2 Diodes and Transistors

Student Group

First Name	Surname	Matrikel Nr.

Table of Contents

2. Diodes and Transistors

A nice introduction can be found in KIT Bridge Course - 4.3.6 Diodes and Transistors (*).
Some of the following passages, videos and pictures are taken from this introduction.

For another introduction, see LEIFIphysics.

Introductory example

The electronics in personal computers, mobile phones, electric toothbrushes, and like all other digital companions, are based on transistor circuits (see at_the_heart_of_a_computer). In fundamentals_of_digital_technology it has already been explained that all logic circuits can be traced back to NAND and NOR gates, respectively, via conjunctive and disjunctive normal forms. These in turn consist of transistors. In the simulation below, the structure of a NAND gate is shown in the current CMOS structure. CMOS here indicates the structure of the circuit and semiconductor structure: **C**omplementary **m**etal-**o**xide-**s**emiconductor - an oppositely complementary circuit of semiconductors of the metal-oxide-semiconductor structure. The complementary structure is shown by the fact that.

- from the digital output ($OUT2$) to ground two transistors of one kind are connected in series and
- from the digital output ($OUT2$) to the 5V supply, two transistors of a different type are connected in parallel.

These two different kinds of MOS-transistors and further used kinds shall be explained in this chapter.

Targets for the bipolar transistor

After this lesson, you should:

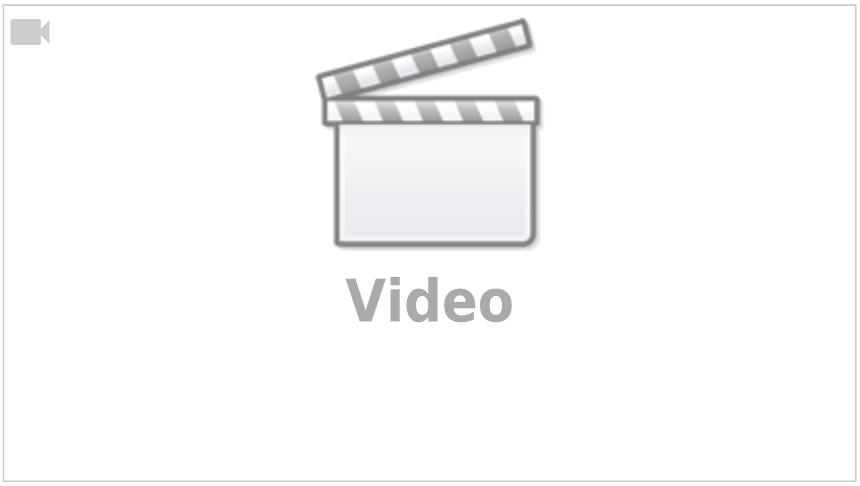
1. Know what types of bipolar transistors there are, what their layer structure and circuit symbol looks like.
2. Know how the two types of bipolar transistors are controlled.
3. Know what are the main characteristics of the bipolar transistor and what they look like.

Targets for the field-effect transistor

After this lesson, you should:

1. Know what types of MOSFETs there are, what their layer structure and circuit symbol looks like.
2. Know what the output characteristic field of the MOSFET looks like.
3. know what the body diode is and where it comes from.
4. know what to look for in the output characteristic field when designing a semiconductor element.

2.6 Functional principle of a (bipolar) transistor



A variable resistor can be developed from the diode or PN junction. With this controlled transition resistor (“transfer resistor” or better transistor) the resistance can be changed by a current and thus the current let through can be adjusted.

Video-Transcript (Alternative to the explanation in the video)

A transistor consists of two diodes connected against each other, which have a common n- or p-layer, e.g. a thin p-doped layer is placed between two n-doped layers. This is an npn transistor, the more common design. However, pnp transistors are also used for special applications. All three layers are electrically contacted, so the transistor has three terminals. The contact to the middle layer is called base (B), the contacts to the two outer layers collector (C) and emitter (E). The circuit symbol of a transistor is shown in figure 2.

A transistor is usually operated as a switch or as a current amplifier. To explain how it works, a typical transistor circuit is shown in the figure below. The circuit containing the consumer, here the incandescent lamp, is called the working circuit. Here, the voltage source must be connected in such a way that the technical direction of current through the transistor runs from the collector to the emitter, i.e. in the direction of the arrow indicated on the emitter.

The second circuit, in which a positive control voltage is applied to the base, is the control circuit. Holes are pumped from the p-doped base into the n-type emitter by the positive control voltage, since a negative voltage is applied to it. Thus, the base-emitter diode is forward biased. On the other hand,

a positive voltage is applied to the collector, so that this diode actually blocks.

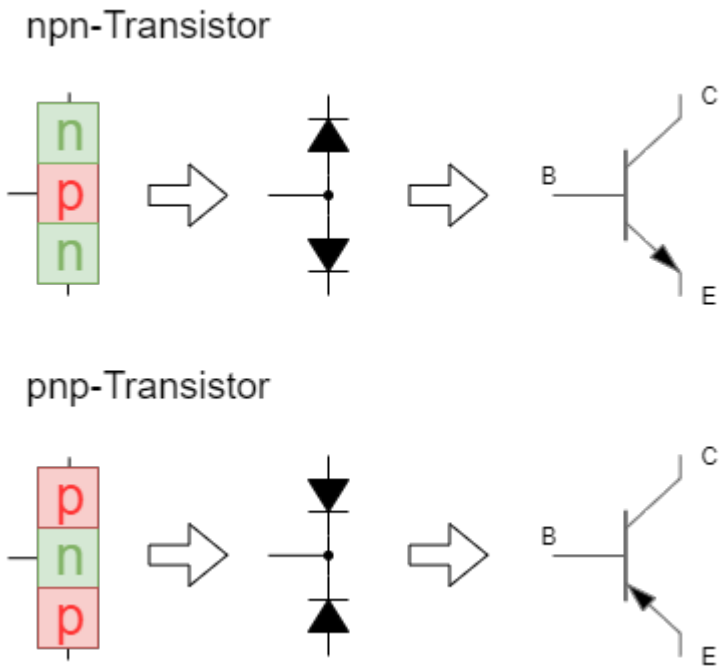
If the voltage U_{BE} in the control circuit exceeds a certain threshold, a current I_C can now flow in the operating circuit. In this respect, the transistor acts as a switch. The small current I_B in the control circuit can therefore be used to control the large current I_C in the operating circuit.

