

## Vorlesung 2

In dieser Vorlesung werden folgende Themen behandelt

Herstellung des MOSFETs

Funktionsweise des MOSFETs und die Herleitung der Stromgleichung

Kanalladung als Funktion von Gate-Spannung

Drain-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.

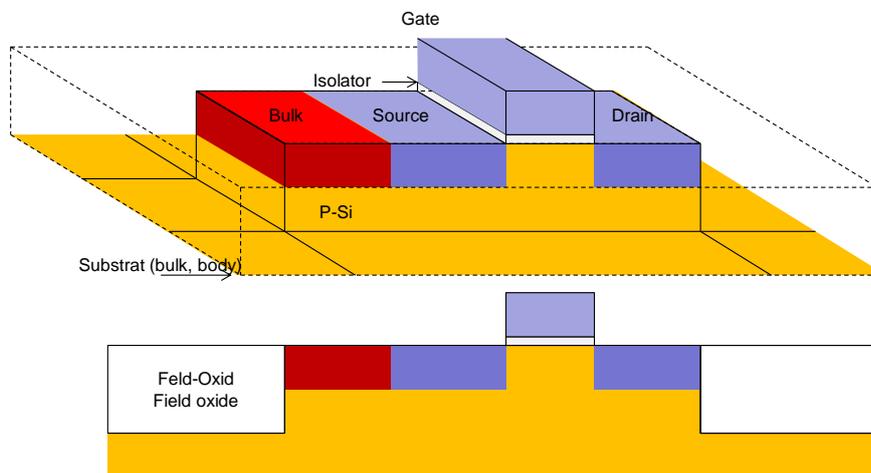


Abbildung 1: 3D-Darstellung eines NMOS Transistors

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

Die Transistoren werden voneinander durch Feld-Oxid isoliert, es ist eine etwa 300nm tiefe polykristalline Silizium-Dioxid- ( $\text{SiO}_2$ -) Schicht.

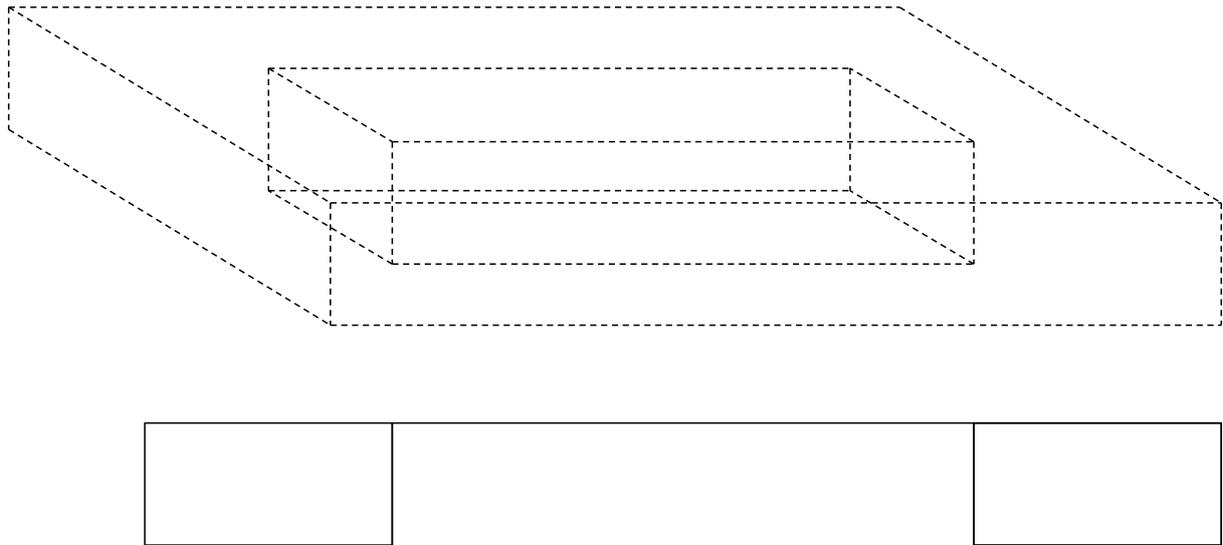


Abbildung 2: Transistor befindet sich im Feld-Oxid

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

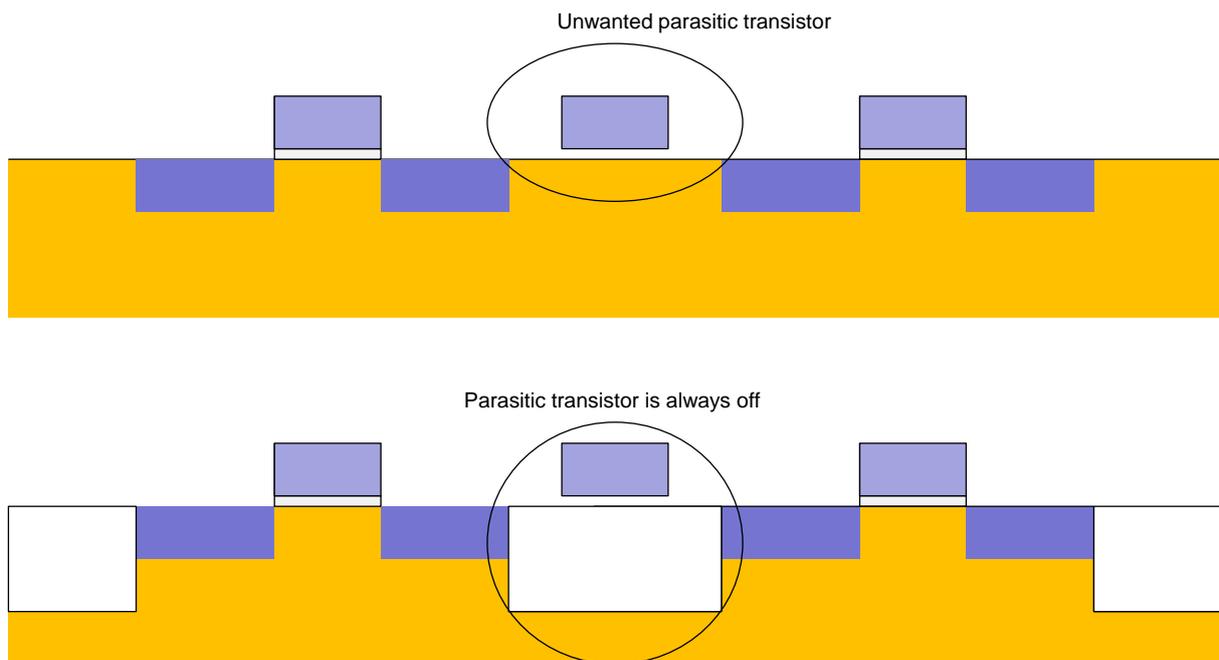


Abbildung 3: Isolierung von Transistoren

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

### **P-Wanne**

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

Eine Maskenlage (PWell) definiert die p-Wanne. Die Fotomaske aus Quarzglas und Metall wird mithilfe vom Fotolack auf Silizium übertragen indem Fotolack mit UV-Licht durch die Maske belichtet und strukturiert wird. Fotolack dient als Maske für die Lagen darunter

(Silizium-Dioxid und Silizium Nitrid). Fotolack wird verwendet um die Bereiche, die nicht geätzt werden sollen zu schützen. Die Silizium-Dioxid und Silizium Nitrid Strukturen werden als Masken für die Dotierung benutzt. Diffusionsverfahren oder Ionenimplantation werden als Dotiertechniken benutzt. ([Dotiertechniken - Waferherstellung - Halbleitertechnologie von A bis Z - Halbleiter.org](#))

Ionenimplantation:

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.

(<https://de.wikipedia.org/wiki/Ionenimplantation>)

Durch den Aufprall werden die Siliziumatome von ihren Gitterplätzen gestoßen (Gitterschäden), die Dotierionen lagern sich meist auf Zwischengitterplätzen an.

Dort sind sie elektrisch nicht aktiv, da keine Bindungen mit anderen Atomen vorliegen, die freie Ladungsträger hervorrufen könnten. Die verschobenen Siliziumatome müssen wieder ins Kristallgitter eingebaut werden (Ausheilen - annealing), und die elektrisch nicht aktiven Dotierstoffe aktiviert werden.

Ausheilen des Kristallgitters und Aktivierung der Dotierstoffe

Durch einen Temperaturschritt bei ca. 1000 °C werden die Dotierstoffe auf Gitterplätze bewegt (vorher befinden sich nur ca. 5 % der Dotieratome auf Gitterplätzen). Die Gitterschäden durch die Zusammenstöße werden bereits bei ca. 500 °C ausgeheilt. Diese Schritte werden sehr kurze Zeit durchgeführt.

([Dotiertechniken - Waferherstellung - Halbleitertechnologie von A bis Z - Halbleiter.org](#))

Die Ionenquelle des Ionen-Implanters besteht aus einem Heizdraht der vom Dotantengas angeströmt wird. Die Elektronen verlassen den Heizdraht durch thermische Emission, werden beschleunigt und ionisieren den Dotantengas.

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

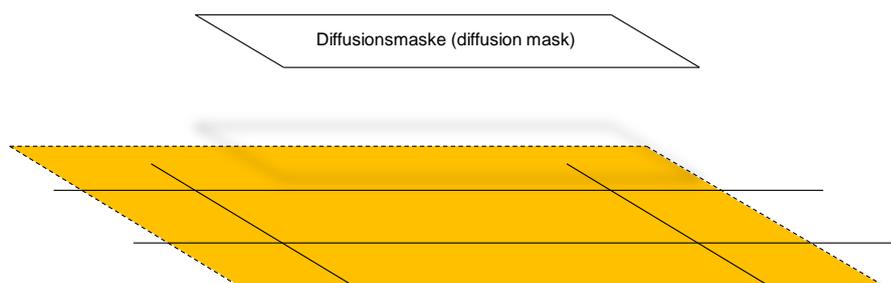


Abbildung 4: „Diffusionsmaske“ definiert die Stelle wo sich Transistor befinden wird

Ätzverfahren (z.B. reaktives Ionenätzen) wird benutzt und den Graben (trench) für das Feldoxid zu erzeugen.

Zur Strukturübertragung sind stark anisotrope Prozesse erwünscht, also Ätzungen nur in vertikaler Richtung, so dass die Lackmaske nicht unterätzt wird. ([Halbleitertechnologie von A bis Z - Halbleiter.org](http://Halbleitertechnologie.von.A.bis.Z.-Halbleiter.org))

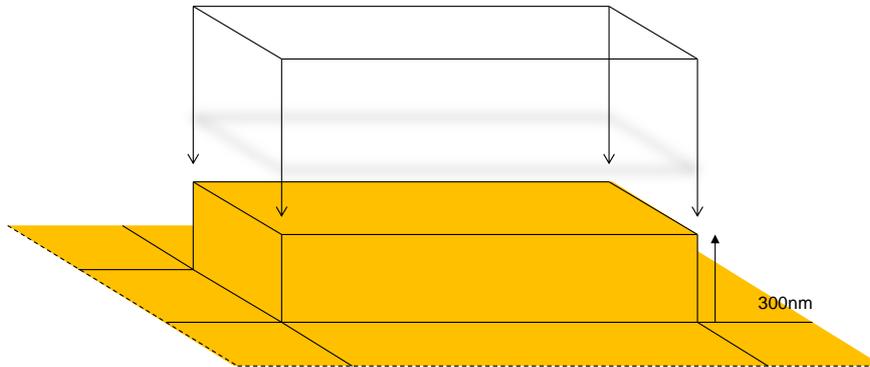


Abbildung 5: Erzeugung von trenches

Dieser Graben wird mit Silizium-Dioxid gefüllt. Dafür wird CVD Verfahren (chemische Dampfabcheidung) verwendet. Eine detailliertere Beschreibung finden Sie hier: <https://de.wikipedia.org/wiki/Grabenisolation>

Betrachten wir ein Standard-CMOS Prozess, wo  $\text{SiO}_2$  als Gate-Isolator dient. (Neuere Technologien verwenden kein  $\text{SiO}_2$  sondern ein Dielektrikum mit größerer relativen Permittivität z.B.  $\text{HfO}_2$ .)

