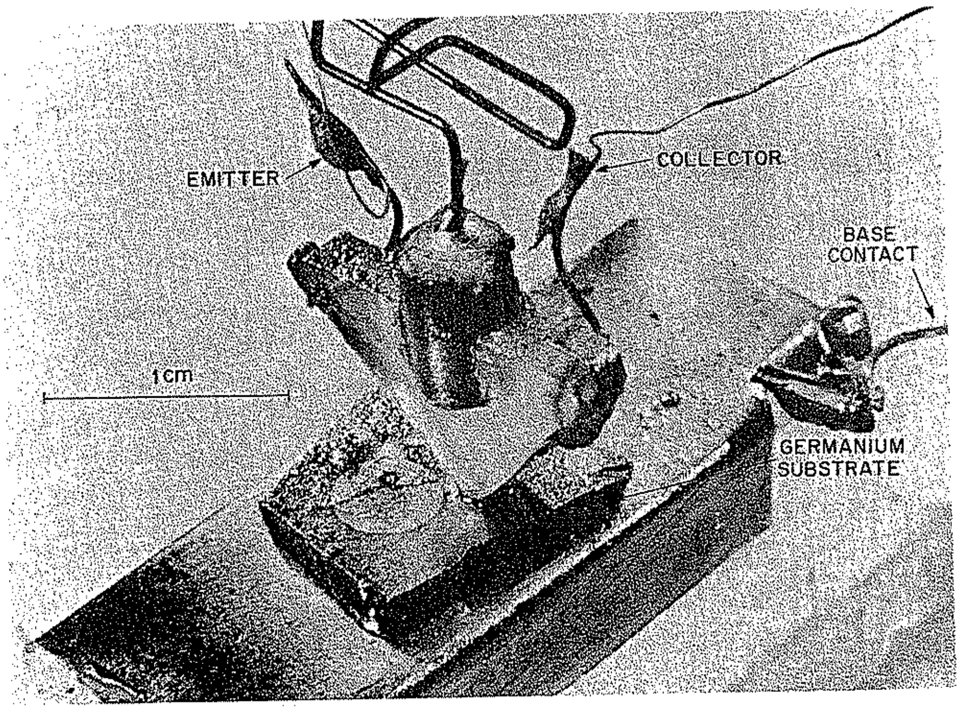
VIII Bipolar Transistors

The adjective bipolar indicates that both electrons and holes participate in the conduction process. In contrast, the FETs discussed in the previous chapter are unipolar devices since channel conduction is dominated by one type of carrier.

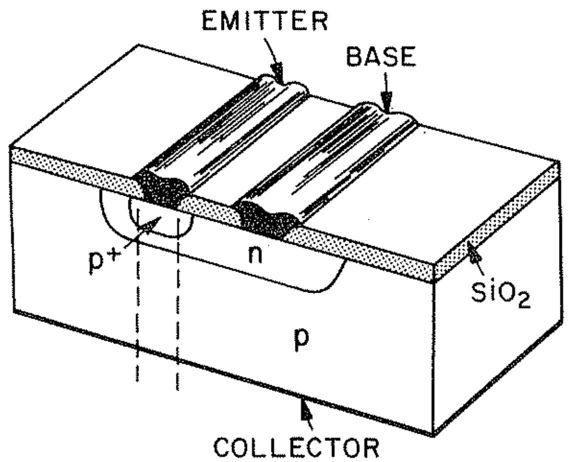
The first bipolar transistor was a point-contact device, which used two metal wires with sharp points contacting a Germanium substrate (W. Brattain and J. Bardeen, 1947). A picture of the first transistor is shown below.

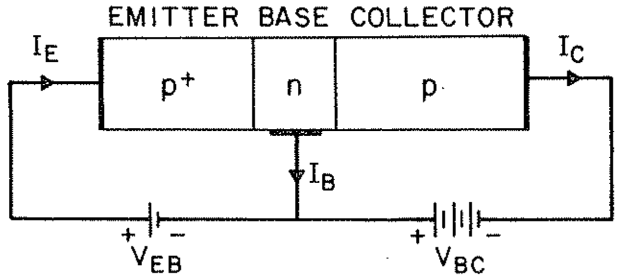


The first "practical" transistor, the Bipolar Junction Transistor (BJT), was invented in 1949 by W. Shockley.
A BJT consists of two closely coupled p-n junctions and thus comes in 2 varieties, a p-n-p or an n-p-n device.
A related bipolar device using three closely coupled p-n junctions is the Thyristor. This p-n-p-n structure can be switched between a high-impedance "off" state and a low impedance "on" state and derives its name from a gas-filled tube with similar bistable characteristics called "Thyratron". Thyristors are frequently used in light dimmers, motor controls and high-voltage switches. The operation of a thyristor will briefly be discussed at the end of this chapter after we have learned about the functionality of the BJT.

8.1 BJT Operation

The figure on the next page shows a perspective view of a p-n-p type (silicon) BJT and an idealized 1-dimensional equivalent structure.

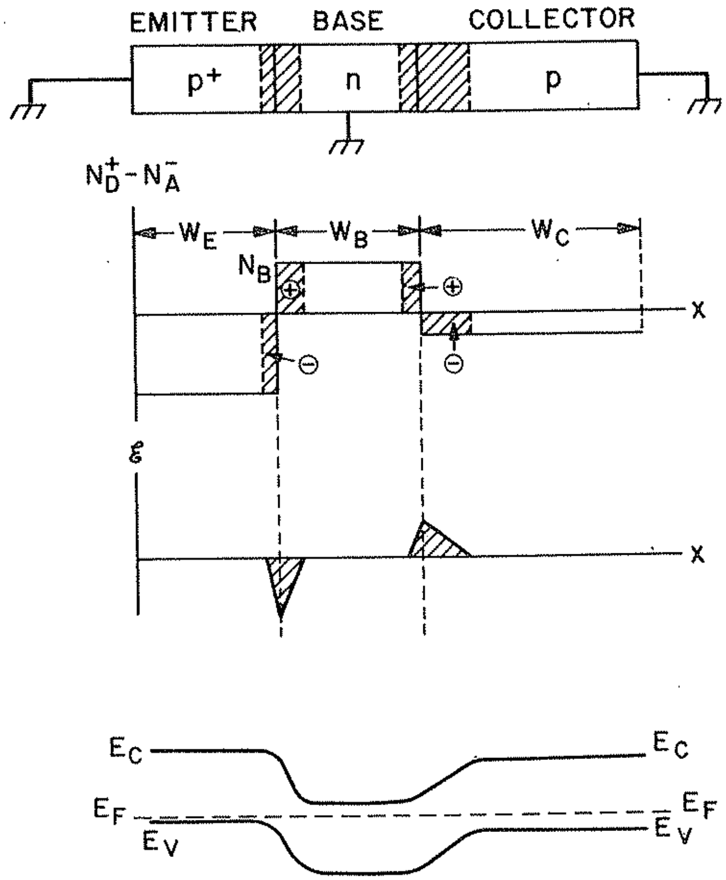




The 1-dimensional p-n-p structure represents a section of the "real" transistor between the dashed lines in the above picture. Each of the three basic regions, Emitter, Base and Collector, is assumed to be uniformly doped. The arrows in the 1-dimensional rendering indicate the current flows under normal conditions, also called "forward active" mode.
The complementary BJT structure, the n-p-n transistor, shows reverse current flows and bias voltages. However, its operation follows the same rules that control the p-n-p device.

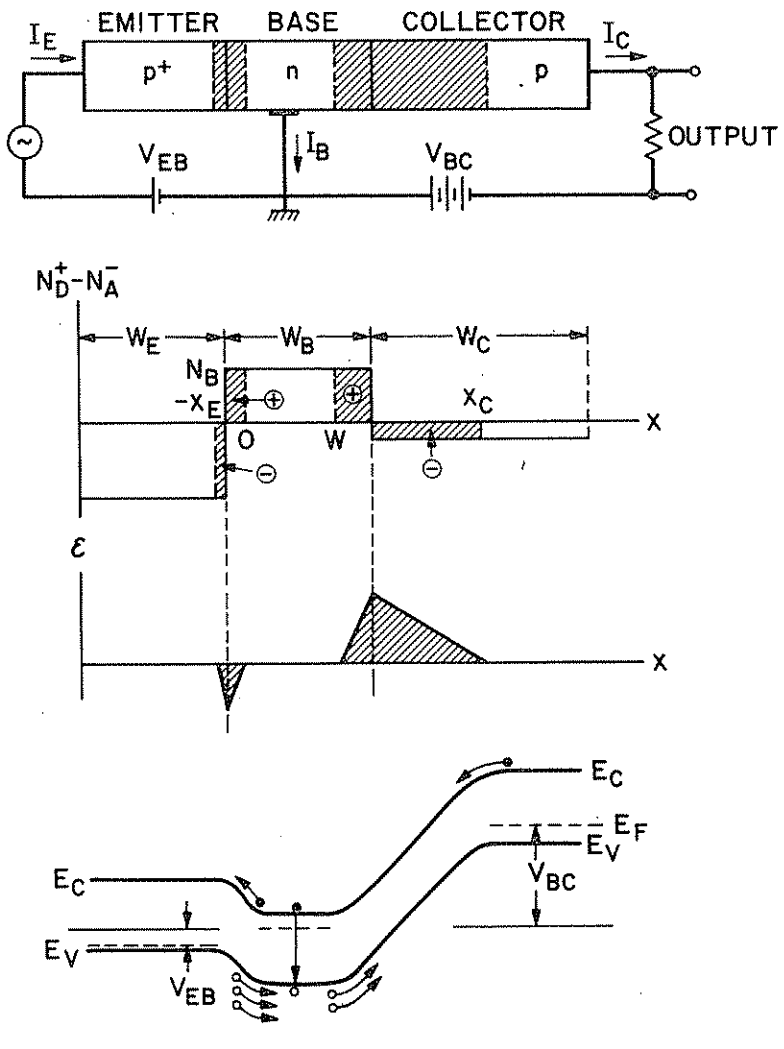
8.1.1 Forward Active Mode

The picture below shows the p-n-p device under thermal equilibrium conditions, i.e., when all 3 leads are grounded.



