

1

TRANSISTORS

1.1 Introduction

1.2 The MOSFET transistor

- 1.2.1 Inversion condition and threshold voltage*
- 1.2.2 Variable resistor*
- 1.2.3 Channel pinch-off and transistor effect*
- 1.2.4 Current modulation with the Gate voltage*
- 1.2.5 The characteristic curves of enhancement MOSFETs*
- 1.2.6 The circuit symbols*
- 1.2.7 Early voltage of the MOSFET*
- 1.2.8 The MOSFET: a continuously evolving device*

1.3 The bipolar junction transistor - BJT

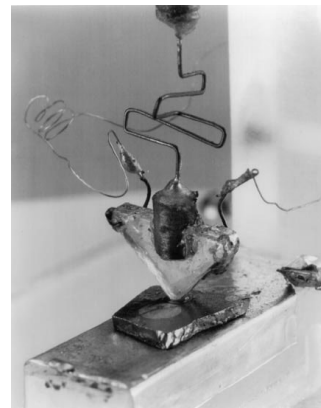
- 1.3.1 The operating principle*
- 1.3.2 The collector current*
- 1.3.3 The base current*
- 1.3.4 The circuit symbols and operation in the active zone*
- 1.3.5 The characteristic curves*
- 1.3.6 The transconductance*
- 1.3.7 Early effect and collector resistance*

1.1 INTRODUCTION

In less than a decade, around the end of the 1940s, research on semiconductor devices made a great leap forward, moving from simple metal oxide rectifier diodes, used for example in radar systems developed in the years before the Second World War, to the invention of semiconductor transistors. Although there was an applicative motivation behind the scientific effort that led to the invention of the new devices (the idea of replacing the mechanical switches of telephone centers with smaller and more reliable controlling devices based on electronic principles), most of the innovative ideas were the result of the enthusiasm and fine speculative skills of a group of researchers from the American telephone company: *Bell Telephone*.

The first experimental confirmation of the possibility of obtaining signal amplification in a semiconductor device was obtained on December 16, 1947, when John Bardeen and Walter H. Brattain made the first transistor work: the transistor with metal-tip contacts.

The device consisted of an n-doped germanium crystal (Base) and two gold contacts very close to each other (Emitter and Collector). By connecting a load to the Collector, the two researchers demonstrated that by varying the current at the Base terminal, it was possible to obtain a proportional variation in the Collector potential. Since the ratio between the variation in Collector potential and the variation in Base current has the dimensions of a resistance, the authors described the device as a “transfer resistor” and coined the term transistor. On December 23, 1947, the results of the experiments were communicated to the management of the laboratories. This is universally considered to be the birth date of the transistor.



But it was another researcher at Bell Laboratories, the physicist William Shockley, who linked his name inextricably to semiconductor devices. In fact, he is responsible for the greatest merits in the study and understanding of the individual physical aspects of the operation of the transistor:

- (i) - the injection of *minority* carriers into pn junctions;
- (ii) - the diffusion transport they undergo in the neutral zone, an idea influenced by his studies with E. Fermi in 1940 on the diffusion of neutrons in graphite;
- (iii) - the band representation of semiconductors and the introduction of the quasi-Fermi level (*imref*);
- (iv) - the extension of the depleted regions in reverse-biased junctions.

The results of his study and a wealth of innovative ideas are summarized in five pages of his laboratory notebook of January 23, 1948. In the paper, all aspects of the operation of the **bipolar junction transistor** are clearly specified. The text formed the basis for the drafting of the patent, presented on June 26, 1948 and accepted on September 25, 1951 (U.S. Patent 2,569,347). Due to the technological difficulties that existed at the time in the growth and doping techniques of semiconductor crystals, it took more than a year before a device with only pn junctions experimentally demonstrated the correctness of the hypothesized operating mechanisms.

The invention of the **unipolar field-effect transistor** is also due to W. Shockley. The idea of using an external electric field perpendicular to the surface of the semiconductor to control the density of carriers in the semiconductor near the surface was in fact the subject of his studies since 1943. However, the difficulty of creating devices in which the idea was not masked by parasitic physical phenomena did not allow real in-depth studies until 1952, when he published the proposal of a junction field-effect transistor, with a conceptually similar structure to that of a modern **JFET**, accompanying it with an in-depth study of its operation. This paper was followed, a few months later, by the experimental confirmation of the described properties.

MOSFET transistors had to wait until the early 1960s before being realized, due to the considerable dependence of their operation on the quality of the crystal used and the characteristics of the oxide-semiconductor interface. The efforts of several research laboratories and numerous scientists from different universities contributed to their realization. W. Shockley, W. H. Brattain and J. Bardeen (pictured aside) were awarded the Nobel Prize in Physics in 1956 for their inventions.



The structure of MOSFETs remained similar to that proposed by the inventors and described in paragraph §1.2, for several decades until new needs emerged that stimulated new solutions, as we will mention in paragraph §1.2.8.

As a seal of the importance of the transistor in contemporary life, one should consider the fact that by 2025 there will be more silicon transistors on planet Earth than human cells.

1.2 The MOSFET transistor

The operation of a MOSFET (*Metal Oxide Semiconductor Field Effect Transistor*) can be illustrated by referring to the schematic structure of an n-channel device, shown in Fig.1.1. The device is fabricated on the surface of a p-doped semiconductor (substrate) and is characterized by two n-contacts (Source -S- and Drain -D-) between which a layer of insulating dielectric (usually silicon oxide, SiO₂) covered by a metal contact called Gate -G- is placed on the crystal. For simplicity, let's assume that there are no electrical charges in the oxide and that the substrate is kept at the same potential as the Source contact, both for example at ground. The two n-contacts of Source and Drain are electrically isolated from each other. In fact, even if a positive potential were applied to electrode D ($V_S=V_{SUB}=V_G=0$ and $V_D>0$) the np junction between Drain and substrate would become inversely polarized and only the very weak reverse current due to the few minority carriers present in the two areas would flow between the two n contacts.

1.2.1 Inversion condition and threshold voltage

The operation of the device is based on the possibility of **varying**, by acting on the Gate potential, **the concentration of free electrons under the oxide** in order to activate a continuous path of electrons between the two n contacts. In fact, if a positive potential is applied to the Gate, a perpendicular electric field is created between the Gate and the substrate (hence the classification of the device as a FET) which pushes the holes away towards the substrate and draws the minority electrons towards the interface between the oxide and the substrate (Fig.1.2). The retreat of the holes determines an area emptied of majority carriers in the region under the oxide which extends further from the surface towards the inside the higher the potential applied to the Gate. If this were the only effect, the isolation between S and D would be even accentuated. But at the same time, the concentration of electrons in a thin layer under the oxide increases, which creates a continuous connection of electrons

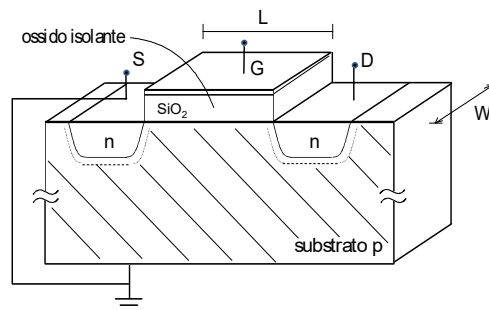


Fig. 1.1 Schematic structure of a prototype n-channel MOSFET.

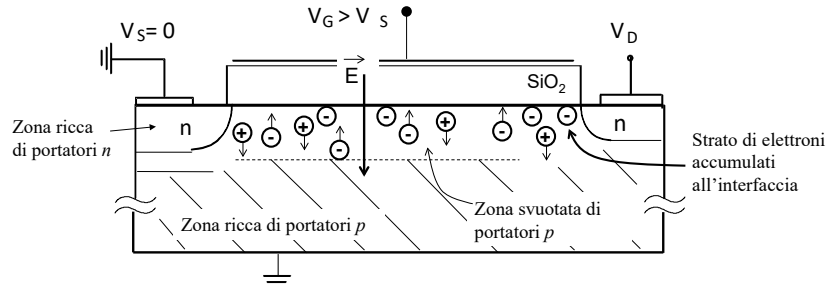


Fig. 1.2 Schematic of the formation of a conductive electron channel at the oxide-semiconductor interface in an nMOSFET when a positive voltage is applied to the Gate with respect to the Source, which pushes holes away and attracts electrons.

(**channel**) between Source and Drain. These two terminals, rich in electrons because they are doped, cease to be electrically isolated and connect to the channel. If a small positive potential V_D were applied to the Drain, current would flow between the two contacts. By further increasing the positive bias of the Gate, the concentration of electrons in the channel increases. It is therefore clear how the Gate potential can act as a control variable for the conductance of the device between the two terminals of Source and Drain.

The conductive channel between Source and Drain is conventionally considered to be formed when the concentration of electrons under the oxide reaches the concentration that the holes had there before the application of the voltage (i.e. the concentration of the majority holes in the neutral substrate). The value of the voltage V_{GS} for which this condition is reached is called the threshold voltage, V_T (**Threshold Voltage**) and when this value is reached, $V_{GS}=V_T$, it is said that the **inversion condition** has been reached: the carriers that guarantee the conductivity of the layer under the oxide are of opposite polarity (in our case electrons) to the majority carriers present there at the time of manufacturing (holes).

The **THRESHOLD VOLTAGE**, V_T , of a MOSFET is usually of the order of Volts (0.3-1V). It depends on the doping of the substrate, the thickness of the oxide and any fixed charges present in it. It can vary up to 20% from one manufacturing batch to another, while between MOSFETs on the same chip the variation is generally much smaller. The value of the V_T can be modified during manufacturing by implanting dopant processes in the semiconductor itself and only under the oxide to locally change the doping.

Note that there is a depleted zone between the conductive electron channel under the oxide and the neutral substrate, ensuring complete electrical isolation between the carriers (electrons) flowing in the channel and the majority carriers (holes) of the substrate.

1.2.2 Variable resistor

The layer of mobile charges in the channel and the gate electrode form the plates of a capacitor, whose dielectric is the oxide (Fig.1.3a). A detailed analysis of the metal-oxide-semiconductor (MOS) system would show that, as the gate voltage increases above the threshold value V_T , the excess voltage ($V_{GS}-V_T$), often called the **overdrive voltage** $V_{ov}=V_{GS}-V_T$, drops almost entirely across the oxide. Therefore, the charge concentration in the channel per unit area grows linearly with the voltage according to the well-known capacitor relationship:

$$Q'_n = C'_{ox} (V_{GS} - V_T) \quad (1.1)$$

where $C'_{ox} = \epsilon_{ox} / t_{ox}$ is the capacitance per unit surface of the gate and t_{ox} the thickness of the oxide. Correspondingly the resistance of the channel between Source and Drain, constituted by the thick electron film S_n would be:

$$R = \rho \cdot \frac{L}{S_n \cdot W} = \frac{1}{\mu_n \cdot q \cdot n} \cdot \frac{L}{S_n \cdot W}$$

where L and W are respectively the length (along the direction of motion of the carriers, between S and D) and the width of the channel (Fig.1.1), μ_n is the mobility

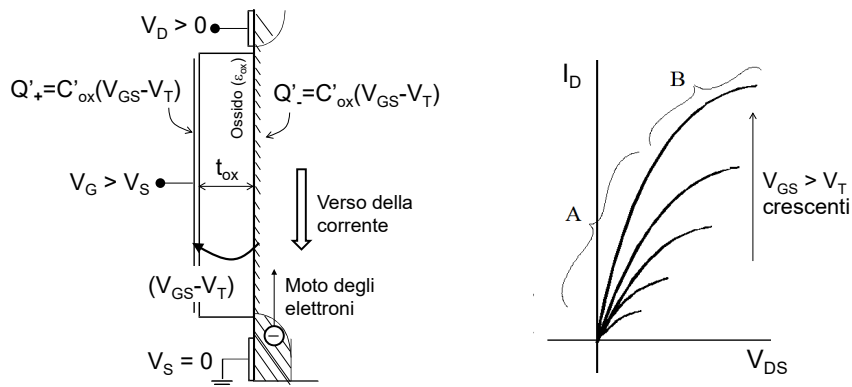


Fig. 1.3 a) Behavior of the MOS, similar to a capacitor with parallel flat faces and b) characteristic curves of an n-channel MOSFET in the ohmic region. The current flows from the highest potential to ground, from top to bottom in the figure.

