

3.4.2 Bipolar Transistors

Basic Concept and Operation

We are not particularly interested in **bipolar transistors** and therefore will treat them only cursory.

Essentially, we have two junctions diodes switched in series (sharing one doped piece of **Si**), i.e. a **npn** or a **pnp** configuration, with the *added condition* that the middle piece (the **base**) is *very thin*. "Very thin" means that the base width d_{base} is much smaller than the diffusion length L .

The other two doped regions are called the **emitter** and the **collector**.

For transistor operation, we switch the emitter - base (**EB**) diode in forward direction, and the base - collector (**BC**) diode in reverse direction as shown below.

This will give us a large forward current and a small reverse current - which we will simply neglect at present - in the **EB** diode, exactly as described for [diodes](#). What happens in the **BC** diode is more complicated and constitutes the principle of the transistor.

In other words, in a **pnp** transistor, we are injecting a lot of holes into the base from the emitter side and a lot of electrons into the emitter from the base side; and vice versa in a **npn**- transistor. Lets look at the two **EB** current components more closely transistor:

For the *hole* forward current, [we have](#) in the simplest approximation (ideal diode, no reverse current; no **SCR** contribution):

$$j_{\text{hole}}(U) = \frac{e \cdot L \cdot n_i^2}{\tau \cdot N_{\text{Acc}}} \cdot \exp - \frac{e \cdot U}{kT}$$

and the relevant quantities refer to the *hole* properties in the *n - doped base* and the doping level N_{Acc} in the *p - doped emitter*. For the electron forward current we have accordingly:

$$j_{\text{electron}}(U) = \frac{e \cdot L \cdot n_i^2}{\tau \cdot N_{\text{Don}}} \cdot \exp - \frac{e \cdot U}{kT}$$

and the relevant quantities refer to the *electron* properties in the *p - doped emitter* and the doping level N_{Don} in the *n - doped base*.

The relation between these currents, i.e. $j_{\text{hole}}/j_{\text{electron}}$, which we call the **injection ratio** κ , then is given by

$$\kappa = \frac{\frac{L_h}{\tau_h \cdot N_{\text{Ac}}}}{\frac{L_e}{\tau_e \cdot N_{\text{Don}}}} = \frac{N_{\text{Ac}}}{N_{\text{Don}}}$$

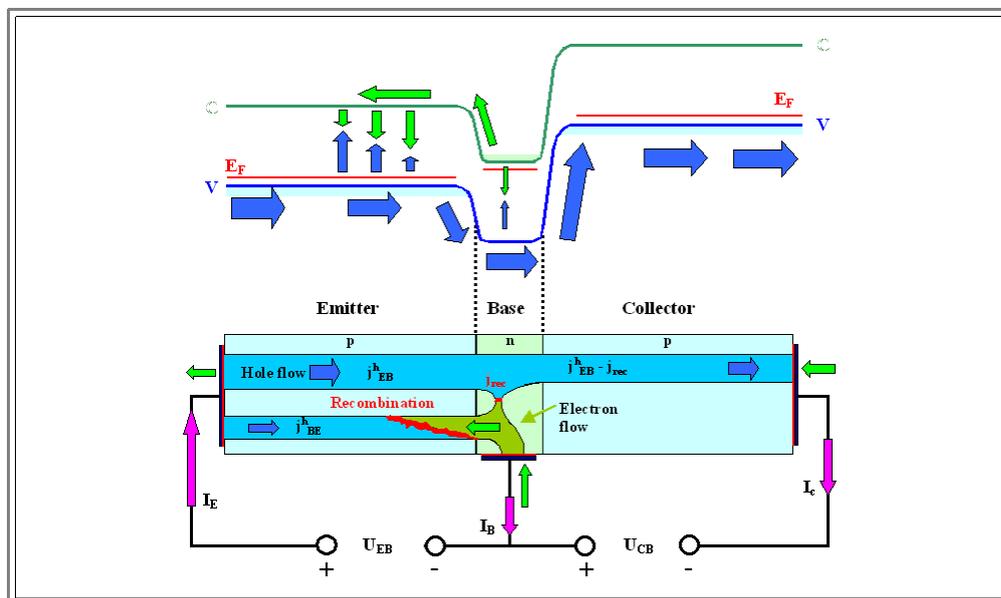
Always assuming that electrons and holes have identical lifetimes and diffusion lengths.

The *injection ratio* κ is a prime quantity. We will encounter it again when we discuss for optoelectronic devices!

