# Chapter 1. Introduction to Solid State Physics.

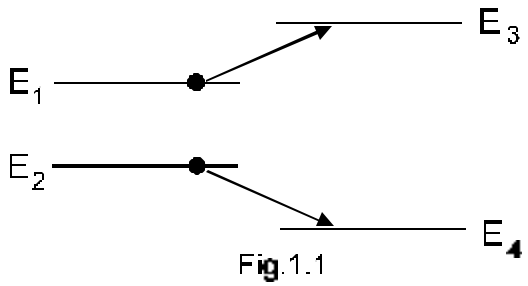


Fig.1.1

## 1.1. Fermi – Dirac Distribution and the Density of Energy States in a Solid

Let be $P(E_1)$ the probability to have an electron in the state characterised by the energy $E_1$, then $1-P(E_1)$ will be the probability to have not an electron in this state (on this energy level). For the energy level configuration, depicted in Fig.1.1, the total probability to have such a state ($E_1$ and $E_2$ filled and $E_3$ and $E_4$ unfilled) is given by the formula:

$$P(E_1)\cdot P(E_2)\cdot(1-P(E_3))\cdot(1-P(E_4))$$

The probability for the complementary situation is:

$$(1-P(E_1))\cdot(1-P(E_2))\cdot P(E_3)\cdot P(E_4)$$

Both probabilities must be equal in the case of thermal equilibrium, therefore we can write the following equality:

$$\left(\frac{1}{P(E_3)}-1\right)\cdot\left(\frac{1}{P(E_4)}-1\right)=\left(\frac{1}{P(E_1)}-1\right)\cdot\left(\frac{1}{P(E_2)}-1\right) \qquad 1.1.1.$$

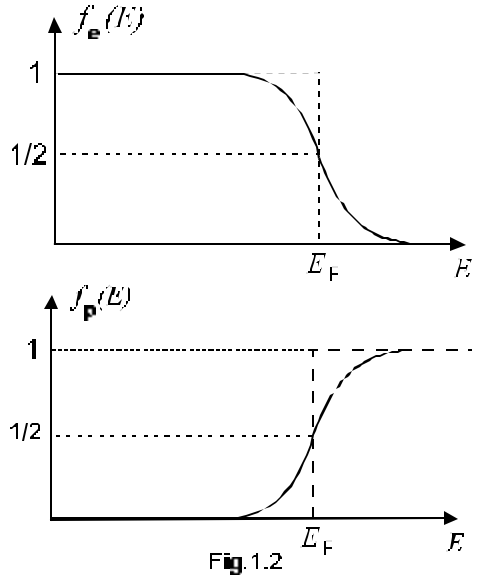


Fig.1.2

But the principle of energy conservation requires that $E_1+E_2=E_3+E_4$ and in this case only the function $A\,exp(\beta E)$ can be identified with $\left(\frac{1}{P(E)}-1\right)$, where $\beta=1/kT$ and $A=exp(-E_F/kT)$.

Then the probability to have an electron in the state characterised by energy E is:

$$P(E)=f_e(E)=\frac{1}{1+e^{\frac{E-E_F}{kT}}} \qquad 1.1.2.$$

and the probability to have an empty state is :

$$f_p(E) = 1 - f_e(E) = \frac{1}{1 + e^{\frac{E_F - E}{kT}}} \qquad 1.1.3.$$

The function ***f*** is known as the "**Fermi-Dirac distribution**" and is represented in fig.1.2.

At T=0 K, the shape of this function is like the shape of a "step function" (see dotted line) ;

At $T \neq 0K$ , for $E = E_F$, the probability to have an electron on this state is 1/2. The shape of Fermi-Dirac distribution for this temperature is represented by an continuous line.

The state characterised by $E = E_F$ is known as **"Fermi level"** and represents a virtual energy level characteristic for any solid state material. *This level is the upper limit of energy levels which can be filled with electrons at* **T=0 K.** (Only in the case of metals exist such situation. For isolators and semiconductors the upper limit been lower, as you will see in next paragraph)

Next problem, in solid state physics, is to obtain the formula for density of such energy states (the number of energy levels in the unit volume). In order to accomplish that, we must work in the "momentum space".

The quantum mechanics asserts that in a stationary state, an electron can be described by a stationary wave function. That means that in a bulk material having the characteristic dimension **L**, only electrons that have the associated wavelengths $\lambda$ verifying $L = n\frac{\lambda}{2}$ can exist, where **n** is a positive integer. This formula must hold on all three coordinates (x, y, z). But the wavelength is linked to the momentum (or impulse) **p** through "de Broglie" formula $\lambda = \frac{h}{p}$. Consequently we have the following relations between momentum (on each space direction) and the dimension of the bulk material:

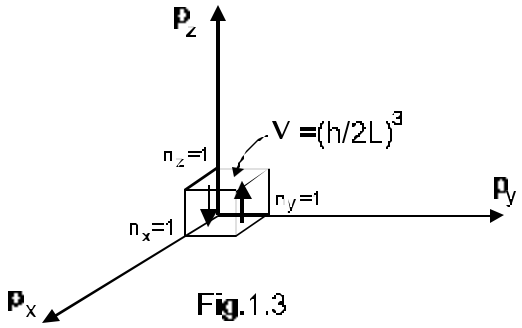


Fig.1.3

$p_x = \frac{n_x h}{2L}$ ; $p_y = \frac{n_y h}{2L}$ ; $p_z = \frac{n_z h}{2L}$. For the unit cell in the space of moments ($n_x = n_y = n_z = 1$), with the volume $\left(\frac{h}{2L}\right)^3$, we can have two states (Pauli's Principle), represented in Fig.1.3 by two arrows (spin quantum number $\pm\frac{1}{2}$. )

Consequently, the density of electrons' states in the unit cell will be:

$$d = \frac{2}{\left(\frac{h}{2L}\right)^3} = 2\left(\frac{2L}{h}\right)^3 \qquad 1.1.4.$$

In this case we can calculate which is the number of electrons which have the momentum lying between *p* and *p+dp* , using the Fig.1.4 which represents only the positive region of the space of momentum, because all components of the electron's momentum are positive.

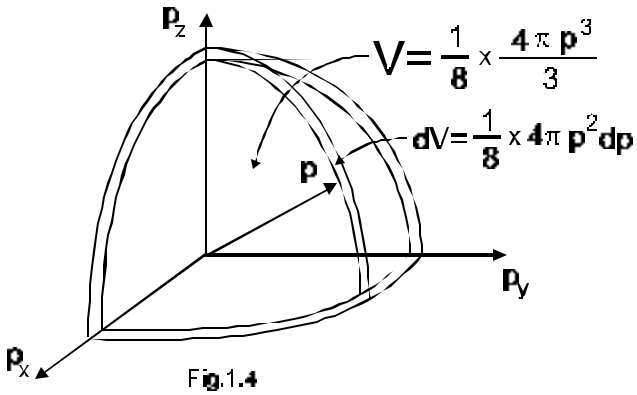



Fig.1.4

$$dn_{p,p+dp} = d \times dV = 2\left(\frac{2L}{h}\right)^3 \times \frac{1}{8} 4\pi p^2 dp$$

Starting from this formula, we can find the density of electronic states which have the momentum between *p* and *p+dp*

$$N(p)dp = \frac{dn_{p,p+dp}}{L^3} = \frac{8\pi p^2 dp}{h^3}$$

1.1.5.

But the kinetic energy for the quasi-free electron can be written as $E_k = \frac{p^2}{2m}$ . Therefore, we can rewrite the 1.4 formula as the density of electronic states which have the energy between E and E+dE

$$N(E_k) = \frac{4\pi}{h^3}\left(2m_n^*\right)^{3/2} E_k^{1/2} \qquad 1.1.6.$$

where $m^*_n$ is the effective mass of the electron.

A similar formula can be found for holes (the density of unfilled electronic states that have the energy between E and E+dE

$$N_p(E_k) = \frac{4\pi}{h^3}\left(2m_p^*\right)^{3/2} E_k^{1/2} \qquad 1.1.7.$$

where $m^*_p$ is the effective mass of the hole.

**1.2. The Density of Charge Carriers in a Pure Semiconductor.**

In an pure semiconductor, as we mention in the introduction, we can represent the energy states of an electron or hole using the model of energy bands. Let's consider for example the case of the Germanium crystalline lattice. As can be seen in Fig.1.5, the bounded valence electrons are in the Valence Band, characterised by the upper energy level $\mathbf{E_V}$, but can exist too in an excited state in Conduction Band, characterised by the lower energy level $\mathbf{E_C}$. In the Conduction Band the electrons are not bounded to the atom and they can have an moving through the crystalline lattice from atom to atom. The same thing can be done by the hole, which represents the empty state which remain in the Valence Band after the jump of the electron from Valence Band to the Conduction Band by thermal excitation.

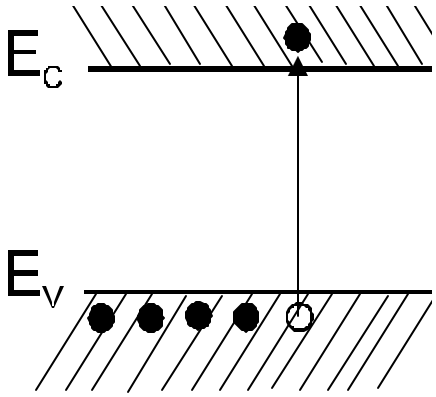


Fig.1.5

But for electrons $E_k = E - E_c$ and for holes $E_k = E_v - E$.

Now we can compute the density of charge carriers in CB or in VB, using next formulas:

$$n_e = n = \int_{E_C}^{+\infty} f_e(E)N(E)dE \qquad 1.2.1.$$

or, for holes,

$$n_p = p = \int_{-\infty}^{E_V} f_p(E)N_p(E)dE \qquad 1.2.2.$$

We will compute the density of charge carriers using the 1.2.1, 1.2.3, 1.1.6 and 1.1.7 substituted into 1.2.1 and 1.2.2:

$$n = \int_{E_c}^{\infty} \frac{4\pi \cdot \left(2m_n^*\right)^{3/2}}{h^3} (E - E_c)^{1/2} \frac{1}{1 + e^{\frac{E - E_F}{kT}}} dE$$

and for holes:

$$p = \int_{-\infty}^{E_v} \frac{4\pi \left(2m_p^*\right)^{3/2}}{h^3} (E_v - E)^{1/2} \frac{1}{1 + e^{\frac{E_F - E}{kT}}} dE$$

In order to be able to integrate the above equations, we have to make certain approximations, allowed by the typical environment conditions. For instance, at room temperature, T=300°K, the index of exponential from the denominator of Fermi-Dirac distribution is very high, and in this case we can neglect the factor 1 from the denominator. In this situation the Fermi-Dirac distribution becomes for electrons:

$f_e(E) = \frac{1}{1 + e^{\frac{E - E_F}{kT}}} \approx e^{-\frac{E - E_F}{kT}}$, while for holes it becomes $f_p(E) = \frac{1}{1 + e^{\frac{E_F - E}{kT}}} \approx e^{-\frac{E_F - E}{kT}}$

Now we can use the next mathematical artifice

$e^{-\frac{E - E_F + E_c - E_c}{kT}} = e^{-\frac{E_c - E_F}{kT}} \times e^{-\frac{E - E_c}{kT}}$, where the first term is a constant for the semiconductor material, that can be moved out of the integral. Finally the expression for density of electrons in Conduction Band is :

$$n = \frac{4\pi \left(2m_n^*\right)^{3/2}}{h^3} e^{-\frac{E_c - E_F}{kT}} \int_{E_c}^{\infty} (E - E_c)^{1/2} e^{-\frac{E - E_c}{kT}} dE \qquad 1.2.3.$$

and, correspondingly, the density of holes we will be:

$$p = \frac{4\pi \left(2m_p^*\right)^{3/2}}{h^3} e^{-\frac{E_F - E_v}{kT}} \int_{-\infty}^{E_v} (E_v - E)^{1/2} e^{\frac{E_v - E}{kT}} dE \qquad 1.2.4.$$

By making a change of variable in both integrals,

$x^2 = \frac{E - E_c}{kT}$ or $x^2 = \frac{E_v - E}{kT}$, the expression for electron density becomes

$$n = \frac{4\pi \cdot \left(2m_n^*\right)^{3/2}}{h^3} e^{-\frac{E_c - E_F}{kT}} (kT)^{3/2} \int_0^{\infty} 2x^2 e^{-x^2} dx$$

correspondingly, the expression for holes density becomes:

$$p = \frac{4\pi \left(2m_p^*\right)^{3/2}}{h^3} e^{-\frac{E_F - E_v}{kT}} (kT)^{3/2} \int_{-\infty}^{0} 2x^2 e^{-x^2} dx$$

Now both integrals can be computed by the parts method, as shown below:

$$\int_0^{\infty} 2x^2 e^{-x^2} dx = -xe^{-x^2}\Big|_0^{\infty} - \int_0^{\infty} \left(-e^{-x^2}\right) dx = \frac{\sqrt{\pi}}{2} \qquad 1.2.5.$$

where the last integral is half of the Poisson integral $\int_{-\infty}^{+\infty} e^{-x^2} dx = \sqrt{\pi}$

Now we have the final expression for both densities of charge carriers if we introduce the 1.2.5. equation in 1.2.3. and 1.2.4. equations:

$$n = \frac{2\left(2\pi m_n^* kT\right)^{3/2}}{h^3} e^{-\frac{E_c - E_F}{kT}} = N_C e^{-\frac{E_c - E_F}{kT}} \qquad 1.2.6.$$

respectively

$$p = \frac{2\left(2\pi m_p^* kT\right)^{3/2}}{h^3} e^{-\frac{E_F - E_v}{kT}} = N_V e^{-\frac{E_F - E_v}{kT}} \qquad 1.2.7.$$

The constants $N_c$ and $N_v$ are so called “density of energy states” in C.B., respectively in V.B.

$$N_C < N_V \quad \text{because} \quad m_n^* < m_p^*$$

**The position of Fermi level in pure semiconductors.**

In an pure semiconductor, the density of the two types of charge carriers is the same $n=p$ because these carriers are generated by thermal excitation from Valence Band to Conduction Band, as shown in Fig.1.5.

Thus, we can write the following equality:

$$N_C e^{-\frac{E_C - E_F}{kT}} = N_V e^{-\frac{E_F - E_V}{kT}} \qquad 1.2.8.$$

Now, we can transform the equation 1.2.8. into:

$$\frac{N_V}{N_C} = e^{\frac{E_F - E_v - E_c + E_F}{kT}},$$

to which we can apply the logarithm and then, with a simple operation we can extract the value of Fermi level:

$$E_F = \frac{E_c + E_v}{2} + \frac{kT}{2} \ln \frac{N_v}{N_c} \qquad 1.2.9.$$

The equation shows that the Fermi level is in the middle of forbidden band, at T=0$^0$K. If temperature is increasing, the Fermi level shifts towards the Conduction Band (see figure 1.6)

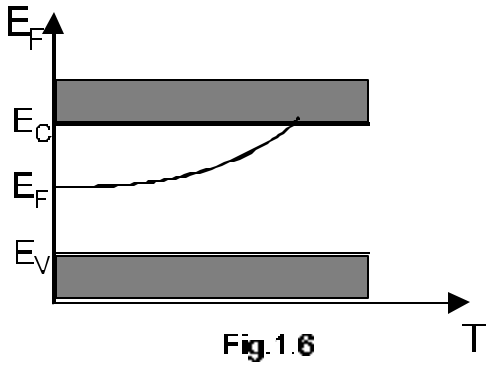


Fig.1.6

The most important equations, valid in any semiconductor are the "law of charge conservation". Based on equation 1.2.8. we can prove that the product of density charge carriers is a constant of semiconductor material, because this product does not depend on Fermi level. This product is the so called "pure density" :

$$n_i^2 = n \times p = N_C N_V e^{-\frac{E_c - E_v}{kT}} \qquad 1.2.10.$$

For the most common semiconductor materials, at room temperature, the values of this constants are:

Germanium: $N_C = 1.04 \times 10^{19}\,cm^{-3}$; $N_V = 6 \times 10^{18}\,cm^{-3}$; $n_i^2 = 2.4 \times 10^{26}\,cm^{-6}$

Silicon: $N_C = 2.8 \times 10^{19}\,cm^{-3}$; $N_V = 1.4 \times 10^{19}\,cm^{-3}$; $n_i^2 = 2 \times 10^{20}\,cm^{-6}$

### 1.3. Extrinsic Semiconductors (Doped Semiconductors).

**p doped Semiconductors.**

If in a material like Silicon or Germanium we introduce atoms like Al, Ga or In , which are atoms from the III [rd] group of Mendeleev's Table , the ionisation potential of this atoms will dramatically decrease. This effect is explained by the dependence of ionisation potential by the $1/\varepsilon_r^2$, where $\varepsilon_r$ is the relative dielectric constant of the medium in which are these atoms, respectively the relative dielectric constant of Germanium or Silicon.

For this materials the relative dielectric constant is $\varepsilon_{r_{Si}} = 12$ and respectively $\varepsilon_{r_{Ge}} = 16$ . The ionisation potential for such atoms from III group of Mendeleev's Table, inserted in Ge or Si, is given in Table 1.

**Table 1**

| | B | Al | Ga | In |
|---|---|---|---|---|
| Si | 0.045eV | 0.057eV | 0.065eV | 0.16eV |
| Ge | 0.0104eV | 0.0102eV | 0.0108eV | 0.0112eV |

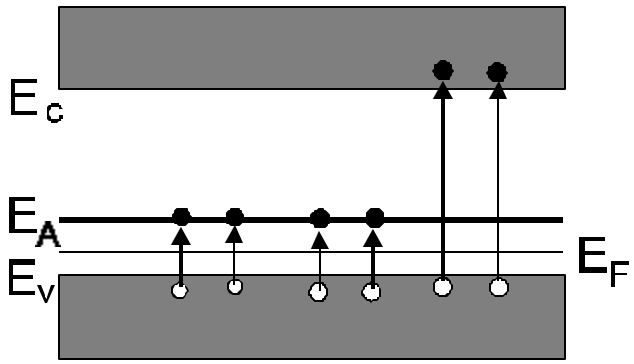



Fig.1.7
Figure legend: ● electron ○ hole
transition; $E_c$ lower level of CB;
$E_v$ upper level of VB; $E_a$ acceptor level;
$E_F$ Fermi level

These energies are represented in the model of band energies by the existence of an acceptor energy level, very close to Valence Band (distance between this acceptor energy level and the upper energy level of Valence Band - $E_v$, is the ionisation energy of impurity atoms) as in figure 1.7.

Arrows indicate transitions of electrons from the Valence Band to acceptor level or to Conduction Band. Because the acceptor level is closer to Valence Band than Conduction Band, the probability to have such a transition in acceptor level is higher than the probability to have such a transition to Conduction Band. For that we will have more electrons on acceptor level than in the Conduction Band, but all these electrons are bounded electrons, ionising the

acceptor impurities. They do not participate to conduction phenomena, but holes, generated by such transitions can participate to conduction phenomena and they are more than the electrons from Conduction Band. The holes are "*majority charge carriers*". For this reason we named these semiconductors "*P semiconductors*". The probability to have an electron on acceptor level has the same form like Fermi-Dirac distribution for electrons in Conduction Band, if we did not take into account the degeneracy factor

$$f_A^-(E) = \frac{1}{1 + e^{\frac{E_A - E_F}{kT}}}$$

then the density of ionised acceptors will be:

$$N_A^- = N_A f_A^-(E) \approx N_A e^{-\frac{E_A - E_F}{kT}} \qquad 1.3.1.$$

and the density of holes obtained by the phenomena of such ionisation will be:

$$p = N_V e^{-\frac{E_F - E_v}{kT}} = N_A^-$$

From this equality we can find the position of Fermi level in the P semiconductor

$$N_V e^{-\frac{E_F - E_V}{kT}} = N_A e^{-\frac{E_A - E_F}{kT}} \qquad 1.3.2.$$

thus, by using the same procedure we applied for pure semiconductors, we find :

$$E_{F_p} = \frac{E_A + E_V}{2} + \frac{kT}{2} \ln \frac{N_V}{N_A} \qquad 1.3.3.$$

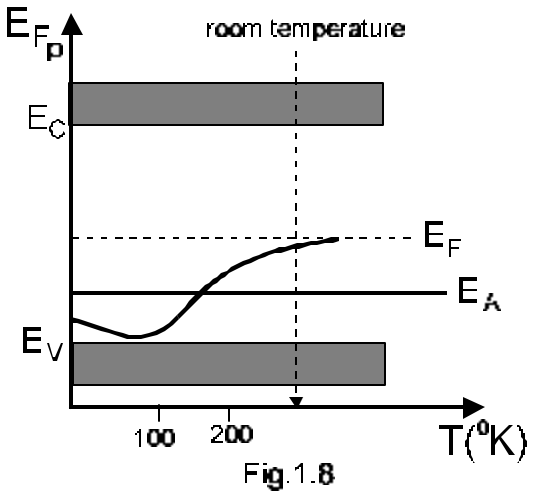


Fig.1.8

This formula shows us than at $T=0^0K$ the Fermi level is at the middle of the distance between acceptor level and upper level of Valence Band. If the temperature is increasing the Fermi level shifts to the middle of the Forbidden Bend (see Figure 1.8), if $N_A<N_V$, as in the case of temperature higher than $100^0K$.

From Fig.1.8 we conclude that at room temperature all impurities are ionised. That means that the density of "*majority carriers*" is

$$p_p \approx N_A$$

At this temperature we have *minority charge carriers* too, generated by band to band transitions of valence electrons (see Fig.7). The density of these carriers can be calculated with the help of 1.2.10. equation

Therefore $n_p \approx \frac{n_i^2}{N_A}$ 1.3.4.

**n Doped Semiconductors.**

The same phenomenon of decreasing of ionisation potential is happening in an pure semiconductor doped with elements from V $^{th}$ group of Mendeleev's Table, like P, As, St . In the Table 2 you can see the modified ionisation potentials of such impurities.

**Table 2**

| | P | As | Sb | Bi |
|---|---|---|---|---|
| Si | 0.045eV | 0.049eV | 0.039eV | 0.067eV |
| Ge | 0.012eV | 0.0127eV | 0.0096eV | - |

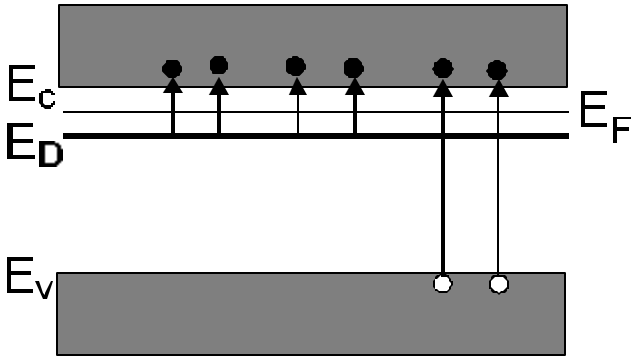



Fig.1.9

Figure legend: ● electron ○ hole ↑ jump

The diagram of energy bands of such semiconductor is shown in Figure 1.9. In this case a donor level represents the energy of ionisation. In this case the "*majority carriers*" are electrons because the probability of a jump from donor level is higher than the probability of a band to band jump.

The probability of ionisation of an impurity will be similar with Fermi-Dirac distribution, if we did not take into account the degeneracy factor

$$f_D^+(E) = \frac{1}{1+e^{\frac{E_F-E_D}{kT}}}$$

therefore the density of ionised donors is

$$N_D^+ = N_D f_D^+(E) \approx N_D e^{-\frac{E_F-E_D}{kT}} \qquad 1.3.5.$$

and the density of electrons obtained by the phenomena of such ionisation will be:

$$n = N_C e^{-\frac{E_c-E_F}{kT}} = N_D^+ \qquad 1.3.6.$$

From this equality we can find the position of Fermi level in the P semiconductor

$$N_C e^{-\frac{E_c-E_F}{kT}} = N_D e^{-\frac{E_F-E_D}{kT}}$$

then, using the same procedure like in the case of pure semiconductors we will find :

$$E_{F_n} = \frac{E_D+E_c}{2} - \frac{kT}{2}\ln\frac{N_C}{N_D} \qquad 1.3.7.$$

This equation shows us than at T=$0^0$K the Fermi level is at the middle of the distance between donor level and lower level of Conduction Band. If the temperature is increasing upper to

100 $^0$K, the Fermi level shifts to the middle of Forbidden Band because $N_c$ becomes higher than $N_D$. (see Figure 1.10).

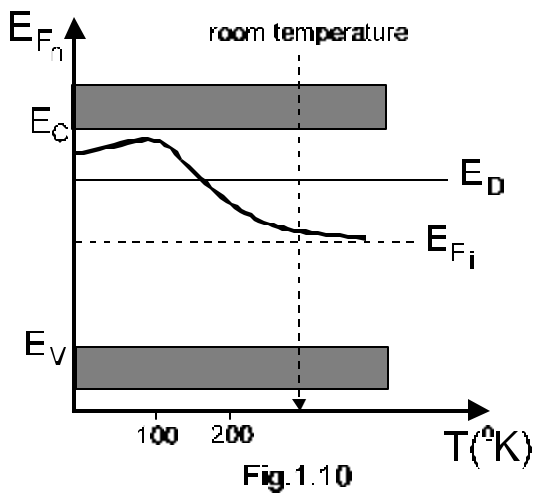


Fig.1.10

By looking at Fig.1.10 we can see that at room temperature, practically all impurities are ionised. That means that the density of "*majority carriers*" is

$$n_n \approx N_D$$

But at this temperature we have too, *minority* carriers generated by band to band jumps of valence electrons (see Fig.1.9). The density of these carriers can be calculated using the 1.2.10. equation. Therefore,

$$p_n \approx \frac{n_i^2}{N_D} \qquad 1.3.8.$$

**1.4. Physical Phenomena in Semiconductors.**

**Conduction.** Different from metals, in semiconductors two different kinds of charge carriers participate to conduction phenomena: negative charge carriers (electrons) and positive charge carriers (holes). In the presence of an electric field both charge carriers will move to the direction of this field (electrons in the opposite way and holes in the same way of the field). Then, in a semiconductor we will have two components of the current density:

$$j_n = env_n = en\mu_n E \qquad 1.4.1.$$

$$j_p = epv_p = ep\mu_p E \qquad 1.4.2.$$

where, $v_n = \mu_n E$ and $v_p = \mu_p E$, are the drift velocities of charge carriers. These velocities are proportional to the intensity of electric field $\boldsymbol{E}$, the constant of proportionality representing "the mobility" of the charge carrier, $\mu$.

Then the total current density can be written as the sum of both components, given by equations 1.4.1 and 1.4.2:

$$j_{tot} = j_n + j_p = en_n\mu_n E + ep_n\mu_p E \qquad 1.4.3.$$

Using equation 1.4.3. we can find the expression for electrical conductivity of the semiconductor material:

$$j_{tot} = \sigma E \quad \text{then,} \quad \sigma = \frac{j_{tot}}{E} = e\left(n\mu_n + p\mu_p\right) \qquad 1.4.4$$

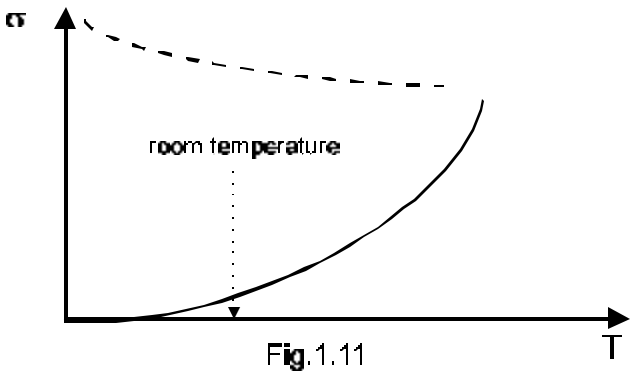


Fig.1.11

An pure semiconductor has $n = p = n_i$, therefore the dependence of σ function of temperature will have the same shape as intrinsic density of charge carriers function of temperature, as it is shown in Fig.1.11, if we neglect the temperature dependence of the mobility of charge carriers.

The dotted line represents the conductivity of metals. Obviously, this figure represents just a qualitative plot of the conductivity. From this plot we can see the difference between metals and semiconductors: in semiconductors, the conductivity increases exponentially with the temperature, and at room temperature the conductivity of semiconductors is lower then the conductivity of metals. This property is used in a number of passive devices, for instance in thermistors.

In the case of extrinsic semiconductors the equation 1.4.3. remains valid, but the density of charge carriers must be renamed in accordance with the semiconductor type.

In the case of N type semiconductors the total current will be:

$$j_{tot} = j_n + j_p = en_n\mu_n E + ep_n\mu_p E$$

and the conductivity in such semiconductors will be predominantly mediated by electrons, since $n_n >> p_n$ ; $n_n \approx N_D$. Therefore:

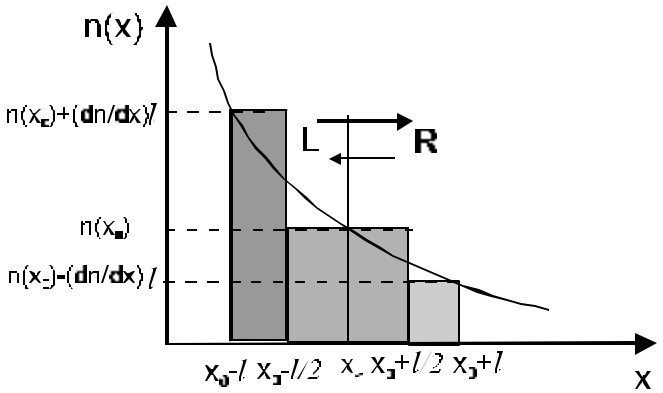
$$\sigma_N \approx eN_D\mu_n \qquad 1.4.5.$$

Correspondingly, for the P type semiconductors the total current will be:

$$j_{tot} = j_n + j_p = en_p\mu_n E + ep_p\mu_p E$$

and the conductivity in such semiconductors will be predominantly mediated by holes, since $p_p >> n_p$ ; $p_p \approx N_A$, therefore:

$$\sigma_P \approx eN_A\mu_p \qquad 1.4.6.$$




**Diffusion.** If there is a density gradient of charge carriers in a semiconductor's region (see Fig.1.12), the carriers in the densely populated region will tend to migrate towards the depleted areas. Therefore, a carrier diffusion current will occur. At thermal balance, the motion of charge carriers (in our example electrons) is random. Then, depending on the density of electrons on each side of the section through the semiconductor at $x_0$, the number of electrons which move through the plane at $x_0$, in the mean free time (the time between two collisions) will be different. The number of electrons passing from right to left, through the plane $x_0$ is:

$$N_{R\to L} = \frac{1}{2}\left[n(x_0 + l) + n(x_0)\right]\cdot\frac{l}{2}\times S \qquad 1.4.7.$$

whereas the number of electrons that pass from the right to the left of the same plane is,

$$N_{L\to R} = \frac{1}{2}\left[n(x_0 - l) + n(x_0)\right]\cdot \frac{l}{2}\times S \qquad 1.4.8.$$

where $l$ is mean free path, which it is assumed to be the same for both carriers, hence most of the collisions occur with the lattice, its defects or impurities, and therefore is independent of the carrier density. Factor $\frac{1}{2}$ is given by the equal probability for the movement from right to left or from left to right.

Then the total number of electrons, which pass through this plane, is the difference between equations 1.4.8 and 1.4.7:

$$N_T = N_{L\to R} - N_{R\to L} = \frac{1}{4}\left[n(x_0 - l) - n(x_0 + l)\right]\cdot l \times S = -\frac{1}{2}\left(\frac{dn}{dx}\right)_{x_0}\cdot l^2 \times S \qquad 1.4.9.$$

This movement of charge carriers creates a current that has the density

$$j_{D_n} = \frac{I}{S} = \frac{Q}{tS} = \frac{-eN_T}{tS} = e\frac{l}{t}\cdot\frac{l}{2}\frac{dn}{dx} = eD_n\frac{dn}{dx} \qquad 1.4.10.$$

where the constant $D$ is the so called “diffusion constant”. Using a similar demonstration we can find the current density of holes:

$$j_{D_p} = -eD_p\frac{dp}{dx} \qquad 1.4.11.$$

The minus sign is determined by the gradient of charge density, which is negative and, at the same time, by the charge of the hole $e$, which is taken as positive.

**Generation and Recombination of charge carriers.** The density of charge carriers can not build up indefinitely in time, because at the same time with the generation phenomena there are the recombination phenomena, which scale with the density of charge carriers. At thermal balance, the generation rate must equal the recombination rate.

The recombination rate is proportional to the product of the densities of charge carriers:

$$R = \text{Const.} \times n_0 p_0 \qquad 1.4.12.$$

In the case of P type semiconductor $p_0 \approx N_A$ ; $n_0 = n_{p_0}$ . Therefore equation 1.4.12 becomes:

$$R_n = \text{Const.} \times N_A n_p = \frac{n_{p_0}}{\tau_n} \qquad 1.4.13.$$

where $\tau_n$ is mean life time of minority carriers generated in excess.

