# Lecture 4

Themes:

Channel length modulation (Early-effect)

Weak inversion

BSIM transistor model

## Channel length modulation (Early-Effekt)

We derived in previous lectures that the transistor current depends on the channel length L and the channel width W ($I_{ds}$ ~ W/L). How big are W and L? To the first approximation, the channel occupies the entire area below the gate oxide. Below the gate, the "attraction" of the positive gate charge is strong enough to form an electron channel.

For transistors that are in the linear- (or triode-) region, the channel is approximately of the same size as the gate oxide.

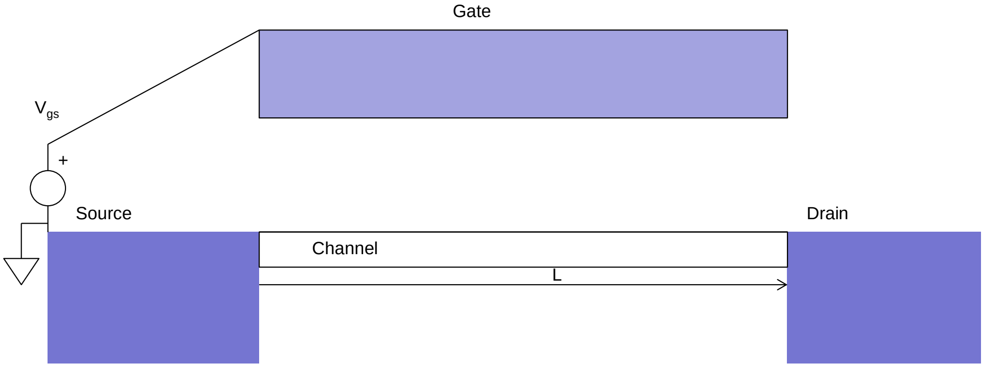



*Figure 1: Channel length*

If $V_{ds}$ is larger than $V_{dssat}$ (transistor is in saturation region), the drain-end of the channel remains nearly at the $V_{dssat}$ potential. Between the drain and the end of the channel, we have a potential difference of $V_{ds} – V_{dssat}$. A depletion zone is formed.

The size of the depletion zone depends on the voltage $V_{ds}$ - $V_{dssat}$. The effective length of the channel is therefore by the size of the depletion zone smaller than the gate oxide length.

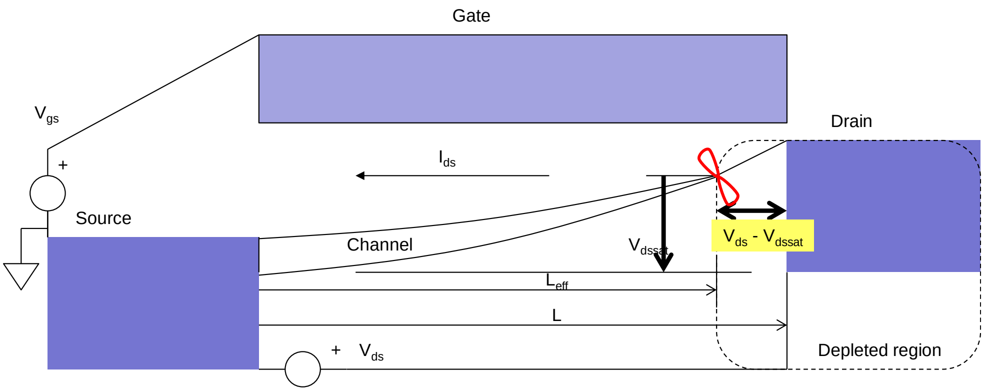



*Figure 2: Channel length modulation*

**Size of the depletion region**

Let us try to calculate the length of the depletion zone.

We can use the same approach as when calculating the depletion zone of a diode. As first, we use the Gauss law to calculate the E-field as a function of the coordinate. We then calculate the voltage from the E field.

The difference in comparison with the diode is that at the edge of the depletion zone (at the point where the channel is pinched off) the E-field is not equal to zero. It is E (0) = $E_{sat}$ (Figure 3).

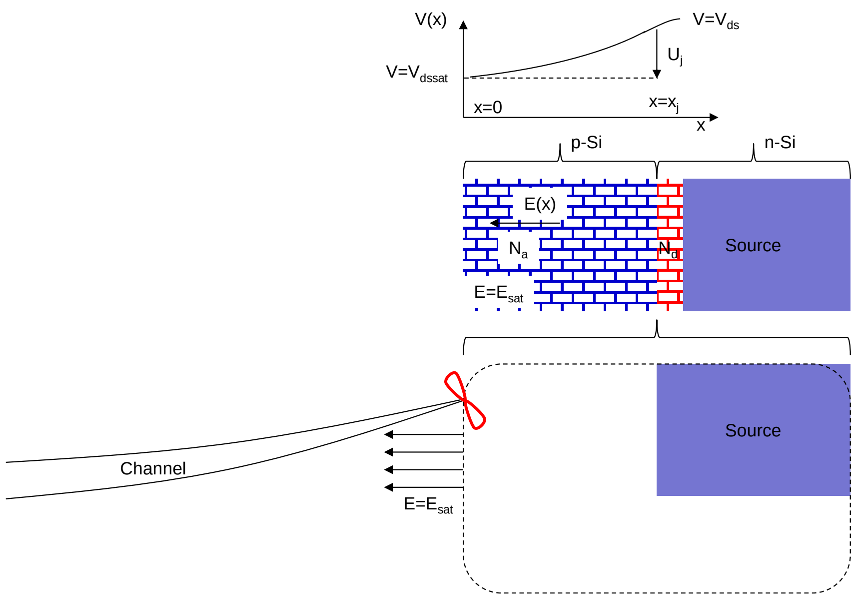



*Figure 3: Calculation of the depleted region length*

The following applies (Gauss law):

$$\frac{dE}{dx} = \frac{eN_a}{\epsilon}$$

E is the horizontal E-field component, x is the coordinate, $N_a$ is the density of the acceptors, $\epsilon$ is the permittivity of silicon, e is the elementary charge. From this, and by taking into account E (0) = $E_{sat}$, it follows:

$$E = \frac{eN_a}{\epsilon} x + E_{sat}$$

The change of potential along the depletion zone $U_j$ is the following integral:

$$U_j = \int_0^{x_j} E\,dx = \frac{eN_a}{\epsilon}\frac{x_j^2}{2} + E_{sat}x_j = V_{ds} - V_{dssat} \quad (A1)$$

$X_j$ is the width of the depletion zone.

**Drain-Source conductance (derivation)**

How does $V_{ds}$ affect the current?

Let us start with the formula for saturation current:

$$I_{dssat} = \frac{1}{2n}\mu C'_{ox}\frac{W}{L_{eff}(V_{ds})}\left(V_{gs} - V_{thsb}\right)^2 \quad (A2)$$

When $V_{ds}$ exceeds $V_{dssat}$, the effective channel length $L_{eff}(V_{ds})$ shortens and the current increases.

We define the drain-source conductance $g_{ds}$ as the slope of the $I_{ds}(V_{ds})$:

$$g_{ds} = \frac{dI_{ds}}{dV_{ds}} = \frac{dI_{dssat}}{dL_{eff}}\frac{dL_{eff}}{dV_{ds}} \quad (A3)$$

By calculating derivative, we get:

$$\frac{dI_{ds}}{dL_{eff}} = -\frac{I_{dssat}}{L_{eff}} \quad (A4)$$

If we put this result in (A3), we get:

$$g_{ds} = -\frac{I_{dssat}}{L_{eff}}\frac{dL_{eff}}{dV_{ds}} \quad (A5)$$

Let us calculate now $\frac{dL_{eff}}{dV_{ds}}$.

The effective length of the channel is:

$$L_{eff} = L - x_j \quad (A6)$$

Therefore:

$$\frac{dL_{eff}}{dV_{ds}} = -\frac{dx_j}{dV_{ds}} \quad (A7)$$

We can calculate the change $dx_j/dV_{ds}$ if we differentiate both sides from A1 ($d/dV_{ds}$):

$$\frac{d}{dV_{ds}}(V_{ds} - V_{dssat}) = \frac{d}{dV_{ds}}\left(\frac{eN_a x_j^2}{2\epsilon} + E_{sat}x_j\right)$$

$$1 = \frac{eN_a x_j}{\epsilon}\frac{dx_j}{dV_{ds}} + E_{sat}\frac{dx_j}{dV_{ds}}$$

From this and from (A7) it follows:

$$\frac{1}{E_{sat}+\frac{eN_a x_j}{\epsilon}} = \frac{dx_j}{dV_{ds}} = -\frac{dL_{eff}}{dV_{ds}} \quad \text{(A8)}$$

If we put A8 in A3, we get the formula for the conductance $g_{ds}$:

$$g_{ds} = \frac{dI_{dssat}}{dV_{ds}} = \frac{I_{dssat}}{L_{eff}(E_{sat}+\frac{eN_a x_j}{\epsilon})} \quad \text{(A9)}$$

$g_{ds}$ increases when $V_{ds}$, and thus $x_j$ increases. At the beginning of saturation, $g_{ds}$ is nearly equal to $I_{dssat}$ / $LE_{sat}$.

