Insulated Gate Bipolar Transistor (IGBT)

Comparison between BJT and MOS power devices:

	BJT			MOS	
pros	cons		pros	cons	
low V_{ON}	thermal instability		thermal stability	high R_{ON} at $V_{MAX} > 400$ V	
high I_C current	complex drive circuits		ease of drive circuit	lower I_D current	
high voltage	no paralleling		can be paralleled	low or medium voltage	
	second breakdown				

It is possible to merge the pros of each device in a new device that has:

ease of drive circuitry (voltage driven)
thermal stability
paralleling capability

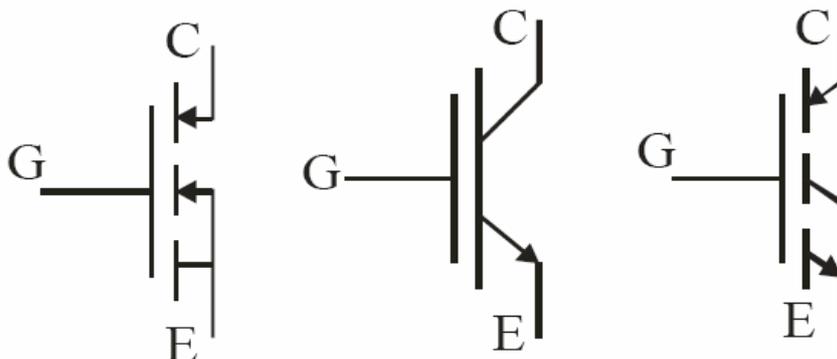:   IGBT

low V_{ON}
high voltage capability
ruggedness

The IGBT is now the preferred power device for voltages from 500V up to 4000V (and more)

The IGBT power device is a device that has a gate as input electrode (insulated by the oxide from the output terminals), as for the power MOS, but has a conduction current given by both electrons and holes, to allow a conductivity modulation of the low doped and thick collector region needed to sustain high output voltages (as for the power BJT).

Its name came from this two main aspects: Insulated Gate and Bipolar Transistor.

The electrical symbol can be one of the following:

The input terminal is named GATE – as in the MOS

The output terminals are named COLLECTOR (V+) and EMITTER (V-) – as in the BJT.

The cross-section of an elementary IGBT cell reveals the strong similarity with the one of a vertical power MOS, but with one (most relevant) difference: the bottom substrate that terminates the low doped N⁻ region is not an N⁺ layer, but a P⁺ layer.
Then a bottom P/N junction is added to the MOS cell.
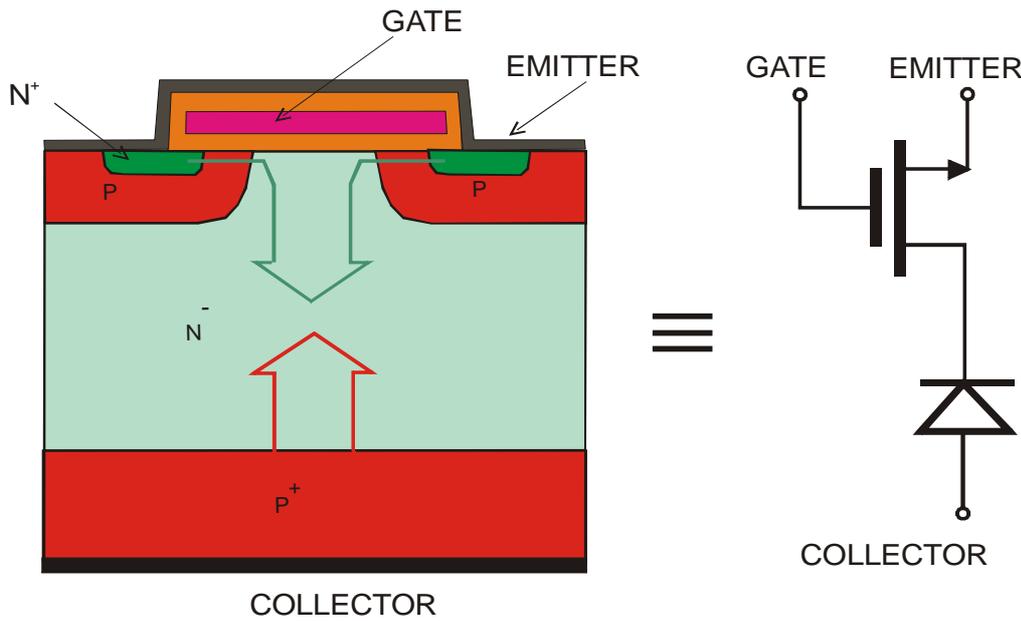
The main role of this P/N junction at the collector, when the device is in ON state, is to inject holes from the bottom, while electrons are injected from the top: that will induce Conductivity Modulation of the epi region and will reduce drastically the resistance R_{ON} of the epi layer.

