Section C2: BJT Structure and Operational Modes

Recall that the semiconductor diode is simply a pn junction. Depending on how the junction is biased, current may easily flow between the diode terminals (forward bias, \( v_D > V_{ON} \)) or the current is essentially zero (reverse bias, \( v_D < V_{ON} \)). Keeping this in mind, we are going to introduce the bipolar junction transistor as nothing more than two diodes placed back to back with the center region common to both diodes. The BJT comes in two “flavors” – the npn and the pnp – the schematic representation and circuit symbols for which are illustrated in Figure 4.1(a) of your text and reproduced below.

As you can see, the diode “sandwich” that creates the BJT results in three distinct regions, each of which is connected to the outside world (hence the three-terminal designation) and are labeled emitter (E), base (B), and collector (C). Each of these regions has a specific purpose and general design considerations as follows:

- The **emitter** is a medium-sized, heavily doped region whose primary purpose is to **inject** its majority carriers (electrons for n-type, holes for p-type), through the base and into the collector.
- The **collector** is a thick, lightly doped region designed to **collect** the majority carriers injected from the emitter.
- The **base** is a thin, medium doped layer whose primary purpose is to provide the **control** of the current between the emitter and collector.

Notice that the base is of the opposite material type from the collector and emitter. This property is what allows the current control of the BJT through the injection of minority carriers into the base region (Note: What is majority in the emitter is minority in the base.). The involvement of both carrier types, electrons and holes, in transistor operation is what gives the **bipolar** junction transistor its name. As a preview of coming attractions, this may be contrasted with the field-effect transistor (FET), which is a **unipolar** device in that it utilizes a single carrier type – either electrons or holes – for operation.
The operational mode of the BJT depends on how the junctions between the regions are biased. Since there are two junctions, the emitter-base junction (EBJ) and the collector-base junction (CBJ), and each of these junctions may be either forward- or reverse-biased, there are four possible modes of operation.

<table>
<thead>
<tr>
<th>Mode</th>
<th>EBJ Bias</th>
<th>CBJ Bias</th>
</tr>
</thead>
<tbody>
<tr>
<td>Active (or Normal Active)</td>
<td>Forward</td>
<td>Reverse</td>
</tr>
<tr>
<td>Cutoff</td>
<td>Reverse</td>
<td>Reverse</td>
</tr>
<tr>
<td>Saturation</td>
<td>Forward</td>
<td>Forward</td>
</tr>
<tr>
<td>Inverted Active</td>
<td>Reverse</td>
<td>Forward</td>
</tr>
</tbody>
</table>

In our studies, we are going to be looking at three of these modes:

- The **normal active** mode is used when the BJT is employed as an amplifier in analog applications and is what we will be concentrating on in this course sequence.
- A combination of **cutoff** and **saturation** modes are used in logic applications and other requirements that involve device switching that may be addressed in other courses or electives you may choose to investigate. For our purposes (analog small-signal amplification), cutoff and saturation will be discussed in terms of introduced nonlinearities – i.e., regions we want to stay away from!

**NOTE:** In the following discussions, we will be concentrating on the npn BJT. The concepts and derivations for the pnp BJT are complimentary, which simply means that the device behaves the same as the npn with the following changes:

- The material types are interchanged (i.e., n-to-p and p-to-n).
- The majority and minority carrier types are interchanged (npn: electrons are majority carriers and holes are minority carriers; pnp: holes are majority carriers and electrons are minority carriers).
- The bias voltage polarities are reversed and the current directions are switched.

**The npn BJT in the Normal Active Mode**

As indicated in the table above, the normal active mode occurs when the EBJ is forward biased and the CBJ is reverse biased. Recall from our discussion of the pn junction that a forward biased junction has a lowered potential barrier, while the barrier for a reverse biased junction is much
larger than equilibrium. This is sketched in the figure below, where the (a) shows the potential hill diagram associated with an unbiased BJT and (b) is representative of the potential hill diagram of a BJT in normal active mode. In both of these figures, the junction depletion regions are shown in yellow. For the EBJ, the depletion region becomes smaller for the applied forward bias and for the CBJ, the depletion region is larger with reverse bias.

Note that the unbiased depletion regions are representative of the relative doping between the two regions of a junction. Recall that the lower the doping, the greater the extent of the depletion region in the material. So...the figure below is in agreement with the earlier discussion of BJT regions; i.e., the doping in the emitter is greater than the doping in the base, and the doping in the base is much greater than the doping in the collector. To lose the multitudinous words and slap this stuff in representative form:

\[ N_{DE} > N_{AB}; \quad N_{AB} >> N_{DC}. \]

Can you see it?

The current flow of the npn BJT is shown in Figure 4.2(a) and is reproduced below in a modified form to agree with the figure above. Keep in mind that we’re talking about an npn BJT so the majority of the current is going to be due to electrons. However, conventional current direction assumes positive charge movement and that’s why the \( i_E \) and \( i_C \) current directions look counterintuitive (The carriers move in
opposite directions, but they also have opposite charge – remember from our current flow discussions in Section A5?).

![Diagram of a transistor](image)

**NOTE:** In the following, only the diffusion current component is considered. Under normal active mode bias conditions, the drift current due to thermally generated minority carriers is essentially negligible.

