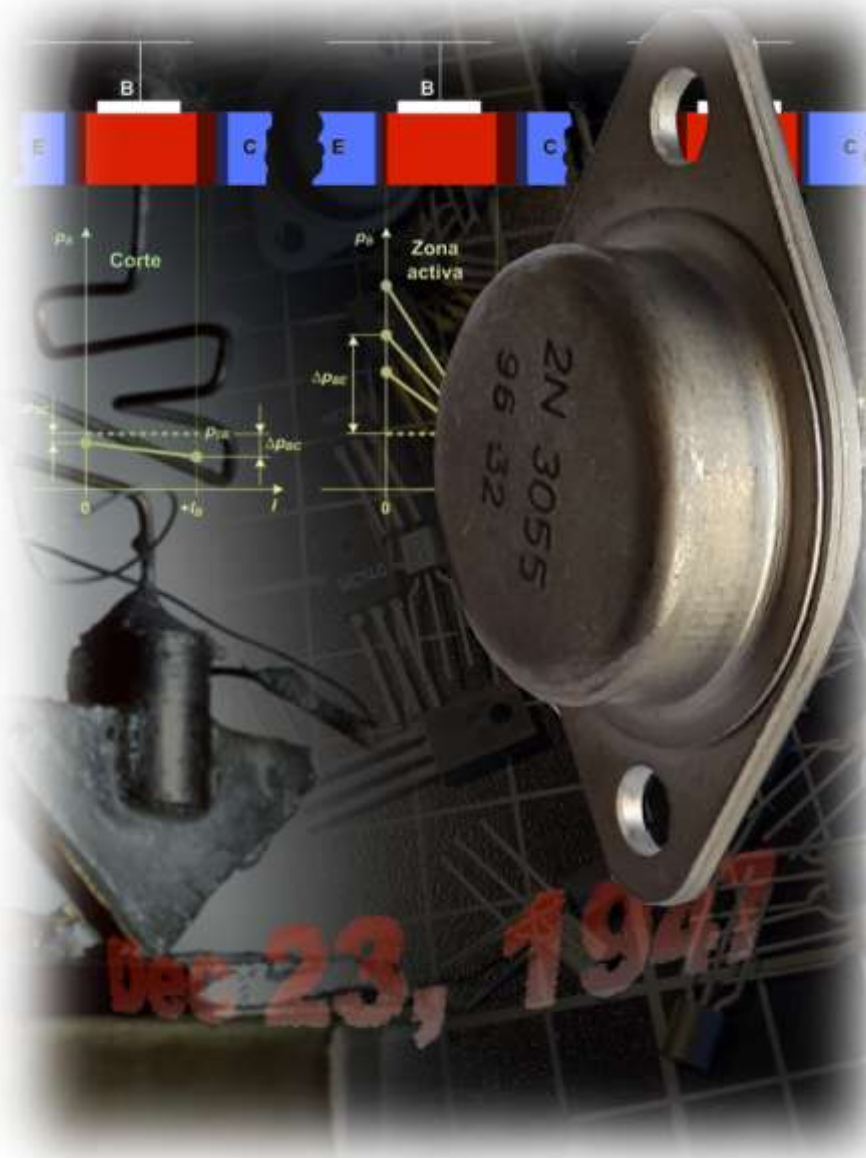


El transistor bipolar

6



En esta lección introducimos los primeros dispositivos electrónicos activos: los transistores bipolares.

Estos dispositivos, basados en una estructura con dos uniones, p-n-p o n-p-n, permiten gobernar señales relativamente intensas en la salida con una acción de control que precisa poca potencia. Ahora estudiaremos las zonas de trabajo y las características generales de estos dispositivos, para pasar en lecciones posteriores a un estudio más centrado en la aplicación al diseño de circuitos.

John Bardeen



John Bardeen (1908-1991) fue un físico estadounidense que participó en el desarrollo del primer transistor bipolar en 1947.

