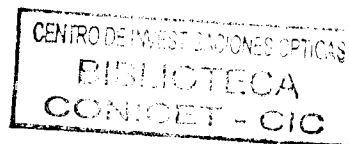


**CURSO
LASER Y OPTICA
EN INGENIERIA**





EL LASER

M. Garavaglia

1. La emisión estimulada

Hacia 1916 estaba perfectamente establecido que los átomos, moléculas e iones sólo pueden encontrarse en estados energéticos definidos, y que pueden interactuar con el campo electromagnético emitiendo o absorbiendo radiación. Estos dos procesos de interacción se cumplen de acuerdo con la relación de Bohr: $E_2 - E_1 = h\nu_{21}$.

Sin embargo, es en 1916 cuando A. Einstein¹, mediante deducciones de índole termodinámica, establece dos conceptos fundamentales: la noción de probabilidad de transición y la de emisión estimulada de radiación.

Para simplificar el análisis, supóngase que un átomo posee sólo dos niveles simples de energía, E_1 y E_2 , siendo $E_2 > E_1$. Si un conjunto N_1 de dichos átomos, cuyos electrones de valencia se encuentran en el nivel fundamental de energía E_1 , interactúa con el campo de radiación de densidad $\rho(\nu_{21})$, se producirá el fenómeno de absorción de radiación, por el cual, según Einstein, un número igual a $N_1\rho(\nu_{21})B_{12}$ pasarán al nivel de energía E_2 . El producto $\rho(\nu_{21})B_{12}$ es la probabilidad de absorción por unidad de tiempo. A su vez, los N_2 átomos cuyos electrones se encuentran en el nivel de energía E_2 podrán emitir radiación según los dos procesos siguientes: emisión espontánea y emisión estimulada. Ambas designaciones se deben a Einstein. La primera se identifica con la única forma de emisión que se conocía antes de 1916 y por la cual los átomos emiten radiación en forma totalmente independiente unos de otros. La emisión estimulada de radiación, en cambio, se producirá por la interacción del campo de radiación de densidad $\rho(\nu_{21})$, de la misma manera que en el caso de la absorción. Una propiedad importantísima de la radiación estimulada es la de estar en fase con la radiación estimulante.

Por los dos procesos de emisión, un número de átomos igual a:

$$N_2 A_{21} + N_2 \rho(\nu_{21}) B_{21}$$

pasarán del nivel excitado de energía E_2 al nivel fundamental de energía E_1 emitiendo al hacerlo radiación de frecuencia ν_{21} . A_{21} es la probabilidad de transición espontánea y el producto $\rho(\nu_{21})B_{21}$ es la probabilidad de transición estimulada. En ambos casos, por unidad de tiempo. A_{21} , B_{21} y B_{12} se llaman coeficientes de Einstein.

Si el conjunto de átomos se encuentra confinado en equilibrio térmico en un recinto adiabático de paredes perfectamente reflectoras y calentadas a la temperatura T , entonces las poblaciones N_2 y N_1 de los niveles de energía E_2 y E_1 estarán relacionados de acuerdo con la ley de Boltzmann, al tiempo que el campo de radiación que interactúa con los átomos tendrá una densidad $\rho(\nu_{21})$ correspondiente a la del cuerpo negro calentado a la temperatura T , según la ley de Planck. Las Figuras 1 y 2 representan la densidad de radiación $\rho(\nu)$ emitida por el cuerpo negro y la distribución de Boltzmann, respectivamente. Los máximos de las sucesivas curvas $\rho(\nu)$ para las diferentes temperaturas T , obedecen a la ley del Corrimiento de Wien, expresada por:

$$\lambda_M T = 2,897 \cdot 10^{-3} \text{ mK} \quad .$$

Luego, si existe equilibrio térmico, la cantidad de energía que los átomos absorben debe ser igual a la que emiten, es decir:

$$N_2 A_{21} + N_2 \rho(\nu_{21}) B_{21} - N_1 \rho(\nu_{21}) = 0 \quad . \quad (1)$$

Las poblaciones N_2 y N_1 se relacionan según:

$$N_2/N_1 = e^{-(E_2-E_1)/kT} \quad , \quad (2)$$

donde k es la constante de Boltzmann.

Combinando las ecuaciones (1) y (2) y recordando que $E_2 - E_1 = h\nu_{21}$, se puede encontrar la expresión de la densidad de radiación:

$$\rho(\nu_{21}) = \frac{A_{21}}{e^{h\nu_{21}/kT} B_{12} - B_{21}} \quad . \quad (3)$$

Comparando esta expresión con la ley de Planck sobre la radiación de un cuerpo negro en la frecuencia ν_{21} , que es:

$$\rho(\nu_{21}) = \frac{8\pi h \nu_{21}^3}{c^3} \cdot \frac{1}{e^{h\nu_{21}/kT_{-1}}} \quad , \quad (4)$$

se puede observar su similitud. De ellas, y para evitar incongruencias, Einstein llegó a la conclusión que los coeficientes A_{21} , B_{21} y B_{12} deben relacionarse así:

$$B_{21} = B_{12} \quad , \quad (5)$$

$$A_{21} = \frac{8\pi h \nu^3}{c^3} \cdot B_{21} \quad .$$

Luego, la introducción del proceso de emisión estimulada de radiación permite interpretar correctamente la interacción entre materia y el campo de radiación. De no ser tenido en cuenta dicho proceso, y con la sola contribución de la emisión espontánea y de la absorción de radiación, se llegaría a incongruencias reñidas con los resultados experimentales.

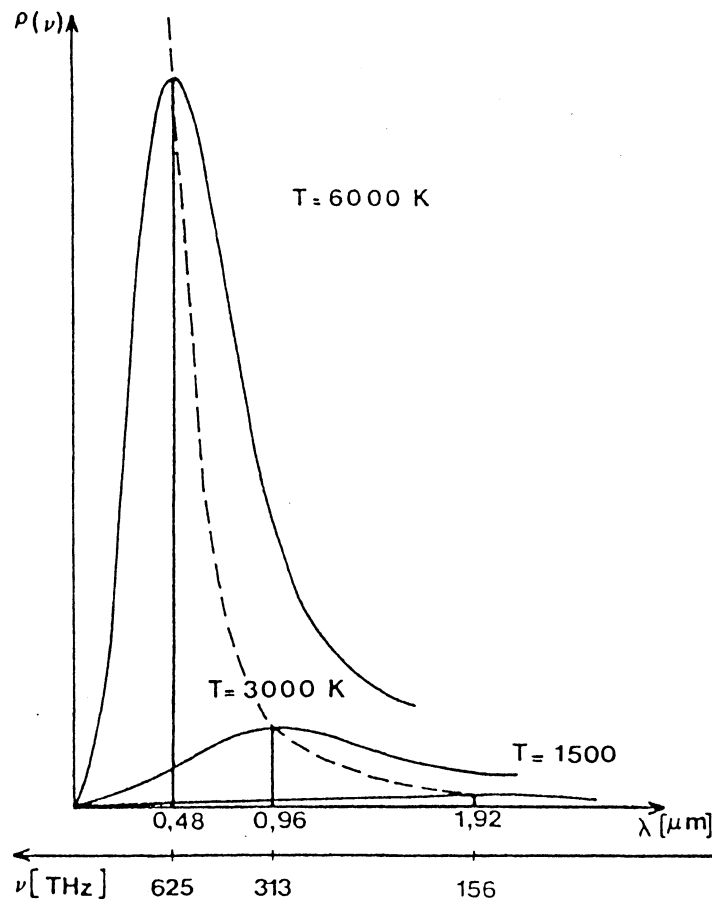


Figura 1. Densidad de radiación $\rho(\nu)$ emitida por el cuerpo negro según la temperatura T y la ley de Wien.

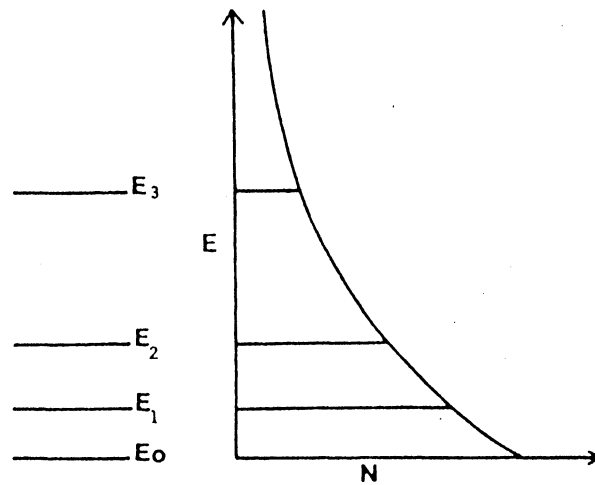


Figura 2. Ley de Boltzmann de la distribución de población.

2. Dificultades para poner en evidencia la emisión estimulada.

Las grandes dificultades encontradas desde la publicación de los trabajos de A. Einstein, en 1917, para poner de manifiesto la emisión estimulada, pueden agruparse en dos categorías: unas dependían de los resultados de la termodinámica y del electromagnetismo, y las otras, del desarrollo de las ideas sobre el átomo. En efecto, aquellos estudios concluían, por un lado, en que la emisión estimulada debía ser extremadamente difícil de observar en razón de que su intensidad sería despreciable frente a la de la emisión espontánea producida en la misma transición electrónica entre los niveles de energía del átomo. Por el otro lado, pesaban las dificultades provenientes del concepto mismo del átomo que por ese entonces estaba en boga, ya que en su esquema resultaba complicado hacer encajar esta forma de emisión estimulada de radiación.

El tratamiento de estas últimas dificultades será abordado en la próxima sección, la que estará destinada a los intentos experimentales por poner de manifiesto la emisión estimulada.

En esta sección se dispensará atención a las dificultades conceptuales planteadas desde los puntos de vista de la termodinámica y del electromagnetismo. Al recinto de paredes perfectamente reflectoras cuya temperatura es T y dentro del cual se produce la interacción entre los átomos y el campo de radiación $\rho(\nu)$ del cuerpo negro, se lo analizará aquí como una cavidad resonante. Entonces, a aquellas frecuencias contenidas

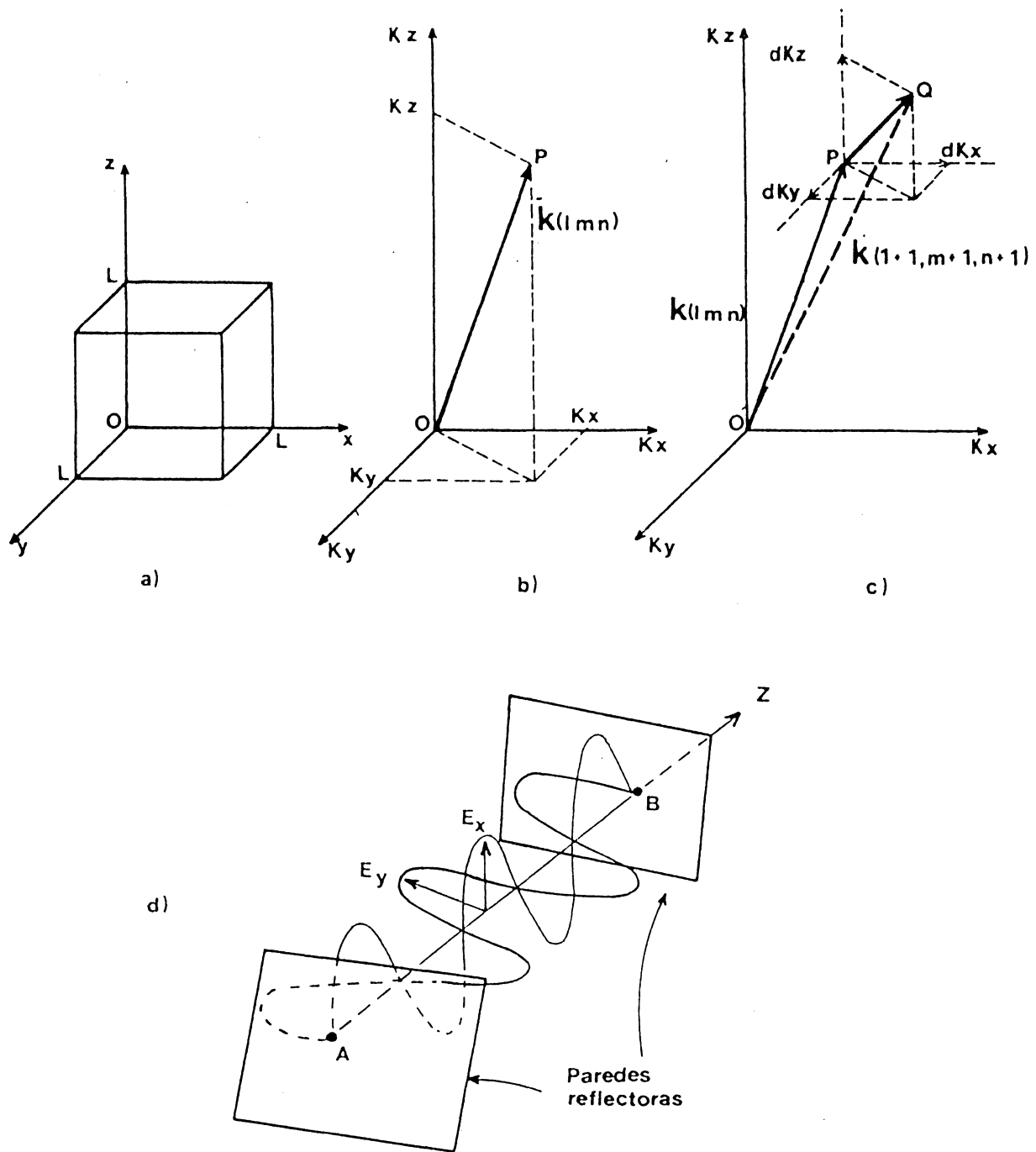


Figura 3. Vínculos entre la geometría de la cavidad y las ondas del campo de radiación. a) Cavidad cúbica de lado L . b) Vector propagación de la onda que excita al modo resonante (l, m, n) . c) Cambios en la longitud de onda de la radiación provocan la excitación del modo $(l+1, m+1, n+1)$. d) Las ondas que se propagan según el eje Z , tienen sus planos de polarización perpendiculares.

en dicho campo de radiación, la cavidad resonará siguiendo simples relaciones geométricas entre sus dimensiones y las longitudes de onda correspondientes a tales frecuencias. Dentro de la cavidad se sustentarán, entonces, ondas estacionarias espaciales. Cada forma de resonar la radiación en las respectivas longitudes de onda con la cavidad, se denomina modo de resonancia.

Supóngase una cavidad cúbica de lado L y sea $K = 2\pi/\lambda$ el vector propagación de las ondas de las radiaciones. La Figura 3 muestra la vinculación entre las características geométricas de la cavidad y las componentes de los vectores de propagación en el espacio de los K . Las componentes cartesianas de K cumplen las siguientes condiciones:

$$K_x = \frac{2\pi}{L} l \quad , \quad K_y = \frac{2\pi}{L} m \quad , \quad K_z = \frac{2\pi}{L} n \quad , \quad (6)$$

las que describen el modo resonante (l,m,n) cuando l , m y n son números enteros. La magnitud del vector K en función de sus componentes es:

$$K = \sqrt{K_x^2 + K_y^2 + K_z^2} \quad . \quad (7)$$

Como cada modo resonante se caracteriza por la terna de números enteros (l,m,n) , resulta que una variación de ellos en una unidad producirá un cambio en las componentes de K que puede calcularse por diferenciación de las expresiones (6):

$$dK_x = \frac{2\pi}{L} dl \quad , \quad dK_y = \frac{2\pi}{L} dm \quad , \quad dK_z = \frac{2\pi}{L} dn \quad ,$$

haciendo $dl=dm=dn=1$. Incrementando con estos valores dK_x , dK_y y dK_z , los de K_x , K_y y K_z de la fórmula (7), puede calcularse el vector K correspondiente al nuevo modo resonante $(l+1, m+1, n+1)$ y, por lo tanto, determinar la frecuencia ν o la longitud de onda λ de la radiación que, dentro de la cavidad, excitará dicho modo de resonancia sustentando una onda estacionaria.

En el espacio K puede asociarse a cada modo un volumen igual a:

$$V_1 = dK_x dK_y dK_z = \left(\frac{2\pi}{L} \right)^3 \quad (8)$$

y, por lo tanto, puede determinarse el número de modos N_K cuyos vectores K adoptan valores entre 0 y K , dividiendo el volumen total que ellos ocupan por aquél que ocupa un solo modo V_1 . Obviamente, el volumen total es igual al de la esfera de radio K en el espacio de los K , el que vale:

$$V_K = \frac{4}{3} \pi K^3$$

Luego, en el espacio de los K , el número de modos resonantes será igual a:

$$N_K = 2 \frac{V_K}{V_1} = \frac{K^3 L^3}{3 \pi^2} \quad (9)$$

El factor 2 que se introdujo para determinar N_K en la expresión (9), responde a la existencia de los dos planos de polarización de la radiación para las ondas estacionarias sustentadas en cada modo resonante. La Figura 3.d) muestra un esquema de dos ondas que se propagan entre las paredes reflectoras y cuyos planos de polarización son perpendiculares entre sí.

A partir de la expresión (9) y teniendo en cuenta que $K = 2\pi\nu/c$, se puede expresar el número de modos resonantes que excitan el volumen $V=L^3$ de la cavidad y cuyas frecuencias se encuentran comprendidas entre $\nu=0$ y ν , por la expresión:

$$\frac{N_\nu}{V} = \frac{8\pi\nu^3}{3c^3} \quad (10)$$

Ahora puede definirse la densidad de modos resonantes $d(\nu)$ por unidad de frecuencia, según:

$$d(\nu) = \frac{1}{V} \frac{dN_\nu}{d\nu} = \frac{8\pi\nu^2}{c^3} \quad (11)$$

Como el recinto de la cavidad está inundado por la radiación del cuerpo negro a la temperatura T y cuya densidad se expresa por $\rho(\nu)$, se puede calcular la energía media por modo \bar{E} , según la definición:

$$\bar{E} = \frac{\rho(\nu)}{d(\nu)} \quad (12)$$

Reemplazando en la fórmula (12) los valores ya obtenidos de $\rho(\nu)$ y $d(\nu)$, resulta:

$$\bar{E} = \frac{\rho(\nu)}{d(\nu)} = \frac{\frac{8\pi h \nu^3}{c^3} \cdot (e^{h\nu/kT} - 1)^{-1}}{\frac{8\pi \nu^2}{c^3}} = \frac{h\nu}{e^{h\nu/kT} - 1} \quad (13)$$

Siguiendo con detalle el desarrollo de la fórmula (3) y su comparación con la (4), resulta que el denominador de la expresión (13) es igual a la diferencia entre el número de transiciones de absorción $1 \rightarrow 2$ y el de emisión estimulada $2 \rightarrow 1$.

Por la ley de Wien resulta que si ν_{12} coincide con la frecuencia ν_M del máximo de la curva $\rho(\nu)$ a la temperatura T correspondiente, el denominador de la ecuación (13) vale:

$$e^{h\nu_M/kT} - 1 = e^{hc/k\lambda_M T} - 1 = e^{10^3 hc/2,897k} - 1 \approx 150 \quad (14)$$

Igualmente, a partir de la fórmula (3), puede escribirse:

$$\frac{A}{\rho(\nu) B} = e^{h\nu/kT} - 1 \approx 150 \quad (15)$$

lo que significa que siempre será mayor - y en un factor aproximadamente a 150 - el número de transiciones espontáneas $2 \rightarrow 1$, que la suma de las transiciones de absorción $1 \rightarrow 2$ y las de emisión estimulada $2 \rightarrow 1$.

El desarrollo de los argumentos considerados hasta aquí, permite resumirlos de la siguiente manera: A medida que aumenta la temperatura T de la cavidad crece la densidad de radiación $\rho(\nu)$ en la frecuencia de su máximo ν_M , de tal suerte que, si su volumen V se mantiene inalterado al igual que el ancho de banda $d\nu$ de las respectivas transiciones, se incrementa el número de modos resonantes sustentados por ella, de acuerdo con las expresiones (10) y (11). Al mismo tiempo, se incrementa la energía media por modo \bar{E} y como la relación entre el número de transiciones de absorción $1 \rightarrow 2$ y de emisión estimulada $2 \rightarrow 1$, es decir, aquéllas que se producen dependiendo de la existencia del campo de radiación $\rho(\nu_M)$ en la cavidad, es constante a pesar de los incrementos de T - según la fórmula (14) - resulta que, si bien dichas transiciones están inmersas en la

masa de transiciones de emisión espontánea en una proporción muy desfavorable para su observación - de acuerdo con la expresión (15) - como la densidad de modos $d(\nu)$ también crece con la temperatura, queda, no obstante, el siguiente balance neto: La relación entre la proporción de transiciones de emisión espontánea y aquéllas que dependen de la existencia de $\rho(\nu)$ con respecto de la energía media por modo, es mayor para las longitudes de onda más largas. De manera que ante una misma cantidad de energía por modo disponible dentro de la cavidad, será más fácil observar las transiciones de emisión estimulada en el infrarrojo que en el visible. En la Tabla I se resumen numéricamente estas conclusiones para tres valores típicos de la temperatura T de las paredes de la cavidad.

Obviamente, la observación de la emisión estimulada será mucho más fácil en la región de las microondas del espectro electromagnético donde la longitud de onda de la radiación puede ser, por ejemplo, de 1 cm. Ello es así, porque la densidad de modos $d(\nu)$ será solamente de $8,3 \cdot 10^{-4} \text{ m}^{-3} \text{ Hz}^{-1}$, si se supone que la cavidad se sigue comportando como un cuerpo negro cuyas paredes se encuentran ahora a temperatura de 0,2 K. En este caso, la relación entre el número de transiciones totales y la energía media por modo será aproximadamente de 10^{27} J^{-1} , es decir, cuatro y cinco órdenes de magnitud mayor que en la región del infrarrojo de $\lambda = 2 \mu\text{m}$ y del visible de $\lambda = 0,5 \mu\text{m}$, respectivamente.

TABLA I

Aspectos comparativos de una cavidad a tres temperaturas típicas.

	<u>Caso 1</u>	<u>Caso 2</u>	<u>Caso 3</u>
T	6000 K	3000 K	1500 K

$\rho(\nu)$

λ_M	0,48 μm	0,96 μm	1,92 μm
ν_M	$6,25 \cdot 10^{14} \text{ Hz}$	$3,13 \cdot 10^{14} \text{ Hz}$	$1,56 \cdot 10^{14} \text{ Hz}$
$d(\nu)$	$363.588 \frac{1}{\text{m}^3 \cdot \text{Hz}}$	$90.897 \frac{1}{\text{m}^3 \cdot \text{Hz}}$	$22.724 \frac{1}{\text{m}^3 \cdot \text{Hz}}$
\bar{E}	$0,2778 \cdot 10^{-20} \text{ J}$	$0,1391 \cdot 10^{-20} \text{ J}$	$0,0694 \cdot 10^{-20} \text{ J}$
$\frac{N^\circ \text{ de Transiciones}}{d(\nu)}$	$= 4,125 \cdot 10^{-4}$	$16,502 \cdot 10^{-4}$	$66,010 \cdot 10^{-4}$
$\frac{N^\circ \text{ de Transiciones}}{\bar{E}}$	$= 5,40 \cdot 10^{22} \text{ J}^{-1}$	$10,78 \cdot 10^{22} \text{ J}^{-1}$	$21,61 \cdot 10^{22} \text{ J}^{-1}$
$\frac{\bar{E}}{d(\nu)}$	$= 0,764 \cdot 10^{-26} \text{ J} \cdot \text{m}^3 \cdot \text{Hz}$	$1,530 \cdot 10^{-26} \text{ J} \cdot \text{m}^3 \cdot \text{Hz}$	$3,054 \cdot 10^{-26} \text{ J} \cdot \text{m}^3 \cdot \text{Hz}$

3. Intentos por poner en evidencia la emisión estimulada.

Como la emisión estimulada de radiación es un concepto precuántico, no se la podía relacionar con ningún fenómeno experimental para comprobar directamente su existencia, debido al modelo de oscilador armónico empleado para describir el átomo.

Recién en 1924, H.A. Kramers², y en 1928 R. Ladenburg³ mostraron que la emisión estimulada de radiación puede relacionarse, mediante el modelo cuántico del átomo, con las expresiones del índice de refracción o del coeficiente de absorción de un gas. La posibilidad de hacer experimentos de dispersión o de absorción con medios ópticos ideales, como los gases a bajas presiones, donde pueden desprejarse las interacciones entre los átomos, y en los que se podía variar la población de los niveles mediante la aplicación de una excitación conveniente, hizo que rápidamente se empezase a trabajar para poner de manifiesto la emisión estimulada.

Con respecto a las investigaciones tendientes a poner en evidencia la emisión estimulada basadas en la determinación del coeficiente de absorción, se deben mencionar a los primeros en abordar el problema: C. Füchtbauer y R. Ladenburg⁸. En los primeros años de la década del 20, ellos y sus respectivos colaboradores hallaron, por separado, una relación entre el coeficiente de absorción $\alpha(\nu)$ de un material y los coeficientes de Einstein para las transiciones entre niveles de energía de los átomos o moléculas que constituyen dicho material y con la distribución de tales átomos o moléculas en sus diferentes niveles de energía. El coeficiente de absorción $\alpha(\nu_0)$ se define teniendo en cuenta la variación de la intensidad de un haz de luz de frecuencia ν_0 que atraviesa un dado espesor de material. La ley de Beer que rige este proceso es la siguiente:

$$I(x, \nu_0) = I(0, \nu_0) e^{-\alpha(\nu_0) \cdot x}, \quad (16)$$

es decir que, si incide una onda de intensidad $I(0, \nu_0)$, después de atravesar un espesor x de un material de coeficiente de absorción $\alpha(\nu_0)$, la intensidad observada será $I(x, \nu_0)$.

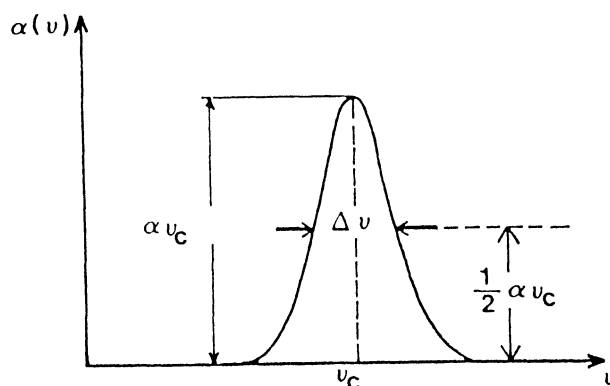


Figura 4. Absorción a través de una línea espectral de frecuencia central ν_C .

Füchtbauer y Ladenburg combinaron esta ley con los coeficientes de Einstein en el entorno de una línea espectral de frecuencia ν . Como las líneas espectrales poseen, como se verá más adelante, un cierto ancho natural y, además, pueden verse ensanchadas por efectos térmicos, de presión o de influencia de campos eléctricos o magnéticos, definieron la absorción integral a través de la línea espectral, así:

$$\int \alpha(\nu) d\nu = h\nu \cdot B_{12}(N_1 - N_2) / c \quad (17)$$

La Figura 4 muestra la forma típica de variación de $\alpha(\nu)$ con ν en una línea espectral de absorción. Cuando el material está en equilibrio térmico a la temperatura T , la distribución de los átomos que lo constituyen entre los diferentes niveles de energía sigue, como ya se dijo antes, la ley de Boltzmann. En el caso particular de los niveles 1 y 2, para los cuales $E_2 > E_1$, las poblaciones se repartirán de acuerdo con la expresión:

$$N_2/N_1 = e^{-(E_2 - E_1)/kT} \quad (18)$$

De esta manera, el término de la derecha de la ecuación (17) resulta siempre positivo. Es decir, un material en equilibrio térmico atenúa siempre por absorción la intensidad de un haz de luz que lo atraviese. Aunque no puede descartarse que no ocurran procesos locales de emisión estimulada de radiación en el interior del material, la absorción será dominante y el resultado neto será la atenuación del haz.

Se hicieron muchos experimentos tratando de invertir el signo de la absorción en busca de una absorción negativa. Sin embargo, la absorción siempre se mantuvo positiva en todos ellos, a pesar de que la relación entre N_2 y N_1 difirió del valor correspondiente al equilibrio térmico.

4. Intentos por amplificar la luz.

Hasta aquí se han considerado los niveles de energía electrónicos de los átomos como simples. Sin embargo, son en general niveles degenerados, o sea compuestos de subniveles susceptibles de observación experimental si se somete a los átomos a campos magnéticos intensos, como ocurre en el efecto Zeeman. La energía de cada uno de los subniveles difiere muy poco de la de los demás, por lo que los factores de Boltzmann $e^{-E/kt}$ son prácticamente iguales para todos ellos. Sin embargo, como cada subnivel tiene distinta probabilidad de ser ocupado por el electrón excitado,

se definen, para cada uno de ellos, el grado de degeneración o multiplicidad, o bien, el peso estadístico de ocupación. Por razones históricas se emplea la letra g para indicar el grado de degeneración de un nivel energético. La introducción de la multiplicidad de cada uno de los niveles 1 y 2, lleva a reescribir la expresión (1) de Einstein, así:

$$\frac{N_2}{g_2} \cdot A_{21} + \frac{N_2}{g_2} \cdot \rho(\nu_{21}) B_{21} - \frac{N_1}{g_1} \cdot \rho(\nu_{21}) B_{12} = 0 \quad ,$$

y la ley de Boltzmann, igualmente, resulta ser:

$$\frac{g_1}{g_2} \cdot \frac{N_2}{N_1} = e^{-(E_2-E_1)/kT} \quad . \quad (19)$$

Una vez introducida la noción de multiplicidad o grado de degeneración de los niveles de energía, conviene recordar las observaciones hechas por V.A. Fabrikant⁹ en su tesis doctoral en 1940. Escribía Fabrikant: "Para lograr amplificación molecular (atómica) es necesario que N_2/N_1 sea mayor que g_2/g_1 . Tal situación no ha sido observada hasta ahora en una descarga, a pesar de que una relación de poblaciones como la señalada es en principio alcanzable..." "Bajo dichas condiciones obtendríamos una salida de radiación mayor que la incidente y, por lo tanto, podríamos hablar de una demostración experimental directa de la existencia de absorción negativa".

Actualmente, la condición puntualizada por Fabrikant catorce años antes de la invención del máser, referente a la distribución de las poblaciones electrónicas entre los niveles de energía atómicos o moleculares y, por extensión, entre las bandas de conducción y de valencia de semiconductores, de tal modo que se cumpla que N_2/N_1 sea mayor que g_2/g_1 se denomina superpoblación o inversión de población de un nivel o banda de energía, con respecto de otro u otra.

El estudio realizado por Fabrikant y los estudiantes que colaboraron con él pone de manifiesto la importancia que le asignaban al método para establecer en el seno de la materia la inversión de población requerida para obtener absorción negativa. También son interesantes las conclusiones especulativas sobre la transformación de un material en el cual se produjo absorción negativa en un amplificador de luz. En esas especulaciones se

consideraba asimismo, la importancia del empleo de una estructura resonante y se anticipaba la posibilidad de modular la radiación producida por el amplificador. Estas novedades puramente teóricas, porque los experimentos hechos por ellos nunca permitieron obtener resultados definitivos, fueron patentados en 1951. La Oficina de Patentes Soviética no dio a conocer el título de la invención hasta 1959. Se trataba del "Método para amplificar radiación electromagnética (ultravioleta, visible, infrarroja y ondas de radio)".

A pesar de que los trabajos de Fabrikant no representaron un aporte decisivo en el desarrollo de las ideas y de las realizaciones del láser, conviene tenerlos en cuenta porque resultan ser el único intento conocido de construir un láser sin pasar por la etapa del máser, como históricamente ocurrió.

Entre los experimentos tendientes a obtener absorción negativa en un material, se describirá aquél en el que se excitan selectivamente átomos de cesio por bombeo óptico empleando la radiación emitida por átomos de helio en la longitud de onda 388,8648 nm. En general, todo proceso de excitación selectiva de niveles de energía se denomina bombeo. En el caso de la excitación de los vapores de Cs, el bombeo es de tipo óptico, por cuanto la radiación emitida por el He excitado en $\lambda=388,8648$ nm es absorbida por el Cs en una transición que se inicia en el nivel fundamental y que corresponde a una longitud de onda de 388,8610 nm. A la temperatura de vaporización del Cs, la totalidad de los átomos tienen sus electrones en el nivel fundamental de energía, por lo que la absorción de radiación adquiere gran eficiencia. Dicha eficiencia se realza notablemente por el hecho de que ambas transiciones, la de emisión del He y la de absorción del Cs, tienen casi la misma longitud de onda, de modo que se produce un fenómeno de resonancia. La Figura 5 muestra, en sus dos partes, los niveles de energía del He y del Cs que intervienen en el proceso y, en forma esquemática, el experimento de Fabrikant.

El propósito del experimento era que los vapores de Cs alcanzasen la condición de absorción negativa, mediante el empleo de las lámparas de He laterales. En tal situación, la lámpara de Cs, situada en el eje del tubo que contiene Cs a una temperatura de 175° C, se encendería y generaría un haz que, al atravesar el medio activado, se amplificaría. De este modo, el detector recibiría una intensidad mayor que la emitida por la lámpara sola. Como ya se dijo antes, los resultados nunca fueron decisivos.

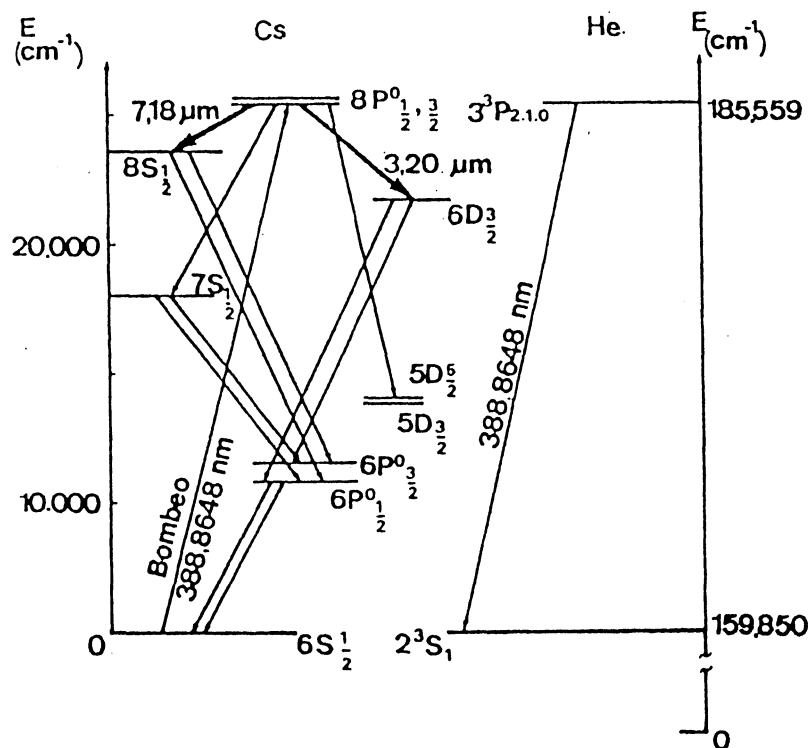


Figura 5. a) Excitación del cesio por bombeo óptico con radiación del helio.

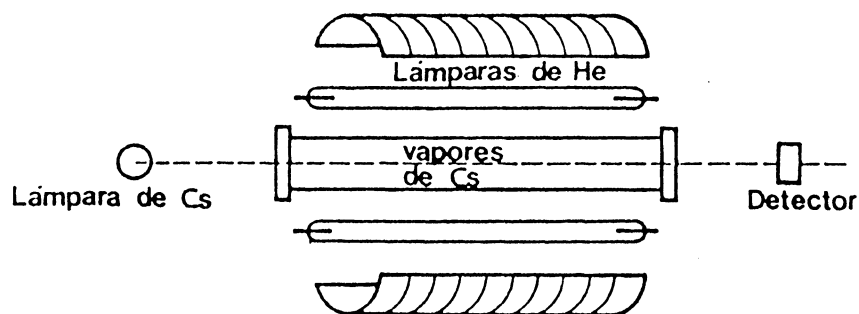


Figura 5. b) Experiencia propuesta por Fabrikant para amplificar radiación.

Las razones pueden encontrarse en los trabajos de Townes y sus discípulos¹⁰ llevados a cabo en la Universidad de Columbia, Estados Unidos de Norteamérica, en 1960. Las conclusiones de los mismos pueden sintetizarse así: La compleja estructura de la fluorescencia del Cs provocada por la excitación de la radiación emitida por el He hace que, en vez de generar una inversión de población entre el nivel $8P^0_{1/2}$ y el nivel fundamental $6S_{1/2}$, más bien se produzca entre dicho nivel superior $8P^0_{1/2}$ y los niveles $8S_{1/2}$ y $6D_{3/2}$. En efecto, esto es así, ya que en 1962, G. Gould y colaboradores¹¹ trabajando en las condiciones experimentales previstas, observaron efecto de amplificación láser en las transiciones sugeridas por C.H. Townes y colaboradores¹⁰, las que corresponden a las longitudes de onda de 3,20 y 7,18 μm , respectivamente.

El láser de Cs es el único láser gaseoso que ha funcionado por bombeo óptico, a pesar de otras sugerencias investigadas y que pueden consultarse en la revisión de C.K.N. Patel¹². Aparte del láser de Cs de G. Gould y colaboradores, no hay información de que se haya intentado construir otro del mismo tipo. Las dificultades experimentales que se presentan son muy grandes y se deben, fundamentalmente, a que los vapores de Cs son reactivos en alto grado y atacan a casi todos los materiales.

El intento de Fabrikant por construir un amplificador óptico sin pasar por la etapa máser se frustra, en principio, por la relación desigual que existe entre las probabilidades de transición estimulada y espontánea en la región óptica ($\lambda = 500 \text{ nm}$) y en la región de las microondas ($\lambda = 1 \text{ mm}$). Dicha relación entre los coeficientes A y B de Einstein es, respecto a las longitudes de onda indicadas, de 10^{13} a favor de la región del espectro de las microondas. El posterior desarrollo de una tecnología, compleja al principio y de uso exclusivo en laboratorios de investigación, permitió, sin embargo, la obtención sistemática de amplificación en la región óptica.

5. El máser.

El primer intento con resultado positivo de obtener amplificación de radiación por el proceso de emisión estimulada se debe a C.H. Townes y a dos de los estudiantes que colaboraban con él en 1954, J.P. Gordon y H.J. Zeiger¹³. La Figura 6 muestra esquemáticamente el efecto de inversión de población en la molécula de amoníaco, así como el montaje experimental del primer máser. Un haz de moléculas de amoníaco excitadas por la acción de una descarga eléctrica posee moléculas en las que el átomo de nitrógeno ha saltado la barrera de potencial que forman los tres átomos de hidrógeno.

El comportamiento de estas moléculas en las que se produjo la inversión de población con respecto de las moléculas en las que el átomo de nitrógeno se mantuvo en su posición normal, es diferente si se hallan en presencia de un campo eléctrico no homogéneo. La consecuencia de que el haz molecular atraviese un campo eléctrico no homogéneo es la formación de un haz divergente de moléculas normales de amoníaco, y de un haz convergente en el que todas las moléculas tienen su átomo de nitrógeno en su nivel de energía superior. Una vez separadas del haz molecular, las moléculas excitadas se recogen en una cavidad que resuena a la frecuencia correspondiente a la transición entre ambos niveles de energía. La frecuencia correspondiente es de 23,87 GHz y su longitud de onda es de 1,25 cm. La cavidad resonante a estas frecuencias es relativamente fácil de construir y operar, porque sus dimensiones físicas son apenas algunos múltiplos más grandes que la longitud de onda. Ya dentro de la cavidad resonante las moléculas comienzan a emitir en forma espontánea su exceso de energía.

Sin embargo, como se aclaró ya, el proceso de emisión estimulada se ve favorecido a estas frecuencias y así es como al penetrar en la cavidad una señal electromagnética proveniente de un generador del tipo Klystron, en la frecuencia de 23,87 GHz, se amplificó notablemente, de modo que a la salida de la cavidad se obtuvo una señal varios órdenes de magnitud mayor que a la entrada. Además, el ancho de línea a la salida resultó ser de unos 6 kHz.

Así pues, las características más importantes del primer máser fueron, en cuanto oscilador, un excelente generador de radiación monocromática en extremo estable, y en cuanto amplificador, un dispositivo cuya relación señal-ruido es incomparablemente mejor que la de cualquier otro dispositivo electrónico.

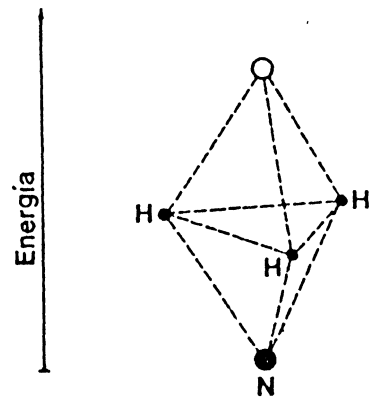


Figura 6. a) Esquema de la molécula de amoníaco.

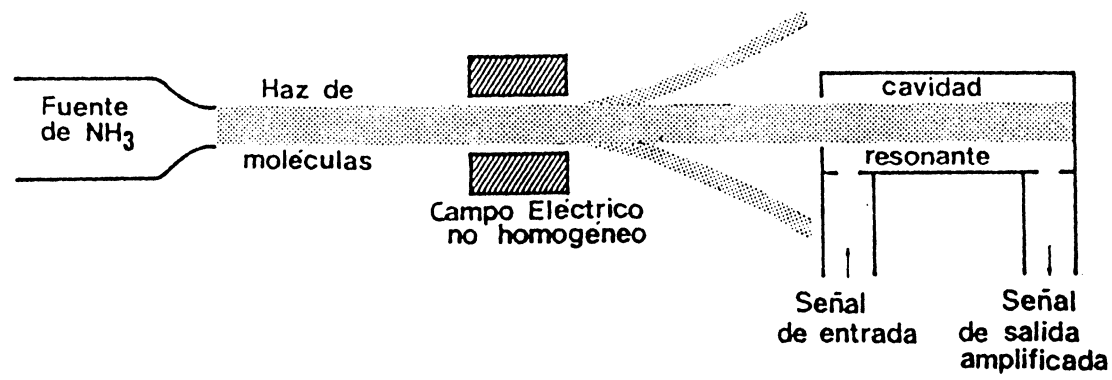


Figura 6. b) Máser de amoníaco.

El grado de monocromaticidad del máser fue investigado cuidadosamente por C.H. Townes, quien obtuvo el siguiente resultado: El ancho medio $\delta\nu$ de la línea espectral amplificada por el máser es proporcional al ruido térmico kT , como ocurre en todos los amplificadores, e inversamente proporcional a su potencia de salida P . La relación que obtuvo fue la siguiente:

$$\delta\nu = 4\pi kT (\Delta\nu)^2 / P, \quad (20)$$

donde $\Delta\nu$ es el ancho medio de la línea espectral espontánea. Cuando un máser opera a temperaturas muy bajas, como las de He líquido, el ancho medio de la línea espectral emitida por acción de la emisión estimulada es varios órdenes de magnitud ($10^4 - 10^6$) menor que el ancho medio de la línea espectral emitida espontáneamente.

La ubicación de la descripción del máser de amoníaco en este lugar cumple dos propósitos. Uno, mostrar que las investigaciones de Fabrikant y de Townes son, en el fondo, equivalentes. Las ideas formuladas por el primero se cumplen exactamente en los estudios teóricos y experimentales del segundo. El otro propósito es indicar cómo las ideas de Fabrikant no influyeron en absoluto en los trabajos de los iniciadores de la física de la generación de radiación en extremo coherente, monocromática e intensa.

Después de la invención del máser de amoníaco, se propusieron otros medios para lograr la inversión de población entre niveles de energía de diversos materiales. En menos de un año, a partir de 1954, se echaron las bases teóricas de los esquemas de inversión de población entre dos, tres y cuatro niveles de energía. En una sección posterior se discutirán más detalladamente estos esquemas.

La influencia de tales estudios permitió la construcción de máseres en los que el medio activo es una sustancia cristalina. Los iones triplemente positivos de gadolinio y de cromo implantados en distintas estructuras cristalinas y excitados convenientemente, fueron empleados por H.E. D. Scovil y colaboradores¹⁴, A.L. McWhorter y J.W. Meyer¹⁵ y G. Makhov y colaboradores¹⁶ en máseres operados a distinta frecuencia dentro del espectro de las microondas. Esto ocurrió entre 1956 y 1958. En especial, el uso del ion Cr^{3+} implantado en varias concentraciones en la red cristalina del óxido de aluminio, llamado comúnmente rubí, hizo posible la construcción de amplificadores de microondas de muy bajo ruido, con variación de la frecuencia de sintonía, gran estabilidad de dicha frecuencia y ancho de banda variable. El costo de cada tipo de amplificador máser dependía de las exigencias impuestas por las características. Su aplicación principal tuvo lugar en los campos de la radioastronomía y en del radar óptico.

6. Camino hacia el máser óptico.

Al llegar la técnica del máser a su más alto grado de perfección y al disponer de equipos de laboratorio sumamente delicados, varios investigadores creyeron llegada la hora de iniciar la búsqueda de materiales y técnicas apropiadas para obtener radiación coherente y monocromática en la región óptica del espectro electromagnético. Se trataba de construir un máser óptico. El comienzo de tal búsqueda estaba supeditado, por lo menos, a la solución de los cuatro problemas decisivos siguientes:

1) Si fuera necesario emplear alguna estructura resonante, como en el caso del máser, para contener al medio activo, dicha estructura tendría que ser capaz de sustentar una gran cantidad de modos de resonancia. Como se explicó ya, una cavidad resonante en la región de las microondas mide unos pocos centímetros, o a lo más unas decenas de centímetros. Una estructura resonante que permita sustentar ondas estacionarias de longitud de onda de 0,5 a 10 μm y que se preste por sus dimensiones al manejo experimental, debe ser concebida de modo que sea compatible con el requerimiento de compensar la reducida absorción negativa en dicha región del espectro y las dimensiones del medio activo. En efecto, tales dimensiones tendrían que ser de alrededor de un metro y, en tal caso, el número de modos resonantes variaría entre 10^5 y 10^7 , según la longitud de onda de la radiación que se quisiera amplificar.

2) La baja absorción negativa $\alpha(\nu)$ en la región óptica del espectro se debe a la desfavorable relación entre los coeficientes A y B de Einstein. La influencia del factor ν^3 haría muy difícil obtener amplificación en la región óptica, y aunque se la obtuviera, la radiación coherente tendría que distribuirse en tan gran número de modos resonantes que, seguramente, no se la podría observar a causa del enmascaramiento provocado por la radiación que el medio activo emitiría espontáneamente y que se constituye en el ruido del amplificador.

3) Las diferencias de energía entre niveles atómicos o moleculares puestas en juego son mayores y en grado sustancial en la región óptica que aquéllos de la de las microondas. Esto trae aparejado requerimientos energéticos de otra magnitud en caso de querer desarrollar una fuente coherente óptica.

4) La posibilidad de excitación del medio activo mediante radiación electromagnética producida por otras fuentes de banda ancha, como el cuerpo negro, o monocromáticas, como las descargas gaseosas, es remota a causa de la baja intensidad de todas ellas.

7. La emisión estimulada en la región óptica.

La contribución decisiva hecha por A.L. Schawlow y C.H. Townes¹⁷ en 1958, permitió abrigar esperanzas de que el salto de la región de las microondas a la región óptica se produjese en corto plazo. Su labor consistió esencialmente en esclarecer la manera en que una cavidad resonante a frecuencias ópticas sólo puede sustentar unos pocos modos de resonancia. Esto es esencial porque si la emisión estimulada se canalizara en todos los modos resonantes posibles su intensidad no se distinguiría de la producida por efecto de la emisión espontánea. En este sentido relacionan, como ya lo habían hecho en el caso del máser, la exigencia instrumental de obtención de radiación estrictamente monocromática con las características de los osciladores atómicos individuales. Así plantean la relación entre el ancho de línea $\delta\nu$ producido en la amplificación de radiación por la acción de la emisión estimulada, caracterizándolo como ancho de línea de la señal, y el ancho de línea de la emisión espontánea $\Delta\nu$, al que por contraposición lo denominan ancho de línea del ruido. La fórmula a la que arribaron por extensión de la (20), es:

$$\delta\nu = 4\pi h\nu (\Delta\nu)^2 / P \quad . \quad (21)$$

Efectivamente, el ruido térmico kT que tiene validez en el caso del máser, fue reemplazado aquí por $h\nu$, que es el ruido cuántico producido por la emisión espontánea de radiación.

Esta expresión es válida cuando la inversión de población entre los estados de energía en los que se produciría la amplificación de radiación es casi despreciable, esto es, cuando sus poblaciones respectivas apenas comienzan a distinguirse de la forma normal de población de Boltzmann. Por ello, resulta que el ancho de línea $\delta\nu$ es inversamente proporcional a la potencia P emitida por el láser.

En la fórmula (21) el ancho de línea del ruido $\Delta\nu$ al parecer se refiere al ancho de la línea espectral emitida en forma espontánea. Sin embargo, este ancho de línea corresponde, más bien, al ancho del modo resonante, que es muchísimo más angosto que el de la línea espectral, como puede apreciarse en la Figura 7.

Para evitar la amplificación de radiación en modos resonantes inconvenientes, Schawlow y Townes proponen el empleo de una cavidad abierta. Y así, llegan a la conclusión que un interferómetro de Fabry-Perot puede comportarse como una cavidad resonante abierta y emplearse, en consecuencia, para alojar el medio material donde debe producirse la inversión de población. Luego, el ancho del modo resonante $\Delta\nu$ señalado anteriormente, depende drásticamente de las condiciones de reflectividad y de

absorción del material con que se confeccionan los espejos del interferómetro.

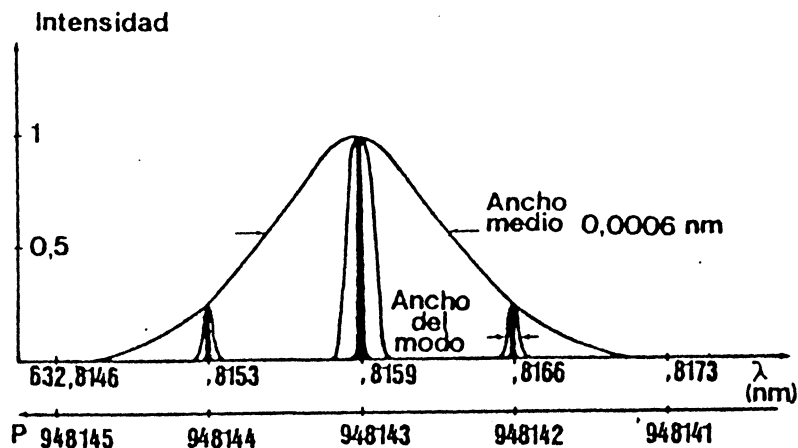


Figura 7. Modos longitudinales de resonancia bajo el perfil de línea espectral de ancho medio 0,0006 nm correspondiente a la transición de $\lambda=633$ nm del Ne. La representación de la estructura modal es la de un láser de He-Ne cuya intensidad tiene una longitud $D=30$ cm.

La descripción hecha por ellos de los materiales que convenía emplear como medio amplificador es breve, aunque hacen mención, además de los gases, de los sólidos cristalinos. En el caso de los gases, comentan los resultados que podrían obtenerse usando potasio. Proponen un experimento semejante al ya comentado del cesio. En cambio, en el caso de los sólidos cristalinos, se encuentran ante la imposibilidad natural de emplear una lámpara para irradiar el cristal y lograr que los electrones sean bombeados al nivel superior de energía para provocar la inversión de población con respecto de algún nivel inferior.

Sin embargo, advierten que tal situación se podría lograr por medio de procesos no radiativos (como las vibraciones de la red cristalina) que permitiesen que los electrones bombeados ópticamente hasta niveles elevados de energía, decayeran hasta acumularse en algún nivel de energía en forma privilegiada.

Schawlow y Townes, antes de enviar sus conclusiones a la revista científica "The Physical Review", donde se publicaron, solicitaron y obtuvieron en 1958 la Patente N^o 2929922¹⁸ de los Estados Unidos de Norte-

américa para proteger su invento, el que consiste en "un medio material en el que se obtiene temperatura absoluta negativa, dispuesto entre dos placas reflectoras paralelas, de modo que dicho medio pueda ser bombeado por fuentes luminosas accesibles y donde pueda ser efectivamente aislado un modo resonante correspondiente a la emisión estimulada de radiación". Este texto se parece más a una aspiración o expresión de deseos, que a la descripción de un invento. En particular, es interesante analizar la forma estereotipada y rimbombante con que señalan cómo el medio excitado puede comportarse a la manera de un amplificador de radiación.

Si se lograran en dicho medio excitado temperaturas absolutas T de signo negativo, la relación de poblaciones N_2/N_1 entre los niveles de energía E_2 y E_1 sería la correspondiente a la inversión, de acuerdo con la expresión (18). Sin embargo, no tiene sentido físico enunciar la posible existencia de tal medio material en el que se produjesen temperaturas absolutas negativas, por cuanto lo único real es que, por acción de algún mecanismo, se obtendría inversión de población entre algunos niveles de energía, y por lo tanto, dentro del esquema de referencia de Boltzmann representado por la fórmula (9), dicho medio no se encontraría ya en equilibrio térmico. En tal caso, o sea, fuera del equilibrio térmico, no tiene sentido referirse a la temperatura.

De todos modos, el contenido de la patente en cuestión muestra que A.L. Schawlow y C.H. Townes prácticamente vislumbraron hasta los detalles de cómo sería un máser óptico, es decir un máser que emitiera radiación coherente en la región de las frecuencias ópticas. Llegan incluso a describir un aparato para comunicaciones consistente en un generador máser de radiación sumamente estable, un amplificador máser y un modulador de las señales basado en el efecto Zeeman.

También G. Gould¹⁹ patentó en 1959, en Inglaterra, un aparato semejante al de Schawlow y Townes, aunque nunca publicó sus estudios y conclusiones. Es interesante comentar otros tipos de mecanismos mediante los cuales Gould vaticina que debe producirse inversión de población en un medio material. Además de la excitación óptica de vapores de un metal alcalino - Gould elige sodio en vez del potasio de Schawlow y Townes y, anteriormente, Fabrikant trató con cesio - lograda por la radiación emitida por una intensísima lámpara del mismo metal, Gould introduce la excitación por transferencia resonante producida en colisiones entre átomos excitados de una especie y átomos sin excitar de otra. Introdujo también la excitación de átomos y moléculas por impacto electrónico directo, como en las descargas eléctricas, así como la excitación óptica por coincidencia de las líneas espectrales de átomos de distinta especie. Tales posibilidades se hicieron realidad más adelante. En efecto, según las investigaciones posteriores, el láser de He-Ne funciona debido a la excitación por transferencia resonante producida en la colisión de átomos de He excitados en niveles de energía metaestables con átomos de Ne en el nivel fundamental; una gran variedad de láseres gaseosos o de vapores, como el de Ar y Hg, funcionan por colisión electrónica directa y tal como lo había intentado

Fabrikant sin ningún éxito, el láser de He-Cs funciona por excitación óptica por coincidencia de la línea espectral de $\lambda=388,8648$ nm emitida por el He y que es absorbida por el Cs en la transición de $\lambda=388,8610$ nm.

La propuesta de Gould de obtener inversión de población por transferencia resonante en colisiones entre átomos de distintos elementos, fue analizada en detalle por A. Javan, en 1959, en el par de átomos de He y Ne. Llegó a la conclusión de que la inversión de población puede alcanzarse en una mezcla excitada de He-Ne, como consecuencia de la acción combinada y competitiva de varios mecanismos de excitación y de pérdida de energía producida a diferentes presiones, porcentajes de mezcla y densidad electrónica en la descarga.

8. El láser de rubí.

Por otra parte, T. Maiman consideró la posibilidad de obtener inversión de población entre niveles de energía del ión triple de cromo embebido en la estructura cristalina del óxido de aluminio.

Maiman, que había dedicado varios años a la construcción y perfeccionamiento de máseres de rubí, estudió en detalle el fenómeno de absorción de luz por el rubí. Esta absorción tiene lugar en dos anchas bandas del azul y verde y da lugar, tras un proceso de decaimiento en el que la red cristalina juega un papel fundamental, a la acumulación de los iones de cromo en dos niveles de energía muy cercanos. Ambos niveles de energía se conectan con el nivel fundamental mediante un lento proceso de emisión espontánea, debido a que la probabilidad de transición espontánea es en extremo pequeña. Estos niveles de energía se comportan como si fuesen metastables. El proceso de emisión se produce en la cercanía de la longitud de onda 694,3 nm. Esta longitud de onda depende del tipo de rubí, esto es, de la concentración de iones de cromo en óxido de aluminio.

De acuerdo con lo dicho, el rubí actúa absorbiendo radiación en dos anchas bandas del azul y del verde, de modo que, después de un cierto tiempo, canaliza esa energía dentro de una estrechísima banda fluorescente de emisión en el rojo. De ahí que el rubí, en potencia, era un excelente candidato a láser.

En 1959 este hecho experimental fue reconocido como definitivo. Sin embargo, la sola idea de obtener inversión de población entre un nivel excitado cualquiera y el nivel fundamental provocaba recelos porque, dada la distribución normal de Boltzmann de las poblaciones, se requeriría que por lo menos la mitad de los iones fueran excitados y desplazados del nivel fundamental a los niveles superiores de interés. En efecto, los niveles que entran en juego son 2E y 4A_2 cuyas multiplicidades valen $g_2=4$ y $g_1=2$, respectivamente. Luego, aplicando el criterio de Fabrikant para que se produzca amplificación en una transición,

$$\frac{N_2}{N_1} > \frac{g_2}{g_1} ,$$

se verifica que para el láser de rubí debe ser $N_2 > \frac{1}{2}N_1$. La Figura 8 muestra el esquema de niveles del ión triple de cromo en óxido de aluminio.

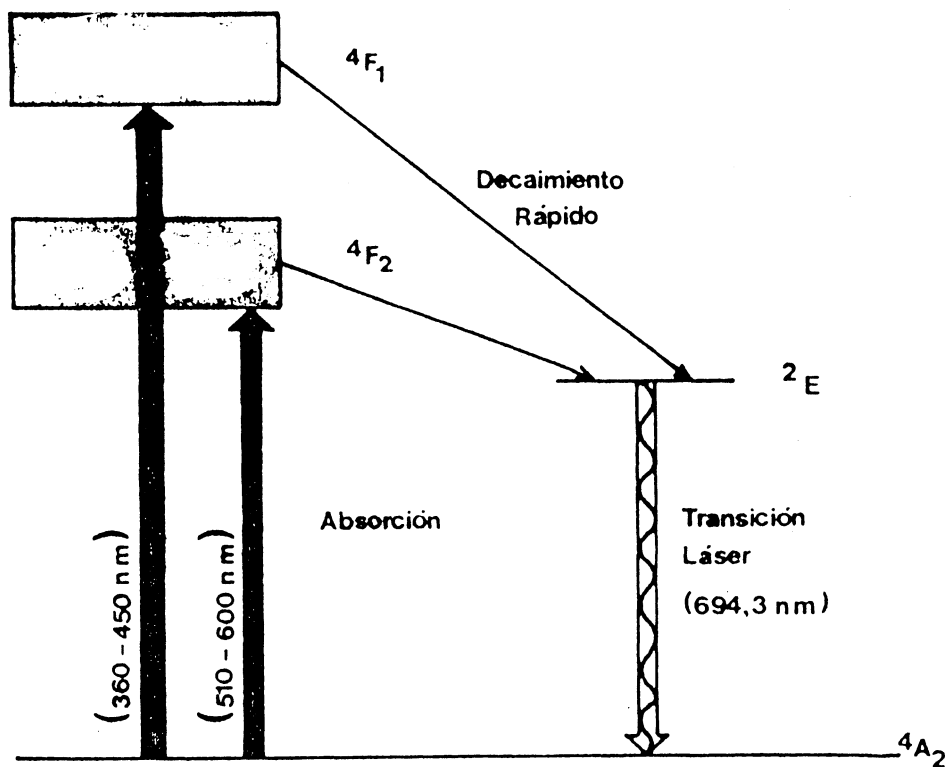


Figura 8 . Esquema simplificado de los niveles de energía del ión Cr^{3+} en óxido de aluminio.

En ella se han señalado las dos anchas bandas de absorción del azul y del verde 4F_1 y 4F_2 , las transiciones no radiativas extremadamente rápidas desde ellas al nivel 2E y la transición láser ${}^2E \rightarrow {}^4A_2$. Teniendo en cuenta estos aspectos, se clasificará al láser de rubí entre los denominados de tres niveles, como se considerará más adelante.

Estos estudios permitieron reconocer la facilidad con que debería obtenerse inversión de población entre niveles de energía excitados, en particular en gases.

Por esto, durante la primera mitad del año 1960, todos los intentos de obtener amplificación en la región óptica del espectro electromagnético

se basaban en la excitación privilegiada de algunos niveles de energía en gases, en especial de los vapores de metales alcalinos bombeados ópticamente, y de los gases nobles excitados por descargas eléctricas. Por la misma razón, el descubrimiento de T. Maiman de la acción amplificadora generada en el rubí fue recibida con gran sorpresa y hasta con incredulidad. Sin embargo, tal como se ha comentado ya, el descubrimiento de Maiman no fue casual. Sus predicciones resultaron ciertas y el desarrollo experimental por él realizado, perfectamente ajustado a ellas. El anuncio periodístico de la invención del láser se hizo el 7 de julio de 1960, y enseguida Maiman envió el resultado de sus estudios y observaciones al editor de "Physical Review Letters", para su publicación. La revista citada rehusó la publicación del artículo de Maiman aduciendo que lo sabido ya sobre el máser era tal que no había razones para insistir sobre el tema. Por fortuna, la revista inglesa "Nature"²¹ dio al artículo su merecido lugar y la invención del láser quedó así registrada oficialmente.

El vuelco provocado en el desarrollo de las investigaciones sobre el máser óptico por el éxito de Maiman, se refleja en que prácticamente todos los laboratorios que intentaban construir un amplificador óptico gaseoso pasaron de inmediato a intentarlo con el rubí. Así es como en un mes el láser de rubí fue construido siguiendo las ideas de Maiman en varios laboratorios de los Estados Unidos de Norteamérica. Efectivamente, a fines de agosto de 1960, "Physical Review Letters" publica la confirmación de que el rubí funciona como un amplificador y emisor de luz coherente basado en la emisión estimulada de radiación²².

Hubo que esperar hasta fines del año 1960 por la noticia de que el tan esperado láser gaseoso era una realidad²³. Este láser era del tipo de mezcla de gases nobles, He-Ne, y actuaba según una secuencia de procesos físicos, entre ellos la excitación preferencial de niveles metastables del He por colisiones electrónicas y la transferencia resonante de la energía acumulada por los átomos de He a los de Ne, por medio de colisiones atómicas. El primer láser de He-Ne emitía radiación coherente en varias longitudes de onda alrededor de 1,1 μm .

9. Las condiciones de bombeo.

9.1. Generalidades.

En las secciones precedentes se citaron algunas técnicas cuyo objeto es producir la superpoblación de ciertos niveles de energía con respecto a la indicación de la ley de Boltzmann y, de esta manera, obtener la inversión de población entre ellos y otros niveles de energía para provocar la emisión estimulada de radiación. El empleo de dichas técnicas depende en alto grado del material de que se ha de construir un láser. Obviamente, un láser de He-Ne, como ya se adelantó, funciona gracias a una serie de mecanismos que favorecen la inversión de población entre los niveles de

energía de los átomos de Ne entre los que se producirá la emisión estimulada y con ello la amplificación de la radiación. La serie de mecanismos por la que, en último término, el láser opera en forma conveniente, debe ser conocida íntimamente para poder controlarla durante los experimentos. Más aún, tal conocimiento y control deben ser perfeccionados de continuo para lograr, como se ha hecho en ciertos casos, la fabricación de láseres empleando métodos seriados de elevada eficiencia como si se tratara de la producción de lámparas de iluminación o de tubos de televisión.

9.2. Esquema de láseres de tres y cuatro niveles de energía.

Hasta aquí, todos los planteos se han iniciado suponiendo que la transición $2 \rightarrow 1$ ocurre entre niveles de energía atómicos, moleculares o iónicos, pero sin relacionarlos efectivamente a los esquemas completos de niveles de energía de dichas especies. La observación experimental ha provisto abundante material informativo para permitir agrupar a los láseres en dos grandes categorías: láseres de tres y láseres de cuatro niveles. La Figura 9 muestra un esquema general de los niveles de energía que entran en juego en cada caso.

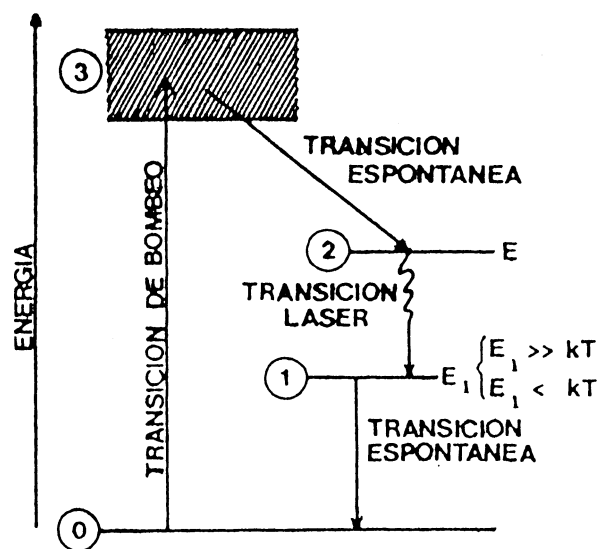


Figura 9 . Esquema de tres y cuatro niveles.

Como puede apreciarse, la idea de Fabrikant, expuesta previamente, queda excluida de estos dos modelos de láseres, ya que propone la excitación preferencial de un nivel energético a fin de lograr inversión de población entre dicho nivel excitado y el nivel fundamental, siguiendo un esquema de dos niveles.

Sin embargo, el análisis del láser de rubí - que, como se recordará, fue el primero - así como de algunos láseres gaseosos - H_2 , N_2 y ON^{24} - de reciente invención, mostrará que es un efecto posible el suscitar inversión de población entre algunos niveles excitados y el nivel fundamental.

El conocimiento de estos factores es imprescindible para determinar si un láser podrá emitir radiación en régimen continuo (cw: continuous wave) o en régimen discontinuo. Por ejemplo, al describir los láseres gaseosos, se podrá apreciar la diferencia esencial que existe entre el láser de He-Ne que emite en forma continua en líneas espectrales del Ne y el láser pulsado de Ne. En el primero, la mezcla gaseosa de He y Ne se excita por una corriente eléctrica y debido a una característica de los niveles 2^1S y 2^3S del He que los hace metastables, es decir, que no pueden conectarse por transiciones radiativas con el nivel fundamental, dichos niveles se pueblan privilegiadamente. Sobrevienen luego las colisiones entre átomos excitados de He y los de Ne que se encuentran en el nivel fundamental, produciéndose la transferencia de energía de los primeros a los últimos, lo que genera la inversión de población entre algunos niveles de energía de los átomos de Ne, como se muestra en la Figura 10.

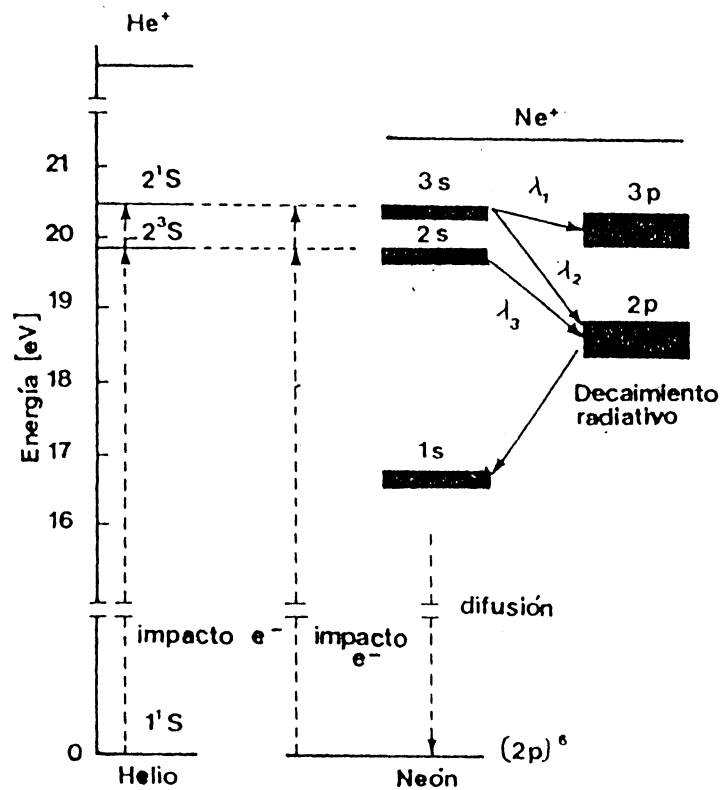


Figura 10. Esquema de los niveles de energía que intervienen en el láser de He-Ne. λ_1 y λ_3 emisión IR y λ_2 emisiones visibles.

Este proceso de bombeo puede ser realizado en forma continua y una vez establecida la condición de equilibrio el láser de He-Ne opera sin dificultad.

Sin embargo, es posible lograr amplificación en las mismas transiciones del Ne obtenidas cuando el láser opera en régimen continuo mediante la excitación de una mezcla de helio y de neón, pero empleando descargas eléctricas pulsadas a través de neón puro. En este caso, al no estar los átomos de neón asociados a los de helio para producir la inversión necesaria en algunos niveles de energía del neón como ya se describió, deben emplearse descargas eléctricas cortas e intensas como para que se sustente la requerida inversión crítica en tiempos que son del mismo orden de magnitud que la vida media de los niveles atómicos involucrados. Por tales razones, el láser de neón puro sólo puede operar en régimen pulsado.

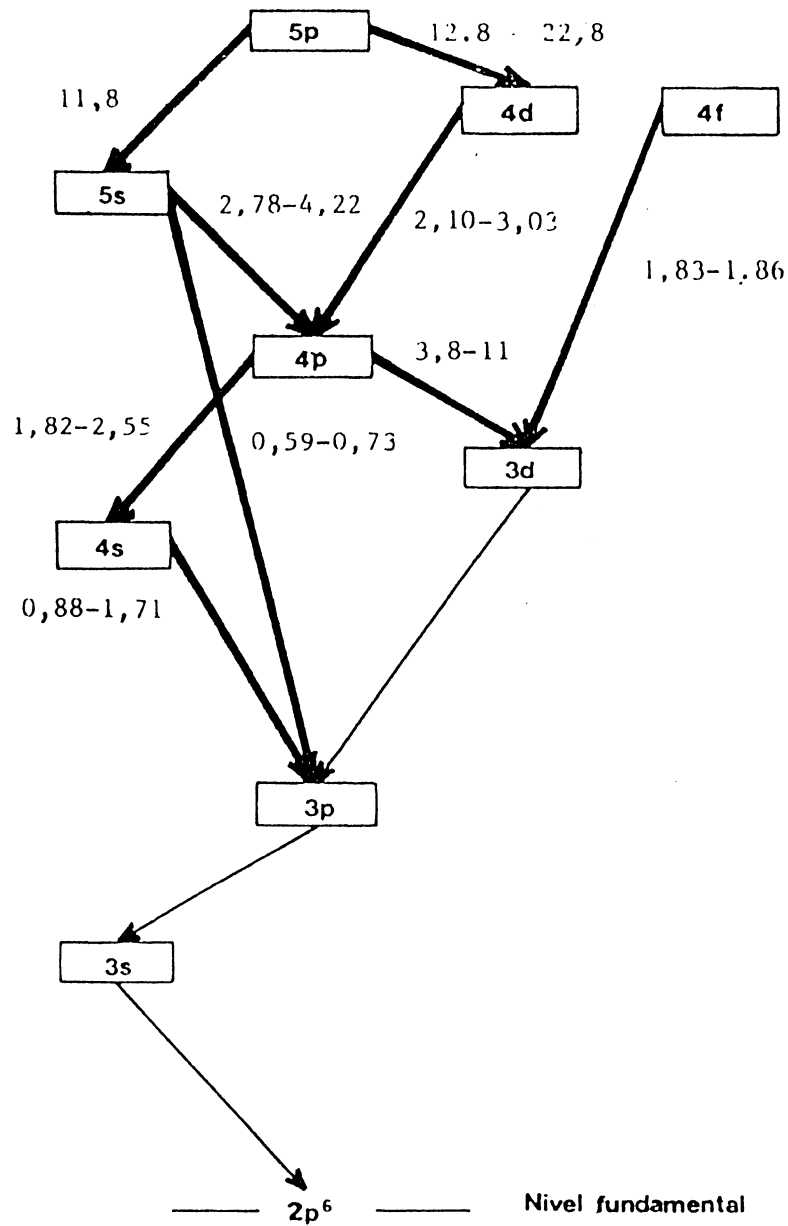


Figura 12. Transiciones en cascada entre niveles de energía del Ne I en las que se observa emisión láser (\longrightarrow) y emisión espontánea (\longrightarrow). Cada bloque de niveles de energía está indicado con su símbolo de Racah. El espaciado entre los niveles no guarda escala uniforme. Las longitudes de onda correspondientes a los grupos de transiciones están expresadas en μm .

9.3. Algunos casos particulares.

La posibilidad teórica de generar emisión estimulada provocando la excitación de niveles de energía de moléculas y logrando la consiguiente inversión crítica entre dichos niveles y el nivel fundamental fue enunciada, en 1965, por P.A. Bazhulin y colaboradores²⁴. La prueba experimental se obtuvo unos años después merced a los trabajos de R.T. Hodgson²⁵, R.W. Waynant y colaboradores²⁶, R.W. Waynant²⁷ y R.T. Hodgson y R.W. Dreyfus²⁸, utilizando N_2 y NO . La Figura 11 muestra la excitación de la molécula de hidrógeno y la emisión estimulada en las bandas de Werner y de Lyman.

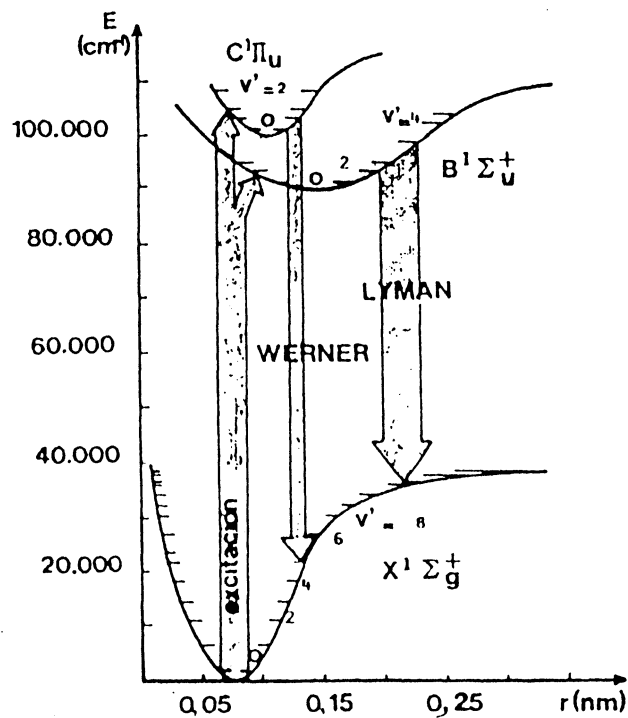


Figura 11. Excitación y emisión del láser de hidrógeno molecular.

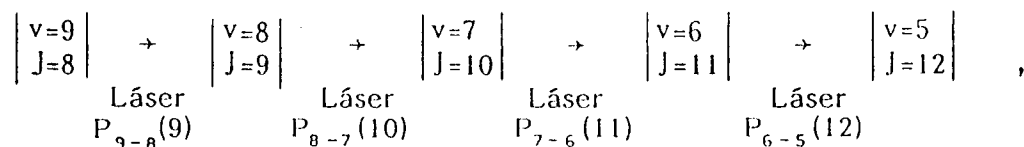
Generalmente se admite que el impacto electrónico directo sobre el nivel fundamental de la molécula es la causa de la producción de la inversión crítica entre los niveles excitados y el nivel fundamental. Sin embargo, recientemente M. Gallardo y asociados^{29,30} enunciaron la hipótesis de que el láser de H_2 resulta excitado mediante un mecanismo rápido en extremo y que ya se había sugerido para el caso de la molécula

de nitrógeno³¹. Dicho mecanismo podría tener la siguiente secuencia: Los electrones de la descarga eléctrica que poseen una energía definida actúan más fácilmente con el H₂ mediante un proceso resonante en virtud del cual un electrón de la descarga queda en una órbita transitoria alrededor de la molécula. El tiempo en que el electrón queda atrapado se estima en 10⁻¹⁶ de segundo. Durante este tiempo se forma un ión negativo del H₂, el H₂⁻, y una vez pasado dicho tiempo, el electrón es expulsado con energía casi nula, y por lo tanto, la molécula H₂ queda con el exceso. Este exceso de energía incorporado a la molécula provoca un rápido reordenamiento de sus propios electrones. En caso de que tal exceso coincida con la energía de un nivel excitado, el electrón del nivel fundamental pasará a ocuparlo y de este modo resultará superpoblado con respecto a lo que indica la distribución de Boltzmann.

En algunas estructuras atómicas o moleculares los procesos de inversión de población entre niveles de energía se desarrollan en el tiempo de modo sucesivo. Así se forman cascadas por las que la inversión de población entre dos niveles una vez cumplida la transición entre ellos dos, es trasladada de manera que resulta invertido el nivel inferior de dicha transición con respecto a otros niveles más bajos que él. Dependiendo de las vidas medias de los respectivos niveles, de sus multiplicidades g y de los factores instrumentales, se pueden observar dichas cascadas en sucesión temporal directa, como ocurre en el caso de las transiciones de los láseres de gases nobles. En efecto, en estos láseres se han establecido varias vías de cascadas⁴².

En la Figura 12 aparecen representadas las más conocidas.

Pero también pueden producirse mecanismos de cascada inversos para establecer la inversión de población entre niveles de energía sucesivos. Tal es el caso del láser de CO⁴³, en el que, una vez que los niveles rotacionales-vibracionales están convenientemente excitados, se produce una primera transición láser entre los niveles más bajos. Este efecto reduce rápidamente la población del nivel superior de dicha transición, por lo que se obtiene inversión crítica entre tal nivel y el inmediatamente más alto, generándose una segunda transición láser entre ellos. Este efecto se propaga subiendo en la escala energética de los niveles rotacionales-vibracionales, mientras se produce la emisión sucesiva de transiciones láser hasta que se agota, en algún momento de la secuencia, la posibilidad de quedar establecida una suficiente inversión de población que asegure la amplificación de la respectiva transición. El caso detallado corresponde a la cascada inversa:



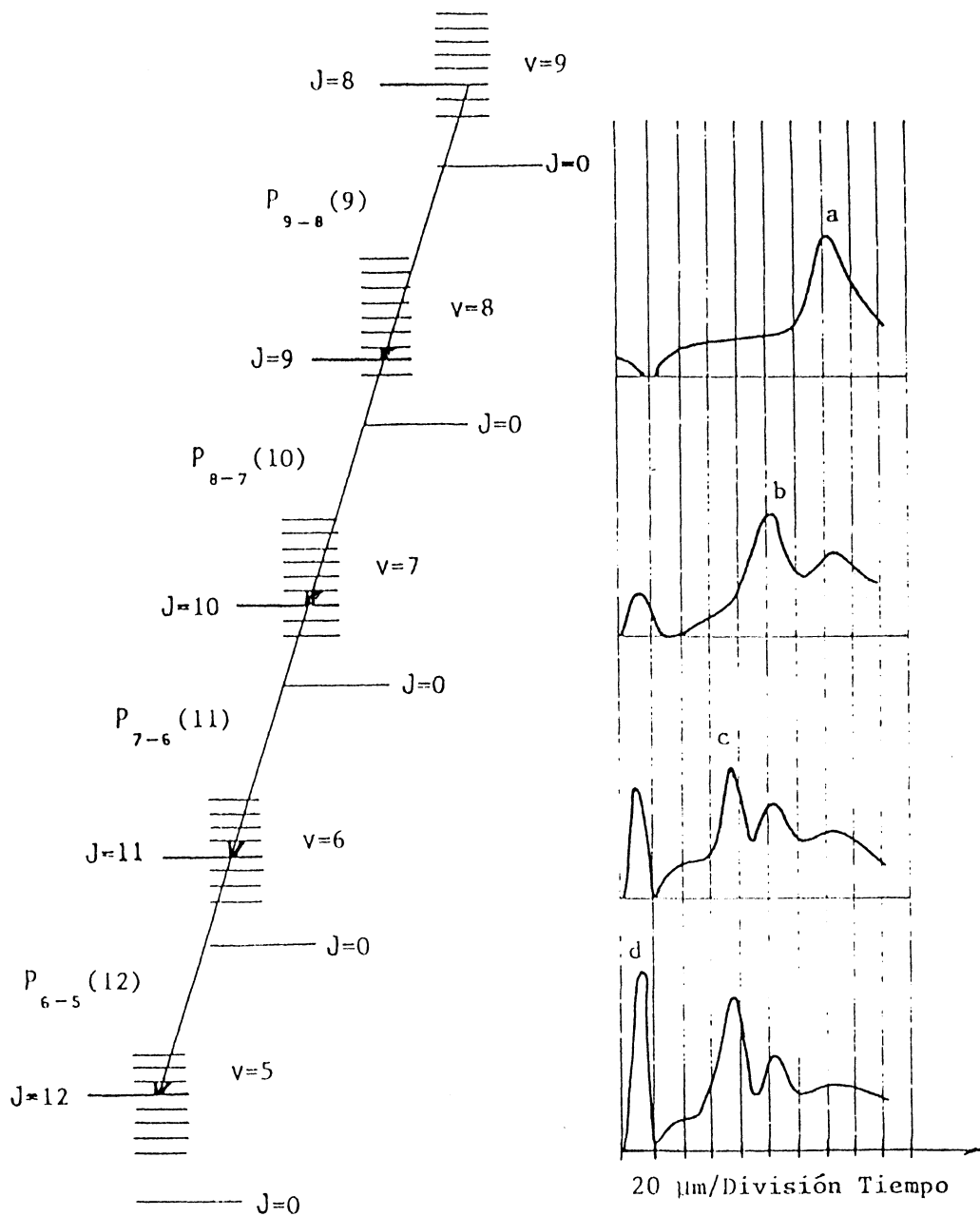


Figura 13. Cascada láser inversa entre niveles rotacionales-vibracionales del CO correspondientes a las transiciones a) $P_{6-5}(12)$ en $5,08845 \mu\text{m}$, b) $P_{7-6}(11)$ en $5,14530 \mu\text{m}$, c) $P_{8-7}(10)$ en $5,20345 \mu\text{m}$, y d) $P_{9-8}(9)$ en $5,26310 \mu\text{m}$. La posición de los niveles no guarda relación de escala. Los registros oscilográficos individualizan la secuencia temporal de las componentes de la cascada láser inversa.

la que se representa esquemáticamente en la Figura 13. En la Figura 13, además, se incluye la observación de los sucesivos pulsos de radiación láser en la pantalla del osciloscopio.

10. Cavidades resonantes ópticas.

10.1. Generalidades.

A.L. Schawlow y C.H. Townes¹⁷, tomando como modelo una cavidad resonante rectangular de paredes altamente reflectoras, como las empleadas en los máseres, sugirieron en 1958 el empleo del interferómetro de Fabry-Perot de placas planoparalelas como resonador multimodal con el fin de extender el efecto de amplificación máser a la región óptica del espectro electromagnético. La misma posibilidad fue señalada por A.M. Prokhorov³² y R.H. Dicke³³, aunque sólo Schawlow y Townes analizaron la forma en que podía reducirse el número de modos resonantes y calcularon la perturbación introducida por la emisión espontánea.

El interferómetro de Fabry-Perot consta de dos espejos planos y perfectamente paralelos entre sí. La normal a los espejos se denomina eje del interferómetro. De este modo, todo rayo de luz que penetre en la cavidad formada por ambos espejos en la dirección de su eje se propagará indefinidamente entre ellos por sucesivas reflexiones. Es decir, que los rayos de luz paralelos al eje de la cavidad se propagarán sin poder salir de ella, en tanto que si aquéllos forman un cierto ángulo con dicho eje, terminarán por abandonar la cavidad. Esta capacidad de mantener la radiación luminosa entre los espejos del interferómetro de Fabry-Perot, se emplea para disponer dentro de la cavidad el tubo de descarga gaseosa o la barra de rubí, por ejemplo, que constituyen el material activo de algunos láseres. Así, por sucesivas reflexiones se simula una longitud virtual del material activo igual al producto de su longitud real por el número de reflexiones. Este número de reflexiones, aún en el caso de rayos luminosos que se propagan paralelamente al eje del interferómetro, no es demasiado grande. Ocurre que no hay espejos perfectos y en cada reflexión parte de la intensidad lumínica se pierde por absorción en el material que constituye el reflector o por transmisión a través de ellos. Estas pérdidas son las que hacen tan difícil la investigación que implique el uso de interferómetros. Sin embargo, en el caso que nos ocupa, esto es, el empleo del interferómetro de Fabry-Perot como cavidad resonante de un láser, las pérdidas pueden controlarse casi a voluntad, de modo que la amplificación generada en el medio activo por la emisión estimulada de radiación las compense siempre, en cuyo caso el dispositivo amplifica. A estas causas que provocan pérdidas deben añadirse otras ya enumeradas en la sección 9. Todas ellas, según se vio, pueden agruparse en un coeficiente de pérdidas por paso a través del medio activo. Cuando el medio alcanza las condiciones experimentales de inversión crítica, entonces el coeficiente de absorción $\alpha(\nu)$

adquiere valores negativos y el dispositivo se vuelve amplificador. La expresión (16) que relaciona la intensidad $I(x,\nu)$, medida después de haber recorrido el camino x en el medio material, con la intensidad incidente $I(0,\nu)$ sigue siendo válida en este caso en que el coeficiente $\alpha(\nu)$ es negativo. La expresión (16) señala claramente ahora que la intensidad $I(x,\nu)$ de la radiación, después de recorrer el camino x a través del medio invertido, resultará ser mayor que la $I(0,\nu)$ de la incidente.

En este sentido dicha expresión (16) permite comprender porqué algunos láseres deben ser contruidos de modo que la longitud del medio activo sea grande y la reflectividad de los espejos prácticamente total a fin de obtener amplificación, como es el caso del láser visible de He-Ne, mientras que otros, con una longitud reducidísima del medio activo y empleando un trozo de papel brillante, a modo de un solo espejo, alcanzan las condiciones de amplificación como es el caso del láser de N_2 que emite radiación ultravioleta. Es decir, si el coeficiente de absorción negativa es bajo para obtener amplificación se requerirá un larguísimo camino a través del medio activo. En general, esto último puede obtenerse con un elevado número de pasajes de la radiación por un medio activo de corta longitud. Por el contrario, si la absorción negativa alcanza valores elevados puede llegarse a la situación experimental en que, sin necesidad de espejos, o quizás empleando sólo uno, el medio amplifique la radiación de modo que el dispositivo emita el característico haz direccional y coherente. En este último caso, el dispositivo se denomina emisor superradiante para distinguirlo del láser. Nótese, sin embargo, que el sentido de la palabra láser no implica en absoluto el empleo de cavidades formadas por espejos. La introducción teórica de tales cavidades por A.L. Schawlow y C.H. Townes tiene por sola justificación el hecho que los máseres funcionaban gracias a ellas. Recientes observaciones astrofísicas mediante radiotelescopios han llevado a los expertos a no descartar la posibilidad de interpretar los resultados obtenidos basándose en la consideración de que la emisión estimulada tiene un papel importante en la génesis de la radiación proveniente de galaxias y nubes lejanas de gases. Como no es concebible que en el espacio cósmico haya cavidades resonantes, el resultado de las observaciones confirman que el efecto de amplificación superradiante está presente en los procesos de emisión de radiación que ocurren en el Universo.

Ahora bien, la cavidad tiene otro papel además del detallado, el que hace referencia a su comportamiento como resonador, sustentando ondas estacionarias.

Conviene recordar que el factor de calidad Q de una cavidad resonante, define el grado de acoplamiento entre el campo electromagnético sustentado en su interior y la propia cavidad en función de la frecuencia de la radiación. La configuración del campo electromagnético dentro de la cavidad será iniciada y sustentada sólo en aquellas frecuencias para las cuales resuena. En las demás frecuencias, en consecuencia, se producirán elevadas pérdidas de energía que impedirán mantener las condiciones de oscilación del campo electromagnético. Una medida del rango $\Delta\nu$ de frecuencias para el cual la cavidad responde a excitaciones externas, es el factor Q , definido como:

$$Q = 2\pi\nu_c \frac{[\text{Energía almacenada}]}{[\text{Potencia disipada}]},$$

donde ν_c es la frecuencia central de resonancia.

La variación de la energía almacenada U con el tiempo, teniendo en cuenta todas las pérdidas que la producen, está gobernada por la expresión siguiente:

$$\frac{dU}{dt} = -\frac{2\pi\nu_c}{Q} U,$$

cuya solución vale:

$$U(t) = U_0 \cdot e^{-(2\pi\nu_c / Q)t}.$$

Si por efecto de la excitación de la cavidad por amplificación de un campo eléctrico, se almacena la cantidad de energía U_0 , su valor irá decreciendo exponencialmente con una constante de amortiguamiento que será inversamente proporcional a Q . Por ello, el campo eléctrico $E(t)$ dentro de la cavidad, decaerá amortiguadamente, según la expresión:

$$E(t) = E_0 e^{-(2\pi\nu_c / Q)t} \cdot e^{-2\pi i\nu_c t},$$

suponiendo que inicialmente se propagaba una onda plana caracterizada por $E_0 e^{-2\pi i\nu_c t}$.

Como la resonancia ocurre en el rango $\Delta\nu$, resulta que la oscilación

amortiguada contendrá no sólo la frecuencia central ν_c , sino todas aquellas que se encuentran a ambos lados de ν_c , las que también resultan excitadas. El contenido espectral de la oscilación amortiguada se determina empleando la técnica del análisis de Fourier, resultando que la amplitud del campo eléctrico según las frecuencias, vale:

$$E(\nu) = \frac{1}{2\pi} \int_0^{\infty} E(t) e^{2\pi i \nu t} dt ,$$

y que la distribución de la intensidad de campo, es decir, de la energía almacenada, resulta tener el perfil de Lorentz:

$$U(\nu) = |E(\nu)|^2 = \frac{1}{2\pi[(\nu-\nu_c)^2 + (\nu_c/2Q)^2]}$$

Comparando con la expresión del perfil de Lorentz para una línea espectral, se obtiene el siguiente resultado:

$$\Delta\nu = \nu_c / Q ,$$

esto es, que el ancho espectral de la resonancia es directamente proporcional a la frecuencia ν_c e inversamente proporcional al factor de calidad Q .

De aquí se sigue otra imposición para las cavidades resonantes ópticas. Como las frecuencias en el rango óptico son más grandes que las de la región de operación del máser en un factor de por lo menos 10^5 , resulta que el Q de las cavidades resonantes ópticas debe ser elevadísimo para compensar y evitar que la radiación estimulada emitida en un ancho de banda espectral $\Delta\nu$, no resulte enmascarada por la radiación espontánea emitida bajo el perfil de la misma línea espectral, como se ve en la Figura 7.

Un factor de calidad Q elevado se obtiene si se cumplen los criterios enunciados por A. Yariv y J.P. Gordon³⁴ en 1962. Dice el primero: Debe existir una familia de rayos luminosos que sufra un número razonable de reflexiones en los espejos que constituyen la cavidad, antes de que se pierda saliendo de ella, de modo que el número de pasajes a través del medio activo baste para lograr el efecto de amplificación. Este número de pasajes se ha estimado entre 20 y 300. Y el segundo criterio indica: Las dimensiones de los espejos circulares deben satisfacer la siguiente re-

lación:

$$\frac{L_1 L_2}{\lambda D} \geq 1 \quad , \quad (22)$$

donde L_1 y L_2 son los radios de los reflectores, D la distancia que los separa y λ es la longitud de onda de la radiación confinada en la cavidad. La ecuación (22) se deduce del tratamiento de la difracción de Fresnel.

Ambos criterios coinciden aunque sus enunciados adopten formas distintas. La primera, surge de la óptica geométrica en tanto que la segunda proviene de la óptica ondulatoria. Tales formas tienen a su vez diferentes posibilidades de aplicación. En ciertos casos el criterio geométrico basta para resolver el problema planteado con una dada configuración de la cavidad resonante. Sin embargo, el criterio ondulatorio resulta siempre más poderoso por cuanto puede asociarse a la teoría electromagnética de las cavidades y, en este caso, se pueden obtener ciertos resultados que con el anterior criterio resultaban imposibles. Como ejemplo pueden darse los siguientes resultados fundamentales obtenidos empleando el criterio ondulatorio: Determinación de Q , determinación de la variación de las pérdidas por pasaje α en función de la frecuencia ν , y variaciones de la fase sobre el frente de onda emitido por el láser.

En caso de ser el radio de los espejos igual a L , la expresión (22) resulta:

$$N = L^2 / \lambda D \geq 1 \quad , \quad (23)$$

que es la definición del número de Fresnel N . Este número desempeña un importante papel en la teoría de los resonadores ópticos y es igual al número de zonas de Fresnel de la figura de difracción que se observa desde el centro de uno de los espejos sobre el otro. La difracción se produce por la incidencia de una onda, cuyo frente tiene fase constante sobre uno de los reflectores. El área del frente de onda se considera para este propósito igual al área del reflector.

Al iniciar la discusión de estructuras o cavidades resonantes a frecuencias ópticas se mencionó el interferómetro de Fabry-Perot. Por supuesto, hay una gran variedad de estructuras que permiten sustentar ondas estacionarias a frecuencias ópticas y, sin embargo, no resulta trivial encontrar la que posea elevado factor de calidad Q para sólo unas pocas frecuencias comprendidas dentro del perfil de la línea espectral que se desea amplificar. Estas frecuencias privilegiadas corresponden a la oscilación o amplificación de los respectivos modos en que resuena la cavidad.

Para ejemplificar estas posibilidades baste mencionar sólo dos antecedentes. C.G.B. Garret y colaboradores³⁵ obtuvieron efecto de amplificación láser dentro de esferas de dieléctricos que constituían a la vez el material activo y la cavidad resonante. El efecto de reflexión total en las paredes de la esfera las hacía excelentes espejos. Sin embargo, no pudieron reducir el número de modos resonantes. Es decir, que el factor Q era aproximadamente de la misma magnitud para todos ellos. Por otro lado, E. Snitzer y H. Osterberg³⁶ lograron obtener efecto láser en fibras cuyo diámetro medía unas pocas longitudes de onda. La radiación correspondía a la amplificación en un solo modo. Sin embargo, como el volumen de material activo de la fibra resultaba ser ínfimo, la intensidad luminosa que obtuvieron fue despreciable.

10.2. El interferómetro de Fabry-Perot como base de las cavidades resonantes ópticas.

La exposición que sigue se iniciará en el interferómetro de Fabry-Perot, tal como ocurrió históricamente, y luego se harán referencias generales a otras estructuras resonantes de empleo frecuente.

La Figura 14 muestra la onda estacionaria sustentada entre los espejos reflectores cuando la condición de resonancia es $D = 5\lambda/2$.

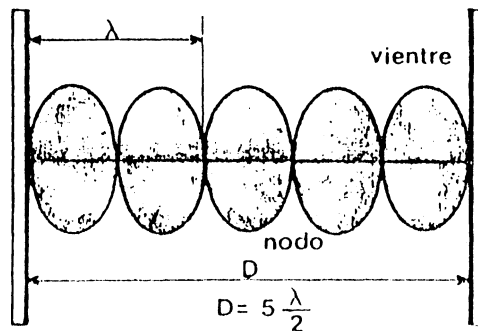


Figura 14. Onda estacionaria.

Es obvio que en la región óptica del espectro electromagnético, al igual que en el campo mecánico, para sustentar ondas estacionarias es indispensable tener en cuenta el desfase y la pérdida de intensidad en cada reflexión. En el caso de las cavidades resonantes ópticas estos factores pueden describirse de la siguiente manera. La onda luminosa - para simplificar puede suponerse plana y de longitud de onda λ - que incide en un espejo, es reflejada por éste, se propaga hasta el opuesto, vuelve

a ser reflejada en el mismo sentido que al comienzo y alcanza la posición inicial, al tiempo que ocurren los siguientes cambios: Su amplitud será ahora menor porque los espejos no son perfectos y parte de la luz se pierde por transmisión y por absorción, y su posición en el espacio no coincidirá con la inicial exactamente en el instante $t = 2D/c$ a contar de la iniciación del proceso, por cuanto en las reflexiones se produce un desfase de las ondas debido a la absorción, que puede interpretarse como si la distancia D entre los espejos fuese $D' \neq D$. Los efectos de estos fenómenos pueden controlarse así: Para que las ondas vuelvan a estar en fase bastará desplazar uno de los reflectores hasta que se cumpla la condición de resonancia y las pérdidas de amplitud se compensarán introduciendo en la cavidad un medio activado convenientemente, que sólo amplifique la luz de longitud de onda λ . En pocas palabras, debe construirse un láser.

La comparación de las dos ondas puede hacerse analizando la configuración de los respectivos campos electromagnéticos siguiendo dos caminos. Uno, resolviendo las ecuaciones de Maxwell correspondientes a un rayo de luz estrecho haciendo que las superficies reflectoras lo intercepten perpendicularmente, según un frente de fase constante, para lograr que el rayo se refleje sobre sí mismo. El otro, adoptado por A.G. Fox y T. Li³⁷, utiliza la formulación escalar del principio de Huygens para calcular el campo electromagnético sobre un espejo después de que la onda recorrió el camino de ida y vuelta dentro de la cavidad, y lo compara con el campo inicial.

Ambos métodos conducen a resultados comparables. Nótese que estas dos formas de abordar el problema provienen de la óptica ondulatoria.

La configuración más simple del campo electromagnético en relación con la capacidad de las cavidades resonantes de sustentar ondas estacionarias es aquella en que los cambios de fase son múltiplos enteros de 2π . A este modo de resonar la onda electromagnética con la cavidad resonante se lo denomina longitudinal o axial. Los modos longitudinales se identifican por la fórmula:

$$p = 2D' / \lambda \quad , \quad (24)$$

de la misma manera que, en la teoría del interferómetro de Fabry-Perot, se expresan las diferentes condiciones de interferencia constructiva. Como ya se vio en el ejemplo de la Figura 14, los únicos posibles valores que adopta p son enteros. Anteriormente se discutió el significado de la distancia D' entre los espejos, al comentar el efecto de la absorción del material de los reflectores. Ahora conviene agregar el efecto del índice de refracción del medio material. En caso de que el medio material no ocupe totalmente la cavidad, como ocurre con los láseres gaseosos, la longitud del camino que la luz debe recorrer para ir de un espejo a otro se puede expresar así:

$$D'' = \int_0^{D'} n(z) dz \quad , \quad (25)$$

donde $n(z)$ es el índice de refracción en función de la posición z a lo largo de la trayectoria. De aquí en adelante, y recordando la definición general implícita en (25), en todas las expresiones se escribirá solamente D , en vez de D'' .

Como en la teoría de Fabry-Perot, la separación entre dos modos longitudinales debida a la variación $\Delta p = \pm 1$, que corresponde a un cambio en D igual a media longitud de onda, se puede calcular diferenciando la ecuación (24) como sigue:

$$dp = - \frac{2D}{\lambda^2} d\lambda \quad .$$

Despejando $d\lambda$ y expresándola en función de $dp = \Delta p = 1$, resulta:

$$\Delta\lambda = \lambda^2 / 2D \quad , \quad (26)$$

o lo que es lo mismo:

$$\Delta\lambda = \lambda / p \quad . \quad (27)$$

A veces se expresa la variación de frecuencia entre dos modos consecutivos en lugar de la separación en longitud de onda. Para esto se debe tener en cuenta la relación $\lambda\nu = c$. Haciendo el cálculo a partir de la expresión (26), resulta:

$$\Delta\nu = c / 2D \quad . \quad (28)$$

La Figura 7 ilustra el caso de la línea espectral del Ne, cuya longitud de onda es $\lambda = 632,8$ nm. El ancho medio espontáneo de dicha línea espectral es del orden de los 0,0006 nm. Para un láser de 30 cm de distancia entre espejos, se sustentarán simultáneamente tres ondas estacionarias, cuyas longitudes de onda se encontrarán debajo del perfil de la línea espectral. La diferencia de longitud de onda entre esas tres ondas es aproximadamente de 0,0006 nm. En este caso se dice que el láser amplifica radiación en sólo tres modos longitudinales y, en efecto, su emisión constará de tres componentes muy cercanas.

Para encontrar que son tres realmente los modos de oscilación de la cavidad resonante en el láser de He-Ne de 30 cm de longitud que emite en 632,8 nm, debe hacerse el siguiente cálculo. La condición de resonancia es, como se dijo antes, $p = 2D/\lambda$. Por tanto, para el valor central de la longitud de onda, resulta $p' = 2 \times 30 \text{ cm} / 632,8160 \text{ nm} = 948142,904$.

Como la condición resonante se cumple para valores enteros de p , debe adoptarse el número entero más cercano al calculado, es decir, $p = 948143$. Para calcular los otros posibles modos, se suma y resta una unidad al modo central y se calculan las longitudes de onda respectivas. Esto lleva a los siguientes valores: $\lambda_{p-1} = 632,8166 \text{ nm}$ y $\lambda_{p+1} = 632,8153 \text{ nm}$. Queda comprobado así que, debajo del perfil de la línea espectral, resultan ser excitados tres modos de oscilación, por cuanto la diferencia entre las longitudes de onda extremas correspondientes a $p+1$ y $p-1$ es menor que el ancho total de línea. Para mayor claridad, en la Figura 7 se han incluido los modos longitudinales $p-2$ y $p+2$, cuyas longitudes de onda caen fuera del perfil de la línea espectral. Un láser de He-Ne de mayor longitud oscilará en un número mayor de modos longitudinales, mientras que uno de 10 cm de longitud oscilará en uno sólo, el central.

Generalmente todos los láseres elevan su temperatura desde el momento en que comienzan a funcionar hasta alcanzar la temperatura de régimen estable. Los láseres de emisión continua se diferencian en este aspecto de los de emisión intermitente o pulsados. En aquéllos el tiempo que tardan en alcanzar la estabilidad térmica puede ser mayor de una hora. Los láseres pulsados alcanzan dicho estado térmico de equilibrio en el transcurso de un pulso, las más de las veces cortísimo.

En algunos láseres pulsados de altísima repetición aparece, además, un efecto acumulativo de calentamiento, que los asemeja a los continuos. Este fenómeno de calentamiento produce variaciones de D , en algunos casos notables. Estas variaciones pueden deberse a la dilatación de los soportes de los espejos, a la del medio material o al cambio de índice de refracción.

Independientemente de su origen, toda variación de D afecta en alto grado la composición modal de la radiación dentro de la cavidad y produce un cambio en la longitud de onda emitida por el láser, de acuerdo con la ecuación (24). Es decir, que dentro de la cavidad resonarán otros modos siempre que la longitud de onda que les corresponda esté presente debajo del perfil de línea espectral.

Los llamados láseres estabilizados emiten siempre radiación de longitud de onda constante. Un sistema detector apropiado controla cualquier variación en la longitud de onda emitida y, en forma proporcional, gobierna mediante un cristal piezoeléctrico la posición de uno de los espejos de la cavidad, de manera que se reestablezca el valor de la longitud de onda inicial.

Además de los modos longitudinales, toda cavidad resonante puede sustentar modos transversales. Estos proceden de cambios de fase menores o mayores que 2π y se los observa con toda facilidad provocando el des-

alineado de los espejos. Otras causas que favorecen la aparición de los modos transversales son la baja calidad de los espejos, la introducción de vibraciones mecánicas en el sistema de soporte de los espejos, la aplicación de gradientes térmicos a las regiones cercanas a los espejos y la saturación de la excitación del medio activo.

La estructura de los modos transversales de resonancia de un Fabry-Perot se puede calcular considerando, como ya se señaló antes, cómo los sucesivos pasajes de una onda electromagnética que se propaga a través del medio activo en su interior, perturban la amplitud y forma de su frente de onda. Es decir, cómo varía la distribución de amplitud y de fase en el frente de onda.

Formalmente, si E_q es la componente eléctrica del campo electromagnético después de q pasajes a través del Fabry-Perot, el campo E_{q+1} en el pasaje siguiente valdrá:

$$E_{q+1} = \epsilon E_q \quad ,$$

donde ϵ es una constante compleja, cuya parte real gobierna la disminución de la amplitud del campo y su parte imaginaria los cambio de fase. La parte real es proporcional al factor γ de pérdidas por pasaje, mientras que la parte imaginaria depende de los desfases provocados en las reflexiones y refracciones.

Ahora bien, el interferómetro de Fabry-Perot se puede interpretar como una sucesión de pupilas difractantes, tal como aparece en la Figura 15.

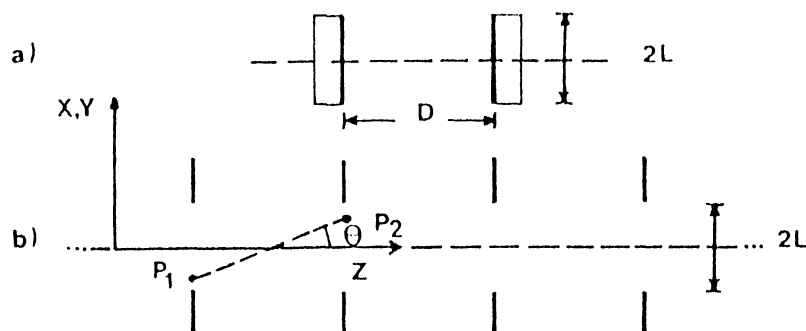


Figura 15 . a) Interferómetro de Fabry-Perot de placas circulares planas y paralelas de diámetro $2L$ separadas una distancia D . b) Sucesión de pupilas circulares difractantes de diámetro $2L$ equiespaciadas a distancias D y que simulan al interferómetro de Fabry-Perot.

En tal caso, si sobre la primera de ellas incide una onda plana, $E_i(x,y,z,t)$, la distribución de la componente de campo eléctrico sobre la siguiente revelará el proceso de difracción de Fresnel que sufrió. La estructura del campo difractado $E_j(x,y,z,t)$ se calcula según la expresión de Kirchoff:

$$E_d(x,y,z,t) = \frac{i}{2\lambda} \iint E_i(x,y,z,t) \frac{e^{-i 2\pi D/\lambda}}{D} (1 + \cos\theta) dx dy, \quad (29)$$

donde θ es el ángulo que forma el segmento de recta que une cada punto de la pupila difractante considerada, $P_1(x,y,0)$, con cada punto de la pupila siguiente, $P_2(x,y,D)$, y el eje de coordenadas z . Después de q pasajes el campo sobre un espejo debido al reflejado en el otro, se expresará por la fórmula (75), reemplazando E_i por E_q y E_d por E_{q+1} .

Puede aceptarse como convincente, además, el que después de un gran número de pasajes - entre 20 y 300, como ya se señaló - la distribución de campo se haga suficientemente estable. Esto es, que la interacción entre el campo electromagnético y la cavidad resonante alcance su estado estacionario, en cuyo caso se pone de manifiesto la estructura de modos transversales que sustenta y que se designa para la componente eléctrica del campo por $E_{m,n}$.

A.G. Fox y T. Li³⁷, basándose en estos antecedentes, calcularon las condiciones de estabilidad que debe satisfacer una onda electromagnética que se propague un elevado número de veces, reflejándose sucesivamente en los espejos del Fabry-Perot, hasta alcanzar la condición estacionaria.

