Lecture 3

Theme of this lecture are

MOSFET equations

Body-effect

Channel length modulation (Early-effect)

MOSFET equations

The electrical state of the transistor is described by two voltages, V_{gs} and V_{ds}, and by two currents I_{ds} and I_{gs}. For DC signals, $I_{gs} = 0$ holds. Gate is just a capacity. (We are neglecting the gate current due to tunnel effect.)

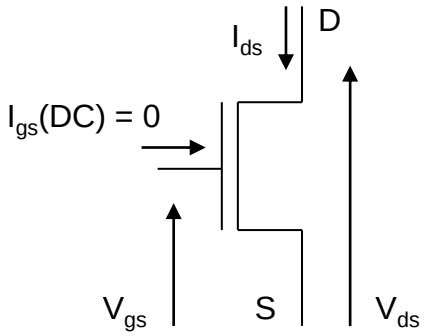


Figure 1: Transistor

Figure 2 shows the transistor I-U characteristics: Transistor behavior for DC signals can be described with the following characteristics. (DC-Signals = slow voltages and currents of any amplitude)

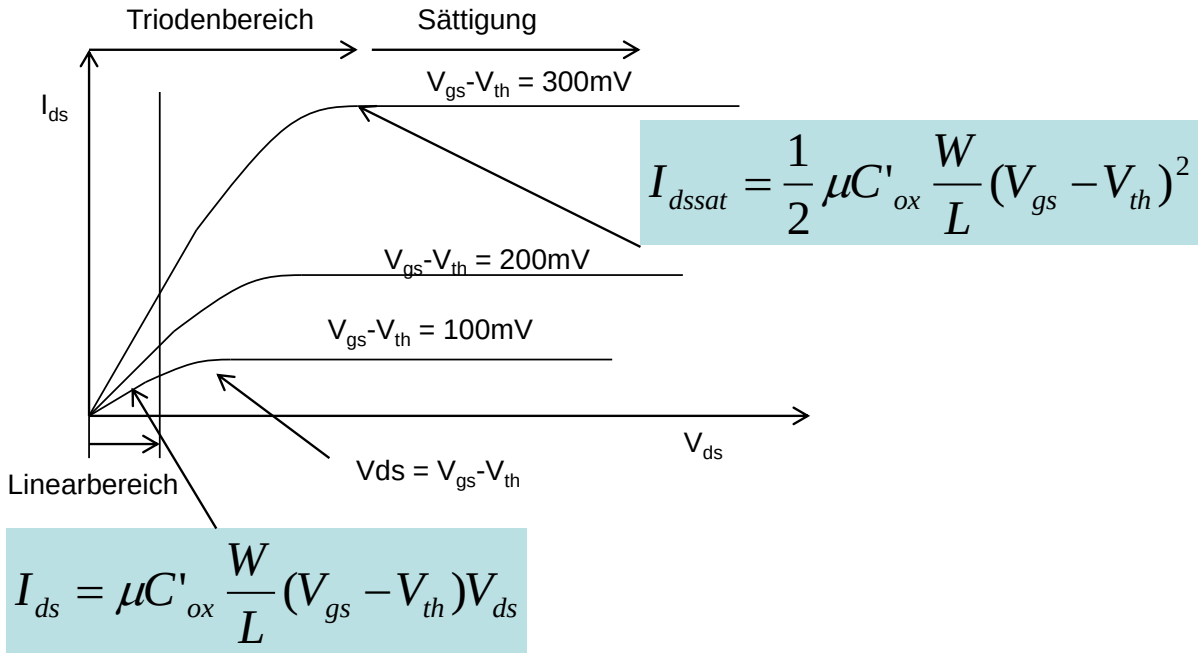


Figure 2: Output characteristics of MOSFET

I_{ds} as a function of V_{ds} for different V_{gs} (output characteristics)

I_{ds} as a function of V_{gs} for different V_{ds} (input characteristics)

Let us consider the output characteristics.

Figure 3Figure 2 shows the I_{ds} - V_{ds} characteristics. The lines are plotted for linearly increased V_{gs}, e.g. 100 mV, 200 mV...

In the right line regions, the current is nearly independent of V_{ds}. We call this region the saturation region. Ideally $I_{ds} = I_{dssat}$ for $V_{ds} > V_{dssat}$. The saturation voltage is nearly $V_{dssat} = V_{gs} - V_{th}$.

The saturation current I_{dssat} is given by the following formula:

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

In the left line regions, the current decreases with the decrease of V_{ds}. We call is triode region. For small V_{ds}, the current voltage is approximately linear function if V_{ds} (linear region). The current in the linear region can be described with the following formula:

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

The current in saturation (1) depends quadratically of V_{gs}. $V_{ds} = V_{dssat}$ at the transition between the saturation and triode regions.

It can be shown that (I_{ds}, V_{ds})-points between the triode and the saturation regions are located at the function $I_{ds} = k/2\ V_{ds}^2$.

This can be derived from the formula for saturation current (1) and the condition $V_{dssat} = V_{gs} - V_{th}$.

In the saturation region, the transistor acts like a voltage-controlled current source. Such current sources are useful. It is possible to implement a voltage amplifier having a high voltage gain using a current source and a large resistance.

In the triode region for small V_{ds}, the transistor acts as a variable resistance.

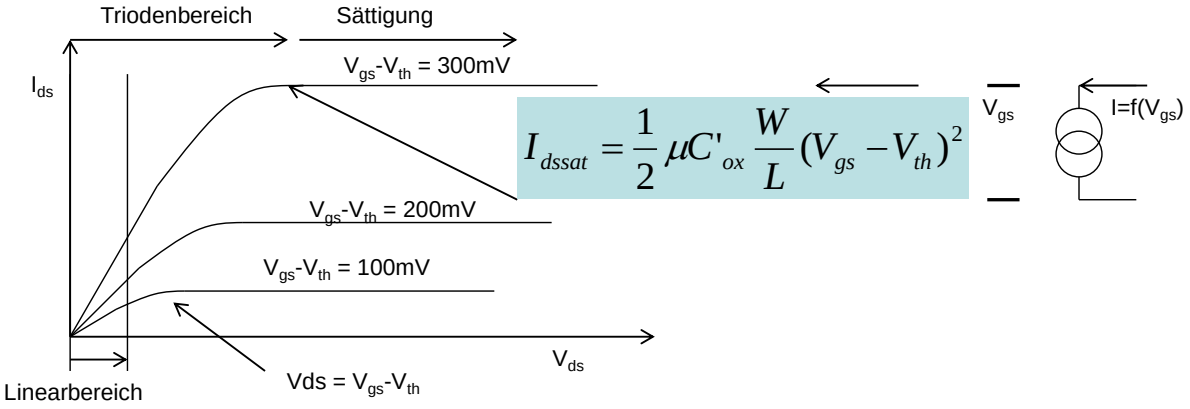


Figure 3: Output characteristice, saturation

Figure 4 shows the characteristic of I_{ds} - V_{gs}. We show in Figure 4 only the currents at the beginning of the saturation: I_{dssat} as function of V_{gs}.