Dünnes Gateoxid wird auf dem aktiven Bereich durch thermische Oxidation erzeugt.

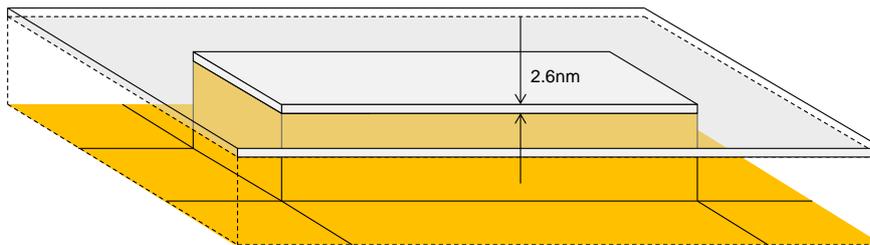


Abbildung 6: Thermische Oxidation

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  $\text{SiO}_2$  etwa 0,5 nm ist.

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

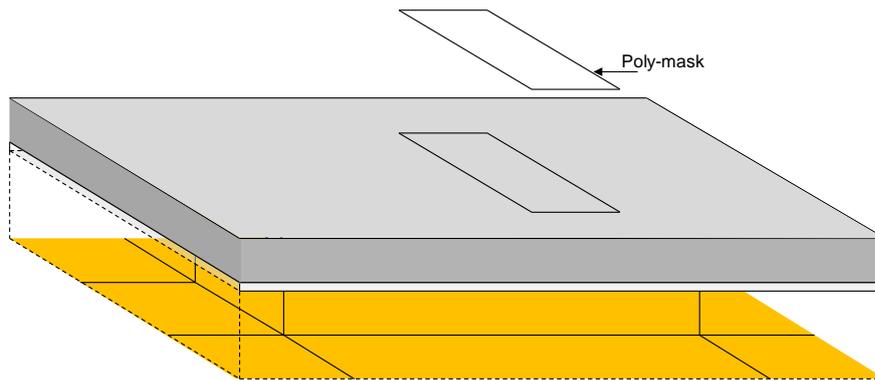


Abbildung 7: Polysilizium Maske definiert die Gate-Elektrode und die Polysilizium-Leitung

Die Polysilizium Strukturen werden durch Ätzen (nasschemisch, Plasma- und Ionenätzen) strukturiert. Auch hier wird Fotolack verwendet um die Bereiche, die nicht geätzt werden sollen zu schützen.

Das dünne Oxid bleibt nur im Überlappbereich von Diffusions- und Polysilizium-Maske.

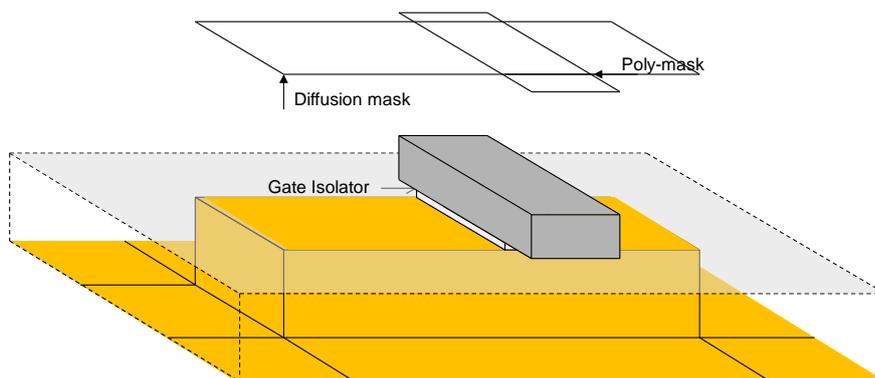
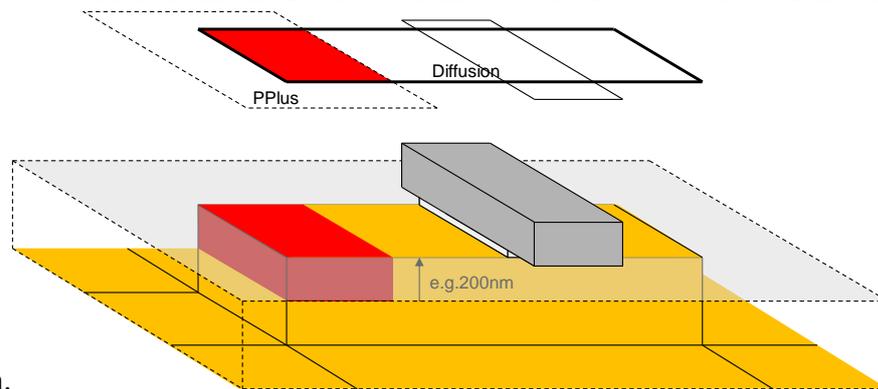


Abbildung 8: Gate-Oxid

Welche sind die Vorteile einer Gate-Elektrode aus Polysilizium?

1. Gate-Elektrode aus Polysilizium kann dotiert werden – dadurch kann die Schwellenspannung angepasst 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 (1400C gegenüber 660C), deswegen können die nachfolgenden Schritte bei höheren Temperaturen erfolgen – insbesondere Annealing von Source und Drain.

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



Ionenimplantation.

Abbildung 9: P-Puls Maske. Überlapp von Diffusions- und P-Plus Masken wird P-dotiert

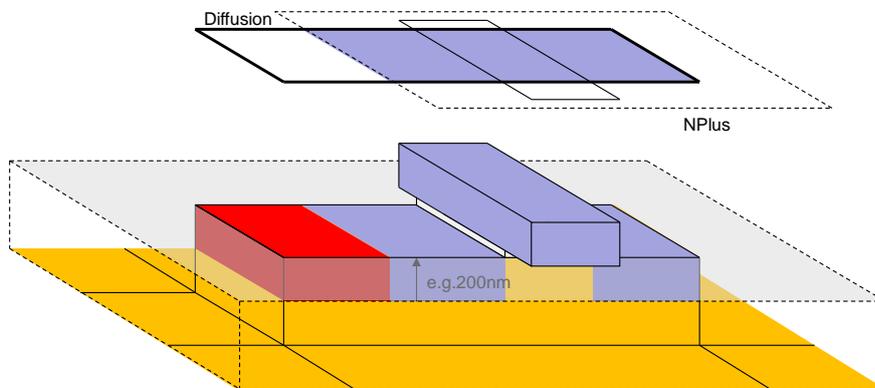


Abbildung 10: N-Puls Maske -> Überlapp von Diffusions- und N-Plus Masken wird N-dotiert (auch Poly-Gate)

Die N-Plus oder P-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/>

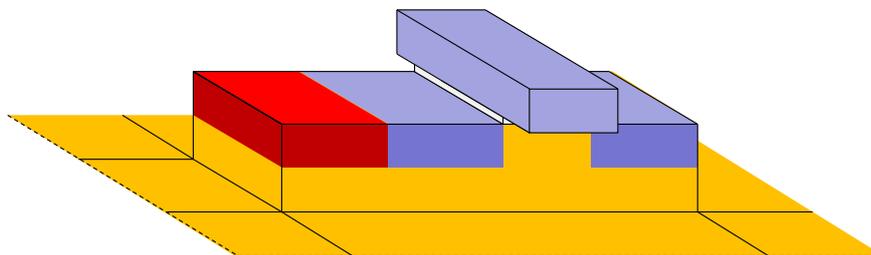


Abbildung 11: Vollständiger Transistor

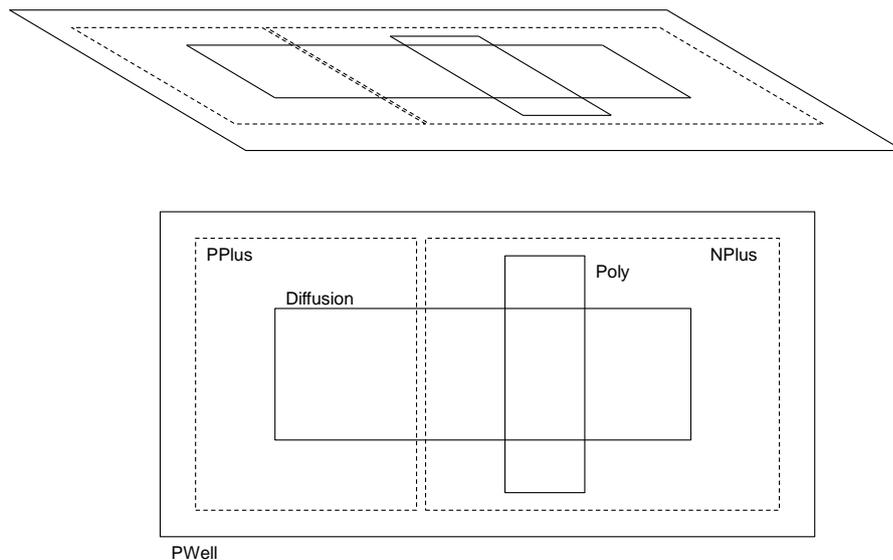


Abbildung 12: Masken, wie sie im Layout-Editor gezeichnet werden

### Salicide (Extra-Thema)

Normalerweise werden die Gate-, Source-, Drain- und Bulkkontakte metallisiert.

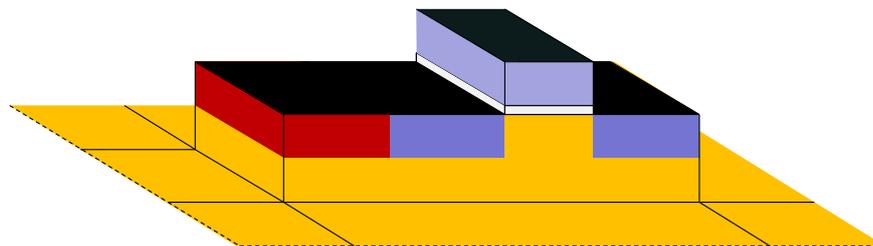


Abbildung 13: Transistorkontakte mit Metallisierung (Salicide)

Dadurch wird der Widerstand von Kontakten verkleinert – das ist für viele Schaltungen vom Vorteil. Man nennt diesen Prozess „Salicide“ – Selbstjustierendes Silicid.