In a certain range, the current I_B in the working circuit is proportional to the current I_C in the control circuit. This ratio is called the current gain $\beta={I_C}/{I_B}$ of the transistor. This behavior can be understood by considering that the p-type base layer is very thin compared to the n-type layers. Electrons supplied through the control circuit diffuse through it rapidly, reaching $99\% $ in the collector connected to the positive terminal, and are pumped back through the working circuit into the emitter. Only a few pass through the emitter directly back into the control circuit. Therefore, the current in the control circuit is much less than the current in the working circuit.

~CLEARFIX~~~

Switching characters

Fig. 1: Switching sign and simplified structure of npn and pnp bipolar transistors



As just described, the bipolar transistor is built by a three-layer alternately doped layer structure, which corresponds to two diodes opposite and connected in series. Depending on the layer sequence (or “direction of the diodes”), pnp or npn transistors result, represented by different circuit symbols with three terminals (see figure 1). In both transistor variants, charge carriers are emitted from the emitter terminal (E) toward the collector terminal (C) if a suitable current flows through the base terminal (B). In simplified terms, the negative charge carriers of the n-doped sides could represent a current through an NPN structure if negative charge carriers were also present in the p-doped layer. The current I_C flowing with it in the technical current direction is illustrated in the circuit symbol by the arrow direction at the emitter. In the NPN transistor, the current I_C flows from the collector to the emitter. Since positive charge carriers enable conductivity in the PNP transistor, the technical current direction here points from the emitter to the collector and the arrow on the emitter points towards the collector. The direction of the arrow is similar to the direction of the diode or the PN

junction. Other mnemonic devices for the direction of the arrow are:

- “If the arrow hurts the base, it´s **pnp**”
- “If the arrow wants to separate from the base, it is **npn**”.

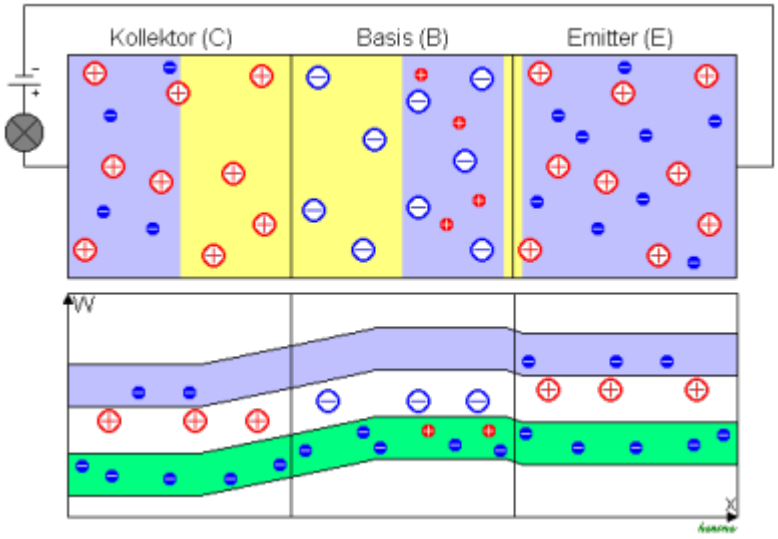
Correct wiring of the transistors

The simulation on the right shows the correct connection of the transistors. In general, the arrow of the symbol of the technical current direction must point at the correct interconnection. The base current I_C is almost always generated in the circuits by a voltage source between base and emitter with a voltage U_{BE}. In this case, a positive voltage with respect to the emitter is required for the NPN transistor and a negative voltage with respect to the emitter is required for the PNP transistor. In practical applications the NPN transistors predominate, among other things because the negative charge carriers used there produce a higher conductivity. For the following explanations only NPN transistors are considered.

A central question that arises from a closer look at the simulation on the right is: Why does a technical current flow into the base, i.e. positive charge carriers into the P-layer, have to be supplied for the NPN transistor? Wouldn't it be more plausible if the negative charge carriers, which are not present and needed for transport, had to be supplied?

Transistor in band model

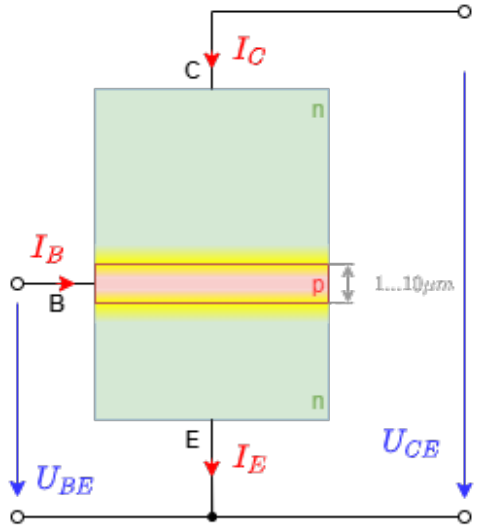
Fig. 2: Transistor in ribbon model



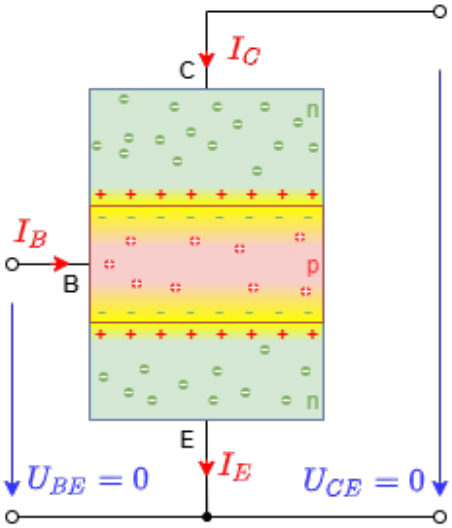
To understand this, the knowledge of the PN junction is needed. In the figure figure 2, the structure of the NPN transistor is shown in the band model. In the n-doped collector and emitter, the free-moving negative charge carriers (blue) and stationary positive charge carriers (red) are drawn, and in the base, correspondingly, the free-moving positive charge carriers (red) and stationary negative charge carriers (blue). Both PN-junctions have formed a junction. A positive voltage U_{CE} is applied to the transistor, which cannot generate any current flow in the situation shown. Due to the positive voltage U_{CE} and the missing potential at the base, the voltage U_{BE} decreases, which leads to a reduction of the junction. In contrast, the voltage $U_{CB} = U_{CE} -U_{BE}$ increases. Thus, the junction between the base and the collector becomes larger. When the external voltage U_{CE} is varied, there will always be at least one PN junction that is reverse biased, i.e. the

transistor will block.

