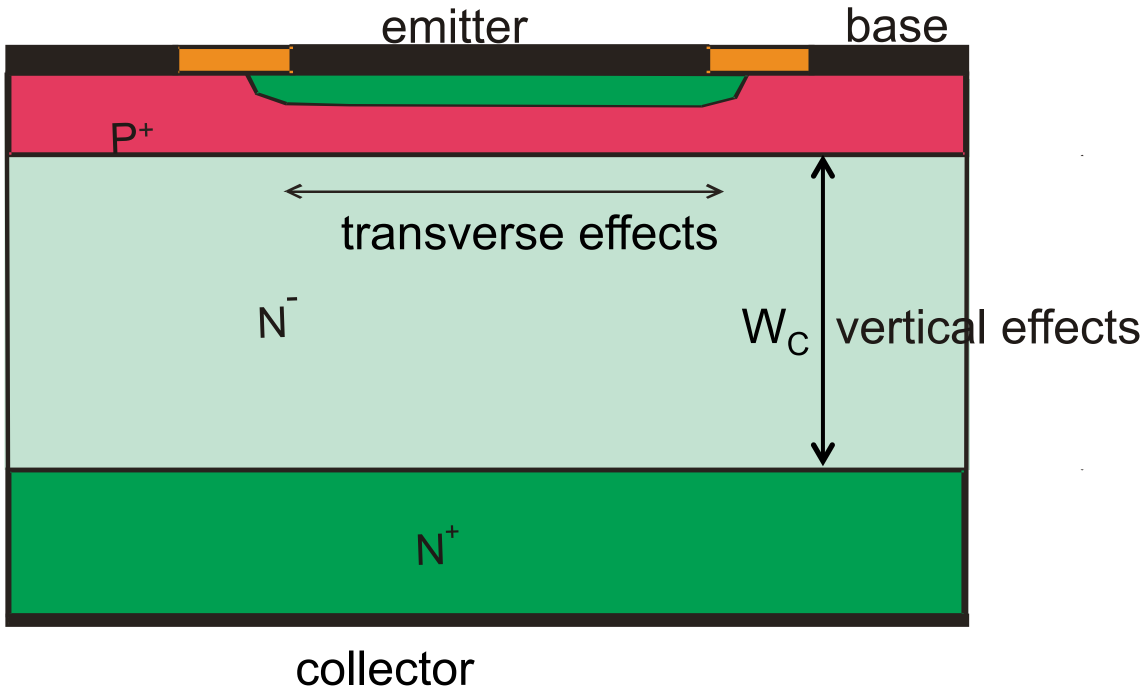
Power bipolar transistor

The Power Bipolar Transistor (BJT) differs from the signal BJT in these main points:

a) a collector region of large thickness and low doping to sustain an high reverse voltage (same as the power diode)
b) a thicker base region than the one of signal BJT, to avoid the base punch-through at high collector voltages
c) a larger doping for the base region to reduce the transverse voltage drop in the base
d) a large area to increase the current (this will lead to transverse effects as we will see)



A cross-section of a typical power transistor is indicated here.
The emitter contact and the two base contacts are located on the top surface, while the collector contact is made on the bottom surface.
The base/collector PN junction must sustain the reverse voltage: the lateral PN junction termination is not indicated here and is done as for the power PIN diode.

With reference to the one-dimensional model (fig. (a)), considering only the vertical doping profile between emitter, base and collector of the cross-section structure indicated before, the typical doping profiles for the three regions are indicate in the plot of fig. (b).
Typical values for the emitter, base and low doped epi collector widths for a BJT with voltage rating of several hundreds volts are: W_E 2 – 3 μm; W_B 2 – 4 μm; W_C 20 – 50 μm

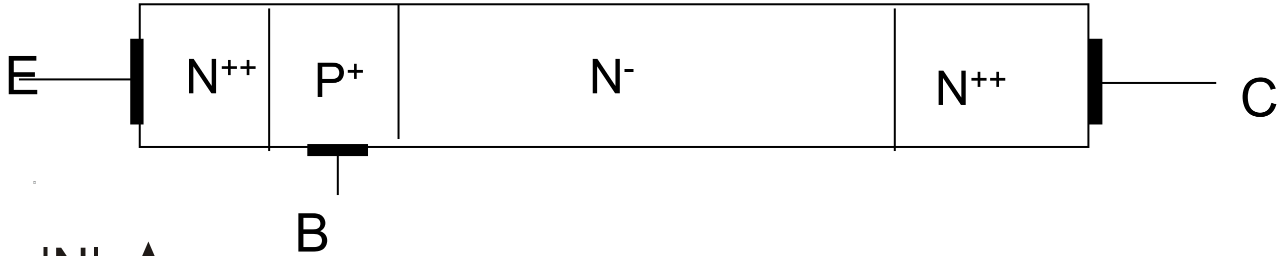


Fig. (a)

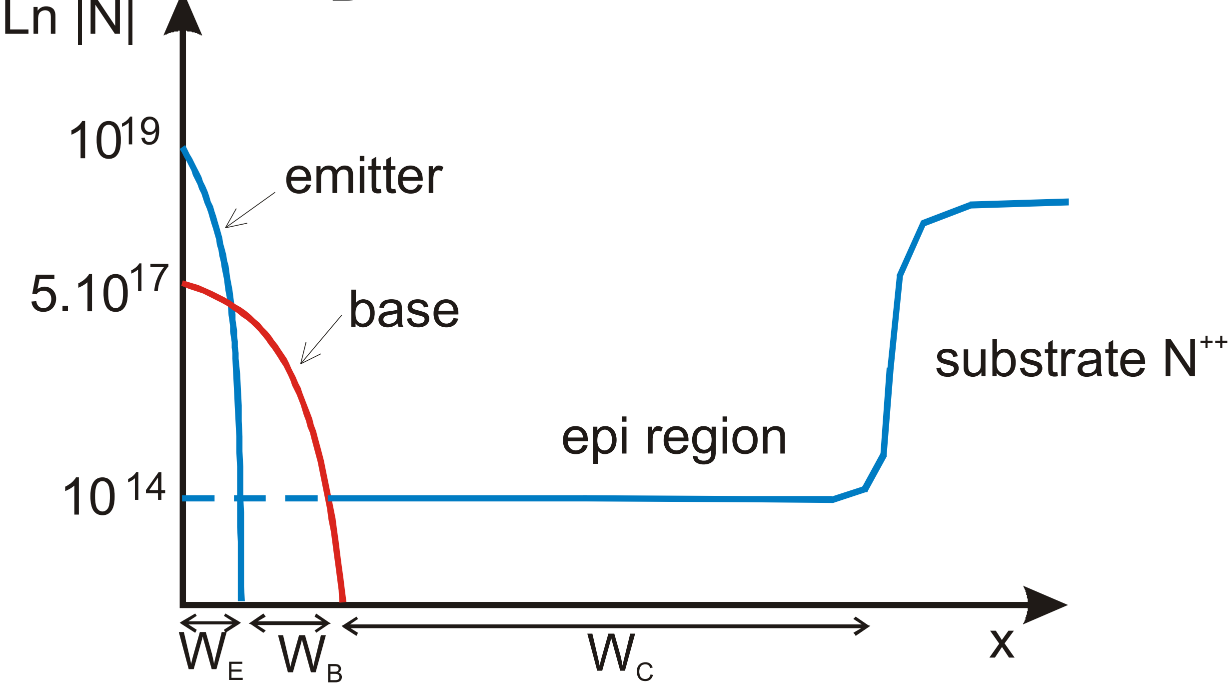


Fig. (b)

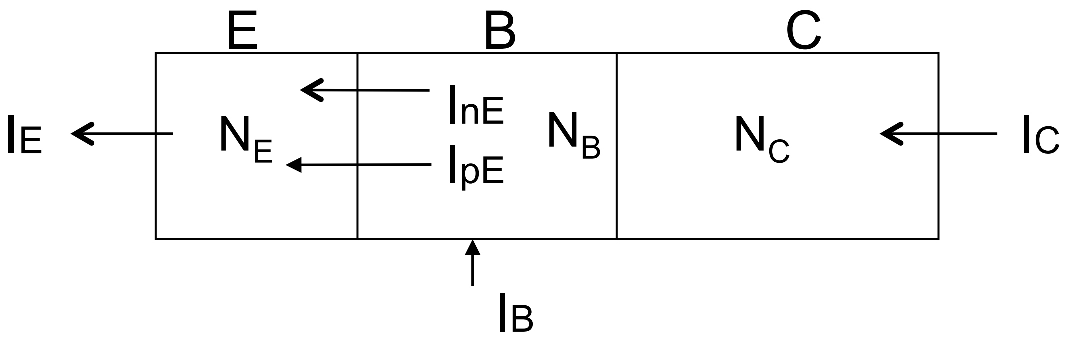
The power BJT has been the first active device to be used in power circuits as a controlled switch, because of its low ON voltage in saturation.
However it has two main drawbacks with respect to other power devices like MOS and IGBT:

a) it needs to be driven by a base current, so it should have an high current gain to reduce the amount of current needed to drive it in the ON state. We will see that while the BJT for signal level applications can have a large current gain, the power BJT for high voltages has a quite low current gain, that makes difficult (and costly) the driving stage required.
b) it presents a thermal instability that will limit the S.O.A at high voltages, and will make difficult the paralleling of more devices if this is needed to allow larger output currents.

In the following we will discuss the vertical effects due to the thick and low doped epi layer (required to withstand reverse voltage), and transverse effects due to the large area (to allow an high current) on the current gain. These constraints will affect both the d.c. and transient behavior of a power BJT.
We will also discuss the causes for thermal instability and the limitation in the max collector voltages in OFF state, to define the S.O.A of the device and its limitations.

Current gain in low injection

Let’s briefly recall the current gain dependence from the emitter and base doping in low injection level, with reference to the simple one-dimensional model in the approximation of constant emitter and base doping:



The current gain β is the ratio I_C/I_B in the active operating region:

$$\beta \cong \frac{I_C}{I_B} = \frac{\alpha}{1-\alpha}; \qquad \left[I_E = I_C + I_B;\ \alpha = \frac{I_C}{I_E} \right]$$

$$\alpha = \frac{I_{nE}}{I_E} \cdot \frac{I_{nC}}{I_{nE}} = \gamma \cdot \alpha_T$$

γ emitter injection factor

α_T base transport factor

Assuming negligible the base recombination (due to the small width $W_B << L_p$):

$$\alpha \cong \gamma; \qquad \beta \cong \frac{\gamma}{1-\gamma} = \frac{I_{nE}}{I_E - I_{nE}} = \boxed{\frac{I_{nE}}{I_{pE}}} \qquad (1)$$

From eq. (1) the current gain can be expressed by the ratio between the the electron current injected into the base $I_n(0^+)$ and the hole current entering into the emitter $I_p(0^-)$

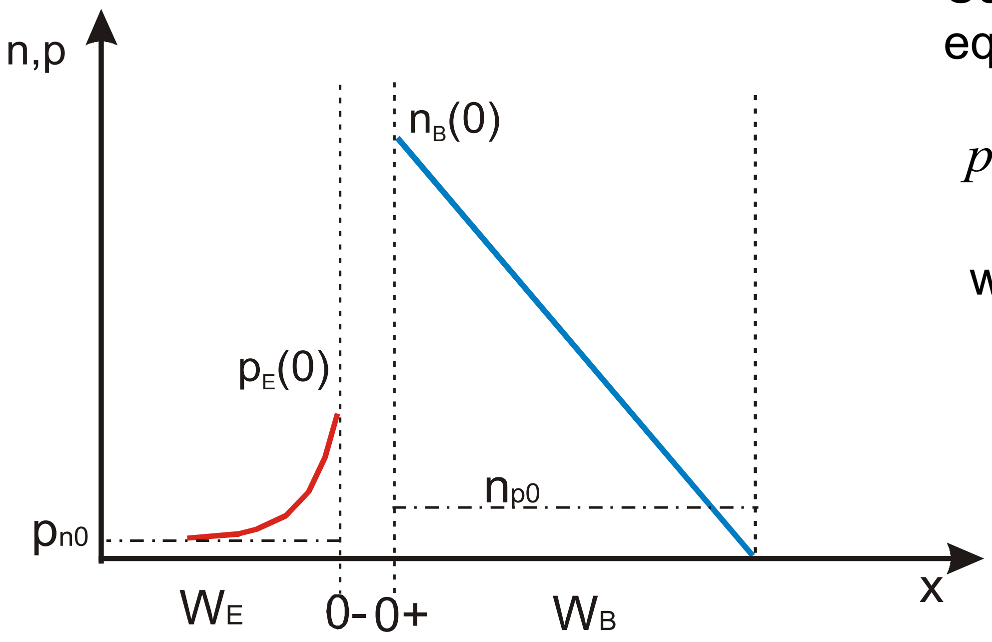
$$\beta \cong \frac{I_n(0^+)}{I_P(0^-)}$$

