Lecture 21: BJT (Bipolar Junction Transistors)

Prof. J. S. Smith

Context

In Monday’s lecture, we discussed minority injection in forward biased PN junctions. Today we will discuss three terminal devices which use this effect for amplification, called:

- BJT (Bipolar Junction Transistors)
Today’s lecture will cover chapter 7, Bipolar Junction Transistors (BJT’s).
Next, we will start looking at amplifiers, chapter 8 in the text.

- Review of minority current injection in PN Diode
- The BJT (7.1)
- BJT Physics (7.2)
- BJT Ebers-Moll Equations (7.3)
- BJT Small-Signal Model
As you increase the forward bias on a PN junction, the barrier keeping the holes from diffusing into the N side, and keeping the electrons from diffusing to the P side, is reduced.

As the barrier height decreases, the diffusion of carriers across the barrier increases exponentially.

Forward bias → reduced barrier height, so more minority carriers on both sides.
Forward bias → Increased population of minority carriers

The minority carrier concentrations at the edges of the depletion region will be given by:

\[
\begin{align*}
    p_n(x = x_n) &= N_A e^{-q(\phi_B - V_D)/kT} \\
    n_p(x = -x_p) &= N_D e^{-q(\phi_B - V_D)/kT}
\end{align*}
\]

Note: \( N_A \) and \( N_D \) are the majority carrier concentrations on the other side of the junction, with the assumption that \( p_n << N_D \) and \( n_p << N_A \).

Also note: This neglects net recombination inside the depletion region.

---

Quasi-Neutrality

- Since the regions outside the depletion regions are going to be very close to electrically neutral (called quasi-neutrality) the number of majority carriers will increase slightly as well (slightly because there are usually many more majority carriers than minority carriers anyway).
Diffuse and Recombine

- Once the minority carriers have been injected across the depletion region, they will diffuse, and they will recombine.
- They will recombine, because now $pn > n_i^2$ and since recombination is proportional to $pn$, it will now cause carriers to recombine at a rate faster than they are generated.
- They will also diffuse into the other side, because there are more of them (the minority carriers) at the edge of the depletion region than there are further in.

Excess injected minority carriers

Once the minority carriers are injected into the other side of the junction, the minority carrier concentration in the bulk region for forward bias is a decaying exponential.
Ambipolar diffusion

- The diffusion constant for minority carriers is complicated by the fact that the minority carriers are dragged on (or pulled!) by the majority carriers, in a mechanism called ambipolar diffusion.
- There is a small $E$ field which keeps the excess holes and electrons together.
- The ambipolar diffusion constant for minority carriers comes out to be the diffusion constant for the other, majority carriers, in the cases where the majority carriers greatly outnumber the minority carriers. In other words, minority holes diffuse with $D_n$ and electrons with $D_p$.
- The minority diffusion length is a function of the ambipolar diffusion constant for the minority carriers, as well as the variability in the lifetime, etc. We will just take it to be a number!

Steady-State Concentrations

In silicon, it is easy for the diffusion lengths to be much longer than the distance to the contacts, if none of the diffusing holes and electrons recombine until they get to the contacts → get a linear concentration gradient.
Diode Current Densities

\[ \frac{dn_p}{dx}(x) \approx n_{p0} e^{qV_p/kT} - n_{p0} \frac{qV_p}{-x_p - (-W_p)} \]

\[ n_{p0} = \frac{n_i^2}{N_x} \]

Under this linear approximation, we can calculate the current due to the holes diffusing into the N side, and the electrons diffusing into the P side. Notice that we are using the electron diffusion constant for the minority holes, and visa versa.

\[ J_{n,\text{diff}} = qD_n \frac{dn_p}{dx} \bigg|_{x=x_a} \approx q \frac{D}{W_p} n_{p0} \left( e^{qV_p/kT} - 1 \right) \]

\[ J_{p,\text{diff}} = -qD_p \frac{dp_n}{dx} \bigg|_{x=x_a} \approx -q \frac{D}{W_p} n_{p0} \left( 1 - e^{qV_p/kT} \right) \]

Notice that the currents due to the injected Minority carriers on Each side are not equal, But are proportional To the carrier concentrations.

Bipolar Junction Transistor (BJT)

- A BJT is physically just two back to back PN diodes, with three contacts, but the current between the emitter and the collector is a minority carrier current in the base.
- Essentially, a forward biased diode is used to create a minority current, most of which then goes all the way across to the depletion region of another, reverse biased diode.
- The geometry can be such that almost all the current goes across to the second diode, so that the controlling electrode doesn’t have to supply much of the current, maybe 1:100 to 1:400
A BJT transistor consists of a pair of diodes which have their junctions very close together, so that the minority currents from one junction go through the thin middle layer to the other junction.

They are called PNP or NPN transistors by the layers they are made up of.

**Currents in a BJT**

- Usually, a BJT is operated in a mode where one of the junctions is forward biased, and the other is reverse biased.
- The reverse biased diode injects minority carriers into the middle layer.
- The minority carriers are then swept through the reversed biased junction.
Currents in the BJT

- A BJT is ordinarily designed so that the minority carrier injection into the base is far larger than the minority carrier injection into the emitter.
- It is also ordinarily designed such that almost all the minority carriers injected into the base make it all the way across to the collector.

Band edge diagram

- The band edge diagram for an NPN transistor in operation.