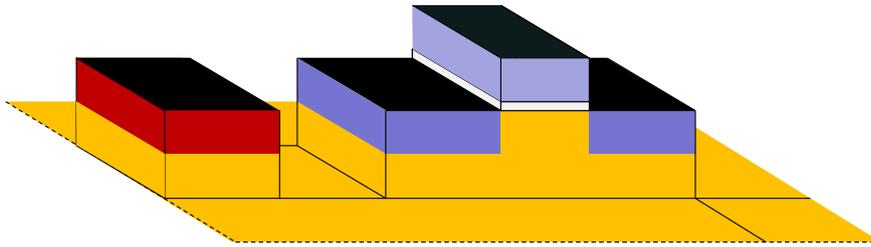
#### [Silicide – Wikipedia](#)

Silicide sind binäre metallische Verbindungen des Siliciums.

Abscheidung einer dünnen Metallschicht wie Nickel (Ni), Titan (Ti), Cobalt (Co) oder Platin (Pt) erfolgt meist durch Sputterdeposition. Nach der Abscheidung folgt ein erster Hochtemperaturschritt – ein sogenanntes Rapid Thermal Processing (dt.: schnelle thermische Bearbeitung) – bei ca. 450–700 °C (je nach Metall) in einer Stickstoffatmosphäre. Dabei diffundieren die Metallatome in das Silizium bzw. umgekehrt (je nach Metall), wobei es zur Silicid-Bildung kommt. Ein wichtiger Aspekt des Salicide-Prozesses ist, dass die Metallionen zwar in das Silizium aber nicht in Siliziumdioxid und -nitrid diffundieren.

Beachten wir, dass durch Metallisierung der Source- und der Substratkontakt kurzgeschlossen werden.

Wenn wir das nicht möchten, müssen wir das Layout modifizieren, indem wir den Substratkontakt vom Transistor trennen.



Ich werde im weiteren Text annehmen, dass die Kontakte keine Silicid-Schicht haben.

### CMOS-Skalierung und Gate-last Prozesse (Extra Thema)

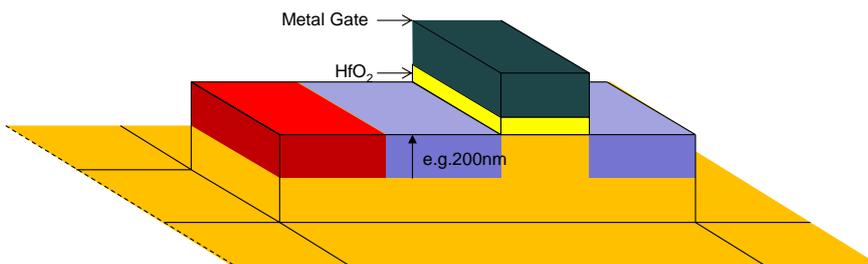
CMOS-Skalierung wird nach dem Prinzip gemacht, dass nicht nur die Transistorgrößen, sondern auch die Spannungen verringert werden.

Kleinere Gate-Source Spannung muss gleiche Ladung im Transistor erzeugen. Deswegen muss die Gate-Kapazität erhöht werden.

Bei einer 65nm Technologie ist die Gate-Isolatordicke nur 2.6 nm und besteht aus  $\text{SiO}_2$ . Gate-Isolator aus  $\text{SiO}_2$  müsste in den Prozessen mit kleineren Transistoren noch dünner werden. Das würde zu viel zu großen Gate-Tunnelströmen führen.

Eine Alternative ist es, nicht die Isolatordicke zu verringern, sondern die dielektrische Konstante zu erhöhen.

Die neuesten Technologien benutzen deshalb ein high K Material als Gate-Isolator, normalerweise ein Oxid wie  $\text{HfO}_2$ . Da  $\text{HfO}_2$  keine gute Verbindung mit Silizium bildet, benutzt man Gate-Elektrode aus Metall.



Bei den neuesten Transistoren ist der Name metal-oxid-semiconductor FET wieder zutreffend, da die Gate-Elektrode aus Metall besteht.

Wir werden die Worte Gate-Isolator und Gate-Oxid weiter als Synonyme benutzen.

Wie wird der Transistor mit dem Gate aus Metall hergestellt?

Es wurde erwähnt, dass die Erzeugung von Source / Drain hohe Temperaturen für Annealing erfordert. Ein Gate aus Metall könnte dabei schmelzen. Man macht deshalb zuerst ein dummy Gate aus Polysilizium. Dieses Gate dient als Maske für Source und Drain. Das dummy Gate

aus Polysilizium wird, nach Herstellung von Source und Drain, weggeätzt. Das finale Gate aus Metall wird dann erzeugt (gate last).

### Funktionsweise des MOSFETs

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

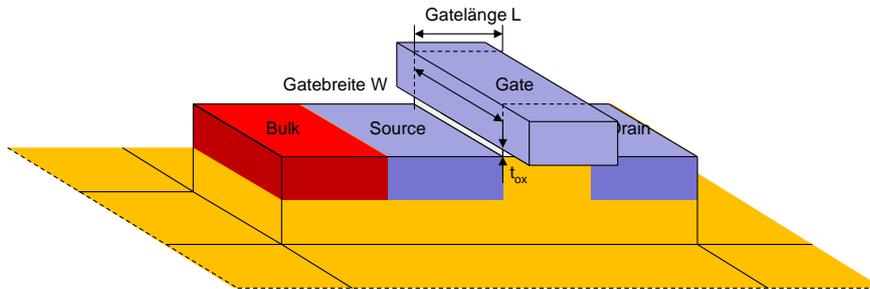


Abbildung 14: Vier Elektroden: Source, Drain, Gate und Substrat (Bulk)

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 – Bulk.

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.

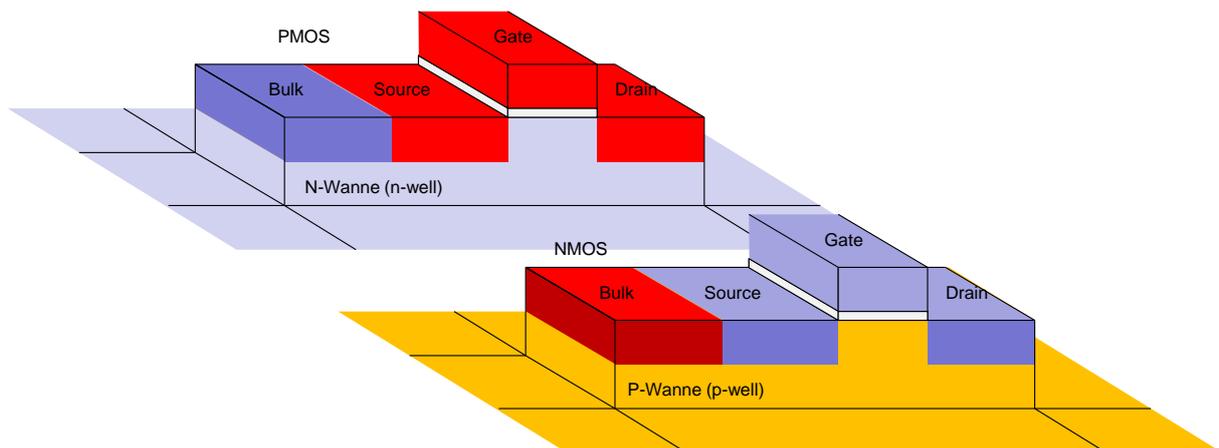


Abbildung 15: NMOS und PMOS

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

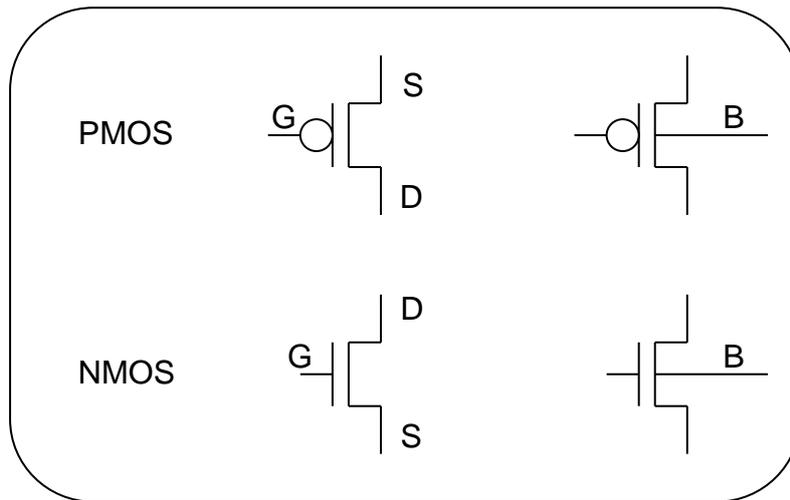


Abbildung 16: Vereinfachte Transistor-Symbole

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.

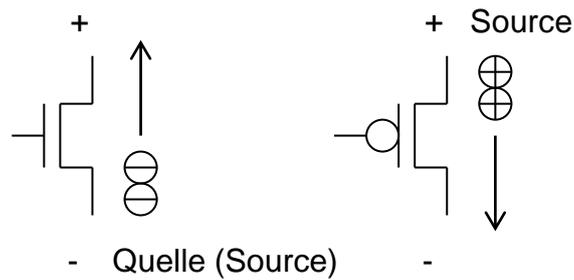


Abbildung 17: Links: NMOS-Source, rechts PMOS-Source

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.

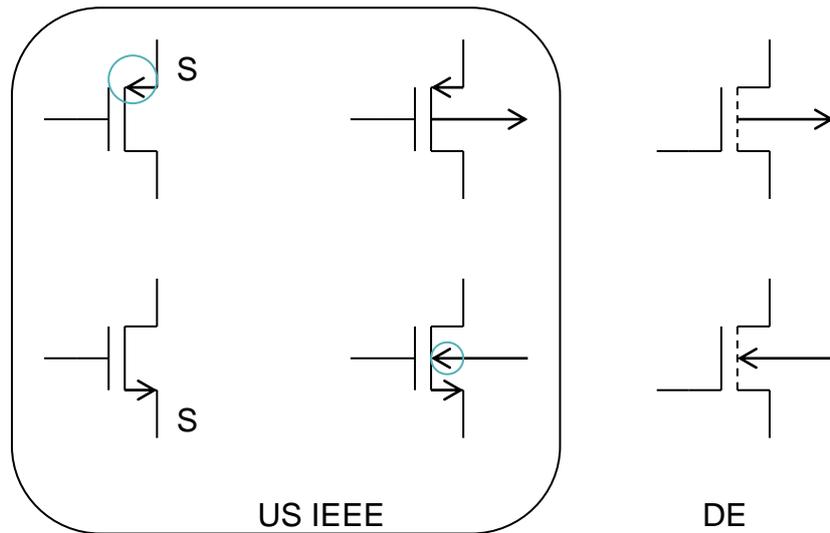


Abbildung 18: Symbole mit Pfeil

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.

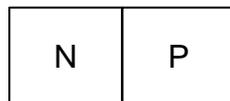
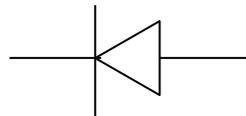


Abbildung 19: Der Pfeil im Symbol einer Diode zeigt von P zu N

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.