The two current components (diffusion currents) can be expressed as a function of the derivative of the carrier concentration at x = 0:

$$J_B \cong J_P(0^-) = qD_{PE}\left|\frac{dp}{dx}\right|_{x=0} \qquad J_C \cong J_n(0^+) = qD_{NB}\left|\frac{dn}{dx}\right|_{x=0} \quad (2)$$

As indicated in the plot, the p(x) distribution is exponential, while the n(x) is linear, assuming $\tau_P \ll \tau_n$ (no recombination in the base):

$$p_E(x) = p_E(0)\exp\left(\frac{-x}{L_P}\right) \qquad n_B(x) = n_B(0)\left(1-\frac{x}{W_B}\right) \quad (3)$$



Using expr. (3) in (2), and recalling the Boltzmann eq. for p(0) and n(0):

$$p_E(0) = p_{E0}\exp\left(\frac{V_{BE}}{V_T}\right); \qquad n_B(0) = n_{B0}\exp\left(\frac{V_{BE}}{V_T}\right)$$

we have for the two currents:

$$J_B \cong J_P(0) = qD_{PE}\frac{p_E(0)}{L_{PE}} = q\frac{D_{PE}}{L_{PE}}p_{E0}\exp\left(\frac{V_{BE}}{V_T}\right) \quad (4a)$$

$$J_C \cong J_n(0) = qD_{NB}\frac{n_B(0)}{W_B} = q\frac{D_{NB}}{W_B}n_{B0}\exp\left(\frac{V_{BE}}{V_T}\right) \quad (4b)$$

From eq. (4) we have for the current gain:

$$\boxed{\beta = \frac{D_{NB} n_{B0} L_{PE}}{D_{PE} p_{E0} W_B}} \quad (5)$$

To relate n_{B0} and p_{E0} to the base and emitter doping using the charge balance eq., we must take into account the effect of bandgap narrowing in the emitter region: in fact for high dopings (> 10^{19}) like in the emitter region the Si bandgap reduces, due to the very high number of donor levels just below the edge of conduction band. Then the value of the intrinsic concentration n_i (function of bandgap) will increases in the emitter above the usual Si value. So we must consider the charge balance eq. in equilibrium $pn = n_i^2$ with two different n_i values for the emitter and base regions:

$$n_{B0} \cdot p_{B0} = n_{iB}^2 \qquad \text{with } p_{B0} = \mathrm{N}_{\mathrm{AB}} \text{ base doping}$$

$$P_{E0} \cdot n_{E0} = n_{iE}^2 \qquad \text{with } n_{E0} = \mathrm{N}_{\mathrm{DE}} \text{ emitter doping}$$

where $n_{iE} > n_{iB}$ due to the bandgap narrowing in the emitter region

From the above relationship we obtain:

$$\frac{n_{B0}}{p_{E0}} = \frac{N_{DE}}{N_{AB}} \left(\frac{n_{iB}^2}{n_{iE}^2} \right)$$

where $n_{iB}/n_{iE} < 1$ is the bandgap narrowing factor, and by substituting in (5) we have :

$$\boxed{\beta = \frac{D_{NB} L_{PE} N_{DE}}{D_{PE} W_B N_{AB}} \left(\frac{n_{iB}^2}{n_{iE}^2} \right)} \quad (6)$$

Eq. (6) gives the current gain β as a function of emitter and base doping, and indicates that the current gain is larger if the ratio N_{DE}/N_{AB} is larger (as well known). It also indicates the β dependence from temperature, that is dependent from the bandgap narrowing factor that is dependent on temperature due to the exponential dependence on $-(E_G/kT)$ of n_i:

$$\frac{n_{iB}^2}{n_{iE}^2} = f(T) \propto \exp-\left(\frac{E_{GB} - E_{GE}}{kT}\right)$$

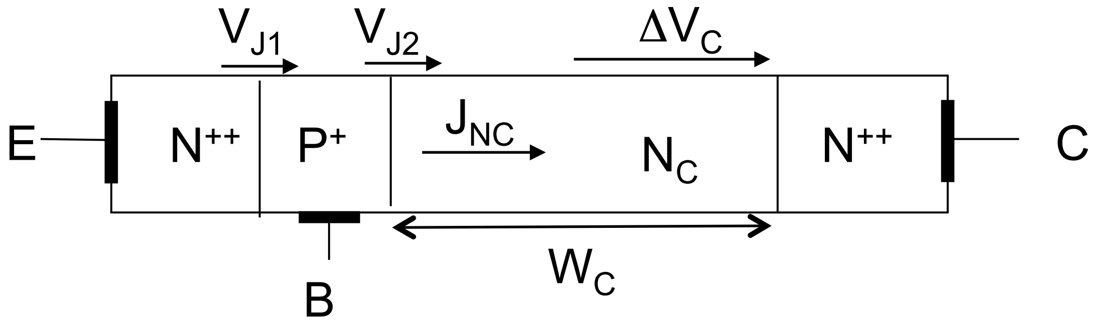
from the above relation we see that β increases with temperature; typically it doubles passing from 25°C to 100°C.

Eq. (6) also indicates that we can not increase the β value by increasing the emitter doping N_E above a value (about $2\ 10^{19}$) for which the bandgap narrowing start to be significant, so the emitter doping range is upper limited.
A deeper analysis of the emitter doping dependence of β will be done by a numerical analysis using PC1D simulator.

The other possibility to have an high value of β is to decrease the base doping, as suggested by eq. (6). This however will lead to a negative effect, meinly dependent on the tranverse effect in the base, as we will see in the following..

Vertical (collector) effects on current gain

Now we will examine the effects of the low doped and thick collector region on the current gain. These are mainly related to the voltage drop that develops on the neutral (not depleted) part of this region by the electron (majority carriers) current J_C flowing into it.
This voltage drop ΔV_C across the N_C region adds to the (reverse) voltage V_{J2} across the base/collector junction to balance the external voltage V_{CB} applied to the base and collector terminals.

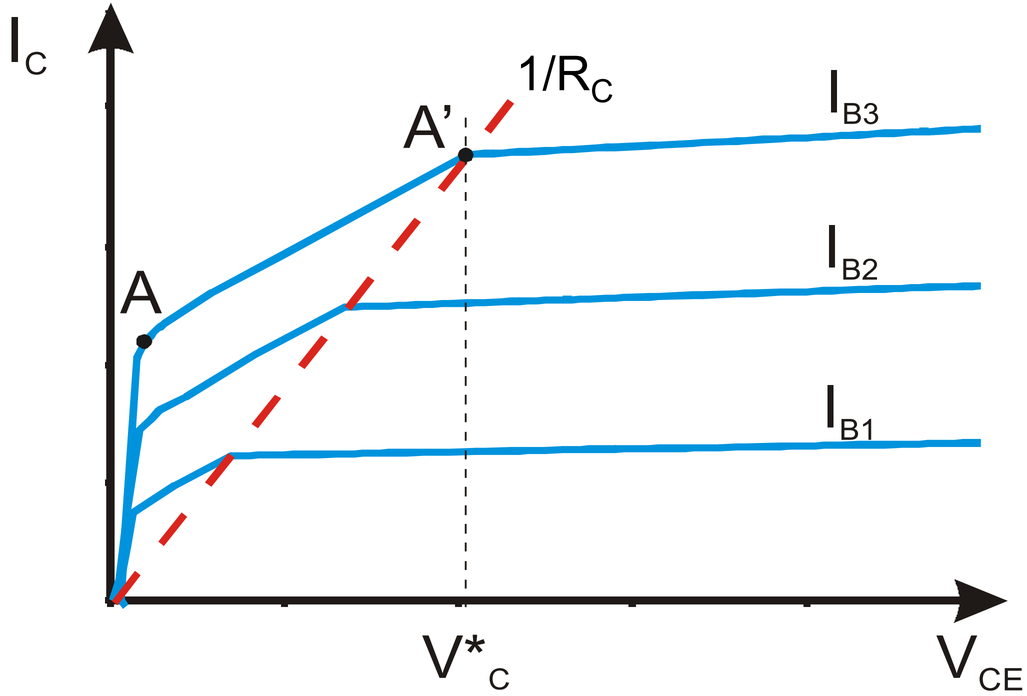


Indicating as W_C the thickness of the neutral part of the collector region (about equal to the epi layer thickness if the collector voltage is quite low, near to saturation voltage), we have for ΔV_C:

$$\Delta V_C = J_C \rho W_C = \frac{J_C W_C}{q \mu_n N_C}$$

When the BJT operates in the active region, with J_2 reverse biased, for a given external voltage V_{BC} the voltage drop ΔV_C actually reduces the reverse bias of the J_2 junction, and it increases with the collector current J_C becoming as large as the external voltage V_{CB} ($\cong V_{CE}$) for a given J_C*.

In this latter case the junction J_2 ceases to be reverse biased. Then the external voltage $V_{CB} \cong V_{CE}$ is almost completely balanced by voltage drop ΔV_C^* (neglecting V_{J1} and V_{J2} of the two junctions in forward bias). Then the J_C^*-V_{CE}^* locus in the output I-V curves plane is given by:



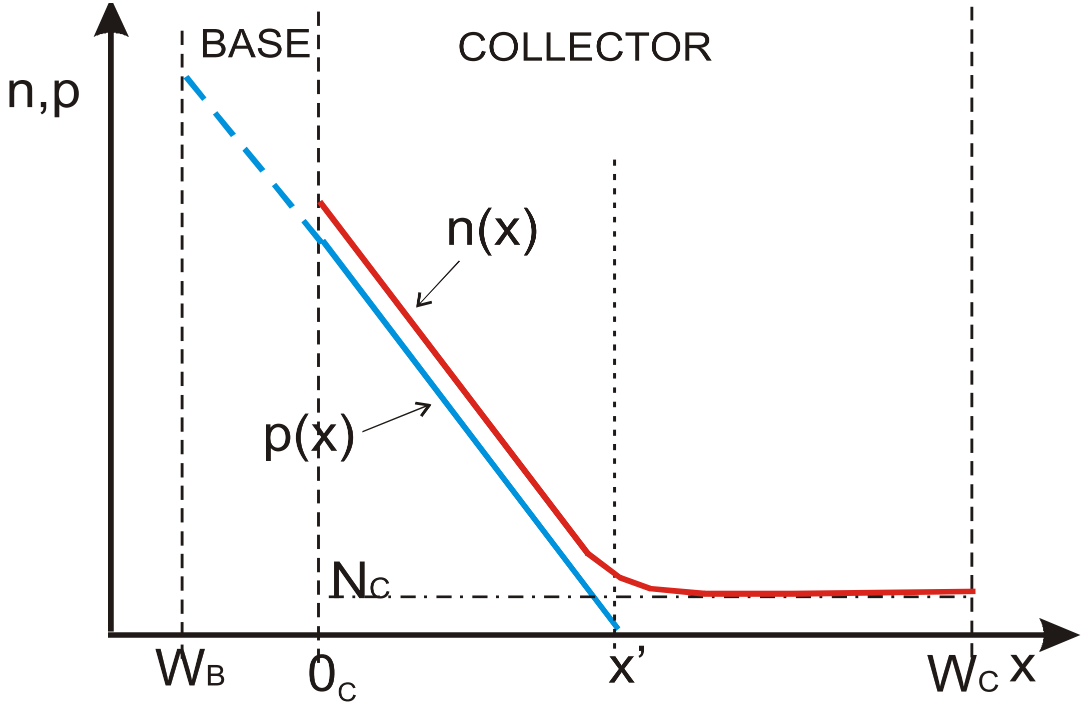
$$V_{CE}^* \cong \Delta V_C^* = \frac{J_C^* W_C}{q \mu_n N_C} = R_C J_C^*$$

and it is represented in the output I-V plane as a straight line of slope equal to the ohmic resistance of the low doped region.

The value of V_C^* can be quite high respect to the saturation voltage $V_{SAT} = V_{J1} - V_{J2}$. For example, assuming: $W_C = 50\ \mu m$, $N_C = 2\ 10^{14}\ cm^{-3}$ $R_C = 0.125\ \Omega cm^2$ for $J_C = 100\ A/cm^2$ we have $V^*_C = 12.5$ V

If the external voltage V_{CB} is less than V_C^* (point A' in the plot for a given I_{B3} current) , the junction J_2 becomes forward biased and it will inject minority carriers (holes) into the collector region.
The region AA' comprised between the saturation region and the 1/R_C line is named:
Quasi-saturation region

In that region, when the holes are injected into the low doped collector, we can have an high injection regime into the collector : p ≈ n >> N_D, in some part (0 – x') of the collector region, that will largely reduce the current gain (as seen in points A – A' of the I-V plot).



For the charge balance in high injection we have:

$$p(0_C) + N_C = n(0_C) \quad \text{with} \quad p, n > N_C$$

$$\text{then} \quad p(x) \cong n(x); \qquad \frac{dp}{dx} \cong \frac{dn}{dx}$$

The width of the 0-x' region in high injection depends on the value of the holes injected from the base and hence from the forward bias of J_2. The voltage drop is now confined on the region x' – W_C that has only majority carriers.