Entró a formar parte del grupo de estado sólido que lideraba Shockley, quien estaba empeñado en sustituir los tríodos de vacío por dispositivos de estado sólido que modularan la conductividad mediante campos eléctricos (siguiendo la idea de Lilienfeld) y que fracasó en todos sus intentos. Fue Bardeen quien propuso cambiar la idea y centrarse en una propuesta diferente que daría lugar al transistor bipolar.

Muestra de su brillantez como científico es el hecho de haber recibido dos veces el Premio Nobel de Física (una de las cuatro personas que han conseguido dos galardones en toda la historia).

El segundo científico del grupo era el físico estadounidense Walter Houser Brattain (1902-1987), con menos nivel teórico que Bardeen, pero de una gran habilidad para resolver los problemas surgidos con los semiconductores. Ellos dos fueron los verdaderos padres del transistor, aunque Shockley, como director del equipo participó del mismo honor y hasta quiso atribuirse el mérito de ser el único verdadero inventor. De hecho, en un alarde de "inteligencia" que precedería a otros del mismo signo, expulsó de su grupo a Bardeen y Brattain.

6.1 Introducción

Cuando en el otoño de 1947 los físicos estadounidenses John Bardeen y Walter Houser Brattain, miembros del equipo de los Laboratorios Bell dirigido por el también físico William Bradford Shockley consiguieron la amplificación de señales mediante un dispositivo semiconductor, aún no sabían qué nombre iba a tener el dispositivo. Más tarde, el ingeniero John R. Pierce sugirió el nombre de "transistor" que seguía los patrones de denominación de otros dispositivos que estaban surgiendo en esa época y que se convirtió en el nombre oficial del invento.

Puede sorprender que el transistor haya aparecido antes que el diodo semiconductor como dispositivo comercial pero, en realidad, lo que ocurría es que el diodo de vacío de la época funcionaba muy bien y la necesidad de sustituirlo por un dispositivo de estado sólido no era acuciante. Aquel descubrimiento tampoco fue un hecho aislado, sino el resultado de un importante esfuerzo investigador de muchas personas desde hacía más de dos décadas. Todo empezó con la patente canadiense de 1925 del físico austro-húngaro Julius Edgard Lilienfeld sobre un dispositivo que funcionaría como un verdadero transistor, pero que no fue probado (ni siquiera se disponía de los materiales necesarios). De hecho, conociendo tal patente, Shockley intentó fabricar el dispositivo sin éxito alguno; tal fracaso indujo a pensar que el invento de Lilienfeld no podría funcionar y así se difundió durante décadas, en buena parte para justificar el poco ético hecho de no haber citado a Lilienfeld en ninguna de las publicaciones y patentes de los Laboratorios Bell sobre el tema.

Casi de forma simultánea e independiente de los americanos, los físicos alemanes Herbert Mataré y Heinrich Welker desarrollaron otro transistor similar al de Bardeen y Brattain, aunque nunca alcanzó la misma repercusión, mucho menos después de que estos últimos recibieran el Premio Nobel en 1956 (junto a Shockley) por su descubrimiento. El círculo relativo al invento del transistor se cerró a finales del siglo XX, cuando se demostró que el transistor de Lilienfeld habría funcionado si en su época hubieran estado disponibles los materiales semiconductores adecuados.

Desde los primeros tiempos del transistor hasta la actualidad se ha avanzado mucho hasta conseguir dispositivos de excelentes prestaciones, de más o menos potencia, de tamaños diminutos, susceptibles de ser integrados en bloques monolíticos y constituir sistemas electrónicos de una complejidad que hubiera sido inimaginable en aquel entonces. La totalidad de los circuitos integrados, desde los que se dedican al mundo analógico hasta los sistemas que trabajan en el mundo digital y configuran el núcleo de nuestros potentes computadores y sistemas de comunicación, están fabricados con transistores. Así, puede decirse sin temor alguno que la invención del transistor marcó el comienzo de una nueva época que ha trascendido el ámbito tecnológico y producido un cambio histórico sin precedentes en la sociedad.