Note that the emitter is more heavily doped than the collector, while the base doping is less than the emitter doping, but greater than the collector doping.

The figure below illustrates the corresponding situations when the transistor is biased in forward active mode.

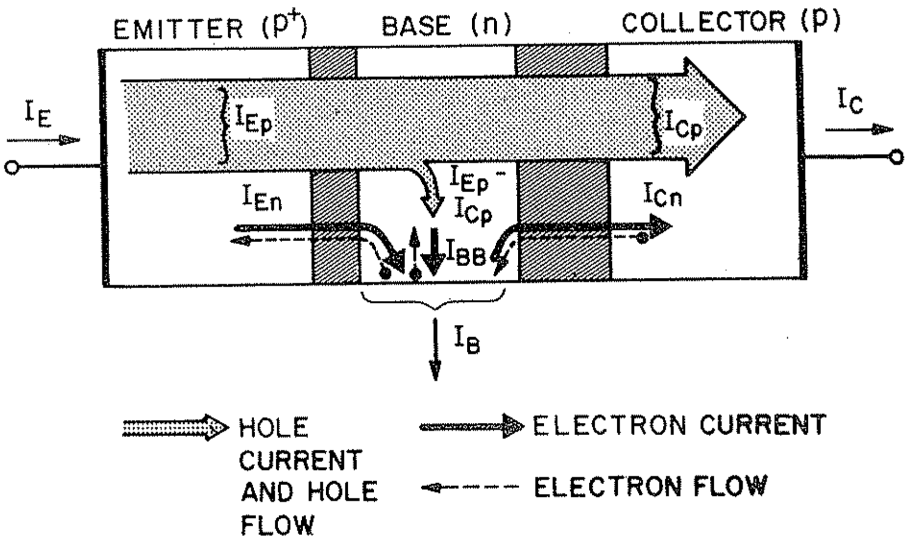


we note that the depletion layer width of the emitter - base junction is narrower while the collector - base depletion layer is wider when compared to the previous equilibrium situation.

Under forward active bias, holes are emitted from the emitter into the base, and electrons are injected from the base into the emitter. If we neglect R-G processes, these two current components constitute the total emitter current. The collector-base junction is reverse biased and a small reverse saturation current will flow across the junction. However, if the base is sufficiently narrow, holes injected (emitted) from the emitter can diffuse through the base to reach the base-collector depletion edge and then "float up" into the collector (holes behave like bubbles). This transport mechanism explains the terminology of "emitter," which emits the carriers, and "collector," which collects the carriers injected from a nearby junction. If most of the injected holes reach the collector without recombining with electrons in the base region, the collector hole current closely matches the emitter hole current. In other words, the carriers injected from the nearby emitter junction can cause a large current flow in the reverse-biased collector junction. This phenomenon is the "transistor action." To be effective, the two junctions involved in this process have to be very closely spaced, otherwise the injected holes recombine in the base before reaching the base-collector junction.

8.1.2 Current Gain

The figure below illustrates the various current components in an ideal p-n-p transistor under forward active bias conditions.



The holes injected from the emitter constitute the current I_{Ep}, the largest current component in a well designed transistor. Most of the injected holes will reach the base-collector junction and give rise to the current I_{Cp}. The difference $I_{Ep} - I_{Cp}$ represents the fraction of the injected holes that recombines with electrons in the base forming the current I_{BB}. Therefore, $I_{BB} = I_{Ep} - I_{Cp}$. The remaining two current components are I_{En}, arising from electrons being injected from the base to the emitter, and I_{Cn}, caused by thermally generated electrons near the collector-base junction edge drifting from the collector to the base. The component I_{En} is not desired, as shown later, and can be minimized by realizing a heavily doped emitter.

According to the previous figure, we can express the three terminal currents as follows:

$$\boxed{I_E = I_{Ep} + I_{En}} \qquad (1)$$

$$\boxed{I_C = I_{Cp} + I_{Cn}} \qquad (2)$$

$$\boxed{I_B = I_E - I_C = I_{En} + (I_{Ep} - I_{Cp}) - I_{Cn}} \qquad (3)$$

We now define the common-base current gain α_0 as

$$\boxed{\alpha_0 \equiv \frac{I_{Cp}}{I_E}} \qquad (4)$$

Substituting (4) in (1) yields

$$\boxed{\alpha_0 = \frac{I_{Cp}}{I_{Ep} + I_{En}} = \left(\frac{I_{Ep}}{I_{Ep} + I_{En}}\right) \frac{I_{Cp}}{I_{Ep}}} \qquad (5)$$

With $\boxed{\gamma \equiv \frac{I_{Ep}}{I_E} = \frac{I_{Ep}}{I_{Ep} + I_{En}}}$ Emitter Efficiency $\qquad (6)$

and $\boxed{\alpha_T \equiv \frac{I_{Cp}}{I_{Ep}}}$ Base Transport Factor $\qquad (7)$

we obtain

$$\boxed{\alpha_0 = \gamma \cdot \alpha_T} \qquad (8)$$

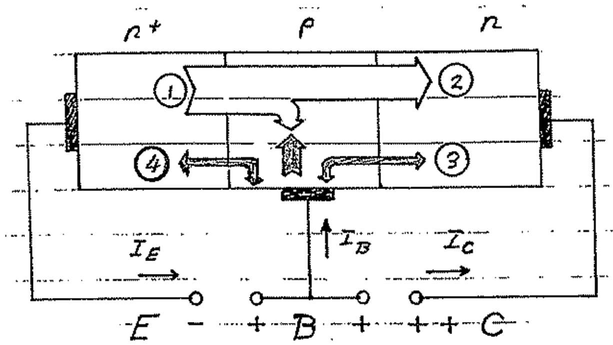
For a well-designed transistor, both γ and α_T approach unity so that α_0 is very close to 1.

Finally, by substituting (6) and (7) into (2), we can describe the collector current I_C as

$$\boxed{I_C = \alpha_T I_{Ep} + I_{Cn} = \alpha_0 I_E + I_{Cn}} \qquad (9)$$

where I_{Cn} corresponds to the collector-base current for an open emitter (i.e., $I_E = 0$). Therefore, I_{Cn} is also called the leakage current I_{CBO} of the reverse biased collector-base junction.

The figure below shows the current flow in a forward biased n-p-n transistor



① I_{En} Current due to electrons injected from emitter

② I_{Cn} Current due to electrons collected by collector

③ I_{Cp} Hole flow from base to collector

④ I_{Ep} Hole flow from base to emitter

8.1.3 Ideal Transistor Currents

In deriving the transistor currents, we will make the following assumptions:

1. Each region is uniformly doped.
2. Low-level injection prevails.
3. RG processes in depletion regions are negligible.
4. There are no series resistances in the device.

Under forward active bias, we can describe the minority carrier distribution in the neutral (field-free) base region as

Continuity eq.
$$\boxed{D_p \frac{d^2 p_n}{dx^2} - \frac{p_n - p_{no}}{\tau_p} = 0} \qquad (11)$$

Diffusion const. (D_p) minority carrier lifetime (τ_p)

If we apply the boundary conditions

$$\boxed{p_n(0) = p_{no}\, e^{\frac{V_{EB}}{V_T}}} \qquad p_{no} = \frac{n_i^2}{N_{Base}} \qquad (12)$$

and

$$\boxed{p_n(W) = 0} \qquad W: \text{Base width} \qquad (13)$$

to eq. (11), we obtain the solution

$$\boxed{p_n(x) = p_{no}\left[\left(e^{\frac{V_{EB}}{V_T}} - 1\right)\frac{\sinh\left(\frac{W-x}{L_p}\right)}{\sinh\left(\frac{W}{L_p}\right)} + 1 - \frac{\sinh\left(\frac{x}{L_p}\right)}{\sinh\left(\frac{W}{L_p}\right)}\right]} \qquad (14)$$

where $0 \le x \le W$ denotes the base region

and $L_p = \sqrt{D_p \tau_p}$ is the hole diffusion length

Minority Carrier Distribution in Base

Generic solution of minority carrier continuity equation: $(d^2 p_n/dx^2 = p_n + C_0)$

$$\left| p_n(x) = C_1 e^{\frac{x}{L_p}} + C_2 e^{-\frac{x}{L_p}} + p_{no} \right| \quad 0 \le x \le W \qquad (18a)$$

Bound. Cond.