As far the external voltage V_{CB} decreases with respect to the V_C* value, J_2 becomes more forward biased, and the values of $p(0_C)$ and $n(0_C)$ increase. As a consequence, the abscissa x' extends into the collector and eventually it reaches the limit W_C of the low doped epi region (point A in the IV curves). In this case all the epi region is in conductivity modulation and the voltage drop is reduced to the one of a BJT in saturation (less than 1 V).

Let’s evaluate how the high injection in the collector will affect the current gain in point A. From the previous comments on the carrier distributions in the collector region we have for the electron and hole currents:

$$J_{nC} = q\mu_n nE + qD_n \frac{dn}{dx} \cong J_C \quad \text{(13a)}$$

$$J_{pC} = q\mu_p pE - qD_P \frac{dp}{dx} \cong 0 \quad \text{(13b)}$$

where J_{nC} is the collector current and $J_{pC} = 0$ because there is non hole flow into the collector, meaning that for J_{pC} the drift component due to field E is balancing the diffusion one due to hole gradient. From (13b) we extract the field E:

$$E = \frac{D_P}{\mu_p p}\frac{dp}{dx} = \frac{V_T}{p}\frac{dp}{dx}$$

and substituting in (13a) we have:

$$J_{nC} = q\mu_n V_T \frac{n}{p}\frac{dn}{dx} + qD_n \frac{dn}{dx} = q2D_n \frac{dn}{dx} = J_C$$

J_C must be constant along x, then the slope *dn/dx* is constant and we have:

$$\boxed{J_C = 2qD_n \frac{n(0_C)}{x'}} \qquad \text{(14)}$$

From (14) we note that J_C decreases if x’ increases, and it gets the smallest value for x’ = W_C

When x’ = W_C all the collector region is in high injection and the voltage drop is very low. This is the true saturation point A in the IV curve, and, from (14), J_C in point can be expressed as:

$$\boxed{J_C(A) = 2qD_n \frac{n(0_C)}{W_C} = 2qD_n \frac{n(0_B)}{W_B + W_C}} \qquad (15)$$

We need to stay in point A to operate the BJT as a closed switch with low voltage drop. But to bring the BJT in point A we need a quite large base current, because the current gain β_{HC} in point A can be much lower than the one β_0 in the active region (A’). The current gain β_{HC} in true saturation can be obtained by using (15) in place of (4b) for the ratio J_C/J_B:

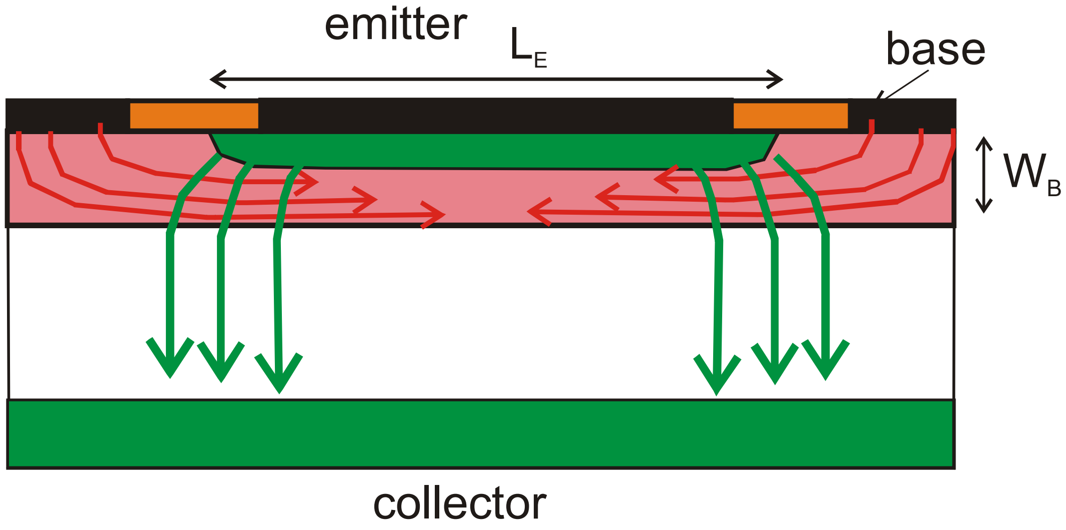
$$\boxed{\beta_{HC} = \frac{J_C(A)}{J_B} = \frac{2W_B}{W_C + W_B} \beta_{.0} < \beta_0} \qquad (16)$$

The decrease of β_{HC} is a drawback for power BJT of high V_{MAX} voltage, that need a thick and low doped collector region to sustain large collector voltages.

A quantitative evaluation of the decrease of β due to the collector high injection will be done by a numerical analysis using PC1D simulator for typical BJT structures for high voltages.

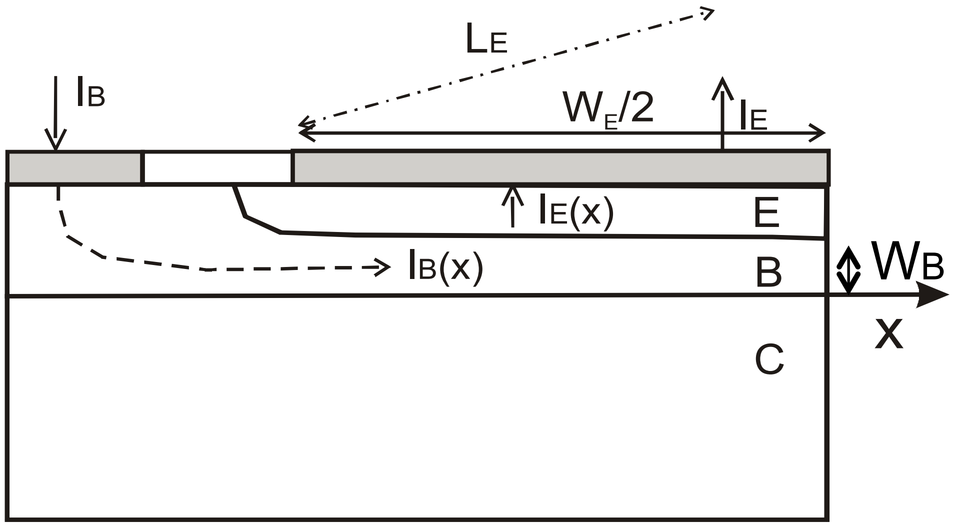
Transverse (base) effects on current gain

The large area required to handle high currents creates a transverse effect in the base.
At large collector (and base) currents, the base current that flows laterally in the thin base layer causes a voltage drop along it, that reduces the V_{BE} voltage across the J_1 junction as we proceed far from the base contact.
As a consequence, the side regions of the emitter junction will have a V_{BE} forward bias larger than the central ones, and the emitter (and collector) current will flow mainly in the lateral regions, as schematically indicated in figure, leaving not fully effective the central region of the emitter.



This effect is called “emitter crowding” because the emitter current is “crowded” in the lateral side of the emitter area, leaving the remaining part quite ineffective. The result is a decrease of the current gain at large I_C, because the “average” emitter current is lower than the one evaluated with one-dimensional analysis.

With reference to the cross section in figure (half emitter width is considered for symmetry), we can evaluate the emitter current distribution along the x axis, due to the "emitter crowding" effect, that makes both the emitter and base currents not uniform along he x axis:



Assuming an emitter length L_E in the z direction, we can express the emitter current $dI_E(x)$ along the x axis as:

$$dI_E(x) = J_E L_E dx$$

In the active region, $dI_B(x)$ can be related to $dI_E(x)$ as:

$$dI_B(x) = \frac{1}{\beta + 1} J_E L_E dx \quad (17)$$

Assuming an (average) conductivity σ_B in high injection for the base layer, the voltage drop $dV_{BE}(x)$ due to the base current flow can be expressed as a function of the base resistance as:

$$dV_{BE}(x) = I_B(x) \frac{dx}{\sigma_B W_B L_E} \quad (18)$$

Assuming high injection in the base, we can express the emitter current density along the x axis as a function of the actual $V_{BE}(x)$:

$$J_E(x) = J_0 \exp\left(\frac{V_{BE}(x)}{2V_T}\right)$$ making the derivative: $$\frac{dJ_E(x)}{dx} = \frac{J_E}{2V_T}\frac{dV_{BE}(x)}{dx} \qquad (19)$$

and, by substituting (18) in (19) we have: $$\frac{dJ_E(x)}{dx} = \frac{J_E(x)}{2V_T}\frac{I_B(x)}{\sigma_B W_B L_E}$$

In high injection we have a conductivity modulation of the base, and, recalling the analysis for the base, eq. (12) we can relate the conductivity σ_B at high injection to the one σ_{B0} due to the doping (low injection) proportional to the increase of the emitter (or collector) current with respect to the critical current J_R where high injection start to develop:

$$\sigma_B = \sigma_{B0}\frac{J_E(x)}{J_R}$$ and we have: $$\frac{dJ_E(x)}{dx} = \frac{J_R}{2V_T}\frac{I_B(x)}{\sigma_{B0} W_B L_E} \qquad (20)$$

Making the derivative: $$\frac{d^2 J_E(x)}{dx^2} = \frac{J_R}{2V_T}\frac{1}{\sigma_{B0} W_B L_E}\frac{dI_B(x)}{dx} \qquad (21)$$

And finally, by substituting the relation (18) that links $dI_B(x)$ to $dI_E(x)$ we have the following differential eq. for $dI_E(x)$:

$$\frac{d^2 J_E(x)}{dx^2} - \frac{J_R}{2V_T}\frac{J_E(x)}{\sigma_{B0} W_B(\beta+1)} = 0$$

this is a 2° order diff. eq. of general type:

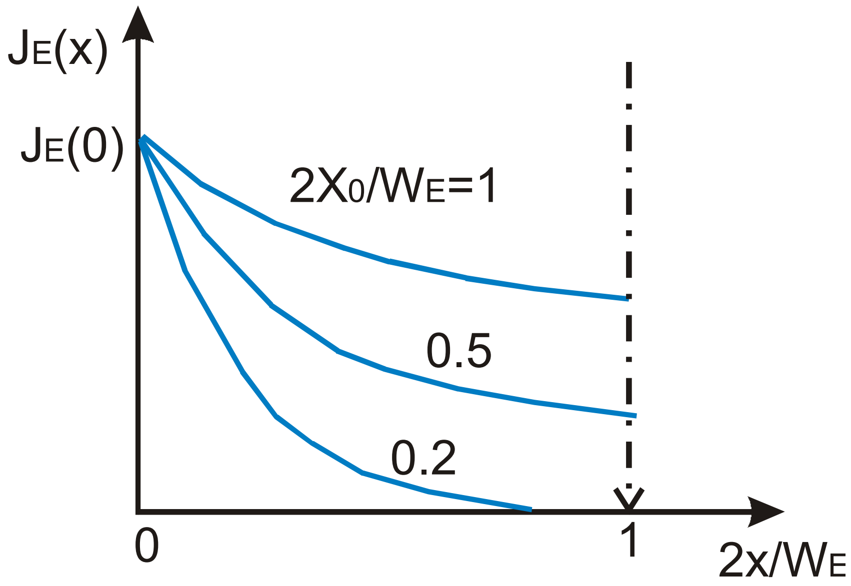
$$\frac{d^2 y}{dx^2} - \frac{y}{X_0^2} = 0$$

and the solution is :

$$\boxed{J_E(x) = J_{E0}\exp\left(-\frac{x}{X_0}\right)} \quad (22)$$

with: $X_0 = \sqrt{\dfrac{2V_T W_B \sigma_{B0}(\beta+1)}{J_R}}$

where J_{E0} is the emitter current density at the edge (x=0) of the emitter area, and X_0 is an “effective” emitter length.



The meaning of X_0 can be understood if one consider that the integral of (22), assuming $J_E(x) \to 0$ for $x<W_E/2$ is:

$$\boxed{\int_0^\infty J_E(x)dx = J_{E0}\cdot X_0}$$

Then X_0 represents the effective half-emitter width (instead of $W_E/2$) if we consider an “ideal” constant current density J_{E0} (with no ΔV_{BE} voltage drop)

A more detailed analysis of the transverse base effects can be done by means of numerical simulation with 2D PISCES II software.

