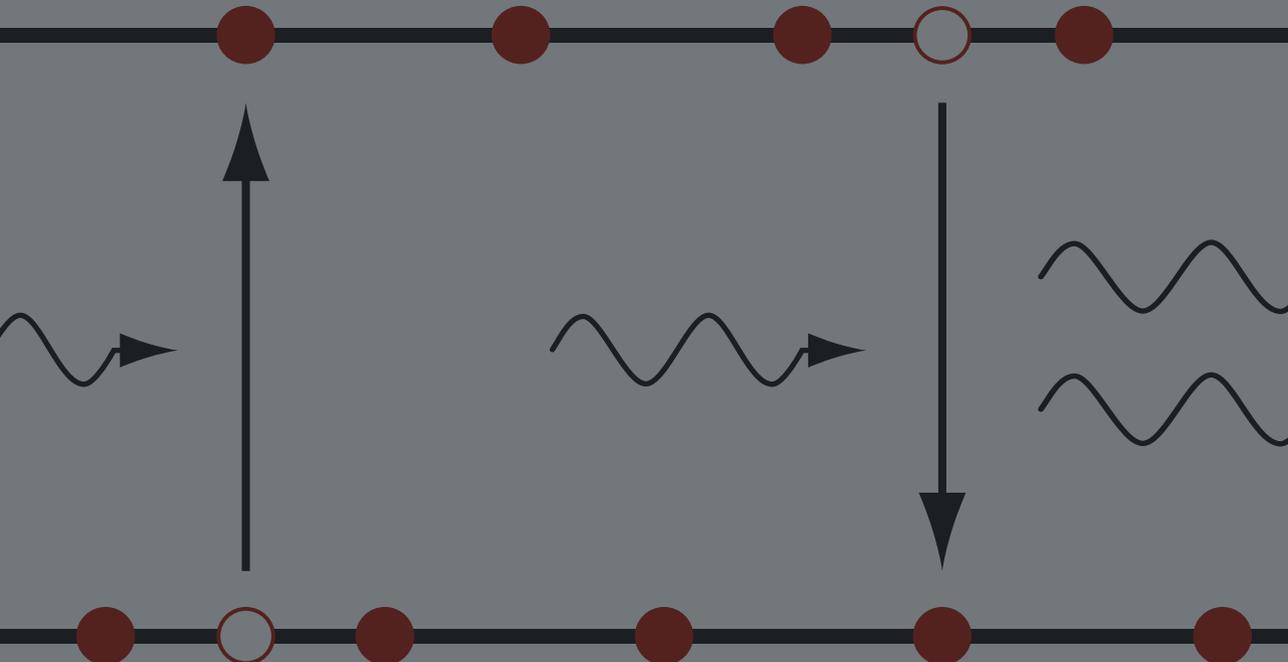




UNIVERSIDAD AUTÓNOMA
DE SAN LUIS POTOSÍ

LÁSERES DE SEMICONDUCTOR

Viatcheslav Andreevich Mishurnyi
Alfonso Lastras Martínez



**LÁSERES DE
SEMICONDUCTOR**

LÁSERES DE SEMICONDUCTOR

Viatcheslav Andreevich Mishurnyi
Alfonso Lastras Martínez

Instituto de Investigación en Comunicación Óptica
Universidad Autónoma de San Luis Potosí

San Luis Potosí, S.L.P., México, 2009

© Derechos Reservados *por*
Viatcheslav Andreevich Mishurnyi
Alfonso Lastras Martínez
Universidad Autónoma de San Luis Potosí

ISBN 978-607-7856-08-5

Editorial Universitaria Potosina

ÍNDICE

Prólogo

I. Principios de funcionamiento

1.1	Introducción	11
1.2	Interacción de la radiación electromagnética con la materia	12
1.3	El sólido cristalino	17
	A. Bandas de energía y brechas prohibidas	17
	B. Interacción del sólido cristalino con el campo eléctrico	18
1.4	Tipos de conductividad en un semiconductor	19
1.5	Huecos en un semiconductor	21
1.6	Generación de radiación en un semiconductor	22
	A. Radiación por recombinación de pares electrón-hueco	22
	B. Unión p-n	23
1.7	Guías de ondas ópticas	25
1.8	Láseres de semiconductor basados en uniones p-n	27
1.9	Conclusiones	29

2. Láseres basados en heterouniones

2.1	Introducción	31
2.2	La heterounión y sus propiedades	32
	A. Modelos de bandas y particularidades del paso de la corriente eléctrica	32
	B. Efecto ventana	34
2.3	La heteroestructura doble	40

2.4	La heteroestructura con confinamientos electrónico y óptico separados	44
2.5	Conclusiones	46
3. Materiales para la construcción de láseres de semiconductor		
3.1	Introducción	47
3.2	La estructura de bandas de los materiales semiconductores	48
3.3	Compuestos binarios III-V	53
3.4	Soluciones sólidas ternarias de $\text{Al}_x\text{Ga}_{1-x}\text{As}$	55
3.5	Soluciones sólidas ternarias GaAsP , GaInP y AlInP	58
3.6	Soluciones sólidas cuaternarias AlGaInP , GaInAsP , GaInAsSb y AlGaAsSb	62
3.7	Conclusiones	69
4. Equilibrios de fase y procesos de cristalización		
4.1	Introducción	71
4.2	Transiciones de fase	72
4.3	Proceso de cristalización por capas	74
4.4	Cristalizaciones volumétrica y epitaxial	77
4.5	Conclusiones	79
5. Tecnologías epitaxiales de crecimiento de cristales		
5.1	Introducción	81
5.2	Epitaxia por fase líquida	82
5.3	Epitaxia por fase gaseosa	86
5.4	Epitaxia por haces moleculares	95
5.5	Conclusiones	101

6. Diferentes diseños de diodos con resonador de Fabry-Perot

6.1	Introducción	103
6.2	Láseres con resonadores de tipo Fabry-Perot	103
	A. Construcción del resonador Fabry-Perot	103
	B. Láseres de contacto ancho	105
	C. Láseres de contacto angosto	106
	D. Características electromagnéticas	108
6.3	Diferentes diseños de láseres de contacto angosto	111
	A. Láser con contacto de stripe	111
	B. Diseño de implantación de iones	113
	C. Láser con stripe de meseta	115
	D. Láseres recrecidos con stripe de meseta	118
	E. Otros diseños	119
6.4	Conclusiones	122

Bibliografía

PRÓLOGO

Este libro está dedicado a los láseres de semiconductor. Debido a su muy alta eficiencia, rapidez, bajo consumo de energía, tamaño diminuto y confiabilidad, así como por su bajo costo de producción en masa, hoy en día estos dispositivos tienen numerosas y variadas aplicaciones. Encontramos a los láseres de semiconductor en campos tan diversos como son, las comunicaciones por fibra óptica, la medición de distancias, la espectroscopía de materiales, la impresión de documentos, la medicina, el control de procesos industriales, los sensores ópticos, la holografía, los sistemas ópticos para el procesamiento de información y las memorias ópticas, por mencionar sólo algunos ejemplos. En consonancia con su amplia gama de aplicaciones, el mercado de los láseres de semiconductor se ha expandido constantemente durante los últimos 20 años.

Los principales fenómenos y efectos físicos que toman lugar dentro de un láser durante su operación son complejos y su descripción rigurosa requiere de un gran aparato matemático. Debido a que el objetivo de los autores del presente libro es discutir el tema de los láseres de semiconductor utilizando un lenguaje accesible y simple, que no involucre cálculos matemáticos complicados, algunos temas son explicados de una forma sencilla aunque en algunas ocasiones no con el suficiente rigor. En estos casos se privilegió la simplicidad en la exposición por sobre la formalidad y la precisión, y los autores esperan condescendencia de parte de los lectores con más conocimiento del tema.

Hay seis capítulos en el libro en los que se consideran: los principios de la operación de los láseres, los láseres de homounión y de heterounión, los materiales utilizados en la producción de heteroláseres,

algunas peculiaridades del equilibrio de fase y las técnicas modernas de crecimiento usadas para la fabricación de láseres, finalizando con una discusión de algunas estructuras empleadas en los láseres modernos. En relación a este último tema, se discuten en particular los láseres con la llamada estructura de *stripe*.

Dada su complejidad y de acuerdo con el enfoque adoptado para el libro, no se incluye una discusión sobre los efectos de cuantización en estructuras de baja dimensionalidad -pozos cuánticos y puntos cuánticos- que juegan un papel esencial en los láseres de semiconductor modernos. No se considera tampoco el empleo de espejos de Bragg en la cavidad óptica de un láser para la operación monomodo longitudinal, ni se incluye una discusión sobre las aplicaciones de los láseres de semiconductor. Esperamos, sin embargo, llenar estos huecos en un futuro próximo en un siguiente libro.

En la actualidad existen muchos y muy populares libros acerca de láseres. Sin embargo, la mayoría de estos libros están publicados en inglés, por lo que creemos que la presente obra encontrará lectores interesados de habla hispana. Para el lector que quiera profundizar en el tema de los láseres de semiconductor, al final del libro se incluye una lista bibliográfica no exhaustiva sobre el tema.

Los autores desean expresar su reconocimiento a todas aquellas personas que participaron en la elaboración de este libro. En particular, expresamos nuestro más grande agradecimiento a Liliana del Carmen Mendoza Tello, Iván Hernández del Castillo, Miguel Ángel Lastras Montaña, y Rocío Lastras Montaña. Queremos igualmente reconocer la cuidadosa revisión del texto realizada por Norberto de la Torre González.

CAPÍTULO I

Principios de funcionamiento

I.1 Introducción

La palabra “láser” es un acrónimo de la expresión en inglés *Light Amplification by Stimulated Emission of Radiation*, que puede ser traducida como: *Amplificación de Luz por Emisión Estimulada de Radiación*. Los láseres de semiconductor son dispositivos que convierten directamente la corriente eléctrica en luz coherente -una luz muy especial, como se explica más adelante. Hoy día estos dispositivos encuentran una amplia aplicación en los más diversos dominios de la ciencia y la tecnología. Esto se debe en primer lugar a que los láseres de semiconductor poseen propiedades únicas como son: una alta eficiencia de conversión de energía eléctrica en energía luminosa -hasta de un 50%-, un tamaño reducido, facilidad para modularse a altas frecuencias, larga vida de servicio y un bajo costo.

Una de las áreas más importantes de aplicación de los láseres de semiconductor es la de los sistemas de comunicación por fibra óptica, diseñados para transmitir grandes volúmenes de información a grandes distancias. Los láseres de semiconductor tienen también una amplia aplicación en la impresión y procesamiento de información óptica, en el bombeo -o excitación- de otros láseres y fibras ópticas “activas”, y en sistemas de medición de distancias. Encuentran igualmente aplicaciones en espectroscopía, en la fabricación de diversos dispositivos ópticos como detectores y sistemas de almacenamiento de datos -memorias ópticas-, en el control de procesos tecnológicos y de contaminación ambiental, en medicina y en la industria militar.

Los láseres semiconductores pueden operar en un amplio rango del espectro electromagnético, desde la región visible hasta la infrarroja, con una potencia de emisión radiativa que puede alcanzar algunos vatios en un régimen continuo. En arreglos especiales de láseres, conteniendo algunas decenas de emisores, se han obtenido potencias ópticas de varios cientos de kilovatios en un régimen de pulsos.

Los factores que determinan la aplicación específica de un láser son fundamentalmente los siguientes: la longitud de onda de la radiación coherente generada, la potencia óptica, la rapidez de operación del dispositivo y el ancho del espectro de emisión. Es por esto que en muchos laboratorios alrededor del mundo se lleva a cabo una intensa investigación en las áreas de aplicación antes mencionadas, mejorándose continuamente los parámetros de operación de los dispositivos existentes, al mismo tiempo que se desarrollan nuevos diseños de láseres de semiconductor.

En lo que sigue se presentan los principios fundamentales de operación de los láseres de semiconductor. De manera específica, en primer lugar se discute la interacción de la radiación con la materia y las características que definen a un material semiconductor; en seguida, se muestra cómo, aprovechando las propiedades de los materiales semiconductores y la forma en que la radiación interactúa con un material, es posible fabricar sistemas generadores de radiación láser; finalmente, se introduce el concepto de guía de onda, elemento fundamental en la fabricación de un láser de semiconductor.

1.2 Interacción de la radiación electromagnética con la materia

Supongamos que tenemos un sistema con N electrones en el que solamente existen dos niveles energéticos E_1 y E_2 , y que además E_2 es mayor que E_1 . Consideremos cómo interacciona este sistema de dos niveles con la radiación electromagnética. En lo que sigue, se empleará el nombre genérico de luz para designar a la radiación electromagnética,

independientemente que ésta sea visible o no. Cuando se ilumina el sistema con luz de energía $\Delta E = E_2 - E_1$ pueden ocurrir tres procesos:

- 1) La energía del fotón es absorbida por el electrón, llevándolo del nivel energético E_1 al nivel E_2 .
- 2) El electrón excitado libera espontáneamente y al azar la energía absorbida emitiendo un fotón, mediante un proceso que se conoce como emisión espontánea.
- 3) Un fotón propagándose en el medio conteniendo un electrón excitado puede interactuar con éste, obligándole a emitir un nuevo fotón. Como resultado de este proceso tenemos dos fotones en lugar de uno, con la misma dirección de propagación e idénticas características de fase y polarización. A la radiación con estas características se le denomina coherente y al proceso de emisión se le conoce como emisión estimulada.

Un sistema alcanza su estado de equilibrio térmico minimizando su energía. Así, en el estado de equilibrio la distribución de electrones en los niveles energéticos puede representarse como se ilustra en la figura 1.1a, en donde todos los electrones se encuentran en E_1 . A esta distribución de electrones se le denomina población normal o de equilibrio. Cuando se ilumina un medio con población normal es muy probable que la luz incidente sea absorbida (figura 1.1a), lo que significa que su intensidad disminuye conforme se propaga en el medio.

Para discutir el caso fuera de equilibrio, supóngase que, a través de una excitación externa a la que se le llama bombeo -una corriente eléctrica, iluminación con una lámpara de gran intensidad, etc.-, la mayoría de los electrones se llevan al nivel superior. A la distribución de electrones en este caso se le conoce como población invertida (figura 1.1b). En un medio con población invertida se absorben menos fotones y en contraste, a través del mecanismo de emisión estimulada (figura 1.1b), se incrementa el número de fotones idénticos, es decir, se aumenta la intensidad de la radiación coherente en el medio.

En contraste, la radiación espontánea se genera como consecuencia de transiciones electrónicas aleatorias desde el estado

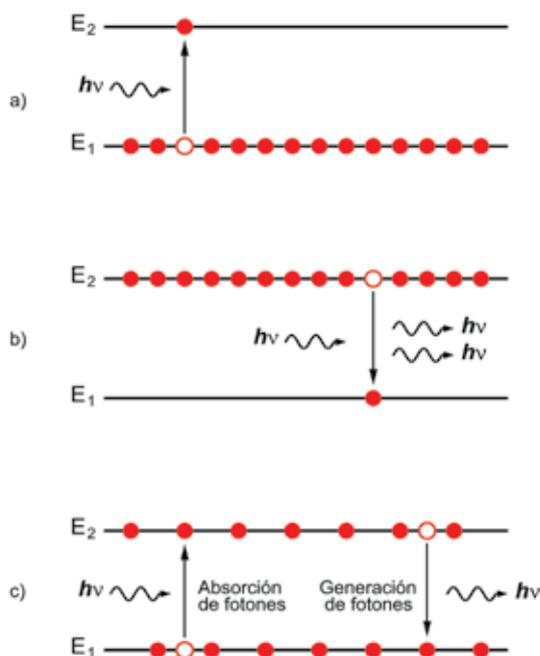


Fig.1.1 Modelo simplificado de un sistema de dos niveles energéticos E_1 y E_2 , con $E_2 > E_1$, y las posibles transiciones entre los mismos. a) Población de niveles en equilibrio; en estas condiciones la luz ocasiona mayoritariamente transiciones desde E_1 hasta E_2 y, en consecuencia, su intensidad se atenúa a medida que se propaga en el medio. b) Población invertida, bajo la cual la luz interacciona con el medio generando fundamentalmente fotones estimulados; en estas condiciones la intensidad de la luz se incrementa a medida que viaja en el medio. c) Condición intermedia, bajo la cual coexisten la absorción, la emisión espontánea y la emisión estimulada.

excitado E_2 hacia el estado con menor energía E_1 (figura 1.1c). Por esta razón la radiación espontánea no es coherente. Este es el tipo de radiación producida por fuentes de luz tales como las lámparas incandescentes, las lámparas de arco y los diodos emisores de luz.

Si ahora se coloca el medio activo -sistema hipotético de dos niveles, con población invertida-, entre dos espejos paralelos como se muestra en la figura 1.2, la radiación que se propaga en dirección perpendicular a los espejos se refleja en estos, y al mismo tiempo se amplifica cada vez que pasa por la región activa. Para que la radiación salga del resonador, el cual está formado por el medio activo y los espejos, se necesita que uno de los espejos sea parcialmente transparente.

Los mecanismos de generación de radiación láser discutidos líneas arriba pueden clarificarse a través del modelo mecánico ilustrado en la figura 1.3. En dicha figura se representa un sistema compuesto por palitas fijadas a un pivote, las cuales pueden girar alrededor del punto O. Las palitas están situadas de forma tal que apenas rozan la superficie del agua. En un inicio (figura 1.3a), las esferas están en el fondo del recipiente con agua. Este estado corresponde a una población en equilibrio termodinámico, similar al ilustrado en la figura 1.1a para el sistema de dos niveles. Inyectando energía al sistema, lo que equivale al proceso de bombeo, se llevan las esferas desde el fondo del recipiente hasta las palitas (figura 1.3b). Este estado con población invertida es equivalente al mostrado en la figura 1.1b. Ahora supóngase que, de manera espontánea, una de las esferas cae al fondo del recipiente. Como resultado de esto, en la superficie del agua aparece una ola que



Fig. 1.2 Diagrama simplificado de un sistema de luz láser, también llamado resonador óptico.

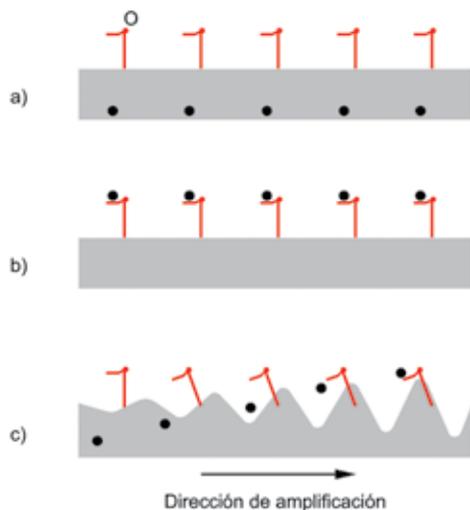


Fig.1.3 Analogía mecánica del sistema electrónico de dos niveles ilustrado en la figura 1.1. a) Estado base con las esferas negras en el fondo del recipiente. b) Estado excitado con las esferas sobre los postes. c) Generación de radiación estimulada.

a medida que se propaga a lo largo del recipiente ocasiona la inclinación de los postes e induce la caída de la siguiente esfera, amplificando la amplitud de la ola como se ilustra en la figura 1.3c.

En este modelo mecánico, la distancia entre los postes y la altura a la que se llevan las bolas se calculan de manera que la caída de cada esfera contribuya a amplificar la ola que se propaga a lo largo del recipiente. Similarmente, la interacción de la radiación con el medio activo induce el decaimiento de los electrones excitados, generando fotones que aumentan -amplifican- la intensidad de dicha radiación.

En esta sección se introdujeron los conceptos fundamentales relativos a la interacción de la radiación con la materia y en particular se discutieron los mecanismos de generación de luz láser. Hacemos notar que dichos conceptos se aplican a todos los sistemas láser, llámense de gas, de semiconductor, de pigmento, de estado sólido, etc.

I.3 El sólido cristalino

A. Bandas de energía y brechas prohibidas

En contraste con los electrones de un átomo libre, en un sólido cristalino y debido a la cercanía entre los átomos que lo componen, los electrones no ocupan niveles discretos sino bandas de energía. En efecto, la interacción de los electrones de un determinado átomo con los núcleos atómicos y con los electrones de los átomos vecinos que ocurre en un sólido, obliga a una combinación de los niveles de energía discretos de los átomos individuales, produciendo lo que se denomina bandas de energía permitida. Estas bandas corresponden a los posibles niveles de energía de los electrones dentro del cristal y están formadas por niveles energéticos situados muy cerca uno del otro. Las bandas de energía permitida están separadas unas de otras por regiones denominadas brechas prohibidas de energía.

Así pues, en un sólido cristalino tenemos que todos los electrones están localizados dentro de las bandas de energía permitida, separadas unas de otras por brechas de energía prohibida.

Los electrones se acomodan en el sólido cristalino llenando las bandas permitidas desde el fondo hacia arriba; es decir, llenando primeiramente la banda con menor energía y ocupando de modo secuencial las siguientes bandas hasta agotar los electrones. Al final de este proceso, la última banda con electrones puede estar en cualquiera de dos condiciones: o bien completamente llena, o bien parcialmente llena, como se ilustra en la figura 1.4. Estos casos corresponden, de manera respectiva, a los aislantes y a los metales. En el caso de un aislante, a la última banda completamente llena se le llama banda de valencia, mientras que la banda permitida inmediatamente arriba de la banda de valencia es la banda de conducción (ver figura 1.4).

De manera estricta, en realidad, la ausencia de electrones en la banda de conducción de un aislante solamente ocurre a una temperatura de cero grados Kelvin. A una temperatura finita, algunos electrones en la banda de valencia tendrán suficiente energía para “saltar” la brecha

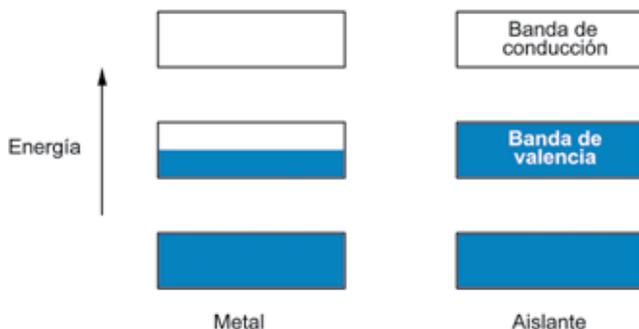


Fig.1.4 Bandas de energía permitidas para un metal y para un aislante. El color azul indica los niveles de energía ocupados por electrones

prohibida y alcanzar la banda de conducción. En un aislante, sin embargo, la brecha prohibida es muy grande, de modo que solamente un número despreciable de electrones son capaces de hacer esto, aun a temperatura ambiente.

B. Interacción del sólido cristalino con el campo eléctrico

Un campo eléctrico externo influye sobre cada uno de los electrones del sólido con una fuerza determinada. Esta fuerza, destruye la simetría en la distribución de velocidades de los electrones, desacelerando a los que se mueven en la dirección del campo eléctrico y acelerando a los que se mueven en contra del mismo. Dichas aceleraciones están relacionadas con la variación de la energía del electrón; esto es, tienen lugar transiciones de los electrones a otros estados permitidos, tanto incrementando como disminuyendo su energía.

Las transiciones podrán, sin embargo, realizarse solamente cuando existan lugares desocupados en la banda de energía en la que se encuentran los electrones. Si dicha banda está completamente llena, no es posible que los electrones se muevan dentro de la misma. Si,

además, la siguiente banda con lugares desocupados está separada de la primera por una brecha energética suficientemente amplia, el campo eléctrico externo no producirá movimiento de electrones. Bajo estas circunstancias, al aplicar un campo eléctrico externo no habrá corriente eléctrica en el cristal.

Los materiales aislantes muestran precisamente esta característica. En los aislantes la brecha prohibida es tan grande que se requeriría de campos eléctricos muy intensos para inducir una corriente eléctrica en ellos. Los aislantes tienen entonces una conductividad eléctrica extremadamente baja. En contraste, los metales tienen una banda parcialmente llena de electrones y es muy fácil generar una corriente eléctrica a través de ellos. Los metales, en consecuencia, están caracterizados por una elevada conductividad eléctrica.

Al igual que los aislantes, en un semiconductor la última banda de energía con electrones está completamente llena -estrictamente, a cero grados Kelvin- y su conductividad eléctrica a bajas temperaturas es muy pequeña. En un semiconductor, sin embargo, la brecha energética entre las bandas de valencia y de conducción no es muy grande, de modo que a una temperatura relativamente alta habrá un número no despreciable de electrones que “saltan” desde la banda de valencia a la de conducción. Cuando esto sucede, en la banda de valencia quedará un agujero o hueco por cada electrón que es transferido a la banda de conducción. Además, con el aumento de temperatura, el número de electrones excitados y de huecos generados se incrementa muy rápidamente, de manera simultánea con la conductividad del cristal.

I.4 Tipos de conductividad en un semiconductor

La conductividad de los semiconductores es muy sensible a la presencia de pequeñas cantidades de impurezas. Para entender esto, tomemos por ejemplo al silicio (Si), el cual se localiza en la cuarta columna de la tabla periódica de los elementos y por tanto tiene cuatro electrones en su última capa electrónica.

Es posible sustituir una parte de los átomos de silicio con átomos de arsénico (As) que pertenece al grupo V de la tabla periódica y que por tanto tiene cinco electrones en su última capa. Para establecer un enlace con sus vecinos más cercanos, el átomo de arsénico necesita sólo cuatro de sus electrones de valencia. El quinto electrón no participa en la formación de enlaces. Resulta además que, como consecuencia de la interacción con el resto de la red cristalina, la fuerza de atracción entre este electrón y su núcleo atómico disminuye notablemente. Bajo estas circunstancias, resulta que la energía térmica, aun a bajas temperaturas, es suficiente para que el electrón se desprenda del átomo de arsénico y pueda de esta forma desplazarse libremente por la red cristalina, convirtiéndose así en un electrón que participa en la conducción. A los átomos que proporcionan electrones extra de conducción al cristal semiconductor, y que permiten cambiar su conductividad por varios órdenes de magnitud, se les llama donadores.

Así pues, para el silicio, o para cualquier otro semiconductor del cuarto grupo de la tabla periódica, se pueden utilizar los elementos del quinto grupo de dicha tabla como átomos donadores. En un cristal de este tipo, los principales portadores de la carga eléctrica son los electrones. Es por esto que a la conductividad de un semiconductor dopado con donadores se le denomina conductividad negativa o tipo n. Consecuentemente, el semiconductor dopado con donadores se denomina semiconductor tipo n.

Ahora, supóngase que parte de los átomos del cristal de silicio es sustituida por átomos de indio (In), perteneciente al grupo III de la tabla periódica y por tanto con tres electrones en su capa de valencia. Para la formación de enlaces con los cuatro átomos de silicio más cercanos, al átomo de indio le falta un electrón, que toma “prestado”, con un gasto mínimo de energía, del átomo vecino de silicio. Este enlace roto, o hueco, no permanece estático sino que se desplaza por la red cristalina del silicio como si fuera una carga eléctrica positiva. A este tipo de impurezas, que inducen en el semiconductor una conductividad positiva, se les llama aceptores. Al semiconductor dopado con aceptores se le denomina semiconductor tipo p.

1.5 Huecos en un semiconductor

Es posible extender fácilmente las ideas tratadas en la sección 1.2 relativas a un sistema de dos niveles a un material semiconductor. En este caso, tal como se muestra en la figura 1.5, el nivel E_1 corresponde a la banda de valencia, mientras que el nivel E_2 de más alta energía corresponde a la banda de conducción.

Supóngase que inicialmente la banda de valencia está completamente ocupada por electrones. Si se estimula el cristal con radiación de energía $h\nu$ mayor que la brecha energética, uno de los electrones en la banda de valencia podrá ser excitado a un nivel energético de la banda de conducción mediante la absorción de un cuanto -una porción- de energía. Debido a que la banda de conducción está casi vacía, al aplicar un campo eléctrico externo el electrón excitado se desplazará a lo

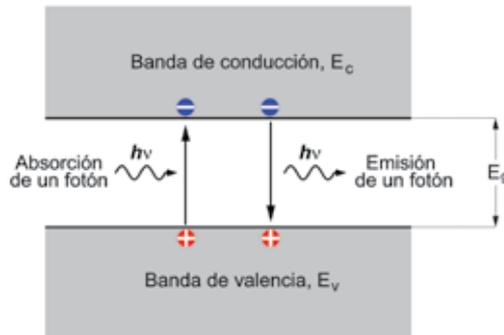


Fig.1.5 Representación esquemática de un material semiconductor visto como un sistema de dos niveles E_c y E_v , separados energéticamente por la brecha prohibida E_g . La absorción de un cuanto de radiación (fotón) con una energía mayor que la brecha prohibida genera un par electrón-hueco. Mediante el proceso inverso de recombinación de pares existe la posibilidad de que se genere un cuanto de radiación con una energía cercana a la de la brecha prohibida.

largo del cristal aumentando su energía, es decir, pasando de un nivel energético a otro más alto. Por otro lado, en la banda de valencia, que antes de la excitación se encontraba llena, el lugar vacío dejado por el electrón excitado podrá ser usado por los demás electrones para moverse en respuesta al campo eléctrico aplicado y contribuir de esta manera a la corriente eléctrica. Es posible demostrar que el lugar vacío se comporta como una partícula en la banda de valencia con carga positiva. A esta partícula ficticia se le denomina hueco.