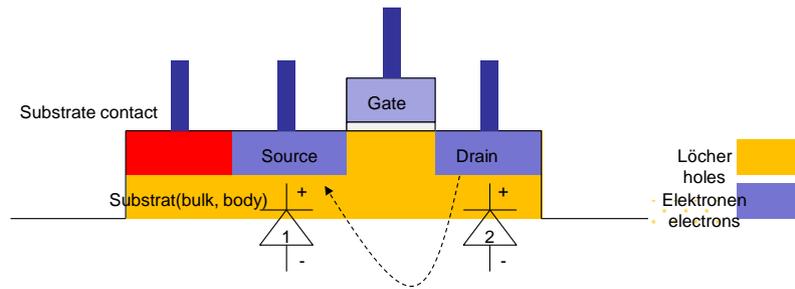


Abbildung 20: Zwei PN-Dioden (Source-Substrat und Drain-Substrat). Die Dioden müssen in Sperrrichtung gepolt werden

### Kontaktspannungen

Silizium und Metall bilden entweder ohmsche Kontakte oder Schottky Dioden. Ein Kontakt, der nur in eine Richtung leitet, wäre für die MOSFET Struktur schlecht.

Kontakte zwischen Metall und *hochdotiertem* Silizium leiten Strom in beide Richtungen, auch dann wenn es sich um Schottky-Dioden handelt. Warum? Die Potentialbarriere (Schottky-Barriere) ist dann sehr dünn. Die Elektronen tunneln durch die Barriere und Strom kann auch in Sperrrichtung fließen. Deswegen wird der Siliziumteil des Substratkontakts hochdotiert (s. Abbildung 21).

Zwischen Silizium und Metallelektroden entstehen Kontaktspannungen.

Die Kontaktspannungen führen dazu, dass sich die Potentiale von Silizium Elektroden ( $V_{g^*}$ ,  $V_{s^*}$ ,  $V_{d^*}$  und  $V_{b^*}$ ) von den Potentialen auf Metallelektroden  $V_g$ ,  $V_s$ ,  $V_d$  und  $V_b$  unterscheiden.

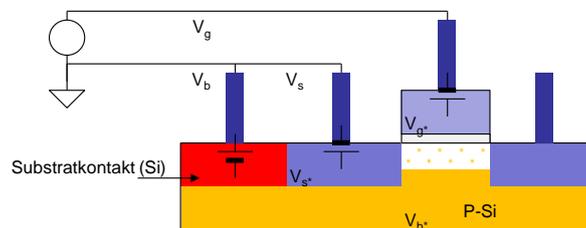


Abbildung 21: Zwischen Silizium und Metallelektroden entstehen Kontaktspannungen

Warum entsteht die Kontaktspannung? (Extra-Thema)

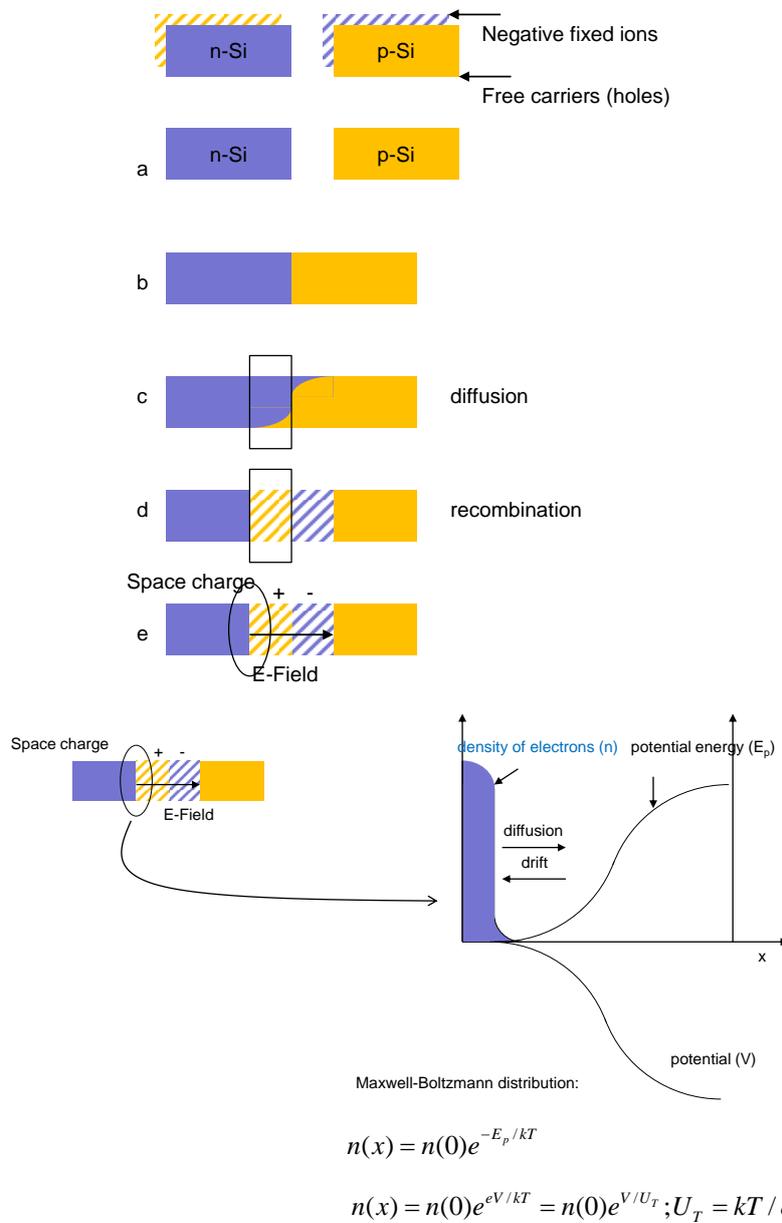


Abbildung 22: Entstehung einer Kontaktspannung

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/Metall mit P-Silizium „verbinden“ (b), entsteht ein Diffusionsstrom von Elektronen in Richtung zum P-Silizium (c). Die Elektronen, die in P-Silizium so gelangen, rekombinieren mit dort vorhandenen Löchern. Die Ladung von negativen Akzeptorionen ist dann nicht mehr durch die positive Ladung der Löcher kompensiert und entsteht eine negativ geladene Zone (d). Auch die Löcher aus P-Silizium diffundieren in das 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 Driftströme die die Diffusionsströme kompensieren – es entsteht Gleichgewichtszustand.

Die Ladungsträgerdichte kann durch die Maxwell-Boltzmann Formel beschrieben werden.

*(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 wie die Dichte von Donatoratomen  $N_d$   
Die Elektronendichte im P-Silizium  $n_p$  ist durch die folgende Gleichung gegeben  $n_p = n_i^2 / N_a$ .

Mithilfe von Maxwell-Boltzmann Formel kann die Kontaktspannung als Funktion von  $n_n$  und  $n_p$  berechnet werden:

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

Eigenleitungsdichte  $n_i$  steigt mit der Temperatur. Deswegen wird die Kontaktspannung kleiner, wenn die Temperatur steigt. Die Kontaktspannungen modellieren wir mit den Spannungsquellen.

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

### Tunneleffekt-Kontakt (Extra-Thema)

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

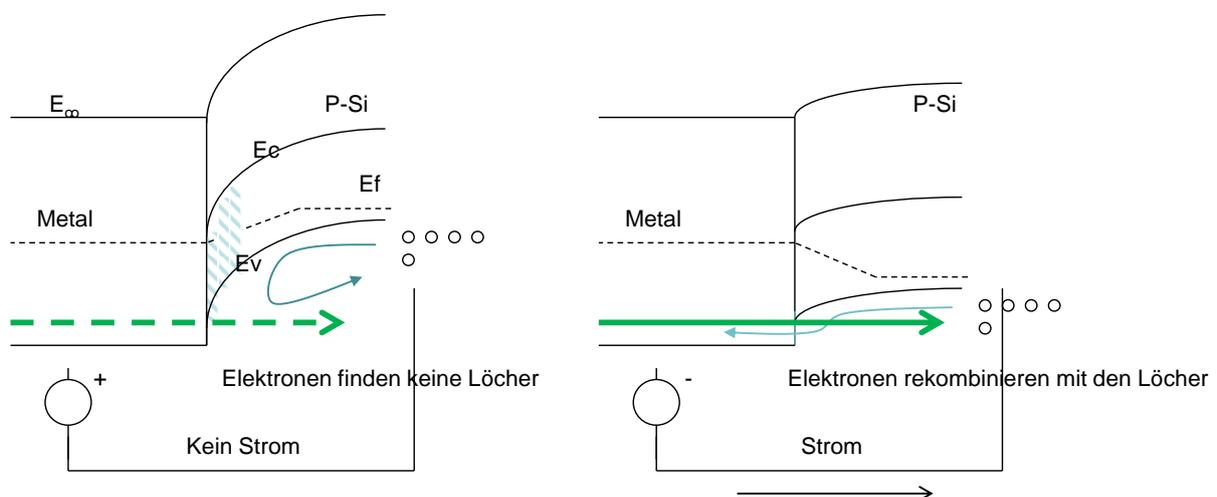


Abbildung 23: Schottky-Diode

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“.

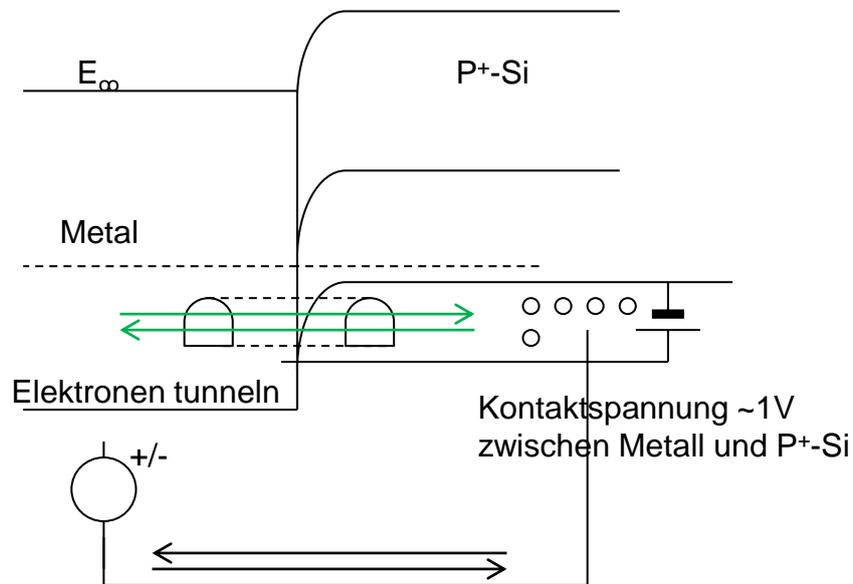


Abbildung 24: Ohmscher Kontakt

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

Folgende Abbildung gibt die Zusammenfassung der wichtigen Formeln für Halbleitertechnik. (Extra Thema)

