

Electrónica física

Principios físicos, materiales y dispositivos

Rafael Quintero Torres





RAFAEL QUINTERO TORRES es egresado de la Universidad Autónoma Metropolitana, Unidad Azcapotzalco. Estudió la maestría en ingeniería eléctrica en el Centro de Investigación y Estudios Avanzados del IPN y el doctorado en ciencias de la ingeniería en Auburn University, Estados Unidos. Ha trabajado en Motorola de México como ingeniero de desarrollo en el área de obleas (fabricación de dispositivos semiconductores de potencia). Actualmente es profesor-investigador de la UAM-Azcapotzalco, en el Departamento de Electrónica.

En los centros educativos en que se formó, el autor trabajó en las propiedades ópticas no lineales y eléctricas de materiales orgánicos. Ha publicado varios trabajos científicos y presentado conferencias en congresos internacionales, los más recientes en el área de optoelectrónica. Su interés científico actual se centra en las propiedades ópticas y eléctricas de los materiales que pueden ser empleados en electrónica.

ELECTRÓNICA FÍSICA

SERIE

Material de apoyo a la docencia

(Teoría y prácticas de laboratorio; problemarios)

Electrónica física

217583

C.B. 2893247

Principios físicos, materiales y dispositivos

Rafael Quintero Torres



2893247

UNIVERSIDAD AUTÓNOMA METROPOLITANA

Rector General
Dr. Julio Rubio Oca

Secretario General
M. en C. Magdalena Fresán Orozco

UNIDAD AZCAPOTZALCO

Rector
Lic. Edmundo Jacobo Molina

Secretario
Mtro. Adrián de Garay Sánchez

Coordinador de Extensión Universitaria
Lic. Alberto Dogart Murrieta

Jefe de la Sección Editorial
Lic. Valentín Almaraz Moreno

UA:
TK 7871
p. 57

Portada:
Adriana Espinosa/Sans Serif Editores

Composición tipográfica, diseño, producción y cuidado editorial:
Sans Serif Editores, telfax 674 60 91

Primera edición 1996

ISBN: 970-620-732-5

© Universidad Autónoma Metropolitana
Unidad Azcapotzalco
Av. San Pablo núm. 180
México, 02200, D.F.

Impreso en México
Printed in Mexico

AGRADECIMIENTOS

MI AGRADECIMIENTO A LOS PROFESORES Jaime Grabinsky y Alex Polo Velázquez por sus múltiples correcciones y sugerencias a lo largo del trabajo, especialmente por el maratón en que consistió la revisión final, así como al profesor Juan Vargas Rubio por sus valiosos comentarios.

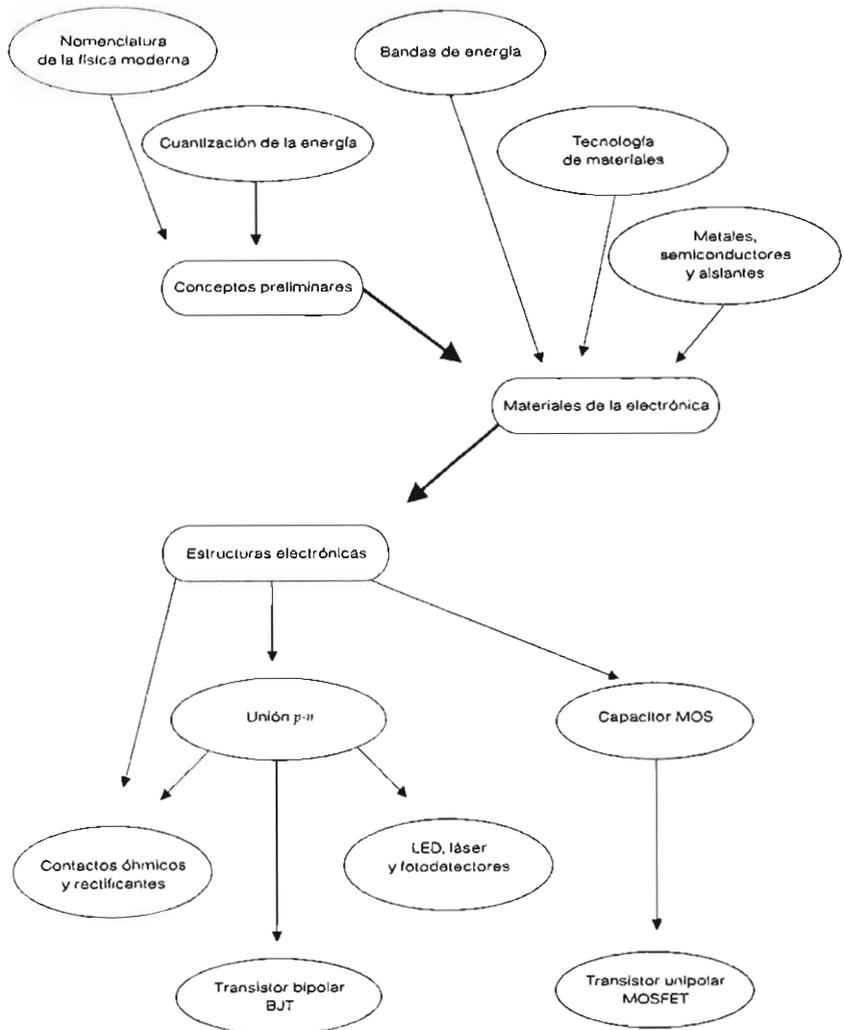
La inspiración para la realización del trabajo surgió de los estudiantes de electrónica física,

además de que su colaboración fue muy importante. Particularmente agradezco la ayuda de la generación de 1995, invierno.

La mayor parte de este trabajo fue realizada a expensas de mi tiempo familiar, así que debo reconocer la comprensión, paciencia y estímulo de mi familia: Rosa María, y Rafael, quien me permitió "jugar a la computadora" innumerables fines de semana.

R. Q. T.

EL MAPA



PRESENTACIÓN

EL PROPÓSITO PRINCIPAL DE ESTE TRABAJO es ayudar en el aprendizaje de la electrónica física, evitando las distracciones que son consecuencia de tener que tomar notas en clase, y dando a los estudiantes el material necesario con anticipación, para que estén preparados antes de presentarse a clase. De esta manera, el aprendizaje y la enseñanza se facilitan en gran medida.

Esta obra se propone cubrir con amplitud su tema, y abarca poco más de lo que se estudia en un trimestre, con el fin de encauzar el interés estudiantil, que con ayuda de investigación bibliográfica permitirá ampliar la información presentada.

La redacción se realizó de manera cuidadosa, a fin de exponer las ideas claramente. Si algo no resulta claro, vuelva al punto e intente de nuevo: puede ser un concepto clave.

Es importante tener en cuenta, para aprovechar mejor el presente texto, lo siguiente:

a) El material de las unidades uno y dos no es exclusivo de nuestra Unidad de Enseñanza-Aprendizaje (UEA); es material básico y común a las áreas de química y física.

b) Si quiere conocer algo acerca de los dispositivos de estado sólido, el origen de los parámetros que aparecen en las hojas de ingeniería, el funcionamiento y las limitaciones de los dispositivos básicos con la menor densidad de milagros, no tiene más remedio que estudiar con detenimiento el texto.

c) Quizá haya otra manera más fácil de estudiar lo mismo, pero no tan divertida. Este asunto debe tomarse con mucha calma, para conocerlo por aproximaciones sucesivas, sacrificando la formalidad en aras de formar ideas claras. Recuerden que, antes que futuros ingenieros, son estudiantes, y es importante que disfruten del placer de aprender, cualquiera que sea el tema.

INTRODUCCIÓN

LOS OBJETIVOS QUE SE INTENTARÁ ALCANZAR CON este texto son:

- Entender cómo se emplean las propiedades de los materiales en dispositivos electrónicos útiles y cómo éstos transfieren sus limitaciones.
- Explicar el comportamiento eléctrico de los materiales, en particular el de los semiconductores homogéneos.
- Conocer cómo surgió el concepto de bandas de energía y manejarlo en el estudio de los semiconductores homogéneos y dispositivos.
- Entender el principio físico del funcionamiento de las estructuras simples (unión *p-n*, estructura MOS, transistor, etcétera).
- Dar cauce a inquietudes relacionadas con la comprensión de nuevos “dispositivos” electrónicos, como los obtenidos de la interacción de la luz y los materiales, entre los cuales se encuentran las memorias magnetoópticas o los láseres.

Probablemente piense usted que este panorama es muy amplio, además de que el desarrollo

se da a gran velocidad, y que por lo tanto le será difícil abarcar el material en 12 semanas, pero no se desanime.

La electrónica física se puede entender como el conocimiento que ha permitido el desarrollo de la electrónica hasta los niveles de complejidad que son comunes hoy día. La electrónica física permite explicar el funcionamiento de los sistemas simples, desde la conducción en un metal hasta el desarrollo de tecnologías, como la que hace posible la fabricación de los modernos circuitos integrados (VLSI) o los cada vez más eficientes diodos láser.

Entre sus campos se pueden mencionar los siguientes:

- Física de semiconductores: Explicación y predicción de las características de los semiconductores.
- Física de dispositivos: Diseño de dispositivos de estado sólido.
- Tecnología de los semiconductores: Fabricación óptima de dispositivos.
- Microelectrónica: Diseño de circuitos integrados.

CAPÍTULO I

CONCEPTOS PRELIMINARES

LA ELECTRÓNICA FÍSICA que aquí estudiaremos se basa en la física moderna. Por ello es necesario partir de conceptos un tanto alejados de lo cotidiano. La primera parte tiene como propósito introducir algunas de las ideas que forman parte de la física moderna. Es claro que entender estos conceptos en una primera presentación es difícil, pero la segunda vez que se escuchan suenan familiares. La segunda parte incluye conceptos relacionados con los dispositivos ópticos, los emisores de luz y los cambios que presenta la luz al interactuar con los materiales. Esto abrirá la puerta al estudio de dispositivos optoelectrónicos. Finalmente se presentan dos de los resultados de la mecánica cuántica que más ayudan en el entendimiento de los materiales, la cuantización de la energía y el efecto túnel.

FÍSICA SUBMICROSCÓPICA

Siempre que hemos emprendido el estudio de un nuevo tema, ha sido necesario recurrir a una

terminología básica que nos permita manejar y entender conceptos e ideas más elaborados. Este curso no es la excepción, y menos ahora que tratamos aspectos tan distantes de nuestro sentido común, lo que no sucedía con la mecánica clásica, por ejemplo. En particular, la estructura interna de los átomos no sigue las leyes clásicas y nuestra intuición no nos ayuda a predecirla.

Los sistemas que se tratarán en adelante, como los electrones, tienen una masa y un tamaño tan pequeño que no siguen las reglas de los cuerpos de gran masa, gran tamaño y pequeñas velocidades, como los autos, las personas, etcétera.

Así, después de un gran esfuerzo intelectual, de muchos años de trabajo y muchas contradicciones y controversias, se propuso una mecánica para los cuerpos del mundo submicroscópico.

Nuestro objetivo ahora no es presentar una introducción de esta mecánica, sino saber de su existencia y conocer algunos de sus conceptos y resultados más elementales, los cuales emplearemos en el transcurso del texto.

Un comentario sobre esta nueva mecánica:

Su utilidad es ampliamente reconocida, ya que gracias a ella se predijo la existencia de dos efectos de particular importancia en electrónica, a saber: el efecto túnel y el efecto láser.

Es muy diferente de la mecánica clásica en su concepción y terminología, pero no en sus resultados, en el dominio de las grandes masas y las pequeñas velocidades, donde las dos mecánicas son esencialmente la misma.

La ventaja de esta nueva mecánica no radica en que nos permite encontrar la solución

analítica a cualquier problema (lo que, estrictamente hablando, no sucede), sino que nos proporciona otra manera de ver los mismos problemas y por lo tanto entender su comportamiento.

Principio de incertidumbre

Este principio no tiene análogo en la mecánica clásica y es uno de los conceptos principales de la nueva mecánica. Una de sus representaciones más frecuentes es:

$$\Delta x \Delta p \geq h \quad (1.1)$$

donde h es la constante universal de Planck, que tiene el valor de 6.626×10^{-34} [J s], y Δx representa la incertidumbre en la medición de la posición, mientras que Δp representa la incertidumbre en la medición de la cantidad de movimiento lineal ($p \equiv mv$).

En mecánica clásica no existe (y no tiene por qué existir) ninguna restricción respecto a la exactitud del conocimiento de la posición y la velocidad de un cuerpo. Intuimos que la exactitud dependerá de nuestra capacidad para diseñar el experimento y de la resolución de los instrumentos utilizados en la medición.

Si estudiamos una partícula puntual que tiene movimiento sólo en la dirección cartesiana x , estaría totalmente caracterizada con el conocimiento de la posición y la velocidad o cantidad de movimiento, como se ve en la figura 1.1.

Y como dijimos antes, no pensamos que Δx o Δv tengan ningún límite predeterminado.

$$x = x_0 \pm \Delta x \quad (1.2)$$

$$v_x = v_{x0} \pm \Delta v_x$$

Para el sentido común, no hay razón por la cual Δx , Δv o ambas no puedan ser cero, pero para el mundo submicroscópico existe un compromiso. No es posible disminuir los errores arbitrariamente, se debe respetar el principio de incertidumbre, en el cual:

$$\Delta x \Delta v_x \geq h \quad (1.3)$$

¿Cómo creer lo increíble?

La medición de cualquier sistema incluye inherentemente su perturbación, esto es, para medirlo hay que interactuar con él, de tal manera que después de medirlo por primera vez, el sistema cambia. Es decir, no se pueden realizar dos mediciones (x, p) a un mismo sistema.

¿Hasta dónde es esto verdadero?

Medir la temperatura con un termómetro es alterar el sistema, ya que se necesita energía del sistema (interna o suministrada desde el exterior) para obtener la lectura de la temperatura, de tal manera que después de medir el sistema éste tiene una temperatura diferente. Si estamos midiendo la temperatura de un litro de agua podríamos no preocuparnos por el cambio que causamos. Pero si estamos interesados en conocer la posición de un electrón en el espacio y para esto hacemos incidir otra partícula que nos dé indicio de su posición, definitivamente el electrón cambió de posición después de la interacción (y quizá también sus otros parámetros), así que el cambio en este caso sí es importante.

Dicho de otra manera, la incertidumbre en la medida de la posición de una masa de 1 kg, cuya velocidad tiene una incertidumbre de 1 m/s, es del orden de 10^{-34} m(1) (distancia que es 10^{-19}

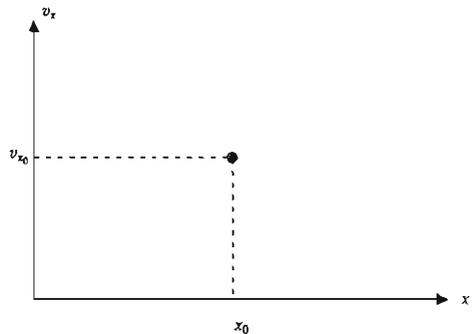


Figura 1.1. Velocidad y posición de un cuerpo definido en una dimensión.

veces más pequeña que el radio del electrón). Sin embargo, para un electrón energético en un átomo de hidrógeno, la incertidumbre es del orden

de 10^{-9} m(2), es decir, 10 veces el tamaño adscrito normalmente a los átomos.

$$(1) \Delta x = \frac{h}{m\Delta v} \approx 10^{-34} \text{ m}$$

$$(2) \Delta x = \frac{h}{m\Delta v} \approx 10^{-9}; \text{ donde } \Delta v = v \approx \sqrt{\frac{2E}{m}} \text{ y } E \approx 13 \text{ eV}^*; m = 0.91 \times 10^{-30} \text{ [kg]}$$

El manejo operacional de este principio puede ayudar a estimar la energía de un electrón confinado a moverse en una dimensión. Este resultado será útil cuando se estudie la distribución de los electrones en un sólido.

¿Qué energía podrá tener un electrón confinado a moverse en una barra unidimensional de longitud L ?

Si suponemos que el electrón está en algún lugar de la barra, entonces $\Delta x = L$.

Empleando el principio de incertidumbre, el error en la medición de la cantidad de movimiento será $\Delta p \geq \frac{h}{L}$. En el mejor de los casos $\Delta p \approx \frac{h}{L}$.

Es posible suponer que el error en la medición es menor que la medición en sí, por lo que es posible escribir: $p \approx \frac{h}{L}$.

Finalmente, con la expresión para la energía cinética se puede escribir la energía que tendrá un electrón.

$$E = \frac{1}{2}mv^2 = \frac{p^2}{2m} = \frac{h^2}{2mL^2} \quad (1.4)$$

Hipótesis de De Broglie

Esta afirmación surgió como un modelo para ajustar algunas suposiciones de la mecánica cuántica, cuenta con una amplia justificación experimental. Su representación habitual es:

$$p = \frac{h}{\lambda} = \hbar k; \quad \lambda = \frac{2\pi}{k} \quad (1.5)$$

donde λ es la longitud de onda que caracteriza a un proceso ondulatorio, p es la cantidad de movimiento lineal y k es proporcional a la cantidad de movimiento (también es un vector). Al mismo tiempo k es una base vectorial muy usada al estudiar los sólidos, por ejemplo en microscopía electrónica, en propiedades de transporte, etcétera.

Esta hipótesis trata de interrelacionar las características intrínsecas de dos fenómenos que son sustancialmente distintos, esto es, relaciona las características de fenómenos ondulatorios (λ) con las características de fenómenos corpusculares (p).

Con esta hipótesis es posible asociar un comportamiento ondulatorio a las partículas y viceversa.

Esta suposición permite manejar con mayor confianza la mecánica cuántica, ya que se basa en asociar a cualquier cuerpo un comportamiento ondulatorio. De ahí el nombre de mecánica ondulatoria.

Bueno, y si esto es cierto, ¿cuál es la longitud de onda de una persona?

Si imaginamos una masa de 75 kg y una velocidad de 10 m/s:

$$\lambda = \frac{h}{p} = \frac{6.6 \times 10^{-34}}{(75)(10)} \left[\frac{\text{Js}}{\text{kgm s}^{-1}} \right] \approx 10^{-37} \text{ m}$$

Se podría afirmar que ésta es una explicación de por qué no vemos que las personas se desplacen como ondas, como cuando se perturba la superficie del agua, ya que no se perciben dimensiones físicas tan pequeñas.

Ahora que se ha mencionado la palabra "ver", una pregunta obligada sería: ¿Qué podemos ver?

Pues bien, en nuestro organismo tenemos diferentes transductores y cada uno permite tener

* 1 eV = 1.60218 × 10⁻¹⁹ J.

acceso a algunos fenómenos que suceden a nuestro alrededor. Por ejemplo:

- Las vibraciones mecánicas de baja frecuencia las detectamos con el tacto.
- Las vibraciones mecánicas de mayor frecuencia las detectamos con el oído.
- La piel nos permite detectar algunas oscilaciones electromagnéticas que corresponden al infrarrojo (calor).
- La vista nos permite discernir entre diferen-

tes oscilaciones de las ondas electromagnéticas, pero en un intervalo muy corto que se denomina "visible".

Así, para detectar el resto del amplio espectro electromagnético es necesario recurrir a un transductor adicional que transforme la señal que deseamos conocer en algo que podamos detectar. Por ejemplo, un radio, un televisor, un detector de centelleo, etcétera.

¿Cuál será la cantidad de movimiento de la luz roja de un láser de He-Ne?

La longitud de onda de este láser es: $\lambda = 632.8\text{nm}$ (un nm es $1 \times 10^{-9}\text{m}$).

$$p = \frac{6.6 \times 10^{-34}}{632.8 \times 10^{-9}} \approx 10^{-27}[\text{kgms}^{-1}]$$

En esencia, ¿qué originó que el comportamiento dual onda-partícula pasara inadvertido durante tanto tiempo?

Claro está, la magnitud de h . ¿Podría usted imaginar al mundo con una h diferente, digamos 10^{-20} , 10^{-10} , o $1[\text{Js}]$?

Hipótesis de Max Planck

Sin temor a equivocarnos, podemos afirmar que con el trabajo de Max Planck (1900) dio principio el vertiginoso desarrollo de la física moderna.

Su hipótesis afirma que, dado que los procesos ondulatorios se pueden comportar como partículas, la luz, al ser una onda electromagnética, también se puede estudiar como un fenómeno corpuscular, y los corpúsculos de luz se denominan fotones.

Como veremos con más detalle en la siguiente sección, una fuente de radiación se caracteriza principalmente por su espectro (longitudes de onda que

radia) y por la radiancia a cada una de esas longitudes de onda. Esa hipótesis relaciona expresamente la energía de los fotones (E) con la frecuencia (ν) o la longitud de onda (λ), ya que la frecuencia y la longitud de onda se relacionan con la velocidad de la luz ($\lambda \nu = c$, $c = 2.9979 \times 10^{10} \text{ cm/s}$).

$$E = h\nu = \hbar\omega \quad (1.6)$$

Esta hipótesis se empleó para explicar los resultados de la radiación del cuerpo negro y del efecto fotoeléctrico.

Si se piensa en un conjunto de osciladores acoplados, entonces la emisión o absorción de energía E es una cantidad proporcional a la frecuencia ν . Así, cuando un oscilador absorbe o emite radiación electromagnética, su energía aumenta o disminuye sólo en una cantidad $h\nu$; para cualquier oscilador con frecuencia ν existe únicamente una constante de proporcionalidad.

¿Qué energía tiene asociada la luz roja del láser de He-Ne, $\lambda = 632.8 \text{ nm}$?

$$E = h\nu = \frac{hc}{\lambda} \approx 3.14 \cdot 10^{-19}[\text{J}], = 1.96[\text{eV}]$$

En este curso es mejor usar electrón volt (eV) como unidad de energía en lugar de joule, ya que proporciona números más razonables. Un eV equivale a la energía que adquiere un electrón después de ser acelerado por una diferencia de potencial de un volt.

$$1 \text{ eV} = 1.60218 \times 10^{-19} \text{ J}$$

NOMENCLATURA ÓPTICA

El ojo humano, en buenas condiciones de iluminación, es capaz de detectar un pequeño intervalo del espectro electromagnético. En la figura I.2 se muestra la respuesta relativa de acuerdo con la Commission Internationale de l'Éclairage en 1924 (CIE). Esto es lo que da origen a dos conjuntos de unidades para el estudio de la radiación electromagnética: las radiométricas, donde la unidad fundamental del flujo radiante es el watt y donde no se consideran algunas longitudes de onda como "especiales", y las fotométricas, donde el flujo luminoso equivalente

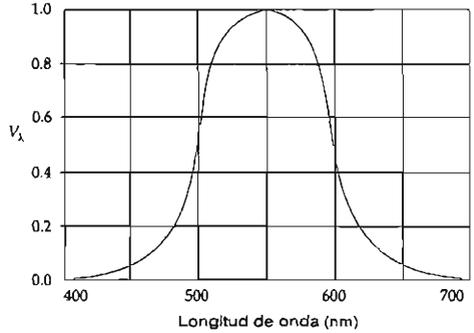


Figura I.2. Curva de respuesta óptica de la CIE.

tiene al lumen como unidad y están determinadas por la visión.

$$(\text{Flujo luminoso en lm}) = 683 \times (\text{flujo radiante en W}) \times V_{\lambda} \quad (1.7)$$

V_{λ} es la respuesta relativa del ojo que tiene el valor de 1 a la longitud de onda de 555 nm que corresponde al color verde. Un watt de energía radiante a 555 nm equivale a 683 lumen.

Términos de radiometría y fotometría

Unidad SI	Término y unidad radiométrica	Término y unidad fotométrica
Q	Energía radiante [J]	Energía luminosa [talbot = lm s]
ϕ	Flujo o potencia radiante [W]	Flujo o potencia radiante [W]
I	Irradiancia [Wm^{-2}]	Illuminancia [$\text{lm m}^{-2} = \text{lux}$]
\mathcal{R}	Intensidad radiante [Wsr^{-1}]	Intensidad luminosa [$\text{lm sr}^{-1} = \text{candela}$]
L	Radiancia [$\text{Wm}^{-2} \text{sr}^{-1}$]	Luminancia o brillantez [$\text{m}^{-2} \text{sr}^{-1} = \text{nit}$]
W	Emitancia radiante [Wm^{-2}]	

Las primeras dos unidades radiométricas no requieren aclaración adicional. La irradiancia es una medida de la cantidad de radiación (potencia radiante) por unidad de área; a menudo el detector no define el área iluminada, así que debe corregirse proporcionalmente. La intensidad radiante es medida en fuentes divergentes y mide la potencia radiada por una fuente pun-

tual como función del ángulo sólido. La radiancia es una propiedad de superficies que emiten o reflejan la radiación y mide la potencia radiada por unidad de área y ángulo sólido definido entre el detector puntual y el área emisora o reflectora. La emitancia radiante mide la potencia total radiada en todas direcciones a partir de un área unitaria.

¿Cuál es la brillantez de un diodo emisor de luz (LED)?

Empezamos por suponer un LED fabricado con un área activa de 0.2 mm de diámetro y que es visto a un metro de distancia. El LED emite a una longitud de onda de 550 nm (verde) y tiene una eficiencia cuántica externa (η) de 0.1% (este número mide la razón de electrones que son convertidos a fotones; los LED amarillos tienen un peor número mientras que los infrarrojos alcanzan eficiencias del 15%). Además, supongamos que emite de manera isotrópica y que el diodo es operado a 2 V y 50 mA.

En estas condiciones el LED se puede trabajar como una fuente puntual, ya que el ángulo creado por el área del LED a un metro de distancia es menor a un minuto de arco.

La potencia radiante total está definida por la energía proporcionada por cada fotón emitido (función de λ , y que posteriormente se verá que depende de la estructura de bandas), por los electrones disponibles para producir fotones (función de la corriente eléctrica) y por la eficiencia de conversión electrón-fotón (η).

$$W = \frac{hc}{\lambda} \frac{I}{e} \eta = \frac{hc}{550 \times 10^{-9}} \frac{50 \times 10^{-3}}{e} 0.001 = 1.13 \times 10^{-4} [\text{W}]$$

De la tabla CIE a esta longitud de onda, un watt es igual a 679 lm, por lo que el LED produce una potencia luminosa de 7.7×10^{-2} lm. Como el flujo luminoso es isotrópico sobre el ángulo sólido 2π la intensidad luminosa a incidencia normal es 1.2×10^{-2} candela.

La definición del ángulo sólido es similar a la definición del radián; el radián es el cociente de la longitud del arco de un círculo entre el radio del círculo. El estereoradián (sr) es el cociente del área de una sección de la esfera entre el radio de la esfera al cuadrado. El ángulo sólido total de una esfera es 4π . Una abertura rectangular $a \times b$, a una distancia h , forma un ángulo sólido (Ω) de la siguiente forma:

$$\Omega = 4 \arctan \frac{ab}{2h\sqrt{a^2 + b^2 + (2h)^2}}$$

Polarización de la luz

El concepto de polarización de la luz es muy importante, ya que muchos dispositivos ópticos controlan y miden la polarización de la luz más que la intensidad. Las memorias magnetoópticas, por ejemplo, se basan en la identificación de pequeños dominios magnéticos analizando la

polarización de la luz que se refleja del material, ya que cambios en el material magnético producen cambios en la polarización de la luz (efecto Kerr magnético).

La radiación electromagnética se caracteriza, además de lo mencionado en la sección anterior, por la propagación del frente de onda y el estado de polarización. La propagación del frente de onda puede ser plana, como la luz del sol vista en una pequeña área en la tierra. Otro caso de propagación puede ser esférica, como las fuentes puntuales, y finalmente las fuentes gaussianas, como la radiación de los láseres de buena calidad, donde la intensidad se distribuye en el espacio como una campana de Gauss.

La polarización de la radiación puede ser plana, en la cual el vector del campo eléctrico vibra paralelo a un plano que incluye la dirección de propagación. La polarización circular se produce cuando en la radiación que se propaga en la dirección cartesiana z , la vibración del

campo eléctrico a lo largo del eje cartesiano x está un cuarto de la longitud de onda fuera de fase respecto a la vibración del campo eléctrico a lo largo del eje cartesiano y , y el campo eléctrico resultante describe un círculo. Si la diferencia de fase es 0 o la mitad de la longitud de onda, se tiene polarización plana. Cualquier diferencia de fase diferente de las anteriores se conoce como polarización elíptica. La reflexión de la radiación de cualquier material modifica la polarización de la radiación. La transmisión de la radiación en un material cristalino altera su polarización por birrefringencia o por dicroísmo, como en los polaroid.

MECÁNICA CUÁNTICA

La mecánica cuántica es la descripción del comportamiento de la materia en todos sus detalles, en particular lo que sucede a escala atómica. A longitudes muy pequeñas las cosas se comportan de manera muy diferente de cualquiera que ustedes hayan visto. No se comportan como ondas ni como partículas ni como nubes ni nada parecido.

¿Cómo sabemos que la mecánica cuántica dice la "verdad"? Recordemos un comentario toma-

do de una idea del profesor Carl Sagan. Los filósofos jónicos, Platón, Demócrito y otros, tenían conceptos científicos similares a los aceptados actualmente como verdaderos. La Edad Media se encargó de destruir ese conocimiento y el Renacimiento de redescubrirlo; desde el punto de vista científico, continuamos en el Renacimiento por sus puntos de vista antidogmáticos. Los jónicos nunca pudieron probar realmente sus afirmaciones acerca de la existencia del átomo, que era un concepto muy revolucionario, pero desde el punto de vista de la verdad, es tan disparatado como suponer que todo está formado por tierra-aire-fuego-agua. Esto fue así porque los jónicos pensaban que el trabajo experimental era trabajo manual, sólo apto para los esclavos. Hoy día muchas sociedades han superado esa etapa. Entonces, ¿las verdades científicas actuales están en peligro de ser superadas y olvidadas? No, al menos desde el punto de vista numérico. Desde el punto de vista filosófico, se espera que la interpretación actual esté incompleta, pero no equivocada. Esta afirmación surgió de la verificación experimental que se hace de sus verdades.

Algunos de los acontecimientos más sobresalientes en el desarrollo de la mecánica cuántica son:

Radiación del cuerpo negro	1901	Planck
Efecto fotoeléctrico	1905	Einstein
Espectros atómicos	1913	Bohr
Dispersión de fotones por electrones	1922	Compton
Principio de exclusión	1924	Pauli
La materia como onda	1925	De Broglie
Ecuación de onda	1926	Schrödinger
Principio de incertidumbre	1927	Heisenberg
Propiedades ondulatorias del electrón	1927	Davisson y Germer
Interpretación física de la función de onda	1927	Born

Ecuación de Schrödinger

Explicar los principios de la mecánica cuántica va mucho más allá de lo que se requiere en este

curso, pero sus ideas son tan completas que es imposible resistirse a presentarlas.

① A toda partícula se le asocia una función de onda compleja:

$$\psi(x, y, z, t) \quad (I.8) \quad \text{hamiltoniano (energía cinética + energía potencial) se transforma en la ecuación básica por medio de los siguientes operadores:}$$

② La expresión clásica de la energía total, el

Variable dinámica		Operador asociado
Posición: x, y, z	\Rightarrow	x, y, z
Cantidad de movimiento lineal: p	\Rightarrow	$\frac{\hbar}{i} \nabla$
Energía: E	\Rightarrow	$-\frac{\hbar}{i} \frac{\partial}{\partial t}$
Ecuación de movimiento: $E = \frac{p^2}{2m} + V(x, y, z)$	\Rightarrow	$-\frac{\hbar}{i} \frac{\partial \psi}{\partial t} = -\frac{\hbar^2}{2m} \nabla^2 \psi + V(x, y, z) \psi$

Donde $V(x, y, z)$ es la energía potencial y ∇^2 representa al operador de Laplace.

③ Las cantidades ψ y $\nabla\psi$ deben ser finitas, univaluadas y continuas, para cualquier x, y, z y t , además de tender a cero en $\pm\infty$.

④ $\psi^*\psi$ es siempre real y se interpreta como la densidad de probabilidad (ψ^* es el complejo conjugado de ψ). Por lo que al integrar sobre todo el espacio debe ser igual a uno:

$$\int_V \psi^* \psi \, dV \equiv 1 \quad (I.9)$$

$\psi^*\psi$ es la probabilidad de que la partícula se encuentre dentro del elemento de volumen dV

en el instante t . Ésta es toda la información que se puede obtener de ψ .

⑤ El valor promedio $\langle \beta \rangle$ (a lo más que se aspira en la nueva mecánica) de cualquier variable dinámica β se calcula de la siguiente forma:

$$\begin{aligned} \langle \beta \rangle &= \int_V \psi^* \beta_{\text{operador}} \psi \, dV \\ \langle x \rangle &= \int_V \psi^* x \psi \, dV \\ \langle p \rangle &= \int_V \psi^* \frac{\hbar}{i} \nabla \psi \, dV \end{aligned} \quad (I.10)$$

Es posible verificar que, al menos en los valores esperados o valores promedio, la mecánica ondulatoria concuerda con las ecuaciones de la mecánica clásica. Resuelva uno de los ejercicios del final del capítulo, para mostrar que la segunda ley de Newton se escribe de la siguiente manera:

$$\langle F_x \rangle = \left\langle -\frac{\partial V}{\partial x} \right\rangle = \frac{\partial \langle p_x \rangle}{\partial t} \quad (I.11)$$

A continuación se presenta una justificación de la ecuación de Schrödinger como una ecuación de conservación de la energía.

Si se parte de la suposición de que cualquier partícula puede ser tratada como una onda, entonces es posible tener una función ondulatoria

que varía en el tiempo y en la posición de la siguiente manera:

$$\psi = Ae^{i(kx - \omega t)} \quad (I.12)$$

Si se sustituyen los valores de k y de ω con las relaciones de Planck y de De Broglie, se obtendrá:

$$\psi = Ae^{i(m\lambda x - Et)} \quad (I.13)$$

En un sistema clásico la energía total es la suma de la energía cinética y potencial, expresada de la siguiente manera:

$$E = \frac{p^2}{2m} + V(x, y, z)$$

Obteniendo la expresión para p^2 y E a partir de la ecuación I.13, derivando la ecuación dos veces respecto a la posición para obtener p^2 y derivando respecto al tiempo para obtener E , finalmente se obtiene la ecuación de movimiento (ésta no es una deducción).

$$-\frac{\hbar}{i} \frac{\partial \psi}{\partial t} = -\frac{\hbar^2}{2m} \nabla^2 \psi + V(x, y, z) \psi \quad (I.14)$$

Cuantización de la energía

Ahora estamos en condiciones de iniciar nuestras experiencias con esta "máquina de procesar información". De inmediato, al analizar la ecuación

de Schrödinger, nos encontramos con que para resolverla es indispensable conocer la dependencia que tiene la energía potencial V de la posición. Este problema es muy fácil de resolver cuando el potencial V es independiente del tiempo. Al emplear la separación de variables para obtener dos nuevas ecuaciones diferenciales, una de ellas se conoce como ecuación de Schrödinger independiente del tiempo, donde ahora E es una constante del sistema. En una dimensión, la ecuación de Schrödinger independiente del tiempo se escribe de la siguiente manera:

$$E\psi(x) = -\frac{\hbar^2}{2m} \frac{\partial^2 \psi}{\partial x^2} + V(x)\psi(x) \quad (I.15)$$

La solución de esta ecuación para una partícula restringida a moverse en un sistema con energía potencial V que adopta la forma que se observa en la figura I.3, dará origen a la cuantización de la energía.

Si introducimos una partícula en esa "caja" restringiremos su movimiento entre $x = 0$ y $x = L$, ya que el potencial es tan alto que se comporta como un muro impenetrable. La energía potencial del muro tiende a infinito y la partícula no puede alcanzar la energía suficiente como para poder saltarlo. Es como estudiar el movimiento de una canica dentro de un popote de longitud L con los extremos sellados, de tal ma-

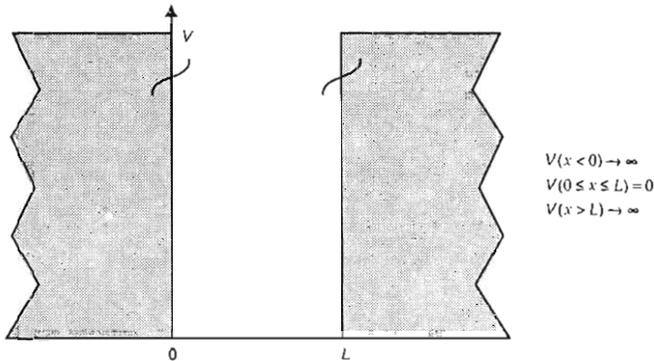


Figura I.3. Partícula confinada en una caja unidimensional de longitud L .

nera que la canica sólo puede estar en algún lugar desde cero hasta L .

Con los comentarios que hicimos podemos fijar las condiciones para la frontera.

$$\psi(0) = 0$$

$$\psi(L) = 0$$

Como a la partícula le asociamos una ψ , podemos resolver la ecuación de Schrödinger entre $x = 0$ y $x = L$, donde la energía potencial V es igual a cero.

$$E\psi(x) = -\frac{\hbar^2}{2m} \frac{\partial^2 \psi}{\partial x^2}$$

Solucionando esta ecuación por los métodos clásicos se obtiene:

$$\psi(x) = A \cos\left(\frac{2mE}{\hbar^2}\right)^{\frac{1}{2}} x + B \sin\left(\frac{2mE}{\hbar^2}\right)^{\frac{1}{2}} x$$

Haciendo que cumpla las condiciones para la frontera se obtiene:

$$A = 0$$

$$E = \left(\frac{n\pi}{L}\right)^2 \frac{\hbar^2}{2m}; \text{ con } n = 1, 2, 3, \dots \quad (1.16)$$

Los números n negativos no son solución porque la cantidad de movimiento k es positiva. El número n no es cero, porque de ser así no habría partícula.

La otra constante B se encuentra a partir de la condición de normalización, esto es:

$$\int_0^L \psi^* \psi dx = 1$$

Y se obtiene que:

$$B = \left(\frac{2}{L}\right)^{\frac{1}{2}}$$

Aquí tenemos pues dos resultados muy interesantes:

1) La energía de la partícula no forma un espectro continuo, sino que es un conjunto discreto. Un valor diferente para cada valor del número cuántico n . A esto se le llama *cuantización de la energía*. La partícula sólo puede tener algunos "estados de energía".

2) La función de distribución nos indica que existe preferencia por encontrar la partícula en algún lugar de la caja.

$$\psi = \left(\frac{2}{L}\right)^{\frac{1}{2}} \sin\left(\frac{n\pi}{L}\right) x$$

Estos dos resultados contradicen nuestro sentido común. Sería de esperar que la partícula tuviera cualquier valor de energía y que existiera la misma probabilidad para hallar la partícula en cualquier lugar de la caja.

Si $n = 1$, es más probable que la partícula esté en el centro de la caja que en las orillas.

Si $n = 2$, es más probable encontrar la partícula en $L/4$ o en $3L/4$.

¿Cuánto debe valer n para que la energía de la partícula sea del orden de las energías macroscópicas?

Las energías macroscópicas se obtienen si $L = 1$ m, $m = 1$ kg y $E = 1$ J, lo que da lugar a tener $n \sim 10^{34}$.

En tales condiciones, es igualmente probable que la partícula esté en cualquier lugar x entre 0 y L . Además, la diferencia en energía entre dos estados contiguos permitidos sería tan pequeña que prácticamente se comporta como un continuo de estados. Esto es lo que se espera en el dominio macroscópico.

¿Por qué una partícula libre no tiene su energía cuantizada?

Una partícula libre, de acuerdo con el ejemplo anterior, no estaría confinada, es decir, la caja tendría una dimensión que tiende a infinito. Conforme L aumenta, la diferencia entre dos estados de energía tiende a cero, así es que los estados de energía son un continuo.

Existe un sistema análogo a la caja de potencial infinito, el átomo de hidrógeno, donde el electrón se encuentra confinado en una determinada región por la atracción que el núcleo (positivo) ejerce sobre el electrón (negativo). Evidentemente, la energía que confina al electrón no es infinita, así que el electrón sí podrá salir de la atracción del núcleo y en ese momento se dice que el átomo está ionizado.

El sistema núcleo-electrón existe sólo cuando se cumple su regla de cuantización:

$$E = -\frac{q^4 m}{8 \epsilon_0^2 h^2 n^2} = -\frac{R}{n^2} \quad (1.17)$$

donde $R = 13.6 \text{ eV}$, y se conoce como constante de Rydberg.

Además, se puede decir que en el sistema atómico real la energía está cuantizada, como en la caja de potencial. Cuando el número cuántico n aumenta, la energía de enlace tiende a cero.

Para cada valor de energía permitido, el sistema se encuentra en un "estado estacionario", esto es, una de las configuraciones posibles del átomo.

Un análogo gravitacional en donde se pueden representar los estados ligados y libres es el siguiente:

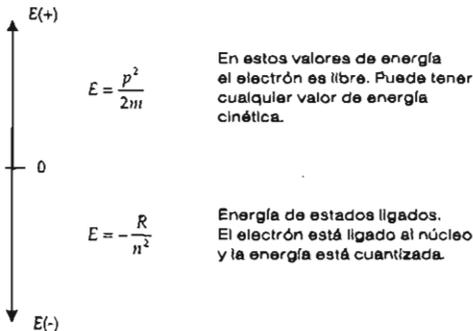


Figura 1.4. Espectro de energía para un átomo simple.

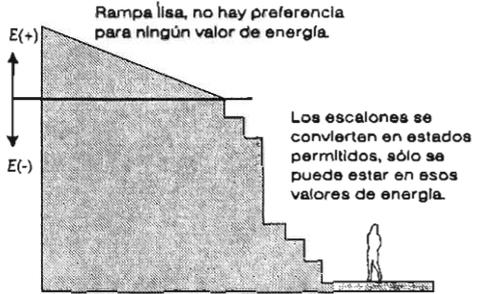


Figura 1.5. Analogía gravitacional de estados libres y ligados.

Efecto túnel

Para encontrar otros resultados elementales de la mecánica cuántica, basta con resolver la ecuación de Schrödinger para cada sistema particular. Por ejemplo, veamos el siguiente sistema, donde la diferencia del problema anterior radica en tener una altura de potencial finita, digamos $V = a$ (véase la figura 1.6).

La solución para la función de onda Ψ ahora se extiende más allá de $x = 0$ y L :

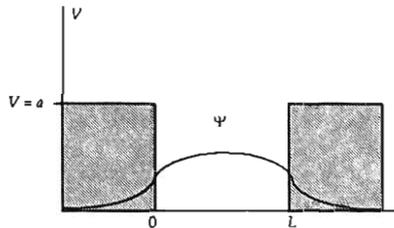


Figura 1.6. Potencial de una caja con muros de potencial finitos.

En este caso existe una región dentro del muro ($x < 0$ y $x > L$) donde la función de onda Ψ no

desaparece. Aquí, además de ser sorprendente, es increíble. ¿Pero qué pasa si hacemos delgados los muros?

CONCLUSIONES

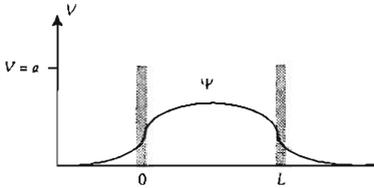


Figura 1.7. Tuneleo cuántico de una partícula.

El mundo submicroscópico está regido por una nueva teoría física, conocida como la ecuación de Schrödinger (análoga a la segunda ley de Newton).

Una serie de conceptos fundamentales en la mecánica cuántica permite explicar la cuantización de la energía para las partículas ligadas.

Los resultados de la mecánica clásica se pueden obtener a partir de la mecánica cuántica.

El tuneleo cuántico es resultado de la extensión de la función de onda más allá de los límites clásicos de los cuerpos.

Entonces nos encontramos con que existe la probabilidad de que una partícula atraviese la pared (y decimos que la atraviese porque es igual de válido si la energía de la partícula es menor a $V = a$, así que no puede brincar la pared). Es posible encontrar la partícula en algunas de las tres cajas. A esto se le conoce como efecto túnel. Se pueden atravesar los muros (de potencial) por efecto túnel cuántico y la probabilidad aumenta conforme el muro es más bajo y su espesor menor.

Hasta el momento se ha pensado en problemas muy simples, propios de una partícula, pero nosotros queremos estudiar sistemas complejos, muchos electrones, electrones moviéndose por una diferencia de potencial, etc. Así que la física tiene que modelar estos problemas complejos, resolver sistemas más o menos simples y, a partir de esto, obtener información sobre problemas complejos.

Ejercicios

Yo sí entendí la teoría,
sólo que no puedo resolver los problemas.
ANÓNIMO

- 1.1.* Explique si un "cuerpo negro" siempre se ve negro.
- 1.2. Explique por qué un pedazo de metal se pone incandescente con un color rojo brillante a 1 100 K; sin embargo, a esta misma temperatura, un pedazo de cuarzo simplemente no brilla. (Sugerencia: relacione este problema con la radiación del cuerpo negro.)
- 1.3. Explique si un tubo de televisión emite rayos X.
- 1.4.* Usando el modelo del átomo de hidrógeno, diga: ¿Cuál es la frecuencia y energía de luz requerida para ionizar al electrón desde el estado base? ¿Cuál sería la frecuencia y energía de la luz para mover el electrón del estado $n = 2$ al siguiente nivel de energía ($n = 3$)? En cada caso, diga el color de la luz.
- 1.5. De una onda que viaja en el agua, si nadie piensa en ella ni nadie la ve, ¿podría decirse que la onda existe? ¿Qué respuesta daría para una función de onda en mecánica cuántica?
- 1.6. ¿Cuál de las siguientes afirmaciones describe mejor la esencia física del principio de incertidumbre?
 - a) Todas las mediciones de x o p son inexactas.
 - b) Es imposible medir con un error arbitrario simultáneamente dos variables conjugadas como x y p .

