

Libros de **Cátedra**

Introducción a la física de semiconductores y dispositivos electrónicos

Eitel Leopoldo Peltzer y Blancá

FACULTAD DE
INGENIERÍA

e
exactas


EDITORIAL DE LA UNLP



UNIVERSIDAD
NACIONAL
DE LA PLATA

INTRODUCCIÓN A LA FÍSICA DE SEMICONDUCTORES Y DISPOSITIVOS ELECTRÓNICOS

Eitel Leopoldo Peltzer y Blancá

Facultad de Ingeniería



UNIVERSIDAD
NACIONAL
DE LA PLATA


EduLP
EDITORIAL DE LA UNLP

*Deseo dedicar este trabajo a quienes con
su esfuerzo me permitieron llegar a este punto,
mi madre y mi padre.
También deseo mencionar en esta dedicatoria a mis hijos,
Julieta, Eitel, Carolina y Federico,
los que permanentemente apoyaron mi tarea.
Especialmente agradezco y dedico este libro
a mi esposa Alicia, que tanto me alentó
y que después de esperar tanto y con tantas
ilusiones puestas, en función de mi anhelo para que se
editara, no va a poder ver la edición de este libro. Todo mi
amor, mi permanente recuerdo y mi admiración a mi esposa,
compañera de mi vida por 48 exquisitos años.*

Agradecimientos

Deseo mencionar y agradecer la tarea realizada por los revisores, a la Ingra. M. González y al Dr. L. Errico, los que pacientemente leyeron los borradores e hicieron muy importantes críticas.

También debo agradecer a los Dres. Ariel Cédola y Marcelo Cappelletti por sus importantes contribuciones a temas específicos desarrollados al final de este libro.

*La inteligencia consiste no sólo en el conocimiento,
sino también en la destreza de aplicar
los conocimientos en la práctica.*
ARISTÓTELES, 384-322AC

Índice

Prólogo	10
----------------	----

PRIMERA PARTE

Herramientas básicas	13
-----------------------------	----

Breve historia	14
----------------	----

Capítulo I

Introducción a la mecánica cuántica	16
-------------------------------------	----

I.1. Cuantización de la energía	16
---------------------------------	----

I.2. Los átomos y los modelos	18
-------------------------------	----

I.3. Dualidad onda – partícula. Teoría de De Broglie	22
--	----

I.4. Experimento de Davisson – Germer y Thomson	24
---	----

I.5. De Broglie y las ondas	26
-----------------------------	----

Referencias	29
-------------	----

Capítulo II

Ecuación de Schrödinger. Casos particulares	30
---	----

II.1 Ecuación de Schrödinger para electrones	31
--	----

II.2 Aplicaciones de la Ecuación de Schrödinger	33
---	----

II.2.1 Pozo infinito	34
----------------------	----

II.2.2 Pozo finito	37
--------------------	----

II.2.3 Barrera de potencial. Efecto Túnel	42
---	----

Referencias	47
-------------	----

Capítulo III

Bandas de energía en sólidos. Modelos cuánticos. Movimiento de electrones en redes periódicas	48
---	----

III.1. El modelo del electrón libre	50
-------------------------------------	----

III.2. Densidad de estados	53
----------------------------	----

III.3. Teorema de Bloch	57
-------------------------	----

III.4. El modelo de Kronig-Penney	59
III.5. El concepto de hueco	67
Referencias	71

Capítulo IV

Breves nociones de mecánica estadística	72
IV.1. Estadística clásica, Maxwell-Boltzmann	73
IV.2. Estadística cuántica, Fermi-Dirac	75
IV.3. Diferencias entre estadísticas y ocupación de estados	78
Referencias	83
Apéndice IV-1. Estadística de Maxwell-Boltzmann. Probabilidad	84
Referencias: Apéndice	88

SEGUNDA PARTE

Herramientas aplicadas	89
-------------------------------	----

Capítulo V

Descripción de un sólido	90
V.1. Descripción de las estructuras cristalinas	91
V.2. Características de los sólidos	94
V.3. Niveles de energía en los sólidos. Estructura de banda	97
V.4. Aisladores, semiconductores y metales	99
Apéndice V-1. Celda en el espacio recíproco	111
Referencias	113

Capítulo VI

Semiconductores	114
VI.1. Semiconductores, introducción de impurezas	115
VI.2. Determinación de los niveles de Fermi en semiconductores intrínsecos y extrínsecos	118
VI.3. Variación con la temperatura	122
VI.4. Cuasi niveles de Fermi	126
VI.5. Neutralidad de carga	128
Referencias	130

Capítulo VII

Conducción en semiconductores	131
VII.1. Movimiento de portadores. Aplicación de campo externo	132
VII.2. Recombinaciones, generación	137
VII.3. Difusión. Relación de Einstein	142

VII.4. Ecuación de continuidad	145
VII.5. Portadores minoritarios	148
VII.6. Longitud de difusión	149
Referencias	151

TERCERA PARTE

La aplicación	152
---------------	-----

Capítulo VIII

Junturas	153
VIII.1. Homojunturas	154
VIII.1.a. La Unión de semiconductores tipo- <i>p</i> y tipo- <i>n</i>	154
VIII.1.a1. El potencial de contacto, V_{bi}	156
VIII.1.a2. El campo eléctrico, \mathcal{E} . Límites y ancho de la región de vaciamiento	159
VIII.2. La unión <i>pn</i> bajo la acción de un potencial externo	165
VIII.2.a. Polarización directa	166
VIII.2.b. Polarización inversa	169
VIII.2.c. La capacidad de la unión	171
VIII.3. Ecuación de continuidad. Distribución de portadores minoritarios	174
VIII.4. Corrientes en la juntura <i>pn</i>	177
VIII.5. El factor temperatura	182
VIII.6. Heterojunturas	185
VIII.7. Unión metal – semiconductor. Unión Schottky	185
VIII.7.a. Unión metal – semiconductor. Aplicación de un potencial externo	190
VIII.7.b. Unión Schottky. Corrientes a través de la unión	192
VIII.8. Unión no rectificante (Metal - semiconductor) Contacto Óhmico	195
VIII.9. Heterojuntura formada por la unión de semiconductores	197
Apéndice VIII-1. Emisión termoiónica. Ecuación de Richardson-Dushman	203
Referencias	206

CUARTA PARTE

Los dispositivos	207
------------------	-----

Capítulo IX

Dispositivos	208
IX.1. El transistor	208
IX.1.a. Las corrientes	212
IX.2. Diodo Túnel (Diodo Esaki)	215
IX.3. Diodo Zener	219

IX.4. Celda Peltier	220
IX.5. Celda solar. Generalidades	224
IX.5.a. Celda solar con perovskitas	228
IX.5.a1. Referencias: Celda solar con perovskitas	230
IX.5.b. Celda solar con puntos cuánticos	232
IX.5.b1. Referencias: Celda solar con puntos cuánticos	237
IX.6. Diodo Laser	240
Referencias	245
Epílogo	246
El autor	247

NOTA: todos los gráficos mostrados a través del libro han sido realizados por el autor del texto correspondiente. Todas las figuras de personalidades han sido extraídas de la carpeta de Imágenes de Google, lugar de libre circulación.

Prólogo

La Física, al igual que la Química y la Matemática son ciencias básicas, esto quiere decir, que ellas aportan sus estructuras, conocimientos, desarrollos, métodos, en general, herramientas, para que las otras ciencias puedan lograr y fundamentar sus avances.

En este sentido, la Ingeniería necesita de los aportes de la Física para abordar sus problemas y resolverlos. Esto implica que se debería entender claramente cual es la importancia que la Física tiene en el mundo de la Ingeniería, en donde sus desarrollos están basados en el campo de esta ciencia. Esta condición hace que sea absolutamente indispensable el conocimiento de los principios fundamentales de la Física, para lograr una mejor aplicación de estos en las diferentes ramas de la Ingeniería.

La Física Clásica explica casi todos los fenómenos de las diferentes ramas de la Ingeniería: Civil, Construcciones, Hidráulica, Aeronáutica, Agrimensura, etc. La que no puede explicar, pues se mueve en otra escala de tiempo y longitudes, es la correspondiente a la electrónica. Los efectos que se producen en los materiales utilizados en este campo, solamente pueden ser explicados por medio de una teoría que se mueva en un mundo microscópico, la *mecánica cuántica*, la cual gobierna la física de las escalas muy pequeñas. Se está hablando de dimensiones del orden de los nanómetros (10^{-9}m) y partículas que se desplazan como ondas. Por lo cual, soslayar el estudio de la mecánica cuántica en la electrónica, es impedir la incorporación del conocimiento que da lugar al propio objeto estudiado. Intentar que este concepto sea entendido, ha sido uno de los objetivos de este libro.

El desarrollo del presente texto tiene otra motivación muy clara, la de generar una obra que cubra las necesidades mínimas del estudiante de ingeniería y de ciencias, para que puedan alcanzar a tener una noción sobre los aspectos científicos que aportan al conocimiento de los materiales y los dispositivos. El libro ha sido desarrollado en cuatro partes, cada una de ellas avanza sobre diferentes temas, haciendo aportes para lograr un saber final que permita entender los procesos que se llevan a cabo en los materiales y/o dispositivos. La primera parte está compuesta por lo que se ha llamado **herramientas básicas**. Estas aportan el conocimiento sobre los fundamentos de los temas esenciales que se deben incorporar para avanzar sobre el objetivo final: materiales semiconductores y dispositivos.

Los puntos a tratar son: la mecánica cuántica y sus aplicaciones y las estadísticas clásica y cuántica. Sin estos conocimientos sería muy difícil entender el comportamiento microscópico de los materiales. Los estudiantes de ciencias pueden saltar estos capítulos, los cuales seguramente han visto con mayor profundidad y detalle que lo que se expone aquí.

Luego el texto conducirá hacia las **herramientas aplicadas**, donde se hace uso de los conocimientos adquiridos previamente, para la descripción de los primeros modelos de materiales semiconductores y metales. Finalmente se llegará a las **aplicaciones**, en este punto se utilizará todo lo visto anteriormente para responder preguntas como las siguientes: ¿cómo es que funcionan las uniones entre dos semiconductores, o un metal y un semiconductor?, ¿Cómo se puede cambiar la condición de conducción en un semiconductor?, etc. Además se mostrará cómo es que las uniones determinan las características del dispositivo. Extenderemos este conocimiento en la cuarta parte, el capítulo de **Los dispositivos**, para mostrar el comportamiento de diferentes dispositivos electrónicos. Tal vez, puedan parecer temas disímiles. Sin embargo, todos ellos están ligados por un hilo conductor, el conocimiento intrínseco de la materia desde el punto de vista microscópico, el cual explica los comportamientos macroscópicos de los dispositivos.

Otro punto que se desea remarcar y que se debe entender claramente es que, las matemáticas nos dan un lenguaje compacto con el cual se puede, en el campo de la ciencia y la técnica, definir un concepto sin ningún tipo de explicación agregada. Si escribo, d^2f/dr^2 eso es entendido universalmente como una derivada segunda de la función f con respecto a la variable r dos veces, y no se necesita escribir: derivada segunda.... en palabras. Si se escribe,

$$\vec{F} = m\vec{a},$$

El lector entiende, que, además de ser una de las leyes de Newton, se está diciendo, en lenguaje compacto, que la fuerza es igual a la masa por la aceleración, que la masa es un escalar que relaciona dos vectores, fuerza y aceleración, con la misma dirección y que además hay una operación indicada como la multiplicación. Todo eso en una pequeña expresión. Por lo tanto, no se puede y no se debe evitar el uso del lenguaje y símbolos matemáticos introducido hace más de tres siglos y cuya importancia no admite discusión. No obstante ello, en virtud de la facilidad que se pretende que tenga la lectura de este libro, se ha reducido a un mínimo la utilización de dilatadas demostraciones matemáticas, y donde ellas necesariamente se deban desarrollar, se harán de manera sencilla, tal que la deducción pueda ser seguida sin dificultad.

Como se verá, el horizonte es muy amplio y para cubrirlo se necesitan casi tantos textos como los puntos que se han mencionado. Encontrar un único libro que abarque estos temas no es un trabajo sencillo. La realización de un texto así, es la tarea desarrollada, con la finalidad de hacer un aporte en ese sentido a la bibliografía existente. Durante años, muchos estudiantes se acercaban a pedirme mis notas de clase o a decirme que escribiera algo sobre los temas de la clase. Sistemáticamente me negué a ello, pues pensaba que debería existir más de un libro que dejara satisfechas estas necesidades académicas. Con el tiempo me fui convenciendo que redactar este ensayo sería de mucha ayuda para los futuros alumnos.

Abarcar tal cantidad de temas es una tarea ardua, esto quiere decir, que el texto será casi seguramente incompleto. Pero se ha tratado de que las omisiones que se detectarán a través de estas páginas, no tengan incidencia o su incidencia sea mínima sobre los restantes temas.

El esfuerzo para el desarrollo de esta tarea ha sido muy grande, pero he tenido una ventaja muy importante para este trabajo, primero, los años de experiencia al frente de la cátedra, se-

gundo, el tema de investigación sobre el cual trabajé fue la Física del Estado Sólido y por último, las muy fructíferas charlas que he tenido con los docentes de la cátedra y las importantes opiniones que ellos han expresado sobre cada uno de los temas. He tratado de desarrollar cada uno de los puntos tocados en la forma lo más armónica y pedagógica posible. A su vez, he introducido ejemplos que sean lo más representativos posible del tema que se está estudiando, también he balanceado las explicaciones con la formulación matemática y sus deducciones, las cuales se han minimizado al máximo, como ya se mencionó.

Espero que los lectores entiendan esto y puedan comprender las ideas que intento transmitir. Es mi deseo lograr a través de estas páginas la claridad suficiente para que, quien las esté leyendo pueda entender los conceptos que trato de verter y lo entusiasme para seguir el camino de indagar sobre nuevos conocimientos. De no ser así, desearía que me hagan llegar los comentarios para mejorar el texto en posibles futuras ediciones. A todos los docentes de la cátedra, los que están aún y los que han pasado y dejado sus opiniones, muchas gracias, y a los lectores el mayor de mis respetos.

Eitel L.Peltzer y Blancá

PRIMERA PARTE

Herramientas básicas

*El que aprende y aprende, y no practica lo que sabe,
Es como el que ara y ara, y no siembra.*

ARISTOCLES (PLATÓN)
427-347AC

Desde el inicio como estudiantes, escuela primaria, secundaria, universidad, etc. se nos ha mostrado, o al menos intentado, el camino que se debe seguir para tener un mejor y más completo entendimiento de las cosas en general. Esto es, especialmente, muy importante en las ciencias, llamadas “duras”, donde es imprescindible avanzar paso a paso, estudiando los fundamentos para llegar a las aplicaciones. La matemática es un ejemplo claro de este orden que se debe imponer en la adquisición de conocimientos.

Para mostrar este necesario orden, se podría actuar en sentido opuesto, imaginando que esto no es verdad, que se puede avanzar sin conocer los fundamentos de las cosas. Para ver esto, desarrollemos un ejemplo muy sencillo, se pueden estudiar todas las tablas de multiplicación de memoria sin saber el proceso de la multiplicación. Se estudiarán las tablas del 1, 2, 3 etc., desde el 1 al 12 o lo que se desee, pero un día, se va a un negocio y se desea comprar 15,50m de tela a \$17,99 el metro. En forma inmediata se debe realizar la siguiente pregunta ¿Cuánto se debe pagar por esta mercadería?, nunca se podrá contestar porque nunca se ha estudiado la tabla del 15,50, o del 17,99. Pero si se tiene el conocimiento del procedimiento de

la multiplicación, la adición, la sustracción, etc, ya no se tendría ningún problema para contestar esa pregunta o para realizar cualquier otra operación matemática de este tipo. Por lo tanto, se ve que el camino que se debe seguir tiene una única dirección: partir de los fundamentos para luego llegar a las aplicaciones. En términos de la ingeniería, para construir una casa se deben poner los cimientos, luego las columnas y finalmente el techo, en ese orden y no en otro.

En este texto, se desarrollarán los temas en el sentido mencionado. Se pretende brindar en estas páginas los fundamentos, las herramientas básicas para entender los complejos procesos que existen en los dispositivos semiconductores y que dan origen a su comportamiento.

Breve historia

Desde los albores de la humanidad, el hombre descubrió que podía mejorar su calidad de vida o asegurar su subsistencia utilizando los materiales que la naturaleza le brindaba: madera, piedras, huesos, etc. Con el paso del tiempo aprendió a procesar algunos de estos materiales, primero de una manera muy rústica, los cuales fue mejorando a través del tiempo. Así pudo confeccionar diferentes utensilios con metales como: cobre, bronce o hierro. La utilización de estos materiales y las tecnologías para su obtención fue un proceso tan importante en el avance de la humanidad, que la historia impuso el nombre de “Edad de los Cobre, Bronce e Hierro” a diferentes períodos entre los años 7000 y 5000 antes de Cristo. A medida que utilizaba los diferentes elementos, el hombre vio que ellos no le brindaban todas las prestaciones que él requería. Obligado por estas circunstancias se dio cuenta que debía modificarlos, y obtener propiedades diferentes, para mejorar su comportamiento. Mediante la experimentación, muchas veces circunstancial, advirtió que podía cambiar lo que la naturaleza le brindaba, por ejemplo, comprendió que agregando al hierro compuestos orgánicos, ramas, hojas de árboles, podía cambiar su dureza, su temple, su elasticidad, etc. haciéndolos más útiles y acordes a sus necesidades. Todo el conocimiento desarrollado durante cientos de años fue totalmente empírico, basado en la experimentación y la intuición.

Hacia fines del siglo XIX se comienzan a explicar algunas propiedades de los materiales en forma teórica, por medio de las investigaciones de Dalton, Maxwell, Boltzmann, etc. por citar a unos muy pocos hombres de ciencia que hicieron aportes extremadamente significativos. Se llega así a los comienzos del siglo XX, donde se empiezan a cimentar los fundamentos teóricos de las propiedades de los materiales a partir de entender que es el comportamiento microscópico el que brinda sus propiedades macroscópicas.

Algunas propiedades de los materiales que el hombre manipulaba habitualmente eran difíciles de explicar con la sola observación o en forma clásica. Durante el primer cuarto de siglo del siglo XX, aparece una nueva visión de la realidad física, se establecen las bases teóricas con las cuales se pueden empezar a entender algunos de los “misteriosos” comportamientos que mostraban ciertos materiales. La llave para comprender y tener una visión más acabada de todos ellos fue la de poderse introducir en el interior de la materia y ver como el hecho microscópico se proyectaba en lo macroscópico.

Se mejoraron las técnicas experimentales, se propusieron y realizaron nuevos experimentos y se desarrollaron nuevas teorías. Pero fue la teoría de Schrödinger la que brindó la herramienta para que se pudiera estudiar el comportamiento de los componentes de la materia, átomos y electrones, a partir de los cuales se tuvo la capacidad de explicar en forma teórica las diferentes propiedades que se observaban en los experimentos. Hoy se entiende sin lugar a dudas, que los átomos, componentes básicos de la materia, deben ser estudiados en el estado de agregación que se encuentren para comprender las propiedades de los materiales.

Finalmente, hemos llegado al siglo XXI, la actualidad. En estos momentos se desarrolla una nueva revolución de conocimientos y se comienza con un nuevo período, estamos en la llamada Edad de la Información, donde todo se traduce a una simple fórmula, manejar información en forma rápida y eficiente. Pero, obviamente, esto requiere realizar diferentes acciones, en general, se hacen con códigos de procesamiento muy potentes, los cuales, para cumplir la premisa de rapidez de operación, necesitan muy altas velocidades para realizar las mismas. Esta tarea se realiza en grandes o pequeños computadores dependiendo del problema planteado, pero en definitiva todos los procesos terminan en extremadamente pequeños dispositivos electrónicos formados por materiales semiconductores en unión con aislantes y metales, que reciben un impulso eléctrico y lo transforman de acuerdo a sus características.

La exigencia está en hacer todas estas tareas cada vez más rápidamente. A esto se llega con mejores códigos de cómputo y con dispositivos más veloces. ¿Cómo se logra esto? La primera de estas tareas no está dentro del alcance de este libro, pero a la segunda se puede llegar por dos vías, una, mejorando los diseños de los dispositivos, es un problema de la Ingeniería. La segunda, utilizando materiales apropiados para la tarea, este es un problema que debe resolver la ingeniería de materiales, es decir la conjunción de la Física, la Química y la Ingeniería. Con esto se vuelve al principio, el conocimiento micro/nanoscópico de la materia para mejorar esas prestaciones. ¿Quién brinda ese conocimiento?, la respuesta es: una teoría que se desarrolle en lo íntimo de la materia y que dé explicaciones sobre el comportamiento de sus componentes, en nuestro caso, los electrones en diferentes escenarios. ¿Cuál es esa teoría?, la respuesta es: la mecánica cuántica.

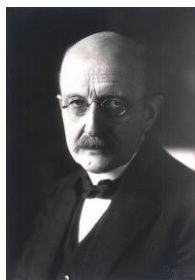
CAPÍTULO I

Introducción a la mecánica cuántica

....3 Les decía: «El sembrador salió a sembrar.
4 Al esparcir las semillas, algunas cayeron al borde del camino y los pájaros las comieron.
5 Otras cayeron en terreno pedregoso, donde no había mucha tierra,
y brotaron en seguida, porque la tierra era poco profunda;
6 pero cuando salió el sol, se quemaron y, por falta de raíz, se secaron.
7 Otras cayeron entre espinas, y estas, al crecer, las ahogaron.
8 Otras cayeron en tierra buena y dieron fruto: unas cien, otras sesenta, otras treinta.
9 ¡El que tenga oídos, que oiga!».
Evangelio según San Mateo, Cap. 13.

En este capítulo se hará una exposición sintética de los acontecimientos que más conmovieron a la física de principios del siglo XX y que abrieron el camino para los posteriores desarrollos que nos han conducido al estado actual de la ciencia y de la técnica. Como he mencionado, no se hará aquí un desarrollo formal introductorio a la Mecánica Cuántica, por lo que la descripción es necesariamente incompleta. Solamente se han desarrollado en forma breve aquellos puntos que servirán para explicar los capítulos posteriores.

I.1. Cuantización de la energía



Max Karl Ernest Ludwig Planck
(23-4-1858, 4-10-1947)
Premio Nobel de Física 1918

En el año 1900, Max Planck estaba trabajando con la radiación del cuerpo negro. Este era un tema de mucho interés en la física de fines del siglo XIX y principios del XX, porque se trataba de la radiación electromagnética que podían emitir los átomos de un material que conforma una cavidad, que simula un cuerpo opaco, que está en equilibrio térmico con el medio ambiente, el cual emite y absorbe iguales cantidades de radiación. Si esto no fuese así, el cuerpo podría estar más frío o caliente que su entorno, lo cual sería contradictorio con la hipótesis de equilibrio térmico. Hasta ese momento, los intentos por explicar el patrón de

radiación de cuerpo negro observado habían sido infructuosos. Pero, en ese año, Planck dio a conocer el hecho de que si hacía hipótesis no ligadas al conocimiento clásico, podía encontrar una función $u(\lambda)$ que se ajustaba a los datos experimentales, ver Figura I-1.

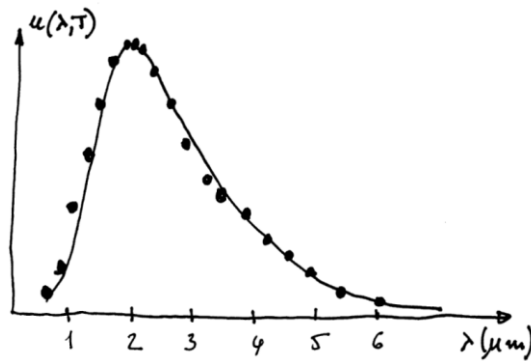


Figura I-1, curva teórica y experimental de la radiación de cuerpo negro.

El conocimiento clásico admitía que la energía era una variable continua, pero para deducir su fórmula empírica, Planck postuló que la energía no era continua sino una variable discreta. Él propuso que las cantidades de intercambio de energía en el proceso de radiación no aumentaban o disminuían en forma continua, sino que estas lo hacían en porciones, en paquetes, a los cuales llamó “cuantos de energía”. Los valores que podían tomar eran: $0, \mathcal{E}, 2 \mathcal{E}, 3 \mathcal{E}, \dots, n \mathcal{E}$, donde n representa un número entero, indicó además que \mathcal{E} la energía, era una cantidad que dependía en forma directamente proporcional con la frecuencia de la radiación, $\mathcal{E} = h\nu$, la dependencia estaba mediada por una constante, por lo que la energía podía escribirse como:

$$E = nh\nu, \text{ con } n=0, 1, 2, \dots$$

La constante que ligaba a la frecuencia con la energía de la radiación se llamó la Constante de Planck, h . Esta describe la energía emitida o absorbida por unidad de tiempo por los átomos de la cavidad.

Finalmente, teniendo en cuenta la naturaleza estadística de la radiación y la ley de distribución de Maxwell-Boltzmann, Planck obtuvo la expresión (ecuación (I-1)) que ajustaba a los puntos experimentales de la Figura I-1, es decir, la densidad de energía de la radiación de cuerpo negro, esta es:

$$u(\nu, T) = \frac{8\pi h\nu^3}{c^3} \frac{1}{e^{h\nu/kT} - 1}, \tag{I-1}$$

donde ν es la frecuencia de onda de la radiación, h la Constante de Planck, c la velocidad de la luz, k la Constante de Boltzmann y T la temperatura de la cavidad. Cuando Planck ajustó esta curva a los datos experimentales obtuvo el valor de la constante h , en la actualidad su valor aceptado es:

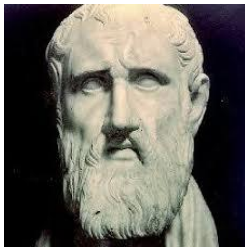
$$h = 6,62606957(29) \times 10^{-34} \text{ Js} = 4,13566733(10) \times 10^{-15} \text{ eVs}$$

Muy frecuentemente, se utiliza la constante de Planck dividida por 2π , se la denomina “ \hbar barra”, o “ \hbar partida”, \hbar , su valor es:

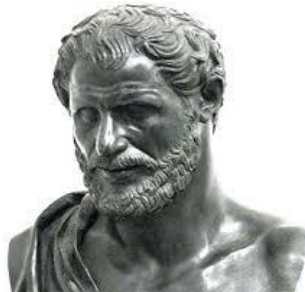
$$\hbar = 1,054571628(53) \times 10^{-34} \text{ Js} = 6,58211899(16) \times 10^{-16} \text{ eVs.}$$

La obtención de la expresión (I-1) por parte de Planck fue por medio de consideraciones estadísticas, él pudo determinar una expresión matemática que se adaptaba a los valores experimentales. Años más tarde, en 1905, Einstein explicó el efecto fotoeléctrico con el modelo de cuantos de luz (fotones), utilizando la teoría de Planck.

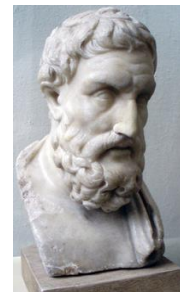
I.2. Los átomos y los modelos



Leucipo de Mileto



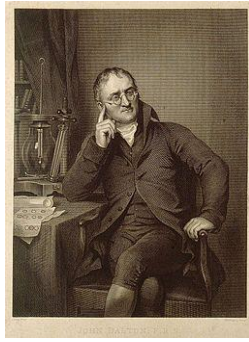
Demócrito de Abdera



Epicuro de Samos

La existencia de elementos muy pequeños constituyendo la materia no es nueva, la historia nos muestra que, unos cinco siglos antes de Cristo, filósofos griegos comprendieron que la materia se encontraba compuesta por pequeños elementos, indivisibles, que llamaron átomos. Estos pensadores fundaron una corriente filosófica llamada “el atomismo”, Leucipo de Mileto (siglo V AC), Demócrito de Abdera (460-370 AC), Epicuro de Samos (341-270 AC), fueron los principales exponentes de esta idea. Demócrito de Abdera dijo: “los átomos son eternos, indivisibles, homogéneos, incompresibles e invisibles”. Desafortunadamente, esta concepción de la materia, fue olvidada por más de veintidós siglos, debido a la aparición de otras corrientes filosóficas. A principios del siglo XIX John Dalton rescata esas ideas y retoma los conceptos del atomismo. En 1808 publica “La Teoría Atómica”, donde en uno de sus postulados dice: “los elementos están constituidos por átomos consistentes en partículas materiales, separadas e indestructibles”

Unos 80 años después de la publicación de Dalton, a fines del siglo XIX, algunos experimentos no podían ser explicados en términos de la teoría clásica, por ejemplo, el realizado por Hertz (Heinrich Hertz, 1857-1894) en el año 1887, donde observó que al iluminar con luz ultravioleta

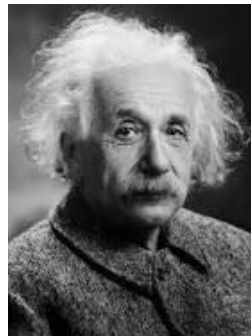


John Dalton
6-9-1766, 27-6-1844



Joseph J. Thomson
18-12-1856, 30-8-1940
Premio Nobel de Física 1906

electrodos de selenio, conectados a alta tensión, se producía un arco de mayor longitud que cuando no se lo iluminaba. La explicación teórica de este fenómeno fue dada más de 20 años después, por Albert Einstein (Premio Nobel de Física, 1921) en 1905 basándose en la teoría propuesta en 1900 por Max Planck sobre los cuantos de energía.



Albert Einstein
14-3-1879, 18-4-1955
Premio Nobel de Física 1921

La física es una ciencia que basa su investigación en el desarrollo de modelos. En el año 1898, J.J. Thomson propuso un modelo de átomo con el cual intentaba explicar los experimentos que se estaban realizando y para los cuales no se encontraba respuesta. El modelo propuesto fue llamado *budín de pasas* (Fig. 1-2), pues establecía que existía una carga positiva distribuida en el espacio en forma esférica y sobre ella se desplazaban los electrones.

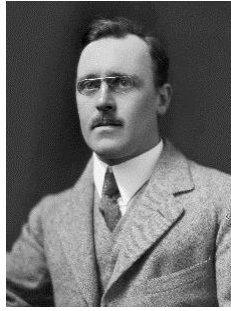
Este modelo de átomo mostró su fracaso con la realización de experimentos de colisiones con *partículas alfa*, sugeridos por E. Rutherford y llevados a cabo por Geiger y Marsden en 1911.

Partiendo del fracaso de los experimentos con el átomo de Thomson, E. Rutherford propuso un nuevo modelo atómico en el cual la carga positiva estaba concentrada en el espacio(núcleo) y la carga negativa se encontraba moviéndose a cierta distancia.

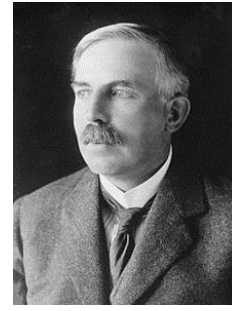
Con él pudo explicar cuáles habían sido las deficiencias del modelo de Thomson y determinar con mucha precisión el tamaño del átomo, 10^{-10}m , como así también la dimensión del diámetro del núcleo atómico, 10^{-14}m .



Hans Wilhelm Geiger
30-9-1882, 24-9-1945



Ernest Madsen
19-2-1889, 15-12-1970



Ernest Rutherford
31-8-1871, 19-10-1937
Premio Nobel de Química 1908

En el año 1913 se propuso un nuevo modelo de átomo el cual corregía las falencias generadas por su antecesor, pero estaba destinado para solucionar los problemas que se presentaban en el átomo de hidrógeno o hidrogenoideos (átomos con un solo electrón en su capa más externa). Permitted predecir todas las series espectrales del hidrógeno. El modelo propuesto por Niels Bohr se basó en trabajos previos de Planck, Einstein, Rutherford, etc.

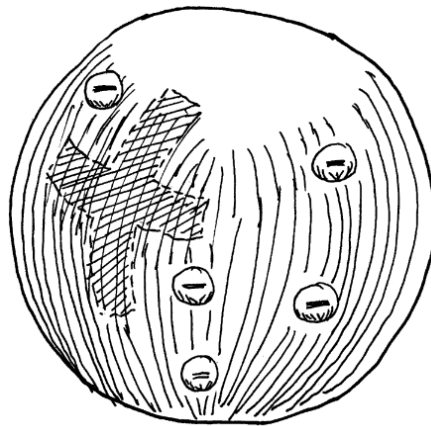


Figura I-2. Modelo atómico de Thomson.

Una de las hipótesis (propuesta en la teoría de Maxwell) que se discutía en aquella época con respecto al nuevo modelo atómico, era que: cuando una partícula cargada (los electrones son partículas con carga) giraban alrededor de un núcleo, estas debían emitir radiación en forma continua, perder energía y caer al núcleo. Clásicamente ese tiempo de caída puede calcularse y además debería poder medirse la radiación emitida.

La hipótesis de Bohr fue que los electrones se movían en orbitas alrededor de un núcleo positivo y que las fuerzas que los ligaban eran electrostáticas. Esto lo llevó a postular, primero, que los electrones se movían alrededor del núcleo en determinadas órbitas estacionarias sin radiar energía. Este fue el primer planteo donde se pone de manifiesto la cuantización del movimiento del electrón.



Niels Henrik David Bohr
7-10-1885, 18-11-1962
Premio Nobel de Física 1922

Años más tarde, de Broglie mostró como debían ser estos movimientos en las orbitas, él hipotetizó el hecho que las partículas podían ser representadas por ondas (como se verá en el próximo subcapítulo). Si se tiene una onda estacionaria, esta no perderá energía, cada onda se continúa con la siguiente sin interferir destructivamente, para ello la longitud de la onda debe ser un número entero de veces la longitud de la órbita del electrón, como se ve en la ecuación (I-2)

$$n\lambda = 2\pi r, \quad n = 1, 2, 3, \dots, \quad (I-2)$$

donde r es la longitud del radio de la órbita del electrón. En otras palabras, una órbita electrónica para que sea estable debe contener un número entero de longitudes de onda de de Broglie, como se ve en la Figura I-3,

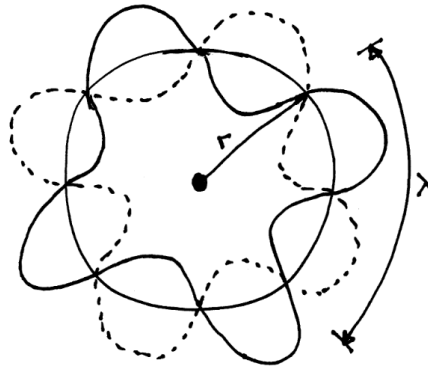


Figura I-3. Órbitas de de Broglie.

El segundo postulado de Bohr fue, que la emisión o absorción de radiación se produciría cuando los electrones se movían de un estado estacionario a otro, de uno de mayor a menor energía o viceversa.

Este postulado indica que el electrón también puede moverse entre estados estacionarios, de una órbita a otra, absorbiendo o emitiendo energía solamente. Este, se puede escribir como:

$$h\nu = E_i - E_f, \quad (I-3)$$

donde h es la Constante de Planck, ν la frecuencia de la radiación, E_i y E_f son la energías totales de las órbitas inicial y final del electrón, respectivamente. El experimento que mostró la

validez del modelo de Bohr fue el realizado por J. Frank y G. Hertz en el año 1914, con el cual se pudo demostrar la cuantización de la energía en los átomos.



James Frank
26-8-1882, 21-5-1964



Gustav Hertz
22-7-1887, 30-10-1975

Premios Nobel de Física 1925

Este modelo atómico fue mejorado posteriormente, se incluyeron órbitas no circulares dando origen a una nueva estructura atómica que se la representó popularmente como un modelo planetario, un núcleo central y los electrones girando a su alrededor como si fueran planetas girando alrededor de una estrella.

Las expresiones mostradas fueron una revolución a principios del siglo XX, pues modificaron el pensamiento clásico. En particular la ecuación (I-3) es muy importante en algunos dispositivos electrónicos, como por ejemplo en celdas solares, LED's o diodos LASER, porque determina la energía o la frecuencia de transiciones entre dos niveles energéticos, por ejemplo niveles de valencia y de conducción.

I.3. Dualidad onda – partícula. Teoría de De Broglie

En 1924 Louis de Broglie sugirió como una especulación, sin pruebas experimentales y/o teóricas, que los cuerpos microscópicos, como por ejemplo, los electrones, podían tener un comportamiento dual, esto es, se podían comportar como partículas o como ondas.

Un par de años antes, Arthur H. Compton había mostrado que una onda podía interactuar con una partícula (electrón) produciéndose la dispersión de la onda incidente bajo un cierto ángulo. La diferencia de la longitud de onda incidente y dispersada se llamó la "longitud de onda de Compton para la partícula dispersora, λ_c (electrón en este caso)". Determinó que la longitud de onda de Compton para un electrón se relaciona con la masa del electrón dispersante por medio de la siguiente expresión:

$$\lambda_c = \frac{h}{m_e c} \quad (I-4)$$



Príncipe LouisVictor Pierre Raymond de Broglie
15 – 8 – 1892, 19 – 3 – 1987
Premio Nobel de Física 1929



Arthur Holly Compton
10-9-1892, 15-3-1962

donde m_e es la masa en reposo del electrón. Planck había encontrado que $E = h\nu$, y Compton que la energía de un fotón es $E = pc$, si despejamos la cantidad de movimiento p , tenemos:

$$p = \frac{E}{c} = \frac{h\nu}{c} = \frac{h}{\lambda}$$

La cantidad de movimiento de un fotón está singularizada por su longitud de onda. De la expresión anterior se puede despejar λ en función de p obteniendo:

$$\lambda = \frac{h}{p} \quad (1-5)$$

Esta última relación se la conoce como la longitud de onda de de Broglie. Con esta, de Broglie, quiere decir que la naturaleza debe ser simétrica y que la expresión (1-4) es completamente general pudiéndose aplicar tanto a partículas materiales como a cuantos de luz (fotones). Si $p = mv$, donde m es la masa relativista y v la velocidad podemos escribir la longitud de onda de de Broglie como:

$$\lambda = \frac{h}{mv} \quad (1-6)$$

Con esta expresión de Broglie establece una relación, ampliamente probada, entre dos aspectos de la física, por un lado, el mundo de características ondulatorias, representado por λ , la longitud de onda, una cantidad que no tiene participación en el mundo material, y por el otro lado, el mundo material, representado por la masa, la cual no tiene contacto con los fenómenos ondulatorios. En esta "simple" fórmula de Broglie une ambos mundos y establece como se vinculan. Esta ecuación nos dice cual es la onda asociada a una partícula en movimiento, y que solamente tiene significación en el mundo microscópico. Se puede ejemplificar esta afirmación con algunos ejemplos.

Ejemplo 1. Imaginemos un electrón ($m=9,1 \times 10^{-31} \text{kg}$) que se mueve a la velocidad de 300 m/s, por lo tanto usando la expresión (I-6) se tiene que $\lambda=2,427 \times 10^{-6} \text{m}$ la cual es una cantidad perfectamente medible.

Ejemplo 2. Ahora, lo que se desea es obtener la longitud de onda asociada a un automóvil que tiene una masa de 100kg y se mueve a 100km/h ($=27,78 \text{m/s}$), reemplazando en la ec. (I-6) se tiene que $\lambda= 2,385 \times 10^{-37} \text{m}$. Esta es una cantidad imposible de medir por cualquier instrumento moderno, no tiene sentido pensar en una longitud de onda de semejante magnitud (Se puede comparar a este valor, el del tamaño nuclear, el cual es del orden de 10^{-14}m).

La idea de estos ejemplos es mostrar que las ondas asociadas a objetos macroscópicos son de una pequeñez tan extrema que no tienen ningún significado físico y que son magnitudes no medibles. Esto no quiere decir que magnitudes muy pequeñas no se puedan medir, por ejemplo la constante de celda de un cristal como el silicio, que conforma una celda cúbica, tiene una longitud de arista de $5,43 \times 10^{-10} \text{m}$. La medida de esta cantidad se hace mediante complejos experimentos con un error extremadamente pequeño.

I.4. Experimento de Davisson – Germer y Thomson

Las hipótesis de de Broglie son difíciles de experimentar en la vida real, en nuestro mundo macroscópico la dualidad entre partícula y onda es imposible de distinguir.



Clinton J. Davisson Lester H. Germer
22-10-1881, 1-2-1958 10-10-1896, 3-10-1971
Premio Nobel de Física 1937



George P. Thomson
3-5-1892, 10-9-1975
Premio Nobel de Física 1937

Pero, ciertas manifestaciones que existen en un campo y no participan del otro, sí son visualizables macroscópicamente. Por ejemplo, la difracción, este un fenómeno estrictamente ondulatorio que no tiene ninguna analogía en el campo de las partículas.

En el año 1927, en Estados Unidos de Norteamérica, Clinton J. Davisson (Premio Nobel de Física 1937) junto con Lester Germer estaban investigando la hipótesis de de Broglie, intentando demostrar que los electrones producían difracción cuando eran dispersados por una red cristalina. En forma independiente George P. Thomson (Premio Nobel de Física 1937) en Inglaterra investigaba sobre el mismo tema.

El esquema del experimento que Davisson y Germer llevaron a cabo se puede ver en la Figura I-4. Ellos bombardeaban una muestra de níquel, con electrones provenientes de un filamento caliente. Tanto la energía de los electrones, el ángulo de incidencia del haz sobre la muestra y la posición del detector eran variables.

La predicción clásica para este tipo de experimento es que la distribución espacial de electrones detectados es uniforme, tal cual era lo que pasaba en el experimento.

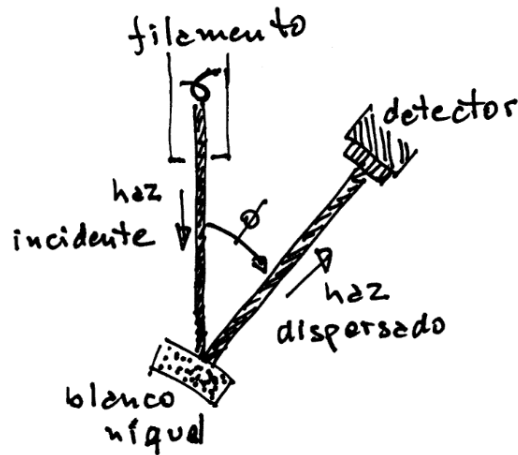


Figura I-4. Diagrama del experimento de Davisson- Germer

Dentro de la campana de vidrio donde se desarrollaba el experimento entró oxígeno y la muestra de níquel se oxidó. El mecanismo para eliminar el óxido es calentar la muestra a muy alta temperatura en una atmosfera inerte (en gas neón) y enfriarla rápidamente. Hicieron esto y la pusieron nuevamente en el detector, los resultados fueron completamente diferentes. Ya no se presentó una variación continua de la intensidad de electrones dispersados, sino que obtuvieron máximos y mínimos de intensidad tal cual los que se presentan en un diagrama de difracción.

La obtención de una figura de difracción hizo que Davisson y Germer se preguntaran el motivo de esta ocurrencia. Si antes de oxidarse no había pasado, ¿porque pasaba después de la oxidación? En realidad la oxidación no tuvo nada que ver con la respuesta posterior, sino que fue el vehículo por medio del cual el cristal se debió calentar y enfriar. Esta acción hizo que en el níquel se produjera una transformación de fase, pasó de una amorfa (distribución aleatoria en el espacio) a una cristalina, los átomos se distribuyeron formando una red regular (distribución periódica en el espacio), por lo que la reflexión de los electrones se realizaba en los planos cristalinos. En muchos libros se pueden encontrar detalles pormenorizados de este experimento y el de Thomson, quien llegó a resultados similares.

Lo interesante de este experimento y los resultados obtenidos por parte de Davisson-Germer y Thomson fue la rotunda confirmación de la validez de la hipótesis de de Broglie, con lo que se mostró claramente que los electrones, las partículas con las cuales se bombardea el blanco de níquel, se comportan como ondas produciendo un diagrama de difracción, que como dijimos es eminentemente ondulatorio. A partir de este resultado, no quedaron dudas que, en una escala microscópica o menor, se puede tratar a las partículas como ondas.

Se puede hacer un rápido cálculo numérico para obtener el valor de la longitud de onda en el experimento de Davisson y Germer. De acuerdo a la expresión de difracción de Bragg vista en óptica, $n\lambda = 2d\sin\theta$, en nuestro caso particular, la distancia entre planos del cristal níquel es $d = 0,91\text{Å}$, el ángulo medido durante el experimento donde se registra un máximo de intensidad fue de $\theta = 65^\circ$. Si suponemos que $n = 1$, se obtiene una longitud de onda, $\lambda = 1,6495\text{Å}$.

Si se utiliza la expresión de de Broglie (Ec. I-4) se debe tener en cuenta que la energía con que se bombardeaba a la muestra de níquel en el experimento fue de 54eV , valor que hace referencia a la energía cinética de los electrones, de esta se puede obtener el valor de la cantidad de movimiento mv . Reemplazando en la ecuación (I-4) se obtiene, $\lambda = 1,6699\text{Å}$. Este valor muestra una gran concordancia con el valor medido, afirmando la validez de la teoría de de Broglie.

I.5. De Broglie y las ondas

El hecho de proponer una dualidad, como la que propuso L. de Broglie, donde la partícula pueda ser representada por una onda, generó una nueva estructura de pensamiento y de razonamiento frente a la nueva realidad confirmada por Davisson y Germer.

Se puede pensar que la partícula con la que se trabajó y teorizó durante más de 20 siglos, es una entidad que tiene a la materia concentrada en un punto del espacio. La onda propuesta por de Broglie, como toda onda, es un fenómeno extendido en el espacio que no posee materia, pero que la representa, es lo que se muestra en la ecuación (I-4). El hecho de representar a una partícula con una onda, abre un interrogante, y es ¿dónde está la partícula?, la respuesta es que, a partir de de Broglie no se puede pensar más desde un punto de vista determinístico, sino desde un punto de vista probabilístico, ya no se puede decir que la partícula está en determinado lugar, sino que, existe una determinada probabilidad de que la partícula esté en ese lugar.

Representemos a la onda de de Broglie por: Ψ , esta puede ser una entidad conformada por números complejos. Como el movimiento del electrón se realiza en un espacio real, la probabilidad de encontrarlo también es real, por lo que, la magnitud que dará la probabilidad de ubicar al electrón por unidad de espacio será Ψ^2 . Esta interpretación fue dada en 1926 por Max Born. Se puede escribir que, la probabilidad de que un electrón se encuentre en un intervalo dx en el espacio es

$$P(x)dx = \Psi^2 dx \quad (I-7)$$

Donde $P(x)$ nos indica la distribución de probabilidades en el intervalo dx .

Debe quedar claro, que el hecho de considerar la teoría de de Broglie abrió otro panorama, el cual, con la caída del pensamiento determinístico, permitió pensar en otras realidades. Por un lado, con el desarrollo probabilístico se abrió el interrogante sobre la localización de la partícula. Por el otro, como hecho natural, se debe pensar que en los experimentos se realizan mediciones en donde intervienen estas partículas/ondas de de Broglie. Si ahora son ondas las partículas, ¿cómo puedo determinar su posición o como puedo medir su velocidad? Como la

onda está ligada a una amplitud, será más probable encontrar a la partícula en aquellas regiones del espacio donde la amplitud de la onda de de Broglie sea más grande. El nuevo escenario propuesto generó marcadas incertidumbres entre los actores de ese momento. En el año 1927 W. Heisenberg planteó el Principio de Incertidumbre, el cual vino a resolver esta cuestión. No desarrollaremos este tema en profundidad, pues supera los alcances de este libro.



Werner Karl Heisenberg
5-12-1901, 1-2-1976
Premio Nobel de Física 1932



Max Born
11-12-1882, 5-1-1970
Premio Nobel de Física 1954

Con este principio se estableció que debido a la naturaleza ondulatoria de las partículas en movimiento, durante el desarrollo de un proceso de medición simultánea de posición y cantidad de movimiento, las indeterminaciones Δx y Δp , de la posición en el espacio y de la cantidad de movimiento, respectivamente, serán en el mejor de los casos mínimos irreductibles, y están relacionados por la expresión

$$\Delta x \Delta p \geq h \quad (1-8)$$

Esta expresión no se plantea como una simple desigualdad debido al hecho que los términos que intervienen, ligados a las determinaciones instrumentales, procesos de medición, o estadísticas siempre incrementarán el valor del producto. Que quiere decir esto, el mejor resultado que se puede tener es que el producto de ambos sea igual a h . Los errores en los procesos de medida de alguna de estas cantidades siempre incrementará el valor del producto $\Delta x \Delta p$ haciendo que la incerteza sea mayor.

El producto de las indeterminaciones significa que no se puede medir con total precisión a ambas magnitudes simultáneamente. Por ejemplo, si tenemos una onda muy compacta, de una longitud muy pequeña, su indeterminación en la posición (Δx) será muy pequeña, por lo que el valor Δp será grande y viceversa. Como se podrá ver esta no es una indeterminación de los aparatos de medida, sino la propia naturaleza es quien la plantea.

Otra forma en que se puede expresar el Principio de Indeterminación, está relacionado con la energía que se emite en un proceso de radiación y la variación en el tiempo del proceso mencionado. La expresión se escribe:

$$\Delta E \Delta t \geq h \quad (1-9)$$

Nuevamente aquí se plantean mínimos irreductibles. Pero a diferencia de la expresión (1-8), existe una excepción en donde Δt , la variación en el tiempo es extremadamente grande, lo que

implica que la variación en la energía ΔE es nula, la excepción está dada en las variaciones de los niveles de energía de los electrones en los átomos, los cuales son inamovibles. Por ejemplo: desde que el universo se generó, y se formaron los primeros átomos, el nivel energético del único electrón del hidrógeno se ubicó en $-13,6 \text{ Ry}$ y permanecerá allí hasta el fin de los tiempos, por lo que al ser esta variación nula, la precisión en la medida será altísima y la variación temporal, infinitamente grande.

Referencias

- Alonso, M., Finn, E. J. *Fundamental University Physics, Vol. III*. 1968 Addison-Wesley series on Physics.
- Beiser, A. *Conceptos de Física Moderna*. 1965 McGraw Hill Book Company.
- Berkeley Physics Course, *Física Cuántica*. Vol. 4. 1972, Ed. Reverte S.A.
- Singh, J. *Modern Physics for Engineers*. 2004 Wiley-VCH Verlag GmbH & Co. KgaA.
- Slater, J. C. *Modern Physics*. 1955. McGraw Hill Book Company.
- Sproull, R. L. *Modern Physics. The quantum physics of atoms, solids and nuclei. 2nd ed.* 1963 J. Wiley & Sons, Inc.
- Tipler, P. A. *Física Moderna*. 2009 Ed. Reverté S.A.
- Tipler, P. A., Mosca, G. *Física para la ciencia y la tecnología. Vol 2C, 5^a ed.* 2010 Ed. Reverté S.A.

CAPÍTULO II

Introducción a la mecánica cuántica.

La ecuación de Schrödinger. Casos particulares

*El tiempo, cuyos dientes roen todas las demás cosas,
es impotente contra la verdad.*

T.E. HUXLEY. Británico, Biólogo,
4-5-1825:29-6-1895

Defensor de la teoría de la evolución.

La electrónica es un espejo donde se reflejan en forma de dispositivos, lo mejores y más sorprendentes desarrollos alcanzados por la física cuántica. No habría dispositivos electrónicos si no hubiera habido una física que explicará cómo se comportan los electrones en el mundo microscópico y como se puede utilizar ese comportamiento para generar diferentes respuestas.

En este capítulo, se introducirá al lector, no en una deducción rigurosa de la ecuación de Schrödinger sino en una idea intuitiva de su obtención. Se avanzará sobre el estudio de aplicaciones de esta ecuación a casos sencillos que están relacionados con temas de la electrónica. El desarrollo de nuevas tecnologías para la construcción de nuevos materiales, especialmente semiconductores, ha mostrado la importancia de estos ejemplos, por medio de los cuales se han podido explicar muchas de las propiedades físicas de los nuevos dispositivos.

II.1. Ecuación de Schrödinger para electrones



Erwin Rudolf Josef Alexander Schrödinger
12-8-1887, 4-1-1961
Premio Nobel de Física 1933

En función de la reciente teoría de de Broglie, la cual asociaba ondas a corpúsculos materiales, era necesario saber cómo se comportarían esas ondas bajo la acción de fuerzas externas. Similarmente a las ecuaciones de Newton, que describen el movimiento de los cuerpos, se necesitaba una ecuación que estableciera como sería el movimiento de estas ondas de de Broglie. En el año 1926, E. Schrödinger postuló una expresión que describe este movimiento. La ecuación de ondas de Schrödinger permite describir la distribución de probabilidad de la partícula en el espacio y en el tiempo. La expresión encontrada en una dimensión espacial y el tiempo, es:

$$-\frac{\hbar^2}{2m} \frac{\partial^2 \Psi(\mathbf{x}, t)}{\partial \mathbf{x}^2} + V(\mathbf{x})\Psi(\mathbf{x}, t) = i\hbar \frac{\partial \Psi(\mathbf{x}, t)}{\partial t} \quad (\text{II-1})$$

Esta es una ecuación de energías, donde en el miembro de la izquierda tenemos el primer término que corresponde a la energía cinética, y el segundo a la energía potencial. En el segundo miembro está expresada la energía total. Aquí se describe el movimiento de una partícula de masa m , en el campo de fuerzas generado por la función potencial $V(\mathbf{x})$. Esta expresión está desarrollada en función de operadores matemáticos que describen “matemáticamente” a los observables físicos, cantidades que se pueden medir en un experimento. Por ejemplo, la cantidad de movimiento en la dirección x , se escribe como: $p_x = -i\hbar(\partial/\partial x)$, mientras que la energía toma la forma: $E = i\hbar(\partial/\partial t)$.

En muchos casos de la física en donde el problema se puede plantear solamente en términos de la posición y no del tiempo, la expresión (II-1) puede simplificarse estableciendo que la función de onda $\Psi(\mathbf{x}, t)$, puede ser separada en una parte espacial y en una parte temporal, esta última se puede expresar como un factor exponencial y la solución se puede escribir como:

$$\Psi(\mathbf{x}, t) = \psi(\mathbf{x})e^{-i\omega t} \quad (\text{II-2})$$

Donde ω es la frecuencia de la onda. Si reemplazamos la ecuación (II-2) en la (II-1) se obtiene la ecuación de Schrödinger independiente del tiempo.

$$-\frac{\hbar^2}{2m} \frac{d^2\psi(x)}{dx^2} + V(x)\psi(x) = E\psi(x) \quad (II-3)$$

Debemos tener en cuenta en (II-3) que $E = \hbar\omega$. Esta es una ecuación diferencial de segundo grado en una variable. La función de onda $\psi(x)$ debe satisfacer a la ecuación (II-3) y debe cumplir otras condiciones: $\psi(x)$ debe ser continua, como la ecuación contiene una derivada segunda, la primer derivada de la función, $\psi'(x)$, también debe ser continua. La última condición sobre la función de onda es que esta no puede ser constante en todo el espacio, no se puede tener la misma probabilidad en cualquier lugar del espacio, por lo que $\psi(x) \rightarrow 0$ cuando la variable x toma valores muy grandes ($x \rightarrow \pm\infty$), esta condición es necesaria para que $\psi(x)$ pueda normalizarse. Para aclarar esto, podemos imaginar que tenemos una partícula en un recinto extremadamente grande, en algún lugar la partícula está, por lo tanto si integramos la densidad de probabilidad de encontrar a la partícula, el resultado debe ser 1, porque la partícula está dentro del recinto, matemáticamente:

$$\int_{-\infty}^{+\infty} \psi^*(x)\psi(x)dx = 1 \quad (II-4)$$

Como $\psi(x)$ puede ser una función compleja, y la densidad de probabilidad es una función real, tomamos el módulo de esta, es decir, el complejo conjugado $\psi^*(x)$ por la función $\psi(x)$.

Por último, para dar una descripción un poco más detallada de la ecuación de Schrödinger y sin ser demasiado específicos, se puede decir que la ecuación (II-3) es una ecuación de autofunciones y autovalores. Las autofunciones son las soluciones $\psi(x)$ de la ecuación y los autovalores son los valores de E . Por supuesto que no cualquier función $\psi(x)$ se corresponde con un valor de E y viceversa, para satisfacer la ecuación de Schrödinger, los que si se corresponden son las soluciones de la ecuación. La función $\psi(x)$ que se denomina autofunción describe un estado del sistema, por lo que también se la puede llamar autoestado del sistema. Por lo general, las autofunciones no son soluciones simples, sino que se forman como una combinación de otras funciones. Estas se llaman funciones “base”, que deben cumplir algunas condiciones. Deben ser un sistema completo, es decir, cualquier estado del sistema debería poder ser descrito por estas funciones, y deberían ser independientes entre si. Esta propiedad se llama “ortogonalidad” y significa que ningún elemento del conjunto de funciones base puede ser expresado en términos de otros elementos de ese conjunto. El alfabeto español es un buen ejemplo de lo que se ha dicho, consta de 27 letras con las cuales se pueden formar todas las palabras que están en el Diccionario de la Real Academia, por lo tanto el conjunto es completo, si nosotros nos refiriéramos a un conjunto de 10 letras (de la a a la j) habría una enorme cantidad de palabras que no podríamos formar, el conjunto no sería completo. Por otro lado, los términos son independientes (ortogonales) pues, por ejemplo, la letra “c” u otra cualquiera, no puede ser escrita en términos de combinaciones de otras letras del conjunto. Si la “ch” o la “ll” o la “rr”, que son composiciones de dos letras, llamadas dígrafos, tienen sus propias

pronunciaciones, si estuvieran dentro del alfabeto ya no se cumplirían las condiciones que hemos mencionado, debido a que son combinaciones de otras letras del alfabeto.

II.2. Aplicaciones de la Ecuación de Schrödinger

En forma rápida se ha mostrado la ecuación de Schrödinger. Cualquier estudio formal sobre ella debería hacerse en espacios mucho más dilatados que el presente, concisamente se ha introducido una de las herramientas que ha permitido ahondar en el conocimiento de la materia, y también ha permitido abrir un sin número de puertas en el conocimiento, cada una de las cuales lleva a campos diferentes de la Física. La comprensión del comportamiento de los dispositivos electrónicos ha sido una de estas puertas.

El comportamiento microscópico de los dispositivos electrónicos dista mucho de poder ser explicado clásicamente, es necesario recurrir a la mecánica cuántica para entender su funcionamiento. La ecuación (II-3) si bien no será la única, si quizás es la más importante para entender el movimiento de los electrones dentro de los materiales que conforman a un dispositivo. En esa ecuación aparece el término de energía potencial $V(x)$ mediante el cual se representa, de alguna manera, el “medio ambiente” donde se mueven los electrones, es el que aporta la condición física a la ecuación. Este, será variable en el espacio, describiendo los diferentes lugares donde los electrones deberán desplazarse. A pesar de las discontinuidades que presente el potencial la función $\psi(x)$ debe ser continua, y también, como mencionamos antes, $\psi'(x)$ también debe ser continua. La continuidad de la función $\psi(x)$ está ligada al hecho de que siendo el cuadrado de esta función la probabilidad de encontrar a la partícula, sería ilógico pensar que la probabilidad pueda ser discontinua entre un punto y el inmediato contiguo. Por lo tanto, lo que decimos es que si en nuestro espacio (unidimensional) en un punto x_a se plantea una discontinuidad del potencial, tenemos que:

$$\psi(x_{a^-}) = \psi(x_{a^+}) \text{ y } \psi'(x_{a^-}) = \psi'(x_{a^+}) \tag{II-5}$$

Aquí, x_{a^-} y x_{a^+} son puntos muy próximos a x_a por izquierda y derecha, respectivamente

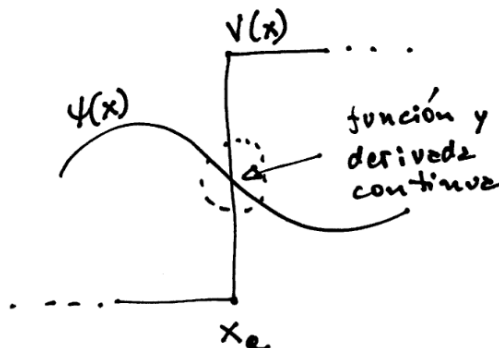


Figura II-1. Continuidad de la función de onda y su derivada.

Se desarrollarán ahora un conjunto de ejemplos que tienen dos finalidades, una introducir al estudiante en la resolución de la ecuación de Schrödinger (ecuación diferencial de 2do orden) con condiciones de contorno que responden a condiciones físicas, y

segundo, comenzar a mostrar algunos fenómenos físicos que influyen muy fuertemente en el comportamiento de los electrones, gracias a los cuales se obtienen dispositivos electrónicos que brindan diferentes respuestas.

II.2.1. Pozo infinito

El pozo infinito tiene la característica de poseer soluciones matemáticas sencillas y las soluciones que se plantean se asemejan a la de una cuerda vibrante sujeta por sus extremos.

Se puede imaginar que existen fuerzas en un lugar del espacio, tan intensas que no permiten que las partículas que hayan caído bajo su influencia puedan librarse de su acción. Como en la física se deben establecer modelos para representar lo que ocurre en la naturaleza, en este caso el modelo que se puede pensar es el de un pozo, para dar idea de que cuando las partículas caen en el no pueden salir. Es decir, se le pueden asignar paredes absolutamente impenetrables. Esto se lo puede modelizar mediante un potencial extremadamente profundo (infinito), y de ancho finito, es decir $V(x) \rightarrow \infty$ entre $x = 0$ y $x = a$, mientras que fuera de estos límites $V(x) \rightarrow 0$ como se ve en la Figura II-2. La ecuación (II-3) se puede escribir de la siguiente manera:

$$-\frac{\hbar^2}{2m} \frac{d^2\psi(x)}{dx^2} = [E - V(x)]\psi(x) \tag{II-6}$$

Como en este modelo las paredes son absolutamente impenetrables, la función de onda $\psi(x)$ debe anularse en los bordes del pozo, $\psi(0) = 0$ y $\psi(a) = 0$. Dentro de la caja se propone una solución general de la ecuación (II-6) de la forma,

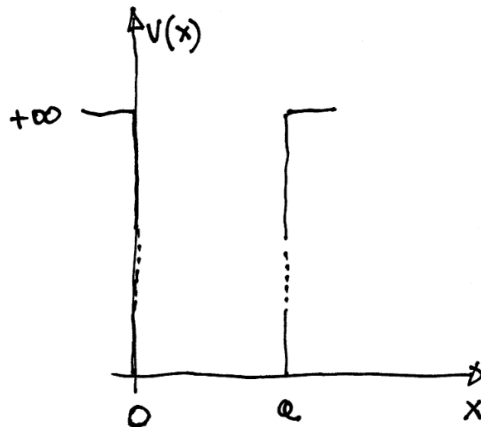


Figura II-2. Pozo infinito

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

$$\psi(x) = A \sin(kx) + B \cos(kx) \tag{II-7}$$

$$\text{con } k = \sqrt{2mE}/\hbar \text{ para ambas.} \quad (\text{II-8})$$

Por simplicidad se utilizará la (II-7), pero se debe aclarar que con ambas propuestas se llega a la misma solución. Debido a la impenetrabilidad de la pared se tiene que $\psi(0) = 0$, en este límite ($x = 0$), las funciones circulares que participan en la solución tiene una de ellas valor nulo, el $\text{sen}(0) = 0$, y la otra, valor unitario, el $\text{cos}(0) = 1$. Esto lleva a determinar que debe ser nula la constante B , por lo tanto la solución será de la forma:

$$\psi(x) = A \text{sen}(kx)$$

Por otro lado, hemos dicho que $\psi(a) = 0$, es decir,

$$\psi(a) = A \text{sen}(ka) = 0 \Rightarrow ka = n\pi \quad (\text{II-9})$$

La constante A no puede anularse, pues sino la solución $\psi(x)$ sería idénticamente nula. Si despejamos k de la ec. (II-9) y lo reemplazamos en (II-8) y de allí despejamos E , tenemos:

$$E = \frac{n^2 \pi^2 \hbar^2}{2ma^2} \quad (\text{II-10})$$

Se puede apreciar que la energía depende de constantes y del número n , el cual es un número entero, que toma los valores 1, 2, 3, 4, ..., el valor 0 está excluido debido a que esto implicaría una solución (función de onda) nula. A esta energía, dependiente de n , la llamaremos E_n . Se puede ver en la expresión (II-10) que si se toma el valor $n = 1$, se tiene:

$$E_1 = \frac{\pi^2 \hbar^2}{2ma^2} \quad (\text{II-11})$$

Por lo que se puede expresar las energías E_n como múltiplos de E_1 , esto es: para $n = 1$, es E_1 , para $n = 2$, $E_2 = 4E_1$, para $n = 3$, $E_3 = 9E_1$, etc. estas energías definen cada uno de los estados que podemos tener dentro de un pozo de estas características, los electrones que se encuentren en estos niveles no podrán salir del pozo y a esos estados se los llama "estados ligados".

Por lo tanto, la solución será ahora:

$$\psi(x) = A \text{sen}\left[\left(\frac{n\pi}{a}\right)x\right] \quad (\text{II-12})$$

Se debe determinar la constante A . Como la partícula está dentro de la caja confinada entre los límites 0 y a , la probabilidad de encontrarla allí va ser igual a 1 (condición de normalización), es decir:

$$\int_{-\infty}^{\infty} \psi^*(x)\psi(x)dx = \int_{-\infty}^{\infty} \psi^2(x)dx = \int_0^a \psi^2(x)dx = 1$$

Desarrollando,

$$\int_0^a \psi^2(x) dx = \int_0^a A^2 \text{sen}^2\left(\frac{n\pi}{a}x\right) dx = A^2 \int_0^a \text{sen}^2\left(\frac{n\pi}{a}x\right) dx = 1$$

La última integral, es una integral tabulada cuyo valor es $a/2$ por lo que tenemos que

$$A^2 \frac{a}{2} = 1 \Rightarrow A = \sqrt{\frac{2}{a}} \quad \therefore \psi(x) = \sqrt{\frac{2}{a}} \text{sen}\left(\frac{n\pi}{a}x\right) \quad (II-13)$$

Como $\psi(x)$ es una función que ha quedado dependiendo de la variable x y del valor n , la llamaremos $\psi_n(x)$. Toda esta descripción puede verse en la Figura II-3, donde se ha indicado el nivel n y su energía E_n , también la función de onda para el nivel correspondiente $\psi_n(x)$, y la densidad de probabilidad (línea de rayas) de encontrar a la partícula en el nivel, $[\psi_n(x)]^2$.

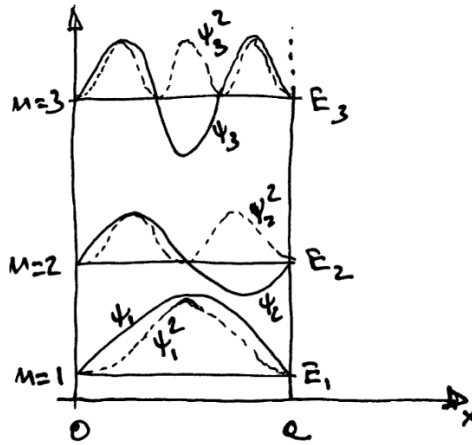


Figura II-3. Niveles de energía, energías, funciones de onda y densidad de probabilidad para una partícula en una caja con potencial infinito.

¿es importante el pozo infinito en una descripción de dispositivos electrónicos, si solamente parece ser una idea con fines académicos?. Si se piensa en una estructura tridimensional formada por dos materiales diferentes, A y B, superpuestos en capas, ABA, como se representa esquemáticamente en la Figura II-4, allí se ve que la estructura energética de los materiales permite que se forme entre ellos un pozo de potencial, cuyas paredes (material A) no pueden ser penetradas por las ondas dentro del material B. Cada uno de los niveles permitidos dentro del pozo, cuando está ocupado por electrones, puede conducir, los cuales se podrían mover solamente en la dirección y . En la actualidad este tipo de estructuras se forman, por ejemplo, con compuestos GaAsAl (material A) y GaAs (material B).

II.2.2. Pozo finito

Se puede imaginar que ahora se está frente a un caso totalmente diferente, un lugar donde las paredes que contiene a las ondas no son absolutamente impenetrables, sino que pueden ser penetradas, como esquemáticamente se muestran en la Figura II-5.

La altura del pozo es V_0 y su ancho es a . El hecho de que las ondas, con una energía aún menor que V_0 tengan la probabilidad de penetrar las paredes es un fenómeno que, exclusivamente, se produce en la mecánica cuántica, no existe en la mecánica clásica un hecho

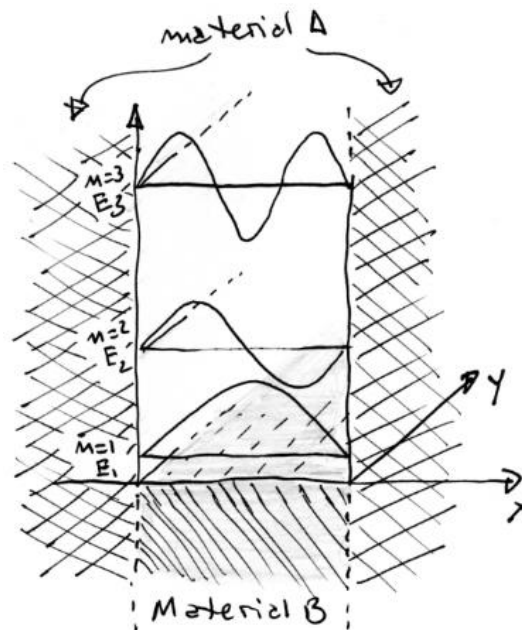


Figura II-4. Pozo con materiales superpuestos.

similar. Clásicamente se podría pensar en una bola de billar que es arrojada contra una pared. Si la bola no tiene la suficiente energía para atravesar la pared, seguramente se producirá un choque elástico, la bola interactuará con la pared y rebotará, pero nunca va a tener la posibilidad de atravesarla, cosa que si ocurre en la mecánica cuántica. Este es uno de los fundamentos que esgrime la física cuántica para diferenciarse de la clásica.

Se debe pensar ahora en resolver la ecuación (II-6), para obtener las funciones de onda y niveles de energía dentro del pozo. Se tiene que hacer una aclaración, un electrón que posea una energía $E > V_0$ será lo que se conoce como electrón libre, ya que su energía le permite alcanzar estados de energía superiores a la altura del pozo, estos podrán moverse en un continuo de energías, mientras que si la energía de los electrones es $E < V_0$ estos serán electrones ligados. Son estos últimos los que, en este momento, nos interesan en particular.

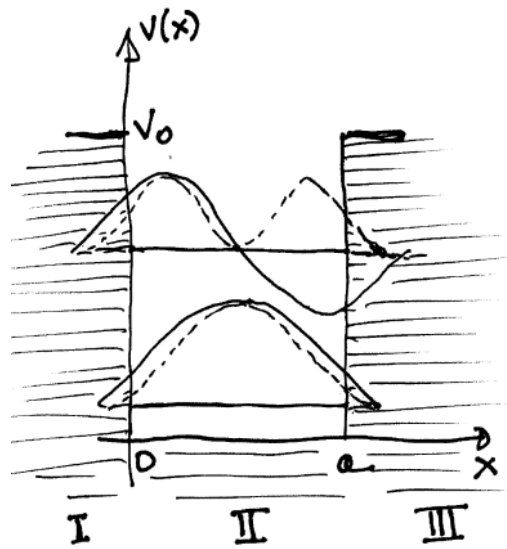


Figura II-5. Pozo finito.

Los límites del pozo están entre 0 y a , donde el potencial vale 0 , fuera de este lugar el potencial vale V_0 , tal como se muestra en la Figura II-5. Matemáticamente:

$$V(x) = 0, \quad 0 < x < a \\ = V_0, \quad x < 0 \text{ y } x > a,$$

como $V(x) = 0$ entre 0 y a (Región II), la expresión (II-6) queda:

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

A esta ecuación la reescribiremos, en forma más compacta, de la siguiente manera

$$\frac{d^2\psi(x)}{dx^2} = -\frac{2m}{\hbar^2} E\psi(x) \rightarrow \psi''(x) = -k^2\psi(x) \quad (\text{II-14})$$

donde: $k^2 = \frac{2mE}{\hbar^2}$ (II-15)

Las soluciones que se propondrán en este caso son:

$$\psi_{II}(x) = A \sin(kx) \quad \text{o} \\ = B \cos(kx) \quad (\text{II-16})$$

La diferencia que se establece con el pozo infinito, donde las funciones de onda se anulaban estrictamente sobre las paredes, es que aquí las funciones no se anulan, por lo tanto

$$\psi(x) \neq 0 \text{ en } x = 0 \text{ y en } x = a.$$

Por otro lado, para $x < 0$ y $x > a$, $V(x) \neq 0$, por lo que la ecuación de Schrödinger quedará:

$$\psi''(x) = \frac{2m}{\hbar^2} (V_0 - E)\psi(x) = \beta^2\psi(x) \quad (\text{II-17})$$

donde

$$\beta^2 = \frac{2m}{\hbar^2} (V_0 - E) \quad (\text{II-18})$$

la soluciones deberán decrecer a medida que nos alejamos del pozo, en este caso, la solución será:

$$\psi_{I,III}(x) = Ce^{-\beta x} + De^{\beta x} \quad (\text{II-19})$$

La solución mostrada en la ecuación (II-19) es válida en las regiones **I** y **III**, pero debemos examinarla cuidadosamente, en la región **I**, $x < 0$, el exponente del exponencial del primer miembro de la derecha será positivo ($-\beta x > 0$) y cuando $x \rightarrow \infty$, este término divergirá, por lo que se debe tomar $C = 0$. La solución en la Región **I**, será:

$$\psi_I(x) = De^{\beta x} \quad (\text{II-20})$$

Similarmente, en la región **III**, cuando x toma valores muy grandes, el exponencial del 2do término de la derecha divergirá, la solución será:

$$\psi_{III}(x) = Ce^{-\beta x}$$

En $x = 0$, el borde entre la Región **I** y **II** se deben cumplir las condiciones de continuidad de la función y su derivada, esto es:

$$\psi_I = \psi_{II} \rightarrow D = 0 \quad (\text{II-21})$$

$$\psi'_I = \psi'_{II} \rightarrow \beta D = kA \quad (\text{II-22})$$

En $x = a$, el límite entre las regiones **II** y **III** se deben cumplir las mismas condiciones,

$$\psi_{II} = \psi_{III} \rightarrow A \sin(ka) = Ce^{-\beta a} \quad (\text{II-23})$$

$$\psi'_{II} = \psi'_{III} \rightarrow kA \cos(ka) = -\beta Ce^{-\beta a} \quad (\text{II-24})$$

si, hacemos el cociente de las ecuaciones (II-24)/(II-23) tenemos:

$$k \cot g(ka) = -\beta \quad (\text{II-25})$$

Para simplificar la obtención de la solución, solamente se ha utilizado en la Región **II**, una parte de la función de onda, la solución completa será la suma de ambas soluciones parciales que se han propuesto en las expresiones (II-16). La ecuación (II-25), es una ecuación trascendente, la incógnita, el valor de E que aparece dentro de la expresiones de k y β , no

puede ser despejado mediante operaciones algebraicas, por lo que debe ser tratada numérica o gráficamente. Para ello, primero “se adimensionalizará” la ecuación, multiplicando ambos miembros por a ,

$$ka \cot g(ka) = -\beta a, \quad (II-26)$$

Se puede hacer la siguiente sustitución, $ka = \xi$ y $\beta a = \eta$, con lo cual (II-26) queda:

$$\xi \cot g(\xi) = -\eta \quad (II-27)$$

Por otro lado, se puede establecer una conexión entre ξ y η mediante la relación:

$$\xi^2 + \eta^2 = \frac{2m}{\hbar^2} E a^2 + \frac{2m}{\hbar^2} (V_0 - E) a^2 = \frac{2m}{\hbar^2} V_0 a^2 \quad (II-28)$$

Esta última expresión define una circunferencia, cuyo radio $r^2 = \left(\frac{2m}{\hbar^2}\right) V_0 a^2$. Para cada valor específico de V_0 , el valor de η (ec. II-27) definirá un punto de la circunferencia con radio r . Como ξ y η están definidas positivas es suficiente considerar sus valores en el primer cuadrante.

El tratamiento numérico se debe hacer mediante algún método recursivo utilizando computadora, dando valores a E , siguiendo el proceso hasta que la igualdad (II-27) se satisfaga y se cumpla la condición (II-28). Gráficamente se puede solucionar determinando los puntos de intersección entre la circunferencia (ecuación (II-28)) y la curva de la ecuación (II-27), como se muestra en la Figura II-6. Se ve en esa figura que el eje ξ es cortado por $-\xi \cot g \xi$ en $\pi/2$ y en $3\pi/2$, por lo que valores del radio de la circunferencia menores de $\sqrt{\pi/2}$ no generarán ningún nivel de energía. A medida que el radio r crezca, esto significa: pozos más profundos, más anchos o ambos, empezarán a aparecer otras curvas, seguramente con un número mayor de intersecciones y por ende los valores de los niveles de energía dentro del pozo (soluciones de la ecuación). Por ejemplo, si se tuviera una circunferencia (ec. (II-28)) cuyo radio r fuese un poco mayor de 2, se tendría un pozo con un único nivel de energía (punto de intersección **a** en la Fig. II-6), por otro lado si el valor de r fuese superior a 5, tendríamos dos niveles de energía dentro del pozo (**b** y **c**).

