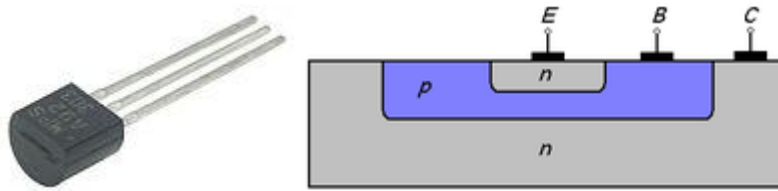


Transistor de unión bipolar



El transistor de unión bipolar (del inglés Bipolar Junction Transistor, o sus siglas BJT) es un dispositivo electrónico de estado sólido consistente en dos uniones PN muy cercanas entre sí, que permite controlar el paso de la corriente a través de sus terminales. La denominación de bipolar se debe a que la conducción tiene lugar gracias al desplazamiento de portadores de dos polaridades (huecos positivos y electrones negativos), y son de gran utilidad en gran número de aplicaciones; pero tienen ciertos inconvenientes, entre ellos su impedancia de entrada bastante baja.

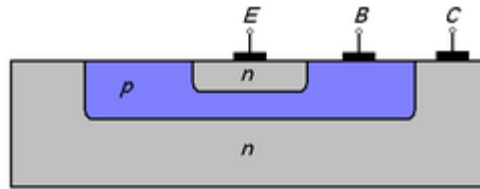
Los transistores bipolares son los transistores más conocidos y se usan generalmente en electrónica analógica aunque también en algunas aplicaciones de electrónica digital, como la tecnología TTL o BICMOS.

Un transistor de unión bipolar está formado por dos Uniones PN en un solo cristal semiconductor, separados por una región muy estrecha. De esta manera quedan formadas tres regiones:

- Emisor, que se diferencia de las otras dos por estar fuertemente dopada, comportándose como un metal. Su nombre se debe a que esta terminal funciona como emisor de portadores de carga.
- Base, la intermedia, muy estrecha, que separa el emisor del colector.
- Colector, de extensión mucho mayor.

En su funcionamiento normal, la unión base-emisor está polarizada en directa, mientras que la base-colector en inversa. Los portadores de carga emitidos por el emisor atraviesan la base, porque es muy angosta, hay poca recombinación de portadores, y la mayoría pasa al colector. El transistor posee tres estados de operación: estado de corte, estado de saturación y estado de actividad.

Estructura



Un transistor de unión bipolar consiste en tres regiones semiconductoras dopadas: la región del emisor, la región de la base y la región del colector. Estas regiones son, respectivamente, tipo P, tipo N y tipo P en un PNP, y tipo N, tipo P, y tipo N en un transistor NPN. Cada región del semiconductor está conectada a un terminal, denominado emisor (E), base (B) o colector (C), según corresponda.

Corte transversal simplificado de un transistor de unión bipolar NPN. Donde se puede apreciar como la unión base-colector es mucho más amplia que la base-emisor.

La base está físicamente localizada entre el emisor y el colector y está compuesta de material semiconductor ligeramente dopado y de alta resistividad. El colector rodea la región del emisor, haciendo casi imposible para los electrones inyectados en la región de la base escapar de ser colectados, lo que hace que el valor resultante de α se acerque mucho hacia la unidad, y por eso, otorgarle al transistor un gran β .

El transistor de unión bipolar, a diferencia de otros transistores, no es usualmente un dispositivo simétrico. Esto significa que intercambiando el colector y el emisor hacen que el transistor deje de funcionar en modo activo y comience a funcionar en modo inverso.

Debido a que la estructura interna del transistor está usualmente optimizada para funcionar en modo activo, intercambiar el colector con el emisor hacen que los valores de α y β en modo inverso sean mucho más pequeños que los que se podrían obtener en modo activo; muchas veces el valor de α en modo inverso es menor a 0.5. La falta de simetría es principalmente debido a las tasas de dopaje entre el emisor y el colector. El emisor está altamente dopado, mientras que el colector está ligeramente dopado, permitiendo que pueda ser aplicada

una gran tensión de reversa en la unión colector-base antes de que esta colapse.

