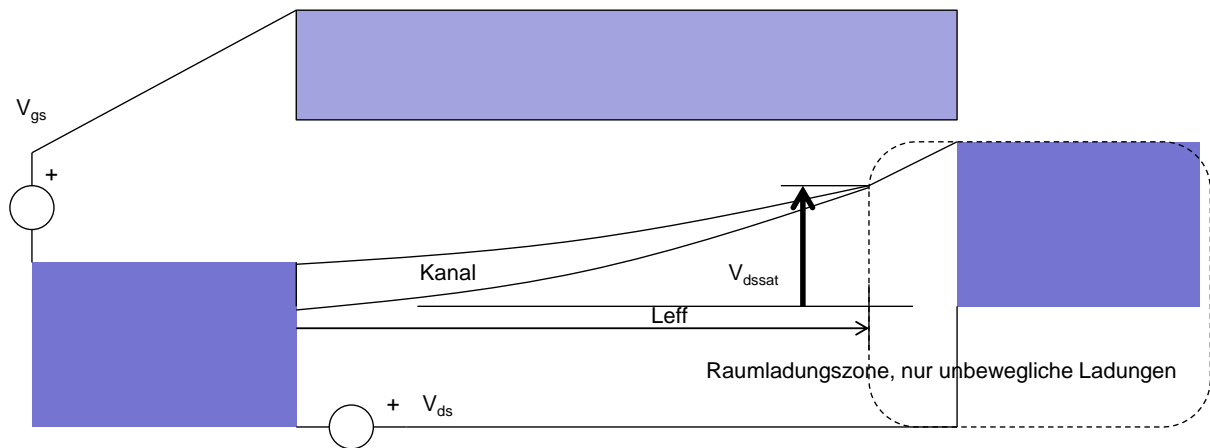


Abbildung 42: Effektive Kanallänge wird kleiner

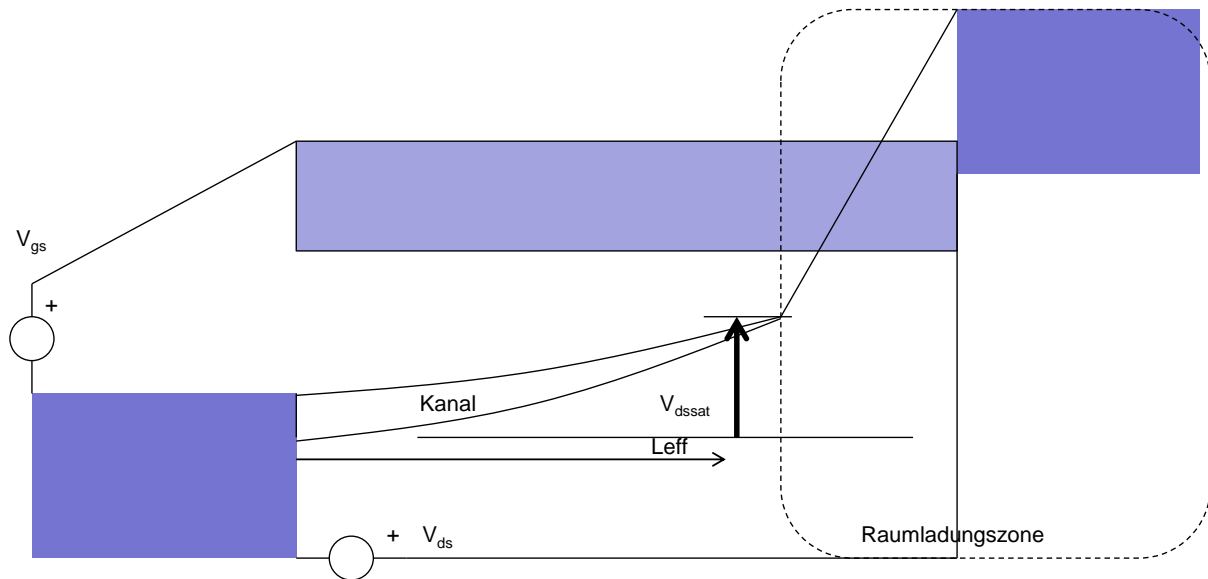


Abbildung 43: Effektive Kanallänge wird kleiner

Wie groß ist der Drain-Source Strom für $V_{ds} = V_{dssat}$, also am Anfang von Sättigung? Machen wir eine nicht ganz korrekte Annahme (A1): die Formel (22), die wir für kleine V_{ds} hergeleitet haben, auch bis $V_{ds} = V_{dssat}$ gilt.

