

El laser de semiconductor

F. Sanz y J.A. Manzanares

Departamento de Termodinámica
Facultad de Física, Universidad de Valencia
46100 Burjasot (Valencia) España

(Recibido el 22 de mayo de 1990; aceptado el 22 de agosto de 1990)

Resumen. Se describe y analiza en forma sencilla el funcionamiento y la evolución de dispositivos que utilizan uniones entre semiconductores para disponer de dos niveles de energía entre los que se puede producir un efecto laser.

PACS: 42.55.Px

1. Introducción

El *laser de semiconductor* o *diodo laser* es, como su nombre indica, un diodo que emite luz laser coherente. Esencialmente, no es más que un bloque de material semiconductor que contiene una unión $p-n$, con las regiones p y n muy densamente dopadas y con una estructura interna más o menos compleja, que se hace funcionar a modo de diodo para producir un efecto laser. En dicha unión los pares electrón-hueco representan los estados excitados y la emisión de fotones se debe a la recombinación de estos pares. En estos materiales, el efecto laser no ocurre por transiciones entre niveles de energía atómicos o de sistemas moleculares sino que, por el contrario, se ha de considerar la estructura de bandas de energía del cristal como un todo.

El laser de semiconductor ha alcanzado hoy día un estado de desarrollo muy avanzado, después de un largo periodo de evolución. Este tipo de láseres había sido discutido ya durante la década de los 50 pero no fue hasta 1961 cuando Basov *et al.* [1] sugirieron el primer laser de semiconductor, en el cual la emisión estimulada de fotones se llevaba a cabo por recombinación de portadores inyectados a través de una unión $p-n$. En 1962, tan sólo tres años después de que Maiman [2] hiciese funcionar el primero de todos los láseres (el laser de rubí), tres grupos de investigadores estadounidenses anunciaron casi simultáneamente que habían conseguido el funcionamiento pulsante del laser de semiconductor a bajas temperaturas [3-6].

La lenta evolución del laser de semiconductor se ha debido sobre todo a problemas específicos relacionados con el desarrollo de una nueva tecnología de semiconductores y, en parte, también está asociado con las características particulares del aparato. Por ejemplo, la emisión estimulada en semiconductores es mucho más intensa, para un determinado grado de inversión, que en láseres de cualquier otro material y el diseño de la cámara resonante es más complejo. A lo largo de este desarrollo, el laser de semiconductor ha sufrido una metamorfosis considerable y ya no se parece nada al primer aparato de 1962. De la simple unión $p-n$ se ha pasado a

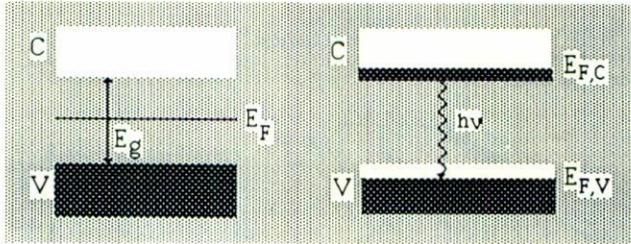


FIGURA 1. Fundamento operativo del laser de semiconductor.

heteroestructuras que contienen varias capas de semiconductores de distintas composiciones. Estas heteroestructuras, introducidas en 1969, causaron un gran impacto y han ido mejorando paulatinamente.

El resultado final de todos los cambios habidos es que se ha pasado de un aparato que funcionaba sólo a bajas temperaturas, a un componente muy práctico, capaz de funcionar a temperatura ambiente en modo continuo.

2. Fundamento teórico

En un semiconductor ideal, a 0 K, entre la última de las bandas de energía que está completamente llena (banda de valencia, V) y la siguiente con mayor energía que está totalmente vacía (banda de conducción, C), existe un gap de energía prohibida de poca anchura, E_g (del orden de 1 eV). Si el cristal se halla dopado con impurezas, tendremos además, dentro del gap de energías, niveles de impurezas donoras cerca del fondo de la banda C y niveles de impurezas aceptoras cerca de la parte superior de la banda V . A una temperatura no nula, en un semiconductor del tipo n , un número considerable de electrones de los niveles de impurezas se hallan excitados en la banda de conducción, mientras que en los de tipo p , los electrones de la banda de valencia se hallan excitados en los niveles de impurezas, resultando así un número considerable de huecos (Fig. 1). La llamada *unión p-n* consiste en una región p y otra n en un monocristal de material semiconductor (Fig. 2).