Some results are reported here for the BJT structure indicated in fig. (a) with doping profiles reported in plot (b)

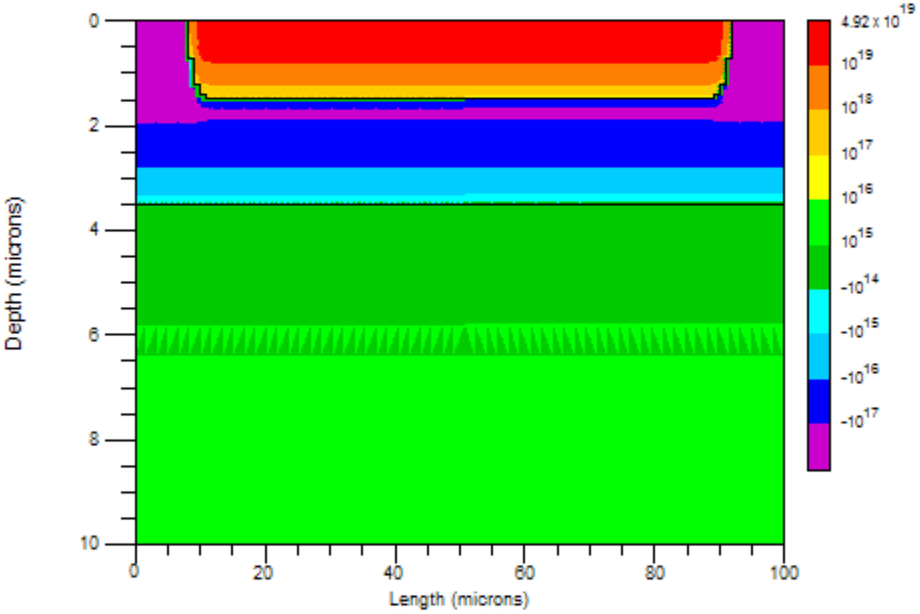
The structure has:

emitter width W_E=90 μm; base depth W_B=2 μm, two base contacts

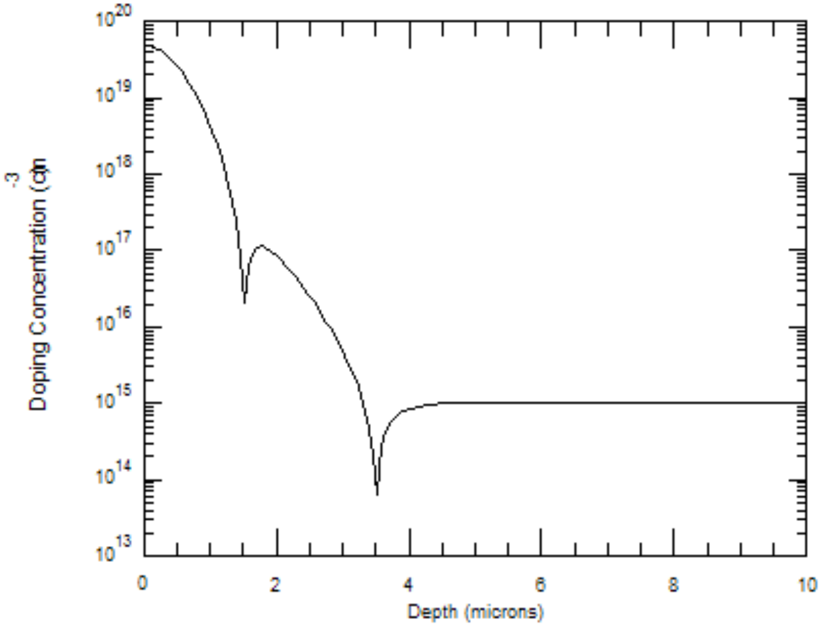
emitter doping: gaussian, peak value 5 10^{19}

base doping: gaussian, peak value 8 10^{17} – sheet resistance $500\Omega/\square$

collector doping: uniform 10^{15}

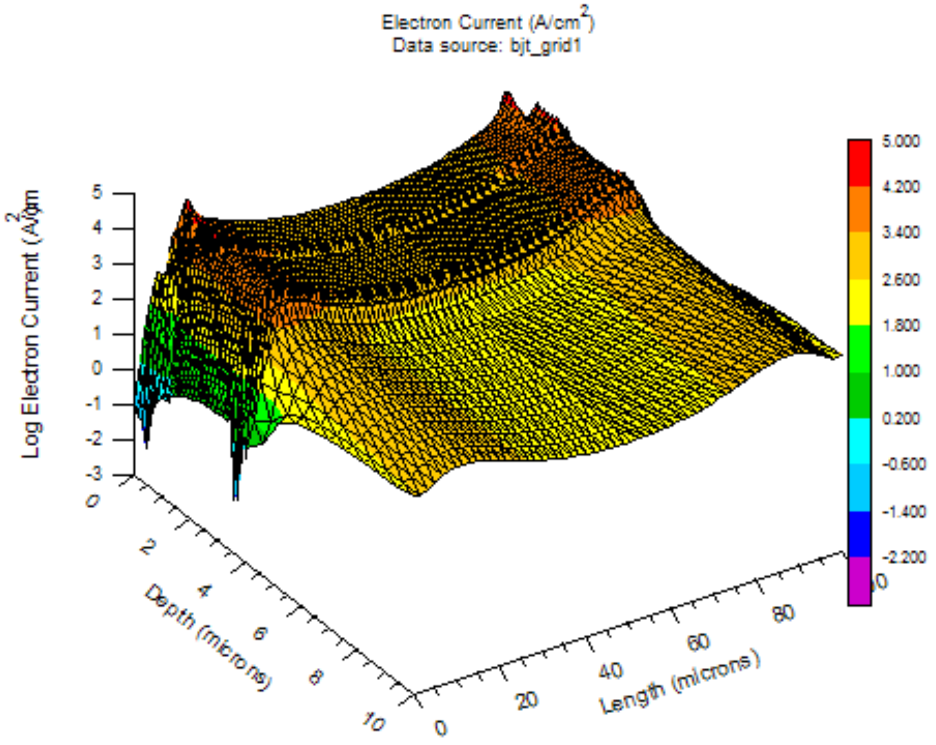


(a)

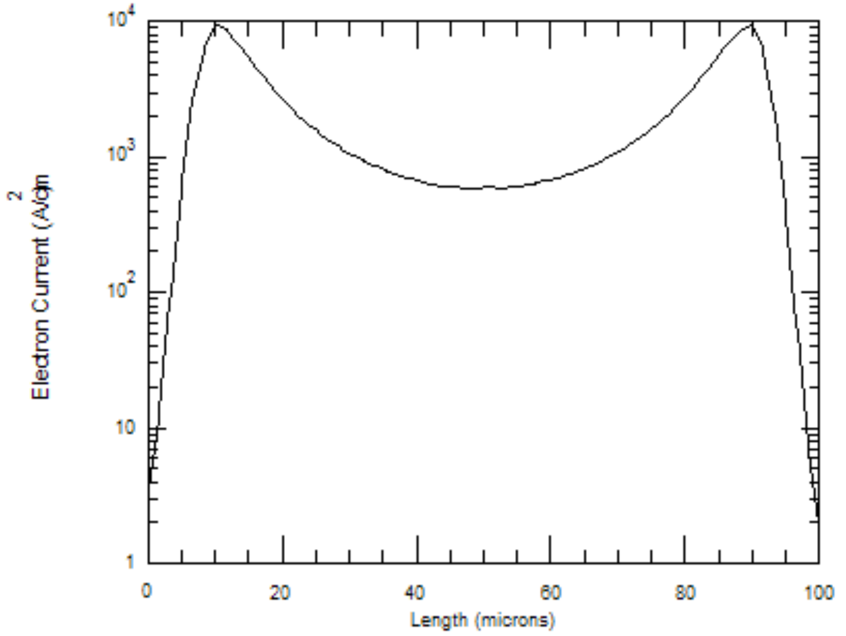


(b)

The 2D distribution of the emitter current density J_E (x,y) (in log scale) is reported in fig. (a) for a V_{BE} voltage of 0.85 V (high injection case), while in fig. (b) there is plotted the $J_E(x^*)$ lateral profile for a depth $x^*=2$ μm (again in log scale).
From the plot (b) we can see that J_E reduces of more than one order of magnitude in about 20 μm from each of the lateral base contacts: the central part of the emitter (20 – 80 μm) is quite ineffective in carrying the output current. For this bias the current gain reduces to less than ½ the value of current gain in low injection.



(a)

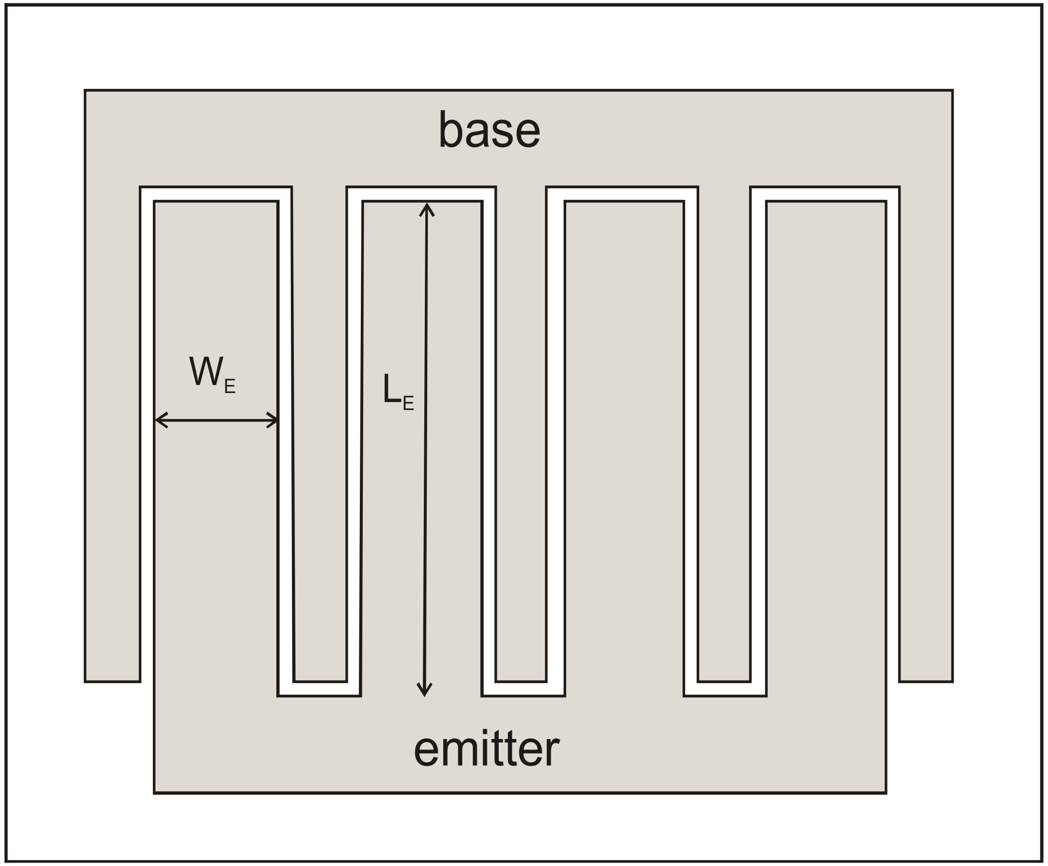


(b)

The result of the above analysis is that at high currents, the “electrically active” emitter area is reduced to $2X_0L_E$ instead of the “geometrical” one W_EL_E.

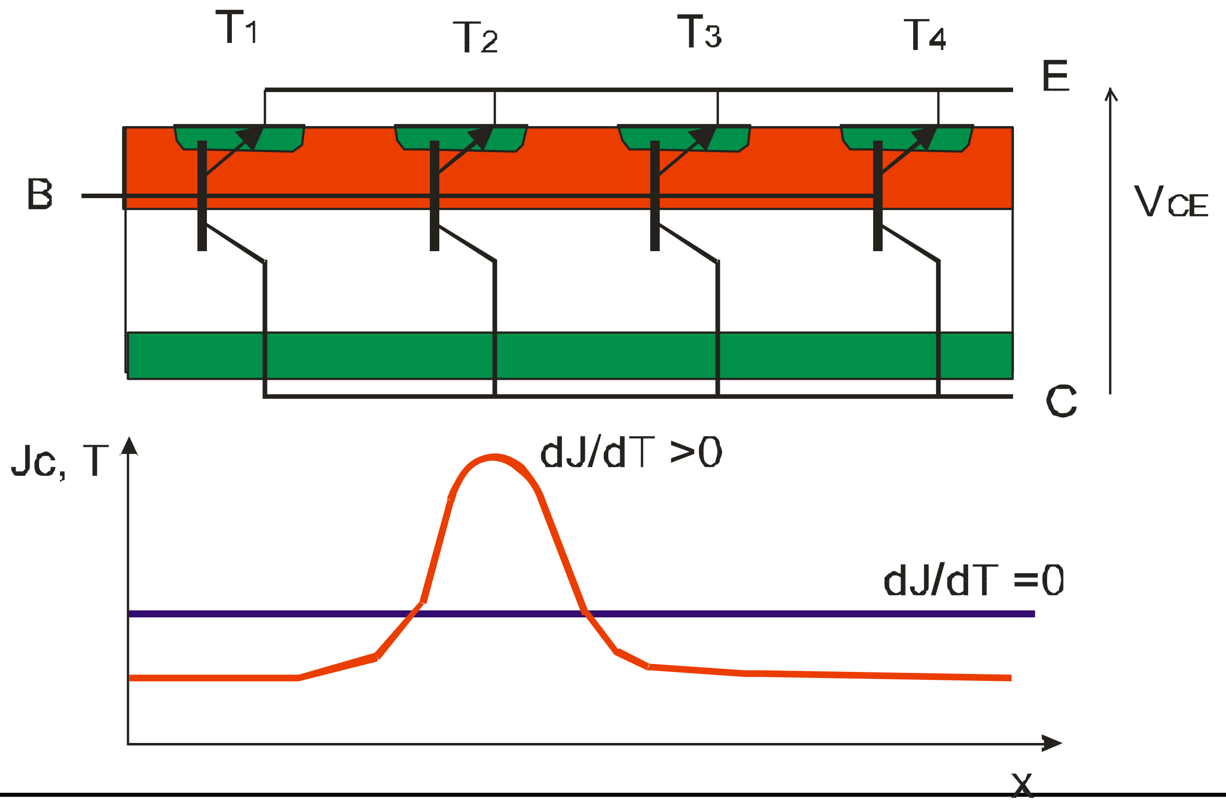
To increase the total BJT emitter area while keeping $W_E < X_0$ for each emitter, we need to use “interdigitated” or “comb” lay-out for the emitter, as indicated in the bottom figure.