1.6 Generación de radiación en un semiconductor

A. Radiación por recombinación de pares electrón-hueco

Al momento de absorberse un fotón con una energía mayor que E_g , se genera en el semiconductor un par electrón-hueco (ver figura 1.5), el cual, después de un cierto tiempo característico, decae mediante un proceso que recibe el nombre de aniquilación o recombinación de pares electrón-hueco. Como resultado de esta recombinación, se puede generar radiación con una energía que corresponde a la diferencia entre los niveles energéticos que ocupaban el electrón y el hueco antes de recombinarse. A la radiación emitida por este proceso se le llama radiación por recombinación. Si la radiación por recombinación se produce como consecuencia de una recombinación espontánea en un medio con población invertida, podrá amplificarse a medida que se propaga en el medio debido a transiciones inducidas.

Para obtener la inversión de la población en un semiconductor es necesario, de manera análoga al sistema de dos niveles estudiado anteriormente, crear una concentración elevada de electrones en la banda de conducción y de huecos en la de valencia. En la siguiente sección se muestra cómo es posible obtener la condición de población invertida en un semiconductor. La manera más simple y efectiva es mediante la inyección de una corriente eléctrica a través de lo que se denomina una unión p-n.

B. Unión p-n

Una unión p-n consiste de la unión íntima entre un semiconductor tipo n y otro tipo p. Al poner en contacto dos semiconductores con estas características, se produce una difusión de electrones desde el semiconductor n al p y de huecos en sentido contrario. Como resultado, ambos semiconductores pierden carga eléctrica formándose una capa volumétrica de cargas positivas -donadores ionizados- en la región tipo n inmediata a la interfaz, así como una correspondiente capa de cargas negativas -aceptores ionizados- en la región tipo p. Las capas de cargas positivas y negativas a ambos lados de la interfaz generan un campo eléctrico de contacto.

El diagrama de bandas de la unión p-n en el estado de equilibrio está representado en la figura 1.6a. Comparando esta figura con la figura 1.5, se observa que el campo eléctrico de contacto establece una diferencia de energía ΔE que actúa como una barrera, tanto para los electrones de la región n, como para los huecos de la región p. La región de campo corresponde en la figura 1.6a a la región sombreada a ambos lados de la interfaz.

Si a la unión p-n se le aplica un voltaje en la dirección directa, el campo eléctrico asociado a este voltaje tendrá una dirección contraria a la del campo eléctrico de contacto. Como consecuencia, se producirá una disminución de la barrera de potencial para los electrones, así como también para los huecos, que será proporcional al voltaje aplicado. Bajo estas condiciones, se induce una inyección de electrones hacia la región tipo p y de huecos hacia la región tipo n, tal como se ilustra en la figura 1.6b.

Los electrones y huecos inyectados se difunden hacia lo profundo del material, más allá de la unión p-n; los electrones se difunden hacia la región p y los huecos hacia la región n. Debido a los procesos de recombinación, sin embargo, la concentración de estos portadores de carga disminuye a medida que se alejan de la unión p-n. Las distribuciones de electrones y de huecos a ambos lados de la misma, están representadas en la figura 1.6b por las regiones sombreadas en color

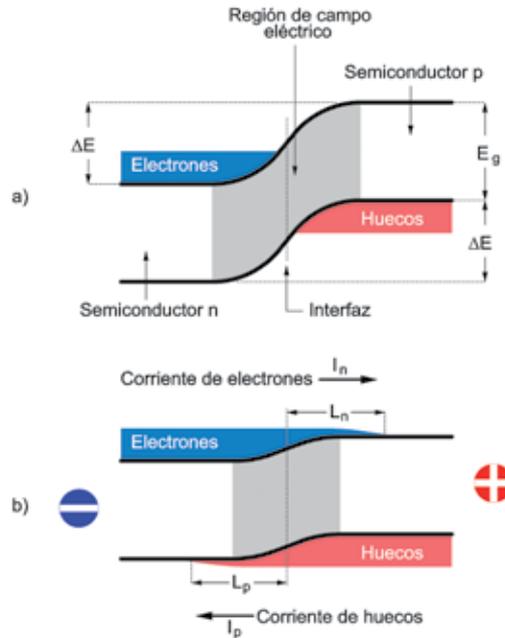


Fig.1.6 La unión p-n. a) en equilibrio. b) con un voltaje de polarización directa. La longitud de difusión L_n es la longitud media que viajan los electrones una vez que son inyectados a la región tipo p. L_p es la longitud de difusión correspondiente a los huecos inyectados hacia la región tipo n. L_n y L_p se miden a partir de la interfaz entre los dos semiconductores por simplicidad (ver texto).

azul y rojo, en forma respectiva. Las longitudes de difusión L_n y L_p indicadas en la figura 1.6b corresponden, respectivamente, a las distancias medias que viajan los electrones y los huecos una vez que son inyectados. Se hace notar que si bien en la figura 1.6b las longitudes de difusión se miden a partir de la interfaz entre los dos semiconductores, de manera estricta L_n y L_p se definen a partir de las fronteras de las regiones de campo eléctrico en los semiconductores p y n, de manera respectiva.

Debido a que la población invertida de portadores de carga tiene lugar sólo cuando existe una elevada concentración de los mismos, la condición de inversión sólo se cumple en la región próxima a la unión p-n. Como consecuencia de esto, los portadores de carga que se difunden hasta distancias lejos de la unión no participan en el proceso de amplificación de la radiación.

1.7 Guías de ondas ópticas

La velocidad de la luz esta determinada por el índice de refracción del medio en que se propaga, el cual se define como el cociente entre la velocidad de propagación de la luz en el vacío y la correspondiente velocidad de propagación en el medio. El índice de refracción también determina el comportamiento de la luz al cruzar una interfaz entre dos medios. En la figura 1.7a se muestra que al viajar la luz desde un material con un índice refracción n_1 , a otro material con un índice menor n_2 , una parte de la misma se refleja y la otra se transmite cambiando su dirección de propagación. El cambio de dirección de la luz transmitida constituye el fenómeno de refracción y la magnitud de dicho cambio depende de la diferencia entre los dos índices de refracción.

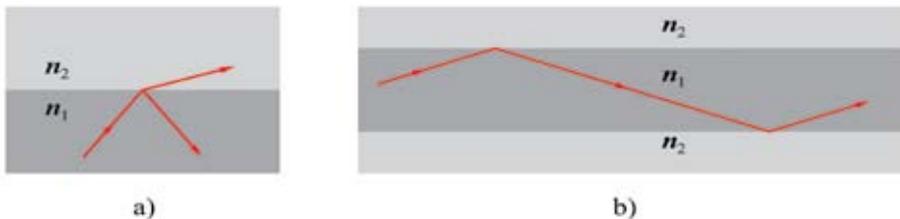


Fig. 1.7 a) Luz incidiendo sobre la interfaz entre dos medios de índices de refracción n_1 y n_2 tal que $n_1 > n_2$. b) Guía de onda formada emparedando un material de índice de refracción n_1 entre dos capas de material de índice de refracción n_2 , tal que $n_1 > n_2$. La guía de onda se basa en el fenómeno de reflexión interna total.

Si en el caso anterior la luz incide de manera rasante sobre la interfaz, ocurre el fenómeno de reflexión interna total, en el que no hay luz transmitida y toda la luz incidente es reflejada hacia el medio con índice de refracción mayor. La reflexión total ocurre para ángulos de incidencia grandes a partir de un cierto valor mínimo. Entre más grande sea la diferencia en índices de refracción, menor será este valor.

Ahora bien, si se coloca una placa de un material de índice de refracción n_1 entre dos medios con índice de refracción menor n_2 , es posible emplear esta estructura como una guía de ondas haciendo uso del fenómeno de reflexión total, tal como se ilustra en la figura 1.7b. La guía de ondas será tanto más efectiva en cuanto mayor sea la diferencia de índices de refracción.

Al polarizar una unión p-n de manera directa (figura 1.8a) se forma de manera espontánea una guía de onda como se explica a continuación. En efecto, dentro de los límites de la región de amplificación de la unión p-n -la parte sombreada en la figura 1.8b- y como consecuencia de la gran concentración de portadores de carga, se observa un aumento en el valor del índice de refracción (figura 1.8c). Debido a esto, la onda luminosa se localiza parcialmente dentro de esta región. Se hace notar, sin embargo, que la diferencia de índices de refracción entre la parte amplificadora y la parte pasiva es sumamente pequeña, de modo que la onda luminosa penetra a la región de absorción de manera apreciable (figura 1.8d), lo que de forma natural resulta en una pérdida de energía.

1.8 Láseres de semiconductor basados en uniones p-n

El primer láser semiconductor se construyó en el año de 1962 con base en una unión p-n de GaAs. En la figura 1.9 se muestra la dependencia de la potencia óptica de este láser en función de la densidad de corriente eléctrica inyectada a la unión p-n, la cual se denomina característica vatios-amperios del láser. En esta misma gráfica se

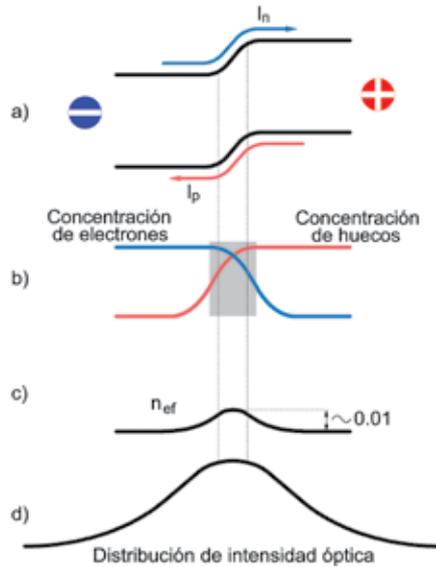


Fig.1.8 a) Unión p-n polarizada de manera directa. b) Concentraciones de electrones y de huecos en las regiones aledañas a la interfaz debidas a la inyección de portadores por la polarización directa. c) Perfil del índice de refracción, el cual muestra un pequeño incremento alrededor de la unión p-n asociado al exceso de electrones y huecos, formando de esta manera una guía de onda. d) Perfil de intensidad de la radiación generada por la unión p-n, la cual tiende a ser confinada por la guía de onda. El confinamiento es, sin embargo, relativamente débil por la pequeña diferencia entre los índices de refracción de las regiones con y sin exceso de portadores.

aprecian dos regiones. En la primera, un aumento en la corriente eléctrica aplicada va acompañado de un crecimiento relativamente pequeño de la potencia óptica -número de fotones extraídos del dispositivo. El espectro de la radiación es bastante ancho y la posición del máximo coincide aproximadamente con el valor del ancho de la brecha prohibida del material. Esta región corresponde a un

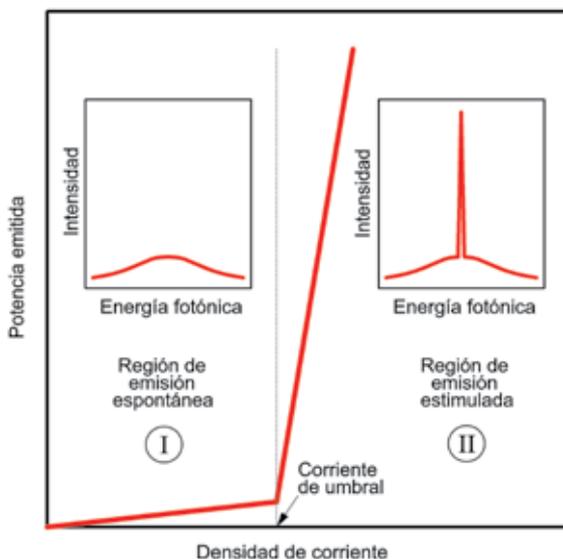


Fig.1.9 Diagrama típico de la curva característica de potencia óptica contra densidad de corriente aplicada en un láser de semiconductor. Se distinguen dos regímenes: la región I de emisión espontánea, por abajo de la corriente de umbral y la región 2 de emisión estimulada, por arriba de la misma.

régimen de emisión espontánea, cuando la inversión de población aún no se ha alcanzado.

A medida que aumenta la corriente eléctrica inyectada a la unión p-n, llega el momento en que se alcanza la inversión de población y el dispositivo entra a un régimen en el que domina la radiación coherente. El ancho del espectro de la emisión se reduce bruscamente (ver figura 1.9) y la potencia emitida crece rápidamente con la corriente inyectada. La transición entre los regímenes de emisión espontánea y emisión estimulada puede entonces ser identificado por un cambio en la pendiente de la curva característica vatios-amperios.

A partir de esta curva característica se obtiene uno de los parámetros fundamentales del láser de semiconductor: la corriente de umbral, la cual se define como la corriente obtenida de la intersección de la extrapolación de la rama de corriente estimulada con el eje horizontal. A fin de posibilitar una comparación entre láseres con diferentes áreas activas, por lo general no se especifican valores absolutos de corriente de umbral, sino densidades de corriente en A/cm^2 .

En los primeros láseres de semiconductor basados en la unión p-n, la densidad típica de corriente de umbral era de 20-30 kA/cm^2 . A tan elevados valores de corriente de umbral, este tipo de láseres podían funcionar solamente a temperaturas muy bajas de alrededor de 77 K, o bien en un régimen de pulsos extremadamente cortos. Esta limitación redujo considerablemente el área de aplicación de los láseres de homounión.

1.9 Conclusiones

En este capítulo se han discutido los principios de funcionamiento de los láseres de semiconductor tomando como ejemplo el láser basado en la unión p-n. Se mostró que para generar luz coherente es necesario inducir un estado de inversión de población en la región activa del dispositivo, la cual se localiza en las inmediaciones de la unión p-n. Debido a las altas densidades de portadores de carga negativa -electrones- y positiva -huecos- en la región activa, el índice de refracción de la misma es mayor que el índice de las regiones adyacentes. En estas condiciones, se produce un confinamiento de la luz generada en una región angosta alrededor de la región activa. La generación de la radiación estimulada y el confinamiento de la misma son dos ingredientes esenciales para la fabricación de un láser de semiconductor. El diseño de láseres de semiconductor estudiado en este capítulo, sin embargo, dista mucho de ser el más eficiente, como se discute en el siguiente capítulo.

CAPÍTULO 2

Láseres basados en heterouniones

2.1 Introducción

En el capítulo anterior se analizaron en detalle los láseres de semiconductor basados en la unión p-n. Estos dispositivos, aunque con una estructura relativamente simple, tienen la desventaja de requerir de altas densidades de corriente eléctrica para obtener radiación coherente. Estas densidades de corriente, como se mencionó en el capítulo anterior, son del orden de decenas de kiloamperios por centímetro cuadrado, lo que limita las posibilidades de aplicación de los láseres de unión p-n, pues sólo permite su operación a bajas temperaturas o bien en un régimen de corriente eléctrica pulsada.

En el año de 1963, en el Instituto Físico-Técnico A.F. Ioffe en Rusia, fue propuesto por primera vez por los físicos Zh. I. Alferov y R.F. Kazarinov un nuevo láser de semiconductor para disminuir considerablemente la densidad de corriente umbral de los láseres de unión p-n. Dicho láser, realizado por Zh. I. Alferov y sus colaboradores en 1967, estaba basado en la así llamada heterounión semiconductor, que se forma al poner en contacto dos materiales semiconductores de diferente composición química.

En este capítulo se introducirán los láseres de heterounión. Se discutirán sus características principales y sus ventajas sobre los láseres de homounión, y cómo estas características han hecho posible la reducción de corriente de umbral, así como la operación del láser en un régimen de corriente continua y a temperatura ambiente.

2.2 La heterounión y sus propiedades

A. Modelos de bandas y particularidades del paso de la corriente eléctrica

En una unión p-n las brechas energéticas de los semiconductores p y n que la forman son iguales, de tal manera que, como se puede apreciar en la figura 2.1, las barreras energéticas ΔE_c y ΔE_v para el flujo de electrones hacia el lado p de la unión y de huecos hacia el lado n, son iguales. En contraste, en una heterounión las alturas de dichas barreras energéticas son en general diferentes, puesto que las brechas de energía prohibida de ambos semiconductores son también en general diferentes.

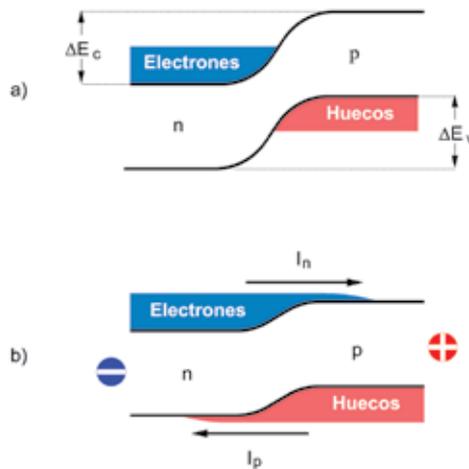


Fig. 2.1 Diagrama de bandas de una unión p-n. a) sin voltaje aplicado, b) con polarización directa. E_g es la brecha energética. ΔE_c y ΔE_v son las barreras de energía que experimentan los electrones y los huecos al cruzar la unión hacia la regiones p y n, respectivamente. En una unión p-n ambas barreras de energía son iguales.

En las figuras 2.2a y 2.2b se muestran los diagramas de bandas de dos heterouniones en equilibrio. Hay que hacer notar que dichos diagramas se presentan en forma simplificada para ayudar a la discusión; de manera específica, como puede observarse en las figuras 2.2a y 2.2b, solamente una de las dos bandas -conducción o valencia- presenta una discontinuidad en la interfaz entre los dos semiconductores. En un caso general, sin embargo, ambas bandas son discontinuas.

En las figuras 2.2a y 2.2b se indica con las letras mayúsculas N o P, según sea el caso, el tipo de conductividad del semiconductor con brecha energética mayor, mientras que las letras minúsculas

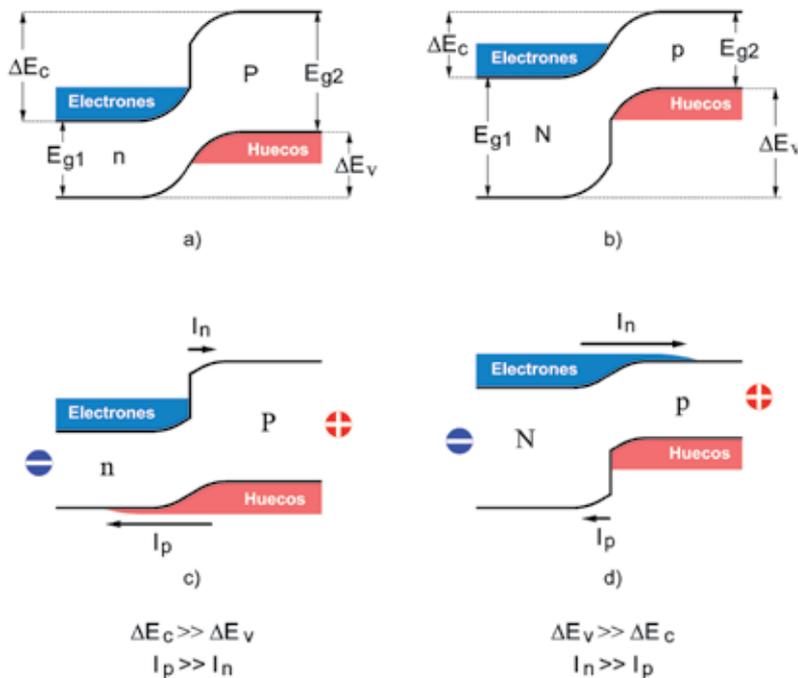


Fig 2.2 Diagrama de bandas de heterouniones n-P y N-p. a), b): sin voltaje aplicado. c), d): con polarización directa.

n y p se refieren al tipo de conductividad del semiconductor con brecha energética menor. En las figuras 2.2c y 2.2d se muestra el efecto que resulta al aplicar un voltaje con polarización directa a las heterouniones de las figuras 2.2a y 2.2b, en forma respectiva. Podemos ver que para la pareja n-P mostrada en la figura 2.2c, la altura de la barrera energética ΔE_c para electrones cruzando la heterounión es mucho mayor que la correspondiente barrera ΔE_v para los huecos. Consecuentemente, con polarización directa la componente de corriente de huecos I_p resulta mucho mayor que la componente de corriente de electrones I_n . Por el contrario, en el caso de la heterounión N-p de la figura 2.2d, la barrera para el flujo de huecos es mucho mayor que la barrera para el flujo de electrones, de tal manera que en este caso I_n supera con mucho a I_p . Para los semiconductores utilizados en las heteroestructuras, la diferencia entre las barreras energéticas debe ser mayor que 0.1 eV.

De lo anterior podemos deducir que, debido a estas diferencias de barreras energéticas para electrones y huecos, al aplicar un campo eléctrico a una heterounión la inyección de portadores desde el semiconductor de brecha energética grande al de brecha energética pequeña constituirá casi por sí misma la corriente total, al ser despreciable la inyección de portadores desde el semiconductor de brecha energética pequeña hacia el semiconductor de brecha energética grande. Esta es una de las propiedades más importantes de las heterouniones, a la cual se le llama efecto de inyección unilateral. Además, las heterouniones son interesantes por una propiedad adicional, relacionada con el llamado efecto ventana.

B. Efecto ventana

Para la comprensión de este efecto, consideremos primeramente los procesos de absorción y generación de radiación en un semiconductor. Supongamos un semiconductor con una brecha de energía E_g , sobre el cual incide radiación con tres diferentes longitudes de onda λ , o

equivalentemente con tres energías fotónicas $h\nu$ -recordemos que la energía fotónica es inversamente proporcional a la longitud de onda. Supongamos con mayor precisión que inciden tres cuantos de radiación -fotones- de energías $h\nu_1$, $h\nu_2$ y $h\nu_3$, tal que $h\nu_1 = E_g$, $h\nu_2 > E_g$ y $h\nu_3 < E_g$, como se muestra esquemáticamente en la figura 2.3. Como es sabido, a temperaturas relativamente bajas la banda de valencia del semiconductor se encuentra llena de electrones mientras que su banda de conducción se encuentra vacía. Recordemos, así mismo, que la banda de valencia y la banda de conducción están formadas por niveles energéticos permitidos situados muy cerca uno de otro, mientras que dentro de la brecha energética del semiconductor E_g dichos niveles no existen.

El proceso de absorción en este caso representa la interacción de los fotones con los electrones en la banda de valencia. Como resultado de esta interacción, la energía del fotón se transmite a uno de estos electrones, el cual es excitado a un nivel energético más elevado. Esto, sin embargo, solamente ocurre bajo ciertas condiciones. En

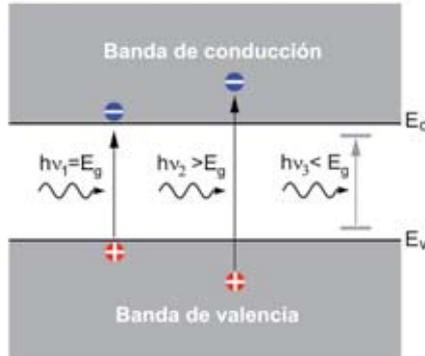


Fig 2.3 Diagrama esquemático de un semiconductor sobre el cual inciden tres fotones de energías $h\nu_1 = E_g$, $h\nu_2 > E_g$ y $h\nu_3 < E_g$. La transición dentro de la brecha prohibida mostrada en color gris es virtual. El diagrama ilustra el efecto ventana.

efecto, si la energía del fotón que interactúa con el semiconductor es igual o mayor que su brecha energética E_g , dicho fotón podrá excitar electrones desde la banda de valencia a la banda de conducción como se muestra en la figura 2.3. Si por el contrario, como se ilustra en la misma figura, el fotón tiene una energía menor que E_g , será incapaz de llevar a cabo la excitación electrónica -notamos que la transición mostrada en la figura 2.3 con color gris es virtual pues no existen niveles energéticos en la brecha prohibida. Como resultado de lo anterior, un semiconductor es transparente para fotones con energía menor que la brecha energética, mientras que resulta opaco para energías iguales o mayores que E_g .

Debemos señalar, no obstante lo anteriormente dicho, que un semiconductor no es perfectamente transparente para $h\nu < E_g$, pues hay procesos de absorción de cuantos de radiación adicionales a los que dan lugar a transiciones de electrones entre las bandas de valencia y de conducción. Uno de estos procesos es la absorción por electrones libres, en el cual la energía del cuanto de radiación es absorbida por un electrón con energía E_c en el fondo de la banda de conducción, excitándolo a un nivel energético más elevado dentro de la misma banda y convirtiéndolo en un electrón “caliente”. Los electrones calientes, así llamados porque tienen una energía mayor que la que corresponde al equilibrio térmico, después de un tiempo muy corto ceden su exceso de energía al cristal y regresan a su energía de equilibrio.

Los procesos de absorción por portadores libres son obviamente más importantes entre más grande sea la concentración de impurezas. En el caso de un láser de heteroestructura, como se discute más adelante, esta absorción es importante en las capas emisoras cuando éstas tienen concentraciones de impurezas mayores a $10^{18}/\text{cm}^3$.

El proceso de absorción de fotones es un proceso con carácter probabilístico. Además, a medida que se aumenta la energía de los fotones por arriba de la brecha energética, la probabilidad de que sean absorbidos aumenta. Esto se manifiesta en un incremento en el coeficiente de absorción, tal como es mostrado esquemáticamente en la figura 2.4 para el caso de GaAs.

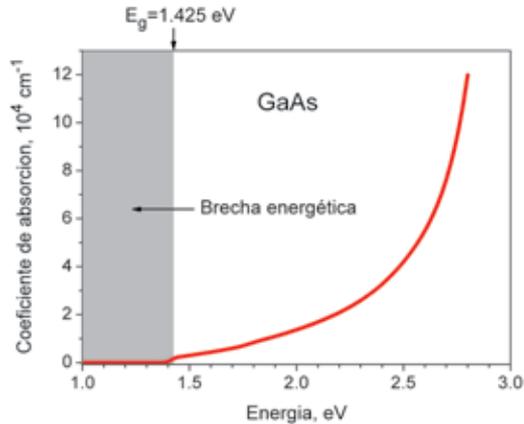


Fig 2.4 El coeficiente de absorción α del GaAs a temperatura ambiente en función de la energía del fotón. Para energías por abajo de la brecha energética E_g el coeficiente de absorción disminuye considerablemente.

A propósito, hacemos notar que el color de los cristales está condicionado precisamente por los procesos de absorción de la radiación. La región del espectro de radiación sensible al ojo humano se extiende en la escala energética de 1.7 eV hasta 3 eV, como se puede apreciar en la figura 2.5, en la que se muestra de manera esquemática la curva de sensibilidad relativa del ojo en función de la energía del fotón $h\nu$ o de su longitud de onda λ . Como se deduce de dicha figura, el ojo humano es más sensible a la radiación con energía fotónica de 2.24 eV, que corresponde a la región verde del espectro. En el intervalo de 1.7 a 2.0 eV se encuentra la región roja de dicho espectro, mientras que a la radiación ultravioleta le corresponden energías mayores que 3.0 eV.

Consideremos un semiconductor con una brecha energética de 1.9 eV. De acuerdo a lo dicho líneas arriba, los fotones con energías $h\nu > E_g$, es decir, los que corresponden a los colores anaranjado, amarillo y verde, serán absorbidos por el semiconductor. Por el contrario, los cuantos con energías $h\nu < E_g$, que corresponden a la región roja del espectro, pasarán por el semiconductor sin ser apreciablemente absor-

bidos. Como consecuencia de esto, el cristal tendrá un color rojo. En el caso de que la brecha energética del semiconductor sea mayor que 1.9 eV, el color del cristal puede variar del amarillo hasta el violeta, dependiendo del valor de E_g . A manera de ejemplo, los cristales de GaP, con una brecha energética de 2.26 eV, tienen un color amarillo.

Cuando se excita un electrón desde la banda de valencia a la banda de conducción por la absorción de un fotón, la banda de valencia queda descompensada con una carga positiva, a la cual le hemos llamado hueco o agujero. De esta manera, podemos decir que la absorción de un cuanto de radiación resulta en la generación de un par electrón-hueco. El proceso inverso al de generación de pares es el de recombinación de un par electrón-hueco. Este último puede ser acompañado por la emisión de un fotón con una energía que aproximadamente corresponde a la brecha energética del semiconductor.

En ciertos semiconductores el proceso de emisión de radiación puede ser muy eficiente. Las condiciones que debe cumplir un semiconductor para que esto suceda serán analizadas en un capítulo posterior dedicado, a los materiales que se utilizan en la fabricación de diferentes láseres de semiconductor.