* Los ejercicios marcados con asterisco están resueltos al final del libro.

- c) Es imposible medir correctamente x y p .
- d) Los incrementos de x y p son mayores o iguales que h .
- e) El error en la medición de x , al igual que en la medición de p , siempre es diferente de cero.

I.7.* Obtenga la ecuación I.11.

I.8.* ¿Cuál de las siguientes restricciones haría válida a Ψ como una Ψ asociada a un sistema físico?, donde $\Psi = ae^{bx}$:

- a) $a = 1, b = 0, x > 0$.
- b) $a = 0, b = 1, \infty > x > -\infty$.
- c) $b = -x^2, \infty > x > -\infty$.
- d) $b = x, x > 0$.
- e) $a = \text{sen } x, b = 1, \infty > x > -\infty$.

I.9.* ¿Cuál es la mínima frecuencia de la luz que causaría fotoemisión de un metal con función trabajo de 2.4 eV? ¿Cuál es la máxima energía cinética de un electrón fotoexcitado con luz de 300 nm de longitud de onda?

I.10.* Un electrón en un microscopio electrónico se acelera por un voltaje de 25 kV. ¿Cuál es la longitud de onda de De Broglie?

I.11. ¿Cuál de las siguientes afirmaciones da cuenta de la explicación para la cuantización de energía?

- a) Si las masas son muy pequeñas, la energía está cuantizada.
- b) Si el sistema está ligado, la energía está cuantizada.
- c) La periodicidad de la función seno.
- d) El valor de h .

I.12. ¿Cuál de las siguientes afirmaciones da cuenta del efecto túnel?

- a) Es resultado de que la partícula esté confinada.
- b) Si se tiene un potencial suficientemente pequeño y una x suficientemente pequeña, entonces existe el efecto túnel.
- c) Todas las partículas confinadas pueden tunelar su confinamiento.
- d) Toda partícula puede atravesar barreras de potencial siempre y cuando su energía sea una fracción del potencial de la barrera.
- e) El valor de la densidad de probabilidad $\psi\psi^*$.

I.13.* Se tiene un arreglo similar al de un cine: fuente de luz, pantalla, espectador. Compare una fuente incandescente isotrópica y un láser que recorre periódicamente toda la pantalla rectangular de 4×8 metros. ¿Qué relación de potencia radiante deben tener las fuentes para que produzcan la misma irradiancia sobre la pantalla? ¿Qué relación existe entre la potencia radiante de una fuente isotrópica y el flujo radiante que recibe un espectador, si la pantalla está a 40 metros de la fuente y el espectador está a 20 metros de la pantalla? Suponga que se refleja 25% en la pantalla y que la fuente y el espectador están en la dirección de la línea normal al plano de la pantalla.

CAPÍTULO II

MATERIALES PARA LA ELECTRÓNICA

ES INTERESANTE NOTAR LA RELACIÓN que guarda la ciencia, entendida como la explicación de un fenómeno, con la tecnología, como el proceso que permite transformar el conocimiento en alguna aplicación. Normalmente, la ciencia surgía y mucho tiempo después lo hacía la tecnología: por ejemplo, la teoría del transistor MOSFET se propuso en 1930, mientras que el primer transistor con buenas características funcionales se obtuvo apenas en 1960. Con el paso del tiempo el lapso entre ciencia y tecno-

logía se ha ido desvaneciendo hasta contar con ejemplos donde el control actual de los procesos de fabricación ha permitido obtener materiales con propiedades aún no entendidas del todo. Esta tendencia de aplicar rápidamente el conocimiento puede explicarse, en parte, porque la electrónica está conducida por innovaciones; como un buen ejemplo están los productos electrónicos actuales. A continuación se presenta una lista incompleta de algunos hechos sobresalientes en la historia de la electrónica.

1874	F. Braun	Inicia la investigación de rectificadores metal-semiconductor.
1879	E.H. Hall	Demuestra experimentalmente que los portadores de carga en los metales son negativos, con algunas excepciones.
1881	Stoney	Llama "electrones" a los portadores negativos.
1904	Bose	Rectifica el contacto puntual.
1907	Round	Descubre la electroluminiscencia.
1908	K. Baedeker	Modifica la conductividad de un material (yoduro de cobre) al impurificarlo (con yodo).
1911	Königsberg y Weiss	Introducen el concepto de semiconductor.
1926	J.E. Lilienfeld	Teoría del transistor MOSFET.
1931	Wilson	Teoría del transporte para semiconductores según la teoría de bandas.
1938	W. Schottky	Teoría de la unión <i>m-s</i> (barrera de Schottky).
1947		Antes de este año los semiconductores sólo se usaban como termistores, fotodiodos y rectificadores.
1948	John Bardeen y Walter Brattain	Desarrollan el primer transistor de contacto puntual.

1949	William Shockley	Teoría del transistor bipolar de unión (BJT).
1950	Hall y Duncap	Obtención de uniones $p-n$ por aleación.
1950	W. Shockley	Teoría básica de la unión $p-n$ (I versus V).
1951	H. Welker	Propone que los materiales compuestos con los grupos III y V de la tabla periódica sean candidatos a semiconductores.
1951	Teal, Sparks y Buelhor	Crecimiento de buenos semiconductores.
1952	W.G. Pfann	Purificación de semiconductores por el principio de refinación por zonas.
1952	W. Shockley	Análisis de la resistencia controlada por voltaje JFET.
1952	SIEMENS	Método industrial para producir cristales.
1954	Chapin, Fuller y Pearson	Desarrollo de la primera celda solar.
1956	J.L. Moll y M. Tanenbaum	Construcción y modelo del primer tiristor.
1956	M. Tanenbaum y Thomas	Difusión en estado sólido.
1956	Premio Nobel	Shockley, Bardeen, Brattain, por su trabajo con transistores.
1957	SIEMENS	Producción en serie de transistores de germanio.
1958	Sah, Noyce, Shockley y Moll	Teoría de la unión $p-n$ (I versus V).
1958	Esaki	Diodo Esaki (túnel).
1959	Moll, Pfann y Garnett	Proponen la estructura MOS como un capacitor controlado por voltaje.
1960	Kilby (Texas Ins) y Noyce (Fairchild)	Invencción de los circuitos integrados.
1960	Kahng y Atalla	Fabricación de un transistor MOSFET.
1960	Theverer y Kleihack	Dispositivos por técnicas epitaxiales.
1962	Hall y Genner	Láser de semiconductores que opera a muy bajas temperaturas.
1963	Gunn	Oscilador de microondas.
1964	Premio Nobel	Townes, Basov, Prokhorov, por el trabajo con máser y láser.
1965		Antes de esta fecha los CI están dominados por bipolares, a partir de este momento salen del laboratorio los CI MOS.
1966	Merd	Teoría de MESFET.
1967	Lehrer y Hooper	Fabricación de MESFET de GaAs.
1970	Boyle Smith	Fabricación de los CCD (dispositivos de acoplamiento de carga).
1970-1979		En esta década empieza el gran avance de los sistemas MOS; surgen los microprocesadores y las DRAM (dynamic random access memories).

1972	Premio Nobel	Bardeen, Cooper, Schieffer, por el trabajo de superconductividad.
1972	Premio Nobel	Esaki, Giaver, Josephson, por el trabajo de tuneleo de semiconductores y superconductores.
1974	INTEL	Desarrollo del primer microprocesador en un CI.
1975	Pickard	Fabricación de uniones $p-n$ por implantación de iones.
1977	Wilmsen	Transistor MOS de GaAs.
1980	V.L.S.I.	Hace posible un millón de dispositivos en un chip de aproximadamente 4 mm.
1982	Bean	Dispositivos por epitaxia de haces moleculares.
1985	IBM	Memoria de computadora de un millón de bits en un trozo de silicio de 5.5 mm por 10.5 mm.
1988	Premio Nobel	Muller, Bednorz, por su trabajo en superconductores de alta temperatura crítica.
1992	IBM, SIEMENS y TOSHIBA	Inicio del programa hacia la memoria DRAM de 256 Megabits.
1995	IBM, SIEMENS y TOSHIBA	Presentan el primer chip de 256 Mb: 286 mm ² y 26 ns.

BANDAS DE ENERGÍA

Antes de discutir un modelo analítico para predecir la existencia de las bandas de energía, detengámonos en algunos de los aspectos cualitativos.

El concepto de bandas de energía es primordial para explicar y entender el comportamiento de todos los fenómenos que estén determinados por los electrones. Las propiedades eléctricas, térmicas, ópticas y magnéticas de los materiales están determinadas en gran medida por los electrones. Expondremos dos ideas que explican la formación de las bandas: la teoría del enlace molecular ampliamente empleada por los estudiosos de la química, y el modelo de Kronig-Penney que permite obtener aspectos cuantitativos.

¿Cómo esperamos que se modifique el diagrama de niveles permitidos (energías para las cuales el sistema puede existir) de un átomo de silicio aislado, respecto a un sólido de silicio (muchos átomos interactuando)?

Pauli propuso un principio en 1925 que ayuda a dar respuesta a esta pregunta: "En un momento dado, únicamente un electrón puede estar en un estado cuántico particular".

Este enunciado de la física puede, por coinci-

dencia, compararse con un aspecto que resulta cotidiano para los estudiantes.

Imaginen un salón de clases en la UAM donde se impartirá una asignatura del primer trimestre. La primera persona en llegar al salón escoge un lugar de su preferencia. La segunda persona en llegar, si es conocida de la anterior, procurará tomar asiento cerca de ella; de no ser así, es probable que busque un lugar de su preferencia, pero alejado una cierta distancia. Al ocupar el salón por completo, se podría identificar a grupos de conocidos o simpatizantes.

¿Qué pasa cuando todos los asientos están ocupados y llega otro alumno? ¿Por qué prefiere quedarse de pie o traer otra silla en vez de sentarse en las piernas de alguien que está sentado? Tal vez, al igual que los alumnos, los electrones tienen más reglas que las impuestas por el espacio físico que ocupan.

Regresemos ahora a los sistemas atómicos. Imaginen que se quiere solucionar un sistema de muchos átomos. Si se pudiera pensar que la solución se obtiene al considerar cada átomo como algo aislado, al resolver cada uno de ellos independientemente se tendría que todos estarían en su estado base y al formar la solución del sistema total, todos los electrones estarían

en el mismo estado, con la misma energía, lo que no es posible.

Después de exponer esto, quizá quede claro que el principio de exclusión de Pauli no es más que una regla que observa la naturaleza cuando interactúan muchos cuerpos. Ésta es una manera muy particular de expresar el comportamiento de la naturaleza.

En el caso de un átomo de silicio, es bien sabido que mientras esté aislado tiene una serie de estados discretos permitidos que se conocen como: $1s, 2s, 2p, 3s, 3p, 3d, 4s, 4p, 4d, 4f, 5s, 5p, 5d, 5f, 6s, 6p, 6d$, etc., y tiene únicamente 14 electrones, que en su estado base (mínima energía) ocupan hasta el nivel $3s, 3p$.

Cabe aclarar en este punto que un átomo se caracteriza por cuatro números cuánticos, tres relacionados con las coordenadas espaciales y uno con el "espín", coordenada interna del electrón. La energía de los electrones se define por esos números cuánticos. Cada orbital atómico ($1s$, por ejemplo) se define por dos números cuánticos. Dentro de cada orbital existen los electrones equivalentes y su número está definido por los otros dos números cuánticos (2 en el s , 6 en el p , etc.). En los problemas del capítulo anterior sólo aparecía un número cuántico n , porque se estudiaron sistemas de una dimensión y de una partícula.

Una de las ideas más simples de la teoría del enlace molecular será empleada para demostrar la existencia de las bandas de energía. De acuerdo con ésta, es posible entender el comportamiento energético de cualquier sistema atómico

o molecular independientemente de su estado de agregación, número de constituyentes o condiciones de presión o temperatura.

En el diagrama que se presenta a continuación se esquematiza la transición de un átomo aislado

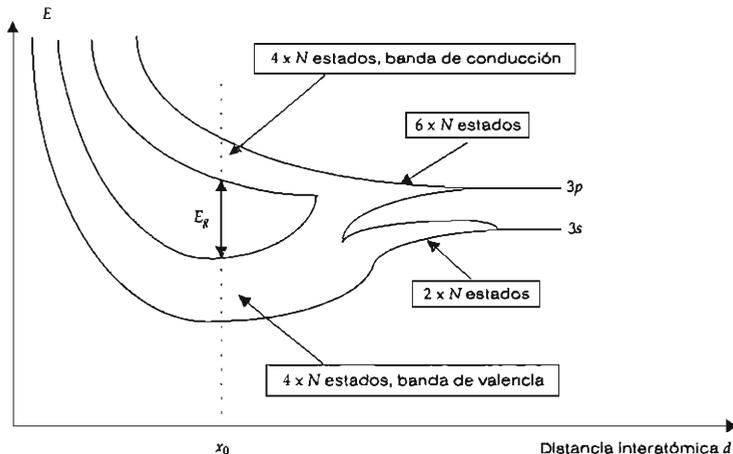


Figura II.1. Diagrama de transición de un material (silicio). Cuando la distancia interatómica es muy grande el material es un gas de poca densidad con diagramas atómicos independientes iguales. Cuando la distancia interatómica es x_0 el sólido está en equilibrio y los niveles $3s$ y $3p$ se hibridizan, dando lugar a las bandas de energía (sp^3).

(la separación entre átomo y átomo, d , es muy grande) a un sólido (la separación atómica $d = x_0$, es pequeña).

En el diagrama se señala el punto en que la distancia interatómica es x_0 , la posición atómica de equilibrio para el sólido de silicio a la temperatura y presión ambientales. El sistema es más estable cuando la distancia interatómica es x_0 (es decir, cuando tiene una energía total mínima). Para modificar esta posición habrá que agregar energía al sistema o extraérsela, en forma de calor o de presión.

Las ideas esenciales de las que se parte para construir el diagrama anterior son más o menos simples. Imagine que tenemos tres átomos idénticos aislados uno del otro (d es muy grande). Los diagramas de energía de estos tres átomos serían idénticos. Sin pérdida de generalidad podríamos suponer que estos espectros surgen del estudio de una partícula en una caja de potencial infinito.

Al disminuir la separación entre los átomos (d tiende a x_0), los ponemos a interactuar. El principio de Pauli establece que dos niveles de un sistema (sea éste un átomo o 10^{23} átomos) no pueden tener la misma energía, así que un nivel atómico se "desdobra" en tantos niveles como átomos formen al sólido.

Esto explica parcialmente el diagrama de la figura II.1. Algunos detalles no están explicados, como por ejemplo, las condiciones requeridas para la formación del sistema más estable con energía total mínima. Los detalles se pueden aclarar consultando las referencias.

El modelo de Kronig-Penney estudia la dinámica de un electrón en un potencial periódico. Éste será nuestro primer intento formal por conocer las características de los electrones en un sólido; hasta ahora conocemos las características de los electrones libres y ligados a un átomo aislado, pero no sabemos qué pasa con ellos al "meterlos" en un sólido.

Un electrón libre se describe por la siguiente función de onda:

$$\psi(x) = Ae^{ikx} + Be^{-ikx} \quad (II.1)$$

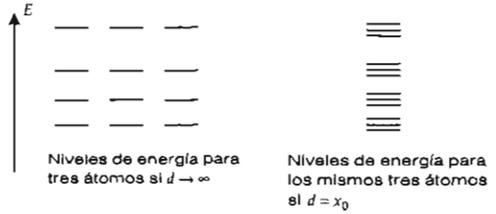


Figura II.2. Formación de las "bandas" de energía para un sólido de tres átomos.

Tiene una energía $E = \frac{\hbar^2 k^2}{2m}$. Lo que se representa en un diagrama de energía contra cantidad de movimiento (E versus k), como una parábola.

Al confinar a la partícula dentro de un sólido, la energía E y la cantidad de movimiento k deben satisfacer alguna regla de cuantización.

Para la solución de este problema, partamos de la hipótesis de que se cumple la ecuación de Schrödinger.

$$\bar{H}\psi = E\psi; \bar{H} = \frac{\bar{p}^2}{2m} + V(x, y, z); \bar{p} = \frac{\hbar}{i}\nabla \quad (II.2)$$

La solución implica el conocimiento de la energía potencial $V(x, y, z)$, que vería un electrón al moverse dentro del sólido, además de las condiciones de la frontera necesarias para resolver la ecuación diferencial.

Las simplificaciones que se emplearán para resolver el problema son:

Consideraremos una red perfectamente periódica unidimensional, sin interrupciones ni fronteras y con periodicidad a . (Los cristales perfectos sólo pueden obtenerse en los libros de texto.)

Se estudiará lo que pasa con un solo electrón en el potencial producido por todos los iones y por el resto de los electrones que no se tengan en cuenta (un electrón en un potencial periódico).

Existe una función simple que mantiene la calidad de periodicidad y todas las propiedades esenciales que hemos comentado, la cual se conoce como potencial de Kronig-Penney.

La periodicidad del potencial se traduce en periodicidad de la función de onda, por medio del teorema de Bloch.

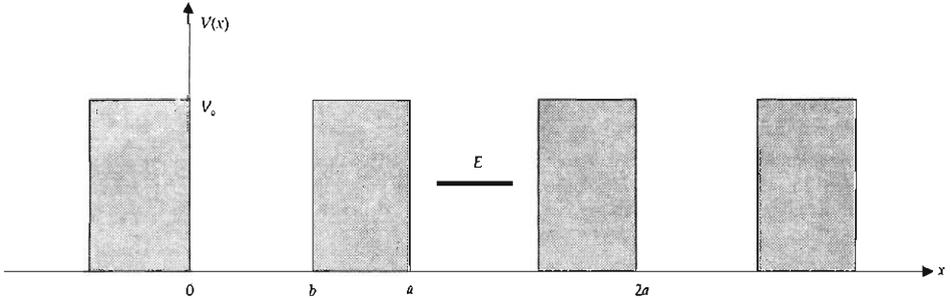


Figura II.3. Potencial de Kronig-Penney.

$$\psi(x) = e^{ikx}u(x); \text{ donde } u(x) = u(x + a)$$

Si ψ se conoce sobre cualquier celda de la red periódica (más generalmente, sobre cualquier región de longitud a), entonces se pueden conocer los valores de ψ en todas las demás celdas.

$$\psi(x) = e^{ikx}u(x)$$

$$\psi(x + a) = e^{-ik(x+a)}u(x+a) = e^{ikx}\psi(x)$$

Para conocer la distribución de los estados de energía de los electrones se aplican las condiciones de periodicidad continua. La manera de eliminar los problemas de las fronteras es suponer que nuestro cristal unidimensional forma un anillo, de radio muy grande, de tal manera que en la frontera se cumpla que:

$$\psi(x) = \psi(x + Na)$$

Donde N es el número de iones que forman el anillo y es del orden del número de Avogadro.

Aplicando la condición anterior, se tiene que:

$$1 = \cos(kNa) + i \operatorname{sen}(kNa)$$

Esta condición se satisface únicamente si $k = \frac{2\pi n}{Na}$ con $n = 1, 2, 3, \dots$ Ésta es la condición para que el sistema exista e indica que la distancia entre dos estados de energía es muy pequeña.

Para conocer la energía (E versus k) que puede tener el electrón en el sistema, debemos resolver la ecuación de Schrödinger con el potencial de Kronig-Penney, en las condiciones de frontera adecuadas. La solución de este problema es simple aunque un poco laboriosa; para las personas interesadas en resolver los detalles, se presentan a continuación algunos resultados parciales que les ayudarán en su desarrollo.

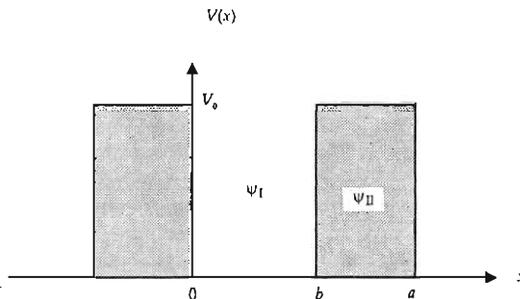


Figura II.4. Condiciones a la frontera para el potencial de Kronig-Penney.

Respetando las condiciones de frontera en b , para la función de onda ψ y su derivada ψ' , se tiene continuidad desde cero hasta a .

Para tener continuidad entre una celda y otra necesitamos que se cumpla la continuidad en la función $u(x)$ y sus derivadas.

$$\begin{aligned} \psi_I(b) &= \psi_{II}(b) \\ \psi'_I(b) &= \psi'_{II}(b) \\ u_-(0) &= u_+(a) \\ u'_-(0) &= u'_+(a) \end{aligned}$$

Explícitamente, la función de onda ψ tiene la forma siguiente:

$$\psi_I(x) = Ae^{ik_1x} + Be^{-ik_1x} \quad (0 < x < b); \quad E = \frac{\hbar^2 k_1^2}{2m}$$

$$\psi_{II}(x) = Ce^{ik_2x} + De^{-ik_2x} \quad (b < x < a); \quad E - V_0 = \frac{\hbar^2 k_2^2}{2m}; \quad E > V_0$$

$$\psi_{III}(x) = Ee^{ik_3x} + Fe^{-ik_3x} \quad (b < x < a); \quad V_0 - E = \frac{\hbar^2 k_3^2}{2m}; \quad E < V_0$$

La solución simultánea de las cuatro ecuaciones donde se han sustituido las condiciones de frontera produce un sistema de cuatro ecuaciones homogéneas. Un sistema de ecuaciones homogéneas tiene solución no trivial para A, B, C y D , sólo si el determinante de la matriz es cero.

El sistema de ecuaciones es el siguiente:

$$\begin{pmatrix} 1 & 1 & -e^{i(k_2-b)} & -e^{-i(k_2-b)} \\ k_1 - k & -k_1 - k & -(k_2 - k)e^{i(k_2-b)} & (k_2 - k)e^{-i(k_2-b)} \\ e^{ik_1a} & e^{-ik_1a} & -e^{ik_2a} & -e^{-ik_2a} \\ k_1 e^{ik_1a} & -k_1 e^{-ik_1a} & -k_2 e^{ik_2a} & k_2 e^{-ik_2a} \end{pmatrix} \begin{pmatrix} A \\ B \\ C \\ D \end{pmatrix} = 0 \quad (II.3)$$

Con las siguientes ecuaciones como solución:

$$\cos(k_1 b) \cos(k_2(b-a)) - \frac{k_1^2 + k_2^2}{2k_1 k_2} \sin(k_1 b)$$

$$\sin(k_2(b-a)) = \cos(ka) \quad (II.4)$$

donde:

$$k_1^2 - k_2^2 = \frac{2mV_0}{\hbar^2}$$

y

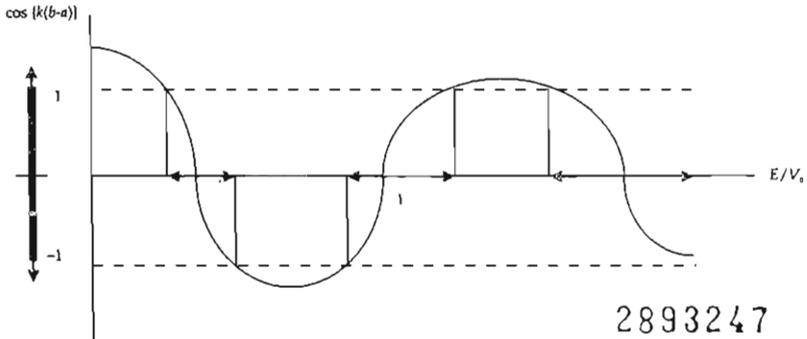
$$\cos(k_1 b) \cosh(k_3(b-a)) - \frac{k_1^2 - k_3^2}{2k_1 k_3} \sin(k_1 b)$$

$$\sinh(k_3(b-a)) = \cos(ka) \quad (II.5)$$

donde:

$$k_3^2 - k_1^2 = \frac{2mV_0}{\hbar^2}$$

Para encontrar las soluciones de las ecuaciones trascendentes anteriores que satisfacen E y k se necesitan las gráficas de las partes izquierda y derecha en un mismo sistema coordenado. Los valores de energía E que hagan válida la igual-



*Figura II.5. Solución gráfica del sistema de ecuaciones de Kronig-Penney.

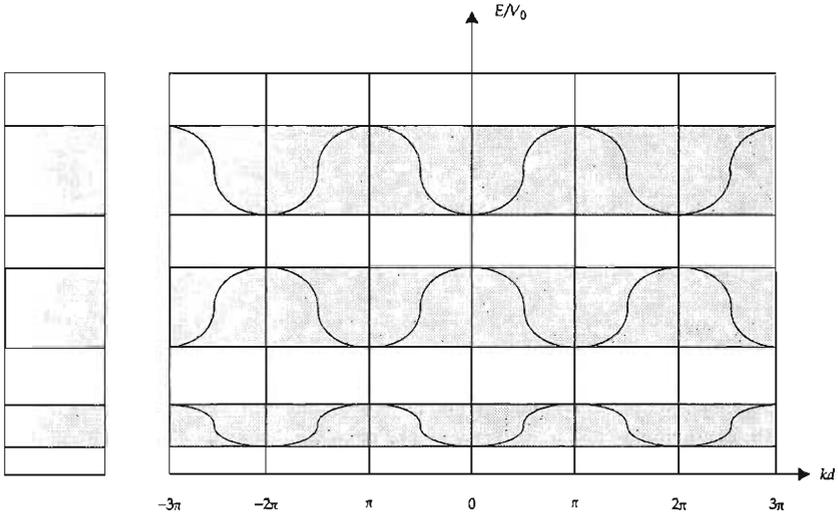


Figura 11.6. Estructura de bandas en un sólido.

dad serán los únicos con los que el electrón se podrá mover en el potencial de Kronig-Penney.

A partir de la figura 11.5 es posible obtener la relación entre la energía de los electrones E y su vector de onda k (recuerde que es proporcional a la cantidad de movimiento).

De esta manera, no todos los valores de energía corresponden a estados permitidos del sistema; existen regiones de energía prohibida para el electrón. Al realizar la gráfica de E versus k tenemos que dentro de una banda permisible la relación es oscilatoria y no parabólica, como la relación para el electrón libre. El ancho de la banda aumenta con el incremento de la energía. En el extremo superior de la banda permitida la relación de E versus k es cóncava hacia abajo, mientras que en el mínimo de la banda la relación es cóncava hacia arriba. Dentro de una banda permitida existe un gran número de estados que pueden ser ocupados por los electrones.

Veamos ahora cómo se ocuparían las bandas de energía cuya existencia hemos predicho.

Si como ya se mencionó, el principio de Pauli se cumple, y si además se supone que el sólido está en su estado base, el cual se obtiene al tener un mínimo en la energía total del sistema, esto es, teniendo al material en el cero absoluto de temperatura (0 K), entonces, como todos los elec-

trones no pueden estar en un mismo nivel de energía (a cada nivel de energía se le asocia un solo número de onda del electrón en el cristal, k). Además, no existe un continuo de estados permitidos en el sólido (k), sino que es un número discreto. Con todo lo anterior, se concluye que los electrones ocupan parcialmente los estados disponibles, el estado de mayor energía ocupado con un electrón en el estado base se llama "k de Fermi" y la energía asociada con ese estado se llama energía de Fermi (E_f).

Volviendo a las teorías anteriores, vemos que cada banda de energía es resultado del "amontonamiento" de niveles iguales de los átomos; este amontonamiento formará la densidad de estados en la banda (niveles por unidad de volu-

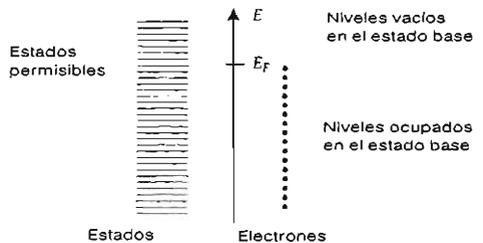


Figura 11.7. Energía de Fermi, E_f .

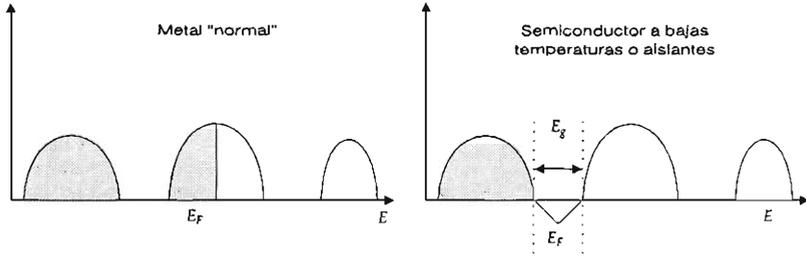


Figura II.8. Bandas de energía en los materiales.

men y por unidad de energía), que se discutirá en la sección de semiconductores. En este momento, lo más importante es encontrar una manera de agrupar y caracterizar los materiales de acuerdo con las diferencias en las bandas de energía.

Se sabe que los electrones no se acomodan en las bandas de manera arbitraria (de esto hablaremos después): al acomodarse los electrones en las bandas del sólido, primero se ocupan los niveles de menor energía y después los de mayor energía (al menos en el estado base). Pero como hay más lugares que electrones, no se ocupan todos los lugares disponibles.

En función de sus bandas de energía y de la manera en que se ocupan las bandas permitidas por los electrones, los materiales se clasifican en metales y aislantes.

En el segundo caso, se ocupó por completo una banda y la inmediata superior está completamente vacía. En estos casos la definición precisa del nivel de Fermi estará determinada por la facilidad para mover un electrón de la banda llena a la vacía, como veremos al estudiar los semiconductores.

La separación entre la banda ocupada (que en adelante llamaremos de valencia) y la banda desocupada (que llamaremos de conducción) será denominada E_g . Y dependiendo de este valor y de las condiciones de uso del material, el material se comporta como aislante o como semiconductor.

Una banda llena de electrones no conduce corriente, ya que no existen estados continuos vacíos que puedan permitir el movimiento de los electrones. Si un electrón quiere cambiar de lugar dentro de la banda, lo hará sólo intercambiando posición con otro electrón.

Una banda vacía de electrones no conduce corriente, ya que no tiene portadores de carga.

Es posible establecer conducción de corriente en una banda llena de electrones quitando algunos de ellos, ya que los electrones que queden podrán moverse entre estos pocos lugares libres. De la misma manera, se puede establecer corriente eléctrica en una banda vacía si se introducen algunos electrones.

¿Es el comportamiento de los electrones en una banda casi vacía el mismo que el producido en una banda casi llena?

La respuesta es no, ya que físicamente se mueven de forma distinta, y por ello se les conoce con diferentes nombres. Los electrones que se mueven en una banda casi vacía se llaman *electrones*; especialmente se mueven en trayectorias grandes y por lo general están asociados con una energía mayor. Los electrones que se mueven en una banda casi llena se llaman *huecos*; especialmente se mueven en trayectorias muy cortas y están asociados con una energía menor. El nombre de hueco no se refiere a una partícula sino a un estado dinámico de los electrones.

Empleando la estructura de bandas que se obtuvo del modelo de Kronig-Penney, es posible describir con más detalle los huecos y los electrones en un material. La relación que controla el movimiento de las partículas en un sólido (E versus k) ya no es una parábola como en los electrones libres. En los extremos de las bandas la función oscilatoria se puede aproximar a una parábola.

$$E = \frac{\hbar^2 k^2}{2m^*} \quad (II.6)$$

Para ajustarse a las características de las bandas, la masa no puede ser una constante, y se llama masa efectiva m^* . En el caso de los huecos, que se mueven en la parte alta de la banda, la masa efectiva debe ser un número negativo, ya que la relación dinámica (E versus k) es cóncava hacia abajo. Para continuar usando masas positivas, se redefine el electrón que se mueve en la parte alta de la banda de valencia como una partícula con masa positiva pero que tiene carga positiva, ya que lo que importa en el transporte es la relación carga-masa.

En la mayoría de los materiales, la masa efectiva m^* se puede considerar como una constante diferente de la masa del electrón libre. La masa del electrón libre es $m = 0.91095 \times 10^{-30}$ kg.

Al "meter" un electrón en un sólido, el electrón es el mismo, pero su movimiento es diferente del de un electrón en el espacio libre. El electrón dentro de un sólido está expuesto a la interacción con los núcleos y con los demás electrones; podría ser menor si los campos eléctricos que forman el sólido facilitan su movimiento, pues parecería que se hubiera disminuido la masa del electrón, cosa que evidentemente no sucede. En general, $m^* \neq m$. Si se piensa en ambas como la masa de un electrón, m^* corresponde a un electrón en un sólido y m corresponde a un electrón libre. Esta diferencia sólo es válida cuando intentamos moverlos. Es como intentar mover un objeto arrastrándolo en una superficie de hielo o de tierra; parece mucho más pesado en el segundo caso.

En la descripción anterior se considera que no todos los electrones del sistema son igualmente activos para definir las propiedades de los materiales; en efecto, sólo una pequeña región de energía, en la cual se da la transición entre estados vacíos y ocupados, definirá a los electrones que determinarán las propiedades de los materiales.

Demos fin a esta sección con una analogía que permite ilustrar la diferencia entre huecos y electrones. Supongamos que existe un cine como el de la figura II.9.

Las personas de la zona A no pasan a la zona B, ni viceversa.

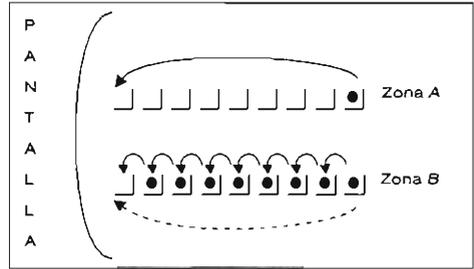


Figura II.9. Cine "con bandas".

Al iniciar la película, en la zona A sólo hay una persona al final de la fila y en la zona B la fila está completamente ocupada, excepto el lugar más cercano a la pantalla.

En ese momento las personas se dan cuenta de que es mejor acercarse a la pantalla, entonces el cinéfilo de la zona A avanza siguiendo la trayectoria indicada por la flecha continua, mientras que en la zona B los cinéfilos son muy ordenados: cada uno avanza un asiento.

Después de los movimientos y si no se puede distinguir a un cinéfilo de otro, la suma de cada uno de los movimientos se puede representar con la flecha discontinua. El movimiento en la zona B es similar al de la zona A, aunque sucedió de diferente manera. La zona A representa a los electrones en la banda de conducción, la zona B representa a los electrones en la banda de valencia, es decir, a los huecos.

METALES

Los materiales metálicos comparten varias características comunes: El diagrama de bandas de energía es similar, el nivel de Fermi está dentro de la banda de conducción. El enlace químico es similar, los electrones libres se comparten en el material formando una red de iones y una nube electrónica. Las propiedades ópticas y eléctricas son parecidas, son buenos conductores del calor y la electricidad, y la penetración de las ondas electromagnéticas es mínima. En esta sección se presentarán varios conceptos macroscópicos, como la ley de Ohm y la de Joule desde un punto

de vista microscópico. Todos los conceptos que se verán en esta sección pueden extenderse a los otros materiales.

Ley de Ohm y de Joule

Una de las relaciones más conocidas en el mundo es la ley de Ohm, que en su forma más simple muestra la proporcionalidad entre la densidad de corriente (\mathbf{J}) y la intensidad del campo eléctrico (\mathbf{E}), a través de la conductividad (σ).

$$\mathbf{J}_i = \sigma_i \mathbf{E}_i \quad (\text{II.7})$$

Esta relación únicamente es válida para materiales homogéneos con condiciones uniformes de temperatura, sin iluminación y con campos eléctricos y magnéticos sin cambio en el tiempo y pequeños.

No todos los materiales o dispositivos son igualmente sensibles a todas las restricciones mencionadas. Un alambre metálico en forma de bobina no cumple esta relación si hay cambios en la corriente eléctrica. El diodo tampoco cumple esta relación, ya que es un dispositivo inhomogéneo, etcétera.

Esta relación, dependiente de la geometría, toma la forma:

$$V = RI \quad (\text{II.8})$$

donde V es la diferencia de potencial, R es la resistencia eléctrica e I es la corriente eléctrica. Las ecuaciones anteriores se pueden relacionar usando una geometría simple, como la de un alambre de longitud L y sección transversal de área uniforme A . Utilizando la definición de la densidad de corriente $\int_A \mathbf{J} d\mathbf{a} = I$ y del campo eléctrico, $\nabla V = -\mathbf{E}$.

$$R = \rho \frac{L}{A}$$

ρ es la resistividad [$\Omega \text{ cm}$] que es el inverso de la conductividad σ . La dependencia de la resisti-

vidad en relación con la estructura microscópica puede obtenerse con el modelo de Drude.

Podemos imaginar que el material está formado por un océano de electrones en condiciones casi de electrones libres. Esta nube de electrones evita que los iones formados de núcleo y los electrones no libres se repelan y permanezcan como un sistema. Este océano de electrones se encuentra en movimiento caótico debido a la energía proporcionada por la temperatura, pero no produce una corriente eléctrica neta porque no es direccional.

Al aplicar un campo eléctrico constante, éste ejerce una fuerza sobre los electrones que tiende a alinear su movimiento, llamada fuerza de Lorentz.

$$\mathbf{F} = q(\mathbf{E} + \mathbf{v} \times \mathbf{B}) \quad (\text{II.9})$$

Esta fuerza constante produce una aceleración constante sobre el electrón, como lo predice la segunda ley de Newton.

$$\mathbf{F} = m\mathbf{a}$$

Si la fuerza del campo eléctrico fuera la única que existe sobre los electrones, entonces se originarían corrientes arbitrariamente grandes, ya que la velocidad aumentaría linealmente. Esto se desprende de la definición de densidad de corriente eléctrica, donde n es la densidad de portadores de carga [cm^{-3}], q es la carga del portador, por ejemplo, electrones, y $\langle \mathbf{v} \rangle$ es la velocidad vectorial promedio de los electrones [cm/s].

$$\mathbf{J} = nq\langle \mathbf{v} \rangle \quad (\text{II.10})$$

Sin embargo, debe existir un factor que limite la velocidad de los portadores y, por lo tanto, la corriente. Este factor surge naturalmente cuando consideramos el efecto de la interacción de los electrones y los iones fijos, o sea, si suponemos que un electrón típico adquiere una velocidad promedio $\langle \mathbf{v} \rangle$ en un tiempo promedio τ (tiempo de relajación). Los electrones chocan con un ion y transfieren su energía de exceso a la red de iones.

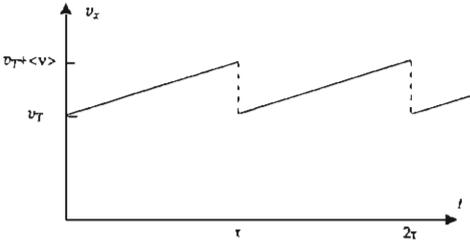


Figura II.10. Modelo de Drude de transferencia de energía por colisiones.

Aplicando este modelo y las leyes anteriores obtenemos una expresión para la velocidad promedio $\langle v \rangle$ que alcanzan los electrones en el sólido. Esto es válido cuando imaginamos un conjunto muy grande de electrones en movimiento.

$$m^* \int_{v_T}^{v_T + \langle v \rangle} dv_x = qE_x \int_0^{\tau} dt \quad (II.11)$$

de donde se obtiene una expresión propia para la velocidad promedio producida por el campo eléctrico:

$$\langle v \rangle = \frac{qE_x \tau}{m^*} \quad (II.12)$$

Al sustituir en la densidad de corriente y comparar con la ley de Ohm se obtiene:

$$\sigma = \frac{nq^2 \tau}{m^*} \quad (II.13)$$

Las colisiones entre los electrones y la red son favorecidas por:

1. Vibraciones térmicas: a mayor temperatura, mayor resistividad.
2. Defectos puntuales: a mayor número de ellos, mayor resistividad:
 - a) Vacantes: ausencia de átomos producidos a altas temperaturas.
 - b) Intersticiales: átomos que están en posiciones incorrectas en la red.
 - c) Impurezas aisladas (cargadas o neutras).
3. Defectos lineales: dislocaciones, grupos de átomos desplazados de su posición de equilibrio.

4. Defectos superficiales.

- a) Superficies externas de los sólidos.
- b) Superficies internas: límite del grano, etcétera.

Definamos otra relación de interés, quizá más importante: la movilidad μ . La movilidad es la relación entre la velocidad media que adquiere un portador en presencia de un campo eléctrico:

$$\mu = \frac{\langle v \rangle}{E} \quad (II.14)$$

Con esta expresión, la conductividad σ se escribiría de la siguiente manera:

$$\sigma = nq\mu \quad (II.15)$$

Cuando pasa corriente por un alambre aumenta su temperatura. La energía térmica que así se obtiene es producida indirectamente por la fuente del campo, a través de los electrones y la red.

De la hipótesis de Drude, que supone que el electrón disminuye su velocidad de $v_T + \langle v \rangle$ a v_T a través de colisiones con la red, se espera que la red se caliente. Este aumento de temperatura es el efecto Joule.

La energía que transfiere un electrón a la red por una colisión es igual a la energía cinética perdida:

$$\omega_e = \frac{1}{2} m^* \langle v \rangle^2 = \frac{q^2 E_x^2 \tau^2}{2m^*}$$

La energía que transfieren todos los electrones a la red es por unidad de volumen y por unidad de tiempo, es la potencia total por unidad de volumen P :

$$P = \frac{2\omega_e n}{\tau} = \frac{q^2 \tau n E_x^2}{m^*} = \sigma E_x^2 \quad (II.16)$$

Para una geometría simple como la empleada anteriormente, se obtiene la forma habitual de la ley de Joule ($P = I V$).

Efecto Hall

$$E_y = \frac{-B_x J_z}{nq} \quad (II.17)$$

Las mediciones de conductividad son insuficientes para conocer el número de portadores por unidad de volumen (n), la movilidad de los portadores (μ) o para conocer la carga del portador que se mueve. Las mediciones de Hall y las de conductividad permiten conocer estas variables.

En este experimento se tiene presente un campo eléctrico que produce una corriente en el material que quiere medirse. Adicionalmente, se aplica un campo magnético transversal a la corriente, como se ve en la siguiente figura:

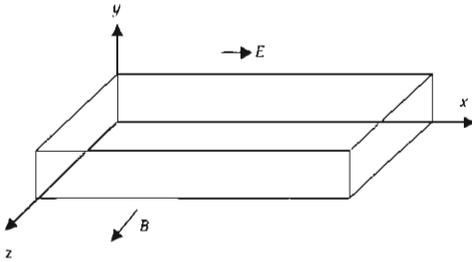


Figura II.11. Geometría para la medición del voltaje de Hall.

Los electrones que se desplazan a través de la muestra por el campo eléctrico E , sienten la presencia del campo magnético B . En la geometría de la figura anterior, la fuerza tendría la siguiente expresión:

$$F = q\mathbf{v} \times \mathbf{B} = q \begin{vmatrix} \hat{i} & \hat{j} & \hat{k} \\ v_x & v_y & v_z \\ 0 & 0 & B_z \end{vmatrix} = q(v_y B_z \hat{i} - v_x B_z \hat{j})$$

La componente de la fuerza que producirá el voltaje de Hall es la que mueve los electrones en la dirección de la "y".

$$F = -qv_x B_z = qE_y$$

Con la expresión misma de la conductividad eléctrica, el campo eléctrico de Hall toma la siguiente forma:

Midiendo el campo magnético aplicado (B_x), la densidad de corriente que pasa por el material (J_z) y el campo eléctrico de Hall (E_y), es posible obtener la densidad de los portadores de carga (n). Por medio del signo del voltaje de Hall podemos conocer qué se mueve, si una partícula positiva o una negativa. Midiendo la conductividad es posible calcular la movilidad.

La anterior descripción puede ilustrarse con las siguientes afirmaciones. La formación de un voltaje de Hall es diferente para cargas de diferente signo. Supongamos que aplicamos un potencial positivo en $x=0$ y uno negativo en $x=L$:

- Si la corriente eléctrica está formada por electrones, el campo eléctrico de la batería tenderá a moverlos hacia $x=0$, mientras que el campo magnético los mueve hacia $y=0$.
- Si la corriente eléctrica está formada por cargas positivas, el campo eléctrico de la batería tenderá a moverlos hacia $x=L$, mientras que el campo magnético los mueve hacia $y=0$.

Estas cargas acumuladas en ambos casos inducen cargas en la otra superficie; esto es, el campo magnético produce una diferencia de potencial a través del material y únicamente con medir su polaridad sabremos si se mueve una carga positiva o negativa.