A further decrease of the emitter area is obtained using cellular topology for BJT, where the BJT is made of many small area BJT connected in parallel by the emitter and base contacts.



Thermal instability in power BJT

The interdigitated or cellular structure of BJT makes it more subject to thermal instability.
In fact for these structures the BJT is actually made of many elementary BJT connected in parallel through the metal of the emitter and collector contacts.
Now, recalling hat the current gain β has a positive temperature coefficient, if for same reason one of the elementary BJT's (T_2) is warmer than the others, the collector current I_C will increase for a given I_B, and that in turn increases the power of that BJT, because the collector voltage V_{CE} is equal for all BJT in parallel. This effect tend to concentrate the current in a small area, and possibly to generate an very high overheating in that area, called "hot spot"



This electro-thermal feedback can give rise to an uneven temperature distribution on the chip area, but not necessarily to an “hot spot” that will genarate a failure of the device. The latter develops if a condition of “thermal instability” is set.

A thermal instability holds if the increase of the thermal power P generated in a device for an increase of temperature T is larger than the increase of the thermal power P_T dissipated, as:

$$\frac{dP_D}{dT} \geq \frac{dP_T}{dT} \qquad (23)$$

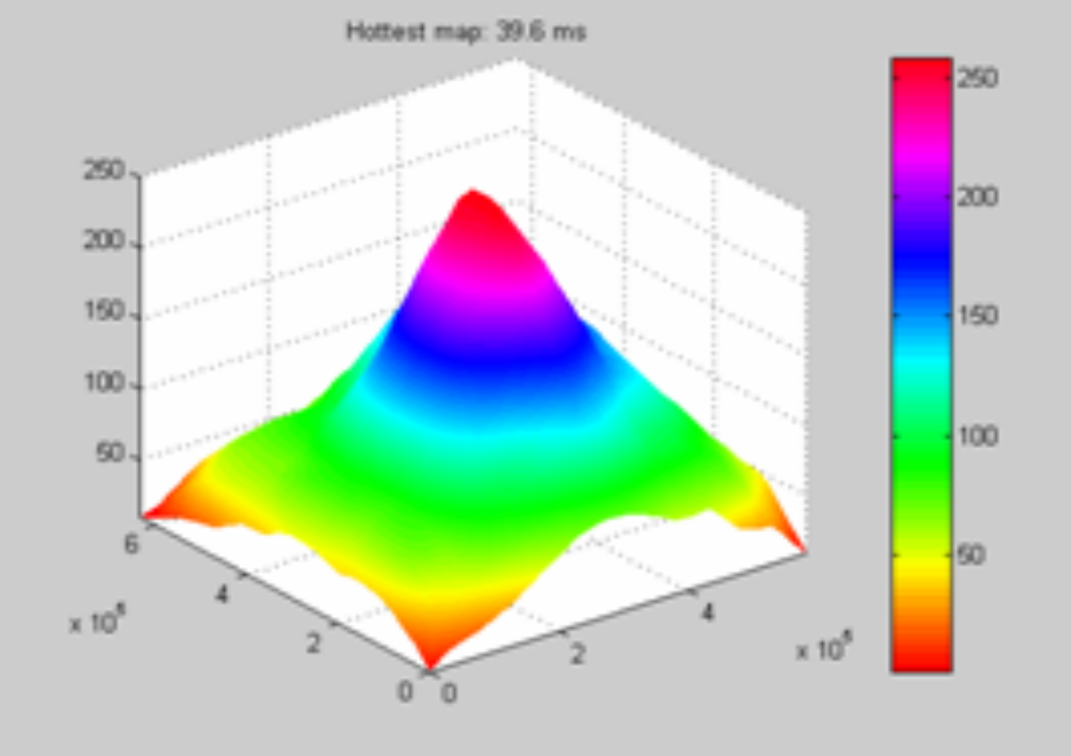
assuming the (electrical) power $P_D = V_{CE} \cdot I_C$, the thermal power $P_T = \Delta T / R_T$, and substituting in (23), we have, for the thermal instability condition (V_{CE} is kept constant by the bias):

$$V_{CE}\frac{dI_C}{dT} \geq \frac{1}{R_T(t)}$$

The derivative dI_C/dT is defined as the temperature coefficient $\alpha(I)$ of the collector current; if α is negative the device is thermally stable, while for α positive the instability condition is:

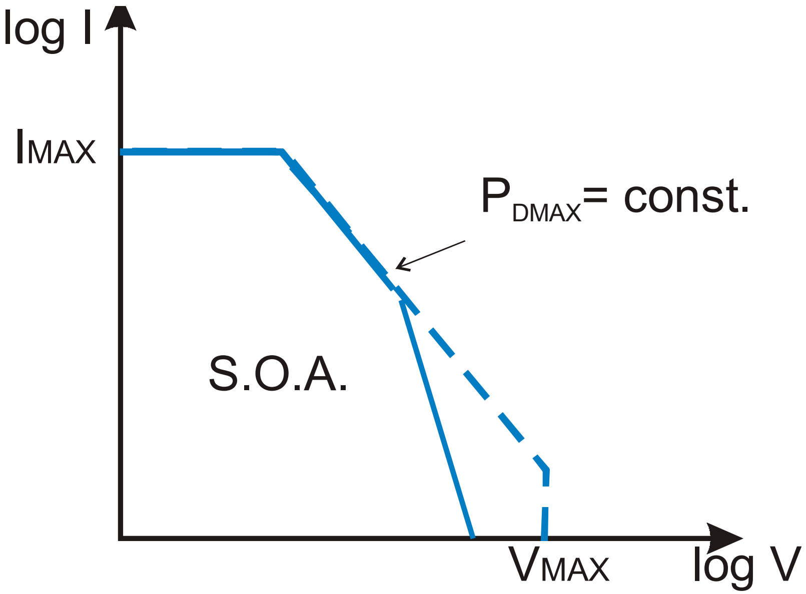
$$\boxed{\frac{dI_C}{dT} \equiv \alpha(I) \geq \frac{1}{V_{CE}R_T(t)}} \qquad (24)$$

We remember that $\alpha(I_C)$ is positive for BJT, so the thermal instability can be avoided with a low thermal resistance R_T. For a given R_T from /24) we see that the onset of instability is increased if V_{CE} increases so it is needed to decrease the electrical power P_D =VI at increasing V_{CE} with respect to the condition P_D = const.



A detailed analysis of temperature distribution on the chip surface of a BJT can be done with numerical 2D simulator using electro-thermal analysis. Here it is plotted the result of such analysis for a cellular BJT biased at a power less than the max power P_{DMAX}.

An hot spot with a T_{MAX} of above 250°C develops in a quite small area respect to the chip size, that eventually leads to a device failure.



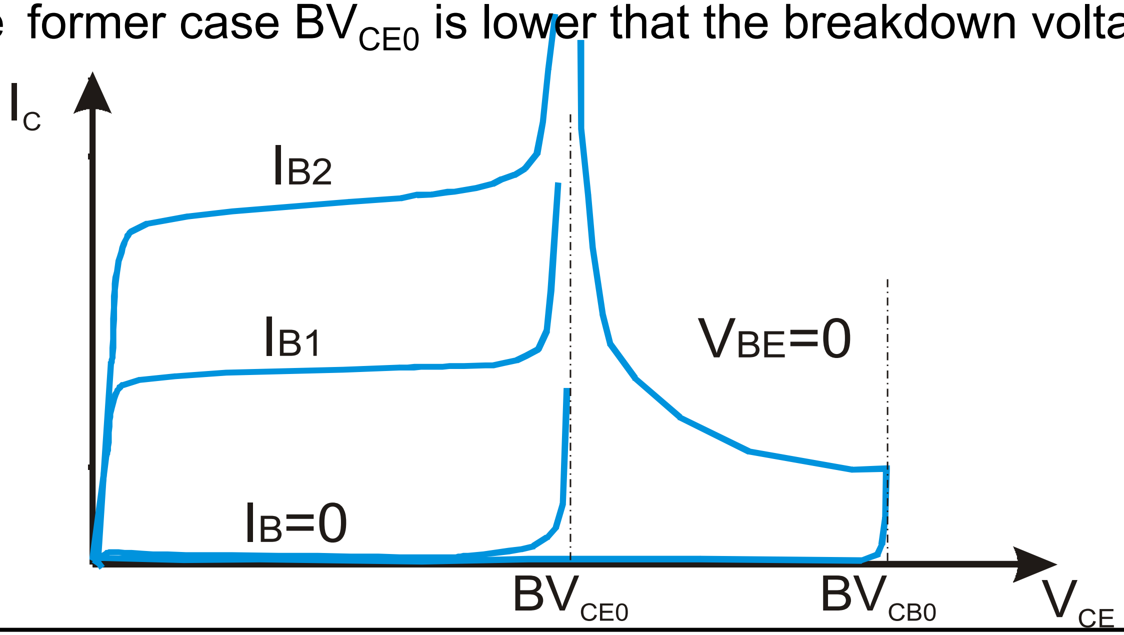
The thermal instability can be avoided by decreasing the collector voltage with respect to the one for uniform power dissipation, in order to reduce the P_D with respect to the P_{DMAX} as indicated in the S.O.A plotted here.

Max voltage ratings of power BJT

The collector voltage of a BJT can be increased up to a max value, indicated as BV. However this max voltage is usually lower than the breakdown voltage of the B/C junction and it is depending from the internal resistance of the drive circuit used. If the drive circuit is considered a a current generator (high output resistance) the max voltage, named BV_{CE0} is much less than the max voltage BV_{CB0} reached when the drive circuit can be considered a voltage generator (low output resistance).

The output IV curves of a power BJT for the two cases are sketched in this plot. All the output curves obtained by the (usual) base current drive (current generator) tend to infinite at a voltage BV_{CE0} lower than the voltage BV_{CB0} reached with a base voltage drive (here only the V_B =0 case is indicated).

In the latter case the BV_{CB0} value corresponds to the breakdown voltage of the base/collector junction, while in the former case BV_{CE0} is lower that the breakdown voltage, as we will see.



a) analysis of BV$_{CE0}$ (with a current drive on base)

The collector current I$_C$ in the active region can be expressed as: $I_C = \alpha I_E + I_{CB0}$

Considering the case of a base current I$_B$ =0 (limit case for the BJT off state), we have:

$$I_{C0} \equiv I_{E0} = \alpha I_{E0} + I_{CB0} \quad \Rightarrow \quad I_{E0} = \frac{I_{CB0}}{1-\alpha} \quad (25)$$

If we consider avalanche multiplication in the collector region (as done for the PIN diode) for high voltages, we can consider both the leakage currents I$_{E0}$ and I$_{CB0}$ being multiplied by M, and (24) becomes:

$$I_{E0}(V_{CE}) = \frac{M \cdot I_{CB0}}{1-\alpha \cdot M} \quad (26) \qquad \text{with } M = \frac{1}{1-\left(\frac{V_{CE}}{V_{BR}}\right)^n} \quad (27)$$

Now from (26) I$_{E0}$ →∞ if $\alpha M \rightarrow 1$ and then the value of BV$_{CE0}$ is given by: M(BV$_{CE0}$) = 1/α (28)

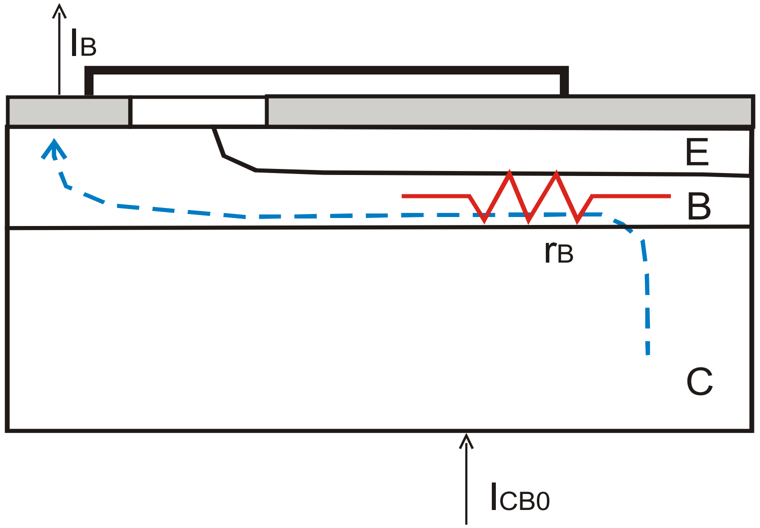
Using (27) in (28) we have: $\left(\frac{BV_{CE0}}{V_{BR}}\right)^n = 1-\alpha$