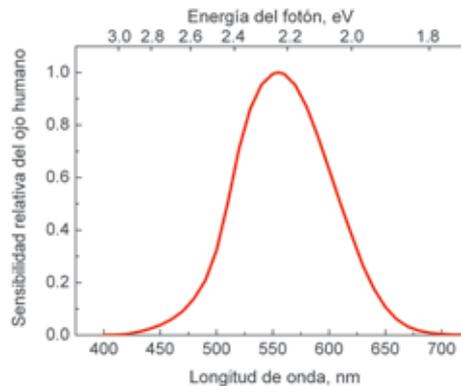


Fig 2.5 Sensibilidad relativa del ojo humano en función de la energía del fotón $h\nu$ y la longitud de onda λ correspondiente.

Para que se lleve a cabo el proceso de recombinación de pares deben existir entonces pares electrón-hueco generados previamente por algún mecanismo. La forma más sencilla y natural para obtener electrones y huecos de manera simultánea es por medio de la inyección de cargas en una unión p-n.

Al inyectar electrones y huecos a ambos lados de una unión p-n polarizada directamente, éstos se pueden recombinar radiativamente. Es en este mecanismo en el que están basados los diodos emisores de luz -LEDs por sus siglas en inglés. Un parámetro que mide la eficiencia de un LED es la eficiencia cuántica interna η_{int} , la cual se define como el cociente entre el número de fotones generados y el número de pares recombinados. De esta manera, η_{int} caracteriza las propiedades radiativas del semiconductor. En algunos materiales como el GaAs y el InP la eficiencia cuántica interna puede alcanzar valores cercanos a la unidad.

Otro parámetro importante para un LED es la eficiencia cuántica externa η_{ext} , definida como el cociente entre el número total de fotones emitidos por el dispositivo y el número total de huecos y electrones inyectados en la unión p-n, siendo este último número proporcional a la corriente que se hace pasar a través del dispositivo. La eficiencia cuántica externa es proporcional a η_{int} . Depende, además, de la construcción y de la geometría del LED, y en general de todos los parámetros que determinan la forma en que los cuantos de radiación son extraídos del interior del dispositivo.

Para entender esto último, consideremos que los pares electrón-hueco recombinados radiativamente en la vecindad de la unión p-n generan fotones con energías cercanas al de la brecha energética del semiconductor y de esta manera, a medida que se desplazan en el cristal rumbo a la superficie del LED, tienen una cierta probabilidad de ser absorbidos creando un nuevo par electrón-hueco. Este proceso recibe el nombre de autoabsorción y es una de las causas de que η_{int} difiera grandemente de η_{ext} . Así, aunque η_{int} sea cercana al 100%, η_{ext} no superará unas pocas unidades porcentuales en un LED basado en una homounión p-n.

Esta situación puede ser superada en buena medida en un LED basado en una heterounión. En efecto, consideremos, por ejemplo, la heterounión mostrada en la figura 2.2c, en la que se mediante una polarización directa se inyectan de manera preferencial huecos al semiconductor tipo n. Estos huecos, al recombinarse radiativamente, generan fotones con energías $h\nu$ cercanas a la brecha energética del semiconductor tipo n. Puesto que esta brecha es menor que la del semiconductor tipo p, el fotón generado puede viajar a través de este último sin ser absorbido apreciablemente. Precisamente en esto consiste el efecto de “ventana de banda ancha”, que es característico de las heterouniones. En los LEDs basados en heterouniones, η_{ext} puede incrementarse apreciablemente gracias a que la radiación generada en la vecindad de la heterounión alcanza la superficie del dispositivo sin ser absorbida apreciablemente. De esta manera, el empleo de heterouniones permite mejorar considerablemente la eficiencia de un LED.

2.3 La heteroestructura doble

El diagrama de bandas de un láser basado en una heterounión se muestra en la figura 2.6. Con letras mayúsculas se indican los emisores de electrones y huecos, los cuales tienen una brecha energética mayor que la de la región activa del láser. El tipo de conductividad de dicha región se denotará con n(p), a fin de considerar el caso general en el que pueda ser tanto n como p.

Al aplicar una polarización directa a esta heteroestructura doble, se inyectan electrones y huecos a la región activa desde los emisores N y P, en forma respectiva. Notamos que la heterounión n(p)-P constituye una barrera para los electrones inyectados, confinándolos a la región activa del láser aun en caso de que su longitud de difusión sea mucho mayor que el ancho d de la misma. Algo similar ocurre con los huecos inyectados, los cuales son confinados por la heterounión n(p)-N. La región activa de una heteroestructura doble constituye entonces un pozo de potencial para electrones y huecos.

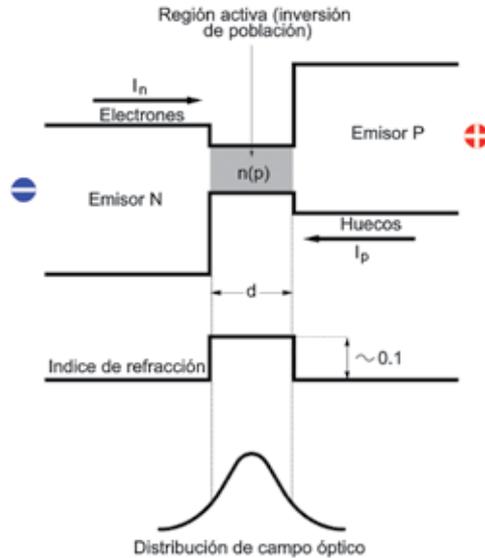


Fig 2.6 Láser de doble heterounión. a) Diagrama de bandas, b) perfil del índice de refracción, c) distribución de radiación óptica.

Así pues, los emisores de brecha energética grande, además de crear las condiciones para la inyección efectiva de portadores de carga a la región activa del láser, dan lugar a barreras de potencial que constriñen a los portadores inyectados a permanecer dentro de los límites de dicha región. Si comparamos un láser basado en una homounión p-n con uno de heteroestructura doble, tenemos que en el primer caso solamente una parte de los portadores de carga inyectados terminan en la región de amplificación del dispositivo, mientras que en el segundo caso todos los portadores de carga inyectados participan en la generación de radiación coherente. Esto es, el heteroláser aprovecha al máximo los portadores inyectados a su región activa para lograr la condición de inversión de población.

Además de lo anterior, el láser de heteroestructura tiene otra ventaja muy importante frente al láser de homounión. Esta ventaja

está relacionada al hecho de que los emisores N y P tienen un índice de refracción menor que el índice de la región activa del láser, debido a que esta última tiene una brecha energética menor que la de dichos emisores. En efecto, como se discutió en el capítulo anterior, la luz que viaja en un medio con índice de refracción n_1 puede experimentar una reflexión total al incidir sobre una interfaz con un medio de índice de refracción n_2 tal que $n_1 > n_2$, si el ángulo de incidencia es mayor que un cierto ángulo crítico. De esta manera, en una heteroestructura doble la región activa y los dos emisores forman una guía de onda altamente efectiva. Podemos, así mismo, afirmar que en un heteroláser la región de inversión de población y la región de propagación de la radiación óptica coinciden completamente, como se muestra esquemáticamente en la figura 2.6c, mientras que en un láser de homounión parte de la radiación óptica generada por la inyección de portadores se propaga fuera de la región activa (ver figura 1.8d), en donde el coeficiente de absorción es positivo.

En resumen, podemos decir que las mayores ventajas de los heteroláseres frente a los láseres de homounión son: 1) la de aprovechar en forma óptima los portadores de carga inyectados a la región activa para la generación de radiación y 2) la reducción de pérdidas debido al confinamiento óptico. La utilización de heterouniones para la fabricación de láseres de semiconductor permitió reducir la corriente de umbral para la generación de radiación coherente hasta valores del orden de 1 kA/cm², lo que a su vez permitió su operación del dispositivo en forma continua y a temperatura ambiente.

Las investigaciones con los primeros láseres de heterounión demostraron que al disminuir el espesor d de la región activa la corriente de umbral disminuye en forma concurrente. No obstante, cuando este espesor alcanza valores $d=0.2-0.3 \mu\text{m}$, la corriente de umbral lejos de disminuir con d aumenta bruscamente. La explicación de este comportamiento es como sigue. Al disminuir el espesor de la región activa será necesario inyectar menos portadores de carga con el objeto de alcanzar el nivel necesario para lograr la inversión de población, y de esta manera la corriente de umbral disminuye

efectivamente con d . Al alcanzar espesores del orden de $0.2\text{-}0.3\ \mu\text{m}$, sin embargo, la onda luminosa ya no puede contenerse dentro de los límites de la región activa, penetrando una cierta distancia en los emisores pasivos adyacentes. Por otro lado, para tener una inyección de portadores efectiva, los emisores N y P deben estar fuertemente dopados, de tal manera que contienen grandes concentraciones de electrones y huecos libres que, como se discutió en el apartado 2.2B, absorben los cuantos de radiación excitando “electrones calientes”. Son estos procesos de absorción de fotones, que representan pérdidas ópticas, los que conducen a un aumento en la corriente de umbral del láser.

De esta manera, para la heteroestructura doble existe un espesor óptimo del orden de $0.2\text{-}0.3\ \mu\text{m}$, por abajo del cual una disminución del mismo llevará a un incremento en corriente de umbral. La dependencia de esta corriente de umbral en función de d para un láser de heteroestructura doble se muestra esquemáticamente en la figura 2.7.

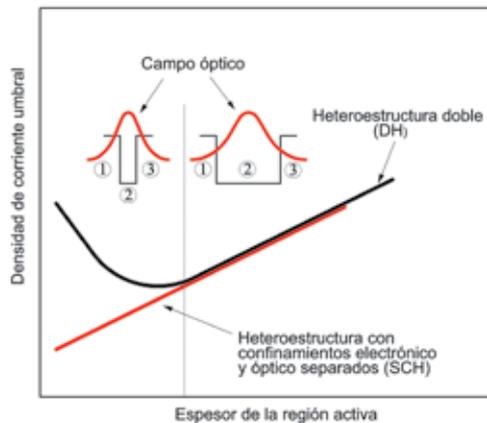


Fig 2.7 Diagrama esquemático de la densidad de corriente de umbral para la generación de radiación coherente para un láser de heteroestructura doble y para un láser con confinamientos electrónico y óptico separados.

2.4 La heteroestructura con confinamientos electrónico y óptico separados

Un refinamiento del láser de doble heteroestructura estudiado en la sección anterior es el láser con confinamientos electrónico y óptico separados, el cual es mostrado esquemáticamente en la figura 2.8. En esta figura, al igual que para el láser de la figura 2.6, los símbolos N y P indican los emisores de portadores de carga, los cuales tienen brechas energéticas mayores que la brecha de la región activa. A diferencia del láser de heteroestructura doble, sin embargo, la región

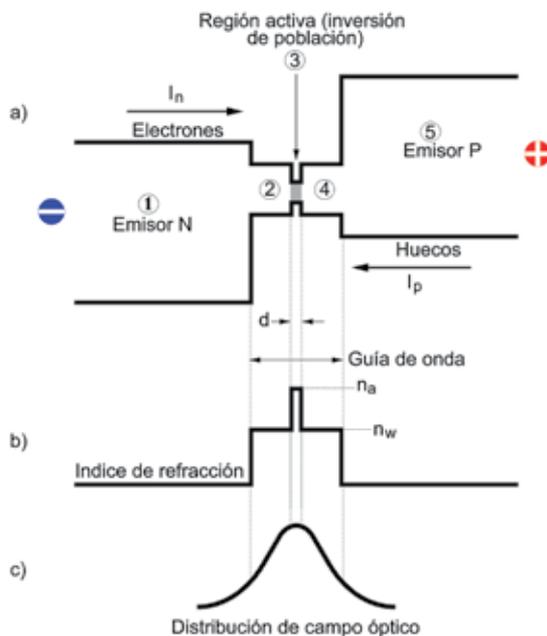


Fig 2.8 Láser con confinamientos electrónico y óptico separados. a) Diagrama de bandas, b) perfil del índice de refracción, c) distribución de radiación óptica.

activa del láser de la figura 2.8 consta no de una, sino de tres capas, indicadas por los números 2, 3, y 4. Estas tres capas forman un pozo de potencial para electrones y huecos, confinando a los portadores inyectados en la región 3, la cual tiene una brecha energética menor que la de las regiones 2 y 4. El espesor de la región 3 puede ser reducido hasta valores inferiores a los 100 Å, de tal manera que para alcanzar la inversión de población se necesitan considerablemente menos portadores de carga que en un láser de heteroestructura doble, lo que resulta en una corriente de bombeo menor.

Para ilustrar esto, consideremos la analogía entre la concentración de portadores y huecos en un pozo de potencial y el nivel de agua en un recipiente con fondo plano. El espesor del pozo en nuestra analogía corresponde al área del fondo del pozo. En la figura 2.9 se muestran dos recipientes con fondos de áreas A y A' , con A mucho mayor que A' . Supongamos que la inversión de población es equivalente a llenar los recipientes hasta un determinado nivel. Obviamente, para lograr esto será necesaria menos agua para el recipiente de área A' que para el recipiente de área A .

En un láser con confinamientos separados, la región en la que se localizan los portadores de carga puede ser mucho más angosta que en el caso de un láser de doble heteroestructura, con la consecuencia de que

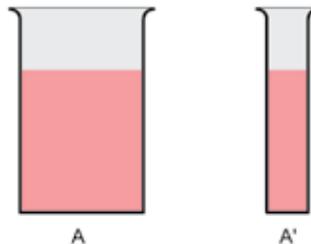


Fig 2.9 Recipientes de diferente área de base llenos de agua hasta un mismo nivel, mostrados como una analogía de dos láseres de diferentes espesores de región activa como se discute en el texto.

el campo óptico desborda dicha región penetrando hacia las capas 2 y 4. No obstante, debido a que la brecha energética de dichas capas es mayor que la de la región 3, y debido también a que sus niveles de impurificación son relativamente bajos, no hay pérdidas por absorción. Por otro lado, el cambio brusco en el índice de refracción entre los emisores N y P y las capas 2 y 4, hace que el campo óptico se localice dentro de las regiones 2, 3 y 4.

El perfil del índice de refracción en la región con inversión de población y la distribución del campo óptico están ilustrados en la figura 2.8. Comparando las figuras 2.6 y 2.8 podemos observar, que mientras que en un láser de heteroestructura doble las regiones de inversión de población y de localización del campo óptico coinciden, en el láser con confinamientos electrónico y óptico separados el campo óptico está confinado por los emisores N y P en la regiones 2, 3 y 4, mientras que los portadores de carga se localizan en la región 3. Precisamente por esta característica, a la estructura presentada en la figura 2.8 se le llama láser con confinamientos separados. En un láser de este tipo la densidad de corriente de umbral para la generación de radiación coherente se ha logrado reducir hasta valores en el rango 50-100 A/cm² y se han obtenido potencias ópticas de varios vatios en un régimen continuo.

2.5 Conclusiones

En la actualidad, en muchos laboratorios y empresas se investigan láseres de semiconductor con los más diversos diseños, empleando en su fabricación una variedad de procesos tecnológicos. Los dispositivos basados en el principio de confinamientos electrónico y óptico separados son los que han alcanzado los valores de corriente de umbral más bajos. Los láseres modernos son fabricados a base de heteroestructuras de decenas e incluso centenas de capas de diferente composición química. El espesor de estas capas es en algunos casos de unas cuantas decenas de ángstroms.

CAPÍTULO 3

Materiales para la construcción de láseres de semiconductor

3.1 Introducción

Un gran número de dispositivos optoelectrónicos modernos, entre los que contamos a los láseres de semiconductor, los diodos emisores de luz, los fotodetectores y las celdas solares, han sido construidos con base en las heterouniones semiconductoras. Como ya se explicó en el capítulo anterior, una heterounión se forma al poner en contacto metalúrgico dos semiconductores con distinta composición química. Las heterouniones semiconductoras ofrecen ventajas notables sobre las homouniones p-n para la fabricación de diferentes dispositivos electrónicos. Dichas ventajas tienen que ver, fundamentalmente, con el fenómeno de inyección preferencial de portadores de carga -electrones o huecos- en una heterounión polarizada positivamente. Esta inyección preferencial ocurre desde el semiconductor de brecha prohibida grande hacia el de brecha prohibida pequeña. Las ventajas de las heterouniones semiconductoras se manifiestan en forma patente en dispositivos que basan su funcionamiento en la interacción de la radiación con la materia; es decir, en los dispositivos optoelectrónicos basados en los procesos de absorción y generación de radiación.

La elección de los materiales semiconductores adecuados para la obtención de heterouniones con las propiedades ópticas y eléctricas deseadas es un problema de actualidad, al cual se le han dedicado numerosos trabajos y publicaciones científicas. Uno de los primeros

estudios sobre la posibilidad de utilizar diferentes sistemas de soluciones sólidas semiconductoras de muchas capas para la obtención de heteroestructuras ideales, fueron realizados en el Instituto Físico-Técnico A.E. Ioffe, en San Petesburgo, Rusia. En el laboratorio de Zh. I. Alferov en este Instituto, se desarrollaron varias tecnologías para la obtención de heteroestructuras semiconductoras que se emplearon en la fabricación de diferentes dispositivos. Entre éstos se encuentran: diodos emisores de luz, fotoreceptores, transistores, tiristores y celdas solares, cuyos parámetros de operación comúnmente exceden a los correspondientes parámetros de los dispositivos fabricados en base a uniones p-n.

El resultado más significativo obtenido en esta área por Alferov y su grupo, sin embargo, fue la creación del primer láser de semiconductor en el mundo operando en un régimen de corriente continua a temperatura ambiente y sin enfriamiento. Por sus investigaciones, a Zh. I. Alferov le fue otorgado el Premio Nobel en Física en el año 2000.

En este capítulo se presenta una introducción a las soluciones sólidas semiconductoras III-V y a su empleo en la fabricación de heteroestructuras láser. Se discute primeramente la estructura de bandas de los semiconductores y las condiciones que debe cumplir en la región activa de un láser. Posteriormente se describen los semiconductores y sus aleaciones ternarias o cuaternarias que en forma práctica se han empleado para las fabricación de láseres. Información adicional sobre este tema podrá ser encontrada en el libro de H.G. Casey y M.B. Panish, *Heterostructure lasers*, incluido en la bibliografía al final de este libro.

3.2 La estructura de bandas de los materiales semiconductores

En nuestra discusión sobre materiales semiconductores para la fabricación de láseres consideraremos primeramente aspectos relativos a la estructura de bandas de estos materiales, los cuales afectan críticamente sus propiedades eléctricas y ópticas. En lo que sigue,

analizaremos esta estructura de bandas, la cual veremos puede ser directa o indirecta.

En la física de los materiales sólidos, la dependencia de la energía del electrón E en función de su momentum P tiene una importancia central. Dicha relación es bastante complicada, mostrando máximos y mínimos para ciertos valores de P . Por otro lado, debido a que cualquier sistema tiende a un estado de equilibrio minimizando su energía, los electrones de la banda de conducción estarán localizados en el nivel más bajo o mínimo absoluto de la dependencia $E=E(P)$, mientras que los huecos se encontrarán en el máximo de la banda de valencia. De aquí que la localización relativa de estos puntos extremos de la relación $E(P)$ influyan de manera determinante en las propiedades electro-físicas y ópticas del material como se discute a continuación.

Si se da el caso que, ambos, el mínimo de la banda de conducción y el máximo de la banda de valencia están situados en un mismo valor de P , se dice que el semiconductor tiene una estructura de bandas directa. Este caso está ilustrado de manera esquemática en la figura 3.1a, en donde a los puntos extremos de ambas bandas están localizados en $P=0$. Si por el contrario, ambos puntos extremos se localizan en diferentes valores de P , tal como se muestra en la figura 3.1b en la que el máximo de la banda de valencia está localizado en $P=0$ mientras que el mínimo de la banda de conducción se localiza en un punto diferente designado por X , entonces el semiconductor tiene una estructura de bandas indirecta. Un ejemplo de un semiconductor con estructura de bandas directa es el arseniuro de galio (GaAs), mientras que el silicio y el germanio son dos ejemplos de materiales con estructura de bandas indirecta.

Es conocido que existen en la naturaleza leyes de conservación de ciertas cantidades físicas. Un ejemplo de esto es la *ley de conservación de la energía*, que se asume tiene una validez universal. En mecánica, al analizar los procesos de interacción de los cuerpos en movimiento con diferentes velocidades y masas se utiliza la *ley de conservación del momentum*. De acuerdo con esta ley, el momentum de un sistema -la

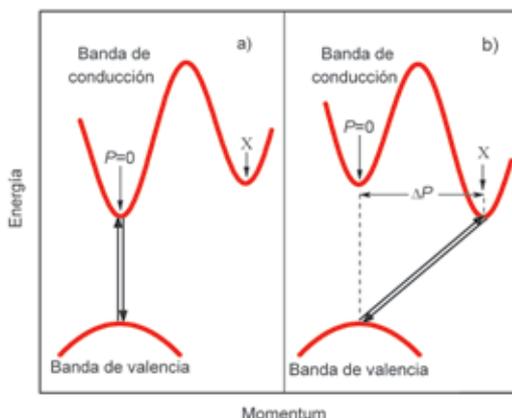


Fig. 3.1 Dependencia de la energía del electrón en función de su momentum $E(P)$ para semiconductores con estructura de bandas a) directa y b) indirecta. En el primer caso los electrones y los huecos están localizados en el mismo punto $P=0$, mientras que en el segundo caso los electrones se localizan en X y los huecos en $P=0$. Las dependencias mostradas son solamente esquemáticas.

suma de los productos masa por velocidad de los cuerpos participantes es el mismo antes y después de la interacción. La ley de conservación del momentum debe de cumplirse también en los semiconductores durante los procesos de interacción entre electrones y huecos. Dos posibles procesos de este tipo son la generación de un par electrón-hueco mediante la absorción de un cuanto de radiación $h\nu$ y el proceso de recombinación de un par electrón-hueco acompañado de la emisión de un cuanto de radiación con una energía aproximadamente igual al ancho de la brecha prohibida del semiconductor. Ambos procesos fueron estudiados en el capítulo uno.

Analicemos los procesos de generación y recombinación de pares a la luz de la ley de conservación del momentum. El proceso de recombinación de un electrón y un hueco de un semiconductor directo se representa, como se muestra en la figura 3.1a, en forma de una flecha vertical dirigida hacia abajo. De manera análoga, con una flecha

vertical dirigida hacia arriba se puede representar la generación de un par electrón-hueco. (figura 3.1a). Se hace notar que en ambos casos, excitación y recombinación de pares, se cumple la ley de conservación del momentum ya que el momentum del electrón permanece invariable.

En contraste, como se ilustra en la figura 3.1b, en un semiconductor de estructura de bandas indirecta los procesos de recombinación y de generación de pares electrón-hueco son acompañados por un cambio en el momentum del electrón, debido a que el mínimo de la banda de conducción y el máximo de la banda de valencia están localizados en diferentes valores de P . Por esta razón, en este caso dichos procesos se representan con flechas inclinadas.

Sin embargo, de acuerdo a la ley de conservación de momentum, el cambio de momentum ΔP del electrón (ver figura 3.1b) al efectuar la transición entre las dos bandas debe ser compensado de alguna manera. Esto se logra mediante la interacción con vibraciones de la red cristalina. En la física de los materiales sólidos estas vibraciones se consideran como cuasipartículas llamadas fonones. De esta manera, para que se cumpla la ley de la conservación del momentum en las transiciones indirectas es necesaria la participación de tres partículas: el electrón, el fotón y el fonón, mientras que, en las transiciones directas, sólo participan dos partículas: el fotón y el electrón. El proceso en el que participan tres partículas siempre es menos probable que el proceso donde toman parte dos, por eso la eficiencia de la recombinación radiativa en los materiales con una estructura de bandas indirecta es sumamente baja.

Por todo lo anteriormente expresado, podemos concluir que para la fabricación de fuentes semiconductoras de radiación electromagnética, tanto coherente como incoherente -láseres de semiconductor y diodos emisores de luz, de manera respectiva-, es necesario elegir materiales con una estructura de bandas directa para la región activa de los dispositivos.

Un láser de heteroestructura está diseñado de tal manera que su región activa, con un cierta brecha energética, quede emparedada entre

dos semiconductores con una brecha prohibida mayor. Para efectos de confinamiento eléctrico, la brecha energética de la región activa debe ser menor que la de las capas adyacentes. Además, el semiconductor que forma la región activa debe tener una estructura de bandas directa como se discutió líneas arriba. Es, además, necesario que las hetero-interfaces estén libres de defectos cristalinos que pudieran degradar las propiedades luminiscentes del láser.

Frecuentemente estos defectos interfaciales surgen por una diferencia entre los parámetros de red -es decir, la distancia entre dos átomos- de la región activa y de las capas de confinamiento. Como una manera de evitar defectos interfaciales en una heteroestructura láser, podríamos entonces escoger materiales con idéntico parámetro de red. Una heterounión entre dos materiales semiconductores de composición química diferente pero de idéntico parámetro de red, recibe el nombre de heterounión ideal*.

A la luz de lo anteriormente expresado, en este capítulo analizaremos la posibilidad de fabricar láseres de semiconductor de heteroestructura atendiendo a los siguientes puntos:

- a) Los materiales que forman la heterounión deben tener un mismo tipo de red cristalina y sus parámetros de red deben coincidir.

*Hacemos notar que en algunos casos es posible obtener heterouniones con un bajo nivel de defectos interfaciales a partir de materiales con parámetros de red diferentes entre sí. Por ejemplo, es factible fabricar láseres de heterounión con una región activa de composición $\text{In}_{0.2}\text{Ga}_{0.8}\text{As}$ emparedada entre dos capas de GaAs. En este caso, puesto que el parámetro de red del $\text{In}_{0.2}\text{Ga}_{0.8}\text{As}$ es más grande que el del GaAs, la película ternaria crece en un estado de compresión. Si el espesor de la película alcanza un cierto valor crítico, la energía elástica acumulada en la película comprimida es lo suficientemente grande para generar defectos interfaciales. Por esta razón, el espesor de la región activa no debe sobrepasar unos 100 Å. En lo que sigue, y no obstante que los láseres de heterounión con semiconductores de parámetros de red diferente juegan en la actualidad un importante papel, limitaremos nuestro tratamiento a los láseres con heterouniones ideales a fin de mantener la discusión en los términos más simples posibles.

- b) El semiconductor con brecha prohibida angosta debe tener una estructura directa de bandas energéticas.

3.3 Compuestos binarios III-V

Dos de los materiales semiconductores más conocidos son el silicio y el germanio, constituidos por átomos situados en la columna IV de la tabla periódica; es decir, por átomos que tienen cuatro electrones en su capa exterior. Esta última característica es la que hace que el silicio y el germanio sean semiconductores.

El silicio y el germanio, sin embargo, tienen una estructura de bandas indirecta. Además, aunque ambos semiconductores tienen una estructura cristalina del tipo “diamante”, sus parámetros de red son sensiblemente diferentes entre si. Por lo tanto, de acuerdo a los puntos a) y b) mencionados al final de la sección anterior, éstos materiales no pueden ser utilizados como componentes de heterouniones ideales para dispositivos emisores de radiación.

Consideremos ahora los compuestos binarios $A^{III}B^V$, donde A^{III} y B^V son, de manera respectiva, elementos de las columnas III y V de la tabla periódica. Puesto que los elementos A^{III} y B^V tienen, respectivamente, 3 y 5 electrones en su última capa, los átomos constituyentes de los compuestos $A^{III}B^V$ tienen en promedio 4 electrones en dicha capa y podemos esperar que tengan propiedades semiconductoras al igual que el silicio y el germanio. De manera análoga, podemos esperar que los compuestos $A^{II}B^{VI}$, constituidos por elementos de los grupos II y VI, sean también semiconductores. Es claro, que utilizando como un criterio la existencia de 4 electrones en la última capa, se puede aumentar la lista de materiales que esperaríamos tuvieran propiedades semiconductoras. Un ejemplo de esto son los compuestos $A^{II}B^{IV}C^V_2$, cuyos átomos constituyentes tienen en promedio 4 electrones en su capa exterior.

Consideremos la formación de semiconductores $A^{III}B^V$ con elementos de las columnas III y V de la tabla periódica, que de manera

específica son: B, Al, Ga, In (columna III) y N, P, As, y Sb (columna V). Se hace notar que no consideramos compuestos que contengan boro o nitrógeno; esto debido a que todos los compuestos con boro son tecnológicamente muy complicados de obtener, mientras que los compuestos con nitrógeno tienen una estructura cristalina de tipo “wurzita”, a diferencia de todos los demás compuestos binarios III-V que tienen una estructura de tipo “esfalerita”. De esta última característica y de acuerdo al punto número uno al final de la sección anterior, relativo a los requisitos necesarios para obtener una heterounión ideal, ésta solamente podrá ser obtenida entre dos nitruros; otras combinaciones, como por ejemplo, nitruro-fosfuro o nitruro-arseniuro, se excluyen por sus diferentes tipos de red. Se hace notar, sin embargo, que ambos compuestos con boro y nitrógeno tienen una gran importancia tecnológica. En particular, el nitruro de galio es la base de los láseres y emisores de luz en la región azul del espectro.