In this structure, one can identify apart from the power MOS, not only a P/N junction, but also a lateral NPN transistor (jellow box), and a vertical PNP transistor (blue box), with its base merged with the collector region of the lateral NPN.
This latter configuration give rise to a NPNP parasitic SCR structure (that we do not want to trigger in latch-up state)
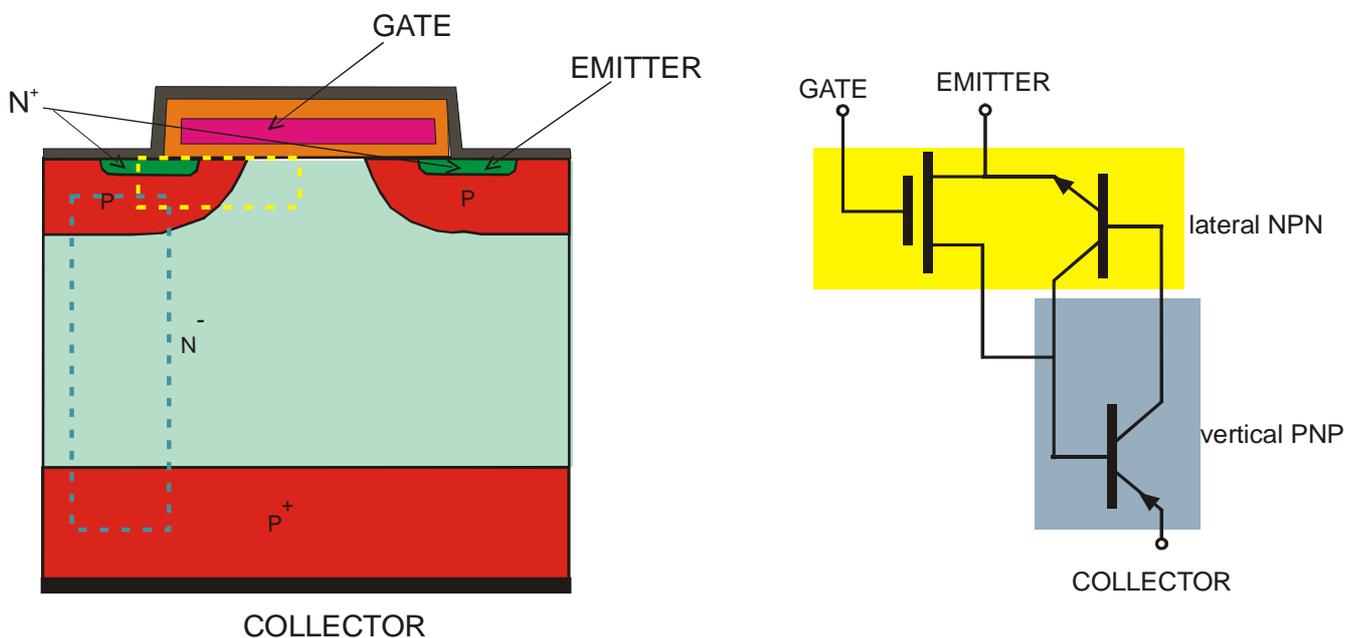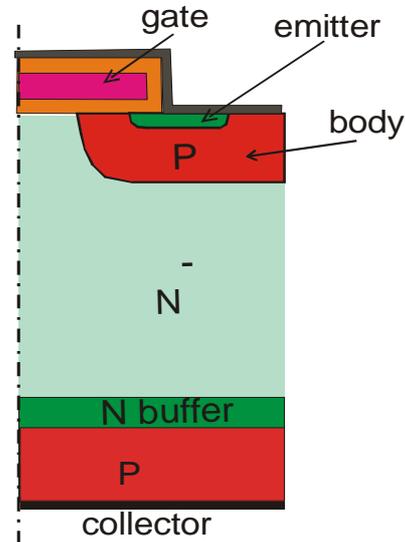
Basically there are two structures for the IGBT:
a) the symmetrical structure named Non Punch Through (NPT) device
b) the unsymmetrical structure named Punch Through (PT) device. This latter has N buffer layer at the boundary of the low doped N⁻ epi layer to avoid the punch through of the body collector P regions
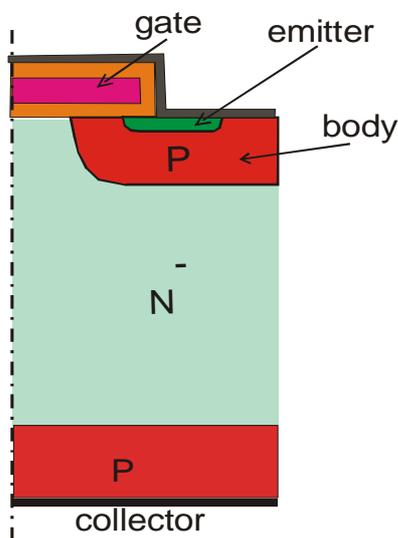
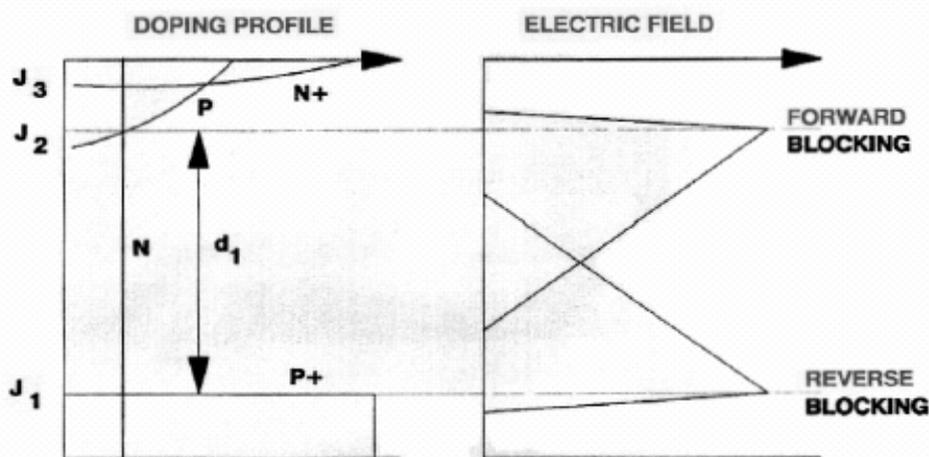a) NPT structure

b) PT structure

Breakdown voltages of NPT and PT devices

a) NPT structure:

For this device, considering the vertical part made by the N emitter, P body, N epi and P collector layers, the collector voltage can be sustained by the J2 body/epi junction in the forward blocking mode, and by the J1 collector/epi junction in the reverse blocking mode.

The field profile is triangular and goes to zero before reaching the opposite junction. Then there are no differences in V_{CEMAX} in forward or reverse bias. V_{CEMAX} can be considered as the BV_{CE0} of the PNP BJT in oper base configuration, butb in this case BV_{CE0} can be very near to BV_{CB0} because α is quite low due to the large width of the N base (epi layer)

b) PT structure:

For these devices, the field profile is trapezoidal, and in forward collector bias it reaches the epi/collector junction.