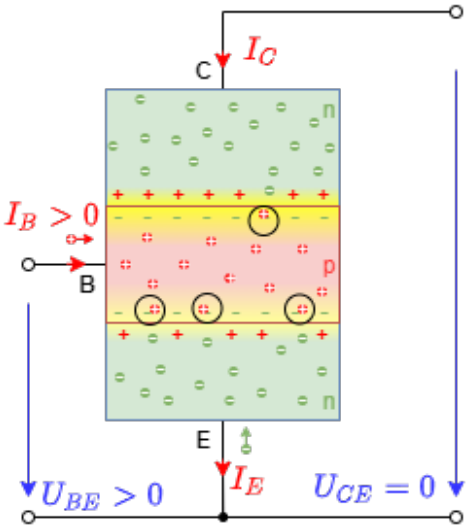
Fig. 3: function of npn bipolar transistor


C
I_C
n
I_B
B
p
1...10µm
n
E
I_E
U_{CE}
U_{BE}

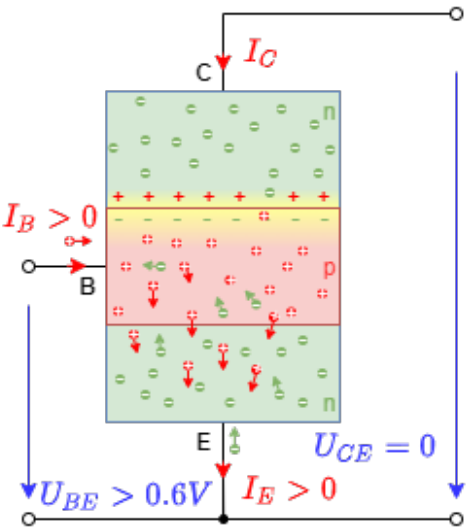
1.


C
I_C
n
I_B
B
p
n
E
I_E
$U_{BE} = 0$
$U_{CE} = 0$

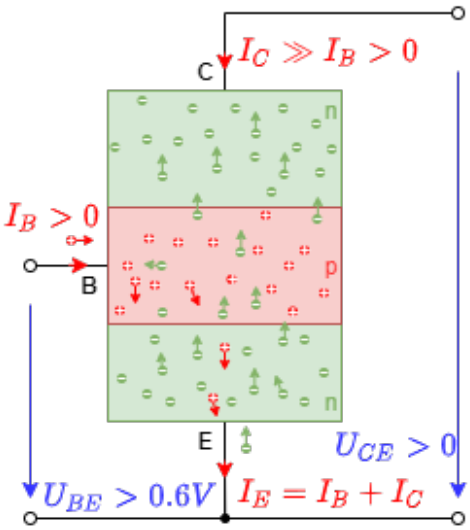
2.


C
I_C
n
$I_B > 0$
B
p
n
E
I_E
$U_{BE} > 0$
$U_{CE} = 0$

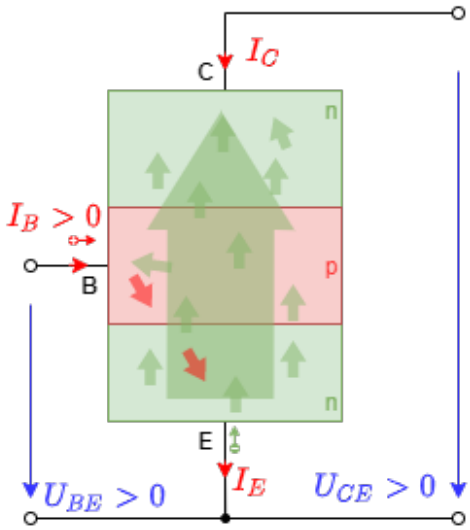
3.


C
I_C
n
$I_B > 0$
B
p
n
E
$U_{CE} = 0$
$U_{BE} > 0.6V$
$I_E > 0$

4.


C
$I_C \gg I_B > 0$
n
$I_B > 0$
B
p
n
E
$U_{CE} > 0$
$U_{BE} > 0.6V$
$I_E = I_B + I_C$

5.


C
I_C
n
$I_B > 0$
B
p
n
E
I_E
$U_{BE} > 0$
$U_{CE} > 0$

6.

To remove the junction between the collector and the base, the latter must be connected in the forward direction. Until the transistor is switched through, this is several steps, which are described below via figure 3:

1. Figure: The physics of this takes place in the narrow p-layer. The following images refer to the highlighted section.
2. Image - Situation $U_{CE}=0V, U_{BE}=0V$: In this picture the unpowered transistor is shown. In it the free charge carriers (electrons in blue, holes in red) and the junction layers between base and emitter, and base and collector in yellow. Only the junction layer shows the stationary charge carriers with their sign. As shown in the band model, the stationary charge carriers are present everywhere in both doped regions.
3. Figure - Situation $U_{CE}=0V, 0V<U_{BE}<0.6V$: First, consider a small, positive voltage U_{BE}. This provides holes in the base with current I_B. This operates the PN junction between the base and emitter in the forward direction. In the figure, it is indicated with black circles that the injected holes compensate some stationary negative charge carriers in both junction layers. Electrons also flow through the emitter into the n-region, which attenuate the junction on the other side.
4. Figure - Situation $U_{CE}=0V, U_{BE}>0.6V$: When the forward voltage of the PN junction between base and emitter is exceeded, the injected holes and electrons cancel the bottom junction. In the simulation below, it can be seen that the circuitry of the transistor is such that in the diode circuit (which is not physically correct), the diode between the base and emitter becomes conductive.
5. Figure - Situation $U_{CE}>0V, U_{BE}>0,6V$: Now with this voltage at the base, the working circuit, i.e. a voltage $U_{BE}>0$ should be present at the output. In the real system, the base is very small compared to the mean free path length of the electrons (“path to recombination with a hole”). This changes the situation at the upper PN junction. In a classical diode, no electrons are present in the p-doped region. However, the electrons present here can cross the base and compensate for the stationary positive charge carriers in the upper junction. The holes injected into the base in turn compensate for the stationary negative charge carriers. Thus, this junction layer is also removed. This is possible as long as enough holes are injected into the base.
6. Figure - Situation $U_{CE}>0V, U_{BE}>0.6V$: Thus, in the NPN bipolar transistor, both holes (to remove the junction layers) and electrons (as the “main agents” responsible for charge transport, the so-called majority carrier charges) contribute to the conductivity. This is where the name bipolartransistor comes from.