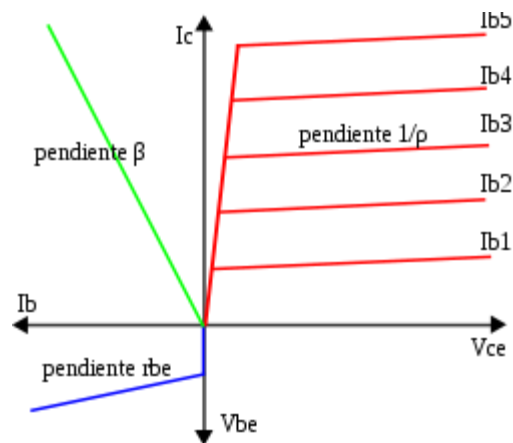
La unión colector-base está polarizada en inversa durante la operación normal. La razón por la cual el emisor está altamente dopado es para aumentar la eficiencia de inyección de portadores del emisor: la tasa de portadores inyectados por el emisor en relación con aquellos inyectados por la base. Para una gran ganancia de corriente, la mayoría de los portadores inyectados en la unión base-emisor deben provenir del emisor.

El bajo desempeño de los transistores bipolares laterales muchas veces utilizados en procesos CMOS es debido a que son diseñados simétricamente, lo que significa que no hay diferencia alguna entre la operación en modo activo y modo inverso.

Pequeños cambios en la tensión aplicada entre los terminales base-emisor genera que la corriente que circula entre el emisor y el colector cambie significativamente. Este efecto puede ser utilizado para amplificar la tensión o corriente de entrada. Los BJT pueden ser pensados como fuentes de corriente controladas por tensión, pero son caracterizados más simplemente como fuentes de corriente controladas por corriente, o por amplificadores de corriente, debido a la baja impedancia de la base.

Los primeros transistores fueron fabricados de germanio, pero la mayoría de los BJT modernos están compuestos de silicio. Actualmente, una pequeña parte de éstos (los transistores bipolares de heterojuntura) están hechos de arseniuro de galio, especialmente utilizados en aplicaciones de alta velocidad.

Funcionamiento



Característica idealizada de un transistor bipolar.

En una configuración normal, la unión emisor-base se polariza en directa y la unión base-colector en inversa. Debido a la agitación térmica los portadores de carga del emisor pueden atravesar la barrera de potencial emisor-base y llegar a la base. A su vez, prácticamente todos los portadores que llegaron son impulsados por el campo eléctrico que existe entre la base y el colector.

Un transistor NPN puede ser considerado como dos diodos con la región del ánodo compartida. En una operación típica, la unión base-emisor está polarizada en directa y la unión base-colector está polarizada en inversa. En un transistor NPN, por ejemplo, cuando una tensión positiva es aplicada en la unión base-emisor, el equilibrio entre los portadores generados térmicamente y el campo eléctrico repelente de la región agotada se desbalancea, permitiendo a los electrones excitados térmicamente inyectarse en la región de la base.

Estos electrones "vagan" a través de la base, desde la región de alta concentración cercana al emisor hasta la región de baja concentración cercana al colector. Estos electrones en la base son llamados portadores minoritarios debido a que la base está dopada con material P, los cuales generan "huecos" como portadores mayoritarios en la base.

La región de la base en un transistor debe ser constructivamente delgada, para que los portadores puedan difundirse a través de esta

en mucho menos tiempo que la vida útil del portador minoritario del semiconductor, para minimizar el porcentaje de portadores que se recombinan antes de alcanzar la unión base-colector. El espesor de la base debe ser menor al ancho de difusión de los electrones.

Control de tensión, carga y corriente

La corriente colector-emisor puede ser vista como controlada por la corriente base-emisor (control de corriente), o por la tensión base-emisor (control de voltaje). Esto es debido a la relación tensión-corriente de la unión base-emisor, la cual es la curva tensión-corriente exponencial usual de una unión PN (es decir, un diodo).

En el diseño de circuitos analógicos, el control de corriente es utilizado debido a que es aproximadamente lineal. Esto significa que la corriente de colector es aproximadamente β veces la corriente de la base. Algunos circuitos pueden ser diseñados asumiendo que la tensión base-emisor es aproximadamente constante, y que la corriente de colector es β veces la corriente de la base. No obstante, para diseñar circuitos utilizando BJT con precisión y confiabilidad, se requiere el uso de modelos matemáticos del transistor como el modelo Ebers-Moll.