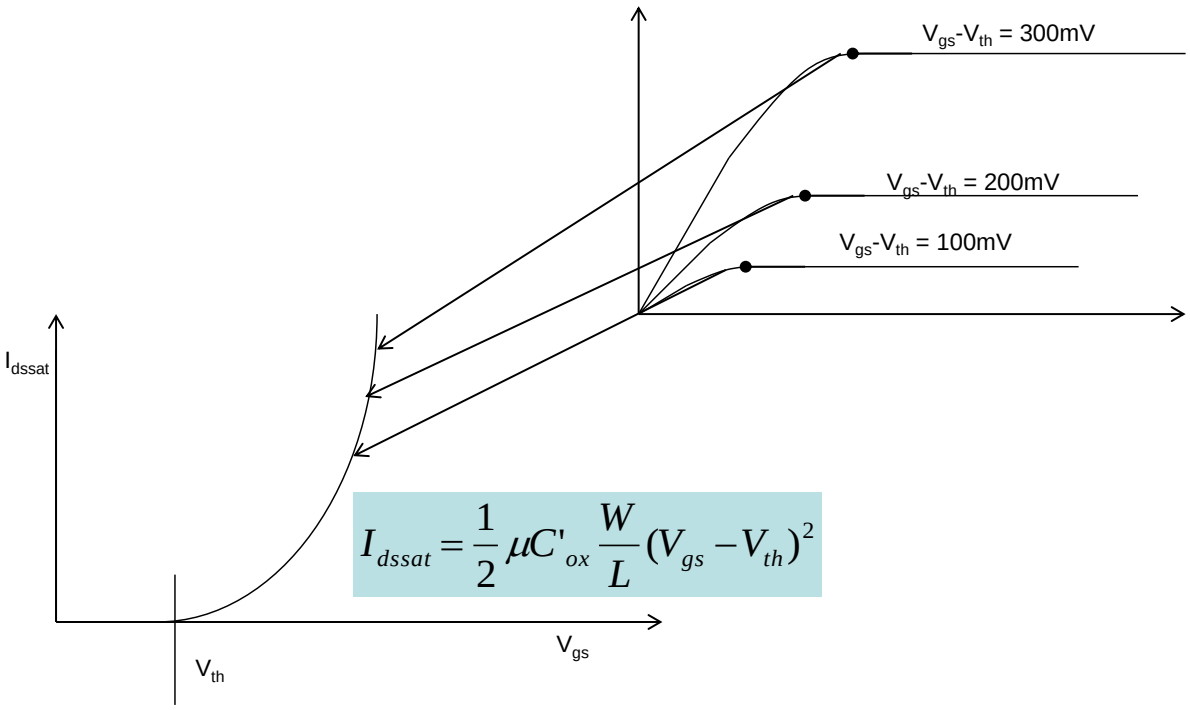


Figure 4: Input characteristics

Small signal model:

The input characteristics is usually linearized in the region around the operation point. The slope dI_{dsat}/dV_{gs} is called the **transconductance** (g_m), Figure 5.

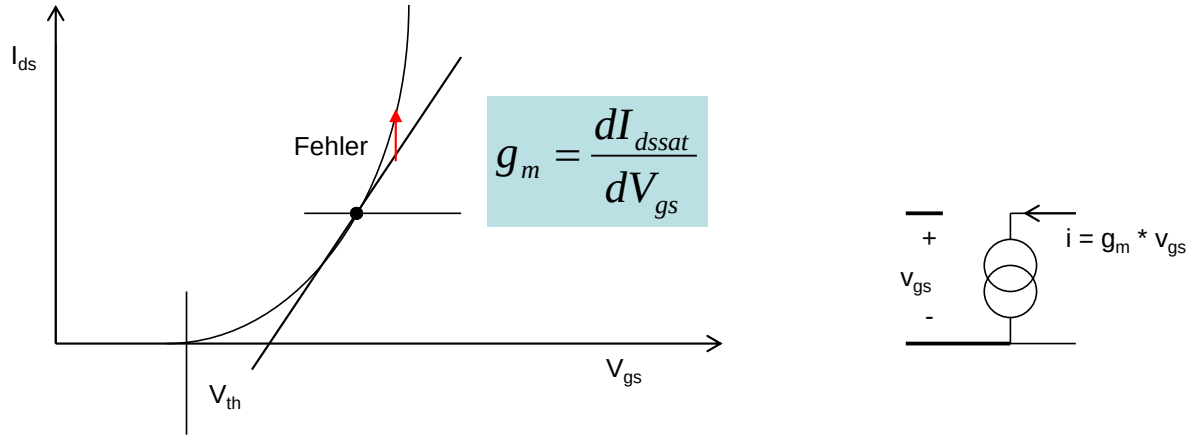


Figure 5: Transconductance

Using g_m as parameter we can derive the small signal model of a MOSFET shown in Figure 5, right.

Notice that the small signal models are valid under certain conditions. The small signal model allows mathematically a large positive and negative v_{gs} and i_{ds} values (small signals). However, the negative small signal current must not exceed the DC current, otherwise the total current would be negative.

Differences between PMOS und NMOS

In the case of the PMOS, the I-V characteristics lines are equal as in the case of the NMOS if the signs of voltages and currents are changed, Figure 6.

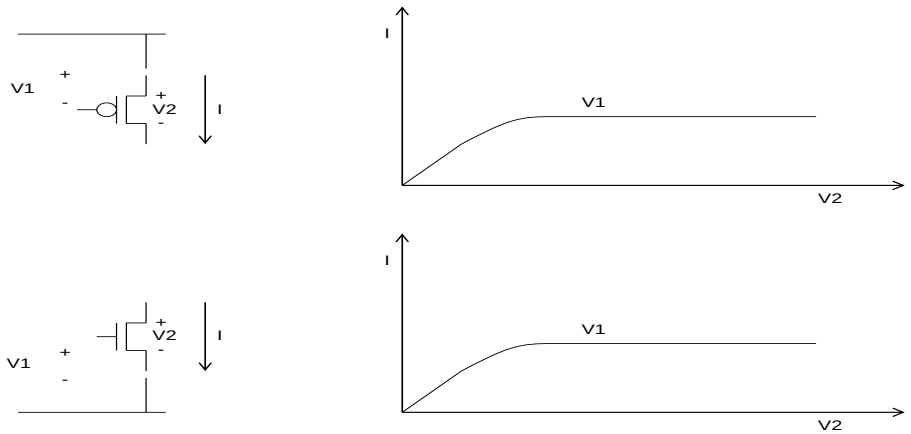


Figure 6: I-V characetristics of PMOS (top) and NMOS (bottom)

PMOS circuits are often a mirror image of NMOS circuits in the way the Figure 7 shows.