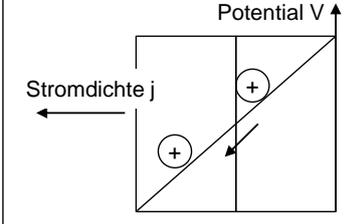
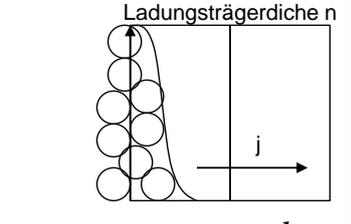
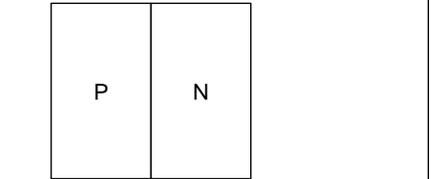
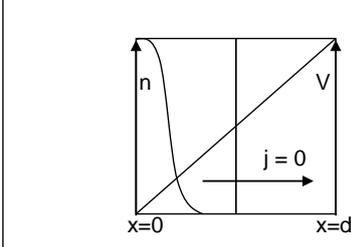
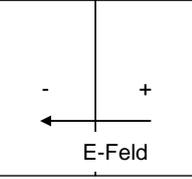
<p><b>Drift</b></p>  <p>Driftstrom <math>j =  q n\mu E</math></p>	<p><b>Diffusion</b></p>  <p>Diffusionsstrom <math>j = -qD \frac{dn}{dx}</math></p>	<p><b>Ladungsträger im Halbleiter</b></p>  <p>(Haupt-)Ladungsträgerdichten  <math>p_p = N_a</math> <math>n_n = N_d</math>  <math>np = n_i^2</math> <math>N_i</math>: Eigenleitungsdichte  <math>p_n = n_i^2 / N_a</math> <math>n_p = n_i^2 / N_d</math>                  Minoritätsladungsträger</p>
<p><b>Gleichgewichtszustand</b></p>  <p>Kontaktspannung  <math>U_T \ln \frac{n(0)}{n(d)} = \pm V</math>  <math>n(x) = n(0)e^{\pm V/U_T}</math>                  Maxwell-Boltzmann Formel  <math>D/\mu = kT/e = U_t</math>                  Einsteingleichung <span style="margin-left: 100px;">U<sub>t</sub> - Thermische Spannung</span></p>		
<p><b>E-Feld, Spannung</b></p> <p>Gaußsches Gesetz <math>E = -\frac{dV}{dx}</math>  <math>\frac{dE}{dx} = \frac{Q}{\epsilon}</math>                  Dicke der Verarmungszone   <math>d = \sqrt{\frac{2\epsilon V}{eN_a}}</math></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 ( $+1.6 \cdot 10^{-19} \text{ C}$ ),  $\mu$  ist Beweglichkeit der Ladungsträger,  $D$  ist die Diffusionskonstante,  $\epsilon$  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 einen Gleichgewichtszustand. Die Ladungsträgerdichte ist innerhalb eines Energiebandes durch die 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 die Ladung  $Q$  das E-Feld erzeugt. E-feld ist die Änderung (Gradient) des Potentials. (Abbildung 25 E-Feld, Spannung)



## Potentialverlauf in der MOSFET Struktur

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 das 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.

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.

Deswegen

$$V_g^* = V_g$$

$$V_s^* = V_s$$

$$V_{gs}^* = V_{gs} \quad (2)$$

Es bleibt nur die Kontaktspannung zwischen Substratkontakt und Substrat.

Es folgt:

$$V_b^* = V_b - V_{\text{cont,np}}; V_{\text{cont,np}} = U_T \ln \left( \frac{N_a N_d}{n_i^2} \right) \quad (3)$$

$N_a$  und  $N_d$  sind die Dichten von Akzeptoren im Substrat (im Kanalbereich) und Donatoren im Source/Gate/Drain,  $n_i$  ist die Eigenleitungsdichte, etwa  $10^{10} \text{ cm}^{-3}$  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:

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

Abbildung 26 zeigt die Potentiale innerhalb des Transistors.

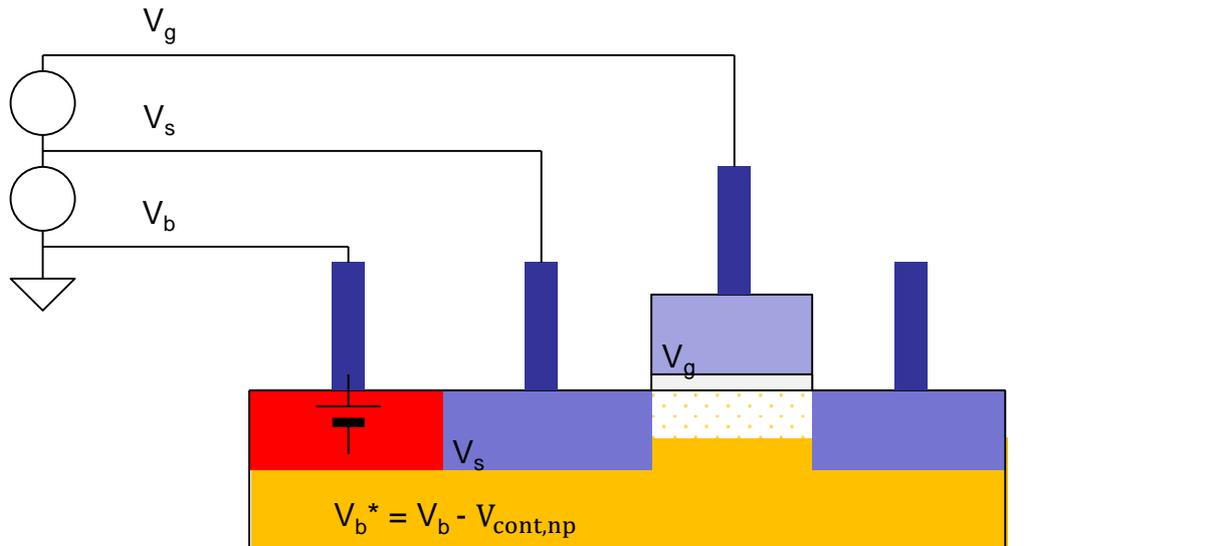


Abbildung 26: Potentiale in verschiedenen Bereichen der MOS-Struktur

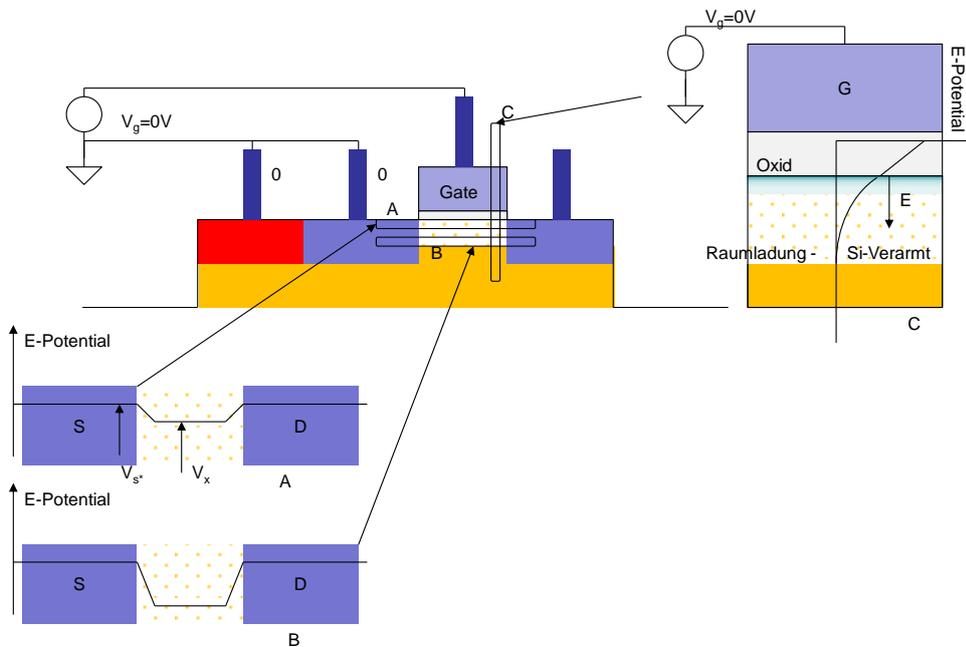


Abbildung 27: Potentiale in verschiedenen Bereichen der MOS-Struktur.

Betrachten wir die Struktur in Abbildung 27. Elektronen sammeln sich in den Bereichen mit höherem elektrischem 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.

Definieren wir die Ausschnitte A, B, C (zwei horizontale Ausschnitte A und B und ein vertikaler C) in der Transistorstruktur (Abbildung 27).

### Vertikaler Potentialverlauf

Betrachten wir die Raumladungzone im P-Bereich (Ausschnitt C). Die Veränderung des Potentials im Substrat führt dazu, dass die Löcherdichte sehr schnell von  $N_a$  auf  $\ll N_a$  fällt.

(Die Löcher werden nach unten verdrängt.) Es entsteht eine Verarmungszone (Raumladungszone) im Substrat, wo die negative Ladung der Akzeptor-Ionen lokal nicht kompensiert ist (Abbildung 28). 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 Silizium-Isolator Grenze ist, da dort das elektrische Potential am größten ist. Wenn wir eine positive Spannung an Gate anlegen (Abbildung 28, unten), werden die Löcher aus dem Substrat noch weiter verdrängt. Die Verarmungszone (Sperrschicht) erweitert sich. Die Elektronendichte an Silizium-Isolator Grenze steigt.

Wir nennen den Effekt, wenn sich in einem P-Siliziumbereich N-Ladungsträger ansammeln „Inversion“. Aus einem p-Bereich wird ein n-Bereich.

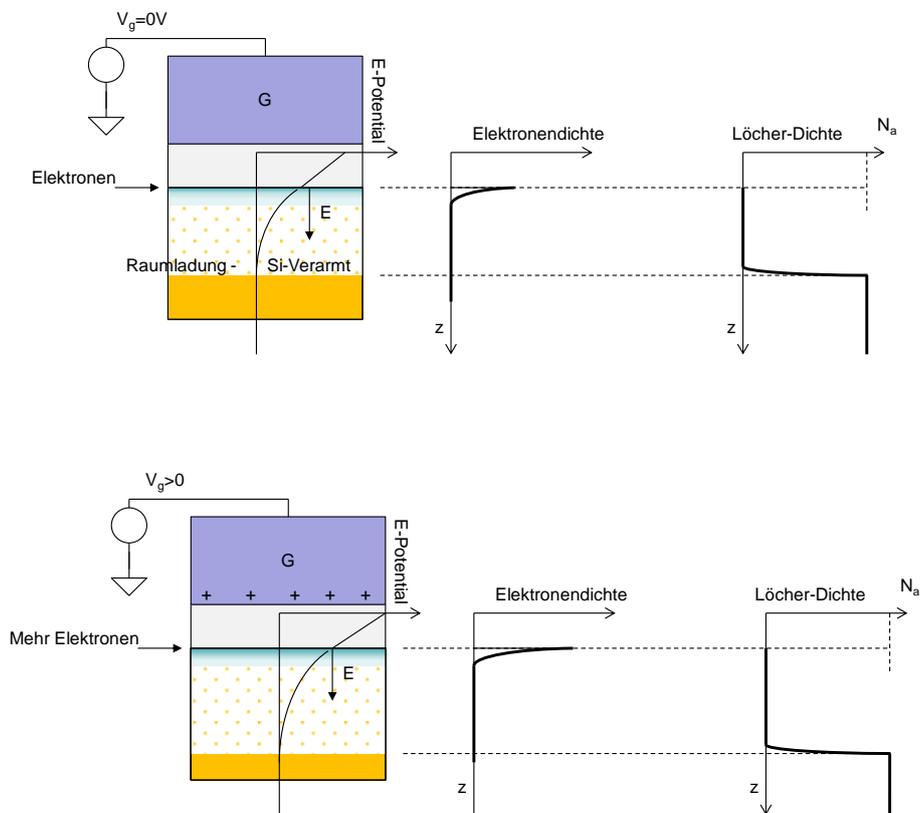


Abbildung 28: Wenn wir eine positive Spannung an Gate anlegen, werden die Löcher aus dem Substrat noch weiter verdrängt