Depending on n factor BV$_{CE0}$ in actual cases is about 0.5 – 0.7 BV$_{CB0}$

b) analysis of BV_{CB0} (with a voltage drive on base)

In the case of a short-circuit between base and emitter terminals, the leakage current I_{CB0} flows into the base layer and exit from the base terminal, because the E/B junction is biased at zero voltage and it can not allow any current flow. As a consequence there is no injection of electrons from the emitter (α is zero) and at high voltages this current is merely increased by the multiplication factor M:

$$I_{C0}(V_{CE}) = MI_{CB0} = \frac{I_{CB0}}{\left(1 - \frac{V_{CE}}{V_{BR}}\right)^n} \quad (29)$$



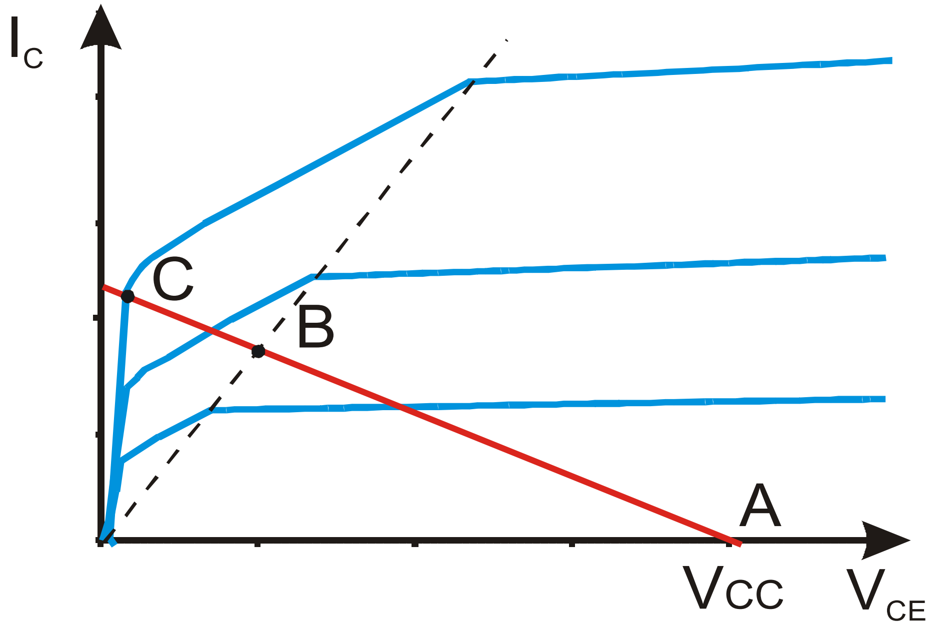
From (29) we see that the current I_{C0} will increase largely only at at the breakdown voltage V_{BR} of the B/C junction, and $BV_{CB0} = V_{CR}$.
However, as indicated in the figure, we must take into account the distributed intrinsic resistance r_b of the base layer: when the current MI_{CB0} (that increases largely at BV_{CB0}) will cause a voltage drop on r_B of the order of 0.6 V, the central part of the emitter area becomes forward biased and will inject electrons into the base. Then we resume the condition of eq. (26), and the current branch will present an asymptote at $V_{CE} = BV_{CE0}$ as in the previous case.

Transient behavior of power BJT

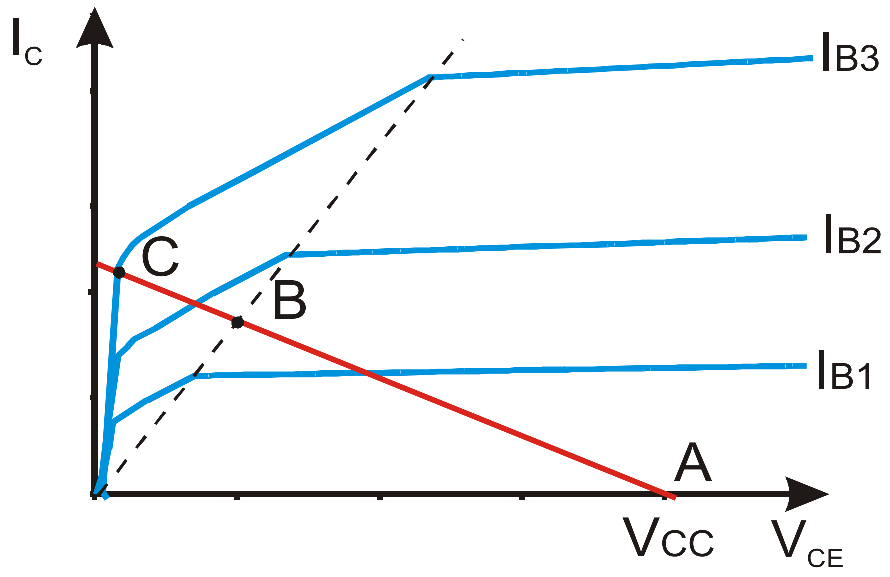
The dynamic behavior (as the steady state) of the power BJT is largely dependent on the presence of the quasi-saturation region that develops at large currents.

Assuming a resistive load as indicated in figure, when the operating point lies in the quasi-saturation (points B – C), the base region extends in the collector layer, and the charge stored in the “electrical” base are much larger than in the active region (points A – B). As a consequence ,both the charge injection (in turn-on) and charge removal (in turn-off) are much slower than in the active region.

We will now discuss both the turn-on and turn-off transients with initial reference to a resistive load.



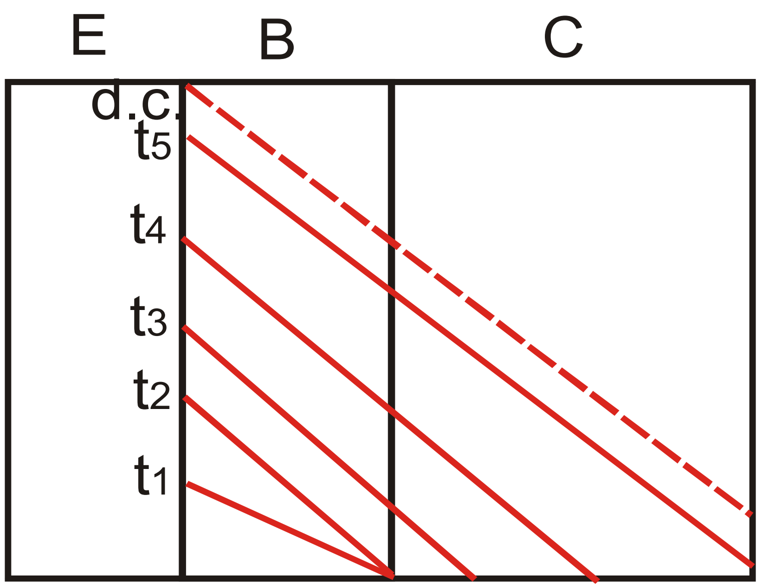
a) Turn-on



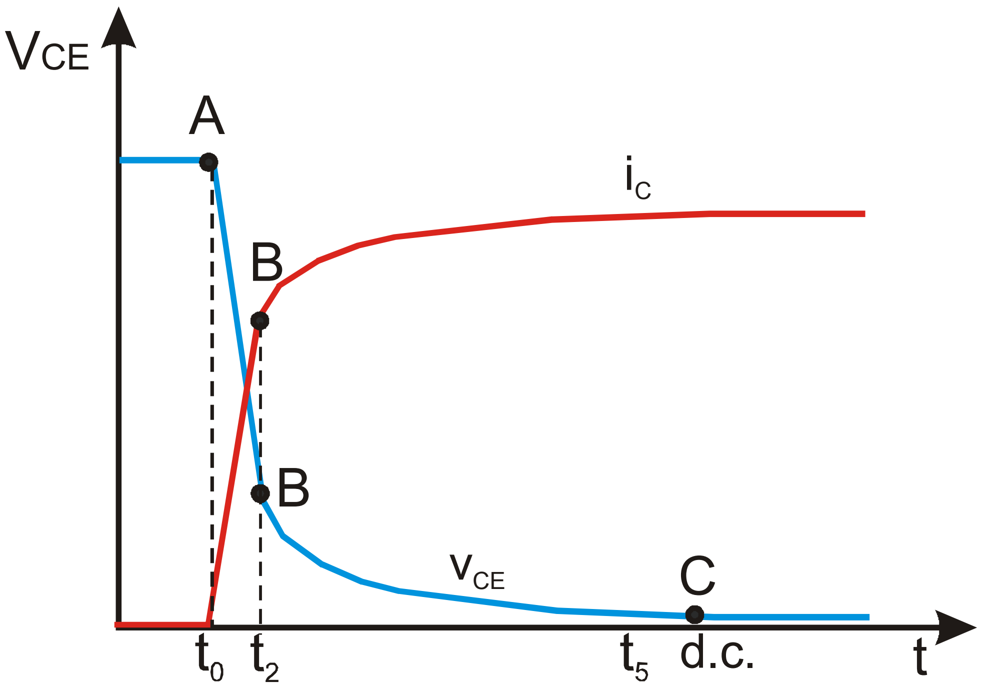
To generate a turn-on, a base drive of a current step from 0 to I_{B3} is assumed to force the BJT in saturation with the assigned load line (remember that we must reach point C of the output IV curves to have a low voltage drop across the BJT when in its conducting stage (switch closed)).

The collector current increases rapidly in passing from point A to point B, and more slowly in passing from point B to point C.

This is due to the increased charge that must be present in a part of the collector region when in quasi-saturation regime; the dynamic is as slow as far the point x' (eq. (14)) of the high injection layer extends into the collector layer.



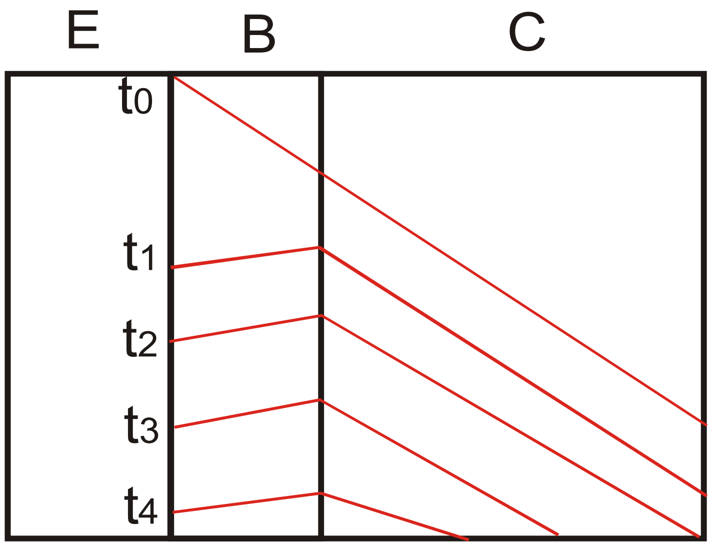
In the top figure the charge distribution during the turn-on is sketched. From time t_0 (point A) to t_2 (point B) the "electrical" base is confined in the "metallurgical" base. From times t_3 to t_5 the base extends into the collector, eventually filling all the low doped layer (point C) if the base current $I_B > I_{B3}$

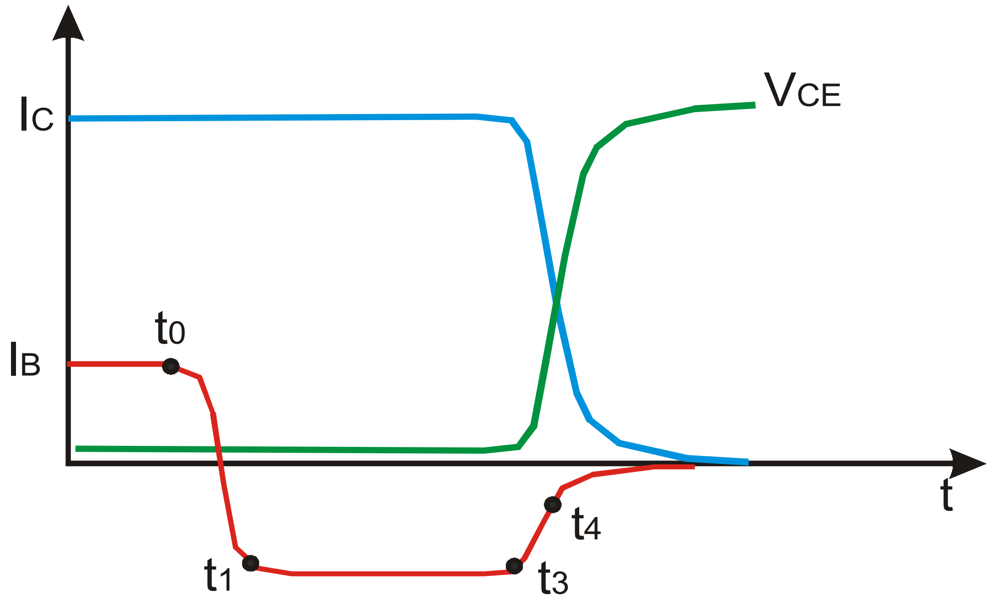


As a result, we have a fast transient from point A to B, while the one from B to C is slowed both in the current and in the voltage transient, with respect to the usual BJT behavior in the active region. The result is an increase of turn-on dynamic loss because of the longer the time (from t_2 to t_5) where both current and voltage are relevant.
This effect is more pronounced for high voltage BJT's where the quasi-saturation region is larger.