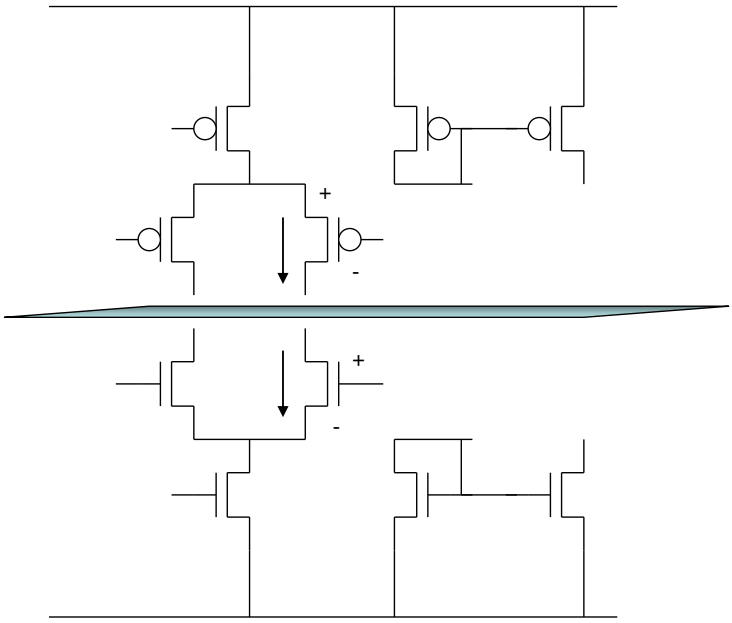


Figure 7: PMOS and NMOS circuits are often symmetrical

The currents and voltages have opposite signs. We will draw circuits in the way that the currents flow from top to bottom and the potentials above in the image are higher than the potentials below.

It is important to determine the operation region (triode-, saturation-region) for every transistor. In analogue circuits, transistors operating is saturation are especially useful. The condition for saturation is $V_{ds} > V_{gs} - V_{th}$. This means for an NMOS that the drain potential may be lower than the gate potential. Figure 8 and Figure 9 show transistors that work in saturation and in linear region.

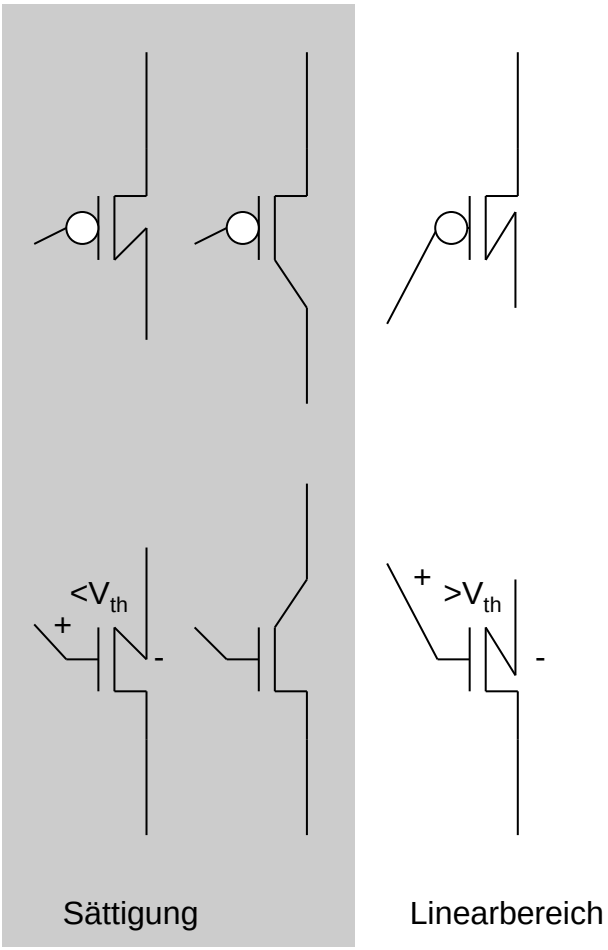


Figure 8: NMOS and PMOS transistors in saturation and linear region. The position of a node or line illustrates its potential. If the position is low, the potential is low as well.

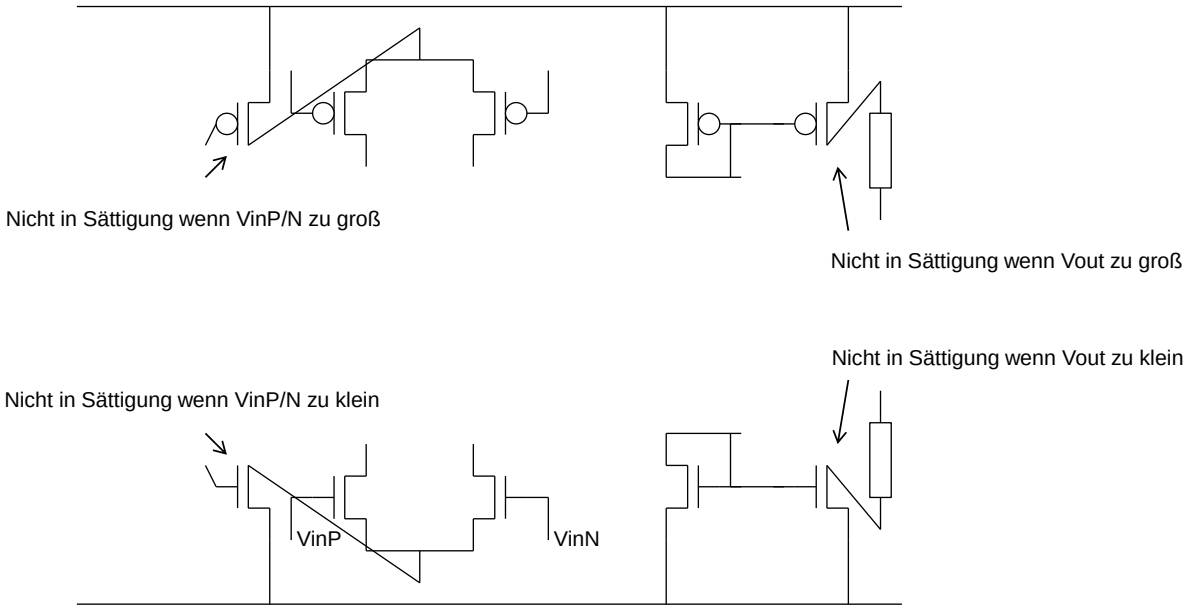


Figure 9: NMOS and PMOS circuits

Why do we need NMOS and PMOS transistors?

An NMOS transistor conducts current, only when its source potential is low. This means, for instance, that an NMOS-based electronic switch cannot be used to short a circuit with VDD (with positive supply). This is one of the reasons why we need PMOS transistors.