Para finalizar esta sección, se puede mencionar un estado de conducción poco usual llamado estado superconductor. La mayoría de los materiales presentan a temperatura ambiente un cierto grado de resistencia al paso de la corriente eléctrica. En unos cuantos materiales, y usualmente a temperaturas bajas, se presenta una pérdida total de la resistencia eléctrica. Es altamente deseable contar con este tipo de materiales, pero en la actualidad no se cuenta con materiales que sean buenos candidatos para aplicaciones comunes. La investigación en estos materiales es muy activa en la actualidad.

Algunas de las propiedades de estos materia-

les son las siguientes: La resistividad eléctrica cae a cero a la temperatura T_c , y no se disipa calor. Como tienen resistencia cero, es posible obtener solenoides con campo magnético elevado y gran densidad de corriente. Con un solenoide de hilo de ciertas aleaciones superconductoras pueden obtenerse campos magnéticos superiores a 10 webers/m² sin disipar potencia, en contraste con los solenoides convencionales de Cu o Ag, que operan a temperatura ambiente y que disipan una cantidad enorme de calor al producir un campo de 10 webers/m². Esto significaría, además, obtener un gran ahorro en las líneas de transmisión de potencia eléctrica.

Los superconductores son materiales diamagnéticos perfectos (véase al respecto la sección Propiedades ópticas y magnéticas), pues no son penetrados por los campos magnéticos. El elevado diamagnetismo de los superconductores ofrece la posibilidad de soportar cargas mediante un campo magnético. De este modo se pueden conseguir cojinetes sin fricción para equipos rotatorios.

La superconductividad puede eliminarse a temperaturas menores de T_c , por medio de un campo magnético intenso.

ASLANTES

Los aislantes usualmente se caracterizan por átomos unidos con enlaces iónicos, como la sal de mesa, o por enlace covalente, como los polímeros o el óxido de silicio. Sus propiedades mecánicas se caracterizan por una gran fragilidad y gran resistencia a la compresión. Usualmente son malos conductores de la electricidad excepto a muy altas temperaturas o en solución, donde los materiales iónicos se disocian dando lugar a electrolitos. También son, por lo general, malos conductores del calor, excepción hecha de ciertos aislantes que conducen el calor sin emplear electrones sino vibraciones atómicas, como el diamante (mejor conductor térmico que los metales a ciertas temperaturas) o algunos óxidos metálicos.

El estudio de los sólidos iónicos no tiene gran importancia en este curso, pero nos permite un

buen acceso a las estadísticas de procesos colectivos, o mejor dicho, al estudio de los sólidos por medio de funciones de distribución estadística.

Estadística de Boltzmann

Para calcular la forma de la curva de conductividad *versus* temperatura de un aislante, es necesario entender cómo se realiza la transferencia de carga eléctrica a través del material.

La corriente eléctrica es el movimiento colectivo de portadores de carga; en un metal usualmente se trata de los electrones, aunque en algunos metales lo realizan las partículas positivas. ¿Qué portador de carga eléctrica se mueve en un material iónico?

Recordemos por un momento las soluciones de electrolitos (sal en agua), un tipo especial de material iónico. De hecho, la única diferencia entre un sólido iónico y un electrolito es que en el electrolito los iones no guardan distancias definidas y tienen una mayor movilidad. Si introducimos en esta solución dos electrodos conectados a una batería, fluirá una corriente que estará determinada por el movimiento de los iones en la solución (incluso éste es un método para obtener algunos materiales; en nuestro ejemplo, cloro e hidróxido de sodio). Así que, en la mayoría de los aislantes, se mueven los iones y no los electrones. En casos muy especiales podrían moverse los electrones, pero no hablaremos aquí de ellos.

Para responder a la pregunta de cómo varía la conductividad σ *versus* la temperatura T , debemos reconocer la variación de cada uno de los términos que definen la conductividad. Como sabemos, la conductividad está determinada por $\sigma = NQ\mu$, donde N será la densidad de iones que participen en la corriente eléctrica, Q la carga de los iones, y μ la movilidad de los iones. Si consideramos a N y Q como constantes (esto es una buena aproximación), la única fuente de variación estará en la movilidad μ .

¿Es posible que un ion se mueva en un cristal? ¿Se puede hablar de movilidad?

La respuesta es sí; aunque, por supuesto, son magnitudes muy diferentes de las que se obtie-

nen para los electrones en un metal, que son usualmente de muchos órdenes de magnitud menores.

La movilidad en los sólidos iónicos dependerá de la facilidad de cambiar de posición: si todas las posiciones de la red cristalina del material están ocupadas, como sucede a muy baja temperatura, muy difícilmente se moverá algún ion. La conductividad tiende a cero a medida que la temperatura tiende a cero.

La movilidad aumentará cuando se aumenten los espacios, que se llaman vacantes. Al responder a la pregunta de cómo varía la cantidad de vacantes con la temperatura, sabremos cómo cambia la conductividad con la temperatura.

¿Cómo se forma una vacante?

En la figura II.12 aparece el proceso de movimiento de un átomo desde el interior de una red cristalina hacia la superficie. La posición final de la red sin un ion es una vacante.

Para calcular la energía de formación de una vacante, veamos cuánta energía tiene el sistema antes y después de su formación.

Para formar la vacante se rompen cuatro enlaces y se reconstruyen solamente dos enlaces, así que hay dos enlaces rotos de diferencia. Si suponemos que la energía de enlace es del orden de 0.5 eV, tendremos que es necesario proporcionar aproximadamente 1 eV para formar una vacante (aquí se calculó para dos dimensiones; para tres dimensiones el procedimiento es parecido).

¿Cuándo tendrá el átomo energía suficiente para formar una vacante?

¿Cómo obtiene energía el átomo?

En este caso nos interesan los cambios de conductividad producidos por la temperatura. La energía del átomo está dada por la temperatura, y aproximadamente cada átomo tendrá una energía promedio igual a kT , donde k es la constante de Boltzmann ($k = 8.6174 \times 10^{-5}$ eV/K), y T representa la temperatura medida en Kelvin ($T[\text{K}] = 273.15 + T[^\circ\text{C}]$).

Entonces, ¿aproximadamente a 15 000 K se tendría la energía necesaria para formar una vacante?

Esto sería verdad si cada átomo fuera independiente de los demás, pero en realidad no es

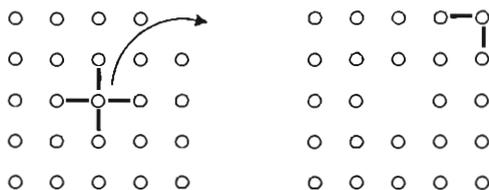


Figura II.12. Formación de una vacante.

así: los átomos se comportan como un sistema acoplado. La figura II.13 presenta la diferencia en la distribución de la energía cuando tenemos átomos aislados y cuando están acoplados.

Así, no es necesario que cada átomo tenga la energía necesaria para formar una vacante. Aun cuando todos los átomos tienen una energía promedio mucho menor, algunos átomos, tomando energía de los demás átomos que forman el sistema, pueden llegar a tener la energía suficiente para formar una vacante.

La ecuación de Boltzmann presenta esta información. Por ser una ecuación básica de todos los procesos activados térmicamente, tiene diferentes nombres en diversas áreas de estudio. Para nosotros será la expresión para la estadística de Boltzmann.

$$N = N_0 \exp\left(-\frac{E_0}{kT}\right) \quad (\text{II.18})$$

N_0 es el número total de vacantes que puede tener un sistema por unidad de volumen, y es del orden de 10^{23} cm^{-3} . E_0 es la energía necesaria para formar una vacante, mientras que kT es la energía promedio que cada átomo tiene, proporcionada por la temperatura. Finalmente, N es el número de vacantes que tiene el sistema por unidad de volumen.

En conclusión, podemos decir que si bien no conocemos los detalles de la curva de conductividad como función de la temperatura en los materiales aislantes, sí podemos decir que su conductividad aumenta exponencialmente con la temperatura.

De manera similar, la estadística de Boltzmann describe la distribución de las partículas en la escala de energía, a todas ellas se les

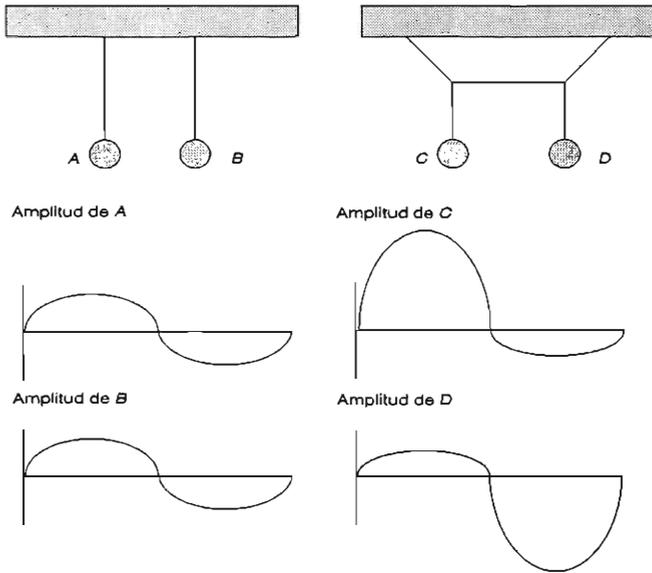


Figura II.13. Un sistema aislado está formado por dos péndulos independientes, la energía total promedio de cada oscilador es la misma. En un sistema acoplado, como el formado por los péndulos compuestos, la energía total del sistema se conserva, pero la energía de cada partícula puede ser mayor o menor que la promedio.

proporciona una energía kT en promedio. De aquí se obtiene que muchas partículas tienen menos energía que kT y unas cuantas tienen ener-

gías mayores que kT . N partículas de N_0 tienen una energía E_0 si en promedio a todas ellas se les da kT .

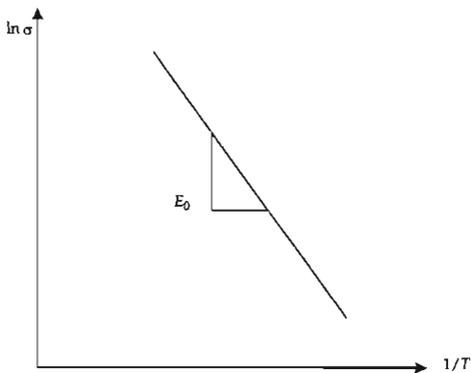


Figura II.14. Variación de la conductividad como función de la temperatura en un material aislante. Estadística de Boltzmann.

SEMICONDUCTORES

Son los materiales que al cero absoluto de temperatura tienen la banda de conducción vacía de electrones y la banda de valencia llena de electrones. Para las aplicaciones en electrónica a temperatura ambiente, el ancho de energía prohibido E_g es aproximadamente de 1.5 eV.

Algunas de las propiedades de los semiconductores son las siguientes:

- Se puede escoger el tipo de carga que transporta la corriente eléctrica, electrones o huecos.
- Se puede modular la concentración de huecos y electrones de tal manera que se puedan obtener propiedades de metal o de aislante, $10^{-7} \leq \rho \leq 10^8 \Omega\text{cm}$.

- Acepta perturbaciones temporales y locales en la concentración de portadores de carga libre (fotorresistencias) y además es posible controlar el tiempo de la perturbación (de fotodiodo a memorias).

El semiconductor que se ha ganado la supremacía en la electrónica moderna de muy alto grado de integración VLSI es el silicio. Algunas propiedades notables que lo han hecho merecedor de esta distinción son:

- Pueden practicarse en él una gran variedad de procesos tecnológicos sin el problema de la descomposición, como sucede con otros semiconductores como el compuesto GaAs (arseniuro de galio).
- Es uno de los semiconductores mejor conocido en cuanto a la preparación y propiedades de los monocristales puros.
- Como su ancho de banda prohibida, E_g , es mayor que el del germanio, pueden fabricarse microcircuitos que trabajan hasta 175°C, temperatura aceptable incluso para aplicaciones militares.
- Su óxido SiO_2 le proporciona una serie de ventajas adicionales:

a) Propiedades pasivadoras: protección en los costados de las uniones.

b) Propiedades dieléctricas: las características conductoras entre el Si puro y su óxido son muy diferentes, ya que éste es un aislante bastante bueno con el que se pueden crear dispositivos con dieléctricos muy fácilmente, como los dispositivos MOS.

c) Facilidad tecnológica: las propiedades intrínsecas del óxido son muy diferentes de las del silicio puro y permiten emplearlo en el proceso fotolitográfico (obtención selectiva de regiones con diferentes cantidades de impurezas), propiedad que desplaza definitivamente al germanio (su óxido tiene muy malas propiedades mecánicas, pues es un polvo sin adherencia).

El silicio, sin embargo, no es óptimo en otros aspectos:

- La movilidad de los portadores es mayor en GaAs y otros compuestos (III-V), lo que permite fabricar dispositivos con respuestas a mayores frecuencias.
- El silicio tiene una transición indirecta entre

bandas (véase más adelante), lo que limita sus aplicaciones electroópticas.

- Algunos dispositivos no se pueden hacer de silicio, como el LED (diodo emisor de luz) y el láser.

El material mejor preparado para suplir al silicio en sus deficiencias es el GaAs que, tecnológicamente hablando, tiene un nivel de desarrollo muy alto. Esto ha permitido que comercialmente se encuentren dispositivos más rápidos que los tradicionales de alta velocidad de silicio, como SGFET (Schottky-Gate-Field-Effect-Transistor).

Estructura de bandas

Cuando se estudió el modelo de Kronig-Penney se encontró la estructura de bandas (E versus k) para un material genérico. Los materiales reales difieren de esta estructura. Hay dos casos que tienen un interés especial para nosotros: 1) La transición directa de bandas, como en el GaAs, donde el mínimo de la banda de conducción tiene el mismo valor de cantidad de movimiento que el máximo de la banda de valencia. Un electrón que pasa de la banda de conducción a la de valencia lo hace sin cambio en la cantidad de movimiento, lo que favorece la generación de fotones (partículas luminosas con cantidad de movimiento despreciable). 2) Los materiales con transición indirecta de bandas, como el silicio o germanio, donde el mínimo de la banda de conducción tiene un valor de k diferente del máximo de la banda de valencia. Una transición electrónica incluye un gran cambio en la cantidad de movimiento y favorece la producción de fonones (partícula asociada a la conducción del calor con gran cantidad de movimiento).

En la figura 11.15 se presenta la estructura de bandas para dos semiconductores típicos. Se incluyen únicamente la banda de conducción y de valencia.

Como señalamos al estudiar la teoría del enlace molecular, el ancho de banda prohibido E_g varía con la temperatura. De acuerdo con la siguiente relación empírica, la temperatura está en Kelvin y E_g en eV:

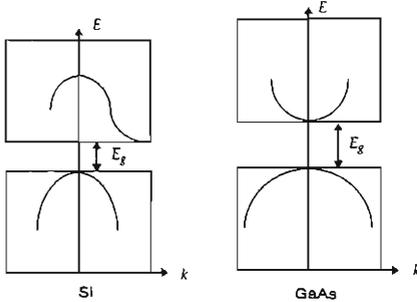


Figura 11.15. Estructura de bandas. Transición directa como en GaAs y transición indirecta como en Si.

$$E_g(T) = E_g(0) - \frac{\alpha T^2}{(T + \beta)}$$

Material	$E_g(0)$ [eV]	$\alpha \times 10^{-4}$	β	$E_g(300)$ [eV]
GaAs	1.519	5.405	204	1.424
Si	1.170	4.73	636	1.12
Ge	0.744	4.77	235	0.66

Densidad de portadores libres

Después de la descripción de los semiconductores, nos interesa entender el comportamiento de sus propiedades eléctricas. Para ello es necesario saber cuáles electrones definen las propiedades de los semiconductores y cómo se calculan éstas. Para alcanzar este objetivo es necesario conocer la densidad de los estados electrónicos $D(E)$, cómo se distribuyen los estados que pueden ocupar los electrones y la función de distribución $f(E)$, de qué manera se acomodan estadísticamente los electrones en estos estados. Al tener ambas expresiones se puede saber cuántos electrones y cuántos huecos definirán las propiedades electrónicas.

Para calcular la expresión de la densidad de estados procederemos como sigue:

Situamos los estados del electrón en un espacio de números de onda k , en donde la energía

de Fermi es el máximo nivel de energía ocupado, tomando como válida la relación de energía cinética de un electrón libre con masa efectiva m^* :

$$E = \frac{\hbar^2(k_x^2 + k_y^2 + k_z^2)}{2m^*} \tag{II.19}$$

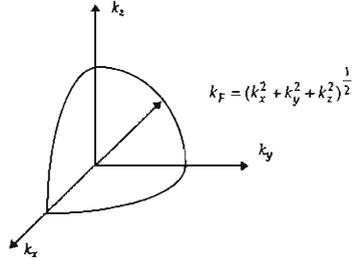


Figura 11.16. Estados ocupados por los electrones en el estado base.

Número de estados =

$$\frac{\text{Volumen de todos los estados}}{\text{Volumen de un estado}} = \frac{\frac{4\pi}{3} k_F^3}{\frac{(2\pi)^3}{L^3}}$$

donde el volumen de todos los estados es el volumen encerrado por k_F en la figura anterior, en tanto que el volumen de un estado se obtuvo en el desarrollo para que diera como resultado E 1-4, donde L es una dimensión del material con volumen L^3 .

Para estimar la densidad de estados $D(E)$, definamos ésta como el cambio en el número de estados en relación con la energía.

$$D(E) \equiv \frac{dN}{dE} = L^3 \frac{\sqrt{2}(m^*)^{\frac{3}{2}}}{\pi^2 \hbar^3} E^{\frac{1}{2}} \tag{II.20}$$

Esta ecuación muestra una dependencia con la energía que crece como la raíz cuadrada de la energía. Esta expresión es muy simple por su origen, pero es bastante útil porque predice analíticamente la distribución de los estados en las bandas. Así podemos al menos suponer que es

válida en la parte alta y baja de las bandas, donde las propiedades electrónicas de los semiconductores se definen.

La densidad de estados en la banda de conducción y de valencia tendría expresamente la siguiente forma, donde las constantes D_C y D_V serían distintas en virtud de que las masas efectivas para los huecos y para los electrones son diferentes. Además hay que incluir la información relacionada con la posición de la banda: E_C es el mínimo de la banda de conducción y E_V es el máximo de banda de valencia.

$$D(E)_{\text{electrones}} = D_C(E - E_C)^{\frac{1}{2}}; E \geq E_C$$

$$D(E)_{\text{huecos}} = D_V(E_V - E)^{\frac{1}{2}}; E \leq E_V \quad (II.21)$$

Conforme aumenta el número de partículas en un sistema, el tratamiento estadístico se vuelve obligatorio. Esto nos lleva a obtener resultados sobre el sistema en general y no sólo sobre el movimiento de cada partícula.

La función de distribución $f(E)$ da la probabilidad de que determinado número de partículas ocupen un determinado estado de energía. Existen diferentes estadísticas de distribución de acuerdo con las características de las partículas que describen. La que describe correctamente el comportamiento de los electrones es la estadística de Fermi-Dirac, que supone partículas indistinguibles e iguales además de cumplir el principio de exclusión de Pauli.

$$f_i(E) = \frac{1}{1 + \exp\left(\frac{E - E_F}{kT}\right)} \quad (II.22)$$

Esta función describe la estadística de ocupación de los estados en un sólido como función de la temperatura, en donde E_F es la energía de Fermi cuando ésta se aproxima a cero, ya que todos los niveles con energía menor a E_F tienen una probabilidad de ocupación igual a 1, $\{f(E) = 1\}$ y todos los niveles con energía mayor a E_F están vacíos,

$\{f(E) = 0\}$. La estadística de Boltzmann que se estudió anteriormente puede usarse para aproximar la función de Fermi-Dirac, pero sólo en el caso en que $E > E_F$ o $E < E_F$ para el caso $1-f(E)$.

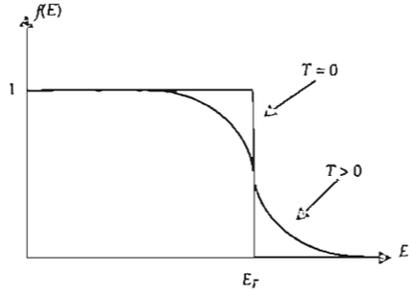


Figura II.17. Función de distribución de Fermi-Dirac.

De manera similar, la función de distribución para los huecos es:

$$f_h(E) = 1 - f_e(E) = \frac{1}{1 + \exp\left(\frac{E_F - E}{kT}\right)} \quad (II.23)$$

En el caso de temperaturas diferentes de cero, la probabilidad de que un nivel con energía δ mayor a E_F esté ocupado es igual a la probabilidad de que un estado con energía δ menor a E_F esté vacío.

Ahora que sabemos cómo se distribuyen los niveles de los electrones $D(E)$ y cómo se acomodan los electrones en los estados $f(E)$, estamos en condiciones de calcular el número de electrones y de huecos por unidad de volumen. Para realizar el cálculo de los electrones, se cuentan únicamente los que se pueden mover en la banda de conducción, es decir, los electrones que tienen una energía mayor a E_C . Para calcular la cantidad de huecos es necesario contar los electrones que se pueden mover en la banda de valencia, los electrones con energía menor a E_V . En lugar de contar todos los electrones se tienen en cuenta los espacios por donde los electrones se pueden mover $f_h(E)$.

	Electrones	Huecos
Partículas móviles por unidad de volumen	$n_0 = \int_{E_c}^{\infty} D_e(E) f_e(E) dE$	$p_0 = \int_{-\infty}^{E_v} D_h(E) f_h(E) dE$
Densidad de estados	$D_e(E) = \frac{\sqrt{2}(m^*)^{\frac{3}{2}}}{\pi^2 \hbar^3} (E - E_c)^{\frac{1}{2}}$	$D_h(E) = \frac{\sqrt{2}(m^*)^{\frac{3}{2}}}{\pi^2 \hbar^3} (E_v - E)^{\frac{1}{2}}$
Función de distribución	$f_e(E) = \frac{1}{1 + \exp\left(\frac{E - E_f}{kT}\right)}$	$f_h(E) = \frac{1}{1 + \exp\left(\frac{E_f - E}{kT}\right)}$

En ambos casos es necesario resolver una ecuación del siguiente tipo:

$$\int_0^{\infty} \frac{x^{\frac{1}{2}}}{\exp(x - \alpha) + 1} dx;$$

donde para los electrones $x \equiv \frac{E - E_c}{kT}$; $\alpha \equiv \frac{E_f - E_c}{kT}$

Esta ecuación sólo se puede resolver en forma numérica, o encontrando algunas soluciones aproximadas válidas en ciertas regiones de energía. Suponiendo que el nivel de Fermi está dentro de la región de energía prohibida, cuando el argumento de la exponencial es positivo, en el numerador puede despreciarse el uno.

Si empleamos esta aproximación para calcular el número de electrones y de huecos, corresponderá a materiales que se llaman semiconductores no degenerados. Los que no cumplen esta condición se llaman semiconductores degenerados, como por ejemplo, los metales donde el nivel de Fermi está dentro de las bandas permitidas. La solución de esta integral es analítica y se conoce como integral de Fermi de un medio o función gama.

$$\int_0^{\infty} x^{\frac{1}{2}} \exp(-x) dx = \frac{\sqrt{\pi}}{2} \quad (II.24)$$

Lo que produce las siguientes expresiones para los electrones n_0 y para los huecos p_0 :

$$n_0 = N_c \exp\left(-\frac{E_c - E_f}{kT}\right); N_c = 2 \left(\frac{2\pi m_e^* kT}{h^2}\right)^{\frac{3}{2}} \quad (II.25)$$

$$p_0 = N_v \exp\left(-\frac{E_f - E_v}{kT}\right); N_v = 2 \left(\frac{2\pi m_h^* kT}{h^2}\right)^{\frac{3}{2}} \quad (II.26)$$

Semiconductores intrínsecos

Definimos un semiconductor intrínseco como aquel en que la agitación térmica transporta un electrón de la banda de valencia a la banda de conducción, quedando así un hueco en la banda de valencia.

En todos estos materiales se cumple la siguiente relación:

$$n_0 = p_0 = n_i \quad (II.27)$$

donde n_i es la concentración de los electrones que pueden pasar de la banda de valencia a la de conducción en equilibrio en un material intrínseco.

En un material intrínseco, usando las expresiones para la densidad de portadores se puede demostrar que la posición de nivel de Fermi está aproximadamente a la mitad del intervalo de energía prohibido y que la densidad de electrones intrínsecos se define únicamente por la mitad del ancho de energía prohibido y la temperatura.

$$E_f = E_v + \frac{E_g}{2} + \frac{kT}{2} \ln \left(\frac{N_v}{N_c} \right) \quad (\text{II.28})$$

$$n_i = \sqrt{N_c N_v} \exp \left(-\frac{E_g}{2kT} \right) \quad (\text{II.29})$$

En cualquier semiconductor no degenerado en equilibrio termodinámico, la densidad de electrones y de huecos no puede fijarse arbitrariamente, ya que las dos están relacionadas mediante la ley de acción de masas ($n_0 p_0 = n_i^2$). Esto resulta de reconocer que la densidad de electrones es proporcional a la distancia entre el nivel de Fermi E_f y el mínimo de la banda de conducción E_c , mientras que la densidad de huecos es proporcional a la distancia entre el máximo de la banda de valencia E_v y el nivel de Fermi E_f , y la suma de estas dos distancias es constante e igual al intervalo de energía prohibido E_g en equilibrio.

Antes de describir el proceso mediante el cual las impurezas aportan a los materiales semiconductores propiedades interesantes, describamos el proceso de conducción y de generación de portadores en un semiconductor intrínseco. Para ello imaginemos un modelo cuasiquímico donde el silicio es un material plano. Considérese sólo el efecto de la temperatura T y del campo eléctrico E (figura II.18).

¿Pero qué se puede hacer para lograr que la concentración de portadores en equilibrio se modifique, más allá del control definido por la temperatura?

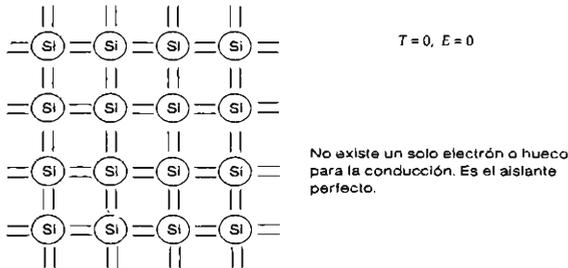


Figura II.18. Modelo de silicio a bajas temperaturas y sin campo aplicado.

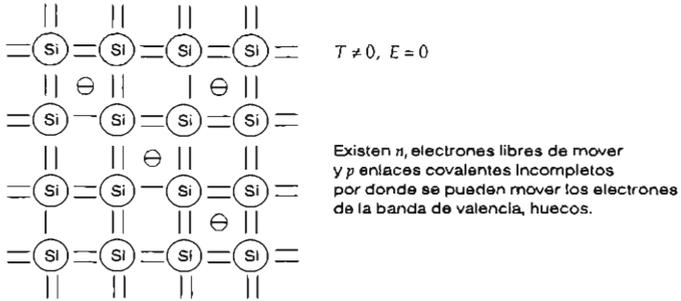


Figura II.19. Modelo de silicio a temperatura diferente de cero y sin campo aplicado.

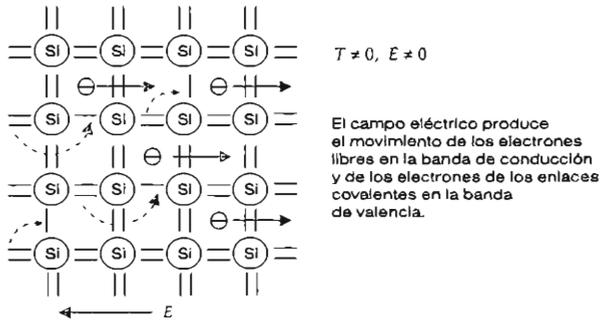


Figura II.20. Modelo de silicio a temperatura diferente de cero y con campo aplicado E .

Semiconductores extrínsecos

La introducción de impurezas en el semiconductor tiene una enorme importancia en electrónica. Aquí la palabra impureza tiene el sentido

de átomos diferentes al que forma la red del semiconductor; por simplicidad supondremos que todas las impurezas son de un solo tipo de átomo. Su importancia en la electrónica radica en que gracias a éstas podemos obtener las gran-

Densidad de átomos de impurezas en el silicio, N_B

$N_B < 10^{10} \text{ cm}^{-3}$	$10^{12} \text{ cm}^{-3} < N_B < 10^{20} \text{ cm}^{-3}$	$N_B > 10^{22} \text{ cm}^{-3}$
Esta densidad de impurezas es tan pequeña que no existe la tecnología para su control, así que se acepta cualquier tipo de impureza.	En este intervalo se encuentran los semiconductores de la industria electrónica, siempre y cuando la cantidad de impurezas esté bajo control. Dependiendo del dispositivo, es necesario obtener cierta concentración en particular.	Estas densidades de impurezas son tan grandes que se trabajan como si fueran aleaciones y no se emplean como material activo.



Figura II.22. Modelo de silicio extrínseco (tipo p).

aceptores se crean cuando los electrones, que normalmente deberían ocupar los estados cercanos a la parte superior de la banda de valencia, se desplazan a niveles aceptores que estaban

vacíos en el principio y cuya energía se localiza a unas cuantas centésimas de eV por encima de E_V . El siguiente diagrama representa los niveles de las impurezas más representativas del silicio:

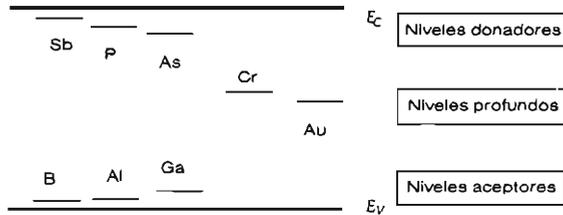


Figura II.23. Niveles de impureza del silicio.

Antes de situar el nivel de Fermi en los semiconductores extrínsecos, veamos cómo representar la condición de neutralidad eléctrica. Si al semiconductor no le hemos agregado carga eléctrica (las impurezas son agregadas como átomos neutros) debe conservarse la condición de que la suma de la carga positiva es igual a la suma de la carga negativa.

En el caso más general, las cargas eléctricas en el semiconductor pueden surgir de los electrones libres n_0 , los huecos p_0 , las impurezas donadoras y las aceptadoras ionizadas p_D y n_A . Conforme la temperatura cambia, la distribución de la carga

eléctrica entre las diferentes fuentes cambia. Las impurezas totales donadoras N_D pueden estar ionizadas p_D o no n_D . Las impurezas totales aceptadoras N_A pueden estar ionizadas n_A o no p_A .

Esto da origen a la siguiente expresión para la ecuación de neutralidad eléctrica:

$$n_0 + n_A = p_0 + p_D$$

$$N_D = p_D + n_D$$

$$N_A = p_A + n_A \quad (II.30)$$

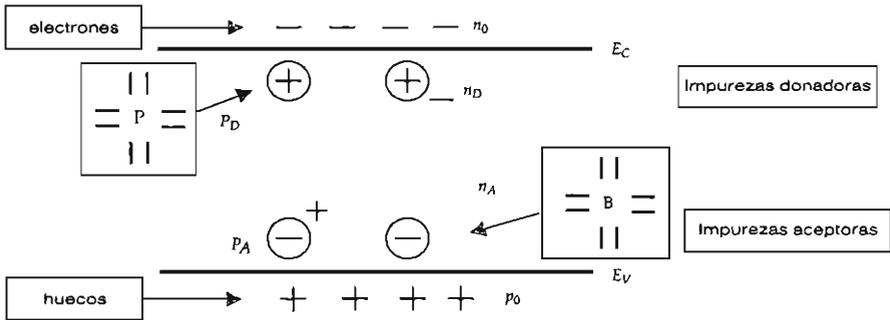


Figura II.24. Diagrama de las diferentes cargas eléctricas en el semiconductor.

Veamos la utilidad de la ecuación de neutralidad eléctrica analizando la cantidad de portadores, huecos y electrones que se tienen en un semiconductor tipo n ($N_A = n_A = p_A = 0$), a una temperatura lo suficientemente alta para que las impurezas estén ionizadas, y suficientemente baja para que la cantidad de portadores intrínsecos no sea muy elevada. El silicio a la temperatura ambiente corresponde a este caso.

Si todas las impurezas están ionizadas:

$$n_D = 0$$

Si sólo tenemos impurezas donadoras:

$$N_A = 0$$

La ecuación de neutralidad eléctrica en este caso es:

$$n_0 = p_0 + N_D$$

Con la ecuación de la ley de acción de masas se obtiene:

$$n_0 = \frac{N_D + \sqrt{N_D^2 + 4n_i^2}}{2} \approx N_D \quad (\text{II.31})$$

La posición del nivel de Fermi es una función del contenido de impureza en el material y de la temperatura. En un semiconductor intrínseco encontramos que el nivel de Fermi permanece aproximadamente en el centro de la banda prohibida de energía para cualquier temperatura, mientras que en un semiconductor extrínseco la

posición del nivel de Fermi depende del estado de ionización de las impurezas.

Si estudiamos un semiconductor tipo n , encontramos que en el cero absoluto de temperatura no hay electrones ni huecos para la conducción. Conforme la temperatura aumenta, la formación de electrones es favorecida por la io-

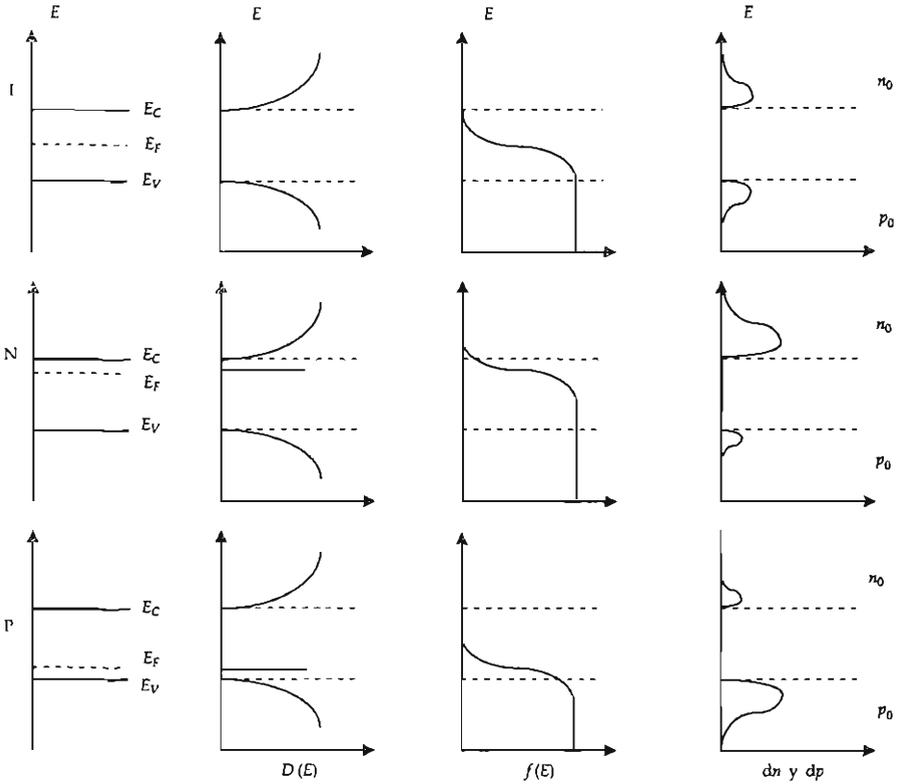


Figura 11.25. Variación de los diagramas de energía E , densidad de estados $D(E)$, función de distribución $f(E)$ y densidad de electrones n_0 y huecos p_0 .

nización de las impurezas, con una energía de activación ($E_c - E_D$). Si la energía necesaria para formar un electrón es ésta, ello indicaría que el nivel de Fermi E_f estaría muy cerca del mínimo de la banda de conducción E_c , ya que, como indicamos anteriormente, la distancia entre E_f y E_c es proporcional al número de electrones.

A una temperatura lo suficientemente alta es factible la transición de electrones de la banda de valencia a la de conducción. Incluso a una temperatura determinada, puede suceder que $n_i > N_D$ y entonces n_0 sea igual a p_0 por lo que E_f estaría al centro de la banda prohibida.

La región de ionización de impurezas se extiende hasta la temperatura de saturación T_s . Esta

temperatura se puede calcular suponiendo que en ella la cantidad de electrones es casi igual a la cantidad de impurezas donadoras.

$$T_s = \frac{E_c - E_D}{k \ln \left(\frac{N_c}{N_D} \right)} \quad (11.32)$$

La región de conductividad intrínseca se obtiene para temperaturas más altas que T_i y puede definirse de una manera similar a la anterior:

$$T_i = \frac{E_g}{k \ln \left(\frac{N_c N_v}{N_D^2} \right)} \quad (11.33)$$

La figura II.25 presenta un resumen de lo visto hasta este momento: I , N , y P representan a un semiconductor intrínseco, un extrínseco tipo n y otro tipo p , respectivamente. $D(E)$ es la densidad de estados disponibles para los electrones y es la misma función para los tres casos. Únicamente difiere para los extrínsecos porque se agregan los niveles de impureza; $f(E)$ es la función de distribución, y es la misma para los tres casos, sólo que el nivel de Fermi que define la probabilidad de ocupación de un medio está en el centro de la banda prohibida en los intrínsecos, cerca del mínimo en la banda de conducción en los semiconductores tipo n y cerca del máximo en la banda de valencia en los tipo p . Los semiconductores intrínsecos tienen la misma cantidad de huecos y electrones, sólo existen más electrones en el semiconductor tipo n y más huecos en un semiconductor tipo p .

Conductividad

Con la información acumulada hasta este momento sabemos que la conductividad en un semiconductor puede cambiar sensiblemente con la temperatura, y esta dependencia se expresa de la siguiente manera:

$$\sigma(T) = q \{ n(T)\mu_n(T) + p(T)\mu_p(T) \} \quad (II.34)$$

La movilidad de los portadores de carga μ , como hemos mencionado anteriormente, depende de las imperfecciones del cristal. En este caso, la perturbación más importante en el movimiento de los electrones son las vibraciones térmicas de la red cristalina. La siguiente es una expresión que muestra la dependencia de la movilidad que surge por las vibraciones térmicas; ya que su deducción está fuera de los objetivos del curso, bien puede considerarse empírica.

$$\mu = A \frac{\sqrt{8 \pi q \hbar^4}}{3 m^{*3/2} (kT)^{3/2}} \propto m^{*-3/2} T^{-3/2} \quad (II.35)$$

A es una constante determinada por parámetros elásticos del material.

En un semiconductor intrínseco es fácil plantear una ecuación que muestre la dependencia de la conductividad respecto a la temperatura. Empleando la conocida expresión sobre la densidad de los portadores, y suponiendo que las masas efectivas de los huecos y electrones no son muy diferentes, se obtiene:

$$\sigma \approx 2q n_i \mu \propto \exp\left(-\frac{E_g}{2kT}\right) \quad (II.36)$$

Esta ecuación muestra una vez más la dependencia general de un proceso activado por temperatura. Al compararla con la ecuación de Boltzmann, se reconoce que la energía de activación es la mitad del ancho de la energía prohibida, ya que, al pasar un electrón de la banda de valencia a la de conducción, se generan dos portadores de carga, un electrón y un hueco.

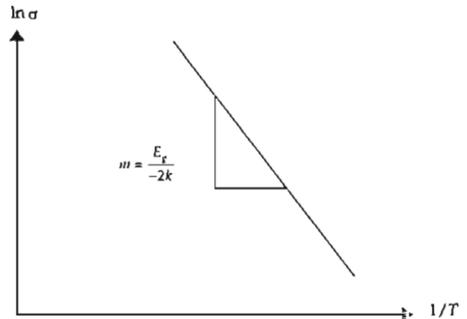


Figura II.26. Variación de la conductividad como función de la temperatura en un semiconductor intrínseco.

La conductividad en un semiconductor intrínseco aumenta exponencialmente con la temperatura, a diferencia de los metales, en los que la conductividad disminuye debido a que la movilidad disminuye y la densidad de portadores de carga es una constante.

Para un semiconductor extrínseco tipo n , vemos que la dependencia de la conductividad respecto a la temperatura tiene la siguiente forma (para un semiconductor tipo p el comportamiento es en esencia el mismo):

$$\sigma(T) \approx q n(T) \mu_n(T) \quad (II.37)$$

En este tipo de materiales, el comportamiento está dominado por los portadores móviles. La densidad de portadores móviles en los semiconductores extrínsecos no tiene una forma simple, sino que es una función del estado de ionización de las impurezas y de la formación de pares electrón-hueco.

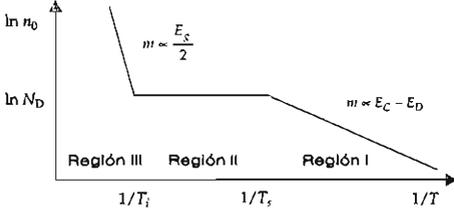


Figura II.27. Variación de densidad de portadores libres por la temperatura en un semiconductor tipo n.

En cualquier región de temperatura, la cantidad total de portadores es la suma de los portadores intrínsecos y los portadores proporcionados por las impurezas.

En la región I la temperatura es tan baja que la energía promedio no es suficiente para crear una cantidad significativa de pares electrón-hueco (energía de activación $= E_g/2 \sim 0.5 \text{ eV}$), mientras que las impurezas requieren una energía de activación mucho menor para producir electrones por ionización (energía de activación $= E_c - E_D \sim 0.05 \text{ eV}$, la energía de activación es ésta y no la mitad, ya que no hay conducción a través de los niveles de las impurezas). Cuando la temperatura en el semiconductor es igual a la temperatura de saturación T_s , todas las impurezas han sido ionizadas. La conductividad en esta región tendría la siguiente expresión, donde A_1 es independiente de la temperatura:

$$\sigma(T) = A_1 \exp\left[-\frac{E_c - E_D}{kT}\right] \quad (II.38)$$

En la región II, que comprende desde la temperatura de saturación T_s hasta la temperatura intrínseca T_i , la densidad de portadores libres se mantiene constante, ya que las impurezas están todas ionizadas y aún no existe suficiente energía térmica promedio para que la generación de pa-

res electrón-hueco sea comparable a la densidad de impurezas ionizadas. Ésta es la región donde los semiconductores son empleados en electrónica para las aplicaciones en dispositivos activos. Aquí la densidad de portadores es una constante, N_D , mientras que la movilidad hace disminuir la conductividad. La expresión de conductividad tiene la siguiente forma, donde A_{11} es otra constante independiente de la temperatura:

$$\sigma(T) = A_{11} T^{-\frac{3}{2}} \quad (II.39)$$

En la región III, cuando la temperatura es mayor que T_i , la energía térmica promedio es suficientemente alta como para que la densidad de pares electrón-hueco sea aún mayor que la densidad de portadores proporcionados por las impurezas, así que la cantidad de portadores aumenta una vez más, pero ahora se debe a los pares electrón-hueco. El comportamiento de cualquier semiconductor en esta región es muy parecido al de los semiconductores intrínsecos. La expresión de conductividad tiene la misma forma que en los semiconductores intrínsecos, donde A_{III} es independiente de la temperatura:

$$\sigma(T) = A_{III} \exp\left[-\frac{E_g}{2kT}\right] \quad (II.40)$$

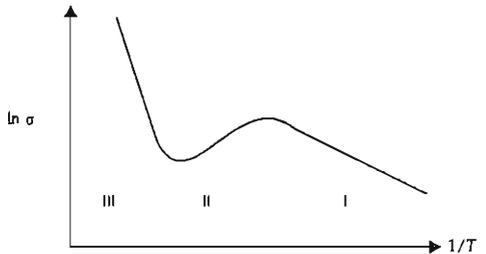


Figura II.28. Variación de la conductividad con la temperatura en un semiconductor extrínseco.

De acuerdo con la figura anterior es posible identificar las regiones I y III, donde la conductividad aumenta con la temperatura debido al aumento en la densidad de portadores, y la región II, donde la conductividad disminuye con la temperatura a causa de la variación de la movilidad de los portadores.