Den Drain-Source Strom für V_{dssat} Spannung (Sättigungsstrom) könnten wir aus der Gleichung (22) durch Einsetzen von $V_{ds} = V_{gs} - V_{th}$ berechnen:

$$I_{dssat} = \mu C'_{ox} \frac{W}{L} (V_{gs} - V_{th})^2 \quad (24)$$

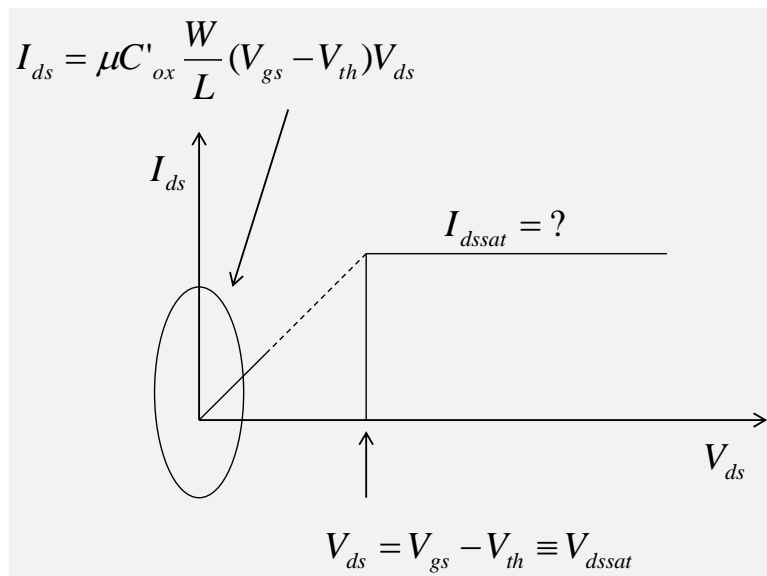


Abbildung 44: I_{ds} Formel – einfache Herleitung

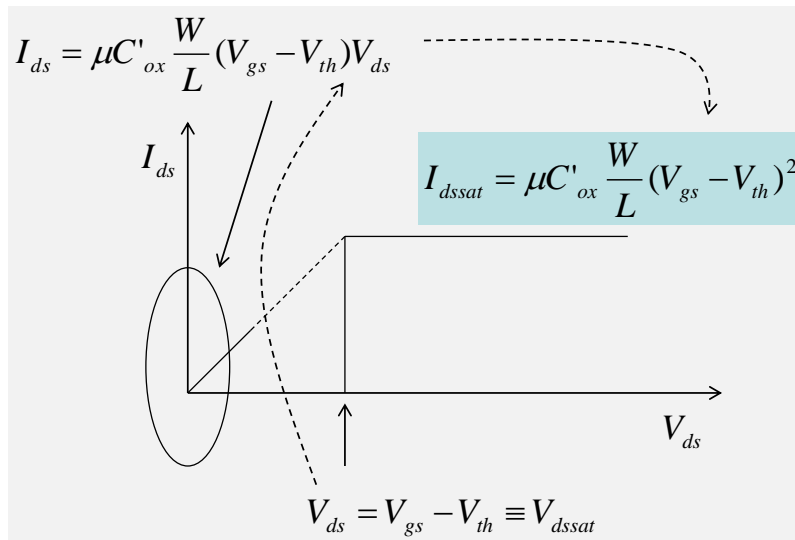


Abbildung 45: I_{ds} Formel – einfache Herleitung

Leider ist die Annahme (A1) nicht ganz korrekt. Der Stromanstieg ist für $V_{ds} > \sim 100\text{mV}$ schwächer als die Formel (22) gibt.

Genauere Berechnung führt zu einem zusätzlichen Faktor $1/2$. Die Formel für Sättigungsstrom lautet mit dieser Korrektur:

$$I_{dssat} = \frac{1}{2} \mu C'_{ox} \frac{W}{L} (V_{gs} - V_{th})^2 \quad (25)$$