Example: A capacitor is discharged with an NMOS transistor T1, Figure 10. The capacitor is then charged with another NMOS transistor T2. The voltage at the capacitor does not reach VDD. When V_{gs} of T2 becomes lower than V_{th}, the transistor becomes off and it cannot charge the capacitor further. (We neglect the subthreshold current.)

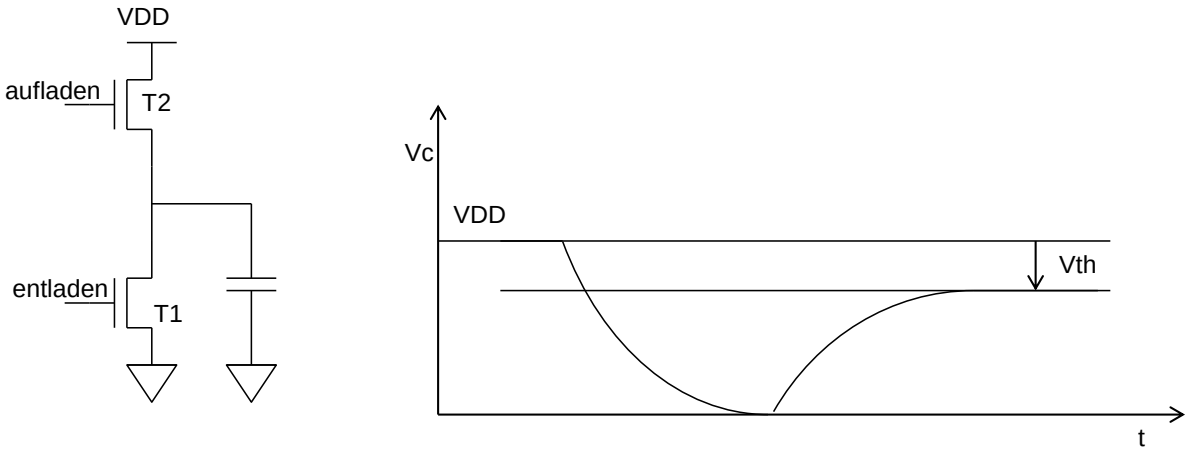


Figure 10: A capacitor is discharged and charged with NMOS transistors.

If the capacitor is charged with PMOS transistor, as Figure 11 shows, the capacitor voltage reaches VDD quickly.

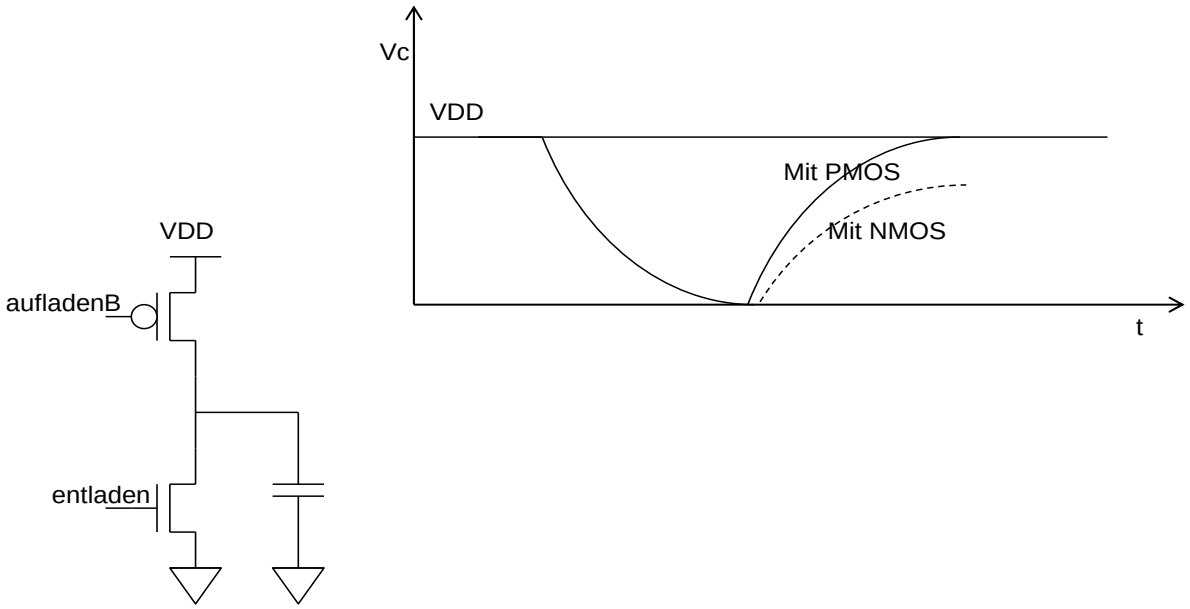


Figure 11: A capacitor is discharged with NMOS transistor and charged with PMOS transistor.

There are several differences when NMOS and PMOS transistors are used. For instance, in the case of a PMOS current source, Figure 12 right, the current flows out of VDD. An NMOS source conducts the current (drains the current) to GND, Figure 12 left.

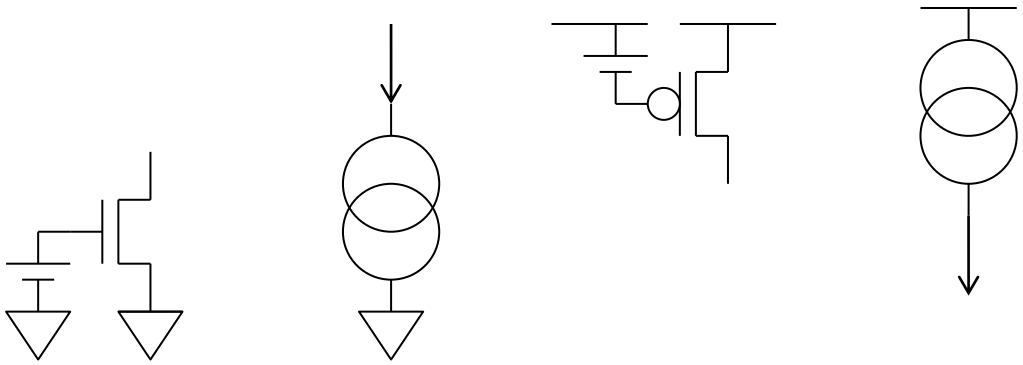


Figure 12: Current sources made with NMOS and PMOS transistors

Body-effect (substrate-effect)

In our previous MOSFET analysis, we had, both, the source and the substrate contact at 0 V.

These two contacts are not always shorted (short circuited) together. The substrate contact is often at a lower potential than the source and drain contacts.

Figure 13 shows the NMOS structure. Substrate contact is initially at 0 V (figure left), the gate source voltage is 0.5 V, which means in our case at the threshold voltage (V_{th}). The potential of the substrate - gate oxide interface (V_x) is 0 V. The interface has equal potential as the source and the drain. For higher V_{gs}, a channel is formed.