Ejercicio sugerido: encontrar los niveles de energía de un pozo de 3eV de profundidad y 3Å (1Å = 10⁻¹⁰m).

Rta: 1er.nivel: ~ 0.6 eV, 2do. Nivel: ~ 2.01eV

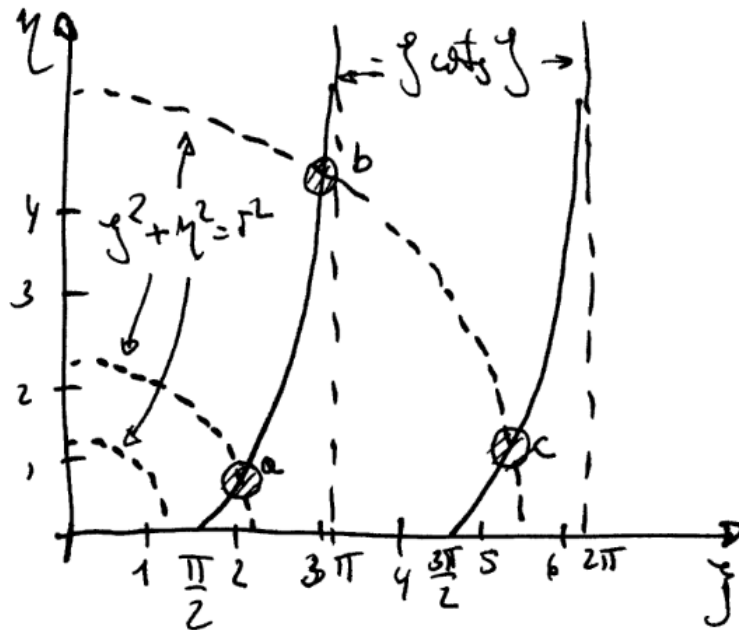


Figura II-6. Gráfico de las soluciones del pozo finito.

Se debe tener en cuenta que para este desarrollo solamente se ha utilizado una parte de la solución, $A \sin(kx)$, para tener un panorama completo de niveles de energía dentro del pozo, también se debe utilizar la otra parte $B \cos(kx)$. Siguiendo un procedimiento completamente análogo al desarrollado, se puede obtener la descripción completa del problema.

Se necesitan encontrar ahora las funciones de onda de cada uno de los estados ligados. De acuerdo a las definiciones que se han utilizado, las funciones de onda tienen comportamiento sinusoidal dentro del pozo y exponencial decreciente fuera de él. La forma de las funciones para dos estados ligados es la mostrada en la Figura II-5. Se deben determinar las constantes A y B , y el problema estará resuelto. Esto se hace mediante la normalización de la función de onda. El exponente del exponencial, β , ec. (II-19), se conoce, cuando se fija V_0 y se determina E .

En este punto, surge la siguiente pregunta, ¿Cuánto penetra la función de onda fuera del pozo? Para la contestación se debe rever la ecuación (II-18), allí lo que se plantea es la diferencia entre la energía potencial y la energía total de la partícula, es decir, su energía cinética. Se ha dicho que a diferencia de la física clásica, uno de los postulados de la física cuántica es que existe la probabilidad de la partícula penetrar la materia, que salga del pozo. En esas regiones $V_0 > E$, lo que estaría indicando una energía cinética negativa. Algo así, no tiene ningún sentido físico. Ninguna de las dos teorías, la clásica o la cuántica, predicen una energía cinética negativa. Es aquí donde interviene el principio de incertidumbre. La función de onda fuera del pozo, para $x > a$, por ejemplo, decae como:

$$\psi(x) \approx e^{-\beta x} \quad \rightarrow \quad \psi^2(x) \approx e^{-2\beta x} \tag{II-29}$$

y la densidad de probabilidad como su cuadrado. Si se supone que nuestra incertidumbre es del orden de $\Delta x \approx \beta^{-1}$, se deduce que ψ^2 es muy pequeño a esa distancia, por lo que se puede

considerar a $\psi(x)$ como despreciable cuando $x > a + \beta^{-1}$. Esto quiere decir que se podrá encontrar la partícula en el rango de la distancia β^{-1} . La expresión de Heisenberg sobre el principio de incertidumbre, relaciona la variación de la cantidad de movimiento y de la posición con \hbar , podemos escribir:

$$\Delta p \Delta x \approx \hbar \quad \rightarrow \quad \Delta p \beta^{-1} \approx \hbar \quad \Rightarrow \quad \Delta p \approx \hbar \beta \quad (\text{II-30})$$

Partiendo de esta expresión, se puede escribir a la energía cinética como:

$$E_c = \frac{(\Delta p)^2}{2m} = \frac{\hbar^2 \beta^2}{2m} = V_0 - E \quad (\text{II-31})$$

En la ecuación (II-31) se ha utilizado la expresión (II-18). Se observa aquí que esta es la energía cinética mínima necesaria para evitar que tome valores negativos. Por lo tanto, el principio de incertidumbre nos dice que en el rango $\Delta x \approx \beta^{-1}$ la energía cinética será positiva y no estaremos violando ningún principio de la física.

La explicación desarrollada sobre cuanto puede penetrar la función de onda en la materia es especialmente importante en el fenómeno túnel que se verá más adelante.

En estos dos puntos previos se ha hablado de cuando las partículas “caen” a un pozo de potencial. Más apropiadamente se debería decir, cuando una partícula queda atrapada en un estado ligado de un potencial atractivo. Eso es lo que se ha descrito brevemente.

Ejercicio sugerido: para el pozo del ejemplo anterior, encontrar la longitud de penetración.

II.2.3. Barrera de potencial. Efecto Túnel

Se brindará en este punto una introducción a uno de los fenómenos de la mecánica cuántica más interesante y con inmensas aplicaciones en el mundo de los dispositivos electrónicos, el efecto túnel. Este tampoco tiene correlato en la Física Clásica. El atravesar una barrera de un potencial repulsivo de ancho finito, solamente puede ser realizado mediante efecto túnel. Como en los casos anteriores la explicación se desarrollará en un espacio unidimensional, en donde se describirá sin pérdida de generalidad el fenómeno en cuestión.

El modelo a utilizar es una barrera rectangular de alto V_0 y ancho a , mostrada en la Figura II-7. Se va a suponer que nuestra barrera está sola en el universo, y las ondas que interactúen con esta, no podrán interactuar con ninguna otra barrera.

La partícula se mueve con una energía $E < V_0$. En la Figura II-7 se indica cada una de las regiones donde se mueven las ondas, como así también, el flujo de densidad de partículas incidentes, reflejadas y transmitidas. Se plantea la solución de la ecuación de Schrödinger para el caso de la barrera de potencial. En la Figura II-7, se ha elegido que las ondas provengan de la izquierda, y estas se moverán en una zona donde $V = 0$, en las regiones I y III, lugares donde se propondrán soluciones oscilatorias para las ondas incidentes y las reflejadas.

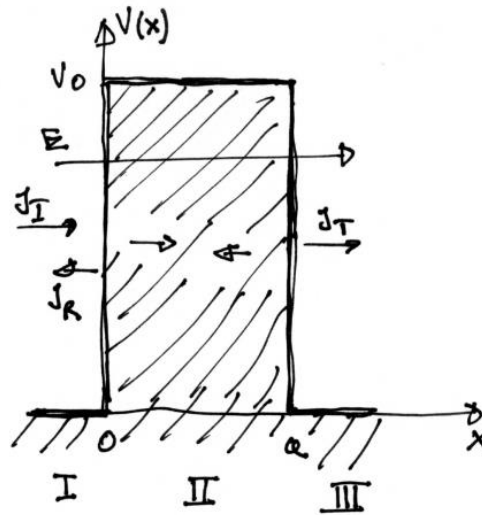


Figura II-7. Barrera de potencial.

Diferente es el caso en la región II, $V(x) = V_0$, aquí el medio es dispersivo, por lo que las ondas tenderán a decaer y se pueden representar por un exponencial decreciente. Como se ha supuesto que no existe otra barrera para que pueda reflejar alguna onda, lo único que se tendrá será una onda transmitida en la región III. Las ecuaciones que a resolver son:

$$-\frac{\hbar^2}{2m} \frac{d^2\psi}{dx^2} = E\psi \quad \text{Región I y III} \quad (II-32)$$

$$-\frac{\hbar^2}{2m} \frac{d^2\psi}{dx^2} + V_0\psi = E\psi \quad \text{Región II} \quad (II-33)$$

Las soluciones serán:

$$\psi_I(x) = Ae^{i\alpha x} + Be^{-i\alpha x} \quad \rightarrow \quad x < 0 \quad (II-34)$$

$$\psi_{II}(x) = Ce^{\beta x} + De^{-\beta x} \quad \rightarrow \quad 0 \leq x \leq a \quad (II-35)$$

$$\psi_{III}(x) = Fe^{i\alpha x} \quad \rightarrow \quad x > a \quad (II-36)$$

en la ecuación (II-34) falta el término reflejado, por lo que se ha mencionado anteriormente. Si la energía de la partícula fuese mayor que V_0 las soluciones en el intervalo desde $x = 0$, hasta $x = a$, también serían funciones ondulatorias. En las ecuaciones (II-34) – (II-36):

$$\alpha = \sqrt{\frac{2m}{\hbar^2} E} \quad y \quad \beta = \sqrt{\frac{2m}{\hbar^2} (V_0 - E)} \quad (II-37)$$

Estas soluciones presentan cinco coeficientes ($A \rightarrow F$) y cuatro ecuaciones, que son las obtenidas de aplicar las condiciones de continuidad para la función y su derivada en los límites $x = 0$ y $x = a$. Ambas condiciones son suficientes para determinar la dependencia de cuatro de los coeficientes, por ejemplo, B, C, D, F con A . En la Figura II-8, se ve la representación de la onda tras pasando la barrera y emergiendo detrás de ella.

Después de aplicar las condiciones de continuidad, y hacer los reemplazos correspondientes, se llega a la expresión que relaciona la amplitud de la onda incidente con la de la onda transmitida, esto es:

$$F = \frac{2\alpha\beta e^{-i\alpha a}}{2\alpha\beta \cosh(\beta a) - i(\alpha^2 - \beta^2) \sinh(\beta a)} A \quad (II-38)$$

El parámetro que interesa, es el coeficiente de transmisión, el cual por ser una cantidad medible no puede ser compleja. El coeficiente de transmisión se define como:

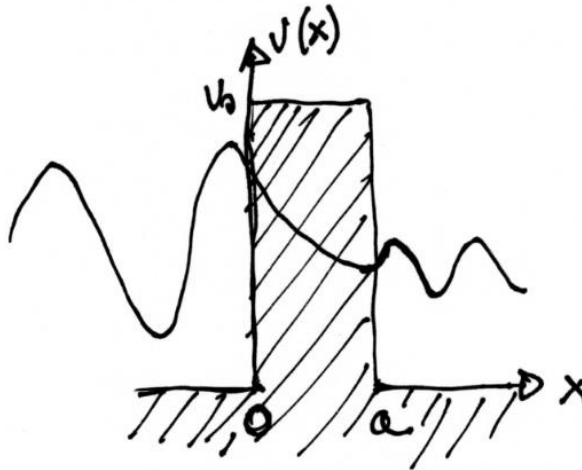


Figura II-8. Efecto túnel.

$$T = \frac{|F|^2}{|A|^2} = \frac{4\alpha^2\beta^2}{4\alpha^2\beta^2 + (\alpha^2 + \beta^2)^2 \sinh^2(\beta a)} \quad (II-39)$$

Ejercicio sugerido: partiendo de las ecuaciones (II-34)-(II-36) y aplicando las condiciones de continuidad, obtener el coeficiente de transmisión, T mostrado en la ecuación (II-39).

En nuestra deducción de la expresión anterior se debe haber encontrado el coeficiente B de la ecuación (II-34) en función de A . Este es:

$$B = \frac{-i(\alpha^2 + \beta^2) \sinh(\beta a)}{2\alpha\beta \cosh(\beta a) - i(\alpha^2 - \beta^2) \sinh(\beta a)} A \quad (II-40)$$

El coeficiente de reflexión está dado por la relación de las ondas reflejadas con las incidentes, es decir:

$$R = \frac{|B|^2}{|A|^2} = \frac{(\alpha^2 + \beta^2)^2 \sinh^2(\beta a)}{4\alpha^2\beta^2 + (\alpha^2 + \beta^2)^2 \sinh^2(\beta a)} \quad (\text{II-41})$$

Se debe notar que en términos generales la cantidad de onda incidente se reparte entre la onda reflejada y la transmitida, en términos de las corrientes, podemos escribir:

$$J_I = J_R + J_T$$

O, en términos de los coeficientes de reflexión y transmisión, se puede escribir:

$$1 = R + T \quad (\text{II-42})$$

Ejercicio sugerido (a): corroborar la ecuación (II-42) con las expresiones (II-39) y (II-41).

Ejercicio sugerido (b): utilizando las expresiones (II-37), encontrar el coeficiente de transmisión para una partícula que llega con una $E = 1,5\text{eV}$ a una barrera de $V_0 = 3\text{eV}$, $a = 3\text{Å}$ ($3 \times 10^{-10}\text{m}$). (Ayuda: todas las variables deben ponerse en función de E y de V_0).

Rta: $T=0,088$

Otro caso que se debería analizar es cuando $E > V_0$. En este, la solución (II-35) debería cambiar a una solución del tipo:

$$\psi_{II}(x) = Ce^{i\beta x} + De^{-i\beta x} \quad \rightarrow \quad 0 \leq x \leq a, \quad \text{donde } \beta = \sqrt{\frac{2m}{\hbar^2}(E - V_0)} \quad (\text{II-43})$$

No se desarrollará esta solución, pero se debe aclarar que la barrera tiene influencia en la solución aun cuando las partículas tengan una energía superior V_0 .

Está claro que el máximo valor que puede tomar T es 1, esto significa que toda la onda ha sido transmitida. Las oscilaciones de la curva de T se deben a interferencias que se producen con las ondas dentro de la barrera. En la Figura II-9, se ven dos curvas, la línea llena corresponde a determinados parámetros de la barrera, por ejemplo los que se han mencionado en el Ejercicio sugerido (b). Si se incrementaría el valor de la altura de la barrera, la curva sería la que se muestra con línea a tramos. La línea punteada vertical corresponde al valor $E = V_0$, debe hacerse notar que en este caso la línea pasa por el valor 1, pues se grafica la relación E/V_0 . Se nota que al incrementarse la altura de la barrera el coeficiente de transmisión disminuye. Un efecto análogo ocurre cuando el ancho de la barrera crece.

Un ejemplo de barrera de potencial es la que se forma en la unión de dos semiconductores dopados con dopantes de cargas opuestas, lo que comúnmente se denominan un diodo tipo P-N. Los portadores deben pasar por esta unión (barrera de potencial) para establecer una corriente entre terminales (este punto lo veremos en detalle más adelante).

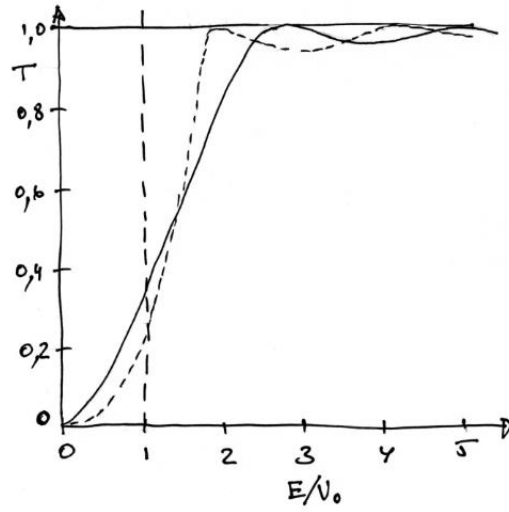


Figura II-9. Coeficiente de transmisión en función de E/V_0 .

Referencias

- Alonso, M., Finn, E. J. *Física*. 1992 Addison-Wesley Iberoamericana S.A.
- Alonso, M., Finn, E. J. *Fundamental University Physics. Vol. III*. 1968 Addison –Wesley Co.
- Beiser, A. *Conceptos de Física Moderna*. 1965 McGraw Hill Book Company.
- Berkeley physics course *Física Cuántica*.– Vol. 4. 1972, Ed. Reverte.S.A.
- Das, A & Melissinos, A. C. *Quantum Mechanics. A modern introduction*. 1988 Gordon and Breach Science Publishers.
- Hlawiczka, P. *Introducción a la electrónica cuántica*. 1977 Editorial Reverté S.A.
- Merzbacher, E. *Quantum Mechanics. Second Edition*. 1970 J. Wiley and Sons.
- Messiah, A. *Mecánica Cuántica*. 1962 Editorial Tecnos S.A.
- Schiff, L. I. *Quantum Mechanics. Third Edition*. 1968 McGraw Hill Kogakusha, LTD.
- Singh, J. *Modern Physics for Engineers*. 2004 Wiley-VCH Verlag GmbH & Co. KgaA.
- Slater, J. C. *Modern Physics*. 1955 McGraw Hill Book Company.Kogakusha Co. LTD.
- Sproull, R. L. *Modern Physics. The quantum physics of atoms, solids and nuclei. 2nd ed.* 1963 J. Wiley & Sons, Inc.
- Tipler, P. A. *Física Moderna*. 2009 Ed. Reverté S.A.
- Tipler, P. A., Mosca, G. *Física para la ciencia y la tecnología. Vol 2C, 5^a ed.* 2010 Ed. Reverté S.A.

CAPÍTULO III

Bandas de energía en sólidos. Modelos cuánticos.

Movimiento de electrones en redes periódicas

*La posesión de la verdad no es más que la lucha
entre las revelaciones impuestas de los hombres.
Unas sucumben y son meras voces de otro espíritu.
Otras triunfan, y añaden nueva voz a la de la naturaleza.*

JOSÉ MARTÍ, poeta Cubano
28-1-1853:19-5-1895

Ya se ha mencionado que la Física no puede interpretar a la naturaleza. Lo que hace la Física es desarrollar modelos que simulen a la naturaleza y estudia esos modelos. Si los resultados concuerdan con las mediciones realizadas del fenómeno que se pretende estudiar, eso quiere decir que el modelo ha sido bien desarrollado y representa a la naturaleza, si esto no ocurre, se deberá modificar el modelo o directamente sustituirlo por otro, y comenzar de nuevo.

En este capítulo se desarrollarán dos modelos que representan el movimiento de los portadores dentro de un sólido. Nuevamente aquí, la explicación proviene de la Mecánica Cuántica, la Física Clásica no está en capacidad de hacerlo. El universo en el que todo esto se mueve es el de las funciones de onda gobernadas por potenciales, los cuales, en el primer caso, el modelo del electrón libre, serán inexistentes y en el segundo, el modelo del electrón dentro de un cristal, serán finitos y tendrán un formato periódico.

Lo importante es, ¿qué conocimiento brindan estos modelos? Y ¿qué conocimientos se pueden extraer de ellos? Para empezar a estudiarlos, se debe comprender que los electrones son partículas ligadas a los átomos. Los átomos en los sólidos, están ligados en general por fuerzas electrostáticas. Se pueden definir dos grandes grupos de sólidos, aquellos en los que los átomos están distribuidos aleatoriamente en el espacio, sin ningún

orden, son los que se llaman sólidos amorfos, como en la Figura III-1, un ejemplo de esta estructura son los vidrios. Luego están los que tiene una distribución espacial periódica de sus átomos en el espacio, estos se llaman sólidos cristalinos. La sal, el ClNa, es un ejemplo muy común de esta clase, como se ve en la Figura III-2. Los electrones que están más débilmente ligados al átomo, que forman las ligaduras con sus vecinos, son llamados electrones de valencia. Los electrones libres son algunos de estos electrones que han abandonado al átomo debido a alguna circunstancia y pueden moverse a través del sólido sin verse afectados por los potenciales existentes.

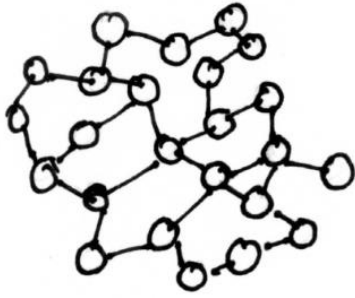


Figura III-1. Sólido amorfo

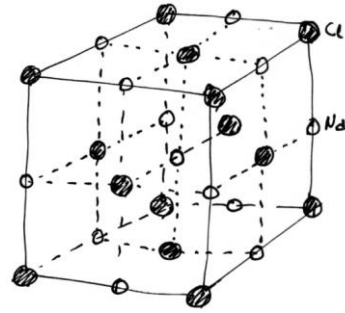


Figura III-2. Sólido cristalino. ClNa

Si bien, los sólidos se encuentran en la naturaleza en forma tridimensional, el estudio de los modelos que los representan se hará en una dimensión. Esto no le quita generalidad a los resultados, sino que brinda una mayor facilidad al tratamiento matemático del problema.

Si se tiene un solo átomo en el espacio, el potencial eléctrico que genera es del tipo coulombiano, por lo que, cualquier carga próxima lo sentirá con una intensidad que variará con la inversa de la distancia ($1/r$), como en la Figura III-3.

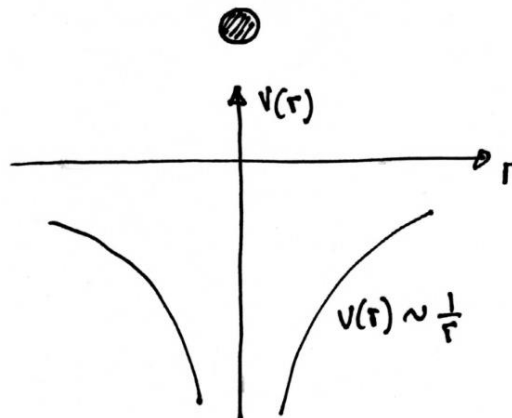


Figura III-3. Potenciales de un átomo.

Si se construyera una cadena de átomos iguales, siendo cada uno de ellos como el hipotético que se tiene en la Figura III-3, y los ubicáramos a una distancia a entre ellos, se generaría una superposición de las colas (parte superior) de los potenciales tal que ellos se anularían de a pares, haciendo que disminuyan los potenciales entre los extremos, dejando un panorama como el de la Figura III-4. En la figura, se nota lo que se recién se mencionó, los

potenciales entre los átomos disminuyen con respecto a los de los extremos. Se debe notar también que entre los átomos los potenciales forman barreras y pozos. Cuando los electrones se encuentran en niveles de energía tales como el E_1 o E_2 se los llama electrones ligados. Eventualmente los electrones en E_2 podrían, por efecto Túnel, pasar a pozos vecinos. Mientras que los que están en los estados E_1 por estar muy profundos, no lo podrán hacer. En cambio los electrones en E_3 no están ligados, son electrones libres, o electrones de conducción. Es muy importante entender con estos ejemplos que algunos sistemas físicos solamente poseen valores determinados de energía para sus electrones, valores discretos, no continuos, esos valores están cuantizados. Las partículas en un pozo de potencial fueron un ejemplo de esto. En general esos son estados ligados.

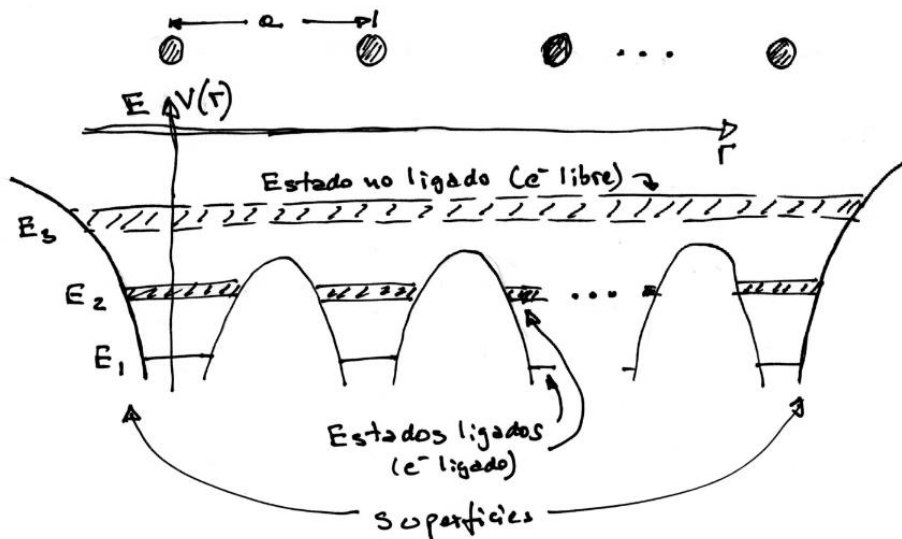


Figura III-4. Cadena lineal de varios átomos, potenciales y niveles de energía generados.

III.1. El modelo del electrón libre

El modelo del electrón libre es muy simple, este representa el movimiento de un electrón cuando no existe ninguna clase de potencial que interactúe con él, por ejemplo, electrones en el nivel E_3 , los cuales se moverán entre los extremos de la cadena de átomos, que esquemáticamente se ven en la Figura III-4. El movimiento de estos electrones a través del cristal, el cual en este caso se representará como una cadena lineal de átomos, se puede deber a causas como excitaciones térmicas o fuerzas externas debidas a campos eléctricos y/o magnéticos que estén aplicados sobre el cristal. La cadena se considera suficientemente larga de manera tal que los extremos no tengan influencia sobre la zona de estudio.

Es necesario considerar el movimiento ondulatorio de los electrones. Su desplazamiento será entre los extremos de la cadena como si fuese una onda estacionaria, la cual se encuentra fija en ambos extremos. Esto es debido a que consideramos que los electrones no pueden abandonar el sólido (la cadena lineal en nuestro caso), por lo que los potenciales de los extremos son, en este caso, impenetrables. La longitud total de la cadena con N átomos separados una distancia a entre ellos, es $L=Na$. Como estamos planteando condiciones muy

parecidas a las del pozo infinito, entre los extremos debe entrar un número entero de semi longitudes de onda,

$$L = n \frac{\lambda_n}{2}, \quad n = 1, 2, 3, \dots, N \quad (III-1)$$

El número de onda es:

$$k_n = \frac{2\pi}{\lambda_n} = \frac{2\pi}{2L/n} = \frac{\pi}{L} n = \frac{\pi}{Na} n \quad (III-2)$$

Debido a la naturaleza ondulatoria del electrón, la cantidad k_n no es continua, sino discreta. En la ecuación (I-3) se vio como se relacionaba la longitud de onda de de Broglie con la cantidad de movimiento, $\lambda = h/p$, por lo que,

$$\lambda = \frac{h}{p} = \frac{2\pi h}{p} \rightarrow \frac{\lambda_n}{2\pi} = \frac{1}{k_n} = \frac{h}{p_n} \Rightarrow p_n = \hbar k_n \quad (III-3)$$

La función de onda para un electrón con esta cantidad de movimiento será una onda plana como la siguiente:

$$\psi(x) \approx e^{ikx} \quad (III-4)$$

La cantidad p_n (en la ecuación (III-3)) puede ser positiva o negativa, indicando que el movimiento del electrón puede ser en cualquiera de las dos direcciones del eje x . Se puede reescribir esta expresión pensando en tres dimensiones, con lo cual queda:

$$\psi(\mathbf{r}) \approx e^{i\mathbf{k} \cdot \mathbf{r}} \quad (III-5)$$

En ambos casos

$$\int |\psi|^2 d\mathbf{r} = 1 \quad (III-6)$$

Lo cual asegura que la partícula se encontrará en algún lugar de la red.

Como se ha supuesto que $V = 0$, entonces toda la energía de la partícula será cinética, por lo que:

$$E = \frac{p^2}{2m} = \frac{\hbar^2 k_n^2}{2m_e} \quad (III-7)$$

Reemplazando k_n tenemos:

$$E = \frac{\hbar^2 \pi^2}{2m_e N^2 a^2} n^2 \quad (III-8)$$

Los valores k_n , indican el número de onda de cada uno de los estados de los electrones, cuyas longitudes de onda son λ_n . Si se piensa en la diferencia entre dos estados sucesivos, podemos escribir:

$$\Delta k = k_n - k_{n-1} = \frac{\pi}{Na} n - \frac{\pi}{Na} (n-1) = \frac{\pi}{Na} \quad (\text{III-9})$$

Si $N \gg 1 \Rightarrow \Delta k_n \rightarrow 0$

La variación de energía entre estados será entonces:

$$\Delta E = \frac{\hbar}{2m_e} (\Delta k_n)^2 \quad \text{si } N \gg 1 \text{ y } \Delta k_n \rightarrow 0 \quad \therefore \Delta E \rightarrow 0 \quad (\text{III-10})$$

Esto indica que la diferencia existente entre dos estados de energía contiguos, cuando existe una gran cantidad de átomos, tiende a ser casi nula, esto quiere decir que se puede considerar que los electrones se mueven en un continuo de energías.

Se ve en las expresiones (III-1), (III-2) y (III-8) que la longitud de onda, número de onda y energía dependen de n , solamente toman valores discretos, son magnitudes que están cuantizadas. Cuando se tenga una cantidad muy grande de átomos, los valores de k y de E habrán llegado a sus valores máximos, como se ve a continuación:

$$n \rightarrow N \quad k = \frac{\pi n}{Na} \rightarrow k_{max} = \frac{\pi}{a}, \quad (\text{III-11})$$

y

$$E = \frac{\hbar^2 k^2}{2m_e} \rightarrow E_{max} = \frac{\hbar^2 \pi^2}{2m_e a^2} \quad (\text{III-12})$$

Como el conjunto de estados de energía conforman una banda de energías, la expresión (III-12) nos indica cuál será su valor máximo (E_{max}). Además se deduce que esta expresión es independiente del número de iones que tenga la red, lo cual es esperado, pues si se agregan iones, se están agregando estados dentro de la banda, pero la periodicidad de la red, que determina el k_{max} (ec.(III-11)) no cambia. Se debe entender, también, que la banda en esta estructura energética va desde el valor 0 hasta su máximo E_{max} , por lo que este valor está dando el ancho de esa banda de energías.

En la ecuación (III-3) se obtuvo la relación que existe entre la cantidad de movimiento y el número de onda ($p = \hbar k$), se sabe que la cantidad de movimiento (p) es un vector, por lo que k es un vector, generalmente utilizamos su módulo. En la ecuación (III-11) aparece el módulo del máximo del número de onda, por lo que los valores que puede tomar son:

$$k_{max} = \pm \frac{\pi}{a} \quad (\text{III-13})$$

Estos números están relacionados con los lugares de reflexión de las ondas en los átomos de la cadena.

Finalmente se representará el comportamiento de la energía de un electrón completamente libre, de acuerdo a la expresión de la ec. (III-7), la cual, la podemos ver en la Figura III-5. Debido a la dependencia cuadrática de E con el número de onda k , la energía tiene una forma parabólica.

Las conclusiones que podemos sacar del modelo del electrón libre ($V = 0$), son: a) los electrones se mueven en un continuo de energías, b) debido a su naturaleza ondulatoria, los parámetros que los caracterizan son discretos, y c) se pueden obtener valores máximos de energía y cantidad de movimiento de los electrones para la banda donde estos se encuentren.

Otra aclaración debe ser hecha: la energía, la cantidad de movimiento y algunas otras magnitudes que se utilizarán son dependientes del número de onda k . Esto significa

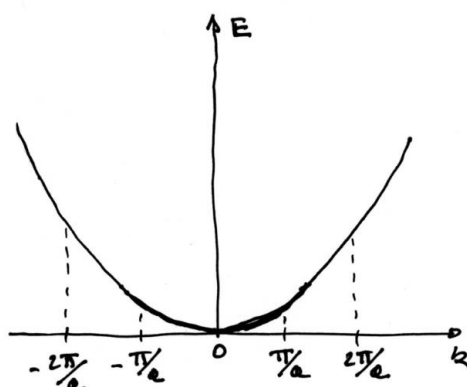


Figura III-5. Curva de energía en función del número de onda.

que el espacio de movimiento de las partículas es diferente al tridimensional (x,y,z) , en el cual la unidad de medida es una longitud (metros, centímetros, nanómetros, etc.), la que corresponda de acuerdo a al sistema estudiado. El movimiento en un espacio dependiente de k , el espacio de los momentos, con coordenadas (k_x, k_y, k_z) tendrá dimensión de $(\text{longitud})^{-1}$. De ahora en adelante se verán muy frecuentemente diferentes magnitudes en este espacio.

III.2. Densidad de estados

La densidad de estados, $N(E)$, es un concepto importante que se comenzará a manejar a partir de este punto. Esta cantidad define el número de estados electrónicos por unidad de energía y unidad de dimensión espacial. De acuerdo al caso que se este tratando, puede ser unidad de volumen, unidad de área o unidad de longitud, dependiendo si se calcula la densidad de estados en una muestra tri, bi o unidimensional, respectivamente. En cada estado electrónico es donde se pueden ubicar los electrones, máximo 2 por estado de acuerdo al principio de exclusión de Pauli. Es necesario establecer como se distribuyen los electrones, y por extensión, los estados electrónicos en una banda, desde su mínimo hasta su máximo.

Como hasta el momento se ha trabajado con sistemas unidimensionales, se seguirá en esa línea y luego se ampliará a más dimensiones. Al principio se mencionó que tratar con sistemas unidimensionales no es una simplificación al problema que se estudia. Desde el punto de vista teórico la unidimensionalidad aporta a la simplicidad matemática, mientras que desde el punto

de vista experimental podemos mencionar que se han desarrollado sistemas con diferentes dimensionalidades, no entraremos en detalle sobre este punto en este libro, pero uno de ellos, los sistemas unidimensionales están representados por los conductores cuánticos “quantum wires” los cuales son ampliamente estudiados y utilizados en todo lo que se refiere a la miniaturización de componentes y circuitos. La densidad de estados para este tipo de sistema la podemos definir como los estados energéticos alrededor de E (en un rango dE) por unidad de longitud, como se ve en la Figura III-6.

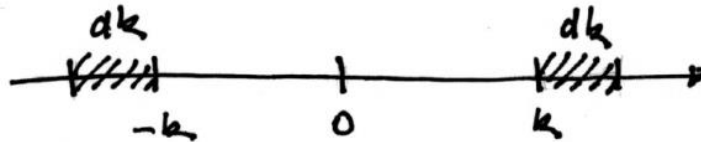


Figura III-6. Distribución de estados en una dimensión.

De acuerdo a esto se tiene:

$$N(E)dE = \frac{dk}{\pi} \tag{III-14}$$

Como la energía es la de una partícula libre, se tiene:

$$E = \frac{\hbar^2 k^2}{2m_e} \rightarrow k = \frac{\sqrt{2m_e E}}{\hbar} \rightarrow dk = \frac{\sqrt{m_e}}{\sqrt{2}\hbar} \frac{1}{\sqrt{E}} dE \tag{III-15}$$

No se debe olvidar que, en cada estado energético y sin violar ninguna regla de la mecánica cuántica, pueden existir 2 electrones, uno con spin hacia arriba y el otro hacia abajo, eso significa que la densidad de estados debe ser multiplicada por 2. Por lo que, haciendo los reemplazos en la ecuación (III-14), con los valores obtenidos en (III-15), queda:

$$N(E) = \frac{\sqrt{2m_e}}{\pi\hbar} \frac{1}{\sqrt{E}} \tag{III-16}$$

En un sistema bidimensional, la definición sería: “el número de estados electrónicos disponibles por unidad de área en el rango $E + dE$ ”. Este rango de energías está representado en el espacio k por un círculo de área πk^2 . En este espacio la variación de energías mencionada se representa en el rango $k + dk$. En este espacio bidimensional es:

$$A = \pi k^2 \rightarrow dA = 2\pi k dk$$

En un espacio bidimensional el área que ocupa cada estado electrónico es $(2\pi\hbar)^2$ como se puede ver en la Figura III-7,

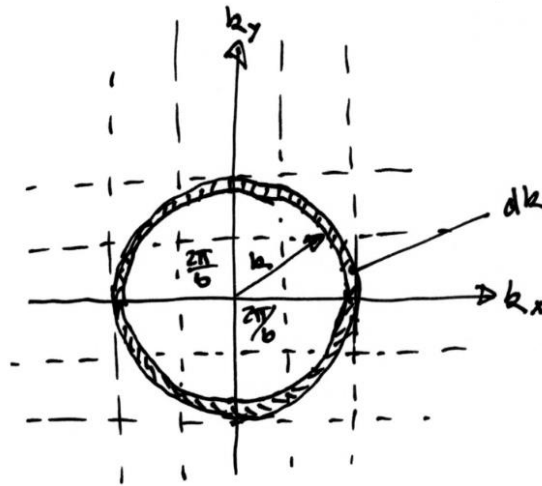


Figura III-7. Desarrollo bidimensional del espacio k . Área ocupada por cada estado electrónico.

En esta estructura se tiene:

$$N(E)dE = \frac{2\pi k dk}{(2\pi)^2} = \frac{k dk}{2\pi} \tag{III-17}$$

En la ecuación (III-15) se han obtenido los valores de k y dk , si los reemplazamos en la ecuación (III-17), se tiene:

$$N(E) = \frac{m_e}{\pi h^2} \tag{III-18}$$

Esta, a diferencia de la anterior, no tiene dependencia con la energía, es constante. Finalmente en un sistema tridimensional, el cual se definiría como la cantidad de estados por unidad de volumen, es:

$$N(E) = \frac{\sqrt{2}m_e^{3/2}}{\pi^2 h^3} \sqrt{E} \tag{III-19}$$

Se debe recordar que se ha trabajado con longitudes, áreas o volúmenes unitarios y se han tenido en cuenta el hecho que por estado electrónico la ocupación puede ser de 2 electrones.

Ejercicio sugerido: deducir la expresión de la densidad de estados para un sistema tridimensional.

Solución: ecuación (III-19).

El número de electrones que se pueden acomodar en un espacio tridimensional por unidad de volumen, entre 0 y la energía E , será:

$$n = \int_0^E N(E)dE = \frac{\sqrt{2}m_e^{3/2}}{\pi^2 h^3} \int_0^E E^{1/2} dE = \frac{\sqrt{2}m_e^{3/2}}{\pi^2 h^3} \frac{2}{3} E^{3/2} = \frac{\sqrt{8}m_e^{3/2}}{3\pi^2 h^3} E^{3/2} \tag{III-20}$$

Como se está describiendo un estado de electrones libres, se supone que se está estudiando un metal. Bajo estas circunstancias, si el metal se encuentra en su estado fundamental, el sistema estará en un estado de mínima energía, es decir todos los electrones estarán ocupando los niveles de energía más bajos, sin violar principios cuánticos. Por lo tanto, dentro de una banda tendremos estados ocupados hasta un cierto nivel de energía y estados desocupados para niveles de energías superiores a este. Ese nivel de energía que marca el límite entre estados ocupados y desocupados, se lo llama **nivel de Fermi**, ϵ_F . Si la cantidad de electrones que se tienen por unidad de volumen es n_e , el valor del nivel de Fermi será:

$$n_e = \frac{\sqrt{8}m_e^{3/2}}{3\pi^2\hbar^3} \epsilon_F^{3/2}, \Rightarrow \epsilon_F = \frac{\hbar^2}{2m_e} (3\pi^2 n_e)^{2/3} \quad \text{(III-21)}$$

gráficamente:

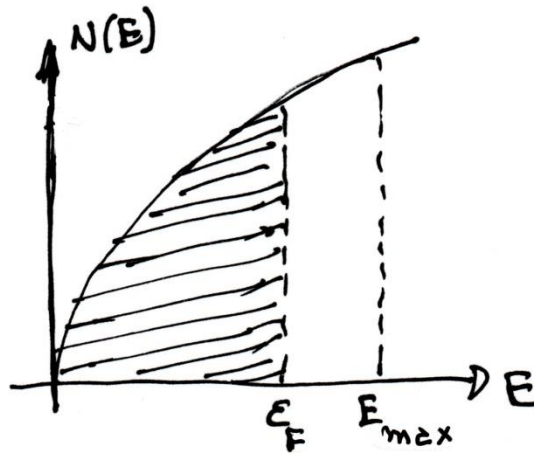


Figura III-8. Densidad de estados en el sólido.

Ejercicio sugerido: calcular la energía de Fermi del Cu, sabiendo que tiene 29 electrones.

Rta: 7.04eV

Así como se ha planteado hasta ahora, todos los electrones se encuentran por debajo del nivel de Fermi (Figura III-9-a). Si se impone sobre el sólido alguna excitación (un campo eléctrico, incremento de temperatura, etc.), esta energía extra que recibe el sistema hace que algunos electrones, los más próximos al nivel de Fermi, puedan cruzarlo ocupando niveles de mayor energía dentro de la banda (Figura III-9-b). Este efecto se verá reflejado en la curva de la densidad de estados, los estados ocupados no son los que se muestran en la Figura III-8, sino que algunos de los electrones han superado el nivel de Fermi ocupando estados superiores a este, como se ve en la Figura III-10.

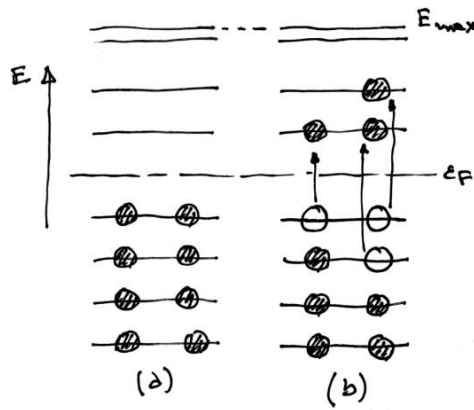


Figura III-9. Ubicación de los electrones en estado fundamental (a) y excitado (b)

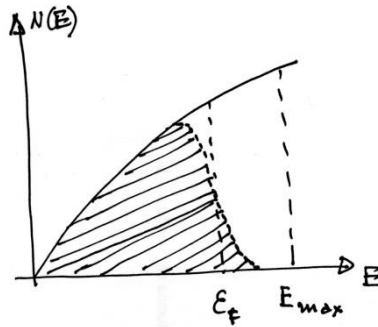


Figura III-10. Densidad de estados en el sólido con estados excitados.

III.3. El Teorema de Bloch



Félix Bloch
23-10-1905, 10-9-1983
Premio Nobel de Física 1952

A partir de ahora, se debe pensar que el espacio en donde se moverán los electrones es el de un cristal, una distribución periódica de átomos en el espacio. Esto quiere decir que se establece el hecho que los potenciales serán periódicos. A diferencia del modelo del electrón libre, donde el potencial era inexistente y no afectaba a la función de onda, en este caso sí se debe pensar en cómo la afecta. Seguramente, la función de onda variará sus amplitudes de acuerdo a la periodicidad de la red. Se puede escribir la función de onda como:

$$\psi(\mathbf{r}) = e^{i\mathbf{k} \cdot \mathbf{r}} u(\mathbf{r}) \tag{III-22}$$

Donde $u(r)$ es una función periódica que modula la amplitud de la onda. Si se volviera al caso unidimensional, donde los electrones se mueven en una red con un espaciado a como el de la Figura III-4, se tendría:

$$\psi(x) = e^{ikx} u(x) \tag{III-23}$$

Como se ha dicho que la función u es periódica, debe cumplir la condición:

$$u(x + na) = u(x) \tag{III-24}$$

Con estas expresiones se está relacionando el valor de la función de onda en un punto de un cristal con otro equivalente en cualquier otro lugar del cristal. Las ecuaciones (III-23) y (III-24) definen el teorema de Bloch. Gráficamente se pueden ver de la siguiente manera:

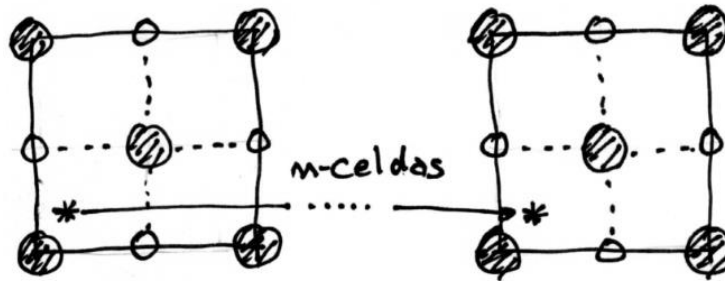


Figura III-11. Teorema de Bloch. Equivalencia de las soluciones en lugares a distancias de número entero de períodos.

En la Figura III-11 se muestran, esquemáticamente, dos caras de alguna celda cristalina separadas por n -celdas. El teorema de Bloch establece que debido a la periodicidad del cristal, las propiedades físicas que se tienen en un punto determinado (por ejemplo, la estrella en la celda de la izquierda), son iguales a las que se tendrán en un punto similar, aún cuando se encuentren separados por n -celdas cristalinas (en la misma figura, la estrella en la celda de la derecha). Se puede definir aquí a la celda unidad, esta es la menor porción de un cristal con la cual se puede generar la totalidad del mismo, aplicando propiedades de simetría.

Ejercicio sugerido: haga la demostración del teorema de Bloch, recordando que la función $u(x)$ es periódica con el mismo período de la red.

Se puede ver como la red influye en el movimiento de los electrones. La función de onda que representa al electrón es la descrita por la ecuación (III-23). Se puede suponer que la función está normalizada a 1, es decir:

$$\int \psi^* \psi dx = \int e^{-ikx} u^* e^{ikx} u dx = \int u^* u dx = 1$$

La cantidad de movimiento promedio será:

$$\langle p \rangle = \int \psi^* \left(-i\hbar \frac{d}{dx} \right) \psi dx \quad (III-25)$$

Si se desarrolla la derivada de la función de onda:

$$\frac{d\psi}{dx} = \frac{d}{dx} \left[e^{ikx} u(x) \right] = ike^{ikx} u(x) + e^{ikx} \frac{du(x)}{dx}$$

y se reemplaza en la integral,

$$\langle p \rangle = \int e^{-ikx} u^* i\hbar \left[ike^{ikx} u + e^{ikx} \frac{du}{dx} \right] dx = \hbar k \int u^* u dx + \int u^* \left(-i\hbar \frac{d}{dx} \right) u dx$$

La primera integral está normalizada y vale 1, mientras que la segunda tiene la misma forma que la ecuación (III-25), por lo tanto queda:

$$\langle p \rangle = \hbar k + \langle p_u \rangle \quad (III-26)$$

El valor $\langle p_u \rangle$ es el valor medio de la cantidad de movimiento de un electrón como resultado de su interacción con la red. Por lo que se puede concluir, que un electrón moviéndose por la red tiene una cantidad de movimiento que está compuesta por dos términos, el del electrón casi libre ($\hbar k$) y el momento debido a su interacción con la red ($\langle p_u \rangle$).

III.4. El modelo de Kronig-Penney

Se podría pensar que nuestros electrones se mueven en un nivel como el E_2 de la Figura III-4, esto quiere decir que ellos, en este momento, no ven a los potenciales como un suelo empedrado debajo de ellos, sino que ahora están entre las piedras. Se describirá ahora el modelo de Kronig-Penney, el cual fue desarrollado en el año 1930 (R. de L. Kronig and W. G. Penney, Proc. Roy. Soc. (London) A 130 (1931) 499).



Ralph Kronig
10-3-1904, 16-11-1995
24-6-1909, 3-3-1991



William George Penney
Baron Penney

Lo que propusieron Kronig y Penney para explicar el comportamiento de los electrones en un cristal, fue una cadena unidimensional periódica de átomos representados por pozos y barreras de potencial, como se muestra en la Figura III-11.

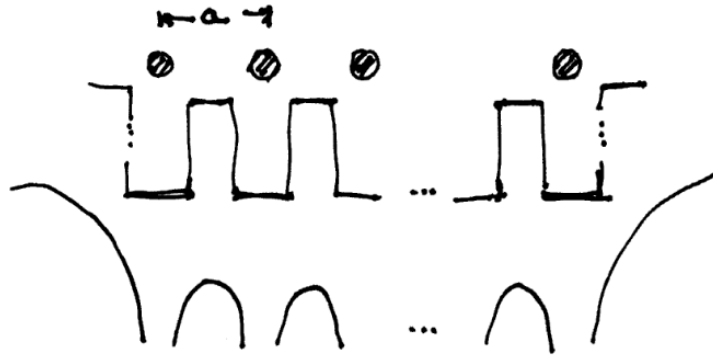


Figura III-11. Diagrama esquemático del modelo de Kronig-Penney.

En la Figura III-11, se ha esquematizado la cadena de átomos desde un extremo hasta el otro. Para desarrollar este modelo nos restringiremos a alguna zona de la cadena lejana a los extremos, de manera tal que estos no interfieran con las soluciones que se obtendrán. La región que se utilizará es como la que se muestra en la Figura III-12.

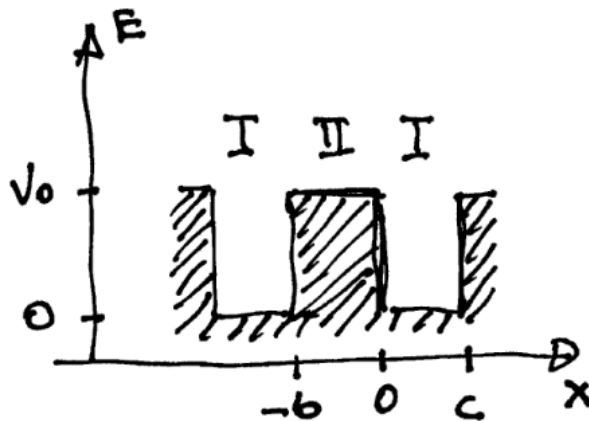


Figura III-12. Modelo de Kronig-Penney, potencial periódico.

El potencial es periódico, con período: $a = c+b$, las partículas se moverán con una energía $E < V_0$. Estas condiciones hacen que el espacio quede dividido en dos regiones,

$$\text{Región I, } V(x) = 0, 0 \leq x \leq c \quad \text{y} \quad \text{(III-27)}$$

$$\text{Región II, } V(x) = V_0, -b \leq x \leq 0. \quad \text{(III-28)}$$

La ecuación que se debe solucionar, que describe el estado del sistema es:

$$\frac{d^2\psi(x)}{dx^2} + \frac{2m}{\hbar^2} [E - V(x)]\psi(x) = 0 \quad \text{(III-29)}$$

En la región I la solución será:

$$\psi_I(x) = Ae^{i\alpha x} + Be^{-i\alpha x} \quad (III-30)$$

Mientras que en la región II, será:

$$\psi_{II}(x) = Ce^{\beta x} + De^{-\beta x}, \quad (III-31)$$

donde,

$$\alpha = \sqrt{2mE / \hbar^2}, \quad \beta = \sqrt{2m(V_0 - E) / \hbar^2} \quad (III-32)$$

La función de Bloch (ec. (65)) es:

$$\psi(x) = u(x)e^{ikx} \quad (III-33)$$

Si sustituimos ψ_I y ψ_{II} en (III-33) y obtenemos los valores u_I y u_{II} , tenemos:

$$\begin{aligned} u_I(x) &= Ae^{i(\alpha-k)x} + Be^{-i(\alpha+k)x} \\ u_{II}(x) &= Ce^{(\beta-ik)x} + De^{-(\beta+ik)x} \end{aligned} \quad (III-34)$$

Para obtener los coeficientes A , B , C y D se deben aplicar las condiciones de continuidad y periodicidad de la función y su derivada primera, esto es:

$$u_I(0) = u_{II}(0), \quad \left. \frac{du_I}{dx} \right|_{x=0} = \left. \frac{du_{II}}{dx} \right|_{x=0} \quad (III-35)$$

$$u_I(c) = u_{II}(-b), \quad \left. \frac{du_I}{dx} \right|_{x=c} = \left. \frac{du_{II}}{dx} \right|_{x=-b} \quad (III-36)$$

De estas ecuaciones se pueden determinar las constantes A , B , C y D , las cuales no deben ser nulas, pero si se debe anular el determinante del sistema. La resolución de este nos lleva a la ecuación:

$$\frac{\beta^2 - \alpha^2}{2\alpha\beta} \sinh(\beta b) \sin(\alpha c) + \cosh(\beta b) \cos(\alpha c) = \cos k(b+c) \quad (III-37)$$

Ejercicio sugerido: desarrollar el determinante y obtener la expresión (III-37).

La ecuación (III-37) relaciona la energía, que es parte de las constantes α y β (ec. (III-32)) con el vector de onda k . Esta expresión tiene dos miembros bien diferenciados, el del lado izquierdo, que se denominará $f(E)$, puede crecer sustancialmente en función de la participación de funciones hiperbólicas en la ecuación. Estas funciones no se encuentran limitadas a un rango específico de valores, además, eventualmente el valor α podría ser muy pequeño comparado con β , y el cociente que multiplica al $\sinh(\beta b) \sin(\alpha c)$, podría tomar valores grandes,

muy superiores a +1 o -1. Por otro lado, el miembro de la derecha es una función circular cuyos valores jamás superaran el rango ± 1 . Por lo que, así como se ha planteado la ecuación (III-37) solamente tendrá significado físico, cuando los valores del miembro de la izquierda se encuentren en el rango de los valores del miembro de la derecha (coseno circular), es decir, entre +1 y -1. Esta situación se muestra en la Figura III-13.

En la Figura III-13 se pueden ver zonas grisáceas y zonas blancas, en la primeras se cumple la condición que la ecuación (III-37) se encuentra entre los valores ± 1 ; estas reflejan condiciones físicas reales, son estados permitidos del sistema. Las zonas blancas representan lugares donde la ecuación (III-37) no satisface la igualdad, son los estados prohibidos del sistema.

En la Figura III-14, se puede ver que los estados permitidos forman bandas de energía que se van ensanchando a medida que la energía E crece y siempre hay estados prohibidos que se intercalan entre los estados permitidos. Esta estructura energética de bandas permitidas seguidas por bandas prohibidas es de suma importancia para entender a los materiales cristalinos. En la Figura III-14 se ha representado la ecuación (III-37) graficando $f(E)$ en función de $k(b+c)$ (argumento del coseno). Esa misma ecuación puede ser graficada poniendo E en función del número de onda k

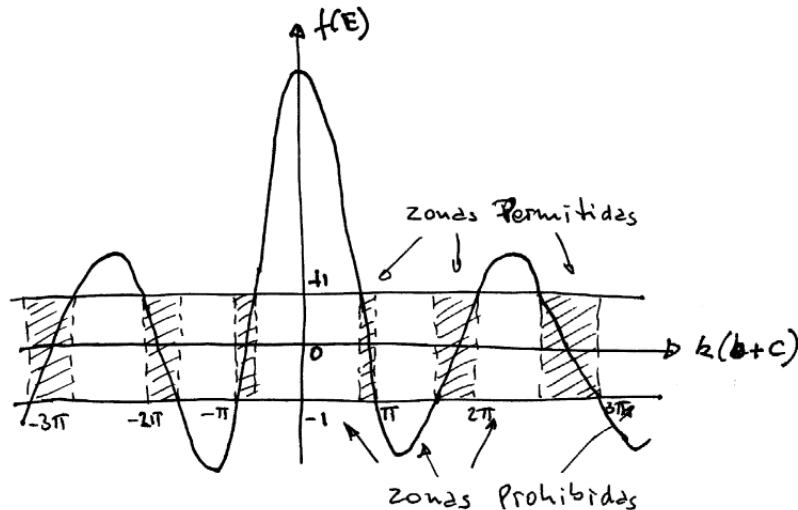


Figura III-13. Curva de la solución propuesta por el modelo de Kronig-Penney.

En la Figura III-14, se han graficado las curvas de energía en función de k , una que describe al electrón libre (parábola continua) y la otra la del modelo de Kronig-Penney (curvas llenas segmentadas), estas últimas se han desplazado ex profeso para ver su comportamiento, sino deberían estar superpuestas, sobresaliendo de la curva del electrón libre, solamente la parte superior e inferior de cada tramo de la curva de Kronig-Penney. En las líneas segmentadas se ve que existen discontinuidades en los puntos $k=\pm \pi/a, \pm 2\pi/a, \pm 3\pi/a$, etc. Estos saltos entre curvas de energía corresponden a las zonas de estados prohibidos, bandas prohibidas, en las cuales los electrones no pueden estar.

La zona que abarca el rango, en el eje k , entre $-\pi/a$ y $+\pi/a$ se llama, la primera zona de Brillouin, la que va desde $+\pi/a$ a $+2\pi/a$ y desde $-\pi/a$ a $-2\pi/a$ es la segunda zona de Brillouin, etc. Debido a la periodicidad de la red, se pueden mover estos segmentos en periodos enteros

hacia la zona central $(-\pi/a, +\pi/a)$ obteniéndose la zona reducida, como se ve en la Figura III-15, en donde hay una descripción completa de los estados de energía en una única zona.

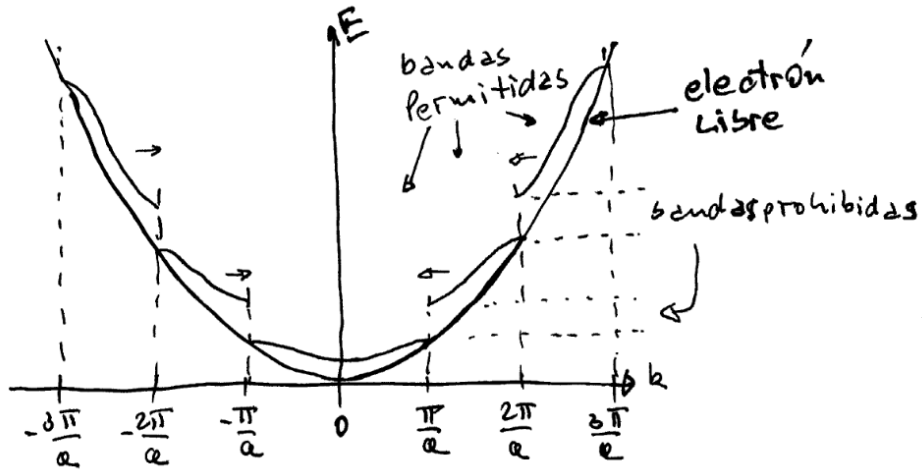


Figura III-14. Modelo de Kronig-Penney (MKP) y del electrón libre. La curva correspondiente al modelo del electrón libre, se ha desplazado verticalmente para diferenciarla de la del MKP, sino estarían superpuestas y solamente sobresaldrían los extremos de la curva del MKP.

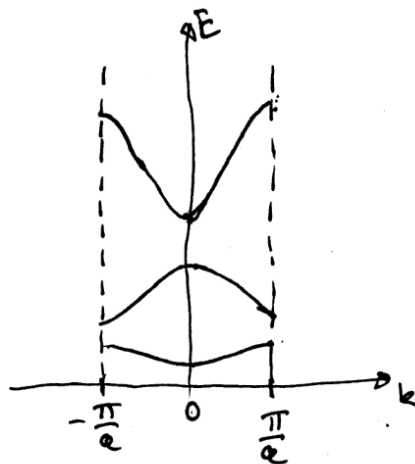


Figura III-15. Modelo de Kronig-Penney, zona reducida.

Se debe entender que los electrones se mueven bajo la acción de fuerzas externas e internas. Las fuerzas externas son debidas, por ejemplo, al campo que apliquemos, pero las internas se deben a los potenciales generados por los átomos, los cuales no son manejables, es necesario desarrollar alguna explicación sobre el movimiento de las partículas, que sea independiente de las fuerzas internas. La velocidad de los electrones en el cristal, será la velocidad del grupo de ondas,

$$v_g = \frac{d\omega}{dk}, \quad E = h\nu = h\omega, \quad \Rightarrow \quad v_g = \frac{1}{h} \frac{dE}{dk} \quad (III-38)$$

La fuerza externa F puede ser descripta por la siguiente expresión:

$$F = \frac{dp}{dt} = \frac{d(hk)}{dt} = h \frac{dk}{dt} \quad (III-39)$$

La aceleración de los electrones será:

$$a = \frac{dv_g}{dt} = \frac{1}{h} \frac{d}{dt} \left[\frac{dE}{dk} \right] = \frac{1}{h} \frac{d}{dk} \left[\frac{dE}{dk} \right] \frac{dk}{dt} = \frac{1}{h} \frac{d^2 E}{dk^2} \frac{dk}{dt} \quad (\text{III-40})$$

Reemplazando (III-39) en (III-40) se tiene:

$$a = \frac{1}{h^2} \frac{d^2 E}{dk^2} F \quad (\text{III-41})$$

Si recordamos la 2da ley de Newton, se tiene que $F = ma$, o $a = F/m$. Aquí es conveniente definir la masa efectiva m^* , la cual tendrá en cuenta la fuerza externa representada por F de acuerdo a Newton, y a es la aceleración real del electrón debida a F y la interacción con la red. Por lo tanto:

$$a = \frac{F}{m^*} = \frac{F}{h^2 / \frac{d^2 E}{dk^2}} \Rightarrow m^* = \frac{h^2}{\frac{d^2 E}{dk^2}} \quad (\text{III-42})$$

El extremo izquierdo de la ecuación (III-42) nos muestra la 2da ecuación de Newton y describe el movimiento del electrón dentro de la estructura cristalina. Se ve allí que si la masa de la partícula es la masa efectiva, la partícula se moverá como si fuera libre. Este es un concepto que se debe remarcar, la masa efectiva es un concepto útil solamente para describir el movimiento de los electrones en una red cristalina periódica. Esta cantidad no es una masa en el sentido ordinario de su definición, debido a que ella depende de la energía, no tiene propiedades gravitacionales ni inerciales, y su valor no será constante. Puede tomar valores mayores o menores que la masa en reposo del electrón e inclusive puede adquirir valores positivos tanto así como negativos, si la derivada segunda fuese negativa. Se puede demostrar que la masa efectiva es la masa gravitacional del electrón cuando este se mueve como un electrón libre. La energía para un electrón libre es (ec. (III-12)):

$$E = \frac{h^2 k^2}{2m_e}$$

Su derivada segunda será:

$$\frac{d^2 E}{dk^2} = \frac{d}{dk} \left[\frac{dE}{dk} \right] = \frac{d}{dk} \left[\frac{h^2}{2m_e} 2k \right] = \frac{h^2}{m_e},$$

Si reemplazamos en la ec. (III-34) da:

$$m^* = \frac{h^2}{\frac{h^2}{m_e}} = m_e \quad (\text{III-43})$$

Se debe ver ahora cómo será el comportamiento de un electrón en la red cristalina. En la Figura III-17 se observa la curva de energía de un electrón libre y la banda de acuerdo al modelo de Kronig-Penney.

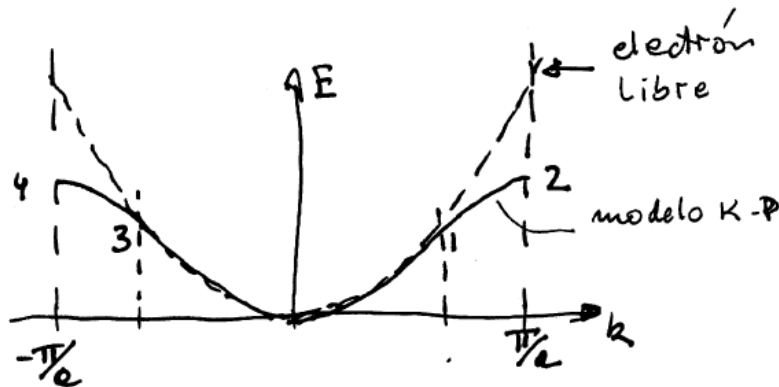


Figura III-16. Variación de la energía de un electrón en el modelo de K-P con respecto al electrón libre.

Utilizando esta curva se hará el análisis del comportamiento de algunas magnitudes relativas al movimiento de los electrones, por ejemplo su velocidad. Claramente, si no existe ninguna fuerza que obligue al electrón a moverse este se quedará en el origen de ambas curvas, por lo que, es necesario aplicar, por ejemplo, un campo eléctrico para que este abandone esta posición y comience a moverse. La velocidad está dada por la ecuación (III-38), en el caso del electrón libre tenemos:

$$v_g = \frac{1}{\hbar} \frac{dE}{dk}, \text{ como } E = \frac{\hbar^2 k^2}{2m_e} \Rightarrow v_g = \frac{\hbar}{m_e} k \tag{III-44}$$

Se puede ver de esta expresión que en el caso del electrón libre, la velocidad tiene una dependencia lineal con k . En el caso de un electrón en una red periódica, la parte central su comportamiento es como la de un electrón libre. A medida que k aumenta el electrón se mueve hacia niveles superiores dentro de la banda y se va alejando del comportamiento anterior. El electrón en el punto 1 comienza a sentir la influencia de las fuerzas internas, se sigue moviendo en la dirección del campo aplicado, pero su velocidad ha llegado a un máximo y comienza a declinar, hasta llegar al punto 2, donde las fuerzas internas obligan a electrón a cambiar de dirección completamente, sufre una reflexión. Aquí su velocidad se anula. Este proceso se grafica en la Figura III-17. Al cambiar de dirección el vector k cambia de signo y de π/a , aparece en $-\pi/a$ (punto 4). Aquí el electrón es repelido, la fuerza interna supera a la generada por el campo, el electrón se mueve en contra del campo externo y alcanza su máximo en el punto 3. En este punto el campo externo comienza a influenciar al electrón y este gradualmente comienza a moverse en la dirección del campo externo, disminuye su velocidad, hasta llegar al origen y allí comienza nuevamente el proceso descrito.

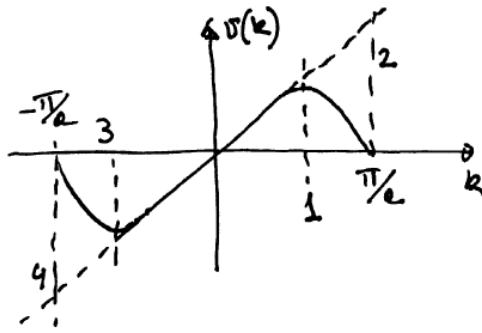


Figura III-17. Gráfico de la velocidad para un electrón libre (rayas espaciadas) y en una red cristalina (curva sólida).

Se debe ver ahora cual será el comportamiento de la masa efectiva de acuerdo a la ecuación (III-42). Se ve que en la parte central, entre las referencias 1 y 3 de la Figura III-16, el comportamiento es como si fuera un electrón libre, esto significa que $m^* = m_e$. Cuando nos acercamos a 1 (3) llegamos a un punto de inflexión de la curva, aquí la $d^2E/dk^2=0$, por lo que el valor de la masa se incrementa tendiendo a infinito, en ese punto se transita de una concavidad positiva a la izquierda de 1, a una negativa a la derecha de esta referencia.

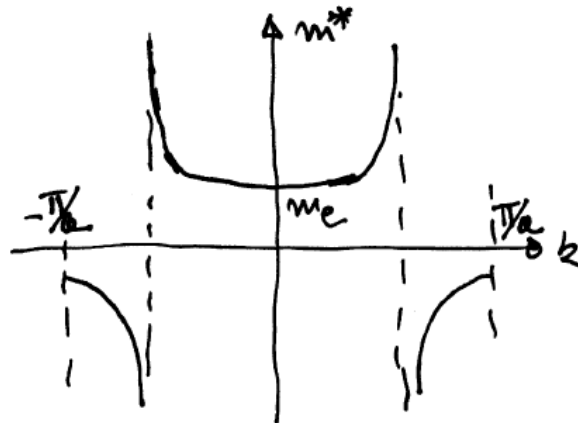


Figura III-18. Comportamiento de la masa efectiva del electrón en un cristal.

Esto indica que la masa efectiva pasa de un valor extremadamente grande positivo a uno tendiente a infinito pero negativo, habiendo superado la referencia 1 la curvatura es negativa y finita, lo cual implicaría una masa efectiva negativa y finita. Un proceso idéntico es el que se registra entre las referencias 4 y 3, como se puede ver en las Figuras III-16 y III-18.

Para finalizar, y fijar conceptos se debe hacer un breve resumen de lo visto hasta este momento, lo cual es muy importante para entender el comportamiento de los semiconductores y metales también. El modelo del electrón libre mostró cómo se comportan los electrones cuando el potencial que actúa sobre ellos es muy pequeño, se vio que ellos se mueven en un continuo de energías, esto se puede asimilar al movimiento de portadores en metales. Contrario es lo que pasa cuando los electrones se mueven en una red periódica. No solamente aparece una masa efectiva, la cual no es constante, sino que además, puede tomar valores negativos. Adicionalmente se originan bandas de energía, las cuales alternan las permitidas con las prohibidas. Este es un concepto muy importante, el ancho de la banda prohibida se determina por la diferencia de energía entre la última banda ocupada (donde hay electrones),

banda de valencia, y la primera vacía (banda de conducción), en un semiconductor esta cantidad define completamente sus propiedades.

III.5. El concepto de hueco

Pensar en una masa como una entidad negativa, es un concepto extraño. En el marco de la mecánica clásica, la energía total de una partícula se define como:

$$E = E_p + E_c$$

la suma de dos términos, una energía potencial, ligada a la posición, la cual puede ser positiva o negativa, dependiendo del sistema de referencia que se utilice. La otra, la energía cinética, relacionada al movimiento. Si no hay movimiento esta tendrá valor nulo, si hay movimiento tendrá un valor diferente de cero y positivo. Si pensamos clásicamente, la energía cinética es:

$$E_c = \frac{1}{2}mv^2$$

la velocidad, si bien es un vector y podría tomar un signo u otro. Como aparece en módulo al cuadrado, será una cantidad siempre positiva, por lo que la masa también debe ser siempre positiva. ¿Qué significa tener una masa efectiva negativa?

Trataremos de dar una explicación cualitativa, en la Figura III-9-b se ve que cuando un electrón es excitado y salta de su nivel a uno superior (de mayor energía, con esto se quiere significar estados excitados) deja su lugar vacío, se podría referir a él como un hueco en la estructura electrónica, es muy probable que este lugar se llene con otro electrón de menor energía que también en este lugar estará excitado, y así siguiendo. Se ve que a medida que los electrones alcanzan niveles de mayor energía, los huecos, si pensamos en ellos como partículas, también deberían moverse hacia lugares de mayor energía (estados excitados). Esto quiere decir que la mayor energía de los huecos se encuentra en sentido opuesto a la de los electrones. Se puede deducir que en estado fundamental, los huecos ocuparan la parte superior de una banda e irán avanzando hacia niveles de mayor energía a medida que se los excite hacia el fondo de la banda, ocupando niveles que anteriormente ocupaban los electrones, entonces la densidad de estados para estas “partículas” tiene un origen diferente al de los electrones, su origen será el tope de la banda E_{max} . Por lo que tendríamos dos sistemas de referencia para la energía, uno para electrones y otro para huecos, como se muestra en la Figura III-19 y III-20.

Podemos considerar ahora la gráfica de la densidad de estados. Hemos visto que la densidad de estados para los electrones depende de la \sqrt{E} , para los huecos dependerá, debido al cambio del origen de coordenadas para estos, de la $\sqrt{E_{max} - E}$, como se ve en la fig. III-20.

Se debe rever ahora el sentido de la masa negativa, se ha visto que huecos y electrones tiene puntos de referencias opuestos para el origen de las energías. El gráfico de la masa del

electrón de la Figura III-17 fue hecho pensando en los electrones, si pensamos desde el punto de vista de los huecos.

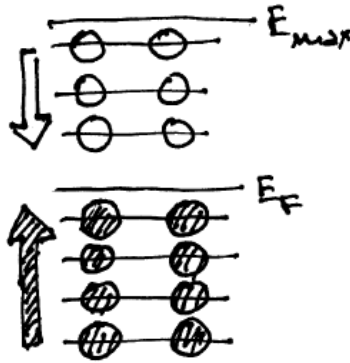


Figura III-19. Las flechas indican el sentido hacia estados excitados (de mayor energía) para los electrones y huecos.

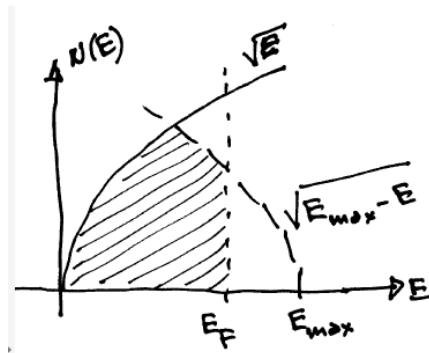


Figura III-20. Densidad de estados para electrones y huecos.

Ese gráfico debería ser visto como su imagen especular, como se muestra en la Figura III-21. La parte superior de la imagen nos muestra la variación de la masa con k para los electrones como la vimos en la Figura III-18. La parte inferior, por debajo de la línea espejo, es ahora lo interesante. Hemos visto que el sistema de referencia para la energía de electrones y huecos tiene sus orígenes opuestos, por lo que podemos considerar que si vemos la variación de masas para los huecos en función de k , esta será como la que se muestra en la parte inferior de la Figura III-21. Podemos ver aquí que lo que era una “masa negativa” en el gráfico de masas electrónicas, es ahora positiva considerando que su sistema de referencia en energías es diferente. Esto quiere decir que a la partícula “hueco” se le debe asignar una masa positiva que variará (al igual que la de los electrones) dependiendo del lugar en el eje k que ocupe. Es de hacer notar que en el eje de masas para huecos, los electrones tiene una masa negativa.

Se debe considerar una segunda cuestión, si se aplica un campo eléctrico los electrones de valencia se moverán en sentido opuesto al campo. Con el campo aplicado, cada electrón se

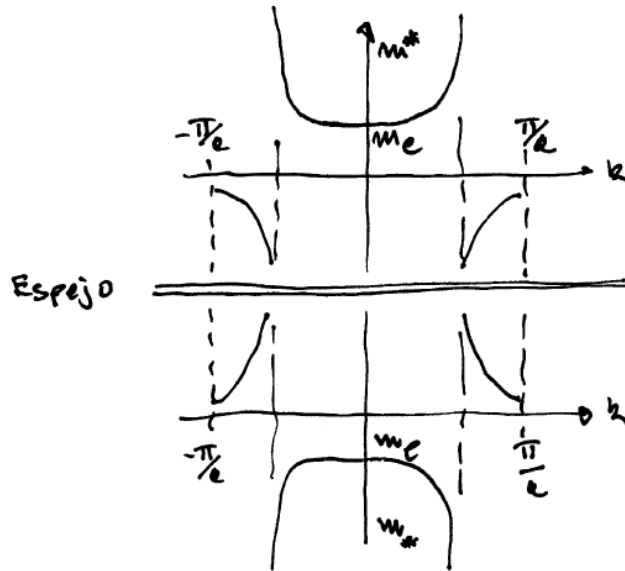


Figura III-21. Imagen especular de la variación de la masa con k .

moverá generando una corriente, que podemos definir como:

$$i = -q \sum_{j=oc.} v_j \tag{III-45}$$

Donde q es la carga elemental, y v_j es la velocidad de cada electrón en la banda. Si la banda está completamente llena (Figura III-22-a, **Banda 1**), habrá tantos electrones en la dirección $+k$ con velocidad v_j como en la dirección $-k$ con velocidad $-v_j$, por lo que $i = 0$. Si la banda está completamente vacía, tampoco habrá una corriente. Si se tiene una banda donde algunos lugares están vacíos, es decir, la banda estará poblada por electrones y huecos, en este caso, estado fundamental, no hay ninguna excitación, $i = 0$, el sistema es simétrico, como en el caso anterior, Figura III-22-a, **Banda 2 y 3**.

Cuando se aplica un campo eléctrico, la simetría existente se rompe, excepto en la **Banda 1**, completamente llena, lo que implica que no habrá movimiento de electrones. En las **Bandas 2 y 3**, de la Figura III-22-b, los electrones irán ocupando los lugares vacíos (huecos) moviéndose en sentido opuesto al campo y haciendo que los huecos se muevan en el sentido del campo aplicado, por lo que para las **Bandas 2 y 3** tendremos que la corriente es:

$$i_2 = -q \sum_{j=oc.} v_j \quad y$$

$$i_3 = -q \sum_{j=oc.} v_j$$

Siempre la suma es sobre los estados ocupados. Para hacer una demostración sencilla, se puede suponer que tenemos un único estado vacío, el l , luego se puede extender y generalizar.

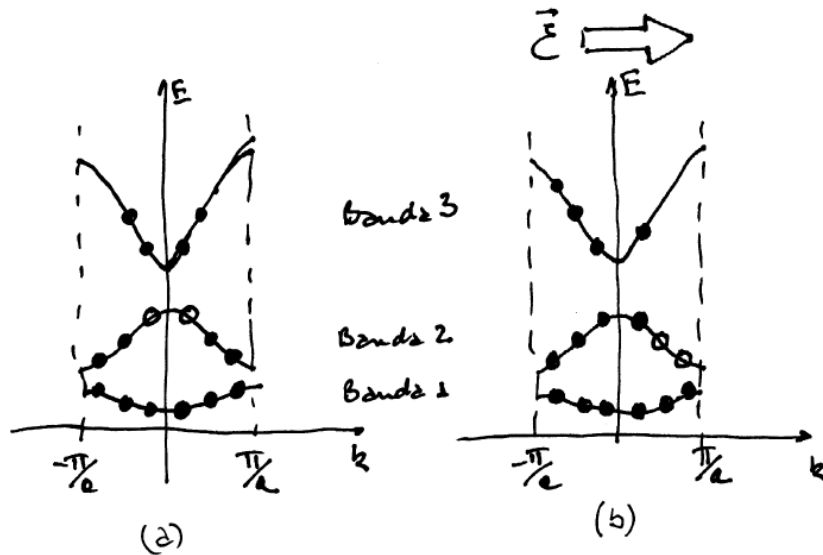


Figura III-22. (a) Sistema simétrico, (b) con campo aplicado.
Electrones: círculos llenos, huecos: círculos vacíos.