**Early voltage (derivation)**

Shortening of the channel leads to the following current change:

$$I_{ds} = I_{dssat} + g_{ds}(V_{ds} - V_{dssat}) = I_{dssat}(1 + \frac{g_{ds}}{I_{dssat}}(V_{ds} - V_{dssat})) \quad \text{(A10)}$$

We define the early voltage $V_A$ as:

$$V_A \equiv \frac{I_{dssat}}{g_{ds}} \quad \text{(A11)}$$

When we put A11 in A10, we get:

$$I_{ds} = I_{dssat}(1 + \frac{V_{ds}-V_{dssat}}{V_A}) \quad \text{(A12)}$$

From A11 and A9, we obtain

$$V_A = L_{eff}E_{sat} + \frac{L_{eff}eN_a\, x_j}{\epsilon} \quad \text{(A13)}$$

The transistors in the 65nm process are described with BSIM4 model.

Contents (iastate.edu)

In this model, the following equation is used:

$$V_A \sim LE_{sat} + L\frac{V_{ds}-V_{dssat}}{0.5\, x_j} \quad \text{(A14)}$$

We can derive this equation using two approximations.

We neglect the term $E_{sat}x_j$ in (A1). It follows:

$$x_j \sim \sqrt{\frac{2\epsilon}{eN_a}(V_{ds} - V_{dssat})}$$

We assume $L_{eff} \sim L$.

In this way, we obtain the equation A14.

The BSIM formula for $g_{ds}$ is:

$$g_{ds} = \frac{I_{dssat}}{V_A} = \frac{I_{dssat}}{L_{eff}(E_{sat} + \frac{V_{ds} - V_{dssat}}{0.5\ x_j})} \quad (A15)$$

**Summary of channel length modulation**

The following equation can be derived:

$$I_{ds} = I_{dssat}(1 + \frac{V_{ds} - V_{dssat}}{V_A})$$

Early voltage $V_A$ is:

$$V_A = L_{eff}E_{sat} + \frac{L_{eff}eN_a\, x_j}{\epsilon}$$

BSIM Formula for Early voltage is

$$V_A \sim LE_{sat} + L\frac{V_{ds} - V_{dssat}}{0.5\, x_j}$$

With

$$x_j = \sqrt{\frac{2\epsilon(V_{ds} - V_{dsssat})}{eN_a}}$$

$E_{sat}$ is the E-field strength at which mobility becomes saturated, $x_j$ is the length of the depletion zone. $N_a$ is the density of the acceptors, $\epsilon$ is the permittivity of silicon, e is the elementary charge.

Drain-Source conductance is:

$$g_{ds} = \frac{I_{dssat}}{V_A} = \frac{I_{dssat}}{L_{eff}(E_{sat} + \frac{V_{ds} - V_{dssat}}{0.5\, x_j})}$$

Figure 4 shows the simulated output characteristics.

The slope $dI_{dssat}/dV_{ds}$ is largest at the beginning of saturation:

$$g_{ds,vds=vdssat} = \frac{I_{dssat}}{L_{eff}\, E_{sat}}$$

For larger $V_{ds}$, $g_{ds}$ decreases (see A15) and the characteristics becomes flatter.

We define the output resistance as:

$$r_{ds} = \frac{1}{g_{ds}}$$

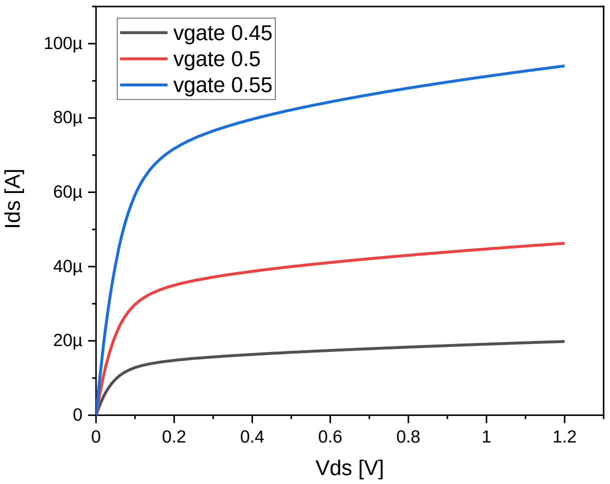



*Figure 4: Early effect leads to current increase for $V_{ds} > V_{dssat}$.*

A small $g_{ds}$ is often an advantage because the transistor behaves more like a current source.

We obtain small $g_{ds}$ values (or large $r_{ds}$ resistance) for long transistors and for small currents (see A9 or A15).

Figure 5 shows a transistor with small $g_{ds}$ and large $r_{ds}$. Such a transistor has small $g_m$. We don't achieve both a good current source and high transconductance.

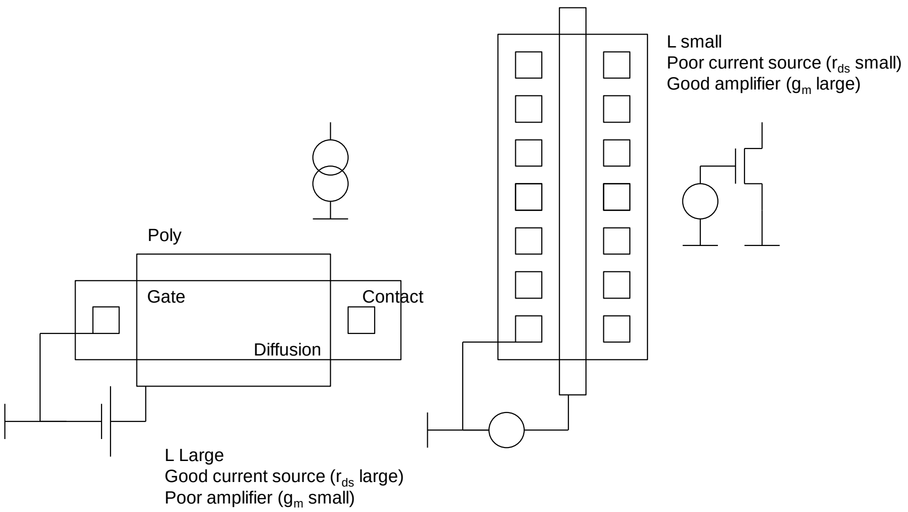



*Figure 5: Two current sources*

There is a trick called cascode (Figure 6) that allows us to maximize both $g_m$ and $r_{ds}$.

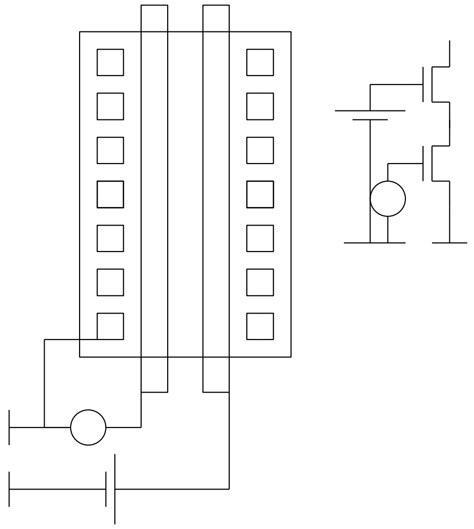


*Figure 6: Current source with cascode*

**Weak inversion**

During previous analysis of MOSFET, we assumed that the channel charge and the $I_{ds}$ current equal zero for $V_{gs} < V_{th}$.

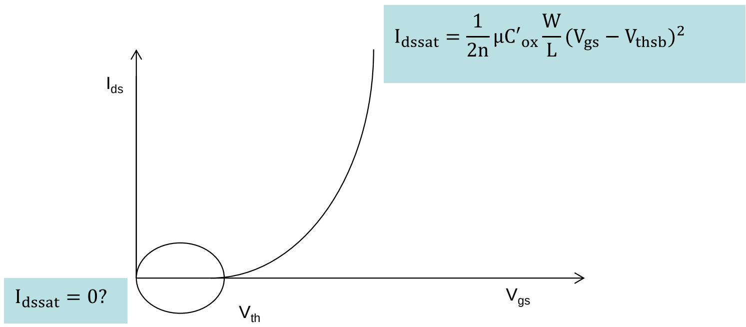
We derived:

$$I_{dssat} = \frac{1}{2n}\mu C_{ox}\frac{W}{L}\left(V_{gs} - V_{thsb}\right)^2 \quad (1)$$

The formulas are based on a simplistic model of channel charge. It is not correct enough for analysis of many low-power circuits.

We will now consider the case when $V_{gs}$ is smaller than $V_{thsb}$.

We call this operation region the weak inversion. In contrast to this, for $V_{gs} > V_{thsb}$ we have a strong inversion.




We mentioned in lecture 2 that the electron density in the channel (at the boundary between the oxide and the silicon) depends on the potential barrier $U_B$ between the channel region and the source/drain. We have also defined the threshold voltage as the $V_{gs}$ voltage leading to $U_B = 0$. For $V_{gs}$ less than the threshold value, the barrier is greater than zero.

If the barrier $U_B$ is small, an electron from the source or from the drain can enter the substrate if it receives additional kinetic energy from the phonons (thermal crystal vibrations). What does *small barrier* mean? This is the barrier, which is roughly equal to the average thermal energy of electrons. Let us remember that the thermal energy at the room temperature (300 K) corresponds to a voltage of $U_T = 26$ mV. ($U_T = kT/e$) Thermal energy is nearly the average kinetic energy of electrons at a certain temperature. (The average kinetic energy is 3/2 kT) Therefore, the electron density is not zero for $V_{gs} < V_{thsb}$ in the channel range. This is shown in Figure 7.