Diagram showing the band edge with N (heavily doped) and P (lightly doped) regions.
Current controlled

- So the current is determined by the minority current across the emitter-base junction:
  \[ I_C \approx I_S e^{\frac{qV_{BE}}{kT}} \]

- But since the majority of the minority current goes right through the base to the collector:
  \[ I_C \approx -I_E \]

- And so the amount of current that must be supplied by the base is small compared to the current controlled:
  \[ I_C \gg I_E \]

Actual BJT Cross Section

- Vertical npn sandwich (pnp is usually a lateral structure)
- n+ buried layout is a low resistance contact to collector
- Base width determined by vertical distance between emitter diffusion and base diffusion
BJT Layout

- Emitter area most important layout parameter
- Multi-finger device also possible for reduced base resistance

BJT Schematic Symbol

- Collector current is controlled by base current linearly, a typical value would be $\beta = 100$, because only one in 100 electrons would stop in the base instead of making it across to the collector
- Collector is controlled by base-emitter voltage exponentially
**Simple NPN BJT model**

- A simple model for a NPN BJT:

\[ I_C(t) = \beta I_B(t) \]

Real diode, not an ideal diode

**BJT Collector Characteristic**

- Ground emitter
- Fix \( V_{CE} \)
- Drive base with fixed current \( I_B \)
- Measure the collector current

\[ I_C = I_C(I_B, V_{CE}) \]
BJT operating modes

- **Forward active**
  - Emitter-Base forward biased
  - Base-Collector reverse biased

- **Saturation**
  - Both junctions are forward biased

- **Reverse active**
  - Emitter-Base reverse biased
  - Base-Collector forward biased
  - Transistor operation is poor in this direction, because $\beta$ is low: lighter doping of the layer designed to be the collector means that there is a lot of minority carrier injection out of the Base.

---

**Collector Characteristics ($I_B$)**

- **Forward Active Region** (Very High Output Resistance)
- **Saturation Region** (Low Output Resistance)
- **Reverse Active Region** (poor Transistor)
- **Breakdown**
- **Linear Increase**
Minority carriers in base form a uniform diffusion current. Since emitter doping is higher, this current swamps out the current portion due to the minority carriers injected from base.
BJT Currents

Collector current is nearly identical to the (magnitude) of the emitter current … define

\[ I_C = -\alpha_F I_E \quad \alpha_F = .999 \]

Kirchhoff:

\[ -I_E = I_C + I_B \]

DC Current Gain:

\[ I_C = -\alpha_F I_E = \alpha_F (I_B + I_C) \]

\[ I_C = \frac{\alpha_F}{1 - \alpha_F} I_B = \beta_F I_B \quad \beta_F = \frac{\alpha_F}{1 - \alpha_F} = .999 \approx 999 \]

Origin of \( \alpha_F \)

Base-emitter junction: some reverse injection of holes into the emitter \( \rightarrow \) base current isn’t zero

Typical: \( \alpha_F \approx .99 \quad \beta_F \approx 100 \)
Collector Current

Diffusion of electrons across base results in

\[ J_{n}^{\text{diff}} = qD_n \frac{dn_p}{dx} = \left( \frac{qD_n n_{PB0}}{W_B} \right) \frac{qV_{BE}}{kT} \]

\[ I_S = \left( \frac{qD_n n_{PB0} A_E}{W_B} \right) \]

\[ I_C = I_S e^{\frac{qV_{BE}}{kT}} \]

Base Current

In silicon, recombination of carriers in the base can usually be neglected, so the base current is mostly due to minority injection into the emitter. Diffusion of holes across emitter results in

\[ J_{p}^{\text{diff}} = -qD_p \frac{dp_{nE}}{dx} = \left( \frac{qD_p p_{nE0}}{W_E} \right) \left( \frac{qV_{BE}}{e^{kT}} - 1 \right) \]

\[ I_B = \left( \frac{qD_p p_{nE0} A_E}{W_E} \right) \left( \frac{qV_{BE}}{e^{kT}} - 1 \right) \]
Current Gain

\[
\beta = \frac{I_C}{I_B} = \frac{\frac{qD_n n_{pB} A_E}{W_B}}{\frac{qD_p p_{nE} A_E}{W_E}} = \left( \frac{D_n}{D_p} \right) \left( \frac{n_{pB}}{p_{nE}} \right) \left( \frac{W_E}{W_B} \right)
\]

Minimize base width

\[
\left( \frac{n_{pB}}{p_{nE}} \right) = \frac{n_i^2}{N_{A,B}} - \frac{n_i^2}{N_{D,E}}
\]

Maximize doping in emitter

Ebers-Moll Equations

Exp. 6: measure E-M parameters

Derivation: Write emitter and collector currents in terms of internal currents at two junctions

\[
I_E = -I_{ES} \left( e^{V_{BE}/V_{th}} - 1 \right) + \alpha_R I_{CS} \left( e^{V_{EC}/V_{th}} - 1 \right)
\]

\[
I_C = \alpha_F I_{ES} \left( e^{V_{BE}/V_{th}} - 1 \right) - I_{CS} \left( e^{V_{EC}/V_{th}} - 1 \right)
\]

\[
\alpha_F I_{ES} = \alpha_R I_{CS}
\]
Ebers-Moll Equivalent Circuit

Building blocks: diodes and $I$-controlled $I$ sources

Diode Currents:
- $I_F = I_G(E^V_{FE} / V_{FE} - 1)$
- $I_R = I_G(E^V_{RE} / V_{RE} - 1)$