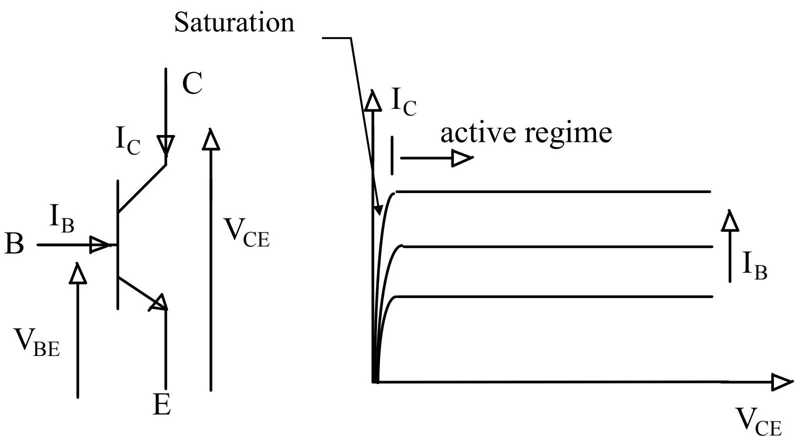
The Bipolar transistor

Introduction and basic operation

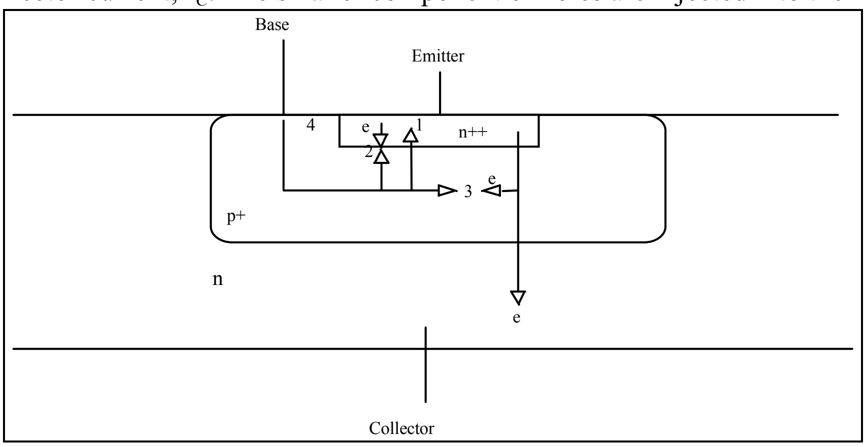
The bipolar junction transistor or BJT, is commonly used in both discrete and integrated circuit applications. Despite its inherent superiority over the MOSFET device, the use of BJTs in modern digital integrated circuit electronics is becoming less widespread and this is because undesirable features of its operation can only be avoided by the use of rather complex circuitry which inevitably reduces the amount of functionality on a chip. Specifically, it is the charge storage associated with the saturated mode of operation that slows down considerably the switching performance of the bipolar transistor. The emitter coupled logic (ECL) family involves circuitry that prevents the transistors saturating and the full potential of the devices can then be realised. ECL is the fastest Si-based logic family.

The symbol for an n-p-n transistor and the corresponding output characteristics are shown below.



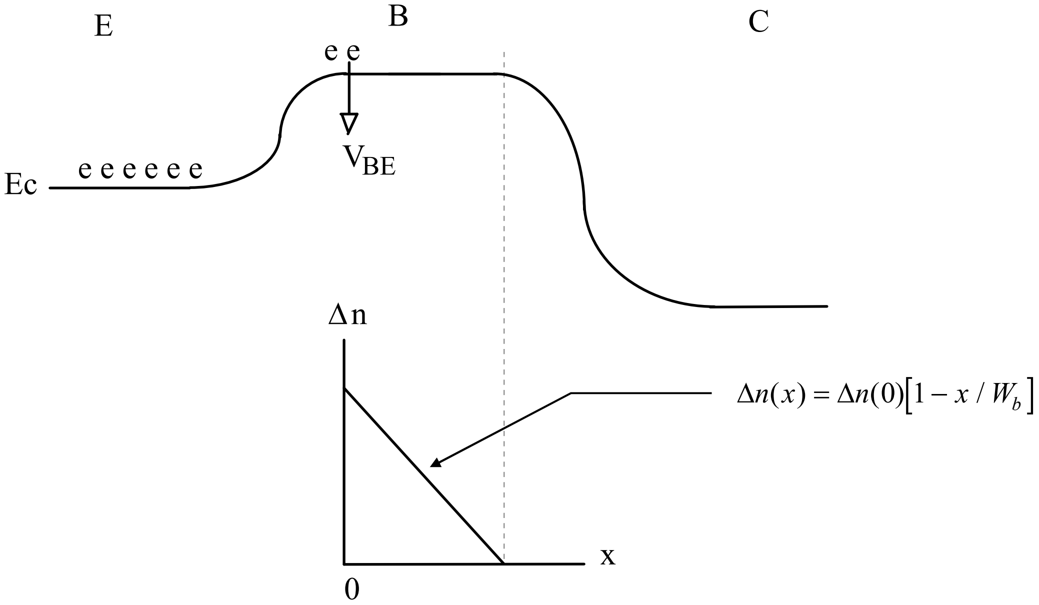
We consider first the D.C. operation of the BJT in the active mode of operation, that is with the emitter-base junction forward biased and the collector base junction reverse biased. The discussion is presented for the n/p/n transistor where electrons carry the collector current and holes the base current, but the treatment is the same for p/n/p device with the opposite type of carriers respectively carrying the currents. A schematic diagram of the BJT is shown below, with the important electron and hole components indicated.