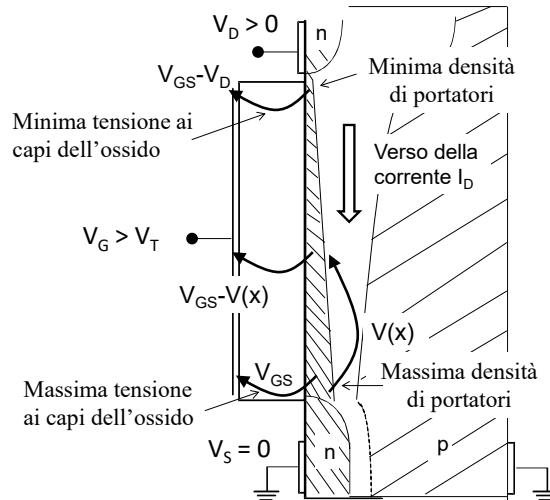


Fig. 1.4 Voltage drops along the channel of a nMOSFET in the ohmic region and the resulting carrier concentration distribution visualized with varying hatching thickness.

of the electrons and n the concentration of electrons in a unit volume. Since $Q_n' = q \cdot n \cdot S_n$ we obtain :

$$R = \frac{1}{\mu_n \cdot Q_n'} \cdot \frac{L}{W} = \frac{1}{\mu_n} \cdot \frac{L}{W} \cdot \frac{1}{C_{ox}} \cdot \frac{1}{V_{GS} - V_T} \quad (1.2)$$

By varying the Gate voltage, according to (1.1), the channel resistance can be increased or decreased. By applying a positive voltage to the Drain, a current flows in the channel that is greater the higher the surface concentration of electrons.

Fig.1.3b shows this: if $V_{GS} > V_T$ and $V_{DS} > 0$, the current in the device varies according to the family of characteristic curves shown. For small voltages V_{DS} (section A of the characteristics) the growth of the current is practically linear because the conductivity of the channel is practically constant along its entire extension between S and D. This operating region is called the **ohmic zone** of the MOSFET precisely because the device behaves like a simple resistor.

However, when the Drain potential is further increased, the further increase in the circulating current follows a less than linear law (section B of the characteristic). To understand this, it is sufficient to concentrate on the voltage across the oxide at various points in the channel (Fig.1.4). In fact, the drop in potential along the channel determined by the passage of the current causes the voltage across the oxide to decrease as one moves towards the Drain: near the Source the voltage across the oxide will be maximum and equal to V_{GS} ; near the Drain the voltage will be minimum and equal to $(V_{GS} - V_T)$. Therefore, in accordance with (1.1), the charge

density in the channel decreases as one proceeds from the Source towards the Drain and reaches its minimum at the end of the channel adjacent to the Drain where the voltage across the oxide is minimum. This decrease means that, by increasing the voltage V_{DS} , the average resistance of the channel increases and therefore the current tends to increase but less than proportionally with V_{DS} , as can be seen in section B of the curves in Fig.1.3b

1.2.3 Channel Pinch-off and transistor effect

By further increasing V_{DS} we arrive at the situation in which the potential drop across the oxide between Gate and Drain becomes equal to or less than V_T and is therefore no longer sufficient to have a conductive channel at the end towards the Drain. This situation, shown in Fig.1.5, is called the **pinch-off condition** of the channel and we indicate with P the point along the channel where this occurs.

The voltage between Drain and Source at which the pinch-off occurs is called the saturation voltage, $V_{Dsat}=V_{GS}-V_T$.

Note that the pinch-off of the conductive channel does not prevent the passage of current. In fact:

- the voltage of point P is $V_P=(V_{GS}-V_T)$ and therefore the voltage across the conductive channel is exactly $(V_{GS}-V_T)$;
- the ratio between the voltage of point P, $(V_{GS}-V_T)$, and the resistance of the channel determines a current I_D that flows between Drain and Source;

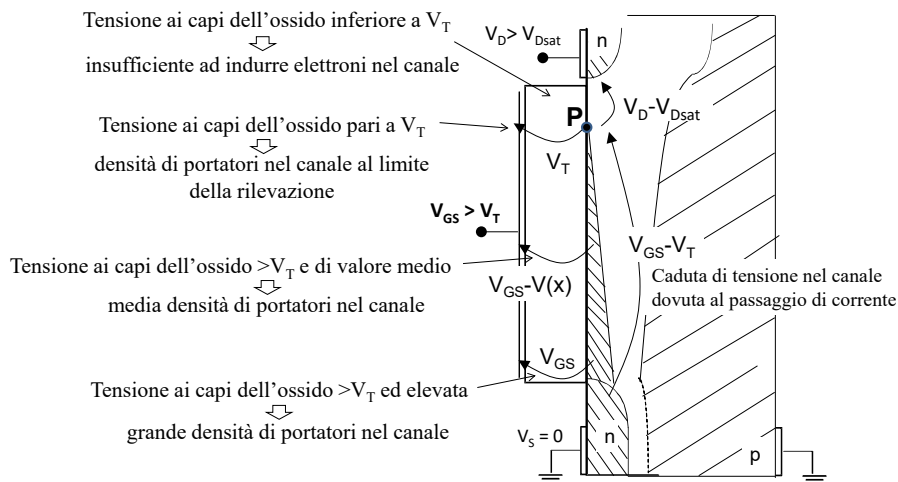
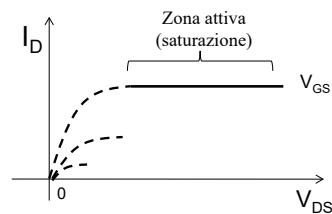


Fig. 1.5 *Distribution of voltage drops across the oxide and corresponding distribution of electron concentration in the channel.*

- the electrons that reach the point P from the Source continue to be accelerated towards the Drain by a potential difference that is still favourable to the continuation of their motion, in a crystal zone where there are no other carriers.

For $V_{DS} > V_{Dsat}$ the situation does not change and there will always be a point P in the channel very close to the Drain at a voltage $V_p = (V_{GS} - V_T)$ (this is the minimum voltage across the oxide to invert the channel), while the excess Drain voltage ($V_{DS} - V_{Dsat}$) falls between the point P and the Drain terminal, possibly causing a local widening of the emptied region and a greater acceleration of the carriers.

The **current** flowing in the transistor therefore **remains almost constant** despite the increase in V_{DS} (the transistor is said to operate in **saturation**) and is always given by the ratio between $V(P) = (V_{GS} - V_T)$ and the channel resistance, both of which are almost constant, as shown in the characteristic curve alongside (and in Fig.1.6).



1.2.4 Modulazione della corrente con la tensione di Gate

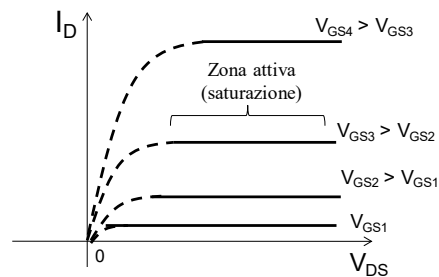
Let us now consider what happens by **increasing the voltage V_{GS}** but leaving V_{DS} at a high value, higher than V_{Dsat} :

1) the charge in the channel increases, decreasing its resistance.

In fact, the increase in voltage at the Gate increases the perpendicular electric field and, according to (1.1), more charge is induced in the channel.

2) increases the voltage across the channel between point P and the Source;

In fact, since with the device in saturation the potential difference between the Gate electrode and point P always remains equal to V_T , each variation of the potential on the Gate is rigidly transmitted as an identical variation of the potential of point P. Ultimately, through an external terminal (the Gate), it is possible to control the potential of an internal point of the device (point P) and consequently the value of the voltage across the channel resistance.



The combination of the two effects determines an increase in I_D proportional to the square of the voltage ($V_{GS} - V_T$), according to the relationship:

$$I_D = \frac{1}{2} \mu_n \cdot C'_{ox} \frac{W}{L} (V_{GS} - V_T)^2 = k(V_{GS} - V_T)^2 \quad (1.3)$$

where the proportionality parameter k has the dimensions $[A/V^2]$. Fig.1.6(b) shows in graphical form the trend of (1.3). This is the relationship to be used to evaluate the current in a MOSFET in saturation as the V_G control varies.

AN EXERCISE IN TECHNOLOGY

Consider using a CMOS technology to create an integrated circuit that allows the definition of geometries with minimum dimensions of 90nm and minimum thicknesses for the gate oxide of 10nm.

Calculate the value of the k coefficient of the nMOSFET and pMOSFET transistors obtainable with this technology and having the minimum possible dimensions.

With reference to Fig.1.1, to calculate the k coefficient of the MOSFETs it is sufficient to remember that:

- the mobility of the carriers at the interface between silicon and silicon oxide is equal to $\mu_n=1000\text{cm}^2/Vs$ and $\mu_p=350\text{cm}^2/Vs$ for electrons and holes respectively;

- the dielectric constant of silicon oxide (SiO₂) is equal to $\epsilon_{ox}=\epsilon_o\epsilon_r$ dove $\epsilon_o=8,8\times 10^{-14}\text{ F/cm}$ is the dielectric constant of vacuum and $\epsilon_r=3.9$ is the relative dielectric constant of silicon oxide. The result is $\epsilon_{ox}=34\times 10^{-14}\text{ F/cm}$.

- the capacitance per unit area, $C'_{ox} = \epsilon_{ox} / t_{ox}$, where t_{ox} is the thickness of the silicon oxide used as a gate dielectric, is therefore equal to 340nF/cm^2 . When doing this calculation, and also the subsequent calculations on the dimensions of the device, be careful to reduce the geometric dimensions to cm.

- the minimum possible dimensions for making the elements of the device (size of an electrode, distance between two electrodes, etc.) with the indicated technology are 90nm, and therefore $W=90\text{nm}$ and $L=90\text{nm}$. Using these data we obtain $k_n=170\mu\text{A}/V^2$ e $k_p=60\mu\text{A}/V^2$.

1.2.5 Characteristic curves of a MOSFET

Fig.1.6(a) shows the **characteristic curves** of an ideal n-channel MOSFET. Note the saturation region (also called the “active region” or “pinch-off region”) in which the current carried by the transistor remains constant despite the voltage between Drain and Source varying. This is the most interesting operating region and is the one in which the MOSFET is biased when used as an amplifier.