b) turn-off

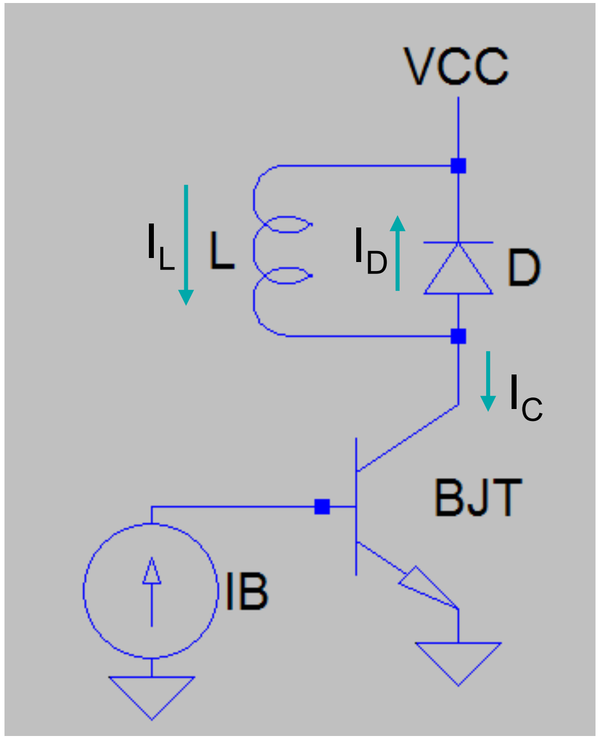
To switch off the BJT, we must eliminate all the stored charge in the “electrical” base, that in deep saturation, extends fully into the collector layer W_C. The result is that the collector current I_C start to decrease after a “storage time” (from t_1 to t_4) needed to remove these charges from the $W_B + W_C$ layer.
This time actually delays the turn-off time, but is not too detrimental in term of energy loss during turn-off, because during that time the collector voltage V_{CE} is still very low.
Both the storage time and turn-off time are reduced if the drive circuit can extract from the base a larger I_{B2} current: we will return on this point discussing the drive circuits.





c) Inductive load switching

What happen in the turn-on and turn-off if the load is an inductive one (we know that is the typical load for power circuits)? With reference to the simplified circuit in figure, with an inductive load with a fly-back diode (to allow the I_L current flow when the BJT is in off state), assuming a constant current I_L into the inductance during the turn-on and turn-off transients, we can do the following analysis.

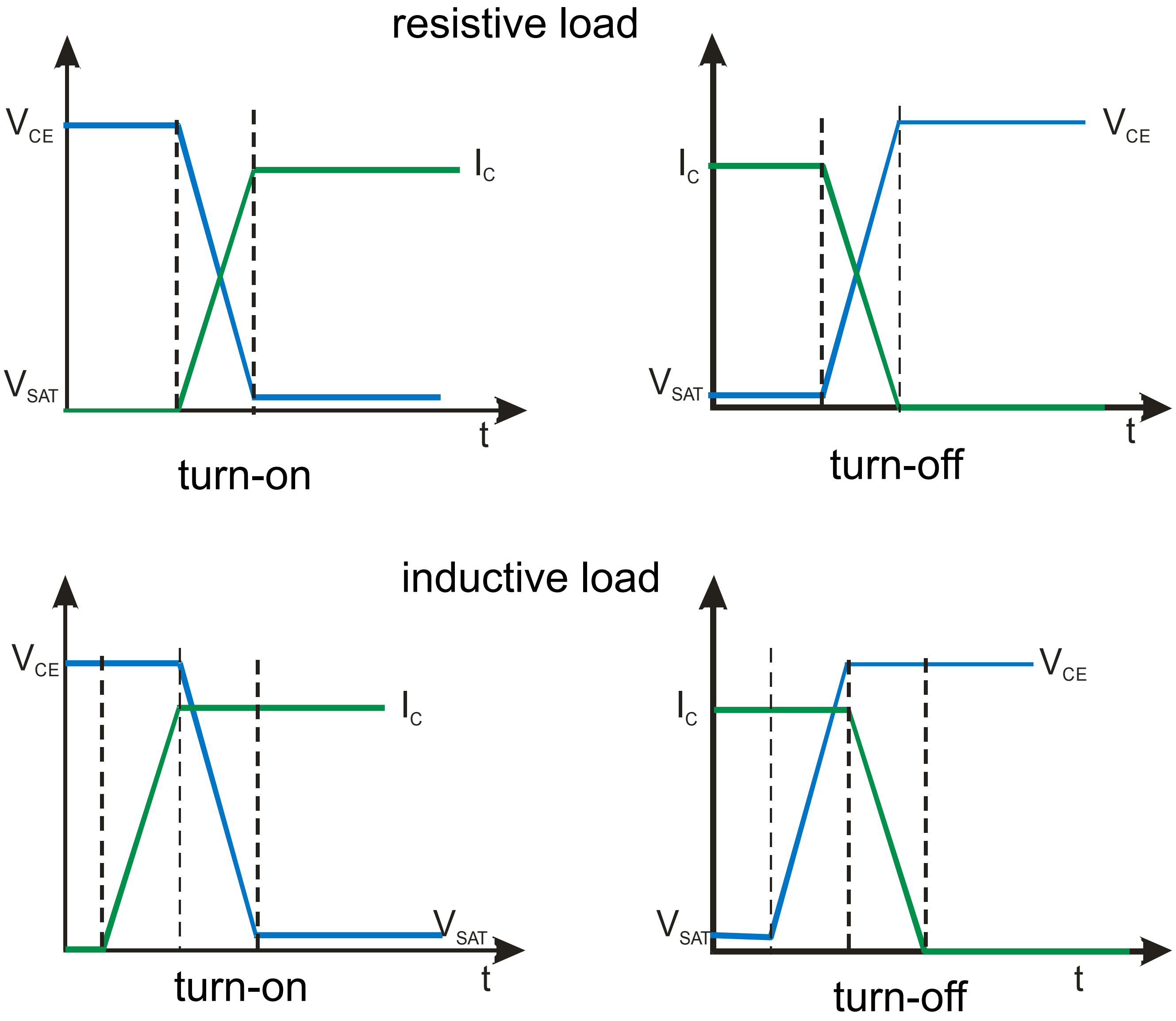


a) turn-on: before transient the BJT is in OFF condition ($I_C = 0$) and all the current I_L is flowing into the diode D; this latter is in ON with a $V_{ON} \cong 1V$. During turn-on the current I_C start to increase, but the diode D is still in ON state until the current I_C becomes equal to I_L, and the voltage V_{CE} start to decrease only at this time (and the diode is in OFF state with a reverse bias across it)

b) turn-off: before transient the BJT is ON and all the current I_L flows into the collector; after the storage time the collector voltage start to increase, but the current I_C is still equal to I_L because the diode is still in OFF state until the V_{CE} increases up to the V_{CC} value.

Only at that time the diode goes in ON state with a forward current I_D and the current I_C start to decrease. The difference I_L-I_C is the diode current I_D and after turn-off $I_D = I_L$.

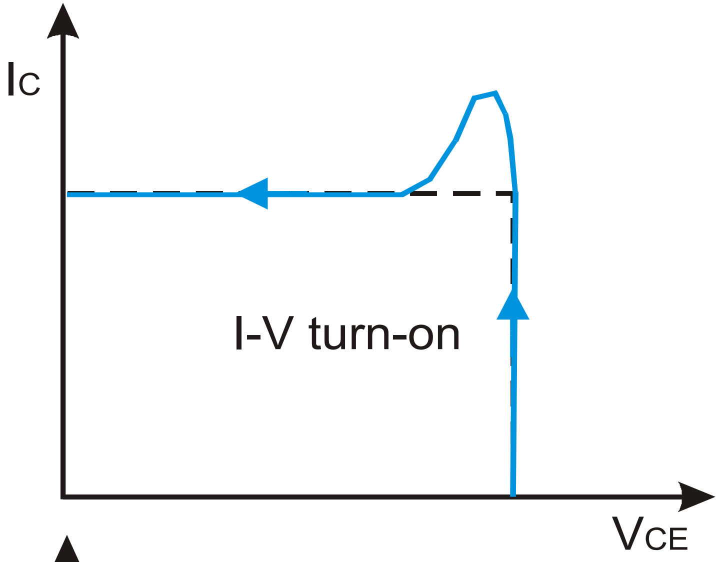
Resistive and inductive load switching



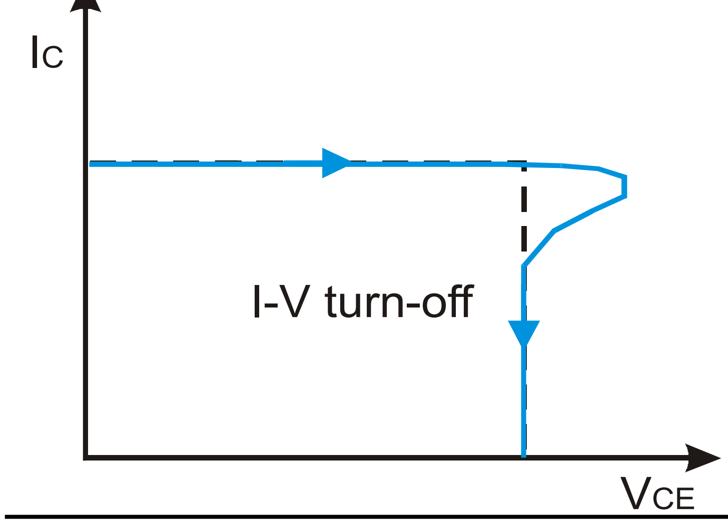
In this figure the dynamics of resistive and inductive switching are ideally sketched according to the analysis done previously.

We can see that in the case of inductive switching the power losses are larger than for resistive switching, because in the latter case, during the transient of one variable the other remains constant at the max value (In the above analysis the diode is assumed ideal, and no parasitic effects are considered).

A different plot of the current-voltage dynamics during turn-on or turn-off losses can be done, by referring to the operating point locus in the IV plane. For the resistive switching this locus is obviously the load line itself, while for the inductive switching the operating point follows an approximately square line, as indicated in the following figures.



For the turn-on locus in the upper plot it is also considered the added effect of the diode reverse recovery: the current overshoot in the IV locus at $V_{CE} = V_{CC}$ is due to the reverse current I_R of the diode that adds to the I_L current, increasing the I_{CMAX} needed for the BJT at turn-on.

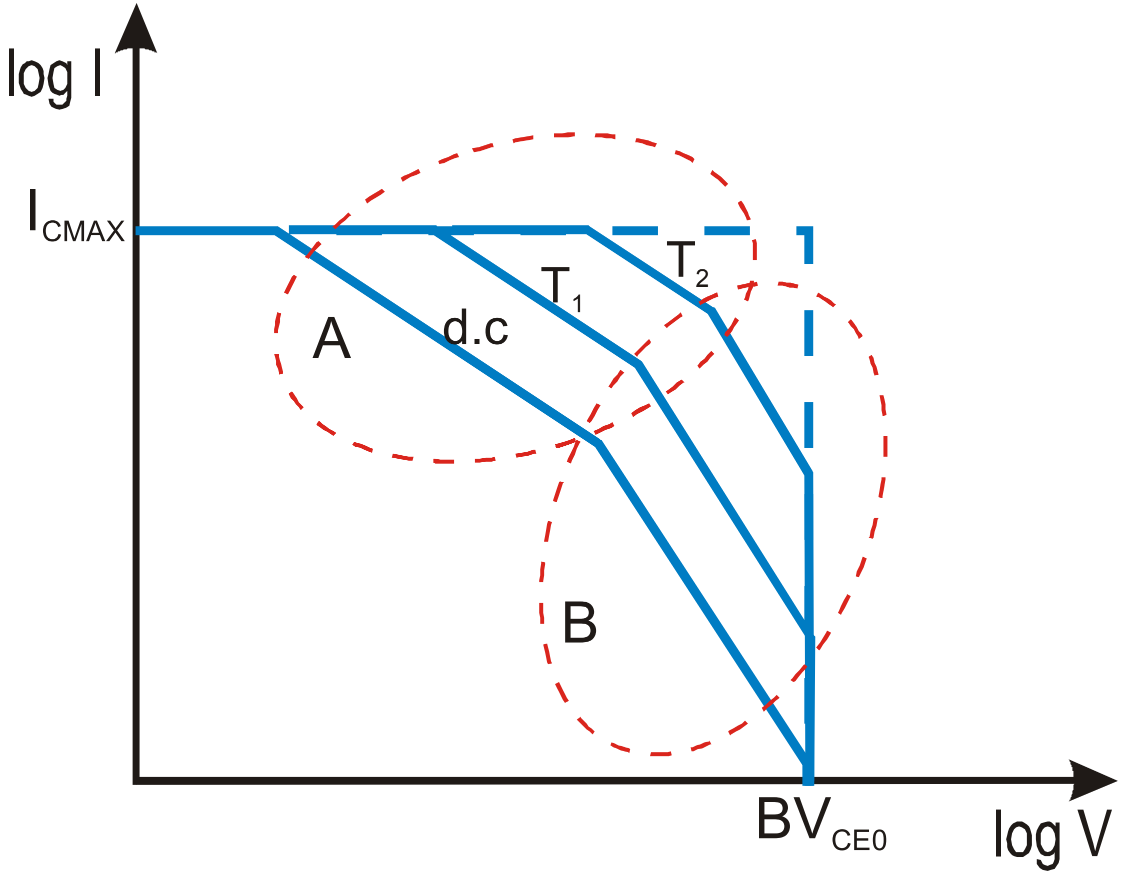


For the turn-off locus it is also considered the added effect of the parasitic stray inductance due to the connection wires between diode and BJT: the voltage overshoot at $V_{CE} = V_{CC}$ is due to the ΔV across this inductance when the current I_C starts to decrease, thus increasing the V_{CEMAX} needed for the BJT at turn-off.