For modern, low and medium power BJTs, acting in the active mode, essentially all the action occurs at the emitter/base junction. The forward bias current across this junction comprises of mainly electrons and some holes. The electrons are injected into the base, which they rapidly transit and then fall into the reverse biased collector where they constitute the collector current, I_C. The smaller component of holes are injected into the emitter and essentially constitute the base current, I_B. Because the n-type doping is much greater than the p-type, the current is carried mainly by the electrons and hence the device has current gain (β) given by the ratio I_C/I_B. The trick has been to initiate a large current flow in the output on the device (collector) with a correlated small current in the control electrode (base).



Mechanism for collector current

As stated above, the collector current arises quite simply from the injection of electrons across the forward biased emitter/base junction, into the neutral base region. This is represented in the band diagram below.



Once in the base, the electrons diffuse across to the collector depletion region. The resulting distribution of the excess (injected) electrons across the base, Δn(x), is shown above. When the electrons arrive at the collector they experience a very high accelerating field and are swept across into the neutral collector. Thus $\Delta n \to 0$ (small but not zero!) at the collector end of the base. The theory for I_C follows closely that of a 'short base' diode.

$$J_C = qD_e \frac{d\Delta n}{dx} \qquad (1)$$

From the above figure, we can write

$$\Delta n(x) = \Delta n(0)\left[1 - x / W_b\right] \qquad (2)$$

where W_b is the neutral base width. From p-n junction theory,

$$\Delta n(0) = \frac{n_i^2}{N_A}\exp\left(\frac{qV_{BE}}{kT}\right) \qquad (3)$$

Differentiating (2) and substituting, together with (3), into (1) gives the expression for collector current:

$$I_C = \frac{AqD_e n_i^2}{W_b N_A}\exp\left(\frac{qV_{BE}}{kT}\right) \qquad (4)$$

Equation (4) is somewhat oversimplified as it assumes a uniform doping density in the base. In most devices, the doping varies with distance (decreases). This gives rise to a built-in field in the base which serves to assist the transport of electrons and thus improves the performance. A more rigorous analysis taking this effect into account, yields:

$$I_C = \frac{AqD_e n_i^2}{\int_{base} N_A(x)dx}\exp\left(\frac{qV_{BE}}{nkT}\right) \qquad (4a)$$

where the integral on the bottom line is known as the Gummel number. The "ideality factor" n, in the argument of the exponential accounts for second order effects and is close to 1. Thus I_C in the active regime, is a well behaved quantity, obeying a near-to-ideal, diode-like equation. This is exploited in design in many ways.

Exercise: By considering equation (4), compare the collector current expected for the device acting in reverse mode (emitter and collector interchanged) with normal mode.

Mechanisms for base current

The main component is due to the reverse injection of holes into the emitter (labelled "1" in the figure in the first section). A similar analysis to that of the collector current yields a diode-like equation of the form

$$I_B = I_{BO} \exp\left(\frac{qV_{BE}}{nkT}\right) \qquad (5)$$

where, for the case of constant doping in the emitter,

$$I_{BO} = \frac{AqD_h n_i^2}{L_h N_D} \qquad (6)$$

The intrinsic carrier concentration in the emitter is written as n_{ie} as it differs from the value for the base, because of the effects of band-gap narrowing in the heavily doped emitter (the very heavy emitter doping introduces strain into the Si crystal and causes the energy band gap to reduce from the value 1.1eV; recall that $n_i \propto \exp\left(E_g / kT\right)$).

There are additional components to the base current and all obey diode-like equations. The net base current is the sum of the following components therefore:-

1: Reverse injection of holes into the emitter (ideality factor $n = 1$)
2: Recombination of electrons and holes in the emitter/base depletion region. ($1 < n < 2$)
3: Recombination of a small fraction of the electrons that constitute I_C, in the base. ($n = 1$)
4: Generation of holes in the collector-base space charge region (varies with collector/base bias).

For modern BJTs used in VLSI circuits, mechanism 1 is the dominant one with mechanism 2 evident at very low current levels.

Current gain

Dividing equations (4) and (5),

$$h_{FE}\left(\equiv \beta\right) = \frac{I_C}{I_B} = \frac{D_e}{D_h}\frac{n_{ib}^2}{n_{ie}^2}\frac{L_h}{W_b}\frac{N_d}{N_a} \qquad (7)$$

Thus $N_d >> N_a$ is required for high current gain (consider above exercise!).