Note also that for $V_{GS} \leq 1V$ in this example the transistor is off (no current flows), a value that therefore constitutes its threshold voltage, $V_T \cong 1V$. Since for $V_{GS} = 0$ the transistor does not conduct, it is said to be *normally off*. Only by applying a Gate voltage greater than the threshold voltage does the channel become enriched with mobile carriers, the conductivity of the channel increases and increasingly intense currents flow in the device. This behavior classifies the device as an **n-channel enhancement MOSFET**.

In Fig.1.6b the graph of the relation (1.3) is shown, corresponding to a parabola branch with vertex in $V_{GS} = V_T$, for a given value of V_{DS} . It is called **transcharacteristic curve** because it links the voltage “at the input” of the MOSFET, V_{GS} , to the current I_D that flows “out” of the device (hence “trans-“)

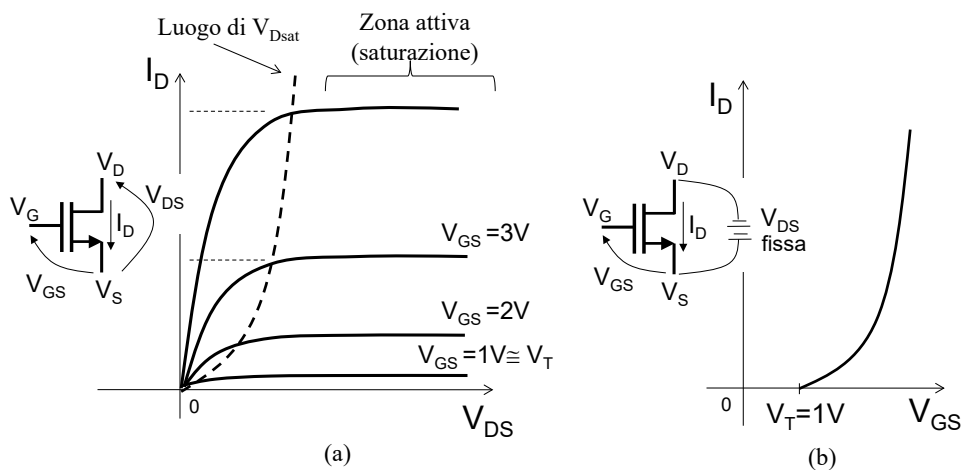


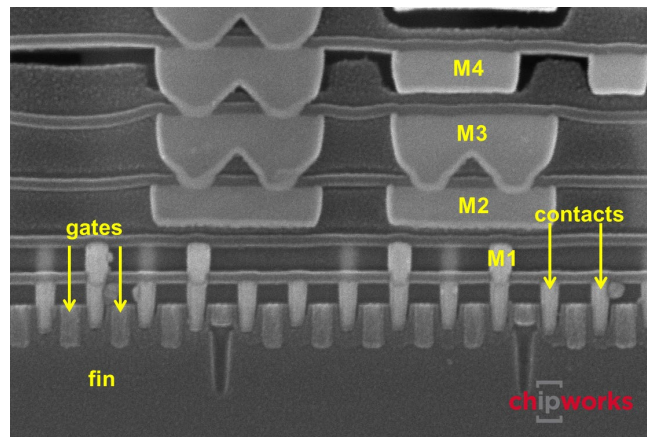
Fig. 1.6 Ideal n-channel enhancement MOSFET with $V_T = 1V$: a) characteristic curves, whose parameter is V_{GS} ; b) transcharacteristic curve (Eq.(1.3)) in the saturation zone for a fixed $V_{DS} > V_{DSsat}$

p-channel MOSFETs, in which the mobile carriers in the channel are holes, are made from an n-substrate on which the Source and Drain contacts are p-type. The threshold voltage is negative because the Gate must be negatively biased with respect to the Source to attract holes under the oxide to form the conductive channel. The characteristic curves of an **enrichment p-channel MOSFET** are given in Exercise E 1.1.

Note that the operation of the MOSFET, like all other existing FETs made with other materials, is linked to the transport of *majority* carriers in a conductive channel. Also note that **the stationary current absorbed by the Gate of a MOS is zero.**

Where is the transistor ?

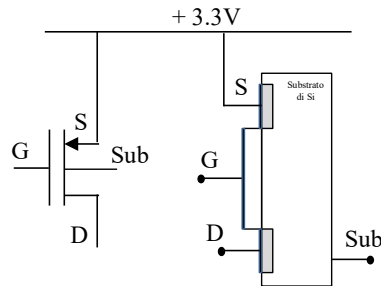
A few years ago, at an international conference, an image similar to the following was shown and the audience was asked “Where is the transistor?”



The electron microscope photo represents the section of a microcontroller chip. It underlines the ever-increasing importance of metal connections (represented in the figure by the metal levels M1,.M4..) needed to connect the many individual transistors present in the chip. You can see some transistors at the bottom under the gates whose small size is negligible compared to the space occupied by the interconnections!

P – MOSFET

Let's think about how a pMOSFET transistor works by answering the following questions:



What type of substrate would you choose to make it, n-doped (n-type) or p-doped (p-type)?

What do you want the Source and Drain contacts to be doped with, n-type or p-type?

What potential should the substrate be connected to? Why?

What do you mean when you say that the transistor has a $V_T=0.5V$?

Which voltage value among these (3.8V; 3.3V; 2.8V) should I apply to the Gate in the example above to start inducing charge in the channel?

In which direction should I move the voltage to the Gate to increase the current in the device, given a small potential difference between S and D, for example 50mV?

At what voltage values can I set the Drain to ensure the operation of the MOSFET in saturation?

At what voltage is the channel pinch-off point when $V_G=2.5V$ (always with MOSFET with $V_T=0.5V$)?

At what voltage is the channel pinch-off point when $V_G=2V$? Is it right that the I_D current increases compared to before? Why?

Why does the current in the MOSFET increase quadratically and not linearly?

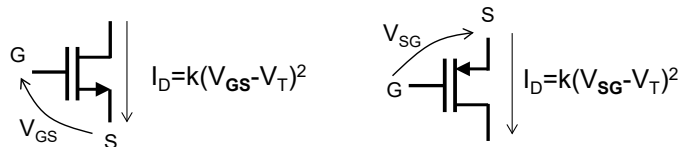
1.2.6 I simboli circuitali

The following figure shows some of the most commonly used circuit symbols for n-channel and p-channel enhancement MOSFETs.



The substrate terminal is often shorted to the Source (to ensure that the Source and Drain junctions never go live with the substrate) and may not be available as an external contact, in which case it may not appear in the device symbol.

By convention borrowed from electrical network theory, many texts assume that the current entering the Drain is positive. Sticking to it, in (1.3) k has a positive sign for nMOSFETs while it has a negative sign for pMOSFETs. In truth, it is much more convenient to think about the real operation of the device, indicating the voltages and currents with their physical direction (V_{SG} in pMOSFETs and V_{GS} in nMOSFETs, I_D from Source to Drain in pMOSFETs and I_D from Drain to Source in nMOSFETs) so as to always deal with positive quantities.



Even k and V_T of pMOSFETs or nMOSFETs are more convenient to always consider positive values related to the physics of the operating mechanisms of the MOSFETs. In this way, the value of k will determine the effective value of the current I_D in the physical direction in which it actually flows in the transistor. As for V_T , it will be enough to take the correct physical direction of the control voltages to always have the correct value of the overdrive voltage, $V_{od}=(V_{GS}-V_T)$ in nMOSFETs and $V_{od}=(V_{SG}-V_T)$ in pMOSFETs, always using positive V_T .

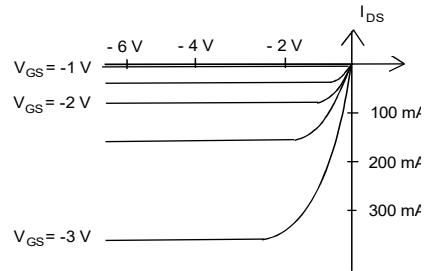
(1.3) would become for the pMOSFET:

$$I_D = \frac{1}{2} \mu_p \cdot C'_{ox} \frac{W}{L} (V_{SG} - V_T)^2 = k(V_{SG} - V_T)^2 \quad (1.4)$$

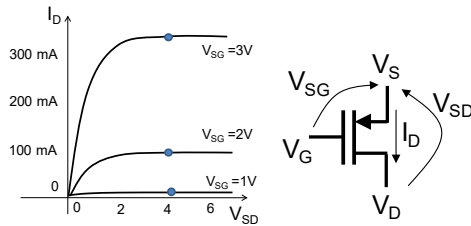
respecting the actual physical operation as shown in the figure above.

In order for the devices to operate in saturation, and therefore (1.3) and (1.4) are valid, it is always necessary to ensure that the voltage between G and D determines the pinch-off of the channel towards the Drain.

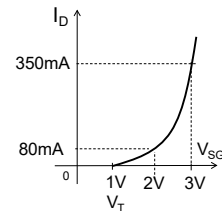
E 1.1 Starting from the characteristic curves of the enriched p-MOSFET in the figure, available on the data sheet of the component supplied by the manufacturer, trace the transcharacteristic curve and obtain the value of V_T



It is convenient to analyze the behavior of the transistor by thinking of the voltages and currents as positive in the direction in which they are actually present in the transistor, and then redraw the characteristic curves as :



By placing yourself at a defined V_{SD} , for example $V_{SD}=4V$, you read the I_D current values at the various V_{SG} . These constitute the values with which to draw the $I_D - V_{SG}$ graph (transcharacteristic). The result is a V_T close to 1V.



Symmetry of MOSFET structure

The structure of the ideal MOSFET (Fig.1.1) is symmetric, in the sense that the Source and the Drain can be interchanged. This aspect is sometimes emphasized by the use of a circuit symbol without the arrow on the Source.



It will then be the connection in the circuit that will define which contact actually acts as Source and which as Drain. Not always easy to understand in a complex circuit!

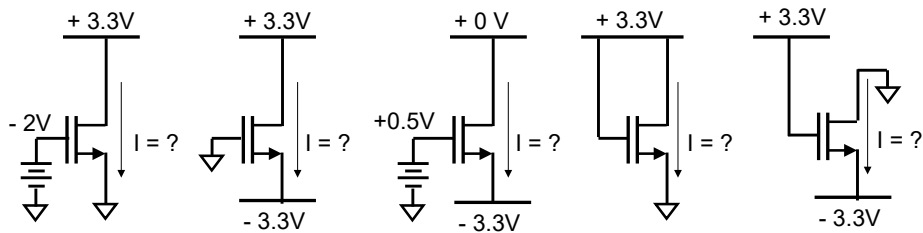
A MOSFET may not be symmetrical if the distances between the gate and the two contacts of S and D are different or even in transistors with circular symmetry, with for example the S at the center surrounded by a Drain ring.

FlashTest on MOSFET : Ohmic, saturated, interdiction

What voltages can there be across the three terminals of a MOSFET for it to function properly?

The following quizzes should help you become familiar with these fundamental aspects of using a MOSFET in an electronic circuit.