In the case of N type semiconductor $n_0 \approx N_D$ ; $p_0 = p_{n_0}$ . Therefore equation 1.4.12 becomes:

$$R_p = \text{Const.} \times N_D p_n = \frac{p_{n_0}}{\tau_p} \qquad 1.4.14.$$

where $\tau_p$ is mean lifetime of minority carriers generated in excess.

Then, if we have an excess of minority carriers, let that be in a P type semiconductor, from any reasons, we can find the time evolution of this excess:

$dp_n = [G - R] \times dt$ but $G = \frac{p_{n_0}}{\tau_p}$ and $R = \frac{p_n(t)}{\tau_p}$ then this relation can be written

$\frac{dp_n}{dt} = -\frac{p_n(t) - p_{n_0}}{\tau_p}$ which is a first order differential equation, which has following solution:

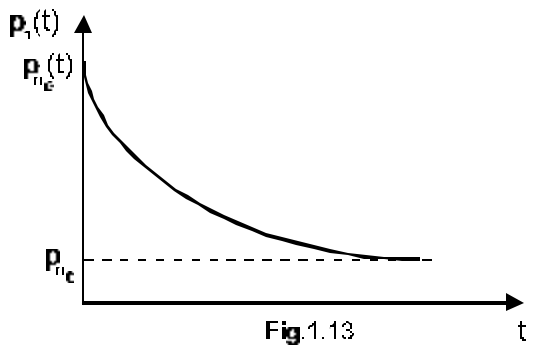


Fig.1.13

$$p_n(t) - p_{n_0} = [p_n(0) - p_{n_0}] \cdot e^{-\frac{t}{\tau_p}} \qquad 1.4.9.$$

The minority carrier excess has therefore an exponential decay in time, as shown in fig.1.13.

**1.5. Equation of continuity (Law of charge conservation).**

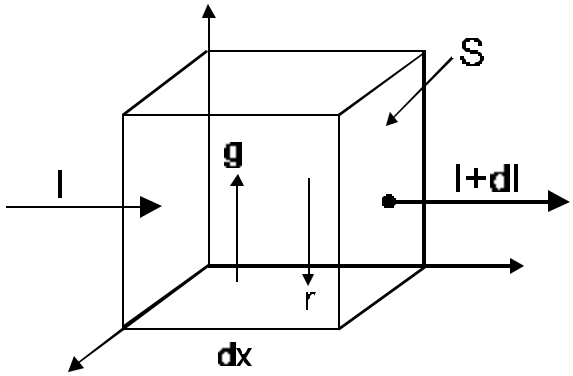


Fig.1.14

Let there be an elementary volume in a semiconductor (see Fig.1.14). Inside this elementary volume we may have generation phenomena, that takes place at a rate **g** and recombination phenomena, at a rate **r** . In this volume enters the current **I** and goes out the current **I+dI**.

In this case the balance equation for the time variation of total charge inside the elementary volume **Sdx** can be written as:

$$eSdx\frac{\partial p}{\partial t} = -\frac{eSdxp}{\tau_p} + eSdxg - dI \qquad 1.5.1.$$

In the stationary case $\frac{\partial p}{\partial t} = 0$ ; $dI = 0 \Rightarrow \frac{p_0}{\tau_p} = g$ , thus the equation 1.5.1. becomes :

$$\frac{\partial p}{\partial t} = -\frac{p - p_0}{\tau_p} - \frac{d}{dx}\left[\mu_p pE - D_p\frac{dp}{dx}\right] \qquad 1.5.2.$$

where we replace the total current by its two components (drift current and diffusion current)

$$I = j_t S = \left[ep\mu_p E - eD_p\frac{dp}{dx}\right]S$$

Therefore, the final form of 1.5.2. equation becomes:

$$\frac{\partial p}{\partial t} = -\frac{p - p_0}{\tau_p} - \mu_p\frac{d(pE)}{dx} + D_p\frac{d^2p}{dx^2} \qquad 1.5.6.$$

Equation 1.5.6. represents the balance equation for minority p carriers in a N type semiconductor, because the density of minority carriers is sensitive to accidental variation of

charge density. For that reason we can replace **p** by $\mathbf{p_n}$ . In a similar way we can find the equation for N type semiconductors:

$$\frac{\partial n_p}{\partial t} = -\frac{n_p - n_{p_0}}{\tau_n} - \mu_n \frac{d(n_p E)}{dx} - D_n \frac{d^2 n_p}{dx^2} \qquad 1.5.7.$$

The minus sign before the diffusion constant appears from the expression of diffusion current for electrons.

**Particular cases of continuity equation.** Let there be a semiconductor of P type. The first particular case is based on the following simplifying assumptions: independence of density to distance (x axis) and null electric field. Accordingly, in the equation 1.5.6 we have:

$\frac{\partial p_n}{\partial x} = 0$ ; $E = 0$ and the equation becomes:

$\frac{\partial p_n}{\partial t} = -\frac{p_n - p_{n_0}}{\tau_p}$ which has the known solution $p_n(t) - p_{n_0} = \left[p_n(0) - p_{n_0}\right] \cdot e^{-\frac{t}{\tau_p}}$ , similar with 1.4.15. equation, plotted in the Fig.1.13

The second particular case is: independence of the carrier density in time and null electric field:

$\frac{\partial p_n}{\partial t} = 0$ ; $E = 0$ and the equation 1.5.6. becomes:

$$D_p \frac{d^2 p}{dx^2} = \frac{p_n - p_{n_0}}{\tau_p}$$

which has the solution $p_n(x) - p_{n_0} = Ae^{\frac{x}{L_p}} + Be^{-\frac{x}{L_p}}$

where we define $L_p = \sqrt{D_p \tau_p}$ which represents the so called “diffusion length” .

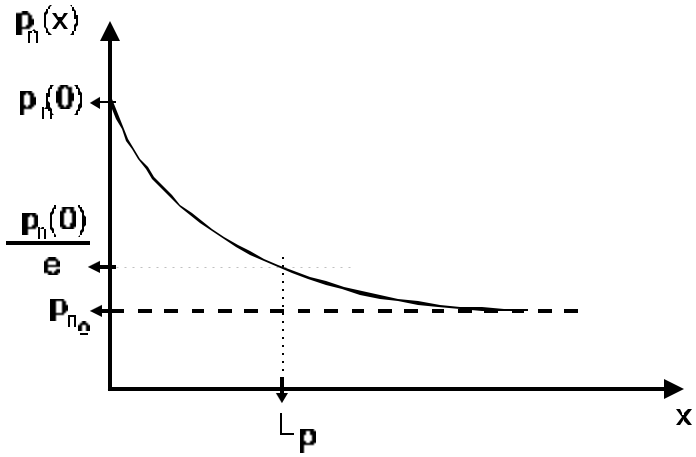


Fig.1.15

The constant A must be zero (since the carrier density cannot increase towards infinity with increasing x), therefore from boundary conditions we can find the value of constant B:

$$B = p_n(0) - p_{n_0}$$

and now we can write the final form of this solution:

$$p_n(x) - p_{n_0} = (p_n(0) - p_{n_0}) \cdot e^{-\frac{x}{L_p}}$$

that has the graphical representation plotted in fig. 1.15.

# Chapter 2. P - N JUNCTION.

### 2.1. Physical Phenomena in P-N Junction

The P - N junction is formed in a bulk semiconductor, which is considered to have the size larger then the diffusion length of charge carriers. Two different regions of doping are created in the structure, one of P type and other one of N type. The boundary between these two regions represents the P - N Junction.

Because this structure has a high gradient of majority charge carriers from P type semiconductor to N type semiconductor, diffusion phenomena will appear at the boundary between these two types of semiconductors. The majority carriers of P type, will diffuse to N type semiconductor whereas the majority carriers of N type will diffuse to P type semiconductor. But in N type semiconductor the holes are minority carriers, therefore a phenomenon of recombination between holes and electrons will occur. The same phenomena will occur in P type semiconductor between electrons and holes.

Following diffusion and recombination, in both sides of the junction, a “depletion layer” will occur due to massive recombination. At the same time, there’s going to be a net electrical charging in the region, because in these regions we will have only the fixed charges, the ion charges. In this region an internal electrical field will appear and, of course, a voltage gradient (see Figure 2.1). In plot a) we plotted the charge density in depletion layer; in plot b) we plotted the intensity of internal electric field function of distance; in plot c) we plotted the voltage gradient function of the distance.

As we can see from plot (a), the charge conservation law can be written as:

$$eN_A L_n S = eN_D L_p S \qquad 2.1.1.$$

which can be further reduced to:

$$N_A L_n = N_D L_p \qquad 2.1.2.$$

To find the expression of electric field in the depletion layer of P type semiconductor we must apply the Gauss law for the any S surface perpendicular to the positive x axis.

$$E(-x)S = \frac{eN_A(-L_n - x)S}{\varepsilon}$$

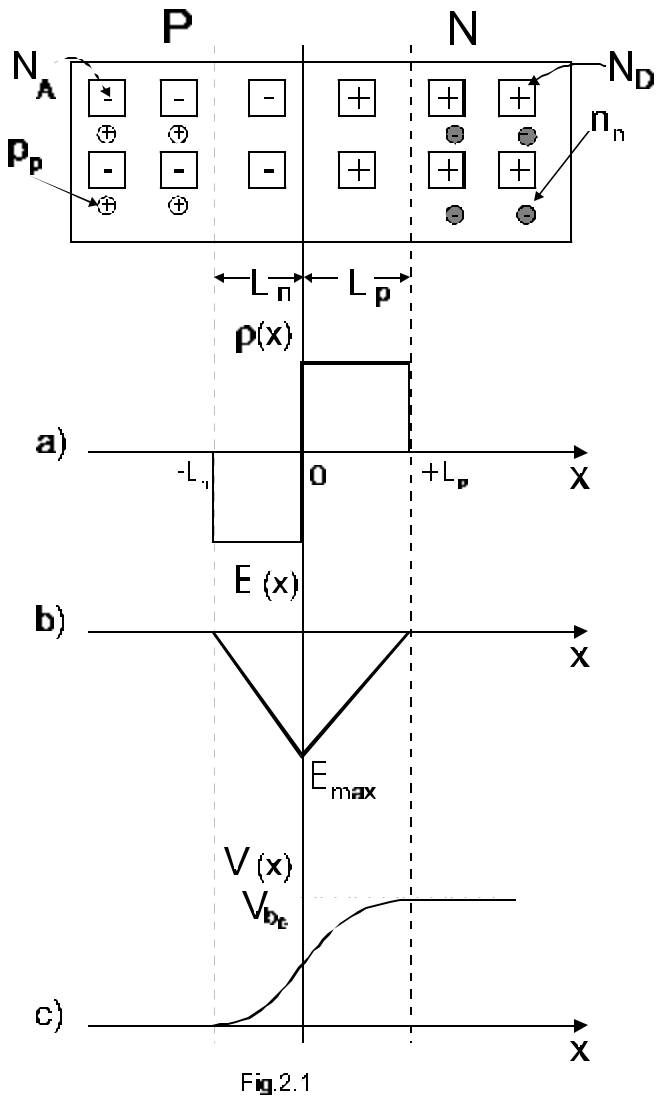



Fig.2.1

Figure legend: ⊖ negative ion: ⊕ positive ion:
○ hole: ● quasi-free electron

resulting in:

$$E(-x) = -\frac{eN_A(L_p + x)}{\varepsilon}$$

2.1.3a.

In the same way we can obtain the equation for the electric field in the direction of the negative x axis.:

$$E(-x) = -\frac{eN_D(L_p - x)}{\varepsilon}$$

2.1.3b.

From the last two equations we can obtain the value of maximum electric field:

$$E_{max} = -\frac{eN_A L_n}{\varepsilon} = -\frac{eN_D L_p}{\varepsilon} = E(0)$$

2.1.3c.

The value of the barrier potential can be obtained simply by integrating the electric field over the length of the junction:

$$V_{b_0} = -\int_{-L_n}^{+L_p} E(x)dx = -\frac{E_{max}(L_p + L_n)}{2}$$

2.1.4.

From equations 2.1.3c. and 2.1.2. we will obtain the final formula for the barrier potential:

$$V_{b_0} = \frac{eN_D L_p(L_p + L_n)}{2\varepsilon} = \frac{eN_A L_n(L_p + L_n)}{2\varepsilon}$$

2.1.5.

All these formulas are calculated at thermal balance.

Relation 2.1.5. is used to determine the diffusion length of majority charge carriers which diffuse in the region where they become minority charge carriers :

$$V_{b_0} = \frac{eN_D L_p^2 (1 + \frac{L_n}{L_p})}{2\varepsilon} = \frac{eN_A L_n^2 (\frac{L_p}{L_n} + 1)}{2\varepsilon} \qquad 2.1.6.$$

From relations 2.1.6. and 2.1.2. we will obtain the final formula for the diffusion length:

$$L_p = \left[ \frac{2\varepsilon}{eN_D \left( 1 + \frac{N_D}{N_A} \right)} \right]^{\frac{1}{2}} V_{b_0}^{\frac{1}{2}} \qquad 2.1.7a.$$

respectively,

$$L_n = \left[ \frac{2\varepsilon}{eN_A \left( 1 + \frac{N_A}{N_D} \right)} \right]^{\frac{1}{2}} V_{b_0}^{\frac{1}{2}} \qquad 2.1.7b.$$

The diffusion of charge carriers will continue until the electric field created by this charge displacement will build up to a value that will completely stop the charges on crossing the junction. Once this equilibrium has been attained, the total current of holes, or electrons, will be zero (assuming also thermal balance):

$$j_{p_t} = j_{p_c} + j_{p_d} = ep\mu_p E - eD_p \frac{dp}{dx} = 0 \qquad 2.1.8a$$

$$j_{n_t} = j_{n_c} + j_{n_d} = en\mu_n E + eD_n \frac{dn}{dx} = 0 \qquad 2.1.8b.$$

If we replace the electrical field with the voltage gradient $\left( E = -\frac{dV_b}{dx} \right)$ we can integrate these formulas. Let take as an example the formula 2.1.8a. :

$$p\mu_p\left(-\frac{dV_b}{dx}\right) = D_p\frac{dp}{dx} \quad , \qquad \text{therefore:} \qquad \frac{dp}{p} = -\frac{\mu_p}{D_p}dV_b$$

The latter differential equation has the following solution:

$$\ln p = -\frac{\mu_p}{D_p}V_b + \text{Const.} \qquad 2.1.9.$$

From the boundary conditions, we will find the value of integration constant:

at $V_b = 0$ $\quad \text{Const.} = \ln p_p \quad$ and at $V_b = V_{b_0}$ $\quad p = p_n$ ,

then equation 2.1.9. becomes:

$$p_n = p_p e^{-\frac{\mu_p}{D_p}V_{b_0}} \qquad 2.1.10.$$

But at room temperature we have the following relations, presented earlier in this chapter, for the density of charge carriers:

$p_n = N_v e^{-\frac{E_{F_n} - E_{V_n}}{kT}}$ ; $p_p = N_v e^{-\frac{E_{F_p} - E_{v_p}}{kT}}$ where, at thermal equilibrium, $E_{F_n} = E_{F_p}$. Then, the ratio between majority carriers and minority carriers can be written as:

$$\frac{p_p}{p_n} = e^{\frac{E_{V_p} - E_{V_n}}{kT}} \qquad 2.1.11.$$

This ratio can be obtained from equation 2.1.10. too, but in order to have an equality between these two ratios it is required that:

$$\frac{\mu_p}{D_p}V_{b_0} = \frac{E_{V_p} - E_{V_n}}{kT} = \frac{eV_{b_o}}{kT} \qquad 2.1.12.$$

From equation 2.1.12. we will obtain the following relations between carrier mobility, diffusion constant, and temperature:

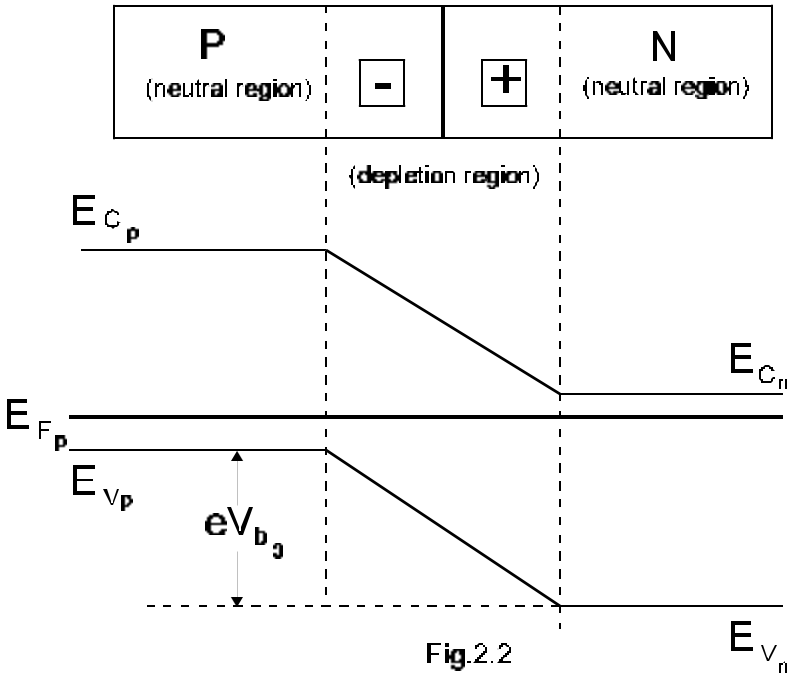



Fig.2.2

$$\frac{D_p}{\mu_p} = \frac{kT}{e} \qquad 2.1.13a.$$

while for electrons, in a similar way,

$$\frac{D_n}{\mu_n} = \frac{kT}{e} \qquad 2.1.13b.$$

Relations 2.1.13. are the so called "Einstein's relations" for semiconductors. The next conclusion extracted from equation 2.1.10. is $eV_{b_0} = E_{V_p} - E_{V_n}$ , which shows us that in a such structure ( PN junction) the energy bands of semiconductor are broken or shifted at the level of the junction, in order to have the same Fermi level on both sides of the junction (see Figure 2.2) at thermal balance.

From equations 2.1.11. and 2.1.12. we can obtain the formula for the barrier potential (which matches the maximum of the voltage gradient):

$$\frac{p_p}{p_n} = e^{\frac{eV_{b_0}}{kT}} \Rightarrow V_{b_0} = \frac{kT}{e}\ln\frac{p_p}{p_n} = \frac{kT}{e}\ln\frac{N_A N_D}{n_i^2} \qquad 2.1.14.$$

The ratio $\frac{kT}{e} = V_T$ is the so called "thermal potential", and at room temperature has the value:

$V_T = 0.026$ Volt $\left(T = 300^0 K\right)$.

## 2.2. The Current-Voltage Characteristic for the P - N Junction.

The density of charge carriers must be a continuos function over the whole length of the semiconductor (see figure 2.3). We also determined, in chapter 2.1., that the diffusion length of majority charge carriers, which diffuse through the junction, depends on the square root of the barrier potential. Then, if we change this barrier potential by applying an external potential

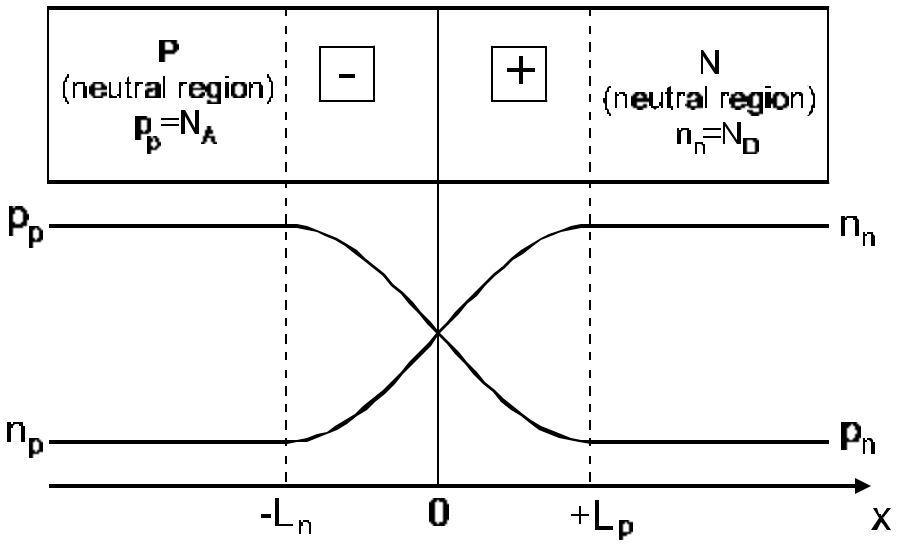



Fig.2.3

(voltage), we will modify these lengths and the height of barrier potential: $L_{p,n} = KV_b^{1/2}$ , where the height of barrier potential is $V_{b_0} = \frac{kT}{e} \ln \frac{N_A N_D}{n_i^2}$ , at thermal balance. But the height of barrier potential can be written as $V_b = V_{b_0} - V_{ext}$, where the convention for external potential is as follows:

$V_{ext} = -V_{ext}$ if this is the so called "**reverse biasing potential/voltage**" (the positive electrode of external source is connected to the N type semiconductor);

$V_{ext} = +V_{ext}$ if this is the so called "**forward biasing potential/voltage**" (the positive electrode of external source is connected to the P type semiconductor).

The changes induced by the external voltage to the length of “depletion region”, the height of “barrier potential” and the currents which flows through the junction are presented in Fig. 2.7.

Now, as you can see in figure 2.4, we can not change the dependence of density charge carriers function on the x axis. Then, at forward polarisation, at the new diffusion length $L_p$ there will be an injection of minority charge carriers in the N type semiconductor, resulting in a carrier density different from the one at thermal balance (a similar process will take place in the P type semiconductor).

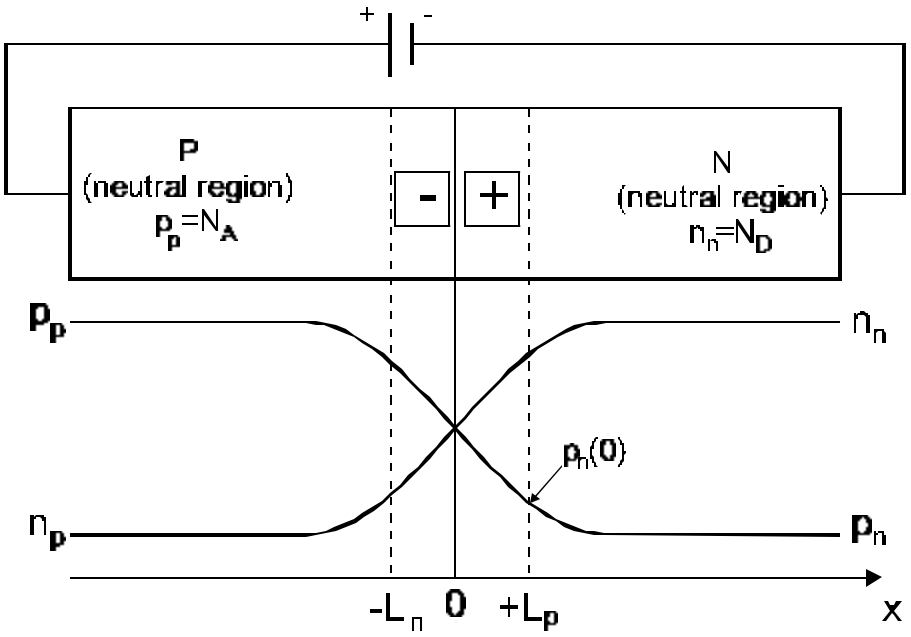


Fig.2.4

Now, if we take the origin of X axes on the boundary of depletion region, we will be in the conditions of the “equation of continuity” particular case $E = 0$ ; $\frac{\partial p}{\partial t} = 0$ , where the solution is:

$$p_n(x) - p_{n_0} = [p_n(0) - p_{n_0}] \cdot e^{-\frac{x}{L_p}} \qquad 2.2.1.$$

Then the diffusion current that will be established can be written as:

$$j_p(x) = -eD_p \frac{dp}{dx} = \frac{eD_p [p_n(0) - p_{n_0}] \cdot e^{-\frac{x}{L_p}}}{L_p} \qquad 2.2.2$$

taking into account the equation 2.2.1.

A similar expression will result for the electron component of the current. But the density of minority charge carriers can also be written as:

$$p_n(0) = p_p e^{-\frac{eV_b}{kT}} = p_p e^{-\frac{e(V_{b0} - V_{ext})}{kT}} = p_{n_0} e^{\frac{eV_{ext}}{kT}} \qquad 2.2.3.$$

Therefore the final formula for the current given by holes being:

$$j_p(x) = \frac{eD_p}{L_p}\left[p_{n_0} e^{\frac{eV_{ext}}{kT}} - p_{n_0}\right] \cdot e^{-\frac{x}{L_p}} \qquad 2.2.4$$

while for electrons it will be:

$$j_n(x) = \frac{eD_n}{L_n}\left[n_{p_0} e^{\frac{eV_{ext}}{kT}} - n_{p_0}\right] \cdot e^{-\frac{x}{L_n}} \qquad 2.2.5$$

The total current is the sum of 2.2.4 and 2.2.5. But this current does not depend of the X abscissa. Then we can compute this current at x=0

$$j_t = \text{const.} = j_p(x) + j_n(x)\Big|_{x=0} = \left[\frac{eD_p p_{n_0}}{L_p} + \frac{eD_n n_{p_0}}{L_n}\right]\left(e^{\frac{eV_{ext}}{kT}} - 1\right) \qquad 2.2.6$$

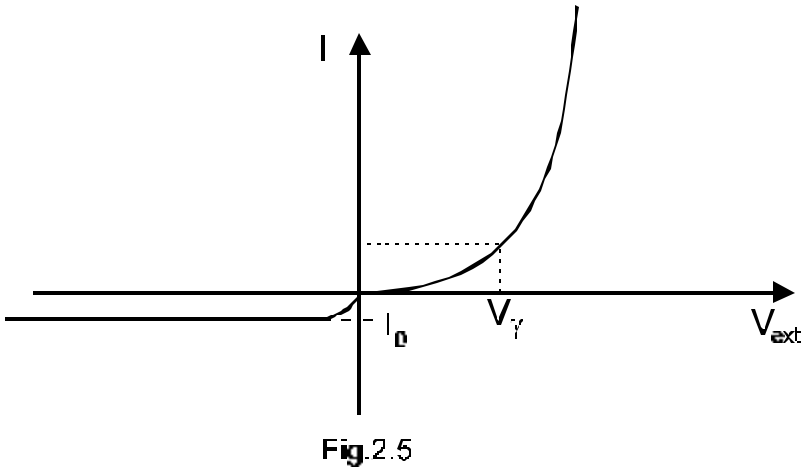



Fig.2.5

If we multiply the equation 2.2.6 with the cross area of the junction, we will obtain the current-voltage characteristic or so called “Volt-Ampere characteristic” of the ideal diode:

$$I = I_0\left(e^{\frac{eV_{ext}}{kT}} - 1\right) \qquad 2.2.7.$$

The plot of relation 2.2.7 is given in Fig 2.5, where the negative current axis has been magnified several orders of magnitude with respect with the positive axis.

The current $I_0$ is the so called the “saturation current” or “reverse current” through the junction. The value of this current depends on the parameters of the crystalline lattice and temperature. For Silicon, values of nanoampers are common, and for Germanium lattice, values of microampers are common for the saturation current. This current is mediated by the minority carriers, and its expression can be determined from equation 2.21. :

$$I_0 = \left( \frac{eD_p p_{n_0}}{L_p} + \frac{eD_n n_{p_0}}{L_n} \right) S \qquad 2.2.8.$$

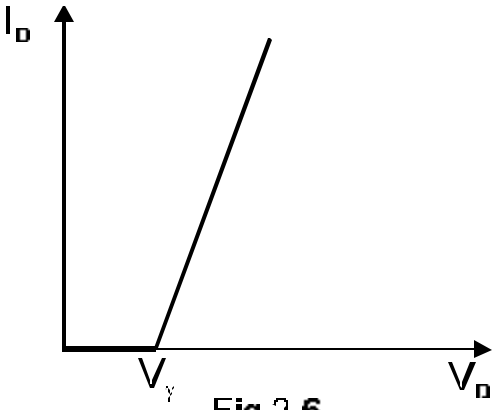


Fig.2.6

The opening potential of the diode, $V_\gamma$, is defined as the forward biasing voltage for which the current is $1\mu A$.

In practical applications and circuit analysis, the plot of current function of forward biasing voltage is approximated by a linear dependence of the voltage on current above a threshold, as you can see in Fig.2.6.

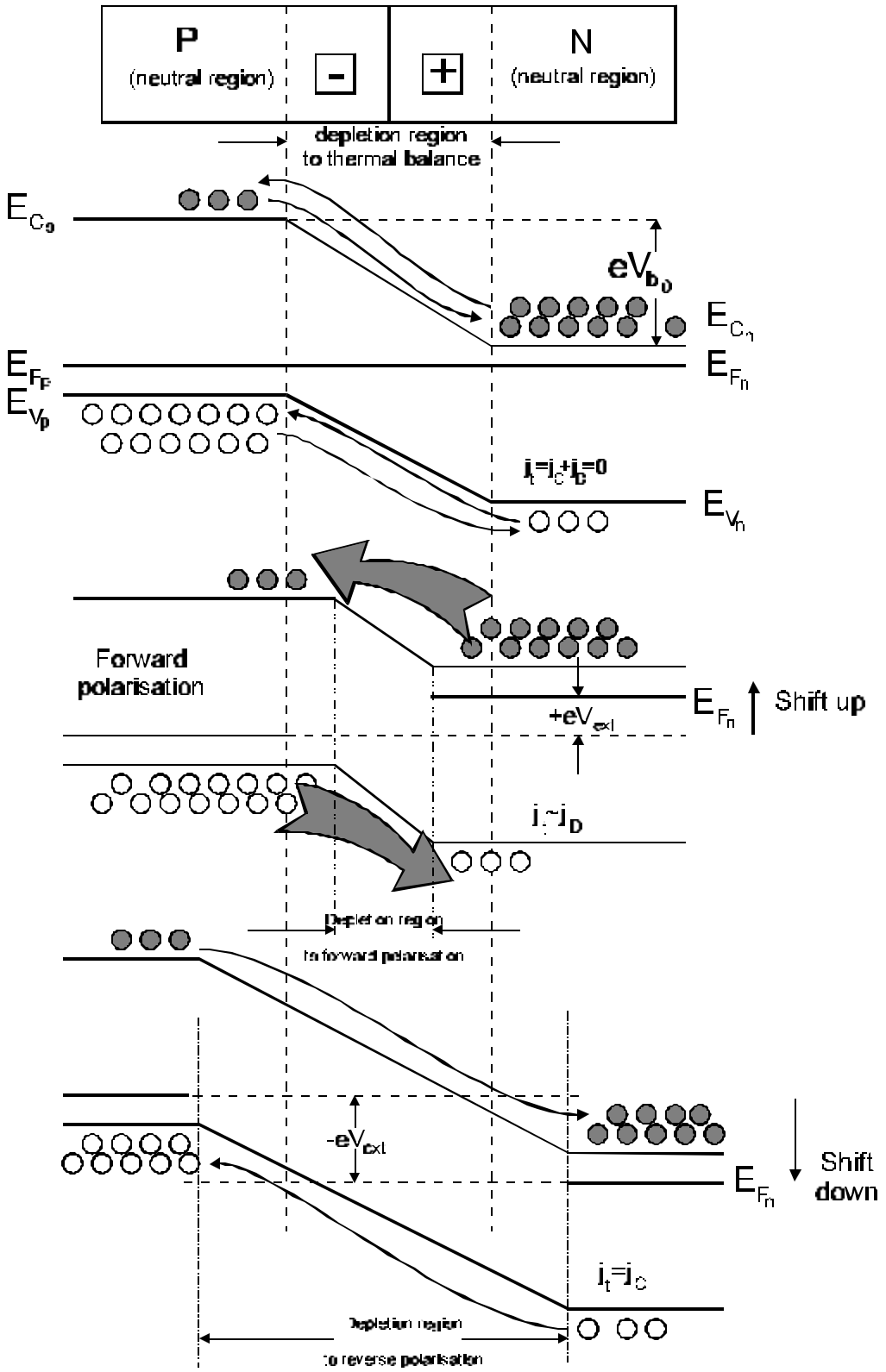



Fig.2.7
**Figure legend:** thin arrows = minority charge carriers currents
large arrows = majority charge carriers currents
○ **hole;** ● quasi-free electron

**2.3. Capacitance Effects in the P-N Junction.**

We determined in paragraph 2.1 that the diffusion length is proportional with the square root of the barrier potential, at thermal balance, as shown in the following formulas:

$$L_{p_0} = \left[ \frac{2\varepsilon}{eN_D\left(1+\frac{N_D}{N_A}\right)} \right]^{\frac{1}{2}} V_{b_0}^{\frac{1}{2}} \text{ and } L_{n_0} = \left[ \frac{2\varepsilon}{eN_A\left(1+\frac{N_A}{N_D}\right)} \right]^{\frac{1}{2}} V_{b_0}^{\frac{1}{2}}$$

If we biased the junction with a reverse potential, the height of barrier potential will increase to a value $V_b = V_{b_0} + V_{ext}$ and the diffusion length will increase:

$$L_p = \left[ \frac{2\varepsilon}{eN_D\left(1+\frac{N_D}{N_A}\right)} \right]^{\frac{1}{2}} (V_{b_0} + V_{ext})^{1/2}$$

equation which can be written, if multiplied by $\left( \frac{V_{b_0}}{V_{b_0}} \right)^{1/2}$, in a new form :

$$L_p = L_{p_0}\left(1+\frac{V_{ext}}{V_{b_0}}\right)^{1/2}$$

(we have considered, in the last two equations, the external voltage as negative).

