The switching regime of the semiconductor devices

Diode

Bipolar Transistor

Unipolar MOS Transistor

Semiconductor Materials

Semiconductors : solid or liquid materials with an electric conductivity somewhere between the conductors and the insulators

conductors (metals): resistivity varies between 10^{-6}- 10^{-4} ohm·cm

insulator's resistivity (diamond, quartz) is between 10^{10}-10^{20}ohm·cm

Semiconductor materials (silicon and germanium): intermediary resistivities (hundreds or thousands ohm·cm); found on the 4th column of the Mendeleev table, having 4 valence electrons

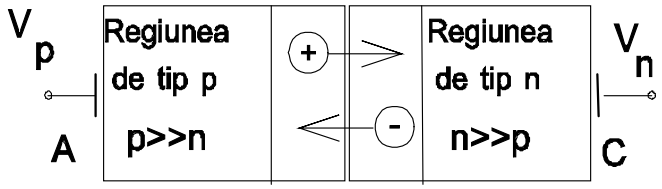
Semiconductor kinds:

n type, where the electrons are in excess; obtained by adding impurities like Phosphorus, Arsenic, elements that are found on the fifth column

p type, where charge carriers in excess are the positive ones (holes); obtained by adding impurities like Boron, Gallium or Aluminum, placed on the third column and having a three valence electrons

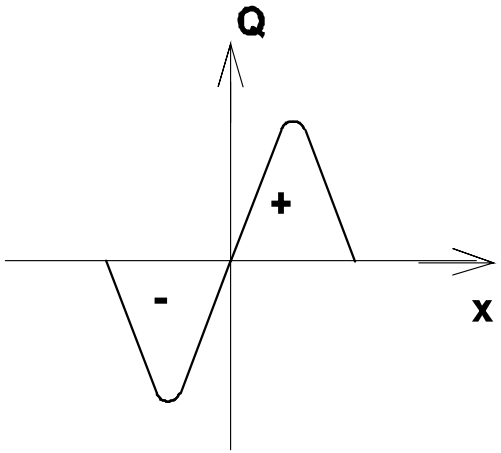
Switching regime of the semiconductor diode

The pn junction at thermal equilibrium



a)

thermal equilibrium: no electrical current shall flow through the semiconductor



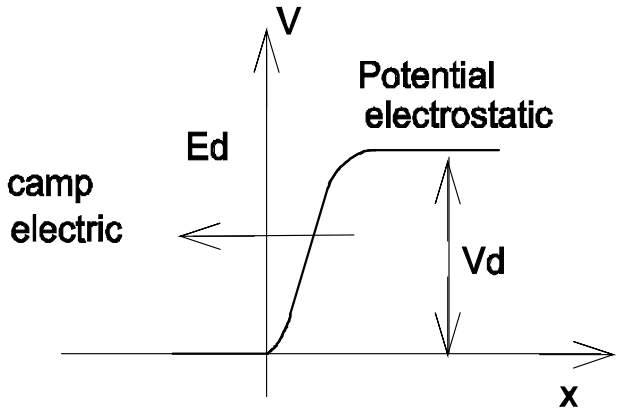
b)

An organized movement of the charge carriers:

-applying an external electrical field

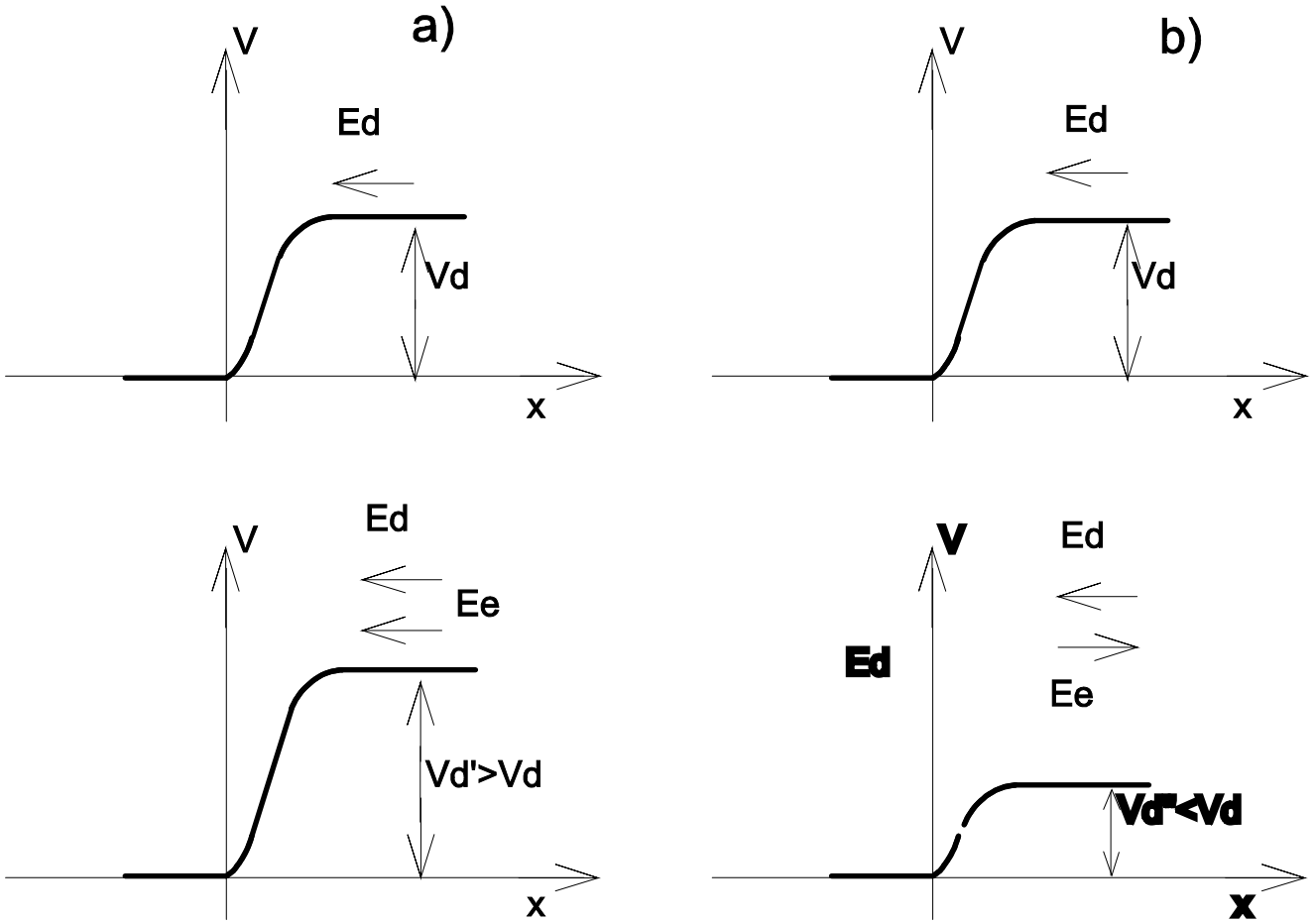
-inducing a non-uniform distribution of the charge carriers (diffusion process)

Near junction plane, electrons from n-region tend to diffuse in p-region and similar for holes from p-region towards n-region

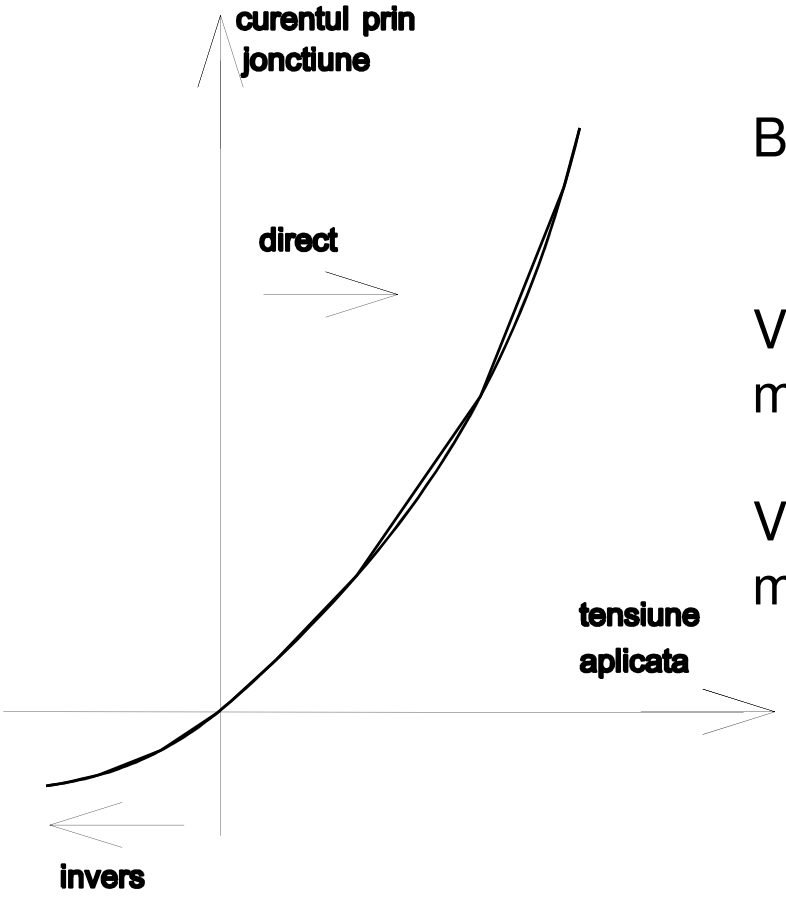


c)