ES.1 Suppose MOSFET have $V_T=0.8V$ and $k=300\mu A/V^2$:



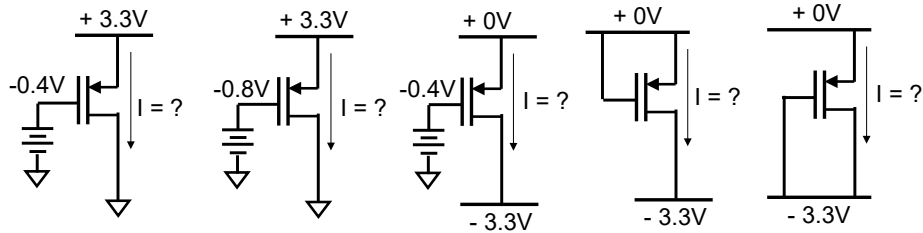
Is MOSFET in saturation (that is in the correct zona di funzionamento) ?

- si no si no si no si no si no

If yes, how much current is flowing ?

I = I = I = I = I =

ES.2 Si supponga che i MOSFET abbiano $V_T=0.7V$ e $k=100\mu A/V^2$:



Is MOSFET in saturation (that is in the correct zona di funzionamento) ?

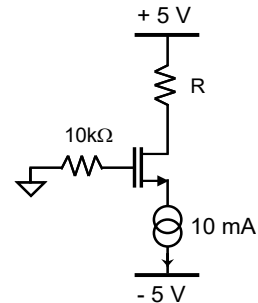
- si no si no si no si no si no

If yes, how much current is flowing ?

I = I = I = I = I =

E 1.2

- (a) Calculate the voltage at the Source of the MOSFET, having $V_T=1.5V$ and $k=10mA/V^2$.
 (b) Indicate which of the following values of resistances connected to the Drain would make the MOSFET exit the saturation zone: $R=80\Omega$, 500Ω , 640Ω , 720Ω
 (c) Calculate by how much the Drain current of the MOSFET would vary if k were $20mA/V^2$.



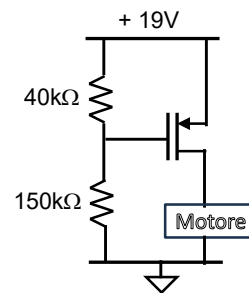
a- Since no current flows in the Gate, there is no voltage drop across the $10k\Omega$ resistor and so $V_G=0V$. The current source forces a current of $10mA$ into the MOSFET and so the Source must drop to $-2.5V$ to let it flow.

b- In order for the MOSFET to operate in the saturation zone, the channel must be pinched off on the Drain side. Therefore the voltage at the Drain must be $V_D > -1.5V$. The only value of R that must be discarded is therefore 720Ω which would determine $V_D = -2.2V$ and would create a conductive channel throughout the transistor, causing it to operate in the ohmic zone. The other values of R instead ensure saturation.

c- The current in the Drain does not vary because it is set by the current generator. By varying k , only the voltage to which the Source is brought to pass will vary to let the $10mA$ set by the current generator pass.

E 1.3

- a) Quale valore di $k = \frac{1}{2}\mu_p \cdot C'_{ox} \frac{W}{L}$ deve avere il MOSFET nel circuito accanto affinché la corrente inviata al motore elettrico sia di $1A$? Si supponga di usare una tecnologia che realizzi MOSFET con $V_T=2V$.
 b) What is the maximum resistance that the motor can show to avoid the MOSFET going out of saturation?



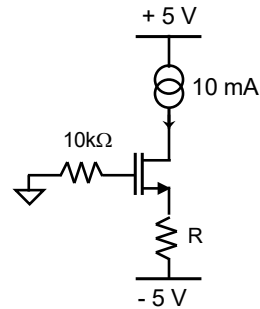
- a) $K=250mA/V^2$.
 b) $R_{motore}=17\Omega$.

E 1.4 Consider the circuit in the figure, whose MOSFET has (same as E1.2) $V_T=1.5V$ e $k=10mA/V^2$. Discuss the following situations:

a) If $R=37\Omega$, MOSFET is in saturazione or not? Find the voltages at Drain, Source and Gate?

b) If $R=1200\Omega$, MOSFET is in saturazione or not? Find the voltages at Drain, Source and Gate?

Speculate on the physical situation in the device and comment on any critical issues found.

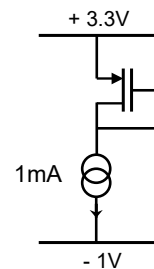


The operating situation of this circuit is completely different from that of the circuit in E1.2.

c) With $R=37\Omega$ the current in the MOSFET, if it were in saturation, would be about 40mA, much greater than 10mA. However, since the current balance at the Drain node must be satisfied, the MOSFET is forced to move into the ohmic zone until it carries exactly 10mA.

d) With $R=1200\Omega$, the situation is different: the current in the MOSFET if it were in saturation would be 2.5mA, less than 10mA, and therefore the balance at the Drain tends to raise its potential up to +5V in an attempt to make the current generator malfunction and make it carry exactly the current of 2.5mA defined by the MOSFET. Note that the MOSFET is in saturation.

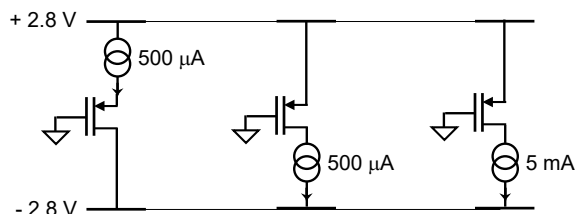
E 1.5 Verify that in this circuit it is possible to force current into the MOSFET Drain and calculate what voltage the Drain is at. ($V_T=0.6V$, $k=4mA/V^2$).



$$V_D=V_G=+2.2V$$

Can you FORCE CURRENT into a MOSFET?

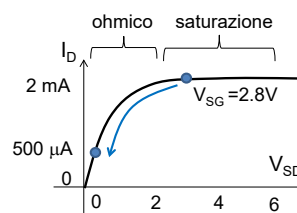
How do the following circuits differ, whose MOSFET have $V_T=0.8V$ and $k=500 \mu A$? Try first before reading the answer.



Let us always remember that in *saturation* the current I_D in the MOSFET is defined by its control voltage V_{SG} according to the relation $I_D = k(V_{SG} - V_T)^2$

In the circuit on the left, V_{SG} is not imposed. The current I , if it were unable to pass through the transistor, would accumulate positive charge at the Source, causing its voltage value to rise. This causes V_{SG} to increase, activating current I_D in the MOSFET. This phenomenon reaches perfect equilibrium when at the Source node $I_D=I$, that is, when V_S has reached exactly $V_S=1.8V$. A posteriori, it is checked that the MOSFET is correctly in saturation, a situation ensured here by the $-2.8V$ power supply. We conclude by saying that **it is possible to force current into the Source of a MOSFET**.

In the circuit in the center, V_{SG} is set and is $V_{SG}=2.8V$. So, if in saturation, the current will be fixed at $I_D=2mA$. But in this way the current balance at the Drain is not satisfied! The additional current from the MOSFET raises the potential of the Drain. However, this causes the MOSFET to come out of saturation and bring it into ohmic. The Drain rises in voltage until the MOSFET in ohmic produces exactly the same current called by the generator, $500\mu A$.



In the circuit on the right the situation is similar to the previous one but now the balance of the currents at the Drain node causes its voltage to drop, decreasing the voltage at the ends of the generator. In this case we must think that the current generator is real and that when the voltage at its ends becomes very small, the generated current can no longer be the nominal one of $5mA$ but decreases, in this example right down to $I=2mA$. We will see in the next chapters that current generators made with transistors allow precisely these adjustments. We conclude by saying that **it is NOT without consequences to force current into the Drain of a MOSFET**.

How to choose the size of a MOSFET

The current carried by the transistor in saturation follows, as seen in (1.3) or (1.4), the relation:

$$I_D = \frac{1}{2} \mu_n \cdot C'_{ox} \frac{W}{L} (V_{GS} - V_T)^2$$

Depending on the application, the current required in a circuit can vary from a few nA (in electronic circuits in cell phones, for example, and in most digital support circuits where the aim is to consume as little electrical power as possible) up to several hundred Amps (in the motor control of electric cars, for example, where, on the contrary, it is essential to manage enormous quantities of energy).

How can we satisfy such different needs?

Certainly not by always using the same transistor and changing the control voltage between Gate and Source. It is usually limited for practical and technological reasons to a few Volts beyond which the fine Gate oxide would be damaged by an excessive electric field.

What is done is to design MOSFETs with the appropriate **W dimension**. This is the actual size that the designer chooses to define the order of magnitude of the current carried by the transistor: low-power circuits with small bias currents use MOSFETs with a very small W, of a few tens of nm; high-power circuits and very large bias currents use large transistors, with W reaching centimeters. All operating at V_{GS} voltages of a few Volts!

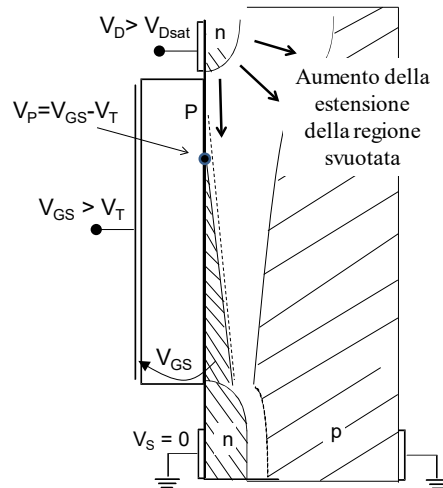
The **L dimension**, that is the length of the channel between S and D, is generally chosen small, often of the minimum size allowed by technology, so as to have the shortest possible transit time, that is the highest possible response speed, and the smallest possible capacities, always to have the highest possible bandwidth. We will see these aspects in the next chapters. The dimension L is chosen large only in particular cases, for example when you want to have a high Early voltage (see the next paragraph) or have a large Drain resistance.

1.2.7 Early voltage of a MOSFET

The ideal situation summarized by (1.3) and represented by the characteristic curves of Fig.1.6 does not find a precise correspondence in the characteristics obtained experimentally on a real device. Normally, in fact, **the real current supplied in the saturation zone is not constant but tends to increase slightly as the voltage V_{DS} increases** (Fig.1.7).

This effect is due to the fact that increasing the V_{DS} voltage increases the reverse voltage of the Drain-Substrate junction and consequently increases the extension of the depleted area at the Drain. It "pushes" the pinch-off point P slightly towards the Source (a phenomenon called "channel length modulation").

As a result, while keeping the V_S and V_G potentials fixed, a slight shortening of the conductive channel decreases its resistance and, while the voltage drop between the Source and the pinch-off point P remains unchanged, increases the current in the channel.