We can obtain, in a similar way, a mathematical formula for diffusion length of electrons:

$$L_n = L_{n_0}\left(1+\frac{V_{ext}}{V_{b_0}}\right)^{1/2} .$$

Now if we define the dynamic capacitance of barrier potential as

$$C_B = \frac{dQ}{dV} \qquad 2.3.1.$$

we can calculate its value in the following steps:

$$C_B = \frac{dQ}{dV} = \frac{dQ}{dL_p}\cdot\frac{dL_p}{dV} \quad , \quad \text{but} \quad Q = eN_AL_pS \quad , \quad \text{then} \quad \frac{dQ}{dL_p} = eN_AS \quad \text{and}$$

$$\frac{dL_p}{dV} = \frac{\varepsilon}{eN_A\left(1+\frac{N_A}{N_D}\right)L_p}$$

then:

$$C_B = \frac{\varepsilon S}{L_p\left(1+\frac{N_A}{N_D}\right)} = \frac{\varepsilon S}{L_p + L_n} = \frac{\varepsilon S}{(L_{p_0} + L_{n_0})}\left(1+\frac{V_{ext}}{V_{b_0}}\right)^{-1/2} = C_{B_0}\left(1+\frac{V_{ext}}{V_{b_0}}\right)^{-1/2} \qquad 2.3.2.$$

The formula 2.3.2 gives us the value of barrier capacitance of the P-N junction, which looks like the formula of the capacitance of a plane capacitor. This capacitance is a characteristic of every diode at reverse polarisation.

This property is used in a class of special devices, called ***varicap***[1] diodes, which are used like capacitors whose capacitance is controlled by the applied reverse voltage.

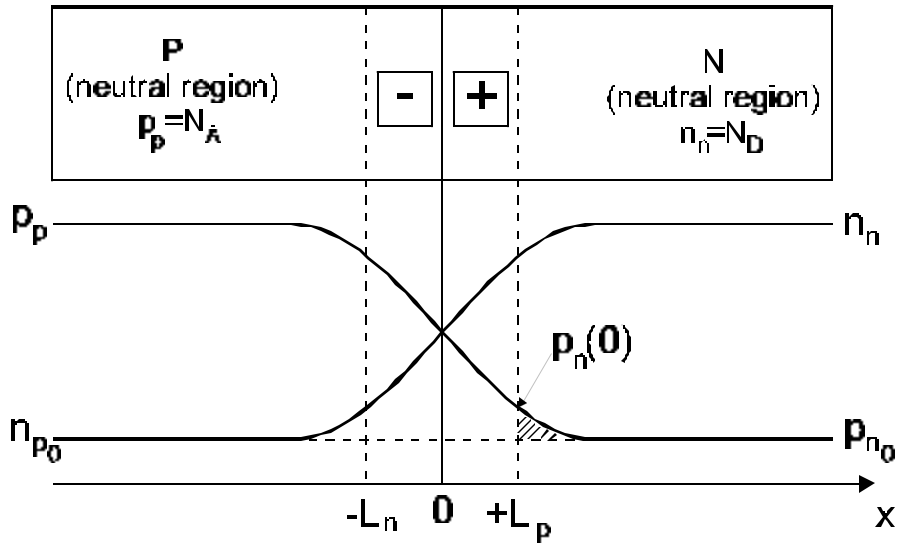


Fig.2.8

The common value of this capacitance is lying in the range 5-20 pF.

[1] VARIable CAPacitance

Now let us see what will happen if we forward bias the P-N junction. In this situation we will have an injection of minority carriers in a region in which they are in excess (see fig.2.8 which is similar with Fig.2.4)

The area below the curve, which describes the density of excess minority carriers in neutral region, will represent the amount of charge injected in these regions. From the equation of continuity we will obtain:

$$p_n(x) - p_{n_0} = (p_n(0) - p_{n_0}) \cdot e^{-\frac{x}{L_p}} \Rightarrow P_n(x) = P_n(0)e^{-\frac{x}{L_p}} \qquad 2.3.3.$$

Then, the amount of charge injected in neutral region of N type semiconductor is:

$$Q = \int_0^{+\infty} eSP_n(x)dx = -L_p eSP_n(0) \cdot e^{-\frac{x}{L_p}}\Big|_0^{+\infty} = eSP_n(0)L_p \qquad 2.3.4.$$

The dynamic capacitance is defined by the formula 2.3.1, therefore the capacitance can be written as:

$$C_D = \frac{dQ}{dV} = eSL_p \frac{dP_n(0)}{dV} \qquad 2.3.5$$

Now, if we take into account only the hole component of the total current which flows through the junction, i.e. $I_p = \frac{eSD_pP_n(0)}{L_p} \Rightarrow P_n(0) = \frac{L_p}{eSD_p} I_p$ , we can calculate the derivative of $P_n(0)$ function of voltage:

$$\frac{dP_n(0)}{dV} = \frac{L_p}{eSD_p} \cdot \frac{dI_p}{dV} \qquad 2.3.6.$$

Now if we replace the expression 2.3.6 in the equation 2.3.5 , we will find the value of the so called "***diffusion/storage capacitance of holes***" :

$$C_{D_p} = \frac{L_p^2}{D_p} \cdot \frac{dI_p}{dV} \qquad 2.3.7.$$

In a similar way we can calculate the "***diffusion/storage capacitance of electrons***":

$$C_{D_n} = \frac{L_n^2}{D_n} \cdot \frac{dI_n}{dV} \qquad 2.3.8.$$

The value of this capacitance is higher than the barrier capacitance. Values of 100 pF, or more, are common for this capacitance.

**2.4. Dynamic resistance of the diode.**

We can define, at forward bias of P-N junction, the dynamic resistance by next formula:

$$\frac{1}{r_d} = g = \frac{dI}{dV} \qquad 2.4.1.$$

The ideal diode equation can be approximated, at forward bias voltage, by:

$$I = I_0\left(e^{\frac{eV_{ext}}{kT}} - 1\right) \approx I_0 e^{\frac{eV_{ext}}{kT}} = I_0 e^{\frac{V_{ext}}{V_T}} \qquad 2.4.2.$$

Then the equation 2.4.1 becomes:

$$g = \frac{I}{V_T} \qquad 2.4.3.$$

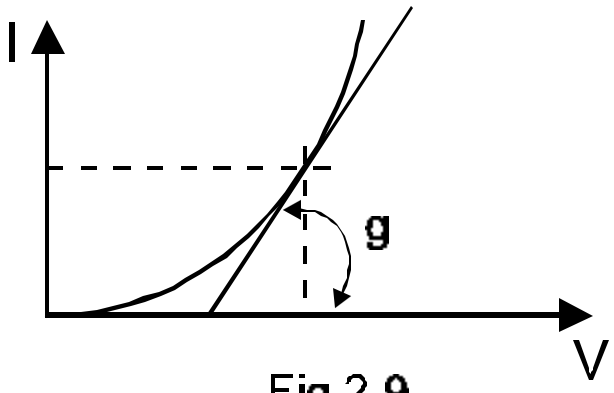



Fig.2.9

By looking at Figure 2.9, you can get a feeling of what is representing this "conductance" on the plot I function of V :

The slope of the tangent line in the point "I" to the curve which represents the current through diode function of biasing voltage, is the characteristic "conductance" for the diode at current "I".

The "conductance" is defined as the inverse of "resistance", as in equation 2.4.1.

**2.5. The Zener Diode.**

In the equation of ideal diode $I = I_0\left(e^{\frac{V_{ext}}{V_T}} - 1\right)$,

the saturation current is an important parameter of the diode and has the following expression: $I_0 = S\left(\frac{eD_P p_{n_0}}{L_P} + \frac{eD_N n_{p_0}}{L_N}\right)$. As you can see, this current is function of minority charge carrier density, which is a constant at a given temperature.

Since $p_{n_0} = \frac{n_i^2}{N_D}$ and $n_{p_0} = \frac{n_i^2}{N_A}$, the expression of saturation current becomes:

$$I_0 = eS\left(\frac{D_P}{L_P N_D} + \frac{D_N}{L_N N_A}\right)n_i^2 = Const \times \left(\frac{D_P}{L_P N_D} + \frac{D_N}{L_N N_A}\right) \times T^3 e^{-\frac{V_{G_0}}{V_T}} \qquad 2.5.1.$$

but the diffusion constants are inversely proportional with the temperature, then we can rewrite 2.5.1 in the final form:

$$I_0 = K_0 \times T^2 e^{-\frac{V_{G_0}}{V_T}} \qquad 2.5.2.$$

Then, for constant temperature we expect to have constant current. In practice, this formula holds only for moderate reverse voltages (see Fig.2.9).

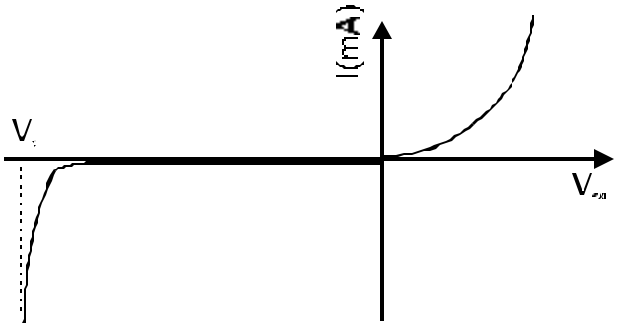


Fig.2.9

While increasing the reverse bias voltage applied to the diode, the current is constant up to a threshold voltage, called the "breakdown voltage". For voltage values larger than breakdown voltage, the saturation current increases abruptly. This behaviour of the saturation current may have two different origins:

- the avalanche multiplication of minority charge carriers (classical phenomena) which can occur at voltage higher than 100 V ;

- the tunnelling of bounded charge, through the barrier potential, from valence band of P type semiconductor, directly to valence band of N type semiconductor (quantum phenomena), which can occur at voltage lower than 100V.

Irrespective of the actually breakdown mechanism, the diodes which work in this regime are called “ZENER diodes”, and are generally used in voltage stabilisation.

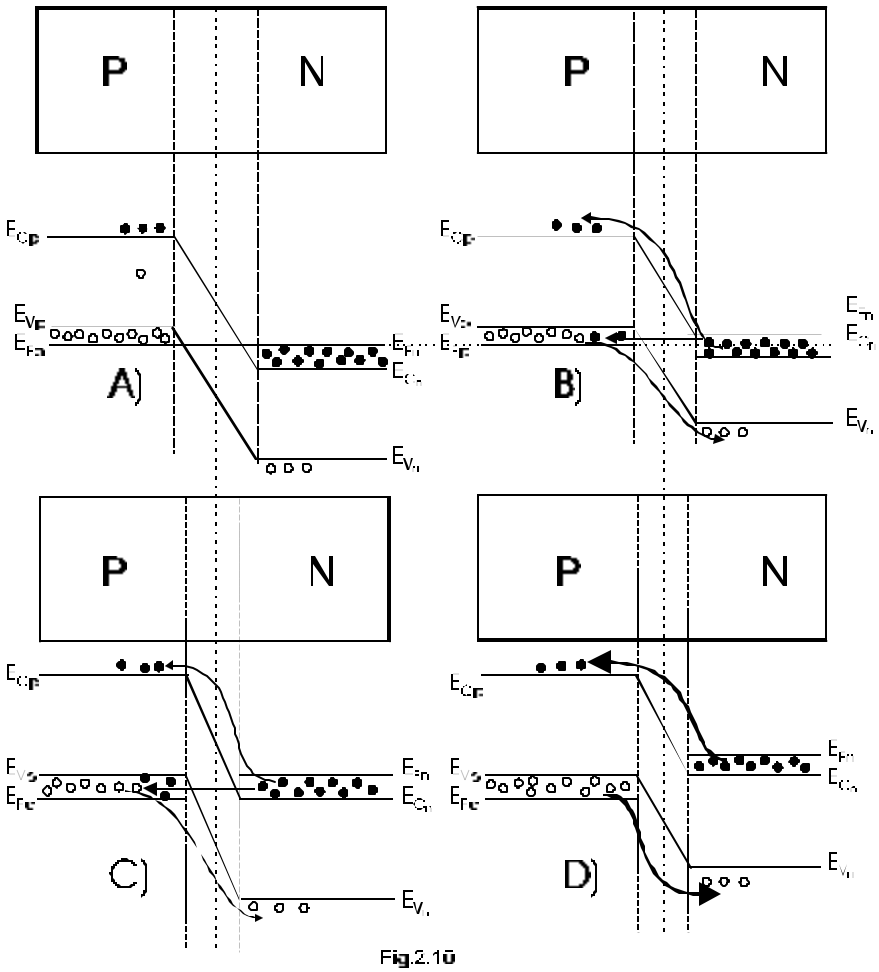



Fig.2.10

Figure legend: curved arrows = diffusion currents; straight arrows = tunneling currents

○ hole; ● quasi-free electron

**2.6. The Tunnel Diode.**

The tunnel diode is a special device that works at very high frequencies (more than 500Mhz). This diode is made in a form of a P-N junction with a heavy doping of both semiconductors ($N_A$ and $N_D$ higher than $10^{19}$ $cm^{-3}$). In these conditions the width of the barrier potential is very short, and the Fermi level, at thermal balance, lies in the valence band of P type semiconductor and, correspondingly, in conduction band of N type semiconductor (see Fig.2.10)

Figure caption:

A) Thermal balance;

B) Forward biasing voltage (tunnelling current is increasing)

C) Forward biasing voltage, higher than in the case B (maximum tunnelling current)

D) Forward biasing voltage, higher than in the case C (tunnelling current is zero)

The current-voltage characteristic is shown in Fig.2.11. We can see in Fig.2.11 four different regions of this plot.

Region 1: here the reverse current increases rapidly, because all electrons which are in valence band of P type semiconductor are tunnelling through the barrier potential because they "see" unfilled states in the conduction band of N type semiconductor, upper than the Fermi level.

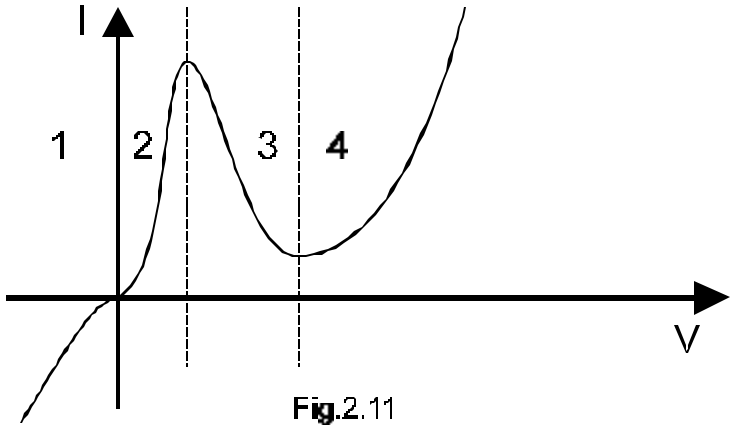


Fig.2.11

Region 2: At a small forward biasing voltage begins a movement of majority carriers over the barrier potential, like in a normal diode, but this current is in concurrence with the "negative current" generated by the "tunnelling effect". When the tunnelling current becomes predominant, the current through the diode moves in region 3 (case B in Fig.2.10)

Region 3: The tunnelling current is higher than the "normal" current, therefore the current through the diode decreases with the bias voltage increase, until a minimum value (at the end of region 3), at which the tunnelling current is maximal (case C in Fig.2.10). In this region the diode is characterised by a "**negative dynamic resistance**" (the slope of this part of the characteristic is negative, as you can see).

Region 4: With the increasing of the voltage, the energy bands of N type semiconductor are shifted more to upper energies, and the region with free electrons from conduction band of N type semiconductor begins to look at a region of forbidden band, then the tunnelling effect decreases, until this effect vanishes (case D in Fig.2.10.) and all the current through the diode will be a normal forward biasing current.

# Chapter 3. The Bipolar Junction Transistor (BJT).

## 3.1. Phenomenological description of Bipolar Transistor.

The Bipolar Junction Transistor (BJT)[2] or simply Junction Transistor, has the structure shown in Fig.3.1.

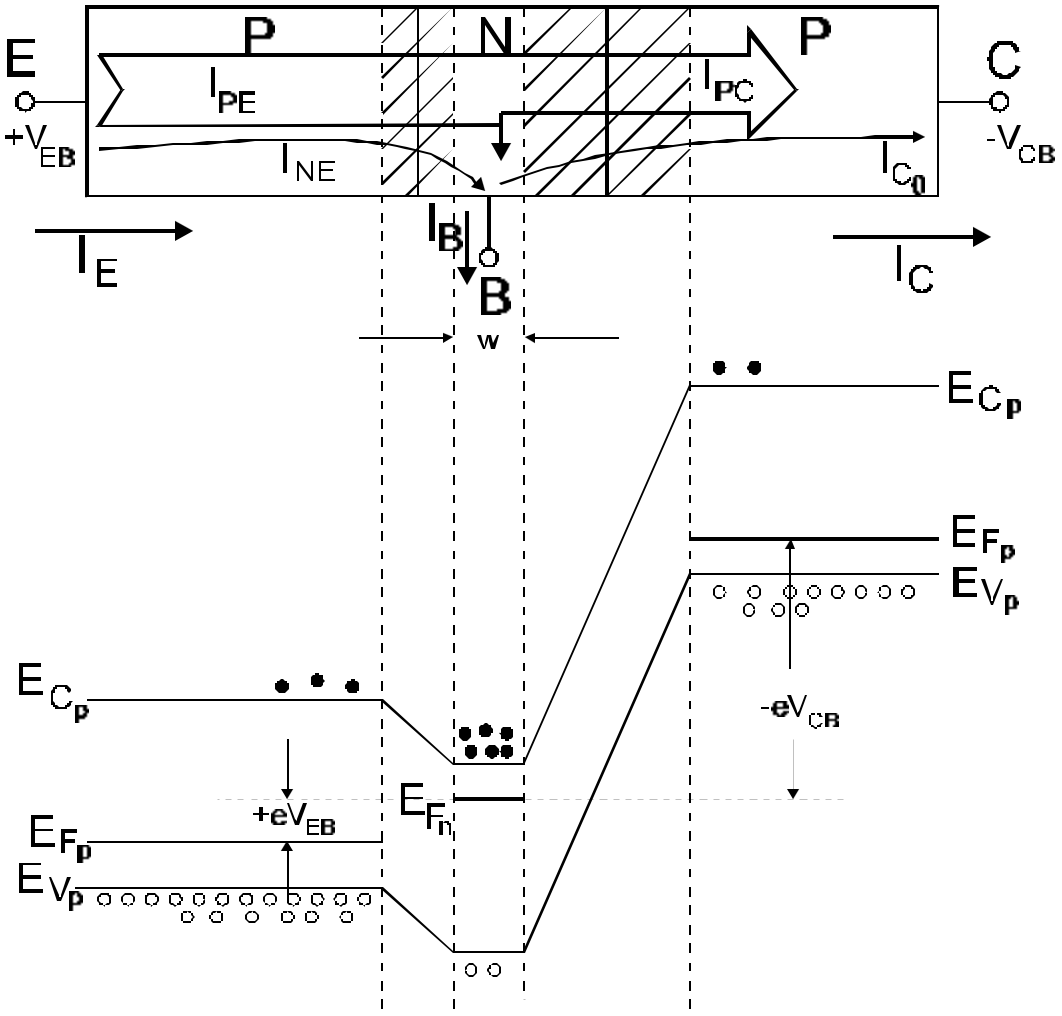



Fig.3.1.
Figure legend: large arrows= majority charge carriers currents
thin arrows= minority charge carriers currents

As you can see in Fig. 3.1, the currents which flow through the bipolar transistor, in the conditions of forward biasing of E-B junction and reverse biasing of C-B junction, are:

$I_E = I_{PE} + I_{NE}$ ; $I_C = I_{PC} + I_{C_0}$ ; $I_B = I_E - I_C$

We can define the following "transistor's constants" :

The efficiency of emitter: $\gamma = \frac{I_{PE}}{I_E}$ (the ideal value of this constant is 1)

[2] Invented by Shockley

The carrier factor: $\beta^* = \frac{I_{PC}}{I_{PE}}$ (the ideal value of this constant is 1)

The current gain : $\alpha = \frac{I_C - I_{C_0}}{I_E}$ (the ideal value of this constant is 1)

From the last relation we can find the "transistor's equation"

$$I_C = \alpha I_E + I_{C_0} \qquad 3.1.1.$$

In practice the current gain constant has values among 0.95 - 0.999 . Then we can use the approximate relation :

$$I_C \approx \alpha I_E \qquad 3.1.2.$$

Like in the case of P-N junction, such a device must be built in the same piece of semiconductor material, in order to assure the continuity of the crystalline lattice. Any defect in the lattice would greatly impair carrier mobility and would distort the energy bands. The mandatory conditions for having such relations between currents, then to have a “transistor behaviour”, are:

- the doping of Emitter is higher than the doping of the Base, i.e. $N_{A(E)} >> N_{D(B)}$

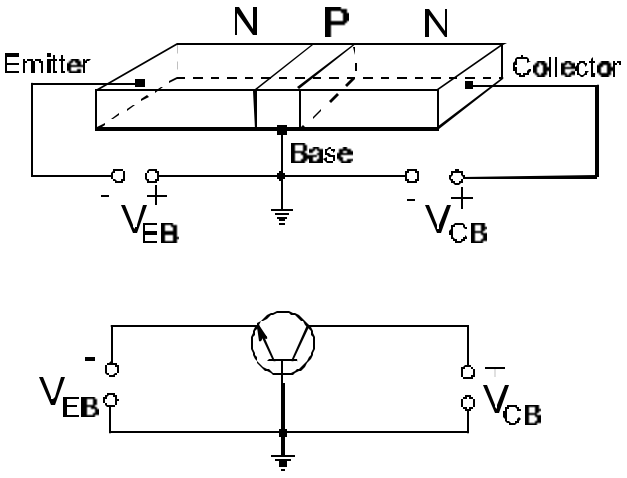



Fig.3.2.

- the base is thin enough that the diffusion length of minority charge carriers which are injected in Base is higher than the width of neutral region of the Base, i.e. $L_p > w$

In Fig. 3.2 we represented the case of a NPN transistor, biased in the active regime (Emitter junction forward biased

and Collector junction reverse biased). In this case all currents have inverse directions compared to the case of PNP transistor. Now, if we take into account the equation 3.1.2., we can approximately calculate the power gain of this device.

Through the junction of emitter, characterised by dynamically resistance $r_d < 10\Omega$ flows the current $I_E$ . Then the input power, dissipated on this junction is

$$P_{in} \approx I_E r_d^2$$

The output power is the power dissipated on collector junction, which is reverse biased, then it is characterised by a higher resistance $R_C > 10^4\Omega$ . Then the output power is approximately

$$P_{out} \approx I_C R_C^2$$

One has to note that these values are not the actual total power dissipated on the emitter and collector, since $r_d$ and $R_C$ are not the static resistances, but the dynamic ones, related to the AC signal. Therefore $P_{in}$ and $P_{out}$ will be AC signal powers at the input and output. The power gain is going to be:

$$G = \frac{P_{out}}{P_{in}} = \frac{I_C R_C^2}{I_E r_d^2} \approx \frac{\alpha I_E R_C^2}{I_E r_d^2} \geq 10^4$$

Now, from the last relation we can understand why this device was called "trans-resistor" or "transistor" . This device makes possible the transfer of a current which flows through a region with low resistance, in a region with high resistance, without a sensitive modification of the current.

### 3.2. The analytical equations of transistor's currents.

The emitter current has two components, as we saw in last section. The electron component of this current must have the same expression like the electron component of an ideal diode, therefore:

$$I_{n_E} = \frac{AeD_n n_{p_0}}{L_n}\left(e^{\frac{V_E}{V_T}} - 1\right) \qquad 3.2.1.$$

In this equation we will annotate the diffusion length of electrons in neutral region of emitter by $L_E$ and the density of minority charge carriers in emitter by $n_{E_0}$. With these new annotations, formula 3.2.1 becomes:

$$I_{n_E} = \frac{AeD_n n_{E_0}}{L_E}\left(e^{\frac{V_E}{V_T}} - 1\right) \qquad 3.2.2.$$

To compute the hole component of the emitter current, we must take into account that this current is a diffusion current in a neutral base region, where it is a minority carrier current. Therefore, the expression of this current is very much alike the formula of a diffusion current:

$$j_{p_E} = -eD_p \frac{dp}{dx} \qquad 3.2.3.$$

But already we know the expression of the density of minority charge carriers injected in a neutral region:

$$p_n(x) - p_{n_0} = Ke^{-\frac{x}{L_p}} \qquad 3.2.4.$$

Because the transit of these charge carriers through the neutral region of the base is fast, due to the small thickness of the base, $(w < L_p)$, we can approximate the equation 3.2.4. by:

$$p_n(x) - p_{n_0} = K_1 + K_2 x \qquad 3.2.5.$$

Then, by replacing 3.2.5 in 3.2.3., we will obtain :

$$j_{p_E} = -eD_p K_2 \qquad 3.2.6.$$

We can obtain the values of constants $K_1$ and $K_2$ from the boundary conditions of equation 3.2.5:

at x=0 , we have

$$p_n(0) - p_{n_0} = K_1 \qquad 3.2.7.$$

and at x=w , we have:

$$p_n(w) - p_{n_0} = K_1 + K_2 w \qquad 3.2.8.$$

But the density of the injected carriers at x=0, which represents the boundary between the space charge region of emitter junction and neutral region of the base, is given by:

$$p_n(0) = p_{n_0} e^{\frac{V_E}{V_T}} \qquad 3.2.9.$$

Now, if we replace the 3.2.9. in 3.2.7. we will obtain the value of the $K_1$ constant:

$$K_1 = p_{n_0}\left( e^{\frac{V_E}{V_T}} - 1 \right) \qquad 3.2.10.$$

In the same way, taking into account that:

$$p_n(w) = p_{n_0} e^{\frac{V_C}{V_T}}$$

we will obtain from equation 3.2.8. the value of the $K_2$ constant:

$$K_2 = \frac{p_{n_0}\left( e^{\frac{V_C}{V_T}} - 1 \right) - p_{n_0}\left( e^{\frac{V_E}{V_T}} - 1 \right)}{w} \qquad 3.2.11.$$

Now, if we replace the value of $K_2$ constant, given by 3.2.11. in 3.2.6. equation we will obtain the value of holes current which flows through the emitter junction:

$$I_{p_E} = -eD_p A \frac{p_{n_0}\left( e^{\frac{V_C}{V_T}} - 1 \right) - p_{n_0}\left( e^{\frac{V_E}{V_T}} - 1 \right)}{w} \qquad 3.2.12.$$

Finally, the total current which flows through the emitter can be expressed as:

$$I_E = I_{p_E} + I_{n_E} = \left[\frac{eAD_n n_{E_0}}{L_E} + \frac{eAD_p p_{n_0}}{w}\right]\left(e^{\frac{V_E}{V_T}} - 1\right) - \frac{eAD_p p_{n_0}}{w}\left(e^{\frac{V_C}{V_T}} - 1\right) \qquad 3.2.13.$$

In the same way we can obtain the expression for the collector current, which is given by:

$$I_C = I_{p_C} + I_{C_0} \qquad 3.2.14.$$

The saturation current is like the electronic component of the current of an ideal diode:

$$I_{C_0} = -\frac{eAD_n n_{C_0}}{L_C}\left(e^{\frac{V_C}{V_T}} - 1\right) \qquad 3.2.15.$$

Now, if we neglect the recombination phenomenon in the neutral region of the base, i.e. $I_{p_E} \approx I_{p_C}$ , the sum of 3.2.15. and 3.2.12. equations will give us the expression for collector current:

$$I_C = I_{p_C} + I_{C_0} = -\left[\frac{eAD_n n_{C_0}}{L_C} + \frac{eAD_p p_{n_0}}{w}\right]\left(e^{\frac{V_C}{V_T}} - 1\right) + \frac{eAD_p p_{n_0}}{w}\left(e^{\frac{V_E}{V_T}} - 1\right) \qquad 3.2.16.$$

The equations 3.2.13. and 3.2.16. represent the analytic expressions for currents which flow through the transistor.

These relations can be written in condensed forms such:

$$\begin{aligned} I_E &= a_{11}\left(e^{\frac{V_E}{V_T}} - 1\right) + a_{12}\left(e^{\frac{V_C}{V_T}} - 1\right) \\ I_C &= a_{21}\left(e^{\frac{V_E}{V_T}} - 1\right) + a_{22}\left(e^{\frac{V_C}{V_T}} - 1\right) \end{aligned} \qquad 3.2.17.$$

where the coefficients $a_{ij}$ are:

$$a_{11} = \frac{eAD_n n_{E_0}}{L_E} + \frac{eAD_p p_{n_0}}{w} \quad ; \quad a_{12} = -\frac{eAD_p p_{n_0}}{w}$$

$$a_{21} = \frac{eAD_p p_{n_0}}{w} \quad ; a_{22} = -\left( \frac{eAD_n n_{C_0}}{L_C} + \frac{eAD_p p_{n_0}}{w} \right) \qquad 3.2.18.$$

The relations 3.2.17. are so called "Ebers-Moll" relations.

### 3.3. Ebers-Moll Model of Bipolar Transistor.

Taking into account the equation 3.1.1. and the general equation for $I_{Co}$, the reverse current of the collector towards the base junction, the general equation 3.1.1 can be written as:

$$I_C = \alpha_N I_E + I_{Co}(\exp \frac{V_C}{V_T} - 1) \qquad 3.3.1.$$

where $\alpha_N$ is the current gain under normal conditions of biasing (emitter to base junction forward biased and base to collector junction reverse biased). Now, if we take the transistor like an reversible device and reversing the biasing, we can rewrite the 3.3.1. as:

$$I_E = -\alpha_R I_C + I_{Eo}(\exp \frac{V_E}{V_T} - 1) \qquad 3.3.2.$$

where $\alpha_R$ is the current gain in reverse conditions of the biasing, having lower value than $\alpha_N$ because the transistor doesn't work in normal regime.

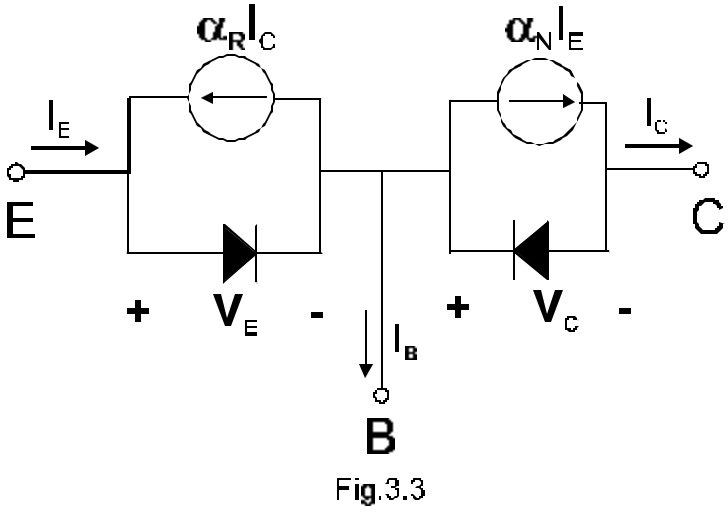


Fig.3.3

The 3.3.1 and 3.3.2 relations can be used to describe a simple model of bipolar transistor, named "Ebers-Moll model", shown in Fig.3.3. where $V_E$ is the forward bias of the emitter and $V_C$ is the reverse bias of the collector. Then the first terms of 3.3.1 and 3.3.2 relations are represented in the Ebers-Moll model as constant current generators and the seconds terms of these relations are represented by the currents which are flowing through two equivalent diodes, first one biased with $V_E$, and the second biased with $V_C$. Now, if we take into account that the current given by constant current generator $\alpha_R I_C$ is lower than the current given by the constant current generator $\alpha_N I_E$, and the reverse current of the equivalent diode of the collector junction is very small in comparison with the current given by forward

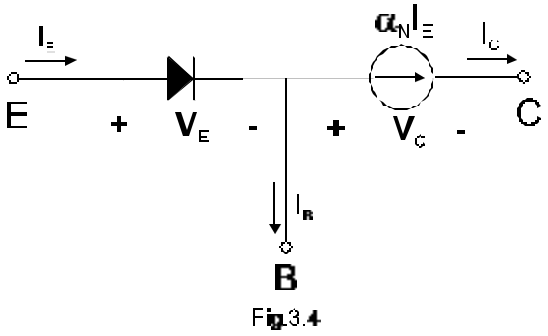


Fig.3.4

biasing equivalent diode of the emitter junction, the Ebers-Moll model can be simplified as in Fig.3.4.

Using the model showed in Fig.3.4, the demonstration of the power gain given at the beginning of this chapter becomes even easier to understand.