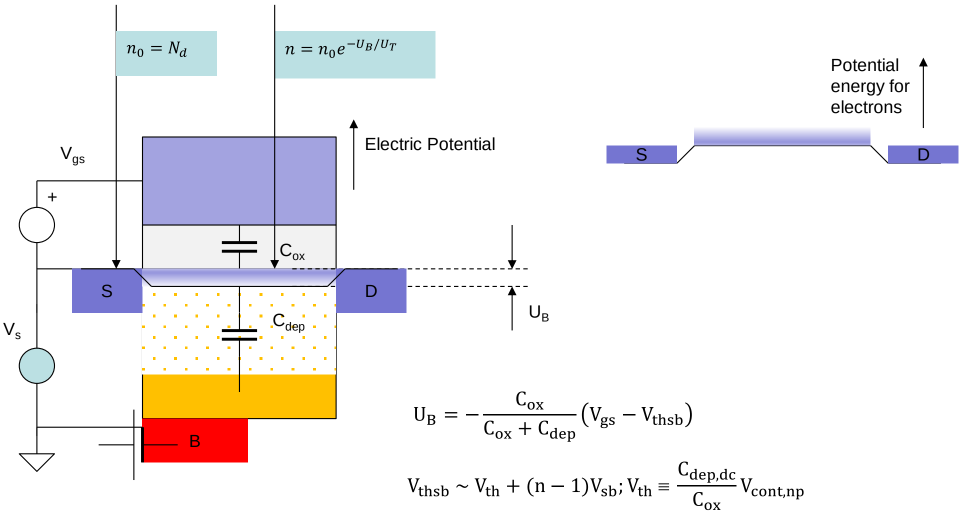



*Figure 7: Charge in the substrate under the oxide depends on barrier $U_B$ . $V_{ds}$ = 0.*

If the drain potential is higher than the source potential, the barrier at the drain side is larger and the density of the electrons decreases there. This is shown in Figure 8. Uneven distribution of electrons leads to diffusion current.

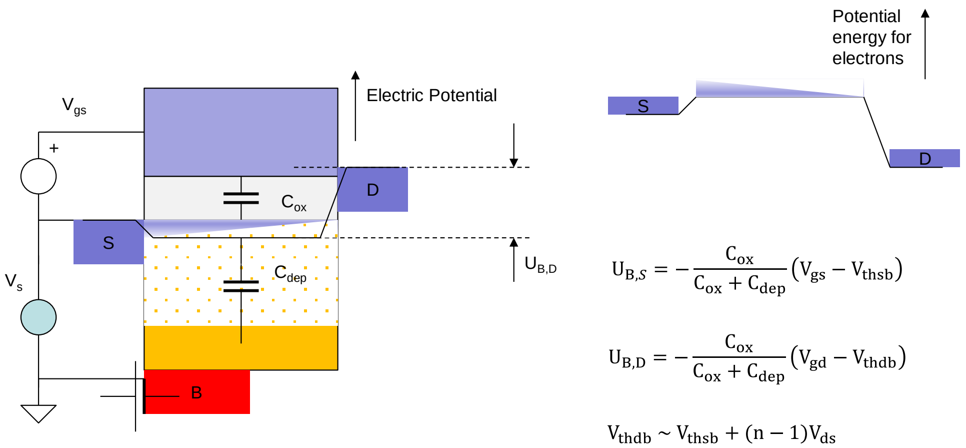



*Figure 8: Potential change and electron density in the substrate under the oxide.* $V_{ds} > 0$.

In the next chapter, we will derive the formula for the drain-source current. This chapter is optional. You can skip it and continue reading starting from the summary.

**Weak inversion – derivation of equations**

The first assumption is that we have a diffusion current. We need to derive the charge carrier density/area[1] in the substrate region between the source and drain. If we know the density we can derive the current using the equation for diffusion current.

We will calculate the density of the charge carriers (in the case of NMOS transistor electrons) in several steps. The first step is the calculation of the electric potential underneath gate oxide as a function of vertical z-coordinate. The second step is the calculation of the charge carrier

[1] The density integrated in z-direction

density as a function of the potential. The third step is the integration of the density function in z-direction. We will not do all steps exactly.

For exact derivation, I refer to document „MOSFET Detailed“.

The height of the potential barrier $U_B$ is (see lecture 2):

$$U_B = -\frac{C_{ox}}{C_{dep}+C_{ox}}(V_{gs} - V_{thsb}) = -\frac{(V_{gs}-V_{thsb})}{n} \quad (3)$$

Now let us calculate the density of the electrons below the oxide. The electron density in Source is $n_0 = N_d$. $N_d$ is the density of the donor atoms. Electron density in thermodynamic equilibrium is described by the Maxwell Boltzmann distribution.

We can assume that the PN junctions between the source and the substrate and between the drain and the substrate are in thermodynamic equilibrium. Therefore, we can use the Maxwell Boltzmann distribution. The electron densities in the source and drain are equal to $N_d$.

For this reason, the electron density at the substrate surface (in the “channel region”) is given by the following equation:

$$n = N_d e^{-U_B/U_T} \quad (4)$$

The derivation of the density per unit area is not easy. We should perform the integration in z-direction. For this we would need the function n(z). As mentioned „MOSFET Detailed“ shows the exact derivation.

The result is quite intuitive.

$$Q' = \int n(z)dz = \int n_0 e^{-V(z)/U_T} dz = \cdots = C'_{dep}\, U_T e^{-U_B/U_T} \quad (5)$$

Q’ is the charge density per area, $C'_{dep}$ is the dynamic capacitance / area of the depleted region. (Precisely $C'_{dep} = C'_{dep,ac}$, the dynamic capacitance of depleted region for $V_{dep} = V_{cont}$.)

The charge density is for $V_{ds} = 0$ uniformly distributed in x-direction. Therefore, the $I_{ds}$ current is zero.

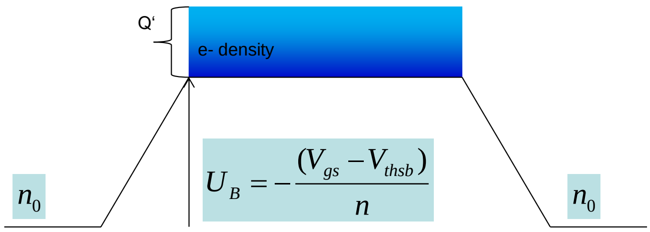



*Figure 9: Charge carrier density in channel for $V_{gs} < V_{th}$ and $V_{ds} = 0$.*

A current will flow only if there is a density gradient. A gradient will be formed when there is a voltage between drain and source.

Let us now consider the case $V_{ds} \neq 0$.

How large is the potential barrier between drain and substrate?

The threshold at the drain end of the channel region is by (n-1) $V_{ds}$ larger than $V_{thsb}$ because of the body effect.

The potential barrier between the drain and the substrate surface is:

$$U_{B,D} = -\frac{(V_{gd} - V_{thsb} - (n-1)V_{ds})}{n} = -\frac{(V_{gs} - V_{thsb})}{n} + \frac{(V_{ds} + (n-1)V_{ds})}{n} = U_{B,S} + V_{ds} \quad (6)$$

$U_{B,S}$ is the potential barrier between drain and substrate.

If we substitute (6) into (5) we obtain the following result: The charge density in the channel region close to drain is by exp ($-V_{ds}/U_T$) smaller than close to source.

This result can be also derived when we apply Maxwell-Boltzmann formula in the drain-substrate junction region. The electron density in the drain is $N_d$ and in the substrate close to drain $N_d$ exp ($-U_{B,D}/U_T$).

Notice the following: Maxwell-Boltzmann Formula does not “work“ in the channel region. The potential is there independent of x coordinate (Figure 9), but the electron density changes from $Q'_s$ to $Q'_d$. The reason for this is that this region is not in the equilibrium state when a current flows.

**Summary**

The charge density close to source is

$$Q'_s = C'_{dep}\, U_T e^{(V_{gs} - V_{thsb})/nU_T} \quad (7)$$

The charge density close to drain is

$$Q'_d = Q'_s\, e^{-V_{ds}/U_T} \quad (8)$$

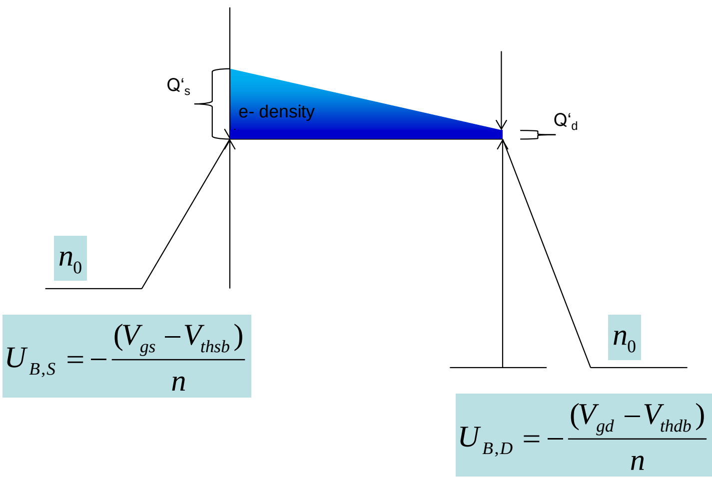



*Figure 10: The density of the charge carriers in the substrate below the gate oxide for $V_{gs} < V_{th}$ and for $V_{ds} > 0$.*

As equations 7 and 8 show, there is a density gradient. It yields to the following diffusion current:

$$|I| = W \cdot D \cdot \frac{dQ}{dx} \quad (9)$$

D is the diffusion constant, for which the Einstein equation holds:

$$D/\mu = U_T \quad (10)$$

It follows:

$$I \approx W \cdot D \cdot \frac{Q'_s - Q'_d}{L} = \frac{W}{L} \mu U_T Q'_s (1 - e^{V_{ds}/U_T}) \quad (11)$$