This effect is undesirable, precisely because it makes the current supplied to the load somewhat dependent on the load itself (through the voltage across its terminals) and not only on the control voltage V_{GS}. Assuming for simplicity that the new

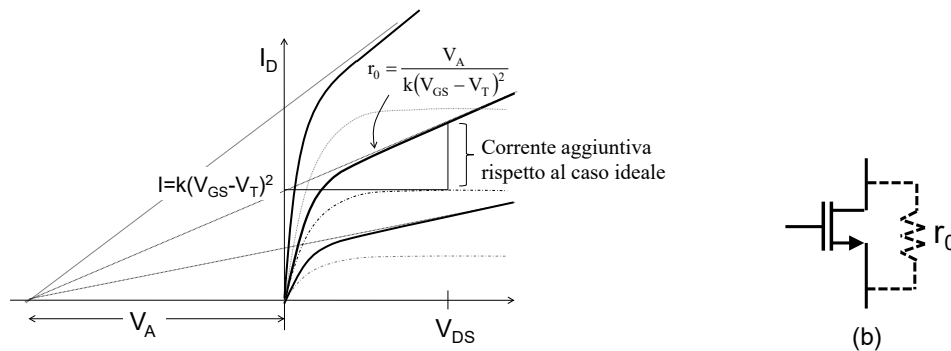


Fig. 1.7 Real characteristic curves of a MOSFET and calculation of its equivalent resistance between Drain and Source; (b) corresponding visualization in the circuit symbol.

characteristic curves (Fig.1.7) are still straight lines in the saturation zone, there is practical evidence that the curves all converge with good approximation to a single point along the negative V_{DS} axis. We call this voltage "Early voltage", V_A . The inverse of the slope of the characteristic curves in the saturation zone has the dimensions of a resistance. It can be considered as a "hidden" resistance inside the transistor that provides an additional contribution of current proportional to the voltage across its terminals, V_{DS} , and for this reason it has been indicated with r_o in the circuit symbol of Fig.1.7b. Its value can be easily estimated by considering the triangle in the second quadrant:

$$r_o = \frac{V_A}{k(V_{GS} - V_T)^2} \quad (1.5)$$

Note that while V_A is a data that characterizes the transistor, and therefore is provided by the manufacturer, the value of r_o varies with the control voltage V_{GS} because the slopes of the curves vary: the more current is circulated in the MOSFET (larger V_{GS}) and the smaller the resistance r_o is. The characteristic curves in fact fan out, as you can see in Fig.1.7.

The fraction of additional current circulating in the transistor compared to the ideal case (horizontal curves) can be easily calculated by quantifying the triangle highlighted in Fig.1.7, so as to obtain the real current circulating in a MOSFET as:

$$I_D \cong k(V_{GS} - V_T)^2 + \frac{k(V_{GS} - V_T)^2}{V_A} \cdot V_{DS} = k(V_{GS} - V_T)^2 \cdot \left(1 + \frac{V_{DS}}{V_A}\right) \quad (1.6)$$

This expression for the Drain current is the one actually used in the **SPICE circuit simulator** for the first level calculation of the current flowing in a MOSFET, where the parameter $LAMBDA=1/V_A$ is required.

1.2.8 The MOSFET: a constantly evolving device

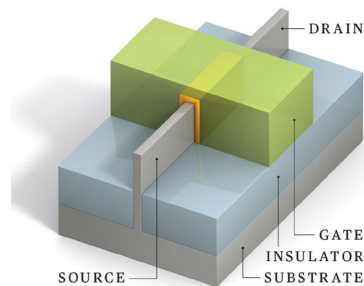
MOS transistors have evolved a lot over the decades, driven by the need to reduce their size in order to speed up their frequency response thanks to a smaller distance between S and D and to be able to have a greater number of them on the same area, thus reducing costs for the same functionality. From the 60s to today, progress in chip technology has been extraordinary and in line with Moore's Law, the observation that the number of transistors on a chip doubles approximately every two years.

But already in the 90s it was understood that we would soon reach the end of this path of size reduction, because we would have reached the size of the same few atoms that would have made up the device. Since then, many ideas have been explored and continue to be proposed to continue to improve the performance of transistors without necessarily reducing their linear dimensions.

An important contribution was made by Hu Chenming, a long-time professor at the University of California, Berkeley (USA). Hu was struggling with the increasing difficulty in turning off transistors as they were scaled up in size and powered with ever lower voltages, in those years now close to 1V. These increasing losses due to currents flowing even when the transistor was off made digital circuits less and less energy efficient. To improve the effectiveness of the Gate in opening and especially closing the channel, Hu proposed creating a Gate that wrapped around a three-dimensional channel on at least three sides so as to control it more effectively.



To do this he had to conceive a construction technology that went beyond the classic planar construction and instead became three-dimensional. The transistor he conceived and created was called FINFET, to recall the shark's fin in its name, which was well evoked by the structure of the device, as evident in the figure alongside. The first industrial product that exploited this technology came out in 2011 thanks to Intel. Today's i7 processor contains billions of FINFETs as small as 14nm.

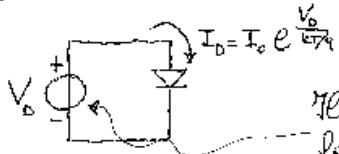


1.3 IL TRANSISTORE BIPOLARE A GIUNZIONE - BJT

TRANSISTORE BIPOLARE

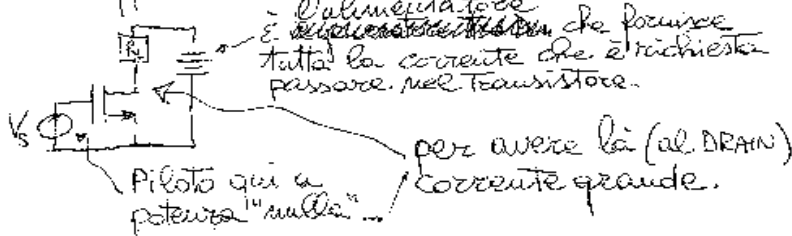
1

① Limite di un bipolo:



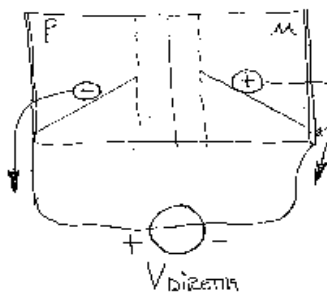
Il generatore di "segnale" deve fornire tutta la potenza utilizzata dal diodo.

Vantaggio di un transistor:



- l'alimentazione è ~~separata~~ che fornisce tutta la corrente che è richiesta passare nel transistor.

Partendo dal diodo:

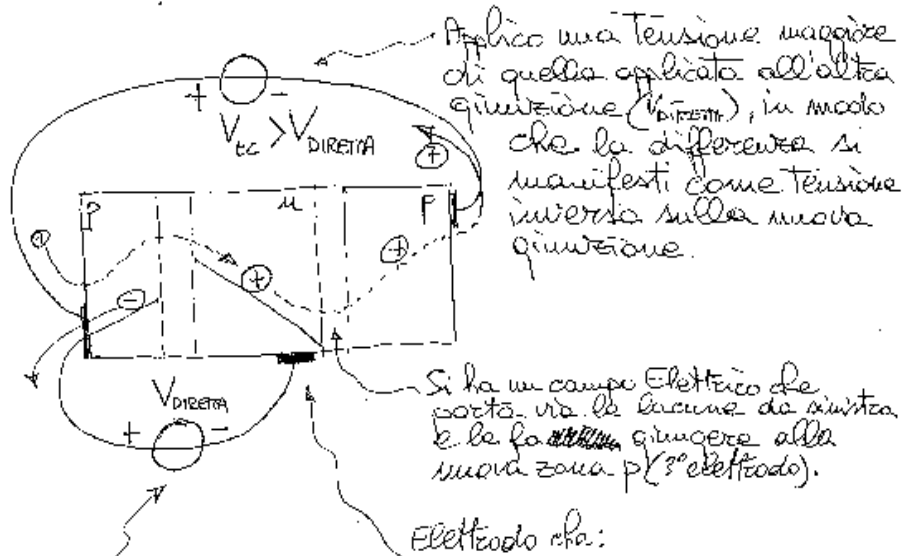


Elettrodo che:

- 1 - definisce il potenziale della zona neutra (che è conduttiva)
- 2 - raccoglie le cariche che stanno diffondendo perché si trova lungo il loro percorso.

Vorrei che i morsetti che modulano la tensione tra p-n NON siano anche quelli che forniscono la corrente. \Rightarrow DEVO TROVARE i POTENZIALI ALTRONVE

② Uso un'altra giunzione, questa volta polarizzata in ~~la~~ inversa, per raccogliere la carica:

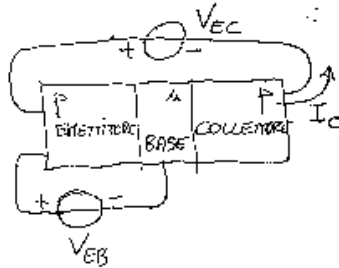


Con questo generatore controllato, come per il diodo, la tensione ai capi della giunzione. Quindi continuo a controllare la quantità di elettroni e di lacune iniettate nelle corrispondenti zone neutre..

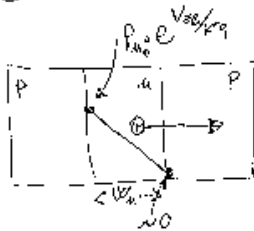
- 1- definire il potenziale di tutta la zona n, grazie all'alto drogaggio n che la rende simile ad un conduttore.
- 2- non raccoglie le lacune perché non è posto lungo il loro percorso di moto per diffusione.

- LE LACUNE FANNO TUTTO IL PERCORSO LUNGO IL NUOVO DISPOSITIVO, però...
- LE LACUNE NON LE FORNISCO CON IL GENERATORE $V_{DIRETTA}$, però...
- IL GENERATORE $V_{DIRETTA}$ TU CONTROLLA IL FLUSSO DI LACUNE.

③ Un po' di nomenclatura:



④ Calcolo la corrente di lacune (I_C):



Come per il diodo:

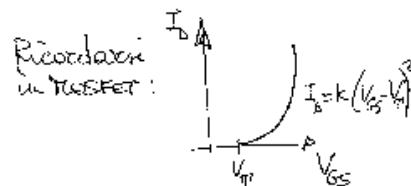
$$I_{\oplus} = q \cdot D_p \cdot \frac{p_{n0} e^{V_{EB}/kT} - n_0}{W_b} \cdot \text{Area}$$

La corrente è unicamente data da questa espressione. Le zone di Emittitore e di COLLETTOR non impongono alcun flusso di lacune, ma semplicemente fanno passare quella dettata dalla DIFFUSIONE NELLA BASE.

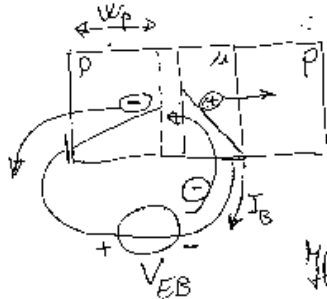
limitato dal costruttore

$$I_C = q D_p \frac{p_{n0}}{W_b} \cdot \text{Area} \cdot e^{\frac{V_{EB}}{kT}}$$

Praticamente indipendente da V_{EC}



⑤ Come lo mettiamo con gli elettroni? 4



Gli elettroni iniettati dalla base nell'emettitore circolano nella maglia del generatore V_{EB} , alla comando! Peccato!

Il comando non è a potenza nulla, come nel MOSFET

$$(V_{GS} \cdot I_{GATE} = 0)$$

N.B. il moto degli elettroni non interferisce in alcun modo con quello della lacuna. primo descritto.

⑥ Calcolo la corrente di base (I_B)

$$I_B = q D_n \frac{\mu_{p0} e^{-V_{EB}/kT} - \mu_{p0}}{W_p} \cdot Area = q D_n \frac{\mu_{p0} \cdot Area \cdot e^{-V_{EB}/kT}}{W_p}$$

Questa corrente è inevitabile, come visto, e deve essere ridotta al minimo con particolari attenzioni costruttive (W_p lunga, μ_{p0} piccola e altro).