De los restantes 6 elementos, Al, Ga, In (columna III) y P, As, Sb (columna V) se obtienen 9 compuestos semiconductores binarios III-V, a saber: AlP*, GaP*, InP, AlAs*, GaAs, InAs, AlSb*, GaSb y InSb, en donde los asteriscos identifican a los 4 compuestos que tienen transiciones indirectas mientras que los 5 restantes tienen transiciones directas. Todos los compuestos anteriores se pueden obtener de una manera relativamente fácil con ambos tipos de conductividad p y n. Los valores de las brechas de energía prohibida a temperatura ambiente y los parámetros de red -denotados por la letra griega α - de estos nueve compuestos III-V se muestran en la tabla 3.1.

En relación a los semiconductores II-VI, aunque algunos de ellos poseen muy buenas propiedades luminiscentes -como por ejemplo, ZnO, ZnSe y ZnTe- en general tienen desventajas desde el punto de vista tecnológico. Se encuentra, por ejemplo, que es muy difícil invertir el tipo de conductividad de los semiconductores II-VI y es por lo tanto complicado obtener uniones p-n. Debido a esta circunstancia, en lo que sigue consideraremos solamente a los semiconductores III-V, los cuales están tecnológicamente bien desarrollados.

Tabla 3.1 Brecha de energía prohibida E_g a temperatura ambiente y parámetro de red α para algunos compuestos binarios III-V. Los asteriscos en la primera columna indican que el compuesto binario tiene una brecha de energía indirecta.

Compuesto binario III-V	Brecha energética E_g , eV	Parámetro de red α , Å
AIP*	2.45	5.45
GaP*	2.26	5.45
InP	1.35	5.86
AlAs*	2.16	5.66
GaAs	1.42	5.65
InAs	0.36	6.06
AlSb*	1.68	6.14
GaSb	0.72	6.10
InSb	0.17	6.48
InSb	0.17	6.48

3.4 Soluciones sólidas ternarias $\text{Al}_x\text{Ga}_{1-x}\text{As}$

Como se observa de la tabla 3.1, para los dos pares de compuestos AlAs-GaAs y AlP-GaP el parámetro de red es α prácticamente idéntico. Los compuestos de aluminio con elementos del grupo V son, sin embargo, inestables al contacto con el vapor de agua de la atmósfera, formando rápidamente el óxido Al_2O_3 . Por esta razón, la utilización de estos compuestos para la fabricación de heterouniones es complicada. Afortunadamente, con casi todos los compuestos del tipo III-V es posible obtener soluciones sólidas químicamente estables en un amplio rango de composiciones. Por ejemplo, para el GaAs, en el que los átomos de galio y arsénico están localizados en lugares predeterminados de la red cristalina formando subredes metálicas y no metálicas, algunos átomos de galio pueden ser sustituidos por átomos de aluminio. A medida que se incrementa el número de átomos sustituidos varía la composición de la solución sólida y de esta manera se modifican sus propiedades semiconductoras.

De acuerdo con la *ley de Vegard*, el parámetro de red de una solución sólida Varía linealmente con la composición entre los dos parámetros de red de los dos semiconductores que la forman. De este modo, en el caso del sistema AlAs–GaAs el parámetro α resulta casi independiente de la composición x , debido a que α_{AlAs} tiene un valor muy cercano a α_{GaAs} .

Las soluciones sólidas basadas en GaAs y AlAs tienen un número de átomos del grupo III igual al número de átomos del grupo V, por lo que pueden expresarse mediante la fórmula química, $\text{Al}_x\text{Ga}_{1-x}\text{As}$, donde x es la composición molar de AlAs en la fase sólida y $1-x$ la correspondiente composición de GaAs. Las soluciones sólidas $\text{Al}_x\text{Ga}_{1-x}\text{As}$ son estables desde el punto de vista corrosivo en el intervalo de composiciones $x < 0.8$.

La dependencia del ancho de la brecha prohibida y del parámetro de red de una solución sólida en función de su composición química tiene una gran importancia práctica. La brecha energética $E_g(x)$ para el sistema $\text{Al}_x\text{Ga}_{1-x}\text{As}$ está dada en la Figura 3.2, en donde el eje horizontal corresponde al parámetro x , y el eje vertical a la brecha energética en unidades de eV. El parámetro x varía entre 0 y 1; $x=0$ corresponde a GaAs y $x=1$ a AlAs. Se hace notar la brecha energética de GaAs es directa mientras que la de AlAs es indirecta.

Las curvas en colores rojo y azul en la figura 3.2 representan, de manera respectiva, las brechas directa e indirecta de la aleación $\text{Al}_x\text{Ga}_{1-x}\text{As}$. Aquí cabe recordar que el valor del ancho de la brecha prohibida del material está dada por la distancia entre el máximo de la banda de valencia y el mínimo de la banda de conducción. De este modo, de acuerdo con la figura 3.2, en el rango de composiciones entre $x=0$ y $x=0.43$ la aleación tiene una estructura de bandas directa y el mínimo en $P=0$ tiene una energía menor que el mínimo en X . En contraste, para composiciones $x > 0.43$ el mínimo en $P=0$ tiene una energía mayor que el mínimo en X y la aleación tiene una estructura de bandas indirecta. Así, la composición $x=0.43$ representa la transición entre las brechas energéticas directa e indirecta de la estructura de bandas.

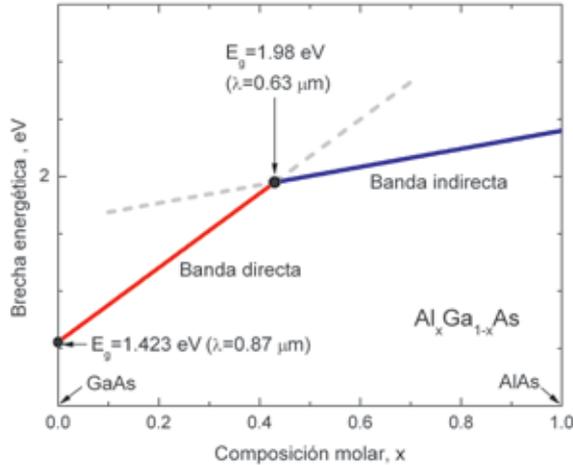


Fig.3.2 Dependencia del ancho de la brecha prohibida con respecto de la composición para el sistema de soluciones sólidas $\text{Al}_x\text{Ga}_{1-x}\text{As}$. Las soluciones sólidas cambian su estructura de bandas de directa a indirecta a una composición $x=0.43$. Por simplicidad las dependencias de la brechas energéticas con la composición se representan con líneas rectas. Las dependencias reales muestran una ligera curvatura que las aparta de las líneas rectas.

Se hace notar que por simplicidad, las dependencias de las brechas energéticas directa e indirecta con respecto a la composición de aluminio se han representado en la figura 3.2 por líneas rectas. En realidad, experimentalmente se encuentra que dichas dependencias tienen ligeras curvaturas que las alejan de una dependencia estrictamente lineal.

De la figura 3.2 podemos ver que la brecha prohibida directa E_g para el GaAs crece con x desde 1.42 eV para GaAs, hasta 1.98 eV para $x=0.43$, cubriendo la región espectral 0.87-0.63 μm . Esto es, ya que el parámetro de red de $\text{Al}_x\text{Ga}_{1-x}\text{As}$ es casi independiente de x , con base en estas soluciones sólidas es posible preparar heterouniones ideales capaces de emitir desde el infrarrojo ($\lambda=0.87 \mu\text{m}$) hasta el

rojo ($\lambda=0.63 \mu\text{m}$). Cabe señalar, que los primeros heteroláseres que funcionaron en un régimen continuo a temperatura ambiente fueron diseñados como heteroestructuras dobles con base en el sistema GaAs-AlGaAs. El diagrama de bandas de esta heteroestructura doble se muestra en la Figura 3.3.

3.5 Soluciones sólidas ternarias GaAsP, GaInP y AlInP

Mientras que las aleaciones $\text{Al}_x\text{Ga}_{1-x}\text{As}$ son muy convenientes para aplicaciones que demanden longitudes de onda en el rango $0.63\text{-}0.87 \mu\text{m}^*$, hay muchas otras aplicaciones que requieren fuentes de radiación en las regiones espectrales del amarillo y del verde que son imposibles de obtener con base en el $\text{Al}_x\text{Ga}_{1-x}\text{As}$. En lo que sigue, analizaremos otros dos sistemas ternarios, $\text{GaAs}_{1-x}\text{P}_x$ y $\text{GaIn}_{1-y}\text{P}$, con los cuales es posible fabricar emisores de radiación en la región verde-amarillo del espectro visible.

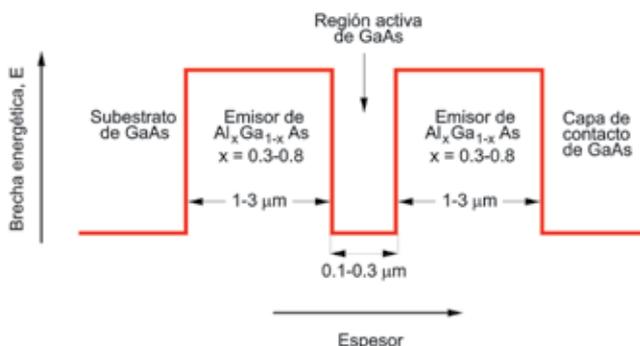


Fig.3.3 Diagrama de bandas de energía de una heteroestructura doble del sistema AlGaAs/GaAs. La solución sólida $\text{Al}_x\text{Ga}_{1-x}\text{As}$ corresponde a los emisores, mientras que el GaAs se utiliza en la región activa y como sustrato.

Consideremos primeramente la solución sólida $\text{GaAs}_{1-x}\text{P}_x$. La línea azul en la figura 3.4(a) muestra la dependencia de la brecha prohibida E_g con la composición x para esta aleación. Como puede observarse, E_g crece con x desde 1.42 eV, valor que corresponde a GaAs ($x=0$), hasta 2.26 eV para GaP ($x=1$). El cambio de pendiente observado en la curva E_g - x a una concentración $x=0.45$ corresponde a la transición de la brecha prohibida de directa ($x<0.45$) a indirecta ($x>0.45$). El valor de la brecha prohibida en el punto de la transición es $E_g=2.06$ eV, lo que hace a la aleación $\text{GaAs}_{1-x}\text{P}_x$ un candidato para la fabricación de fuentes de radiación en la región roja del espectro.

Consideremos ahora la solución sólida $\text{Ga}_y\text{In}_{1-y}\text{P}$. La línea roja de la figura 3.4(a) muestra la dependencia de su brecha prohibida E_g con respecto de la composición y . Debido a que la aleación $\text{Ga}_y\text{In}_{1-y}\text{P}$ tiene una estructura de bandas directa hasta $y=0.7$, en donde tiene una brecha energética de 2.25 eV, dicha aleación puede ser utilizada para la fabricación de fuentes de radiación en un intervalo de longitudes de onda que comprende desde el rojo hasta el amarillo-verde del espectro visible.

Lamentablemente, ni el $\text{GaAs}_{1-x}\text{P}_x$ ni el $\text{Ga}_y\text{In}_{1-y}\text{P}$ pueden ser utilizados para la obtención de heterouniones ideales, debido a que el parámetro de red de ambas aleaciones depende considerablemente de su composición. En efecto, para $\text{Ga}_y\text{In}_{1-y}\text{P}$ este parámetro varía en el rango desde $\alpha=5.86 \text{ \AA}$ (InP) hasta $\alpha=5.45 \text{ \AA}$ (GaP), mientras que para $\text{GaAs}_{1-x}\text{P}_x$ lo hace desde $\alpha=5.65 \text{ \AA}$ (GaAs) hasta $\alpha=5.45 \text{ \AA}$ (GaP). En la figura 3.4b se muestran las dependencias α -(x,y) para ambas aleaciones. De manera analítica, dichas relaciones están dadas por: $\alpha_{\text{GaInP}} = -0.41y+5.86$ para $\text{Ga}_y\text{In}_{1-y}\text{P}$ y $\alpha_{\text{GaAsP}} = -0.20x+5.65$ para $\text{GaAs}_{1-x}\text{P}_x$. De la figura 3.4b puede verse que la aleación $\text{Ga}_{0.51}\text{In}_{0.49}\text{P}$

*Hay que hacer notar que para soluciones sólidas con brechas energéticas cerca de la transición directa-indirecta en $0.63 \mu\text{m}$, la población de electrones excitados se reparte entre los valles directo e indirecto de la banda de conducción, y que la eficiencia cuántica interna se reduce en la medida en que más electrones ocupan el valle indirecto.

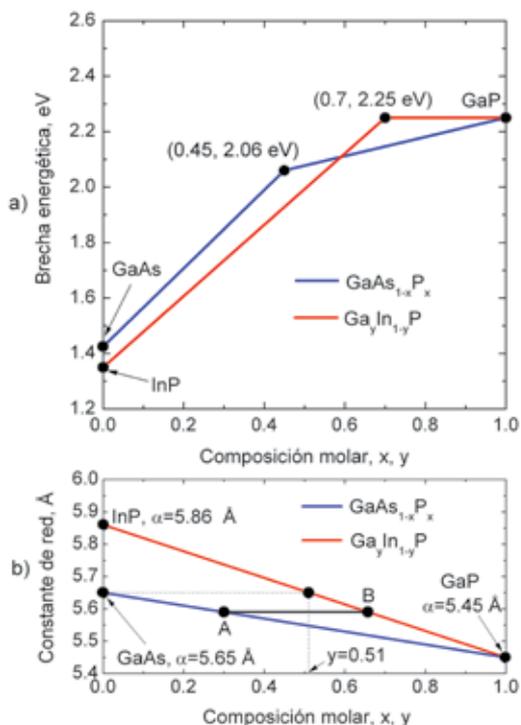


Fig.3.4 Sistemas de soluciones sólidas GaAsP y GaInP. a) Dependencia del ancho de la brecha prohibida en función de la composiciones x , y de las soluciones sólidas $Ga_{1-x}As_xP_x$ y $Ga_yIn_{1-y}P$, en forma respectiva. b) Variación del parámetro de red con x , y para las mismas soluciones.

tiene el mismo parámetro de red que el GaAs. Será entonces posible obtener heterouniones ideales entre el GaAs y el $Ga_{0.51}In_{0.49}P$.

Consideremos ahora las dos composiciones x , y unidas por el segmento A-B en la figura 3.4b. Es claro que para estas composiciones las aleaciones $Ga_{1-x}As_xP_x$ y $Ga_yIn_{1-y}P$ tienen un mismo parámetro de red y es por lo tanto posible obtener heterouniones ideales con tales composiciones. A partir de la identidad $\alpha_{GaInP} = \alpha_{GaAsP}$, dedu-

timos que la condición para la igualdad de parámetros de red es $y=0.49x+0.51$. Dicho de otra manera, es posible obtener heterouniones ideales $\text{GaAs}_{1-x}\text{P}_x$ - $\text{Ga}_y\text{In}_{1-y}\text{P}$ siempre que se cumpla la condición $y=0.49x+0.51$.

Consideremos finalmente las soluciones sólidas $\text{Al}_x\text{In}_{1-x}\text{P}$. De todos los sistemas III-V, excluyendo los compuestos de B y N, esta aleación es el que tiene la brecha prohibida con estructura directa más grande. La dependencia E_g - x para $\text{Al}_x\text{In}_{1-x}\text{P}$ se representa con una línea azul en

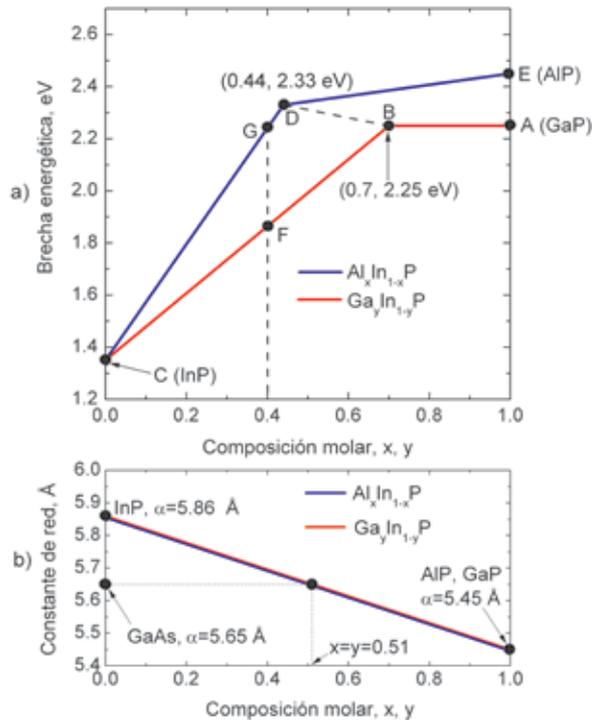


Fig.3.5 Las soluciones sólidas AlInP y GaInP. a) Dependencia de la brecha energética con la composición de las soluciones sólidas $\text{Al}_x\text{In}_{1-x}\text{P}$ y $\text{Ga}_y\text{In}_{1-y}\text{P}$. b) Dependencia del parámetro de red con la composición para estas mismas soluciones sólidas.

la figura 3.5a. Podemos ver de esta figura que el punto de transición de la estructura directa a la indirecta se encuentra a una composición $x=0.44$, a la que corresponde una brecha prohibida $E_g=2.33$ eV. En la figura 3.5a se representa también, con una línea roja, la dependencia de E_g y para el sistema $Ga_yIn_{1-y}P$ ya anteriormente analizado. Debido a que los parámetros de red de AIP y de GaP son muy parecidos y a que el InP se utiliza como segundo componente binario en ambos sistemas AIP-InP y GaP-InP, las dependencias $\alpha_{AlInP}(x)$ y $\alpha_{GaInP}(y)$ para las soluciones sólidas $Al_xIn_{1-x}P$ y $Ga_yIn_{1-y}P$ son muy parecidas entre sí, como se muestra en la figura 3.5b.

Tenemos entonces que será posible obtener heterouniones ideales entre $Al_xIn_{1-x}P$ y $Ga_yIn_{1-y}P$ si coincide la concentración de InP en las dos aleaciones. En la figura 3.5a esta condición se representa por la línea recta FG -para el caso particular $x=y=0.4$ - a la cual se le llama sección isoperiódica. Por medio de las secciones isoperiódicas se pueden encontrar los valores de los anchos de las brechas prohibidas de $Al_xIn_{1-x}P$ y $Ga_yIn_{1-y}P$ para los que coinciden sus parámetros de red. A partir de la figuras 3.4b y 3.5b es posible deducir que para $x=y=0.51$ las constantes de red de las aleaciones $Ga_{0.51}In_{0.49}P$ y $Al_{0.51}In_{0.49}P$ coinciden con la constante de red de GaAs. Esto es, el GaAs puede ser empleado como sustrato para la fabricación de heterouniones ideales $Ga_{0.51}In_{0.49}P-Al_{0.51}In_{0.49}P$ con brechas energéticas $E_g=1.91$ eV ($Ga_{0.51}In_{0.49}P$) y $E_g=2.35$ eV ($Al_{0.51}In_{0.49}P$).

3.6 Soluciones sólidas cuaternarias AlGaInP, GaInAsP, GaInAsSb y AlGaAsSb

Consideremos la aleación ternaria $Ga_yIn_{1-y}P$. Suponga que, manteniendo el contenido de indio constante, parte de los átomos de galio son sustituidos por átomos de aluminio. Este proceso de sustitución nos lleva de la aleación ternaria $Ga_yIn_{1-y}P$ a la cuaternaria $Al_xGa_yIn_{1-x-y}P$, en donde x representa la composición de AIP de la solución sólida. Notamos que el parámetro de red de la aleación cuaternaria es

independiente de x , ya que $\alpha_{\text{GaP}} = \alpha_{\text{AlP}}$. En contraste, la brecha energética se incrementa con el contenido de aluminio de la aleación. En efecto, de la figura 3.5a podemos deducir que, manteniendo $x=y$, dicha brecha se incrementa desde el valor de E_g para $\text{Ga}_y\text{In}_{1-y}\text{P}$ hasta el correspondiente valor para $\text{Al}_x\text{In}_{1-x}\text{P}$.

Si en las soluciones sólidas ternarias la dependencia del ancho de la brecha prohibida y del parámetro de red en función de la composición se representa en una línea, para las soluciones cuaternarias ésta se representa en un plano. Las soluciones sólidas $\text{Al}_x\text{Ga}_y\text{In}_{1-x-y}\text{P}$ se encuentran situadas, como se observa en la figura 3.5a, en la región limitada por $E_g(y)_{\text{GaInP}}$, por $E_g(x)_{\text{AlInP}}$ y por la línea vertical que une a los puntos A y E, los cuales corresponden a los valores del ancho de las brechas prohibidas de GaP y de AlP, en forma respectiva. La brecha prohibida de la solución sólida ternaria $\text{Al}_x\text{Ga}_{1-x}\text{P}$ varía con x a lo largo del segmento AE.

En los sistemas ternarios de soluciones sólidas semiconductoras la variación de la composición lleva consigo cambios, tanto en la brecha prohibida E_g , como en el parámetro de red α . En los sistemas cuaternarios se tiene un grado adicional de libertad, de modo que es posible variar E_g y α de manera independiente. En estos sistemas existen secciones isoperiódicas a lo largo de las cuales el parámetro de red es constante. Por ejemplo, para $\text{Al}_x\text{Ga}_y\text{In}_{1-x-y}\text{P}$ una de estas secciones está dada por la condición $x+y=0.51$ y corresponde al parámetro de red de GaAs. Esto es, el GaAs puede ser utilizado como substrato para la obtención de heterouniones ideales con base en el sistema de soluciones sólidas AlGaInP. En particular, en esta sección isoperiódica se crecen, con ayuda de diferentes métodos tecnológicos, heteroláseres que funcionan en la región amarillo-verde del espectro visible.

La dependencia del ancho de la brecha prohibida con respecto de la composición mostrada en la figura 3.5a para el $\text{Al}_x\text{Ga}_y\text{In}_{1-x-y}\text{P}$, puede ser representada en forma tridimensional con la ayuda de un triángulo equilátero como se ilustra en la figura 3.6. En los lados del triángulo se representan los sistemas ternarios GaInP, AlInP y AlGaP, mientras que el eje vertical corresponde al valor de la brecha prohibida. Las

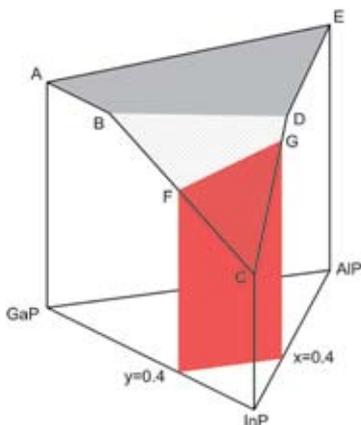


Fig.3.6 El sistema GaP-InP-AIP se representa mediante un triángulo equilátero. Los lados del triángulo corresponden a soluciones sólidas ternarias, mientras que los planos proyectados representan soluciones sólidas cuaternarias.

líneas ABC, EDC y AE muestran la dependencia de E_g con la composición de las aleaciones ternarias GaInP , AlInP y AlGaP , de manera respectiva. Las soluciones sólidas cuaternarias están representadas por planos. De este modo, los planos ABDE y CBD representan las regiones en donde la aleación cuaternaria tiene estructura de bandas indirecta y directa, en forma respectiva. La línea BD de intersección de estas superficies corresponde a la frontera entre los dos tipos de estructura de bandas. La figura 3.5a puede obtenerse de la figura 3.6, si los planos en los que están representados el GaInP y el AlInP se traslapan girándolos sobre el eje de InP .

En el triángulo de concentraciones GaP-AIP-InP, que se localiza en el plano horizontal en la figura 3.6, una línea paralela a uno de los lados del triángulo corresponde a aleaciones con una concentración fija del semiconductor binario del vértice opuesto. Por ejemplo, una línea paralela al eje GaP-AIP corresponde a soluciones sólidas con un contenido fijo de InP . La intersección del plano vertical en color rojo en la figura

3.6 con el plano base del triángulo de composiciones corresponde a las aleaciones $\text{Al}_x\text{Ga}_y\text{In}_{1-x-y}\text{P}$ con $x+y=0.4$. A lo largo de la intersección FG de este plano vertical con en plano BDC encontraremos, entonces, a todas las soluciones cuaternarias con una brecha prohibida intermedia entre las de $\text{Ga}_{0.4}\text{In}_{0.6}\text{P}$ y $\text{Al}_{0.4}\text{In}_{0.6}\text{P}$, pero con una misma composición de InP y por tanto con una misma constante de red.

En la tabla 3.2 se listan diferentes variantes de heterouniones ideales en el sistema AlGaInP y las condiciones de isoperiodicidad de las aleaciones que las constituyen.

En la Figura 3.7 se grafican las brechas prohibidas de diferentes aleaciones ternarias de semiconductores III-V en función de su constante de red. Los círculos negros en dicha figura corresponden a los nueve semiconductores binarios III-V que generan las aleaciones ternarias. Los semiconductores binarios con una estructura indirecta de bandas, tal como se hizo anteriormente, se señalan con un asterisco. Los puntos están unidos por líneas que describen la dependencia de la brecha prohibida con respecto al parámetro de red de las aleaciones ternarias. Como se mencionó anteriormente, de acuerdo a la ley de Vegard el parámetro de red varía linealmente con la composición de dichas aleaciones.

Las funciones $E_g(\alpha)$ mostradas en la figura 3.7 representan la dependencia del ancho de la brecha prohibida respecto de las composición

Tabla 3.2 Heterouniones ideales con base en las soluciones sólidas cuaternarias $\text{Al}_x\text{Ga}_y\text{In}_{1-x-y}\text{P}$.

Aleaciones	Condición de isoperiodicidad
$\text{Ga}_x\text{In}_{1-x}\text{P} - \text{Al}_x\text{In}_{1-x}\text{P}$	$x=x'$
$\text{Ga}_x\text{In}_{1-x}\text{P} - \text{Al}_x\text{Ga}_y\text{In}_{1-x-y}\text{P}$	$x=x'+y$
$\text{Al}_x\text{In}_{1-x}\text{P} - \text{Al}_x\text{Ga}_y\text{In}_{1-x-y}\text{P}$	$x=x'+y$
$\text{Al}_x\text{Ga}_y\text{In}_{1-x-y}\text{P} - \text{Al}_x\text{Ga}_y\text{In}_{1-x-y}\text{P}$	$x+y=x'+y'$

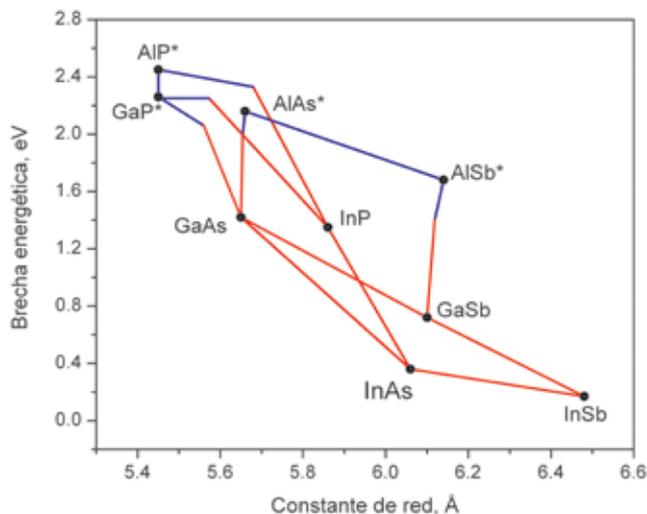


Fig.3.7 Brecha energética E_g de 9 aleaciones binarias con semiconductores III-V en función de su constante de red a . Con color azul se ilustran las transiciones indirectas y con el rojo las directas.

para diferentes sistemas de soluciones sólidas ternarias, parte de las cuales ya fueron analizadas en las figuras 3.4 y 3.5. Si se da el caso que la solución sólida esté formada por dos semiconductores con igual tipo de estructura de bandas, directa o indirecta, la función $E_g(\alpha)$ se representa con una línea recta ente los puntos de los semiconductores binarios que forman la aleación. Esto debido a que su estructura de bandas de las soluciones sólidas ternarias tiene la misma naturaleza que las de sus componentes binarias. Si, por el contrario, la aleación ternaria está formada por dos semiconductores binarios, uno de los cuales tiene una estructura de bandas directa y el otro indirecta, la dependencia $E_g(\alpha)$ se representará por dos segmentos rectos que se unen en el punto en donde la brecha de la aleación ternaria cambia de directa a indirecta.

En la figura 3.7 los segmentos en rojo corresponden a brechas energéticas directas y los segmentos en azul a brechas indirectas.

Puede verse que las aleaciones GaInAs, InAsP, InAsSb, GaAsSb y GaInSb tienen una estructura de bandas directa en todo el rango de composición, mientras que la brecha energética de las aleaciones AlAsSb y AlGaP es indirecta, también en todo el rango de composición. En contraste, las aleaciones AlGaAs, AlInP, GaAsP y GaInP muestran una transición entre una brecha energética directa a una indirecta a una composición característica.