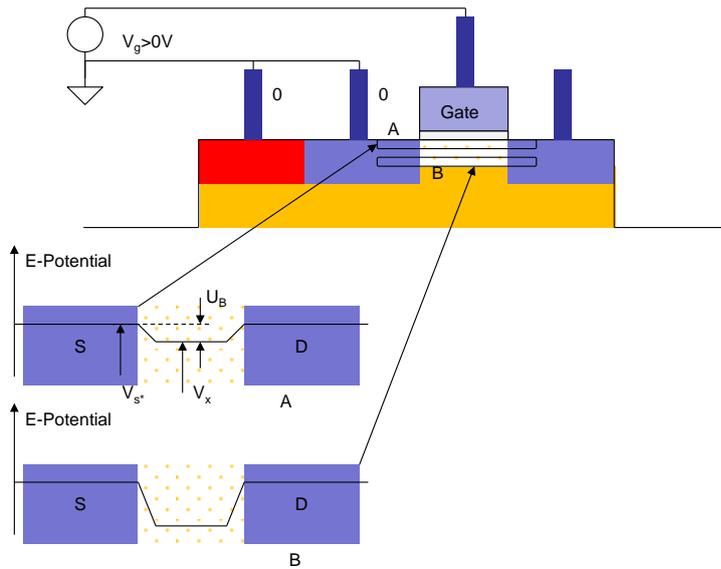


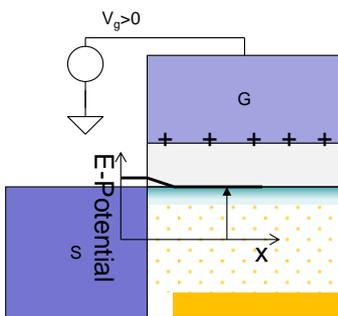
Abbildung 29: Horizontale Ausschnitte, Potentialbarriere

### Horizontaler Potentialverlauf

Betrachten wir die Ausschnitte A und B (Abbildung 29). Eine positive Potentialänderung ist eine Barriere für die positive Ladung (Löcher) und eine negative Potentialänderung ist die Barriere für Elektronen. Die Raumladungszone unter dem Gate stellt eine Potentialbarriere für Elektronen dar. Ein Elektron kann schwer von Source zum Drain durch die tiefere P-Silizium Schichten gelangen, da dort die Potentialbarriere groß ist. Wenn sich Elektron knapp unterhalb der Isolator-Schicht bewegt, sieht es die kleinste Potentialbarriere. In diesem flachen Bereich entsteht der Source-Drain Strom. Man nennt es Kanal. Die Höhe der Potentialbarriere ist:

$$U_B = |V_x - V_{s^*}|$$

$V_x$  ist das Potential des Siliziums unterhalb des Isolators.



Man kann zeigen, dass nicht nur das Gate Potential die Dichte von Elektronen im Kanal beeinflusst, sondern auch das Source Potential. **Elektronendichte im Kanal hängt von der Höhe der Barriere  $U_B$  ab.** Das mag seltsam aussehen. Wir haben gesagt die Elektronendichte durch Maxwell-Boltzmann Formel approximiert werden kann. Das Potential  $V_x$  hängt in der Regel nicht von  $V_s$  ab. Aber, die die Dichte der Elektronen am unteren Ende der Verarmungszone im

Substrat hängt von  $V_s$  ab. Deswegen wird auch die Elektronendichte unterhalb des Isolators von  $V_s$  beeinflusst.

Die Ladungsträgerdichte an der Silicium-Oxid Grenze ist für  $U_B < 0$  durch folgende Formel beschrieben:

$$n = n_0 e^{-U_B/U_T}$$

$n_0$  ist die Elektronendichte in Source. Diese ist gleich wie die Dichte der Donator-Ionen:  $n_0 = N_d$ .

### Potentialbarriere als Funktion von $V_{dep}$

Wir werden im nachfolgenden Text den Transistorstrom als Funktion von  $V_{gs}$  herleiten.

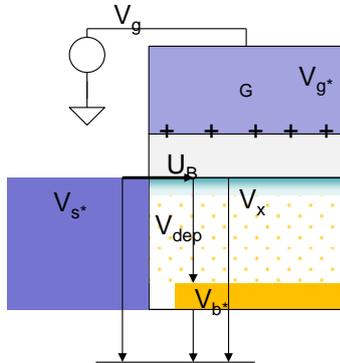


Abbildung 30

Um den Strom zu berechnen, sollen wir zuerst die Elektronendichte im Kanal bestimmen. Die Elektronendichte hängt von der Höhe der Potentialbarriere. Diese hängt von  $V_{gs}$ , und  $V_{sb}$  ab. Wir leiten es in mehreren Schritten her:

$$-U_B = V_x - V_{s*} \quad (4)$$

Es gilt (s. Abbildung 30):

$$V_x = V_{b*} + V_{dep} \quad (5)$$

$V_{b*}$  ist das Potential im Substrat (p-Silizium),  $V_{dep}$  ist die Potentialänderung innerhalb der Verarmungszone.

### Kapazitäten $C_{ox}$ und $C_{dep}$

Um  $V_{dep}$  als Funktion von  $V_g$  zu berechnen, werden wir die MOS-Struktur mit einer Ersatzschaltung modellieren.

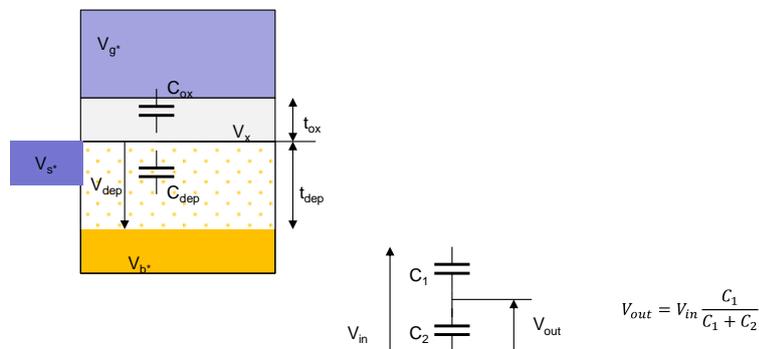


Abbildung 31: Die Kapazitäten in der MOS-Struktur bilden einen Spannungsteiler. Wir haben zwei Kapazitäten 1) die Oxidkapazität  $C_{ox}$  – sie ist durch die Oxiddicke bestimmt und 2) die Kapazität der Verarmungszone  $C_{dep}$ .  $C_{ox}$  und  $C_{dep}$  bilden einen Kapazitiven Spannungsteiler.

Warum?

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 31.  $C_{ox}$  ist als  $Q_{gate}/V_{ox}$  definiert ( $V_{ox}$  ist die Spannung im Gate-Oxid). Es gilt:

$$C_{ox} = \epsilon_0 \epsilon_{SiO_2} \frac{A}{t_{ox}} \sim 8.854 \cdot 10^{-12} \frac{As}{Vm} \times 3.9 \times \frac{A}{t_{ox}} \quad (6)$$

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

$C_{dep}$  kann als normale DC Kapazität,  $C_{dep,dc} = Q_{dep}/V_{dep}$  oder als dynamische AC Kapazität  $C_{dep,ac} = dQ_{dep}/dV_{dep}$  definiert werden. In beiden Fällen hängt  $C_{dep}$  von  $V_{dep}$  ab.

Die AC-Kapazität ist mit folgender Formel beschreiben (Beweis im Kasten):

$$C_{dep,ac} \equiv \frac{dQ_{dep}}{dV_{dep}} = \epsilon_0 \epsilon_{Si} \frac{A}{t_{dep}} \quad (7)$$

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

Man kann zeigen (s. Beweis unten) dass:

$$C_{dep,dc} = Q_{dep}/V_{dep} = 2 C_{dep,ac} \quad (8)$$

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:

$$\frac{dE_z}{dz} = -\frac{eN_a}{\epsilon_0 \epsilon_{Si}} \Rightarrow E_z = -\frac{eN_a}{\epsilon_0 \epsilon_{Si}} z \quad (B1)$$

Potentialgleichung:

$$-\frac{dV_z}{dz} = E_z \Rightarrow V_z = \frac{eN_a}{\epsilon_0 \epsilon_{Si}} \frac{z^2}{2} \quad (B2)$$

Daraus folgt:

$$t_{dep} = \sqrt{\frac{2\epsilon_0 \epsilon_{Si} V_{dep}}{eN_a}} \quad (B3)$$

Die Ladung in Verarmungszone ist (Absolutwert):

$$Q_{\text{dep}} = AeN_a t_{\text{dep}} = A\sqrt{eN_a 2\epsilon_0 \epsilon_{\text{Si}} V_{\text{dep}}} \quad (\text{B4})$$

Die dynamische Kapazität der Verarmungszone ist:

$$C_{\text{dep,ac}} \equiv \frac{dQ_{\text{dep}}}{dV_{\text{dep}}} = A\sqrt{\frac{eN_a \epsilon_0 \epsilon_{\text{Si}}}{2V_{\text{dep}}}} \quad (\text{B5})$$

Es gilt also, wegen (B5) und (B3):

$$C_{\text{dep,ac}} = A\frac{\epsilon_0 \epsilon_{\text{Si}}}{t_{\text{dep}}} \quad (\text{B6})$$

Es gilt auch:

$$Q_{\text{dep}} = A\sqrt{eN_a 2\epsilon_0 \epsilon_{\text{Si}} V_{\text{dep}}} = V_{\text{dep}} A\sqrt{\frac{2\epsilon_0 \epsilon_{\text{Si}} eN_a}{V_{\text{dep}}}} = 2C_{\text{dep,ac}} V_{\text{dep}} \quad (\text{B7})$$

und

$$C_{\text{dep,dc}} = \frac{Q_{\text{dep}}}{V_{\text{dep}}} = 2C_{\text{dep,ac}} \quad (\text{B8})$$

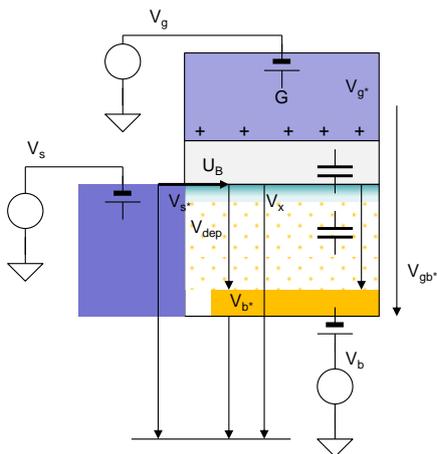
Die Dotierung des Siliziums und die Oxiddicke werden üblicherweise so gewählt, dass  $C_{\text{ox}}$  etwa  $4 \times$  größer als  $C_{\text{dep,ac}}$  ist.