El tratamiento matemático hecho por A.G. Fox y T. Li³⁷ les permitió expresar el valor de la componente eléctrica del campo electromagnético de una dada polarización sobre la superficie de uno de los reflectores, como:

$$E_{m,n} = E H_m \left[\left(\frac{2\pi}{\lambda D} \right)^{1/2} x \right] H_n \left[\left(\frac{2\pi}{\lambda D} \right)^{1/2} y \right] e^{-\pi(x^2 + y^2)/\lambda D}, \quad (30)$$

donde E es el campo eléctrico sobre el eje Z , y H_m y H_n son los polinomios de Hermite de orden m y n , respectivamente.

De la expresión (30) se puede escribir la distribución de campo eléctrico sólo para la componente x , según:

$$E_m = E H_m \left[\left(\frac{2\pi}{\lambda D} \right)^{1/2} x \right] e^{-\pi x^2/\lambda D}$$

Los primeros polinomios de Hermite, esto es, para $m = 0, 1$ y 2 , adoptan los valores siguientes:

$$H_0(\xi) = 1 \quad ,$$

$$H_1(\xi) = 2\xi \quad ,$$

$$H_2(\xi) = 4\xi^2 - 2 \quad .$$

De modo que, haciendo el cambio de variable $\xi = (2\pi/\lambda D)^{1/2} x$, en la Figura 16.a) se han graficado las amplitudes del campo eléctrico E_m en función de ξ para $m = 0, 1$ y 2 .

Obviamente, debe recordarse que la observación del campo electromagnético se realiza experimentalmente siguiendo las variaciones de su intensidad. Por ello, a partir de la distribución de campo de la Figura 16.a), puede obtenerse la distribución de intensidad calculando los cuadrados de E_0 , E_1 y E_2 , cuyas representaciones aparecen en la Figura 16.b).

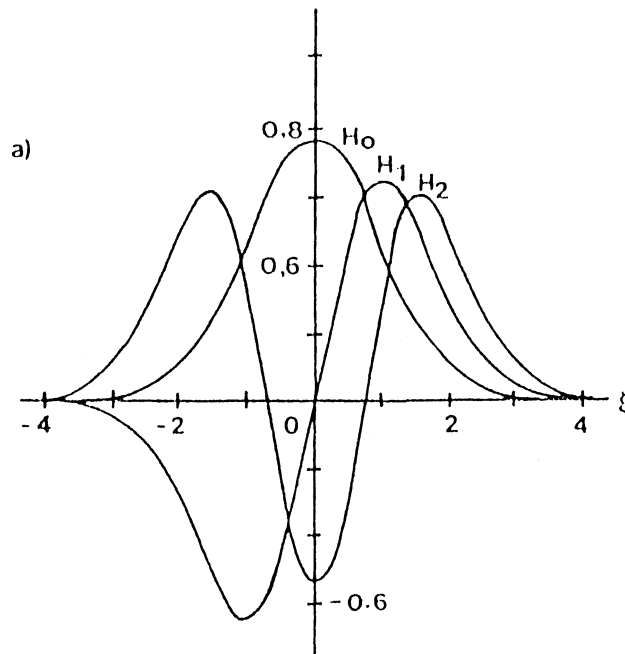


Figura 16 .a) Amplitudes de la componente eléctrica del campo electromagnético para los modos $m = 0, 1$ y 2 .

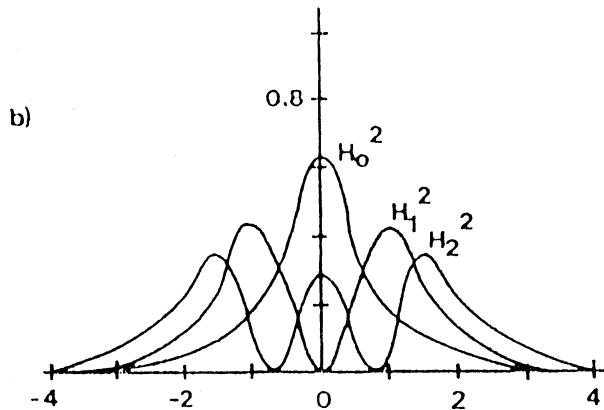


Figura 16 .b) Intensidades de los modos $m = 0, 1$ y 2 .

La condición resonante general para los modos longitudinales y modos transversales es la siguiente:

$$\left(\frac{2}{\lambda}\right)^2 = \left(\frac{p}{D}\right)^2 + \left(\frac{m^2+n^2}{4L^2}\right) \quad (31)$$

Si sólo se excita el modo transversal más bajo dentro de la cavidad resultan $m = 0$ y $n = 0$, con lo que la expresión (31) se reduce a la (24) que gobierna las condiciones relativas a los modos longitudinales. En este caso, la ecuación (30) se reduce a la parte exponencial, con lo que se obtiene un resultado verificado perfectamente por las determinaciones experimentales, en el sentido de que el modo transversal más bajo, $m = 0$ y $n = 0$, tiene una distribución de intensidad gaussiana. La excitación de otros modos transversales en los cuales m y n son siempre números enteros positivos, da lugar a toda una serie de longitudes de onda emitidas por el láser. Sin embargo, como la distancia óptica D entre los espejos es muchísimo mayor que sus dimensiones transversales L , resulta que p es, en general, un número centenares de miles de veces mayor que m y n . En el caso típico ya comentado del láser de He-Ne, de emisión $\lambda = 632,8 \text{ nm}$ y distancia entre los espejos $D = 30 \text{ cm}$, p asciende a 10^6 en tanto que m y n superan apenas el valor de 10. Esto hace que los modos transversales emitan longitudes de onda muy cercanas y a ambos lados de las que corresponden a los modos longitudinales. La siguiente expresión, derivada de la (31), permite calcular la diferencia en longitud de onda entre los modos transversales excitados

en un láser:

$$\Delta\left(\frac{1}{\lambda}\right) = \frac{1}{2D} \left[\Delta p + \frac{1}{8} \frac{D\lambda}{L^2} (m \Delta m + n \Delta n) \right] \quad . \quad (32)$$

En el caso de ser $m = 0$ y $n = 0$, se obtiene la ecuación (28) para los modos longitudinales.

A pesar de todo, éstas no son las únicas conclusiones que se pueden extraer del análisis de las ecuaciones (30) y (31).

La expresión (31) indica que si se excita el modo transversal más bajo, $m = 0$ y $n = 0$, la radiación emitida por el láser se confina en un estrechísimo haz de escasísima divergencia. En el láser de He-Ne ya mencionado, la divergencia del haz alcanza a 1 mrad, es decir que, tras propagarse 1000 m, el círculo que iluminará tendrá un diámetro de 1 m. Si se excitan otros modos transversales elevados, $m > 0$ y $n > 0$, el haz emitido por el láser pierde definición espacial y se convierte en un manajo de haces de distinta longitud de onda.

La forma de caracterizar estos modos es la siguiente. Como se trata de modos electromagnéticos transversales de resonancia, se emplea la sigla TEM, del inglés "Transverse Electromagnetic Modes", seguida de tres subíndices que expresan los valores de m , n y p , por ejemplo, TEM_{mnp}. La emisión del modo TEM_{00p}, el más bajo de los transversales, lleva varios modos longitudinales p . Es decir, existe una degeneración, ya que no se puede catalogar con cuántos modos longitudinales y con qué índices p está emitiendo el láser. La determinación de estos datos requiere la aplicación de delicadas técnicas interferométricas y electrónicas. La pérdida de definición espacial permite de manera sencilla distinguir el modo TEM_{00p} del TEM_{10p}, del TEM_{21p}, o del TEM_{50p}, a pesar de que las diferencias de longitudes de onda son ínfimas, empleando la observación directa sobre una pantalla donde se proyecta el haz luminoso procedente del láser. La Figura 17 muestra los cuatro casos señalados. Las fotografías se han tomado a unos 5 m del láser. Para simplificar la escritura, se suele suprimir el subíndice p , escribiendo TEM_{mn} en vez de TEM_{mnp}. Estos modos transversales fueron observados por vez primera por H. Kogelnik y W.W. Rigrod³⁸, en 1961. Claramente se distinguen los modos unos de otros porque la sección del haz luminoso emitido por el láser se compone, como se dijo, de un manajo de haces distinto en cada configuración. A pesar de que el láser emita una gran variedad de modos transversales, es posible reducir su emisión a sólo el TEM₀₀ con la introducción de un filtro que atenúe por absorción a los restantes modos, evitando que se amplifique, o bien un diafragma que lo seleccione.

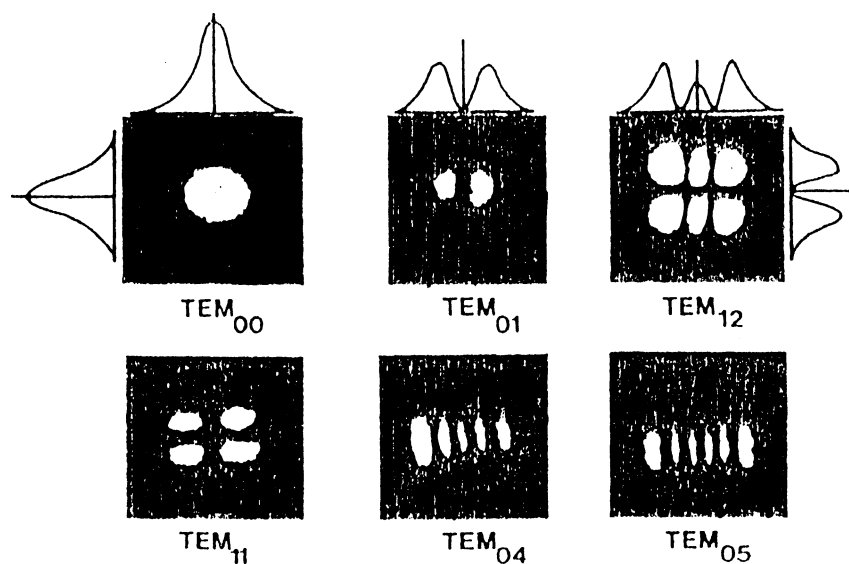


Figura 17. Fotografía de algunos modos transversales emitidos por el láser de He-Ne, $\lambda=632,8$ nm, y distribución de la intensidad de campo para aquéllos de $m = 0, 1$ y 2 , y $n = 0$ y 1 .

Otro resultado valioso por su índole didáctica y que puede comprobarse cuando un láser emite en régimen modal elevado con componentes en x e y , por ejemplo en el TEM_{21} , es el de la ortogonalidad de la solución expresada por la ecuación (30). Si en la cavidad resonante se introduce una tarjeta con mucha delicadeza de modo que uno de sus bordes corte el haz impidiendo que se amplifique su componente superior derecha, por ejemplo, entonces resultará anulada simultáneamente la amplificación en la componente inferior izquierda.

10.3. Geometría y estabilidad de las cavidades resonantes.

La radiación, en su trayectoria de vaivén dentro de la cavidad, a causa de la falta de uniformidad de los reflectores y de la acción del medio material, va sufriendo difracciones que deforman el frente de onda. Luego, si se ha supuesto que el frente de onda era inicialmente plano, las sucesivas difracciones lo convertirán en un frente esférico. Si el frente de onda plano se sitúa en $z=0$ al comienzo del proceso, se propagará, convirtiéndose en esférico, según líneas hiperbólicas hasta alcanzar uno de los reflectores colocados en $z = +D/2$. Es posible calcular el radio de la sección del haz para el modo TEM_{00} en el centro de la cavidad,

$z=0$, y sobre los reflectores, $z=\pm D/2$. Como la distribución de la intensidad del haz es gaussiana, se adopta calcular su radio para el valor e^{-1} del máximo de intensidad. De esta manera, resulta el radio del haz para $z=0$:

$$W_0 = (D \lambda / \pi)^{1/2} \quad (33)$$

y para $z = \pm D/2$:

$$W = W_0 [1 + (z/D)^2]^{1/2} \quad (34)$$

Al valor W_0 se denomina cintura del haz. Nótese que W_0 y W no dependen de las dimensiones de los espejos L . La Figura 18 ilustra convenientemente lo aquí expuesto.

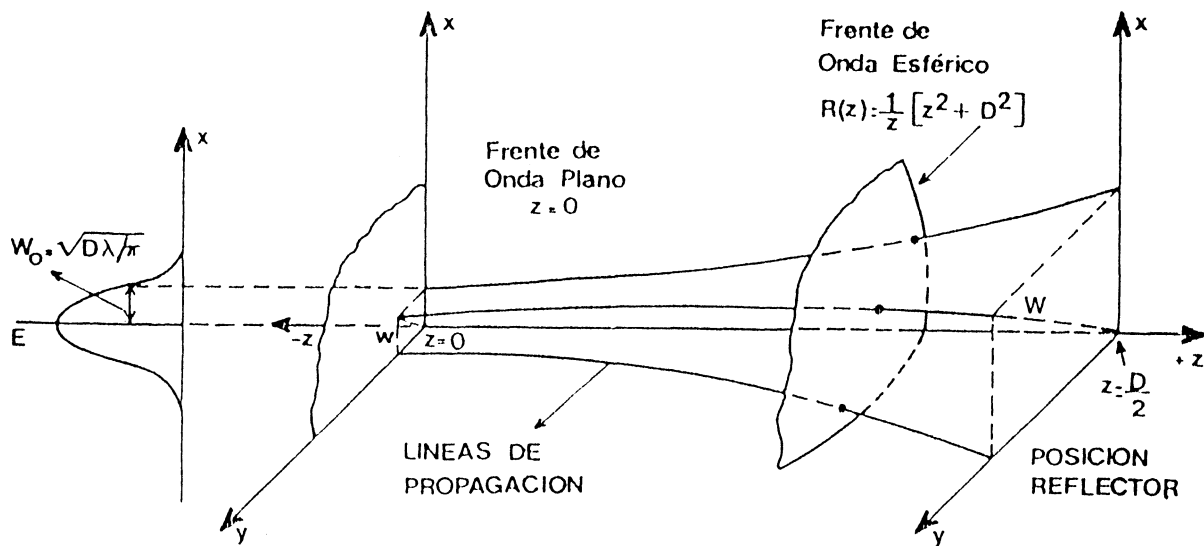


Figura 18. Parámetros de la cavidad resonante.

Considerando que las líneas de propagación son hiperbólicas, para que haya reflexión normal en los espejos es conveniente que la superficie de los mismos sea esférica. Luego, el frente de onda de un cierto radio de curvatura que arriba a un espejo conviene que encuentre que tal espejo tenga su mismo radio de curvatura. El radio de curvatura del frente de onda a medida que se propaga entre los espejos separados una distancia D , vale:

$$R(z) = \frac{1}{z} (z^2 + D^2) \quad . \quad (35)$$

Por lo expuesto, se sigue la conveniencia de que los espejos de las cavidades resonantes sean esféricos. Estas cavidades eran bien conocidas antes de la invención del láser. P. Connes³⁹ las había utilizado ya en 1955 en trabajos de interferometría de muy alta resolución. Se las suele llamar interferómetro de Fabry-Perot esférico. El interferómetro de Connes es mucho más fácil de alinear que el de Fabry-Perot, y presenta notables ventajas cuando la separación entre los espejos es mayor de 10 cm.

De la expresión (35) se sigue como corolario una regla simple para seleccionar los radios de curvatura de los espejos y la separación necesaria entre ellos a fin de que las pérdidas sean mínimas, porque los rayos serán reinyectados en la cavidad después de cada reflexión. Según dicha regla, el centro de curvatura de un espejo o el espejo mismo, pero no ambos, deben encontrarse entre el otro espejo y su respectivo centro de curvatura. Los casos límites que se dan son los tres siguientes: Cuando $R \rightarrow \infty$ (interferómetro de Fabry-Perot plano), cuando los focos de los espejos coinciden (cavidad confocal) y cuando los centros de curvatura coinciden (cavidad concéntrica). Esta regla para determinar cuándo una cavidad resonante es o no estable, se la puede reducir, a partir de la expresión (81), a la siguiente:

$$0 \leq \left(1 - \frac{D}{R_1}\right) \left(1 - \frac{D}{R_2}\right) \leq 1 \quad . \quad (36)$$

Si los espejos tienen distintos radios de curvatura, cuando R_1 , o bien R_2 , sea igual a D se tendrá el caso límite de igualación a cero de la fórmula (36). En el otro caso límite, el de igualación a la unidad, debe descartarse la solución trivial que implicaría aceptar que la distancia D fuera nula, porque no tiene sentido físico. La Figura 19 ilustra, con varios ejemplos, las configuraciones de cavidades resonantes incluso dos de elevada pérdida o inestables porque no cumplen la ecuación (36).

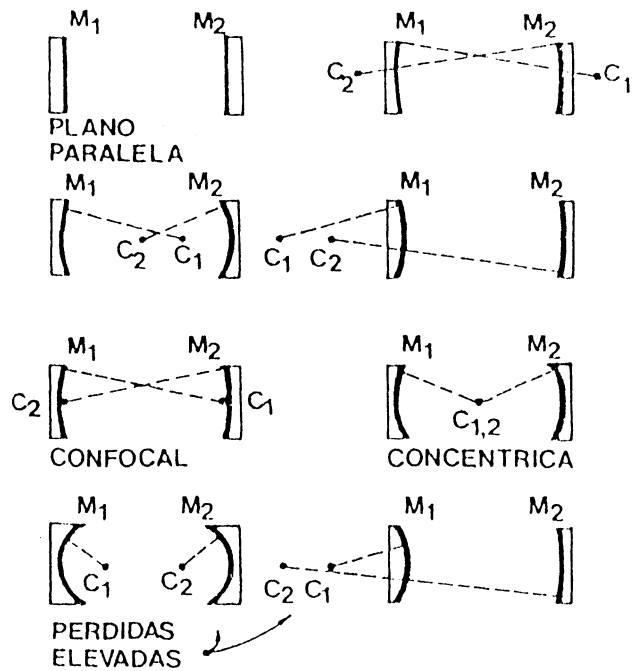


Figura 19 . Diferentes configuraciones de cavidades resonantes estables e inestables, constituidas por espejos planos y esféricos.

La Figura 20 representa la regla enunciada e indica cuándo la correcta selección de los radios de curvatura R_1 y R_2 de los dos espejos configuran cavidades estables o inestables.

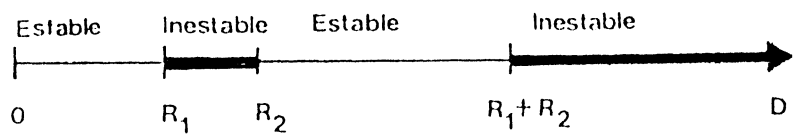


Figura 20. Caracterización de las configuraciones estables e inestables de cavidades resonantes.

Una cavidad sumamente estable, derivada de la confocal, es la denominada hemifocal. Se construye colocando un espejo plano en el foco del espejo cóncavo. El espejo plano produce una imagen simétrica del espejo cóncavo y constituye así una cavidad confocal. La cavidad hemifocal es tan estable como la confocal y, además, tiene la ventaja de la cavidad plano-paralela, ya que, a través del espejo plano, el haz de luz emitido por el láser tiene un frente de onda plano. La Figura 18 ayuda en esta discusión si se considera que en $Z=D/2$ se coloca el espejo cóncavo de radio $R=D$ y en $Z=0$ se ubica el espejo plano.

Las cavidades de pérdidas elevadas o inestables se emplean también, aunque sólo en casos especiales, como el de los láseres de CO_2 de altas potencias y energías, en los que la sección del tubo del láser puede alcanzar diámetros de hasta 1 m.

Si el radio W del haz de luz es mayor que el radio L de los espejos, las pérdidas por difracción serán tan elevadas que la cavidad así constituida no podrá sustentar modos resonantes.

En cambio, si $L \gg W$, el campo electromagnético será despreciable en los bordes de los espejos y la cavidad resultará estable y sustentará modos resonantes. Es decir, cuando el número de Fresnel aumenta, las pérdidas por difracción disminuyen y favorece sólo a un ínfimo número de modos transversales de oscilación. Las pérdidas por pasaje de la radiación dentro de las cavidades pueden estudiarse en función del número de Fresnel. G.D. Boyd y J.P. Gordon⁴⁰ calcularon estas relaciones para distintos tipos de cavidades. Algunos resultados aparecen representados en la Figura 21. La observación de la misma muestra claramente que la cavidad confocal es la mejor.

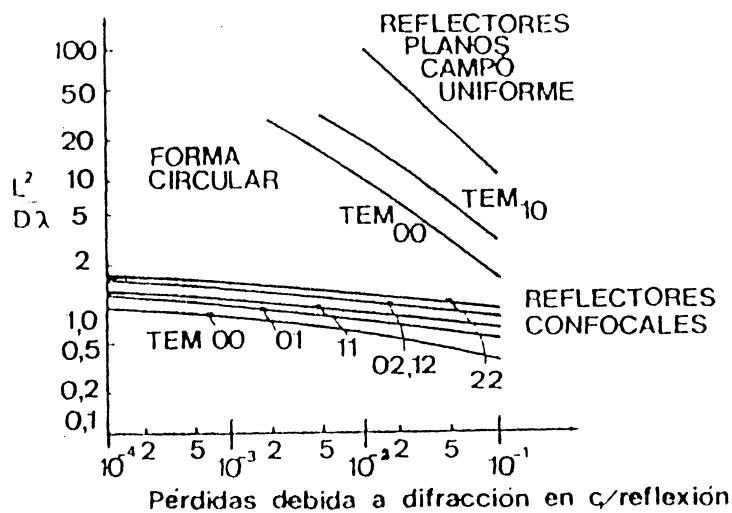


Figura 21. Pérdidas en las cavidades resonantes.

Bibliografía

- 1 - Einstein, A. - Phys. Z., 18, 121 (1971).
- 2 - Kramers, H.A. - Nature (Londres), 113, 310 y 114, 673 (1924).
- 3 - Ladenburg, R. - Z. Phys., 48, 15 (1928).
- 4 - Kopfermann, H. y Ladenburg, R. - Nature (París), 122, 438 (1928) y Z. Phys. Chem. Abt. A., 139, 375 (1928).
- 5 - Kuhn, W. - Z. Phys., 33, 408 (1925).
- 6 - Thomas, W. - Naturwiss., 13, 627 (1925).
- 7 - Heisenberg, W. - Z. Phys., 33, 879 (1925).
- 8 - (a) Füchtbauer, C., Joos, G., y Dinkelacker, O. - Ann. Phys., 71, 204 (1923). (b) Füchtbauer, C. y Bartels, H. - Z. Phys., 4, 337 (1921). c) Ladenburg, R. - Phys. Z., 27, 789 (1926).
- 9 - Fabrikant, V.A. - Tesis Doctoral, Universidad de Moscú, U.R.S.S. (1940).
- 10 - Cummins, H.Z., Abella, I., Heavens, O.S., Knable, N. y Townes, C.H. - Advances in Quantum Electronics, J.R. Singer (editor), Columbia University Press, New York, N.Y., pp.12-17 (1961).
- 11 - Rabinowitz, P., Jakobs, S. y Gould, G. - Appl. Opt., 1, 513 (1962).
- 12 - Patel, C.K.N. - Lasers, A.K. Levine (editor), Dekker, New York, N.Y., Vol.2, pp. 1-190 (1968).
- 13 - Gordon, J.P., Zeiger, H.J. y Townes, C.H. - Phys. Rev., 95, 282 (1954).
- 14 - Scovil, H.F.D., Feher, G. y Seidel, H. - Phys. Rev., 105, 762 (1957).
- 15 - McWhorter, A.L. y Meyer, J.W. - Phys. Rev., 109, 312 (1957).
- 16 - Makhov, G., Kikuchi, C., Lambe, J. y Terhune, R.W. - Phys. Rev., 109, 1399 (1958).
- 17 - Schawlow, A.L. y Townes, C.H. - Phys. Rev., 112, 1940 (1958).
- 18 - Schawlow, A.L. y Townes, C.H. - Patente N° 2929922 de los Estados Unidos de Norteamérica (30 de julio de 1958).
- 19 - Gould, G. - Patentes N° 953721 a N° 953727 de Inglaterra (6 de abril de 1959).
- 20 - Javan, A. - Phys. Rev. Lett., 3, 87 (1959).
- 21 - Maiman, T.H. - Nature (Londres), 187, 493 (1960) y Brit. Commun. Electron., 7, 674 (1960).
- 22 - Collins, R.J., Nelson, D.F., Schawlow, A.L., Bond, W., Garret, C.G.B. y Kaiser, W. - Phys. Rev. Lett., 5, 303 (1960).

- 23 - Javan, A., Bennett, W.R. y Herriott, D.R. - Phys. Rev. Lett., 6, 106 (1961).
- 24 - Bazhulin, P.A., Knyazev, I.N. y Petrash, G.G. - Sov. Phys. JETP., 23, 649 (1965).
- 25 - Hodgson, R.T. - Phys. Rev. Lett., 25, 494 (1970).
- 26 - Waynant, R.W., Shipman Jr., J.D., Elton, R.C. y Ali, A.W. - Appl. Phys. Lett., 17, 383 (1970).
- 27 - Waynant, R.W. - Phys. Rev. Lett., 28, 533 (1972).
- 28 - Hodgson, R.T. y Dreyfus, R.W. - Phys. Rev. Lett., 28, 536 (1972).
- 29 - Tagliaferri, A.A., Gallardo, M., Massone, C.A. y Garavaglia, M. - Phys. Lett., 45A, 211 (1973).
- 30 - Gallardo, M., Massone, C.A. y Garavaglia, M. - IEEE J. Quantum. Electron., QE-10, 525 (1974).
- 31 - Massone, C.A., Garavaglia, M., Gallardo, M., Calatroni, J.A.E. y Tagliaferri, A.A. - Appl. Opt., 11, 1317 (1972).
- 32 - Prokhorov, A.M. - Sov. Phys. JEPT. (Traducción al inglés), 7, 1140 (1958).
- 33 - Dicke, R.H. - Patente Nº 2851652 de los Estados Unidos de Norteamérica (1968).
- 34 - Yariv, A. y Gordon, J.P. - Proc. IEEE., 51, 4 (1963).
- 35 - Garrett, C.G.B., Kaiser, W. y Bond, W.L. - Phys. Rev., 124, 1807 (1961).
- 36 - Snitzer, E. y Osterberg, H. - J. Opt. Soc. Am., 51, 499 (1961).
- 37 - Fox, A.G. y Li, T. - Bell System Tech. J., 40, 453 (1961).
- 38 - Kogelnik, H. y Rigrod, W.W. - Proc. IRE., 50, 220 (1962).
- 39 - Connes, P. - Rev. d'Optique, 35, 37 (1956).
- 40 - (a) Boyd, G.D. y Gordon, J.P. - Bell Syst. Tech. J., 40, 489 (1961).
(b) Boyd, G.D. - Advances in Quantum Electronics, J.R. Singer (editor), Columbia University Press, New York, N.Y., pp. 318-328 (1961).
- 41 - Laures, P. - Phys. Lett., 10, 61 (1964).
- 42 - Lengyel, B.A. - Lasers. 2nd. edition. Wiley-Interscience, New York, N.Y., p.314 (1971).
- 43 - Patel, C.K.N. y Kerl, R.J. - Appl. Phys. Lett., 5, 81 (1964).

LASER II

J. O. Tocho

1. Introducción

En principio cualquier fuente de radiación puede convertirse en un láser. Es necesario invertir suficientemente la población de algún sistema material que interactúe con radiación de manera tal que la ganancia óptica producida por la emisión estimulada sea mayor que las pérdidas.

La emisión de un láser consiste de radiación, como la radiación emitida por cualquier otra fuente de ondas electromagnéticas. Sin embargo hay varias diferencias importantes y fundamentales entre la luz "incoherente" emitida por cualquier fuente térmica de luz, como una lámpara incandescente, y la luz "coherente" emitida por un láser. A menudo se describe a los láseres como fuentes espacial y temporalmente coherentes. Estas frases son todavía vagas y serán aclaradas a lo largo de este curso. Pero un punto importante a tener en cuenta es que todas estas propiedades de coherencia surgen principalmente de la presencia de un resonador o cavidad en la operación del láser y por supuesto de las características de la emisión estimulada.

Trataremos en este capítulo de mostrar como las transiciones atómicas (usaremos "átomo" como sinónimo de sistema material; los láseres funcionan con átomos, moléculas, iones, electrones en semiconductores, y aún con electrones libres) que proveen la ganancia necesaria y la cavidad con sus modos resonantes se combinan en el funcionamiento de un láser.

Nos referiremos a un láser ideal; otros capítulos describirán los láseres más utilizados en aplicaciones prácticas o en tareas de investigación. Cuando decimos un láser ideal nos referimos a aquel donde hemos optimizado todos los criterios de diseño y hemos reducido a un mínimo la infinidad de procesos que ocurren en un dispositivo real. Veremos sin embargo como se ponen de manifiesto las propiedades esenciales que distinguen a un láser de una fuente convencional.

2. Modos de una cavidad

Cualquier campo electromagnético arbitrario puede describirse en términos de su desarrollo en serie de Fourier.

Por ejemplo, para una cavidad prismática de volumen V , la componente cartesiana i del vector campo eléctrico \vec{E} , podrá escribirse,

$$E_i = \sum_m E_{i,m} \cdot \exp [i(\omega_m \cdot t - \vec{k}_m \cdot \vec{r})] + c.c., \quad (1)$$

donde los vectores de onda \vec{k}_m están condicionados, por las dimensiones de la cavidad, a tomar valores discretos,

$$k_x = n_x \cdot (\pi / L_x), \quad k_y = n_y \cdot (\pi / L_y), \quad k_z = n_z \cdot (\pi / L_z), \quad (2)$$

donde los $n_{x,y,z}$ son valores enteros 1, 2, ... y los $L_{x,y,z}$ representan los lados de la cavidad. (Figura 1 (a)).

Cada conjunto $m = \{n_x, n_y, n_z\}$ se denomina "modo de la cavidad" y tiene asociada una frecuencia ω_m . Entonces puede calcularse el número de modos que tienen frecuencia entre ω y $\omega + d\omega$,

$$N(\omega)d\omega = \frac{n^3 \omega^2 V}{\pi^2 c^3} \cdot d\omega = 8 \pi n^3 \frac{V}{\lambda^3} \cdot \frac{d\omega}{\omega}, \quad (3)$$

donde n representa el índice de refracción del material. (Figura 1 (b)).

En conclusión para describir la radiación en un volumen V y con frecuencia comprendida en un intervalo $d\omega$ alrededor de ω necesitamos conocer solamente $N(\omega) d\omega$ valores de las amplitudes $E_{i,m}$.

Ejemplos:

- 1) Láser de He-Ne. La línea espectral del neon utilizada tiene un ancho espectral de aproximadamente 1500 MHz y una longitud de onda de 633 nm. Para una cavidad cilíndrica de 1 mm de diámetro y 30 cm de longitud, son necesarios $4,67 \times 10^{10}$ modos para describir la radiación.
- 2) Diodo láser visible. En este caso el ancho espectral de emisión es mucho mayor ($\Delta\nu = 1.45 \times 10^{13}$ Hz) pero el volumen de confinamiento de modos es mucho

menor. Para un láser prismático de $1 \times 2 \times 200 \mu\text{m}$ de AsGa ($n = 3,6$) el número de modos es del orden de 5×10^5 .

3. Fotones

Podemos calcular la energía e.m. almacenada por unidad de volumen y en cada modo, dividiendo la cantidad de energía por el número de modos en el mismo intervalo de frecuencia,

$$\text{Energía del modo "m"} = \frac{V \cdot \rho(\omega) \cdot d\omega}{N(\omega) \cdot d\omega}, \quad (4)$$

y, si esta energía la pensamos acumulada en fotones de energía $\hbar\omega$ (que significa $h\nu$), podemos calcular el número de fotones por modo,

$$n_m = \frac{\text{Energía en el modo "m"}}{\hbar\omega} = \frac{\pi^2 c^3}{n^3 \hbar \omega^3} \rho(\omega). \quad (5)$$

Ejemplo: Si $\rho(\omega)$ corresponde a radiación e.m. en equilibrio con un cuerpo a temperatura T , el número de fotones por modo está representado en la Figura 2.

4. Emisión estimulada versus emisión espontánea.

Supongamos un sistema atómico en un campo de radiación con densidad espectral $\rho(\omega)$. El número de transiciones espontáneas por unidad de tiempo estará dado por el coeficiente de Einstein A_{21} y el número equivalente de transiciones inducidas será $B_{21} \cdot \rho(\omega)$. (Figura 3).

Utilizando la relación entre los coeficientes A y B de Einstein, podemos escribir:

$$B_{21} = \frac{c^3}{8n^3 \pi \hbar \nu^3} A_{21}, \quad (6)$$

$$W_I = B_{21} \cdot \rho(\omega) = \frac{c^3}{n^3 8 \pi \hbar \nu^3} \cdot A_{21} \rho(\omega) = \frac{\pi^2 c^3}{n^3 \hbar \omega^3} \rho(\omega) \cdot A_{21} = n_m \cdot W_E \quad (7)$$

Emisiones inducidas por unidad de tiempo al modo "m"	=	Emisiones espontáneas al modo "m" por unidad de tiempo	×	Número de fotones en el modo "m"
--	---	--	---	-------------------------------------

5. Número efectivo de modos.

El número efectivo de modos que interactúan con una transición atómica de frecuencia ω_a y ancho a media altura $\Delta\omega_a$ (para una forma de línea lorentziana) es el número de modos de cavidad comprendidos en un intervalo de frecuencias igual a $(\pi/2)\Delta\omega_a$,

$$p = N(\omega)(\pi/2)\Delta\omega_a = 4\pi^2 n^3 \frac{V}{\lambda^3} \cdot \frac{\Delta\omega_a}{\omega_a} . \quad (8)$$

El ancho de una distribución rectangular equivalente que tiene la misma altura y la misma área que una lorentziana es $(\pi/2)\Delta\omega_a$, en lugar de simplemente $\Delta\omega_a$ (Figura 4).

6. Cinética de fotones.

La población de los modos de cavidad (número de fotones en cada modo) se altera por la interacción con los átomos y por la conexión con el exterior de la cavidad (pérdidas). Para cada modo el balance de población responde a,



$$\frac{dn_i}{dt} = \frac{A_{21}}{p} \cdot N_2 + \frac{A_{21}}{p} \cdot n_i \left[N_2 - \frac{g_2}{g_1} N_1 \right] - \gamma_i n_i . \quad (9)$$

que, para un láser de 4 niveles se simplifica ya que $N_2 \gg g_2 / g_1 \cdot N_1$, y por lo tanto se puede despreciar la absorción,

$$\frac{dn_i}{dt} = (A_{21} N_2 / p)(n_i + 1) - \gamma_i n_i . \quad (10)$$

En el estado estacionario esta ecuación tiene solución simple,

$$n_i = \frac{N_2}{p\gamma_i / A_{21} - N_2} = \frac{N_2}{N_{iu} - N_2} \quad (11)$$

donde hemos definido la población umbral como aquella necesaria para hacer significativa la población de fotones en el modo correspondiente (Figura 5),

$$N_{iu} = p\gamma_i / A_{21} \quad (12)$$

La Figura 6 representa el caso en que se alcanza el umbral para muy pocos modos alrededor del llamado 1 (láser) y en lo que sigue supondremos que es sólo uno de ellos.

La discriminación se consigue a través del factor de pérdida γ_i y más finamente por la variación del coeficiente A_{21} con la frecuencia (ver más abajo).

$$\gamma = \text{coeficiente de pérdida} = \text{pérdida por paso} / \text{tiempo de tránsito}$$

En una cavidad abierta, que consiste en un par de espejos paralelos, la mayoría de los modos que no se propagan perpendicularmente a éstos pierden su energía en un solo paso. Tienen por lo tanto un factor de pérdida alto y un umbral muy elevado. Los modos que se propagan en la dirección apropiada se conectan con el exterior a través de la transmisión de los espejos, para ellos,

$$\gamma_e = [\alpha L + (1 - \sqrt{R_1 R_2})] / (nL / c) \quad (13)$$

donde α es la pérdida distribuida y R_1, R_2 son las reflectividades de cada espejo (Figura 7).

7. Cinética de las poblaciones.

Consideremos un modelo simple que cumpla con las siguientes condiciones:

- i) la población del nivel inferior de la transición es despreciable ($N_2 \gg g_2 N_1 / g_1$);
- ii) la desexcitación espontánea del nivel superior es totalmente radiativa ($A_{21} = 1 / \tau_2$);

entonces,

$$\frac{dN_2}{dt} = \text{bombeo externo} - \text{emisión eea} - \text{spontán emisión estimulada}$$

$$\frac{dN_2}{dt} = B - \left(N_2 \frac{A_{21}}{p} \cdot \sum_i n_i \right) - N_2 A_{21} \quad (14)$$

$$\frac{dN_2}{dt} = B - N_2 \frac{A_{21}}{p} \cdot \sum_{i=1}^p (n_i + 1) \quad (15)$$

$$\frac{dN_2}{dt} = B - N_2 A_{21} - N_2 \frac{A_{21}}{p} \cdot m \begin{cases} p-1 \text{ modos tienen } n_i \ll 1 \\ 1 \text{ modo puede tener } n_i \gg 1 \end{cases} \quad (16)$$

La población interactúa con los p modos. De éstos, $p-1$ no alcanzan el umbral y por lo tanto no tienen nunca un número importante de fotones. El modo 1 alcanza el umbral. De todas maneras para bombeos bajos, o sea muy lejos del umbral, el número de fotones por modo es muy pequeño ($n_i \ll p$) y la solución, en estado estacionario, de la ecuación (16) es simple,

$$N_2 = B / A_{21} \quad , \quad (17)$$

$$N_2 = r \cdot N_{2u} \quad , \quad (18)$$

siempre que $B \ll B_u$ y si definimos $N_{2u} = p \cdot \gamma_l / A_{21}$, $B_u = p \cdot \gamma_l$ y el bombeo normalizado adimensional, $r = B / B_u$ (Figura 8).

8. Fotones y poblaciones.

La definición del bombeo umbral B_u , y del bombeo normalizado, $r = B / B_u$, permite reescribir las ecuaciones (10) y (16) de la siguiente manera,

$$\frac{dn_i}{dt} = \frac{A_{21}}{p} N_2 (n_i + 1) - \frac{B_u}{p} n_i \quad , \quad (19)$$

$$\frac{dN_2}{dt} = r B_u - N_2 A_{21} - N_2 \frac{A_{21}}{p} n_i \quad , \quad (20)$$

despejando N_2 de la ecuación (20) para el estado estacionario,

$$N_2 = \frac{r B_u}{A_{21} \cdot (1 + n_i / p)} \quad , \quad (21)$$

y que sustituyendo en (19) conduce a,

$$\frac{n_l^2}{p} + (1-r)n_e - r = 0 \quad (22)$$

que admite como única solución,

$$n_l = \frac{p}{2} \cdot \left[(r-1) + \sqrt{(r-1)^2 + 4r/p} \right] , \quad (23)$$

Veamos algunos comportamientos simples:

I) Muy por debajo del umbral, $r \ll 1$

$$n_l = r , \quad (24)$$

$$N_2 = r \cdot N_{2u} , \quad (25)$$

II) Cerca del umbral, $r < 1$

$$n_l = r / (1-r) , \quad (26)$$

$$N_2 = r \cdot N_{2u} , \quad (27)$$

III) En el umbral, $r = 1$

$$n_l = \sqrt{p} , \quad (28)$$

$$N_2 = N_{2u} , \quad (29)$$

IV) Por encima del umbral, $r \gg 1$, $n_l \gg 1$,

$$n_l = p(r-1) , \quad (30)$$

$$N_2 = N_{2u} , \quad (31)$$

La Figura 9 representa los comportamientos de la población y del número de fotones por modo (láser y no-láser) antes y después del umbral.

9. Discriminación de modos.

El comportamiento tan agudo del número de fotones en función del bombeo normalizado y la brusca saturación del crecimiento de la población en la zona del umbral explica el hecho de que habiendo $p = 10^5 - 10^{10}$ potenciales modos de oscilación, un láser pueda extraer toda potencia suministrada por encima del umbral en un solo modo, el modo "l".

El modo "l" será el que tenga menores pérdidas y por lo tanto alcance primero el umbral. Recordar que A_{21} depende de la frecuencia,

$$A_{21}(\omega) = A_{21} \cdot g(\omega) \quad , \quad (32)$$

si $N_{2u} = p \cdot \gamma / A_{21}(\omega)$, es claro que alcanzará primero el umbral el modo que esté más próximo al centro de la línea, donde $g(\omega) = 1$.

Podemos estimar un poco más realísticamente si otros modos alcanzan a oscilar de la siguiente manera. Supongamos un modo i con $A_{21}(\omega) = A_i$, el n_i será:

$$n_i = \frac{N_2}{N_{iu} - N_2} = \frac{N_{2u}}{N_{iu} - N_{2u}} = \frac{A_i}{A_{21} - A_i} \quad , \quad (33)$$

Si el modo oscila $\eta_i = \sqrt{p}$, entonces,

$$\frac{A_i}{A_{21} - A_i} = \sqrt{p} \quad , \quad (34)$$

$$A_i = A_{21} (1 - 1/\sqrt{p}) \quad , \quad (35)$$

Oscilarán, entonces, solamente los modos con $A_{21}(\omega) \geq A_{21} (1 - 1/\sqrt{p})$.

El valor muy grande de p asegura que sólo muy pocos modos cumplan esta condición dentro del ancho de línea de la transición (Figura 10).

10. Emisión espontánea y emisión estimulada.

Debajo del umbral el modo l no es diferente a ninguno de los p modos de la cavidad comprendidos en la transición. Podemos calcular la intensidad emitida en cualquier dirección,

$$I_{esp} = \hbar\omega c r p (\Delta\Omega / 4\pi) , \quad r < 1 . \quad (36)$$

Por encima del umbral, en cambio, es un solo modo el que concentra el mayor número de fotones,

$$I_{ind} = \hbar\omega c (r - 1) p , \quad r > 1 , \quad (37)$$

$$I_{total} = \hbar\omega c p \Delta\Omega / 4\pi + \hbar\omega c (r - 1) p . \quad (38)$$

Las medidas experimentales de este comportamiento son difíciles porque los detectores utilizados no captan todos los modos isotrópicamente emitidos. Un detector de ángulo sólido $\Delta\Omega$ colocado enfrente de un láser muestra un abrupto cambio en la señal al llegar al umbral provocado por el cambio de emisión de p modos a un modo (Figura 11). Es más fácil observar el cambio espectral en la emisión al pasar el umbral (Figura 12).

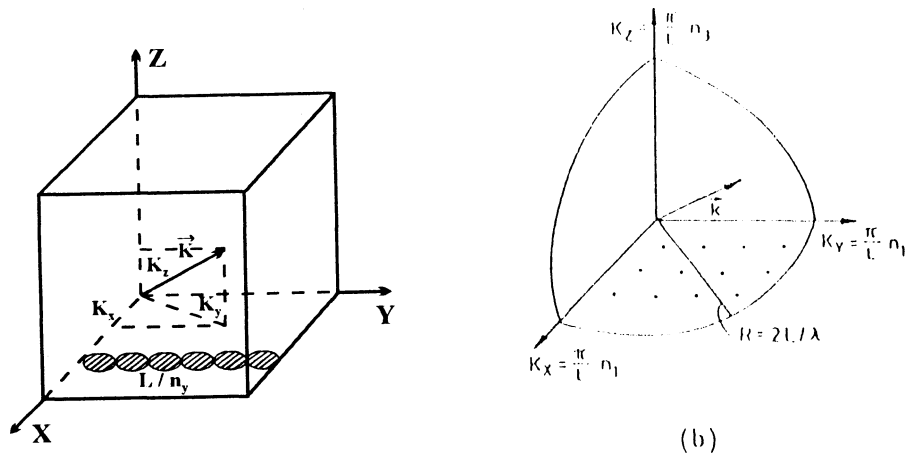


Figura 1: Modos de una cavidad prismática. **(a)** Los valores de k son discretos. **(b)** Como se cuentan los modos que tienen frecuencia determinada.

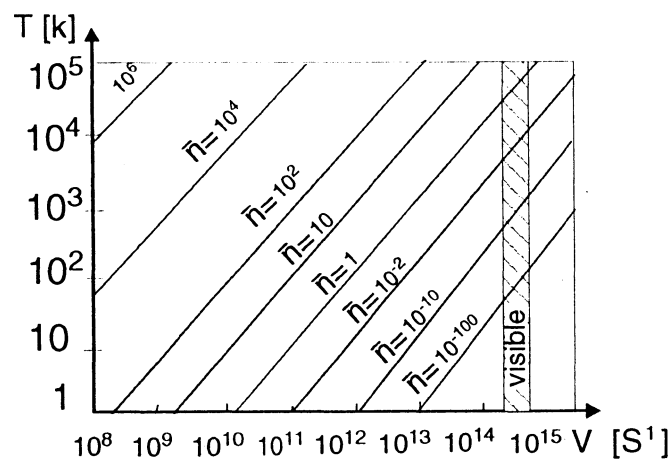
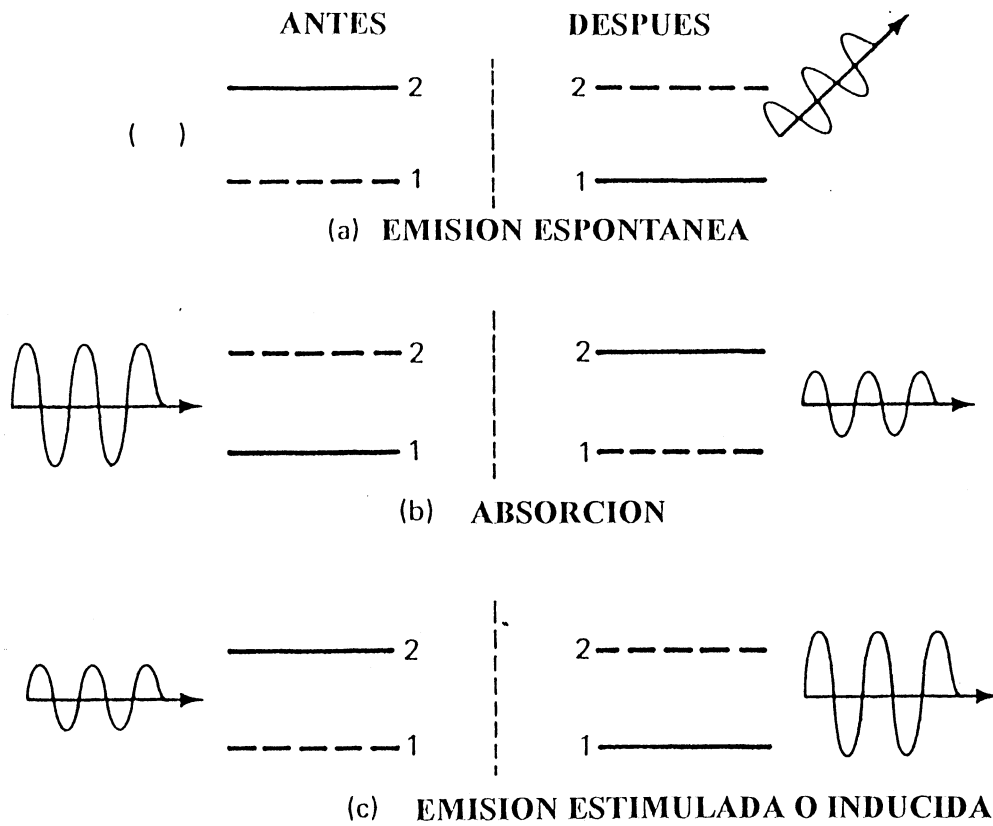


Figura 2: Número medio de fotones por modo en equilibrio térmico (cuerpo negro) en función de la temperatura T y de la frecuencia.



Interacción de la radiación y los átomos.

La emisión estimulada añade un fotón con:

1. *la misma frecuencia*
2. *la misma polarización*
3. *la misma dirección de propagación y*
4. *la misma fase que el fotón original.*

Figura 3: Las 3 maneras de interacción entre los átomos y la radiación. La emisión estimulada añade un fotón con idéntica frecuencia, polarización, dirección de propagación y fase que el fotón original.

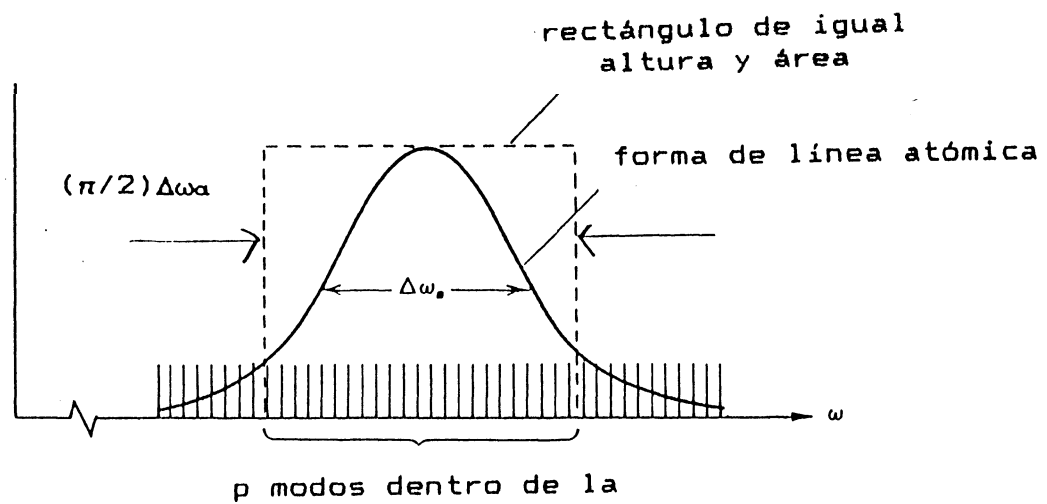


Figura 4: Modos y forma de línea.

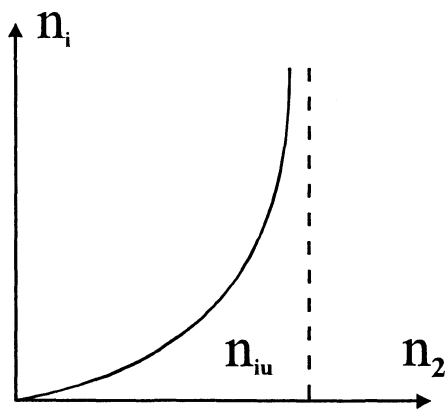


Figura 5: Número de fotones en el modo i cuando crece la inversión de población hasta llegar al umbral.

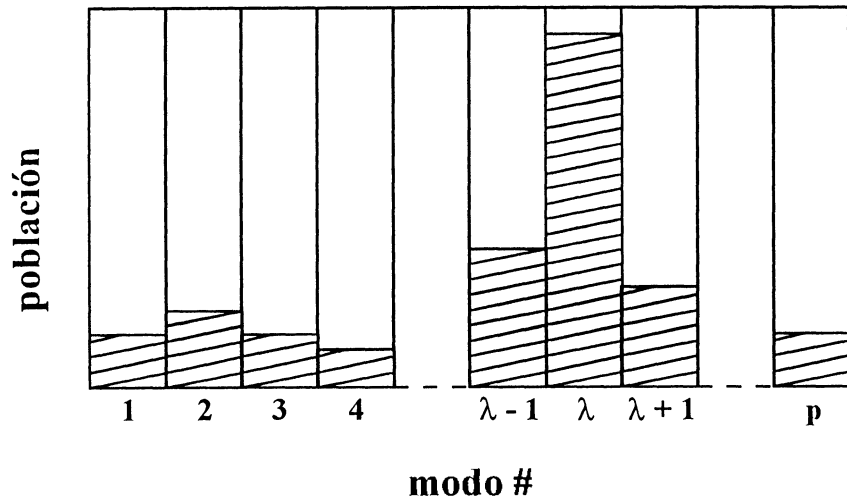


Figura 6: Número de fotones por modo.

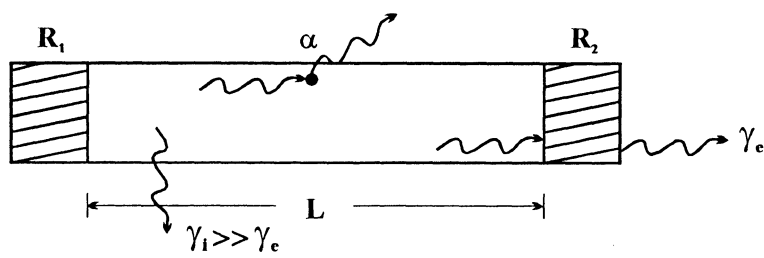


Figura 7: Pérdidas en una cavidad.

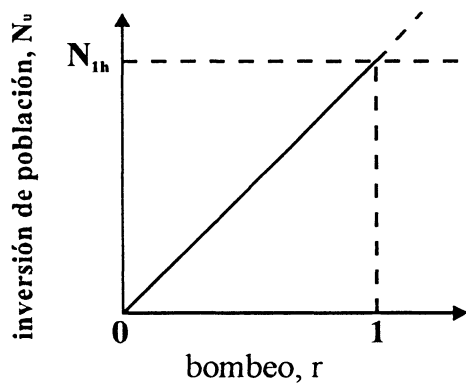


Figura 8: Población en función del bombeo por debajo del umbral.

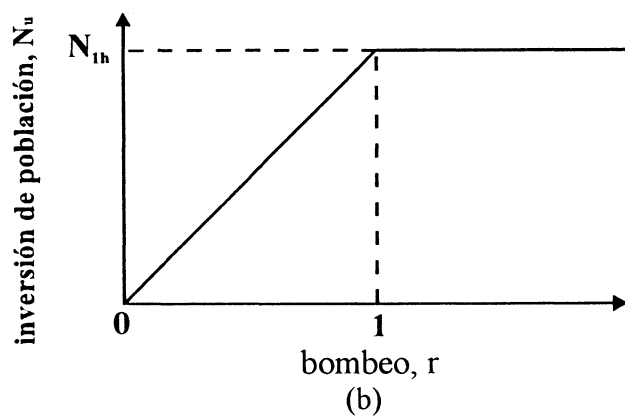
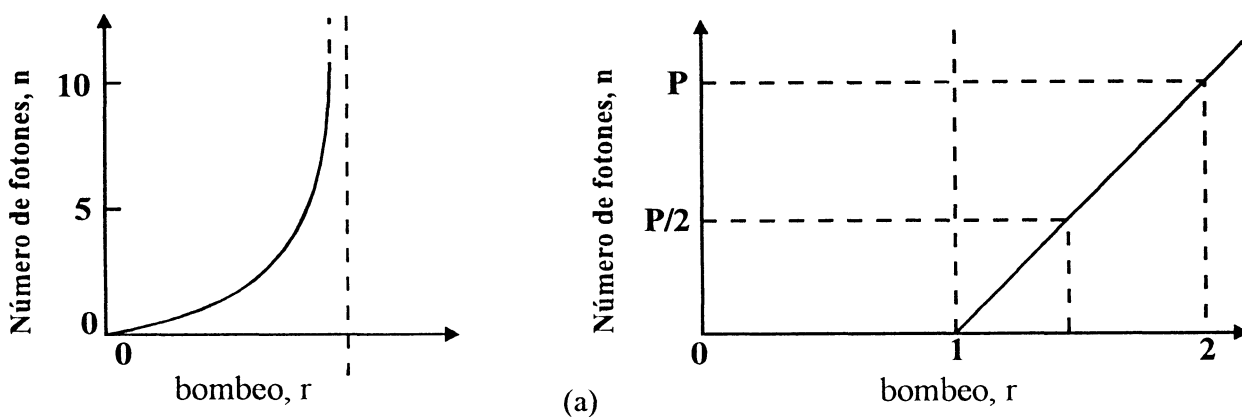


Figura 9: (a) Número de fotones por modo. (b) Población.

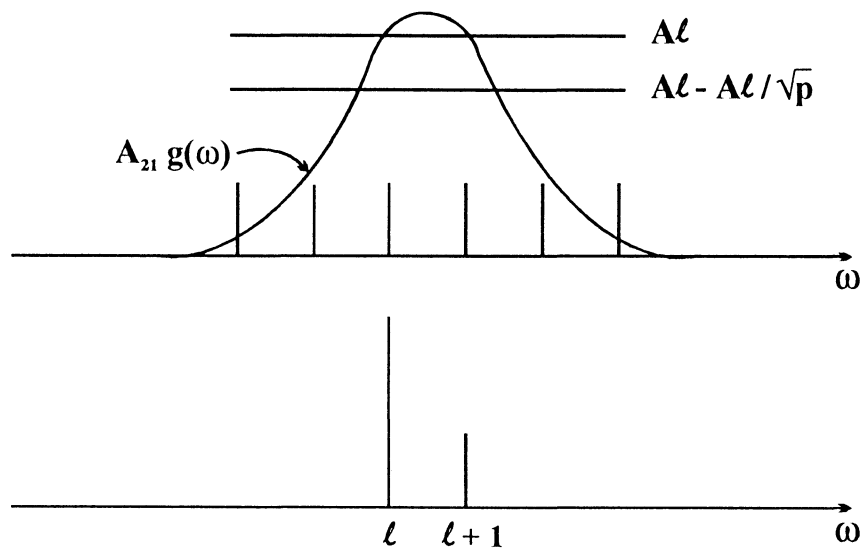


Figura 10: Una pequeña disminución en $A_{21} (\omega)$ hace que modos cercanos al modo " ℓ " no oscilen.

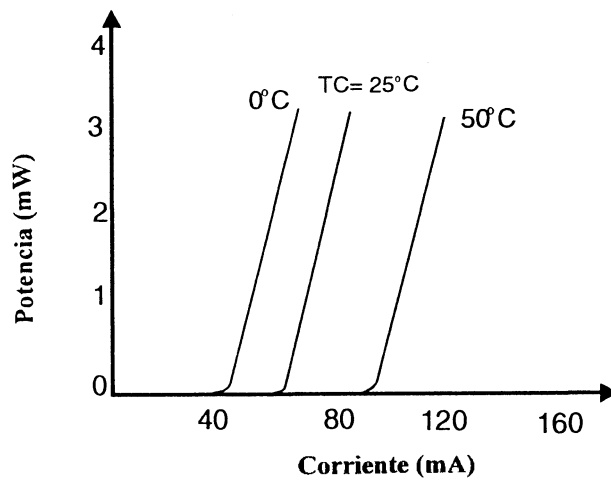


Figura 11: Potencia versus corriente en un láser de diodo semiconductor.

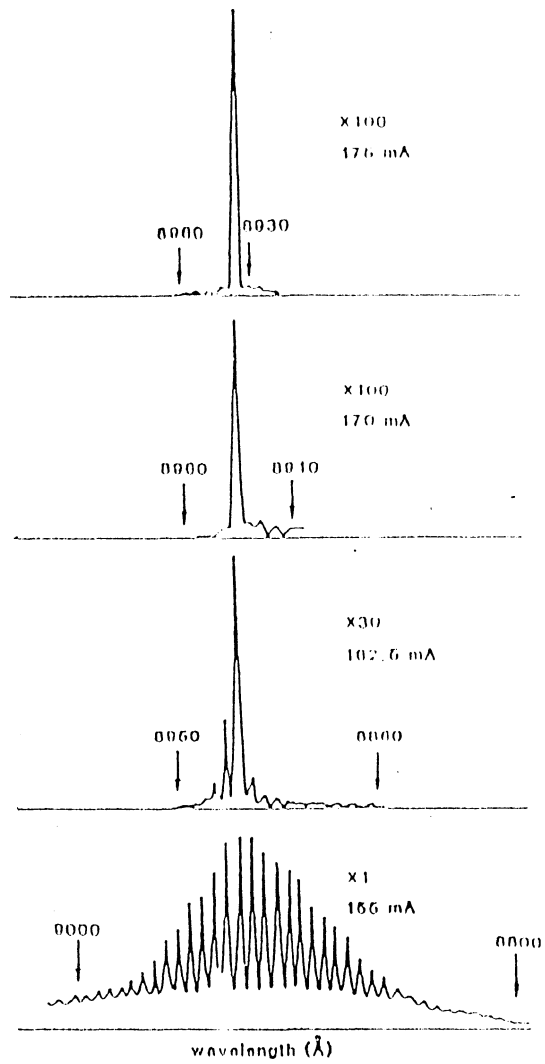


Figura 12: Espectro de emisión de un láser de diodo semiconductor antes y después del umbral.

TIPOS DE LASER. ESTADO ACTUAL Y TENDENCIAS.

L.B. Scaffardi
D.C. Schinca

Introducción

Esta clase tiene por objeto dar una breve introducción al tema de los diversos tipos de láseres que pueden encontrarse en la actualidad. Debido a la gran aplicación de estos dispositivos en una diversidad de áreas, resultaría imposible una enumeración exhaustiva de todos los tipos de láseres, ya sea clasificándolos por naturaleza, modificaciones de un mismo láser o láseres diseñados para alguna aplicación específica. Desde el funcionamiento exitoso del primer láser en 1960, se inició una espiral siempre creciente de interacción directa entre: aplicaciones → láser adecuado → nuevas aplicaciones → láseres más perfeccionados. Se puede decir que siempre se puede adaptar mejor un láser a los problemas planteados.

En campos de aplicabilidad que van desde la alineación de pequeños sistemas ópticos a la fusión inducida por láser, pasando por la medicina, la ingeniería genética, la espectroscopía nuclear, la ingeniería civil, las comunicaciones, sin contar las aplicaciones militares, es razonable pensar que los láseres que puedan adaptarse a ellas serán de características muy diversas. Por ello, con el fin de ordenar la descripción de los distintos tipos de láseres, resulta necesario recurrir a determinadas clasificaciones según frecuencia de emisión, tipo de emisión (continua o pulsada), nivel de potencia, tipo de material activo, etc. El criterio para elegir alguna de éstas depende en gran medida del tipo específico de aplicación requerida. Así, para poder dar un pantallazo general del tema, vamos a clasificarlos según el estado del material activo, es decir, en sólidos, líquidos y gaseosos. Dentro de cada clase, nos limitaremos a describir los láseres más representativos, que son los que se utilizan con mayor frecuencia tanto en laboratorios como en la industria y que se han ido optimizando luego de un extenso proceso de investigación y desarrollo. Los lectores interesados en mayores detalles pueden consultar la bibliografía listada al final de este apunte.

El esquema que seguiremos es el siguiente: en la sección I haremos un breve comentario sobre los sistemas de excitación o de "bombeo", poniendo énfasis en aquellos empleados en los láseres que vamos a describir más adelante. En la sección II se describirán brevemente ciertas técnicas o mecanismos por los cuales se puede modificar las características de la salida de un láser para hacerla más versátil y adecuada a determinadas aplicaciones. En las secciones III a VI entraremos ya en la descripción de los distintos tipos de láseres, comentando brevemente sus aplicaciones más comunes. Finalmente, en la sección VII trataremos las tendencias actuales en cuanto a los tipos de láseres como herramientas de aplicación.

Cabe destacar que esta clase resulta de una revisión y actualización de una versión previa escrita por el Dr. Hector F. Ranea Sandoval y la Dra. Lucía B. Scaffardi.

I. Procedimientos de excitación

El medio activo de un láser puede ser sólido (iones en matrices cristalinas o amorfas), líquido (moléculas en solución) o gaseoso (átomos, iones o moléculas en estado de gas o vapor). Para que estos sistemas puedan emitir radiación, deben ser excitados a partir de sus niveles inferiores de emisión hacia los llamados niveles superiores y para que tal radiación pueda ser considerada emisión estimulada, debe existir un desbalance neto entre ambas poblaciones, por lo que la excitación debe ser lo más selectiva que se pueda.

Existen dos métodos que se aplican en prácticamente todos los láseres, a saber, la excitación óptica y la excitación por descarga eléctrica. En general, la primera es especial para sólidos y líquidos, mientras que la segunda es la indicada para láseres gaseosos. Cabe destacar un caso especial, que son los láseres de semiconductor (sólido) que se bombean por inyección de portadores (corriente eléctrica). Vamos a describir brevemente los dos primeros, dejando el último para cuando veamos los láseres de diodo.

I. 1- Excitación óptica

Cuando un fotón posee la energía suficiente como para provocar una transición óptica desde un nivel inferior a uno excitado en un sistema atómico o molecular, se dice que el mismo se ha excitado ópticamente. Si esta excitación invierte la población de los niveles involucrados, el método resulta el adecuado para que el material activo opere como láser.

El método de bombeo óptico es particularmente adecuado para materiales sólidos y / o líquidos debido a que los niveles de energía de los iones en matrices cristalinas se encuentran ensanchados por el campo eléctrico del cristal o por los numerosas colisiones con las moléculas de solvente en el caso de los colorantes en solución (dyes). Este ensanchamiento aumenta mucho la probabilidad de que un fotón sea absorbido por la transición en cuestión.

Un sistema práctico muy utilizado para el bombeo óptico son las lámparas de flash. En éstas, un condensador cargado a tensiones entre 2 kV y 8 kV se descarga (gracias a un disparo auxiliar de alta tensión para preionización) a través de la lámpara generando un pulso de corriente de aproximadamente 1 kA de pico en tiempos que varían entre 0,1 μ s y 10 μ s (dependiendo de la carga del condensador y de la impedancia del plasma formado en la lámpara). Las numerosas colisiones de electrones con los átomos del gas (usualmente Xe o Kr a presiones entre 100 Torr y 400 Torr) provocan la excitación de estos con la posterior emisión de fluorescencia, que genera un espectro continuo entre los 200 nm y los 2500 nm (figura 1).

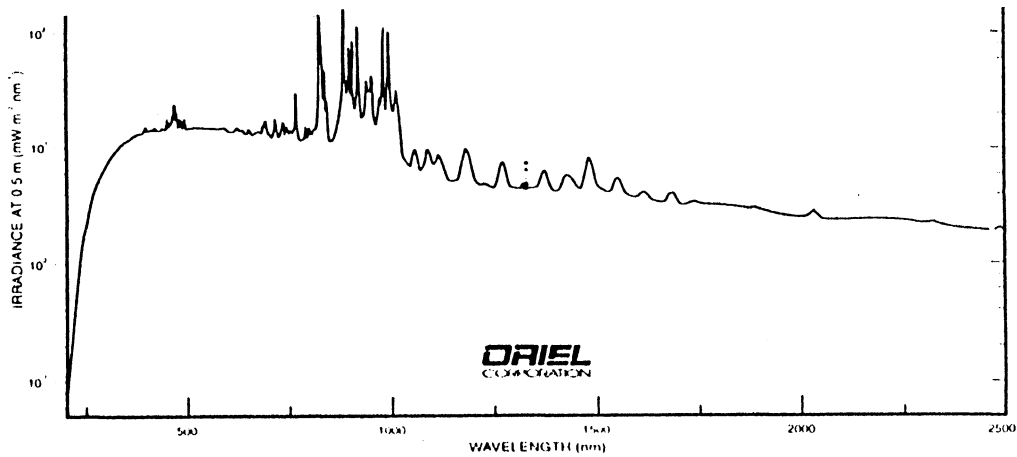
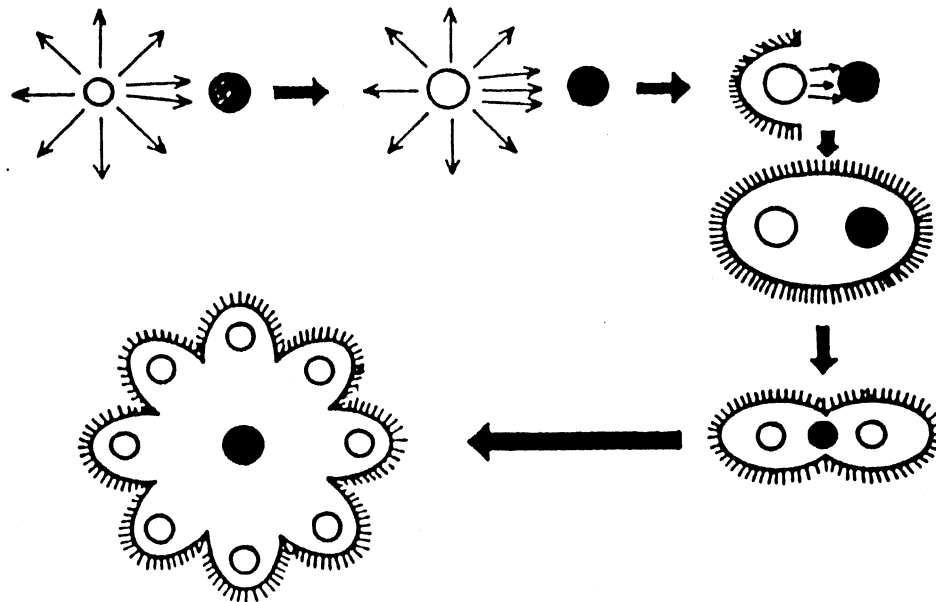


FIGURA 1

Dicha distribución espectral debería ser lo más similar (en condiciones ideales) a la curva de absorción del material a ser bombeado, de tal forma de aprovechar la mayor cantidad de potencia emitida por la lámpara. Además, es necesario que esta radiación sea canalizada eficientemente al material de bombeo, hecho que depende fuertemente de la geometría en la que se dispongan lámpara y material. Así, existen tres tipos de cuestiones (que llamaremos eficiencias) a optimizar y que pueden definirse de la siguiente manera:

i) Eficiencia de transferencia: es el cociente entre la potencia total absorbida por el material (todos los niveles y no solo los de interés láser) y la potencia emitida por la lámpara. Existen muchos modos de transferir lo emitido por la lámpara al material, cada uno mejor adaptado a determinados láseres. En general se utilizan cavidades reflectoras elípticas en las que la lámpara y el material se colocan en los focos de la misma, de modo que toda la radiación emitida por la primera se refleje en las paredes e incida sobre el segundo (figura 2).






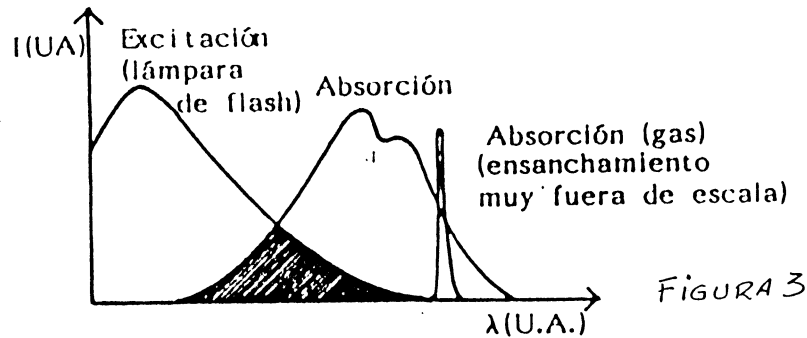
-  superficies espejadas
-  muestra, colocada en un foco de la elipse
-  lámpara de flash, ubicada en el foco conjugado

FIGURA 2

ii) Eficiencia espectral: el material absorberá ópticamente solo una parte de la potencia total absorbida. Este número depende del solapamiento entre la curva de absorción del nivel de interés y la de emisión del sistema de bombeo. Para optimizar esta eficiencia, se recurre a diversos métodos que van desde el diseño de las lámparas de flash a la adición de sustancias que transformen convenientemente la emisión para una mejor adaptación espectral (figura 3).



iii) Eficiencia cuántica de bombeo: resulta del cociente entre el número de átomos excitados al primer nivel del sistema y los que realmente pueblan el nivel láser superior. El hecho de que esta eficiencia no sea la unidad se justifica debido a que existen otros canales de desexcitación del nivel al que fue bombeado el material. La figura 4 ilustra el significado de este tipo de eficiencia para el caso de los láseres de colorantes (que veremos más adelante): no todas las moléculas excitadas del nivel 0 al 1 terminan poblando el nivel de interés (el 2) y del 2 no todas se transfieren al 3 en forma radiativa. Los niveles T son por lo general metaestables y su decaimiento radiativo al nivel 3 suele tener una vida media del orden del segundo.

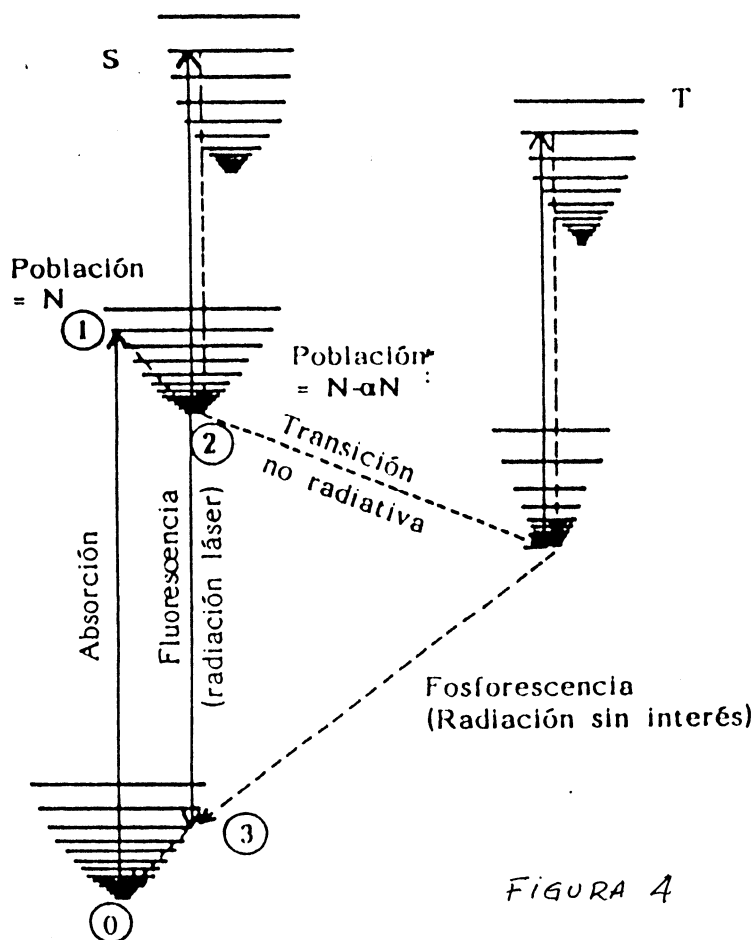


FIGURA 4

El bombeo óptico también suele ser efectuado por medio de láseres que emiten en zonas espectrales de gran absorción del material. Tal es el caso de los láseres de colorantes continuos o pulsados bombeados por un láser de argón o de nitrógeno respectivamente. Estos sistemas, aunque son más caros que los bombeados por lámparas de flash, poseen una mayor eficiencia espectral junto con una mejor calidad de haz de salida. Otro caso de destacar, y que comentaremos con detalle más adelante, es el de los láseres sólidos bombeados por láseres de semiconductor, que ha permitido, entre otras cosas, la “miniaturización” de sistemas láser de potencia media (dimensión típica de unos 10 a 15 cm).

I. 2- Excitación eléctrica.

Por descarga eléctrica en un gas entendemos la generación de una corriente en el seno del mismo mediante la aplicación, ya sea en forma continua o pulsada, de una determinada tensión sobre él. La aplicación de un voltaje continuo suficientemente elevado sobre un tubo cargado con un gas a una presión apropiada, produce un fenómeno conocido como ruptura del gas, que pasa a conducir con características de baja resistencia. En ese instante, la tensión cae abruptamente a valores bajos para estabilizar la corriente que circula por el tubo. Se genera así un plasma luminoso estable que puede llegar a invertir la población de un par de niveles en determinados gases.

En caso de descargas pulsadas, el fenómeno es análogo, excepto que puede diseñarse de modo de optimizarla para los fines buscados (baja impedancia, corto tiempo de crecida del pulso de corriente, etc). Los esquemas típicos para excitaciones eléctricas, tanto continuas como pulsadas, pueden verse en la figura 5. En el caso de las primeras, una alta tensión rectificadora (A.T.), descarga a través de una resistencia (R, R_B) sobre un tubo de descarga (T), fijando el condensador (C) la tensión sobre el canal de descarga. En el caso de descargas pulsadas, un condensador cargado a determinada tensión, descarga a través de una llave rápida (S) sobre el canal de descarga (T).

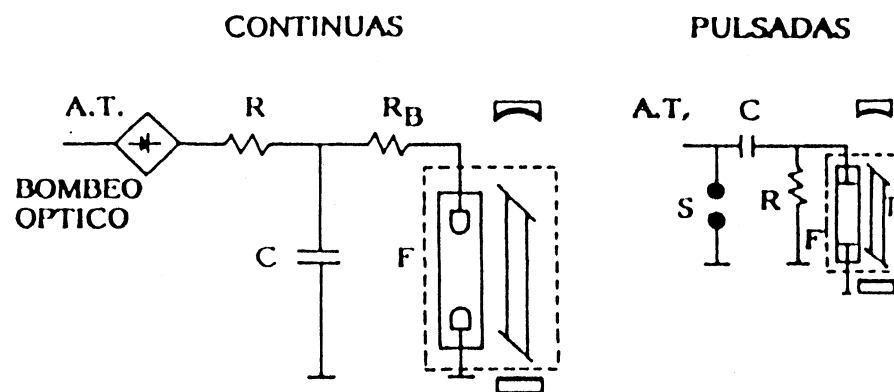
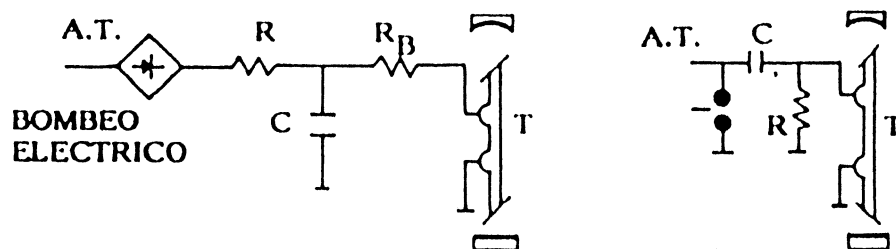


FIGURA 5



II. Técnicas de modificación de la salida láser.

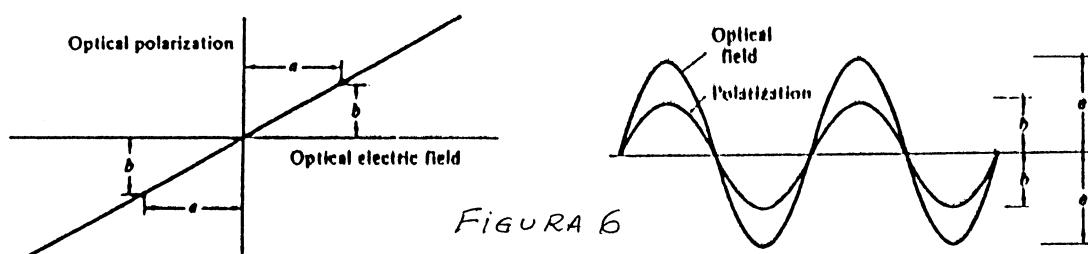
Desde la entrada en funcionamiento del primer láser, las características temporales como de potencia de la salida del mismo fueron siempre un tema de interés tanto de investigación como práctico. Es sabido que los láseres pueden funcionar en modo continuo o pulsado, dependiendo del sistema de excitación y del material activo mismo. Sin embargo, la alta intensidad inherente a la emisión estimulada, los modos de oscilación permitidos por la cavidad láser y las características espectroscópicas de ciertos materiales activos fueron la base para el desarrollo de técnicas que modifican la estructura del haz de salida y le confieren al láser una gran flexibilidad para el estudio de diferentes problemas. En esta sección vamos a repasar muy brevemente las tres técnicas más importantes, a saber: generación de armónicas, conmutación del Q de la cavidad (Q-switch) y puesta en fase de modos (mode-locking).

II. 1 Generación de armónicas

La generación de armónicas es un hecho largamente conocido (evitado a veces y deseado otras) dentro del ámbito de la ingeniería electrónica. Dado un circuito eléctrico, es posible lograr que oscile no solo a la frecuencia ω de una señal de entrada, sino también al doble de esa frecuencia utilizando componentes no lineales. Ejemplos conocidos son los circuitos eléctricos con bobinados con núcleo de ferrite o la respuesta tensión-corriente de un diodo en la zona de los 0,6 V.

En Óptica, es posible observar un fenómeno parecido cuando un haz de luz intenso atraviesa un cristal. Es de destacar que los efectos ópticos no lineales se conocían (al menos teóricamente) desde fines del siglo pasado, cuando el físico escocés James C. Maxwell desarrolló la teoría del campo electromagnético y demostró que los fenómenos ópticos (refracción, difracción, interferencia, propagación en medios transparentes, etc) podían interpretarse desde el punto de vista de la propagación de ondas electromagnéticas, de frecuencia adecuada, tanto en el vacío como en medios materiales. Sin embargo, los efectos de segundo orden resultan totalmente despreciables frente a los de primer orden para las intensidades de campo eléctrico conocidas en esa época, de tal modo que fueron olvidados por casi un siglo. El advenimiento del láser con su alta intensidad de radiación hizo resurgir los efectos de segundo orden y los hizo pasar de curiosidades de laboratorio a aplicaciones prácticas.

La propagación de luz a través de un medio transparente puede considerarse como un proceso en el cual los dipolos eléctricos (formados por los electrones ligados al núcleo de los átomos constituyentes del cristal) son forzados a oscilar por el campo eléctrico incidente y por lo tanto a emitir radiación. Si este campo es pequeño, la amplitud de oscilación de los dipolos también lo será y éstos seguirán en forma lineal la oscilación externa. Así, si entra una onda de frecuencia ω , saldrá del cristal luz de la misma frecuencia (efecto de primer orden o lineal), según se esquematiza en la figura 6.



Si ahora la amplitud del campo eléctrico incidente es lo suficientemente grande como para desplazar los electrones mucho más lejos del núcleo, estos salen fuera de la región lineal de atracción coulombiana y, al oscilar, generan una polarización que no sigue en amplitud al campo externo, tal como se representa en la figura 7.

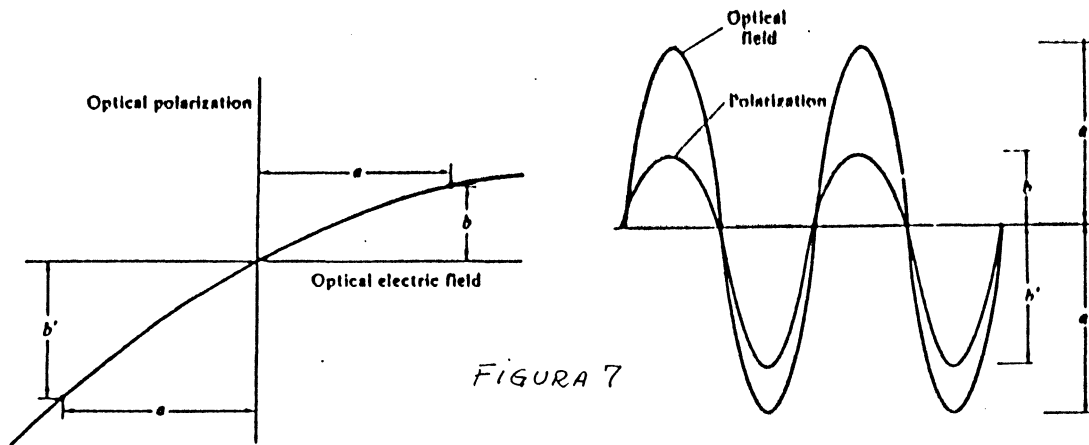


FIGURA 7

Esta polarización no lineal puede desarrollarse en serie de potencias en función del campo incidente E :

$$P = \chi_1 \cdot E + \chi_2 \cdot E^2 + \chi_3 \cdot E^3 + \dots = P_1 + P_2 + P_3 + \dots$$

donde χ_i son las susceptibilidades eléctricas de primer, segundo y tercer orden respectivamente. Si suponemos que la onda incidente es plana y monocromática, se puede escribir como:

$$E = E_0 \text{sen}(\omega \cdot t)$$

y si reemplazamos en el segundo término del desarrollo en serie de la polarización, nos queda que:

$$P_2 = \chi_2 \cdot E_0^2 \text{sen}^2(\omega \cdot t) = (1/2) \chi_2 \cdot E_0^2 (1 - \cos(2 \cdot \omega \cdot t))$$

Vemos entonces que aparece un término de polarización que contiene una frecuencia de 2ω , lo que significa que los electrones oscilarán, y por lo tanto irradiarán, una onda de frecuencia doble de la de la onda incidente (figura 8). Como la frecuencia está relacionada en forma inversa con la longitud de onda (λ) según la expresión:

$$\omega = 2 \cdot \pi \cdot c / \lambda$$

una duplicación de frecuencia significa una reducción de la longitud de onda a la mitad. Así, si nos manejamos con este último parámetro, y para tener una idea intuitiva de lo que estamos diciendo, si se hace incidir sobre un cristal apropiado un haz intenso de por ejemplo 800 nm (extremo rojo del espectro visible), tendremos a la salida, además del haz incidente, un haz de 400 nm (extremo violeta del espectro visible). Más adelante veremos casos concretos de duplicación de frecuencia en láseres usados comúnmente.

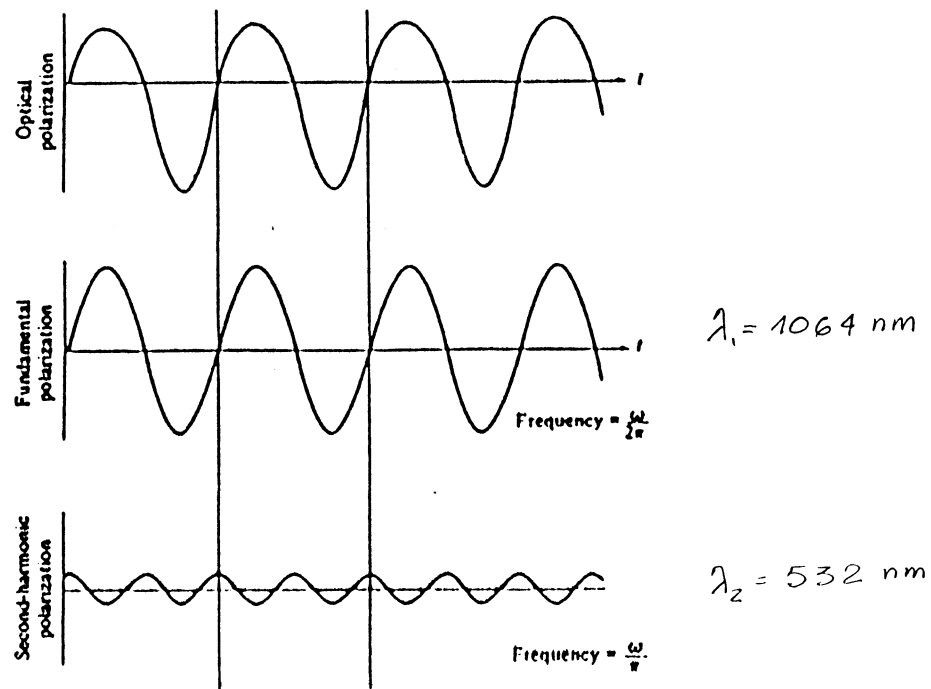


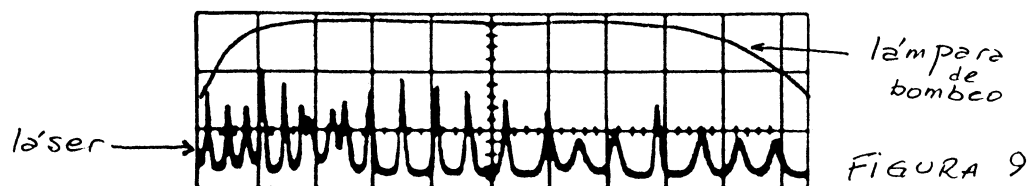
FIGURA 8

Debe destacarse finalmente dos requisitos esenciales para la generación de segunda armónica. En primer lugar, es posible demostrar que, por razones de simetría, un cristal que posea centro de simetría tendrá todos los coeficientes de susceptibilidad χ_i de orden par nulos, por lo que no serán capaces de generar este efecto. Así, solo los cristales no-centrosimétricos serán potencialmente útiles como generadores de segunda armónica. En segundo lugar, se debe considerar el hecho de que dos ondas de frecuencia distinta que viajan dentro de un material tienen velocidad de propagación diferente, ya que en general el índice de refracción depende de la frecuencia de la onda. Así, para que la generación de segunda armónica sea un fenómeno coherente, es necesario que ambas ondas se propaguen con igual velocidad dentro del material. Esto se consigue utilizando cristales birrefringentes, en los que la luz viaja a diferente velocidad según la dirección de propagación. En éstos, es posible encontrar una dirección particular para la cual la onda de frecuencia ω como la de 2ω se propagan en fase; a este ajuste de velocidades se lo conoce como "phase matching" y resulta esencial para obtener el doblado de frecuencia con una eficiencia importante. Cabe destacar también que es posible triplicar y cuadruplicar la frecuencia de emisión de un láser utilizando estos efectos no lineales. Estos procesos sirven en la práctica para ampliar el espectro de longitudes de onda que se pueden obtener de determinados tipos de láser.

II. 2 Conmutación del Q de la cavidad (Q-switch).

Esta técnica permite la generación de pulsos láser de corta duración (aproximadamente entre 5 y 50 ns) y altas potencias pico (del orden de las decenas de MW). El mecanismo en el que se basa es el siguiente: supongamos que se introduce un obturador dentro de la cavidad láser. Si el material activo está siendo bombeado mientras el obturador está cerrado, no se podrá generar acción láser, ya que éste impide que los pocos fotones espontáneos presentes en la cavidad comiencen el proceso de emisión estimulada. Sin embargo, la inversión de población alcanza valores muy por encima del valor umbral al que hubiera llegado de no haber estado presente el obturador. Si ahora este se abre bruscamente (típicamente en un tiempo corto comparado con la vida media del nivel superior), el medio tendrá una ganancia que excede en mucho a las pérdidas (ganancia muy por encima del umbral) y toda la energía almacenada en el nivel superior es liberada rápidamente en forma de un pulso de corta duración, lo cual hace aumentar la potencia liberada en varios órdenes de magnitud. Como esta técnica involucra la conmutación rápida del factor de calidad de la cavidad láser se la conoce como “Q-switching”.

Es claro que uno de los primeros requerimientos de esta técnica es que el nivel láser superior sea capaz de mantener su población durante un tiempo comparable al de excitación, lo que significa que debería tener una vida media relativamente larga (centenas de microsegundos a algunos milisegundos). Este es el caso de transiciones dipolares eléctricas prohibidas, que son comunes en muchos láseres de estado sólido cristalino como el rubí, neodimio, alexandrita, etc, que veremos más adelante. En estos casos, la salida temporal láser sin Q-switch esta formada por una serie de pulsos de amplitud aleatoria e irregularmente espaciados en tiempo, como se muestra en la figura 9. Aquí el pulso de bombeo de la lámpara (trazo superior) invierte la población entre niveles y da lugar al primer pulsito de emisión estimulada, ya que la cavidad está “abierta”. Sin embargo, como la lámpara aun sigue bombeando, vuelve a invertir la población entre los niveles láser y se vuelve a generar otro pulsito láser, en una secuencia que termina al llegar a su fin el pulso de excitación de la lámpara.



Otro requerimiento es que la apertura del obturador sea rápida comparada con el tiempo de generación del pulso láser dentro de la cavidad, de tal forma que los primeros fotones espontáneos encuentren un medio con una muy alta ganancia y se produzca un proceso de avalancha que genere un pulso estimulado de corta duración (figura 10). Como caso típico, citaremos el ejemplo de un láser Q-switch de Nd:YAG (sólido) cuyo nivel láser superior tiene una vida media de aproximadamente 250 μ s es bombeado por un pulso de lámpara de unos 150-200 μ s y con una apertura de obturador de unos pocos nanosegundos se obtienen pulsos de entre 20 y 50 ns de duración y potencias del orden de los MW.

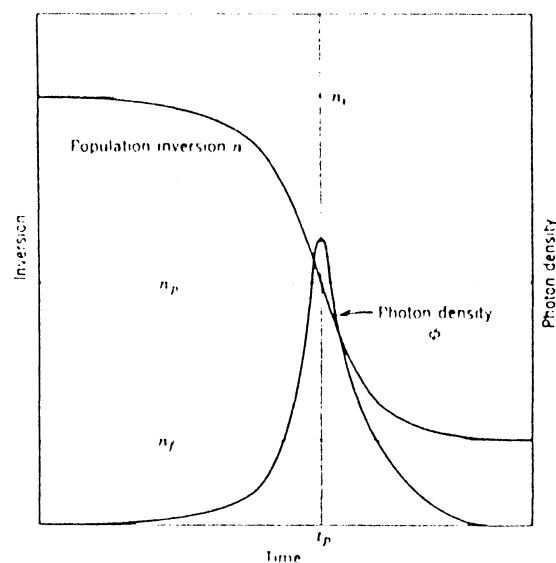


FIGURA 10

Desde luego que la palabra “obturador” se refiere aquí a un mecanismo de abre-cierre rápido más que a un sistema mecánico como el de una cámara de fotografía. La mayoría de los láseres Q-switch modernos utilizan la llamada celda de Pockels, que hace uso del efecto electroóptico consistente en inducir, sobre un cristal no lineal adecuado, un cambio en su birrefringencia proporcional a un determinado voltaje aplicado. Este cambio convierte a la celda de Pockels en un polarizador con dirección de polarización variable a voluntad según la tensión aplicada. Un esquema típico de este método puede verse en la figura 11, donde el polarizador intracavidad hace que la radiación esté polarizada en dirección vertical (por ejemplo). Si a la celda se le aplica

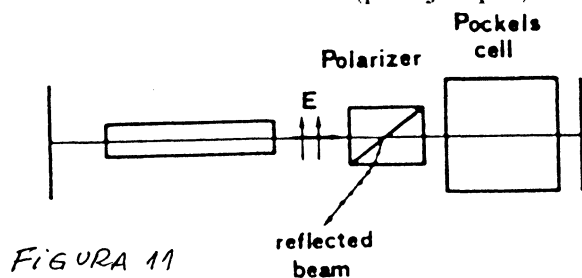


FIGURA 11

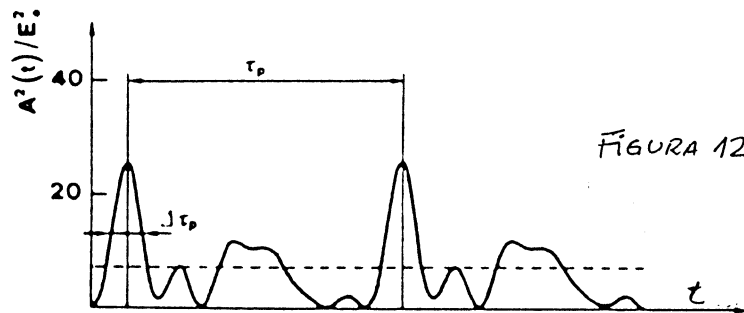
una tensión tal que su eje es perpendicular al del polarizador, la luz no pasará y el “obturador” estará cerrado, permitiendo la acumulación de población en el nivel superior. Una vez pasado el tiempo adecuado, si se conmuta la tensión de celda a un valor para el cual su dirección de polarización es paralela a la del polarizador, los fotones espontáneos que estén en la cavidad podrán llegar al espejo posterior (“obturador”abierto) y reinyectarse en el medio activo de alta ganancia, generando el pulso Q-switch.

Uno de los primeros sistemas de Q-switch consistía en reemplazar el espejo posterior por un prisma retroreflector rotatorio, que aumentaba el Q de la cavidad cada vez que se alineaba con el espejo de salida. Sin embargo, aun girando a 24000 rpm la duración de alto Q era de 400ns, un tiempo de conmutación suficientemente lento como para producir en muchos casos pulsos múltiples. Debido a esto, este sistema fue reemplazado por la conmutación electroóptica.

II. 3 Generación de pulsos ultracortos

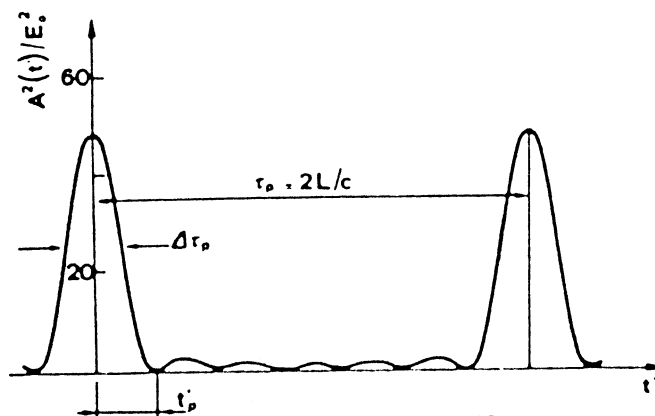
En muchas aplicaciones, resulta interesante contar con un láser capaz de emitir pulsos de muy corta duración con el fin de utilizarlos para estudios de cinéticas ultrarrápidas. Desde ya que existen láseres de tipo pulsado, pero en cualquier caso, éstos poseen una duración no inferior al nanosegundo. Sin embargo, existe una técnica basada en el acoplamiento de los modos de un láser que permite generar un tren de pulsos cada uno con un ancho temporal entre unas pocas decenas de picosegundo (10^{-12} seg) y unas pocas decenas de femtosegundos (10^{-15} seg).

Recordemos que los modos de un láser son formas de oscilación del campo electromagnético dentro de la cavidad resonante. Estos modos oscilan, en general, con fases aleatorias entre sí, de tal forma que la salida de un láser en función del tiempo es aproximadamente como se ve en la figura 12. Cada modo longitudinal puede



representarse como un vector rotante con velocidad angular ω_j y fase ϕ_j ; la amplitud del mismo en cada instante es la proyección de ese vector en el eje \underline{x} , y la salida del láser en cualquier instante de tiempo es la suma de esas proyecciones. Supongamos que ahora, de alguna manera, logramos que las fases ϕ_j de los modos se ajusten de tal forma que se cumpla: $\phi_j - \phi_{j-1} = \phi$ (constante). Como los modos tienen frecuencias distintas, separadas por una diferencia $\Delta\omega = c/2L$ (donde c es la velocidad de la luz en el medio y L la longitud de la cavidad), la frecuencia del modo j será:

$\omega_j = \omega_{j-1} + j \Delta\omega$ y es posible demostrar, aunque no lo haremos aquí, que el campo eléctrico instantáneo resultante surge de la interferencia de los modos individuales y posee máximos notables a intervalos de tiempo regulares. Esto puede entenderse a partir de la representación anterior, observando que existirán instantes de tiempo para los cuales los vectores rotantes coincidirán en la misma posición y por lo tanto sumarán en fase sus amplitudes para dar una alta intensidad (proporcional al cuadrado de la amplitud total), según puede verse en la figura 13. En cambio, para un instante



posterior, los modos habrán dejado de coincidir y se habrá destruido la interferencia constructiva entre los mismos, haciendo que la amplitud total (y consecuentemente la intensidad) caiga prácticamente a cero. Cuanto mayor sea el número de modos involucrados, más estricta será la condición temporal de interferencia constructiva y el efecto de refuerzo será más notable. Así, es de esperar que el ancho temporal de los pulsos sea inversamente proporcional al número de modos, o sea al ancho de banda total de oscilación. En efecto, se demuestra que $\Delta\tau_p = 1/\Delta\nu_{osc}$. Cuando un láser trabaja bajo este régimen, se dice que está enganchado en fase o "mode-locked", en cuyo caso, la salida consiste en un tren de pulsos de muy corta duración, separados por un cierto intervalo de tiempo. Desde un punto de vista formal, este hecho se ve claramente ya que el comportamiento temporal del pulso es justamente la transformada de Fourier de su espectro de frecuencias. Como el ancho de banda de oscilación es del orden del ancho de ganancia del medio ($\Delta\nu_0$), es posible esperar pulsos cortos (del orden de unos pocos picosegundos, 10^{-12} seg.) de láseres de estado sólido en mode-locking. En cambio, para láseres de colorantes líquidos que tienen un ancho de banda de ganancia alrededor de 100 veces mayor que los láseres de estado sólido, es posible obtener pulsos mucho más cortos (del orden de algunas decenas de femtosegundos 10^{-15} seg.). En láseres gaseosos, el ancho de banda es mucho más angosto y solo se pueden generar con esta técnica pulsos más bien "largos" de alrededor de 400 ps.

Los métodos prácticos para conseguir mode-locking pueden dividirse en activos y pasivos. Los primeros consisten en introducir dentro de la cavidad un dispositivo que modula las pérdidas (obturador) a una frecuencia igual a la inversa del tiempo de tránsito de la cavidad. En los segundos, se reemplaza el modulador por un absorbente saturable que produce efectos similares sin necesidad de acción externa (figura 14).

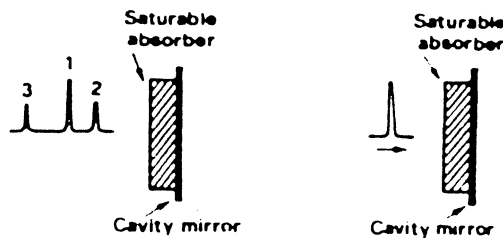
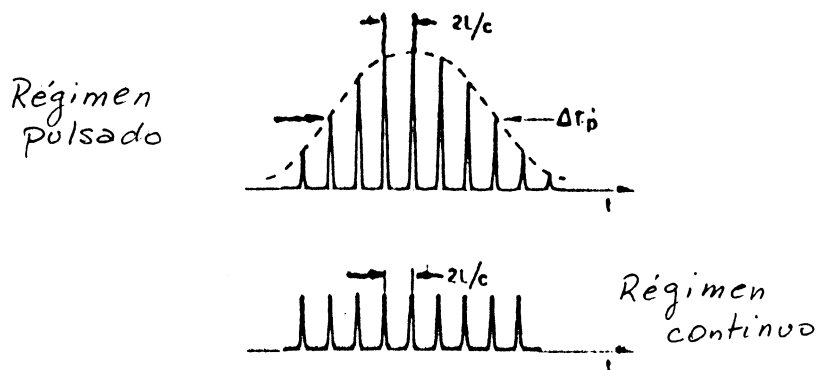


FIGURA 14



III. Láseres de estado sólido

Se suele llamar láseres de estado sólido a aquellos que tienen como medio activo un cristal ya sea puro o dopado con determinado elemento. Usualmente, éste resulta ser un ion perteneciente a la serie de elementos de transición de la tabla periódica (por ejemplo Cr) o perteneciente a la llamadas tierras raras (por ejemplo Nd o Er). En estos casos, son estas impurezas las que forman el medio activo propiamente dicho, siendo el cristal el huésped para estos elementos. Existe una clase especial de láseres de estado sólido en los que el medio activo es un cristal iónico puro (en general un haluro alcalino) y se trabaja sobre las imperfecciones naturales o inducidas que pueda presentar el medio. Estas son defectos en la red cristalina tal como vacancias de iones, exceso de electrones, etc. Estos láseres, que reciben el nombre de centros de color (por el aspecto de opacidad que presentan los defectos a la luz natural), fueron muy estudiados en la década del 70 por sus características de emisión en el IR cercano y por sus propiedades de sintonía de la misma; sin embargo, la necesidad de mantenerlos refrigerados a temperatura de nitrógeno líquido (-190° C) para su correcto funcionamiento los hizo poco prácticos (aunque una importante compañía ha comercializados ciertos modelos) y lentamente han dejado de tener interés práctico.

Resultaría muy extenso tratar en esta sección los distintos tipos de láseres de estado sólido que han demostrado un funcionamiento exitoso durante la últimas tres décadas. Vamos a concentrarnos solo en aquellos que por su alta versatilidad, utilidad o tradición han quedado como los más representativos de esta clase de láseres. Ellos son el láser de rubí, el de neodimio y el de titanio-zafiro.

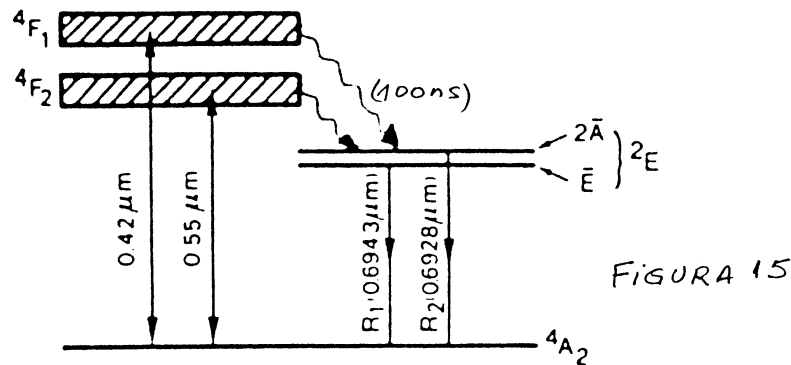
Los láseres de semiconductor son, sin lugar a dudas, láseres de estado sólido, aunque no del tipo cristalino definido más arriba. Debido a que los mecanismos de excitación de los semiconductores son diferentes de los de estado cristalino y a la enorme importancia que han desarrollado los primeros, los trataremos dentro de esta misma sección en forma más extensa.

III. 1 Láser de Rubí

El láser de rubí fue el primer láser en funcionar con éxito. Su demostración fue llevada a cabo el 16 de mayo de 1960 por T. Maiman en los laboratorios de la compañía Hughes. Hoy, a 37 años de ese evento, se podría pensar que este láser es un fósil viviente del mundo de los láseres; sin embargo, con algunas modificaciones menores respecto de su diseño original, este láser ha encontrado áreas específicas de aplicación: sus pulsos de alta energía en el rojo del espectro visible se utilizan para el tratamiento de lesiones pigmentarias en la piel y borrado de tatuajes, para ensayos no destructivos por interferometría holográfica y para diagnóstico de plasmas en estudios de combustión. Aunque este láser fue superado y reemplazado por el láser de neodimio como láser multipropósito, en ciertas aplicaciones donde se necesita alta fluencia (energía por unidad de área), se prefiere la mayor capacidad de su material activo para almacenar energía de bombeo durante más tiempo que el neodimio.

El medio activo en el láser de rubí consiste en una barra cilíndrica de zafiro sintético (Al_2O_3) dopada con cromo en un porcentaje que oscila entre el 0,01% y el 0,5%. Este nivel de dopaje corresponde a una concentración de alrededor de 10^{19} iones de cromo por cm^3 . Este reemplaza al ion de aluminio en la red de zafiro con la misma valencia (Cr^{3+}), dando al cristal su típica coloración rojiza. El material es fácil de crecer, fabricándose barras de 3 a 25 mm y de hasta 20 cm de largo, posee un alta umbral de daño óptico y conduce el calor generado por el bombeo en forma más eficiente que el vidrio.

La estructura de niveles de este material puede verse en la figura 15. El Cr^{3+} posee dos bandas de absorción centradas alrededor de 400nm y 550nm, que decaen rápidamente (100ns) a un par de niveles metaestables de 3 ms de vida media.



A partir de allí se produce la emisión láser hacia el estado fundamental en una transición en 694.3nm. La ubicación de las bandas de absorción imponen la necesidad de bombeo con lámparas de flash pulsadas, tanto de xenon como de mercurio. Como puede verse de la figura anterior, este láser funciona como sistema de tres niveles, pues el estado fundamental coincide con el nivel láser inferior, lo que implica que se necesitan altas potencias de bombeo para llevar más de la mitad de los átomos desde el nivel inferior al superior y generar así la inversión de población requerida para la emisión estimulada. Esto aumenta el umbral y baja la eficiencia, lo que constituye una de las desventajas de este láser.