Tomando en cuenta que, según la ley de Vegard, el parámetro de red de una solución sólida varía linealmente con la composición, concluimos que la dependencia de E_g con respecto a α es funcionalmente la misma que la dependencia E_g con respecto a x .

Por otro lado, es necesario mencionar que las dependencias $E_g(x)$ mostradas en la figura 3.7 como líneas rectas son solamente esquemáticas, pues en general dichas dependencias muestran una curvatura que puede ser pronunciada, como es el caso de las aleaciones GaAsSb y InAsSb. No obstante y a fin de facilitar la discusión, las funciones $E_g(x)$ se representan en la figura 3.7 de manera conveniente como líneas rectas.

Las secciones isoperiódicas corresponden en la figura 3.7 a líneas rectas verticales. Las soluciones cuaternarias se localizan a lo largo de estas líneas entre las intersecciones con las dos soluciones ternarias correspondientes. Las secciones isoperiódicas con los semiconductores binarios empleados como substratos son, por supuesto, particularmente importantes. En la tabla 3.3 se listan algunas soluciones ternarias y cuaternarias cuyo parámetro de red puede acoplarse al de los semiconductores más comúnmente empleados como substratos.

Vamos a analizar con detalle el sistema cuaternario $\text{Ga}_x\text{In}_{1-x}\text{As}_y\text{P}_{1-y}$, que es una de las soluciones sólidas de más importancia desde el punto de vista práctico. En la figura 3.8 se muestran las dependencias de las brechas energéticas de las soluciones sólidas ternarias formadas a partir de los semiconductores binarios GaP, GaAs, InAs y InP. Hacemos notar que las curvas mostradas en dicha figura muestran las dependencias $E_g-\alpha$ reales, en contraste con la figura 3.7 en donde dichas dependencias se

Tabla 3.3 Algunas aleaciones ternarias y cuaternarias cuyo parámetro de red puede ser acoplado al de los semiconductores binarios III-V empleados como sustratos.

Substrato	GaAs	InP	GaSb	InAs	GaP
Aleaciones isoperiódicas	AlGaAs	GaInAsP	AlGaAsSb	GaInAsSb	AlGaP
	AlGaInP	AlGaAsSb	GaInAsSb	AlGaAsSb	
	GaInAsP	GaInAsSb	InAsSbP	InAsSbP	

representan por líneas rectas, como se apuntó con anterioridad. En la figura 3.8, las secciones isoperiódicas del sistema $Ga_x In_{1-x} As_y P_{1-y}$ con GaAs y InP corresponden, de manera respectiva, a los segmentos verticales $GaAs-Ga_{0.51} In_{0.49} P$ y $InP-Ga_{0.51} In_{0.49} As$. La primera de estas secciones cubre el rango espectral 0.65-0.87 μm y nos permite construir heteroláseres de inyección sobre sustratos de GaAs. De manera específica, es posible

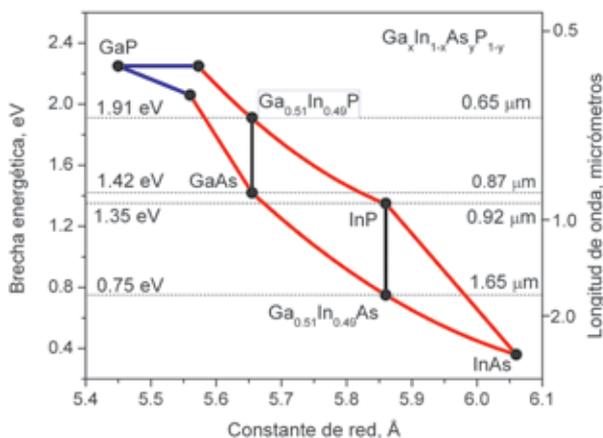


Fig.3.8 Secciones isoperiódicas de la solución sólida cuaternaria $Ga_x In_{1-x} As_y P_{1-y}$ con los sustratos GaAs y InP.

fabricar láseres con una longitud de onda de emisión alrededor de la línea de absorción del láser YAG:Nd, debido a lo cual pueden ser utilizados para bombear láseres de estado sólido. La segunda sección isoperiódica (InP) cubre el rango espectral de 0.92-1.65 μm , el cual es muy importante para los sistemas de comunicación por fibra óptica, pues en esta región espectral las fibras ópticas modernas alcanzan un mínimo de pérdidas por transmisión, además de que es posible obtener una dispersión igual a cero.

La tecnología para la construcción de láseres de alta potencia con base en los sistemas GaInAsP/GaAs y GaInAsP/InP ha sido desarrollada con éxito en el Instituto Físico-Técnico A.F. Ioffe mencionado con anterioridad. Uno de los autores de la presente obra (V. Mishurnyi) ha hecho contribuciones, tanto en dicho Instituto como en la Universidad Autónoma de San Luís Potosí, al desarrollo de este tipo de láseres, empleando heteroestructuras con confinamientos electrónico y óptico separados.

Por último, se mencionarán las soluciones sólidas, GaInAsSb y AlGaAsSb, las cuales tienen importantes aplicaciones prácticas. Con base en estos materiales, que tienen secciones isoperiódicas con el GaSb y con el InAs, se pueden construir dispositivos optoelectrónicos tales como láseres, fotodiodos y fotodetectores, que operan en el rango de longitudes de onda $\lambda > 2 \mu\text{m}$. Estos dispositivos son utilizados para el monitoreo de la contaminación ambiental.

3.7 Conclusiones

Como conclusión del presente capítulo, podemos señalar que en la actualidad, dada su gran importancia tecnológica, diferentes sistemas de soluciones sólidas de los grupos III-V son estudiados en compañías y laboratorios de investigación de todo el mundo. Con base en estos materiales y con la ayuda de diversas tecnologías de crecimiento epitaxial de cristales -por haces moleculares, a partir de una fase de vapor o a partir de una fase líquida- se ha logrado obtener

la más diversa gama de heteroestructuras semiconductoras que a su vez son la base de una gran variedad de dispositivos, en particular los que se emplean en aplicaciones optoelectrónicas.

CAPÍTULO 4

Equilibrios de fase y procesos de cristalización

4.1 Introducción

Como se discutió en los capítulos anteriores, los láseres de semiconductor hoy en día, al igual que otros dispositivos optoelectrónicos, se construyen con base en heteroestructuras formadas por varias capas de materiales semiconductores de diferente composición química. Típicamente un láser de semiconductor está constituido por 5-8 capas cuyos espesores varían desde decenas de nanómetros hasta algunos micrómetros, aunque en algunos casos la región activa de un láser puede llegar a tener un espesor que no sobrepasa algunas decenas de ángstroms. Existen también láseres de semiconductor con estructuras sensiblemente más complicadas y que en ocasiones llegan a contener cientos de capas de varias composiciones químicas.

Las estructuras láser de capas múltiples se preparan con base en diferentes soluciones sólidas cuyo parámetro de red depende de su composición. Por otro lado, como se mencionó en el capítulo anterior, los parámetros de red de dichas soluciones deben ser ajustados cuidadosamente a fin de minimizar la formación de defectos interfaciales que podrían afectar negativamente las propiedades luminiscentes de la heteroestructura láser. De este modo, durante el crecimiento epitaxial de una heteroestructura se deben controlar de manera precisa las composiciones de las diferentes capas que la forman.

4.2 Transiciones de fase

Antes de discutir los procesos de crecimiento de las heteroestructuras de muchas capas, consideremos el concepto de transición de fase. En la naturaleza existen cuatro tipos de fases: sólida, líquida, gaseosa y de plasma. En el crecimiento de cristales, sin embargo, sólo se consideran las tres primeras. Es conocido que en determinadas condiciones externas pueden coexistir dos o tres fases. Tomemos como ejemplo un vaso con agua a una temperatura de 80°C a la que se le han añadido cristales de sal común (cloruro de sodio o NaCl). Una vez colocados en el agua, los cristales de NaCl se disolverán poco a poco hasta que la solución se sature con sal. El concepto de solución saturada está relacionado con el así llamado límite de solubilidad. Supongamos para mayor precisión que el límite de solubilidad del cloruro de sodio en un cierto volumen de agua a la temperatura de 80°C es de 80 gr. Esto significa que si añadimos 80 gr de sal al agua, éstos se disolverán completamente si esperamos por el tiempo suficiente. Si ahora agregamos una cantidad adicional de sal, ésta ya no se disolverá, aun siendo muy pequeña. El término solución saturada se aplica entonces a una solución que alcanzó la solubilidad límite a una determinada temperatura.

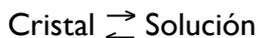
Consideremos una solución de agua a la que se ha añadido una cantidad de sal más allá del límite de saturación. Tendemos en este caso dos fases en equilibrio: la solución saturada (fase líquida) y los cristales de sal sin disolver (fase sólida). Si mantenemos constantes las condiciones externas (temperatura y presión), las dos fases coexistirán indefinidamente. A este estado, alcanzado entre dos o más fases en condiciones externas constantes -de temperatura, presión, concentración de los componentes en las fases, etc.- se le llama estado de equilibrio.

El estado de equilibrio de nuestra solución se perderá si cambiamos la temperatura del sistema. Esto es debido a que la solubilidad límite de la sal en agua se incrementa con la temperatura. De este modo, al aumentar la temperatura del agua parte del cristal de NaCl se disuelve hasta que el agua alcanza el nuevo límite de solubilidad. La curva que

describe la dependencia de la solubilidad límite con la temperatura se le conoce como curva de solubilidad. A partir de la misma es posible determinar cual será en nuestro ejemplo la concentración de sal en el agua cuando se alcance el equilibrio.

El ejemplo anterior nos ilustra el equilibrio entre una fase sólida y una solución saturada. Otro ejemplo de equilibrio entre fases sólida y líquida es el sistema agua-hielo a presión atmosférica y a una temperatura de 0°C. Es igualmente posible que una fase sólida y una gaseosa, o una líquida y una gaseosa coexistan en equilibrio. En todos estos casos se mantendrá el equilibrio entre las fases en la medida que la temperatura se mantenga constante; de otra manera, el equilibrio se perderá.

Al cambiar la temperatura de un sistema de dos fases en equilibrio se produce la transformación de una fase en otra. Esto lo podemos indicar de la siguiente manera:



En la expresión anterior, al elevar la temperatura se produce una disolución de la fase sólida, aplicándose la flecha que apunta a la derecha. Si por el contrario se reduce la temperatura, la flecha que se aplica es la dirigida a la izquierda, ya que en este caso parte de la sal disuelta en el agua es expulsada de la misma precipitándose en forma de cristales de NaCl.

En la tabla 4.1 se listan las posibles transformaciones de fase en un sistema de dos fases, juntamente con los nombres que reciben. Notamos que una transición de una fase sólida a una líquida recibe el nombre de fusión o disolución dependiendo de cómo ocurre el cambio de estado. Si un cristal es llevado hasta su temperatura de fusión y a partir de ahí el material sólido se convierte gradualmente en líquido inyectándole calor, se habla de un cambio de fase por fusión. Si, en contraste, se tiene el mismo cristal en equilibrio a una determinada temperatura con una solución saturada con el mismo material del cristal y en un determinado momento se incrementa

Tabla 4.1. Nomenclatura de las transiciones de fase.

Transición de fase	Nombre recibido
Fase gaseosa → Fase líquida	Condensación
Fase líquida → Fase gaseosa	Evaporación
Fase gaseosa → Fase sólida	Cristalización
Fase sólida → Fase gaseosa	Sublimación
Fase líquida → Fase sólida	Cristalización
Fase sólida → Fase líquida	Fusión o disolución

dicha temperatura, parte del material en la fase sólida se incorporará a la fase líquida por un proceso de disolución. Notamos que la temperatura de disolución del cristal es inferior a su temperatura de fusión.

4.3 Proceso de cristalización por capas

Los procesos tecnológicos involucrados en la fabricación de heteroestructuras de muchas capas corresponden a procesos de cristalización. Existen varios modelos físicos que ilustran y explican estos procesos. Vamos a estudiar uno de estos modelos, el cual recibe el nombre de *modelo de crecimiento de cristales por capas*. En la figura 4.1 se muestra esquemáticamente la superficie del cristal durante el crecimiento, representando a los átomos por cubos. Un modelo de crecimiento por capas supone que cada átomo del cristal tiene 6 enlaces que se denotarán, cada uno de ellos, por el símbolo ϕ , correspondiendo un enlace a cada lado del cubo tal como se muestra esquemáticamente en la parte superior derecha de la figura 4.1. Por otro lado, como se deduce de cálculos teóricos y de observaciones experimentales, la superficie de un cristal real no es plano sino que tiene una cierta rugosidad o relieve. Este relieve se representa en la figura 4.1 en forma de

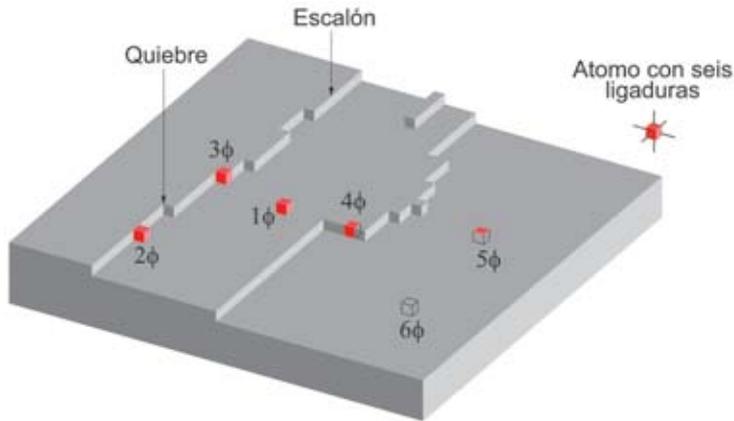


Fig. 4.1 Crecimiento epitaxial por capas y energías de ligadura de átomos sobre la superficie de un cristal. Los átomos se representan por cubos. Se muestran átomos en varias posiciones sobre la superficie o en el interior del cristal. 6ϕ : en el interior del cristal; 5ϕ : inmediatamente por debajo de la superficie; 4ϕ : integrado a un escalón; 3ϕ : en un quiebre de un escalón; 2ϕ : adyacente a un escalón; 1ϕ : sobre la superficie.

escalones con una altura de un átomo. Los escalones tienen además una cantidad considerable de quiebres o discontinuidades como se ilustra en la figura 4.1. De cálculos teóricos se deduce que en el momento de estar creciendo en los cristales semiconductores típicos surge un quiebre cada 5-10 distancias interatómicas a lo largo del escalón.

Notamos de la figura 4.1 que un átomo que se encuentra en el interior del cristal (átomo marcado con 6ϕ en la figura 4.1) empleará todos sus seis enlaces para ligarse con sus átomos vecinos y estará de este modo enlazado con la red cristalina de manera óptima. Por su parte, un átomo colocado inmediatamente debajo de la superficie plana del cristal (marcado con 5ϕ) solamente empleará cinco enlaces para este propósito, dejando uno libre y por tanto estará más débilmente ligado. Un átomo que forma parte de un escalón de la superficie (marcado con 4ϕ) se ligará solamente con cuatro enlaces,

mientras que si se localiza en el quiebre de un escalón (átomo marcado con 3ϕ) lo hará únicamente con tres. Finalmente, podemos ver que los átomos más débilmente ligados al cristal son aquellos que se encuentran adyacentes a un escalón (marcado con 2ϕ) y que cuentan con cuatro enlaces libres, y los localizados sobre una superficie plana (marcado con 1ϕ) con cinco enlaces libres. Tenemos entonces que el átomo sobre la superficie de cristal que está mejor sujeto a la misma es el que se encuentra en el quiebre de un escalón.

Imaginemos que el cristal se encuentra en equilibrio con una fase líquida o gaseosa y que en un momento dado, debido a una disminución de su temperatura, se crean las condiciones necesarias para que ocurra una cristalización. Como un primer paso en este proceso, algunos átomos o partículas que se encuentran en la fase líquida o gaseosa se adsorben (se adhieren) a la superficie del cristal con un enlace ϕ . Se sabe que la energía de una partícula depende de la temperatura. De este modo, ya que el crecimiento de cristales se realiza a una temperatura considerablemente alta, las partículas adsorbidas tienen una energía relativamente grande, lo que les permite desplazarse por la superficie del cristal. Cálculos teóricos indican que la partícula adsorbida a una temperatura típica de crecimiento adquiere la energía suficiente para desplazarse por la superficie del cristal por algunos cientos de periodos de la red cristalina.

Durante su desplazamiento, la partícula puede desprenderse de la superficie del cristal y regresar a la fase líquida o gaseosa de donde provino. Esta posibilidad, sin embargo, se reduce si la partícula al desplazarse encuentra un escalón. En tal caso, se liga a la superficie con dos enlaces ϕ de modo que su adherencia a la misma se incrementa. Una vez en esta posición, si la partícula tiene suficiente energía puede continuar desplazándose, sólo que ahora lo hace de manera preferencial a lo largo del escalón manteniendo todos sus enlaces con el cristal. Esta situación se prolongará hasta que la partícula encuentre un punto de quiebre en el escalón, con lo que su número de enlaces con el cristal se incrementará de dos a tres. Después de esto, la posibilidad de que la partícula se mueva se verá reducida de manera considerable por no

tener la energía suficiente para romper los enlaces. Así, a medida que se reúnen más y más partículas en los puntos de quiebre de los escalones, éstos últimos se desplazan a lo largo de la superficie del cristal hasta formar una nueva capa. De este modo, el cristal crece capa por capa y de aquí que a este modelo se le llame *modelo de crecimiento por capas*. Se hace notar que existen otros mecanismos de crecimiento de cristales. Mencionamos, sin embargo, solamente el crecimiento por capas en virtud de que es el más simple, así como el más gráfico

4.4 Cristalizaciones volumétrica y epitaxial

Imaginemos que calentamos cristales de cloruro de sodio por arriba de su temperatura de fusión $T_f \approx 800^\circ\text{C}$, hasta que la sal se funda totalmente. Si ahora disminuimos la temperatura hasta alcanzar el punto de fusión de la sal, en un momento dado y de manera espontánea se formará en el seno de la fase líquida un primer cristal de NaCl. Una vez formado, este cristal comenzará a crecer siguiendo, por ejemplo, el mecanismo de crecimiento por capas descrito anteriormente. El proceso de crecimiento terminará cuando toda la fase líquida se haya cristalizado.

Un proceso de cristalización como el descrito es eficiente desde un punto de vista de costos de fabricación, ya que todo el volumen de la fase líquida se incorpora al cristal. Estos métodos se emplean cuando se tiene la necesidad de obtener cristales de gran volumen y por esta razón reciben el nombre de *métodos volumétricos de cristalización*. Aunque adecuados para obtener cristales masivos, los métodos volumétricos de cristalización presentan algunos problemas. Uno de éstos es el asociado a las altas temperaturas necesarias para fundir el material a cristalizar, lo que lleva a una posible interacción de la fase líquida con el material del recipiente que la contiene con la consecuente contaminación del material fundido. Así, utilizando los métodos volumétricos de cristalización no siempre es posible obtener cristales lo suficientemente puros.

En contraste con el método de crecimiento volumétrico de un cristal de sal, en el que es necesario llevar el cloruro de sodio hasta su temperatura de fusión cercano a los 800 °C, la técnica de crecimiento a partir de una solución saturada discutido líneas arriba involucra temperaturas relativamente bajas. En efecto, consideremos el siguiente experimento. Supongamos, como lo hicimos anteriormente, que la solubilidad límite de la sal en el vaso de agua a una temperatura de 80 °C es igual a 80 gr, mientras que a una temperatura de 100 °C es de 100 gr. Introduzcamos en un vaso con agua calentada hasta una temperatura de 100 °C, un cristal de NaCl con un peso de 150 gr. Después de algún tiempo, 2/3 de este cristal, es decir 100 gr, se disolverán en el agua, con lo que la solución alcanzará la saturación. Si se mantiene una temperatura constante y se evita la evaporación del agua encerrando todo el sistema en un depósito hermético, la solución saturada y los 50 gr remanentes del cristal parcialmente disuelto se encontrarán en equilibrio. Si ahora enfriamos el sistema hasta una temperatura de 80 °C la solubilidad decrecerá a 80 gr y los 20 gr de sal de exceso en la solución se incorporarán la fase sólida depositándose sobre el cristal de NaCl. De este modo, en nuestro experimento de crecimiento de cristales de cloruro de sodio se emplearon temperaturas del orden de los 100 °C, en contraste con los cerca de 800 °C que son necesarios para el mismo propósito empleando el método de crecimiento volumétrico.

Como una variante del experimento anterior, preparemos una solución saturada de sal en agua a una temperatura de 100 °C y en seguida, manteniendo constante la temperatura, introduzcamos cuidadosamente en dicha solución un cristal de cloruro de sodio al que llamaremos sustrato. Puesto que la solución está saturada, el cristal no se disolverá, manteniéndose por el contrario en equilibrio con la solución. Si ahora se enfría todo el sistema, se depositará sobre el sustrato una película cristalina de NaCl.

Al crecer cristales a partir de soluciones saturadas, el volumen del material crecido será menor que el que se obtendría si se utilizara todo el material disuelto. Supongamos a manera de ejemplo que la

solubilidad de la sal en un vaso de agua a una temperatura ambiente de 20 °C es de 20 gr. Si se enfría una solución saturada a 100 °C hasta la temperatura ambiente, se cristalizarán entonces 80 gr de sal. Los 20 gr de sal restantes quedarán disueltos en agua al final del proceso y no se incorporarán al cristal crecido.

De este modo, al contrario de lo que sucede cuando se obtiene un cristal a partir de un material fundido, en el proceso de cristalización a partir de una solución saturada no se emplea todo el material disuelto. Esto implica que los métodos de crecimiento de cristales basados en soluciones saturadas son menos efectivos desde el punto de vista productivo que los métodos volumétricos. Es importante señalar, no obstante, que el solvente empleado en las técnicas de solución saturada no se consume durante el proceso y es susceptible de ser utilizado nuevamente, lo que contribuye a disminuir costos de fabricación.

Por otro lado, el que las temperaturas involucradas en los procesos de solución saturada sean considerablemente menores que aquellas empleadas con las técnicas volumétricas, simplifica considerablemente los problemas técnicos relacionados con el diseño y construcción del equipo necesario para el crecimiento de cristales. Lo más importante, sin embargo, es que los cristales crecidos a bajas temperaturas tienen menos defectos y son más puros.

El crecimiento de un cristal de NaCl sobre la superficie de un sustrato estudiado en los párrafos anteriores, recibe el nombre de *Epitaxia por fase líquida*, en virtud de que el proceso de cristalización ocurre a partir de una fase líquida. Si la cristalización ocurriera a partir de una fase gaseosa se hablaría de *Epitaxia por fase gaseosa*.

4.5 Conclusiones

En este capítulo se consideraron las fases sólida, líquida y gaseosa en las que puede existir un sistema. Se discutió, además, el equilibrio entre dos de estas fases bajo condiciones externas determinadas. Tratamos en detalle el equilibrio entre una fase sólida de cloruro de sodio y una

fase líquida formada por una solución de agua con NaCl, así como el proceso crecimiento de cristales de NaCl que ocurre al disminuir la temperatura a partir del estado de equilibrio. Este proceso de crecimiento nos ejemplifica la técnica de epitaxia por fase líquida, la cual es empleada para el crecimiento de películas cristalinas de materiales semiconductores. En el siguiente capítulo se discutirá en detalle esta técnica, así como otros métodos de crecimiento epitaxial que son esenciales en la tecnología actual de los láseres de semiconductor.

CAPÍTULO 5

Tecnologías epitaxiales de crecimiento de cristales

5.1 Introducción

El término epitaxia (del griego *epi*: sobre, *taxis*: orden) apareció por primera vez hace aproximadamente 50 años. Este término se refiere al depósito o crecimiento de una película delgada sobre un substrato cristalino siguiendo el ordenamiento atómico del mismo. Distinguiamos dos casos: homoepitaxia y heteroepitaxia. En el primer caso la película y el substrato tienen una misma composición química, mientras que en el segundo ambas composiciones son diferentes. A los procesos de cristalización de películas delgadas sobre la superficie de un substrato se les llama en forma genérica *métodos de crecimiento epitaxial de cristales*.

Las técnicas de crecimiento epitaxial ocupan un lugar muy importante en la tecnología de los dispositivos semiconductores, desde los circuitos integrados hasta los dispositivos optoelectrónicos. En el caso particular de los láseres de semiconductor, su desarrollo estuvo íntimamente ligado a los avances en las técnicas de crecimiento epitaxial. Dependiendo de la forma de transportar el material a crecer desde la fuente hasta el substrato, los procesos epitaxiales se dividen en tres grupos:

- Epitaxia por Fase Líquida.
- Epitaxia por Fase Gaseosa.
- Epitaxia por Haces Moleculares.

Hoy en día estos tres procesos son utilizados para la fabricación de los más diversos dispositivos y estructuras semiconductoras, así como en la búsqueda de nuevos materiales semiconductores. En lo que sigue se discutirán los aspectos básicos de cada una de estas tres técnicas de crecimiento epitaxial.

5.2 Epitaxia por Fase Líquida

La epitaxia por fase líquida (LPE, por sus siglas en inglés) es una técnica para el crecimiento de películas cristalinas a partir de una fase líquida. En el capítulo anterior se describieron los puntos esenciales de esta técnica, tomando como ejemplo el crecimiento de cristales de cloruro de sodio a partir de una solución de agua saturada con NaCl. En el caso de los semiconductores, la cristalización de las capas epitaxiales por LPE tiene lugar a partir de una solución líquida saturada con el semiconductor a depositar. Dos solventes muy comúnmente empleados para el crecimiento de los semiconductores III-V y sus aleaciones son el galio y el indio. Estos elementos tienen temperaturas de fusión bajas y son, de manera ventajosa, parte de la composición de las películas crecidas. Además, tanto el galio como el indio tienen una presión de vapor muy baja en el intervalo de temperatura que generalmente es utilizado en el proceso de crecimiento. Como ventajas adicionales, estos dos elementos se pueden obtener muy puros y a las temperaturas usuales de crecimiento ninguno de los dos interaccionan con el grafito, material del cual están generalmente hechos los contenedores de las fases líquidas.

Para el empleo de la técnica LPE es esencial conocer los diagramas de fase que determinan la relación entre las composiciones de las fases líquida y sólida. Estos diagramas deben conocerse a todas las temperaturas de interés y para todos los semiconductores empleados en el crecimiento. Dadas las condiciones adecuadas, la fase sólida de un compuesto binario III-V -formado por igual proporción de átomos III y V- podrá establecer un equilibrio termodinámico a temperatura

dada con la fase líquida formada por átomos del grupo V disueltos en el metal del grupo III. La composición de la fase líquida es dependiente de la temperatura a la que se encuentre, debido a que la solubilidad del elemento del grupo V disminuye con dicha temperatura. Notamos, por otro lado, que la concentración del elemento del V grupo en la fase líquida es de sólo algunas unidades porcentuales y que a las temperaturas relativamente bajas que se utilizan en los crecimientos por fase líquida, las soluciones se encuentran bien homogeneizadas.

El método de crecimiento epitaxial por LPE fue utilizado por primera vez en 1963 por H. Nelson para el crecimiento de uniones p-n de GaAs. Consideraremos a continuación el arreglo y el procedimiento experimental empleado por este investigador.

Es primeramente necesario preparar una solución saturada de arsénico en galio y la manera más sencilla de hacerlo es poniendo en contacto galio fundido con un cristal de GaAs a una determinada temperatura. El contenido de arsénico del cristal de GaAs puede ser mayor que el que fija el diagrama de fases para saturar la solución de arsénico en galio, ya que sólo los átomos de arsénico estrictamente necesarios para alcanzar la saturación de dicha solución son finalmente incorporados a la misma. Una vez saturada, la solución se pone en contacto con el substrato de GaAs y se inicia un lento descenso de temperatura, lo que da inicio al crecimiento epitaxial de GaAs sobre el substrato. Se finaliza el crecimiento retirando la solución de la superficie del substrato.