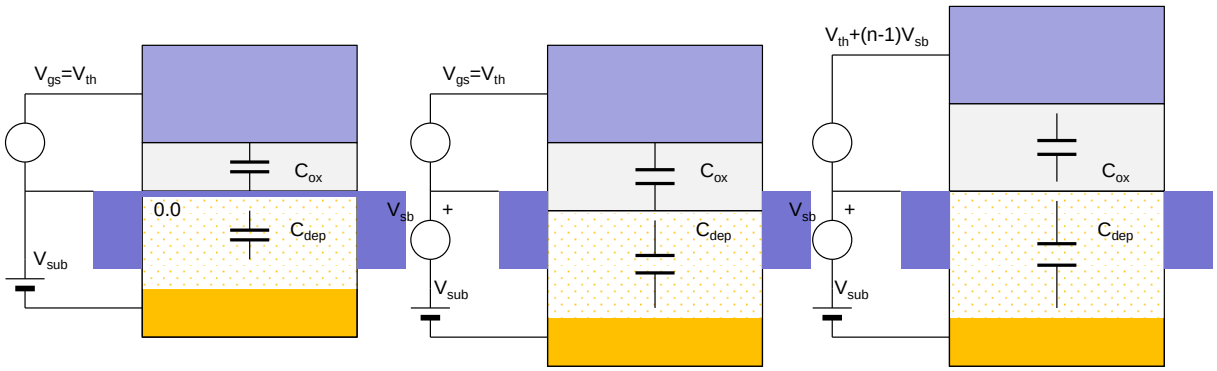


Figure 13: Illustation of the body effect

Let us now reduce the substrate potential by a small value V_{sb}, as shown in Figure 13 middle. As explained in previous lecture, there is a capacitive voltage divider in the MOS structure. The potential of the substrate – SiO_2 interface becomes:

$$V_x = -\frac{C_{dep}}{C_{dep}+C_{ox}} V_{sb} \quad (1)$$

V_x now too low for channel formation.

The choice of C_{dep} was a bit tricky in lecture 2. In this lecture it is easier. Since we calculate just a small difference of V_x voltage with respect to the value for $V_{gs} = V_{th}$ (and $V_{dep} = V_{cont}$) it is justified to take for C_{dep} the dynamic capacitance

$C_{dep} = C_{dep,min}$. (2)

$C_{dep,min}$ is the dynamic capacitance of the depletion region for the case $V_{dep} = V_{cont.}$

If the potential of the substrate-SiO_2 interface is $-|V_x|$. V_{gs} must be increased by $+|V_x|$ x $(C_{dep} + C_{ox})/C_{ox}$ in order to obtain the interface potential of 0 V. For the interface potential = 0V the channel is formed because in this case there are no potential barriers for the electrons from the source and the drain.

The additional V_{gs}-voltage required to remove potential barrier $|V_x|$ x $(C_{dep} + C_{ox})/C_{ox}$ can be understood as threshold voltage increase, denoted as ΔV_{thsb}.

It holds:

$$\Delta V_{thsb} = |V_x| \frac{C_{dep}+C_{ox}}{C_{ox}} = V_{sb} \frac{C_{dep}}{C_{dep}+C_{ox}} \frac{C_{dep}+C_{ox}}{C_{ox}} = V_{sb} \frac{C_{dep}}{C_{ox}} = (n-1)V_{sb} \quad (3)$$

Please take a look at Body-effect summary.

Check (optionally)

We have derived in lecture 2 the formula (17) for $C_{dep,min}$:

$$C_{dep,min} = A\sqrt{\frac{eN_a\varepsilon_0\varepsilon_{Si}}{2V_{cont}}} \qquad (4)$$

If we substitute (4) in (3), we obtain:

$$\Delta V_{thsb} = V_{sb}\frac{\sqrt{\frac{eN_a\varepsilon_0\varepsilon_{Si}}{2V_{cont}}}}{C'_{ox}} (5)$$

C'$_{ox}$ is the capacity per unit area.

Let us try to derive the same formula starting from the formula for threshold voltage (lecture 2 - 18):

$$V_{th} = \frac{\sqrt{2eN_a\varepsilon_0\varepsilon_{Si}V_{cont}}}{C'_{ox}} \quad (6)$$

Additional voltage V_{sb} causes increase of the voltage across depletion region from V_{cont} to $V_{cont} + V_{sb}$.

The threshold voltage increases to (we substitute V_{cont} with $V_{cont} + V_{sb}$ in (6)):

$$V_{thsb} = \frac{\sqrt{2eN_a\varepsilon_0\varepsilon_{Si}(V_{cont}+V_{sb})}}{C'_{ox}} \quad (7)$$

For $V_{sb} < V_{cont}$, (7) can be simplified as follows by using the linear term of Talyor series:

$$V_{thsb} = \frac{\sqrt{2eN_a\varepsilon_0\varepsilon_{Si}V_{cont}(1+V_{sb}/V_{cont})}}{C'_{ox}} = V_{th}\sqrt{1+V_{sb}/V_{cont}} = V_{th} + \frac{V_{th}V_{sb}}{2V_{cont}} \quad (8)$$

It follows:

$$\Delta V_{thsb} = \frac{V_{th}V_{sb}}{2V_{cont}} = V_{sb}\frac{\sqrt{2eN_a\varepsilon_0\varepsilon_{Si}/2V_{cont}}}{C'_{ox}} \qquad (9)$$

This is the same formula as we derived using voltage divider (5).

Body-effect summary

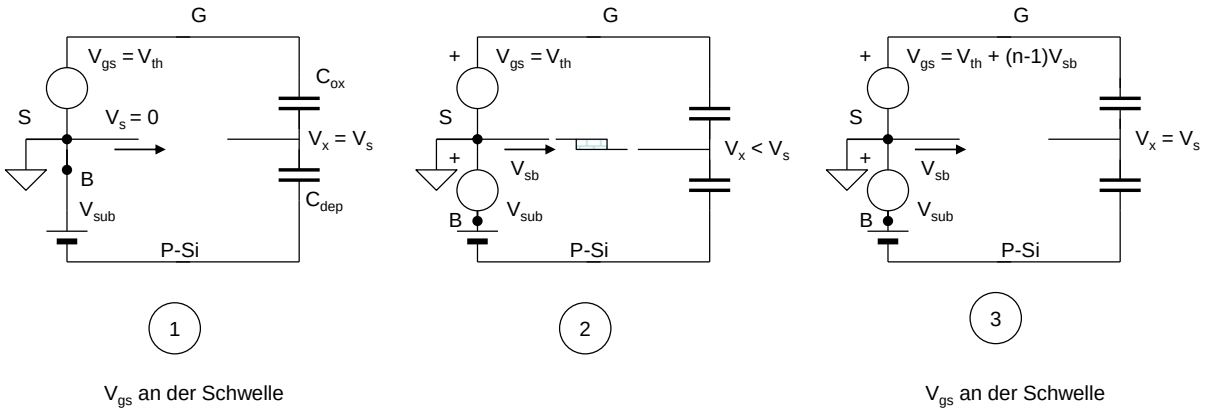


Figure 13b: G - gate, S - source, B – substrate contact, V_{sub} – contact potential, X – interface between Si und SiO_2, P-Si – p-silicon substrate (deep, undepleted region). We define C_{ox} und C_{dep} in order to derive various formulas (e.g. for the threshold voltage) by means of circuit analysis methods. The threshold is defined as the V_{gs} voltage when: $V_s = V_x$. If $V_s = V_x$, there