The simulation on the right shows the simplified model of the opposing diodes. The necessary input current I_C and the corresponding input voltage U_{BE} resemble the ratios of the diode between base and emitter. In figure 4 the principle of operation is shown. The current I_B across the diode between base and emitter regulates the current I_C in the working circuit. This regulation is done by the variable resistor R_{CE}.

Fig. 4: function of npn bipolar transistor

Characteristic, characteristic curves, characteristic diagrams

In the previous chapter 1 Fundamentals of Amplifiers the characteristics of a black box have already been discussed, there especially for an amplifier. The methodology can also be applied here. In the video above, the first parameter has already been described: The **current gain** $\beta=\frac{d I_C}{d I_B}$, or in the form of a graph, the **current control characteristic** $I_C(I_B)$.[1)].

Another characteristic is the **input characteristic field** $U_{BE}({I_B})$ or as differential characteristic (=slope in the characteristic) the **differential input resistance** $r_{BE}=\frac{d U_{BE}}{d I_B}$. As described earlier, the structure between the base and emitter resembles a diode. Accordingly, the input characteristic field resembles that of a diode. Since the current flow I_B is very small (a few microamps or smaller), the input resistance r_{BE} is large.

The following simulation shows the current control characteristic $I_C(I_B)$ and input characteristic(nfield) $U_{BE}({I_B})$ by varying U_{BE} (or I_B).

For the description of the transistor, the **output characteristic field** $U_{CE}({I_C})$ and the **differential collector-emitter resistance** $r_{CE}=\frac{U_{CE}}{I_C}$ present in it as a slope is particularly important. This can be seen in the following simulation for different input voltages U_{BE} (and thus different control currents I_B). The output characteristic field can be divided into different ranges:

1. junction: at low input voltages $U_{BE}< 600mV$, the junction is not degraded. Accordingly, the entire transistor becomes non-conducting. In the output characteristic field, this can be seen by the fact that when the output voltage U_{CE} is positive, the output current I_C becomes very small. In this case, the transistor on the output side corresponds to a high-impedance resistor, or an open switch.
2. Gain region (or active region): at larger input voltages $U_{BE}> 600mV$, the junction is degraded. In the gain region, the output characteristic behaves as a straight line. The output current I_C is thus only dependent on I_B, as defined by the current gain $\beta=I_C/I_B$.
3. Saturation region: The saturation region is found at larger input voltages $U_{BE}> 600mV$ and only small output voltage U_{CE}. At constant input voltage U_{BE} the output voltage behaves to the output current like a high non-linear resistor. in this case the transistor on the output side corresponds to a low impedance resistor, or a conducting switch.

In the datasheet, a different nomenclature is occasionally found, resulting from the so-called H-characteristic of quadrupole theory[2)]:

- current gain $h_{fe}=\beta(I_C, U_{CE})=\frac{I_C}{I_B}$
- input resistance $h_{ie}=r_{BE}(I_C, U_{CE})=\frac{U_{BE}}{I_B}$
- output resistance $h_{oe}=r_{CE}(I_B, U_{BE})=\frac{U_{CE}}{I_C}$

The bipolar transistor is used where a low threshold voltage or current amplifier is required. This is advantageous in various amplifier circuits, for example. Bipolar transistors are also found in some simple power supplies. The most common bipolar transistor circuit is the so-called collector circuit. This is characterized by the fact that a constant voltage - the supply voltage - is applied to the collector. Several collector circuits can be operated by a common voltage supply. This means that the

same voltage is applied to all collector connections. Because of the wide use that bipolar transistors have had, even today the common voltage supply of electronic circuits is called V_{CC}, where CC stands for **Common Collector**. This is often seen even when bipolar transistors are no longer used.

A major disadvantage of the bipolar transistor is that a control current is required for switching. Especially in digital circuits, but also in power electronics, this results in a non-negligible input power $P=U_{BE}\cdot I_B$. This leads to losses and waste heat, which must be taken into account in the power supply and thermal design. For this reason, bipolar transistors are no longer used in current microcontrollers. In these fields, the bipolar transistor has been displaced by the field-effect transistor.

Note: Bipolar transistors

There are 2 different types of bipolar transistors. These differ in the type of layer structure, or majority carrier charges:

- **npn bipolar transistors**: Major conduction occurs via electrons. These cannot pass through the p-doped region without current I_B across the base. By $I_B>0$ holes are introduced into the base, which remove junction layers.
- **pnp bipolar transistors**: The main conduction occurs through holes. These cannot pass through the n-doped region without current I_B across the base. By $I_B<0$ electrons are introduced into the base, which degrade junction layers.

In the bipolartransistor, both types of charge carriers are involved in the transport.

2.7 Operating principle of a field-effect transistor

Fig. 6: Function of the MOSFET

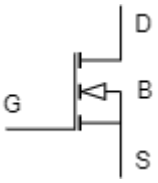


Fig. 5: FET circuit symbols

A field effect transistor (FET) also consists of two diodes connected against each other, which have a common n- or p-layer. However, the conductivity of the field-effect transistor is not generated by applying a controlcurrent, but solely by a controlvoltage. In the case of the bipolar transistor, the control current was also generated by a control voltage. However, the control current must flow continuously to drive the bipolar transistor, since the charge carriers introduced via the base recombine internally.

In figure 5 a special field-effect transistor is drawn, the so-called “metal-oxide-semiconductor field-effect transistor”. This will be explained in more detail below. The figure 6 outlines the principle of operation: the control voltage U_{GS} (in English V_{GS}) regulates the current I_D in the working circuit. This is done by the resistance between R_{DS}.