Para realizar las operaciones anteriores, Nelson construyó un reactor epitaxial capaz de girar alrededor de un eje horizontal, tal como se muestra en la figura 5.1. El reactor está compuesto de un tubo de cuarzo y de un horno tubular en el que se coloca dicho tubo. Para proceder al crecimiento, se introduce primeramente al tubo del reactor el crisol que contiene tanto a la solución saturada como al substrato de GaAs. Durante la introducción el reactor se mantiene inclinado a un cierto ángulo, de modo que la solución y el substrato permanezcan separados, tal como se muestra en la figura 5.1. En esta posición se inicia el calentamiento hasta alcanzar la temperatura de

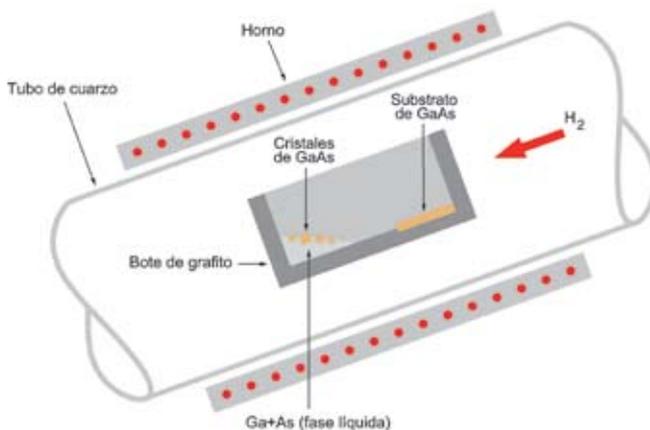


Fig.5.1 Reactor para el crecimiento epitaxial de GaAs por LPE, según el arreglo empleado por H. Nelson. El sistema mostrado fue uno de los primeros diseños utilizados para este tipo de crecimientos.

crecimiento. Después de esperar el tiempo necesario para que la solución de Ga se sature con As, el reactor se inclina en sentido opuesto de tal modo que la solución cubra al substrato. A partir de este punto se disminuye la temperatura de manera lenta y progresiva, lo que da inicio al crecimiento de una capa epitaxial de GaAs sobre el substrato de GaAs. Una vez que se ha obtenido el espesor epitaxial deseado, el proceso de crecimiento se interrumpe regresando el horno a su posición inicial, retirando de esta manera la solución de la superficie del substrato.

El proceso de crecimiento descrito en el párrafo anterior se lleva a cabo en una atmósfera de hidrógeno. Para esto, se hace circular por el tubo del reactor un flujo de hidrógeno de muy alta pureza durante todo el proceso de crecimiento. El hidrógeno previene la formación de películas de óxido en la superficie de la solución de galio, así como en las superficies del substrato y del material utilizado para saturar la fase líquida. La existencia de estas películas de óxido dificulta el

contacto entre el sustrato y la solución; es decir, la solución no *moja* al sustrato.

El diseño anteriormente descrito sirve para realizar crecimientos epitaxiales de una sola capa. Para la obtención de estructuras más complicadas, que contengan varias capas con diferentes composiciones químicas, es necesario utilizar un montaje más elaborado. En la figura 5.2 se ilustra un diseño de un crisol muy ampliamente usado para este propósito. Este crisol consiste de un bote de grafito con una estructura tipo panal compuesto de dos partes: una parte fija que contiene las diferentes soluciones de crecimiento en los orificios del “panal” y una reglilla móvil que aloja al sustrato en una cavidad. Manipulando la posición de la reglilla móvil es posible poner en contacto al sustrato con cada una de las fases líquidas. Si se disminuye la temperatura al estar en contacto una de las fases líquidas con el sustrato, sobre la superficie del mismo crecerá una película epitaxial con una composición que dependerá de la temperatura y de la composición de la fase líquida.

Como se mencionó anteriormente, para llevar a cabo un crecimiento epitaxial con una composición determinada por LPE es necesario conocer los diagramas de fase correspondientes, ternarios o

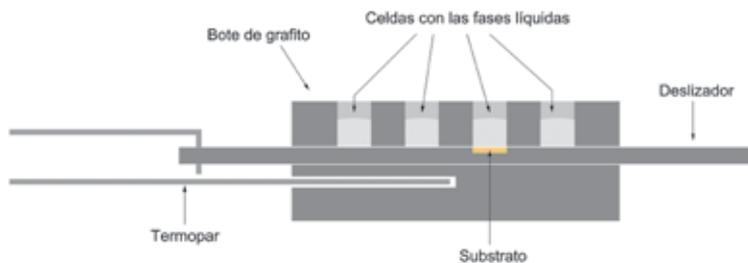


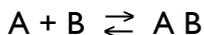
Fig.5.2 Bote de grafito para el crecimiento de estructuras de capas múltiples por epitaxia en fase líquida. Los sistemas modernos de LPE emplean este tipo de botes. El número de compartimientos para alojar las soluciones líquidas depende del número de capas de diferente composición química que se pretende crecer.

cuaternarios según sea el caso. Estos diagramas han sido estudiados en detalle, tanto experimental como teóricamente, para casi todos los sistemas de materiales que representan algún interés práctico. A partir de dichos diagramas, que establecen la relación de composiciones entre las fases líquida y sólida a diferentes temperaturas, se pueden controlar las propiedades físicas de la capa a crecer. El espesor de dicha capa se puede regular variando el intervalo de temperatura en el que se realiza el proceso de cristalización, o bien variando el volumen de la fase líquida.

Cabe señalar que la epitaxia por fase líquida desde un punto de vista técnico es bastante simple, lo que permite su implementación con un costo relativamente reducido. Notamos además que, no obstante su simplicidad técnica, la epitaxia por fase líquida revolucionó en su momento el campo del crecimiento de cristales semiconductores, lo que llevó a mejorar los parámetros de dispositivos ya existentes y a la creación de otros nuevos basados en heteroestructuras de capas múltiples, como es el caso de los heteroláseres.

5.3 Epitaxia por fase gaseosa

La epitaxia por fase gaseosa se conoce también con el nombre de método de reacción por transporte químico. Una gran cantidad de las más diversas reacciones químicas son reversibles. Esto significa, que si ocurre una reacción entre las sustancias A y B formando un compuesto AB, entonces también puede tener lugar el proceso inverso, es decir, la descomposición de AB en sus componentes originales A y B. Tal reacción reversible se escribe de la siguiente manera:



La reacción de izquierda a derecha, indicada por la flecha apuntando hacia la derecha, se denomina directa, mientras que la reacción de derecha a izquierda recibe el nombre de inversa. Si las reacciones directa e inversa ocurren a una misma velocidad, entonces el sistema alcanzará

un estado de equilibrio químico si esperamos un tiempo suficientemente largo. En este estado, las concentraciones de los componentes que participan en el proceso son constantes.

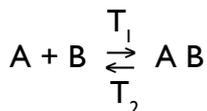
Los parámetros fundamentales que inciden en el equilibrio químico de un sistema son: la concentración de los componentes, la temperatura y la presión. Para entender esto, suponga que aumentamos la concentración de las sustancias iniciales A y B en el sistema químico en equilibrio descrito arriba. Como resultado de este incremento, la velocidad con que se forma el producto AB también aumenta a fin de mantener el equilibrio. De este modo, el sistema desplaza su equilibrio de izquierda a derecha, es decir, hacia la formación del producto de reacción AB.

Imaginemos ahora que la descomposición del producto AB en el sistema considerado, o sea la reacción inversa, tiene lugar con absorción de calor -reacción endotérmica. Suponga que el sistema está inicialmente en un estado de equilibrio a una temperatura dada y que en un determinado momento este equilibrio se pierde por un incremento en dicha temperatura. Como resultado de esto, la reacción inversa incrementa su velocidad, de modo que el sistema desplaza su equilibrio hacia la formación de las sustancias originales, es decir, de derecha a izquierda. De la misma manera, un aumento en la presión en el sistema desplaza el equilibrio hacia las sustancias que se forman cuando se reduce la distancia entre las moléculas de la fase gaseosa.

De lo anterior se desprende que un cambio en los parámetros externos de un sistema en equilibrio químico puede desplazar su equilibrio, ya sea hacia la formación del producto de reacción, o bien hacia la formación de las sustancias originales.

Las reacciones químicas se pueden clasificar por la fase de los componentes reactantes. Si todos estos componentes se encuentran en una sola fase la reacción se denomina homogénea; si por el contrario se encuentran en diferentes fases, la reacción recibe el nombre de heterogénea. Supongamos que en una reacción heterogénea el componente B se encuentra en la fase sólida, mientras que la componente A y el producto AB están en una fase gaseosa. Si las velocidades de

las reacciones directa e inversa alcanzan sus máximos valores a diferentes temperaturas, la reacción se puede representar de la siguiente manera:



En el caso en que la diferencia entre las temperaturas T_1 y T_2 no sea muy grande, lo que es frecuente, es posible obtener un transporte químico. En la figura 5.3 se ilustra este proceso. Suponga un tubo de vidrio dentro del cual se hace pasar un flujo de la componente A en fase gaseosa. En la región de la fuente se encuentra una cierta cantidad de la componente B en una fase sólida a una temperatura tal que favorece la formación de producto AB, también en fase gaseosa. Este producto es transportado hacia la región de condensación, en donde la temperatura favorece la disociación del producto AB en sus componentes originales, A en fase gaseosa y B en fase sólida. De este modo, se produce un transporte de la componente B desde la región de la fuente a la región de condensación. Podemos entonces afirmar que las reacciones químicas de transporte son procesos heterogéneos reversibles, con la participación de una fase gaseosa que conduce a la formación de un producto gaseoso intermedio,

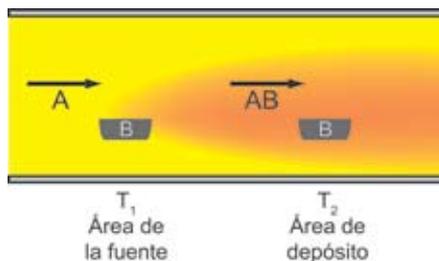


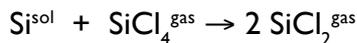
Fig.5.3 Ilustración del mecanismo de transporte de materia aprovechando las reacciones químicas directas e inversas que ocurren a diferentes temperaturas.

con ayuda de la cual se puede realizar el transporte de una sustancia entre dos regiones reactivas.

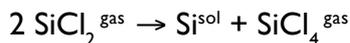
Naturalmente que la sustancia B puede contener diferentes impurezas. Es frecuente que el producto AB de la reacción se evapore con más facilidad que las impurezas, de tal manera que, como resultado de la reacción química de transporte, en la región de condensación se depositará el componente B en un estado más puro que aquel en la región de la fuente. En otras palabras, con ayuda de las reacciones químicas de transporte es factible purificar una sustancia.

Si durante el proceso de transporte en una reacción química se introduce un sustrato en la zona de condensación, entonces parte del componente transportado B se depositará en dicho sustrato. Una vez absorbidas en el sustrato, en ciertas condiciones las partículas de B se organizan de tal manera que generan un crecimiento epitaxial.

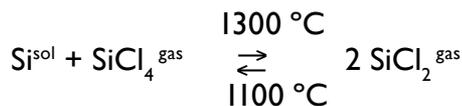
Como ejemplo de lo anterior, analicemos un proceso muy conocido de transporte de silicio, el cual utiliza como medio gaseoso el tetracloruro de silicio (SiCl_4). A la temperatura de $1300\text{ }^\circ\text{C}$, el SiCl_4 reacciona con el silicio formando cloruro de silicio como sigue:



Una vez formado el producto gaseoso de SiCl_2 , se disocia en una zona más fría a una temperatura de $1100\text{ }^\circ\text{C}$ liberando silicio sólido según la reacción:



En general esta es una reacción heterogénea reversible que de manera más completa se puede escribir como sigue:



Con frecuencia, en la práctica para realizar este tipo de transporte heterogéneo se utiliza un gas inerte como medio de transporte de los componentes reactantes.

Para el crecimiento de películas epitaxiales con base en los compuestos III-V se utilizan técnicas similares a las descritas líneas arriba con algunas modificaciones. Consideraremos una de estas técnicas, la cual recibe el nombre de método hidroclórico por razones que se harán claras más adelante, tomando como ejemplo el crecimiento de GaAs. En este caso, el gas que más comúnmente se emplea como fuente de arsénico es el tricloruro de arsénico (AsCl_3); como fuente de galio se usa galio metálico y como gas de transporte el hidrógeno. El AsCl_3 es un líquido a una temperatura en el rango entre -8.5°C y 130°C .

El proceso de crecimiento de cristales de GaAs se ilustra en la figura 5.4, en donde podemos ver que la región de la fuente está a una temperatura de $800\text{-}850^\circ\text{C}$, mientras que la región de depósito se mantiene a una temperatura de $700\text{-}750^\circ\text{C}$. En la región de la fuente se tiene galio fundido y en la región receptora se coloca el sustrato de GaAs. Por facilidad de manejo el hidrógeno se introduce a través de tres canales: en el primero se inyecta mezclado con la impureza o

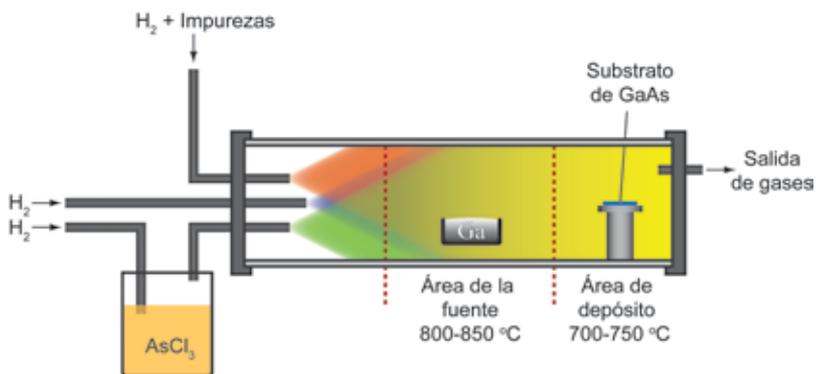
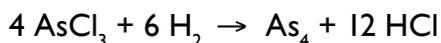
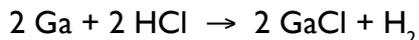


Fig.5.4 Método hidroclórico para el crecimiento de GaAs.

dopante que se quiere incorporar al cristal, a través del segundo canal se hace fluir hidrógeno puro y por el tercero se introduce hidrógeno saturado con AsCl_3 . En este último caso el hidrógeno se satura de AsCl_3 por medio de un burbujeador que se mantiene a temperatura constante. Esta temperatura determina la concentración de AsCl_3 en el gas. Una vez saturado con AsCl_3 , el hidrógeno llega a la región de la fuente, en donde se produce arsénico y ácido clorhídrico (HCl) de acuerdo a la reacción:

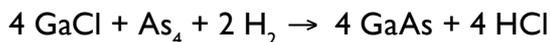


El arsénico producto de esta reacción se combina con el galio en la región de la fuente, formando pequeños cristales de GaAs en la superficie del galio. El proceso es más reproducible cuando se establece un equilibrio dinámico entre los cristales de GaAs en la fuente y el flujo de gas, pues de otro modo la relación Ga/As en dicho flujo es inestable. Una vez que se alcanza este equilibrio, el contenido de GaCl en el flujo de hidrógeno alcanza a su vez un equilibrio de acuerdo con la siguiente reacción:



en la que el GaCl -que también es un gas- sirve como vehículo de transporte para el galio.

En la región de depósito tiene lugar la siguiente reacción:



que conduce a la formación de GaAs. Esta técnica recibe el nombre de método hidrocórico por el papel que desempeña el ácido clorhídrico en la reacción.

Otra técnica por fase gaseosa para el crecimiento de cristales de semiconductores III-V, muy utilizada tanto en laboratorios de investigación como en la industria, es la que emplea hidruros como fuente

del elemento V. Para el crecimiento de GaAs, el hidruro empleado como fuente de arsénico es la arsina (AsH_3). La fuente de galio, al igual que en el caso del método hidroclórico anteriormente discutido, es el GaCl, que se forma al interactuar el HCl con el galio. Dada su alta toxicidad, la arsina se emplea en mezclas muy diluidas en hidrógeno, típicamente en el rango 1-5%. Tenemos entonces que en un proceso hidroclórico empleando hidruros, se introducen al reactor de manera separada los siguientes gases: $\text{AsH}_3 + \text{H}_2$, $\text{HCl} + \text{H}_2$ y H_2 ; además, en caso necesario, se agrega un dopante.

Naturalmente que si se pretende crecer una solución sólida ternaria o cuaternaria por un método de epitaxia por fase gaseosa será necesario introducir al reactor un mayor número de gases, tantos como demande la composición del cristal a crecer. Por ejemplo, para el crecimiento de las soluciones sólidas $\text{GaAs}_{1-x}\text{P}_x$ por el método hidroclórico se emplean mezclas de AsCl_3 y PCl_3 con hidrógeno, como fuentes de arsénico y fósforo, de manera respectiva. Variando el flujo del hidrógeno a través de cada burbujeador se puede controlar la composición de la solución sólida a crecer.

Las soluciones sólidas $\text{GaAs}_{1-x}\text{P}_x$ también pueden ser crecidas por el método hidroclórico empleando hidruros. Para esto, se inyectan al reactor, por canales independientes, mezclas de hidrógeno con AsH_3 y PH_3 que proporcionarán, de manera respectiva, el arsénico y el fósforo necesarios. Adicionalmente, se introduce una mezcla de hidrógeno y vapor de HCl, a fin de obtener GaCl para el transporte del galio. En la figura 5.5 se muestra de manera esquemática este proceso. Son necesarias tres regiones de temperatura: 1) la región de la fuente ($T=800-850$ °C), donde se forma el GaCl, 2) una región intermedia ($T=880-930$ °C) en donde se mezclan los gases inyectados y 3) la región del depósito ($T=750-830$ °C) donde se coloca el sustrato.

Los métodos hidroclóricos de transporte de gas, con o sin hidruros, se encuentran entre los primeros utilizados para el crecimiento epitaxial de materiales semiconductores. Hoy en día, sin embargo, la técnica más ampliamente empleada para este propósito, tanto en laboratorios de investigación como en la industria, es la técnica conocida

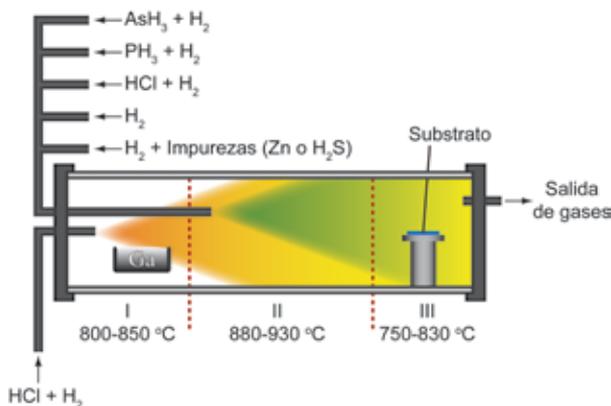
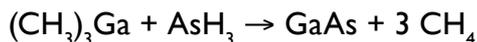


Fig.5.5 Esquema de un proceso de crecimiento de GaAsP por la técnica hidroclórica empleando hidruros.

como epitaxia por fase gaseosa empleando compuestos organo-metálicos (MOCVD, por sus siglas en inglés). La técnica MOCVD se basa en la descomposición térmica (pirólisis) de hidruros de elementos del grupo V y de compuestos organo-metálicos de elementos del grupo III, y por tanto involucra reacciones químicas irreversibles. Para el depósito de GaAs por la técnica MOCVD, se emplea arsina y un compuesto organo-metálico que contenga galio, tal como el (CH₃)₃Ga (trimetilo de galio). El proceso de pirólisis en este caso es el siguiente:



Aunque las reacciones de pirólisis con compuestos organo-metálicos se han usado por largo tiempo para obtener una gran variedad de recubrimientos metálicos, en la tecnología de materiales semiconductores la técnica MOCVD comenzó a utilizarse hasta fechas relativamente cercanas. Una de las ventajas que ofrece esta técnica sobre el método hidroclórico es la posibilidad de crecer compuestos y soluciones sólidas que contengan aluminio, lo que es difícil de realizar empleando la técnica hidroclórica debido a que los compuestos volátiles de aluminio, tales como el AlCl₃, son muy activos químicamente

y reaccionan con una gran cantidad de materiales, en particular con la superficie del tubo de cuarzo del reactor epitaxial en contacto con los gases reactantes.

Los materiales organo-metálicos más ampliamente usados para el crecimiento de GaAs son el trimetilio de galio, ya anteriormente mencionado, y el $(C_2H_5)_3Ga$ (trietilio de galio). Estos materiales pueden ser adquiridos en cilindros con salida especial para que el gas portador -hidrógeno de muy alta pureza- pueda ser burbujeado a través de los mismos. Los flujos gaseosos inyectados al reactor epitaxial son controlados por flujómetros electrónicos de alta precisión y , a fin de estabilizar la concentración del material organo-metálico en el flujo de hidrógeno, el burbujeador se mantiene a una temperatura fija en el rango 10-20 °C.

En la figura 5.6 se ilustra un montaje para el crecimiento de soluciones sólidas $Al_xGa_{1-x}As$ por MOCVD. El sistema emplea trimetilio de galio y trimetilio de aluminio como fuentes de galio y aluminio, en forma respectiva. El porta-substrato está hecho de grafito de alta pureza y es calentado por inducción a alta frecuencia. De este modo, las paredes de cuarzo del reactor se mantienen durante el crecimiento a una temperatura relativamente baja, lo que ayuda a minimizar la reacción química de dichas paredes con los gases en el interior. Además, con el objeto de obtener una mejor homogeneidad en las propiedades físicas de las películas epitaxiales, durante el proceso de crecimiento se hace girar el porta-substrato.

El crecimiento de heteroestructuras con capas de diferentes composiciones químicas se realiza manipulando la composición del flujo gaseoso inyectado al reactor. Debido a que los cambios tienen que hacerse de manera precisa, el control de los flujómetros electrónicos se lleva a cabo por computadora.

En la actualidad, la epitaxia por fase gaseosa por medio de compuestos organo-metálicos es uno de los principales métodos de crecimiento de cristales empleados en la tecnología de materiales y dispositivos semiconductores. Con base en esta técnica se fabrican heteroláseres que emiten en una amplia región del espectro electromagnético, desde el infrarrojo hasta el azul.

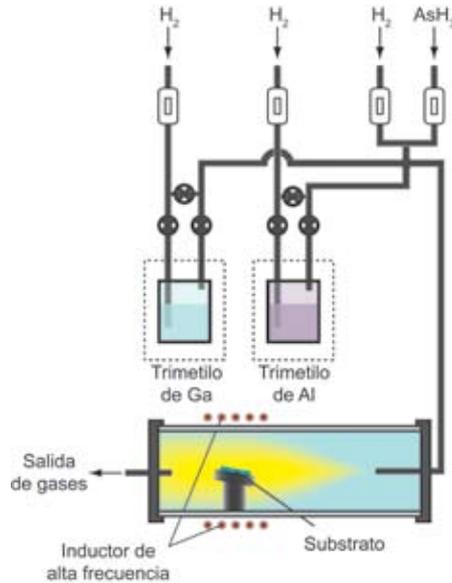


Fig.5.6 Esquema de un sistema de crecimiento de aleaciones de AlGaAs por medio de la técnica de epitaxia por fase gaseosa empleando compuestos organo-metálicos (MOCVD).

5.4 Epitaxia por haces moleculares

La epitaxia por haces moleculares (MBE, por sus siglas en inglés) es una técnica de crecimiento de películas epitaxiales semiconductoras a partir de haces moleculares generados por evaporación en celdas de efusión. La técnica MBE se desarrolló a partir del método de depósito de películas metálicas por evaporación en vacío, que ha sido ampliamente usado por un largo tiempo. Es bien conocido que es posible depositar metales por evaporación en un vacío en el rango 10^{-5} - 10^{-7} mm Hg. Para el depósito de películas semiconductoras, sin embargo, se requieren vacíos varios órdenes de magnitud más elevados. Dada esta circunstancia, el desarrollo de la técnica MBE

tuvo que esperar a que hubiese disponibilidad de equipo capaz de alcanzar niveles de vacío superiores a 10^{-10} mm Hg; esto ocurrió al inicio de los años 70.

En la figura 5.7 se muestran de manera esquemática los elementos esenciales de un sistema de MBE. El corazón del mismo es una cámara de crecimiento capaz de alcanzar un vacío superior a 10^{-10} mm Hg. En el centro de dicha cámara se coloca el sustrato epitaxial fijo a un porta-sustrato de molibdeno, mientras que su periferia se fijan las celdas de efusión, una por cada elemento a evaporar. Dichas celdas apuntan hacia el sustrato -por ejemplo, GaAs o InP- en el centro de la cámara, de modo que los flujos de moléculas generados por las celdas convergen sobre dicho sustrato en donde son adsorbidas.

Cuando hay necesidad de abrir a la atmósfera la cámara de crecimiento -para recargar las celdas de efusión o para alguna otra labor de mantenimiento-, las paredes de la misma absorben gases atmosféricos los cuales tienen que ser eliminados para reestablecer las condiciones de ultra alto vacío necesarias para el proceso de crecimiento epitaxial. Esto se logra calentando la cámara bajo condiciones de ultra alto vacío, por el tiempo necesario para obtener un nivel de vacío adecuado. Este proceso puede llevar días e incluso semanas, por lo que se hace necesario minimizar la frecuencia de exposición a la atmósfera de las superficies internas de la cámara de crecimiento. Con este propósito, los sistemas de MBE están provistos de una cámara de vacío anexa de volumen reducido, a través de la cual se introducen los sustratos epitaxiales sin perder el alto vacío de la cámara principal.

El proceso de crecimiento del cristal depende fuertemente de la temperatura del sustrato, parámetro que debe entonces controlarse con gran precisión. Esto se lleva a cabo con la ayuda de un termopar que se coloca en el interior del porta-sustrato. Las condiciones óptimas de cristalización ocurren cuando el sustrato tiene una temperatura entre 300 y 700 °C, dependiendo de los materiales que se pretendan crecer. Para obtener mayor homogeneidad en las capas crecidas, el porta-sustrato se hace girar durante el crecimiento.

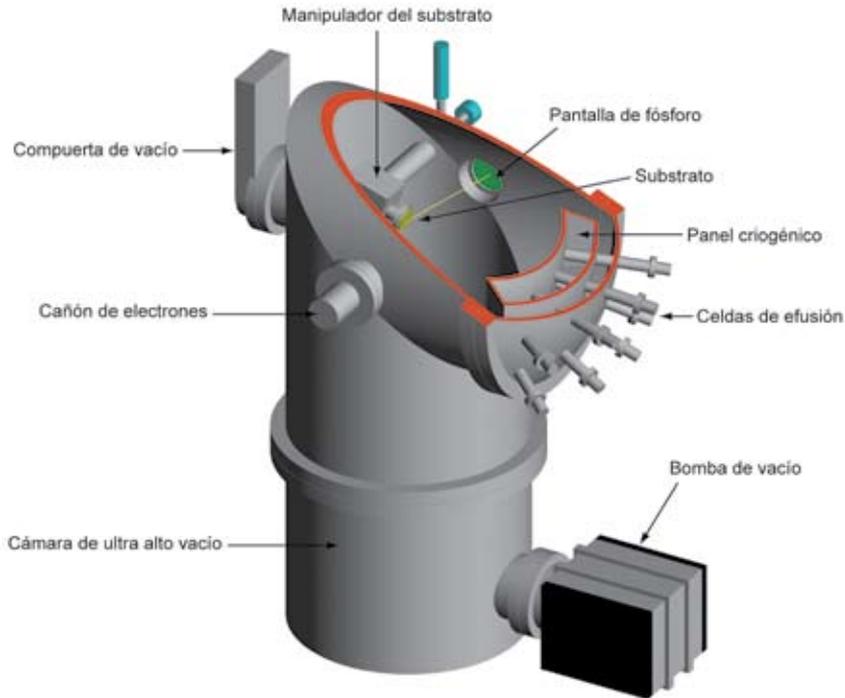


Fig.5.7 Corte esquemático de un sistema de epitaxia por haces moleculares. El sistema está basado en una cámara de ultra vacío capaz de alcanzar un vacío superior a 10^{-10} mm Hg. Las celdas de efusión están colocadas en la periferia de la cámara de crecimiento apuntando al centro de la misma, en donde se coloca al sustrato sobre un manipulador con varios grados de libertad. El sistema permite seguir en tiempo real el crecimiento de la película epitaxial mediante difracción de electrones. El panel criogénico proporciona un aislamiento térmico entre las celdas de efusión.

En la figura 5.8 se muestra un diagrama esquemático de una celda de efusión. Se pueden distinguir las siguientes componentes: 1) un crisol de BN de ultra alta pureza en donde se colocan los materiales a depositar, 2) un calentador para dicho crisol y 3) un termopar para

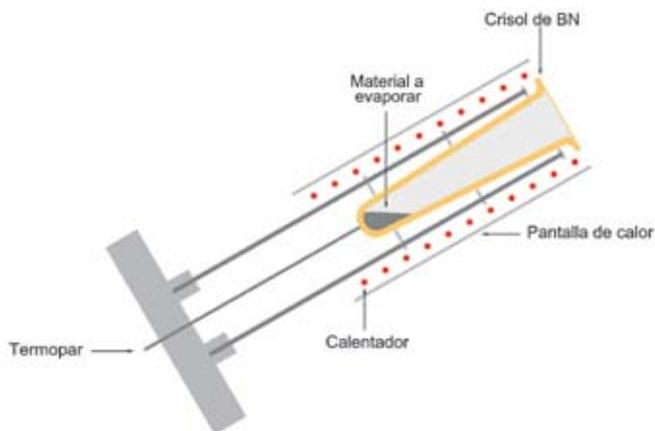


Fig.5.8 Corte esquemático de una celda de efusión. El material a evaporar se coloca en un crisol de nitruro de boro de ultra alta pureza. En la parte inferior del crisol se coloca un termopar para la medición de la temperatura del mismo. El crisol se calienta por medio de una resistencia enrollada a su alrededor. En la boca de la celda se coloca un obturador (no mostrado) que permite interrumpir o permitir el flujo de moléculas según sea necesario.

medir su temperatura. Además-no mostrado-, enfrente de la boca de la celda de coloca un obturador mecánico móvil para interrumpir o permitir el flujo molecular generado en el interior de la misma.

