

2 La unión PN

La mayoría de los dispositivos semiconductores utilizados en la electrónica contienen regiones del tipo P y regiones del tipo N. Las propiedades de los contactos entre los dos tipos de zona, denominados *uniones PN*, son fundamentales para el funcionamiento del dispositivo. Como ejemplos, un diodo es un dispositivo formado por una sola unión; un transistor, tanto si es bipolar como MOS, contiene al menos dos uniones, etcétera. Se dedica este capítulo al estudio del comportamiento de la unión PN.

Después de una descripción cualitativa del funcionamiento de una unión PN, se dedica la mayor parte del capítulo al análisis cuantitativo empleando un modelo simple. Se comienza por el estudio del sistema en equilibrio, se continúa con la evaluación de las corrientes bajo la aplicación de tensiones continuas, y se pasa luego a ver el efecto de las tensiones dependientes del tiempo —los efectos dinámicos—. Aunque este estudio se centra en las uniones en las que la región P y la región N son del mismo semiconductor —homouniones—, se discute también las características de las heterouniones, formadas por dos semiconductores distintos. Se acaba el capítulo presentando el comportamiento de los contactos entre un metal y un semiconductor, y discutiendo las características que deben tener los contactos que actúan como terminales de un dispositivo —los contactos óhmicos—.

2.1 La unión PN: bandas de energía y efecto rectificador

2.1.1 Hipótesis iniciales del modelo

Una unión PN es un cristal semiconductor único, con una región dopada con impurezas aceptoras y otra con impurezas donadoras. Cuando los dopajes son homogéneos en el interior de cada región y existe un plano de separación entre ambas, se habla de *unión abrupta*. Si el cambio de dopaje es progresivo, se habla de *unión gradual*. Este estudio se centra en las uniones abruptas; se denomina N_A la concentración de impurezas aceptoras de la región P, mientras que N_D es el nivel de dopaje de la región N. Se supone que para esas impurezas se cumple la condición de ionización total.

Se consideran válidas las principales hipótesis que han permitido llevar a cabo el estudio de los semiconductores del primer capítulo; particularmente, que los semiconductores son no degenerados y que se mantienen las condiciones de baja inyección. Por otro lado, este trabajo se simplifica empleando un modelo unidimensional: las variables que se utilizan dependen de una sola coordenada, medida en la perpendicular al plano de la unión, pero no de las otras dos coordenadas.

2.1.2 El diagrama de bandas

La figura 2.1 representa el símbolo, la estructura y el diagrama de bandas en equilibrio de la unión PN. Debe recordarse aquí que el nivel de Fermi, E_f , es constante en equilibrio térmico. La deformación de

los niveles E_c y E_v —y junto con ellos, de E_f — indica que hay un campo eléctrico en sentido de derecha a izquierda en la región de transición, es decir, un campo que va de la región N a la P. Tal como se ha descrito en el capítulo 1, ese campo eléctrico confina los portadores mayoritarios en las respectivas regiones: los electrones en la región N y los huecos en la P. Solo pueden pasar a la otra región los portadores que tengan suficiente energía como para superar el escalón representado por la curvatura de las bandas —en equilibrio, una energía cinética superior a qV_{bi} —. Dado el valor que suele tener esa barrera, el porcentaje de portadores que la pueden superar es muy pequeño.

Se distinguen tres regiones en la unión PN: la *región neutra P*, en la que el campo eléctrico es nulo y en la que hay, por lo tanto, neutralidad de carga; la *zona de carga espacial (ZCE) o región de transición entre P y N*, en la que hay un campo eléctrico intenso producido por un dipolo de carga espacial; y, finalmente, la *zona neutra N*. Se supone que en la zona neutra P la concentración de huecos es N_A y $n \ll p$, mientras que para la región neutra N se supone que $n = N_D$ y $p \ll n$.

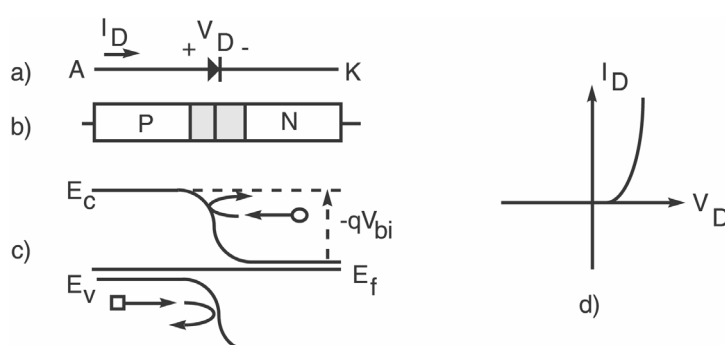


Fig. 2.1 La unión PN: a) Símbolo circuital. b) Estructura física. c) Diagrama de bandas en equilibrio térmico. d) Característica $I(V)$ mostrando el efecto rectificador.

2.1.3 El efecto rectificador

El efecto rectificador consiste en la propiedad de permitir el paso de la corriente en un sentido —de P a N, en el caso de una unión PN— y de bloquearlo en el sentido contrario (v. fig. 2.1.d). El dispositivo que presenta ese efecto se denomina *diodo*.

Más adelante se justifica que si se aplica una tensión positiva de valor V_D en la región P con respecto a la región N —se denomina *polarización directa*—, el escalón de energía en la ZCE pasa a valer $q(V_{bi} - V_D)$. La polarización directa causa la disminución del campo eléctrico en la región de transición, porque reduce la pendiente del perfil de E_f . En polarización inversa —región N positiva con respecto a la P— la altura del escalón $q(V_{bi} - V_D)$ aumenta.

En polarización directa, la disminución del escalón de energía en la región de transición permite el paso de muchos portadores mayoritarios de cada una de las regiones a la otra, porque precisan menos energía cinética para ello. En efecto: según la ley de Fermi, la distribución energética de los electrones dentro de la banda de conducción disminuye exponencialmente a medida que su energía se aleja de E_c . De la misma manera, el número de huecos dentro de la banda de valencia disminuye exponencialmente a medida que su energía se aleja de E_v . Así pues, en polarización directa hay un flujo muy intenso de huecos de P a N, y de electrones de N a P. En consecuencia, la corriente a través del diodo —en sentido de P a N— aumenta exponencialmente con la tensión de polarización.