To avoid the punch-through of the body-collector P regions, a thin N+ buffer layer is inserted between N epi and P collector, to allow the field to go to zero before reaching the P layer. Then the forward blocking voltage V_{CEMAX} differs largely form the reverse blocking voltage: this latter is very low (some tens of volt) because it is actually the breakdown voltage of the N+/P+ buffer/collector J1 junction
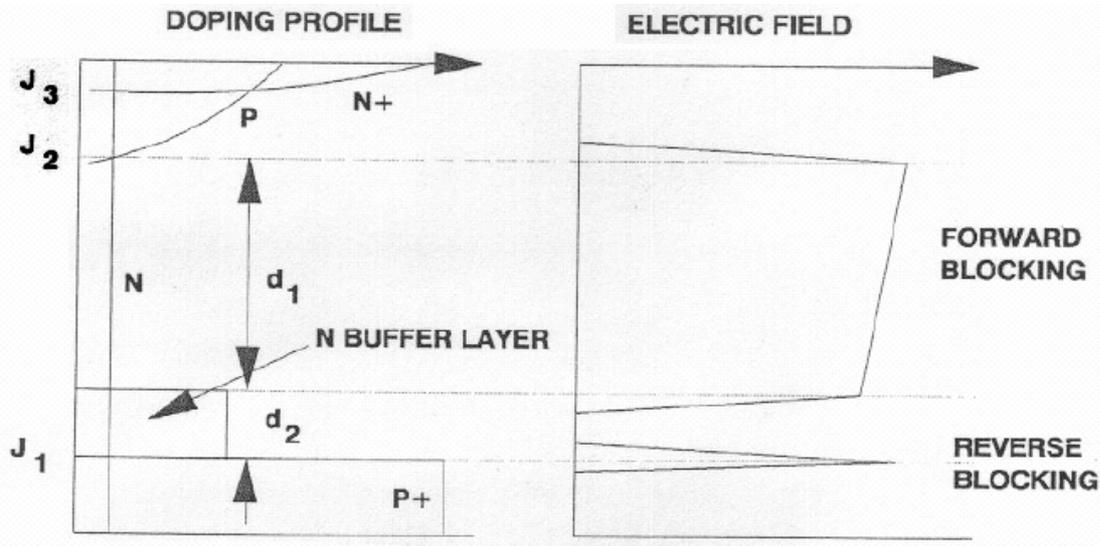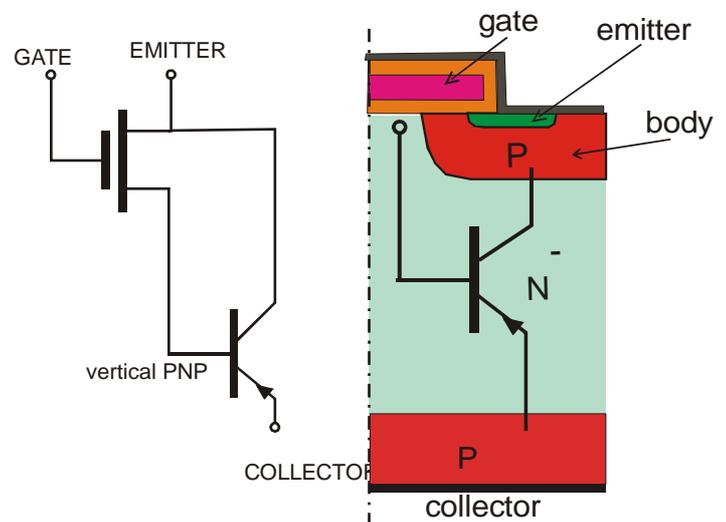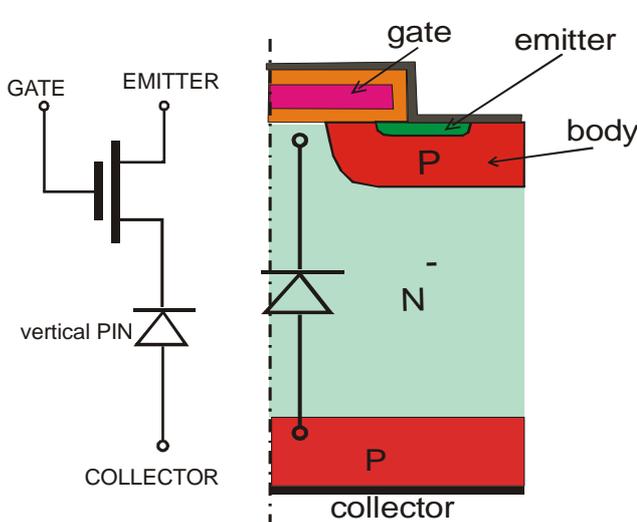
IGBT modeling

Two different basic models are used for the IGBT modeling in compact models:

a) MOS + PIN diode

b) MOS + BJT

The model MOS + PIN is more adequate for the NPT (symmetrical) IGBT structure, where the thickness of the epi layer is quite high and the doping not so low, so the interaction between the two junctions P/N body/epi and N/P epi/collector is negligible.
In this case we can consider the PNP vertical BJT having a current gain very low, so the two junction do not interact each other.

The model MOS + BJT is more adequate for PT IGBT structure, where the epi layer has a reduced thickness and the doping quite low. Then the two junctions body/epi and epi/collector interact each other, and we must consider these junction as making a vertical PNP transistor, and thaking into account the current gain resulting from this bipolar structure.