Si conseguimos excitar algunos pares electrón-hueco, es decir, si conseguimos que algunos electrones de los niveles de más alta energía de la banda V pasen a ocupar los niveles de más baja energía de la banda C , esto implicará una inversión de poblaciones entre ambas bandas. Los ahora electrones de la banda de conducción pueden volver a la de valencia, ocupando uno de los huecos que se han producido, emitiendo un fotón (radiación de recombinación). Esta emisión estimulada es la

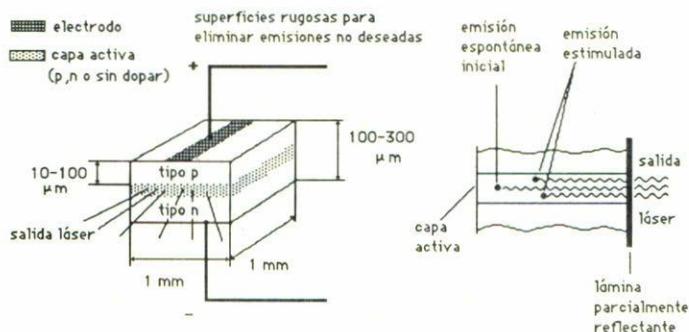


FIGURA 2. Diagrama esquemático de un laser de semiconductor y detalle del proceso de emisión.

que producirá un efecto laser cuando el semiconductor se sitúe en un resonador adecuado.

La excitación de pares electrón-hueco puede realizarse por varios métodos, por ejemplo, por irradiación con luz o por bombardeo con electrones. Sin embargo, el modo más conveniente es por inyección de electrones y huecos en la unión $p-n$ mediante una diferencia de potencial eléctrico (funcionamiento de la unión $p-n$ como diodo). Si aplicamos a la región p un potencial negativo y a la región n uno positivo (polarización inversa) no pasa casi corriente eléctrica, pero si aplicamos a la región p un potencial positivo y a la n uno negativo (polarización directa), la corriente fluirá muy fácilmente. En este último caso, podría decirse que los huecos positivos de la región p se inyectan en la n , al tiempo que los electrones de la n se inyectan en la p . Estos electrones y huecos se difundirán ahora por el material semiconductor hasta llegar a la región de la unión, donde se encuentran y recombinan.

La aniquilación de un par electrón-hueco con emisión de un fotón puede implicar o no, en función de la estructura de bandas del material que forma la unión $p-n$, la emisión simultánea de un fonón. Si se da dicha emisión de un fonón se dice que sucede una *transición indirecta* o, en caso contrario, una *transición directa*.

La recombinación con emisión espontánea es muy efectiva en semiconductores con transiciones directas, como GaAs, pues en ellos la vida media de emisión espontánea es muy corta (10^{-10} - 10^{-9} s) y, por tanto, la amplificación que se obtiene en un sistema con inversión de población como éste es muy alta. Por el contrario, en un semiconductor de transiciones indirectas, como Si o Ge, la emisión espontánea tiene una vida media relativa muy alta (10^{-5} s) y predomina el efecto de los procesos

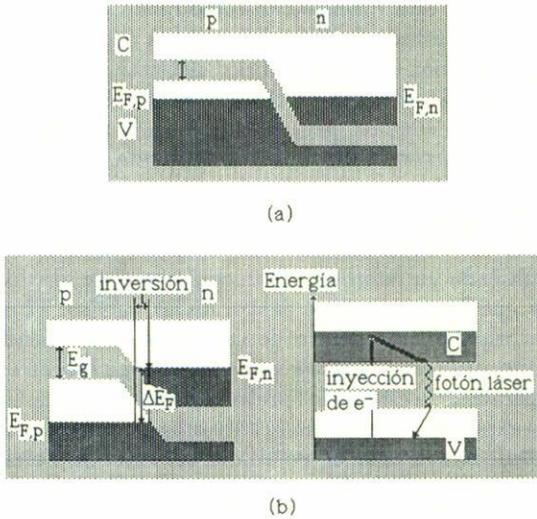


FIGURA 3. Diagrama de energía en una unión p-n: (a) a la que no se aplica ningún voltaje y (b) a la que se aplica un voltaje V en polarización directa.

no radiativos, de modo que no es posible obtener un laser a partir de este tipo de semiconductores.