Diffusion electric field (Ed): associated to a difference of potential, named *barrier potential* near the plane of the junction (due to presence of electrons in p-region and holes in n-region; opposes to the tendency of diffusion, so have a process with self-limitation



A voltage which increases the potential barrier is called **reverse polarity voltage**
An external voltage of an opposite polarity is called **direct voltage**



Current-voltage characteristics

Be a voltage applied on diode electrodes:

$$U = V_p - V_n$$

$V_p > V_n$ – great diode current due to the majority carriers; the junction is forward biased

$V_p < V_n$ – current is negligible due to the minority carriers; the junction is reversely biased

Static switching parameters of the semiconductor diode

For an ideal **pn** junction, the current-voltage relation is:

$$I = I_0(e^{U / c_r U_T} - 1)$$

I_0 is the reverse saturation current of the diode $I_0 = eS\frac{c_{dg} \cdot p_{m0}}{w}$

where:

e: electrical charge of the electrons ($e=1,6 \cdot 10^{-19} C$)

S: transversal section through the junction

c_{dg} : the diffusion coefficient for holes

p_{m0} : the concentration of the minority carriers.

w : width of the recombining zone

c_r - recombining factor of the holes (has value 1 for Ge, value 2 for Si)

U_T is the thermal voltage: $U_T = \frac{KT}{e} = \frac{T}{11800}$

Where:

K is the Boltzmann constant ($K= 1,38 \cdot 10^{-23} J/^oK$)

T is the absolute temperature

Static (dc) resistance of the diode, value depends on the functioning point

$$R_{cc} = \frac{U}{I}$$

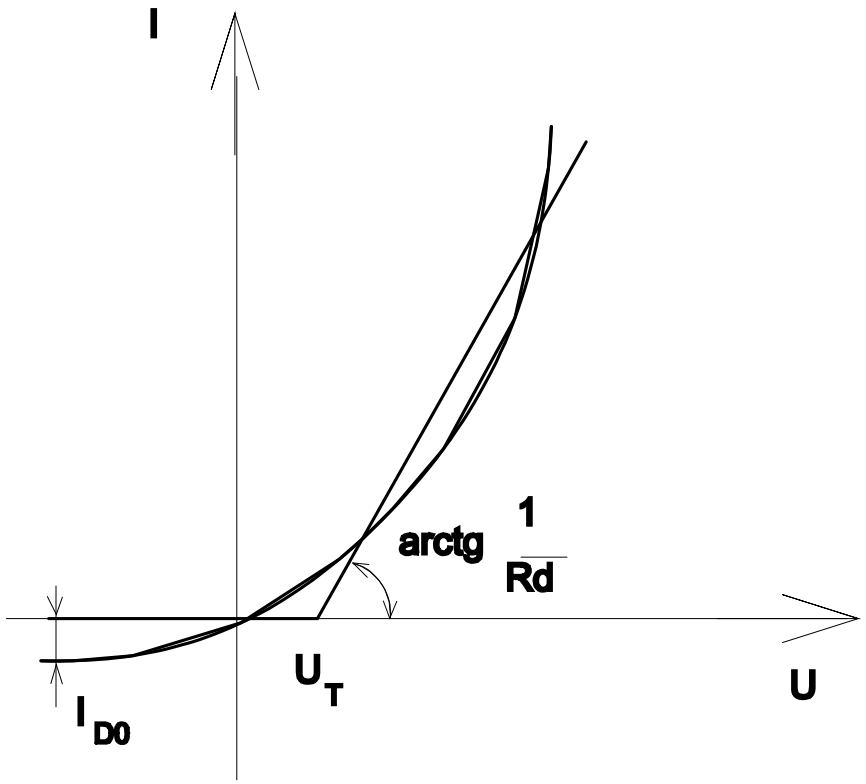
Dynamic (ac, differential) resistance

$$R_d = \frac{\Delta U}{\Delta I}$$

Barrier capacity C_b, it depends nonlinear with the voltage

Diffusion capacity C_d, results from the stored charges through the diffusion of the minority carriers. It depends on the forward bias voltage

$\mathbf{U_T}$ threshold voltage,
$\mathbf{R_d}$ differential resistance,
$\mathbf{I_0}$ residual current
are giving
the linear approximation of the volt-ampere diode characteristics



Typical values:

$I_0 = 0{,}05A$

$U_T = 0{,}5V$

$R_d = 15ohm$

Dynamic parameters of the semiconductor diodes

Forward switching time

Necessary time for the diodes to pass from the blocked (off) state to the conduction state

Smaller than the reverse switching (recovery) time

Reverse switching time

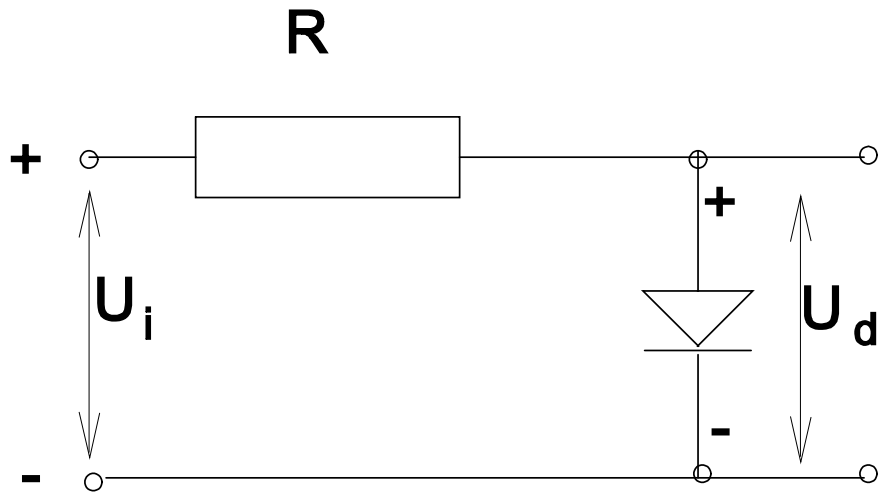
Time when diode is passing from the conducting state to the blocking (off) state

Switch analysis using circuit below

Two time components:

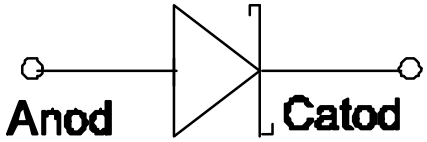
Storage time: $t_s = \tau \ln(1 + \frac{I_D}{I_I})$

Fall down time: $t_c = 2{,}3 \cdot R \cdot C_b$



Schottky barrier diodes (*hot-carrier* diodes)
Based on a metal-semiconductor contact
Electrical current is created by the movement of the majority carriers

Manufactured from a semiconductor material (silicon)
of type **n**, in contact with a metal, gold or Aluminum



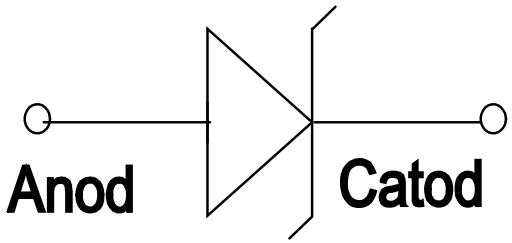
Forward biasing of the metal: a junction current is created: movement through the junction of the electrons within the semiconductor
The current is created from the movement of electrons, the majority carriers
Have a greater energy than the one of the free electrons, name of hot carriers.