La corriente para la **Banda 2**, será:

$$i_2 = -q \sum_{j \neq l} v_j, \quad \text{se puede reescribir}$$

$$i_2 = -q \sum_{j \neq l} v_j + qv_l - qv_l = -q \sum_j v_j + qv_l$$

Al incluir el término l en la suma, se está sumando sobre todos los vectores k , todos los electrones en la banda, donde existen igual cantidad de valores positivos como negativos, por lo que la suma se anula, solo queda que la corriente es:

$$i_2 = +qv_l \tag{III-46}$$

Esto quiere decir que la corriente en una banda (en este caso la 2) transportada por todos los electrones, se puede expresar como la corriente llevada por un hueco, al cual le asignamos una carga positiva, del mismo valor que la que tiene los electrones. Si esto se generaliza, se puede escribir:

$$i_2 = +q \sum_{j=vac.} v_j \tag{III-47}$$

A su vez esto se puede generalizar a todas las bandas.

Se pueden sacar un par de conclusiones de lo visto en este capítulo que nos servirá de aquí en más. Primera, se ha visto que a los huecos se les puede asignar una masa efectiva positiva, en los lugares donde la masa efectiva de los electrones aparecía como negativa, en la parte superior de la banda. Como los huecos son vacancias electrónicas en la banda de valencia, para estos, se debería medir la energía en esta banda, en sentido positivo hacia abajo. Segunda, se les puede asignar a los huecos una carga eléctrica positiva, $+q$, con el mismo valor absoluto que tiene para un electrón, es decir $+e$.

Referencias

- Alonso, M., Finn, E. J. *Física*. 1992 Addison-Wesley Iberoamericana S.A.
- Alonso, M., Finn, E. J. *Fundamental University Physics. Vol. III*. 1968 Addison –Wesley Co.
- Beiser, A. *Conceptos de Física Moderna*. 1965 McGraw Hill Book Company.
- Berkeley Physics Course, *Física Cuántica*.– Vol. 4. 1972 Ed. Reverte S.A.
- Das, A. & Melissinos, A. C. *Quantum Mechanics. A modern introduction*. . 1988 Gordon and Breach Science Publishers.
- Hlawiczka, P. *Introducción a la electrónica cuántica*. 1977 Editorial Reverté S.A.
- Hofmann, P. *Solid State Physics. An Introduction*. 2008 Wiley-VCH Verlag GmbH & Co
- Merzbacher, E. *Quantum Mechanics. Second Edition*. 1970 J. Wiley and Sons.
- Messiah, A. *Mecánica Cuántica*.. 1962 Editorial Tecnos S.A.
- Schiff, L. I. *Quantum Mechanics. Third Edition*. 1968 McGraw Hill Kogakusha, LTD.
- Singh, J. *Quantum Mechanics. Fundamentals & Applications to Technology*. 1996 John Wiley & Sons, Inc. A Wiley-Interscience Publication.
- Slater, J. C. *Modern Physics*. 1955 McGraw Hill Book Company. Kogakusha, LTD
- Yepifanov, G. *Physical Principles of Microelectronics*. 1974 Mir Publishers. Moscow.

CAPÍTULO IV

Breves nociones de mecánica estadística

“La estadística es la primera de las ciencias inexactas”

EDMOND GOUNCOURT (1822-1896)

Novelista, naturalista Francés

Hasta ahora se han visto un conjunto de teorías en las cuales se describe el movimiento de un electrón en diferentes escenarios, con pozos de potencial, o barreras, afectado por campos aplicados o sin estos, pero siempre, un solo electrón. Cuando se comenzó con el estudio de la mecánica cuántica se hubo de abandonar la creencia de la ocurrencia determinística de un fenómeno y se debió adoptar el hecho probabilístico que la rige, eso obligó a un cambio en la estructura del pensamiento de aquella época. Ahora se propone un nuevo cambio, el hecho de aceptar el panorama probabilístico del comportamiento de las partículas implica que debemos pensar en un tratamiento diferente para los sistemas de partículas. La mecánica estadística no trata bajo ningún aspecto con el comportamiento individual de ellas sino con su aspecto grupal, mostrando cual es el comportamiento más probable de un conjunto de partículas. Si bien es posible, por ejemplo, ver la evolución de un electrón solucionando la ec. de Schrödinger, sería imposible hacerlo cuando estamos tratando con un gas de electrones compuesto por el número de Avogadro, $6,02 \times 10^{23}$ electrones. Esto debería llevar el pensamiento hacia otro panorama, y es el hecho de que se ve que cada uno de los elementos del conjunto se moverá en forma particular pero el conjunto seguramente lo hará de manera diferente. Lo que se está diciendo es que el comportamiento microscópico es el que dará lugar a un comportamiento macroscópico que es sobre el cual se puede realizar una medida. Si confeccionamos un simple circuito y ponemos un conductor, una batería y una resistencia, habrá una corriente, sobre la resistencia se generará una tensión. Obviamente, si ponemos un medidor, podremos medir una corriente. Es el efecto macroscópico, pero eso no quiere decir que un electrón que sale de un borne de la batería es el mismo que llega al otro borne habiendo pasado por el medidor. Lo que ha ocurrido es que a través de innumerables interacciones con otros electrones estos han ido moviéndose para establecer una corriente, este es el efecto macroscópico. La corriente medida es el comportamiento más probable que el conjunto de los comportamientos microscópicos podía tener.

En física podemos distinguir tres tipos de estadísticas ligadas a tres tipos de partículas diferentes. Cuando: podemos distinguir una partícula de otra, están suficientemente separadas, las podemos rotular, pueden tener cualquier spin y se pueden seguir por el espacio, ese conjunto es un conjunto clásico. Las moléculas de un gas son de este tipo. Estas obedecen a la estadística clásica de Maxwell-Boltzmann.

Por otro lado, existen partículas que tienen spin semientero ($\frac{1}{2}$). Estas no pueden distinguirse una de otra y obedecen al principio de exclusión de Pauli (se describirá en el punto IV.2.). Estas partículas son llamadas fermiones y actúan de acuerdo a la estadística de Fermi-Dirac.

Por último, se tienen partículas que cumplen las mismas condiciones que las anteriores pero con una única diferencia frente a estas: tienen spin entero, estas son llamadas bosones, pueden ser tratadas mediante la estadística de Bose-Einstein. Las partículas estudiadas por esta estadística son aquellas ligadas a núcleos atómicos y a partículas fundamentales como fotones, gluones, mesones, etc. Por exceder los alcances de este libro no nos dedicaremos a esta estadística.

IV.1. Estadística clásica, Maxwell-Boltzmann



James Clerk Maxwell
13-6-1831, 5-11-1879



Ludwig Eduard Boltzmann
20-2-1844, 5-9-1906

Se puede pensar en un conjunto de P partículas que se mueven a diferentes velocidades y con diferentes energías, se pueden diferenciar a estos y llamarlos p_1 al conjunto de partículas con energía E_1 , p_2 al que tiene energía E_2 , etc, donde la energía E_1 es menor que la E_2 , y la E_2 menor que la E_3 , y así siguiendo. Se puede empezar a establecer algunas relaciones fundamentales. Primero el número de partículas es constante, esto es: si una partícula pasa del conjunto p_k al p_{k+1} , habrán variado las cantidades p_k y p_{k+1} pero no P .

$$\sum_i p_i = p_1 + p_2 + \dots + p_n = P \tag{IV-1}$$

La conservación del número de partículas está indicando la conservación de la masa, uno de los principios fundamentales de la Física. El otro principio de conservación es el de la energía interna del sistema,

$$\sum_i p_i E_i = p_1 E_1 + p_2 E_2 + \dots + p_n E_n = E \quad (IV-2)$$

Si se realizan variaciones sobre la cantidad de partículas en las ecuaciones (IV-1/2), se tiene:

$$\sum_i \delta p_i = \delta p_1 + \delta p_2 + \dots + \delta p_n = 0 \quad (IV-3)$$

Si hay una variación del número de partículas en el conjunto i y varía en δp_i , la variación de energía interna será $E_i \delta p_i$, por tanto:

$$\sum_i E_i \delta p_i = E_1 \delta p_1 + E_2 \delta p_2 + \dots + E_n \delta p_n = 0 \quad (IV-4)$$

Se deben introducir multiplicadores de Lagrange como cantidades independientes de p_i , multipliquemos (IV-3) por $-\alpha$ y (IV-4) por $-\beta$ (Ver apéndice IV-1 para el completo desarrollo de este punto), se tiene:

$$\sum_i [-\ln p_i + \ln f_i - \alpha - \beta E_i] \delta p_i = 0 \quad (IV-5)$$

donde f_i es la probabilidad a priori, para cualquier valor de i . El valor entre corchetes debe ser idénticamente nulo, por lo que:

$$-\ln p_i + \ln f_i - \alpha - \beta E_i = 0 \quad (IV-6)$$

Despejando p_i , se puede escribir:

$$p_i = f_i e^{-\alpha} e^{-\beta E_i} = \frac{f_i}{e^{\alpha} e^{\beta E_i}} \quad (IV-7)$$

Esta expresión es la ley de distribución de Maxwell-Boltzmann. Si no se consideran las variaciones de la energía en forma discreta, sino como un continuo, se puede escribir la expresión (IV-7) como:

$$p(E) dE = \frac{f(E)}{e^{\alpha} e^{\beta E}} dE \quad (IV-8)$$

Como se está tratando con un gas de electrones, tal como dijimos al principio, esta ecuación describe la cantidad de partículas que se mueven en el intervalo de energía entre E y $E+dE$. Se puede demostrar que

$$\frac{1}{e^{\alpha}} = \frac{Ph^3}{V} \left(\frac{\beta}{2\pi m} \right)^{3/2} \quad (IV-9)$$

Donde $\beta=1/kT$, y k la constante de Boltzmann, reemplazando se tiene:

$$f(E) = \frac{2\pi V (2m)^{3/2} \sqrt{E}}{h^3} \quad (IV-10)$$

Con lo cual se obtiene:

$$p(E)dE = \frac{2P}{\sqrt{\pi}} \beta^{3/2} \sqrt{E} e^{-\beta E} dE \quad (IV-11)$$

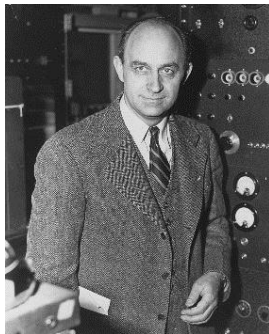
Como se dijo previamente, esta es la distribución de partículas en el intervalo E y $E+dE$.

Ejercicio sugerido 1: como se trata con un gas de electrones libres, se puede decir que $E = p^2/2m$. Partiendo de la ec. (IV-11), encontrar la distribución $p(p)dp$.

Ejercicio sugerido 2: si se supone que la energía de los electrones es solamente cinética, encontrar $p(v)dv$, donde v es la velocidad de la partícula.

Ejercicio sugerido 3: a partir de la distribución del ejercicio anterior, encontrar cual es la velocidad más probable. **Rta:** $v = \sqrt{2kT / m}$.

IV.2. Estadística cuántica, Fermi-Dirac



Enrico Fermi
29-9-1901, 28-11-1954
Premio Nobel de Física 1938



Paul Adrien Maurice Dirac
8-8-1902, 20-10-1984
Premio Nobel de Física 1933

En este caso se debe pensar en partículas las cuales no se pueden distinguir una de otra, y deben obedecer el principio de Exclusión de Pauli.

Brevemente puede ser descrito en forma cualitativa, como sigue: cuánticamente una partícula está descrita por cuatro números, que se llaman números cuánticos, no se entrará en detalle para explicar este punto, serían algo así como sus coordenadas cuánticas, la función de onda de un electrón se debería escribir:

$$\Psi_{n,l,m_l,m_s}(r) \quad (IV-12)$$

donde, n , l y m_l dan una descripción del movimiento orbital del electrón, n es el *número cuántico total*, que describe la cuantización de la energía, l es el *número cuántico orbital*, y hace

referencia a la cuantización del momento angular. El *número cuántico magnético* es m_l , que corresponde a la relación entre el momento magnético generado por el electrón y la dirección del campo magnético externo. Estos tres valores para cada electrón son únicos, no se pueden repetir entre electrones de un mismo nivel. Finalmente el *número cuántico magnético de spin*, m_s , es la descripción de la cuantización espacial del spin del electrón, puede tomar solamente dos valores $m_s=+\frac{1}{2}$ y $m_s=-\frac{1}{2}$. Esto indica que sin violar ninguna regla cuántica en un nivel electrónico, podemos tener dos estados cuánticos, dos electrones, uno con $(n, l, m_l, +\frac{1}{2})$ y otro con $(n, l, m_l, -\frac{1}{2})$. Esto es lo mismo que decir que dos electrones en un mismo nivel no pueden tener las mismas coordenadas cuánticas, la única diferencia que pueden tener, para no violar ningún principio, es que uno debe estar con su spin “para arriba”, mientras que el otro lo debe tener “para abajo”, popularmente hablando.

Wolfgang E. Pauli
25-4-1900, 15-12-1958
Premio Nobel de Física 1945



En la estadística de Maxwell-Boltzmann cada nivel puede contener un número indefinido de partículas, aquí como se ha visto, eso no ocurre, por estado cuántico podemos tener una partícula. Se puede suponer un sistema que puede albergar f_i estados cuánticos con la misma energía, y se tienen p_i partículas, por lo tanto habrá p_i estados cuánticos ocupados con un partícula, los restantes estados, hasta llegar al número f_i estarán vacíos. Debido al hecho de que las partículas son idénticas, si permuto una partícula por otra, el sistema no cambiará, por lo tanto, se va a tener $p_i!$ permutaciones que no cambian al sistema. Lo mismo ocurre con los estados no ocupados, su permutación no cambia el estado del sistema, es decir existen $(f_i - p_i)!$ permutaciones que no tienen significación. Esto quiere decir que las permutaciones que si modifican al sistema son:

$$P_i = \frac{f_i!}{p_i!(f_i - p_i)!} \tag{IV-13}$$

Para determinar la distribución final de las partículas se debe realizar el producto de estas ocupaciones para cada una de las partículas, es decir:

$$P = \prod_i \frac{f_i!}{p_i!(f_i - p_i)!} \tag{IV-14}$$

Esta expresión nos dice cuál va a ser el resultado final de la distribución del conjunto completo de partículas y estados desocupados, pero debemos saber cuál será la distribución de las partículas p_i , para lo cual, tomaremos logaritmos a ambos lados de la ecuación (IV-14)

$$\ln P = \sum_i [\ln f_i! - \ln p_i! - \ln(f_i - p_i)!]$$

Utilizando la expresión de Stirling (apéndice IV-1), tenemos:

$$\ln P = \sum_i [f_i \ln f_i - p_i \ln p_i - (f_i - p_i) \ln(f_i - p_i)] \quad (\text{IV-15})$$

Se debe encontrar cual es la distribución más probable de todas las posibles, eso se puede hallar variando el $\ln P$, esa distribución se logra cuando (IV-15) alcanza su máximo:

$$\delta \ln P = \sum_i [-\ln p_i + \ln(f_i - p_i)] \delta p_i = 0 \quad (\text{IV-16})$$

Teniendo en cuenta las leyes de conservación de masa, número de partículas, y de energía, junto con los multiplicadores de Lagrange, se tiene:

$$-\alpha \sum_i \delta p_i = 0$$

y

$$-\beta \sum_i E_i \delta p_i = 0$$

La ecuación (IV-16) queda:

$$\sum_i [-\ln p_i + \ln(f_i - p_i) - \alpha - \beta E_i] \delta p_i = 0 \quad (\text{IV-17})$$

La variación sobre el número de partículas es una cantidad independiente, entonces lo que debe idénticamente anularse es la cantidad entre corchetes para cada valor i .

$$-\ln p_i + \ln(f_i - p_i) - \alpha - \beta E_i = 0$$

operando,

$$\ln \left[\frac{f_i - p_i}{p_i} \right] = \alpha + \beta E_i \Rightarrow \frac{f_i - p_i}{p_i} - 1 = e^{\alpha + \beta E_i} = e^{\alpha} e^{\beta E_i}$$

De aquí, recordando que $\beta = 1/kT$

$$p_i = \frac{f_i}{e^{\alpha} e^{E_i/kT} + 1} \quad (\text{IV-18})$$

Se obtiene la ley de distribución de Fermi-Dirac. Se puede ver en la (IV-18) que por el solo hecho de poner un único electrón por estado, el denominador cambia con respecto a la estadística clásica (IV-7) donde se describen sistemas con muchísimos electrones participando en cada estado.

Como se ha dicho que los niveles en una banda se encuentran muy próximos entre si, nos referiremos ahora al sistema como si fuera un continuo cambiando E_i por E . Reescribimos la ec.(IV-18) de la siguiente manera:

$$p(E) = \frac{f(E)}{e^\alpha e^{E/kT} + 1} \quad (\text{IV-19})$$

Que nos dice como se distribuyen los electrones en un sistema cuántico.

IV.3. Diferencias entre estadísticas y ocupación de estados

Una de las cosas que se puede hacer es intentar comparar el comportamiento de las dos estadísticas que hemos visto, cuando se varía la energía. Si bien, las teorías que hemos descripto parten de fundamentos diferentes, una es la clásica y la otra la cuántica, si pueden compararse sus resultados cuando se las normaliza, p refiere al número de partículas bajo estudio, mientras que f , hace referencia a número de estados, por lo tanto, la relación p/f dice que al dividir la cantidad de partículas por el número de estados se está obteniendo la ocupación del estado i . Esto es de las ecuaciones (IV-8) y (IV-19):

$$s(E) = \frac{p}{f} = \frac{1}{e^\alpha e^{E/kT} + 1} \quad (\text{IV-20})$$

Para la estadística de Maxwell-Boltzmann, mientras que para la de Fermi-Dirac, es:

$$s(E) = \frac{p}{f} = \frac{1}{e^\alpha e^{E/kT} + 1} \quad (\text{IV-21})$$

Como el valor e^α no depende de E , actúa como un factor constante y no modificará la tendencia con la variación de energía, reescribiremos la expresión (IV-20) de la siguiente manera:

$$s(E) \approx \frac{1}{e^{E/kT}} \quad (\text{IV-22})$$

y la ecuación (IV-21) como sigue:

$$s(E) \approx \frac{1}{e^{E/kT} + 1} \quad (\text{IV-23})$$

Se puede hacer un análisis rápido de ambas ecuaciones. En la (IV-22), cuando E es un valor pequeño frente a kT ($E \ll kT$) el exponencial toma valores muy pequeños y el cociente crece, tendiendo a valores muy grandes, son muchas las partículas que se pueden encontrar para valores de energía pequeños. Cuando la energía aumenta, el cociente decrece y va tendiendo a cero monótonamente. Esto significa que hay pocas partículas que tengan energías elevadas.

Analicemos la ecuación (IV-23). Cuando $E \ll kT$, el valor del exponencial es muy pequeño, mucho menor que 1, por lo que en el denominador de esta ecuación el exponencial es despreciable frente al 1, por esta razón, el valor de $s(E) = 1$. Cuando la energía crece el exponencial toma valores mucho más grandes que 1, el 1 no tiene significación en el denominador, el valor que tome el denominador serán los valores que vaya tomando el exponencial, por lo tanto el comportamiento será como el de la ecuación (IV-22),

$$s(E) \approx \frac{1}{e^{E/kT}} \tag{IV-24}$$

Esto quiere decir que para pequeñas energías la ocupación del sistema será 1 como se ha predicho, mientras que a grandes energías los electrones se comportan bajo la estadística de Maxwell-Boltzmann, como un gas de electrones. La descripción gráfica de este comportamiento se ve en la Figura (IV-1). Se ha mencionado pequeñas energías o grandes energías, y no se ha dicho cuál es el punto de referencia para establecer la comparación que se ha mencionado, la referencia es E_α , es un parámetro que se debe determinar.

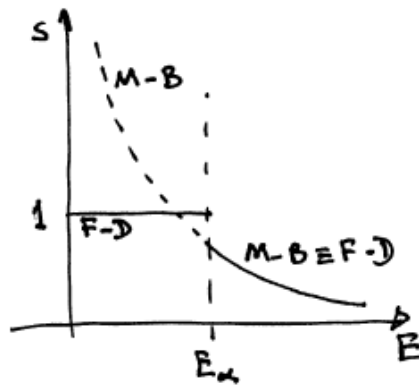


Figura IV-1. Comparación de las curvas de Maxwell-Boltzmann y Fermi-Dirac.

Un concepto muy importante que se debe utilizar ahora es el que se introdujo en el capítulo III sobre la energía de Fermi. Gráficamente se representó, por ejemplo en la Figura III-21, que a temperatura de 0K, todos los electrones se encuentran por debajo de la energía de Fermi. No se entrará en mayores detalles, pero mediante argumentos termodinámicos y utilizando la mecánica estadística de sistemas reales, se puede llegar al siguiente resultado:

$$\alpha = -\frac{E_F}{kT} \tag{IV-25}$$

La cual lleva a que la expresión (IV-21) tome la siguiente forma:

$$s(E) = \frac{1}{e^{\frac{(E-E_F)}{kT}} + 1} \tag{IV-26}$$

Se puede rehacer el análisis, semejante al realizado más arriba. Supongamos que $T = 0K$, viendo que si, $E < E_F$, en el denominador el exponencial desaparece por lo tanto

$$s(E) = 1 \tag{IV-27}$$

Por otro lado, si $E > E_F$, en el denominador el exponencial es dominante

$$s(E) = 0 \tag{IV-28}$$

Podemos ver en la Figura (IV-2) este comportamiento.

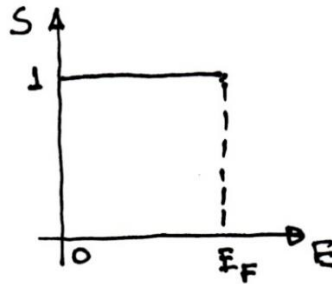


Figura IV-2. Distribución de estados ocupados a $T=0K$.

¿Qué nos indica esta Figura, a $T=0K$? Todos los estados por debajo del nivel de Fermi se encuentran ocupados, mientras que cuando se supera esa energía, todos los estados se encuentran vacíos. Esto es lo que se ha indicado esquemáticamente en las Figuras III-9-a y III-21. ¿Qué sucede cuando $E = E_F$? Si se introduce esto en la ecuación (IV-26) se verá que el exponente del exponencial se anula, $E - E_F = 0$, por tanto:

$$e^{\frac{E-E_F}{kT}} \Big|_{E=E_F} = 1 \Rightarrow s(E) = \frac{1}{2} \tag{IV-29}$$

Esto significa, que el nivel puede estar ocupado o no, la probabilidad es $\frac{1}{2}$. Se debe ver que es lo que pasa cuando $T \neq 0$. En este caso, también cuando $E = E_F$, $s(E) = \frac{1}{2}$. Bajo estas condiciones ocurre lo que se muestra en la Figura III-9-b, los electrones adquieren más energía y algunos pueden superar el nivel de Fermi, con lo cual niveles próximos inferiores al de Fermi comienzan a despoblarse, mientras que algunos superiores se han poblado, como muestra la Figura IV-3.

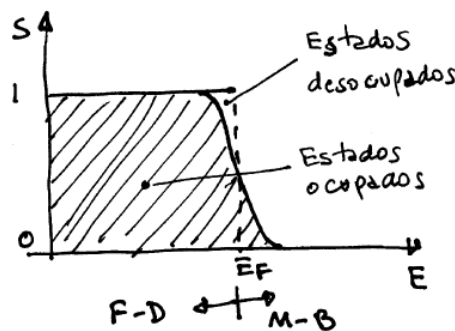


Figura IV-3. Distribución de estados ocupados a $T \neq 0K$

Debe quedar claro que cuando un electrón es excitado y abandona un nivel de energía por debajo de la energía de Fermi, pasa a ocupar un nivel por encima del de Fermi. Si obtuviéramos las áreas de estados desocupados mostrados en la Figura IV-3, y la de los estados ocupados para energías mayores que la de Fermi, se vería que son exactamente iguales. Esto es debido al hecho que cada electrón que abandona un nivel de energía inferior a E_F deja un lugar vacío y pasa a ocupar otro estado de energía superior a E_F .

Ejercicio sugerido: se puede suponer que la energía se desplaza una cantidad δE con respecto a E_F en ambos sentidos, ¿cuál será el nivel de ocupación?

Retomemos la ecuación (IV-19), en términos continuos y habiendo reemplazado el valor de α , quedará:

$$p(E) = \frac{f(E)}{e^{(E-E_F)/kT} + 1} \tag{IV-30}$$

Reemplazando $f(E)$ (ec.(IV-10))

$$p(E) = \frac{2\pi V(2m)^{3/2}}{h^3} \frac{\sqrt{E}}{e^{(E-E_F)/kT} + 1} \tag{IV-31}$$

En la Figura IV-4 se puede ver la gráfica de la ecuación (IV-31) donde se ha oscurecido el área ocupada.

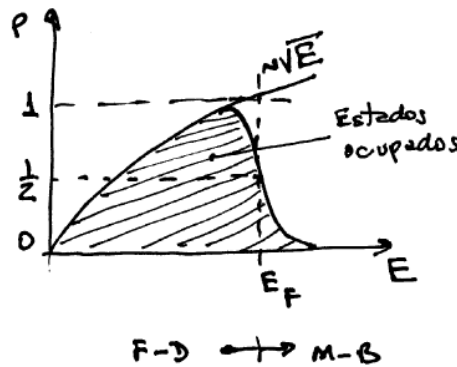


Figura IV-4. Gráfica de $p(E)$ en función de E .

Para finalizar se debe obtener la energía de Fermi. Para ello se debe recordar que a $T=0K$ todos los electrones se encuentran por debajo del nivel de Fermi que es la máxima energía que ellos pueden alcanzar a esta temperatura. Por lo tanto,

$$F = \int_0^{E_F} f(E)dE = \frac{2\pi V(2m)^{3/2}}{h^3} \int_0^{E_F} E^{1/2}dE = \frac{4\pi V}{3h^3} (2m)^{3/2} E_F^{3/2} \tag{IV-32}$$

Como cada estado puede albergar a dos electrones, con spines opuestos, por lo que la (IV-32) debería multiplicarse por 2, quedando:

$$F = \frac{8\pi V}{3h^3} (2m)^{3/2} E_F^{3/2} \quad (\text{IV-33})$$

De aquí podemos despejar el valor de E_F :

$$E_F = \left[\frac{F}{V} \frac{3}{8\pi} \right]^{2/3} \frac{h^2}{2m} = \frac{h^2}{2m} \left[\frac{3n}{8\pi} \right]^{2/3} \quad (\text{IV-34})$$

En donde F/V es el número de electrones por unidad de volumen, que se denominará n .

Ejercicio sugerido: Encontrar el nivel de Fermi para el litio (Li).

Rta: 4,72eV

Referencias

- Alonso, M., Finn, E. J. *Física*. 1992 Addison-Wesley Iberoamericana S.A.
- Alonso, M., Finn, E. J. *Fundamental University Physics. Vol. III*. 1968 Addison –Wesley Company.
- Beiser, A. *Conceptos de Física Moderna*. 1965 McGraw Hill Book Company.
- Dalven, R. *Introduction to Applied Solid State Physics* 1990 Plenum Press
- Hlawiczka, P. *Introducción a la electrónica cuántica*. 1977 Editorial Reverté S.A
- Pierret, F. R. *Advanced Semiconductor Fundamentals, Volume VI*. 1987 Modular Series on Solid State Devices, R.F Pierret and G. W. Neudeck, Editors.1987 Addison-Wesley Pub. Comp.
- Sears, F. W. *Termodinámica*. 1959 Ed. Reverté, S.A.
- Singh, J., *Quantum Mechanics. Fundamentals & Applications to Technology*. 1996 John Wiley & Sons, Inc. A Wiley-Interscience Publication.
- Slater, J. C. *Modern Physics*. 1955 McGraw Hill Book Company. Kogakusha Co. LTD
- Tipler, P. A. *Física Moderna*. 2009 Ed. Reverté, S.A.
- Yepifanov, G. *Physical Principles of Microelectronics*. 1974 Mir Publishers. Moscow.

Apéndice IV-1

Estadística de Maxwell-Boltzmann. Probabilidad

Clásicamente, se pueden tratar a los electrones como un gas. Donde habrá conjuntos de partículas que poseen diferentes energías unos de otros, pero dentro de cada conjunto las partículas poseen la misma energía. Para tratar este problema, se puede hacer una analogía, pensando que se tienen esferitas de iguales radios. No obstante ser las esferitas iguales, como la descripción se realiza dentro de una estadística clásica, se puede nominar cada una de las esferitas. Además se tiene una gran caja de superficie S , subdividida en diferentes pequeñas cajitas de menor superficie, s_i . El conjunto de cajitas cumple con:

$$S = \sum_i s_i \quad (\text{A-1})$$

Se arrojan las esferitas sobre la caja y éstas empezarán a caer en las diferentes cajitas. Se repite el proceso una gran cantidad de veces. En cada tiro habrá una distribución de esferitas en cajitas que puede ser igual o diferente a las anteriores, con lo cual se puede realizar una estadística de los eventos y se podrá determinar, ¿cuál de las diferentes distribuciones es la de mayor probabilidad?

Se debe establecer que existen dos probabilidades a considerar, una va a depender de ¿en cuántas formas diferentes se distribuirán las esferitas en las cajitas?, y la otra estará relacionada a la conformación propia de la cajita. La primera se llama la “probabilidad termodinámica” y en forma más extensa, nos habla de la cantidad de maneras diferentes en que se pueden distribuir las esferitas en las cajitas sin modificar la cantidad de esferitas en cada una de ellas. La segunda se denominará “probabilidad a priori”.

La probabilidad de que un evento ocurra es la relación entre el número probable sobre el número total de casos. Si tiramos un dado una sola vez, la probabilidad de que obtengamos el número tres es $1/6$. Como aquí se está trabajando con áreas, la probabilidad de que una esfera caiga en una cajita será la relación entre el área de la cajita y el área total.

$$f_i = \frac{s_i}{S} \quad (\text{A-2})$$

Esta es la que se ha llamado la “probabilidad a priori”. La suma de todas ellas será:

$$\sum_i f_i = 1 \quad (\text{A-3})$$

esto es debido a que la esferita debe estar en alguna de las cajitas dentro de la caja. Si la probabilidad de que caiga una esfera dentro de una cajita i es f_i , la probabilidad de que caigan 2 esferas dentro de la cajita es $f_i \times f_i$, es decir f_i^2 , por lo tanto la probabilidad a priori de que n_1 esferas caigan en la cajita **1**, n_2 en la cajita **2**, etc. será:

$$f_1^{n_1} \cdot f_2^{n_2} \cdot f_3^{n_3} \dots f_k^{n_k} \tag{A-4}$$

Donde la condición es la conservación de la cantidad de esferitas, estas no se crean ni se destruyen, pueden cambiar, si existiera algún mecanismo para que lo hagan, de cajita, pero su número será siempre el mismo.

$$\sum_i n_i = N \tag{A-5}$$

El número total de permutaciones que pueden hacerse con las N de esferas es: $N!$ Pero se debe considerar que si en una cajita se tienen n_i esferas, que se permuten ellas dentro de la cajita, no tiene ningún efecto ni significado, por lo que, esas $n_i!$ permutaciones no tendrán sentido. La cantidad de maneras diferentes en que las esferas pueden distribuirse en las cajitas, según se definió más arriba, es la probabilidad termodinámica, y la podemos obtener como el cociente entre el número total de permutaciones $N!$ y la cantidad total de permutaciones sin sentido, es decir:

$$\frac{N!}{n_1! n_2! \dots n_k!} = \frac{N!}{\prod_i n_i!} \tag{A-6}$$

Se debe encontrar ahora la probabilidad total de la distribución D . Esta vincula a las probabilidades de los diferentes eventos y la propia conformación de las cajitas, la primera se refiere a la probabilidad estadística y la segunda a la probabilidad a priori,

$$D = \frac{N!}{n_1! n_2! \dots n_k!} f_1^{n_1} f_2^{n_2} \dots f_k^{n_k} = \frac{N!}{\prod_i n_i!} \prod_i f_i^{n_i} \tag{A-7}$$

Como paso siguiente se debe determinar cuál es la distribución más probable de las esferas. Cuando los factoriales que se utilizan son sobre números muy grandes, se puede aproximar el valor de n por el de $\ln n$, y el $\ln n!$ se puede aproximar mediante la expresión de Stirling,

$$\ln n! = n \ln n - n \quad \text{si } n \gg 1 \tag{A-8}$$

Para determinar el valor más probable de D , se debe encontrar su máximo. Para ello se debería operar sobre la expresión (A-7), si se aplica logaritmo natural a ambos miembros, se obtiene:

$$\ln D = \ln N! - \sum_i \ln n_i! + \sum_i n_i \ln f_i \tag{A-9}$$

y si a continuación se utiliza la expresión de Stirling, se tiene:

$$\ln D = N \ln N - N - \sum_i (n_i \ln n_i - n_i) + \sum_i n_i \ln f_i \tag{A-10}$$

Se debe observar que, de acuerdo a la ecuación (A-5) el término N se anula con el término n_i dentro de la suma (3er. Término, 2do miembro), con lo cual (A-10) queda:

$$\ln D = N \ln N - \sum_i n_i \ln n_i + \sum_i n_i \ln f_i \quad (A-11)$$

Normalmente, si se deseara encontrar el máximo de una función continua g cualquiera, se debería obtener su derivada con respecto a la variable x e igualarla a cero,

$$\frac{dg(x)}{dx} = 0$$

Como el conjunto que se estudia es discreto, no se puede utilizar la derivada sino las variaciones con respecto a la variable. Esto quiere decir que si se varía n_i en δn_i , el valor de D , no debe cambiar. Si se establece esta variación en la ecuación (A-11) se podrá encontrar el valor máximo de D , el primer término de (A-11) se anula debido a que N el número total de partículas, por definición es constante, por lo tanto:

$$\delta \ln D = - \sum_i (\delta n_i \ln n_i + n_i \delta \ln n_i) + \sum_i \delta n_i \ln f_i = 0 \quad (A-12)$$

De acuerdo a la definición de variación de un logaritmo,

$$\delta \ln n_i = \frac{\delta n_i}{n_i} \quad (A-13)$$

si se reemplaza esta en (A-12) se tiene:

$$- \sum_i \delta n_i \ln n_i + \sum_i \delta n_i + \sum_i \delta n_i \ln f_i = 0 \quad (A-14)$$

El segundo término se anula, debido al hecho que el número de partículas es constante, por lo que las variaciones de las esferas en las cajitas son de tal manera que las que salen de una cajita van a otra u otras, pero la suma total de las δn_i es cero. La ecuación (A-14) queda:

$$- \sum_i \delta n_i \ln n_i + \sum_i \delta n_i \ln f_i = 0 \quad (A-15)$$

Esta da la distribución más probable de esferas. Pero ambos términos contienen la variación δn_i . Estas variaciones no son independientes, pues como se vio en la ecuación (A-5) el número total de esferas es fijo. Se deben utilizar los multiplicadores de Lagrange para poder obtener la distribución más probable. Estas constantes deben ser independientes de n_i . Sea α esa cantidad, se tiene:

$$\sum_i \alpha \delta n_i = 0 \quad (A-16)$$

Como este valor es cero, se puede sumar a la (A-15) sin que por ello sufra modificaciones, se tiene:

$$-\sum_i \delta n_i \ln n_i + \sum_i \delta n_i \ln f_i + \sum_i \alpha \delta n_i = \sum_i [-\ln n_i + \ln f_i + \alpha] \delta n_i = 0 \quad (\text{A-17})$$

Para todo i , la expresión dentro del corchete se debe anular, de allí se puede determinar el valor de n_i ,

$$-\ln n_i + \ln f_i + \alpha = 0 \Rightarrow n_i = f_i e^\alpha \quad (\text{A-18})$$

Si al resultado de (A-18) se lo extiende, realizando la suma sobre todas las cajitas y todas las probabilidades “a priori”, y viendo que α no depende de i , se tiene:

$$\sum_i n_i = e^\alpha \sum_i f_i \quad (\text{A-19})$$

Se debe tener presente que la suma de las probabilidades “a priori” es igual a 1, esto es debido al hecho que las esferitas están dentro de la caja, distribuidas entre las diferentes cajitas por lo que la probabilidad de cada cajita, dependiendo de sus características tendrá un valor menor a la unidad (ver ecuación (A-2)), pero la suma de todas ellas será 1.

$$\sum_i f_i = 1 \quad (\text{A-20})$$

De acuerdo a (A-5) se tiene:

$$N = e^\alpha \quad (\text{A-21})$$

Si se opera sobre (A-18) y se tienen en cuenta las ecuaciones (A-20) y (A-21),

$$n_i = e^\alpha f_i \rightarrow n_i = e^\alpha \frac{S_i}{S}, \quad (\text{A-22})$$

Se obtiene:

$$n_i = \frac{N}{S} S_i \quad (\text{A-23})$$

Esta es una expresión muy importante, ella establece como será la distribución de esferitas en la caja, pero además nos dice que esa distribución depende de la densidad de las partículas (N/S) que se tengan, este es un concepto nuevo, pues independiza la distribución de las características propias de las esferitas. Además establece que la distribución depende de la superficie de cada cajita. Aquí es donde se introduce el concepto de las características físicas del lugar donde se ubicarán las esferitas.

Referencias: apéndice

Alonso, M., Finn, E. J. *Fundamental University Physics. Vol. III.* 1968 Addison–Wesley Company.

Beiser, A. *Conceptos de Física Moderna.* 1965 McGraw Hill Book Company.

Sears, F. W., *Termodinámica.* 1959 Ed. Reverté, S.A.

SEGUNDA PARTE

Herramientas aplicadas

CAPÍTULO V

Descripción de un sólido

*Hay herramientas de pensamiento tan romas que no sirven casi para nada,
otras de filo tan aguzado que se vuelven peligrosas.
Pero el hombre sabio hará uso de ambas*

GREGORY BATESON. Antropólogo
9/5/1904-4/7/1980

La materia se manifiesta a nuestros sentidos mediante agregados de una enorme cantidad de átomos, presentándose en tres diferentes estados: gaseoso, líquido o sólido. Cada una de estas fases posee propiedades que diferencian a unas de otras. En lo que respecta a los gases, la distancia entre moléculas es mucho mayor que el tamaño propio de la molécula, en el caso de las fuerzas, las intermoleculares son mucho más pequeñas que las intramoleculares que son aquellas que mantienen cohesionados a los átomos de las moléculas. Por otro lado, en un gas, la molécula mantiene su individualidad. Además, los gases adoptan completamente la forma y el volumen de su contenedor.

Los sólidos se encuentran en el otro extremo de este tipo de comportamiento, los átomos o moléculas se encuentran en posiciones fijas en el espacio, las fuerzas que los mantienen en estas posiciones son semejantes a las existentes en sus ligaduras. El volumen y la forma que adoptan los sólidos son diferentes a la de su recipiente. Las variaciones que sufren bajo los efectos de la temperatura y/o la presión son poco perceptibles en comparación a la que sufren los gases. También, a diferencia de los gases, los sólidos presentan propiedades completamente diferentes a los átomos que los constituyen.

Las propiedades de los líquidos se encuentran entre estos dos mencionados, y ellas exceden el alcance de este texto, no se hará ningún desarrollo sobre ellos.

Hasta aquí se han visto los modelos que representan el movimiento de electrones en el sólido, también se ha visto el tratamiento estadístico que se le podía dar a estos para su estudio. Al principiarse el Capítulo III se han mencionado los sólidos cristalinos y amorfos dando una muy breve descripción de ellos. Debido a que nuestro interés está centrado en los semiconductores y dentro de este grupo, en los que presentan estructura cristalina, se desarrollará una breve descripción de la conformación y propiedades de los sólidos cristalinos.

Los cristales, sólidos cristalinos, se desarrollan en base a una distribución periódica en el espacio, como se mostró en el capítulo III. Para estudiarlos, para entender la estructura del sólido, solamente es necesario estudiar la “celda unidad”, esta es la porción mínima del cristal que conserva las propiedades del conjunto, la cual distribuida mediante operaciones de simetría, traslaciones y/o rotaciones en el espacio, forma el cristal en su totalidad.

Ejemplos de sólidos cristalinos son el diamante (carbono en la fase cúbica de cara centrada), el rubí (óxido de aluminio con cromo en fase trigonal), el topacio (óxido de aluminio con silicio y flúor en fase ortorrómbica), el hierro (hierro en fase cúbica centrada en el cuerpo), el cobre (átomos de cobre en fase cúbica de cara centrada), etc, la lista es casi interminable, los tres primeros son transparentes, los siguientes son opacos. Hay cristales que conforman metales, o aislantes, o semiconductores. Cada uno de ellos, de una u otra forma ha impactado en la vida, en las civilizaciones, de manera que han dejado su huella en forma indeleble.

Las propiedades mencionadas y otras, fueron lentamente determinadas con el correr del tiempo, en forma totalmente empírica durante siglos, hasta el advenimiento de la mecánica cuántica, la cual pudo escudriñar en el mundo microscópico, entre átomos y moléculas, y pudo explicar cada una de las propiedades de los materiales. La aplicación de esta teoría ha sido fundamental en las últimas décadas para apoyar y guiar el avance tecnológico hasta lo que conocemos hoy en día. Y más importante aún, la mecánica cuántica ha permitido determinar fehacientemente las diferencias existentes entre los materiales, los que son metálicos, de los que son aislantes y de los que son semiconductores, etc. además ha sido capaz de establecer la estructura energética de los diferentes materiales, los cuales son la “huella dactilar” de ellos, no hay dos materiales diferentes que tengan la misma distribución de energías.

V.1. Descripción de las estructuras cristalinas

En algún momento de la historia, cientos de millones de años atrás, los átomos se encontraban separados, las fuerzas de atracción entre ellos empezaron a trabajar y los ubicaron en posiciones próximas. Estas fuerzas se incrementan a medida que las distancias se acortan. Pero, cuando las partículas se encontraban a cortas distancias, un nuevo actor se hizo presente, las fuerzas de repulsión. A una determinada distancia ambas se balancearon, es decir la repulsión y la atracción se igualaron con direcciones opuestas. Es allí donde las partículas que constituyen el sólido (átomos o moléculas) se encuentran en equilibrio. Con el agregado de más partículas, provenientes de todas direcciones, en donde todas participan de este juego de fuerzas, se formó una estructura tridimensional, que se denominará cristal.

Si no media ninguna acción que pueda alterar las posiciones de las partículas en la estructura, estas permanecerán en sus posiciones de equilibrio. Si por algún motivo las partículas se apartan de estas posiciones, ellas ganan energía, y aparecen fuerzas que tienden a restituir las a su posición original. La única forma posible para que las partículas se muevan son vibraciones alrededor de sus posiciones de equilibrio.

Dependiendo del tipo de átomo, y sus relaciones con los ‘átomos vecinos, la naturaleza ha conformado distintos tipos de cristales, los que se pueden clasificar en siete (7) sistemas cristalinos: *triclínicos*, *monoclínicos*, *ortorrómbicos*, *tetragonales*, *trigonales (o romboédricas)*, *hexagonales* y *cúbicos*. Cada uno de estos sistemas cristalinos está ligado a operaciones de simetría que permiten generar al cristal, por ejemplo, si tomamos un átomo que se encuentra en el origen de una celda cúbica (0,0,0), Fig. V-1 (l), y lo trasladamos una distancia “a”, obtenemos el átomo en la posición (a,0,0), si lo trasladamos sobre el eje y se tendría un átomo en la posición (0,a,0) y así siguiendo. Estas posiciones pertenecen al cristal. Pero si se lo desplazara en “a/2” esa no sería una operación permitida, no hay ningún átomo o partícula en la posición (a/2,0,0) en esta estructura. Cuando se modifican algunas operaciones de simetría, como la mencionada, se obtienen los catorce (14) grupos espaciales que estos cristales pueden tener. Los cuales se describen a continuación.

Triclínico: $\alpha \neq \beta \neq \gamma, \alpha \neq \beta \neq \gamma \neq 90^\circ$ (Fig. V-1 (a))

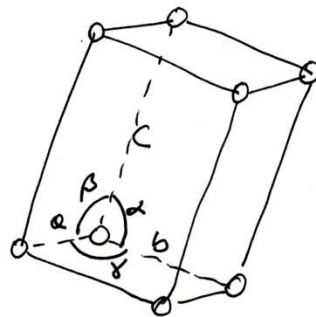


Fig. V-1 (a)

A) Monoclínico: $a \neq b \neq c, \alpha = \gamma = 90^\circ, \beta \neq 90^\circ$ (Fig. V-1, monoclínico simple (b) y monoclínico de cuerpo centrado (c))

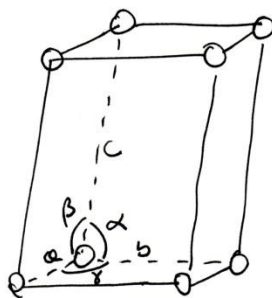


Fig. V-1 (b)

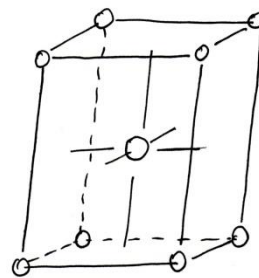


Fig. V-1 (c)

B) Ortorrómbico: $a \neq b \neq c, \alpha = \beta = \gamma = 90^\circ$ (Fig. V-1, ortorrómbica simple (so) (d), ortorrómbica de base centrada (e), ortorrómbica de cuerpo centrado (f), ortorrómbica de cara centrada (g))

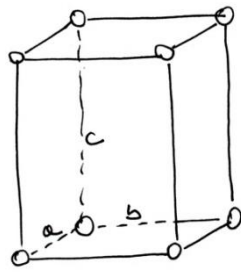


Fig. V-1 (d)

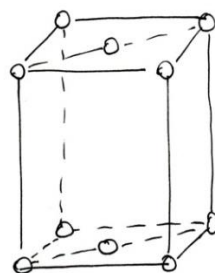


Fig. V-1 (e)

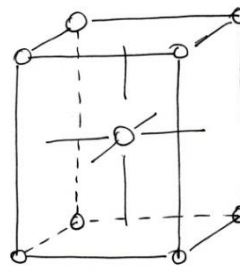


Fig. V-1 (f)

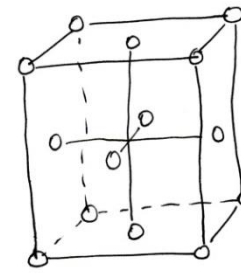


Fig. V-1 (g)

C) Tetragonal: $a = b \neq c$, $\alpha = \beta = \gamma = 90^\circ$ (Fig. V-1, tetragonal simple (st) (h), tetragonal de cuerpo centrado (bct) (i))

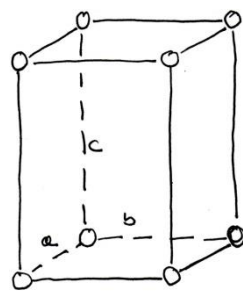


Fig. V-1 (h)

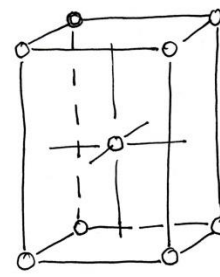


Fig. V-1 (i)

D) Trigonal: $a = b = c$, $\alpha = \beta = \gamma \neq 90^\circ$ (Fig. V-1 (j))

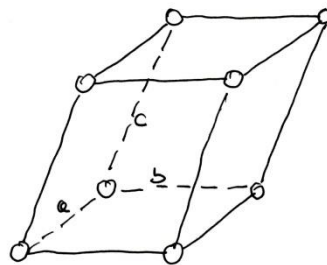


Fig. V-1 (j)

E) Hexagonal: $a = b \neq c$, $\alpha = \beta = 90^\circ$, $\gamma = 120^\circ$ (Fig. V-1 (k))

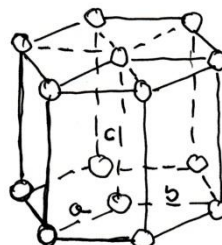


Fig. V-1 (k)

F) Cúbico: $a = b = c$, $\alpha = \beta = \gamma = 90^\circ$ (Fig. V-1, cúbico simple(sc)(l), cúbico de volumen centrado(bcc) (m), cúbico de cara centrada (fcc)(n))

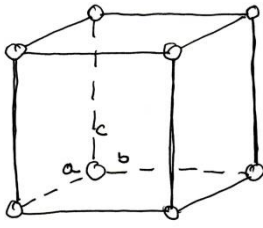


Fig. V-1 (l)

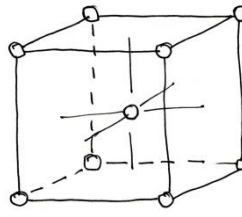


Fig. V-1 (m)

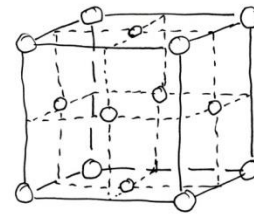


Fig. V-1 (n)

Variaciones y combinaciones entre las diferentes operaciones de simetría dan finalmente un total de 230 grupos cristalinos que son todos los que existen en la naturaleza.

V.2. Características de los sólidos

Se ha mencionado que una de las características de los sólidos es la estructura de posiciones fijas que los átomos que los componen adoptan en el espacio, a diferencia de lo que ocurre con los gases o los líquidos. El motivo que los lleva a adoptar esta distribución es el vínculo que establecen los átomos con sus vecinos, estos vínculos son los diferentes tipos de ligaduras que ellos poseen. Esta es una de las formas de clasificar a un sólido. Se describirán brevemente las diferentes ligaduras que los sólidos pueden adoptar.

a) Sólidos con ligaduras covalentes.

Estos presentan como característica relevante la alta direccionalidad de la unión. Casos típicos de estos son los cristales cúbicos de C (carbono), Si (silicio), GaAs (arseniuro de galio), etc., todos estos presentan una estructura tipo diamante, como se puede ver en la Figura V-2(a), mientras que el Ge (germanio), tiene una estructura cúbica de cara centrada, Fig. V-2(b). En la primera de ellas, se observan los átomos (esferas blancas y grises) que conforman parte de la red cristalina. En esta estructura, cada uno de los átomos está ligado a cuatro átomos vecinos. Las líneas dobles indican las ligaduras que mencionamos. Los átomos involucrados comparten dos electrones, uno de cada uno de ellos, estableciendo un intercambio de carga eléctrica, donde los electrones participantes tienen spines antiparalelos no violando ninguna ley física. Las líneas a trazos largos delimitan la celda cristalina que se forma trasladando y rotando la celda unidad, indicada con líneas de trazos pequeños. La alta direccionalidad de las ligaduras hacen que su estructura sea rígida, difícil de deformar. El C en su estructura diamante (ligaduras covalentes) tiene una dureza de 10 Mohs^(*), mientras que el Fe (hierro) que tiene estructura cristalina cúbica de cuerpo centrado pero no tiene ligaduras covalentes, posee una dureza de 4 Mohs.

^(*) En 1825 el geólogo alemán Carl Frederich Christian Mohs (29-1-1773, 29-9-1839) propuso una escala de durezas de los materiales sobre la base de que un material pueda rayar a otro pero no lo contrario. Así ordenó 10 materiales con durezas de mayor a menor, el mayor es el diamante con dureza 10, mientras que el talco con dureza 1.

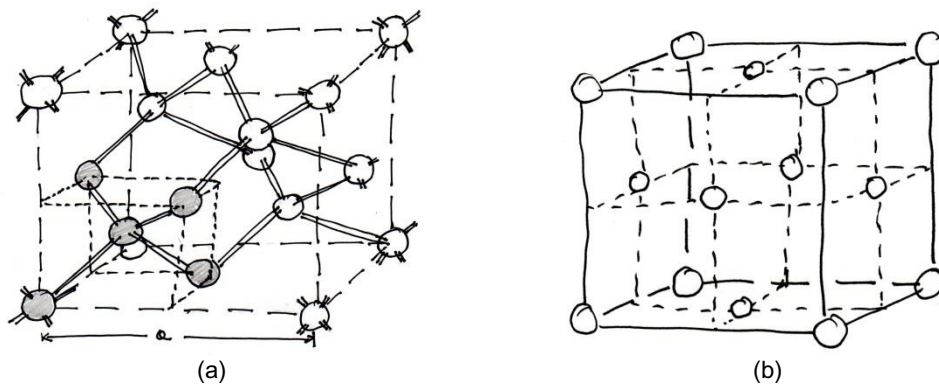


Figura V-2. Estructuras (a) del C, Si, GaAs, tipo diamante, con ligaduras covalentes (los átomos agrisados conforman la celda unidad), y (b) del Ge cúbica de cara centrada (fcc).

En general, estos cristales son pobres conductores de la electricidad y de la temperatura, debido a que poseen pocos electrones libres. Muchos de ellos son cristales transparentes, especialmente el diamante debido a que el primer estado electrónico está a aproximadamente 5,5 eV por encima del correspondiente estado de valencia, y como los fotones del espectro visible se encuentran entre 1,8 y 3,1 eV, no pueden ser absorbidos. Todos los que no cumplen esta condición son opacos.

b) Sólidos formados por cristales iónicos.

Estos sólidos se encuentran en las antípodas de los descritos recientemente. Ellos se forman alternando en el espacio iones positivos y negativos. El ejemplo típico es el cloruro de sodio, la sal (ClNa) como se puede ver en la Figura V-3. La unión entre los átomos es debido a las fuerzas electrostáticas entre ellos. La atracción que presentan los iones de cargas opuestas es del tipo q^+q^-/R (R es la distancia de separación de los iones en equilibrio), la principal contribución a la energía de la ligadura es la llamada energía de Madelung (Erwin Madelung, 18-5-1881, 1-8-1972) (En esta clase de cristales se puede calcular la energía de cohesión mediante la suma de las energías de las interacciones coulombianas (Madelung), teniendo en cuenta solamente la disposición geométrica y las cargas relativas entre iones, en el cristal mencionado, ClNa, cada ion Na^+ y cada ion Cl^- , está rodeado por seis átomos de la otra especie, esto genera un potencial particular entre el ion y sus vecinos y da lugar a la energía de Madelung que conforma una gran parte de la energía del cristal).

Estos son duros, quebradizos, pobres conductores de la temperatura y de la electricidad. Los iones tienen capas completas, el Na: $1s^22s^22p^63s^1$ y el Cl: $1s^22s^22p^63s^23p^5$, por lo que presentan características diamagnéticas, los electrones se encuentran apareados, ellos no presentan un momento magnético neto.

Esta distribución de cargas, hace que presenten una simetría casi esférica, por lo que las ligaduras no tienen direcciones preferenciales.

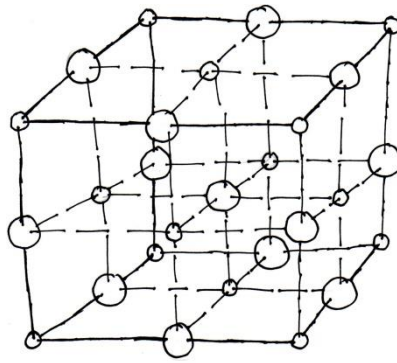
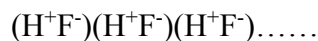


Figura V-3. Cristal del cloruro de sodio. Las esferas grandes corresponderían a los iones cloro (Cl) y las pequeñas a los iones sodio (Na).

c) Sólidos con ligadura de hidrógeno.

Esos, están relacionados con los cristales iónicos. Están caracterizados por fuertes ligaduras polares en las cuales participan uno o más átomos de hidrógeno, por ejemplo el H₂O o el ácido hidrofúrico (HF).

Los iones de hidrógeno positivos, son bastante pequeños y pueden atraer a otras moléculas negativas formando cadenas, como:



Este tipo de unión es particularmente interesante en el caso del hielo. Las moléculas de agua (H₂O) adquieren una disposición tetraédrica.

d) Sólidos metálicos.

Estos, históricamente han sido de suma importancia para el desarrollo de la humanidad por la gran versatilidad y cantidad de aplicaciones. En estos materiales, los electrones se encuentran débilmente ligados a sus capas más externas, lo cual permite que puedan ser fácilmente liberados, conformando un "gas" de electrones libres responsables de las ligaduras de los átomos. En el "gas", los electrones pueden moverse libremente indicando que se encuentran "no localizados". Esta movilidad de los electrones hace que los metales sean buenos conductores térmicos y eléctricos. Las estructuras cristalinas que poseen los metales son diversas, los hay cúbicos, hexagonales, tetraédricos, etc., como también son diversos los tipos de metales existentes, por ejemplo, metales alcalinos, metales de transición, etc. Un ejemplo de esta estructura es el hierro (Fe) cuya estructura es la que se muestra en la Figura V-1(m).

e) Sólidos moleculares.

Estos sólidos están constituidos en general por sustancias cuyas moléculas no son polares. Los electrones en estas moléculas están apareados, debido a este hecho, no se presentan ligaduras covalentes entre los átomos de dos moléculas diferentes. Las moléculas mantienen su individualidad. Las fuerzas que las mantienen ligadas son las de van der Waals, las cuales establecen ligaduras débiles. Los sólidos moleculares no son buenos conductores ni del calor ni de la electricidad, tiene un bajo punto de fusión, son compresibles y deformables. Los gases inertes, cuando son solidificados, forman sólidos moleculares. Otros ejemplos son los cristales de H₂, CO₂, CH₄ y el HCl.

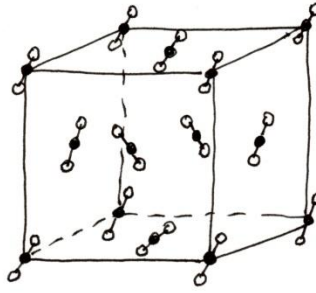


Fig. V-6. Sólido molecular, CO_2 . En cada vértice y en el centro de las caras se encuentra una molécula de CO_2 .

V.3. Niveles de energía en los sólidos. Estructura de banda

Al principio del Capítulo III se ha visto como se forman las bandas en una cadena lineal de átomos. En la realidad el panorama es un poco más complejo, debido a que las estructuras con las cuales se trabajó, se desarrollarán ahora en un espacio tridimensional, por lo que habría que imaginarse los potenciales que se propusieron en una dimensión, en tres dimensiones.

La determinación de los niveles de energía no es un proceso sencillo, hay que solucionar la ecuación de Schrödinger para los potenciales reales, con bases para las funciones de onda que en general son complejas, teniendo en cuenta las condiciones de contorno, etc. Se eliminará este proceso y se hará solamente una descripción fenomenológica.

Se podría comenzar pensando en los átomos como si estuvieran muy alejados, de manera que los electrones de los diferentes átomos no interactúen. En esta circunstancia ellos conservan sus niveles atómicos, y los niveles de energía del sólido serían los del átomo aislado. Si la distancia entre ellos se reduce, los electrones comienzan a interactuar, los niveles atómicos que originalmente contenían los pocos electrones que tenía el átomo, ahora se ven superpuestos. ¿Cuál es la respuesta de la naturaleza a esta nueva situación?, siguiendo el principio de Pauli (ver punto IV.2.) los niveles de energía se dividen en tantas partes como átomos existan en el sólido. La separación entre estos niveles de energía, impuesta por la naturaleza, es tan pequeña que puede considerarse como si fuera un continuo de energía. Esquemáticamente puede verse en la Figura V-7.

En la figura se puede ver que, cuando las distancias entre átomos son grandes (los electrones no interactúan), punto a' , el valor de energía corresponde al nivel atómico. Mientras que cuando se acercan, punto a , ese nivel se ha ensanchado dando cabida a todos los átomos del sólido que a esa distancia, tiene sus electrones interactuando. Cada uno de los niveles, entre los extremos B y C, da lugar a dos electrones con espines opuestos $+\frac{1}{2}$ y $-\frac{1}{2}$.

Si originalmente el nivel atómico poseía 2 electrones y la cantidad de átomos en el cristal es N , la banda alojará $2N$ electrones, si eran 6, ahora serán $6N$, etc. todo dependerá de a que orbital pertenece la banda generada.

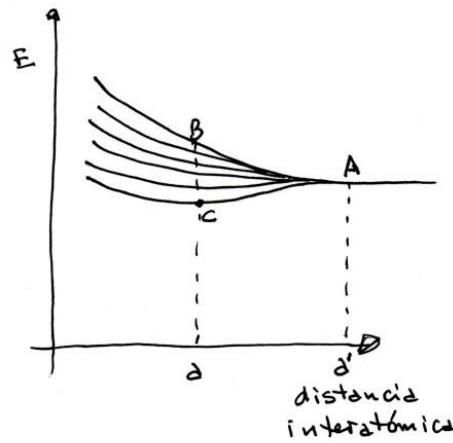


Fig. V-7. Comparación de los niveles de energía en el sólido de acuerdo a la distancia interatómica.

En forma general, en un sólido existen muchas bandas de energía, cada una correspondiente a los orbitales de los átomos que lo conforman. Nuevamente aquí, cuando los átomos se encuentran alejados, no existe interacción entre los electrones y los niveles energéticos se conservan como niveles, mientras que cuando los electrones empiezan a interactuar los niveles pierden esa característica y se transforman en bandas de energía, como se ve esquemáticamente en la Figura V-8.

En esta figura se puede ver la hipotética estructura de banda de dos sólidos distintos dependiendo de sus distancias interatómicas. En el punto a' , más a la derecha, podemos ver dos niveles (1 y 2), son capas más internas, se mantienen como niveles, los electrones no interactúan, están más o menos localizados, probablemente estos niveles pertenezcan a los electrones que están más ligados al núcleo. Siguiendo hacia energías más altas, se pueden ver, ahora sí, dos bandas, la 3 y 4. Si esta estructura perteneciese a un semiconductor, la banda 3 estaría llena, sería la banda de valencia y la 4 vacía a $T=0K$, sería la banda de conducción.

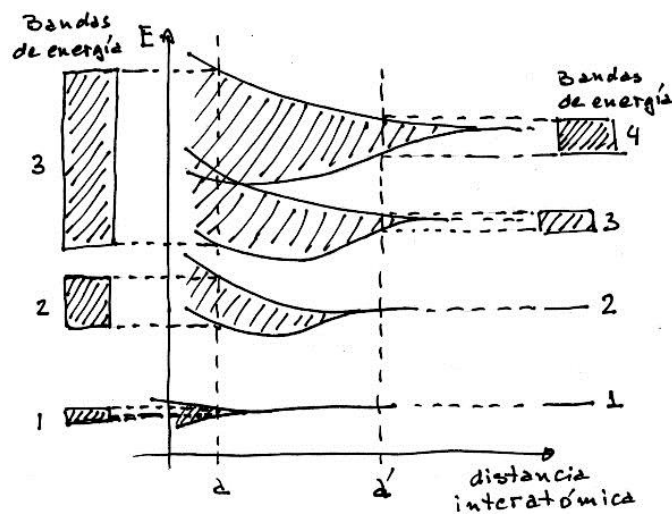


Figura V-8. Bandas de energía en función de la distancia interatómica.

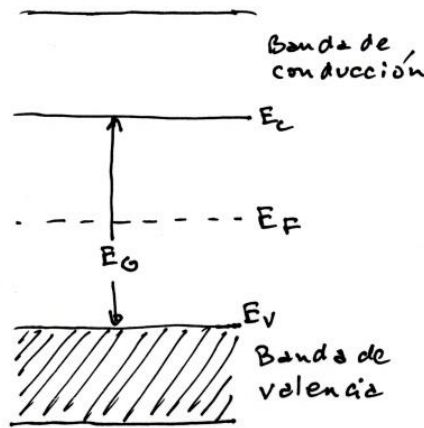
Sobre la izquierda de la gráfica, distancia interatómica a , se tiene una estructura energética diferente, el estado 1 conforma un nivel de energía, pertenece a electrones muy profundos en energía,

seguramente próximos al núcleo. Las bandas 2 y 3 podrían describir a un metal, donde por ejemplo, la banda 2 estaría llena y la 3 podría estar parcialmente llena. Se ve que la banda 3 se forma por la superposición de otras dos bandas, los electrones que participan de ella tienen todos los estados disponibles desde el fondo hasta la parte superior de la banda. En nuestros próximos desarrollos los niveles 1 y 2 sobre la derecha o 1 sobre la izquierda no serán considerados. Los electrones que están en este último nivel (nivel 1 de la izquierda) no interactúan con los vecinos, no son parte de los niveles de valencia. Los electrones que ocupan los niveles 2 y 3 son los que utilizan los sólidos para establecer las uniones químicas con los otros átomos de la red.

V.4. Aisladores, semiconductores y metales

Una de las propiedades sobre la que se trabajara es la conductividad eléctrica. Los materiales, en general, tienen diferente comportamiento con respecto a este parámetro, están los que son buenos conductores, los que dependiendo de condiciones externas pueden conducir o no, y los que no lo hacen bajo ningún aspecto. Hasta principios del siglo XX, la determinación de estas propiedades se hacía por el método de prueba y error. Con el advenimiento de la mecánica cuántica y la teoría de bandas se pudo determinar a ciencia cierta las características de los materiales con respecto a su comportamiento eléctrico. La teoría de bandas permitió determinar cuándo un material, era un aislante o un semiconductor o un metal, con total certeza. Teorías como la del modelo del electrón libre (véase capítulo III), la teoría desarrollada por Kronig-Penney y el hecho de tener en cuenta la estructura periódica de la red con el teorema de Bloch, son las que han permitido llevar a cabo esta identificación en forma precisa.

Se verán algunos ejemplos de materiales característicos de las diferentes clases de sólidos. Se comenzará por un aislador. Estos materiales tienen la particularidad que entre la última banda ocupada, banda de valencia, y la siguiente desocupada, banda de conducción, hay una zona de energías prohibidas muy ancha (varios eV), esto hace que los electrones que se encuentran en la banda de valencia necesiten mucha energía para pasar a la de conducción. Un ejemplo típico de material aislante es el carbono en la estructura cúbica diamante y también en la hexagonal grafito. Esquemáticamente tendríamos una distribución de energías como la que se puede ver en la Figura V-9, en ella, se puede ver la brecha existente entre la parte superior de la banda de valencia, E_V , y la parte inferior de la banda de conducción, E_C , esta diferencia, denominada E_G , energía del ancho de la banda prohibida (gap en inglés) cuando es amplia (del orden de 2,5 eV o más), estamos en presencia de un aislador. En la Figura V-10 se puede ver la estructura de bandas de energía del carbono. El rango de energías se extendió hasta los -30.0 eV con la finalidad de incluir a los electrones más próximos al núcleo, los cuales no participan de los enlaces químicos con sus vecinos. La configuración electrónica del carbono es: $1s^2 2s^2 2p^2$, se puede ver que el nivel 1s está completo, el 2s también, pero el 2p, el cual admite 6 electrones, solamente tiene 2. En la Figura V-10 y V-11 se muestra la estructura de banda y las densidades de estados, todos estos valores han sido calculados en forma teórica con el código de primeros principios Full Potential Local Orbital (FPLO) (Ver Referencias al final del capítulo), en el marco



FiguraV-9. Diagrama de bandas de energía en un aislador.

de la Teoría de la Funcional Densidad, teoría que se desarrolló para encontrar las propiedades físicas de los materiales en sus estados fundamentales (temperatura del sólido 0K), por lo que necesariamente la diferencia de energía entre la banda de valencia y la de conducción (estados excitados) no está bien calculada, el valor experimental de $E_G = 5,5\text{eV}$, mientras que el calculado es de: $4,23\text{eV}$. En el gráfico se puede ver que el nivel de Fermi, E_F , ha sido puesto en la parte inferior de la banda de conducción, por arriba de él, todos los estados están desocupados, mientras que por debajo se encuentran todos ocupados.

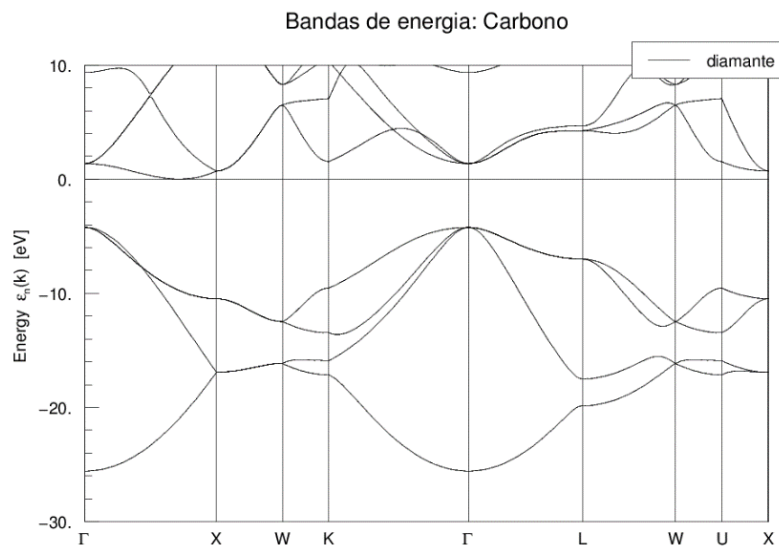


Figura V-10. Bandas de energía del carbono en la estructura diamante. La línea horizontal en el valor cero del eje de la energía corresponde al Nivel de Fermi, E_F .

El eje horizontal de la Figura V-10 se encuentra en unidades de $2\pi/a$ (el parámetro a es la longitud del lado de la celda (en este caso cúbica) del carbono. Los puntos Γ , X, W, etc. (Ver Apéndice V-1) pertenecen a puntos de alta simetría de la celda en el espacio de puntos k (espacio de los momentos) (ver capítulo III ítem 4). En la Figura V-11 Se encuentran las densidades de estados (DOS) las cuales muestran como es la distribución de los electrones en función de la energía. El gráfico superior corresponde a la densidad total, allí se están las contribuciones de todos los orbitales del átomo. En el gráfico medio, está la contribución de

los orbitales “s” solamente, correspondiente al orbital 2s, pues el 1s se encuentra en niveles muy profundos de energía, por debajo de los -30eV.

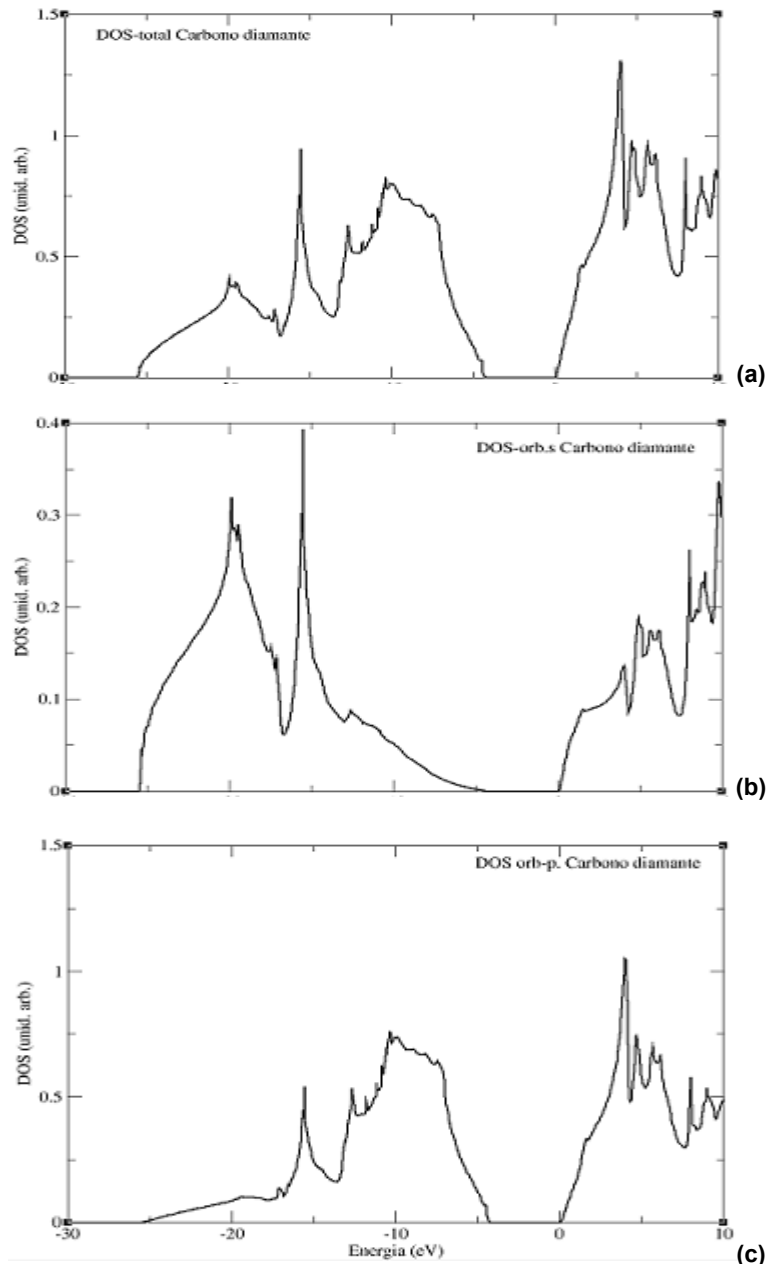


Figura V-11. Densidad de estados (DOS) para el carbono con estructura diamante. La superior (a) es la correspondiente a la densidad total, la del medio (b) a los orbitales “2s” y la inferior (c) a los orbitales “2p”. El valor cero del eje de la energía corresponde al Nivel de Fermi, E_F .

La parte inferior del gráfico corresponde a los orbitales 2p. En la Figura V-12 (a) y (b) se han graficado las bandas ocupadas por cada uno de estos orbitales, las cuales se indican con un engrosamiento de la línea de energía. Las Figuras V-11 y V-12 son complementarias.

Por un lado, el eje de energía en las Figuras V-11 se encuentra en el eje horizontal, mientras que en las V-12 está en el vertical. Ambos ejes abarcan el mismo rango de energías. Se podrían adosar estos ejes (el vertical con el horizontal de las figuras V-10 y la V-11(a)) para ver como varía la ocupación de las diferentes bandas en función de la energía. Sin hacer esto se puede hacer una análisis rápido de orbitales y ocupaciones. Se puede

ver en la Figura V-12 (a) que los electrones 2s ocupan la parte inferior de las bandas de valencia, lo que concuerda con la Figura V-11 (b) donde en el rango entre -25eV y -12eV está la parte principal de estos orbitales, luego la densidad de estados va disminuyendo hasta anularse en el tope de la banda de valencia (-4,23eV). Se puede ver en esa Figura, los estados que ocuparían los electrones 2s si fuesen excitados y pasaran a la banda de conducción. En la Figura V-12 (b) se pueden ver los estados 2p, estos se encuentran en el rango de -15eV a -4,23eV. Estos electrones se encuentran en la parte superior de la banda de valencia, mientras que por debajo de este valor (-15eV) tienen pocos estados ocupados (Figura V-11 (c)). Queda claro que son estos electrones los que actúan como electrones de valencia, estableciendo vínculos con los átomos vecinos. Nuevamente aquí en la banda de conducción se ven los estados que serían ocupados por estos electrones si ellos fueran excitados y pudieran superar el ancho de la banda prohibida.

