

5 El transistor bipolar

5.1 Introducción

El transistor bipolar, conocido en inglés por el acrónimo BJT (*Bipolar Junction Transistor*, transistor bipolar de unión) es uno de los dispositivos más utilizados en los circuitos electrónicos. El nombre de *transistor* procede de la condensación de dos vocablos ingleses, *transfer* y *resistor*, y hace referencia al hecho de que la corriente que circula entre dos terminales está controlada por una señal aplicada al tercer terminal; el término *bipolar* es debido al hecho de que la corriente es transportada por portadores de ambas polaridades: electrones y huecos. La electrónica moderna, basada en circuitos integrados (CI) se inició de hecho con el descubrimiento de este dispositivo. Actualmente sigue siendo el dispositivo amplificador por excelencia, y el que más se emplea en los CI analógicos. En este capítulo se hace una introducción al transistor bipolar y se presenta la teoría de su funcionamiento, su proceso de fabricación, el comportamiento en continua y en señal, y sus modelos circuitales.

El transistor bipolar fue descubierto casualmente en diciembre de 1947 por Bardeen, Brattain y Shockley en los Laboratorios Bell, cuando intentaban realizar un «amplificador de estado sólido» basado en lo que más adelante se denominaría *transistor MOS*. Ese descubrimiento fue seguido casi inmediatamente de la teoría que explicaba su funcionamiento, y condujo a una revolución tecnológica que significó la desaparición, en pocos años, de la tecnología de las válvulas de vacío, que hasta entonces había proporcionado soporte físico a los circuitos electrónicos.

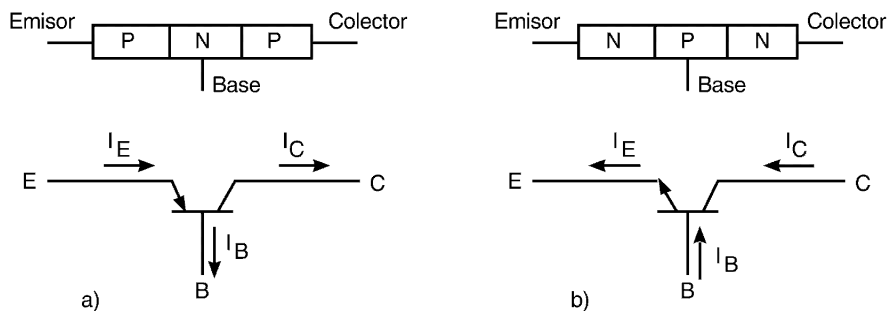


Fig. 5.1 Estructura, símbolo y sentido positivo de las corrientes del transistor bipolar.
a) PNP. b) NPN.

El transistor bipolar es un dispositivo de tres terminales denominados *emisor*, *base* y *colector*. Hay dos tipos de transistores bipolares: los NPN y los PNP, cuyas denominaciones responden a su estructura básica, la cual se esquematiza en la figura 5.1. El transistor bipolar tiene dos uniones PN:

una entre el emisor y la base —la unión *emisora*— y otra entre la base y el colector —la unión *colectora*—. En el símbolo de estos transistores se incluye una flecha en el terminal del emisor, que apunta siempre en el sentido de P a N, y sirve para identificar el tipo de transistor. Los sentidos de las corrientes indicadas en la figura 5.1. se consideran positivos. Estos sentidos se basan en la asignación al emisor del sentido de la corriente de su símbolo —de P a N—, y a la base y al colector de los derivados de la suposición de que la corriente de emisor es la suma de la corriente de base más la corriente de colector.

El transistor más utilizado es el NPN, y es el que se estudia en este texto. Según se ve más adelante, el comportamiento del PNP es dual con respecto al del NPN, lo cual significa que si se cambian electrones por huecos, junto con los sentidos de las corrientes y las polaridades de las tensiones, el comportamiento resulta idéntico.

5.1.1 Principio de funcionamiento

Cuando el transistor se emplea como amplificador, se polariza el diodo del emisor en directa y el del colector en inversa. Esa polarización se denomina *activa* y es la que se considera aquí. Conviene recordar de la teoría de la unión PN que con una polarización directa cada región inyecta sus portadores mayoritarios a la región adyacente, de forma que en el inicio de esa segunda región, justo en la frontera con la zona de carga espacial —la ZCE—, la concentración de los minoritarios inyectados es $m_0 \exp(V_D/V_T)$, siendo m_0 la concentración de minoritarios en esa región en condiciones de equilibrio térmico, y V_D la tensión de polarización de la unión. Cuando una unión se polariza en inversa, el campo eléctrico de la ZCE aumenta y domina el transporte de minoritarios de una región a la otra. Si la polarización es suficientemente negativa, la concentración de éstos en la frontera con la ZCE se anula.

Se analizan ahora las corrientes que circulan por el transistor. Al estar la unión emisora en directa, el emisor N inyecta electrones a la base P, y ésta inyecta huecos al emisor N. La teoría de la unión PN muestra que si el dopaje de la región N es mucho mayor que el de la P, la corriente de electrones a través de la ZCE de la unión es muy superior a la de huecos. Esa relación entre corrientes es la que la figura 5.2 representa. Por otro lado, al estar la unión colectora polarizada en inversa, la concentración de electrones en el punto I_B de la región P —frontera con la ZCE del colector— es cero. Por lo tanto, en la región neutra de la base P del transistor hay una diferencia de concentraciones de electrones entre los puntos 0_B —la frontera con la ZCE del emisor— y I_B , diferencia que origina un flujo de electrones por difusión desde la parte del emisor hacia la parte del colector. Cuando esos electrones llegan a la ZCE del colector, el campo eléctrico presente en esa región los arrastra desde la base hacia el colector. La corriente de huecos en esa segunda unión es prácticamente nula, porque está polarizada inversamente y casi no hay huecos en la parte N del colector.

Una parte de los electrones que se trasladan por difusión desde el emisor hacia el colector a través de la base se recombinan. Es la corriente denominada I_r en la figura 5.2. En régimen estacionario debe entrar por el terminal de la base una corriente de huecos igual a la de electrones que se recombinan, ya que de otro modo los huecos de la base terminarían por agotarse. Por la misma razón, por el terminal de la base deben entrar los huecos que la base inyecta al emisor, y deben salir los pocos huecos que el colector inyecta a la base.

Partiendo de esas corrientes elementales se puede escribir:

$$I_E = I_{En} + I_{Ep} \quad I_C = I_{Cn} + I_{Cp} \quad I_B = I_r + I_{Ep} - I_{Cp} \quad (5.1)$$