ON voltage drop with MOS + PIN model

In this case the ON voltage drop is the sum of the voltage drop across the lateral MOS and the one across the PIN diode
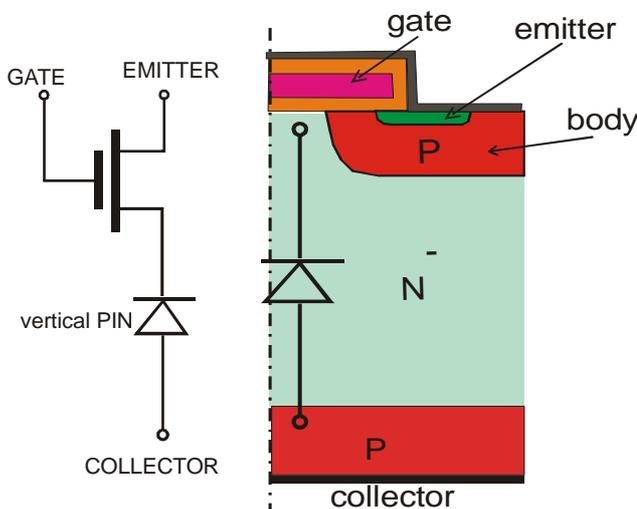
The IV expression for the MOS at low V_{DS} is:

$$I_D = \frac{\mu C_{OX} W}{L}(V_G - V_T)V_{MOS} \Rightarrow V_{MOS} = \frac{I_D L}{\mu C_{OX} W(V_G - V_T)}$$

The IV expression for the PIN diode is:

$$I_{PIN} = I_0(\exp\frac{qV_{PIN}}{2V_T}) \Rightarrow V_{PIN} = 2V_T \ln(\frac{I_{PIN}}{I_0})$$

The IGBT ON voltage, referred to the common collector current I_C is then:

$$V_{ON,IGBT} = \frac{I_C L}{\mu C_{OX} W(V_G - V_T)} + 2V_T \ln(\frac{I_C}{I_0})$$

The ON voltage at very low low currents is the threshold diode voltage V_γ. At increasing gate voltages V_{ON} is still the PIN diode voltage plus the (low) voltage drop due to the R_{chan} of the lateral MOS.
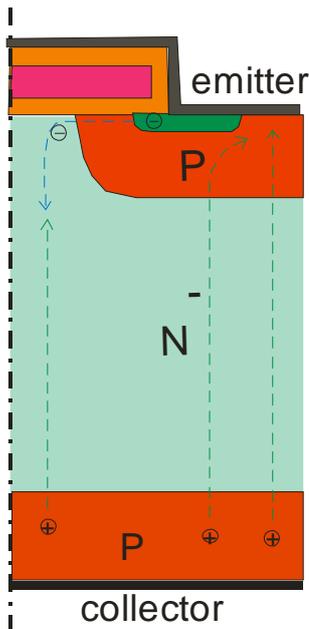
In comparison, for a MOS of the same voltage rating, the voltage drop is the one of the $R_{ON,EPI}$, much higher that the one of the IGBT.
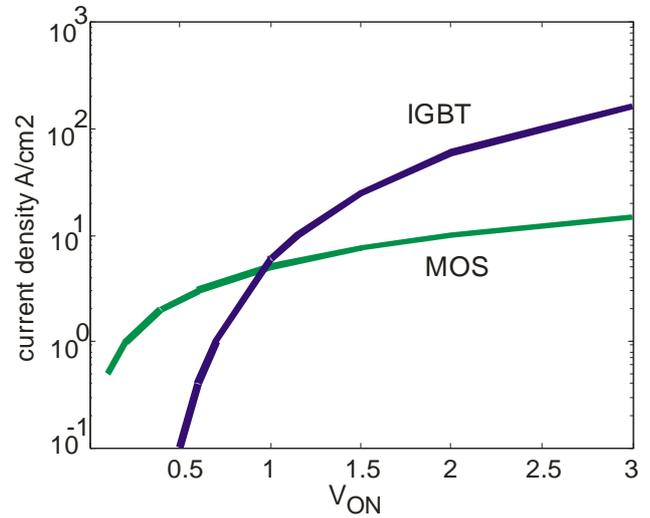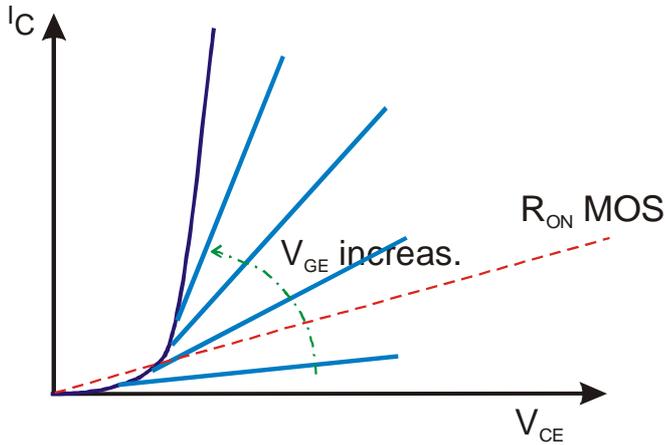
When the collector voltage V_{CE} increases, the MOS reaches the pinch-off current (max current for a given V_{GE}) and the IGBT behaves as a current controlled device, like the MOS.
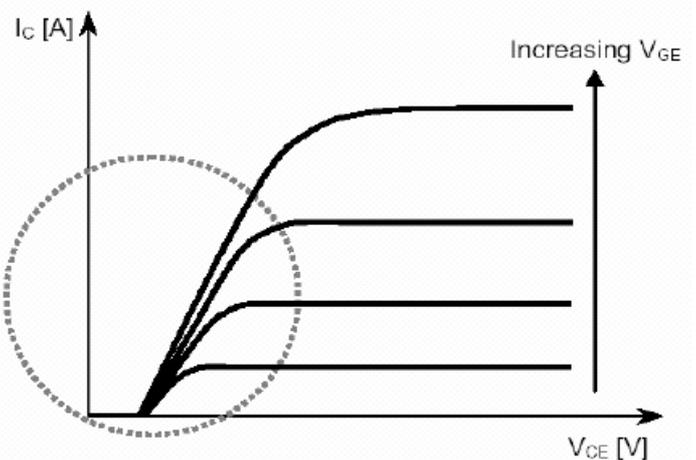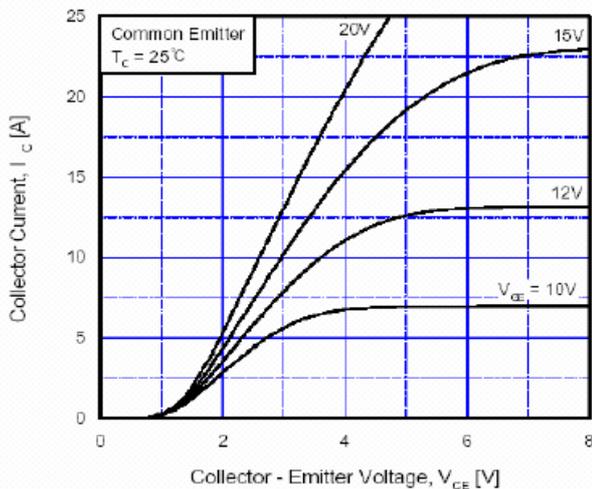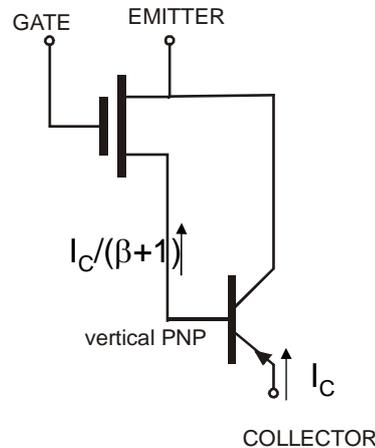
ON voltage drop with MOS + BJT model