### 3.4. Static Characteristics of Bipolar Transistor.

The most common connections for the bipolar transistor are the "Common Base Connection" and the "Common Emitter Connection", named this way because the Base,

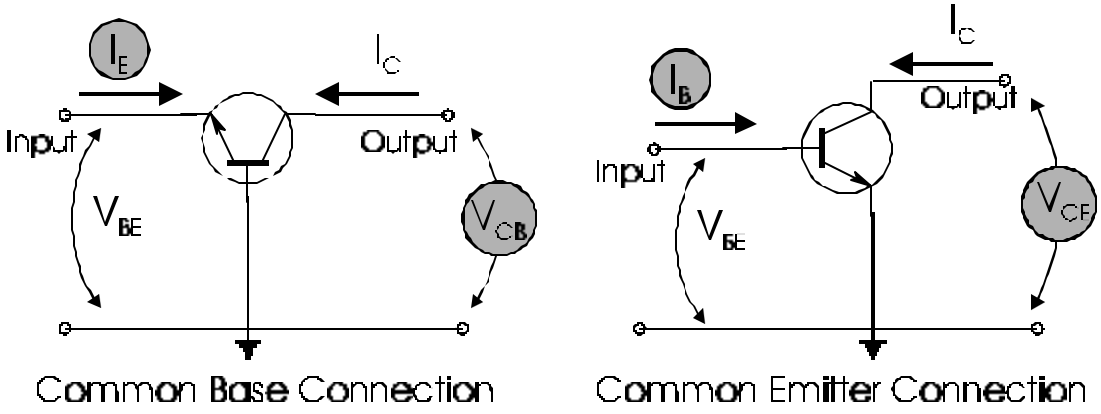


Fig. 3.5

respectively the Emitter, are connected to the common connection between input and output, connection which is conventionally taken as ground.

In Fig. 3.5 are shown these two basic connections of bipolar transistor. Each one is characterised by two input connections and two output connections. Then we have four terminals, of which two of them connected to the common ground. For that reason, such a device is named a "**four-terminal network**". The behaviour of such device can be characterised by the input and output currents and voltages. Usually we take as independent

variables the current of input and the voltage of output. In this way we can write the dependent variable (voltage of input and current of output) function of the independent variable.

$V_{BE} = f_1(I_E, V_{CB}); I_C = f_2(I_E, V_{CB})$ will be the relations for CBC four-terminal network, and

$V_{BE} = f_1(I_E, V_{CB}); I_C = f_2(I_E, V_{CB})$ will be the relations for CEC four-terminal network.

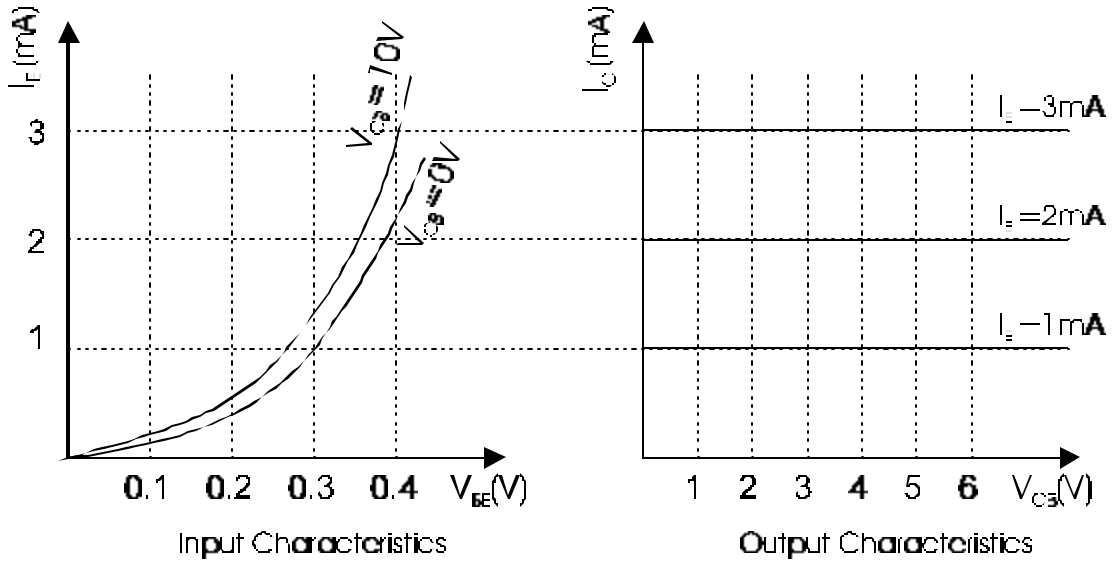



**Fig. 3.6**

In figure 3.6 you can see such characteristics for CBC circuits, and in figure 3.7 for CEC circuits.

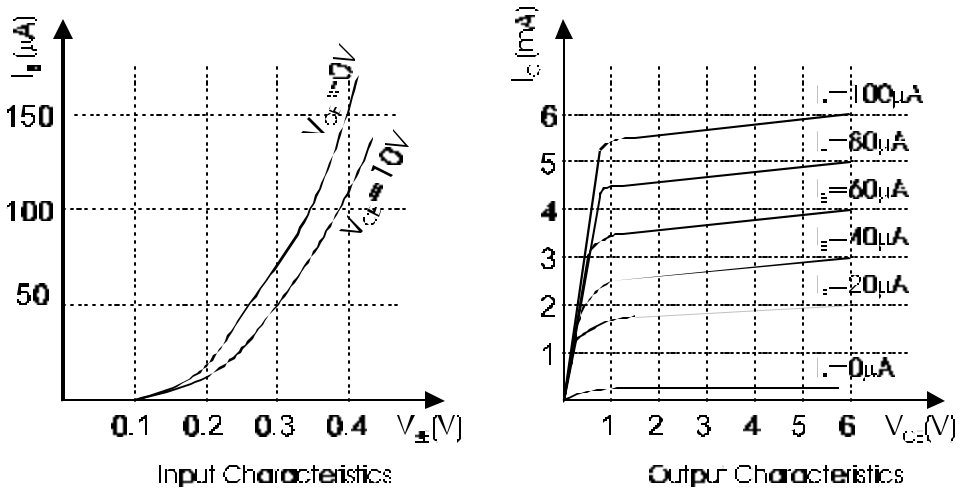



Fig. 3.7

The way we defined the current gain for CBC connection, the relation between the input current and the output current is provided by the following equation:

$$I_C = \alpha I_E + I_{C_0} \qquad 3.4.1.$$

This relation is so called "the device equation" for the transistor in CBC connection.

In the case of CEC connection the "device equation" can be found by replacing the emitter current by : $I_E = I_C + I_B$ . In this case the equation 3.4.1. becomes:

$$I_C = \beta I_B + (\beta + 1) I_{C_0} \qquad 3.4.2.$$

where $\beta = \dfrac{\alpha}{1-\alpha}$ is the current gain in the CEC connection. A typical value for **b** is 100.

**The biasing circuits for bipolar transistor.**

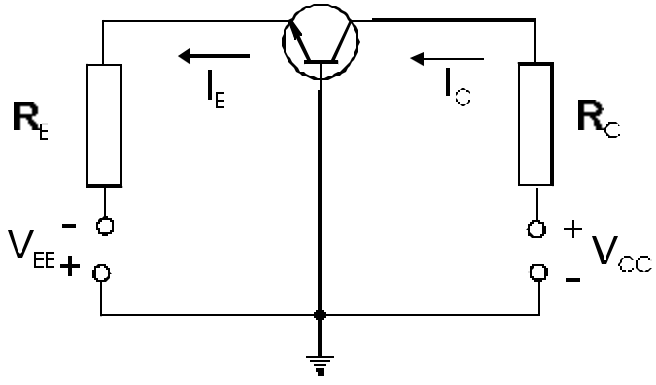



Fig.3.8

In the case of CBC connection we may have a biasing circuit comprising two d.c. sources, like in Fig. 3.8

In the transistor's active region, the junction E-B is forward biased and the junction C-B is reverse biased. Under

the assumption $V_{EB}>>kT/e=V_T$ , we can infer that the emitter current is higher than the reverse collector current $I_{C0}$ , then we have in a wide range of values for the $I_E$ the relation

$$V_{EB} \cong \text{const.} \qquad 3.4.3.$$

and the equation 3.3.1. becomes

$$I_C \approx \alpha I_E \cong I_E \qquad 3.4.4.$$

The equations 3.4.3. and 3.4.4. represent the "device equations". Now we will write the "circuit equations", which will be the based on second Kirchhoff's law for the input and output circuits:

$$V_{EE} = V_{EB} + I_E R_E \qquad 3.4.5.$$

$$V_{CC} = V_{CB} + I_C R_C \qquad 3.4.6.$$

From 3.4.5. we can calculate the value of $I_E$ for a given circuit:

$$I_E \cong \frac{V_{EE}}{R_E} = \text{const.} \qquad 3.4.7.$$

assuming that $V_{EE}>>V_{EB}$ .

Then, if we take into account the relation 3.4.4. , we can assert that $I_E \approx I_C = \text{const.}$

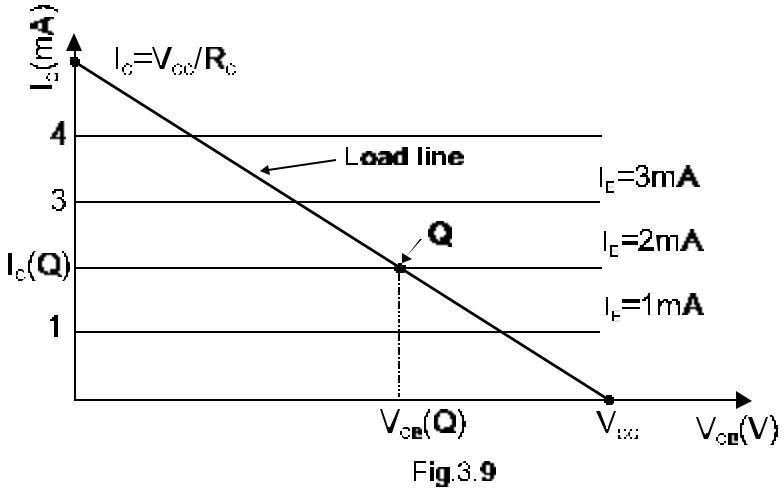


Fig.3.9

The equation 3.4.6. represents the so called "load line equation", from which we can calculate the bias voltage of the C-B junction.

$$V_{CB} = V_{CC} - I_C R_C \qquad 3.4.8.$$

The intersection between the load line and the output static characteristic corresponding to the emitter current calculated by equation 3.4.7. represents the so called "static operating point" for the transistor. This point is marked in fig.3.9 with a **Q** letter**.** In this point the bias voltage for C-B junction is provided by equation 3.4.8.

***Equations 3.4.7. and 3.4.8. prove that the CBC connection is the most stable operating configuration of bipolar transistor. This is provided by the fact that we control the output current $I_C$ with a current $I_E$ , higher than the residual current of the collector and the gain current $\alpha$ is approximately constant, having values in the range 0.98-0.99.***

The circuit for the CEC connection, which is by the way the most usual circuit, is shown in Fig. 3.10. The biasing circuit is commonly called "voltage divider biasing" or “universal biasing” circuit, because the resistors $R_{B1}$ and $R_{B2}$ provide the biasing of E-B and C-B junctions using a single power supply.

In this configuration, the situation is quite different as compared to the CBC connection because the current gain $\beta$ may have a wide dispersion over individual transistors. If $\alpha$ is 0.98, $\beta$ is 49 while if $\alpha$ is one percent higher, 0.99, $\beta$ is 99.

For this configuration the "device equation" is given by eq. 3.4.2, but usually for Silicon transistors is used the simplified formula

$$I_C \cong \beta I_B \qquad 3.4.9.$$

The voltage divider biasing circuit showed in Fig.6 has an equivalent d.c. circuit which is used to find the circuit equations. To obtain this equivalent circuit we must use next steps:

1. We will assume all capacitors having "infinite resistance";

2. We will apply the "Thevenin's theorem" for the voltage divider biasing circuit;

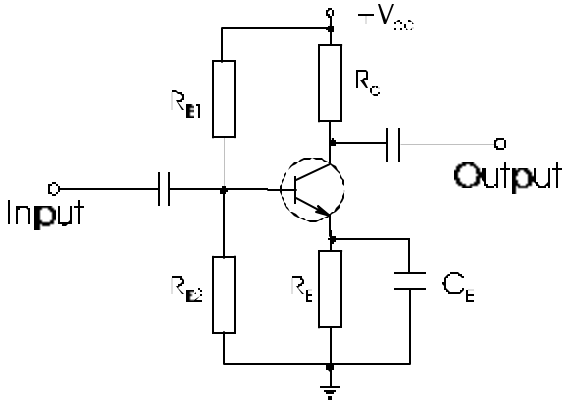


Fig.3.10

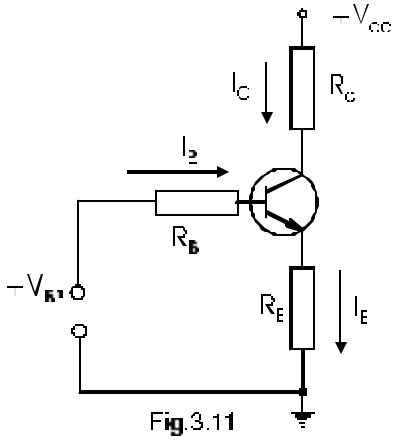


Fig.3.11

In this case the equivalent circuit of Fig.3.10 is the circuit showed in Fig.3.11.

Now if we write the second Kirchhoff's law for input and output circuits of Fig.3.11, we will find next equations:

$$V_{BB} = I_B R_B + V_{BE} + I_E R_E \qquad 3.4.10.$$

$$V_{CC} = I_C R_C + V_{CE} + I_E R_E \qquad 3.4.11.$$

where $R_B = \dfrac{R_{B1}R_{B2}}{R_{B1} + R_{B2}}$ and $V_{BB} = \dfrac{V_{CC}}{R_{B1} + R_{B2}} R_{B2}$ as result from Thevenin's theorem. The equations 3.4.10 and 3.4.11 are the "circuit equations" for CEC connection. If the resistance $\mathbf{R_E=0}$ **,** from relations 3.4.10 we will find that

$$I_B = \frac{V_{BB} - V_{BE}}{R_B} \qquad 3.4.12.$$

and using relation 3.4.9, results that:

$$I_C = \frac{\beta(V_{BB} - V_{BE})}{R_B} \qquad 3.4.13.$$

But $\beta$ may have a wide dispersion, then for a given base current we can have a lot of output currents. In this case the "operating point" of the transistor is not stable. To prevent this

situation it is necessary to have the condition $R_E \neq 0$. In this case the equation 3.4.12. becomes:

$$I_B = \frac{V_{BB} - V_{BE}}{R_B + R_E(\beta + 1)} \qquad 3.4.14.$$

(in equation 3.4.14 we take into account that $I_E = I_C + I_B = I_B(\beta + 1)$, if we used the relation 3.4.9.)

In this case equation 3.4.13. becomes:

$$I_C = \frac{\beta(V_{BB} - V_{BE})}{R_B + R_E(\beta + 1)} \qquad 3.4.15.$$

Now, if we have met the criterion

$$R_E(\beta + 1) >> R_B \quad , \qquad 3.4.16.$$

the output current $I_c$ **becomes independent of** $\beta$**, then the "operating point" becomes stable,** and the equation 1.4.15. becomes:

$$I_C = \frac{\beta(V_{BB} - V_{BE})}{R_B + R_E(\beta + 1)} \approx \frac{\beta(V_{BB} - V_{BE})}{R_E(\beta + 1)} \approx \frac{V_{BB} - V_{BE}}{R_E} = \text{const.} \qquad 3.4.17.$$

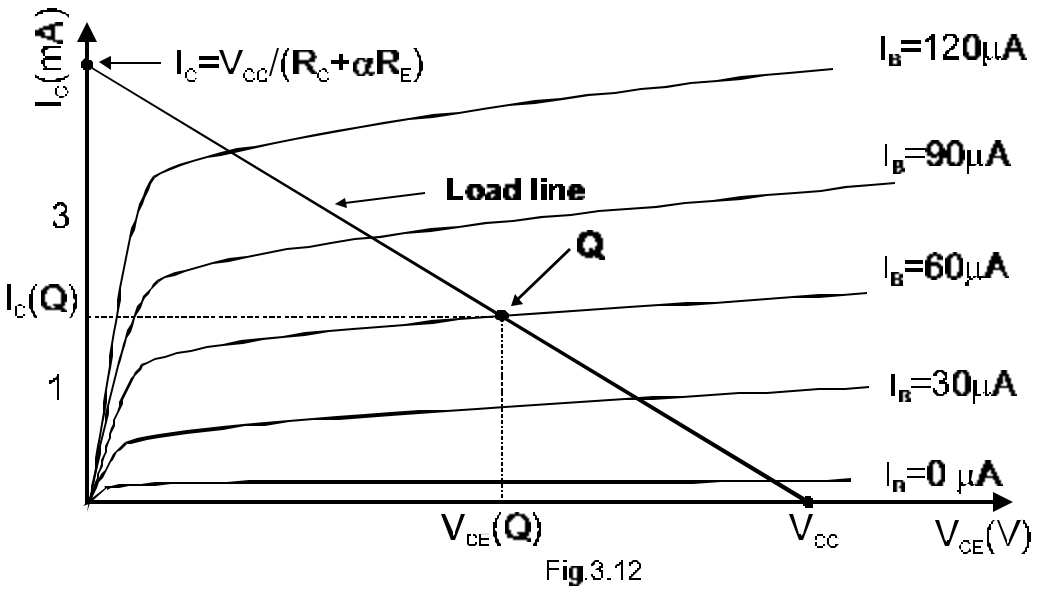



Fig.3.12

Because the current gain in CEC is large, we can approximate $\beta + 1 \approx \beta$.

The equation 3.4.11 represents the "load line equation" for CEC of bipolar transistor. In Fig.3.12 you can see the output characteristics for CEC of the bipolar transistor and the operating point, obtained in the same way like in the case of CBC.

**3.5.The stabilisation of working conditions for bipolar transistor.**

The output current of CBC or CEC circuits is the collector current. The value of this current is function of temperature by his dependence on: saturation collector current $I_{C0}$, bias voltage of the E-B junction $V_{BE}$ and the current gain $\beta$.

Then we can write that

$$I_C = I_C(I_{C0}, V_{BE}, \beta) \qquad 3.5.1.$$

where each variable depends on temperature following a common law:

$$I_{C0}(T) = I_{C0}(T_0)e^{a\cdot(T_1-T_2)}$$

where $T_0$=300$^0$K; the value of the constant **a** depends on the nature of the semiconductor, being higher for Germanium which has the gap energy (width of the forbidden band) lower than Silicon.

$$\beta(T) = \beta(T_0)\cdot\left[1+\frac{T_1-T_2}{K}\right]$$

where the value of constant **K** is 100 for Germanium and 50 for Silicon.

$$\frac{\partial V_{BE}}{\partial T} = -2.2\text{mV}/^0\text{C}$$

The strongest dependence on temperature is for the saturation current, because this current is provided by minority carriers, and the their concentration depends exponentially on temperature.

Now, if we take the derivative of expression 3.5.1. with respect to temperature, we will obtain the following equation:

$$\frac{\partial I_C}{\partial T} = \frac{\partial I_C}{\partial I_{C0}} \times \frac{\partial I_{C0}}{\partial T} + \frac{\partial I_C}{\partial V_{BE}} \times \frac{\partial V_{BE}}{\partial T} + \frac{\partial I_C}{\partial \beta} \times \frac{\partial \beta}{\partial T} \qquad 3.5.2.$$

where the coefficients of the temperature derivatives are the so called “sensitivity factors”:

$$\frac{\partial I_C}{\partial I_{C0}} = S_I \; ; \frac{\partial I_C}{\partial V_{BE}} = S_U \; ; \frac{\partial I_C}{\partial \beta} = S_\beta$$

The sensitivity factor of current $\mathbf{S_I}$ is the most important, its minimisation leading to the minimisation of all other factors.

In the aim to find the expression of $\mathbf{S_I}$, we must calculate the next derivative:

$$\frac{\partial}{\partial I_C}[I_C = \beta I_B + (\beta + 1) I_{C0}]$$

from which we can obtain :

$$S_I = \frac{\beta + 1}{1 - \beta \dfrac{\partial I_B}{\partial I_C}} \qquad 3.5.3.$$

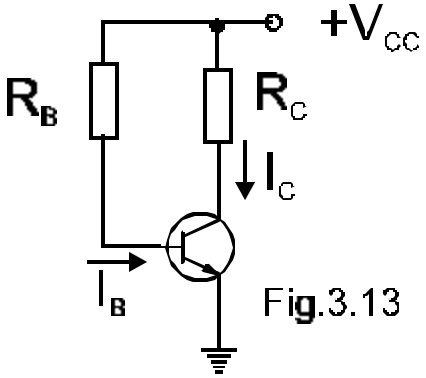


Fig.3.13

The value of the derivative $\dfrac{\partial I_B}{\partial I_C}$ depends on the type of circuit used for biasing the transistor.

The simplest biasing circuit is shown in figure 3.13. The base current is given by the next relation :

$I_B = \dfrac{V_{CC} - V_{BE}}{R_B}$ , which can be found by writing the second Kirchhoff's law for the input circuit. Because $V_{CC} >> V_{BE}$ we can ignore the value of $V_{BE}$ in the expression of $I_B$, therefore it results that the base current is constant, and, as consequently, its derivative with respect to $I_C$ is zero. In this case the equation 3.5.3. becomes

$$S_I = \beta + 1$$

which is a large value, therefore the sensitivity with the temperature is very high. This circuit has a bad stability function of temperature.

A good stability with the temperature has the circuit shown in Fig.3.14.

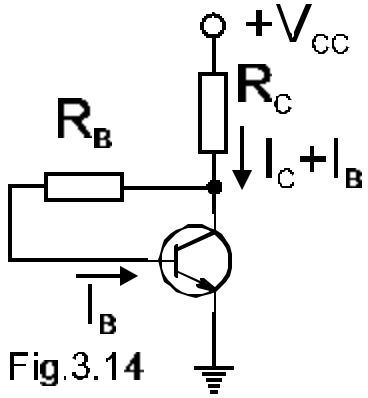


Fig.3.14

Using the second Kirchhoff law, we can write the next equation:

$V_{CC} = R_C(I_C + I_B) + R_B I_B + V_{BE}$ from which, ignoring the $V_{BE}$ voltage because it is much smaller than $V_{CC}$, we will find:

$$I_B = -\frac{R_C}{R_C + R_B} I_C \Rightarrow \frac{\partial I_B}{\partial I_C} = -\frac{R_C}{R_C + R_B}$$

Than, if we replace this derivative in 3.5.3. equation, we will find:

$$S_I = \frac{\beta + 1}{1 + \beta \dfrac{R_C}{R_C + R_B}} \approx \frac{R_C + R_B}{R_C}$$ which has values in the range 3 to 10, depending on

the values of resistors used in the circuit. This is a low sensitivity, resulting in a good stability of the circuit with respect to the temperature variations. In the particular case $\mathbf{R_B=0}$ we get the best sensitivity value, $S_I=1$, but in this case the transistor has the C-B junction shunted. However, such a circuit is used to stabilise the second transistor. This method of stabilisation is called "current mirror stabilisation".

The best value, that means the lowest value for $S_I$, is obtained in the case of voltage divider biasing circuit shown in Fig.3.15

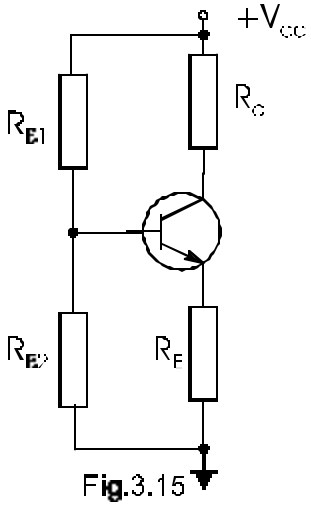


Fig.3.15

From the d.c. equivalent circuit, discussed earlier in this chapter, we can find the value of $I_B$ as a function of the $I_C$ current

$$I_B = -\frac{R_E}{R_B + R_E} I_C \ .$$

Then the expression 3.5.3. becomes:

$$S_I = \frac{\beta + 1}{1 + \beta \dfrac{R_E}{R_E + R_B}} \approx \frac{R_E + R_B}{R_E}$$

which may have typical values in the range 2 to 6.

# Chapter 4. Small Signals Operating Regime.

**Notation Conventions for the Dynamic Regime.**

In the dynamic regime we have in our circuit both currents and voltages to look at. In general currents and voltages are designated by capital letters **I** or **V** having a subscript which represents the letter characteristic for the transistor terminal (E for emitter, C for collector and B for base). The a.c. components are denoted by italic small letters having small letters as subscripts, designating transistor terminal, as in the d.c. case (e for emitter, c for collector and b for base). Then, the d.c. collector current is noted by $\mathbf{I_C}$ and the a.c. collector current by $\boldsymbol{i_c}$ .

The sum of both components is denoted by italic capital letters as you can see in the following example: $\boldsymbol{I_c = i_c + I_C}$

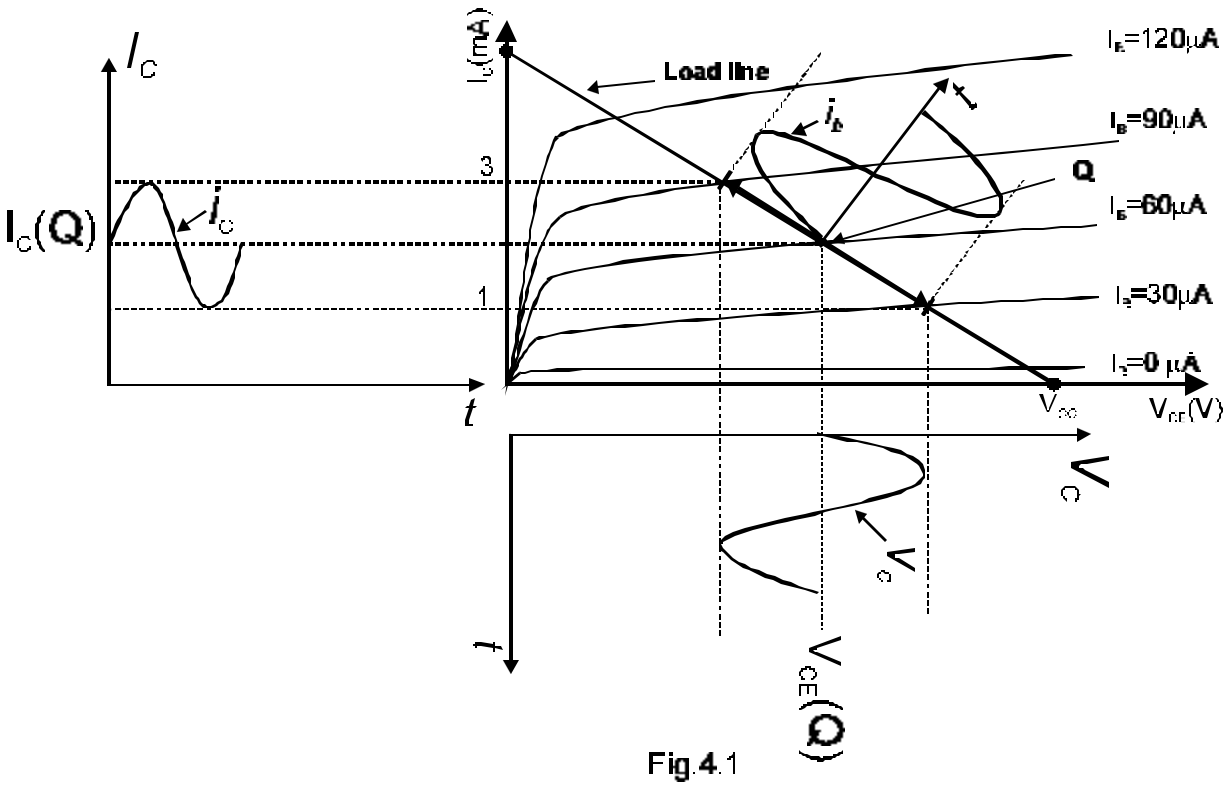



Fig.4.1

In figure 4.1 you can see a graphical analysis of the common emitter amplifier in dynamic regime. The input signal is applied in the base of the transistor, which has the static

operating point Q given by the intersection of load line with the static characteristic for the base current 60 μA. This case will result in a total output current/voltage signal $\boldsymbol{I_c}$ and $\boldsymbol{V_c}$ as a function of time ***t*** that can be seen on lower left panel of fig.4.1.

As you can see in Fig.4.1, a small input signal (current or voltage signal) is amplified by the transistor. The usual unit for the gain of an amplifier is the *decibel*, defined as:

$$\text{Number of decibels} = 10\log P_{out}/P_{in}$$

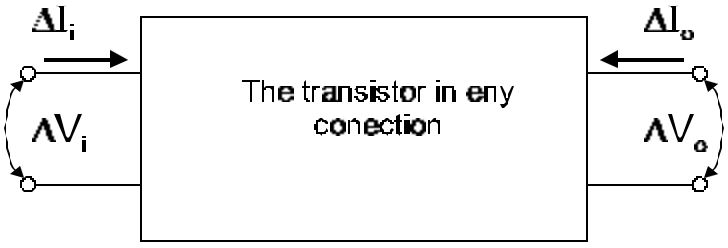
If the input and output power of an amplifier is measured on the same resistor, the definition for the decibel becomes:

$$\text{Number of decibels} = 10\log P_{out}/P_{in} = 10\log (v_{out})^2/(v_{in})^2 = 20\log v_{out}/v_{in}$$

But, in spite of not being technically correct, it has become customary to define the *decibel voltage gain* of an amplifier in terms of the voltage gain, even that the input and output resistances are not equals. Therefore, last formula becomes:

$$\text{Decibel voltage gain} = G_v = 20 \log A_v$$

**4.1 The Small Signals Model for Bipolar Transistor.**




The behaviour of four-terminal network, as a general class of circuits, can be characterised by linear equations *only for small signals*. This affirmation can be demonstrated using the Ebers-Moll equations of the transistor, written for total current (both components):

$$I_e = a_{11}\left(e^{\frac{V_E}{V_T}} - 1\right) + a_{12}\left(e^{\frac{V_C}{V_T}} - 1\right)$$

$$I_c = a_{21}\left(e^{\frac{V_E}{V_T}} - 1\right) + a_{22}\left(e^{\frac{V_C}{V_T}} - 1\right)$$

where $I_e$=$I_E$ + $i_e$ ; $V_e$= $V_E$ + $v_e$ *and* $V_c$=$V_C$ + $v_c$ according to convention adopted in the beginning of this chapter. For small signals, the amplitude of a.c. components is small enough to allow us to keep only the first order terms from the Taylor (power) series in which can be approximated the exponential of a.c. components. Let's take as an example the first Ebers-Moll equation:

$$e^{\frac{V_E+v_e}{V_T}} \cong e^{\frac{V_E}{V_T}}\left(1+\frac{v_e}{V_T}+\ldots\right) \text{ respectively } e^{\frac{V_C+v_c}{V_T}} \cong e^{\frac{V_C}{V_T}}\left(1+\frac{v_c}{V_T}+\ldots\right)$$

Grouping the terms of d.c. and a.c. components we will obtain the following equation:

$$I_E = a_{11}\left(e^{\frac{V_E}{V_T}}-1\right)+a_{12}\left(e^{\frac{V_C}{V_T}}-1\right)+a_{11}e^{\frac{V_E}{V_T}}\frac{v_e}{V_T}+a_{12}e^{\frac{V_C}{V_T}}\frac{v_c}{V_T}$$

where the first two terms represent just the d.c. component of the emitter current and the next two terms represent the a.c. component of the emitter current.

Now, by differentiating the above equation, we obtain:

$$\Delta i_e = y_{11}\Delta v_e + y_{12}\Delta v_c \qquad 4.1.1.$$

where $y_{11} = \frac{a_{11}e^{\frac{V_E}{V_T}}}{V_T}$ and $y_{12} = \frac{a_{11}e^{\frac{V_C}{V_T}}}{V_T}$

Using a similar method, we can obtain the equation for small variations of the collector current given by the second Ebers-Moll equation:

$$\Delta i_c = y_{21}\Delta v_e + y_{22}\Delta v_c \qquad 4.1.2.$$

The equations 4.1.1. and 4.1.2. define the so called "admittance parameters", determined using the next equations:

$$y_{11} = \left.\frac{\Delta i_e}{\Delta v_e}\right|_{\Delta v_c = 0}$$ , which represents the "input admittance"

$$y_{12} = \left.\frac{\Delta i_e}{\Delta v_c}\right|_{\Delta v_e = 0}$$ , which represents the “reverse transfer admittance”

$$y_{21} = \left.\frac{\Delta i_c}{\Delta v_e}\right|_{\Delta v_c = 0}$$ , which represents the “forward transfer admittance”

$$y_{22} = \left.\frac{\Delta i_c}{\Delta v_c}\right|_{\Delta v_e = 0}$$ , which represents the “output admittance”

The equivalent circuit described by relations 4.1.1 and 4.1.2 is drawing in Fig.4.2

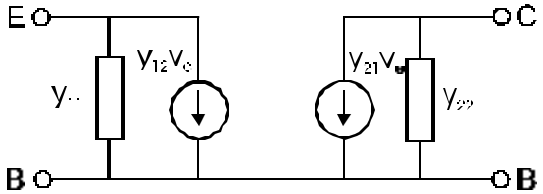



Fig.4.2

This “four-terminal network” represent the equivalent circuit for CBC of the bipolar transistor using admittance parameters. As you can see, the equivalent input circuit of the transistor according to the equation 4.1.1. comprises the input admittance $y_{11}$ and the constant current generator $y_{12}v_c$ which represents the influence (feedback) of the output circuit to the input circuit. Similarly, the equivalent output circuit comprises the output admittance $y_{22}$ and the constant current generator $y_{21}v_e$ which represents the influence of the input circuit to the output circuit.