Pero, ¿qué es un transistor? La respuesta es amplia pero, para simplificar —tiempo habrá de complicar las ideas a lo largo de esta lección y de las siguientes— podemos imaginar el transistor como un sistema que es capaz de gobernar señales de cierta potencia con otras señales mucho más débiles, como

se muestra en la Figura 6.1.

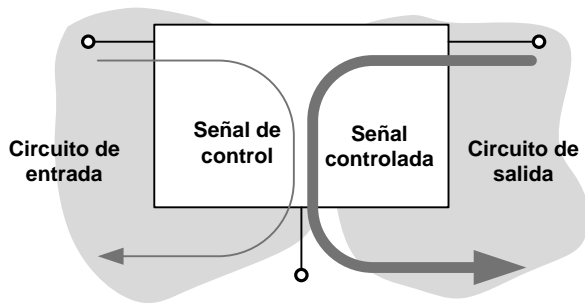


Figura 6.1. Idea de funcionamiento de un transistor: una señal pequeña es capaz de controlar una señal grande.

Existen dos tipos básicos de transistores, en función de si su funcionamiento se explica con uno o con dos tipos de portadores. Los segundos, que será por los que empezamos, siguen la idea del desarrollado por Bardeen y Brattain, un dispositivo que trabaja con los dos tipos de portadores y que, por ello, se denominan **transistores bipolares** de unión, BJT (*Bipolar Junction Transistor*) y a los que nos referiremos en adelante como transistores bipolares. Los primeros, basados en la conducción mediante uno de los dos tipos de portadores, son transistores unipolares (“similares” al que ideó Lilienfeld), aunque dado que su funcionamiento se debe a la presencia de un campo eléctrico que modula la conducción, se denominan **transistores de efecto campo**, FET (*Field Effect Transistor*). De ellos hablaremos en lecciones posteriores.

6.2 El transistor bipolar

Un transistor bipolar es un dispositivo formado por dos uniones entre materiales de tipos n y p, por lo que hay dos opciones: poner el material n en el medio y formar un dispositivo cuya secuencia de materiales sea p-n-p, o hacer justo lo contrario, lo que lleva a una secuencia de materiales n-p-n. Estas secuencias son las que dan nombre a los dos tipos de transistores bipolares, de modo que tendremos, respectivamente, dispositivos pnp y npn, cuyas características resultarán duales. Sin embargo, una sucesión de materiales como esa no tiene que desencadenar necesariamente ningún efecto nuevo, diferente a los ya conocidos. Consideremos, por ejemplo, un sistema como el de la Figura 6.2, en el que se han dispuesto tres trozos de semiconductor unidos de una forma similar a la que se explicó en el caso de los diodos y de longitud suficiente como para considerarlos largos en relación con la longitud de difusión.

El sistema así diseñado no parece tener ninguna influencia de una unión en la otra puesto que la zona intermedia, constituida con semiconductor de tipo n tiene longitud suficiente como para que los efectos de inyección de portadores que supone la presencia de cada una de las uniones haya desaparecido mucho antes de acercarse a la otra. Cada unión puede estar polarizada directa o inversamente y su situación es perfectamente compatible con cualquier polarización

El transistor europeo



El físico alemán Herbert Franz Mataré (1912-2011) patentó en agosto de 1948 el primer transistor europeo, de forma independiente al desarrollado por Bardeen y Brattain, junto al también físico alemán Heinrich Johann Welker (1912-1981).



Aunque no compartieron el Premio Nobel con sus colegas americanos, la Ciencia ha reconocido la invención del transistor también para estos dos científicos y, de hecho, la patente fue concedida igualmente.

La empresa fundada por Mataré en 1951, Intermetall, fue la primera en comercializar transistores como dispositivos electrónicos.

1952: el año clave

Apenas pasaron cuatro años desde que el primer transistor había sido probado con éxito en un laboratorio y ya aparecieron las primeras aplicaciones.