To distinguish the transistor types, and to emphasize the physics behind them, the terminals are labeled differently for the field-effect transistor:

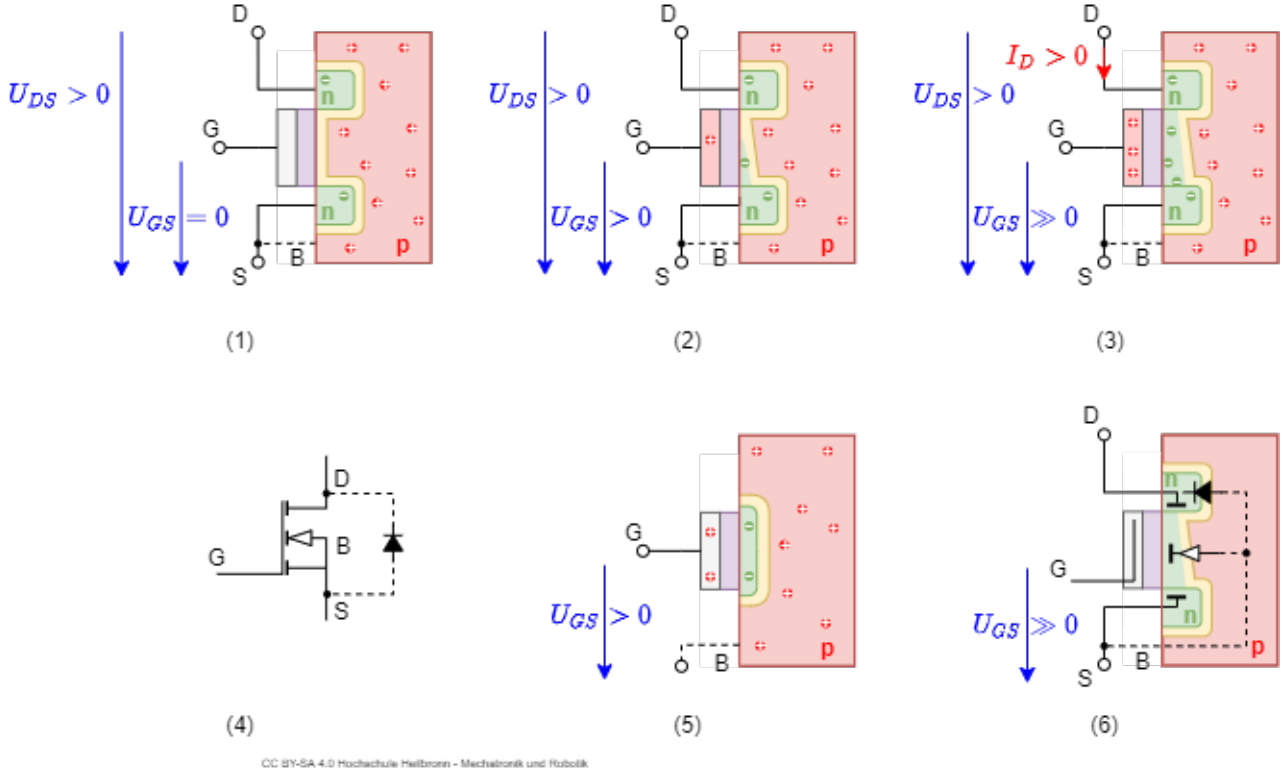
- **(S) Source**: Terminal from which the charge carriers pass through the transistor (roughly corresponds to the emitter).
- **(G) Gate**: Terminal at which a voltage can be used to change the conductivity (roughly corresponds to the base, with control currents being injected there).
- **(D) Drain**: Terminal at which the charge carriers arrive and leave the transistor (corresponds approximately to the collector).

In addition, there is the **“Bulk” (B)** in the structure, which refers to the basic substrate of the transistor. This is usually not led out separately, but shorted to the source terminal. In some FETs, the bulk is represented by the middle connection.

In the simulation on the right, you can see that the field-effect transistor behaves much like a switch, which is controlled by a voltage. No current seems to flow on the gate, but when the voltage on the gate changes, the behavior changes from “conductive” to “open”.

Metal Oxide Semiconductor Field Effect Transistor (MOSFET)

Fig. 7: MOSFET layering



The structure of the metal oxide semiconductor field effect transistor (**Metal Oxide Semiconductor Field Effect Transistor: MOSFET)** resembles the bipolar transistor at first glance. In figure 7, the individual figures (1)...(3) show the layering of an n-channel (English n-channel) MOSFET, and in (4) the circuit symbol is shown again. In contrast to the npn-bipolar transistor, the middle p-doped layer (bulk) is not directly connected to the control electrode. Rather, the metal layer of the gate (figure 7, Fig. (5), gray), the insulating layer of the oxide (shown in purple), and the conductive p-doped layer of the bulk (shown in red) form a capacitor. It should be noted that the bulk is at the potential of the source connection (dotted line in the picture).

Without voltage difference U_{GS} between gate and source, a (small) junction is formed at the p-n junctions. If the voltage difference U_{GS} is increased, the capacitor between gate and bulk is charged. This accumulates electrons opposite the gate electrode (figure 7, Fig. (2), dark blue “wedge”). If the voltage difference U_{GS} exceeds a certain threshold voltage, the enriched electrons form a channel between source and gate. This allows a current $I_D \gg 0$ to flow through the MOSFET (figure 7 Fig. (3)).

The switching symbol (figure 7, figure (4)) can also be described as follows: Capacitors form between gate and source, between gate and base, and between gate and drain, respectively, in the off state because of the oxide layer (purple in Fig. (1))[3]. To drive the MOSFET, the voltage at the gate U_{GS} must be such that a PN junction forms in the bulk, indicated by the white filled triangle in figure (4). Since the apex of the triangle (or the diode symbol sketched with it) points toward the gate, it is clear that we are dealing with an n-channel MOSFET.

In the simulation on the right, the same voltage ratios are shown as in figure 7 (1)...(3). The toggle switch on the left makes it possible to invert the voltage U_{DS} across the transistor. If this becomes negative, a slightly different situation arises: The MOSFET appears to become conductive regardless of what voltage U_{GS} assumes. This is due to the fact that another diode has been hidden in the layer structure: a junction has formed between the bulk (p) and drain (n), which is operated at $U_{DS}<0$ and with the bulk and source connected in the forward direction. This so-called body diode is explicitly built into the simulation at (3b).

Output characteristics of the MOSFET

The **output characteristic field** $U_{DS}({I_D})$ is also to be considered for the MOSFET. This is also similar to the bipolar transistor, but now the different characteristics are adjustable by different control voltages U_{GS} and not by a control current.

Unfortunately, the naming of the different operating ranges of a MOSFET differs from that of the bipolar transistor:

1. Blocking range: at low input voltages U_{GS}, no channel can be formed. Accordingly, the entire transistor becomes nonconducting. In the output characteristic field this can be seen by the fact that at positive output voltage U_{DS} the output current I_D becomes very small.

