Power Bipolar Junction Transistors (BJTs)

Zheng Yang

(ERF 3017, email: yangzhen@uic.edu)
Power Bipolar Junction Transistor (BJT) Background

The bipolar junction transistor was invented in 1947 by Shockley, Brattain, and Bardeen. The multifold advantages of replacing vacuum tubes with a solid-state device rapidly promulgated its use in applications. The availability of single-crystal silicon technology with higher purity levels together with the capability to form diffused junctions enabled the fabrication of transistors with high blocking voltage capability. Several decades of effort were undertaken to improve the photolithography techniques and epitaxial deposition capability culminating in the availability of 500V power transistors in the late 1970s. However, during this effort, it was recognized that one of major shortcomings of the power bipolar transistor is its poor current gain. Consequently, these devices were soon eclipsed by the invention and rapid commercialization of the insulated gate bipolar transistor (IGBT). The high input impedance of the IGBT simplified the gate control circuit enabling its integration. Together with the larger on-state current density and improved ruggedness of the IGBT when compared with the power bipolar transistor, this enabled major improvements in the size, cost, and performance of power electronic systems for numerous applications.

The bipolar power transistor is primarily used as a switch. The switching time intervals when turning on and off the structure are therefore of interest from an application’s viewpoint. The presence of minority carrier stored charge within the device is demonstrated to limit the operating frequency of these devices.

Although the power bipolar transistor has been supplanted by the IGBT in all power applications, a good understanding of the physics of current transport in the bipolar transistor is essential for power device specialists. In addition, the IGBT contains a bipolar transistor within its structure. Consequently, the analysis of the IGBT structure requires the application of the physics and concepts of power bipolar junction transistor.
How (conventional low-power) BJT works?

BJTs come in two types, or polarities, known as PNP and NPN based on the doping types of the three main terminal regions. An NPN transistor comprises two semiconductor junctions that share a thin p-doped region, and a PNP transistor comprises two semiconductor junctions that share a thin n-doped region. Charge flow in a BJT is due to diffusion of charge carriers across a junction between two regions of different charge concentrations. The regions of a BJT are called emitter, collector, and base. A discrete transistor has three leads for connection to these regions. Typically, the emitter region is heavily doped compared to the other two layers, whereas the majority charge carrier concentrations in base and collector layers are about the same (collector doping is typically ten times lighter than base doping). By design, most of the BJT collector current is due to the flow of charges injected from a high-concentration emitter into the base where they are minority carriers that diffuse toward the collector, and so BJTs are classified as minority-carrier devices.

(cited from Wikipedia)
In typical operation, the base-emitter junction is forward-biased, which means that the p-doped side of the junction is at a more positive potential than the n-doped side, and the base-collector junction is reverse-biased. In an NPN transistor, when positive bias is applied to the base-emitter junction, the equilibrium is disturbed between the thermally generated carriers and the repelling electric field of the n-doped emitter depletion region. This allows thermally excited electrons to inject from the emitter into the base region. These electrons diffuse through the base from the region of high concentration near the emitter towards the region of low concentration near the collector. The electrons in the base are called minority carriers because the base is doped p-type, which makes holes the majority carrier in the base. To minimize the fraction of carriers that recombine before reaching the collector–base junction, the transistor's base region must be thin enough that carriers can diffuse across it in much less time than the semiconductor's minority-carrier lifetime. In particular, the thickness of the base must be much less than the diffusion length of the electrons. The collector–base junction is reverse-biased, and so little electron injection occurs from the collector to the base, but electrons that diffuse through the base towards the collector are swept into the collector by the electric field in the depletion region of the collector-base junction. The thin shared base and asymmetric collector–emitter doping are what differentiates a bipolar transistor from two separate and oppositely biased diodes connected in series.