Reverse bias: electrons that broke through the metal are not distinguished from the free electrons of the metal; no residual current
Reverse switching time of the Schottky diode very small (≈10ps)

Schottky diodes to create high speed switching circuits

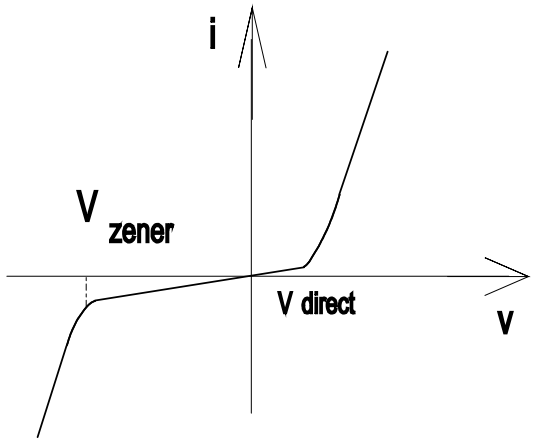
The forward voltage fall on a Schottky diode is ≈0,4V, unlike the 0,75V for the Si based diodes and the 0,3V for the Ge ones

Zener diodes



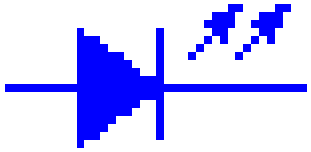
Reverse biasing applied on a **pn** junction: if the existing electric field possesses the necessary energy to extract the electrons from their covalent rotating orbits, there will be a sudden increase of current

The electron-hole pairs newly created contribute to the generation of an important current, and it is said that the diodes work in the Zener region



Value of the threshold Zener voltage depends on the dopping degree and can range from values of 2V to hundreds of volts

LED - Light Emitting Diode



At a forward bias of junction, electrons from **n** region get recombined with holes from **p** region, generating energy as heat and light

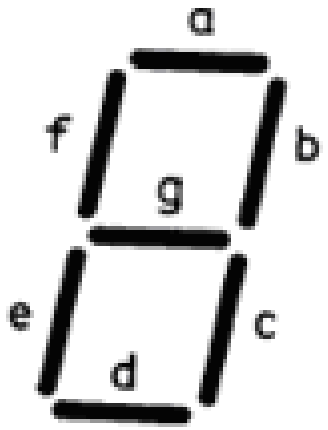
Doping regions with impurities, the emitted light's wavelength may be controlled, obtaining different LED colors

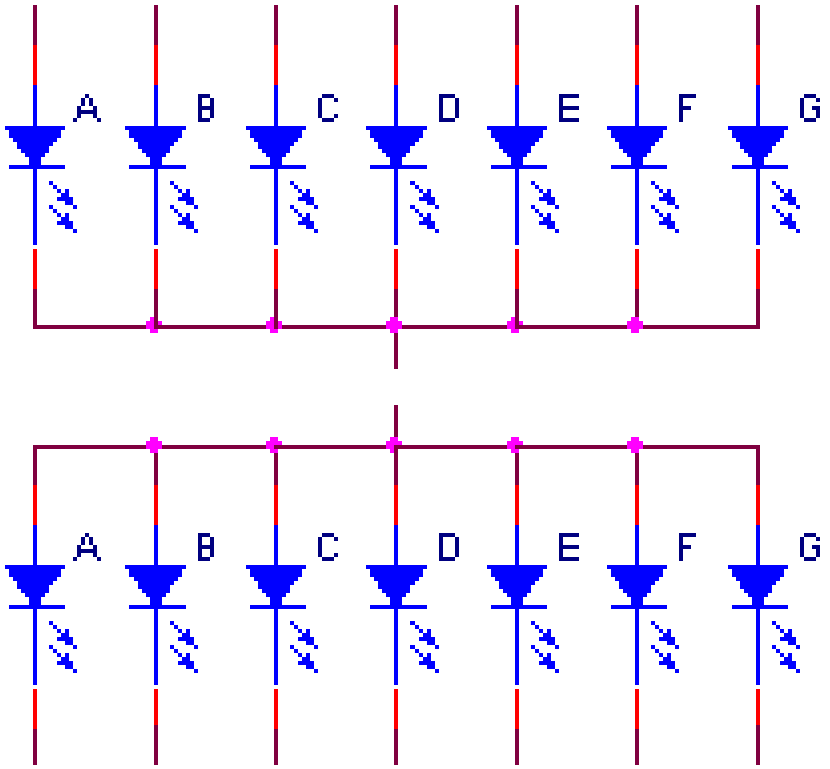
- GaAsP galium-arsen-phosphorus – red/yellow light
- GaP – red/green light

Direct voltage is higher than for Si diodes (1,2V - 3,2V)

Direct current is higher also ($\geq$10mA)

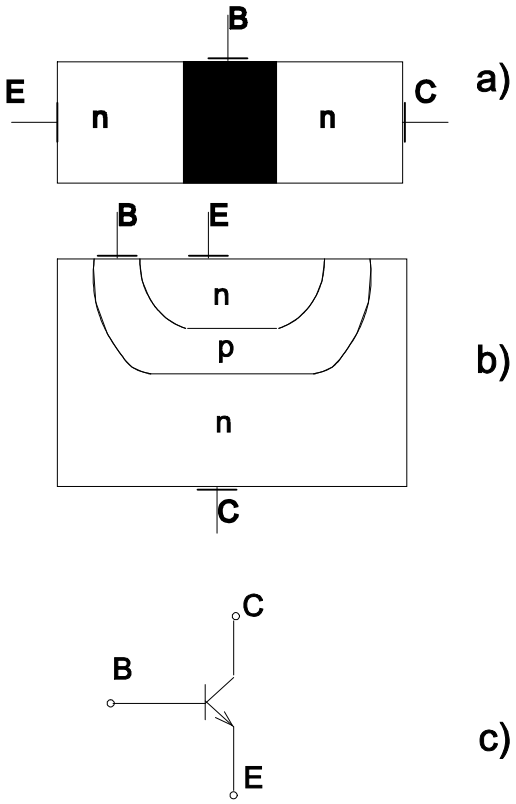
- 7 segments display: 7 LEDs in an array displaying the decimal digits
- Two types:
  - Common cathode – LEDs are driven in anode with an active ‘1’ signal
  - Common anode – signal applied in LEDs cathode, with active low level





Bipolar junction transistor during the switching regime

Functioning regimes of a bipolar transistor

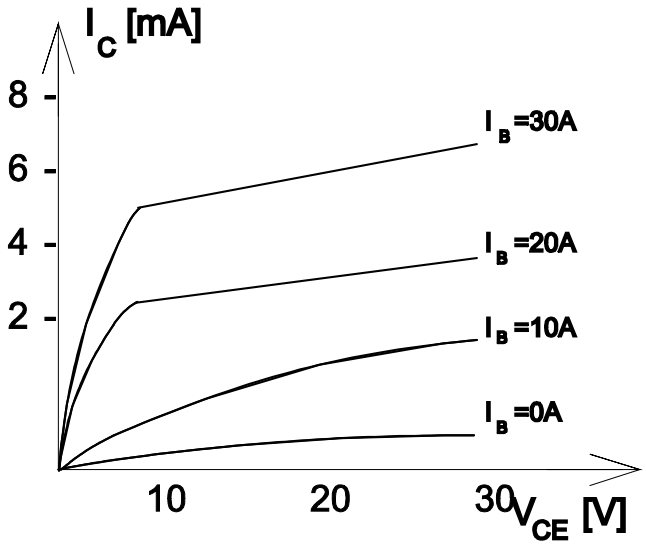


	Base-emitter junction, forward bias	**Base-emitter junction, reverse bias**
The base-collector junction, forward bias	The saturation region	The reverse active region
The base-collector junction, reverse bias	The forward active region	The cut-off region

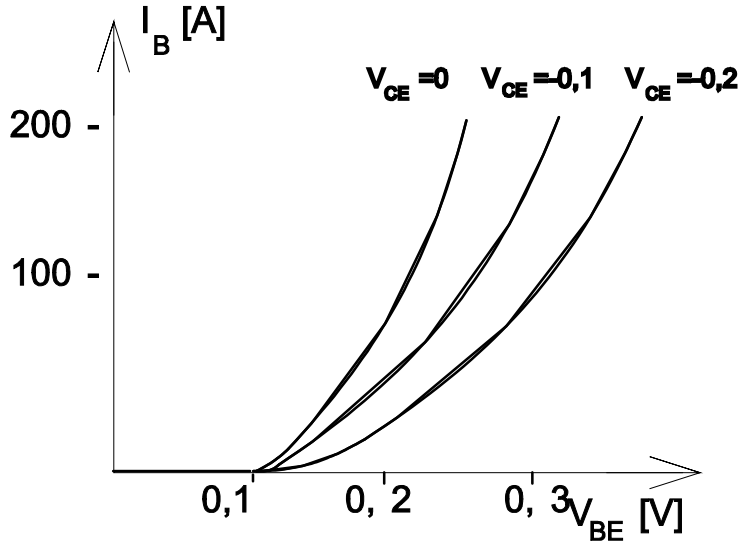
The structure of a bipolar transistor with junctions

The most important relations for **forward active region** are:
$I_C = \alpha \cdot I_E$.