In equations 4.1.1 and 4.1.2 we have taken as independent variables the input voltage $v_i$ and the output voltage $v_c$, by writing the input current $i$ and the output current $i_c$ using linear relations allowed by the small a.c. signal approximation.

Another transistor model can be found by taking as independent variables the input and output currents. In this case the modelling parameters are “impedances”, defined by the following relations:

$$z_{11} = \left.\frac{\Delta v_e}{\Delta i_e}\right|_{\Delta i_c = 0}$$ , which represents the “input impedance”

$$z_{12} = \left.\frac{\Delta v_e}{\Delta i_c}\right|_{\Delta i_e = 0}$$ , which represents the “reverse transfer impedance”

$$z_{21} = \left.\frac{\Delta v_c}{\Delta i_e}\right|_{\Delta i_c = 0}$$ , which represents the “forward transfer impedance”

$$z_{22} = \left.\frac{\Delta v_c}{\Delta i_c}\right|_{\Delta i_e = 0}$$ , which represents the “output impedance”

The equivalent circuit using impedance parameters is shown in Fig.4.3.

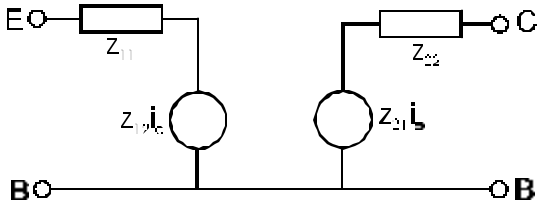



Fig.4.3

As in the case of admittance parameters model, this four-terminal network which modelled the CBC of bipolar transistor, has an input circuit made by the input impedance $z_{11}$ and a constant voltage generator $z_{12}i_c$, and an output circuit made by the output impedance $z_{22}$ and a constant voltage gene rator $z_{12}i_e$.

Then, the equations 4.1.3 and 4.1.4 will approximate the behaviour of the bipolar transistor:

$$\Delta v_e = z_{11}\Delta i_e + z_{12}\Delta i_c \qquad 4.1.3.$$

$$\Delta v_c = z_{21}\Delta i_e + z_{22}\Delta i_c \qquad 4.1.4.$$

Now, if we take as independent variables the input current and the output voltage, like in the case of static characteristics, the variations of the input voltage and output current can be written as follows:

$$\Delta V_i = h_{11} \Delta I_i + h_{12} \Delta V_o \qquad 4.1.5.$$

$$\Delta i_o = h_{21} \Delta I_i + h_{22} \Delta V_o \qquad 4.1.6.$$

where parameters $h_{ij}$ , named hibrid parameters, are defined by following relations:

$h_{11} = \left.\frac{\Delta V_i}{\Delta I_i}\right|_{V_o = ct.}$ represents the “input impedance”. Usually is noted by $\boldsymbol{h_i}$ .

$h_{12} = \left.\frac{\Delta V_i}{\Delta V_o}\right|_{I_i = ct.}$ represents the “reverse transfer factor”. Usually is noted by $\boldsymbol{h_r}$ .

$h_{22} = \left.\frac{\Delta I_o}{\Delta I_i}\right|_{V_o = ct.}$ represents the “forward transfer factor”. Usually is noted by $\boldsymbol{h_f}$ .

$h_{21} = \left.\frac{\Delta I_o}{\Delta V_o}\right|_{I_i = ct.}$ represents the “output admittance”. Usually is noted by $\boldsymbol{h_o}$ .

All these factors have a second index to characterise the transistor’s connection; the letter “e” for CEC of the tansistor, “b” for CBC of the transistor, respectively “c” for CCC of the transistor.

Equations 4.1.5 and 4.1.6 allow us to imagine a new four terminal network for the bipolar transistor, which is most used in electronics. This circuit, named “hybrid parameters model of the transistor” is shown in Fig.4.4

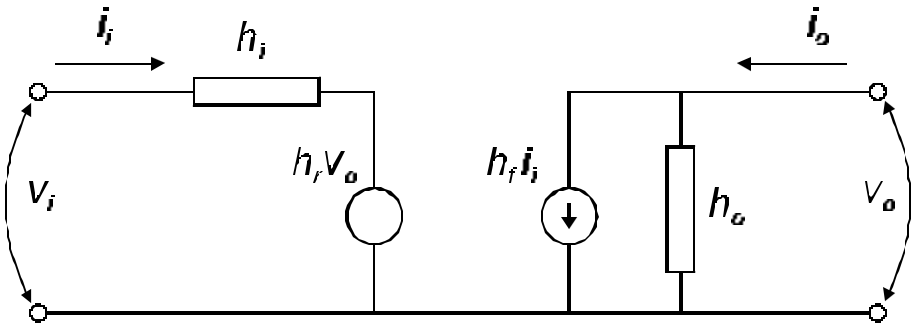



Fig.4.4

As you can see, this circuit is a mixture between “impedance parameters” and “admittance parameters” circuits, for that reason being named “hybrid parameters model”.

**Typical values for hybrid parameters.**

| Second index → / Parameter | (CEC) e | (CBC) b | (CCC) c |
|---|---|---|---|
| $h_i$ | $1.1 \times 10^3\ \Omega$ | 10-100 Ω | $20\text{-}50 \times 10^3\ \Omega$ |
| $h_r$ | $10^{-4}$ | $10^{-4}$ | 1 or less |
| $h_f$ | 100 (medium) | 0.99 (medium) | 100 (medium) |
| $h_o$ | $10^{-5}\ \Omega^{-1}$ | $10^{-7}\ \Omega^{-1}$ | $10 - 10^3\ \Omega^{-1}$ |

**4.2 General Characteristics of an Amplifier.**

Every amplifier is characterised by voltage amplification $A_v$ , current amplification $A_i$ , input impedance $z_i$ and output admittance $y_o$ . In order to be able to calculate these parameters it is necessary to transform the actual circuit in it’s a.c. equivalent. Here are two rules to be followed:

- **every capacitance is a short-circuit in a.c.**
- **the d.c. biasing source is a short-circuit to the ground in a.c.**

Let's take the most usual amplifier circuit using transistors in practical applications, the Common Emitter Connection amplifier, showed in Fig.4.5.

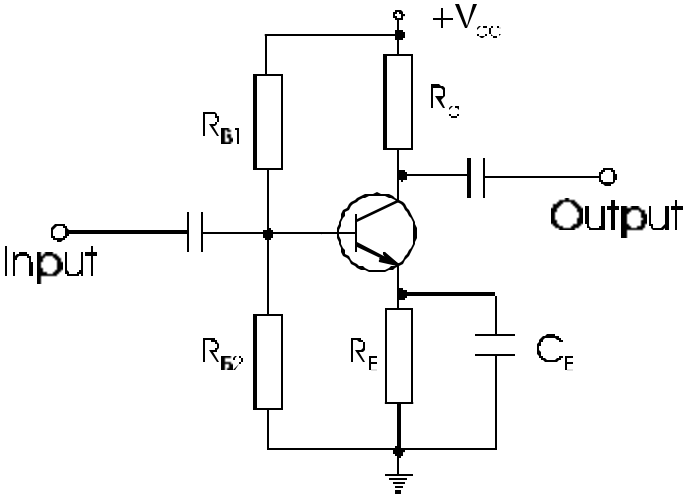


Fig.4.5

The equivalent a.c. circuit, taking into account the above rules, is shown Fig.4.6 . By dotted lines are represented in the circuit of Fig.4.5 the input signal source ($v_g$ and $R_g$ in the output circuit the load resistor ($R_L$). Now, in fig.4.6 we must replace the transistor with its equivalent hybrid circuit, in this particular case with the hybrid circuit for the CEC transistor. It will result the final equivalent circuit, showed in fig.4.7.

The resistor $R_B$ is the equivalent resistance of the voltage divider bias circuit (resistors $R_{B1}$ and $R_{B2}$ in parallel connection).

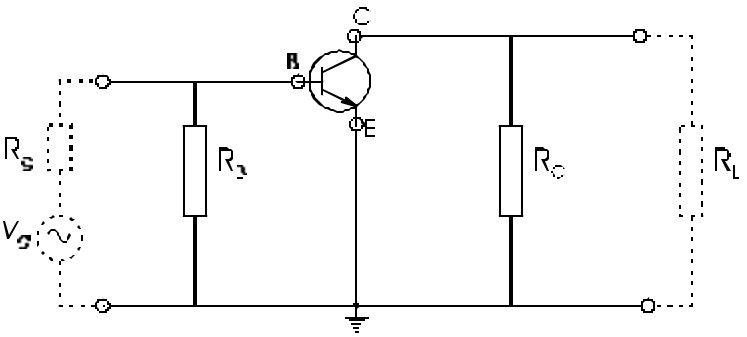


Fig.4.6

By definition, the current gain $\mathbf{A_i}$ is

$$A_i = \frac{i_L}{i_i} \qquad 4.2.1.$$

For a simplified calculation, in fig.4.7 we will take into account the load resistor as the equivalent resistor $R'_L=R_CR_L/(R_C+R_L)$, thus the fig.4.7 becomes fig.4.8.

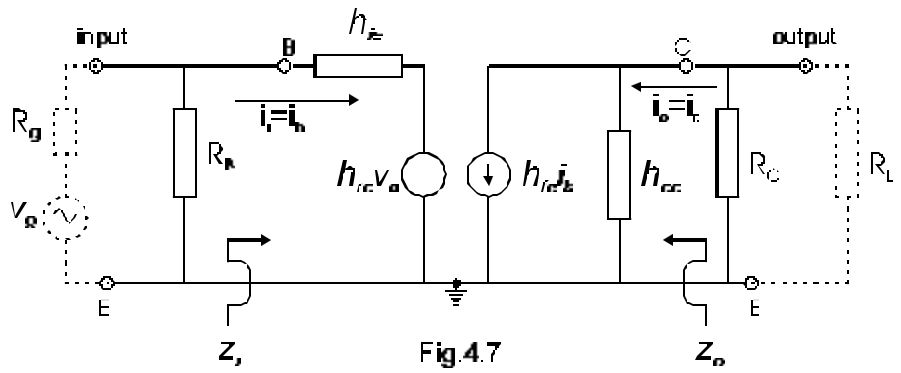


Fig.4.7

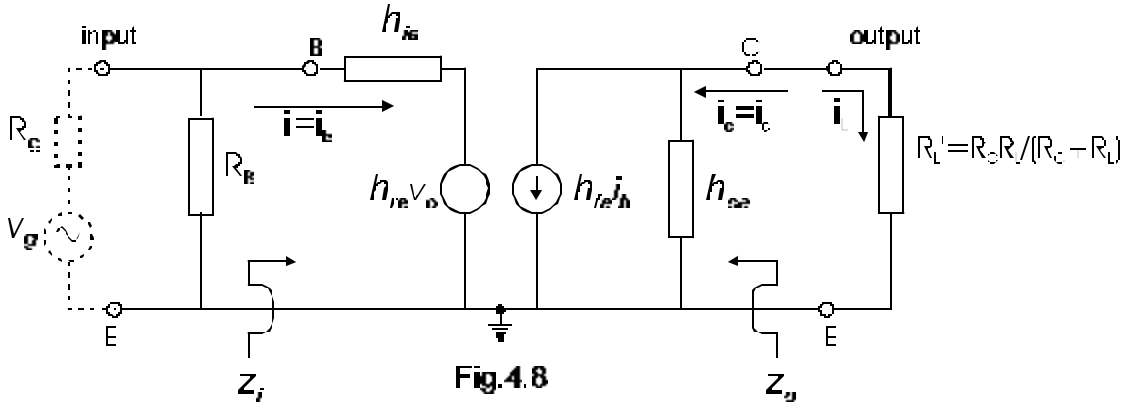



Fig.4.8

Then the current gain is:

$$A_i = -\frac{i_c}{i_b} = -\frac{h_{fe}i_b + h_{oe}v_o}{i_b} = -h_{fe} - h_{oe}\frac{i_L R_L'}{i_b} = -h_{fe} - h_{oe}A_i R_L' \qquad 4.2.2.$$

$$A_i = -\frac{h_{fe}}{1 + h_{oe}R_L'}$$

The input impedance is defined by

$$z_i = \frac{v_i}{i_i} \qquad 4.2.3.$$

then, from fig.4.8 we can calculate this impedance by applying the second Kirchhoff's law for the input circuit:

$$z_i = \frac{v_i}{i_i} = \frac{h_{ie}i_b + h_{re}v_o}{i_b} = h_{ie} + h_{re}A_i R_L' \qquad 4.2.4.$$

Now, we can calculate easily the voltage gain, which is defined by:

$$A_v = \frac{v_o}{v_i} \qquad 4.2.5.$$

then, in the same way used in last demonstrations, we will find:

$$A_v = \frac{v_o}{v_i} = \frac{i_L R_L'}{i_i z_i} = A_i\frac{R_L'}{z_i} \qquad 4.2.6.$$

The last parameter which we must know is the output admittance (impedance). This is defined by the next relation:

$$y_0 = \left.\frac{i_o}{v_o}\right|_{v_g=0} \qquad 4.2.7.$$

This parameter is defined in condition of input signal source in shot-circuit ($v_g$=0). Then

$$y_{oe} = \frac{h_{oe}v_o + h_{fe}i_i}{v_o} = h_{oe} + \left.\frac{h_{fe}i_i}{v_o}\right|_{v_g=0} \qquad 4.2.8.$$

But from input circuit applying second Kirchoff's Law we can write

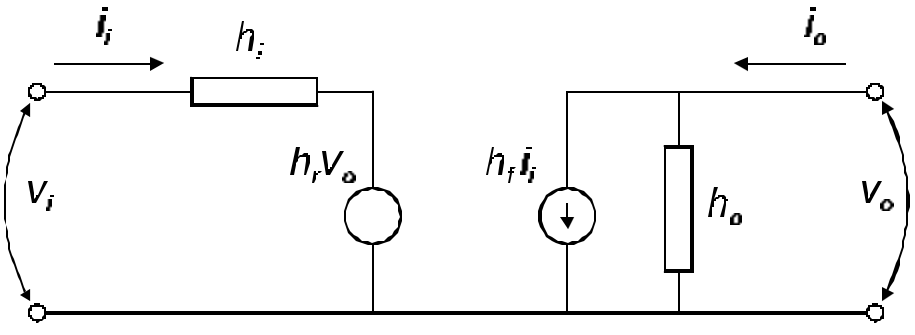
$$0 = R_g^{'}i_i + h_i i_i + h_r v_o \Rightarrow \left.\frac{i_i}{v_o}\right|_{v_g=0} = -\frac{h_{re}}{R_g^{'} + h_{ie}}$$

and if we replace this expression in 4.2.8. we will find

$$y_{oe} = h_{oe} - \frac{h_{fe}h_{re}}{R_g^{'} + h_{ie}} \qquad 4.2.9.$$

**4.3. The Simplified Hybrid Circuit for Bipolar Transistor.**

For an CEC amplifier having an equivalent circuit as the one depicted in the figure below we found earlier the following equations:




Current gain (4.2.2)

$$A_i = \frac{h_{fe}}{1 + h_{oe}R_L^{'}}$$

and input impedance (4.2.4.)

$$z_i = h_{ie} + h_{re}A_i R_L^{'}$$

In the case where we have satisfied the condition $h_{oe}R_L^{'} < 10^{-1}$ (in particular that means a maximum value for the load resistor of $10^4$ ohms), the second term from the denominator of $\mathbf{A_i}$

can be ignored and the current gain of the amplifier can be approximated by the hybrid factor $\mathbf{h_{fe}}$.

*That means, from a practical standpoint of view, that in the hybrid model of the transistor, we can neglect the output admittance* $\boldsymbol{h_{oe}}$ *.*

Now, if we take into account the actual typical values for the hybrid parameters in the equation for the input impedance ($\mathbf{h_{re}=10^{-4}}$ ; $\mathbf{A_i=h_{fe}=10^2}$ ; $\mathbf{R_L'=10^4\text{-}10^3}$ ; $\mathbf{h_{ie}=10^3}$), we will see that the second term in the expression of input impedance can have values in the range $10^2$-10. In this case we can ignore this term versus the first term which has a value 10 to 100 times higher.

*That means, from a practical standpoint of view, that in the hybrid model of the transistor, we can neglect the reverse transfer factor* $\boldsymbol{h_{re}}$ *.*

Taking into account these two approximations, the hybrid model of the transistor becomes more simple, like the circuit shown in Fig.4.9.

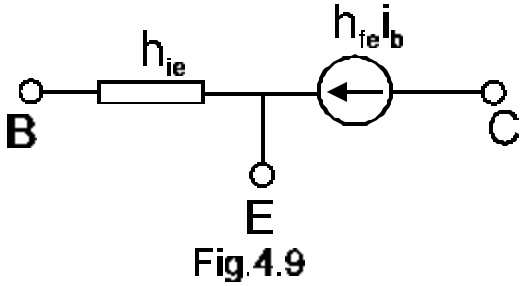


Fig.4.9

In the case of the CEC amplifier, using the simplified equivalent circuit in Fig.4.9, the general parameters of a common emitter amplifier are:

$A_i=h_{fe}$ ; $z_i=h_{ie}$ ; $y_{oe}=0$ ; $A_u=h_{fe}R_L'/h_{ie}$ 4.3.1

The error in calculating these parameters, using the simplified hybrid model, is around 4%. This error is less than the dispersion in the values of commonly used resistors, which makes it acceptable.

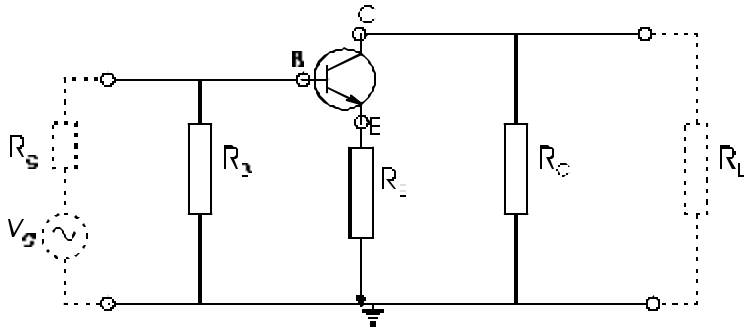



Fig.4.6 bis

A special case of common emitter amplifier is the *common emitter amplifier with emitter resistor*. In this case the $C_E$ capacitor doesn't exist. Then the equivalent a.c. circuit must take into account the presence of this resistor and Figure 4.6 becomes 4.6bis. Replacing the transistor now with his simplified hybrid circuit we will obtain figure 4.7bis.

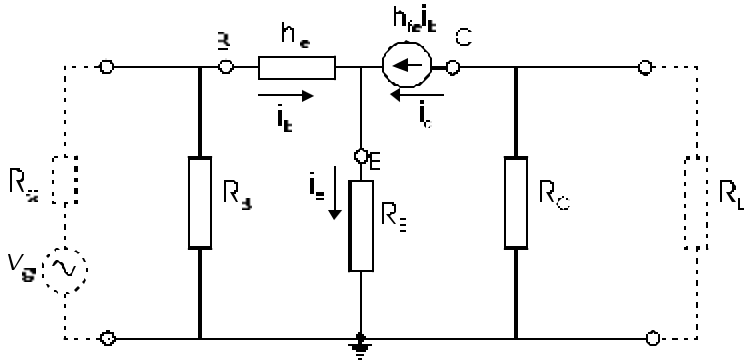



Fig.4.7 bis

The current gain remains the same like for common emitter amplifier, but the input impedance and voltage gain will be dramatically changed.

$v_i = h_{ie} \cdot i_b + R_E \cdot (i_b + i_c) = i_b [h_{ie} + R_E (1 + h_{fe})]$ and using the definition formula for input impedance we will obtain

$$z_i = h_{ie} + R_E (1 + h_{fe})$$

which has a higher value than the value of common emitter impedance.

Now, using the general formula for voltage gain we will obtain the new value of $A_v$.

$$A_V = -\frac{h_{fe} R_L^{'}}{h_{ie} + R_E (1 + h_{fe})} \approx -\frac{h_{fe} R_L^{'}}{R_E (1 + h_{fe})} \approx -\frac{R_L^{'}}{R_E}$$

As you can see the voltage gain, in this case, doesn't depend on the transistor performances, the amplifier having a stable voltage gain.

The simplified circuit shown in Fig.4.9 can be used too, for computing the parameters of any amplifier, independent of the type of the transistor connection. Let's test that right now.

**4.4. Common Base Amplifier.**

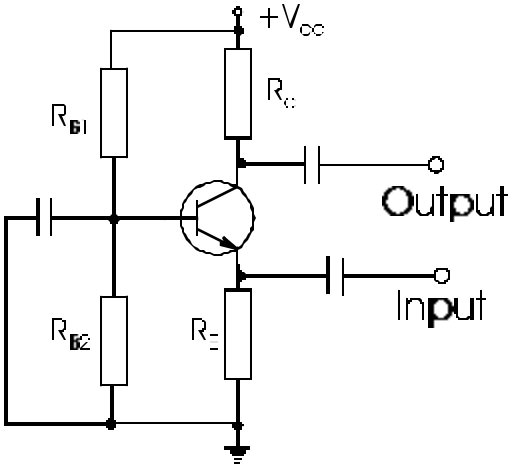


Fig.4.10

In Fig.4.10 is depicted the Common Base Amplifier with only one biasing source $+V_{CC}$ . To obtain the equivalent c.a. circuit we must follow the same rules as in the general case.

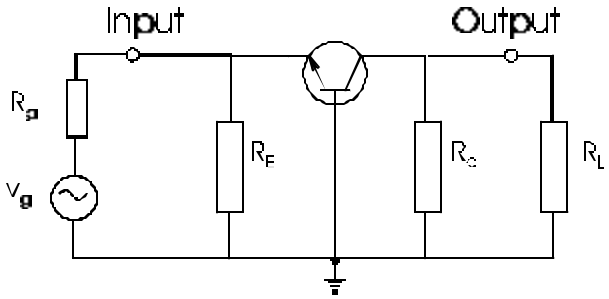


Fig.4.11

Using these rules we will obtain the equivalent a.c. circuit shown in Fig.4.11. If we replace the transistor with its simplified circuit from Fig.4.9, we will obtain the final equivalent circuit, shown in Fig. 4.12, from which we will be able to compute the general characteristics of this amplifier.

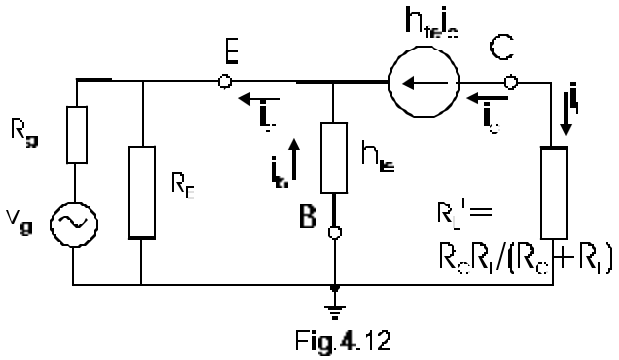


Fig.4.12

Because $i_i=-i_e=-(i_b+i_c)$ the current gain will be:

$$A_i = \frac{i_L}{i_i} = \frac{-i_c}{-i_e} = \frac{h_{fe} i_b}{i_b(1+h_{fe})} = \frac{h_{fe}}{1+h_{fe}} \qquad 4.4.1.$$

whose value is a little bit smaller than 1, which is in accordance with the definition of **a** factor.

The input impedance is:

$$z_i = \frac{v_i}{i_i} = \frac{h_{ie} i_b}{i_b(1+h_{fe})} = \frac{h_{ie}}{1+h_{fe}} \qquad 4.4.2.$$

which has a value around 10 ohms, smaller than the input impedance of CEC amplifier.

The voltage gain is similar with $\mathbf{A_v}$ of CEC amplifier, and output admittance is more close to zero than in the case of CEC amplifier because $\mathbf{h_{ob}=10^{-7}\ w^{-1}}$, a value two orders of magnitude lower than $\mathbf{h_{oe}}$.

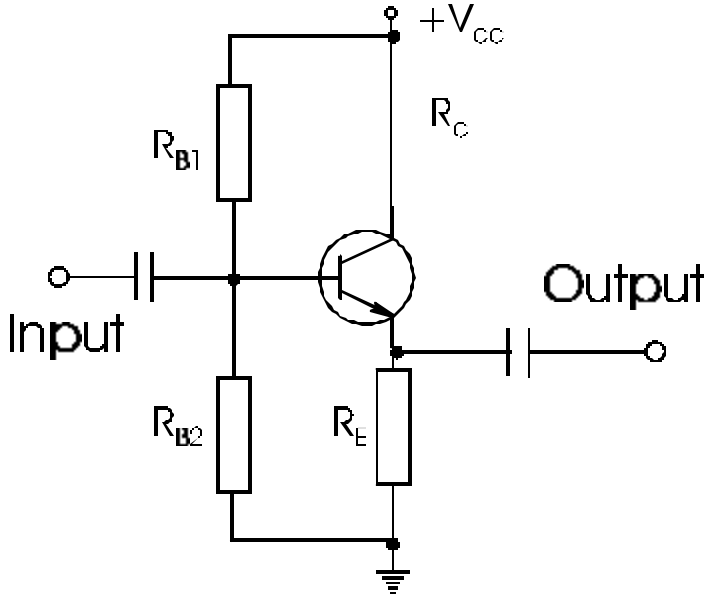


Fig.4.13

**4.5. Common Collector Amplifier.**

This amplifier is shown in Fig.4.13. The differences between this amplifier and CEC amplifier are: the absence of a resistor in the collector circuit and the output point in the emitter of the transistor. Now, by applying the rules for a.c. equivalent circuit and replacing the transistor with its simplified hybrid circuit, we will find the circuit drown in Fig. 4.14.

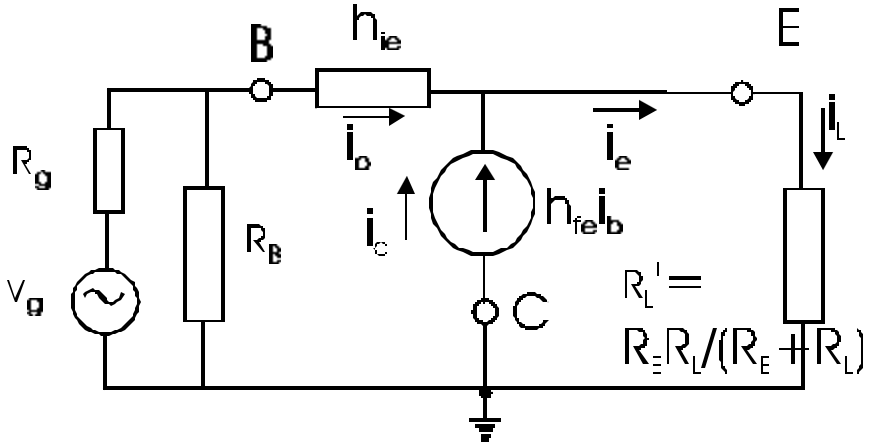


Fig.4.14

In this case the current gain has next formula:

$$A_i = \frac{i_L}{i_i} = \frac{i_e}{i_b} = 1 + h_{fe}$$

4.5.1.

The input impedance can be calculated from two standpoints of view, the input impedance of transistor $\mathbf{z_{iT}}$ and the input impedance of amplifier $\mathbf{z_{iA}}$.

$$z_{iT} = \frac{v_i}{i_{iT}} = \frac{h_{ie} i_b + i_L R_L}{i_b} = h_{ie} + (1 + h_{fe}) \cdot R_L' \approx h_{ie} + (1 + h_{fe}) \cdot R_E \qquad 4.5.2.$$

where we assumed that the load resistor is much higher than the $R_E$ , therefore we approximated $R_L'$ as $R_E$ .

$$z_{iA} = R_B \left\| z_{iT} \right. = \frac{R_B z_{iT}}{R_B + z_{iT}} \qquad 4.5.3.$$

The value given by the formula 4.5.3 is lower, in general, than the value given by formula 4.5.2. This is the so called **"problem of the biasing circuit"** in the case of CCC amplifier.

The voltage gain of this amplifier is given by following formula:

$$A_v = \frac{v_o}{v_i} = \frac{i_L \cdot R_E}{i_b \cdot [h_{ie} + (1 + h_{fe}) \cdot R_E]} = \frac{(1 + h_{fe}) \cdot R_E}{h_{ie} + (1 + h_{fe}) \cdot R_E} \leq 1 \qquad 4.5.4.$$

As you can see, the voltage gain is almost equal with **1**. This means that the amplitude of the output signal is the same with the amplitude of the input signal. This is the reason for which this amplifier is called "Emitter Follower".

The output admittance (impedance) can be calculated based on the following definition:

$$y_o\Big|_{v_g = 0} = \left. \frac{i_o}{v_o} \right|_{v_g = 0} \qquad 4.5.5.$$

but $i_o = -i_b (1 + h_{fe})$, and in the particular condition $\mathbf{v_g=0}$ , we have the next relation $0 = (R_g' + h_{ie}) i_b + v_o$ , then results :

$$y_0\Big|_{v_g = 0} = \frac{1 + h_{fe}}{R_g' + h_{ie}} \qquad 4.5.6.$$

where $\mathbf{R_g'=R_gR_B/(Rg+R_B)}$ **.** Then the output impedance is:

$$z_o\Big|_{v_g = 0} = \frac{R_g' + h_{ie}}{1 + h_{fe}} \qquad 4.5.7.$$

The output impedance is very low as compared to the input impedance (four orders of magnitude). For this reason CCC amplifier is used like *impedance adapter.*

**The Problem of the Biasing Circuit for CCC Amplifier.**

We defined in last section the input impedance of Common Collector Amplifier as:

$$z_{iA} = R_B \| z_{iT} = \frac{R_B z_{iT}}{R_B + z_{iT}}$$

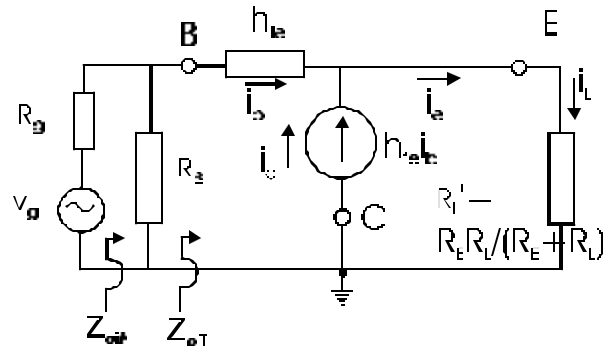



Fig.4.15

As we said, the amplifier input impedance (see figure 4.15), is modified by the presence of biasing resistor $\mathbf{R_B}$ , as in the above formula. If the resistor $\mathbf{R_B}$ has a lower value than the input impedance of the transistor, the input amplifier impedance can be dramatically changed as compared to the input transistor impedance alone. To prevent this negative influence of the biasing circuit, we must modify it in such a way as to provide the same d.c. biasing current, but have a higher a.c. impedance. The modified circuit is shown in Fig 4.16.

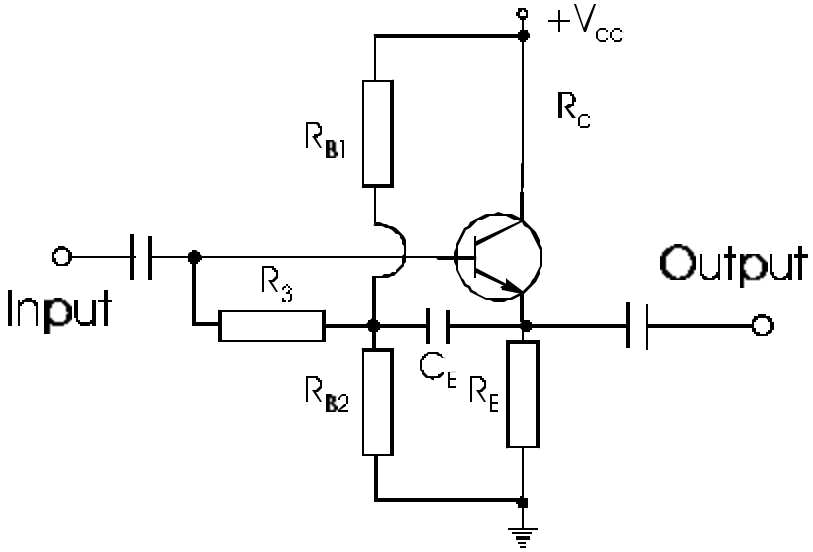



Fig.4.16

The a.c. equivalent value of resistor $\mathbf{R_3}$ is given by Miller's theorem

$$R_3' = \frac{R_3}{1 - A_v}$$

For Common Collector Amplifier an usual value for the voltage gain is **0.999** , then the value of $\mathbf{R_3'}$ becomes few orders of

magnitude bigger than the value of input impedance of the transistor. In this case the input impedance of the amplifier is determined mainly by the value of the transistor input impedance. The equivalent a.c. circuit of amplifier shown in Fig.4.16 is presented in Fig.4.17

The current which flows through the resistor $R_3$ can be calculated using the following formula:

$$i_3 = \frac{V_i - V_o}{R_3} = \frac{V_i\left(1 - \frac{V_o}{V_i}\right)}{R_3} = \frac{V_i}{\frac{R_3}{1 - A_V}} = \frac{V_i}{R'_3}$$

then, the resistor $R_3$, connected between input and the output of the amplifier, via $C_E$ capacitor, is equivalent with a $R'_3$ resistor in series with $R_B$ equivalent resistor, which has an increased value only for a.c. signals.