Components of, and contributions to, the total device current may be described as:

- Under forward bias ($V_{BE} > 0$), majority carriers (electrons) are “pushed away from” the negative terminal of $V_{BE}$ over the lowered potential barrier and injected into the base region.
- Since the EBJ is forward biased, the depletion region is relatively narrow and most of the carriers injected from the emitter will diffuse across the EBJ into the base region.
- Once in the base region, there are two possibilities for the injected majority carriers.
  1. They may make it through the base region into the CBJ depletion region, attracted by the positive potential of the collector connection. Since the base region is very thin, this is indeed what happens to most of the carriers injected from the emitter.
  2. A small fraction of the injected majority carriers are lost to the base. There are two ways to look at this: the carriers are attracted to the positive terminal of $V_{BE}$ and are lost through diffusion or, the carriers injected into the base from the emitter (electrons) are annihilated by the free carriers in the base (holes) through recombination. However it
is more comfortable for you to look at it, this portion of the majority carriers are lost and no longer participate in the current flow of the device. This is where the control comes in – manipulation of this “loss” effect determines the amount of charge (and therefore the current) that makes it from the emitter to the collector!

- The majority carriers that make it to the CBJ depletion region (option 1 above) are rapidly swept through the depletion region and into the collector contact. This may also be looked at as if the electrons “fall down” the very large potential hill created by the CBJ reverse bias as shown in Figure 4.1(b) of your text or as sketched in the figure above.

Looking at the above figure, and keeping in mind that even though electrons and holes are shown moving in opposite directions, they both contribute to positive current (defined in the figure as into the device) since they carry opposite charge, we can define the currents for the npn BJT as follows:

- $I_{CBO}$: leakage current across the CBJ, due to reverse bias of the junction.
- $i_B$: base current due to positive base potential of $V_{BE}$ (this accounts for the holes that are injected across the forward biased EBJ and the recombination losses in option 2 above).
- $i_C$: collector current due to majority carriers that made it through the base region.
- $i_E$: emitter current that can be expressed by KCL at the top node as:

$$i_E = i_C + i_B \quad \text{(Equation 4.1)}$$

Since the device currents depend on several factors, AND we’re going to have to have models for the transistor in order to design and analyze circuits, the interrelationships that occur in the transistor will be modeled by dependent sources. To establish these dependencies for device models, several gain constants are defined.

The first of these gain constants, the common-base current gain $\alpha$, is defined as the rate of change of the collector current with the change in the emitter current with the voltage between the collector and base held constant, or

$$\alpha = \frac{\Delta i_C}{\Delta i_E} \bigg|_{V_{CB}=\text{const}} \quad \text{(Equation 4.2)}$$
The internal device current flows, as detailed in Figure 4.2(a) or the figure above, are simplified and redefined in Figure 4.2(b) (presented to the right) in terms of the gain term $\alpha$.

Ideally, all of the emitter current would be collector current and $\alpha$ would be unity (=1). Equation 4.2 allows us to define a proportional relationship between the emitter and collector currents in terms of this gain or, $i_c \propto \alpha i_E$. From Figure 4.2(b), it may be seen that the collector current is the sum of this contribution and the leakage current across the reverse biased CBJ, or

$$i_c = \alpha i_E + I_{CBO} \approx \alpha i_E \quad \text{(Equation 4.3, modified)}$$

Practically, the values for $\alpha$ usually range between 0.9 and 0.999. By combining Equations 4.1 and 4.3, solving for $i_B$, and approximating $1/\alpha$ as equal to one, an approximation of the base current may be expressed as

$$i_B \approx \left(\frac{1-\alpha}{\alpha}\right)i_c - I_{CBO} \approx \left(\frac{1-\alpha}{\alpha}\right)i_c \quad \text{(Equation 4.5, modified)}$$

The coefficient of $i_C$ in the simplification (also found by taking the partial derivative of the first expression with respect to $i_C$), is the second gain constant we will be discussing. By definition, the large signal amplification factor also known as the direct current amplification factor, is given the symbol $\beta$ and is defined as the rate of change of the collector current with the change in the base current, or:

$$\beta = \frac{\Delta i_C}{\Delta i_B} = \frac{\alpha}{1-\alpha} \quad \text{(Equation 4.6)}$$

Rewriting the simplified version of Equation 4.5 and making the assumption that $\alpha$ is approximately equal to one in Equation 4.3, we can express the relationships between the terminal currents as:

$$i_C \approx \beta i_B \approx i_E \quad \text{(Equations 4.8 & 4.9)}$$
Although there is some inherent inaccuracy in the above approximations, in most practical circumstances (i.e., if $\beta \gg 1$ or, equivalently, $\alpha < 1$) these simple relationships will be more than adequate to effectively design and analyze amplifier circuits.

However and, although your author does discuss it too much, if $\beta$ is not much greater than one we must use

$$\alpha = \frac{\beta}{\beta + 1}; \quad i_B = \frac{i_C}{\beta}; \quad i_C = \beta i_B = \alpha i_E; \quad i_E = \left(\frac{\beta + 1}{\beta}\right)i_C = \frac{i_C}{\alpha} = (\beta + 1)i_B.$$ 

Nothing has changed; these are the real relationships between the gain constants and the device currents.

Keep this in mind though – no approximation is 100% valid!

Usually, we will be dealing with $\beta$ on the order of 100 or greater. For these cases, the simplified approximations are sufficient and any discrepancy is usually lost in the round off. Prove it to yourself!