α is the gain in current of the transistor with the common base
Applying a current I_B into the base, a collector curent will result:
$I_C = I_B \cdot (\alpha/(1-\alpha)) = ß \cdot I_B$.
ß is the current gain for a common emitter configuration, or
amplification of the current (values from 10 to 1000)



Output characteristic



Input diagram

Cut-off region when both junctions are reverse biased;
State is defined by relations:

$V_{BE} \leq 0$

$V_{CE} - V_{BE} > 0$

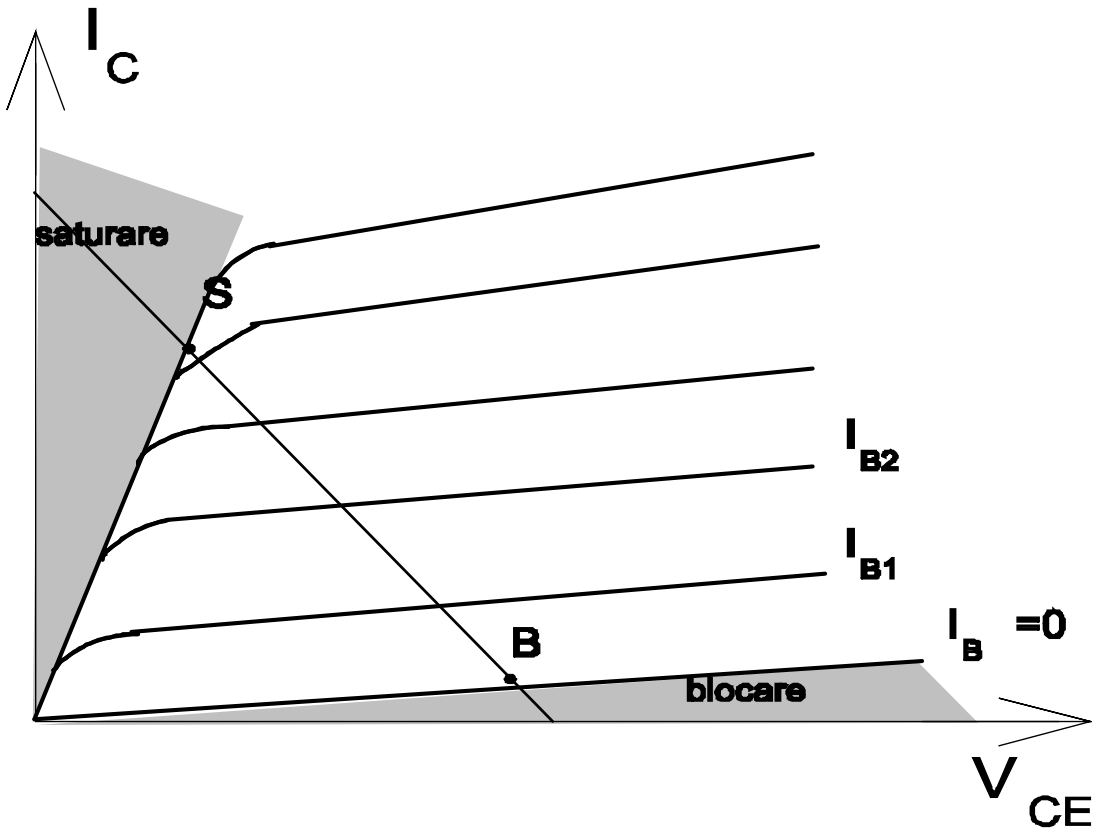
Saturation region implies that both junctions to be forward biased
The relations for the saturation regime are:

$V_{BE} > V_{CE}$

$I_C < \beta \cdot I_B$

$V_{CEs} \approx 0{,}2V$

Reverse active region is equivalent with a normal behavior in active region, in which the roles of the emitter and the collector are reversed
Rarely used because the current amplification has a very small value ($\alpha_i \approx 0{,}1$)



Operation points of the switching regime:
S is the saturation driving point
B is the cut-off driving point

Dynamic switching parameters of the bipolar transistor

Forward switching time

tcd is defined as the necessary time for a transistor to switch from cut-off to conduction (this includes saturation)

Two components:

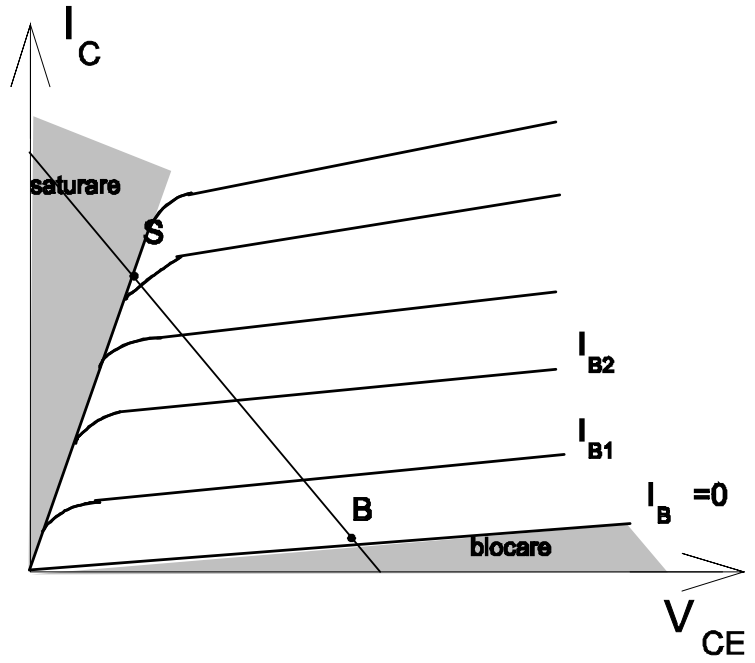
t_i, - delay time

t_r, - raising time

The delay time t_i is composed by more components:

- Necessary time for the minority carriers to pass through the base (reach of B point)
- Necessary time for the value of the collector current to increase from I_{C0} (or 0), to $0,1 \cdot I_{Cs}$

The values of these times (so of delay) are small, negligible, the rising time being the most important



Raising time, denoted by t_r is defined by the time interval in which the collector current increases in value from $0{,}1 \cdot I_{Cs}$ to the value $0{,}9 \cdot I_{Cs}$

Time is driven by the applied direct base current, and usual this is the direct base current for approaching point S (start of saturation)

$$I_{Bds} = \frac{I_{Cs}}{\beta_{N0}}$$

To foster unblocking, use of over-driving currents, defining *over-driving factor* :

$$N_d = \frac{\beta_{N0} I_{Bd}}{I_{Cs}}$$

Reverse switching time

Switch of transistor from a conduction state (forward active/saturation) to a blocking (off) state

Time called t_{ci} it has two components:

storage time t_s,

falling down time t_c

Storage time t_s has components:

t_{s1} represents the time necessary to eliminate the excess of charges from the base, such that the transistor would be in the forward active region

t_{s2} represents the time needed for the collector current to decrease from I_{Cs} to $0,9 \cdot I_{Cs}$

$$t_{s1} = \tau_s \ln \frac{I_{Bd} - I_{Bi}}{I_{Bs} - I_{Bi}} \qquad t_{s2} = \tau_r \ln \frac{1 + \frac{1}{N_b}}{1 + \frac{0,9}{N_b}}$$