En cuanto a su estructura interna, la cavidad puede estar formada simplemente por depósitos reflectores en las caras de la barra, aunque los láseres modernos incorporan espejos externos parcialmente reflectores separados de la barra para compensar efectos de lente térmica. Las lámparas de bombeo son de tipo lineal, ubicadas en cavidades elípticas (ver figura 2) y pueden trabajar en configuración oscilador-amplificador que mejora la calidad del frente de onda, lo que resulta esencial en aplicaciones de holografía interferométrica. Operando como oscilador, puede producir pulsos de entre 50 y 100 J de energía con una duración de algunos milisegundos. La eficiencia de conversión de potencia de salida sobre potencia de entrada es de entre 0,1 a 1 %. Como el nivel superior tiene una vida media del orden de los milisegundos, este láser puede trabajar en modo Q-switch, que al comprimir los pulsos a una duración de entre 10 y 30 ns, aumenta la potencia pico al rango de los 100 MW. Si además opera como oscilador-amplificador, la potencia pico llega a un GW, con una energía por pulso de 10 J. El láser de rubí también puede trabajar en régimen mode-locking, generando trenes de 20 a 30 pulsos de 3 ps de duración cada uno, con energías de 1 mJ.

La aplicación típica de este láser ha sido el registro de hologramas de gran volumen debido a la alta energía por pulso que es capaz de generar en una región del espectro donde las placas holográficas de registro son sensibles (690 nm). Los láseres de rubí Q-switch pueden producir uno, dos o tres pulsos coherentes de 10 a 30 ns de duración por disparo de lámpara de flash. El sistema de doble pulso suele usarse para “congelar” el movimiento de un cuerpo en dos instantes cercanos, de modo que si se superponen ambos en una misma placa es posible observar microdeformaciones por interferometría holográfica.

Este láser ha encontrado una reciente aplicación en medicina para el tratamiento de lesiones de piel (manchas de origen hematológico o por exceso de melanina) y borrado de tatuajes. Según puede verse en la figura 16, tanto la melanina como los

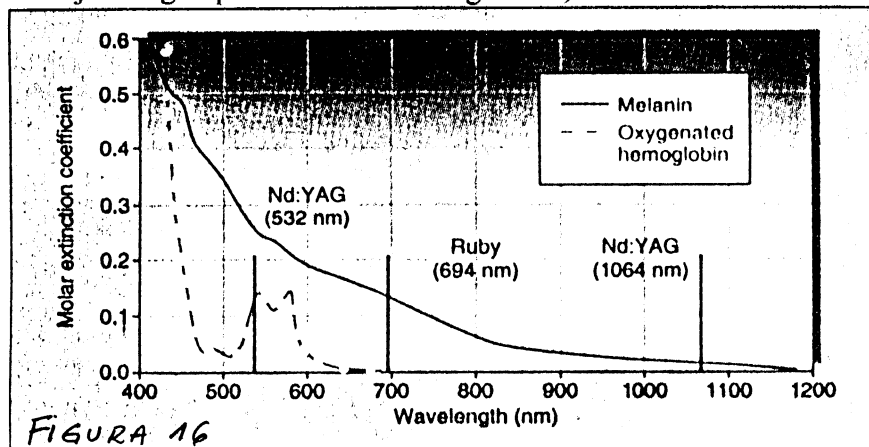


FIGURA 16

pigmentos usados en tatuajes tienen una alta absorción en la longitud de onda del rubí, de tal forma que al depositar sobre estos elementos una alta energía en un tiempo corto es posible fragmentarlos y destruir las células que los contienen, dejando intactas a las vecinas no pigmentadas y permitiendo que el organismo remueva los detritos y los reemplace por tejido nuevo.

III. 2. Láser de Neodimio

El láser de neodimio es tal vez el láser más popular e importante dentro de los del tipo de estado sólido. El medio activo consiste generalmente en un cristal de $Y_3Al_5O_{12}$ (comúnmente llamado YAG por su acrónimo en inglés de yttrium aluminium garnet), en el que algunos de los iones de Y^{3+} son reemplazados por iones de Nd^{3+} . Los niveles típicos de dopaje son del orden del 1%, pues valores superiores tienden a inhibir la fluorescencia de los niveles láser además de provocar una tensión mecánica en el cristal, ya que el radio del ion Nd^{3+} es 14% más grande que el del Y^{3+} . Esta impurificación hace que el cristal adopte una coloración púrpura debido a la absorción del neodimio en el rojo. También es posible utilizar como huésped vidrios en los que el nivel de dopaje es mayor que en el caso del YAG (alrededor del 3%).

Un esquema simplificado de niveles de energía del Nd:YAG se muestra en la figura 17. Los mismos corresponden a los electrones de la capa $4f$ del Nd dentro del campo cristalino. Posee dos bandas de absorción centradas en 730nm y 800nm que poseen un decaimiento no radiativo rápido (100 ns) al nivel $^4F_{3/2}$, que tiene una vida

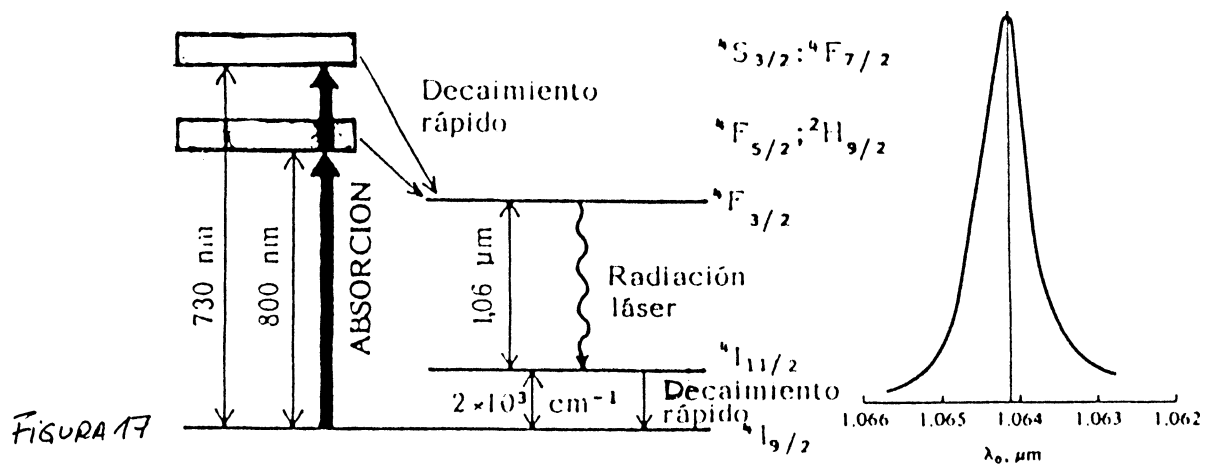


FIGURA 17

media larga (aprox. 0,23 ms), lo que le permite acumular buena parte de la energía de bombeo como nivel láser superior. Desde este nivel, la transición más fuerte es hacia el nivel $^4I_{11/2}$ el que, al estar acoplado al nivel fundamental $^4I_{9/2}$ por un decaimiento no radiativo muy rápido, se puede considerar que aquel permanece vacío, favoreciendo así la inversión de población entre ambos. Además, la sección eficaz de absorción para la transición láser es aproximadamente 75 veces mayor que en el láser de rubí, lo que implica que, para una dada inversión, la ganancia del Nd es casi dos órdenes de magnitud mayor, lo cual explica el bajo umbral de oscilación de este láser y su facilidad para operar en régimen continuo. Si bien la transición en 1.064 nm es la más fuerte, existen otras transiciones a niveles del multiplete inferior que corresponden a otras longitudes de onda láser, como por ejemplo 950 nm, 1100 nm y 1319 nm, que se obtienen por sintonía o por cambio de huésped. Volviendo a la transición de 1064 nm, está homogéneamente ensanchada a temperatura ambiente ($\Delta\nu=200\text{GHz}$) debido a interacciones con fonones de la red, lo que favorece la operación tipo mode-locking de este láser. Debido también a la larga vida media del nivel superior puede funcionar en modo Q-switch. De hecho, en el primero de los modos de operación puede entregar pulsos de 20 ps, mientras que en el segundo la potencia pico de salida puede llegar a los 50 MW. Además puede operar en régimen continuo (con potencia de salida de 200 W multimodo) o pulsado (potencia promedio de 500 W para una frecuencia de repetición de 50 Hz). En ambos casos la excitación es de tipo óptica a través de lámparas de flash lineales dispuestas en cercanías o en simple o múltiple eclipse.

La alta potencia que se obtiene de este láser permite utilizar las técnicas de generación de armónicas para extender el rango de longitudes de onda de salida por doblado, triplicado y cuadruplicado de frecuencia. Así, es posible tener en el laboratorio un láser de Nd:YAG Q-switch que emita en su frecuencia fundamental 1064 nm y en sus armónicas 532 nm, 355 nm y 266 nm respectivamente, con energías por pulso entre 650 mJ y 100 mJ.

El hecho de que el láser de Nd:YAG tenga sus bandas de absorción en la región de emisión de los láseres de semiconductor (que veremos más adelante) y el gran desarrollo de estos últimos, ha hecho que surgiera, en la segunda mitad de la década de '80, una nueva generación de Nd:YAG, que son los bombeados por diodos. Este tipo de láser presenta una mayor eficiencia de bombeo pues recibe radiación solo en la región espectral de absorción, disminuyendo además la carga térmica sobre el medio activo. Estos diseños tienden a la miniaturización de los láseres de Nd (de hecho hay compañías que ofrecen láseres que caben en la palma de la mano) que por su alta potencia y versatilidad se utilizan en dispositivos para procesamiento de materiales, aplicaciones científicas y militares.

III. 3. Láser de Titanio:Zafiro

El primer láser de Ti:zafiro funcionó en 1982 y su versión comercial de operación continua apareció en 1988. Este láser utiliza como material activo un cristal de zafiro (Al_2O_3), similar al láser de rubí, pero dopado con alrededor del 0,1% de Ti, que reemplaza al ión de aluminio en el cristal como Ti^{3+} . Este elemento pertenece a la misma serie de elementos de transición de la tabla periódica como el Cr y el Co, también dopantes en otras matrices. El cristal de Ti:zafiro posee buenas propiedades ópticas, es resistente y puede crecerse con facilidad (barras de 1 a 30 cm de largo).

Un esquema simplificado de niveles puede verse en la figura 18. La transición láser toma lugar entre el estado excitado 2E y el estado fundamental 2T_2 . El bombeo óptico lleva al ión Ti^{3+} al nivel electrónico superior, que relaja luego al nivel más bajo de la banda vibrónica, desde donde realiza la transición láser. Esta deja al ión en un subnivel vibracional excitado del estado fundamental, desde donde relaja al nivel vibracional inferior del mismo. La transición láser muestra un gran ancho de banda debido a una fuerte interacción entre el átomo de titanio y la red cristalina.

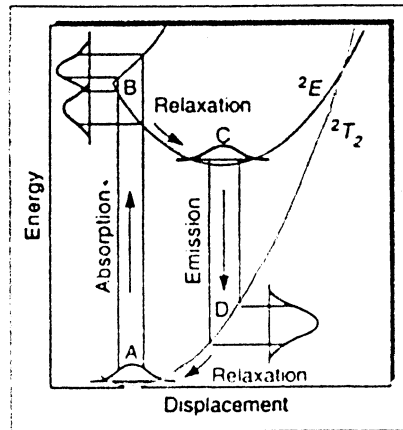


FIGURA 18

Las bandas de absorción y de emisión de este láser se encuentran levemente solapadas. La primera tiene su máximo en los 500 nm, por lo que se favorece su excitación con láser de argón (514 nm) o láser de Nd doblado en frecuencia (532 nm). Este láser no puede bombearse directamente con láseres de semiconductor, aunque se ha demostrado el funcionamiento de un dispositivo en el que un láser de Nd excitado por diodos y doblado en frecuencia bombea a un Ti:zafiro, en una configuración de todo-estado-sólido. Por otro lado, la banda de fluorescencia se extiende entre 700 nm y 900 nm con máximo en 780 nm (figura 19). El láser es fácilmente sintonizable en este rango, con ancho de línea de hasta 1 kHz en configuración de anillo (ring cavity).

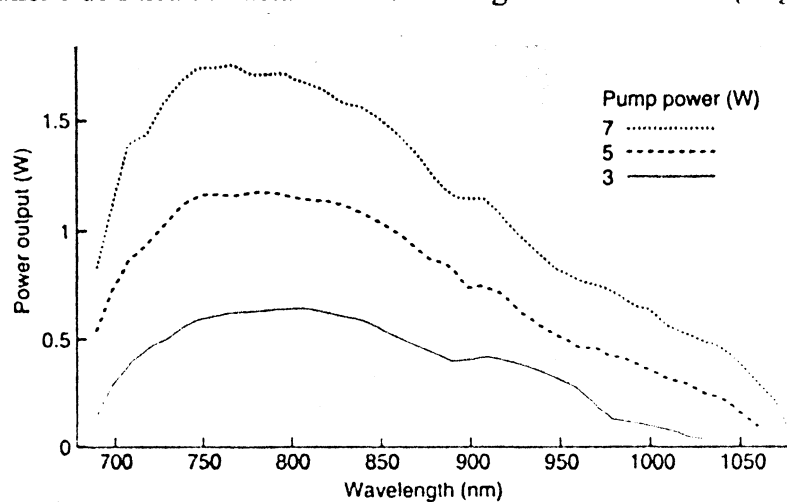


FIGURA 19

El gran ancho de banda de ganancia del Ti:zafiro se utiliza para generar en forma directa pulsos ultracortos por técnicas de mode-locking. Así, se han logrado pulsos de 17 fs de ancho temporal, superando el record logrado con láseres de colorantes en similares condiciones. Los láseres de Ti:zafiro comerciales pueden operar en régimen continuo con potencias de 1 a 2 W, aunque dispositivos de laboratorio han llegado a más de 5W. Estas intensidades se utilizan para generar radiación coherente en 350-470 nm por segunda armónica, en 235-300 nm por tercera y cerca de 210 nm por generación de cuarta armónica. La principal aplicación del Ti:zafiro es en la investigación básica y aplicada y en el reemplazo de los láseres de colorantes en el cercano IR, donde éstos se tornan químicamente inestables e ineficientes.

IV. Láseres de estado líquido

Dentro de esta categoría entran los láseres de soluciones líquidas de colorantes orgánicos, llamados comúnmente láseres de colorante. Aquí, el material activo son moléculas orgánicas de un gran número de átomos, cuya característica más importante es la aparición de ligaduras dobles y simples alternadas entre átomos de carbono (ligadura no saturada). Estos compuestos absorben luz en forma muy eficiente en regiones espectrales que van desde el UV al IR cercano (220 a 1500 nm) según su composición. La energía absorbida en un cierto rango es luego reemitida en longitudes de onda mayores, dando lugar a la intensa coloración que los caracteriza. Un esquema de niveles puede verse en la figura 20.

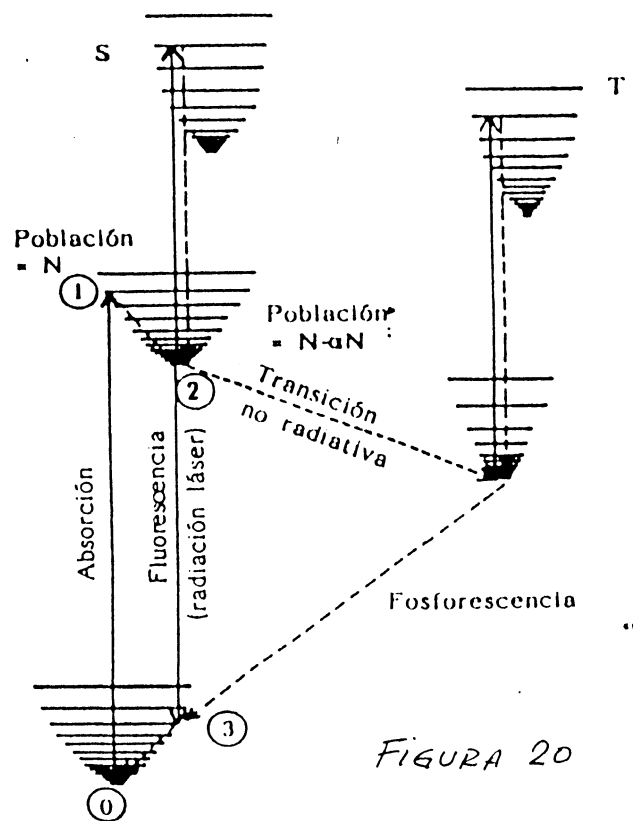


FIGURA 20

Se presentan dos sistemas de niveles electrónicos, singuletes y tripletes, subdivididos a su vez en un número muy grande de niveles vibro-rotacionales. La transición de singulete excitado al fundamental es de una alta probabilidad y da lugar a la fluorescencia, por la cual se obtiene acción láser. En cambio, la transición entre el triplete excitado y el fundamental posee baja probabilidad debido a las reglas de transición que se deducen de la mecánica cuántica (transición prohibida), por lo que la vida media de este nivel es larga (del orden de segundos). De la figura anterior puede verse que este láser es de cuatro niveles y que la radiación emitida es de menor energía que la absorbida. Esto puede observarse claramente en la figura 21, donde se muestran los espectros de absorción y de fluorescencia para el caso de la Rodamina 6G, colorante este de mayor eficiencia y por lo tanto el más estudiado.

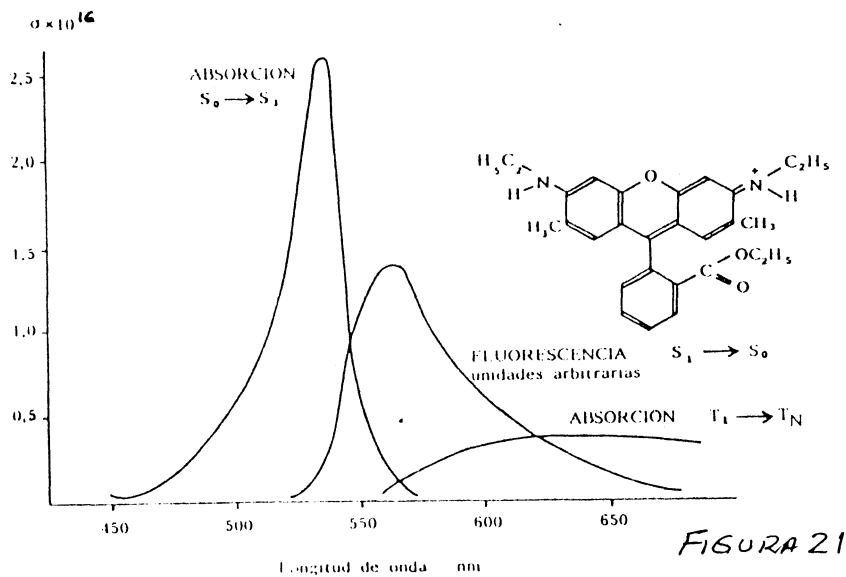
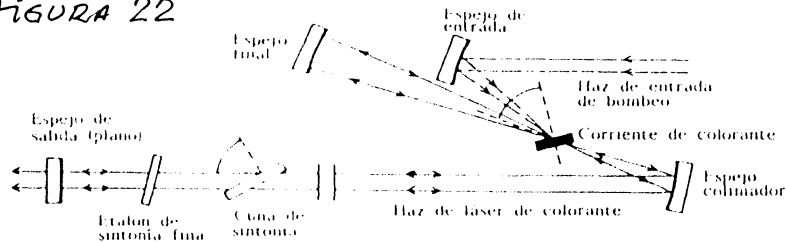


FIGURA 21

Por la ubicación de la banda de absorción, el bombeo puede llevarse a cabo con lámparas de flash o láser de argón (488-514 nm). Aunque no se muestra en la figura, esta banda tiene un pequeño máximo en la región del UV ($\sim 300-350$ nm), que es aprovechado para el bombeo con láseres pulsados de emisión en esa región como el de nitrógeno (337 nm) y el de excímeros (308-340 nm). Por otro lado, la radiación emitida abarca una banda ancha que va entre los 530 nm y los 650 nm. Cuando se llega al umbral y se produce la emisión estimulada, se produce un estrechamiento del ancho de banda a unos 100 nm. La característica más importante de este láser es su propiedad de sintonía, que permite obtener una salida de reducido ancho de banda (mucho más angosta que su emisión "libre"), sin pérdida apreciable de energía. Esto se debe a que el ensanchamiento de los niveles de energía es de tipo homogéneo, es decir, por colisiones con las moléculas de solvente, de forma que si reinyectamos dentro del medio activo una banda angosta de frecuencias, con una red de difracción por ejemplo, todas las moléculas contribuirán a amplificar dicha banda, de forma que casi toda la energía de la banda ancha se canalizará en la banda estrecha seleccionada.

Los colorantes pueden operar en régimen pulsado o continuo, según sea la fuente de bombeo. En el primer caso, la sintonía se logra colocando un prisma o una red de difracción en reemplazo del espejo de reflexión total. En el segundo caso, el diseño más común es el de cavidad de tres espejos que se muestra en la figura 22.

FIGURA 22



El colorante fluye perpendicular al eje de la cavidad en un determinado punto de la misma, libre de celdas a través de una tobera. El haz de bombeo es enfocado en un punto del "jet", que será la región de ganancia. La cavidad está diseñada para compensar pérdidas por astigmatismo. Con láseres de este tipo pueden obtenerse varios watts de potencia sintonizable en 50-90 nm con anchos de línea de 0,02 nm.

Debido al gran ancho de banda natural que tienen estos láseres, es posible obtener pulsos extremadamente cortos (de hasta 6 fs) con técnicas de mode-locking y de compresión de pulsos. Sin embargo, no operan en modo Q-switch debido al corto tiempo de vida del nivel superior. Como dijimos más arriba, la utilización de distintos colorantes permite abarcar (lease sintonizar) todo el espectro visible (figura 23), lo que hace de estos láseres importantes herramientas de aplicación en áreas como la investigación básica, sensado remoto de sustancias, detección de contaminantes, LIDAR, fluorescencia de muestras biológicas, etc.

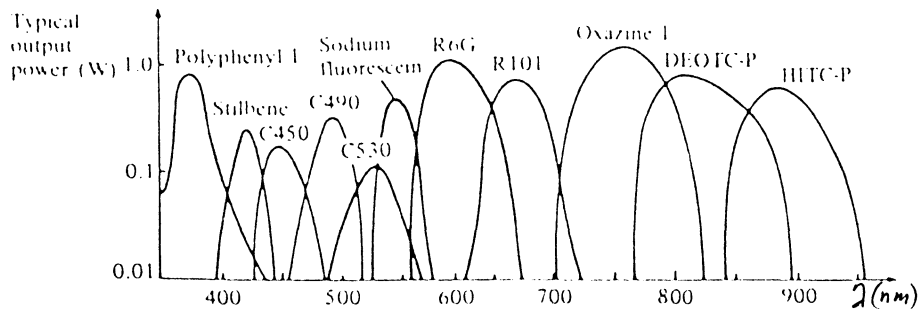
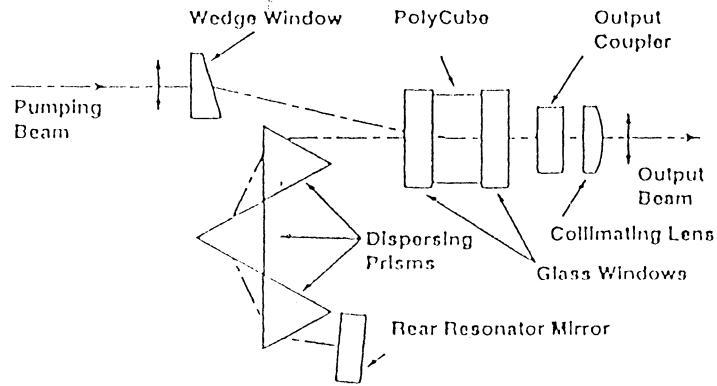


FIGURA 23

Es de destacar otra modalidad de estos láseres: colorantes en matrices poliméricas. Si bien la emisión estimulada de polímeros dopados con colorantes orgánicos era conocida desde hace más de veinte años, solo recientemente ha tomado carácter comercial, principalmente debido al avance producido en la fabricación de estos materiales amorfos. El medio activo consiste en una pequeña barra de polímero dopado con un determinado colorante que se introduce en uno de los focos de una cavidad elíptica excitada por lámpara de flash. La compañía Coherent comercializa un sistema bombeado por un Nd:YAG doblado en frecuencia y que posee tres prismas dispersores intracavidad para sintonizar la salida girando adecuadamente los mismos. Cambiando la barra de polímero dopada con otro colorante se puede obtener emisión sintonizable entre 550 nm y 750 nm (figura 24). Sin embargo, estos materiales degradan relativamente rápido, por lo cual es necesario reemplazar la barra con cierta frecuencia. Es claro que esta tecnología no está aun madura como para competir con los láseres de colorantes en solución.

PolyChrom Optical Layout



PolyChrom Tuning Curves

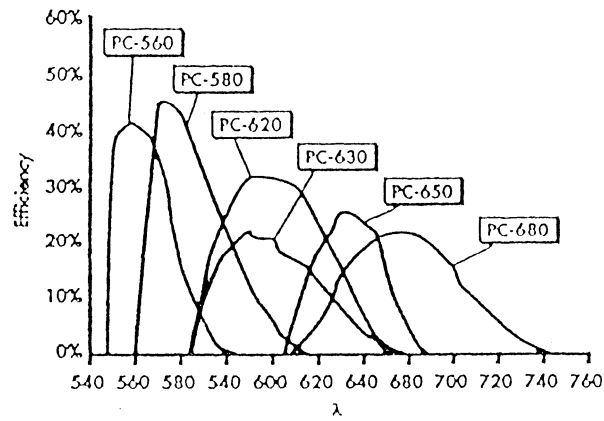


FIGURA 24

V. Láseres gaseosos

Dentro de esta clasificación se encuentran todos aquellos láseres que tienen como medio activo un gas, o mezcla de gases o vapor a determinada presión. Es así que en este grupo se encuentran comprendidos varios láseres importantes como los de gases nobles, de nitrógeno, de dióxido de carbono, de vapores metálicos (oro, cobre, cadmio), etc. En todos los casos, la excitación se lleva a cabo por medio de descargas eléctricas de determinadas características y aun de bombeo por cañón de electrones. En esta sección solo vamos a describir los tres láseres que, por sus características, suelen ser los más utilizados además de cubrir los rangos bajo, medio y alto de potencia. Estos son: el láser de He-Ne, el láser de Ar y el láser de excímeros. Dejamos de lado al láser de CO₂ pues se trata extensamente en otra clase del curso.

V. 1. Láser de He-Ne

Este fue el primer láser gaseoso en funcionar en forma continua en el año 1962. Desde entonces, ha tenido un desarrollo constante y se ha convertido en uno de los más populares. El medio activo es una mezcla de helio y neon en una proporción 10:1, donde este último es el que genera la acción láser. La mezcla, a una presión de un Torr, se confina en un tubo capilar de 1,5 mm de diámetro por el que se hace circular una corriente continua de entre 5 y 50 mA. Con referencia al esquema de niveles que muestra la figura 25, dicha descarga excita al He hasta dos niveles metaestables (larga vida media) que se encuentran prácticamente a la misma "altura energética" que los niveles 2s y 3s del Ne. Esta resonancia permite poblar selectivamente estos niveles a

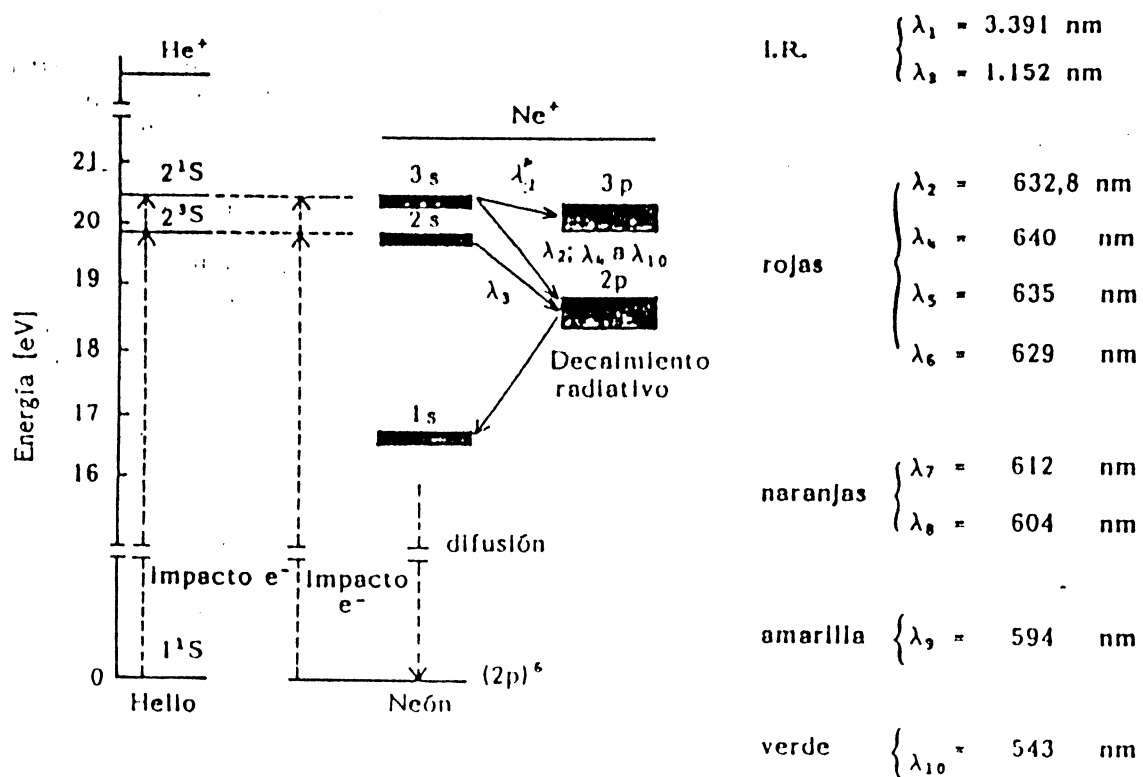
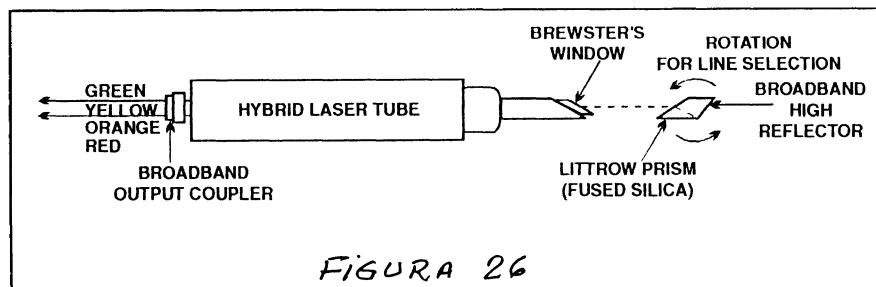


FIGURA 25

través de colisiones, de forma que se genera inversión de población entre éstos y los niveles de la configuración p . Así, este láser mejora su funcionamiento con el aumento de temperatura, pues aumenta el número de colisiones. De todas las longitudes de onda detalladas en la figura anterior, la más popular y la que más usos ha tenido es la de 632.8 nm. Debido al movimiento térmico de los átomos, esta transición tiene un ancho de unos 1700 Mhz. Si el tubo del láser se hace de unos 30 cm de longitud, solo tres modos longitudinales oscilarán simultáneamente. Si la longitud se lleva a 10 cm, es posible obtener emisión monomodo longitudinal, lo que aumenta la longitud de coherencia del haz por disminución del ancho de línea.

Como todas las transiciones mostradas (excepto la IR) poseen baja ganancia, los espejos que forman la cavidad resonante son de multicapas dieléctricas (hasta 23 capas) capaces de llegar a una reflectividad del 99,9% para una determinada longitud de onda y del 99% para el espejo de acople exterior. Las potencias típicas de esta línea van entre 3 y 5 mW. Actualmente se pueden adquirir dispositivos comerciales capaces de emitir láser en 13 diferentes líneas que cubren desde 611nm a 1520 nm y que pueden ser seleccionadas individualmente por medio de un prisma dispersor colocado intracavidad (figura 26), con potencias que van entre los 0,1 mW y 1 mW.

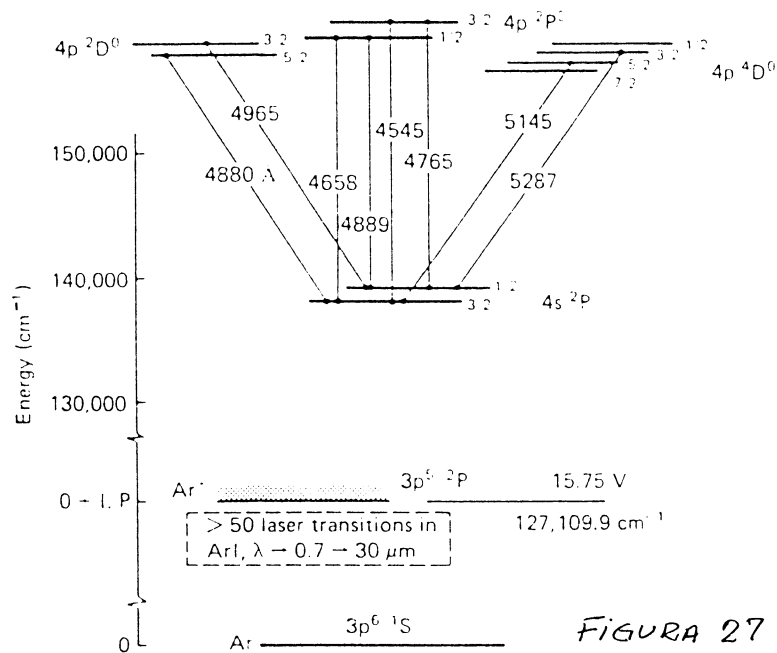


Este láser puede ser considerado una fuente barata de luz coherente de alta calidad. Tiene aplicaciones industriales y comerciales como alineación de componentes, interferometría, lectura de código de barras, "data-display", shows artísticos, etc.

V. 2. Láser de argón

El láser de argón es un representante de los llamados láseres iónicos, en los que el medio activo no es un gas neutro (como en el caso del He-Ne) sino que es una especie ionizada de cierto gas, como también es el caso del láser de kriptón o el de xenon. La operación es continua en el primero y segundo y pulsada en el tercero. En el caso del argón, la acción láser tiene lugar en varias transiciones que se encuentran mucho más arriba del nivel fundamental del primer ión como se ve en la figura 27. Las dos principales líneas de este láser se encuentran en 514,5 nm y 488 nm, que representan cerca del 80% de la potencia total del láser. La potencia total de todas las líneas varía entre menos de 1 W y 20 W, dependiendo de la corriente que circule por el canal de descarga.

La excitación de los láseres iónicos se realiza mediante una descarga eléctrica de alta corriente en un tubo conteniendo argón a una presión de alrededor de 1 Torr. La descarga se concentra en un capilar central, que es el lugar donde se produce la emisión estimulada. Inicialmente, un pulso de sobretensión de unos pocos miles de voltios



producen la ruptura dieléctrica del gas, entonces la tensión cae a valores entre 100 y 400 V y la corriente toma valores entre 10 y 70 A en una descarga autosustentada. Un campo magnético externo paralelo al tubo ayuda a confinar la descarga en el capilar. La colisión de los electrones con los átomos de argón hace que éstos se ionicen y lleguen al nivel fundamental del ión (ver figura anterior). Luego una segunda colisión electrónica bombea al ión hasta el nivel láser superior. Por esto, la potencia de este láser depende del cuadrado de la corriente, de modo que para aumentar la potencia es necesario aumentar mucho más la corriente. Esto impone severas condiciones al canal de descarga: resistencia a altas corrientes eléctricas, stress térmico y bombardeo de fotones energéticos de UV de vacío de 74 nm (ver figura anterior) que dañan severamente las paredes interiores y la óptica. Así, los láseres modernos poseen el tubo de descarga construido de cerámica terminado en ventanas de cuarzo de alta calidad orientadas en el ángulo de Brewster. Solo una pequeña fracción de la energía depositada en el plasma por la descarga eléctrica es utilizada en la excitación de los niveles láser correspondientes, siendo el resto canalizado en la excitación de otros niveles y en el aumento de la energía cinética de los átomos. Esta ineficiencia de bombeo, típica de las descargas eléctricas, hace que la llamada eficiencia de “wall-plug” (potencia de salida láser dividido potencia eléctrica de entrada) no supere el 0,1%. El exceso de calor generado por esta baja eficiencia debe ser removido rápidamente del tubo, por lo que es necesario un sistema de refrigeración por circulación de agua en el caso de los láseres de potencias mayores a 1 W o por aire forzado para dispositivos de baja potencia. El láser de argón tiene emisión en varias líneas (como se resume en la tabla al final de esta sección), por lo que los espejos de la cavidad son de alta reflectividad (multicapas de dieléctrico) en una región amplia del espectro que abarca todo el rango de emisión, con los que se obtienen las mayores potencias. Sin embargo, con la introducción de un prisma dispersor dentro de la cavidad es posible reinyectar en el medio activo una línea determinada, de forma que solo ella experimente ganancia y se tenga salida láser en esa longitud de onda. De esta forma se pueden sintonizar las distintas líneas de emisión, aunque a expensas de una menor potencia de salida. En la figura 28 puede verse un esquema simplificado de la estructura del láser de argón.

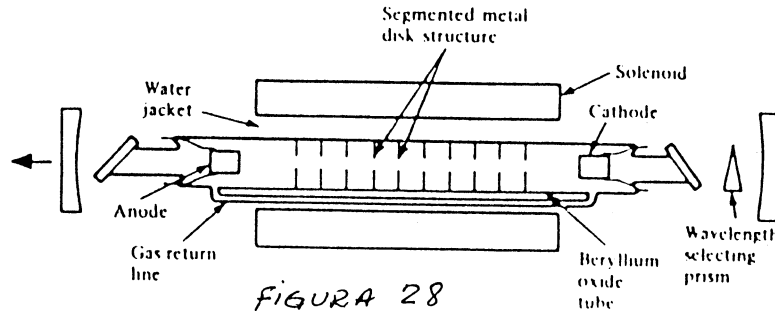
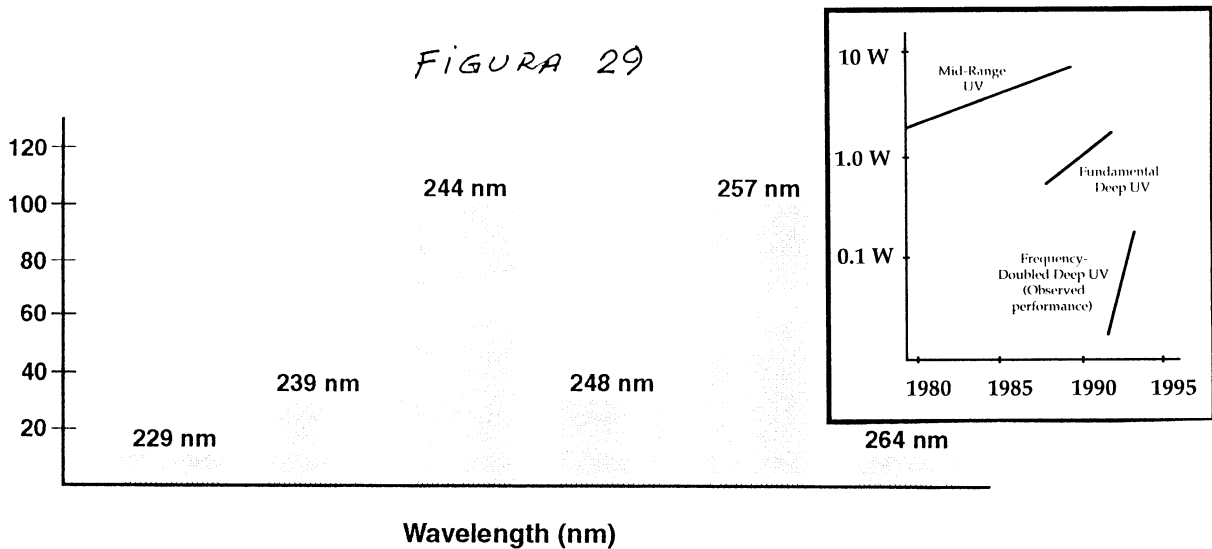


FIGURA 28

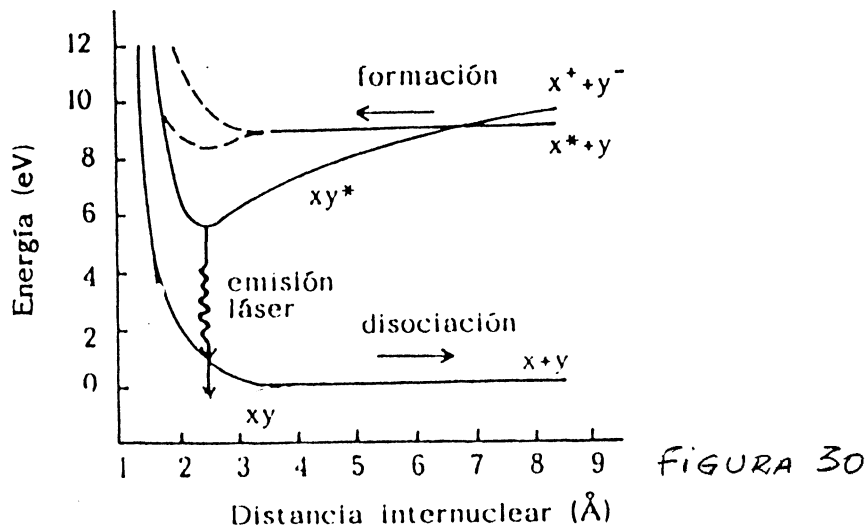
Estos láseres pueden trabajar tipo mode-locking, produciendo pulsos de unos pocos picosegundos de duración. Gracias al avance en la tecnología del tubo de plasma, ventanas de cuarzo resistentes y diseño del cátodo, es posible obtener emisión en el UV con láseres iónicos, ya sea en forma directa o por doblado de frecuencia por generación de segunda armónica intracavidad, obteniéndose potencias de 100 mW en 244 nm y 257 nm. La figura 29 resume la potencia de salida para diversas longitudes de onda y el incremento en la performance de estos láseres en el UV a lo largo de los años.



La vida media de estos láseres está especificada en el orden de las 10.000 hs., aunque en realidad sean operables por el doble de horas. La salida azul-verde del láser de argón tiene una alta interacción con la materia, ya que muchas sustancias absorben en ese rango. Las grasas que se encuentran en los dedos pueden hacerse fluorescer con este láser, por lo que se lo utiliza en análisis forense. La hemoglobina absorbe fuertemente en esta zona, por lo que este láser es muy utilizado en el tratamiento de retinopatías diabéticas y soldaduras de retinas. También se lo utiliza en aplicaciones de velocimetría Doppler y determinación de tamaño de partículas. Cuando el medio activo es una mezcla de argón y kriptón, el láser posee salida en las líneas combinadas de ambos, lo que se utiliza como efecto visual en shows artísticos y de entretenimiento. Sin embargo, la aplicación típica del láser de argón es como fuente de bombeo estándar para láseres de colorante continuos sintonizables y, más recientemente, como bombeo para el láser de Ti:zafiro.

V. 3 Láser de excímero

Los láseres de excímeros son una familia de láseres gaseosos que emiten pulsos UV de muy alta potencia con una duración de decenas de nanosagundos. El medio activo consiste en una mezcla de un gas raro y un halógeno que se combinan para dar una molécula de haluro de gas raro de corta vida media. La palabra excímero surge de la contracción (en inglés) de “*excited dimer*”, que es una molécula formada por dos átomos idénticos que existe solamente en estado excitado, tal como He_2 y Xe_2 . Ahora, esta palabra se usa para designar cualquier molécula diatómica en la que los átomos están ligados solo cuando se encuentran en estado excitado pero no en el estado fundamental. Esta propiedad los hace buenos candidatos láser pues automáticamente despueblan el nivel inferior de la transición. El esquema de niveles se ve en la figura 30



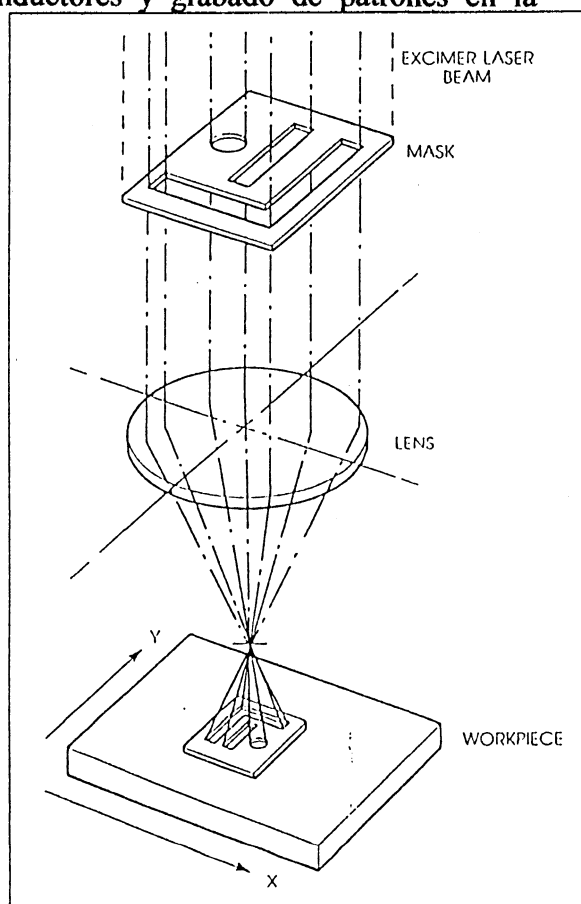
Cuando son excitados, los dos átomos se atraen y permanecen ligados dentro de los niveles vibracionales del estado superior; cuando efectúan la transición al nivel inferior se repelen y vuelven a quedar como átomos separados, despoblando automáticamente el nivel inferior de la transición láser. Este proceso contribuye a la alta ganancia de estos láseres. Los primeros excímeros fueron construidos a mediados de los '70 en laboratorios de la universidad, dentro de la avalancha de estudios de medios láser que había en la época. Las primeras versiones comerciales se conocieron poco tiempo después, en parte debido a que la forma de excitación era muy similar a la de los láseres de dióxido de carbono de campo transversal.

Los láseres de excímeros más importantes son los compuestos fluorados y clorados: ArF , KrF , XeF y XeCl , que emiten en 193, 248, 308 y 350 nm respectivamente. En todos ellos, la presión total de trabajo es menor a 5 atmósferas, donde el 90% al 99% corresponde a un gas buffer (comúnmente helio o neon) que colabora en la transferencia de energía. El gas raro está presente en una proporción del 1% al 9% y el halógeno lo está entre el 0,1% y 0,2%. La energía es depositada en el gas por medio de una descarga eléctrica rápida de alta tensión (~30 kV) y alta corriente pico (~3 kA) que se logra con capacitores de 50 a 70 nF. Esto debe ser así pues el nivel superior posee una vida media corta, de pocas decenas de nanosagundos, por lo que resulta necesario invertir la población en un tiempo muy corto de un láser de tres niveles. Debido a esto, es frecuente que el sistema de descarga utilice pulsos de preionización que prepara al gas en la cámara para la descarga principal, que le sigue

unos nanosegundos después. Esta se produce en dirección perpendicular al eje longitudinal de la cavidad. El gas se degrada con el uso y se pierde materia activo, de modo que es necesario renovar el volumen del canal de descarga entre pulsos. Es posible adicionar periódicamente pequeñas cantidades de halógeno para extender la vida útil del gas. Actualmente se diseñan cavidades selladas con posibilidad de purgado que permiten operar hasta 10^8 disparos sin cambiar el gas. Tanto la cavidad como la óptica y los electrodos deben ser resistentes a la corrosión impuesta por los halógenos presentes en el gas. Debido a la alta ganancia de este láser, la cavidad solo necesita un espejo de alta reflexión; el 5% de reflectividad de una superficie óptica común provee la necesaria retroalimentación para obtener emisión estimulada. Debido a que esta emisión tiene un ancho de banda de unos 5 nm, es posible introducir una red de difracción intracavidad que permite una sintonía fina dentro de la curva de ganancia.

Las aplicaciones del láser de excímero se han incrementado desde su creación debido a su capacidad de generar pulsos UV de alta energía a repeticiones elevadas como para obtener potencias promedio que superan los 100 W. Las primeras aplicaciones fueron en el campo científico, como fuentes de bombeo de láseres de colorante y como instrumento de investigación en temas de fotoquímica. También se lo utiliza en estudios de espectroscopía no lineal, plasmas, fusión inducida por láser, sensado remoto de componentes atmosféricos y medición de concentración de ozono. La alta energía de los fotones UV de un excímero destruye las ligaduras de las moléculas componentes de una sustancia, generando la ablación del material en la zona de incidencia. Esta propiedad, junto con el gran avance producido en materia de seguridad, confiabilidad y durabilidad de este láser, ha hecho que pasara de instrumento de laboratorio a herramienta de producción industrial. Así, estos láseres han encontrado aplicación en la fabricación de superconductores de alta temperatura por deposición de films delgados y uniformes a partir de vapores que surgen de la ablación de materiales cerámicos. Por otro lado, la corta longitud de onda de emisión permite enfocar el haz en dimensiones menores que un micrómetro, lo que se utiliza en procesos de microlitografiado en la industria de semiconductores y grabado de patrones en la fabricación de circuitos integrados (figura 31).

FIGURA 31



VI. Láseres de semiconductor

Según nuestra clasificación previa, los láseres de semiconductor se inscriben dentro de los láseres de estado sólido. Sin embargo, como dijimos al tratar éstos, los primeros han crecido y madurado enormemente desde su nacimiento a principios de los '60. A esto contribuyó el notable progreso observado en la década del '80 sobre deposición de capas finas y fabricación de estructuras ultrapequeñas (del orden de los nanómetros) en materiales cristalinos (óptica integrada) y sólidos en general. También contribuyó el interés en la utilización de estos láseres como fuentes emisoras compactas en dispositivos optoelectrónicos y en sistemas de comunicaciones por fibra óptica. Si a esto le sumamos el hecho de que los láseres de semiconductor son bombeados eléctricamente (inyección de portadores) y no ópticamente como los láseres de cristal, debemos dedicarle una sección especial aparte de la correspondiente a láseres sólidos. La amplia variedad de dispositivos láser de semiconductor y a la diversidad de estructuras hace imposible aquí una descripción detallada del tema por su gran extensión. Intentaremos dar un pantallazo de las distintas posibilidades, comenzando por un breve repaso de la física de semiconductores para entrar luego en la descripción de la estructura y funcionamiento específico de los láseres de semiconductor.

Cuando varios átomos de un dado elemento se agrupan en determinada región espacial para formar un sólido, se generan diversas interacciones debido al campo coulombiano, de forma que las fuerzas de atracción y repulsión entre ellos encuentran un balance a una determinada distancia interatómica. En este proceso se producen importantes cambios en los niveles de energía de los electrones debido a que el llamado principio de exclusión de Pauli (que indica la imposibilidad de que dos electrones posean el mismo conjunto de números cuánticos que definen su nivel de energía) comienza a dominar la distribución de éstos. A medida que la distancia interatómica disminuye, las funciones de onda de los electrones (cuyo módulo al cuadrado indica la probabilidad de encontrar al electrón en determinado punto del espacio) comienzan a solaparse y, para cumplir con el principio de Pauli, se produce un desdoblamiento de los niveles de energía discretos de los átomos aislados. Estos nuevos niveles corresponden a los electrones como parte integrante del conjunto de átomos como un todo. Dicho desdoblamiento será mayor cuanto mayor sea el número de átomos que se agrupan. Así, en un sólido real, este número es muy grande (del orden de 10^{18}) y estos niveles finamente desdoblados forman bandas de energía permitidas separadas por espacios de energía prohibida (figura 32). las bandas permitidas más bajas están siempre

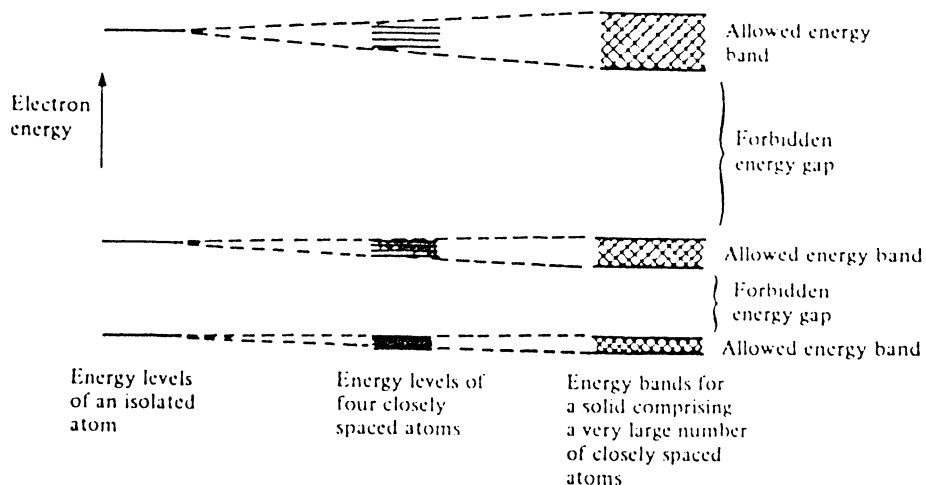


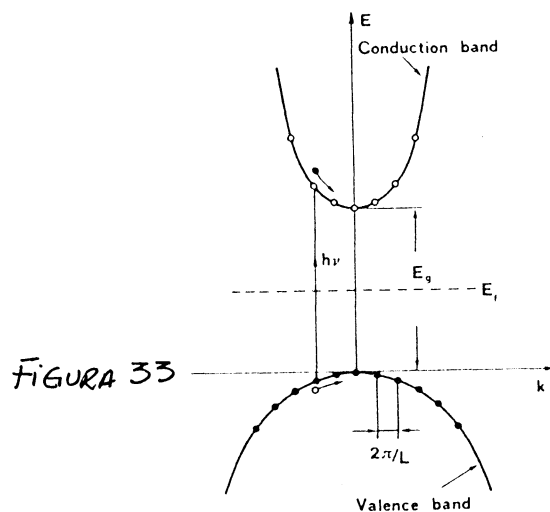
FIGURA 32

ocupadas con electrones y no juegan un rol importante en cuanto a las propiedades eléctricas y ópticas se refiere. En cambio si lo hacen los electrones que ocupan las bandas superiores, en particular las dos más altas llamadas de valencia y de conducción, separadas por una región prohibida (“gap”). En el caso de materiales aislantes, la banda de conducción está totalmente vacía y el gap es grande (~ 4 eV) comparado con la energía térmica a temperatura ambiente (~ 0,02 eV), de forma que prácticamente ningún electrón podrá ser transferido a la banda de conducción y la aplicación de un campo eléctrico no producirá movimiento de cargas. El caso de los conductores es opuesto: las bandas se solapan levemente de modo que siempre hay electrones disponibles en la banda de conducción para generar una corriente. Existe un caso intermedio donde el gap posee valores entre 1,5 eV y 2,7 eV aproximadamente en los que es posible “bombear” electrones a la banda de conducción y de allí producir una corriente eléctrica: estos son los semiconductores.

Dentro de cada banda, los valores permitidos de energía se pueden relacionar con los correspondientes valores permitidos de k (constante de propagación) por

$$E_c = \hbar^2 k^2 / 2 m_c \quad \text{y} \quad E_v = \hbar^2 k^2 / 2 m_v$$

Los valores de E vs k se grafican en la figura 33 como puntos negros y blancos, separados a intervalos regulares de $2\pi/L$, siendo L la dimensión del cristal.



Como la función de onda electrónica se extiende por todo el cristal, se debe aplicar el principio de exclusión de Pauli para asegurar que cada nivel de energía sea ocupado a lo sumo por dos electrones de spin opuesto. Así, la probabilidad de ocupación de un dado nivel de energía E viene expresada por la estadística de Fermi-Dirac:

$$f(E) = \{1 + \exp[(E - E_f)/kT]\}^{-1}$$

donde E_f es el llamado nivel de Fermi y representa el límite entre estados ocupados y vacíos para $T=0^\circ$ K. Para cualquier temperatura, $f(E_f) = 1/2$. Una representación de esta función puede verse en la figura 34, donde se aprecia la forma que tiene la distribución de Fermi-Dirac para temperatura de cero absoluto y mayores.

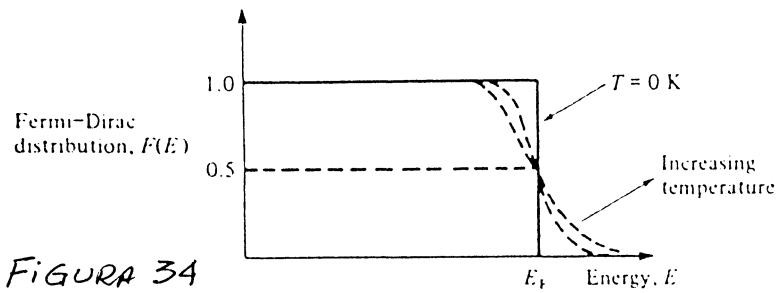


FIGURA 34

Supongamos ahora que tenemos un semiconductor ideal a $T=0^\circ\text{K}$ con la banda de valencia totalmente ocupada. Si de alguna manera podemos llevar electrones de la banda de valencia a la de conducción, veremos que después de un corto tiempo ($\sim 10^{-13}$ seg) los electrones en la banda de conducción caen en forma no radiativa a los niveles más bajos de la misma, y cualquier electrón cercano al límite superior de la banda de valencia relajará a los niveles desocupados más bajos, quedando la parte superior de esta banda llena de “huecos”. Esto significa que puede conseguirse inversión de población entre las bandas para obtener acción láser si el material es colocado en un resonador apropiado. La recombinación electrón-agujero da lugar a un fotón cuya energía es mayor que el gap interbanda. Esto puede verse en la figura 35.

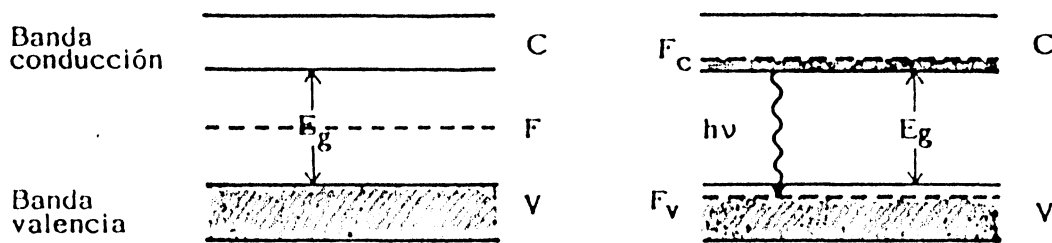


FIGURA 35

Usualmente el ancho de banda de ganancia en un semiconductor es bastante grande ($\sim 400\text{ cm}^{-1}$), aunque no tan grande como el de un colorante ($\sim 2000\text{ cm}^{-1}$). La forma más conveniente de operación de un semiconductor es utilizando una configuración de diodo en la que la excitación se produce por circulación de corriente en sentido directo. En este caso, la inversión de población tiene lugar en una región estrecha ($< 1\mu\text{m}$) entre los lados p y n de la juntura. Se pueden distinguir dos clases de diodos láser: los de homojuntura (que fueron los primeros en demostrar su funcionamiento) y los de heterojuntura doble (inventados en 1969). Estos últimos hicieron posible la operación a temperatura ambiente de los láseres de semiconductor en régimen continuo. Describiremos brevemente los mecanismos de operación de ambos para ver las ventajas de los segundos sobre los primeros.

En el láser de homojuntura, el proceso de bombeo se logra en una juntura p-n donde ambas regiones están hechas del mismo material semiconductor (p.ej. arseniuro de galio, GaAs). Ambas regiones están fuertemente dopadas ($\sim 10^{18}$ átomos/cm³) de modo que los niveles de Fermi caen dentro de la banda de valencia para la región tipo p y dentro de la banda de conducción para la región tipo n. Cuando no hay un voltaje aplicado en la juntura, ambos niveles de Fermi yacen en el mismo nivel de energía (ver figura 36(a)). Cuando se aplica una tensión V , los niveles se separan en $\Delta E = e.V$, y la

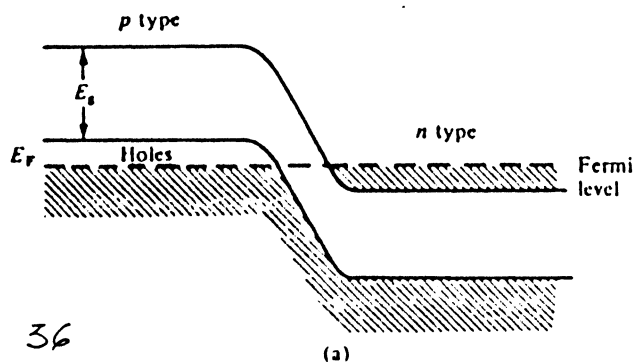
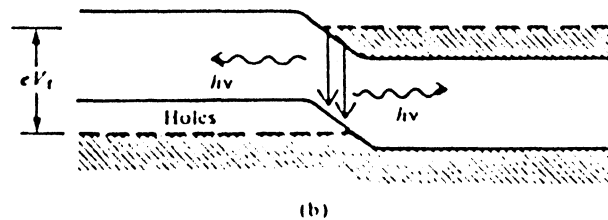


FIGURA 36



estructura de bandas se verá como en la figura 36(b), donde se aprecia que se ha producido inversión de población en la región de la juntura. La tensión aplicada inyecta en la capa activa electrones de la banda de conducción del material tipo n y agujeros de la banda de valencia del material tipo p. Cuando un electrón llega a la región p, se convierte en portador minoritario y difunde hasta recombinarse con un agujero, emitiendo un fotón. Así, el espesor d de la región activa es del orden de la distancia media que viaja un electrón antes de recombinarse. Para GaAs, $d = 1 \mu\text{m}$, y $V = \Delta E / e \approx E_g / e = 1,5 \text{ V}$. Debido a efectos de difracción, la dimensión transversal del haz láser que se propaga por la región activa es del orden de $5 \mu\text{m}$, mucho mayor que d , de forma que se extiende considerablemente dentro de las regiones p y n, donde es fuertemente absorbido. Por esto, la densidad de corriente de un láser homojuntura a temperatura ambiente es alta ($\sim 10^5 \text{ A/cm}^2$) y debe operarse a temperaturas criogénicas para bajar la corriente umbral, lo que constituyó el más serio obstáculo para el desarrollo de estos láseres en la década del '60.

A fines de dicha década se desarrolló el láser de heterojuntura doble (HD). La figura 37 muestra un ejemplo de un láser de GaAs de HD. En este diodo hay dos junturas [AlGaAs(p)-GaAs y GaAs-AlGaAs(n)] entre diferentes materiales. La región

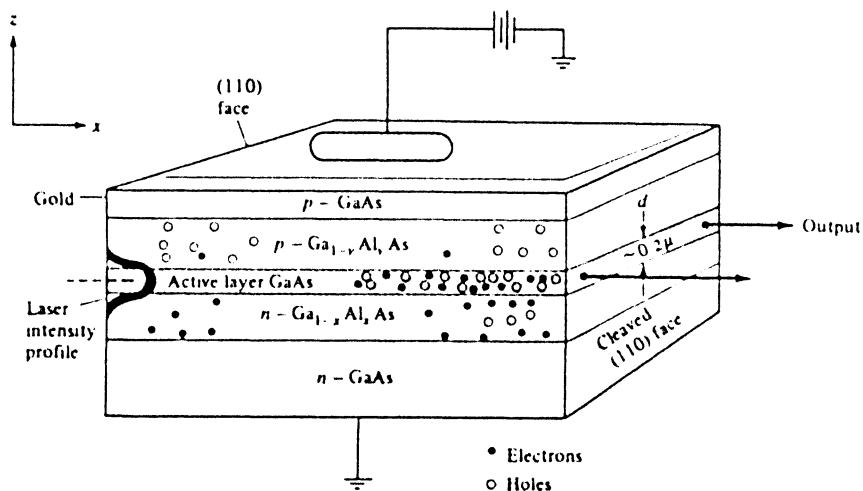


FIGURA 37

activa consiste de una capa delgada de GaAs de $0.2 \mu\text{m}$. Con esta estructura la densidad de corriente umbral a temperatura ambiente puede reducirse en dos órdenes de magnitud comparada con el láser homojuntura, haciéndose posible la operación continua a temperatura ambiente. La reducción de corriente umbral se debe al efecto combinado de tres circunstancias: (a) el índice de refracción del GaAs ($n=3,6$) es bastante mayor que el correspondiente a AlGaAs ($n=3,4$) generando de esta forma una estructura de guía de onda que obliga al láser a estar confinado en la región de GaAs, que es donde hay ganancia; (b) el gap de la región activa ($1,5 \text{ eV}$) es menor que el de AlGaAs ($1,8 \text{ eV}$). Esto genera barreras de energía en las dos junturas que confinan a los portadores inyectados dentro de la región activa, aumentando la concentración de electrones y huecos y por lo tanto la ganancia; (c) como el gap de la región activa es apreciablemente menor que el del material vecino, los extremos del perfil del modo que se extienden fuera de la región activa no son tan fuertemente absorbidos (figura 38)

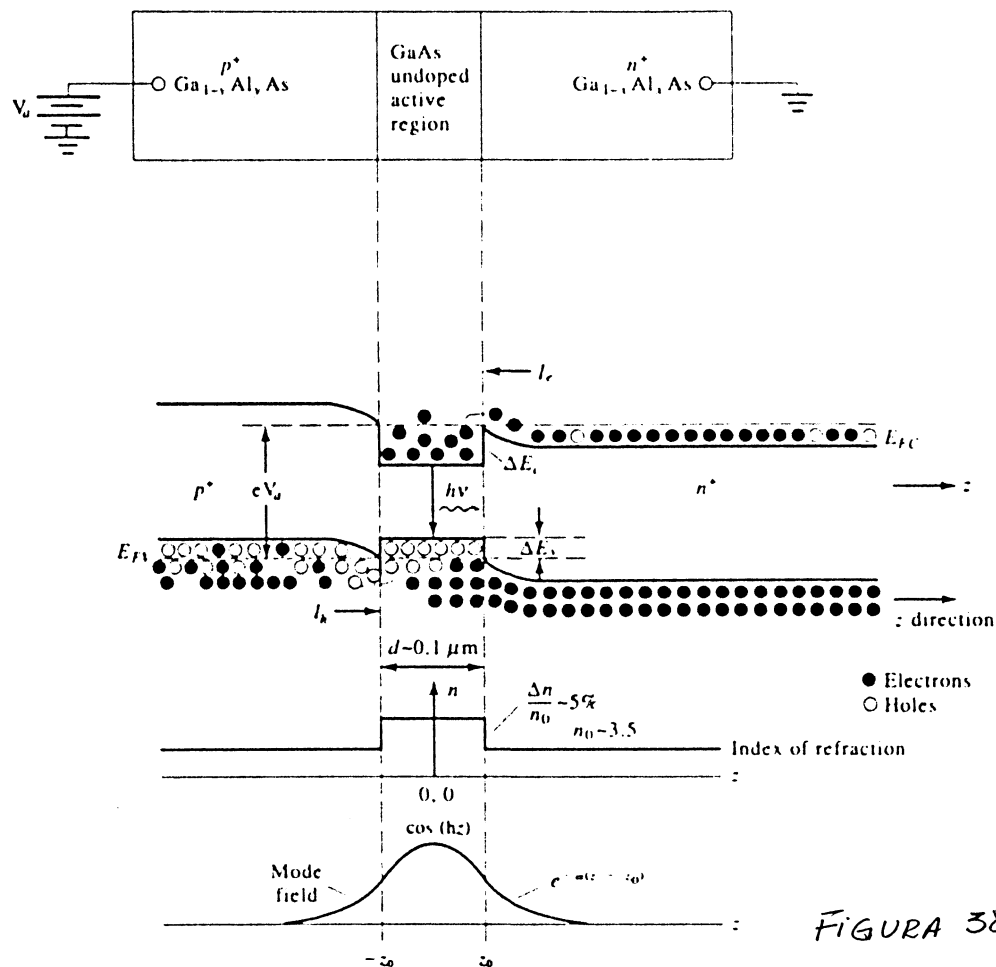


FIGURA 38

El láser de GaAs tiene una longitud de onda de emisión de $0,85 \mu\text{m}$, que coincide con un mínimo local de la curva de pérdida de las fibras ópticas de silicio (primera ventana de transmisión). Como esta curva tiene otros dos mínimos en $\lambda = 1,30 \mu\text{m}$ y $1,55 \mu\text{m}$ (segunda y tercera ventana) se han desarrollado láseres de heterojuntura doble con emisión en esas longitudes de onda. Para ello se utiliza la aleación cuaternaria $\text{In}_{1-x}\text{Ga}_x\text{As}_y\text{P}_{1-y}$, mientras que las regiones p y n de la juntura se hacen del compuesto binario InP. En este caso se añade una nueva condición: el período de la red de la aleación debe coincidir, dentro del 0,1%, con el del material InP, para que no se desarrollen tensiones mecánicas que destruirían la juntura, Esto se logra para una

relación de índices de composición $y = 2,2 x$. Eligiendo apropiadamente x , puede variarse la longitud de onda dentro del rango $0,92$ y $1,5 \mu\text{m}$.

Una configuración muy difundida es la que se muestra en la figura 39. La introducción de una capa aisladora de geometría apropiada debajo del contacto metálico positivo obliga a la corriente a fluir en una banda estrecha ("tira") de ancho entre $5-10 \mu\text{m}$. Esto trae dos efectos benéficos: (a) como el área de la tira es pequeña,

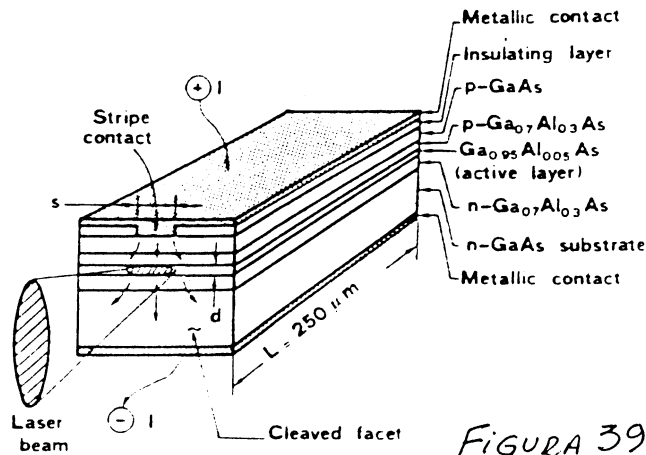


FIGURA 39

la corriente umbral $I_u = J_u A$ también lo es (p.ej. si $J_u = 2 \cdot 10^3 \text{ A/cm}^2$ y $s = 10 \mu\text{m}$, entonces $I_u = 50 \text{ mA}$). (b) como el ancho de la región de ganancia en el plano de la juntura es aproximadamente igual al ancho s de la tira, el confinamiento de la ganancia permite circunscribir al haz dentro de los límites del modo fundamental transversal si $s \leq 10 \mu\text{m}$. A esta configuración en la que hay confinamiento de la distribución de ganancia, se la conoce como *láser de ganancia guiada*. Existe otra estructura en la que el confinamiento del haz dentro de la región de ganancia se realiza por un apropiado perfil de índice de refracción dentro del plano de la juntura y se lo conoce como *láser de índice guiado*. En ambos casos, el astigmatismo del haz de salida es compensado con sistemas ópticos adecuados.

Un espectro de emisión típico de un diodo láser se muestra en la figura 40. Los picos igualmente espaciados corresponden a diferentes modos longitudinales de la cavidad. Típicamente para un diodo láser de GaAs suponiendo que $\lambda = 0,85 \mu\text{m}$ y que la

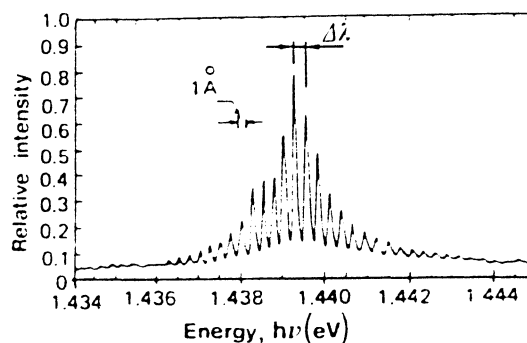


FIGURA 40

longitud de la cavidad L es $250 \mu\text{m}$, $\Delta\lambda_p = 3,9 \text{ \AA}$. Así, la emisión espectral abarca un amplio rango de longitudes de onda ($5-10 \text{ nm}$), formada por varios modos longitudinales. Para obtener anchos de línea menores se utiliza el dispositivo de

realimentación distribuida (distributed feedback, DFB). En este esquema el diodo láser se fabrica de tal forma de tener una variación periódica del índice de refracción de la capa activa a lo largo de la dirección de propagación. En cada plano donde cambia el índice de refracción se produce una retroreflexión que inyecta parte de la onda propagante al medio con ganancia. Dependiendo del período de esa variación, las ondas reflejadas estarán en fase (interferencia constructiva) para una determinada longitud de onda, que se verá amplificada en su paso por el medio. Esto reduce notablemente el ancho de línea de la emisión, usualmente a un solo modo longitudinal de la cavidad (figura 41).

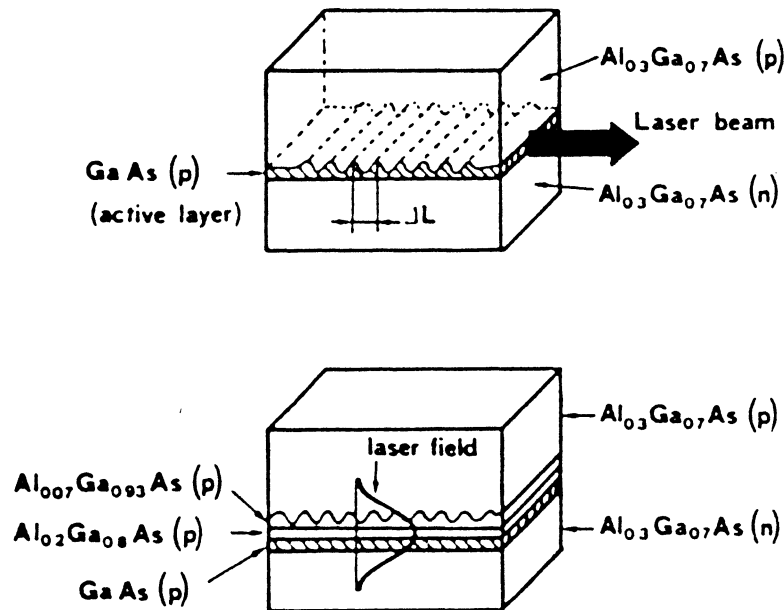


FIGURA 41

Recientemente, se han desarrollado los llamados láseres de pozo cuántico. En los láseres normales de HD, el espesor de la región activa va de 0,1 a 0,3 μm . Sin embargo, la habilidad de crecer capas ultradelgadas (entre 40 y 100 \AA) comparables con la longitud de onda de deBroglie del electrón, trajo aparejado la observación de efectos de tamaño cuántico. Estos se manifiestan principalmente en la restricción del número de niveles de energía en el fondo de la banda de conducción (y de la banda de valencia), lo cual hace que las transiciones sean más selectivas (monocromáticas). Por otro lado, como la corriente de inyección para mantener la inversión de población es proporcional a

$$I \sim (e V_a \times 10^{18}) / \tau$$

donde V_a es el volumen de la región activa. Un angostamiento de ésta reduce el volumen y por lo tanto la corriente umbral, por lo que los láseres de pozo cuántico poseen corrientes umbrales de unos pocos miliamperes. Uno de los problemas de los láseres de pozo cuántico simple es que, debido a la estrechez de la región de ganancia, el confinamiento óptico es muy pobre, lo que aumenta las pérdidas. Una forma de solucionar esto, es crear varios pozos cuánticos uno al lado del otro de forma de tener un confinamiento mucho mayor de la radiación. A estos dispositivos se los llama

láseres de pozo cuántico múltiple. La figura 42 muestra los detalles de la región activa de un diodo láser de pozo cuántico múltiple, donde los electrones son inyectados desde la izquierda y los huecos desde la derecha. Los pozos son tan delgados que los portadores inyectados tienen la capacidad de atravesar las regiones de GaAs debido a la ausencia de colisiones en ese trayecto. Consecuentemente, la concentración de aluminio debe expandirse sobre distancias mayores para permitir el atrapamiento de los portadores en la región central donde tienen lugar las transiciones radiativas de recombinación. Estos pozos cuánticos múltiples se forman alternando capas delgadas de GaAs con capas aun más delgadas de AlAs. Estas últimas crean barreras de potencial a la función de onda del electrón dentro del pozo de GaAs, desde donde se produce la transición láser. Nótese también que en la región del pozo, el índice de refracción es mayor que en las vecindades, de modo que se favorece el guiado óptico de la radiación.

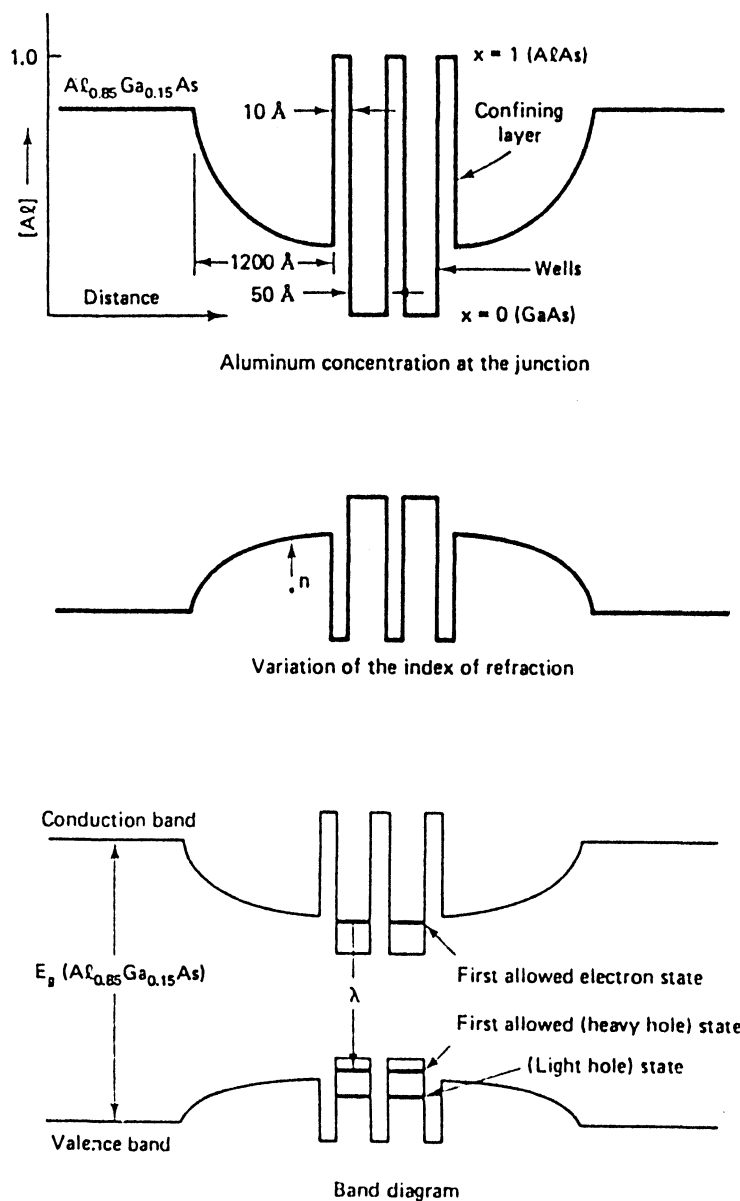


FIGURA 42

VII. Tendencias actuales

Se puede afirmar sin lugar a dudas que las actuales tendencias en materia de láseres es hacia la utilización y desarrollo extensivo de dispositivos de estado sólido. La experiencia adquirida en el crecimiento de cristales, en el dopaje de materiales y en la fabricación de nanoestructuras en semiconductores como así también la alta confiabilidad de estos materiales han hecho que se vuelquen grandes esfuerzos hacia el desarrollo de este tipo de láser.

En cuanto a los semiconductores, la tendencia se divide en dos líneas; una es la fabricación de diodos láser de potencia para bombeo de láseres de cristal. Estos dispositivos toman la forma de arreglos uni o bidimensionales que trabajan en fase y pueden llegar a entregar potencias continuas entre 10 y 20 W, con circuito de refrigeración por agua. Conjuntamente se han diseñado ópticas de colimación y enfoque de la radiación emitida sobre el medio activo. La segunda línea se dirige hacia el desarrollo de diodos láser con emisión en longitudes de onda cortas (500-400 nm), basada en el estudio de aleaciones de compuestos II-VI de la tabla periódica. pues tienen un gap mayor y consecuentemente una menor longitud de onda de emisión. La figura 43 muestra la variación de la constante de red con la energía del gap (o longitud de onda) para varios compuestos III-V. El área rayada indica el compuesto cuaternario

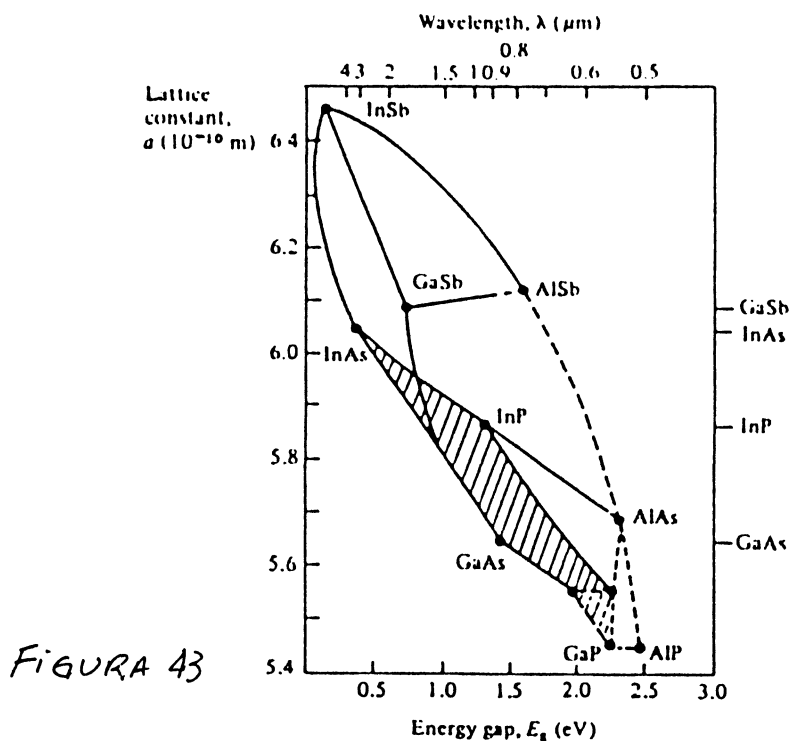


FIGURA 43

GaInAsP, con el que se obtienen láseres que emiten entre 1,1 y 1,6 μm . Ajustando las concentraciones de los componentes, es posible ajustar las constantes de red de ambos semiconductores y evitar las tensiones mecánicas de un desajuste que compromete la estabilidad a corto plazo. En el año 1994 se desarrolló un dispositivo semiconductor basado en una aleación II-VI (ZnSe) que logró emisión láser en 483 nm en régimen pulsado a temperatura ambiente, aunque solo por unos minutos. Con un dispositivo similar se logró emisión continua en 508 nm también por poco tiempo. Estos compuestos presentan dificultades de dopaje y de ajuste de red. Otros dispositivos que

se están investigando son los compuestos nitrurados como el InGaN de pozo cuántico múltiple.

La firma Coherent comercializa un láser de 10 mW que opera en régimen continuo con emisión en 430 nm, basado en el doblado directo de frecuencia de un diodo láser de potencia media en 860 nm. Por otro lado, la firma Spectra Physics ha puesto en el mercado un láser continuo de potencia media todo-estado-sólido que tiende a reemplazar al láser de argón. Este láser se basa en una barra de Nd:YVO bombeada por una barra de diodos de potencia, cuya salida es duplicada en frecuencia por un cristal no lineal, siendo capaz de entregar entre 2 y 5 W (figura 44). Un joint-venture entre Hitachi y Melles Griot ha desarrollado un láser todo-sólido de Cr:LiSAF bombeado por diodos y con doblado de frecuencia intracavidad que opera en forma continua, entregando 10 mW en 430 nm. Estos ejemplos sirven para mostrar la fuerte tendencia a utilizar láseres de estado sólido aun en aquellos casos donde los láseres gaseosos eran los láseres de elección (bombeo de colorantes , etc)

Verdi 532 nm Output Power vs. Diode Pump Power

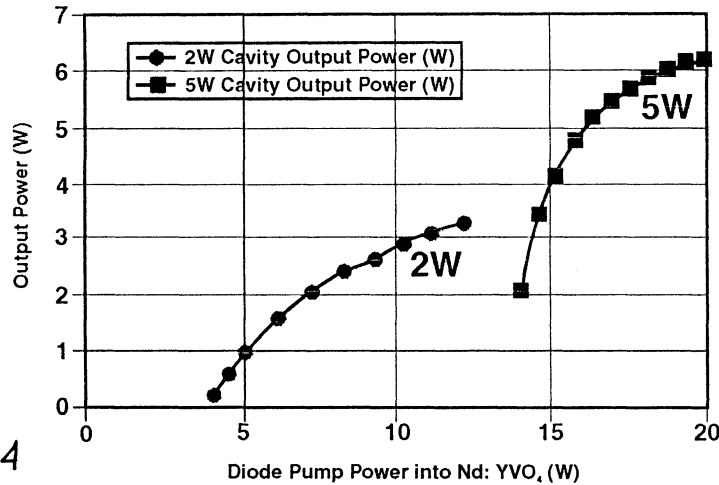


FIGURA 44

Finalmente, digamos que actualmente se ha avanzado sobre los llamados osciladores paramétricos ópticos (OPO). Estos sistemas se basan en un cristal no lineal sobre el que incide un haz láser intenso de frecuencia ω_3 y se genera dentro del mismo amplificación en las frecuencias ω_2 y ω_1 tal que se cumple: $\omega_3 = \omega_2 + \omega_1$ (figura 45).

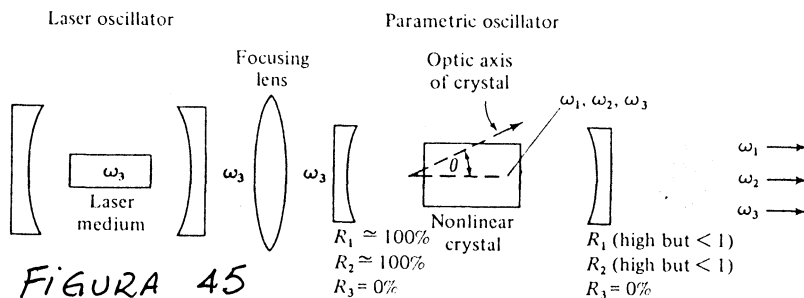


FIGURA 45

Si la frecuencia de bombeo ω_3 es fija, es posible sintonizar las otras dos de forma que siempre se cumpla la condición anterior. Esta sintonía se logra por cambio de temperatura o por rotación mecánica del cristal, de forma que se puede obtener salida láser sintonizable en un muy amplio rango de longitudes de onda (p.ej. desde 355 nm a

TIPO DE LASER	MEDIO ACTIVO	FORMA DE EXCITACION	REGIMEN DE OPERACION	POTENCIA TIPICA DE EMISION	LONGITUD DE ONDA DE EMISION (nm)	REGION DEL ESPECTRO DONDE EMITEN
EXCIMEROS	GAS	ELECTRICO	PULSADO	MEDIA/ALTA	152-351	ULTRAVIOLETA
NEODIMIO (Nd-Yag)	SOLIDO	OPTICO	PULSADO CONTINUO	MEDIA/ALTA	266; 355 532 1064; 1320	ULTRAVIOLETA VERDE INFRARROJO
HELIO CADMIO (He-Cd)	GAS	ELECTRICO	CONTINUO	BAJA	325 442	ULTRAVIOLETA AZUL
NITROGENO (N ₂)	GAS	ELECTRICO	PULSADO	BAJA/MEDIA	337	ULTRAVIOLETA
RUBI	SOLIDO	OPTICO	PULSADO	MEDIA/ALTA	347 694	ULTRAVIOLETA ROJO
COLORANTES	LIQUIDO	OPTICO	PULS./CONT.	BAJA/MEDIA	205 - 1000	DESDE EL ULTRAVIOLETA AL INFRARROJO
IONICOS DE GASES NOBLES (Ar, Kr, Xe)	GAS	ELECTRICO	PULSADO CONTINUO	MEDIA/BAJA	350 457, 488, 514; 540 647	ULTRAVIOLETA AZUL-VERDE ROJO
VAPORES METALICOS (ORO - COBRE)	GAS	ELECTRICO	PULSADO	MEDIA	510; 578 628	VERDE ROJO
HELIO - NEON (He-Ne)	GAS	ELECTRICO	CONTINUO	BAJA	543, 594 611, 632 1152, 1524, 3392	VERDE ROJO INFRARROJO
DIODOS SEMI CONDUCTORES	SOLIDO	ELECTRICO	PULSADO CONTINUO	BAJA	635 - 900 1100 - 1600 2700 - 3000	ROJO A INFRARROJO
ALEXANDRITA	SOLIDO	OPTICO	PULSADO	MEDIA	700 - 800	INFRARROJO
TITANIO ZAFRO	SOLIDO	OPTICO	PULS./CONT.	MEDIA/ALTA	670 - 1100	ROJO A INFRARROJO
ERBIO / HOLMIO	SOLIDO	OPTICO	PULSADO	BAJA	850; 1730, 1540; 2060	INFRARROJO
CENTROS DE COLOR	SOLIDO	OPTICO	CONTINUO	BAJA	1400 - 1600 2300 - 3300	INFRARROJO
LASERES QUIMICOS (DF, H, F)	GAS	OPTICO	PULSADO CONTINUO	ALTA	1300; 2600 - 3000 3600 - 4000	INFRARROJO
MONOXIDO DE CARBONO (CO)	GAS	ELECTRICO	CONTINUO	MEDIA/ALTA	5000 - 6000	INFRARROJO
DIOXIDO DE CARBONO (CO ₂)	GAS	ELECTRICO	PULSADO CONTINUO	ALTA	9000 - 11000	INFRARROJO

Referencias: 1 nm = 10⁻⁹ m

PULS./CONT. = PULSADO CONTINUO

2000 nm) con intensidad prácticamente constante. Si bien el principio de funcionamiento de los OPO se conocía desde poco después de la aparición del láser, la gran mejora en el crecimiento de cristales no lineales y en la calidad espacial del haz de bombeo introducidas hace pocos años, hizo resurgir estos dispositivos que seguramente llegaron para quedarse.

Bibliografía

- 1.- Ranea Sandoval H.F, Scaffardi L.B. *Algunos tipos de láseres* , Curso Láser y Óptica en Ingeniería, Ciop, 1985
- 2.- Verdeyen J.T *Laser Electronics* (2nd ed.), Prentice Hall, 1989
- 3.- Wilson J., Hawkes J.F.B. *Optoelectronics* (2nd ed), Prentice Hall, 1989
- 4.- Yariv A. *Optical Electronics* (4th ed), Saunders College Publ. 1991
- 5.- Svelto O. *Principles of Lasers* (3rd ed.) Plenum Press, 1989
- 6.- *Laser Focus World* Pennwell Publishing, monthly publication, 1996-1997

ESPECTROSCOPIA ATOMICA Y MOLECULAR

J.G. Reyna Almandos

Introducción

En el año 1666 Newton realizó la observación del espectro visible de la luz del sol, haciéndola pasar a través de un prisma de vidrio y observando la figura característica sobre una pantalla. A comienzos del siglo XIX, este mismo espectro fue estudiado con más detalle por Wollaston y también por Fraunhofer, encontrándose que en la composición del mismo había líneas oscuras, las que posteriormente fueron atribuidas al espectro de absorción de la luz, debido a las capas más externas de la atmósfera solar. Particularmente relevante resulta luego el trabajo realizado por Kirchoff, quien en 1859 establece las bases del análisis espectroquímico, y que a partir de entonces es utilizado por varios investigadores para profundizar estudios sobre los espectros provenientes de diferentes elementos o sustancias analizadas.

Posteriormente, los trabajos de Balmer, Rydberg, Ritz, y la interpretación teórica de Bohr sobre el espectro del átomo de hidrógeno constituyen, junto con el desarrollo de la mecánica cuántica, el fundamento teórico sobre el que se basan actualmente las espectroscopías atómica y molecular.

Si analizamos la radiación emitida por una determinada fuente luminosa, encontraremos que el espectro obtenido es característico de los átomos o moléculas que se excitan para producir la descarga luminosa. En el caso de los espectros atómicos observamos que los mismos están constituidos por líneas que se agrupan de acuerdo a regularidades características del elemento excitado (Figura 1). Se dice entonces que estamos en presencia de un espectro de líneas o característico. Si nos referimos a los espectros moleculares, observaremos en este caso que las líneas se han agrupado de manera tal que forman lo que se denominan bandas (Figura 2), como en el caso de los gases moleculares, o simplemente un continuo como en el caso de los espectros emitidos por sólidos. Estos espectros son denominados entonces, espectros de bandas o espectros continuos.

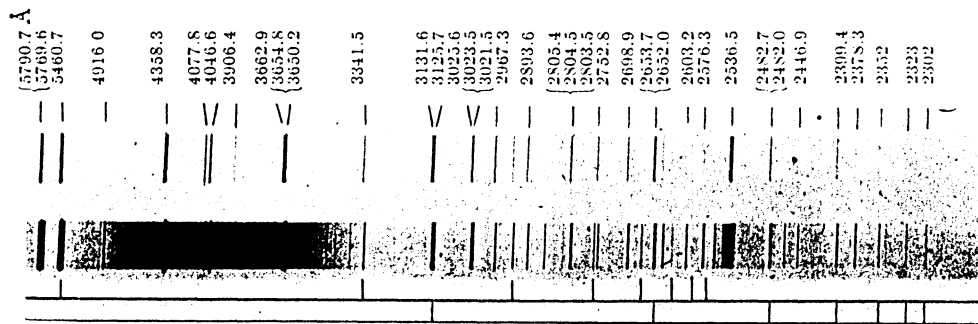


Figura 1. Parte del espectro del Fe.

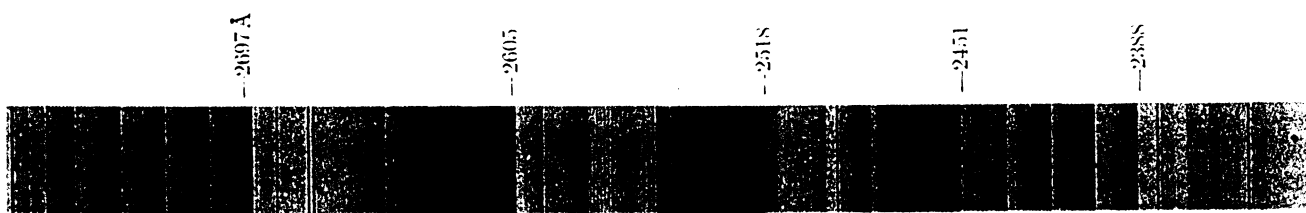


Figura 2. Espectro de banda de la molécula de PN.

Este tipo de figuras o regularidades características de los espectros, nos permitirán obtener, luego de un estudio apropiado, información sobre los niveles de energía del medio excitado, temperatura, presión, densidades electrónicas, etc., así como otras magnitudes físicas que nos darán información sobre la estructura íntima de los átomos o moléculas estudiadas.

Con cierta generalidad se podría decir entonces, que la espectroscopía es aquella rama de la Física que está relacionada con el estudio de la materia a través de su interacción con la radiación electromagnética. La espectroscopía luego, puede ser usada para determinar la identidad, estructura y entorno de los átomos y moléculas, por el análisis de la radiación emitida o absorbida por ellos. Esto nos permite adentrarnos en el conocimiento de la materia, de acuerdo al rango espectral considerado y a las diversas técnicas de trabajo que correspondan a las frecuencias involucradas en éste. Es por ello que resulta común hoy en día hablar de la espectroscopía óptica, de microondas, nuclear, de superficies, etc.

En las últimas tres décadas se han producido algunos hechos significativos que le han dado a los estudios espectroscópicos un resurgimiento y una mayor profundización. Uno de ellos es la posibilidad de obtener, mediante la utilización de cohetes y satélites, información espectroscópica por debajo de los 200 nm, zona en donde la atmósfera terrestre comienza a absorber la radiación electromagnética proveniente del exterior de la Tierra y que es conocida como región espectral del ultravioleta de vacío.

Un segundo hecho es el interés de la comunidad científica internacional en contar con datos espectroscópicos provenientes de plasmas sometidos a muy altas temperaturas, relacionados con los estudios de fusión controlada.

Por último, y debido al desarrollo tecnológico recientemente comenzado en algunos laboratorios, cabe mencionar la necesidad de datos espectroscópicos para el estudio de los mecanismos de excitación intervinientes en el funcionamiento de los láseres con emisión de radiación en la zona del ultravioleta de vacío y de los rayos X.

También en los últimos años, el desarrollo de las distintas espectroscopías ha sido enorme, no sólo en lo atinente a su contribución a las áreas básicas de la Física, la Química, la Astronomía, etc., sino también en su aplicación a la resolución de problemas tecnológicos. Ejemplos de esto último pueden ser la aplicación de láseres al estudio de la contaminación del medio ambiente, o la muy utilizada técnica del análisis espectral para la determinación del grado de pureza de productos en la industria siderúrgica.

En este texto, el tratamiento que se dará estará restringido a la espectroscopía óptica, entendiéndose por tal aquella que cubre la región fotográfica del espectro electromagnético, o sea, desde la región del ultravioleta de vacío hasta el infrarrojo cercano.

Espectros atómicos

El espectro atómico más simple corresponde al del hidrógeno. En éste, las distintas líneas espectrales, que surgen de transiciones entre los niveles involucrados, se agrupan de acuerdo a series de líneas cuya separación e intensidad decrece regularmente hacia las longitudes de onda cortas. La regularidad de estas series espectrales puede ser interpretada de acuerdo a las expresión matemática

$$\sigma = \frac{1}{\lambda} = R \left(\frac{1}{n_2^2} - \frac{1}{n_1^2} \right) ,$$

en donde R es la constante de Rydberg, n_1 y n_2 son números naturales no nulos con $n_1 > n_2$. Los números n_2 y n_1 adoptan los siguientes valores para darnos las diferentes series:

<u>n_2</u>	<u>n_1</u>	<u>Serie</u>
1	2,3,4,...	Lyman
2	3,4,5,...	Balmer
3	4,5,6,...	Paschen
4	5,6,7,...	Brackett
5	6,7,8,...	Pfund

Si para cualquier serie se hace tender $n_1 \rightarrow \infty$, entonces ésta se aproxima a su límite $\sigma_\infty = R/n_2^2$. Este es el llamado límite de la serie y por encima de él se supone un continuo de energía que se corresponde con la presencia del átomo ionizado.

Al valor $T(n) = R/n^2$ se lo denomina término espectral y se ve que cualquier línea del espectro puede ser indicada como la diferencia entre dos términos de la serie

$$\sigma = T_2 - T_1 .$$

Aunque para otros elementos las expresiones de los términos son un poco más complicados que para el caso del hidrógeno, la fórmula anterior sigue siendo válida como consecuencia del principio de combinación de Rydberg-Ritz, el que establece que la diferencia entre cualquier par de términos de un átomo dado, es igual al número de onda de una línea

espectral del átomo.

Transiciones entre dos estados estacionarios i y j pueden ser mostrados en un diagrama como el de la Figura 3.

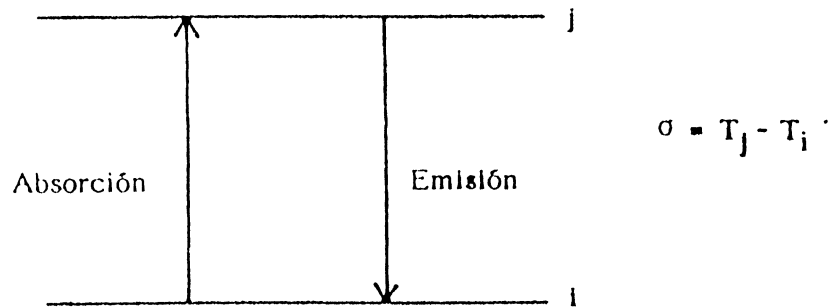


Figura 3

Utilizando este tipo de diagramas podemos ver que en los espectros atómicos simples, las líneas espectrales se agrupan a veces de a dos, formando dobletes, en otros casos de a tres, llamados tripletes, o en el caso más simple, singuletes, como se muestra en la Figura 4, en donde se reproduce una porción del espectro del CaI .

El agrupamiento de las líneas depende de la cantidad de electrones fuera de las capas cerradas que posee el elemento analizado, notándose además que los elementos con la misma cantidad de electrones fuera de las capas cerradas del átomo muestran espectros similares entre sí, como por ejemplo ocurre con los metales alcalinos.

En el caso de elementos más pesados, se observan mayores agrupamientos de líneas, las que reciben el nombre genérico de multipletes; y cuando estos son interpretados en términos de niveles de energía, los niveles muestran también un agrupamiento en multipletes.

Para poder explicar este comportamiento particular de los espectros, se deben tener en cuenta los distintos tipos de acoplamientos que se producen en el átomo entre los momentos angulares individuales de los electrones (l_i) en su rotación alrededor del núcleo, y el momento angular intrínseco o de spin (s_i) de cada electrón en su rotación sobre sí mismo.

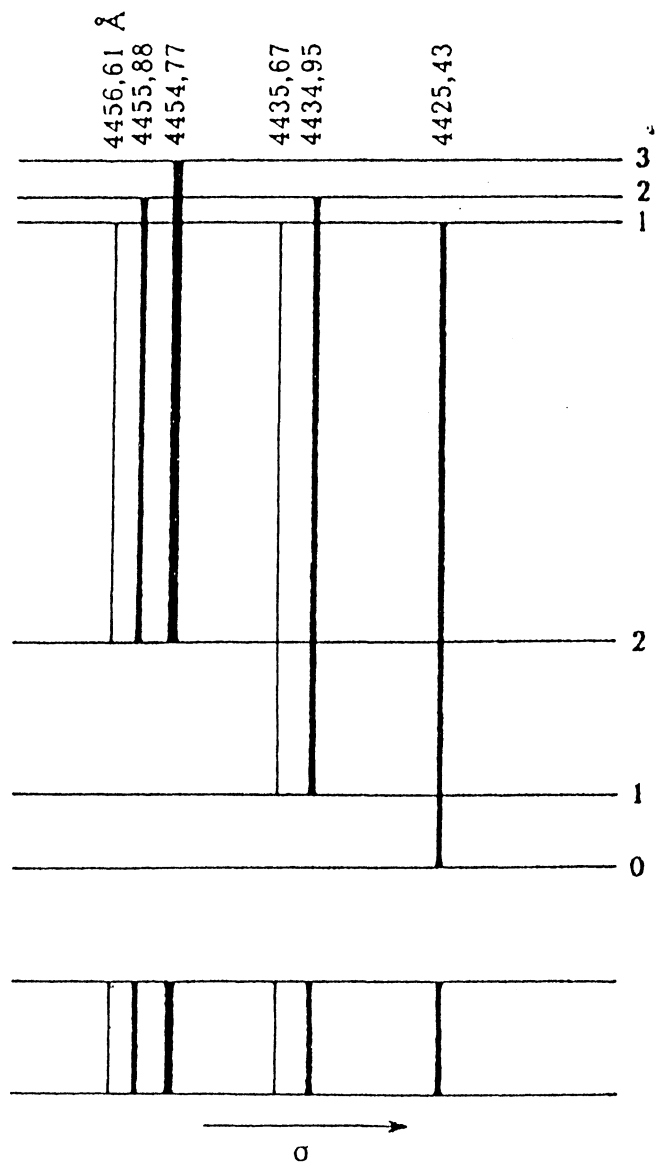


Figura 4

Acoplamiento L-S

Este acoplamiento, también llamado spin-órbita o de Russell-Saunders, se produce generalmente en átomos livianos. En este caso, los diferentes momentos angulares l_i de cada electrón se acoplan entre sí para dar un L resultante, y lo mismo ocurre con los momentos angulares de spin de cada electrón que dan un S resultante.

Estos momentos angulares L y S se acoplan entre sí resultando un momento angular total del átomo $J = L + S$.

Los valores que puede tomar el vector J varían desde

$$J = L + S \dots\dots |L - S|$$

Se ve que cuando $L > S$ el número de posibles valores de J para un dado L es $2S+1$. Este valor es llamado la multiplicidad de un término.

Veamos un ejemplo estudiando el problema de un átomo con dos electrones no equivalentes fuera de capa cerrada, esto es, electrones con distintos valores de n y l . En este caso consideraremos los electrones $3d$ y $4p$. Para esta configuración se obtienen los valores $L = 3, 2, 1$ dado que este número cuántico puede tomar los valores

$$L = (l_1 + l_2) \dots\dots |l_1 - l_2|$$

Los electrones de las capas cerradas (el "carozo"), no son tenidos en cuenta porque para ellos los valores del momento angular orbital L y de spin S son iguales a cero.

Los posibles valores de S van desde

$$S = (s_1 + s_2) \dots\dots |s_1 - s_2|$$

que en este caso serán $S = 1, 0$.

De acuerdo a la nomenclatura utilizada para denominar los términos electrónicos, se escribe el valor de $L(0, 1, 2, \dots)$ en su correspondiente notación (S, P, D, F, G, ...).

De lo expuesto surgen los siguientes términos:

$${}^3F, {}^1F, {}^3D, {}^1D, {}^3P, {}^1P$$

en donde como superíndice anterior se coloca la multiplicidad.

Finalmente, obtenemos la denominación de los niveles de energía colocando como subíndice del término el valor de J .

En la Figura 5 se puede ver una representación de los niveles de energía para un estado excitado del átomo de calcio, cuya configuración es:

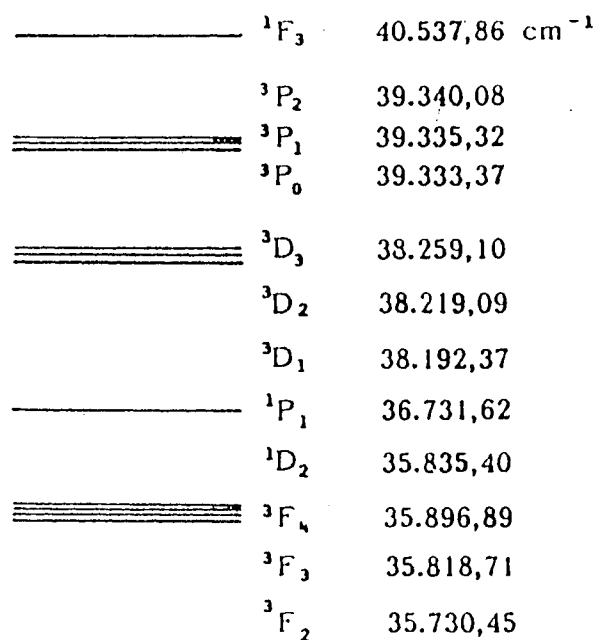
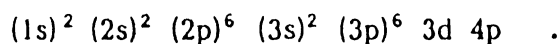


Figura 5

En este tipo de acoplamiento las reglas de selección para transiciones dipolares eléctricas son:

$$\Delta L = 0, \pm 1$$

$$\Delta S = 0$$

$$\Delta J = 0, \pm 1 \quad \text{con } J=0 \neq J=0$$

Vale también estrictamente la regla de Laporte referida a la simetría de los términos espectrales, la que dice que los términos pares pueden combinar sólo con términos impares, y los impares sólo con pares.

Acoplamiento J-J

En átomos pesados la carga nuclear llega a ser lo suficientemente grande como para producir interacciones spin-órbita comparables en magnitud a las interacciones electrostáticas entre los l_i y los s_i , y el esquema de acoplamiento L-S comienza a romperse. En ese caso, los momentos angulares totales j_i de los electrones se acoplan, formando momento angular total J del átomo.