La más madrugadora fue la de la empresa Intermetall que presentó en Dusseldórf un receptor de radio con transistores de carácter experimental.

Enseguida surgió la segunda aplicación que sacaba provecho del reducido tamaño del dispositivo: el Sonotone 1010, aunque esa primera versión de audífono fue un dispositivo híbrido que aún mantenía válvulas junto al único transistor que empleaba.



La fotografía del Sonotone 1010 es cortesía de The Hearing Aid Museum (<http://hearingaidmuseum.com>)

Finalmente, también la empresa RCA se sumó al uso de transistores y presentó su primer amplificador de audio que incluía una etapa final en push-pull que evitaba el acoplamiento con transformador típico de las válvulas, un diseño que se iría propagando hasta formar parte de la mayoría de las etapas de salida de los amplificadores actuales.

Así pues, 1952 fue el año en el que aparecieron los primeros productos comerciales con transistores. En un contexto de guerra fría, marcada por el comienzo de una rivalidad entre bloques, el transistor actuó como factor desequilibrante. Las empresas estadounidenses vieron la oportunidad y armados con el nuevo dispositivo destacaron una separación tecnológica con el resto del mundo que les ha permitido mantener hasta la actualidad una posición privilegiada.

en la otra. En resumen, el sistema de la Figura 6.2 se comporta como un conjunto de dos diodos independientes que comparten el cátodo, y que se podría usar con éxito como rectificador en medio puente. A una conclusión semejante se llegaría en el caso de tener una disposición n-p-n de similares características, excepto porque ahora se tendría un bloque de dos diodos de ánodo común.

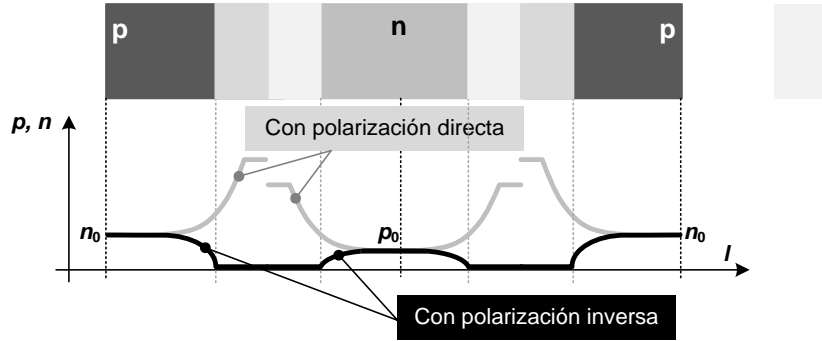


Figura 6.2. Disposición de materiales semiconductores p-n-p de longitud suficiente de modo que hay independencia entre las dos uniones. Dos simples diodos y ningún efecto nuevo...

Luego la disposición de dos uniones sin más no permite ningún efecto diferente a los ya estudiados para los diodos. ¿Dónde está el efecto de amplificación de señales que caracteriza al transistor?

Consideremos ahora un caso diferente, en el que la zona intermedia formada por el material n se ha construido mucho más estrecha, tanto que su longitud es bastante menor que la longitud de difusión de portadores. Esto produce dos cambios; el primero y más importante es que los efectos de la inyección de portadores de la zona de transición de una unión llegan hasta la zona de transición de la otra unión y el segundo, que se modifica el aspecto de la solución de la ecuación de continuidad puesto que cambian las condiciones de contorno. Con este diseño vamos a polarizar una de las uniones directamente y la otra inversamente según se muestra en la Figura 6.3.