Como los semiconductores empleados están altamente dopados, el nivel de Fermi del semiconductor del tipo p, $E_{F,p}$, cae dentro de la banda de valencia, y el nivel de Fermi del semiconductor del tipo n, $E_{F,n}$, cae dentro de la banda de conducción. Si no se aplica ningún voltaje, los dos niveles de Fermi tienen la misma energía. Cuando se aplica un voltaje V en polarización directa aparece una separación $\Delta E_F = eV$ entre los niveles, de forma que se produce una inversión de población, dentro de la llamada capa activa (Fig. 3).

La radiación emitida espontáneamente puede ahora ser amplificada por la existencia de otros pares electrón-hueco que se aniquilan (emisión estimulada), o bien ser absorbida en la creación de un nuevo par electrón-hueco (Fig. 2).

La recombinación de pares electrón-hueco puede también llevarse a cabo a través de procesos no radiativos en los que la energía se cede a los fonones de la red o a portadores libres. Si bien, en los materiales empleados en los diodos laser (basta con que sea así en los que forman la capa activa) los procesos radiativos dominan ampliamente sobre los no-radiativos.

En un semiconductor en equilibrio a una $T > 0$ K, las probabilidades de ocupación de las bandas V y C vienen dadas por la expresión

$$F = \left[1 + \exp \left(\frac{E - \mu}{kT} \right) \right]^{-1}$$

donde μ es el nivel de Fermi.

Cuando la unión $p-n$ está funcionando como diodo en polarización directa, ya no está en equilibrio y el estudio teórico de la misma debería realizarse a través de probabilidades de ocupación de no-equilibrio, distintas a las expresadas anteriormente. Ahora bien, cuando se excita un electrón dentro de la banda de conducción, las interacciones de dicho electrón con los fonones de la red disminuirán su energía hasta que se corresponda con la del fondo de la banda de conducción. Esta transición dentro de la banda ocurre en un tiempo muy pequeño, del orden de 10^{-13} s. Evidentemente, lo mismo le sucede al hueco excitado dentro de la banda de valencia. Una vez situados en los extremos de las bandas respectivas, el electrón y el hueco se recombinan emitiendo un fotón en un proceso que dura alrededor de 10^{-9} s, es decir, mucho más que las transiciones dentro de las bandas. El equilibrio térmico se reestablece tan rápidamente dentro de cada una de las bandas que podemos considerar que, aunque el semiconductor como un todo no esté en equilibrio térmico, las bandas por separado sí lo están y las probabilidades de ocupación de cada una de las bandas pueden expresarse tal y como se hace para un semiconductor en equilibrio sin más que sustituir el concepto de equilibrio del nivel de Fermi por su análogo de no-equilibrio, el cuasinivel de Fermi. Así, al sustituir en la expresión anterior μ por $E_{F,C}$ (cuasinivel de Fermi de la banda V), se obtendría la probabilidad de ocupación de la banda V y, al sustituir μ por $E_{F,V}$ (cuasinivel de Fermi de la banda C), se obtendría la probabilidad de ocupación de la banda C .

Una condición necesaria para el funcionamiento del laser es que el número de emisiones estimuladas debe ser mayor que el número de fotones absorbidos (siendo un exceso necesario para compensar las pérdidas en la cavidad resonante). La velocidad de emisión estimulada será proporcional al producto de la probabilidad de ocupación del nivel superior, F_C , y la probabilidad de no-ocupación del nivel inferior, $(1 - F_V)$. Recíprocamente, la velocidad de absorción de fotones será proporcional al producto de la probabilidad de ocupación de nivel inferior, F_V , y la probabilidad de no-ocupación del nivel superior, $(1 - F_C)$. Así, para conseguir la emisión estimulada, se debe satisfacer