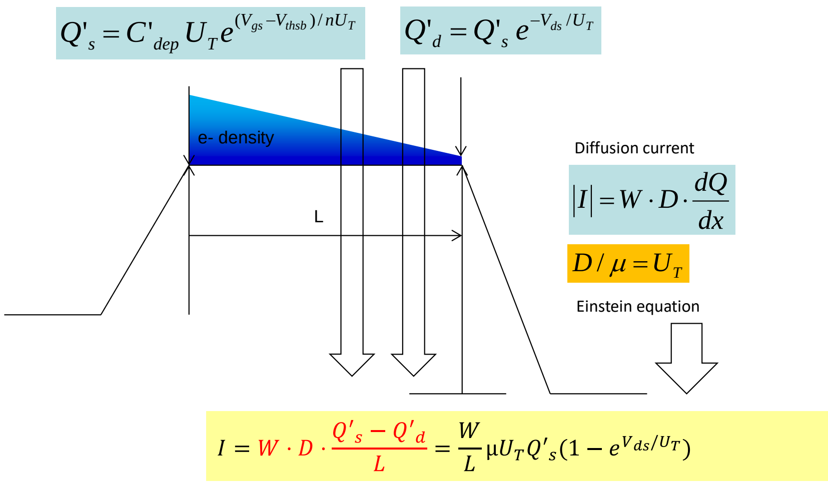



*Figure 11: Diffusion current for $V_{gs} < V_{th}$*

The characteristics $I_{ds} = f(V_{ds})$ (11) shows saturation behaviour for $V_{ds} >$ a few $U_T$ (Figure 12):

$$I_{ds} = const \cdot e^{(V_{gs} - V_{thsb})/nU_T} \left(1 - e^{-V_{ds}/U_T}\right) \quad (12)$$

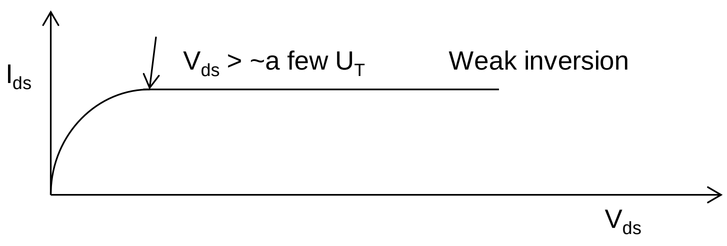



*Figure 12: Saturation of the diffusion current für $V_{ds} > U_T$.*

When we substitute $V_{ds} >$ a few $U_T$ in (11), we get:

$$I_{dssat} = \frac{W}{L} \mu U_T Q'_s \quad (13)$$

It holds also (7):

$$Q'_s = C'_{dep}\, U_T e^{(V_{gs}-V_{thsb})/nU_T} \qquad (14)$$

and

$$C'_{dep} = (n-1)C'_{ox} \quad (15)$$

By substituting (14) and (15) into (13) we obtain the equation for saturation current for weak inversion:

$$I_{dssat} = \frac{W}{L}\cdot \mu \cdot C'_{ox} \cdot U_T{}^2 \cdot (n-1)\cdot e^{(V_{gs}-V_{thsb})/nU_T} \quad (16)$$

We can conclude

1) A transistor is never fully off. Assume $V_{gs} = V_{thsb}$. From equations for strong inversion we would expect $I_{dssat} = 0$. If we use equation (16) we obtain $I_{dssat} \sim W/L \times 100$ nA.

2) The condition for saturation in weak inversion is $V_{ds}$ > a few $U_T$. Notice that $V_{dssat}$ does not depend of $V_{gs}$, as it does in strong inversion.

Strong inversion: $V_{ds} > (V_{gs} - V_{thsb})/n$.

This is an interesting result that influences some circuits (e.g. current mirrors).

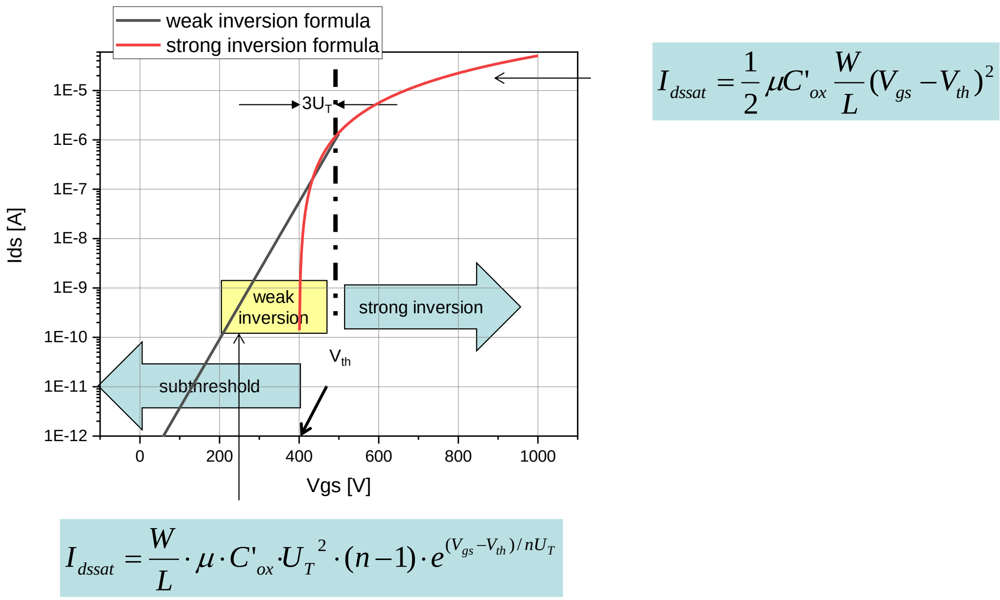



*Figure 13: Simulated current as function of $V_{gs}$*

A current of 100 nA may sound small but for many applications it is significant.

Let us assume a DRAM cell with a capacitance of 10 fF. In the case of a 100 nA current, the DRAM cells gets discharged in about 100 ns.

Weak inversion is one of the reasons for DC current consumption of CMOS logic gates.

We can divide the $V_{gs}$ voltage range in weak inversion ($V_{gs} < V_{thsb}$ + a few $U_T$) and in strong inversion $V_{gs} > V_{thsb}$ + a few $U_T$, as shown in Figure 13.

For strong inversion following equation is valid:

$$I_{dssat} = \frac{1}{2n} \mu C'_{ox} \frac{W}{L} (V_{gs} - V_{thsb})^2 \quad (17)$$

For weak inversion we derived (16):

$$I_{dssat} = \frac{W}{L} \cdot \mu \cdot C'_{ox} \cdot U_T^{\ 2} \cdot (n-1) \cdot e^{(V_{gs} - V_{thsb})/nU_T}$$

A further consequence of weak inversion is that we cannot increase the transconductance by increasing W/L ratio beyond some value if the bias current is kept constant.

Let us calculate the transconductance as $dI_{dsat}/dv_{gs}$:

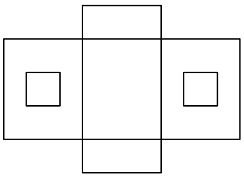
From the formula for strong inversion (17) we get:

$$g_m = \sqrt{2kI_{dssat} \cdot (W/L)}; k = \mu C'_{ox} /n \quad (18)$$

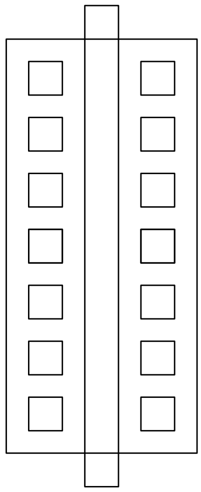
From the formula for weak inversion (16) we get:

$$g_m = I_{dssat}/nU_T \quad (19)$$

Equation (18) would imply, that we can increase $g_m$ as much as we wish when the transistor is made shorter and wider. However, this is not true. If we have a constant bias current $I_{ds}$, increase of W/L lead to a decrease of $V_{gs}$ - $V_{thsb.}$ This yields from (17). This puts a transistor into weak inversion where the transconductance does not depend of W/L. The increase of $g_m$ is stopped. The value $I_{dssat}/n \times U_T$ is the maximum transconductance that we can get from an MOS transistor for a given $I_{dssat}$ bias current.



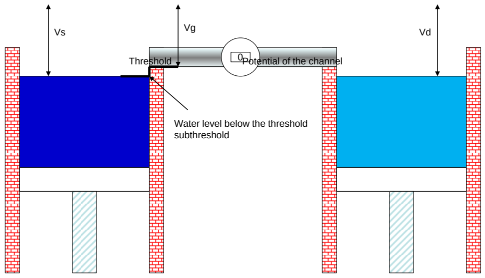
$$g_m = \sqrt{2kI_{dssat}(W/L)}$$



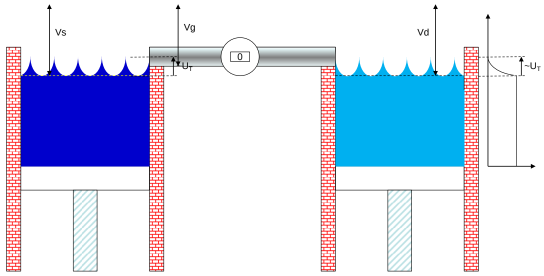
$$g_m = I_{dssat} / nU_T$$

*Figure 14: Transconductances of a transistor working in strong inversion (left) and weak inversion (right) respectively.*

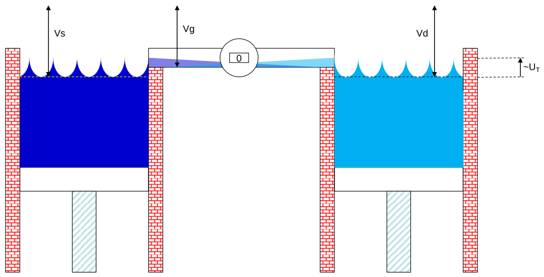
Following figures show an analogy for weak inversion.