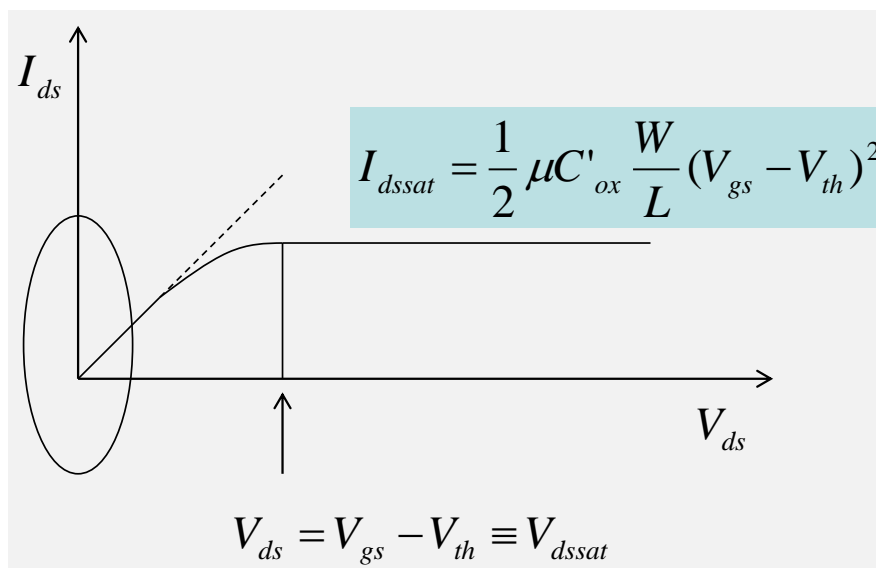


Abbildung 46: I_{ds} Formel – genauer

Auch diese Formel ist nicht besonders korrekt. Genauere Modelle zeigen (Substrateffekt und Sättigung der Driftgeschwindigkeit), dass der Faktor $1/2$ durch $1/(2n\alpha)$ ersetzt werden soll. Für V_{dssat} gilt:

$$V_{dssat} = \frac{V_{gs} - V_{th}}{n\alpha} \quad (26)$$

$$\alpha = 1 + \frac{V_{gs}}{nE_{sat}L} \quad (27)$$

Faktor α beschreibt die Sättigung der Driftgeschwindigkeit und Faktor n (slope factor – Formel 6) Substrateffekt (s. nächste Vorlesung).

$$n = (C_{dep,min} + C_{ox})/C_{ox} \sim 1.25 \quad (28)$$

α ist für lange Transistoren (Transistoren mit $L > 1\mu\text{m}$) ~ 1 . Für sehr kurze Transistoren oder für große Gate-Source Spannungen ist α deutlich größer als 1 und führt zu viel niedrigeren I_{dssat} Werten als erwartet. In der Regel gilt „kleinere Transistoren“ brauchen komplexere Formeln.

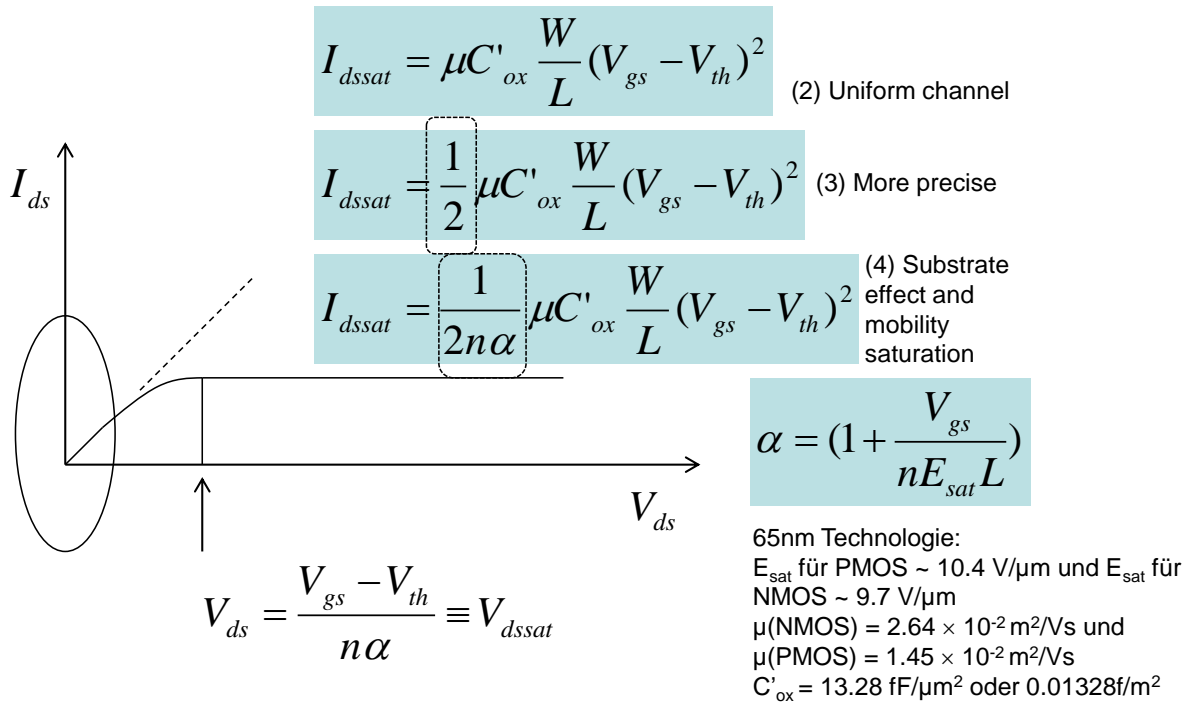


Abbildung 47: I_{ds} Formel – Überblick