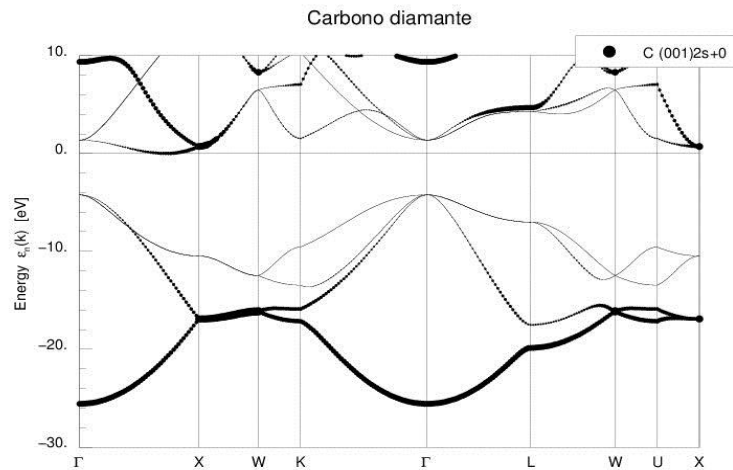


Figura V-12. (a)

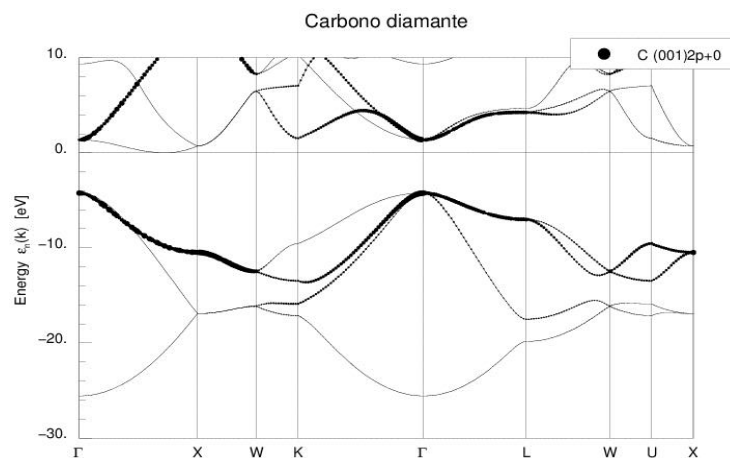


Figura V-12. (b)

Otros materiales que tienen semejanza con los aisladores son los semiconductores. La semejanza proviene del hecho que estos también tienen una estructura de banda con una zona de energías prohibidas entre la banda de valencia ocupada y la de conducción vacía, pero en este caso, es mucho más estrecha que la de los aisladores, la estructura de energías coincide con el esquema que se muestra en la Figura V-9. Uno de los semiconductores más conocido y sobre el que se basa la industria electrónica es el silicio (Si, $Z=14$), este tiene la misma estructura cristalina que el carbono diamante (Figura V-2 (a)) pero aun teniendo iguales estructuras sus propiedades son diferentes. El carbono diamante es la estructura natural conocida con el índice de dureza más alto, en la escala de Mohs, el diamante tiene un valor de 10 (tope de la escala), mientras que para el silicio ese índice vale 6,5. Otra diferencia visible es que el diamante es transparente, mientras que el silicio no lo es, este comportamiento radica en el hecho que los fotones de la luz, en el caso del carbono, no tienen suficiente energía para excitar a los electrones de la banda de valencia y enviarlos a la de conducción, por lo tanto, pasan a través del material sin ser absorbidos. En el caso del silicio, son absorbidos, los electrones son promovidos a la banda de conducción y el material es no transparente. El silicio es un semiconductor de ancho de banda prohibida $E_G = 1,12\text{eV}$, tiene la banda 3s completa con 2 electrones y en la banda 3p, que admite 6 electrones, posee solamente 2 electrones, por lo tanto es un material cuya configuración electrónica es $1s^2 2s^2 2p^6 3s^2 3p^2$. El silicio en los orbitales 3s y 3p podría acomodar 8 electrones, $2s+6p = 8$ estados, pero solamente tiene $2s+2p = 4$ estados ocupados. Cuando se conforma el sólido, estos estados se dividen y quedan los 4 estados ocupados en la banda de valencia, mientras que los 4 restantes se encuentran en la de conducción. En la Tabla V-1 se pueden ver los elementos que tiene una configuración electrónica parecida, todos estos cristales tiene estructura cúbica, por lo que su volumen es: $V = a \times a \times a$, donde a es la arista del cubo.

Elemento	Número Atómico (Z)	Configuración electrónica	Tipo de estructura	Long. Arista de la celda a (Å) (10^{-10}m)	Ancho banda prohibida (eV)
Carbono (C)	6	[He] $2s^2 2p^2$	Cúbica diamante	3,57	5,5
Silicio (Si)	14	[Ne] $3s^2 3p^2$	Cúbica diamante	5,43	1,12
Germanio (Ge)	32	[Ar] $3d^{10} 4s^2 4p^2$	Cúbica Cara centrada (FCC)	5,66	0,66

Tabla V-1. Características físicas de los elementos Carbono, Silicio y Germanio.

Se puede ver de la Tabla V-1 que a medida que aumenta el tamaño de la celda, disminuye el ancho de la banda prohibida. Las bandas s y p se superponen, en un sólido con N átomos, estas ofrecen $8N$ estados, al disminuir la distancia interatómica a , la banda se separa en dos de $4N$ estados cada una. La inferior, que contiene $4N$ electrones es la banda de valencia, la superior con $4N$ estados vacíos es la de conducción.

En la Figura V-13 se ha graficado, en forma esquemática, las bandas s y p del Carbono, Silicio y Germanio, se debe tener en cuenta que cada una de ellas pertenece a niveles de energía diferentes, 2 para el C, 3 para el Si y 4 para el Ge.

Se pueden ver ahora las densidades de estado del silicio Figuras V-14 (a), (b) y (c) y las bandas de energía (Fig. V-15) correspondientes a los orbitales s y p . En las Figuras (a) se encuentran las densidades o bandas totales y sin ningún carácter resaltado, mientras que en las (b) y (c) están las correspondientes a los orbitales s y p , respectivamente. En el eje horizontal, donde se encuentran las energías, el valor cero (0) corresponde a la Energía de Fermi, en este caso se encuentra en la parte superior de la banda de valencia.

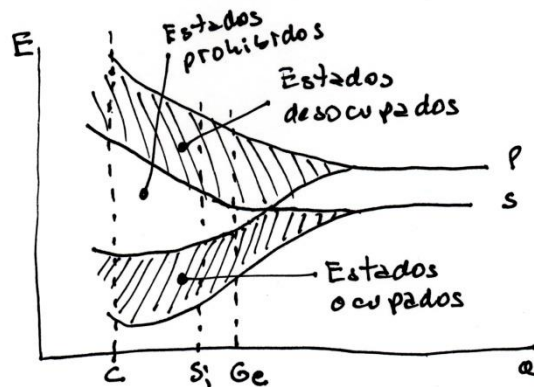


Figura V-13. Energía de las bandas s y p en función de la distancia interatómica.

Los tres elementos mencionados, cuando forman un sólido tienen propiedades diferentes. El Carbono (C), es un aislante, mientras que el Silicio (Si) y el Germanio (Ge) son semiconductores, con anchos de bandas prohibidas diferentes. No obstante ello, tienen similitud en cuanto a la ocupación que generan en los orbitales s y p , los orbitales más externos que estos sólidos poseen. Esto hace que desde el punto de vista de sus estructuras de bandas, estas sean parecidas.

Se puede observar, como en el caso del Silicio, que el ancho calculado de la banda prohibida, $\approx 1\text{eV}$, es inferior al experimental, $1,12\text{eV}$, lo cual es de esperar, de acuerdo a la teoría en cuyo marco, este valor fue calculado. En las Figuras V-15 se muestran las bandas de energía calculadas, para el silicio y en donde, gráficos (b) y (c) se han remarcado los lugares que ocupan los electrones en esos orbitales.

Vemos en esta figura que el silicio es un semiconductor de gap indirecto, esto quiere decir que para el mismo valor de k , el tope de la banda de valencia (punto Γ) no se enfrenta con el fondo de la banda de conducción (próximo al punto X). El mínimo en la banda de conducción se encuentra en $0,8X$, aproximadamente.

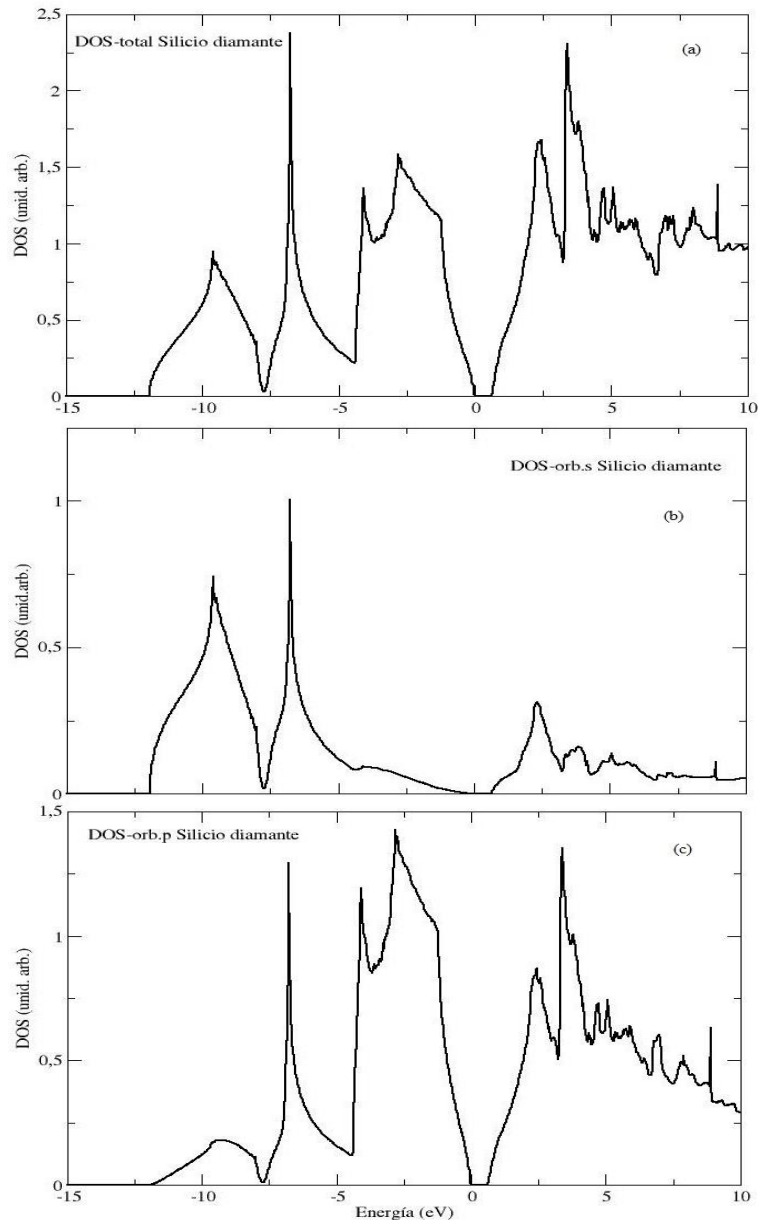


Figura V-14. Densidad de estados del silicio, total (a), orbital s (b) y orbital p (c). En el valor cero del eje de la energía (eje horizontal) se encuentra la Energía de Fermi, E_F

Si se comparan estas gráficas con las del carbono podemos notar una similitud muy grande, cambian los valores de las energías en juego, pero las formas de las bandas son muy parecidas, esto se debe a que se está hablando de los electrones s y p en diferentes estados energéticos. En este caso el nivel de Fermi se ha indicado en el tope de la banda de valencia. En la Figura V-15 (c) se puede ver que los estados correspondientes a los orbitales 3p, se encuentran más próximos al nivel de Fermi que los pertenecientes a los orbitales 3s, esto quiere decir que aquellos electrones serán más fácilmente excitados, y podrán pasar a la banda de conducción con menores cantidades de energía puestas en juego. Los mecanismos para excitar a los electrones y promoverlos a las bandas superiores (de conducción) son variados, pueden ser campos eléctricos, magnéticos, incremento de la temperatura, etc. Los electrones que pasen a la banda de conducción por medio de alguno de estos mecanismos, en semiconductores, que como en este caso no han sido dopados, se los llama electrones intrínsecos y la conducción que se logre con ellos es la conducción intrínseca.

En el caso de los aisladores, el ancho del gap es muy grande, se necesita mucha energía para que un electrón de la banda de valencia pueda pasar a las de conducción. Ese no es el caso de los semiconductores, de allí su nombre. Si bien estos a 0K son aislantes, a medida que la temperatura se incrementa, los electrones ganan energía y pueden empezar a saltar a la banda de conducción y comportarse como en un metal, formando un gas de electrones el cual

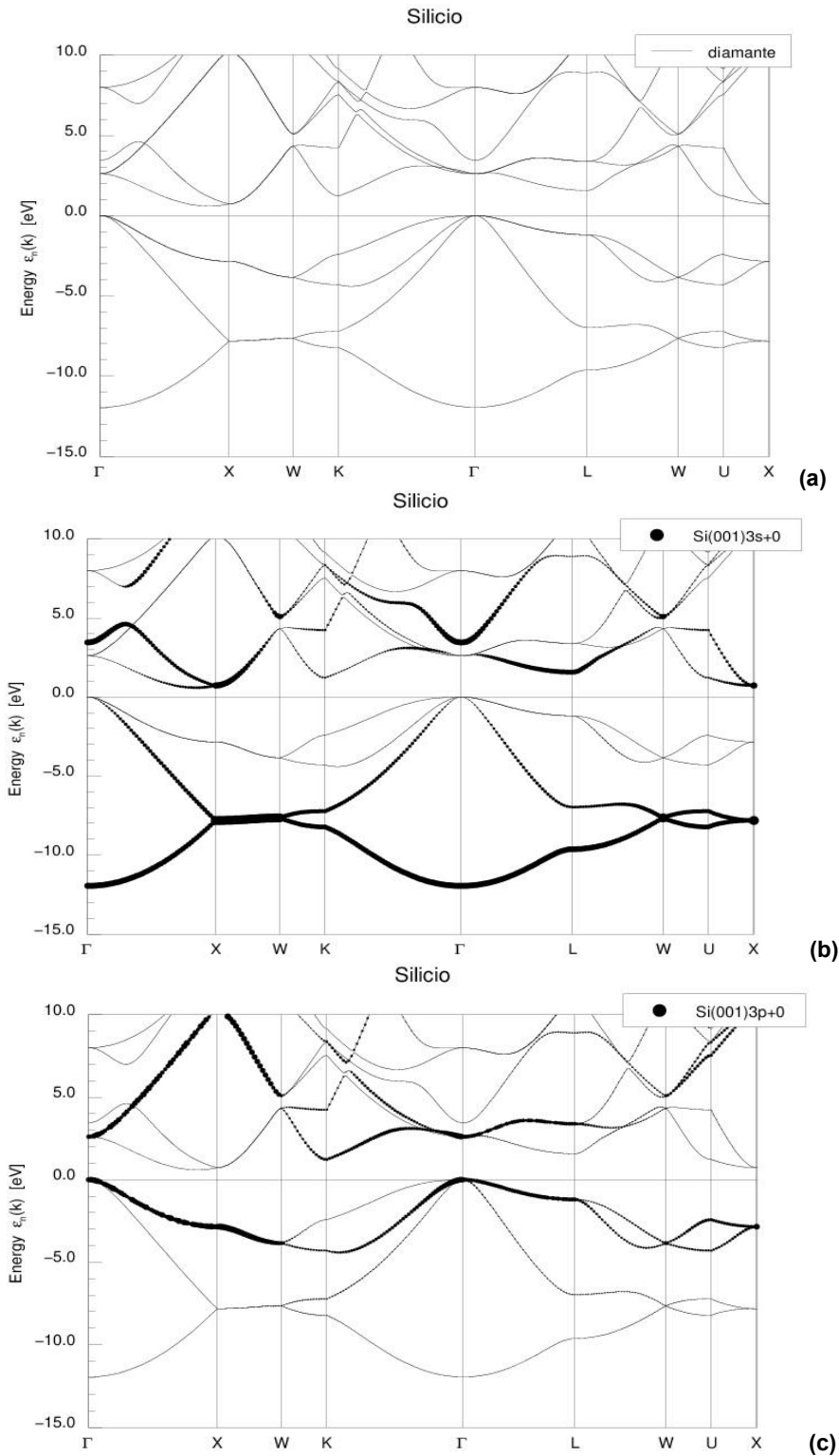


Figura V-15. Estructura de energía de bandas para el Si (a), remarcando los estados 3s (b) y remarcando los estados 3p (c). La línea horizontal representa la Energía de Fermi.

puede ser manejado aplicando potenciales sobre ellos o por medio de campos externos.

En el otro extremo de esta descripción se encuentran los metales, los cuales no presentan el salto que se ha mencionado entre la banda de valencia y la de conducción, sino que es un continuo de energías con estados ocupados por debajo del nivel de Fermi y estados vacíos por encima de él a 0K. A medida que la temperatura se incrementa, los electrones pasan rápidamente por sobre E_F y el metal comienza a conducir. Por este motivo, los metales son buenos conductores. Estos se dividen en diferentes clases, alcalinos, alcalinotérreos, de transición, lantánidos, etc. todos ellos, como se mencionó son buenos conductores eléctricos y también térmicos. Esquemáticamente la estructura energética de un metal es como se ve en la Figura V-16.

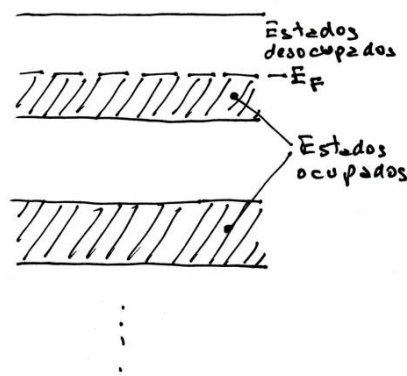


Figura V-16. Estructura de bandas de un metal.

Esta descripción gráfica del metal indica como es la distribución de bandas de energía y su ocupación. Como se mencionó, a 0K todos los electrones del metal se encuentran por debajo del nivel de Fermi (E_F) hasta el final de la banda. Todas las bandas por debajo de esta, desde el punto de vista energético, están totalmente ocupadas. Por encima de E_F , esta banda se encuentra desocupada, de la misma manera que todas las que están por encima de ella. Hay otros metales que superponen sus bandas, brindando a los electrones muchos más estados posibles de ocupar. La excitación de los electrones, para que crucen E_F , en los metales puede lograrse por medio de diferentes mecanismos. Se los puede excitar térmicamente, esto produce un desorden térmico que no aporta a la conducción de corriente. También se los puede excitar por medio de la aplicación de un campo eléctrico, en este caso la nube electrónica que se forma, se mueve (movimiento colectivo) en forma opuesta al campo aplicado. Esto constituye una corriente eléctrica. Estos electrones ocupan lugares desocupados de la banda, por encima de E_F . La excitación también puede ser combinando ambos efectos. Se tomará como ejemplo de metal al aluminio (Al, $Z=13$), uno de los elementos más abundantes en nuestro planeta. En la Figura V-17 (a), (b) y (c), se puede distinguir claramente las diferencias que existe entre los semiconductores y aisladores con los metales si comparamos con las figuras V-11 y V14.

En los primeros se podía ver que en el valor 0eV (nivel de Fermi) había una separación entre bandas, mientras que en este caso, se plantea un continuo de estados, a 0K los que se encuentran a la izquierda del 0eV estarán ocupados y por encima de él desocupados.

Se puede ver también en las Figuras V-18 (a), (b) y (c) la estructura de bandas general, la estructura de banda donde se han resaltado los lugares que ocupan los electrones s y los que ocupan los electrones p, respectivamente. En estas figuras se puede notar que el nivel de Fer-

mi corta las bandas, indicando que estamos en presencia de un metal. En la Figura V-18 (c) la leyenda dice SUM, eso indica que se está graficando la suma de los orbitales p , $p_x + p_y + p_z$.

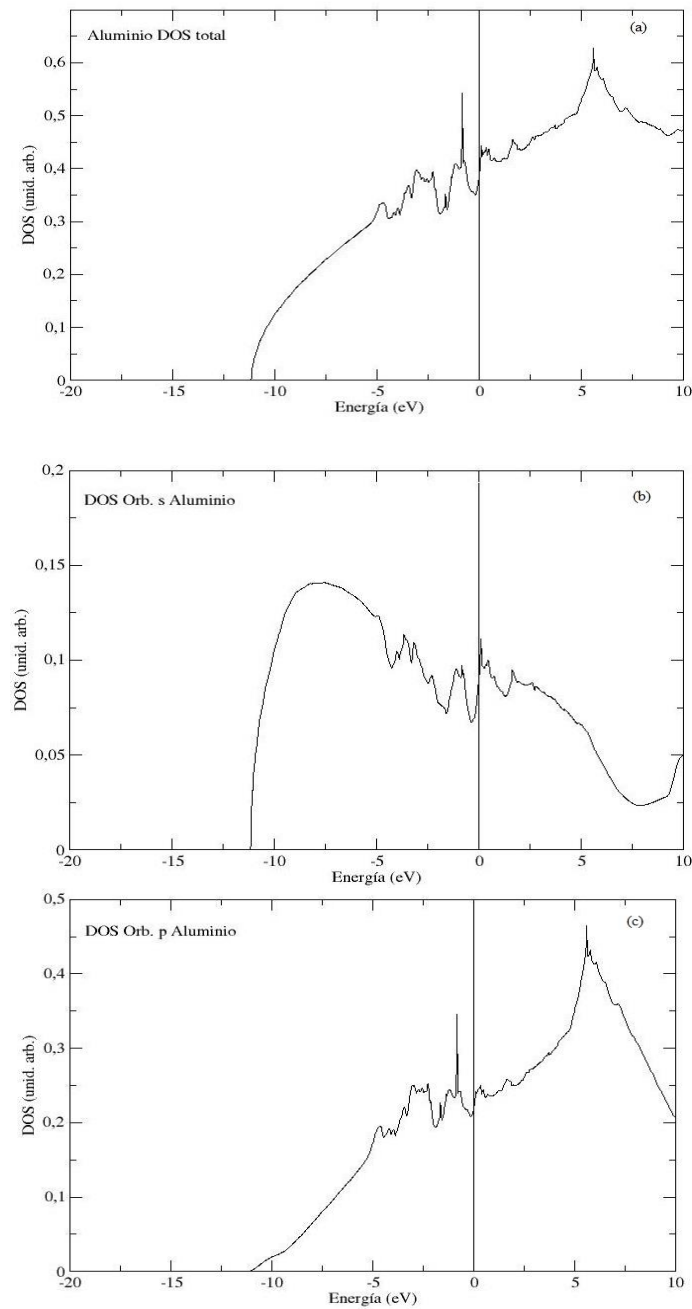


Figura V-17. Densidad de estados del Aluminio, total (a), orb. S (b) y orb. P (c). El valor 0 en el eje de energía corresponde a la Energía de Fermi, E_F .

La determinación del comportamiento eléctrico de los materiales dejó su carácter empírico con la aparición de la teoría de bandas de sólidos. El gran mérito de la teoría fue que, mediante explicaciones simples sobre comportamientos eléctricos tan diferentes entre unos y otros materiales, se pudo determinar fehacientemente la clase de material que se tenía bajo estudio.

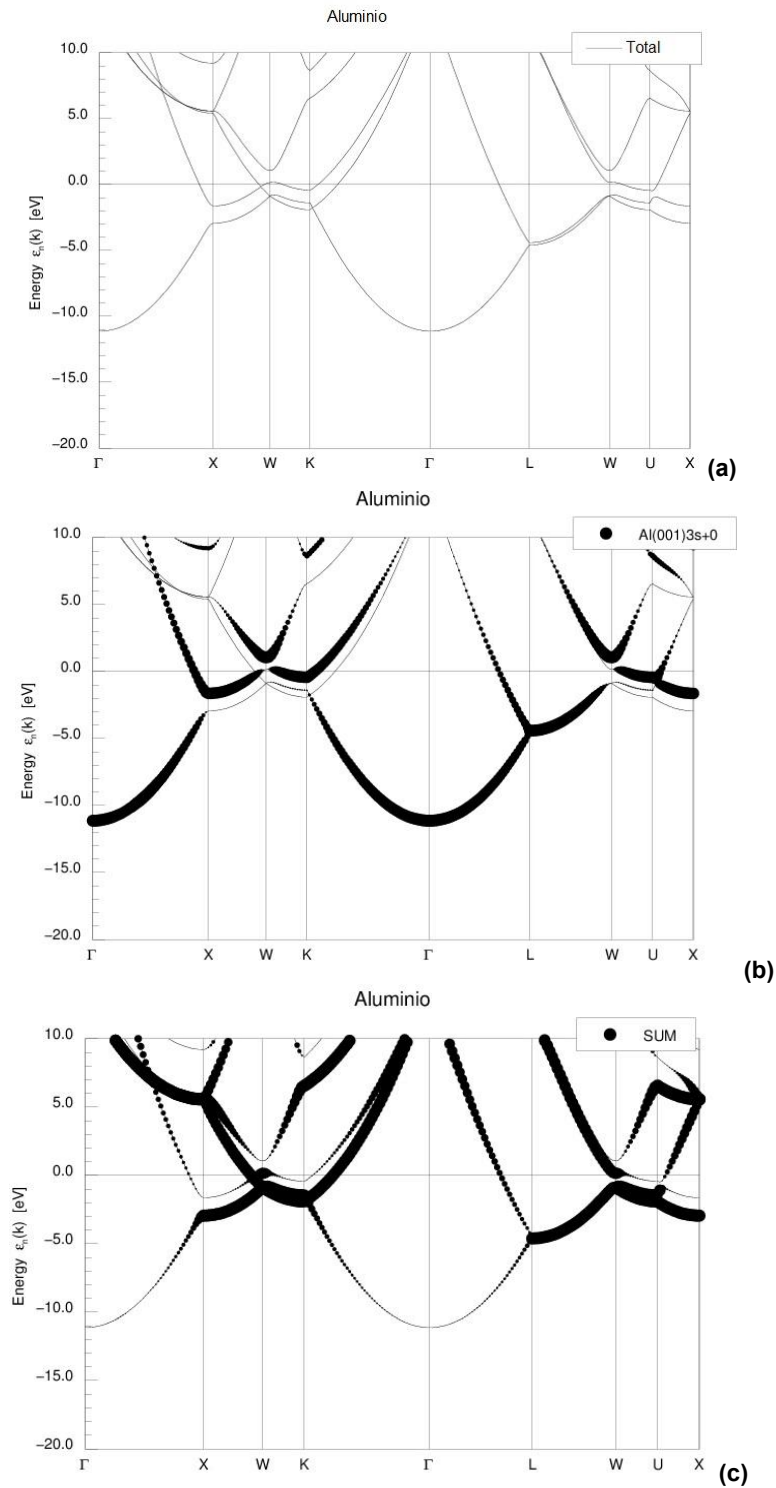


Figura V-18. Bandas de energía sin indicación de orbitales (a), orbitales 3s (b) y orbitales 3p (c). La línea horizontal corresponde a la Energía de Fermi, E_F

Se puede mediante un esquema simple sintetizar los conceptos que se han expuesto, en la Figura V-19 se ha graficado esta idea.

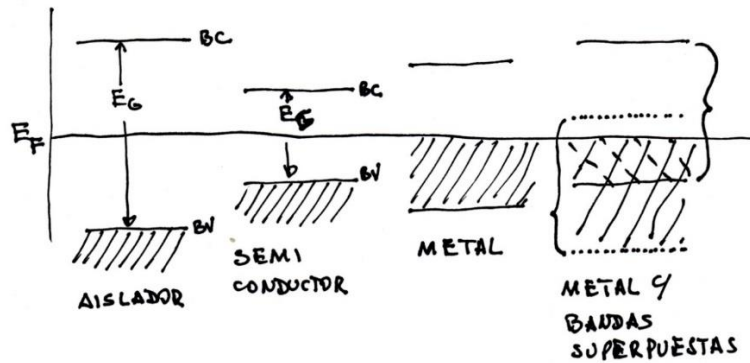


Figura V-19. Esquema de diagrama de bandas para materiales eléctricamente diferentes.

Se ve en la Fig. V-19 un esquema de la estructura de bandas de los materiales que hemos mencionado previamente. En los dos primeros lugares de la izquierda se tienen al aislador y al semiconductor, en ambos, entre la última banda llena (banda de valencia, B_V) y la primera vacía (banda de conducción, B_C), hay una zona de energías prohibidas (E_G), donde no existen estados electrónicos del material. Esta banda prohibida es mucho más ancha en los aisladores que en los semiconductores. A continuación se encuentra un metal, una de sus bandas se encuentra dividida por el nivel de Fermi, los estados ocupados van desde el fondo de esta banda hasta el E_F , y desde allí hasta la finalización de la banda, están los estados desocupados, por ejemplo, si la banda es de carácter "s", hay $2N$ estados posibles, si el material tuviera un solo electrón s, quiere decir que hay $1N$ estados ocupados y $1N$ estados vacíos. En el último lugar hay otro metal ligeramente diferente al anterior, en el, hay dos bandas que se superponen, los estados ocupados van desde el fondo de la banda más profunda hasta el nivel de Fermi, incluyendo el límite inferior de la banda menos profunda. Superando el E_F , los electrones encontrarán más estados electrónicos libres, indicados por la suma de estados de cada una de las bandas menos las que se encuentran ocupadas. Para ejemplificar esto se puede suponer que se tiene un material con orbitales s y p que se superponen, los s pueden alojar $2N$ electrones y los p, $6N$ electrones, como las bandas están superpuestas se tiene $2N + 6N = 8N$ estados. Se supondrá que el material tiene $2e$ en la banda s, quiere decir que $2N$ estados ocupados, quedarán $6N$ estados vacíos que los electrones podrán ocupar.

Apéndice V

Celda en el espacio recíproco

Se ha visto en el capítulo III que la energía de un electrón libre es una función que depende del vector de onda k de la función de onda del electrón, cuya dimensión es:

$$[k] = \frac{2\pi}{\text{long.}} \quad (\text{B-1})$$

Se debe pensar que la celda del cristal se encuentra en el espacio real, en el caso de la celda cúbica como la del Carbono, cada lado de la celda tiene dimensión a . Para pasar al espacio de las k -es, llamado espacio recíproco, se debe hacer la transformada de Fourier de la celda real.

El caso de la energía del electrón libre fue una introducción para mostrar que, para obtener la estructura de banda de un material se debe hacer el cálculo en el espacio recíproco (espacio transformado). A través de este mecanismo, la celda cúbica que se ve en la Figura (V-2)(a) (la indicada con líneas punteadas) se transforma en la figura que vemos abajo.

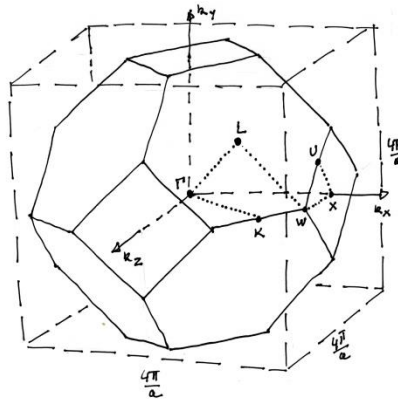


Figura B-1. Diagrama de una celda cúbica de cara centrada en el espacio recíproco con la inclusión de puntos de simetría.



Jean-Baptiste Joseph Fourier
21-3-1768, 16-5-1830

La transformada de Fourier hace pasar una celda cúbica de lado a , a una celda cúbica en el espacio recíproco de lado $4\pi/a$. Además en ese espacio se pueden definir, de acuerdo a las simetrías existentes en la celda, puntos que tiene características especiales (no se abordará este tema en este texto). El cálculo de energías en las direcciones que vinculan estos puntos es lo que constituye un diagrama de energía de bandas. En el caso de los diagramas de bandas que se mostraron para el carbono, el recorrido elegido fue el Γ -X-W-K- Γ -L-W-U-X, primero

se recorrió la superficie de energías Γ -X-W-K- Γ y luego valores internos como el Γ -L (en esta dirección se encuentra en el espacio real la ligadura de dos átomos vecinos de C), y luego se vio como eran las energías sobre la superficie W-U-X. En el modelo de Kronig-Penney se mencionaron las zonas de Brillouin, ese modelo se desarrolló en forma unidimensional, ahora se está viendo la primera zona de Brillouin para un cristal tridimensional. En función de las simetrías que posee este tipo de cristal, es suficiente realizar un análisis de 1/48ava parte de su volumen (delimitado por los puntos Γ -X-W-K- Γ -L-U-X) en el espacio recíproco (1era.zona de Brillouin), para obtener información de la totalidad del cristal.



Léon Brillouin
7-8-1889, 4-10-1969

Referencias

- Albela, J. M., Martínez-Duart, J. M., *Fundamentos de electrónica física y microelectrónica*. 1996 Addison-Wesley Universidad Autónoma de Madrid.
- Alonso, M., Finn, E. J. *Fundamental University Physics. Vol. III*. 1968 Addison –Wesley Company.
- Ashcroft, N. W., Mermin, N. D., *Solid State Physics*. 1981 Holt-Saunders International Editions.
- Dalven, R. *Introduction to Applied Solid State Physics* 1990 Plenum Press
- Ibach, H., Lüth, H. *Solid State Physics An Introduction to Theory and Experiments* 1991 Springer-Verlag.
- Kittel, C. *Introduction to Solid State Physics Sixth Edition* 1986 John Wiley and Sons Inc. New York.
- Shalimova, K. V. *Física de los Semiconductores* 1975 Editorial MIR
- Van der Ziel, A. *Electrónica Física del Estado Sólido*. 1968 Editorial Prentice/Hall Inc., Englewood Cliff, New Jersey
- Método Full Potential Local Orbitals. FPLO (Desarrollado en las 3 siguientes publicaciones)*
- Koepnik, K, Eschrig H, Phys. Rev. B, 59, 1743:1775 (1999)
- Koepnik, K., Velicky, B., Hayn, R., Eschrig, H. Phys. Rev.B 55, 5717 (1997)
- Opahle, I., Koepnik, K., Eschrig, H., Phys. Rev. B 60, 14035 (1999)

CAPÍTULO VI

Semiconductores

*Llegará un momento en que creas
que todo ha terminado.*

Ese será el principio

EPICURO

341 - 270AC

En el capítulo anterior se ha hecho una breve introducción sobre los materiales semiconductores, específicamente sobre aquellos a los que no se les ha introducido ninguna impureza en su estructura, estos han sido llamados intrínsecos. Ahora se avanzará sobre el tema y se estudiarán aquellos a los cuales, para mejorar alguna de sus características se les agregan átomos diferentes (impurezas, dopantes) a los del material receptor (semiconductor).

La historia de la humanidad está jalonada por hechos singulares que dejan alguna marca en el mundo, el comienzo de la utilización del hierro, hace cientos de años, o la utilización del carbón durante la era industrial del siglo XIX, por ejemplo. Estos logros permitieron que la humanidad avanzara significativamente en épocas pasadas. Durante el siglo XX el gran descubrimiento, los semiconductores, produjeron cambios significativos en muchos y diferentes aspectos. Desde el punto de vista científico, porque se pudo explicar el comportamiento de un material no utilizado como tal en otras épocas. Desde el punto de vista tecnológico, el impacto ha sido monumental. Desde mediados del siglo XX en adelante, prácticamente no ha habido desarrollo tecnológico que no involucrase algún dispositivo electrónico, los cuales por supuesto, están basados en semiconductores.

Una nueva revolución se ha planteado hace algunos años: la miniaturización. Este proceso involucra no solamente a los materiales, sino a la misma teoría de la Física, para saber dónde está el límite al cual se puede llegar reduciendo tamaños. Las dimensiones de los objetos que se estudian indican en que marco teórico uno se debe mover. Sobre esto hemos hablado muy brevemente en algún capítulo anterior, con el modelo de Kronig-Penney. Es claro que si deseamos estudiar los aspectos atómicos de algún material no podemos utilizar las leyes de Newton, y si queremos estudiar el movimiento de los planetas, no se puede utilizar la teoría de Schrödinger. Lo mismo pasa para la nanoescala que se está proponiendo y desarrollando en los nuevos dispositivos. Se debe buscar una teoría que brinde respuestas dentro del nuevo marco de escalas que se está desarrollando actualmente.

Pero para llegar a estos nuevos conocimientos, se deben haber estudiado, entendido, asentado y madurado todos los conocimientos previos. Es como querer empezar a correr cuando uno aún no sabe caminar.

VI.1. Semiconductores, introducción de impurezas

Hasta ahora se ha trabajado con materiales puros. Esto significa que no se ha incluido ningún tipo de impureza en la estructura cristalina. Se verá ahora como se modifican algunas de sus propiedades cuando se introduce un átomo diferente en el cristal. Se hará especial hincapié en los semiconductores que son los materiales esenciales en los dispositivos electrónicos. ¿Para qué se introducen impurezas en los semiconductores? La respuesta inmediata sobre la cual nos abocaremos, pero no es la única, es que con las impurezas, se pueden modificar, por ejemplo, sus propiedades de conducción. Además, de acuerdo al tipo de impureza que se introduzca, se puede definir el tipo de corriente que tendrá el dispositivo.

La introducción de un dopante produce cambios en la estructura cristalina del receptor, modifica la periodicidad del cristal. También genera variaciones en su estructura de bandas.

Para desarrollar el tema, se puede partir pensando en una estructura como la del silicio, este es un átomo con valencia cuatro (tetraivalente), esto quiere decir que comparte sus cuatro electrones con sus cuatro vecinos, y cada uno de ellos con sus cuatros vecinos y así siguiendo, en la Figura VI-1 se representa esta estructura.

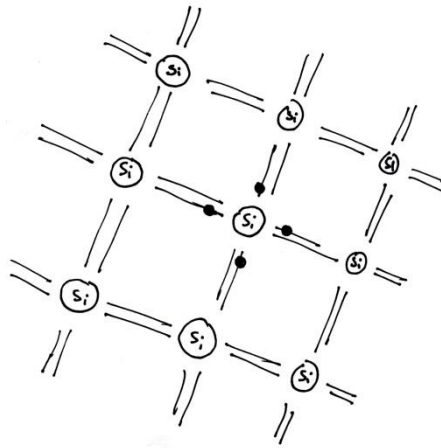


Figura VI-1. Diagrama esquemático de las uniones de los átomos de Si con sus vecinos.

¿Cómo debe entenderse la Figura VI-1? Esta es un esquema bidimensional de las uniones de los átomos de silicio con sus vecinos. La estructura real es la que se muestra en la Figura V-2(a) donde cada átomo de Si tiene 4 electrones para compartir en la unión. Este tipo de ligadura donde se comparten (intercambian) electrones entre átomos, son llamadas covalentes (misma que tiene el carbono en la estructura diamante) forman uniones muy rígidas. Como ejemplo se puede tomar el átomo central de la Figura VI-1, él está rodeado

por ocho (8) líneas que representan electrones, cuatro (4) de ellas, las que van a sus cuatro (4) vecinos que poseen un círculo negro, son los electrones propios, las otras 4 son las que aportan cada uno de sus 4 vecinos. Esto se repite en toda la estructura cristalina. El diagrama de bandas de energía es del tipo al que se muestra en la Figura V-9, con un E_G mucho más pequeño que el de un aislador.

Para introducir una impureza, uno de estos átomos de silicio debe ser reemplazado por uno diferente, por ejemplo, un elemento pentavalente, el Arsénico (As, $Z = 33$). Este es un átomo que tiene la siguiente configuración electrónica $[Ar]4s^23d^{10}4p^3$, como la capa 3d está completa con los 10 electrones, esta no aporta a la valencia, por lo que quedan los 5 electrones de valencia que son provistos por la capa 4s y 4p. En el ejemplo anterior se podría tomar el átomo del centro de la figura y reemplazarlo, con lo cual se tendría un diagrama como el de la Figura VI-2 (a).

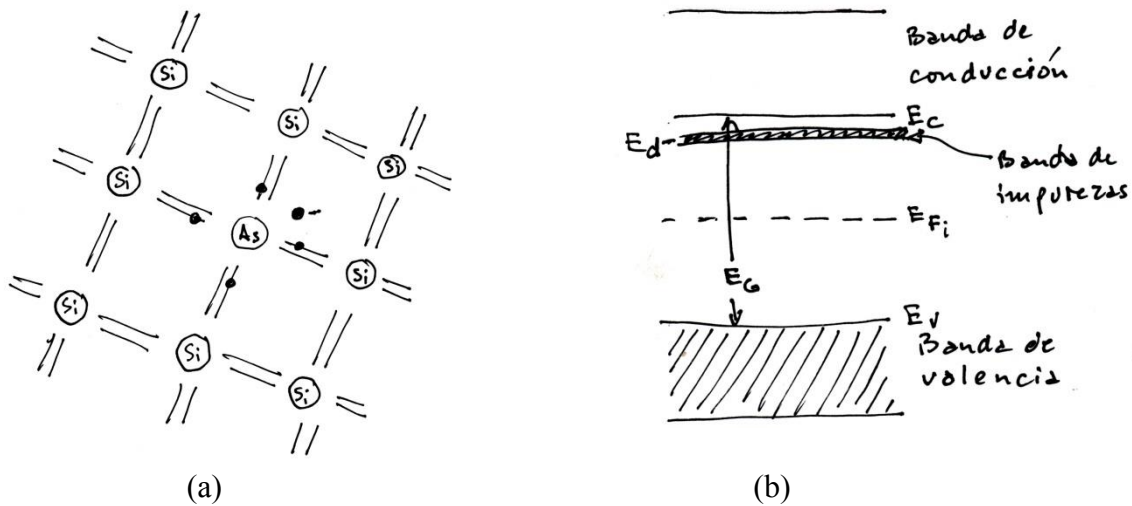


Figura VI-2 (a). Diagrama esquemático de la introducción de una impureza del grupo V, en la red de Silicio. (b) Estructura de bandas con impureza pentavalente.

En la Figura VI-2 (a) se puede ver que al introducir una impureza con un electrón más que los que tiene el Silicio, este átomo completa las cuatro valencias del Silicio con sus vecinos y queda una carga negativa (un electrón) débilmente ligado. Esta impureza se llama “donora”, brinda un electrón al material. El semiconductor que posee este tipo de impurezas es un semiconductor tipo “n” (n por la carga negativa del electrón). En este ejemplo, no se está considerando el reacondicionamiento de la estructura por la inclusión de un átomo diferente. Este electrón que no pertenece al sistema receptor genera un nivel de energía muy próximo a la banda de conducción, dentro de la banda prohibida del Silicio, como puede verse en la Figura VI-2(b). Este nivel se encuentra a una energía E_d por debajo del límite inferior de la banda de conducción (E_c). Estos electrones, en general, están muy débilmente ligados, por lo que se necesita muy poca energía para excitarlos y que salten a la banda de conducción. La conducción con los portadores cedidos por átomos de impurezas, que han sido ionizados, se llama extrínseca. Por ejemplo, en el Silicio, si el átomo de impureza fuera el Antimonio (Sb, $Z = 51$), $E_d = 0.039\text{eV}$ y si fuese Arsénico (As, $Z = 33$), $E_d = 0.054\text{eV}$, ambos por debajo de la banda de conducción.

Por otro lado, si el reemplazo del átomo de Silicio fuese por un átomo trivalente (valencia = 3, átomo del grupo III, por ejemplo Boro, (B, $Z = 5$, $[\text{He}]2s^22p^1$) la situación, aunque diferente, sería análoga a la recién vista. Estos átomos podrían restablecer los vínculos con solamente 3 átomos vecinos de Silicio, por lo que una de las uniones quedaría defectuosa. Ese defecto generado por falencia de un electrón se llama “huevo”, estos sitios pueden ser fácilmente ocupados por electrones. Esto hace que la estructura de banda se vea modificada, pues se introduce un nivel de energía en la banda prohibida que es la que aloja a todos estos huecos generados. Este nivel se llama “aceptor”, se encuentra a una distancia de E_a por sobre el borde superior de la banda de valencia del material receptor. El semiconductor que posee este tipo de impureza se denomina semiconductor tipo “p” (p por la carga positiva del huevo). Esta descripción se puede ver en la Figura VI-3 (a) y (b). Como en el caso anterior, en el gráfico (a), estamos haciendo una representación bidimensional de la nueva situación generada, sin tener en cuenta el acomodamiento de la estructura ante la introducción de un átomo diferente.

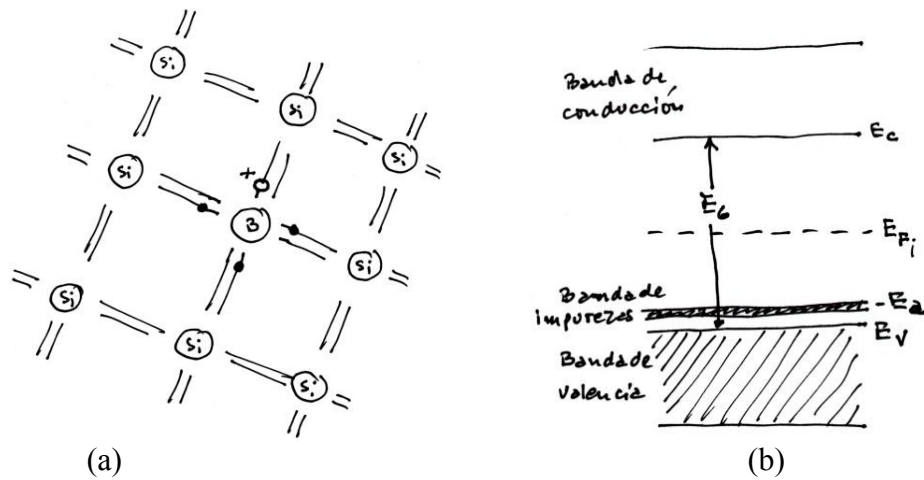


Figura VI-3. (a). Diagrama esquemático de la introducción de una impureza del grupo III, en la red de Silicio. (b) Estructura de bandas con impureza trivalente.

En la Figura VI-3 (a) se puede observar que el átomo de impureza aporta 3 electrones, el cuarto electrón que debería proveer para restablecer todas las uniones que tenía el Silicio no está, por lo tanto, hay un electrón faltante, indicado en la figura con un círculo vacío. Al no existir la carga negativa que brinda el electrón, se le asigna a este portador una carga positiva. De manera similar al caso anterior, los huecos creados generan una banda aceptora que se encuentra muy cerca (Figura VI-3(b)), en términos de energía, de la banda de valencia. Si el material receptor fuera el Silicio y la impureza Boro, $E_a = 0.045\text{eV}$ y si fuera Aluminio (Al, $Z = 13$), $E_a = 0.067\text{eV}$, ambos por encima del límite superior de la banda de valencia.

VI.2. Determinación de los niveles de Fermi en semiconductores intrínsecos y extrínsecos

En un semiconductor intrínseco, donde no existen dopantes, los electrones que puedan excitarse y pasar a la banda de conducción, dejarán en la banda de valencia un lugar vacío, un hueco. Se deduce de este razonamiento, que habrá igual número de electrones en la banda de conducción que de huecos en la de valencia. Esto quiere decir que las concentraciones de portadores en cada una de las bandas es la misma. Se puede encontrar la distribución de portadores mediante la estadística de Fermi-Dirac y las funciones densidad de estados para cada una de las bandas. Si un electrón es excitado y pasa a la banda de conducción, se comportará como un electrón libre, lo mismo pasará con el hueco que ha quedado en la banda de valencia, ambos tendrán una energía como la que se indica en la ecuación III-7, dependiente de k^2 , y la curva de energía tendrá forma parabólica.

La densidad de estados en la banda de conducción se puede escribir:

$$g_C(E)dE = \frac{8\sqrt{2}\pi}{h^3} m_n^{*3/2} \sqrt{E - E_C} dE \quad (VI-1)$$

Mientras que para la banda de valencia, será:

$$g_V(E)dE = \frac{8\sqrt{2}\pi}{h^3} m_p^{*3/2} \sqrt{E_V - E} dE \quad (VI-2)$$

Donde m_n^* y m_p^* son las masas efectivas de los portadores en la banda de conducción y de valencia respectivamente. En el intervalo, $E_V < E < E_C$, la región prohibida, se tiene que $g(E) dE = 0$.

En condiciones de equilibrio térmico, el número de electrones en la banda de conducción está dado por la densidad de portadores multiplicada por la probabilidad de ocupación (función de distribución de Fermi), esto es:

$$dn_0 = f_0(E)g_C(E) = \frac{8\sqrt{2}\pi}{h^3} m_n^{*3/2} \frac{\sqrt{E - E_C}}{1 + e^{-\frac{E - E_F}{kT}}} dE \quad (VI-3)$$

Como se ha visto en el Capítulo IV, a temperaturas alrededor de la ambiente y más altas, la función de Fermi, puede aproximarse por la de Boltzmann. Debido a esto f_0 queda:

$$f_0(E) \cong \frac{1}{e^{\frac{E - E_F}{kT}}} \quad (VI-4)$$

Si se reemplaza este valor de f_0 en (VI-3) y se integra entre E_C e ∞ se tendrá:

$$n_0 = \int_{E_C}^{\infty} dn_0 = \frac{8\sqrt{2}\pi}{h^3} m_n^{*3/2} e^{E_F/kT} \int_{E_C}^{\infty} \sqrt{E - E_C} e^{-E/kT} dE \quad (VI-5)$$

Para integrar esta expresión se debe hacer la sustitución $x_C = (E - E_C)/kT$, con lo cual la (VI-5) queda:

$$n_0 = 2 \left[\frac{2\pi m_n^* kT}{h^2} \right]^{3/2} e^{-(E_C - E_F)/kT} = N_C e^{-(E_C - E_F)/kT} \quad (VI-6)$$

n_0 es la cantidad de portadores, (e^-), electrones intrínsecos, en la banda de conducción, provenientes de átomos térmicamente ionizados, N_C es la densidad efectiva de estados en la banda de conducción. Si $m_n^* = m_0$ (masa en reposo del electrón) y $T = 300K$, entonces, $N_C = 2,51 \times 10^{19} \text{cm}^{-3}$.

El número de huecos, h^+ , en la banda de valencia se encuentra en forma totalmente análoga,

$$dp_0 = f_{p0}(E)g_V(E)dE = [1 - f_0(E)]g_V(E)dE \quad (VI-7)$$

Integrando la expresión (VI-7) entre los límites $-\infty$ y E_V , se encuentra que:

$$p_0 = 2 \left[\frac{2\pi m_p^* kT}{h^2} \right]^{3/2} e^{-(E_F - E_V)/kT} = N_V e^{-(E_F - E_V)/kT} \quad (VI-8)$$

Como se ha mencionado, la concentración de portadores en un material intrínseco es la misma en la banda de valencia como en la de conducción, en ambos casos los portadores son intrínsecos, esto se puede escribir como, $n_0 = n_i$ y $p_0 = p_i$, igualando las expresiones (VI-6) y (VI-8), se tiene:

$$N_C e^{-(E_C - E_F)/kT} = N_V e^{-(E_F - E_V)/kT} \quad (VI-9)$$

Operando se obtiene:

$$E_F = \frac{1}{2}(E_C + E_V) + \frac{kT}{2} \ln \left(\frac{N_V}{N_C} \right) \quad (VI-10)$$

Si se reemplazan N_C y N_V por sus valores, se tiene:

$$E_F = \frac{1}{2}(E_C + E_V) + \frac{3}{4}kT \ln \left(\frac{m_p^*}{m_n^*} \right) \quad (VI-11)$$

Si la temperatura fuese $T = 0K$, el nivel de Fermi se encuentra en la mitad de la banda prohibida. Por ser un material intrínseco, lo llamaremos:

$$E_{Fi} = \frac{1}{2}(E_C + E_V) \quad (VI-12)$$

Si las masas efectivas de electrones y huecos fuesen iguales, $m_n^* = m_p^*$, nuevamente el nivel de Fermi se encontraría en la mitad del gap. Si $m_n^* > m_p^*$ o si $m_n^* < m_p^*$ el nivel de Fermi se moverá ligeramente hacia abajo o hacia arriba, respectivamente, con respecto a la posición central. La densidad de estados depende en forma directa de las masas efectivas de los porta-

dores, grandes o pequeñas masas efectivas implican grandes o pequeñas densidades de estados. El nivel de Fermi debe ubicarse dentro de la banda prohibida de manera tal de mantener igualados el número de electrones y huecos.

A los fines prácticos, debido a que la variación establecida por la relación de masas efectivas es muy pequeña, el nivel de Fermi intrínseco se considera que se encuentra en la mitad de la banda prohibida.

Ejemplo: Si se toma Silicio y se supone que las masas efectivas fueran una el doble de la otra, $m_n^* = 2 m_p^*$, tendríamos a $T = 300K$, que

$$E_{Fi} - \frac{1}{2}(E_C + E_V) = \frac{3}{4}kT \ln\left(\frac{1}{2}\right) = \frac{3}{4}(0,026)\ln(0,5) = -0,014eV$$

La diferencia entre el nivel de Fermi y el centro de la banda prohibida es -14meV, si se compara con la posición del E_{Fi} en el Silicio que es: 0,56eV, se ve que la diferencia es del 2,5%.

También se puede expresar las concentraciones de equilibrio de electrones y huecos en función de parámetros intrínsecos. Se puede partir de la ecuación (VI-6) y se suma y resta la energía de Fermi intrínseca E_{Fi} ,

$$n_0 = N_C e^{\frac{[-(E_C - E_{Fi}) + (E_F - E_{Fi})]}{kT}} \tag{VI-13}$$

o,

$$n_0 = N_C e^{\frac{-(E_C - E_{Fi})}{kT}} e^{\frac{(E_F - E_{Fi})}{kT}} \tag{VI-14}$$

En esta expresión se puede distinguir la concentración intrínseca, la cual tiene la forma:

$$n_i = N_C e^{\frac{-(E_C - E_{Fi})}{kT}} \tag{VI-15}$$

Por lo que la ecuación (VI-14) puede ser escrita:

$$n_0 = n_i e^{\frac{(E_F - E_{Fi})}{kT}} \tag{VI-16}$$

Siguiendo un proceso similar se puede escribir la concentración intrínseca de huecos:

$$p_0 = n_i e^{\frac{-(E_F - E_{Fi})}{kT}} \tag{VI-17}$$

A diferencia de lo que se había visto, que el nivel de Fermi cambiaba con la modificación de los valores del dopado, en estas dos últimas ecuaciones se puede ver que los valores n_0 y p_0 cambiarán de acuerdo a los niveles de Fermi. Las variaciones de E_F con respecto a E_{Fi} determinarán si el material es tipo- n o p .

Se puede ver ahora como se ubica el nivel de Fermi cuando un semiconductor está dopado con alguna impureza donora, tipo n o aceptora, tipo p ; el procedimiento para ambos es equivalente, se desarrollará para un semiconductor dopado tipo n . Se debe recordar que cuando el semiconductor tiene una impureza tipo- n , esta genera un nivel de energía dentro de la banda prohibida con una energía E_d , próxima y por debajo de la banda de conducción, como en la Figura VI-4.

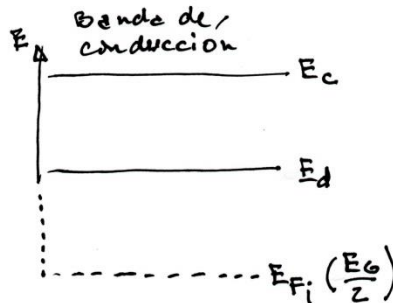


Figura VI-4. Diagrama de bandas de semiconductor dopado con impureza donora.

Como en el caso del semiconductor intrínseco, se puede determinar la densidad de ambos portadores. Debido a que el nivel de Fermi debe describir la situación a la que se ha llevado al semiconductor con el dopado realizado, quiere decir que como es un material tipo- n , el nivel de Fermi deberá estar cerca de la banda de conducción. Se puede imaginar que E_F está por debajo de E_c y por encima de E_d , se debe aclarar que el eje de energías se están tomando positivo hacia arriba, como en la Figura VI-5.

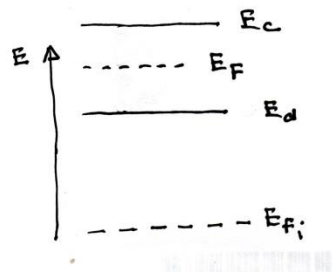


Figura VI-5. Diagrama de bandas indicando el nivel de Fermi debido a impurezas.

Por lo tanto, la cantidad de portadores extrínsecos en la banda de conducción será

$$n = N_c e^{-(E_c - E_F) / kT} \tag{VI-18}$$

Por otro lado, la cantidad de donores en la banda de impureza es:

$$p = N_d e^{-(E_F - E_d) / kT} \tag{VI-19}$$

Donde N_d es la densidad efectiva de estados en la banda donora. Nuevamente aquí, cuando un electrón salte de la banda de impureza a la de conducción, esta última tendrá un electrón extrínseco y habrá quedado un átomo ionizado en la banda de impurezas. Por este motivo las

poblaciones de electrones y átomos ionizados será iguales, se pueden igualar n y p de las ecuaciones (VI-18) y (VI-19)

$$N_C e^{-(E_C - E_F)/kT} = N_d e^{-(E_F - E_d)/kT}$$

De esta igualdad se puede determinar el valor de E_F .

$$E_F = \frac{(E_C + E_d)}{2} + \frac{kT}{2} \ln \left(\frac{N_d}{N_C} \right) \quad (\text{VI-20})$$

Se puede observar aquí que si $T = 0\text{K}$, el nivel de Fermi se encuentra en la mitad entre la banda de conducción y la de impurezas, lo mismo ocurre si la relación en el argumento del logaritmo es 1, es decir, si ambas densidades efectivas fuesen iguales.

VI.3. Variación con la temperatura

Tanto N_C como N_d son expresiones que dependen de la temperatura ($\sim T^{3/2}$). Si al aumentar la temperatura, el valor de N_C crece más rápidamente que el de N_d , todos los electrones que se liberen de la banda de impureza hacia la de conducción encontrarán estados vacíos y podrán ser alojados en estos estados. Se está en presencia de un semiconductor no degenerado. El nivel de Fermi irá decreciendo lentamente hacia el nivel de Fermi intrínseco (E_{Fi}). En cambio, si la banda de impureza libera muchos más electrones que los que pueda aceptar N_C , el semiconductor se llama degenerado, $N_d > N_C$. Si la temperatura sigue aumentando, el nivel de Fermi deberá aumentar para dar cabida a todos los electrones provenientes de E_d . Como la cantidad de impurezas es finita, en algún momento todos los átomos habrán sido ionizados, no proveyendo de más portadores a la banda de conducción, en ese momento la curva que describe el movimiento de E_F , habrá llegado a una meseta y comenzará a bajar. Matemáticamente:

$$\frac{dE_F}{dT} = 0 = \frac{k}{2} \ln \left(\frac{N_d}{N_C} \right) \Rightarrow N_d = N_C \quad (\text{VI-21})$$

Si hacemos que la temperatura siga creciendo, N_C seguirá aumentando, no así N_d , E_F se empieza a comportar como una función lineal de T , ver Figura (VI-6). La concentración de portadores en la banda de conducción permanecerá constante, la banda de impureza ha entregado todos los portadores que tenía.

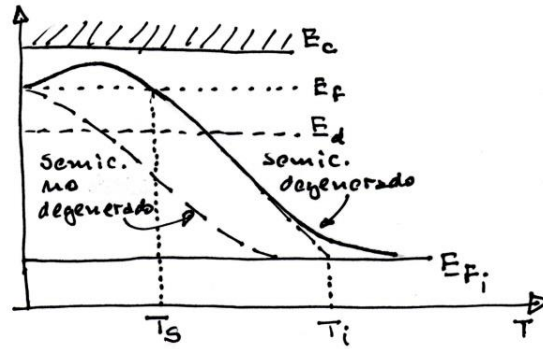


Figura VI-6. Variación del nivel de Fermi con la temperatura.

Cuando el nivel de Fermi cruza la banda de impurezas, $E_F = E_d$, se define la temperatura de extinción, T_s . La cantidad de portadores que hay en la banda de conducción es N_d . Si se utiliza la expresión (VI-18) reemplazando la T por T_s , se tiene:

$$N_d = N_C e^{E_d/kT_s} \Rightarrow T_s = E_d / k \ln \left(\frac{N_d}{N_C} \right) \tag{VI-22}$$

Se puede ver de esta ecuación que la temperatura de extinción decrece con la disminución de la energía E_d . Todas las deducciones realizadas para portadores “n” es absolutamente similar a la que se puede hacer con los “p”.

Ejemplo VI-1: Si tenemos Silicio dopado con antimonio (Sb) el nivel de la banda de impureza del Sb en Si se encontrará 0.039eV por debajo de la banda de conducción. Se puede suponer que $N_d = 10^{16} \text{ cm}^{-3}$. Por lo tanto $T_s = 57,82\text{K}$. Esto quiere decir que a partir de los 57,82K, estadísticamente el Antimonio estará exhausto, todos los átomos estarán ionizados, y hasta que la banda de valencia empiece a producir portadores, la corriente en el Silicio estará conformada por portadores extrínsecos.

Si la temperatura sigue incrementándose, en algún momento se alcanza una excitación térmica suficiente para que los electrones de la banda de valencia puedan pasar a la de conducción. A partir de allí, los portadores que haya en la banda de conducción serán los extrínsecos, provenientes de las impurezas, n_{imp} , y los intrínsecos, cuya cantidad es p que es la densidad de huecos dejada en la banda de valencia. Es decir:

$$n = p + n_{imp} \tag{VI-23}$$

Los procesos que se dan en los semiconductores cuando se produce el pasaje de portadores de una banda a otra, dopados o no, son los de generación y de recombinación, los cuales producen cambios en los semiconductores. En una banda, el aumento de un tipo de portador implica la disminución del otro tipo, debido a la recombinación. Se puede pensar

(válido para ambos portadores) que el proceso de recombinación, R , es proporcional al número de portadores tipo- n , (en este caso electrones,) en la banda de conducción, y tipo- p en la banda de valencia.

$$R = \rho np \quad (\text{VI-24})$$

Esta ecuación indica que la recombinación es proporcional ($\rho \rightarrow$ coeficiente de proporcionalidad, coeficiente de captura) al producto np , el cual es constante una vez que se establece el equilibrio termodinámico. En el equilibrio, los valores n y p son los parámetros que se han llamado n_0 y p_0 en el párrafo anterior.

Si se multiplican las concentraciones de portadores (VI-18) y (VI-19), se tiene:

$$np = n_i^2 = N_C N_V e^{-\frac{E_G}{kT}} \quad (\text{VI-25})$$

El producto np sólo depende de T y de propiedades propias del material receptor, E_G y masas efectivas. Esta relación, que se mantiene constante, es la llamada Ley de Acción de Masas, y establece el vínculo entre las concentraciones relativas de electrones y huecos. Al ser np una constante, está claro que si n crece, p decrece y viceversa. De la ecuación (VI-25) se puede obtener:

$$n_i = \sqrt{N_C N_V} e^{-E_G/2kT} \quad (\text{VI-26})$$

En función de la dependencia que tiene el producto np , es también válido para semiconductores extrínsecos. Si se vuelve a la ecuación (VI-23) vemos que, la cantidad de electrones provenientes de las impurezas es N_d , y que la cantidad de portadores intrínsecos se puede determinar a partir de la Ley de Acción de Masas,

$$n = p + N_d \quad (\text{VI-27})$$

Si reemplazamos la ec. (VI-27) en la ec. (VI-25), se tiene:

$$(p + N_d)p = n_i^2 \quad (\text{VI-28})$$

Operando se tiene una ecuación de segundo grado:

$$p^2 + N_d p - n_i^2 = 0$$

Cuya solución es:

$$p = -\frac{N_d}{2} + \sqrt{\frac{1}{4}N_d^2 + n_i^2} = \frac{N_d}{2} \left(-1 + \sqrt{1 + 4 \frac{n_i^2}{N_d^2}} \right) \quad (\text{VI-29})$$

Si se reemplaza (VI-29) en (VI-27) se tiene:

$$n = \frac{N_d}{2} \left(1 + \sqrt{1 + 4 \frac{n_i^2}{N_d^2}} \right) \quad (\text{VI-30})$$

Se puede pensar aquí en dos posibles opciones, una donde $n_i \ll N_d$ este sería el caso extrínseco, bajas temperaturas, si operamos con esta opción en la ecuación (VI-30) se tiene que:

$$n = N_d \tag{VI-31}$$

Por otro lado, si se toma la otra opción: $n_i \gg N_d$ estaríamos en el caso intrínseco, altas temperaturas, nuevamente si se introduce esta desigualdad en la ecuación (VI-30) se tiene:

$$n = n_i \tag{VI-32}$$

Esta última relación muestra que a altas temperaturas el semiconductor tiende a ser intrínseco, como se puede ver en la Figura VI-6, donde el nivel de Fermi a altas temperaturas tiende al nivel de Fermi intrínseco. La cantidad de portadores provenientes de la banda de valencia es muy superior a la brindada por la banda de impureza.

Lo que se debería hacer ahora es encontrar la temperatura a la cual la banda de valencia comienza a producir portadores intrínsecos. Para ello se debe tomar la ec. (VI-10) rescribiéndola en función de E_G , el ancho de energía del gap, y la ecuación (VI-18) teniendo en cuenta que el nivel de Fermi varía linealmente con T desde el comienzo del comportamiento de extinción, lo que indica que la pendiente $(\ln(N_d/N_C))$ es constante, por lo tanto se tiene:

$$-\frac{E_G}{2} + \frac{kT_i}{2} \ln\left(\frac{N_V}{N_C}\right) = kT_i \ln\left(\frac{N_d}{N_C}\right) \tag{VI-33}$$

Si se despeja T_i de (VI-33) se tiene:

$$T_i = \frac{E_G}{k \ln\left(\frac{N_C N_V}{N_d^2}\right)} \tag{VI-34}$$

Se ve aquí que a diferencia con la temperatura de extinción, la temperatura intrínseca depende de las características propias del semiconductor y del dopado establecido.

Ejemplo VI-2: siguiendo el ejemplo de la página 111, con los valores:

$$E_G = 1,12\text{eV},$$

$$N_C = 2,51 \times 10^{19} \text{cm}^{-3},$$

$$N_V = 1,04 \times 10^{19} \text{cm}^{-3},$$

$$N_d = 1 \times 10^{16} \text{cm}^{-3},$$

$$k = 8,617 \times 10^{-5} \text{ eV/K}$$

se determinará la temperatura intrínseca para el silicio dopado con Antimonio. Reemplazando estos valores en la ecuación (VI-34) se tiene que $T_i = 879,7\text{K}$.

Debe quedar claro que entre el rango de los 57,82K que era la temperatura de extinción y los 879,7K, el semiconductor funcionará, estadísticamente, con los portadores extrínsecos.

Se puede ver claramente la dependencia que tiene la variación de la cantidad de portadores en función de la inversa de la temperatura. En el grafico (Figura VI-7) se pueden distinguir 3 regiones, la a-b es la zona extrínseca, solamente los átomos de impureza han cedido electrones, la conducción es extrínseca. La zona b-c representa el rango de extinción, los dopantes están completamente ionizados, han brindado sus electrones a la banda de conducción, no hay suficiente energía para que los electrones de la banda de valencia pasen a la de conducción. Finalmente, la zona c-d es la zona intrínseca, hay suficiente energía térmica para que los electrones de la banda de valencia pasen a la de conducción y puedan conducir. Se tiene ahora en la banda de conducción ambos portadores, los extrínsecos y los intrínsecos, debido a que el número de los segundos es muy superior a los primeros, estos pueden ser tenidos en cuenta como una pequeña variación de los intrínsecos.

VI.4. Cuasi niveles de Fermi

Se ha visto que el nivel de Fermi y la temperatura determinan unívocamente las concentraciones de electrones y huecos en un semiconductor, como se ve en las ecuaciones (VI-6) y (VI-8)

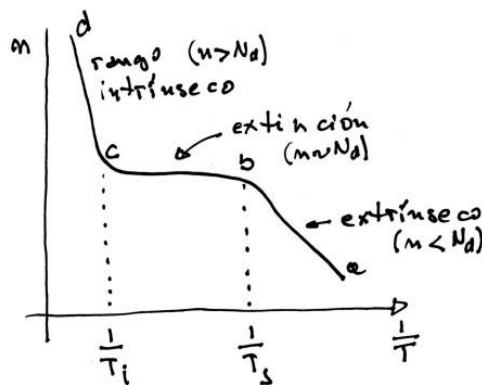


Figura VI-7. Variación de la cantidad de portadores en función de $1/T$

$$n_0 = N_C e^{-(E_C - E_F)/kT} \quad \text{y} \quad p_0 = N_V e^{-(E_F - E_V)/kT}$$

Estas mismas ecuaciones se reescribieron, ec. (VI-16) y ec. (VI-17)

$$n_0 = n_i e^{\frac{(E_F - E_{Fi})}{kT}} \quad \text{y} \quad p_0 = n_i e^{\frac{-(E_F - E_{Fi})}{kT}}$$

En función del nivel de Fermi y el nivel de Fermi intrínseco, niveles que pueden verse en la Figura VI-5, con un semiconductor dopado n , donde $E_F > E_{Fi}$, para este caso, se puede ver de las dos últimas ecuaciones que $n_0 > n_i$ y que $p_0 < n_i$. Si se introduce una perturbación (inyección de portadores) y se lleva el sistema a un estado de no equilibrio, las concentraciones de portadores diferirán de aquellas en equilibrio. Claramente esta nueva situación debe ser descripta

por el nivel de Fermi. Para ello es conveniente escribir las concentraciones de portadores del no equilibrio en función de nuevos niveles de Fermi, que se llamarán Cuasi Niveles de Fermi o Pseudo Niveles de Fermi, y se denotarán como E_{Fn} o E_{Fp} , dependiendo a qué tipo de portador se está haciendo referencia, electrones o huecos, respectivamente. Por lo tanto se pueden reescribir las ecuaciones (VI-16) y (VI-17) de la siguiente manera:

$$n = n_0 + \Delta n = n_i e^{\frac{E_{Fn} - E_{Fi}}{kT}} \quad (\text{VI-35})$$

$$p = p_0 + \Delta p = n_i e^{\frac{E_{Fi} - E_{Fp}}{kT}} \quad (\text{VI-36})$$

Se observa en estas ecuaciones que las concentraciones dependen de los cuasi niveles de Fermi.

Ejemplo VI-3: se tratará de ejemplificar las diferencias de niveles de Fermi entre sistemas en equilibrio y no equilibrio. Para ello se supondrá que se tiene Silicio a $T = 300\text{K}$ con un dopado, $N_d = 10^{14} \text{ 1/cm}^3$, $n_i = 1,5 \times 10^{10} \text{ 1/cm}^3$. El ancho de banda del Si, $E_G = 1,12\text{eV}$, por lo que $E_{Fi} = 0.56\text{eV}$. Si se supone que todos los átomos donores han sido ionizados, entonces $n_0 = 10^{14} \text{ 1/cm}^3$. De la ec. (VI-16) se puede determinar la relación de energías de Fermi, en la cual se reemplazarán los valores mencionados:

$$E_F = E_{Fi} + kT \ln\left(\frac{n_0}{n_i}\right) = 0.78893\text{eV}$$

Si se tiene en cuenta este valor, se puede hacer un diagrama de energías como el de la Figura VI-8 (líneas continuas).

Si utilizamos la Ley de Acción de Masas, se obtiene la densidad de portadores minoritarios:

$$p_0 = n_i^2 / n_0 = 2,25 \times 10^6 \text{ cm}^{-3}$$

Se introduce una perturbación, $\Delta n = \Delta p = 10^{11} \text{ cm}^{-3}$, se tiene:

$$n = n_0 + \Delta n \cong n$$

la cantidad de donores inyectados es mucho menor que los donores en equilibrio. Por otro lado,

$$p = p_0 + \Delta p \cong \Delta p$$

La cantidad de aceptores inyectados es mucho mayor que los que están en equilibrio. De la ecuación (VI-35) se puede obtener:

$$E_{Fn} = E_{Fi} + kT \ln\left(\frac{n}{n_i}\right) = 0,78895 eV$$

Mientras que de la ec.(VI-36)

$$E_{Fp} = E_{Fi} - kT \ln\left(\frac{p}{n_i}\right) = 0.5107 eV$$

Estos niveles pueden verse en la Figura VI-8 (líneas a trazos)

Se puede observar que la inyección de portadores no modifica significativamente el E_F , la diferencia entre este y E_{Fn} se encuentra en la 5ta cifra decimal ($2 \times 10^{-5} eV$), pero si lo ubica por sobre el nivel de Fermi del sistema en equilibrio. No ocurre lo mismo con el nivel de Fermi para los portadores minoritarios, E_{Fp} , este cambia en casi 0,01eV, ubicándose por debajo de E_{Fi} , acercándose a la banda de valencia.

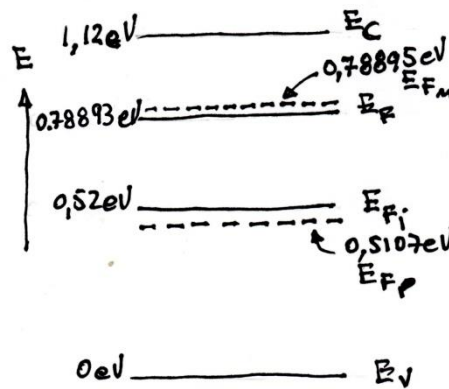


Figura VI-8. Niveles de Fermi, en equilibrio (líneas continuas), y en no equilibrio (líneas a trazos)

VI.5. Neutralidad de carga

Cuando el semiconductor se encuentra en equilibrio térmico, el cristal es eléctricamente neutro. Esta condición puede utilizarse para determinar las concentraciones de equilibrio de huecos y electrones en función del dopado.

La neutralidad de carga implica que no haya ningún exceso de una carga con respecto a la otra, es decir que las cargas negativas y positivas se encuentren equiparadas. Esto se puede escribir de la siguiente manera:

$$n_0 + N_a^- = p_0 + N_d^+ \tag{VI-37}$$

donde n_0 y p_0 son las concentraciones en equilibrio térmico de electrones y huecos, respectivamente. Las cantidades N_a^- y N_d^+ son las densidades totales de átomos aceptores y donores ionizados, respectivamente.

Si se supone que se ha producido una ionización total de los dopantes, la ec. (VI-37) queda:

$$n_0 + N_a = p_0 + N_d \quad (\text{VI-38})$$

De acuerdo a la Ley de Acción de Masas, $p_0 = n_i^2/n_0$, reemplazando en la ec. (VI-38)

$$n_0 + N_a = \frac{n_i^2}{n_0} + N_d \quad (\text{VI-39})$$

La ec. (VI-39) se puede escribir como una ecuación de segundo grado,

$$n_0^2 - (N_d - N_a)n_0 - n_i^2 = 0 \quad (\text{VI-40})$$

La solución es:

$$n_0 = \frac{(N_d - N_a)}{2} + \sqrt{\left(\frac{(N_d - N_a)}{2}\right)^2 + n_i^2} \quad (\text{VI-41})$$

La raíz cuadrada debería tener el doble signo “±”, pero el signo – no tiene significado físico aquí. Si $N_a = N_d$, la concentración de equilibrio térmico será $n_0 = n_i$.

Referencias

- Albella, J. M., Martínez-Duart, J. M. *Fundamentos de electrónica física y microelectrónica*. 1996 Addison-Wesley Universidad Autónoma de Madrid.
- Ashcroft, N. W., Mermin, N. D. *Solid State Physics*. 1981 Holt-Saunders International Editions.
- Kittel, C. *Introduction to Solid State Physics Sixth Edition* 1986 John Wiley and Sons Inc. New York.
- Pierret, F. R. *Advanced Semiconductor Fundamentals, Volume VI*. 1987 Modular Series on Solid State Devices, R.F Pierret and G. W. Neudeck, Editors. Addison-Wesley Pub. Comp.
- Sapoval, B., Hermann, G. *Physics of Semiconductors*. 1993 Springer-Verlag. New York.
- Singh, J. *Semiconductor Devices an Introduction*. 1994 Mc Graw Hill International Editions.
- Van der Ziel, A. *Electrónica Física del Estado Sólido*. 1968 Editorial Prentice/Hall Inc. Englewood Cliff, New Jersey.
- Yepifanov, G. *Physical Principles of Microelectronics*. 1974 Mir Publishers. Moscow.

CAPÍTULO VII

Conducción en semiconductores

Todo fluye, nada permanece.

HERÁCLITO DE ÉFESO
540 - 480AC

Cuando los electrones se encuentran en la banda de conducción, ellos conforman un gas. Si el sistema es no degenerado (ver sección VI.3), la estadística con la que se los puede estudiar es la de Maxwell-Boltzmann, si fuese degenerado habría que hacerlo con la de Fermi-Dirac.

Cuando se trata con un gas térmicamente excitado, la energía transmitida desde la red es $\frac{1}{2}kT$ por cada grado de libertad. Como el sistema que se propone es tridimensional, y la energía transmitida se transforma en energía cinética, se tendrá:

$$\frac{1}{2} m_e^* \langle v^2 \rangle = \frac{3}{2} kT \quad (\text{VII-1})$$

Donde m_e^* es la masa efectiva del electrón. Se puede despejar la velocidad de esta expresión y se puede ver que para la temperatura ambiente, $T = 300\text{K}$, se obtiene $\langle v \rangle \approx 10^5 \text{ m/s}$. La velocidad puede parecer elevada, pero hay que considerar que el movimiento térmico es totalmente desordenado, los electrones cambian permanentemente de dirección, si reemplazamos v por $-v$ la función permanece inalterada, esto quiere decir que la distribución de velocidades es simétrica con respecto al eje de las poblaciones (eje vertical).

Esto es lo que se muestra en la Figura VII-1, con línea llena, las poblaciones en ambos sentidos de la velocidad con respecto al eje f , son iguales en ambos lados, esto indica que la velocidad promedio es cero, no hay ninguna corriente en el material.

Para que haya una corriente, se debe romper la simetría que muestra el sistema. La ruptura se puede lograr con la aplicación de un campo externo como se ve en la Figura VII-1, el cual mueve a los portadores en una u otra dirección. Se debe recordar que en los

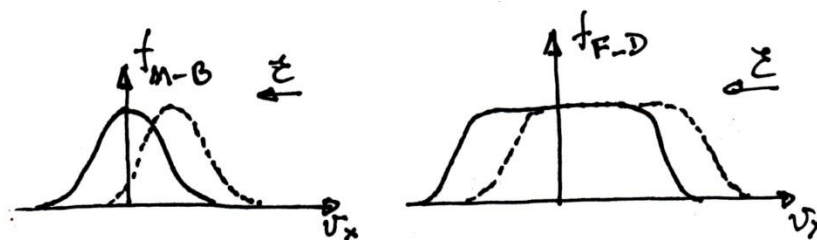


Figura VII-1. Distribuciones de velocidades de Maxwell-Boltzmann y de Fermi-Dirac, con línea llena sin campo aplicado, línea punteada, con campo aplicado.

semiconductores tenemos portadores de carga negativa y positiva.