τ_s is the storage time constant

I_{Bd} is the forward base current

I_{Bi} is the reverse base current

I_{Bs} is the base current at the boundary of the saturation and the forward active regions

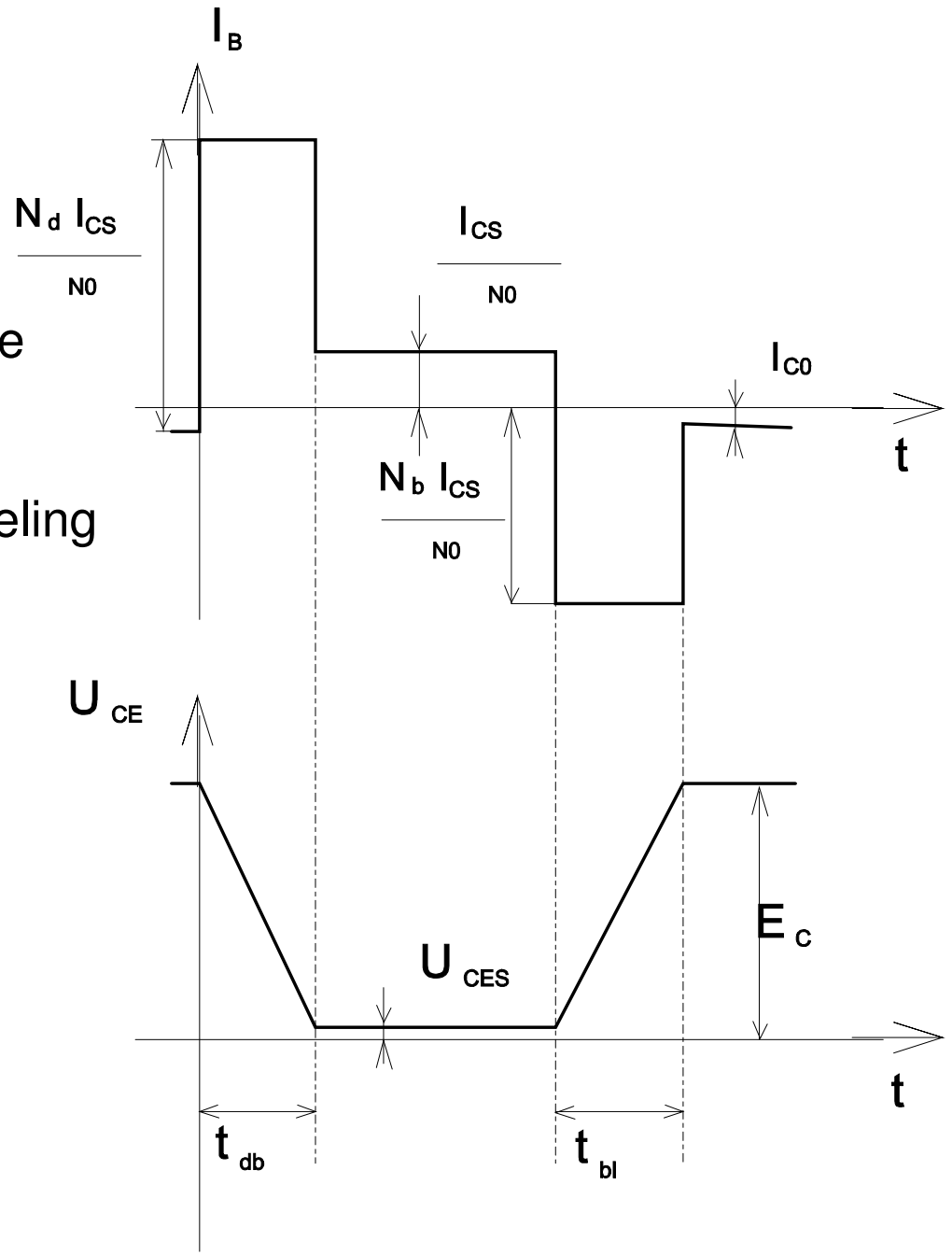
t_c, falling down time (necessary for the collector current to decrease from $0.9I_{Cs}$ to $0.1\ I_{Cs}$, reaching vicinity of point B):

$$t_c = \tau_r \ln \frac{1 + \frac{9}{10\,N_b}}{1 + \frac{1}{10\,N_b}}$$

N_b represents the blocking over-driving coefficient $\qquad N_b = \dfrac{\beta_{N0}\, I_{Bi}}{I_{C0}}$

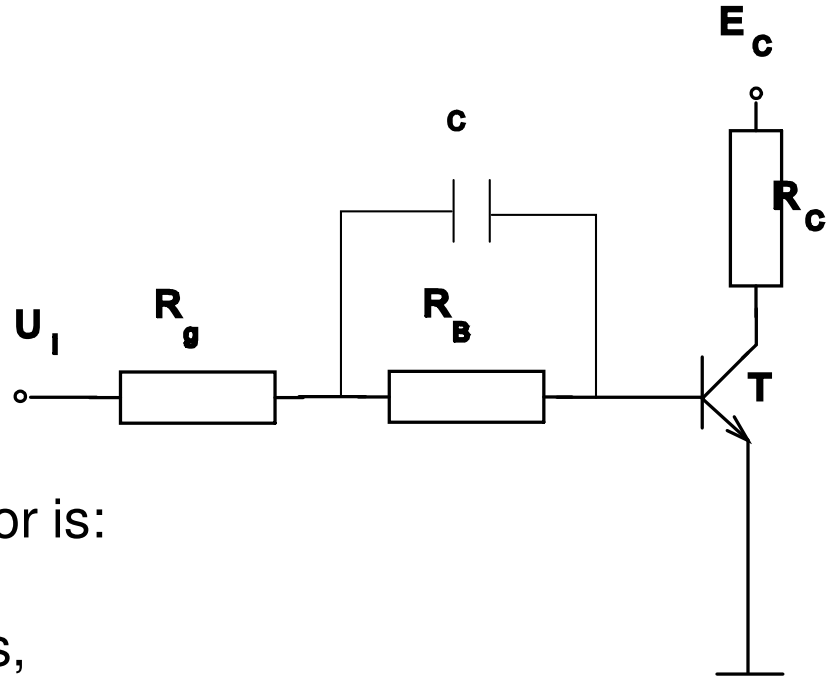
Speed-up methods

Following methods to reduce the switching time for a transistor:
-over-drive at unblocking,
to decrease the direct switching time
-over-drive at blocking, to decrease the reverse switching time
-avoid the saturation stage by canceling the storage time



Speed-up capacitances

Apply of a pulse voltage on circuits input
U_1 high level, U_2 low



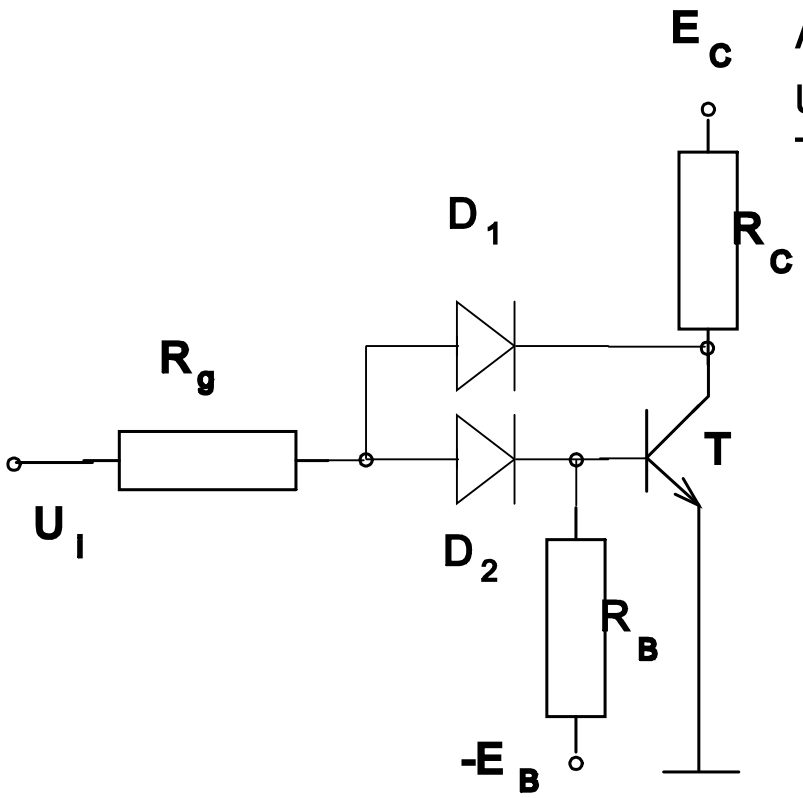
For a positive voltage swing:

Time constant for charging up the capacitor is:
$\tau_{inc} = C \cdot (R_g + R_{in})$
base current now exponentially decreases,
with the same time constant, towards a constant
value which does not allow the saturation to take place:
$I_{Bd}=U_1/(R_g+R_B+R_{in})$

For a negative voltage drop:

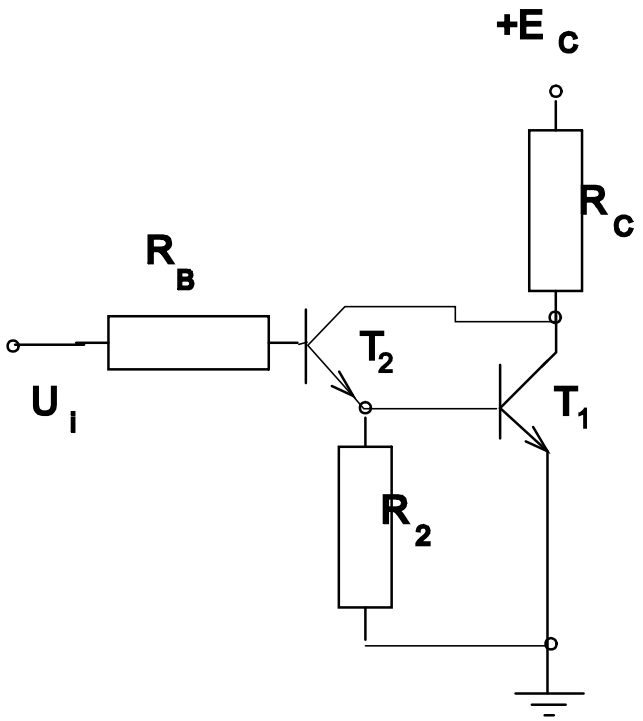
Base current falls to I_{Bi} whose value is determined by U2 and Rin' – the high resistance of the blocked transistor. Then, once the transistor is blocked and the capacitor is unloaded, the reverse current exponentially decreases towards a constant value: $I_{Bi} = I_{C0}$
Capacitor time constant now: $\tau_{disc} \approx C \cdot R_B$

Using a negative, non-linear voltage reaction to avoid saturation



Avoid the saturation of the T transistor by using the diodes D1 şi D2

The diodes have different junction potentials



Using a Darlington circuit

T2 – the control transistor and T1 – the output transistor

T1 never gets saturated as long as the collector potential is always greater than the base potential:

$$U_{C1}=U_{CE2}+U_{B1}$$

Therefore, $V_{C1}>V_{B1}$, and the base-collector junction is reversed biased

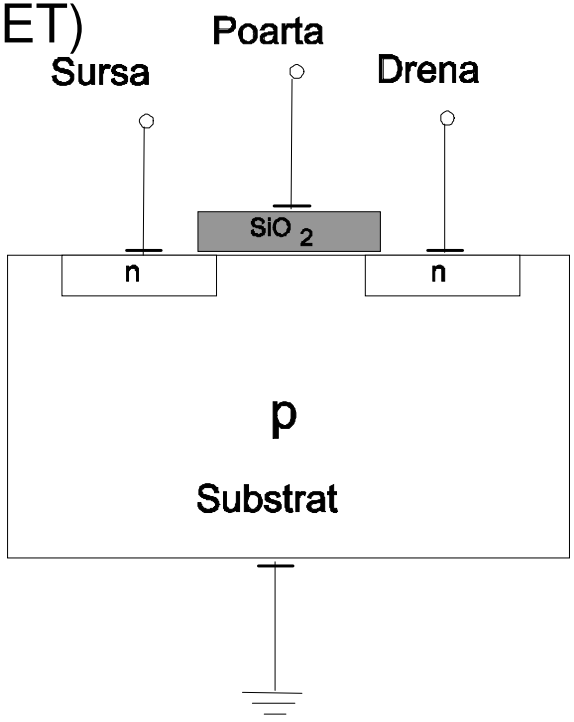
The Field Effect Transistor

Classification

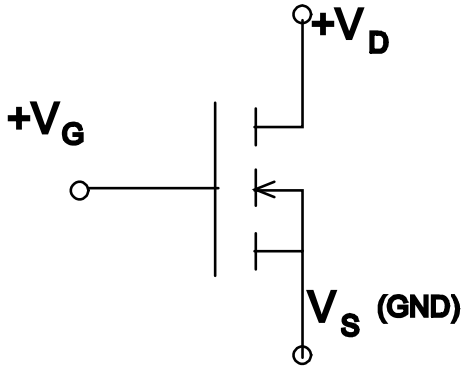
-gate junction transistor
-insulated gate transistor
-thin layers transistor

Study of IGFET (also known as MOSFET (*metal-oxide* FET)) or insulated gate transistors:
Grouped by the manufacturing technology in:
- MOS transistors with induced channel (enhancement MOSFET)
- MOS transistors with initial channel (depletion MOSFET)

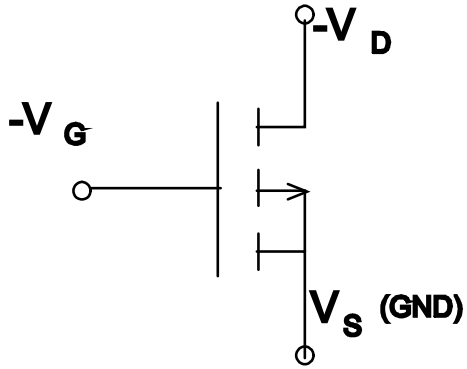
The physical structure of a MOS transistor is presented
Shown a cross section of a MOS transistor with
n induced channel, also called NMOS



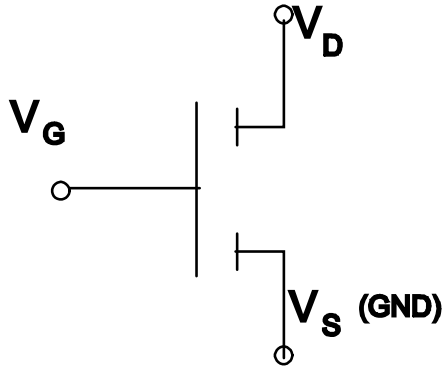
Components:
-basis substrate
-source
-drain
-gate (grid)



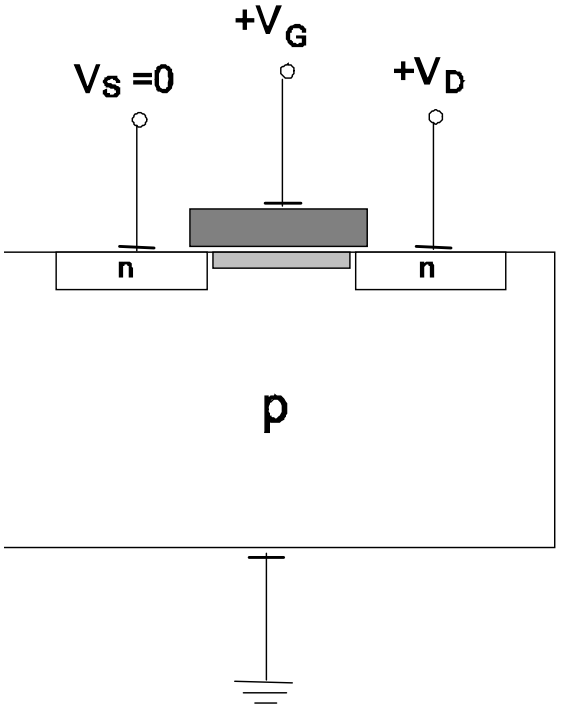
NMOS



PMOS



CMOS



NMOS Electrodes

NMOS Operation presented next slide

If a positive voltage, relative to the drain and the source regions, is applied on the gate electrode (on the grid) then the electric charges of **p** type from the basis substrate will be rejected and the electrons from drain and source regions will be attracted to the surface of the silicon substrate laying under the gate.

