

TEMA 3: FÍSICA DE LA UNIÓN P-N

Introducción.

3.1. La unión p-n en equilibrio térmico.

Formación de la región espacial de carga.
Barrera de potencial y bandas de energía.
Corrientes de arrastre y de difusión.

3.1.1. Relaciones analíticas en equilibrio térmico.

3.1.1.1. *Electrostática de la región espacial de carga.*
3.1.1.2. *Constancia del nivel de Fermi.*
3.1.1.3. *Potencial de contacto.*

3.2. Resolución de la ecuación de Poisson en la región espacial de carga.

3.2.1. La unión abrupta.
3.2.2. La unión gradual.

3.3. La unión P-N polarizada: diodo de unión.

Polarización directa.
Polarización inversa.

3.3.1. Capacidad de transición de la unión polarizada.

3.4. Modelo matemático de la unión P-N: ecuación de Shockley.

3.4.1. Características ideales.

- Concentración de portadores en los límites de la región espacial de carga.
- Distribución de portadores minoritarios en las zonas neutras.

3.5. Fenómenos de ruptura.

- a) Ruptura por inestabilidad térmica.
- b) Ruptura por efecto túnel.
- c) Ruptura de multiplicación por avalancha.

3.5.1. *Cálculo de la tensión de ruptura por avalancha.*

3.5.2. *Efecto de la curvatura de la unión.*

SOLUCIONES:

1. Disminución de la concentración ligera de impurezas.
2. Técnicas de biselado y de anillos de campo.
3. Implantación iónica.

3.6. Almacenamiento de carga y capacidad de difusión.

Circuito equivalente en régimen de pequeña señal.

3.7. Comportamiento transitorio: tiempo de paso a corte.

3.8. Unión metal-semiconductor.

3.8.1. *Diodos Schottky.*

3.8.2. *Contacto óhmico metal-semiconductor.*

3.9. Estructura $p^+ - n - n^+$.

3.10. Construcción del diodo.

TEMA 3: FÍSICA DE LA UNIÓN P-N.

Introducción.

En los temas anteriores se han estudiado las propiedades de los semiconductores en equilibrio térmico y fuera del equilibrio. Los semiconductores intrínsecos tienen un uso muy limitado, sin embargo los semiconductores dopados con impurezas son la base de los dispositivos que a partir de ahora vamos a estudiar.

Si sobre la superficie de un semiconductor previamente dopado con impurezas aceptoras se difunden átomos donadores se forma una unión de propiedades muy interesantes llamada unión p-n. La unión p-n desempeña un importante papel en las aplicaciones de la electrónica moderna, así como en la construcción y aplicación de otros dispositivos semiconductores. Se utiliza por ejemplo en aplicaciones de rectificación, conmutación, ... Además es un dispositivo fundamental en la construcción de otros dispositivos semiconductores tales como los transistores bipolares, tiristores, transistores de efecto de campo o dispositivos para aplicación en microondas o fotónicos.

En esta lección se van a presentar las características ideales estáticas y dinámicas de la unión p-n en base a las ecuaciones obtenidas en la lección anterior.

3.1. La unión p-n en equilibrio térmico.

Fig. 1.a muestra dos materiales semiconductores, uno de tipo P y otro de tipo N separados entre sí. En ellos el nivel de Fermi E_F está cerca de la banda de valencia para la muestra de tipo P y cerca de la banda de conducción para el material de tipo N. Por otra parte, el material de tipo P tiene una concentración de huecos mucho mayor que de electrones y al revés ocurre en el material de tipo N.

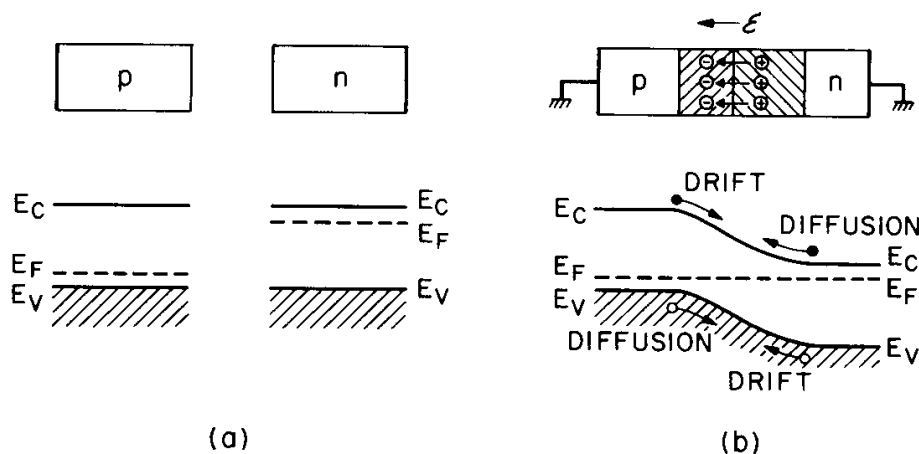


Fig. 1: (a) Semiconductores de tipo N y tipo P uniformemente dopados antes de la unión. (b) Campo eléctrico en la región espacial de carga y diagrama de bandas de energía en la unión p-n en equilibrio térmico.

Formación de la región espacial de carga.

Una unión p-n se forma cuando se unen estas dos regiones. En la práctica, los procesos de fabricación de uniones p-n son los de epitaxia, difusión e implantación de iones.

TEMA 3 : Física de la unión P-N.

Para simplificar vamos a suponer el caso en que el límite entre las regiones P y N represente una *unión escalón* o *unión abrupta*. En este tipo de unión la transición entre las regiones P y N tiene lugar en una distancia extremadamente pequeña. Por lo tanto, hay un cambio brusco en el dopaje yendo desde P hasta N.

Aunque la unión abrupta no es una estructura típica de los dispositivos modernos, sí es, en cambio, una buena aproximación de la unión p-n, que pone de manifiesto sus características de funcionamiento y de tensión-corriente. Otro argumento a favor de utilizar dicha aproximación es que el comportamiento físico interno y las propiedades eléctricas de la unión varían muy poco con el método que se emplee en su obtención.

El hecho de unir ambos tipos de material provoca un elevado gradiente de concentración de portadores en las proximidades de la unión. De hecho, lo que se observa es por un lado, una corriente de difusión de huecos de la región P hacia la región N y por otro, una corriente de difusión de electrones de la región N hacia la P. Pero además, la marcha o difusión de estos portadores de su región inicial a la otra deja al descubierto algunos iones fijos en la red cristalina, iones aceptores negativos N_A^- en la región P e iones dadores positivos en la región N, N_D^+ . Dicha zona es llamada *zona de agotamiento* o *región espacial de carga* y es una región en que no existen ni electrones ni huecos pero contiene átomos dadores ionizados positivamente a un lado y átomos aceptores ionizados negativamente al otro lado. En consecuencia, se forma una zona de cargas negativas fijas en la cara P de la unión y una zona con cargas positivas fijas en la cara N. La presencia de estas cargas fijas da lugar a la aparición de un fuerte campo eléctrico cuyas líneas de campo se dirigen desde la zona N hacia la zona P, es decir, desde la zona de carga positiva a la zona de carga negativa (Fig. 1b). La orientación, por tanto, de este campo eléctrico es siempre en contra de la corriente de difusión del portador de carga considerado.

Las condiciones después del contacto se muestran en Fig. 2:

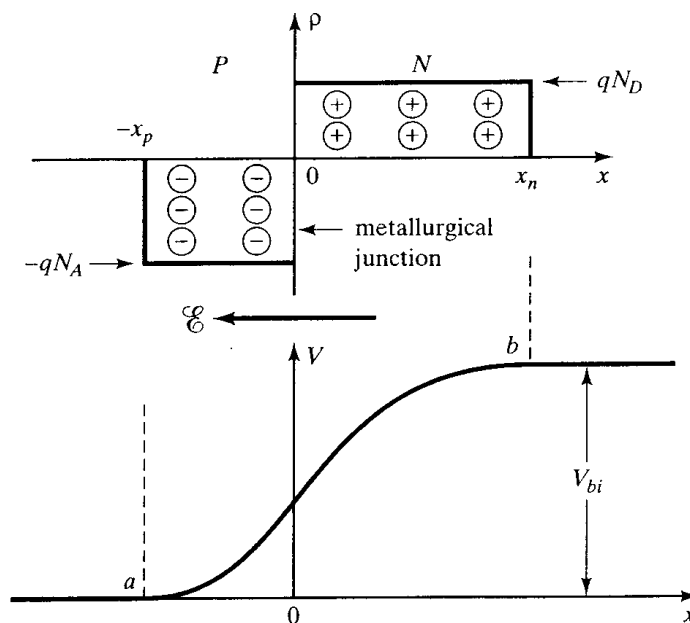


Fig. 2: Región espacial de carga, sentido del campo eléctrico y potencial de contacto en equilibrio térmico.

Se ha supuesto que la región espacial de carga contiene únicamente átomos de impurezas ionizados (en realidad existen unas pequeñas concentraciones de portadores

móviles -electrones y huecos- despreciables frente a las concentraciones de impurezas) y que más allá de esta región los materiales (P y N) son neutros.

Si consideramos, por ejemplo, el semiconductor de tipo N, en todo este semiconductor existen N_D impurezas por unidad de volumen pero en la zona neutra la carga de los iones es compensada con la de los portadores móviles mayoritarios (electrones) mientras que en la zona próxima a la unión la carga de los iones dadores cargados positivamente no es compensada.

Observamos como la zona N de la región espacial de carga se extiende desde $x = 0$ hasta $x = x_n$, mientras que la zona P se extiende desde $x = 0$ hasta $x = -x_p$. La línea que separa las dos regiones está en $x = 0$ y corresponde a la unión metalúrgica de las regiones P y N.

Barrera de potencial y bandas de energía.

La presencia de cargas en la región espacial de carga da lugar a un campo eléctrico, el cual causa una diferencia de potencial en dicha región que recibe el nombre de *potencial de contacto*, V_{bi} y es mostrado en Fig. 2.

Este potencial de contacto viene acompañado por un desnivel de las bandas de energía como se muestra en Fig. 3. Este desnivel se explica como sigue: la caída de tensión entre dos puntos a y b se define como la energía empleada o adquirida para mover una carga positiva unidad desde **a** hasta **b**. Si la carga positiva (huevo) está en **a** (zona P) y **b** (zona N) está a una mayor tensión V_{bi} , se realiza un trabajo para mover la carga. Al llegar a **b**, la carga positiva ha ganado una energía potencial igual a la energía empleada que viene dada por $(q \cdot V_{bi})$, donde q es la carga del electrón.

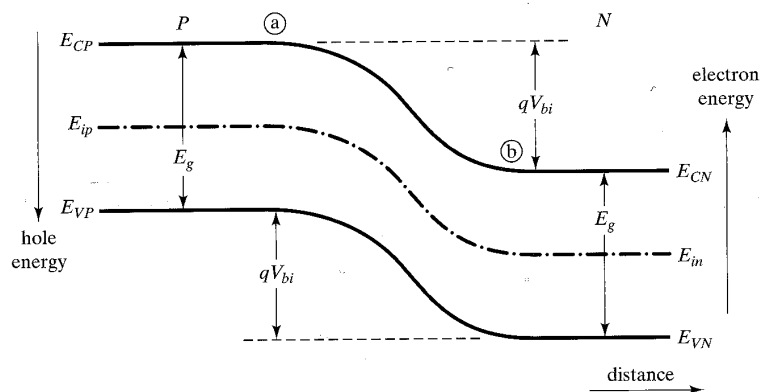


Fig. 3: Diagrama de bandas de energía en la zona de la unión.

Si un electrón es llevado desde **a** hasta **b**, con **b** a una tensión superior V_{bi} , el electrón (carga negativa) pierde una energía potencial $-q \cdot V_{bi}$.

En consecuencia, un electrón en $x \geq x_n$ tiene menos energía potencial que un electrón en $x \leq x_p$ (zonas neutras). Los niveles energéticos de la zona N son menores que los de la zona P en una cantidad $q \cdot V_{bi}$.

Recordemos que los electrones residen en la banda de conducción y los huecos en la de valencia y que la energía del electrón aumenta al moverse hacia arriba en la banda de conducción siendo E_c la mínima energía de un electrón libre (corresponde a la energía de un electrón en reposo y toda ella es energía potencial –energía cinética cero– aumentando la energía cinética al desplazarse hacia arriba en la banda de conducción), mientras que la del huevo aumenta al moverse hacia abajo en la banda de valencia siendo E_v la energía de un huevo en reposo (toda ella es energía potencial –energía

TEMA 3 : Física de la unión P-N.

cinética cero- aumentando la energía cinética al desplazarse hacia arriba en la banda de conducción)

Al potencial de contacto también se le conoce como *barrera de potencial*.

Corrientes de arrastre y de difusión.

La presencia del campo eléctrico en la zona de la unión debido a la presencia de cargas fijas, orientado en sentido contrario al movimiento de cada uno de los portadores por difusión, crea a su vez una corriente de arrastre. Por un lado hay una corriente de difusión de huecos de la zona P a la N debida al gradiente de concentración y una corriente de arrastre también de huecos (debida al campo eléctrico) de la zona N a la P. Con respecto a los electrones, la corriente de difusión debida al gradiente de concentración va de la zona N a la P y la de arrastre debida al campo eléctrico de la zona P a la N.

En condiciones de equilibrio térmico, sin ninguna excitación externa, la corriente total, suma de las dos corrientes de arrastre (de electrones y de huecos) y de las dos corrientes de difusión ha de ser cero. Una corriente total que no fuese cero causaría una disipación energética, lo cual no es posible ya que no hay ninguna fuente en el diodo que proporcione dicha energía.

Por otra parte la corriente total de electrones y la corriente total de huecos son ambas cero. La corriente neta correspondiente a cada tipo de portador (electrones o huecos) ha de ser cero debido al siguiente razonamiento: como la unión p-n está en equilibrio térmico, el producto n-p debe ser igual a n_i^2 . Supongamos que la corriente neta de electrones (diferencia entre las corrientes de difusión y arrastre) no fuese cero, entonces la corriente neta de huecos debería ser igual en magnitud pero sentido contrario a la de los electrones (pues la corriente total es cero). Esto implicaría que habría una transferencia neta de electrones y huecos en el mismo sentido provocando un incremento de las concentraciones de ambos en una de las regiones y violando por tanto la condición de equilibrio térmico.

3.1.1. Relaciones analíticas en equilibrio térmico.

A partir de ahora se van a establecer las relaciones analíticas para la unión p-n en equilibrio térmico. Determinaremos primeramente la electrostática de la región espacial de carga y con posterioridad la localización del nivel de Fermi y una expresión para el potencial de contacto en la unión.

3.1.1.1. Electrostática de la región espacial de carga.

Si recordamos la definición de campo eléctrico, este se define como la fuerza ejercida sobre la unidad de carga positiva. Cuando un campo eléctrico ϵ se aplica al semiconductor, dicho campo realiza una fuerza $-q \cdot \epsilon$ sobre cada electrón, donde q es el valor de la carga del electrón. También sabemos que dicha fuerza ejercida es igual al valor negativo del gradiente de la energía potencial:

$$F = -q \cdot \epsilon = - (\text{gradiente de la energía potencial del electrón})$$

Sabemos que el valor inferior de la banda de conducción E_c corresponde a la energía potencial de un electrón. Como estamos interesados en el gradiente de la energía potencial, podemos utilizar cualquier parte del diagrama de bandas que sea paralelo a E_c

TEMA 3 : Física de la unión P-N.

(es decir, E_F , E_i o E_v). Es conveniente utilizar E_i pues dicha magnitud es ampliamente utilizada en el caso de uniones P-N . Por tanto y para el caso unidimensional:

$$-q \cdot \varepsilon = -\frac{dE_i}{dx} \Rightarrow \varepsilon = \frac{1}{q} \cdot \frac{dE_i}{dx}$$

Ya se calculó el valor de dicho campo eléctrico para el caso de un semiconductor de tipo N sujeto a una diferencia de potencial constante.

El potencial electrostático, ψ , se define como aquella magnitud cuyo gradiente negativo iguala al campo eléctrico. Por tanto, para el caso unidimensional:

$$\varepsilon \equiv -\frac{d\psi}{dx} \quad (\text{O sea } \psi \equiv -\int \varepsilon \cdot dx)$$

Con lo que:

$$\psi = -\frac{E_i}{q}$$

La ecuación de Poisson para el caso unidimensional:

$$\frac{d\varepsilon}{dx} = \frac{\rho_s}{\varepsilon_s} \Rightarrow \begin{cases} \frac{d^2\psi}{dx^2} = -\frac{\rho_s}{\varepsilon_s} \\ \frac{d^2E_i}{dx^2} = q \cdot \frac{\rho_s}{\varepsilon_s} \end{cases}$$

nos permite relacionar el potencial electrostático, la energía potencial y el campo eléctrico para una unión p-n.

Moviéndonos desde cualquiera de las dos regiones neutras hacia la unión p-n aparecen distintas zonas. En una primera zona, la carga aportada por los iones de las impurezas es parcialmente compensada con la carga de los portadores móviles, es una región estrecha de poco espesor y recibe el nombre de *región de transición* (Fig. 4c).

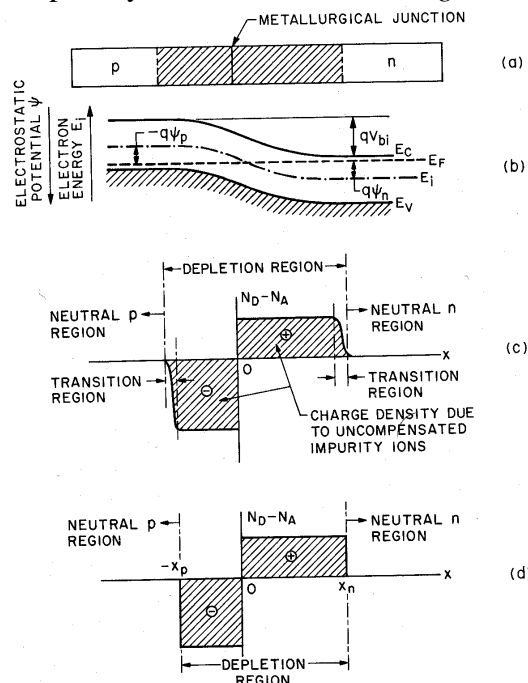


Fig. 4 : (a) Unión p-n con cambios bruscos de dopado en la unión metalúrgica. (b) Diagrama de bandas de energía de la unión abrupta en equilibrio térmico. (c) Distribución espacial de carga. (d) Aproximación rectangular de la distribución espacial de carga.

Más allá de esta región aparece la región espacial de carga. Para uniones p-n típicas, la anchura de la región de transición es pequeña comparada con la de la región espacial de carga, de forma que puede despreciarse dicha región de transición y aproximar la región espacial de carga por una distribución espacial (Fig. 4d) donde x_p y x_n representan las anchuras de esta región en las zonas P y N respectivamente.

En las regiones de transición, la distribución espacial de carga vendrá dada a partir de la contribución de todas las cargas que intervienen:

$$\frac{d^2\psi}{dx^2} \equiv -\frac{d\varepsilon}{dx} = -\frac{\rho_s}{\epsilon_s} = -\frac{q}{\epsilon_s}(N_D - N_A + p - n)$$

En el caso general de una unión abrupta, en el semiconductor de tipo N pueden coexistir ambos tipos de impurezas cumpliéndose la relación $N_D \gg N_A$ e igualmente en el semiconductor de tipo P se cumplirá $N_A \gg N_D$. En la práctica vamos a suponer que la concentración de impurezasceptoras N_A en el semiconductor de tipo N es nula al igual que la concentración de impurezas dadoras N_D en el de tipo P:

$$\begin{aligned} \frac{d^2\psi}{dx^2} &= -\frac{\rho_s}{\epsilon_s} = \frac{q}{\epsilon_s} \cdot N_A & -x_p \leq x \leq 0 \\ \frac{d^2\psi}{dx^2} &= -\frac{\rho_s}{\epsilon_s} = -\frac{q}{\epsilon_s} \cdot N_D & 0 \leq x \leq x_n \end{aligned}$$

En las expresiones anteriores se ha tenido en cuenta que en la región espacial de carga no existen portadores móviles ($n = p = 0$). Como ya se ha dicho con anterioridad, en realidad existen unas pequeñas concentraciones de portadores móviles -electrones y huecos- despreciables frente a las concentraciones de impurezas.

En regiones alejadas de la unión metalúrgica (zonas neutras de tipo N y P), se mantiene la neutralidad de carga y la densidad espacial de carga total ρ_s es cero. En estas regiones:

$$\begin{aligned} \frac{d^2\psi}{dx^2} &= 0 \\ \text{y} \quad N_D - N_A + p - n &= 0 \end{aligned}$$

Así en la zona neutra de tipo P, suponiendo $N_D = 0$ y teniendo en cuenta que $p \gg n$ se tiene que $p \cong N_A$. Para la zona neutra de tipo N, , suponiendo $N_A = 0$ y teniendo en cuenta que $n \gg p$ se tiene que $n \cong N_D$.