La aplicación de un campo no es el único mecanismo para establecer una corriente. Dentro del semiconductor se pueden generar variaciones locales de la concentración de portadores, por ejemplo: mediante el dopado. Esto hace que se ponga en funcionamiento otro mecanismo de desplazamiento de portadores, la difusión.

VII.1. Movimiento de portadores. Aplicación de campo externo

La presencia de un campo eléctrico externo sobre el gas de electrones existente en la banda de conducción hace que exista una corriente eléctrica. El campo aplicado rompe la simetría de las distribuciones electrónicas, como muestra la línea punteada en la Figura VII-1, la cual describe esquemáticamente la desigualdad de las poblaciones con respecto al origen de coordenadas, lo que marca la tendencia de generación de la corriente.

Al mencionar “rotura de la simetría”, estamos simplificando un proceso complejo de interacciones entre los electrones y las vibraciones de la red (fonones), las impurezas, la red cristalina, etc. En el desarrollo del texto, solamente cuando sea necesario se dará la explicación de estos procesos. La rigurosidad que se requiere para explicar a estos mecanismos, excede el alcance de este texto.

La percepción macroscópica que se tiene de algunos efectos en los materiales (por ejemplo la corriente eléctrica) es producida por fenómenos microscópicos. Cuando se mide una corriente con algún dispositivo, se está midiendo la suma de todos los efectos que sufren los portadores en su movimiento a través del material. El movimiento de los electrones dentro del cristal no es lineal, ellos van interactuando con el medio ambiente que los rodea de diferentes formas. Sí se puede decir que su movimiento promedio, la que será llamada velocidad de arrastre, \vec{v}_d , es en la dirección opuesta al campo.

Se utilizarán algunas aproximaciones para el desarrollo. Se supone que los electrones no interactúan entre ellos (aproximación de Drude). Debido a que los portadores interactúan con átomos, impurezas, dislocaciones, etc., se debe tener alguna cantidad de referencia que indique cuán frecuentes, en tiempo y espacio son sus interacciones en el camino que realizan dentro del material. Para ello se determinan dos parámetros. Uno de ellos es el “camino libre medio”, este indica la distancia que recorren entre colisiones sucesivas. Entre colisiones los electrones se mueven como si fuesen partículas libres,

$$\hbar d\vec{k}/dt = q\vec{\mathcal{E}}$$

El otro es el “tiempo de vida media, τ ”, que indica el tiempo entre colisiones. Ambos varían de acuerdo al material, al dopado existente, al tipo de portador que se mueve, etc. Una colisión significa una interacción del portador con alguna de las perturbaciones mencionadas.

Cuando un campo externo de baja intensidad es aplicado, dentro del cristal se establecerá una corriente eléctrica de densidad j , llamada densidad de corriente de arrastre, la cual se relaciona con el campo por medio de la ley de Ohm,

$$\vec{J}_{arr} = \sigma \vec{\mathcal{E}} \quad (\text{VII-2})$$

En el segundo miembro, multiplicando al campo, está el factor de proporcionalidad σ que es llamado la conductividad de un conductor. La inversa de la conductividad es la resistividad del mismo conductor,

$$\rho = 1/\sigma \quad (\text{VII-3})$$

La existencia de un flujo de portadores o corriente es una indicación de que el campo que se ha aplicado ha sacado a los electrones de su estado original.

Se puede determinar la velocidad promedio de un portador por la acción de un campo. La relación entre el campo aplicado y la fuerza ejercida sobre el electrón, de carga q , será:

$$\vec{F} = m_n^* d\vec{v}/dt = -q\vec{\mathcal{E}} \quad (\text{VII-4})$$

Aquí la velocidad es debida al campo eléctrico y no a efectos térmicos. Si se despeja v de la ec. (VII-4), se tiene:

$$\vec{v} = -\int_0^t \frac{q\vec{\mathcal{E}}}{m_n^*} dt = -\frac{qt}{m_n^*} \vec{\mathcal{E}} \quad (\text{VII-5})$$

Se debe entender que esta velocidad representa la velocidad de arrastre, \vec{v}_d , que se ha mencionado previamente. Entre colisiones, durante el tiempo τ , es el momento en que el electrón gana velocidad. Como la densidad de corriente es la carga transportada por todos los portadores, multiplicada por la velocidad promedio de estos, se puede escribir:

$$\vec{J}_{arr} = -nq\vec{v}_d = \frac{nq^2\tau}{m_n^*} \vec{\mathcal{E}} \quad (\text{VII-6})$$

Si se compara la ec. (VII-2) con la ec. (VII-6) se ve que la conductividad se puede escribir:

$$\sigma = nq^2\tau/m_n^* \quad (\text{VII-7})$$

Por otro lado, en la ec. (VII-5) se puede ver que entre la velocidad de arrastre y el campo existe también un factor de proporcionalidad, el que es llamado movilidad μ , cuyo valor es:

$$\mu = q\tau/m_n^* \quad (\text{VII-8})$$

Por lo tanto,

$$\vec{v}_d = -\mu\vec{\mathcal{E}} \quad (\text{VII-9})$$

Las unidades de μ en el sistemas MKS son m^2/Vs .

Las ecuaciones (VII-7) y (VII-8) pueden ser relacionadas,

$$\sigma = nq\mu \quad (\text{VII-10})$$

La movilidad para los electrones es negativa y para los huecos es positiva. Se ve en la ec. (VII-10) que la obtención de la conductividad se reduce a determinar concentración y movilidad de los portadores. Si se está estudiando un semiconductor donde se encuentren ambas concentraciones de portadores, n electrones y p huecos, la ec. (VII-10) se puede generalizar a:

$$\sigma = q(n\mu_n + p\mu_p) \quad (\text{VII-11})$$

Ejemplo VII-1: Un semiconductor de silicio, dopado con $N_d = 10^{16} \text{ cm}^{-3}$ y $N_a = 10^{15} \text{ cm}^{-3}$ se encuentra bajo la acción de un campo de $\mathcal{E} = 10 \text{ V/cm}$. Esto, generará una corriente a través del material, se desea determinar el valor de la corriente.

Datos útiles:

$$\mu_n = 1350 \text{ cm}^2/\text{Vs}$$

$$\mu_p = 480 \text{ cm}^2/\text{Vs}$$

$$n_i = 1,5 \times 10^{10} \text{ cm}^{-3}$$

$$q = 1,6 \times 10^{-19} \text{ C}$$

De acuerdo a la ec. (VI-39) se puede determinar la concentración de portadores

$$n = 9,0 \times 10^{15} \text{ cm}^{-3}$$

Utilizando la Ley de Acción de Masas podemos determinar la cantidad de portadores p

$$p = \frac{n_i^2}{n} = \frac{(1,5 \times 10^{10})^2}{9,0 \times 10^{15}} = 2,5 \times 10^4 \text{ cm}^{-3}$$

La diferencia existente entre estos dos valores determina que el material es tipo n , los portadores mayoritarios son los electrones. Eso también se podría haber dicho si la diferencia entre los dopantes hubiera sido mayor ($N_d \gg N_a$) que la propuesta aquí.

De acuerdo a las ec. (VII-2) y (VII-11), la corriente se puede escribir

$$j_{arr} = q(n\mu_n + p\mu_p)\mathcal{E} = 19,44 \text{ A/cm}^2$$

Frente a la poca incidencia que tienen los huecos en la corriente, la ecuación anterior se podría haber escrito:

$$j_{arr} = q(n\mu_n)\mathcal{E} = 19,44 \text{ A/cm}^2$$

Se ve que no existe diferencia en el resultado, los portadores mayoritarios son los que conforman la corriente de arrastre que existe en el semiconductor.

No obstante se puede demostrar que aunque la cantidad de aceptores propuesta en el ejemplo es pequeña, su participación en la corriente tiene peso. Si se supone ahora, que $N_a = 0$, no hay dopado de aceptores, se debería notar la diferencia con respecto al caso anterior. Se pueden utilizar las mismas expresiones anteriores, se tendrá:

$$n = 1,0 \times 10^{16} \text{ cm}^{-3}, \text{ y}$$

$$p = \frac{n_i^2}{n} = \frac{(1,5 \times 10^{10})^2}{1,0 \times 10^{16}} = 2,25 \times 10^4 \text{ cm}^{-3}$$

La corriente será: $j_{arr} = q(n\mu_n)\mathcal{E} = 21,60 \text{ A/cm}^2$

La diferencia de corriente de arrastre entre ambos casos es $\Delta j = 2,16 \text{ A/cm}^2$. Como estamos en un semiconductor tipo n , mayor será al corriente de electrones cuanto menor sea el dopado aceptor que se tenga. Se podría ver cuál sería la cantidad de portadores que generarían esa corriente, eso sería:

$$\Delta j_{arr} = q\Delta n\mu_n\mathcal{E}$$

De esta expresión se puede despejar n . Haciendo las cuentas, se tiene que $\Delta n = 10^{15} \text{ cm}^{-3}$.

Esta cantidad es exactamente el dopado N_a de la primera parte del ejemplo, es la cantidad de electrones que se quitan de la corriente por recombinación con los huecos provistos en el dopado de aceptores.

Hasta ahora se ha hablado del movimiento de portadores en un semiconductor, sin tener en cuenta los mecanismos que afectan su movimiento. Principalmente se pueden mencionar dos que revisten gran importancia, la dispersión con la red o con fonones, y la dispersión con impurezas ionizadas. Estos dos mecanismos actúan sobre la movilidad. Si bien no se profundizará las explicaciones sobre estos, se mencionará la forma en que la afectan. Se ha visto que los átomos vibran alrededor de sus posiciones de equilibrio. Las vibraciones térmicas producen la interacción entre los electrones o huecos y estos átomos. Estas son las dispersiones a través de la red o fonones. Por lo tanto, la temperatura hace que la movilidad se modifique. Esta movilidad, si la única dispersión existente fuera la debida a la red (*Lattice* en inglés), será:

$$\mu_L \propto T^{-3/2} \tag{VII-12}$$

Esta es la movilidad debida a procesos de dispersión en la red, se ve que a medida que decrecen las vibraciones térmicas, la movilidad se incrementa.

El otro mecanismo que se mencionó fue el de la dispersión con impurezas ionizadas. A temperatura ambiente, las impurezas están ionizadas, la interacción electrostática de los electrones y huecos con las impurezas modifica su velocidad y trayectoria. La movilidad resultante como consecuencia de la interacción con impurezas ionizadas es llamada μ_I , y tiene la forma:

$$\mu_I \propto T^{3/2} / N_I \tag{VII-13}$$

Aquí, $N_I = N_d^+ + N_a^-$ es la totalidad de impurezas ionizadas. Si aumenta el número de impurezas ionizadas, aumenta el número de lugares para dispersar, se reduce el valor de la movilidad.

En una estructura cristalina ideal, los electrones se moverán mucho más libres que en una estructura real. Cada una de las perturbaciones que pueda haber en un cristal es una generadora de colisiones. Cada una de estas colisiones tiene tiempos característicos, τ_1, τ_2, τ_3 , etc. El tiempo equivalente será:

$$\frac{1}{\tau} = \frac{1}{\tau_1} + \frac{1}{\tau_2} + \dots \tag{VII-14}$$

Esto quiere decir que el resultado final no puede ser mejor que los valores parciales, la suma de las inversas de los tiempos de vida media significa que la vida media total disminuye. Para los metales esta es la llamada regla de Matthiensen. De la ec. (VII-8) se puede ver que la movilidad y los tiempos de vida media tienen una dependencia lineal. Se puede escribir:

$$\frac{1}{\mu} = \frac{1}{\mu_1} + \frac{1}{\mu_2} + \dots \tag{VII-15}$$

En nuestro caso, μ_I sería, por ejemplo, μ_L y μ_2 sería μ_s , se están considerando solamente, los dos mecanismos principales de colisión, μ es la movilidad neta.

Al principio de este módulo, se decía que si aplicáramos un potencial circulaba una corriente eléctrica de densidad j . Se puede pensar, con un modelo simplificado, que los electrones se mueven en un material de forma de un paralelepípedo recto y que la corriente atraviesa una sección de área A del paralelepípedo, perpendicular a su movimiento, como puede verse en la Figura (VII-2).

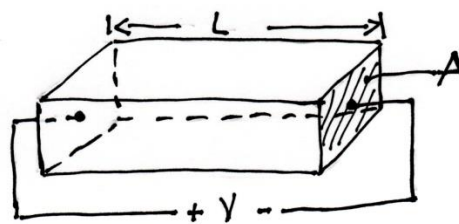


Figura VII-2. Barra de un semiconductor con potencial aplicado.

De acuerdo al modelo, se puede descomponer la densidad de corriente en flujo de corriente y área de la siguiente manera.

$$j_{arr} = i/A \tag{VII-16}$$

Por otro lado, podemos escribir al campo:

$$\mathcal{E} = V/L \tag{VII-17}$$

Donde V es el potencial aplicado al material y L su longitud como puede verse en la Figura VII-2. De acuerdo a la ec. (VII-2), la ec. (VII-16) se puede reescribir:

$$\frac{i}{A} = \sigma \frac{V}{L} \quad (\text{VII-18})$$

si de esta ecuación se despeja V , se tiene:

$$V = \frac{L}{\sigma A} i = \frac{\rho L}{A} i = iR \quad (\text{VII-19})$$

Esta es la Ley de Ohm, la resistencia es una función de la resistividad y de la geometría del material.

VII.2. Recombinaciones, generación

El hecho de que un electrón pueda abandonar su lugar original, por ejemplo en una ligadura, dejando un lugar vacío, un hueco, susceptible de ser ocupado por otro electrón, interactuando con el hueco, da lugar a la existencia de recombinaciones.

Existen varios tipos de recombinaciones, las cuales describiremos brevemente, pero, dentro de ellas también se debe mencionar que las hay radiativas, las cuales al saltar los electrones emiten radiación electromagnética, y no radiativas, donde no hay radiación involucrada.

Las imperfecciones que se puedan presentar en un cristal y los dopantes que se introduzcan, generan niveles de energía en la banda prohibida, estos pueden actuar como centros de captura de electrones y pueden ser profundos o poco profundos. Si el semiconductor fuera Si y la impureza fuera B (Boro), se ha visto en el parágrafo VI.1 que forma un nivel aceptor a 0.045eV sobre la banda de valencia, ese es un nivel poco profundo, en cambio si la impureza fuera oro (Au), el nivel aceptor que genera en la banda se encuentra aproximadamente a 0.54eV por debajo de la banda de conducción, casi en la mitad del gap del Si. Este es un nivel profundo. Con esto se puede ver, que aquellos niveles introducidos por impurezas y que se encuentren próximos al centro de la banda, se los denomina niveles profundos. Contrario a esto, cuando los niveles de energía generados por las impurezas son cercanos a las bandas, de valencia o conducción, a estos niveles se los llaman niveles poco profundos.

1. Recombinación banda a banda

Los electrones saltan directamente de la banda de conducción a alguna posición libre (hueco) en la banda de valencia, como se ve en la Figura VII-3.

En el salto de una banda a la otra existe la emisión de radiación, un fotón de energía $h\nu$.

Donde:

$$h\nu = E_C - E_V \quad (\text{VII-20})$$

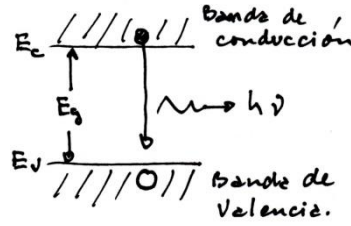


Figura VII-3. Esquema de la transición banda a banda.

Si en la transición se verifica que la cantidad de movimiento del electrón y el hueco son diferentes, la transición se llama indirecta; si esto no ocurre la transición será directa. Ejemplo del primero son las transiciones en el Si, el fondo de la banda de conducción no se enfrenta al tope de la banda de valencia (diferente k , diferente cantidad de movimiento). Como se ve en la Figura VII-4 (a). En el caso del GaAs, el proceso es distinto, este corresponde al caso directo, el fondo de la banda de conducción y tope de la de valencia están alineados (mismo k , misma cantidad de movimiento), como se ve en la Figura VII-4 (b).

La transición que se produce en el GaAs produce la emisión de un fotón. Este material por tener gap directo es muy utilizado en dispositivos optoelectrónicos.

Las transiciones no radiativas involucran vibraciones de la red llamadas fonones. En general, este tipo de transiciones se da en semiconductores con bandas prohibidas amplias. En semiconductores con grandes anchos de bandas prohibidas, los procesos multifonónicos son poco probables, estos decrecen cuando aumenta el gap de energía. Pero esta hipótesis se opone a la realidad, pues se ve que a medida que aumenta el ancho del gap, aumenta la cantidad de transiciones fonónicas. Esta diferencia queda zanjada cuando se muestra que al aumentar el gap, aumenta la probabilidad de transmisión fonónica a través de centros de recombinación en la banda prohibida.

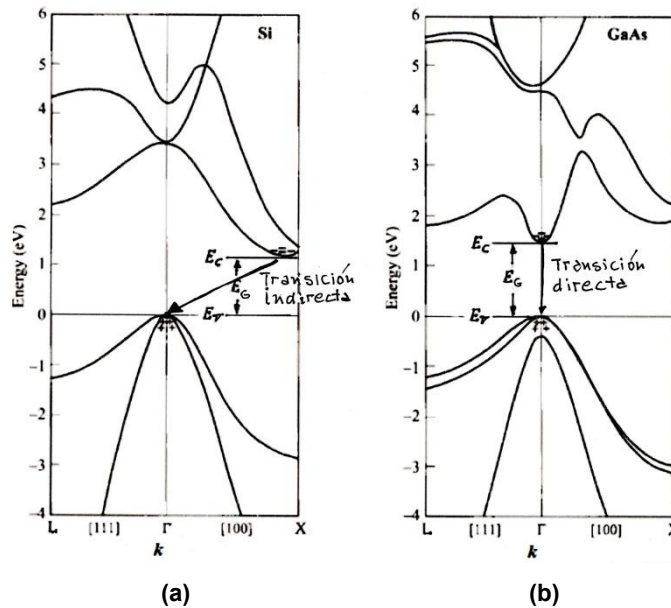


Figura VII-4. Transiciones indirecta (a) y directa (b), en Si y GaAs, respectivamente.

En el caso de una transición indirecta (emisión en este caso) un electrón va desde un estado en la banda de conducción (E_C) hasta uno en la banda de valencia (E_V), con la emisión de un fonón de energía $\hbar\Omega$ y un fotón de energía $\hbar\omega$, como se ve esquemáticamente en la Figura VII-5.

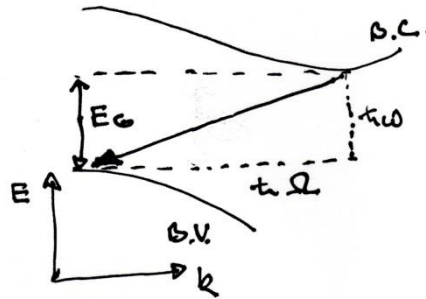


Figura VII-5. Esquema de transición indirecta.

Por lo tanto, de acuerdo a lo mencionado se tiene:

$$E_G = \hbar\Omega + \hbar\omega \tag{VII-21}$$

2. Recombinación mediante centros de captura. Teoría de Shockley-Read-Hall

Para desarrollar la teoría de Shockley-Read-Hall se puede suponer que se tiene un semiconductor tipo n (para un semiconductor p el tratamiento sería idéntico) con un nivel profundo de trampas neutras o centros de recombinación dentro de la banda prohibida con energía E_{CR} , con iguales probabilidades de capturar electrones o huecos. Esta teoría establece que en el centro de recombinación pueden ocurrir diferentes procesos, como se muestra esquemáticamente en la Figura VII-6.

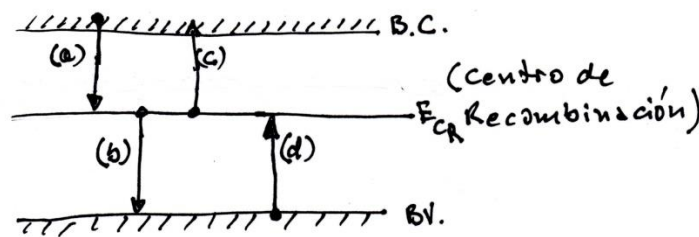


Figura VII-6. Esquema de recombinaciones de Shockley-Read-Hall.

- a) Una trampa captura un electrón de la banda de conducción, quedando negativamente cargada.
- b) Desde la trampa el electrón salta a la banda de valencia; se puede pensar que la trampa captura un hueco de la banda de valencia.
- c) Se puede producir un salto inverso (transporte térmico inverso) debido a que el nivel se encuentra saturado, el electrón puede saltar a la banda de conducción, liberando un electrón en esta banda.

- d) Se puede producir, si la concentración de electrones es pequeña, la captura por los centros de recombinación de electrones de la banda de valencia, o se puede pensar en la captura de huecos por parte de la banda de valencia.

La tasa de recombinación de electrones y huecos se determina por:

$$R_{n,p} = \frac{C_n C_p N_{CR} (np - n_i^2)}{C_n \left(n + N_C e^{\left[\frac{-(E_C - E_{CR})}{kT} \right]} \right) + C_p \left(p + N_V e^{\left[\frac{-(E_{CR} - E_V)}{kT} \right]} \right)} \quad (\text{VII-22})$$

En la ecuación anterior:

C_n es una constante proporcional a la probabilidad de que un electrón sea capturado por una trampa.

C_p es una constante proporcional a la probabilidad de que un hueco sea capturado por una trampa.

N_{CR} es la densidad de trampas.

E_{CR} es la energía del nivel del centro de recombinación (trampas, centros de impurezas).

Si el semiconductor se encuentra en equilibrio térmico, $np = n_i^2$, la ec. (VII-22) se anula, esto significa que esta ecuación da la relación de recombinación de electrones y huecos en exceso.

3. Recombinación mediante niveles poco profundos

Cuando los niveles de las trampas se encuentran cercanos a la banda de conducción o de valencia, son como vimos mas arriba: niveles poco profundos. La probabilidad de una re-emisión desde estos niveles hacia sus propias bandas (transferencia inversa de electrones y huecos), es mayor que la probabilidad de captura de portadores que tienen ellos. Por este motivo, estos niveles son muy ineficientes en cuanto a la mecánica de los centros de generación y recombinación de portadores.

El proceso en general produce un importante intercambio de electrones (o huecos) entre el nivel de las trampas y la banda de conducción (o de valencia). El esquema puede verse en la Figura VII-7.

Si hubiese una transferencia desde estos niveles hacia las bandas más alejadas y/o hacia los niveles de trampas opuestos, habría transiciones, que en general, podrían ser radiativas.

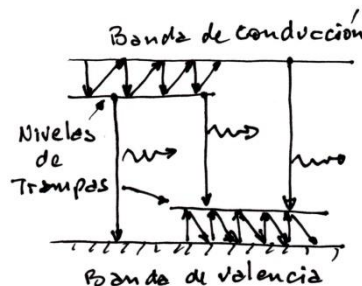


Figura VII-7. Esquema de transiciones en niveles poco profundos.

4. Recombinación mediante excitones

Desde el comienzo del texto se han mencionado a los electrones y a los huecos como partículas que pueden interactuar entre ellas recombinándose. No obstante esto, las interacciones coulombianas entre ambos portadores llevan a establecer un estado ligado, haciendo que se establezca un par electrón-hueco llamado excitón. Los excitones tienen tiempos de vida muy cortos, son del orden de los nano (10^{-9}) segundos. Su existencia es frecuente en procesos optoelectrónicos.

Los excitones establecen niveles de energía temporarios (ver Figura VII-8), muy próximos a las bandas de conducción y/o de valencia (niveles poco profundos). La recombinación del electrón y el hueco del excitón da lugar a un proceso radiativo. Este es uno de los mecanismos de producción de luz en los Light Emitting Diode (LED).

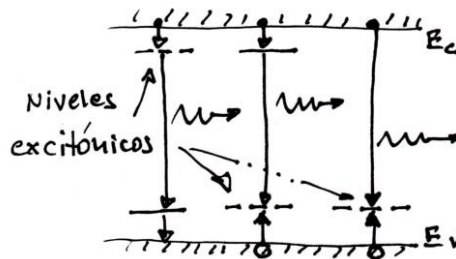


Figura VII-8. Esquema de transiciones en niveles excitónicos (indicados con línea punteada), los indicados con línea llena son niveles poco profundos permanentes.

5. Recombinación Auger

Hasta el momento, no se ha hablado de las energías puestas en juego durante la generación o recombinación de electrones y huecos. Por ejemplo, en un proceso como el descrito recientemente, la energía de ligadura en un excitón es de unos 5meV.

La recombinación Auger es un proceso en el que toman parte más de un electrón y no hay emisión de radiación. Mediante un choque, un electrón le transfiere energía a otro electrón, el primero se recombina con un hueco. El segundo electrón queda en un estado excitado por sobre los restantes portadores. Necesariamente, como está en un estado inestable, debe perder energía para compatibilizar con los restantes portadores, esta compatibilización se llama "termalización". La pérdida de energía se realiza en un proceso tipo cascada a través de choques con la red de átomos del semiconductor (Ver Figura VII-9). El número de colisiones aumenta si aumenta la densidad de portadores. Este tipo de recombinaciones es un mecanismo importante en semiconductores altamente dopados, el incremento de las recombinaciones es proporcional al cuadrado de la densidad de portadores, por ejemplo: uniones lasers.

El mecanismo de recombinación electrón-hueco será a través de centros de recombinación. En la Figura VII-9 se encuentra la descripción esquemática gráfica del proceso Auger.

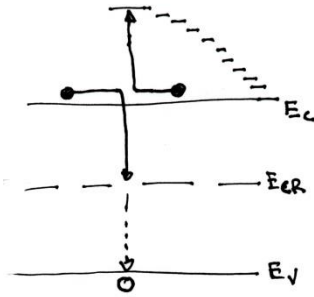


Figura VII-9. Proceso Auger.

6. Generación

Se han presentado diferentes casos de recombinación de portadores, el proceso inverso del cual no se ha hablado hasta el momento es la generación de portadores. Si se toman los mecanismos inversos a la recombinación se tiene la generación, esto quiere decir que se puede plantear la generación banda a banda, etc.

Los procesos de generación están ligados a los fenómenos que los producen. Si es por medio de energía térmica, la generación será térmica. Si el mecanismo es la excitación por medio de radiación lumínica se tendrán procesos de fotogeneración. El proceso inverso de la recombinación Auger es la ionización por impacto, a este se lo verá más adelante.

VII.3. Difusión. Relación de Einstein

La difusión es un proceso por el cual se tiende a uniformizar el medio, las partículas se mueven en forma aleatoria en un proceso donde la temperatura juega un papel preponderante. Esto es debido a que la naturaleza, sin mediar ningún contenedor, no puede mantener una concentración de partículas en un lugar. Esta concentración debe ser difundida de manera tal que después de cierto tiempo, el medio sea uniforme. El proceso de difusión es irreversible.

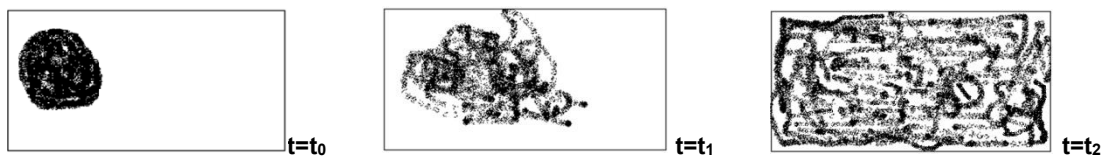


Figura VII-10. Esquema del proceso de difusión.

En la Figura VII-10, se muestra un breve esquema del proceso de difusión, al principio, $t = t_0$ se encuentra un gran cantidad de partículas acumuladas en una región del espacio. Transcurrida una cierta cantidad de tiempo ($t_1 > t_0$) el gas de partículas ha comenzado a difundirse ocupando una gran parte del medioambiente. Habiendo pasado más tiempo ($t_2 > t_1$) el gas prácticamente ha ocupado todo el volumen, y el proceso de difusión está por llegar a su fin. El fin será una distribución uniforme de partículas a través de todo el volumen.

El fenómeno de difusión es muy común en la vida cotidiana, cuando un globo de gas se desinfla, está presente la difusión. Existe dentro del globo una concentración muy grande de moléculas de un gas, contenidas por la delgada pared del globo, debido a las diferencias de presiones dentro y fuera, las moléculas empiezan a escapar a través de los intersticios de la pared, y el globo se desinfla. Cuando se deja caer una gota de alguna tintura en un volumen de agua, primero la gota está concentrada en algún lugar del volumen, tiempo después todo el volumen tiene el color de la tintura.

Si se tuviera una concentración grande de portadores, por ejemplo electrones, se puede imaginar que en dos puntos próximos del espacio las concentraciones serán diferentes, como los electrones transportan carga, y se están moviendo para uniformizar el medio, esto hace que se genere una corriente de difusión,

$$j_{dif} \sim -q \left(\frac{n_2 - n_1}{x_2 - x_1} \right) = -q \frac{dn}{dx} \quad (\text{VII-23})$$

La expresión anterior indica que la corriente de difusión se forma con la diferencia entre dos puntos (genéricos en este caso) 1 y 2. Esa definición no es nada menos que la definición de la derivada. La expresión correcta tendría que contemplar, por el tipo de fenómeno que se estudia, el gradiente de densidad, por simplicidad se considerará el problema en forma unidimensional (dirección x solamente), pero no se debe olvidar que n también depende del tiempo.

En la ec. (VII-23) se ha puesto que j_{dif} es proporcional a la derivada. Para que la proporcionalidad pase a ser una igualdad, se debería agregar un parámetro, coeficiente de difusión, que relacione a ambos miembros de la ec. (VII-23). El fenómeno de difusión está ligado a procesos térmicos, por lo tanto este coeficiente deberá tener en cuenta ese factor. También se deberá tener presente alguna característica del material donde se produce la difusión. Además el coeficiente de difusión debería indicar a que portadores se está refiriendo. En este caso como estamos tratando con electrones se denominará D_n . Si se hubieran propuesto huecos el coeficiente sería D_p . Por lo tanto la ecuación (VII-23) quedará:

$$j_{dif} = -qD_n \frac{dn}{dx} \quad (\text{VII-24})$$

Se debería ahora tratar de determinar el valor de D_n .

El movimiento de portadores dentro del semiconductor se debe a dos efectos diferentes: la aplicación de un campo eléctrico y la difusión. Bajo condiciones de equilibrio, no debería haber ninguna corriente en el material, es decir, la densidad de corriente de electrones debería anularse (lo mismo debería pasar para la densidad de corriente para huecos):

$$j_n = j_{n,arr} + j_{n,dif} = q\mu_n n \mathcal{E} + qD_n \frac{dn}{dx} = 0 \quad (\text{VII-25})$$

Por otro lado se tiene:

$$\mathcal{E} = -\frac{dV}{dx} = \frac{1}{q} \frac{dE_C}{dx} \quad (\text{VII-26})$$

$$\text{y,} \quad n = N_C e^{-\frac{(E_C - E_F)}{kT}} = N_C e^{-\frac{(E_F - E_C)}{kT}} \quad (\text{VII-27})$$

si se sustituye en la ec. (VII-27), $(E_F - E_C)/kT$ por B, la derivada de n será,

$$\frac{dn}{dx} = N_C e^B \frac{dB}{dx} = -N_C e^B \frac{1}{kT} \frac{dE_C}{dx} = -n \frac{1}{kT} \frac{dE_C}{dx} = -n \frac{1}{kT} q\mathcal{E} \quad (\text{VII-28})$$

En la ec. (VII-28) se ha tenido en cuenta la ec. (VII-26). Reemplazando en la ec. (VII-25) queda:

$$q\mu_n n \mathcal{E} - \frac{q^2}{kT} n D_n \mathcal{E} = q\mathcal{E} n \left(\mu_n - \frac{q}{kT} D_n \right) = 0 \quad (\text{VII-29})$$

Los valores que multiplican al paréntesis, ninguno es nulo, por lo tanto la que debe ser nula es la operación indicada dentro del paréntesis, es decir:

$$\mu_n - \frac{q}{kT} D_n = 0 \quad (\text{VII-30})$$

Es decir:

$$\frac{D_n}{\mu_n} = \frac{kT}{q} \quad (\text{VII-31})$$

Esta relación nos da el coeficiente de difusión en función de la temperatura y de la movilidad de los electrones en el semiconductor. Si los portadores hubieran sido huecos la relación sería:

$$\frac{D_p}{\mu_p} = \frac{kT}{q} \quad (\text{VII-32})$$

Con lo cual, el coeficiente de la expresión dada por la ec. (VII-24) queda determinado.

A las ecuaciones (VII-31) o (VII-32) se las llama "Relación de Einstein". Se puede demostrar que estas son válidas aun cuando el sistema se encuentre fuera del equilibrio.

Ejemplo VII-2: se puede calcular el coeficiente de difusión para el Silicio a $T = 300\text{K}$, la movilidad para electrones y huecos es, $\mu_n = 1350 \text{ cm}^2/\text{Vs}$ y $\mu_p = 480 \text{ cm}^2/\text{Vs}$, respectivamente, $q = 1,602 \times 10^{-19} \text{C}$, $k = 8.617 \times 10^{-5} \text{eV/K}$. Con estos valores se obtiene que los coeficientes de difusión son: $D_n = 35 \text{ cm}^2/\text{s}$ y $D_p = 12,4 \text{ cm}^2/\text{s}$, para electrones y huecos, respectivamente.

VII.4. Ecuación de continuidad

El conocimiento de las condiciones estáticas y dinámicas de los semiconductores es muy importante. La ecuación de densidad de corriente para electrones, ec. (VII-25), y la correspondiente para los huecos, generadas por la existencia de campos aplicados o fenómenos de difusión, causan desviaciones de las condiciones de equilibrio térmico. A estas se les debe sumar los diferentes mecanismos de generación y recombinación, los cuales también producen desviaciones ya que cambian las poblaciones de portadores de las corrientes mencionadas cambiando las densidades de carga transportada. Los cambios producidos se pueden ver tanto desde el punto de vista temporal como espacial.

Debido a la necesidad de conocer las variaciones temporales y espaciales de las densidades de carga de los portadores en exceso se necesita desarrollar una ecuación que contemple este fenómeno, esa ecuación es la de continuidad. Estas ecuaciones describen el comportamiento del exceso de electrones y huecos. Cuando se realiza el balance total de los portadores dentro del semiconductor se debe tener en cuenta que está formado por los portadores moviéndose, debido al arrastre y la difusión, menos aquellos que son restados de ese flujo por las recombinaciones.

Las ecuaciones que se desarrollarán son equivalentes para electrones y huecos. Las densidades de portadores son una función que depende del espacio y del tiempo, es decir $n = n(x,y,z,t)$ y $p = p(x,y,z,t)$. Por lo tanto, variación temporal se puede escribir como:

$$\frac{\partial n}{\partial t} = \sum_{\text{todas las variaciones}} \frac{\partial n}{\partial t} = \frac{\partial n}{\partial t_{arr}} + \frac{\partial n}{\partial t_{dif}} + \frac{\partial n}{\partial t_{R-G}} + \frac{\partial n}{\partial t_{\text{otros procesos}}} \quad (\text{VII-33})$$

$$\frac{\partial p}{\partial t} = \sum_{\text{todas las variaciones}} \frac{\partial p}{\partial t} = \frac{\partial p}{\partial t_{arr}} + \frac{\partial p}{\partial t_{dif}} + \frac{\partial p}{\partial t_{R-G}} + \frac{\partial p}{\partial t_{\text{otros procesos}}} \quad (\text{VII-34})$$

Los dos últimos términos de ambas ecuaciones, que indican *otros procesos*, están dirigidos a procesos como la fotogeneración u otro mecanismo que pueda participar en la generación o recombinación de portadores; a estos términos se los llamará G_L . La ec. (VII-33) queda (el procedimiento para la ec. (VII-34) es idéntico)

$$\frac{\partial n}{\partial t} = \frac{\partial n}{\partial t_{arr}} + \frac{\partial n}{\partial t_{dif}} + \frac{\partial n}{\partial t_{R-G}} + G_L \quad (\text{VII-35})$$

Los términos de arrastre y difusión se pueden reescribir imaginando un elemento de volumen de área unitaria y longitud dx en un semiconductor, como se ve en la Figura VII-11.

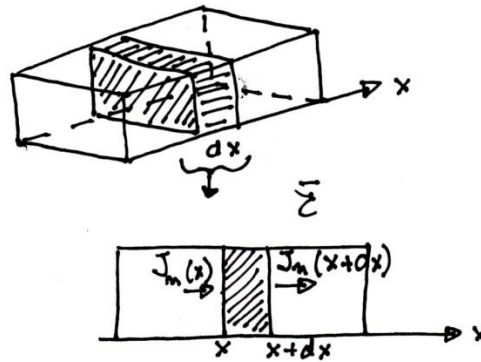


Figura VII-11. Diagrama de términos de la ecuación de continuidad.

Se debe pensar que la concentración de portadores tiene un determinado valor al tiempo t y un valor diferente cuando ha transcurrido un dt , con esto se quiere decir que en el elemento de volumen la cantidad de carga existente en el lapso dt , será:

$$[n(x, t + dt) - n(x, t)]dx = \frac{\partial n}{\partial t} dt dx \tag{VII-36}$$

La ecuación (VII-36) indica que la variación de portadores en el volumen de la Figura VII-11 está dada por los procesos de generación y recombinación; también de acuerdo a las corrientes de difusión y arrastre. Se puede escribir:

$$j_n(x) - j_n(x + dx) = -\frac{\partial j}{\partial x} dx \tag{VII-37}$$

La ecuación anterior indica que la corriente que entra no es la misma que la que sale del elemento de volumen. La carga transportada cambia en el dx , debido a la generación y recombinación que ocurre en ese espacio. Si multiplicamos ambos miembros por dt queda:

$$[j_n(x) - j_n(x + dx)]dt = -\frac{\partial j}{\partial x} dx dt \tag{VII-38}$$

En esta última ecuación se ve el cambio en el número de electrones en la sección de volumen durante el período dt , el cual depende de las corrientes entrante y saliente del elemento de volumen.

Por lo tanto podemos sustituir los dos primeros términos del segundo miembro de la ec. (VII-35) por:

$$\frac{\partial n}{\partial t_{arr}} + \frac{\partial n}{\partial t_{dif}} = \frac{1}{q} \frac{\partial j_n}{\partial x} dx dt \tag{VII-39}$$

Esta ecuación indica que una variación temporal de las densidades de portadores producirá un cambio en la corriente dentro y fuera del volumen seleccionado, la cual, en una dimensión, tendrá la forma, para electrones y huecos:

$$\frac{\partial n}{\partial t} = \frac{1}{q} \frac{\partial j_n}{\partial x} + \frac{\partial n}{\partial t_{R-G}} + G_L \quad (\text{VII-40})$$

$$\frac{\partial p}{\partial t} = -\frac{1}{q} \frac{\partial j_p}{\partial x} + \frac{\partial p}{\partial t_{R-G}} + G_L \quad (\text{VII-41})$$

En el proceso de recombinación, muchos electrones serán removidos del volumen unitario, por unidad de tiempo y de volumen dx . Por lo tanto, la cantidad de electrones removidos será:

$$\frac{\partial n}{\partial t_{R-G}} = -\frac{n-n_0}{\tau_n} dxdt \quad (\text{VII-42})$$

En la ecuación anterior n_0 es la concentración de electrones en equilibrio. $n = n_0 + \Delta n$, es la concentración de electrones bajo condiciones de no equilibrio, Δn son los cambios en las concentraciones de equilibrio debidos a la generación o recombinación, τ_n vida media del portador. Con lo cual las ecs. (VII-40) y (VII-41) se pueden escribir:

$$\frac{\partial n}{\partial t} = \frac{1}{q} \frac{\partial j_n}{\partial x} - \frac{\Delta n}{\tau_n} + G_L \quad (\text{VII-43})$$

$$\frac{\partial p}{\partial t} = -\frac{1}{q} \frac{\partial j_p}{\partial x} - \frac{\Delta p}{\tau_p} + G_L \quad (\text{VII-44})$$

Si se deseara escribir estas ecuaciones en 3 dimensiones, tomarían la forma:

$$\frac{\partial n}{\partial t} = \frac{1}{q} \nabla \cdot \vec{J}_n - \frac{\Delta n}{\tau_n} + G_L \quad (\text{VII-45})$$

$$\frac{\partial p}{\partial t} = -\frac{1}{q} \nabla \cdot \vec{J}_p - \frac{\Delta p}{\tau_p} + G_L \quad (\text{VII-46})$$

En una dimensión las ecuaciones (VII-43) y (VII-44) o en tres dimensiones las ecuaciones (VII-45) y (VII-46) conforman las ecuaciones de continuidad.

Si se tiene en cuenta la densidad de corriente de la ec. (VII-25) y la correspondiente para los huecos, las ec. (VII-43) y (VII-44), se pueden escribir:

$$\frac{\partial n}{\partial t} = D_n \frac{\partial^2 n}{\partial x^2} + \mathcal{E}_x \mu_n \frac{\partial n}{\partial x} - \frac{\Delta n}{\tau_n} + G_L \quad (\text{VII-47})$$

$$\frac{\partial p}{\partial t} = D_p \frac{\partial^2 p}{\partial x^2} - \mathcal{E}_x \mu_p \frac{\partial p}{\partial x} - \frac{\Delta p}{\tau_p} + G_L \quad (\text{VII-48})$$

Las cuales indican que las variaciones temporales de la densidad de portadores son debidas a las corrientes de difusión, las corrientes de arrastre, la generación y/o recombinación y eventualmente procesos lumínicos. Bajo condiciones estacionarias,

$$\frac{\partial n}{\partial t} = \frac{\partial p}{\partial t} = 0 \quad (\text{VII-49})$$

Las ecuaciones (VII-47) y (VII-48) quedan:

$$-\left(D_n \frac{\partial^2 n}{\partial x^2} + \mathcal{E}_x \mu_n \frac{\partial n}{\partial x}\right) = -\frac{\Delta n}{\tau_n} + G_L \quad (\text{VII-50})$$

$$-\left(D_p \frac{\partial^2 p}{\partial x^2} - \mathcal{E}_x \mu_p \frac{\partial p}{\partial x}\right) = -\frac{\Delta p}{\tau_p} + G_L$$

Estas indican que la cantidad de portadores dejando el volumen unitario es igual al número de partículas recombinadas en el volumen más aquellas generadas por algún factor ionizante (G_L).

VII.5. Portadores minoritarios

Como se verá más adelante, cuando se comience a ver el desarrollo de las junturas, los portadores minoritarios adquieren una relevancia fundamental en el funcionamiento de los dispositivos electrónicos. Las ecuaciones de continuidad y su aplicación a los portadores minoritarios son ampliamente utilizadas para el estudio de dispositivos electrónicos. Para ese fin se harán algunas suposiciones:

- 1) El análisis será unidimensional.
- 2) Se hará el estudio sobre los portadores minoritarios.
- 3) En las zonas volumétricas se considerará que el campo eléctrico es nulo, $\mathcal{E} = 0$.
- 4) Las concentraciones de portadores en equilibrio no son función de la posición, no dependen de la variable x , $n_0 \neq n_0(x)$ y $p_0 \neq p_0(x)$.
- 5) Se supondrá que se tiene un bajo nivel de inyección, es decir, la perturbación que se introduzca es mucho menor que el nivel de equilibrio de los portadores,
 $\Delta p \ll n_0$ y $\Delta n \ll p_0$.
- 6) La única fuente externa será la fotogeneración, si la hubiera.

En este módulo se desarrollarán las ecuaciones para los electrones, el tratamiento para los huecos es absolutamente similar.

Si se parte de la ecuación (VII-25) y se aplica la condición 3) recién enumerada, se puede escribir:

$$j_n = qD_n \frac{dn}{dx} \quad (\text{VII-51})$$

Esto se puede justificar en función de la aproximación, si no existe el campo aplicado, tampoco existe la corriente de arrastre. Por otro lado, la concentración de portadores minoritarios es pequeña y el campo es pequeño, por lo que el producto $n\mathcal{E}$ es pequeño.

Se ha dicho que n_0 es constante (no depende de x), y además, como $n = n_0 + \Delta n$, entonces, la dn/dx de la ec. (VII-51) se puede escribir (considerando que n es una función del espacio y del tiempo):

$$\frac{\partial n}{\partial x} = \frac{\partial n_0}{\partial x} + \frac{\partial \Delta n}{\partial x} = \frac{\partial \Delta n}{\partial x} \quad (\text{VII-52})$$

Utilizando este resultado, la ec. (VII-51) queda:

$$j_n = qD_n \frac{d\Delta n}{dx} \quad (\text{VII-53})$$

La concentración de equilibrio de portadores no es una función del tiempo, $n_0 \neq n_0(t)$, por lo que podemos escribir:

$$\frac{\partial n}{\partial t} = \frac{\partial n_0}{\partial t} + \frac{\partial \Delta n}{\partial t} = \frac{\partial \Delta n}{\partial t} \quad (\text{VII-54})$$

Si sustituimos en las ecuaciones de continuidad (VII-47) y (VII-48) se tiene:

$$\frac{\partial \Delta n_p}{\partial t} = D_n \frac{\partial^2 \Delta n_p}{\partial x^2} - \frac{\Delta n_p}{\tau_n} + G_L \quad (\text{VII-55})$$

$$\frac{\partial \Delta p_n}{\partial t} = -D_p \frac{\partial^2 \Delta p_n}{\partial x^2} - \frac{\Delta p_n}{\tau_p} + G_L \quad (\text{VII-56})$$

En este caso G_L representa el número de pares electrón-hueco generados por unidad de tiempo y de volumen por la absorción de fotones.

VII.6. Longitud de difusión

Se puede tomar la ecuación (VII-55) y suponer que no existe excitación externa, $G_L = 0$, y que el proceso es estacionario, se tiene:

$$D_n \frac{\partial^2 \Delta n_p}{\partial x^2} = \frac{\Delta n_p}{\tau_n} \Rightarrow \frac{\partial^2 \Delta n_p}{\partial x^2} = \frac{\Delta n_p}{D_n \tau_n} \Rightarrow \frac{\partial^2 \Delta n_p}{\partial x^2} = \frac{\Delta n_p}{L_n^2} \quad (\text{VII-57})$$

El parámetro $L_n^2 = D_n \tau_n$ tiene unidades de longitud al cuadrado. L_n es llamado la longitud de difusión de los portadores minoritarios. La solución de la ecuación (VII-57) es del tipo:

$$\Delta n_p = Ae^{-x/L_n} + Be^{x/L_n} \quad (\text{VII-58})$$

Para valores muy grandes de x ($x \rightarrow \infty$) la concentración de portadores debe ser la de equilibrio, $\Delta n_p = 0$, mientras que en $x = 0$, $\Delta n_p = \Delta n_0$. Por otro lado, cuando x crece, el segundo término de (VII-58) crece exponencialmente no describiendo la situación física que se presenta (ver Figura VII-12), por lo tanto el coeficiente B debe ser nulo.

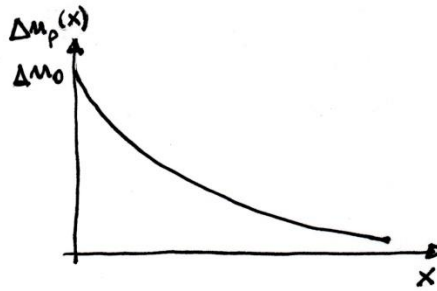


Figura VII-12. Longitud de difusión de portadores minoritarios.

Finalmente la solución tiene la forma:

$$\Delta n_p(x) = \Delta n_0 e^{-x/L_n} \quad (\text{VII-59})$$

Donde $L_n = \sqrt{D_n \tau_n}$
(VII-60)

Se puede escribir también,

$$L_p = \sqrt{D_p \tau_p} \quad (\text{VII-61})$$

A las longitudes de difusión de los portadores minoritarios, electrones o huecos, se las puede interpretar como la distancia media que difundirá un electrón antes de recombinarse con un hueco, o también, la distancia media que la concentración en exceso decrece al valor $1/e$ de su valor original, en este último caso la longitud es llamada la *Longitud de Debye*.

Ejemplo VII-3: se puede tomar Silicio a $T = 300\text{K}$ según el ejemplo del parágrafo VII.3 se tiene que $D_n = 35\text{cm}^2/\text{s}$ y que $D_p = 12,4\text{cm}^2/\text{s}$. Si se supone que la vida media es igual para electrones y huecos y es aproximadamente 1ns ($1 \times 10^{-9}\text{s}$), se encuentra que:

$$L_n = 1,87 \times 10^{-4}\text{cm} \text{ y } L_p = 1,11 \times 10^{-4}\text{cm}.$$

Referencias

- Albella, J. M., Martínez-Duart, J. M. *Fundamentos de electrónica física y microelectrónica*. 1996 Addison-Wesley Universidad Autónoma de Madrid.
- Ashcroft, N. W., Mermin, N. D. *Solid State Physics*. 1981 Holt-Saunders International Editions.
- Dalven, R. *Introduction to Applied Solid State Physics. Second Edition*. 1980 Plenum Press. New York and London.
- Ghandhi, S. K. *Semiconductor Power Devices*, 1977 Wiley, New York
- Kittel, C. *Introduction to Solid State Physics Sixth Edition* 1986 John Wiley and Sons Inc. New York.
- Hall, R. N. *Germanium rectifier characteristics*, Physical Review 83, 228(1951).
- Hall, R. N. *Electron-hole recombination in germanium*, Physical Review 87, 387(1952).
- Ihm, T. *Semiconductor Nanostructures Quantum States and Electronic Transport*. 2010 Oxford University Press, Great Clarendon Street, Oxford OX2 6DP
- Pierret, F. R. *Advanced Semiconductor Fundamentals, Volume VI*. 1987 Modular Series on Solid State Devices, R.F Pierret and G. W. Neudeck, Editors. Addison-Wesley Pub. Comp.
- Sapoval, B., Hermann, G. *Physics of Semiconductors*. 1993 Springer-Verlag. New York.
- Singh, J. *Semiconductor Devices an Introduction*. 1994 Mc Graw Hill International Editions.
- Shockley, W.S. Read, W. T. *Statistics of the recombination of holes and electrons*, Physical Review 87, 835(1952).
- Van der Ziel, A. *Electrónica Física del Estado Sólido*. 1968 Editorial Prentice/Hall International
- Yepifanov, G. *Physical Principles of Microelectronics*. 1974 Mir Publishers. Moscow.
- Yu, P. Y. Cardona, M. *Fundamentals of Semiconductors Physics and Material Properties*, 1996 Springer-Verlag Berlin Heidelberg

TERCERA PARTE

La aplicación

CAPÍTULO VIII

Junturas

*El que escucha estas palabras mías y las cumple,
es como el sabio que edifica su casa sobre piedra.
Cae la lluvia, vienen los torrentes, soplan los vientos
y se echan sobre esa casa; pero no cae, porque está cimentada sobre roca.
Y el que escucha estas palabras mías y no las cumple,
es como el necio, que edifica su casa sobre arena;
cae la lluvia, vienen los torrentes, soplan los vientos
y se echan sobre esa casa: y cae y es grande su ruina.*

MATEO 7:24-27 (año 80-90 DC)

Se ha visto en el capítulo precedente como se modifican las propiedades eléctricas de los semiconductores, cuando se introduce algún dopante. Para desarrollar un dispositivo electrónico no es suficiente tener un material dopado, se necesita establecer una unión de materiales con diferentes impurezas tal que sus comportamientos eléctricos sean diferentes. En este sentido, se pueden diferenciar dos clases de uniones, aquellas que ponen en contacto un mismo material con dopados diferentes, por ejemplo, Silicio dopado p de un lado y dopado n del otro. Esta clase de unión es llamada *homojuntura* (también *homounión*). Por otro lado, se pueden considerar la *heterojuntura* (heterounión) donde se ponen en contacto dos materiales diferentes con dopados diferentes, por ejemplo la unión metal-semiconductor, Tungsteno (W, Wolframio, $Z=74$)-Silicio con dopado tipo n (n -Si), o la unión de dos semiconductores p -GaAs en contacto con n -GaAlAs. También puede darse una heterounión entre dos semiconductores diferentes con el mismo tipo de dopante, pero en cantidades diferentes (n^+ -GaAs/ n -GaAlAs), el superíndice + en n^+ indica un cantidad de dopante es mayor que la que tiene un material con dopado n (sin el signo +).

En general se mencionará que los semiconductores están juntos o pegados, esto no se debe tomar literalmente, sino como una imagen, la unión se realiza de manera tal que los átomos de uno y otro material puedan “verse”. En el texto, no se avanzara sobre las técnicas experimentales con las que se realizan las uniones.

La gran mayoría de los dispositivos electrónicos tiene al menos una unión de materiales con dopados diferentes. Estos muestran comportamientos que permiten generar: rectificadores, amplificadores, llaves, etc. Por la importancia que las uniones tienen, es que se hace necesario tener un buen conocimiento de ellas y de su funcionamiento.

VIII.1. Homojunturas

Como se mencionó previamente una homojuntura es una unión de un mismo material con un dopado p de un lado y n del otro lado. Se supondrá que el dopado de ambos es equivalente. Las concentraciones de dopantes pueden tener dos diferentes características, pueden ser abruptas, en cuyo caso las concentraciones de donores y aceptores, N_d y N_a , son constantes. Si esto no ocurre y las concentraciones son funciones de la distancia, la unión se llama graduada. En la actualidad estas últimas no son las más utilizadas, por lo que se desarrollará el estudio sobre las uniones abruptas, que son las más utilizadas y las que pueden ser más fácilmente descriptas desde el punto de vista matemático.

VIII.1.a La unión de semiconductores tipo- p y tipo- n

Una homojuntura tipo $p-n$ se realiza poniendo en contacto un mismo material con dos dopados diferentes. Esquemáticamente la unión se muestra en la Figura VIII-1 allí se puede ver su conformación, una parte del conjunto está dopada con átomos aceptores, la región- p , mientras que la otra parte tiene átomos donores, la región- n . En medio de ellas hay una interface que separa ambas regiones, esta zona es la llamada unión metalúrgica. En esta estructura la unión metalúrgica definirá el origen de coordenadas de nuestro sistema.

El hecho de poner a estos semiconductores en contacto hace que comiencen a funcionar mecanismos que hemos descripto previamente. En la región n existe una gran concentración de electrones (portadores mayoritarios), lo que no ocurre, obviamente en la región p , donde los portadores mayoritarios son los huecos. Por este motivo empieza un proceso de difusión, haciendo que los electrones del lado n se muevan hacia el p . Como el semiconductor no tiene ninguna conexión externa, el proceso no puede durar indefinidamente. A medida que los electrones van ocupando lugares en la zona p , los átomos aceptores van siendo ionizados negativamente. Mientras que los átomos que abandonaron los

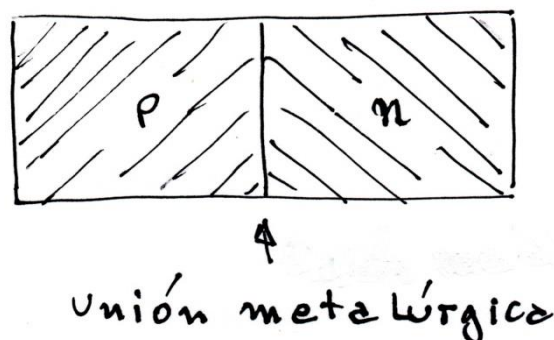


Figura VIII-1. Esquema de una unión $p-n$.

electrones para moverse a la zona p han quedado ionizados con una carga positiva. Esto hace que se genere un campo eléctrico, el cual se va intensificando a medida que los electrones van pasando de una región a la otra. La dirección del campo es desde la zona n a la p , como se

puede (Figura VIII-2), se va oponiendo en forma cada vez más intensa al pasaje de los electrones de n hacia p .

Se podría también ver este fenómeno desde el punto de vista de los huecos, se puede pensar que los huecos del lado p , donde hay una gran concentración de ellos, difunden hacia el lado n dejando átomos ionizados negativamente en el lado p e ionizando al lado n en forma positiva. Como puede verse en la Figura VIII-2.

Como se mencionó previamente, el proceso no puede seguir indefinidamente, en algún momento posterior a la unión de los materiales se llegará al equilibrio termodinámico, esto quiere decir que los electrones que abandonaron el lado n han dejado una zona cargada positivamente y se han movido hacia la zona p cargándola negativamente. Obviamente, el campo eléctrico generado ha llegado a su valor de equilibrio e impide que más electrones puedan pasar de la zona n a la p o huecos de la p a la n . El movimiento descrito ha sido realizado moviendo portadores mayoritarios, electrones del lado n al lado p y huecos del lado p al n . Por otro lado, este mismo campo mueve portadores minoritarios en sentido opuesto al mencionado, esto es, electrones desde el lado p hacia el lado n , y simultáneamente, huecos del lado n al p , claro está que este ya no es un proceso debido a la difusión, sino, como se mencionó, debido al campo existente.

El equilibrio termodinámico se establece cuando ambas corrientes, la de difusión, que mueve portadores mayoritarios, se iguala a la de portadores minoritarios arrastrados por el campo existente. Esto significa que la fuerza debida a la difusión sobre los portadores mayoritarios, electrones y huecos, es igual y de sentido opuesto a la fuerza debida al campo eléctrico sobre electrones y huecos. La corriente sobre la unión es cero.

A izquierda y derecha de la unión metalúrgica quedan zonas cargadas, las cuales han sido vaciadas de portadores móviles. Esta es la llamada zona de vaciamiento, en el tratamiento que se dará para hacer una descripción de este proceso, el campo existe solamente sobre esta zona.

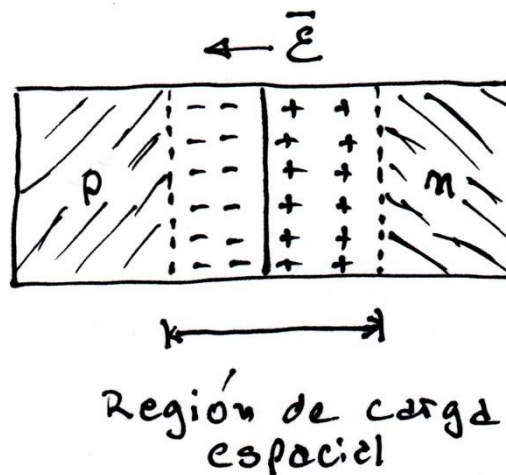


Figura VIII-2. Esquema de la región cargada positiva y negativamente.

Otra suposición que se debe hacer es que el sistema se encuentra en equilibrio térmico y no hay ningún potencial externo aplicado sobre la unión pn . En función de esta breve descripción, se ve que existen cantidades que deben ser determinadas para conocer en forma más profun-

da a la unión, para esto se debe conocer como dato previo, los dopados de cada región. Con esta información se podrá determinar la densidad de carga de cada signo en las inmediaciones de la unión metálica, el potencial generado por las cargas existentes, el campo eléctrico en esta región y los anchos de estas zonas que conforman la región de vaciamiento.

VIII.1.a1. El potencial de contacto, V_{bi}

Se ha visto en el capítulo VI que en los semiconductores dopados el Nivel de Fermi cambia su posición dentro de la banda prohibida. Se puede definir el Nivel de Fermi intrínseco (E_{Fi}) ubicado en la mitad de esta banda del semiconductor y además se tiene al Nivel de Fermi (E_F) que hace referencia al tipo de dopado del semiconductor, como se mencionó en el capítulo VI, y permanece constante a través del sistema. Cuando semiconductores con diferentes dopados se unen, ocurre el proceso mencionado más arriba, pero como tomamos como referencia al Nivel de Fermi, E_F , las bandas del semiconductor se deben curvar, modificándose como se muestra en la Figura VIII-3. El nivel relativo de las bandas de conducción y de valencia con respecto a E_F cambia dependiendo si se observa en la región p o n . Del lado de la región n el Nivel de Fermi se encontrará más cerca del fondo de la banda de conducción, debido a que en esta zona la población preponderante de portadores es la de electrones. Caso contrario es lo que ocurre en la zona p , aquí el Nivel de Fermi se encontrará más cerca de la parte superior de la banda de valencia debido a que son pocos los electrones en esta zona comparados con los huecos existentes. Si se toma la diferencia que hay entre los bordes de una y otra banda, valencia y conducción, en ambas regiones, se obtendrá el valor del potencial generado por la unión de estos materiales, el potencial de la unión o potencial de contacto V_{bi} . La curvatura que adquiera cada una de las bandas en la zona de vaciamiento, dependerá de cuán abrupta es la transición entre una región y la otra. Esto está relacionado con la cantidad de carga acumulada en la zona de vaciamiento.

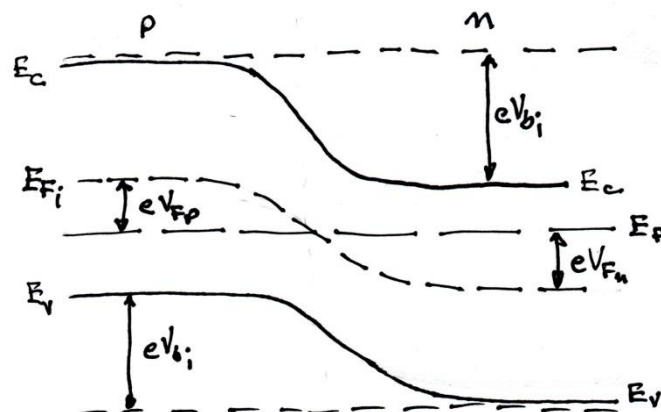


Figura VIII-3. Estructura de bandas de una unión pn .

De acuerdo a la ecuación de Poisson, se establece una relación entre el potencial y la densidad de carga como:

$$d^2V/dx^2 = -\rho(x)/K_s\epsilon_0$$

Se debe recordar que $\rho_i = qN_i$. Se puede relacionar al potencial con la energía,

$$V = E/q$$

Si se reemplaza en la ecuación anterior, se tiene:

$$d^2E/dx^2 = -q\rho(x)/K_s\epsilon_0$$

De aquí se puede establecer que a la izquierda de la unión metalúrgica la carga es de signo negativo, q representa a la carga del electrón, por lo que el producto del segundo miembro será negativo, eso indica una concavidad hacia abajo. Por otro lado, a la derecha de la unión, la densidad de carga es positiva, por lo que la derivada segunda indicará una concavidad hacia arriba. Esta descripción concuerda con las curvaturas mostradas en la Figura VIII-3.

El potencial de contacto (V_{bi}) formado es igual para electrones y huecos, es lo que ven estos portadores como barrera de potencial cuando tratan pasar de una región a la otra por efecto de la difusión.

Como se mencionó, el potencial de contacto se puede determinar mediante los valores de las bandas y el nivel de Fermi, de la siguiente manera:

$$qV_{bi} = (E_c^p - E_F) - (E_c^n - E_F) = E_c^p - E_c^n$$

Finalmente, el valor del potencial de contacto es la diferencia entre el valor del borde inferior de la banda de conducción en la región p y la región n , o la parte superior de la banda de valencia entre la región p y la n . Es decir

$$V_{bi} = -\frac{(E_c^n - E_c^p)}{q} = -\frac{(E_v^n - E_v^p)}{q} \quad \text{(VIII-1)}$$

También puede ser escrito como la suma de los potenciales con respecto a E_F , tomando a este como origen de energías, se tiene:

$$V_{bi} = V_{Fp} - V_{Fn} \quad \text{(VIII-2)}$$

La concentración de portadores minoritarios en cada una de las regiones son:

$$n_n = N_c e^{-(E_c^n - E_F)/kT} \quad \text{(VIII-3)}$$

$$p_p = N_v e^{-(E_F - E_v^p)/kT} \quad \text{(VIII-4)}$$

De acuerdo a la Ley de Acción de Masas,

$$n_i^2 = n_n p_n = N_v N_c e^{-(E_c^n - E_v^n)/kT} \quad (\text{VIII-5})$$

Donde n_n fue definido en la ecuación (VIII-3) y

$$p_n = N_v e^{-(E_F - E_v^n)/kT} \quad (\text{VIII-6})$$

Si en esta última ecuación se aplica el logaritmo, queda:

$$\ln p_n = \ln N_v - (E_F - E_v^n)/kT \quad (\text{VIII-7})$$

Si se despeja E_v^n de la ecuación (VIII-7), se tiene:

$$E_v^n = E_F + kT \ln(p_n/N_v) \quad (\text{VIII-8})$$

Siguiendo el mismo procedimiento de aplicar logaritmo a la ec. (VIII-4), se puede despejar el valor de E_v^p ,

$$E_v^p = E_F + kT \ln(p_p/N_v) \quad (\text{VIII-9})$$

Reemplazando estos valores de la energía en la ecuación (VIII-2), se puede escribir:

$$qV_{bi} = -E_F - kT \ln(p_n/N_v) + E_F + kT \ln(p_p/N_v) = -kT \ln(p_n/p_p) \quad (\text{VIII-10})$$

o,

$$qV_{bi} = kT \ln(p_p/p_n) \quad (\text{VIII-11})$$

De acuerdo a la ec. (VIII-5), se puede poner a p_n en función de n_i y de n_n ,

$$p_n = n_i^2/n_n$$

Si se reemplaza en la ec. (VIII-11) queda:

$$qV_{bi} = kT \ln(p_p n_n/n_i^2) \quad (\text{VIII-12})$$

Si se considera que la ionización de los portadores es total, la ec. (VIII-12) se puede escribir:

$$V_{bi} = \left(\frac{kT}{q} \right) \ln(N_a N_d/n_i^2) \quad (\text{VIII-13})$$

Con lo cual se ha determinado el potencial de contacto. El valor kT/q se denomina “tensión térmica”.

Ejemplo VIII-1: se desea calcular el valor de V_{bi} .

Los datos del problema que se tiene que tener en cuenta son:

$$T = 300^\circ\text{K},$$

$$N_a = 1 \times 10^{17} \text{ cm}^{-3}$$

$$N_d = 1 \times 10^{15} \text{ cm}^{-3}$$

$$n_i = 1,5 \times 10^{10} \text{ cm}^{-3}$$

si reemplazamos estos valores en la ec. (VIII-13) se tiene:

$$V_{bi} = 0.026 \ln \left(\frac{10^{17} \times 10^{15}}{(1,5 \times 10^{10})^2} \right) = 0,82 \text{ V}$$

Se puede realizar otro cálculo cambiando alguno de los dopados, se puede tomar ahora

$$N_d = 1,10^{14} \text{ cm}^{-3}$$

Con lo cual

$$V_{bi} = 0,76 \text{ V}$$

Con esto se desea mostrar que los cambios en las densidades se ven atenuados por la función logaritmo, como se ve, un orden de magnitud en los dopados modifica en centésimas al potencial de contacto.

VIII.1.a2. El campo eléctrico, \mathcal{E} . Límites y ancho de la región de vaciamiento.

Debido a la existencia de zonas cargadas en las inmediaciones de la unión metalúrgica, la que se ha llamado zona de vaciamiento, sobre la cual se supone que las densidades de átomos ionizados en cada región son constantes, y que los límites de esta zona son abruptos (no son funciones de la distancia). Como se ha tomado la unión metalúrgica como origen de coordenadas, se tiene que la zona de carga negativa en la región p termina en $-x_p$, mientras que la región de carga positiva, en la región n lo hace en x_n , como se ha esquematizado en la Figura VIII-4.

Se sabe de electrostática que la existencia de cargas eléctricas da lugar a la aparición de un campo eléctrico, la forma de representar a este fenómeno es por medio de la ecuación de Poisson. La cual establece la relación entre el campo eléctrico o el potencial y la densidad de carga, afectada por una constante que describe la propiedad del espacio circundante a la carga, llamada la "permitividad del medio". La descripción generalmente se hace en un espacio tridimensional, pero como se ha mencionado previamente, se lo describirá en forma unidimensional.

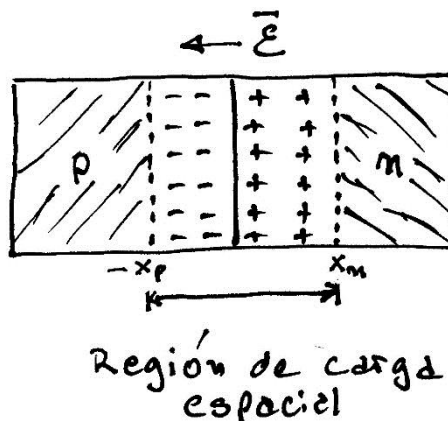


Figura VIII-4. Límites de la región de carga espacial.

De acuerdo a lo dicho, la ecuación de Poisson queda:

$$d\mathcal{E}(x)/dx = \rho(x)/\epsilon_s = -d^2V(x)/dx^2 \quad (\text{VIII-14})$$

Donde ρ es la densidad volumétrica de carga, ϵ_s es la permitividad del semiconductor, la cual se puede escribir como $K_s\epsilon_0$, siendo K_s la constante dieléctrica relativa del semiconductor (constante adimensional) y ϵ_0 la permitividad del vacío (Farad/cm). $\mathcal{E}(x)$ es el campo eléctrico y $V(x)$ es el potencial eléctrico. Se debe mencionar que la densidad de carga neta ρ está dada por la diferencia entre los portadores de carga móviles y por los iones, es decir:

$$\rho(x) = q(p - n + N_d - N_a) \quad (\text{VIII-15})$$

Aquí q es la carga eléctrica elemental, no se utiliza e para que no se confunda con la función exponencial, si en algún momento se utiliza para tal fin, se aclarará convenientemente. Como se puede ver en la Figura VIII-4 las densidades de carga se pueden definir, de acuerdo al espacio que ocupan,

$$\rho(x) = q N_d \quad 0 < x < x_n \quad (\text{VIII-16})$$

$$\rho(x) = -q N_a \quad -x_p < x < 0 \quad (\text{VIII-17})$$

Partiendo de estas definiciones se puede determinar el valor de $\mathcal{E}(x)$. De la ec. (VIII-14) se tiene:

$$d\mathcal{E}(x)/dx = qN_d/\epsilon_s$$

Si se integra:

$$\mathcal{E}(x) = \frac{qN_d}{K_s\epsilon_0} \int_x^{x_n} dx = \frac{qN_d}{K_s\epsilon_0} (x_n - x) \quad (\text{VIII-18})$$

Por otro lado, se tiene para la zona $(-x_p, 0)$

$$\mathcal{E}(x) = -\frac{qN_a}{K_s\epsilon_0} \int_{-x_p}^x dx = -\frac{qN_a}{K_s\epsilon_0} (x + x_p) \quad (\text{VIII-19})$$

Se puede ver que las expresiones del campo eléctrico en ambas zonas depende linealmente de la variable x . Si se hace una gráfica de $\mathcal{E}(x)$, se obtiene la Figura VIII-5.

En el intervalo $(-x_p, 0)$ la pendiente es negativa, mientras que en $(0, x_n)$ es positiva, pero el campo es siempre negativo, en esta disposición, apunta en la dirección de los x negativos (de la región n a la p), esto significa que siempre se opone al pasaje de portadores mayoritarios. El campo eléctrico es continuo y existe aun cuando no haya ningún potencial externo aplicado. Este se genera espontáneamente cuando el material dopado p se pone en contacto con el material n , empieza el proceso de difusión y van apareciendo las zonas con cargas opuestas.

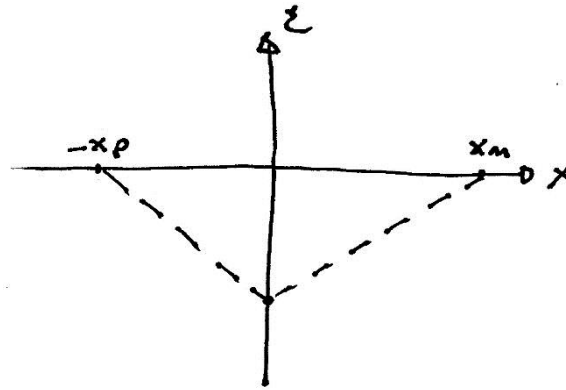


Figura VIII-5. Gráfica del campo eléctrico en función de la distancia.

Este tiene su máxima intensidad en $x = 0$, por lo que si igualamos las ecuaciones (VIII-18) y (VIII-19) para este valor de x , se tiene:

$$\mathcal{E}_{max} = -\left(\frac{qN_d}{K_s\epsilon_0}\right)x_n = -\left(\frac{qN_a}{K_s\epsilon_0}\right)x_p \quad \text{(VIII-20)}$$

Tomando la igualdad de los dos últimos miembros y simplificando términos:

$$N_d x_n = N_a x_p \quad \text{(VIII-21)}$$

Esto significa que las áreas en un gráfico densidad-distancia deben ser iguales, como se muestra en la Figura VIII-6.

En la ecuación (VIII-21) se observa que esa relación balancea el espacio ocupado por la carga con el dopado existente. Esto quiere decir que si

$$N_d > N_a \text{ se tiene que } x_n < x_p$$

Contrario a esto, si $N_d < N_a$ esto implica que $x_n > x_p$.

Como en este desarrollo el campo eléctrico está delimitado por x_n y x_p , se ve que el campo se extiende siempre sobre la parte menos dopada de la unión semiconductor.

Todos los desarrollos llevados a cabo han sido hecho bajo determinadas aproximaciones. El conjunto de ellas se denomina "aproximación de vaciamiento". Esta se basa en que:

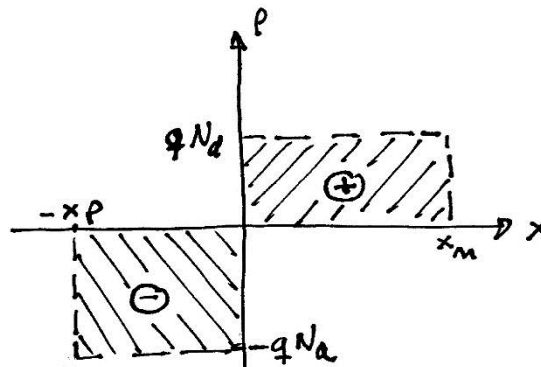


Figura VIII-6. Densidad de carga en función de la distancia.

- a) el campo eléctrico solamente está presente en la región $(-x_p, x_n)$, fuera de esta el campo se considera nulo.
- b) Los bordes de la zona de carga terminan en forma abrupta, como se ve en la Figura (VIII-6). La densidad de carga fuera de estos límites es cero (estas regiones son llamadas regiones masivas), pero en el interior del rango $(-x_p, x_n)$, cada una de ellas es constante.
- c) Los contactos que pueda tener el semiconductor están alejados de la unión metalúrgica ($x = 0$).
- d) No hay portadores móviles en la región de vaciamiento $(-x_p, x_n)$.
- e) En la zona de vaciamiento no hay ni generación ni recombinación de portadores.

A partir del conocimiento del campo eléctrico se puede obtener una expresión para el potencial de contacto. Si se parte de la ecuación de Poisson (VIII-14):

$$dV/dx = -\mathcal{E} \quad (\text{VIII-22})$$

Se puede escribir:

$$V(-x_p) = 0, \quad dV/dx = \frac{qN_a}{K_s\epsilon_0} (x + x_p)$$

Integrando,

$$V(x) = \int_0^{V(x)} dV = \frac{qN_a}{K_s\epsilon_0} \int_{-x_p}^x (x + x_p) dx$$

Se tiene,

$$V(x) = \frac{qN_a}{2K_s\epsilon_0} (x + x_p)^2 \quad (\text{VIII-23})$$

Ecuación válida en el rango $-x_p \leq x \leq 0$.

Para el rango $0 \leq x \leq x_n$ sabemos que $V(x_n) = V_{bi}$. Esto es debido a que el potencial en la región masiva del lado p se eligió nulo, entonces el potencial en la región masiva n debe tener la altura V_{bi} . Siguiendo el mismo procedimiento anterior, se tiene:

$$dV/dx = \frac{qN_d}{K_s\epsilon_0} (x_n - x)$$

Integrando,

$$\int_{V(x)}^{V_{bi}} dV = V_{bi} - V(x) = \frac{qN_d}{K_s\epsilon_0} \int_x^{x_n} (x_n - x) dx$$

Se obtiene:

$$V(x) = -\frac{qN_d}{2K_s\epsilon_0} (x_n - x)^2 + V_{bi} \quad (\text{VIII-24})$$

Se puede ver de las ecuaciones (VIII-23) y (VIII-24) que el potencial es una función cuadrática con respecto a la distancia. En el rango $(-x_p, x_n)$ el potencial adquiere diferentes comportamientos. Si desarrolláramos la derivada segunda entre los valores $(-x_p, 0)$ se vería que esta es positiva, esto significa que la concavidad de la curva es hacia arriba. Mientras que en el rango $(0, x_n)$ la derivada segunda es negativa, esto implica la concavidad hacia abajo como se ven en la Figura (VIII-7)

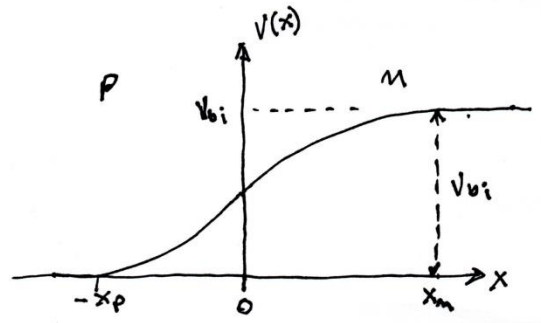


Figura VIII-7. Potencial sobre la unión pn.