Es gilt etwa:  $C_{\text{ox}} = 4 \times C_{\text{dep,ac}}$  (5) und  $C_{\text{ox}} = 2 \times C_{\text{dep,dc}}$

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

$$\mathbf{n = (C_{\text{dep,ac}} + C_{\text{ox}})/C_{\text{ox}} \sim 1.25 \text{ (9)}}$$

**Potentialbarriere als Funktion von  $V_{\text{gs}}$**



Berechnen wir nun die Barrieren-Spannung:  $-U_B = V_x - V_{s^*}$  als Funktion von  $V_{\text{gs}}$

Es gilt:

$$-U_B = V_x - V_{s^*} = V_{b^*} + V_{\text{dep}} - V_{s^*}$$

Wir können die Formel für Spannungsteiler verwenden. Daraus folgt:

$$V_{\text{dep}} = V_{\text{gb}^*} \frac{C_{\text{ox}}}{C_{\text{dep}} + C_{\text{ox}}}$$

und dementsprechend:

$$-U_{\text{B}} = V_{\text{b}^*} + V_{\text{gb}^*} \frac{C_{\text{ox}}}{C_{\text{ox}} + C_{\text{dep}}} - V_{\text{S}^*} =$$

$$V_{\text{g}^*} \frac{C_{\text{ox}}}{C_{\text{ox}} + C_{\text{dep}}} + V_{\text{b}^*} \frac{C_{\text{dep}}}{C_{\text{ox}} + C_{\text{dep}}} - V_{\text{S}^*} =$$

$$V_{\text{gs}^*} \frac{C_{\text{ox}}}{C_{\text{ox}} + C_{\text{dep}}} + V_{\text{bs}^*} \frac{C_{\text{dep}}}{C_{\text{ox}} + C_{\text{dep}}} =$$

$$\frac{C_{\text{ox}}}{C_{\text{ox}} + C_{\text{dep}}} \left( V_{\text{gs}^*} - \frac{C_{\text{dep}}}{C_{\text{ox}}} V_{\text{sb}^*} \right)$$

Wegen (2) und (3)

$$-U_{\text{B}} = \frac{C_{\text{ox}}}{C_{\text{ox}} + C_{\text{dep}}} \left( V_{\text{gs}} - \frac{C_{\text{dep}}}{C_{\text{ox}}} (V_{\text{cont,np}} + V_{\text{sb}}) \right) \quad (10)$$

### Schwellespannung

Wir definieren die Schwellespannung als  $V_{\text{gs}}$ , bei der es keine Barriere mehr gibt. Fangen wir mit (10) an:

$$-U_{\text{B}} = \frac{C_{\text{ox}}}{C_{\text{ox}} + C_{\text{dep}}} \left( V_{\text{gs}} - \frac{C_{\text{dep}}}{C_{\text{ox}}} (V_{\text{cont,np}} + V_{\text{sb}}) \right)$$

Es folgt:

$$V_{\text{gs}} = \frac{C_{\text{dep}}}{C_{\text{ox}}} (V_{\text{cont,np}} + V_{\text{sb}}) \Rightarrow U_{\text{B}} = 0$$

Wir bezeichnen den Schwellenwert für  $V_{\text{sb}}$  ungleich 0 als  $V_{\text{thsb}}$ .

$$V_{\text{thsb}} \equiv \frac{C_{\text{dep}}}{C_{\text{ox}}} (V_{\text{cont,np}} + V_{\text{sb}}); C_{\text{dep}} = C_{\text{dep,dc}} \text{ für } V_{\text{dep}} = V_{\text{cont,np}} + V_{\text{sb}}$$

Man kann die Schwelle bei  $V_{\text{sb}} = 0$  als „Nullwert“  $V_{\text{th}}$  verstehen und für  $V_{\text{thsb}}$  folgende Approximation schreiben:

$$V_{\text{thsb}} \sim \frac{C_{\text{dep,dc}}}{C_{\text{ox}}} V_{\text{cont,np}} + \frac{C_{\text{dep,ac}}}{C_{\text{ox}}} V_{\text{sb}}; C_{\text{dep,ac}} C_{\text{dep,dc}} \text{ für } V_{\text{dep}} = V_{\text{cont,np}}$$

Oder

$$V_{\text{thsb}} \sim V_{\text{th}} + (n - 1) V_{\text{sb}}; V_{\text{th}} \equiv \frac{C_{\text{dep,dc}}}{C_{\text{ox}}} V_{\text{cont,np}} \quad (11)$$

Es gilt dann

$$-U_B = V_x - V_{s^*} = \frac{C_{ox}}{C_{ox} + C_{dep}} (V_{gs} - V_{thsb}) \quad (12)$$

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

Es ist (siehe Gleichung B5):

$$C_{dep,ac} = A \sqrt{\frac{eN_a \epsilon_0 \epsilon_{Si}}{2V_{cont}}} \quad (13)$$

Deswegen ist es:

$$V_{th} = \frac{C_{dep,dc}}{C_{ox}} \times V_{cont} = \frac{2C_{dep,ac}}{C_{ox}} \times V_{cont} = \frac{A\sqrt{2eN_a \epsilon_0 \epsilon_{Si} V_{cont}}}{C_{ox}} \quad (14)$$

Die Kontaktstension ist mit (3) beschrieben. Da wir  $N_d = N_a$  angenommen haben, gilt:

$$V_{cont} = 2U_T \ln\left(\frac{N_a}{n_i}\right) \quad (15)$$

Formel (14) wird normalerweise in Literatur hergeleitet.

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

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 (14), da das Kontaktpotential sinkt (15). Es ist die Folge des Anstiegs von Eigenleitungsdichte in Silizium  $n_i$ .

### Starke Inversion

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 32).

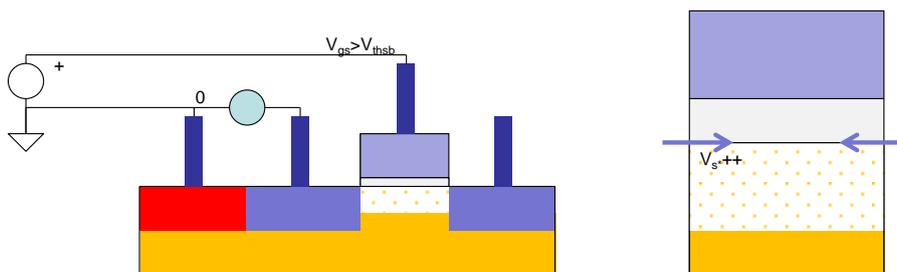


Abbildung 32: Potentiale in der MOS Struktur für  $V_{gs} > V_{thsb}$

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 33). Wir nennen solch einen Arbeitsbereich „starke Inversion“.

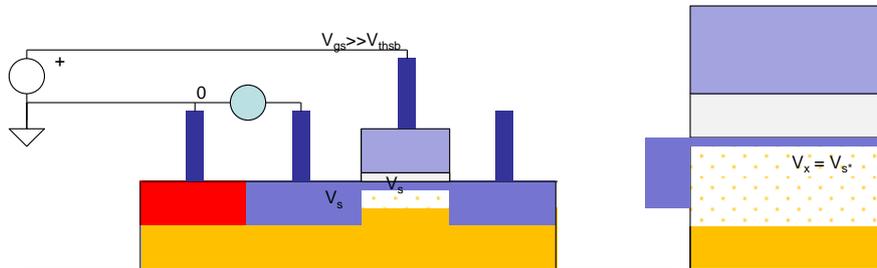


Abbildung 33: Elektronen aus der Source und dem Drain bilden einen leitenden Kanal. Source, Drain und die Substratoberfläche werden kurzgeschlossen ->  $V_x = 0$

### Ladung im Kanal

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  $dV_{gs}$  ändert, fließt die Ladung  $C_{ox} dV_{gs}$  durch die Spannungsquelle. Genau dieselbe Ladung bildet sich im Kanal.

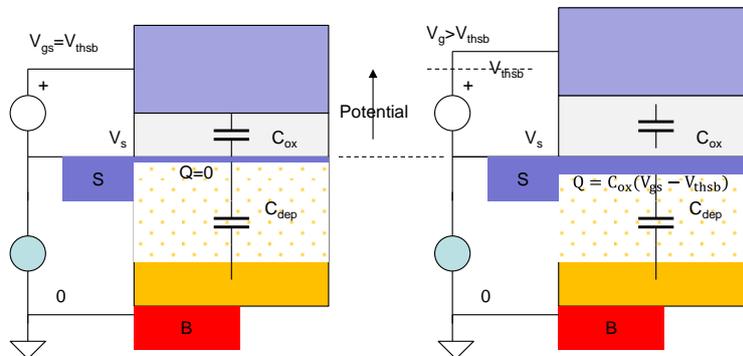


Abbildung 34: Ladung im Kanal

Wir machen folgende Annahme:

Für  $V_{gs} = V_{thsb}$  gibt es noch keine Ladung im Kanal.

Diese Annahme ist nicht für jede Anwendung korrekt (s. nächste Vorlesung – „schwache Inversion“)

Für  $V_{gs} > V_{thsb}$  gilt  $dQ = C_{ox} dV_{gs}$ . Daraus folgt für die Ladung im Kanal:

$$Q = C_{ox}(V_{gs} - V_{thsb}) \quad (16)$$

Zusammenfassung:

Wir definieren die Schwellenspannung in den Nähe von Source  $V_{thsb}$  (11) als die Gate-Source Spannung für welche die Potentialen in Source und auf der Substratoberfläche (Si-Oxid Grenze) etwa gleich sind.

Wenn die Gate-Source Spannung über die Schwelle steigt, sammeln sich die Elektronen im Kanal. Wir reden dann von einer starken Inversion.

## Drain-Strom Strom als Funktion von Drain-Source Spannung

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

Falls  $V_{gs} > V_{thsb}$  entsteht ein Kanal aus Elektronen zwischen der Source und dem Drain, also eine Ohmsche Verbindung oder einen Widerstand. (Abbildung 35).

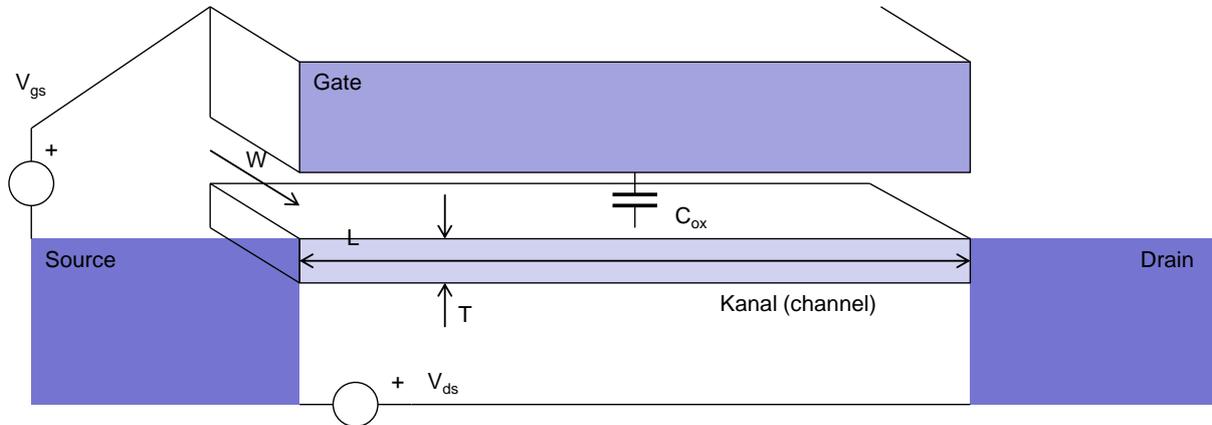


Abbildung 35: Ohmsche Verbindung zwischen der Source und dem Drain

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 35):

$$I_{ds} = A e \mu n E \quad (17),$$

$e$  ist die elementare Ladung,  $\mu$  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:

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

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

Wir bekommen:

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

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

E-Feld ist in erster Näherung:

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

Deshalb gilt:

$$I_{ds} = \mu C'_{ox} \frac{W}{L} (V_{gs} - V_{thsb}) V_{ds} \quad (18).$$

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:

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

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

Wegen Symmetrie ist die Kanalladung in der Nähe von Source

$$Q'(0) = C_{ox} (V_{gs} - V_{thsb})$$

und in der Nähe von Drain

$$Q'(L) = C_{ox} (V_{gd} - V_{thdb}).$$

Es gilt

$$V_{thdb} = V_{th} + (n - 1)V_d = V_{thsb} + (n - 1)V_{ds}$$

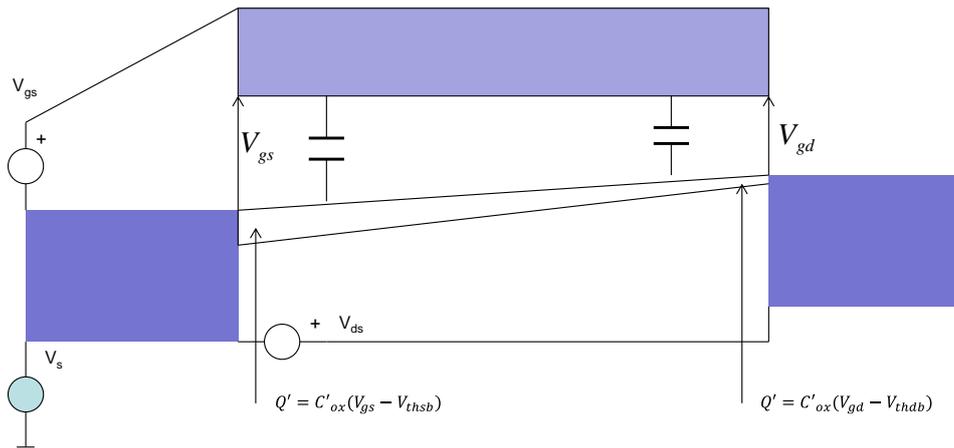


Abbildung 36: Strom als Funktion von  $V_{ds}$ .  $V_{ds} > 0$

Deshalb:

$$Q'(L) = C'_{ox}(V_{gd} - V_{thdb}) = C'_{ox}(V_{gst} - nV_{ds}) ; V_{gst} \equiv V_{gs} - V_{thsb}$$

Für  $V_{gst} = nV_{ds}$  haben wir praktisch keinen Kanal an der Drain-Seite mehr. Wir sagen, dass der Kanal abgeschnürt ist (pinched-off). Der Strom fließt, weil das E-Feld an der Drain Seite groß ist.

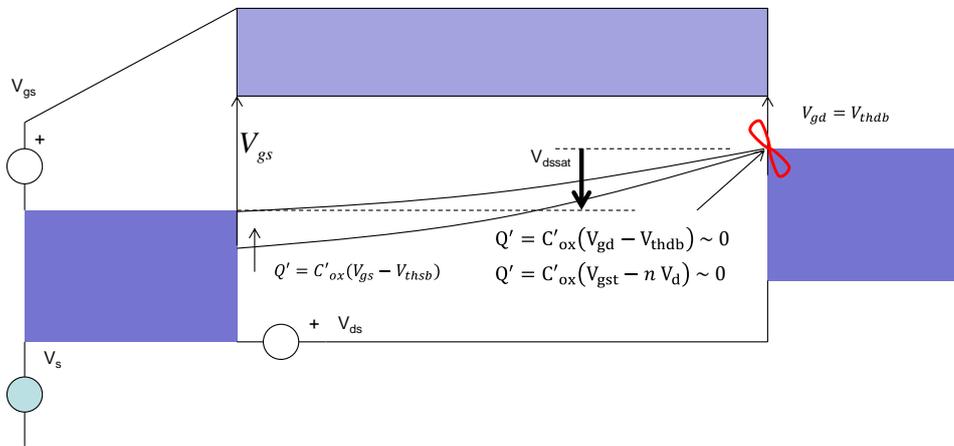


Abbildung 37: Strom als Funktion von  $V_{ds}$ .  $V_{ds} = V_{dssat}$



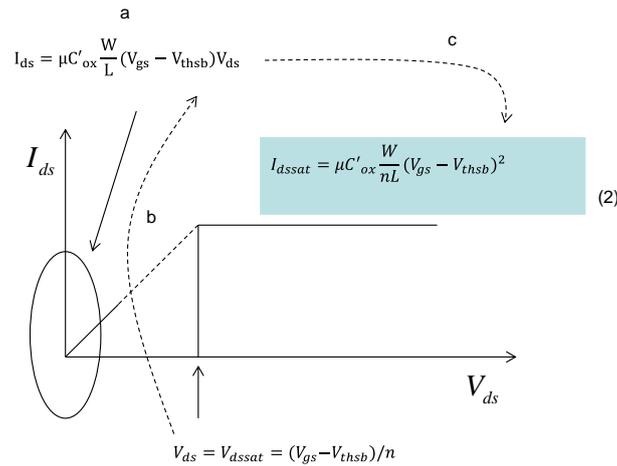


Abbildung 40:  $I_{ds}$  Formel – einfache Herleitung

Leider ist die Annahme (A1) nicht ganz korrekt. Der Stromanstieg ist für  $V_{ds} > \sim 100\text{mV}$  schwächer als die Formel (20) gibt.

Wenn wir die Variation der Kanalladung berücksichtigen, bekommen wir zusätzlichen Faktor  $\frac{1}{2}$ . Die Formel für Sättigungsstrom lautet mit dieser Korrektur:

$$I_{dssat} = \frac{1}{2n} \mu C'_{ox} \frac{W}{L} (V_{gs} - V_{thsb})^2 \quad (21)$$

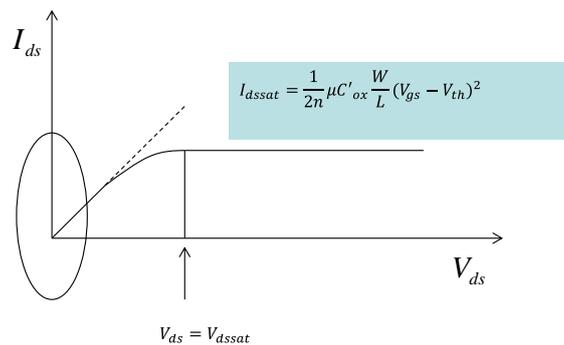


Abbildung 41:  $I_{ds}$  Formel – genauer

Herleitung der Formel (21):

Wir haben folgende Formel für die Kanalladung pro Fläche:

$$Q' = C'_{ox} (V_{gs} - V_{thsb}) \quad (A1)$$

Die Formel gilt in starker Inversion in der Nähe von Source. Die Kanalladung pro Fläche ändert sich von dem Wert gegeben durch (A1) an der Source-Seite bis null an der Drain-Seite, da dort der Kanal abschnürt. Bezeichnen wir das Potential im Kanalbereich (an der Grenze zwischen

Si und Oxid) als  $V_x(x)$ . Koordinate  $x$  verläuft von Source zum Drain. Definieren wir  $V_x$  in Bezug auf das Source-Potential  $V_s$ . Das bedeutet:

$$V_x(0) = 0$$

und

$$V_x(L) = V_{ds}$$

Die Kanalladung pro Fläche in einem beliebigen Punkt  $x$  ist:

$$Q'(x) = C_{ox}(V_{gx} - V_{thxb}) = C_{ox}(V_{gs} - V_{thsb} - nV_x) \quad (A2)$$

Oder vereinfacht

$$Q'(x) = C_{ox}(V_{gst} - nV_x); \quad V_{gst} \equiv V_{gs} - V_{thsb} \quad (A3)$$

Leiten wir zuerst die Gleichung für  $I_{ds}$  Strom für beliebige  $V_{ds}$  her.

Fangen wir mit der Formel für Driftstrom an:

$$I_{ds} = \mu W Q'(x) |E_x| = \mu C'_{ox} W (V_{gst} - nV_x) |E_x| \quad (A4)$$

$E_x$  ist die E-Feld Komponente in  $x$ -Richtung,  $W$  ist Gate-Breite,  $\mu$  ist Beweglichkeit der Ladungsträger.

Es gilt:

$$|E_x| = \frac{dV_x}{dx} \quad (A5)$$

Wenn wir (A5) in (A4) einsetzen, bekommen wir:

$$I_{ds} = \mu C'_{ox} W (V_{gst} - nV_x) \frac{dV_x}{dx}$$

oder

$$I_{ds} dx = \mu C'_{ox} W (V_{gst} - nV_x) dV_x \quad (A6)$$

Wir können die beiden Seiten integrieren:

$$\int_0^L I_{ds} dx = \int_0^{V_{ds}} \mu C'_{ox} W (V_{gst} - nV_x) dV_x$$

Es folgt:

$$I_{ds} = \frac{1}{2n} \mu C'_{ox} \frac{W}{L} \left[ V_{gst}^2 - (V_{gst} - nV_{ds})^2 \right]$$

Nachdem wir den zweiten Ausdruck quadrieren, bekommen wir die allgemeine Formel für Transistorstrom:

$$I_{ds} = \mu C'_{ox} \frac{W}{L} \left[ V_{gst} V_{ds} - n \frac{V_{ds}^2}{2} \right] \quad (A7)$$

In Sättigung gilt:

$$V_{ds} = V_{dssat} = \frac{V_{gst}}{n} \quad (A8)$$

Wenn wir dies in (A7) einsetzen, bekommen wir die Formel für Sättigungsstrom:

$$I_{dssat} = \frac{1}{2n} \mu C'_{ox} \frac{W}{L} V_{gst}^2 \quad (A9)$$

Auch die Formel (A9) / (21) beschreibt die Messergebnisse nicht besonders gut. Genauere Modelle zeigen (Sättigung der Driftgeschwindigkeit), dass der Faktor  $\frac{1}{2}$  durch  $\frac{1}{2n\alpha}$  ersetzt werden soll. Für  $V_{dssat}$  gilt:

$$V_{dssat} = \frac{V_{gs} - V_{thsb}}{n\alpha} \quad (22)$$

$$\alpha = 1 + \frac{V_{gs}}{nE_{sat}L} \quad (23)$$

Faktor  $\alpha$  beschreibt die Sättigung der Driftgeschwindigkeit.

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

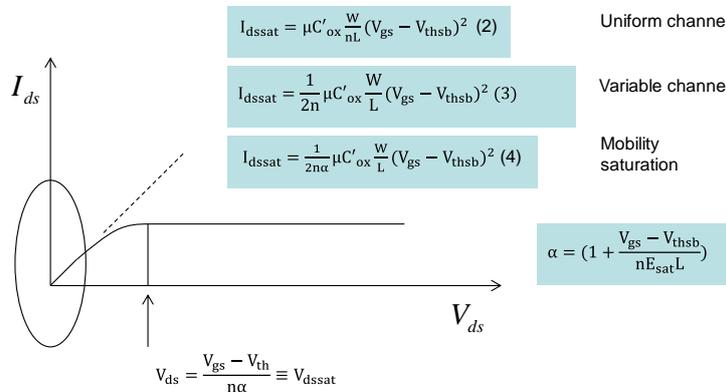


Abbildung 42:  $I_{ds}$  Formel – Überblick