El Alfa y Beta del transistor

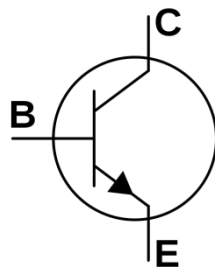
$$\alpha_F := \frac{I_C}{I_E} = \frac{\beta_F}{1 + \beta_F} \quad ; \quad \beta_F = \frac{I_C}{I_B} = \frac{\alpha_F}{1 - \alpha_F}$$

Una forma de medir la eficiencia del BJT es a través de la proporción de electrones capaces de cruzar la base y alcanzar el colector. El alto dopaje de la región del emisor y el bajo dopaje de la región de la base pueden causar que muchos más electrones sean inyectados desde el emisor hacia la base que huecos desde la base hacia el emisor. La ganancia de corriente emisor común está representada por β_F o por h_{fe} . Esto es aproximadamente la tasa de corriente continua de colector a la corriente continua de la base en la región activa directa y es

típicamente mayor a 100. Otro parámetro importante es la ganancia de corriente base común, α_F . La ganancia de corriente base común es aproximadamente la ganancia de corriente desde emisor a colector en la región activa directa. Esta tasa usualmente tiene un valor cercano a la unidad; que oscila entre 0.98 y 0.998. El Alfa y Beta están más precisamente determinados por las siguientes relaciones (para un transistor NPN):

Tipos de Transistor de Unión Bipolar

NPN

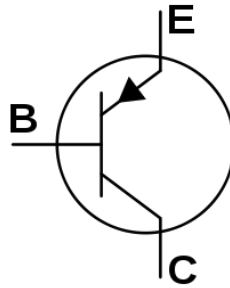


NPN es uno de los dos tipos de transistores bipolares, en los cuales las letras "N" y "P" se refieren a los portadores de carga mayoritarios dentro de las diferentes regiones del transistor. La mayoría de los transistores bipolares usados hoy en día son NPN, debido a que la movilidad del electrón es mayor que la movilidad de los "huecos" en los semiconductores, permitiendo mayores corrientes y velocidades de operación.

Los transistores NPN consisten en una capa de material semiconductor dopado P (la "base") entre dos capas de material dopado N. Una pequeña corriente ingresando a la base en configuración emisor-común es amplificada en la salida del colector.

La flecha en el símbolo del transistor NPN está en la terminal del emisor y apunta en la dirección en la que la corriente convencional circula cuando el dispositivo está en funcionamiento activo.

PNP



PNP es el otro tipo de transistor de unión bipolar es el PNP con las letras "P" y "N" refiriéndose a las cargas mayoritarias dentro de las diferentes regiones del transistor. Pocos transistores usados hoy en día son PNP, debido a que el NPN brinda mucho mejor desempeño en la mayoría de las circunstancias.

Regiones operativas del transistor

Los transistores de unión bipolar tienen diferentes regiones operativas, definidas principalmente por la forma en que son polarizados:

- **Región activa:**
Cuando un transistor no está ni en su región de saturación ni en la región de corte entonces está en una región intermedia, la región activa. En esta región la corriente de colector (I_c) depende principalmente de la corriente de base (I_b), de β (ganancia de corriente, es un dato del fabricante) y de las resistencias que se encuentren conectadas en el colector y emisor. Esta región es la más importante si lo que se desea es utilizar el transistor como un amplificador de señal.
- **Región inversa:**
Al invertir las condiciones de polaridad del funcionamiento en modo activo, el transistor bipolar entra en funcionamiento en modo inverso. En este modo, las regiones del colector y emisor intercambian roles. Debido a que la mayoría de los BJT son diseñados para maximizar la ganancia de corriente en modo activo, el parámetro beta en modo inverso es drásticamente menor al presente en modo activo.

- Región de corte: Un transistor está en corte cuando:
corriente de colector = corriente de emisor = 0, ($I_c = I_e = 0$)
En este caso el voltaje entre el colector y el emisor del transistor es el voltaje de alimentación del circuito. (como no hay corriente circulando, no hay caída de voltaje, ver Ley de Ohm). Este caso normalmente se presenta cuando la corriente de base = 0 ($I_b = 0$)
- Región de saturación: Un transistor está saturado cuando:
corriente de colector = corriente de emisor = corriente máxima, ($I_c = I_e = I_{\text{maxima}}$)
En este caso la magnitud de la corriente depende del voltaje de alimentación del circuito y de las resistencias conectadas en el colector o el emisor o en ambos, ver Ley de Ohm. Este caso normalmente se presenta cuando la corriente de base es lo suficientemente grande como para inducir una corriente de colector β veces más grande. (recordar que $I_c = \beta * I_b$)