La función potencial es continua, por lo que en $x = 0$ se puede escribir:

$$\frac{qN_a}{2K_s\epsilon_0} x_p^2 = -\frac{qN_d}{2K_s\epsilon_0} x_n^2 + V_{bi}$$

Despejando el potencial de contacto, se obtiene:

$$V_{bi} = \frac{q}{2K_s\epsilon_0} (x_p^2 N_a + x_n^2 N_d) \quad (\text{VIII-25})$$

De la ecuación (VIII-21) se puede despejar uno de los límites de zona,

$$x_p = \frac{N_d}{N_a} x_n \quad (\text{VIII-26})$$

Esta expresión se reemplaza en la ecuación (VIII-25) la cual queda:

$$V_{bi} = \frac{q}{2K_s\epsilon_0} \left(\frac{N_d^2}{N_a} + N_d \right) x_n^2$$

Si despejamos x_n , se obtiene:

$$x_n = \left[\frac{2K_s\epsilon_0}{q} V_{bi} \frac{N_a}{N_d(N_a+N_d)} \right]^{1/2} \quad (\text{VIII-27})$$

Mediante un procedimiento análogo se puede obtener:

$$x_p = \left[\frac{2K_s\epsilon_0}{q} V_{bi} \frac{N_d}{N_a(N_a+N_d)} \right]^{1/2} \quad (\text{VIII-28})$$

El ancho total de la región de vaciamiento, w , es la suma de estas dos últimas cantidades,

$$w = x_n + x_p = \left[\frac{2K_s\epsilon_0}{q} V_{bi} \frac{(N_a+N_d)}{N_a N_d} \right]^{1/2} \quad (\text{VIII-29})$$

Ahora que se ha determinado el valor de x_n y de x_p , se pueden reemplazar en la ecuación (VIII-20) y determinar una nueva expresión par el campo máximo, esta es:

$$\mathcal{E}_{max} = 2V_{bi}/w \quad (\text{VIII-30})$$

Con la cual se escribe al campo máximo (en $x = 0$) en función del potencial de contacto y las densidades de dopantes introducidos, el primero también depende de la densidad de portadores intrínsecos, es decir en función de parámetros conocidos como son N_a y N_d y que dependen del propio material dopado como lo es n_i .

Ejemplo VIII-2: Si se toman los dopados mencionados en el Ejemplo VIII-1 del subcapítulo anterior, se tiene: $N_a = 1 \times 10^{17} \text{ cm}^{-3}$, $N_d = 1 \times 10^{15} \text{ cm}^{-3}$, y el $V_{bi} = 0,82 \text{ V}$, se puede obtener la intensidad máxima del campo eléctrico si se utilizan las ecuaciones (VIII-20).

Se determinarán primero los valores de x_n y x_p .

Se debe recordar que, $K_s = 11,7$ (para el Silicio), $\epsilon_0 = 8,85 \times 10^{-14} \text{ F/cm}$.

$$x_n = [(2 \times 11,7 \times 8,85 \times 10^{-14} / 1,6 \times 10^{-19}) \times 0,82 \times (10^{17} / (10^{15} \times (10^{17} + 10^{15})))]^{1/2}$$

$$x_n = 3,242 \times 10^{-5} \text{ cm}$$

Un cálculo semejante es el que se debe hacer para conocer x_p .

$$x_p = [(2 \times 11,7 \times 8,85 \times 10^{-14} / 1,6 \times 10^{-19}) \times 0,82 \times (10^{15} / (10^{17} \times (10^{17} + 10^{15})))]^{1/2}$$

$$x_p = 3,242 \times 10^{-7} \text{ cm}$$

El valor del ancho total será la suma de estas dos cantidades,

$$w = x_n + x_p = 3,242 \times 10^{-5} + 3,242 \times 10^{-7} = 3,274 \times 10^{-5} \text{ cm}$$

Con estos valores se puede ver que la zona menos dopada es la más extensa.

Para corroborar estos valores, se puede calcular directamente el valor de w mediante la ecuación (VIII-29).

$$w = [(2 \times 11,7 \times 8,85 \times 10^{-14} / 1,6 \times 10^{-19}) \times 0,82 \times (10^{17} + 10^{15}) / (10^{17} \times 10^{15})]^{1/2}$$

$$w = 3,274 \times 10^{-5} \text{ cm.}$$

Se puede ahora calcular la intensidad máxima del campo, la cual se localizará en $x = 0$.

$$\mathcal{E} = qN_d / K_s \epsilon_0 x_n = (1,6 \times 10^{-19} \times 10^{15}) / (11,7 \times 8,85 \times 10^{-14}) \times 3,242 \times 10^{-5}$$

$$\mathcal{E} = 50.096,1 \text{ V/cm}$$

Similarmente

$$\mathcal{E} = qN_a / K_s \epsilon_0 x_p = 50.096,1 \text{ V/cm}$$

Lo cual muestra igual valor del campo eléctrico en el límite entre ambas zonas ($x = 0$) en cada una de las zonas.

VIII.2. La unión *pn* bajo la acción de un potencial externo

Hasta el momento no se ha aplicado a la unión *pn* ningún potencial externo, el único con el que se hicieron los desarrollos pasados es el potencial de contacto, el cual se genera por el hecho mismo de poner en contacto un mismo material con dopados diferentes.

Se puede aplicar ahora un potencial externo y ver cómo responde la unión. A la unión se le pueden aplicar un potencial de dos maneras diferentes. Uno es cuando el terminal negativo del potencial externo se conecta a la región *n* y el positivo a la *p*, este caso es llamado “polarización directa”. El otro es el opuesto, el terminal negativo se conecta a la región *p* y el positivo a la *n*, esta circunstancia es llamada “polarización inversa”. La ilustración de estas dos opciones se muestra en la Figura VIII-8.

En las uniones del semiconductor con el metal se generan potenciales V_n y V_p , que como veremos más adelante tiene valores fijos, solamente dependen del tipo de metal y de semiconductor utilizado en la unión. Por lo tanto, el potencial sobre la unión se puede escribir como:

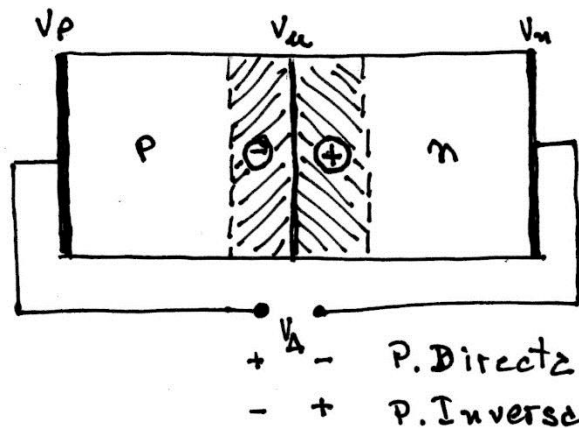


Figura VIII-8. Esquema de polarizaciones de la unión *pn*.

$$V_u = V_n - V_A + V_p \tag{VIII-31}$$

Si ocurriera como en el subcapítulo anterior, que $V_A = 0$, entonces

$$V_u = V_n + V_p = V_{bi} \tag{VIII-32}$$

Esta última expresión refleja el estado de equilibrio del sistema. Cuando se aplica una polarización externa el sistema sale del equilibrio. Se puede definir que el sistema está sometido a una polarización directa cuando $V_A > 0$. En caso contrario, si $V_A < 0$, se tiene aplicada una polarización inversa. Tener una u otra polarización significa que se debe insertar una batería en el lugar V_A de la Figura VIII-8. Esto hace que portadores positivos o negativos sean inyectados en la región *n* o *p*, y viceversa cambiando las conexiones.

VIII-2.a. Polarización directa

Se comenzará el análisis con una polarización directa sobre la unión. En este caso la parte negativa de la batería está conectada a la región n , como puede verse en la Figura VIII-9, esto significa que se están aportando portadores negativos a esta región, provocando la recombinación de portadores en las inmediaciones de la zona de vaciamiento.

La consecuencia de esta acción es una reducción de esta zona. Como $V_A > 0$, el potencial en la unión, será:

$$V_u = V_{bi} - V_A \tag{VIII-33}$$

Con lo cual se ve que el potencial sobre la unión se encuentra reducido con respecto al sistema en equilibrio. Como puede verse en la Figura VIII-10. Utilizando esta última expresión y remplazándola en las ecuaciones (VIII-27), (VIII-28) y (VIII-29), se tiene:

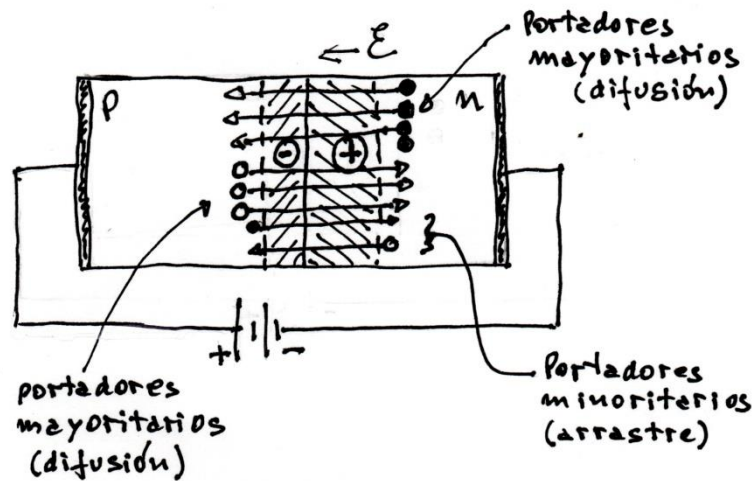


Figura VIII-9. Esquema de conexión en “polarización directa” y del movimiento de portadores por arrastre y difusión.

$$x_n = \left[\frac{2K_S \epsilon_0}{q} (V_{bi} - V_A) \frac{N_a}{N_d(N_a + N_d)} \right]^{1/2} \tag{VIII-34}$$

$$x_p = \left[\frac{2K_S \epsilon_0}{q} (V_{bi} - V_A) \frac{N_d}{N_a(N_a + N_d)} \right]^{1/2} \tag{VIII-35}$$

$$w = \left[\frac{2K_S \epsilon_0}{q} (V_{bi} - V_A) \frac{(N_a + N_d)}{N_a N_d} \right]^{1/2} \tag{VIII-36}$$

Se puede ver de estas expresiones, que la zona de vaciamiento se ha reducido en función de la aplicación de un potencial en directo sobre la unión.

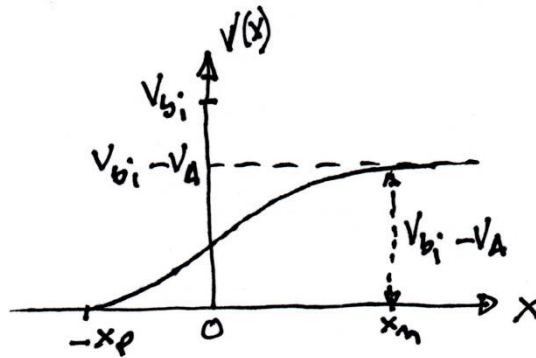


Figura VIII-10. Curva del potencial con polarización directa aplicada.

El campo eléctrico que existe en la zona de vaciamiento es una función que depende en forma directa con el ancho de las zonas, por lo tanto, este también se reduce. El valor del campo máximo en este caso será, de acuerdo a las ecuaciones (VIII-30) y (VIII-33),

$$\mathcal{E}_{max_pd} = 2(V_{bi} - V_A)/w$$

La variación relativa del campo eléctrico con respecto a la no existencia de potencial aplicado, en función de la distancia, se puede ver en la Figura VIII-11.

Ejemplo VIII-3: Se puede retomar el Ejemplo VIII-2 aplicando un potencial externo

$V_A = 0,7V$, y calcular nuevamente los valores de x_n , x_p , w y \mathcal{E} .

$$x_n = [(2 \times 11,7 \times 8,85 \times 10^{-14} / 1,6 \times 10^{-19}) \times (0,82 - 0,7) \times (10^{17} / (10^{15} \times (10^{17} + 10^{15})))]^{1/2}$$

$$x_n = 1,240 \times 10^{-5} \text{ cm}$$

$$x_p = [(2 \times 11,7 \times 8,85 \times 10^{-14} / 1,6 \times 10^{-19}) \times (0,82 - 0,7) \times (10^{15} / (10^{17} \times (10^{17} + 10^{15})))]^{1/2}$$

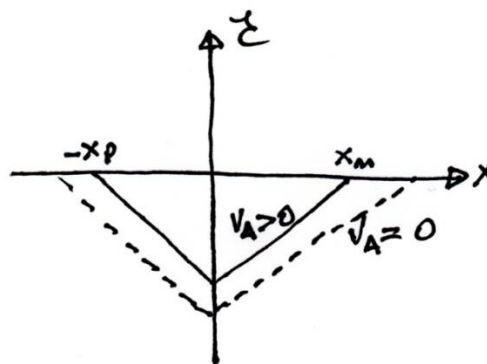


Figura VIII-11. Campo eléctrico sobre la zona de vaciamiento, con línea llena, $V_A > 0$, con línea de trazos, cuando $V_A = 0$.

$$x_p = 1,240 \times 10^{-7} \text{ cm}$$

$$w = [(2 \times 11,7 \times 8,85 \times 10^{-14} / 1,6 \times 10^{-19}) \times (0,82 - 0,7) \times (10^{17} + 10^{15}) / (10^{17} \times 10^{15})]^{1/2}$$

$$w = 1,252 \times 10^{-5} \text{ cm.}$$

En el Ejemplo VIII-2 se ha visto que el valor del campo no cambia si se calcula con el valor de x_n o x_p , por lo tanto,

$$\mathcal{E} = qN_d / K_S \epsilon_0 x_n = (1,6 \times 10^{-19} \times 10^{15}) / (11,7 \times 8,85 \times 10^{-14}) \times 1,240 \times 10^{-5}$$

$$\mathcal{E} = 19.160,8 \text{ V/cm}$$

Con la aplicación de un potencial externo en forma directa de 0,7V, ambas zonas se han reducido un 61,8%. En igual porcentaje se ha reducido el campo eléctrico.

La aplicación de un potencial externo $V_A > 0$ produce modificaciones en cantidades sensibles como las que se han calculado, pero la modificación fundamental que da lugar a estas que se han visto es la variación de la estructura energética de la unión, la estructura de bandas, como puede verse en la Figura VIII-12.

Un potencial de este tipo, lo que produce es reducir la altura de la barrera de potencial impuesta por el potencial de contacto (V_{bi}) en la cantidad V_A , como se indica en la ec. (VIII-33).

Esto hace que los portadores mayoritarios vean una barrera de menor altura, permitiendo que el proceso de difusión sea mayor frente al que ocurrió cuando,

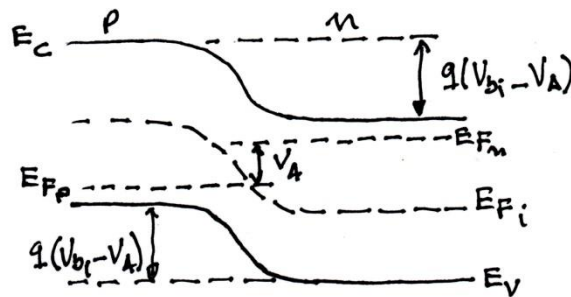


Figura VIII-12. Estructura de banda con polarización directa.

originalmente se tuvo, $V_A = 0$. Esto significa que una gran cantidad de portadores mayoritarios puedan pasar de una región a la otra, transformándose en portadores minoritarios. Es decir, la mayor cantidad de electrones (portadores mayoritarios) que pasan de la región n a la p , se contabilizan en esta última como una gran afluencia de portadores minoritarios, lo mismo pasa con los portadores positivos (huecos) que pasan de la zona p a la n . Este proceso se denomina “inyección de portadores minoritarios”.

Como se ha visto, el ancho de la zona de vaciamiento se reduce. Si bien este hecho cambia la intensidad del campo eléctrico, no afecta sensiblemente la cantidad de portadores minoritarios pasando a través de la zona de vaciamiento, pues estos son generados térmicamente, por lo que mientras no cambie sustancialmente la temperatura, la cantidad de portadores minoritarios se mantendrá casi constante.

VIII-2.b. Polarización inversa

La otra posibilidad para conectar a la unión pn es poner el polo positivo de la batería a la región n y el negativo la región p . La polarización que se logra con este tipo de conexión es la denominada “polarización inversa”, donde, de acuerdo a lo que se definió previamente, el potencial que se aplica sobre la unión es negativo, es decir $V_A < 0$. Como se ve en la Fig. VIII-13.

Por lo tanto, el potencial que siente la unión será:

$$V_u = V_{bi} + V_A \tag{VIII-37}$$

Como se hizo con la polarización directa, se vuelve sobre las ecuaciones (VIII-27), (VIII-28) y (VIII-29) en las cuales se reemplaza el valor del potencial por el que aparece en la ecuación (VIII-37), de esta manera los valores de x_n , x_p y w quedan:

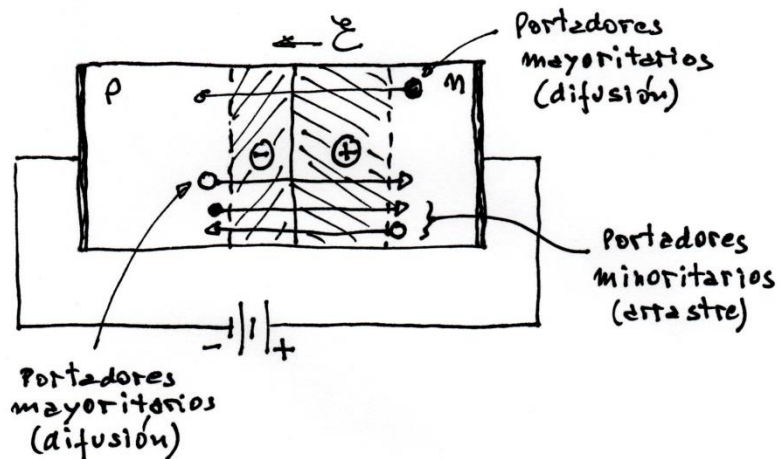


Figura VIII-13. Esquema de conexión en “polarización inversa” y del movimiento de portadores por arrastre y difusión.

$$x_n = \left[\frac{2K_s \epsilon_0}{q} (V_{bi} + V_A) \frac{N_a}{N_d(N_a + N_d)} \right]^{1/2} \tag{VIII-38}$$

$$x_p = \left[\frac{2K_s \epsilon_0}{q} (V_{bi} + V_A) \frac{N_d}{N_a(N_a + N_d)} \right]^{1/2} \tag{VIII-39}$$

$$w = \left[\frac{2K_s \epsilon_0}{q} (V_{bi} + V_A) \frac{(N_a + N_d)}{N_a N_d} \right]^{1/2} \tag{VIII-40}$$

En este caso la barrera de potencial que ven los portadores mayoritarios se ve aumentada en la cantidad V_A , por lo que menos portadores serán capaces de sobrepasarla por difusión. Nuevamente aquí, la cantidad de portadores minoritarios, que son generados térmicamente, pasando a través de la zona de vaciamiento por efecto del campo no ha cambiado, pues la temperatura no ha cambiado. La conexión del lado positivo de la batería a la región n y negativo a la región p , hace que portadores de signo negativo sean extraídos de las inmediaciones de la unión sobre el lado n incrementando el ancho de la zona de vaciamiento, aumentando la carga positiva proceso que es llamado de “extracción de portadores”. Similar es lo que pasa en el lado p , con respecto a los portadores positivos.

Ejemplo VIII-4. Se retomará el ejemplo del subcapítulo anterior para poder comparar los valores obtenidos.

Los datos que se utilizaron fueron: $N_a = 1 \times 10^{17} \text{ cm}^{-3}$, $N_d = 1 \times 10^{15} \text{ cm}^{-3}$, y $V_{bi} = 0,82 \text{ V}$, se debe recordar que ahora $V_A = -0,7 \text{ V}$, por lo que:

$$x_n = [(2 \times 11,7 \times 8,85 \times 10^{-14} / 1,6 \times 10^{-19}) \times (0,82 + 0,7) \times (10^{17} / (10^{15} \times (10^{17} + 10^{15})))]^{1/2}$$

$$x_n = 4,413 \times 10^{-5} \text{ cm}$$

$$x_p = [(2 \times 11,7 \times 8,85 \times 10^{-14} / 1,6 \times 10^{-19}) \times (0,82 + 0,7) \times (10^{15} / (10^{17} \times (10^{17} + 10^{15})))]^{1/2}$$

$$x_p = 4,413 \times 10^{-7} \text{ cm}$$

$$w = [(2 \times 11,7 \times 8,85 \times 10^{-14} / 1,6 \times 10^{-19}) \times (0,82 + 0,7) \times (10^{17} + 10^{15}) / (10^{17} \times 10^{15})]^{1/2}$$

$$w = 4,458 \times 10^{-5} \text{ cm.}$$

Se puede ver de estos valores que las regiones se han extendido un 36,1%, con respecto a la longitud de la zona de vaciamiento cuando ningún potencial había sido aplicado.

Por último, se puede calcular el valor del campo eléctrico, de acuerdo a la ecuación (VIII-20) se tiene:

$$\mathcal{E} = qN_d / K_s \epsilon_0 x_n = (1,6 \times 10^{-19} \times 10^{15}) / (11,7 \times 8,85 \times 10^{-14}) \times 4,413 \times 10^{-5}$$

$$\mathcal{E} = 68.190,6 \text{ V/cm}$$

Con una polarización inversa de $-0,7 \text{ V}$, el campo eléctrico se incrementa en un 36,1%, como se ve en la Figura VIII-14. Aquí el campo máximo (en $x = 0$) es:

$$\mathcal{E}_{max_pi} = 2(V_{bi} + V_A) / w \tag{VIII-41}$$

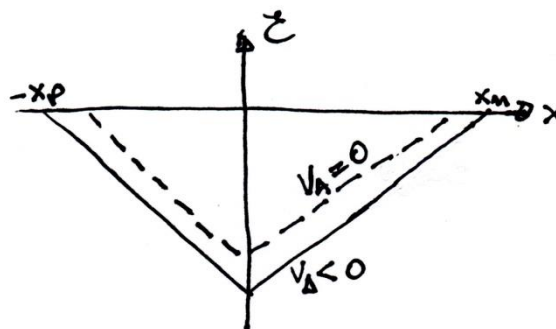


Figura VIII-14. Variación del campo eléctrico bajo polarización inversa. Con línea punteada $V_A = 0 \text{ V}$, con línea llena $V_A < 0 \text{ V}$.

Contrario a lo que pasaba con la polarización directa, en este caso el potencial se incrementa como se ha mencionado y se esquematiza en la Figura VIII-15.

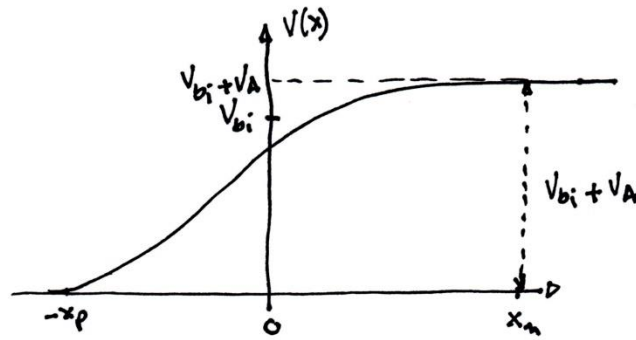


Figura VIII-15. Curva del potencial con polarización inversa aplicada.

La aplicación de un potencial de estas características sobre la unión, como se ha visto, hace que se modifiquen parámetros como por ejemplo las distancias a los límites de la zona de vaciamiento desde la unión metalúrgica, es decir, los valores $-x_p, x_n$.

Como se dijo en el subcapítulo previo, todas estas variaciones provienen de la modificación de la estructura de bandas de energía de la unión. El hecho de establecer un incremento al potencial de contacto con el potencial aplicado, hace que los portadores mayoritarios vean un aumento de la barrera de potencial que deben atravesar por difusión, con respecto a la no aplicación de algún potencial externo reduciendo la cantidad de estos que pueden atravesar la barrera.

Esta situación puede verse en la Figura VIII-16 donde se ha realizado un esquema de la estructura de bandas.

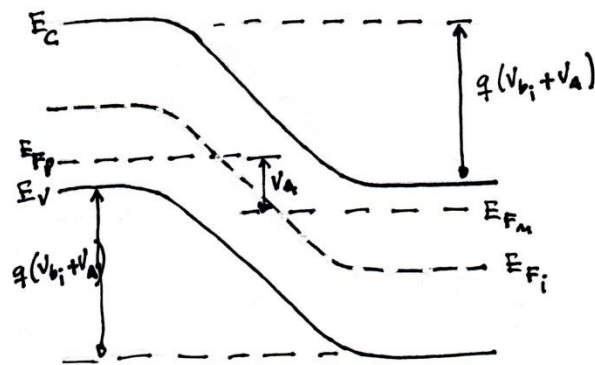


Figura VIII-16. Estructura de banda con polarización inversa.

VIII.2.c. La capacitancia de la unión

Se ha dicho al comienzo de este libro que la Física es una ciencia que basa su desarrollo en modelos, los cuales tratan de emular al objeto que está bajo estudio.

En este punto se podría modelizar la unión de estos semiconductores a los cuales les hemos asignado dopados ideales. Se puede imaginar a la unión como si las cargas que existen en la vecindad de la unión metalúrgica y que ocupan los anchos x_n y x_p , respectivamente, pudieran moverse y concentrarse a una distancia x_n y x_p desde la unión metalúrgica, de esta manera se tendría a las cargas positivas y negativas a una distancia w . Esta disposición se la puede

imaginar como si fuera un capacitor de placas paralelas con las cargas correspondientes a las que hay en la zona de vaciamiento. Esquemáticamente puede verse en la Figura VIII-17.

Esta capacitancia, denominada capacitancia de barrera, está ligada directamente a las cargas fijas existentes en la zona de vaciamiento. Se debe tener en cuenta que se debe introducir otra capacitancia que es la que da cuenta del fenómeno de difusión, en donde se debe tener en cuenta el aporte hecho por los portadores minoritarios inyectados. Debido a que ambas capacidades tienen en cuenta circunstancias diferentes, se debería distinguir dada una cierta polarización, cuál de las dos capacidades predomina. Cuando se tiene aplicada una polarización directa, la altura de la barrera de potencial disminuye y este efecto favorece el pasaje de

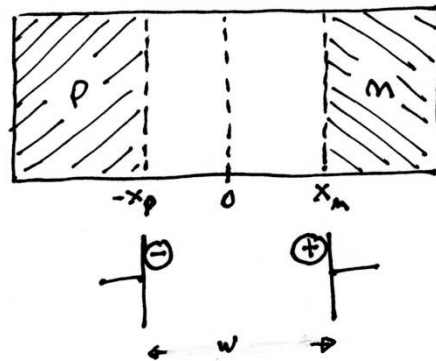


Figura VIII-17. Modelización de la unión *pn* como un capacitor de placas paralelas.

portadores, por lo tanto, favorece la inyección de portadores, obviamente, la capacitancia que domina este estadio del sistema es la capacitancia de difusión. Mientras que si se tiene una polarización inversa, se está incrementando la barrera de potencial y la zona de vaciamiento, más carga de ambos signos habrá en esta zona, dominará la capacitancia de barrera. Se encontrará primero la capacitancia de barrera.

La definición general de la capacitancia está dada por la relación entre la variación de la carga acumulada y la variación del potencial, se puede escribir

$$C = \left| \frac{dQ}{dV} \right| \tag{VIII-42}$$

Esto quiere decir que si se realiza una variación del potencial, se va a generar una variación de la carga, por ejemplo, si x_n es el límite de la zona de vaciamiento en la región *n*, la cual se obtiene debido a la aplicación del potencial inverso V_A . Si se realiza una variación del potencial aplicado dV_A , esto generará una variación de la carga acumulada dQ en la zona de vaciamiento, involucrando una variación en la ocupación espacial, de magnitud dx_n , de la carga. Gráficamente se puede mostrar como en la Figura VIII-18. Un proceso idéntico ocurre para la zona *p*.

Como se trata con un sistema unidimensional se puede utilizar la expresión (VIII-21)

$$N_d x_n = N_a x_p$$

La carga en cada una de las regiones será:

$$Q = qAN_d x_n = qAN_a x_p \tag{VIII-43}$$

Donde A es el área transversal de la unión. Se puede encontrar la variación de la carga,

$$dQ = qAN_d dx_n = qAN_a dx_p \tag{VIII-44}$$

De acuerdo a la ecuación (VIII-38) se tiene:

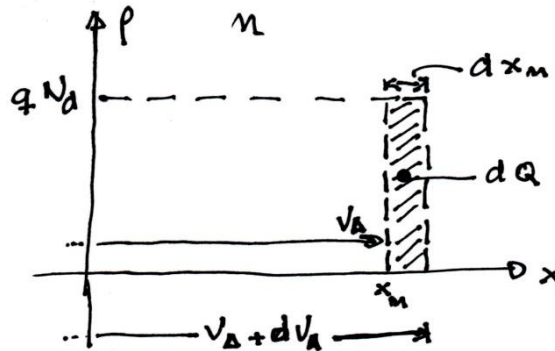


Figura VIII-18. Variación de la carga en función del potencial.

$$x_n = \left[\frac{2K_S \epsilon_0}{q} (V_{bi} + V_A) \frac{N_a}{N_d(N_a + N_d)} \right]^{1/2}$$

Bajo estas condiciones se tiene:

$$C = \left| \frac{dQ}{dV} \right| = qAN_d \left| \frac{dx_n}{dV_A} \right| \tag{VIII-45}$$

Operando queda:

$$C = \frac{A}{2} \left[\frac{2qK_S \epsilon_0 N_a N_d}{(V_{bi} - V_A)(N_a + N_d)} \right]^{1/2} = K_S \epsilon_0 / w \tag{VIII-46}$$

En donde para obtener la última igualdad se ha considerado un área unitaria. Esta expresión describe la capacitancia de la unión cuando la juntura pn se encuentra polarizada en forma inversa. También se ha considerado que el dopado es uniforme.

La polarización en directa se produce debido a la inyección de portadores, como se mencionó más arriba, este punto se tratará más adelante.

Ejemplo VIII-5. Siguiendo con los datos que se han utilizado para los anteriores ejemplos, ahora se calculara la capacitancia de barrera de la unión propuesta.

Cuando no hay un potencial externo aplicado (Ejemplo VIII-2), $w = 3,274 \times 10^{-5}$ cm, y considerando un área de 10^{-4}cm^2 , se tiene:

$$C = [11,7 \times 8.85 \times 10^{-14} / 3,27 \times 10^{-5}] \times 10^{-4} = 3,167 \times 10^{-12} \text{F} = 3,167 \text{pF}$$

Cuando la polarización es directa (Ejemplo VIII-3), $w = 1,252 \times 10^{-5}$ cm. Por lo que:

$$C = [11,7 \times 8.85 \times 10^{-14} / 1,252 \times 10^{-5}] \times 10^{-4} = 8,27 \times 10^{-12} \text{F} = 8,27 \text{pF}$$

Si la polarización fuera inversa (Ejemplo VIII-4), se tiene: $w = 4,458 \times 10^{-5}$ cm.

$$C = [11,7 \times 8.85 \times 10^{-14} / 4,458 \times 10^{-5}] \times 10^{-4} = 2,32 \times 10^{-12} \text{F} = 2,32 \text{pF}$$

VIII.3. Ecuación de continuidad. Distribución de portadores minoritarios

Como se ha visto en los subcapítulos anteriores, dependiendo del tipo de polarización que apliquemos a la unión, los portadores mayoritarios se verán afectados en forma diferente. Si esta es directa, su movimiento de una región hacia la otra, debido al proceso de difusión se verá favorecido. Mientras que la difusión de portadores, bajo una polarización inversa, verá un gran escollo en la gran barrera de potencial que se les presenta para su pasaje a la región opuesta, por lo que son pocos los portadores en este caso que tiene la energía suficiente para pasar la barrera de potencial.

El solo hecho de poner un semiconductor con dopados diferentes en contacto, hace que se genera un campo eléctrico que siempre favorece el pasaje de los portadores minoritarios. Es decir, electrones de la región p hacia la región n , y huecos de la n a la p . Se entiende de esta descripción que el movimiento de estos últimos portadores, de una región a la otra, no puede ser evitado, como si lo puede ser el de los mayoritarios.

Por este motivo es importante estudiar como es el comportamiento de los portadores minoritarios, su influencia en el comportamiento de un dispositivo electrónico es fundamental. Estos comportamientos, los estudiamos bajo la influencia de algún tipo de polarización, por lo que hacia donde apunta el estudio que se hará con la ecuación de continuidad, es sobre el exceso de portadores.

Se ha visto en el capítulo VII la ecuación de continuidad, y las aproximaciones con las cuales se ha desarrollado la teoría. Se utilizara ahora para ver como es el comportamiento de los portadores minoritarios.

Otra aproximación que haremos es que se tomará un régimen estacionario, es decir $dVariable/dt = 0$.

Si se toman la ecuaciones (VII-55) y (VII-56) (aquí se repiten)

$$\frac{\partial \Delta n_p}{\partial t} = D_n \frac{\partial^2 \Delta n_p}{\partial x^2} - \frac{\Delta n_p}{\tau_n} + G_L \quad (\text{VII-55})$$

$$\frac{\partial \Delta p_n}{\partial t} = D_p \frac{\partial^2 \Delta p_n}{\partial x^2} - \frac{\Delta p_n}{\tau_p} + G_L \quad (\text{VII-56})$$

Como se dijo anteriormente, nos encontramos en un proceso estacionario,

$$\frac{\partial \Delta n_p}{\partial t} = \frac{\partial \Delta p_n}{\partial t} = 0 \quad (\text{VIII-47})$$

Y no existe generación óptica, $G_L = 0$. Por lo que las ecuaciones (VII-55) y (VII-56) quedan:

$$D_n \frac{\partial^2 \Delta n_p}{\partial x^2} - \frac{\Delta n_p}{\tau_n} = 0 \tag{VIII-48}$$

$$D_p \frac{\partial^2 \Delta p_n}{\partial x^2} - \frac{\Delta p_n}{\tau_p} = 0 \tag{VIII-49}$$

Se tomará la ecuación (VIII-49) para obtener la solución bajo una polarización directa. Esta queda:

$$D_p \frac{\partial^2 \Delta p_n}{\partial x^2} = \frac{\Delta p_n}{\tau_p}$$

Se puede reescribir:
$$\frac{\partial^2 \Delta p_n}{\partial x^2} = \frac{\Delta p_n}{D_p \tau_p} = \frac{\Delta p_n}{L_p^2} \tag{VIII-50}$$

Donde $L_p = \sqrt{D_p \tau_p}$.

Se analizará el comportamiento de los huecos (portadores minoritarios) en la región n , similar análisis es para los electrones en la región p , y para ambos con una polarización inversa. Para este análisis se definirá un nuevo sistema de coordenadas, como se ve en la Figura VIII-19.

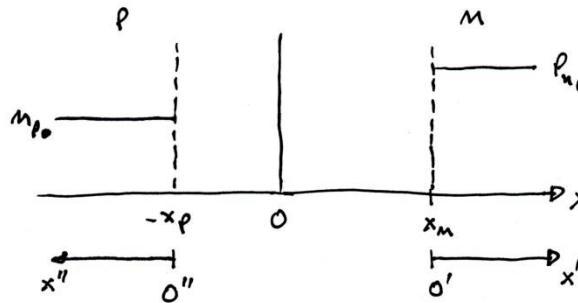


Figura VIII-19. Sistema de referencia para movimiento de portadores minoritarios

Aquí, se definen dos orígenes de coordenadas, los cuales son positivos en el sentido de las flechas de los ejes, esto quiere decir que el eje doble primado es positivo hacia la izquierda. El simplemente primado es positivo hacia la derecha. La ventaja de esta definición del sistema de coordenadas se debe a la simplicidad que introduce en el tratamiento matemático del problema. Para conocer el comportamiento de los portadores se debe obtener la solución de cada una de las ecuaciones (VIII-48) y (VIII-49). El tratamiento para ambas es idéntico, se puede tomar una de las ecuaciones para mostrar el proceso de solución. Se tomara la ecuación (VIII-49). La ecuación es:

$$\frac{\partial^2 \Delta p_n}{\partial x^2} = \frac{\Delta p_n}{L_p^2}$$

Como nos movemos en una dimensión en el eje simplemente primado, esta queda:

$$\frac{d^2 \Delta p_n(x')}{dx'^2} = \frac{\Delta p_n(x')}{L_p^2} \quad (\text{VIII-51})$$

La solución es:

$$\Delta p_n(x') = A_1 e^{x'/L_p} + A_2 e^{-x'/L_p} \quad (\text{VIII-52})$$

Debido a que la cantidad de portadores minoritarios va decreciendo por efecto de la recombinación, es que la constante A_1 del segundo término debe anularse, pues a medida que x' crece, ese término tiende a divergir, por lo que nuestra primera condición de contorno es:

$$\text{Cuando } x' \rightarrow \infty \Rightarrow A_1 = 0$$

Por otro lado, la cantidad de portadores que pasan por difusión y llegan al origen $0'$ depende del potencial externo aplicado. La variación entre los que han pasado por difusión (p_n) y los que han sido generados térmicamente (p_{n0}) definen la segunda condición de contorno

$$\Delta p_n(0') = p_n - p_{n0} = p_{n0} \left(e^{qV_A/kT} - 1 \right) = A_2 \quad (\text{VIII-53})$$

Aplicando estas condiciones a la ecuación (VIII-52) se tiene

$$\Delta p_n(x') = p_{n0} \left(e^{qV_A/kT} - 1 \right) e^{-x'/L_p} \quad (\text{VIII-54})$$

A esta ecuación se la puede escribir de la siguiente manera:

$$p_n(x') = p_{n0} + p_{n0} \left(e^{qV_A/kT} - 1 \right) e^{-x'/L_p} \quad (\text{VIII-55})$$

La cual permite establecer que los portadores minoritarios a partir del origen $0'$ son la suma de aquellos generados térmicamente y los que han sido aportados por la existencia de un potencial externo, los cuales irán decayendo con una ley exponencial. Esto puede verse en la figura VIII-20. La ecuación (VIII-55) también se puede escribir como:

$$p_n(x') = p_{n0} + \Delta p_n(x') \quad (\text{VIII-56})$$

Un tratamiento absolutamente similar es para los electrones en la región p . En cuyo caso la ecuación (VIII-56) quedaría:

$$n_p(x'') = n_{p0} + \Delta n_p(x'') \quad (\text{VIII-57})$$

Donde

$$\Delta n_p(x'') = n_{p0} \left(e^{qV_A/kT} - 1 \right) e^{-x''/L_n} \quad (\text{VIII-58})$$

Se debe recordar que se había mencionado que el eje x'' hacia la izquierda era positivo. Con lo cual, la descripción gráfica del comportamiento de ambos tipos de portadores en las regiones n

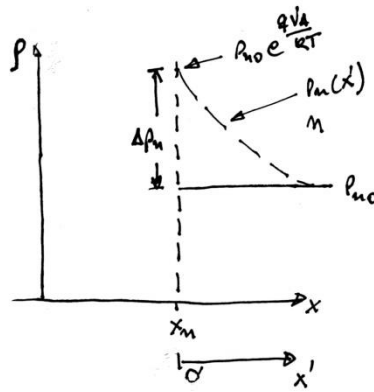


Figura VIII-20. Comportamiento de los portadores minoritarios.

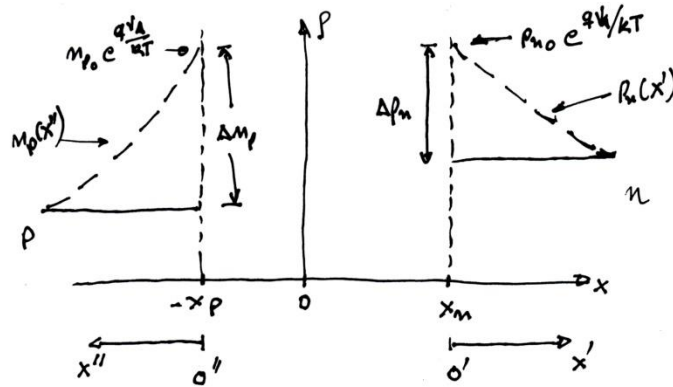


Figura VIII-21. Comportamiento de los portadores minoritarios en ambas regiones.

y p tiene la forma que se muestra en la figura VIII-21.

La obtención de las soluciones para la distribución de los portadores minoritarios es lo que permitirá determinar las corrientes en la unión.

VIII.4. Corrientes en la juntura pn

Una de las primeras condiciones que se deben tener en cuenta para la determinación de las corrientes de los portadores mayoritarios y minoritarios es que, la corriente a través de la unión es constante y función de la distancia,

$$J = J_n(x) + J_p(x) = cte. \tag{VIII-59}$$

La constancia de la corriente indica que la suma de los dos tipos de portadores es siempre la misma, pero no indica que la cantidad de portadores de cada tipo sea constante en la unión. En nuestra aproximación se ha mencionado que no hay procesos de recombinación ni de generación en la zona de vaciamiento, el campo eléctrico es constante y el proceso de difusión está termodinámicamente en equilibrio, por lo que dentro de la zona de vaciamiento, las corrientes de portadores mayoritarios y minoritarios son constantes. Se debe determinar el valor de cada una de ellas.

La densidad de corriente de difusión en la región n , la podemos expresar como:

$$J_p(x') = -qD_p \frac{d\Delta p_n}{dx} = \frac{qD_p}{L_p} p_{n0} (e^{qV_A/kT} - 1) e^{-x'/L_p} \quad \text{(VIII-60)}$$

Debido a que consideramos al sistema en equilibrio, en el límite x_n se tendrá:

$$J_p(x_n) = J_p(0') = \frac{qD_p}{L_p} p_{n0} (e^{qV_A/kT} - 1) \quad \text{(VIII-61)}$$

Como se ha venido suponiendo, la densidad de dopantes aceptores es mayor que la de donores, $N_a > N_d$, la representación de las corrientes es como se ve en la figura VIII-22.

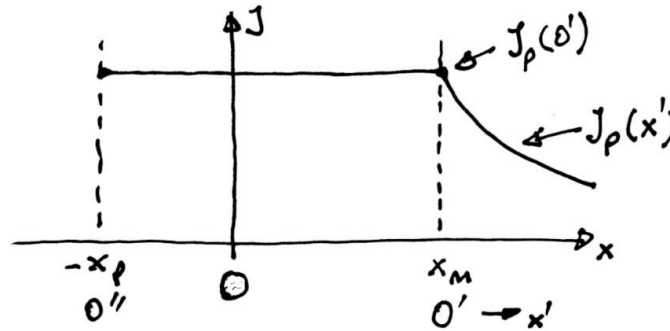


Figura VIII-22. Densidad de corriente de portadores minoritarios en la región n.

De manera similar, en la región p, la variación de portadores minoritarios indicará el valor de la corriente de difusión, esta es:

$$J_n(x'') = -qD_n \frac{d\Delta n_p}{dx} = \frac{qD_n}{L_n} n_{p0} (e^{qV_A/kT} - 1) e^{-x''/L_n} \quad \text{(VIII-62)}$$

En $-x_p$ el valor será:

$$J_n(-x_p) = J_n(0'') = \frac{qD_n}{L_n} n_{p0} (e^{qV_A/kT} - 1) \quad \text{(VIII-63)}$$

Si se hace la representación de esta corriente, queda como se ve en la Figura VIII-23.

Como fue supuesto al principio, en la zona de vaciamiento la corriente es constante, por lo que se puede asumir que el valor de la corriente de portadores minoritarios de carga positiva en la zona de vaciamiento es $J_p(0')$ o $J_p(x_n)$ y la negativa es $J_n(0'')$ o $J_n(-x_p)$, como se ve en la figura VIII-24.

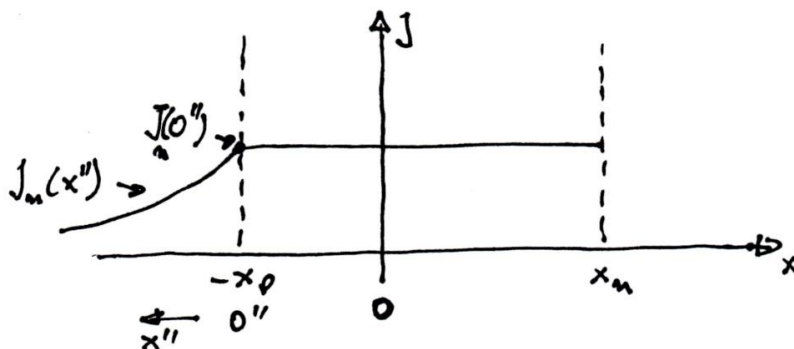


Figura VIII-23. Densidad de corriente de portadores minoritarios en la región p.

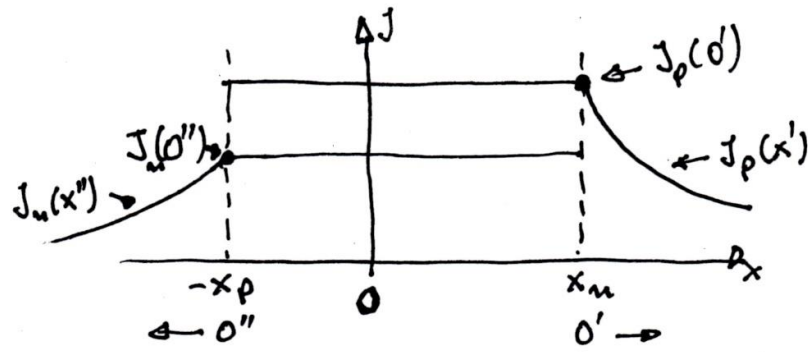


Figura VIII-24. Densidad de corriente de portadores minoritarios en las regiones p y n .

Una vez determinado el comportamiento de las corrientes de portadores minoritarios, se pueden determinar las corrientes de los mayoritarios. Se ha visto que la corriente total es constante, esta está formada por la suma de las corrientes debidas a los portadores mayoritarios y los minoritarios,

$$J = J_{\text{mayoritario}} + J_{\text{minoritario}}$$

Por lo que $J_{\text{mayoritario}} = J - J_{\text{minoritario}}$.

El grafico de las corrientes quedará como se ve en la Figura VIII-25.

Si se realiza la suma de ambas cantidades se obtiene:

$$J = J_p(0'') + J_n(0') = q \left[\frac{D_n}{L_n} n_{p0} + \frac{D_p}{L_p} p_{n0} \right] (e^{qV_A/kT} - 1) \quad (\text{VIII-63})$$

La ecuación (VIII-63) define la densidad de corriente total en la unión pn . Esta se puede reescribir de la siguiente manera:

$$J = J_s (e^{qV_A/kT} - 1) \quad (\text{VIII-64})$$

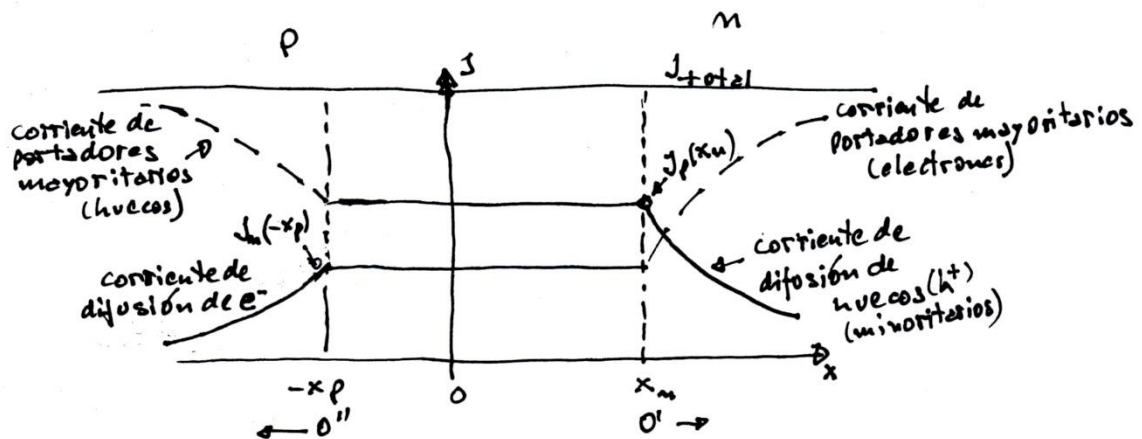


Figura VIII-25. Densidad de corriente, portadores mayoritarios y minoritarios en las regiones p y n .

Donde

$$J_s = q \left[\frac{D_n}{L_n} n_{p0} + \frac{D_p}{L_p} p_{n0} \right] \quad (\text{VIII-65})$$

La ecuación anterior es la densidad de corriente de saturación inversa. Si multiplicamos a esta por el área A transversal de la unión, se obtiene el flujo de corriente,

$$I_0 = qA \left[\frac{D_n}{L_n} n_{p0} + \frac{D_p}{L_p} p_{n0} \right] \quad \text{(VIII-66)}$$

La ecuación (VIII-64) es llamada la Ecuación del Diodo Ideal o Ecuación de Shockley. Esta define la relación entre la corriente y la tensión aplicada a la unión. Se puede ver en la ecuación (VIII-63) que cuando el sistema se encuentra bajo la aplicación de una polarización directa ($V_A > 0$) el término exponencial domina ampliamente frente al valor -1 . Por lo que la corriente crece en forma exponencial. Se puede escribir

$$J = J_s e^{qV_A/kT} \quad \text{(VIII-67)}$$

En el caso en que la unión se vea sometida a una polarización inversa ($V_A < 0$) el término exponencial tiene exponente negativo y a medida que crece el valor $-V_A$ y se va haciendo más negativo, menos incidencia tiene el exponencial frente al -1 , por lo que la ecuación (VIII-64) queda:

$$J = -J_s \quad \text{(VIII-68)}$$

El comportamiento de las corrientes lo podemos ver en la figura VIII-26.

Si se tiene en cuenta la Ley de Acción de Masas, y se considera que todos los átomos donores o aceptores han sido ionizados, la ecuación (VIII-65) se puede escribir de la siguiente manera:

$$J_s = q \left[\frac{D_n}{L_n N_a} + \frac{D_p}{L_p N_d} \right] n_i^2 \quad \text{(VIII-69)}$$

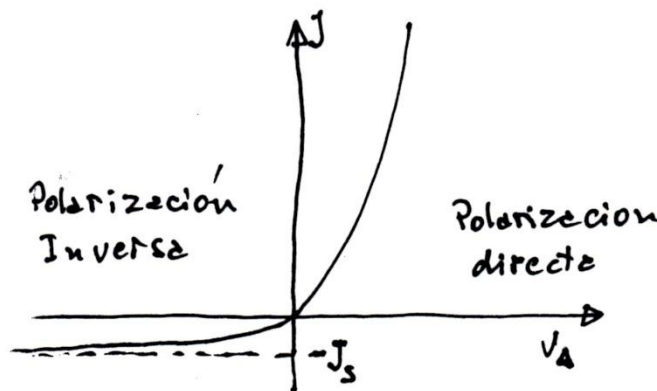


Figura VIII-26. Curva característica de corriente de una unión. Las escalas de la corriente están distorsionadas.

Una transformación semejante se puede hacer con la ecuación (VIII-66).

Ejemplo VIII-6. Con los parámetros que se han estado utilizando se puede determinar la corriente en la unión a temperatura ambiente ($T = 300\text{K}$). Se debe recordar que $N_a = 1 \times 10^{17} \text{cm}^{-3}$ y $N_d = 1 \times 10^{15} \text{cm}^{-3}$. Si el material de la unión es silicio (Si), entonces, $D_n = 25 \text{cm}^2/\text{seg}$, $D_p =$

$10\text{cm}^2/\text{seg}$, $n_i = 1,5 \times 10^{10}\text{cm}^{-3}$. Se considera que la vida media de ambos portadores es la misma, $\tau_n = \tau_p = 1 \times 10^{-7}\text{seg}$. La longitud de difusión puede ser determinada

$$L_n = \sqrt{D_n \tau_n} = 0,00158\text{cm}, \quad L_p = \sqrt{D_p \tau_p} = 0,00316\text{cm}$$

De la ecuación (VIII-69) se tiene:

$$J_s = 1.196 \times 10^{-11} \text{A/cm}^2$$

Si se tuviera un polarización inversa de $V_A = -0.1\text{V}$, se debe tener en cuenta que el valor de $kT/q = 0.0259\text{V}$ a $T = 300\text{K}$, por lo que el exponencial en la ecuación (VIII-64) vale,

$$e^{qV_A/kT} = 0,021, \quad e^{qV_A/kT} - 1 = -0.979 \quad \text{y} \quad J = -1,171 \times 10^{-11} \text{A/cm}^2$$

Mientras que si el potencial aplicado fuese positivo, $V_A = 0.1\text{V}$, estaríamos bajo una polarización en directa, el exponencial y la corriente de saturación valdrían:

$$e^{qV_A/kT} = 47,513, \quad e^{qV_A/kT} - 1 = 46,516 \quad \text{y} \quad J = 5,563 \times 10^{-10} \text{A/cm}^2$$

Se puede ver la diferencia existente entre los valores del exponencial con las distintas polarizaciones aplicadas, para una variación en sentido negativo y positivo del mismo valor de V_A , con respecto al eje de las corrientes ($V_A = 0$), los valores del exponencial son completamente diferentes. Si $V_A = -0.2\text{V}$, el valor del exponencial será $-4,43 \times 10^{-4}$, el valor de $J = -1,195 \times 10^{-11} \text{V/cm}^2$. Por otro lado, si $V_A = 0.2$, el exponencial valdrá: $2257,487$ y $J = 2.70 \times 10^{-8} \text{V/cm}^2$. Se puede ver que la corriente inversa se mantiene en valores extremadamente pequeños y semejantes, mientras que la corriente debida a la polarización directa aumenta a medida que aumenta el potencial en directa. En el ejemplo desarrollado, se ve que una décima de volt de aumento en el potencial hace que la corriente aumente dos órdenes de magnitud.

Otra conclusión que se puede extraer de estas expresiones y que es muy importante, es que si miramos la ecuación (VIII-69) se ve que ambos miembros del término de la derecha dependen de los dopados existentes, N_a y N_d , si $N_a \gg N_d$, el primer término de la ecuación (VIII-69) pierde significado frente al segundo, la ecuación quedaría:

$$J_s = q \left[\frac{D_p}{L_p N_d} \right] n_i^2 \quad (\text{VIII-70})$$

Si ocurriera al revés, $N_d \gg N_a$, entonces:

$$J_s = q \left[\frac{D_n}{L_n N_a} \right] n_i^2 \quad (\text{VIII-71})$$

Esto quiere decir que la corriente es transportada por los portadores minoritarios del material menos dopado.

Se puede tomar la ecuación (VIII-67) y aplicarle logaritmos, la cual queda:

$$\ln J = \ln J_s + \frac{q}{kT} V_A \quad (\text{VIII-72})$$

Se ve que el $\ln J$ queda como una función lineal dependiendo de la ordenada al origen, el $\ln J_s$, el logaritmo de la corriente de saturación, y el término lineal que depende del potencial aplicado V_A , cuya coeficiente determina el valor de la pendiente de la recta. En la Fig. VIII-27 se encuentra la representación gráfica de la ec. (VIII-72).

Como se puede ver todo el análisis está centrado sobre los portadores minoritarios. Como se ha mostrado previamente, dependiendo del tipo de polarización aplicada, la densidad de portadores minoritarios puede hacerse variar drásticamente. Se puede ver de aquí que los portadores en exceso o en defecto, y por lo tanto todas las características en el comportamiento de la unión que dependen de estas variaciones, están reguladas por los portadores minoritarios.

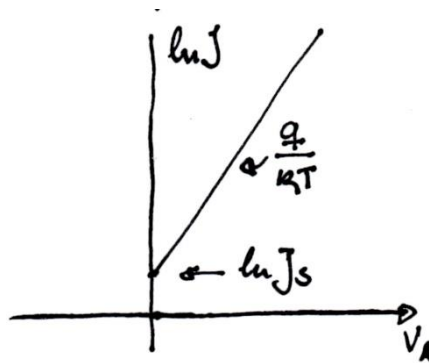


Figura VIII-27. Gráfica del logaritmo de la ecuación del diodo ideal.

VIII.5. El factor temperatura

En todos los desarrollos realizados hay un parámetro que está presente siempre y que hasta ahora no ha sido mencionado explícitamente, la temperatura. En los ejemplos mostrados hasta este momento, siempre se ha tomado a este parámetro como la temperatura ambiente, se menciona $T = 300\text{K}$, una constante.

En la realidad la temperatura juega un papel preponderante, en muchos casos la temperatura se presenta en forma directa, pero en otros casos se la encuentra dentro de otros parámetros.

Si se observa la ecuación (VIII-65) junto con al (VIII-69), la ecuación de Shockley, se tiene

$$J = q \left[\frac{D_n}{L_n N_a} + \frac{D_p}{L_p N_d} \right] n_i^2 (e^{qV_A/kT} - 1) \quad (\text{VIII-73})$$

Aquí la temperatura figura directamente en el exponente del exponencial, pero no se debe olvidar que también está presente en muchos de los parámetros mencionados.

En la ecuación (VI-26) se puede ver la expresión de n_i

$$n_i = \sqrt{N_C N_V} e^{-E_G/2kT}, \quad n_i^2 = N_C N_V e^{-E_G/kT}$$

La cual depende de un exponencial en cuyo exponente se encuentra T . De las ecuaciones (VI-6) y (VI-8) se ve que en forma genérica la densidad efectiva se puede escribir:

$$N_{C,V} = 2 \left[\frac{2\pi k m_{e,h}^*}{h^2} T \right]^{3/2}$$

Se observa que aquí la dependencia es con $T^{3/2}$. También el coeficiente de difusión es un parámetro que depende de T , de acuerdo a las ecuaciones (VII-31) y (VII-32), nuevamente en forma genérica

$$D_{n,p} = \mu_{n,p} (k/q) T$$

depende en forma lineal con T . Por último, la longitud de difusión, que se puede escribir, de acuerdo a las ecuaciones (VII-60) y (VII-61),

$$L_{n,p} = \sqrt{D_{n,p} \tau_{n,p}}$$

La cual depende de T a través del parámetro D_i .

Ejemplo VIII-7. Encontrar el valor de la corriente cuando $T = 400\text{K}$.

En este caso se tiene: $q/kT = 0.0345\text{V}^{-1}$. Se ejemplificará con ambas polarizaciones para comparar con el ejemplo anterior. Se tomará un potencial en directa, y los mismos valores de vidas medias que en el ejemplo anterior $\tau_n = \tau_p = 1 \times 10^{-7}\text{seg}$ como en el ejemplo VIII-6 de: $V_A = 0.2\text{V}$, esto hace que:

$$D_n = 33,33\text{cm}^2/\text{seg}, \quad D_p = 13,33\text{cm}^2/\text{seg}.$$

Esto hace que:

$$L_n = 0,0058\text{cm}, \quad L_p = 0,0037\text{cm}.$$

La densidad efectiva de estados N_C y N_V depende de las masas efectivas y de la temperatura. En el capítulo VI, para simplificar las deducciones se consideraron que las masas efectivas de huecos y electrones eran iguales, no obstante ello se mencionó que sucedería si fuesen diferentes. Como estas masas son diferentes, de debe considerar que las densidades efectivas son diferentes, se tomó el valor de $N_C = 2,51 \times 10^{19}\text{cm}^{-3}$ a $T = 300\text{K}$, para esta misma temperatura, de acuerdo al valor de la masa efectiva, $N_V = 1,04 \times 10^{19}\text{cm}^{-3}$. Por lo tanto, a $T = 400\text{K}$,

$$N_C = 3,86 \times 10^{19}\text{cm}^{-3}, \quad N_V = 1,60 \times 10^{19}\text{cm}^{-3}.$$

Por último, considerando $E_G = 1,12\text{eV}$, $n_i = 6,4 \times 10^9\text{cm}^{-3}$ a $T = 300\text{K}$

Se debe hacer una aclaración en este punto. El valor aceptado para este parámetro es $1,5 \times 10^{10} \text{cm}^{-3}$. La divergencia surge por varios motivos, una muy buena discusión sobre esto se encuentra en el libro de Neamen (cap. 3).

Tomado el valor aceptado para n_i , se tiene a $T = 400\text{K}$, $n_i = 3,37 \times 10^{12} \text{cm}^{-3}$.

Si se calcula la corriente de acuerdo a la ecuación (VIII-73) (se ha despreciado el -1) se tiene:

$$J = 0,00022 \text{ A/cm}^2$$

Comparando con el valor a $T = 300\text{K}$ del ejemplo anterior, $J = 2,70 \times 10^{-8} \text{ V/cm}^2$, se ve que la corriente ha crecido notablemente, como se ve en la figura VIII-28.

Si se aplica una polarización inversa con $V_A = -0,2\text{V}$, la corriente a $T = 400\text{K}$, es (se ha despreciado el valor del exponencial frente al -1):

$$J = 6,618 \times 10^{-7} \text{ A/cm}^2$$

La corriente a $T = 300\text{K}$, que se calculó fue de $J = -1,195 \times 10^{-11} \text{ V/cm}^2$, se puede ver que la corriente

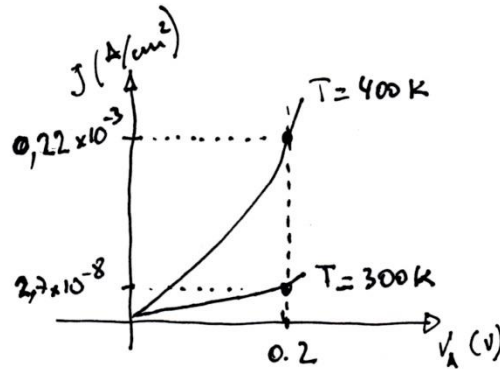


Figura VIII-28. Comparación de corrientes con polarización directa en función de la temperatura. Gráfico fuera de escala.

inversa con el incremento de la temperatura también ha incrementado su valor en casi 4 órdenes de magnitud, estas características han sido graficadas en la figura VIII-29.

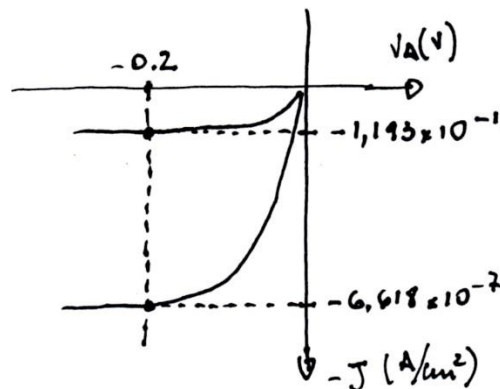


Figura VIII-29. Comparación de corrientes con polarización inversa en función de la temperatura. Gráfico fuera de escala.

VIII.6. Heterojunturas

En capítulos previos se estudió la unión de un mismo material semiconductor en toda la estructura con dos regiones de dopados diferentes, lo que se llama homounión (ejemplo Si-p/Si-n, o Ge-p/Ge-n). A continuación, en los subcapítulos anteriores se ha empezado a ver el problema de las heterouniones, esto es cuando dos materiales diferentes se ponen en contacto. Estas uniones tienen propiedades totalmente diferentes de las que tienen las homouniones. Se verá a continuación la unión metal-semiconductor, estos dos materiales presentan propiedades físicas no análogas, lo que hace que la unión tenga otras propiedades que las hacen atractivas para su uso en la industria electrónica. Se verán en los subcapítulos siguientes, las uniones Schottky y la no rectificante u óhmica, que son ejemplos de heterouniones. También, en este subcapítulo se desarrollará la unión de dos semiconductores diferentes con dopados diferentes. Cuando se piensa en uniones con diferentes semiconductores se debe ser cuidadoso, pues no cualquier semiconductor se puede unir con otro, sino que se deben cumplir algunas condiciones, deben tener anchos de bandas prohibidas diferentes, la variación relativa de las constantes de red de las redes cristalinas de uno y otro semiconductor debe ser pequeña para que una y otra estructura puedan unirse, tal es el caso de la unión GaAs/AlGaAs, muy utilizada en dispositivos optoelectrónicos, el primero tiene un constante de red de $5,65\text{Å}$ ($5,65 \times 10^{-10}\text{m}$), mientras que en el segundo semiconductor es de $5,658\text{Å}$ ($5,658 \times 10^{-10}\text{m}$), la relación porcentual entre constantes de red es del 0,14%.

VIII.7. Unión metal – semiconductor. Unión Schottky

Toda operación que se realiza dentro de un dispositivo semiconductor es para dar respuesta a algún proceso que lo involucra. Claramente, se modifica el estado del dispositivo mediante alguna señal de entrada, variaciones de corrientes o del potencial externo. Esto hace que de acuerdo a las características del semiconductor la respuesta tenga alguna particularidad. Necesariamente, el resultado de esta operación debe conectarse con el medioambiente exterior al dispositivo. En general la unión metal – semiconductor no tiene las características de la unión semiconductor – semiconductor. El transporte de portadores es unipolar y su respuesta es más rápida. Este tipo de unión es llamada diodo Schottky, puede verse en la figura VIII-30.

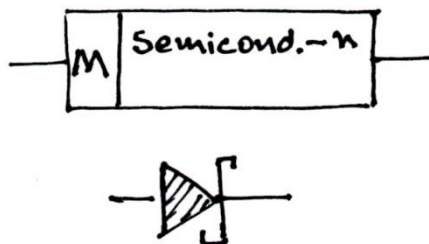
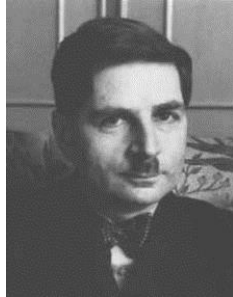


Figura VIII-30. Esquema del diodo Schottky y símbolo que lo representa.

La conexión se debe hacer mediante un medio conductor, está claro que el semiconductor no es la mejor respuesta a este requerimiento, la respuesta está en que la conexión debe ser realizada mediante un metal, los cuales, en general, son conductores naturales debido a sus características físicas.



Walter Hermann Schottky (Suizo-Alemán)
 23-6-1886, 4-3-1976
 Dr. en Física, 1915- inventó: pantalla de la red de tubo de vacío, 1919- válvula termoiónica tetrodo, trabajando en Siemens.

Para llevar adelante este punto, se pondrá un contacto un semiconductor tipo n con una concentración de impurezas del orden de $10^{21}m^{-3}$ (los valores mencionados son sugeridos, solamente para tener idea del orden de las magnitudes con que se está trabajando) y un metal cuya función trabajo será mayor que la del semiconductor, es decir $\phi_M > \phi_S$, como se ve en la figura VIII-31. La magnitud χ_s es la afinidad electrónica de semiconductor. En la Tabla 1, pueden verse las funciones trabajo de algunos elementos, así como también las afinidades electrónicas. Se debe tener en cuenta que el nivel de vacío es el nivel de referencia. Bajo estas circunstancias, para que el sistema metal-semiconductor llegue al equilibrio, habrá un pasaje de electrones del semiconductor al metal, el nivel de Fermi del metal aumentará y el del semiconductor disminuirá, debido a la ganancia de electrones del primero y la pérdida del segundo. El proceso continuará hasta que los niveles de Fermi de ambos materiales se equiparen. Cuando esto ocurre, el metal tendrá un exceso de carga negativa debido a los electrones que fueron entregados por el semiconductor, mientras que el semiconductor habrá quedado positivamente cargado por la pérdida de portadores negativos.

A medida que se produce el pasaje de portadores del semiconductor al metal, se va generando una diferencia de potencial entre los materiales que da lugar a una barrera de potencial. Cuando se llega al equilibrio termodinámico, la altura de la barrera de potencial será:

Elemento	Función Trabajo (V)
Ag Plata	4,26
Al Aluminio	4,28
Au Oro	5,1
Cr Cromo	4,5
Fe Hierro	4,81
Mo Molibdeno	4,6
Ni Níquel	5,15
Pd Paladio	5,12
Pt Platino	5,65
Ti Titanio	4,33
W Tungsteno	4,55

Elemento	Afinidad electrónica (V)
Ge Germanio	4,13
Si Silicio	4,01
GaAs Arseniuro de Galio	4,07
AlAs Arseniuro de Aluminio	3,5

Tabla 1. Funciones trabajo y afinidades electrónicas.

$$\varphi_0 = eV_C = e\phi_S - e\phi_M \tag{VIII-74}$$

Donde V_C es el potencial de contacto (generalmente del orden de 1V).

El contacto que se establece entre ambos materiales es de una distancia del orden de una constante de red, $d = 5 \times 10^{-9} \text{m}$. Al hacer esto se está enfrentando a dos materiales que tiene cantidades de electrones totalmente diferentes. En un metal, la densidad superficial de electrones es de aproximadamente $n = \epsilon_0 \times V_C / e \times d \sim 1,1 \times 10^{16} \text{e/m}^2$, mientras que en el semiconductor la densidad superficial de átomos donores es, $n_s \sim 10^{21} / 2 \times 10^9 \sim 5 \times 10^{11} \text{m}^{-2}$ (donde 2×10^9 es la cantidad de capa atómicas que habría en un cubo de 1m de lado de silicio, considerando el espaciado entre capas de $5,43 \times 10^{-10} \text{m}$). Se considerará que esta cantidad es equivalente en las capas atómicas subsiguientes en el semiconductor. Para que ambos niveles de Fermi se igualen, la cantidad de electrones que ponen en juego ambos, metal y semiconductor deben ser iguales, ver Fig. VIII-32.

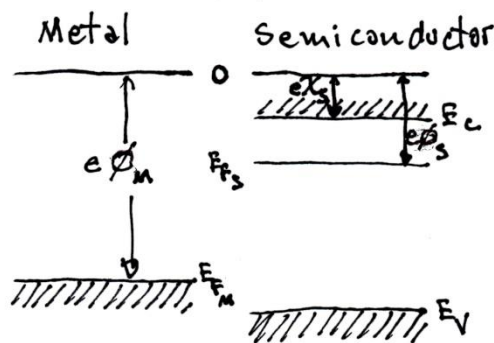


Figura VIII-31. Diagrama de funciones trabajo del metal y del semiconductor-n.

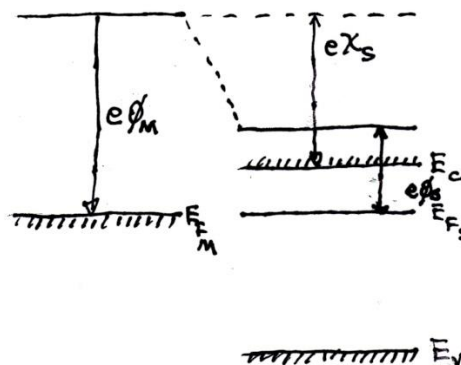


Figura VIII-32. Diagrama de funciones trabajo del metal y del semiconductor-*n* con niveles de Fermi igualados.

Para que esto ocurra, el semiconductor deberá ionizar 2×10^5 capas atómicas para igualar los 10^{17} electrones por unidad de superficie que tiene el metal. Esto implica que el semiconductor será “vaciado” de electrones en una profundidad d_s . El esquema de esta distribución se ve en la Fig. VIII-33. Esta circunstancia es similar a lo que ocurría con la unión *pn* cuando uno de los semiconductores estaba más dopado que el otro, la zona de carga era mucho más extensa en el menos dopado. En este tipo de unión, el metal pone en juego los electrones de una capa superficial, mientras que el semiconductor debe involucrar una cantidad importante de capas atómicas, la carga positiva generada por la ionización hace que los niveles de Fermi se igualen.

Esta capa de carga positiva es llamada “capa de barrera” o como en los semiconductores “capa de vaciamiento”. La diferencia de cargas entre ambos materiales

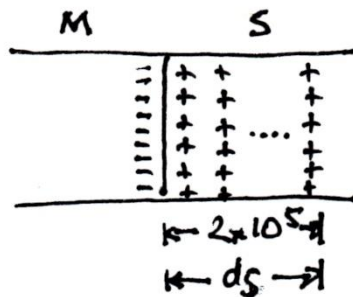


Figura VIII-33. Esquema de distribución de cargas eléctricas en la unión.

genera, como se mencionó, un potencial de contacto y un campo eléctrico, el cual caerá mayoritariamente sobre el semiconductor. Esto hace que las bandas de energía se curven, tal como ocurría en la unión de semiconductores, indicando que tipo de dopado existe en el semiconductor, como se ve en la Figura VIII-34.

El campo de contacto se puede escribir como:

$$\epsilon_c = \varphi_0 / qd_s \tag{VIII-75}$$

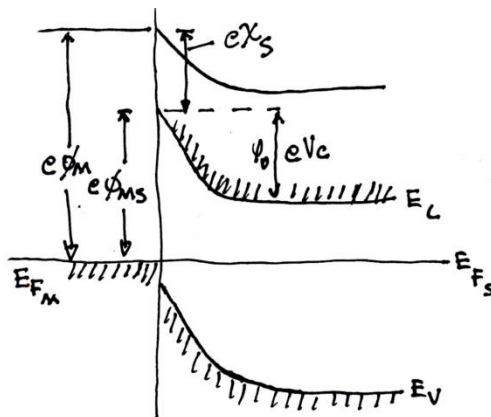


Figura VIII-34. Esquema de la estructura de bandas de la unión metal-semiconductor-*n*.

Con valores como los mencionados, con $\varphi_0 = 1eV$, y suponiendo un $d_s = 10^{-7}m$, se tiene que $\varepsilon_c \sim 10^{-7} V/m$, un campo de este orden de magnitud no afecta la estructura energética del sistema. Solamente, como se observa en la Fig. VIII-34, produce una inclinación en las bandas. Si este no existiera, las bandas serían planas como se ve en la figura VIII-31. La existencia del campo hace que las bandas en el semiconductor en las inmediaciones de la unión sean dependientes de la distancia (en nuestro modelo unidimensional dependerán de la variable x). Debido a la relación de funciones trabajo del metal y el semiconductor, los electrones del semiconductor tienden a ir hacia el metal. A medida que el electrón se acerca a la unión ($x = 0$) empieza a ser repelido, para llegar a este punto debe ejercer una fuerza que venza a la repulsión, esto hace que se genere un trabajo, e incremente su energía potencial, $\varphi(x)$, la cual es nula lejos de la unión y máxima sobre ella. Para determinar la forma de $\varphi(x)$ se debe utilizar la ecuación de Poisson.

$$\frac{d^2V}{dx^2} = -\frac{\varphi(x)}{\epsilon_0\epsilon} \tag{VIII-76}$$

Se debe recordar que en el segundo miembro $\varphi(x)$ representa la carga eléctrica existente. La ecuación (VIII-76) se puede reescribir, recordando que $\varphi(x) = -eV(x)$,

$$\frac{d^2\varphi}{dx^2} = -\frac{e\varphi(x)}{\epsilon_0\epsilon} \tag{VIII-77}$$

Si se supone que todos los átomos donores están ionizados, es decir: $\varphi = eN_d$, tenemos:

$$\frac{d^2\varphi}{dx^2} = \frac{e^2N_d}{\epsilon_0\epsilon} \tag{VIII-78}$$

Para obtener la solución se deben establecer condiciones de contorno,

$$\varphi(d_s) = 0, \quad \frac{d\varphi}{dx} = 0 \text{ (en } x = d_s) \tag{VIII-79}$$

Bajo estas condiciones, si se integra la ec. (VIII-78) dos veces, se tiene:

$$\varphi(x) = \frac{e^2}{2\epsilon_0\epsilon} N_d (d_s - x)^2 \tag{VIII-80}$$

Se puede ver aquí, que esta es una ecuación de segundo grado, que existe un decrecimiento parabólico a medida que el valor de x aumenta. En $x = 0$,

$$\varphi(0) = \varphi_0 = e\phi_S - e\phi_M$$

Sustituyendo en la ec. (VIII-80) se tiene:

$$d_s = \left(\frac{2\epsilon_0\epsilon\varphi_0}{N_d e^2} \right)^{1/2} = \left(\frac{2\epsilon_0\epsilon V_c}{N_d e} \right)^{1/2} \tag{VIII-81}$$

Este parámetro dice cuanto penetra el campo dentro del semiconductor, se puede ver que depende de otros dos parámetros, φ_0 y N_d . En la medida que φ_0 aumente y N_d decrezca, la penetración aumentará.

Utilizando la relación entre el campo y el potencial electrostático

$$\mathcal{E} = -dV/dx$$

De la ecuación (VIII-80) se obtiene:

$$\mathcal{E} = -\frac{e}{\epsilon_0 \epsilon} N_d (d_s - x) \quad \text{(VIII-82)}$$

Esta ecuación muestra, en un semiconductor uniformemente dopado, que el campo es lineal con la distancia, y alcanza su valor máximo en la interface metal-semiconductor. Aunque el campo eléctrico es nulo dentro del metal, existe necesariamente una carga negativa en el metal en la unión metal-semiconductor.

Ejemplo VIII-8. Obtener los valores de d_s para diferentes N_d .