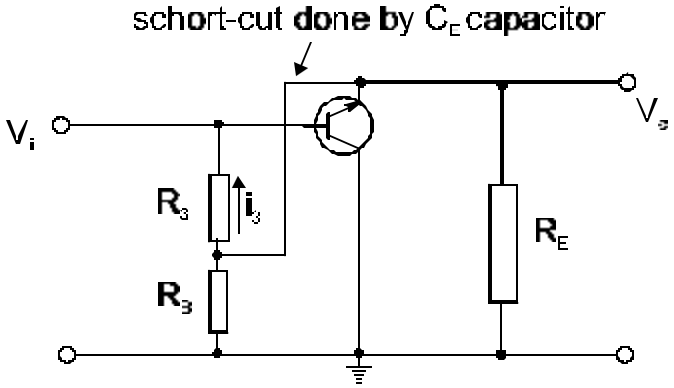



Fig.4.17

**4.6. The DARLINGTON Pair.**

The Darlington pair is obtained by directly connecting two transistors as shown in Fig.4.18. We will exemplify its use in a CCC amplifier.

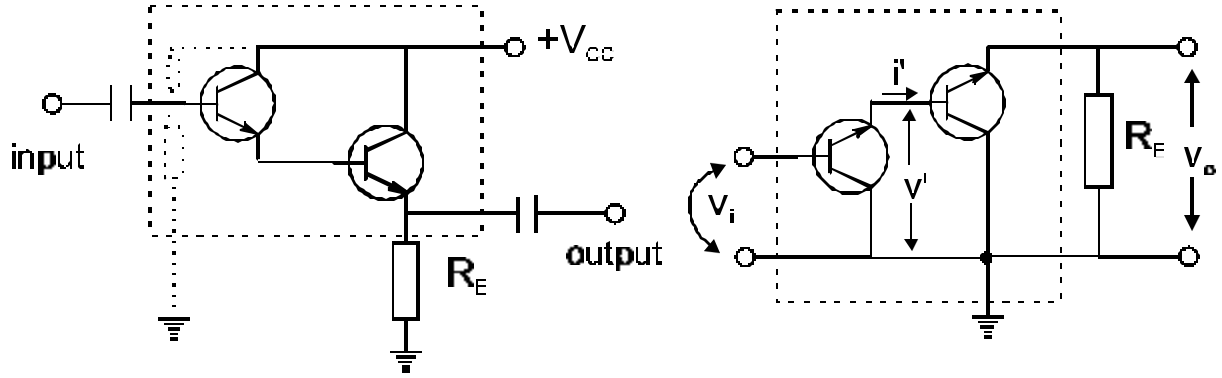



Fig.4.18

From the a.c. equivalent circuit, using the simplified hybrid model we can calculate the performances of this amplifier.

$$A_i = \frac{i_L}{i_i} = \frac{i'}{i_i}\frac{i_L}{i'} = A_{i_1}A_{i_2} \approx (1+h_{fe})^2 \qquad 4.6.1.$$

$$z_i = h_{ie} + (1+h_{fe})z' \approx (1+h_{fe})^2 R_E \qquad 4.6.2.$$

$$A_v = \frac{v_0}{v_i} = \frac{v'}{v_i}\frac{v_o}{v'} = A_{v_1}A_{v_2} \le 1 \qquad 4.6.3.$$

$$z_o = \frac{h_{ie} + z'}{1+h_{fe}} = \frac{h_{ie} + \dfrac{h_{ie} + R_g'}{1+h_{fe}}}{1+h_{fe}} \qquad 4.6.4.$$

The most important conclusions are: Darlington pair amplifier has higher current gain and input impedance than the Common Collector Amplifier. The pair is sometimes called a supertransistor whose current transfer factor $h_{fe}$ is the product of the individual current transfer factors.

**4.7. The CASCODE Amplifier.**

This amplifier is obtained by directly coupling two transistors, one in Common Emitter Connection and the second in Common Base Connection as in Fig.4.19

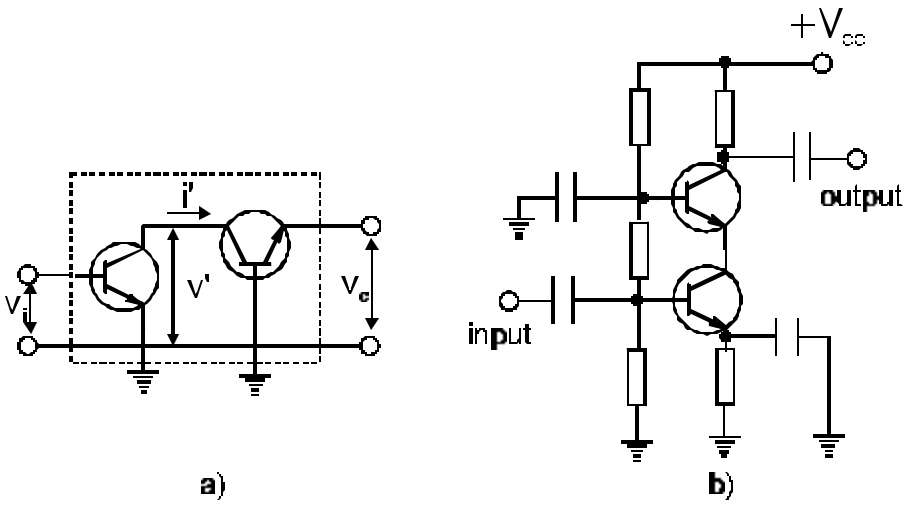



Fig.4.19

$$A_i = \frac{i_L}{i_i} = \frac{i'}{i_i}\frac{i_L}{i'} = h_{fe}h_{fb} \approx h_{fe} \qquad 4.7.1.$$

$$z_i \approx z_{ie} = h_{ie} \qquad 4.7.2.$$

$$h_{re} = \frac{v_i}{v_o} = \frac{v_i}{v'}\frac{v'}{v_o} \approx h_{re}h_{rb} \approx 10^{-8} \qquad 4.7.3.$$

The equivalent a.c. circuit of the CASCODE amplifier draw in figure 4.19b is shown as a four terminal network in the figure 4.19a.

The most important conclusion is: the reverse transfer factor is 4 orders of magnitude lower than in the case of Common Emitter Amplifier or Common Base Amplifier. Then this amplifier has the lowest reverse transfer factor.

**4.8. The Differential Amplifier.**

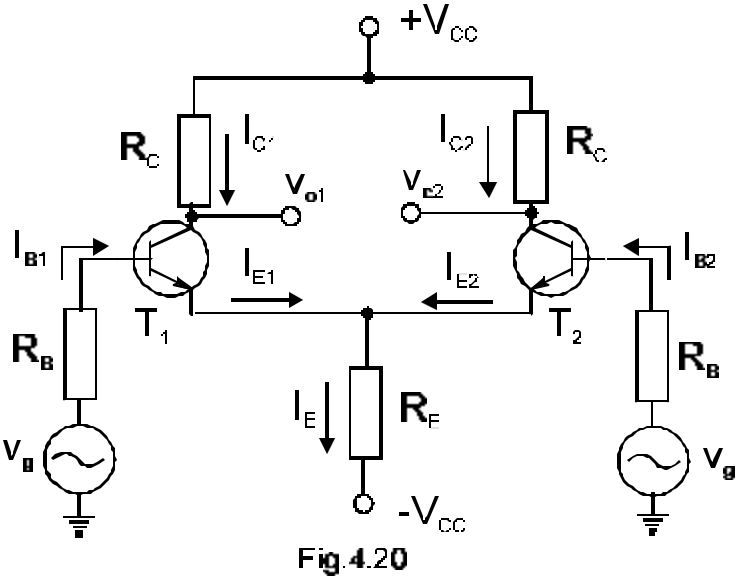



Fig.4.20

This amplifier is obtained by connecting two transistors, with very similar static characteristics, as in the Fig.4.20.

From second Kirchhoff's Law we can write the next equation:

$$2V_{CC} = R_C I_{C1} + V_{CE_1} + R_E I_E$$

If we have satisfied the condition

$$R_C I_{C1} << R_E I_E ,$$

we can approximate the current which flows through resistor $R_E$ as an **constant current**. This assertion is true too, for the dynamic component of this current. If this current is constant, we can conclude that when its component $I_{E1}/i_{e1}$ increases, the component $I_{E2}/i_{e2}$ must decrease.

The relation between input and output signals must be linear in the case of small signals. Therefore :

$$\begin{aligned} v_{01} &= A_{11} v_{i1} + A_{12} v_{i2} \\ v_{02} &= A_{21} v_{i1} + A_{22} v_{i2} \end{aligned} ,$$

but due to the symmetry of the circuit, $A_{11} \approx A_{22} \approx A_1$ and $A_{12} \approx A_{21} \approx A_2$.

Now, if we define the differential input signal as $v_{id} = v_{i1} - v_{i2}$ , and the common input signal as $v_{ic} = \frac{v_{i1} + v_{i2}}{2}$, by decomposing the input signals in the differential input signal and the common input signal and replacing these new values in the definition of output signals, we will obtain:

$$v_{o1} = \frac{A_1 - A_2}{2} v_{id} + (A_1 + A_2) v_{ic} \qquad 4.8.1.$$

$$v_{o2} = \frac{A_2 - A_1}{2} v_{id} + (A_1 + A_2) v_{ic} \qquad 4.8.2.$$

Now, we can define the differential output signal as :

$$v_{od} = v_{o1} - v_{o2} = (A_1 - A_2) v_{id} \qquad 4.8.3.$$

and the common output signal as:

$$v_{oc} = \frac{v_{o1} + v_{o2}}{2} = (A_1 + A_2) v_{ic} \qquad 4.8.4.$$

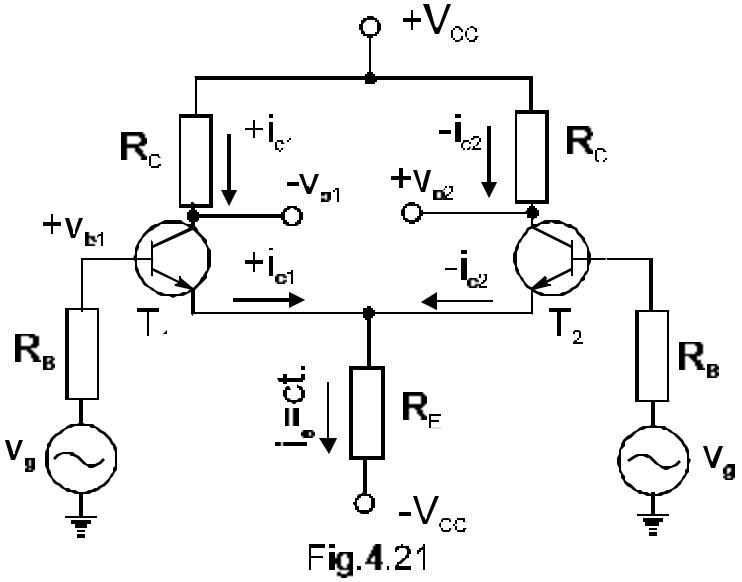


Fig.4.21

As you can see in fig.4.21, the dynamic components of currents follow the rule of constant current through the resistor $R_E$ , then if the input signal in the base of $T_1$ increases (+) the $T_1$ emitter current will increase, too, and consequently the emitter current of $T_2$ will decrease. That drives an increase of the output signal of $T_2$ . Like in a mirror, basically the same thing, will happen if the input signal of $T_2$ will increase.

Then we can conclude that the output of $T_2$ is in phase with the input of $T_1$ , and the output of $T_1$ is in opposite phase with its input.

Let be now the input signals in opposite phase, having following values :

$v_{i1} = \frac{v_g}{2}$ and respectively $v_{i2} = -\frac{v_g}{2}$ , then the relations 4.8.3 and 4.8.4 become

$$v_{od} = v_{o1} - v_{o2} = (A_1 - A_2)v_{id} = A_d v_g$$

$$v_{oc} = \frac{v_{o1} + v_{o2}}{2} = (A_1 + A_2)v_{ic} = 0$$

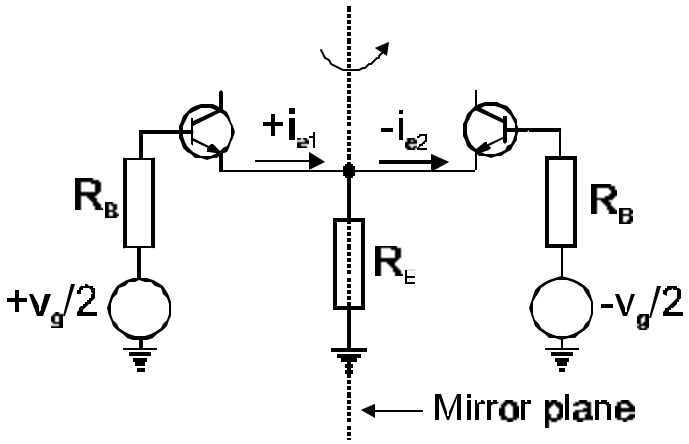



Fig.4.22

That will drive the amplifier in pure "differential" mode (no common-mode input). The dynamic components of emitter currents will be as shown in the figure 4.22.

Then the total a.c. current that flows through the resistor $R_E$ is zero. That means the emitters have constant potential, as being grounded, from the a.c. standpoint of view. Therefore, both transistors operate as a common emitter amplifier, the equivalent a.c. circuit being showed in fig.4.23.

From Fig.4.23 we can compute the gain for the differential amplifier, taking into account that in fact we have two transistors:

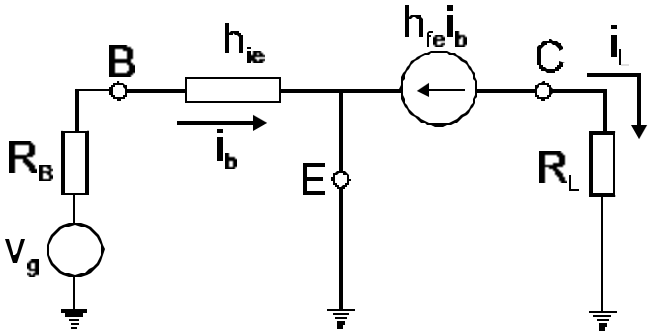



Fig.4.23

$$2A_d = \frac{v_o}{v_i} = \frac{-h_{fe}R_L}{R_B + h_{ie}} \qquad 4.8.5.$$

Let's consider now a pure common mode input signal, with no differential component. The input signals in phase having next values:

$v_{i1} = v_g$ and respectively $v_{i2} = v_g$ . In this case the 4.8.3 and 4.8.4 relations become:

$$v_{od} = v_{o1} - v_{o2} = (A_1 - A_2)v_{id} = 0$$

$$v_{oc} = \frac{v_{o1} + v_{o2}}{2} = (A_1 + A_2)v_{ic} = A_c v_g$$

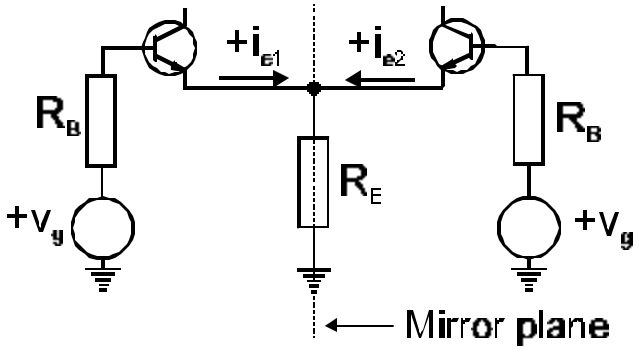


Fig.4.24

As you can see, in this case the total emitter current that flows through the resistor $R_E$ is $2i_{e1}$ , then the dynamic components of the emitter currents being as in Fig.4.23.

Then the a.c. equivalent circuit will be much like a "common emitter" amplifier with a resistor $2R_E$ connected in the emitter (see Fig.4.25)

Then the $A_c$ gain (common mode gain) is:

$$A_c = \frac{v_o}{v_i} = \frac{-h_{fe}R_L}{h_{ie} + R_B + 2R_E(1 + h_{fe})} \qquad 4.8.6$$

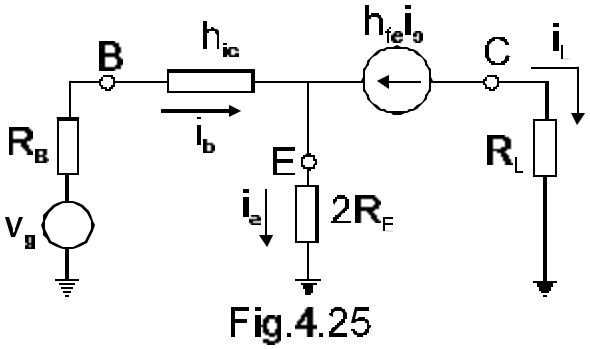


Fig.4.25

As you can see, the common mode gain is lower than differential gain. In general is desired to have this common mode gain as lower as possible. The ratio $A_d/A_c$ is called Common Mode Rejection Ratio (CMRR) and serves as a figure of merit to measure how close is the actual amplifier to an ideal differential amplifier, that is considered to have $A_c$=0 and infinite CMRR. In view of further increasing the CMRR, we can use constant current source, that has very high output impedance in a.c. instead of the $R_E$ resistor. A transistor in the common base configuration is known to have the highest output impedance of all three possible configurations. Its input is

shorted to ground in a.c. (the capacitor connected to its base), whereas its collector is connected to the emitters of the transistors that form the differential amplifier, as in fig.4.26.

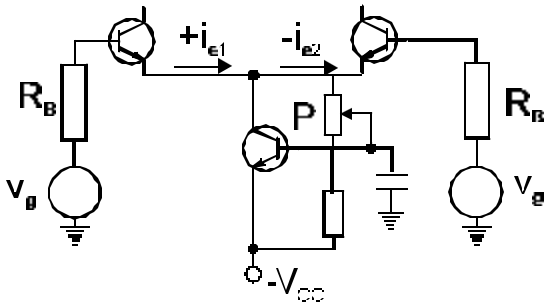


Fig.4.26

In this case the value of $R_E$ resistor is substituted by the output impedance of common base amplifier, which is in the order of megaohms. This will bring the common mode gain close to zero. The potentiometer P is used to set the desired d.c. emitter current of the differential amplifier, and bring the $T_1$ and $T_2$ transistor in the optimal operating point.

# Chapter 5. Power Amplifiers.

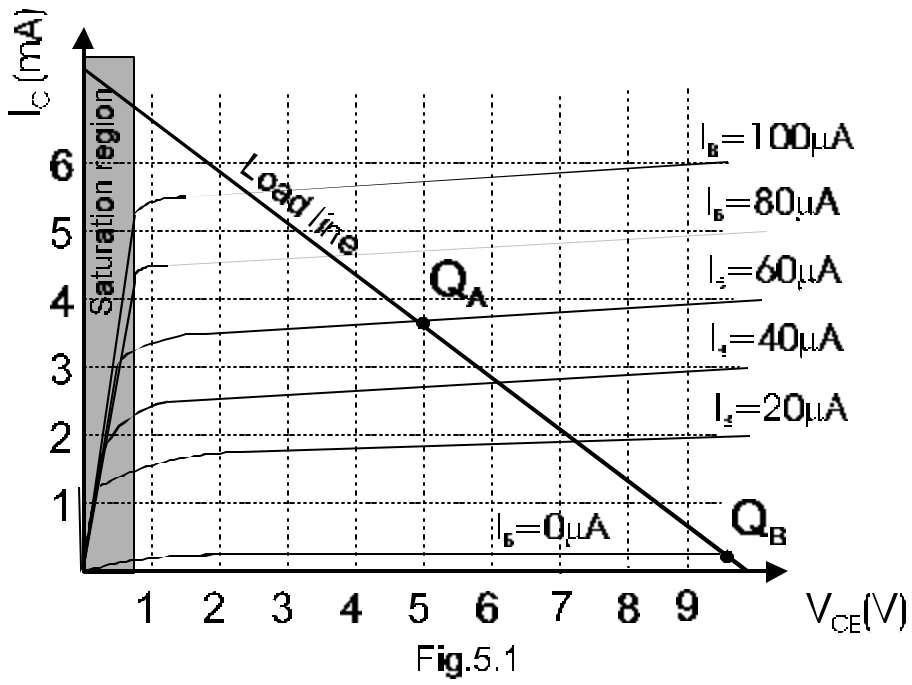


Fig.5.1

There are two main types of power amplifiers: class A power amplifiers that have the operating point at the middle of the load line, and class B power amplifiers, which have the operating point close to the cut-off limit of the transistor (see Fig.5.1)

The first problem of these amplifiers is the signal distortions, or in terms of Fourier spectrum, the problem of second harmonics generation. This problem originates from the fact that these amplifiers work with large input signals, and in this case the hybrid parameters do not remain constant. Therefore the output signal is not a linear function of the input signal, and can be written as a power series:

$x_O = G_1 x_i + G_2 x_i^2 + \ldots$ where $x_o$ is the output signal and $x_i$ is the input signal. All the power stages discussed in this chapter are "current amplifiers" and therefore the output signal is the collector current or the emitter current, which is about the same thing ($I_C \approx I_E$) when the base current can be neglected. Let be the input signal a periodic function: $x_i = X_m \cos \omega t$ . In this case, the dynamic component of the output current will be

$$i_c = G_1 X_m \cos \omega t + G_2 X_m^2 \cos^2 \omega t + \ldots \qquad 5.1.$$

But we can replace in this equation $\cos^2 \omega t = \frac{1 + \cos 2\omega t}{2}$, which then it becomes:

$$i_c = G_1 X_m \cos\omega t + G_2 X_m^2 \frac{1+\cos 2\omega t}{2} + \ldots = B_0 + B_1 \cos\omega t + B_2 \cos 2\omega t + \ldots \quad 5.2.$$

where $B_i$ are constant coefficients. You can see now, from equation 5.2. that the output signal has the second harmonics of the input signal, too. This handicap of such amplifiers becomes important only for high-power amplifiers. In the case of low power amplifiers, such as the Class A amplifiers, this disadvantage is not important since the ratio $B_2/B_1 << 1$.

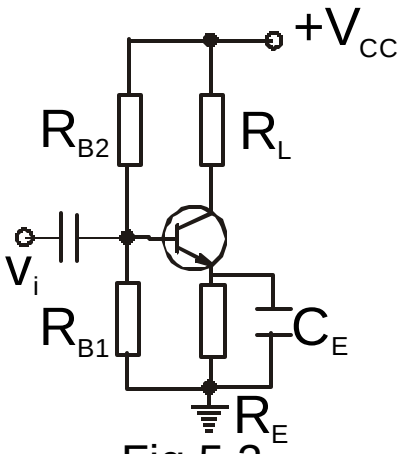


Fig.5.2

A class A amplifier is shown in Fig.5.2. Such amplifier is a typical common emitter (or a common collector amplifier), with the difference that the load resistor $R_L$ has much lower impedance (a few ohms) than in a standard amplifier. Accordingly, the currents flowing through the circuit are much larger.

The main drawback of the class A amplifier is that in order to insure a maximal output voltage swing, the operating point has to be approximately the midpoint of the load line (see fig.5.3). That means that the quiescence current $I_Q$ is approximately $I_Q = 0.5*(V_{max}-V_{min})/R_L$, which is half of the maximum current. The efficiency of the class A amplifier, will be the ratio between the maximum signal output power and the average DC power consumption of the amplifier:

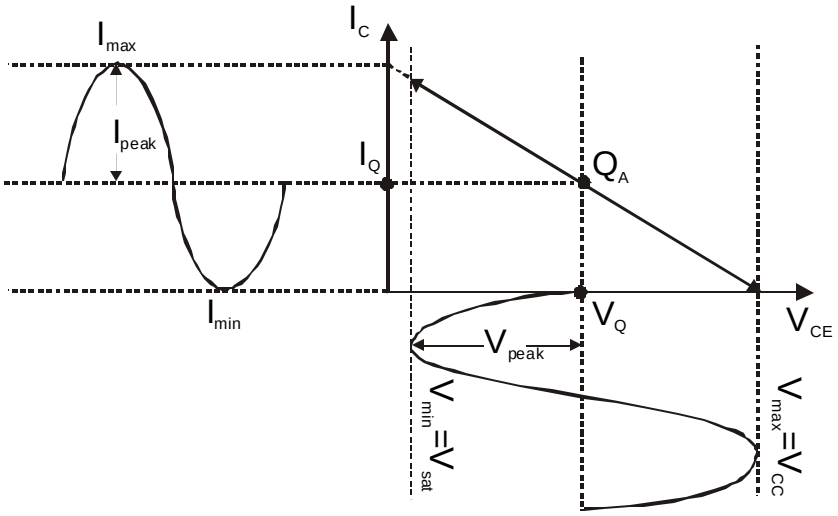


Fig.5.3

$$\eta_A = \frac{P_{out}}{P_{DC}} = \frac{I_{RMS}^{out} V_{RMS}^{out}}{I_Q V_{CC}} = \frac{\frac{I_{peak}}{\sqrt{2}} \frac{V_{peak}}{\sqrt{2}}}{I_Q V_{CC}} ,$$

but, $I_{peak} \approx I_Q$ , $V_{CC} = V_{max}$ , and $V_{peak} = \frac{V_{max} - V_{min}}{2}$ ,

therefore

$$\eta_A = 25 \frac{V_{max} - V_{min}}{V_{max}} \%$$

Since $V_{max} >> V_{min}$ , the efficiency is close to 25% only, that means that 75% of the power is dissipated in form of heat on the circuit.

**Class B power amplifier.**

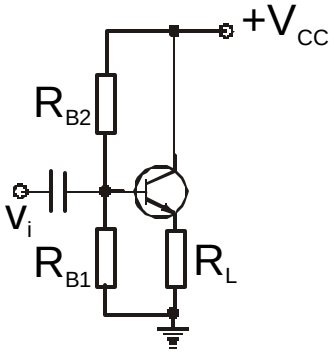


Fig.5.4

The main problem of the class A amplifier is the large quiescence current. By biasing the transistor near cut-off, as shown in fig 5.5, this current can be made almost null. This is the called a Class B operation. Such an amplifier with bipolar transistor in Common Collector Configuration is shown in fig. 5.4. However, when input is a sine waveform, only its positive alternation will drive the transistor in the active region, whereas the negative alternation will bring the transistor into even deeper cut-off. As we will see later on, this suppression of half of the signal can be avoided by pairing two complementary transistors in such a way that each one is active on an opposite signal alternation. The efficiency of a Class B amplifier is:

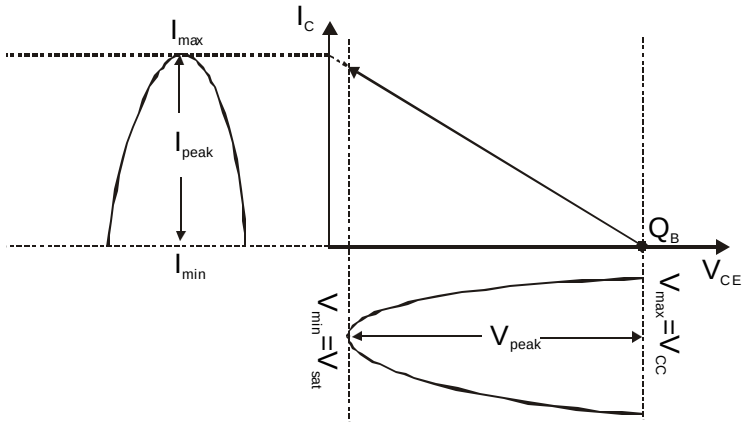


Fig.5.5

$$\eta_B = \frac{P_u}{P_c} = \frac{\frac{1}{2} I_{RMS}^{out} V_{RMS}^{out}}{I_{avg} V_{CC}} = \frac{\frac{1}{2} \frac{I_{peak}}{\sqrt{2}} \frac{V_{peak}}{\sqrt{2}}}{I_{avg} V_{CC}}$$

The factor 1/2 from the nominator is due to the fact that the transistor is active for a single alternation only. For a similar reason, the average DC supply current (see the diode rectifier, presented earlier in this book) is $I_{avg} = \frac{I_{peak}}{\pi}$. Also $V_{CC} \approx V_{max}$, therefore the efficiency of the Class B amplifier becomes:

$$\eta_B = \frac{P_u}{P_c} = \frac{\frac{1}{2} \frac{I_{peak}(V_{max} - V_{min})}{2}}{\frac{I_{peak} V_{max}}{\pi}} = \frac{\pi}{4} \frac{V_{CC} - V_{min}}{V_{CC}}$$

For circuits in which $V_{min} << V_{CC}$, the efficiency is close to 78%, which is a much better value than the efficiency of Class A amplifiers.

However, a simple Class B amplifier that is active during one signal alternation only, suppressing the other one, is of very little use, since it introduces a major distortion of the original signal. In order to have the output active for both alternations while preserving the Class B operating point, we can pair two transistors in such a way that they are active on opposite alternations, in a so called "push-pull" configuration, shown in fig 5.6.

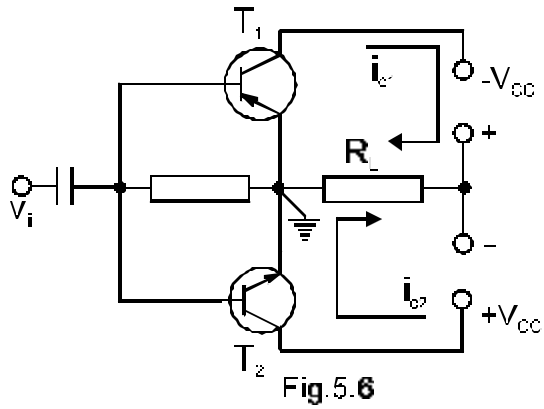


Fig. 5.6

In this case the negative alternation is amplified by $T_1$ transistor, because it is a pnp transistor, and the positive alternation is amplified by $T_2$ transistor because it is a npn transistor, and all the voltages and current flows have opposite polarity.

In all figures, the resistor $R_L$ designates the load resistor that may be in the real world a speaker, for instance.

**Crossover distortion.**

The simplified push-pull circuit in fig. 5.6 is missing the bias resistors in the bases of the transistors (as are $R_{B1}$ and $R_{B2}$ in fig 5.4), therefore with zero input signal both transistors are at cut-off. This is the part of the desired Class B operating regime, and would not hamper too much amplifier's operation, unless an arbitrarily small signal would take the transistors out of the cut-off region. However, in the circuit in fig. 5.6, the signal has to be larger than the threshold voltage $V_{be}=V_{\gamma}$ of the base-emitter junction, which is typically 0.6 - 0.7 V for a silicon bipolar transistor. The result is a symmetrical distortion of the input signal, as shown in fig. 5.7.