⑦ Importante il rapporto:

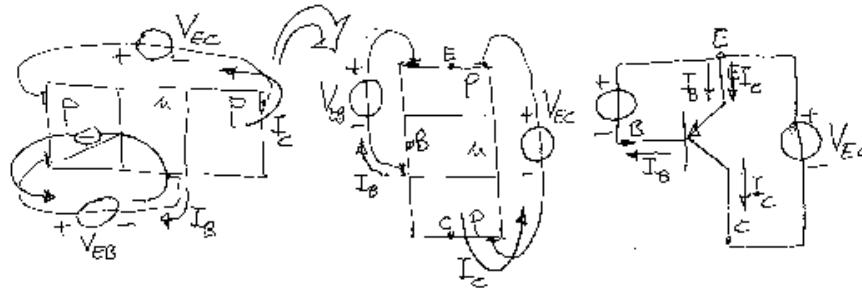
$$\beta = \frac{I_C}{I_B} = \frac{D_p \mu_{n0} W_p}{D_n \mu_{p0} W_n}$$

che dipende solo da termini legati alla costruzione.

Benché le singole correnti (I_C, I_B) siano definite da termini (μ, W) molto variabili da esemplare a esemplare (anche di parecchi ordini di grandezza), il loro rapporto è molto meglio controllato (vari di poche decine di % tra esemplari dello stesso lotto).

8) Simbolo circuitale e modo d'uso

5



9) Verificare sempre che:

- giunzione EMETTITORE-BASE sia in DIRETTA
- " " BASE-COLLEITORE sia in INVERSA.

1.3.1 Il principio di funzionamento

Il transistor bipolare a giunzione (BJT - *Bipolar Junction Transistor*) è un dispositivo costituito da tre zone di semiconduttore contigue, chiamate **Elettore**, **Base** e **Collettore**, in cui quella centrale è drogata di segno opposto alle altre due in modo da formare due giunzioni *pn*, una tra Elettore e Base e l'altra tra Base e Collettore. Ognuna delle 3 zone di semiconduttore è collegata ad un proprio elettrodo metallico in modo da poter applicare localmente la tensione e raccogliere la corrente. Nel funzionamento normale del BJT, la giunzione Elettore-Base è polarizzata direttamente e quella Base-Collettore è polarizzata inversamente.

Il principio di funzionamento di un transistor *npn* è illustrato nella Fig.1.8 in cui è riportata la sezione di un dispositivo realizzato con tecnologia planare. Gli elettroni, iniettati nella Base dalla giunzione Elettore-Base polarizzata direttamente, diffondono verso la giunzione di Collettore (dall'alto verso il basso

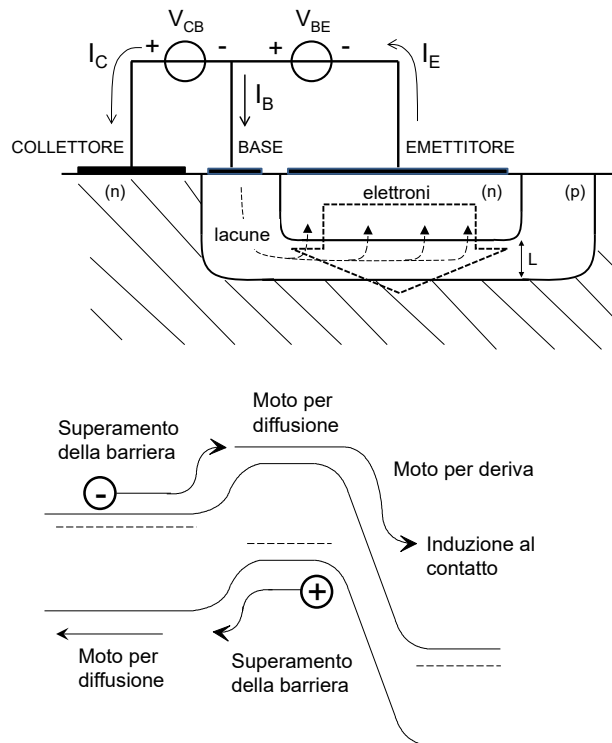


Fig. 1.8 Percorsi dei singoli portatori (elettroni e lacune) sia nella sezione che nel diagramma a bande di energia potenziale di un transistor *npn* prototipo.

nella sezione). Se lo spessore della Base neutra, L , è molto più piccolo della lunghezza di diffusione degli elettroni, la quasi totalità degli elettroni iniettati raggiunge la giunzione Base-Collettore. Essendo questa giunzione polarizzata inversamente, il campo elettrico nella sua regione svuotata è tale da continuare a favorire il moto degli elettroni e trascinarli effettivamente nel Collettore. Nella sezione della Fig.1.8 il Collettore è costituito da tutta la regione tratteggiata: essendo essa drogata, può essere considerata tutta conduttiva e quindi gli elettroni raccolti al Collettore sono immediatamente disponibili al contatto di Collettore in superficie. In un buon transistor bipolare, quindi, il flusso di elettroni raccolti al Collettore (che definisce la corrente I_C) è praticamente uguale al flusso di elettroni partiti dall'Emettitore.

Questo movimento degli elettroni dall'Emettitore al Collettore può anche essere illustrato facendo riferimento ai diagrammi energetici lungo le giunzioni (Fig.1.8 in basso). La polarizzazione diretta, V_{BE} , della giunzione Base-Emettitore causa un corrispondente abbassamento della barriera energetica vista dagli elettroni che si trovano nell'Emettitore. Essi, quindi, sono iniettati nella regione di Base. Essendo questa una zona drogata, essa è conduttiva e pertanto non vi è lungo di essa una caduta di potenziale significativa, come messo in luce dalla costanza del livello di energia nel grafico. Essendo il campo elettrico quasi inesistente, nella Base i portatori iniettati dall'Emettitore si muovono per diffusione, stimolati cioè dal gradiente di concentrazione lungo l'asse Emettitore-Collettore (tanti elettroni nella parte di Base adiacente all'Emettitore e pochi elettroni nella parte di Base adiacente al Collettore). Alla giunzione Base-Collettore infatti, come già accennato, gli elettroni sentono la ulteriore diminuzione dell'energia potenziale, dovuta alla polarizzazione inversa, V_{CB} , che li trascina via dalla Base dentro nel Collettore, dove diventano maggioritari così da indurre immediatamente il segnale nel filo esterno di contatto del Collettore.

La diminuzione della barriera di energia tra Emettitore e Base consente contemporaneamente anche alle lacune maggioritarie della Base di essere iniettate dalla Base nell'Emettitore. Quando in Emettitore, le lacune si trovano ad essere minoritarie in una zona altamente drogata (e quindi senza presenza di campo elettrico come è effettivamente l'Emettitore) e si muoveranno per diffusione (al lato verso la Base le lacune sono tante, al lato verso il contatto ohmico esterno la densità è prossima a zero) attraverso l'Emettitore fino a concludere il loro moto al contatto di Emettitore, inducendo un ugual segnale sul filo del contatto. Ad ogni lacuna iniettata dalla Base nell'Emettitore corrisponde una lacuna richiamata dal filo di contatto nella Base per conservare la neutralità elettrica della regione di Base, dando luogo a quella che chiameremo **corrente di Base**, I_B . Poiché normalmente l'Emettitore è molto più drogato della Base, il flusso di portatori (I_B) iniettati dalla Base verso l'Emettitore è molto minore del flusso (I_C) di elettroni iniettati dall'Emettitore verso la Base che poi raggiungono il Collettore.

Tra le correnti ai morsetti del dispositivo vale quindi la seguente relazione di bilancio:

$$I_E = I_C + I_B \quad (1.7)$$

Per quanto visto, il rapporto I_C/I_E ha un valore molto vicino all'unità ed è indicato con il simbolo α

$$\alpha = \frac{I_C}{I_E} \quad (1.8)$$

Esso rende conto del trasporto principale di portatori lungo la direttrice Emettitore-Collettore del transistor.

Il rapporto tra la corrente di Collettore e la corrente di Base rende conto invece del guadagno di corrente del transistor bipolare ed è indicato con il simbolo h_{FE} oppure β :

$$\beta = \frac{I_C}{I_B} \quad (1.9)$$

In commercio si trovano transistori con valori di β compresi tra 10 (transistori di potenza) e 5000 (transistori di segnale a superbeta). Normalmente sono impiegati transistori con $\beta=100 \div 400$. Da ultimo, in base alla (1.8) e ricordando la definizione di α è facile ricavare che $\beta = \alpha/(1-\alpha)$ e $\alpha = \beta/(\beta+1)$. Se $\beta=100$ si ha $\alpha=0.99$.

1.3.2 La corrente di Collettore

Per ricavare le relazioni funzionali tra le tensioni applicate ai morsetti del transistor e le correnti circolanti si faccia riferimento alla Fig.1.9 in cui sono riportati i profili di concentrazione degli elettroni nella Base e delle lacune nell'Emettitore di un transistor *npn* quando la giunzione Emettitore-Base è polarizzata in diretta e la giunzione Base-Collettore è polarizzata in inversa, come deve essere nella condizione normale di funzionamento. Trascurando la eventuale (e comunque piccolissima) ricombinazione degli elettroni con le lacune maggioritarie della Base durante il loro moto da Emettitore a Collettore, il profilo della densità di elettroni è rettilineo lungo la Base con:

i) un valore di concentrazione al bordo della giunzione con l'Emettitore pari a:

$$n(0) = \frac{n_i^2}{N_A} e^{V_{BE}/V_{th}} \quad (1.10)$$

dove n_i è la concentrazione intrinseca e N_A è la concentrazione dei droganti accettori nella Base;

ii) un valore di concentrazione all'estremità della Base verso il Collettore quasi nulla, perché gli elettroni vengono trascinati via verso il Collettore dal campo elettrico della giunzione polarizzata inversamente.

La corrente di Collettore è pari alla corrente di diffusione degli elettroni nella Base e quindi è proporzionale alla pendenza del loro profilo $n(x)$ secondo la relazione:

$$I_C = qAD_n \frac{dn(x)}{dx}$$

dove A è la sezione del dispositivo e D_n è il coefficiente di diffusione degli elettroni. Sostituendo si ottiene:

$$I_C = qAD_n \frac{n_i^2}{W_B N_A} e^{V_{BE}/V_T} = I_S e^{V_{BE}/V_T} \quad (1.11)$$

dove W_B è la lunghezza della zona neutra della Base. La relazione (1.11) ci dice che al variare della tensione diretta V_{BE} del transistor, la corrente di Collettore varia seguendo la legge esponenziale tipica del diodo Emettore-Base. Per come abbiamo ricavato la (1.11), essa ci dice anche un'altra cosa fondamentale e cioè che la corrente I_C è sostanzialmente indipendente dal valore della tensione del Collettore: infatti, al variare della tensione V_{BC} inversa varia l'entità del campo elettrico nella regione svuotata tra Base e Collettore ma non modifica la capacità di raccolta degli elettroni da parte del Collettore stesso.

Si noti come la corrente I_S , detta **corrente di saturazione inversa**, sia proporzionale all'area del transistor. Quindi, transistori realizzati con gli stessi processi tecnologici e polarizzati a pari V_{BE} , conducono correnti proporzionali alle relative aree di Emettore.