is no potential barrier for electrons coming from the source. Fig. 1: Source is the reference point (labeled with ground symbol). Bulk and source are shorted. It holds: $V_s = V_x$, V_{gs} is at the threshold ($V_{gs} = V_{th}$). Fig. 2: Substrate contact is at lower potential $-V_{sb}$ then source. There is a potential barrier for electrons between S and X. Fig. 3: If we increase V_{gs} by $(n-1)V_{sb}$, we obtain $V_s = V_x$ and V_{gs} is again at the threshold. The threshold voltage is now higher: $V_{gs} = V_{th} + (n-1)V_{sb}$.

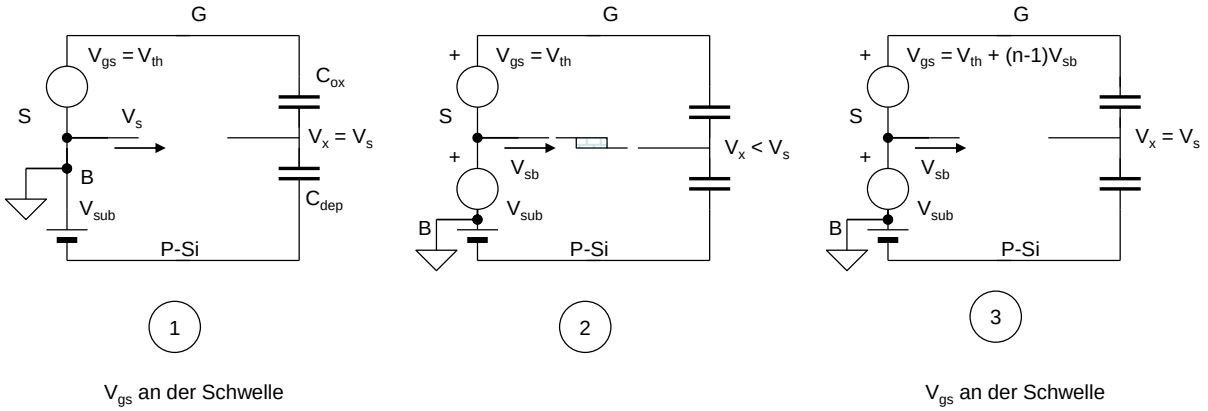


Figure 13c: The same holds when we chose the substrate as reference point. Fig 1: There is no voltage between B and S. The threshold voltage is V_{th}. Fig 2: Source is at a higher potential than bulk. A potential barrier is induced. The threshold voltage is higher than V_{th}. Fig. 3: We increase V_{gs} in order to remove potential barrier. ($V_s = V_x$). The new threshold is $V_{th} + (n-1) V_{sb}$.

Influence of the body effect to the saturation of the drain current (optionally)

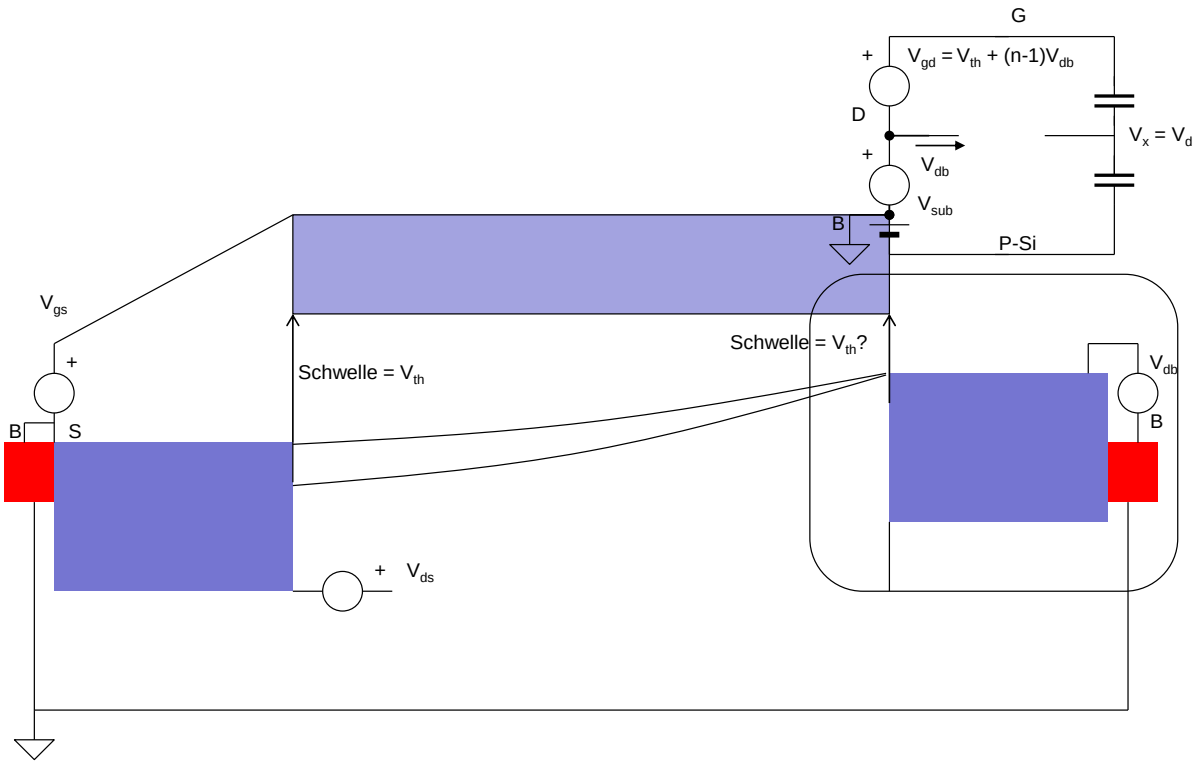


Figure 14: *In lecture 2 we have defined the saturation voltage V_{dssat}. It is the voltage when V_{gd} becomes equal to V_{th}. The channel becomes pinched off. We have assumed that the threshold close to the drain is equal as the threshold voltage close to source. However, the drain potential is higher as the potential of the substrate (B). For this reason, there is a body effect at the drain side of the channel. The threshold voltage close to drain is by $(n - 1)\,V_{db}$ higher than V_{th}. Therefore, the saturation happens earlier and the saturation voltage is smaller when we take body effect into account.*

Notice that there is a body effect at the drain side of the channel as well (Fig 14). The threshold increase close to drain is:

$$\Delta V_{thdb} = (n-1)V_{db}$$

It holds:

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

V_{thsb} und V_{thdb} are the threshold voltages at the source- und drain-sides when body effect is taken into account.