Los haces moleculares de los elementos metálicos del grupo III (Al, In y Ga) se generan a partir de la evaporación del elemento correspondiente, el cual debe tener una pureza extrema. El sistema debe contar también con celdas de efusión para los elementos empleados como dopantes, tanto tipo n como tipo p.

Además del control de la temperatura del sustrato, es igualmente muy importante controlar la temperatura de la celda de efusión, ya que de esto depende la intensidad del flujo molecular de una especie química dada. Cabe también señalar que la temperatura necesaria para obtener un flujo determinado depende fuertemente del material a evaporar, y de manera específica de su presión de vapor. De este

modo, para obtener un flujo adecuado de un elemento de grupo III puede ser necesario elevar la temperatura de la celda de efusión por arriba de los 1000 °C, ya que los metales tienen una presión de vapor baja. En contraste, para evaporar los elementos del grupo V (As, P), que tienen una presión de vapor considerablemente más alta, sólo son necesarias temperaturas de la celda de efusión de algunos cientos de grados centígrados.

Dada esta disparidad en temperaturas, las celdas de efusión se aíslan térmicamente por medio de paneles criogénicos enfriados con nitrógeno líquido a una temperatura de -196 °C, como se muestra en la figura 5.7. Los paneles criogénicos, además, ayudan a conservar el alto vacío dentro de la cámara de crecimiento, congelando en su superficie las impurezas que se desprenden de las superficies interiores que están expuestas a altas temperaturas.

Es posible controlar la velocidad de crecimiento de las capas epitaxiales mediante el control de la temperatura de las celdas de efusión. La velocidad típica de crecimiento por MBE es del orden de una capa atómica por segundo, de modo que una película de GaAs de un micrómetro de espesor tomaría una hora en crecer. Dadas estas bajas velocidades de crecimiento, es posible tener un gran control del espesor de la película epitaxial con la ayuda del obturador mecánico con que cuenta cada celda de efusión, el cual abre o cierra el flujo molecular a voluntad. El tiempo de conmutación de dichos obturadores es de una fracción de segundo. Así, es posible controlar el espesor de la película epitaxial con una precisión de una capa atómica e incluso menos.

Con la ayuda de los obturadores de las celdas de efusión se puede también controlar la composición de la película epitaxial abriendo o cerrando los flujos moleculares según la composición deseada. De esta manera, es posible crecer estructuras de muchas capas con diferentes composiciones químicas, algunas de ellas muy delgadas y con fronteras bien definidas. Un ejemplo ilustrativo de las posibilidades de la técnica MBE es la fabricación de las llamadas superredes, que consisten de hetroestructuras con cientos de capas epitaxiales de

composición química alternada y espesores de decenas de ángstroms. Las superredes semiconductoras se emplean en la fabricación de dispositivos optoelectrónicos de una gran complejidad, entre ellos láseres de semiconductor,

Además de lo anterior, el método de MBE tiene una ventaja adicional sobre otros métodos de crecimiento epitaxial. Esta ventaja tiene que ver con el hecho de que el proceso de crecimiento por MBE se realiza en alto vacío, de modo que es posible emplear técnicas de difracción de electrones para caracterizar el crecimiento en el momento en que está ocurriendo. La difracción de electrones nos permite obtener información sobre la calidad cristalina de la película epitaxial, así como acerca de la velocidad con que ocurre el crecimiento con una precisión de una fracción de capa atómica.

Para llevar a cabo tales determinaciones es necesario acoplar a la cámara de crecimiento una cañón de electrones y una pantalla fluorescente que permita visualizar el patrón de difracción producido por la película epitaxial. El haz de electrones se hace incidir sobre el sustrato a un ángulo raso, y después de interactuar con la superficie de la capa epitaxial se dirige sobre la pantalla fluorescente, tal como se muestra en la figura 5-7.

El control de crecimiento epitaxial puede llevarse a cabo con la ayuda de una computadora, la cual manipula la posición abierto/cerrado de los obturadores de las celdas de efusión y fija las temperaturas tanto del sustrato como de dichas celdas, de acuerdo con un programa preestablecido. De este modo es posible fabricar estructuras de una gran complejidad, como las superredes mencionadas con anterioridad.

Por otro lado, debido a que las celdas de efusión tienen un volumen limitado, se tiene la necesidad de recargarlas de manera periódica, para lo cual hay que abrir la cámara de crecimiento. Tal circunstancia puede considerarse como una desventaja de la técnica MBE.

Este problema puede subsanarse, sin embargo, mediante el empleo de fuentes gaseosas para los diferentes elementos químicos a depositar, en lugar de celdas de efusión con fuentes sólidas. Los

compuestos gaseosos empleados en este caso son similares a los utilizados en los sistemas MOCVD y son mezclados externamente en las proporciones requeridas e introducidos a las celdas en donde se disocian a altas temperaturas. Así, es posible hacer el cambio de un contenedor de gas sin necesidad de abrir la cámara de crecimiento y por lo tanto ésta puede permanecer con un alto vacío por un tiempo más prolongado. Esta variante de la técnica MBE recibe el nombre de epitaxia por haces moleculares con fuente gaseosa y es, de hecho, la unión de las dos tecnologías epitaxiales: MBE y MOCVD.

Es interesante mencionar que el nivel de vacío en el espacio es lo suficientemente alto para realizar un crecimiento por MBE sin necesidad de una cámara de vacío. En la nave espacial norteamericana *Space Shuttle* se llevaron a cabo hace algunos años crecimiento epitaxiales por MBE empleando el vacío espacial.

De todo lo anterior se puede concluir que el equipo que se requiere para la tecnología MBE es de una gran complejidad y tiene un elevado costo, el cual ronda a un millón de dólares.

5.5 Conclusiones

Las tres tecnologías analizadas en este capítulo, epitaxia por fase líquida, epitaxia por fase gaseosa y epitaxia por haces moleculares, son ampliamente utilizadas hoy en día para fabricar láseres de semiconductor, entre muchos otros dispositivos electrónicos y optoelectrónicos. La primera de estas tres tecnologías que fue empleada para la obtención de materiales semiconductores fue la epitaxia por fase gaseosa. Al inicio de los años 60s, mediante esta técnica se crecieron películas de GaAs y de GaP, así como soluciones sólidas de $\text{GaAs}_{1-y}\text{P}_y$. El material atractivo para la fabricación de heteroláseres, sin embargo, resultó ser el $\text{Al}_x\text{Ga}_{1-x}\text{As}$, ya que este semiconductor forma una heteroestructura casi ideal con el GaAs en todo el rango de composición x . Aunque el $\text{Al}_x\text{Ga}_{1-x}\text{As}$ no pudo crecerse por epitaxia por fase gaseosa mediante el proceso de transporte de cloro en boga en los años 60, debido

a la alta reactividad química del aluminio, sí pudo sintetizarse por la técnica de epitaxia por fase líquida. Así, a finales de los años 60s se realizaron con éxito los primeros crecimientos de la solución sólida $\text{Al}_x\text{Ga}_{1-x}\text{As}$ por esta técnica, la cual posteriormente se convirtió en el principal método de crecimiento de heteroestructuras empleando el sistema AlAs-GaAs .

A mediados de los años 70s se intensificaron las investigaciones dirigidas al desarrollo de la tecnología de epitaxia por haces moleculares, incluyendo la construcción del equipo necesario. De manera simultánea se llevaron a cabo trabajos de investigación sobre el crecimiento de materiales semiconductores por epitaxia por fase gaseosa empleando compuestos organo-metálicos, los cuales lograron éxitos considerables. Los avances obtenidos con las técnicas MBE y MOCVD llevaron a que la técnica LPE fuera desplazada poco a poco. Las ventajas de los métodos MBE y MOCVD sobre LPE radican en que con las primeras es posible crecer películas muy delgadas, hasta de algunas capas atómicas, además de que es más sencillo controlar la composición química y la concentración de las impurezas de dichas películas. Es factible, además, realizar crecimientos heteroepitaxiales sobre sustratos de diferente tipo de red cristalina; por ejemplo, un semiconductor III-V sobre un sustrato de silicio.

Por otro lado, el método de LPE involucra procesos más cercanos al equilibrio termodinámico que aquellos procesos característicos de las técnicas MBE y MOCVD. Debido a esta circunstancia, los materiales obtenidos por LPE pueden mostrar propiedades optoelectrónicas superiores a aquellas de los materiales crecidos por MBE y MOCVD. Este es el caso, por ejemplo, de la eficiencia de recombinación radiativa. Otra ventaja de la técnica LPE es el costo del equipo necesario para su instalación, que es mucho menor que el correspondiente costo de las técnicas MBE y MOCVD.

Se mencionará, finalmente, que las tecnologías epitaxiales descritas en este capítulo tienen todas ventajas y desventajas y que en nuestra opinión estas técnicas continuarán perfeccionándose y aplicándose para el crecimiento de las más diversas estructuras semiconductoras.

CAPÍTULO 6

Diferentes diseños de diodos láser con resonador de Fabry-Perot

6.1 Introducción

En los capítulos 1 y 2 se consideraron los principios generales de funcionamiento de los láseres de semiconductor. Entre otras cosas, aprendimos que para obtener luz coherente es necesario contar con un resonador óptico formado por dos espejos paralelos. Debido a que el índice de refracción de los semiconductores es suficientemente grande, estos dos espejos pueden simplemente estar formados por las superficies clivadas del cristal, que por naturaleza son paralelas y lisas. Los resonadores formados de esta manera se denominan resonadores o cavidades de *Fabry-Perot*. En este capítulo vamos a discutir diferentes diseños de láseres con cavidades de Fabry Perot. Hacemos notar que si bien habremos de considerar solamente este tipo de dispositivos, existen láseres con resonadores diferentes a los de Fabry-Perot.

6.2 Láseres con resonadores de tipo Fabry-Perot

A. Construcción del resonador de Fabry-Perot

Para obtener clivados de buena calidad es necesario que el espesor d de la oblea, que incluye el substrato y las diferentes capas epitaxiales de la heteroestructura, no sea mayor de 100-120 μm . Tenemos, por otro lado, que durante el proceso de crecimiento epitaxial, ya sea por

fase líquida o gaseosa, el sustrato debe tener un espesor de 300-500 μm a fin de que su resistencia mecánica sea lo suficientemente grande para no fracturarse con las manipulaciones -limpieza, ataques químicos, etc.- a las que es sometido. Por esta razón, después del crecimiento epitaxial el sustrato debe adelgazarse, ya sea mediante un pulido mecánico o por un ataque químico, hasta que el espesor total d de la oblea alcance el valor deseado. Notamos que el espesor total de las capas crecidas por lo general no sobrepasa las 10 μm , por lo que d está determinado por el espesor del sustrato. Después de la operación de adelgazado se forman contactos metálicos -óhmicos- sobre la superficie inferior del sustrato y sobre la capa epitaxial superior. Esto se realiza generalmente evaporando en vacío diferentes metales, los cuales posteriormente se someten a un tratamiento térmico o recocido en un ambiente de hidrógeno o de algún gas inerte. El siguiente paso consiste en clivar la oblea en pequeñas barras de ancho b como se muestra en la figura 6.1b. El ancho de estas barras determina el largo de los reso-

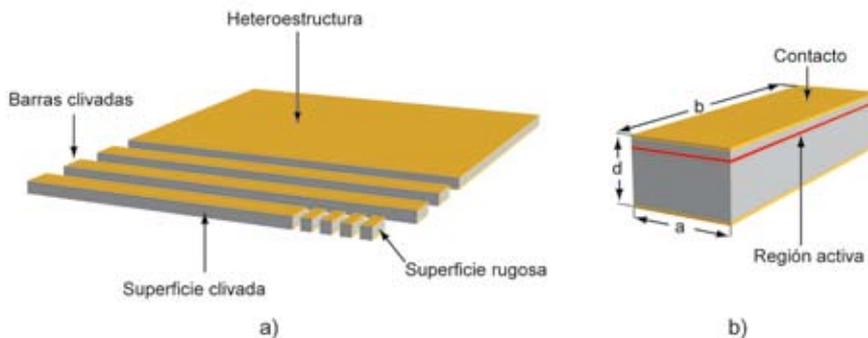


Fig. 6.1 Fabricación de láseres con resonador de Fabry-Perot . a) Después de formar los contactos metálicos, la oblea se cliva en delgadas barras, las cuales a su vez se cortan en dispositivos con una relación entre largo y ancho de 2.5-3. b) Láser de contacto ancho en el que el contacto superior cubre por completo la superficie del dispositivo.

nadores de Fabry Perot. En seguida, las barras se subdividen en láseres individuales de ancho a con ayuda de un disco con recubrimiento de diamante. Por lo general, la relación entre el largo b y el ancho a del resonador tiene un valor en el rango 2.5-3.

B. Láseres de contacto ancho

Los dispositivos obtenidos con el procedimiento descrito en el párrafo anterior tienen dos superficies clivadas y reflejantes de área $A=ab$, que constituyen los extremos del resonador óptico, y dos superficies laterales rugosas de largo b generadas con el disco de diamante. Esta geometría permite al láser generar radiación sólo en su dirección longitudinal. Lo anterior debido a que la luz que se desplaza lo largo de la región activa sufre múltiples reflexiones en las superficies clivadas del cristal sin cambiar su dirección, mientras que la luz que viaja transversalmente se dispersa en las superficies rugosas cambiando dirección de manera aleatoria con cada reflexión. Si fuera el caso de que las cuatro superficies del cristal estuvieran clivadas, se perdería parcial o totalmente la selectividad en la dirección de propagación de la radiación, algo indeseable para un láser.

Para su operación, el láser se fija con indio a un disipador de calor de cobre, el cual también funciona como el electrodo inferior del dispositivo. El otro electrodo lo constituye un delgado alambre de oro soldado al contacto superior del diodo (ver figura 6.2).

Los láseres descritos en la sección anterior se clasifican como de *contacto ancho*. Al aplicar un voltaje entre ambos electrodos en este tipo de láseres, se hace circular una corriente eléctrica a través del dispositivo que es uniforme a lo largo y ancho del contacto, al igual que lo es la densidad de portadores de carga inyectados. El área activa, en consecuencia, se extiende a toda el área A del láser, como se ilustra en la figura 6.1b.

Notamos que en un láser de heteroestructura de contacto ancho la región activa esta confinada en la dirección perpendicular a la unión

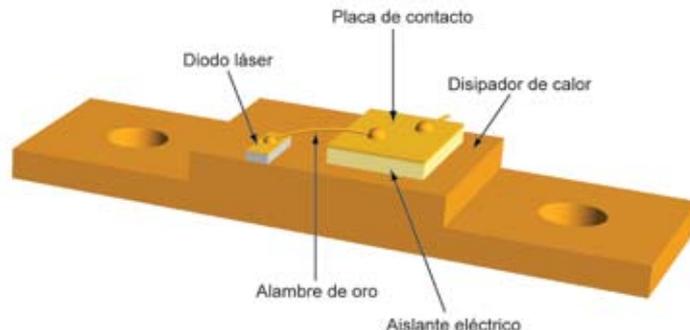


Fig. 6.2 Láser de semiconductor montado en un disipador de calor de cobre, el cual funciona además como electrodo inferior.

p-n, lo que le da ventajas sobre los láseres de homounión tal como se discutió el capítulo 2. Los láseres de contacto ancho, a su vez, pueden mejorarse substancialmente por medio de estructuras que confinan la radiación generada en dos dimensiones. Una manera de realizar esto -aunque no la más eficiente como se discutirá más tarde- consiste simplemente en sustituir al contacto superior de ancho a por un contacto longitudinal con un ancho mucho menor -contacto angosto-, el cual tiende a confinar a la corriente que circula en el láser, y por lo tanto a su región activa, al área cubierta por dicho contacto. Esto se ilustra de manera esquemática en la figura 6.3a. En lo que sigue discutiremos con detalle estos dispositivos.

C. Láseres de contacto angosto

Confinar a los portadores de carga inyectados y a la radiación generada en dos dimensiones tiene una serie de ventajas. La primera es que se obtiene una disminución del volumen de la región activa, lo que conduce a una reducción de la corriente de bombeo necesaria para operar el láser. Así, tenemos que el valor característico de corriente

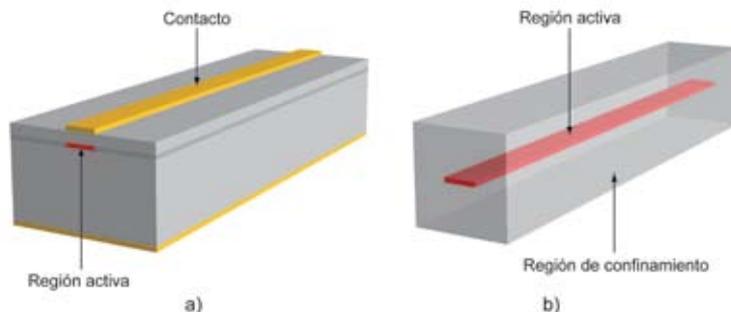


Fig. 6.3 a) Láser de contacto ancho que confina lateralmente la región activa del dispositivo. b) Región activa del láser ideal, la cual está limitada tanto vertical como lateralmente. Dicha limitación se logra rodeando a la región activa con un semiconductor, o semiconductores, de brecha energética mayor -confinamiento eléctrico- e índice de refracción menor -confinamiento óptico- que los de dicha región.

de operación de un láser de contacto ancho de dimensiones típicas es del orden de algunos amperes, mientras que en un láser de contacto angosto puede ser tan pequeña como algunos miliamperes. Las bajas corrientes de umbral de los láseres de contacto angosto permiten su operación en un régimen continuo de generación de luz coherente a temperatura ambiente.

Una de las aplicaciones más importantes de los láseres de semiconductor es la transmisión de información en sistemas de comunicación por fibra óptica. En esta aplicación, la intensidad de la luz láser se modula variando la amplitud de la corriente de bombeo. La capacidad informática de un canal de transmisión, por otro lado, está determinada por el volumen de información transmitida por unidad de tiempo y esta característica, a su vez, depende de la velocidad de modulación de la corriente que pasa por el láser. Es obvio que a menor magnitud absoluta de corriente ésta se podrá modular a mayor velocidad, y en consecuencia los láseres de contacto angosto tienen mejores características de frecuencia que los dispositivos de contacto ancho.

Los láseres de contacto angosto tienen una ventaja más frente a los láseres de contacto ancho. Esta ventaja consiste en que todo el perímetro de la unión p-n se encuentra lejos de la superficie del cristal, lo que retarda el proceso de degradación de los dispositivos de manera significativa y aumenta su periodo de vida útil.

En la figura 6.3b se representa de manera abstracta la región activa de un láser ideal. En esta representación, dicha región está contenida dentro una caja rectangular con paredes bien definidas situada en el interior del dispositivo. La región activa de un láser diseñado de forma correcta se acerca a la representación ideal de la figura 6.3b. Para lograrlo se deben satisfacer dos condiciones: 1) los portadores de carga inyectados a la región activa deben confinarse por medio de barreras de potencial y 2) la radiación generada debe ser confinada empleando barreras de índice de refracción. Esto último se logra rodeando la región activa de índice de refracción n_1 con un semiconductor con un índice de refracción $n_2 < n_1$. De esta manera, la radiación incidiendo sobre las paredes de la región activa será reflejada totalmente, tal como se discutió en la sección 1.7. En la sección 6.3 se discuten diversas estructuras láser que buscan acercarse al ideal de la figura 6.3b.

D. Características electromagnéticas

Se encuentra que la distribución de la radiación electromagnética, o campo óptico, dentro de la región activa depende de las dimensiones de la misma. En efecto, dicha radiación se estabiliza cuando se producen interferencias constructivas y para que estas ocurran la región activa debe contener un número entero de semi longitudes de onda.

De este modo, en la dirección perpendicular a las capas de la heteroestructura láser, la distribución de la intensidad del campo óptico tiene un solo máximo debido al pequeño espesor de la capa activa. Si la dimensión transversal de la región activa, es decir su ancho, no sobrepasa los 3-5 μm , entonces también en ésta dirección la intensidad del campo óptico tendrá un sólo máximo, en cuyo caso se dice que el láser emite

en un solo modo transversal -el modo es transversal porque la dirección del campo electromagnético de la radiación es perpendicular a su dirección de propagación. Por otro lado, en la medida en que incrementa el ancho de la región activa, aparecerán nuevos máximos en la intensidad de campo óptico y nuevos modos transversales generados por el láser, tal como se ilustra en la figura 6.4.

En el caso del régimen de generación de un solo modo transversal, la radiación del láser se concentra sólo en un lóbulo. Si por el contrario, la distribución de la intensidad del campo óptico tiene varios máximos, la radiación láser se distribuirá en diferentes lóbulos, el número de los cuales coincidirá con el número de los máximos de la intensidad del campo. Esto se ilustra en la figura 6.5, en donde se muestran de manera esquemática los patrones de radiación de los llamados campo cercano y campo lejano del láser, para tres anchos de la región activa que coinciden con los del la figura 6.4. Notamos que mientras que el patrón de campo cercano corresponde a la distribución de intensidad del campo de la figura 6.4, el patrón de campo lejano está fuertemen-

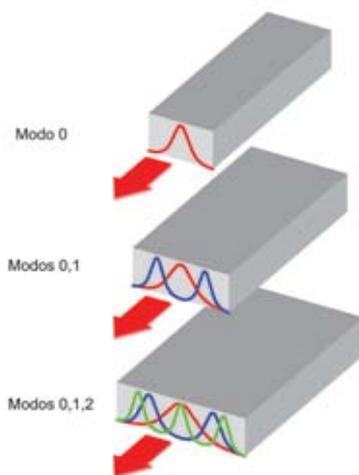


Fig. 6.4 Modos transversales en un láser para tres diferentes anchos de la región activa.

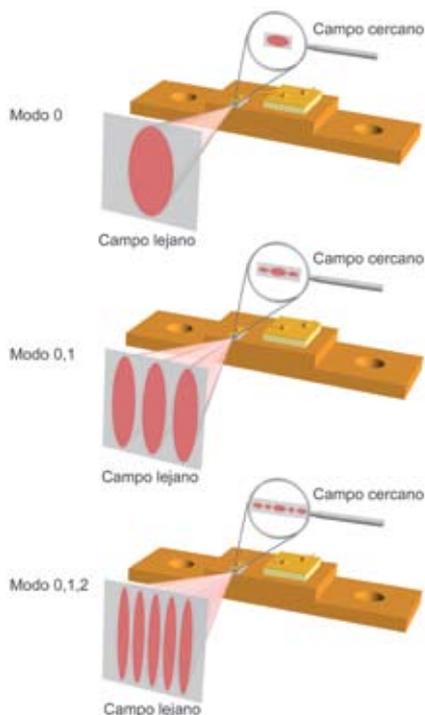


Fig. 6.5 Patrones de campo lejano y de campo cercano de un láser semiconductor.

te determinado por fenómenos de difracción. En efecto, una vez que deja la región activa, la radiación de láser diverge con un ángulo que depende de las dimensiones de dicha región, las cuales son comparables con la longitud de onda de la radiación láser y por lo tanto producen marcados efectos difractivos. De este modo, el patrón de campo lejano tiene dimensiones mayores en la dirección perpendicular al plano de la heteroestructura que en la dirección paralela a la misma, como se muestra de manera esquemática en la figura 6.5, debido a que la dimensión perpendicular de la región activa es del orden de $1\ \mu\text{m}$, mientras que ésta es de varios micrómetros en la dirección horizontal.

Para una serie de aplicaciones prácticas, que incluyen la comunicación por fibra óptica, es mucho más conveniente utilizar los láseres de un solo modo transversal en lugar de los láseres multimodo. Estos láseres pueden ser fabricados en base a estructuras de *stripe* -término en inglés que significa franja-, como se discute en la siguiente sección.

Hay que hacer notar que adicionalmente a los modos transversales, en la cavidad de Fabry-Perot de un láser existen un gran número de modos longitudinales. En una cavidad de longitud b , estos modos corresponden a longitudes de onda λ_m que cumplan con la condición de interferencia constructiva $m\lambda_m = 2nb$, donde m es un número entero y n es el índice de refracción de la cavidad. Existen aplicaciones en las cuales, además de una operación monomodo transversal, es necesario que un láser de semiconductor emita en un solo modo longitudinal. En estos casos es menester recurrir a cavidades con espejos de Bragg integrados, los cuales seleccionan uno de los modos longitudinales de la cavidad de Farby-Perot. Este tipo de estructuras serán discutidas en un libro futuro.

6.3 Diferentes diseños de láseres con estructura de *stripe*

Existe una gran variedad de láseres de estructura de *stripe*, con ventajas y desventajas unos respecto de otros. En lo que sigue consideraremos los diseños más importantes.

A. Láser con contacto de *stripe*

En la figura 6.6 se muestra un láser de estructura de *stripe* con uno de los diseños más simples posibles. Este tipo de dispositivos se fabrica como sigue. Como primer paso, una vez crecida la heteroestructura láser se deposita sobre la última capa epitaxial de la oblea una película dieléctrica -SiO₂, por ejemplo. A continuación, utilizando técnicas de

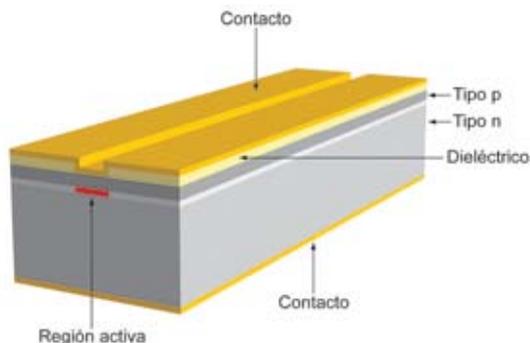


Fig. 6.6 Diagrama del láser con contacto de franja o *stripe*. El dieléctrico aísla eléctricamente a la heteroestructura del contacto superior de oro, excepto en la región de la franja libre de dieléctrico. La región activa queda entonces confinada eléctricamente por dicha franja.

fotolitografía, se abren franjas en la capa de dieléctrico. Por lo común, el ancho de estas franjas es de algunos micrómetros, con distancias entre ellas de 200-400 micrómetros. Como siguiente paso, sobre toda la estructura se deposita una fina película de oro que juega el papel de contacto eléctrico.

Con este arreglo, la heteroestructura queda aislada eléctricamente del contacto de oro excepto en la región de la franja libre de dieléctrico, de modo que al aplicarse un voltaje en polarización directa la corriente tenderá a confinarse en dicha región. Este confinamiento, sin embargo, no es perfecto, pues la corriente eléctrica que circula en el dispositivo no lo hace en una dirección estrictamente perpendicular a la unión p-n. Por el contrario, dicha corriente tiene un componente paralela a la unión p-n, que es más importante cerca de los bordes de la franja y cuya magnitud relativa a la componente perpendicular crece con la corriente total. De este modo, al incrementar la corriente a través del dispositivo se incrementa el ancho de la región de inversión de población y por lo tanto el ancho de la región activa. Esto constituye una desventaja para los láseres con esta construcción.

Un problema adicional con este tipo de láseres es que no existe un mecanismo que limite el campo óptico en la dirección horizontal, de tal manera que al variar la corriente del láser -por ejemplo al modularla- la estructura de modos de la radiación generada se modificará.

B. Diseño de implantación de iones

Un diseño de láser con estructura de *stripe* que resuelve parcialmente los problemas de los láseres discutidos en el apartado anterior, es aquel que emplea la implantación de partículas de alta energía para definir la región activa del dispositivo. El proceso de fabricación de un láser por implantación de iones se explica a continuación.

Se parte de una oblea con una heteroestructura láser completamente formada, incluyendo todas sus capas epitaxiales y el contacto superior de oro. Como primer paso, la oblea se somete a un flujo de protones o de iones de oxígeno como se muestra esquemáticamente en la figura 6.7. Las partículas cargadas penetran -se implantan- hasta una cierta profundidad. Esta profundidad depende de la energía de las partículas, la

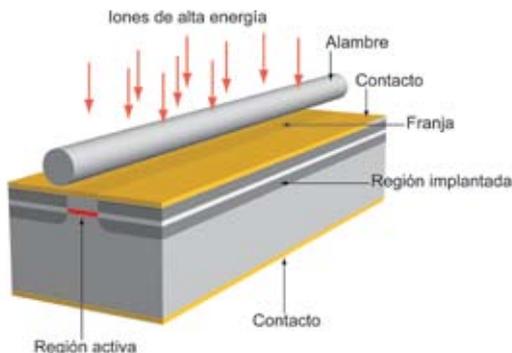


Fig. 6.7 Fabricación de un láser de contacto estrecho empleando la implantación de partículas de alta energía para definir la región activa.

cual se escoge de modo que la penetración rebase el plano de la unión p-n. Al implantarse, las partículas de alta energía generan defectos en el semiconductor, los cuales le confieren una alta resistividad.