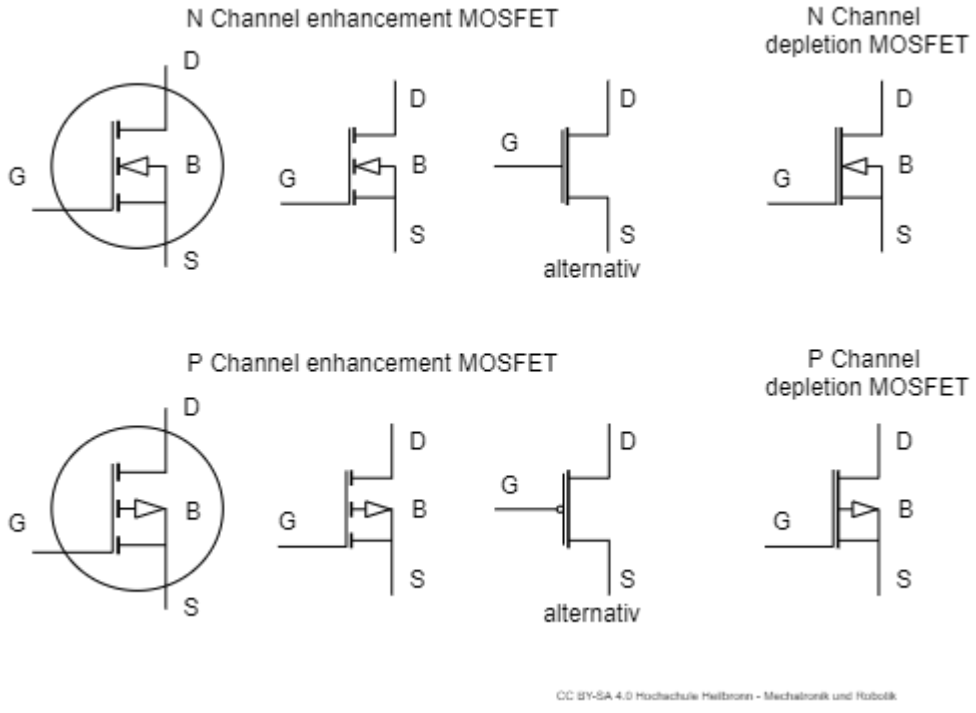
In this case, the transistor corresponds to a high-impedance resistor, or an open switch, on the output side.

2. Saturation region: for larger input voltages $U_{GS}> U_{th}$ above a threshold, a conductive channel is formed. In the saturation region, the output characteristic behaves like a straight line. The output current I_D is thus only dependent on U_{GS}.
3. linear region (active region) : The linear region is found at larger input voltages $U_{GS}> U_{th}$ and only small output voltage U_{DS}. At constant input voltage U_{GS}, the output voltage to output current behaves like a high non-linear resistor. In this case, the transistor on the output side corresponds to a low-impedance resistor, or a conducting switch.

It should be noted that the saturation region for MOSFET and bipolar transistor characterizes different operating ranges.

Variants of MOSFETs

Fig. 8: FET circuit symbols



The so far considered (and also most frequently used) field effect transistor is the so-called “**n-channel enhancement type MOSFET**”. The part “n-channel” comes from the type of the current-forming charge carrier and was already given above. The part “enhancement type” represents, that the charge carriers are not present at first and have to be accumulated in the bulk by means of the voltage U_{GS} for conductivity.

Some circuits (especially digital circuits) also use “**p-channel enhancement type MOSFET**”, where holes are the current-forming charge carriers. In the simulation on the right, this type of MOSFET is shown. Most clearly, when the p-channel enhancement type MOSFET is connected, the drain and source are generally reversed. Thus, the numerical values of U_{DS} and I_D in the output characteristic field become negative. To enrich holes in the p-channel, a negative voltage must be applied to the gate $U_{DS}<0$.

In the figure 8 the circuit symbols of different variants of MOSFETs are shown. In the MOSFETs in the top row, an n-channel is formed for charge transport, and in the bottom row, a p-channel is formed.

Three variations of an **n-channel enhancement type MOSFET** are shown in figure 8 in the upper left. In the first circuit symbol, the circle represents that it is a discrete device, i.e., a single MOSFET not integrated with others in a chip. The second circuit symbol has already been used in the previous chapters. The third circuit symbol of the same n-channel enhancement type MOSFET is the reduced version (i.e., without bulk). This representation is used for simplification in digital circuits.

In figure 8 on the lower left, three variations of a **p-channel enhancement type MOSFET** are shown. Again, the circle on the first circuit symbol indicates that it is a discrete device, but now the direction of the arrow on the bulk is rotated. The second switching symbol is used in the same way as for the n-channel MOSFET - in integrated circuits. The third symbol is again the reduced version (without bulk). For the digital circuit it is only important whether the switch closes or opens at high signal ($= 5V$). Since the p-channel enhancement type MOSFET opens, this is drawn with a negation sign (small circle) at the gate.

In figure 8 on the right, so-called **n-channel and p-channel depletion-type MOSFET** are shown. The MOSFETs considered so far were not conductive in the off state (i.e. $U_{DS}=0$). However, in some applications, it would be good if the MOSFET resembled a conductive switch when off. Looking at the layer structure (figure 7, Figure (1)…(3)), this is possible via selective redoping of the region opposite the gate. The doping can be used to dislocate a conductive channel. The charge carriers of this channel can be displaced or depleted by a suitable field - and thus suitable gate voltage U_{GS}. Thus, the MOSFET becomes non-conducting in the presence of a reverse voltage U_{GS}. In the circuit symbol, the “short circuit” between source and drain is also drawn in pictorially.

Remember: MOSFETs

There are 4 different types of MOSFETs. On the one hand, these differ in the type of current-forming charge carriers:

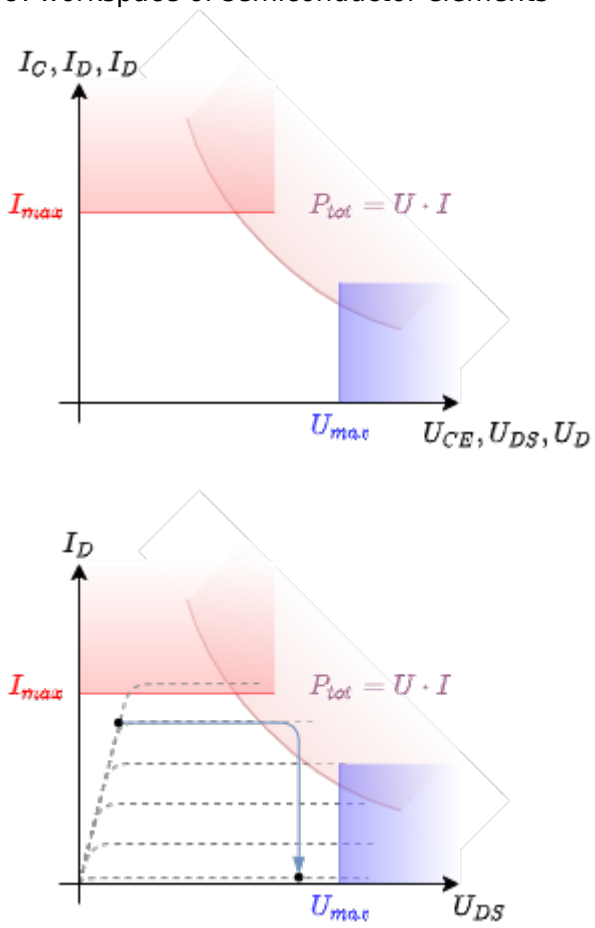
- **n-channel**: The current-forming charge carriers are electrons.
- **p-channel**: The current-forming charge carriers are holes.

The second distinguishing feature is the off-state conductivity ($U_{GS}=0$):

- **enhancement type**: When the gate voltage is $U_{GS}=0$, no conductive channel is present. Only by charging the gate bulk capacitor, the channel is formed or the carriers are enriched.
- **depletion type**: At a gate voltage of $U_{GS}=0$, a conductive channel is present. By charging the gate bulk capacitor, the channel is reduced or the charge carriers are displaced (“depleted”).

In the field-effect transistor, the electric field of the gate-bulk capacitor enriches or depletes only those charge carriers that contribute to charge transport.

Fig. 9: workspace of semiconductor elements



Semiconductor element design

For all transistors and diodes, various limit values must be observed for the circuit design. These can be entered directly in the output characteristic field (figure 9, above). Due to the heating of the component and the resulting increase in intrinsic conduction, two limit values result:

- In the conducting state, the power dissipation $P_{loss}=R(T)\cdot I^2$ forms a direct reference to the current through the semiconductor element I_C, I_D, I_D (bipolar transistor, MOSFET, diode). This results in current I_{max}, which should not be exceeded.
- In the state where there is both a noticeable current and voltage, there is a maximum allowed power $P_{tot}=const.=U\cdot I$. This is a hyperbola in the output characteristic. If the output current exceeds this hyperbola, the semiconductor element heats up to such an extent that, due to the increasing intrinsic conductivity, the conductivity drops, which in turn leads to an increasing current. This effect leads to the thermal destruction of the component.

In addition, a maximum voltage U_{max} must not be exceeded. This is usually due to the (internal) dielectric strength of the component.

These limits are especially important if, for example, a MOSFET is to be used as a switch (example: figure 9, below). In this case, there are two states:

- Switch is conductive: a low voltage U_{DS} is applied, at which a large current $I_D<I_{max}$ flows.
- Switch is non-conductive: A high voltage $U_{DS}<U_{max}$ is applied, at which no current flows.

When switching from “conductive” to “non-conductive”, even if the individual current and voltage limits are taken into account, this can destroy the switch. In figure 9, this case can be seen in the diagram below. Current flow I_D is initially maintained (or are only small), although voltage U_{DS} increases (blue line). In this case, P_{tot} may be exceeded and the MOSFET is destroyed due to thermal overload.

To speed up the switching process (especially for power MOSFETs, e.g. for motor drivers), so-called **driver circuits** generate the voltage U_{GS}. With these driver circuits, the control voltage can be made available and reset very quickly. For this purpose, currents in the range of several amperes must be provided for a short time for charging and discharging the gate capacitor.

Remember: Maximum output values of a semiconductor element

> For each semiconductor element, there are three maximum values to consider at the output:
>
> - a maximum voltage limit U_{max},
> - a maximum current limit I_{max},
> - a maximum power limit $P_{tot}=U \cdot I$

2.8 Applications for bipolar transistors

Darlington-Transistor

The Darlington circuit or the Darlington transistor (as a discrete element) is a simple construction, which makes it possible to control the output voltage U_{BE} with a considerably lower base current I_B. On the right is the Darlington circuit compared to a simple bipolar transistor. Details can be found in Wikipedia under Darlington circuit.

Internal life of an operational amplifier

The operational amplifier as an “almost ideal” differential voltage amplifier represents a central component of electronic circuit technology from the next chapter on. In the chapter basics to amplifiers - feedback an ideal differential voltage amplifier was already used. In the simulation on the right, the core of the differential voltage amplifier is simplified. Accordingly, there is no differential voltage at the input, but a small sinusoidal voltage. This is first applied to the base of the first bipolar transistor, which is a high impedance input amplifier stage. The current I_C regulated by this in turn leads to a base of another bipolar transistor and then to the output amplifier stage. In simulation, this setup achieves a differential gain of about $A_D=10,000,000$. In real differential amplifiers, this is more in the range $A_D ≈100'000$. Details can be found in Wikipedia under operational amplifier.

2.8 Applications for field effect transistors

NOT gate

Just about all consumer electronics products have field-effect transistors at their core. In detail, this is based on CMOS technology (CMOS: Complementary metal-oxide-semiconductor) is used. The MOSFETs on the ground side and the MOSFETs on the power supply side behave in opposite ways, i.e. complementary. The simulation on the right shows the simplest gate, the NOT gate. Another gate was considered in an introductory way.

Reverse polarity protection

Many chips (such as microcontrollers) can be destroyed by an incorrectly polarized power supply. Battery powered electronics should have an active protection circuit for this. A diode is not practical for the power supply (why?). Instead, a MOSFET can be used, which does not pass negative voltages. Details are well explained on the page of Lothar Miller.

Level converter

During electronics development it can happen that several integrated circuits (e.g. intelligent light sensor, microcontroller, intelligent LED) require different voltage levels. This can lead to problems especially during data exchange, if logic High has to be in a certain voltage range. This problem can be solved by a level converter. The level converter (also logic level converter, level shifter) enables the bidirectional connection of digital connections of different voltage levels, e.g. 5 V to 3.3 V.

For the level converter, any n-channel enhancement MOSFET whose threshold voltage is below $1.8...2.0 V$ can be used. This limit is due to the minimum logic level of $2.0 V$ for logic high. For simplicity, “logic level enhancement mode MOSFET” are used, which are just optimized for the logic voltage of $3.3V$.

The way it works is well explained on Wikipedia and can be derived with simulation.