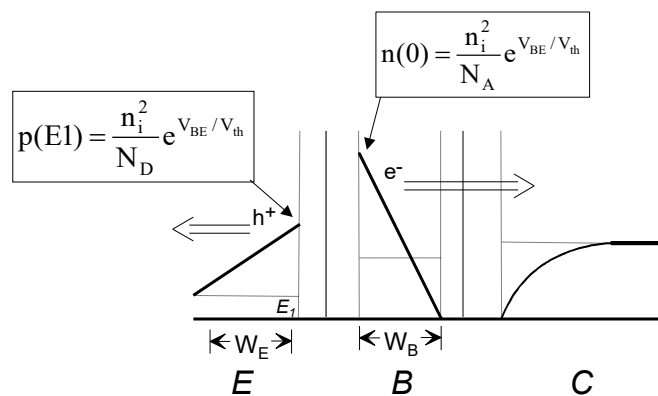


Fig. 1.9 Distribuzione della concentrazione dei portatori minoritari in un transistor npn polarizzato in zona attiva diretta.

1.3.3 La corrente di Base

Come già accennato, contemporaneamente alla iniezione di elettroni dall'Emettitore alla Base, quando si applica una tensione V_{BE} diretta si ha anche una iniezione di lacune dalla Base nell'Emettitore. La concentrazione di lacune appena entrati in Emettitore (E_1 nella figura) è pari, in analogia con la (1.10), a:

$$p(E_1) = \frac{n_i^2}{N_D} e^{V_{BE}/V_{th}} \quad (1.12)$$

Al contatto metallico di Emettitore invece non ci sarà alcun accumulo di lacune, perché tutte quelle che lì vi giungono vengono subito portate via dal filo del contatto. La distribuzione di lacune lungo l'Emettitore segue quindi il profilo lineare visibile nella Fig.1.9 e permette di calcolare il valore della corrente di lacune che si sta muovendo per diffusione:

$$I_B = qAD_p \cdot \text{grad}(p(x)) = qAD_p \frac{n_i^2}{W_E N_D} e^{V_{BE}/V_T} = I_K e^{V_{BE}/V_T} \quad (1.13)$$

dove W_E è lo spessore dell'Emettitore. Questo contributo costituisce la Corrente di Base, I_B , del transistor perché le lacune iniettate in Emettitore devono essere costantemente rifornite dal contatto metallico della Base. La (1.13) ci mostra che anche la corrente di Base ha un andamento esponenziale con la tensione di comando applicata tra Base ed Emettitore, V_{BE} . Il β del transistor:

$$\beta = \frac{I_C}{I_B} = \frac{D_n N_D W_E}{D_p N_A W_B}$$

dipende solo da termini legati alla costruzione, quali dimensioni e drogaggi. Benché ognuno di questi termini può variare molto da lotto a lotto di fabbricazione, il loro rapporto tende ad essere ragionevolmente controllato, per cui il valore di β di transistori dello stesso tipo differisce al più di qualche decina di %.

1.3.4 I simboli circuitali ed il funzionamento in zona attiva diretta

Nella Fig.1.10 sono riportati i simboli circuitali dei transistori BJT *npn* ed *npn*. La freccia è sempre sul terminale di Emettitore ed indica il verso della corrente. In un transistor *npn* la corrente è equiversa al flusso di lacune dall'Emettitore verso il Collettore, mentre in un transistor *npn* la corrente scorre dal Collettore all'Emettitore nel verso opposto all'effettivo flusso di elettroni.

Nella figura sono anche indicati i versi delle tensioni tra i terminali quando il dispositivo funziona in **zona attiva diretta** in cui la giunzione Base-Emettitore è polarizzata direttamente e la giunzione Base-Collettore è polarizzata inversamente.

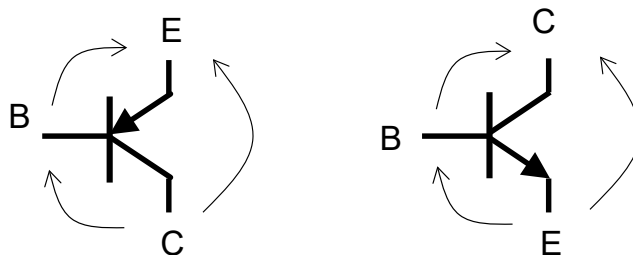
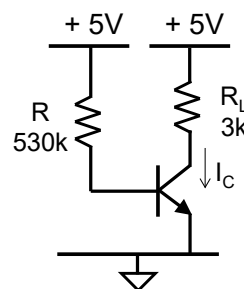


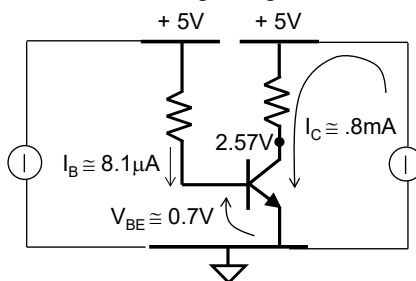
Fig. 1.10 Simboli circuitali dei transistori bipolari a giunzione: a) pnp; b) npn.

E 1.4

Ricavare le tensioni ai capi del transistor collegato come nella figura e le correnti in esso circolanti. Verificare che il transistor effettivamente si trovi a funzionare in zona attiva diretta. Si supponga che il BJT abbia $\beta=100$.



La Base è connessa all'alimentazione positiva attraverso un resistore R, per cui la giunzione Base-Elettore è certamente polarizzata direttamente e, come accade ai capi di un semplice diodo, la tensione V_{BE} si porta a circa 0.7V. La corrente di Base è quindi $I_B \approx 4.3V/R = 8.1\mu A$ e la corrente di Collettore $I_C = \beta I_B = 0.81mA$. Il potenziale del morsetto di Collettore è quindi pari a $V_C = 5V - 2.43V = 2.57V$.



È facile verificare che la giunzione Base-Collettore è polarizzata inversamente di $2.57V - 0.7V = 1.87V$ e che quindi il BJT opera effettivamente in zona attiva diretta. Si noti come una differente scelta del valore di V_{BE} (ad esempio $V_{BE} = 0.67V$ o $V_{BE} = 0.72V$, perché in effetti esso non ci è rigorosamente noto) non avrebbe condotto ad una valutazione significativamente diversa delle correnti circolanti.

Si supponga ora di far variare il potenziale dell'alimentatore connesso alla resistenza R di Base dell'esercizio precedente. Si pensi di diminuire il valore della tensione erogata da 5V a 4V. La corrente di Base diminuisce da $8.1\mu\text{A}$ a $6.2\mu\text{A}$. La variazione $\Delta I_B = 1.9\mu\text{A}$ determina una diminuzione $\beta \Delta I_B = 0.19\text{mA}$ della corrente di Collettore ed una diminuzione di $\beta \Delta I_B R_L = 0.57\text{V}$ della caduta di tensione ai capi del resistore di carico. In definitiva la variazione ΔI_B determina un aumento del potenziale del nodo di Collettore che passa da 2.57V a 3.14V.

Questo è stato l'effetto inizialmente osservato da J. Bardeen e W. H. Brattain: variando la corrente al morsetto di Base, varia la corrente che fluisce nel circuito principale, costituito dalla maglia che contiene R_L , il Collettore, l'Emettitore e l'alimentazione da 5V. Il dispositivo attivo si comporta quindi come una valvola in un condotto idrico: la posizione della valvola controlla la portata del condotto principale. In questo caso la variazione della piccola corrente di Base fa variare la corrente ben maggiore che fluisce tra Collettore ed Emettitore.

Come variabile di comando del transistor si può anche pensare alla tensione V_{BE} tra Base e Collettore: variando di poco la tensione V_{BE} infatti si riesce a variare di molto la corrente I_C e quindi la tensione V_C , con un effetto di "amplificazione" molto interessante.

I transistori sono quindi dispositivi con cui si può controllare e modulare la corrente che fluisce tra Emettitore e Collettore agendo su un opportuno elettrodo di comando, la Base del BJT. Ai loro morsetti si ha una amplificazione di potenza, nel senso che spendendo una piccola potenza per il controllo si modula l'erogazione di una potenza ben maggiore ai capi del carico.

Si noti come, fissata la corrente che deve fluire in un carico, all'aumentare del β del transistor bipolare diminuisce la corrente di Base e quindi la potenza dissipata nel ramo di controllo. Ecco quindi perchè il β è un parametro di merito del transistor.

1.3.5 Le curve caratteristiche

Zona attiva diretta - La relazione esponenziale (1.11) che lega la tensione di comando V_{BE} alla corrente al terzo morsetto (Collettore) è riportata nella Fig. 1.11a. Poichè essa lega la corrente erogata ad una grandezza della maglia di controllo questa curva è detta *curva transcaratteristica* del BJT e ricalca l'andamento tipico della relazione corrente-tensione di una giunzione *pn*. Diversamente, fissata la corrente di Base (o equivalentemente la tensione V_{BE}) è possibile tracciare la dipendenza della corrente I_C con la tensione V_{CE} . Questa è chiamata *curva caratteristica* ed un suo esempio è mostrato nella Fig. 1.11b.

Nelle applicazioni in cui è impiegato per amplificare dei segnali, il transistor è polarizzato nella zona attiva diretta, ovvero nella regione in cui le curve caratteristiche sono quasi piatte (I_C praticamente indipendente dalla tensione V_{CE}) e la corrente circolante è βI_B . In questa regione il transistor si comporta come un

generatore di corrente comandato dalla corrente di Base I_B o dalla tensione V_{BE} . La massima corrente erogabile dal transistor è solo limitata dalla massima potenza termica dissipabile dal dispositivo, mentre la massima tensione V_{CE} deve essere inferiore alla tensione di rottura (*breakdown*) della giunzione Collettore-Base indicata dal costruttore.

Zona di saturazione - La primissima parte delle curve caratteristiche, per tensioni V_{CE} inferiori a 0.7V circa, corrisponde al caso in cui la giunzione Base-Collettore è polarizzata direttamente. In questo caso l'iniezione dei portatori minoritari nella Base avviene da entrambe le giunzioni. I minoritari in Base hanno un profilo ancora rettilineo come quello della Fig.1.9, ma la concentrazione di minoritari alla giunzione Base-Collettore non è più nulla. La pendenza del profilo è minore rispetto a quella che si avrebbe se il transistor funzionasse in zona attiva, e quindi anche la corrente di Collettore è minore. Questa condizione di funzionamento, in cui la Base è invasa (*saturata*) dai portatori minoritari provenienti anche dal Collettore, costituisce la cosiddetta zona di saturazione di un BJT, che è bene evitare quando si progettano circuiti analogici. Considerando una singola curva caratteristica di Fig.1.11b ad I_B costante, la condizione di saturazione è quindi identificata dalla diseuguaglianza:

$$I_C < \beta I_B$$

Zona di interdizione - Quando la tensione V_{BE} è molto piccola, o addirittura si ha una polarizzazione inversa della giunzione Emettore-Base, la corrente I_C è praticamente nulla e si dice che il transistor bipolare è nella zona di interdizione.