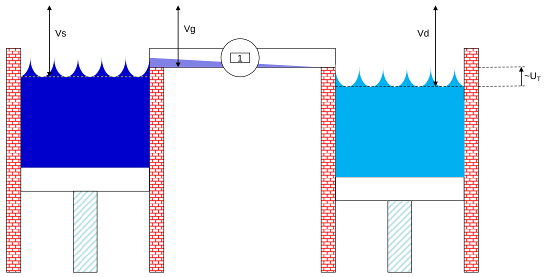

*Two water tanks (source and drain), whose bottom heights (potential energies) can be adjusted. A pipe connects the two tanks at different heights, corresponding to the gate voltage.*
*Let us imagine the case where the gate voltage $V_g$ is below the threshold. (The water level is below the pipe.)*
*In our simplified analogy, the threshold voltage is zero, and $V_g$ equals the channel potential. The transistor conducts when $V_g=V_s$.*
*No current flows.*



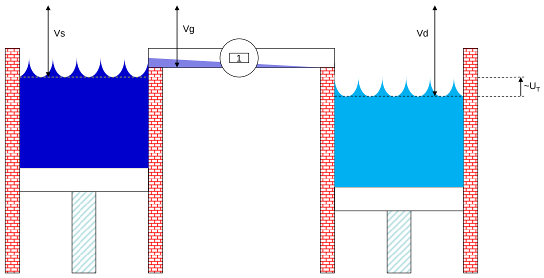

*Now let us assume that there are waves. The waves symbolize thermal energy; therefore, the "height of the waves" corresponds to $U_T=kT/e$.*
*Because of the waves, water (electrons) can reach the pipe even when the water level (without waves) is below the pipe (i.e., $V_s$ and $V_d < V_g$).*



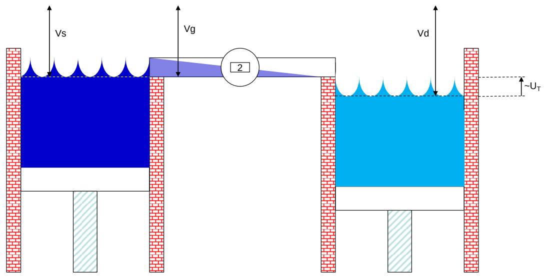

*The water that reaches the pipe spreads out by diffusion (without any change in potential). Since the same amount of water sloshes in from both the source and drain tanks, the net flow is zero.*




*This changes when the drain potential is increased so that the drain potential is higher by one thermal voltage compared to the channel (or gate) potential ($V_d>V_g$).*
*Even with waves (thermal energy), the electrons from the drain can no longer reach the channel. There is now a current flow from source to drain.*




*The current does not change when $V_d$ is increased further.*




*However, if $V_{gs}$ is increased, more water enters the channel, and the current increases.*

Let us now summarize the transistor equations that we derived so far:

**Strong inversion**

Simple model with body effect

$$I_{ds} = \frac{\mu_0 C'_{ox} W}{L}\left((V_{gs} - V_{thsb})V_{ds} - n\frac{V_{ds}^2}{2}\right)$$

Saturation voltage

$$V_{ds} = \frac{V_{gs} - V_{thsb}}{n} \equiv V_{dssat}$$

Saturation current

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

Symmetrical equation:

$$I_{ds} = I_{dssat}(V_{gs}, V_{thsb}) - I_{dssat}(V_{gd}, V_{thdb})$$

Body effect

$$V_{thsb} = V_{th} + (n-1)V_{sb}$$

n is the slope factor, $n = 1 + C_{dep,min} / C_{ox} \sim 1.25$.

If we do not want to take body effect into account we can set n = 1.

**Model that assumes velocity saturation**

$$I_{ds} = \frac{\mu_0 C'_{ox} W}{L\left(1 + \frac{V_{ds}}{LE_{sat}}\right)}\left((V_{gs} - V_{thsb})V_{ds} - n\frac{V_{ds}^2}{2}\right)$$

Saturation voltage

$$V_{ds} = \frac{V_{gs} - V_{thsb}}{n\alpha} \equiv V_{dssat}$$

$$\alpha = \left(1 + \frac{V_{gs} - V_{thsb}}{nE_{sat}L}\right)$$

Saturation current

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

**Weak inversion**

$$I_{ds} = \frac{W}{L} \cdot \mu \cdot C'_{ox} \cdot U_T^{\,2} \cdot (n-1) \cdot e^{(V_{gs}-V_{thsb})/nU_T} (1 - e^{-Vds/U_T})$$

Saturation voltage

z.B. $V_{ds} > 3U_T$

Saturation current

$$I_{dssat} = \frac{W}{L} \cdot \mu \cdot C'_{ox} \cdot U_T^{\,2} \cdot (n-1) \cdot e^{(V_{gs}-V_{thsb})/nU_T}$$

Symmetrical equation holds also in weak inversion:

$$I_{ds} = I_{dssat}(V_{gs}, V_{thsb}) - I_{dssat}(V_{gd}, V_{thdb})$$

$$V_{thsb} = V_{th} + (n-1)V_{sb}$$

$$V_{thdb} = V_{th} + (n-1)V_{db}$$

**Early-effect**

$$I_{ds} = I_{dssat}(1 + (V_{ds} - V_{dssat})/E_{sat}L)$$

$$V_A = L_{eff}E_{sat} + \frac{L_{eff}eN_a\, x_j}{\epsilon}$$

Or in BSIM Model

$$V_A \sim LE_{sat} + L\frac{V_{ds}-V_{dssat}}{0.5\, x_j};\; x_j = \sqrt{\frac{2\epsilon(V_{ds}-V_{dsssat})}{eN_a}}$$

**Threshold voltage**

$$V_{th} = \frac{2C'_{dep,min}}{C'_{ox}} \times V_{cont} = \frac{\sqrt{2eN_a\varepsilon_0\varepsilon_{Si}V_{cont}}}{C'_{ox}}$$

$$V_{cont} = 2U_T \ln(\frac{N_a}{n_i})$$

$N_a$ is the density of acceptor atoms in the channel region, $n_i$ is the intrinsic carrier density

$$n_i = \sqrt{N_c N_v} e^{\frac{-E_g}{2eU_T}}$$

$N_c$ und $N_v$ are the effective densities of quantum states in conduction band and valence band, $E_g$ is the band gap energy. $n_i = 10^{10}/cm^3$, density of silicon atoms is: $n_{si} = 5 \times 10^{22}/cm^3$. We assume $N_a = 10^{18}/cm^3$.

Following parameter values are realistic for a 65 nm technology

NMOS:

$V_{th}$ = 400mV

n = 1.25

$E_{sat}$ ~ 5.29 V/µm

$\mu_0 = 4.78 \times 10^{-2}\ m^2/Vs$

$C'_{ox}$ = 13.28 fF/µm$^2$ = 0.01328f/m$^2$

PMOS:

$V_{th}$ = 380mV

n = 1.22

$E_{sat}$ ~ 33.8 V/µm

$\mu_0 = 1.19 \times 10^{-2}\ m^2/Vs$

$C'_{ox}$ = 12.33 fF/µm$^2$

Comment:

$\mu_0$ is the low field mobility u0 from model file

$C_{ox}$ calculated from toxe = 2.6nm(NMOS)/2.8nm(PMOS) (model file) and $\epsilon_r = 3.9$

$E_{sat} = 2v_{sat}/u0$ ($v_{sat}$ from model file)

$V_{th}$ simulated – voltage when the current has the value from weak inversion formula

n calculated from subthreshold slope (simulation)

**Reference Voltage Generator (Reference Voltage Source)**

In this chapter, we will present a realization of the generator of reference voltage (the reference voltage source).

The reference voltage source should generate a voltage that is relatively independent of the supply voltage $V_{IN}$ and of temperature.

Figure *1* shows the schematic.

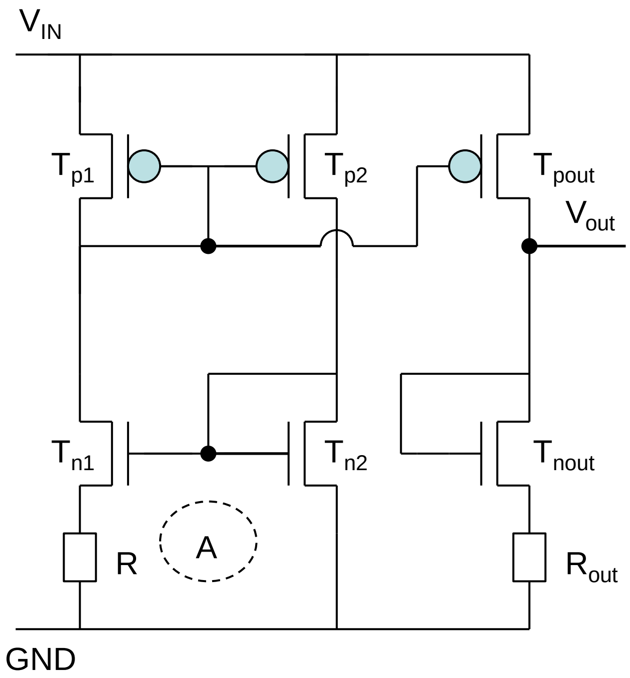



Figure 1: Reference Voltage Generator

Transistor $T_{n1}$ has W/L ratio N times greater than $T_{n2}$.