3.1.1.2. Constancia del nivel de Fermi.

Anteriormente se ha visto que, en equilibrio térmico, la corriente neta tanto de electrones como de huecos a través de la unión es cero. En consecuencia, para cada tipo de portador, la corriente de arrastre debida al campo eléctrico debe igualarse a la de difusión causada por el gradiente de concentración. Para la densidad de corriente de huecos en el caso unidimensional se cumplirá que:

$$J_p = J_p(\text{arrastre}) + J_p(\text{difusión})$$

TEMA 3 : Física de la unión P-N.

$$\begin{aligned} &= q \cdot \mu_p \cdot p \cdot \varepsilon - q \cdot D_p \cdot \frac{dp}{dx} \\ &= q \cdot \mu_p \cdot p \cdot \left[\frac{1}{q} \cdot \frac{dE_i}{dx} \right] - K \cdot T \cdot \mu_p \cdot \frac{dp}{dx} = 0 \end{aligned}$$

donde se ha hecho uso de la ecuación del campo eléctrico y de la relación de Einstein:

$$D_p = \frac{K \cdot T}{q} \cdot \mu_p$$

Sustituyendo la expresión para la concentración de huecos:

$$p = n_i \cdot e^{(E_i - E_F)/K \cdot T}$$

y de su derivada:

$$\frac{dp}{dx} = \frac{p}{K \cdot T} \cdot \left[\frac{dE_i}{dx} - \frac{dE_F}{dx} \right]$$

en la ecuación anterior, se obtiene para la densidad de corriente de huecos la expresión:

$$J_p = \mu_p \cdot p \cdot \frac{dE_F}{dx} = 0 \Rightarrow \frac{dE_F}{dx} = 0$$

De forma análoga, para la densidad de corriente de electrones:

$$\begin{aligned} J_n &= J_n(\text{arrastre}) + J_n(\text{difusión}) \\ &= q \cdot \mu_n \cdot n \cdot \varepsilon + q \cdot D_n \cdot \frac{dn}{dx} \\ &= \mu_n \cdot n \cdot \frac{dE_F}{dx} = 0 \end{aligned}$$

La conclusión primera que se obtiene es que, derivado de la condición de flujo neto nulo de portadores (huecos y electrones) en la unión, el nivel de Fermi deberá ser constante (es decir, independiente de x), a lo largo de la muestra semiconductor p-n (Fig. 1b y Fig. 4b).

3.1.1.3. Potencial de contacto.

Tal y como se ha comentado, a través de la unión se establece un campo eléctrico y una diferencia de potencial, y las bandas de energía tanto en la región N como en la P se presentan un desnivel de valor $q \cdot V_{bi}$. También se ha visto como, aunque E_c , E_v y E_i presentan dicho desnivel, el nivel de Fermi E_F permanece constante. En este apartado se va a responder a la pregunta de cual es la magnitud de dicha desnivel o, dicho de otra forma, cual es la diferencia de potencial a través de la unión.

Dicha diferencia de potencial aparece a través de la región espacial de carga, la cual se ha supuesto que está enteramente constituida por átomos dadores o aceptores ionizados. En realidad, hay una pequeña cantidad de portadores de carga (electrones y

TEMA 3 : Física de la unión P-N.

huecos) que están continuamente atravesando dicha región, pero su concentración es despreciable comparada con las concentraciones de impurezas ionizadas.

Como los valores del nivel de Fermi intrínseco (E_i) son diferentes en la zona P y en la zona N mientras que el nivel de Fermi (E_F) es constante en la región espacial de carga se definen los potenciales electrostáticos de las regiones P y N con respecto al nivel de Fermi de la siguiente manera (Fig. 4b):

$$\psi_p \equiv -\frac{1}{q} \cdot (E_i - E_F) \Big|_{x \ll -x_p} \text{ (zona P)}$$

$$\psi_n \equiv -\frac{1}{q} \cdot (E_i - E_F) \Big|_{x \gg x_n} \text{ (zona N)}$$

En la región P, E_i es mayor que E_F por lo que ψ_p es negativo, mientras que, en la región N, E_F está por encima de E_i por lo que ψ_n es positivo.

Para la región neutra de tipo P:

$$p = n_i \cdot e^{(E_i - E_F)/K \cdot T} \Rightarrow \frac{p}{n_i} = e^{(E_i - E_F)/K \cdot T} \Rightarrow E_i - E_F = K \cdot T \cdot \ln \frac{p}{n_i} \cong K \cdot T \cdot \ln \frac{N_A}{n_i}$$

donde se ha hecho uso de que $p = N_A \gg n_i$. Y por tanto:

$$\psi_p \equiv -\frac{1}{q} \cdot (E_i - E_F) \Big|_{x \ll -x_p} = -\frac{K \cdot T}{q} \cdot \ln \left(\frac{N_A}{n_i} \right)$$

De forma similar podemos obtener el potencial electrostático de la región de tipo N con respecto al nivel de Fermi teniendo en cuenta que, para dicha región, $n = N_D \gg n_i$:

$$\psi_n \equiv -\frac{1}{q} \cdot (E_i - E_F) \Big|_{x \gg x_n} = \frac{K \cdot T}{q} \cdot \ln \left(\frac{N_D}{n_i} \right)$$

Como E_F es constante la separación total existente entre los niveles de Fermi intrínsecos en las zonas neutras de tipo P y de tipo N vendrá dado por $\psi_n + |\psi_p|$ y corresponde al valor del potencial de contacto o altura de la barrera de potencial. Las líneas de E_c , E_i y E_v están “curvadas” en una cantidad $q \cdot V_{bi}$.

$$V_{bi} = \psi_n + |\psi_p| = \frac{K \cdot T}{q} \cdot \ln \left(\frac{N_A \cdot N_D}{n_i^2} \right)$$

Fig. 5 muestra las magnitudes de ψ_n y $|\psi_p|$ en función de la concentración de impurezas (N_A o N_D) para el Si y el GaAs. La representación corresponde a una curva que es idéntica para ψ_n y $|\psi_p|$ para un mismo tipo de material semiconductor y a temperatura fija (300K), teniendo en cuenta que para la obtención de ψ_n la concentración de impurezas será la correspondiente a N_D mientras que para la obtención de $|\psi_p|$ la concentración de impurezas será N_A . Como se puede observar el potencial

electrostático para el GaAs es mayor debido a que la concentración intrínseca (n_i) es más pequeña.

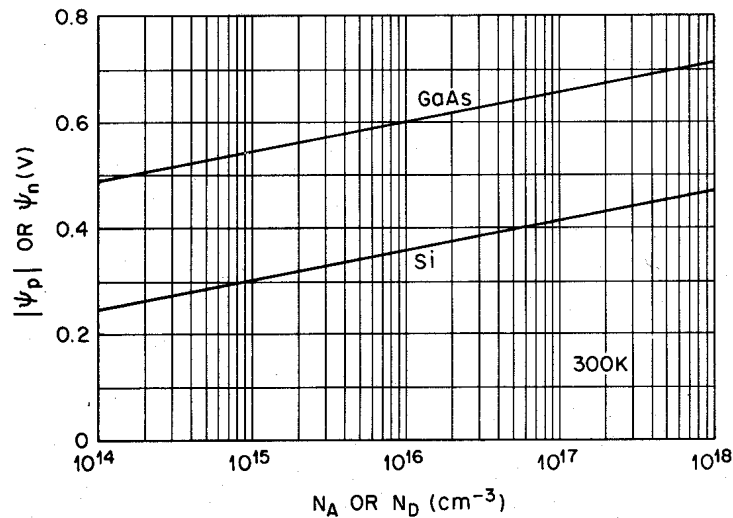


Fig. 5: Potenciales de contacto en la zona P y en la zona N para uniones abruptas de Si y GaAs en función de la concentración de impurezas.

3.2. Resolución de la ecuación de Poisson en la región espacial de carga.

Hasta este momento sólo se ha resuelto el problema electrostático en las regiones neutras alejadas de la unión. Considerando despreciable la región de transición, debe resolverse ahora el problema electrostático (ecuación de Poisson) en las regiones situadas a ambos lados de la unión en las cuales sí que hay una distribución espacial de carga. La ecuación a considerar es la anteriormente escrita:

$$\frac{d^2\psi}{dx^2} = -\frac{q}{\epsilon_s} \cdot (N_D - N_A) = \frac{q}{\epsilon_s} \cdot (N_A - N_D)$$

lo cual obliga a conocer cual es la distribución neta de impurezas $N_D - N_A$. Así, suponiendo $N_D = 0$ en la zona P ($-x_p \leq x \leq 0$) y $N_A = 0$ en la zona N ($0 \leq x \leq x_n$):

$$\frac{d^2\psi}{dx^2} = +\frac{q}{\epsilon_s} \cdot N_A \text{ para } -x_p \leq x \leq 0$$

$$\frac{d^2\psi}{dx^2} = -\frac{q}{\epsilon_s} \cdot N_D \text{ para } 0 \leq x \leq x_n$$

Se van a resolver dos casos importantes, correspondientes al caso de tener una unión abrupta o una unión gradual.

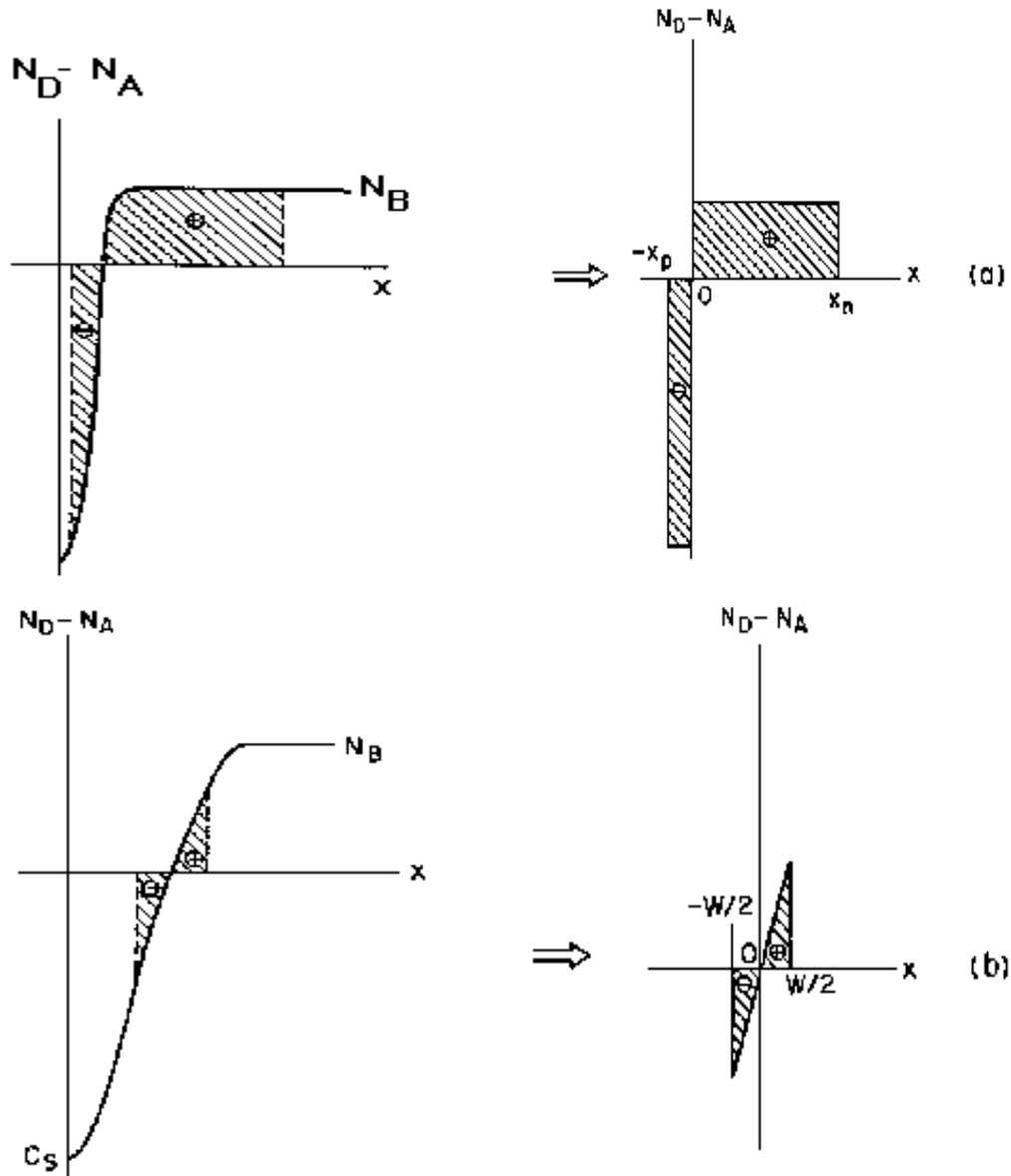


Fig. 6: Perfiles aproximados de dopado. (a) Unión abrupta. (b) Unión gradual.

Fig 6a representa una *unión escalón* o *unión abrupta*, es decir, una unión p-n formada por difusión poco profunda o por implantación de iones de baja energía. La distribución de impurezas puede aproximarse por transiciones abruptas de concentración de impurezas entre las regiones P y N y con valor constante a ambos lados de la unión.

Fig. 6b muestra la aproximación efectuada cuando se trata de uniones p-n que presentan una concentración de impurezas en la región espacial de carga que varía de forma lineal a lo largo de la unión. Este tipo de uniones se obtienen mediante difusiones profundas o técnicas de implantación de iones de alta energía.

Hasta el momento habíamos supuesto siempre, para simplificar que estábamos trabajando con una unión abrupta. En este apartado se analizará también el caso de unión gradual obteniendo las expresiones del campo eléctrico en la unión y el potencial de contacto para este tipo de unión aunque, como se podrá comprobar, el

comportamiento físico interno y las propiedades eléctricas de la unión variarán muy poco.

3.2.1. La unión abrupta.

En la unión abrupta se pasa de una densidad constante de impurezas aceptoras N_A en la zona P a otra de valor constante también N_D de impurezas dadoras en la zona N. La unión abrupta se dice que es simétrica si se cumple la igualdad $N_A = N_D$ y asimétrica o si $N_A \neq N_D$, en cuyo caso, si $N_A \ll N_D$ se dice que la unión es p-n⁺ y si $N_A \gg N_D$ la unión es p⁺-n. Cuando la concentración de impurezas en una zona de la unión es mucho mayor que en la otra (uniones p-n⁺ o p⁺-n) se dice que la unión es *abrupta a un lado*, siendo el caso más frecuente el de una unión p⁺-n.

La unión p-n se caracterizará finalmente especificando su potencial de contacto V_{bi} , la anchura de la región espacial de carga W , el valor máximo del campo eléctrico interno ϵ_{max} y la densidad de carga eléctrica Q .

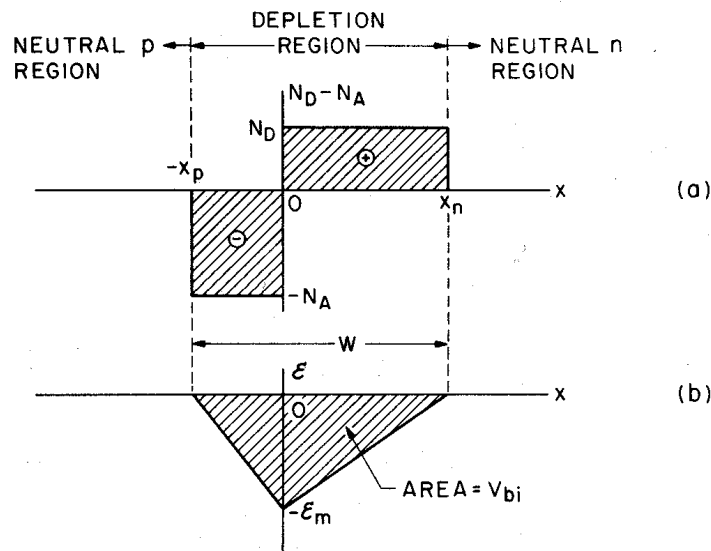


Fig. 7: (a) Distribución espacial de carga en la zona de la unión en equilibrio térmico.
(b) Distribución de campo eléctrico.

Fig. 7a muestra la distribución espacial de carga en una unión abrupta. En la zona espacial de carga no hay portadores móviles y la ecuación de Poisson tiene la forma:

$$\frac{d^2\psi}{dx^2} = +\frac{q}{\epsilon_s} \cdot N_A \text{ para } -x_p \leq x \leq 0 \text{ (zona P)}$$

$$\frac{d^2\psi}{dx^2} = -\frac{q}{\epsilon_s} \cdot N_D \text{ para } 0 \leq x \leq x_n \text{ (zona N)}$$

La condición de neutralidad de carga en el conjunto del semiconductor requiere que el área positiva y el área negativa encerradas por la distribución espacial de carga sea la misma (Fig. 7a).

$$N_A \cdot x_p = N_D \cdot x_n$$

La anchura W de la región espacial de carga es:

TEMA 3 : Física de la unión P-N.

$$W = x_p + x_n$$

El campo eléctrico mostrado en Fig. 7b se obtiene a partir de la definición del potencial electrostático de la siguiente forma:

Para $-x_p \leq x \leq 0$ (zona P):

$$\varepsilon \equiv -\frac{d\psi}{dx} = -\int_{-x_p}^x \frac{q}{\epsilon_s} \cdot N_A \cdot dx = (-) \frac{q \cdot N_A \cdot (x + x_p)}{\epsilon_s} + cte$$

Sabiendo que $\varepsilon(-x_p) = 0 \Rightarrow cte = 0$, por lo que:

$$\varepsilon(x) = (-) \frac{q \cdot N_A \cdot (x + x_p)}{\epsilon_s}; \quad -x_p \leq x \leq 0 \text{ (zona P)}$$

Para $0 \leq x \leq x_n$ (zona N):

$$\varepsilon(x) \equiv -\frac{d\psi}{dx} = -\int_{x_n}^x (-) \frac{q}{\epsilon_s} \cdot N_D \cdot dx = \frac{q \cdot N_D \cdot (x - x_n)}{\epsilon_s} + cte$$

Sabiendo que $\varepsilon(x_n) = 0 \Rightarrow cte = 0$, por lo que:

$$\varepsilon(x) = \frac{q \cdot N_D \cdot (x - x_n)}{\epsilon_s}; \quad 0 \leq x \leq x_n \text{ (zona N):}$$

El campo eléctrico tiene un valor máximo ε_m que está localizado en $x = 0$ que viene dado por la expresión:

$$\varepsilon_m = \frac{q \cdot N_D \cdot x_n}{\epsilon_s} = \frac{q \cdot N_A \cdot x_p}{\epsilon_s}$$

En función de este valor máximo el campo eléctrico en la zona N de la región espacial de carga puede reescribirse de la siguiente manera:

$$\varepsilon(x) \equiv (-)\varepsilon_m + \frac{q \cdot N_D \cdot x}{\epsilon_s}; \text{ para } 0 \leq x \leq x_n \text{ (zona N)}$$

Este campo eléctrico es negativo pues tal y como dicta la expresión $\varepsilon \equiv -\frac{d\psi}{dx}$, tiene sentido contrario al de aumento del potencial electrostático (Fig. 2).

Para obtener el potencial de contacto V_{bi} entre las dos regiones de distribución espacial de carga, se ha de integrar el campo eléctrico entre los valores $-x_p$ y x_n :

$$\begin{aligned} V_{bi} &= (-) \int_{-x_p}^{x_n} \varepsilon(x) \cdot dx = (-) \int_{-x_p}^0 \varepsilon(x) \cdot dx \Big|_{cara p} - \int_0^{x_n} \varepsilon(x) \cdot dx \Big|_{cara n} = \\ &= (-) \int_{-x_p}^0 \frac{q \cdot N_A}{\epsilon_s} \cdot (x + x_p) \cdot dx - \int_0^{x_n} \frac{q \cdot N_D}{\epsilon_s} \cdot (x - x_n) \cdot dx = \\ &= \frac{q \cdot N_A \cdot x_p^2}{2 \cdot \epsilon_s} + \frac{q \cdot N_D \cdot x_n^2}{2 \cdot \epsilon_s} = \frac{1}{2} \cdot \varepsilon_m \cdot W \end{aligned}$$

TEMA 3 : Física de la unión P-N.