(cited from Wikipedia)
Power BJT Structure

The basic structure for an N-P-N bipolar power transistor is illustrated in figure above. In addition to the N⁺ emitter and P-base region for the conventional (low-power) bipolar transistor, the power transistor contains a lightly doped collector (N-drift) region to support high blocking voltages. When a positive bias is applied to the collector terminal of the device, the collector-base junction (J₁) becomes reverse biased and supports the voltage. The blocking voltage capability for the device is determined by the doping concentration and thickness of the N-drift region.
Current flow through the N-P-N bipolar transistor is induced by forward biasing the emitter-base junction ($J_2$) to initiate the injection of electrons. The injected electrons diffuse through the narrow P-base region and arrive at the collector-base junction. When this junction is reverse biased, the electrons collected by the junction are swept through its depletion region producing a collector current. It can be shown that a small base current can produce a large collector current resulting in a substantial current gain. In conjunction with the much larger collector voltage when compared with the base voltage, this produces a large power gain as well.
The bipolar transistor can be switching from its on-state to the blocking state by reversing the bias applied to the base region. The reverse bias not only stops the injection of minority carriers from the emitter-base junction but also removes some of the charge stored in the base region. In the case of the power bipolar transistor, a substantial amount of charge is stored within the thick N-drift region as well. This prolongs the time taken for the transistor to begin supporting voltage, limiting its maximum frequency of operation. The vertical doping profile, extending through one of the N⁺ emitter regions, for the power bipolar transistor is illustrated in the figure above. This profile is achieved by using starting material consisting of an N-type epitaxial layer grown on a heavily doped N-type substrate. The doping concentration and thickness of the epitaxial layer are chosen to obtain the desired voltage-blocking capability for the power bipolar transistor.
The P-type base region is then ion implanted and diffused across the entire active area as shown in the figure above. However, the P-base region must be patterned at the edges of the device to form floating field rings to enhance the breakdown voltage. A mask is now used to define the locations of the N⁺ emitter regions. The doping concentration and depth for the emitter must be carefully chosen to obtain a high-current gain. The doping and thickness of the P-base region, determined by the combination of the P-type and N ion implants, are critical to both the blocking capability and the current gain for the power bipolar transistor. It is customary to interdigitate the emitter and base contacts as illustrated in the figure above because the emitter current tends to concentrate at the periphery of the emitter regions due to a current crowding phenomenon.
Basic Operating Principles of Power BJT

The power bipolar transistor is most often used in the common-emitter circuit configuration as illustrated in the figure above. In this case, the emitter terminal of the transistor is a common element between the input and output side of the circuit. The input side of the circuit is controlled by the drive circuit, which contains two power supplies that can be used to turn on and turn off the transistor. The voltage source $V_{BS1}$ is used to drive the transistor when it is operating in the current conduction mode while the voltage source $V_{BS2}$ is used to turn off the transistor and maintain it in its voltage-blocking mode. The output side of the circuit contains a high-voltage source ($V_{CS}$) that delivers power to the load. The transfer of power from the voltage source to the load is controlled by the drive circuit using switches $S_1$ and $S_2$. 
The bipolar transistor is operated in its current conduction mode by opening switch S\textsubscript{2} and closing switch S\textsubscript{1}. This connects the input voltage source $V_{BS1}$ across the base-emitter terminals of the bipolar power transistor through resistor $R_{B1}$. If the input source voltage exceeds the built-in potential ($V_{bi} \sim 0.7$ V) of the base-emitter junction, it becomes sufficiently forward biased to produce a base current given by

$$i_B = \frac{V_{BS1} - V_{bi}}{R_{B1}}$$
The base current flow is accomplished by the injection of minority carriers (electrons for an N-P-N transistor) from the N⁺ emitter region into the P-base region. These minority carriers diffuse from the emitter-base junction (J₂) through the P-base region and are collected by the reverse-biased base-collector junction (J₁). The electrons captured by the base-collector junction are swept through its depletion region producing a collector current \( i_C \).