Safe Operating Area (S.O.A) of BJT

From the previous analysis of the static and dynamic BJT behavior we can now understand the limits indicated in the SOA plane of power BJT's, as reported in the datasheets.

a) Forward bias SOA (FBSOA)



The SOA in forward bias (FBSOA), both in d.c. and pulsed operation, is limited:

a) by a max current I_{CMAX} for wich the current gain in saturation has still a reasonable value (about 10)

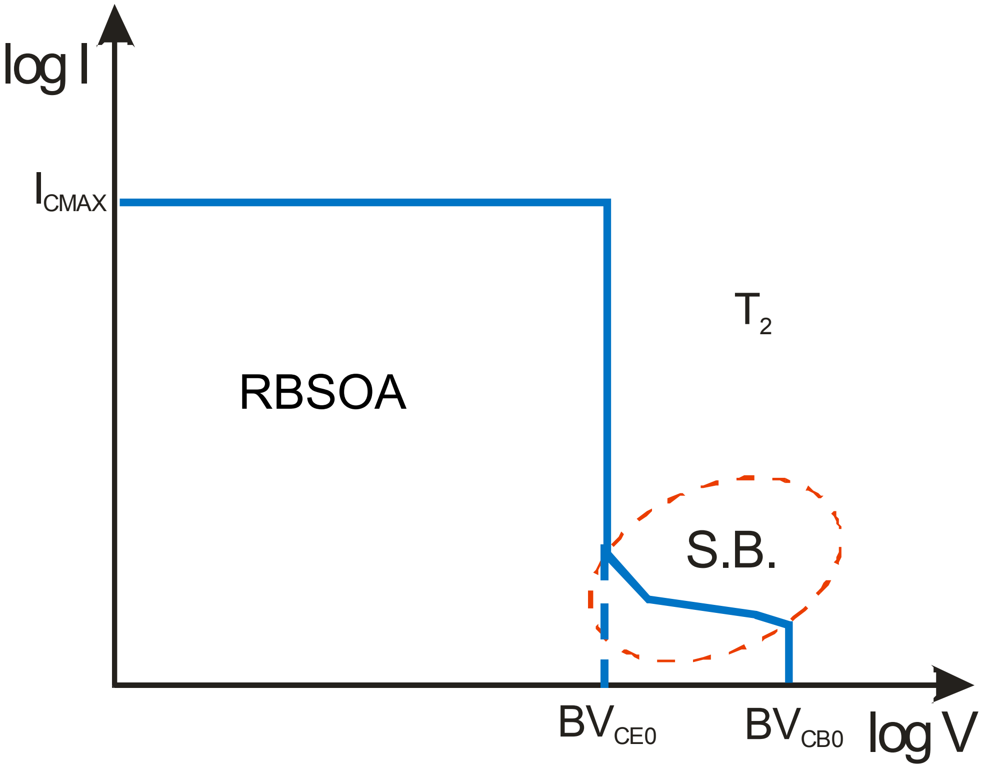
b) by a max voltage BV_{CEO} lower than the breakdown voltage BV_{CB0} (this is a limitation of BJT)

c) in reg. A by the max power line P_{DMAX} = const. (related to the max junction temperature T_{JMAX} and to the thermal impedance Z_T that decreases for pulse times $T_2 < T_1$)

d) in reg. B by a line with increased slope due to the onset of thermal instability that ask for a lower I_C at larger V_{CE} to prevent the hot spot formation.

b) Reverse Bias SOA (RBSOA)

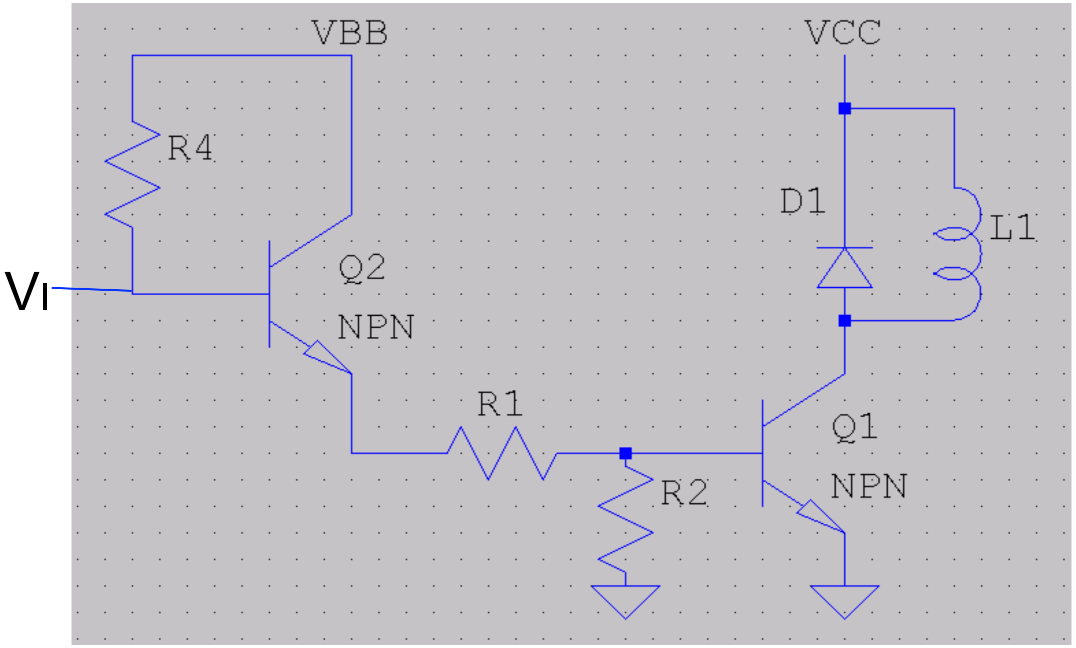
The reverse bias SOA is valid only during turn-off switching, and due to the very low switching time the thermal limitations do not apply. A typical RBSOA is reported in the figure.



The voltage limitation for collector currents of the order of the operating currents is given by the BV_{CE0} value, because at high currents we know that the asymptote for both current and voltage drive is the same.
At very low currents, and with voltage drive, we can exceed this value up to the BV_{CB0} value. However the operating locus must be contained in the SOA, because, as we noted from the BV_{CB0} analysis, if we exceed a critical current that brings the emitter/base junction in forward bias, we enter in a negative resistance region, called second breakdown, that will destroy the device, due a strong current constriction in the center of the emitter area.

Drive circuits for power BJT

The drive circuit must give a base current I_{B1} to the power BJT, to bring it in saturation when in ON state to reduce the static power losses, and must allow a negative base current I_{B2} flow to accelerate the turn-off. Recalling that for high voltages BJT the current gain β is strongly reduced by the quasi-saturation region, we need a quite high base current to bring it in saturation: with a β of 10 we need an I_{B1} = 5 A to allow a collector current I_C = 50 A in on state.



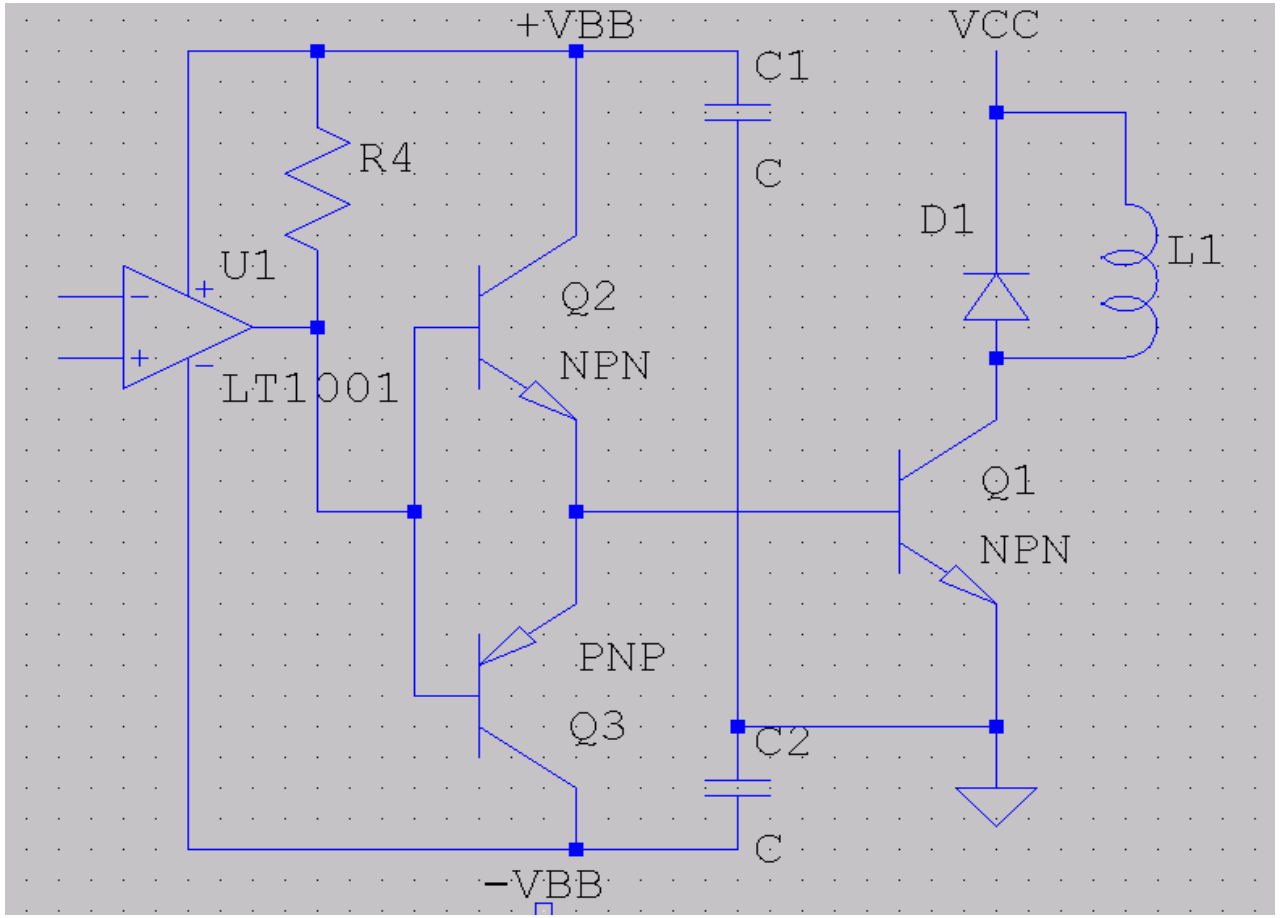
A first solution is to use as a drive stage a BJT in a common-collector configuration, that as a large current gain (we do not need a voltage gain), as reported in the figure: in this scheme, the resistance R_1 will limit the base current I_{B1} to avoid excessive charge storage during ON state, while R_2 will allow a negative base current $-I_{B2}$ due to the V_{BESAT} :

$$-I_{B2(Q1)} = -\frac{V_{BEsat(Q1)}}{R_2}$$

However, I_{B2} can not be made quite high because if R_2 is reduced , also I_{B1} will be reduced:

$$I_{B1(Q1)} = I_{E(Q2)} - \frac{V_{BEsat(Q1)}}{R_2}$$

A better solution is the use of a push-pull configuration for the drive stage. The circuit requires two complementary BJTs (NPN and PNP) but it allows an independent control of both I_{B1} and I_{B2} currents, that can be easily made equal (in absolute values), if the two BJT of the driving stage have equal current gains.



The drawback of that circuit is the need of two power supplies +VBB and –VBB to bias the two complementary BJT, but the double voltage bias is usually needed in analog circuits, as for the operational amplifiers used for signal processing.
(a further analysis of the driving stages will be done by using the SWCADIII simulator in a BJT exercise)