Es decir, el potencial de contacto corresponde al área encerrada del campo eléctrico (área del triángulo).

$$V_{bi} = \frac{1}{2} \cdot \epsilon_m \cdot W$$

Combinando las expresiones correspondientes a la neutralidad de carga y del potencial de contacto se obtiene la anchura de la región espacial de carga W.

$$\begin{aligned} V_{bi} &= \frac{q \cdot N_A}{2 \cdot \epsilon_s} \cdot x_p^2 + \frac{q \cdot N_D}{2 \cdot \epsilon_s} \cdot x_n^2 = \frac{q}{2 \cdot \epsilon_s} \cdot (N_A \cdot x_p \cdot x_p + N_D \cdot x_n \cdot x_n) = \frac{q}{2 \cdot \epsilon_s} \cdot N_A \cdot x_p \cdot (x_p + x_n) = \\ &= \frac{q}{2 \cdot \epsilon_s} \cdot N_A \cdot x_p \cdot \frac{(x_p + x_n)}{(x_p + x_n)} \cdot (x_p + x_n) = \frac{q \cdot N_A}{2 \cdot \epsilon_s} \cdot (x_p + x_n)^2 \cdot \frac{x_p}{(x_p + x_n)} = \frac{q \cdot N_A}{2 \cdot \epsilon_s} \cdot W^2 \cdot \frac{1}{1 + \frac{x_n}{x_p}} = \\ &= \frac{q \cdot N_A}{2 \cdot \epsilon_s} \cdot W^2 \cdot \frac{1}{1 + \frac{N_A}{N_D}} = \frac{q}{2 \cdot \epsilon_s} \cdot W^2 \cdot \frac{N_A \cdot N_D}{N_D + N_A} \end{aligned}$$

y despejando W:

$$W = \sqrt{\frac{2 \cdot \epsilon_s}{q} \cdot \left[\frac{N_A + N_D}{N_A \cdot N_D} \right] \cdot V_{bi}} = \sqrt{\frac{2 \cdot \epsilon_s}{q} \cdot \left[\frac{1}{N_A} + \frac{1}{N_D} \right] \cdot V_{bi}}$$

Fijarse que la anchura de la región espacial de carga (W) disminuye al aumentar la concentración de sustancias dopantes (N_A y N_D).

Fig. 8b representa la distribución de carga para una unión abrupta a un lado p⁺-n en la que se cumple que $N_A \gg N_D$.

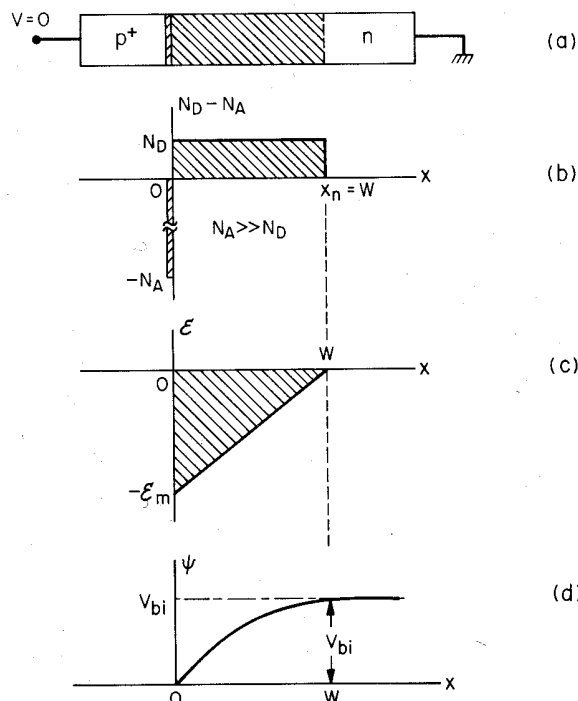


Fig. 8: (a) Unión abrupta a un lado (con $N_A \gg N_D$) en equilibrio térmico. (b) Distribución espacial de carga. (c) Distribución de campo eléctrico. (d) Potencial electrostático en función de la distancia, donde V_{bi} es el potencial de contacto.

TEMA 3 : Física de la unión P-N.

En este caso para que se mantenga la neutralidad de la carga, la anchura de la región espacial de carga en la zona P debe ser mucho menor que la de la zona N ($N_A \cdot x_p = N_D \cdot x_n$ con $N_A \gg N_D \Rightarrow x_p \ll x_n$), y en consecuencia, la anchura total W de la región espacial de carga puede simplificarse como:

$$W \approx x_n = \sqrt{\frac{2 \cdot \epsilon_s \cdot V_{bi}}{q \cdot N_D}}$$

donde se ha hecho uso de que:

$$\frac{N_A + N_D}{N_A \cdot N_D} \cong \frac{1}{N_D}$$

La expresión del campo eléctrico en función de x es la misma que para el caso anterior en la región N:

$$\epsilon(x) = (-)\epsilon_m + \frac{q \cdot N_B \cdot x}{\epsilon_s} \quad 0 \leq x \leq W$$

donde la cantidad N_B se denomina concentración ligera de impurezas (concentración de impurezas que está en menor cantidad), la cual en el caso de una unión p⁺-n es N_D . El valor máximo del campo:

$$\epsilon_m = \frac{q \cdot N_B \cdot W}{\epsilon_s}$$

Y en consecuencia:

$$\epsilon(x) = (-)\epsilon_m \cdot \left[1 - \frac{x}{W} \right] \quad 0 \leq x \leq W$$

El cual está representado en Fig. 8c. Integrando el campo eléctrico se obtiene la distribución de potencial electrostático $\psi(x)$.

$$\psi(x) = (-)\int_0^x \epsilon(x) \cdot dx = \epsilon_m \left[x - \frac{x^2}{2 \cdot W} \right] + cte \quad 0 \leq x \leq W$$

Tomando como origen de potenciales el potencial nulo existente en la región neutra de tipo P, $\psi(0) = 0 \Rightarrow cte = 0$; y empleando la relación para el potencial de contacto V_{bi}

$$V_{bi} = \frac{1}{2} \cdot \epsilon_m \cdot W \Rightarrow \epsilon_m = \frac{2 \cdot V_{bi}}{W}$$

se puede obtener que:

$$\psi(x) = \frac{2 \cdot V_{bi}}{W} \cdot \left[x - \frac{x^2}{2 \cdot W} \right] = \frac{V_{bi} \cdot x}{W} \cdot \left[2 - \frac{x}{W} \right]$$

que es la distribución de potencial representada en Fig. 8b. Fijarse que en $x = W$, $\psi(W) = V_{bi}$.

3.2.2. La unión gradual.

En este caso, la distribución de impurezas en la región espacial de carga aparece en Fig. 9a.

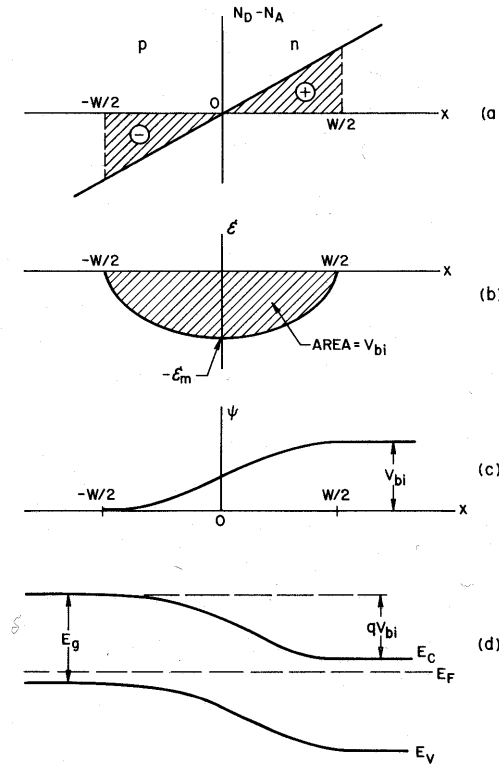


Fig. 9: Unión gradual en equilibrio térmico. (a) Distribución espacial de carga. (b) Distribución de campo eléctrico. (c) Potencial electrostático en función de la distancia. (d) Diagrama de bandas de energía.

La variación espacial que experimenta la concentración neta de impurezas es de forma lineal:

$$N_D - N_A = a \cdot x$$

donde $a \equiv$ Gradiente de impurezas $= \frac{d(N_D - N_A)}{dx}$ [cm^{-4}]

Multiplicando la concentración neta de impurezas en la región espacial de carga por la carga del electrón obtenemos la distribución espacial de carga:

$$\rho_s \equiv q \cdot (N_D - N_A) = q \cdot a \cdot x \quad \text{para } -\frac{W}{2} \leq x \leq \frac{W}{2}$$

$$x < 0 \Rightarrow \rho_s < 0$$

$$x > 0 \Rightarrow \rho_s > 0$$

La condición de neutralidad de carga en el conjunto del semiconductor requiere que el área positiva y el área negativa encerradas por la distribución espacial de carga sea la misma por lo que si W es la anchura de la región espacial de carga, ésta estará igualmente repartida en la zona N y en la zona P ($x_n = W/2$ y $-x_p = -W/2$).

Para este tipo de distribución de impurezas, la ecuación de Poisson es:

TEMA 3 : Física de la unión P-N.

$$\frac{d^2\psi}{dx^2} = (-)\frac{d\varepsilon}{dx} = (-)\frac{\rho_s}{\epsilon_s} = (-)\frac{q}{\epsilon_s} \cdot a \cdot x \quad \text{para } -\frac{W}{2} \leq x \leq \frac{W}{2}$$

siendo **a** el gradiente de impurezas (en cm^{-4}) y **W** la anchura de la región espacial de carga.

Integrando una vez para obtener el campo eléctrico, se obtiene la distribución $\varepsilon(x)$ para este campo, representada en Fig. 9b:

$$\begin{aligned} \varepsilon(x) \equiv \frac{d\psi}{dx} &= -\int_{-\frac{W}{2}}^x (-)\frac{q}{\epsilon_s} \cdot a \cdot x \cdot dx = \frac{q \cdot a}{\epsilon_s} \cdot \left[\frac{x^2}{2} \right]_{-\frac{W}{2}}^x + cte = \frac{q \cdot a}{\epsilon_s} \cdot \left[\frac{x^2 - \left(-\frac{W}{2}\right)^2}{2} \right] + cte = \\ &= -\frac{q \cdot a}{\epsilon_s} \cdot \left[\frac{\left(\frac{W}{2}\right)^2 - x^2}{2} \right] + cte \quad \text{para } -\frac{W}{2} \leq x \leq \frac{W}{2} \end{aligned}$$

Con las condiciones de contorno de que el campo en $-\frac{W}{2}$ y $\frac{W}{2}$ ha de ser cero:

$$\varepsilon\left(-\frac{W}{2}\right) = \varepsilon\left(\frac{W}{2}\right) = 0 \Rightarrow cte = 0$$

Con lo que:

$$\varepsilon(x) = -\frac{q \cdot a}{\epsilon_s} \cdot \left[\frac{\left(\frac{W}{2}\right)^2 - x^2}{2} \right] \quad \text{para } -\frac{W}{2} \leq x \leq \frac{W}{2}$$

cuyo valor máximo está localizado en $x = 0$ y tiene un valor (haciendo $x = 0$ en la anterior expresión):

$$\varepsilon_m = \frac{q \cdot a \cdot W^2}{8 \cdot \epsilon_s}$$

Integrando nuevamente se pueden obtener por un lado la distribución de potencial electrostático y, por otro lado, la disposición de bandas de energía (Fig. 9c y 9d respectivamente).

$$\psi(x) = -\int_{-\frac{W}{2}}^x \varepsilon(x) \cdot dx \quad \text{para } -\frac{W}{2} \leq x \leq \frac{W}{2}$$

con la condición de contorno $\psi\left(-\frac{W}{2}\right) = 0$.

El potencial de contacto viene dado por la siguiente expresión:

$$V_{bi} = -\int_{-\frac{W}{2}}^{\frac{W}{2}} \varepsilon(x) \cdot dx \Rightarrow V_{bi} = \frac{q \cdot a \cdot W^3}{12 \cdot \epsilon_s}$$

Y la anchura de la región espacial de carga (despejando de la anterior ecuación):

$$W = \left(\frac{12 \cdot \epsilon_s \cdot V_{bi}}{q \cdot a} \right)^{1/3}$$

Utilizando la expresión definitoria del potencial de contacto V_{bi} (diferencia de potencial electrostático entre las zonas neutras) y para el valor de N_A y N_D en los extremos de la región espacial de carga ($-\frac{W}{2}$ y $\frac{W}{2}$), que es $a \cdot \frac{W}{2}$ en ambos casos, se obtiene la expresión:

$$V_{bi} = \frac{K \cdot T}{q} \cdot \ln \frac{N_A \cdot N_D}{n_i^2} = \frac{K \cdot T}{q} \cdot \ln \left[\frac{a \cdot W}{2} \cdot \frac{a \cdot W}{2} \cdot \frac{1}{n_i^2} \right] = \frac{2 \cdot K \cdot T}{q} \cdot \ln \left[\frac{a \cdot W}{2 \cdot n_i} \right]$$

Si se elimina W en esta última expresión empleando la anterior para W se obtiene una ecuación trascendente que permite resolver el potencial de contacto V_{bi} en función de a . Los resultados para el Si y el GaAs están reflejados en Fig. 10.

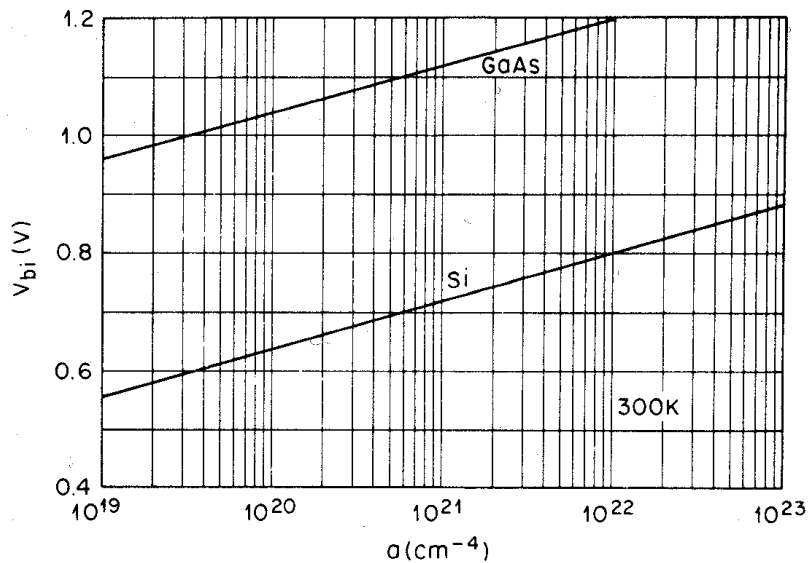


Fig. 10: Potencial de contacto para uniones graduales de Si y GaAs en función del gradiente de impurezas.

Como se ha podido comprobar, el comportamiento físico interno y las propiedades eléctricas de la unión variarán muy poco de considerar una unión abrupta a considerar una unión gradual.

3.3. La unión P-N polarizada: diodo de unión.

Hasta ahora sólo se ha considerado la unión p-n sin ningún tipo de polarización (Fig. 11a). Sin embargo, cuando la unión es sometida a una diferencia de potencial externa supuesta constante aparece un nuevo comportamiento. Se tienen dos casos a considerar, según la tensión externa aplicada sea en sentido directo o inverso.

Polarización directa: En este caso, la parte P se somete a un potencial positivo respecto de la parte N, $V_F > 0$. El efecto es que el potencial electrostático de contacto en la unión decrece en dicha cantidad V_F , es decir, $V_{bi} - V_F$, reduciéndose además la región espacial de carga, cuya anchura depende de la diferencia de potencial a la que está sometida la unión (Fig. 11b). Los niveles energéticos en el semiconductor de tipo N se desplazan hacia arriba en una cantidad $q \cdot V_F$, que es la misma diferencia que existe entre los niveles de Fermi de los semiconductores de tipo N y P al existir una diferencia de potencial entre ellos de V_F .

La reducción de la barrera de potencial cuando la polarización es directa hace que la corriente de difusión predomine sobre la de arrastre. Esto permite que los portadores mayoritarios atraviesen la unión, convirtiéndose entonces en minoritarios y dando lugar a una corriente mensurable. Se tiene, por tanto, una inyección de portadores minoritarios, es decir, de electrones en la zona P y de huecos en la N.

Polarización inversa: En este caso, la parte N tiene una tensión positiva aplicada $V_R > 0$ con respecto de la P. A diferencia del anterior caso, el potencial electrostático se ve aumentado en la cantidad V_R , es decir, $V_{bi} + V_R$, aumentando la anchura de la región espacial de carga (Fig. 11c). El consiguiente aumento del campo eléctrico en la región espacial de carga hace que las corrientes de difusión se vean muy reducidas, resultando únicamente una pequeña corriente inversa de arrastre.

Sustituyendo dichos valores de tensión en la unión $V_j = V_{bi} - V$ donde V es positivo para el caso de polarización directa (V_F) y negativo para el caso de polarización inversa ($-V_R$), en lugar de V_{bi} en la expresión que nos da la anchura de la región espacial de carga en el equilibrio podemos obtener una nueva expresión para W , ahora ya fuera del equilibrio, en el caso de una unión abrupta:

$$W = \sqrt{\frac{2 \cdot \epsilon_s}{q} \cdot \left(\frac{N_A + N_D}{N_A \cdot N_D} \right) \cdot (V_{bi} - V)}$$

Darse cuenta de que la anchura de la región espacial de carga W varía con la raíz cuadrada de la diferencia de potencial electrostático total a través de la unión, $V_j = V_{bi} - V$.

Otra relación importante es con el dopado de ambas zonas. Cuanto mayor sea el dopado de ambas zonas (N_A y N_D), más estrecha será la región espacial de carga (esto es cierto también en ausencia de polarización externa).

Si se trata de una unión p^+n ($N_A \gg N_B = N_D$):

$$W \approx x_n = \sqrt{\frac{2 \cdot \epsilon_s \cdot (V_{bi} - V)}{q \cdot N_B}}$$

En este caso es la concentración ligera de impurezas la que determina la anchura de la región espacial de carga, disminuyendo dicha anchura al aumentar N_D .

Para una unión gradual:

$$W = \left(\frac{12 \cdot \epsilon_s \cdot (V_{bi} - V)}{q \cdot a} \right)^{1/3}$$

En cuanto a la relación entre el valor máximo del campo eléctrico en la zona de la unión, en el caso de una unión abrupta, y la anchura de la región espacial de carga se verá modificada a:

$$V_j = V_{bi} - V = \frac{1}{2} \cdot \epsilon_m \cdot W$$

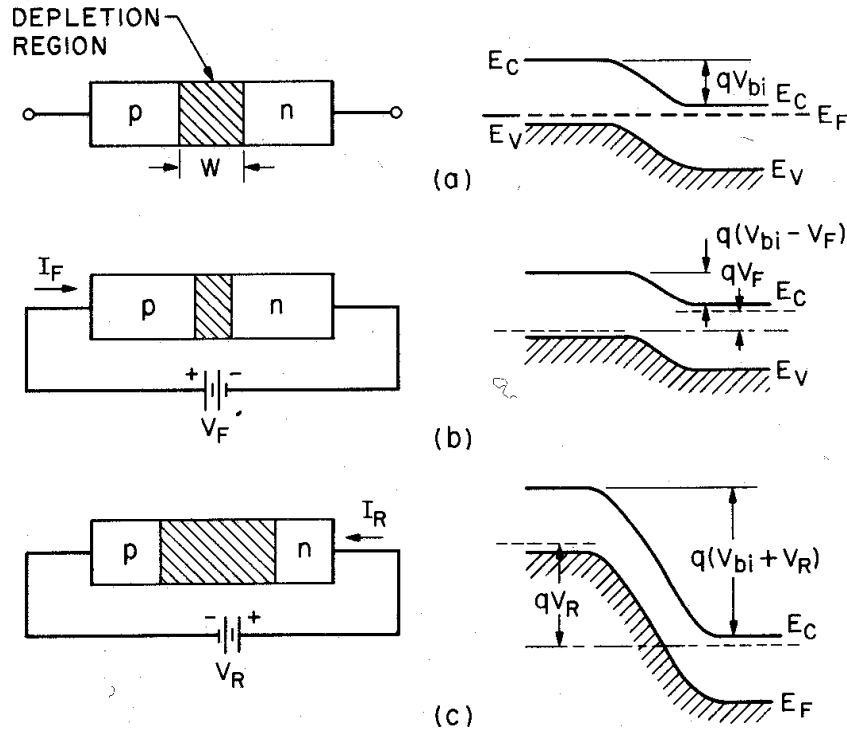


Fig. 11: Representación esquemática de la anchura de la región espacial de carga y de los diagramas de bandas de la unión p-n bajo distintas condiciones de polarización. (a) En equilibrio térmico. (b) En polarización directa. (c) En polarización inversa.

3.3.1. Capacidad de transición de la unión polarizada.

Hay un parámetro de tipo capacitivo que se puede obtener considerando cual es la variación de la carga en la región espacial de carga por unidad de superficie cuando se realiza un incremento de potencial en la unión p-n.