When the bipolar transistor is operating in its forward active mode with a reverse-biased base-collector junction, the collector and base currents are related by the **common-emitter current gain** called beta (\( \beta \)):

\[
\beta = \frac{i_C}{i_B}
\]

Based upon the application of Kirchhoff’s current law for the bipolar transistor as a node,

\[
i_E = i_B + i_C
\]

Combining this relationship with that for the common-emitter current gain,

\[
i_E = \left(1 + \frac{i_C}{i_B}\right)i_B = (1 + \beta)i_B
\]
When the bipolar power transistor is operated as a switch, the power delivered to the load is proportional to \((i_C V_{CS})\), while that utilized from the input control circuit is proportional to \((i_B V_{BS1})\). The **common-emitter power gain** is then given by

\[
G_{P,CE} = \frac{i_C V_{CS}}{i_B V_{BS1}} = \beta \left( \frac{V_{CS}}{V_{BS1}} \right)
\]

It is desirable to control a large collector current, which flows through the load, with a small base current to achieve a large power gain. This requires optimization of the bipolar transistor structure to obtain a large common-emitter current gain. The physical parameters within the bipolar transistor that determine the current gain.
The bipolar transistor is sometimes used in the **common-base** circuit configuration as illustrated in the figure above. In this case, the base terminal of the transistor is used as a common element between the input and output side of the circuit. The input side of the circuit is controlled by the drive circuit, which contains two power supplies that can be used to turn on and turn off the transistor. The voltage source $V_{BS1}$ is used to drive the transistor when it is operating on the current conduction mode while the voltage source $V_{BS2}$ is used to turn off the transistor and maintain it in its voltage-blocking mode. The output side of the circuit contains a high-voltage source ($V_{CS}$) that delivers power to the load. The transfer of power from the voltage source to the load is controlled by the drive circuit using switches $S_1$ and $S_2$. 
When the bipolar transistor is operating in its forward active mode with a reverse-biased base-collector junction, the collector and emitter currents are related by the common-base current gain called alpha ($\alpha$):

$$\alpha = \frac{i_C}{i_E}$$

Since only a fraction of the emitter current is delivered to the collector, the common-base current gain is always less than unity. When the bipolar power transistor is operated as a switch in the common-base configuration, the power delivered to the load is proportional to ($i_C V_{CS}$), while that utilized from the input control circuit is proportional to ($i_E V_{BS1}$). The common-base power gain is then given by

$$G_{P, CB} = \left( \frac{i_C V_{CS}}{i_E V_{BS1}} \right) = \alpha \left( \frac{V_{CS}}{V_{BS1}} \right)$$
Basic Operating Principles of Power BJT (cont’d)

The common-base current gain can be related to the common-emitter current gain by using

$$\alpha = \frac{i_C}{i_B + i_C} = \frac{\beta i_B}{i_B + \beta i_B} = \frac{\beta}{1 + \beta}$$

In a similar manner, the common-emitter current gain can be related to the common-base current gain:

$$\beta = \frac{i_C}{i_E - i_C} = \frac{\alpha i_E}{i_E - \alpha i_E} = \frac{\alpha}{1 - \alpha}.$$
The physics of operation of the bipolar power transistor has been described in this chapter. The bipolar transistor is a current-controlled structure whose current gain depends upon the doping concentrations of the emitter and base regions as well as the width of the base region. In the case of the bipolar power transistor, a falloff in the current gain occurs due to high-level injection in the base region, which degrades its power gain. The bipolar power transistor is also prone to failure modes associated with emitter current crowding during the on-state and turn-off modes.

All of these complexities have motivated the development of alternative power switches for applications. The most successful among these alternatives is the IGBT. Its development in the early 1980s has resulted in the complete displacement of the bipolar power transistor leading to its extinction from the power semiconductor landscape. However, the physics of operation for the bipolar power transistor continues to have relevance because it is incorporated within the IGBT structure.

Sections discussed: 7.1, 7.2, and 7.11

Sections not discussed: 7.3-7.10