Las reglas de selección $\Delta S=0$ y $\Delta L=0, \pm 1$ pierden validez y sólo se considera $\Delta J=0, \pm 1$ con $J=0 \neq J=0$, así como la regla de Laporte para los términos.

Acoplamientos intermedios

Los acoplamientos L-S y J-J constituyen casos límites. En el caso de los espectros correspondientes a los iones pesados, en los que ellos que aparecen a medida que crece el número atómico en la tabla periódica, observamos que el tipo de acoplamiento que se produce no corresponde exactamente ni al caso L-S ni al J-J. En este caso, decimos que son acoplamientos intermedios, y consideraremos el caso del tipo J-K o acoplamiento de a pares, ya que el mismo ocurre particularmente en configuraciones excitadas de gases nobles y en elementos del grupo del carbono, así como en otros casos. Se produce cuando la interacción electrostática es pequeña en relación a la interacción spin-órbita del carozo atómico, pero grande en relación al acoplamiento de spin del electrón óptico. Eso significa que existe un momento angular resultante definido del carozo atómico j_c , el cual se acopla con el momento angular orbital l del electrón óptico, dando un momento resultante K ,

$$l + j_c = K$$

A su vez, K se acopla con el valor S del momento angular de spin, dando un momento angular total

$$K + S = J$$

Espectros moleculares

El estado energético molecular está determinado no sólo por los cambios en las configuraciones electrónicas, como en el caso de los átomos, sino también por las vibraciones de los átomos que constituyen la molécula y por la rotación del conjunto

$$E = E_e + E_v + E_r \quad .$$

Para conocer la contribución que hacen a la energía total la energía rotacional y la vibracional, consideremos el caso simple de la molécula diatómica.

Energía rotacional

En este caso se considera a la molécula como constituida por dos masas separadas por una distancia r fija, rotando alrededor de un eje que pasa por su centro de masa. En este modelo los niveles de energía rotacionales quedan determinados por

$$E = \frac{1}{2I} J(J+1) \hbar^2 \quad .$$

en donde I es el momento de inercia de la molécula y J es el denominado número cuántico rotacional.

En la práctica, la energía rotacional es expresada en número de onda y designada por la letra F ,

$$F(J) = B J(J+1) \quad ,$$

en donde B es la denominada constante rotacional.

Vemos que los estados de energía del rotador rígido forman una sucesión de niveles discretos, como se muestra en la Figura 6.

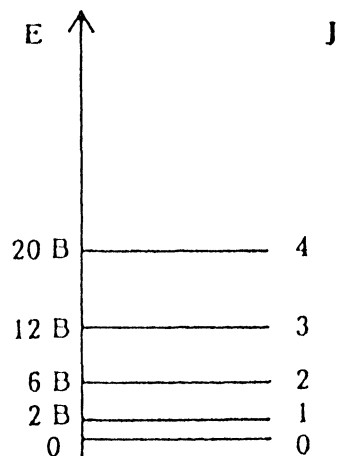


Figura 6

En realidad, la distancia interatómica r no es fija y se incrementa a una mayor frecuencia de rotación. Este efecto es tenido en cuenta escribiendo la energía de rotación como

$$F(J) = B J(J+1) - D J^2(J+1)^2 .$$

Las reglas de selección en este caso indican que $\Delta J = \pm 1$.

Energía vibracional

En el caso de una molécula diatómica, las fuerzas elásticas recíprocas que actúan luego de la separación con relación a la posición de equilibrio, pueden ser consideradas como proporcionales a la variación de la distancia interatómica. Teniendo en cuenta esto y resolviendo la ecuación de Schrödinger, obtenemos

$$E_v = h f_{\text{vib}} \left(v + \frac{1}{2} \right) .$$

en donde f es la frecuencia de la molécula y v puede tomar los va-

lores 0,1,2,... .

En la práctica se expresa a la energía vibracional por

$$G(v) = \frac{f_{\text{vib}}}{c} \left(v + \frac{1}{2}\right) = w_e \left(v + \frac{1}{2}\right) .$$

en donde w_e es la denominada constante vibracional.

Una mejor aproximación al modelo de molécula diatómica se tiene cuando se considera, en la resolución de la ecuación de Schrödinger, al potencial de un oscilador anarmónico. En este caso es

$$G(v) = w_e \left(v + \frac{1}{2}\right) - X_e w_e \left(v + \frac{1}{2}\right)^2 ,$$

en donde X_e es la constante de anarmonicidad.

Se ve entonces que si v aumenta, la separación entre los niveles vibracionales disminuye. Las reglas de selección en este caso son

$$\Delta v = \pm 1, \pm 2, \pm 3, \dots .$$

Energía electrónica

Si por un instante consideramos fijos a los núcleos, entonces la energía electrónica dependerá de la distancia internuclear. Por lo tanto, para cambiar la posición de los núcleos, estos deben realizar un trabajo igual a la suma de la energía electrónica más la fuerza de repulsión coulombiana. En otras palabras: la suma de la energía electrónica y el potencial de Coulomb de los núcleos actúa como energía potencial bajo cuya influencia los núcleos llevan a cabo sus vibraciones.

Las curvas que representan esta energía potencial son conocidas como curvas de potencial y cada estado electrónico está caracterizado por una curva de potencial, la que tiene una forma como la mostrada en la Figura 7, en donde r es la distancia internuclear y r_0 corresponde a la posición de equilibrio.

Los estados electrónicos son designados por las letras X, A, B, ... , en donde X se refiere al estado fundamental. Además se utiliza el valor del momento angular que representa la componente del momento angular orbital electrónico a lo largo del eje internuclear. El mismo puede tomar los valores 0, 1, 2, ... , los que son designados por $\Sigma, \Pi, \Delta, \Phi, \dots$,

análogamente a la designación para los átomos.

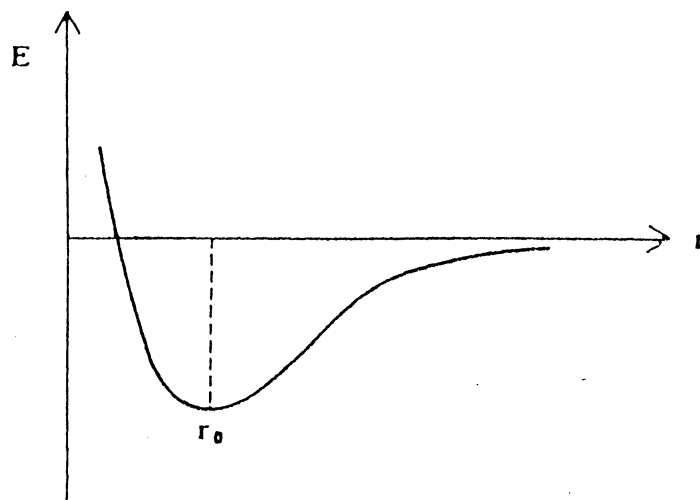


Figura 7

Espectros de moléculas diatómicas

Los números de onda de las líneas espectrales correspondientes a las transiciones entre dos estados electrónicos están dados por

$$\sigma = T' - T'' = (T'_e - T''_e) + (G' - G'') + (F' - F'') = \sigma_e + \sigma_v + \sigma_r$$

en donde las letras primadas corresponden al estado superior y las doble primadas al inferior.

Debido a la estructura rotacional, cada transición vibracional es, en efecto, una banda, o un grupo de líneas espaciadas a intervalos aproximadamente iguales a $2B$ sobre cada lado del origen de bandas σ_0 . Líneas para las cuales $J' = J'' + 1$, dando lugar a la denominada rama R, mientras que aquellas para las que $J' = J'' - 1$, corresponden a la rama P. La transición $J'=0 \leftrightarrow J''=0$ está prohibida. La Figura 8 muestra un diagrama de los niveles involucrados en un espectro de vibración-rotación, el cual es mostrado también en la misma figura.

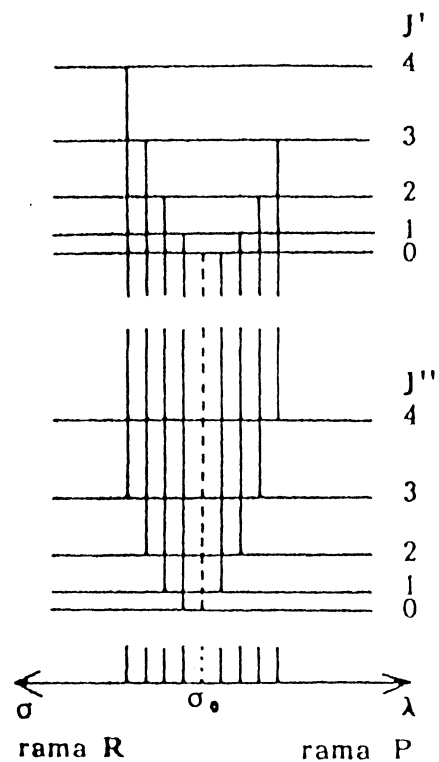


Figura 8

El espectro electrónico toma la forma de un sistema de bandas. La Figura 9 muestra esquemáticamente parte de uno de tales sistemas entre dos estados electrónicos X y A. Cada par de niveles vibracionales v' y v'' puede dar lugar a una banda, y cada banda está compuesta de transiciones entre diferentes niveles rotacionales $J'-J''$. Para transiciones electrónicas no hay restricciones sobre Δv , de manera que en la figura las flechas representan los orígenes de banda, correspondientes a líneas verticales en el esquema del espectro dibujado abajo.

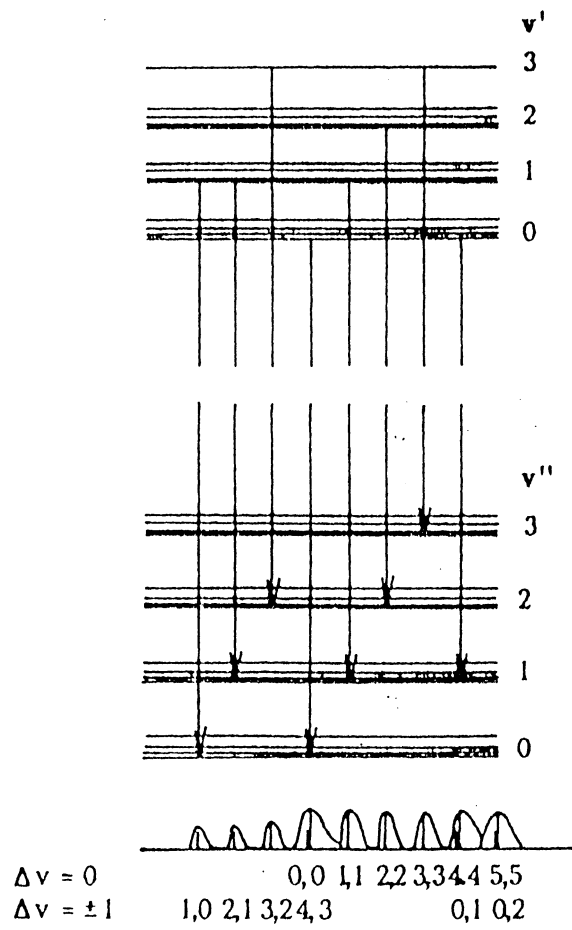


Figura 9

Sistemas espectroscópicos

La Figura 10 muestra los elementos básicos que posee un sistema espectroscópico utilizado en estudios de emisión o absorción de la radiación electromagnética.

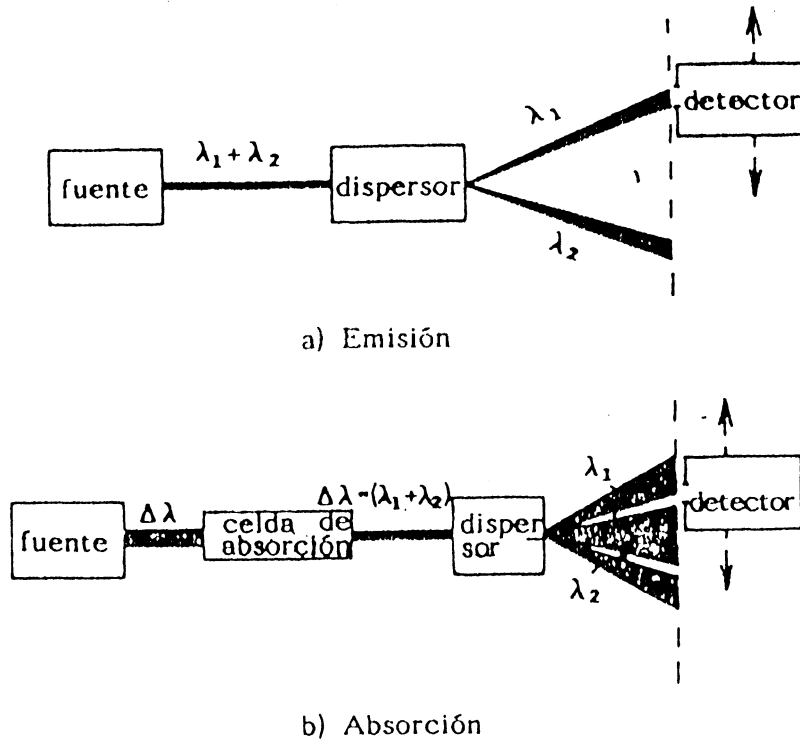


Figura 10

La fuente luminosa puede ser de diversos tipos, respondiendo a distintas geometrías, formas de excitación y al tipo de elemento bajo estudio.

El detector puede ser la placa fotográfica o un detector fotoeléctrico u otro dispositivo que registre la intensidad de la luz incidente como función de la posición.

El dispersor puede ser clasificado de acuerdo al dispositivo utilizado, y entonces hablaremos de instrumentos de prisma, instrumentos de red de difracción, e interferómetros. En nuestro caso distinguiremos fundamentalmente los instrumentos que utilizan el prisma y los instrumentos que utilizan red de difracción como elemento dispersor.

Cabe agregarse aquí que, utilizando los términos apropiadamente, uno debería mirar a través de un espectroscopio, tomar fotografías con

un espectrógrafo, medir (longitudes de onda e intensidades) con un espectrómetro, y aislar luz de una determinada longitud de onda con un monocromador. Sin embargo, el término espectrógrafo es a menudo usado y aceptado como nombre genérico para este tipo de instrumentos.

Existen varios montajes diferentes de espectrógrafos y en las Figuras 11 y 12 se muestran dos montajes típicos, el primero correspondiente al denominado montaje Babinet-Bunsen para instrumentos con prisma como elemento dispersor, y el segundo, correspondiente al montaje Ebert, propio de instrumentos que utilizan una red plana de difracción, como elemento dispersor.

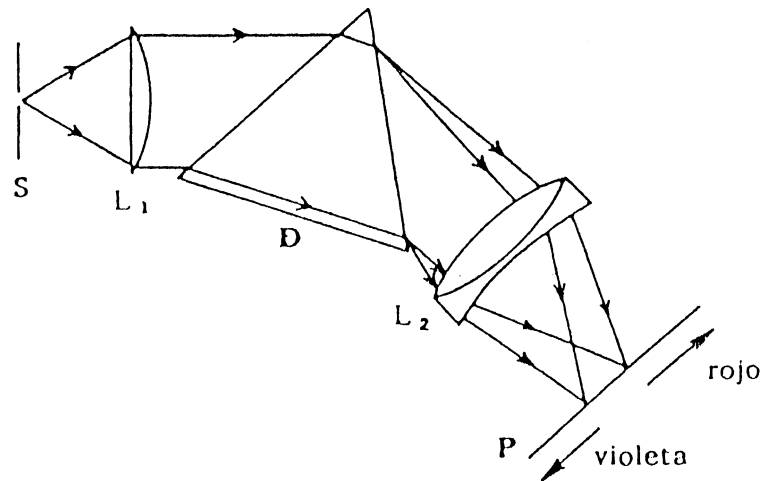
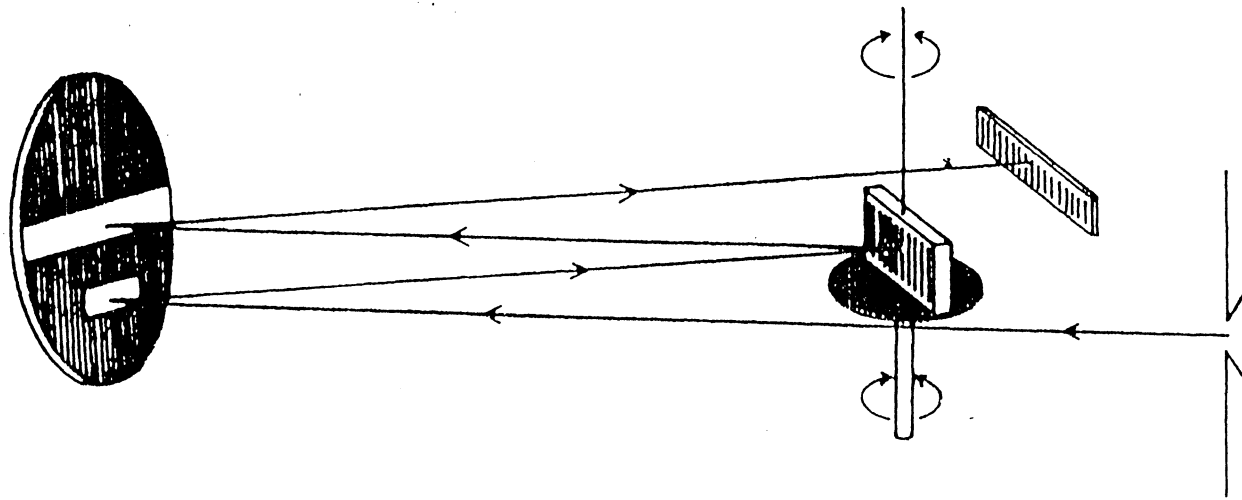
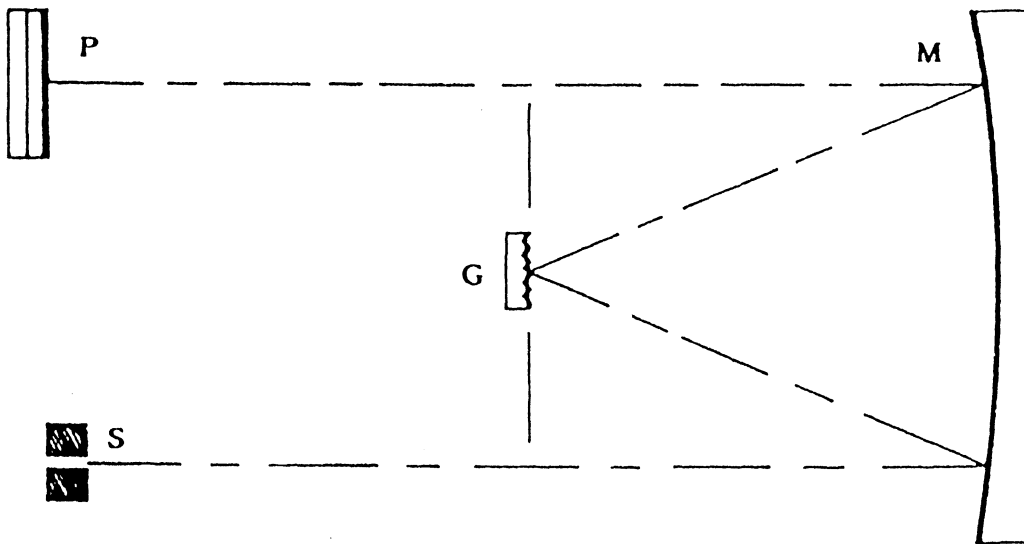


Figura 11. Componentes de un espectrógrafo de prisma convencional. S es la ranura, L_1 y L_2 las lentes colimadoras y de cámara, respectivamente, D el prisma y P el plano focal de L_2 .



a)



b) P: placa. M: espejo. G: red de difracción. S: rendija de entrada.

Figura 12

Es interesante mencionar que en las regiones del espectro correspondientes al infrarrojo, así como por debajo de los 200 nm, los dispositivos experimentales y también los montajes de los espectrógrafos, tienen algunas características particulares, a fin de mejorar la detección de los espectros en dichas regiones. En la bibliografía que se adjunta al final del capítulo el lector puede ampliar más sobre este punto.

Unidades

Las longitudes de onda en el rango óptico del espectro electromagnético son expresadas actualmente en nanómetros ($1 \text{ nm} = 10^{-9} \text{ m}$), pero el Ångström (Å) es aún una unidad reconocida y es una de las más encontradas en la literatura espectroscópica.

Dado que $1 \text{ Å} = 10^{-10} \text{ m}$, $1 \text{ nm} = 10 \text{ Å}$.

A pesar de la clasificación común de la radiación por longitudes de onda, la cantidad importante desde el punto de vista de la estructura atómica y molecular, es la frecuencia, $f = c/\lambda_{\text{vacío}}$, porque está relacionada con la diferencia de energía ΔE entre dos estados estacionarios de un sistema por $hf = \Delta E$.

A longitudes de onda cortas, es más conveniente reemplazar f por una unidad proporcional a ella: el número de onda $\sigma = f/c = 1/\lambda_{\text{vacío}}$.

La unidad del número de onda es el cm^{-1} , llamado también Kayser (K) . A veces es usado un submúltiplo del mismo, el miliKayser (mK) .

Niveles de energía atómicos y moleculares, al igual que los números de onda, son usualmente expresados en cm^{-1} , aunque en algunos casos, y para los primeros, es utilizado el electrón-volt (eV) .

En los estudios de espectroscopía óptica de alta resolución, es común la utilización del MHz ($1 \text{ mK} = 29.979 \text{ MHz}$) .

Las relaciones entre todas estas unidades surgen de considerar las ecuaciones que las ligan y el valor de las constantes. O sea:

$$\sigma(\text{cm}^{-1}) = 1/100.\lambda_{\text{vacío}}(\text{m}) = f/100.c = \Delta E(\text{Joule})/100.h.c \quad .$$

La siguiente tabla muestra relaciones entre estas unidades:

	Joule	eV	erg	cm
1 Joule	1	$6,242 \times 10^{18}$	10^7	$5,034 \times 10^{22}$
1 eV	$1,602 \times 10^{-19}$	1	$1,602 \times 10^{-12}$	8065
1 erg	10^{-7}	$6,242 \times 10^{11}$	1	$5,034 \times 10^{15}$
1 cm^{-1}	$1,986 \times 10^{-23}$	$1,240 \times 10^{-8}$	$1,986 \times 10^{-16}$	1

Bibliografía

- "Introduction to Atomic Spectra", H.E. White, McGraw-Hill (1934).
- "Atomic Spectra and Atomic Structure". G. Herzberg, Dover (1944).
- "Spectrophysics", A.P. Thorne, Science Paperbacks (1974).
- "Experimental Spectroscopy", R.A. Sawyer, Dover (1963).
- "Técnica y Práctica de Espectroscopía", A.N. Zaidel, G.V. Ostrovskaya, y Yu.I. Ostrovski, MIR (1979).
- "The Spectroscope", K.I. Tarasov, Adam Hilger (1974).
- "Atomic Spectra", H.G. Kuhn, Longmans (1969).
- "Molecular Spectra and Molecular Structure", G. Herzberg, Van Nostrand (1950).
- "Introduction to Molecular Spectra", G.M. Barrow, McGraw-Hill (1962).
- "Current Trends in Atomic Spectroscopy", Report on a Workshop Held in Tucson, Arizona, October 24-26, 1982. National Academy Press, Washington, D.C. (1984).
- B.C. Fawcett, J. Opt. Soc. Am., B 1, 195 (1984).

LOI 1996

ESPECTROSCOPIA OPTICA APLICADA AL MONITOREO DE CONTAMINANTES ATMOSFERICOS

Dr. Jorge Reyna Almandos

INTRODUCCION

Se puede definir la contaminación del aire como la presencia en la atmósfera exterior de uno o más contaminantes o sus combinaciones, en cantidades tales y con tal duración que sean o puedan afectar la vida humana, de animales, de plantas, o de la propiedad, que interfiera el goce de la vida, bienes de uso o el ejercicio de las actividades.

Esta contaminación puede incluir la presencia en el aire de sustancias abióticas o bióticas (en estado gaseoso, líquido o sólido o una combinación de ellas) o de formas de energía (radiaciones, ruido, etc) en cantidad y tiempo suficientes para provocar daños en el ecosistema.

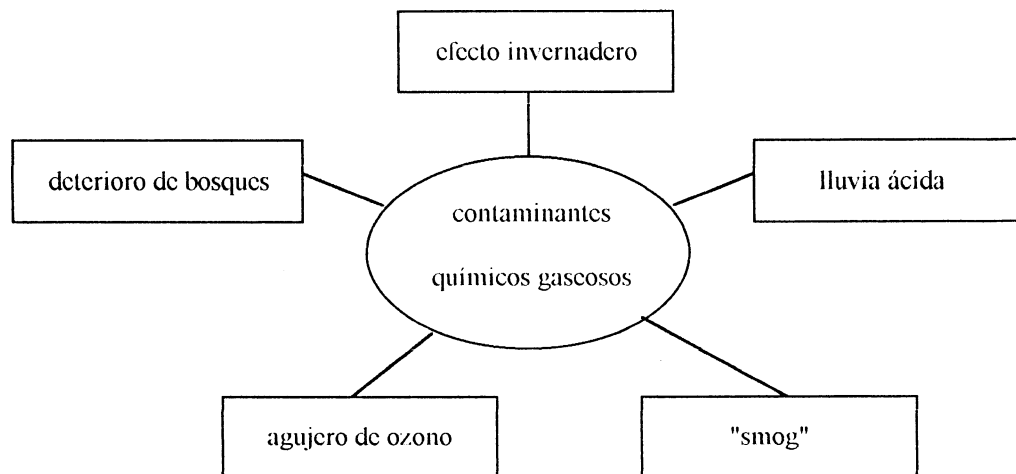
Aplicando esta definición a una zona determinada se puede observar que la misma queda condicionada por los siguientes elementos:

- fuentes de emisión de los contaminantes
- condiciones meteorológicas que determinen la dispersión de dichos contaminantes
- receptores animados o inanimados sobre los cuales inciden y afectan los contaminantes.

Restringiendo esta exposición a los contaminantes químicos del aire, estos son emitidos por fuentes puntuales y distribuidas, e inician su ciclo de estancia aérea con la emisión seguido de su transporte y difusión en la atmósfera. El ciclo se completa cuando los contaminantes pudiéndose transformar en otros o no, se depositan sobre la vegetación, los animales, la superficie del suelo y cuerpos de agua y otros objetos; cuando son arrastrados por lluvia o cuando se escapan al espacio. En algunos casos los contaminantes depositados pueden volver a introducirse en la atmósfera por la acción del viento.

De lo expuesto arriba puede observarse que en cuanto a contaminación del aire se refiere los parámetros meteorológicos del lugar juegan un papel muy importante.

Los principales efectos de la contaminación atmosférica actual se pueden ver en el siguiente esquema:



La toma de conciencia de la contaminación atmosférica como un fenómeno global a nivel planetario y su consecuente traslado a los sistemas normativos ha desembocado tanto en la mejora o sustitución de procesos como en el desarrollo de métodos de detección más eficientes.

Se indicaran a continuación algunos de los contaminantes gaseosos que se consideran más representativos en zonas industriales o grandes ciudades

Dióxido de azufre (SO₂)

Por ser un contaminante procedente en una gran proporción de los procesos actuales de combustión, se lo ha considerado como indicador del estado general de contaminación de un área. Los derivados del petróleo, principalmente las gasolinas, contienen cantidades proporcionales de SO₂. Vemos así que los principales focos emisores de SO₂ serán las calefacciones domésticas y los quemadores industriales que empleen carbón bituminoso, las centrales térmicas, las industrias petroquímicas, derivados del petróleo, productoras de ácido sulfúrico, etc. Este contaminante es un gas incoloro, que produce, en elevadas proporciones, un efecto irritante en la vista y en las vías respiratorias y que combinado con el vapor de agua o una atmósfera saturada, da lugar a la formación de ácido sulfúrico, el que se precipita, con todas las consecuencias sobre personas, tierras y bienes en uso.

Oxidos de nitrógeno (NO_x)

Los dos más característicos son el NO y NO₂, midiéndose generalmente su suma y obteniéndose una medida de los óxidos totales NO_x. Los focos emisores de estos gases son los escapes de vehículos automotores, procesos de combustión en la industria del acero, petroquímicas, centrales termoeléctricas, etc. El gran peligro que presentan los óxidos de nitrógeno está vinculado con la acción de la luz solar que incide sobre una atmósfera muy contaminada, dando lugar a procesos fotoquímicos con la producción de contaminantes secundarios de mayor toxicidad, como son el ozono (O₃) y el nitrato de peroxibenzoilo (PBzN), éste último sospechoso de ser un agente cancerígeno.

Es un contaminante que, emitido por chimeneas industriales, tiene un color rojizo, y al descargarse en una atmósfera saturada de vapor de agua puede dar lugar a la formación de ácido nítrico, muy perjudicial para el suelo y el agua.

Monóxido de carbono (CO)

Es, como el SO₂, un contaminante muy generalizado, producido principalmente por los procesos de combustión de vehículos automotores. En forma menos frecuente se lo puede encontrar en emanaciones de minas carboníferas y emanaciones volcánicas, y en plantas metalúrgicas en ciertos procesos de tratamiento del acero.

Es un gas incoloro e inodoro, muy tóxico para las personas, por sus efectos directos sobre los sistemas circulatorio y respiratorio.

Ozono (O₃)

Es un oxidante de origen fotoquímico, producido por la acción de la luz solar al incidir sobre las capas de la baja atmósfera terrestre. El ozono es considerado un contaminante que puede ser muy peligroso en altas concentraciones, produciendo una calcificación acelerada en los huesos del cuerpo humano, afectando asimismo la visión, sistemas respiratorio y circulatorio, llegando a lesionar, en dosis altas y continuadas, al sistema nervioso central.

Hidrocarburos

Corresponde a la familia de compuestos que contienen carbono e hidrógeno, tales como metano y acetileno, de los cuales los hidrocarburos no saturados corresponden al grupo más peligroso, por su facilidad para reaccionar con la radiación solar, originando el conocido smog fotoquímico.

Los vehículos automotores, sobre todo aquellos con el motor fuera de punto o muy gastado, son los principales contribuyentes a la contaminación por hidrocarburos.

Material particulado

Comprenden los polvos, humos, nieblas y neblinas, sistemas que en buena parte se hallan comprendidos dentro de la denominación general de aerosoles. También el material particulado se origina desde fuentes naturales continentales como el polvo levantado por tormentas, erupciones volcánicas y otras.

Su influencia se manifiesta en la pérdida de visibilidad del área, la generación de enfermedades o la predisposición a ellas, y el impacto sobre el medio ambiente construido.

Los efectos adversos de los contaminantes son prácticamente evidentes en los grandes centros urbanos, donde innumerables fuentes contribuyen a las emisiones. Como consecuencia de la mezcla en la atmósfera, los efectos de cada una de las emisiones en particular no son identificables. Por otra parte, es frecuente el "sinergismo", término que indica que los efectos observados por la acción simultánea de dos a más contaminantes son de mayor intensidad que la suma de los efectos producidos por los contaminantes cuando actúan por separado. También se presenta sinergismo cuando un contaminante sufre transformaciones en la atmósfera y genera otros con mayores efectos adversos que el inicial.

METODOS DE DETECCION

La atmósfera tiene una constitución química compleja a lo que se le suma la variación de las condiciones meteorológicas con el lugar y el tiempo. Esto hace que definir un método que cubra todas las necesidades sea un verdadero desafío:

Según M.A. Sigrist (1), una buena técnica de detección debe cumplir con los siguientes requisitos en mayor o menor grado dependiendo del caso particular:

- Factibilidad de detectar un gran número de compuestos con un mismo instrumento
- Alta sensibilidad para permitir la detección de bajas concentraciones
- Alta selectividad para diferenciar entre los diferentes componentes en mezclas multicomponentes complejas
- Rango de trabajo grande para poder monitorear bajas y altas concentraciones con un mismo aparato
- Buena resolución temporal para poder trabajar *on line*
- Fácil de transportar para permitir realizar mediciones *in situ*

Las diferentes técnicas de monitoreo se pueden agrupar de distintas maneras, por ejemplo, según el método de muestreo o el método analítico empleado.

Según el primer método la muestra de aire puede ser efectuada por

- a) método extractivo
- b) método no extractivo

Los métodos extractivos están basados en la colección de la muestra de aire en un recipiente adecuado y su posterior análisis en el laboratorio. Tienen la ventaja es que se puede analizar la muestra con alta sensibilidad y selectividad pero se corre riesgo de representatividad de la misma y no se tiene un monitoreo en tiempo real (medidas en el tiempo y momento de muestreo), ni monitoreo continuo.

Los métodos no extractivos pueden ser analizando la muestra que está siendo tomada en ese momento *in situ* o a través de un dispositivo remoto donde el análisis se produce a una distancia considerable de donde se está tomando la muestra. Ambos ofrecen la ventaja de trabajar en tiempo real si bien se puede perder algo de sensibilidad o selectividad frente a las técnicas extractivas.

En el caso de dispositivos que trabajan con una considerable porción de aire como muestra, las medidas dependen fuertemente de las condiciones de dispersión atmosférica .

Según el método analítico empleado las técnicas de monitoreo podrían ser agrupadas en espectroscópicas y no espectroscópicas.

Hay una gran variedad de técnicas no espectroscópicas para la medición de contaminantes del aire: cromatografía gaseosa (CG) (de muy buena sensibilidad y selectividad), espectrometría de masa (buena sensibilidad y selectividad), coulombimetría (buena sensibilidad), colorimetría (buena sensibilidad), conductividad electrolítica (buena sensibilidad), quimiluminiscencia (buena sensibilidad), detección térmica, etc.

La CG combinada con detector de ionización de llama (FID) o con espectroscopía de masa (MS) es una herramienta muy poderosa para el monitoreo de contaminantes del aire. A excepción de la CG los restantes métodos no espectroscópicos son solamente adecuados en la práctica para la detección de una especie simple.

TECNICAS ESPECTROSCOPICAS

En los últimos años las técnicas espectroscópicas han atraído mucho interés debido a que ellas ofrecen características únicas que las vuelven herramientas muy versátiles para el estudio de gases contaminantes.

Los métodos espectroscópicos pueden ser separados en aquellos que usan fuentes de luz convencionales y aquellos que usan dispositivos con láseres, si bien la base de todas las técnicas espectroscópicas es el mismo tipo de interacción entre radiación y materia.

Veamos ahora como la espectroscopía puede proveer información sobre el medioambiente haciendo medidas a distancia. Mediciones remotas (Remote sensing), es un término general utilizado para técnicas, por medio de las cuales la información sobre condiciones físicas o químicas en una posición espacial distante, puede ser obtenida usando el equipamiento adecuado. El principio de funcionamiento se basa en considerar que por medio de ondas electromagnéticas el equipo de medición remota (generalmente llamado el sensor), está en contacto con el elemento de volumen a medir, el que está caracterizado por ciertos parámetros físicos o químicos. El proceso de medición involucra un análisis de la radiación que alcanza el sensor desde el objeto y que trae información sobre los valores de los parámetros analizados.

Los sistemas de medición remota pueden tener diferentes montajes dependiendo del tipo de medición a llevar a cabo. Montajes sobre satélites, aviones, y sistemas fijos utilizados para el monitoreo urbano de áreas industriales constituyen algunos ejemplos.

Vamos a distinguir a aquellos sistemas que utilizan espectroscopía de absorción de los que utilizan espectroscopía de emisión.

En el primer caso mencionaremos el sistema experimental basado en la espectroscopía de absorción óptica diferencial (DOAS), que permite detectar y medir contaminantes atmosféricos como el dióxido de azufre, óxidos de nitrógeno, monóxido de carbono, hidrocarburos, etc.

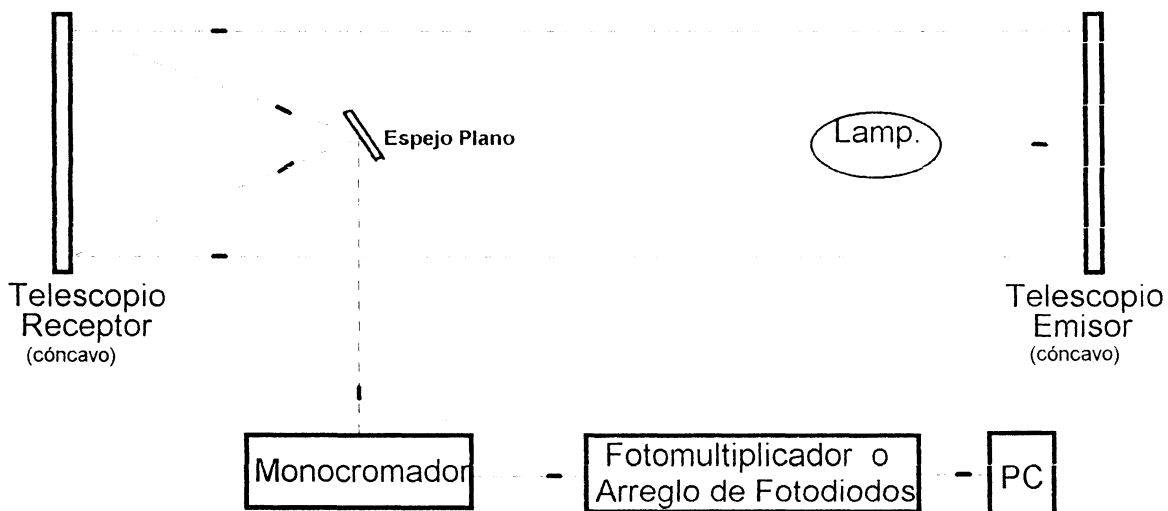
La medición de contaminantes por la técnica arriba citada se basa en la absorción selectiva de las moléculas contaminantes cuando son iluminadas con una fuente de luz de espectro continuo. Considerando que si se conoce la intensidad de radiación emitida (I_0) y la recibida (I), luego de recorrer un camino óptico (x), es posible determinar la concentración promedio del contaminante (K), la que está relacionada con la expresión:

$$I/I_0 = \exp(-Kx)$$

En nuestro caso el sistema consta de un telescopio emisor que tiene un espejo paraboidal de aproximadamente 15 cm de distancia focal y 30 cm de diámetro. En el foco del mismo se

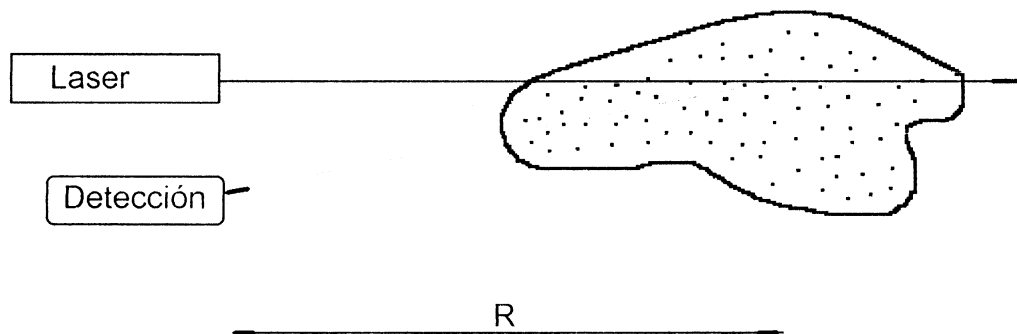
coloca una lámpara de xenón de alta presión, que posee un espectro muy rico en frecuencias y que cubre la región del ultravioleta cercano al límite de absorción del aire (aproximadamente 200 nm) y hasta el infrarrojo cercano del espectro electromagnético.

La luz emitida por la lámpara, en este caso de 150 W, es recibida por un sistema óptico reflector tipo telescopio newtoniano de 1,5 m de distancia focal, que posee un espejo paraboidal de 30 cm de diámetro y un espejo plano de 7,5 cm de diámetro. La luz recibida por este dispositivo es analizada con un monocromador de barrido, montaje Czerny-Turner, de 50 cm de distancia focal. A la salida de este la luz es detectada por un fotomultiplicador, generando una corriente proporcional a la intensidad de la luz, la que amplificada es a su vez tomada por una computadora personal (PC) a través de una tarjeta digitalizadora .



TECNICAS LIDAR

En el caso en que estemos utilizando como fuente luminosa luz proveniente de un laser, podremos utilizar la técnica LIDAR. Esta técnica, cuyo nombre resulta un acrónimo de Light Detection and Ranging, utiliza pulsos de radiación laser que son transmitidos en la atmósfera y la luz retrodispersada es colectada por un telescopio y detectada como se muestra en la figura:



En el principio de funcionamiento del LIDAR la luz del laser que es retrodispersada desde una distancia R arriba al receptor después de un tiempo $t = 2R/c$ (en donde c es la velocidad de la luz) luego de la transmisión del pulso. Información resuelta en distancia puede entonces ser obtenida del corrimiento en el tiempo.

Una característica particularmente útil del método LIDAR (también llamado radar laser), es su capacidad de monitorear grandes áreas. Por supuesto que el tamaño del área cubierta depende del tipo de dispositivo sobre el que hemos montado todo el sistema LIDAR. Desde un laboratorio fijo un área industrial o una sección de una ciudad puede ser cubierta o barrida. Un sistema móvil puede ser utilizado para cubrir medidas en el campo o en áreas urbanas industriales. Sistemas montados sobre aviones pueden ser utilizados exitosamente para medidas que cubran amplias regiones. Finalmente, los satélites pueden resultar plataformas para sistemas LIDAR espaciales que se utilicen en estudios globales de nuestro planeta.

Si la longitud de onda del laser es variada desde una línea de absorción de un gas contaminante a otra, cercana a la anterior, el cambio en la detección de la intensidad de la luz retrodispersada puede ser utilizado para evaluar la concentración estudiada. Esto se conoce como técnica DIAL (Differential Absorption Lidar).

La misma permite el monitoreo remoto del aire ambiente, de emisiones industriales o naturales debidas a varios fenómenos geofísicos (como volcanes o otras fuentes geotérmicas) en donde los estudios también pueden ser de tipo troposférico o estratosférico.

En el caso de sistemas que utilizan métodos de espectroscopía de emisión, existen dispositivos comerciales que trabajan en forma no dispersiva utilizando filtros para seleccionar las regiones espectrales de interés. Por ejemplo, un analizador de fluorescencia ultravioleta diseñado para medir bajas concentraciones de dióxido de azufre aprovecha que el SO_2 tiene una fuerte absorción en el espectro ultravioleta entre 200 y 240 nm. La absorción a esta longitud de onda da como resultado la emisión de fluorescencia en longitudes de onda alrededor de 350 nm y la cantidad de fluorescencia emitida es directamente proporcional a la concentración de dióxido de azufre.

Dispositivos experimentales como este último y otros con algunas modificaciones que hacen uso de técnicas espectroscópicas en distintas regiones del espectro electromagnético, abundan en el mercado comercial debido a sus vastas aplicaciones para el monitoreo de diversos gases presentes en emisiones de tipo industrial y otras.

BIBLIOGRAFIA

- (1) Markus A. Sigrist, "Air Monitoring by Spectroscopic Techniques"
1994, JOHN WILLEY & SONS INC.
- (2) Sune Svanberg, "Atomic and Molecular Spectroscopy"
1992, SPRINGLER VERLAG.
- (3) Anne P. Thorne, "Spectrophysics"
1988, CHAPMAN AND HALL.
- (4) Kennet Wark and Cecil F. Warner, "Contaminación del Aire"
1990, LIMUSA GRUPO NORIEGA
- (5) J.A.del Giorgio, "Contaminación Atmosférica"
1977, ALHAMBRA
- (6) Manual de Instalación y Uso del Equipo ML 9800
1992, LEGAR SIEGLER, Measurment Controls Corporation
- (7) Juan Morettón, "Contaminación del Aire en la Argentina"
1996, EDICIONES UNIVERSO
- (8) H.Sandoval L., M. Prendez B., P. Ulriksen U. "Contaminación Atmosférica de Santiago"
1993, EDITORES

FIBRAS OPTICAS PARA COMUNICACIONES

UNA BREVE RESEÑA

Dr. Daniel Schinca

**Curso Laser y Optica en Ingenieria
Centro de Investigaciones Opticas 1995**

I. FUNDAMENTOS

Por más de un siglo, las ondas electromagnéticas han demostrado su utilidad para transmisión de información, principalmente debido al hecho de que no resulta necesario un conductor metálico para su propagación. De hecho, pueden propagarse en el vacío o por medios dieléctricos, es decir, en materiales no conductores. Dentro del ancho espectro electromagnético, la luz visible ocupa un muy angosto rango entre 400 nm (violeta) y 750 nm (rojo). Los intentos de transmitir información modulando la intensidad de una fuente luminosa se remontan a más de 200 años, contándose el telégrafo óptico del francés C. Chappe y el fotófono del estadounidense A.G.Bell. En 1870, el físico inglés J.Tyndall demostró que la luz podía ser guiada a través de un chorro de agua. Basado en este hecho y en las experiencias de Bell, el estadounidense N.R.French describe en 1934 un sistema por medio del cual una señal de audio podía transmitirse por un llamado cable óptico, consistente en una varilla sólida de vidrio de baja atenuación, utilizando el principio de reflexión total interna de la luz. Desde esa fecha hasta nuestros días se han realizado trascendentales avances tanto en el área específica de las fibras ópticas como en el de las fuentes emisoras y detectores que han permitido la implementación de este sistema de comunicaciones tanto a nivel local como internacional.

Es sabido que la luz (al igual que cualquier onda electromagnética) se propaga en el vacío a una velocidad constante de muy aproximadamente 300.000 Km/s (c_0). Sin embargo, cuando la luz se propaga por un medio material (aire, agua, vidrios, plásticos transparentes, etc), se observa que lo hace a una velocidad menor (v), distinta para cada material. Al cociente entre la velocidad de la luz en el vacío y la velocidad de la luz en cierto material se lo llama *índice de refracción* de ese material:

$$n = \frac{c_0}{v}$$

Cuando la luz llega a la interfaz entre dos medios con distintos índices de refracción, un cierto porcentaje se refleja mientras que el resto continúa propagándose en el nuevo medio, aunque en una dirección diferente, tal como puede verse en la figura 1.

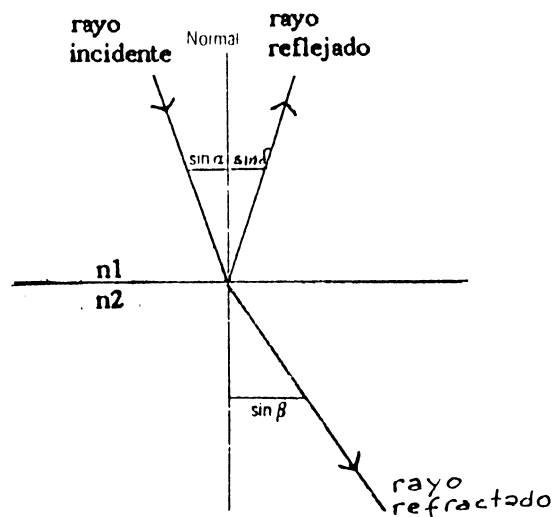


fig.1

El término "rayo de luz" se utiliza aquí para describir el camino que siguen los frentes de onda de la luz al propagarse por los distintos medios. Es posible verificar que se cumplen las siguientes relaciones entre el rayo reflejado y el refractado:

$$\alpha = \delta$$

$$\frac{\text{sen } \alpha}{\text{sen } \beta} = \frac{n_2}{n_1}$$

El índice de refracción no solo depende del materia sino que, para un cierto medio, es función de la longitud de onda de la luz que lo atraviese. En la figura 2 puede verse $n(\lambda)$ para el caso de sílica pura (SiO_2).

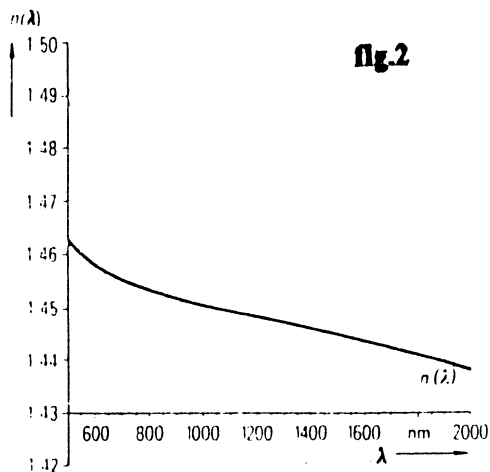
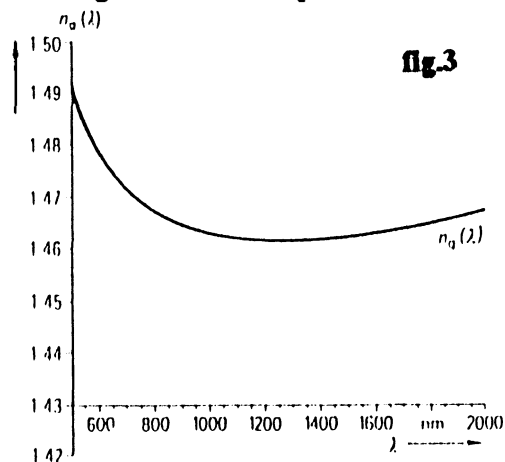


fig.2

Este índice se aplica para luz de determinada longitud de onda y de amplitud constante y por lo tanto incapaces de transmitir información alguna. Para hacerlo es necesario modular esta onda portadora (transmisión analógica) o utilizar trenes de pulsos de luz (transmisión digital). En ambos casos se generan grupos mas o menos angostos de ondas que contienen varias frecuencias o longitudes de onda (paquetes de onda). Dentro de éstos, las distintas ondas componentes se desplazan a distintas velocidades debido a la dependencia de n con λ y a su vez, el paquete se desplaza con la llamada *velocidad de grupo* a la que se asocia un *índice de refracción de grupo* que viene dado por:

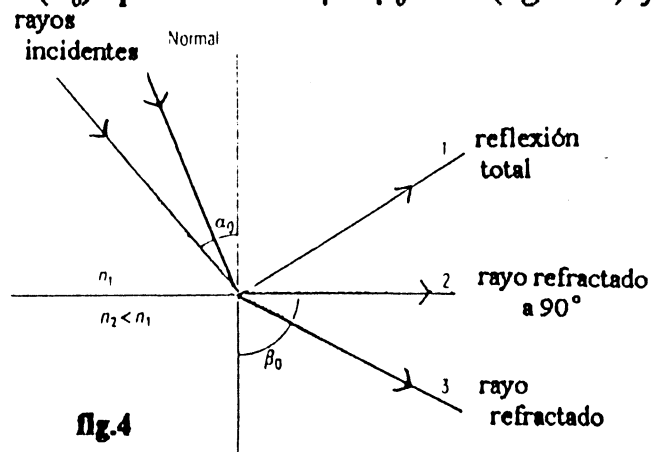
$$n_g = n - \lambda \frac{dn}{d\lambda},$$

cuya dependencia con la longitud de onda puede verse en la figura 3:



Cuando una onda que viaja por un medio de índice n_1 llega a la interfaz con un medio de índice n_2 menor, el rayo refractado se aleja de la normal. A medida que el ángulo de incidencia α aumenta, también lo hace el de refracción β . Existirá así un *ángulo de incidencia crítico* (α_0) para el cual $\beta = \beta_0 = 90^\circ$ (figura 4) y se verifica que:

$$\sin \alpha_0 = \frac{n_2}{n_1}$$



Para todos los rayos con ángulo de incidencia mayor que el crítico no existe rayo refractado en el medio ópticamente menos denso. A este fenómeno se lo conoce como **reflexión total interna** y es el mecanismo en el que se basa la transmisión de luz por fibras ópticas. En éstas, el efecto de reflexión total interna se logra encerrando a un núcleo de vidrio (core) de índice n_1 con un recubrimiento de vidrio (cladding) de índice ligeramente menor n_2 . De la condición antedicha que $\text{sen } \alpha_0 = n_2/n_1$, se sigue que los rayos que no diverjan del eje de la fibra en más de un ángulo de $(90^\circ - \alpha_0)$ serán guiados dentro del núcleo (figura 5).

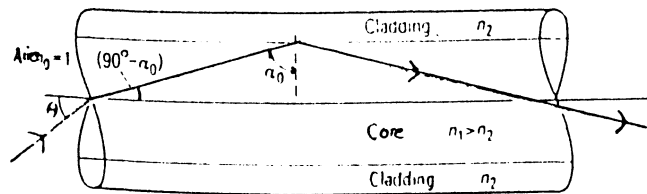


fig.5

Para lograr este guiado, es necesario introducir luz en el extremo de la fibra con un ángulo θ tal que se cumpla:

$$\text{sen } \theta = \sqrt{n_1^2 - n_2^2}$$

El máximo ángulo de entrada θ_{max} se llama **ángulo de aceptación** y el seno del mismo se llama **apertura numérica** : $\text{NA} = \text{sen } \theta_{\text{max}}$.

Debido a las condiciones de contorno que la teoría electromagnética impone a los campos eléctrico y magnético en la interfaz entre dos medios, es posible demostrar (aunque no lo haremos aquí) que solo se permiten ciertas formas de propagación (**modos**) que poseen la característica de mantener constante la distribución transversal de intensidad a lo largo de la longitud de la fibra. Algunos modos típicos en fibras se muestran en la figura 6. Puede verse que los modos de orden más alto tienden a distribuirse más hacia la periferia del núcleo y por lo tanto tienden a penetrar más en el cladding que los modos de orden más bajo, como el LP_{01} o el LP_{02} . La terminología LP significa que las ondas planas componentes son linealmente polarizadas. El primer subíndice se llama **número modal azimutal** y equivale a la mitad de lóbulos luminosos que aparecen en cada anillo de luz. El segundo subíndice se llama **número modal radial** y equivale al número de anillos luminosos concéntricos del modo.

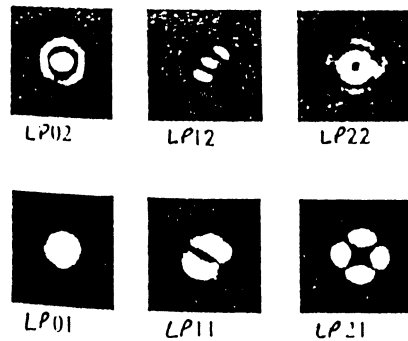


Fig.6

Es posible fabricar guías de onda óptica en las que el índice de refracción del núcleo es función de la distancia radial desde el eje de la fibra hacia la periferia ,en dirección a la cubierta.La dependencia general viene dada por:

$$n^2(r) = \left[1 - 2\Delta \left(\frac{r}{a} \right)^g \right]$$

donde n_1 es el índice en el eje de la fibra,

$\Delta = (n_1^2 - n_2^2) / 2n_1^2$ es la diferencia normalizada de índices,

r es la distancia radial a partir del eje de la fibra,

a es el radio del núcleo

g es el llamado exponente de perfil.

En la figura 7 se pueden ver algunos perfiles de índice de refracción:

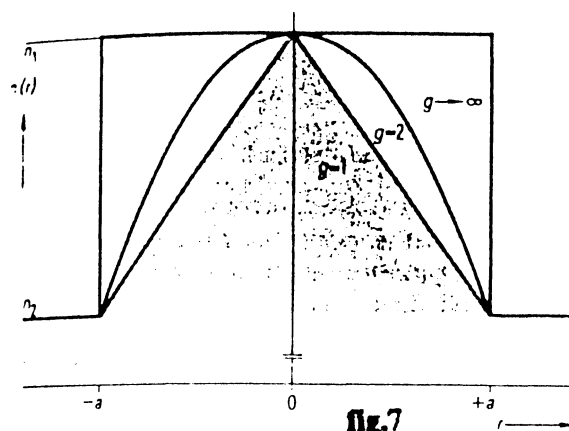


Fig.7

Los perfiles más útiles corresponden a $g \rightarrow \infty$ (*perfil de índice escalonado*) y a $g=2$ (*perfil de índice parabólico o gradual*). La propagación de los modos en una fibra depende de la forma de estos perfiles. También es posible demostrar que el número de modos guiados (N) depende del radio del núcleo, de la longitud de onda y de la apertura numérica con las siguientes relaciones:

para perfil escalonado: $N \approx V^2/2$

para perfil gradual: $N \approx V^2/4$

donde $V = 2\pi (a/\lambda) AN$

Se llaman *fibras multimodo* a aquellas que pueden guiar varios modos sin pérdidas apreciables. Estas pueden ser de perfil escalonado (figura 8.a) o de perfil gradual (figura 8.b) y pueden sostener entre 2000 y 4000 modos. En la figura pueden verse las dimensiones típicas de estas fibras y más adelante veremos algunas de sus características de transmisión. Las llamadas *fibras monomodo* solo pueden sostener el modo guiado más bajo, tienen dimensiones típicas menores que las anteriores y son las preferidas para telecomunicaciones (figura 9).

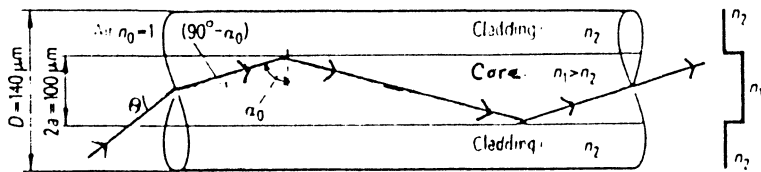


Fig.8 (a)

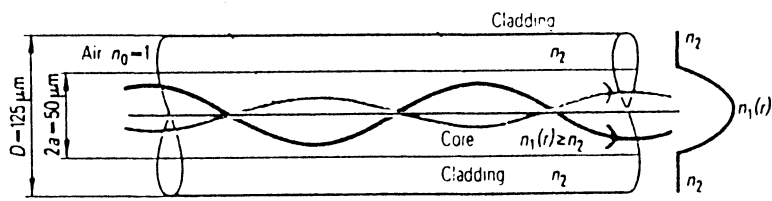


Fig 8 (b)

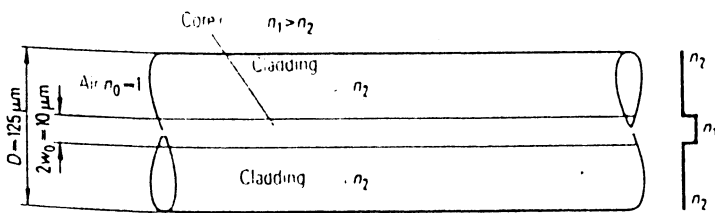


Fig.9

II . PRINCIPALES CARACTERISTICAS DE TRANSMISION

Las principales características de transmisión de fibras ópticas son la *atenuación*, la *dispersión* y el *ancho de banda*.

II . 1 ATENUACION.

La atenuación es la pérdida de señal a lo largo de la fibra. Reconoce dos causas principales: pérdidas por absorción y pérdidas por scattering.

La absorción del material está relacionada con la composición del mismo y resulta en la disipación de una parte de la potencia óptica en calor dentro de la fibra. La absorción puede ser de dos tipos: 1) la absorción intrínseca, causada por la interacción con los componentes mayoritarios de la fibra y 2) la absorción extrínseca causada por interacción con impurezas dentro de la fibra.

La primera se produce por dos mecanismos : a) estimulación de transiciones electrónicas por absorción de fotones de alta energía (UV, $0.3 \mu\text{m}$) y b) estimulación de transiciones vibracionales de la molécula de SiO_2 (B-O, P-O, Ge-O) en la región por encima de los $7 \mu\text{m}$.

Las colas de estas anchas bandas de absorción dejan una región casi libre de atenuación ($<0.1 \text{ dB/Km}$) entre 0.8 y $1.7 \mu\text{m}$, como puede verse en la figura 10.

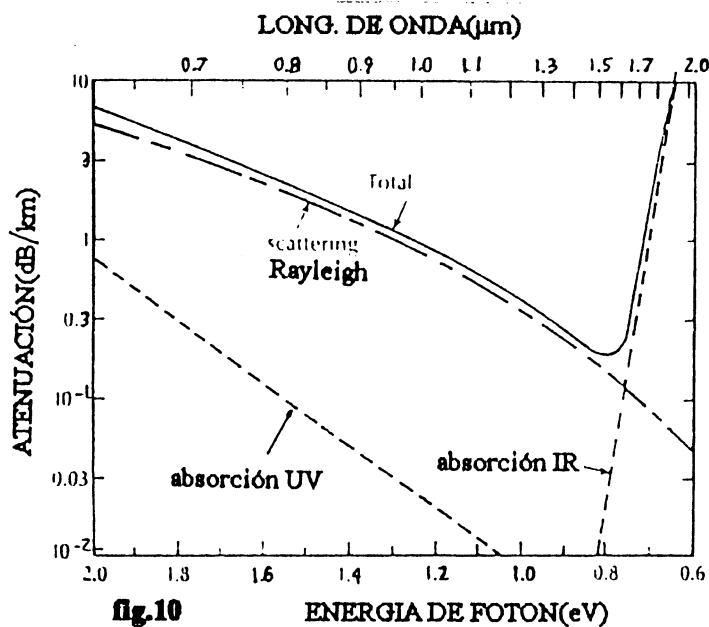


fig.10

ENERGIA DE FOTON(eV)

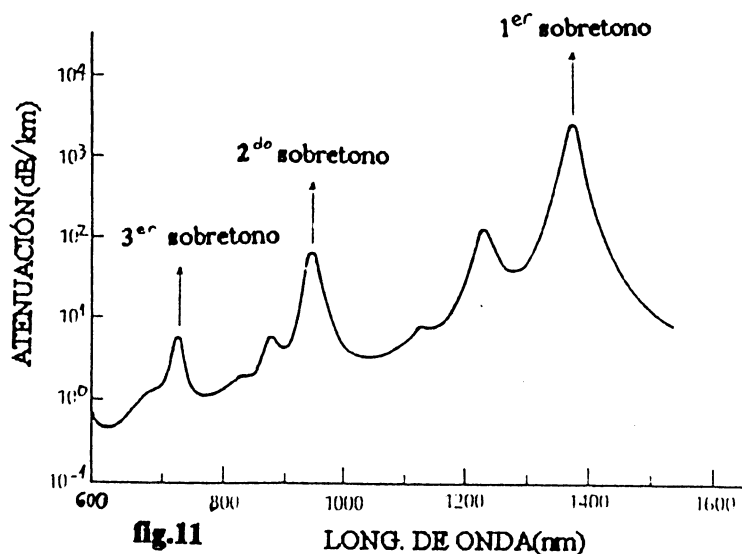
La absorción extrínseca es debida a la presencia de metales de transición como impurezas dentro de la fibra. La tabla I muestra algunas de las impurezas metálicas más comunes encontradas en fibras junto con la longitud de onda pico de absorción y la atenuación que produce una determinada concentración en unidades de ppb.

	λ pico (nm)	$1:10^9$ (dB/km)
Cr ³⁺	625	1.6
C ²⁺	685	0.1
Cu ²⁺	850	1.1
Fe ²⁺	1100	0.68
Fe ³⁺	400	0.15
Ni ²⁺	650	0.1
Mn ³⁺	460	0.2
V ⁴⁺	725	2.7

TABLA I

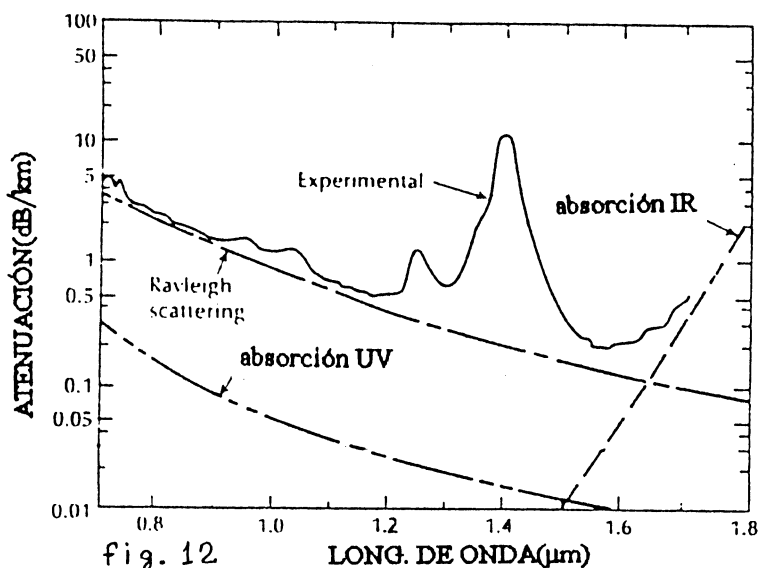
Esta atenuación puede reducirse ($<10^{-10}$) refinando las técnicas de fabricación, tal como deposición en fase de vapor, que elimina en gran parte este tipo de impurezas.

Otro mecanismo de pérdida extrínseca es la absorción del grupo oxidrilo (OH). Las vibraciones resonantes de esta molécula están en la región de 2.7 a 4.2 μm . Poseen armónicos (sobretonos) en la región visible-cercano IR. Como estas resonancias son angostas, las regiones alrededor de 1.31 y 1.55 μm presentan valles ("ventanas") que no se ven afectados por la absorción del OH ($c < 10^{-7}$). Estas bandas de absorción pueden verse en la figura 11.



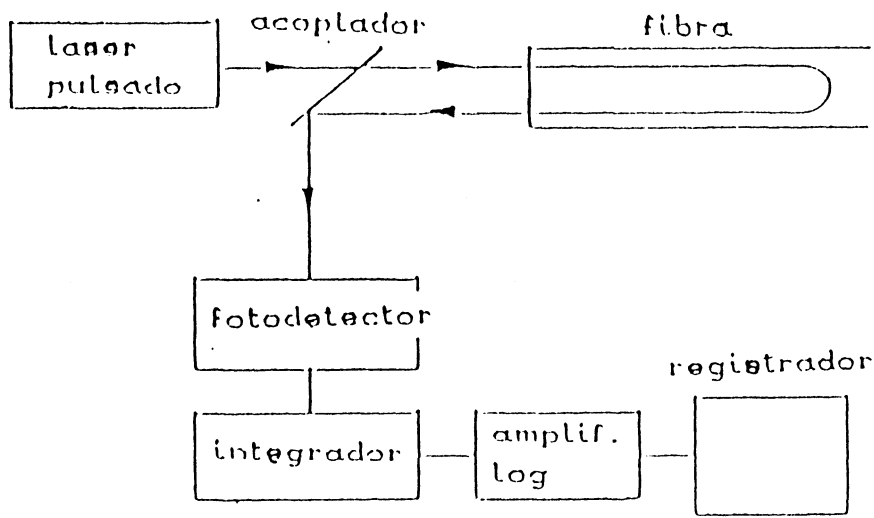
Cuando un haz luminoso alcanza un obstáculo ó inhomogeneidad presente en un medio, se reemite una onda secundaria (*scattering*) en direcciones diferentes a la de incidencia , llevándose parte de la energía incidente.

En fibras opticas existen inhomogeneidades aleatorias dadas por fluctuaciones del índice de refracción que surgen de variaciones en la composición y densidad de componentes. Hay dos tipos de scattering: tipo Rayleigh y tipo Mie. El primero ocurre cuando la dimensión de las inhomogeneidades es pequeña comparada con la longitud de onda incidente. Ocurre en todas las direcciones y la atenuación resulta proporcional a λ^{-4} . El scattering Mie ocurre cuando la dimensión de las inhomogeneidades es del orden de la longitud de onda incidente . Surge de imperfecciones en la estructura cilíndrica de la guía, irregularidades en la interfaz core-cladding, variaciones en la diferencia de índice de refracción, tensiones y burbujas. Muestra un gran lóbulo frontal y puede causar pérdidas importantes. El espectro de atenuación debido a todas las causas antes mencionadas de una fibra monomodo de baja pérdida (≈ 0.2 dB/Km) @ 1.55 μm se esquematiza en la figura 12.



El método no destructivo más popular para medir atenuación, tanto en el laboratorio como en el campo, se basa en la medida de la luz retrodispersada por la fibra y se la conoce como reflectometría óptica temporal (OTDR). Esta resulta ser una técnica sofisticada que ha encontrado amplia aplicación tanto en laboratorio como en campo y es llamada frecuentemente método de medida de "backscattering". Es capaz de determinar la atenuación en un enlace óptico a lo

largo de su longitud, dando información de la dependencia de las pérdidas con la misma. En este sentido es superior a los métodos de medida descritos anteriormente, que sólo proveen pérdidas promediadas en toda la longitud de la fibra. Cuando la atenuación del enlace varía con la longitud, esta pérdida promedio resulta inadecuada. La OTDR también permite evaluar las pérdidas en conectores y empalmes. La técnica se basa en la medida y análisis de la fracción de luz que es retroenviada dentro de la apertura numérica de la fibra debido al scattering Rayleigh y posee la ventaja de requerir el acceso a solo un extremo de la fibra. El principio de funcionamiento se esquematiza en la figura 12 bis (a). Un pulso de luz de un diodo láser se inyecta en la fibra. La luz retroscatteredeada es detectada por un fotodiodo de avalancha que se conecta a un integrador Box-car para mejorar la relación señal-ruido a través de un promedio aritmético de un gran número de medidas tomadas en un determinado punto de la fibra. Esto es necesario debido a que la potencia óptica recibida es en general de un nivel muy bajo comparada con la potencia inyectada (del orden de entre 45 y 60 dB). La señal del integrador alimenta un amplificador logarítmico y las medidas promediadas para sucesivos puntos de la fibra se grafican en un registrador o se muestran en un CRT. Esto provee valores de atenuación en función de la distancia, que da un panorama general de la pérdida óptica a lo largo del enlace. Un gráfico de retroscattering típico se muestra en la figura 12 bis (b), donde se observa el pulso inicial causado por la reflexión en el inicio de la fibra, seguido de una larga cola causada por el scattering Rayleigh distribuido a medida que el pulso de entrada viaja por la fibra. El segundo pulso corresponde a una reflexión discreta en un conector debido a una discontinuidad en el índice de refracción. También se observa un salto debido a una región de alta atenuación (falla o empalme). El final de la fibra está indicado por un pulso correspondiente a una reflexión de Fresnel en la discontinuidad fibra-aire.



(a)

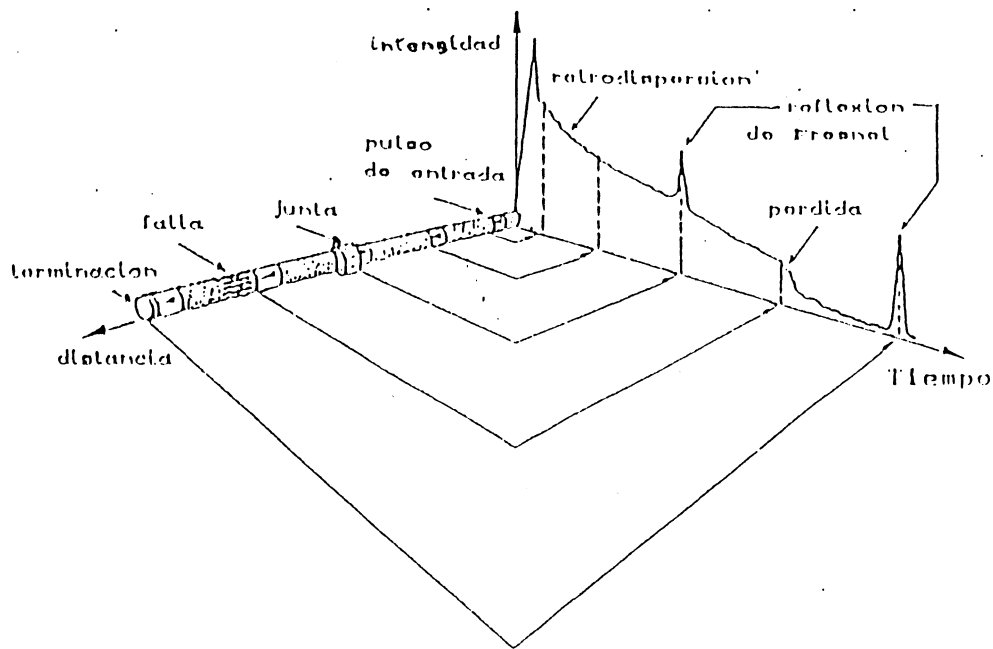


fig.12 bis (b)

II .2 DISPERSION

El fenómeno de dispersión se relaciona con la dependencia del índice de refracción de un medio con la frecuencia de la onda electromagnética que lo atraviesa $n(\lambda, \nu)$. Así, cuando se considera la propagación de un pulso de luz temporalmente angosto, se sabe que posee un espectro de frecuencias cuyo ancho es inversamente proporcional al ancho temporal del mismo. De modo tal que si este pulso se transmite a través de una fibra óptica, las distintas frecuencias componentes tendrán distintos índices de refracción y por lo tanto distintas velocidades de propagación, lo que causa un ensanchamiento del pulso con la longitud de propagación. Cuando la transmisión involucra algún tipo de modulación digital, este fenómeno causa el progresivo ensanchamiento de los pulsos individuales hasta hacerlos solapar entre sí, causando distorsión en el receptor . Este efecto se conoce como interferencia intersímbolo y es una de las causas de aumento del bit error rate.

El fenómeno de dispersión limita el ancho de banda máximo accesible a través de la frecuencia más allá de la cual los símbolos individuales se hacen indistinguibles. Para que no haya superposición de pulsos a lo largo del enlace con fibra, el digital bit rate no debe exceder al recíproco de la duración del pulso ensanchado por dispersión: $B_T \leq 1/2\tau$. El valor total del ensanchamiento del pulso depende de la distancia que este viaje a lo largo de la fibra. Así, dado un cierto enlace por fibra óptica, el ancho de banda usable está determinado por la distancia entre repetidores. Las medidas de las propiedades dispersivas suele darse entonces como el ensanchamiento temporal por intervalo de distancia recorrida (por ejemplo: ns/km). Debido a esto, el número de pulsos ópticos que pueden ser transmitidos en un dado período, y por lo tanto la capacidad de transporte de información de la fibra, está restringida al valor de dispersión por unidad de longitud. Esto lleva a la adopción de un parámetro más adecuado para representar la capacidad antes mencionada que es el producto del ancho de banda por longitud ($B_{opt} * L$) . En la figura 13 se muestran tres estructuras típicas de fibras y se esquematizan los respectivos ensanchamientos. Valores característicos del producto ancho de banda-distancia para esta tres fibras son 20 MHz km, 1 GHz km, y 100 GHz km respectivamente.

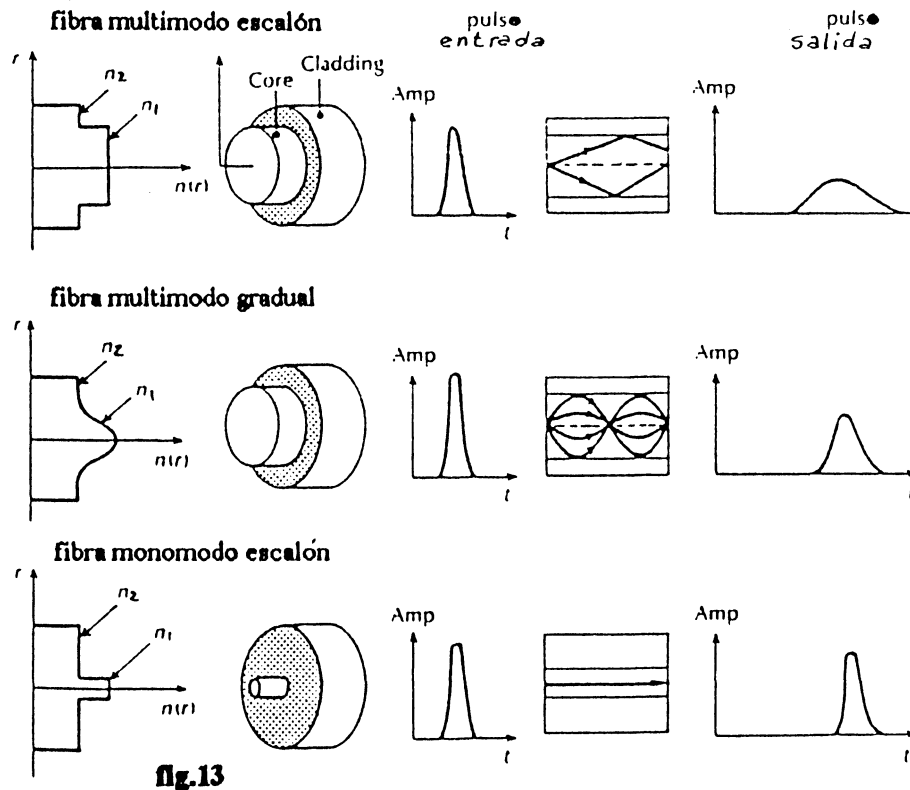


Fig.13

Consideremos ahora brevemente los diferentes mecanismos de dispersión que pueden encontrarse en los distintos tipos de fibras ópticas. Estos pueden ser dispersión intramodal (material y guiado) y dispersión intermodal.

II.2.1 DISPERSION INTRAMODAL

También llamada dispersión cromática, puede ocurrir en todo tipo de fibras y resulta del ancho de línea espectral finito de la fuente óptica. Como éstas no emiten una sola frecuencia sino más bien una banda de frecuencias (más o menos estrechas alrededor de la frecuencia central), existirán diferentes retrasos en la propagación de las distintas componentes espectrales de la señal transmitida. Esto causa un ensanchamiento del modo transmitido y por lo tanto dispersión intramodal. Las diferencias temporales pueden ser causadas por las propiedades dispersivas del material y por efectos de guiado dentro de la estructura de la fibra o guía de onda.

El ensanchamiento del pulso debido a la dispersión cromática resulta de las distintas velocidades de grupo de las diferentes componentes espectrales que son transmitidas a través de la fibra por la fuente óptica. Esto ocurre cuando la velocidad de fase de una onda plana que se propaga en un medio varía en forma no lineal con la longitud de onda y se dice que un material posee dispersión cuando la derivada segunda del índice de refracción con respecto a la longitud de onda es diferente de cero. Se define el parámetro de dispersión del material :

$$D_M = \frac{\lambda}{c} \left| \frac{\partial^2 n_1}{\partial \lambda^2} \right|$$

La dispersión de guiado es causada por la dependencia con la longitud de onda de la distribución espacial de luz del modo fundamental dentro de la fibra. A mayor longitud de onda, corresponde mayor penetración del modo en el cladding y mayor proporción del guiado en región con índice de refracción menor. Por lo tanto esta parte tendrá una mayor velocidad, lo que genera diferencias de tiempo dentro del ancho espectral. Un tratamiento con teoría electromagnética permite definir el coeficiente de dispersión por guiado :

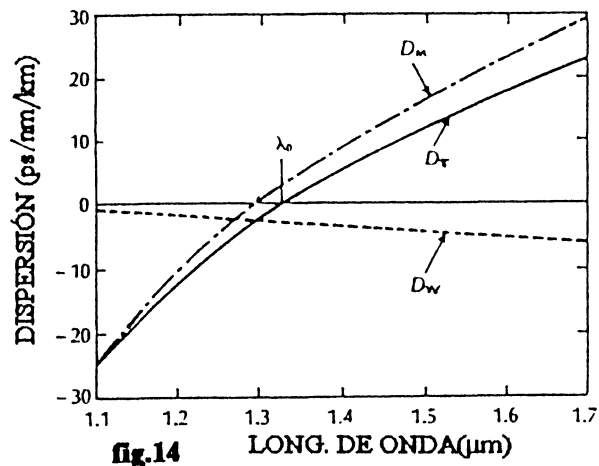
$$D_W = - \left(\frac{n_1 - n_2}{\lambda c} \right) V \frac{\partial^2 (Vb)}{\partial V^2} \quad (*)$$

Así, el coeficiente de dispersión total para fibras monomodo puede escribirse como :

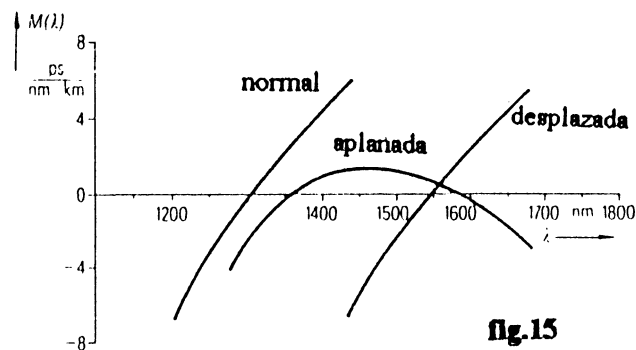
$$D_T = D_M + D_W$$

Para la región de longitud de onda de interés, estos parámetros se comportan como muestra la figura 14.

(* donde V es la llamada *frecuencia normalizada* y depende del diámetro del core y b es la *constante de propagación normalizada* y es función de la diferencia de índices de refracción. Ambos parámetros son adimensionales y surgen de resolver las ecuaciones de propagación de ondas en un dieléctrico.)



DM depende de las impurezas y DW de la estructura del perfil del índice. Así, variando el contenido de impurezas y el formato del perfil del índice de refracción es posible desplazar la curva de DT para que la región de mínima dispersión coincida con la de mínima atenuación (1.55 μm) (dispersión desplazada) o para crear una zona de baja dispersión entre 1.3 y 1.6 μm (dispersión aplanada) como se ve en la fig15.



II.2.2 DISPERSIÓN INTERMODAL

Es consecuencia de las diferentes velocidades de propagación de los modos de una fibra (multimodo). Se generan diferencias en los tiempos de propagación que ensanchan el pulso. Es la mayor causa de dispersión en fibras multimodo. En el caso de fibras de perfil escalonado, el retraso temporal entre el rayo axial y el meridional causa el ensanchamiento del pulso al propagarse por la fibra. Para fibras multimodo de índice gradual, el rayo axial sigue una trayectoria recta en la región de mayor índice y por lo tanto viaja a menor velocidad. El rayo meridional en cambio, sigue una

trayectoria sinusoidal por una región de menor índice y viaja entonces a mayor velocidad. Existe entonces una compensación de efectos que hace que los tiempos de propagación sean similares y produce una reducción de los retrasos temporales de los modos. Teóricamente es posible reducir esta dispersión en un factor cercano a 1000: 15 ns/nm.Km (escalonada) → 16 ps/nm.Km (gradual)

II . 3 ANCHO DE BANDA

El conocimiento de la función de transferencia de banda de base es esencial para determinar la capacidad de transporte de información de una fibra. Esta función de transferencia está principalmente determinada por la dispersión modal y la cromática. A potencias bajas, la fibra puede considerarse como un dispositivo lineal y la función de transferencia o respuesta en frecuencia $H(\omega)$ se define como la transformada de Fourier de la respuesta impulsiva $h(t)$ de la fibra: $H(\omega) = F[h(t)]$. Normalmente, toda la información acerca de la respuesta en frecuencia de la fibra se concentra en un parámetro: el ancho de banda, que se define como la frecuencia para la cual el módulo de $H(\omega)$ se reduce a la mitad del valor que toma para frecuencia cero.

$$|H(\omega_B)| = \frac{1}{2} |H(\omega = 0)|$$

Suponiendo una respuesta impulsiva de tipo gaussiana, se puede escribir el ancho de banda como:

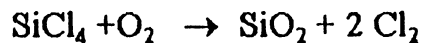
$$B = \frac{0,187}{\sigma_t}$$

con $\sigma_t^2 = \sigma_m^2 + \sigma_c^2$ y σ_m, σ_c son los anchos de respuesta impulsiva con dispersión modal y cromática respectivamente.

III. FABRICACION DE FIBRAS OPTICAS

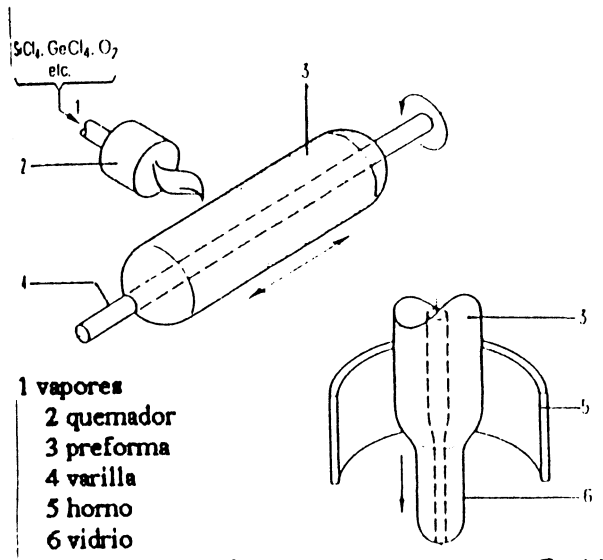
La fabricación de fibras ópticas se realiza en general en varias etapas de producción, posibilitando así optimizar en forma selectiva las propiedades ópticas, mecánicas y geométricas de la fibra. Las distintas técnicas utilizadas se basan en la fabricación de la llamada *preforma*, que consiste en una varilla de vidrio del cual surgirá tanto el núcleo como la cubierta. La sección transversal de la preforma es un modelo a escala de las dimensiones geométricas y del perfil de índice de refracción de la fibra a la que finalmente dará lugar. Mientras un extremo de la varilla es calentado a una temperatura mas alta, un cilindro extrude la fibra a partir de la preforma y simultaneamente se le aplica un recubrimiento plástico de protección.

El hito fundamental en la producción de fibras ópticas con muy baja atenuación lo constituye el método de *deposición en fase de vapor*, usado por primera vez en 1970 por la Corning Glass Works. En esta técnica, la deposición de dióxido de silicio (cuarzo) se produce por reacción de compuestos volátiles ultrapuros con oxígeno en la llama de un mechero:

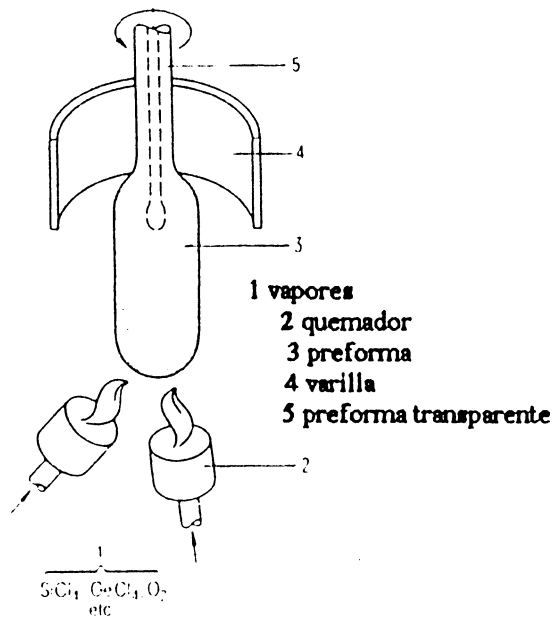


Se elige el tetracloruro de silicio ya que se puede producir por destilación en una forma mucho más pura que el dióxido de silicio.

Existen varias variantes del método de deposición en fase de vapor: el llamado OVD, donde la deposición se realiza sobre la superficie externa de una varilla rotante, el VAD, donde la deposición se hace sobre la cara de un extremo de una varilla de silica o el IVD donde se hace sobre la superficie interna de un tubo hueco rotante de silica. En esta última variante, la energía necesaria para la deposición se obtiene desde el exterior a través de un quemador de hidrógeno-oxígeno (MCVD) o desde el interior con una llama de plasma (PCVD). Estas técnicas están ilustradas en la figura 16 (a)-(d).

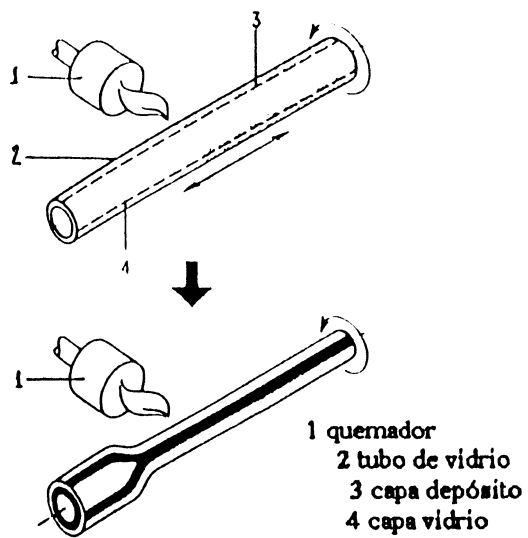


OVD (a)

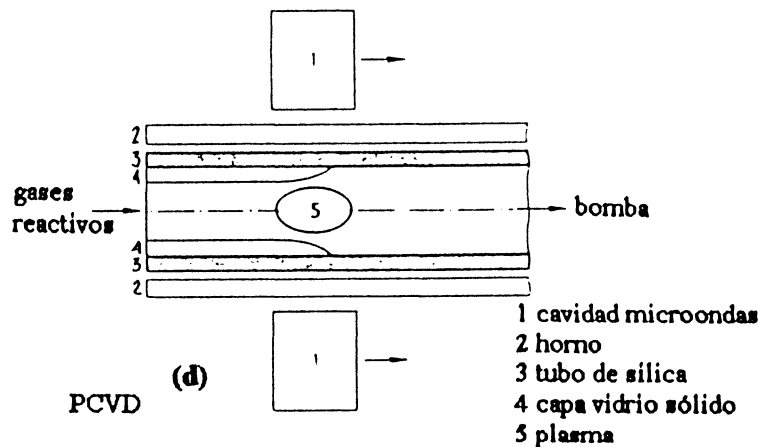


VAD (b)

fig.16



MCDV (c)



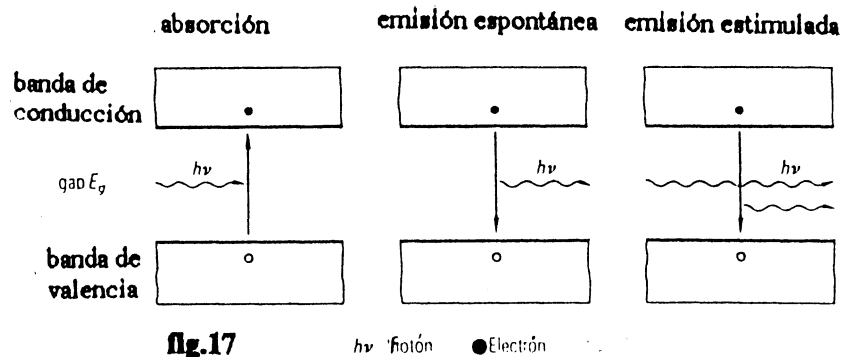
PCVD (d)

Estos métodos se utilizan corrientemente para la producción masiva de preformas para fibras multi y monomodo con atenuación extremadamente baja (0.2 dB/km @1550 nm) y gran ancho de banda (> 2 GHz.Km @ 1310 nm).

IV. DETECTORES Y EMISORES

Para poder transmitir señales luminosas a través de una fibra es necesario contar con dispositivos transmisores (que conviertan señales eléctricas en ópticas a través de la modulación de la intensidad de una fuente luminosa) y receptores (que realicen la conversión inversa). Estos transductores electroópticos se fabrican a partir de materiales semiconductores de los grupos III-IV-V ,junto con sus aleaciones.

Es sabido que los semiconductores poseen dos bandas de energía disponibles para los electrones: la **banda de valencia** y la de **conducción**, que permanecen separadas una de la otra por una banda de energías prohibidas E_g (figura 17).



Un fotón incidente sobre un semiconductor transfiere su energía a un electrón en la banda de valencia llevándolo a la banda de conducción. En el proceso el fotón desaparece y el electrón deja un lugar desocupado en la banda inferior llamado agujero positivo. Este proceso se conoce como **absorción**. Por otro lado, si en la banda de conducción hay una distribución de electrones mayor que la de equilibrio, entonces éstos pueden volver espontáneamente a la banda de valencia liberando la diferencia de energía en forma de fotones. A este proceso se lo conoce como recombinación radiativa de portadores de carga en exceso o **emisión espontánea**. Además existe un tercer proceso por medio del cual un fotón incidente

sobre un semiconductor causa que los portadores en exceso se recombinen radiativamente, es decir, el fotón incidente induce una transición electrónica desde la banda de conducción a la de valencia. En este caso, la radiación así inducida tiene la misma fase y longitud de onda que la de excitación y al proceso se lo llama *emisión estimulada*. Estos tres procesos ocurren, en general, simultáneamente, y es posible hacer que uno de ellos predomine sobre los otros de forma de poder hacer un uso técnico del mismo: el fotodiodo hace uso de la absorción, el diodo emisor de luz (LED) utiliza la emisión espontánea y el diodo láser la emisión estimulada.

Para que cualquiera de estos procesos ocurra, es necesario introducir un exceso de portadores en el semiconductor, lo que se logra inyectándolos a través de la llamada *juntura-pn* (zona de unión entre un semiconductor con defecto de electrones, positivo, y uno con exceso de electrones, negativo).

IV.1 EMISORES

Si la juntura pn se polariza en directa, se inyectan electrones adicionales a la capa p y huecos adicionales a la capa n, exceso que se usa para generar procesos radiativos en los dispositivos emisores (LED y diodo láser).

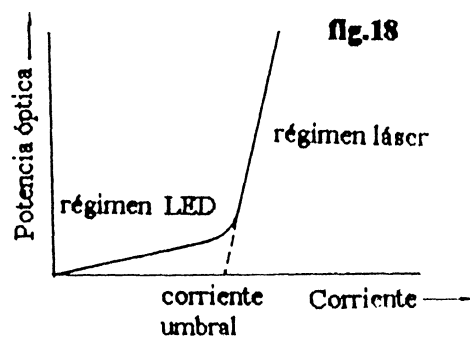
El LED es un dispositivo que emite luz por el proceso de emisión espontánea. La calidad de la conversión de corriente eléctrica en luz está descrita por la llamada eficiencia cuántica, que da el número de fotones emitidos en la unidad de tiempo por número de portadores que atraviesan la juntura-pn (para GaAs varía típicamente entre 0.5 y 1%). Debido a que esta eficiencia es inversamente proporcional a la temperatura, debe evitarse un recalentamiento de la zona de recombinación, proveyendo una adecuada disipación del calor para el LED. La longitud de onda de la luz emitida está determinada principalmente por E_g y es posible calcular su valor a través de la relación:

$$\lambda = \frac{1.24}{E_g}$$

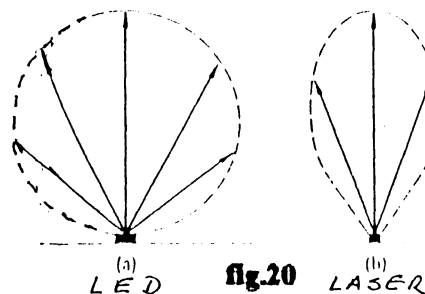
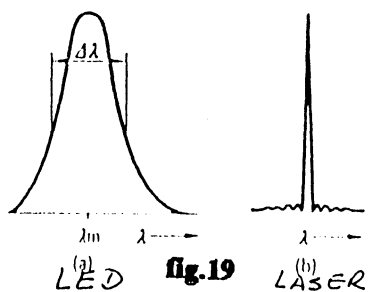
Por ejemplo, para GaAs, $E_g=1,43$ eV y $\lambda =0,89$ μm ; para InP, $E_g=1,35$ eV y $\lambda =0,92$ μm . El ancho espectral de un LED es aproximadamente proporcional al cuadrado de la longitud de onda y por lo tanto aumenta considerablemente hacia λ mayores.

La vida media promedio del exceso de portadores es la principal causa de emisión de luz luego de haber cesado de circular corriente. Así, por debajo de este valor de tiempo, el diodo no puede seguir los cambios en la corriente inyectada. Las vidas medias típicas están en el orden de unos pocos nanosegundos, lo que corresponde a un **ancho de banda de modulación de alrededor de 100 MHz**.

El diodo láser utiliza, a diferencia del LED, el proceso de emisión estimulada descrito más arriba. Si se aumenta la densidad de corriente de inyección en una juntura-pn, se genera un gran exceso de portadores en la banda de conducción y por lo tanto una situación de inversión de población entre dicha banda y la de valencia. La recombinación radiativa estimulada de pares electrón-hueco da lugar a acción láser, la que se ve reforzada por acción del resonador óptico formado por las caras facetadas del mismo cristal semiconductor. La figura 18 muestra la diferencia de operación de un LED y un diodo láser. A medida que se aumenta la corriente en el diodo, se llega a un valor umbral por encima del cual la ganancia (o amplificación)



supera a las pérdidas dentro del cristal debidas a atenuación, difusión, etc, y se genera una fuerte acción láser. El ancho espectral de esta emisión se reduce en alrededor de un orden de magnitud con respecto al ancho del LED (figura 19). En contraste con la emisión de éste, la radiación de un diodo láser posee coherencia espacial debido al proceso de emisión estimulada y muestra una divergencia angular notablemente reducida, tal como se esquematiza en la figura 20.



Para que un diodo láser opere en la práctica, se requieren otros grupos funcionales que normalmente vienen encapsulados, junto con el diodo, en un pack hermético llamado módulo láser. La figura 21 muestra un esquema en bloques de uno de estos módulos.

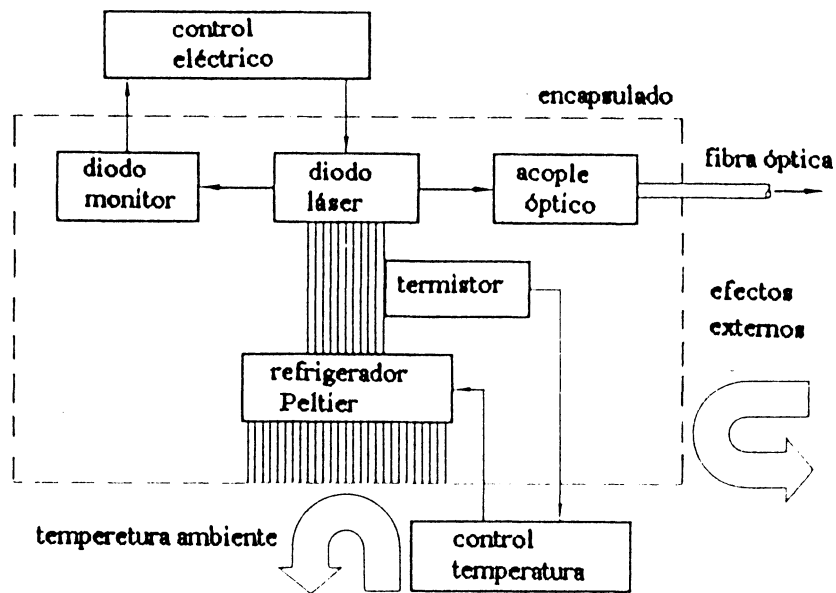


fig.21

La temperatura del diodo se monitorea a través de un termistor que realimenta el control eléctrico de un refrigerador Peltier acoplado al diodo. La salida láser se monitorea con un fotodetector semiconductor, cuya salida eléctrica es tomada por la fuente de alimentación del diodo, que controla la corriente de inyección de portadores a la juntura. La salida láser exterior puede realizarse a través de un sistema de lentes o acoplarla dentro mismo del pack a una fibra óptica que sale fuera del módulo y permite un acople conveniente a la ruta de transmisión elegida.

IV.2 DETECTORES

Ya hemos visto que cuando un fotón cuya energía es mayor que el rango de energías prohibidas E_g incide sobre un semiconductor, es absorbido por el mismo, excitando un electrón de la banda de valencia a la de conducción y generando así un par electrón-hueco. Si esto se produce en una juntura-pn, es posible remover los fotoportadores generados. La figura 22 muestra un esquema de este proceso para un fotodiodo.

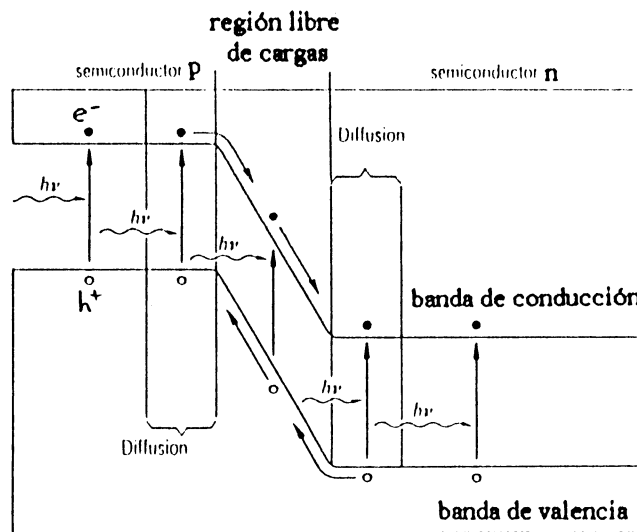


fig. 22

Tanto los portadores generados en la región de depletion como en la de difusión son separados por el campo eléctrico imperante en el juntura (aunque los de la segunda región tardan un poco más en ser arrastrados por el campo debido a que tienen que difundir a través de la juntura), generándose así una corriente eléctrica proporcional a la intensidad de la radiación incidente.

Los materiales utilizados para la construcción de estos fotodiodos son el silicio, el germanio o compuestos de los grupos III-IV-V. El coeficiente de absorción de estos materiales es función de la longitud de onda (figura 23). Puede verse que el silicio es adecuado para el rango por debajo de 1000 nm, mientras que el germanio es adecuado para la región entre 1100 y 1600 nm. En el rango entre 1300 y 1600 nm, los materiales más adecuados resultan ser el InGaAs y el InGaAsP.

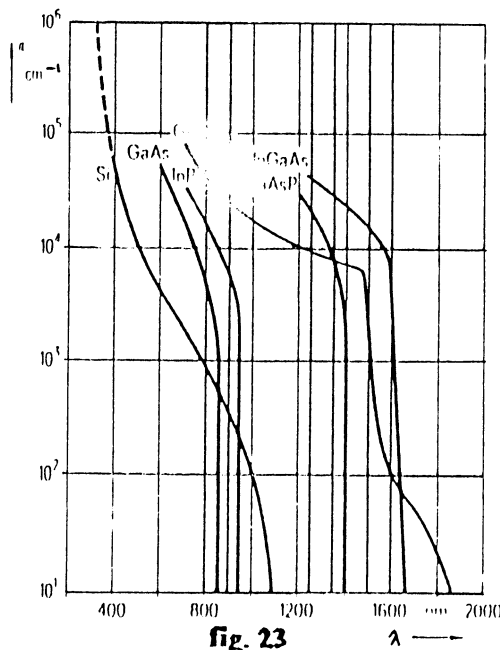


fig. 23

Cuando la aceleración de los portadores debida al campo eléctrico les imprime tal velocidad que estos producen portadores secundarios por colisiones ionizantes, se genera una fotocorriente notablemente más fuerte que en otros casos. A este proceso se lo conoce como breakdown de avalancha y al dispositivo que lo utiliza se lo llama fotodiodo de avalancha. Poseen mayor sensibilidad que los fotodiodos comunes y son los utilizados preferentemente en el extremo de detección en un tendido de cable óptico.

V. CONEXIONADO Y CABLES DE FIBRA ÓPTICA

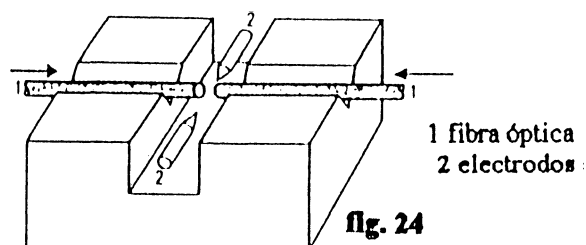
En este capítulo describiremos brevemente los elementos relativos al conexionado de fibras ópticas, tanto en lo referente a empalmes como a conectores discretos. También haremos una sucinta referencia a los llamados cables ópticos, sus características y tipos.

En lo que se refiere al conexionado, es importante destacar que son fuentes importantes de atenuación (pérdidas por inserción), por lo que resulta necesario optimizar los mismos. Existen dos tipos de conexionado o uniones: las llamadas permanentes o empalmes y las llamadas no-permanentes o conectores.

V.1 EMPALMES

Los empalmes pueden ser de tipo mecánico o de fusión. En los de tipo mecánico, las fibras a conectar se alinean cuidadosamente y luego se enfrentan sobre una canaleta en forma de "V" practicada sobre una planchuela metálica. Cuando la distancia entre ambas es pequeña, son pegadas a forma permanente con un adhesivo rápido que ajusta los índices de refracción. Una planchuela similar se ajusta por encima para fijar el conexionado. La pérdida promedio en este tipo de empalme es de 0,2 dB.

El empalme por fusión se utiliza para fibras con índice de refracción gradual hechas de silica o de vidrio con multicomponentes. La figura 24 muestra un esquema de este tipo de empalme :



Una tensión alterna de alta frecuencia produce un arco eléctrico entre dos electrodos enfrentados, generando un microplasma de alta temperatura entre ellos. Cuando los extremos de las fibras a empalmar son alcanzadas por este plasma se funden y la tensión superficial del vidrio así fundido tiende a autoalinearse los ejes de las fibras. Este efecto logra compensar un desalineado previo a la fusión de hasta $10\ \mu\text{m}$. El dispositivo de empalme contiene un cortador de fibra de precisión que asegura obtener fracturas con errores angulares inferiores a 3° . También posee un microscopio de proyección que permite observar el procedimiento de enfrentado y alineación de las fibras.

Para fibras monomodo, cuyo diámetro de núcleo es aproximadamente de $9\ \mu\text{m}$, se requiere una precisión y cuidado durante el proceso de empalme superiores a la que se puede obtener con los dispositivos para fibra multimodo de índice gradual. Para aquel tipo de fibra se han desarrollado posicionadores de precisión en los tres ejes. El proceso de fusión se lleva a cabo automáticamente en el mismo orden que para fibras de índice gradual y donde a través de un potenciómetro es posible ajustar la distancia entre fibras de $0,20\ \text{mm}$. En contraste con fibras multimodo, el efecto de autocentrado causado por la tensión superficial del vidrio fundido es pernicioso para fibras monomodo, ya que el movimiento del vidrio en la zona de empalme puede causar un doblez del núcleo, que puede llevar a altas pérdidas. Esto puede evitarse aortando el tiempo de fusión, cortes rectos de la fibra ($< 1^\circ$) y baja corriente de fusión. Las pérdidas promedio de empalme por este método son menores que $0,1\ \text{dB}$.

Para optimizar el proceso de alineado se desarrolló un sistema de inyección local de luz y detección tanto para fibras multimodales como para monomodales. A cierta distancia de la zona de empalme, la fibra es doblada por sendos mandriles. En una de dichas curvas se inyecta luz utilizando las pérdidas que introduce el doblado de la fibra mientras que se detecta en la otra curva la intensidad de salida de la misma. Los extremos de las fibras a empalmar se posicionan de tal forma que maximicen la intensidad de salida de luz (figura 25).

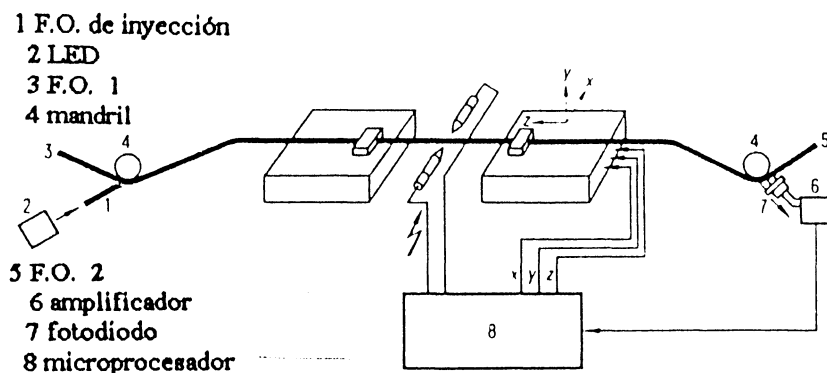


Fig. 25

V.2 CONECTORES

En las rutas de transmisión óptica, los conectores se utilizan para unir fibras de modo que sea fácil desconectarlas en caso de reparaciones y que además aseguren una baja pérdida de conexión. Estos conectores poseen severas restricciones en cuanto a sus tolerancias mecánicas, dependiendo del diámetro del núcleo de la fibra en cuestión y de la pérdida por inserción que se requiera. Existen dos tipos de conectores dependiendo de su principio de operación: los que utilizan acoplamiento por lentes y los que usan acoplamiento por extremo.