$$\left| p_n(0) = p_{no} e^{\frac{V_{EB}}{V_T}} \right| \qquad p_{no} = \frac{n_i^2}{N_B} \qquad (18b)$$

$$\left| p_n(W) = 0 \right| \qquad (18c)$$

From (18b) we obtain

$$\left| C_1 + C_2 = p_{no}\left(e^{\frac{V_{EB}}{V_T}} - 1\right) \right| \qquad (18d)$$

while (18c) yields

$$\left| C_1 e^{\frac{W}{L_p}} + C_2 e^{-\frac{W}{L_p}} = -p_{no} \right| \qquad (18e)$$

$$\left| C_2 = p_{no}\left(e^{\frac{V_{EB}}{V_T}} - 1\right) - C_1 \right. \qquad (18f)$$

$$\left| C_1\left[e^{\frac{W}{L_p}} - e^{-\frac{W}{L_p}}\right] = -p_{no}\left[\left(e^{\frac{V_{EB}}{V_T}} - 1\right)e^{-\frac{W}{L_p}} + 1\right] \right| \qquad (18g)$$

$$\left\| C_1 = -p_{no} \frac{\left(e^{V_{EB}/V_T} - 1\right)e^{-W/L_p} + 1}{\left(e^{W/L_p} - e^{-W/L_p}\right)} \right\| \qquad (18h)$$

$$\left\| C_2 = p_{no} \frac{\left(e^{V_{EB}/V_T} - 1\right)e^{W/L_p} + 1}{\left(e^{W/L_p} - e^{-W/L_p}\right)} \right\| \qquad (18i)$$

$$\left\| p_n(x) = p_{no}\left(e^{\frac{V_{EB}}{V_T}} - 1\right)\frac{\sinh\left(\frac{W-x}{L_p}\right)}{\sinh\left(\frac{W}{L_p}\right)} + p_{no}\left[1 - \frac{\sinh\left(\frac{x}{L_p}\right)}{\sinh\left(\frac{W}{L_p}\right)}\right] \right\| \qquad (18j)$$

For $W/L_p \ll 1$ (thin base) eq. (14) approaches a straight line $\rightarrow p_n(x) \cong p_{no}\, e^{V_{EB}/V_T}\left(1-\frac{x}{W}\right)$.

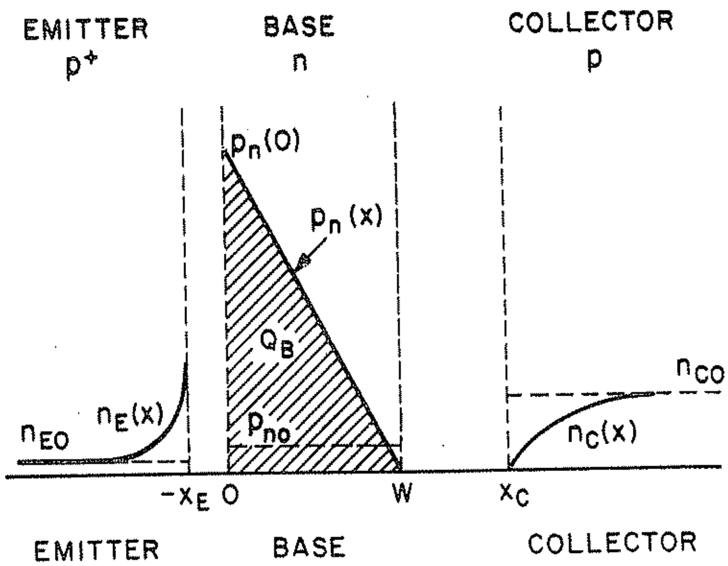
The minority carrier distributions of electrons in the emitter and collector regions, respectively, are given by the following equations

$$\left| n_E(x) = n_{EO}\left[1+\left(e^{\frac{V_{EB}}{V_T}}-1\right)e^{\frac{x+x_E}{L_E}}\right]\right| \quad x \leq -x_E \qquad (15)$$

and

$$\left| n_C(x) = n_{CO}\left[1-e^{-\frac{x-x_C}{L_C}}\right]\right| \qquad x \geq x_C \qquad (16)$$

The figure below illustrates typical minority carrier distributions across the emitter, base and collector regions of a p-n-p transistor.



The stored minority carrier charges in the base are of particular practical interest (switching behavior).

If A represents the cross-sectional area of the device, we can express the minority carrier charge stored in the base as

$$p_n(0) = p_{no}\, e^{\frac{V_{EB}}{V_T}}$$

$$\boxed{Q_B = qA\int_0^W [p_n(x) - p_{no}]\,dx \cong \frac{1}{2}\, qAW p_n(0)} \qquad (17)$$

Having established all relevant minority carrier distribution functions, we can now calculate the various (diffusion) current components via the gradient of the carrier distribution functions.

$$\boxed{\begin{aligned} I_{Ep} &= A\left(-qD_p \frac{dp_n}{dx}\Big|_{x=0}\right) \\ &= qA\frac{D_p p_{no}}{L_p}\coth\left(\frac{W}{L_p}\right)\left[e^{\frac{V_{EB}}{V_T}} - 1 + \frac{1}{\cosh\left(\frac{W}{L_p}\right)}\right] \end{aligned}} \qquad (18)$$

Assuming $W/L_p \ll 1$ and replacing p_{no} by $\frac{n_i^2}{N_B}$ yields

$$\boxed{I_{Ep} \cong qA\frac{D_p n_i^2}{W N_B}\left(e^{\frac{V_{EB}}{V_T}} - 1\right) + qA\frac{D_p n_i^2}{W N_B}} \qquad (19)$$

Following the same procedure, we can express the hole current I_{Cp} at $x = W$ by

$$\boxed{\begin{aligned} I_{Cp} &= qA\frac{D_p n_i^2}{L_p N_B}\frac{1}{\sinh\left(\frac{W}{L_p}\right)}\left[\left(e^{\frac{V_{EB}}{V_T}} - 1\right) + \cosh\left(\frac{W}{L_p}\right)\right] \\ W/L_p \ll 1 \qquad &\cong qA\frac{D_p n_i^2}{W N_B}\left[\left(e^{\frac{V_{EB}}{V_T}} - 1\right) + 1\right] \end{aligned}} \qquad (20)$$

Similarly, the electron current I_{En} (base-emitter) and I_{Cn} (collector-base) can be expressed as

$$\left| I_{En} \cong qA \frac{D_E n_{E0}}{L_E} \left(e^{\frac{V_{EB}}{V_T}} - 1\right) \right| \tag{21}$$

and

$$\left| I_{Cn} \cong qA \frac{D_C n_{C0}}{L_C} \right| \tag{22}$$

where D_E and D_C are the diffusion constants in the emitter and collector region, respectively. The terminal currents can now be obtained by combining eq. (19)-(22) appropriately.

$$\frac{W}{L_p} \ll 1 \qquad \left| \begin{aligned} I_E &= I_{Ep} + I_{En} \qquad n_{E0} = \frac{n_i^2}{N_E} \\ &\cong qA\left[\frac{D_p n_i^2}{W N_B} + \frac{D_E n_{E0}}{L_E}\right]\left[e^{\frac{V_{EB}}{V_T}} - 1\right] + qA\frac{D_p n_i^2}{W N_B} \end{aligned} \right| \tag{23}$$

$$\frac{W}{L_p} \ll 1 \qquad \left| \begin{aligned} I_C &= I_{Cp} + I_{Cn} \qquad n_{C0} = \frac{n_i^2}{N_C} \\ &\cong qA\frac{D_p n_i^2}{W N_B}\left[e^{\frac{V_{EB}}{V_T}} - 1\right] + qA\left[\frac{D_p n_i^2}{W N_B} + \frac{D_C n_{C0}}{L_C}\right] \end{aligned} \right| \tag{24}$$

$$\frac{W}{L_p} \ll 1 \qquad \left| \begin{aligned} I_B &= I_E - I_C \\ &\cong qA\frac{D_E n_{E0}}{L_E}\left[e^{\frac{V_{EB}}{V_T}} - 1\right] - qA\frac{D_C n_{C0}}{L_C} \end{aligned} \right| \tag{25}$$

It is obvious from the previous treatment that all three terminal currents are related to the minority carrier distribution in the base region. For a well designed transistor, the static emitter and collector currents reduce to terms proportional to the minority carrier gradient at $x=0$ and $x=W$, respectively. For example, if we neglect I_{En}, we can approximate the emitter current by

$$\boxed{I_E \cong -qAD_p \left.\frac{dn_p}{dx}\right|_{x=0} \cong qAD_p \frac{n_p(0)}{W} = \frac{2D_p}{W^2} Q_B} \quad (26)$$

In other words, the emitter current (and so the collector current) under forward active biasing conditions is directly proportional to the minority carrier charge in the base.

The derived current relationships can be summarized by the following statements:

1) The applied bias voltages control the boundary densities through the term $e^{\frac{V}{V_T}}$ $(V_T = \frac{kT}{q})$.
2) The emitter and collector currents are given by the minority carrier gradient at the junction boundaries ($x=0$ and $x=W$).
3) The base current is the difference between the emitter current and the collector current (recall KCL).

By combining eq. (19) and (21) we can approximate the emitter efficiency by

$$\gamma \equiv \frac{I_{Ep}}{I_{Ep} + I_{En}} \cong \frac{1}{1 + \frac{D_E}{D_p} \cdot \frac{n_{E0} N_B}{n_i^2} \frac{W}{L_E}} = \frac{1}{1 + \frac{D_E}{D_p} \frac{N_B}{N_E} \frac{W}{L_E}} \qquad (27)$$

where $N_B = \frac{n_i^2}{p_{n0}}$ is the base doping concentration and $N_E = \frac{n_i^2}{n_{E0}}$ denotes the emitter doping concentration. Consequently, selecting $N_E \gg N_B$ keeps the γ value close to unity.

The base transport factor α_T can be approximated by combining eq. (18) and (20). This yields

$$\alpha_T \equiv \frac{I_{Cp}}{I_{Ep}} \cong \left[\cosh\left(\frac{W}{L_p}\right) \sinh\left(\frac{W}{L_p}\right)\right]^{-1} \cong 1 - \frac{W^2}{2L_p^2} \qquad (28)$$

Since we cannot alter L_p significantly, keeping α_T close to unity requires a very small base width W.