The output voltage is formed at the $T_{nout}$ and $R_{out}$.

The current mirror (transistors $T_{p1}$, $T_{p2}$ and $T_{pout}$) ensures that the currents through $T_{n1}$, $T_{n2}$ and $T_{nout}$ are equal:

$$I_{n1} = I_{n2} = I_{nout} = I \qquad (1)$$

We refer to the supply voltage as $V_{IN}$.

Let us calculate the current I.

We use the Kirchhoff's law for the contour A. The following applies:

$$RI + V_{gs,n1} = V_{gs,n2} \qquad (2)$$

Let us assume that all transistors are in saturation and in weak inversion. The transistor current is then:

$$I_{ds,sat} = \frac{W}{L}\mu C'_{ox}U_T^2(n-1)e^{(V_{gs}-V_{th})/nU_T} = \frac{W}{L}I_0 e^{(V_{gs}-V_{th})/nU_T} \quad (3)$$

with

$$I_0 \equiv \mu C'_{ox}U_T^2(n-1) \sim 100nA$$

μ is the mobility of the charge carriers, n is slope factor, $U_T$ the thermal voltage: $U_T = kT/e$, $C'_{ox}$ is the oxide capacitance per area ($\varepsilon_0\,\varepsilon_{sio2}/t_{ox}$).

From (3) we get for $T_{n1}$, $T_{n2}$ and $T_{nout}$:

$$V_{gs,n1} = V_{th} + nU_T \ln\left(\frac{I}{I_0 W_{n1}/L_{n1}}\right)$$

$$V_{gs,n2} = V_{th} + nU_T \ln\left(\frac{I}{I_0 W_{n2}/L_{n2}}\right)$$

$$V_{gs,n2} = V_{th} + nU_T \ln\left(\frac{I}{I_0 W_{nout}/L_{nout}}\right) \quad (4)$$

Note that for weak inversion $I_{ds,sat} = I < I_0\,W/L$. The logarithms in (4) are negative.

If we substitute (4) in (2), we get:

$$RI + V_{th} + nU_T \ln\left(\frac{I}{I_0 W_{n1}/L_{n1}}\right) = V_{th} + nU_T \ln\left(\frac{I}{I_0 W_{n2}/L_{n2}}\right)$$

It follows:

$$RI = nU_T \ln\left(\frac{W_{n1}/L_{n1}}{W_{n2}/L_{n2}}\right)$$

Since the transistor $T_{n1}$ has W/L by factor N larger:

$$I = \frac{nU_T}{R}\ln(N) = \frac{nkT}{e}\ln(N) \quad (5)$$

The current I is independent of supply voltage $V_{IN}$. As a first approximation, the current increases linearly with temperature. We neglect the temperature dependence of the slope factor n.

Now let us calculate the output voltage. The following applies (we use the result (5)):

$$V_{out} = R_{out}I + V_{gs,nout} = nU_T\frac{R_{out}}{R}\ln(N) + V_{th} + nU_T \ln\left(\frac{I}{I_0 W_{nout}/L_{nout}}\right) \quad (6)$$

The temperature dependence of $V_{out}$ is complicated. However, it is relatively easy to see that the first term:

$$nU_T\frac{R_{out}}{R}\ln(N) = \frac{nkT}{e}\frac{R_{out}}{R}\ln(N) \quad (7)$$

increases linearly with the rise of T.

Let us look at the third term:

$$nU_T \ln\left(\frac{I}{I_0 W_{nout}/L_{nout}}\right) \quad (8)$$

The logarithm is negative and $U_T = kT/e$ increases with temperature. Therefore, the third term causes $V_{out}$ to become smaller as the temperature rises.

How does the threshold voltage depend on temperature?

We will use the following approximate formula for $V_{th}$:

$$V_{th} = \frac{2C_{dep,ac}}{C_{ox}} \times V_{cont} \sim \frac{1}{2} V_{cont} \quad (9)$$

$N_a$ is the density of acceptor atoms in the channel region, $V_{cont}$ is the contact voltage between n and p silicon (for equal n and p doping strengths):

$$V_{cont} = 2U_T \ln\left(\frac{N_a}{n_i}\right) \quad (10)$$

$N_i$ is the intrinsic density of charge carriers:

$$n_i = \sqrt{N_c N_v} e^{\frac{-E_g}{2eU_T}} \quad (11)$$

$N_c$ und $N_v$ are the effective densities of quantum states in conduction band and valence band:

$$N_c = 2\left(\frac{2\pi m_e^* kT}{h^2}\right)^{3/2} \sim 2.4 \times 10^{19}\ cm^{-3}$$

$$N_v = 2\left(\frac{2\pi m_h^* kT}{h^2}\right)^{3/2} \sim 1.5 \times 10^{19}\ cm^{-3}$$

https://www.tf.uni-kiel.de/matwis/amat/mw_for_et/kap_8/backbone/r8_3_3.html

$E_g$ is the energy of the bandgap.

We will assume that $\frac{2C_{dep,ac}}{C_{ox}} \sim 0.5$ and is independent of temperature.

The threshold voltage

$$V_{th} = \frac{2C_{dep,ac}}{C_{ox}} \times V_{cont} \sim \frac{1}{2} V_{cont} \quad (12)$$

becomes smaller with temperature increase, because $V_{cont}$ (10) becomes smaller.

Let us try to roughly calculate the temperature dependence of contact voltage:

We use:

$$V_{cont} = 2U_T \ln(\frac{N_a}{n_i}) \quad (13)$$

and

$$n_i = \sqrt{N_c N_v} e^{\frac{-E_g}{2eU_T}} \quad (14)$$

It follows:

$$V_{cont} = \frac{2kT}{e}\left[\ln\left(N_a/\sqrt{N_c N_v}\right) + \frac{E_g}{2eU_T}\right] = 2U_T \ln\left(\frac{N_a}{\sqrt{N_c N_v}}\right) + \frac{E_g}{e}$$

$\ln\left(N_a/\sqrt{N_c N_v}\right)$ ist negative! Why?

The contact voltage is smaller than the bandgap expressed in volts and the density of ionized acceptors is smaller than the density of quantum states in valence and conduction band.

It follows that the threshold voltage decreases with temperature increase:

$$V_{th} = \sim \frac{1}{2} V_{cont} = \frac{1}{2}\,\frac{2kT}{e} \ln\left(\frac{N_a}{\sqrt{N_c N_v}}\right) + \frac{E_g}{2e}$$

Let us insert this result into (6)

$$V_{out} = \frac{E_g}{2e} + nU_T \frac{R_{out}}{R} \ln(N) + 2U_T \ln\left(\frac{N_a}{\sqrt{N_c N_v}}\right) + nU_T \ln\left(\frac{I}{I_0 W_{nout}/L_{nout}}\right)$$

We rewrite it as follows:

$$V_{out} = \frac{E_g}{2e} + nU_T \frac{R_{out}}{R} \ln(N) - 2U_T \ln\left(\frac{\sqrt{N_c N_v}}{N_a}\right) - nU_T \ln\left(\frac{\frac{I_0 W_{nout}}{L_{nout}}}{I}\right)$$

Both logarithms are now positive. It is therefore possible to set the parameters of the circuit ($R_{out}$, R, N, $W_{out}/L_{out}$) in such a way that the temperature increase of the first term is compensated by the second and third terms. Correct dimensioning can be determined in simulations. Often it is also necessary to adjust the circuit again after production. After that, a second design iteration is made and tested.

To find the initial values, the derivative of $V_{out}$ over T can be calculated.

$$\frac{dV_{out}}{dT} = \frac{nk}{e}\frac{R_{out}}{R} \ln(N) - \frac{2k}{e} \ln\left(\frac{\sqrt{N_c N_v}}{N_a}\right) - \frac{nk}{e} \ln\left(\frac{\frac{I_0 W_{nout}}{L_{nout}}}{I}\right) \quad (15)$$

If the derivative is 0, the function $V_{out}(T)$ has a saddle point and a weak temperature dependence.

This is fulfilled under the following conditions (we neglect $\frac{2k}{e} \ln\left(\frac{\sqrt{N_c N_v}}{N_a}\right)$):

$$R_{out} = R$$

and

$$\frac{I_0 W_{nout}/L_{nout}}{I} \sim N \Rightarrow I \sim \frac{1}{N}\frac{W_{nout}}{L_{nout}} I_0$$

The output voltage in this case is approximately:

$$V_{out} \sim \frac{E_g}{2e} \sim 0.5V$$

The output voltage is about 1/2 voltage of the band gap. This is the reason why this type of reference voltage generator is called band-gap reference. It is interesting that the output voltage of one circuit reflects one quantum property of the semiconductor, the band gap energy.

**BSIM transistor model**

Up to now, we have used different equations for the various operating regions of a MOS transistor, for example:

$V_{GS} > V_{DS,sat}$

$V_{GS} > V_{th}$

This approach works well when performing calculations on paper.
Typically, you first make an assumption about the operating region, then apply the equation that is valid for that region, and finally verify whether your assumption was correct.