Este parámetro recibe el nombre de capacidad de la región espacial de carga ó capacidad de transición, C_j bajo polarización inversa.

Fig. 12 muestra dicha capacidad para una unión p-n con una distribución arbitraria de impurezas. La línea continua corresponde a las distribuciones de carga y campo eléctrico cuando una tensión V es aplicada en la zona N. Si esta tensión se incrementa en una cantidad dV , las distribuciones de carga y campo eléctrico se expanden hasta las regiones delimitadas por la línea discontinua. En Fig. 12b la carga incremental dQ corresponde al área marcada entre las dos curvas de distribución de carga a cada lado de la región espacial de carga. Las cargas incrementales en las zonas N y P de la región espacial de carga son de igual magnitud pero con polaridad de la carga opuesta, pues se ha de mantener la neutralidad de carga.

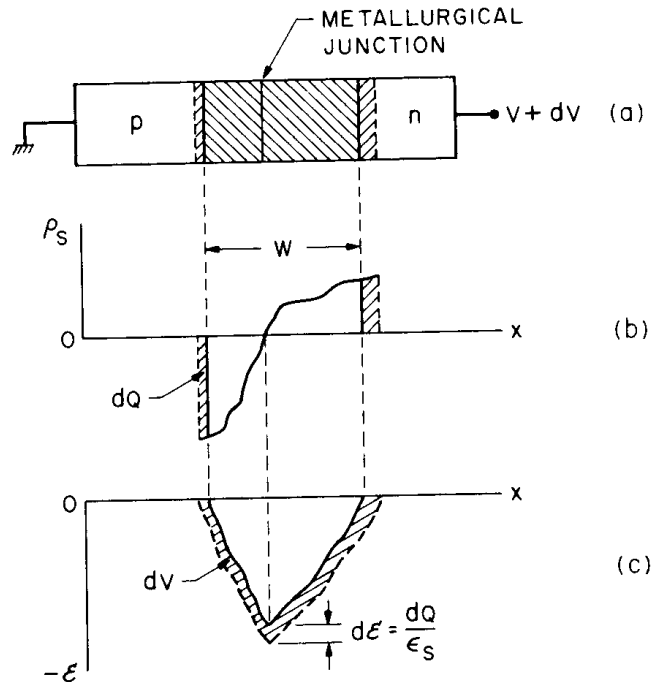


Fig. 12: (a) Unión p-n con un perfil de impurezas arbitrario bajo polarización inversa. (b) Cambio en la distribución espacial de carga debida al cambio en la tensión de polarización. (c) Cambio correspondiente en la distribución de campo eléctrico.

El cambio en la distribución de carga dQ provoca un incremento en el campo eléctrico de $d\mathcal{E} = dQ/\epsilon_s$ (obtenido a partir de la ecuación de Poisson). El correspondiente cambio en la tensión aplicada dV , representado por el área rayada en Fig. 12.c, es aproximadamente $W \cdot d\mathcal{E}$, el cual es igual a $W \cdot dQ/\epsilon_s$. Por lo tanto, la capacidad de transición por unidad de área viene dada por la expresión:

$$C_j \equiv \frac{dQ}{dV} = \frac{dQ}{W \cdot d\mathcal{E}} = \frac{dQ}{W \cdot \frac{dQ}{\epsilon_s}} = \frac{\epsilon_s}{W} \text{ (F/cm}^2\text{)}$$

Esta expresión para la capacidad de la zona de carga espacial tiene la misma expresión que la de la capacidad de un sistema de placas plano-paralelas, donde la distancia entre las dos placas es la anchura de la región espacial de carga. Esta expresión es válida cualquier arbitraria distribución de impurezas.

Ha de tenerse en cuenta además que se ha considerado que sólo interviene en el conjunto de la capacidad, la distribución espacial de carga en la región de carga espacial. Ello sólo ocurre si la **polarización** es **inversa**.

En el caso de que se tenga una polarización directa, la variación de la distribución de carga en la zona espacial de carga tendrá una contribución adicional procedente de los portadores mayoritarios de carga en cada una de las regiones P y N pues este tipo de polarización disminuye la barrera de potencial para los portadores mayoritarios contribuyendo a un incremento adicional en la corriente de difusión. Dicho fenómeno contribuye a la llamada capacidad de difusión que será tratada con posterioridad.

Utilizando la expresión anterior y la dada para la anchura W en términos de la tensión externa V , se pueden obtener las siguientes expresiones para la capacidad de transición:

TEMA 3 : Física de la unión P-N.

- Si la unión es abrupta:

$$C_j = \sqrt{\frac{\epsilon_s \cdot q}{2 \cdot (V_{bi} - V)} \cdot \frac{N_A \cdot N_D}{(N_A + N_D)}}$$

Aquí se ha empleado la relación:

$$W = \sqrt{\frac{2 \cdot \epsilon_s}{q} \cdot \left(\frac{N_A + N_D}{N_A \cdot N_D} \right) \cdot (V_{bi} - V)}$$

para la anchura de la región espacial de carga fuera del equilibrio.

- Si la unión es gradual:

$$C_j = \left[\frac{q \cdot a \cdot \epsilon_s^2}{12 \cdot (V_{bi} - V)} \right]^{1/3}$$

donde se ha empleado la relación:

$$W = \left[\frac{12 \cdot \epsilon_s \cdot (V_{bi} - V)}{q \cdot a} \right]^{1/3}$$

para la anchura de la región espacial de carga fuera del equilibrio.

Dado que la capacidad de transición obtenida con estas ecuaciones lo es por unidad de área, para obtener la capacidad de transición total habría que multiplicar por la sección transversal del diodo (A):

$$C_{j_total} = A \cdot \frac{\epsilon_s}{W} (F)$$

Muchos circuitos emplean esta propiedad de cambio de la capacidad de transición de una unión p-n polarizada inversamente con la tensión. La unión p-n utilizada con este propósito recibe el nombre de varactor, el cual es una palabra procedente del término reactancia variable. El varactor es un dispositivo particularmente útil en la sintonización de receptores de FM.

3.4. Modelo matemático de la unión P-N: ecuación de Shockley.

Cuando una tensión es aplicada a la unión p-n, las corrientes de difusión y de arrastre se ven alteradas rompiéndose la situación existente en el equilibrio (Fig. 13).

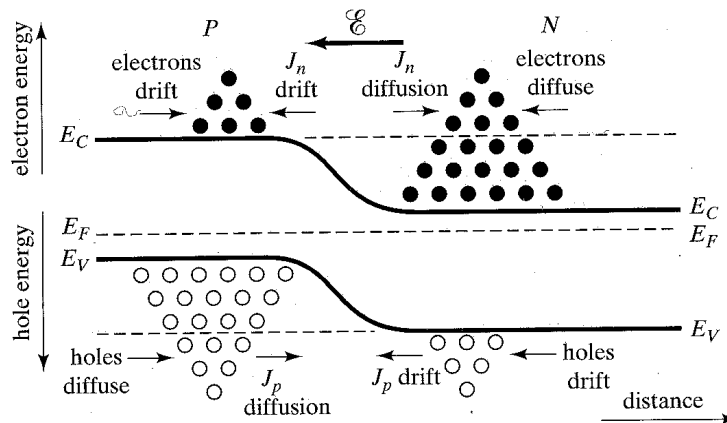


Fig. 13: Movimiento de los portadores en equilibrio.

Cuando una tensión de polarización directa V_a se aplica a la unión, la caída de tensión a través de la unión y el campo eléctrico en ésta se reducen. Si suponemos que los niveles energéticos en la zona P permanecen fijos, entonces los niveles energéticos en la zona N aumentan en $q \cdot V_a$, como se muestra en Fig. 14.

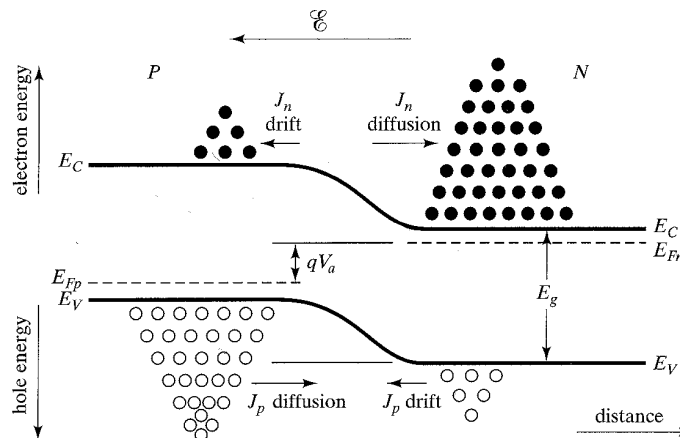


Fig. 14: Movimiento de los portadores con polarización directa ($V_a > 0$).

El número de electrones, portadores minoritarios en la zona P, por encima de E_c es el mismo que en la situación de equilibrio. Estos constituyen la corriente de arrastre de electrones. El número de electrones en la zona N que tiene energías por encima de E_c de la zona P es considerablemente mayor que en equilibrio. Estos electrones se difundirán hasta P. Por tanto, se tiene una corriente de difusión de electrones a través de la unión que es mayor que la corriente de arrastre causada por los electrones que se desplazan desde P hasta N. De forma análoga, los huecos (con energía superior a E_v de la zona N) se difundirán desde P hasta N y esta corriente de difusión será mucho mayor que la de arrastre debida al movimiento de los huecos desde N hasta P.

Cuando se aplica una tensión de polarización inversa a la unión ($V_a < 0$), las corrientes de arrastre no cambian ya que la concentración de portadores minoritarios no ha cambiado. Pero las corrientes debidas a la difusión son pequeñas debido a que hay un menor número de portadores mayoritarios, electrones en la zona N y huecos en la zona P, con energías superiores a E_c de la zona P y E_v de la zona N respectivamente. Esto se muestra en Fig. 15:

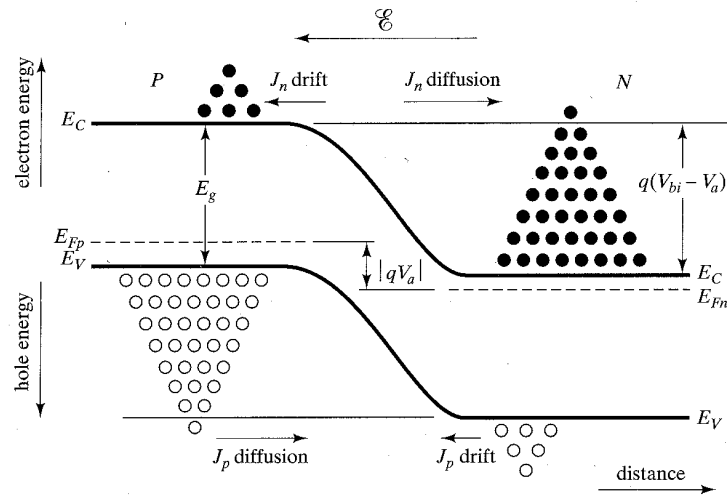


Fig. 15: Movimiento de los portadores con polarización inversa ($V_a < 0$).

Dicho de otra manera, los electrones en la zona N y los huecos en la P tienen una barrera energética mayor que superar, comparada con la existente en la condición de equilibrio, de manera que la corriente neta es negativa. Esta corriente es negativa porque está compuesta por un número mayor de electrones cruzando desde P hasta N (arrastre) que cruzando desde N hasta P (difusión) y, de forma similar, un número mayor de huecos cruzando desde N a P que desde P a N, comparando con la situación de equilibrio.

Por tanto, la corriente está limitada a la posibilidad de generar portadores minoritarios térmicamente y es, de hecho, independiente de la tensión de polarización inversa para decenas de voltios o superior.

3.4.1. Características ideales.

A continuación se va a desarrollar un modelo matemático de la unión p-n en ausencia de equilibrio y que explica el comportamiento interno del dispositivo. Finalmente, el modelo matemático dará lugar a las características macroscópicas eléctricas corriente-tensión.

Para la obtención del modelo matemático se realizarán las siguientes hipótesis:

1. La región espacial de carga tiene límites abruptos (modelo de *unión abrupta*).
2. No existen portadores en la *región espacial de carga*, estos sólo la atraviesan (en la práctica las concentraciones de portadores móviles -electrones y huecos- que atraviesan la unión son despreciables frente a las concentraciones de impurezas). Por tanto, se supondrá que esta región consta *únicamente de impurezas ionizadas*.
3. *Fuera* de los límites de la región espacial de carga, el semiconductor es *neutro*.
4. El funcionamiento es a una temperatura tal que todos los *átomos de impurezas* están *ionizados*.
5. Las concentraciones de portadores en los límites de la región espacial de carga vienen dados en función del potencial electrostático de contacto V_{bi} .
6. Los contactos al final de las regiones P y N son *contactos* perfectamente *óhmicos*. Un contacto perfectamente óhmico es aquel que tiene una resistencia cero de manera que la caída de tensión entre sus extremos es cero.

7. Cuando el diodo está polarizado directamente, se supone *condición de baja inyección*. Esto significa que cuando los electrones desde N o los huecos desde P son inyectados en las regiones opuestas, la concentración de dichos portadores en el límite con la región espacial de carga de la nueva región, donde estos ahora son portadores minoritarios, es mucho menor que la concentración de portadores mayoritarios en el equilibrio en dicha región.
En otras palabras, la concentración de portadores mayoritarios en los contactos con las zonas neutras no sufre cambio apreciable por el hecho de aplicar una polarización.
8. Se considerarán inexistentes los procesos de recombinación y de generación de portadores en la región espacial de carga. En consecuencia, las corrientes de huecos y de electrones serán constantes a lo largo de la región espacial de carga.

Concentraciones de portadores en los límites de la región espacial de carga.

Se pretende calcular, en primer lugar, las concentraciones de portadores en los límites de la región espacial de carga (es decir en x_n y $-x_p$). El cálculo de estas concentraciones nos permitirá obtener condiciones de contorno para la resolución de la ecuación de continuidad en las regiones neutras y, por tanto, obtener la variación espacial que experimentan las concentraciones de portadores minoritarios en las regiones neutras.

Así por ejemplo, en condiciones de polarización directa se produce una inyección de portadores minoritarios, de electrones en la zona P y de huecos en la zona N. La obtención de la concentración de electrones en $-x_p$ ($n_p(-x_p)$) y de huecos en x_n ($p_n(x_n)$) nos proporciona la cantidad de electrones inyectados en la zona P y de huecos inyectados en la zona N. Estos valores serán condiciones de contorno para resolver la ecuación de continuidad en cada una de las regiones neutras, en las cuales tendrá lugar un proceso de difusión-recombinación de los portadores minoritarios inyectados, tal y como sucedía en el caso de inyección de portadores desde una cara (visto en Tema 2).

En condiciones de equilibrio térmico (ausencia de polarización externa), las concentraciones de portadores mayoritarios se iguala a las de sustancias dopantes. Llamaremos por un lado n_{no} y n_{po} a la concentración de electrones en las regiones N y P respectivamente en equilibrio térmico, y por otro lado p_{no} y p_{po} a las de huecos en las regiones N y P respectivamente, también bajo la misma situación de equilibrio térmico. Haciendo uso de la condición de ionización total en el semiconductor de tipo N se tendrá $n_{no} = N_D$, mientras que en el semiconductor de tipo P $p_{po} = N_A$.

Bajo esta situación, recordando la expresión obtenida para el potencial electrostático de contacto V_{bi} :

$$V_{bi} = \frac{K \cdot T}{q} \cdot \ln \left[\frac{N_A \cdot N_D}{n_i^2} \right] \approx \frac{K \cdot T}{q} \cdot \ln \left[\frac{p_{po} \cdot n_{no}}{n_i^2} \right] = \frac{K \cdot T}{q} \cdot \ln \left[\frac{n_{no}}{n_{po}} \right]$$

donde se ha hecho uso de la ley de acción de masas en el equilibrio térmico:

$$p_{po} \cdot n_{po} = p_{no} \cdot n_{no} = n_i^2$$

La ecuación anterior de V_{bi} también se puede escribir de la forma:

TEMA 3 : Física de la unión P-N.

$$n_{no} = n_{po} \cdot e^{qV_{bi}/K \cdot T}$$

expresión que nos da la concentración de electrones a ambos lados de la unión en términos del potencial electrostático de contacto en equilibrio térmico.

De forma análoga, también se obtiene para la concentración de huecos a ambos lados de la unión en términos del potencial electrostático de contacto en equilibrio térmico:

$$V_{bi} = \frac{K \cdot T}{q} \cdot \ln \left[\frac{N_A \cdot N_D}{n_i^2} \right] \approx \frac{K \cdot T}{q} \cdot \ln \left[\frac{p_{po} \cdot n_{no}}{n_i^2} \right] = \frac{K \cdot T}{q} \cdot \ln \left[\frac{p_{po}}{p_{no}} \right]$$

$$p_{po} = p_{no} \cdot e^{qV_{bi}/K \cdot T}$$

A partir de la suposición número cinco realizada, se puede esperar que, bajo condiciones de polarización, la expresión que nos relaciona las concentraciones de electrones a ambos lados de la región espacial de carga (x_n y $-x_p$) sea similar a la obtenida en equilibrio térmico, considerando ahora $V_j = V_{bi} - V$, en lugar de V_{bi} . Es decir:

$$n_n = n_p \cdot e^{q(V_{bi}-V)/K \cdot T}$$

donde ahora n_n es la concentración de electrones en $x = x_n$ y n_p la concentración de electrones en $x = -x_p$ ya fuera del equilibrio, con V representando la tensión de polarización (directa o inversa, es decir $V_{bi}-V_F$ ó $V_{bi}+V_R$), fuera del equilibrio.

Suponiendo ahora que se cumple la condición de baja inyección de portadores minoritarios y, en consecuencia, que la concentración de portadores minoritarios inyectados es mucho más pequeña que la de portadores mayoritarios, se cumplirá que:

$$n_n \approx n_{no}$$

Utilizando esta nueva condición, junto con la expresión obtenida anteriormente para n_{no} se obtiene la concentración de electrones en los límites de la región de distribución de carga espacial con la zona P ($x = -x_p$):

$$n_p = n_{po} \cdot e^{qV/K \cdot T} \text{ para } x = -x_p$$

O bien:

$$n_p - n_{po} = n_{po} \cdot (e^{qV/K \cdot T} - 1) \text{ para } x = -x_p$$

De forma similar se puede actuar para obtener la concentración de huecos en la zona N en el límite de la región de carga espacial con la zona N ($x = x_n$). A partir de la suposición número cinco realizada, se puede esperar que, bajo condiciones de polarización, la expresión que nos relaciona las concentraciones de huecos a ambos lados de la región espacial de carga (x_n y $-x_p$) sea similar a la obtenida en equilibrio térmico, considerando ahora $V_j = V_{bi} - V$, en lugar de V_{bi} . Es decir:

$$p_p = p_n \cdot e^{q(V_{bi}-V)/K \cdot T}$$

TEMA 3 : Física de la unión P-N.

donde ahora p_p es la concentración de huecos en $x = -x_p$ y p_n la concentración de huecos en $x = x_n$ ya fuera del equilibrio, con V representando la tensión de polarización (directa o inversa, es decir $V_{bi}-V_F$ ó $V_{bi}+V_R$), fuera del equilibrio.

Suponiendo ahora que se cumple la condición de baja inyección de portadores minoritarios y, en consecuencia, que la concentración de portadores minoritarios inyectados es mucho más pequeña que la de portadores mayoritarios, se cumplirá que:

$$p_p \approx p_{po}$$

Utilizando esta nueva condición, junto con la expresión obtenida anteriormente para p_{po} se obtiene la concentración de huecos en los límites de la región de distribución de carga espacial con la zona N ($x = x_n$):

$$p_n = p_{no} \cdot e^{qV/KT}$$

$$p_n - p_{no} = p_{no} \cdot (e^{qV/KT} - 1) \text{ para } x = x_n$$

Fig. 16a y 16b muestran los cambios producidos en la disposición de las bandas de energía y en las concentraciones de portadores con respecto a la situación de equilibrio y para los casos de polarización directa e inversa respectivamente.