- 1.- Para $N_d = 10^{21} \text{m}^{-3}$ y $\varphi_0 = 1 \text{eV}$, $\epsilon = 10$, $d_s = 1,05 \times 10^{-6} \text{m}$.
- 2.- Si se toma $N_d = 10^{15} \text{m}^{-3}$, $d_s = 1,05 \times 10^{-3} \text{m}$

Conociendo la variación del potencial en función de x , se puede determinar cómo varía la concentración de electrones en la banda de conducción con la distancia.

$$n(x) = N_c e^{-((E_F + \varphi(x))/kT)} = N_c e^{-E_F/kT} e^{-\varphi(x)/kT} = n_{n0} e^{-\varphi(x)/kT} \quad \text{(VIII-83)}$$

Donde n_{n0} es la concentración de equilibrio de electrones en el semiconductor.

El caso que se planteó fue cuando $\chi_M > \chi_S$, en el caso en que ocurriera lo inverso, es decir $\chi_M < \chi_S$, los electrones se moverían del semiconductor al metal, se produciría una acumulación de portadores en el semiconductor, la densidad estaría representada por:

$$n(x) = n_{n0} e^{\varphi(x)/kT} \quad \text{(VIII-84)}$$

Esta capa es llamada “antibarrera”.

VIII.7.a. Unión metal – semiconductor. Aplicación de un potencial externo.

Hasta este punto se ha visto una unión metal – semiconductor sobre la cual no se había aplicado ningún potencial externo para modificar las condiciones de la unión. En este caso, la barrera de potencial que ven los electrones (ver Figura VIII-34) en la banda de conducción del semiconductor para moverse hacia el metal es V_C , en cambio la barrera de potencial que ven los electrones del metal que tratan de ir hacia el semiconductor es:

$$\phi_{MS} = \phi_M - \chi_S \tag{VIII-85}$$

Ahora se verá cuáles son los efectos de la aplicación de un potencial externo (V_A) tal cual se lo vio en la unión semiconductor – semiconductor. Si se aplica una tensión positiva o negativa al semiconductor con respecto al metal, la altura de la barrera de potencial disminuirá o aumentará, pero el valor de ϕ_{MS} se mantendrá casi constante.

El potencial aplicado puede ser en forma directa o inversa. En la conexión en directa, se aplica un voltaje positivo del metal con respecto al semiconductor, con lo cual se cargará más negativamente al semiconductor, haciendo que los valores de los niveles de energía aumenten como se ve en la Figura VIII-35a y se reduzca la barrera de potencial.

Por el contrario, si se aplica un voltaje positivo al semiconductor con respecto al metal la altura de la barrera se incrementa, mientras que como se dijo anteriormente, ϕ_{MS} se mantendrá casi constante, como se ve en el diagrama en la figura (VIII-35b).

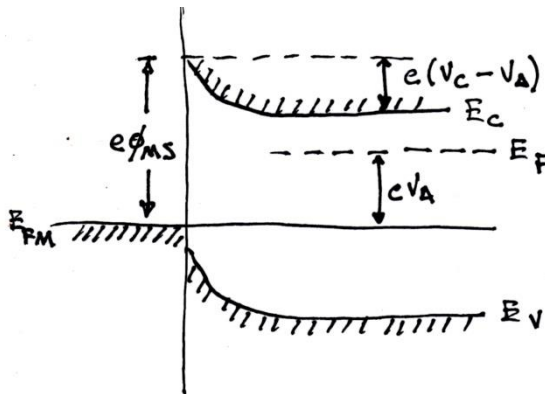


Figura VIII-35a. Estructura de bandas de una unión metal-semiconductor-n con un potencial en directa aplicado.

Debido a la aplicación de un potencial externo, V_A , se reduce o incrementa la altura de la barrera de potencial. Esto hace, como en el caso de la unión de semiconductores, que la capa de vaciamiento se reduzca o se incremente. En la ecuación (VIII-86) se muestra la relación con los potenciales existentes:

$$d_S = (2\epsilon_0\epsilon(V_C \mp V_A) / eN_d)^{1/2} \tag{VIII-86}$$

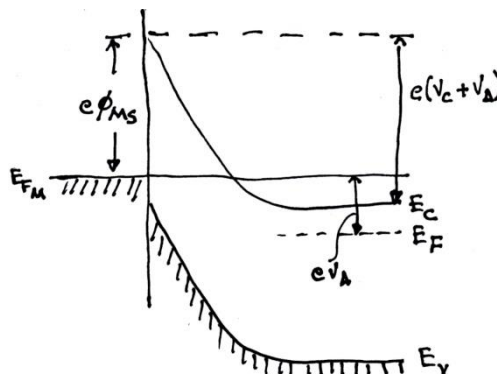


Figura VIII-35b. Estructura de bandas de una unión metal-semiconductor-n con un potencial en inversa aplicado.

Cuando se aplica un potencial en directa (signo negativo) el ancho de la zona de vaciamiento se reduce, mientras que cuando el potencial aplicado esta en inversa (signo positivo en la ecuación (VIII-86)), la zona de vaciamiento se incrementa.

VIII.7.b. Unión Schottky. Corrientes a través de la unión.

Cuando la unión metal-semiconductor se encuentra en equilibrio, el flujo de electrones del metal al semiconductor y viceversa está balanceado, ambas corrientes son iguales, la corriente total es nula. Cuando se aplica un potencial, el principal mecanismo para que los electrones atraviesen la barrera de potencial es el de la emisión termoiónica (Ver apéndice VIII-1). Los electrones que se mueven por medio de este mecanismo, deben tener una energía mayor que la barrera existente $e(V_C - V_A)$ para poder sobrepasarla, como se ha visto V_C es el potencial de contacto y V_A el potencial aplicado. Este último puede ser positivo o negativo dependiendo del tipo de polarización aplicada, esto implica que la altura de la barrera cambiará, y eso hará que la cantidad de electrones inyectados también cambie. Estas variaciones pueden verse en las Figuras VIII-35a y b.

En la Figura VIII-35a el metal es positivo con respecto al semiconductor, hay un flujo neto de portadores desde el semiconductor hacia el metal. En la Figura VIII-35b, el metal es negativo con respecto al semiconductor, el flujo de corriente se mueve en sentido opuesto. Se va a suponer el pasaje de los portadores por efecto túnel es despreciable. El proceso por el cual un electrón se mueve desde el semiconductor hacia el metal, es el de la difusión, la altura de la barrera permitirá el pasaje a aquellos electrones con mayor energía que la barrera.

La cantidad de electrones con suficiente energía para saltar la barrera es:

$$n_B = n_0 e^{\left[-\frac{e(V_C - V_A)}{kT}\right]} \quad (\text{VIII-87})$$

Se puede poner la densidad electrónica en la región neutra en función de la densidad efectiva de estados en la banda de conducción, N_C

$$n_0 = N_C e^{\left[-\frac{(E_C - E_{Fs})}{kT}\right]} \quad (\text{VIII-88})$$

La densidad en la zona de contacto (barrera) es:

$$n_B = N_C e^{\left[-\frac{e(\phi_{MS} - V_A)}{kT}\right]}$$

En 1944 Hans A. Bethe propone una teoría donde la corriente en el diodo metal-semiconductor está limitada por la emisión termoiónica, con esta teoría se obtiene la dirección correcta de la corriente. La expresión que propuso fue:

$$I_{sm} = \frac{e}{4} A \langle v \rangle N_C e^{\left[-\frac{e(\phi_{MS} - V_A)}{kT}\right]} \quad (\text{VIII-89})$$



Hans Albrecht Bethe
2-7-1906, 6-3-2005

Premio Nobel de Física 1967. “por sus contribuciones sobre la teoría de las reacciones nucleares, especialmente, acerca de sus descubrimientos sobre la producción de energía en las estrellas”

Donde A es el área transversal que atraviesa la corriente (área del dispositivo). Para esto consideró que los electrones se mueven aleatoriamente y la cantidad que impacta sobre la barrera metal-semiconductor es: $\langle v \rangle n_B / 4$, $\langle v \rangle$ es la velocidad promedio de los electrones.

Como se mencionó más arriba, si $V_A = 0$ las corrientes se encuentran balanceadas, es decir:

$$I_{ms} = I_{sm} = \frac{e}{4} A \langle v \rangle N_c e^{-e\phi_{MS}/kT} \quad (\text{VIII-90})$$

Al aplicar un potencial externo, los electrones del metal, que siempre ven una barrera de potencial que prácticamente no cambia, generan un flujo que tampoco cambia, la corriente del metal hacia el semiconductor es la corriente generada por la emisión termiónica, la corriente $I_{ms} = I_s$. Si se combinan las ecuaciones (VIII-89) y (VIII-90) se tiene:

$$I = I_{sm} - I_{ms} = I_s \left[e^{(eV_A/kT)} - 1 \right] \quad (\text{VIII-91})$$

Pero a diferencia de los electrones que van del metal al semiconductor, los del semiconductor hacia el metal ven una barrera variable.

En el Capítulo IV se vio como se podía determinar la velocidad media de una partícula de acuerdo a la ley de distribución de Maxwell-Boltzmann. Como se supone que los electrones en la banda de conducción se están moviendo en este marco, se puede determinar su velocidad media,

$$\langle v \rangle = \left(\frac{8kT}{\pi m^*} \right)^{1/2} \quad (\text{VIII-92})$$

Si reemplazamos el valor de N_c de la ecuación (VI-6) en (VIII-90), teniendo en cuenta la ecuación (VIII-92), se tiene:

$$I_s = 2A \left(\frac{4em^*k^2}{\pi^2 \hbar^3} \right) T^2 e^{-e\phi_{MS}/kT} \quad (\text{VIII-93})$$

La cantidad entre paréntesis es la constante de Richardson, R , cuyo valor es:

$$R = \frac{4em^*k^2}{\pi^2\hbar^3} = 1,2 \times 10^6 \text{ A/m}^2\text{K}^2 \quad (\text{VIII-94})$$

Donde se ha considerado a m^* como la masa en reposo del electrón.

La ecuación (VIII-93) muestra que la corriente de saturación es la corriente producida por la emisión termiónica. Sobre la barrera, esta es mucho más intensa que la corriente de saturación en la unión $p-n$ para iguales tensiones aplicadas. Como resultado de esto, se pueden tener corrientes inversas mayores con la aplicación de potenciales menores.

La curva característica $I-V$ es como la que se muestra en la figura VIII-36.

La curva de la figura VIII-36 está dada por:

$$I = I_s \left[e^{-\left(\frac{eV_A}{nkT}\right)} - 1 \right] \quad (\text{VIII-95})$$

La cual es idéntica a la que se vio en la unión pn . La corriente inversa es la corriente de emisión

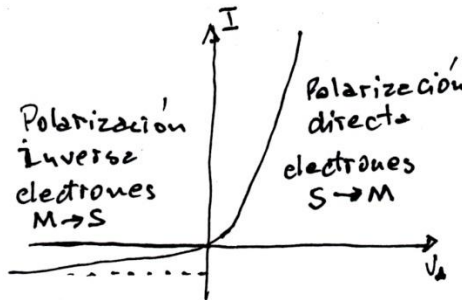


Figura VIII-36. Curva característica de la unión metal-semiconductor- n .

termiónica y el parámetro n es el factor de idealidad, el cual hace referencia a cuán cerca o lejos se está de la unión ideal. En el caso de estas uniones, el factor de idealidad se encuentra muy próximo a la unidad debido a que la participación de portadores minoritarios en la corriente es casi inexistente, son los portadores mayoritarios los que dominan la situación.

Es bueno remarcar lo que se ha mencionado hasta ahora, hechos que establecen las diferencias entre ambas uniones, Semiconductor-Semiconductor (S-S) y Metal-Semiconductor (M-S). En la primera la corriente de saturación está dominada por la difusión de portadores minoritarios. En la segunda, por la corriente termiónica formada por portadores mayoritarios. También se debe recordar que para una misma corriente de saturación, los potenciales en la unión M-S son mucho menores.

Ejemplo VIII-9. Se desea ver cuál es la diferencia del potencial aplicado para tener una corriente de 5 A/cm^2 para una unión pn y para una unión Schottky.

En el ejemplo VIII-6 se obtuvo la corriente de saturación, con los datos allí mencionados, de una unión pn , $J_s = 1.196 \times 10^{-11} \text{ A/cm}^2$.

Si tuviéramos una unión metal-semiconductor de oro-silicio, la altura de la barrera de potencial es $\phi_{MS} = 0,79 \text{ V}$.

La corriente de saturación generada por la corriente termoiónica será (sin tener en cuenta la existencia de los spines del electrón, se elimina el 2 que multiplica a toda la expresión):

$$J_s = RT^2 e^{(-e\phi_{MS}/kT)} = 120x(300^2)e^{(-0.79/0.026)} = 6,87x10^{-7} \text{ A/cm}^2$$

Como la polarización aplicada es directa, se puede omitir el -1 de las expresiones de la corriente. Por este motivo utilizaremos al ecuación (VIII-67) de la cual despejaremos el valor de V_A y lo mismo se hará con la ecuación (VIII-95). Para la unión pn se tiene:

$$V_{A-pn} = (kT/e)\ln(J/J_s) = 0,0259x\ln(5/1,196x10^{-11}) = 0,693V$$

En el caso de la unión metal-semiconductor:

$$V_{A-MS} = (kT/e)\ln(J/J_s) = 0,0259x\ln(5/6,87x10^{-7}) = 0,409V$$

Como se mencionó más arriba, para obtener una misma corriente se necesita aplicar un potencial mucho menor que en la unión pn .

VIII.8. Union no rectificante (Metal-Semiconductor). Contacto Óhmico

Las uniones vistas hasta ahora eran rectificantes, esto quiere decir que podían modificar la corriente que atravesaba la unión. Tanto en el caso de la unión pn o de la metal-semiconductor vista recientemente, en las junturas se generan barreras de potencial que impiden el libre pasaje de huecos y electrones, o electrones o huecos, dependiendo del tipo de unión.

Sin lugar a dudas es deseable tener una unión que permita el pasaje de portadores en forma libre. Para esto se debe buscar una unión que tenga una curva característica que se comporte en forma lineal, como se ve en la figura VIII-37. Este tipo de unión se llama óhmica.

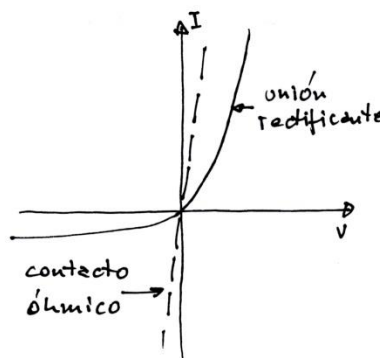


Figura VIII-37. Comparación curvas características, unión rectificante y contacto y contacto óhmico.

Existen varios mecanismos para construir uniones óhmicas, se ha visto que el campo en junturas metal semiconductor se desarrolla completamente sobre el semiconductor, la expresión (VIII-86) da la extensión de la zona de vaciamiento en el semiconductor, se ha puesto $V_A = 0$ por simplicidad,

$$d_s = \left(\frac{2\epsilon_0\epsilon V_c}{eN_d} \right)^{1/2} \tag{VIII-96}$$

Si el semiconductor está altamente dopado $N_d \gg 1$, el valor d_s , adquiere valores sumamente pequeños, la zona es extremadamente delgada. No obstante ello, la barrera de potencial existe, pero al ser tan delgada, los electrones pueden atravesarla por efecto túnel sin demasiado inconveniente.

La descripción realizada para la figura VIII-34 bajo estas condiciones de alto dopado se modificaría para quedar como en la figura VIII-38.

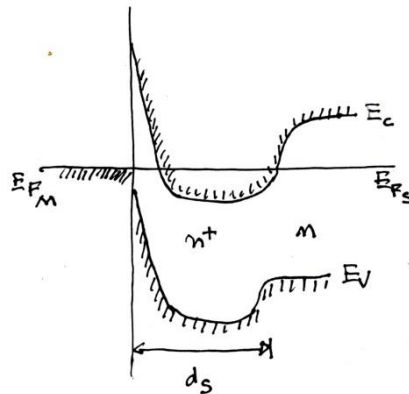


Figura VIII-38. Diagrama de bandas de la unión metal-semiconductor para un contacto óhmico.

Ejemplo VIII-10. Retomemos el ejemplo VIII-8 se obtuvo:

Para $N_d = 10^{21} \text{m}^{-3}$ y $\phi_0 = 1 \text{eV}$, $\epsilon = 10$, $d_s = 1,05 \times 10^{-6} \text{m} = 1,05 \times 10^{-4} \text{cm}$.

El dopado que se menciona no es muy alto, si lo traducimos a cm^{-3} , ($1 \text{m}^3 = 10^6 \text{cm}^3$), el dopado $N_d = 10^{15} \text{cm}^{-3}$. Se dopará al semiconductor con una densidad $N_d = 10^{18} \text{cm}^{-3}$.

$$d_s = \left(\frac{2 \times 11,8 \times 8,85 \times 10^{-14} \times 1}{(1,62 \times 10^{-19} \times 10^{18})} \right)^{1/2} = 3,59 \times 10^{-6} \text{cm}$$

$$= 3,59 \times 10^{-8} \text{m} = 359 \times 10^{-10} = 359 \text{ \AA}$$

Se ve que el incremento en el dopado hizo reducir en dos órdenes de magnitud el ancho de la zona de vaciamiento. Se puede ver qué pasa cuando se incrementa el dopado a 10^{20}cm^{-3} . En este caso, $d_s = 35,9 \text{ \AA}$, se ve que el ancho de la zona se reduce en un orden de magnitud.

VIII.9. Heterojuntura formada por la unión de semiconductores

Las uniones entre semiconductores diferentes tienen características variadas, en este texto solamente se hará referencia a la que uno de los semiconductores tiene la banda prohibida completamente superpuesta al otro (straddling), como se indica en la figura VIII-39.

Claro está que el semiconductor A puede estar dopado n y el B puede ser p , o al revés, $A-p$, $B-n$. También puede ocurrir que ambos A y B estén dopados con dopante n o p pero uno de ellos con un dopado mucho mayor que el otro, es decir, $A-n^+$ / $B-n$, o al revés y lo mismo puede ocurrir con dopados tipo- p .

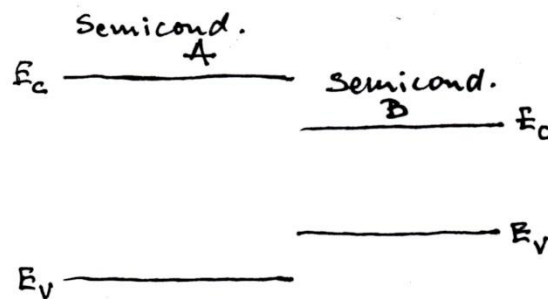


Figura VIII-39. Bandas de semiconductores.

Cuando se tiene este tipo de materiales que se desea unir, se debe saber cómo se conforman las bandas del dispositivo después que se establece la unión. En la figura VIII-40 se grafica una secuencia de gráficos donde se muestra como sería el proceso de conformación.

En la Figura VIII-40, se puede ver como se forma la estructura de banda de la unión de dos semiconductores con anchos de banda disimiles. En la parte superior de la Figura se ve cual sería la estructura de banda de los semiconductores A y B, donde A tiene un ancho de banda mayor que B, la línea punteada indica cuales serían los niveles de conducción y valencia si no estuviera el otro semiconductor. La línea a trazos vertical indica la zona de unión. En el gráfico del medio se muestran las bandas de energía de los semiconductores, previo a la unión. En la parte inferior, se ve la estructura de banda final que adopta el sistema cuando los semiconductores han sido unidos y tienen anchos de banda diferentes.

En la Figura VIII-41, se muestran los parámetros que intervienen en estas uniones, en la (a) antes y en la (b) después de la unión.

Se debe pensar que cada uno de los materiales tiene propiedades diferentes, el semiconductor A tiene una banda prohibida E_{gA} , una afinidad electrónica χ_A (diferencia entre la

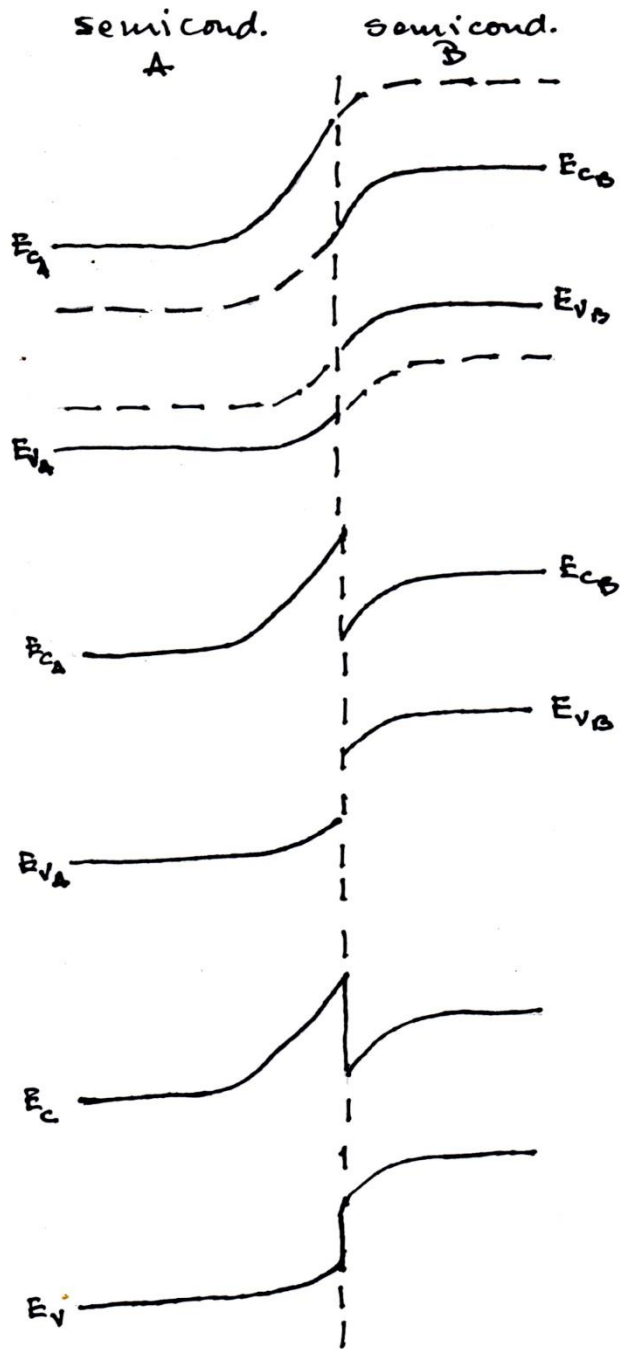


Figura VIII-40. Secuencia de formación de la estructura de banda de la unión de dos semiconductores.

banda de conducción y la energía del electrón libre) y una constante dieléctrica ϵ_A , mientras que para el semiconductor B se tiene: E_{gB} , χ_B y ϵ_B .

Como los materiales tienen distintos anchos de banda, se plantean dos discontinuidades, una en la banda de conducción ΔE_c y la otra en la de valencia, ΔE_v . Se puede determinar la diferencia entre las bandas prohibidas,

$$\Delta E_g = E_{gA} - E_{gB} = \Delta E_c + \Delta E_v \quad (\text{VIII-97})$$

Por otro lado, si se observa la Figura VIII-41a, se puede ver que

$$\Delta E_g = q(\chi_B - \chi_A) \tag{VIII-98}$$

Se puede establecer una similitud entre estas uniones y la *pn* de manera de obtener cantidades como el potencial de contacto o el ancho de la zona de vaciamiento. Si se remite a los inicios de este capítulo, se verá que el potencial de contacto se puede determinar en función de las diferencias de los niveles de conducción, valencia y de Fermi. Esto es para este caso:

$$qV_{bi} = E_{gB} - (E_F^p - E_v^p) - (E_c^n - E_F^n) + \Delta E_c \tag{VIII-99}$$

Si se observan las ecuaciones (VIII-8), (VIII-9) con los índices adecuados a la presente situación, se pueden hacer sustituciones de los valores entre paréntesis, obteniéndose:

$$V_{bi} = \frac{1}{q}(\Delta E_c + E_{gB}) - \left(\frac{kT}{q}\right) \ln \left(\frac{N_{cA}N_{vB}}{n_A p_B}\right) \tag{VIII-100}$$

Donde N_{cA} y N_{vB} son las densidades en la banda de conducción y de valencia de los semiconductores A y B, y los parámetros n_A y p_B , son las concentraciones electrónicas de los portadores (electrones y huecos) de los semiconductores A y B, respectivamente. En caso de haber una total ionización, los valores n_A y p_B coincidirían con los N_{cA} y N_{vB} , respectivamente.

Debido a la discontinuidad de la estructura de bandas, en la unión de ambos materiales ($x = 0$) los valores del campo eléctrico ε no serán continuos.

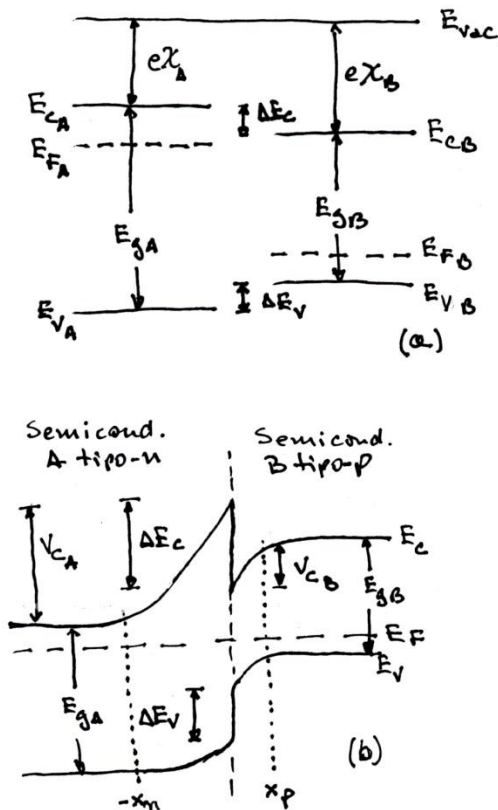


Figura VIII-41. Parámetros de la unión S-S (no se respetan las relaciones entre las medidas) (a) antes de la unión, (b) después de la unión.

En el caso de la unión pn , la continuidad del valor del campo en la unión estaba dada por el hecho de que a ambos lados, el material semiconductor era el mismo y tenía la misma constante dieléctrica. Por esta razón, lo que si debe ser continuo es el producto $\epsilon\mathcal{E}$, es decir:

$$\epsilon_A \mathcal{E}_{AM} = \epsilon_B \mathcal{E}_{BM} \tag{VIII-101}$$

Como la igualdad se plantea en la línea divisoria (interface) entre ambos semiconductores ($x = 0$), lo que se está escribiendo en la ecuación (VIII-101) es el valor máximo de cada uno de los campos eléctricos. Lo que aquí se plantea es la continuidad del producto Constante Dieléctrica x Campo, Como ocurría en el caso de la unión pn , el campo máximo se encuentra en la unión de ambos semiconductores, aca ocurre lo mismo, el campo es máximo sobre la unión. Si se supone que el material B está más dopado que el A, el gráfico de los campos será como en la Figura VIII-42.

Aquí se continúa utilizando al aproximación de unión abrupta, se supone que la carga acumulada a ambos lados de la unión finaliza en x_p y en $-x_n$, por lo que como en la unión pn , las cargas eléctricas a ambos lados deben ser iguales, es decir:

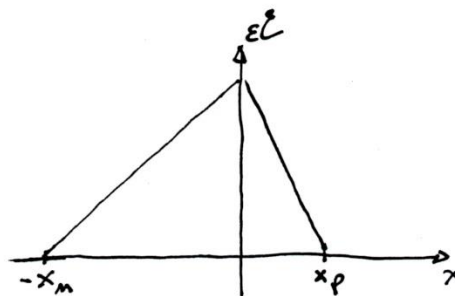


Figura VIII-42. Campo eléctrico en la heterounión.

$$N_d x_n = N_a x_p \tag{VIII-102}$$

Se había visto en la unión pn que el campo era una función lineal de x , ecuaciones (VIII-18 y 19). Como nos encontramos en $x = 0$, se tiene:

$$\mathcal{E}_{AM} = \left(\frac{qN_d}{\epsilon_A} \right) x_n \quad \text{y} \quad \mathcal{E}_{BM} = \left(\frac{qN_a}{\epsilon_B} \right) x_p \tag{VIII-103}$$

Los potenciales que actúan sobre cada uno de los semiconductores se puede encontrar integrando la expresión del campo eléctrico en la variable x , entre $-x_n$ y 0 para el semiconductor A y entre 0 y x_p para el B. En $x = 0$ se tiene:

$$V_{bin} = \frac{qN_d x_n^2}{2\epsilon_A} \tag{VIII-104}$$

$$V_{bip} = \frac{qN_a x_p^2}{2\epsilon_B} \tag{VIII-105}$$

El potencial de contacto total será:

$$V_{bi} = V_{bin} + V_{bip} = \frac{qN_d x_n^2}{2\epsilon_A} + \frac{qN_a x_p^2}{2\epsilon_B} \quad (\text{VIII-106})$$

Si de la ecuación (VIII-102) despejamos x_p , lo reemplazamos la ec. (VIII-106) y operamos, se tiene como resultado,

$$x_n = \left[\frac{(2\epsilon_A \epsilon_B N_a V_{bi})}{(qN_d (\epsilon_A N_d + \epsilon_B N_a))} \right]^{1/2} \quad (\text{VIII-107})$$

Si se realiza el mismo procedimiento, pero despejando x_n de la ec. (VIII-102), se tiene:

$$x_p = \left[\frac{(2\epsilon_A \epsilon_B N_d V_{bi})}{(qN_a (\epsilon_A N_d + \epsilon_B N_a))} \right]^{1/2} \quad (\text{VIII-108})$$

El ancho total de la zona de vaciamiento será

$$w = x_n + x_p = \left[\frac{(2\epsilon_A \epsilon_B (N_d + N_a)^2 V_{bi})}{(qN_a N_d (\epsilon_A N_d + \epsilon_B N_a))} \right]^{1/2} \quad (\text{VIII-109})$$

Si existiese algún potencial externo, como en el caso de la unión pn , el potencial de contacto V_{bi} deberá ser reemplazado por el nuevo potencial existente, potencial de contacto más el aplicado, $V_{bi} - V_A$ (V_A puede ser positivo o negativo).

Los cambios de potencial, con la aplicación de potenciales en directa o inversa, implican cambios en la extensión de la zona de vaciamiento, esto hace que se genere una capacitancia en la unión, que como lo hicimos anteriormente. La capacitancia en la unión es mayor cuando el sistema se encuentra bajo la acción de un potencial en inversa (V_I).

La capacitancia se define, como se ha visto al principio de este capítulo:

$$C = \frac{dQ}{dV_I}, \text{ donde } dQ = qN_d dx_n = qN_a dx_p$$

Por lo que,

$$C = \frac{dQ}{dV_I} = qN_d \frac{dx_n}{dV_I}$$

Haciendo las operaciones indicadas, se tiene:

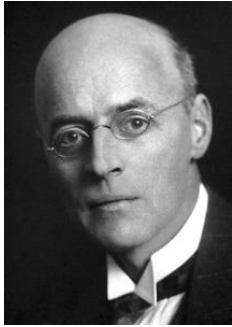
$$C = \left[\frac{qN_a N_d \epsilon_A \epsilon_B}{(2(\epsilon_A N_d + \epsilon_B N_a)(V_{bi} + V_I))} \right]^{1/2} \quad (\text{VIII-110})$$

La unidad del valor de C viene dada en F/cm^2

En la Figura VIII-41b en la parte superior de la interface, la banda de conducción toma una forma de V, se pueden obtener los niveles de energía solucionando la ecuación de Schrödinger, tomando como aproximación para el potencial, un pozo de potencial triangular, no se entrara en la discusión de este tema, pero se debe aclarar que en ese lugar se forma un gas bidimensional de electrones que no interactua con los electrones de conducción. Esa muesca en la banda de conducción no captura electrones ni perturba la corriente de electrones. El salto de potencial que existe en la banda de valencia evita la inyección de huecos del material dopado p al dopado n . La idea de las heterouniones es incrementar la barrera que deben sortear los huecos en relación a la que deben pasar los electrones. En las uniones pn las barreras eran iguales tanto para los electrones como para los huecos, eso no pasa en las heterouniones, esas barreras son desiguales.

Apéndice VIII-1

Emisión termoiónica. Ecuación de Richardson-Dushman



Owen Willans Richardson
26-4-1879, 15-2-1956
Premio Nobel de Física 1928

Por su trabajo sobre el fenómeno termoiónico y, la ley que lleva su nombre.



Saul Dushman (3ero de izq. A der.)
12-7-1883, 7-7-1954

Se debe entender que existe algún mecanismo que impide que los electrones de cualquier material, en particular metales y semiconductores, puedan escaparse de este. El efecto fotoeléctrico llevado adelante por A. Einstein demostró que para sacar un electrón del material hay que darle una cierta cantidad de energía para que pueda vencer las fuerzas que lo impiden. Estas fuerzas están representadas por los potenciales de superficie. Esto quiere decir, que hay un límite energético, el cual una vez alcanzado por los electrones, estos pueden encontrarse fuera del material, este límite se llama *energía de electrón libre*, E_a .

Por lo tanto, si un electrón se encuentra en el fondo de la banda de conducción, E_c , debe ganar una cierta cantidad de energía para que llegue al nivel E_a . La diferencia entre ambos se llama *trabajo real de salida*, como se puede ver en la Figura A1.

$$V_0 = E_a - E_c \quad (1-A)$$

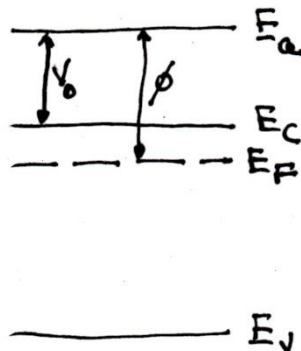


Figura A1. Niveles de energía en el semiconductor con respecto a E_a .

Para que un electrón pueda salir del material necesita superar la barrera de potencial de la superficie, por lo que solo unos pocos tienen la energía suficiente para hacerlo. Nuevamente

aquí, la temperatura juega un papel importante, cuanto más alta sea, mas electrones podrán escapar del semiconductor. Este efecto es lo que se llama *emisión termiónica*.

Los electrones que podrán escapar del material son aquellos que posean una energía cinética mínima, superior al valor de la altura de la barrera de potencial de la superficie,

$$\frac{1}{2}m^*v_{x,min}^2 \geq V_0 \quad (2-A)$$

Se debe determinar el número de estados cuánticos por unidad de volumen en el cristal para aquellos electrones que se mueven en un rango de energía v y $v + dv$. Esto es:

$$dZ = 2\left(\frac{m^*}{h}\right)^3 dv_x dv_y dv_z \quad (3-A)$$

Los electrones que pueden abandonar al semiconductor son los que tiene una energía,

$$E - E_F \geq kT \quad (4-A)$$

Por lo que, la ecuación (3-A) queda:

$$dn = f_0 dZ = 2\left(\frac{m^*}{h}\right)^3 e^{-(E-E_F)/kT} dv_x dv_y dv_z \quad (5-A)$$

Se debe evaluar la cantidad de electrones que pueden sobrepasar la barrera de potencial de altura V_0 . Para esto, se puede imaginar que los electrones se mueven en todas las direcciones en un material que tiene grandes dimensiones en las direcciones y y z , pero no en x , en esta dirección el material termina en $x = 0$. Por lo tanto, solamente interesan los electrones que se mueven en la dirección x hacia $x = 0$, con velocidad v_x . La densidad de corriente de los electrones que se mueven hacia la superficie es:

$$J = q \int_{v_{x,min}}^{\infty} \int_{-\infty}^{\infty} v_x dn = 2q \left(\frac{m^*}{h}\right)^3 e^{E_F/kT} \int_{v_{x,min}}^{\infty} \int_{-\infty}^{\infty} e^{-E/kT} v_x dv_x dv_y dv_z \quad (6-A)$$

En esta expresión E es la energía total del electrón. Como se está suponiendo que los electrones que se encuentran en la banda de conducción pueden lograr sobrepasar la barrera de potencial cuando alcanzan la energía cinética suficiente, se tiene que

$$E = E_c + \frac{1}{2}m^* \sum_i v_i^2 \quad (i \text{ corre sobre los ejes } x, y \text{ y } z) \quad (7-A)$$

Por lo tanto,

$$J = 2q \left(\frac{m^*}{h}\right)^3 e^{-(E_c-E_F)/kT} \int_{v_{x,min}}^{\infty} e^{-m^*v_x^2/2kT} v_x dv_x \int_{-\infty}^{\infty} e^{(m^*/2kT)(v_y^2+v_z^2)} dv_y dv_z \quad (8-A)$$

Para obtener la resolución de esta expresión, se debe recordar que algunas integrales ya están tabuladas, esto es:

$$\int_{-\infty}^{\infty} e^{-\alpha x^2} dx = \sqrt{\frac{\pi}{\alpha}}$$

Por otro lado, en la ecuación (8-A) la integral en dv_x es del tipo

$$\int e^{\alpha f} f' \quad \text{donde} \quad \alpha = -m^*/2kT \quad \text{y} \quad f = v_x^2$$

Por lo que:

$$\int_{v_{x,min}}^{\infty} e^{-m^*v_x^2/2kT} v_x dv_x = \frac{kT}{m^*} e^{-m^*v_x^2/2kT} = \frac{kT}{m^*} e^{-V_0/kT} \quad (9-A)$$

Si en la ecuación (8-A) se reemplazan los valores de las integrales, se tiene:

$$J = \left(4\pi q m^* k^2 / h^3 \right) T^2 e^{-(E_a - E_F)/kT} = AT^2 e^{-\phi/kT} \quad (10-A)$$

Donde ϕ es el *trabajo termodinámico de salida*, al valor $A = 1,2 \times 10^6 \text{ Am}^{-2}\text{K}^{-2}$, es la constante de Richardson. Si esta última ecuación es dividida por el área transversal lo que tenemos es la expresión de la densidad de corriente termoiónica, conocida como ecuación de Richardson-Dushman.

En el capítulo IV se vio como era la distribución de los electrones en función de la energía. La curva (Figura IV-4) mostraba un corte brusco cuando $T = 0\text{K}$ y decaía superando el nivel de Fermi, reflejando los estados ocupados más allá de E_F . En esa misma curva se pueden indicar cuáles serán los electrones que podrán superar la barrera de potencial V_0 , tal como se muestra en la Figura A2.

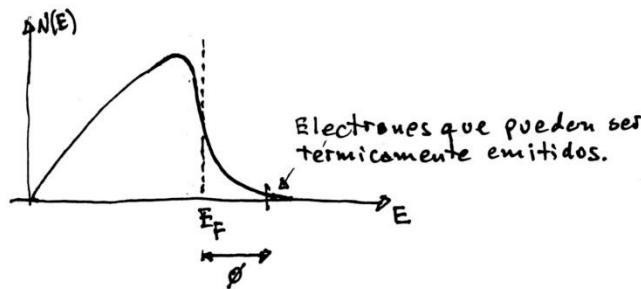


Figura A2. Distribución electrónica en función de la energía.

Si en lugar de trabajar con un semiconductor se hubiera trabajado con un metal, la diferencia establecida en la ecuación (1-A) se hubiera desarrollado entre el nivel de Fermi y la energía de salida, es decir,

$$\phi = V_a - V_F$$

El resto de la deducción es idéntica. Lo que esta ecuación establece es que la densidad de corriente termoiónica es muy dependiente de la temperatura de la superficie emisora y de la función trabajo.

Referencias

- Albella, J. M., Martínez-Duart, J. M. *Fundamentos de electrónica física y microelectrónica*. 1996 Addison-Wesley Universidad Autónoma de Madrid.
- Ashcroft, N. W., Mermin, N. D. *Solid State Physics*. 1981 Holt-Saunders International Editions.
- Dalven, R. *Introduction to Applied Solid State Physics. Second Edition*. 1980 Plenum Press. New York and London.
- Ihm, T. *Semiconductor Nanostructures Quantum States and Electronic Transport*. 2010 Oxford University Press, Great Clarendon Street, Oxford OX2 6DP
- Ghandhi, S. K. *Semiconductor Power Devices*, 1977 Wiley, New York
- Kittel, C. *Introduction to Solid State Physics Sixth Edition* 1986 John Wiley and Sons Inc. New York
- Mishra, U. K., Singh, J. *Semiconductor Device Physics and Design*. 2008 Springer The Netherlands 2008.
- Neamen, D. A. *Semiconductor Physics & Devices Basic Principles 2nd Edition* 1997 Irwin
- Pierret, F. R. *Advanced Semiconductor Fundamentals, Volume VI*. 1987 Modular Series on Solid State Devices, R.F Pierret and G. W. Neudeck, Editors. Addison-Wesley Pub. Comp.
- Sapoval, B., Hermann, G. *Physics of Semiconductors*. 1993 Springer-Verlag. New York.
- Shalímová, K. V. *Física de los Semiconductores*. 1975 Editorial MIR, Moscú.
- Singh, J. *Semiconductor Devices an Introduction* 1994 Mc Graw Hill International Editions.
- Sze, S. M. *Physics of Semiconductor Devices 2nd Edition*. 1981 Wiley-Interscience Publication
- Sze, S. M., Ng, K. K., *Physics of Semiconductor Devices Third Edition*. 2006 Wiley-Interscience Publication
- Van der Ziel, A. *Electrónica Física del Estado Sólido*. 1968 Editorial Prentice/Hall International
- Yepifanov, G. *Physical Principles of Microelectronics*. 1974 Mir Publishers. Moscow
- Yu, P. Y., Cardona, M. *Fundamentals of Semiconductors Physics and Material Properties*, 1996 Springer-Verlag Berlin Heidelberg.

CUARTA PARTE

Los dispositivos

CAPÍTULO IX

Dispositivos

Las ideas aisladas valen poco. El valor de la innovación reside en su implementación práctica

WERNER VON SIEMMENS
1816-1892

El desarrollo de los capítulos previos ha sido realizado para llegar a este punto. Todo lo anterior brinda las herramientas y el conocimiento necesario para entender como es el comportamiento de diferentes dispositivos con la mirada puesta en sus aplicaciones.

El aporte de la física a los dispositivos electrónicos ha sido muy importante, ayudando a la generación de dispositivos electrónicos que han cambiado las perspectivas de la humanidad. El ejemplo de este impresionante desarrollo es el transistor, por el cual recibieron el premio Nobel sus creadores, pero que además, su aparición en la escena mundial fue tan impactante e importante que hasta el día de hoy, multiplicado en infinidad de versiones, su uso continúa siendo esencial en la electrónica. En la actualidad la aparición de nuevos descubrimientos y tecnologías ha permitido observar que se abre un nuevo camino, esto es, dispositivos basados en la proposición de una nueva física basada en dimensiones mucho más pequeñas, el mundo nanoscópico. Esta puerta que se ha abierto conduce a novísimos dispositivos de dimensiones impensadas una década atrás.

En este capítulo, además se han incluido dos subcapítulos desarrollados por especialistas en ambos, los Dres. M. Cappelletti, A.Cédola y el Ing. A. Casas. Ellos han aceptado escribir sobre sus temas, celdas solares con perovskitas y celdas solares con puntos cuánticos. Dos temas absolutamente actuales y de gran perspectiva futura.

Comenzaremos haciendo una descripción del dispositivo que cambió la electrónica, el transistor.

IX.1. El transistor

El transistor, cuya denominación proviene de dos palabras del idioma inglés “transfer resistor” o mejor dicho la unión bipolar transistor, y bipolar porque maneja ambos portadores, fue desarrollada en los Laboratorios Bell por el John Bardeen y Walter Houser Brattain, liderados por William Shockley.



John Bardeen
23-5-1908, 30-1-1991



Walter H. Brattain
10/2/1902, 13/10/1987



William B. Shockley
13/2/1910, 12/8/1989

En el año 1956, recibieron el premio Nobel "por sus investigaciones sobre semiconductores y el descubrimiento del efecto transistor".

La historia del transistor es más extensa que esta breve reseña, comienza en el primer cuarto del siglo XX, no se desarrollará aquí, pero es ilustrativo interiorizarse de la misma. John Bardeen es uno de los pocos científicos en recibir dos premios Nobel.

El transistor original es el que se puede ver en la siguiente figura: IX-1

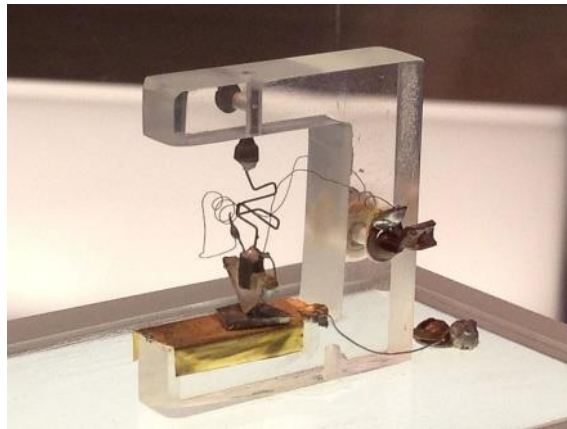


Figura IX-1. 1er. Transistor, año 1948, modelo original.

En este subcapítulo se verán los fundamentos del funcionamiento del transistor, para una descripción más detallada, referirse a los libros de la bibliografía incorporada al final del capítulo.

Se ha visto que la unión pn , es una rectificadora natural, elimina de la onda la parte del semiciclo negativo. Análogamente, el transistor es un amplificador natural, de acuerdo a como se lo conecte con otros elementos, puede amplificar la corriente, el voltaje (la tensión) o la potencia de una señal. Es un dispositivo de tres terminales y dos uniones, cuya función básica es controlar la corriente en una terminal mediante las tensiones aplicadas en las otras dos.

El transistor está formado por tres regiones dopadas en forma diferente, estas pueden ser nnp o pnp . Como se ve en la Figura IX-2, los dopados se mueven en forma descendente desde el emisor al colector, el emisor está más dopado la base y está más que el colector, es decir $N_{emisor} \gg N_{base} \gg N_{colector}$, lo cual nos muestra que no es una estructura simétrica. Claramente se puede ver que existen siempre dos uniones pn , por este motivo se podrá usar gran parte de lo que se desarrolló para estas uniones en el capítulo VIII. Como la región del medio en cada

una de los dos diferentes tipos debe ser muy estrecha, las dos uniones interactúan entre ellas. La estructura del transistor, en cada una de sus versiones, se puede ver en la Figura IX-2

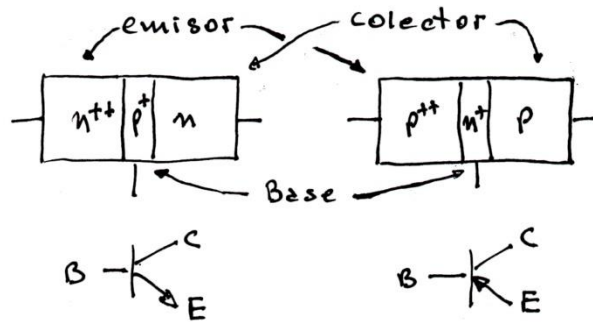


Figura IX-2. Esquema de los transistores npn y pnp, y los símbolos que los representan.

Debe entenderse que lo que se muestra en la Figura IX-2 es solamente un esquema del transistor, un dispositivo real dista mucho de la construcción mostrada en la figura.

Con cualquiera de las dos conformaciones que se elija, el desarrollo de las ecuaciones es similar, habrá que tener en cuenta los signos de corrientes y tensiones en cada una de ellas. Para nuestra explicación de su funcionamiento, tomaremos un transistor *nnp*.

Antes de proseguir se debe hacer un comentario, la convención indica que cuando se menciona a una tensión, por ejemplo $V_{BE} > 0$ se está indicando que se conectará el lado positivo a la base y el negativo al emisor, esto indica que $V_B - V_E > 0$, la base es positiva respecto del emisor, el segundo índice indica la tensión de referencia, etc. Tal como se muestra en la Figura IX-3.

El transistor puede funcionar en cuatro modalidades diferentes: activa, saturación, inversa y corte, dependiendo de las tensiones aplicadas entre emisor y colector. La modalidad más utilizada es la activa. Estas pueden verse en la Figura IX-4.

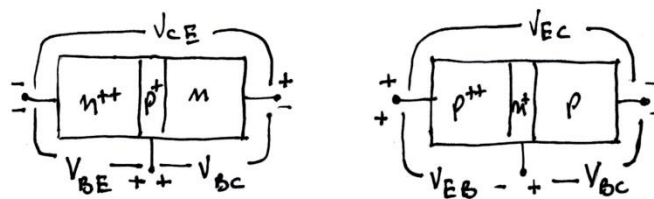


Figura IX-3. Definición de tensiones en el transistor.

El signo de las tensiones aplicadas se extrae automáticamente de la Figura IX-4. En la región activa, la más utilizada, es la que desarrolla mayor amplificación. En la región inversa, se intercambian emisor y colector, posee muy baja ganancia. En las otras dos regiones, saturación y corte, el transistor funciona como una resistencia entre emisor y colector.

Como se mencionó, se desarrollará la teoría para un transistor n^+p^+n actuando en la región activa. Como se ve en la Figura IX-3, la unión n^+p^+ (en adelante *np*), está polarizada en directa, mientras que la p^+n (en adelante *pn*) lo está en inversa. Cuál es la idea para el funcionamiento del transistor, esta es, intentar inyectar en el colector una elevada corriente proveniente del emisor, pasando por la base, aun cuando la unión *pn* está polarizada en inversa, de manera que domine ampliamente la corriente de portadores mayoritarios.

Como en el diodo de unión pn , se deben conocer las distribuciones de portadores, es decir resolver las ecuaciones de difusión como se hizo en el capítulo anterior. Para esto, se debe tener en cuenta para cada una de las zonas existentes el comportamiento de los portadores, en la zona del emisor, de la base y del colector.

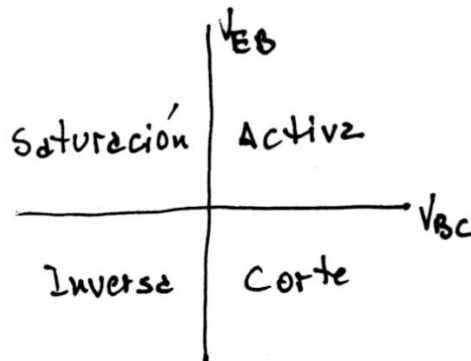


Figura IX-4. Regiones de funcionamiento del transistor.

Como se hizo en el tratamiento para el diodo, se deben redefinir las zonas estableciendo diferentes sistemas de coordenadas, para plantear desde el punto de vista matemático, ecuaciones menos complejas. No repetiremos esas deducciones, pero si incluiremos el comportamiento de las distribuciones de portadores, como se puede ver en la Figura IX-5. Antes de avanzar sobre esta distribución se deben recordar un par de cosas. Se ha dicho que la unión np se encontraba en directa, esto significa que se reduce la barrera de potencial entre ambos materiales al igual que el ancho de esta zona de vaciamiento. Por el contrario, la unión pn se encuentra bajo la acción de una polarización en inversa, lo que significa que la barrera de potencial se incrementa, y lo mismo pasa con la zona de vaciamiento entre la base y el colector. Las distribuciones de portadores tendrán un comportamiento como el de la Figura IX-5. En un caso como el que se ha mencionado, las bandas tendrán la forma como se ve en la Fig.IX-6, Este tipo de distribución de bandas pertenece a una conexión en el modo activo.

Por ser el transistor un dispositivo de 3 terminales, las formas de conexión se definen de acuerdo a cuál es la terminal en común que se tiene, puede ser base común, emisor común o colector común, en el caso del tipo de transistor que se ha elegido para su desarrollo, el npn , las conexiones son como las que se ven en la Figura IX-7.

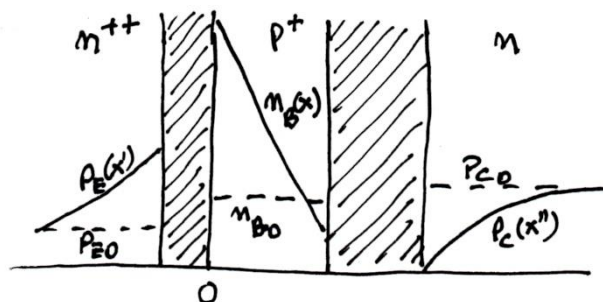


Figura IX-5. Distribución de portadores minoritarios para un transistor en el modo activo.

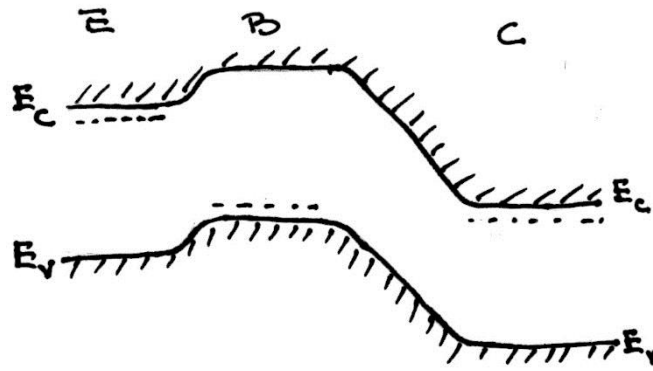


Figura IX-6. Estructura de banda de un transistor.

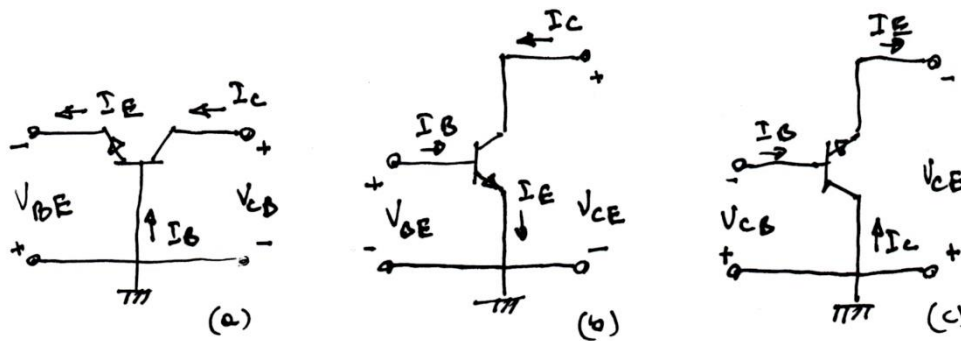


Figura IX-7. Esquema de conexiones. (a) base común, (b) emisor común y (c) colector común.

IX.1.a. Las corrientes

La importancia del transistor, como se mencionó más arriba, radica en el manejo de las corrientes. Como se ha impuesto una polarización en directa sobre la unión np , habrá una fuerte corriente de portadores mayoritarios (I_{En}) saliendo del emisor que atravesarán la base y llegarán al colector (I_{Cn}). Por otro lado, también en la unión emisor-base, habrá corriente de portadores minoritarios en la dirección base-emisor (I_{Ep}). Similarmente, en la unión pn (base-colector) polarizada en inversa se genera una corriente de portadores minoritarios en dirección a la base ($I_{Cp}=I_{CB}$) generados térmicamente. En todo esto, no se debe olvidar que los portadores (electrones) que pasan por la base, provenientes del emisor, se recombinarán con los huecos (portadores mayoritarios en la base), haciendo que disminuya la cantidad de electrones en la corriente emisor-colector ($I_{BR} = I_{En} - I_{Cn}$). Debido a este hecho es que se necesita que la base sea lo más estrecha posible. Se tiene también una corriente de recombinación (I_{ER}) en la unión base-emisor.

Los portadores mayoritarios que se pierden en la base debido a la recombinación, son reemplazados por la corriente de la base, I_B . Se puede esquematizar este movimiento de portadores como se muestra en la Figura IX-8.

Todas las corrientes mencionadas son producto de la difusión de portadores, excepto la de recombinación.

La forma de conexión que se ha establecido es la llamada de base común. La ganancia de corriente está establecida por la relación entre la corriente de emisor y la de colector,



Gustav Robert Kirchhoff
12/3/1824, 17/10/1887

Se puede ver de la Figura IX-8, que, utilizando la ecuación (IX-1), la corriente de colector está conformada por,

$$I_C = I_{Cn} + I_{CB} = \alpha_0 I_E + I_{CB} = \alpha_0 I_E + I_{CB0} \quad (IX-7)$$

Aquí se relaciona α_0 con la corriente de emisor. En esta ecuación, $I_{CB0} = I_{CB}$, cuando la corriente de emisor es nula ($I_E = 0$), esto quiere decir emisor abierto (para facilitar la comprensión del sentido de las corrientes se han cambiado estos dos símbolos anteriores por los de I_{CB} e I_{CB0} , solamente en la notación, no es su significado, corriente de portadores minoritarios de Colector a Base).

El mismo procedimiento se puede realizar si se toma la figura IX-7(b) emisor común, en este caso el factor de ganancia de corriente es β_0 y se lo puede relacionar con la corriente de base, es decir:

$$I_C = \beta_0 I_B + I_{CE0} \quad (IX-8)$$

Donde I_{CE0} es I_{C0} cuando $I_B = 0$, o base abierta. Si en la ecuación (IX-7) se reemplaza I_E por la ec. (IX-6), y se opera, se tiene:

$$\begin{aligned} I_C &= \alpha_0(I_C + I_B) + I_{CB0} \\ I_C(1 - \alpha_0) &= \alpha_0 I_B + I_{CB0} \\ I_C &= \left(\frac{\alpha_0}{1 - \alpha_0}\right) I_B + \left(\frac{1}{1 - \alpha_0}\right) I_{CB0} \end{aligned} \quad (IX-9)$$

Si se compara esta última ecuación con la ec. (IX-8), se puede ver que los coeficientes de ganancia de corriente β_0 y α_0 están relacionados,

$$\beta_0 = \frac{\alpha_0}{1 - \alpha_0} \quad (IX-10)$$

Lo mismo sucede con las corrientes de saturación,

$$I_{CE0} = I_{CB0} / (1 - \alpha_0) \quad (\text{IX-11})$$

Se había visto que (ec. (IX-1)) si la cantidad de portadores que llegan al colector desde el emisor es muy parecida a la corriente de emisor (base estrecha, poca recombinación) el diseño del transistor es bueno porque se inyecta una gran parte de la corriente de emisor en el colector, esto quiere decir que α_0 es muy próximo a 1. Si esto ocurre el denominador de la ec. (IX-10) va a ser muy pequeño y β_0 tendrá un valor muy grande.

Ejemplo IX-1. Se puede suponer que el valor de $\alpha_0 = 0,991$, esto implica que $\beta_0 = 110,1$. Por lo que la corriente de colector será al menos 110 veces mayor que la de la base.

Con este simple ejemplo se puede ver el efecto amplificador que tiene el transistor.

Se ha dado aquí una descripción muy rápida sobre la estructura del transistor y sobre su funcionamiento. Al igual que el diodo, el transistor también puede modelizarse mediante códigos de computadora y estudiar todas las variaciones que se puedan realizar hasta encontrar el dispositivo apropiado. Hay varios modelos que simulan al transistor, el Modelo de Ebers-Moll, el Modelo Híbrido-Pi entre otros. En este texto no abordaremos ese tema. Existen descripciones mucho más detalladas sobre el transistor que el alumno debe conocer, las cuales se encuentran en los libros que se dan como referencia.

IX.2. Diodo Túnel (Diodo Esaki)

El diodo Túnel fue desarrollado en el año 1958, en el año 1973, Leo Esaki, Ivar Giaever y Brian David Josephson recibieron el Premio Nobel por sus trabajos, los dos primeros “por sus descubrimientos experimentales en cuanto a los fenómenos de túnel en semiconductores y superconductores”, mientras que el tercero lo recibió: “por sus predicciones teóricas acerca de las propiedades de los superflujos a través de una barrera de túnel y, en particular, por aquellos fenómenos que son generalmente conocidos como efecto Josephson”.



Leo Esaki
12/3/1925,-



Ivar Giaever
5/4/1929,-



Brian D. Josephson
4/1/1940,-

Premios Nobel de Física, 1973, Temas: Diodo Túnel, Túnel en Semiconductores y superconductores y Efecto Josephson, respectivamente

El funcionamiento del diodo se basa en el “tuneleo” mecánico cuántico de los electrones. La característica de esta unión pn , es que ambas partes están altamente dopadas, haciendo que el nivel de Fermi se encuentre dentro de las bandas, en la de conducción del lado n y en la de valencia del lado p , y no en la banda prohibida como se ha visto hasta ahora. Esquemáticamente se ve en la Figura IX-9a.

Cuando el sistema ha llegado a su equilibrio termodinámico, las dos corrientes que existen, la túnel directa (I_{TD}) y la túnel inversa (I_{TI}) son iguales y opuestas, la primera va desde la banda de valencia a la de conducción y la segunda al revés, la corriente total en el sistema es cero. Ambas pasan de una zona a la otra por efecto túnel. Por encima del nivel de Fermi a ambas zonas son estados vacíos los que eventualmente podrían ser ocupados.

Si a un dispositivo con las características de dopado de este diodo, lo sometemos a un potencial en inversa, la corriente túnel inversa se incrementa mucho con respecto a la directa. En este caso, estaríamos en presencia del efecto Zener, al cual lo trataremos en el próximo subcapítulo.

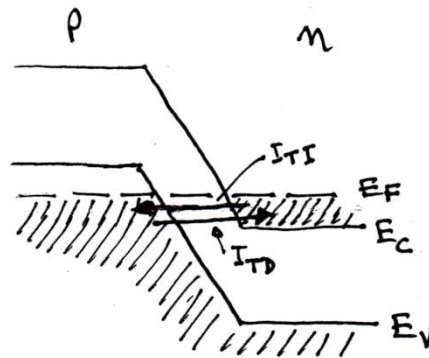


Figura IX-9a. Esquema de las bandas del diodo Túnel en equilibrio.

Debido a que este dispositivo está pensado para que trabaje con un potencial aplicado en directa, el estudio que se desarrollará será bajo este tipo de potencial.

La aplicación de un potencial en directa, como vimos en la unión pn la altura de la barrera de potencial disminuye, esto hace que el flujo de electrones que pasa por efecto túnel de la zona n a la p , aumente, en detrimento de la corriente opuesta. Si se sigue aumentando este potencial, la barrera disminuirá aún más y se llegará a un máximo cuando el fondo de la banda de conducción se encuentre a la misma altura del nivel de Fermi en la banda de valencia, como se ve en la Figura IX-9b.

Entre el estado de equilibrio y este de máxima, la parte ocupada de la banda de conducción va enfrentando diferentes estados de la banda de valencia. Al principio, todos

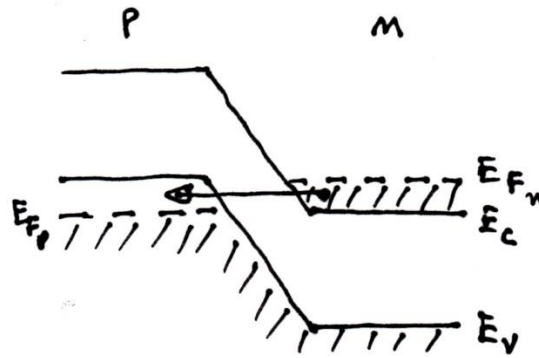


Figura IX-9b. Esquema de las bandas del diodo Túnel en su estado de máxima.

ocupados y a medida que aumenta el potencial va enfrentando estados vacíos superiores al nivel de Fermi. En este último en donde todos los estados ocupados enfrentan estados vacíos, esto implica un aumento de la corriente túnel directa cuyo máximo se logra con un esquema de bandas como el de la Figura IX-9b y donde la curva característica del diodo para este estadio es el que se muestra en la Figura IX-9c.

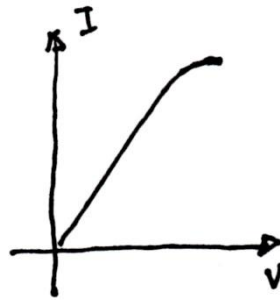


Figura IX-9c. Curva característica cuando se llega a la máxima corriente.

Si se sigue aumentando el potencial aplicado, la parte ocupada de la banda de conducción deja de enfrentar completamente a la parte desocupada de la banda de valencia y empieza a avanzar sobre la banda prohibida del diodo, de esta manera la corriente túnel directa se reduce (Figura IX-9d), puede haber transferencia de electrones hacia la banda de valencia pero no hacia la banda prohibida.

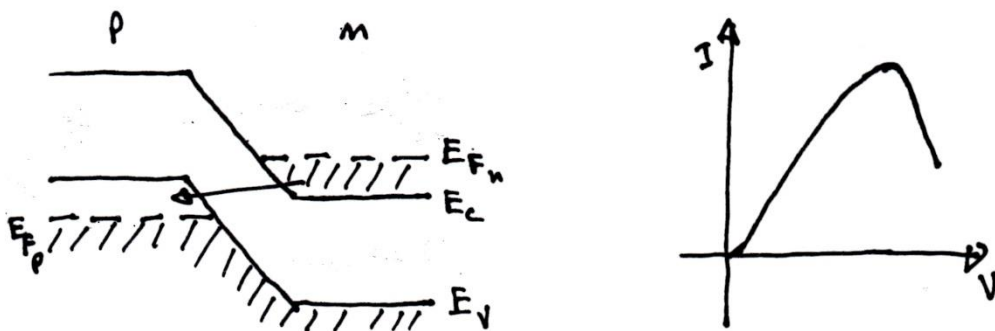


Figura IX-9d. Esquema de bandas y curva característica con potencial superando el máximo.

Se puede notar que habiendo superado el máximo, aun cuando se aumente el potencial, los valores de la corriente siguen bajando, presentando una curva que tiene pendiente negativa, en contra de lo que es la relación clásica entre corriente y tensión. Esta región hacia abajo de la curva, se la llama resistencia diferencial negativa. Es una característica de este tipo de diodo y de los generadores de microondas (efecto Gunn) (John Battiscombe Gunn, 13 May 1928 – 2 December 2008)

Con un nuevo aumento del potencial aplicado, la parte ocupada por los portadores en la banda de conducción, enfrenta totalmente a la banda prohibida, la corriente ha llegado a su mínimo, y el diodo Túnel se comportará a partir de este momento como un diodo común donde los portadores que dominan la escena son los minoritarios. Este comportamiento se muestra en la Figura IX-9e.

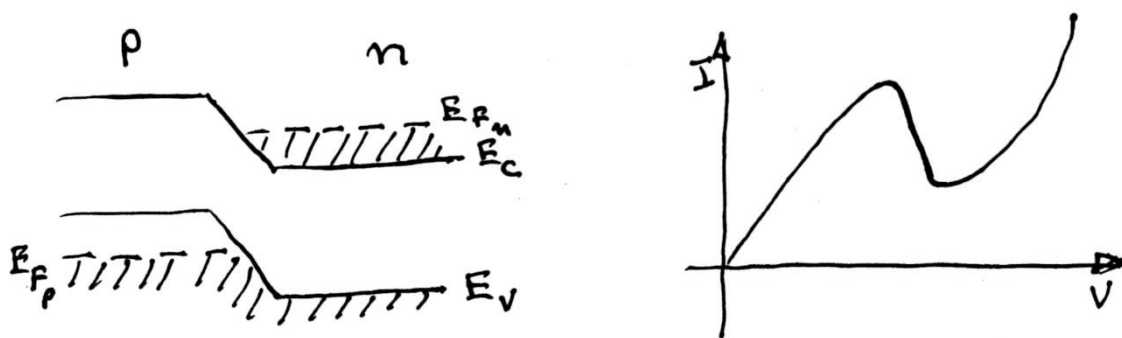


Figura IX-9e. Esquema de bandas y curva característica con potencial superando el mínimo.

La gran virtud del diodo Túnel es que la corriente se debe a los portadores mayoritarios, los cuales no son alcanzados por las limitaciones que tienen los minoritarios, inyección, acumulación, etc. Por este motivo, el diodo Túnel, entre otras cosas, tiene una mejor respuesta en frecuencia que el diodo común.

En un diodo Túnel, el ancho típico de la zona de vaciamiento es de aproximadamente uno 100\AA o menos. Para que haya efecto túnel, los estados vacíos deben estar a la misma energía que los ocupados. La barrera de potencial debe ser baja y angosta para que haya pasaje de electrones por efecto túnel. Durante el efecto túnel, la cantidad de movimiento se debe conservar.

En el Figura IX-9e, se muestra al curva característica del diodo Túnel, se puede ver que el mínimo de la curva no llega al valor $I = 0$, lo cual se podría suponer pues es un punto de cambio de régimen, del comportamiento basado en portadores mayoritarios al de minoritarios. La curva característica del diodo Túnel muestra un comportamiento diferente, antes y después del potencial de valle, el cual se registra en el mínimo de la curva, como se ve en la Figura IX-10.

Aproximadamente, hasta llegar al potencial de valle, donde se registra un mínimo de corriente túnel, incluyendo el potencial de pico, donde se registra la máxima corriente, en toda esa zona, el efecto predominante es el pasaje de los electrones del material dopado n ocupando lugares vacíos en el material p , por medio del efecto túnel. Superado ese potencial, el atravesar la barrera mediante "tuneleo" no es casi posible y comienzan a dominar los efectos térmicos, tanto huecos como electrones, portadores de uno y otro signo, pasaran por difusión sobre la barrera de potencial, de acuerdo a la energía térmica que posean.

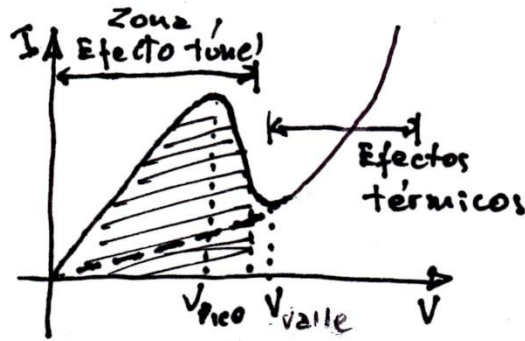
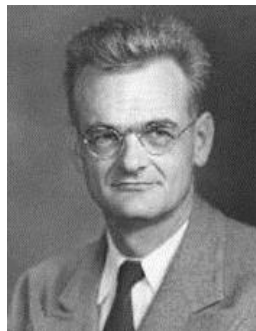


Figura IX-10. Curva característica del diodo Esaki.

IX.3. Diodo Zener

Un diodo con un alto dopado semejante al del diodo Túnel, con una zona de vaciamiento muy estrecha y sometido a una tensión en inversa relativamente baja, del orden de unos 5V, presenta la característica de la rotura Zener.

Un potencial de ese orden genera un campo crítico ϵ_{cr} el cual es suficientemente intenso para romper las ligaduras covalentes de los átomos en la banda de valencia y sacar esos electrones y forzarlos a atravesar la barrera por efecto túnel, campos críticos típicos son del orden de 10^5V/cm o 10^6V/cm .



Clarence Melvin Zener
1/12/1905, 15/7/1993

Si se escribe el campo en función del potencial y la distancia, se puede expresar como:

$$\epsilon_{cr} \cong V_B/d \tag{IX-12}$$

Donde

$$d = (2\epsilon\epsilon_0 V_B/qn_{n0})^{1/2} \tag{IX-13}$$

Es el espesor de la unión *pn*. Reemplazando la ecuación (IX-13) en la (IX-12), se puede obtener el potencia de ruptura V_B :

$$V_B = 2\epsilon\epsilon_0\epsilon_{cr}^2/qn_{n0} = 2\epsilon\epsilon_0\epsilon_{cr}v_d\rho_n \quad (IX-14)$$

Donde v_d y ρ_n son la velocidad de arrastre y la resistividad del semiconductor. Se puede ver de la ec. (IX-14) que la ruptura Zener es proporcional a la resistividad de la región óhmica del semiconductor. La curva que la representa es la de la Figura IX-11. En la Figura se puede ver que la parte interesante del comportamiento está cuando se tiene un potencial negativo. La corriente se comporta como el diodo ideal, hasta llegar al potencial de ruptura, en donde se despega completamente de la curva ideal y con pequeñas variaciones de la tensión aplicada, se producen grandes variaciones de la corriente.

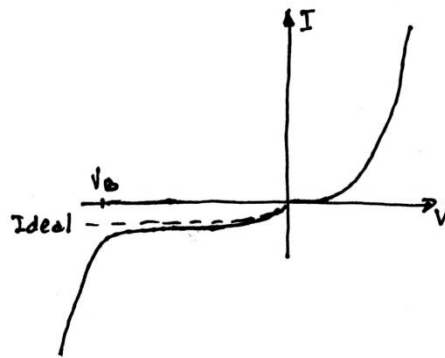


Figura IX-11. Curva característica del diodo Zener,

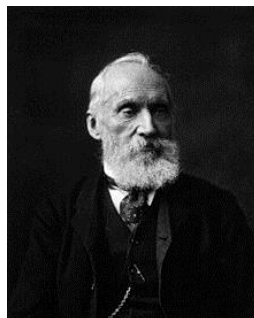
A medida que se incrementa la zona de vaciamiento, la probabilidad de que exista “tuneleo” decrece, desapareciendo el efecto Zener.

IX.4. Celda Peltier



Jean Charles Athanase Peltier

22/2/1785, 27/10/1845



William Thomson
Lord Kelvin

26/6/1824, 7/12/1907



Thomas Johann Seebeck

9/4/1770, 10/12/1831

Los fenómenos termoelectricos son aquellos donde se establece una interacción entre energía térmica y eléctrica. Los efectos Peltier, Thomson y Seebeck están basados en este fenómeno.

La celda Peltier, basada en el efecto Peltier establece que cuando una corriente pasa a través de la unión, y dependiendo del sentido de la corriente, se puede incrementar o decrementar la temperatura de ella.

Para realizar el desarrollo de la celda Peltier, se comenzara haciendo algunos comentarios sobre el efecto Seebeck. Si se tiene dos conductores o semiconductores unidos, se formaran dos juntas, si a estas se las pone a diferentes temperaturas, circulara una corriente en el circuito como se ve en la Figura IX-12.

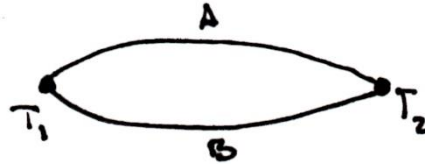


Figura IX-12. Esquema del efecto Seebeck.

Si este circuito es abierto, como se ven en la Figura IX-13, entre las dos terminales se genera una diferencia de potencial llamado Potencial Seebeck, V_s . Se puede demostrar que el potencial Seebeck es una medida de la diferencia de temperatura entre ambas uniones.

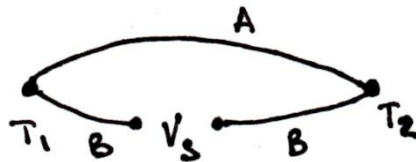


Figura IX-13. Esquema del circuito para generar el potencial Seebeck

El efecto descubierto por Peltier fue 13 años después que Seebeck descubriera el efecto que lleva su nombre. Originalmente Peltier trabajo con metales, observando que cuando una corriente eléctrica circula por un circuito formado por dos metales diferentes, las temperaturas en las inmediaciones de las juntas cambian, el esquema se muestra en la Figura IX-14.

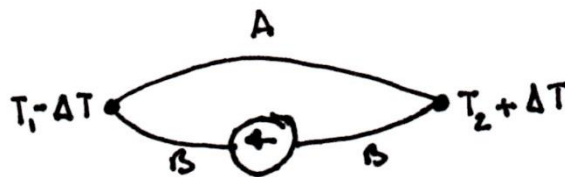


Figura IX-14. Esquema del circuito Peltier

En el presente caso, como no se va a trabajar con metales sino con semiconductores, se toman dos materiales semiconductores diferentes con dopados n y p , respectivamente. Los materiales deben tener una conexión única en uno de sus extremos, se podría llamar la parte de arriba, y contactos separados en el extremo opuesto, la parte de abajo, como se ven en la Figura IX-15.

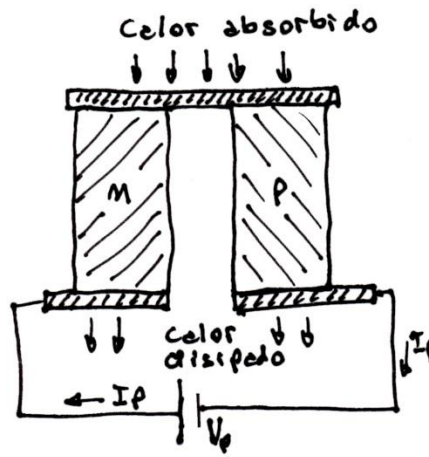


Figura IX-15. Esquema de conexión de la celda Peltier.

De acuerdo a la polaridad que tiene la fuente en la conexión de la Figura IX-15, los electrones se mueven en sentido opuesto a la corriente, lo hacen yendo hacia los contactos inferiores, los huecos que se mueven en el mismo sentido de la corriente también se mueven hacia los contactos inferiores. Por esta razón la parte superior queda des poblada de electrones, generando una superficie fría, contrario a lo que sucede en los contactos inferiores donde la gran cantidad de electrones hace que la temperatura aumente. Si la polaridad se invierte, también se invierten los focos frío y caliente.

