Lecture 10: Bipolar Junction Transistor Construction. NPN Physical Operation.

For the remainder of this semester we will be studying transistors and transistor circuits.

The transistor is a three terminal device. It is probably only the second such device you've encountered in electrical engineering, after the op amp. Roughly speaking, the transistor acts an electronic valve: the node current or voltage at one terminal controls the current entering the second terminal and exiting the third.

It would be difficult to overstate the importance of the transistor to electronics. Some of its uses are:
- Digital logic
- Memory circuits
- Amplifiers
- Electronic switches.

The two basic families of transistors are bipolar junction transistors (BJTs) and field effect transistors (FETs). We'll start by discussing BJTs for approximately 13 lectures followed by FETs for the remaining 12 lectures. BJTs are covered in Chapter 5 of your text.

BJT Physical Structure

BJTs are formed from three doped regions on a silicon crystal. These can either be *npn* doped regions or *pnp*. A simplified sketch of an *npn* transistor is:
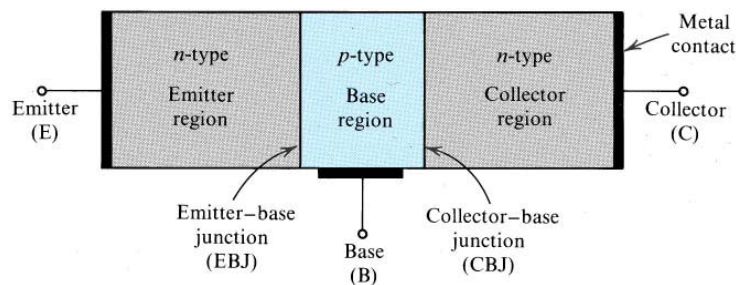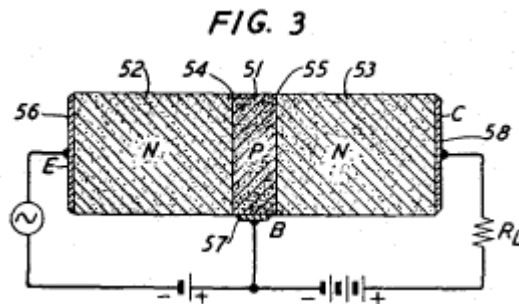

(Fig. 5.1)

As can be seen, the BJT is formed from two back-to-back *pn junctions*:

- Emitter-base junction (EBJ)
- Collector-base junction (CBJ).

This specific way of drawing the BJT has been around from the very beginning of these transistors. This figure:

is from William Shockley's U.S. patent 2,569,347 issued in 1951.

BJTs can also be fabricated from two *p*-type regions and one *n*-type. This is called a *pnp* transistor:
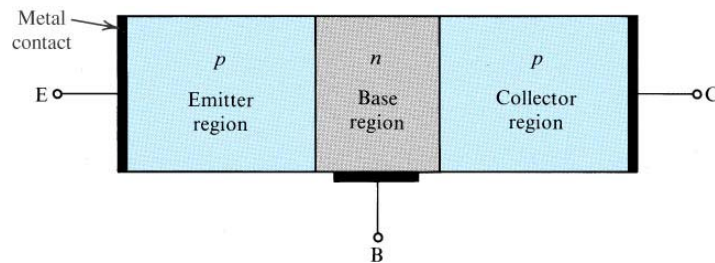
(Fig. 5.2)

While the BJT might appear to be symmetrical by looking at Fig. 5.1, the actual devices are not.

For example, the cross section below of an *npn* transistor clearly shows that the EBJ and CBJ, for example, have very differently sized surface contact areas, which will greatly change their relative behaviors.
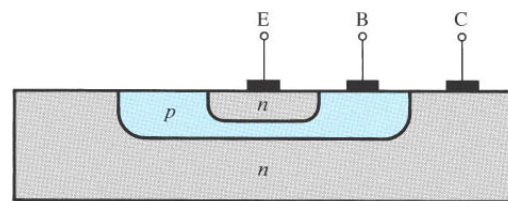
(Fig. 5.6)

There are four basic modes of operation for a BJT depending on the states of the two *pn* junctions of the transistor:

Mode	Emitter-Base Jct.	Collector-Base Jct.
Cutoff	Reverse	Reverse
Active	Forward	Reverse
Saturation	Forward	Forward
Reverse Active	Reverse	Forward

In digital logic applications, the transistor switches between the cutoff and saturation modes. As a linear amplifier in a communication circuit, the transistor would operate in the active mode.

Because of the asymmetrical physical construction, the reverse active mode is not the same as interchanging the collector and emitter leads.

The states of the two *pn* junctions can be altered by the external circuitry connected to the transistor. This is called biasing the transistor.