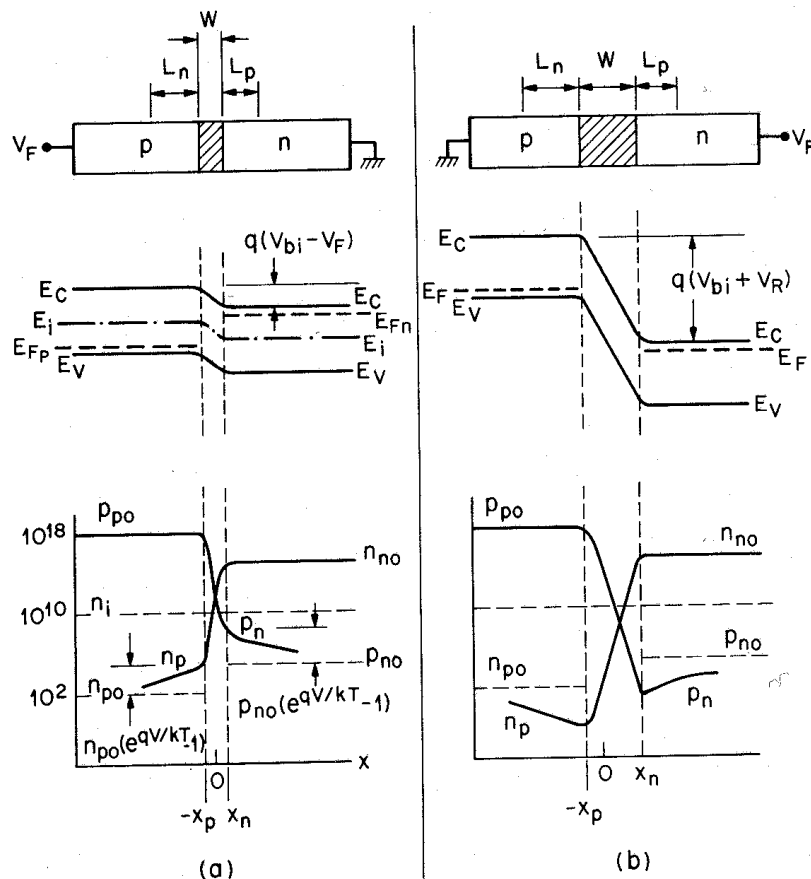


Fig. 16: Región espacial de carga, diagrama de bandas de energía y distribución de portadores. (a) En polarización directa. (b) En polarización inversa.

El hecho a resaltar es que, bajo polarización directa, la concentración de portadores minoritarios en los límites de la región espacial de carga ($-x_p$ o x_n) ha aumentado con respecto a la del equilibrio, y bajo polarización inversa ha disminuido con respecto de dicho nivel.

TEMA 3 : Física de la unión P-N.

Las ecuaciones anteriores, que expresan las concentraciones de portadores minoritarios en los límites de la región espacial de carga, son además importantes condiciones de contorno que serán empleadas para obtener la característica ideal corriente - tensión. Son conocidas bajo el nombre de **leyes de la unión**, aunque no deben olvidarse las restricciones bajo las cuales han sido obtenidas.

Distribución de portadores en las zonas neutras.

A continuación se va a calcular la distribución de portadores en las zonas neutras, en concreto se va a calcular la distribución de portadores minoritarios pues, como se ha visto, en condiciones de baja inyección de portadores las concentraciones de portadores mayoritarios en las zonas neutras apenas varían frente al caso de ausencia de polarización (equilibrio térmico). Debido a las hipótesis anteriores, no existe generación de corriente dentro de la región espacial de carga, todas las corrientes generadas proceden de las regiones neutras.

Para la región neutra de tipo N, no hay campo eléctrico en ella ($\varepsilon=0$) y la ecuación de continuidad para los huecos (portadores minoritarios) en el caso unidimensional y en el estado estacionario:

$$\frac{\partial p_n}{\partial t} = (-) p_n \cdot \mu_p \cdot \frac{\partial \varepsilon}{\partial x} - \mu_p \cdot \varepsilon \cdot \frac{\partial p_n}{\partial x} + D_p \cdot \frac{\partial^2 p_n}{\partial x^2} + G_p - \frac{p_n - p_{no}}{\tau_p} = 0$$

La cual se reduce a:

$$\frac{d^2 p_n}{dx^2} - \frac{p_n - p_{no}}{D_p \cdot \tau_p} = 0$$

Se trata de una ecuación diferencial de segundo orden y necesita para su resolución de dos condiciones de contorno. La primera de ellas ha sido obtenida anteriormente en las leyes de la unión, y la segunda se obtiene de considerar que el límite de la región N está lo suficientemente alejado de la zona espacial de carga como para que la concentración de portadores minoritarios en este punto sea igual a la del equilibrio, es decir, las condiciones de contorno son:

$$p_n = p_{no} \cdot e^{qV/KT} \text{ para } x = x_n \text{ (ley de la unión)}$$

$$p_n = p_{no} \text{ para } x \rightarrow \infty \text{ (límite alejado)}$$

Con estas dos condiciones, la ecuación diferencial anterior ofrece la solución:

$$p_n - p_{no} = p_{no} \cdot (e^{qV/KT} - 1) \cdot e^{-(x-x_n)/L_p} \text{ para } x > x_n$$

donde $L_p \equiv \sqrt{D_p \cdot \tau_p}$ es la longitud de difusión de los huecos (portadores minoritarios en la zona N). Como puede observarse, esta expresión es muy similar a la que se obtuvo en el Tema 2 en el ejemplo de resolución de la ecuación de continuidad correspondiente a la inyección de portadores minoritarios desde una cara en un semiconductor de tipo N, salvo que en este caso los portadores minoritarios (huecos) inyectados provienen del semiconductor de tipo P en el cual son mayoritarios.

Por otro lado, la densidad de corriente de huecos será:

TEMA 3 : Física de la unión P-N.

$$J_p(x) = (-)q \cdot D_p \cdot \frac{dp_n}{dx} = \frac{q \cdot D_p \cdot p_{no}}{L_p} \cdot (e^{qV/KT} - 1) \cdot e^{-(x-x_n)/L_p} \text{ para } x > x_n$$

en $x = x_n$ será:

$$J_p(x_n) = (-)q \cdot D_p \cdot \left. \frac{dp_n}{dx} \right|_{x=x_n} = \frac{q \cdot D_p \cdot p_{no}}{L_p} \cdot (e^{qV/KT} - 1) \text{ para } x = x_n$$

Únicamente se ha considerado de las dos aportaciones posibles (arrastre y difusión) la correspondiente a la corriente de difusión pues en las zonas neutras no existe campo eléctrico siendo nula la corriente de arrastre.

De forma análoga, la ecuación de continuidad para los electrones en la zona P en el caso unidimensional y en el estado estacionario:

$$\frac{\partial n_p}{\partial t} = n_p \cdot \mu_n \cdot \frac{\partial \varepsilon}{\partial x} + \mu_n \cdot \varepsilon \cdot \frac{\partial n_p}{\partial x} + D_n \cdot \frac{\partial^2 n_p}{\partial x^2} + G_n - \frac{n_p - n_{po}}{\tau_n} = 0$$

La cual se reduce a:

$$\frac{\partial^2 n_p}{\partial x^2} - \frac{n_p - n_{po}}{D_n \cdot \tau_n} = 0$$

Se trata de una ecuación diferencial de segundo orden, la cual se puede resolver utilizando las condiciones de contorno:

$$n_p = n_{po} \cdot e^{qV/KT} \text{ para } x = -x_p \text{ (ley de la unión)}$$

$$n_p = n_{po} \text{ para } x \rightarrow -\infty \text{ (límite alejado)}$$

Obteniendo:

$$n_p - n_{po} = n_{po} \cdot (e^{qV/KT} - 1) \cdot e^{(x+x_p)/L_n} \text{ para } x < -x_p$$

donde $L_n \equiv \sqrt{D_n \cdot \tau_n}$ es la longitud de difusión de los electrones (portadores minoritarios en la zona P). Por otro lado, la densidad de corriente de electrones:

$$J_n(x) = q \cdot D_n \cdot \frac{dn_p}{dx} = \frac{q \cdot D_n \cdot n_{po}}{L_n} \cdot (e^{qV/KT} - 1) \cdot e^{(x+x_p)/L_n} \text{ para } x < -x_p$$

En $x = -x_p$ será:

$$J_n(-x_p) = q \cdot D_n \cdot \left. \frac{dn_p}{dx} \right|_{x=-x_p} = \frac{q \cdot D_n \cdot n_{po}}{L_n} \cdot (e^{qV/KT} - 1) \text{ para } x = -x_p$$

Fig.17 muestra por un lado la distribución de portadores minoritarios en las zonas neutras y por otro la densidad de corriente en estas regiones bajo condiciones de polarización directa e inversa (casos a y b respectivamente).

TEMA 3 : Física de la unión P-N.

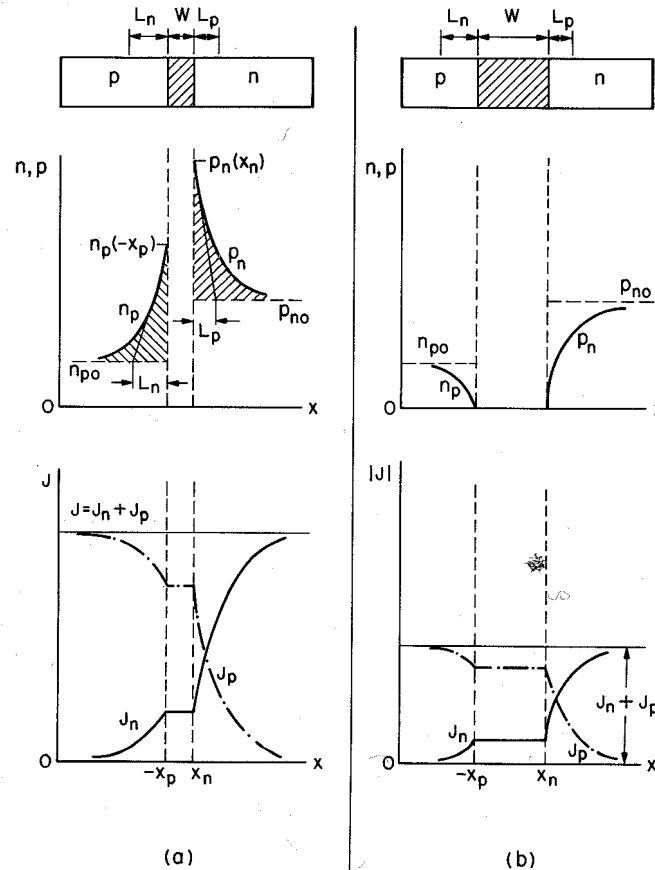


Fig. 17: Concentración de portadores minoritarios y corrientes de electrones y de huecos. (a) En polarización directa. (b) En polarización inversa.

Se observa que, por un lado, la corriente de difusión de huecos en la región N decae de forma exponencial con una longitud de difusión L_p y por otro como la corriente de difusión de electrones decae exponencialmente en la región P con longitud de difusión L_n .

Debido a la hipótesis 8, el número de portadores que entran en un borde de la zona de transición por unidad de tiempo es el mismo que la abandonan en el borde opuesto. Es decir:

$$J_n(x_n) = J_n(-x_p)$$

$$J_p(x_n) = J_p(-x_p)$$

Luego la densidad de corriente neta –suma de las densidades de corrientes de electrones y huecos- será uniforme y constante en régimen estacionario:

$$J = J_p(x_n) + J_n(-x_p) = J_s (e^{qV/KT} - 1)$$

siendo:

$$J_s \equiv \frac{q \cdot D_p \cdot p_{no}}{L_p} + \frac{q \cdot D_n \cdot n_{po}}{L_n} *$$

el término J_s recibe el nombre de la densidad de corriente de saturación. La ecuación anterior recibe el nombre de Ecuación del Diodo Ideal o de Shockley.

* Se ha supuesto en el tratamiento seguido hasta el momento el caso denominado de *diodo largo*, en que la anchura de la zona neutra de tipo P - W_p - es mucho mayor que L_n (longitud de difusión de los electrones en la zona P) y que la anchura de la zona neutra de tipo N - W_n - es mucho mayor que L_p (longitud de difusión de los huecos en la zona N). En dicho caso, la recombinación hace que se anule el exceso de portadores minoritarios inyectados en las regiones neutras, antes de llegar a los contactos metálicos. Un caso diferente se presenta cuando $W_n \ll L_p$ y $W_p \ll L_n$. Este es el caso del denominado *diodo corto*. En dicho caso la recombinación en las regiones neutras se considera nula y se puede demostrar que:

$$J_s \equiv \frac{q \cdot D_p \cdot p_{no}}{W_n} + \frac{q \cdot D_n \cdot n_{po}}{W_p}$$

La característica ideal corriente-tensión se muestra en Fig. 18a y b, en escala lineal y semilogarítmica, respectivamente.

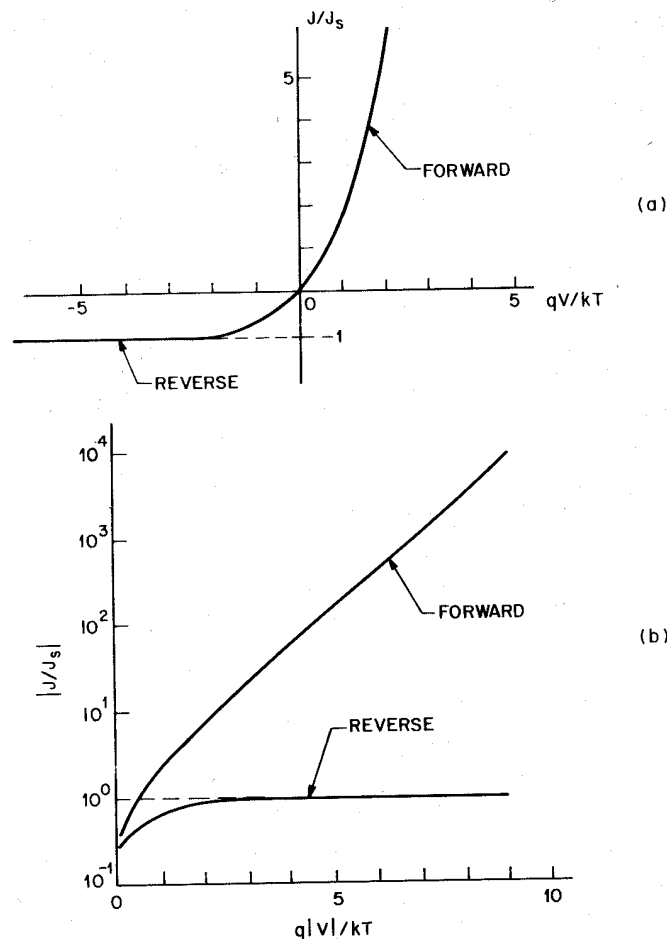


Fig. 18: Característica corriente-tensión ideal. (a) En escala lineal. (b) En escala semilogarítmica.

En el caso de polarización directa, para $V \geq (3 \cdot K \cdot T/q)$, la velocidad de incremento de la corriente es constante como se muestra en Fig. 17b. A 300K por cada década de cambio de la corriente, el cambio en la tensión es de 60mV ($= 2.3 \cdot K \cdot T/q$). En el caso de polarización inversa, la densidad de corriente satura en un valor $-J_s$.

TEMA 3 : Física de la unión P-N.

Teniendo en cuenta que $J = I/A$, siendo A el área de la sección recta del dispositivo, también puede obtenerse otra expresión en términos de corriente y de tensión para la unión p-n.

$$I = I_s \cdot (e^{qV/KT} - 1)$$

siendo:

$$I_s \equiv A \cdot \left[\frac{q \cdot D_p \cdot p_{no}}{L_p} + \frac{q \cdot D_n \cdot n_{po}}{L_n} \right]$$

La ecuación anterior de I puede ser considerada como la ley ideal del diodo. La relación I-V ideal es la característica representada en Fig. 19.

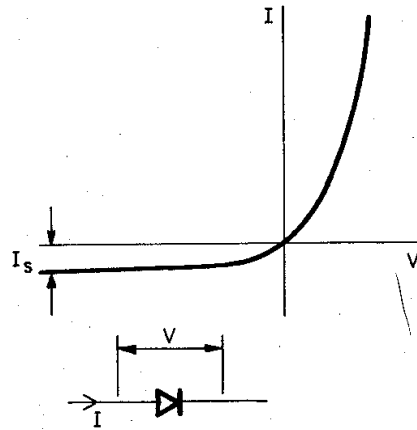


Fig. 19: Característica corriente-tensión de un diodo ideal y representación simbólica del diodo de unión.

Se observa que para el caso de polarización directa ($V > 0$), pequeñas variaciones en la tensión aplicada V provocan grandes variaciones en la corriente I , ya que en este caso, se puede aproximar la corriente por el término exponencial:

$$I \approx I_s \cdot e^{qV/KT}$$

pues $\frac{K \cdot T}{q}$ es muy pequeño ($\frac{K \cdot T}{q} \approx 26mV$ a 300K). Si, por el contrario, $V < 0$ y si

$|V| \gg \frac{K \cdot T}{q}$, entonces $I \approx -I_s$, valor muy pequeño como es el que se requiere en muchas

aplicaciones de los diodos. En particular I_s puede hacerse tan pequeño como se quiera consiguiendo que la concentración intrínseca n_i sea pequeña, es decir, con un material semiconductor que posea una ancha banda de energía prohibida y evitando altas temperaturas.

A efectos de aplicación, el diodo de unión presenta un marcado carácter de dispositivo rectificador pues sólo deja pasar la corriente en un único sentido. La corriente directa según el tipo de unión toma valores comprendidos entre 10^{-4} y $10^2 A$. En polarización inversa, la I_s puede llegar a valer, a temperatura ambiente entre algunos μA para uniones de Ge, y la milésima parte de esos μA para uniones de Si.

La expresión para la corriente en el diodo ideal:

$$I = I_s \cdot (e^{qV/KT} - 1)$$

configura el modelo analítico del diodo de unión. Hay que recordar que ha sido obtenida efectuando una serie de hipótesis previas. El modelo describe con precisión el comportamiento interno del dispositivo para corrientes directas y polarizaciones inversas moderadas. La anterior ecuación describe correctamente la característica corriente-tensión en las uniones p-n de Ge en régimen de baja densidad de corriente.

Para uniones p-n de Si o GaAs, la ecuación sólo describe el comportamiento de forma cualitativa y hay que considerar adicionalmente los efectos de generación y recombinación de portadores en la región espacial de carga. Para estos casos, considerando bajo nivel de corriente predominan los efectos de recombinación y, en régimen de alto nivel de corriente, es el fenómeno de la difusión el que predomina.

3.5. Fenómenos de ruptura.

El fenómeno de ruptura aparece en las uniones p-n cuando se las somete a una tensión inversa de elevado valor. En principio, para pequeños valores de la tensión inversa aplicada, la corriente que circula por la unión es aproximadamente constante e igual a $-I_s$ (Fig. 20). Sin embargo, cuando la tensión inversa aplicada toma valores relativamente elevados se llega a un momento en el cual la corriente crece bruscamente alcanzando valores muy grandes, es el fenómeno de la ruptura. Al valor de la tensión inversa aplicada para el cual ocurre este aumento brusco de la corriente recibe el nombre de tensión de ruptura V_r .

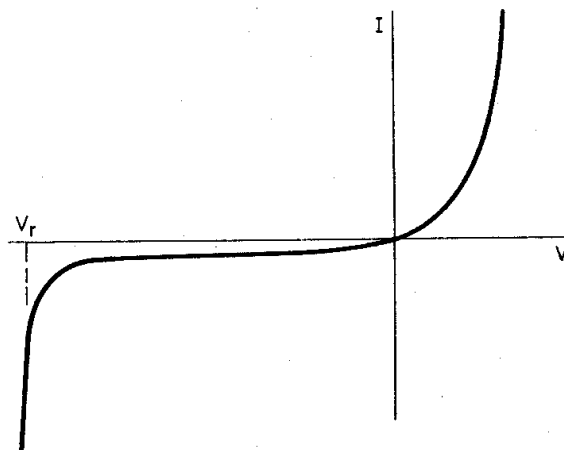


Fig. 20: Característica I-V de un diodo de unión real, con la tensión de ruptura V_r .

El proceso de ruptura por si solo es destructivo y el aumento de corriente debe ser limitado por un circuito externo a la unión p-n para evitar el aumento excesivo de calor en la unión. Fundamentalmente se dan dos tipos de fenómenos de ruptura: a) ruptura por efecto túnel o zéner y b) ruptura por avalancha. Consideraremos brevemente el primero de ellos y se entrará en más detalles para el caso de multiplicación por avalancha, ya que la ruptura por avalancha impone un límite máximo para la tensión inversa en la mayor parte de diodos. La ruptura por avalancha también limita la tensión de colector en un transistor bipolar y la de drenador de un MOSFET. Además los mecanismos de multiplicación por avalancha pueden generar potencia a nivel de microondas (diodo IMPATT) y pueden ser utilizados para detectar señales ópticas (fotodetector por avalancha).

- a) Ruptura de multiplicación por avalancha: Fenómeno de ruptura provocado por la creación en cadena de pares e^-h^+ en condiciones de campo eléctrico intenso.