emitter

P

-

N

P

collector

In this case, the MOS current is only a component of the total IGBT , and in particular it is the base current of the PNP transistor, as indicated in the circuit schematic of the structure.

Then the MOS drain current can be expressed as: $I_D = \dfrac{I_C}{\beta+1} = (1-\alpha)I_C$

GATE   EMITTER

$I_C/(\beta+1)$

vertical PNP

I_C

COLLECTOR

and the ON voltage expression obtained by the MOS + PIN diode, expressed as a function of the total collector current IC can be modified as:

$$V_{ON,IGBT} = \frac{(1-\alpha_{PNP})I_C L}{\mu C_{OX} W(V_G - V_T)} + 2V_T \ln(\frac{I_C}{I_0})$$

The MOS component of the total V_{ON} voltage is then reduced because of the lower MOS current with respect to the total I_C current.

PT structures vs NPT structures

We know that the PT structures with N buffer layer do not have capability of reverse blocking voltages, so why they are used?

From the model MOS+BJT that is suitable for PT structures, it comes out that the V_{ON} voltage drop pf PT structures,for the same I_C current, is lower than for the case of NPT structure ,where the MOS+PIN model holds.

Furthermore, the reduced epilayer width will make the switching from ON to OFF faster, because of the lower stored charge in this layer (the base of the PNP transistor).

Finally, for the same thickness of the epi layer, with a quasi rectangular field profile, the breakdown voltage is about double that in the case of a triangular field profile.

From both models, the main features of the IGBT become clear:
A minimum voltage drop of about .7 V is present in the IGBT even at negligible currents, so the device is not suitable for low supply voltages.
The IGBT voltage drop at high currents (above 100 A/cm^2) is much lower than the one of a MOS device for V_{MAX} above some hundreds of volts, as reported in this figure.
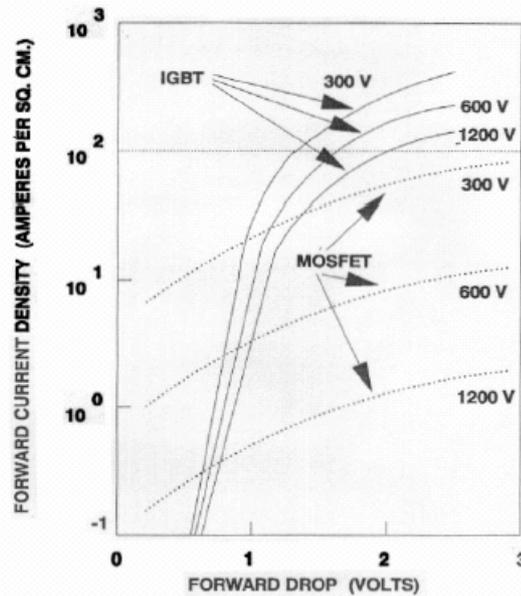
# Temperature effects on the D.C. curves

At low currents the bipolar effect (either PIN or BJT) is predominant, and the collector current shows a positive temperature coefficient. At high current the MOS effect become predominant and the collector current presents a negative temperature coefficient.
The negative temperature coefficient allows the paralleling of devices and prevents from thermal instability.
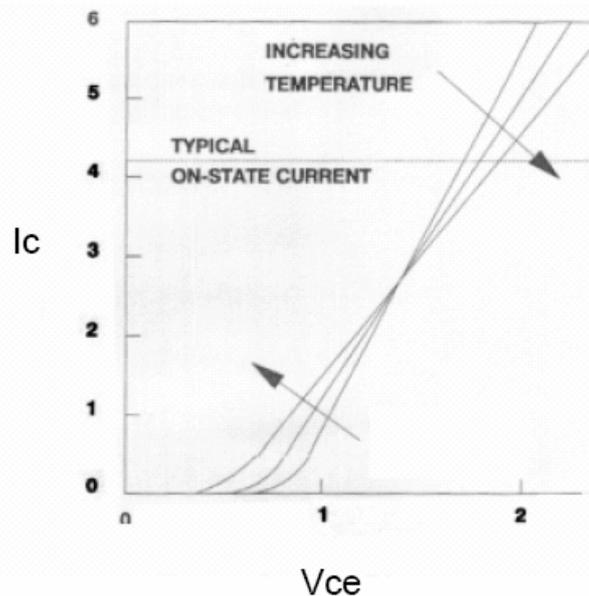
Switching behavior of the IGBT