Voltage doubler/inverter

As a power supply for electronics, $5V$ or $3.3V$ is often used. In the following chapter, we will see that a bipolar power supply is often used for operational amplifier circuits. To be able to generate $-5V$ at low currents from a $5V$ supply, charge pumps are often used. One such can be seen on the right in the simulation. In the oscilloscope (in the simulation below), the voltage U_{C1} is displayed at the input capacitor C1 and U_{C2} at the storage capacitor C1. This circuit can be

found, for example, in IC ICL7660 (Renesas), LMC7660 (TI), TC7660 (Microchip) integrated. Details on how it works can be found in this video, for example.

Study Questions:

- In which state is the voltage U_{C1} equal to $1 V$? * In which state is the difference between the voltages $U_{C2}-U_{C1}$ across the two capacitors equal to $1V$?
- What happens if the voltage sources for $0V$ and $1V$ are reversed?
- How can this circuit be implemented with diodes instead of changeover switches?

Voltage inverter in the microcontroller

In some microcontrollers a negative voltage is required internally (e.g. for operational amplifiers). Since this voltage is not supplied externally, the microcontroller must provide it via an internal circuit. The simulation on the right shows a circuit that can be integrated into a microcontroller in this way. The ring oscillator generates a high frequency clock signal, which drives an inverter stage (logical NOT gate). The charge can then be shoveled down via the two capacitors in such a way that the capacitor provides a negative voltage at the output. For more information, see Wikipedia under charge pump and “Inside the 8087's substrate bias circuit”.

four-quadrant

In many applications, current and voltage must be controlled independently of each other. This is the case, for example, with a motor (= ohmic-inductive load). There, the current is essentially proportional to the torque and the voltage to the speed. If voltage and current are to be output bipolar (or in the application: Torque and speed are to be controlled in both directions), a four-quadrant controller made of transistors is suitable. In modern integrated circuits, these are made of MOSFETs, directly equipped with the MOSFET driver, and several four-quadrant controllers can be found next to each other (e.g. the stepper motor driver DRV8835). Details can be found on Wikipedia under four-quadrant actuators.

Other MOSFET Applications

MOSFETs are not only used for pure switching of currents. Further applications are also:

1. as a display element in TFT screens (TFT ... Thin Film Transistor).
2. as memory element e.g. in SD cards (Floating Gate Transistor, or also new approaches, like Ferroelectric_Random_Access_Memory)
3. as an integrated “upstream” element for power bipolar transistors, especially in the Bipolar_transistor_with_insulated_gate_electrode (IGBT)
4. as a chemical sensor for various materials (see Chemical_sensitive_field_effect_transistor)
5. as a link between photonics/optoelectronics and classical electronics

Learning questions

for self-study

- Describe the function of a transistor.
  - Sketch the layered structure of a bipolar transistor. Explain the switching through of a PNP bipolar transistor with the help of the sketch drawn.
  - Draw the simplified diode equivalent circuit of an NPN transistor describe the working.
- Explain the difference between a PNP and NPN transistor.
  - Draw a circuit each with the respective switch connected to U+ = 5V and ground in such a way that switching through is possible with a voltage between U+ and ground at the base.
  - Name the respective connections of the transistors in the drawing.
  - What voltage must be applied to the base in each case for the transistor to switch through?
  - How should the sign of the control current be chosen in each case?
  - In what size range is a typical current gain?
- Current-controlled and voltage-controlled transistors
  - Explain the difference between a current controlled transistor and a voltage controlled transistor.
  - Which type of transistor is current controlled and which is voltage controlled?
  - Draw a circuit diagram each for a current controlled transistor and a voltage controlled transistor.
  - What is the doping order of the transistors drawn?
- What are the two basic types of transistors?
- MOSFET
  - What are the advantages of a MOSFET over a bipolar transistor?
  - How is a MOSFET constructed? (layer structure, connections)
- H-bridge
  - Draw an H-bridge with switches (ideal switch), a resistive/inductive load and an external voltage source with V+ and GND.
  - How can the various switches be controlled to have any voltage between V+ and V- applied to the load? What is the technical term for the method of control?
- Draw the PWM signal necessary to generate a sinusoidal output when a full bridge is used.
- What are the uses for transistors
  - What are some uses for transistors?
  - Draw a voltage doubler.
  - What is a level converter?
  - Why is it preferred to use field effect transistors rather than bipolar transistors nowadays?

with answers

Looking at the picture above, which of the following statement(s) is/are correct?

- ☐ The transistor has an npn structure internally
- ☐ The collector terminal is at the bottom
- ☐ It is a bipolar transistor
- ☐ In order to make I_C flow, the voltage U_BE must become positive

Which statement(s) about bipolar transistors is/are correct?

- ☐ The current I_C or the voltage U_BC controls the current flow I_B.
- ☐ The input characteristic of a bipolar transistor corresponds to that of a diode.
- ☐ The disadvantage of the bipolar transistor is the continuous current flow required in the conductive state.
- ☐ VCC stands for Voltage Common Connector

Which statement(s) about MOSFETs is/are correct?

- ☐ MOSFET stands for the structure of the field-effect transistor made of metal oxide and semiconductor
- ☐ Due to the body diode the MOSFET acts in one direction like a diode
- ☐ Enrichment type MOSFET are conductive with $U_{GS} =0V$
- ☐ In n-channel MOSFETs, holes are the current-carrying charge carriers.

Check answers
You Scored % - /

Image references

References to the media used

Element	License	Link
Video: Circuit Elements - Diodes and Transistors - Part 4	CC-BY (Youtube)	https://www.youtube.com/watch?v=KjyHta5p9WE
figure 4: Function of the npn bipolar transistor	(c) Open Music Lab, with permission for further use	Source: Mail of the illustrator
figure 6: function of the MOSFET	(c) Open Music Lab, with permission to reuse	CrowdSupply

---

[1] In practice, a distinction is still made between small-signal current gain $\beta = h_{fe}$ and large-signal current gain $B = h_{FE}$. In small-signal behavior, a relatively small change around a fixed operating point (e.g., around certain values I_C and U_{CE}) is considered. In large-signal behavior, a change between 0 and a given value is considered. For nonlinear characteristics, the two quantities may differ. In this course, only the small-signal behavior is described. The large-signal behavior and the distinction between the two considerations are not considered in this course
[2], [3] In field-effect transistors, an additional capacitor forms between source and drain, which can lead to overvoltages at the MOSFET, especially during fast switching of inductors