Desde un punto de vista cualitativo, el proceso de multiplicación por avalancha ocurre cuando en presencia de un campo eléctrico suficientemente elevado se genera un electrón de procedencia térmica (denotado por 1). Dicho electrón adquiere energía cinética a expensas de la del campo eléctrico. Si este campo es suficientemente elevado, la energía cinética puede ser la suficiente como para producir, mediante colisión con un átomo de la red cristalina, un par electrón-hueco (denotado por 2 y 2'). Este hecho se denomina ionización por impacto. El par electrón-hueco también puede aumentar su energía cinética a expensas del campo aplicado y generar otro par electrón-hueco (3 y 3'). Este proceso en presencia del campo eléctrico puede ser multiplicativo aumentando la conductividad de las regiones P y N y, en consecuencia, la corriente que pasa a su través. Este es el llamado fenómeno de ruptura a través de la multiplicación por avalancha (Fig. 21b).

Existe un factor M llamado factor multiplicativo que gobierna la corriente inversa I_i una vez iniciado el proceso de avalancha. Este factor viene dado por:

$$M = \frac{1}{1 - \left(\frac{|V|}{V_r}\right)^\mu}$$

donde V_r = tensión de ruptura
 V = tensión inversa aplicada anterior a la ruptura
 μ = factor experimental que depende del dopado de las regiones P y N.

Con este factor multiplicativo, la corriente inversa en las regiones de la característica corriente-tensión cercanas a la ruptura viene dada por:

$$I_R = \frac{I_s}{1 - \left(\frac{|V|}{V_r}\right)^\mu}$$

Como se ve, para $V = V_r$, la corriente se hace infinita debido al fenómeno de avalancha, aunque en la práctica se limita con una resistencia de un circuito externo.

Condiciones usuales son las siguientes: unión p-n con dopado moderado (por ejemplo uniones p⁺n de tipo abrupto con concentración $N_D \approx 10^{17} \text{ cm}^{-3}$ o inferior), y bajo polarización inversa.

Por otra parte, la tensión de ruptura V_r por avalancha tiene un coeficiente de temperatura positivo, es decir, aumenta con la temperatura.

- b) Ruptura por efecto túnel. El efecto túnel o zéner se produce cuando un campo eléctrico intenso se aplica a la unión p-n. Tiene su explicación en las propiedades mecánico-cuánticas de los electrones. En concreto bajo estas condiciones de campo eléctrico un electrón situado en la banda de valencia puede efectuar una transición a la banda de conducción penetrando o saltándose la barrera de potencial (Fig. 21a).

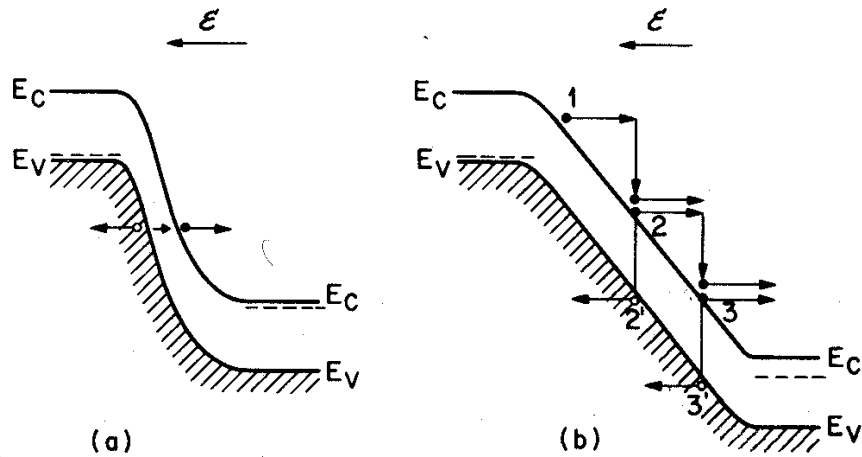


Fig. 21: Diagrama de bandas de energía en condiciones de ruptura. (a) Efecto túnel.
(b) Multiplicación por avalancha.

Desde un punto de vista cuantitativo y atendiendo a la mecánica cuántica existe cierta probabilidad de que un electrón con energía E atraviese una barrera de potencial de altura E_0 y anchura W con $E < E_0$, esta probabilidad viene dada por la expresión:

$$\text{Probabilidad} = \left[1 + \frac{E_0^2 \cdot \sinh^2(\beta \cdot W)}{4 \cdot E \cdot (E_0 - E)} \right]^{-1}; \text{ con } \beta \equiv \left[\frac{2 \cdot m \cdot (E_0 - E)}{h^2} \right]^{1/2}$$

El efecto túnel se da cuando hay presente en la unión p-n altos valores de campo eléctrico, para el caso del Si y el GaAs, el valor del campo para manifestarse el efecto túnel es de 10^6V/cm o mayor. Esto sólo es posible cuando las regiones P y N están fuertemente dopadas ($> 5 \cdot 10^{17} \text{cm}^{-3}$), de forma que la anchura W de la región espacial de carga es muy pequeña.

Hablando de términos macroscópicos, el fenómeno de ruptura por efecto túnel se da para valores de la tensión de ruptura del orden de $4 \cdot E_g/q$ o menores donde E_g es la anchura de la banda prohibida de energía. Para valores de la tensión de ruptura V_R superiores a $6 \cdot E_g/q$, el fenómeno de ruptura es una consecuencia directa de la multiplicación por avalancha. Si la tensión de ruptura V_R está comprendida entre los valores $4 \cdot E_g/q$ y $6 \cdot E_g/q$ el fenómeno de ruptura es debido a la presencia de ambos efectos, el de multiplicación por avalancha y el efecto túnel.

Finalmente, para el Si y el GaAs, debido a que el valor de E_g decrece con la temperatura T , la tensión de ruptura debida al efecto túnel decrece cuando la temperatura aumenta (coeficiente de temperatura negativo).

En los casos en que ambos mecanismos influyen por igual (ruptura por avalancha y zéner), la tensión de ruptura V_r es casi independiente de la temperatura y muy estable. Esta última propiedad es aplicada por determinados diodos (llamados diodos zéner) para conseguir referencias de tensión muy estables.

3.5.1. Cálculo de la tensión de ruptura por avalancha.

Al aumentar la tensión de polarización inversa, esta provoca un incremento de la corriente provocando la ruptura. La tensión inversa aplicada genera un elevado campo eléctrico que provoca el efecto de avalancha acelerando a los portadores que, de hecho, producen nuevos portadores.

Existe un determinado valor crítico del campo eléctrico, ϵ_{cr} , el cual causa la ruptura. El valor de dicho campo eléctrico se supone constante para un determinado semiconductor, aunque en realidad varía ligeramente con el dopado. Suponiendo constante el valor del campo eléctrico crítico vamos a obtener un valor para la tensión de ruptura por avalancha.

Para una unión abrupta p-n el se cumple la siguiente relación:

$$V_j = V_{bi} - V = \frac{1}{2} \cdot \epsilon_m \cdot W$$

con

$$W = \sqrt{\frac{2 \cdot \epsilon_s}{q} \cdot \left(\frac{N_A + N_D}{N_A \cdot N_D} \right) \cdot (V_{bi} - V)}$$

En consecuencia, el valor máximo del campo eléctrico en la zona de la unión (el cual ocurre en $x = 0$), ϵ_m , será:

$$\epsilon_m = \sqrt{\frac{2 \cdot q}{\epsilon_s} \cdot \left(\frac{N_A \cdot N_D}{N_A + N_D} \right) \cdot (V_{bi} - V)}$$

Cuando el valor de dicho campo eléctrico máximo alcance el valor crítico ocurrirá la ruptura de la unión. Esto ocurrirá para $V = -V_r$.

$$\epsilon_{cr} = \sqrt{\frac{2 \cdot q}{\epsilon_s} \cdot \left(\frac{N_A \cdot N_D}{N_A + N_D} \right) \cdot (V_{bi} + V_r)}$$

Esta expresión nos permite, conocido el valor de ϵ_{cr} , obtener la tensión de ruptura por avalancha para el caso de una unión abrupta.

Para una unión abrupta a un lado p⁺-n ($N_A \gg N_D$):

$$\epsilon_m = \sqrt{\frac{2 \cdot q \cdot N_D}{\epsilon_s} \cdot (V_{bi} - V)}$$

El campo eléctrico máximo en la zona de la unión queda determinado en función de la concentración ligera de impurezas N_D . La ruptura ocurre cuando, para $V = -V_r$, el campo eléctrico máximo iguala su valor crítico $\epsilon_m = \epsilon_{cr}$. Suponiendo $V_{bi} \ll V_r$, se obtiene:

$$\epsilon_{cr} = \sqrt{\frac{2 \cdot q \cdot N_D}{\epsilon_s} \cdot V_r}$$

Y, por tanto:

$$V_r = \frac{\epsilon_s \cdot \mathcal{E}_{cr}^2}{2 \cdot q \cdot N_D}$$

Suponiendo constante el campo eléctrico crítico, el valor de la tensión de ruptura disminuye al aumentar la concentración ligera de impurezas.

3.5.2. Efecto de la curvatura de la unión.

Un efecto interesante relacionado con la tensión de ruptura en la unión p-n es el llamado efecto de la curvatura de la unión. Cuando se forma una unión p-n por difusión a través de una ventana en una capa de SiO₂, las impurezas se difunden no sólo de forma vertical y hacia abajo sino también hacia los lados (Fig. 22).

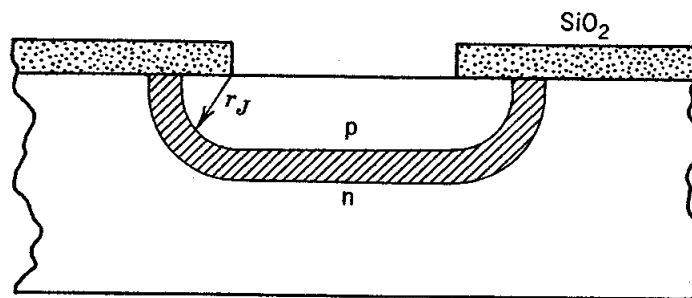


Fig. 22: Formación de la curvatura de la unión mediante el proceso de difusión planar.

Por ello, la unión p-n presenta una zona plana con laterales de tipo cilíndrico y las esquinas conformadas de tipo esférico (Fig. 23).

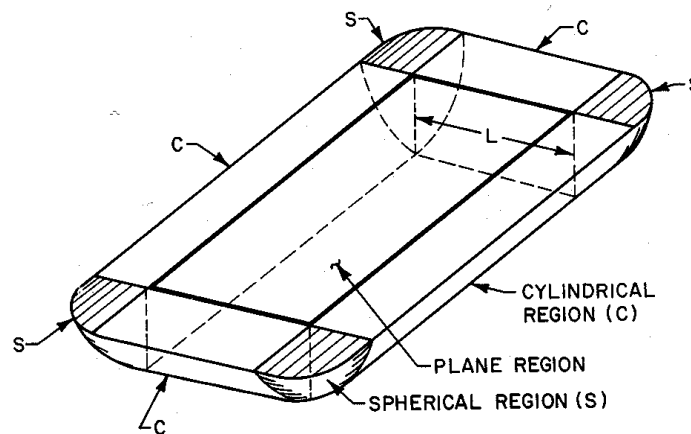


Fig. 23: Regiones cilíndricas y esféricas formadas.

Estas dos últimas regiones acusan una mayor intensidad de campo eléctrico que la región plana de la unión, siendo, para un mismo radio de curvatura r_j , mayor en una región de tipo esférico que cilíndrico. Si recordamos, la tensión de ruptura aparecía cuando la tensión inversa aplicada era suficiente para que el campo eléctrico máximo en la zona de la unión igualase su valor crítico. Como en las zonas esféricas, el campo eléctrico creado por la tensión de polarización inversa será mayor que en las demás zonas, se llegará antes (con la misma tensión de polarización) a igualar el valor crítico y son, por tanto, las que fijarán el fenómeno de ruptura.

SOLUCIONES:

1) Disminución de la concentración ligera de impurezas.

La concentración de impurezas en cada zona de la unión influye en la tensión de ruptura por avalancha. A partir de Fig. 24 observamos como, para el caso de una unión p^+n , una solución para aumentar la tensión de ruptura es disminuir la concentración ligera de impurezas ($N_B=N_D$). En esta observamos la variación que experimenta la tensión de ruptura por avalancha en función de la concentración ligera de impurezas para distintos tipos de semiconductores:

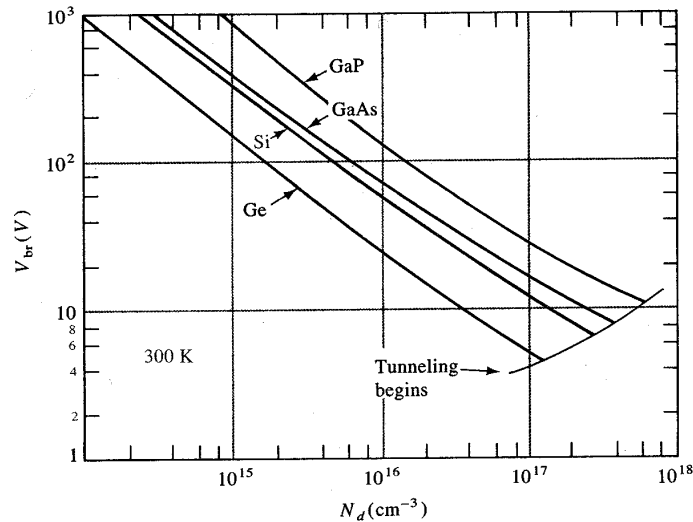


Fig. 24: Variación de la tensión de ruptura por avalancha en una unión abrupta p^+n , en función de la concentración ligera de impurezas para varios semiconductores.

Este resultado está de acuerdo con la expresión obtenida anteriormente para V_r en el caso de una unión p^+n :

$$V_r = \frac{\epsilon_s \cdot \mathcal{E}_{cr}^2}{2 \cdot q \cdot N_D}$$

Considerando la geometría de la unión, para el caso de una unión p^+n de Si (Fig. 25):

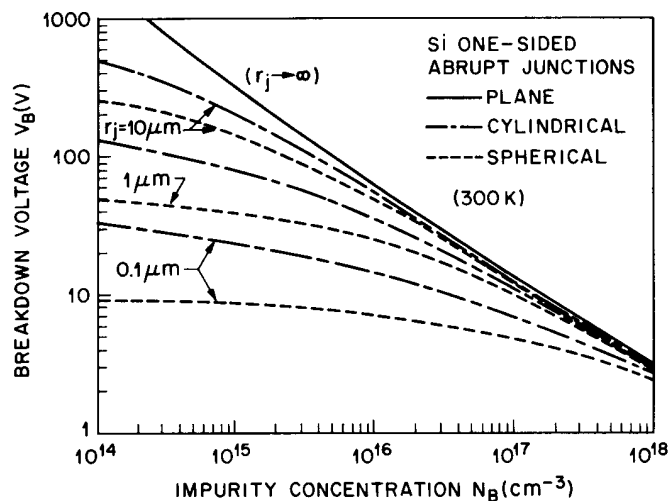


Fig. 25: Variación de la tensión de ruptura por avalancha en una unión abrupta p^+n de Si, en función de la concentración ligera de impurezas considerando la geometría de la unión, donde r_j es el radio de curvatura.

TEMA 3 : Física de la unión P-N.

La línea continua es para el caso considerado previamente de unión plana ($r_j \rightarrow \infty$). Cuando el radio de curvatura disminuye, la tensión de ruptura disminuye, especialmente para el caso de las zonas esféricas para alto nivel de dopado.

Sin embargo, una disminución de $N_B = N_D$ haría aumentar la resistividad de la región ligeramente dopada:

$$\rho_n = \frac{1}{q \cdot N_D \cdot \mu_n}$$

lo cual tendería a incrementar la resistencia directa del diodo y, por tanto, podría producir efectos térmicos de calentamiento I^2R . Para reducir la resistencia de la región ligeramente dopada, es necesario hacer su área grande y reducir su longitud ($R = \frac{\rho \cdot L}{A}$).

Por tanto, la geometría física del diodo es otra variable importante de diseño.

Pero la región ligeramente dopada no puede ser arbitrariamente corta. Uno de los principales problemas que se presentan cuando se tiene una unión ligeramente dopada corta es el llamado efecto “*punch-through*”. Dado que la anchura de la región espacial de carga (W) se incrementa con la tensión inversa aplicada y se extiende dentro de la región ligeramente dopada (recordar que en una unión $p^+ - n$, $W \cong x_n$) es posible que W se incremente de tal forma que complete toda la longitud de esta región. Si la región ligeramente dopada (tipo N) es corta, esto es fácil que se produzca en condiciones de polarización inversa.

Este fenómeno es lo que se llama “*punch-through*” y tiene como consecuencia la ruptura de la unión por debajo del valor esperado de V_r .

2) Técnicas de biselado y de anillos de campo.

En dispositivos diseñados para ser utilizados con altas tensiones inversas se debe tener mucho cuidado para evitar una ruptura prematura. A fin de aumentar dicha tensión de ruptura se tienen dos posibles soluciones: biselado y anillos de campo.

La técnica de biselado consiste en el corte perpendicular (con una sierra con diente de diamante) de aquellas regiones en que la unión presenta una geometría cilíndrica o esférica. El campo eléctrico máximo disminuye de esa forma y por tanto aumenta la tensión de ruptura.

Fig. 26a muestra el diodo biselado y 26b muestra la zona lateral vista desde arriba.

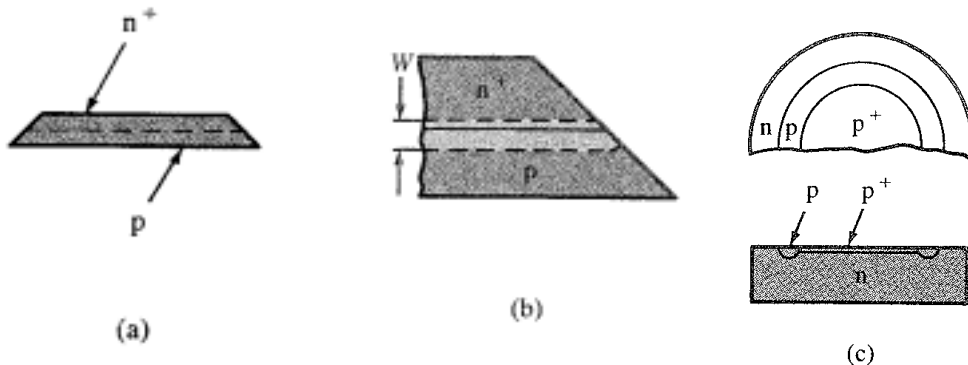


Fig. 26: Técnicas de biselado y anillos de campo para aumentar la tensión de ruptura inversa. (a) Diodo biselado. (b) Zona lateral vista desde arriba. (c) Técnica de anillos de campo.

En la solución conocida como de anillos de campo, el proceso de difusión o implantación de impurezas se hace de forma selectiva utilizando una unión p^+-n en la zona central y una unión $p-n$ en las laterales (Fig. 26c). Dado que la anchura de la región espacial de carga es mayor en el anillo p que en la región p^+ , el campo eléctrico es menor en el anillo para una tensión inversa aplicada dada. Por tanto, la ruptura de la unión $p-n$ se produce a una tensión mayor que la de la unión p^+-n y, por tanto, será la zona central y no la lateral la que determinará la tensión de ruptura.

3) Implantación iónica:

La utilización de la técnica conocida como implantación iónica (bombardeo de iones) en lugar de difusión evita la formación de las zonas de tipo cilíndrico y esférico en la unión $p-n$ evitando los efectos no deseados sobre la tensión de ruptura debidos a estas zonas.

3.6. Almacenamiento de carga y capacidad de difusión.

Los fenómenos anteriormente descritos ocurren en casos de polarización inversa. Sin embargo, hay dos efectos que están relacionados con la unión $p-n$ cuando ésta está sometida a polarización directa. Bajo estas condiciones ocurre una inyección de portadores minoritarios en las dos regiones de la unión, concretamente de huecos en la zona N y de electrones en la zona P .

La distribución de portadores minoritarios en cada una de las regiones decae exponencialmente a medida que nos alejamos de la unión (Fig. 27).

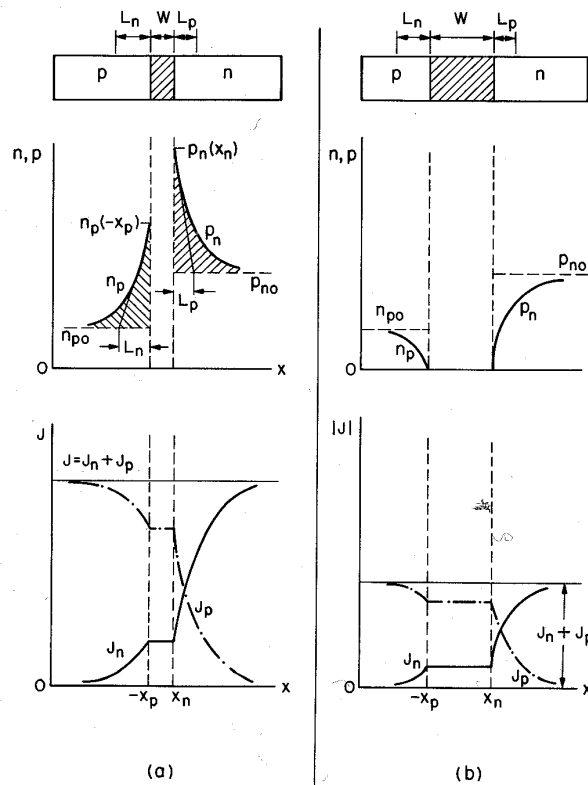


Fig. 27: Concentración de portadores minoritarios y corrientes de electrones y de huecos. (a) En polarización directa. (b) En polarización inversa.