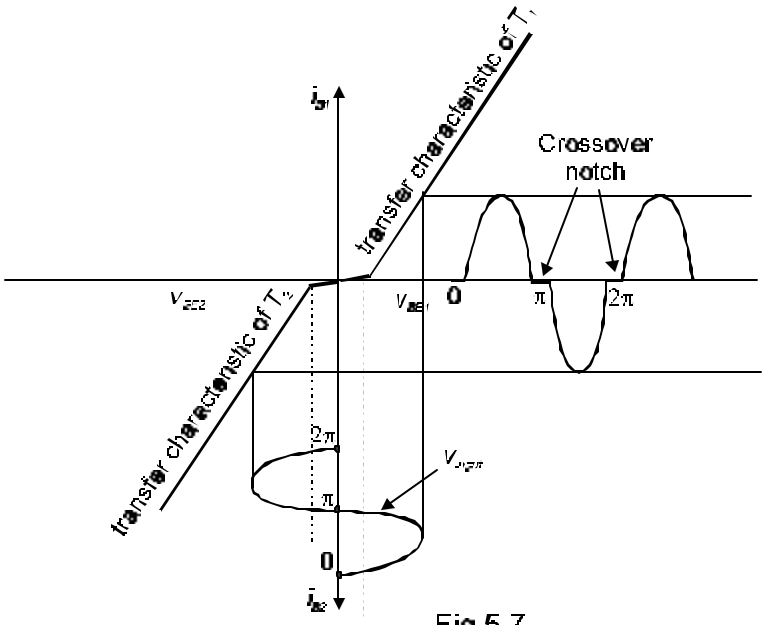



Fig.5.7

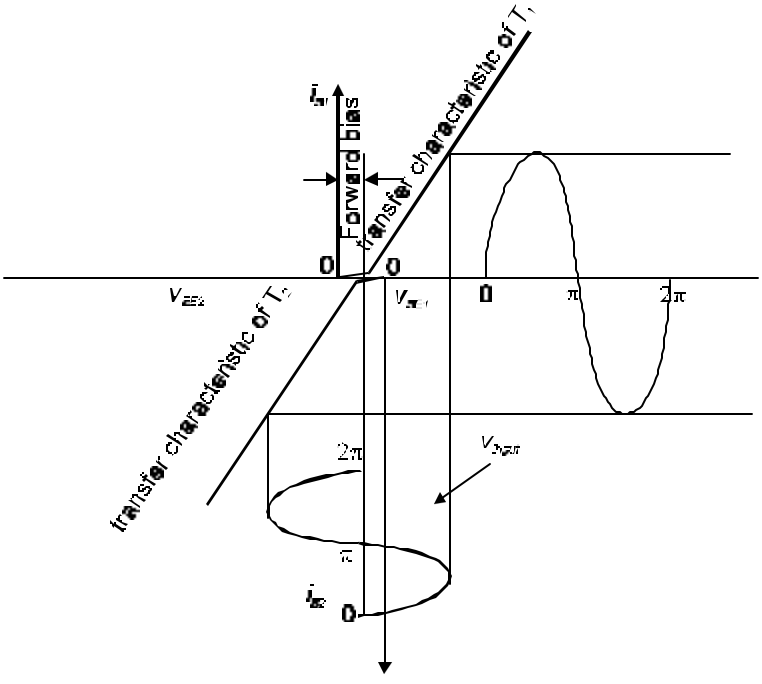



Fig.5.8

The transfer characteristic of the second transistor $T_2$ is plotted along the negative axis and reversed with respect to the characteristic of first transistor $T_1$. The input signal is composed of two opposite alternations, each one driving a

different transistor. The distortion notch near the crossover point on the current axis, due to the 0.5-0.6 V voltage threshold required to bring the transistor into normal operation, is called *crossover distortion*. And it is more important at low levels of the input signal.

This kind of distortion can be prevented by introducing a biasing voltage on the bases of the transistors in such a way that cut-off is not completely reached at point Q, and the transfer characteristics will look as in the Fig.5.8.

In this case the power amplifier works in class AB because at quiescence the operating

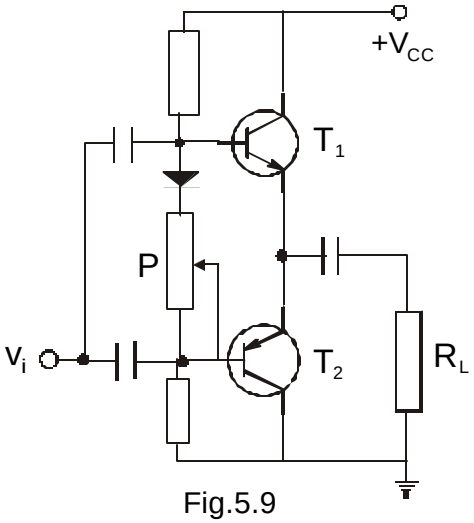


Fig.5.9

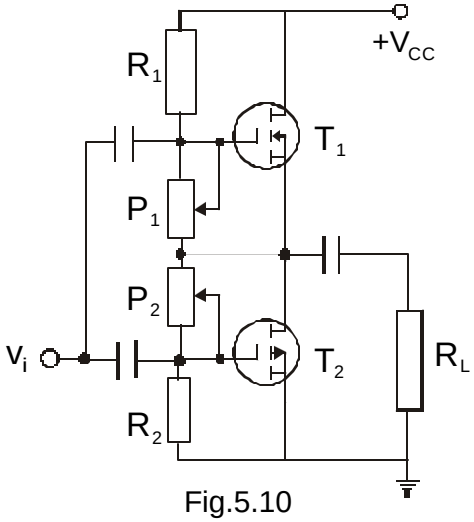


Fig.5.10

point Q is a little bit shifted from the cut off region into the active region. The typical circuit which avoids this distortion is shown in Fig.5.9. With the help of the potentiometer P it is possible to adjust the required bias voltage and consequently the quiescence current. This current is of the order of miliamperes, which is typically much smaller than the normal operating current, of the order of amperes, therefore it will not affect significantly the efficiency of the amplifier.

In practice, the current amplification factor of bipolar power transistors is much smaller than for small-signal transistor. Therefore, the bias circuit must supply relatively important currents to the transistor bases, and has to be implemented with relatively low values of the resistors,

increasing the DC power consumption at quiescence and decreasing the amplifier efficiency. One practical solution would be to use Darlington pairs instead of single power transistors. However, while this solves the biasing problem, it doesn't solve the problem of poorer linearity of bipolar transistors with respect to other transistor types. Today, most of the highly linear power amplifier are implemented using power MOSFETS, using a simplified circuit shown in figure 5.10.

**Class C power amplifier.**

Further increase of the power amplifiers efficiency can be obtained by biasing a transistor well below cutoff, so that it is active for only a small portion of the input signal (for instance for the "tip" of the sine shown in fig. 5.5). These amplifiers have limited use, since they introduce important distortions of the AC signal, and pairing (as for Class B amps) cannot solve this problem. However, in radio-frequency (RF) power circuits, most of the loads are resonant circuits, that are selective for a single frequency. If we restrict the operation of a Class C amplifier to a frequency that matches the one of the resonant load circuit, we can use it in this high efficency (up to 90%) regime without distorting the output signal.

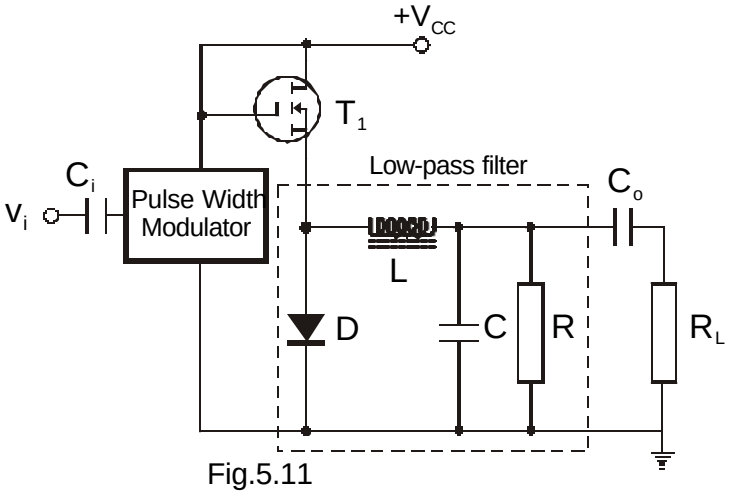


Fig.5.11

**Class D power amplifier.**

The ultimate solution for driving up the efficiency would be to operate the transistor at either full cutoff or full saturation. Ideally, if the cutoff current and the saturation voltage were both null, the efficiency would be 100%. In practice, efficiencies in the 90% 's are common. The transistor does not operate any more in a "linear" mode, but switches continuously between saturation and cutoff. In the past, the power transistors had poor switching characteristics, therefore class B push-pull amplifiers were the

most efficient circuits available. The advances in the semiconductor technology made possible fast-switching power transistors that can operate in the MHz range. However, in most applications a continuous range of output voltages is desired, not just “all-or-none”. This can be achieved in a Class D switching amplifier by modulating the duration of a high-frequency rectangular signal that drives the output transistors, then low-pass filtering (or averaging) the high frequency output component of the signal using an LC filter, as shown in fig. 5.11. The key to understanding the circuit operation is the Pulse Width Modulation (PWM) process, depicted in fig 5.12. When the input is positive, the PWM circuit modifies the duty cycle of the fixed frequency oscillator in such a way that its positive alternation is wider than the negative one, and, after low-pass filtering (averaging), the net output value will be positive. A similar encoding scheme is applied for the negative signal alternation. In summary, the signal encoding is changed from amplitude modulation to the duty cycle f the high-frequency signal, by PWM, amplified, then reconstructed back through low-pass filtering.

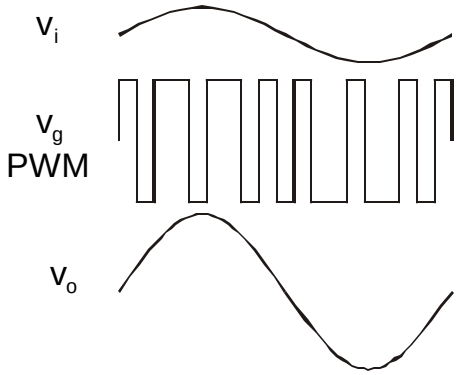


Fig.5.12

## Chapter 6. Negative feedback.

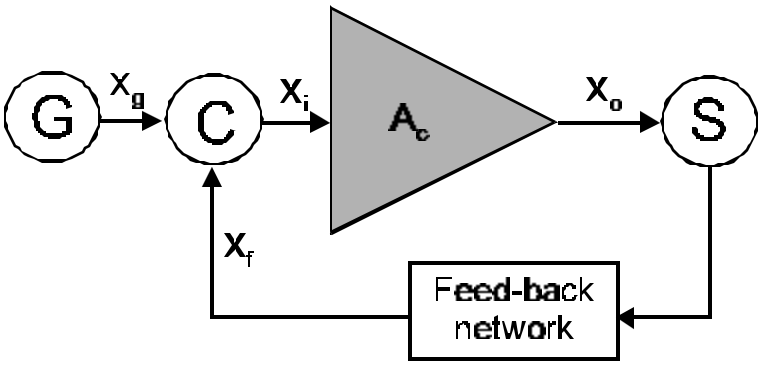



Fig.6.1

Most devices used in signal amplification generally do not have the desired gain, or there is a large dispersion in their parameters. In order to make an amplifier circuit more stable with the dispersion of individual components or enhance its parameters, a portion of the output signal is fed back to its input. The principle of feedback amplifiers is shown in Fig.6.1. It comprises five functionally distinct blocks. The letter **G** stands for the "Signal source"; the letter **C** stands for the "Comparing circuit"; while the letter **S** stands for the "Sampling circuit". The amplifier has the gain $\mathbf{A_o}$ defined by the classical relation:

$$A_o = \frac{x_o}{x_i} \qquad 6.1.$$

The feedback signal $\mathbf{x_f}$ is first sampled from the output signal $\mathbf{x_o}$, then it is fed back through a network that has the transfer function $\beta$:

$$x_f = \beta x_o \qquad 6.2.$$

The input signal results by substracting (or comparing) the source and feedback signals:

$$x_i = x_g - x_f \qquad 6.3.$$

By definition the gain of feed-back amplifier is given by:

$$A_f = \frac{x_o}{x_g} = \frac{x_o}{x_i + x_f} = \frac{A_o}{1 + \beta A_0} \qquad 6.4.$$

Formula 6.4. represents the general formula for the gain of feedback amplifiers. The signal **x** can be a current or a voltage signal.

Base on equation 6.4., we can define two kinds of feed-back:

**positive feed-back** when $\mathbf{A_f > A_o}$ or **negative feed-back** when $\mathbf{A_f < A_o}$

We will now focus our attention on the negative feedback. We can classify the feed-back amplifiers based on the sampling and comparing circuits. If the output signal is a current, the "sampling circuit" must be a **series circuit,** while if the output signal is a voltage signal the "sampling circuit" must be a **shunt circuit.** In the case of "comparing circuit" we have the reverse situation. If the input signal is a voltage signal the "comparing circuit" must be a **series circuit**, while if the input signal is a current, the "comparing circuit" must be a **shunt circuit.** These situations are shown in fig.2 for two ideal amplifiers. You can imagine the next two possible situations, in which:

a. the input signal is a voltage and the output signal is a current (the case of ideal trans-admittance amplifier $i_o = A_y v_i$) or

b. the input signal is a current and the output signal is a voltage (the case of ideal trans-resistance amplifier $v_o = A_z i_i$) .

Of course, in reality such ideal situations do not exist. The current amplifier has an input impedance different from zero and a finite output impedance, while the voltage amplifier has a finite input impedance and an output impedance greater than zero. Based on the feedback type, these impedances are modified, for the advantage of the feedback amplifiers.

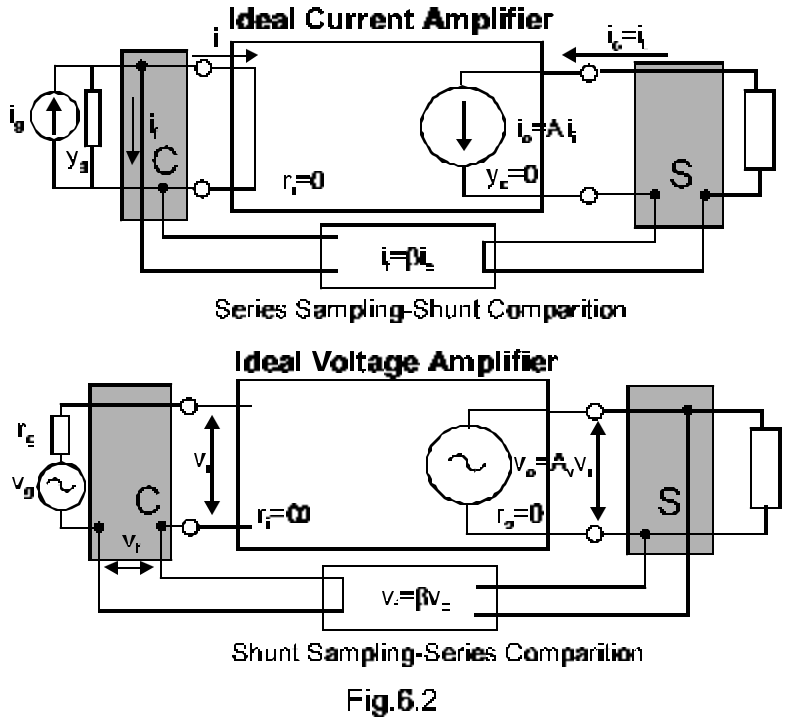



Fig.6.2

**Advantages of Negative Feedback Amplifiers.**

**1. Stability.**

When applying strong negative feedback we have $\beta A_o >> 1$. In this case we can neglect the term 1 to the denominator of equation 1.4., and thus the gain becomes:

$$A_f \approx \frac{1}{\beta} \qquad 6.5.$$

This formula proves that the negative feedback amplifiers have a constant amplification, independent of the active device (transistor, OPAMP, etc.).

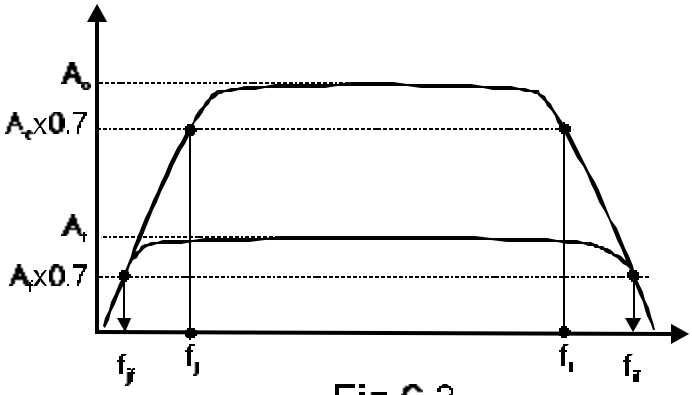



Fig.6.3

**2. Expansion of bandwidth.**

In Fig.6.3 is represented the amplification as a function of the frequency of input signal for an amplifier ($\mathbf{A_o}$) and for the same amplifier with a negative feedback ($\mathbf{A_f}$). As you can see the bandwidth is largest for the negative feeback amplifier. We will mathematically demonstrate in the next paragraph that for high cut-off frequency, this increase is:

$$f_{if} = f_i\left(1 + \beta A_o\right)$$

**3. Noise reduction.**

Every amplifier is characterised by a noise gain $\mathbf{A_N}$**.** This gain decreases by applying a negative feedback as in 6.4. formula.

$$A_{Nf} = \frac{A_N}{1+\beta A_N}$$

**4. Total Harmonic Distortion (THD) reduction.**

Every amplifier has a certain degree of non-linearity. Its output cannot be written any more as a simple linear function of the input signal $x_0=A_0x_i$ , but it can be written as a power series. In terms of the power spectrum, this is reflected in the generation of harmonics when the input signal is a pure sine signal having a fixed frequency. The cumulated amplitude of these harmonics expressed as a percentage of the amplitude of the fundamental frequency is called total harmonic distortion (THD). In most cases, the second harmonic is largest then the sum of all others, and therefore finding its relative amplitude $\mathbf{D_2}$ provides a fair estimation of the overall THD rate. In transistor circuits the distortions arise form the fact that the signal cannot be made arbitrarily small to make it verify the linearity hypothesis, therefore for real signals the hybrid parameters cannot be considered constant any more. By applying a negative feed-back, the second harmonic generation is reduced following the same equation 6.4. :

$$D_{2f} = \frac{D_2}{1+\beta D_2}$$

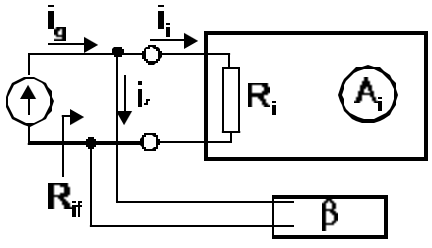



Fig.6.4

**5. Input and output impedance changing.**

Fig.6.4 represents a typical shunt (parallel) comparing circuit for a real current amplifier. By definition, the input impedance is $R_i = \frac{v_i}{i_i}$ and the current gain is $A_i = \frac{i_o}{i_i}$.

The input impedance, in the case of negative feed-back becomes :

$R_{if} = \frac{v_i}{i_g} = \frac{v_i}{i_i + \beta i_o} = \frac{R_i}{1 + \beta A_i}$, then $\mathbf{R_{if} < R_i}$ in this case.

Fig.6.5 represents a typical series comparing circuit for a real voltage amplifier. By definition the input impedance is $R_i = \frac{v_i}{i_i}$ and the voltage gain is $A_v = \frac{v_o}{v_i}$.

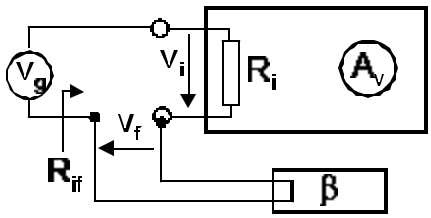


Fig.6.5

The new input impedance, for negative feed-back amplifier becomes :

$R_{if} = \frac{v_g}{i_i} = \frac{v_i + \beta v_o}{i_i} = R_i(1 + \beta A_v)$, then $\mathbf{R_{if} > R_i}$ in this case. You can try to demonstrate what happens with output impedance (see fig.6.2 and take into account that you have a real amplifier).

**Positive Feed-back.**

As we have seen, the gain of amplifier with feed-back, is given by the equation 6.4 :

$$A_f = \frac{A_o}{1 + \beta A_0}$$

From this formula one can see that the gain goes towards infinity (and so does the amplitude of the output signal) if the denominator vanishes to zero. In this situation, even in the absence of an input signal, any noise or fluctuation at the input leads in the generation of a large output signal, which is further fed back to the input, being perpetually regenerated. The device is called "oscillator" or signal source generator. The condition:

$$\beta A_o = -1 \qquad 6.6.$$

is named "Barkhausen's criterion". One has to note that in general both the transfer function $\beta$ and the open-loop amplification factor $A_0$ are complex (in order to accurately describe both amplitude and phase properties), therefore the Barkhausen criterion can be written as two different criterions for the real and imaginary part.

**The LC oscillator.**

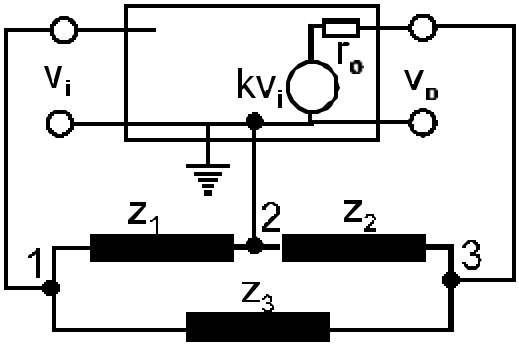



Fig.6.6

The voltage gain of the voltage amplifier depicted in Fig.6.6, having the load impedance $\mathbf{z_L}$ is :

$$A_v = \frac{v_o}{v_i} = \frac{z_L i_L}{v_i} = \frac{k z_L}{z_L + r_o} \qquad 6.7.$$

where $\mathbf{z_L}$ is :

$$z_L = \frac{z_2 (z_1 + z_3)}{z_1 + z_2 + z_3} \qquad 6.8.$$

The feedback signal is the input signal:

$$v_i = v_f = \beta v_o$$

From this relation results the feedback factor:

$$\beta = \frac{z_1}{z_1 + z_3} \qquad 6.9.$$

Now, in relation 6.6. we can substitute $A_o$ as given by equation 6.7. and $\beta$ as given by equation 6.6. Under these conditions, the Barkhausen criterion becomes:

$$-1 = \frac{k z_1 z_2}{r_o (z_1 + z_2 + z_3) + z_2 (z_1 + z_3)} \qquad 6.10.$$

Now, if all $z_i$ impedances are reactances ($z_i = jX_i$), the formula 6.10 can be a real number only if

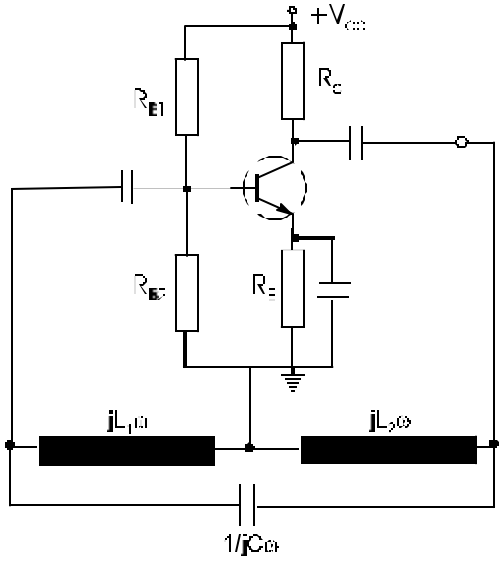
$$X_1 + X_2 + X_3 = 0 \qquad 6.11.$$

The relation 6.11 is the "oscillator condition" for LC oscillators. From equations 6.10 and 6.11, we can obtain the second condition for LC oscillators:

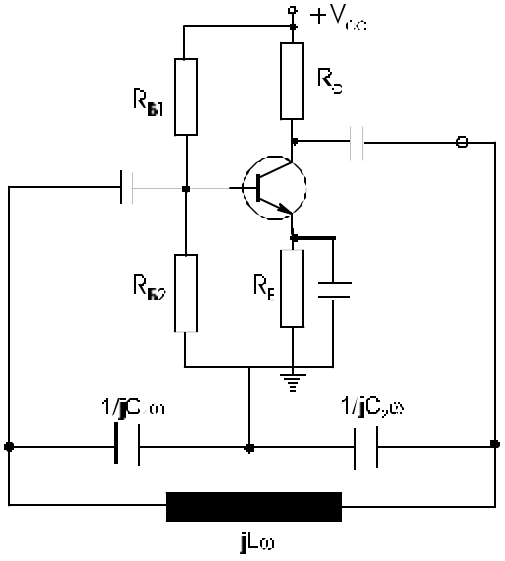
$$k = \frac{X_2}{X_1} \qquad 6.12.$$

From equation 6.11 results that if $X_1$ and $X_2$ are inductances, $X_3$ must be a capacitance. This is the so called **Hartley** oscillator. Conversely, if $X_1$ and $X_2$ are capacitances, $X_3$ must be an inductance. This is the so called **Colpitts** oscillator. In next figures are presented the two main types of LC oscillators. The operating frequencies are deduced from equation 1.6. Then, for the Hartley oscillator we have:

$$jL_1\omega + jL_2\omega - j\frac{1}{C\omega} = 0 \Rightarrow \omega^2 = \frac{1}{(L_1 + L_2)C} \qquad 6.13.$$




**Hartley oscillator**




**Collpitz oscillator**

and for Colpitts oscillator we have:

$$jL\omega - j\frac{1}{C_1\omega} - j\frac{1}{C_2\omega} = 0 \Rightarrow \omega^2 = \frac{1}{\left(\frac{C_1C_2}{C_1 + C_2}\right)L} \qquad 6.14.$$

The usual range of operating frequencies for the Hartley oscillator is 100kHz – 10MHz, and for Colpitts oscillator is 1MHz – 100MHz. For lower frequencies, RC oscillators are used, while for higher frequencies oscillators with Leher line (distributed LC constants oscillators) or wave guide oscillators are used. For very high frequencies, special electronic devices such as tunnelling diodes, Impatt diodes, clystrons and magnetrons (up to 1Ghz) are used.

**RC oscillators.**

The positive feedback in these oscillators is performed by the "Wien network", as in fig.6.7. ,

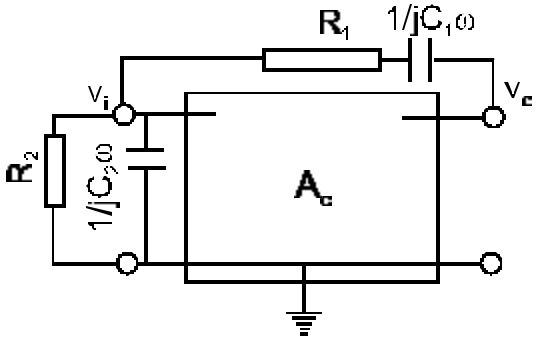


Fig.6.7

where:

$$z_1 = R_1 + \frac{1}{jC_1\omega} \text{ and } \frac{1}{z_2} = \frac{1}{R_2} + jC_2\omega$$

$$v_i = \frac{v_o}{z_1 + z_2} z_2 = \beta v_o \Rightarrow \beta = \frac{z_2}{z_1 + z_2}$$

Then:

$$\beta = \frac{1}{1 + \frac{R_1}{R_2} + \frac{C_2}{C_1} + j\left(R_1 C_2 \omega - \frac{1}{C_1 R_2 \omega}\right)} \qquad 6.15.$$

Now, if we want to have the Barkhausen criterion satisfied, the feedback factor must be a real number, therefore the parenthesis in the denominator of 6.15 formula must be null. Based on that condition, we can calculate the frequency of the RC oscillator as:

$$\omega^2 = \frac{1}{R_1 R_2 C_1 C_2} \qquad 6.16.$$

If we have $R_1$=$R_2$ and $C_1$=$C_2$ , the gain of amplifier must be $A_o$=3, according to equation 6.15.

# Chapter 7. The Bipolar Transistor Behaviour at High Frequency Signals.

In the case of high frequency signal, applied to the input of a bipolar transistor, the output signal will depend on the junctions' capacitance, that cannot be neglected as we did for the low frequency signal analysis. Then for high frequencies signals the hybrid model of the transistor must be replaced by other model that take into account the junctions' capacitance.

**The "Π" Hybrid Model. (Giacoletto Model)**

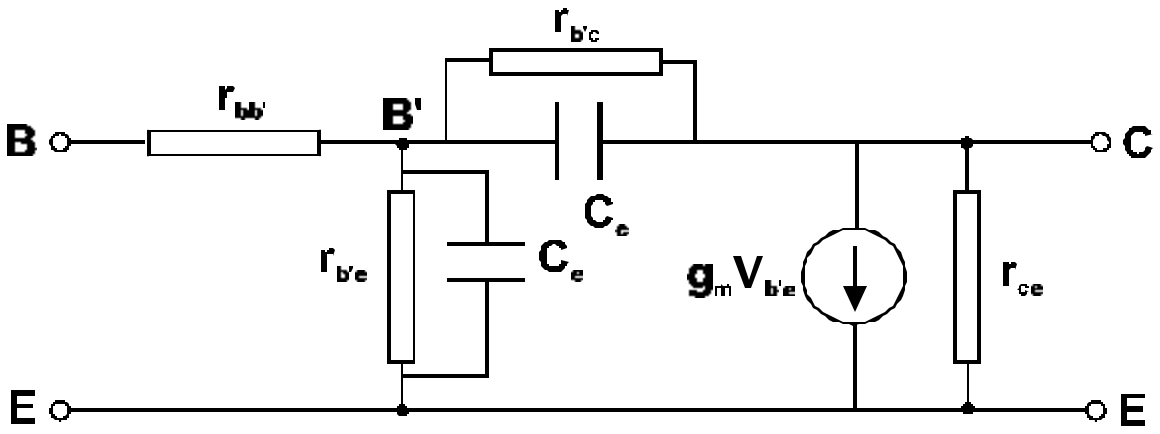



Fig.7.1

In Fig.7.1 is shown the model proposed by Giacoletto for the bipolar transistor operating with high frequency input signals. The point **B'** represents the so called **"virtual base"** which represents a point inside the semiconductor which forms the base of transistor. Between this point (**B'**) and the pinch of the base (**B** point) exists the so called **"distributed resistance"** of the transistor base ($r_{bb'}$). This "distributed resistance" takes into account the real resistance of the thin longitudinal section of the semiconductor which forms the base of transistor. In this case, the input signal applied to the base **B** is not the real input signal, due to the voltage drop that occurs on resistor $r_{bb'}$ . The real input signal will be the voltage signal which goes up the virtual base **B'**, therefore the forward transfer factor $g_m$ (the transfer-

conductance) is multiplied with the voltage between virtual base **B'** and the emitter of transistor **E**, aiming at providing a model for the **"constant current generator"** of the collector ($g_m V_{b'e}$).

The input impedance of the transistor is modelled by resistor $r_{b'e}$ , while the **diffusion capacitance** of the emitter junction is modelled by capacitor $C_e$ connected in parallel with the $r_{b'e}$ resistor.

The "Early effect", or base width modulation effect, is modelled by the resistor $r_{b'c}$ , while the **barrier capacitance** of the collector is modelled by the $C_c$ capacitor connected in parallel with the resistor $r_{b'c}$ . Finally, the output impedance is modelled by the $r_{ce}$ resistor.

The common values for all these model components are:

**$r_{bb'}$ = 100 ohms ; $r_{b'e}$ = 1 Kohms ; $r_{b'c}$ = 4 Mohms ; $r_{ce}$ = 80 Kohms ;**

**$C_e$ = 100 pF ; $C_c$ = 5 pF ; $g_m$ = 50 mA/V**

**The main advantage of this high-frequency model is that it allows to compute all these parameters using the well-known hybrid parameters of the transistor ($h_{ie}$ , $h_{re}$ , $h_{fe}$ , $h_{oe}$) .**

**The transfer - conductance $g_m$ .**

The most important factor of the "Π" Hybrid Model is the transfer - conductance $g_m$ . The definition relation of $g_m$ is as follows:

$g_m = \frac{\partial I_C}{\partial V_{b'e}} = -\alpha \frac{\partial I_E}{\partial V_{b'e}} = \alpha \frac{\partial I_E}{\partial V_E}$ , where we took into account that $I_C \approx -\alpha I_E$ .

But , from the P-N junction we know that $\frac{\partial V_E}{\partial I_E} = r_e'$ , then $g_m = \alpha \frac{\partial I_E}{\partial V_E} = \frac{\alpha}{r_e'}$ . Because the emitter junction behaves like a forward biased P-N junction, then we can use the relation for

the ideal diode $I_E \approx const. \times e^{\frac{V_E}{V_T}}$. Therefore the derivative of $I_E$ versus $V_E$ is going to be:

$\frac{\partial I_E}{\partial V_E} \approx \frac{I_E}{V_T}$, and the expression of $\mathbf{g_m}$ becomes :

$$g_m = \alpha \frac{\partial I_E}{\partial V_E} \approx \alpha \frac{I_E}{V_T} = \frac{|I_C|}{V_T} \qquad 7.1.$$

As you can see, the equation 7.1. is the first one connecting the main transfer factor in the transistor model to temperature ($V_T = kT/e$).

**The connection between the Giacoletto Model and the Hybrid Model.**

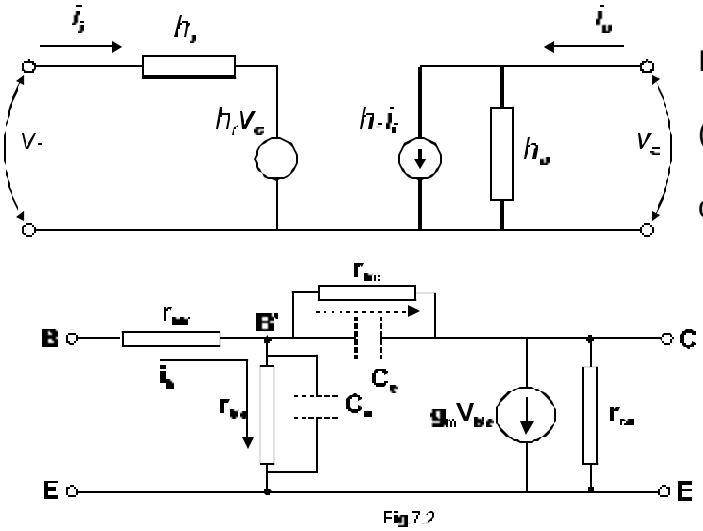


Fig 7.2

In fig.7.2 we present both models (hybrid and $\Pi$ hybrid) of transistor. In order to draw a correspondence between them, we must consider only the case of low frequency input signals, because the hybrid model is valid only for low frequencies.