Numerical Example

A p-n-p transistor is characterized by the following physical parameters: $N_E = 10^{25} m^{-3}$, $N_B = 10^{23} m^{-3}$, $N_C = 5 \cdot 10^{21} m^{-3}$, $D_E = 10^{-4} \frac{m^2}{Vs}$, $D_p = 10^{-3} \frac{m^2}{Vs}$, $L_E = 10^{-6} m$, $L_p = 10^{-5} m$, $W = 5 \cdot 10^{-7} m$

$\therefore$ $\gamma = 0.9995$, $\alpha_T = 0.9987$, $\alpha_0 = \gamma \cdot \alpha_T = 0.9982$

8.1.4 Modes of Operation

A BJT has 4 modes of operation. They are

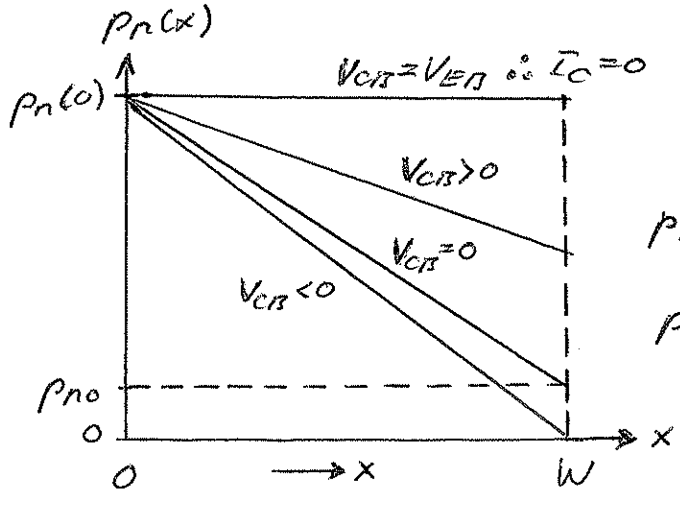
- Forward active
- Saturation
- Cutoff
- Reverse active

Forward active

In this mode the emitter-base junction is forward biased and the collector-base junction is reverse polarized. In actuality the CB junction can be slightly forward biased. For silicon $V_{CB} < 0.5\,V$

Saturation

Both junctions are forward biased and the minority carrier concentration at $x = W$ becomes $p_n(W) = p_{no}\, e^{V_{CB}/V_T}$. This creates an accumulation of excess minority carriers in the base as illustrated below.



$$p_n(0) = p_{no} \cdot e^{\frac{V_{EB}}{V_T}}$$

$$p_{no} = \frac{n_i^2}{N_B}$$

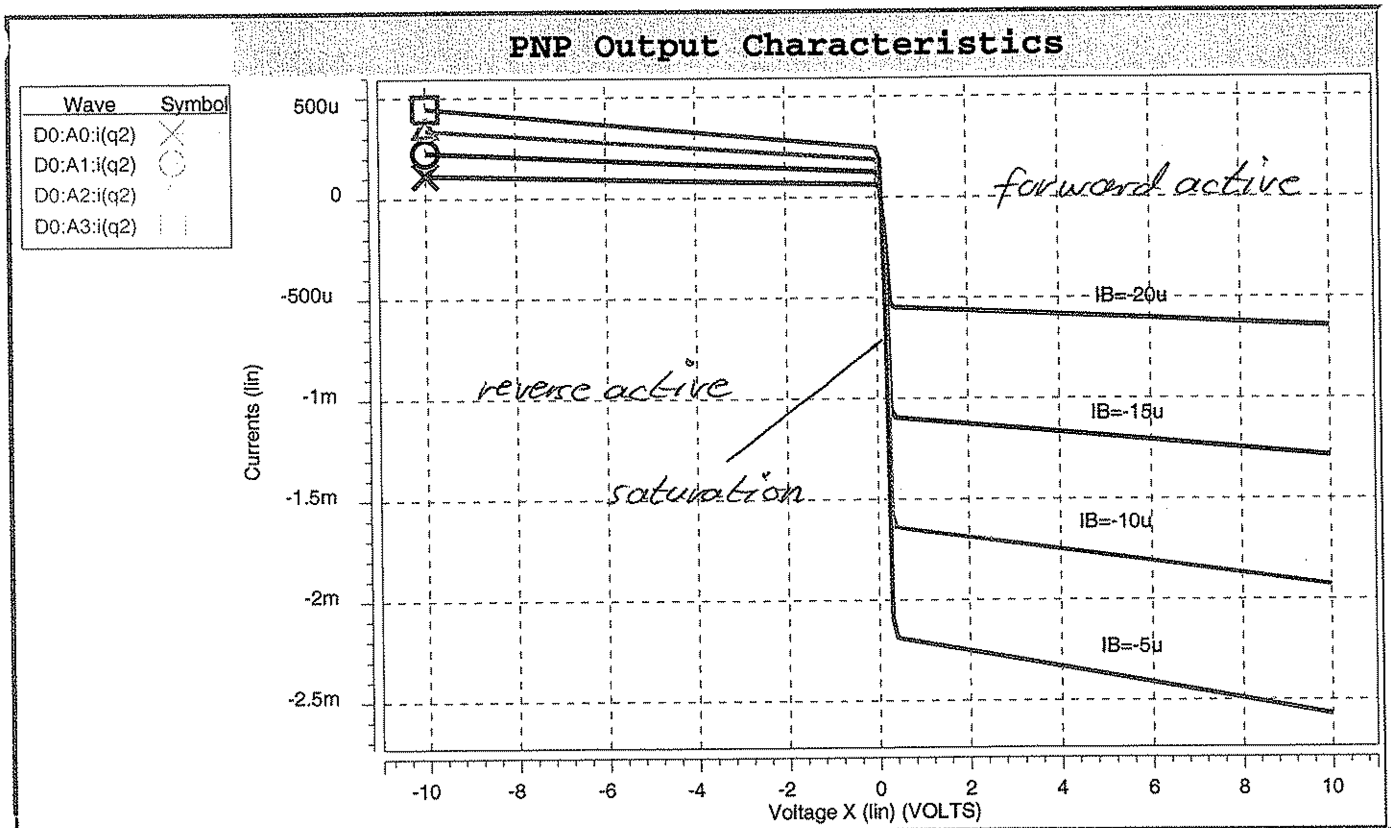
Cutoff

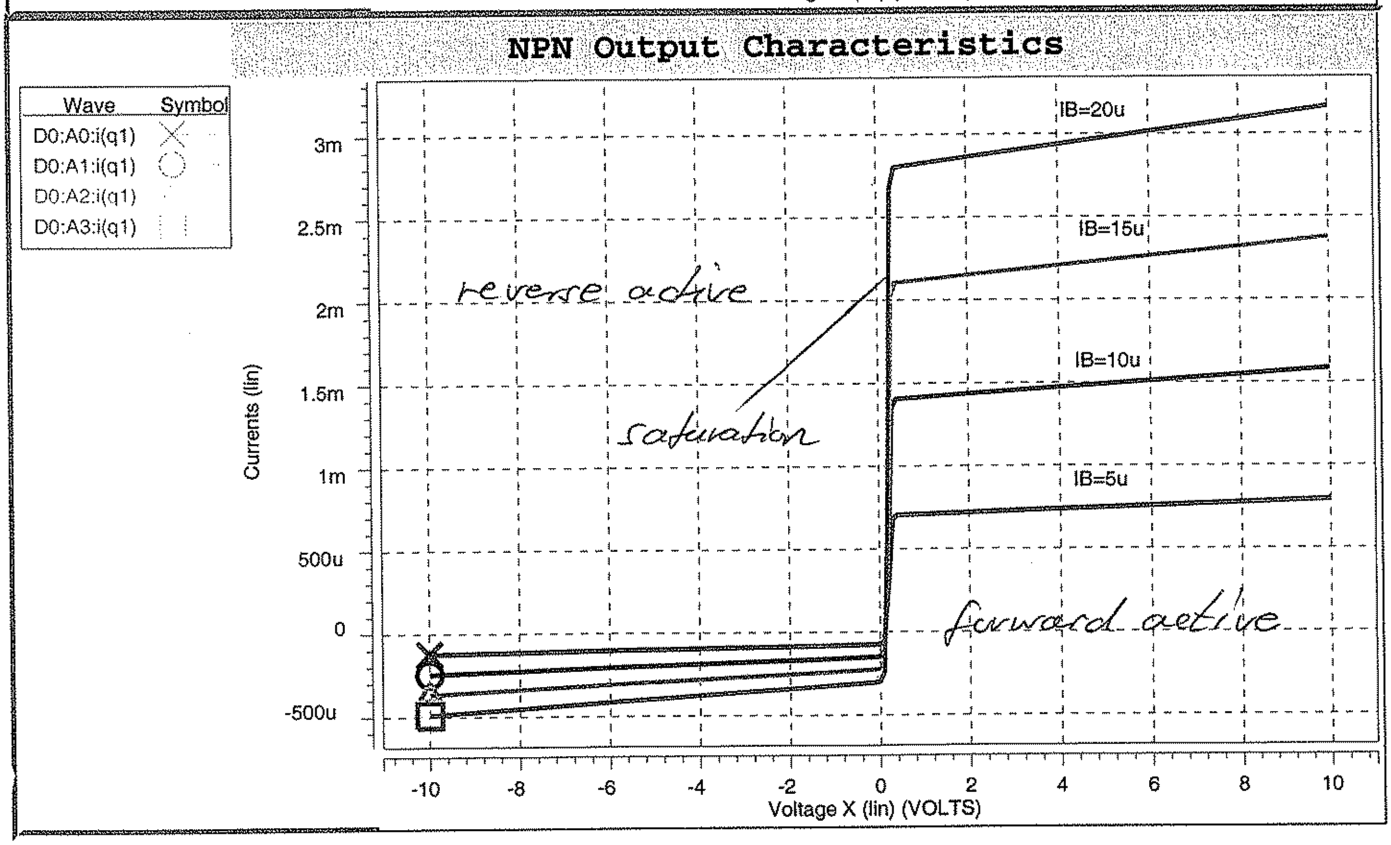
Both junctions are reverse biased and no minority carriers are stored in the base. Consequently, the collector current approaches zero. The emitter-collector terminal can thus be considered open, i.e., the device is off.