TEMA 3 : Física de la unión P-N.

Esta distribución de portadores minoritarios inyectados tiene asociada una determinada carga almacenada. Esta carga se puede obtener para el caso de la región N mediante integración del exceso de portadores minoritarios (huecos) en la zona neutra (zona sombreada de Fig. 27). Es decir, empleando la relación anteriormente obtenida para la zona neutra de tipo N:

$$p_n - p_{no} = p_{no} \cdot (e^{qV/KT} - 1) \cdot e^{-(x-x_n)/L_p} \text{ para } x > x_n$$

se tiene que:

$$\begin{aligned} Q_n &= q \cdot \int_{x_n}^{\infty} (p_n - p_{no}) \cdot dx = q \cdot \int_{x_n}^{\infty} p_{no} \cdot (e^{qV/KT} - 1) \cdot e^{-(x-x_n)/L_p} \cdot dx = \\ &= q \cdot L_p \cdot p_{no} \cdot (e^{qV/KT} - 1) = q \cdot L_p \cdot p_n(x_n) \end{aligned}$$

donde se ha utilizado la Ley de la Unión que nos dice que:

$$p_n(x_n) = p_{no} \cdot (e^{qV/KT} - 1)$$

Una expresión similar puede ser obtenida para la carga almacenada en la región P debida a la inyección de portadores minoritarios (electrones).

$$Q_p = q \cdot L_n \cdot n_{po} \cdot (e^{qV/KT} - 1) = q \cdot L_n \cdot n_p(-x_p)$$

Se observa que la carga almacenada depende por un lado de la longitud de difusión y por otro de la densidad de carga presente en los límites de la región espacial de carga.

También puede expresarse la carga almacenada en términos de la corriente inyectada. Empleando la ecuación de la densidad de corriente debida a los portadores minoritarios anteriormente obtenida para los huecos en la zona N, es decir:

$$J_p(x_n) = (-)q \cdot D_p \cdot \left. \frac{dp_n}{dx} \right|_{x=x_n} = \frac{q \cdot D_p \cdot p_{no}}{L_p} \cdot (e^{qV/KT} - 1)$$

y la anterior expresión para Q_n puede obtenerse:

$$Q_n = \frac{L_p^2}{D_p} \cdot J_p(x_n) = \tau_p \cdot J_p(x_n)$$

donde se ha hecho uso de que $L_p = \sqrt{D_p \cdot \tau_p}$.

Esta expresión nos dice que la cantidad de carga almacenada debido a la inyección de portadores minoritarios depende de la corriente de estos portadores en el límite de la región espacial de carga y de su tiempo de vida media. Los procesos de almacenamiento de carga son importantes cuando el tiempo de vida media de los portadores minoritarios es alto, es decir, predomina la difusión frente a la recombinación. (transcurre mucho tiempo entre la inyección y la recombinación completa de todos los portadores minoritarios inyectados).

Es importante darse cuenta de que la carga almacenada a la cual nos estamos refiriendo es debida al exceso de portadores minoritarios. La presencia de un exceso de portadores mayoritarios provenientes de los contactos metálicos garantiza la neutralidad de las zonas P y N. No hay, por tanto, una carga espacial neta en las regiones neutras, cosa que sí sucede en la región espacial de carga.

Recuérdese que aparecía una capacidad asociada a la región espacial de carga C_j en condiciones de polarización inversa. Cuando la unión p-n está sometida a polarización directa, el fenómeno de almacenamiento de carga eléctrica en las regiones neutras asociado a la inyección de portadores minoritarios también lleva asociado una

TEMA 3 : Física de la unión P-N.

capacidad, es la llamada *capacidad de difusión*, C_d . Este término se obtiene a partir de considerar una unión p-n ideal cuando los portadores minoritarios se almacenan en las regiones neutras por difusión.

La expresión que permite obtener la capacidad de difusión C_d es:

$$C_d \equiv A \cdot \frac{dQ}{dV} \text{ con } Q = Q_n + Q_p$$

donde Q_n es la carga almacenada debida a los huecos almacenados en la zona neutra de tipo N y Q_p es la carga almacenada debida a los electrones almacenados en la zona neutra de tipo P.

En el caso de una unión abrupta a un lado p⁺-n:

$$\left. \begin{array}{l} \text{Ley de acción de masas: } p_{no} \cdot n_{no} = n_{po} \cdot p_{po} = n_i^2 \\ \text{Semiconductor tipo N: } n_{no} \cong N_D \\ \text{Semiconductor tipo P: } p_{po} \cong N_A \end{array} \right\} \begin{array}{l} p_{po} (=N_A) \gg n_{no} (=N_D) \\ \downarrow \\ p_{no} \gg n_{po} \Rightarrow Q_n \gg Q_p \end{array}$$

por lo que la capacidad de difusión es aproximadamente igual a la debida a los huecos almacenados en la zona neutra de tipo N - Q_n , es decir:

$$C_d = \frac{A \cdot q^2}{K \cdot T} \cdot (L_p \cdot p_{no} + L_n \cdot n_{po}) \cdot e^{qV/K \cdot T} \cong \frac{A \cdot q^2 \cdot L_p \cdot p_{no}}{K \cdot T} \cdot e^{qV/K \cdot T}$$

En condiciones de polarización inversa C_d no tiene sentido debido a que la carga de portadores minoritarios almacenada es despreciable.

En muchas aplicaciones se representa la unión p-n por medio de un circuito equivalente, el cual además de recoger las capacidades de difusión y de transición C_d y C_j incluye un nuevo elemento llamado conductancia del diodo que nos proporciona la variación que experimenta la corriente a través del diodo al variar la tensión de polarización. En el diodo ideal esta conductancia se obtiene a partir de la ecuación del diodo ideal o de Shockley:

$$J = J_s \cdot (e^{\frac{qV}{K \cdot T}} - 1) \Rightarrow \left\{ \begin{array}{l} \frac{dJ}{dV} = \frac{q}{K \cdot T} \cdot J_s \cdot e^{\frac{qV}{K \cdot T}} \\ J + J_s = J_s \cdot e^{\frac{qV}{K \cdot T}} \text{ y } J + J_s \cong J = \frac{I}{A} \end{array} \right.$$

$$G \equiv \frac{dI}{dV} = \frac{A \cdot dJ}{dV} = \frac{q \cdot A}{K \cdot T} \cdot J_s \cdot e^{qV/K \cdot T} = \frac{q \cdot A}{K \cdot T} \cdot (J + J_s) \approx \frac{q \cdot I}{K \cdot T} = \frac{I}{V_T} \text{ (A/V)}$$

con:

$$V_T = \frac{K \cdot T}{q}$$

El circuito equivalente total que representa a la unión p-n queda reflejado en Fig. 28, en el cual C_j representa la capacidad de transición total (es decir, multiplicada por la sección transversal de diodo, A). Este circuito, que contiene las capacidades y la conductancia puede ser utilizado para reemplazar al diodo en *variaciones incrementales de poca amplitud* en la tensión o la corriente. Cuando, por ejemplo, se realiza un incremento de la tensión de polarización aplicada, analizando el circuito se pueden

determinar tanto el correspondiente cambio que experimenta la corriente como el tiempo de respuesta del diodo ante dicho cambio, a través de por ejemplo la medida de cuanto tiempo tarda la corriente a través del diodo en alcanzar una determinada fracción de su valor final. Analizando el circuito es posible obtener también la respuesta ante tensiones o corrientes sinusoidales de baja amplitud. Dicho circuito es conocido como circuito equivalente del diodo en régimen de pequeña señal y satisface con exactitud el comportamiento eléctrico del circuito teniendo en cuenta que para polarización directa $C_j \ll C_d$ y para polarización inversa $C_j \gg C_d$.

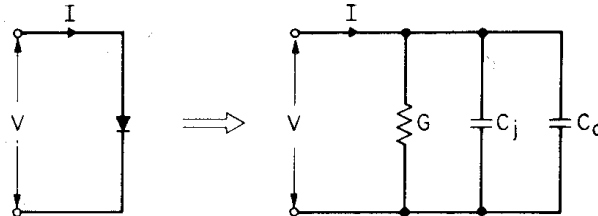


Fig. 28: Circuito equivalente en pequeña señal de la unión p-n.

En la bibliografía al término V_T/I también se le suele llamar resistencia incremental o dinámica del diodo:

$$r_d \equiv \frac{dV}{dI} = \left(\frac{dI}{dV} \right)^{-1} \approx \frac{V_T}{I} \quad (\Omega)$$

donde $V_T \cong 26\text{mV}$ (a 300K).

Dicha resistencia es la que aparece en Fig. 28.

3.7. Comportamiento transitorio: tiempo de paso a corte.

Los diodos son utilizados como conmutadores en gran cantidad de aplicaciones. Cuando se utiliza como conmutador, el diodo se dice que está a ON cuando está directamente polarizado y a OFF cuando está inversamente polarizado en la región de la curva característica que precede a la ruptura. Cuando la tensión de polarización del diodo cambia bruscamente de polaridad se dice que el diodo trabaja en conmutación.

En esta explicación vamos a suponer que trabajamos con un diodo de unión p⁺-n de forma que la concentración en exceso de huecos en la zona N es la que determina el almacenamiento de carga. Vamos a centrarnos en el paso a corte (OFF).

Para cortar el diodo es necesario eliminar las cargas almacenadas en exceso. El diodo estará cortado cuando la corriente a su través sea $-I_s$.

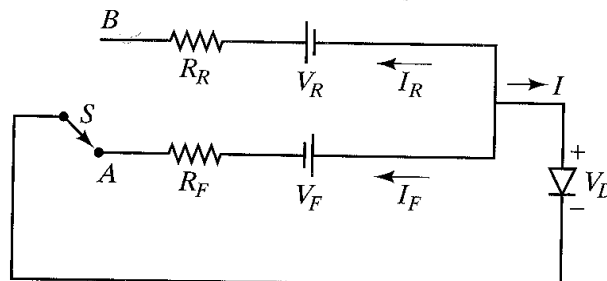


Fig. 29: Circuito de conmutación del diodo.

En $t=0$, el conmutador de Fig. 29 es llevado de forma casi instantánea desde la posición A a la posición B, de forma que se aplica al diodo una polarización inversa. En el instante antes de mover el conmutador, en $t=0^-$, el diodo está conduciendo una

TEMA 3 : Física de la unión P-N.

corriente I_F y la tensión entre extremos del diodo es muy pequeña, V_D , y, por tanto mucho menor que V_R . Como puede observarse en Fig. 30a, para $t=0^-$ al tratarse de una unión p^+-n la concentración en exceso de huecos en la zona N es mucho mayor que la concentración en exceso de electrones en la zona P y por lo tanto $Q_n \gg Q_p$ (recordar que estas magnitudes estaban relacionadas con las áreas sombreadas en Fig. 27).

En $t=0$, el conmutador es llevado de forma casi instantánea desde la posición A a la posición B aplicando al diodo una polarización inversa. La corriente cambia de forma instantánea desde su valor en polarización directa I_F hasta un valor constante de corriente inversa I_R que viene dado por:

$$I_R = \frac{V_R + V_D}{R_R} \cong \frac{V_R}{R_R} \text{ pues } V_D \ll V_R$$

Mientras haya un gran número de huecos cerca del límite de la región espacial de carga con la zona N, la corriente inversa será grande y los huecos se difundirán en el sentido desde N hasta P, con una concentración de huecos en x_n que tiene una pendiente positiva:

$$I_D = A \cdot (J_p(x_n) + J_n(-x_p)) \cong A \cdot J_p(x_n) = (-)q \cdot D_p \cdot A \cdot \left. \frac{dp_n}{dx} \right|_{x=x_n} = -I_R \text{ con } \left. \frac{dp_n}{dx} \right|_{x=x_n} > 0$$

En esta expresión se ha hecho uso de que, al tratarse de una unión p^+-n , la concentración en exceso de huecos es la que determina el almacenamiento de carga y, por tanto, también la densidad de corriente ($J_p(x_n) \gg J_n(-x_p)$).

Este proceso continúa hasta que la concentración de huecos en exceso en la unión se reduce a cero, de forma que la concentración de huecos se iguala a su valor en equilibrio térmico, haciéndose V_D igual a cero. El tiempo que tarda la concentración de huecos en exceso en llegar a ser cero ($p_n = p_{n0}$) es conocido como tiempo de almacenamiento. Durante todo este tiempo la corriente inversa se mantiene en un valor constante $-I_R$.

Una expresión aproximada del tiempo de almacenamiento, t_s , para un diodo largo p^+-n es:

$$t_s = \tau_p \cdot \ln \left[1 + \frac{I_F}{I_R} \right]$$

Conforme la concentración de huecos en $x=x_n$ va haciéndose mucho más pequeña que su valor en el equilibrio térmico, V_D se hace negativa y grande, y la corriente se aproxima a $-I_s$. El tiempo en que la corriente cae desde $-I_R$ hasta $-0.1 \cdot I_R$ es llamado tiempo de recuperación, t_r . El tiempo de recuperación t_r es el tiempo que se tarda en cargar la capacidad de transición C_j a la tensión inversa, V_R . Una estimación de t_r viene dada por la constante de tiempo $2.3 \cdot R_R \cdot C_j$.

El tiempo total de paso a OFF, $t_{tr} = t_r + t_s$, es el llamado tiempo de recuperación inversa del diodo.

TEMA 3 : Física de la unión P-N.

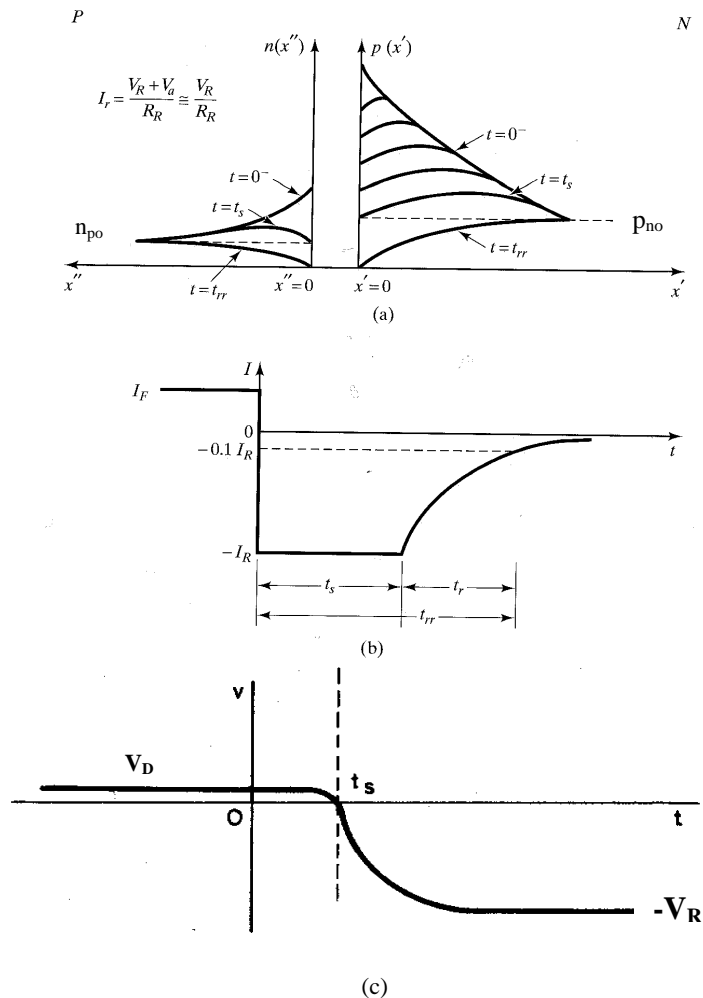


Fig. 30: (a) Caída de la concentración de portadores minoritarios. (b) Variación que experimenta la corriente debido a la caída de la concentración de los huecos. (c) Evolución de la tensión en el transitorio de corte del diodo.

3.8. Unión metal-semiconductor.

3.8.1. Diodos Schottky.

Los diodos llamados Schottky son dispositivos formados mediante la unión de un metal y un semiconductor. En esta unión se crea una barrera de potencial (barrera Schottky) debido a la diferente concentración de electrones libres entre ambos elementos. La conducción se efectúa, por tanto, mediante portadores mayoritarios (electrones), al contrario que en los diodos de unión, lo cual permite una inmediata respuesta, muy útil en aplicaciones de conmutación rápida (lógicas digitales de alta velocidad). Por otra parte, también poseen una menor altura de la barrera de potencial con respecto al diodo de unión (aproximadamente 0.3V para los diodos Schottky frente a los 0.6V de los de unión).

Fig. 31 muestra el proceso de formación de la barrera de acuerdo con la teoría de Schottky-Mott. Vamos a suponer que el semiconductor es de tipo N y está uniformemente dopado. Fig. 31a muestra los dos componentes (metal y semiconductor tipo N) aislados el uno del otro. Se ha tomado como nivel de referencia de las energías el nivel superior que es el llamado *nivel de vacío*, el cual representa la energía de un electrón en reposo fuera de los materiales. La función de trabajo ϕ_m (del metal) y ϕ_{sc} (del semiconductor), son la diferencia de energía entre el nivel de vacío y el nivel de

TEMA 3 : Física de la unión P-N.

Fermi del material y, por tanto, representan la energía necesaria para llevar un electrón que está en el nivel de Fermi del material al nivel de vacío fuera de este. La afinidad electrónica del semiconductor $q\chi$ denota la diferencia energética entre el nivel inferior de la banda de conducción E_C y el nivel de vacío. En un semiconductor determinado, $q\chi$ es independiente del nivel de dopado (magnitud constante), pero sin embargo $q\Phi_{sc}$ no lo es.

Para que la unión metal-semiconductor tipo N se comporte como una unión rectificadora se ha de cumplir la relación $q\Phi_m > q\Phi_{sc}$. En caso contrario, los electrones pueden ser transferidos de un lado a otro con facilidad y el contacto es óhmico.

Cuando el metal es puesto en contacto con el semiconductor, los electrones de la banda de conducción del semiconductor (como $\Phi_m > \Phi_{sc}$, los electrones de la banda de conducción del semiconductor tienen una energía superior a la de los del metal) fluyen hacia el metal hasta que los niveles de Fermi en los dos materiales se alinean. Este paso se debe por tanto a una diferente densidad de electrones libres en los dos materiales. El diagrama de bandas resultante puede observarse en Fig. 31b.

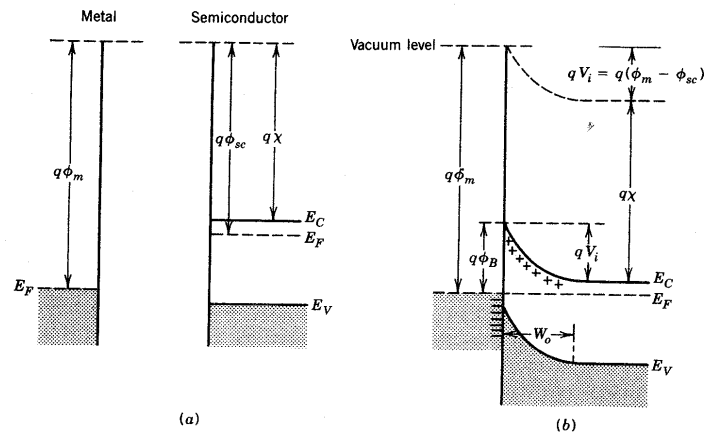


Fig. 31: Diagrama de bandas de energía de un contacto metal-semiconductor de tipo N con $\Phi_m > \Phi_{sc}$. (a) De los dos materiales aislados el uno del otro. (b) Situación de equilibrio térmico después del contacto.

Esta transferencia de electrones deja en el semiconductor una carga positiva fija, debida a los iones sin compensar N_D^+ , creándose una carga negativa en el metal (debida a los electrones transferidos) que iguala a la positiva.

Se forma, por tanto, una región de carga espacial de anchura W_0 en el semiconductor de tipo N. Sin embargo, la carga producida por los electrones en el metal es una carga cuyo espesor es despreciable y es, esencialmente, una carga superficial. La presencia de estas cargas establece un campo eléctrico en el sentido desde el semiconductor al metal.

Una vez alcanzada la situación de equilibrio térmico, el flujo de electrones del semiconductor al metal es igualado por otro flujo de electrones del metal al semiconductor.