Por ser la celda Peltier un elemento termoelectrico, está gobernado por las teorías desarrolladas por Joule y Thomson sobre este tipo de fenómenos. Se deben aplicar ellas a los semiconductores, la potencia termoelectrica obtenida de un semiconductor tipo n , es:

$$-Q_n T = (E_C - E_F) + 2kT \tag{IX-15}$$

La de un semiconductor tipo p , es:

$$Q_p T = (E_F - E_V) + 2kT \tag{IX-16}$$

Donde las E_i , son las energías de los bordes de las bandas de conducción y valencia y la energía de Fermi, las Q son las energías caloríficas producida por la temperatura T y k la constante de Boltzmann. Como se mencionó más arriba, si se mantienen los contactos a diferentes temperaturas,

$$\Delta T = T_C - T_F \tag{IX-17}$$

donde T_C y T_F son las temperaturas de la fuente caliente y de la fría respectivamente. Como consecuencia de esto, se genera una diferencia de tensión V_S . Esta diferencia de potencial surge del rápido movimiento de los portadores de la fuente caliente a la fría. El coeficiente Seebeck se define como la relación entre el potencial Seebeck (fuerza termo electromotriz) y la temperatura,

$$S = V_S / \Delta T \tag{IX-18}$$

La potencia calorífica que se intercambia entre la fuente fría y la caliente, por unidad de área, es

$$Q = I_p \Pi = I_p S \Delta T \quad (\text{IX-19})$$

Donde Π es el coeficiente Peltier, $\Pi = S \Delta T$. El cual define el calor intercambiado entre la fuente caliente y la fría en la unión, por unidad de tiempo y de corriente circulando a través de la misma.

Cuando a la celda se le aplica una tensión, que genera una corriente, la potencia absorbida responde a dos términos, uno es el correspondiente al efecto Joule y el otro al Peltier, por lo tanto la potencia absorbida es:

$$P = I_p^2 R_p + I_p S \Delta T \quad (\text{IX-20})$$

Finalmente, el efecto Peltier esta generado por el movimiento de los portadores dentro del material, los cuales generan diferencias de temperatura, es decir, por la conducción térmica normal y por el efecto Joule. Si se establece el balance de energías, se tiene:

$$\begin{aligned} Q_F &= \text{corriente efecto Peltier} + \text{corriente efecto Joule} + \text{corriente térmica Thomson} \\ &= S \Delta T I_p - \frac{1}{2} I_p^2 R_p - \kappa \Delta T \end{aligned} \quad (\text{IX-21})$$

En esta expresión, Q_F , S , I_p , R_p y κ , son: calor absorbido en la fuente fría, coeficiente Seebeck, corriente que circula por el circuito, resistencia eléctrica de los semiconductores y coeficiente de conductividad térmica, respectivamente. La ecuación (IX-21) se puede reescribir de la siguiente manera.

$$Q_F = \Pi I_p - \frac{1}{2} I_p^2 R_p - \kappa \Delta T \quad (\text{IX-22})$$

Las corrientes en la celda se representan por el esquema de la Figura IX-16.

Para saber cuán alta es la calidad de una celda en cuanto a su respuesta, no sus materiales, se utiliza normalmente un parámetro llamado Figura de Mérito (FOM en Inglés, Figure of Merit). En el caso de la celda Peltier, la Figura de Mérito, Z , es:

$$Z = S^2 \sigma / \kappa \quad (\text{IX-23})$$

Donde, σ es la conductividad. El valor ZT es una cantidad adimensional.

La capacidad de refrigerar (o eventualmente de calentar) de una celda Peltier con materiales semiconductores, depende de diferentes efectos, el coeficiente Seebeck, que relaciona el potencial con las variaciones de temperatura, la conductividad eléctrica, relacionada con el efecto Joule y la conductividad térmica. Este conjunto de parámetros que describen el

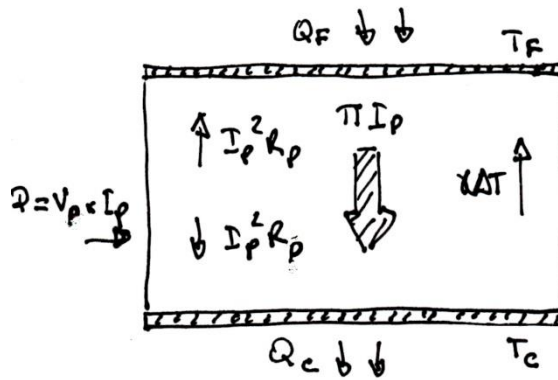


Figura IX-16. Esquema de corrientes en la celda Peltier.

comportamiento físico de la celda se reúnen en la Figura de Mérito, para un buen rendimiento, este valor debería ser maximizado.

IX.5. Celda solar. Generalidades

La tarea de una unión *pn* es modificar la señal que recibe transformándola en algo diferente, es decir, lo que egresa de la unión es distinto a lo que ingresó. Las celdas solares realizan justamente esa tarea, expuestas a la radiación solar, transforman esa energía incidente en energía eléctrica. La idea que subyace a este desarrollo es que con la energía generada se pueda realizar algún trabajo, por ejemplo: mover un motor. Por este motivo, en el desarrollo que se hará de la celda solar, se adicionará una carga, en general resistiva, la cual simula el dispositivo en el cual se aplicará la energía generada.

Se avanzará sobre la explicación del mecanismo físico que produce esta transformación.

En los dispositivos sensibles a la luz, sin lugar a dudas, la incidencia de fotones debe afectar alguna de las propiedades de los semiconductores, lo cual los hace útiles para su uso. Teóricamente, cuando un fotón llega al semiconductor y es absorbido, este produce la generación de un par electrón-hueco. Este proceso en forma masiva hace que haya exceso de portadores, los cuales se moverán bajo la influencia del campo existente en la zona de generación y se sumarán a alguna corriente o la crearán.

La radiación solar incidente está compuesta por fotones de frecuencia ν y energía $E = h\nu$, h es la constante de Planck, de acuerdo a lo que vimos en el primer capítulo. Se puede escribir:

$$\nu = E/h \Rightarrow 1/\nu = h/E \Rightarrow c/\nu = hc/E = \lambda \tag{IX-24}$$

Donde λ es la longitud de onda del fotón y c es la velocidad de la luz en el vacío. Con esta expresión se ha relacionado la longitud de onda con la energía del fotón. Como h y c son constantes, el valor aproximado de λ es:

$$\lambda = 1,24/E \mu m \quad (\text{IX-25})$$

Cuando un fotón incide sobre un semiconductor, lo ideal sería que un electrón del borde de la banda de valencia sea excitado con la energía suficiente para que alcance la banda de conducción, en cuyo caso la energía necesaria sería la del ancho de la banda prohibida del semiconductor, E_g . Esto determina, de acuerdo al semiconductor que se tenga, cuál será la longitud de onda necesaria excitar electrones y producir pares electrón-hueco,

$$\lambda = 1,24/E_g \mu m \quad (\text{IX-26})$$

Claro está, que pueden ocurrir otros casos, puede ser que la energía del fotón sea menor que la de la banda prohibida, $E < E_g$, en este caso, el fotón no es absorbido, atraviesa el material y emerge. Esta situación es la que define que un material sea transparente. Si el fotón tiene la misma energía que la banda prohibida, el caso que se ha comentado en líneas anteriores, $E = E_g$, un electrón de la banda de valencia es promovido a la banda de conducción, creando un hueco en la de valencia, par electrón-hueco. Por último, si se da el caso $E > E_g$, se crea un par electrón-hueco, pero el exceso de energía es disipado en forma de calor dentro del material, esquemáticamente puede verse en la Figura IX-17.

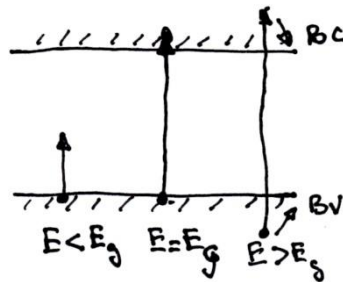


Figura IX-17. Esquema de salto de portadores de la BV a la BC.

Se puede considerar ahora la unión pn a la cual se la expone a radiación lumínica y se le agrega una carga. Cuando se vio en el capítulo VIII, en una unión pn aun cuando no exista potencial aplicado, existe un campo eléctrico debido a las cargas acumuladas en las inmediaciones de la unión metalúrgica, que se opone al pasaje de portadores mayoritarios. La dirección de este campo será siempre la misma. Los fotones que llegan a la zona de vaciamiento generan pares electrón-hueco, estos portadores son arrastrados por el campo a cada una de las zonas correspondientes, esto quiere decir que los electrones son llevados hacia la zona n y los huecos hacia la p . El movimiento de los portadores genera una corriente que se mueve hacia la zona p . Este flujo de portadores genera una fotocorriente I_{FC} . La fotocorriente genera una caída de potencial en la carga, tal que el polo negativo esté conectado a la región n y el positivo a la p . Esto hace que la celda solar vea a este potencial como un potencial en directa aplicado sobre sus terminales. Como se vio en la unión pn una tensión de este tipo genera una corriente directa I_D . En este punto se ve que en el circuito circula una corriente total que será la suma de ambas. El esquema del circuito se muestra en la Figura IX-18,

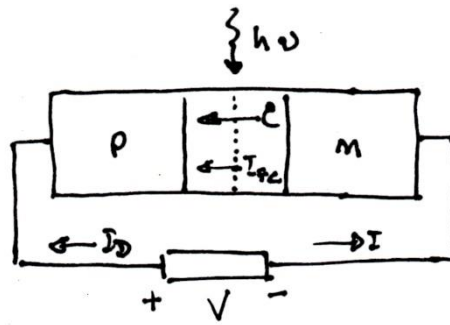


Figura IX-18. Esquema de circuito con celda solar.

La expresión de la corriente total en el circuito será:

$$I = I_{FC} - I_D = I_{FC} - I_s \left(e^{qV/kT} - 1 \right) = I_{FC} - I_s e^{qV/kT} \quad (\text{IX-27})$$

Se debe pensar que puede existir la circunstancia donde se supone que la carga tiene un valor muy grande (la resistencia tiende a ∞) es como si el circuito estuviera abierto, en este caso la corriente tendrá valor cero, $I = 0$ y el potencial será el de circuito abierto, V_{oc} . Por otro lado, se puede tener que la carga sea cero, esto implicaría que $V = 0$, en este caso la corriente es llamada corriente de cortocircuito, I_{sc} . En este caso:

$$I = I_{sc} = I_{FC} \quad (\text{IX-28})$$

Si nos referimos al primer caso, se tiene:

$$I = I_{FC} - I_s e^{qV_{oc}/kT} = 0 \quad (\text{IX-29})$$

Si de aquí despejamos el potencial de circuito abierto,

$$V_{oc} = \left(\frac{kT}{q} \right) \ln \left(\frac{I_{FC}}{I_s} \right) \quad (\text{IX-30})$$

Ejemplo IX-2. Determinar el valor de V_{oc} .

En el ejemplo VIII-6 se obtuvo que $J_s = 1,196 \times 10^{-11} \text{ A/cm}^2$. Se puede suponer que la fotocorriente, para la misma área que en el ejemplo mencionado, es: 20 mA/cm^2 , es decir:

$$\begin{aligned} V_{oc} &= \left(\frac{kT}{q} \right) \ln \left(\frac{I_{FC}}{I_s} \right) = \left(\frac{kT}{q} \right) \ln \left(\frac{J_{FC}}{J_s} \right) = 0,0259 \times \ln \left(\frac{20 \times 10^{-3}}{1,196 \times 10^{-11}} \right) \\ &= 0,55 \text{ V} \end{aligned}$$

La potencia que la carga genera es,

$$P = IV = \left[I_{FC} - I_s e^{qV/kT} \right] V \quad (\text{IX-31})$$

El potencial, V_M , que determina la máxima potencia que pueda erogarse el dispositivo está dada por $dP/dV = 0$,

$$dP/dV = I_{FC} - I_s e^{qV_M/kT} - I_s V_M (q/kT) e^{qV_M/kT} = 0 \quad (IX-32)$$

A partir de esta ecuación también se puede determinar la corriente que determina la máxima potencia. Operando algebraicamente a la ecuación (IX-32) se puede escribir:

$$e^{qV_M/kT} [1 + V_M (q/kT)] = I_{FC}/I_s \quad (IX-33)$$

Como el valor V_M se encuentra en el exponente del exponencial y a su vez, multiplica al exponencial, no se puede despejar directamente y la solución debe ser encontrada en forma numérica. En la Figura IX-19 se muestra la curva característica de una celda solar y además el punto Q determinado por la máxima corriente y el máximo potencial obtenido mediante la Ec. (IX-33).

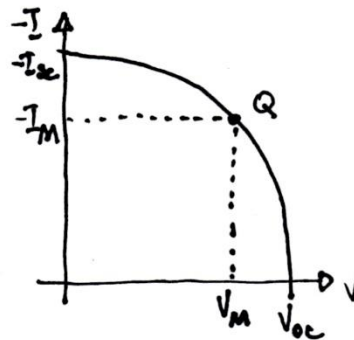


Figura IX-19. Curva característica de la celda solar y el punto Q definido por los valores máximos I_M y V_M .

Un parámetro muy importante para determinar la calidad de conversión que tiene las celdas solares es su eficiencia η , este define la cantidad de pares electrón-hueco que se generan frente al número de fotones que llegan al dispositivo, dicho de otra manera es la relación entre potencia eléctrica generada frente a potencia lumínica recibida. Se la puede expresar:

$$\eta = P_M/P_{inc} (\%) = (I_M V_M)/P_{inc} (\%) \quad (IX-34)$$

En la actualidad, la eficiencia de las celdas solares está lejos de ser la óptima, ese es un punto de investigación, que se debe realizar para obtener un mejor aprovechamiento de la energía que entrega el Sol.

A continuación se desarrollarán dos ejemplos de celdas solares donde se apunta a mejorar su eficiencia. La idea de estos dos subcapítulos es mostrar al alumno los estudios

que se hacen sobre temas básicos para su mejor entendimiento. Ambos subcapítulos son contribuciones por los propios investigadores.

IX.5.a. Celda solar con perovskitas

Contribución del Dr. Marcelo Cappelletti y el Ing. Guillermo Casas

En los últimos años, ha cobrado especial relevancia un nuevo tipo de dispositivo fotovoltaico, denominado celdas solares basadas en materiales con estructura de perovskita (CSP), la cual es una tecnología que promete revolucionar la industria fotovoltaica. Se trata de dispositivos de heterojunturas híbridos orgánico/inorgánico, conformada con materiales con diferente ancho de la banda prohibida (*bandgap*), de bajo costo, cuya eficiencia ha sido incrementada sustancialmente desde su aparición, desde 3.8% en 2009 [1] hasta 22.1% en 2016 [2].

Si bien la perovskita es un trióxido de titanio y de calcio (CaTiO_3), mineral descubierto en Rusia en el año 1839, también se denomina perovskita a una familia de materiales con la misma estructura cristalina que el CaTiO_3 . Una característica que los hace muy atractivos a algunos de estos materiales y que los ha puesto a la vanguardia de la innovación en cuanto a las aplicaciones fotovoltaicas, es su capacidad para absorber la radiación lumínica incidente.

Hoy en día, los prototipos de laboratorio CSP han demostrado competir en eficiencia frente a tecnologías ya establecidas en el mercado mundial, tales como las celdas solares de película delgada basadas en telurio de cadmio (CdTe) o en cobre-indio-galio-selenio (CIGS), o como las celdas solares de silicio poli-cristalino. Además, a diferencia de los materiales semiconductores rígidos como el silicio que requieren de una gran cantidad de procesos para convertirlos en celdas solares, las CSP son más económicas, son flexibles y tienen un proceso de fabricación más simple. Es por ello que las CSP poseen menor tiempo de retorno energético que las celdas solares de silicio, es decir el tiempo necesario para que un material ahorre la misma cantidad de energía que se gasta en su producción.

No obstante, más allá de sus cualidades, las CSP presentan aún una serie de limitaciones que desafían a la comunidad científica abocada en esta línea de estudio a continuar y profundizar las investigaciones, tanto teóricas como experimentales, sobre estos dispositivos para hacerlos aptos para una futura comercialización de ellos [3-4].

En este sentido, los estudios no solamente apuntan a maximizar la eficiencia de las CSP, sino también a utilizar materiales más económicos, a reducir la toxicidad en la fabricación de estos dispositivos y fundamentalmente a aumentar su estabilidad. Respecto a este último punto, la poca estabilidad para mantener en el tiempo la eficiencia (hasta 1000 horas de sol), es uno de los principales inconvenientes que presentan actualmente las CSP, dado que cuando entran en contacto con la humedad, se destruyen sus capacidades de captación de energía solar.

Una estructura utilizada en los prototipos experimentales actuales de CSP está conformada por tres regiones de la forma: HTM/Perovskita/ETM (Figura IX-20), donde la capa HTM (*Hole Transporting Material*) debe corresponder a un material tipo p que permita el transporte

de huecos y bloquee el flujo de electrones hacia el ánodo, y la capa ETM (*Electron Transporting Material*) debe ser un material tipo n que posibilite el transporte de electrones libres y bloquee el flujo de huecos hacia el cátodo. La luz ingresa por la capa ETM y es absorbida por la capa de perovskita, en la cual se generan los pares electrón-hueco.

Los materiales que han sido utilizados inicialmente como HTM y ETM son el spiro-OMeTAD y el TiO_2 , respectivamente [5]. El primero de ellos es un compuesto orgánico, relativamente caro, con el que se han obtenido en la práctica las mayores eficiencias hasta el momento. En los últimos años, diferentes investigaciones han sido llevadas a cabo mediante herramientas de simulación, para analizar los efectos de las variaciones en las características de las capas que conforman las CSP, entre los que se incluye un trabajo reciente realizado por los autores de este libro, en colaboración con el Dr. Bernabé Marí Soucase (Depto. de Física Aplicada-ETSID, Universidad Politécnica de Valencia, España), referente internacional en el estudio teórico y experimental de las CSP [6-12]. Entre los materiales candidatos a ser utilizados como capa HTM en reemplazo del spiro-OMeTAD se tienen diferentes materiales inorgánicos, tales como Cu_2O , CuSCN , NiO y CuI . Por su parte, los candidatos a reemplazar al TiO_2 para la capa ETM son ZnO , Al_2O_3 y Fe_2O_3 (Hematita). Y finalmente, existe una gran cantidad de materiales posibles a ser utilizados en la región central de las CSP, todos con estructura de perovskita

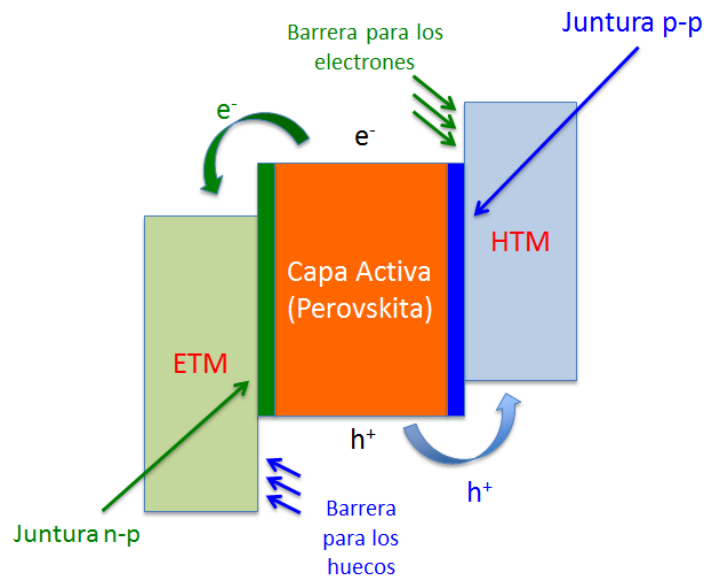


Figura IX-20. Estructura planar de una CSP de tipo p-p-n.

de la forma MA-MX_3 , donde MA es el metilamonio CH_3NH_3 , M es un metal, por ejemplo: Pb, Sn; y X representa I, Br, Cl o F. También pueden ser utilizados mezclas tales como $\text{MA-Pb}_x\text{Br}_{3-x}$, $\text{MA-Pb}_x\text{I}_{3-x}$, $\text{MA-PbI}_x\text{Br}_{3-x}$, entre otras.

Una manera de aumentar la eficiencia de las celdas solares, es mediante la utilización de dos o más celdas apiladas una encima de la otra (celdas solares tándem), para favorecer la habilidad de captar rangos de energía más amplios del espectro solar. Con esta idea es que surgieron hace tiempo las celdas solares multijunturas basadas en semiconductores compuestos III-V, las cuales presentan las mayores eficiencias hoy en día [2]. Debido a su alto costo de fabricación y buena tolerancia a la radiación de partículas de altas energías (protones y elec-

trones), su campo de aplicación principal está en los vehículos y satélites destinados a aplicaciones espaciales [13-15].

En las celdas multijuntura la subcelda ubicada en la parte frontal del dispositivo captura los fotones de más alta energía y deja pasar el resto, que serán absorbidos por las subceldas con menores bandgap. Con este mismo criterio es que últimamente han comenzado a estudiarse celdas solares tándem basadas en perovskita, como por ejemplo, las perovskita/silicio [16-18] o bien las perovskita/perovskita utilizando dos tipos de perovskita diferentes [19-20]. De esta manera, cada celda solar absorbe una porción diferente del espectro solar, mejorando el rendimiento en comparación con las celdas individuales separadas.

Un hecho que también las hace sumamente interesantes para su estudio, es que las celdas solares tándem perovskita/perovskita han sido estudiadas para la obtención de hidrógeno a partir de la disociación del agua, por medio de la foto-electrólisis conducida por energía fotovoltaica [21]. El uso del hidrógeno como combustible es una de las alternativas más importantes para hacer frente a la disminución de reservas de combustibles fósiles y a los altos índices de contaminación. Sin embargo, al día de hoy continúa siendo un gran desafío desarrollar sistemas solares de disociación de agua para producir hidrógeno en forma limpia, que sean simultáneamente de bajo costo y lo suficientemente eficientes como para generar combustible a un precio competitivo con los combustibles fósiles.

La disociación del agua por medio de la foto-electrólisis requiere generar una diferencia de potencial de aproximadamente 1.5 V [22]. Los electrolizadores comerciales operan típicamente a una tensión de 1.8 a 2.0 V [23]. Esta diferencia de potencial es la que debe proveer la celda solar. No es posible alcanzar estos valores de tensión con una única celda solar convencional de silicio mono- y poli-cristalino debido a sus bajas tensiones en circuito abierto. Para poder conducir la electrólisis con estos dispositivos, se deben conectar en serie tres o cuatro celdas para lograr una eficiencia razonable [24], tal como se implementó con celdas solares multijunturas de silicio [25] y con multijunturas basadas en semiconductores compuestos III-V [26]. Las CSP han alcanzado tensiones de circuito abierto de al menos 0,9 V y hasta 1,5 V, lo cual es suficiente para una eficiente separación del agua conectando sólo dos celdas en serie [21].

La eficiencia de las celdas solares tándem perovskita/perovskita puede ser aún mejorada mediante la optimización del diseño de cada celda por separado.

IX.5.a1. Referencias: Celda solar con perovskitas:

- [1] A. Kojima et al., Organometal halide perovskites as visible-light sensitizers for photovoltaic cells, *Journal American Chemical Society* 131 (2009) 6050-6051.
- [2] National Renewable Energy Laboratory (NREL) Efficiency Chart.
http://www.nrel.gov/pv/assets/images/efficiency_chart.jpg (February, 2016).
- [3] M. Green, et al., The Emergence of Perovskite Solar Cells, *Nature Photonics* 8 (2014) 506-514.
- [4] C. Zuo et al., Advances in Perovskite Solar Cells, *Advanced Science* (2016) 1500324 (16pp).

- [5] H. Snaith, M. Gratzel, Electron and Hole Transport through mesoporous TiO₂ infiltrated with Spiro-MeOTAD, *Adv. Mater.* 19 (2007) 3643-3647.
- [6] G.A. Casas, M.A. Cappelletti, A.P. Cédola, B. Marí Soucase y E.L. Peltzer y Blancá, Analysis of the power conversion efficiency of perovskite solar cells with different materials as Hole-Transport Layer by numerical simulations, *Superlattices and Microstructures, Elsevier.* 107 (2017) 136-143 <http://dx.doi.org/10.1016/j.spmi.2017.04.007>
- [7] K. Taretto, M. Soldera y A. Koffman-Frischknecht, Material Parameters and Perspectives for Efficiency Improvements in Perovskite Solar Cells Obtained by Analytical Modeling, *IEEE Journal of Photovoltaics*, 7 (2017) 206-213.
doi: 10.1109/JPHOTOV.2016.2627398
- [8] L. Peng et al., Device simulation of solid-state perovskite solar cells, 31st European Photovoltaic Solar Energy Conference and Exhibition (2015).
- [9] Y. Zhou, A. Gray-Weale, A numerical model for charge transport and energy conversion of perovskite solar cells, *Phys. Chem. Chem. Phys.*, 18 (2016) 4476-4486.
- [10] K. Adhikari et al., Comparative study on MAPbI₃ based solar cells using different electron transporting materials, *Phys. Status Solidi C* (2015) 1-5.
- [11] M. Javad Taghavi et al., Modeling of optical losses in perovskite solar cells, *Superlattices and Microstructures* 97 (2016) 424-428.
- [12] F. Liu et al., Numerical simulation: Toward the design of high-efficiency planar perovskite solar cells, *Applied Physics Letters* 104 (2014) 253508 (4pp).
- [13] Sato S et al 2009 Degradation modeling of InGaP/GaAs/Ge triple-junction solar cells irradiated with various-energy protons. *Sol. Energy Mater. Sol. Cells* 93 768–73
- [14] Rong W et al., Effects of 0.28–2.80 MeV proton irradiation on GaInP/GaAs/Ge triple-junction solar cells for space use
Nucl. Instrum. Methods Phys. Res., Sect. B 266 (2008) 745–749
- [15] M.A. Cappelletti, G.A. Casas, D.M. Morales, W. Hasperué y E.L. Peltzer y Blancá. Study of the electrical parameters degradation of GaAs sub-cells for triple junction space solar cells by computer simulation. *Semiconductor Science and Technology* 31 (2016) 115020.
doi:10.1088/0268-1242/31/11/115020
- [16] K. Bush et al., 23.6%-efficient monolithic perovskite/silicon tandem solar cells with improved stability, *Nature Energy* 2 (2017) 17009. doi:10.1038/nenergy.2017.9
- [17] M. Filipi et al., CH₃NH₃PbI₃ perovskite / silicon tandem solar cells: characterization based optical simulations, *Opt. Express* 23, (2015) A263-A278
- [18] K. Jäger et al., Numerical optical optimization of monolithic planar perovskite-silicon tandem solar cells with regular and inverted device architectures, *Opt. Express* 25 (2017) A473-A482
- [19] Y. Hu et al., Hybrid Perovskite/Perovskite Heterojunction Solar Cells, *ACS Nano*, 10 (2016) 5999–6007. doi: 10.1021/acsnano.6b01535
- [20] D. Forgács et al., Efficient Monolithic Perovskite/Perovskite Tandem Solar Cells, *Adv. Energy Mater.* 7 (2017) 1602121
- [21] J. Luo et al., Water photolysis at 12.3% efficiency via perovskite photovoltaics and Earth-abundant catalysts, *Science* 345 (2014) 1593.
- [22] I. Castelli et al., *Energy & Environmental Science* 5, 5814-5819 (2012).

- [23] K. Zeng, D. Zhang, *Pror. Energy Combust. Sci.* 36, 307–326 (2010).
 [24] T. Jacobsson et al., *Energy Environ. Sci.* 6, 3676–3683 (2013).
 [25] S. Y. Reece et al., *Science* 334, 645–648 (2011).
 [26] O. Khaselev, J. A. Turner, *Science* 280, 425–427 (1998).

IX.5.b. Celda solar con puntos cuánticos.

Contribución del Dr. Ariel Cédola.

Las celdas solares nanoestructuradas son ampliamente estudiadas en la actualidad por su capacidad de ofrecer eficiencias de conversión de potencia muy por encima del límite teórico determinado por Shockley-Queisser para dispositivos de una juntura [1]. En principio, las nanoestructuras semiconductoras incrementan la eficiencia de las celdas solares ya que posibilitan la absorción de fotones de baja energía, que de otro modo son ignorados. Las implementaciones más relevantes de celdas solares nanoestructuradas están basadas en nanohilos [2, 3], pozos cuánticos (QWs) [4, 5] y puntos cuánticos (QDs) [6-8]. Las celdas de puntos cuánticos (QDSCs) particularmente han despertado gran interés por tratarse de una de las más viables implementaciones del concepto de celda solar de banda intermedia (IBSC) [9]. Bajo máxima concentración teórica de luz solar (46211 soles) la predicción de la eficiencia de una IBSC óptima excede 63% [10], un valor mucho mayor al límite asociado a celdas convencionales operando bajo idénticas condiciones, no superior al 41% de acuerdo a la teoría del balance detallado.

El concepto de IBSC se muestra en la Figura 1 [11]. Una minibanda de energía introducida en la banda prohibida de la región de absorción de una estructura p-i-n huésped, denominada barrera, facilita la creación de pares electrón-hueco adicionales por acción de fotones con energías por debajo de la banda prohibida de la barrera, a partir de un proceso de dos pasos, indicados en la figura como los mecanismos 1 y 2. Los portadores extra se colectan, la fotocorriente aumenta, mientras que la tensión de salida de la celda, en principio, se conserva. Los QDs apilados en capas uniformemente espaciadas son las estructuras más convenientes para la implementación de una banda intermedia (IB) debido a la naturaleza cero-dimensional de su densidad de estados y la característica discreta de sus niveles de energía.

Los QDs son nanoestructuras semiconductoras que confinan electrones en la banda de conducción y huecos en la banda de valencia en las tres dimensiones espaciales, a una escala nanométrica [12]. La estructura de bandas de los QDs consiste de un nivel de energía fuertemente confinado, el estado base (GS), uno o más estados excitados (ES), ambos con confinamiento cero-dimensional, y un estado de energía bidimensional asociado a la wetting layer (WL), que es una fina capa del mismo material que los QDs, creada durante la etapa inicial del crecimiento y sobre la cual se forman los QDs.

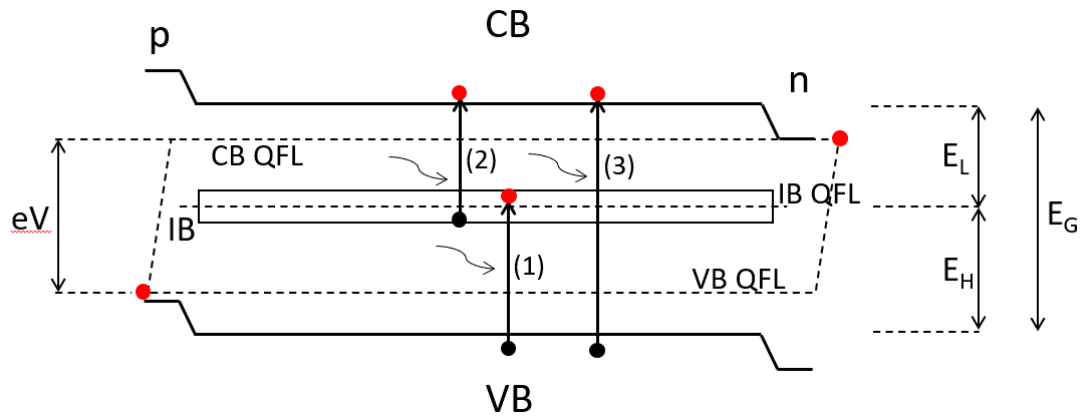


Fig. 1. Diagrama de bandas de energía de una celda solar de banda intermedia mostrando absorción de fotones con energías por debajo (1 y 2) y por encima (3) del bandgap.

La técnica de Stranski-Krastanow es una de las más utilizadas para la fabricación de QDs. Consiste básicamente en la deposición del material semiconductor de los QDs sobre un sustrato con constante de red levemente diferente, a partir de procesos bien conocidos como Molecular Beam Epitaxy (MBE) o Metal-Organic Vapor Phase Epitaxy (MOVPE). Producto de la tensión (strain) inducida por el desajuste entre ambas redes cristalinas en la interface, una vez que el espesor de la capa bidimensional WL supera un valor crítico de algunas monocapas atómicas, los átomos del material depositado tienden a auto-organizarse formando pequeñas islas tridimensionales, que constituyen los denominados puntos cuánticos auto-ensamblados [13]. La altura y el diámetro de los QDs son del orden de algunos nanómetros y decenas de nanómetros, respectivamente.

La separación de energía entre niveles, y en consecuencia el espectro de absorción de los QDs, están estrechamente relacionados a su tamaño y forma (pirámide, lente, cono) [14]. La no uniformidad en las características de los QDs, inherente al proceso de fabricación, es lo que da lugar al ensanchamiento del espectro de fotoluminiscencia. A mayor tamaño de los QDs mayor profundidad en el confinamiento de los niveles de energía, es decir, mayor longitud de onda pico de GS. Dado que el coeficiente de absorción de la WL es mucho mayor que los de GS y ES a relativamente bajas densidades de QDs, del orden de 10^{10} cm^{-2} , la WL juega un rol clave en la fotogeneración de portadores por luz solar.

Básicamente, las QDSCs son estructuras semiconductoras p-i-n con varias capas de QDs apiladas dentro de la región intrínseca, separadas unas de otras por una distancia apropiada a fines de reducir la generación de dislocaciones y el deterioro de la homogeneidad en el tamaño de los QDs de capa a capa, inducidos por la acumulación de tensión interna en la red [16]. Para contrarrestar los efectos adversos del apilamiento de QDs, la técnica del balanceo de tensión (strain-balancing) es comúnmente aplicada. Consiste en crecer capas separadoras utilizando materiales que ejercen una tensión extensiva, que compensa la tensión compresiva introducida por las nanoestructuras [17-19]. Las celdas solares de GaAs con capas de QDs de InAs (QDSCs de InAs/GaAs) son los prototipos más estudiados, debido a la madurez tecnológica de los procesos involucrados en el crecimiento de dispositivos basados en estos materiales, y el vasto conocimiento que se ha

adquirido en las últimas décadas sobre nanoestructuras de InAs, y principalmente sobre sistemas InAs/GaAs, tanto teórica como experimentalmente [14, 20-26]. El método de balanceo de tensión se ha aplicado con éxito a sistemas de InAs/GaAs a partir del uso de capas de GaNAs, dado que este compuesto posee un constante de red menor que el GaAs [18], y también del crecimiento de delgadas capas de GaAsSb cubriendo los QDs de InAs, lo que ha demostrado incrementar la densidad de QDs, reducir la fusión de QDs y mejorar la calidad del cristal [27, 28].

Los QDs absorben fotones con longitudes de onda mayores al límite intrínseco impuesto por la banda prohibida de la barrera, lo que conduce a un ensanchamiento de la respuesta espectral de la celda solar y, si los portadores fotogenerados en los niveles de los QDs son extraídos eficientemente, a un incremento de la corriente de cortocircuito (J_{sc}). Una noción acerca de la expansión de la respuesta espectral de la celda, en este caso de la eficiencia cuántica externa (EQE) versus longitud de onda, como función del número de capas y densidad de QDs, puede verse en la Figura 2a. La estructura *p-i-n* sin QDs, denominada celda de referencia, muestra una EQE angosta con longitud de onda umbral cercana a los 880 nm para el GaAs (curva azul). Al introducir una cierta cantidad de capas de QDs dentro de la región de vaciamiento de la estructura (región intrínseca o no dopada), la EQE se expande como consecuencia de la absorción de fotones de baja energía por parte de los QDs (curva violeta). Al incrementar el número de capas embebidas la EQE vuelve a ensancharse debido a la ampliación del espectro de absorción de los QDs (curva verde), y si la densidad de QDs en cada capa es multiplicada, la EQE se ensancha aún más producto del incremento de la absorción en los niveles confinados (curva roja). La Figura 2b muestra la superposición de todas estas curvas en un único gráfico.

Mediciones experimentales de EQE de celdas de GaAs con 0, 10, 20 y 40 capas de QDs de InAs han sido publicadas en [29].

En ensanchamiento de la EQE demuestra que los fotones con bajas energías absorbidos por los QDs contribuyen a la fotocorriente total de la celda solar. Esto significa que los portadores fotogenerados en los niveles confinados y en la WL son capaces de escapar hacia la barrera, donde son separados y arrastrados por el campo eléctrico hasta alcanzar los contactos. De este modo la fotocorriente de la celda tiene dos componentes, asociadas a los portadores generados

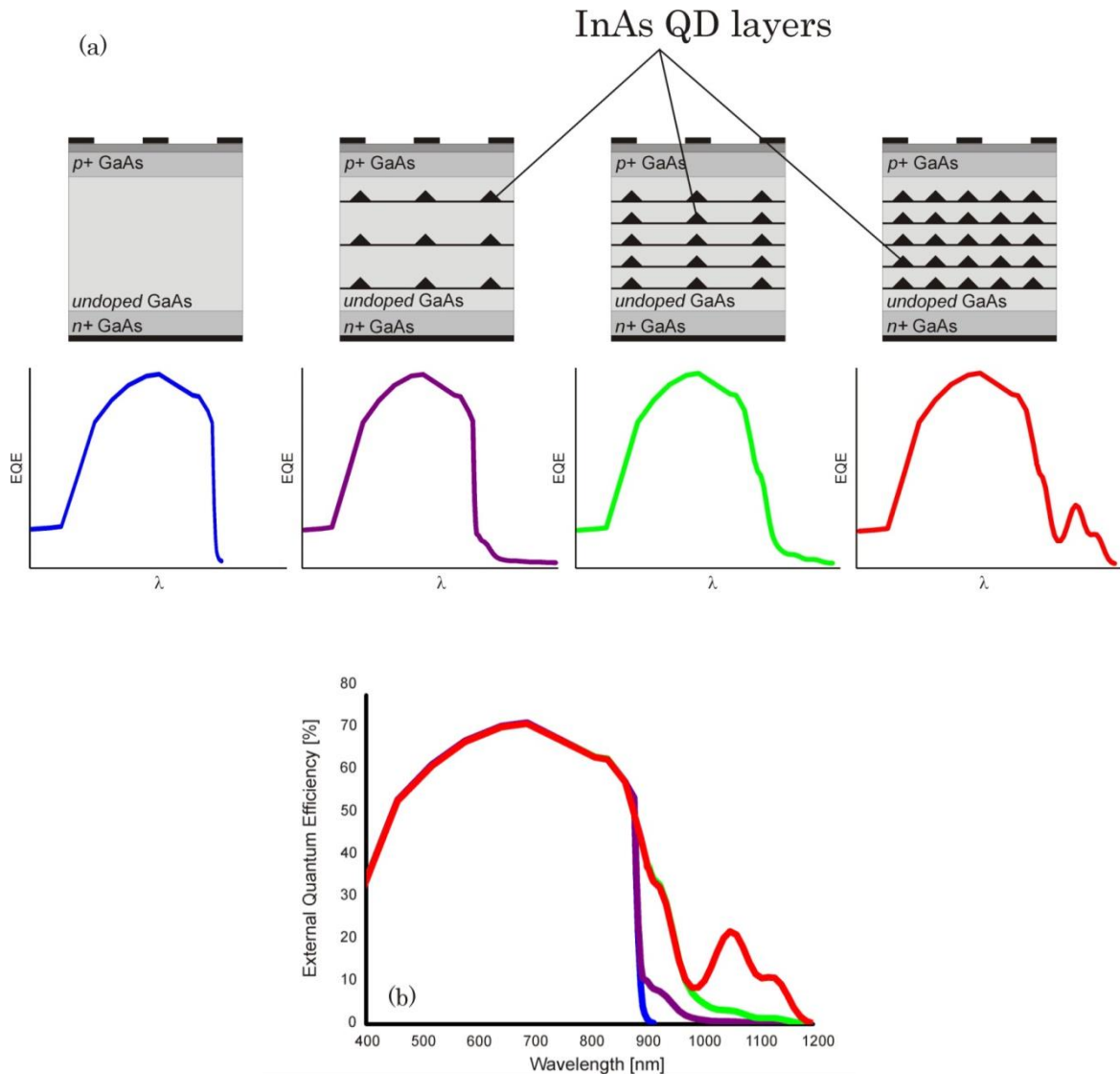


Fig. 2. Ensanchamiento de la respuesta espectral de QDSCs de InAs/GaAs en función del número de capas y densidad de QDs: (a) mejora paso a paso de la EQE versus longitud de onda (λ); (b) superposición de todas las curvas.

por la luz solar tanto en la barrera como en los QDs. Estas dos componentes pueden o no adicionarse linealmente una a otra para constituir la corriente total del dispositivo. El escape de los portadores fotogenerados en los QDs hacia la barrera es un proceso netamente térmico a temperatura ambiente, lo que representa una desventaja ya que este acoplamiento conduce a una banda prohibida más angosta y en consecuencia a una tensión de circuito abierto (V_{OC}) reducida. El escape inducido por fotones de baja energía ha sido observado experimentalmente solo a muy bajas temperaturas [30].

La dinámica de los portadores en los niveles de energía de los QDs es compleja, y no todos los electrones y huecos creados en los estados de los QDs son capaces de alcanzar la barrera y ser colectados. Se establece un balance entre varios procesos, como lo son la captura desde la barrera hacia los QDs, la relajación inter-subbanda desde un estado confinado de los QDs hacia uno más bajo, el escape hacia estados de energía más altos, fotogeneración y recombinación. Los electrones y los huecos siguen en principio diferente dinámica, aunque bajo ciertas condiciones ambos portadores pueden mostrar un comportamiento exci-

tónico [31]. Los tiempos de captura y relajación están en el rango de los picosegundos, mientras que los tiempos de escape dependen de la separación en energía entre los estados inicial y final [32]. Los procesos de dispersión son asistidos por múltiples fonones ópticos longitudinales (LO) [33], dado que la población de portadores en los estados de los QDs en QDSCs no es suficientemente elevada como para permitir una ocurrencia significativa de eventos de dispersión portador-portador [34]. La Figura 3 muestra un esquema idealizado de las bandas de energía de las capas de QDs de InAs en GaAs. Como puede apreciarse el alineamiento de bandas es de tipo I y el pozo de potencial para huecos es claramente mucho más chico que aquel para electrones. Las flechas indican los procesos de transferencia de carga inter-subbanda y banda a banda arriba mencionados.

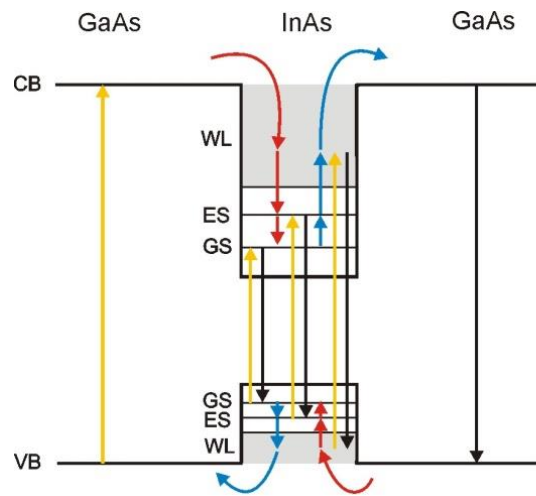


Fig. 3. Diagrama simplificado de bandas de energía de una capa de QDs de InAs embebida en GaAs. El área sombreada denota la densidad de estados 2D asociada a la WL. Las flechas de colores representan los procesos de portadores: captura/relajación (rojo), escape (azul), fotogeneración (amarillo) y recombinación (negro).

Los portadores extra generados en los QDs que son extraídos eficientemente hacia los estados de la barrera contribuyen a incrementar la fotocorriente de la QDSC, es decir, J_{sc} aumenta. Sin embargo, como ocurre prácticamente con todas las implementaciones experimentales de QDSCs reportadas hasta el momento, el incremento de J_{sc} es débil y siempre es acompañado por una reducción en V_{oc} [15, 35-40]. Estas características pueden apreciarse en una serie de mediciones de curvas tensión-corriente (J - V) representativas, publicadas recientemente en la literatura para QDSCs de InAs/GaAs con 5 capas de QDs de diferentes tamaños [15], QDSC de InGaAs/GaAs con 10 capas de QDs [37], QDSCs de InAs/GaAs con 10 capas de QDs y reducción del strain [39] y QDSCs de 20 capas de QDs de InAs/GaAs con WL reducida y dopado [40]. La reducción en V_{oc} es el resultado de la captura de portadores desde la barrera hacia los estados confinados de los QDs y subsecuente recombinación, la cual empeora a medida que el confinamiento de portadores en GS se incrementa [15, 41-43]. Los estados confinados de los QDs se convierten en efectivos centros de atrapamiento cuando el dispositivo es polarizado en directa.

El incremento débil de J_{sc} y la no despreciable caída de V_{oc} observados experimentalmente impiden un incremento apreciable de la eficiencia de conversión de potencia de las QDSCs con respecto a las celdas solares convencionales. A pesar de la implementación de numerosas

estrategias de diseño como por ejemplo la inserción de capas de dopado tipo delta [36, 44], el dopado directo de los QDs [45], el balanceo de tensión [46], la optimización de la cobertura del InAs [29] e ingeniería de WL [40], la eficiencia de las QDSCs de InAs/GaAs se ha mantenido siempre por debajo del 20%. Esta relativamente baja eficiencia, en comparación con la máxima predicha (63%), ha motivado el estudio profundo de estos dispositivos por medio de la modelización y simulación numérica [42]. Se ha demostrado a partir de sofisticados modelos electro-ópticos acoplados, que la implementación de mecanismos de atrapamiento de fotones (light trapping) y reciclado de fotones (photon recycling) en celdas solares de película delgada (thin film) con QDs, podría conducir a eficiencias superiores respecto a celdas convencionales sin nanoestructuras, operando bajo luz solar concentrada [47, 48]. Estos desarrollos se encuentran actualmente en fase experimental. Asimismo, capas de QDs con alineamiento de bandas de energía de tipo II surgen como alternativa a las heterojunturas de tipo I por su reducida recombinación, producto de la ausencia de solapamiento entre las funciones de onda de electrones y huecos. Recientemente se han reportado tiempos de recombinación en QDs de InAs embebidos en GaAsSb, del orden de 100 ns - 200 ns, obtenidos a partir de la inserción de pozos de potencial de GaAs entre los estados de los QDs y la barrera [49].

IX.5.b1. Referencias: Celda solar con puntos cuánticos

- [1] W. Shockley, H. Queisser, Detailed balance limit of efficiency of p-n junction solar cells, *J. Appl. Phys.* 32 (1961) 510-519.
- [2] K. Sun, A. Kargar, N. Park, K. N. Madsen, P. W. Naughton, T. Bright, Y. Jing, D. Wang, Compound semiconductor nanowire solar cells, *IEEE J. Sel. Topics Quantum Electron.* 17 (2011) 1033-1049.
- [3] R. R. LaPierre, A. C. E. Chia, S. J. Gibson, C. M. Haapamaki, J. Boulanger, R. Yee, P. Kuyanov, J. Zhang, N. Tajik, N. Jewell, K. M. A. Rahman, III-V nanowire photovoltaics: review of design for high efficiency, *Phys. Status Solidi RRL* 7 (2013) 815-830.
- [4] K.W.J. Barnham, I. Ballard, J.P. Connolly, N.J. Ekins-Daukes, B.G. Kluffinger, J. Nelson, C. Rohr, Quantum well solar cells, *Physica E* (2002) 27-36.
- [5] M. Mazzer, K.W.J. Barnham, I.M. Ballard, A. Bessiere, A. Ioannides, D.C. Johnson, M.C. Lynch, T.N.D. Tibbits, J.S. Roberts, G. Hill, C. Calder, Progress in quantum well solar cells, *Thin Solid Films* 511-512 (2006) 76-83.
- [6] V. Aroutiounian, S. Petrosyan, A. Khachatryan, and K. Touryan, Quantum dot solar cells, *J. Appl. Phys.* 89 (2001) 2268-2271.
- [7] A.J. Nozik, Quantum dot solar cells, *Physica E* (2002) 115-120.
- [8] A. Martí, E. Antolín, C. R. Stanley, C. D. Farmer, N. López, P. Díaz, E. Cánovas, P. G. Linares, A. Luque, Production of photocurrent due to intermediate-to-conduction-band transitions: a demonstration of a key operating principle of the intermediate-band solar cell, *Phys. Rev. Lett.* 97 (2006) (247701-1-247701-4).

- [9] A. Luque, A. Martí, Increasing the efficiency of ideal solar cells by photon induced transitions at intermediate levels, *Phys. Rev. Lett.* 78 (1997) 5014-5017.
- [10] A. Luque, A. Martí, The intermediate band solar cell: progress toward the realization of an attractive concept, *Adv. Mater.* 22 (2010) 160–174.
- [11] A. Luque, A. Martí and C. Stanley, Understanding intermediate-band solar cells, *Nature Photonics* 6 (2012) pp. 146–152.
- [12] D. Bimberg, M. Grundmann, N. N. Ledentsov, *Quantum Dot Heterostructures*, Wiley, New York, 1999.
- [13] S. Franchi, G. Trevisi, L. Seravalli, and P. Frigeri, Quantum dot nanostructures and molecular beam epitaxy, *Prog. Cryst. Growth Charact. Mater.*, vol. 47, no. 23, pp. 166–195, 2003.
- [14] M. Grundmann, O. Stier, D. Bimberg, InAs-GaAs pyramidal QDs: strain distribution, optical phonons and electronic structure, *Phys. Rev. B* 52 (1995) 11969-11981.
- [15] D. Guimard, R. Morihara, D. Bordel, K. Tanabe, Y. Wakayama, M. Nishioka, Y. Arakawa, Fabrication of InAs/GaAs quantum dot solar cells with enhanced photocurrent and without degradation of open circuit voltage, *Appl. Phys. Lett.* 96 (2010) (203507-1-203507- 3).
- [16] Z.R. Wasilewski, S. Fafard, J.P. McCaffrey, Size and shape engineering of vertically stacked self-assembled quantum dots, *Journal of Crystal Growth*, Volumes 201–202, (1999) pp. 1131-1135.
- [17] Zhang, Wei and Uesugi, Katsuhiko and Suemune, Ikuo, The application of an In-GaAs GaAsN strain-compensated superlattice to InAs quantum dots, *Journal of Applied Physics* 99 (2006) 103103.
- [18] Oshima, Ryuji and Hashimoto, Takayuki and Shigekawa, Hidemi and Okada, Yoshitaka, Multiple stacking of self-assembled InAs quantum dots embedded by GaNAs strain compensating layers, *Journal of Applied Physics* 100 (2006) 083110.
- [19] Laghumavarapu, R. B. and El-Emawy, M. and Nuntawong, N. and Moscho, A. and Lester, L. F. and Huffaker, D. L., Improved device performance of InAs GaAs quantum dot solar cells with GaP strain compensation layers, *Applied Physics Letters* 91 (2007) 243115.
- [20] C. Pryor, Eight-band calculations of strained InAs/GaAs quantum dots compared with one-, four-, and six-band approximations, *Phys. Rev. B*, 57, 12 (1998) pp. 7190-7195.
- [21] R. Heitz, I. Mukhametzhanov, O. Stier, A. Hoffmann, A. Madhukar and D. Bimberg, Electronic Properties of InAs/GaAs Quantum Dots, in *Physics and Applications of Semiconductor Quantum Structures*, ed. J.-C. Woo and T. Yao, IOP Publishing, Bristol (2001) pp. 1-20.
- [22] W. Yang, R. R. Lowe-Webb, H. Lee and P. C. Sercel, Effect of carrier emission and retrapping on luminescence time decays in InAs/GaAs quantum dots, *Phys. Rev. B* 56, 20 (1997) pp. 13314-13320.
- [23] D. A. Yarotski, R. D. Averitt, N. Negre, S. A. Crooker, A. J. Taylor, G. P. Donati, A. Stintz, L. F. Lester, and K. J. Malloy, Ultrafast carrier relaxation dynamics in self-assembled InAs/GaAs quantum dots, *J. Opt. Soc. Am. B* 19 (2002) pp. 1480-1484.
- [24] W. -H. Chang, W. Y. Chen, T. M. Hsu, N. -T. Yeh, and J. -I. Chyi, Hole emission processes in InAs/GaAs self-assembled quantum dots, *Phys. Rev. B* 66, 19 (2002) p. 195337.
- [25] T. Piwonski, I. O'Driscoll, J. Houlihan, G. Huyet, R. J. Manning and A. V. Uskov, Carrier capture dynamics of InAs GaAs quantum dots, *Appl. Phys. Lett.* 90 (2007) p. 122108.

- [26] I. Saidi, K. Sellami, M. Yahyaoui, C. Testelin and K. Boujdaria, Electron and hole energy levels in InAs/GaAs quantum dots: Size and magnetic field effects, *J. Appl. Phys.* 109 (2011) p. 033703.
- [27] H.Y. Liu, M.J. Steer, T.J. Badcock, D.J. Mowbray, M.S. Skolnick, P. Navaretti, K.M. Groom, M. Hopkinson, R.A. Hogg, Long-wavelength light emission and lasing from InAs/GaAs quantum dots covered by a GaAsSb strain-reducing layer, *Applied Physics Letters* 86 (2005) 143108-1–143108-3.
- [28] Yu-An Liao, Wei-Ting Hsu, Pei-Chin Chiu, Jen-Inn Chyi, and Wen-Hao Chang, Effects of thermal annealing on the emission properties of type-II InAs/GaAsSb quantum dots, *Appl. Phys. Lett.* 94 (2009) 053101.
- [29] C. G. Bailey, D. V. Forbes, S. J. Polly, Z. S. Bittner, Y. Dai, Ch. Mackos, R. P. Raffaele, S. M. Hubbard, Open-circuit voltage improvement of InAs/GaAs quantum-dot solar cells using reduced InAs coverage, *IEEE J. Photovolt.* 4 (2012) 269-275.
- [30] T. Nozawa, H. Takagi, K. Watanabe and Y. Arakawa, Direct observation of two-step photon absorption in an InAs/GaAs single quantum dot for the operation of intermediate-band solar cells, *Nano Lett.* 15 (2015) pp. 4483-4487.
- [31] A. Markus, J. X. Chen, O. Gauthier-Lafaye, J. -G. Provost, C. Paranthoën, A. Fiore, Impact of intraband relaxation on the performance of a quantum-dot laser, *IEEE J. Sel. Topics Quantum Electron.* 9 (2003) 1308-1314.
- [32] I. O'Driscoll, T. Piwonski, C.-F. Schlessner, J. Houlihan, G. Huyet, and R. Manning, Electron and hole dynamics of InAs/GaAs quantum dot semiconductor optical amplifiers, *Appl. Phys. Lett.*, vol. 91 (2007.) p. 071111.
- [33] I. Magnúsdóttir, A. V. Uskov, S. Bischoff, B. Tromborg, and J. Mørk, One- and two-phonon capture processes in quantum dots, *J. Appl. Phys.*, vol. 92 (2002) pp. 59821-1–59821-9.
- [34] N. Majer, K. Ludge, J. Gomis-Bresco, S. Dommers-Volkel, U. Woggon and E. Scholl, Impact of carrier–carrier scattering and carrier heating on pulse train dynamics of quantum dot semiconductor optical amplifiers, *Appl. Phys. Lett.*, vol. 99 (2011) pp. 131102-1–131102-3.
- [35] X. J. Shang, J. F. He, H. L. Wang, M. F. Li, Y. Zhu, Z. C. Niu, Y. Fu, Effect of built-in electric field in photovoltaic InAs quantum dot embedded GaAs solar cell, *Appl. Phys. A* 103 (2011) pp. 335-341.
- [36] K. Sablon, J. W. Little, V. Mitin, A. Sergeev, N. Vagidov, K. Reinhardt, Strong enhancement of solar cell efficiency due to quantum dots with built-in charge, *Nano Lett.* 11 (2011) 2311-2317.
- [37] G. Jolley, L. Fu, H. F. Lu, H. H. Tan, C. Jagadish, The role of intersubband optical transitions on the electrical properties of InGaAs/GaAs quantum dot solar cells, *Prog. Photovolt: Res. Appl.* 21 (2013) 736-746.
- [38] S. Hubbard, D. Wilt, S. Bailey, D. Byrnes, R. Raffaele, OMVPE grown InAs quantum dots for application in nanostructured photovoltaics, in: *Proceedings of 4th IEEE World Conference on Photovoltaic Energy Conversion 2006* , 118-121.
- [39] Ch. G. Bailey, D. V. Forbes, R. P. Raffaele and S. M. Hubbard, Near 1 V open circuit voltage InAs/GaAs quantum dot solar cells, *Appl. Phys. Lett.* 98 (2011) p. 163105.

- [40] K. Sablon, J. Little, N. Vagidov, Y. Li, V. Mitin, A. Sergeev, Conversion of above- and below-bandgap photons via InAs quantum dot media embedded into GaAs solar cell, *Appl. Phys. Lett.* 104 (2014) (253904-1-253904-5).
- [41] G. Jolley, H. F. Lu, L. Fu, H. H. Tan, Ch. Jagadish, Electron-hole recombination properties of In_{0.5}Ga_{0.5}As/GaAs quantum dot solar cells and the influence on the open circuit voltage, *Appl. Phys. Lett.* 97 (2010) p. 123505.
- [42] M. Gioannini, A. Cedola, N. Di Santo, F. Bertazzi, F. Cappelluti, Simulation of quantum dot solar cells including carrier intersubband dynamics and transport, *IEEE J. Photovolt.* 3 (2013) 1271-1278.
- [43] K. Tanabe, D. Guimard, D. Bordel and Y. Arakawa, High-efficiency InAs/GaAs quantum dot solar cells by metalorganic chemical vapor deposition, *Applied Physics Letters*, 100 (2012) p. 193905.
- [44] S. J. Polly, D. V. Forbes, K. Driscoll, S. Hellstrom, S. M. Hubbard, Delta-doping effects on quantum-dot solar cells, *IEEE J. Photovolt.* 4 (2014) 1079-1085.
- [45] X. Yang, K. Wang, Y. Gu, H. Ni, X. Wang, T. Yang, Z. Wang, Improved efficiency of InAs/GaAs quantum dots solar cells by Si-doping, *Sol. Energy Mater. Sol. Cells* 113 (2013) 144-147.
- [46] N. J. Ekins-Daukes, K. W. J. Barnham, J. P. Connolly, J. S. Roberts, J. C. Clark, G. Hill, M. Mazzer, Strain-balanced GaAsP/InGaAs quantum well solar cells, *Appl. Phys. Lett.* 75 (1999) 4195–4197.
- [47] F. Cappelluti, D. Kim, M. van Eerden, A.P. Cédola, T. Aho, G. Bissels, F. Elsehrawy, J. Wu, H. Liu, P. Mulder, G. Bauhuis, J. Schermer, T. Niemi, M. Guina, Light-trapping enhanced thin-film III-V quantum dot solar cells fabricated by epitaxial lift-off, *Solar Energy Materials and Solar Cells*, Available online 3 January 2018.
- [48] Ariel Cedola, Dongyoung Kim, Alberto Tibaldi, Mingchu Tang, Arastoo Khalili, Jiang Wu, Huiyun Liu, and Federica Cappelluti, Physics-based modeling and experimental study of Si-doped InAs/GaAs quantum dot solar cells, *INDAWI Journal of Photoenergy*, in press.
- [49] D. Sato, J. Ota, K. Nishikawa, Y. Takeda, N. Miyashita and Y. Okada, Extremely long carrier lifetime at intermediate states in wall-inserted type II quantum dot absorbers, *J. Appl. Phys.* 112 (2012) p. 094305.

IX-6. Diodo Laser

La cara opuesta a lo que se acaba de ver, donde los dispositivos mediante una excitación externa generan una respuesta eléctrica, son los dispositivos emisores, a los cuales se los excita eléctricamente y su respuesta es una emisión de radiación electromagnética, en alguna de sus frecuencias. Si la frecuencia se encuentra dentro del espectro visible, la emisión será luz de algún color.

En estos casos, lo que ocurre es que un electrón que se encuentra en la banda de conducción, pierde su excitación y pasa a la banda de valencia recombinándose, como se ve en la Figura IX-26.

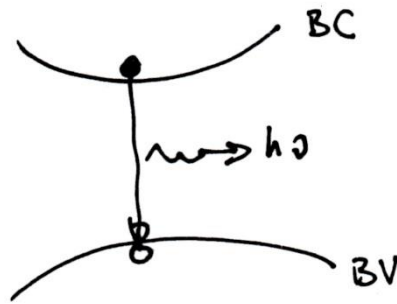


Figura IX-26. Esquema de la emisión de un fotón.

Esta acción implica la emisión de un fotón de energía $h\nu$ igual al ancho de la banda prohibida, es decir:

$$h\nu = E_g$$

Ejemplo IX-3. Determinar la longitud de onda en el Arseniuro de Galio cuando un electrón pierde su excitación.

El Arseniuro de Galio tiene un ancho de banda prohibida de 1,4eV, si se produjera la pérdida de la excitación, la frecuencia y longitud de onda serían:

$$h\nu = 1,4 \Rightarrow \nu = 1,4/h = 3,38 \times 10^{14} \text{ s}^{-1}$$

$$\lambda = c/\nu = 0,8 \mu\text{m}$$

La emisión se encuentra en la zona del infrarrojo.

Estos simples dispositivos se denominan LED por su nombre en inglés “*Light Emitting Diode*” (Diodos Emisores de Luz). Debe quedar claro, que la frecuencia de la emisión depende del ancho de la banda prohibida del material que conforme al diodo. Por lo que la investigación sobre materiales es un hecho imprescindible para obtener diferentes emisiones.

Un dispositivo interesante en el tema de las emisiones es el diodo laser, del inglés “*light amplification by stimulated emission of radiation*” (amplificación de luz por emisión estimulada de radiación). La emisión laser tiene la característica de ser luz monocromática, coherente y no divergente. La condición para que haya emisión estimulada es que en las regiones neutras se produzca inversión de población. Para lo cual debe haber una fuerte inyección de portadores minoritarios desde las regiones opuestas. También debe haber una gran cantidad de fotones generados a través de la recombinación de portadores minoritarios en las regiones neutras. Las zonas de recombinación deben ser estrechas, la cavidad donde se produce la recombinación debe estar confinada por materiales que tengan un alto índice de refracción, de manera que los fotones generados no penetren en estos materiales y todos ellos puedan salir del diodo evitando pérdidas por absorción en los materiales lindantes a la zona de recombinación.

En la Figura IX-27 se ve el esquema de bandas para que se desarrolle el proceso laser.

Si a la unión se le aplica una potencial en directa, se comportará como un diodo normal inyectando electrones en el lado p y huecos en el lado n , Figura IX-27(a). La condición para que haya inversión de población es que los estados excitados están más ocupados que los del nivel fundamental, en un diodo común esto es a la inversa. Si se incrementa el potencial aplicado, lo que ocurre es que más portadores mayoritarios de la zona n ocuparán estados vacíos en la banda de conducción de la zona p , como portadores minoritarios, logrando una concentración de portadores minoritarios mayor que en condiciones de equilibrio térmico. Lo mismo ocurre para los huecos en la banda de valencia. En general, se ha podido determinar que los semiconductores con “gap directo” son los que generan transiciones radiativas más probables y más útiles para ser usados en la optoelectrónica. Los semiconductores que mejor se adaptan a esta circunstancia son los que pertenecen a la familia de los compuestos III-V, llamados así porque son compuestos formados por elementos de la columna III y V de la tabla periódica. El más comúnmente mencionado es el GaAs y compuestos terciarios con estos elementos, por ejemplo el $Al_yGa_{1-y}As$, etc. estos compuestos presentan anchos diferentes de la banda prohibida con respecto al GaAs.

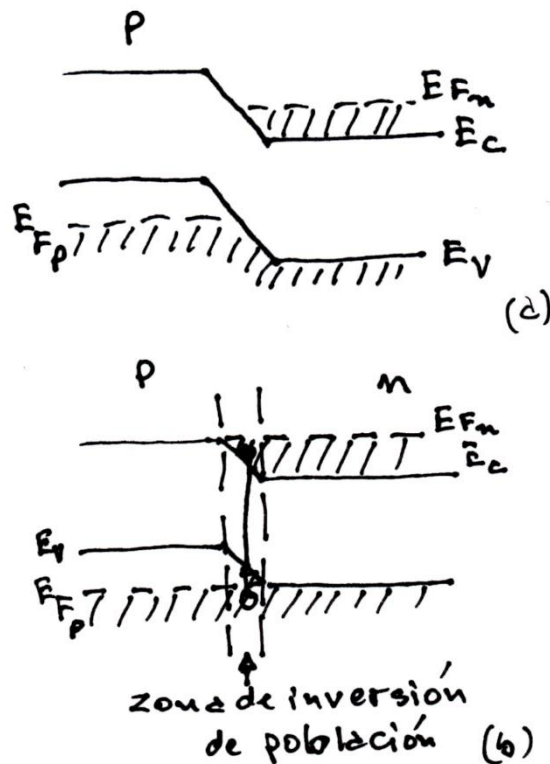


Figura IX-27. (a) unión pn altamente dopada con potencial en directa aplicado, (b) unión en condiciones de inversión de población y emisión.

Se puede hacer una breve descripción del laser $GaAs/Al_yGa_{1-y}As$. Una capa muy delgada de $GaAs$ se encuentra entre dos capas de $Al_yGa_{1-y}As$ altamente dopadas, una será p y la otra será n . Como se puede ver, bajo estas condiciones se tienen dos uniones de materiales diferentes, $Al_yGa_{1-y}As(p)/GaAs$ y $GaAs/Al_yGa_{1-y}As(n)$, se han conformado dos heterouniones. El funcionamiento del diodo está basado en este tipo de estructuras. Estas dos heteroestructuras, generadas por las uniones de los materiales $p/GaAs/n$, forman dos pozos de potencial, de alturas ΔE_c y ΔE_v , para los electrones y para los huecos, respectivamente. Como se puede ver en la Figura IX-28.

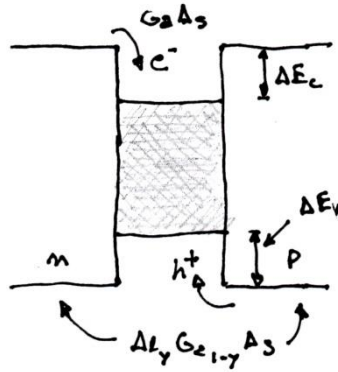


Figura IX-28. Esquema del diagrama de bandas del diodo laser.

Estos pozos que se forman, sirven para el confinamiento de los portadores. El valor del ancho de la banda prohibida del compuesto $Al_xGa_{1-y}As$ (sea n o p) depende de la fracción molar “ y ” la cual también determina la diferencia entre los anchos de bandas prohibidas del GaAs y de sus vecinos. El diodo está fuertemente polarizado en directa, esto hace que gran cantidad de electrones y huecos (alta inyección) “caigan” en los pozos correspondientes, provocando la inversión de población. La zona del GaAs, que es la capa interna entre materiales dopados, es donde tiene lugar la recombinación y la emisión estimulada, esta región es llamada “región activa”, cuyo esquema puede verse en la Figura IX-29.

Para tener un buen rendimiento del efecto (alta ganancia), la luz generada por las transiciones de la banda de conducción a la de valencia, debe estar lo más confinada posible a la región activa, evitando que la radiación se disperse hacia el exterior de esta región, perdiéndose fotones y no aportando a la emisión estimulada. Para lograr esto, los compuestos que limitan a la región activa, deben tener un alto índice de refracción, tal que estos se comporten como una guía de ondas. Para incrementar el índice de refracción se debe disminuir el ancho de la banda prohibida, esto hace que se deba tomar una solución de compromiso, pues el decremento de la banda prohibida incrementa el índice de refracción, efecto buscado, pero reduce la diferencia de energía entre la región activa y los laterales reduciendo la altura del pozo. El esquema del diodo laser se muestra en la Figura IX-29(a), en la Figura IX-29(b) se muestra esquemáticamente un corte transversal del diodo.

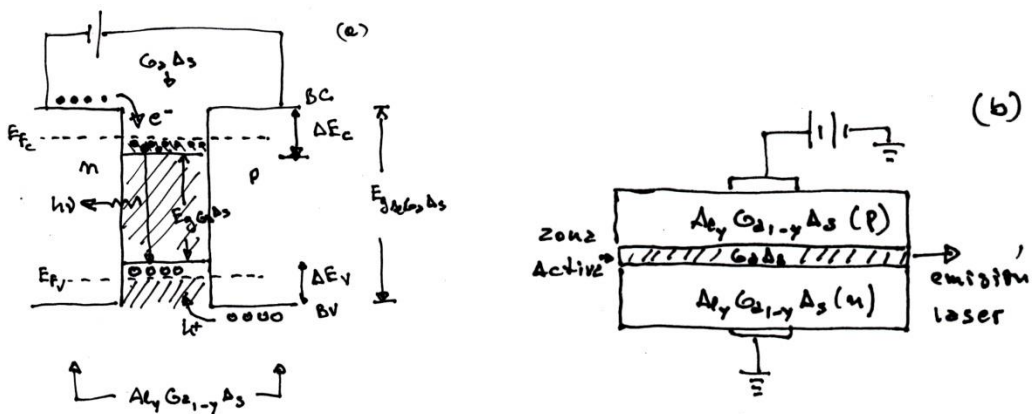


Figura IX-29(a). Esquema de la transición de por medio de la emisión espontánea, (b) esquema transversal del diodo laser.

Se ha mencionado que tanto electrones como huecos se acumularán en el pozo de potencial formado en la zona central, un ancho típico de esos pozos es del orden de 100\AA .

Los diodos que involucran pozos de potencial, movimiento de ambos tipos de portadores, índices de refracción, etc., presentan complicaciones teóricas para su descripción. Hay teorías que abarcan solamente una parte de los efectos, no la totalidad. Necesariamente, el problema debe ser atacado numéricamente, el cálculo computacional hace un gran aporte a la solución del problema, describiendo físicamente al diodo como un todo, con ecuaciones que permiten optimizar la ingeniería del sistema.

Referencias

- Albella, J. M., Martínez-Duart, J. M. *Fundamentos de electrónica física y microelectrónica*. 1996 Addison-Wesley Universidad Autónoma de Madrid.
- Ashcroft, N. W., Mermin, N. D. *Solid State Physics*. 1981 Holt-Saunders International Editions.
- Dalven, R. *Introduction to Applied Solid State Physics. Second Edition*. 1980 Plenum Press. New York and London.
- Ghandhi, S. K. *Semiconductor Power Devices*, 1977 Wiley, New York
- Hess, K. *Advanced Theory of Semiconductor Devices*., 1999 Wiley-Interscience Publication
- Hlawiczka, P. *Introducción a la Electrónica Cuántica*. 1977 Ed. Reverte, S.A.
- Ihm, T. *Semiconductor Nanostructures Quantum States and Electronic Transport*.. 2010 Oxford University Press, Great Clarendon Street, Oxford OX2 6DP
- Kittel, C. *Introduction to Solid State Physics Sixth Edition* 1986 John Wiley and Sons Inc. New York
- Mishra, U. K., Singh, J. *Semiconductor Device Physics and Design*. 2008 Springer The Netherlands
- Pierret, F. R. *Advanced Semiconductor Fundamentals, Volume VI*. 1987 Modular Series on Solid State Devices, R.F Pierret and G. W. Neudeck, Editors. Addison-Wesley Pub. Comp.
- Sapoval, B., Hermann, G. *Physics of Semiconductors*. 1993 Springer-Verlag. New York.
- Shalimova, K. V. *Física de los Semiconductores*. 1975 Editorial MIR, Moscú
- Singh, J. *Semiconductor Devices an Introduction* 1994 Mc Graw Hill International Editions. Electrical Engineering Series.
- Sze, S. M. *Physics of Semiconductor Devices 2nd Edition*. 1981 Wiley-Interscience Publication.
- Sze, S. M., Ng, K. K., *Physics of Semiconductor Devices Third Edition*. 2006 Wiley-Interscience Publication.
- Van der Ziel, A. *Electrónica Física del Estado Sólido*. 1968 Editorial Prentice/Hall International
- Yariv, A. *Quantum Electronics. Third Edition*. 1989 John Wiley and Sons.
- Yepifanov, G. *Physical Principles of Microelectronics*.. 1974 Mir Publishers. Moscow
- Yu, P. Y., Cardona, M. *Fundamentals of Semiconductors Physics and Material Properties*, 1996 Springer-Verlag Berlin Heidelberg

Epílogo

Se ha llegado al final del libro, en el he invertido mucho tiempo de trabajo. Ha sido una tarea gratificante, porque además de aprender muchas cosas nuevas, he podido completar temas para redondear ideas. Todas ellas las he tratado de transmitir en el texto, espero haberlo logrado.

La idea que se mencionó en el prólogo sobre la realización de este libro, tratar de hacer algo ameno, conceptual y que muestre la importancia de la Física en la Ingeniería, siempre estuvo presente, esto también espero haberlo logrado.

La importancia que tiene la Física en la Ingeniería no es un hecho menor, ni tampoco la utilización que la Ingeniería hace de la Física. Sin lugar a dudas, el progreso y el bienestar de la humanidad necesitan que ambos universos se unan, un ejemplo bien claro de esto es el desarrollo del transistor, los tres premios Nobel fueron: John Bardeen, Ingeniero Eléctrico y Físico, ganador de dos premios Nobel, Walter Brattain, Físico y William Shockley, Físico. El trabajo de los tres, junto con la Física y la Ingeniería, dieron lugar a un dispositivo que ha cambiado la vida de las personas por décadas. El impacto generado por este desarrollo ha sido tremendo, en todos los órdenes de la vida.

Espero haber logrado el objetivo que me he propuesto para el desarrollo del libro, esto es, que quede clara la idea sobre la importancia de conocer los fundamentos de la ciencia, en este caso la física, que nos explica el comportamiento de los materiales, los semiconductores, por ejemplo, para poder generar un dispositivo y que nos brinde las propiedades que buscamos, o entender cuales son los alcances del rendimiento de un dispositivo para obtener una correcta aplicación del mismo.

Mis respetos a los que hayan llegado hasta este punto.

El Autor

Peltzer y Blancá, Eitel Leopoldo

Dr. en Física, Lic. en Física Universidad Nacional de La Plata (UNLP). Profesor Titular Dedicación Exclusiva (Profesor Jubilado Activo) Facultad de Ingeniería, UNLP.

1.- Calculated Static and Dynamical Properties of β -Sn and Sn-O Compounds

E.L.Peltzer y Blancá, A.Svane, N.E. Christensen, C.O.Rodríguez and S. Moreno.

Phys. Rev. B48, 15712 (1993).

2.- The magnetization of γ -Fe₄N: theory vs. experiment.

E. L. Peltzer y Blancá, J. Desimoni, N. E. Christensen, H. Emmerich, S. Cottenier.

Physica Status Solidi, 246, 5, 906-928 (2009)

3.- Extraction of the minority carrier transport properties of solar cells using the Hovel model and genetic algorithms

M.A. Cappelletti, A.P. Cédola, L.M.Olivera, G. Casas, J.Osio, E.L. Peltzer y Blancá

Measurement Science and Technology, 31, 25601(9pag) (2020).

Física del estado sólido, Semiconductores, Materiales magnéticos, Hidruros metálicos, Reservorios de Hidrógeno, Estudio de dispositivos electrónicos, modelización. Director "Grupo de Estudio de Materiales y Dispositivos Electrónicos". Facultad de Ingeniería, CONICET.

Secretario de Ciencia y Técnica, UNLP 2007-2010. Secretario Asociación Física Argentina 2002-2006. Coordinador Red COMPUMAT (2009-). Organizador de 10 Workshops Internacionales "Workshop on novel methods for electronic structure calculations". Subsidios recibidos de la Secretaria de Ciencia y Técnica (PICT), CONICET (PIP), UNLP, SECyT-FWO (Bélgica), Beca de la Comunidad Económica Europea. Director de Investigadores, de tesis de doctorado, de trabajos finales.

Peltzer y Blancá, Eitel Leopoldo

Introducción a la física de semiconductores y dispositivos electrónicos / Eitel Leopoldo Peltzer y Blancá ; contribuciones de Ariel P. Cédola ; Marcelo A. Cappelletti ; coordinación general de Eitel Leopoldo Peltzer y Blancá. - 1a ed. - La Plata : Universidad Nacional de La Plata ; EDULP, 2021.

Libro digital, PDF/A - (Libros de cátedra)

Archivo Digital: descarga
ISBN 978-950-34-1986-1

1. Física. 2. Electrónica. I. Cédola, Ariel P., colab. II. Cappelletti, Marcelo A., colab. III. Título. CDD 537.5

Diseño de tapa: Dirección de Comunicación Visual de la UNLP

Universidad Nacional de La Plata – Editorial de la Universidad de La Plata
48 N.º 551-599 / La Plata B1900AMX / Buenos Aires, Argentina
+54 221 644 7150
edulp.editorial@gmail.com
www.editorial.unlp.edu.ar

Edulp integra la Red de Editoriales Universitarias Nacionales (REUN)

Primera edición, 2021
ISBN 978-950-34-1986-1
© 2021 - Edulp

e
exactas

**Edulp**
EDITORIAL DE LA UNLP



UNIVERSIDAD
NACIONAL
DE LA PLATA