Ecuación de continuidad: generación y recombinación

$$n(x, y, z, t) = n_0 + \Delta n(x, y, z, t)$$

$$p(x, y, z, t) = p_0 + \Delta p(x, y, z, t) \quad (II.42)$$

Esta sección trata de las ecuaciones que describen el comportamiento de semiconductores fuera del equilibrio. Hasta este punto las características principales de los semiconductores estaban determinadas al conocerse n_0 . Algunos factores externos ocasionan que el semiconductor esté fuera de equilibrio térmico, como los gradientes de temperatura, los campos eléctricos y magnéticos estáticos y variables, la radiación luminosa, etc. Estas variaciones pueden ser locales, como las causadas al iluminar una parte de un semiconductor, o temporales, como las perturbaciones que persisten en un intervalo de tiempo. Al apartarse el semiconductor del equilibrio, el comportamiento dinámico y estático de la densidad de portadores debe ser descrito por un nuevo conjunto de ecuaciones. Las ecuaciones básicas de los semiconductores que describen la operación dinámica de los portadores bajo la influencia de campos externos que ocasionan que el semiconductor pierda el equilibrio térmico pueden clasificarse en los siguientes tres grupos:

- Ecuaciones de Maxwell
- Ecuaciones de densidad de corriente
- Ecuación de continuidad

Para materiales isotrópicos y homogéneos, una de las ecuaciones de Maxwell toma la siguiente forma:

$$\nabla \cdot \mathbf{E} = \frac{\rho(x, y, z)}{\epsilon_s} \quad (II.41)$$

donde \mathbf{E} es el vector del campo eléctrico, ρ es la densidad de carga eléctrica total y ϵ_s es la permisividad del semiconductor

En los semiconductores es posible que agentes externos modifiquen la concentración de portadores de su valor de equilibrio en su posición o en el tiempo, por lo que la concentración de portadores libres es, en general, una función de las coordenadas espaciales y del tiempo, es decir:

Las variaciones en la concentración de portadores se deben a cualquiera de los siguientes fenómenos:

- Generación
- Recombinación
- Difusión
- Deriva

Por generación se entienden los procesos que permiten pasar electrones de la banda de valencia a la de conducción o de los niveles de impureza a la banda de conducción. La generación tiene dos componentes: la generación térmica G_0 , que impone la densidad de portadores en equilibrio, y la generación fuera del equilibrio ΔG , que puede ser causada por luz, partículas nucleares, impacto ionizante (avalancha), etcétera.

$$G = G_0 + \Delta G$$

Por recombinación entendemos el proceso que modifica la cantidad de portadores, aniquilando los pares electrón-hueco. Es el proceso inverso de la generación, y si se tiene una condición de bajo nivel de inyección, es decir, si la cantidad de portadores generados es mucho menor que la de los portadores mayoritarios ($\Delta n < n_0$, por ejemplo), la recombinación de los electrones puede ser aproximada por $\Delta n/\tau$, donde τ es el tiempo de vida de los portadores minoritarios y representa el tiempo promedio en que los electrones fuera del equilibrio tienden a permanecer en la banda de conducción. En equilibrio térmico, la generación es idéntica a la recombinación.

$$R = R_0 + \Delta R = R_0 + \frac{\Delta n}{\tau} \quad (II.43)$$

Los procesos de difusión y de deriva también ocasionan una variación de la cantidad de portadores en el semiconductor. La difusión es el movimiento de los portadores de las regiones de mayor concentración de portadores a las regio-

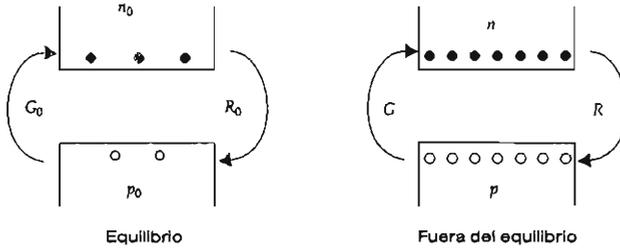


Figura II.29. Generación y recombinación en semiconductores.

nes de menor cantidad de portadores. La fuerza impulsora del movimiento es el gradiente de la concentración.

La deriva es el movimiento de los portadores por la presencia de un campo eléctrico, y la fuerza impulsora es el gradiente de potencial eléctrico, el potencial.

La corriente de deriva fue mencionada al analizar los metales, pero la corriente de difusión no; examinemos un semiconductor heterogéneo en el cual la concentración de electrones (y huecos) cambia de punto a punto. Imaginemos que esto se obtiene contaminando al semiconductor de manera no uniforme.

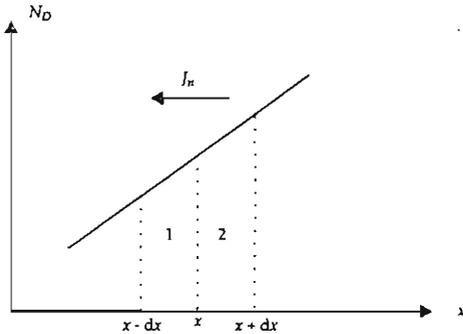


Figura II.30. Movimiento de carga por difusión en un semiconductor inhomogéneo.

Al analizar el movimiento de los portadores de carga de las capas de espesor dx (1 y 2), encontramos que debe existir una corriente de electrones que va de la capa 2 a la capa 1. Para explicar esta corriente, recordemos que el movi-

miento de los electrones es en principio aleatorio y, por lo mismo, puede suceder que de los electrones de la capa 1 la mitad se muevan hacia la izquierda y la otra mitad hacia la derecha, así que la mitad va de la capa 1 a la capa 2. Pero lo mismo es válido para la capa 2 y, como en la capa 2 hay más electrones, es posible que la corriente de electrones se comporte como se señala en la figura. Análíticamente, este proceso se expresa de la siguiente manera:

$$J_n(\text{difusión}) = qD_n \frac{dn}{dx}$$

$$J_p(\text{difusión}) = -qD_p \frac{dp}{dx} \quad (\text{II.44})$$

donde $\frac{dn}{dx}$ es el gradiente de la cantidad de electrones y D_n es la constante de difusión para los electrones [$\text{cm}^2 \text{s}^{-1}$]. Esta corriente de difusión, que existe únicamente en los semiconductores, origina una separación espacial de las cargas por un lado, los iones que están fijos en el material, y por el otro, los electrones que se desplazan por difusión, lo que genera un campo eléctrico estático entre las cargas. El campo eléctrico producido origina corrientes de deriva: en equilibrio termodinámico ambas corrientes se anulan, produciendo una corriente neta igual a cero.

En equilibrio termodinámico es posible relacionar las constantes que aparecen en la corriente de difusión y en la corriente de deriva. Esta relación se conoce como la relación de Einstein, y relaciona la movilidad de los portadores μ con la constante de difusión D a una temperatura T .

$$\frac{\mu_n}{D_n} = \frac{q}{kT} \quad (II.45)$$

Existe una relación similar para los huecos.

La densidad de corriente en los semiconductores puede, en general, tener componentes de huecos y de electrones y provenir de una corriente de difusión y de una corriente de deriva. Así, en general, las ecuaciones de densidad de corriente son:

$$\begin{aligned} J_n &= q\mu_n n E + qD_n \nabla n \\ J_p &= q\mu_p p E - qD_p \nabla p \end{aligned} \quad (II.46)$$

El último grupo de ecuaciones que permiten estudiar los semiconductores fuera del equilibrio son las ecuaciones que expresan las relaciones de continuidad de la cantidad de portadores en el semiconductor con respecto al tiempo y a la posición.

$$\begin{aligned} \frac{\partial n}{\partial t} &= G_n - R_n + \frac{1}{q} \Delta \cdot J_n \\ \frac{\partial p}{\partial t} &= G_p - R_p - \frac{1}{q} \Delta \cdot J_p \end{aligned}$$

Estas relaciones permiten determinar $n(x, y, z, t)$ y $p(x, y, z, t)$, ya que tienen todas las posibles causas de su variación: generación, recombinación, campos eléctricos alternos, campos eléctricos estáticos y difusión.

$$\begin{aligned} \frac{\partial n}{\partial t} &= \Delta G_n - \frac{n - n_0}{\tau_n} + n\mu_n \frac{\partial E}{\partial x} + \mu_n E \frac{\partial n}{\partial x} + D_n \frac{\partial^2 n}{\partial x^2} \\ \frac{\partial p}{\partial t} &= \Delta G_p - \frac{p - p_0}{\tau_p} - p\mu_p \frac{\partial E}{\partial x} - \mu_p E \frac{\partial p}{\partial x} + D_p \frac{\partial^2 p}{\partial x^2} \end{aligned}$$

La siguiente figura trata de proporcionar una imagen hidráulica de los procesos que se han mencionado, en particular la generación que agrega portadores, la recombinación que los desaparece y la difusión y deriva que los mueve de lugar.

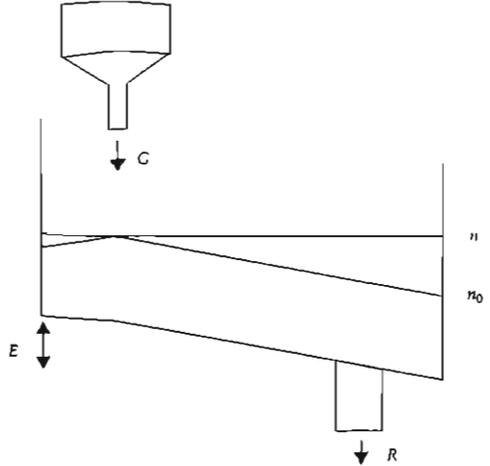


Figura II.31. Imagen que trata de representar los procesos que modifican la densidad de portadores en un semiconductor.

Al conocer la densidad de portadores fuera del equilibrio, se reconoce que los métodos usados en las secciones anteriores no son aplicables en este caso, en el cual parámetros tan importantes como el nivel de Fermi o la ley de acción de masas son inoperantes. Afortunadamente existe un recurso para reunir la información de equilibrio con los conceptos aprendidos al estudiar el estado fuera de equilibrio. Al introducirse el concepto del seudonivel de Fermi, la densidad de electrones en equilibrio o fuera de él se calcula midiendo la distancia entre el seudonivel de Fermi de los electrones E_{Fn} y el mínimo de la banda de conducción E_c . La densidad de huecos en equilibrio o fuera de él se calcula midiendo la distancia entre el seudonivel de Fermi de los huecos E_{Fp} y el máximo de la banda de valencia E_v . En el equilibrio, los dos seudoniveles de Fermi convergen en el nivel de Fermi. A continuación se presentan las ecuaciones en equilibrio y fuera de él, para calcular la densidad de electrones, la densidad de huecos y la ley de acción de masas.

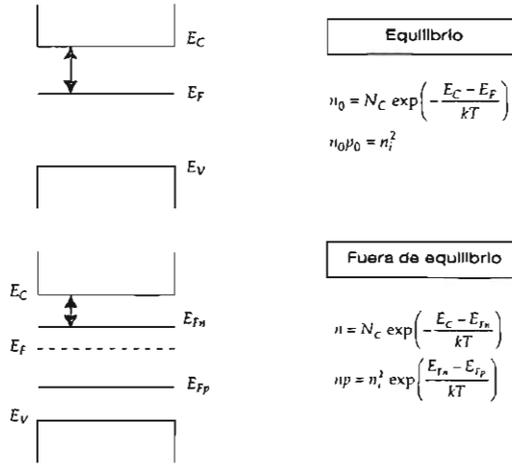


Figura 11.32. Diagrama de energía para un semiconductor en equilibrio y fuera de él, seudoniveles de Fermi.

PROPIEDADES ÓPTICAS Y MAGNÉTICAS

Esta sección tiene un interés informativo, así que su contenido es un tanto superficial. Al hablar de propiedades ópticas, el objetivo principal es comentar cómo se realiza el proceso de las señales ópticas controlando la velocidad de la luz. En relación con los materiales magnéticos, intentaremos proporcionar las ideas esenciales sobre las memorias magnéticas.

Propiedades ópticas

Describir las propiedades ópticas a un estudiante de ingeniería es bastante fácil, ya que está familiarizado con señales que cambian periódicamente en el tiempo, e incluso ha visto en el osciloscopio el efecto de la diferencia de fase de dos señales de voltaje. Las señales ópticas son idénticas a una señal cosenoidal de voltaje, sólo que de mayor frecuencia, $\sim 5 \times 10^{14}$ Hz. En la sección de nomenclatura óptica se habló de la polarización de la luz y se describió como el ángulo de fase entre las vibraciones del campo eléctrico ortogonales a la dirección de propagación. Si la luz se propaga en la dirección cartesiana z, entonces, en general, el ángulo de fase entre las

vibraciones del campo eléctrico en la dirección x y y definirá la polarización.

La velocidad de propagación de la luz cambia en los diferentes medios. En el vacío se propaga a su velocidad máxima y en los medios transparentes la propagación es siempre menor: $v = \frac{c}{n}$, donde n es el índice de refracción del medio. En el aire n es casi uno, en el vidrio de ventana es 1.4, etc. El índice de refracción cambia en las diferentes longitudes de onda, lo que da origen al arco iris. En general, en un material, n aumenta al disminuir la longitud de onda (por cierto, ¿saben por qué el cielo es azul de día y negro de noche?).

De la misma manera en que el ángulo de fase en un circuito eléctrico se modifica al cambiar la impedancia (frecuencia o elemento físico), así también la polarización de la luz puede cambiar al pasar por un material adecuado. Imaginen un haz de luz monocromático (una longitud de onda) y polarizado linealmente (ángulo de fase cero). Cuando la luz incide en un material adecuado, uno que tenga dos diferentes valores de n en diferentes direcciones, n_x y n_y , (o sea, un material anisotrópico; muchos cristales en la naturaleza tienen esta propiedad, a diferencia de los vidrios, que son materiales no cristalinos que

tienen propiedades ópticas isotrópicas), al salir el haz de luz del material, la componente x del campo eléctrico que viajó a una velocidad definida por n_x , estará adelante o atrás de la componente y del campo eléctrico que viajó a una velocidad diferente definida por n_y . El ángulo de fase (ϕ) entre los dos componentes del campo eléctrico se puede escribir de la siguiente manera:

$$\phi = \frac{2\pi |n_x - n_y| t}{\lambda}$$

donde t es el espesor del material y λ es la longitud de onda de la luz. Existen dispositivos ópticos que modifican las propiedades de la luz. Un polarizador es un material que cambia el ángulo de fase a cero después de pasar por el material; una placa $\lambda/2$ rota el campo eléctrico de la luz polarizada linealmente en un ángulo que depende de la orientación inicial del campo eléctrico y de la placa. Una placa $\lambda/4$ cambia la polarización de lineal a elíptica o circular. Estos dispositivos y algunos otros son empleados en las cabezas de detección de los discos compactos o en las memorias magnetoópticas. En el último caso son esenciales para convertir la informa-

ción del ángulo de polarización de la luz causada por la reflexión en el material magnético en intensidad, que es lo único que miden los fotodiodos.

Existen también dispositivos ópticos de conmutación que emplean estas propiedades para provocar interferencia constructiva o destructiva del campo eléctrico de la luz, y muchos otros fenómenos que tratan de imitar lo que se puede hacer con las señales eléctricas, con la ventaja de que trabajan a mayor frecuencia.

Propiedades magnéticas

Todos los campos magnéticos surgen a partir de circuitos de corriente. La relación fundamental entre dos de las variables de Maxwell: \mathbf{H} , intensidad del campo magnético y \mathbf{B} , inducción magnética, es la siguiente:

$$\mathbf{B} = \mu \mathbf{H}$$

donde μ es la permeabilidad, la cual tiene para el vacío un valor de $4\pi \times 10^{-7} \text{ Wb A}^{-1} \text{ m}^{-1}$ [o henry/metro].

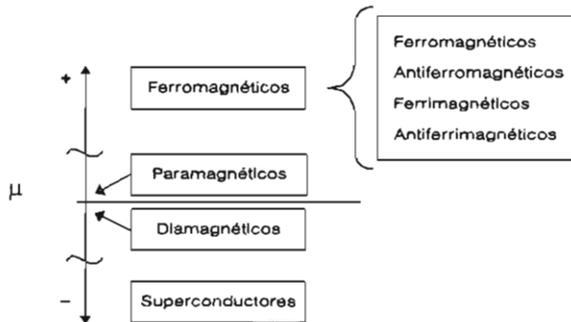


Figura 11.33. Clasificación de los materiales magnéticos.

Los materiales magnéticos se clasifican en:

a) *Diamagnéticos*. Donde μ es una constante independiente del campo y de valor negativo muy pequeño. Este efecto se origina por los electrones que giran alrededor del núcleo y no tiene

aplicación práctica excepto en los superconductores, que son diamagnéticos perfectos donde la constante μ negativa es muy grande. Esta constante se tiene en cuenta cuando se buscan aplicaciones, como la sustentación magnética.

b) *Paramagnéticos*. Algunos átomos o iones presentan momentos magnéticos permanentes, que en ausencia de campos magnéticos se orientan aleatoriamente de manera independiente. μ es positiva pero muy pequeña. El paramagnetismo tiene su origen en el espín electrónico y compite con el diamagnetismo que está presente en todos los materiales.

c) *Ferromagnéticos*. Son los materiales que presentan magnetismo en ausencia de campos magnéticos. Entender su origen es un poco más complicado; se basan en la formación de dominios magnéticos, donde los espines atómicos se alinean en la misma dirección, lo que indica un proceso colectivo, a diferencia de los dos procesos anteriores. Varios nuevos conceptos surgen al estudiar estos materiales; entre ellos están el antiferromagnetismo, que presenta dominios magnéticos pero con los espines en dirección contraria, los ferrimagnéticos, un grupo especial de materiales; las ferritas, cuyas propiedades ferromagnéticas son intermedias de las dos anteriores, pero muestran una μ positiva y muy grande, además de depender del campo magnético.

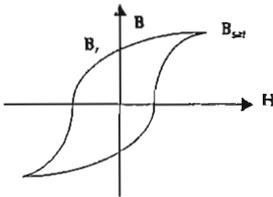


Figura 11.34. Ciclo de histéresis magnética.

La diferencia en el comportamiento magnético entre diamagnéticos y paramagnéticos, por un lado, y ferromagnéticos por el otro alcanza la magnitud de 10^9 . Los ferromagnéticos se caracterizan por su dependencia de μ con el campo magnético, llamada ciclo de histéresis. El origen de este ciclo de histéresis magnética es la energía necesaria para modificar la orientación de los dominios magnéticos. En este punto podemos aclarar que las propiedades magnéticas colectivas (como ésta) desaparecen a determinadas temperaturas que comúnmente se llaman tem-

peraturas de Curie. Una curva de **H versus B** nos permite visualizar el ciclo de histéresis magnética.

Se parte de un material virgen sin inducción magnética presente; es decir, en el origen de las coordenadas. Un aumento en la intensidad del campo magnético **H** producirá un aumento en la inducción magnética hasta que se satura B_{sat} , y no aumenta más. Aunque se continúe aumentando el campo magnético, los dominios magnéticos del material están ordenados en forma paralela al campo externo. Si se disminuye el campo magnético **H** la inducción magnética no desaparece y se tiene una inducción magnética remanente B_r . El ciclo de histéresis magnética indica la energía necesaria para cambiar la orientación de los dominios magnéticos de manera cíclica.

Las diferencias en el ciclo de histéresis indican las diferentes aplicaciones. Los núcleos de los transformadores causan las pérdidas al transformar la energía. Entre mayor sea el ciclo de histéresis (área incluida en el ciclo), mayores son las pérdidas y, conforme menor sea su resistividad, mayores corrientes eléctricas se inducirán, desperdiciando potencia. Los núcleos de los transformadores requieren materiales magnéticos suaves y de gran resistividad, como los que se obtienen al laminar aleaciones de hierro silicio para aplicaciones a bajas frecuencias, así como ferritas para aplicaciones a altas frecuencias.

Por otro lado, los requisitos en las memorias magnéticas son diferentes para los distintos tipos de presentaciones, desde cintas magnéticas de audio, hasta discos magnéticos de alta densidad. Los parámetros, como distancia entre cabeza grabadora/lectora y material magnético, son importantes, ya que la transferencia de la información disminuye exponencialmente con la distancia. El tamaño de la zona con información y la facilidad para grabar y borrar dependerá de la permeabilidad μ . Valores de μ pequeños definirán una baja densidad de información y alta sensibilidad a campos extraños; valores grandes de μ definirán materiales de alta densidad y casi insensibles a campos externos, ya que se requiere un campo magnético mayor para definir las zonas magnéticas que forman la información.

TECNOLOGÍA

En las pasadas dos décadas la industria electrónica se ha caracterizado por una evolución en la tecnología de los semiconductores que, entre otras cosas, ha reflejado una tendencia opuesta al proceso inflacionario (desde el punto de vista de una economía como la de Estados Unidos). Este comportamiento se debe en parte a la posibilidad sostenida de fabricar más y más elementos del circuito en un pequeño pedazo de silicio. El aumento en la complejidad de los procesos se ha llamado "very large scale integration" (VLSI). Las compañías de semiconductores son capaces de manufacturar más de un millón de componentes por circuito gracias a la tecnología y al diseño de componentes. El 15 de febrero de 1985 IBM anunció la fabricación de la memoria para computadora DRAM de un millón de bits con dimensión de 5.5 mm × 10.5 mm. En su momento fue la más rápida y densa jamás imaginada. Hoy día, en 1995, se fabrican en serie memorias de 64 megabits y tiene dos años el proyecto de la memoria de 256 megabits entre IBM, Siemens y Toshiba.

La estructura de un circuito integrado es compleja, tanto en la topografía superficial como en su composición interna. De hecho, cada componente tiene una arquitectura tridimensional que deberá ser la misma para cada circuito. La estructura del circuito se logra por medio de muchas capas. Cada una forma parte de un patrón detallado; algunas se forman en el silicio, mientras que otras lo hacen en su superficie. Las operaciones del proceso de la oblea se pueden clasificar como aditivas (epitaxia, oxidación, contaminación y metalizado), selectivas (litográficas) y sustractivas (ataque químico o físico). El número de repeticiones de cada proceso depende de la tecnología y del grado de complejidad del circuito.

El punto de partida en la tecnología del silicio es contar con obleas de silicio monocristalino de gran pureza, o "grado electrónico". Este material se obtiene a partir del óxido de silicio SiO₂; por una reacción de óxido-reducción se obtiene el silicio metalúrgico (98% de pureza). La purificación se completa al formar un compuesto gaseoso de silicio que se purifica para después obtener

el silicio sólido policristalino. Es común que al crecer el lingote se modifiquen las características eléctricas, agregando impurezas que transfieren propiedades eléctricas del tipo *n* o *p*. El lingote se obtiene por técnica Czochralski, en la cual el silicio fundido es solidificado al jalar un trozo de cristal perfecto mientras se gira. De ahí se obtienen obleas de silicio con orientación cristalina definida y con un diámetro desde 5 hasta 30 cm. La oblea debe ser monocristalina y no tener imperfecciones mecánicas o cristalográficas que degraden las propiedades eléctricas.

La microelectrónica, que ha permitido el enorme desarrollo de la electrónica, puede ilustrarse a través de dos de sus estructuras básicas: un transistor de efecto de campo con compuerta formada por un capacitor MOSFET, y un transistor de unión BJT. El MOSFET de canal *n* parte de un sustrato formado por una oblea de silicio tipo *p*. Las regiones de fuente y de drenaje se forman mediante la creación de una zona superficial tipo *n*. El óxido de silicio delgado y de alta calidad se emplea para formar la compuerta, mientras que el óxido grueso se emplea para aislar dispositivos. La zona de silicio policristalino permite tener acceso a la compuerta, en tanto que el aluminio se emplea para conectar la fuente y el drenaje. El BJT se construye alternando regiones de conductividad tipo *n* y *p*, partiendo de un semiconductor tipo *p*. El óxido se emplea para aislar las diferentes zonas y se emplea un metal para formar los contactos en el emisor, la base y el colector.

Para cualquier desarrollo es necesario emplear varios procesos:

- Oxidación: el óxido de silicio se forma al exponer el silicio a altas temperaturas (1 100°C) en presencia de oxígeno. Se realizan así varias funciones: enmascarar y limitar la difusión en áreas específicas, pasivar y proteger la superficie del silicio, aislar un dispositivo eléctrico de otro, y fabricar dieléctricos para las compuertas MOS.
- Fotolitografía: es el proceso que permite seleccionar diferentes áreas que se emplearán en los dispositivos, transfiriendo patrones definidos en el diseño. El óxido de silicio puede detener la penetración de impurezas,

por lo que en su fabricación se emplea un polímero fotosensible que permite, con ayuda de exposición selectiva y del ataque químico, definir las estructuras sobre el óxido de silicio para después transferirlas en estructuras de concentración de impurezas.

- Ataque químico o físico: es el proceso mediante el cual los materiales se remueven selectivamente. En un principio, el ataque en fase líquida fue suficiente, pero delinear estructuras de un micrómetro o menores sólo es posible por medio del ataque químico por plasmas.
- Difusión: es el proceso mediante el cual se introducen las impurezas químicas. A menudo se realiza a muy altas temperaturas en presencia de las especies contaminantes en fase de vapor y tiene el propósito de modificar las características eléctricas y el tipo de portadores de carga y densidad.
- Evaporación: es la manera de depositar metales. El metal se evapora y se transporta hasta el silicio con el propósito de formar líneas de interconexión.
- Depósito de sólidos a partir de fases gaseosas (CVD): ésta es la técnica para depositar óxidos o silicio policristalino a bajas temperaturas (500°C).
- Implantación de iones: se emplea para introducir impurezas al semiconductor a temperaturas relativamente bajas. Su principal ventaja radica en el control que se tiene sobre la cantidad y el perfil de las impurezas.
- Epitaxia: es la técnica que permite crecer silicio sobre silicio conservando las propiedades de cristalinidad. Usualmente se emplea para mejorar las propiedades del sustrato, ya que sobre esta capa se construirá el circuito.

Al terminar la oblea, se monta en una computadora que realiza la verificación de cada dado: las que no pasan las pruebas eléctricas son marcadas y posteriormente desechadas. Las obleas montadas en un bastidor de maylar y se cortan para separar los dados. Los dados aceptables se montan en un soporte que contiene las conexiones externas (pins) y, por medio de alam-

bre de oro, se conecta el dado al soporte y se sella e identifica el circuito integrado.

Los procesos tecnológicos más importantes son:

- NMOS (*n-channel devices*): tecnología que emplea como dispositivo activo el MOSFET de canal de electrones.
- CMOS (*complementary MOS*): tecnología donde un transistor NMOS y otro PMOS trabajan simultáneamente como dispositivos activos.
- Bipolar

El proceso NMOS empieza con la formación de un óxido delgado sobre la oblea de silicio tipo *p*, y sobre éste se deposita una película de nitruro de silicio a baja temperatura. Empleando la mascarilla 1 se define el área activa del transistor removiendo el nitruro y el óxido excepto donde se formará el transistor. El implante de boro permitirá aislar los diferentes transistores; posteriormente se oxida toda la oblea. Hasta este momento el nitruro protege la zona que formará el transistor; se remueve entonces el nitruro y se crece el óxido de alta calidad que constituirá la compuerta del transistor. El ajuste del voltaje de encendido puede realizarse en este momento.

El silicio policristalino se deposita sobre la oblea por medio de CVD y la mascarilla 2 se emplea para definir las regiones que formarán la compuerta del transistor y su conexión. Después de remover el silicio policristalino no deseado se procede a formar las regiones de fuente y de drenaje por implante de iones, y después se aplicará un proceso a alta temperatura que redistribuye las impurezas y crece un óxido. La mascarilla 3 se emplea para remover el óxido sobre las regiones donde se conectará eléctricamente el transistor. El depósito del aluminio cubre la oblea por completo, así que la mascarilla 4 se emplea para definir las regiones donde debe removerse el metal para formar así las pistas metálicas. Finalmente, se recubre la oblea de nuevo con algún vidrio protector y la mascarilla 5 se emplea para descubrir las regiones metálicas donde se interconectarán los alambres externos.

El proceso CMOS se emplea para formar dos transistores adyacentes, un NMOS y un PMOS que

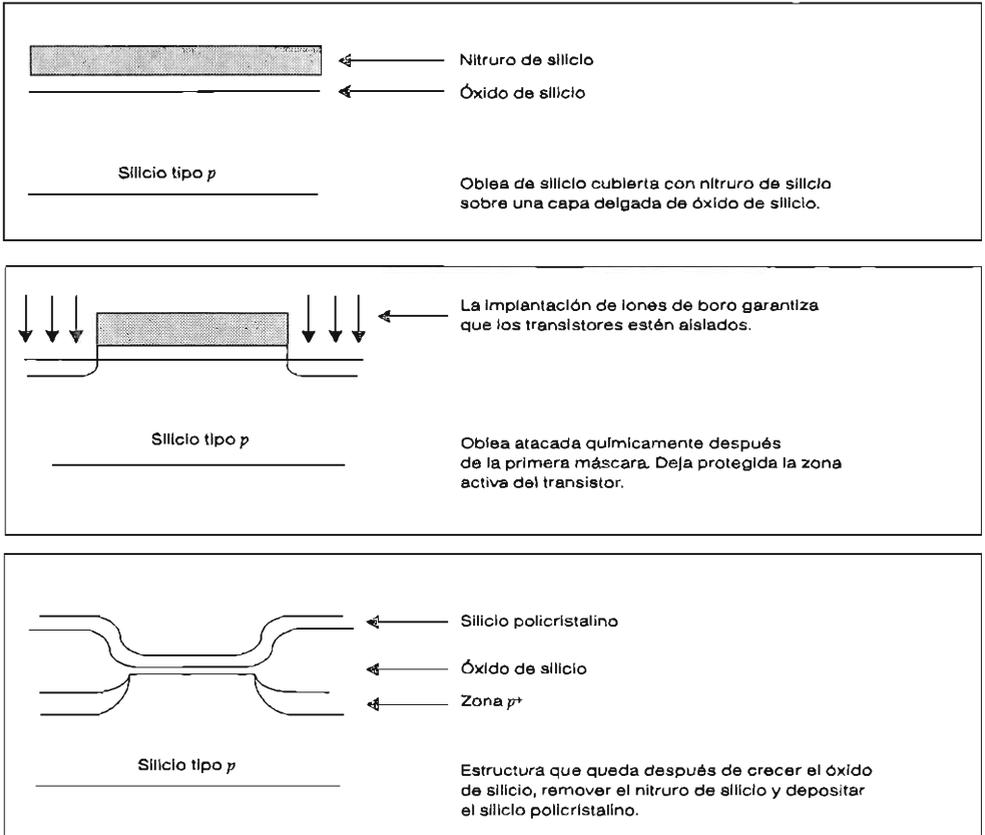


Figura II.35. Primeras secuencias en el proceso CMOS.

llamaremos gemelo. Aquí se necesita una máscara adicional que dará origen a la región que formará el transistor gemelo, y otra para la formación de la fuente y el drenaje del transistor gemelo.

El proceso bipolar es un poco más complejo y su control es más difícil. La intención es tener verticalmente tres regiones de diferente conductividad para formar un transistor npn , pero la conexión será horizontal y los accesos eléctricos están uno después de otro en la superficie. Para disminuir la resistencia del colector (la capa más profunda y por lo tanto con menor densidad de impurezas) es necesario enterrar una región de

n' cercana al colector. Además, los procesos de cambio de conductividad n a p (base) y p a n (emisor) se realizan a altas temperaturas sobre sustratos epitaxiales. El número de mascarillas es el mismo que en el CMOS.

El primer paso en el desarrollo de un circuito microelectrónico son las especificaciones y características funcionales. Después de este paso empieza el diseño propiamente dicho. El diseño asistido por computadora CAD es el más común, ya que además de permitir la construcción del circuito, hace posible simular sus características, diseñar el proceso tecnológico y construir las mascarillas. La importancia de la simulación en

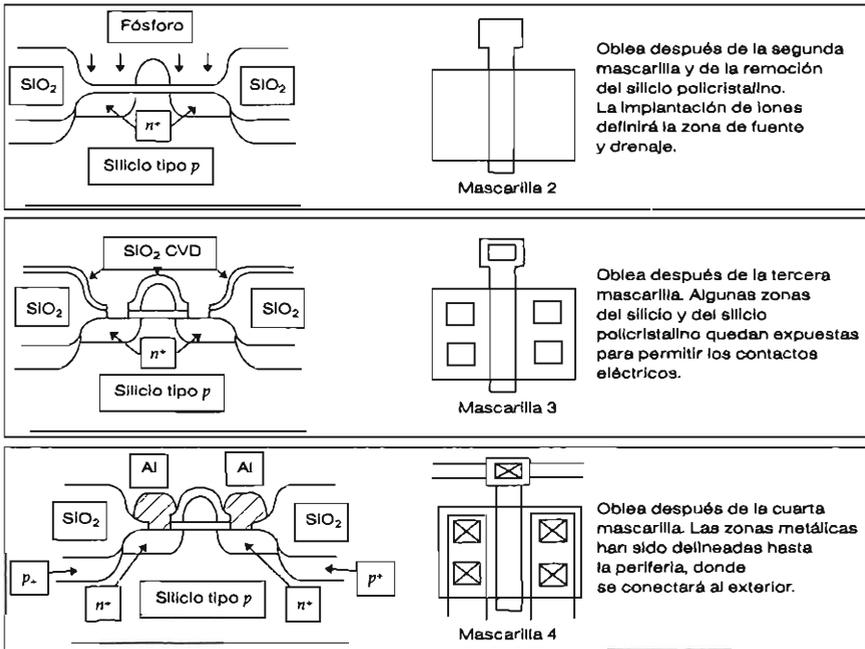


Figura II.36. Secuencias finales en el proceso NMOS.

todos los detalles es primordial. Baste recordar que un lote virgen de obleas (que incluye 20) puede costar 20.00 dólares y después de procesado (con las semanas o meses de esfuerzo que ello implica) llega a costar 10 000.

A continuación presentamos tres ilustraciones sobre el proceso de diseño asistido por computadora. La primera ilustración (figura II.37) muestra el diagrama de una compuerta NAND usando la tecnología NMOS (dos transistores y un resistor) por medio de la herramienta MAGIC. Este programa contiene las reglas de diseño y permite la construcción de las mascarillas; en este caso la distancia mínima es de cinco micrómetros. En el diagrama aparecen, además, cinco transistores con diferente geometría. En este diagrama es

posible diferenciar las regiones como las de metal, de silicio policristalino en la compuerta, etcétera.

La simulación del proceso tecnológico se realiza por medio de SUPREM III, que determina los parámetros eléctricos a partir de la simulación de los procesos tecnológicos. En la figura aparece la simulación de la fabricación de la compuerta del transistor NMOS, el programa que describe el proceso y las condiciones de fabricación y los resultados de la distribución de impurezas obtenidas por la simulación con la gráfica del perfil de impurezas. Finalmente se presenta el análisis de los parámetros de uno de los transistores MOSFET, el más grande, usando el analizador de parámetros HP4145DC.

MICROELECTRONICS DESIGN

Auburn University

DATE: Jun 7, 1991

Scale: 140 microns/inch

Filename: rqtcell.cif

User: rtorres

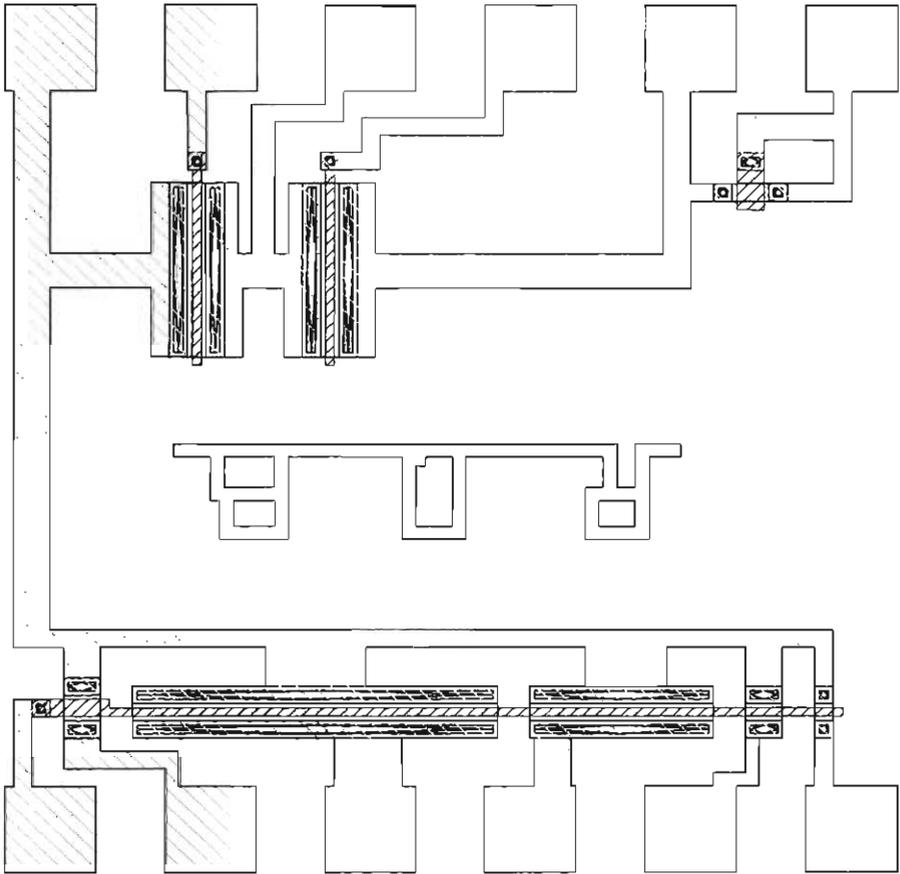


Figura II.37. Gráfica obtenida del simulador MAGIC. En la parte superior hay una compuerta NAND y en la parte inferior cinco transistores con diferente geometría en el canal.

PROGRAMA PARA EL SIMULADOR SUPREM III

```

TITLE           Electrónica Física NMOS GATE

COMMENT        Oblea para el sustrato
INITIALIZE     <100> SILICON, BORON CONCENTRATION=1E15
               THICKNESS=3.0 XDX=1.0 DX=.005 SPACES=150

COMMENT        Crecimiento de óxido de 400 A
DIFFUSION      TEMP=1000 TIME=40 DRYO2

COMMENT        Depósito de nitruro de silicio por CVD
DEPOSIT        NITRIDE THICKNESS=.08 SPACES=15

COMMENT        Crecimiento de óxido de campo
DIFFUSION      TEMP=1000 TIME=180 WET02

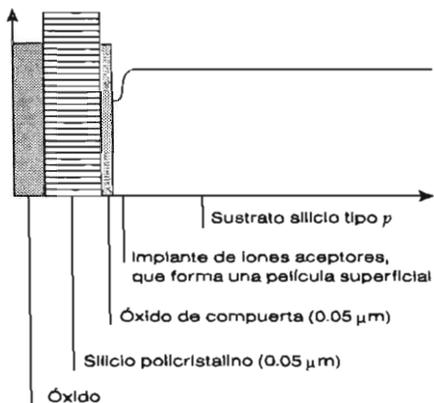
COMMENT        Remove el óxido hasta el silicio
ETCH           OXIDE ALL
ETCH           NITRIDE ALL
ETCH           OXIDE ALL

COMMENT        Implante de iones para ajustar el voltaje de encendido del transistor
IMPLANT        Boron DOSE=4E11 ENERGY=50

COMMENT        Crecimiento de óxido de la compuerta
DIFFUSION      TEMP=1100 TIME=300 DRYO2 HCL%=3

COMMENT        Depósito de silicio policristalino que hace las veces de metal
DEPOSIT        POLY THICKNESS=0.5 TEMO=600
DIFFUSION      TEMP=1000 TIME=30 PHOSPHORUS SOLIDSOLUBILITY

COMMENT        Gráfica de la distribución de materiales, normal a la compuerta
PLOT           CHEMICAL NET LP.PLOT
STOP
    
```



GRAPHICS PLOT

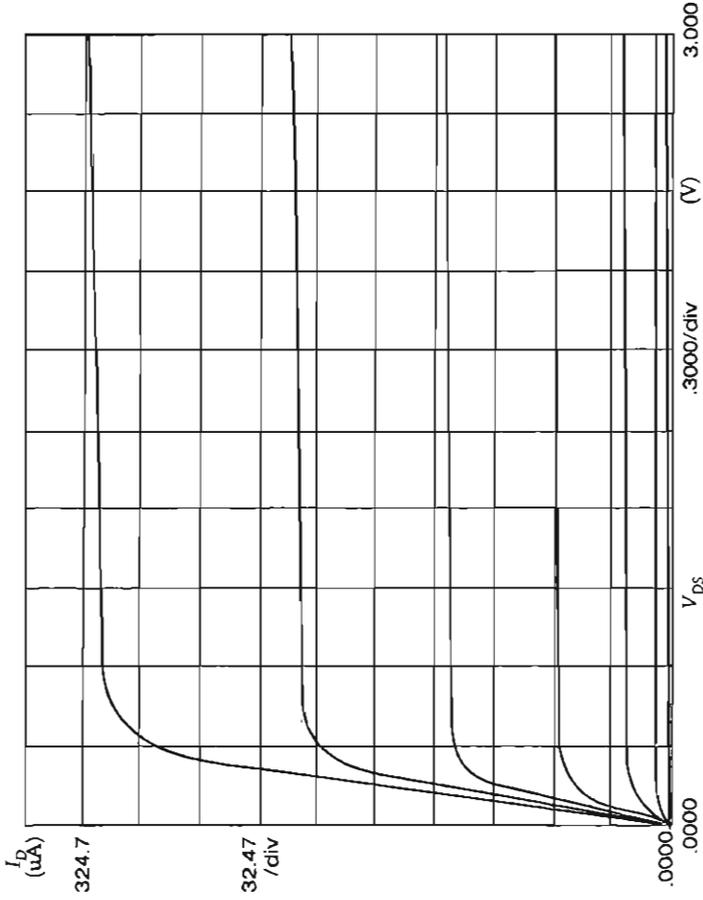


Figura II.38. Mediciones eléctricas de un transistor MOSFET, corriente de drenaje I_D versus voltaje de drenaje V_{DS} teniendo al voltaje de compuerta V_G como parámetro desde 2.5 V hasta 4.5 V.

Variable 1:

V_{DS} -Ch2

Linear sweep

Start .0000V

Stop 3.0000V

Step .0500V

Variable 2:

V_G -Ch3

Start 2.5000V

Stop 4.5000V

Step .2500V

Constante:

V_S -Ch1

.0000V

CONCLUSIONES

En este capítulo se presentó el panorama general de los materiales y las propiedades que son más importantes en el estudio de la electrónica, culminando con una introducción a la tecnología empleada en la manufactura de los circuitos integrados.

El concepto de bandas de energía es esencial para entender las propiedades electrónicas de la materia. Se emplea también para dividir el estudio de los materiales en metales y aislantes. Entre los aislantes, los que tienen el intervalo de energía prohibida cercano a 1.5 eV se llaman semiconductores. Empleando a los metales como ejemplo, se pueden entender conceptos como el de la movilidad de los electrones; los aislantes, por su parte, nos permiten comprender la importancia de la temperatura en los procesos electrónicos, además de ilustrar la mecánica de los procesos colectivos.

Los semiconductores se presentaron con más detalle, ya que son empleados intensivamente en el resto del curso. Los semiconductores intrínsecos, en los cuales la cantidad de huecos y electrones es la misma, así como el efecto de las impurezas sobre el número de portadores y la conductividad, se presentaron con el propósito de entender la dinámica y el control de la contaminación para modificar las propiedades electrónicas.

Se proporcionó asimismo una pequeña introducción a las propiedades ópticas y magnéticas, donde se comentó el control de la polarización de la luz y la clasificación de los materiales magnéticos, identificando las propiedades prácticas de los materiales ferromagnéticos. La introducción a la tecnología permite integrar los conceptos anteriores y permite poner en perspectiva las ideas científicas al reconocer lo que se puede lograr con la tecnología, a fin de producir dispositivos electrónicos.

Ejercicios

II.1. ¿Con qué objetivo, en el modelo de Kronig-Penney, surge la necesidad de la periodicidad de la red?

- a) Para simplificar el potencial de la red.
- b) Para resolver la ecuación $H\psi = E\psi$.
- c) Porque se estudia un cristal.
- d) Para verificar que la energía está cuantizada.
- e) Para aplicar las condiciones de frontera a ψ .

II.2. Explique el comportamiento de un semiconductor en el cual la cantidad de impurezas donadoras N_D es igual a la cantidad de impurezas aceptadoras N_A . Realice la gráfica de la conductividad *versus* temperatura.

II.3. Explique el efecto que la luz con energía de 2 eV tiene sobre la conductividad en el aluminio, el silicio intrínseco y el cuarzo.

II.4.* ¿Cuál es la temperatura de saturación y la temperatura intrínseca para silicio con 10^{18} cm^{-3} impurezas de fósforo si $E_C - E_D = 0.012 \text{ eV}$?

II.5. Es posible que electrones energéticos sean emitidos por los materiales. ¿Puede un hueco ser emitido por un material?

II.6.* Es común que los elementos que tienen un número impar de electrones en sus niveles, al formar un sólido, sean buenos conductores, por ejemplo, los metales Na, Al, Au, Ag, etc. ¿Sucede lo mismo con el hidrógeno?

II.7. ¿Qué diferencia existe entre la velocidad térmica de los electrones $v_T = \sqrt{\frac{3kT}{m^*}}$ y la velocidad de deriva

$$v_D = \frac{qE\tau}{m^*}?$$

* Los ejercicios marcados con asterisco están resueltos al final del libro

II.8. ¿Cuántos niveles de energía hay en la banda de conducción en un gramo de silicio?

II.9. Explique la ley de acción de masas.

II.10. Un semiconductor intrínseco de silicio a temperatura ambiente tiene aproximadamente 10^{10} electrones/cm³ y 10^{10} huecos/cm³, mientras que un semiconductor tipo p con $N_A = 10^{18}$ tiene aproximadamente 10^{18} huecos/cm³ y 100 electrones/cm³. ¿Cómo se explica esta diferencia entre los huecos?