NPN Transistor in the Active Mode

We'll begin the discussion of the BJT physical operation by considering an *npn* transistor in the active mode. To bias it in the active mode, we need to forward bias the EBJ and reverse bias

the CBJ (notice that the emitter and collectors have swapped positions from Fig. 5.1 shown earlier):
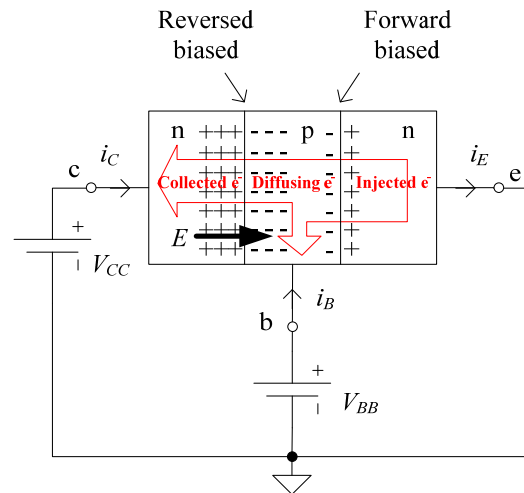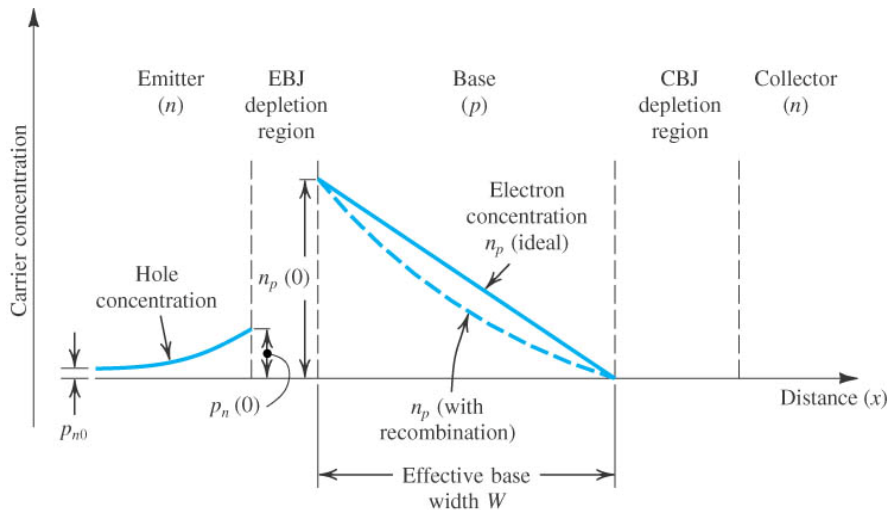

(Fig. 1)

The overall objective of this circuit is to create a current flowing from the collector to the emitter terminals in the transistor that is "controlled," so to speak, by the base voltage V_{BB}.

How does this transistor operate in this circuit?

- Because of the forward bias on the EBJ, charges can flow across this junction giving rise to i_E. This current is primarily electrons that are injected from n to p.

- The electrons injected in the base diffuse across the **thin** base region towards the collector. Some of the e^- recombine in the base, but this region is manufactured to be thin and lightly doped compared to the emitter so this recombination is kept small. Otherwise, the BJT would just operate as two back-to-back diodes and no current would flow.

A representative "minority carrier" concentration profile is shown below in Fig. 5.4. (Note that C and E are switched with reference to Fig. 1 above.)

(Fig. 5.4)

- The e^- that reach the reverse-biased CBJ encounter a large electric field. This **E** sweeps them into the collector forming the collector current i_C as shown in Fig. 1 above.

- A small base current i_B is present largely due to recombination in the base with the small amount of injected holes from the base to the emitter. This is an important current, though.

Discussion About BJT in the Active Mode

1. The proportion of electrons from the emitter that "make it" to the base is called the collector efficiency, α:

$$\alpha \equiv \frac{i_C}{i_E} \quad \text{or} \quad \boxed{i_C = \alpha i_E} \qquad (5.16),(1)$$

Typically it has values of near 0.99. Note that α is called the common-base current gain in the text.

From KCL in the circuit of Fig. 1:

$$i_E = i_B + i_C \qquad (5.13),(2)$$

or

$$i_B = i_E - i_C = (1-\alpha)i_E \qquad (3)$$

We can deduce from (3) that with $\alpha \approx 1$, then i_B will be much, much smaller than i_E.

2. The ratio of collector current to base current is called the current gain, β:

$$\beta \equiv \frac{i_C}{i_B} \quad \text{or} \quad \boxed{i_C = \beta i_B} \qquad (5.10),(4)$$

3. Dividing (1) by (4) we find that

$$\frac{i_C}{i_B} = \frac{\alpha}{1-\alpha}$$

Equating this to (4) we find

$$\boxed{\beta = \frac{\alpha}{1-\alpha}} \qquad (5.19),(5)$$

and solving this equation for α
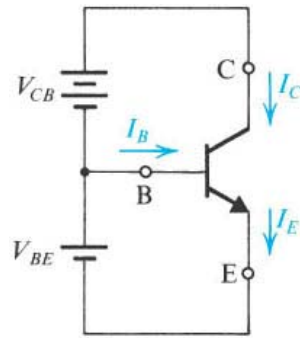
$$\boxed{\alpha = \frac{\beta}{\beta+1}} \qquad (5.17),(6)$$

With $\alpha \approx 0.99 \implies \beta \approx 100$.

4. One can think of the base current in the BJT as controlling the collector current: $i_C = \beta i_B$. Since β is large, then a small change in i_B produces a large change in i_C. If the base were an input signal and the collector the output, then this would be signal amplification! Awesome!

5. The circuit symbol and current conventions for the *npn* BJT are
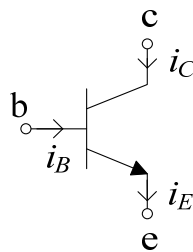
The arrows indicate the assumed directions for positive current for the *npn* BJT. The filled arrow is always located on the emitter and helps us to remember the direction of the emitter current.

6. For biasing in the active mode as shown in Fig. 1, one biasing circuit might be

(Fig. 5.14a)