$$F_C(1 - F_V) - F_V(1 - F_C) > 0$$

la cual implica que

$$F_C > F_V$$

es decir,

$$E_{F,C} - E_{F,V} > h\nu > E_g.$$

La frecuencia de la radiación emitida debe estar comprendida entre las que corresponden a la anchura del gap de energía entre bandas y la diferencia entre las energías de los cuasiniveles de Fermi de las bandas C y V , los cuales dependen del número de electrones excitados. Esto implica la necesidad de excitar un mínimo número de electrones para asegurar que $E_{F,C} - E_{F,V} > E_g$, ya que, de no ser así, la ganancia del semiconductor sería nula y no se podría observar efecto laser.

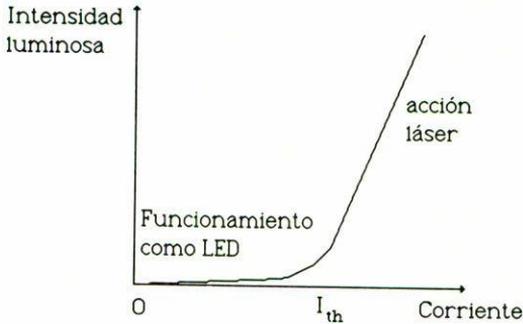


FIGURA 4. Curva de respuesta del laser de semiconductor: intensidad de luz emitida frente a la corriente de inyección.

3. Homoestructuras

Los primeros láseres de semiconductor constaban simplemente de una *homounión* u *homoestructura*, es decir, en ellos la unión laser se formaba con las regiones del tipo *p* y *n* del mismo material (normalmente GaAs). A esta unión se fijaban dos láminas paralelas entre sí y perpendiculares a la unión, que reflejaban parte de la luz generada de forma que ésta volvía al semiconductor. Se formaba así un resonador del tipo Fabry-Perot. Realmente, el índice de refracción del semiconductor era tan alto que ya poseía una gran reflectividad ($\sim 33\%$) en la interfase aire-semiconductor y algunos diodos laser no requerían de estas láminas. La diferencia principal con otros tipos de láseres, en lo que al resonador se refiere, estaba en que la longitud del Fabry-Perot podía hacerse mucho más pequeña (del orden de décimas de milímetro) gracias a que la emisión estimulada es mucho más rápida en semiconductores.

En la Fig. 4 se muestra la forma en que varía la intensidad de la luz emitida con la corriente de inyección. En esta figura se ve cómo la intensidad crece muy rápidamente a partir de cierta corriente umbral, I_{th} , mientras que por debajo de I_{th} es muy débil. Cuando la corriente es menor que I_{th} , la directividad de la luz es pobre y la anchura espectral grande, funcionando de forma similar a un LED.* Por encima de I_{th} , por el contrario, la luz laser tiene una anchura espectral muchísimo menor y una directividad mucho mayor.

Para el GaAs, I_{th} es del orden de 10^9 A/m² a temperatura ambiente y de 10^7 A/m² a la temperatura de nitrógeno líquido. Aproximadamente I_{th} decrece con T como $\exp(T/T_0)$, dependiendo T_0 del semiconductor. Esto se debe a las altas pérdidas que causa la forma de la cavidad pues el haz, al abrirse, entra en las regiones *p* y *n* donde la absorción es dominante. Además, la región donde se da la

*Un *diodo emisor de luz* (LED) es una diminuta fuente de luz incoherente que se basa en este mismo fenómeno. Su estructura es similar a la de un diodo laser, si bien suele carecer de los elementos estructurales y de las láminas parcialmente reflectantes.

inversión no está perfectamente definida sino que presenta unos límites muy poco definidos en la dirección en que se difunden los portadores de carga.

Recordemos que la ganancia del semiconductor depende de $F_C(1 - F_V) - F_V(1 - F_C)$ y que, al disminuir la temperatura, $F_C(1 - F_V)$ aumenta y $F_V(1 - F_C)$ disminuye. Resulta así que la ganancia aumenta rápidamente al disminuir la temperatura. Ésta es la razón de que los láseres de homoestructura sólo funcionen en modo continuo a bajas temperaturas.