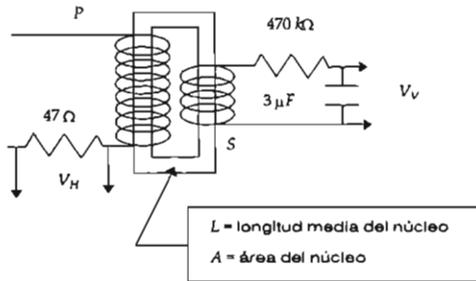
II.11.* Dibuje la densidad de portadores como función de la posición en un trozo de silicio en forma de cubo, cuando está siendo iluminado por una de sus caras con luz de energía de 1.12 eV. ¿Algo cambia si la radiación tiene una energía de 3 eV?

II.12. ¿Qué diferencia existe entre el material empleado para un imán y para un núcleo de transformador, y qué diferencia existe entre el material en un disco flexible de computadora de baja densidad y de alta densidad?

II.13. ¿Qué elementos pasivos podría integrar en el proceso NMOS? ¿Cuál sería la mejor manera de tener un resistor de 10 kΩ?

II.14. Si se tienen dos haces de luz polarizados en la misma dirección e interfieren, ¿qué diferencia de fase es necesaria para que no exista luz a la salida? ¿A dónde fue la energía de los haces incidentes?

II.15.* ¿Explique por qué el siguiente circuito, al mostrar en el osciloscopio V_V (en el canal y) y V_H (en el canal x) muestra un ciclo parecido a la histéresis magnética? ¿En qué condiciones las relaciones de voltaje y características magnéticas son válidas?



$$B = \frac{R_s C}{N_s A} V_V$$

$$H = \frac{N_p}{L R_p} V_H$$

CAPÍTULO III

ESTRUCTURAS ELECTRÓNICAS BÁSICAS

ESTA SECCIÓN TRATA DEL FUNCIONAMIENTO de algunas estructuras electrónicas básicas, los semiconductores heterogéneos, que forman los dispositivos elementales de la electrónica moderna. Nuestro objetivo es obtener modelos que permitan entender y describir su funcionamiento; por ejemplo, la unión entre un metal y un semiconductor permite entender el mecanismo que explica los contactos eléctricos a los semiconductores. La unión entre un semiconductor tipo n y otro tipo p permite explorar la rectificación de las señales, mientras que el capacitor formado por un metal, un óxido y un semiconductor (MOS) permite el estudio del efecto del campo eléctrico interno en un semiconductor.

Adicionalmente se describe el funcionamiento de otros dispositivos un poco más complejos, como los transistores BJT y MOSFET, así como el de dispositivos optoelectrónicos, LED, láser y fotodetectores.

Iniciemos el estudio para entender los procesos que permiten formar una unión entre dos materiales. Dos sistemas en contacto alcanzan el equilibrio cuando:

- $\Delta T = 0$ (al tener la misma temperatura, se alcanza el equilibrio térmico).
- $\Delta P = 0$ (al tener la misma presión, se alcanza el equilibrio mecánico).
- $\Delta E_f = 0$ (al tener un nivel de Fermi constante, se alcanza el equilibrio químico).

Las dos primeras condiciones anteriores son muy generales. Para explicar la tercera, imagine

dos semiconductores aislados iguales que sólo se diferencian por la cantidad de portadores libres que tienen, la cantidad de electrones en el semiconductor A (n_A), mayor que la cantidad de electrones en el semiconductor B (n_B) y, por consiguiente, la cantidad de huecos tiene la relación, $p_A < p_B$.

Esquemáticamente esto se representa así:

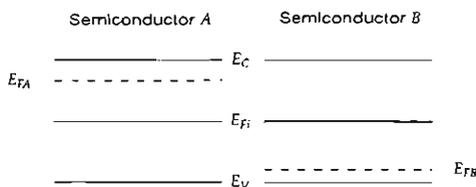


Figura III.1. Diagrama de bandas para dos semiconductores iguales, pero con diferente cantidad de impurezas que cambian la conductividad.

En este caso, un sistema formado por los dos semiconductores en contacto, se sabe que llega al equilibrio cuando el flujo de electrones del semiconductor A al B sea igual al flujo de electrones del semiconductor B al A .

$$J_{A \rightarrow B} = J_{B \rightarrow A}$$

Es importante subrayar que los semiconductores son iguales, es decir, que el mínimo de la banda de conducción en los dos semiconductores es $E_{CA} = E_{CB}$; al igual que el máximo de la banda de valencia $E_{VA} = E_{VB}$, cuando los semiconductores están aislados el uno del otro. Al formar

el contacto metalúrgico, como existen más electrones en el semiconductor *A* que en el *B*, debe existir un flujo de electrones de *A* hacia *B*. Pero por cada electrón que pasa del semiconductor *A* al *B* aumenta la energía potencial electrostática del semiconductor *B* (porque se ha aumentado el número de electrones) y disminuye la energía potencial electrostática del semiconductor *A* (porque se ha disminuido el número de electrones).

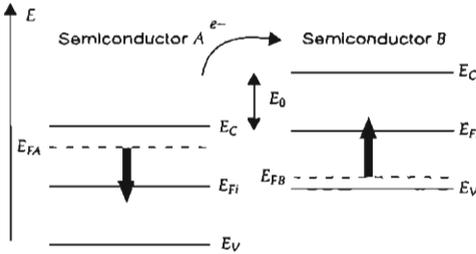


Figura III.2. Diagrama de bandas de energía en un estado intermedio antes de alcanzar el equilibrio.

En todo momento el flujo de electrones de $A \rightarrow B$ será:

$$j_{A \rightarrow B} \propto n_{0A} e^{-\frac{E_0}{kT}} \quad (III.1)$$

Esta ecuación indica que, de la cantidad total de electrones n_{0A} , solamente participan en la corriente eléctrica aquellos que tienen una energía mayor a E_0 (donde $E_0 = E_{CB} - E_{CA}$ y representa la energía que deben tener los electrones para pasar del semiconductor *A* al *B*). El flujo total de electrones que va del semiconductor *B* al *A* será:

$$j_{B \rightarrow A} \propto n_{0B} \quad (III.2)$$

Como los electrones del semiconductor *B* no tienen ninguna barrera que subir, todos los electrones pueden pasar de *B* a *A*. En el equilibrio, las densidades de corriente son iguales. Como las constantes de proporcionalidad son las mismas para el cálculo de n_{0A} , la densidad de estados efectiva N_{CA} es idéntica a N_{CB} , lo que permite concluir que:

Dos materiales *A* y *B* alcanzan el equilibrio cuando los niveles de Fermi se igualan.

$$E_{FA} = E_{FB} \quad (III.3)$$

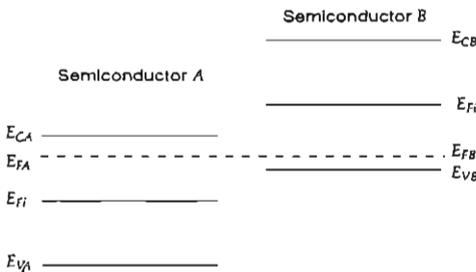


Figura III.3. Bandas de energía al alcanzar el equilibrio.

En el caso de que uno de los materiales sea un metal, es necesario introducir dos constantes que permitan definir los cambios del sistema de una manera tan precisa como antes. La función tra-

bajo en los metales ϕ_m , tiene esa propiedad, donde $q\phi_m$ es la energía necesaria para mover un electrón que está en el metal en el nivel de Fermi fuera de él, hacia el nivel de vacío del metal (próxima al metal), con energía cinética cero. Ésta es una constante en los metales y no tiene significado físico en los semiconductores, ya que en su nivel de Fermi no hay electrones.

La afinidad electrónica es igual a χ , donde $q\chi$ es la energía necesaria para mover un electrón desde el mínimo de la banda de conducción E_C y llevarla al nivel de vacío con energía cinética cero. Ésta es una constante en los semiconductores y no tiene sentido físico en los metales.

Con estas nuevas definiciones, la condición de equilibrio tiene la siguiente expresión:

Al unir dos materiales cualesquiera, existe un flujo de portadores. Este flujo se puede caracterizar de las siguientes dos maneras:

- Los electrones fluyen del material que tiene mayor nivel de Fermi al que tiene menor nivel de Fermi.
- Los electrones fluyen del material que tiene menor función de trabajo al material que tiene mayor función de trabajo.

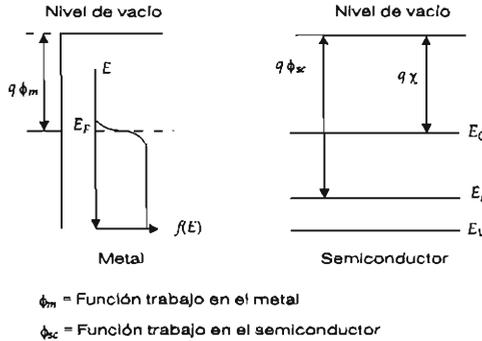


Figura III.4. Definición de función trabajo y afinidad electrónica.

DIODO DE UNIÓN

Uno de los dispositivos básicos en la electrónica es el que se forma al unir dos semiconductores, uno tipo *p* y el otro tipo *n* (unión *p-n*). El análisis de la unión *p-n* incluirá un estudio fenomenológico; después se presentará un análisis electrostático y finalmente un modelo de transporte. Cada uno de éstos se realiza con detalle por dos razones: es la primera vez que se presentan en este trabajo y el tratamiento es similar en todos los otros dispositivos.

El análisis fenomenológico requiere estudiar paso a paso lo que sucede con los electrones en los semiconductores tipo *n* y *p* al hacer la unión y llegar al equilibrio. Con objeto de que se logre la unión entre los semiconductores, ¿es suficiente tomar dos trozos de semiconductor diferentes y juntarlos o pegarlos con “Kola Loca”?

Veamos las diversas posibilidades. Si juntamos dos trozos de semiconductor nunca podremos lograr el contacto por varias razones: no

podemos garantizar que las superficies sean perfectamente lisas, además de que el silicio, por ejemplo, está cubierto por una película de óxido que hace que los dos trozos estén aislados. Ahora bien, si optamos por pegarlos, a los inconvenientes anteriores hay que agregar la interfaz con el adhesivo. Así que la única posible respuesta es emplear las técnicas de la microelectrónica, la difusión de impurezas, por ejemplo, y transformar un semiconductor tipo *n* en *p* y formar la unión en el proceso.

El equilibrio de la unión de los semiconductores se obtiene siguiendo el procedimiento que a continuación se describe:

- Al hacer la unión metalúrgica existe un flujo de electrones del semiconductor tipo *n* al *p* (y un flujo de huecos del semiconductor tipo *p* al *n*).
- Esto ocasiona que se rompa la neutralidad eléctrica de cada material y se tenga una carga neta positiva en el lado *n* y una carga neta negativa en el tipo *p*.

- Esto origina una variación de la energía potencial electrostática del semiconductor tipo p y del tipo n .
- Aumenta la energía potencial electrostática del semiconductor tipo p y disminuye la energía potencial electrostática del semiconductor tipo n . Esto sucede por la entrada de electrones al semiconductor tipo p y la salida de electrones del semiconductor tipo n .
- Aparece una diferencia de potencial entre los dos materiales.
- Aparece un campo eléctrico entre los dos materiales que se opone al movimiento de los portadores de carga.
- Se detiene el flujo neto de las partículas cuando se igualan los dos niveles de Fermi. La fuerza que mueve a los electrones del material n al p se iguala a la fuerza del campo eléctrico interno que se produce por el movimiento de las cargas.
- En adelante, por cada electrón que va del semiconductor tipo n al tipo p hay otro electrón que va del tipo p al tipo n .

El diagrama de energía que se obtiene es el de la figura III.3, donde nos falta aclarar varias dudas:

- ¿Cómo cambia el mínimo de la banda de conducción del semiconductor tipo n al tipo p y como cambia el máximo de la banda de valencia?
- El diagrama anterior es de energía potencial electrostática E , contra posición x . ¿Qué orden de magnitud en x cubre la transición del semiconductor tipo n al tipo p ?
- ¿Cuál región es la que produce el efecto rectificante y cuál es sólo accesoria?

El análisis electrostático dará respuestas a estas preguntas. Empecemos por calcular una expresión para la diferencia entre los mínimos de la banda de conducción de los materiales, $E_{c_p} - E_{c_n} = E_{bi}$. Esta expresión se obtiene de las

expresiones para la cantidad de electrones en equilibrio en los materiales, n_0 .¹

$$E_{bi} = kT \ln \frac{n_{i0}}{n_{p0}} = kT \ln \frac{N_D N_A}{n_i^2} \quad (III.4)$$

Es necesario determinar varios parámetros que caracterizan a la unión $p-n$ y por añadidura todos los dispositivos, además de terminar el diagrama de energía potencial electrostática contra posición.

Para simplificar el problema: suponga una unión $p-n$ abrupta, esto es, cuando la concentración de impurezas es uniforme en los semiconductores y súbitamente cambia de ser tipo n a tipo p .

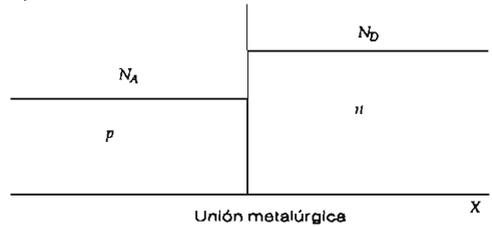


Figura III.5. Unión $p-n$ abrupta.

Además, emplearemos la aproximación de la "zona desértica". Esto es, imaginaremos que los electrones, al pasar del semiconductor tipo n al tipo p , salen ordenadamente en forma de "placas de carga (ρ) negativa" que al llegar al semiconductor tipo p ocupan los enlaces covalentes no saturados también de manera ordenada.

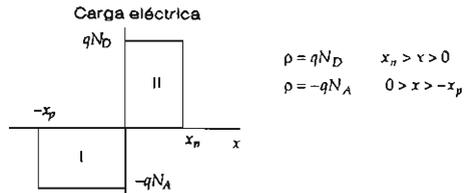


Figura III.6 Aproximación de la zona desértica en la unión $p-n$.

¹ Los subíndices empezarán a aparecer cada vez con mayor frecuencia. 0 siempre indicará una propiedad de equilibrio. Los subíndices p y n se refieren al semiconductor tipo p y tipo n respectivamente; el subíndice bi indica el potencial interno o de contacto (*built-in*). De aquí en adelante será necesario agregar el subíndice n o p (de la región), puesto que podemos tener electrones en el semiconductor tipo n (n_n) o p (n_p).

La región I en el semiconductor tipo p no es eléctricamente neutra: tiene carga negativa por los electrones que emigraron del semiconductor tipo n , para igualar los niveles de Fermi.

La región II en el semiconductor tipo n no es eléctricamente neutra: tiene carga positiva ya que los electrones que la hacían neutra se fueron al semiconductor tipo p .

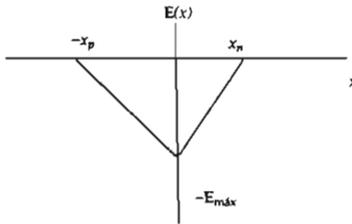
En esta aproximación, la carga de las dos regiones es igual y se puede escribir de la siguiente manera:

$$N_D x_n = N_A x_p \quad (III.5)$$

Por la poca cantidad de los portadores móviles que existen entre $-x_p$ y x_n , se le llama "zona desértica".

Al hacer el análisis electrostático del efecto de esta distribución de carga interna con la ecuación de Poisson se encuentran expresiones para el campo eléctrico en la unión $p-n$, el potencial eléctrico y la energía potencial electrostática.

$$\frac{\partial E}{\partial x} = \frac{\rho}{\epsilon \epsilon_0}; E = -\frac{\partial V}{\partial x}$$



$$E(x) = \frac{qN_D}{\epsilon \epsilon_0} (x - x_n) \quad x_n > x > 0$$

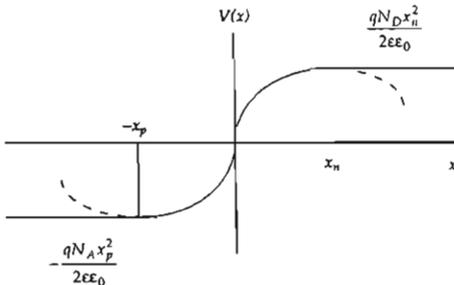
$$E(x) = \frac{qN_A}{\epsilon \epsilon_0} (-x - x_p) \quad 0 > x > -x_p$$

Figura III.7. Campo eléctrico interno en una unión $p-n$.

El campo eléctrico E es cero fuera de la zona desértica y tiene un valor máximo en la unión metalúrgica E_{max} .

El potencial eléctrico siempre es relativo a un

punto. Por comodidad, se puede hacer que el potencial sea cero en el origen $V(x=0) = 0$. Los resultados son idénticos con cualquier otro punto de referencia.



$$V(x) = -\frac{qN_D}{\epsilon \epsilon_0} \left(\frac{x^2}{2} - x_n x \right) \quad x_n > x > 0$$

$$V(x) = +\frac{qN_A}{\epsilon \epsilon_0} \left(\frac{x^2}{2} + x_p x \right) \quad 0 > x > -x_p$$

Figura III.8. Variación de potencial en las bandas por la formación de la unión $p-n$.

Para encontrar la gráfica de la energía potencial electrostática sólo tenemos que multiplicar por el negativo de la carga eléctrica, lo que invierte la figura anterior, $E(x) = -qV(x)$. La energía del potencial interno es igual a la diferencia de energía potencial fuera de la zona desértica. Esta

identificación permite relacionar el potencial interno V_{ni} con el ancho de la zona desértica $W(W = x_n + x_p)$. Esto da como resultado una serie de relaciones para los parámetros electrostáticos con las características de fabricación.

En equilibrio, el ancho de la zona desértica W

y el campo eléctrico interno máximo E_{max} tienen las siguientes expresiones:

$$E_{max} = - \left(\frac{2qV_b N_D N_A}{\epsilon \epsilon_0 (N_D + N_A)} \right)^{\frac{1}{2}}$$

$$W = \left(\frac{2V_b \epsilon \epsilon_0 (N_D + N_A)}{q N_D N_A} \right)^{\frac{1}{2}} \quad (III.6)$$

De esta manera, el diagrama de energía potencial electrostática E contra posición x , de la unión $p-n$ será:

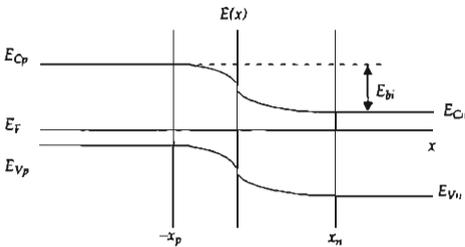


Figura III.9. Diagrama de la energía potencial electrostática de la unión $p-n$.

Veamos con detalle el efecto de la polarización:

El acceso eléctrico a cualquier material se realiza a través del nivel de Fermi, ya que todo sistema alcanza el equilibrio cuando tiene el mismo nivel. Las diferencias de potencial eléctrico

son proporcionales a la diferencia de los niveles de Fermi.

Un semiconductor tipo n homogéneo como el que forma un resistor normal puede representarse con los diagramas de energía de la figura III.10.

En ella se representa una variación lineal del potencial aplicado a lo largo del resistor, ya que el campo eléctrico es uniforme. Es importante observar la relación que guarda el potencial eléctrico V con la energía potencial electrostática E . Si la energía está en eV, ambos son numéricamente iguales pero de signo contrario. La diferencia de potencial aplicado es idéntica a la diferencia de los niveles de Fermi.

¿Cómo cambia el diagrama de energía potencial electrostática de una unión $p-n$ cuando se aplica una diferencia de potencial en sus extremos?

En la figura III.11 se observa que:

- El ancho de la zona desértica (W) cambia, en polarización directa disminuye y en polarización inversa aumenta.
- El nivel de Fermi es una constante y es horizontal cuando $V_0 = 0$. Además, es continuo.
- El nivel de Fermi es discontinuo al polarizar la unión con V_0 diferente de cero.
- El nivel de Fermi fuera de W (aproximadamente) sigue siendo continuo y horizontal.
- En W (aproximadamente) no se dibujó el nivel de Fermi, porque éste es un concepto de sistemas en equilibrio y al pasar corrientes

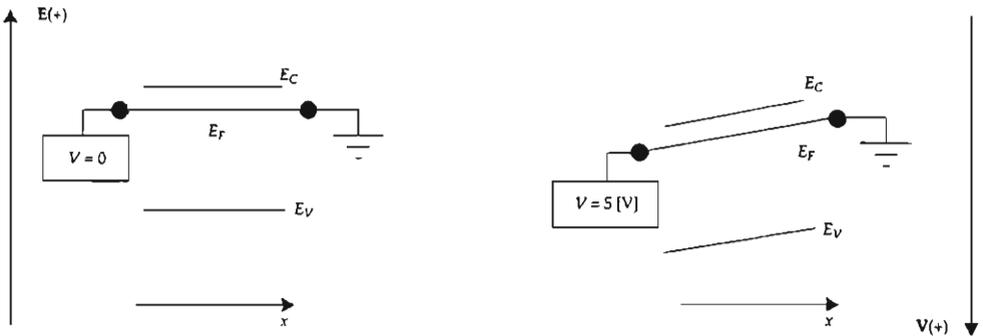


Figura III.10. Cambios en el diagrama de energía por efecto de la polarización.

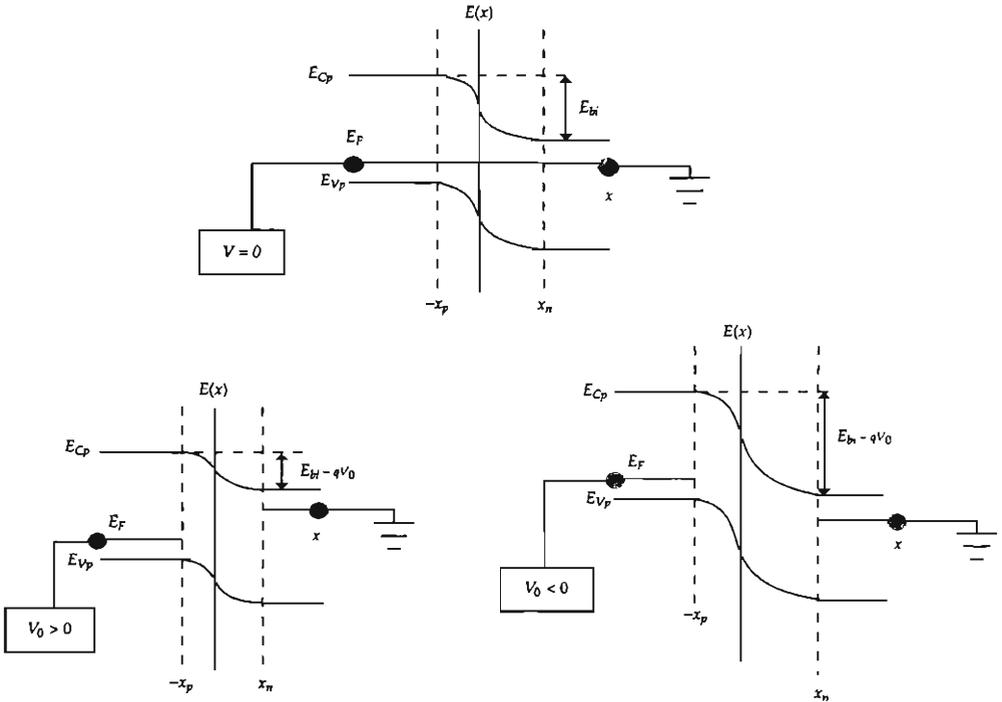


Figura III.11. Diagrama de energía potencial electrostática contra posición de una unión $p-n$. En equilibrio ($V = 0$), en directa ($V_0 > 0$) y en inversa ($V_0 < 0$).

te por la unión el sistema está fuera de equilibrio.

- Todo el potencial aplicado aparece en W .
¿Por qué “cae” todo el potencial en W ?

De manera preliminar, se puede decir que la zona desértica es mucho más resistiva que las zonas neutras. Es como tener tres resistores en serie, y uno de ellos de muchos órdenes de magnitud más resistivo. Para todos los fines prácticos, el potencial aplicado aparecerá en el resistor de mayor resistencia. Por eso los niveles de Fermi se separan en W y no se modifican o inclinan en las regiones neutras.

Esta afirmación tiene una validez limitada si se polariza la unión en polarización directa. Con un voltaje suficientemente grande, la resistividad de la zona desértica disminuye a tal grado que en algún momento la suposición deja de ser

válida. Para entender sus limitaciones, es necesario analizar el mecanismo de conducción en la unión al ser polarizada.

El concepto de seudoniveles de Fermi es particularmente útil al analizar la unión $p-n$ fuera de equilibrio. El diagrama de la figura III.12 representa una unión $p-n$ polarizada en inversa, donde se emplea el concepto de seudoniveles de Fermi.

Al dividir al semiconductor en cinco regiones como aparece en la figura, podemos hacer las siguientes observaciones:

La región I es la región neutra del semiconductor tipo p . Esta región se llama así porque no siente la presencia de la unión $p-n$. De hecho, es una zona parásita porque no es eléctricamente activa. En esta zona se cumplen todas las propiedades de los semiconductores homogéneos en equilibrio.

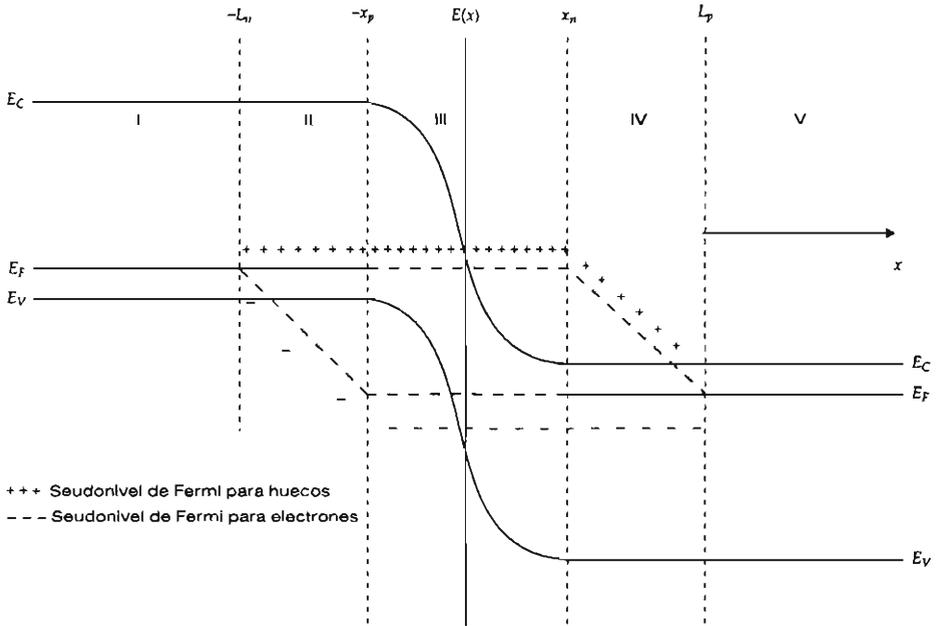


Figura III.12. Unión p-n en polarización inversa con los pseudoniveles de Fermi para los huecos (+) y los electrones (-).

La región II es la región de semiconductor tipo p, que ve perturbada su concentración de electrones por la polarización de la unión p-n. En polarización inversa pasan electrones del semiconductor tipo n y los huecos del semiconductor p no sienten el efecto de la polarización, porque los huecos que pasan del lado n son insignificantes comparados con los que aquí existen en equilibrio. En la polarización directa esta región es perturbada por los electrones que llegan del semiconductor tipo n. Así que todos los cambios suceden por la variación de los portadores minoritarios, los electrones, y sólo existen cuando se polariza la unión p-n.

La región III es la región que corresponde a la zona desértica, que por efecto de la polarización tiene una cantidad de portadores diferente. El cálculo de la densidad de portadores se debe realizar empleando los dos pseudoniveles de Fermi, como en las regiones II y IV.

La región IV es la región del semiconductor tipo n. Si bien es cierto que es similar a la región

II, en esta región todos los cambios suceden en los huecos.

La región V es la región neutra del semiconductor tipo n y es similar a la región I. Las regiones II, III y IV están fuera de equilibrio, mientras que las regiones I y V están en equilibrio.

El análisis para entender el transporte se puede realizar después de identificar cuántos portadores existen en las fronteras de las zonas anteriores, como se muestra en el cuadro (p. 79).

En este cuadro se han empleado dos condiciones. Una de ellas es la condición de equilibrio, en donde por cada electrón que pasa la zona desértica en una dirección existe otro que viaja en dirección contraria.

$$n_{i0} \exp\left(-\frac{E_{hi}}{kT}\right) = n_{p0}$$

$$p_{i0} \exp\left(-\frac{E_{ln}}{kT}\right) = p_{n0} \tag{III.7}$$

*	Densidad de electrones	Densidad de huecos
Zona 1 $x \leq -L_n$	$n_{p0} = N_C \exp\left(-\frac{E_{Cp} - E_{Fp}}{kT}\right)$	$p_{p0} = \frac{n_i^2}{n_{p0}} \approx N_A$
$x = -x_p$	$n = n_{i0} \exp\left(-\frac{q(V_{bi} - V_0)}{kT}\right) = n_{p0} \exp\left(\frac{qV_0}{kT}\right)$	$p = p_{i0}$
$x = x_n$	$n = n_{i0}$	$p = p_{i0} \exp\left(-\frac{q(V_{bi} - V_0)}{kT}\right) = p_{n0} \exp\left(\frac{qV_0}{kT}\right)$
Zona V $x \geq L_p$	$n_{n0} = N_C \exp\left(-\frac{E_{Cn} - E_{Fn}}{kT}\right) \approx N_D$	$p_{n0} = \frac{n_i^2}{n_{n0}}$

* Este cuadro incluye las condiciones de frontera generales para ambas polarizaciones. V_0 es positivo para polarización en directa y negativo para polarización en inversa.

Donde E_{bi} es la barrera de potencial que ven los electrones que están en el semiconductor tipo n y quieren pasar al semiconductor tipo p . La otra condición se refiere a la convención usada para indicar polarización inversa y directa, donde V_0 es negativa en el primer caso y positiva en el segundo, el cual seguiremos usando a lo largo del capítulo.

En polarización directa existe un flujo de electrones en exceso del material n al p y un flujo de huecos en exceso del material p al n . La corriente aumenta rápidamente al aumentar el voltaje.

En polarización inversa existe un flujo de electrones en exceso del material p al n y un flujo de huecos en exceso del material n al p . La corriente casi no aumenta al aumentar el voltaje, siempre que no existan procesos de ruptura.

Modelo de Shockley

Este modelo describe analíticamente el transporte de carga a través de la unión p - n fuera de equilibrio, y se basa en la hipótesis de que las ecuaciones de transporte son válidas.

$$J_n = q\mu_n n E + qD_n \nabla n \quad \text{y} \quad J_p = q\mu_p p E - qD_p \nabla p$$

$$\frac{\partial n}{\partial t} = G_n - R_n + \frac{1}{q} \nabla \cdot J_n \quad \text{y} \quad \frac{\partial p}{\partial t} = G_p - R_p + \frac{1}{q} \nabla \cdot J_p \quad (\text{III.8})$$

Como se recordará, estas ecuaciones fueron introducidas al final del capítulo anterior.

La solución de estas ecuaciones se simplifica en gran medida con las siguientes aproximaciones:

- Todo el potencial aplicado aparece en la zona desértica (W). El campo eléctrico es cero fuera de ella.
- Las regiones n y p son infinitas; en general, es suficiente con que sean algunas veces mayores que la longitud de difusión de los portadores minoritarios (L_n y L_p).
- Es importante tener un bajo nivel de inyección; que la cantidad de portadores minoritarios en exceso sea mucho menor que la de los portadores mayoritarios en equilibrio. También se puede decir que en directa $V_0 < V_{bi}$ y que en inversa no exista ruptura.
- Que los semiconductores sean no degenerados.
- Que no exista recombinación en W .

Si no existe campo eléctrico en las zonas cuasineutras, ¿cómo se produce entonces la corriente eléctrica?

Para dar respuesta a esta pregunta, veamos con detenimiento la distribución de los portadores y la resistividad a lo largo de la unión p - n .

Cualitativamente, con ayuda del diagrama de la figura III.13 se puede ver que la densidad de portadores minoritarios cambia cerca de la

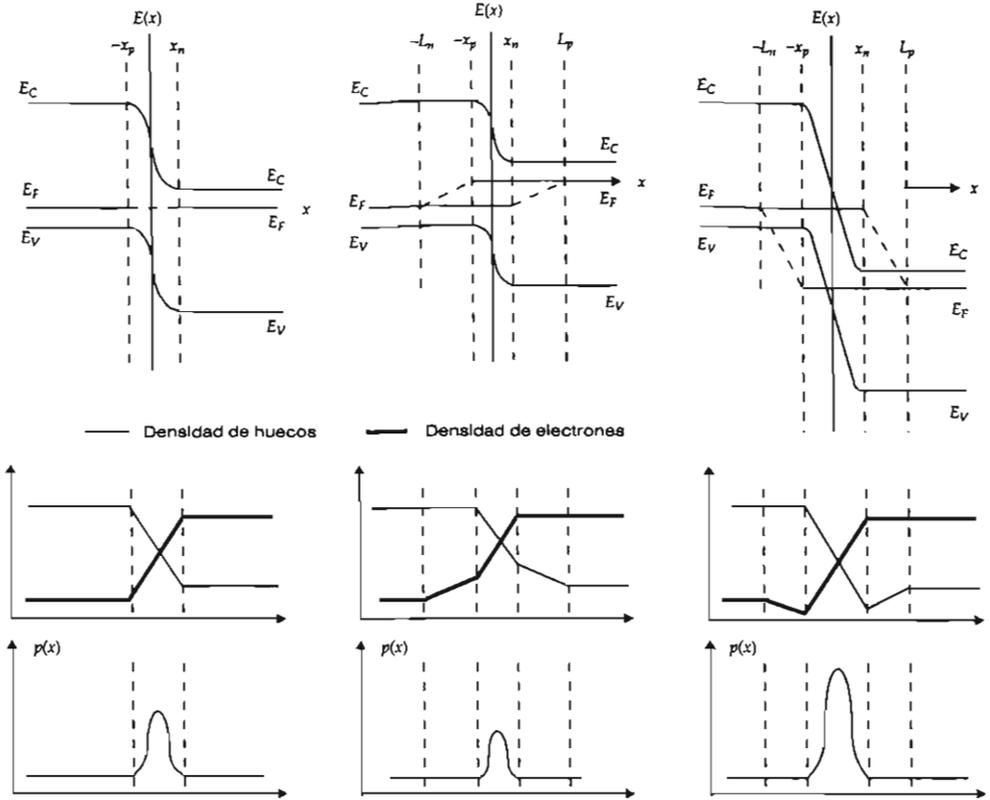


Figura III.13. Diagrama de la densidad de portadores en la unión p-n al polarizar.

zona desértica cuando se polariza la unión: los portadores minoritarios aumentan en directa y disminuyen en inversa. También es claro que la corriente es constante a lo largo del dispositivo y que no existen campos eléctricos fuera de la zona desértica, de donde se concluye que la corriente es controlada únicamente por los portadores minoritarios.

Los portadores mayoritarios no contribuyen a la corriente eléctrica porque son muchos, no se cambia su número localmente (no existe difusión de mayoritarios) y no existe campo eléctrico que los acelere (no existe deriva). Los portadores minoritarios tampoco son acelerados por el campo eléctrico, pero al presentar una variación local en

su número cerca de la zona desértica provocan difusión y controlan la corriente del dispositivo. Cuanto mayor sea el gradiente de concentración de portadores minoritarios, mayor será la corriente eléctrica que pasa por el dispositivo. El gradiente se puede controlar mediante dos parámetros:

El gradiente de concentración es mayor cuanto mayor sea la densidad de portadores en la frontera de la zona desértica, controlado únicamente por la polarización. El gradiente de concentración aumenta de acuerdo con la rapidez de recuperación de la concentración de portadores de equilibrio, lo que está determinado por el proceso de generación y recombinación de portadores minoritarios.

Ahora plantearemos todo lo anterior cuantitativamente.

Considere inicialmente sólo al semiconductor tipo n , en donde los huecos son minoritarios. La zona desértica no contribuye a la corriente externa, ya que la corriente producida por el campo eléctrico interno es contrarrestada por la corriente de difusión. En la región cuasineutra del semiconductor ($x \geq x_n$), la ecuación de continuidad toma la siguiente forma:

$$D_p \frac{\partial^2 (p_n - p_{n0})}{\partial x^2} - \frac{p_n - p_{n0}}{\tau_p} = 0 \quad (\text{III.9})$$

Dado que el campo eléctrico, E , es cero, y la generación, G , es cero, se impone una condición de estado estable $\frac{\partial \Delta p}{\partial t} = 0$ y además no existen cambios locales del campo eléctrico.

La solución general de esta ecuación diferencial tiene la forma típica siguiente:

$$p_n - p_{n0} = A \exp\left(\frac{x}{L_p}\right) + B \exp\left(-\frac{x}{L_p}\right) \quad (\text{III.10})$$

donde la longitud de difusión de los portadores minoritarios, en este caso huecos, L_p , tiene la expresión $\sqrt{D_p \tau_p}$.

La solución particular se obtiene al aplicar las condiciones a la frontera, en $x = x_n$; y cuando $x \rightarrow \infty$, la densidad de portadores minoritarios en esos puntos está indicada en la tabla anterior y se obtiene:

$$p_n(x) - p_{n0} = p_{n0} \exp\left(\frac{x_n - x}{L_p}\right) \left[\exp\left(\frac{qV_0}{kT}\right) - 1 \right] \quad (\text{III.11})$$

La densidad de corriente de huecos en el lado n se obtiene por medio de la ecuación de densidad de corriente adecuada, donde sólo existe corriente de difusión ya que el campo eléctrico es cero. Esta corriente de difusión depende de la posición, ya que la recombinación o generación tratan de producir el equilibrio en la densidad de portadores. La contribución de la corriente realizada por los huecos se calcula en el punto en el que adquiere el máximo valor, esto es en $x = x_n$. Así, tenemos que:

$$J_p(x) \Big|_{x=x_n} = J_p(V_0) = \frac{qD_p}{L_p} p_{n0} \left[\exp\left(\frac{qV_0}{kT}\right) - 1 \right] \quad (\text{III.12})$$

Con el resultado análogo para los electrones en el lado p , se obtiene la densidad de corriente total:

$$J(V_0) = q \left[\frac{D_n n_i^2}{L_n N_A} + \frac{D_p n_i^2}{L_p N_D} \right] \left[\exp\left(\frac{qV_0}{kT}\right) - 1 \right] \quad (\text{III.13})$$

En la figura III:14 se visualiza la contribución de la corriente eléctrica realizada por los portadores minoritarios en inversa y directa; trate de relacionarlo con la ecuación anterior.

Existe una serie de factores que modifican el comportamiento ideal de la unión $p-n$. Algunos son geométricos, como el que se produce al tener la superficie del material muy cerca de la zona desértica, que modifica la recombinación de los portadores minoritarios, como es el caso de las celdas solares. Otros factores son tecnológicos, como los que resultan en un resistor parásito en paralelo con la unión $p-n$, debido a la imposibilidad de eliminar los efectos superficiales de la

unión (por pasivarla), o en el resistor en serie por la unión que se produce de las zonas que no forman el dispositivo propiamente dicho (las regiones cuasineutras y los contactos óhmicos semiconductor-metal). Un último factor que modifica el comportamiento ideal es la suposición de la ausencia de recombinación en la zona desértica de la unión: en algunos casos existe recombinación accidental o intencional, que aumenta la corriente que circula por el dispositivo a un voltaje dado. Este comportamiento es no lineal y se parece al del modelo anterior.

Los efectos reales modifican la característica ideal de la unión $p-n$ de la siguiente manera:

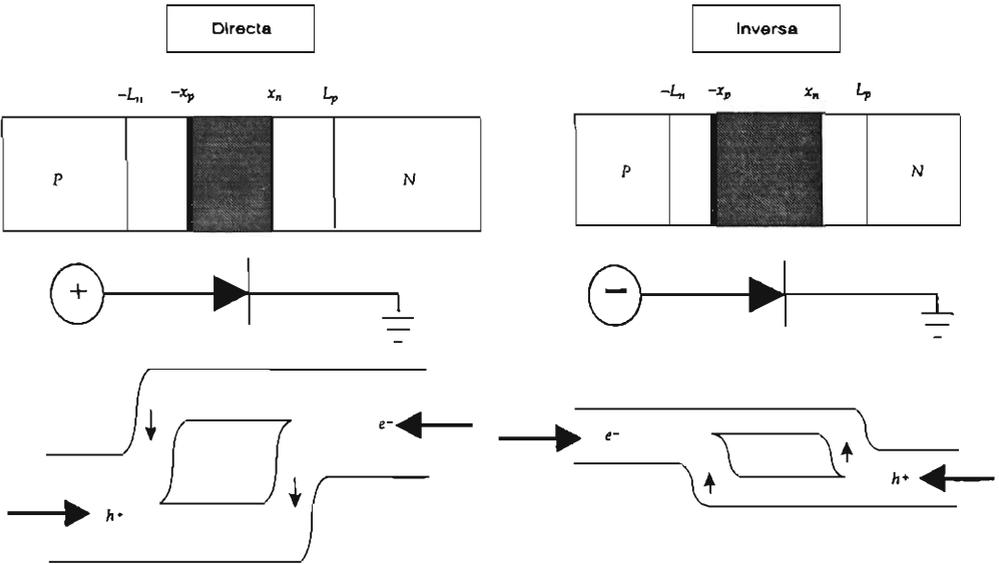


Figura III.14. Representación de la corriente eléctrica en una unión $p-n$ polarizada, en directa está dominada por la recombinación de los electrones y en inversa por la generación de los electrones.

- Aumento en la corriente a bajos niveles de voltaje en directa por la corriente de recombinación y la resistencia en paralelo.
- Disminución en la corriente a altos niveles de voltaje en directa por el alto nivel de inyección.
- Disminución en la corriente a altos niveles de voltaje en directa por la resistencia en serie.
- Aumento en la corriente a bajos niveles de voltaje en inversa por la recombinación y los efectos de superficie.
- Aumento en la corriente a variados niveles de voltaje en inversa debido a la ruptura de la unión.

Capacitancia de la zona desértica

La capacitancia por unidad de área se define como:

$$C = \frac{dQ}{dV} \tag{III.14}$$

Esta ecuación cuantifica el cambio de la carga en la zona desértica por unidad de área (Q) conforme cambia el voltaje aplicado (V).

Para una unión $p-n$, bajo la suposición de una unión abrupta, se tiene:

$$C = \frac{d(qNW)}{d\left(\frac{qNW^2}{2\epsilon\epsilon_0}\right)} = \frac{\epsilon\epsilon_0}{W} = \sqrt{\frac{q\epsilon\epsilon_0(N_A N_D)}{2(N_A + N_D)(V_{bi} - V_0)}} \tag{III.15}$$

Esta ecuación presenta dos resultados interesantes: primero, la capacitancia de una unión $p-n$ no es constante y disminuye conforme la polarización de la unión en inversa aumenta y, segundo, permite contar con un modo experimental de medir la diferencia de las funciones de trabajo de la unión (V_{bi}), midiendo la capacitancia de la unión $p-n$ como función del voltaje inverso. En directa no es usual, ya que la corriente que circula por la unión la puede destruir.

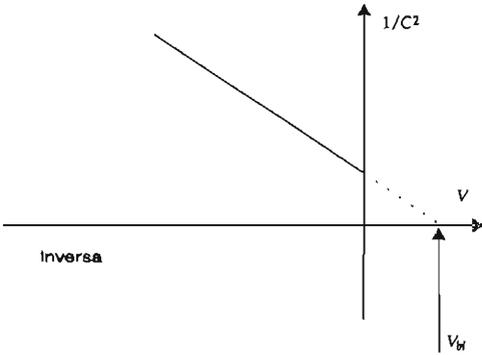


Figura III.15. Curva de capacitancia de la unión p-n como función de la polarización.

Ruptura de la unión p-n

Los principales mecanismos de ruptura en la unión p-n son los siguientes:

- a) Ruptura térmica (inestabilidad térmica).
- b) Ruptura túnel (Zenner).
- c) Ruptura por avalancha.

a) En semiconductores con ancho de banda pequeño, como el germanio, la corriente inversa puede llegar a calentar la unión. Esto ocasiona que la corriente de equilibrio sea mayor, lo que produce un posterior calentamiento de la unión y así sucesivamente. De tal manera que se forma un ciclo, aumento de la temperatura = aumento de la corriente. Si no se controla la temperatura, este proceso puede ser destructivo. La corriente en directa o inversa puede destruir a la unión por el efecto térmico de Joule.

b) En uniones p-n altamente impurificadas puede presentarse este mecanismo, que no es destructivo, al menos no en la unión, ya que las transiciones son a energía constante. Los electrones del semiconductor tipo p por efecto túnel pasan a la zona desértica y aparecen en el semiconductor tipo n.

Una imagen análoga se obtiene al pensar en un sistema formado por dos cajas de potencial separadas por un barrera de ancho W y altura H ; el proceso es más probable cuando W y H son

pequeños. Como W es función de la cantidad de portadores, si esta cantidad aumenta W se hace más pequeña y es probable el efecto túnel. La muestra no se calienta y puede suceder en directa y en inversa.