For only one diode, that would be all. But we have a second diode right after the first one. The holes injected into the base from the emitter, will diffuse around in the base and long before the die a natural death by recombination, they will have reached the other side of the base

There they encounter the electrical field of the base-collector **SCR** which will sweep them rapidly towards the collector region where they become majority carriers. In other words, we have a large hole component in the reverse current of the **BC** diode (and the normal small electron component which we neglect).

A band diagram and the flow of carriers is shown schematically below in a band diagram and a current and carrier flow diagram.



Lets discuss the various currents going from left to right.

- At the **emitter contact**, we have two hole currents, j_{EB}^h and j_{BE}^h that are converted to electron currents that carry a negative charge away from the emitter. The technical current (mauve arrows) flows in the opposite direction by convention.

For the **base current** two major components are important:

- An electron current j_B^e , directly taken from the **base contact**, most of which is injected into the emitter. The electrons are minority carriers there and recombine within a distance L with holes, causing the small hole current component shown at the emitter contact.
- An internal recombination current j_{rec} caused by the few holes injected into the base from the emitter that recombine in the base region with electrons, and which reduces j_B^e somewhat. This gives us

$$j_{BE}^h = j_B^e - j_{rec}$$

- Since all holes would recombine within L , we may approximate the fraction recombining in the base by

$$j_{rec} = j_{EB}^h \cdot \frac{d_{base}}{L}$$

Last, the current at the **collector contact** is the **hole** current $j_{EB}^h - j_{rec}$ which will be converted into an **electron** current at the contact.

The external terminal **currents** I_E, I_B , and I_C thus are related by the simple equation

$$I_E = I_B + I_C$$

A bipolar transistor, as we know, is a **current amplifier**. In black box terms this means that a small current at the the **input** causes a large current at the **output**.

- The input current is I_B , the output current I_C . This gives us a current amplification factor γ of

$$\gamma = \frac{I_C}{I_B} = \frac{I_E}{I_B} - 1$$

- Lets neglect the small recombination current in the base for a minute. The emitter current (density) then is simply the total current through a **pn-junction**, i.e. in the terminology from the picture $j_E = j_{BE}^h + j_B^e$, while the base current is just the electron component j_B^e .
- This gives us for I_E/I_B and finally for γ :

$$\frac{I_E}{I_B} = \frac{j_B E^h + j_B E^e}{j_B E^e} = \kappa + 1$$

$$\gamma = \frac{I_E}{I_B} - 1 = \kappa + 1 - 1 = \kappa = \frac{N_{Ac}}{N_{Don}}$$

Now *this is really easy!* We will obtain a large current amplification (easily **100** or more), if we use a lightly doped base and a heavily doped emitter. And since we can use large base - collector voltages, we can get heavy power amplification, too.

- Making better approximations is not difficult either. Allowing somewhat different properties of electrons and holes and a finite recombination current in the base, we get

$$\gamma = \frac{\frac{L_h}{\tau_h \cdot N_{Ac}}}{\frac{L_e}{\tau_e \cdot N_{Don}}} \cdot \left(\frac{1 - \frac{d_{base}}{L}}{1} \right) \approx \frac{N_{Don}}{N_{DAc}} \cdot \left(\frac{1 - \frac{d_{base}}{L}}{1} \right)$$

- The approximation again is for identical life times and diffusion lengths.

Obviously, you want to make the base width d_{base} small, *and* keep L large.

Real Bipolar Transistors

Real bipolar transistors, especially the very small ones in integrated circuits, are complicated affairs; for a quick glance on [how they are made and what the pnp or npn part looks like](#), use the link.

Otherwise, everything mentioned in the context of [real diodes](#) applies to bipolar transistors just as well. And there are, of course, some special topics, too.

- But we will *not* discuss this any further, except to point out that the "small device" topic introduced for a simple p-n-junction now becomes a new quality:
- Besides the length of the emitter and collector part which are influencing currents in the way discussed, we now have the **width of the base region** d_{base} which introduces a new quality with respect to device dimensions and device performance.
- The numerical value of d_{base} (or better, the relation d_{base}/L), does not just change the device properties somewhat, but is the *crucial* parameter that brings the device into existence. A transistor with a base width of several **100 μm** simply is not a transistor, neither are two individual diodes soldered together.

The immediate and unavoidable consequence is that at this point of making semiconductor devices, *we have to make things real small.*

- Microtechnology - typical lengths around or below **1 μm** (at least in one dimension) - is mandatory. There are no big transistors in more than two dimensions.
- Understanding *microscopic* properties of materials (demanding quantum theory, statistical thermodynamics, and so on) becomes mandatory. *Materials Science and Engineering was born.*