En el primer caso se utilizan lentes u otros sistemas formadores de imágenes para transformar la luz que sale de una de las caras de la fibra en un haz paralelo de rayos con diámetro grande para luego concentrarla nuevamente hacia la cara de la fibra receptora (figura 26). La ventaja de este tipo de acoplamiento es que permite mayor tolerancia de separación en la región de conexión, aunque se aumenta la pérdida debida a reflexiones en las interfaces entre medios con diferentes índices de refracción. Por ello este tipo de conectores se reserva para conexiones de fibras en el interior de instrumentos de medición (como por ej. atenuadores).

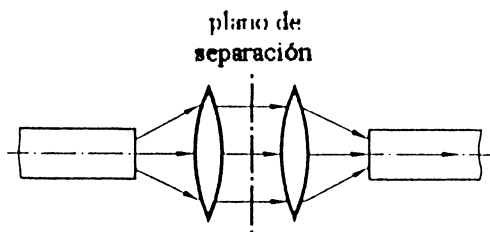


fig. 26

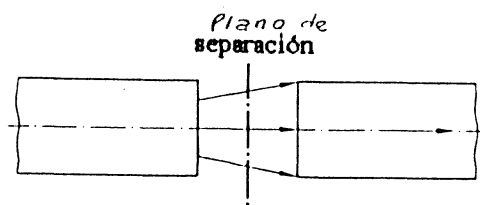


fig. 27

En los conectores de acoplamiento por extremo (figura 27), la fibra de salida se posiciona en paralelo y cercanamente a la fibra de entrada. Ambas caras deben estar correctamente cortadas y pulidas de forma que se produzca la menor pérdida posible por inserción. Para evaluar la calidad de transmisión en un conector de este tipo es necesario determinar la magnitud del aumento en la atenuación de la

línea óptica cuando se adiciona el conector. Las pérdidas por inserción son debidas en general a tres causas : desajuste axial de la fibra, desalineado angular y exceso de separación entre las caras de las fibras (figura 28). Estos defectos deben llevarse a un mínimo, lo cual requiere un procedimiento de alta precisión en el mecanizado del conector. La figura 29 muestra un esquema de un conector desmontable de este tipo. En general las pérdidas por inserción son del orden de 1 dB.

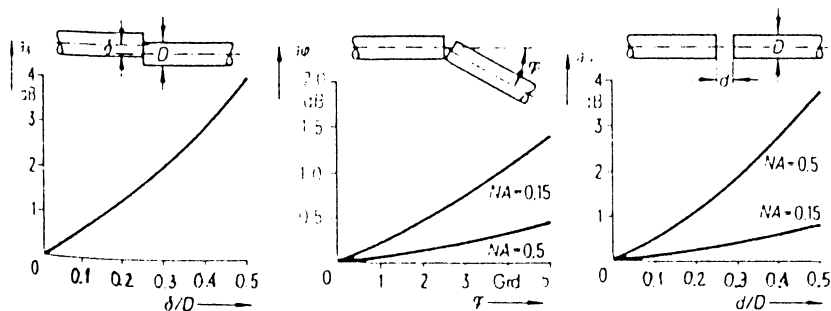


fig. 28

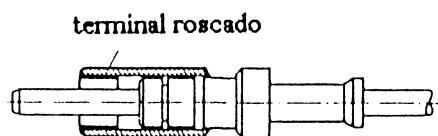


fig. 29

V.3 ACOPLADORES ÓPTICOS

Cuando la comunicación de datos debe ser transmitida en ambas direcciones via enlace óptico o en el caso de que varios transmisores y receptores deban unirse a traves de una estructura de bus óptico, se necesitan los llamados acopladores ópticos. Estos son de tipo multipuerta, es decir que tienen por lo menos tres puertas ópticas a traves de las cuales puede acoplarse luz hacia el interior o exterior de la fibra.

Los acopladores de tres puertas (figura 30) contiene una fibra continua 1-2 con un semiespejo depositado en e camino de los rayos. La fibra 3 se acopla con la continua a través del semiespejo, que refleja parte de la luz incidente. Si éste divide la señal de entrada en una relación 1:1, la atenuación debida al acoplador y a los posibles conectores que lo unen a la línea óptica es menor que 6,5 dB. Es posible obtener otras relaciones de división hasta un máximo de 1:10, y pueden operarse tanto con fibras de índice escalonado como de índice gradual.

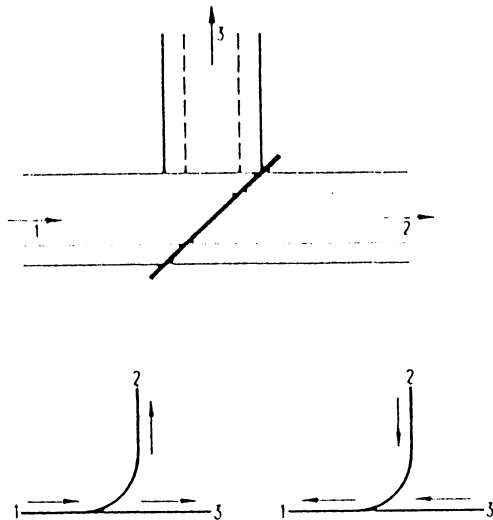


fig. 30

Los acopladores de cuatro puertas se construyen con dos tipos de tecnologías. En una de ellas, dos pares de fibras se disponen en forma de cruz. En la intersección de los cuatro brazos se deposita un espejo parcialmente transmisor ubicado a 45° (figura 31). Este acoplador sin lentes no muestra dependencia modal y funciona tanto para fibras de índice escalonado como de índice gradual.

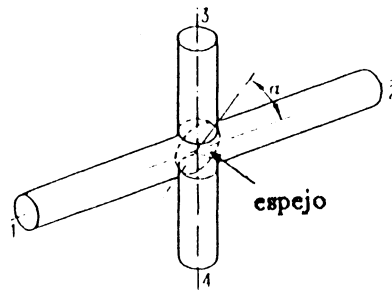


fig. 31

El otro tipo de acoplador de cuatro puertas es el llamado de onda evanescente (o bicónico). En este, dos pares de fibras se retuercen levemente y se estiran bajo la acción de una llama (figura 32). Los cambios en la sección transversal así obtenidos causan que, por ejemplo, la luz guiada en la fibra 1 incida sobre el lado interior de la cubierta en un ángulo pronunciado (inferior al ángulo crítico) y por lo tanto se transfiera dentro de la cubierta. Cuando la sección transversal vuelve a ensancharse lo suficiente, la luz se introduce en la región del núcleo de las fibras 2 y 3. La proporción de división de la luz depende de las variaciones de diámetro $D2/D1 < 1$ y de la longitud de la zona de acople.

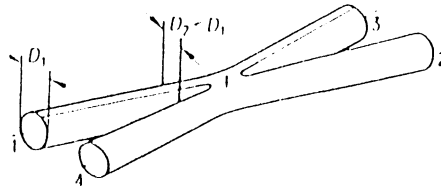


Fig. 32

V.4 CABLE ÓPTICO

En la mayoría de las aplicaciones, las fibras (tanto monomodo como multimodo) no pueden usarse sin una adecuada infraestructura de protección, debido principalmente a su naturaleza frágil y quebradiza y al notable aumento de la atenuación que pueden producir los esfuerzos de tracción, de torsión y de doblado en el tendido de una línea óptica de comunicación. Para este fin existen, por un lado, un dispositivo especial llamado buffer y por otro métodos que incorporan dentro del cable elementos tensores y antibucle que permiten optimizar el diseño del mismo para que cumpla con la prestación requerida.

Se llama buffer a un pequeño tubo de plástico que lleva dentro una (o más) fibra óptica suficientemente protegida contra deformaciones y/o fricciones. El tubo está formado por una capa interna de muy bajo coeficiente de roce y una capa externa de protección contra influencias mecánicas. Normalmente la fibra posee unas décimas de milímetro de espacio libre dentro del buffer, de tal forma de tener algún tipo de movilidad tanto radial como longitudinal. Posee además la ventaja de poder abrirse fácilmente y permitir un acceso rápido a la fibra. La figura 33 ilustra los aspectos arriba mencionados. Para prevenir la posible introducción de agua que pudiera congelarse y causar daños por expansión, el buffer se llena con un com-

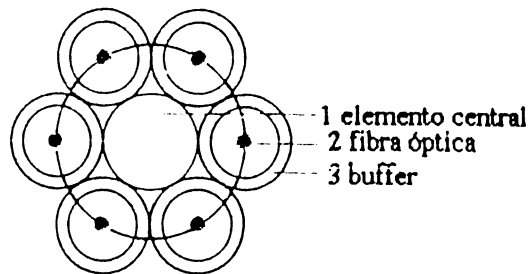


Fig. 33

puesto tipo gel que mantiene su consistencia en un amplio rango de temperaturas entre -30° y $+70^{\circ}$ C, no es corrosivo y es fácil de limpiar a la hora de acceder al interior del buffer para reparar o conexas la fibra.

Existe otro tipo de diseño llamado de cinta o encintado por medio del cual es posible hacer un arreglo de fibras individuales dentro de un cable como muestra la figura 34. Una cinta de fibras consiste de 12 fibras pegadas una paralela a la otra. Es posible apilar 12 de éstas cintas una arriba de la otra y formar un arreglo de sección rectangular. Este arreglo forma el núcleo del cable y puede doblarse sobre sí

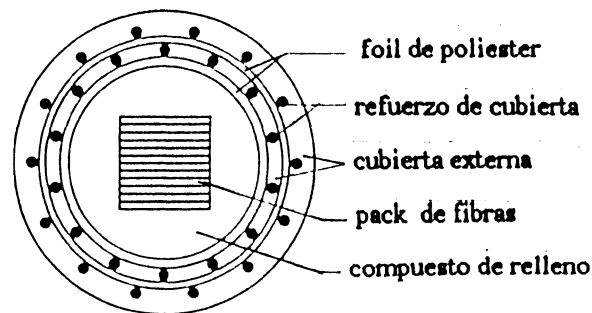
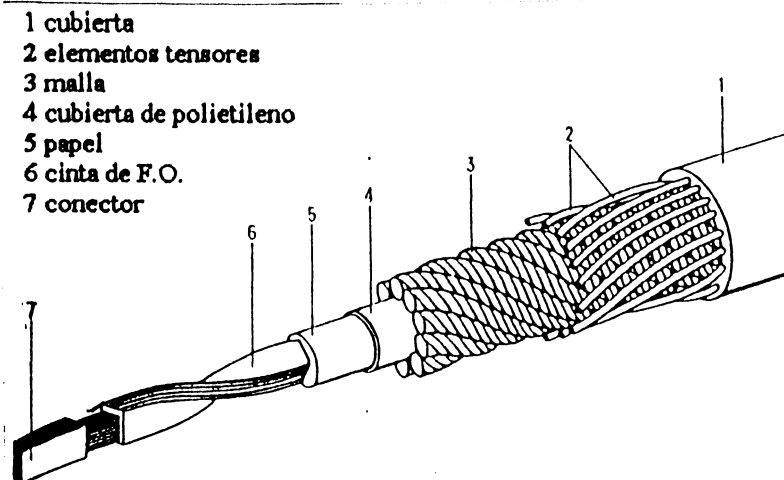


fig. 34

mismo en el eje longitudinal del cable a lo largo del tendido del mismo (figura 35). Una doble cubierta reforzada con alambres de acero es estrudada sobre el núcleo, al que se llena con un gel a base de petróleo. La ventaja de este tipo de cable con un gran número de fibras radica en la alta densidad de empaquetamiento y en la simplificación de la técnica de conectores debido a la posición ordenada de las fibras dentro del cable.



- 1 cubierta
- 2 elementos tensores
- 3 malla
- 4 cubierta de polietileno
- 5 papel
- 6 cinta de F.O.
- 7 conector

fig. 35

A partir de la estructura del buffer de la fibra descrita mas arriba, se debe seleccionar el núcleo, la cubierta y la armadura del cable óptico que mejor favorezcan las prestaciones planificadas durante una prolongada vida útil, prestando especial atención a la protección del mismo contra influencias ambientales como cambios de temperatura y esfuerzos mecánicos.

Para aumentar la resistencia mecánica del núcleo de los cables ópticos, ya sea con buffers monofibra o multifibra, éstos se despliegan en forma helicoidal a lo largo de un elemento tensor central que sirve a la vez de sostén y relajador de posibles tensiones. En general el despliegue de estos buffers se realiza por capas concéntricas alrededor del elemento central (cable de capas) o varias de estas unidades alrededor del elemento central (cable de unidades), lo que puede verse en la figura 36. En las

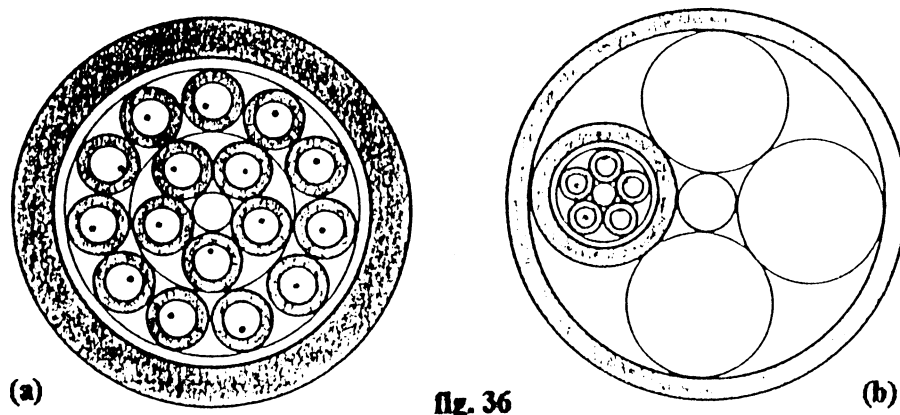


fig. 36

redes troncales de telecomunicaciones predominan los cables multifibra del primer tipo, mientras que en las redes locales se prefieren los del segundo tipo debido a que permiten aumentar la densidad de empaquetamiento.

Existe otro tipo de núcleo llamado acanalado o ranurado y cuya sección se ilustra en la figura 37. En éste, las fibras yacen en finos canales pretallados en forma helicoidal en la superficie del elemento central. Estos canales pueden contener, según su dimensión, una o más fibras, y, al igual que los buffers, se llenan con un gel protector. Para asegurar la hermeticidad longitudinal del cable a posibles filtraciones de agua, se rellenan los intersticios vacíos del núcleo con un compuesto de bajo coeficiente de expansión térmica a muy alta presión (15 bar). Una capa de adhesivo termoplástico extrudado alrededor del núcleo sirve como barrera de contención a dicho compuesto.

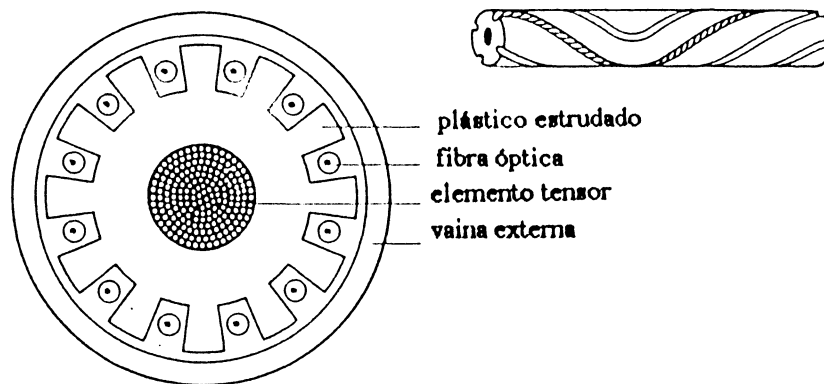


fig. 37

Finalmente, digamos que la cubierta protectora del cable se construye de polietileno o de cloruro de polivinilo (PVC), éste último especialmente para cables interiores. Frecuentemente, los cables se tienden entubados bajo tierra sin armadura, ya que los elementos tensores, firmemente unidos al cable por el adhesivo termoplástico, son suficientes para soportar los esfuerzos típicos a los que se ve sometido. Sin embargo, para aplicaciones especiales como por ejemplo cables submarinos o aéreos, se provee una armadura consistente en una cinta de acero arrollada alrededor del cable o una cubierta de PVC laminada en acero. Según la protección requerida, el arrollado o laminado puede ser simple o múltiple.

SENSORES DE FIBRAS OPTICAS

R. Duchowicz

La tecnología de fibras ópticas ha posibilitado durante las décadas pasadas la revolución producida en las telecomunicaciones y es uno de los factores esenciales del desarrollo de otras áreas como televisión por cable y realización de enlaces locales. La mayor parte de los trabajos sobre el tema fue estimulada en el pasado por programas militares, como el de la Armada de Estados Unidos. Si bien parte de estos programas aun continúan, principalmente en las áreas de sistemas de sonar y giroscopos de fibras, en los últimos años se han hecho significativos progresos en aplicaciones de tipo industrial y medica.

Una revolución similar fue producida en la industria optoelectronica, con los diodos emisores de luz (LED) y los láseres de semiconductores, los que jugaron roles claves en el desarrollo de dispositivos como discos compactos, impresora láser, señaladores y, mas recientemente, en la digitalización de imágenes (scanners). La tecnología de los sensores por fibras ópticas se ha desarrollado en paralelo con dichas industrias y ha sido beneficiada en gran medida del aprovechamiento de componentes de bajo costo y alta performance asociados con producciones masivas. Esta conjunción ha resultado en la introducción en el mercado de sensores de fibras en la forma de giroscopos ópticos, sensores biomédicos y sensores para procesos de control, entre otros. Como el costo de la mayor parte de los componentes que se utilizan continua en descenso, es de esperar un mayor desarrollo de esta tecnología y de los campos de aplicaciones en el futuro.

En muchos procesos es necesario monitorear ciertas cantidades o parámetros tales como desplazamiento, temperatura, presión, nivel de liquido, etc. La tabla 1

muestra distintas áreas de aplicación de estos sensores y las magnitudes a medir en cada caso.

Tabla 1: Varios campos de aplicaciones y parámetros a medir de los sensores de fibras ópticas.

Aplicación	Parámetro a medir
Líneas de producción (papel, acero, etc.)	Posición, espesor, switch de límites, detección de roturas, velocidad.
Control de procesos	Temperatura, presión, flujo, análisis químicos.
Automatización	Temperatura, presión, torque, detección de gases, aceleración.
Maquinas herramientas	Desplazamiento, detección de rotura de herramientas.
Aviación	Temperatura, presión, rotación, desplazamiento, nivel de líquidos.
Calefacción, ventilación aire acondicionado	Temperatura, presión, flujo.
Petroquímica	Detección de pérdidas de gases inflamables y tóxicos, nivel de líquidos.
Militar	Sonido, rotación, radiación, vibración, posición, temperatura, presión, nivel de líquidos.
Geofísica	Campo magnético, tensión.

Siendo una fibra óptica un medio físico, esta sujeta a ser afectada por perturbaciones de distintos tipos. Puede experimentar cambios geométricos (tamaño, forma) y de en sus propiedades ópticas (índice de refracción, distribución de la luz en los modos) de mayor o menor medida dependiendo de la naturaleza y magnitud de la perturbación. En aplicaciones como Comunicaciones Ópticas, uno trata de minimizar tales efectos de manera que no perturben la transmisión y recepción de señales. Por el contrario en los sensores de fibra se trata de maximizar la respuesta a estímulos externos, de manera que los cambios en la radiación óptica pueda ser utilizada como una medida de dicha perturbación. En Comunicaciones la señal es modulada, mientras que en nuestro caso la fibra bajo el estímulo externo puede actuar como modulador.

La figura 1 muestra dos configuraciones particulares de los sensores de fibras ópticas: en el primero la fibra cumple dos roles distintos, transportar la luz y modularla en su interacción con un estímulo externo; en el segundo, en cambio, solamente transporta la luz la que en un momento sale de la fibra, es enviada a una región donde alguna de sus propiedades es alterada y luego es colectada nuevamente en una fibra óptica (la misma u otra). En el primer caso se habla de sensores de tipo intrínseco y en el segundo, de tipo extrínseco. El elemento que permite la interacción se

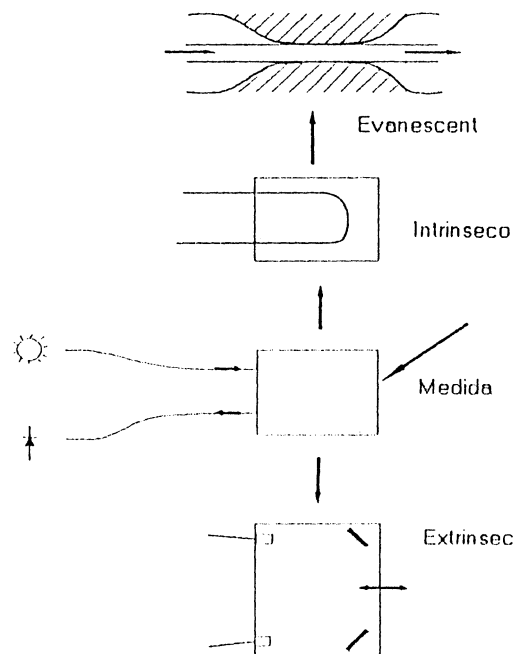


Figura 1: Tipos de sensores

denomina transductor y es el que permite convertir variaciones de temperatura, presión, rotación, corriente eléctrica, etc. en cambios de las características de la radiación óptica. Dado que la luz se caracteriza por su amplitud (intensidad), fase, frecuencia y polarización, la interacción externa puede actuar sobre cualquiera de dichos parámetros. La utilidad de los sensores de fibra depende de la magnitud de este cambio y de sus posibilidades de cuantificar adecuadamente dicho cambio.

Los argumentos en favor de los sensores de fibras ópticas son similares a aquellos para Comunicaciones por fibras, tales como:

- a) Inmunidad a interferencias frecuencias de radio (RFI) y electromagnéticas (EMI).
- b) Aislación eléctrica.
- c) Pequeño tamaño y peso.
- d) Posibilidad de control a distancia lo que permite el acceso en áreas inaccesibles.
- e) Gran sensibilidad potencial.
- f) Posibilidad de realizar interfaces de gran cantidad de mediciones con sistemas de comunicación de datos.
- g) Seguridad en transmisión de datos.
- h) Alta exactitud y confiabilidad de un equipo de estado sólido.
- i) Potencialmente resistente a radiación ionizante.
- j) Economía.

Otra ventaja es que mucho de los elementos ópticos utilizados (separadores de haces, multiplexores, filtros, líneas de retardo, etc.) pueden ser realizados con fibras permitiendo la construcción de sensores totalmente realizados con tecnología de fibras ópticas. Circuitos fotonicos (óptica integrada) pueden elaborarse también como elementos de procesamiento de señales de un solo chip lo que posibilita una producción en serie y económica de dichos elementos.

Tipos de sensores de fibras

De acuerdo a lo expresado anteriormente los sensores pueden ser clasificados, según sean los procesos de modulación/demodulación de la luz, como sensores de amplitud (intensidad), fase, frecuencia o polarización. Dado que los cambios de fase o frecuencia se presentan al utilizar técnicas interferométricas, los sensores de fase son denominados también interferométricos. Estos casos implican una la detección de tipo coherente o heterodina. Los sensores de intensidad son básicamente incoherentes. Mientras que estos últimos son en general de construcción sencilla, los

interferométricos presentan una mayor complicación en su diseño pero poseen una mayor sensibilidad.

Los sensores de fibras ópticas también pueden ser clasificados de acuerdo a su función: sensores físicos (miden temperatura, presión, etc.), sensores químicos (medida de pH, análisis de gases, estudios espectroscópicos, etc.), sensores biomédicos (insertados vía catéteres o endoscopios para medir flujo sanguíneo, contenido de glucosa, etc.).

Características del diseño y análisis de funcionamiento.

La configuración más general consiste de una fuente luminosa, un transductor y un detector óptico interconectado por fibras. Las fuentes utilizadas habitualmente son emisores de tipo espontáneo o coherentes tales como LEDs, láseres de semiconductores, fuentes gaseosas, láseres gaseosos (argón, HeNe), láseres de cristales sólidos (ej. Nd-YAG), etc. tanto continuos como pulsados. Como detectores se utilizan principalmente fotodiodos y fotomultiplicadores. Las fibras utilizadas también cubren un amplio espectro, incluyendo fibras de silicio convencionales monomodos y multimodos y fibras de plástico. Las pérdidas de intensidad debidas a la atenuación en las fibras no son en general significativas, debido las cortas longitudes empleadas. Sin embargo la influencia del medio ambiente sobre estas sí pueden ser importantes dado que pueden modificar sus propiedades y por lo tanto, las características de la señal luminosa. Muchos de los componentes de Óptica Integrada como lentes Grin, acopladores, etc. son también empleados. A fin de ilustrar todo lo expresado, veremos a continuación distintos ejemplos de sensores por fibras ópticas.

Sensores de cambio de la intensidad

Sensores de desplazamiento

Los sensores pasivos más simples, que utilizan fibra óptica multimodal (FO-MM), son los de posición o desplazamiento. Por ejemplo, cuando los terminales de dos fibras inicialmente enfrentadas y alineadas se mueven lateralmente entre sí, la pérdida

introducida en el acoplamiento depende del desplazamiento. Diversos sensores que utilizan la misma idea, están graficados en la figura 2. En estos casos se utilizan distintos obturadores los cuales se desplazan lateralmente. En los casos (a), (b) y (c) simplemente cortan el paso de la luz operando con distinta sensibilidad y rango (uno se contrapone con el otro). El diseño de la figura 2(c) permite tener un rango mayor aunque la sensibilidad es menor a la de los anteriores.

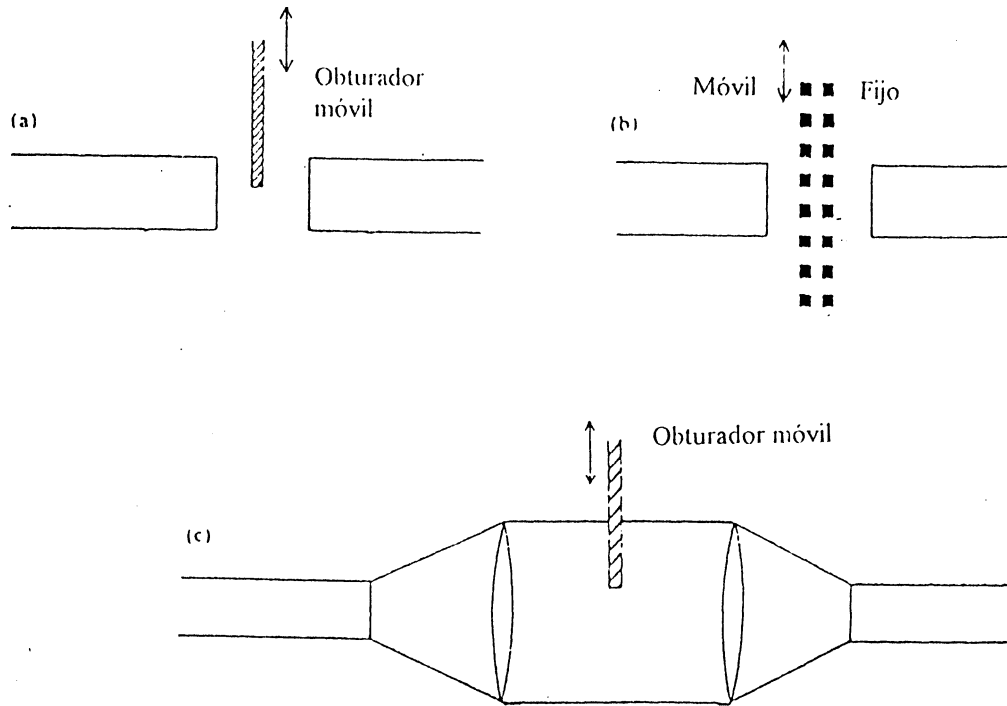


Figura 2: Sensores de desplazamiento simples: en (a) un obturador móvil varía la luz acoplada; en (b) el uso de dos redes aumenta la sensibilidad y en (c) un sistema de expansión del haz permite aumentar el rango de medida.

El mostrado en la figura 3 es uno de los primeros sensores de desplazamiento de tipo comercial (denominado Fotonic sensor). Utiliza un manojo de fibras mitad de las cuales están conectadas a una fuente de radiación, mientras que la otra mitad del mismo extremo están conectadas a un detector. El otro extremo del conjunto de fibras es colocado en las proximidades del elemento cuyo desplazamiento se quiere determinar. La iluminación que producen parte de las fibras es reflejada hacia atrás por el objeto y recogida por las fibras restantes las cuales la dirigen al detector. La cantidad de luz detectada depende de la distancia del terminal de fibras a la superficie. La relación no es, en general, lineal salvo en algún intervalo. Debido a ello, no es apropiado para medir

largos desplazamientos. Este sensor fue desarrollado inicialmente para medidas de vibraciones.

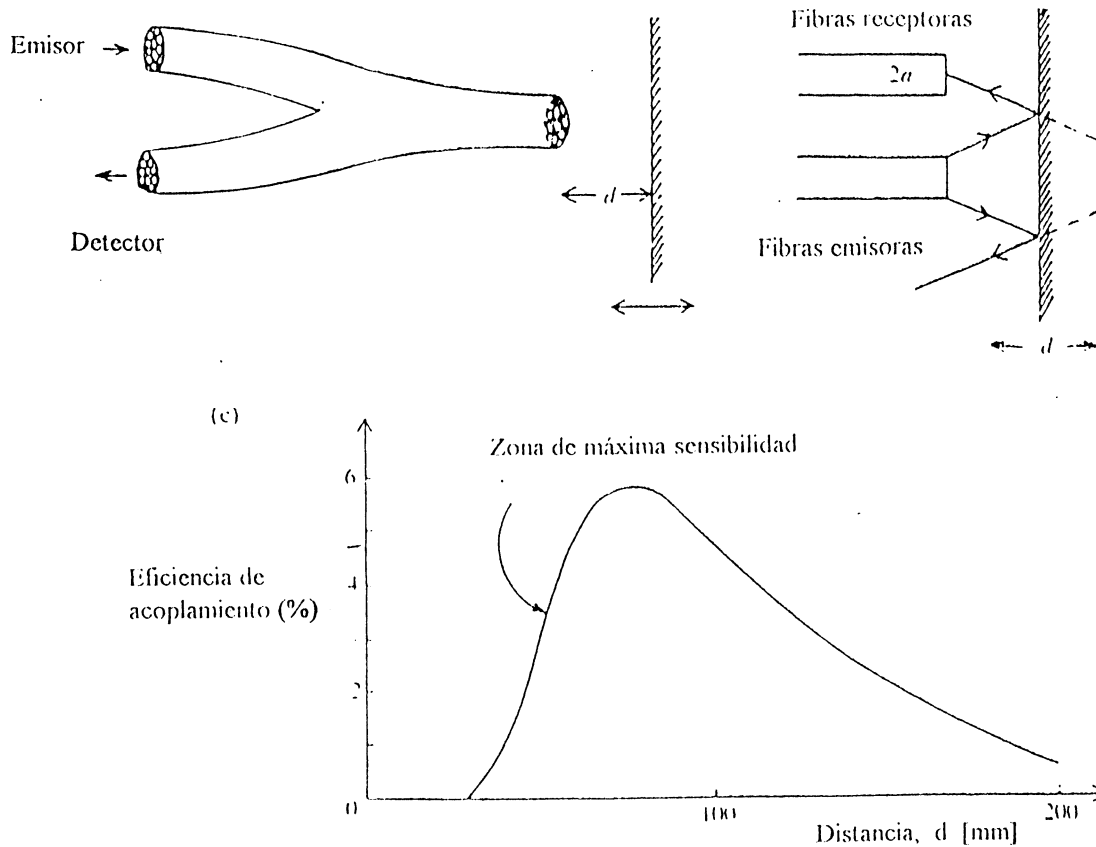


Figura 3: "Fotonic sensor". La curva en (c) fue calculada con $a=s=100 \mu\text{m}$ y $NA=0.4$.

Sensores de presión

Cualquier técnica de medida de desplazamientos puede ser adaptada para medir presiones. Los obturadores utilizados en los diseños mostrados en la figura 2 pueden ser montados sobre una membrana deformable de acuerdo a la presión que soporta. El "Fotonic sensor" puede, a su vez, ser aplicado a una pared con las mismas características que la anterior. Debe recordarse, sin embargo, que los sistemas descritos no son lineales.

Un gran avance en este campo fue logrado en un programa de la NASA para el desarrollo de sensores de fibra. El diseño se observa en la figura 4 y el transductor puede medir presiones absolutas en forma lineal y sin problemas de histerisis entre 0 y 40 psi. El transductor consiste de un diafragma el cual es deformado en proporción a la presión aplicada. La deformación del diafragma modula directamente la luz transmitida por la

fibra por el grado de microdoblez producido. Un microdoblez genera un mayor acople de la luz del núcleo a la cubierta aumentando la pérdida de la fibra. Así, una medida de la variación de la potencia transmitida produce una señal proporcional a la presión aplicada.

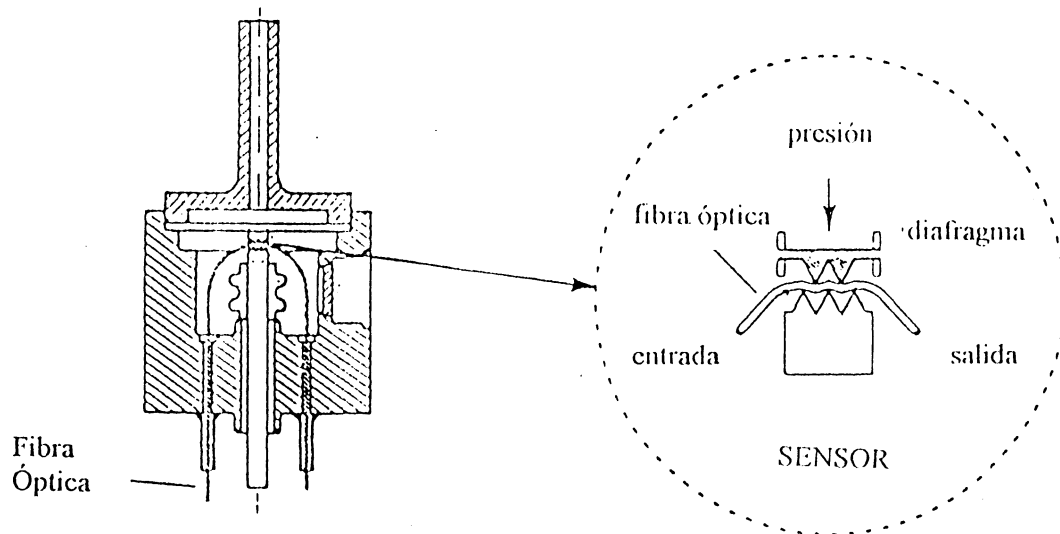


Figura 4: Sensor de presión por microdoblez.

Sensores de temperatura

Existen aquí también diversos diseños de sensores. El indicado en la figura 5 se basa en la variación de la luz transmitida por la fibra debido a la pérdida causada por la onda evanescente en el microdoblez. Dicho cambio es producido por la dependencia con la temperatura del material con que esta construido el núcleo o, mas corrientemente, la cubierta de la fibra. Este sistema fue diseñado para el monitoreo de la temperatura en fuentes de alta tensión, debido a su insensibilidad a las interferencias electromagnéticas.

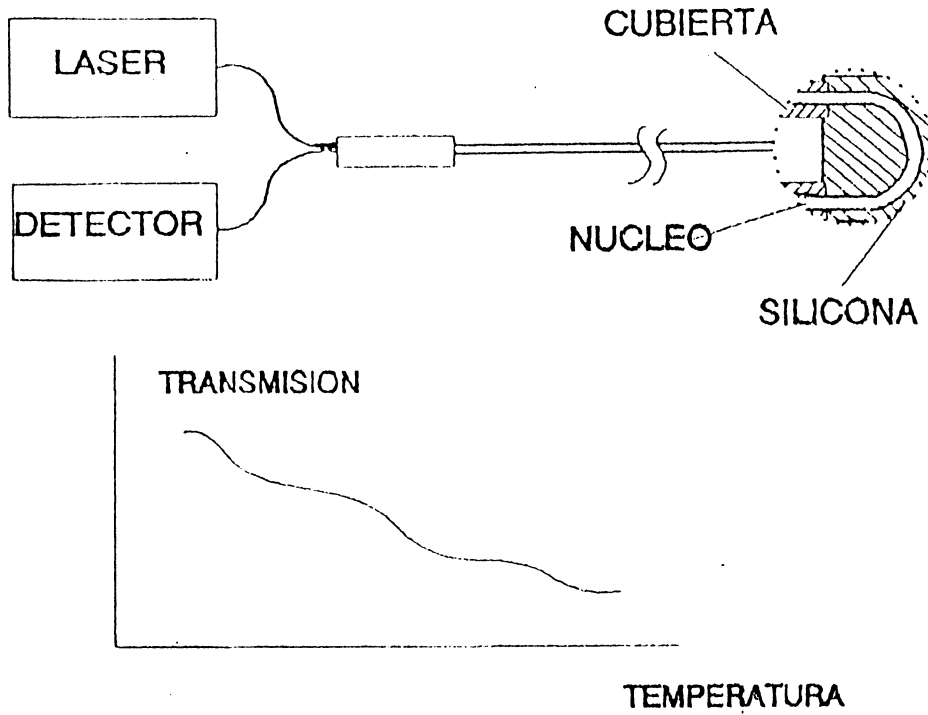


Figura 6: Sensor de temperatura por pérdidas de guiado

Otro tipo de sistema utiliza la dependencia del coeficiente de absorción del AsGa con la longitud de onda y la temperatura. La fuente de luz es un láser de AsGa estabilizado térmicamente.

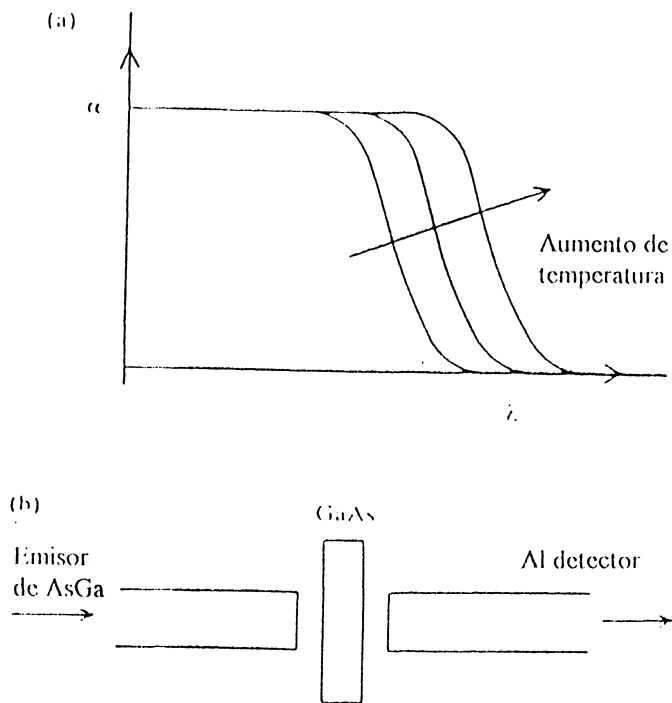


Figura 6: Sensor de temperatura por variaciones de la absorción del AsGa.

Sensores polarimétricos.

Sensor de corriente

De acuerdo a lo expresado, una de las ventajas de estos instrumentos es su insensibilidad a campos electromagnéticos. Bajo situaciones muy particulares, sin embargo, esto no es así. La figura 7 muestra un sensor de corriente basado en el efecto Faraday o efecto magneto-óptico no recíproco, debido al cual la corriente que circula produce un campo magnético cuya dirección coincide con la de propagación de la luz. Dicho campo afecta al material componente de la fibra modificando sus propiedades para la transmisión del haz. Como resultado el plano de polarización del haz de luz es rotado un ángulo θ que es proporcional a la componente del campo magnético paralela a la dirección de propagación de la luz.

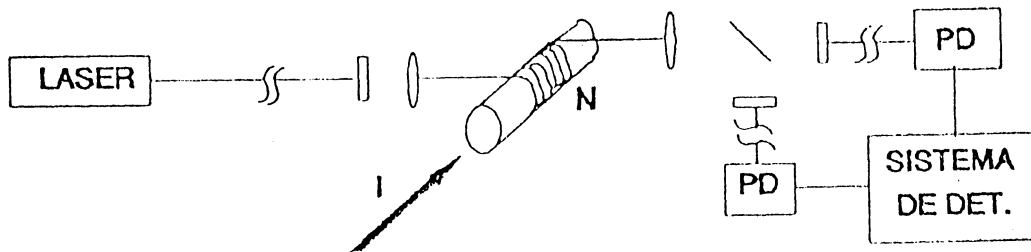
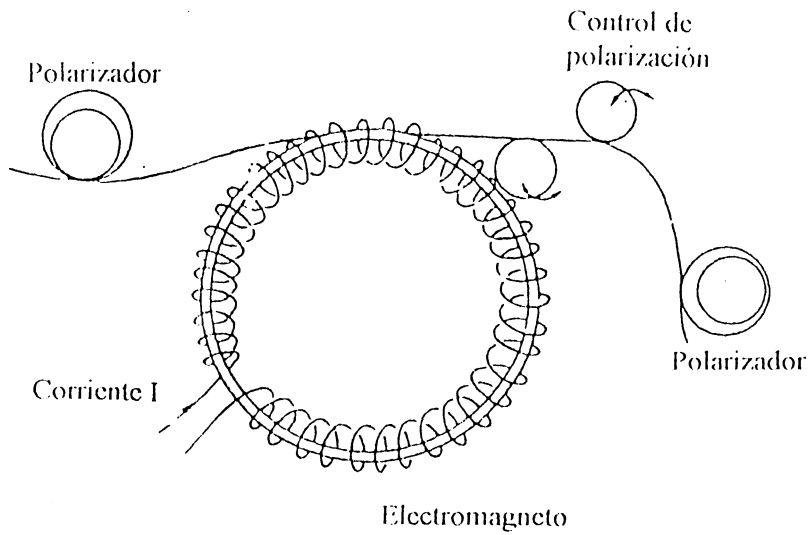


Figura 7: Sensor de corriente (1)

$$\theta = V \cdot B \cdot L = \mu \cdot V \cdot I \cdot N$$

En la ecuación V es la constante de Verdet del material ($V=4 \mu\text{rad}\cdot\text{A}$ para el silicio), B es el campo magnético producido por la corriente I que circula en el conductor y N es el número de vueltas de fibra en el conductor. Dado que el número de Verdet es muy pequeño, este esquema es válido solo para grandes corrientes ($I > 30 \text{ A}$).

El diseño de la figura 8 tiene mayores ventajas pues es un sistema totalmente de fibra y la sensibilidad es aumentada por un factor n igual al número de vueltas del conductor en el rollo de fibra.



$$\theta = \mu \cdot V \cdot I \cdot N \cdot n$$

Figura 8: Sensor de corriente (2)

Sensores de tensión eléctrica

Dos ejemplos de este tipo de sensores, también de tipo polarimétrico, se observan en la figura 9 y ambos utilizan luz inicialmente polarizada. El de la figura 9(a) está basado en el efecto piezoeléctrico. Al aumentar la tensión se expande el piezo comprimiendo la fibra lo cual produce una birrefringencia en esta. El de la figura 9(b) utiliza el efecto Pockels, que origina una birrefringencia sobre el cristal de acuerdo a la tensión aplicada. La birrefringencia producida rota en ambos casos el plano de polarización de la luz la que es analizada a la salida. Se calibra inicialmente la dependencia de esta rotación con la tensión.

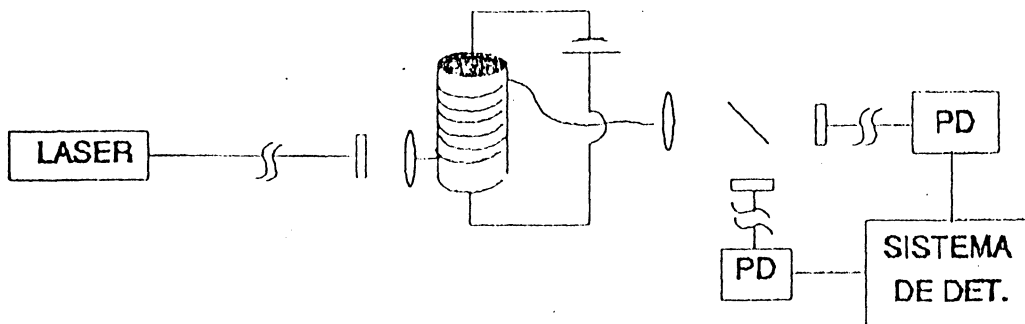


Figura 9(a): Sensor de tensión eléctrica por efecto piezoeléctrico.

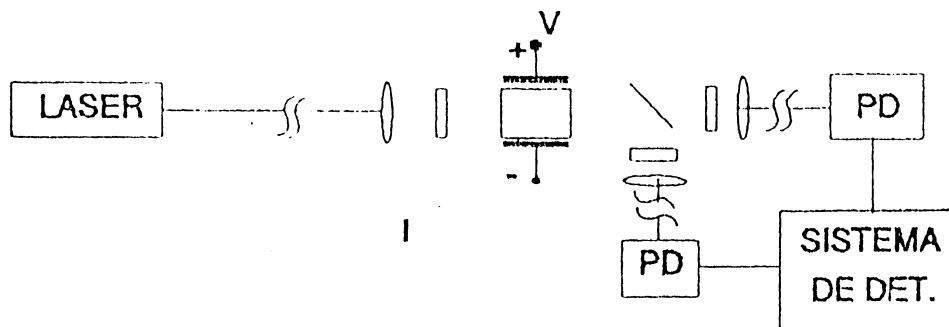


Figura 9(b): Sensor de tensión eléctrica por efecto Pockels.

Sensores interferométricos

Los sensores interferométricos difieren de los anteriores es que es necesario el uso de fibras monomodales y en el hecho de que en ves de la intensidad de la luz, deben ser determinados cambios de fase óptica. Son mucho mas sofisticados que los descritos anteriormente pero ofrecen la potencialidad de ser tremendamente sensibles a perturbaciones externas.

Interferómetro Mach-Zehnder

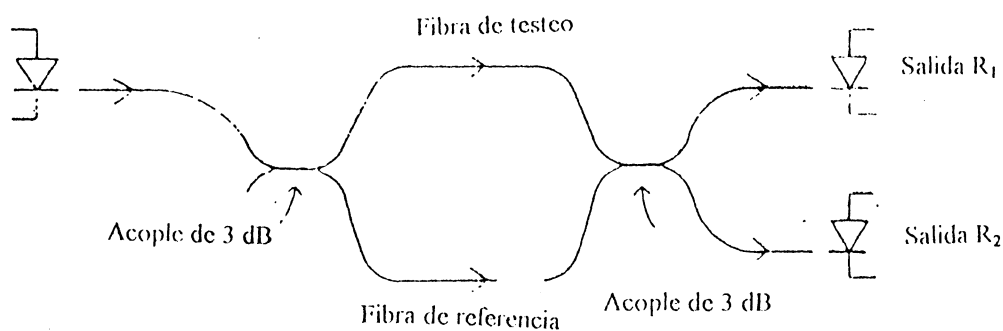


Figura 10: Configuración típica de un interferómetro Mach-Zehnder.

El interferómetro mas común es el Mach-Zehnder (figura 10). La luz colectada inicialmente de la fuente por una fibra es dividida en dos ramas de fibras iguales, una

de las cuales se mantiene aislada de la perturbación externa (referencia) mientras que la otra se diseña para que responda con la mayor sensibilidad posible a la influencia externa. Dicha perturbación induce un cambio de fase en la segunda rama que es observada cuando los dos haces coherentes son recombinados en la salida formando un patrón de franjas cuyo desplazamiento puede ser determinado.

Este tipo de sensores permite medir prácticamente todos los parámetros mencionados anteriormente, como temperatura, presión, corriente, voltaje, etc. Valores típicos son los siguientes, considerando fibra monomodo de silicio fundido y luz proveniente de un láser de HeNe (632.8 nm):

temperatura $\delta\phi / \delta T.L = 107 \text{ radianes/}^\circ\text{C-m}$

presión $\delta\phi / \delta p.L = -4.09 \times 10^{-5} \text{ radian/Pa-m}$

Interferómetro de Sagnac (giroscopo).

En el interferómetro de Sagnac (figura 11), el corrimiento de fase es producido entre dos haces que viajan en direcciones opuestas por un rollo de fibras cuando el rollo está rotando respecto a un eje perpendicular al plano del rollo. La diferencia de fase se relaciona con la diferencia de tiempos que demoran los haces en recorrer el rollo de fibra. Como este está rotando, el tiempo será mayor en el sentido del giro del rollo y menor para el haz contrario.

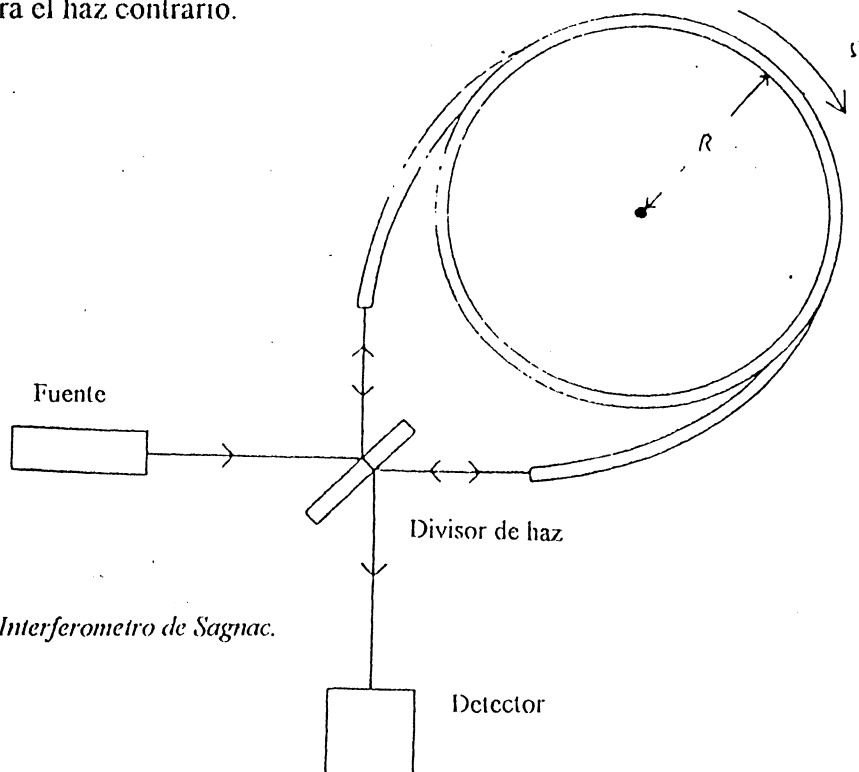


Figura 11: Interferómetro de Sagnac.

La diferencia de fase esta expresada por:

$$\phi_s = \frac{8\pi\Omega\Lambda N}{\lambda_0 \cdot c}$$

donde Ω es la velocidad del rollo, Λ la sección, N el numero de vueltas de fibra y λ_0 la longitud de onda. La medida de la diferencia de fase debe ser mejor que 10^{-7} radianes. Los requerimientos en cuanto a la reciprocidad para ambos haces son fuertes. No se requiere la utilización de láseres muy coherentes debido a que cualquier dispersión en la fibra produciría interferencia que llevaría a señales espurias en el detector. Por ello suelen utilizarse diodos superluminiscentes (diodos láser a los cuales se coloca una perdida alta de energía para hacer que no funcione totalmente como láser).

Sensores de fibra óptica multiplexados

Un ancho rango de técnicas de multiplexado han sido diseñadas y desarrolladas en los últimos diez años. Ello aumenta su potencialidad a la vez que disminuye costos permitiendo una mayor competitividad respecto de las tecnologías convencionales.

Existen diversas aproximaciones en el diseño de estos sistemas, encontrándose también la división entre sensores de intensidad y sensores interferometricos. El multiplexado puede producirse en el dominio temporal (OTDR), en el de frecuencia o en el de longitud de onda. Veamos dos ejemplos.

Multiplexado por división temporal

Están basados en la reflectometría óptica (figura 12). Un pulso láser transmitido por una fibra es parcialmente dispersado hacia atrás por minúsculas imperfecciones en la fibra. Esta característica, que representa una perdida para las comunicaciones ópticas, es aquí aprovechada construyendo trasductores que amplifiquen el fenómeno respondiendo a estímulos externos. Dichos trasductores pueden ser ubicados en distintas regiones de la fibra testeando distintos ambientes. Como las señales detectadas provienen de regiones de la fibra ubicadas a distintas distancias, resultan temporalmente resueltas. La figura muestra un transductor de presión constituido por un sistema de

microdoblado de la fibra. Diversos diseños de estos elementos permiten medir temperatura, acción de agentes químicos como oxígeno, etc. (figura 13 (a) y (b)).

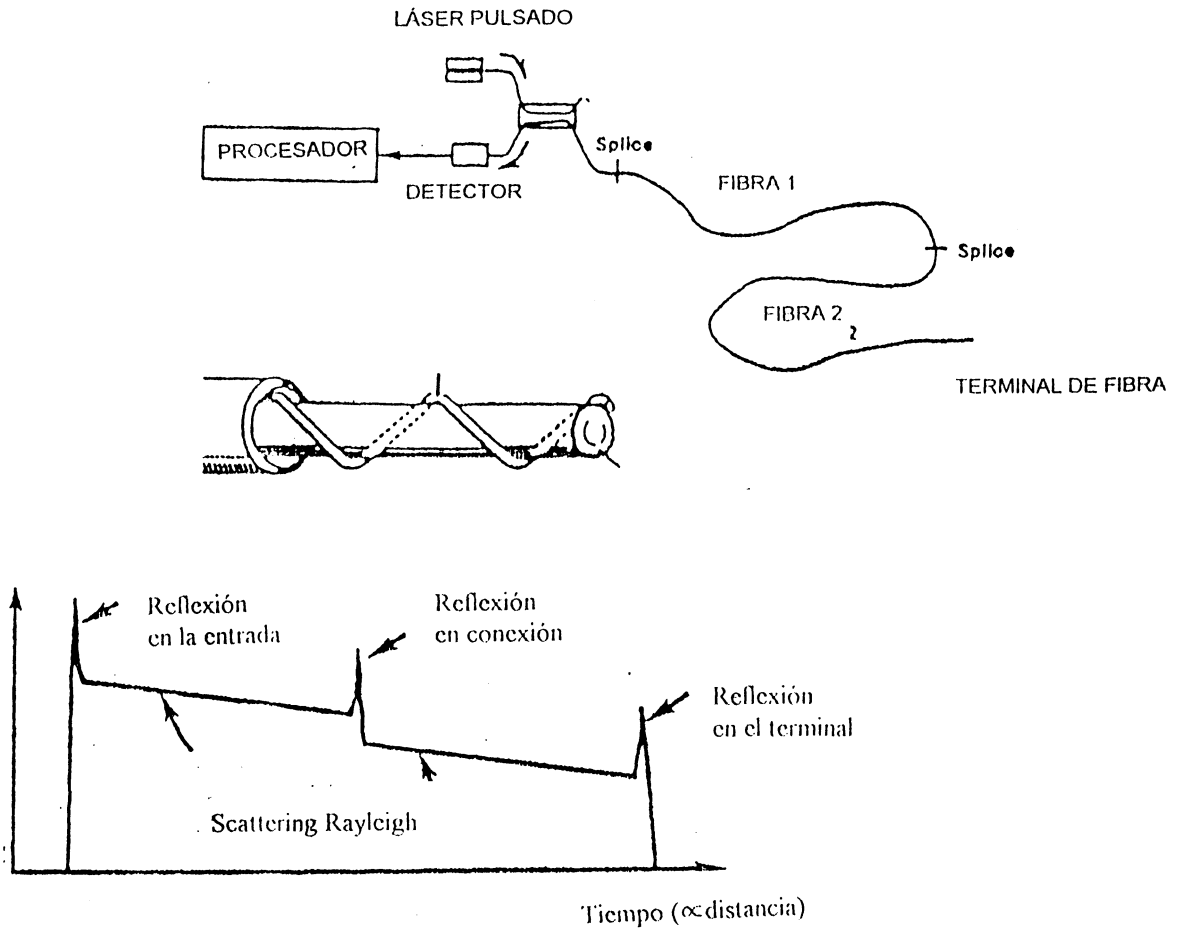


Figura 12: Multiplexado por división temporal

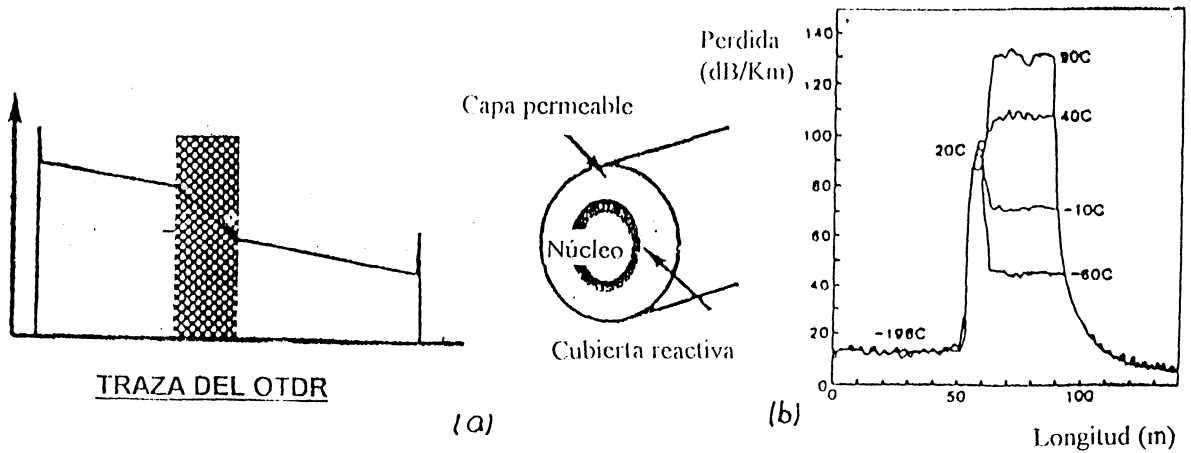


Figura 13: a) Medición de acción de tipo químico y b) medición de temperatura a través de la determinación del nivel de pérdida óptica.

ÓPTICA FÍSICA, COHERENCIA, INTERFERENCIA, DIFRACCION.

L. M. Zerbino

Introducción

Las características especiales de la luz láser son las que le han hecho ocupar un lugar destacado en la investigación científica y en la industria.

La elevada *intensidad* y la gran *direccionalidad* de la luz láser la hacen de por sí interesante para un gran número de aplicaciones.

Otra característica importante es la apariencia granulosa, centelleante, de la luz reflejada por un objeto difusor iluminado con luz láser. Este fenómeno, conocido habitualmente como "speckle", se debe al elevado grado de *coherencia* que tiene este tipo de luz. Es ese elevado grado de coherencia, también manifestado en su generalmente alta monocromaticidad, el que hace del láser una herramienta particularmente poderosa en todas aquellas aplicaciones que involucran mediciones precisas y a distancia, entre las que se destacan los controles por métodos interferométricos.

En el presente capítulo se desarrollarán los tópicos especiales que permitirán una adecuada comprensión del concepto de coherencia y de los fenómenos de interferencia y difracción, que son la base de un gran número de dispositivos de medición y control actualmente en uso en la industria y en la investigación científica y tecnológica, y que permiten vislumbrar un amplio espectro de nuevas técnicas, con futuro en esas mismas áreas.

Ondas luminosas

Según la teoría electromagnética, la luz es de naturaleza ondulatoria, y puede ser interpretada como resultado de la propagación simultánea de una combinación de campos eléctricos y campos magnéticos alternativos. Se puede demostrar que la vibración luminosa en la región del espacio en la que se propaga la luz, puede expresarse como la vibración del campo eléctrico correspondiente.

Utilizando las ecuaciones de Maxwell, las ondas luminosas en el espacio libre se pueden representar por una misma ecuación matemática, representando ya sea el campo eléctrico **E** como el magnético **B** de la onda, ya que, como las componentes de esos campos son complementarias, alcanza con estudiar uno

solo de ellos. La ecuación de propagación para el campo eléctrico es de la forma:

$$\nabla^2 \mathbf{E} = \frac{1}{v^2} \frac{\partial^2 \mathbf{E}}{\partial t^2} \quad (1)$$

donde $(\nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2})$ es el operador laplaciano y t el tiempo.

En los fenómenos que vamos a estudiar, la vibración luminosa puede expresarse como una magnitud escalar U , para la cual podremos escribir la ecuación de onda:

$$\nabla^2 U = \frac{1}{v^2} \frac{\partial^2 U}{\partial t^2} \quad (2)$$

cuya solución es una función $U(x,y,z,t)$ que representa el campo luminoso del punto (x,y,z) en el instante t .

La onda electromagnética más simple que se puede considerar es la onda plana monocromática, que es una onda sinusoidal, de extensión infinita, propagándose en una dirección perfectamente determinada. El campo luminoso para una onda plana monocromática que se propaga en la dirección del eje z puede escribirse:

$$U(x,y,z,t) = A \cdot \cos[2\pi(\nu t - z/\lambda) + \phi] \quad (3)$$

donde A es la amplitud de la onda, en este caso constante en el plano xy , ν es la frecuencia, expresada en hertz, λ la longitud de onda (nm, μm) y ϕ la constante de fase, la cantidad entre corchetes es lo que se denomina fase de la onda, y depende a la vez del tiempo t y de la distancia z sobre el eje de propagación. Representando esta función con el tiempo t , o la coordenada z como parámetro, pueden hacerse más evidentes sus principales características:

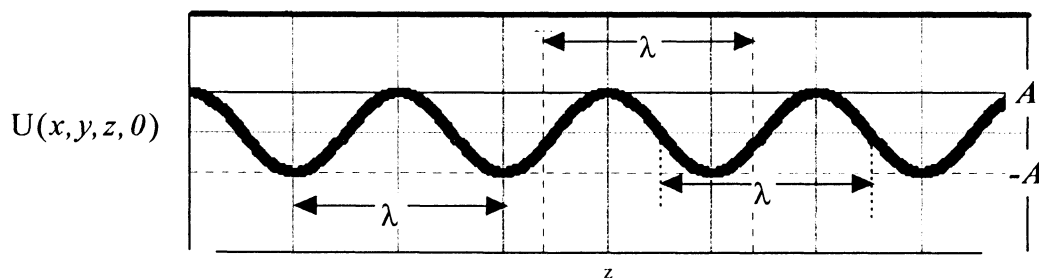


Figura 1

Para simplificar, se ha puesto aquí $\phi = 0$. En el plano $z = 0$ y en todos los planos paralelos a él y que están a una distancia $z = n\lambda$, n entero, el campo luminoso $U(x,y,z,0)$ es máximo y su valor es A . En todos los planos intermedios, según sea el argumento del coseno, el campo toma valores entre A y $-A$. La

cantidad $1/\lambda$ se denomina *frecuencia espacial* de la onda. Poniendo $k = 2\pi/\lambda$, a la cual se designa como *frecuencia espacial angular*, y que se mide en radianes por unidad de longitud, el campo luminoso para $t = 0$ puede escribirse:

$$U(x,y,z,0) = A \cdot \cos(kz) \quad (4)$$

Si en cambio se considera la evolución temporal del campo luminoso U en el plano $z = 0$.

$$U(x,y,0,t) = A \cdot \cos(2\pi\nu t + \phi) \quad (5)$$

y, para $\phi = 0$, se obtiene la figura:

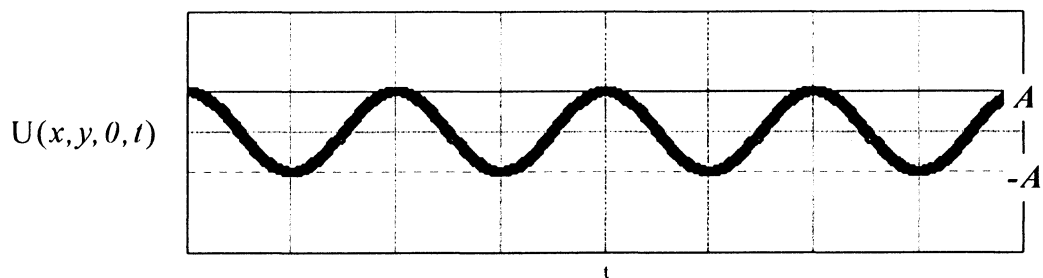


Figura 2

donde ν es la frecuencia de oscilación de U en un punto fijo. Si definimos la frecuencia angular temporal $\omega = 2\pi\nu$, que se mide en radianes por segundo, el campo luminoso en el plano $z = 0$ se puede escribir:

$$U(x,y,0,t) = A \cdot \cos(\omega t) \quad (6)$$

Con estas mismas definiciones, el campo luminoso en la ecuación (3) puede escribirse como:

$$U(x,y,z,t) = A \cdot \cos[\omega t - kz + \phi] \quad (7)$$

A medida que el tiempo transcurre, las superficies de fase constante se desplazan en el espacio. Se llaman superficies de onda, y en este caso simple son planos paralelos al plano $z = 0$, que se desplazan en la dirección z ; la expresión de U es función explícita sólo de z y de t , describe una onda plana de extensión infinita $A(x,y) = cte$ para todo (x,y) , que se propaga en la dirección z .

Si queremos generalizar esta expresión para una dirección cualquiera del espacio, podemos definir un vector propagación \mathbf{k} , cuya dirección y sentido son aquéllos de la propagación de la onda que queremos describir, y cuyo módulo es $k = \sqrt{\mathbf{k} \cdot \mathbf{k}} = 2\pi/\lambda$, y considerando además que \mathbf{r} representa el vector que va desde el origen de coordenadas al punto (x,y,z) , la expresión del campo luminoso de una onda plana monocromática que se propaga en la dirección \mathbf{k} es:

$$U(x,y,z,t) = A \cdot \cos[\omega t - \mathbf{k} \cdot \mathbf{r} + \phi]$$

Recordemos, además que la frecuencia ν y la longitud de onda λ de un campo luminoso que se propaga en el vacío están ligadas por la relación $\lambda\nu = c$, donde c es la velocidad de la luz en el vacío. En cualquier otro medio, $\lambda\nu = c/n = v$, donde n es el índice de refracción y v la velocidad de la luz en el medio en cuestión.

Cuando una fuente luminosa es de pequeñas dimensiones, o se encuentra muy alejada, podemos considerarla como una fuente puntual, es decir, podemos considerar que el campo luminoso producido por esa fuente tiene frentes de onda que son esferas concéntricas, con centro en el punto en el cual está ubicada la fuente. En ese caso decimos que las ondas emitidas por la fuente son ondas esféricas. Si la fuente está en el origen, el campo luminoso que emite se puede escribir:

$$U(x,y,z,t) = \frac{A}{r} \cdot \cos[\omega t - \mathbf{k} \cdot \mathbf{r}]$$

Muchas veces es cómodo escribir las ecuaciones anteriores en notación compleja, de la forma:

$$\begin{aligned} U(x,y,z,t) &= A(x,y) \cdot (\cos[\omega t - \phi(x,y)] + i \cdot \text{sen}[\omega t - \phi(x,y)]) = \\ &= A(x,y) \cdot \exp(i[\omega t - \phi(x,y)]) \end{aligned}$$

En todas estas expresiones, el factor $\exp(i[\omega t])$ indica la naturaleza ondulatoria de la luz, y aparece en todos los cálculos. En los casos en que se consideren ondas luminosas de la misma frecuencia ω , se puede trabajar simplemente con la amplitud del campo luminoso y escribir:

$$U = A \cdot \exp(-i \cdot \phi(x,y)) \text{ para la onda plana;}$$

$$U = (A/r) \cdot \exp(-i \cdot \phi(x,y)) \text{ para la onda esférica.}$$

Esta notación es cómoda ya que, si las operaciones de que se trata son lineales, todos los cálculos pueden realizarse mediante la función compleja. La magnitud física que se desea será en ese caso la parte real de las expresiones finales encontradas mediante el cálculo. Por ejemplo, la amplitud real del campo luminoso es el módulo de la amplitud compleja y la intensidad de la vibración es proporcional al cuadrado del módulo de la amplitud compleja, es decir,

$$U = \text{Re}(U), \text{ donde Re quiere decir parte real, y}$$

$$I = U \cdot U^*, \text{ donde } * \text{ quiere indicar conjugado.}$$

Una fuente luminosa capaz de emitir vibraciones monocromáticas tales como las que hemos descrito, es decir ilimitadas, es una fuente de radiación **temporalmente coherente**. Es un caso límite, y en la práctica no existen fuentes capaces de emitir tales vibraciones. Todas las fuentes luminosas reales emiten vibraciones de duración finita y no pueden ser rigurosamente monocromáticas.

Una fuente de luz blanca emite todas las radiaciones del espectro, es decir, todas las frecuencias, todos los colores. Haciendo pasar esta luz a través de filtros, adquiere un cierto grado de monocromaticidad que depende del filtro. La descarga eléctrica en un gas emite una luz que está compuesta, en general, por una serie de frecuencias determinadas, que son sus componentes espectrales. El grado de monocromaticidad de una fuente de este tipo, o de una de sus rayas espectrales, está caracterizado por su ancho espectral $\Delta\nu$.

La emisión de una fuente láser es, en general, más monocromática que cualesquiera de estas rayas espectrales. Un láser monomodo emite la radiación más monocromática que se pueda hasta ahora construir, pero tiene, sin embargo, un cierto ancho de línea.

En todas las experiencias en las que intervienen fuentes luminosas que se comportan como fuentes puntuales, se dice que existe *coherencia espacial*. Las fuentes luminosas reales tienen siempre dimensiones finitas, no son fuentes puntuales. Una fuente puntual ideal sería una fuente luminosa esférica, en cuyo interior la intensidad es constante y en cuyo exterior la amplitud luminosa disminuye, a medida que nos alejamos de la fuente, y cuyo radio tendiera a cero. A una cierta distancia r de esta fuente, suficientemente grande frente a las dimensiones de la fuente, las ondas esféricas que emite tienen un radio muy grande, y puede ser asimilada a ondas planas. Todos los puntos de una sección normal se encuentran en un mismo frente de onda, están en fase, y son espacialmente coherentes.

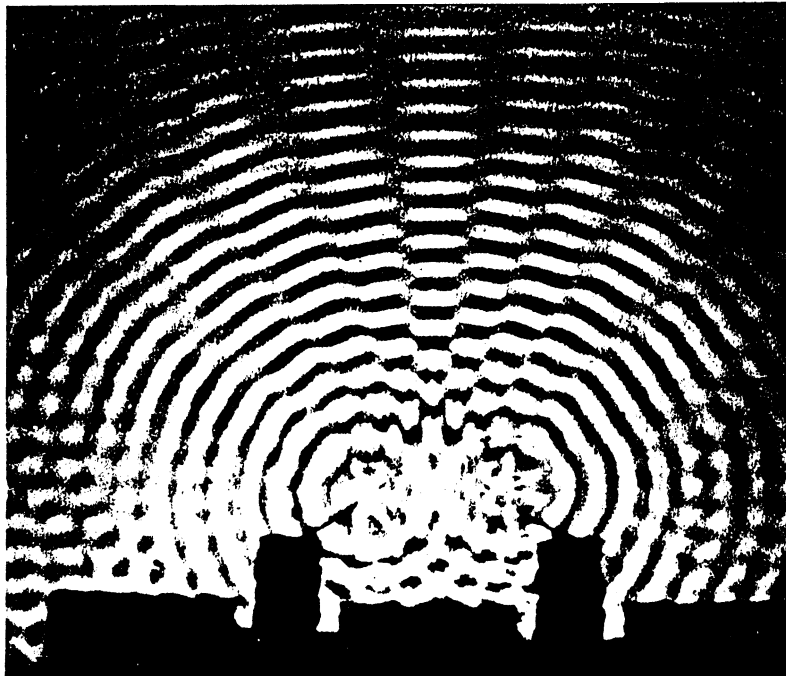
La duración temporal τ de la vibración se llama tiempo de coherencia, y su longitud $c\tau$ es la longitud de coherencia. Las lámparas de descarga ordinarias tienen longitudes de coherencia de unos pocos milímetros, las fuentes luminosas monocromáticas clásicas como las lámparas espectrales de isótopos de baja presión llegan a tener longitud de coherencia de hasta 1 m, son cuasimonocromáticas, mientras que para los láseres, la longitud de coherencia puede pasar los kilómetros. Este elevado grado de coherencia de la radiación láser es una de las características que la hace más atractiva para las aplicaciones metrológicas.

Superposición de ondas

Como dijimos al principio, cada componente del campo de una onda electromagnética $(E_x, E_y, E_z, B_x, B_y, B_z)$ satisface la ecuación de onda diferencial escalar como (2). Una característica muy importante de esa expresión es que es *lineal*, es decir, que U y sus derivadas aparecen solamente con la primera potencia. Esto implica que *cualquier combinación lineal de soluciones individuales* de la ecuación también será una solución de la ecuación. Esta propiedad se conoce como el *Principio de Superposición*: *La perturbación resultante en cualquier punto de un medio es la suma algebraica de sus ondas constitutivas separadas. Si varias ondas luminosas se superponen o coexisten en una cierta región del espacio, el campo eléctrico en cada punto de superposición será el resultante de la suma de los campos eléctricos correspondientes a cada onda presente. De la misma manera, los*

campos luminosos producidos por cada una de las ondas individuales en un punto dado del espacio y en un instante dado, se sumarán para dar como resultado el campo luminoso en ese instante y en ese punto.

Si suponemos contar por un momento con dos fuentes luminosas puntuales ideales que emiten vibraciones ilimitadas, monocromáticas, paralelas, podemos considerar los campos escalares en la región del espacio en la que las ondas se superponen. Si analizamos la amplitud de campo o la intensidad luminosa sobre una pantalla colocada delante de ellas, podemos encontrar dos tipos distintos de fenómenos, según que las frecuencias de la luz emitida por las fuentes sean diferentes o que las dos emitan en la misma frecuencia. En el primer caso se trata de *batidos ópticos*, que solamente se pueden detectar si las frecuencias son muy parecidas y se encuentran en la zona de las ondas de radio. En el segundo caso, se trata del *fenómeno de interferencia*: cuando las frecuencias son iguales, como las velocidades de propagación son las mismas, en un punto cualquiera del espacio la diferencia de fase entre las vibraciones que provienen de cada una de las fuentes que emiten ondas ilimitadas permanece constante en el tiempo, y sólo depende del punto considerado. En las regiones en las que esas vibraciones llegan en fase, las amplitudes se suman y hay un máximo de intensidad. En otras regiones, las vibraciones están en oposición de fase y se restan destructivamente, de manera de producir allí un mínimo de intensidad. Esto constituye el fenómeno de interferencia, y su constatación es una medida de la *coherencia* entre las fuentes que la producen.



Superposición de ondas en agua.
Las perturbaciones provienen de fuentes puntuales

Fotografía 1

En realidad, podría decirse que el fenómeno de interferencia se produce siempre que se dé la superposición de ondas en el espacio, pero la detección del fenómeno sólo es posible en la medida en que las regiones de máxima y mínima intensidad estén definidas en el espacio un lapso de tiempo suficientemente prolongado, compatible con los instrumentos de detección de que se disponga. Independientemente de esta discusión, casi filosófica, se acostumbra a definir el grado de interferencia que existe entre dos ondas, o lo que es lo mismo, el grado de coherencia de las fuentes que las producen, en función de la visibilidad o el contraste de las franjas, definido por medio de la expresión

$$\gamma = \frac{I_{max} + I_{min}}{I_{max} - I_{min}}$$

INTERFERENCIAS

Una de las maneras de explicar el fenómeno de "speckle" es a partir del fenómeno de interferencia. El primero en observar y estudiar este fenómeno fue Thomas Young, quien en 1801 enunció el *principio de interferencia* a partir de sus observaciones.

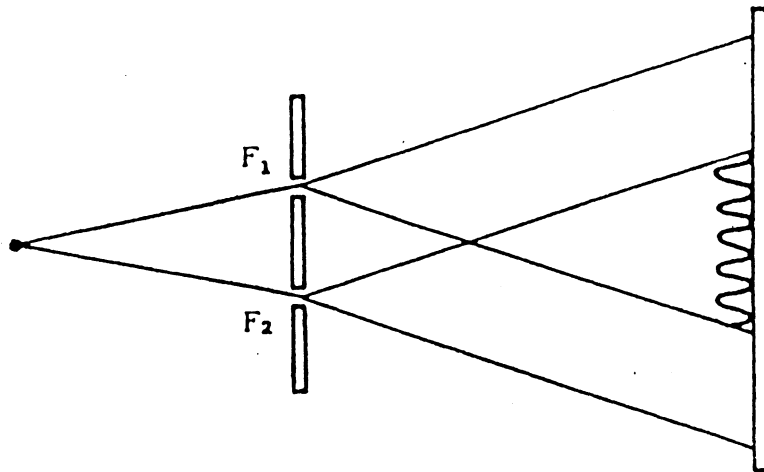


Figura 3

La experiencia de Young consiste esencialmente en iluminar con luz supuestamente monocromática una pantalla opaca en la que se han realizado dos pequeñas perforaciones F_1 y F_2 separadas una distancia d .

A una cierta distancia de la pantalla es posible observar que la distribución de intensidad luminosa sigue bandas paralelas alternativamente oscuras y brillantes, que se llaman *franjas de interferencia*. En los puntos en los que la intensidad luminosa es máxima, el campo luminoso es el doble del que habría si en la pantalla hubiera un solo orificio. En los puntos en los que la intensidad es mínima (franjas oscuras), los campos luminosos están en oposición de fase y por tanto el campo resultante es nulo. Es decir, esta distribución de intensidad es debida a la superposición de las ondas

provenientes de cada uno de los dos orificios existentes en la pantalla y que, llegando a ciertos puntos de observación con una diferencia de fase de π , se suman destructivamente, haciendo que luz+luz= oscuridad.

Habiendo descrito el fenómeno, escribiremos las ecuaciones matemáticas por medio de las cuales es posible analizar sus características fundamentales.

Los dos campos, que por simplicidad supondremos de igual amplitud, pueden escribirse:

$$U_1 = A \cdot \exp(-i \cdot [\omega t - kz + \phi_1])$$

$$U_2 = A \cdot \exp(-i \cdot [\omega t - kx \cdot \text{sen}(\theta) + \phi_2])$$

Por el Principio de superposición, el campo total en el punto (x,y,z) al tiempo t , será:

$$U(x,y,z,t) = U_1 + U_2$$

Como ni el ojo ni los detectores más rápidos pueden seguir las variaciones del campo (del orden de los 5×10^4 Hz), la magnitud observable es la intensidad luminosa, que se obtiene haciendo el valor medio temporal del cuadrado del campo, es decir:

$$\begin{aligned} I(x,y,z) &= \frac{1}{T} \int_0^T U^2(x,y,z,t) \cdot dt = U \cdot U^* = \\ &= 2 \cdot A^2 (1 + \cos[kx \text{sen} \theta + (\phi_1 - \phi_2) + kz[\cos \theta - 1]]) \end{aligned}$$

Entonces, si la pantalla está en $z = 0$

$$I(x,y,0) = 2 \cdot A^2 [1 + \cos(kx \text{sen} \theta + \Delta\phi)], \quad \text{donde hemos puesto } \Delta\phi = (\phi_1 - \phi_2)$$

La intensidad luminosa varía como una función $1 + \cos[\cdot]$ y tiene un gráfico como el de la Figura 4, que consiste en una sucesión de franjas claras y oscuras.

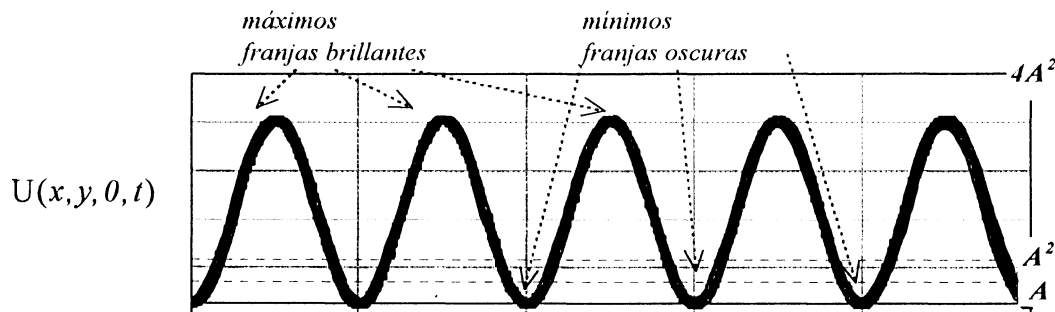


Figura 4

La interferencia óptica se puede definir como una interacción de dos o más ondas luminosas que se superponen en un mismo lugar del espacio y que producen allí una intensidad resultante que no es la suma de las intensidades de las ondas individuales.

El conjunto de estas franjas constituye la *figura de interferencia*. Se produce porque las dos ondas tienen sus fases ligadas por una relación que es independiente en el tiempo. Por esto son importantes para comprender los fenómenos de interferencia, los conceptos de *diferencia de marcha* y de *fase relativa*.

Cuando en un punto determinado del espacio dos ondas oscilan como $\omega \cdot t + \phi_1$ y $\omega \cdot t + \phi_2$, decimos que entre ellas hay una diferencia de fase o fase relativa $\phi_1 + \phi_2$ radianes. Si esta diferencia de fase es igual a un múltiplo par de π , decimos que en ese punto las ondas están en fase, y la interferencia es *constructiva*. Si la diferencia de fase es un número impar de π radianes, decimos que las ondas están en oposición de fase, y la interferencia es *destructiva*. En un punto dado del espacio, la fase relativa depende de dos factores:

1. La fase relativa de las ondas en las fuentes de origen, y
2. La diferencia de camino entre los caminos ópticos recorridos por las ondas para llegar al punto que se está considerando.

Si suponemos por un momento que la fase relativa entre las ondas emitidas por las fuentes en el origen es cero, y la distancia óptica entre punto considerado y la fuente F_1 difiere de la distancia óptica entre el punto y F_1 en una cantidad δ , la fase relativa entre las dos ondas en el punto considerado es $k\delta = 2\pi\delta / \lambda$. Una diferencia de marcha o de camino óptico de $\lambda/2$ provoca un cambio de fase relativa de π , y hace cambiar la intensidad en un punto de un valor máximo a un valor mínimo. Es por esto que las técnicas de interferencia son la base de excelentes métodos de medida de variaciones de distancia.

Por lo dicho anteriormente, la posibilidad de medir la coherencia, parcial o total, se basa en las relaciones entre *coherencia e interferencia*.

La luz emitida por los láseres presenta diferentes grados de coherencia. Si la fuente que ilumina los dos orificios en una experiencia de Young es suficientemente direccional, como para ser considerada como una fuente puntual, y si los dos orificios están suficientemente cerca uno del otro, las ondas luminosas que emergen de ellos presentan las mismas variaciones temporales, y es posible observar franjas de interferencia sobre una pantalla colocada delante de ellos. Si la visibilidad de las franjas, que nunca alcanza a 1, es diferente de cero, decimos que la luz que llega a los orificios es parcialmente coherente. Si las franjas dejan de verse, o si su visibilidad es cero, decimos que la luz es totalmente incoherente,

La coherencia de una onda en la dirección perpendicular a la de su propagación puede inspeccionarse desplazando uno de los orificios sobre el plano de la pantalla y manteniendo el otro fijo. La superficie en la cual puede

desplazarse el orificio móvil sin que las franjas de interferencia desaparezcan, es lo que se llama *área de coherencia*. Este valor mide la *coherencia espacial o transversal* de la onda en cuestión.

Los átomos de una fuente luminosa emiten vibraciones que no son ilimitadas. La emisión se hace por *trenes de onda* y hay una relación entre su longitud y la composición espectral de la luz emitida. En la Figura 5 se muestra un esquema que podría servir para visualizar este concepto.

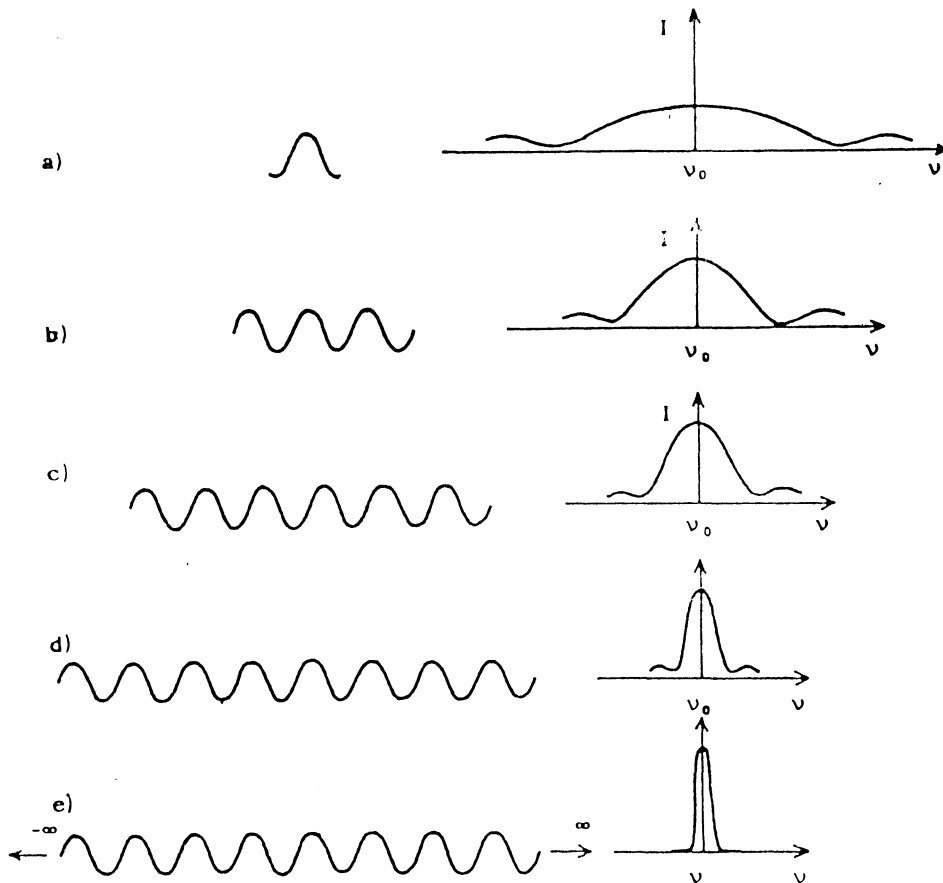


Figura 5

Las sinusoides graficadas representan trenes de onda finitos de distinta longitud, salvo en el caso e) que representa un tren de onda cuya longitud tiende a ∞ . En la parte derecha se representan las curvas de distribución espectral correspondientes a cada uno de los trenes. La frecuencia ν_0 es la frecuencia medida del espectro emitido. En el límite, si un tren, de ondas fuera infinito, tendría una sola frecuencia, es decir, sería monocromático. El grado de monocromaticidad de una fuente está relacionado entonces con su coherencia temporal, es decir, con su coherencia en la dirección de propagación.

La coherencia de una onda en la dirección de su propagación se mide comparando el estado de la onda en un punto en un instante dado y el estado de la onda en el mismo punto en un instante posterior. Experimentalmente puede controlarse mediante un dispositivo como el de la Figura 6, en el que por medio de separadores de haz SH y espejos móviles E_1 y E_2 es posible superponer una

fracción de la onda con otra de la misma onda, pero que ha recorrido una distancia $dz = c\Delta t$.

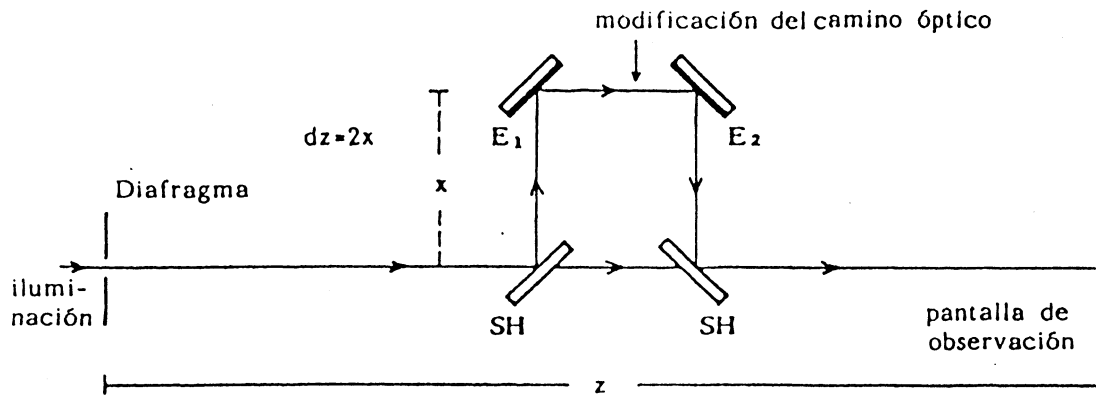


Figura 6

Una parte de la luz llega a la pantalla luego de un recorrido que dura un tiempo z/c y la otra luego de un recorrido $(z+dz)/c$, y variando dz se puede estudiar el grado de coherencia temporal de la luz para diferentes retardos, midiendo la visibilidad de las franjas de interferencia. A medida que dz aumenta, la visibilidad de las franjas disminuye, y el valor $dz=\ell$ para el cual las franjas comienzan a desaparecer se llama *longitud de coherencia*. Conociendo la velocidad de la luz en el medio de que se trate, es posible calcular el tiempo de coherencia como $\tau=\ell/c$.

Para ondas cuasimonocromáticas, la longitud de coherencia es muy grande. El ancho espectral $\Delta\nu$, la longitud de coherencia ℓ y el tiempo de coherencia τ se relacionan a través de:

$$\Delta\nu = \frac{c}{\ell} \quad \text{y} \quad \tau = \frac{1}{\Delta\nu}$$

Interferómetros

En realidad, la coherencia se mide habitualmente por medio de interferómetros. De entre los muchos diferentes sistemas ópticos que producen o se basan en la interferencia, comentaremos sólo algunos de los más importantes. Para estudiarlos, vamos a dividirlos en dos grandes grupos: aquellos en los que se produce una *división del frente de onda* y aquellos en los que la *división es de amplitud*. Por otra parte, haremos la diferencia entre aquellos en los que la superposición es entre dos haces luminosos, y aquellos en los que se superponen más de dos haces de luz. Como ejemplos de ellos analizaremos en particular el interferómetro de Michelson, el de Max-Zender y el de Twyman-Green, entre los que superponen sólo dos haces luminosos, y el de Fabry-Perot como ejemplo de superposición de múltiples haces. Por último, describiremos los interferómetros de "shearing".

Interferómetro de Michelson:

Este interferómetro, cuyo esquema está representado en la Figura 7, permite una medida muy directa de la longitud de coherencia. La fuente se supone de pequeñas dimensiones y la luz que emite se colima por medio de una lente L y se divide en dos haces de igual intensidad por medio de un semiespejo SH. Los espejos E₁ y E₂ reflejan estos haces luminosos haciéndolos incidir nuevamente sobre el semiespejo SH, del cual emergen en la misma dirección y sentido, superponiéndose en toda la región del espacio a partir de SH. Es evidente que la diferencia de fase entre estos dos haces en la zona de superposición es la que corresponde a una diferencia de camino óptico $ds=2n(L_2-L_1)$, donde n es el índice de refracción del medio para la longitud de onda utilizada.

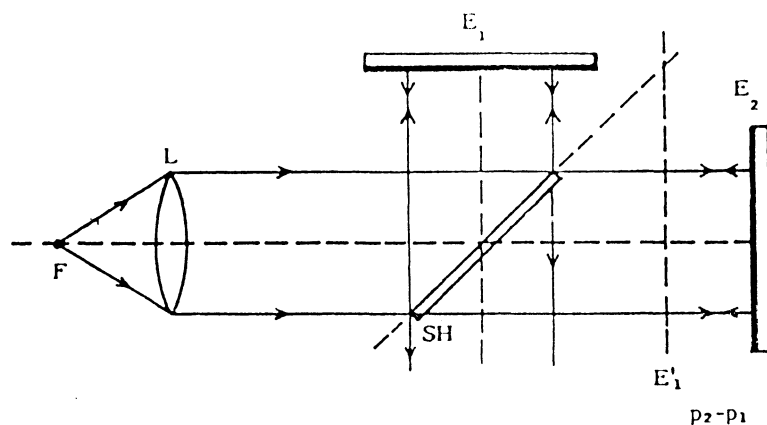


Figura 7

Si el sistema está perfectamente alineado, y el haz de iluminación es en buena aproximación una onda monocromática, la amplitud del campo luminoso y la intensidad luminosa a la salida tendrán valores constantes, para todo (x,y,z) , que dependerán solamente de la diferencia de camino óptico ds , que serán máximos para $ds=0$ o $ds= m\lambda$ y mínimos para $(2m+1)\lambda/2$.

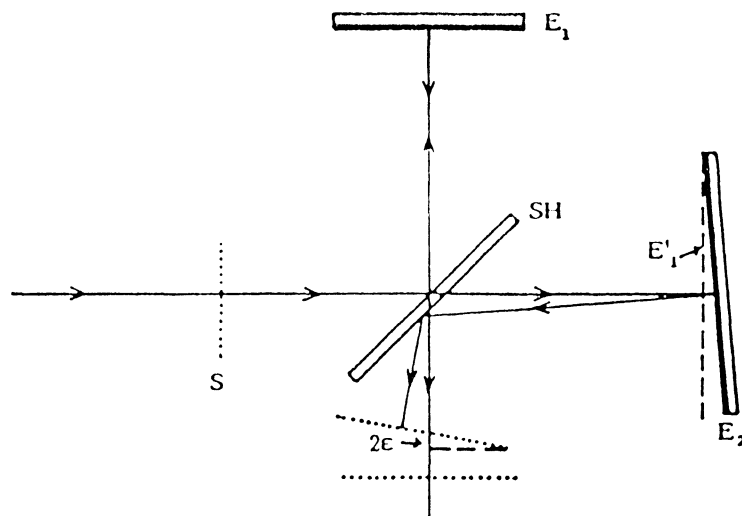


Figura 8

Si $ds=0$ y el sistema está ligeramente desalineado, como en la figura anterior, de manera que los haces emergentes del dispositivo formen entre ellos un pequeño ángulo ϵ , sobre la pantalla de observación existirá un sistema de franjas paralelas cuya repartición sinusoidal de intensidad dará para visibilidad un valor 1, ya que hemos supuesto que las amplitudes de los haces eran iguales y que, por haberse obtenido de un mismo haz original y recorrido el mismo camino, son perfectamente coherentes. Si se desplaza uno de los espejos, este desplazamiento introduce una diferencia de camino óptico ds entre los dos haces, que produce un desplazamiento global del sistema de franjas sobre la pantalla. Si la luz utilizada no es perfectamente monocromática, a medida que la diferencia de camino aumenta se observará también una disminución de la visibilidad de las franjas, ya que si bien las ondas que interfieren provienen siempre de la misma fuente original, han sido emitidas en tiempos diferentes. Es así como esta técnica constituye una medida directa de la longitud de coherencia de las fuentes utilizadas. Otra aplicación directa de este dispositivo es la medida de desplazamientos. Con el advenimiento del láser, cuya longitud de coherencia alcanza muchas veces valores elevados, el campo de aplicaciones del interferómetro de Michelson se amplió considerablemente, ya que, iluminado con luz láser, no necesita de una alineación demasiado cuidadosa y se pueden introducir diferencias de camino considerables sin que la visibilidad de las franjas decaiga demasiado. Es importante citar su rutinaria utilización en la medida precisa de pequeños desplazamientos y pequeñas distancias, tanto en el laboratorio como en la industria .

Interferómetro de Mach-Zender :

Es un interferómetro de dos haces, como el de Michelson, con la particularidad que, como puede verse en el esquema de la Figura 9, los brazos están espacialmente bien separados. Esto lo hace práctico en el estudio de variaciones de índice de refracción y perturbaciones ópticas del medio. De hecho, una de las aplicaciones de este interferómetro es en investigaciones aerodinámicas.

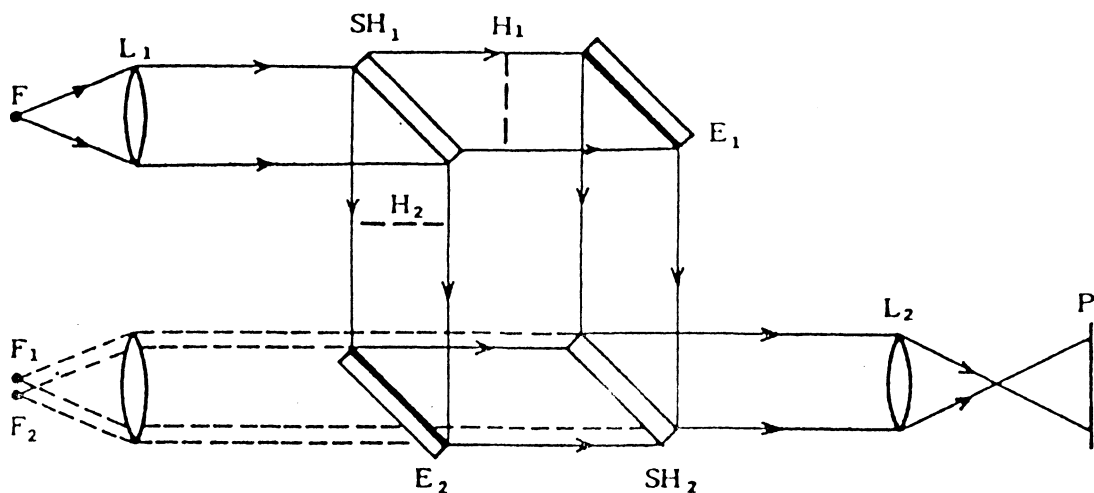


Figura 9

Una lente colimadora L_1 envía un frente de onda plano sobre un semiespejo SH_1 , y se obtienen de esta forma dos ondas planas S_1 y S_2 que se guían mediante espejos E_1 y E_2 y se recombinan mediante el semiespejo SH_2 . Los dos haces emergentes de SH_2 se enfocan mediante la lente L_2 de manera de observar la interferencia sobre una pantalla P . Estos dos haces emergentes parecen provenir de dos fuentes: F_1 y F_2 que son las imágenes virtuales de la fuente F dadas por los brazos del interferómetro. Orientado adecuadamente F_1 y F_2 es posible lograr una separación conveniente entre F_1 y F_2 de manera tal que sobre la pantalla P aparezca un sistema de franjas de interferencia rectas equiespaciadas, alternativamente brillantes y oscuras, como las que se obtienen al hacer una experiencia de Young con dos orificios de pequeño diámetro.

Cuando se utiliza este tipo de dispositivo en estudios aerodinámicos en túneles de viento, lo que interesa investigar es el flujo de aire alrededor de un obstáculo, como por ejemplo un modelo de ala de avión. El problema es averiguar la distribución de presión alrededor de un obstáculo. Como el índice de refracción del aire η es una función de la presión p , según la expresión:

$$\eta = 1 + 0,0003 \frac{p}{H} \quad ,$$

donde H es la presión atmosférica, es posible utilizar este dispositivo para hacer un estudio de la presión, investigando la distribución de índice de refracción.

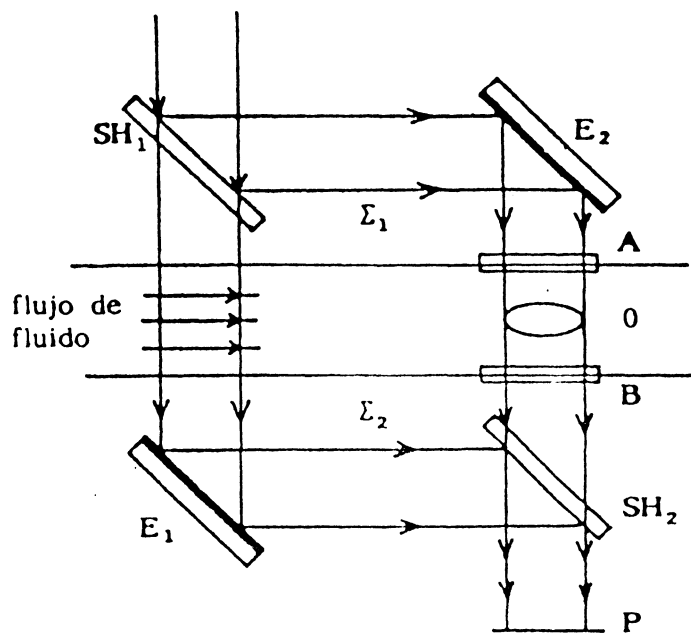


Figura 10

La Figura 10 muestra un esquema de una adaptación del interferómetro de March-Zender para analizar la distribución de presión alrededor de un obstáculo 0 colocado en el interior de un túnel de viento. Para esto se intercala dicho túnel en uno de los brazos del interferómetro, entre dos láminas de vidrio planoparalelas de alta calidad, que sirven como ventanas de observación. Las variaciones de presión introducidas por el obstáculo y que se traducen en variaciones del índice de refracción y, por consiguiente, del camino óptico

recorrido por la luz en uno de los brazos, y de las diferencias de camino entre los dos haces, se manifiestan en un cambio en la distribución de las franjas de interferencia que se observan en P. Del análisis de los interferogramas y la aplicación de la fórmula dada, es posible deducir la distribución de presión alrededor de 0.

La sensibilidad de este instrumento, así como la de todos los interferómetros que se basan en la superposición de dos haces, depende de la geometría elegida. De todas maneras siempre alcanza valores altos, y varía entre algunas longitudes de onda y algunas décimas de la longitud de onda, y dan resultados precisos en el rango de la décima de micrómetro. Ya que la longitud de onda constituye el patrón en el que se basa la unidad de medida, es fundamental utilizar como fuentes de iluminación láseres estabilizados en frecuencia, para hacer los resultados confiables. Al mismo tiempo, son altas las exigencias de calidad de los instrumentos ópticos utilizados y del poder resolvente de los medios de detección y registro. Las técnicas de interferometría holográfica que se describen en otro capítulo son una alternativa para la medición en el mismo rango de trabajo y con precisión comparable, y que implican menores exigencias en cuanto a la calidad de los instrumentos utilizados, pero grandes requerimientos de estabilidad del ámbito donde se realiza la experiencia.

Interferómetro de Twyman-Green

En el interferómetro de Michelson, la interferencia se debe a la superposición de frentes de onda planos; en consecuencia, la figura de interferencia que se obtiene consiste en una sucesión de franjas rectilíneas. Desde este punto de vista, el diagrama de interferencia así obtenido, puede ser utilizado para medir la planitud de los frentes de onda que se superponen, que puede ser afectada en cada una de las reflexiones totales o parciales que sufren dentro del dispositivo. Si se conoce perfectamente la forma del frente de onda de uno de los dos haces, es posible, a partir del diagrama de interferencia, deducir la forma del otro haz. Este comentario puede servir para dar una idea de cómo modificaciones convenientes del interferómetro de Michelson pueden servir para controlar la curvatura y calidad de superficies ópticas, tanto refractoras como reflectoras. Un interferómetro de Michelson así modificado, recibe el nombre de interferómetro de Twyman-Green y tiene un esquema como el de la Figura 11.

Como en el interferómetro de Michelson, la fuente luminosa F se coloca en el plano focal de un buen objetivo O_1 , de manera que la onda luminosa que incide sobre el semiespejo separador de haz SH tenga en una buena aproximación, un frente de onda plano. Uno de los brazos del interferómetro, provisto de un buen espejo plano E_1 , provee el haz de referencia que, superpuesto al haz proveniente del otro brazo del interferómetro, permitirá obtener sobre una pantalla de observación, un diagrama de interferencia que dará cuenta de la forma del frente de onda del haz que emerge del interferómetro luego de haber recorrido este segundo brazo. En el caso de la Figura 11, se ha representado el dispositivo aplicado al estudio de la calidad de un objetivo O_2 . Este elemento se ha colocado en el brazo del interferómetro y el espejo E_2 es

esférico, convexo, y esta colocado de manera tal que su centro de curvatura coincida con el foco de O_2 . De esta forma, si el objetivo O_2 fuera perfecto, la curvatura del frente de onda introducida por él en el haz que incide en E_2 sería justamente compensada por la sucesiva reflexión en E_2 y refracción en O_2 , y así el haz incidente en SH, luego de recorrer el brazo (2) del interferómetro, tendría frente de onda plano, y en la pantalla P la iluminación sería uniforme. Si el objetivo O_2 bajo test presentara aberraciones, el frente de onda que lo atravesara sufriría deformaciones que darían como resultado una iluminación no uniforme en la pantalla de observación P. La figura de interferencia sobre esta pantalla describe la forma de un frente de onda afectado del doble de las deformaciones que el objetivo O_2 introduce a una onda plana, debido a que el haz correspondiente atravesó dos veces dicho elemento refractor. De esta manera, las franjas de interferencia que se observan, resultan apartarse de las rectas en una cantidad que es el doble de la deformación introducida por O_2 . Esto último implica que el dispositivo tiene más sensibilidad que el interferómetro de Mach-Zender, pero que la interpretación de los diagramas de interferencia es más complicada.

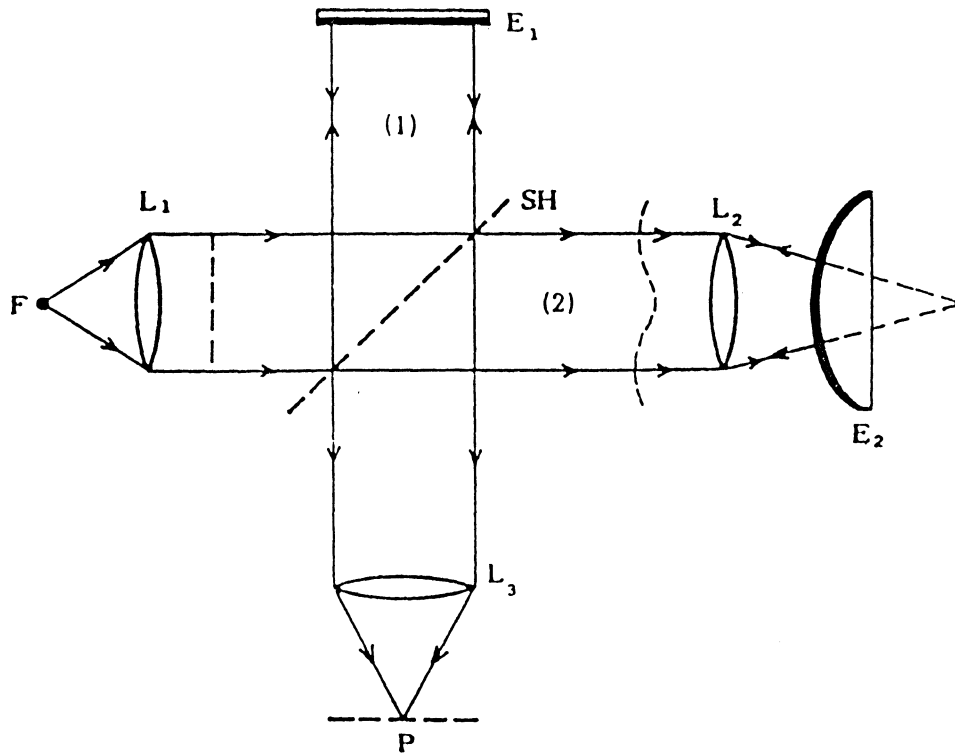


Figura 11

Interferencia de haces múltiples

Estudiaremos el caso en el que se produce interferencia entre un gran número de ondas, originadas por sucesivas reflexiones en el interior de una cavidad formada por ventanas semi-reflectantes.

Para analizar el fenómeno, comenzaremos por estudiar las reflexiones múltiples en el interior de la cavidad formada por dos superficies planas semi-reflectantes paralelas, separadas por una distancia e como la que se esquematiza en la Figura 12.