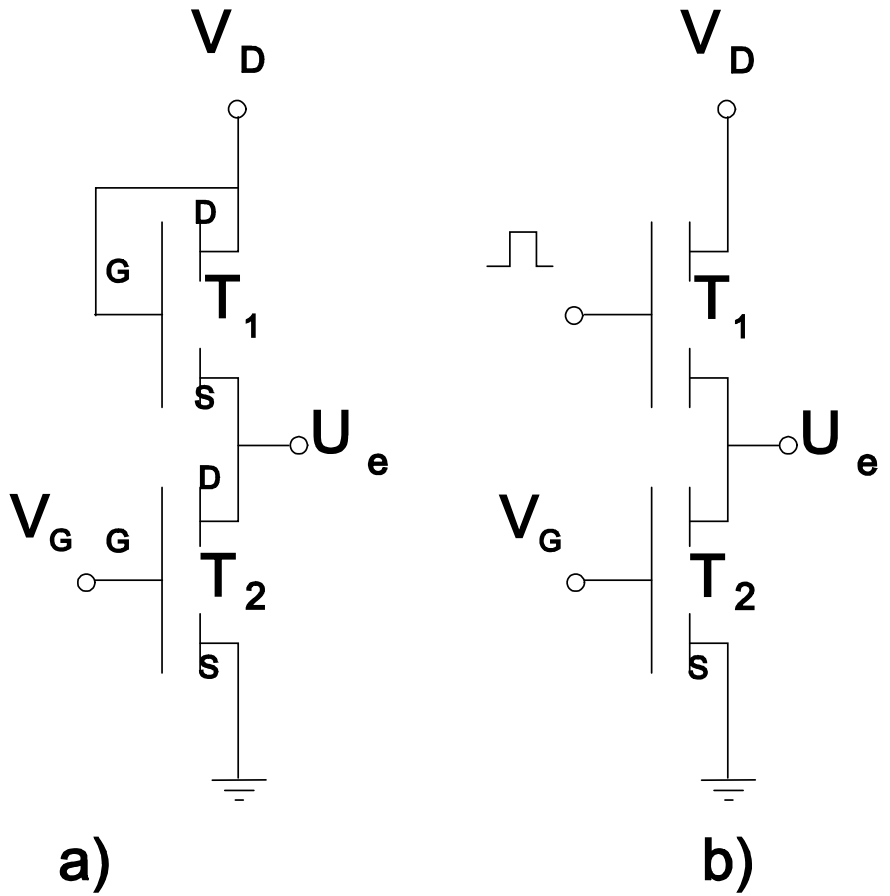
Assume that V_{DS}, the voltage between the drain and the source, is much lower than V_{GS}, the voltage between the gate and the source. Therefore, the potential voltage from this region, laying under the silicon dioxide insulator, is relatively constant. By increasing the grid potential, the intensity of the existing electric field is also increased, as the number of electrons from the region multiplies. We may assume that there exists a channel between the drain and the source whose depth increases proportionally with the voltage applied on the gate. For a **threshold voltage** of V_{GS}, denoted by $\mathbf{V_T}$, the concentration of the electrons will exceed the concentration of the holes. Consequently the region will reverse its type: from a **p** type region it will become an **n** type region. It follows that on the surface of the **p** type semiconductor an **n** type channel has been created. The conductance of the drain-source region and the drain current will increase proportionally with V_{GS}.

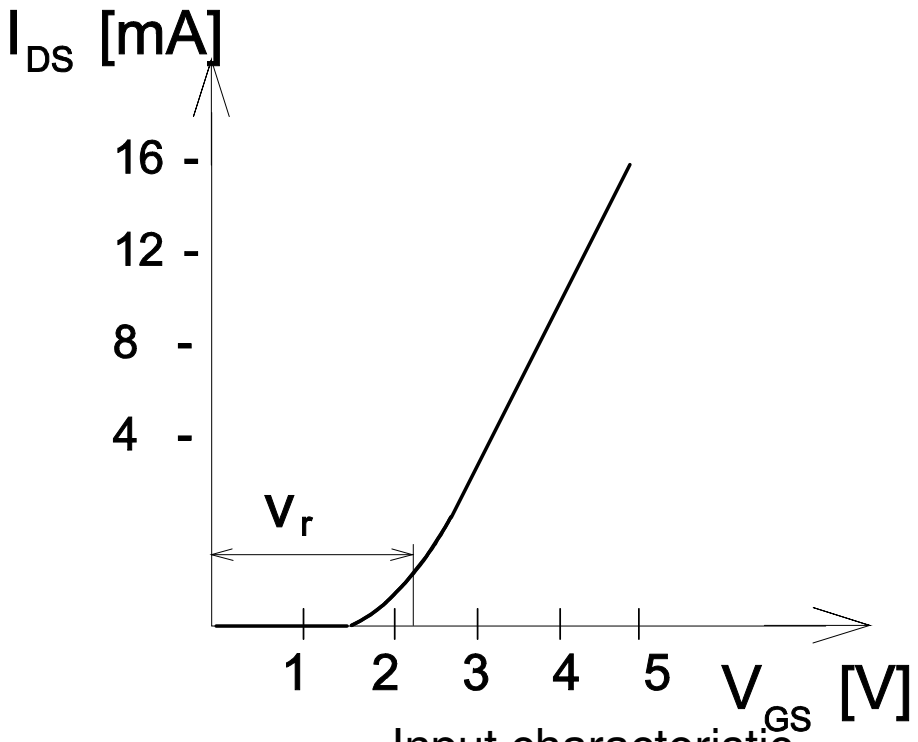
If the voltage between the drain and the source decreases, the voltage between the gate and the drain will also decrease. The channel will present an irregular distribution of the electric charges. This denotes a non-linear dependence of I_{DS} with respect to V_{DS}

Advantages of using MOS transistors:
more easily manufactured than bipolar transistors
higher density of integration
high input impedance (10^{14}-$10^{16}\Omega$), a low driving current
use in MOS structure of an active resistance,

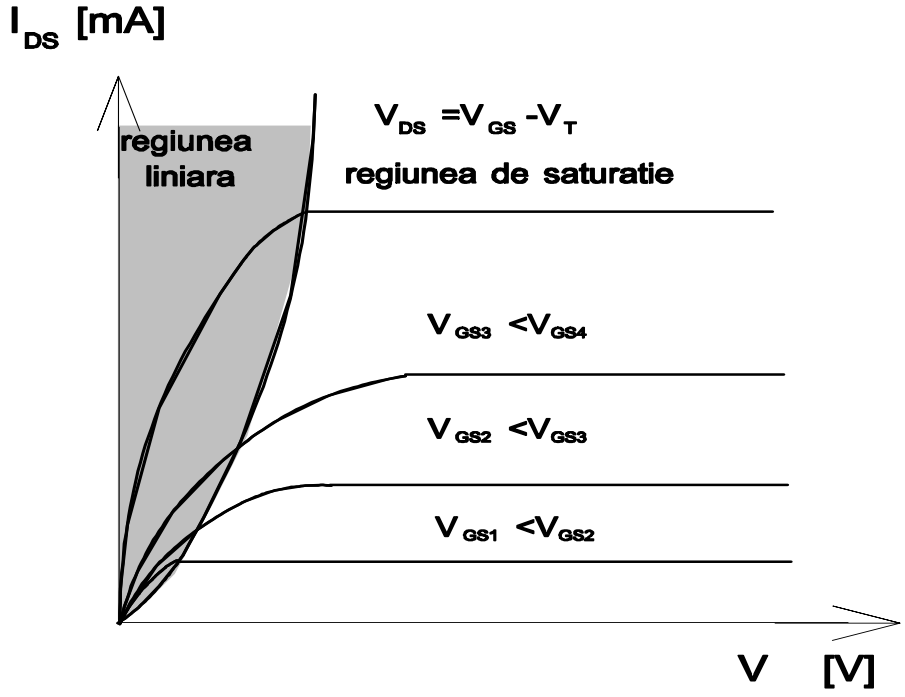
Some disadvantages implied by caution to storage and transport

Ways to implement active resistance





Input characteristic



Output characteristics

From the output characteristic, three regions can be deduced and analyzed:
the cut-off stage: the output current, the drain-source current I_{DS} is almost null and the input voltage, V_{GS}, is lower than the threshold voltage: $V_{GS} < |V_T|$

the linear (triode) stage: the region situated at the left of the current characteristic, when $V_{DS} = V_{GS} - V_T$; the drain current I_{DS} quickly increases as a function of the drain-source potential V_{DS} :

$$I_{DS} = K((V_{GS} - |V_T|)V_{DS} - \frac{V_{DS}^2}{2})$$

Also: $0 <= V_{DS} <= V_{GS} - |V_T|$

-**the saturation stage**: the region situated at the right of the current characteristic, when $V_{GS}-|V_T|=V_{DS}$