Si el voltaje de ruptura en inversa es menor a $4E_g/q$, entonces el proceso de ruptura es por túnel.

c) Este proceso se debe a la generación de pares electrón-hueco por impacto ionizante. Los parámetros necesarios para que ocurra son un alto campo eléctrico interno y una pequeña energía de ionización de los átomos para obtener pares electrón-hueco.

De tal manera, el primer electrón se multiplica y esto produce un aumento abrupto en la corriente eléctrica.

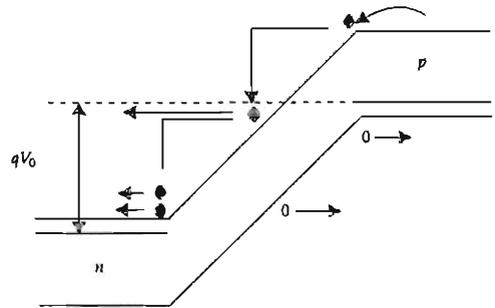


Figura III.16. Proceso de multiplicación de portadores por avalancha.

Transitorios en la unión p-n

Veamos qué sucede con la concentración de portadores al conmutar rápidamente la conducción en directa a la conducción en inversa. Este comentario cualitativo se sustenta en la solución de la ecuación de continuidad sin ignorar la dependencia en relación con el tiempo, como se hizo en el análisis del estado estable.

Se obtendría la siguiente gráfica si se dibuja la corriente eléctrica contra el tiempo, al cambiar rápidamente la polarización de directa a inversa. Cabe notar que todo sistema capacitivo tiene almacenamiento de carga y, de acuerdo con la resistividad del circuito, presenta transitorios. Lo

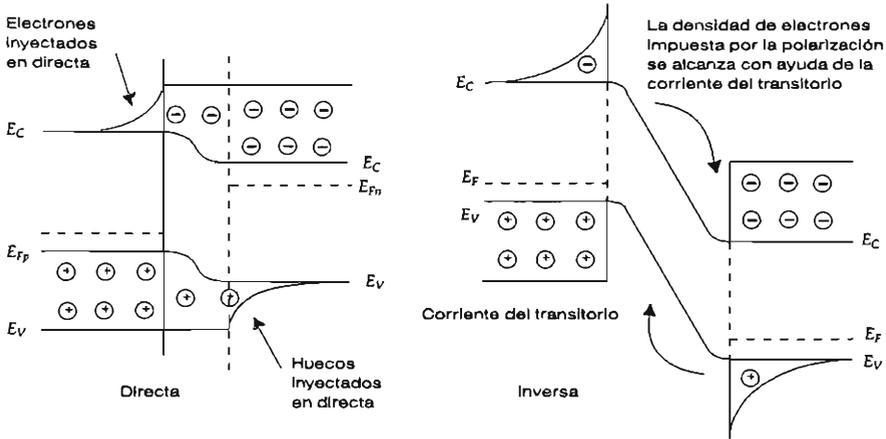


Figura III.17. El transitorio se forma al conmutar de directa a inversa por los minoritarios que antes de recombinar prefieren regresar al semiconductor de donde provenían.

que planteamos aquí es adicional y único en los dispositivos semiconductores, que basan su funcionamiento en la difusión de los portadores minoritarios.

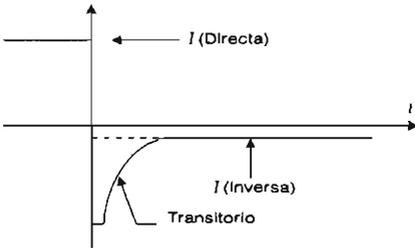


Figura III.18. El aumento de corriente en inversa se debe en gran parte a los portadores minoritarios.

Este transitorio se debe a que algunos de los portadores minoritarios que se inyectaron y que se encuentran cerca de la zona desértica tienden a regresar al lugar de donde partieron (región *n*), por el enorme gradiente que tienen ahora (este gradiente se tiene por la nueva polarización), y la corriente transitoria dura todo el tiempo que tardan en desaparecer estos portadores.

Si en el semiconductor existe una gran cantidad de trampas (niveles que favorezcan el des-

censo de los portadores), como las que provocan la corriente de recombinación en la zona desértica, entonces los transitorios serán menos importantes.

Si tenemos en cuenta este parámetro, se puede hacer la siguiente clasificación:

Los diodos de Si normales pueden conmutar de directa a inversa en 50 ns.

Los diodos rápidos de Si (con centros de recombinación) pueden conmutar de directa a inversa de 1 a 5 ns.

Los diodos GaAs (por tener tiempo de vida muy corto) pueden conmutar de directa a inversa en 0.1 ns o menos.

Las uniones metal-semiconductor (cuando conducen por mayoritarios) pueden conmutar en tiempos aun menores.

CONTACTOS ÓHMICOS Y RECTIFICANTES

Esta sección trata de la formación de efectos óhmicos (comportamiento lineal entre voltaje y corriente) y rectificantes (comportamiento no lineal entre voltaje y corriente), al unir un metal y un semiconductor. Es claro que la unión de dos metales usualmente no causa problemas eléctricos especiales, excepto en los casos de formación

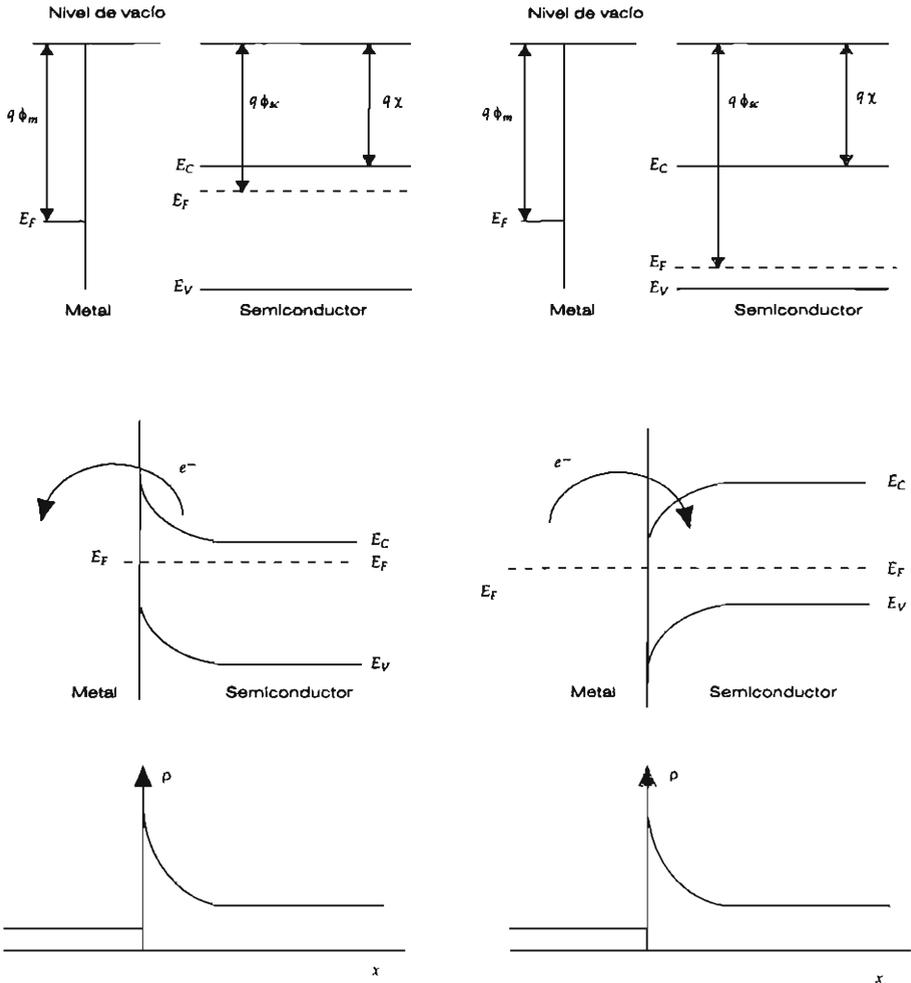


Figura III.19. Diagramas de energía potencial electrostática antes y después de la formación de la unión y diagrama de resistividad a lo largo de la unión. Primeros dos casos.

de termopares o cuando se intentan medir señales extremadamente débiles, donde la conexión entre los alambres se vuelve un asunto de especial cuidado. El caso de la unión de dos semiconductores fue descrito en la sección anterior. El estudio de la unión entre un metal y un semiconductor es de gran importancia, ya que permite entender cómo se realizan los accesos eléctricos

a los semiconductores, además de ser la parte esencial en los dispositivos ultrarrápidos llamados MESFET (transistor de efecto de campo por medio de la estructura metal-semiconductor).

Los metales se caracterizan de manera energética con el conocimiento de la función trabajo ϕ_m . Ésta es una constante y difícilmente se modifica en virtud de la gran cantidad de electrones

que tienen los metales, mientras que los semiconductores se caracterizan por la afinidad electrónica χ , una constante que no se modifica ya que está definida por distancias entre las bandas de energía. En los semiconductores es posible visualizar la función trabajo ϕ_{sc} , pero es una variable, ya que la cantidad de electrones puede variar localmente.

Al hacer la unión metal-semiconductor se producen los mismos fenómenos que durante la unión $p-n$, sólo que aquí, para identificar de qué material fluyen los electrones, necesitamos conocer la relación entre las funciones de trabajo antes de hacer la unión, si recordamos en la unión $p-n$ dio origen a V_{bi} :

El sistema de la unión entre un metal y un semiconductor puede formarse mediante la siguiente combinación de materiales:

- a) $\phi_m > \phi_{sc}$ (semiconductor tipo n): dará lugar a una unión rectificante.
- b) $\phi_m < \phi_{sc}$ (semiconductor tipo p): dará lugar a una unión rectificante.
- c) $\phi_m < \phi_{sc}$ (semiconductor tipo n): dará lugar a una unión óhmica.
- d) $\phi_m > \phi_{sc}$ (semiconductor tipo p): dará lugar a una unión óhmica.

Los valores de la función trabajo para algunos materiales metálicos son: platino $q\phi_m = 5.8$ eV, oro $q\phi_m = 5.2$ eV, cromo $q\phi_m = 4.6$ eV, aluminio $q\phi_m = 4.1$ eV. La afinidad electrónica para el silicio $q\chi = 4.05$ eV; por lo tanto, puede producir, con el cambio en la cantidad de impurezas, valores de la función trabajo en el siguiente intervalo $q\phi_{sc} = 4.05$ eV (Si n^+) hasta 5.20 eV (Si p^+).

Diagramas de energía en los contactos rectificantes

De los diagramas de la figura III.19 es posible concluir que las estructuras presentan un efecto no lineal entre el voltaje y la corriente, ya que las curvas de resistividad se parecen a las de la unión $p-n$. Las similitudes con la unión $p-n$ son limitadas. El análisis electrostático es válido, suponiendo solamente que el metal es un semicon-

ductor muy impurificado, de tal manera que la zona desértica existe sólo en el semiconductor, al igual que la densidad de carga, el campo eléctrico y la energía potencial electrostática. Y más aún, las expresiones son iguales eliminando las impurezas que corresponden al metal; por ejemplo:

$$x^2(m - sc) = \lim_{N_D \rightarrow \infty} W^2(p - n) \quad (III.16)$$

La estructura rectificante formada por un metal y un semiconductor con las características $\phi_m > \phi_{sc}$ (semiconductor tipo n), al hacer el análisis electrostático, daría:

$$x_n^2(n - sc) = \frac{V_{bi} 2\epsilon\epsilon_0}{qN_D} \quad (III.17)$$

La diferencia esencial entre la unión $p-n$ y la unión metal-semiconductor es que la conducción se lleva a cabo por portadores mayoritarios (a diferencia de la unión $p-n$, en que la conducción es por minoritarios). Veamos cualitativamente el mecanismo de transporte en una unión rectificante de este tipo:

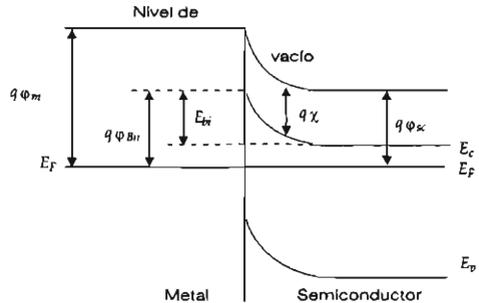


Figura III.20. Diagrama de energía potencial electrostática de una unión metal-semiconductor rectificante.

Al polarizar el sistema, la gran resistividad de la unión provoca que toda la caída de potencial sea en la unión; esto no modifica el valor de $\phi_{b,m}$, que está definido por dos constantes, la función trabajo del metal y la afinidad electrónica del semiconductor ($\phi_{b,m} = \phi_m - \chi$). Al no cambiar esta

barrera de potencial, los electrones que van del metal al semiconductor no cambian con la polarización. La barrera de energía que perciben los electrones del semiconductor y quieren pasar al metal en equilibrio es E_{bi} ($E_{bi} = q\phi_m - q\phi_s$), que

aumenta al polarizar la unión en inversa (voltaje negativo en el metal respecto al semiconductor) y disminuye al polarizar la unión en directa, $E_{bi} - qV_0$.

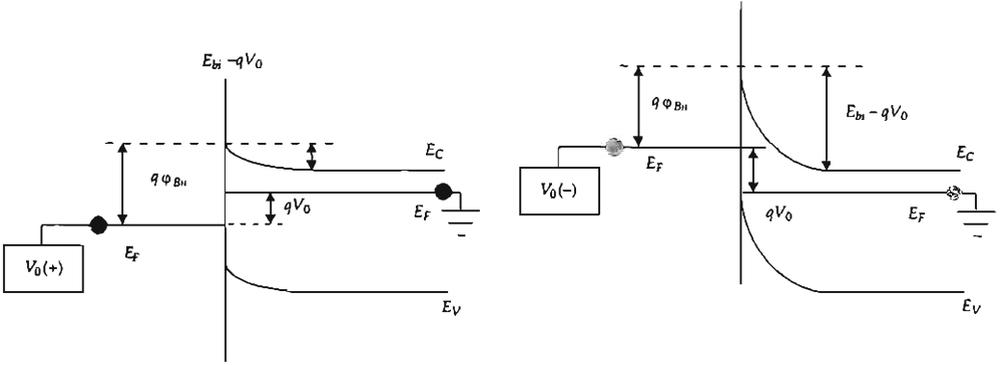


Figura III.21. Efecto de la polarización sobre la unión metal-semiconductor rectificante.

Con la ayuda de los diagramas anteriores es posible concluir que, en equilibrio, la corriente que circula del metal al semiconductor (que no se ha dicho cómo calcular, pero se verá que no es importante) es igual a la corriente que circula del semiconductor al metal y es proporcional a:

$$J(V_0 = 0) \propto n_{i0} \exp\left(-\frac{E_{bi}}{kT}\right) \quad (\text{III.18})$$

Al polarizar en directa, la cantidad de electrones que pasan del metal al semiconductor no se modifica ya que ϕ_{bi} no cambia, y la cantidad de electrones que pasan del semiconductor al metal aumenta conforme se aumenta el voltaje.

$$J(V_0) \propto n_{i0} \exp\left(-\frac{E_{bi}}{kT}\right) \exp\left(\frac{qV_0}{kT}\right) \quad (\text{III.19})$$

Al polarizar en inversa, la corriente del metal al semiconductor continúa siendo una constante, mientras que la del semiconductor al metal ha disminuido con el potencial. De tal manera que,

al menos cualitativamente, la relación V versus I se comporta como en la unión $p-n$.

$$I = I_{0(m-s)} \left[\exp\left(\frac{qV_0}{kT}\right) - 1 \right] \quad (\text{III.20})$$

La densidad de corriente de saturación J_0 depende de parámetros como la corriente que circula del metal al semiconductor, que en general es mucho mayor que la corriente de saturación de la unión $p-n$, llamada corriente termoiónica. Este parámetro es conocido a partir del estudio de las válvulas termoiónicas, ya que el proceso de pasar un electrón del metal al vacío es similar al paso hacia un semiconductor:

$$I_{0(m-s)} = A^* T^2 \exp\left(-\frac{q\phi_{bi}}{kT}\right); A^* = \frac{4\pi q m^* k^2}{h^3} \quad (\text{III.21})$$

donde A^* es la constante efectiva de Richardson (el término se refiere a la masa efectiva de los electrones en el sólido), h es la constante de Planck y k es la constante de Boltzmann.

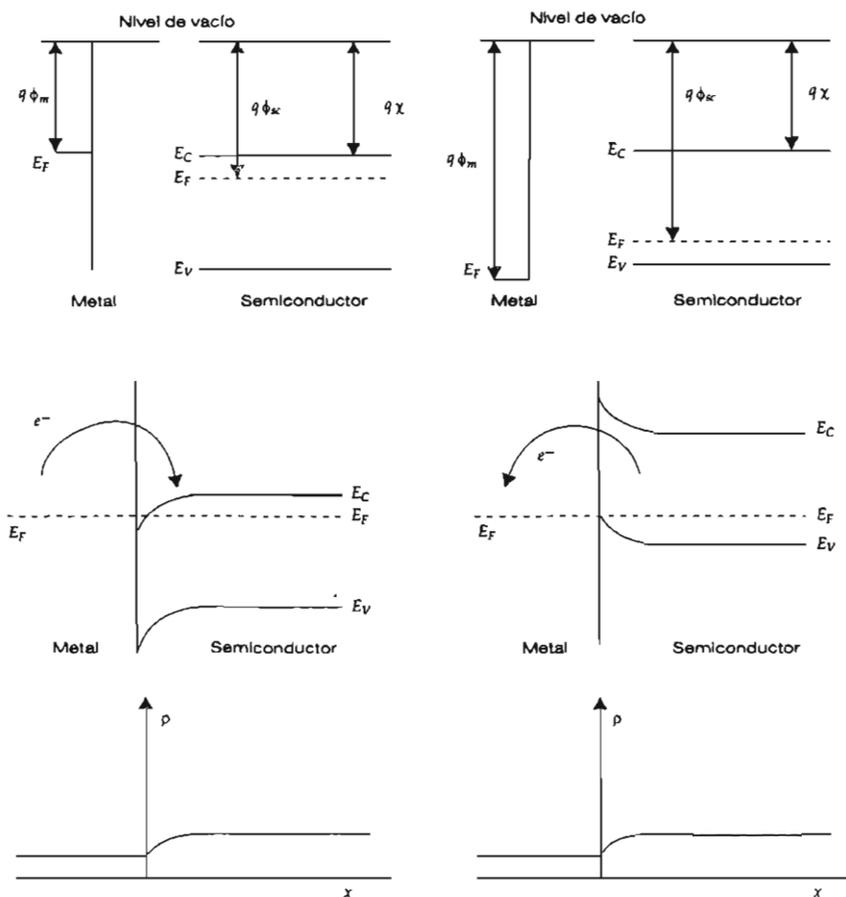


Figura III.22. Diagramas de energía potencial electrostática antes y después de formar la unión entre el metal y el semiconductor que establecen contactos óhmicos de acuerdo con el diagrama de resistividad contra posición. Éstos son los últimos dos casos.

En este dispositivo también existe la capacitancia de unión, procesos de ruptura, efectos de componentes parásitos, etc. La diferencia más interesante consiste en que los transitorios en esta estructura disminuyen en forma considerable, puesto que al conducir por portadores mayoritarios no existe difusión (caso ideal) y por lo tanto al conmutar rápidamente de directa a inversa no existen portadores que puedan regresar y formar el transitorio. Una vez que todos los electrones llegan al metal ya no pueden regresar por difusión: por el abismo $q\phi_{Bn}$ únicamente pue-

den pasar del metal al semiconductor por emisión termoiónica.

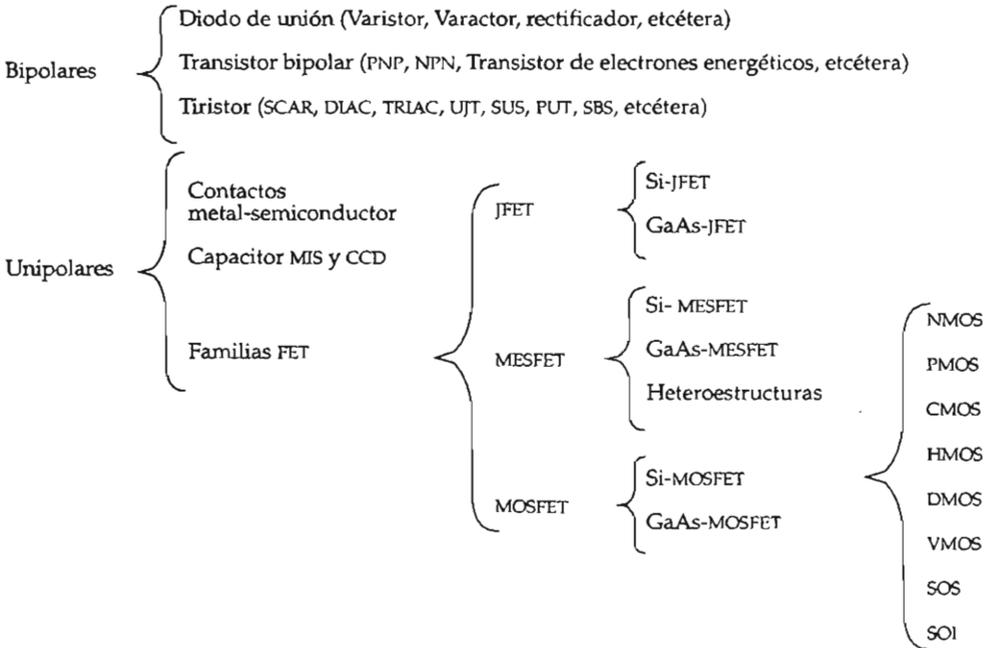
Diagramas de energía en los contactos óhmicos

La figura III.22 ilustra la formación de contactos óhmicos al unir un metal y un semiconductor. La diferencia de funciones trabajo desplaza electrones de tal manera que destruye la barrera de energía y los electrones o huecos pueden pasar libremente por la unión.

CAPACITOR METAL-ÓXIDO-SEMICONDUCTOR

quizá valga la pena resumir en un cuadro las principales divisiones en los dispositivos electrónicos.

Si bien este tema tiene como única finalidad dar una introducción a la física de la estructura MOS,



Algunos de los acrónimos y nombres empleados se describen a continuación:

- **Varistor** (*Variable-resistor*). Dispositivo de dos terminales con comportamiento no óhmico.
- **Varactor** (*Variable-reactor*). Dispositivo de dos terminales en el que la reactancia puede ser controlada por el voltaje. Es el caso de la variación de la capacitancia con el voltaje.
- **Transistor de electrones energéticos** (Hot-Electron Transistor). Se llama electrón caliente al electrón que tiene una energía kT mayor que el nivel de Fermi. Los primeros transistores que funcionaron con estos electrones que no están en equilibrio con la red

- son los MIMIM (Metal-Insulator-Metal-Insulator-Metal) y los más recientes son los llamados transistores balísticos.
- **Tiristor**. Es el nombre genérico que se da a los dispositivos de más de tres capas (los transistores bipolares sólo tienen tres).
- **SCR** (Semiconductor-Controlled-Rectifier).
- **DIAC** (*Diode AC switch*). Es un SCR bidireccional.
- **TRIAC** (*Triode AC switch*). Es un SCR bidireccional.
- **UJT** (UniJunction-Transistor). Es un dispositivo que puede conmutar un estado de alta impedancia OFF, a otro de baja impedancia ON.

- **PUT** (Programable-Unijunction-Transistor). Más versátil que el UJT, pero estructuralmente similar al SCR, sólo cambiando la conexión de compuerta.
- **SUS** (Silicon-Unilateral-Switch). Función similar al PUT, pero estructuralmente diferente.
- **SBS** (Silicon-Bilateral-Switch). Corresponde a dos SUS conectados y funciona como un TRIAC.
- **MIS** (Metal-Insulated-Semiconductor). Siglas generales de todos los dispositivos que emplean aislantes. MOS es un caso particular donde el aislante es un óxido. En general, las familias unipolares en Europa se llaman de efecto de campo y poco a poco ha ido predominando ese nombre.
- **JFET** (Junction Field Effect Transistor). Básicamente consiste en una resistencia controlada por voltaje y equivale a controlar la cantidad de portadores por medio de una zona desértica.
- **MESFET** (Metal-Semiconductor Field Effect Transistor). Básicamente es igual al anterior, sólo que ahora la zona desértica se forma con una unión metal-semiconductor. En el caso anterior es por la unión $p-n$.
- **MOSFET** (Metal-Oxide-Semiconductor-Field-Effect-Transistor). Hay quien le llama erróneamente **ICFET** (Insulated-Gate Field Effect Transistor) o **MOST** (Metal-Oxide-Semiconductor Transistor). El principio físico es la variación de la cantidad de portadores superficiales por el efecto del campo eléctrico. Hoy por hoy es el dispositivo más importante en **VLSI**.
- **N-MOS** es un MOSFET de canal tipo n . Se refiere al tipo de portadores que forman el canal para conectar la conductividad entre S (Source) y D (Drain). El sustrato es tipo p .
- **P-MOS** es un MOSFET de canal tipo p y el sustrato es tipo n .
- **C-MOS** (MOS complementario). Es una tecnología en la que dos transistores, uno de canal n y otro de canal p , se acoplan en un dispositivo.
- **HMOS** (High-performance-MOS). Es un transistor en el que, por medio de la tecnología, se logra controlar el voltaje de encendido, impurificando desigualmente el sustrato.
- **DMOS** (Double-diffused-MOS). Por medio de la tecnología se controla la longitud del canal; pueden alcanzarse longitudes muy cortas.
- **VMOS** (Vertical o V-shaped-grooved-MOS). Estructura geométrica del transistor que puede manejar alta corriente.
- **SOI** (Silicon-On-Insulated). Tecnología que reduce capacitancias parásitas.
- **SOS** (Silicon-On-Shapphire). Tecnología que reduce capacitancias parásitas.
- **FAMOS** (Floating-gate-Avalanche-Injection-MOS). Dispositivo empleado en la construcción de memorias electrónicas ROM que se graban empleando un voltaje relativamente alto, para forzar el efecto túnel de los electrones sobre el óxido de silicio; se pueden borrar con luz ultravioleta, ya que ésta permite relajar la carga acumulada durante el grabado al aumentar la conductividad del óxido de silicio.

Diagrama de energía del capacitor MOS

Las características eléctricas del grupo de dispositivos que más importancia comercial tiene entre los dispositivos electrónicos se pueden entender a través del estudio del capacitor MOS ideal. La estructura física del dispositivo se muestra en la figura III.23 y consta de un metal y un semiconductor separados por un aislante. Si la distancia entre el semiconductor y el metal es del orden de 500 \AA , las propiedades de la interfaz entre el semiconductor y el metal son de muy alta calidad, sin imperfecciones o impurezas, y el semiconductor tiene una pequeña densidad de impurezas del orden de 10^{14} a 10^{16} cm^{-3} . Entonces, por efecto de la polarización entre el metal y el semiconductor, se puede modificar la conductividad cerca de la superficie del semiconductor con el aislante, dando lugar al llamado efecto de campo. Este efecto tiene una gran cantidad de aplicaciones en la actualidad, como los capacitores, que cambian su capacitancia con la polarización y la frecuencia. Varios de estos elementos puestos en paralelo, muy cerca uno de otro, forman un dispositivo que funciona

carga (CCD, Charge-Coupled Devices) y se emplea como sensor de imágenes, procesos de señales y operaciones lógicas. El efecto es también empleado en el funcionamiento del MOSFET, don-

de el capacitor MOS forma la compuerta del transistor que controla la conductividad entre los electrodos de fuente y drenaje.

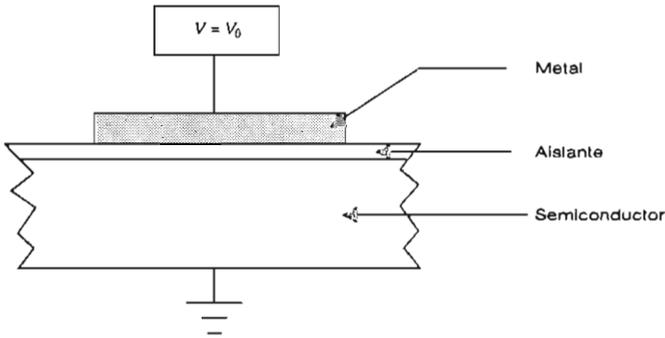


Figura III.23. Estructura física del capacitor MOS.

El capacitor MOS tiene varias diferencias respecto a las estructuras tratadas hasta este momento: por un lado no existe transporte de carga eléctrica en el capacitor y, por lo tanto, no hay necesidad de considerar seudoniveles de Fermi. Además, todo el potencial aplicado aparece entre las superficies del óxido. Si el campo eléctrico en el óxido es muy elevado, $7 \times 10^6 \text{ Vcm}^{-1}$, el dieléctrico conduce destruyendo al dispositivo. El transporte de carga para alcanzar el equilibrio termodinámico (al eliminar la diferencia de la función trabajo del metal y el semiconductor $|\phi_b - \phi_s|$), se alcanza únicamente por medio de alguna conexión externa entre el metal y el semiconductor, que fácilmente se consigue mientras se están fabricando los dispositivos. Desde el punto de vista de la descripción del funcionamiento del capacitor, es indiferente cuál es el origen de la diferencia de potencial entre el metal y el semiconductor, puede ser un potencial externo, la diferencia de función trabajo de los materiales o cargas internas en el aislante. De hecho, al calcular el voltaje de encendido del transistor MOSFET todas estas contribuciones están incluidas.

El parámetro que ahora determina el funcionamiento del dispositivo es el efecto de campo que se caracteriza por ψ ; $q\psi$ es la energía que separa el mínimo de la banda de conducción en equilibrio

(sin efecto de campo) y el mínimo de la banda de conducción cuando existe efecto de campo. El valor máximo es $q\psi$, y mide la diferencia de energía entre el mínimo de la banda de conducción lejos de la interfaz con el aislante y el mínimo de la banda de conducción en la interfaz (como las bandas son paralelas, es igualmente correcto usar como referencia el máximo de la banda de valencia o el centro del intervalo de energía prohibida, el nivel de Fermi intrínseco E_{Fi}).

En general, el voltaje aplicado a la compuerta sumado o restado al potencial interno del capacitor (diferencia de las funciones trabajo, cargas internas, etc.), se muestra en parte en el óxido (definido por su capacitancia); el resto se emplea en el semiconductor para el efecto de campo ψ . Más adelante se expresará explícitamente la relación entre el voltaje aplicado y ψ .

Existen varias condiciones impuestas por el campo eléctrico interno que definen las características de la superficie del semiconductor en el capacitor:

1) La *acumulación* de portadores mayoritarios en la superficie del semiconductor y del aislante. El voltaje aplicado, la diferencia de funciones trabajo y las cargas internas son tales que ψ es menor que 0 en un semiconductor tipo *p*, y es mayor que 0 en un semiconductor tipo *n*.

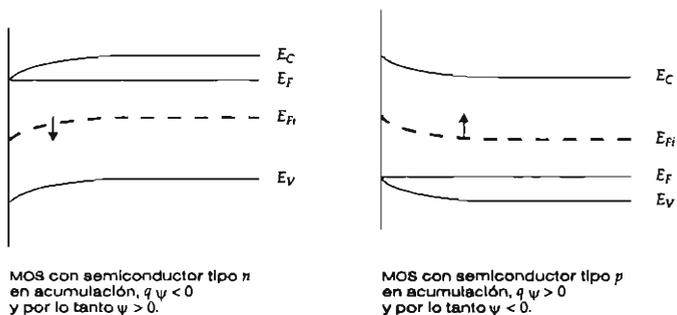
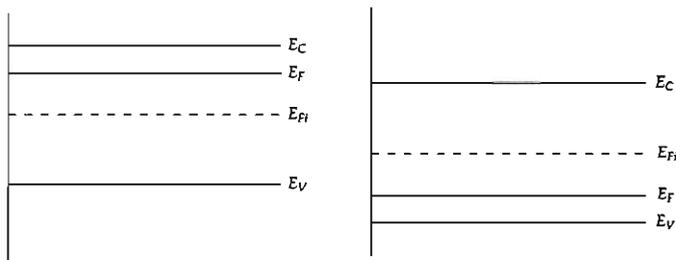


Figura III.24. Representación gráfica de la condición de acumulación de portadores.

2) La *condición de banda plana* es aquella en la que no existe campo eléctrico interno; el voltaje

aplicado y las otras condiciones son tales que $\psi_s = 0$.



La condición de banda plana en un capacitor MOS presenta las bandas de energía sin cambios. No existe efecto de campo.

Figura III.25. Representación gráfica de la condición de banda plana.

3) La condición de formación de *zona desértica* se consigue cuando el campo eléctrico interno producido por el voltaje aplicado y las otras condiciones disminuyen la conductividad del semiconductor a niveles cercanos a la del semiconductor intrínseco, ψ_s es mayor que 0 pero menor que ψ_B para un semiconductor tipo p , y ψ_s es menor que 0 pero mayor que $-\psi_B$ para un semiconductor tipo n .

4) La condición de *inversión* de la población se tiene cuando el campo eléctrico interno modifica la población de portadores de carga de tal manera que existen más portadores minoritarios que la densidad de portadores intrínsecos en la superficie, esto se consigue cuando ψ_s es igual a

ψ_B en un semiconductor tipo p y ψ_s es igual a $-\psi_B$ en un semiconductor tipo n (figura III.26)

5) La condición de *fuerte nivel de inversión* se consigue cuando el campo eléctrico interno es tal que la densidad de portadores minoritarios es mayor o igual que la densidad de portadores mayoritarios. Ésta es la condición necesaria para el inicio de la conducción en un transistor MOSFET. Al voltaje necesario para tener esta condición se le llama de encendido V_t (*threshold*). Se consigue cuando ψ_s es mayor o igual que $2\psi_B$ en un semiconductor tipo p y ψ_s menor o igual que $-2\psi_B$ en un semiconductor tipo n .

En las definiciones anteriores se empleó la constante ψ_B para identificar el estado de carga

superficial. Esto es posible porque $q\psi_B$ es la energía entre el nivel de Fermi y el nivel de Fermi intrínseco. Analíticamente, con la ecuación para el cálculo de la densidad de portadores se obtiene la ecuación:

$$\psi_B = \frac{kT}{q} \ln \left(\frac{N_B}{n_i} \right) \quad (III.22)$$

donde N_B es la densidad de impurezas donadoras N_D en un semiconductor tipo n , o la densidad de impurezas aceptadoras N_A en un semiconductor tipo p .

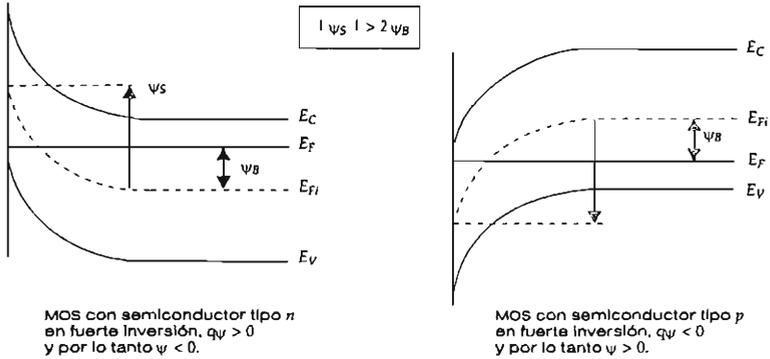


Figura III.26. Diagrama de energía potencial electrostática en el semiconductor de un mos en condición de fuerte inversión.

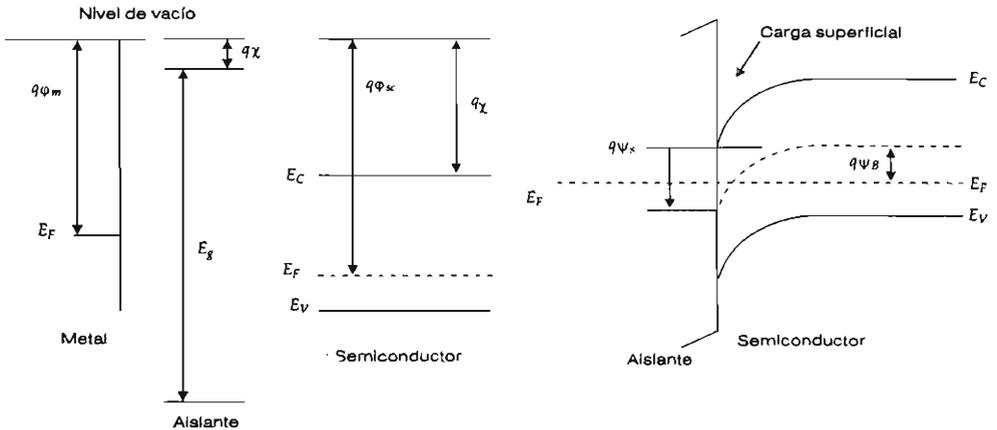


Figura III.27. Diagrama de energía potencial electrostática para un capacitor mos antes y después de la unión.

En el diagrama de la figura III.27 se puede ver que los electrones pasan del metal al semiconductor para llegar al equilibrio (por alguna trayectoria externa), y como ψ_s es positivo y mayor que ψ_b , la superficie se encuentra en inversión de población.

La identificación del origen de la carga superficial es muy importante, ya que no cualquier tipo de carga es igualmente útil en los dispositivos. De hecho, el control de la diferente carga superficial a través del campo eléctrico es lo que

determina el funcionamiento de los diferentes dispositivos.

Los siguientes diagramas de energía y diagramas de carga para un capacitor MOS formado con un semiconductor tipo *p* como el de la figura anterior pueden ayudar a su visualización. Nótese que la variación del estado de carga se consigue con la polarización externa, pero también es función de la diferencia de funciones de trabajo entre el metal y el semiconductor y otros parámetros no considerados, como las cargas internas en el óxido.

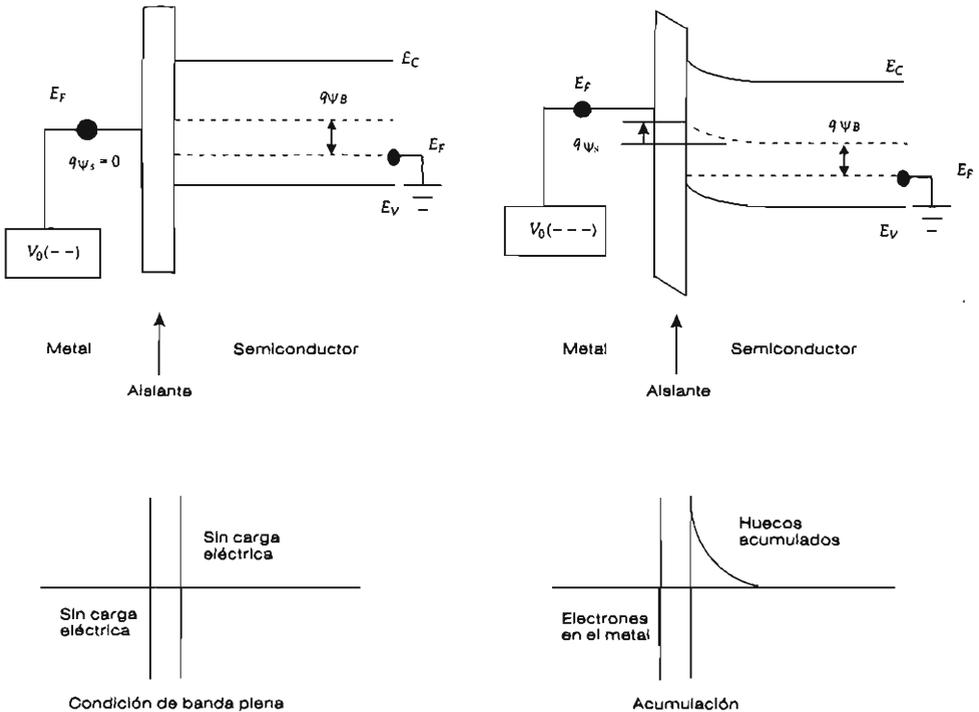
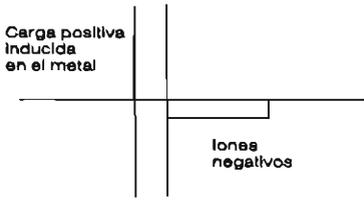
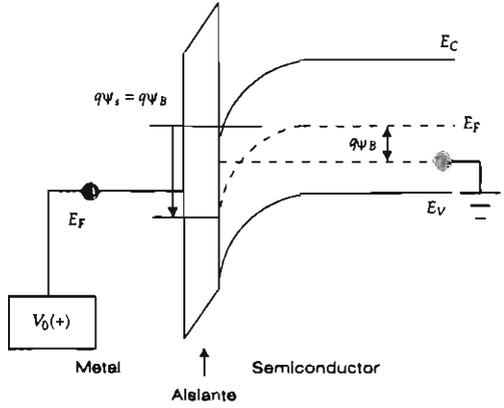
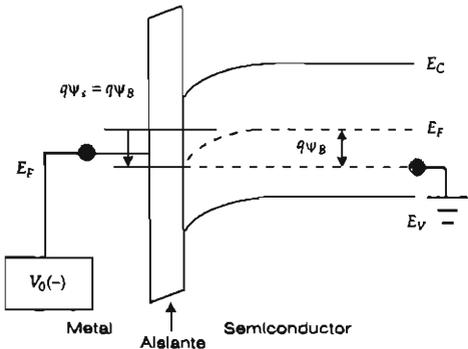


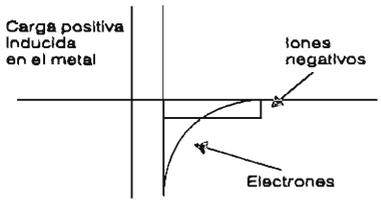
Figura III.28. Diagrama de energía y distribución de carga en condición de banda plana (es necesario aplicar un potencial externo negativo) y acumulación (es necesario un potencial externo aún más negativo, ya que es un semiconductor tipo *p* y $\psi_m < \psi_{sc}$.)

Un resumen de las condiciones que se producen por el campo eléctrico y la carga superficial asociada se puede ilustrar en el siguiente diagrama que muestra la carga superficial $|Q_s|$ en

relación con el potencial superficial ψ_s . El dispositivo tiene características similares al que se ha estado analizando; silicio tipo *p* con 4×10^{15} impurezas aceptoras a 300 K (figura III.30).



Condición de zona desértica



Fuerte inversión

Figura III.29. Diagrama de energía y distribución de carga en condición de formación de zona desértica (es necesario aplicar un potencial externo negativo) y fuerte inversión (es necesario un potencial externo positivo). El dispositivo es el mismo que el de la figura anterior.

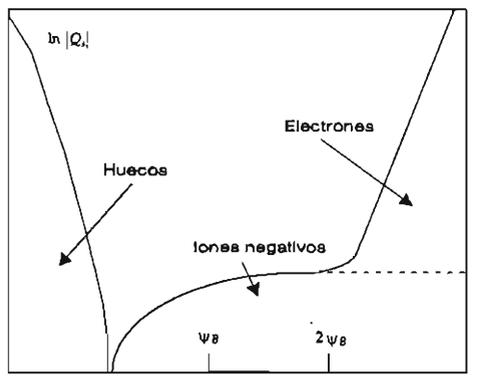
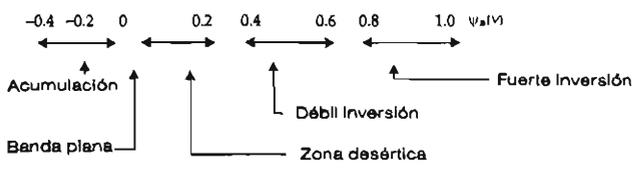


Figura III.30. Carga superficial en un capacitor mos como función del potencial superficial ψ_s .



Esta variación de la carga eléctrica en la zona del semiconductor cercana al aislante, por el efecto de campo que se produce al polarizar la estructura, es lo que produce un comportamiento de capacitor modulado por voltaje, donde la estructura MOS puede modelarse de la siguiente manera:

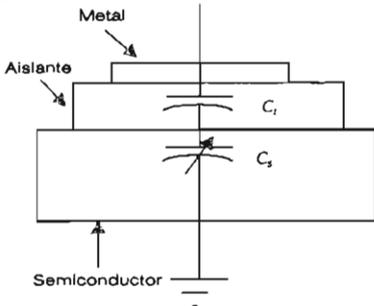


Figura III.31. Modelo capacitivo de un mos.

El capacitor del óxido es una constante y depende únicamente del espesor del óxido. El capacitor del semiconductor puede calcularse con base en la pendiente de la gráfica de la carga superficial $|Q_s|$ versus el potencial superficial ψ_s .

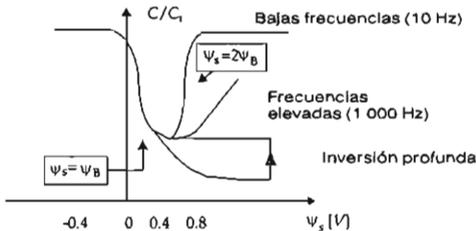


Figura III.32. Variación de la capacitancia en un mos como función de la polarización y de la frecuencia.