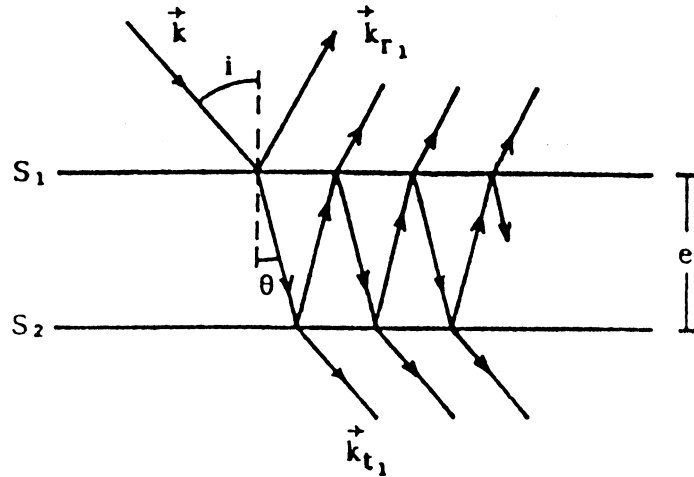


Figura 12

Sea k el vector propagación de una onda plana de amplitud A que incide bajo un ángulo i sobre la superficie S_1 , y sean k_{rj} y k_{tj} los vectores propagación correspondientes a las R_j ondas reflejadas y T_j ondas transmitidas luego de haber sufrido múltiples reflexiones dentro de la cavidad. Las amplitudes correspondientes de cada uno de estos campos luminosos se pueden escribir:

Para R_0 :	A	$\cdot r'_1$	
Para R_1 :	$A \cdot t_1 \cdot t'_1$	$\cdot r_2$	$\cdot \exp(-i2\phi)$
Para R_2 :	$A \cdot t_1 \cdot t'_1$	$\cdot r_2^2 \cdot r_1$	$\cdot \exp(-i4\phi)$
Para R_3 :	$A \cdot t_1 \cdot t'_1$	$\cdot r_2^3 \cdot r_1^2$	$\cdot \exp(-i6\phi)$
Para R_j :	$A \cdot t_1 \cdot t'_1$	$\cdot r_2^j \cdot r_1^{(j-1)}$	$\cdot \exp(-i2j\phi)$

y

Para T_1 :	$A \cdot t_1 \cdot t_2$		$\cdot \exp(-i\phi)$
Para T_2 :	$A \cdot t_1 \cdot t_2$	$\cdot r_2 \cdot r_1$	$\cdot \exp(-i3\phi)$
Para T_3 :	$A \cdot t_1 \cdot t_2$	$\cdot r_2^2 \cdot r_1^2$	$\cdot \exp(-i5\phi)$
Para T_j :	$A \cdot t_1 \cdot t_2$	$\cdot r_2^{j-1} \cdot r_1^{(j-1)}$	$\cdot \exp(-i(2j-1)\phi)$

donde los coeficientes complejos de amplitud $t_1, t'_1, t_2, r'_1, r_1, r_2$ corresponden a:

- t_1 transmisión de la superficie S_1 para la luz incidente desde el exterior;
- t'_1 transmisión de la superficie S_1 para la luz incidente desde el interior;
- t_2 transmisión de la superficie S_2 para la luz incidente desde el interior;
- r'_1 reflexión de la superficie S_1 para la luz incidente desde el exterior;
- r_1 reflexión de la superficie S_1 para la luz incidente desde el interior,
- r_2 reflexión de la superficie S_2 para la luz incidente desde el interior.

Recordemos que t_1, t'_1 y t'_2 son de la forma :

$$t_j = \tau \exp(-i\psi), \quad \text{donde } \tau \text{ es la fracción transmitida de la amplitud, y } \psi \text{ el retardo de fase al atravesar la superficie}$$

$$T_j = t_j \cdot t_j^* \quad \text{transmitancia de la superficie } S_j \text{ definida como:}$$

$$t_j = \text{energía transmitida} / \text{energía incidente,}$$

De modo similar, r_1 y r_2 son de la forma:

$r_j = r \exp(-i\delta)$, donde r es la fracción reflejada de la amplitud y δ el retardo de fase en la reflexión;

$R_j = r_j r_j^*$ la reflectancia de la superficie S_j , definida como:
 $R_j = \text{energía reflejada/energía incidente.}$

además,

$\phi = 2\pi n e \cos \theta / \lambda$ atraso de fase entre las dos superficies,
 (θ es el ángulo de transmisión).

En la reflexión en el vidrio r es real (dado que $\delta = 0$ o $\delta = \pi$)

En la reflexión metálica r es complejo.

Si el medio es transparente t es real, o sea $t = t$

En films metálicos t es complejo.

Si el medio es transparente, usando la ley de conservación de la energía:
 podemos escribir: $R+T=1$,

En la reflexión metálica, una porción de la energía se pierde por absorción, y tenemos: $R+T+A=1$

Tanto los rayos transmitidos como los reflejados son paralelos entre sí y su superposición dará un diagrama de interferencia en el infinito, que de observarse en el plano focal de una lente colectora. Entonces, en un punto cualquiera P sobre el plano focal de la lente o plano de observación, la amplitud resultante es la suma de las amplitudes de todas las N ondas transmitidas:

$$U_p = \sum_{j=1}^N A \cdot t_1 \cdot t_2 \cdot r_2^{j-1} \cdot r_1^{j-1} \exp(-i(2j-1)\phi)$$

Si el número de reflexiones es suficientemente elevado, esta suma se convierte en una serie geométrica, cuyo resultado es:

$$U_p = \sum_{j=1}^N \frac{A \cdot t_1 \cdot t_2}{1 - r^2 \exp(-2i\phi)} \exp(-i\phi)$$

y la intensidad correspondiente será: $I_M = \frac{I_0}{1 + F \cdot \sin^2(\frac{\phi'}{2})}$

que es la llamada formula de Airy,

donde $I_0 = \frac{T^2}{(1-R)^2}$; $F = \frac{4R}{(1-R)^2}$ y $\phi' = 2\phi + 2\delta$

El contraste, o visibilidad de las franjas es, en estos casos:

$$\gamma = \frac{I_{max} + I_{min}}{I_{max} - I_{min}} = \frac{F}{1+F} = \frac{4R}{(1-R)^2}$$

Graficando los perfiles de las franjas de interferencia para distintos valores de la reflectancia R se obtiene:

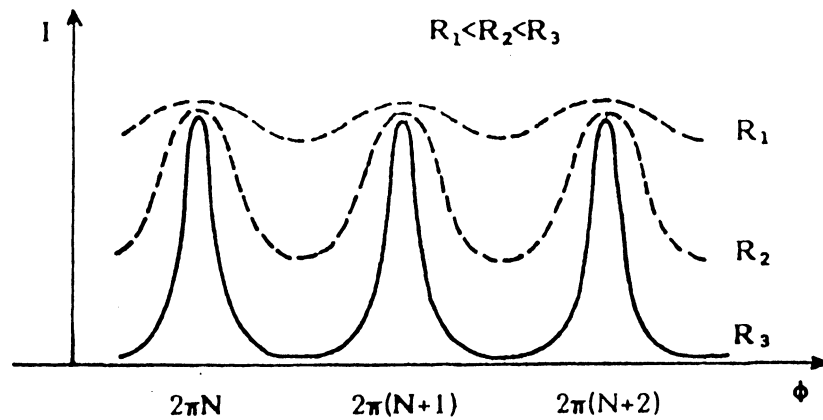


Figura 13

Si la onda que llega a la cavidad es una onda esférica, todas las ondas transmitida también lo serán, y el diagrama de interferencia consistirá en una familia de anillos concéntricos.

Interferómetro de Fabry-Perot :

Este interferómetro se basa en el fenómeno de reflexiones múltiples que acabamos de describir.

El dispositivo adopta varias formas, según sea el medio óptico en el interior de la cavidad, y según que se desee que el espaciado entre las placas sea fijo o regulable, pero responde esencialmente el esquema representado en la Figura 14

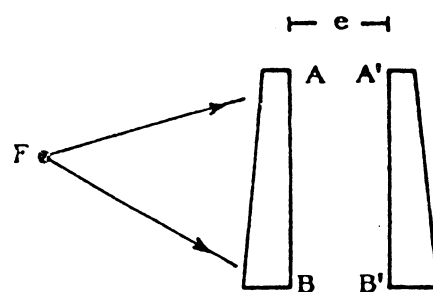


Figura 14

Las superficies internas AB y A'B' tienen una alta reflectancia. Las placas se fabrican de forma levemente prismática, de manera tal que las múltiples reflexiones que pudieran producirse en sus caras externas, y que darían lugar a la aparición de otros dos sistemas de anillos de interferencia caigan fuera del campo de observación y no se superpongan con la familia de anillos que se desea analizar.

Si se ilumina el dispositivo con luz policromática, para cada longitud de onda habrá un sistema de anillos, y es posible observar ninguna figura de interferencia en la superposición. El diagrama de interferencia puede hacerse manifiesto cuando la luz emergente del interferómetro es analizada en un espectrógrafo; el diagrama resultante es el de los llamados espectros acanalados.

Si se ilumina un interferómetro de Fabry-Perot con luz de dos longitudes de onda diferentes λ_1 y λ_2 el diagrama de interferencia consistirá en dos familias de anillos concéntricos. Si las placas del interferómetro son suficientemente reflectantes, los anillos serán muy finos y podrán separarse con ayuda de un espectrógrafo. Este es el principio de la espectroscopía interferencial con un Fabry-Perot, que permite investigar la estructura fina de las líneas espectrales emitidas por una fuente luminosa.

Si en cambio el haz de iluminación es un haz láser expandido, la figura de interferencia consistente en anillos concéntricos, podrá observarse en el plano focal de una lente colocada inmediatamente después del interferómetro. La elevada coherencia de los láseres permite hacer grande el espaciado entre las placas, sin que la visibilidad de las franjas decaiga.

El interferómetro de Fabry-Perot, además de su tradicional aplicación a la determinación de longitudes de onda, es particularmente adecuado para medir con gran precisión espesores ópticos, índices de refracción de sólidos, líquidos y gases, y estudiar rugosidad de superficies.

Interferómetros de "shearing" o desplazamiento:

En esta clase de interferómetros, la interferencia se debe a la superposición de dos frentes de onda, pero no los hemos incluido en el primer grupo ya que estos dos frentes de onda no son independientes, sino que son dos versiones diferentes o modificadas del mismo frente de onda, de manera que no podemos considerar a uno de ellos como un frente de onda de referencia. Podemos distinguir dos tipos de interferómetros de "shearing": de "shearing" lateral y de "shearing" radial.

Como su nombre lo indica, en un interferómetro de "shearing" lateral hacen interferir dos versiones del frente de onda que se desea estudiar, desplazadas lateralmente una con respecto a la otra. Si el desplazamiento se realiza en dirección x , podemos llamar Δx a este corrimiento a "shear" y se puede ver que habrá una franja de interferencia brillante si:

$$s(x+\Delta x,y) - s(x,y) = m\lambda \quad ; \quad \text{con } m \in \mathbb{Z} \text{ (enteros)}$$

donde $s(x+\Delta x,y)$ y $s(x,y)$ son los caminos ópticos recorridos por cada haz.

De manera similar, habrá una franja de interferencia oscura si

$$s(x+\Delta x,y) - s(x,y) = (2m+1)\lambda/2; \quad \text{con } m \in \mathbb{Z} \text{ (enteros)}$$

Estas relaciones pueden interpretarse como el promedio de la derivada del camino óptico respecto a la dirección del desplazamiento multiplicado por ese mismo desplazamiento, es decir, que puede escribirse

Esto pone en evidencia que los interferómetros de "shearing" lateral sólo dan información de la derivada del frente de onda en la dirección del "shear" o desplazamiento. Por lo tanto, si los frentes de onda que se desean estudiar no tienen simetría axial (en el plano xy), deberán realizarse interferogramas con direcciones de desplazamiento diferentes (obviamente, la interpretación será mas sencilla si son ortogonales), a fin de describirlos completamente.

La Figura 15 muestra esquemas de dos ejemplos de interferómetros de "shear" lateral.

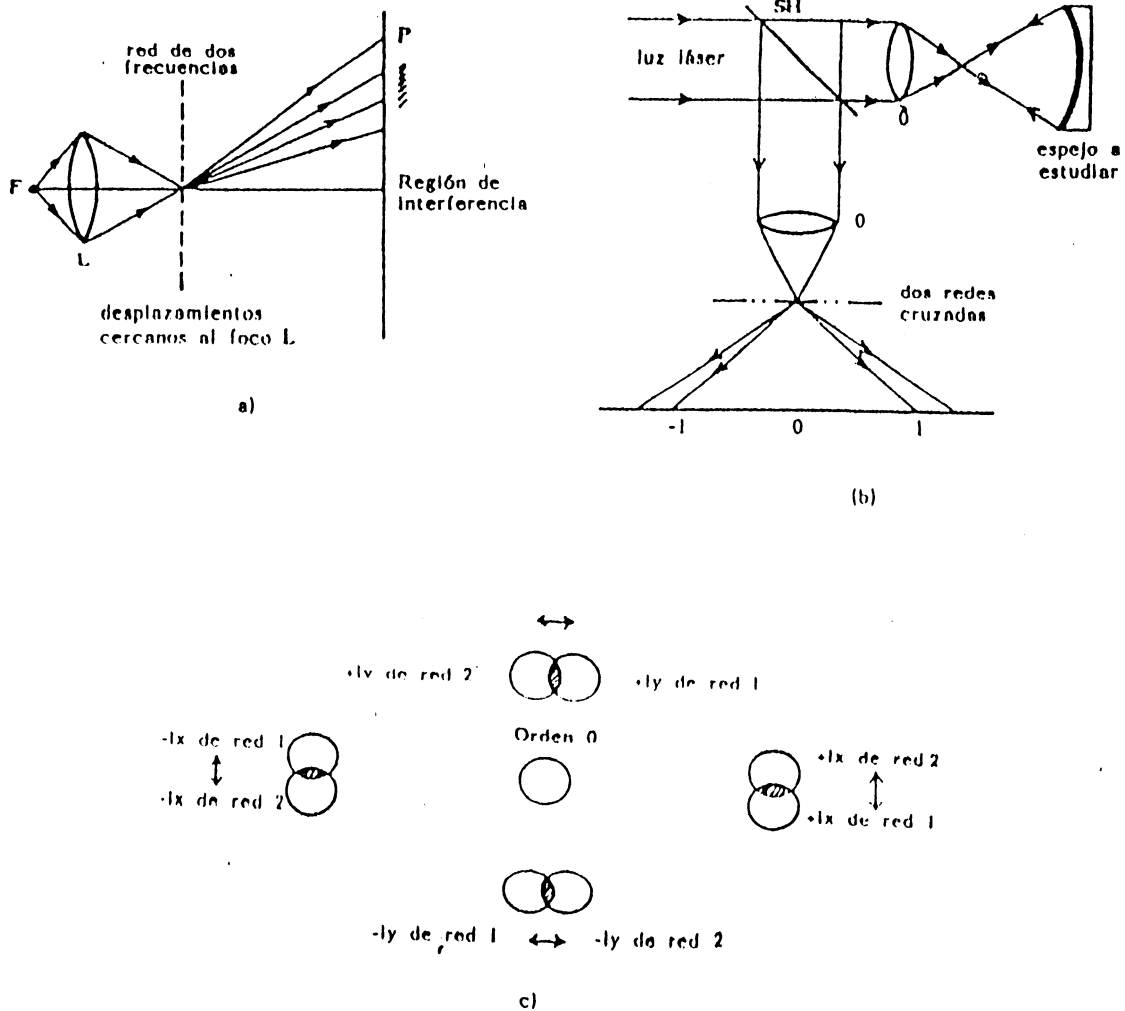


Figura 15

En el de la parte a), la interferencia se realiza por superposición de frentes de onda difractados en el primer orden por dos redes paralelas de frecuencias espaciales distintas o, lo que es lo mismo, por una red que contenga

dos frecuencias espaciales. Si se desea investigar el frente de onda en la dirección perpendicular, habría que superponer dos redes dobles cruzadas perpendicularmente. El dispositivo esquematizado en la parte b) utiliza dos redes cruzadas de la misma frecuencia espacial, rotadas un ángulo pequeño. De esta manera, los cuatro primeros órdenes de una red se superponen parcialmente con los cuatro primeros órdenes de la otra, para dar interferogramas de "shear" lateral en direcciones ortogonales, como se muestra en el diagrama presentado en c).

En el caso de "shear radial", se hace interferir un frente de onda con su versión expandida. Es decir, que cada punto del frente de onda es comparado con su correspondiente, que ha sufrido un desplazamiento en la dirección radial. Puede esquematizarse como en la Figura 16.

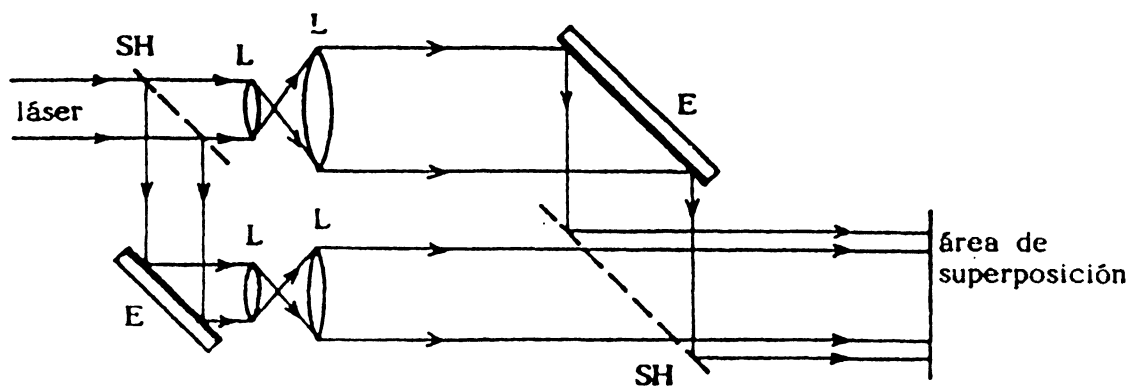


Figura 16

Como se desprende de los esquemas analizados, los interferómetros de "shear" son de implementación simple; y en general, de bajo costo. El hecho de que no utilicen un frente de onda de referencia es en algunos casos una ventaja, pero trae como consecuencia cierta dificultad para deducir, a partir de los interferogramas, la forma del frente de onda que se desea analizar.

Interferometría. Aplicaciones.

Para que el fenómeno de interferencia pueda ser visualizado, el sistema de franjas de interferencia debe ser estacionario, y para esto la relación de fase entre los haces interfirientes debe mantenerse constante en el tiempo. Esto significa que la luz utilizada debe tener un alto grado de coherencia.

Las fuentes luminosas clásicas tienen poca coherencia, tanto espacial como temporal, de manera que su utilización en interferometría implica la necesidad de un filtrado de frecuencias para aumentar su coherencia temporal y un diafragmado, o filtrado de frecuencias espaciales, a fin de hacerlas espacialmente coherentes. Estas operaciones involucran una considerable pérdida de intensidad luminosa, lo que dificulta, en general, la experiencia.

El láser, con longitudes de coherencia que pueden alcanzar varios metros, y con gran monocromaticidad constituye una herramienta sin igual en este tipo de dispositivos. Sin embargo este elevado grado de coherencia de la radiación láser puede traer aparejado algunos inconvenientes o, al menos, imponer ciertas condiciones sobre las características de las componentes ópticas utilizadas en los dispositivos interferométricos ya que, además de los diagramas de interferencia buscados, la superposición de frentes de onda reflejados por las diferentes superficies ópticas atravesadas por la luz dentro del interferómetro, puede dar lugar a la aparición de franjas de interferencia adicionales no deseadas. Es el caso de las reflexiones internas entre las caras de los espejos semi-reflectantes. Así, cuando se trabaja con iluminación láser, en lugar de semiespejos, es conveniente usar prisma-cubos o separadores "pellicle", o recubrimientos reflectantes son aconsejables en una de las caras del separador de haz, a fin de evitar las franjas de interferencia parásitas. Similares precauciones de recubrimientos antirreflectantes son aconsejables en cada una de las superficies ópticas que atraviese la luz. Estos condicionamientos encarecen la experiencia.

Además, es de destacar que ya que la relación $\lambda/\Delta\lambda$ da cuenta del número de franjas con buen contraste que puede esperarse en un fenómeno de interferencia, para una lámpara de vapor de Hg, cuya emisión se filtre con un filtro de ancho de banda de 3,5 nm centrado en 436 nm (violeta), este número es aproximadamente 125, y para un láser de He-Ne ($\lambda=633$ nm, $\Delta\lambda=2 \cdot 10^{-3}$ nm), es del orden de $1,2 \cdot 10^5$, así que, con la utilización del láser, el buen contraste resultante está ampliamente garantizado.

Las técnicas interferenciales son métodos de medida extremadamente precisos. Se las emplea para determinar la planitud o la calidad del pulido de superficies, la forma y calidad de frentes de onda, el paralelismo de superficies, el radio de curvatura de superficies esféricas, las variaciones de índice de refracción del medio, el contenido espectral de las radiaciones luminosas, las variaciones de presión y temperatura en fluidos, la horizontalidad de superficies y el control de desplazamientos entre otras muchas aplicaciones.

Los microscopios interferenciales son instrumentos de uso corriente en los laboratorios y en la industria, para estudiar y analizar objetos transparentes o reflectantes.

Las medidas de longitud por métodos interferenciales son la base de la determinación del metro en longitudes de onda, pero también del funcionamiento de dispositivos para la medición y control de velocidades, aceleraciones y aceleraciones angulares (girómetro), mediante el conteo de franjas.

La gama de dispositivos es amplia, y la selección del método y las componentes que mejor se adapten para cada problema particular depende tanto de las condiciones de trabajo como de las precisiones que se necesite alcanzar.

Combinado con el de difracción, que comentaremos en el siguiente párrafo, el fenómeno de interferencia es la base de la holografía, de la holografía interferométrica, de la interferometría speckle, que se desarrollan en otros capítulos de esta publicación.

Fenómenos de difracción.



Difracción de ondas en una cubeta de agua, ejemplificando el principio de Huygens.
a) Difracción por una rendija. b) Varias fuentes puntuales igualmente espaciadas cubren la apertura y generan a distancia un diagrama similar.

Fotografía 2

Al igual que los fenómenos de interferencia, los fenómenos de difracción dan cuenta de la naturaleza ondulatoria de la luz, y se manifiestan en la medida en que los resultados de la experiencia difieran de los predichos por la óptica geométrica o de rayos.

Supongamos que realizamos experiencias como las esquematizadas en la Figura 17.

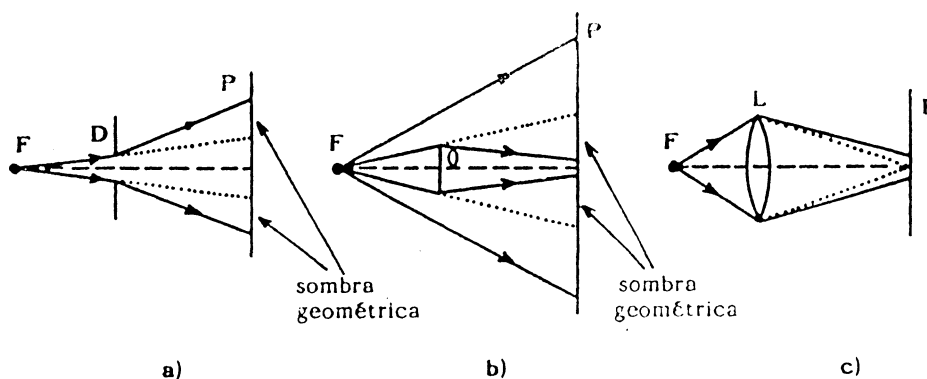


Figura 17

- Parte de la luz proveniente de una fuente F atraviesa un orificio practicado en el diafragma D y es recogida en una pantalla P .
- Parte de la luz proveniente de una fuente F es interceptada por un objeto opaco O y el resto es recogida en una pantalla P .
- La luz proveniente de una fuente F es enfocada por medio de la lente L en un punto P .

Hemos dibujado la marcha de rayos predicha por la óptica geométrica por una parte (línea de puntos) y por otra lo que en realidad puede observarse; que es posible encontrar que la intensidad luminosa no es cero en los lugares en que según la óptica geométrica debería haber

sombras, y que, en general, la distribución de intensidad luminosa difiere de la que podría predecirse de la óptica geométrica. En estos casos podemos decir que tanto el diafragma como los bordes de la lente, como el objeto opaco, difractan la luz.

Así como los fenómenos de interferencia se producen cuando existe una superposición de frentes de onda, los fenómenos de difracción ocurren cuando existe una limitación del frente de onda. El efecto es una característica general de los fenómenos ondulatorios que ocurren siempre que una porción del frente de onda es obstruida u alterada.

En los casos en los que la observación de los fenómenos de difracción se realiza a distancia finita se dice que se trata de difracción de Fresnel, y cuando la observación se realiza en el infinito, o en el plano focal de una lente, se dice que se trata de difracción de Fraunhofer.

Principio de Huygens-Fresnel:

La propagación de una onda luminosa puede analizarse haciendo uso del llamado principio de Huygens (1678), que puede ser enunciado como sigue:

La propagación de un frente de onda puede obtenerse considerándolo como formado por un conjunto de fuentes secundarias puntuales sincronas, es decir, que emiten ondas esféricas en fase. La envolvente, en el instante $t + \Delta t$ de las ondas esféricas emitidas en el instante t por las fuentes secundarias constituirá el frente de onda en el instante $t + \Delta t$.

Este modelo de Huygens se complementa con la hipótesis de Fresnel (1818) según la cual pueden existir interferencias entre las diferentes ondas secundarias. Haciendo uso de este modelo podemos calcular la deformación del frente de onda a medida que se propaga en el espacio, es decir, la difracción. Sin embargo, es importante señalar que el modelo predeciría la existencia de ondas y frentes de onda de retroceso, que no se verifican en la realidad. Desarrollos matemáticos apropiados permiten justificar el principio de Huygens-Fresnel y evitar la existencia de las ondas de retroceso que son contrarias a la experiencia. El esquema de la Figura 18 b) muestra como el modelo Huygens-Fresnel permite explicar la difracción a través de un diafragma.

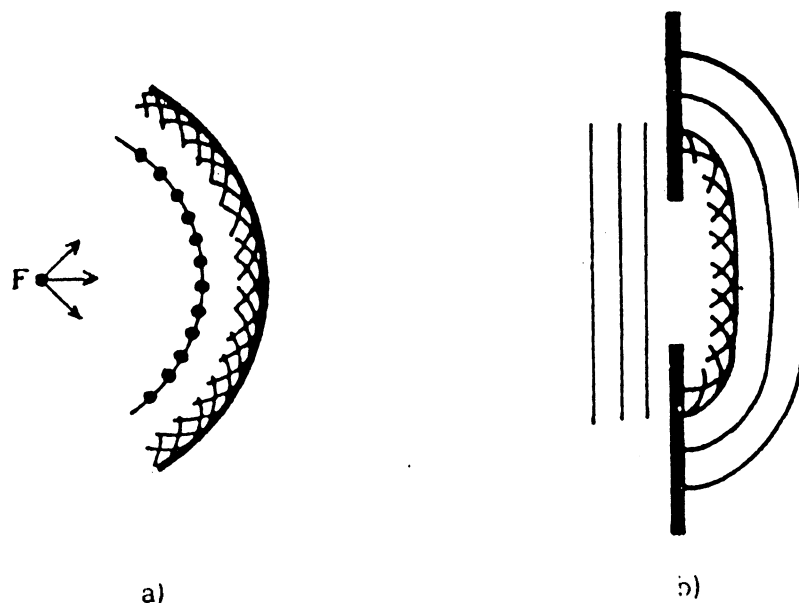


Figura 18

Formulación escalar de Fresnel-Kirchhoff. Aproximaciones:

El principio de Huygens-Fresnel es la base de la formulación matemática desarrollada por Kirchhoff (1882), y permite calcular la amplitud del campo luminoso en un punto (x_0, y_0, z_0) a partir de la amplitud del campo luminoso en un punto $(x, y, 0)$ como los que se señalan en la Figura 19.

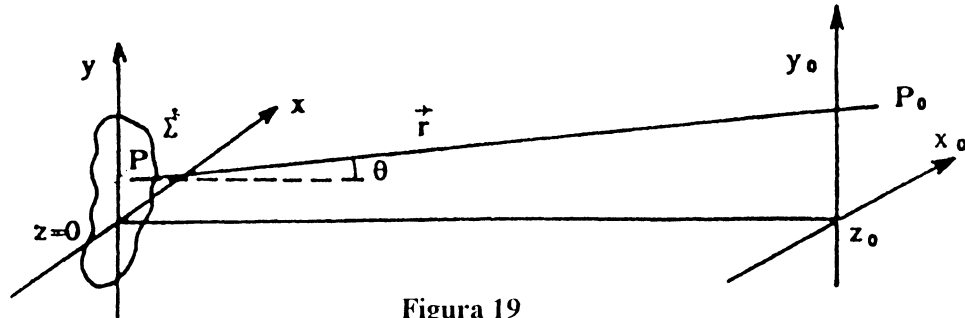


Figura 19

Consideremos que una pantalla opaca plana colocada en el plano $z=0$, que tiene una pupila Σ , es iluminado desde la izquierda por un frente de onda de forma arbitraria, monocromático, de longitud de onda λ . El plano de la pantalla y de la pupila es el de coordenadas $(x, y, 0)$, y el plano de observación está a una distancia z de él y tiene coordenadas (x_0, y_0, z_0) .

Haciendo uso del principio de superposición, el campo escalar en el punto (x_0, y_0, z_0) se obtendrá sumando las contribuciones de cada una de las fuentes individuales en el plano $z=0$, es decir,

$$U(x_0, y_0, z_0) = \frac{1}{i 2 \lambda} \iint_{\Sigma} U(x, y; 0) \frac{\exp(i k r)}{r} (1 + \cos \theta) dx dy$$

donde $U(x_0, y_0; 0)$ es el valor del campo en cada punto del plano $z=0$,

$\frac{\exp(i k r)}{r}$ corresponde a las ondas esféricas secundarias provenientes de las fuentes individuales en el plano $z=0$

$r^2 = (x_0 - x)^2 + (y_0 - y)^2 + z_0^2$ corresponde a la distancia entre fuente puntual y los puntos del plano de observación.

θ es el ángulo entre $\vec{r} = \vec{PP}_0$ y la normal en la dirección z , $k = \frac{2\pi}{\lambda}$.

y donde se ha considerado que las dimensiones de la abertura difractante son mucho menores que z_0 y mucho mayores que λ . La contribución de las fuentes secundarias incluyen un factor de oblicuidad $(1 + \cos \theta)$, una amplitud constante $1/2\lambda$, y un factor de fase de $90^\circ(1/i)$.

Si tanto las dimensiones más grandes en la región de observación como en la apertura difractante son mucho menores que z_0 , es posible simplificar la expresión de manera de obtener

$$U(x_0, y_0, z_0) = \frac{\exp(i k z_0)}{i 2 \lambda z_0} \iint_{\Sigma} U(x, y; 0) \exp \left\{ \frac{i k}{2 z_0} [(x_0 - x)^2 + (y_0 - y)^2] \right\} dx dy$$

que es la llamada **Aproximación de Fresnel**.

Si la transmitancia en el plano $z=0$ vale 0 fuera de la abertura, es posible extender los límites de integración entre $-\infty$ e ∞ y de esta manera esta última integral podría interpretarse como la convolución entre la función exponencial y el campo $U(x,y;0)$.

Para distancias z_0 aun mayores, para las que es posible escribir $z_0 \gg \frac{k[x^2 + y^2]_{max}}{2}$

Se puede escribir la llamada **Aproximación de Fraunhofer** o de difracción en el infinito:

$$U(x_0, y_0; z_0) = C(x_0, y_0; z_0) \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} U(x, y; 0) \exp \{-i2\pi(\mu x + \nu y)\} dx dy$$

donde hemos puesto $\mu = \frac{x_0}{\lambda z_0}$ y $\nu = \frac{y_0}{\lambda z_0}$ las frecuencias espaciales,

y el factor
$$C(x_0, y_0; z_0) = \frac{1}{i 2 \lambda z_0} \exp \left\{ i k \left[\frac{x_0^2 + y_0^2}{2 z_0} + z_0 \right] \right\}$$

Los cálculos se simplifican aquí notablemente ya que, a menos de un factor complejo, el campo difractado está relacionado con el campo en el plano $z=0$ por una transformación de Fourier bidimensional y, como se ve en esta última expresión, las variaciones de la posición de la abertura sobre el plano $z=0$ no cambian el valor del campo en el plano $z=z$ salvo en un factor complejo constante, y por consiguiente, la intensidad no cambia.

Recordemos que U representa la amplitud compleja del campo luminoso en cada punto, de manera que, incluye información sobre la amplitud y fase en cada punto.

La intensidad luminosa estará dada por la expresión

$$I = U(x_0, y_0; z_0) \cdot U(x_0, y_0; z_0)^*$$

Ya que en general lo que interesa calcular es la intensidad luminosa, que es la magnitud observable, se acostumbra hacer caso omiso del factor complejo C en los cálculos de campos luminosos, y se acostumbra a escribir, para los casos de difracción de Fraunhofer,

$$U(\nu, \mu, z) = \mathfrak{F}\{U(x, y)\}_{(\nu, \mu)} \quad \text{donde } \mathfrak{F}\{ \} \text{ significa transformada de Fourier}$$

y la Intensidad
$$I(\nu, \mu; z) = \left| \mathfrak{F}\{U(x, y)\}_{(\nu, \mu)} \right|^2$$

es decir, que la intensidad se calcula como la transformada de Fourier de la función de autocorrelación de $U(x,y)$.

Efecto producido por una lente:

Analizaremos el efecto que produce una lente en la propagación de haz luminoso .

Cuando la luz pasa de un medio transparente (1) a otro (2), su velocidad de propagación cambia, dependiendo del índice de refracción del medio en cuestión. Si el índice de refracción del medio (1) es menor que el del medio (2), la velocidad de propagación se reduce.

Si se intercala una lente en el camino de un haz luminoso, y es posible considerar que la lente es suficientemente delgada como para suponer que la luz que la atraviesa emerge de ella en las mismas coordenadas que en las que entró, la única acción que producirá sobre el campo luminoso será la de retrasar el frente de onda incidente en una cantidad proporcional al espesor del material óptico en cada punto. Esto puede expresarse matemáticamente mediante un factor de fase cuadrático M de forma

$$M(x,y) = \exp \left[-\frac{ik}{2f}(x^2 + y^2) \right], \text{ donde } f \text{ es la distancia focal de la lente.}$$

Si en la expresión del campo difractado que da la aproximación de Fresnel, se reemplaza $U(x,y)$ por $U(x,y) M(x,y)$, lo que implicaría que la lente está en el mismo plano que la pupila difractante o que la pupila difractante contiene tanto el borde de la lente como el cambio de fase introducido por la lente, el campo difractado en el plano focal de la lente se obtendrá poniendo en esa expresión x_0/f y la expresión del campo queda de la forma

$$U(\mu, \nu; f) = C(x_f, y_f; f) \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} U(x, y; 0) \exp \{-i2\pi(\mu x + \nu y)\} dx dy$$

donde las frecuencias espaciales en el Plano focal de la lente son: $\mu = \frac{x_f}{\lambda f}$; $\nu = \frac{y_f}{\lambda f}$

La expresión encontrada a menos del factor complejo C es nuevamente la transformada de Fourier bidimensional del campo $U(x,y)$. Es decir, que en el plano focal de la lente a menos de ese factor de fase se obtiene el espectro de la distribución luminosa bidimensional ubicada delante de la lente. Esto es coherente con el hecho de que habíamos encontrado que la transformación de Fourier nos daba la difracción de Fraunhofer en el infinito, o a muy grandes distancias de la apertura, y que en términos de óptica, el plano focal de una lente corresponde a hacer infinita la distancia de observación. En la figura se esquematiza este resultado.

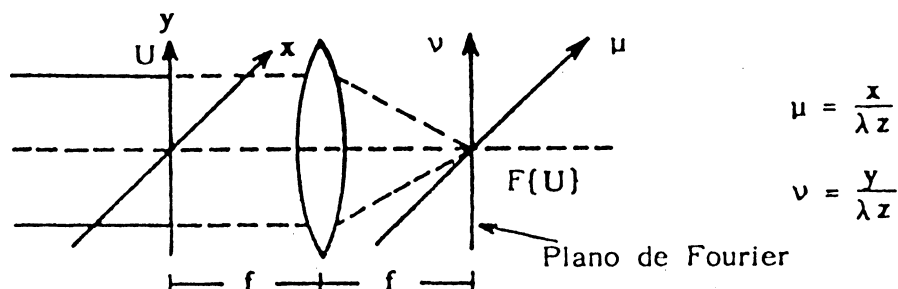


Figura 20

Este resultado es importantísimo y es la base de gran cantidad de dispositivos, ya que hemos encontrado que el plano focal de una lente es el plano de las frecuencias espaciales correspondientes a las coordenadas sobre el plano anterior a la lente. De esta manera en ese plano, llamado también por esto plano de Fourier, es posible realizar operaciones (procesado) filtrando adecuadamente con obstáculos opacos (filtros de amplitud) y/o transparentes (filtros de fase), las distintas componentes del espectro espacial del campo luminoso a la entrada. Es factible entonces, por medio de un dispositivo como el de la Figura 21, reconstruir una imagen

del objeto con las componentes espectrales restantes, no filtradas, y de esta forma realzar detalles: o realizar modificaciones en el contraste de la imagen original.

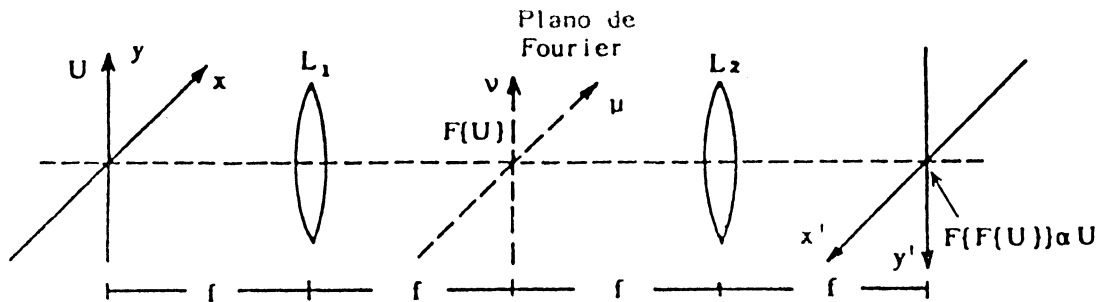


Figura 21

Estas técnicas de procesamiento óptico tienen aplicación en los campos más diversos, como la inspección de circuitos integrados (fotomáscara), detección de defectos en objetos espacialmente periódicos y procesamiento de imágenes de todo tipo.

Difracción de Fraunhofer de una abertura rectangular:

Si la abertura rectangular tiene dimensiones $x'y'$ y está iluminada por una onda plana, tenemos:

$$U(x, y) = \begin{cases} \text{rect}\left(\frac{x}{x'}\right) \text{rect}\left(\frac{y}{y'}\right) = 1 & \text{dentro de la abertura} \\ 0 & \text{fuera de la abertura} \end{cases}$$

Entonces el campo difractado en el infinito, o en el plano focal de una lente puede calcularse como

$$U(\mu, \nu; f) = Q \text{sinc}\left(\pi \frac{\mu x'}{\lambda}\right) \text{sinc}\left(\pi \frac{\nu y'}{\lambda}\right)$$

El factor Q es proporcional a la superficie de la abertura. En general cualquiera sea la forma de la abertura, la amplitud del diagrama de interferencia en un punto, es proporcional a la superficie de la abertura, y por consiguiente, la intensidad es proporcional al cuadrado de la superficie de la abertura.

En el caso de la abertura rectangular, la figura de difracción estará formada por una serie de rectángulos, en los ejes μ y ν , y el perfil de intensidad estará dado por

$$\begin{aligned} \text{sinc}^2\left(\pi \frac{\mu x'}{\lambda}\right) & \text{ en eje } \mu \\ \text{sinc}^2\left(\pi \frac{\nu y'}{\lambda}\right) & \text{ en eje } \nu \end{aligned}$$

El rectángulo central tiene dimensiones (λ/x') (λ/y') y se llama orden central de difracción, y sus dimensiones son inversamente proporcionales a las dimensiones de la apertura. Este también es un resultado general, independientemente de la forma de la apertura.

A medida que nos alejamos del orden central de difracción, la intensidad decae considerablemente. Esto puede verse en el gráfico de la función $\text{sinc}^2\left(\pi\frac{ux'}{\lambda}\right)$

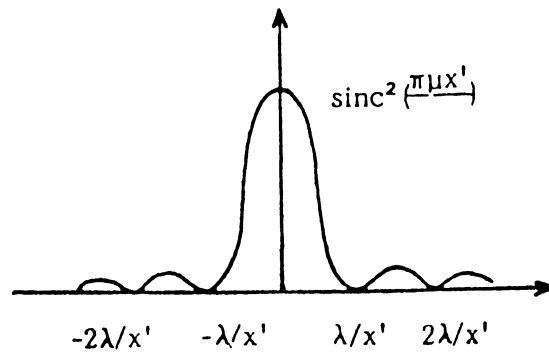


Figura 22

Pantallas complementarias : Teorema de Babinet

Sean U_1 el campo difractado por una pantalla s_1 y U_2 el difractado por una pantalla s_2 , complementaria de s_1 , es decir, cuya transmitancia es opuesta a la s_1 en cada punto. Decimos que s_2 es el "complemento" de s_1 : es opaca para todos los puntos para los que s_1 es transparente y transparente para todos los puntos para los que s_1 es opaca.

Según el principio de Babinet, el campo $U = U_1 + U_2$ será el campo en el plano de observación cuando no se intercale ninguna pantalla difractante. De esta manera, conocido el campo difractado por s_1 es posible calcular el campo difractado por s_2 .

En la Figura 23 se presenta un ejemplo en el que se compara el campo de difracción de una rendija con el de su complemento. La parte a) de la figura muestra el perfil en amplitud de un haz láser y las correspondientes amplitud e intensidad del campo difractado. En b) se representan respectivamente una rendija y la amplitud e intensidad del campo difractado por ella. El esquema en c) representa lo que sería el complemento de la rendija, es decir, la amplitud de un haz láser en el que se ha interpuesto un obstáculo complementario a la rendija y los correspondientes perfiles de amplitud e intensidad del campo difractado. Puede verse que en cada punto, la suma de las amplitudes en c) y en b) tiene como resultado la amplitud en a). La intensidad del diagrama de difracción, que es lo que se detecta, es la misma para el caso de la rendija transparente que para la rendija opaca (obstáculo), salvo en el lóbulo central. Para que el lóbulo central no afecte el mínimo de la rendija en $x_0 \approx \lambda/z$, dicho lóbulo debe ser menor en ancho. Esto sería así si la sección del haz láser fuera muchas veces mayor que el ancho de la rendija. De esta manera, es posible medir el ancho de rendijas opacas, midiendo la separación de los mínimos del diagrama de difracción.

Esta técnica se aplica a la medida precisa de diámetros de objetos opacos como alambres, o espaciado de tamices. Poniendo detectores fotoeléctricos que midan el lugar en que aparecen los mínimos en el plano de observación es posible generar un lazo de control automático

para garantizar tolerancias preespecificadas . La precisión en estos casos es del orden de 0,3% para luz visible, y la técnica es útil entre 2,5 mm y 2 , 5 um,

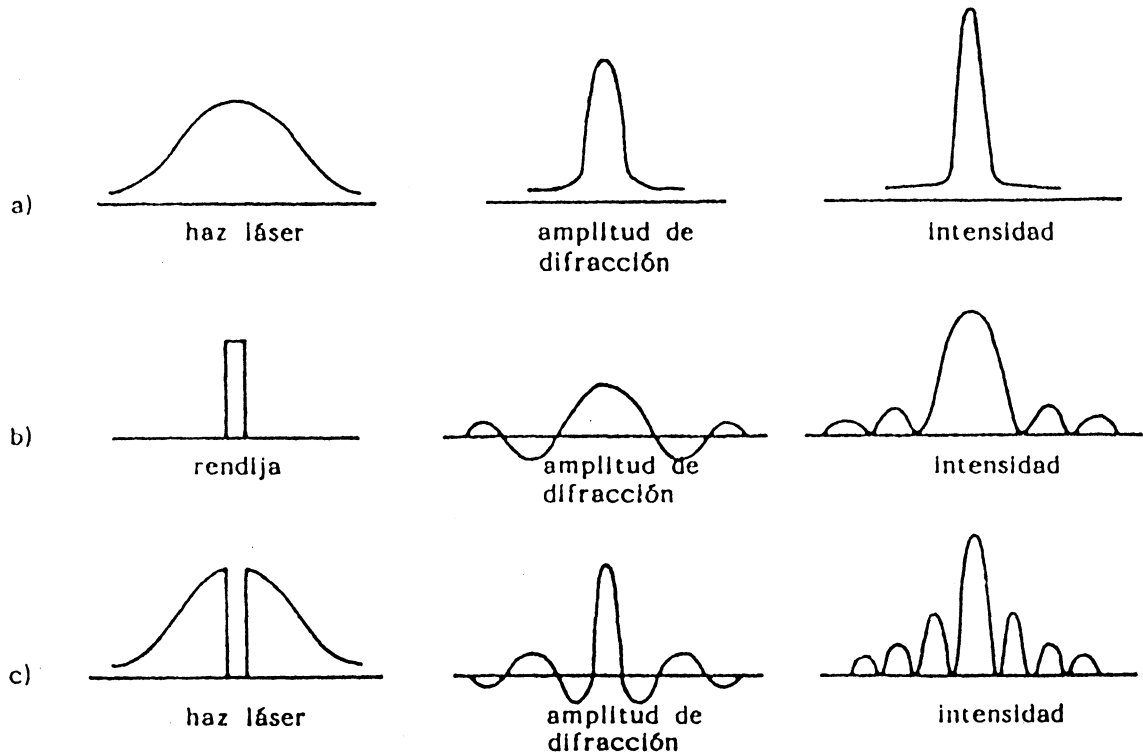


Figura 23

Difracción por muchas rendijas. Red de difracción

A partir de los resultados obtenidos en los párrafos anteriores, resulta relativamente sencillo analizar la difracción producida por una pantalla con varias rendijas o , en el límite , una red de difracción. La transformación de Fourier resulta para esto una herramienta poderosa.

La expresión matemática resultante para la amplitud del campo difractado en el caso de N+1 rendijas rectangulares se obtiene sumando las amplitudes de campo para una sola rendija. Si el ancho de las rendijas individuales es a y su separación es d , la intensidad del campo difractado resulta ser

$$I(x_0) = \frac{I_0}{(N + 1)^2} \left\{ \frac{\text{sen}(N + 1) \left(\frac{k}{2f} \right) (x_0 d)}{\frac{k x_0 d}{2f}} \right\}^2 \left\{ \frac{\text{sen} \left(\frac{k x_0 a}{2f} \right)}{\frac{k x_0 a}{2f}} \right\}^2$$

El primer factor entre llaves representa la figura de interferencia entre las N+1 rendijas (ver Figura 24 b)), y el segundo factor entre llaves es una función que modula esta distribución de intensidad y corresponde a la difracción por una sola rendija (ver Figura 24 c)). El producto de ambos explica la distribución de intensidad del campo difractado que está representado en d) y que corresponde a una sucesión de franjas brillantes (órdenes de difracción) modulado por el diagrama de difracción de una rendija. La separación de los órdenes depende de d y si el ancho a tiende a cero, el factor tiende a un valor constante.

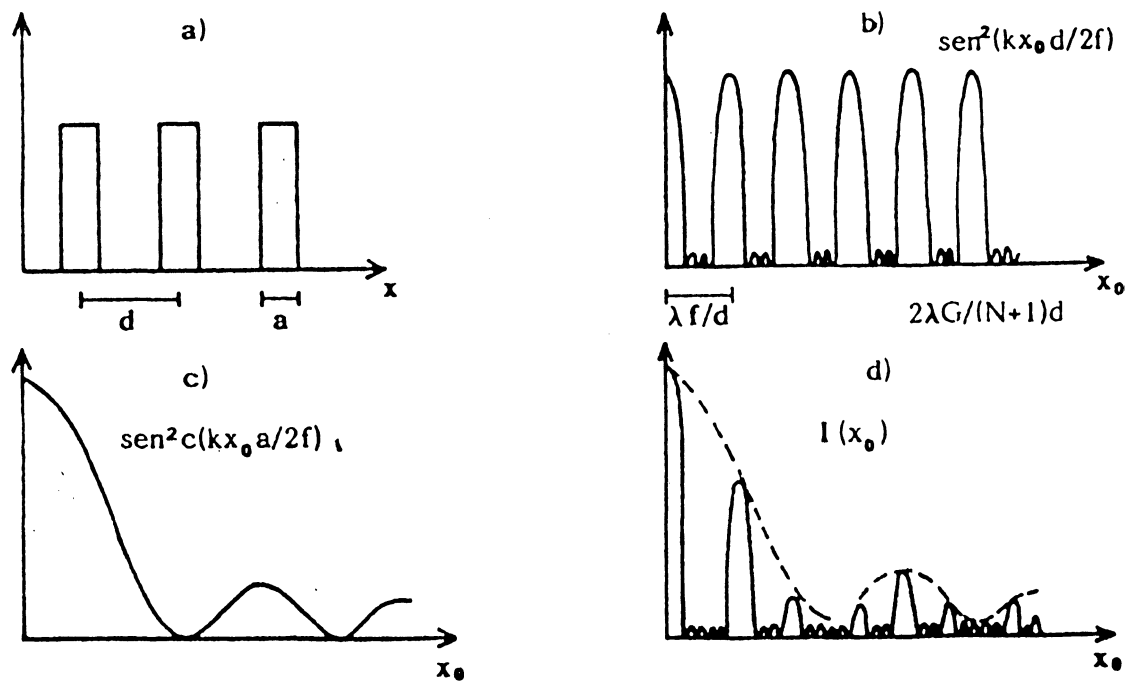


Figura 24

En la tabla de la Figura 25 que se presenta al final, se grafican diferentes funciones y sus correspondientes transformadas de Fourier. Si las funciones de la primera columna representan amplitudes complejas del campo luminoso en la pantalla difractante, las correspondientes de la segunda columna darán cuenta de la forma de la distribución del campo difractado en el infinito, o en el plano focal de una lente, es decir representan el espectro espacial de las funciones de la primera columna. Se han representado sucesivamente una rendija unidimensional, una función coseno limitada por una rendija, una función seno y una función coseno² limitadas por la misma rendija. La función cos² correspondería a una red de difracción cosenoidal en amplitud, mientras que las dos anteriores corresponderían a redes de fase. Por último, se presenta el caso de varias rendijas rectangulares (red de Ronchi) y su correspondiente diagrama de difracción de Fraunhofer.

La información contenida en un arreglo periódico está concentrada en regiones discretas y muy reducidas del plano de frecuencias. De esta manera el filtrado espacial puede resultar una herramienta muy útil para discriminar estructuras periódicas y no periódicas, ya que las últimas tendrán un espectro mucho más extendido que las primeras en el plano de Fourier

BIBLIOGRAFÍA

- *Introduction to Fourier Optics*, J.W. Goodman, McGraw-Hill, New York (1968)
- *Les Lasers et leurs Applications*, A. Orszag y G. Hepner, Masson, Paris (1980).
- *Laser in Industry*, S.S. Charschan ed., Van Nostrand (1972).
- *Introduction to Lasers and their Applications*, DC. O'Shea, W.R. Callen y W.T. Rhodes, Addison-Wesley (1977).
- *Optique Formation et Traitement des Images* M. Françon, Masson, Paris (1972).
- *Determination of the Diameter of Small Fibers by Diffraction of a Laser Beam*, L.A. Jeffers IEEE Conf. on Laser Engineering and Applications, paper THAM 11.5, Washington D.C. (1967).

- *Fundamentos de Optica*, F.A. Jenkins y H.E. White, McGraw-Hill, New York(1970).
- *Optical Interferometry*, M. Françon, Academic Press, New York (1966).
- *Aversatile Laser Interferometer*, J.H. Bruning, Appl.Optics 9,2180 (1970),
- *Inspection of Periodic Patterns with Intensity Spatial Filters*, L.S. Watkins , Solid State Technology 12, N°2 , 29
- *Optica*, E. Hecht and A. Zajac, Fondo Educativo Interamericano, Mexico, 1986.
- *Optics and Lasers*, M. Young, Springer-Verlag, Berlin,1977.

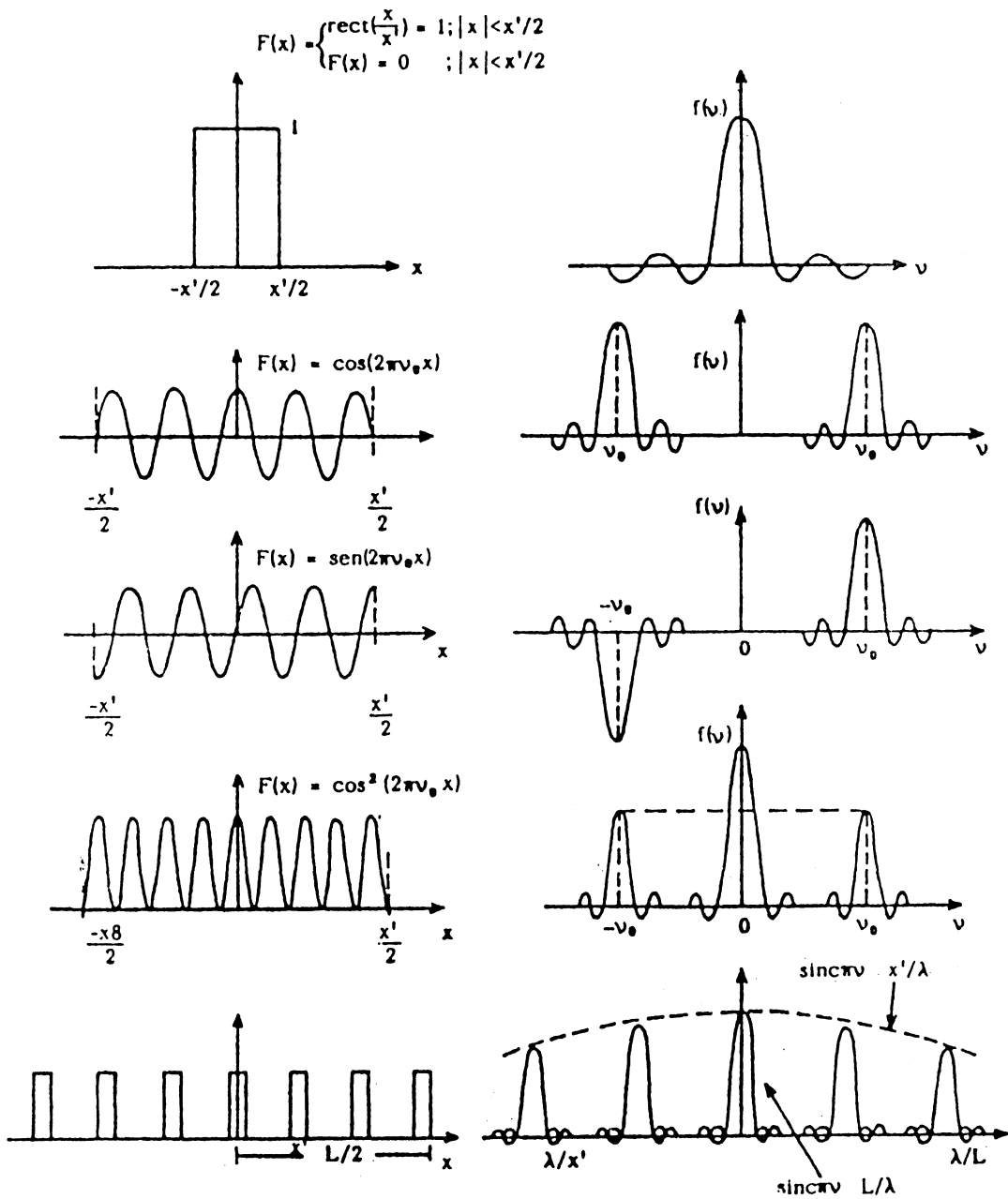


Figura 25

HOLOGRAFIA

R.D. Torroba

Introducción

La holografía, del griego holos (todo) y graphos (registro), es la rama de la ciencia que estudia el registro en un material apropiado, de toda la información de un frente de ondas, tanto en amplitud como en fase. El registro, denominado holograma, preserva la verdadera tridimensionalidad de la escena, al contrario de la fotografía convencional, que registra una escena tridimensional en un formato bidimensional. Asimismo, la holografía no requiere lentes ni ningún otro dispositivo formador de imágenes; en su lugar se registra un diagrama en intensidad, relacionado con la amplitud y fase de las ondas reflejadas por el objeto, haciendo uso de un proceso interferométrico. Este diagrama no guarda parecido con el objeto original, y sin embargo contiene toda la información que sobre el objeto tendría una fotografía corriente, además de la fase, la cual es requerida para visualizar la tercera dimensión.

La generación de una imagen a partir del holograma se denomina "proceso de reconstrucción". En este proceso, el holograma es iluminado con luz coherente y monocromática. Esta luz es difractada por el holograma en frentes de onda, que en esencia son indistinguibles de aquéllos que fueron generados por el objeto original. Estas ondas difractadas reproducen todos los fenómenos ópticos que caracterizan a las ondas originales (se puede enfocar y formar una imagen del objeto, aún cuando éste ya no esté). Un observador que intercepta estas ondas tiene la sensación de ver al objeto en su verdadera forma tridimensional. Más aún, si cambia de posición, también cambia la perspectiva de la escena con efectos de paralaje evidentes, debiendo además reenfocar la vista cuando se mueve el punto en observación desde cerca a un lugar distante en la escena. Es como si el holograma fuera una ventana a través de la cual se visualiza al objeto.

Además de estas características, el registro holográfico tiene otras propiedades interesantes. Cada porción de holograma reproduce toda la escena. Por otra parte, la copia por contacto de un holograma (las áreas opacas se vuelven transparentes y viceversa) reconstruye la misma imagen

que el holograma original.

Otra propiedad de un holograma es que siempre se reconstruyen dos imágenes, una real y otra virtual.

En la práctica, un holograma se registra mediante el montaje ilustrado en la Figura 1.

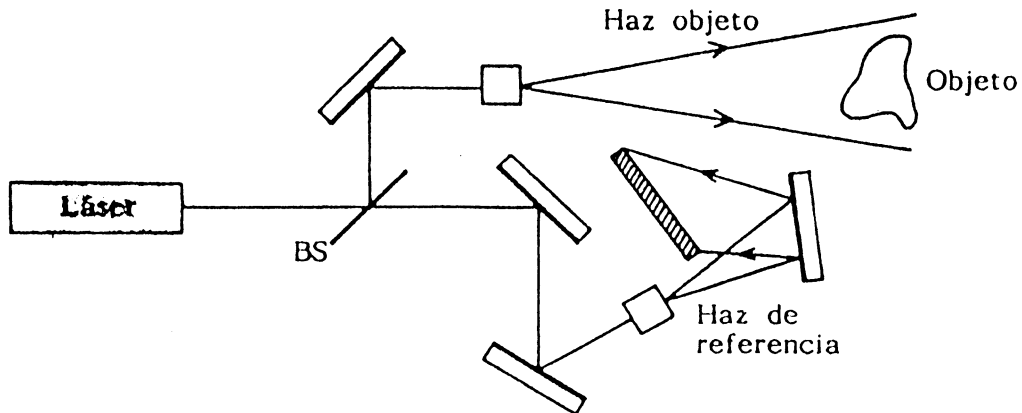


Figura 1

La luz proveniente del láser es dividida en dos por medio del divisor del haz BS. Uno de los haces es expandido y usado para iluminar al objeto, mientras que el otro haz, conocido con el nombre de haz de referencia, es expandido y dirigido por medio de espejos hacia una placa fotográfica de alta resolución H. La superposición de estos haces da lugar a un diagrama de franjas que queda registrado en la película. Las franjas de interferencia no son visibles al ojo desnudo, debido al pequeño espaciado interfranja ($\sim 5 \cdot 10^{-4}$ mm). Las franjas visibles se deben a partículas de polvo existentes en el montaje.

Una vez revelada, si la placa es iluminada con el mismo haz de referencia, para el observador aparecerá una imagen virtual del objeto original situada detrás del holograma, ocupando la misma posición del objeto.

Análisis teórico general

Como se indicó, el proceso holográfico requiere de dos operaciones: una de registro y otra de reconstrucción. Por el momento nos restringimos a la primera de ellas. Recordemos que la luz empleada es coherente.

Es necesario almacenar la información tanto de la fase como de la amplitud de las ondas. Pero los medios de registro comunes responden sólo a intensidad luminosa. Entonces, las variaciones de fase deben ser convertidas en variaciones de intensidad con el fin de poder registrarlas. La técnica que permite este paso es la interferometría, o sea que una onda de amplitud y fase conocida es sumada a la onda que proviene del objeto.

Matemáticamente, el frente de onda del objeto puede ser representado en el plano de la placa fotográfica como

$$a(x,y) = a_0(x,y) \exp[-i\phi(x,y)],$$

y la del haz de referencia

$$A(x,y) = A_0(x,y) \exp[-i\psi(x,y)].$$

Los dos frentes se suman en forma coherente, dando una intensidad

$$I(x,y) = |a_0(x,y)|^2 + |A_0(x,y)|^2 - 2a_0A_0\cos[\phi(x,y)-\psi(x,y)].$$

Los primeros dos términos representan las intensidades de cada frente de onda por separado. El tercer término es el efecto de la interferencia y el resultado directo de que las ondas sean coherentes una con la otra. Este término es función de las amplitudes individuales y, por medio del coseno, hay una explícita dependencia con la fase del objeto.

Procesamiento fotográfico

En este punto es conveniente recordar algunas definiciones con respecto a las emulsiones fotográficas.

La densidad de plata presente en la emulsión fotográfica luego del revelado y fijado, es una función de la distribución de intensidad y el tiempo de exposición. Esta relación está caracterizada por la curva de Hurter-Driffield (curva H-D) de las películas. Esta curva representa la densidad óptica D en función del logaritmo de la exposición E , que a su vez es el producto de la intensidad recibida durante la exposición por el tiempo que ésta duró. La Figura 2 representa la curva típica.

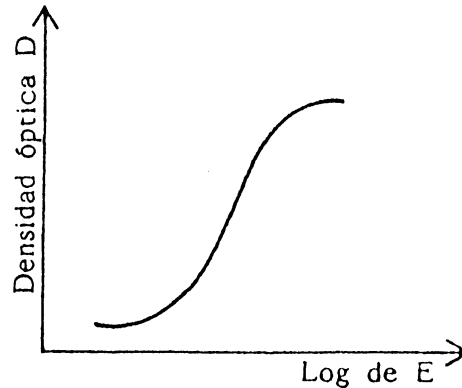


Figura 2

La densidad óptica se define

$$D = -\log \tau,$$

donde τ es la transmitancia en intensidad de la película revelada. Como se ve en la curva, las regiones de baja y alta exposición son no lineales. Sin embargo, entre estos extremos la curva es relativamente lineal, pudiendo ser representada en una buena aproximación por

$$D = \gamma \log E + K,$$

donde γ es la pendiente de la curva y refleja las propiedades del film y el proceso de revelado, y K es una constante. Sustituyendo en la ecuación que

define D , la relación entre transmitancia en intensidad y exposición es

$$\tau \propto E^{-\gamma}.$$

A su vez, la transmitancia en intensidad está asociada, en forma natural para la óptica coherente, a una transmitancia en amplitud tal que $\tau = tt^*$ (el asterisco denota complejo conjugado).

Si suponemos despreciable al cambio de fase introducido por la emulsión, entonces $t = \sqrt{\tau}$ y de esta manera

$$t \propto E^{-\gamma/2} \quad \delta \quad t \propto I^{-\gamma/2} \propto |U|^{-\gamma},$$

donde U es la amplitud compleja del campo incidente durante la exposición.

El material fotográfico provee una correspondencia entre la intensidad incidente durante la exposición y la amplitud transmitida luego del revelado. Basta trabajar con una película de $\gamma = -2$ para que esta correspondencia sea lineal, asegurando fidelidad en el holograma procesado.

Finalmente, retornando a nuestro ejemplo y asumiendo que la intensidad del haz de referencia $|A|^2$ es uniforme en la superficie de registro, la amplitud de transmitancia de la placa es

$$t(x,y) = t_0 + \beta(|a|^2 + A^*a + Aa^*),$$

siendo t_0 una transmitancia uniforme, y β un parámetro que tiene en cuenta las características del film y el tiempo de exposición.

Reconstrucción del frente de onda original

Supongamos que el holograma es iluminado por una onda de reconstrucción coherente descrita por $B(x,y)$. La luz transmitida es

$$B(x,y)t(x,y) = t_0B + \beta aa^*B + \beta A^*Ba + \beta ABA^* = U_1 + U_2 + U_3 + U_4.$$

Notemos que si B fuese una réplica exacta de A , el tercer término es

$$U_3(x,y) = \beta |A|^2 a(x,y) .$$

Recordando que $|A|^2$ fue supuesto uniforme, U_3 es, salvo por un factor, la duplicación exacta del frente de onda del objeto original.

Análogamente, si B fuera el conjugado del haz de referencia, o sea $A^*(x,y)$, el cuarto término es

$$U_4(x,y) = \beta |A|^2 a^* ,$$

que es proporcional al conjugado de la onda objeto original. En cualquier caso, la componente de interés está acompañada de tres términos adicionales, que hacen las veces de ruido óptico. Es evidente entonces que se requiere de algún método para separar estas varias componentes de luz transmitida.

El holograma de Gabor

Históricamente, el primer holograma fue hecho por un proceso desarrollado por Denis Gabor. La geometría de registro se ilustra en la Figura 3.

Supondremos que el objeto es muy transparente, o sea posee una amplitud de transmitancia de la forma

$$t(x',y') = t_0 + \Delta t(x',y') ,$$

donde t_0 es la transmitancia media y Δt representa las variaciones con respecto a esta media y

$$\Delta t \ll |t_0| .$$

Cuando tal objeto es iluminado por una fuente coherente, transmite dos componentes: una onda muy intensa, que se corresponde con t_0 ; y

una débil generada por las variaciones de transmitancia Δt . Sobre la placa H, se registra una intensidad

$$I(x,y) = |A + a(x,y)|^2 = A^2 + |a(x,y)|^2 + Aa(x,y) + Aa^*(x,y),$$

donde A es la amplitud de la onda plana (ver Figura 3) y $a(x,y)$ es la amplitud de la luz dispersada. En cierto sentido, se puede decir que la onda de referencia fue suministrada por t_0 .

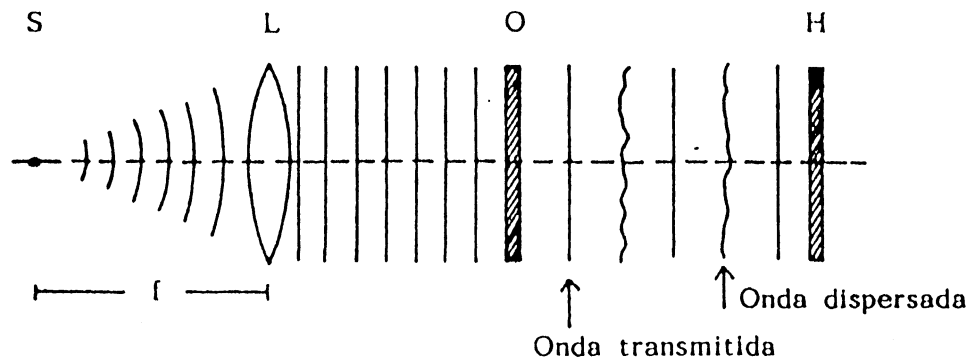


Figura 3

La placa reveladora tendrá una amplitud transmitida de la forma

$$t(x,y) = t_b + \beta (|a|^2 + Aa + Aa^*).$$

Si es iluminada por una onda plana, en incidencia normal, y de amplitud B , el campo resultante es

$$Bt = Bt_b + \beta B |a(x,y)|^2 + \beta ABa(x,y) + \beta ABa^*(x,y).$$

El primer término es una onda plana, que es atenuada uniformemente, sin dispersar. El segundo término se puede despreciar en virtud de nuestra suposición que $|a(x,y)| \ll A$. El tercer término es proporcional a la onda original $a(x,y)$, y aparece como la imagen virtual del objeto. En forma similar, el término restante representa la imagen real del objeto, en el lado opuesto de la transparencia. Lo antedicho se esquematiza en la Figura 4.

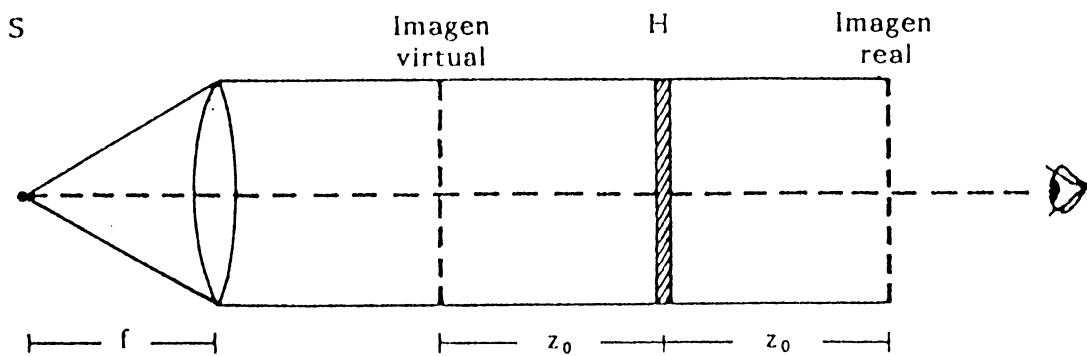


Figura 4

Como vemos, el holograma de Gabor genera simultáneamente dos imágenes del objeto, una real y otra virtual, ambas centradas en el eje del sistema óptico. Estas "imágenes gemelas" están separadas una distancia $2z_0$ y acompañadas de un fondo coherente B_t . Esto es una seria limitación, debido a su inseparabilidad. Cuando se enfoca una imagen, siempre se ve la otra fuera de foco, con una consecuente pérdida de calidad.

Se desarrollaron muchas técnicas para solucionar este inconveniente, pero la más exitosa es la originada por Leith y Upatnieks, que será discutida en detalle.

Holograma de Leith-Upatnieks

La mayor diferencia entre este tipo de holograma y el de Gabor, reside en la utilización de un haz de referencia por separado y no la luz directamente transmitida por el objeto. Más aún, este haz no es colineal con el eje objeto-película. La geometría de registro se ve en la Figura 5.

La fuente puntual S es colimada por la lente. Una parte de la onda plana así generada da en el objeto y otra parte incide en el prisma, siendo deflectada un ángulo θ con respecto a la película. Entonces, la distribución de amplitud en la placa de registro se escribe

$$U(x,y) = A \exp(-2\pi i \alpha y) + a(x,y) ,$$

donde la frecuencia espacial del haz de referencia es $\alpha = \text{sen } \theta / \lambda$. El primer término denota a la onda plana inclinada y el segundo la luz transmitida por el objeto. La intensidad es

$$I(x,y) = A^2 + |a(x,y)|^2 + Aa(x,y)\exp(2\pi i \alpha y) + Aa^*(x,y)\exp(-2\pi i \alpha y) .$$

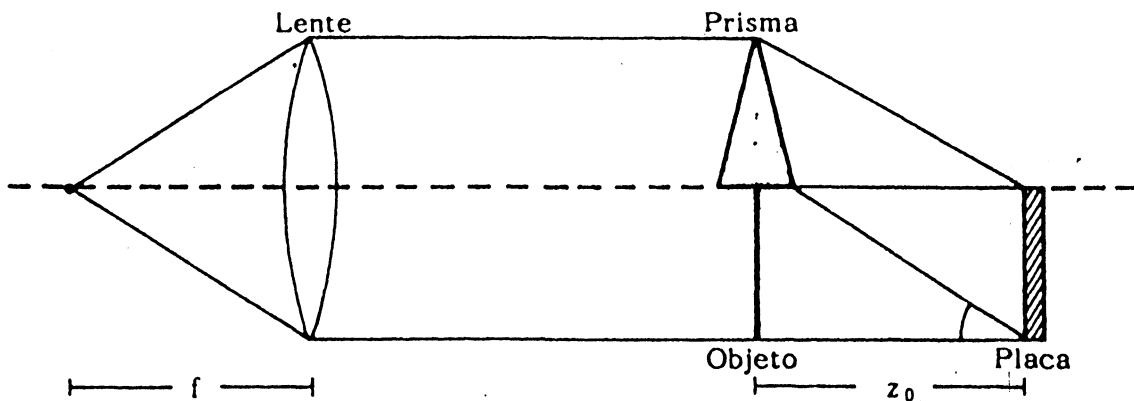


Figura 5

De la manera usual supondremos que el holograma revelado es iluminado por una onda plana uniforme de amplitud B , componiéndose el campo transmitido por cuatro términos, identificados como

$$U_1 = Bt_b \quad , \quad U_3 = \beta B A a(x,y)\exp(2\pi i \alpha y) ,$$

$$U_2 = \beta B |a(x,y)|^2, \quad U_4 = \beta B A a^*(x,y) \exp(-2\pi i c y).$$

La componente U_1 es una versión atenuada de la onda de reconstrucción y representa a una onda plana que viaja en el eje de la transparencia. El segundo término U_2 depende de las coordenadas y se puede tomar como la superposición de ondas planas que viajan, aproximadamente, con ángulos que no se separan mucho del eje óptico.

La onda U_3 genera una imagen virtual del objeto, a una distancia z_0 de la transparencia, y el factor exponencial indica que la imagen está deflectada un ángulo θ con respecto al eje de la transparencia, como se aprecia en la Figura 6.

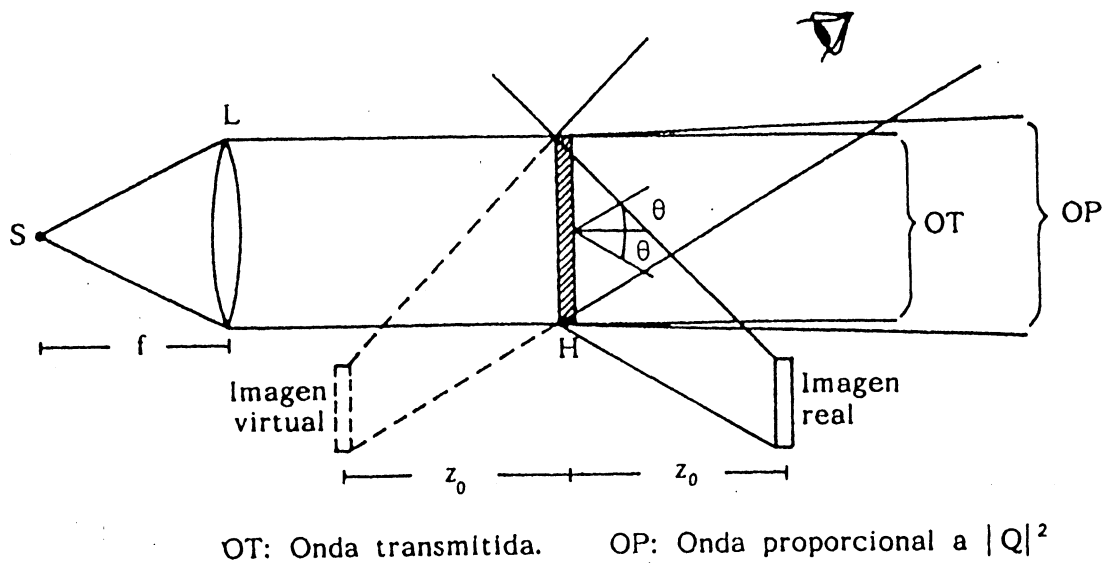


Figura 6

Análogamente, la onda U_4 da lugar a la imagen real del objeto formada a z_0 de la transparencia en el lado opuesto de la imagen virtual, indicando el factor exponencial que la deflexión es a un ángulo $-\theta$ con respecto al eje de la placa.

El resultado más importante de este proceso es que, si bien las imágenes real y virtual son generadas al mismo tiempo, ambas están separadas

angularmente una de la otra y también con respecto al fondo. Que una imagen haya sido aislada de la otra, fue el resultado de usar un haz de referencia no colineal con el haz del objeto. Con respecto a la naturaleza de la onda de reconstrucción, ésta no necesita ser una réplica exacta de la onda de referencia para obtener un resultado aceptable.

Hasta ahora hemos tratado objetos transparentes. Pero si nos retrotraemos a la Figura 1, vemos una configuración que nos permite registrar un holograma de una escena tridimensional. Basta recordar que el proceso holográfico es la formación de imágenes mediante la duplicación del frente de ondas generado por la escena original.

Holografía Interferométrica

La interferometría holográfica es similar a la holografía convencional, salvo que se registran en la misma placa dos exposiciones, y por ende, dos hologramas del mismo objeto, siendo éste deformado entre ambas exposiciones. Al ser reconstruido, se formarán dos imágenes tridimensionales, y por ser ambas coherentes y al estar ubicadas aproximadamente en el mismo lugar en el espacio, van a interferir, produciendo un conjunto de franjas claras y oscuras sobre la imagen reconstruida. Las franjas representan contornos de igual desplazamiento, donde cada franja sucesiva representa un desplazamiento de media longitud de onda de la luz coherente empleada en la reconstrucción.

Hay tres variaciones sobre esta técnica, cada una con ventajas específicas sobre las otras, dependiendo del problema a resolver. Ellas son:

- 1) Doble exposición estática.
- 2) Promedio temporal dinámico.
- 3) A tiempo real.

Describiremos a continuación cada una de ellas.

1) Doble exposición estática: es el registro en dos exposiciones sucesivas de un objeto en dos posiciones diferentes. Cuando se reconstruye al holograma, se produce una figura de interferencia que indica el desplazamiento entre las exposiciones. Supongamos que se registran dos frentes de onda U_1 y U_2 con una onda de referencia U_r . La intensidad registrada será

$$I = |U_1|^2 + |U_2|^2 + |U_r|^2 + U_r^*(U_1+U_2) + U_r(U_1^*+U_2^*) .$$

Si la placa procesada es iluminada por una onda réplica del haz de referencia, el término U_i que lleva información de la imagen virtual será

$$U_i \propto (U_1 + U_2) ,$$

que describe al vector suma de los campos correspondientes. Por otro lado, cada campo se puede representar por

$$U_1 = A_0 \exp(ik\phi_1) \quad , \quad U_2 = A_0 \exp(ik\phi_2) .$$

La intensidad para la doble exposición será

$$I_i = |U_i|^2 \propto 2A_0^2 [1 + \cos k(\phi_2 - \phi_1)] ;$$

aquí $\phi_2 - \phi_1$ da una medida directa de la longitud de camino adicional que sufrió el haz objeto como resultado del desplazamiento.

Para simplificar la idea, supongamos que el objeto se desplazó una distancia z_1 perpendicular a la superficie. Si llamamos α al ángulo que forma la luz incidente con la normal a la superficie en el proceso de construcción, entonces

$$I_i \propto 2A_0^2 [1 + \cos kz_1(\cos\alpha)] ;$$

usando la identidad $\cos^2(x/2) = (1 + \cos x)/2$, queda

$$I_i \propto 4A_0^2 \cos^2 \frac{kz_1}{2} (\cos\alpha) .$$

Esta ecuación describe la geometría de las franjas de interferencia, con la distancia entre franjas sucesivas directamente ligada a z_1 .

2) Promedio temporal dinámico: consiste en hacer un solo registro holográfico de un objeto sujeto a un movimiento vibratorio cíclico, en la suposición que el tiempo de exposición es largo comparado con un período del ciclo vibratorio. De esta forma el holograma almacena un conjunto de imágenes que corresponden al promedio en el tiempo de todas las posiciones del objeto durante su vibración. Consecuentemente, al reconstruir el holograma, la interferencia entre estas imágenes produce un diagrama resultante con información sobre la deformación de la pieza.

Si el objeto vibrante se mueve con una frecuencia w y amplitud z_2 , el campo se puede representar como

$$U = A \exp[ik\gamma z_2 \cos wt],$$

donde γ es una función del ángulo de iluminación de la luz incidente en el registro. Como el holograma registra todas las posiciones del objeto en un período T del ciclo vibratorio, es necesario promediar esta ecuación en un período para obtener la forma correcta de la onda que será reconstruida,

$$U_i \propto (1/T) \int_0^T U dt = (A/T) \int_0^T \exp(ik\gamma z_2 \cos wt) dt,$$

$$I_i = |U_i|^2 \propto A^2 (1/T)^2 \left| \int_0^T \exp(ik\gamma z_2 \cos wt) dt \right|^2.$$

La representación de las funciones de Bessel en forma integral es

$$J_n(x) = (i^{-n}/2\pi) \int_0^{2\pi} \exp(ix \cos t) \exp(int) dt.$$

Para el caso $n=0$, nuestra ecuación puede escribirse

$$I_i \propto A^2 J_0^2(k\gamma z_2),$$

donde J_0 es la función de Bessel de orden cero. Esta ecuación guarda similitud con la hallada en el ítem anterior, salvo que ésta varía con el cuadrado de la función de Bessel. Esta variación se caracteriza por tener un máximo cuando el argumento es cero, con máximos sucesivos pero que decrecen en valor. Esto indica que cuando el desplazamiento es cero, la intensidad en la imagen es la más alta. Ello significa que las partes que no se mueven, o sea los nodos, tienen la mayor intensidad y se pueden detectar fácilmente mirando la imagen reconstruida.

La técnica se puede extender al estudio de otros movimientos cíclicos temporales compuestos por varios movimientos sinusoidales. Sus funciones características serán de Bessel de orden superior combinadas con exponenciales.

Existe un movimiento no-cíclico que se encuentra a menudo y es debido a un movimiento a velocidad constante, que puede ser, por ejemplo, el resultado de cambios por variación de temperatura. En este caso,

$$I_i \propto A^2 (1/T)^2 \left| \int_0^T \exp(ikVt) dt \right|^2,$$

donde T es ahora el tiempo de exposición, y V la velocidad. Integrando y elevando al cuadrado,

$$I_i \propto A^2 [(2 - 2 \cos kVt) / (kVT)^2].$$

Los ceros de esta función (o franjas oscuras) aparecen cuando VT/λ es un entero.

3) A tiempo real: esta técnica consiste en tomar una sola exposición fotográfica de un objeto sin tensar, revelar la placa y reposicionarla en exactamente el mismo lugar en que fue registrada. La precisión debe ser del orden de la longitud de onda. En la práctica, el revelado se hace "in situ", en tanques especiales diseñados para esta aplicación. A continuación se reconstruye al holograma usando el mismo haz de referencia que se empleó en la construcción, de tal suerte que la imagen holográfica está superpuesta al objeto original, que está iluminado con la misma luz del registro. Si el objeto sufre una deformación, aparecerán franjas de interferencia. Esta comparación entre ambos estados del objeto se hace en el instante en que esta deformación ocurre, constituyendo un ejemplo de holografía interferométrica a tiempo real.

Supongamos que el objeto sufre un desplazamiento como un todo y al mismo tiempo una vibración armónica. Sea A el frente de onda del objeto original sin perturbar

$$U_i \propto A.$$

La superposición con el frente perturbado será

$$U_{ir} \propto A + A \exp(iky z_1 + ik y z_2 \cos \omega t),$$

donde z_1 es el desplazamiento como un todo independiente del tiempo y

$z_2 \cos wt$ el movimiento armónico,

$$I_{ir} \propto A^2 [2 + \exp ik\gamma(z_1+z_2 \cos wt) + \exp -ik\gamma(z_1+z_2 \cos wt)] .$$

Promediando en el tiempo de manera similar a lo desarrollado en el ítem anterior

$$I_{ir} \propto 2A^2 [1 + \cos(k\gamma z_1) J_0(k\gamma z_2)] .$$

Si hubiese sido sólo un desplazamiento como un todo ($z_2=0$), la ecuación queda idéntica al del caso de doble exposición, con las mismas franjas características. Si el objeto sólo vibrara ($z_1=0$),

$$I_{ir} \propto 2A^2 [1 + J_0(k\gamma z_2)] ,$$

que no es igual al caso del promedio temporal. El resultado es una reducción en el contraste de las franjas para el caso de tiempo real y además el número total de franjas será la mitad que en el caso de promedio temporal.

En resumen, la interferometría holográfica a tiempo real ofrece la ventaja de poder estudiar varios movimientos con un solo holograma, a expensas de una disminución de tanto el contraste como la sensibilidad, en el caso de movimientos vibratorios. Sin embargo, estos inconvenientes son superados por el considerable ahorro en tiempo que esta técnica permite.

Para concluir, la interferometría holográfica, en combinación con otros mecanismos (técnicas de excitación ultrasónica, tensión inducida por efectos térmicos, presurización, fuerzas mecánicas, etc), constituye una técnica promisoría en el área de ensayos no destructivos. Además, permite explorar áreas que por métodos más convencionales eran inaccesibles o de acceso restringido.

Topografía de un objeto tridimensional

Una técnica óptica de considerable potencial en el campo de ensayos no destructivos es el contorneado óptico (líneas de nivel). Sus aplicaciones incluyen el control de calidad en línea, el pulido de superficies de modo de aproximarlas a la forma que se desea, la descripción de la topografía del

objeto, etc.

Existen métodos no holográficos basados en técnicas de proyección de líneas, tal como el moiré. Entre los métodos holográficos, discutiremos en detalle, a modo de ejemplo, el llamado de inmersión o múltiple índice de refracción.

Consideremos un objeto difusor colocado en una caja transparente que contiene un fluido con un índice de refracción diferente al del medio exterior y conocido. Los rayos de luz más simples son aquellos perpendiculares a la superficie frontal de la caja. Por ello consideremos la geometría de la Figura 7, la más apta para el registro.

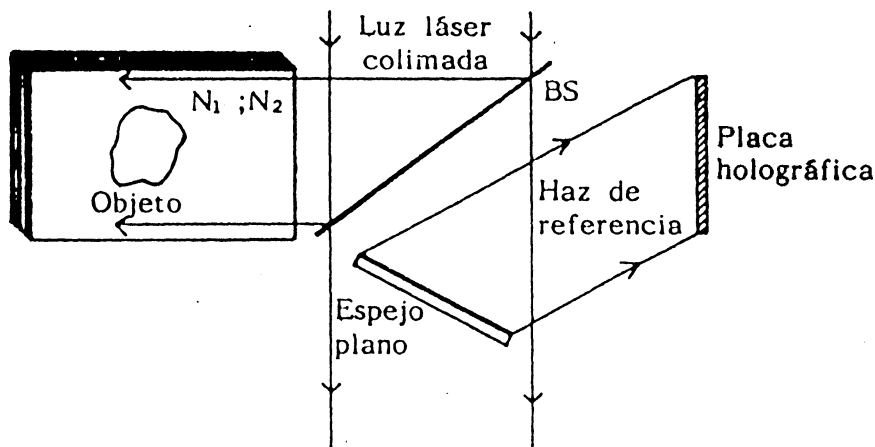


Figura 7

El divisor del haz BS permite tanto la iluminación del objeto como el paso de la luz reflejada por el objeto hacia la placa de registro. El espejo provee del haz de referencia. Primero se hace un holograma con el objeto inmerso en un líquido de índice conocido n_1 . Luego, manteniendo la misma geometría, se reemplaza el fluido de la caja por otro de índice conocido n_2 . La interferencia resultante resulta en un sistema de franjas que representan contornos de igual profundidad.

Como resultado del cambio de índice de n_1 a n_2 , la fase cambia en una cantidad igual a

$$\Delta\phi = (2\pi/\lambda) |n_1 - n_2| 2z .$$

La distancia entre dos contornos sucesivos está dada cuando la fase cambia en 2π . De esta forma el espaciado resulta

$$D = \lambda / (2 |n_1 - n_2|) .$$

Técnicas holográficas especiales

Hasta este punto, el intento fue proveer de una teoría general de la holografía óptica. Además del simple registro y reconstrucción de objetos tridimensionales, se han descrito técnicas de interferometría holográfica y contorneado. Sin embargo, el énfasis fue puesto en los principios básicos, métodos de laboratorio y experimentos generales. A partir de ahora, consideremos algunos procedimientos que hacen de la holografía un proceso útil en problemas de inspección ingenieriles. Los ejemplos elegidos son para mostrar la viabilidad y versatilidad de la técnica.

En general, una vez que se ha determinado que una técnica holográfica experimental en particular, provee una solución adecuada a un problema específico, el siguiente paso es diseñar un sistema apropiado (o prototipo) que permita su uso rutinario.

Uno de los requerimientos condicionantes asociado a la holografía es la aislación contra vibraciones externas (menos que $\lambda/8$), lo que hace necesario el empleo de mesas antivibratorias. Esto aumenta el costo y la complicación en el sistema, haciéndolo, por supuesto, no portátil. También se perderían las ventajas de bajo mantenimiento y un mínimo de conocimiento por parte del operador. Por esta razón, se han desarrollado técnicas de compensación de movimientos vibratorios, que se pueden clasificar en tres categorías: generación de un haz local de referencia, compensación por uso de servo sistemas electrónicos de retroalimentación y reducción del tiempo de exposición.

El primero es el más fácil de implementar, y se basa en afectar al haz de referencia con el movimiento del propio objeto. Así, el movimiento del objeto está compensado automáticamente. El segundo, sensa la vibración del objeto para luego mover electrónicamente el espejo del haz de referencia. El tercero se basa en incrementar la intensidad de la luz recibida por el material de registro, lo que permite la reducción en el tiempo de exposición, y por ende la reducción en el mantenimiento de la estabilidad.

Medición de esfuerzos

Aquí mostraremos cómo los métodos para medir pequeños desplaza-

$$\delta = (\vec{k}_1 - \vec{k}_2) \cdot \vec{d} .$$

Esto indica que para el sistema holográfico más simple, o sea aquél en el que $(\vec{k}_1 - \vec{k}_2)$ está cerca de la normal a la superficie, se medirá el movimiento en esta dirección normal y no el desplazamiento en el plano. Por esta razón se desarrollan métodos para obtener el movimiento en el plano y poder medir el esfuerzo.

Si aplicamos una técnica de tiempo real, y reconocemos las franjas de orden cero, podemos asignar a cada franja sin ambigüedad un número de orden n , obteniendo el valor de δ a partir de

$$\delta = n[\lambda / (2 \cos \phi)] ,$$

con $\phi = k_1 - k_2$. Si las direcciones de iluminación y observación están a ángulos iguales con respecto a la normal a la superficie de prueba (dirección z), las franjas medirán el desplazamiento real z . De la teoría elástica, el momento de pando B en un punto en la superficie de un objeto que está siendo flexionado, está dado por

$$B = YI \partial^2 z / \partial x^2 ,$$

donde Y es el módulo elástico del material, I es el momento de la sección transversal y x es la dirección a lo largo de la superficie en la cual se mide el momento de pando. Graficando se puede obtener la derivada segunda de z . Si se conoce la distancia z_0 de la superficie al eje neutral y sabiendo que $\epsilon = Bz_0/YI$, entonces

$$\epsilon = -z_0 \partial^2 z / \partial x^2 ,$$

y se obtiene el esfuerzo ϵ .

Esta técnica se usa para objetos que sufren sólo pando, por ejemplo una viga empotrada en voladizo, pero no en un puente con armazón.

DetECCIÓN DE FISURAS

El diagrama de franjas de un holograma interferométrico da la distribución de desplazamientos en una superficie. La presencia de una fisura de magnitud suficiente causará una anomalía localizada en el diagrama. Esta aplicación es particularmente útil cuando es imposible o no deseable el pulir o tratar la superficie a inspeccionar, o cuando es imposible el contacto directo con ella. Aquí es conveniente destacar una ventaja aún no mencionada de la holografía, cual es la habilidad de inspeccionar el campo objeto como un todo, sin necesidad de una detección punto a punto. Un ejemplo es la detección de fisuras por corrosión, que se puede implementar a tiempo real, y pudiendo examinar toda la superficie de la muestra, monitoreando en forma continua, sin necesidad de sacarla del medio corrosivo. La ventaja fundamental es la detección de la fisura antes de que crezca hasta un tamaño que la permita observar fácilmente.

En el caso de una probeta de una aleación de titanio, inmersa en metanol, cuyas dimensiones eran 2 x 7/8 de pulgada y 0,1 de pulgada de espesor, sometida a tensión en tres puntos, se pudo detectar el comienzo de una fisura de $0,3\mu$ al cabo de 7 horas de ser sumergida la muestra en el líquido. Se deben notar, sin embargo, ciertas limitaciones. La estabilidad mecánica debe ser la apropiada para el registro holográfico y el ambiente de la muestra debe ser transparente para la luz láser. La temperatura debe mantenerse constante. Cada aplicación propuesta de esta naturaleza debe ser investigada individualmente para ver si el procedimiento es viable.

INSPECCIÓN DE ÁLABES DE TURBINAS

Tanto los modos de vibración de torsión como de flexión pueden ser inducidos en álabes de turbinas por varios métodos. En general, el análisis vibracional simple no es un problema si los diagramas modales de interés están restringidos a aquellas frecuencias que se pueden esperar en el medio común de operación del álabes. Por otro lado, es útil el estudio de hasta cientos de kilociclos en la evaluación de las características internas de los álabes. Esto se logra induciendo acústicamente la vibración por medio de transductores piezoeléctricos fijados a su superficie. Estudios holográficos intensivos encontraron que todos los tipos de álabes exhiben el mismo diagrama básico a bajas frecuencias (100 a 15.000 Hz), independiente de su forma o tamaño. Es para frecuencias más altas que los diagramas de modos se hacen función de las características físicas y en el caso de álabes huecos, de su estructura interna. Aquí se pueden desprestigiar los efectos de dispersión por granos, atenuación en metales y rugosidad normal de superficies, factores éstos que limitan la resolución de las pruebas ultrasónicas más convencionales, como ser ecos pulsados o transmisión de ultrasonido.

La holografía con láseres pulsados (20 nseg) demostró ser una herramienta efectiva al permitir detectar componentes de vibración no periódicas. A pesar de habernos centrado en álabes individuales, también es posible examinar todo un grupo a fin de estudiar su efecto interactivo.

Bibliografía

- "Optical Holography", R.J. Collier, C.B. Burckhardt, L.H. Lin, Academic Press, 1976.
- "Holographic Nondestructive Testing", Robert Erf, Academic Press, 1974.
- "Introduction to Fourier Optics", J.W. Goodman, McGraw-Hill, 1968.
- "Coherent Optical Techniques in Experimental Mechanics", Optical Engineering, Vol. 21, N°3, 1982.
- "Incoherent Optical Techniques in Experimental Mechanics", Optical Engineering, Vol. 21, N°4, 1982.
- "Optical Metrology", Optical Engineering, Vol. 18, N°5, 1979.
- "The Engineering Use of Holography", G. Groth, Cambridge Univ. Press, London, 1968.
- "Principles of Optics", M. Born y E. Wolf, Pergamon, 1965.
- "Theory and Applications of Holography", J.B. De Velis y G. Reynolds, Addison-Wesley, 1967.

ALINEACION, POSICIONAMIENTO, CONTROL Y MEDIDAS

M. Gallardo

Introducción

Elementos de la óptica geométrica, física e instrumental, complementados con cualidades que posee la radiación generada por el láser, constituyen herramientas de invaluable potencial para solucionar problemas que plantea la ingeniería actual.

Esta nueva forma de encarar los problemas, reemplaza en muchos casos con ventajas el uso de viejas técnicas, aportando un mayor grado de precisión y confiabilidad y extendiendo la posibilidad de control a distancias mayores o a puntos de difícil acceso. En otros casos, los métodos de la óptica y el láser generan nuevas técnicas que permiten atacar problemas insolubles hasta hace no mucho tiempo.

En general, la cantidad y calidad del instrumental a emplear en la realización de un trabajo, depende del grado de complejidad de dicho trabajo; pero cuando se utilizan ideas y métodos básicos de la óptica moderna, la implementación del instrumental tiende a simplificarse y los diversos componentes del mismo pueden ser fácilmente provistos y además, ser utilizados en otras experiencias con sólo efectuarles cambios mínimos.

En este artículo describiremos métodos generales que emplean radiación láser y técnicas ópticas y que son utilizados para resolver problemas de posicionamiento, alineación, medidas y control de grandes componentes o para generar rectas, planos y superficies de luz, que puedan ser tomados como referencia para construcción o verificación de otras.

Es importante hacer notar que muchas de las ideas básicas que se desarrollarán en este artículo, han sido aplicadas directamente en el país por integrantes de nuestro equipo de trabajo y que algunos de los métodos que se mencionarán son originales. Esto se verá reflejado en la bibliografía consignada en el texto.

Optica de haces gaussianos

Antes de introducirnos en las aplicaciones tecnológicas propiamente dichas, es importante entender la naturaleza de la radiación láser a los efectos de conocer principios y posibilidades de la óptica física.

Las características principales del láser que se tendrán en cuenta son las siguientes: la direccionalidad del haz, la facilidad para observar la luz aún en condiciones adversas y la coherencia de la radiación emitida, importante en aplicaciones en donde es necesario focalizarla en puntos pequeños.

Con respecto a los instrumentos ópticos a utilizar, es conveniente hacer notar que los diseños y tolerancias son más críticos que para sistemas convencionales, aunque éstos tienden a ser simples ya que problemas asociados con sistemas de imágenes no están en general presentes. El único inconveniente podría surgir de aberraciones esféricas, pero los radios de los elementos ópticos son relativamente grandes, haciendo el problema insignificante, ya que se trabaja en la zona paraxial.

En general, en los trabajos de alineación, posicionamiento, medidas y control, se emplean láseres de He-Ne de baja potencia (1-6 mW) operando a 632,8 nm en el modo TEM₀₀. En este modo el haz emitido tiene a la salida del láser un frente de onda plano y una distribución de intensidad gaussiana mostrada en la Figura 1 y representada por la expresión

$$I(r) = I_0 \exp \left(-\frac{2r^2}{w^2} \right)$$

en la que r es la distancia radial a partir del centro del haz y w el radio de la circunferencia en donde la intensidad cae a $1/e^2$ (13,5%) del valor I_0 en el centro del haz.

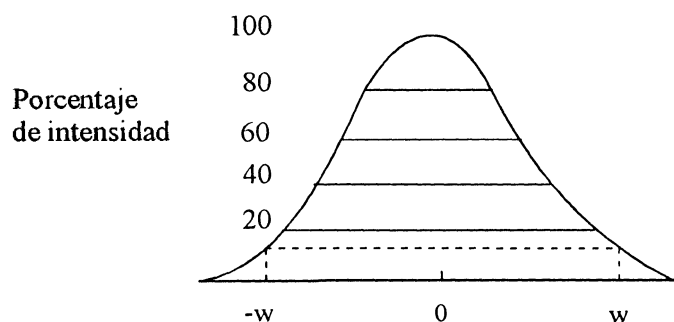


Figura 1

Este haz puede ser "colectado" y enfocado en áreas pequeñas a cualquier distancia a partir de la salida del láser, pero toda la óptica a utilizar debe funcionar limitada sólo por la difracción.

A pesar de que la distribución de intensidad del haz del láser permanece gaussiana al propagarse, las leyes de la óptica dictaminan que debe abrirse transversalmente. Por lo tanto es imposible tener un haz perfectamente colimado aunque bajo ciertas circunstancias la divergencia puede ser tan insignificante que es difícil decir que ésta esté presente.

El hecho de que el haz del láser diverja, está de acuerdo con las predicciones que marca la teoría de la difracción y como las aberraciones son totalmente insignificantes, las fórmulas que utilizaremos describirán con exactitud la divergencia del haz láser.

Supongamos tener en un determinado momento y lugar un frente de onda plano correspondiente a un haz láser TEM_{00} gaussiano con todos sus elementos moviéndose en direcciones paralelas precisas. Rápidamente este frente de onda se curvará y el haz tenderá a ser divergente de acuerdo con las siguientes fórmulas [1]

$$R(z) = z \left[1 + \left(\frac{\pi w_0^2}{\lambda z} \right)^2 \right]$$

$$w(z) = w_0 \left[1 + \left(\frac{\lambda z}{\pi w_0^2} \right)^2 \right]^{1/2}$$

donde z es la distancia que se propagó el haz tomando como referencia el punto donde el frente de onda era plano, λ es la longitud de onda de la luz, w_0 y $w(z)$ son los radios, tomados a $1/e^2$ del pico de intensidad a las distancias 0 y z del punto de referencia respectivamente y $R(z)$ es el radio de curvatura del frente de onda después de propagarse una distancia z .

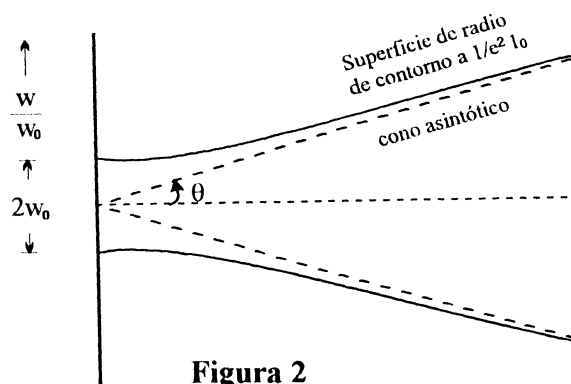
La posición $z = 0$ marca la ubicación de la "cintura del haz" gaussiana o lugar donde el frente de onda es plano y w_0 se llama "radio de la cintura del haz". Simultáneamente $R(z)$ es infinito a $z = 0$, pasa por un mínimo para algún valor finito de z y crece nuevamente hacia infinito asintóticamente como lo hace z , aproximándose al valor mismo de z . En estas circunstancias $w(z)$ se aproxima también asintóticamente al valor

$$w(z) \cong \frac{1z}{\pi w_0}$$

Como z se supone mucho más grande que $\pi w_0^2 / \lambda$, la superficie que marca el $1/e^2$ de la intensidad tiende a un cono de radio angular (divergencia del haz gaussiano)

$$q = \frac{w(z)}{z} = \frac{1}{\pi w_0}$$

con vértice en el centro del haz. La Figura 2 muestra el crecimiento (debido sólo a difracción) del radio del contorno $1/e^2$ de la intensidad en función de la distancia (a partir de la cintura del haz gaussiano).



Observando la Guía de Óptica 3 de Melles Griot podemos, como ejemplo, decir que un láser de He-Ne de 4 mW tiene un diámetro de haz de salida

$$2w_0 = 0,8\text{mm}.$$

De acuerdo con esto, la misma guía nos dice que la divergencia del mismo será

$$q = \frac{1}{\pi w_0} = \frac{632,8 \cdot 10^6 \text{ mm}}{\pi 0,4 \text{ mm}} = 5,04 \cdot 10^{-4} \text{ rad}.$$

A la distancia de $z = 100\text{m}$

$$w(z) = zq = 10^5 \text{ mm} \cdot 5,04 \cdot 10^{-4} = 50,4\text{mm}$$

valor 126 veces más grande que w_0 .

Como para algunas aplicaciones la divergencia del haz conspira con la precisión de los métodos a utilizar, es necesario el aporte de óptica adicional para disminuir θ . El instrumento a tener en cuenta es un telescopio invertido, comúnmente llamado "expansor de haz" (Figura 3), con el que se obtiene una magnificación angular m definida como

$$m = \frac{w_{\text{sal}}}{w_{\text{ent}}},$$

en donde w_{sal} y w_{ent} son mostrados en la Figura 3.

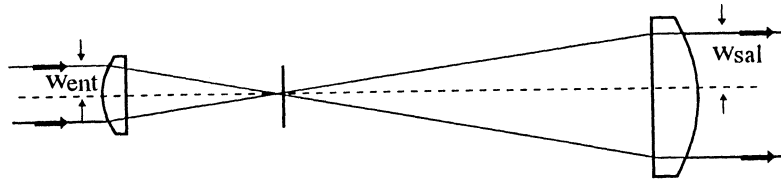


Figura 3

Supongamos que el láser anterior se acopla un expansor con $m = 5$. Esto significa que el sistema Láser-Expansor tendrá a la salida una cintura de radio $5w_0 = 2$ mm. Por lo tanto la divergencia, inversamente proporcional al radio de la cintura, se verá reducida 5 veces ($\sim 10^{-1}$ m rad) y para $z = 100$ m,

$$w(z) = \frac{10^{-5} \cdot 5,04 \cdot 10^{-4}}{5} = 10,07 \text{ mm.}$$

A partir del expansor, el radio sólo crece aproximadamente 5 veces en 100 m. Si la magnificación hubiera sido el doble ($10 w_0 = 4$ mm) el crecimiento del radio del haz sería 1,26 veces.

Los ejemplos expuestos fueron realizados tomando como base un sistema óptico fijo, que genera una cintura de haz exactamente a la salida del mismo. Pero la óptica instrumental puede, con simples cambios estructurales, ofrecernos sistemas también fijos en donde la cintura se encuentre en otro valor de Z distinto de cero o sistemas mecánicamente más complejos por medio de los cuales la posición de la cintura del haz puede variarse a voluntad.

En este punto, cabe preguntarse durante qué distancia la sección del haz permanece con dimensiones parecidas a la de la cintura. Sobre el particular existen varios criterios: uno de ellos, dado por Siegman [1], restringe la variación del w en no más de $\sqrt{2}$ veces, lo que condiciona al intervalo de longitud en una magnitud dada por:

$$2z_R = \frac{2\pi w_0^2}{\lambda}$$

donde w_0 es la cintura del haz en la mitad del intervalo, z_R es el llamado rango de Rayleigh y $w(\pm z_R) = w_0 \sqrt{2}$ es el radio a $1/e^2$ del pico de intensidad en cualquiera de los dos extremos del rango.

Si suponemos que a la salida de un sistema óptico como los ejemplificados anteriormente.

$$w(z) = w_0 \sqrt{2} = 4 \text{ mm} ,$$

la cintura del haz a la distancia dada por el rango de Rayleigh será

$$w_0 = \frac{4}{\sqrt{2}} = 2\sqrt{2} \text{ mm}$$

y

$$2z_R = \frac{2\pi w_0^2}{\lambda} = \frac{16\pi}{632,8 \cdot 10^{-6}} = 7,944 \cdot 10^4 \text{ mm} \cong 80 \text{ m.}$$

Esto implica que a lo largo de esa distancia el cambio en la sección del haz puede ser casi imperceptible para el ojo, constituyéndose dicho haz en una recta de luz perfecta.

Otras consideraciones deben ser tenidas en cuenta en el caso de utilizar sistemas ópticos que permitan enfocar la radiación del láser a distancias variables (con la condición final de tener un haz colimado a $Z = \infty$).

La forma más simple de encarar el problema es determinar la cintura del haz formada por una lente libre de aberraciones cuando incide sobre ella radiación láser. Si el haz incidente es perfectamente colimado en la expresión asintótica

$$w(z) = \frac{\lambda z}{\pi w_0} ,$$

la distancia z debe ser cambiada por la distancia focal f de la lente y $w(z)$ por $w(f) = w$, radio del contorno del haz a $1/e^2$ del pico de intensidad en la lente. Con estos cambios, la cintura del haz en el foco de la lente será:

$$w_0 = \frac{\lambda f}{\pi w} .$$

Un caso no tan simple se presenta cuando radiación láser no colimada atraviesa una lente de distancia focal f . Si la cintura w_0 del haz incidente se ubica a la distancia d_0 de la lente, la radiación resultará enfocada de manera que un nuevo valor de cintura w se obtendrá a la distancia d al otro lado de la lente. Las relaciones entre dichos parámetros serán [3]:

$$\frac{1}{w^2} = \frac{1}{w_0^2} \left(1 - \frac{d_0}{f}\right)^2 + \frac{1}{f^2} \left(\frac{\pi w_0}{\lambda}\right)^2$$

y

$$d = f + \frac{(d_0 - f)f^2}{(d_0 - f)^2 + (\pi w_0^2 / \pi)^2}$$

La profundidad de foco, que en este caso reemplaza al llamado rango de Rayleigh, se define arbitrariamente como la distancia entre los puntos en los que la cintura w se ha incrementado en un 5% y se obtiene a partir de las expresiones previas, siendo la misma

$$\Delta f = \frac{0,32\pi w^2}{\lambda}$$

El análisis referido a la acción de una lente puede extenderse para analizar el acoplamiento de dos lentes con el objeto de formar un sistema telescópico. Para ello la cintura del haz de salida de la primera lente (lente expansora) debe tomarse como cintura de entrada de la segunda lente (lente colimadora).