However, for **computer-based modeling**, it is preferable to use models that apply **a single equation across all operating regions**.

The simulators in **Cadence Spectre** use the **BSIM**, which was developed at **University of California, Berkeley** by **Chenming Hu** and his colleagues.

Reference:
https://www.chu.berkeley.edu/modern-semiconductor-devices-for-integrated-circuits-chenming-calvin-hu-2010

BSIM replaces the piecewise MOSFET equations with smooth transitions so circuit simulators can compute derivatives and converge reliably.

Unfortunately, the full BSIM equations are quite complicated. For teaching purposes, I have **significantly simplified them** by removing most correction terms while keeping the main physical ideas.

In the BSIM model, **a single current equation is used for both weak and strong inversion**.

The most important equation describes the drain current:

$$I_{ds} = \mu \frac{W}{L} \frac{1}{1+\frac{V_{dseff}}{E_{sat}L}} Q_{eff}\, V_{dseff}\left(1 - \frac{V_{dseff}}{2V_b}\right)\left(1 + \frac{V_{ds}-V_{dseff}}{V_A}\right) \qquad (1)$$

$$Q_{eff} = C_{ox}V_{gsteff}$$

$$V_b = \frac{V_{gsteff}+nU_T}{n}$$

$$V_{dssat} = \frac{(V_{gsteff}+nU_T)}{n\alpha}$$

$$\alpha = \left(1 + \frac{V_{gsteff}}{nE_{sat}L}\right)$$

The equations for $V_{dseff}$ and $V_{gsteff}$ will be introduced in the next section.

The **term highlighted in red** describes **channel-length modulation**, which we will ignore in the following discussion.

The **terms highlighted in blue** describe **mobility saturation**.

n is the **slope factor**.

**Equation for $V_{dseff}$ and $V_{gsteff}$**

The effective voltages $V_{gsteff}$ and $V_{dseff}$ are defined by the following equations:

$$V_{gsteff} = \frac{2nU_T \ln\left(1+\exp\left(\frac{V_{gst}}{2nU_T}\right)\right)}{1+n\frac{C_{ox}}{C_{dep}}\exp\left(-\frac{V_{gst}}{2nU_T}\right)}$$

With $V_{gst} = V_{gs} - V_{th}$

The effective drain–source voltage is

$$V_{dseff} = V_{DS,sat} - \frac{1}{2}\left(V_{DS,sat} - V_{DS} - \delta + \sqrt{(V_{DS,sat} - V_{DS} - \delta)^2 + 4\delta V_{DS,sat}}\right)$$

where

$\delta = 0.001V$ to $0.05V$

The parameter δ\deltaδ ensures a **smooth transition between the linear and saturation regions**, which is important for circuit simulators.

**Simplified Expressions for $V_{gsteff}$ and $V_{DS,sat}$**

To better understand the behavior of the model, we can examine the limiting cases for

$V_{gst} \gg U_T$ (strong inversion)

$V_{gst} \ll U_T$ (weak inversion)

**Case 1: $V_{gst} \gg U_T$**

In this case, the exponential term dominates and the expression simplifies to

$$V_{gsteff} = \frac{2nU_T \ln\left(1+\exp\left(\frac{V_{gst}}{2nU_T}\right)\right)}{1+n\frac{C_{ox}}{C_{dep}}\exp\left(-\frac{V_{gst}}{2nU_T}\right)} \sim V_{gst}$$

And therefore

$$V_{dssat} = \frac{(V_{gsteff}+nU_T)}{n\alpha} \sim \frac{V_{gst}}{n\alpha}$$

**Case 2: $V_{gst} \ll U_T$**

For small arguments we use the approximation

$\ln(1+x) \approx x$ for $x \ll 1$

Applying this to the expression for $V_{gsteff}$ gives

$$V_{gsteff} = \frac{2nU_T \ln\left(1+\exp\left(\frac{V_{gst}}{2nU_T}\right)\right)}{1+n\frac{C_{ox}}{C_{dep}}\exp\left(-\frac{V_{gst}}{2nU_T}\right)} \sim 2U_T \frac{C_{dep}}{C_{ox}} \exp\left(\frac{V_{gst}}{nU_T}\right)$$

and therefore

$$V_{dssat} = \frac{(V_{gsteff}+nU_T)}{n\alpha} \sim U_T$$

For **strong inversion**, the model reduces approximately to the familiar expression

$$V_{gsteff} \approx V_{gs} - V_{th}.$$

For **weak inversion**, $V_{gsteff}$ shows an **exponential dependence**, which is responsible for the exponential current behavior in the subthreshold region.

**Equations for $V_{dseff}$ for $V_{gst} \gg U_T$**

We now calculate $V_{dseff}$ for the case $V_{gst} \gg U_T$ (strong inversion) and consider both operating regions of $V_{DS}$: the **triode region** and the **saturation region**.

For $V_{gst} \gg U_T$, the effective gate voltage simplifies to

$$V_{gsteff} = \frac{2nU_T \ln(1+\exp(\frac{V_{gst}}{2nU_T}))}{1+n\frac{C_{ox}}{C_{dep}}\exp(-\frac{V_{gst}}{2nU_T})} \approx V_{gst}$$

and therefore

$$V_{DS,sat} = \frac{V_{gsteff}+nU_T}{n\alpha} \approx \frac{V_{gst}}{n\alpha}$$

**Case 1: $V_{DS} < V_{DS,sat}$ (Triode Region)**

$$V_{dseff} \approx V_{DS,sat} - \frac{1}{2}\left(V_{DS,sat} - V_{DS} + \sqrt{(V_{DS,sat} - V_{DS})^2}\right)$$

Since

$$\sqrt{(V_{DS,sat} - V_{DS})^2} = V_{DS,sat} - V_{DS}$$

we obtain

$$V_{dseff} \approx V_{DS}$$

Thus, in the triode region the effective drain voltage is essentially equal to the applied drain voltage.

**Case 2: $V_{DS} > V_{DS,sat}$ (Saturation Region)**

$$V_{dseff} \approx V_{DS,sat} - \frac{1}{2}\left(V_{DS,sat} - V_{DS} + \sqrt{(V_{DS,sat} - V_{DS})^2}\right)$$

Since $\sqrt{(V_{DS,sat} - V_{DS})^2} = V_{DS} - V_{DS,sat}$

we obtain

$$V_{dseff} \approx V_{DS,sat}$$

Thus, in saturation the effective drain voltage **saturates at** $V_{DS,sat}$

**Equations for $V_{dseff}$ for $V_{gst} \ll U_T$**

Next, we calculate $V_{dseff}$ for the case $V_{gst} \ll U_T$ (weak inversion), again considering both operating regions of $V_{DS}$

The effective gate voltage simplifies to

$$V_{gsteff} = \frac{2nU_T \ln\left(1+\exp\left(\frac{V_{gst}}{2nU_T}\right)\right)}{1+n\frac{C_{ox}}{C_{dep}}\exp\left(-\frac{V_{gst}}{2nU_T}\right)} \sim 2U_T\frac{C_{dep}}{C_{ox}}\exp\left(\frac{V_{gst}}{nU_T}\right)$$

and therefore

$$V_{DS,sat} = \frac{V_{gsteff}+nU_T}{n\alpha} \approx U_T$$

**Case 1: $V_{DS} < V_{DS,sat} = U_T$**

$$V_{dseff} = V_{DS,sat} - \frac{1}{2}\left(V_{DS,sat} - V_{DS} - \delta + \sqrt{(V_{DS,sat} - V_{DS} - \delta)^2 + 4\delta V_{DS,sat}}\right)$$

For $V_{DS} \ll U_T$, this expression can be approximated as

$$= V_{dssat} - \frac{1}{2}\left(V_{dssat} - V_{ds} - \delta + ((V_{dssat} - \delta)^2 + 4\delta V_{dssat})^{0.5}\right) = \frac{1}{2}V_{ds}$$

**Case 2: $V_{DS} > V_{DS,sat}$**

In this case the expression simplifies to

$$V_{dseff} \approx V_{DS,sat}$$

δ = 0.001V to 0.05V

The parameter δ ensures a **smooth transition between the linear and saturation regions**, which is essential for stable circuit simulation.

In **strong inversion**, the BSIM model behaves similarly to the **classical MOSFET model**.

In **weak inversion**, $V_{DS,sat}$ becomes approximately $U_T$, meaning the transistor enters saturation at **very small drain voltages (~25 mV at room temperature)**.

**Saturation Current Formula for Weak and Strong Inversion**

We now derive the formula for the **saturation current**.
This expression applies when

$V_{DS} > V_{DS,sat}$

and is valid for **both weak and strong inversion**.

We start from Equation (1) and substitute

$V_{dseff} = V_{DS,sat}$.

The drain current equation is

$$I_{DS} = \mu \frac{W}{L} \frac{1}{1+\frac{V_{dseff}}{E_{sat}L}} \ Q_{eff} \, V_{dseff}(1 - \frac{V_{dseff}}{2V_b}) \qquad (1)$$

with

$Q_{eff} = C_{ox}V_{gsteff}$

$V_b = \frac{V_{gsteff}+nU_T}{n}$

$V_{dssat} = \frac{(V_{gsteff}+nU_T)}{n\alpha} = \frac{V_b}{\alpha}$

$\alpha = (1 + \frac{V_{gsteff}}{nE_{sat}L})$

**Saturation Condition**

For $V_{DS} > V_{DS,sat}$ we have $V_{dseff} = V_{DS,sat}$.

