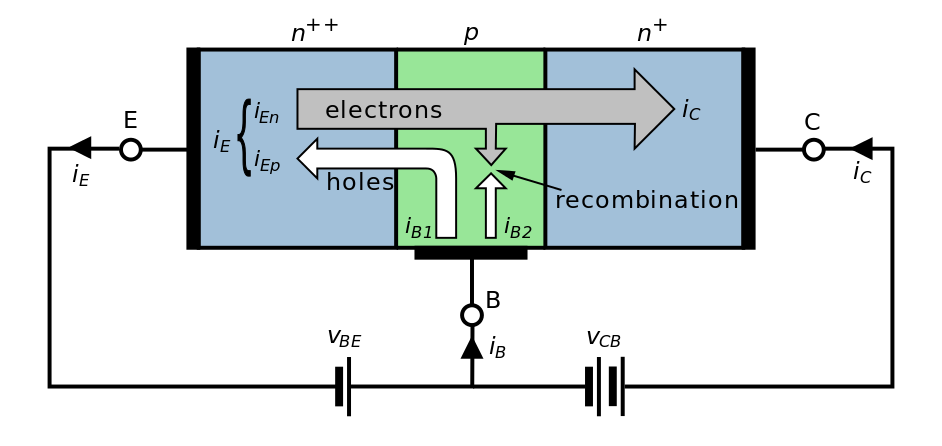
**Bipolar junction transistor**

A **bipolar junction transistor** (**bipolar transistor** or **BJT**) is a type of transistor that uses both electron and [hole](https://en.wikipedia.org/wiki/Electron_hole) charge carriers. For their operation, BJTs use two junctions between two [semiconductor](https://en.wikipedia.org/wiki/Semiconductor) types, n-type and p-type.

The basic function of a BJT is to amplify current. This allows BJTs to be used as [amplifiers](https://en.wikipedia.org/wiki/Electronic_amplifier) or switches, giving them wide applicability in electronic equipment, including computers, televisions, mobile phones, audio amplifiers, industrial control, and radio transmitters.

**Function**

BJTs come in two types, or polarities, known as PNP and NPN based on the [doping](https://en.wikipedia.org/wiki/Doping_%28semiconductors%29) types of the three main terminal regions. An NPN transistor comprises two [semiconductor junctions](https://en.wikipedia.org/wiki/P%E2%80%93n_junction) that share a thin p-doped [anode](https://en.wikipedia.org/wiki/Anode) region, and a PNP transistor comprises two semiconductor junctions that share a thin n-doped [cathode](https://en.wikipedia.org/wiki/Cathode) region.

[](https://en.wikipedia.org/wiki/File:NPN_BJT_Basic_Operation_(Active).svg)

NPN BJT with forward-biased E–B junction and reverse-biased B–C junction

Charge flow in a BJT is due to [diffusion](https://en.wikipedia.org/wiki/Diffusion) of [charge carriers](https://en.wikipedia.org/wiki/Charge_carriers_in_semiconductors) 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.

By design, most of the BJT collector current is due to the flow of charges injected from the high-concentration emitter into the base, that diffuse toward the collector.

In typical operation, the base–emitter [junction](https://en.wikipedia.org/wiki/P-n_junction) is [forward biased](https://en.wikipedia.org/wiki/P-n_junction#Forward_bias) (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](https://en.wikipedia.org/wiki/P-n_junction#Reverse_bias). Electrons [diffuse](https://en.wikipedia.org/wiki/Diffusion) 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*](https://en.wikipedia.org/wiki/Minority_carrier) because the base is doped p-type.

To minimize the percentage of carriers that [recombine](https://en.wikipedia.org/wiki/Carrier_generation_and_recombination) before reaching the collector, the transistor's base region must be thin enough that carriers can diffuse across it before recombining.

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.

**Voltage, current, and charge control**

The collector–emitter current can be viewed as being controlled by the base–emitter current (current control), or by the base–emitter voltage (voltage control).

In [analog circuit](https://en.wikipedia.org/wiki/Analog_circuit) design, the current-control is sometimes used because it is approximately linear. That is, the collector current is approximately β times the base current.