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

$$V_{thdb} \equiv V_{th} + (n-1)V_{db} = V_{thsb} + (n-1)V_{ds} \quad (1)$$

In lecture 2, we have derived the following condition for saturation:

$V_{dssat} = V_g - V_{th}$

Let us now correct this formula by taking body effect into account:

$V_{dssat} = V_{gs} - V_{thdb}$ (2)

If we substitute (1) in (2) and if we solve the equation to obtain V_{ds} (=V_{dssat}), we get:

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

The equation for transistor current can be then modified:

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

or

$$I_{ds} = \frac{1}{2n}\mu C'_{ox}\frac{W}{L}\left(V_{gs} - V_{th} + (n-1)V_{sb}\right)^2$$

The substrate potential influences the I_{ds} transistor current in a similar way as the gate potential, only by C_{dep}/C_{ox} ~ 0.25 weaker.

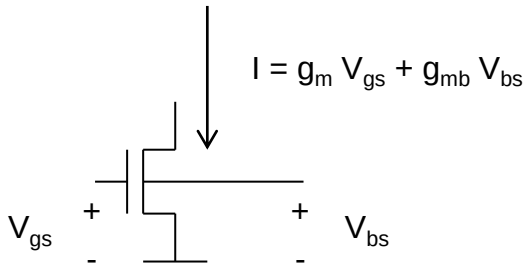


Figure 15: Body effect – smal signal model

Body effect can be modelled with the circuit in Figure 15.

Channel length modulation (Early-Effekt)

We have seen that the channel length L and the channel width W determine the transistor current (I_{ds} ~ W/L). How big are W and L? In the first approximation, the channel takes 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 the same size as the gate oxide.

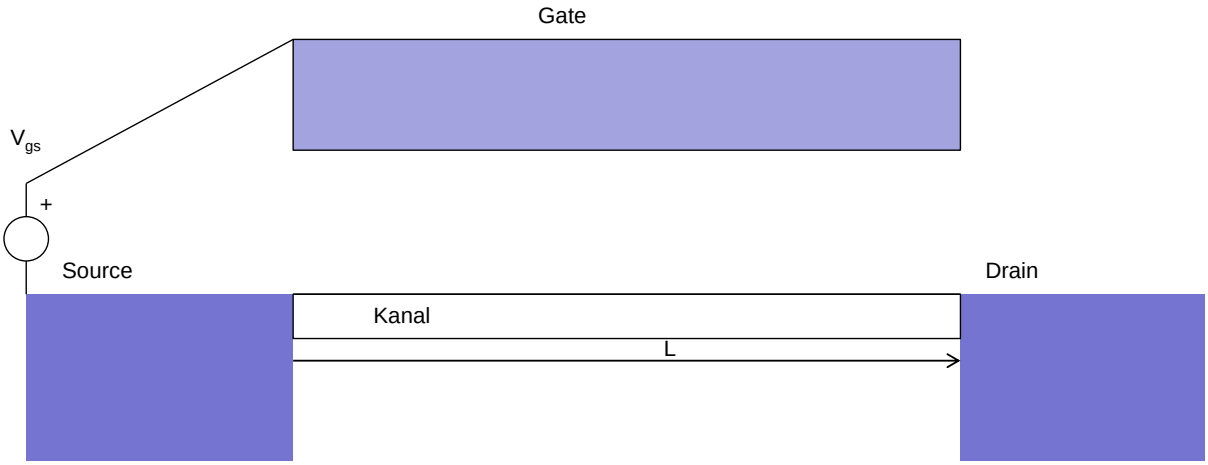


Figure 16: Channel length

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.

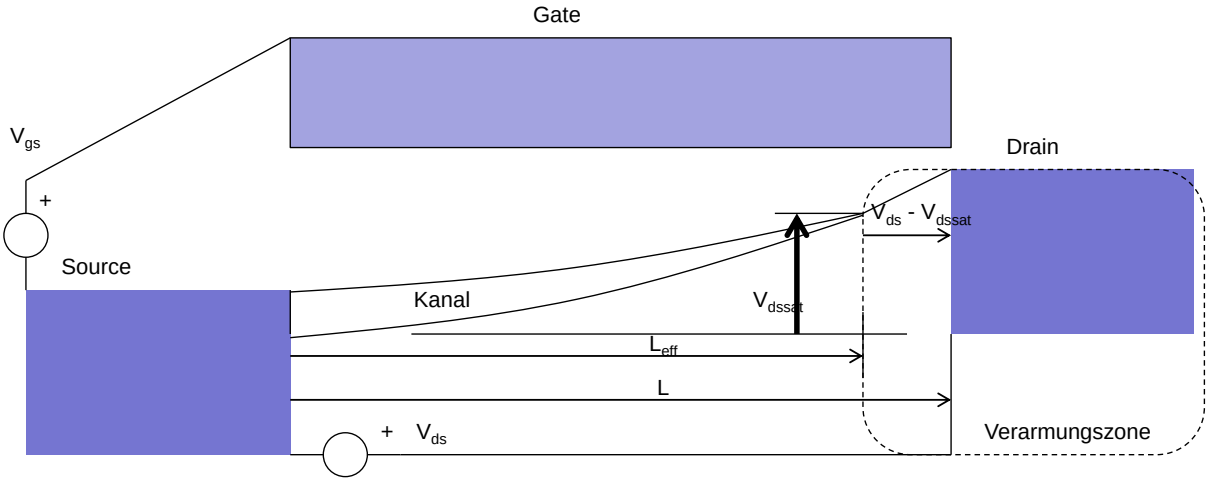


Figure 17: Early effect

The current in saturation region is given by the equation:

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

When V_{ds} becomes greater than V_{dssat}, the channel length L gets smaller and the current increases according to following formula:

$$I_{ds} = I_{dssat} + \frac{\partial I_{dssat}}{\partial L}\Delta L$$

It can be derived:

$$I_{ds} = I_{dssat}(1 + \frac{\Delta L}{L})$$

When we take into account the following relation:

$$\Delta L \sim (V_{ds} - V_{dssat})/E_{sat}$$

(E_{sat} is a constant), we get the formula for I_{ds} that takes the channel length modulation into account:

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

The slope dI_{dssat}/dV_{ds} is:

$$g_{ds} = \frac{\partial I_{ds}}{\partial V_{ds}} = \frac{I_{dssat}}{E_{sat}L}$$