La implantación de iones en la oblea, por otro lado, no se hace de manera uniforme sobre toda su superficie, sino que se dejan sobre ésta franjas delgadas libres de partículas implantadas. Para lograrlo, se enmascara la oblea con franjas metálicas depositadas sobre su superficie. De manera alternativa, dicha superficie se puede enmascarar con alambres de tungsteno colocados en su proximidad cercana, como se ilustra en la figura 6.7. Las partículas de alta energía son fuertemente absorbidas por las franjas metálicas o por los alambres de tungsteno, según sea el caso, y por lo tanto no penetran a la oblea. Se forman de este modo regiones no implantadas con forma de *stripe* delgada -que definen la región activa de los láseres-, rodeadas por regiones implantadas de alta resistividad. La región activa queda entonces bien definida desde el punto de vista eléctrico, ya que la corriente del láser necesariamente se canalizará por la región del *stripe* de baja resistividad.

Se hace notar que al penetrar los iones implantados hasta la región activa del láser, los defectos que generan en la misma necesariamente afectan sus propiedades luminiscentes y por lo tanto aumentan las pérdidas ópticas del dispositivo. Esto lleva a un incremento en la densidad de corriente de umbral del láser. Los defectos generados por la implantación de iones, sin embargo, se pueden reducir mediante un tratamiento térmico o recocido, conservando al mismo tiempo la alta resistividad eléctrica de la región implantada.

Notamos que si bien un láser fabricado por implantación de iones supera el problema de la falta de definición de la región activa del que adolecen los láseres de *stripe* estudiados en el apartado anterior, en cuanto al confinamiento lateral de la radiación generada el diseño no aporta nada nuevo. En efecto, el láser de implantación no cuenta con barreras de índice de refracción para confinar lateralmente la radiación generada, así que esta puede penetrar la región implantada con las consecuentes pérdidas ópticas. Este problema puede ser superado con el llamado láser con *stripe* de meseta que se discute en el siguiente apartado.

C. Láser con *stripe* de meseta

En la figura 6.8 se muestra esquemáticamente la secuencia de fabricación de un láser con *stripe* de meseta. La secuencia incluye varias etapas tecnológicas que involucran técnicas de fotolitografía, ataque químico, depósito de dieléctricos y depósito de metales. Como un primer paso se deposita una capa de un material dieléctrico -por ejemplo, SiO_2 - que sea químicamente estable a los ataques químicos que posteriormente

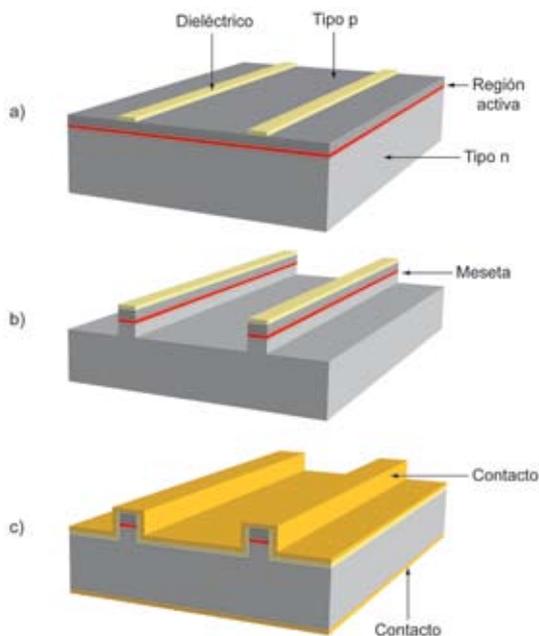


Fig. 6.8 Esquema de Fabricación de un láser con *stripe* de meseta. a) Formación de franjas de SiO_2 por medio de técnicas de fotolitografía. b) Ataque químico para decapar las áreas de la estructura que no se encuentran protegidas por las franjas del SiO_2 . El ataque se lleva hasta una profundidad que rebasa la región activa del láser. c) Depósito de contactos metálicos.

se llevarán a cabo. En seguida, se definen *stripes* de material dieléctrico empleando técnicas fotolitográficas, como se muestra en la figura 6.8a. Como siguiente paso se realiza un ataque químico para remover las partes de la heteroestructura que no se encuentran protegidas por las franjas de dieléctrico. Este ataque se prolonga hasta alcanzar una profundidad que rebasa a la región activa del láser. Como resultado, sobre la superficie de la oblea se forman estructuras con una altura correspondiente a la profundidad del ataque, tal como se ilustra en la figura 6.8b. Estas estructuras son conocidas como mesetas, de donde toman el nombre los láseres descritos en esta sección.

Una vez formadas las mesetas, se les retira la capa dieléctrica protectora mediante un ataque químico selectivo, después de lo cual se deposita nuevamente una capa dieléctrica sobre toda la superficie de la oblea. En seguida, empleando técnicas fotolitográficas, se retira la película dieléctrica de la región de las franjas, como se muestra en la figura 6.8c y finalmente, se deposita una capa metálica sobre la oblea para formar el contacto eléctrico. Notamos que toda la estructura, con excepción de la parte superior de las mesetas, se encuentra aislada del contacto. La estructura láser resultante se ilustra en la figura 6.9.

Este láser cuenta con una delimitación tanto eléctrica como óptica de la región activa, ya que: 1) la corriente eléctrica está limitada a la

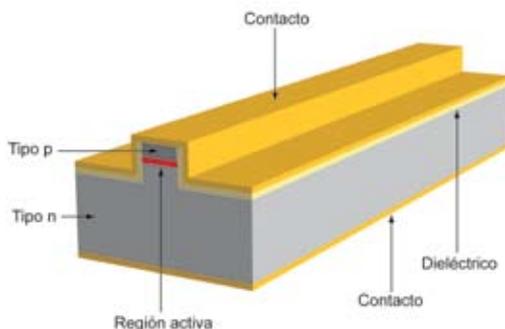


Fig. 6.9 Diagrama del láser con *stripe* de meseta.

región de la meseta y 2) la radiación generada es confinada por las paredes laterales de la misma. Escogiendo un ancho de meseta suficientemente pequeño, es posible generar radiación láser en un solo modo transversal.

Desafortunadamente, este diseño de láser también tiene desventajas, las cuales están fundamentalmente relacionadas con los defectos cristalinos que se generan en las superficies laterales de la región activa, los cuales forman canales de recombinación no radiativa y de fugas de corriente. Es bien conocido que los defectos en la superficie de los semiconductores constituyen un grave problema en un gran número de dispositivos. La densidad de defectos superficiales puede reducirse, sin embargo, recubriendo la superficie del semiconductor con un material que le sea químicamente afín y que tenga una constante de red cristalina similar. En el caso de láser de *stripe* de meseta, desafortunadamente, la estructura cristalina y la composición química de la capa dieléctrica son enteramente diferentes de las de la región activa, lo que no contribuye a la reducción de los defectos en las superficies laterales de dicha región.

Lo anterior sugiere que para fabricar un láser con estructura de *stripe* que se aproxime al ideal representado en la figura 6.3b, la región activa debe estar delimitada en la dirección paralela a la unión p-n por un material semiconductor con el mismo parámetro de red cristalina que el de dicha región. Esto, con el fin de minimizar la generación de defectos cristalinos. Además, para posibilitar los confinamientos eléctrico y óptico, este material semiconductor debe tener una brecha prohibida más grande y un índice de refracción menor que los de la región activa. En otras palabras, las ventajas y particularidades de las heterouniones, analizadas en detalle en el capítulo 2, deben de ser aprovechadas no solamente en la dirección perpendicular a la unión p-n, sino también en su dirección transversal.

Para crear un confinamiento de portadores de carga y de radiación en la dirección lateral se requieren de procedimientos tecnológicos especiales que frecuentemente involucran un segundo crecimiento epitaxial. A los dispositivos en los que se ha llevado a cabo este se-

gundo crecimiento se les conoce como láseres de estructura de *stripe* recrecidos.

Existe una gran variedad de este tipo de láseres, así como de procedimientos tecnológicos para su fabricación, y no es nuestra intención hacer una discusión exhaustiva de los mismos. En lugar de esto, vamos a considerar solamente algunos casos particulares, que son los más importantes según nuestro criterio. Uno de estos casos es el llamado láser recrecido con *stripe* de meseta, cuya estructura se muestra en la figura 6.10. En el siguiente apartado se discutirá en detalle la fabricación de este tipo de láseres.

D. Láseres recrecidos con *stripe* de meseta

Para la fabricación de estos dispositivos se parte de la estructura con mesetas de *stripe* mostrada en la figura 6.8b, la cual hubo de ser obtenida siguiendo los procedimientos discutidos en la sección anterior. Como primer paso, a dicha estructura se le remueve la capa de

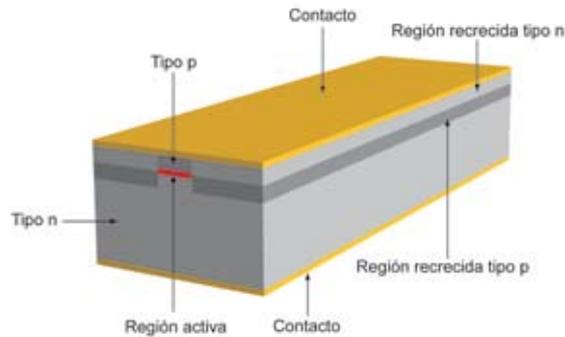


Fig. 6.10 Diagrama de un láser recrecido con *stripe* de meseta. Al polarizar directamente la región activa, la unión p-n de la región recrecida se polariza inversamente impidiendo el paso de corriente a través de la misma. La corriente circula solamente por la región de la meseta.

dieléctrico remanente y se lleva a cabo un recrecimiento por epitaxia en fase líquida de un material de brecha prohibida mayor que la de la región activa. Es importante hacer notar que la forma de las mesetas, así como las particularidades del proceso de recrecimiento sobre las mismas, dependen en gran medida de la orientación cristalográfica del sustrato.

Consideremos láseres de recrecimiento fabricados con base en el sistema AlAs-GaAs, con una región activa de GaAs. En este caso, se recrecen dos capas epitaxiales de $\text{Al}_x\text{Ga}_{1-x}\text{As}$, una tipo p y la otra tipo n, con una composición de aluminio en el rango $x=0.3-0.6$. Por la diferencia en índices de refracción entre la región activa y las capas recrecidas, el campo óptico se confina dentro de los límites de las mesetas. El confinamiento eléctrico, por otro lado, se realiza con ayuda de la unión p-n de $\text{Al}_x\text{Ga}_{1-x}\text{As}$ recrecida, la cual se polariza inversamente al polarizar directamente la región activa. En relación a esto, recordamos que en una unión p-n polarizada inversamente se induce una región de alta resistividad que de manera efectiva impide el paso de corriente a través de ella. De esta manera, la corriente solamente circulará a través de la meseta y la extensión lateral de la región activa del láser coincidirá con el ancho de la misma. Este tipo de dispositivos puede ser fabricado con secciones transversales de región activa tan pequeñas que en algunos casos aislados se han obtenido corrientes de umbral del orden de 10-15 μA .

E. Otros diseños

Discutiremos dos diseños adicionales de láseres con estructura de *stripe*: los láseres aislados por una unión p-n y los láseres de canal. En la discusión vamos a considerar dispositivos basados en soluciones sólidas $\text{Al}_x\text{Ga}_{1-x}\text{As}$, aunque pueden ser igualmente fabricados con otros semiconductores.

Consideremos primeramente los láseres aislados por una unión p-n. Para fabricar un láser de este tipo, primeramente se crean franjas de

un dieléctrico $-\text{SiO}_2$ o Si_3N_4 sobre un sustrato de GaAs tipo n por medio de técnicas fotolitográficas, tal como se muestra en la figura 6.11a. Estas franjas son las que dan lugar al *stripe* una vez terminado el dispositivo. Posteriormente, el sustrato de GaAs se expone a una

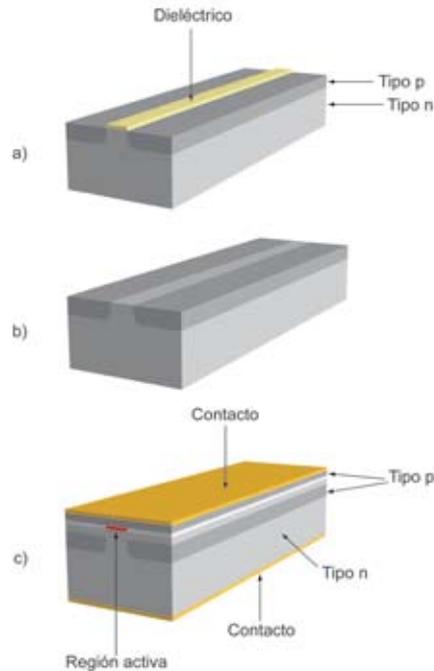


Fig. 6.11 Diagrama de fabricación de un láser aislado por una unión p-n. a) Difusión selectiva de un impureza tipo p (cinc) en un sustrato de GaAs tipo n. La difusión selectiva se lleva a cabo depositando una película dieléctrica (SiO_2 o Si_3N_4) sobre la oblea semiconductora y definiendo franjas de dieléctrico por técnicas fotolitográficas. Dichas franjas son impermeables a la difusión de cinc. b) Remoción de las franjas de dieléctrico. c) Estructura completa fabricada recreciendo las diferentes capas epitaxiales que forman la heteroestructura láser sobre la oblea con difusión selectiva.

alta temperatura y a un flujo de átomos de cinc, los cuales se difunden hacia el interior excepto en las regiones protegidas por la franja de dieléctrico. El cinc es una impureza tipo p que invierte el tipo de conductividad del sustrato en las regiones hasta donde penetra. De esta manera, sobre la superficie del GaAs se forman regiones tipo p aisladas por regiones tipo n. Como siguiente paso, se retira el dieléctrico (figura 6.11b) y se crecen todas las capas epitaxiales de la heteroestructura láser. Finalmente, se depositan los contactos eléctricos superior e inferior. La estructura terminada se muestra esquemáticamente en la figura 6.11c.

Como es aparente de la figura 6.11c, al aplicar un voltaje positivo al contacto superior con respecto al contacto inferior, la unión p-n cerca de la superficie se polariza directamente. La corriente circula, no obstante, solamente a través del área de la franja o *stripe*, debido a que la unión p-n enterrada, formada entre la capa tipo p del sustrato y la primera capa epitaxial tipo n, se polariza en dirección inversa y por lo tanto impide el flujo de corriente. Ésta fluirá, entonces, solamente a través del área en la cual no hubo difusión de átomos de cinc.

Consideraremos ahora los láseres de canal. La secuencia de fabricación de estos láseres es la siguiente. Primeramente se crece sobre un sustrato de GaAs tipo n una capa epitaxial de GaAs tipo p. A continuación, se deposita sobre esta última una película dieléctrica a la cual se abren franjas que dejan al descubierto la superficie del GaAs. El siguiente paso consiste en atacar químicamente el sustrato en las áreas libres de dieléctrico, como se muestra en la figura 6.12a. El régimen de ataque se elige de tal manera que la profundidad del surco formado por ataque químico rebase el espesor de la capa de GaAs tipo p. Dicho con otras palabras, el surco debe dividir la capa epitaxial de conductividad tipo p. Posteriormente, después de remover el dieléctrico de la superficie, sobre la misma se crecen las capas epitaxiales que forman la heteroestructura láser. Eligiendo el régimen de crecimiento correcto, se puede crecer una heteroestructura láser de modo que la región activa del GaAs sea definida y rodeada por $\text{Al}_x\text{Ga}_{1-x}\text{As}$, como se muestra en la figura 6.12 b.

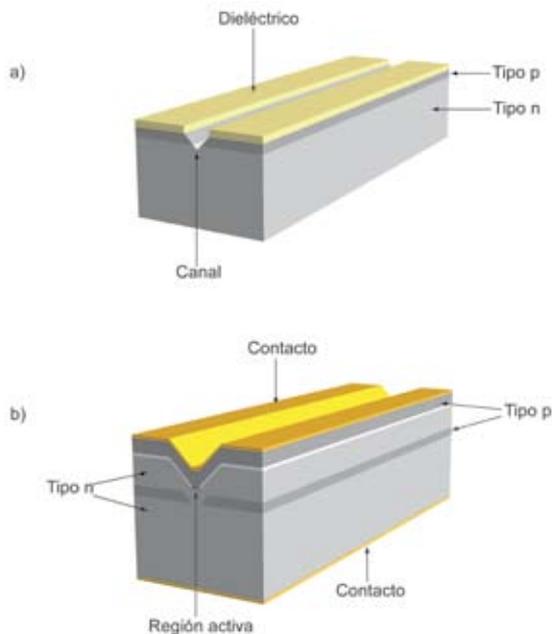


Fig. 6.12 Diagrama del proceso de fabricación de un láser de canal. a) Ataque preferencial en forma de V en una capa epitaxial de GaAs tipo p crecida sobre un sustrato de GaAs tipo n. El ataque preferencial se realiza depositando una película dieléctrica sobre la oblea y definiendo franjas libres de dieléctrico por técnicas fotolitográficas. b) Estructura completa fabricada recreciendo las diferentes capas epitaxiales que forman la heteroestructura láser sobre la oblea con canales, después de remover el remanente de la película dieléctrica.

6.4 Conclusiones

En este capítulo se han analizado varios diseños de láseres que buscan optimizar parámetros de operación tales como la corriente umbral y la frecuencia de modulación, así como obtener un patrón de emisión

monomodo transversal. Se consideraron láseres con estructura de *stripe* basados en resonadores de Fabry-Perot. Tanto con los láseres de recrecimiento con *stripe* de meseta, como con los dos láseres analizados en el apartado 6.3E, se han obtenido excelentes resultados en cuanto a los parámetros de operación.

A lo largo de los últimos decenios se han fabricado y caracterizado un número muy grande de los más diversos tipos de láseres de semiconductor. En este capítulo no se ha pretendido ofrecer una discusión exhaustiva de los mismos. Por el contrario, nuestro objetivo fue simplemente describir algunos de los dispositivos más comunes, a fin de proporcionar al lector información básica sobre los procesos tecnológicos involucrados en su fabricación.

BIBLIOGRAFÍA

Capítulo I

1. Zh. I. Alferov, R. F. Kazarinov, USSR Patent #28448 (1967).
2. D. C. O'Shea, W. R. Callen and W. T. Rhodes, *Introduction to Lasers and Their Applications*, Addison-Wesley (1977).
3. H. Kressel, J. K. Butler, *Semiconductor Lasers and Heterojunction LEDs*, Academic Press (1977).
4. B. Sapoval, C. Hermann, *Physics of Semiconductor Devices*, segunda edición, John Wiley & Sons (1981).
5. F. L. Pedrotti and L. S. Pedrotti, *Introduction to Optics*, Prentice-Hall (1987).
6. M. V. Klein and T. E. Furtak, *Optics*, segunda edición, John Wiley & Sons (1986).
7. P. Bhattacharya, *Semiconductor Devices*, segunda edición, Prentice-Hall (1997).
8. W. Koechner, *Solid State Laser Engineering*, quinta edición, Springer Series in Optical Sciences, Vol. I (1999).
9. A. Rogers, *Essentials of Photonics*, segunda edición, CRC Taylor & Francis Group (2009).
10. M. A. Parker, *Physics of Optoelectronics*, CRC Taylor & Francis Group (2005).
11. K. A. Jones, *Introduction to Optical Electronics*, John Wiley & Sons (1987).
12. H. Ghafouri-Shiraz, *The Principles of Semiconductor Laser Diodes and Amplifiers*, Imperial College Press (2004).
13. Shun Lien Chuang, *Physics of Optoelectronic Devices*, John Wiley & Sons (1995).

14. Louis Desmarais, *Applied Electro-Optics*, Prentice-Hall (1998).
15. P. Vasil'ev, *Ultrafast Diode Lasers, Fundamentals and Applications*, Artech House (1995).
16. G. H. B. Thompson, *Physics of Semiconductor Laser Devices*, John Wiley & Sons (1980).

Capítulo 2

1. Zh. I. Alferov, R. F. Kazarinov, *USSR Patent #28448* (1967).
2. Zh. I. Alferov, V. M. Andreev, E. L. Portnoi and M. K. Trucan. *Fiz. Tekh. Poluprovodn.* **3**, 1328, 1969. (*Sov. Phys. Semicond.* **3**, 1107, 1970).
3. Zh. I. Alferov, V. M. Andreev, D. Z. Garbuzov, Yu. V. Zhilyaev, E. P. Morozov, E. L. Portnoi and V. G. Trofim. *Fiz. Tekh. Poluprovodn.* **4**, 1826, 1970. (*Sov. Phys. Semicond.* **4**, 1573, 1971).
4. H. Kressel, J. K. Butler, *Semiconductor Lasers and Heterojunction LEDs*, Academic Press (1977).
5. A. G. Milnes and D. L. Feucht, *Heterojunction and Metal-Semiconductor Junctions*, Academic Press (1978).
6. H. C. Casey, Jr. and M. B. Panish, *Heterostructure Lasers*, Academic Press (1978).
7. J. K. Butler, editor, *Semiconductor Injection Lasers*, IEEE (1980).
8. A. E. Siegman, *Lasers*, University Science Books (1986).
9. J. T. Luxon and D. E. Parker, *Industrial Lasers and Their Applications*, Prentice-Hall (1987).
10. P. W. Milonni and J. H. Eberly, *Lasers*, John Wiley & Sons (1988).
11. J. T. Verdeyen, *Laser Electronics*, segunda edición, Prentice-Hall (1989).
12. Eli Kapon, *Semiconductor Lasers I: Fundamentals*, Academic Press (1998).
13. Piter S. Zory Jr., editor, *Quantum Well Lasers*, Academic Press (1993).
14. Cheng T. Wang, editor, *Introduction to Semiconductor Technology. GaAs and Related Compounds*, John Wiley & Sons (1990).

15. K. A. Jones, *Introduction to Optical Electronics*, John Wiley & Sons (1987).
16. H. Ghafouri-Shiraz, *The Principles of Semiconductor Laser Diodes and Amplifiers*, Imperial College Press (2004).
17. S. L. Chuang, *Physics of Optoelectronic Devices*, John Wiley & Sons (1995).
18. P. Vasil'ev, *Ultrafast Diode Lasers: Fundamentals and Applications*, Artech House (1995).
19. G. H. B. Thompson, *Physics of Semiconductor Laser Devices*, John Wiley & Sons (1980).

Capítulo 3

1. S. L. Chuang, *Physics of Optoelectronic Devices*, John Wiley & Sons (1995).
2. G. P. Agrawal, editor, *Semiconductor Lasers: Past, Present and Future*, American Institute of Physics (1995).
3. H. C. Casey and Jr., M. B. Panish, *Heterostructure Lasers*, Academic Press (1978).
4. P. K. Das, *Lasers and Optical Engineering*, Springer-Verlag (1990).
5. W. Streiffer and M. Ettenberg, editor, *Semiconductor Diode Lasers*, IEEE (1991).
6. N.N. Chow, S.N. Koch, M. Sargent, *Semiconductor Laser Physics*, Springer (1994).
7. Eli Kapon, *Semiconductor Lasers I: Fundamentals*, Academic Press (1998).
8. Eli Kapon, *Semiconductor Lasers II: Materials and Structures*, Academic Press (1999).
9. A. Yariv, *Optical Electronics in Modern Communications*, quinta edición, Oxford University Press (1997).
10. C. T. Wang, editor, *Introduction to Semiconductor Technology. GaAs and Related Compounds*, John Wiley & Sons (1990).
11. M. A. Parker, *Physics of Optoelectronics*, CRC Taylor & Francis Group (2005).

Capítulo 4

1. A. Mesmann, *Crystallization Technology Handbook*, John Wiley & Sons (1994).
2. A. W. Vere, *Crystal Growth*, Springer Verlag (2003).
3. A. Pimpinelli and J. Villain, *Physics of Crystal Growth*, Cambridge University Press (1999).
4. I. V. Markov, *Crystal Growth for Beginners*, Springer (2003).
5. M. Herman, W. Richter and H. Sitter, *Epitaxy: Physical Foundation and Technical Implementation*, Springer Verlag (2004).

Capítulo 5

1. M. Herman, W. Richter and H. Sitter, *Epitaxy: Physical Foundation and Technical Implementation*, Springer Verlag (2004).
2. A. Mesmann, *Crystallization Technology Handbook*, John Wiley & Sons (1994).
3. A. W. Vere, *Crystal Growth*, Springer Verlag (2003).
4. P. Capper and M. Mauk, editor, *Liquid Phase Epitaxy of Electronic, Optical and Optoelectronic Materials*, John Wiley (2007).
5. Aicha A. R. Elshabini-Riad and F. D. Barlow III, *Thin Film Technology Handbook*, McGraw-Hill (1998).
6. P. Kordos and J. Novak, editor, *Heterostructure Epitaxy and Devices – HEAD'97. NATO Science Series 3*, High Technology, v.48 (1997).
7. Mahmoud Fallahi S. C. Wang, *Fabrication, Testing and Reliability of Semiconductor Lasers II*, SPIE-International Society for Optical Engineering (1997).
8. G. B. Stringfellow, *Organometallic Vapor-Phase Epitaxy: Theory and Practice*, Academic Press (1989).
9. J. S. Ford, G. J. Davies and W. T. Tsang, *Chemical Beam Epitaxy and Related Techniques*, John Wiley & Sons (1997).
10. Peter S. Zory, Jr, editor, *Quantum Well Lasers*, Academic Press (1993).

11. Eli Kapon, *Semiconductor Lasers I: Fundamentals*, Academic Press (1998).
12. Eli Kapon, *Semiconductor Lasers II: Materials and Structures*, Academic Press (1999).
13. R. F. C. Farrow, *Molecular Beam Epitaxy - Applications to Key Materials*, Noyes Publishing (1995).
14. D. Smith, *Thin-Film Deposition: Principles and Practice*, McGraw-Hill (1995).
15. A. Brednikhina, V. A. Debelov. *Journal of Crystal Growth*. **311**, 15, 666 (2009).

Capítulo 6

1. Piter S. Zory, Jr, editor, *Quantum Well Lasers*, Academic Press (1993).
2. Mahmoud Fallahi S. C. Wang, *Fabrication, Testing and Reliability of Semiconductor Lasers II*, SPIE-International Society for Optical Engineering (1997).
3. Yoshihisa Yamamoto, editor, *Coherence, Amplification and Quantum Effects in Semiconductor Lasers*, John Wiley & Sons (1991).
4. J. T. Verdeyen, *Laser Electronics*, Prentice-Hall (1994).
5. L. A. Coldren, S. W. Corzine, *Diode Lasers and Photonic Integrated Circuits*, John Wiley & Sons (1995).
6. O. Svelto, *Principles of Lasers*, Plenum Press (1998).
7. R. Diehl, editor, *High-Power Diode Lasers. Fundamentals, Technology, Applications*, Springer Verlag (2000).
8. V. L. Granatstein, *Handbook of Lasers*, CRC Press (2000).
9. Takahiro Numai, *Fundamentals of Semiconductor Lasers*, Springer-Verlag (2004).
10. Mitra Dutta, Michael A. Stroschio, Charles H. Townes, *Advances in Semiconductor Lasers and Application to Optoelectronics*, Selected Topics in Electronics and Systems - Vol. 16
11. K. A. Jones, *Introduction to Optical Electronics*, John Wiley & Sons (1987).

12. Peter Vasil'ev, *Ultrafast Diode Lasers. Fundamentals and Applications*, Artech House (1995).
13. G. H. B. Thompson, *Physics of Semiconductor Laser Devices*, John Wiley & Sons (1980).

Por acuerdo del Señor Rector
de la Universidad Autónoma de San Luis Potosí
Lic. Mario García Valdez,
se ordenó la impresión del libro
Láseres de Semiconductor,
cuya edición se terminó de imprimir el de diciembre de 2009
en los Talleres Gráficos de la UASLP.
Se imprimieron 500 ejemplares

Este libro está dedicado a los láseres de semiconductor. Debido a su muy alta eficiencia, rapidez, bajo consumo de energía, tamaño diminuto y confiabilidad, así como por su bajo costo de producción en masa, hoy en día estos dispositivos tienen numerosas y variadas aplicaciones.

Los principales fenómenos y efectos físicos que toman lugar dentro de un láser durante su operación son complejos y su descripción rigurosa requiere de un gran aparato matemático. Debido a que el objetivo de los autores del presente libro es discutir el tema de los láseres de semiconductor utilizando un lenguaje accesible y simple, que no involucre cálculos matemáticos complicados, algunos temas son explicados de una forma sencilla aunque en algunas ocasiones no con el suficiente rigor. En estos casos se privilegió la simplicidad en la exposición por sobre la formalidad y la precisión.