Todo lo expuesto en este párrafo, que ha sido desarrollado teniendo en cuenta las cualidades que posee la radiación emitida por un láser, será tomado como básico en gran parte de este artículo, que trata fundamentalmente sobre aplicaciones concretas de la óptica y el láser en ingeniería.

Alineación

La más común de las aplicaciones industriales de la óptica y el láser, es la solución de problemas de alineación de componentes. Su empleo es más simple y sencillo que el uso de cuerdas de piano, autocolimadores y otros métodos antiguos y muy elaborados.

El problema central en trabajos de este tipo, se basa en el conocimiento del desplazamiento, a partir de una línea de referencia, de las componentes a alinear. El uso del láser, cuyo rayo de luz se toma como línea de referencia, reemplaza otras formas de operación, proporcionando resultados más precisos y con la participación de menos personal.

Métodos de alineación

Existen distintos métodos que combinan la simpleza o complejidad del sistema a utilizar con el menor o mayor grado de precisión de los resultados a obtener. El más sencillo consiste en proyectar libremente el haz del láser y medir el desplazamiento del objeto a partir del centro del rayo, mediante el empleo de blancos adecuados, solidarios con los elementos a alinear (Figura 4).

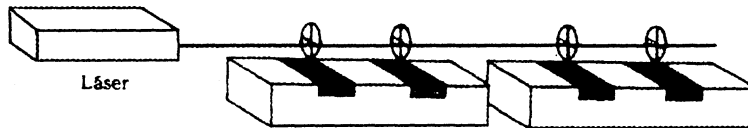


Figura 4

Esto tropieza con las desventajas inherentes al tamaño del haz y su natural divergencia, lo que hace al método poco preciso a partir de cierta distancia, tomando como referencia la boca de salida de radiación.

El empleo del expansor o de un sistema de enfoque para el haz, elimina los dos problemas anteriores y hace al método más preciso y aplicable a casos en donde se deban salvar distancias mayores. La Figura 5 muestra un esquema del sistema completo y de los elementos a utilizar.

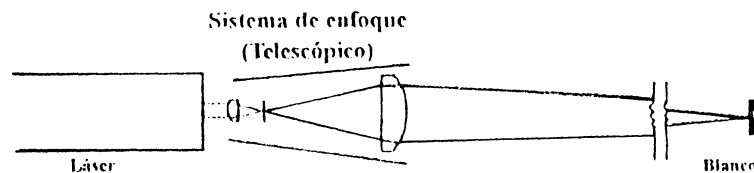


Figura 5

Si la óptica adicional (telescopio invertido) es de buena calidad y cuenta con sistemas mecánicos de precisión, el haz del láser puede ser enfocado continuamente a distancias variables, en puntos pequeños, los que determinan una recta perfecta.

La detección del máximo de intensidad se realiza mediante blancos adecuados (Figura 6) o sobre marcas fiduciaras ubicadas a distancias variables y solidarias a las componentes que se deban alinear. Ejemplos de este último método para alinear puntos en el espacio, se presentan en la bibliografía [3, 4].

El método más completo de alineación es el que emplea un láser y un detector de centrado (detector de cuadrante) como blanco, para determinar el centro exacto del haz [5].

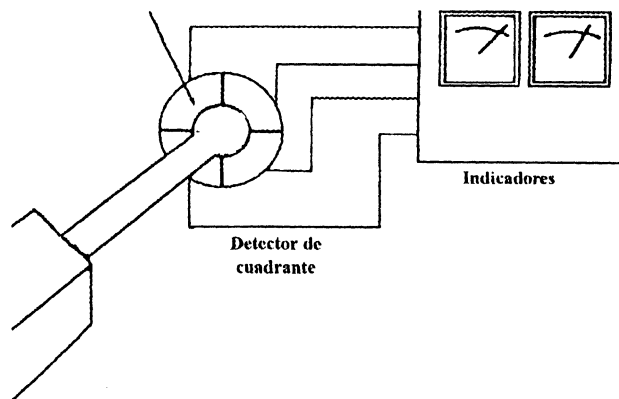
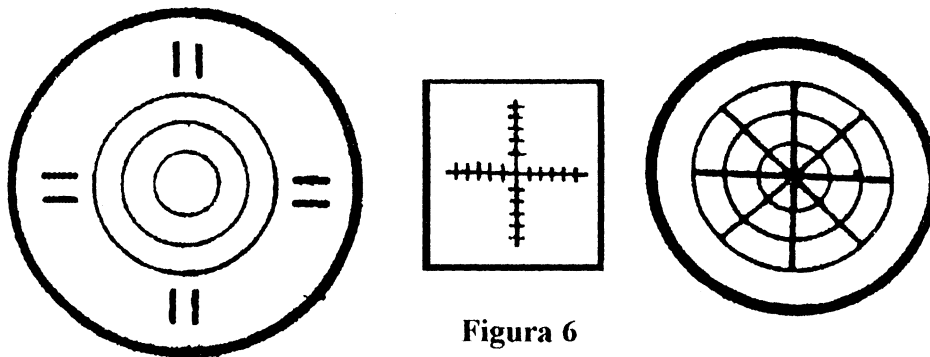
El detector consta de cuatro fotodetectores de silicio conectados en forma de puente, como indica la Figura 7a.

Esta forma de conexión permite comparar horizontal y verticalmente cada una de las mitades del haz con la otra. Si cada cuadrante recibe la misma cantidad de luz, el haz está perfectamente centrado.

En el caso de que el puente, formado por los cuatro detectores, no esté perfectamente balanceado a cero, es decir, que el centro del cuadrante no coincida con su similar del haz del láser, los indicadores x, y previamente calibrados, mostrarán el desplazamiento a partir de la posición de centrado.

El esquema de la Figura 7a representa una versión primaria del sistema, en la que el movimiento central del haz y la escala de sensibilidad eran controlados por medio de perillas manuales convencionales, las que en condiciones de trabajo adversas (movimientos, vientos, etc.) generaban errores en las medidas, que se sumaban a las de observación de los indicadores.

La Figura 7b presenta una reproducción fotográfica de una nueva versión, en la que todo el comando del equipo se realiza por medio de una computadora, la que -además- muestra en la pantalla un diagrama x, y en donde se observa la posición del centro del haz, un detalle de los datos que determinan la medida y opciones que permiten al programa un manejo adecuado de los resultados obtenidos.



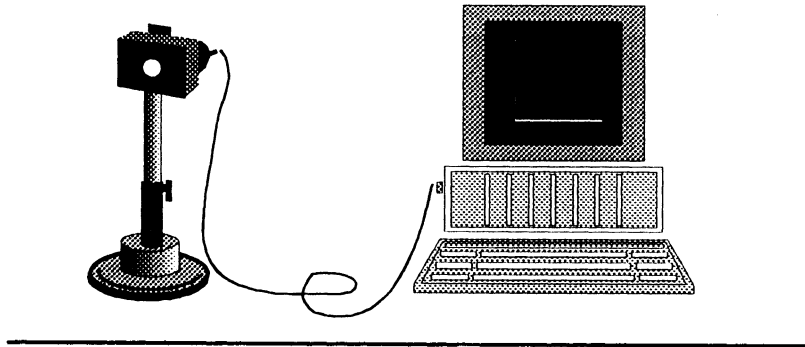


Figura 7 b)

La divergencia del haz para un pequeño láser de He-Ne, estimada en 10^{-3} rad, representa una limitación en la precisión de la medida realizada por métodos simples. Para dar una idea de lo que significa un haz con esa divergencia, podemos decir que precisiones de 1 mm en 10 m requieren $0.1 \cdot 10^{-3}$ rad como máximo de apertura angular. Sin embargo, este tipo de inconvenientes puede ser obviado, ya que la óptica instrumental provee sistemas sencillos para expandir y colimar el haz hasta valores que oscilan entre 10^{-5} y 10^{-6} rad en la divergencia de la radiación.

El uso de un sistema de enfoque continuo puede incrementar aún más las posibilidades del láser y aumentar la precisión del método ya que, como dice Siegman [1], el centro del haz puede ser localizado fotoeléctricamente con un error no mayor de $w_0 / 1000$. Prácticamente se considera [5], al usar este tipo de sistema óptico, un error del orden de $1 \mu\text{m}$ por metro lo que permite determinar la recta con grados de precisión no alcanzables por ningún otro método. Cálculos más precisos sobre posición, cintura del haz enfocado, profundidades de foco y errores estimados surgen de la aplicación de fórmulas descriptas en el párrafo anterior.

Esto se establece como cierto, siempre que el láser opere en el modo TEM_{00} gaussiano; la precisión disminuye si la radiación es multimodal.

También aparecen otros inconvenientes cuando, al usar láseres comunes, cambia la temperatura del medio exterior, ya que 1°C de variación puede implicar un desvío de hasta $10''$ de arco en la dirección del haz. Aunque este inconveniente puede ser salvado en parte, operando el láser luego de un cierto tiempo de encendido o trabajando el mismo en un lugar climatizado, existen láseres contruidos especialmente para aplicaciones en alineación. Ellos son fabricados de tal manera que el haz emerge del tubo con una precisión de $25 \mu\text{m}$. Este láser especialmente diseñado provee una estabilización en la dirección del haz de aproximadamente $1''$ de arco por cada grado en el cambio de temperatura.

El ejemplo más común de la aplicación de estas técnicas es en el control de máquinas herramientas. Esto quiere decir, un detector de cuadrante solidario con el portaherramienta de la máquina para medir punto a punto del movimiento, el apartamiento, según dos ejes perpendiculares, de la recta teórica que debería seguir su recorrido y graficarlo como muestra la Figura 8.

Algunas aplicaciones de la óptica y el láser en problemas de alineación o control requieren un montaje especial para el detector. El caso típico es el de colocar en línea recta los ejes de diferentes superficies cilíndricas o verificar el grado de colinealidad de las mismas. Para solucionar este problema, los técnicos ya han diseñado y construido montajes llamados "araña" ajustables de diferentes diámetros (Figura 9), los que colocan con gran precisión el blanco o detector, en el centro de la superficie. A partir de este momento, el sistema a láser alinea o verifica los diferentes centros.

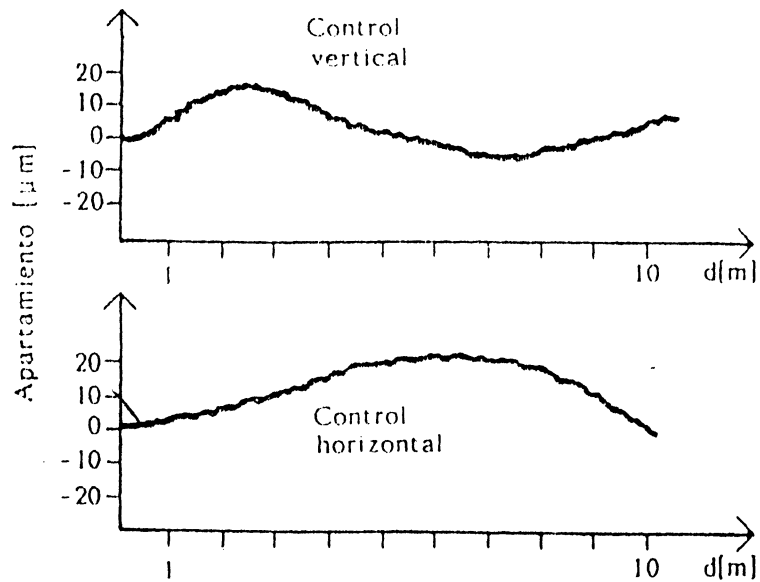


Figura 8 a)

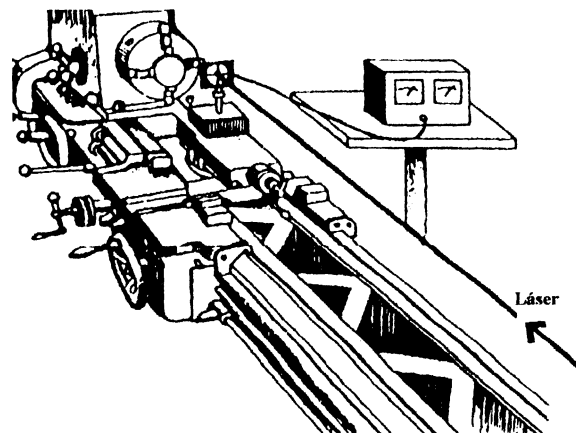


Figura 8 b)

La calidad del sistema de arañas es determinante para la exactitud del centrado o la precisión en el control de alineación.

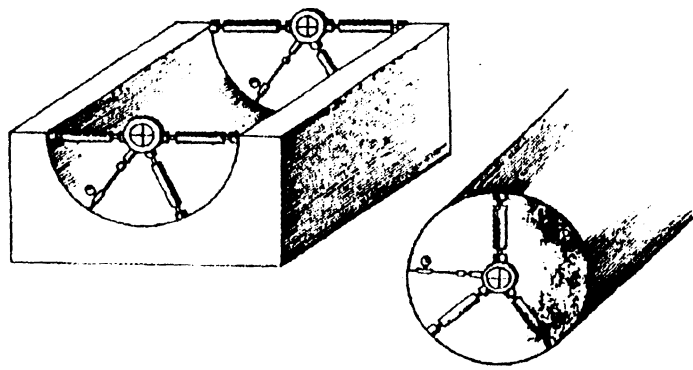


Figura 9

Posicionamiento y control de movimientos

Observando las cualidades de los distintos sistemas descritos en el punto anterior y utilizando los mismos elementos, se pueden generar nuevas aplicaciones. La primera de ellas simplemente consiste en el posicionamiento de algún elemento, tomando como base la línea que genera el haz del láser.

A partir de la determinación de la posición de un punto mediante el uso del láser y de un detector de cuadrante solidario con el cuerpo bajo estudio, se puede controlar el estado de movimiento en un plano perpendicular al haz, tomando la posición según el eje x o y a intervalos de tiempo.

Para el control del movimiento de la dirección z , sería necesario otro sistema láser-detector, y para otros grados de libertad, rotaciones, etc., pueden ser verificados mediante el uso de espejos-blancos (o semiespejos), montados sobre el elemento a controlar. La observación del haz reflejado por el espejo permite, mediante simples relaciones de la óptica geométrica, realizar el control del movimiento angular.

Un detector de cuadrante o algún otro sistema de medida, colocado de manera adecuada, controlará a distancia el movimiento angular del blanco, midiendo la desviación en cualquier dirección, a partir de una posición tomada como punto de partida (Figura 10).

Una aplicación concreta de estas técnicas está descrita en el Informe CIC número 19 [6].

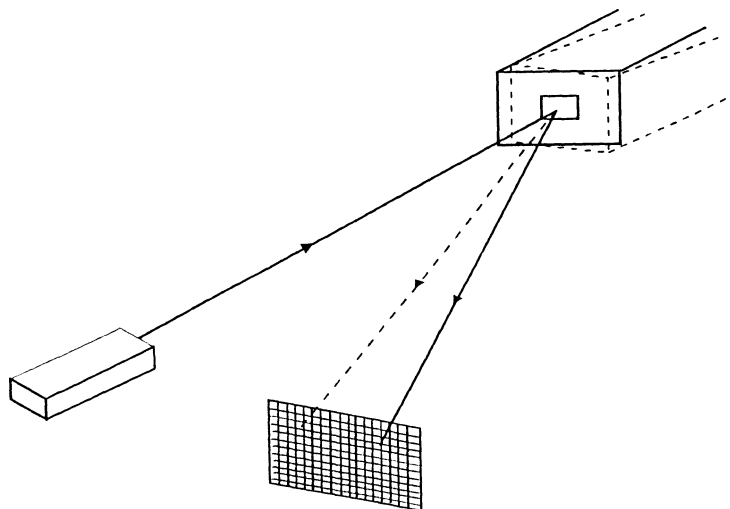


Figura 10

El uso como blanco de un detector de cuadrante, tiene un límite a partir del hecho que el desvío, desde una posición inicial, puede superar las posibilidades de observación al caer el haz de luz fuera del área del instrumento. Los más comunes permiten desplazamientos de hasta 1,75 cm., llegando en casos especiales a 5 cm. pero con precisiones algo menores.

Para detectar y medir desplazamientos, en una dirección y hasta de 10 cm., se cuenta con arreglos especiales, que determinan la posición del centro del haz de luz con precisiones, en condiciones óptimas, comparables al detector de cuadrante. Dichos arreglos consisten en dos detectores de Silicio en forma de cuña, colocados de manera inversa, de los que se extrae una señal que tiene en cuenta la intensidad que recibe cada cuña.

El arreglo lineal no permite la medida de desplazamientos laterales, los que deberán ser controlados por otros arreglos paralelos o cruzados.

El control a grandes distancias de pequeños desplazamientos, puede ser tratado por medio de los procedimientos ya enunciados en la primera parte de este párrafo.

El inconveniente se presenta, cuando los desplazamientos superan las dimensiones de los detectores de cuadrante o lineales y deben ser observados desde grandes distancias.

Una forma adecuada para resolver el problema, tiene en cuenta el uso de telescopios en las que en su punto de observación se coloca una cámara que determina, a tiempo real, el desvío o que permite almacenar la información para su posterior evaluación. Ambas acciones requieren una calibración previa.

El método, que no tiene el grado de precisión de los dispositivos anteriores, pero que se acerca al mismo si se toma en cuenta el error relativo, consiste en observar a través del telescopio, el desplazamiento de una "marca" colocada en el lugar bajo estudio. Dicha "marca" debe ser visible y precisa en todo momento de la experiencia y puede materializarse mediante el empleo de líneas y/o círculos, fuentes luminosas convencionales, leds o láseres.

Aplicación en casos de inestabilidad atmosférica

La inestabilidad atmosférica presenta inconvenientes para el uso de sistemas ópticos de alineación, posicionamiento y control.

Cambios en la temperatura y turbulencia del aire, dentro del camino que sigue la luz, implican fluctuaciones en el índice de refracción. Esta inestabilidad del medio origina movimientos al azar del haz que se propaga a través del él y, si la distancia a recorrer es grande, la variación en la posición final puede cambiar en el orden de los centímetros.

Existen varios métodos para solucionar este tipo de inconvenientes. El primero, usado en distancias cortas, es muy simple, consistiendo solamente en enviar el haz a través de un tubo o el de soplar aire a lo largo de la dirección de alineación, con la que en ambos casos se reducirá el problema de la turbulencia del medio.

El segundo método se basa en el hecho que, estadísticamente, la distribución de luz promediada a lo largo de un cierto tiempo, es simétrica alrededor de la posición del haz sin perturbar. Por lo tanto, con la ayuda de los detectores de cuadrante y de electrónica adicional, es posible obtener con gran precisión la posición del centro del haz, integrado en el intervalo de tiempo.

Diversos equipos han sido construidos para resolver este tipo de problemas [5] y los resultados muestran que se puede alcanzar una desviación máxima de 0,2" de arco en una distancia de algunos kilómetros, lo que significa un error en la medida de 1 mm/Km.

Generación y control de rectas y superficies

Como se observó en párrafos anteriores, existen varios métodos ópticos para materializar una recta en el espacio y poder, a partir de ella, alinear componentes.

Sin embargo, utilizando el mismo sistema y con el agregado de óptica adicional, se pueden realizar otras materializaciones relacionadas [3,4], que tienen su punto de partida en las siguientes construcciones geométricas:

a) Recta perpendicular a otra dada

Se proyecta el haz de luz sobre un pentaprisma y luego de realizar dos reflexiones en las superficies espejadas (Figura 11), el mismo emerge formando un ángulo de 90° con la dirección original.

Debemos hacer notar que, en general, el ángulo no es exactamente 90° , pero sin recurrir a una óptica extremadamente elaborada, se puede lograr que el error no exceda de los 5" de arco, lo que implica una gran precisión en la materialización de la recta perpendicular.

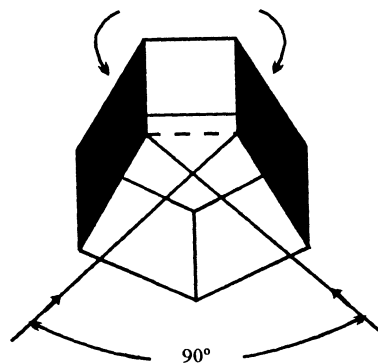


Figura 11

b) Plano perpendicular a una recta dada

Haciendo rotar el pentaprisma alrededor de un eje coincidente con la dirección del haz incidente, la radiación emergente genera un plano perpendicular a la recta que contiene al rayo inicial.

c) Determinación de planos

Cuando un haz de luz incide sobre un espejo plano que rota, y si el eje de rotación es normal a la dirección del haz, la superficie generada por la radiación emergente, es un plano

La aplicación de estos métodos está limitada a aquellos casos donde el haz luminoso puede propagarse libremente a partir de un punto. Cuando la zona de trabajo presenta obstáculos que se interponen en el camino de la luz, deben utilizarse otros arreglos experimentales que permitan llegar con el plano de referencia a todos los puntos posibles. Este tipo de problemas es muy común en el control de turbinas, generadores, grandes motores, etc., que se encuentran ya instalados con equipos suplementarios propios o externos y en recintos que no cuentan con espacio libre.

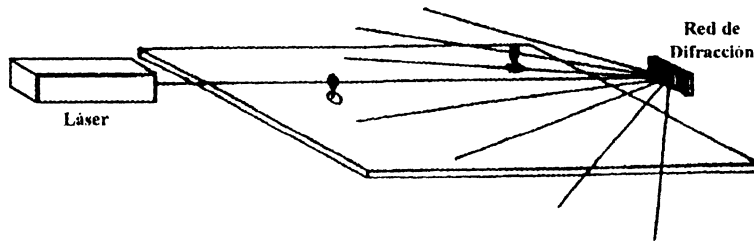


Figura 13

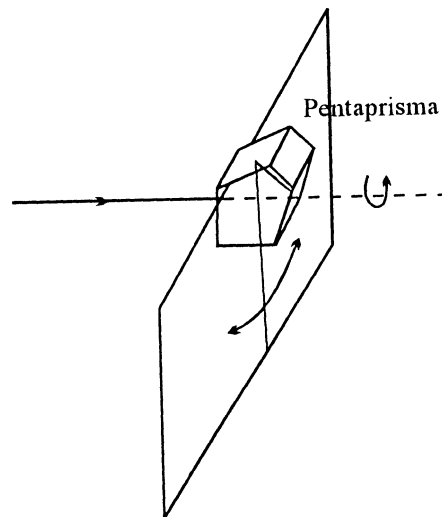


Figura 14

Para este tipo de problemas, describimos aquí un método para generación de planos bajo condiciones de libertad restringida para el paso del haz, basado en el trazado de dos rectas que se cortan. Estas generan un plano; luego, cualquiera recta que las corte simultáneamente, yacará en el mismo plano [8].

Las rectas originales se determinan de la siguiente manera: una, mediante un rayo de luz proveniente de un láser, y la otra, por su reflexión en un espejo fijo E_1 , Figura 15a. De esta manera es muy sencillo generar la parte básica del plano y lograr su ubicación más conveniente, variando a voluntad las posiciones y orientaciones del láser y del espejo E_1 .

Para determinar otras rectas pertenecientes a ese plano, las que a su vez definirán los puntos del mismo, se necesita cierta instrumentación adicional. Un detector de cuadrante colocado en el punto B de la Figura 15b, montado sobre un soporte que gira alrededor de un eje colineal con el eje del detector, determina con precisión la posición del plano en dicho punto. Y un segundo espejo E_2 ,

interpuesto en el haz original en el punto A (Figura 15b), es orientado de modo que la radiación reflejada sea dirigida al punto B. El detector allí ubicado es rotado para enfrentar el rayo, comandando el ajuste fino del espejo hasta lograr que la radiación esté concentrada exactamente en dicho punto. Este rayo simulará la intersección de la recta secundaria con las dos principales, según lo propuesto.

A partir de la construcción de la primera recta secundaria, se pueden materializar otras rectas pertenecientes al plano, con sólo cambiar las posiciones de A y/o B, manteniendo siempre en su posición original el espejo E_1 . De esta manera se pueden determinar todos los puntos de interés posibles sorteando, con sólo cambiar las rectas secundarias, cualquier obstáculo que impida el uso de técnicas más convencionales.

La inclusión en el punto A de la Figura 15b de una red de difracción plana (por transparencia o por reflexión) en lugar del espejo E_2 permitirá tener líneas pertenecientes al plano, que transpongan los límites marcados en la Figura 15a, por el haz láser y su reflexión en E_1 . Estas líneas adicionales (Figura 16) pueden ser utilizadas para controlar puntos o superficies que se encuentren en zonas interiores, sin salida para el rayo de luz de referencia.

Los arreglos experimentales mostrados en las Figuras 15 y 16 conforman sistemas sencillos pero con un alto grado de aplicabilidad para controlar o generar distintos planos.

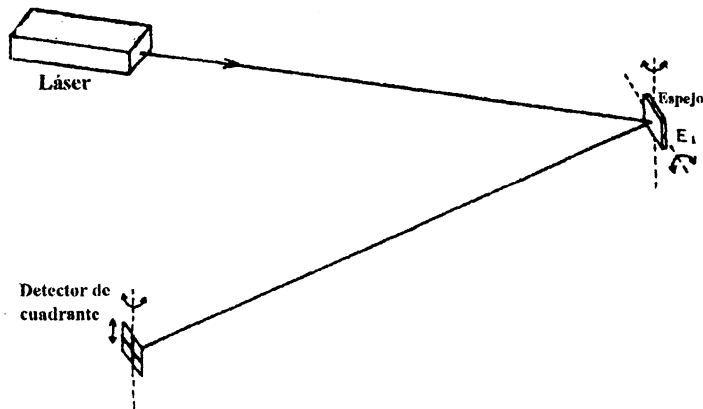


Figura 15 a)

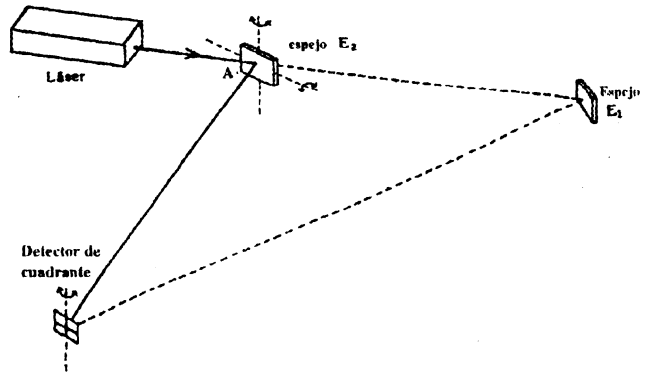


Figura 15 b)

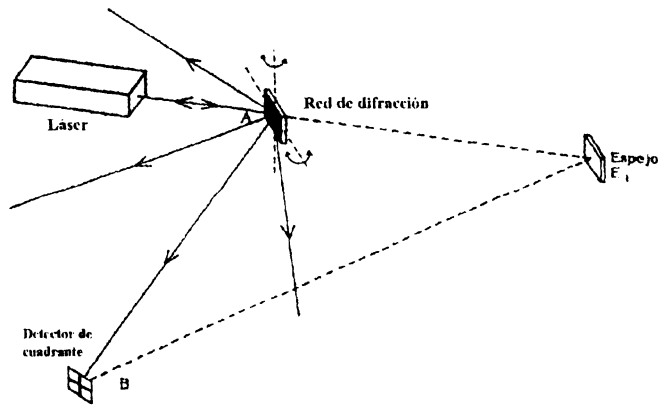


Figura 16

Otras posibilidades de la óptica y el láser

Otras aplicaciones del láser en alineación serían las concernientes a establecer la dirección vertical del lugar y la materialización de planos de nivel.

Para el primer problema, es necesario proveerse de una pequeña batca de mercurio, que forma un espejo plano horizontal, base del experimento [5]. Luego, tal cual se observa en la Figura 17a, se coloca el espejo normal al eje del autocolimador y se mueve el láser hasta que las reflexiones en el espejo y en el mercurio converjan en el retículo del telescopio. En ese momento, la dirección del haz del láser marca la vertical del lugar con una precisión que puede alcanzar a ser menor que 1" de arco.

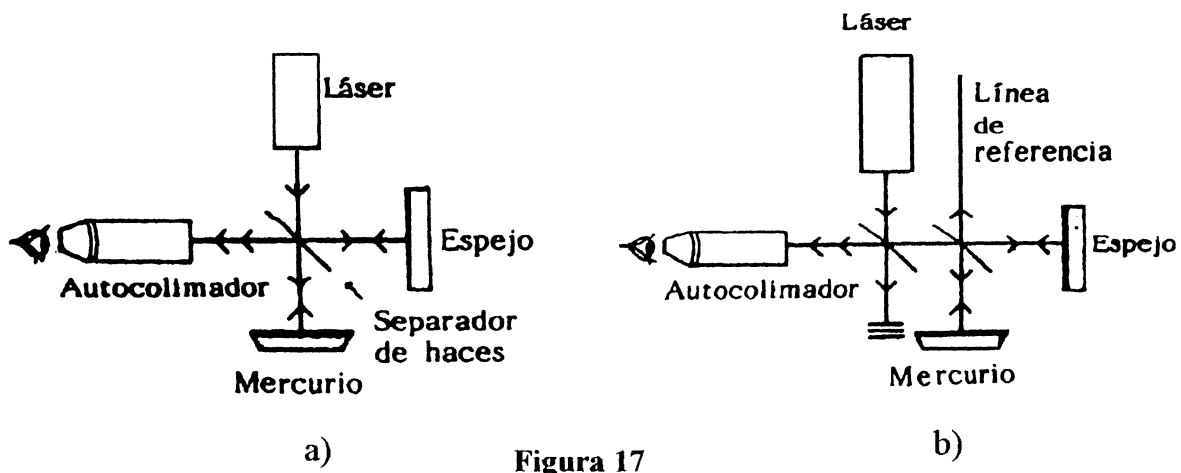


Figura 17

Se puede presentar otro arreglo experimental, eliminando el autocolimador, pero observando la interferencia de los dos haces en una pantalla, en la forma clásica de un interferómetro de Michelson.

Una variante del mismo método provee un haz vertical hacia arriba. La Figura 17b muestra el dispositivo en el que el uso de un segundo separador de haces genera el haz vertical a

partir del cual, utilizando lo visto en puntos anteriores, se pueden construir planos perpendiculares a diferentes alturas, los que constituyen en sí planos de nivel.

Existen sistemas diseñados especialmente para trabajos de campo, que se usan para controlar la maquinaria que construye caminos, nivela superficies, etc. Uno de estos sistemas consta de un trípode sobre el que está montado un láser en posición vertical o un sistema como el descrito en el punto anterior, que proyecta su luz sobre un prisma de 90° rotante, el que genera un plano que puede ser materializado a la altura y con la inclinación deseada. Máquinas provistas de sensores adecuados, alertan al operario, mediante indicadores, luces o sonidos especiales, si la posición no es correcta o si no se alcanzó el nivel adecuado (Figura 18).

Obviamente, sistemas más complejos y costosos pueden operar la maquinaria automáticamente, pero la precisión del orden del centímetro en cientos de metros que se logra con aparatos sencillos, vuelve innecesario el gasto.

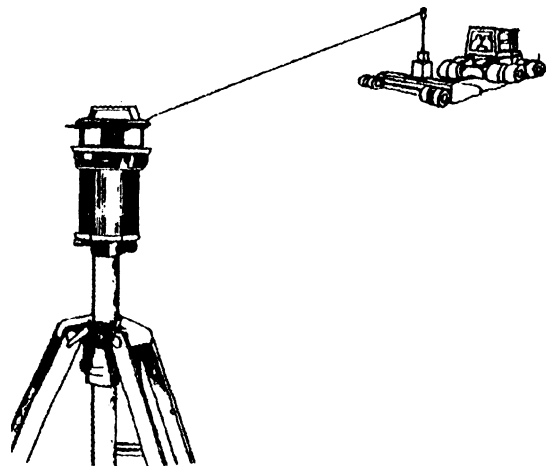


Figura 18

La óptica y el láser en problemas de mediciones

Los métodos ópticos de medidas son conocidos y usados desde hace muchos años, pero siempre han estado limitados por las características de las fuentes luminosas.

Desde la invención del láser en 1960, la óptica primero y las aplicaciones tecnológicas de la misma después, sufrieron un impulso vertiginoso, ya que el láser provee la luz con que soñaron los ópticos de todos los tiempos. Por lo tanto, los antiguos métodos de la óptica, respaldados por la presencia de una emisión luminosa de potencia, direccional y de alta coherencia, resurgen como importantes herramientas en aplicaciones industriales.

El empleo de los métodos de la óptica y el láser que describiremos, involucra el conocimiento de:

- el uso del interferómetro de Michelson en problemas de medición;

- la producción de pulsaciones cuando interfieren en un punto del espacio dos frecuencias muy cercanas;
- la generación de corrimientos en frecuencia usando el efecto Doppler.

Medidas de distancia por métodos interferométricos

Las características principales de la radiación que va a ser utilizada, son su gran intensidad y su alta coherencia temporal. Pero, además, debemos asegurar su estabilidad en frecuencia.

Si consideramos un láser de He-Ne operando a 632,8 nm, eso significa $5 \cdot 10^{14}$ Hz ($5 \cdot 10^5$ Ghz) y su variación dentro del ancho de línea puede llegar a 10^9 Hz (10^3 MHz). Por lo tanto, para no sufrir variaciones que puedan alcanzar a ser del orden de una parte en 10^5 , importantes en este tipo de medidas, es necesario trabajar con láseres estabilizados. Esto involucra la construcción de un dispositivo térmicamente controlado, con espejos rigidamente montados y con un servocontrol de la longitud de la cavidad. Esto asegura la frecuencia estabilizada en el centro de la línea de emisión, aunque con el decrecimiento de la potencia del láser, que en estas condiciones no supera el milivatio.

Como espaciado entre modos longitudinales es $C/2L$, donde C es la velocidad de la luz y L la longitud de la cavidad del láser, acortando L, los modos estarán más separados y si esta separación llega a ser el ancho de la línea de emisión, sólo un modo resonará en la cavidad. Controlando la posición de éste, en el centro de la línea podemos tener la frecuencia estabilizada dentro de una parte en 10^9 .

Si las tolerancias del trabajo a realizar así lo requieren, se pueden utilizar láseres menos controlados, no perdiendo los métodos a usar sus características de interferométricos ni un grado adecuado de precisión.

Los métodos de medida de distancia que describiremos en este párrafo, se basan esencialmente en el interferómetro de Michelson y su rango de aplicación no va más allá de algunas decenas de metros en condiciones ambientales adecuadas.

La Figura 19 muestra un diagrama de la disposición de los elementos a emplear para la medida interferométrica de distancias.

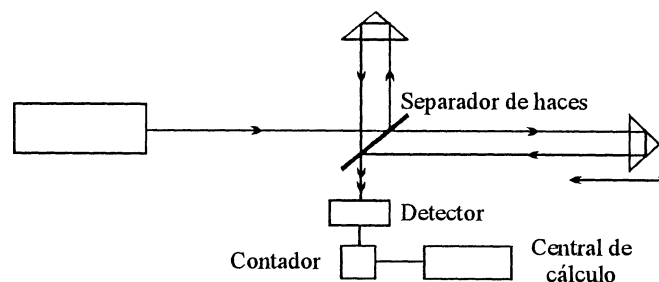


Figura 19

El diagrama nos muestra un interferómetro de Michelson adaptado, mediante retrorreflectores, a las exigencias del problema, ya que éstos aseguran el retorno del haz en la misma dirección aunque las condiciones angulares del reflector cambien. Para observar esta particularidad, la Figura 20 muestra un esquema de un retrorreflector, que en esencia no es más que un triedro reflector, en el que la luz incidente, después de reflejarse en las tres caras del mismo, sale siempre en la dirección opuesta.

Por lo tanto, considerando nuevamente la Figura 1, los dos rayos obtenidos a partir del divisor de haces, después de reflejarse retornan, son recombinados formando la figura de interferencia y viajan hacia el detector. La amplitud de la luz frente al mismo, depende de la diferencia de fase (o de camino óptico) entre los rayos provenientes de las dos ramas del interferómetro.

Suponiendo que la dirección de observación coincide con el eje del haz, la diferencia de fase será dada por

$$\Delta = \frac{2\pi}{\lambda} s$$

donde λ es la longitud de onda de la luz empleada y S la diferencia de camino entre las ramas del interferómetro.

Cuando la rama móvil se desplaza $\lambda/2$, S será igual a λ y la fase de la luz en el detector cambiará 2π . Esto implica que el sistema de franjas se moverá un periodo, es decir, que comprenderá un cambio de oscuridad por luz en el detector. Esta modulación en la amplitud es registrada por el detector y tomada en cuenta a través de un circuito electrónico de conteo. Llevada la información a una central de cálculo, ésta determinará el camino recorrido.

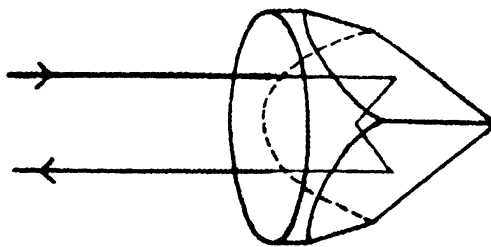


Figura 20

Es importante hacer notar que el interferómetro no determina una medida absoluta de la distancia, sino relativa, a partir de un cero arbitrario y que además, la misma constituye el camino óptico que difiere del real por un factor que es el índice de refracción.

Dicho índice tiene, para aire seco a 15°C, presión normal y con 0,03% de CO₂, un valor

$$n = 1,0002765$$

a la longitud de onda del láser de He-Ne, que en vacío es

$$\lambda = 632,99138 \text{ nm.}$$

Este valor del índice de refracción debe ser corregido por variación de humedad, temperatura y presión [9]. Por lo tanto en el momento de trasladar de camino óptico a camino físico, la central de cálculo debe tener en cuenta todas estas consideraciones.

Este método interferométrico es sensible a cambios en la intensidad del láser y como ésta puede variar por alteraciones atmosféricas, la medida estará sujeta a posibles errores que surgirán del mal conteo de las franjas.

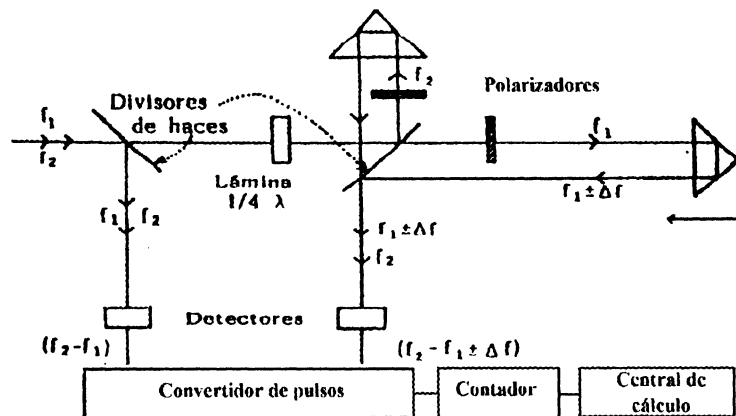


Figura 21

La Figura 21 muestra el esquema de un sistema que mide esencialmente, la velocidad del retroreflector móvil, independiente de la intensidad de la radiación, la que luego integrada nos determinará el camino recorrido. Para este caso, el láser de He-Ne emite luz de dos diferentes frecuencias y propiedades de polarización, que se logran aplicando un campo magnético axial que separa la línea del Ne en dos componentes de frecuencias f_1 y f_2 .

Las frecuencias f_1 y f_2 , separadas aproximadamente por 1 GHz, viajan por caminos diferentes debido a la presencia de polarizadores colocados en el camino de ida de cada una de las ramas. Luego de la reflexión, los dos rayos son recombinados por el separador de haces y producen un "batido" de frecuencias $f_2 - f_1 \pm \Delta f$ donde Δf es la variación de frecuencia dada por el efecto Doppler. Su valor es

$$\Delta f = \frac{f_1 v}{c}$$

y v representa la velocidad de movimiento del retroreflector.

Usando como señal de referencia $f_2 - f_1$, obtenida del primer separador de haces, el convertidor de pulsos extrae Δf y produce un pulso por cada $\lambda/4$ de movimiento del reflector.

Para estos métodos de medida interferométrica tenemos como fuentes generadoras de error, las siguientes:

- a) Error debido al instrumento. Este puede en principio medir dentro de la fracción de longitud de onda, a menos de un error en el conteo de las franjas. Algunos instrumentos comerciales miden dentro de los 50 m, en entornos cerrados, con errores del orden de $1\mu\text{m}$.
- b) Error debido a problemas atmosféricos y de efectos de temperatura sobre el material. Estos errores son generalmente compensados en forma automática y caen dentro del error instrumental.
- c) Errores debidos a una imperfecta alineación entre el interferómetro y la dirección de movimiento, por ejemplo, de la máquina de controlar. El uso de detectores de cuadrante, ya explicado con detalle, soluciona totalmente este tipo de problemas.
- d) Error debido a la posición arbitraria del cero. Si esta posición corresponde a una distancia d_1 desde el interferómetro y debemos medir d_2 , la compensación automática actúa sobre d_2 y no sobre $d_1 + d_2$, que es la distancia que realmente recorrió la luz. El error puede ser compensado utilizando expresiones simples que tienen en cuenta el camino muerto d_1 [10], pero la solución más elemental consiste en colocar el interferómetro lo más cerca posible de la posición cero de medida.

Una gran cantidad de ejemplos muestran la eficacia de la interferometría láser para la medida, control de movimiento y velocidad, y varias otras aplicaciones industriales, dentro de un rango que puede llegar a los 50 m.

La limitación de distancia puede depender de la coherencia de la radiación, ya que el sistema de franjas tiende a "desdibujarse" cuando se aproxima al límite de la longitud de coherencia.

La diferencia de camino S debe ser menor que $C\Delta f$. Ahora bien, si tenemos un láser estabilizado, Δf es el orden de 10^5 Hz, lo que implica S aproximadamente igual a 300 m. Este valor calculado en el ejemplo es sólo teórico, ya que las franjas también sufren los efectos atmosféricos en los recintos de trabajo y pierden definición para distancias superiores a 40 m.

En este párrafo nos hemos propuesto dar sólo algunas ideas acerca de las ventajas de la interferometría láser, aunque la bibliografía muestra una larga lista de posibles aplicaciones [10, 11].

Conclusiones

Esta presentación acerca de las posibilidades de la óptica y el láser en problemas de alineación, posicionamiento, control y medidas, en muchos casos puede ser considerada como propuestas teóricas de laboratorio. A partir de este punto y con la experiencia de un especialista en

temas de ingeniería, se pueden generar un sinfín de aplicaciones a casos que en épocas anteriores presentaban una difícil e imprecisa solución.

Las referencias bibliográficas presentadas y otras que se pueden consultar, toman la forma de testigos para confirmar la eficacia de las ideas, y de los métodos expuestos, que tienen a la óptica y a la radiación láser como herramientas básicas para el tratamiento de problemas tecnológicos.

Bibliografía

- [1] Introduction to Lasers and Masers. A.E. Siegman. Mc Graw-Hill, (1986).
- [2] Melles Griot. Optics Guide 3, (1985).
- [3] Alineación empleando un sistema óptico de enfoque y un láser. Aplicaciones a la Industria Naval. L.M. Zerbino, H.J. Rabal, H.F. Ranea Sandoval, J.O. Tocho, M. Gallardo y M. Garavaglia. Ingeniería Naval (España) 524, 48, (1979).
- [4] Aplicaciones del láser a la metrología. L.M. Zerbino, H.J. Rabal, H.F. Ranea Sandoval, J.O. Tocho, A.P. Laquidara, M. Gallardo y M. Garavaglia. Carta Metrológica 3, 9, (1981).
- [5] New laser applications in Geodetic and Engineering Surveys. A. Chrzanowsky, F. Ahmed y B. Kurz. Applied Optics 11, 319, (1972).
- [6] Measuring flatness of large tables: New Methods. H.F. Ranea Sandoval and M. Gallardo. Applied Optics 25, 1370, (1986).
- [7] Boletín Informativo. CIC Provincia de Buenos Aires, Argentina, pág. 6, Septiembre, (1984).
- [8] Método de generación de superficies. M. Gallardo. Informe interno CIOp, (1986).
- [9] Optical Refractive Index of Air: Dependence of Pressure, Temperature and Composition. J.C. Owens. Applied Optics 6, 51, (1967).
- [10] Industrial Applications of Lasers. J.F. Ready ed. Academic Press, (1978).
- [11] Laser Applications. I.M. Ross ed. Academic Press, (1971).

METROLOGIA SPECKLE

E.E. Sicre

Introducción

Cuando una superficie difusora, como por ejemplo un metal despulido, se ilumina con un láser, parece estar cubierta por una estructura granular muy fina. A diferencia de lo que sucede con luz blanca, todos los puntos de la superficie son coherentes, es decir, emiten ondas luminosas capaces de interferir. La imagen de cada uno de estos puntos formada en la retina de un observador es una figura de difracción propia del sistema óptico del ojo. La interferencia de dichas figuras de difracción origina el aspecto granular aleatorio característico, denominado "speckle". El mismo fenómeno se produce si se reemplaza el ojo por una cámara fotográfica. En realidad, no es necesario formar una imagen de la superficie difusora para obtener un diagrama de speckle. En efecto, la superficie rugosa genera ondas secundarias con fase aleatoria, que interfieren entre sí en todo el espacio circundante al difusor. La superposición coherente tridimensional de tales ondas, a una distancia finita de la superficie, se denomina speckle de Fresnel.

Los diagramas de speckle, considerados como interferencia de múltiples haces, no poseen una posición de referencia en el espacio (como en el caso de la holografía) y por consiguiente no son, por sí mismos, de utilidad metrológica. Por esta razón, los primeros trabajos sobre holografía interferométrica no reconocían la importancia del speckle, tratándose más bien de reducir su efecto a fin de mejorar la calidad de las imágenes reconstruidas. Sin embargo, a partir de la observación de que el speckle era sensible al movimiento de la superficie difusora, se comenzaron a desarrollar técnicas que utilizaban al diagrama de speckle como una portadora aleatoria, a la que se podía modular información sobre cambios locales de la superficie. En las secciones que siguen se analizarán algunas de estas técnicas y sus correspondientes aplicaciones metrológicas.

Propiedades físicas del speckle

En esta sección se analizarán los factores que determinan la forma de los granos individuales del speckle, como así también la distribución de intensidad de la envolvente del diagrama.

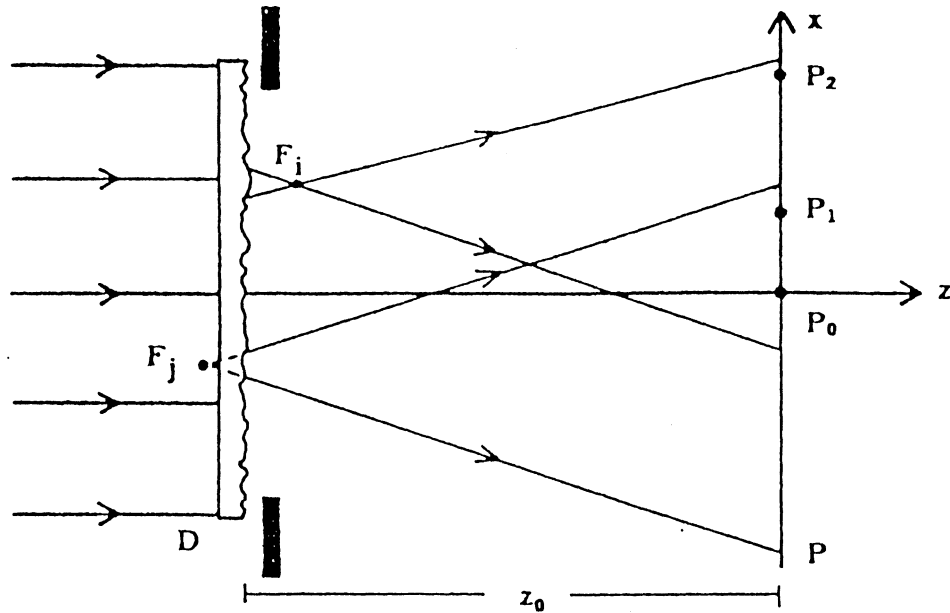


Figura 1

En la Figura 1 se muestra un objeto difusor por transmitancia (por ejemplo, un vidrio despulido) iluminado por una onda plana coherente proveniente de un láser. El análisis que sigue puede asimismo extenderse a objetos difusores reflectantes. Un tratamiento simple del fenómeno consiste en considerar las irregularidades superficiales como un conjunto aleatorio de microlentes, convergentes y divergentes, de apertura numérica variable. En la Figura 1 se ilustran dos de dichas lentes elementales, una convexa (positiva) que enfoca la luz en el punto F_i y otra cóncava (negativa) de la cual la luz diverge como proveniente del punto F_j . De esta manera, se puede reemplazar la superficie del difusor D por un conjunto aleatorio de puntos $\{F_i\}$, representando los focos de las lentes elementales, los que a su vez originan ondas esféricas coherentes cuya interferencia mutua producirá la distribución de amplitud en el plano de observación P . Estas ondas tendrán entre sí fases aleatorias debido a la variación también aleatoria del espesor óptico del difusor D .

En un punto del plano de observación tal como P_1 , próximo al punto central P_0 del diagrama, existirán contribuciones de amplitud luminosa de todos los elementos del difusor. Por el contrario, en un punto tal como P_2 más alejado del centro, existirán contribuciones sólo de aquellos elementos de D que originen difusión de luz ("scattering") de ángulo amplio. Por esta razón, es de esperar que la envolvente de intensidad del diagrama de difracción tenga un valor mayor en P_1 que en P_2 . La forma específica de la curva envolvente dependerá de la desviación standard de los tamaños y curvaturas de los elementos difusores. Por ejemplo, si todos fueran idénticos, las ondas difundidas tendrían la misma extensión angular y por consiguiente habría una envolvente de intensidad que sería casi-constante dentro de una región de P igual a la del difusor, disminuyendo muy rápidamente en los bordes. Cuanto mayor sea el rango de tamaños de los elementos difusores, mayor será el rango angular de la distribución de intensidad producida. En este caso, existirá una región central de intensidad constante menor, y una más gradual transición de intensidad hacia los bordes del diagrama. Esta dependencia de la figura de intensidad del diagrama de speckle con la microestructura de la superficie difusora puede ser utilizada para medir los principales parámetros de la rugosidad superficial, como se verá más adelante.

La estructura detallada de los granos individuales del speckle está determinada por la pupila del difusor. Considérese el diagrama de speckle producido en el plano P como una figura de interferencia de las diversas ondas esféricas coherentes. De esta forma, la distribución de amplitud luminosa en P puede interpretarse como una superposición de franjas de Young, de diferentes orientaciones y espaciados, que resultan de la interferencia de las ondas provenientes de dos elementos difusores cualesquiera. El valor mínimo del espaciado corresponderá a las franjas de Young producidas por los elementos situados en los bordes del difusor. Por lo tanto, si el difusor tiene un diámetro a , el espaciado mínimo Δ de las franjas será:

$$\Delta = \lambda z_0 / a, \quad (1)$$

siendo z_0 la distancia entre el difusor y el plano de observación. La escala de los granos individuales en el diagrama de speckle está dada predominantemente por el valor de Δ . El análisis efectuado permitió derivar la forma general de la envolvente y el tamaño de speckle más pequeño del diagrama de difracción. Así, si el área iluminada del difusor se reduce, la envolvente del diagrama de difracción es poco afectada, pero el tamaño promedio de los speckle individuales aumenta inversamente proporcional al diámetro de D , de acuerdo con la Ecuación(1). Sin embargo, para analizar otros parámetros importantes de los diagramas de speckle, como por ejemplo el contraste, deben emplearse consideraciones estadísticas. En el caso de que la superficie difusora posea una estadística gaussiana, puede demostrarse que el contraste del speckle vale uno. Para superficies con otras estadísticas, el contraste dependerá de la rugosidad superficial, y esta depen-

dencia puede usarse como método para medir la rugosidad.

Finalmente, en el caso de los diagramas de speckle formados a través de una lente, se pueden utilizar los mismos argumentos dados hasta aquí con la única condición de reemplazar la distribución de amplitud luminosa sobre el difusor, por la amplitud del diagrama de speckle de Fresnel en el plano de la lente. De esta forma, la pupila de la lente actúa como una fuente luminosa secundaria con fase aleatoria que sintetiza en el plano imagen del difusor, el diagrama de speckle correspondiente. El tamaño promedio de los granos individuales del speckle imagen está también dado por la Ecuación (1), pero donde ahora z_0 indica la distancia entre la lente y el plano imagen de D , y a el diámetro de la pupila de la lente.

Medida de desplazamientos y deformaciones

Las técnicas para medir deformaciones de superficies que se basan en el fenómeno de speckle pueden clasificarse en dos categorías: la interferometría speckle, que permite medir directamente el desplazamiento local de la superficie, y la interferometría speckle de "shearing" por medio de la cual se determina la derivada del desplazamiento. En esta sección se analizarán algunos métodos correspondientes a la primera de las técnicas mencionadas.

1) Interferometría speckle por un solo haz de iluminación.

El esquema básico de este método se ilustra en la Figura 2. Una superficie difusora D , cuyas deformaciones se desean determinar, al aplicársele una carga exterior T , es iluminada por un haz de láser. Antes de aplicar la carga, se registra el diagrama de speckle imagen de D en una placa fotográfica H . Al aplicar la carga, el campo de deformaciones que se origina desplaza la microestructura superficial del difusor según el vector $\vec{\zeta}(\xi, \eta)$. En estas condiciones, el diagrama de speckle imagen se trasladará: $\vec{d} = m\vec{\zeta}$, siendo m la magnificación de la lente L . Si se registra en la misma placa H este nuevo diagrama de speckle, la distribución de intensidad total almacenada puede considerarse como una distribución aleatoria de pares de granos de speckle, cuya separación y orientación en cada punto (x, y) dan cuenta de la magnitud y dirección de la deformación local en cada punto del difusor (ver Figura 2).

Una vez revelada, la placa H es iluminada por una onda plana como se muestra en la Figura 3.

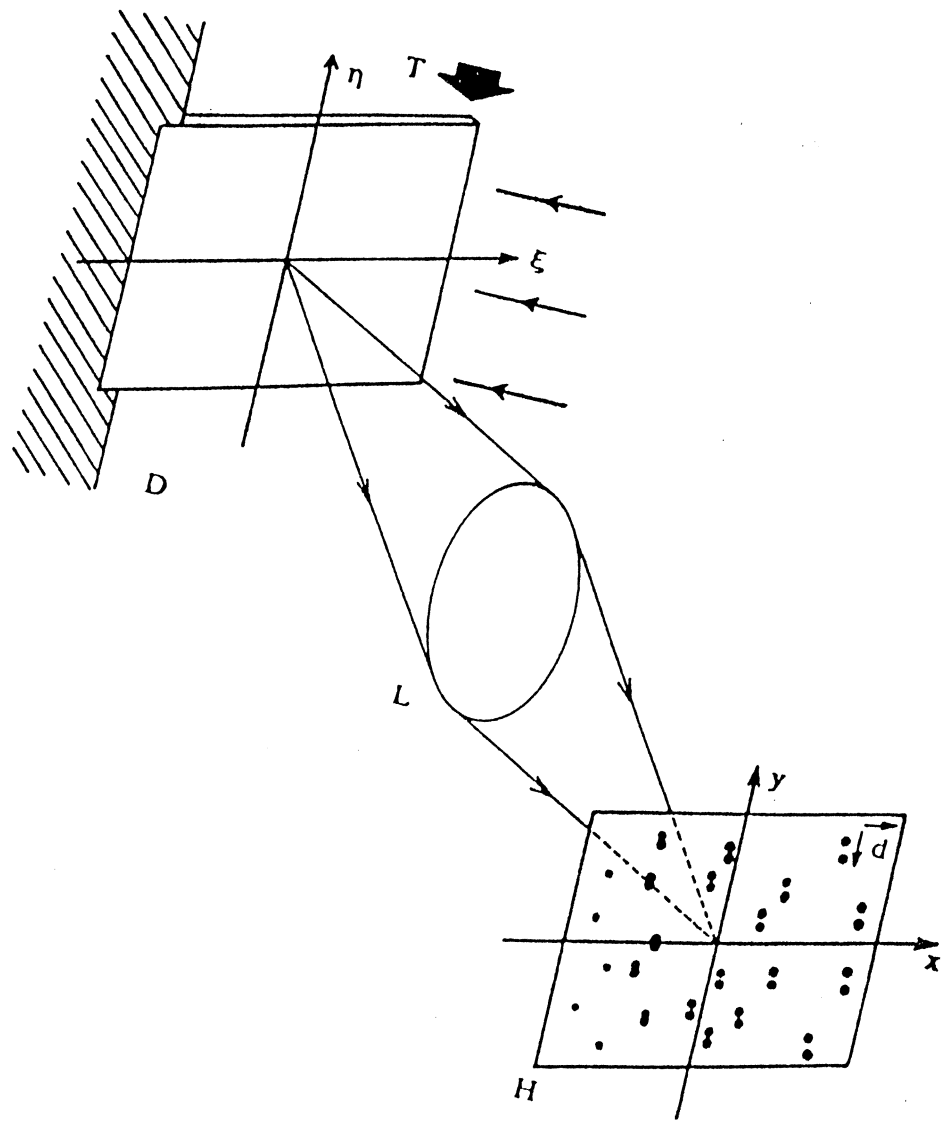


Figura 2

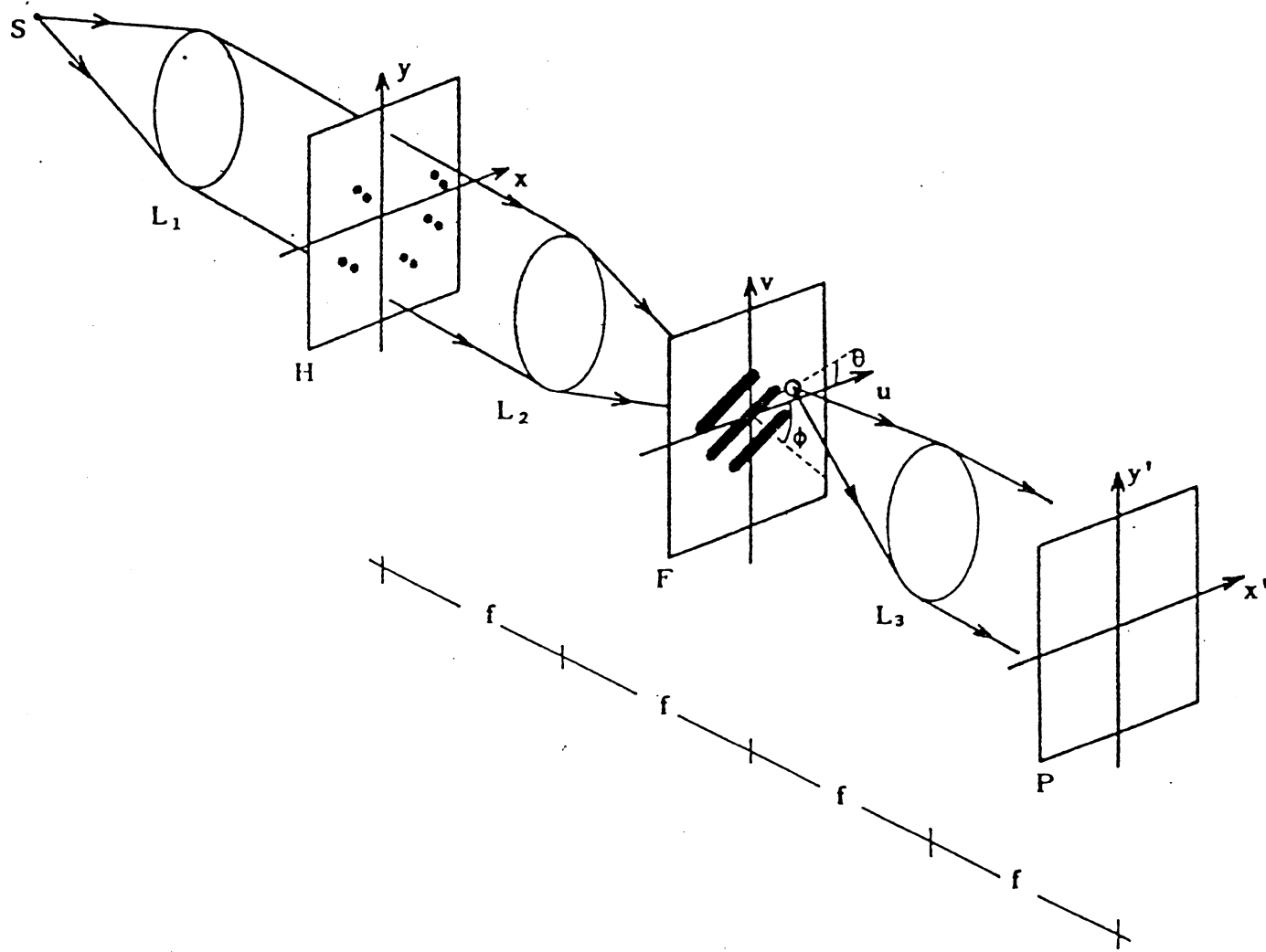


Figura 3

La transmitancia en amplitud de H corresponderá a una colección de múltiples dobles aberturas (debido a los pares de granos de speckle), que originarán en el plano focal posterior de L_2 franjas de Young de diferentes espaciados y orientaciones. En la Figura 3 se ilustra una de estas familias de franjas. La orientación queda especificada a través del ángulo ϕ que la dirección normal a las franjas forma con el eje horizontal u , mientras que el espaciado Δ está dado por la expresión:

$$\Delta = \lambda f/d = \lambda f/ms \quad . \quad (2)$$

En el plano de coordenadas (u,v) , se coloca un filtro espacial F consistente en una máscara opaca con una pequeña abertura circular. De esta manera, la luz que forma la imagen final de H en el plano (x',y') deberá provenir de las porciones de los máximos de intensidad de las franjas que transmita la abertura de F . Si ésta se encuentra sobre el eje u (es decir, se elige $\theta=0$), a una distancia u_0 del eje óptico, la condición que se debe satisfacer para que la abertura intercepte un máximo de una familia de franjas de espaciado Δ y orientación ϕ (ver Figura 4), resulta:

$$u_0 \cos \phi = n\Delta \quad ; \quad n = 1,2,3,\dots \quad (3)$$

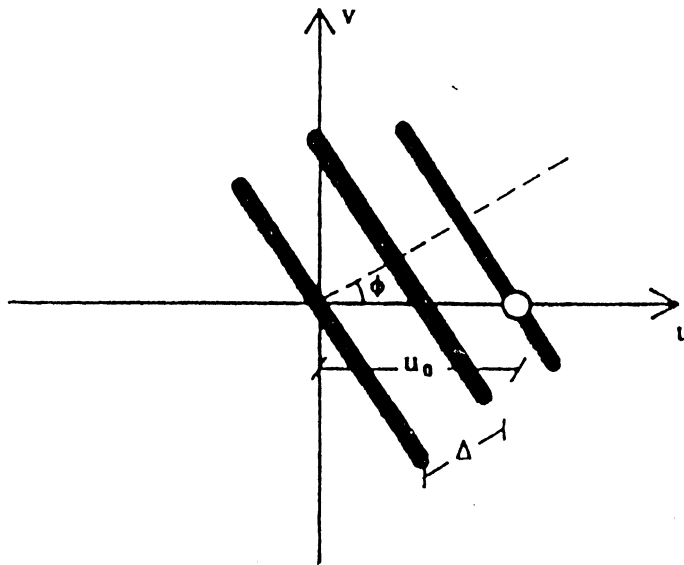


Figura 4

Teniendo en cuenta que $s_x = s \cos \phi$, la Ec.(3) se puede reescribir como:

$$s_x = n \lambda f / m u_0 \quad (4)$$

Análogamente, si la abertura de F se encuentra sobre el eje vertical v ($\phi=90^\circ$), a una distancia v_0 del eje óptico, la condición equivalente a la Ec.(4) será:

$$s_y = n' \lambda f / m v_0 \quad ; \quad n' = 1, 2, 3, \dots \quad (5)$$

Por consiguiente, la distribución de intensidad en el plano (x', y') de salida del sistema óptico corresponderá a una imagen de H , y por ende del difusor D , modulada por franjas brillantes y oscuras cuya geometría dependerá de la posición del filtro espacial F . Estas franjas representan las regiones de la superficie del difusor con igual valor de la componente de deformación según el eje definido por la posición de la abertura. Así, si se desean determinar las deformaciones de D según un cierto eje q , que forma un ángulo θ con el eje horizontal, se coloca la abertura de F según dicho ángulo θ y a una distancia q_0 del eje óptico. La franja de orden n observada, contada a partir del borde sostén no-deformado de D , representa las zonas de la superficie con un valor de deformación local:

$$s_q = n \lambda f / m q_0$$

Por lo tanto, entre dos franjas adyacentes cualesquiera de la imagen, existirá un incremento del valor de la deformación:

$$\Delta s_q = \lambda f / m q_0$$

El valor específico de q_0 [o de u_0 y v_0 en las Ecs.(4) y (5)] se elige dependiendo de la sensibilidad requerida para medir deformaciones. Un valor grande de q_0 implica que Δs_q sea pequeño, aumentando la sensibilidad del método. Para medidas cuantitativas de las deformaciones, es necesario conocer con mucha precisión los valores de m , f y q_0 . En estos casos, se utilizan procedimientos de calibración que permiten

independizar las mediciones de los valores de dichos parámetros.

Un procedimiento alternativo para medir directamente las deformaciones locales en el plano (ξ, η) del difusor, consiste en iluminar la placa H mediante un haz de láser sin expandir. La zona iluminada de H producirá por difracción las franjas de Young correspondientes al vector separación \vec{d} entre los granos de speckle de esa región de la imagen de D. De esta forma, midiendo el espaciado Δ de las franjas y el ángulo ϕ que ellas forman con el eje x, cuando la placa H se ilumina en el punto de coordenadas (x_0, y_0) , resulta:

$$|\hat{s}(x_0/m, y_0/m)| = \lambda z / m \Delta(x_0, y_0) \quad (6a)$$

$$\hat{s}(x_0/m, y_0/m) \cdot \hat{x} = \cos[\phi(x_0, y_0)] \quad ; \quad (6b)$$

donde z es la distancia entre H y el plano de observación de las franjas de Young. A través de las Ecs.(6), y midiendo en forma continua Δ y ϕ en todos los puntos (x_0, y_0) cuando éstos son secuencialmente iluminados por el haz de láser, se pueden determinar la magnitud y dirección de la deformación local \vec{s} para todos los puntos del difusor. Este método de decodificación "punto-por-punto" tiene la ventaja respecto al método anterior de proporcionar tanto la magnitud como la dirección del desplazamiento y de no requerir la identificación de órdenes de franjas.

2) Interferometría speckle por dos haces de iluminación.

El objeto difusor en estudio es iluminado por dos haces coherentes, y simétricos respecto al eje óptico, como se muestra en la Figura 5.

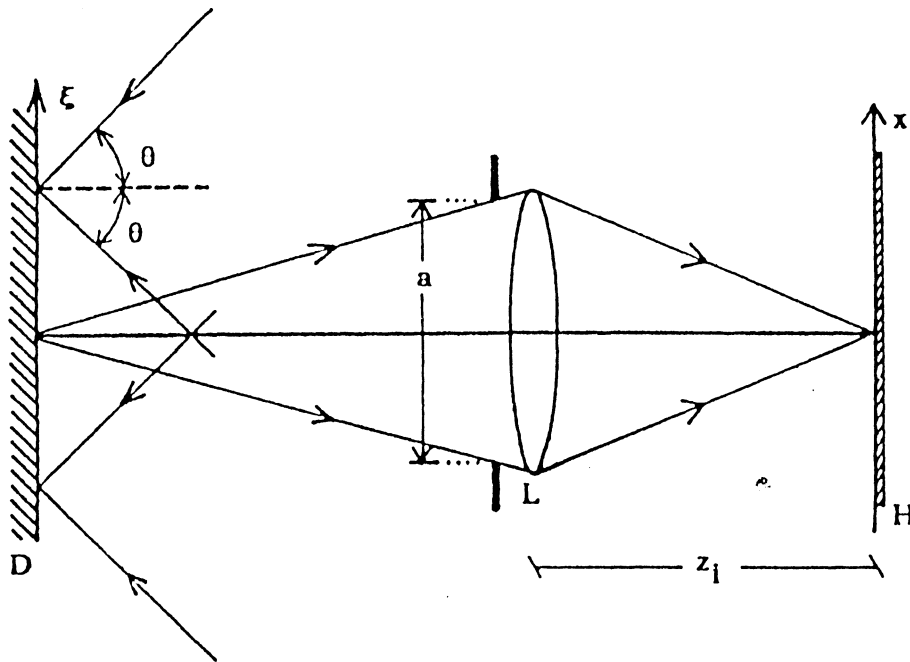


Figura 5

El dispositivo que se analizará permite medir desplazamientos de la superficie D en su propio plano. Sin embargo, modificando la geometría del sistema, se pueden determinar también desplazamientos de la superficie fuera de su plano (por ejemplo, según la dirección del eje óptico).

La imagen del difusor D (sin deformaciones) se registra en la placa fotográfica H , y luego de revelada, se la ubica exactamente en su posición original. El diagrama de speckle imagen registrado en H puede considerarse resultante de la interferencia de dos diagramas de speckle idénticos, cada uno originado por un haz de iluminación. Los granos de speckle registrados en H representan las regiones de transmitancia cero (ya que son puntos oscuros en la placa revelada). Por consiguiente, al estar la placa reposicionada, coincidirán espacialmente estas regiones con los máximos de intensidad (granos de speckle) del diagrama suma incidente sobre H , y no habrá luz transmitida por la placa fotográfica.

Cuando el objeto se deforma, la microestructura superficial de D se desplaza en su propio plano según el vector $\vec{s}(\xi, \eta)$. Para analizar el

efecto de este cambio, se considerará un punto P del difusor al cual se le asocia un grano de speckle como imagen en el plano de H (ver Figura 6).

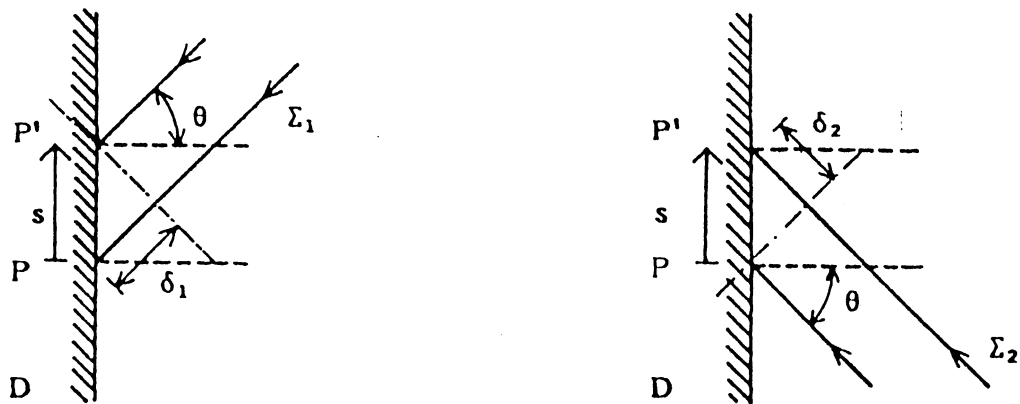


Figura 6

Es decir, los rayos luminosos originados en P por cada uno de los haces Σ_1 y Σ_2 de iluminación llegarán al punto imagen P' con una diferencia de fase igual a cero y por lo tanto interfieren constructivamente produciendo un máximo de intensidad. Al deformarse la superficie, el punto P se desplaza una cantidad s hasta ocupar la posición P' . En la medida que se verifique: $ms < \lambda z_1/a$, o sea, que la magnitud del desplazamiento del punto imagen de P' respecto a P , sea menor que el tamaño del grano de speckle, puede considerarse que el grano de speckle imagen resultante no cambia de posición pero sí la fase relativa de los rayos que llegan al mismo provenientes de P' debidos a los haces de iluminación Σ_1 y Σ_2 . Como puede derivarse de la Figura 6, el cambio de P a P' implica para los rayos originados en P' debidos a Σ_1 un cambio de fase: $\phi_1 = -(2\pi/\lambda)s \text{ sen } \theta$, mientras que para los rayos debidos a Σ_2 el cambio de fase será: $\phi_2 = (2\pi/\lambda)s \text{ sen } \theta$. Por lo tanto, el cambio de fase relativo será:

$$\Delta\phi(\xi, \eta) = (4\pi/\lambda) \text{ sen } \theta \cdot s(\xi, \eta) \quad ; \quad (7)$$

siendo $s(\xi, \eta)$ la componente de la deformación \vec{s} según el eje ξ en

cada punto del difusor.

De esta forma, en las regiones de la imagen del difusor donde $\Delta\phi = 2\pi n'$ (condición de interferencia constructiva), el diagrama de speckle suma resultante es idéntico al ya registrado en H correspondiente a la superficie sin deformaciones. Por consiguiente, no habrá luz transmitida por H, apareciendo estas regiones como franjas oscuras. Para las regiones que no satisfagan la condición anterior, los diagramas de speckle estarán descorrelacionados, apareciendo brillantes por transmisión. Del mismo modo que el método por un solo haz de iluminación, se obtiene una imagen de la superficie difusora modulada por franjas brillantes y oscuras, donde ahora el orden de franja n se relaciona con la deformación local s a través de la expresión:

$$s = n \lambda/2 \text{ sen } \theta \quad . \quad (8)$$

De la Ec.(8) se deduce que el incremento de la deformación superficial entre dos franjas adyacentes es $\lambda/2 \text{ sen } \theta$, sirviendo así este método para medir deformaciones muy pequeñas. Si bien la sensibilidad puede ser reducida disminuyendo el valor del ángulo θ , el máximo desplazamiento que se puede medir estará limitado por el tamaño de los granos de speckle.

3) Discusión comparativa de los métodos utilizados.

El método por un solo haz de iluminación tiene la ventaja con respecto a la holografía interferométrica y a la interferometría speckle por dos haces de iluminación, que no requiere de aislación especial contra vibraciones mecánicas. Además, la interpretación de las franjas observadas no depende (como sí sucede en las otras dos técnicas) de la dirección del haz de iluminación, permitiendo así emplear una onda divergente, lo cual elimina la restricción del tamaño para el objeto. Los valores de los desplazamientos superficiales que se pueden medir por esta técnica son mayores que por las otras dos, normalmente en el rango $1 \mu\text{m} - 50 \mu\text{m}$.

Una comparación entre la holografía interferométrica y la interferometría speckle por dos haces es apropiada, ya que ambas técnicas determinan deformaciones muy pequeñas (en el rango $0,1 \mu\text{m} - 10 \mu\text{m}$). La ventaja principal de la interferometría speckle con respecto a la holografía interferométrica consiste en que las franjas están siempre localizadas sobre la superficie del objeto. Además, la interferometría speckle requiere en general de placas fotográficas de menor resolución espacial que las requeridas en las técnicas holográficas. Sin embargo, la calidad de las franjas obtenidas empleando holografía interferométrica es superior a la de los métodos de interferometría speckle.

Análisis de vibraciones

1) Método visual utilizando un haz de referencia.

El interferómetro speckle que se emplea en este caso se muestra en la Figura 7.

Un observador O forma la imagen del objeto difusor D , cuyas vibraciones se desean analizar, a través de un sistema óptico telescópico. Este sistema posee un diafragma variable F por medio del cual se controla el tamaño del speckle imagen. Por intermedio de un semiespejo SE , se introduce un haz de referencia (derivado del mismo láser que el usado para iluminar la superficie de D), el que se superpone coherentemente con el diagrama de speckle originado por el difusor. Enfrente del telescopio, se coloca un polarizador P orientado de forma tal que transmita sólo la luz polarizada en el mismo plano que el correspondiente al haz de referencia.

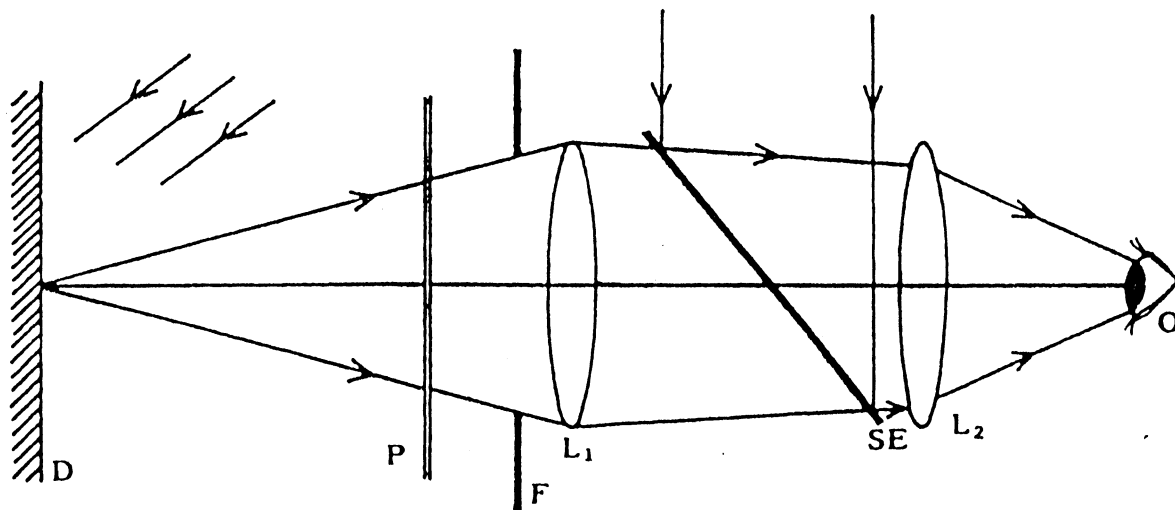


Figura 7

La figura de interferencia formada entre el diagrama de speckle y la onda de referencia de amplitud uniforme tendrá una distribución de intensidad que dependerá de la coherencia mutua entre ambos frentes de ondas. Para las regiones asociadas con las partes estacionarias (no-vibrantes) de la superficie D , la coherencia mutua será alta, originando como resultante de la figura de interferencia un diagrama de speckle imagen con muy buen contraste. Para las regiones vibrantes, la coherencia mutua

se reduce debido a que el ojo (detector) integra las variaciones de intensidad sobre el ciclo de la vibración, produciendo de esta forma una figura de interferencia con variaciones de intensidad más suaves (speckle de bajo contraste). La sensibilidad visual para distinguir vibraciones de la superficie paralelas a la línea de visión es equivalente a la obtenida empleando técnicas holográficas, y puede ser ajustada a través del parámetro α , cociente entre las intensidades de los haces de referencia y objeto. La principal ventaja respecto a la holografía consiste en poder efectuar el análisis a tiempo real (sin etapas de revelado fotográfico). Sin embargo, si se requiere un análisis cuantitativo del estado vibratorio de la superficie, es conveniente el empleo adicional de técnicas holográficas.

El efecto de "borroneado" (pérdida de contraste) que la vibración le produce al speckle imagen está relacionado con la amplitud de vibración. Realizando un análisis estadístico puede demostrarse que el contraste del speckle C satisface la ecuación:

$$C = [1 + 2\alpha J_0^2(4\pi\rho/\lambda)]^{1/2} / (1+\alpha), \quad (9)$$

donde ρ es la amplitud de vibración en cada punto del difusor y J_0 es la función de Bessel de orden cero y primera clase. La Ec.(9) predice que el contraste del speckle variará periódicamente a partir de un valor máximo en las regiones nodales de D , con un valor progresivamente decreciente hacia las regiones ventrales. En el caso de un interferómetro speckle optimizado a través de los diversos parámetros involucrados (valor de α , tamaño de la pupila F , etc.), es posible distinguir el primero, y a veces hasta el segundo, máximo secundario del contraste del diagrama de speckle.

2) Método visual sin haz de referencia.

Este método que detecta exclusivamente líneas nodales de la superficie, se esquematiza en la Figura 8. La superficie vibrante D es iluminada por un haz láser a través de un difusor. La diferencia entre las regiones vibrantes y en reposo de la superficie D es detectada a través de la variación del contraste, promediado temporalmente por el observador, del speckle imagen originado por la superficie. Para las zonas en reposo, este contraste vale aproximadamente uno, y así las líneas nodales aparecerán con alto contraste. Para las zonas vibrantes, el contraste del speckle disminuye gradualmente, desapareciendo eventualmente para grandes amplitudes de vibración. La sensibilidad del método puede ser ajustada utilizando diferentes tipos de difusores, o variando la distancia entre el difusor de iluminación y el objeto vibrante.

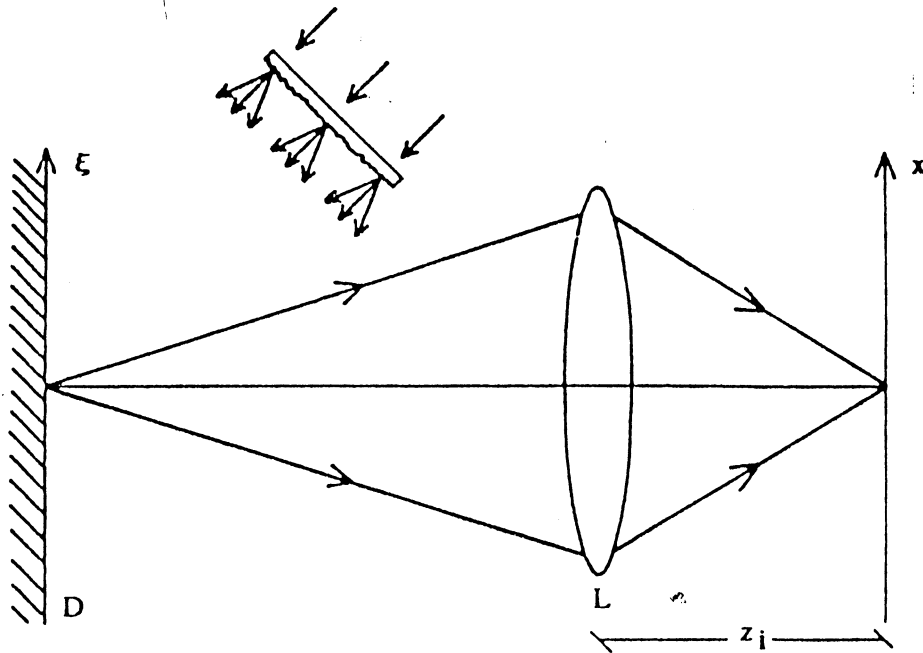


Figura 8.

Medida de la rugosidad superficial

La técnica standard para medir la rugosidad de una superficie se basa en utilizar un palpador o perfilómetro mecánico. Una punta de diamante, ligeramente en contacto con la superficie, se desliza sobre la misma, originando un voltaje de salida proporcional a la altura del perfil de la superficie en cada punto de contacto. La información suministrada es confiable siempre que el diámetro de la punta sea pequeño comparado con la longitud promedio de las irregularidades de la superficie. Sin embargo, a pesar de ser ampliamente aceptada, esta técnica adolece de dos desventajas: (i) puede dañar la superficie en examen; (ii) usualmente mide el promedio de las desviaciones del nivel superficial, independientemente de la longitud de correlación de la rugosidad. En el caso de aplicaciones industriales, se requiere de métodos de medida no-contactantes que provean un mayor nivel de automatización. En esta dirección, se han desarrollado diversos métodos ópticos que funcionan en diferentes rangos de valores de la rugosidad superficial. Para valores de la rugosidad media cuadrática R_a menores que $0,1 \mu\text{m}$ (superficies reflectoras), se emplean métodos basados en la interferometría con medición electrónica de la

fase. En el rango $0,1 \mu\text{m} < R_a < 3 \mu\text{m}$ (superficies difusoras de baja rugosidad), generalmente se utilizan métodos que analizan la forma de la figura de difusión o "scattering" de la superficie, cuando ésta se ilumina con un haz láser. Finalmente, para el rango $3 \mu\text{m} < R_a < 50 \mu\text{m}$ (superficies difusoras muy rugosas), los métodos más empleados analizan el contraste de la estructura fina de la figura de "scattering", que es precisamente el diagrama de speckle originado por la superficie. En esta sección se discutirán dos de estos métodos: el primero, que involucra la medición directa del contraste del diagrama de speckle imagen de la superficie; y el segundo, que se basa en detectar la correlación de dos diagramas de speckle.

- 1) Medida de la rugosidad media cuadrática a través del contraste del diagrama de speckle imagen.

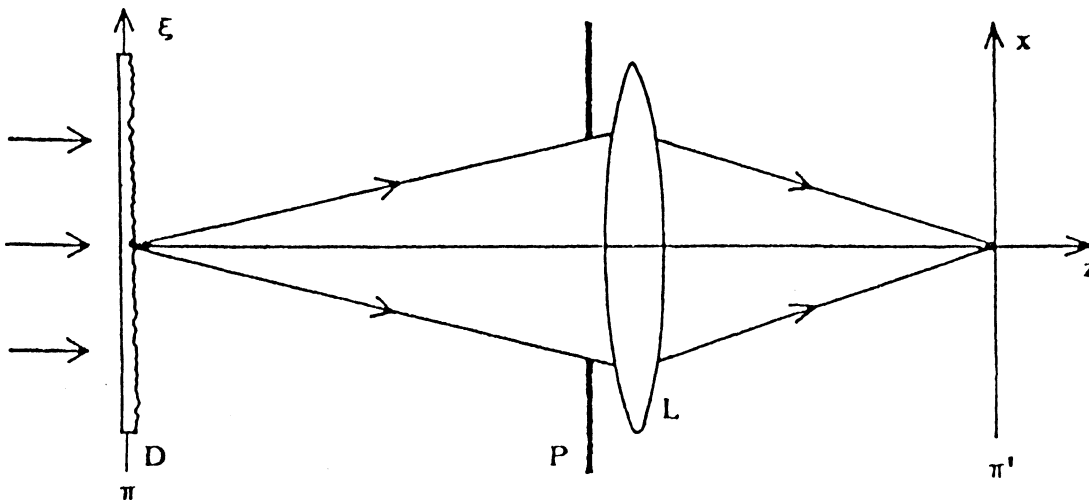


Figura 9

En el sistema óptico ilustrado en la Figura 9, un haz de láser ilumina un difusor plano D cuya rugosidad se desea analizar. Una lente L, diafragmada por una pupila P, forma la imagen del difusor en un cierto plano π'. Considerando que el proceso de formación de imágenes es un fenómeno lineal, la distribución de amplitud luminosa en el plano imagen $U_i(x,y)$ estará relacionada con la amplitud $U_0(\xi,\eta)$ en el plano del difusor a través de una integral de superposición:

$$U_i(x,y) = \iint_{-\infty}^{\infty} K(x,y;\xi,\eta) U_0(\xi,\eta) d\xi d\eta \quad (10)$$

siendo $K(x,y;\xi,\eta)$ la función de punto extendido del sistema óptico; es decir, es la amplitud que se obtiene en el plano imagen (x,y) al colocar una fuente luminosa puntual en las coordenadas (ξ,η) . Esta función, característica de los sistemas lineales, depende exclusivamente del sistema óptico. Si el difusor posee una transmitancia de amplitud uniforme, la alteración que introduce al frente de onda plano proveniente del láser se reducirá a un término de fase aleatorio, debido a las variaciones irregulares de camino óptico $\Delta(\xi,\eta)$ producidas por la rugosidad superficial. De esta forma: $U_0(\xi,\eta) = e^{ik\Delta(\xi,\eta)}$, y la Ec.(10) se puede reescribir:

$$U_i(x,y) = \iint_{-\infty}^{\infty} K(x,y;\xi,\eta) e^{ik\Delta(\xi,\eta)} d\xi d\eta . \quad (11)$$

Por lo tanto, la intensidad imagen $I_i(x,y)$ será entonces:

$$I_i(x,y) = \left| \iint_{-\infty}^{\infty} K(x,y;\xi,\eta) e^{ik\Delta(\xi,\eta)} d\xi d\eta \right|^2 . \quad (12)$$

Esta distribución de intensidad, de naturaleza estadística, representa el diagrama de speckle imagen originado por el difusor D . Una interpretación física de la Ec.(12) surge de considerar que la función $K(x,y;\xi,\eta)$ es no-nula dentro de una cierta región A centrada en el punto imagen (x,y) . Recíprocamente, también se puede interpretar que la amplitud luminosa en el punto (x,y) resulta de superponer las ondas originadas dentro de un área A ubicada alrededor del punto (ξ,η) en el plano del difusor. Por lo tanto, si las variaciones de fase $\Delta(\xi,\eta)$ dentro de esta área aumentan, la correspondiente variación de intensidad en el plano imagen también aumenta, de acuerdo con la Ec.(12), originando entonces un mayor contraste del diagrama de speckle. Por otra parte, si las variaciones de $\Delta(\xi,\eta)$ disminuyen (por ejemplo, en el caso de difusores más pulidos), el contraste del speckle disminuye tendiendo a cero (distribución de intensidad uniforme) para superficies reflectantes para las cuales $\Delta(\xi,\eta) \approx 0$. En consecuencia, existe una relación directa entre la rugosidad superficial y el contraste del diagrama de speckle. En la Figura 10 se muestran las fotografías obtenidas al registrar los diagramas de speckle imagen producidos por dos vidrios despulidos de diferente rugosidad. En la Figura 11 se muestran los correspondientes registros densitométricos. Se observa que en el caso del difusor de menor rugosidad, la intensidad media de fondo es mayor, disminuyendo por lo tanto el contraste del speckle.

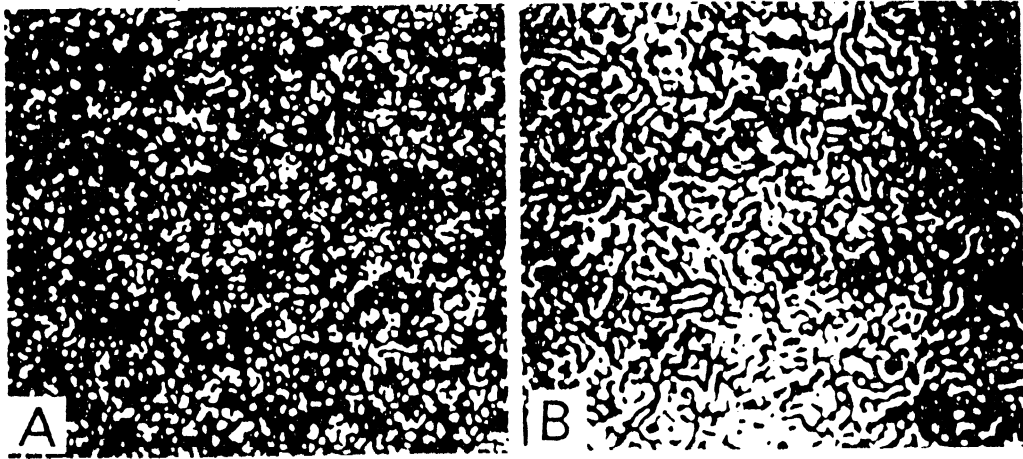
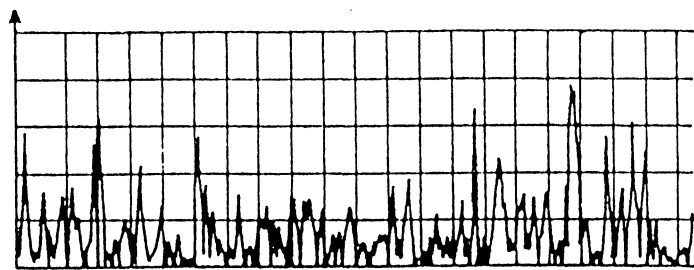
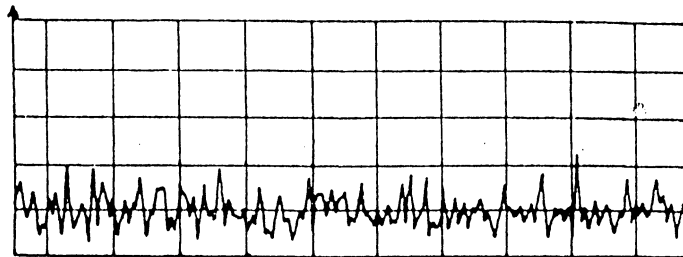


Figura 10



(a)



(b)

Figura 11

Para expresar cuantitativamente las variaciones de intensidad del speckle, se introduce el contraste promedio C , definido como:

$$C = \sqrt{\langle I_i^2(x,y) \rangle - \langle I_i(x,y) \rangle^2} / \langle I_i(x,y) \rangle \quad (13)$$

donde $\langle \dots \rangle$ indica un promedio estadístico. Si se introduce en la Ec. (13) la expresión de $I_i(x,y)$ dada por la Ec.(12), y teniendo en cuenta que $R_g = (\langle \Delta^2(\xi,\eta) \rangle)^{1/2}$, se obtiene una relación funcional: $C = C(R_g)$, a partir de la cual midiendo experimentalmente con un fotomultiplicador el contraste del speckle, se puede derivar el valor de la rugosidad media cuadrática de la superficie difusora.

2) Medida de la rugosidad superficial por correlación de diagramas de speckle.

Como ya se discutió, dos diagramas de speckle producidos por una misma superficie pero en condiciones diferentes (por ejemplo, efectuándole un pequeño desplazamiento a la superficie) pueden estar correlacionados. Esta propiedad se utiliza tanto en holografía interferométrica como en interferometría speckle para generar franjas de interferencia o franjas de correlación entre los dos diagramas de speckle. La geometría de estas franjas proporciona información respecto al campo de desplazamientos aplicado a la superficie, mientras que la visibilidad o contraste de franjas dependerá del grado de correlación existente entre los dos diagramas de speckle. A su vez, este grado de correlación es función (entre otros parámetros), de la rugosidad superficial.

La Figura 12 esquematiza un dispositivo experimental que utiliza esta propiedad para medir la rugosidad media cuadrática de la superficie D .

El método se basa en iluminar la superficie con dos haces láser que forman ángulos de incidencia θ_1 y $\theta_1 + \Delta\theta_1$. Si se analiza en cada caso el diagrama de speckle al infinito (o en el plano focal posterior de la lente L_3), se observarán las siguientes dos propiedades:

1 - El diagrama de speckle difundido en la dirección media θ_2 (originado por el haz incidente según θ_1) rota angularmente $\Delta\theta_2$ cuando se ilumina la superficie según la dirección $\theta_1 + \Delta\theta_1$, de tal forma que:

$$\Delta\theta_2 = (\cos \theta_1 / \cos \theta_2) \Delta\theta_1 \quad (14)$$

2 - La distribución estadística de la intensidad en el diagrama de speckle también cambia, de forma que el grado de correlación entre los dos diagramas disminuye cuando $\Delta\theta_1$ aumenta.

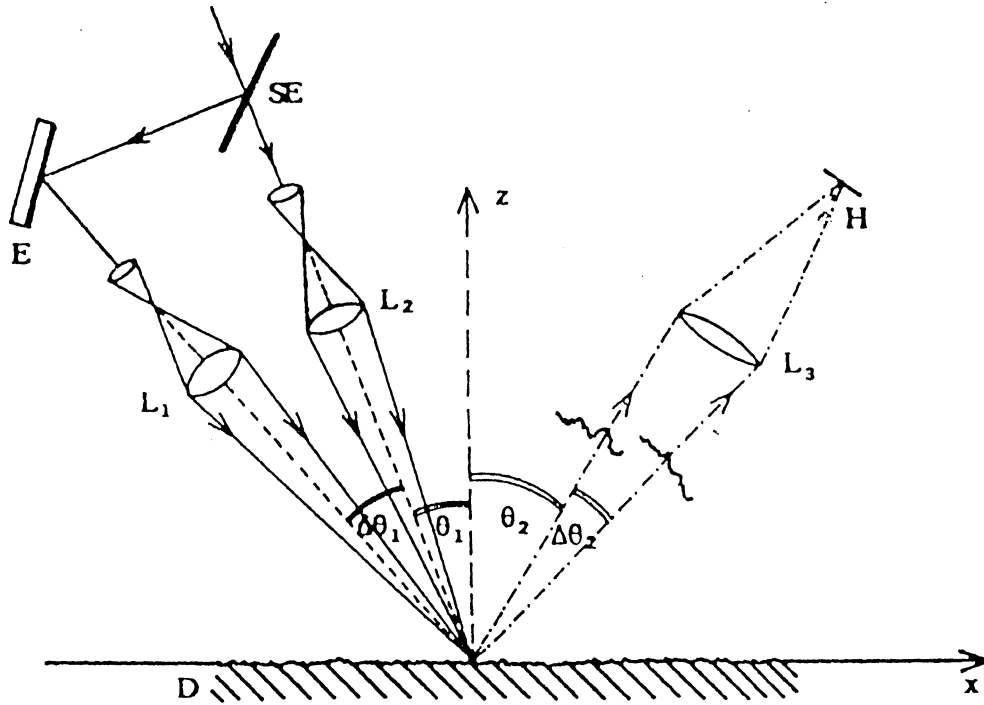


Figura 12

Si se registran los dos diagramas de speckle por doble exposición en una misma placa fotográfica H, y se la ilumina luego de revelada con un haz láser, aparecerán franjas de Young en el plano de Furler. La medida del contraste de estas franjas permite deducir el grado de correlación existente entre ambos diagramas de speckle. Por otra parte, un análisis teórico desarrollado en la referencia 4 (página 1210), demuestra que dicho contraste C está relacionado con la rugosidad superficial a través de la expresión:

$$C = e^{-(2\pi R_a \text{ sen } \theta_1 \Delta\theta_1 / \lambda)^2} ; \quad (15)$$

donde por simplicidad se eligió $\theta_2=0$ (es decir, dirección de observación normal a D).

La Figura 13 muestra la curva teórica, dada por la Ec.(15), para tres valores diferentes de la rugosidad media cuadrática R_a . Asimismo, se muestran los valores experimentales hallados midiendo directamente el

contraste de las franjas de Young, los que concuerdan muy bien con el resultado teórico. Por lo tanto, si se desea investigar la rugosidad de una cierta superficie, conociendo los valores de θ_1 , $\Delta\theta_1$ y λ , y midiendo experimentalmente el contraste C de las franjas originadas a partir del doble registro H , se puede utilizar la Ec.(15) para obtener el valor de R_a .

Una modificación del método descrito, que evita el registro de los diagramas de speckle, permite medir la rugosidad a tiempo real. Según esta variante, la superficie D es iluminada simultáneamente por los dos haces coherentes formando ángulos de incidencia θ_1 y $\theta_1 + \Delta\theta_1$. Los dos diagramas de speckle difundidos por la superficie tendrán un cierto grado de correlación, el que es puesto de manifiesto por medio de un interferómetro de dos haces tipo Michelson (ubicado entre la superficie y la lente L_3) que origina una figura de interferencia entre ambos diagramas de speckle en el plano focal posterior de L_3 donde se coloca un fotodetector (en lugar de la placa fotográfica H). La distribución de intensidad procesada electrónicamente permite obtener directamente el contraste de las franjas, el que se relaciona con la rugosidad superficial por una expresión similar a la Ec.(15).

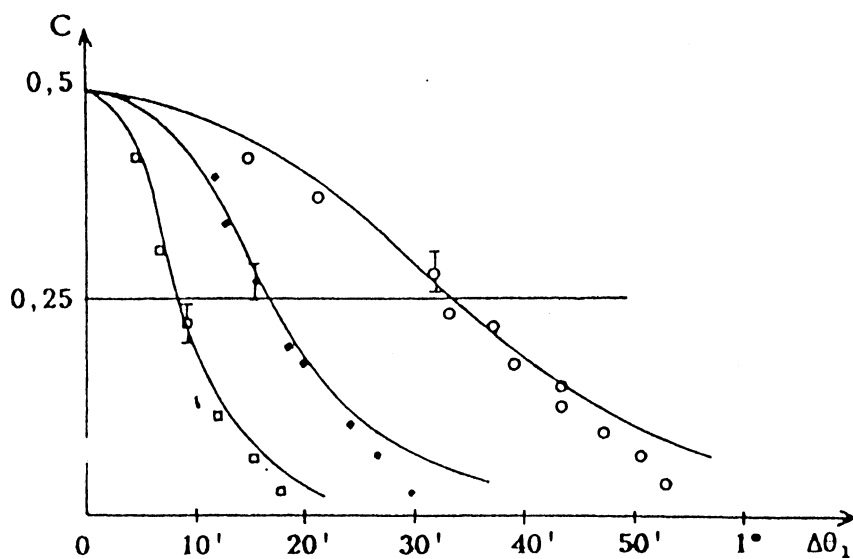


Figura 13

Bibliografia

1. "La granularité laser (speckle) et ses applications en optique", M. Francon, Masson, 1975.
2. "Speckle metrology", R.K. Erf ed., Academic Press, 1978.
3. "Laser speckle and related phenomena", J.C. Dainty ed., Springer, 1975.
4. "Speckle in optics", Journal of the Optical Society of America, vol.66, n°11, 1976.
5. "Optical metrology", Optical Engineering, vol.18, n°5, 1979.

INTERFEROMETRIA SPECKLE - INTERFEROMETRIA SPECKLE DIGITAL

1 - INTRODUCCION

La Metrología Speckle puede definirse como la rama de la óptica en la que el fenómeno de speckle es utilizado para determinar y cuantificar desplazamientos, deformaciones y velocidades de objetos difusores, variaciones o gradientes de índice de refracción de un medio, etc. Ella comprende una cantidad relativamente extensa de técnicas, las que generalmente resultan incluidas en una de las dos siguientes categorías: a) **Fotografía Speckle** y b) **Interferometría Speckle**.

La Fotografía Speckle está fundamentada en la experiencia de Burch y Tokarsky^[1]. A continuación serán analizados los conceptos básicos y las principales aplicaciones de la Interferometría Speckle. Se discutirá, además, la técnica de **Corrimiento de Fase (Phase stepping)**, la cual, conjuntamente con los métodos digitales, constituye una poderosa herramienta en metrología óptica.

En primer lugar, realizar una distinción cortante de las dos categorías mencionadas (Fotografía e Interferometría Speckle) no resulta trivial, ya que ambas incluyen técnicas fotográficas (como medio de registro) y ambas involucran al fenómeno de interferencia, presente en el speckle, que es en sí mismo el resultado de interferencias múltiples. Podría pensarse entonces, a primera vista, que no existen diferencias entre ambos métodos. En la práctica, sí hay tales diferencias, una prolija caracterización de las cuales fue propuesta por Stetson^[2] en 1975. Stetson sugirió que si, en una doble exposición, en ambas imágenes los diagramas de speckle están correlacionados, el proceso se denomina **Fotografía Speckle**. Este es el caso de la experiencia de Burch y

Tokarsky, donde las franjas de interferencia surgen, al iluminar la placa doblemente expuesta desde uno de los lados, como franjas de Young producidas por pares congruentes de granos de speckle. Si, por otro lado, las franjas se forman como resultado de la fluctuación en la correlación de ambos diagramas de speckle, haya o no traslación entre las porciones correlacionadas de ellos, el proceso se denomina Interferometría Speckle.

Existen además, por supuesto, diferencias en los arreglos experimentales usuales para cada una de estas técnicas y otras formas de distinguir ambos procesos. De hecho puede considerarse a la Interferometría Speckle como Holografía Interferométrica en el plano imagen, debido a la gran similitud que presentan ambos arreglos experimentales.

2 - INTERFEROMETRIA SPECKLE

El punto de partida de la Interferometría Speckle puede considerarse representado en el trabajo de Leendertz^[3], quien utilizó un arreglo interferométrico similar al de un interferómetro de Michelson, pero donde los espejos fueron reemplazados por superficies difusoras, S_1 y S_2 , como se muestra en la Figura 1. Un sistema óptico formador de imágenes (lente en la Figura) es requerido de modo tal que las dos superficies se superponen en el plano imagen, posibilitándose, de este modo, un registro fotográfico simultáneo de las mismas. Una de las superficies dará origen a un diagrama de speckle $D_1(x,y)$ y la otra a otro diagrama diferente $D_2(x,y)$. La adición coherente de estos dos diagramas de speckle aleatorios dará, en general, como resultado un nuevo diagrama de speckle $D_3(x,y)$, con estadística similar a los anteriores pero distinto en los detalles.

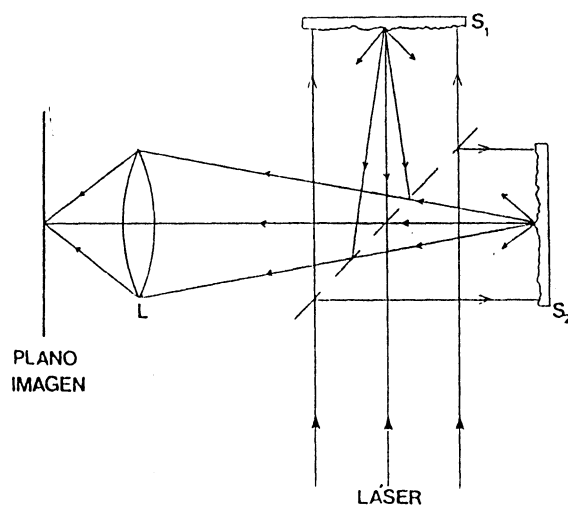


FIGURA 1
La experiencia de Leendertz
Interferómetro Speckle combinando dos diagramas de speckle

Considérese a la superficie S_2 de la Figura 1 como "fija" o "de referencia". Si en estas condiciones la superficie S_1 es desplazada en la dirección de su normal, la fase de la luz dispersada por ella cambiará en todas partes una cantidad δ . El nuevo diagrama de speckle resultante, $D_3(x,y,\delta)$, diferirá del anterior, $D_3(x,y,0)$ de acuerdo al valor que tome δ . Esto significa que ha ocurrido una reducción en el grado de correlación de los diagramas de speckle, obtenidos antes y después de desplazar a S_1 . Si $\delta = (2n+1)\pi$ no habrá correlación entre ambos diagramas, pero si $\delta = 2n\pi$, (con n entero), entonces $D_3(x,y,\delta) = D_3(x,y,0)$ y los diagramas estarán nuevamente correlacionados. Conceptualmente, es posible pensar que, en este último caso, cada grano individual de speckle ha completado un número entero de ciclos *brillante - oscuro* para terminar en un estado idéntico al de partida. Si es posible detectar estas "posiciones de correlación", el interferómetro puede utilizarse para medir el cambio de fase o, equivalentemente, el desplazamiento sufrido por la

superficie.

En la práctica puede procederse, por ejemplo, obteniendo un registro fotográfico del diagrama $D_3(x,y,0)$ antes de provocar el desplazamiento de la superficie S_1 y, luego de revelar la placa, reposicionarla donde se encontraba al momento de obtener la fotografía. Como se trata de un negativo, las zonas oscuras de la placa fotográfica impedirán el paso de luz proveniente de las zonas claras que les corresponden y viceversa. Como resultado se obtiene una muy baja transmisión de luz por parte de la placa. La misma condición de baja transmisión ocurrirá siempre que la fase δ haya variado un número entero de veces 2π radianes, ya que en este caso los diagramas $D_3(x,y,0)$ y $D_3(x,y,\delta)$ serán idénticos. Por otro lado, para variaciones de fase distintas de $2n\pi$, el diagrama de speckle $D_3(x,y,\delta)$ diferirá del original registrado en la placa fotográfica y, como consecuencia de ello, habrá luz transmitida por la placa, ya que el nuevo campo de speckle y la máscara negativa no son coincidentes entre sí.