We can derive the following small signal resistance:

$$r_{ds} = 1/g_{ds} = \frac{E_{sat}L}{I_{dssat}}$$

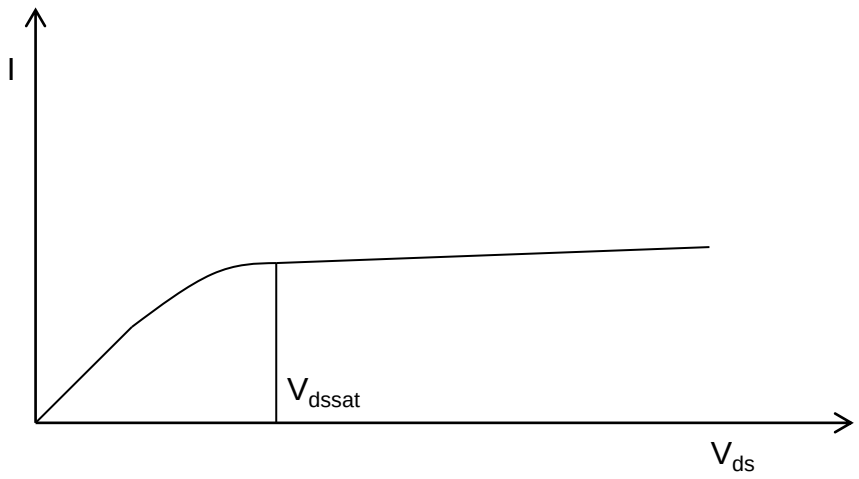
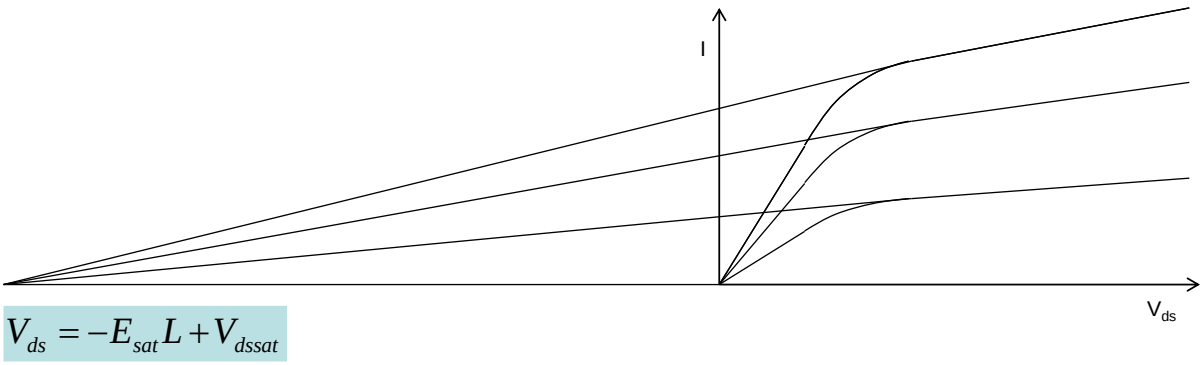


Figure 18: Early effect leads to current increase for $V_{ds} > V_{dssat}$.

One interesting result follows from the equation:

$$I_{ds} = 0 => \frac{(V_{ds} - V_{dssat})}{E_{sat}L} = -1$$

All extensions of the lines $I_{ds}(V_{ds})$ intersect x-axis in one point.



$$V_{ds} = -E_{sat}L + V_{dssat}$$

Figure 19: Extensions of the lines intersect in one point

A small g_{ds} is usually advantageous because the transistor behaves more like an ideal current source. Long transistors that conduct small currents usually have small g_{ds} values (or a large $r_{ds} = 1/g_{ds}$ resistances).

Figure 20 (left) shows the transistor with small g_{ds}, i.e. large r_{ds}. Notice that this transistor has small g_m. We often cannot get both from a single transistor: a good current source and a high transconductance. There is a trick called cascode that allows to optimize both properties. A cascode, Figure 21, has both a large g_m and a large r_{ds}.

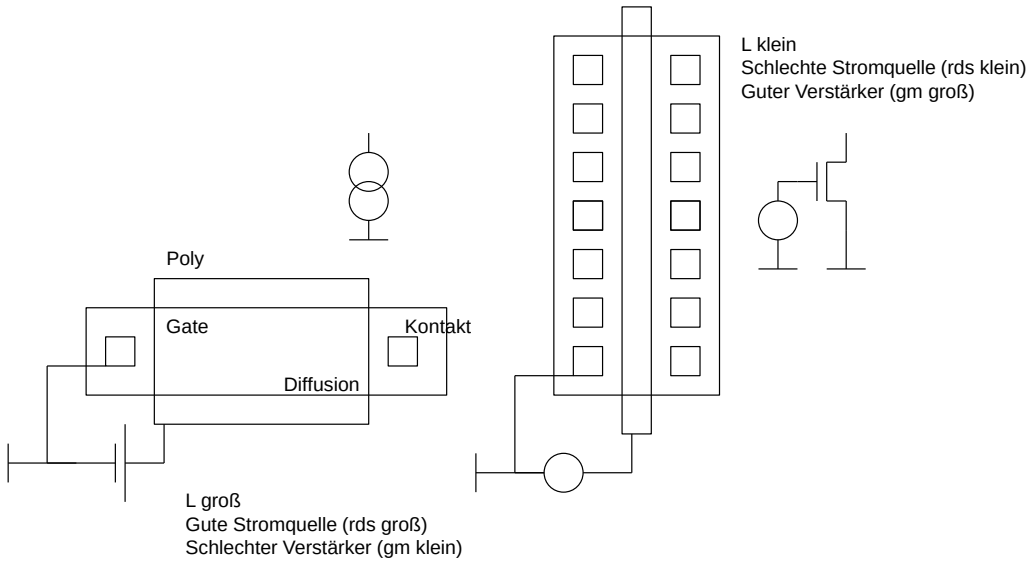


Figure 20: Different current sources

Kaskode (wird später erklärt)
Gute Stromquelle (rout groß)
Guter Verstärker (gm groß)

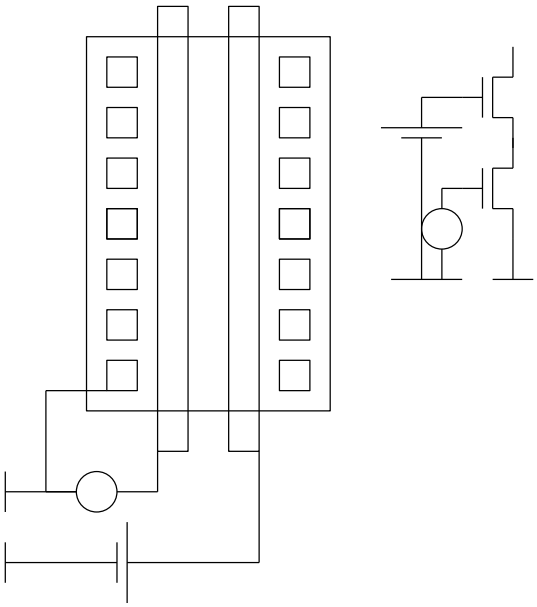


Figure 21: Cascode