4. Heteroestructuras

Para salvar las dificultades mencionadas en la Sec. 3, se utilizan láseres con heterouniones. Si la capa de semiconductor donde se produce la inversión se sitúa entre dos capas de otro material semiconductor con mayor gap de energías, los portadores inyectados y los fotones emitidos quedan confinados por completo dentro de la capa activa. El hecho de que los portadores queden confinados se debe simplemente a las barreras de potencial que existen en las interfases con el material de mayor gap. El confinamiento de la luz se debe a que la capa activa tiene un mayor índice de refracción que las capas del semiconductor de mayor gap y, por tanto, se forma una guía de ondas. Esta estructura de multicapas puede complicarse muchísimo con el objeto de mejorar las características de operación del laser pero, en general, se trata de crear una guía de ondas para la luz y reducir el tamaño de la capa activa [8].

Los láseres de *heterounión simple* son aquéllos en los que la capa activa limita por una parte con un material que tiene una banda prohibida distinta. Es decir, la capa activa se halla situada entre dos capas de diferente composición química, usualmente GaAs y GaAlAs, que tienen diferentes bandas prohibidas. Los láseres de heterouniones simples tienen mejores propiedades que los de homouniones y son muy usados en modo pulsante para producir altas potencias de pico. Sin embargo, estos láseres son relativamente ineficientes y tienen altas I_{th} , lo cual los hace inadecuados para funcionamiento continuo.

Los láseres de *heterounión doble* tienen una capa activa limitada por dos capas de materiales diferentes, por ejemplo, GaAs entre dos capas de GaAlAs. Estos láseres han demostrado ser los más adecuados para funcionamiento continuo y son muy usados en sistemas con fibras ópticas. No pueden producir pulsos de altas potencias de pico, pero pueden operar en modo continuo a potencias del orden de 10^{-2} W a temperatura ambiente. Sus capas activas son mucho más delgadas que en las heterouniones simples ($0.1 \mu\text{m} - 0.3 \mu\text{m}$), de forma que sus I_{th} son lo suficientemente bajas ($\sim 10^7$ A/m²).

En uno de estos láseres con heterounión doble y estructura plana, la luz en la dirección perpendicular a la unión se concentra cerca de la capa activa, pero la luz en la dirección transversal se dispersa mucho y su distribución de intensidades es variable incluso con pequeños cambios en la corriente. La anchura del electrodo superior se reduce con el objeto de concentrar la acción laser en el centro, de forma que nos queda lo que se denomina una *estructura de banda*. Esto no consigue impedir la dispersión horizontal, pero sí mejora la calidad del haz, en cuanto que limita el número de modos espaciales y, además, disminuye más I_{th} .

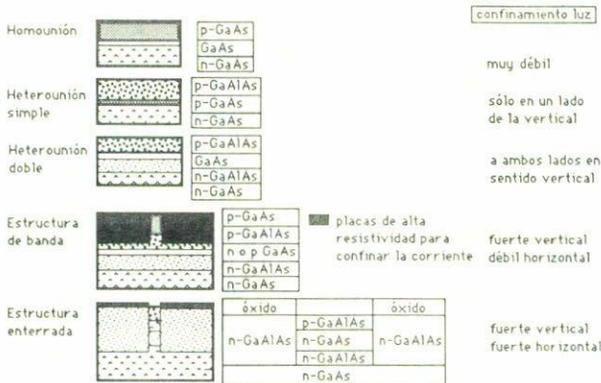


FIGURA 5. Ejemplos de estructuras de diodos laser.

Por estos motivos, se han buscado muchas estructuras diferentes de laser que confinan la capa activa no sólo a una pequeña profundidad sino también a una pequeña anchura. Como ejemplo de tal estructura se muestra en la Fig. 5, junto con otras estructuras más simples de diodos laser, una *heteroestructura enterrada*. Después de formar la heterounión doble de GaAlAs con capa activa de GaAs, sólo se deja la parte central, cubierta por una máscara de unas cuantas micras de anchura, y se elimina el resto por corrosión. Las partes corroídas se reemplazan por crecimiento de un cristal de GaAlAs(n). De esta forma la capa activa está rodeada por todos los lados (arriba, abajo, derecha e izquierda) por GaAlAs y, como la inyección de portadores se da sólo en la parte central, el modo laser dentro de la heterounión es como una radiación de microondas en una guía rectangular.