En el diagrama de bandas de la Fig. 31b, se ha supuesto que Φ_m , χ y la anchura de la banda prohibida E_g del semiconductor permanecen invariables después de producirse el contacto entre los dos materiales. Por tanto, lejos de la zona de transición y en el semiconductor podemos dibujar este diagrama colocando el nivel de Fermi en el equilibrio E_F y después dibujando las posiciones relativas de E_C , E_V y el nivel de vacío con respecto a E_F (y que por tanto, disminuye en una cantidad $q\cdot(\Phi_m - \Phi_{sc})$). Finalmente, en la zona de transición, el nivel de vacío del semiconductor se ha de aproximar de forma gradual al del metal para mantener la continuidad. E_C y E_V han de

TEMA 3 : Física de la unión P-N.

seguir la misma variación que el nivel de vacío porque $q\cdot\chi$ permanece invariable. Este procedimiento es bastante general y puede ser utilizado para dibujar el diagrama de bandas cualesquiera que sean los dos materiales en contacto. La diferencia entre las dos funciones de trabajo (del metal y del semiconductor) es el potencial de contacto $V_i = (\Phi_m - \Phi_{sc})$. La altura de la barrera para los electrones que fluyen desde el semiconductor al metal es $q\cdot V_i$ mientras que la altura de la barrera vista desde el metal hasta el semiconductor viene dada por:

$$q\cdot\Phi_B = q\cdot(\Phi_m - \chi) = q\cdot V_i + (E_c - E_F)$$

ya que $q\cdot\Phi_{sc} = q\cdot\chi + (E_c - E_F)$. La ecuación $q\cdot\Phi_B = q\cdot(\Phi_m - \chi)$ se conoce como aproximación de Schottky. En la mayor parte de los casos, el potencial Φ_B es varios ordenes de magnitud mayor que $K\cdot T/q$, y la región espacial de carga en el semiconductor se asemeja bastante a la de una unión p-n.

En equilibrio térmico a una temperatura T, sólo una pequeña fracción de electrones de la banda de conducción del semiconductor tendrá energía suficiente para sobrepasar la barrera. Estos electrones que fluyen desde el semiconductor hacia el metal dan lugar a una corriente I_{ms} (recordar que la corriente tiene sentido contrario al movimiento de los electrones). Esta corriente es equilibrada exactamente por una igual y de sentido contrario I_{sm} causada por el flujo de electrones del metal al semiconductor (Fig. 32a).

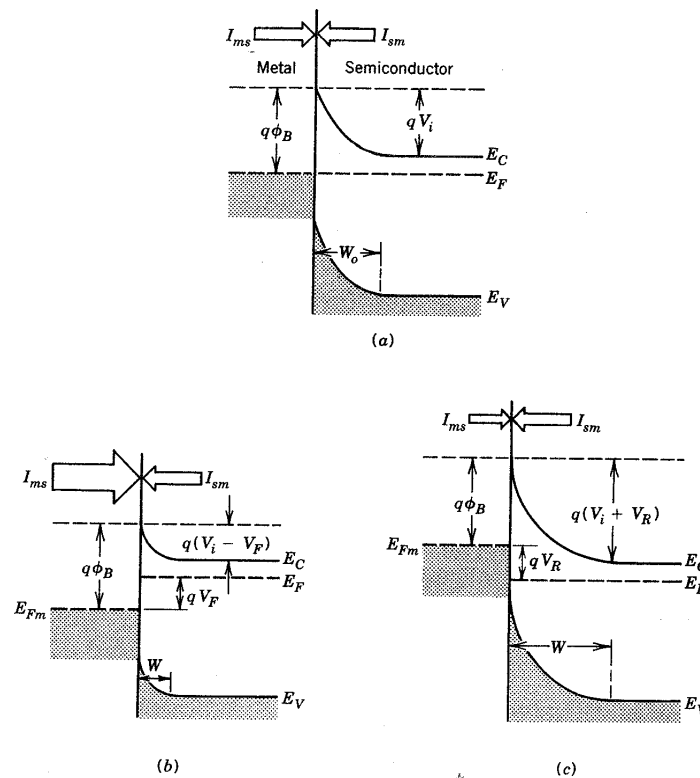


Fig. 32: Diagrama de bandas de energía de una unión rectificadora metal-semiconductor de tipo N en (a) equilibrio térmico, (b) polarización directa y (c) polarización inversa.

Cuando se aplica una tensión negativa al semiconductor con respecto al metal (Fig. 32b), V_F , la altura de la barrera para los electrones del semiconductor disminuye desde $q\cdot V_i$ hasta $q\cdot(V_i - V_F)$. Más electrones podrán por tanto fluir desde el semiconductor hasta

el metal, y I_{ms} se incrementa por encima de su valor en equilibrio térmico. Sin embargo, I_{sm} permanece invariable ya que no hay caída de tensión a través del metal y $q\Phi_B$ permanece casi invariable. Por tanto, hay un flujo neto de corriente del metal al semiconductor.

La aplicación de una tensión de polarización inversa V_R (Fig. 32c) reduce el flujo de electrones desde el semiconductor hasta el metal, y I_{ms} se reduce por debajo de su valor en equilibrio térmico mientras I_{sm} permanece casi inalterada. Por tanto, hay un flujo de una pequeña corriente inversa.

3.8.2. Contacto óhmico metal-semiconductor.

La anterior discusión se ha aplicado a un semiconductor de tipo N con $\Phi_m > \Phi_{sc}$. En el caso de un semiconductor de tipo N con $\Phi_m < \Phi_{sc}$, el diagrama de bandas de energía de metal y semiconductor con los dos materiales aislados el uno del otro se muestra en Fig. 33a. Después del contacto, los electrones fluyen desde el metal hasta la banda de conducción del semiconductor hasta que se alcanza la situación de equilibrio térmico. Esto provoca una caída de potencial ($\Phi_{sc} - \Phi_m$) a través del semiconductor (Fig. 33b). La carga negativa de electrones que se acumulan en el semiconductor se confina en una zona de muy poco espesor y es, en esencia, una carga superficial. Idéntico resultado se obtiene para la carga positiva en el metal. Como no hay una región espacial de carga en el semiconductor, no existe ninguna barrera para la circulación de electrones bien del semiconductor al metal, bien del metal al semiconductor. Cuando se aplica una polarización, toda la tensión aparece aplicada al semiconductor (neutro) como se observa en Fig. 33c y d. Este tipo de contacto no rectificador es llamado a menudo contacto óhmico.

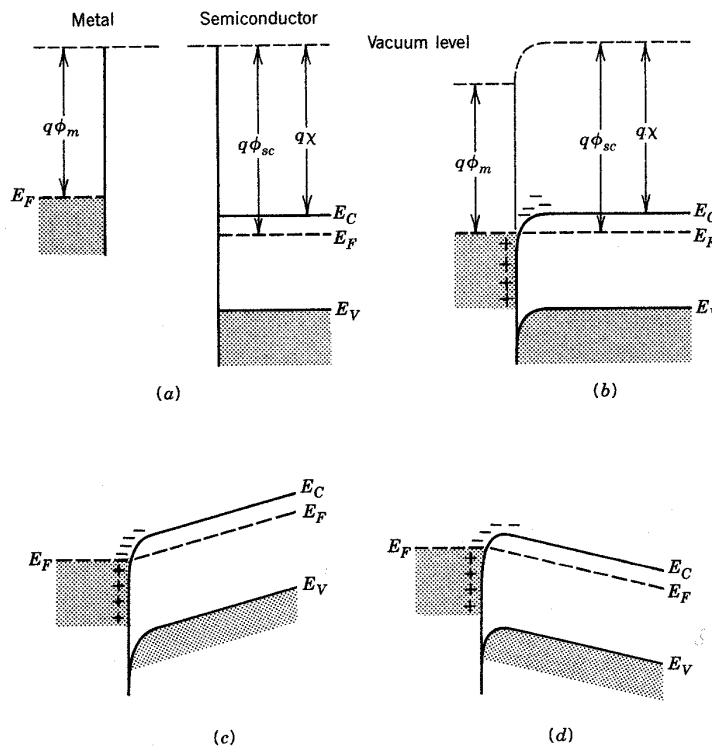


Fig. 33: Diagrama de bandas de energía de una unión metal-semiconductor de tipo N con $\Phi_m < \Phi_{sc}$ (a) Materiales aislados el uno del otro (b) contacto en equilibrio térmico, (c) tensión negativa aplicada al semiconductor y (d) tensión positiva aplicada al semiconductor.

Para un contacto metal-semiconductor de tipo P se tiene por el contrario que:

$$\Phi_m < \Phi_{sc}: \text{Unión rectificadora}$$

$$\Phi_{sc} < \Phi_m: \text{Contacto óhmico}$$

3.9. Estructura p⁺-n-n⁺.

Una vez implementado el diodo de unión p⁺-n ó p-n⁺ se han de colocar los contactos metálicos a cada una de estas regiones semiconductoras a fin de poder polarizar el diodo. Tal y como se ha comentado con anterioridad, la unión metal-semiconductor puede comportarse como una unión rectificadora lo cual en nuestro caso no nos interesaría pues lo que pretendemos obtener es un diodo de unión. Es por ello que en fabricación de uniones p⁺-n ó p-n⁺ es habitual, a fin de facilitar el contacto óhmico (no rectificador), terminar la región ligeramente dopada (n ó p) con una capa muy dopada del mismo tipo. El hecho de tener un contacto metal-semiconductor altamente dopado facilita el contacto óhmico frente a una unión metal-semiconductor débilmente dopado. De esta forma, a pesar de existir una barrera de potencial en la unión metal-semiconductor, la anchura de la región espacial de carga es lo suficientemente estrecha como para permitir a los portadores atravesar dicha barrera por efecto túnel. La utilización de una región altamente dopada no es estrictamente necesaria en todos los casos, así por ejemplo, el aluminio (Al) presenta un buen contacto óhmico con un semiconductor de tipo p de Si, pero no con uno de tipo n por lo cual se requiere una zona de tipo n⁺ antes de contactar con el aluminio.

En la fabricación de las uniones p⁺-n ó p-n⁺ si utilizamos una capa muy dopada para terminar la región ligeramente dopada (n ó p) el resultado es una estructura p⁺-n-n⁺ con una unión activa p⁺-n ó una estructura p⁺-p-n⁺ con una unión activa p-n⁺. Fig. 34a muestra el caso de la estructura p⁺-n-n⁺.

Para dicho caso, la resistencia serie completa del diodo es la suma de las resistencias correspondientes de las zonas p⁺, n y n⁺. Las zonas p⁺ y n⁺ al ser altamente dopadas presentan una baja resistencia. Para la resistencia de la zona débilmente dopada:

$$\text{Resistividad: } \rho_n = \frac{1}{q \cdot N_D \cdot \mu_n} \text{ y resistencia } R_n = \rho_n \cdot \frac{L}{A}$$

A fin de disminuir la resistencia serie del diodo y, por tanto, las pérdidas en conducción, se puede disminuir la longitud de esta zona débilmente dopada, lo cual puede ser de gran utilidad en dispositivos que manejen altas corrientes. Fig. 34b muestra la estructura p⁺-n-n⁺ en ausencia de polarización. En esta puede observarse como la región espacial de carga se extiende fundamentalmente sobre la región débilmente dopada (zona n). También se muestra el diagrama de bandas de energía en ausencia de polarización donde podemos ver como el nivel de Fermi es constante a lo largo de la muestra semiconductor y como E_F está más próximo a E_c en la zona n⁺ (altamente dopada) que en la zona n.

Pero una región ligeramente dopada corta también puede conducirnos al “punch-through” bajo condiciones de polarización inversa, como puede observarse en Fig. 34c en que la anchura de la región espacial de carga sobrepasa la zona n, produciéndose la ruptura de la unión p⁺-n antes de lo esperado.

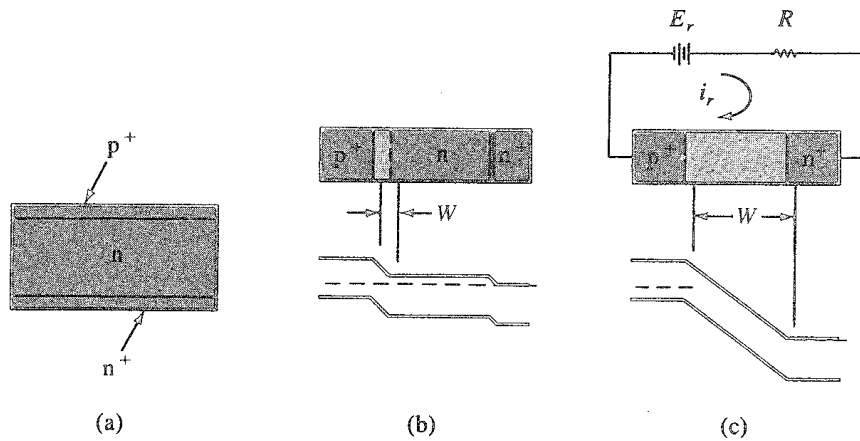


Fig. 34: Unión p^+-n-n^+ . (a) Configuración del dispositivo. (b) En ausencia de polarización. (c) "Punch-through" en condición de polarización inversa.

La longitud de la zona n debe ser escogida como una solución de compromiso. Por una parte debe ser corta para disminuir la resistencia serie. Por otra parte, su longitud mínima está limitada por las especificaciones máximas de tensión inversa que deba ser capaz de soportar el dispositivo.

En la práctica (Fig. 35), suele utilizarse una zona n^+ de mucha mayor longitud que la zona n (alrededor de $500\mu\text{m}$ frente a los de 2 a $10\mu\text{m}$ de la zona n). De esa forma la zona n^+ puede ser utilizada como soporte físico (para dar consistencia al dispositivo) y, por otra parte, se disminuye la resistencia serie del dispositivo. En el diseño y construcción se ha de prever cual va a ser la anchura máxima de la región espacial de carga para que ésta no sobrepase la longitud de la zona n produciéndose el efecto "Punch-through".

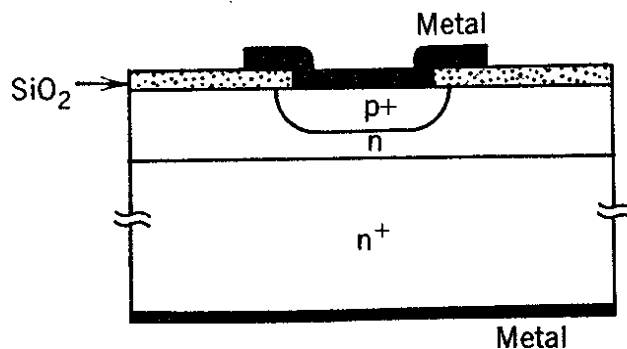
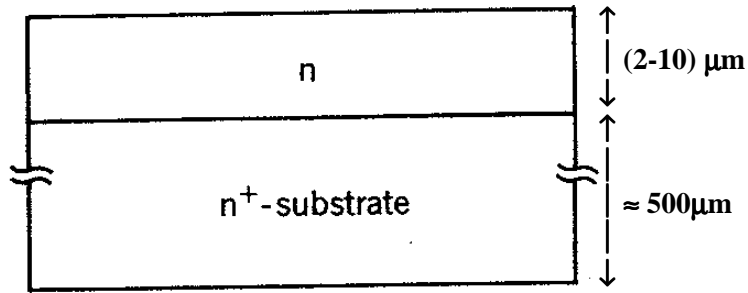


Fig. 35: Unión p^+-n-n^+ vertical.

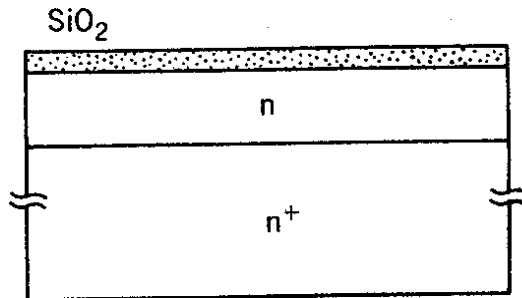
3.10. Construcción del diodo

A continuación se muestran las distintas etapas de fabricación de una unión p^+-n-n^+ vertical.

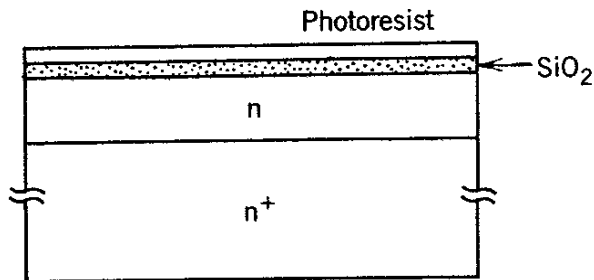
- (a) Se obtiene un sustrato n^+ de aproximadamente $500\mu\text{m}$ de espesor (para dar consistencia física al diodo).
Por crecimiento epitaxial se crea sobre el sustrato una capa de silicio de entre 2 y $10\mu\text{m}$ de espesor.



(b) Se deposita por oxidación térmica una capa de dióxido de silicio (SiO_2).

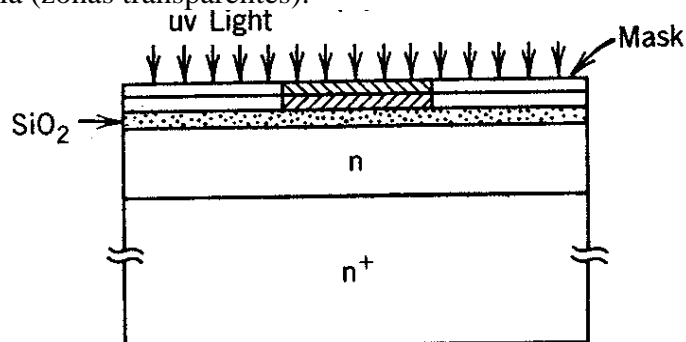


(c) La superficie es cubierta con una capa de material fotosensible (fotoreistencia).



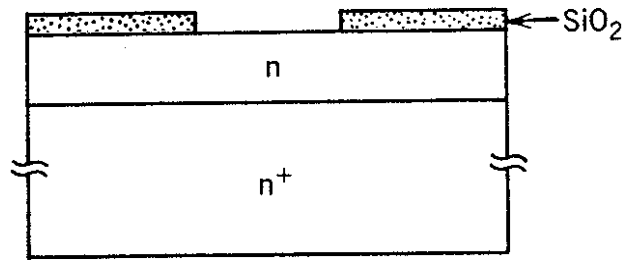
(d) Se sitúa una máscara sobre la fotoreistencia, alineada y expuesta a radiación ultravioleta (la máscara es un cristal que tiene zonas transparentes y zonas opacas en una secuencia predefinida).

- La fotoreistencia es polimerizada en las regiones en que se expone a la luz pero no en las regiones opacas.
- En las zonas en que la resistencia es polimerizada queda una fina película de resistencia (zonas transparentes).

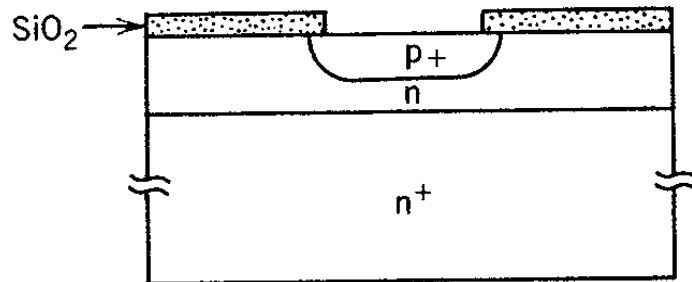


(e) La máscara es eliminada. La capa de óxido en las zonas no cubiertas por la resistencia son eliminadas y después esta resistencia es eliminada.

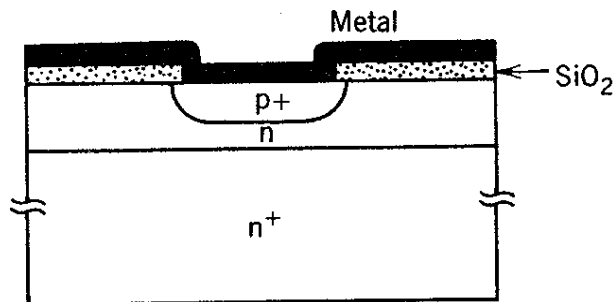
TEMA 3 : Física de la unión P-N.



- (f) Se difunde boro (o cualquier otro tipo de sustancia trivalente) para formar la zona de tipo p^+ a través de la ventana abierta en la capa aislante de SiO_2 . El boro se difunde fácilmente en el silicio pero no en el SiO_2 .



- (g) Una fina película de aluminio se deposita sobre la superficie.



- (h) El área metalizada es cubierta con material fotosensible (fotoresistencia), y se utiliza otra máscara para identificar las áreas donde el metal tiene que ser preservado.

Se elimina la zona metálica no deseada y disuelve la resistencia. En la parte inferior se coloca un contacto metálico.