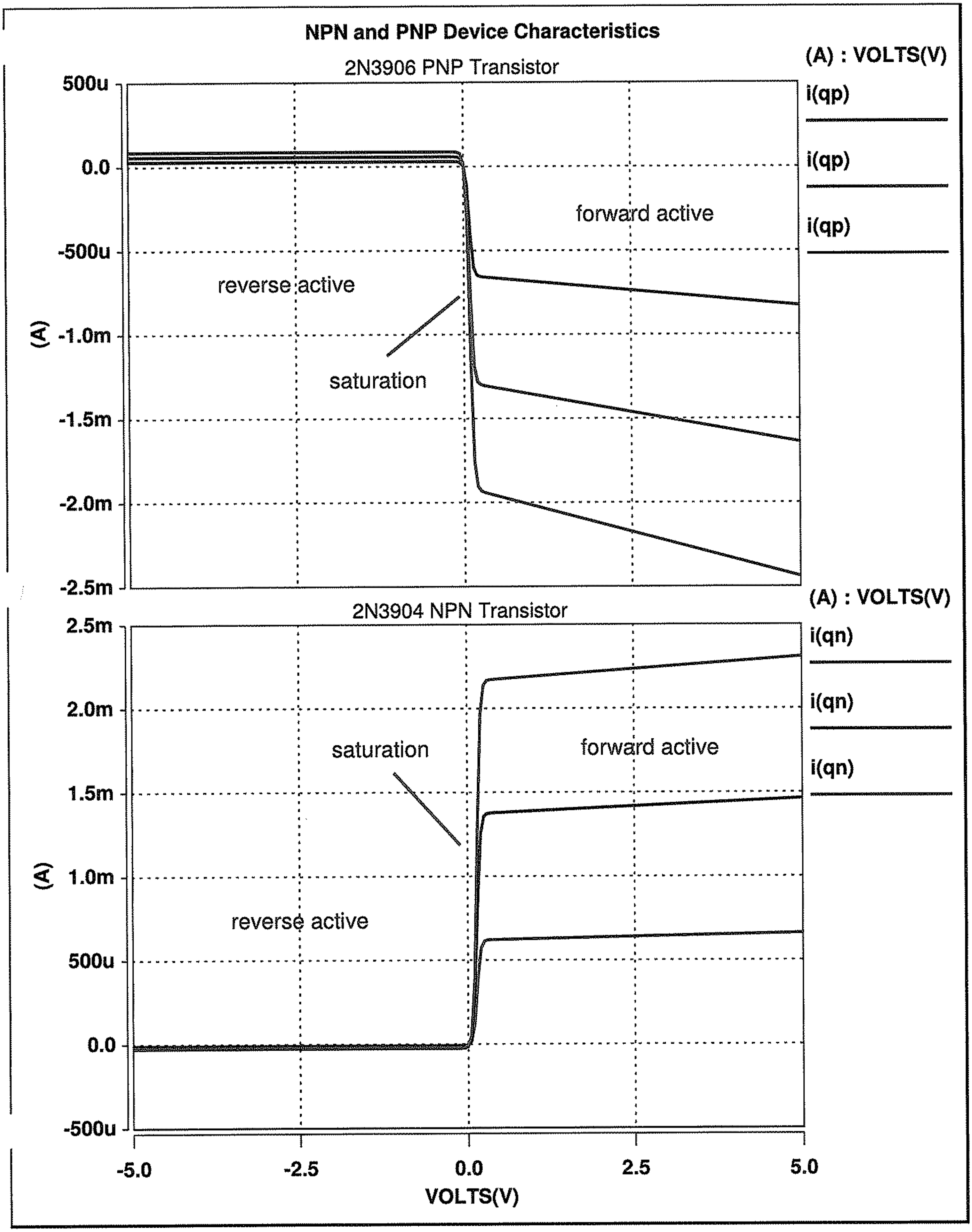
Reverse active

In this mode, the emitter-base junction is reverse biased and the collector-base junction is forward biased. In other words, the collector acts like the emitter and visa versa. However, due to the asymmetric doping of the emitter and collector region, the current gain, i.e., the ratio I_C/I_B is considerably smaller in this mode when compared to the forward active case (N_B/N_C is typically less than 1, which significantly degrades the emitter efficiency γ).

In linear applications (analog circuits) BJT's are preferably operated in the forward active mode, while digital applications switch the device from "on" (forward active) to cutoff. Due to the excess minority carrier storage in the base, digital applications should avoid saturation (e.g. Emitter-Coupled Logic).


PNP Output Characteristics
Wave Symbol
D0:A0:i(q2)
D0:A1:i(q2)
D0:A2:i(q2)
D0:A3:i(q2)
500u
0
-500u
-1m
-1.5m
-2m
-2.5m
Currents (lin)
-10
-8
-6
-4
-2
0
2
4
6
8
10
Voltage X (lin) (VOLTS)
forward active
IB=-20u
IB=-15u
IB=-10u
IB=-5u
reverse active
saturation

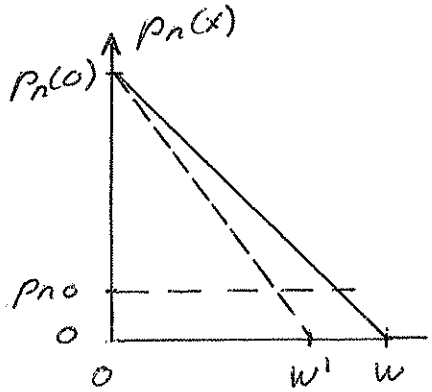

NPN Output Characteristics
Wave Symbol
D0:A0:i(q1)
D0:A1:i(q1)
D0:A2:i(q1)
D0:A3:i(q1)
3m
2.5m
2m
1.5m
1m
500u
0
-500u
Currents (lin)
-10
-8
-6
-4
-2
0
2
4
6
8
10
Voltage X (lin) (VOLTS)
IB=20u
IB=15u
IB=10u
IB=5u
reverse active
saturation
forward active


NPN and PNP Device Characteristics
2N3906 PNP Transistor
(A) : VOLTS(V)
i(qp)
i(qp)
i(qp)
500u
0.0
-500u
-1.0m
-1.5m
-2.0m
-2.5m
(A)
forward active
reverse active
saturation
2N3904 NPN Transistor
(A) : VOLTS(V)
i(qn)
i(qn)
i(qn)
2.5m
2.0m
1.5m
1.0m
500u
0.0
-500u
(A)
saturation
forward active
reverse active
-5.0
-2.5
0.0
2.5
5.0
VOLTS(V)

8.1.5 Nonideal Effects

Base Width Modulation

Increasing the emitter-collector bias voltage of a p-n-p device results in a reduction of the effective base width since the base-collector depletion zone widens. This increases the gradient of the minority carrier density in the base region and, according to eq. (26), the emitter (and collector) current.



Recall

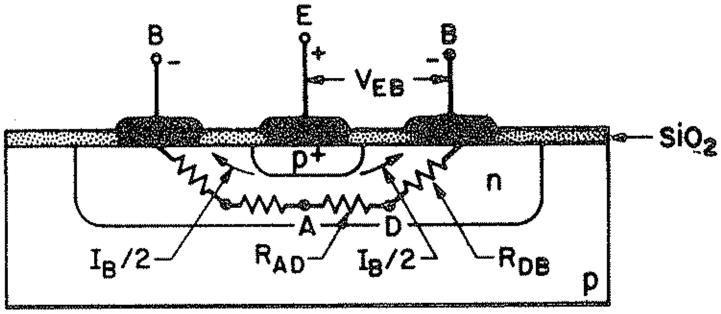
$$I_E \cong I_C \cong q A D_p \frac{p_n(0)}{W}$$

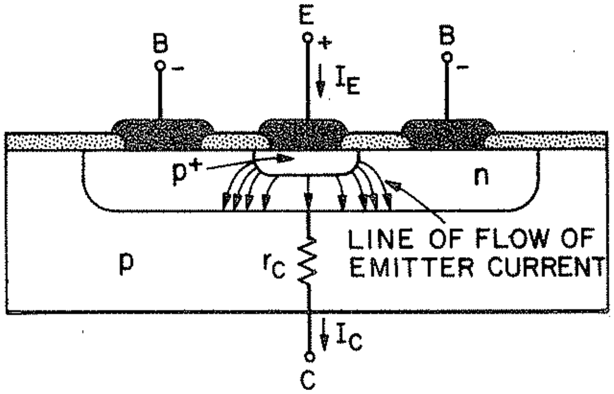
$$\therefore \frac{\Delta I}{I} \cong \frac{\Delta W}{W}$$

This base width modulation is known as the "Early Effect" and is expressed via the "Early Voltage". The latter is the (negative) intersection of the linearly extrapolated collector current in forward active mode with the emitter-collector voltage V_{EC}.