Nelle applicazioni in cui il BJT è usato per amplificare segnali, le condizioni di saturazione e di interdizione sono da evitare. Il funzionamento in saturazione ed

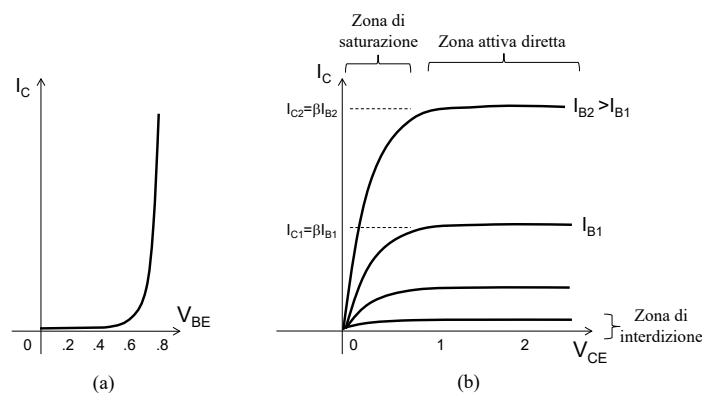


Fig. 1.11 Esempio di (a) curva transcaratteristica e (b) di famiglia di curve caratteristiche di un transistor bipolare npn.

in interdizione è invece utilizzato nelle applicazioni digitali ed in generale quando il transistor è utilizzato come interruttore. Si intuisce infatti come la condizione di interdizione corrisponda allo stato di interruttore aperto, mentre la condizione di saturazione, in cui i potenziali dell'Emettitore e del Collettore sono tra loro vicini (circa 0.2V), corrisponda allo stato di interruttore chiuso.

Attenzione !

Purtroppo si indica con lo stesso termine di *saturazione* due condizioni di funzionamento nel BJT e nel FET completamente diverse, corrispondenti a due tratti diversi delle relative curve caratteristiche.

Nei MOSFET la saturazione corrisponde al tratto di curva caratteristica in cui la corrente di Drain non varia più il suo valore (è saturata) nonostante la variazione della tensione di Drain. È questa la zona di funzionamento di maggior interesse per le applicazioni circuitali analogiche lineari.

Nei BJT invece, la saturazione corrisponde ad un accumulo eccessivo (saturazione) di portatori nella Base provenienti anche dal Collettore ed è in genere una zona di funzionamento da evitare.

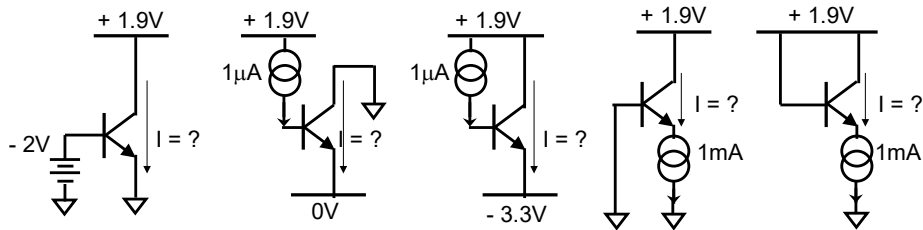
PRATICA di POLARIZZAZIONE

Quali tensioni possono esserci ai tre morsetti di un BJT affinché funzioni correttamente ?

Quali tensioni non possono essere applicate ai tre morsetti in quanto farebbero lavorare il BJT in una zona diversa dalla zona attiva diretta ?

I seguenti quiz dovrebbero aiutarvi a prendere confidenza con questi aspetti fondamentali di un BJT in un circuito elettronico.

ES.1 Si supponga che i BJT abbiano $\beta=200$:



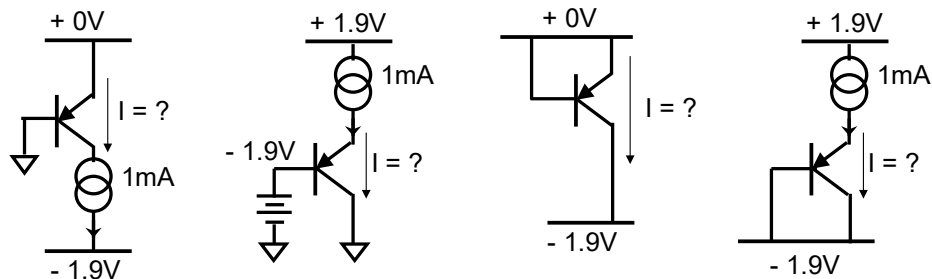
Il BJT è nella corretta zona di funzionamento ?

- si no si no si no si no si no

Se si, quanta corrente I scorre ?

- I = I = I = I = I =

ES.2 Si supponga che i BJT abbiano $\beta=200$:



Il BJT è nella corretta zona di funzionamento ?

- si no si no si no si no

Se si, quanta corrente I scorre ?

- I = I = I = I =

1.3.6 La transconduttanza

Differenziando la relazione (1.11) rispetto alla tensione di comando V_{BE} , si ottiene la transconduttanza del dispositivo in zona attiva diretta:

$$g_m = \frac{\partial I_C}{\partial V_{BE}} = \frac{I_C}{V_{th}} = \frac{I_C}{V_{th}} \quad (1.14)$$

che definisce quanto varia la corrente che fluisce nel Collettore a causa di una piccola variazione della tensione di comando V_{BE} . La transconduttanza dipende dal valore della corrente di polarizzazione I_C ed aumenta proporzionalmente con I_C . Questo è comprensibile perché la transconduttanza è la pendenza della curva esponenziale e quindi aumenta all'aumentare di V_{BE} , cioè di I_C come graficamente evidente dalla Fig.1.12. Ad esempio, con $I_C=1mA$ e ricordandosi che $V_{th}=25mV$ si ottiene $g_m=40mA/V$.

1.3.7 Effetto Early e resistenza di Collettore

Le curve caratteristiche nella zona attiva di un BJT reale non sono perfettamente parallele all'asse delle tensioni ma leggermente inclinate come nella Fig.1.13. Il Collettore quindi non si comporta come un generatore di corrente ideale, erogando una corrente dipendente solo dalla giunzione Emittitore-Base, ma fornisce una corrente che risente, anche se poco, del suo stesso valore di tensione. Per illustrare le ragioni di questa non idealità del transistor BJT si faccia riferimento al transistor *npn* prototipo della Fig.1.14. Si tenga fissa la tensione V_{BE} (equivalente a muoversi lungo una definita curva caratteristica nella Fig.1.13). All'aumentare della tensione V_{CE} , aumenta la polarizzazione inversa della giunzione Base-

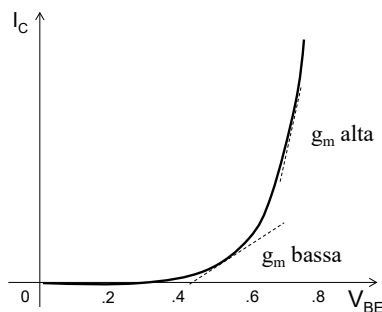


Fig. 1.12 Esempio di (a) curva transcaratteristica e (b) di famiglia di curve caratteristiche di un transistor bipolare *npn*.

Collettore e quindi aumenta lo spessore della zona di carica spaziale della giunzione stessa. Corrispondentemente diminuisce lo spessore neutro della Base, W . Questo produce una aumentata pendenza della distribuzione dei minoritari nella Base che comporta un aumento della corrente di collettore anche se V_{BE} resta costante, come evidenziato nella curva caratteristica. Questo effetto di modulazione della lunghezza della Base al variare della tensione V_{CE} è noto come **effetto Early**.

Quantitativamente esso si esprime nel fornire un valore di tensione V_A , detta **tensione di Early**, che identifica il punto nella Fig.1.13 dove con buona approssimazione convergono le curve caratteristiche. Poiché il grafico ha sugli assi una tensione ed una corrente, è intuitivo definire la "resistenza di collettore" equivalente del BJT come:

$$r_o \cong \frac{V_A}{\beta \cdot I_B} = \frac{V_A}{I_0 \cdot e^{\frac{qV_{BE}}{kT}}} \quad (1.15)$$

e rappresentarla nel simbolo circuitale della Fig.1.13 da una resistenza di uscita r_o tra Collettore ed Emettore. Si noti come il valore di r_o dipenda anche dalla corrente a cui il transistor opera: maggiore è la corrente circolante maggiore è la pendenza della curva e quindi minore è il valore di r_o corrispondente. Ciò proprio perché V_A è un valore fisso, in prima approssimazione definito dalla tecnologia costruttiva. La tensione di Early di un BJT ha un valore tipico compreso tra 50V÷200V che, per correnti di collettore intorno ad 1mA, corrisponde ad una resistenza di Collettore di 50 kΩ÷200kΩ.

Per ricavare la effettiva corrente circolante nel transistor date le tensioni ai morsetti, basta sommare al valore ideale (quello che si avrebbe per $V_A=\infty$) il contributo aggiuntivo dovuto alla pendenza (per comodità si considera il triangolo)

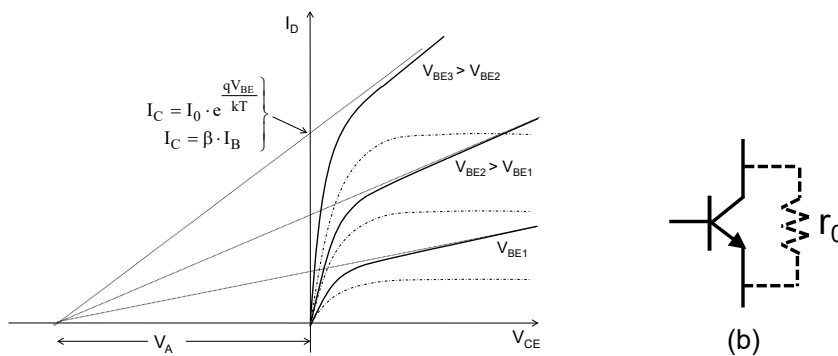


Fig. 1.13 (a) Curva caratteristica di un transistor bipolare reale con in evidenza la tensione di Early V_A e (b) visualizzazione della resistenza equivalente prodotta.

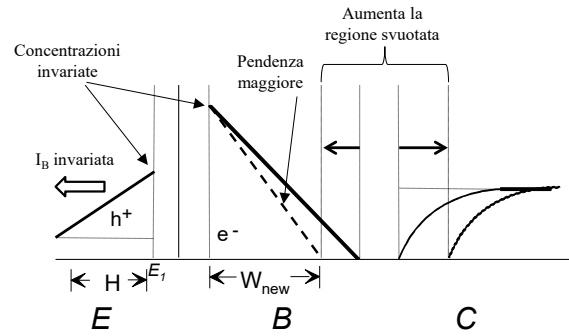


Fig. 1.14 Effetto di un aumento della tensione inversa tra Collettore e Base in un transistor npn.

$$I_C = I_0 \cdot e^{\frac{q(V_{BE} + v_{be})}{kT}} + I_0 \cdot e^{\frac{q(V_{BE} + v_{be})}{kT}} \frac{V_{CE}}{V_A} = I_0 \cdot e^{\frac{q(V_{BE} + v_{be})}{kT}} \cdot \left(1 + \frac{V_{CE}}{V_A} \right)$$

Concludendo questo primo capitolo hai concluso la tua prima fatica. Sentiti come Ercole che per espiare il fatto di essersi reso colpevole della morte della sua famiglia, viene convinto dall'oracolo a mettersi al servizio di Euristeo ed a compiere una serie di imprese, ottenendone alla fine come compenso l'immortalità.

La prima fatica di Ercole fu catturare e uccidere il leone di Nemea, una bestia dalla pelle durissima che non poteva essere scalfita, che sbranava e divorava animali e pastori. Ercole scelse di battersi a mani nude contro la belva, cingendogli il collo con forza e riuscendo a stringerla fino a soffocarla. Dopo aver ucciso il leone Ercole conció la pelle invincibile del leone e ne fece un mantello per se stesso, mentre della testa ne fece un elmo. La prima fatica era compiuta.