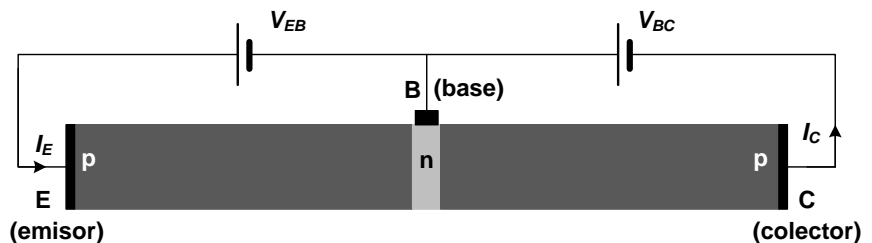


Figura 6.3. Dispositivo pnp con polarización directa entre base y emisor y polarización inversa entre colector y base.

Aunque podríamos realizar el análisis sin ninguna denominación especial para las tres zonas del dispositivo, vamos a utilizar el nombre habitual para cada una de ellas y, de esa forma, facilitar la referencia a cada una de ellas. Llama-

remos **base** (*base*) y la denotaremos por “B” a la zona de material n y nombraremos como **emisor**, E (*emitter*) y **colector**, C (*collector*) a cada una de las otras dos zonas, ambas de material p.

Para entender el funcionamiento podemos pensar que si no existiese la zona de base, la conducción entre emisor y colector sería buena ya que se realizaría sobre un trozo de semiconductor de tipo p que se extendería entre ambos. Pero en medio está la base y eso evita esa conducción franca. Sin embargo, también es fácil suponer que si jugamos adecuadamente con los portadores que inyectamos en la base podríamos llegar a facilitar el paso de los del lado de emisor al lado de colector o impedirlo por completo. Así pues, parece que la clave de funcionamiento del dispositivo está en lo que ocurre en los alrededores de la base, de modo que concentraremos nuestro estudio en esa zona.

Con las condiciones de polarización que tenemos indicadas en la Figura 6.3, el estado en los alrededores de la base supone una unión polarizada directamente (base-emisor) con una tensión V_{EB} y otra, inversamente (base-colector) con una tensión V_{BC} , con lo que se establecerán las correspondientes distribuciones de portadores a ambos lados de cada una de las zonas de transición según se estudió en la Lección 3. Comenzando por la unión base-emisor (Figura 6.4) tenemos que como unión polarizada directamente, se produce una inyección de huecos en la zona de base y de electrones en la zona de emisor.

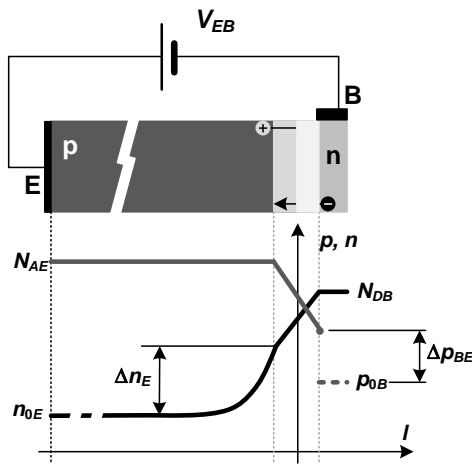


Figura 6.4. Estado de la unión base-emisor polarizada directamente (las escalas de concentraciones son logarítmicas).

Según vimos en el Apartado 3.2, la concentración de minoritarios en los extremos de la zona de transición es:

$$n_E = n_{0E} \cdot e^{\frac{V_{EB}}{V_T}} \quad p_{BE} = p_{0B} \cdot e^{\frac{V_{EB}}{V_T}}$$

Se ha añadido el subíndice “E” para denotar que se refiere al emisor y los subíndices “BE” para indicar que es en la zona de base, pero en el extremo más próximo al emisor. La polarización directa introduce un incremento de minoritarios en los extremos de la zona de transición, que se traduce en un incre-

Regency TR-1

Fue el primer receptor de radio realizado íntegramente con transistores (incluía cuatro transistores y un diodo) y se caracterizaba por su reducido tamaño y precio (menos de 50 \$), así como por un sonido de no muy buena calidad.

El diseño fue patentado por Richard C. Koch (US Patent 2892931) y salió al mercado en 1954 por las empresas IDEA y Texas Instruments con un notable éxito.