Base Resistance

High current gain requires a very thin base. This, in turn, can result in a significant base resistance. As illustrated by the figure below, electrons supplied from the base contacts flow toward the center of the emitter and cause the base-emitter voltage drops to vary with position along the BE junction.





The EB junction expiriences the largest forward bias at the edge of the emitter (point D). This causes the highest injection of holes and so the highest current flow at the edge → "emitter crowding." While this crowding can reduce the base resistance, it can give rise to high-injection effects such as reduced emitter efficiency. To minimize crowding, the emitter region should exhibit a high perimeter-to-area ratio. This can be achieved with an interdigitated geometry between emitter and base.

R-G Current and High-Injection

In a real transistor, R-G processes in the reverse biased BC junction are contributing to additional current leakage.

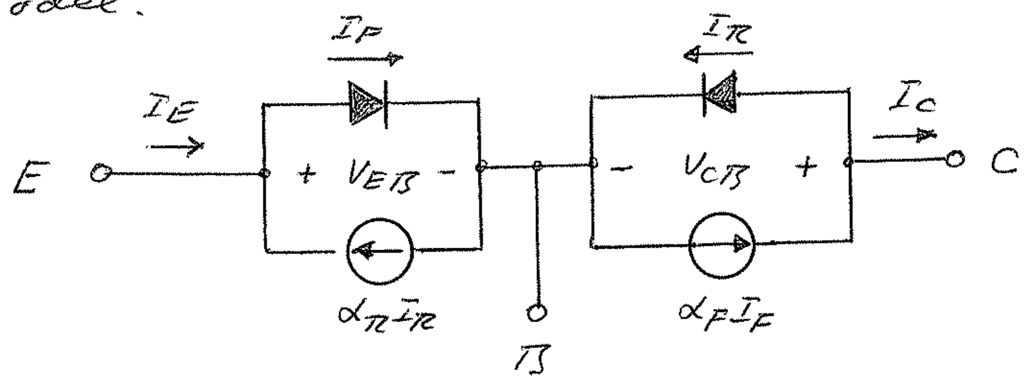
R-G processes in the forward biased EB junction can have a profound effect on the current gain. At low current levels, the recombination current can dominate so that I_B varies more like $\exp\left(\frac{V_{EB}}{mV_T}\right)$, with $m \cong 2$. The collector current I_C is not affected by the EB recombination current, since I_C is primarily due to holes injected into the base that diffuse to the collector. This causes a drop in the

current gain ratio I_C/I_B. The current gain at low current levels can be improved by minimizing R-G centers (impurities) in the device.
At very high current levels, the (injected) minority carrier density in the base approaches N_B, which increases the effective base doping and thus lowers the emitter efficiency according to eq. (27).

8.2 BJT Frequency Response

8.2.1 The Ebers-Moll Model

By replacing the BJT by two (coupled) back-to-back connected diodes, one obtains the following model.



The two diode currents are

$$\boxed{I_F = I_{F0}\left(e^{\frac{V_{EB}}{V_T}} - 1\right)} \qquad (29)$$

$$\boxed{I_R = I_{R0}\left(e^{\frac{V_{CB}}{V_T}} - 1\right)} \qquad (30)$$

The two terminal currents, I_E and I_C, can now be written as

$$\boxed{I_E = I_{F0}\left(e^{\frac{V_{EB}}{V_T}}-1\right) - \alpha_R I_{R0}\left(e^{\frac{V_{CB}}{V_T}}-1\right)} \qquad (31)$$

$$\boxed{I_C = \alpha_F I_{F0}\left(e^{\frac{V_{EB}}{V_T}}-1\right) - I_{R0}\left(e^{\frac{V_{CB}}{V_T}}-1\right)} \qquad (32)$$

Comparing these two current expressions with the approximations in eq. (23) and (24)* yields the following equivalencies

$$\frac{W}{L_p} \ll 1 \qquad \boxed{I_{F0} \cong q A n_i^2 \left[\frac{D_p}{W N_B} + \frac{D_E}{L_E N_E}\right]} \qquad (33)$$

$$\frac{W}{L_p} \ll 1 \qquad \boxed{\alpha_R I_{R0} = \alpha_F I_{F0} \cong q A \cdot n_i^2 \frac{D_p}{W N_B}} \qquad (34)$$

$$\frac{W}{L_p} \ll 1 \qquad \boxed{I_{R0} = q A n_i^2 \left[\frac{D_p}{W N_B} + \frac{D_C}{L_C N_C}\right]} \qquad (35)$$

* While deriving the two approximate equations (23) and (24), we have assumed a strongly reverse biased CB junction so that $e^{\frac{V_{CB}}{V_T}} \ll 1$. The above current expressions are more generic since they also cover the reverse mode of operation, where $V_{CB} > 0$ and conversely $V_{EB} < 0$.

Note: Since $\alpha_R I_{R0} = \alpha_F I_{F0}$, the Ebers-Moll model, eq.(31) and (32), comprises only 3 device parameters, i.e., I_{F0}, I_{R0} and $\alpha_R I_{R0} = \alpha_F I_{F0}$.

$\alpha_F \cong \left(1 + \frac{D_E}{D_P}\frac{N_B}{N_E}\frac{W}{L_P} + \frac{1}{2}\frac{W^2}{L_P^2}\right)^{-1}$ and $\alpha_R \cong \left(1 + \frac{D_C}{D_P}\frac{N_B}{N_C}\frac{W}{L_C} + \frac{1}{2}\frac{W^2}{L_P^2}\right)^{-1}$ represent the forward and reverse common-base current gain, respectively (cf. eq.(8), $\alpha_0 = \gamma \cdot \alpha_T$).

By relating the two terminal currents to each other, we can rewrite the Ebers-Moll equations as

$$\boxed{I_E = (1-\alpha_R\alpha_F) I_{F0}\left(e^{\frac{V_{EB}}{V_T}} - 1\right) + \alpha_R I_C} \qquad (36)$$

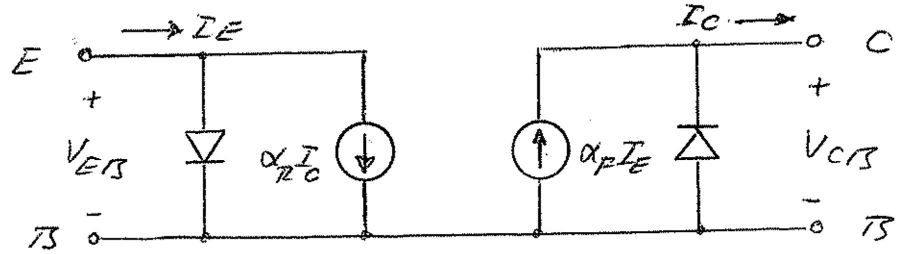
$$\boxed{I_C = \alpha_F I_E - (1-\alpha_R\alpha_F) I_{R0}\left(e^{\frac{V_{CB}}{V_T}} - 1\right)} \qquad (37)$$

The terms $(1-\alpha_R\alpha_F) I_{F0}$ and $(1-\alpha_R\alpha_F) I_{R0}$ can be abbreviated as I_{E0} and I_{C0}, respectively, where I_{E0} denotes the magnitude of the emitter saturation current when $I_C = 0$, and I_{C0} stands for the magnitude of the collector saturation current for $I_E = 0$. The Ebers-Moll equations thus become

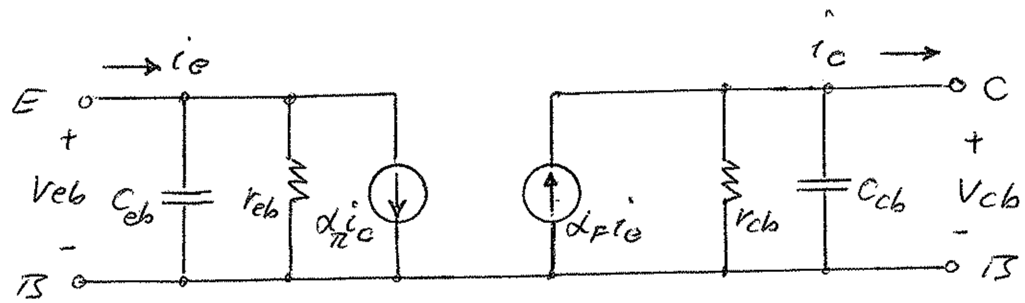
$$\boxed{I_E = I_{E0}\left(e^{\frac{V_{EB}}{V_T}} - 1\right) + \alpha_R I_C} \qquad (38)$$

$$\boxed{I_C = \alpha_F I_E - I_{C0}\left(e^{\frac{V_{CB}}{V_T}} - 1\right)} \qquad (39)$$

We can thus model the BJT (in common base configuration as follows



In the forward active mode of operation, we can replace the EB diode by a linear equivalent circuit consisting of a voltage source in series with the effective resistance $r_{eb} \cong \frac{V_T}{I_E}$. The reverse biased CB diode can be modeled by its saturation current I_{CBO}. The ac equivalent circuit of the BJT in CB mode thus becomes



where $\quad r_{eb} \cong \frac{V_T}{I_E} \qquad r_{cb} \cong \frac{\alpha_F}{(1-\alpha_F)} \frac{V_A}{I_C}$

$C_{eb} \cong \frac{\tau_B}{r_{eb}} \qquad C_{cb} \cong \sqrt{\frac{\varepsilon_s q N_C}{2(V_{bi}+V_{BC})}} \cdot A \quad N_B \gg N_C$

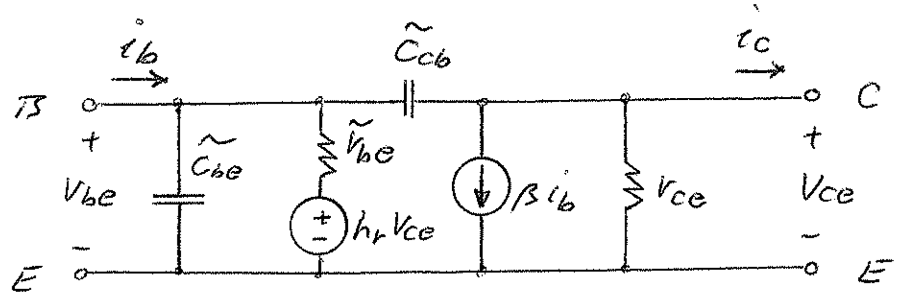
Base Transit Time - $\tau_B \approx \frac{1}{2} \frac{W^2}{D_p} \qquad V_A$: Early voltage

Note: α_F and α_R only act in the forward and reverse bias mode, respectively.