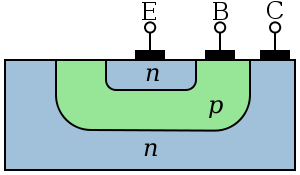
**Transistor parameters: alpha (α) and beta (β)**

The proportion of electrons able to cross the base and reach the collector is a measure of the BJT efficiency. The heavy doping of the emitter region and light doping of the base region causes many more electrons to be injected from the emitter into the base than holes to be injected from the base into the emitter.



where IC is the collector current, IB is the base current, IE is the emitter current.

**Real physical structure**

[](https://en.wikipedia.org/wiki/File:NPN_BJT_(Planar)_Cross-section.svg)

Simplified cross section of a planar *NPN* bipolar junction transistor

A BJT consists of three differently [doped](https://en.wikipedia.org/wiki/Doping_%28semiconductor%29) semiconductor regions: the *emitter* region, the *base* region and the *collector* region. These regions are, respectively, *p* type, *n* type and *p* type in a PNP transistor, and *n* type, *p* type and *n* type in an NPN transistor. Each semiconductor region is connected to a terminal, appropriately labeled: *emitter* (E), *base* (B) and *collector* (C).

The *base* is physically located between the *emitter* and the *collector* and is made from lightly doped material. The collector surrounds the emitter region, making it almost impossible for the electrons injected into the base region to escape without being collected. A cross-section view of a BJT indicates that the collector–base junction has a much larger area than the emitter–base junction.

The bipolar junction transistor is usually not a symmetrical device. This means that interchanging the collector and the emitter makes the transistor leave the forward active mode and start to operate in reverse mode. Because the transistor's internal structure is usually optimized for forward-mode operation, interchanging the collector and the emitter makes the values of α and β in reverse operation much smaller than those in forward operation.

The reason the emitter is heavily doped is to increase the emitter injection efficiency: the ratio of carriers injected by the emitter to those injected by the base. For high current gain, most of the carriers injected into the emitter–base junction must come from the emitter.

**Regions of operation**

Bipolar transistors have four distinct regions of operation, defined by BJT junction biases.

|  |  |  |  |
| --- | --- | --- | --- |
| **Applied voltages** | **B-E**  **Junction bias (NPN)** | **B-C**  **Junction bias (NPN)** | **Mode (NPN)** |
| E < B < C | Forward | Reverse | Forward-active |
| E < B > C | Forward | Forward | Saturation |
| E > B < C | Reverse | Reverse | Cut-off |
| E > B > C | Reverse | Forward | Reverse-active |

**Forward-active (or simply *active*)**

The base–emitter junction is forward biased and the base–collector junction is reverse biased. Most bipolar transistors are designed to afford the greatest common-emitter current gain, βF, in forward-active mode. If this is the case, the collector–emitter current is approximately [proportional](https://en.wikipedia.org/wiki/Proportionality_%28mathematics%29) to the base current, but many times larger, for small base current variations.

* This is the operating condition used to obtain signal amplification.

-------------------------------------------------------------------------------------------------------------------------

**Saturation**

With both junctions forward-biased, a BJT is in saturation mode and facilitates high current conduction from the collector to the emitter (with negatively charged carriers flowing from emitter to collector). This mode corresponds to a logical "on", or a closed switch.

**Cut-off**

In cut-off, biasing conditions opposite of saturation (both junctions reverse biased) are present. There is very little current, which corresponds to a logical "off", or an open switch.

* These two operating conditions are used to obtain a digital switch, capable to give an on/off signal, or to convey a digital 0/1 information.

-------------------------------------------------------------------------------------------------------------------------

**Reverse-active (or *inverse-active* or *inverted*)**

By reversing the biasing conditions of the forward-active region, a bipolar transistor goes into reverse-active mode. In this mode, the emitter and collector regions switch roles. Because most BJTs are designed to maximize current gain in forward-active mode, the βF in inverted mode is several times smaller (2–3 times for the ordinary germanium transistor). This transistor mode is seldom used, usually being considered only for failsafe conditions.

* This is the operating condition used to obtain very precise current control.

-------------------------------------------------------------------------------------------------------------------------