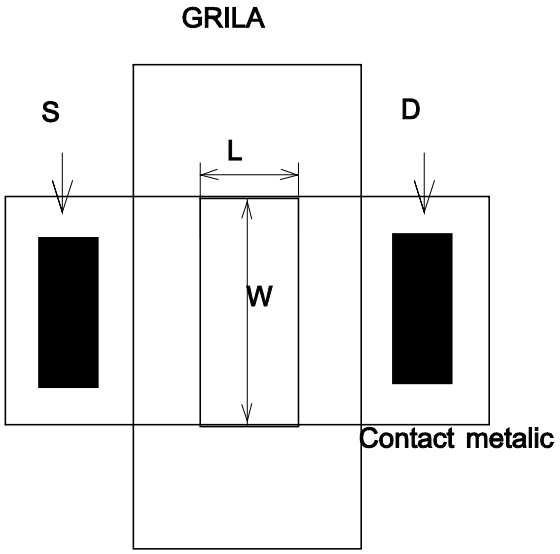
The following relations take place:

$0 <= V_{GS} - |V_T| <= V_{DS}$

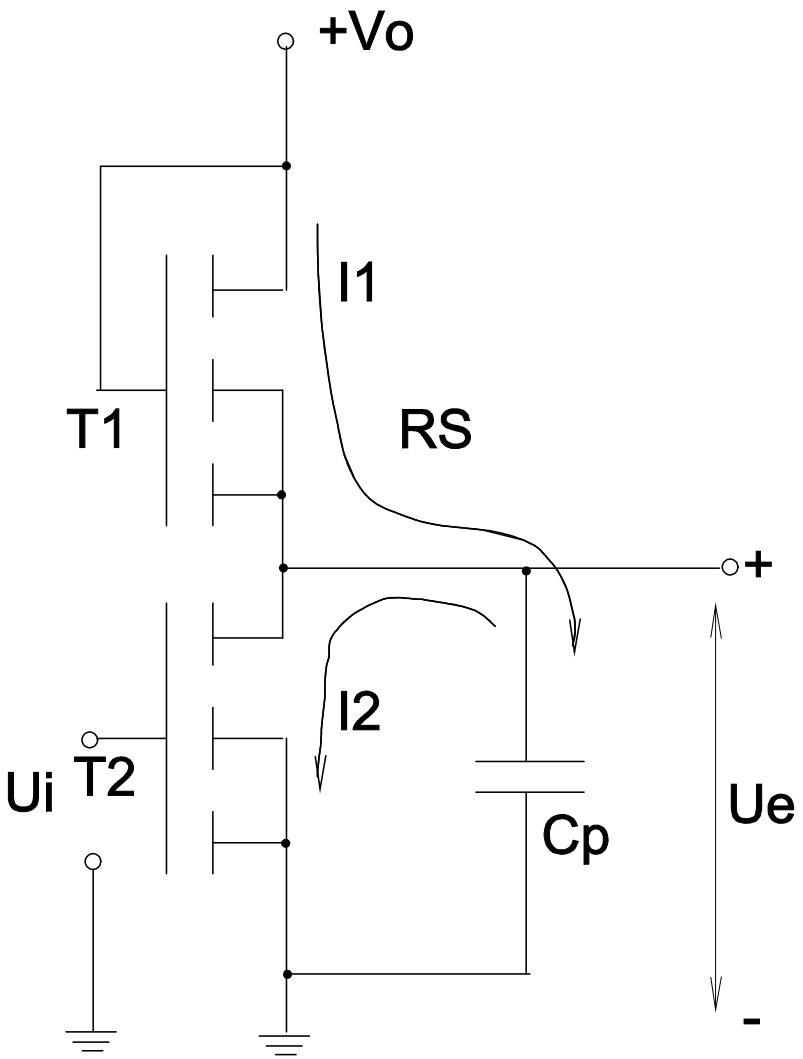
$$I_{DS} = \frac{K}{2}(V_{GS} - V_T)^2$$

K, conduction factor ≈ß·(W/L), where:

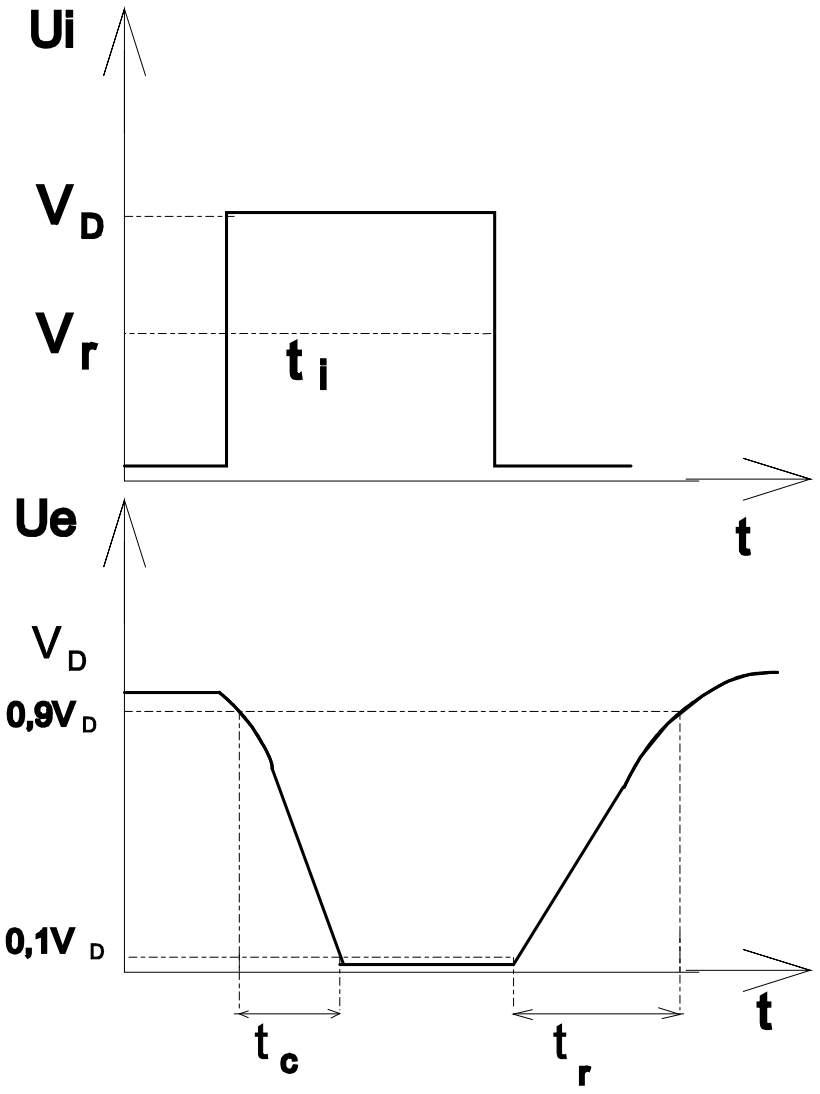
- ß is the internal conduction factor; it has an approx value of 10μA/V²
- W is the channel width; it takes values in the range of 10-200μm
- L is the channel length; it taks values between 1-10μm



Dynamic Parameters for MOS Transistor



Layout of a MOS Inverter



Switching Times

Assume that a MOS transistor has to drive the gate of one or more MOS transistors. It must guarantee the charging and discharging of the input capacitances for the driven transistors. Denote by C_p the sum of all input capacitances of the driven transistors. We may associate the switching times of the driving transistor with the charging/discharging times for the capacity C_p. The formula for the time constants of the RC circuits are:

$\tau_{inc} = R_s \cdot C_p$
$\tau_{desc} = R_T \cdot C_p$

where, R_T is the direct resistance of the conducting transistor
R_s is the load resistance
We choose that the output voltage for the low logic level being as close as possible to the ground potential, $R_s >> R_T$.
Thus, the raising and falling times for MOS transistors are:
$t_r = 2{,}2 \cdot R_s \cdot C_p$
$t_c = 2{,}2 \cdot R_T \cdot C_p$