Explicar el comportamiento de la curva anterior es relativamente fácil: en la región de acumulación (voltajes superficiales negativos, $\psi_s < 0$), al cambiar el potencial cambia la cantidad de portadores tan rápidamente como el potencial, ya que sólo cambia la cantidad de portadores mayoritarios que se pueden agrupar fácilmente en su superficie. La capacitancia total es únicamente la del óxido.

Para voltajes superficiales positivos, lo que se

tiene en la superficie del semiconductor son iones, y si se da suficiente tiempo al semiconductor se pueden generar los portadores minoritarios. La capacitancia mide la rapidez de cambio en la carga eléctrica, los portadores minoritarios pueden variar por recombinación y generación, así que la capacitancia en esta zona depende de la frecuencia. Los portadores deben ser generados o recombinados para aumentar o disminuir. A baja frecuencia, los portadores pueden seguir la variación del potencial y la capacitancia es nuevamente la del óxido. Si la frecuencia es mayor, los portadores minoritarios no pueden cambiar tan rápidamente y la capacitancia estará determinada por los dos capacitores en serie. El capacitor del semiconductor medirá la modulación de la zona desértica. La zona indicada como inversión profunda se produce cuando el voltaje de polarización (no el que mide la capacitancia, que es muy pequeño) cambia rápidamente produciendo una zona desértica especialmente grande, ya que es más fácil descubrir más iones que generar portadores minoritarios. Esta zona es inestable y la capacitancia rápidamente toma el valor de la capacitancia a la frecuencia de medición.

La relación entre el voltaje aplicado V y el voltaje superficial ψ_s es la siguiente:

$$V = \phi_{ms} + V_i + \psi_s - \frac{Q_B}{C_i} = \phi_{ms} - \left(\chi + \frac{E_g}{2q} \pm \psi_B \right) - \frac{Q_s}{C_i} + \psi_s - \frac{Q_B}{C_i} \quad (III.23)$$

donde ϕ_{ms} es la diferencia de funciones-trabajo entre el metal y el semiconductor. Si este número es positivo, provoca un corrimiento de la curva de capacitancia hacia la derecha. Para un semiconductor tipo p se usa el signo positivo, para un tipo n el negativo. V_i es la caída de potencial en el óxido y depende de la capacitancia en el óxido C_i y de la carga superficial Q_s . Q_B son las cargas positivas fijas del óxido.

La cantidad de carga en el óxido es un parámetro difícil de conocer para cualquier condición de polarización. En fuerte inversión, la que indica el voltaje de encendido de un MOSFET, se

tienen las siguientes expresiones para ψ_s , y para V_i o Q_s :

$$\psi_s = 2\psi_B$$

$$V_i = \frac{\sqrt{2\epsilon_s \epsilon_0 N_A (2\psi_B)}}{C_i} \quad (\text{III.24})$$

donde se supone que la carga en el semiconductor está formada por los iones de zona desértica. Además, se han empleado las ecuaciones de la zona desértica obtenidas al estudiar la unión $p-n$.

TRANSISTOR BIPOLAR DE UNIÓN (BJT)

El transistor (*transfer resistor*) bipolar de unión es un dispositivo construido al unir en serie tres semiconductores de diferente conductividad $n-p-n$ o $p-n-p$, que toman el nombre de emisor, base y colector. Por condiciones tecnológicas de fabricación, el dispositivo no es simétrico y no puede intercambiarse el emisor por el colector, ya que las características de funcionamiento son diferentes. La operación normal del dispositivo se consigue al polarizar la unión emisor-base en directa y la unión base-colector en inversa. La estructura de un transistor $p-n-p$ ideal se presenta en la figura III.33:

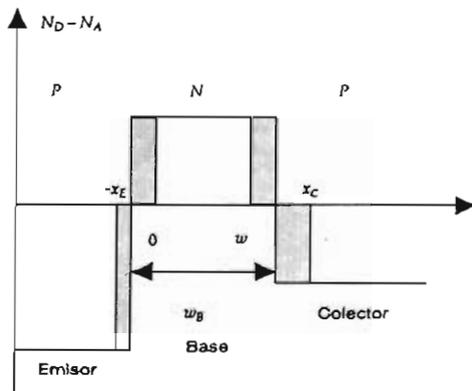


Figura III.33. Estructura simplificada de un transistor $p-n-p$.

Ecuaciones de transporte

Las relaciones de transporte en este dispositivo se pueden obtener, de manera similar a la unión $p-n$, con base en la hipótesis de que las ecuaciones de continuidad son válidas.

$$\frac{\partial p_n}{\partial t} = G_p - \frac{p_n - p_{n0}}{\tau_p} - p_n \mu_p \frac{\partial E}{\partial x} - \mu_p E \frac{\partial p_n}{\partial x} + D_p \frac{\partial^2 p_n}{\partial x^2}$$

$$\frac{\partial n_p}{\partial t} = G_n - \frac{n_p - n_{p0}}{\tau_n} + n_p \mu_n \frac{\partial E}{\partial x} + \mu_n E \frac{\partial n_p}{\partial x} + D_n \frac{\partial^2 n_p}{\partial x^2} \quad (\text{III.25})$$

Al igual que las ecuaciones de la densidad de corriente:

$$J_p = q\mu_p p E - qD_p \nabla p$$

$$J_n = q\mu_n n E + qD_n \nabla n \quad (\text{III.26})$$

Las simplificaciones son las mismas que en la unión $p-n$. Las dos uniones son uniones $p-n$ ideales: se evita la recombinación superficial, no existe resistencia en serie y se impone un bajo nivel de inyección. El diagrama de energía de la estructura se muestra en la figura III.34.

En este dispositivo, la densidad de portadores mayoritarios no cambia. El control del transporte es realizado por los portadores minoritarios y las condiciones de frontera para los portadores minoritarios en el transistor $p-n-p$ son las siguientes:

Para el emisor, definido para toda x desde $-\infty$ hasta $-x_E$:

$$n(-\infty) = n_{j0} \equiv n_E$$

$$n(-x_E) p(-x_E) = n_i^2 \exp\left(\frac{qV_{EB}}{kT}\right)$$

$$n(-x_E) = n_E \exp\left(\frac{qV_{EB}}{kT}\right) \quad (\text{III.27})$$

Para la base definida desde $x = 0$ hasta $x = w$, se tiene:

$$p(0) = p_n \exp\left(\frac{qV_{EB}}{kT}\right)$$

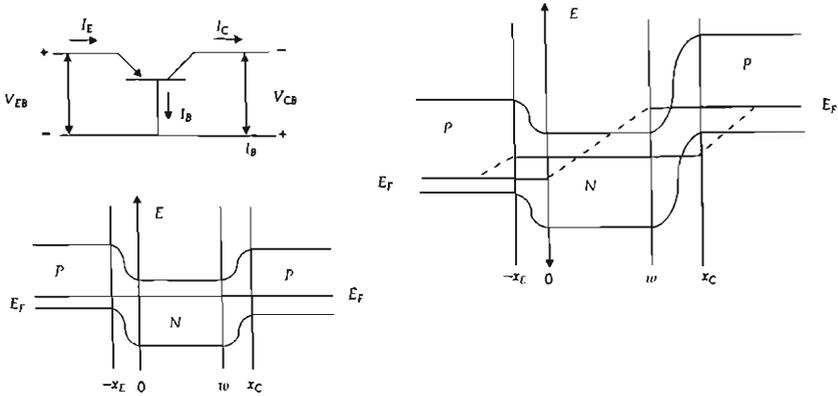


Figura III.34. Diagrama de energía potencial para un transistor $p-n-p$ en equilibrio y polarizado. V_{EB} es el voltaje de la unión del emisor en directa y V_{CB} es el voltaje de la unión del colector en inversa.

$$p(w) = p_b \exp\left(\frac{qV_{CB}}{kT}\right) \quad (III.28)$$

Para el colector, definido para toda x mayor que x_C :

$$n(x_C) = n_c \exp\left(\frac{qV_{CB}}{kT}\right) \quad (III.29)$$

$$n(\infty) = n_c$$

La obtención de estas ecuaciones es bastante fácil, como se puede observar en la segunda ecuación para el emisor. Basta con aplicar la ley de acción de masas empleando los seudoniveles de Fermi.

Al resolver las ecuaciones de continuidad empleando las condiciones de frontera correspondientes y usando la notación $\Delta p(0) \equiv p_b(0) - p_b$ para los minoritarios en $x = 0$ y de manera similar para los otros casos, se obtiene:

$$n_e(x) = n_e + \Delta n(-x_e) \exp\left(\frac{(x + x_e)}{L_e}\right) \quad \text{para } x \leq -x_e \quad (III.30)$$

$$p_b(x) = p_b + \left[\frac{\Delta p(w) - \Delta p(0) \exp\left(-\frac{w}{L_b}\right)}{2 \sinh\left(\frac{w}{L_b}\right)} \right] \exp\left(\frac{x}{L_b}\right) - \left[\frac{\Delta p(w) - \Delta p(0) \exp\left(\frac{w}{L_b}\right)}{2 \sinh\left(\frac{w}{L_b}\right)} \right] \exp\left(-\frac{x}{L_b}\right) \quad (III.31)$$

para $w > x > 0$

$$n_c(x) = n_c + \Delta n(x_c) \exp\left(-\frac{(x - x_c)}{L_c}\right) \quad \text{para } x \geq x_c \quad (III.32)$$

Las densidades de corriente se obtienen de manera similar a la unión $p-n$. La siguiente descripción se puede visualizar con la ayuda de las figuras III.35 y III.36. Para la corriente del emisor se tiene: $I_e = I_e(\text{BASE}) + I_e(\text{EMISOR})$, donde el primer término representa la inyección de huecos del

emisor a la base (que se obtiene derivando la densidad de portadores en la base en el punto $x = 0$) y la segunda es la inyección de electrones de la base al emisor (se obtiene derivando la densidad de portadores en el emisor en el punto $x = -x_e$).

$$I_E(\text{BASE}) = \frac{AqD_B p_B}{L_B} \coth\left(\frac{w}{L_B}\right) \left[\exp\left(\frac{qV_{EB}}{kT}\right) - 1 - \frac{1}{\cosh\left(\frac{w}{L_B}\right)} \left(\exp\left(\frac{qV_{CB}}{kT}\right) - 1 \right) \right] \quad (\text{III.33})$$

$$I_E(\text{EMISOR}) = \frac{AqD_E n_E}{L_E} \left(\exp\left(\frac{qV_{EB}}{kT}\right) - 1 \right) \quad (\text{III.34})$$

Para la corriente del colector se tiene que: $I_C = I_C(\text{BASE}) + I_C(\text{COLECTOR})$, donde el primer término representa la inyección de huecos de la base al colector (se obtiene derivando la densidad de portadores en la base en el punto $x = w$)

y la segunda es la inyección de electrones del colector a la base (se obtiene derivando la densidad de portadores en el colector en el punto $x = x_C$).

$$I_C(\text{BASE}) = \frac{AqD_R p_B}{L_R} \frac{1}{\sinh\left(\frac{w}{L_R}\right)} \left[\exp\left(\frac{qV_{EB}}{kT}\right) - 1 - \cosh\left(\frac{w}{L_R}\right) \left(\exp\left(\frac{qV_{CB}}{kT}\right) - 1 \right) \right] \quad (\text{III.35})$$

$$I_C(\text{COLECTOR}) = - \frac{AqD_C n_C}{L_C} \left(\exp\left(\frac{qV_{CB}}{kT}\right) - 1 \right) \quad (\text{III.36})$$

Finalmente, la corriente de base se obtiene de la diferencia entre la corriente del emisor y la corriente del colector.

gen cualitativa que ayuda a tener una mejor comprensión del dispositivo.

$$I_B = I_E - I_C \quad (\text{III.37})$$

Ganancia en corriente

Como el interés en esta sección es solamente presentar una introducción general, la parte restante de este capítulo tratará de la ganancia en corriente del transistor y finalmente de una ima-

En el análisis anterior se identificaron las contribuciones a la densidad de corriente en el transistor. A continuación, presentamos un resumen.

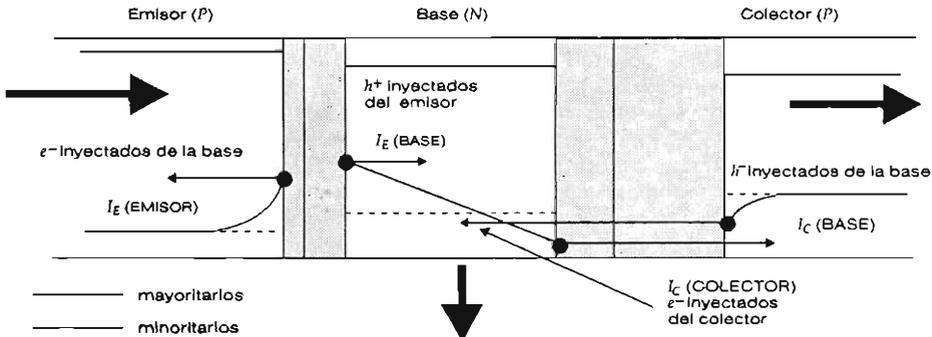


Figura III.35. En la unión E-B la corriente de huecos es muchos órdenes de magnitud mayor que la corriente de electrones, si se diseña el transistor correctamente. La corriente de electrones en la unión B-C se debe a la corriente de saturación de la unión en Inversa. La corriente de base debe ser tan pequeña como sea posible en un buen transistor. Aquí se ha supuesto que no existe recombinación en las zonas desérticas (áreas sombreadas). La unión B-E está en directa y la unión B-C está en inversa.

La ganancia en corriente a base común α_0 se define de la siguiente manera:

$$\alpha_0 \equiv h_{FB} = \frac{\partial I_C}{\partial I_E} = \frac{\partial I_E(\text{BASE})}{\partial I_E} \frac{\partial I_C(\text{BASE})}{\partial I_E(\text{BASE})} \frac{\partial I_C}{\partial I_C(\text{BASE})} = \gamma \alpha_T M \quad (\text{III.38})$$

donde γ es la eficiencia de emisión, α_T es el factor de transporte en la base y M es el factor de multiplicación en la unión base-colector.

De manera similar, la ganancia en corriente a emisor común es:

$$\beta_0 \equiv h_{FE} = \frac{\partial I_C}{\partial I_B} = \frac{\alpha_0}{1 - \alpha_0} \quad (\text{III.39})$$

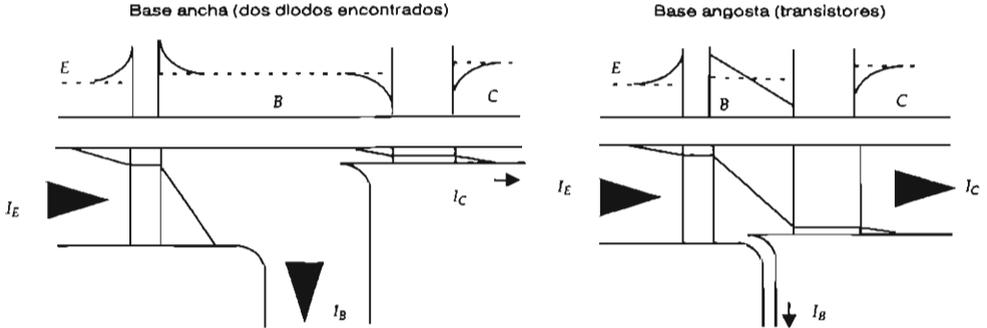


Figura III.36. Las curvas superiores representan la densidad de portadores minoritarios y las curvas inferiores la corriente eléctrica. Los diagramas de la izquierda corresponden a dos diodos encontrados donde la corriente de emisor es muy parecida a la de base. Los diagramas de la derecha corresponden a un transistor, donde la corriente de emisor es muy parecida a la corriente de colector, al estar la unión E-B en directa y la unión B-C en inversa.

En condiciones normales de operación, es posible obtener una expresión simplificada para la ganancia de un transistor, ya que M es aproximadamente igual a uno cuando el voltaje en inversa de la unión base-colector es menor al voltaje de ruptura, lo que significa que por cada portador que se presenta en $x = w$, el campo eléctrico lo traslada a $x = x_C$. α_T también se aproxima a la unidad, siempre y cuando el ancho de la base (w) sea más pequeño que la longitud de difusión de los portadores minoritarios, lo que significa que los portadores no desaparecen en la base: por cada portador que es inyectado en $x = 0$, todos llegan hasta $x = w$. Ésta es la esencia del transistor: las dos uniones $p-n$ deben estar lo suficientemente cerca para que exista el efecto transistor. Con estas condiciones, la ganancia en corriente a base común es aproximadamente igual a la eficiencia de emisión, que después de algunas simplificaciones se puede escribir de la siguiente manera:

$$\alpha \equiv \gamma = \frac{\partial A J_E(\text{BASE})}{\partial I_E} = \frac{\frac{\partial A J_E(\text{BASE})}{\partial e} \frac{V_{EB}}{kT}}{\frac{\partial I_E}{\partial e} \frac{V_{FB}}{kT}} = \frac{p_B D_B L_E}{n_E D_E L_B} \coth\left(\frac{w}{L_B}\right) \sim \frac{N_E}{N_B w} \quad (\text{III.40})$$

La β es mayor conforme mayor es la cantidad de impurezas en el emisor N_E , menor la cantidad de impurezas en la base N_B y menor el ancho de la base. Queda como ejercicio trabajar las expresiones para justificar este resultado.

Otro resultado interesante se obtiene cuando se averigua si la ganancia en corriente es una constante para cualquier valor de corriente de colector. A bajos niveles de corriente de colector, la ganancia en corriente aumenta con la corriente de colector. Este comportamiento se debe a que pequeñas corrientes de colector son ineficientes para lograr la inyección de portadores, y una

gran parte no produce inyección debido a la recombinación en la zona desértica de la unión base-emisor y a las corrientes de fuga superficiales. Con altos niveles de corrientes de colector, la

ganancia disminuye conforme aumenta la corriente, debido al aumento en la recombinación a altos niveles de inyección.

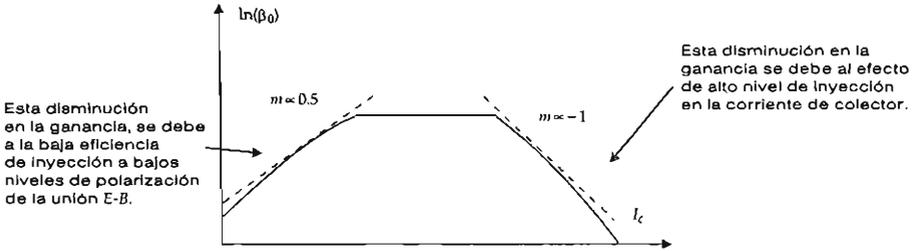


Figura III.37. Disminución de la ganancia en un BJT a bajo nivel de corriente de colector por la corriente de fuga, y a alto nivel de corriente de colector por el alto nivel de inyección.

En resumen, un análisis cualitativo de las ecuaciones presentadas es el siguiente:

- El voltaje aplicado controla la densidad de portadores en la frontera, a través de los términos $\exp(V_{EB}/kT)$ y $\exp(V_{CB}/kT)$.
- Las corrientes de emisor y colector están

dadas por los gradientes en las densidades de los portadores minoritarios en las fronteras de la base, esto es en $x = 0$ y $x = w$.

- La corriente de base es la diferencia entre las corrientes de emisor y colector.

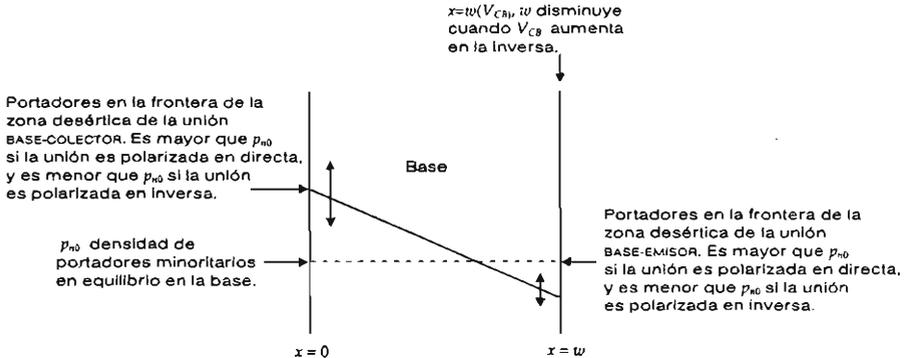


Figura III.38. El efecto transistor se obtiene al modificar la concentración de portadores minoritarios en la base.

Los puntos representativos de polarización, como los representados en las curvas de emisor común de un BJT, modifican la concentración de portadores en la base. En la figura III.39 se representan la saturación, activa directa y el corte o apagado.

MOSFET

El transistor de efecto de campo controlado por un capacitor MOS, MOSFET (*Metal-Oxide-Semiconductor Field-Effect-Transistor*) es el transistor de mayor importancia en la fabricación de circuitos integrados y recientemente en aplicaciones de potencia. Este dispositivo presenta conducción

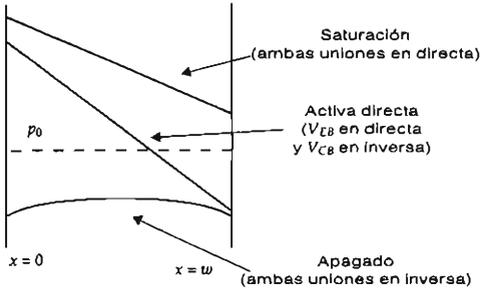


Figura III.39. Cambio en la concentración de portadores minoritarios en la base por las diferentes condiciones de polarización.

por portadores de un solo tipo, por lo que los dispositivos son llamados unipolares. La construcción física de este transistor se muestra en la figura III.40. Un capacitor MOS forma la compuerta G y en el semiconductor se difunden dos

zonas de conductividad diferentes a la del sustrato a cada lado del capacitor MOS. Una se llama fuente S y la otra se llama drenaje D . Usualmente, el sustrato y la fuente se conectan eléctricamente, lo que forma un dispositivo de tres terminales que no es simétrico en relación con el intercambio de los electrodos de fuente y de drenaje. Si el semiconductor es de tipo p y no existe condición de inversión de la población en la compuerta cuando no está polarizado, se conoce como MOSFET de canal n normalmente apagado, ya que las zonas tipo n (S y D) están aisladas eléctricamente por las dos zonas desérticas que se forman y no existe conducción entre S y D . Cuando se polariza la compuerta hasta conseguir un fuerte nivel de inversión, se forma un canal de electrones que pone en contacto la fuente y el drenaje, permitiendo el paso de la corriente eléctrica.

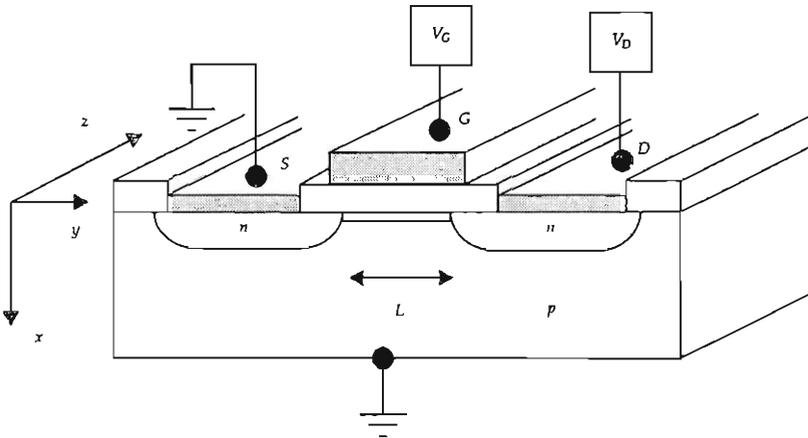


Figura III.40. Diagrama físico de un MOSFET de canal n .

Para entender los diagramas de energía potencial electrostática contra posición es necesario hacer algunos comentarios, ya que se obtendrían diagramas en cuatro dimensiones: x , y , z y E . Como esto es imposible de presentar en una gráfica y además poco útil, intentaremos visualizar las diferentes condiciones por separado.

Si en el transistor se polariza la compuerta de tal manera que se obtiene la condición de banda

plana y no existe voltaje en el drenaje, entonces el diagrama de energía potencial es idéntico a un transistor BJT en equilibrio, donde el emisor hace las bases de la fuente, el colector es el drenaje y la base es el sustrato.

Cuando el voltaje del drenaje es cero, la compuerta se polariza hasta llegar a una fuerte inversión V_T . El diagrama de energía en la superficie con el aislante es semejante a un semiconductor

tipo n desde el electrodo de fuente hasta el electrodo de drenaje. Alejándose de la superficie con el aislante el diagrama de energía se parece más y más al caso anterior.

Cuando la compuerta está en condición de banda plana y el drenaje se polariza con un voltaje positivo respecto a la fuente, entonces se tienen dos uniones $p-n$: la unión de fuente-sus-

trato está en directa y la unión sustrato-drenaje está en inversa, así que la corriente de drenaje es igual a la corriente de saturación de un diodo en inversa.

El diagrama de energía potencial electrostática que se presenta en la siguiente figura es para el caso en que V_G es mayor que V_T y V_D es mayor que cero.

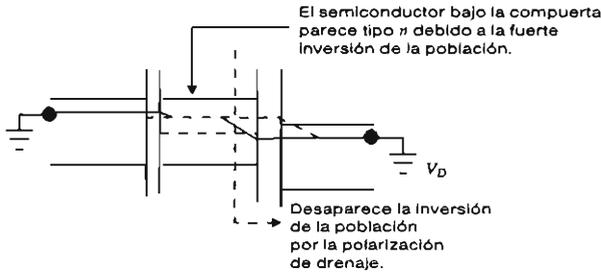


Figura III.41. Diagrama de energía potencial electrostática para la superficie del semiconductor en un transistor MOSFET, cuando V_G es mayor que V_T y V_D es positivo.

El semiconductor que está frente a la compuerta es diferente del diagrama clásico del ca-

pacitor MOS, ya que existe una corriente circulando entre el drenaje y la fuente.

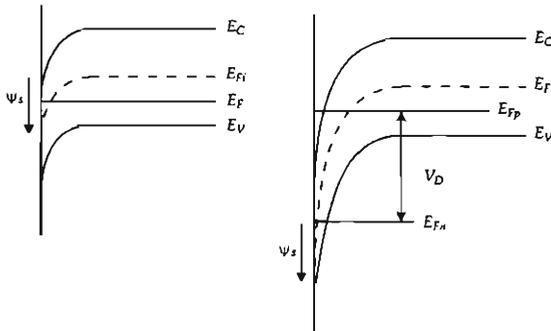


Figura III.42. Diagrama de energía en el semiconductor situado frente a la compuerta; fuerte inversión de la población en equilibrio y fuera de equilibrio.

En la construcción del diagrama anterior se hacen las siguientes suposiciones:

El seudonivel de Fermi de los mayoritarios no varía del volumen a la interfaz con el óxido.

El seudonivel de Fermi de los minoritarios está separado del seudonivel de Fermi de los mayoritarios por el voltaje aplicado V_D .

Ecuaciones de transporte

Las ecuaciones de transporte de un MOSFET se pueden obtener haciendo algunas simplificaciones adicionales:

- La estructura de la compuerta corresponde a una estructura MOS ideal.

- Sólo se considera la corriente de deriva en la estructura.
- La movilidad de los portadores en la zona de inversión es constante.
- La contaminación en el semiconductor es uniforme.
- La corriente en inversa en las uniones es muy pequeña.
- El campo eléctrico transversal (E_x) es mucho mayor que el campo eléctrico longitudinal (E_y).

La corriente de drenaje I_D como función de los voltajes de compuerta V_G y de drenaje V_D se puede encontrar al aplicar la ley de Ohm a la carga superficial formada por la compuerta, donde x es la coordenada que se interna en el semiconductor alejándose de la compuerta y la superficie del semiconductor es en $x = 0$. y es la coordenada que va de la fuente al drenaje, donde $y = 0$ es el electrodo de la fuente. L es la longitud del canal y x_i es el grosor de la zona de inversión.

La conductancia del canal formado por la compuerta es:

$$g = \frac{Z}{L} \int_0^{x_i} \sigma(x) dx = \frac{qZ\mu_n}{L} \int_0^{x_i} n(x) dx = \frac{qZ\mu_n |Q_n|}{L} \quad (\text{III.41})$$

donde Z , el ancho de canal, es uniforme para cualquier z . La cantidad de carga móvil encerrada desde $x = 0$ hasta $x = x_i$ es $|Q_n(y)|$. La caída de voltaje en un tramo del canal de ancho dy al pasar una corriente I_D es:

$$dV = I_D dR = I_D \frac{dy}{gL} = \frac{I_D dy}{Z\mu_n |Q_n(y)|} \quad (\text{III.42})$$

Por supuesto, I_D es independiente de y , ya que es una constante.

La carga a lo largo del canal $|Q_n(y)|$ cambia de acuerdo con los voltajes aplicados. Una manera más sencilla de expresarlo es con la ayuda de las simplificaciones de la compuerta MOS ideal, sin trampas interfaciales ni cargas eléctricas en el óxido y sin diferencias de funciones trabajo.

$$Q_s(y) = [-V_G + \psi_s(y)] C_i$$

Esta carga en el semiconductor tipo p está formada por los iones de la zona desértica $Q_b(y)$ y por los electrones de acumulación $Q_n(y)$.

$$Q_s(y) = Q_n(y) + Q_b(y)$$

El voltaje superficial en el semiconductor en condición de fuerte inversión es:

$$\psi_s(y) = 2\psi_B + V(y)$$

donde $V(y)$ es el voltaje entre un punto y y el electrodo de fuente (tierra). Combinando las ecuaciones anteriores se obtiene la carga a lo largo del canal:

$$Q_n(y) = \frac{-(V_G - V(y) - 2\psi_B)C_i}{\sqrt{2\epsilon_s q N_A [V(y) + 2\psi_B]}} \quad (\text{III.43})$$

donde ψ_B está definido en la sección del capacitor MOS y es función únicamente del semiconductor empleado en el sustrato, C_i es la capacitancia del óxido y ϵ_s es la permisividad del semiconductor.

Al integrar desde $y = 0$ hasta $y = L$ con $V = 0$ hasta $V = V_D$.

$$I_D = \frac{Z}{L} \mu_n C_i \left\{ \left(V_G - 2\psi_B - \frac{V_D}{2} \right) V_D - \frac{2\sqrt{2\epsilon_s q N_A}}{3C_i} \left[(V_D + 2\psi_B)^{3/2} - (2\psi_B)^{3/2} \right] \right\} \quad (\text{III.44})$$

Para un valor dado de V_G , la corriente de drenaje aumenta linealmente con el voltaje de drenaje (régimen lineal u óhmico) y paulatinamente deja de aumentar hasta que alcanza un valor de saturación (región de saturación) idéntico a la curva experimental del transistor MOSFET presentado en la sección de tecnología.

Cuando los voltajes de polarización del drenaje son pequeños, la ecuación anterior se puede escribir de la siguiente forma:

$$I_D = \beta \left((V_G - V_T) V_D - \gamma V_D^2 \right)$$

$$\beta = \frac{Z}{L} \mu_n C_i$$

$$\gamma = \left(\frac{1}{2} + \frac{\sqrt{\epsilon_0 \epsilon_s q N_A / \psi_B}}{4C_i} \right)$$

$$V_T = 2\psi_B + \frac{\sqrt{2\epsilon_0 \epsilon_s q N_A} (2\psi_B)}{C_i} \quad (\text{III.45})$$

Y con una corriente de saturación:

$$I_{D, \text{sat}} \approx \frac{Z}{2L} \mu_n C_i (V_G - V_T)^2 \quad (\text{III.46})$$

Funcionamiento

A continuación presentamos un resumen cualitativo del funcionamiento del MOSFET, usando una analogía hidráulica.

La región lineal del MOSFET se consigue con un voltaje de compuerta mayor que el voltaje de encendido y voltajes de drenaje pequeños. Podemos imaginar dos depósitos con agua, uno fijo —el depósito de fuente— y el otro móvil —el de drenaje—. Si el voltaje de drenaje es positivo, el depósito baja como la energía potencial. Los dos depósitos están comunicados por un conducto que tiene una articulación en la unión con el depósito de fuente y está sostenido sobre el depósito de drenaje con una particularidad: existe un pivote que impedirá que el conducto baje aun después de que el depósito de drenaje continúe bajando.

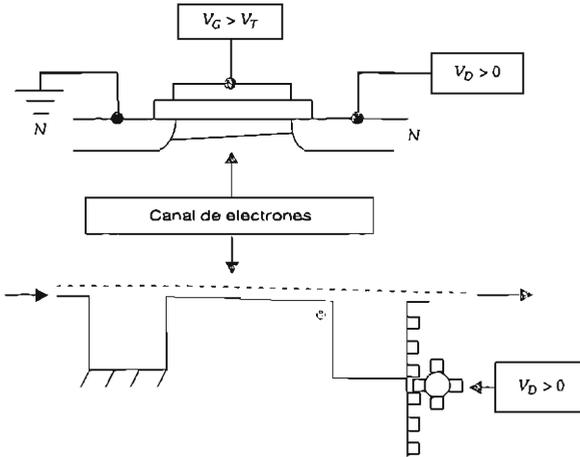


Figura III.43. Región lineal en un MOSFET e Imagen hidráulica.

En la región de saturación el voltaje de drenaje ha aumentado y es aproximadamente igual al voltaje de la compuerta menos el voltaje de encendido ($V_D = V_G - V_T$). En estas condiciones, la imagen hidráulica muestra el depósito de drenaje en un nivel inferior respecto al depósito de fuente, y el conducto que los une ha tocado el pivote que impide que continúe tocando el depósito de drenaje (figura III.44).

Más allá de la saturación, el voltaje de drenaje

continúa aumentando pero no la corriente de drenaje. En la imagen hidráulica no aumenta el flujo de agua porque la inclinación del conducto es constante. La explicación eléctrica es que la diferencia de potencial desde la fuente hasta el final de la zona de inversión es una constante y es el potencial que acelera los electrones. Los electrones que han llegado a este punto se desplazan por la unión en inversa, como en la zona base colector de un bjt (figura III.45).

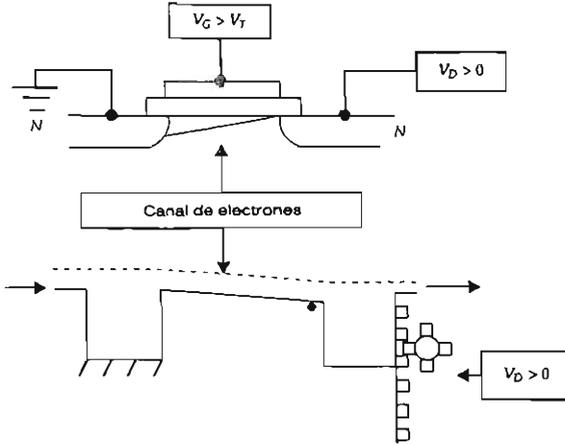


Figura III.44. Región de saturación en un MOSFET e imagen hidráulica.

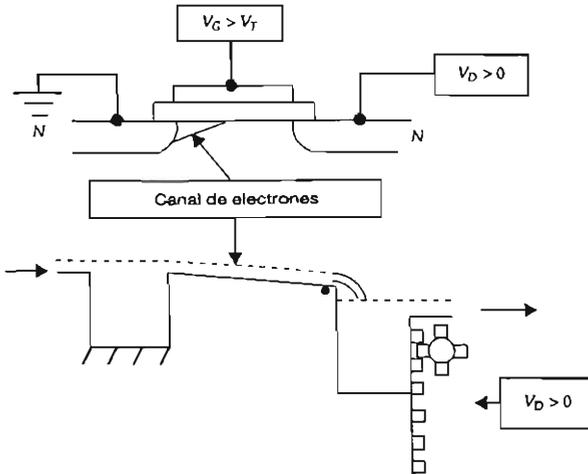


Figura III.45. Región posterior a saturación en un MOSFET e imagen hidráulica.

LED, LÁSER Y FOTODETECTORES

Esta sección se ocupa de la descripción de los dispositivos optoelectrónicos que basan su funcionamiento en los semiconductores y en los dispositivos semiconductores. Estos dispositivos deben llamarse en realidad dispositivos fotónicos, ya que en ellos la partícula que desem-

peña el papel más importante es el fotón, es decir, la luz.

Empecemos diferenciando dos conceptos relacionados con los dispositivos electroópticos: la incandescencia y la luminiscencia. La primera es la radiación que resulta solamente de la temperatura del material, mientras que la luminiscencia es la emisión de radiación (ultravioleta, visi-

ble o infrarroja) como resultado de la excitación electrónica. Ciertas formas de energía permiten que electrones se desplacen a niveles de mayor energía y, en consecuencia, emitan radiación al regresar a su estado base. Al suspender la forma de energía, se espera que la luminiscencia desaparezca; si es así, se llama fluorescencia. Si por alguna razón la luminiscencia persiste, se llama fosforescencia (usualmente ésta se debe a estados metaestables que atrapan electrones y los liberan lentamente).

Existen varios dispositivos exhibidores electrónicos que no basan su funcionamiento en los dispositivos semiconductores. Algunos emplean materiales que permiten emitir radiación, mientras que otros sólo modulan la radiación que incide en ellos. Así, por ejemplo, los tubos de televisión de color emplean la cátodoluminiscencia, donde

tres cañones de electrones ligeramente inclinados el uno hacia el otro coinciden al pasar por una máscara y golpear en tres diferentes regiones el material, que emite luz al ser golpeado por los electrones. Cada región emite luz en un color primario, azul (ZnS:Ag), verde (Zn,Cd_{1-x}S:Cu) y rojo (Y₂O₃S:Eu, Tb) y constituye un punto de color de la imagen.

Otro ejemplo son los dispositivos electroluminiscentes clásicos, donde a un material (ZnS:Cu) en polvo o en película se le aplica una gran diferencia de potencial DC o AC y emite luz con eficiencia cercana a 1 por ciento.

En los exhibidores de plasma, la emisión de luz se produce por la neutralización atómica después de ser ionizados al pasar una corriente eléctrica.

	Cercano UV		Violeta	Azul	Verde	Amarillo	Anaranjado	Rojo	Cercano IR
λ [μm]	0.30	0.39	0.455	0.492	0.577	0.597	0.622	0.77	1.50
Energía [eV]	4.1	3.19	2.73	2.52	2.15	2.08	2.00	1.61	0.83

Los exhibidores de cristal líquido no emiten luz, únicamente la modulan. Un material capaz de modificar el ángulo de polarización de luz cuando es energizado (cristal líquido) se coloca entre dos polarizadores a 90° con un espejo al final. Cuando no está energizado el sistema, la luz incidente no es reflejada debido al arreglo de los polarizadores, y se ve negro. Cuando se energiza el sistema, el cristal líquido modifica el ángulo de polarización 90° y entonces la luz se refleja y se ve brillante.

Los materiales optoelectrónicos que basan su funcionamiento en los semiconductores se pueden dividir en tres grupos:

1. Dispositivos que convierten energía eléctrica en radiación.
 - LED (*Light-Emitting Diode*).
 - Diodo láser (*Light amplification by stimulated emission of radiation*).
2. Dispositivos que detectan señales ópticas a través de procesos electrónicos.

- Fotodetectores.
3. Dispositivos que convierten radiación óptica en energía eléctrica.
 - Celdas solares (dispositivos fotovoltaicos).
- Para mostrar aún más la diferencia entre los dispositivos, se puede ver la gráfica de la intensidad de radiación como función de la energía de un material incandescente, un LED y un láser (figura III.46).

Los LED son uniones *p-n* polarizadas en directa que emiten espontáneamente luz por la recombinación de los portadores cerca de la unión metalúrgica, siempre y cuando la transición de los electrones de la banda de conducción a la de valencia sea directa, por ejemplo en semiconductores, como el GaAs.

Si bien los LED para el visible son en general de aleaciones de arseniuro de galio y fósforo (GaAs_{1-x}P_x con $0 < x < 0.45$ molar), en este semiconductor el ancho de banda (E_g) para la recombinación directa varía de 1.424 eV ($x = 0$) hasta

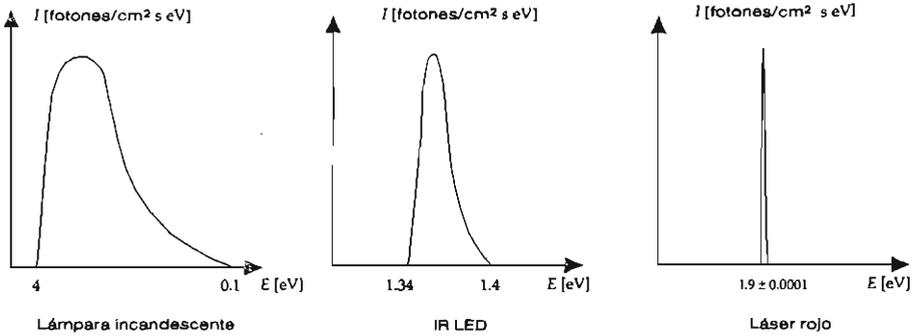


Figura III.46. Espectros de emisión para varias fuentes luminosas.

1.977 eV ($x = 0.45$). Cuando la cantidad molar de fósforo es superior a 45% no existe emisión de luz, porque la recombinación es indirecta. Los LED más eficientes se construyen en el infrarrojo, alrededor de 1.5 μm .

En la actualidad existe una intensa búsqueda de materiales que permitan la emisión en el azul. Existen LED azules comerciales, pero aún no hay láseres azules de semiconductores.

- Mucha mayor direccionalidad: la luz no diverge.
- En general es monocromático o tiende a serlo. En los sólidos es difícil que sea perfectamente monocromático.
- Una gran concentración de la intensidad: en un área menor a un micrómetro cuadrado puede concentrar toda su potencia.

El fenómeno láser se basa en la emisión estimulada de la radiación. En los procesos luminiscentes podemos tener dos situaciones al emitir luz: la emisión espontánea, como en el LED, donde no existe relación entre la luz producida por la recombinación de dos electrones, y la emisión estimulada, como en los láseres, donde los fotones tienen una relación muy particular, pues son una copia idéntica uno del otro. La figura III.48 trata de ilustrar estos casos.

Para que exista la emisión estimulada es necesario tener una inversión de población, esto es, tener más electrones en los niveles de la banda de conducción que en los niveles de la banda de valencia. Para obtener esta situación se bombean electrones de la banda de valencia a la banda de conducción.

En los semiconductores es posible obtener la inversión de población por medio de una unión $p-n$ polarizada en directa. La única diferencia con un LED es que aquí los semiconductores n y p están mucho más impurificados y las uniones deben hacerse con mucho cuidado, ya que las dimensiones del dispositivo son muy pequeñas. Además, es común emplear una caja de resonan-

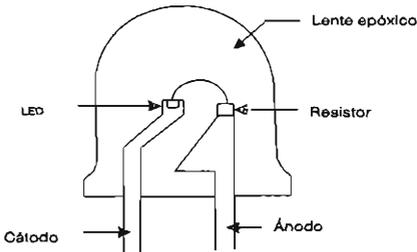


Figura III.47. Estructura de un LED.

El láser semiconductor de estado sólido es un caso muy especial de LED, donde el principio electroluminiscente es el mismo, sólo que la construcción del dispositivo suele ser más compleja para reunir las características que definen a los láseres.

Las principales características de un rayo láser comparadas con las de un rayo luminoso ordinario son las siguientes:

- Mucha mayor coherencia espacial y temporal, las ondas conservan la misma fase.

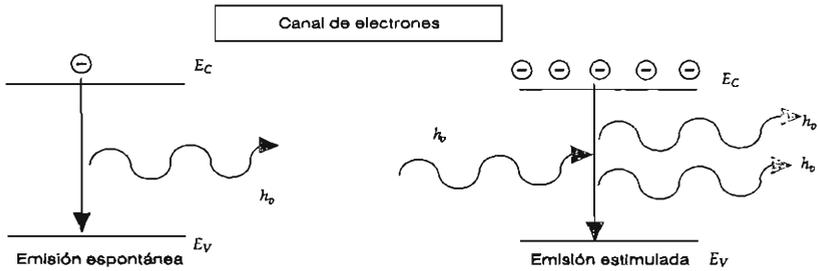


Figura III.48. Diagrama de la emisión espontánea y estimulada.

cia para fomentar la producción con una sola longitud de onda. En una caja de resonancia sólo sobrevivirán los fotones que tengan una longitud de onda que sea un múltiplo de la longitud de la caja de resonancia, la longitud del dispositivo. Para que sirva como caja de resonancia debe tener una geometría perfecta y se debe tener un control muy especial sobre los índices de refracción, para que los fotones se reflejen como si existieran espejos (los fotones deben ver la caja).

Las aplicaciones del láser semiconductor son

enormes. Entre las más interesantes por ahora se encuentran las comunicaciones ópticas y los innumerables productos de consumo, máquinas lectoras de discos compactos, impresoras, *scanners*, etc. Muchas de sus aplicaciones se deben a la alta eficiencia, bajo consumo de potencia y escaso peso comparados con los equivalentes de gas como los de helio-neón.