a) Turn-on on inductive load
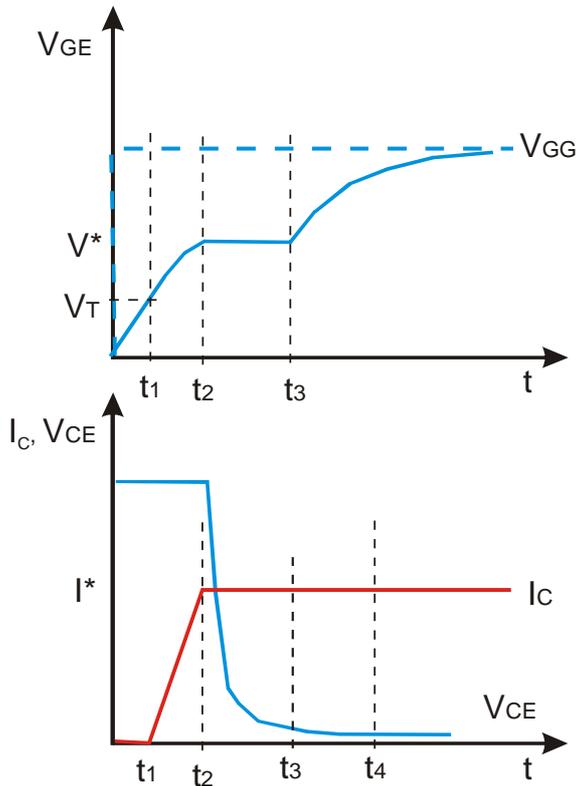
The behavior is similar to the one of power MOS:

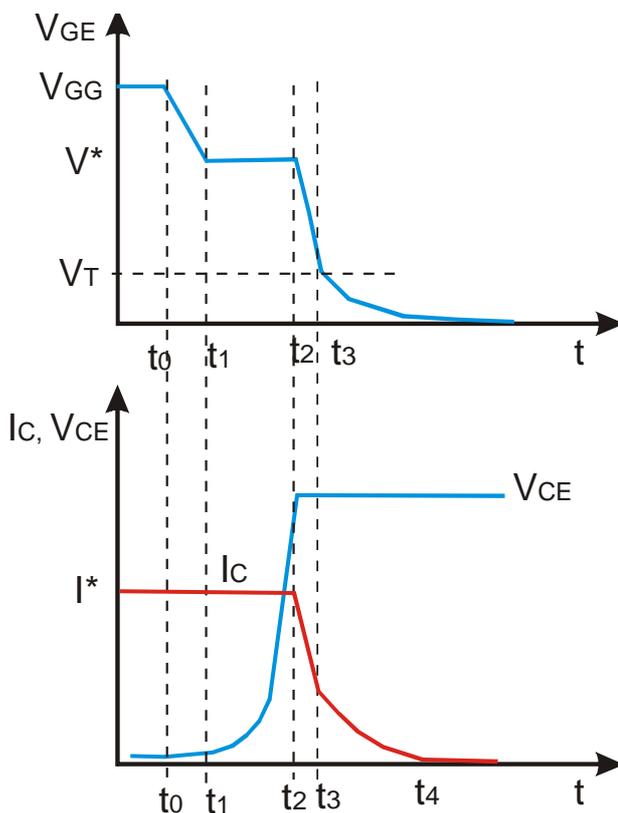$0-t_1$=delay time – the gate voltage has to reach the threshold voltage

t_1-t_2 = current rise time – the collector current rises to the total inductance current I*

t_2-t_3 = voltage fall time - the voltage drops due to the turn-on of the MOS component

t_3-t_4 = this is the time required to bring in saturation the PNP vertical transistor (relatively high time constant due to the thicker base of the PNP)

b) Turn-off on inductive load

In this case the behavior differs from the one of a power MOS:

$0-t_1$=delay time – the gate voltage has to reach the pinch-off value V* after which the voltage of the MOS component start to depend on V_G

t_1-t_2 = voltage rise time – the collector voltage rises to the supply voltage

t_2-t_3 = first part of the current fall time: the MOS current component (electrons) will go to zero, but the total collector current is not going to zero, because of the BJT component. This latter is still active because the charge (holes) stored in the PNP base can not be extracted by a base terminal.

t_3-t_4 = this is the time required to bring the PNP vertical transistor in the OFF state, due to the turn-off time of the PNP that can be quite high. This part of the current fall is named current tail

Current tail in IGBT turn-off
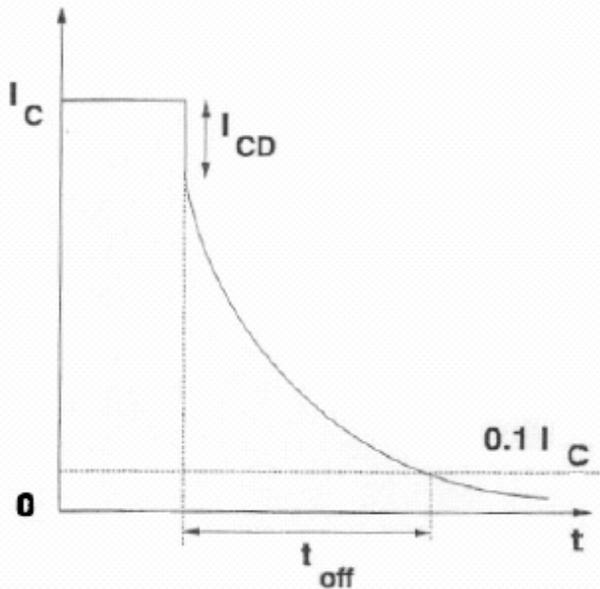
The current tail due to PNP BJT component (or to the PIN diode stored charge in NPT structures) was a significant drawback in IGBT devices with respect to the MOS ones.

Due to the exponential fall-off of the current in the tail, it can largely increase the turn-off time t_{off} - defined as the time for the current to drop from 90% to 10% - if the MOS component – indicated as I_{CD} in this plot, is much less than the total current I_C.
To reduce the current tail one must reduce the BJT component of the total collector current, so that I_{CD} >> I_{CBJT}.