Substituting this into Equation (1) gives

$$I_{DS,sat} = \mu \frac{W}{L} \frac{1}{1+\frac{V_{DS,sat}}{E_{sat}L}} \ Q_{eff}V_{DS,sat}(1 - \frac{V_{DS,sat}}{2V_b})$$

Using $V_{DS,sat} = \frac{V_b}{\alpha}$

we obtain

$$1 - \frac{V_{DS,sat}}{2V_b} = 1 - \frac{1}{2\alpha}$$

and

$$\frac{1}{1+\frac{V_{DS,sat}}{E_{sat}L}} = \frac{1}{1+\frac{V_{gsteff}}{n\alpha E_{sat}L}}$$

Substituting these relations yields

$$I_{DS,sat} = \mu\frac{W}{L}\frac{1}{1+\frac{V_{gsteff}}{n\alpha E_{sat}L}}\ Q_{eff}V_{DS,sat}(1-\frac{1}{2\alpha})$$

For typical conditions, the factors involving α approximately cancel, leading to the simplified expression

$$I_{DS,sat} = \frac{\mu}{2}\frac{W}{L}Q_{eff}V_{DS,sat} \qquad (2)$$

**Interpretation**

Equation (2) is a **compact saturation current expression** that is valid for **both weak and strong inversion**.

The effective charge term

$$Q_{eff} = C_{ox}V_{gsteff}$$

allows the model to smoothly transition between:

**weak inversion**, where the current depends exponentially on $V_{GS}$, and

**strong inversion**, where the classical quadratic MOSFET behavior is recovered.

---

The key idea of the model is that the same current equation works for **both weak and strong inversion** because the effective gate voltage $V_{gsteff}$ smoothly connects the two regimes.

**Strong inversion**

We now calculate the drain current for **strong inversion**, i.e. $V_{gst} \gg U_T$

and consider both operating regions of $V_{DS}$: the **triode region** and the **saturation region**.

We start from the general current equation

$$I_{ds} = \mu\frac{W}{L(1+\frac{V_{dseff}}{E_{sat}L})}Q_{eff}V_{dseff}(1-\frac{V_{dseff}}{2V_b}) \ (1)$$

Approximations for Strong Inversion

For $V_{gst} \gg U_T$, the following approximations apply:

$V_{gsteff} \sim V_{gst}$

$Q_{eff} = C_{ox} V_{gst}$

$V_b = \frac{V_{gsteff} + nU_T}{n} \sim \frac{V_{gst}}{n}$

$V_{DS,sat} = \frac{V_{gsteff} + nU_T}{n\alpha} \approx \frac{V_{gst}}{n\alpha} = \frac{V_b}{\alpha}$

with

$\alpha = 1 + \frac{V_{gst}}{nE_{sat}L}$

**1. Triode Region**

For $V_{DS} < V_{DS,sat}$

We have

$V_{dseff} = V_{DS}$.

Substituting this into Equation (1) yields

$I_{ds} = \mu C_{ox} \frac{W}{L} \frac{1}{(1+\frac{V_{ds}}{E_{sat}L})} (V_{gst} V_{ds} - n \frac{V_{ds}^2}{2})$

This expression corresponds to the **triode-region current**, including the effect of **velocity saturation** through the factor

$\frac{1}{1+\frac{V_{DS}}{E_{sat}L}}$

**2. Saturation Region**

For

$V_{DS} > V_{DS,sat}$

we have

$V_{dseff} = V_{DS,sat}$.

Using the previously derived saturation current expression

$I_{DS,sat} = \frac{\mu}{2} \frac{W}{L} Q_{eff} V_{DS,sat}$

and substituting

$Q_{eff} = C_{ox} V_{gst}$

$V_{DS,sat} = \frac{V_{gst}}{n\alpha}$

gives

$I_{DS,sat} = \frac{\mu C_{ox}}{2n}\frac{W}{L}\frac{V_{gst}^2}{1+\frac{V_{gst}}{nE_{sat}L}}$

**Interpretation**

This result shows that in **strong inversion**

the current approximately follows the **classical quadratic dependence** on $V_{gst}$, but it is reduced by the factor

$\frac{1}{1+\frac{V_{gst}}{nE_{sat}L}}$

which describes **velocity saturation** in short-channel devices.

For **long-channel devices** where $E_{sat}L \gg V_{gst}$, this factor approaches 1, and the familiar square-law MOSFET equation is recovered.

---

The classical square-law MOSFET current still appears, but modern short-channel devices limit the current through **velocity saturation**, which reduces the quadratic growth of current.

**Weak inversion**

We now calculate the drain current for **weak inversion**, i.e. $V_{gst} \ll U_T$ and consider both operating regions of $V_{DS}$ the **triode region** and the **saturation region**.

We start again from the general current equation

$$I_{DS} = \mu\frac{W}{L}\frac{1}{1+\frac{V_{dseff}}{E_{sat}L}}\, Q_{eff}V_{dseff}(1-\frac{V_{dseff}}{2V_b}) \qquad (1)$$

**Approximations for Weak Inversion**

For $V_{gst} \ll U_T$, the following approximations apply:

$V_{gsteff} \sim 2U_T \frac{C_{dep}}{C_{ox}} \exp\left(\frac{V_{gst}}{nU_T}\right)$

$Q_{eff} = C_{ox} V_{gsteff}$

$V_b = \frac{V_{gsteff} + nU_T}{n} \sim U_T$

$V_{dssat} = \frac{(V_{gsteff} + nU_T)}{n\alpha} \sim U_T$

$\alpha = 1 + \frac{V_{gst}}{nE_{sat}L} \approx 1$

**1. Triode Region**

For

$V_{DS} < V_{DS,sat}$

we have

$V_{dseff} = \frac{1}{2} V_{DS}.$

Substituting this into Equation (1) gives

$I_{ds} = \mu U_T^2 C_{dep} \frac{W}{L} \exp\left(\frac{V_{gst}}{nU_T}\right) \frac{V_{ds}}{U_T}\left(1 - \frac{V_{ds}}{2U_T}\right)$

Using the approximation

$e^x \sim 1 + x + \frac{1}{2} x^2$

this expression can be rewritten as

$I_{ds} \sim \mu \frac{W}{L} U_T^2 C_{dep} \exp\left(\frac{V_{gst}}{nU_T}\right)\left(1 - \exp\left(\frac{-V_{ds}}{U_T}\right)\right)$

**2. Saturation Region**

For

$V_{DS} > V_{DS,sat}$

we have

$V_{dseff} = U_T.$

Substituting this into the general expression yields

$I_{DS,sat} = \frac{\mu}{2} \frac{W}{L} Q_{eff} V_{DS,sat}$

which simplifies to

$$I_{dssat} = \mu U_T^2 C_{dep} \frac{W}{L} \exp\left(\frac{V_{gst}}{nU_T}\right)$$

**Final Weak-Inversion Current Expression**

We obtain the same equation that was derived previously for the **subthreshold current**:

$$I_{ds} = \frac{W}{L} \cdot \mu \cdot C'_{ox} \cdot {U_T}^2 \cdot (n-1) \cdot e^{(V_{gs} - V_{thsb})/nU_T}\left(1 - e^{-Vds/U_T}\right)$$

**Comparison of Current Formulas**

The transistor current can therefore be described using different formulas depending on the operating regime.

Weak Inversion:

$$I_{DS} = \frac{W}{L} k_{wi} \exp\left(\frac{V_{gs} - V_{thsb}}{nU_T}\right)$$

Strong Inversion:

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

Mixed (BSIM) Formula:

$$I_{DS,sat} = \frac{\mu}{2} \frac{W}{L} Q_{eff} V_{DS,sat}$$

The **BSIM expression smoothly connects weak and strong inversion**, allowing circuit simulators to use **one continuous model across all operating regions**.

The following figure shows the typical relationship between $I_{DS}$ and $V_{GS}$

**Exponential behavior** in weak inversion

**Quadratic behavior** in strong inversion

**Smooth transition** between the two regions

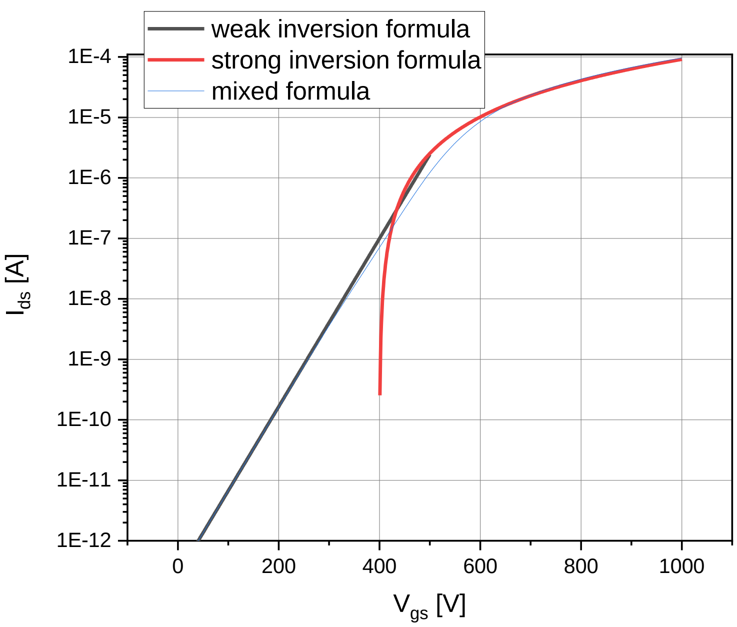



*Figure 4: $I_{ds}$ vs $V_{gs}$ described by different models*