In this case, we can neglect all capacitors in $\Pi$ hybrid model and the current that flows through resistor $r_{b'c}$ .

Then, according to hybrid parameters definition, we can write following relations for the input circuits of these models:

$i_c\big|_{R_L=0} = g_m v_{b'e} = g_m i_b r_{b'e}$ , but $\left.\frac{i_c}{i_b}\right|_{R_L=0} = h_{fe}$, then we have following relation

between trans-conductance of $\Pi$ hybrid model and forward transfer factor ($h_{fe}$) of the hybrid model:

$$h_{fe} = g_m r_{b'e} \qquad 7.2.$$

**The Gain Current at High Frequency.**

In the case of high frequency input signal and low load resistor ($R_L < r_{ce}$), we can neglect a number of elements of the Giacoletto model, as you can see from the figure 7.3.

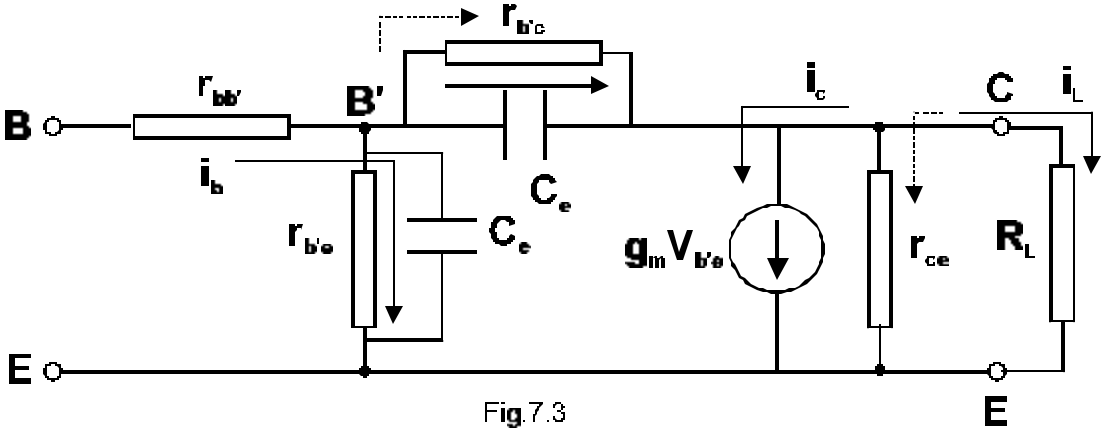



Fig.7.3

The currents that can be neglected are represented by doted lines. Also, the devices

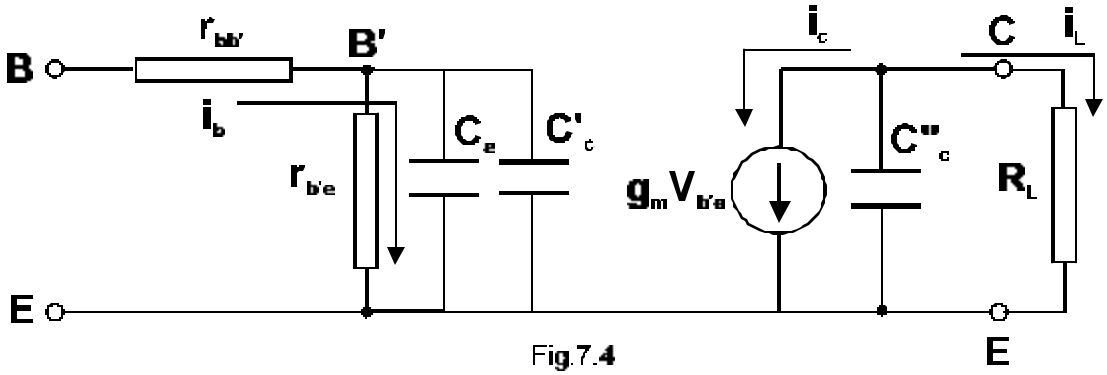



Fig.7.4

through which such currents flow can be neglected, too. In this situation we have following elements:

$$r_{b'c} >> \frac{1}{\omega C_c} \qquad ; \qquad r_{ce} >> R_L$$

Using Miller's theorem, we can replace the $C_c$ capacitor as you can see in figure 7.4.

were $\frac{1}{\omega C'_c} = \frac{\frac{1}{\omega C_c}}{1 - A_v} = \frac{1}{\omega C_c (1 + g_m R_L)}$, because

$$A_v = \frac{i_L R_L}{v_{b'e}} = \frac{-g_m v_{b'e} R_L}{v_{b'e}} = -g_m R_L$$

and C"$_c$=C$_c$ , because $z'' = \frac{zA_v}{A_v - 1} \approx z\big|_{A_v >> 1}$

Then the current gain is:

$$A_i = \frac{i_L}{i_i} = -\frac{g_m v_{b'e}}{\frac{v_{b'e}}{z_i}} = -\frac{g_m r_{b'e}}{1 + j\omega C_i r_{b'e}} \qquad 7.3.$$

where $C_i = C_e + C'_c = C_e + C_c (1 + g_m R_L)$

but, $A_i = -\frac{h_{fe}}{1 + j\frac{f}{f_{\sqrt{2}}}}$ , the we can define

$$f_{\sqrt{2}} = \frac{1}{2\pi r_{b'e} \left[ C_e + C_c (1 + g_m R_L) \right]} \qquad 7.4.$$

as the high frequency for which the real current gain decreases to $\left| A_{\sqrt{2}} \right| = \frac{|A_i|}{\sqrt{2}}$.

We can also define the cut off frequency ($f_{off}$), as the frequency for which the real gain becomes equal to one,

$$|A_i|_{f = f_{off}} = 1 = \frac{h_{fe}}{\sqrt{1 + \frac{f_{off}^2}{f_{\sqrt{2}}^2}}} = \frac{h_{fe} f_{\sqrt{2}}}{\sqrt{f_{\sqrt{2}}^2 + f_{off}^2}} \approx \frac{h_{fe} f_{\sqrt{2}}}{f_{off}} \qquad 7.5.$$

from 7.4. and 7.5. relations we can obtain

$$f_{off} = h_{fe} f_{\sqrt{2}} = \frac{h_{fe}}{2\pi r_{b'e} (C_e + C'_c)} = \frac{h_{fe} g_{b'e}}{2\pi (C_e + C'_c)} \qquad 7.6.$$

Relation 7.6. connect the cut off frequency to main parameters that define the Giacoletto model ($g_m$, $C_e$, $C_c$).

If we take the load resistor as being null, $R_L$=0, the capacitance $C'_c$ becomes $C_c$ , then the relation 7.6 becomes, if we neglect $C_c$ versus $C_e$ ($C_c << C_e$):

$$f_{off} = \frac{h_{fe} g_{b'e}}{2\pi C_e} \qquad 7.7.$$

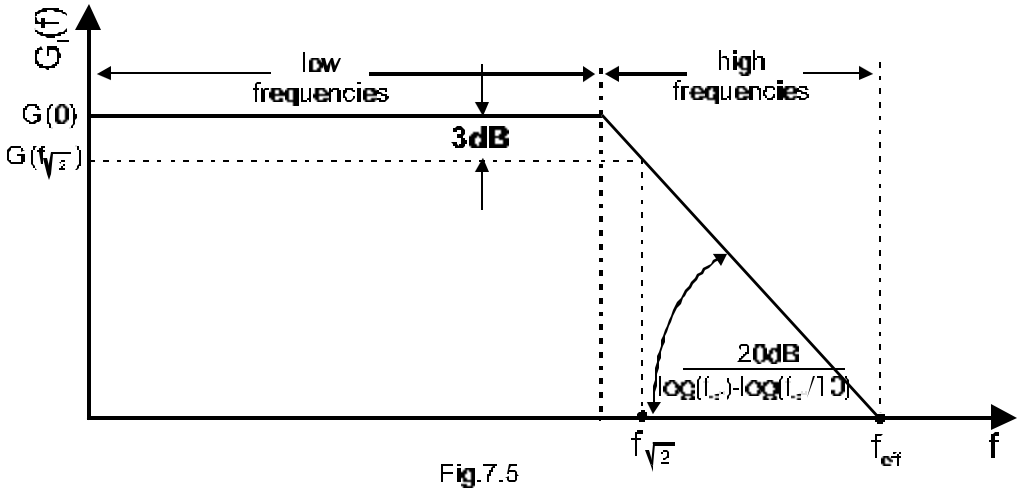



Fig.7.5

Equation 7.6. can be used in order to determine experimentally the value of $C_e$ , the most important capacitance of $\Pi$ hybrid model.

Now, if we define the current gain using the attenuation of the signal, that means to use the decibels way of expressing it:

$$G_i = 20\log|A_i| = 20\log\frac{h_{fe}}{\sqrt{1+\frac{h_{fe}^2 f^2}{f_{off}^2}}} \approx 20\log\frac{h_{fe} f_{off}}{h_{fe} f} = 20\log\frac{h_{fe} f_{\sqrt{2}}}{f} = 20\log\left(h_{fe} f_{\sqrt{2}}\right) - 20\log f$$

you can see that **$G_i$(f)** is a linear function of the frequency **f.** Therefore this function can be represented as in figure 7.5.

# Chapter 8. Multivibrators

## 8.1. The Astable (free-running) Multivibrator.

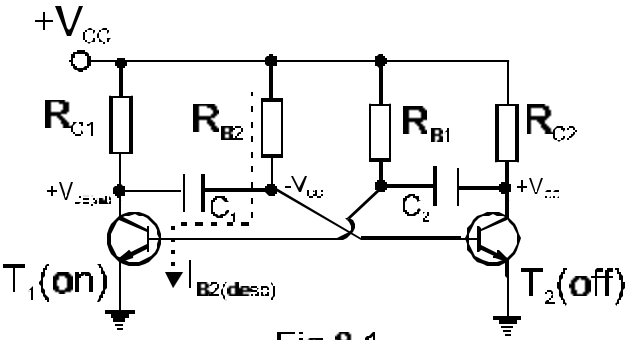


Fig.8.1

The astable or free-running multivibrator is presented in Fig.8.1. In fact this circuit is realised by two simply common emitter amplifiers (both with fixed biasing circuits) coupled between them by two capacitors $C_1$ and $C_2$. The biasing circuit ($R_{B1}$ and $R_{B2}$ resistors) has such a value that the transistor works in saturation region of output current (see fig.8.2 where are represented the possible working points).

The signals generated by such circuit are shown in Fig.8.3. As you can see, if one transistor is on, the second transistor is off. The transistor is switched on when his base voltage becomes a little bit positive (npn transistor). That is possible because the capacitor connected in his base will discharge via

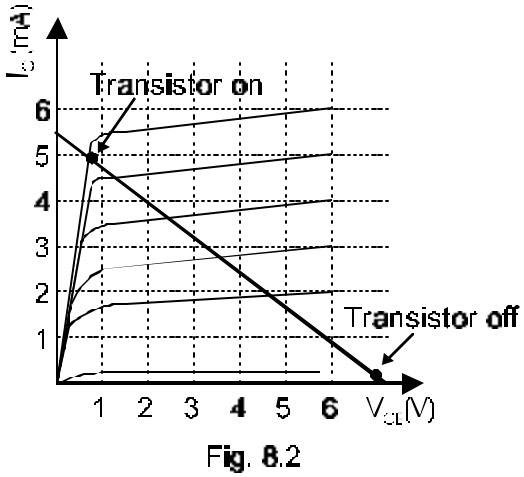


Fig. 8.2

the transistor which is in conduction (on).

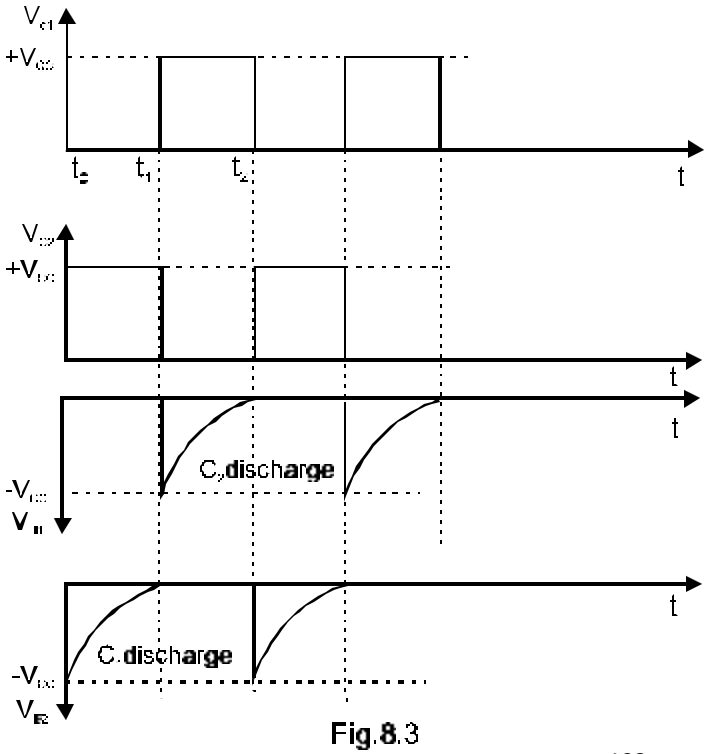


Fig.8.3

If we start from the initial conditions as in Fig.8.1 ($T_1$ on and $T_2$ off) at time $t_0$ , the collector voltage of $T_1$ is very close to zero ($V_{CEsat}$), whereas the collector voltage of $T_2$ is $V_{cc}$ , this transistor being turned off. The base voltage of $T_1$ is a little positive, this transistor being turned on, and the base voltage of $T_2$ being $-V_{CC}$ (transistor turned off) as in fig.8.3. Because the transistor $T_1$ is on, the capacitor $C_1$ can not remain charged, then it will discharge through $R_{B2}$ resistor and collector-emitter circuit of $T_1$. At the time $t_1$ the base voltage of $T_2$ becomes a little bit positive (the $C_1$ capacitor will try to recharge to maximum potential which exist in system) and in this moment $T_2$ is switched on. Following that, the polarity of $C_2$ capacitor plates will change because the drop voltage can not change instantly on a capacitor[3], then the base voltage of $T_1$ transistor become $-V_{CC}$ and this transistor will be switched off. From this moment ($t_1$) becomes the discharge of $C_2$ capacitor trough $R_{B1}$ resistor and the collector-emitter circuit of $T_2$ , as in the first case, till the moment $t_2$, when the system come back to the state of $t_0$ moment.

**8.2. The Monostable Multivibrator.**

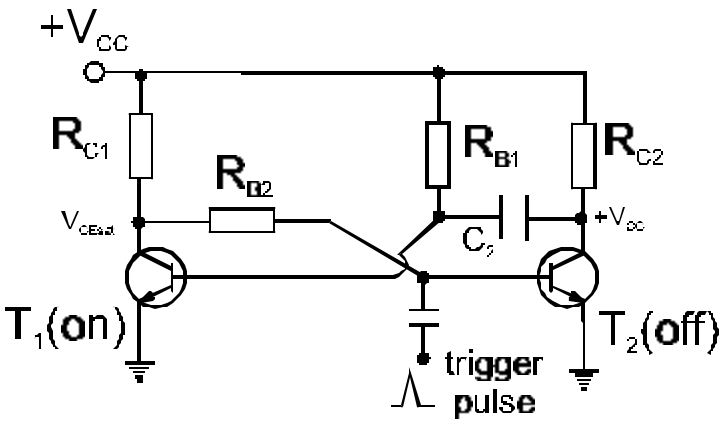



Fig.8.4

As you can see in fig.8.4, the biasing circuit for **$T_1$** is a typical circuit via **$R_{B1}$** resistor. The biasing circuit for **$T_2$** transistor is connected in the collector of **$T_1$** (the $R_{B2}$ resistor). Then, if transistor **$T_1$** is "on", that means in conduction, the biasing voltage in the base of **$T_2$** is too low and this transistor is "off", that means no conduction for this one and the voltage on his collector is just $+V_{CC}$ . This state

[3] $V_{C2} = +0 - (+V_{CC}) = -V_{CC} - (+V_{CE2sat}) \approx -V_{CC}$

represents the stable state of this circuit and for that this circuit is called “monostable circuit”. Now, if we apply a positive pulse (trigger pulse) on his base we will force this transistor to enter in conduction regime, that means a change in the polarity of the plates of **$C_2$** capacitor, because the voltage across a capacitor can not change suddenly ($U_{c_2} = V_A - V_B = +0 - V_{CC} = -V_{CC} - 0 = -V_{CC}$). Then, after applying the trigger pulse, transistor **$T_1$** will be turned off, because the A plate of **$C_2$** capacitor will change his potential from +0 Volts to $-V_{CC}$ Volts. Because under these circumstances **$T_2$** remain in conduction state, the capacitor $C_2$ will discharge via the emitter of **$T_2$** to the ground. Then **$T_1$** will remain in cut off state untill the $C_2$ capacitor discharging is complete.

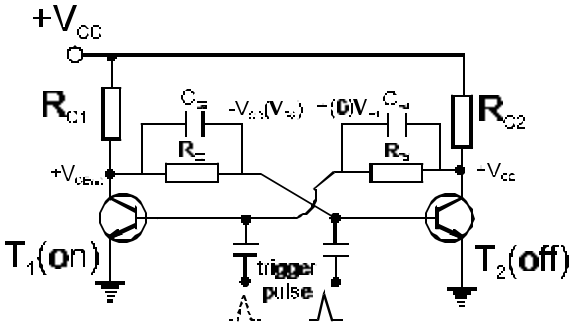


Fig.8.5

### 8.3. The Bistable Multivibrator.

In Figure 8.5 is shown the Bistable Multivibrator, or the Flip-Flop. The name of this circuit means that both on/off states of the transistor are stable states. Then, in order to bring transistors to their complementary states, we must trigger the transition of the circuit from an external signal. If, such as in Fig.8.5, the $T_1$ transistor is in the “on” state, then the $T_2$ transistor is in “off” state, the switching to the complementary state ($T_1$ off and $T_2$ on), can be triggered by a positive pulse applied to the base of $T_2$. In this

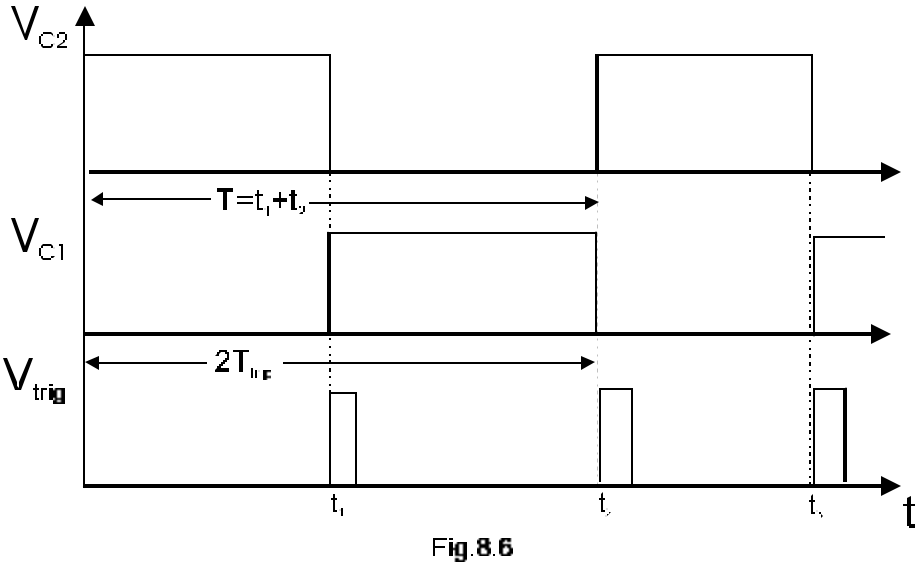


Fig.8.6

case, the polarity of both capacitor plates connected to the transistor bases, will be reversed, because the drop voltage on a capacitor can not change instantly. The new state will be $T_1$ off and $T_2$ on, which is a stable state too. In the aim to come back to the initial state, $T_1$ on and $T_2$ off, we must applied a second trigger pulse, now in the base of $T_1$ transistor. The shape of collector signals looks like in Fig.8.6. However, there's no need to apply two different pulses to trigger the switching. Indeed, if we connect the two trigger inputs together, the positive pulse applied to the base of the transistor which is already in the "on" state will have no effect, whereas the pulse delivered to the base of the transistor that's in the "off" state will actually trigger state switching.

As you can see in Fig.8.6, we need two trigger pulses to come back to the initial state. That means that the bistable circuit is a divider with 2 of the original trigger frequency.

# Chapter 9. The Operational Amplifier.

Historically speaking, the name Operational Amplifier comes from their use in early analogic computers, for performing different mathematical operations. In general an Operational Amplifier (OPAMP) is a high gain ($A_V \sim 10^5$) differential voltage amplifier, with relatively high input impedance ($Z_i \sim 10^6$). Following these characteristics, one can quickly extract the following rules for the normal operation of the OPAMP:

A. Due to the high input impedance, the current through the inputs in negligible and can be considered 0 in most applications.
B. The voltage between the differential inputs is very close to zero. This is a result of the fact that if the output is not at saturation, its normal voltage range is of the order of volts or tens of volts. Since $A_V \sim 10^5$, the input differential voltage is $v_i = v_o/A_V \sim 10^{-4}$ V, in the sub-mV range.

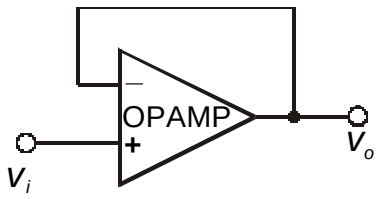


Fig. 9.1

The OPAMPS are never operated in open loop configuration, and in order to stabilise their behaviour, a negative feed-back is applied in various configurations, the simplest being shown in fig. 9.1, in which there's a "total feedback" applied from the output to the inverting input. Following rule B, the voltage between the inputs (of which one is now shunted to the output) is null, therefore: $v_o = v_i$. Since the voltage output closely follows the input voltage, this configuration is called "voltage follower". However, this stage differs from a simple shunt, which provides the same unitary voltage gain, in that that the input impedance is very high.

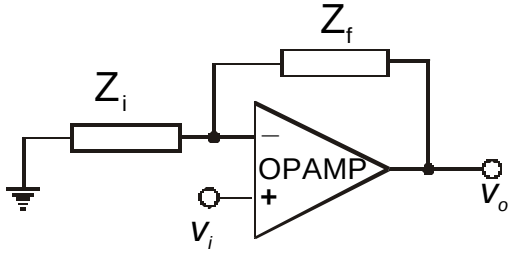


Fig. 9.2

If one desires a voltage gain greater than one, it has to apply a less stronger negative feedback. This can be achieved by dividing the output voltage using a resistive divider before applying the negative feedback, as shown in fig. 9.2. The feedback

voltage at the output of the divider is $V_o \cdot \frac{Z_i}{Z_f + Z_i}$, which, following rule B is practically the input voltage. From here, the voltage gain can be calculated as:

$$A_V = 1 + \frac{Z_f}{Z_i}$$

Since this configuration preserves the polarity of the DC input signals or the phase of the AC input signals, it is called “noninverting amplifier”.

In some applications, one may want not only to amplify the signal, but also to change its polarity or phase. This is achieved by the configuration shown in fig. 9.3, named “inverting amplifier”. The negative feedback propagates through the $Z_f$ impedance from output to input. It results in an equivalent input impedance at the inverting input which will be shown to vanish to zero. By means of the Miller theorem, the feedback impedance impedance is equivalent to two impedances connected in parallel with the input circuit, respectively with the output circuit, as shown in fig. 9.4. The values of these impedances as given by Miller’s theorem are:

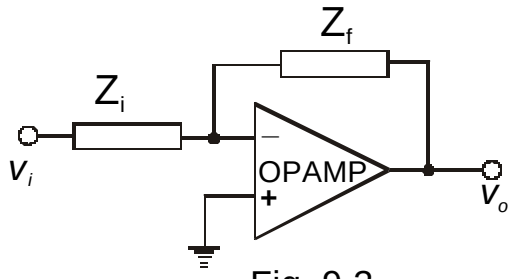


Fig. 9.3

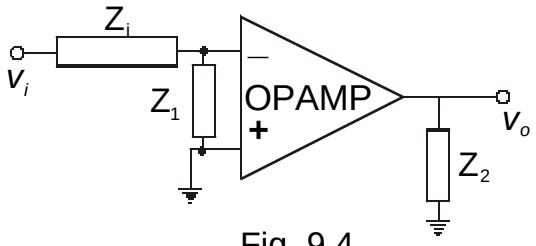


Fig. 9.4

$$Z_1 = \frac{Z_f}{A_v - 1} \;\; ; \;\; Z_2 = \frac{Z_f}{1 - \frac{1}{A_v}}$$

Since $A_v$ is very high, the input impedance $Z_1$ vanishes to zero and acts like a “virtual short-circuit” or shunt to the ground. This could have also been quickly inferred using rule B. However, this virtual shunt differs from a real one in that that the current through it is null, given

the rule A, therefore all the current flow is directed through the feedback impedance $Z_f$. The equivalent circuit of Fig. 9.3 results in being as simple as shown in Fig. 9.5

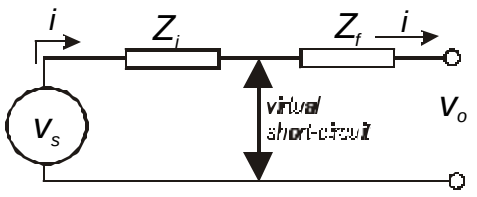


Fig. 9.5

Based on this transformation, the voltage gain of the inverting amplifier can be calculated as :

$$A_f = \frac{v_o}{v_s} = \frac{-iZ_f}{iZ_i} = -\frac{Z_f}{Z_i}$$

Accordingly, the input impedance is simply going to be the impedance inserted in series with the input. This is going to be much less than the input impedance of the OPAMP itself, or the input impedance of the voltage follower or the inverting amplifier. Moreover, if the signal source does not have a negligible internal (or output) impedance, its impedance will add to the input impedance of the inverting amplifier. This will result in a change in the voltage gain, which for the other amplifier configurations is fairly constant with respect to the signal source impedance.

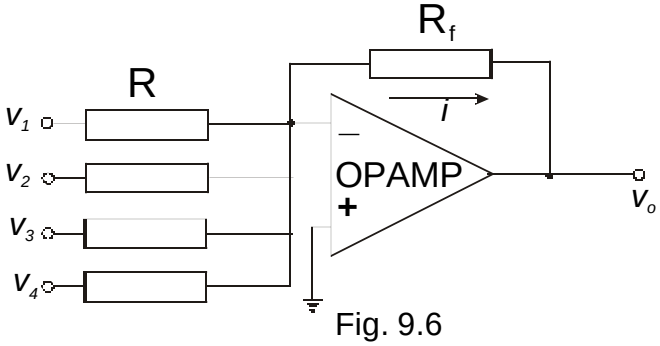


Fig. 9.6

## Basic Circuits using OPAMPS.

### 9.1 . Summation Circuit.

As mentioned earlier, one of the first uses of OPAMPS was in performing mathematical operations. Summation is the most basic one, and the circuit that performs this operation is based on the inverting amplifier configuration, as shown in fig. 6.

The input currents sum at the virtual ground point following one of the Kirchhoff's current laws.

$$i = \frac{v_1}{R} + \frac{v_2}{R} + \frac{v_3}{R} + \frac{v_4}{R} + \ldots.$$

but since this current flows entirely through the feedback resistor $R_f$ (following rule A), we have:

$$v_0 = -iR_f$$

then

$$v_o = -\frac{R_f}{R}(v_1 + v_2 + v_3 + v_4 + ....)$$

equation that shows that the output voltage is proportional with the summation of input voltages.

A substracting circuit can be built in two stages, first by inverting the voltage to be substracted using an inverting amplifier having unity gain ($R_f$=$R_i$), then adding this "negated" voltage to the second one by using the summation circuit discussed above.

**9.2. Integration Circuit.**

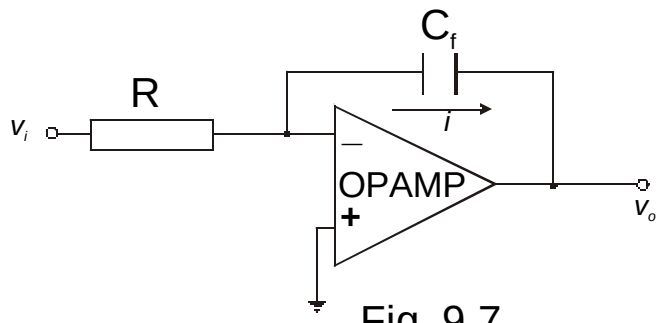


Fig. 9.7

The impedances used in the basic inverting and noninverting applications can be purely capacitive or inductive, not only resistive. In this case the OPAMP can perform complex mathematical functions such as integration and differentiation. If in the inverting configuration we use as feedback impedance a capacitor, the voltage across it will be the integral of the current that charges it:

$$v_o = -\frac{1}{C_f}\int i(t)dt$$

But the charging current is actually the input current, $i(t) = \frac{v_i(t)}{R}$,

therefore

$$v_o = -\frac{1}{C_f R}\int v_i(t)dt\ ,$$

equation that shows that the output voltage is proportional to the integral of input voltage.

**9.3. Differentiation Circuit**

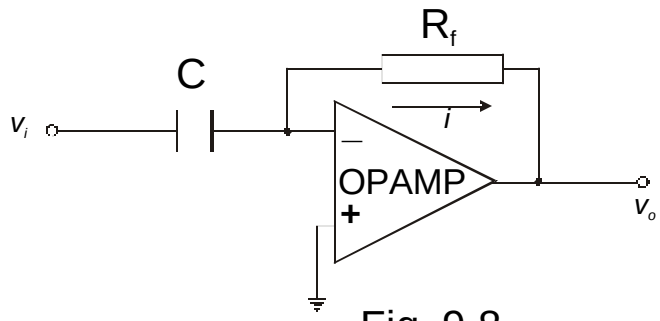



Fig. 9.8

Conversely, if the input impedance is a capacitor, the current that passes through it is going to be:

$$i = C\frac{dv_i}{dt}$$

but $v_o = -R_f i$ and therefore

$$v_o = -R_f C\ \frac{dv_i}{dt},$$

equation that shows that the output voltage is proportional to the derivative of the input voltage.

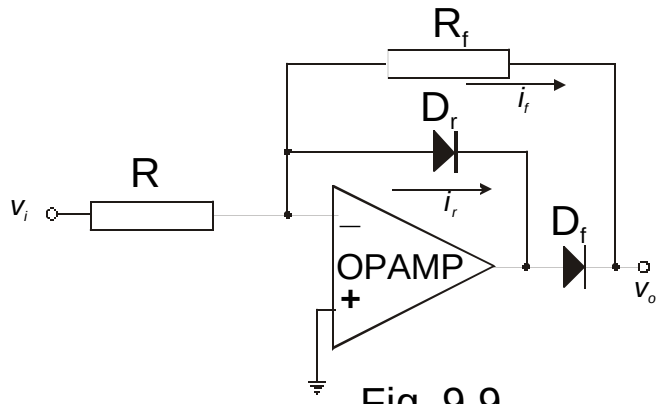



Fig. 9.9

### 9.4. Ideal Rectifier

Voltage rectification is primarily performed using rectifier diodes. However, the diodes present the flaw of a nonzero rectifying voltage threshold (or forward voltage). Signals smaller than the forward voltage simply cannot be rectified. Since we have seen that the OPAMP in appropriate configurations can do a number of magic things such as creating virtual shunts with no current through them or changing circuit impedances, it's worth checking if it can do something about creating a "virtual diode" with a null threshold voltage (ideal rectifier). Let's take a look at the schematics outlined in figure 9.9.

The situations in which the input voltage is positive or negative will have to be dealt with separately, since the conduction states of the of the diodes $D_r$ and $D_f$ will be different, and therefore the input current will branch in a different way. In this inverting configuration, for an arbitrarily small negative input, the very high voltage gain will drive the output of the OPAMP positive enough to bring the diode $D_f$ into conduction, above its threshold. Conversely, the diode Dr will be reverse biased, and therefore there's going to be no current flow through it.

All the current $I_f$ is going to flow from the output of the OPAMP through the diode Df and the feedback resistor Rf, with a direction opposite to the conventional one depicted in fig. 9.9. Since the inverting input is a virtual ground point, the voltage at the output is going to be:

$v_0 = -i_f R_f$ , or $v_0 = -v_i \frac{R_f}{R}$, that is inversely proportional to the negative input voltage.

For a positive input voltage, the OPAMP output is going to be driven negative, in which case the diode $D_f$ is going to be reverse polarized, acting as a circuit breaker, letting the output voltage $v_0$ being driven to the ground through the load resistor or through the feedback resistor. All the input current will flow through the diode $D_r$, which is forward polarized to the output of the OPAMP. The output voltage of the rectifier stage will be null. This dual behavior is approaching the ideal rectifier characteristics.

*Finish: Congratulations for those who got to finish this course!*

*The authors.*