Then we have again a trade-off between low voltage drop in ON state – that requires a large bipolar component of the total current – and fast turn-off time – that requires that the BJT current component is reduced with respect to the MOS one.

The current tail has the following negative aspects:

-increase of dynamic power dissipation during the turn-off
-limitations of the maximum operating frequency of the IGBT

Most efforts were made in the early IGBT devices to reduce the current tail. This tail is due to the quite slow turn-off of the vertical PNP transistor, that will turn-off at open base (it has no base contact to be used to extract the stored charges through a negative base current).
The best solution has been the use of lifetime control processes to the epi layer to reduce the time need for the stored charge in the epi layer to recombine.
Now the actual IGBT for fast switching applications present a quite limited current tail due to suitable control of the lifetime.

The techniques most used use are electron helium irradiation to create recombination centers in the epilayer.
The decrease of lifetime has two positive effects on the t_{off}:

-it will reduce the turn-off time of the PNP transistor, and hence the duration of the current tail

-it will reduce the BJT current component with respect to the MOS one because it reduces the current gain of the BJT, and then the peak value of the current tail is decreased.

As an example, in the following figure there are reported the I_C current turn-off waveforms, for IGBT without (fig. a), and with electron irradiation (fig. b) to reduce the lifetime. In the latter case the I_{CD} MOS component is increased, so the current tail peak value is decreased, and the time for the tail to go to zero is decreased, so the t_{off} time is also decreased.

before electron irradiation   After electron irradiation

(a)   (b)
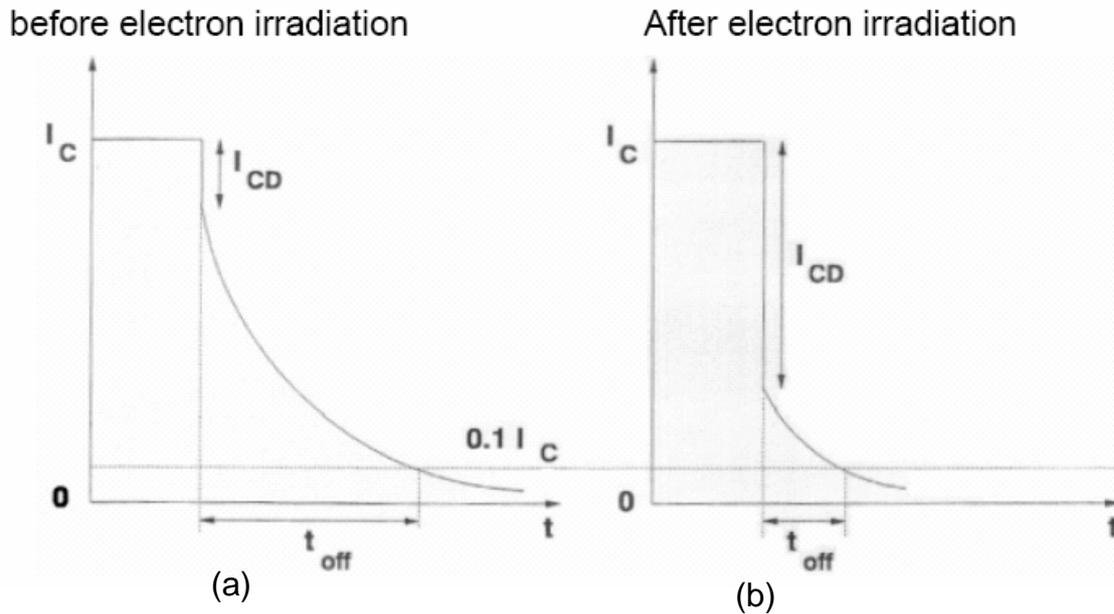
## Temperature effects on turn-off behavior

When temperature increases, the carrier lifetime increases; this has a twofold effect on current tail:

a) the current gain of the PNP BJT increases, and the bipolar current component become larger

b) the recombination of the stored charge takes a longer time.
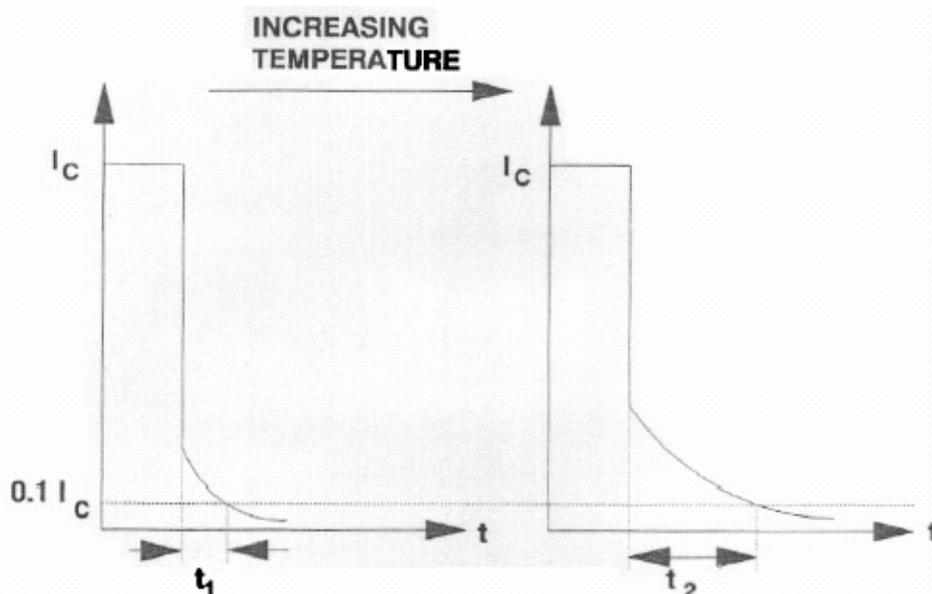
Latch-up of the IGBT