To improve the frequency response, C_{cb} must be kept as small as possible. Therefore, high-frequency transistors are designed with a small base width W. Furthermore, because the electron Diffusion const. in silicon is about 3 times larger than that of holes, high-frequency transistors are of the n-p-n type.

Due to the small input impedance, the common-base configuration is only used in high-frequency applications or in conjunction with a common-emitter common-base cascade → cascode amplifier.

The ac equivalent circuit for the CE configuration looks as follows.



where $\tilde{C}_{be} \cong \frac{Q_B}{V_{be}}$ $\qquad \tilde{C}_{cb} \cong \sqrt{\frac{\varepsilon_s q N_C}{2(V_{bi}+V_{BC})}} \cdot A$

$$\tilde{r}_{be} \cong \frac{\alpha_F}{(1-\alpha_F)} \frac{V_T}{I_C} \qquad r_{ce} \cong \frac{V_A}{I_C}$$

$$h_r \cong \frac{r_{be}}{r_{cb}} = \frac{V_T}{V_A} \qquad \beta = \frac{\alpha_F}{(1-\alpha_F)}$$

Most often, the high-frequency behavior of a BJT is indicated by the device's transit frequency f_T, where the common-emitter (forward) current gain $\beta_F = \frac{\alpha_F}{(1-\alpha_F)}$ becomes unity. This frequency is almost identical with the 3dB corner frequency of α_F. In fact, the latter is only $(1-\alpha_{F0})f_T$ larger, where α_{F0} denotes the low frequency value of α_F (recall that α_{F0} is only slightly smaller than 1). As a final note regarding the BJT's frequency behavior, it should be restated that the most limiting factor defining the frequency response is the transit time τ_B of the minority carriers across the base region. If we approximate the minority carrier distribution in the base region by a triangular function as depicted on page VIII-16, we can express the base transit time as

$$\boxed{\tau_B \cong \frac{1}{2}\frac{W^2}{D_p} = \frac{1}{2}\frac{W^2}{\mu_p V_T}} \quad p\text{-}n\text{-}p \qquad (40)$$

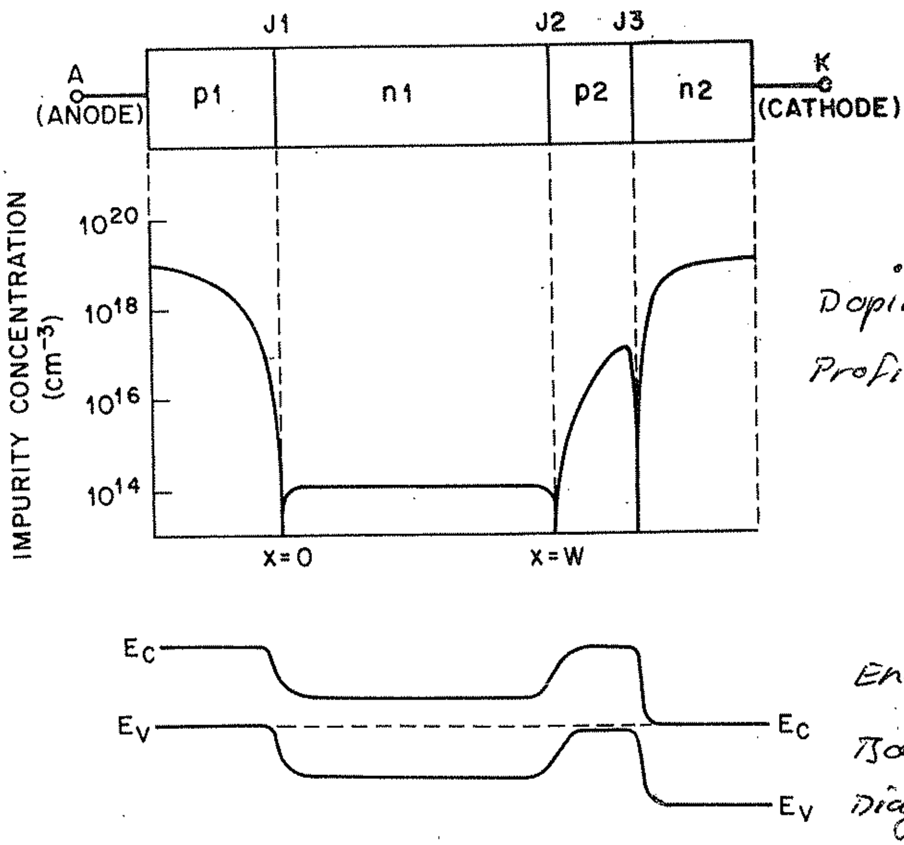
This relationship underlines the importance of minimizing the base width W in the design of high-frequency transistors, since the response time reduces as the square of W.

8.3 The Thyristor

The thyristor is closely related to the BJT, but features a different switching mechanism. Due to the more robust design, thyristors can switch currents ranging from a few mA to several thousand Amperes, and the voltage ratings extend beyond 10,000V.

8.3.1 Basic Operation

The figure below depicts a cross-sectional view of a thyristor structure with three pn junctions in series (J1, J2 and J3).

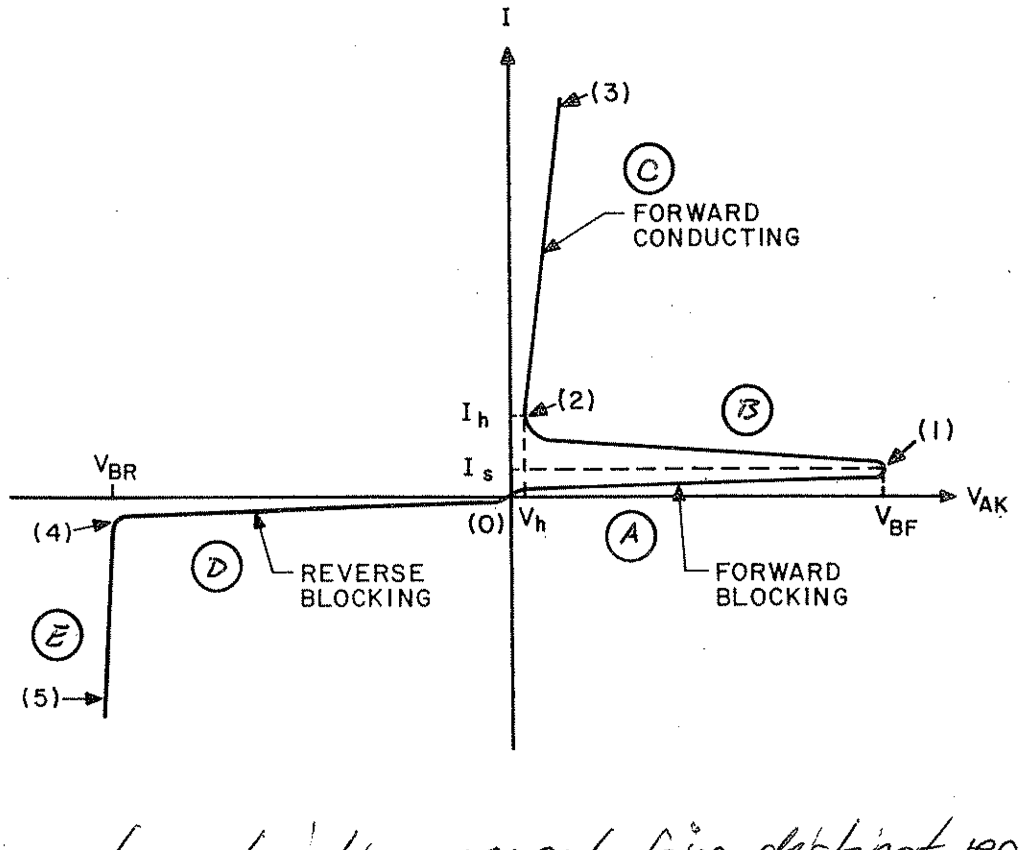


Doping Profile

Energy Band Diagram

If the depicted structure only features two external contacts (Anode and Cathode), it is referred to as a p-n-p-n diode. If a third contact, the Gate, is added to the inner p-layer, the three terminal device is called a Semiconductor Controlled Rectifier (SCR) or thyristor.

The I-V characteristics of a thyristor are shown below.