En la fotografía de Theoprakt, un modelo del Regency TR-1 que se conserva en el Deutsche Museum de Munich.

En poco tiempo, esta clase de dispositivo pasaría a denominarse popularmente como “transistor”.

mento de los electrones en el emisor y un incremento de los huecos en la base. Dichos excesos respecto de las concentraciones basales (n_{0E} y p_{0B}) en los extremos de la zona de transición serán:

$$\Delta n_E = n_{0E} \cdot \left(e^{\frac{V_{EB}}{V_T}} - 1 \right) \quad \Delta p_{BE} = p_{0B} \cdot \left(e^{\frac{V_{EB}}{V_T}} - 1 \right)$$

Sin embargo, en la otra unión –polarizada inversamente– la situación es justo la contraria, con disminución de electrones en el colector y de huecos en la base (Figura 6.5), según se estudió en el Apartado 3.3:

$$n_C = n_{0C} \cdot e^{\frac{-V_{BC}}{V_T}} \quad p_{BC} = p_{0B} \cdot e^{\frac{-V_{BC}}{V_T}}$$

La disminución de minoritarios en los extremos de la zona de transición es:

$$\Delta n_C = n_{0C} \cdot \left(e^{\frac{-V_{BC}}{V_T}} - 1 \right) \quad \Delta p_{BC} = p_{0B} \cdot \left(1 - e^{\frac{-V_{BC}}{V_T}} \right)$$

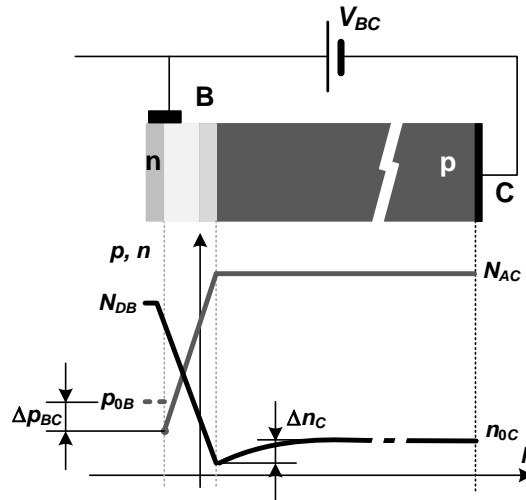


Figura 6.5. Situación de la unión base-colector polarizada inversamente (las escalas de concentraciones son logarítmicas).

Ya hemos determinado en la Lección 3 cómo varía la concentración de minoritarios en un semiconductor largo (en relación con la longitud de difusión correspondiente) y lo hemos representado de forma aproximada en las figuras 6.4 y 6.5; sin embargo, no hemos calculado cómo es ese cambio en la base, donde su tamaño es mucho menor que la longitud de difusión (Figura 6.6).

En este caso, al resolver la ecuación de continuidad no nos encontramos con una evolución exponencial en el espacio, sino que obtenemos:

$$p_B(l) = p_B(l_B) + (p_B(0) - p_B(l_B)) \frac{\sinh \frac{l_B - l}{L_P}}{\sinh \frac{l_B}{L_P}}$$

El primer transistor de silicio

Contrariamente a lo que mis colegas les han dicho acerca de las sombrías perspectivas de los transistores de silicio, tengo unos cuantos aquí, en mi bolsillo. Esas fueron las palabras desafiantes pronunciadas por Gordon Kidd Teal (1907-2003) de Texas Instruments, durante su presentación en la National Conference on Airborne Electronics, Dayton, Ohio (EE.UU. 10 de mayo de 1954).

Era la presentación en sociedad del primer transistor de silicio. Curiosamente, en los Laboratorios Bell se acababa de desarrollar un dispositivo del mismo tipo, cuyos detalles serían publicados por Morris Tannenbaum, L.B. Valdes, E. Buehler y N.B. Hannay, *Silicon n-p-n Grown Junction Transistors*, Journal of Applied Physics, Volume 26, Issue 6, p.686-692. 1955.