La ganancia en los diodos laser es muy alta, de forma que las cavidades resonantes son pequeñas (de unos 100 μm de largo). Estas cavidades tienen formas variadas en función de la naturaleza de la capa activa. Los diodos laser pueden producir haces por ambas caras aunque normalmente tienen una lámina totalmente reflectante que limita la salida a una sola cara.

5. Aplicaciones

El campo de aplicación para los diodos laser se ha ampliado rápidamente y hoy día ya no se restringe sólo a su empleo como fuente de luz en las comunicaciones por fibra óptica, sino que incluye además muchos otros campos del tratamiento de información [9]. El interés en las comunicaciones por fibras ópticas fue una motivación para el desarrollo de la tecnología de los diodos laser durante los 70

y el comienzo de los 80. Los primeros sistemas de fibras ópticas operaban entre 0.8 y 0.9 μm (láseres de GaAlAs), pero progresivamente se han ido usando sistemas de mayores longitudes de onda para disminuir la atenuación. Por ejemplo, la longitud de onda de 1.3 μm se usa ampliamente en los sistemas comerciales habituales, ya que las fibras actuales tienen una amplia anchura de banda y muy pocas pérdidas sobre dicha longitud.

Las comunicaciones por fibras ópticas siguen siendo el uso dominante de los láseres de InGaAsP, y parece ser que esto va a continuar así. Sin embargo, a esta aplicación sólo se dedica una pequeñísima parte de los láseres de GaAlAs, pues éstos tienen otras aplicaciones en tratamiento de información, incluyendo la lectura de información previamente grabada y su grabación en impresoras o sistemas de almacenamiento de datos.

Por supuesto, también es importante su aplicación en compact-disk, ya que sólo en 1985 se vendieron más de un millón de estos diodos y hoy día los diodos laser dominan las aplicaciones en almacenamiento de datos y sistemas compact-disk y videodisk. Asimismo, las impresoras laser (que siguiendo una técnica similar a la de las fotocopiadoras, escriben sobre un cilindro electrostático para después transferirlo a papel normal) también usan diodos laser. Asimismo, también se utilizan diodos laser para la lectura de códigos de barras y en una gran variedad de sistemas de comunicación atmosféricos.

Las aplicaciones militares, incluidas la simulación de guerras y el rastreo de objetivos, emplean láseres de heteroestructuras simples y baterías de láseres, en modo pulsante. Los misiles aire-aire incluyen diodos laser de búsqueda de objetivos, que detectan las distancias a los blancos militares, y sincronizan la detonación de la cabeza nuclear cuando el misil está lo suficientemente cerca del blanco como para destruirlo.

Referencias

1. N.G. Basov, O.N. Kroklin and Y.M. Popov, *Sov. Phys. JETP*, **13** (1961) 1320.
2. T.H. Maiman, *Phys. Rev.* **123** (1961) 1145.
3. R.N. Hall, G.E. Fenner, J.D. Kingsley, T.J. Soltys and R.O. Carlson, *Phys. Rev. Lett.* **9** (1962) 366.
4. M.I. Nathan, W.M. Drunke, G. Burns, F.H. Dill and G.J. Lasher, *Appl. Phys. Rev.* **1** (1962) 62.
5. T.M. Quist, R.H. Rediker, R.J. Keyes, W.E. Krag, B. Lax, A.L. McWhorter and H.J. Zeiger, *Appl. Phys. Rev.* **1** (1962) 1145.
6. N. Holonyak Jr. and S.F. Bevacqua, *Appl. Phys. Rev.* **1** (1962) 82.
7. O. Svelto, *Principles of lasers*, Plenum Press (1982), pp. 254-255.
8. G.H.B. Thompson, *Physics of semiconductor laser devices*, John Wiley & Sons (1985), pp. 1-2.
9. J. Hecht. *The laser guidebook*, McGraw-Hill Co. (1986), p. 247.

Abstract. The working principles and the development of devices using semiconductor junctions in order to generate two energy levels which can be used to produce a laser effect, are described and analysed in an easy way.