These characteristics reveal five distinct regions:

(A) The device is in the forward blocking or "off" state (high impedance state). Switching occurs where $dV/dI = 0$ (point 1 with $V = V_{BF}$ and $I = I_S$). (0-1)

(B) The device is in a negative resistance region, where the current increases while the voltage decreases sharply. (1-2) CB junction of npn Tr. breaks down

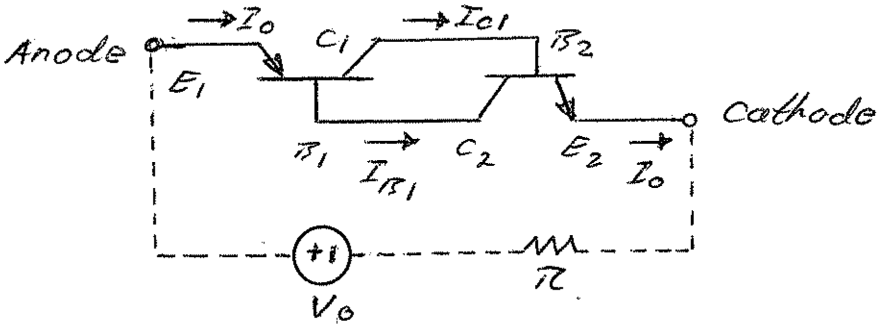
(C) The device is in the forward conducting or "on" state (low impedance state). Point 2, where $dV/dI = 0$, is defined by the holding voltage V_h and the holding current I_h. (2-3)

(D) The device is in the reverse blocking state. (0-4)

(E) The device is in the reverse breakdown region (4-5)

The p-n-p-n diode operated in the forward region is a bistable device that can switch from a high-impedance, low current state (off) to a low-impedance, high-current state (on) and visa versa.

To understand the forward-blocking characteristics, we will consider the device as two bipolar transistors, a p-n-p and an n-p-n transistor connected with the base of one transistor serving also as the collector of the other and visa versa as illustrated below.



The base current of the p-n-p transistor is

$$\left| \begin{aligned} I_{B1} = I_{E1} - I_{C1} &= (1-\alpha_1) I_{E1} - I_{CBO1} \\ &= (1-\alpha_1) I_0 - I_{CBO1} \end{aligned} \right| \qquad (41)$$

and the collector current of the n-p-n device is

$$\left| I_{C2} = \alpha_2 I_{E2} + I_{CBO2} = \alpha_2 I_0 + I_{CBO2} \right| \qquad (42)$$

where I_{CBO} denotes the collector-base leakage current while α_1 and α_2 are the common-base current gains, respectively.

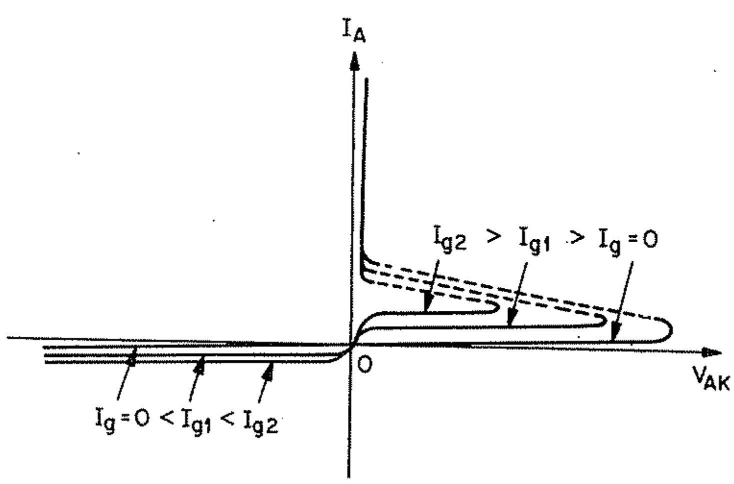
Equating I_{B1} to I_{C1} yields

$$\left| (1-\alpha) I_0 - I_{CBO1} = \alpha_2 I_0 + I_{CBO2} \right| \qquad (43)$$

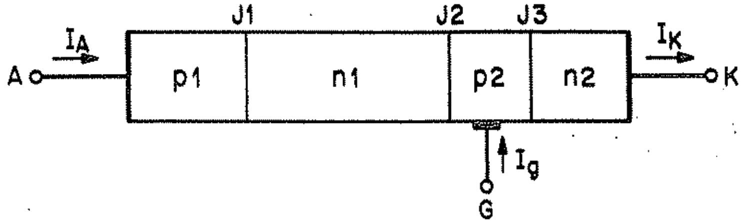
or

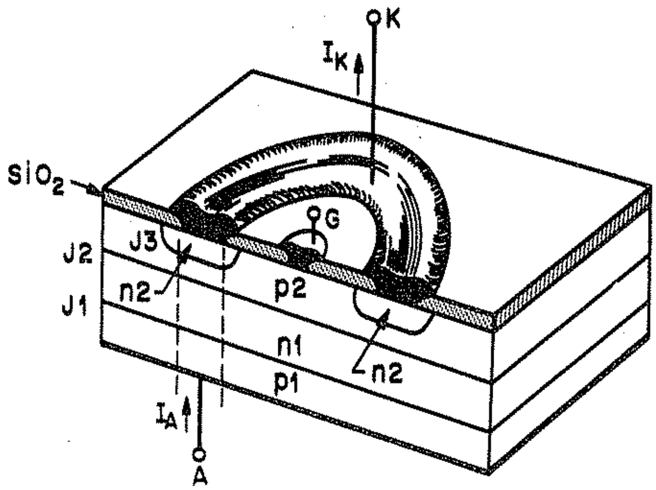
$$\left| I_0 = \frac{I_{CBO1} + I_{CBO2}}{1-(\alpha_1 + \alpha_2)} \right| \qquad (44)$$

As shown in section 8.1, the common-base current gain varies between 0 and 1 and generally increases with the current. At low current levels, both α_1 and α_2 approach zero and $I_0 \cong I_{CBO1} + I_{CBO2}$. As the bias voltage increases, I_0 increases, as do α_1 and α_2. This in turn causes I_0 to increase further (positive feedback or regenerative behavior). Eventually, the sum $\alpha_1 + \alpha_2$ approaches unity and the current increases drastically (forward breakover). Apart from the forward voltage excursion, domains Ⓐ and Ⓑ on page VIII-29, the p-n-p-n device or thyristor acts like a rectifying single p-n junction. By injection an additional bias current through the third "Gate" electrode, one can control the forward breakdown voltage V_{BF} as illustrated below.

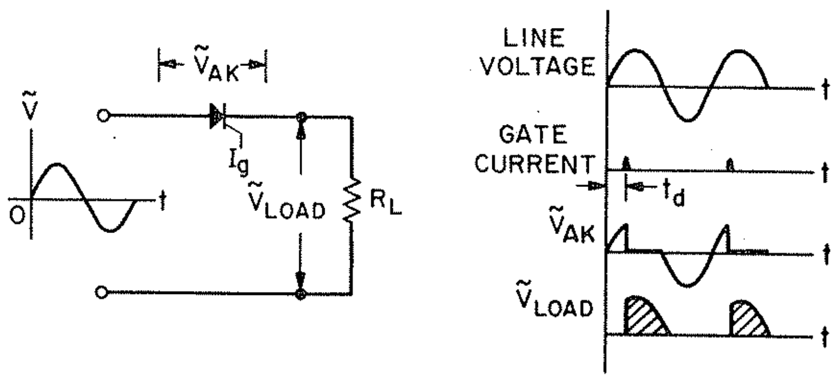


The practical fabrication of a planar three-terminal thyristor is shown below.





A simple application of a thyristor is depicted below.



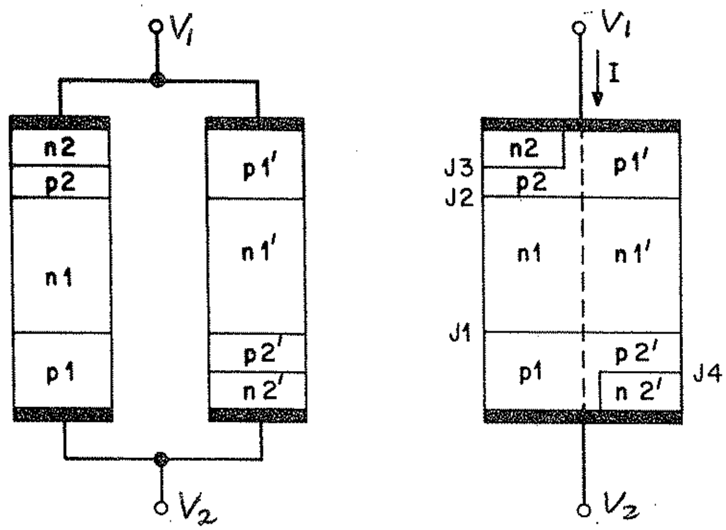
The thyristor is used here to deliver a variable power to a resistive load R_L from an ac source.

Note that the amount of power delivered to R_L depends on the timing of the gate current pulses. If the I_G pulse occurs near the positive transition of the source voltage (but after the zero crossing) almost all of the positive half-wave of the source voltage will drop across the load R_L. Conversely, if the current pulse arrives shortly before the negative zero crossing of the source voltage, almost no power will be delivered to the load.

8.5.2 Bidirectional Thyristors

A bidirectional thyristor is capable of switching positive and negative voltages and is thus very useful in ac (switched) power regulators.
A bidirectional p-n-p-n diode without a gate electrode is called a "DIAC" (DIode AC switch). It behaves like two conventional thyristors in parallel configuration, whereby the Anode of one device is connected to the cathode of the other and vica versa. The symmetrical design of the "diac" gurantees that the device performance is identical for either polarity of the applied voltage.

The figure below depicts a possible implementation of a "diac":



From a practical perspective, it is more interesting if the breakdown voltages can be controlled by a third gate electrode. Such a device is called "Triac" (Triode AC switch). A cross-sectional view of such a six-layer structure featuring 5 pn junctions is shown below.