A continuación presentamos una lista incompleta de familias de dispositivos usados como fotodetectores:

Fotodetector	Proceso físico	Ganancia	Tiempo de respuesta (s)
Fotoconductor	Cambio en la conductividad.	1 a 10^6	10^{-3} a 10^{-8}
Unión <i>p-n</i>	Cambio de la concentración de portadores. Gran polarización en inversa, que reduce la capacitancia y el tiempo de transporte.	1	10^{-11}
Diodo <i>p-i-n</i>	Similar a la unión <i>p-n</i> , con la ventaja de tener el ancho de la zona desértica fija.	1	10^{-8}
Diodo de avalancha	Multiplica los fotones por ionización de electrones, multiplica el ruido porque amplifica igual a los electrones y a los huecos.	10^2 a 10^4	10^{-10}
Transistores	Cambio de la densidad de portadores en la base.	10^2	10^{-7}
Fotomultiplicadores de estado sólido	Como los diodos de avalancha, pero multiplican sólo un portador.	10^5	10^{-10}

En general, cualquier dispositivo con base en semiconductores es un candidato a fotodetector. Algunos de los parámetros que determinan la capacidad de un dispositivo en particular son:

- Sensibilidad a diferentes longitudes de onda.
- Ganancia electrónica: cuántos electrones se producen por cada 100 fotones (eficiencia cuántica).

- Ruido: a una temperatura dada, siempre existen electrones y huecos generados. Este parámetro indica cuántos fotones son indispensables para ser detectados.
- Velocidad de respuesta: indica qué tan rápido puede cambiar un evento óptico que cambia las características eléctricas.
- Tamaño.
- Requerimientos de potencia.
- Confiabilidad.

CONCLUSIONES

En este capítulo presentamos una introducción al funcionamiento de los dispositivos a base de semiconductores más comunes. Iniciamos con la unión de dos materiales arbitrarios que alcanzan el equilibrio al igualar sus niveles de Fermi;

el control de la corriente eléctrica en los dispositivos bipolares como la unión $p-n$, con efectos como rectificador o capacitor controlado por voltaje, hasta llegar al transistor de unión, que funciona como un resistor de resistencia variable por la inyección de portadores en la base. También vimos los dispositivos unipolares que basan su funcionamiento en el movimiento de la carga eléctrica, como el capacitor MOS, que se puede emplear para mover carga eléctrica o como sensor de imágenes, o el MOSFET, que controla la conductividad entre fuente y drenaje con el voltaje de compuerta.

Finalmente, se presentó una breve descripción de los dispositivos optoelectrónicos de semiconductores, los de generación de luz como en los LED, o láseres, o detección de luz como es el caso de los fotodetectores.

Ejercicios

Sección I

- III.1.1.* Demuestre la validez de la ecuación III.3.
- III.1.2. Demuestre la validez de la ecuación III.6.
- III.1.3. Demuestre la validez de las ecuaciones de la figura III.8.
- III.1.4. Demuestre la validez de la ecuación III.13.

Sección II

Mencione si las siguientes afirmaciones son falsas (F) o verdaderas (V):

- III.2.1. Siempre que se unen dos semiconductores, para llegar al equilibrio, fluyen electrones del semiconductor que tiene más electrones al que tiene menos. ()
- III.2.2. Existe potencial de contacto en las uniones $p-n$ de un mismo semiconductor. ()
- III.2.3. Al aumentar el potencial aplicado en una unión $p-n$ en inversa, aumenta el ancho de la zona desértica. ()
- III.2.4. Los electrones fluyen del material que tiene menor función trabajo al material que tiene mayor función trabajo para llegar al equilibrio. ()
- III.2.5. Fluye carga eléctrica a través de la estructura MOS ideal. ()

Sección III

Responda las siguientes preguntas:

- III.3.1. ¿Qué es un seudonivel de Fermi?
- III.3.2. ¿Qué es tiempo de vida de los portadores y cómo se modifica?
- III.3.3.* ¿Cómo puede medirse el potencial de contacto de una unión $p-n$ y por qué?

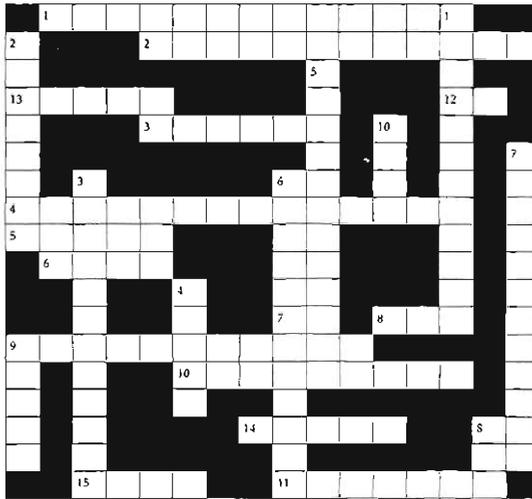
* Los ejercicios marcados con asterisco están resueltos al final del libro.

III.3.4. ¿Qué relación existe entre el voltaje de encendido de una unión $p-n$ y el potencial de contacto?

III.3.5. ¿Cómo puede realizarse un acceso eléctrico en un semiconductor tipo n ?

Sección IV

III.4. Resuelva el siguiente crucigrama. Cuando la respuesta es una palabra o símbolo compuesto, omita los espacios. El símbolo \leftarrow indica que debe escribirse primero la última letra y terminar con la primera.



Horizontales

1. Proceso físico en los semiconductores, que en ciertas aproximaciones se escribe fuera del equilibrio como n/τ .
2. Intervalo de tiempo que permanece un electrón en la banda de conducción, entre la generación y la recombinación.
3. Estructura básica de los semiconductores heterogéneos.
4. Distancia energética entre el nivel de Fermi y el nivel de vacío ϕ . \leftarrow
5. Semiconductor tal que en el equilibrio $n_0 \gg p_0$. \leftarrow
6. Semiconductor que supera al Si en las aplicaciones de alta frecuencia y electroópticas.
7. Valor mínimo de energía en la banda de conducción.
8. Diferencia de potencial entre los mínimos de las bandas de conducción de una unión $p-n$ en equilibrio.
9. Semiconductor en que las propiedades eléctricas pueden cambiar de región en región.
10. Zona de transición que se forma al unir los semiconductores n y p .
11. Semiconductor más importante en la electrónica moderna. \leftarrow
12. Símbolo del nivel de Fermi.
13. Energía que limita los estados vacíos y ocupados a bajas temperaturas.
14. Fuente de radiación monocromática, coherente y direccional.
15. Nombre de la primera letra del alfabeto griego.

Verticales

1. Magnitud energética esencial en el estudio de los semiconductores. Entre otras cosas, permite determinar n_0 y p_0 .
2. Corriente que se obtiene sólo en los semiconductores cuando existe un gradiente en su concentración de portadores.
3. Característica distintiva de las uniones $p-n$ y MOS, dQ/dV .
4. Nombre común de los dispositivos que presentan efecto rectificante.
5. Nombre del efecto que se tiene cuando la relación entre corriente y voltaje es muy asimétrica. ←
6. Nombre genérico de los semiconductores en los que el nivel de Fermi está dentro de la banda prohibida.
7. Portadores que existen en menor cantidad en un semiconductor en equilibrio.
8. Siglas de los dispositivos más comunes que controlan las propiedades eléctricas del semiconductor a través del efecto del campo eléctrico E.
9. Portador de carga que se mueve en la banda de valencia.
10. Material magnético con inducción magnética remanente muy elevada.

Sección V

III.5. Relacione las dos columnas.

1. Unión $p-n$ real.	() () ()	$\Delta E_f = 0$
2. MOS.	() () ()	$\phi_m > \phi_{sc}(p)$
3. Unión MS rectificante.	() () ()	n/τ
4. Unión MS óhmica.	() () ()	χ
5. Equilibrio.	() () ()	$n = N_c \exp[(E_f - E_c)/kT]$
6. Metal.	() () ()	$C = ct(1/\sqrt{V})$
7. Semiconductor.	() () ()	$J_o = q(D_n n_{p0}/L_n + D_p p_{n0}/L_p)$
8. Unión $p-n$ ideal.	() () ()	ψ_s
9. Recombinación.	() () ()	$J_o'' = \frac{q\sigma v_r N_i n_i W}{2}$
	() () ()	$\phi_m > \phi_{sc}(n)$
	() () ()	ϕ_{Bn}
	() () ()	V_{bi}
	() () ()	$\phi_m < \phi_{sc}(n)$
	() () ()	Inversión de población
	() () ()	ψ_B
	() () ()	$\phi_m < \phi_{sc}(p)$

Sección VI

III.6.* Estudio de la unión $p-n$.

Una unión $p-n$ puede ser empleada para detectar luz. Considere los siguientes dos dispositivos:

EG&G VACTEC, VTS3028	EG&G VACTEC, VTP100
Celda solar	Fotodiodo
$V_{oc} = 0.57 \text{ V @}100 \text{ mW/cm}^{-2}$	$V_{BR} = 30 \text{ V}$
$I_{sc} = 86 \text{ mA @}100 \text{ mW/cm}^{-2}$	$I_{sc} = 70 \text{ }\mu\text{A}$
Área = 392 mm ²	$I_0 = 30 \text{ nA}$
	$C_j = 50 \text{ pF @}3\text{V}$
	Área = 7.45 mm ²
	Responsividad = 0.05 AW ⁻¹ cm ² @ 940 nm
	$R_{sh} = 0.25 \text{ G}\Omega$

- Señale las diferencias entre una celda solar y un fotodiodo.
- Para la celda solar, indique la carga en la cual se obtendría la máxima potencia. Suponga una curva corriente *versus* voltaje de forma parabólica.
- Explique qué parámetros determinan la rapidez de respuesta de un fotodiodo.
- Para el fotodiodo, diseñe un circuito eléctrico que permita detectar un tren de pulsos en un osciloscopio con un MΩ de resistencia de entrada y una señal óptica de 100 μW de luz verde y 50% Duty Cycle señal cuadrada con un periodo de 1 μs.
- ¿Qué le pasará a la curva corriente *versus* voltaje de un diodo normal, si se aumenta la temperatura? Si se supone válido el modelo de Shockley, demuestre que al aumentar 5°C se duplica la corriente en inversa.

Sección VII

III.7.* Estudio de la estructura MOS.

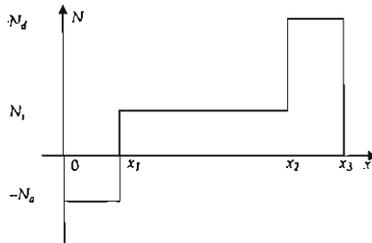
Con los siguientes materiales se construye un sistema MOS ideal.

Metal	Óxido	Semiconductor
Silicio tipo <i>n</i>	SiO ₂	Silicio tipo <i>p</i>
$q_x = 4.15 \text{ eV}$	$q_x = 0.95 \text{ eV}$	$q_x = 4.15 \text{ eV}$
$q\phi = 4.15 \text{ eV}$	$E_g \sim 8 \text{ eV}$	$q\phi = 5.0 \text{ eV}$
		$E_g = 1.12 \text{ eV}$

Suponga que V_a y V_{0x} se miden respecto al metal y que en el equilibrio $V_{0x} = 0.4 \text{ V}$

- Construya el diagrama de energía potencial electrostática en el equilibrio.
- Explique cómo alcanzaría la condición de fuerte inversión de población.

III.8. Dibuje el diagrama del campo eléctrico para la siguiente distribución de carga eléctrica.



RESPUESTAS A LOS EJERCICIOS SELECCIONADOS

I.1. Una gran cantidad de fuentes de radiación pertenecen al grupo de fuentes térmicas, que radian energía electromagnética en proporción directa a su temperatura. Las fuentes que presentan radiación de cuerpo negro pertenecen a este grupo y tienen las siguientes propiedades:

- Cumplen la ecuación de Stefan-Boltzmann:

$$W = \epsilon \sigma T^4$$

donde W es la irradiancia o potencia total radiada por unidad de área, σ representa la constante de Stefan ($\sigma = 5.67 \times 10^{-8} \text{Wm}^{-2}\text{K}^{-4}$), T es la temperatura absoluta y ϵ es la emisividad. La emisividad tiene valores de cero a uno, el valor de uno corresponde al emisor más eficiente de radiación térmica, y se conoce como cuerpo negro.

- Cumplen la ley de desplazamiento de Wien:

$$\lambda_m T = 2898 [\mu\text{mK}]$$

donde λ_m es la longitud de onda a la cual la potencia radiada es máxima a una temperatura dada, lo que indica un corrimiento del máximo de radiación hacia menores longitudes de onda cuando se aumenta la temperatura. Empleando esta ecuación y el hecho de que el máximo de la radiación solar se encuentra a $0.47 \mu\text{m}$, podemos decir que el Sol radia como si fuera un cuerpo negro a 6000K aproximadamente.

- Las dos ecuaciones anteriores fueron explicadas correctamente por Planck y están contenidas en su ecuación fundamental para la radiación del cuerpo negro:

$$W_\lambda = \frac{2\pi hc^2}{\lambda^5} \left(\frac{1}{\exp(hc/\lambda kT) - 1} \right) \left[\frac{\text{W}}{\text{m}^2 \text{m}} \right]$$

donde W_λ es la potencia radiada por unidad de área para una longitud de onda dada.

Así, un cuerpo negro cambia su color con la temperatura. Un cuerpo negro a 7°C tiene su máximo de radiación a una longitud de onda de $10.24 \mu\text{m}$; a 47°C tiene su máximo de radiación a $9.06 \mu\text{m}$; 727°C a $2.9 \mu\text{m}$; 3227°C a 724nm (color rojo) y 5727°C a 483nm (color azul).

- I.4. Para el modelo del átomo de hidrógeno se cumple que la energía permitida toma la siguiente forma:

$$E_n = \frac{-13.6}{n^2} [\text{eV}]$$

La frecuencia se puede calcular de la ecuación I.6.

a) Para ionizar un átomo de hidrógeno se requiere aplicar una energía suficiente para mover el electrón del nivel $n = 1$ hasta $n \rightarrow \infty$. Esto es, $E = 13.6 \text{ eV}$ y a esta energía le corresponde una frecuencia

$$\nu = \frac{E}{h} = \frac{13.6}{4.14 \times 10^{-15}} = 3.29 \times 10^{15} \text{ Hz}$$

que en el espectro electromagnético corresponde a radiación ultravioleta y no es visible.

b) Para desplazar el electrón del nivel $n = 2$ al nivel $n = 3$ se necesita aplicar una energía

$$E = 13.6 \left(\frac{1}{2^2} - \frac{1}{3^2} \right) = 1.89 \text{ eV}$$

A esta energía le corresponde una frecuencia $n = 4.57 \times 10^{14}$ Hz, que en el espectro electromagnético corresponde al rojo.

1.7. Para obtener esta relación, se puede iniciar expresando el valor promedio de la cantidad de movimiento lineal:

$$\langle p_x \rangle = \int_{-\infty}^{\infty} \varphi^*(x, t) \left(\frac{\hbar}{i} \frac{\partial}{\partial x} \right) \varphi(x, t) dx$$

Obteniendo la derivada respecto al tiempo se obtiene:

$$\frac{\partial \langle p_x \rangle}{\partial t} = \frac{\hbar}{i} \int_{-\infty}^{\infty} \frac{\partial \varphi^*}{\partial t} \frac{\partial \varphi}{\partial x} dx + \frac{\hbar}{i} \int_{-\infty}^{\infty} \varphi^* \frac{\partial}{\partial x} \frac{\partial^2 \varphi}{\partial t} dx$$

Sustituyendo la expresión para la derivada de φ en el tiempo, obtenida de la ecuación de Schrödinger, se tiene:

$$\frac{\partial \langle p_x \rangle}{\partial t} = \frac{\hbar}{i} \int_{-\infty}^{\infty} \left[-\frac{i\hbar}{2m} \frac{\partial^2 \varphi^*}{\partial x^2} + \frac{i}{\hbar} V(x) \varphi^* \right] \frac{\partial \varphi}{\partial x} dx + \frac{\hbar}{i} \int_{-\infty}^{\infty} \varphi^* \frac{\partial}{\partial x} \left[\frac{i\hbar}{2m} \frac{\partial^2 \varphi}{\partial x^2} - \frac{i}{\hbar} V(x) \varphi \right] dx$$

Pasando a una de las integrales los términos que incluyen la dependencia en la energía potencial V y realizando el álgebra inmediata, se tiene:

$$\frac{\partial \langle p_x \rangle}{\partial t} = -\frac{\hbar^2}{2m} \int_{-\infty}^{\infty} \frac{\partial}{\partial x} \left[\frac{\partial \varphi}{\partial x} \frac{\partial \varphi^*}{\partial x} - \varphi^* \frac{\partial^2 \varphi}{\partial x^2} \right] dx - \int_{-\infty}^{\infty} \varphi^* \frac{\partial V}{\partial x} \varphi dx$$

La primera es una integral exacta que al evaluar en $\pm\infty$ desaparece y la segunda integral es el valor promedio de la fuerza, que es lo que se quería demostrar.

$$\frac{\partial \langle p_x \rangle}{\partial t} = \left\langle -\frac{\partial V}{\partial x} \right\rangle = \langle F_x \rangle$$

1.8

a) $\psi = 1$ para todo $x > 0$. No es una ψ asociada a un sistema físico ya que no cumple el tercer principio de la mecánica cuántica $\psi(x \rightarrow \infty) = 0$.

b) $\psi = 0$. No es una ψ asociada a un sistema físico, ya que no representa ninguna partícula.

c) $\psi = ae^{-x}$. a deberá definirse de la condición de normalización, pero ésta sí satisface las propiedades de la función de onda para un sistema físico.

d) $\psi = ae^x$ para todo $x > 0$. No es una ψ asociada a un sistema físico, ya que no cumple el tercer principio de la mecánica cuántica $\psi(x \rightarrow \infty) = 0$, entre otros.

e) $\psi = \sin x$ para todo x . No es una ψ asociada a un sistema físico, ya que no cumple varios de los principios de la mecánica cuántica, el número tres en particular.

1.9

a) Empleando la misma ecuación que en el ejercicio 1.4, con $E = 2.4$ eV, ya que por definición de función trabajo define la energía de ionización en un sólido, se obtiene $v = 5.8 \times 10^{14}$ Hz.

b) La energía asociada a esta longitud de onda es: $E = \frac{hc}{\lambda} = \frac{4.14 \times 10^{-15} \text{ J} \cdot \text{s} \cdot 3 \times 10^8 \text{ s}^{-1}}{300 \times 10^{-9} \text{ m}} = 4.13$ eV, donde se ha empleado la relación de la velocidad de la luz $c = \lambda v$. Esta energía de la onda electromagnética al incidir en el sólido, 2.4 eV, se emplea para sacar al electrón del material, lo que produce una energía cinética resultante de 1.73 eV.

1.10. Un electrón acelerado por una diferencia de potencial de 25 kV adquiere una energía de 25 keV. La longitud de onda de De Broglie se obtiene con la ecuación del ejercicio anterior, y es:

$$\lambda = \frac{hc}{E} = \frac{4.14 \cdot 10^{-15} \cdot 3 \times 10^8}{25\,000} = 4.96 \times 10^{-11} \text{ m}$$

I.13

a) La irradiancia del láser ($I_{\text{láser}}$) como sólo incide en la pantalla del cine por un sistema de desviación, es igual a la potencia del láser ($P_{\text{láser}}$) sobre el área de la pantalla. La irradiancia de la fuente isotópica incandescente ($I_{\text{lámpara}}$), radia su potencia ($P_{\text{lámpara}}$) en un ángulo sólido de 4ϕ de la que se emplea sólo la potencia que incide en la pantalla, que se extiende en un ángulo sólido Ω_1 . En realidad, el sistema de proyección aumenta en gran medida la eficiencia.

$$\Omega_1 = 4 \arctan \frac{4 \times 8}{2.40 \cdot \sqrt{4^2 + 8^2} + (2.40)^2} = 0.0199$$

$$I_{\text{láser}} = \frac{P_{\text{láser}}}{4 \cdot 8}; \quad I_{\text{lámpara}} = \frac{P_{\text{lámpara}} \cdot \Omega_1}{4 \cdot \pi \cdot 4 \cdot 8}; \quad \frac{P_{\text{lámpara}}}{P_{\text{láser}}} = \frac{4 \cdot \pi}{\Omega_1} = 600$$

b) De acuerdo con el resultado del inciso anterior, la pantalla tiene una radiancia L_{pantalla} :

$$L_{\text{pantalla}} = k \frac{I_{\text{lámpara}}}{2 \cdot \pi} = \frac{P_{\text{lámpara}} \cdot \Omega_1}{8 \cdot \pi^2 \cdot 4 \cdot 8} \text{ [Wm}^{-2} \text{ sr}^{-1}\text{]}$$

donde k es la eficiencia de reflexión de la pantalla. La pantalla radia con un ángulo sólido igual a 2π , ya que sólo radia hacia el frente. La cantidad de potencia que puede recibir un observador ($\Phi_{\text{observador}}$) dependerá del área de radiación y del ángulo sólido (Ω_2) que forman el observador y la pantalla.

$$\Omega_2 = 4 \arctan \frac{4 \cdot 8}{2.20 \cdot \sqrt{4^2 + 8^2} + (2.20)^2} = 0.078$$

$$\Phi_{\text{observador}} = L_{\text{pantalla}} \cdot 4 \cdot 8 \cdot \Omega_2 = \frac{k \cdot P_{\text{lámpara}} \cdot \Omega_1 \cdot \Omega_2}{8 \cdot \pi^2} \approx 5 \times 10^{-6} \cdot P_{\text{lámpara}}$$

II.4. Para determinar las temperaturas críticas, es necesario contar con expresiones que muestren su dependencia con la temperatura.

$$T_s = \frac{E_C - E_D}{k \ln \left(\frac{N_C}{N_D} \right)} \quad T_i = \frac{E_g(T_i)}{k \ln \left(\frac{N_C \cdot N_V}{N_D^2} \right)}$$

donde es posible suponer que $E_C - E_D = 0.012 \text{ eV}$ y, además, que es independiente de la temperatura, $k = 8.614 \times 10^{-5} \text{ eVK}^{-1}$, mientras que N_C y N_V muestran una dependencia de la temperatura (ecuaciones II.25 y II.26) de la forma $\propto T_i^3$. Con los valores para estas funciones a temperatura ambiente, es posible obtener una función con la temperatura de la siguiente forma:

$$N_C = 5.39 \times 10^{15} \cdot T_i^3 [\text{cm}^{-3}]$$

Respuestas a los ejercicios seleccionados

$$N_v = 2 \times 10^{15} \cdot T_2^3 [\text{cm}^{-3}]$$

$$N_D = 1 \times 10^{18} \text{ cm}^{-3}$$

$$E_g = 1.17 - \frac{4.75 \times 10^{-4} \cdot T^2}{T + 636}$$

Con estas ecuaciones se obtienen las siguientes relaciones que se pueden resolver mediante algún método numérico:

$$T_S (\ln T_i^{3/2} - 5.223) = 139.3, \quad T_S = 90.6 \text{ K}$$

$$T_i (\ln T_i^3 - 11.44) = 11\,609 \cdot \left(1.17 - \frac{4.75 \times 10^{-4} \cdot T_i^2}{T_i + 636} \right), \quad T_i = 1\,054.4 \text{ K}$$

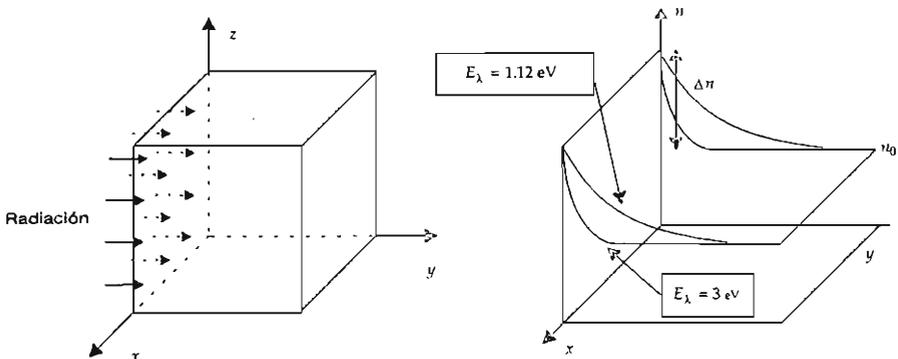
II.6. Sí, ya que al formar el sólido de hidrógeno es posible esperar que se forme una banda s que estaría ocupada a la mitad, ya que existe un electrón por átomo y en los niveles s pueden estar dos electrones. Lo difícil es formar el hidrógeno sólido.

Los metales monovalentes, litio, sodio, potasio, etc., son los más simples de estudiar y su estructura de bandas, de una manera simplista, se puede ver como una banda de conducción que se forma del amontonamiento de los niveles s .

Los elementos, cobre Cu (NA = 29), plata Ag (NA = 47) y oro Au (NA = 79) en su estado base también tienen un electrón por átomo en una banda medio llena, obtenida de los niveles $4s$, $5s$ y $6s$ respectivamente, lo que los caracteriza como buenos conductores de la electricidad, de manera similar al caso anterior.

Un grupo de metales más complicado son los metales de transición, hierro, cobalto y níquel, donde existe una superposición de niveles s y d al formar la banda de conducción, que les transfiere menor conductividad eléctrica y propiedades magnéticas interesantes.

II.11 Cualitativamente, si el cubo de silicio está siendo iluminado por una de sus caras, se espera que la generación sea mayor en la superficie expuesta y cada vez menor dentro del material, ya que la absorción del material (generación de pares electrón-hueco) hace disminuir la intensidad. Si la energía de la radiación aumenta, la absorción es mucho mayor, lo que evitaría que la radiación penetre en el semiconductor. El gradiente de concentración estaría determinado por la densidad de portadores en la superficie del material y en la difusión de los portadores.



II.15. Ciclo de histéresis magnética.

El área del ciclo de histéresis magnética es proporcional al trabajo por unidad de volumen necesario para cambiar el material ferromagnético alrededor de un ciclo y es disipada en forma de calor.

En el secundario y de acuerdo con las propiedades del transformador, se tiene que:

$$V_s = N_s \frac{d\Phi}{dt} = N_s A \frac{dB}{dt}$$

donde N_s es el número de vueltas en el secundario, A es el área del núcleo, y Φ es el flujo magnético a través del secundario.

En la malla del secundario se tiene, al aplicar la ley de Kirchhoff de voltajes:

$$V_s = V_C + RI = \frac{Q}{C} + RC \frac{dV_C}{dt}$$

Si se cuenta con un capacitor suficientemente grande, e igualando las dos ecuaciones anteriores, se tiene:

$$B = \frac{RC}{N_s A} V_C$$

Para el primario, aplicando la ley de Ampère, se tiene:

$$H = \frac{N_p I_p}{L}$$

donde N_p es el número de vueltas en el primario, I_p es la corriente del primario y L es la circunferencia del núcleo. Empleando la ley de Ohm al resistor del primario, se tiene:

$$H = \frac{N_p}{LR_p} V_R$$

III.1.1. El equilibrio se alcanza cuando no existe corriente neta en el dispositivo:

$$J_{A \rightarrow B} = J_{B \rightarrow A}$$

Con las ecuaciones III.1 y III.2, y III.25 se obtiene el resultado esperado:

$$n_{0A} \exp\left(-\frac{E_0}{kT}\right) = n_{0B}$$

$$n_{0A} = N_C \exp\left(-\frac{E_{CA} - E_{FA}}{kT}\right) \text{ y } n_{0B} = N_C \exp\left(-\frac{E_{CB} - E_{FB}}{kT}\right)$$

de donde se puede obtener:

$$-E_0 - E_{CA} + E_{FA} = E_{CB} + E_{FB} \text{ y } E_0 = E_{CB} - E_{CA}$$

que produce finalmente $E_{FA} = E_{FB}$.

III.3.3. El potencial de contacto no se puede medir con un multímetro, ya que el multímetro sólo mide diferencias de niveles de Fermi y los sistemas en equilibrio no tienen diferencia de niveles de Fermi. Éste desaparece al desplazar carga eléctrica para alcanzar el equilibrio. Existen métodos indirectos que miden el potencial de contacto; particularmente, la capacitancia en inversa es útil para medir y caracterizar el contacto entre los semiconductores, como se señala en la figura III.15.

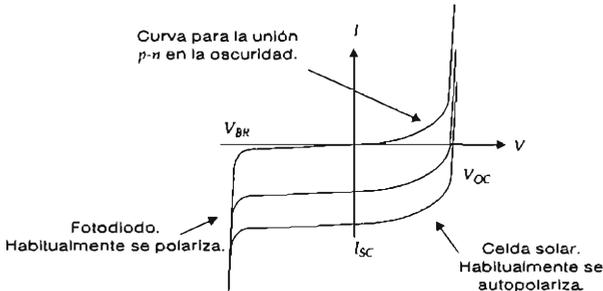
Las curvas reales de capacitancia *versus* voltaje en inversa son un tanto diferentes de la figura III.15, en virtud de que la aproximación de unión abrupta no se cumple en realidad, y la capacitancia disminuye al aumentar el voltaje en inversa pero de una manera diferente.

La capacitancia a voltaje cero es del orden de 5 pF/mm² en los dispositivos rápidos en área pequeña, a 500 pF/mm² en dispositivos de potencia de área grande.

III.6. Unión p-n.

a) Los fotodiodos y las celdas solares son uniones p-n que por efecto de la luz modifican la curva de corriente *versus* voltaje. Los fotones de la luz generan pares electrón-hueco que por efecto del campo eléctrico interno de la unión p-n separan a los portadores. Los electrones se desplazan al semiconductor tipo n y los huecos al semiconductor tipo p.

Si las terminales de la unión p-n están desconectadas se genera una autopolarización conocida como voltaje de circuito abierto V_{oc} . Si existe un alambre entre las terminales de la unión p-n pasa una corriente de cortocircuito I_{sc} . Estos parámetros son función de la irradiancia y se caracterizan por su responsividad.



b) En la celda solar, se tiene que $V_{oc} = 0.57 \text{ V}$ y $I_{sc} = 86 \text{ mA} @ 100 \text{ mW/cm}^2$ en un área de 92 mm^2 , lo que produce en la aproximación de una función parabólica la curva siguiente. Se puede ignorar el signo si se reconoce que la celda solar aporta potencia al circuito:

$$I = 86 - 264.7 V^2 \text{ [mA]}$$

La potencia máxima se calcula de la manera habitual, ya que en la corriente de cortocircuito y en el voltaje de circuito abierto es cero.

La potencia máxima se obtiene cuando $V = 0.3291 \text{ V}$, que se alcanza con un resistor de 5.74Ω .

c) El tiempo de respuesta de un fotodiodo se define como el tiempo que toma a los portadores generados por la luz alcanzar la unión p-n.

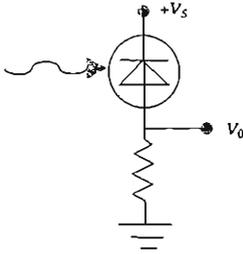
Básicamente, existen tres componentes que determinan esta cantidad:

- Tiempo de colección: tiempo que les toma a los portadores desplazarse en la zona desértica. Este número es muy pequeño y por lo común es del orden de un nanosegundo.
- Tiempo carga: es el tiempo necesario para cargar o descargar el capacitor de la zona desértica, y es proporcional a la multiplicación de la capacitancia de la unión y la resistencia de carga. Si $C = 100 \text{ pF}$ y $R = 50 \Omega$ $t = 5 \text{ ns}$, si $R = 2 \Omega$, entonces $t = 200 \text{ ns}$.
- Tiempo de difusión: es el tiempo necesario para que los portadores generados fuera de la zona desértica alcancen la zona desértica. Si la radiación se absorbe poco, como en el infrarrojo, este tiempo llega a ser de $2 \mu\text{s}$; si se tiene gran absorción, como en el color rojo, se puede llegar a tener una respuesta de 50 ns .

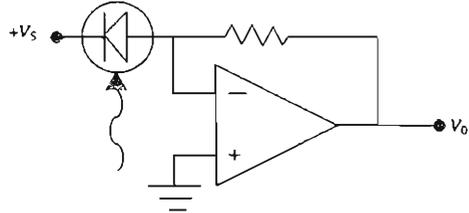
La velocidad de respuesta se incrementa si se emplea un dispositivo de área pequeña, una pequeña capacitancia de unión, un voltaje en inversa grande, radiación visible y una pequeña resistencia de carga.

d) En estado estable la corriente del fotodiodo en la oscuridad es de 30 nA . Al iluminarlo con $200 \mu\text{W}$ de radiación visible (ya que en promedio recibe 100 mW), probablemente mande la corriente del fotodiodo a saturación, ya que

al multiplicar la responsividad por la potencia luminosa y dividir entre el área se obtienen $134 \mu\text{A}$. La velocidad para cambiar de 30 nA a $100 \mu\text{A}$ está limitada únicamente por el tiempo de carga; éste se minimiza al polarizar la unión en inversa, cerca de V_{BR} y poniendo una carga de $\tau/C \sim 500 \text{ ns}/50 \text{ pF} \sim 10 \text{ k}\Omega$ o menor en paralelo con el fotodiodo. En un osciloscopio podría medirse la caída de potencial en un resistor de $1 \text{ k}\Omega$ debido a la fotocorriente, sin cambiar el tiempo de respuesta. Esto produciría cambios de 100 mV a $30 \mu\text{V}$. Una alternativa más versátil se consigue con un amplificador de corriente a voltaje, como se ilustra en la siguiente figura:



Convertidor de corriente a voltaje



Amplificador de corriente a voltaje

e) El coeficiente de temperatura de un dispositivo se define como:

$$TC_f = \frac{1}{X} \frac{dX}{dT}$$

donde X puede ser resistencia, capacitancia u otra variable.

Una unión $p-n$ en inversa tiene una corriente con la siguiente expresión:

$$I_s = qA \left(\frac{D_n n_i^2}{L_n N_A} + \frac{D_p n_i^2}{L_p N_D} \right) \approx \frac{qADn_i^2}{LN} = BT^3 \exp\left(-\frac{E_g}{kT}\right)$$

donde B es una constante independiente de la temperatura. La dependencia con la temperatura es un recuento de la densidad de portadores intrínsecos.

$$TC_f \equiv \frac{1}{I_s} \frac{dI_s}{dT} = \frac{3}{T} + \frac{E_g}{kT^2} = 0.154 \text{ [K}^{-1}\text{]}$$

donde se ha puesto $T = 300 \text{ K}$, $E_g = 1.12 \text{ eV}$ y $kT = 0.0259 \text{ eV}$. La corriente en inversa se duplica cuando TC_f toma el valor de uno, o sea, cada 6.5°C .

Una unión $p-n$ en directa tiene una corriente con la siguiente expresión:

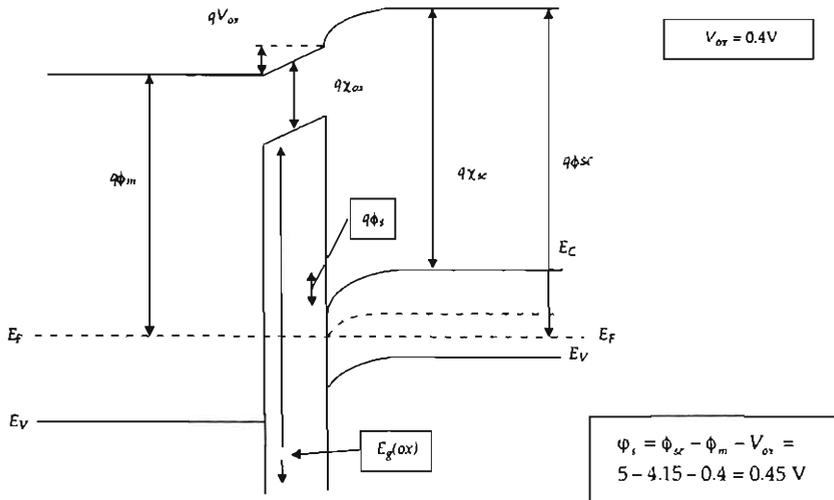
$$I = I_s \exp\left(\frac{qV_0}{kT}\right) = BT^3 \exp\left(\frac{E_g - qV_0}{kT}\right)$$

$$TC_f \equiv \frac{1}{I} \frac{dI}{dT} = \frac{3}{T} + \frac{E_g - qV_0}{kT^2} = 0.077 \text{ [K}^{-1}\text{]}$$

donde se ha puesto $T = 300 \text{ K}$, $E_g = 1.12 \text{ eV}$, $kT = 0.0259 \text{ eV}$ y $V_0 = 0.6 \text{ V}$. La corriente en directa se duplica cuando TC_f toma el valor de uno, o sea, cada 13°C .

III.7. MOS

a)



b) Se alcanza una fuerte inversión cuando $|\varphi_s| = 2\varphi_B$. En este caso

$$\varphi_B = \varphi_{sc} - \chi_{sc} - E_g/2q = 5 - 4.15 - 1.12/2 = 0.29V.$$

Para tener una fuerte inversión es necesario aplicar un voltaje positivo al metal respecto al semiconductor, para lograr que las bandas del semiconductor se doblen 0.13 V más.

Constantes

Nombre	Símbolo	Valor
Constante de Avogadro	N_{avo}	$6.022 \times 10^{23} \text{ mol}^{-1}$
Constante de Boltzmann	k	$1.38 \times 10^{-23} \text{ J/K}$
Carga elemental	q	$1.602 \times 10^{-19} \text{ C}$
Masa del electrón	m	$0.911 \times 10^{-30} \text{ Kg}$
Electrón-volt	eV	$1 \text{ eV} = 1.60218 \times 10^{-19} \text{ J}$
Permeabilidad en el vacío	μ_0	$1.2566 \times 10^{-8} \text{ H/cm}$
Permisividad del vacío	ϵ_0	$8.854 \times 10^{-14} \text{ F/cm}$
Constante de Planck	h	$6.626 \times 10^{-34} \text{ Js}$
Velocidad de la luz en el vacío	c	$2.99792 \times 10^{10} \text{ cm/s}$
Voltaje térmico (300 K)	kT/q	0.0259 V

Propiedades de Ge, Si y GaAs a 300 K

Propiedad	Ge	Si	GaAs
Átomos/cm ³	4.42×10^{22}	5×10^{22}	4.42×10^{22}
Densidad (g/cm ³)	5.3267	2.328	5.32
Constante dieléctrica	16	11.9	13.1
N_c (cm ⁻³)	1.04×10^{19}	2.8×10^{19}	4.7×10^{17}
N_v (cm ⁻³)	6.0×10^{18}	1.04×10^{19}	7×10^{18}
Masa efectiva electrones m^*/m_0	1.64	0.98	0.067
Masa efectiva huecos m^*/m_0	0.044	0.16	0.082
Afinidad electrónica χ [V]	4	4.05	4.07
Ancho de banda E_g [eV]	0.66	1.12	1.424
Portadores intrínsecos (n, cm^{-3})	2.4×10^{13}	1.45×10^{10}	1.79×10^6
Resistividad intrínseca ρ ($\Omega \text{ cm}$)	47	2.3×10^5	10^8
Tiempo de vida minoritario (s)	10^{-3}	2.5×10^{-3}	10^{-8}
Movilidad (cm ² /Vs), e y h	3 900 y 1 900	1 500 y 450	8 500 y 400

BIBLIOGRAFÍA

- DeWitt, G. Ong, *Modern MOS technology: Processes, devices and design*, McGraw-Hill, 1986.
- Feynman, R.P. et al., *The Feynman lectures on physics I, II y III*, Addison-Wesley, 1966.
- Jaeger, R.C., *Introduction to microelectronics fabrication*, Addison-Wesley Publishing Company, 1990.
- McKelvey, J.P., *Solid state and semiconductors physics*, Harper & Row Publishers, 1966.
- Muller, R.S. y T.I. Kamins, *Device electronics for integrated circuits*, 2a. ed., John Wiley & Sons, 1986.
- Sze, S.M., *Physics of semiconductor devices*, 2a. ed., John Wiley & Sons, 1981.
- , *VLSI technology*, 2a. ed., McGraw-Hill, 1988.
- Wert, C.A. y R.M. Thomson, *Physics of solids*, 2a. ed., McGraw-Hill, 1970.
- Wilson, J. y J.F.B. Hawkes, *Optoelectronics*, 2a. ed., Prentice Hall, 1989.
- Yang, E.S., *Microelectronics devices*, McGraw-Hill, 1988.

ÍNDICE

<i>Agradecimientos</i>	7
<i>Presentación</i>	9
<i>Introducción</i>	11
<i>Capítulo I. Conceptos preliminares</i>	13
Física submicroscópica	13
Principio de incertidumbre	14
Hipótesis de De Broglie	15
Hipótesis de Max Planck	16
Nomenclatura óptica	17
Polarización de la luz	18
Mecánica cuántica	19
Ecuación de Schrödinger	19
Cuantización de la energía	21
Efecto túnel	23
Conclusiones	24
Ejercicios	24
<i>Capítulo II. Materiales para la electrónica</i>	27
Bandas de energía	29
Metales	36
Efecto Hall	39
Aislantes	40
Estadística de Boltzmann	40
Semiconductores	42
Estructura de bandas	43
Densidad de portadores libres	44
Conductividad	53
Ecuación de continuidad: generación y recombinación	55
Propiedades ópticas y magnéticas	58
Propiedades ópticas	58
Propiedades magnéticas	59

Tecnología	6
Conclusiones	6
Ejercicios	6
<i>Capítulo II. Estructuras electrónicas básicas</i>	7
Diodo de unión	7
Modelo de Shockley	7
Capacitancia de la zona desértica	8
Ruptura de la unión $p-n$	8
Transitorios de la unión $p-n$	8
Contactos óhmicos y rectificantes	8
Diagramas de energía en los contactos rectificantes	8
Diagramas de energía en los contactos óhmicos	8
Capacitor metal-óxido-semiconductor	8
Diagrama de energía del capacitor MOS	9
Transistor bipolar de unión (BJT)	9
Ecuaciones de transporte	9
Ganancia en corriente	9
MOSFET	10
Ecuaciones de transporte	10
Funcionamiento	10
LED, láser y fotodetectores	10
Conclusiones	11
Ejercicios	11
<i>Respuestas a ejercicios seleccionados</i>	11
<i>Constantes</i>	12
<i>Bibliografía</i>	12

Electrónica física

se terminó de imprimir en febrero de 1996
en los talleres de Editorial Ducere,
S.A. de C.V., Rosa Esmeralda 3 bis,
col. Molino de Rosas, 01470 México, D.F.

El tiro consta de 1 000 ejemplares más sobrantes
para reposición.

La composición tipográfica, la formación y el cuidado
editorial estuvieron a cargo de Sans Serif Editores,
S.A. de C.V., telfax 674 60 91.

Otros títulos en esta colección

Margarita Alegría, *Curso de lectura y redacción*
Victor Cuevas Ahumada, *México: revolución
de las políticas de estabilización*

Adolfo Jiménez Otamendi, *Átomos, enlaces
y reacciones.*

Abelardo Mariña Flores, *Insumo-producto*

Saúl Holguín et al., *Química inorgánica II*

Patrick Staelens, *El trabajo de los menores*

Luis Rodríguez, *El diseño preindustrial*

José Dolores Juárez Cervantes, *Sistemas
de distribución de energía eléctrica*

Textos de próxima aparición

Mariem Henaine-Abed, *Planación y control
de la producción*

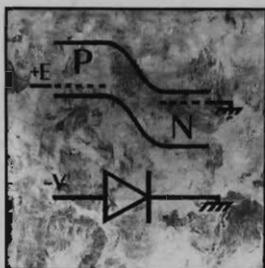
José Vega Luna y Gerardo Salgado Guzmán,
Prácticas de laboratorio de sistemas digitales

Ana Lilia Laureano, *Programación orientada
a objetos: un enfoque con tipos abstractos*

Juan González Márquez, *Introducción al derecho
bancario mexicano*

Violeta Múgica y José de Jesús Figueroa,
Contaminación ambiental, causas y control

Adalberto Cantú Chapa, *Electrónica II. Análisis
de diseño con diodos y transistores*



Este libro se elaboró para los estudiantes de ingeniería electrónica, con el propósito de facilitar el proceso de enseñanza-aprendizaje de los principios físicos, los materiales y los dispositivos semiconductores básicos empleados en electrónica. Partiendo de ideas simples de la física moderna, se explica la formación de las bandas de energía, lo que permite describir los materiales y sus propiedades (particularmente de los semiconductores). Se presentan algunos dispositivos semiconductores como el transistor bipolar (BJT) y el transistor de efecto de campo (MOSFET). A lo largo del texto se ofrecen numerosas analogías con el fin de formar ideas claras.

El material está asociado normalmente con una introducción al diseño y la fabricación de circuitos integrados, lo que en ciertos países es una elección natural. Aquí se ha preferido presentar el material como sustento de los conocimientos básicos de la electrónica; se busca con ello una aplicación más juiciosa de los dispositivos electrónicos y hacer más eficiente su uso. Desde esta perspectiva, se espera que el libro sea una buena alternativa a los textos comúnmente utilizados en los Estados Unidos o sus traducciones.

Finalmente, como el tema aquí tratado forma parte de las ciencias de la ingeniería más que de la ingeniería aplicada, y es material básico en el que se sustenta una parte del desarrollo de la electrónica, con toda seguridad esta obra ayudará a la comprensión de algunos de los vertiginosos avances que se han dado y de los que se darán en el futuro.