A problem, faced by the early IGBT, was the undesired turn-on of the parasitic SCR structure present between emitter and collector in the basic IGBT cell. The turn-on of the SCR will lead to the failure of the device, that will stay at the latching voltage (very low) with a current limited only by the external load.
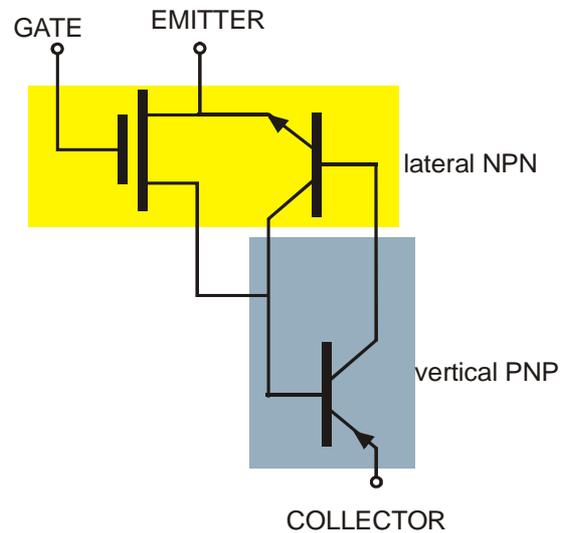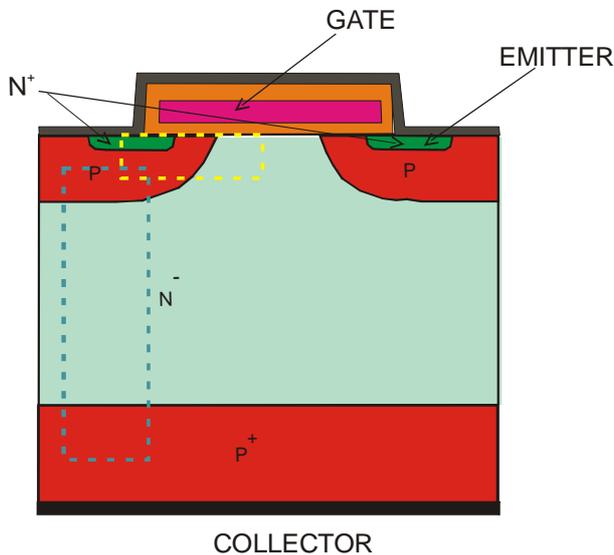
The solutions to prevent latch-up of the PNPN structure are basically two:

a) to reduce the current gain of the lateral NPN bjt
b) to reduce the current gain of the vertical PNP bjt.

The decrease of the current gain of the vertical PNP device has a negative effect on the V_{ON} voltage of the IGBT, and reduces the conductivity modulation of the epi layer
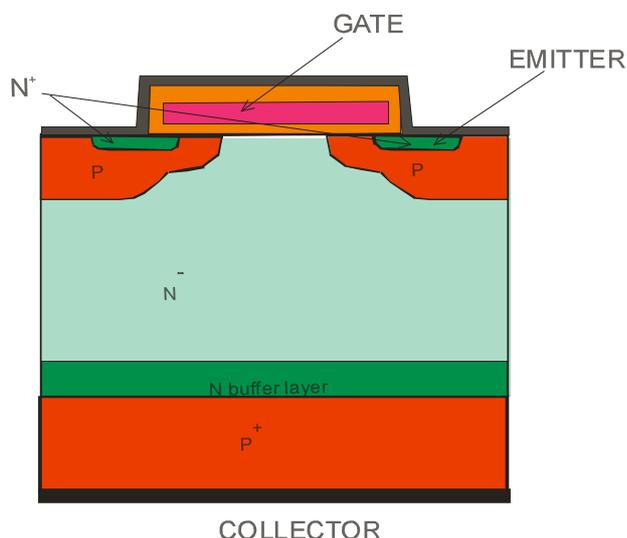
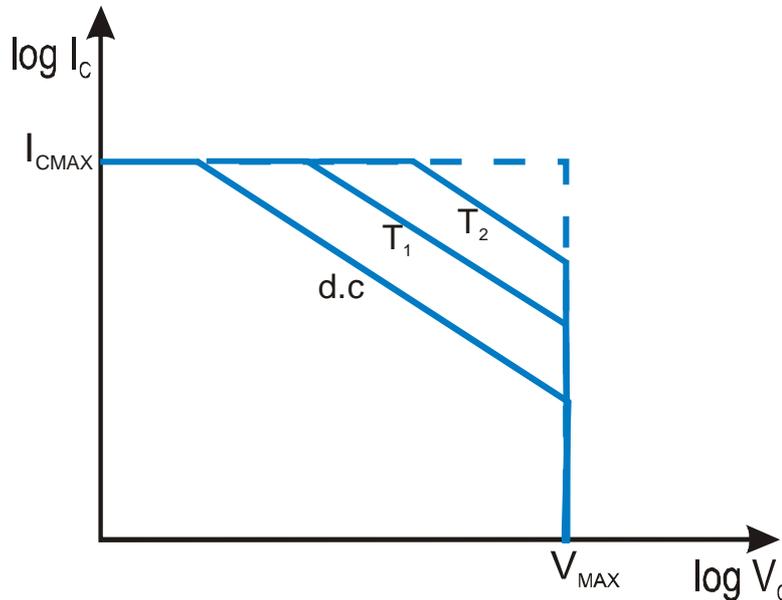It is preferable to reduce the current gain of the lateral NPN bjt .
It could be done (as for the Power MOS) by using a double implant of the body region: a lateral one to define the channel length, and a deeper one of high dose under the emitter. This latter will reduce the resistivity of this region (the base of the lateral NPN) and will increase the thickness of this layer. It will reduce the spreading resistance between base and the body contact, and avoids the premature triggering of the SCR.

IGBT S.O.A

The S.O.A. of the IGBT does not presents the low voltage limitation (of the Power MOS) due to the R_{ON}, neither the current limitation at high voltage (of the BJT) due to the thermal instability.

The max power limitation curves, due to a max operating temperature of the device, are a series of constant slope lines with P_D = const. that depend on the power pulse duration T.

Datasheet of an IGBT