Más aún, si δ varía sobre la superficie S_1 , por ejemplo porque ésta ha sido deformada, la máscara transmitirá luz en aquellas zonas donde el nuevo diagrama de speckle no está correlacionado con el que se encuentra registrado en ella, pero no transmitirá luz en las zonas donde exista correlación. Esto hace que la superficie aparezca cubierta por un diagrama de franjas que recorren zonas con δ constante. Estas franjas se denominan, por lo tanto, **Franjas de Correlación de Speckle**.

Sin embargo, debido a las dificultades que presenta el reposicionado del negativo, el método descrito (en el cual las franjas se obtienen "en vivo"), resulta un tanto engorroso. La obtención de las franjas de correlación de speckle puede mejorarse aplicando la técnica de "restado fotográfico".

Leendertz, en su experiencia, procedió de éste modo, registrando los dos estados $D_3(x,y,0)$ y $D_3(x,y,\delta)$ en placas separadas. Estos correspondían a una membrana metálica difusora, sujeta por los bordes, en reposo y con un gradiente de presión aplicado sobre ella respectivamente. A continuación realizó una copia positiva de una cualquiera de las placas y colocó cuidadosamente una sobre la otra. Finalmente, iluminando al conjunto de placas por transmisión obtuvo un diagrama de franjas de correlación de speckle consistente en anillos concéntricos. El espaciado de estos anillos estaba en dependencia con la presión ejercida y, por lo tanto, con la deformación sufrida por la membrana. Sus resultados fueron corroborados por la (en esa época) más asentada técnica de interferometría holográfica.

El trabajo de Leendertz demostró, por lo tanto, la posibilidad de extraer información metrológica por restado óptico de dos diagramas de speckle.

Es importante destacar el paralelismo existente entre esta técnica de correlación de diagramas de speckle y la Holografía Interferométrica en el plano imagen^[4]. La diferencia fundamental a favor de la primera de estas dos consiste en que no es necesario que el medio de registro pueda resolver muy altas frecuencias ($\cong 2000$ líneas/mm) como se requiere para la holografía. El precio a pagar por ello es que las franjas de correlación aparecen inmersas en el ruido generado precisamente por el speckle; de este modo las franjas más finas se "pierden" en el ruido de fondo, limitando así la resolución del método a una menor que la alcanzable holográficamente.

3 - APLICACIONES DE LA INTERFEROMETRIA SPECKLE

La Interferometría Speckle ha encontrado, desde su iniciación en el tra-

bajo de Leendertz; numerosas aplicaciones en metrología óptica. Las más destacadas de ellas, sea por razones históricas o por su importancia práctica se describen a continuación.

3.1 - Detección de vibraciones

La más sencilla de las aplicaciones de la interferometría speckle, que fue también la primera, consiste en la detección y observación, por inspección visual directa, de las líneas nodales en la superficie de un objeto sometido a vibraciones.

Considérese un arreglo interferométrico que consista básicamente en un interferómetro de Michelson en el cual uno de los espejos se reemplaza por una superficie difusora, como por ejemplo una lámina metálica. Si se utiliza luz coherente para iluminar y se realiza la observación de los haces superpuestos, provenientes de la superficie difusora y del espejo del interferómetro, a través de una abertura suficientemente pequeña, se observará a la superficie cubierta por un diagrama de speckle. Si se provoca una pequeña rotación de la superficie o ésta es desplazada en dirección de la luz incidente, no existirá una diferencia evidente entre la nueva situación y la anterior. Pero si la superficie está moviéndose con suficiente lentitud, cada speckle cambiará su brillo cíclicamente de claro a oscuro, independientemente de sus vecinos, de la misma forma que los puntos del espejo en el interferómetro de Michelson equivalente cambian su brillo a medida que es recorrido por franjas de interferencia cuando se lo mueve. Si la superficie del objeto se desplaza periódicamente hacia adelante y hacia atrás, con un recorrido de al menos un cuarto de longitud de onda, en esas zonas, el diagrama de speckle aparecerá, (si el movimiento es suficientemente rápido) borroneado

por la múltiple superposición de granos que cambian alternativamente de claros a oscuros. Su textura visible se alterará como si los haces objeto y referencia (proveniente este último del espejo) se hubiesen combinado en forma incoherente; así, en estas zonas ya no habrá speckles oscuros.

Por otro lado, en las zonas donde la superficie se encuentra en reposo (líneas nodales de la vibración) se observará un diagrama de speckle de alto contraste, con puntos oscuros y brillantes.

Este tipo de interferómetro, que no utiliza un diagrama de speckle como referencia, sino un haz libre de ruido, proveniente de un espejo, fue introducido originalmente por Archbold y sus colaboradores^[5] en 1970 y luego fue mejorado por Stetson^[6] en el mismo año, quien diseñó la versión que se muestra en la Figura 2.

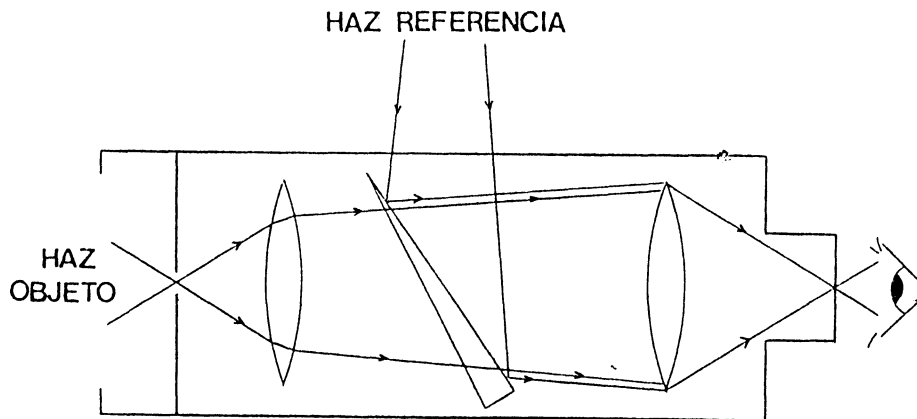


FIGURA 2
Esquema del Interferómetro Speckle utilizado por Stetson para la detección visual de vibraciones.

3.2 - Medida de desplazamientos en el plano

Si bien la Holografía Interferométrica^[4] es capaz de determinar con gran precisión desplazamientos y deformaciones superficiales en tres dimensiones, en la práctica no siempre resulta sencillo obtener separadamente las componentes de estos desplazamientos o deformaciones en el plano y en la dirección perpendicular a él.

Si se trata de medir esfuerzos superficiales, sólo es necesario medir el desplazamiento en el plano de la superficie y encontrar su variación con la posición en la dirección de interés. Esto puede realizarse utilizando una técnica de Interferometría Speckle, sin que los desplazamientos en dirección normal a la superficie afecten a las mediciones.

El montaje básico para llevar a cabo estas medidas se esquematiza en la Figura 3 y se basa en la propuesta original de Leendertz^[3] de utilizar como haz de referencia también un diagrama de speckle, como ya se vio en la sección 2.

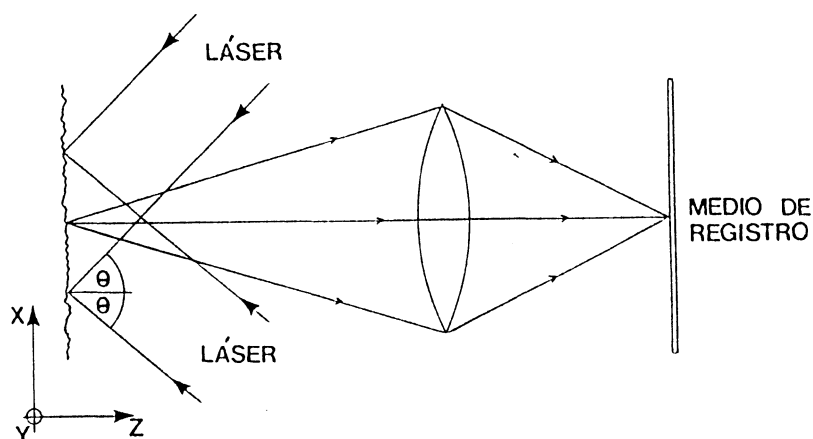


FIGURA 3
Esquema de un Interferómetro Speckle de doble iluminación para la medición de desplazamientos en el plano.

En este caso, la superficie que se desea investigar es iluminada con dos haces coherentes entre sí dispuestos simétricamente, formando ángulos θ iguales a ambos lados de la normal a la superficie. Cada haz genera su propio diagrama de speckle y la combinación coherente de ambos, que también es un diagrama de speckle, es registrada por la placa fotográfica H con ayuda de la lente L . Si la superficie se mueve en la dirección Z , normal a la superficie, los dos haces que la iluminan sufrirán idénticos cambios de fase y, por lo tanto, el diagrama de speckle resultante de la combinación de ellos permanecerá inalterado. Lo mismo ocurrirá en la dirección Y , perpendicular al plano determinado por los haces de iluminación.

Considérese ahora la dirección X , definida por la intersección del plano generado por los haces de iluminación con el plano de la superficie que se desea estudiar. Un desplazamiento en ésta dirección provocará una alteración en la fase relativa de ambos haces de iluminación y, por lo tanto, el diagrama de speckle resultante se alterará. En efecto, un desplazamiento d_x en la dirección X provocará un incremento del camino óptico de $d_x \sin\theta$ para uno de los haces y un decremento de la misma magnitud para el otro, siendo el cambio total del camino óptico $2 d_x \sin\theta$. El nuevo diagrama de speckle combinado estará correlacionado con el existente antes de producirse la traslación cuando se cumpla la condición:

$$2 d_x \sin\theta = n \lambda \quad , \quad (1)$$

donde n es un entero.

La distancia entre franjas adyacentes ($n = 1$) corresponde a un desplazamiento

$$d_x = \frac{\lambda}{2 \operatorname{sen}\theta} ,$$

de modo tal que la sensibilidad del sistema puede ajustarse variando el ángulo θ de incidencia de los haces de iluminación. Graficando las posiciones de las franjas como función de x , puede ser obtenido el esfuerzo lineal $\delta d_x / \delta x$.

Hay algunas características del experimento que acaba de describirse que resultan de gran importancia para su aplicación. En primer lugar, el poder resolutivo del sistema óptico (lente en la Figura 3) no resulta de mayor importancia para la operación del instrumento. El solo debe ser capaz de resolver los granos de speckle y no el conjunto de franjas de interferencia de alta frecuencia espacial que ambos haces de iluminación producen al encontrarse sobre la superficie. En segundo lugar, la existencia de franjas de correlación de speckle presupone que el desplazamiento d_x de la superficie no excede (en rigor es mucho menor) el tamaño de un grano de speckle proyectado sobre la superficie por el sistema óptico. El rango de desplazamientos medibles puede incrementarse entonces por medio de una desmagnificación de la superficie, teniendo en cuenta que ésto provoca una pérdida en la cantidad de luz que llega a la placa fotográfica.

En 1972, Duffy^[7] propuso una alternativa para medir desplazamientos en el plano de objetos rugosos utilizando, en lugar de dos haces de iluminación simétricamente dispuestos, un sólo haz y una lente con dos aberturas circulares simétricamente dispuestas en su pupila, como indica la Figura 4. De este modo, dos diagramas de speckle (uno correspondiente a cada abertura) se formarán en el plano imagen y se combinarán en forma coherente dando lugar a una

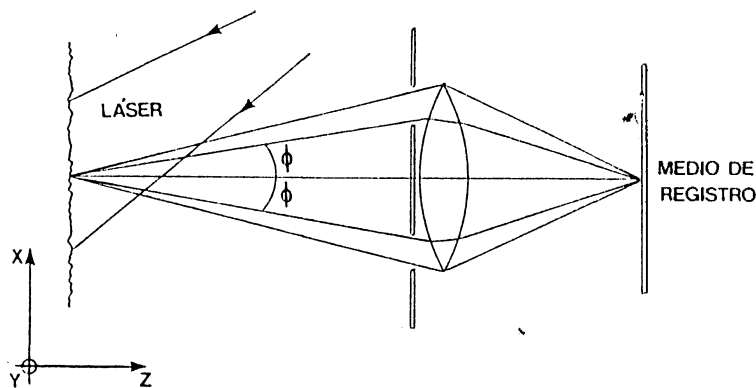


FIGURA 4
Esquema de un Interferómetro Speckle de doble apertura para la medición de desplazamientos en el plano.

modulación periódica en forma de red a cada grano. Para desplazamientos en la dirección X, paralela a la línea que une los centros de las aberturas, la fase resultante hará que ésta modulación periódica se desplace paralelamente a sí misma, avanzando un período cada vez que el desplazamiento d_x corresponda a $\lambda/2 \sin \phi$, siendo ϕ el ángulo subtendido por cada una de las aberturas de la pupila. En este caso ambos diagramas de speckle se encuentran correlacionados. Para desplazamientos de la mitad del valor que los anteriores, la modulación invertirá su contraste, cambiando las franjas brillantes por oscuras respecto al patrón original y viceversa. Los dos diagramas de speckle combinados no estarán, en este caso, correlacionados.

Posteriormente, Chiang^[8] extendió la versión de Duffy colocando en la pupila de la lente cuatro aberturas en forma de cruz, pudiendo obtener de este modo información simultánea sobre los desplazamientos en las dos direcciones ortogonales a la normal al plano de la superficie.

3.3 - Contorneado de superficies

En el ítem precedente se mostró la capacidad de la Interferometría Speckle para determinar desplazamientos de una superficie en el plano. La superficie misma no necesita ser plana o casi plana ya que, siempre que la longitud de coherencia del láser sea suficientemente grande, se formará un diagrama de speckle como resultado de la combinación coherente de la luz de cada haz de iluminación dispersada por ella. Por lo tanto, el método de la sección 3.2 es también aplicable a objetos no planos para determinar desplazamientos en las superficies.

Por otro lado, si se pretende obtener información acerca de la forma en profundidad del objeto, debe provocarse un cambio de fase que involucre a esta dimensión. En general este cambio de fase puede escribirse

$$\Delta\phi = \frac{2\pi}{\lambda} n \Delta Z \quad , \quad (2)$$

donde λ es la longitud de onda de iluminación, n es el índice de refracción del medio donde se trabaja y ΔZ es una variación de distancia en la dirección de la normal a la superficie.

Por lo tanto, para lograr el contorneado por curvas de nivel, el cambio de fase necesario puede lograrse provocando un desplazamiento ΔZ , alterando el índice de refracción o variando la longitud de onda de iluminación entre ambas exposiciones. Si el objeto es difícil de acceder de tal modo que Z no puede variarse y si resulta engorroso modificar el índice de refracción del medio donde se trabaja (como suele ser el caso), el contorneado se realiza cambiando la longitud de onda de iluminación como se describe a continuación.

Cambiando la longitud de onda utilizada para iluminar al objeto, el dia-

grama de speckle perderá, en general, correlación. Por otro lado, para dos longitudes de onda λ_1 y λ_2 la correlación volverá a hacerse notable a intervalos regulares tales que la diferencia de camino entre los haces objeto y referencia cumplan la condición

$$\Delta = m \lambda_1 = k \lambda_2 \quad (3)$$

donde m y k son enteros.

La diferencia de camino entre posiciones sucesivas donde la ecuación (3) se satisface está dada por

$$d\Delta = m \lambda_1 = (m+1) \lambda_2$$

ó,

$$d\Delta = \lambda_1 \lambda_2 / (\lambda_1 - \lambda_2) = \frac{\lambda_1 \lambda_2}{\Delta\lambda} \quad (4)$$

Si ahora, una superficie no plana es analizada en un interferómetro speckle por el método de doble exposición, cambiándose la longitud de onda de iluminación entre ambas exposiciones, se generarán franjas de correlación que contornean la superficie siguiendo zonas de igual "profundidad". El intervalo de contorneado, o distancia entre curvas de nivel se obtiene de la ecuación (4) y está dado por $\lambda_1 \lambda_2 / 2 \Delta\lambda$ ya que la luz debe viajar dos veces la distancia de variación de profundidad. Por ejemplo utilizando las dos líneas más intensas de un láser de Argón, que son de $\lambda_1 = 488\text{nm}$ y $\lambda_2 = 514.5\text{nm}$ pueden generarse líneas de nivel separadas $4.7\mu\text{m}$ en profundidad.

El contorneado de objetos tridimensionales utilizando dos longitudes de onda diferentes fue propuesto originalmente por Hildebrand y Haines^[9] utilizando técnicas holográficas y recientemente mejorado por Rastogui y

Pflug^[10]. La ventaja de utilizar un Interferómetro Speckle en lugar de un sistema holográfico consiste en que, en el primer caso, una lente forma una imagen del objeto y la posición de ésta no se altera al cambiar la longitud de onda. En Holografía, por el contrario, la necesidad de utilizar un haz de referencia fuera de eje hace que la imagen reconstruída se desplace lateralmente al cambiarse la longitud de onda, dificultándose la interpretación de los resultados.

4 - INTERFEROMETRIA SPECKLE ELECTRONICA

4.1 - Introducción

Para cualquier implementación de los sistemas de **Interferometría Speckle**, como los que fueron tratados en las dos últimas secciones, el registro de la distribución de intensidades de los diagramas de speckle resulta un paso fundamental en el proceso global.

Las técnicas fotográficas de registro poseen ciertas ventajas como alta sensibilidad, buena resolución y capacidad de retener la información registrada durante largos períodos sin que ésta se altere considerablemente. Por otro lado, la necesidad de utilizar procedimientos químicos para el revelado, constituye la principal desventaja de las técnicas fotográficas, convirtiéndolas en esencialmente lentas.

Algunos métodos alternativos, como la utilización de láminas termoplásticas han sido propuestos y utilizados en reemplazo de las técnicas fotográficas. Sin embargo la resolución alcanzable con estos materiales es notablemente inferior.

En el intento de simplificar los procedimientos para aplicaciones metro-lógicas surge a comienzos de los años 70 en la Universidad de Loughborough (Inglaterra)^[11,12] y en forma paralela en Stanford (USA)^[13] un cambio importante en la metrología speckle cuando se reconoce la posibilidad de implementar las técnicas básicas de la Holografía Interferométrica utilizando un sistema de video en reemplazo del procesamiento fotográfico y la factibilidad de procesar en forma electrónica las señales adquiridas ópticamente por dicho sistema de video. El principal objetivo de estos trabajos consistía en proveer sistemas rápidos (capaces de funcionar en régimen continuo o próximo a él) de análisis para la inspección de piezas "en línea", mientras éstas eran fa-

bricadas.

La técnica resultante, que incluía un haz de referencia fijo, proveyendo así la sensibilidad interferométrica de la Holografía, surge entonces de la fusión y evolución de la Interferometría Speckle y la Holografía Interferométrica y fue denominada Interferometría Speckle Electrónica (Electronic Speckle Pattern Interferometry = ESPI, en el idioma original).

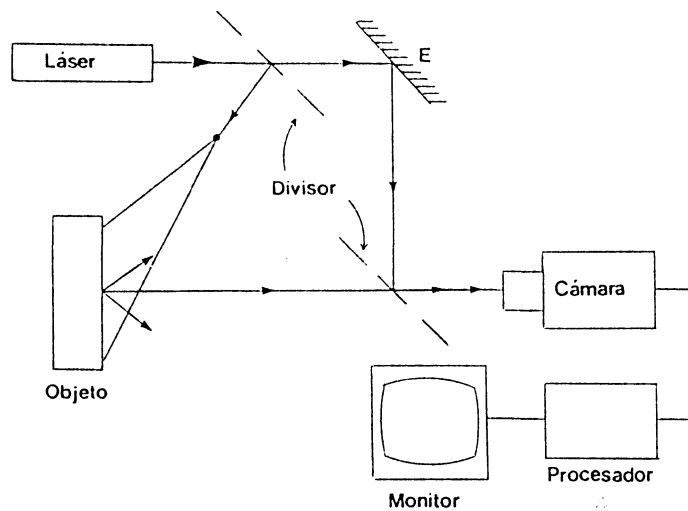


FIGURA 5
Representación esquemática de un Interferómetro Speckle Electrónico (ESPI)

El esquema básico de los sistemas ESPI originales se muestra en la Figura 5. El objeto a ser estudiado es iluminado por un haz láser convenientemente expandido y una imagen de dicho objeto es formada por el sistema óptico de una cámara de video sobre el elemento foto-sensible de ella. Un haz de referencia es introducido por medio de un divisor de haz de forma tal que su fuente efectiva se encuentre en el plano focal anterior del sistema óptico de

la cámara; de este modo el haz de referencia llega al detector de la cámara en forma de una onda plana. El sistema estándar de detección lo constituye una cámara de vidicon o, actualmente, una cámara de estado sólido de CCD; el poder de resolución típico de ambas es de aproximadamente 500 elementos por línea. Por esta razón la pupila del sistema óptico no puede elegirse arbitrariamente grande, sino que debe reducirse de modo tal que el tamaño de los granos de speckle crezca hasta que éstos puedan ser resueltos por la cámara. De otro modo la señal en cada elemento corresponderá a un promedio sobre más de un grano de speckle, originándose así una pérdida en el contraste del diagrama. Tanto en un sistema vidicon como uno de CCD la señal luminosa recibida en un área elemental (un pixel en el caso de la cámara de CCD) es integrada durante el tiempo que el sistema electrónico demora en realizar un barrido completo de todo el detector, siendo este tiempo típicamente de 1/30seg. Las imágenes pueden ser actualizadas entonces muy rápidamente (a razón de 30 por segundo) pero, simultáneamente el tiempo máximo de exposición queda fijado en aproximadamente 33ms, de modo tal que el brillo de la superficie que se está examinando debe ser alto, para que la relación señal/ruido a la salida del detector sea apreciable. Este problema de la relación señal/ruido en los sistemas ESPI ha sido profundamente estudiado, entre otros autores, por C. Wykes^[14] quien encontró un valor máximo para esta relación cuando la apertura de la lente es tal que el diagrama de speckle está justo en el límite de ser completamente resuelto y cuando el cociente de las intensidades correspondientes a los haces objeto y referencia es tal que este último no supera en intensidad al primero por más de un factor de 1,7. Asimismo C. Wykes demostró que debido a los limitados rangos dinámicos de las cámaras, para evitar pérdidas de información debidas a la saturación del detector por excesiva

iluminación, es necesario mantener valores de iluminación total tales que la intensidad media del haz objeto en el detector sea del orden del 15% del valor requerido para saturarlo. Este valor aparentemente bajo surge del hecho de que en un diagrama de speckle ocurren fuertes variaciones de intensidad, que pueden saturar localmente al detector de la cámara aún cuando, en promedio, la intensidad del haz objeto sea baja.

Finalmente, la señal proveniente de la cámara es amplificada y mostrada en un monitor. El paso importante es aquí que, antes de que la señal alcance el monitor, puede ser procesada electrónicamente para mostrar finalmente el resultado de alguna operación deseada. En los sistemas ESPI que reemplazan a los sistemas de Interferometría Speckle con registro fotográfico, luego de que una imagen del estado inicial del objeto a analizar ha sido almacenada, las imágenes registradas a continuación son restadas punto a punto de la imagen inicial de referencia antes de que el resultado aparezca en el monitor, siguiendo siempre la idea original de Leendertz^[3] descrita en la sección 2. Si existe algún cambio de fase en porciones o puntos equivalentes de la imagen de referencia respecto a la que es registrada en un determinado instante posterior, debido a que el objeto se corrió, se deformó o fue variada la dirección de iluminación del haz objeto, la resta de las dos imágenes dará, en estos puntos un resultado variable según sea la magnitud del cambio de fase. Así, donde la fase haya variado múltiplos enteros de 2π radianes, los diagramas de speckle de ambas imágenes coincidirán y la resta dará un resultado nulo que aparecerá oscuro en el monitor. Para aquellas zonas donde el cambio de fase no resulte múltiplo de 2π radianes, la intensidad correspondiente en el monitor variará según sea el caso. La imagen del objeto en el monitor aparecerá entonces modulada por franjas que recorren zonas para las cuales se ha

producido el mismo cambio de fase.

Actualmente, la denominación *Interferometría Speckle Electrónica* ha sido reemplazada por *Interferometría Speckle Digital*^[15] (Digital Speckle Pattern Interferometry = DSPI en el idioma original) que indica que el almacenamiento y procesamiento de las imágenes se realiza digitalmente en una computadora auxiliar dotada de un sistema dedicado de procesamiento de imágenes. En este caso, la aplicación de técnicas de filtrado no-lineal^[16] permite mejorar el contraste de las franjas en casos en los que, por ejemplo, el tamaño medio de los granos de speckle es menor que la resolución espacial de la cámara.

En lo que sigue se utilizará el término DSPI, a menos que razones históricas indiquen que la sigla ESPI es más adecuada.

4.2 - Comparación entre DSPI y Holografía Interferométrica

Antes del desarrollo de la Interferometría Speckle, la Holografía Interferométrica^[17] se mostró como una técnica muy adecuada para la medición de desplazamientos, deformaciones y modos de vibración de estructuras.

La Holografía Interferométrica se basa en la detección de diagramas de interferencia secundarios. Para obtener estos diagramas secundarios, un diagrama primario es registrado en un holograma del objeto a analizar. Posteriormente el objeto es desplazado, deformado o sometido a un esfuerzo que corresponde a las condiciones de trabajo reales y, a continuación, un segundo diagrama de interferencia primario (un nuevo holograma) es registrado sobre la misma placa. Cuando el holograma doblemente expuesto es reconstruido, se observa la interferencia entre los dos diagramas primarios, esto es los dos frentes de onda correspondientes a los dos estados del objeto, la cual resulta en una reconstrucción tridimensional del mismo cubierto con franjas.

En DSPI, por otro lado, el diagrama secundario es formado por procesamiento electrónico, libre de revelado y las franjas se observan en el monitor de video.

Cuando se utiliza Holografía Interferométrica, la adquisición de datos registra información excesiva que codifica los aspectos tridimensionales del objeto. El precio a pagar por ello es que el medio de registro debe ser capaz de resolver muy altas frecuencias espaciales, típicamente del orden de 2000 líneas/mm.

Como contraparte, en un sistema DSPI, sólo se obtiene información acerca de desplazamientos resueltos en alguna dirección particular, paralela o perpendicular a la dirección de observación. Los granos de speckle son los portadores de información que dan origen a la formación de franjas y deben, por lo tanto, ser resueltos por el sistema óptico. A causa de esto, la claridad de las franjas obtenidas en un sistema DSPI es considerablemente más pobre que la de las franjas secundarias en un sistema de Holografía Interferométrica, donde los granos de speckle son apenas visibles o, en ocasiones, absolutamente imposibles de reconocer.

Los sistemas de video usuales utilizados para DSPI y las memorias digitales asociadas almacenan, por lo general, imágenes de 512x512 elementos (pixels) de extensión. Esto implica que, independientemente de la geometría del sistema óptico, los speckles aparecerán de un tamaño de, por lo menos, 1/500 de la extensión de la imagen en una dirección.

Por último, tal vez la mayor ventaja de los sistemas DSPI, consiste en que la información en forma de franjas puede ser obtenida y actualizada en tiempos del orden de 1/30seg, dando la posibilidad de implementar sistemas de monitoreo que operan prácticamente a tiempo real, permitiendo su inserción en

cadena de control automatizado.

4.3 - Obtención de franjas de correlación en DSPI

Considérese nuevamente el esquema de la Figura 2.5, en donde el objeto a ser estudiado es iluminado por un haz (el haz objeto), coherente con un haz de referencia que es expandido y dirigido por medio de un divisor de haz hacia la cámara de video. De este modo, la intensidad $I(r)$ en el punto r del plano imagen esta representada por:

$$I_1(r) = U_o(r) U_o^*(r) + U_R(r) U_R^*(r) + U_o(r) U_R^*(r) + U_R(r) U_o^*(r) \quad (5)$$

en donde

$$U_o(r) = u_o(r) e^{i\phi_o(r)}$$

y

$$U_R(r) = u_R(r) e^{i\phi_R(r)}$$

representan las amplitudes complejas de los haces objeto y referencia respectivamente.

Si la superficie es ópticamente rugosa, como es el caso generalmente en DSPI, los valores de u_o y ϕ_o varían en forma aleatoria, dando lugar a la formación de un diagrama de speckle. Esto es, $I_1(r)$ varía sobre la imagen aleatoriamente, en forma de un diagrama de speckle.

Luego de que la imagen correspondiente a $I_1(r)$ ha sido almacenada en una memoria de imágenes, se introduce una variación de fase en el haz objeto, ya sea deformando al objeto mismo o modificando la iluminación. La amplitud com-

pleja que representa al haz objeto es, en estas condiciones:

$$U_o(r) = u_o(r) e^{i\phi_o(r) + i\psi(r)} = U_o(r) e^{i\psi(r)} \quad (6)$$

en donde $\psi(r)$ representa el cambio de fase introducido en r .

La nueva intensidad registrada por la cámara en el punto r esta dada ahora por

$$I_2(r) = U_o U_o^* + U_R U_R^* + U_o U_R^* e^{i\psi(r)} + U_o^* U_R e^{-i\psi(r)}, \quad (7)$$

donde, por simplicidad, ha sido eliminada la dependencia explícita de $U_o(r)$ y $U_R(r)$ con la posición r .

Restando, a continuación, la intensidad de la segunda imagen de la de la primera, esto es $I_2 - I_1$ se obtiene:

$$I_2(r) - I_1(r) = U_o U_R^* \left[1 - e^{i\psi(r)} \right] + U_R U_o^* \left[1 - e^{-i\psi(r)} \right] \quad (8)$$

De la ecuación (8) puede verse que en aquellos puntos de la superficie donde el cambio de fase $\psi(r)$ introducido es nulo o múltiplo entero de 2π radianes, la intensidad resultante será nula, mientras que para otros valores de $\psi(r)$ la intensidad resultante será distinta de cero, presentando una variación sinusoidal.

En un sistema DSPI el resultado de la diferencia calculada en la ecuación (8) es mostrado en módulo en un monitor de video, ya que no es posi-

ble representar intensidades resultantes menores que cero. Se observa así una imagen del objeto cubierto por un diagrama de speckle (dado por la presencia de U_o y U_R en la ecuación (8)) modulado por un sistema de franjas sinusoidales, cuyos mínimos recorren las zonas donde $\psi(r) = 2n\pi$ radianes. El contraste del diagrama de speckle será máximo justamente entre dos mínimos de las franjas, esto es donde $\psi(r) = (2n+1)\pi$ radianes, lo que puede interpretarse como los lugares donde la fase es tal que las franjas presentan máximos de intensidad.

Las franjas así obtenidas en DSPI no son "franjas de interferencia" en el sentido de la Holografía Interferométrica sino "franjas de correlación" entre dos diagramas de speckle. La información que da origen a estas franjas está contenida en la interferencia primaria de los haces objeto y referencia de cada uno de los dos diagramas de speckle (antes y después de que se produzca el cambio de fase $\psi(r)$) que intervienen en el proceso.

Los sistemas modernos de procesamiento de imágenes están provistos de software específico que permite restar dos imágenes de 512×512 pixels cada una de ellas y presentar el resultado de esta operación (en módulo) tan sólo en el tiempo que el sistema de video tarda en adquirir una nueva imagen (típicamente $1/30$ seg.). De este modo es posible observar en continuo o a tiempo real, como evoluciona el sistema de franjas a medida que la fase $\psi(r)$ cambia en el tiempo, por ejemplo a causa de que el objeto se está deformando.

4.4 - Aplicaciones de la Interferometría Speckle Digital

a) Análisis de vibraciones:

De un modo similar al caso de la Interferometría Speckle, que encontró

su primera aplicación en la detección de vibraciones, por inspección visual, como ya se describió en la sección 3.1, la primera aplicación de la Interferometría Speckle Digital (en ese entonces todavía denominada Interferometría Speckle Electrónica) estuvo dirigida a la observación directa de líneas nodales en superficies vibrantes^[18]. En ese sistema la señal de salida de una cámara de vidicon es amplificada y sometida a un filtrado pasa-altos que mantiene presentes a las frecuencias espaciales mayores o iguales que las existentes en el diagrama de speckle, antes de ser mostrada en el monitor. De este modo, las zonas del objeto que se encuentran en reposo y están, por lo tanto, "cubiertas" por un diagrama de speckle aparecerán, brillantes en el monitor, mientras que aquellas zonas en donde el speckle presenta menor contraste (esta borroneado) a causa del movimiento del objeto se presentarán oscuras en el monitor ya que el filtrado no permitirá el paso de las frecuencias más bajas correspondientes a dichas zonas de bajo contraste. La imagen resultante en el monitor consiste de un sistema de franjas brillantes que recorren las líneas nodales del objeto vibrante.

b) Medida de desplazamientos y deformaciones:

Si se pretende realizar mediciones de desplazamientos o deformaciones, ya no es posible utilizar tan sólo las técnicas de filtrado en frecuencia que se utilizan para el análisis de vibraciones sino que es necesario implementar alguna técnica de correlación que permita la comparación del estado del objeto con algún estado de referencia.

La forma de realizar ésto en Interferometría Speckle Digital es, análogamente al caso de doble exposición fotográfica, registrar un estado inicial del objeto y luego restar la imagen correspondiente a él de la de un

estado posterior a analizar. En la sección 4.3 ya se vio como este procedimiento conduce a la formación de franjas de correlación que recorren al objeto siguiendo zonas que sufrieron igual cambio de fase debido a la deformación o traslación del objeto.

c) Ensayos no destructivos:

En estos ensayos se busca usualmente controlar o visualizar la respuesta de un componente a una determinada perturbación sin que el mismo resulte inutilizado en el proceso. El resultado "bueno" o "malo" se obtiene por comparación con el correspondiente para otro componente del mismo tipo que ya ha sido aceptado como "bueno" y constituye, por lo tanto, una referencia. Una imagen de este objeto de referencia es almacenada en la memoria de un sistema de Interferometría Speckle Digital y restada de la del objeto que se desea controlar obtenida en idénticas condiciones. Las diferencias de forma que existan entre ellos corresponderán a diferencias de fase que darán lugar a formación de franjas de correlación, las que se aglomerarán más en las zonas donde la fase varíe más rápidamente, esto es, donde la diferencia de forma respecto al componente de referencia es mayor. En base a un criterio preestablecido, por ejemplo el máximo número de franjas permitidas por unidad de área de la imagen, puede decidirse si el componente es aceptado o no.

Los sistemas de Interferometría Speckle Electrónica o Digital son muy adecuados para este tipo de aplicación ya que constituyen dispositivos de identificación de fallas que son a la vez rápidos y versátiles.

En conclusión, las técnicas de Interferometría Speckle convencionales, discutidas ya en este capítulo, son siempre factibles de ser implementadas en

sistemas digitales, con la ventaja a favor de estos últimos de poder operar en continuo y poder recomenzar el proceso (cambiando por ejemplo el estado del componente de referencia) con gran facilidad tantas veces como las condiciones de trabajo así lo requieran.

d) Contorneado de superficies utilizando DSPI:

Desde los comienzos de la Holografía Interferométrica han sido propuestos diversos métodos de contorneados de superficies utilizando técnicas holográficas, de speckle o moiré. Trabajos posteriores mostraron la posibilidad de implementar técnicas de contorneado de superficies en procesadores de imágenes digitales^[19-21].

A continuación se discuten dos variantes de un nuevo método de contorneado de superficies^[22], basadas en técnicas DSPI. Ambas versiones pueden funcionar mostrando las líneas de contorneado en continuo, esto es no existe un tiempo de espera apreciable para el observador visual entre la adquisición de los datos y la obtención de los resultados (el objeto recorrido por franjas de contorneado) en el monitor de video. En caso de que la operación en continuo o a tiempo real no sea de mayor importancia, ella puede sacrificarse para obtenerse, como compensación, franjas de contorneado que surgen de múltiples exposiciones, siendo estas mas delgadas (aumentando así la resolución del método) o de ancho variable, conformando un patrón estructurado que permite simplificar el seguimiento de las franjas individuales en el caso de superficies complejas o de perfil rápidamente variable.

Se trata, asimismo, de dos métodos que no requieren acceso mecánico al objeto ya que las líneas de contorneado se obtienen al intersectar al mismo con planos de franjas no localizadas. Esto presenta una gran ventaja en los

casos en que, por ejemplo, los objetos a contornear forman parte de un proceso de varias etapas interdependientes y se requieren controles parciales de la mismas sin alterar la posición del objeto durante este procedimiento de control.

Discusión teórica

Para la descripción teórica del proceso de obtención de franjas de contorneado pueden ser utilizados indistintamente los montajes experimentales de las Figuras 6 y 7. En ambos casos, el haz luminoso coherente proveniente de un láser es separado en dos haces por medio de un divisor de haz. Uno de ellos es utilizado para iluminar al objeto difusor (en el sentido de no especular para la longitud de onda con la que se trabaja) que se desea contornear.

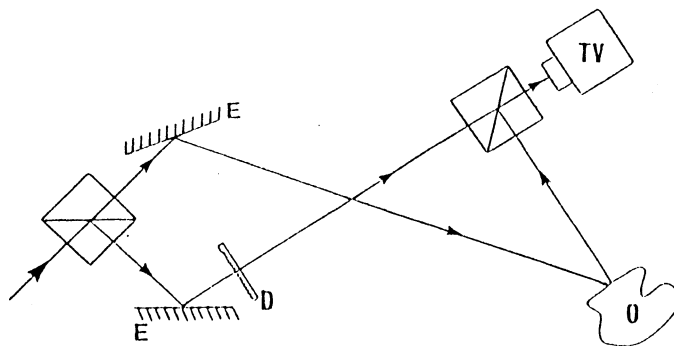


FIGURA 6
Arreglo experimental para el contorneado de superficies
utilizando DSPI.

La luz por él difundida es registrada por el sensor de la cámara de video a través de su sistema óptico, elegido de tal modo que los granos de

speckle sean resueltos por el elemento foto-sensible de la misma; para esto se requiere un tamaño promedio de los granos de speckle del orden del tamaño de un pixel del sensor de CCD. El tamaño medio de los granos de speckle puede ajustarse variando el diafragma del sistema óptico de la cámara. El segundo haz es superpuesto al primero como haz de referencia y permite, por lo tanto, obtener un registro no sólo de la amplitud luminosa proveniente del objeto sino también de la fase en cada punto del mismo.

En el montaje de la Figura 6 el haz de referencia es provisto por el diagrama de speckle que produce un difusor plano (placa de vidrio despulido) D, que es captado por la cámara simultáneamente con el haz objeto con ayuda de un divisor de haz.

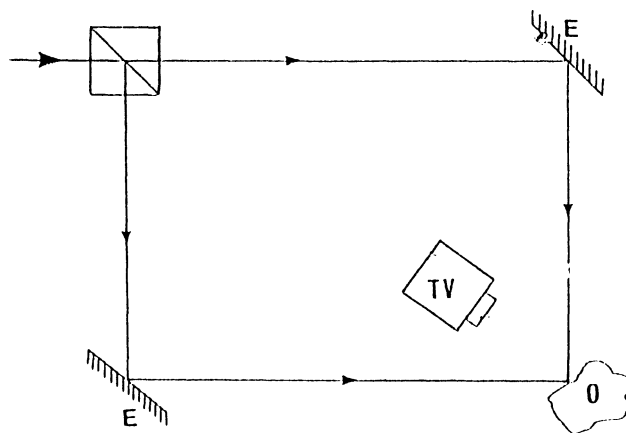


FIGURA 7

Variante del sistema DSPI para el contorno de superficies. Tanto el haz objeto como el haz referencia son generados por el objeto que se desea contornear.

En el montaje de la Figura 7 los dos haces que se obtienen al dividir el haz proveniente del láser se utilizan para iluminar el objeto. El haz referencia corresponde, en este caso, al diagrama de speckle generado por el ob-

jeto debido a la iluminación de cualquiera de los dos haces. El diagrama de speckle generado por el objeto debido a la iluminación del segundo haz constituye el haz objeto. Por supuesto, ambos haces son intercambiables.

Para obtener las curvas de contorneado, una primera imagen del objeto es almacenada en la memoria digital del procesador de imágenes y, seguidamente, la dirección del haz que ilumina al objeto (o la de cualquiera de los haces que lo iluminan si se trata del montaje de la Figura 7) es rotada un pequeño ángulo $\Delta\theta$. A continuación se recommienza el proceso de adquisición de imágenes. La imagen registrada por la cámara en un instante cualquiera t es restada pixel a pixel de la previamente almacenada y el resultado se muestra, en módulo, en el monitor de video, actualizándose cada vez que una nueva imagen es leída por el procesador; ésto ocurre a razón de 30 imágenes por segundo.

La expresión para el campo eléctrico correspondiente a la imagen inicial puede escribirse

$$E_1 = E_o + E_r \quad , \quad (9)$$

donde E_o es la contribución debida al haz objeto y E_r la correspondiente al haz de referencia. La intensidad total registrada por la cámara para esta primera imagen es entonces

$$I_1 = E_1 E_1^* \quad . \quad (10)$$

Para el campo y la intensidad de las imágenes subsiguientes se obtienen expresiones análogas.

La substracción de dos imágenes da como resultado, en módulo

$$|I_2 - I_1| = 4 |E_o E_r \text{ sen} \left[\phi(x,y) + \Psi/2 \right] \text{ sen}(\Psi/2)| \quad , \quad (11)$$

siendo $\phi(x,y)$ una fase aleatoria y Ψ es el cambio de fase que se introduce para cada punto a causa de la rotación del haz objeto. La expresión (11) es deducible de la ecuación más general (8) obtenida en el ítem 4.3. El primer factor sinusoidal en la ecuación (11) describe al objeto modulado por un diagrama de speckle mientras que el segundo representa a un conjunto de franjas que indican la intersección del objeto con superficies definidas por $\Psi = \text{constante}$. El espaciado de los planos $\Psi = \text{constante}$ está dado por la expresión^[23]

$$s = \frac{\lambda}{2 \text{ sen} \frac{\Delta\theta}{2}} \quad , \quad (12)$$

siendo λ la longitud de onda de iluminación y $\Delta\theta$ el ángulo que giró el haz objeto entre ambas exposiciones.

Resultados experimentales

En la Figura 8b se muestran los resultados del contorneado de dos objetos curvos utilizando la técnica DSPI recientemente analizada. Para comparación, la Figura 8a presenta al mismo par de objetos con líneas de contorneado obtenidas por Holografía Interferométrica, aplicando el procedimiento propuesto por Hildebrand y Haines^[9]. Puede observarse, salvo un factor de escala,

la similitud de los resultados; sin embargo en la Figura 8b está presente el ruido de speckle inherente a la técnica DSPI.

Para obtener líneas de contorneado más delgadas se realizó la experiencia utilizando múltiples exposiciones^[1]. En este caso debe ser sacrificada la operación a tiempo real, ya que no es posible para el procesador de imágenes con que se trabajó (ITEX 100) realizar operaciones sobre más de dos

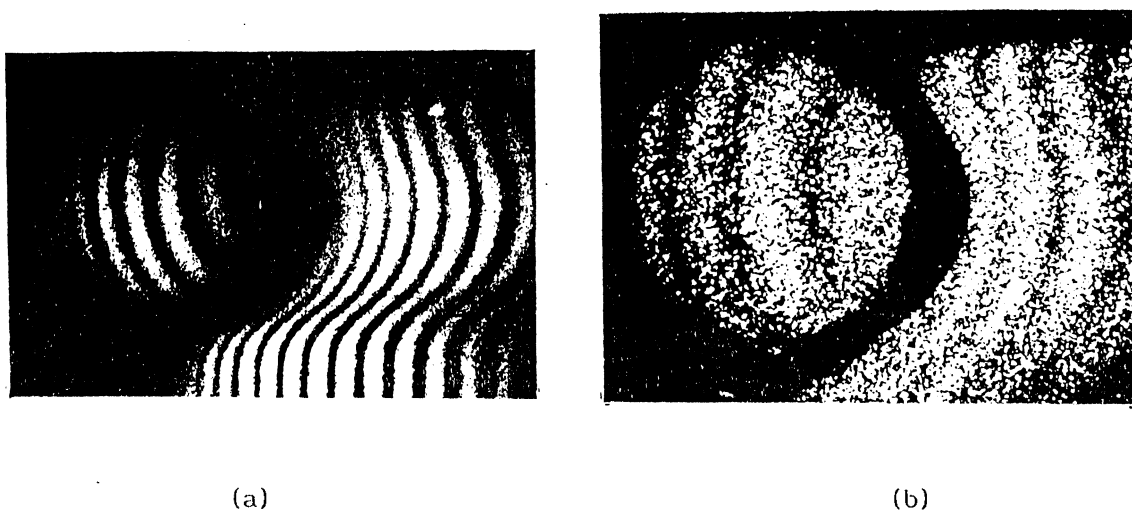


FIGURA 8

a) Resultado de contornear dos objetos curvos con el método holográfico propuesto por Hildebrand y Haines. b) El mismo conjunto de objetos contorneados utilizando la técnica DSPI. Nótese la presencia del ruido de speckle.

imágenes simultáneamente. Utilizando el arreglo experimental de la Figura 7 se adquirieron cuatro imágenes sucesivas con incrementos progresivos de igual magnitud en el ángulo de iluminación de uno de los haces (el que se eligió como haz objeto) y las mismas fueron almacenadas en archivos auxiliares de acceso secuencial.

Sea $Z_i(x,y)$ el valor de intensidad del pixel en la posición (x,y) de la i -ésima imagen. Por medio del software adecuado, trabajando con los cuatro archivos secuenciales correspondientes a las cuatro imágenes, se asigna al pixel en la posición (x,y) el valor resultante Z , obtenido de la operación

$$Z = | Z_1 - Z_2 + Z_3 - Z_4 | + | Z_1 - Z_3 | + | Z_1 - Z_4 | \quad ; \quad (13)$$

en donde se ha eliminado, por simplicidad, la dependencia explícita de cada valor Z_i con las coordenadas (x,y) .

En la ecuación (13), el primer término da lugar a la formación de franjas de correlación de frecuencia espacial baja dado que contiene diferencias entre pares de imágenes separadas por un único incremento angular $\Delta\theta$; la ecuación (12) da, en este caso, el mayor espaciado entre planos de corte. Al aumentar $\Delta\theta$ el espaciado de los planos que cortan al objeto disminuye en consecuencia; así, por ejemplo, entre la imagen inicial (representada por los valores de intensidad Z_1) y la tercera imagen de la serie (a la que le corresponden los valores de intensidad Z_3) el incremento angular total en el haz objeto es el doble que para el caso de imágenes con subíndices consecutivos y la frecuencia espacial resultante para las franjas de correlación es mayor debido a que la diferencia de fase introducida es un función más rápidamente variable con la posición. El último término de la ecuación (13) da origen a las más altas frecuencias espaciales compatibles con el caso planteado ya que involucra imágenes entre las cuales existe el máximo incremento angular de la serie.

La Figura 9 muestra el resultado de aplicar la operación de la ecuación (13) a una porción ampliada del conjunto de objetos de la Figura 8. Además de

la presencia de más de una frecuencia espacial en las franjas de correlación es posible observar una reducción en el ruido de speckle como resultado de la operación sobre cuatro imágenes.

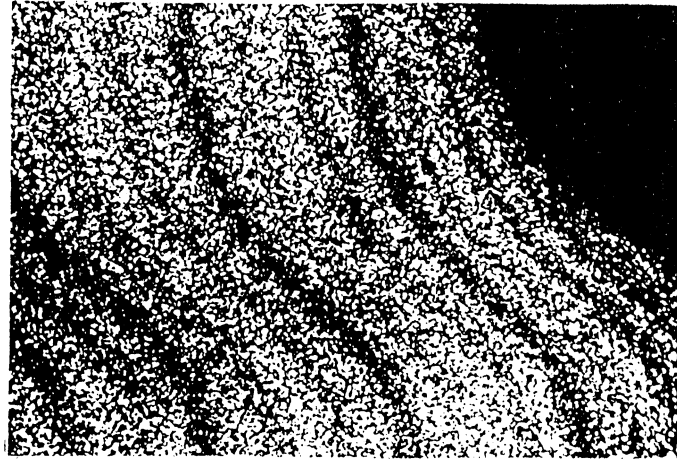


FIGURA 9

Obtención de franjas de contorneado por multiple exposición
Se aprecia la presencia de mas de una frecuencia espacial

Variando los "pesos" con que los valores de intensidad de las distintas imágenes intervienen en la ecuación (13) es posible obtener sistemas de franjas con diferentes perfiles, realzando algunas frecuencias espaciales y atenuando otras. En particular, si los pesos se eligen en las proporciones de los coeficientes binomiales, las franjas resultantes no presentan máximos secundarios^[1], resultando un conjunto de franjas muy delgadas y de alto contraste. El problema es equivalente al de interferencia por N rendijas idénticas, donde la transmitancia de cada una de ellas puede ser ajustada independientemente de la de las demás.

Finalmente, si se utiliza el montaje de la Figura 7, una vez que se han obtenido las franjas es posible alterar la dirección de iluminación del haz

de referencia de modo tal que los cambios de fase que no se deban exclusivamente a la topografía del objeto pueden ser compensados^[10], obteniéndose planos de contorneado perpendiculares a la dirección de observación. En la práctica esto puede realizarse colocando junto al objeto que se desea contornear un objeto plano auxiliar; cuando al rotar el haz de referencia se logra eliminar las franjas sobre este objeto plano no existen en el sistema variaciones de fase en la dirección de la superficie del mismo y sólo se observan franjas (en el objeto curvado) debidas a cambios de fase provocados por variaciones de profundidad.

e) Detección de la "Luz en Vuelo" utilizando DSPI

En 1978, Nils Abramson^[24] demostró un método que denominó "Holografía de la luz en vuelo" (Light-in-flight Holography) que permitía obtener en un montaje holográfico la ubicación de los puntos donde los haces objeto y referencia presentan una diferencia de camino óptico nula. Métodos muy similares fueron propuestos por Boden y coautores^[25] y por Denysiuk y coautores^[26].

Por otro lado, como ya se vio en secciones anteriores, DSPI se ha revelado como una poderosa herramienta para numerosas aplicaciones asociadas normalmente con la Holografía Interferométrica en el plano imagen, por lo cual DSPI es llamada también por algunos autores "Holografía Digital".

Las tendencias actuales en física de láseres conducen a la obtención de pulsos ultracortos (del orden de unos pocos picosegundos o incluso de femtosegundos) con miras a depositar grandes cantidades de energía sobre un blanco en tiempos muy reducidos. Si bien la generación de estos pulsos constituye el problema fundamental, la detección y medición de los mismos no es un problema trivial ya que lleva a la electrónica al límite de sus posibilidades,

debiéndose utilizar técnicas sofisticadas e instrumental de detección muy costoso.

A su vez, un pulso de esta duración producirá en un sistema holográfico o DSPI, en forma equivalente a lo que lo haría un haz láser de baja longitud de coherencia, al iluminar un objeto en forma no perpendicular al mismo, una franja brillante de ancho reducido; la duración del pulso puede determinarse entonces a partir del ancho de la misma.

A continuación, teniendo presentes las ventajas de los sistemas DSPI, se presenta una adaptación de uno de tales sistemas digitales para obtener, en continuo, un registro de la luz en vuelo^[27].

Discusión Teórica y Resultados Experimentales

El principio de operación del método es el usual para un sistema DSPI: la imagen de un objeto iluminado por un láser, es registrada por una cámara de video de CCD, superponiéndosele un haz de referencia coherente que permite obtener información acerca de la fase de los puntos del objeto, y almacenada en la memoria digital de un procesador de imágenes. Luego de que esta primera imagen ha sido adquirida y almacenada se produce un cambio de fase en el haz proveniente del objeto, las imágenes que le suceden son restadas de ella y el resultado es mostrado, en módulo, en un monitor de video.

El restado produce, en las zonas donde el cambio de fase es de múltiplos enteros de 2π radianes (el diagrama de speckle se repite) un resultado nulo, generando áreas oscuras en el monitor. En los puntos donde ocurrieron otros cambios de fase el diagrama de speckle de ambas exposiciones no se repetirá y el restado dará un resultado no nulo, obteniéndose zonas de nivel de gris variable en el monitor. Se obtienen, de este modo, franjas que recorren al ob-

jeto siguiendo las zonas de diferencia de fase constante.

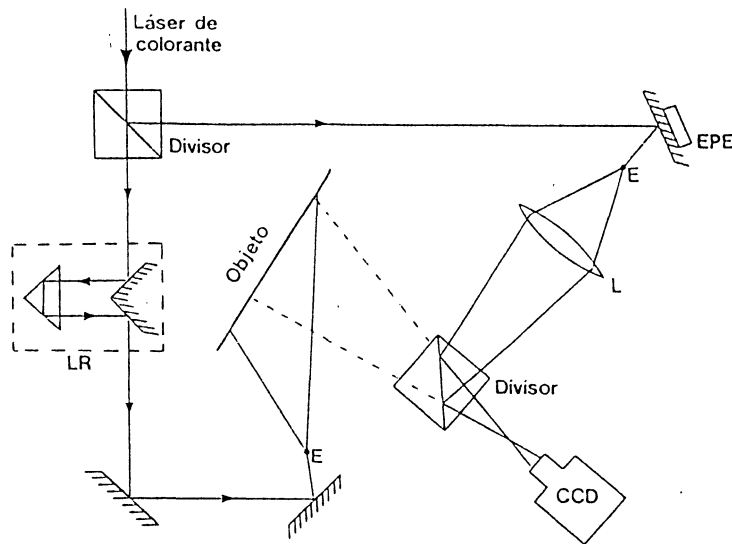


FIGURA 10
Interferómetro Speckle Digital para registro de la luz en vuelo.

El montaje experimental utilizado se muestra en la Figura 10. Un láser de colorante (Rodamina 110, $\lambda=550$ nm) bombeado por un láser continuo de Argón, de 5W de potencia, se utilizó como fuente de luz de longitud de coherencia reducida (de algunos milímetros). Un espejo comandado por un cristal piezoeléctrico (EPE) se utilizó para provocar un cambio de fase de π radianes ($\lambda/2$) entre ambas exposiciones. La lente L conjuga el punto focal del expansor E en el plano focal anterior de la lente de la cámara, produciendo así una onda plana sobre el detector de la misma. La imagen que registra el elemento sensible de la cámara es la superposición de un diagrama de speckle (proveniente del objeto) con un fondo uniforme proveniente de haz de referencia.

En todos los puntos de la imagen donde la diferencia de camino óptico

(DCO) sea mayor que la longitud de coherencia de la luz, el cambio de fase no es de importancia. En efecto, en estos puntos, la superposición de los haces objeto y referencia es incoherente y un cambio de fase no alterará la intensidad total en el plano del detector, que resulta ser la suma de las intensidades individuales. Por lo tanto, el procedimiento de restado dará en estas zonas de la imagen un resultado nulo, representado como oscuro en el monitor.

En las regiones donde la superposición sea coherente, esto es donde se cumpla que la DCO sea menor que la longitud de coherencia del láser, un cambio de fase de π radianes modifica al diagrama de speckle reemplazando los granos de speckle "claros" por "oscuros" y viceversa. El restado produce en el monitor una zona angosta (para una longitud de coherencia reducida como la que se utilizó) de alto brillo que intersecta al objeto y que indica, precisamente, los lugares geométricos donde la DCO está dentro de la longitud de coherencia del láser.

Una línea de retardo (LR) como la que se muestra en la Figura 10 permite elegir la zona del objeto donde se cumpla la condición: $DCO < \text{longitud de coherencia del láser}$, para obtener la franja brillante en esta zona del mismo. Variando la posición de la LR puede hacerse "viajar" a la franja brillante sobre el objeto, permitiendo visualizar el avance del frente de ondas sobre el mismo.

Los resultados experimentales se muestran en la Figura 11 para la intersección de una onda esférica con un objeto plano.

En una segunda experiencia, una lámina de vidrio de caras planas paralelas se colocó en la región central del objeto plano de modo tal que el frente de ondas que ilumina a dicho objeto debía atravesarla. La Figura 12 muestra el retardo que la lámina de vidrio introduce en la porción del frente de on-

das que pasa por ella. Para la lámina utilizada, de un milímetro de espesor e índice de refracción 1,5, el retardo que se observa es de aproximadamente 5 ps, mostrando la sensibilidad del método.

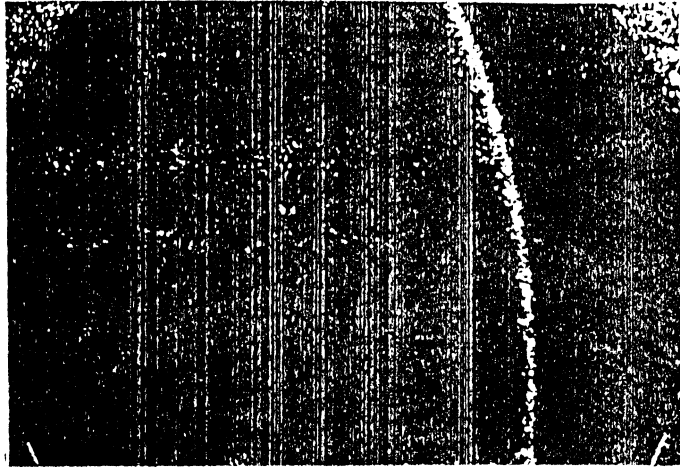


FIGURA 11

Registro de un frente de ondas (línea brillante) curvo cuando este intersecta a una superficie plana. El ancho de la línea brillante se corresponde con la longitud de coherencia del laser.

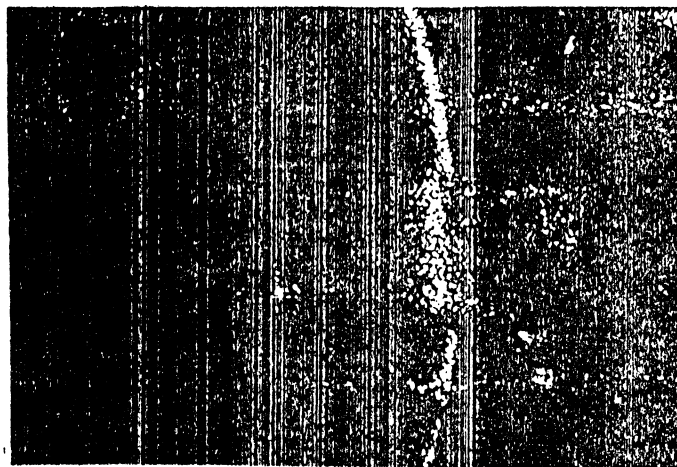


FIGURA 12

Fotografía del mismo frente de ondas de la Figura 11 cuando se interpone, en su zona central, una lámina de vidrio delgada. Se puede apreciar el retraso de la luz en dicha zona central debido al índice de refracción del vidrio.

f) Sistema DSPI para medir desplazamientos:

Como otra aplicación metrológica de los sistemas DSPI se presenta un sistema apto para la medición de desplazamientos, cuyo montaje experimental se

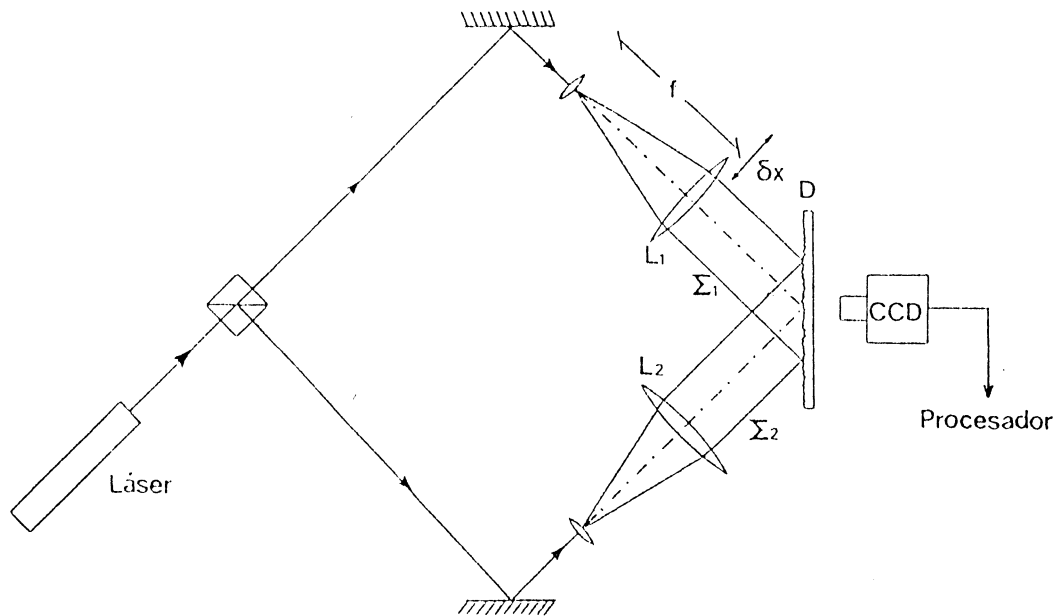


FIGURA 13
Esquema del Interferómetro Speckle Digital utilizado
utilizado para la medición de desplazamientos.

muestra en la Figura 13. Este consiste en un interferómetro speckle digital modificado^[23] en el cual dos haces provenientes de un mismo láser interfieren sobre el difusor D , que genera el diagrama de speckle. Una cámara de CCD produce una imagen del mismo que es enviada en forma analógica al sistema de digitalización y procesamiento de imágenes. Las lentes L_1 y L_2 se ubican de tal modo que la luz emergente de los expansores E lo hace en forma de ondas planas, Σ_1 y Σ_2 respectivamente, luego de pasar por ellas. La lente L_1 se

monta, a su vez, sobre un sistema que permite desplazarla lateralmente en dirección perpendicular a la del haz Σ_1 . Los desplazamientos laterales, δx , son producidos por un motor paso a paso (de un desplazamiento mínimo de $2\mu\text{m}$) y controlados con un sistema piezoeléctrico.

Con ayuda de la Figura 14 puede observarse que un desplazamiento lateral de L_1 equivale a desplazar el punto focal del expansor (de donde emerge la luz), ubicado, a su vez, en el plano focal de la lente, en la misma cantidad pero en sentido contrario. Como resultado de ello, la onda plana Σ_1 , emergente de

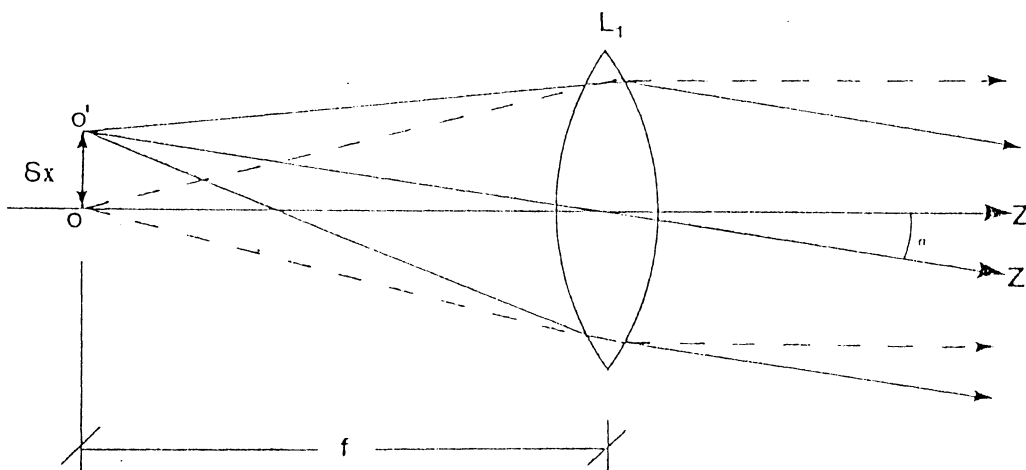


FIGURA 14
Representación esquemática del efecto de desplazar lateralmente una fuente puntual ubicada en el plano focal de una lente.

L_1 rota un ángulo $\alpha = \text{arctg}(\delta x/f)$ respecto a su dirección original, siendo f la distancia focal de la lente L_1 . Se obtiene así una nueva onda plana, que llamaremos Σ_1' . Esta rotación del haz Σ_1 , identificado como "haz objeto" producirá sobre el difusor D un cambio de fase proporcional a dicha rotación y, por lo tanto al desplazamiento δx , que puede ser registrado por un sistema

DSPI en formas de franjas de correlación, las que codifican dicho desplazamiento en su espaciado. Siguiendo el procedimiento que ya se explicó, luego de almacenar una primera imagen de D, se provoca un cambio de fase (en este caso desplazando lateralmente a L_1) y el nuevo diagrama de speckle generado por el difusor es registrado por la cámara, transmitido al procesador de imágenes y restado pixel a pixel del primero. El resultado, mostrado en módulo en el monitor de video son franjas de perfil cosenoidal que presentan zonas oscuras donde el cambio de fase resultó ser un múltiplo entero de 2π radianes.

De la Figura 15 puede verse que si α_0 es el ángulo con que Σ_1 incide sobre el difusor, entonces, luego de girar un ángulo α , la diferencia de camino óptico (DCO) entre Σ_1 y Σ_1' (luego de la rotación) en algún punto del difusor puede expresarse como:

$$DCO = b - a = Z \text{ sen } (\alpha_0 + \alpha) - Z \text{ sen } \alpha_0, \quad (14)$$

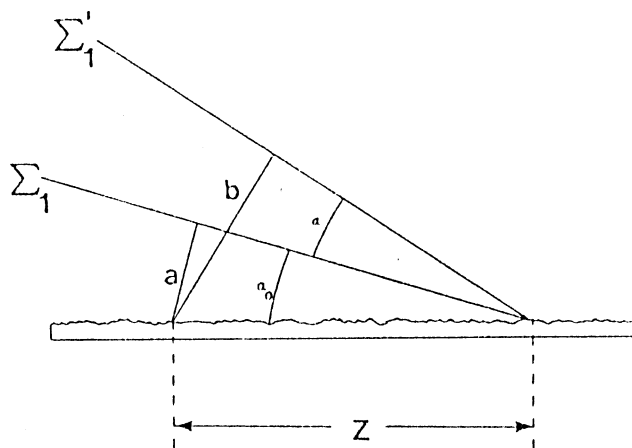


FIGURA 15
Diferencia de Camino Óptico para dos ondas planas que alcanzan una pantalla viajando en distintas direcciones.

siendo Z la distancia, medida sobre el plano del difusor, entre el punto en cuestión y el punto origen, donde ambos frentes de onda se cortan.

Más explícitamente, la ecuación (14) puede escribirse

$$DCO = Z (\sin \alpha \cos \alpha_0 + \cos \alpha \sin \alpha_0) - Z \sin \alpha_0 \quad (15)$$

Si el cociente $\delta x/f$ es pequeño, entonces lo es el ángulo α y, por lo tanto, puede considerarse $\cos \alpha \cong 1$ y $\sin \alpha \cong \tan \alpha = \delta x/f$.

Esta condición no es difícil de satisfacer, ya que la distancia focal de la lente puede elegirse suficientemente grande (del orden de 30 o 40 cm) y los desplazamientos que se desean medir son inferiores a un milímetro.

Luego, la DCO resulta:

$$DCO \cong Z \left[\frac{\delta x}{f} \cos \alpha_0 \right] \quad (16)$$

Además, si la DCO es igual a una longitud de onda λ de la luz con que se ilumina al difusor, entonces Z será igual al espaciado entre franjas d . Esto es,

$$\lambda = d \frac{\delta x}{f} \cos \alpha_0,$$

de donde, conocidos λ , f y α_0 , puede obtenerse el valor del desplazamiento δx midiendo el espaciado d de las franjas. Este resulta:

$$\delta x = \frac{1}{d} \frac{\lambda f}{\cos \alpha_0} \quad (17)$$

El valor de d puede obtenerse, por ejemplo, por aplicación de un algo-

ritmo de Transformada Rápida de Fourier (FFT) al sistema de franjas. Sin embargo, existen otras alternativas de evaluación de la interfranja d , menos complejas desde el punto de vista computacional y que brindan resultados tan buenos como la FFT. Una de estas alternativas la constituye el algoritmo basado en autocorrelación discreta que se describe a continuación^[28].

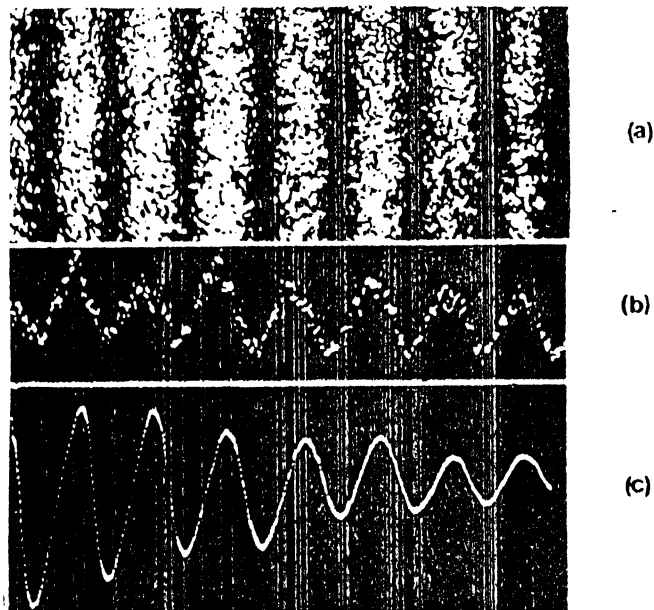


FIGURA 16

a) Sistema de franjas obtenidas en un sistema DSPI. b) Vector obtenido del sistema de franjas promediando 100 pixels en dirección paralela a las franjas. c) Autocorrelación discreta del vector representado en b). Se puede observar la fuerte reducción de ruido obtenida luego de esta operación.

Para implementar el algoritmo propuesto se supone un sistema de franjas, con ruido de speckle, orientadas en el monitor de video del sistema DSPI en dirección paralela a la dirección cartesiana \underline{Y} , que se encuentra almacenado en una memoria de imágenes (Frame Grabber) de una extensión de 512x512 ele-

mentos de imagen o pixels. La Figura 16a muestra un sistema de franjas como el propuesto.

El parámetro a medir es el espaciado d de las franjas de correlación, esto es, la distancia que separa a dos puntos para los cuales la diferencia de fase $\Psi(r)$ en la ecuación (8) es de 2π radianes. Esto no puede realizarse directamente dado que el fuerte ruido de speckle enmascara la ubicación real de las máximos y mínimos del sistema de franjas, a partir de los cuales es posible hallar d . Por lo tanto se proponen los siguientes pasos:

a) A partir del diagrama de franjas se construye un vector en dirección X de 512 posiciones, en cada una de las cuales se almacena el valor medio de la intensidad de un cierto número de pixels (típicamente 100) en la dirección Y, en torno a una posición arbitraria y_0 . De este modo se logra un primer paso en la reducción del ruido presente debido al speckle^[29]. En símbolos, se obtiene una expresión

$$g[x] = \frac{\sum_{y=y_0 - n/2}^{y_0 + n/2} g[x, y]}{n}, \quad (18)$$

donde $g[x, y]$ representa al nivel de gris del pixel ubicado en las coordenadas (x, y) del sistema de video y $g[x]$ es su valor medio sobre n pixels.

La Figura 16b muestra una gráfica del vector obtenido del sistema de franjas de la Figura 16a aplicando la ecuación (18) con $n=100$.

b) Como puede apreciarse de la Figura 16b, el vector $g[x]$ es aún demasiado "ruidoso" como para que de él pueda obtenerse sin ambigüedad el valor

de \underline{d} .

Por lo tanto, tomando como base la definición de autocorrelación de una función continua definida en el intervalo $(-\infty, \infty)$

$$\Lambda(x) = \int_{-\infty}^{\infty} f(\alpha) f(\alpha-x) d\alpha \quad (19)$$

se propone calcular el equivalente discreto para el vector $g[x]$, dado por:

$$\Lambda[x] = \frac{1}{K} \left[\sum_{i=0}^{512} g[i] g[x+i] \right] + X \quad (20)$$

La razón básica para proponer la ecuación (20) es la periodicidad de $g[x]$, que no se alterará (como se verá) por medio de esta operación, mientras que el ruido será reducido fuertemente ya que para la obtención de cada punto de $\Lambda[x]$ se utilizan todos los elementos de $g[x]$.

Para poder llevar a cabo el cálculo de la ecuación (20), el vector $g[x]$ debe ser extendido (se le deben agregar elementos) con ceros, de modo tal que las posiciones con índice de suma $\underline{x+i}$ mayor que 512 (longitud real del vector) no queden sin definir.

Al operar según (20), la contribución del primer término es tal que $\Lambda[x]$ oscila debido a la superposición alternada de máximos con máximos y máximos con mínimos del vector $g[x]$, a medida que x aumenta. El efecto de los ceros agregados a $g[x]$ o, más exactamente, el hecho de que $g[x]$ tenga una extensión finita hace que $\Lambda[x]$ presente, además, una disminución en amplitud a medida que se reduce la cantidad de ciclos de $g[x]$ que se superponen. Este efecto es tan importante que la amplitud de los máximos que suceden al primero puede no

superar por mucho al mínimo precedente, dificultando, de este modo, la posterior ubicación de los extremos relativos de intensidad de $A[x]$. El segundo término en la ecuación (20) corrige este problema sin alterar la posición relativa de dichos valores extremos ya que tiene derivada segunda nula. Más formalmente puede decirse que, como $g[x]$ es de extensión finita, resulta del producto de un vector $g'(x)$, de extensión infinita y coincidente con $g[x]$ en sus 512 elementos, con una función rectángulo que vale cero en todo lugar excepto en los 512 elementos de $g[x]$, donde toma el valor uno. La autocorrelación de $g[x]$ se presenta entonces modulada por la de esta función rectangular, que tiene la forma triangular ilustrada en la Figura 17. Por lo tanto $A[x]$ debe disminuir gradualmente su amplitud, según indica este perfil triangular.

Por medio de la corrección introducida por el segundo término de (20) se produce una autocorrelación discreta $A[x]$ con los máximos y mínimos claramente diferenciables, modulada por una envolvente que la hace decrecer gradualmente en amplitud.

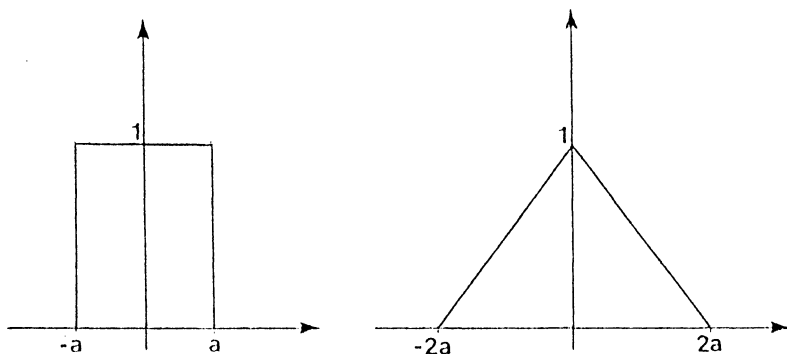


FIGURA 17
Función rectángulo (a) y su autocorrelación (b).

El resultado final de la operación indicada en (20) es entonces una función del mismo período que $g[x]$ (ya que los máximos de $A[x]$ ocurren cuando se superponen los máximos de $g[x]$) y para la cual se ha reducido drásticamente el ruido, ya que todos los puntos de $g[x]$ aportan información para cada punto de $A[x]$. La Figura 16c ilustra dicho resultado.

El factor de normalización $1/K$ se elige de tal modo que cuando x es cero (la suma de los productos se realiza con el vector sin desplazar) $A[0]$ tenga un valor representable por el procesador de imágenes en el monitor de video.

c) Controlando los cambios en la pendiente o primera derivada discreta del vector $A[x]$ (dada por $A[x+1] - A[x]$) puede obtenerse la ubicación de los mínimos del mismo. En una señal prácticamente libre de ruido como lo es $A[x]$, un extremo puede ser localizado con un error no superior a unos pocos (3 a 5) pixels, según sea la "sensibilidad" elegida en el software.

d) Calculando la distancia promedio (en pixels) entre dos mínimos consecutivos se obtiene el espaciado de las franjas \bar{d}

$$\bar{d} = \frac{\sum_{k=1}^{N-1} d_k}{N - 1} \quad (21)$$

donde N es el número de mínimos que han sido detectados en la extensión total de $A[x]$ y d_k es la distancia, en pixels, entre el k -ésimo mínimo y el que le antecede.

Para calibrar al sistema en unidades de longitud se ejecutó el algoritmo

sobre una red de Ronchi, cuyo espaciado se midió previamente con una precisión de $1\mu\text{m}$. Al período en pixels así obtenido se le asigna el valor en μm correspondiente y, de este modo, la ecuación (17) da el desplazamiento δx en μm si se miden λ en μm y f y d en la mismas unidades.

Utilizando el arreglo experimental de la Figura 13, con $\lambda = 0,633\mu\text{m}$, $\alpha_0 = 19^\circ$ y una distancia focal de 30cm para la lente L_1 , se provocaron distintos desplazamientos laterales (entre 80 y $220\mu\text{m}$) a la misma, los que fueron controlados utilizando el transductor piezoeléctrico, con una precisión de $2\mu\text{m}$.

Los resultados obtenidos para dichos desplazamientos por medio de la ecuación (17) se presentan en la Figura 18. La línea recta de pendiente unidad representa las posiciones que los puntos medidos utilizando la información contenida en las franjas de correlación ocuparían si la concordancia con los desplazamientos medidos mecánicamente fuese ideal.

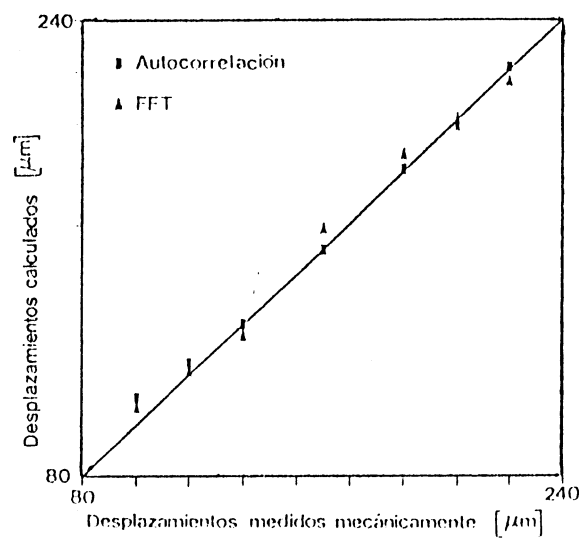


FIGURA 18

Resultado de la medición de desplazamientos a partir de la información contenida en las franjas de correlación.

Por comparación se han representado en la misma gráfica los resultados dados por la ecuación (17) cuando \underline{g} es obtenido del diagrama de franjas utilizando un algoritmo de transformada rápida de Fourier (FFT).

5 - TECNICA DE CORRIMIENTO DE FASE (PHASE STEPPING)

5.1 - INTRODUCCION

Como ya ha sido mencionado, la interferometría, además de ser un punto de interés académico, ha contribuido a resolver problemas de control de calidad en componentes ópticos^[30], a las ciencias médicas^[31] y a diversas ramas de la ingeniería, principalmente cuando se trata de realizar ensayos sobre objetos que sufren desplazamientos o deformaciones muy pequeñas (del orden de algunas longitudes de onda) y que se pretende conservar intactos luego de las mencionadas pruebas^[32,33].

Por otro lado, como se vio, el resultado de toda experiencia interferométrica, holográfica o de speckle, se presenta en forma de franjas. En todos los casos, las franjas codifican la información metrológica que se desea obtener.

En muchas aplicaciones es deseable que los resultados interferométricos sean expresados de modo tal que personal no especializado pueda comprenderlos o para que puedan ser combinados con otras técnicas no interferométricas. Es necesario, entonces, un análisis de los interferogramas cuyo producto final sea la expresión numérica o gráfica de la información contenida en las franjas.

En los inicios de la interferometría, hacia los comienzos de 1800, Thomas Young realizó manualmente el análisis de franjas de interferencia cuando trataba de medir, a partir de la información que ellas codifican en su espaciado, la longitud de onda de la luz.

Durante los años 60, aún cuando la mayor parte del proceso seguía siendo manual, se fueron incorporando diversos equipamientos electrónicos que permi-

tieron, por ejemplo, la localización más exacta de las franjas dentro de un interferograma^[34,35] y, a partir de principios de los años 80, cuando se comenzaron a utilizar equipos de procesamiento de imágenes, se dedicó un esfuerzo considerable a la obtención de algoritmos de análisis automático de franjas^[36]. De este modo fue posible eliminar los tediosos pasos manuales del proceso, favoreciendo su incorporación a sistemas automatizados.

5.2 - ANALISIS DE FRANJAS POR CORRIMIENTO DE FASE

Una de las técnicas de mayor importancia para el análisis automatizado de franjas es la conocida como de "Corrimiento de Fase" (Phase stepping en el idioma original). Esta técnica, propuesta inicialmente por Carré^[37] ha adquirido gran importancia en la obtención de la fase de un frente de ondas en medidas interferométricas debido a que posibilita un mapeo punto a punto de la misma. En la literatura, esta técnica, que consiste esencialmente en la adquisición y posterior evaluación de varios sistemas de franjas que difieren solamente en el hecho de que entre ellos se ha producido un cambio de fase discreto, suele denominarse, en ocasiones, "cuasi-heterodina", ya que una variación de la fase en el tiempo puede interpretarse también como debida a un cambio en frecuencia.

Existen diversas maneras de implementar prácticamente el método de corrimiento de fase. A modo de ejemplo puede considerarse un interferómetro del tipo Twyman-Green, utilizado normalmente para el control de superficies ópticas, como el mostrado en la Figura 19. En esta configuración, el espejo E, que constituye la superficie de referencia se monta sobre un trasladador piezoeléctrico (TPE) que puede, a su vez, ser controlado por una computadora y que permite introducir los cambios de fase necesarios. Cuando se opera este

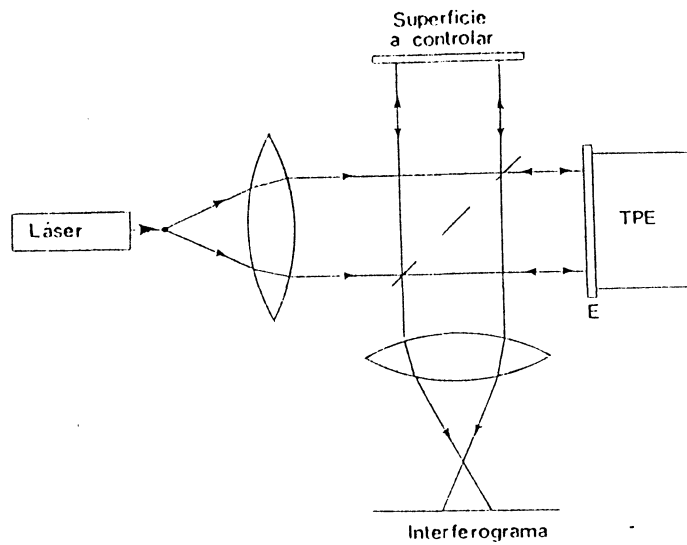


FIGURA 19
Esquema de un interferómetro de Twyman-Green

sistema en un montaje de DSPI, un primer interferograma es registrado por una cámara de video, digitalizado y almacenado en la computadora. A continuación, la fase del haz de referencia, proveniente del espejo en la Figura 19, es variada una cantidad $2\pi/k$, siendo k un entero; para lograr esto, el espejo es desplazado una distancia $\lambda/2k$, donde λ es la longitud de onda de iluminación. El nuevo interferograma así obtenido es almacenado en la computadora y se produce un nuevo cambio de fase, igual al primero. El proceso se repite hasta que el espejo de referencia se ha desplazado $(k-1)$ pasos iguales y se han adquirido k imágenes. De este modo se dispone de k niveles de intensidad registrados para cada pixel del sistema. Para calcular la fase $\phi(x,y)$ en la posición correspondiente al pixel de coordenadas (x,y) se evalúan los coeficientes de la serie de Fourier

$$\alpha(x,y) = \sum_{N=0}^{k-1} I_N(x,y) \cos\left(\frac{2\pi N}{k}\right)$$

(22)

$$\beta(x,y) = \sum_{N=0}^{k-1} I_N(x,y) \sin\left(\frac{2\pi N}{k}\right)$$

donde I_N es el valor de intensidad del pixel de coordenadas (x,y) en la N -ésima imagen. La fase puede entonces obtenerse de la expresión:

$$\phi(x,y) = \operatorname{arctg} \frac{\beta(x,y)}{\alpha(x,y)}$$

(23)

Si bien en teoría es posible utilizar tantos interferogramas como se desee, usualmente y, esencialmente por razones de tiempo de cálculo, se utilizan sólo cuatro (tres desplazamientos del espejo). En este caso, la ecuación (23) se reduce a la forma sencilla:

$$\phi(x,y) = \frac{I_1 - I_3}{I_0 - I_2}$$

(24)

que permite calcular rápidamente la fase en cada punto del interferograma.

En resumen, los valores de la fase están disponibles para cada pixel, las características o detalles cóncavos y convexos son resueltos sin ambigüedades, el ruido estático es eliminado (ya que es el mismo en todas las imágenes y la ecuación (24) sólo incluye diferencias entre ellas) y el método

es más preciso que aquellos basados en las intensidades de una única imagen.

Para el caso particular de la Interferometría Speckle Digital (DSPI), en el cual, como ya se vio, los diagramas de speckle se utilizan como portadores de información, el nivel de ruido presente en un sistema de franjas o interferograma es muy elevado y los métodos de análisis basados en intensidad son difíciles de aplicar. En estos casos, la técnica de corrimiento de fase resulta muy adecuada para la obtención de los valores de la fase punto a punto y, relacionadamente, los desplazamientos o deformaciones sufridos por objetos difusores.

En el caso de DSPI la evaluación de la fase se realiza en forma ligeramente distinta de la que se describió^[38] para el caso del interferómetro de Twyman-Green. Al igual que se muestra en la Figura 19, un espejo de referencia se monta sobre un trasladador piezoeléctrico, responsable de generar los cambios de fase requeridos. En el primer specklegrama (o imagen) del objeto a analizar se provocan los cambios de fase deseados y, utilizando la ecuación (23) se obtienen los valores $\phi(x,y)$ de la fase en cada punto (x,y) . De la misma manera, luego de someter al objeto al esfuerzo cuyo efecto se desea medir o controlar, se obtienen los valores $\phi'(x,y)$ de la fase en cada punto para esta nueva situación. Seguidamente, al restar punto a punto ambos conjuntos de fases se obtiene una medida de la deformación de la superficie del objeto, expresada en términos de cambio de fase. Cada vez que la diferencia de fase acumulada alcanza el valor de 2π radianes se produce un salto discontinuo, de modo tal que las nuevas "franjas de fase" que se obtienen de la operación de restado y cuya intensidad en cada punto de la imagen es proporcional al cambio de fase en ese punto, presentan un perfil modulado por una función "diente de sierra".

6 - REFERENCIAS

- 1 - J. M. Burch y J. M. J. Tokarsky, *Optica Acta*, 15, 101 (1968).
- 2 - K. A. Stetson, *Opt. Eng.*, 14, 482 (1975).
- 3 - J. A. Leendertz, *J. Phys. E (Sci. Instrum.)*, 3, 214 (1970).
- 4 - J. P. Waters "Interferometric Holography" en "Holographic Nondestructive Testing", editado por R. K. Erf. Academic Press, New York y London, 1974.
- 5 - E. Archbold, A. E. Ennos y P.A. Taylor. "Optical Instruments and Techniques". Pág. 265. Oriel Press, Newcastle-upon Tyne, 1970.
- 6 - K. A. Stetson, *Opt. Laser Technol.*, 2, 179 (1970).
- 7 - D. E. Duffy, *Appl. Optics*, 11, 1778 (1972).
- 8 - F. P. Chiang, Proc. Conf. on "Speckle Phenomena and their Applications", Loughborough University, Inglaterra (1974).
- 9 - B. P. Hildebrand y K. A. Haines, *JOSA*, 57, 155 (1967).
- 10 - P. K. Rastogui y L. Pflug, *Journal of Modern Optics*, 38, 1673 (1991).
- 11 - J. N. Butters y J. A. Leendertz, Proc. Electro Optics '74 Intern Conf., pág. 43, Brighton, Inglaterra (1974).
- 12 - J. N. Butters y J.A. Leendertz, *Measurement and Control*, 4, 349 (1971).
- 13 - A. Macovski, S. D. Ramsey y L. F. Schaefer, *Appl. Optics*, 10, 2722 (1971).
- 14 - C. Wykes, *Journal of Modern Optics*, 34, 539 (1987).
- 15 - K. Creath y G. Slettemoen, *JOSA*, 2, 1629 (1985).
- 16 - S. Nakadate, T. Yatagai y H. Saito, *Appl. Optics*, 19, 1879 (1980).
- 17 - A. E. Ennos. "Advances in Quantum Electronics". Pág. 199-258. Ac. Press, Londres (1970).
- 18 - J. N. Butters y J.A. Leendertz. Proc. Electro Optics '71 Intern. Conf.,

pág. 189, Brighton, Inglaterra (1971).

19 - C. Joenathan, B. Pfister y H. Tiziani, *Appl. Optics*, 29, 1905 (1990).

20 - N. Bolognini, H. Rabal y R. Torroba, *Appl. Optics*, 31, 1009 (1992).

21 - X. Peng, H. Y. Diao, Y. L. Zou y H. Tiziani, *Optik*, 90, 129 (1992).

21 - J. Pomarico, R. Arizaga, H. Rabal y R. Torroba. "Digital Holographic Contouring", *Optics Communications*, 98, 257 (1993).

23 - J. R. Varner "Holographic and Moiré Surface Contouring" en "Holographic Nondestructive Testing", editado por R. K. Erf. Academic Press, New York y London, 1974.

24 - Nils Abramson, *Opt. Letters*, 3, 121 (1978).

25 - F. Boden, H. D. Reidenbach y G. Köppen-Borgstedt, *Optik*, 51, 199 (1978).

26 - Recientemente (*Appl. Optics*, 31, 1682 (1992)), Y. N. Denisyuk y D. I. Staselko argumentan que el método fue desarrollado por ellos en 1969.

27 - H. Rabal, J. Pomarico y R. Arizaga. "Digital Holographic display of Light-in-Flight". Enviado para su publicación a *Applied Optics*. Marzo de 1993.

28 - J. Pomarico, R. Arizaga, R. Torroba y H. Rabal. "Algorithm to compute Spacing of Digital Speckle Correlation Fringes". Aceptado para su publicación en *Optik*. Abril de 1993.

29 - B. R. Frieden. "Computational Methods of Probability and Statistics", pág. 162 - 164 en "The computer in Optical Research", editado por B. R. Frieden, Springer Verlag, 1980.

30 - D. Malacara, Editor. "Optical shop testing", J. Wiley, New York, 1978.

31 - G. Van Bally, Editor. "Holography in Medicine and Biology", Springer Verlag, Berlin, 1979.

32 - R. K. Erf, Editor. "Holographic non-destructive testing", Academic

Press, New York and London, 1974.

33 - R. K. Erf, Editor. "Speckle Metrology". Academic Press, New York, San Fco. y London, 1978.

34 - J. Dyson, Appl. Optics, 2, 487 (1963).

35 - G. D. Dew, J. Scient. Instruments, 41, 160 (1964).

36 - J. J. Snyder, Appl. Optics, 19, 1223 (1980).

37 - P. Carré, Metrologia, 2, 13 (1966).

38 - J. Schwider, R. Brow, K. E. Elssner, J. Gryanna y R. Spolaczyk, Appl. Optics, 22, 3421 (1983).

MEDICIÓN DE TAMAÑO DE PARTÍCULAS POR MEDIOS ÓPTICOS

Dres. Lucía B. Scaffardi y Jorge O. Tocho

De todos los problemas a los que se enfrenta la industria moderna, la determinación precisa de las características del material particulado en un dado producto, ha tomado gran importancia. La industria farmacéutica depende de una detallada información sobre el tamaño y distribución de partículas en polvos y emulsiones para asegurar la calidad de sus compuestos. Asimismo, los fabricantes de semiconductores confían a los contadores de partículas el monitoreo de la pureza del aire dentro de sus áreas limpias de fabricación. Por otro lado, la industria de la pintura debe tener información acerca del tamaño de las partículas componentes para un mejor control de calidad de las mismas. En la industria alimenticia (fabricación de chocolates, harinas, etc.) el conocimiento del tamaño de partículas permite dotar a los alimentos de una consistencia y sabor óptimos. El conocimiento del tamaño de las partículas componentes en la industria cerámica permite sinterizar estos materiales sin defectos. La industria del cemento necesita conocer la distribución de tamaños y porcentajes de los distintos componentes en un dado cemento para asegurar la mejor calidad del mismo. Para la industria del cuero es imprescindible conocer el tamaño de las partículas componentes de las emulsiones con que son tratados los mismos con el objeto de mejorar la penetración en los poros del material y con esto asegurar una mejor calidad. La lista podría continuar aún. Estos son sólo algunos ejemplos que nos indican la gran importancia que tiene el conocimiento del tamaño de partículas y la potencial aplicación a un vastísimo campo tanto en áreas industriales como de investigación básica.

El estudio del "scattering" de luz debido a las partículas presentes en esos medios es el método más apropiado para la detección y caracterización del material particulado.

1. Conceptos de Scattering, absorción y extinción

La materia está compuesta de cargas eléctricas discretas: electrones y protones. Si un obstáculo (como por ej. un electrón simple, un átomo o una molécula, una partícula sólida o líquida) es iluminado por una onda electromagnética, las cargas eléctricas de aquél serán inducidas a realizar un movimiento oscilatorio debido al campo eléctrico de la onda incidente. Como sabemos, una carga eléctrica acelerada irradia energía en todas las direcciones. Esta radiación secundaria es la que denominamos **radiación dispersada** o **scattered** por el obstáculo (Fig.1). El término inglés "scattering" significa desparramamiento o dispersión y nos atreveremos a usar este vocablo de aquí en más para indicar el fenómeno descrito arriba:

Scattering = excitación + reirradiación

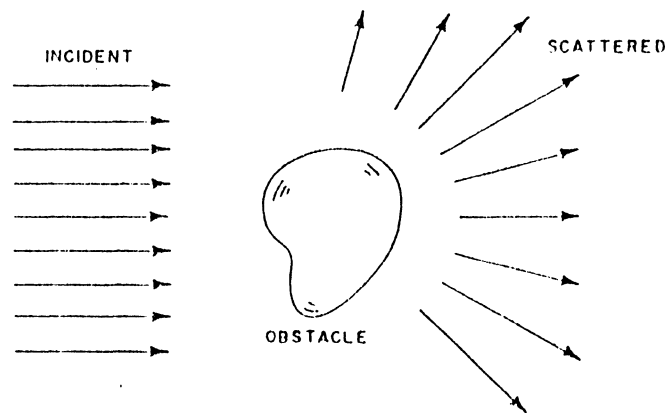


Fig. 1 Scattering por un objeto

Existe otro proceso que puede ocurrir cuando un objeto es alcanzado por una onda electromagnética. Las cargas excitadas pueden transformar parte de la energía incidente en otras formas (energía térmica, por ej.) por un proceso denominado **absorción**. De manera que si por un momento observamos la intensidad de la onda electromagnética a lo largo de su dirección de propagación, veremos que se habrá atenuado (extinción) luego de alcanzar el obstáculo debido a dos procesos: uno, el de scattering, que distribuye energía hacia otras direcciones distintas de la incidente, y otro, el de absorción, que transforma la energía incidente en calor:

Extinción = Scattering + Absorción

El color celeste del cielo, los colores del arco iris, los atardeceres rojizos, los colores de las rocas y el follaje, entre otros ejemplos, involucran complicadas interacciones de la luz con densos centros de scattering y absorción. Por ejemplo, el verde de las hojas se debe al scattering de esta zona del espectro visible mientras que la radiación roja es absorbida por las mismas. El humo negro presenta absorción en todo el espectro visible al igual que el caso del carbón.

Todos los medios, salvo el vacío, son heterogéneos en algún sentido. Aún en medios que usualmente son considerados homogéneos (como por ej. gases puros, líquidos o sólidos) es posible distinguir heterogeneidades individuales (átomos y moléculas) con un estudio suficientemente fino. De manera que todos los medios pueden dispersar luz. Muchos fenómenos que no son usualmente tratados como de scattering, son el resultado de esto último, por ejemplo reflexión difusa por superficies rugosas, difracción por ranuras, redes o bordes y reflexión y refracción especular por interfaces ópticamente suaves.

Existen numerosas técnicas destinadas a la medición del tamaño de partículas por medio del estudio del scattering óptico. Generalmente pueden dividirse en dos grandes

clases: las que operan con partículas simples o aisladas y aquellas que operan con un conjunto de partículas. Dentro de esta última clase, vamos a referirnos brevemente a las técnicas más importantes, de acuerdo a los rangos de tamaños involucrados, para detenernos más extensamente en las técnicas destinadas a medición de radios promedio.

Medición de tamaño promedio

En la zona comprendida entre 0.05 y 10 μm , las técnicas de medición de radio promedio parecen ser muy apropiadas, sobre todo en el caso de partículas poliméricas usadas en la industria del cuero; en estudios de partículas de látex vinculadas a la industria de la pintura; en el estudio de polvos superficiales (silicatos, sulfatos, calcitas, óxidos de hierro y minerales arcillosos); en el estudio de aerosoles estratosféricos (cenizas volcánicas, polvo cósmico) cuyos tamaños oscilan entre 0.05 y 5 μm . También son utilizadas para el estudio de partículas marinas en suspensión, como por ejemplo, sales y partículas orgánicas.

A continuación detallaremos tres técnicas interesantes: 1) Distribución Angular de Scattering (DAS), y dos técnicas basadas en la medición de extinción: 2) Atenuación (AT) y 3) Transmitancia Espectral (TS). Para simplificar el tratamiento, plantearemos algunas limitaciones importantes, como por ejemplo: a) consideraremos que el scattering es elástico, es decir, ocurre a la misma frecuencia que la luz incidente; b) es independiente, es decir, que si en un material hay M centros dispersores, la energía de scattering será M veces la energía dispersada por una única partícula y c) es único, condición que se cumple cuando el campo total de scattering es pequeño comparado con el campo incidente.

1- Distribución Angular de Scattering (DAS)

Este método está basado en la medición de la intensidad de luz dispersada para distintos ángulos cuando la muestra es iluminada por un haz de luz polarizado y monocromático. El arreglo experimental que se muestra en la Fig. 2 está conformado

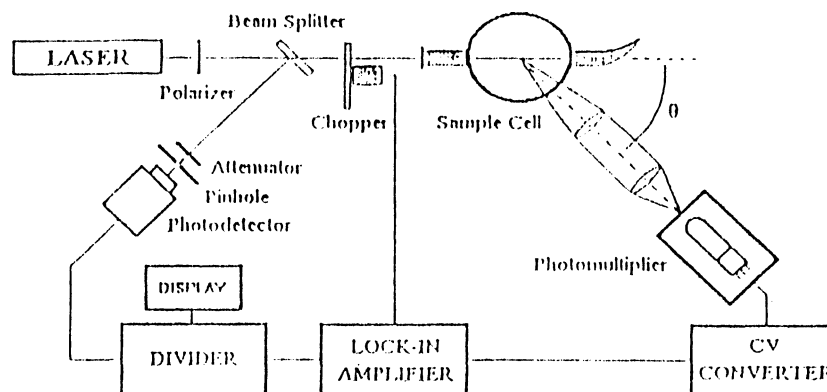


Fig. 2 Esquema experimental del método DAS

básicamente por un nefelómetro polar que contiene una fuente de luz colimada (láser de He-Ne, 5 mW, $\lambda = 632.8$ nm), una celda para la muestra y un detector montado en un brazo móvil que incluye un sistema óptico. Un polarizador permite iluminar la muestra con luz polarizada en diferentes planos. Un amplificador sensible a la fase (lock-in) y un chopper permiten trabajar bajo condiciones de iluminación ambiente.

El método consiste de dos series de mediciones angulares: en la primera, el patrón de scattering corresponde a luz incidente polarizada en forma paralela al plano de observación, y en la segunda, el patrón de scattering fue determinado para polarización perpendicular. El ajuste simultáneo de las dos series de medidas con las predicciones de la teoría de Mie permite calcular el radio promedio de las partículas con mucha precisión si se satisfacen las siguientes condiciones: uniformidad en el tamaño, forma esférica de las partículas y un cierto conocimiento del índice de refracción de las mismas.

Este método presenta la ventaja de emplear sólo una longitud de onda proveniente del láser de He-Ne. Se requiere una simple dilución del producto y no es necesario conocer la concentración de sólido en la muestra. Este método resulta ser el menos sensible a los cambios en la parte compleja del índice de refracción de las partículas.

Esta técnica implementada en el CIOp, fue utilizada para la medición de tamaño de partículas uniformes de látex de 89 nm de radio utilizadas como patrones. Los resultados experimentales pueden observarse en la Fig. 3 (a). Se muestra en línea llena el mejor ajuste teórico y en líneas de puntos las curvas de ajuste para radios con un 10% de desviación. Mediciones repetitivas dieron como resultado un radio de 89 ± 4 nm. La Fig.3 (b) muestra los resultados obtenidos con este método para polímeros comerciales utilizados en la industria del cuero (polímero de Bayer, 50-B). Uno de los más importantes factores que gobiernan las características físicas y mecánicas de las emulsiones utilizadas en la industria del cuero es el tamaño de las partículas. Se determinó un radio de 65 ± 5 nm, que coincidió con los resultados obtenidos con otros métodos.

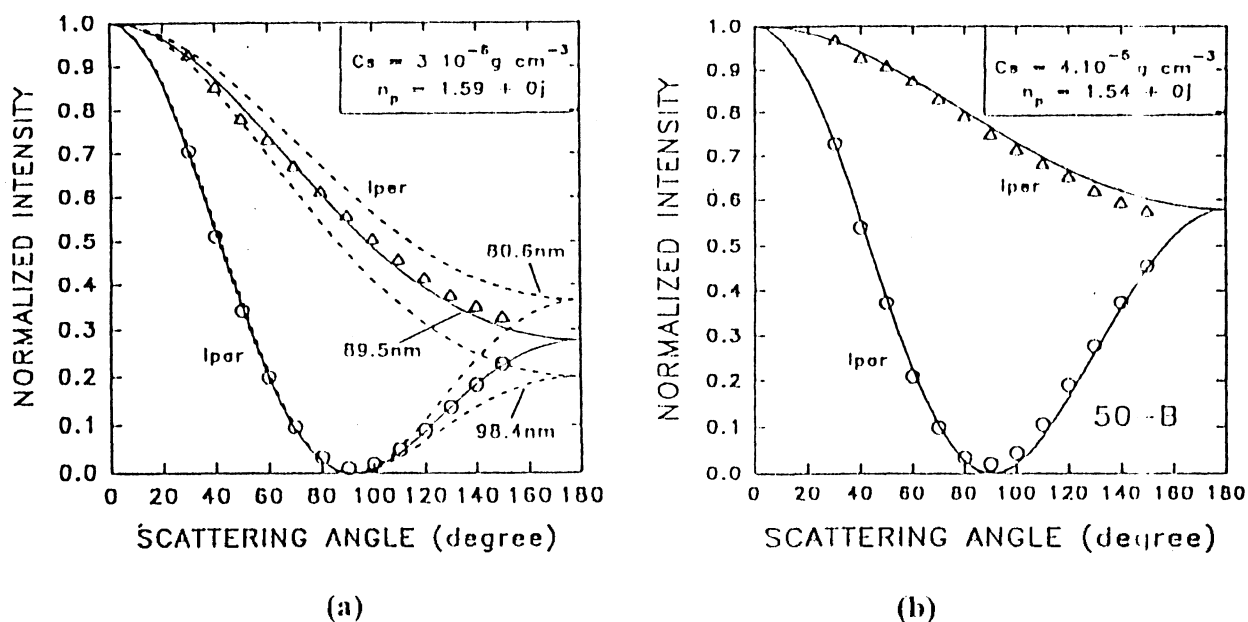


Fig. 3 Resultados con el método DAS

Métodos de extinción

La intensidad de un haz luminoso es exponencialmente atenuado a una distancia l en un medio particulado. Esta atenuación es causada por dos procesos simultáneos : absorción y scattering. Si N es el número de partículas por unidad de volumen, la ley de Lambert-Beer permite expresar la intensidad transmitida I_t como función de la intensidad incidente I_i y la absorbancia A ,

$$\begin{aligned} I_t &= I_i \exp (-A) \\ A &= N l C_{ext}, \\ C_{ext} &= C_{abs} + C_{scat} \end{aligned}$$

donde C_{ext} , C_{abs} y C_{scat} son las secciones eficaces para los efectos de extinción, absorción y scattering. Para esferas, la teoría de Mie permite calcular las secciones eficaces en términos del índice de refracción relativo [$m = (n' + i n'')_{part} / n_m$] y el parámetro de tamaño ($x = 2\pi r m / \lambda$) de la partícula. De esta forma se pueden conocer los valores exactos de C_{scat} y C_{abs} .

Una forma apropiada de conocer el tamaño de las partículas es a través de la comparación entre las predicciones teóricas y las medidas experimentales de la absorbancia. La incerteza en el conocimiento del valor de los parámetros ópticos tiene, en los métodos de extinción, un efecto más importante que en el de scattering angular. En particular, la parte imaginaria del índice de refracción tiene un efecto directo sobre el valor de C_{abs} y una influencia notoria sobre la absorbancia. Como dijimos antes, describiremos dos versiones del método de medida de extinción: atenuación (AT) y transmitancia espectral (TS).

2- Atenuación (AT)

El método AT está basado en la medición de la intensidad de luz monocromática no polarizada que atraviesa una muestra de longitud l en función de la concentración C_x . El dispositivo experimental (Fig. 4) consiste de una fuente de luz colimada (láser de He-Ne, 5 mW), una celda cilíndrica para colocar la muestra, un fotodetector para medir la

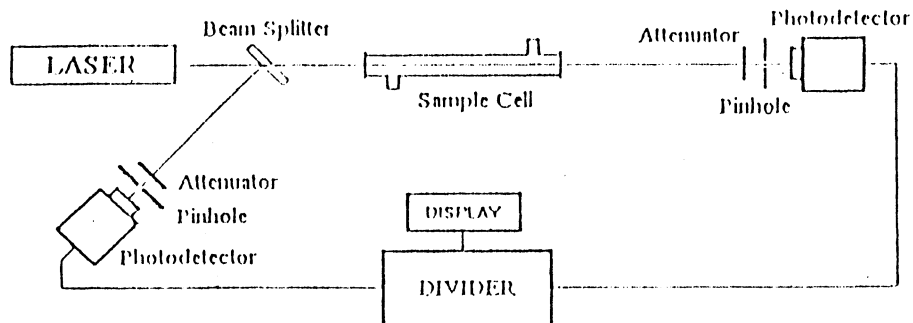


Fig. 4 Esquema experimental del método AT

intensidad de luz que atraviesa la muestra, un divisor de haz y otro fotodetector que toma una muestra de la intensidad incidente, para compensar variaciones de la fuente de luz.

El método consiste en la medición de la intensidad transmitida para varias concentraciones conocidas y el posterior ajuste de los resultados experimentales con una regresión lineal teórica, a partir de la cual puede ser determinado el radio promedio. Este método tiene la ventaja de emplear una sola longitud de onda y de trabajar con relativamente baja intensidad y fotodetectores muy simples.

Los resultados obtenidos con esta técnica implementada en el CIOp para concentraciones entre 2 y 15 $\mu\text{g}/\text{cm}^3$, se muestran en la Fig. 5. Las líneas punteadas indican un apartamiento de $\pm 5\%$ respecto del valor medio del radio, mostrando la sensibilidad del método. Para el caso de las partículas de látex usadas como patrones de medida, el óptimo ajuste (línea llena) muestra un resultado para el radio de 85 ± 2 nm.

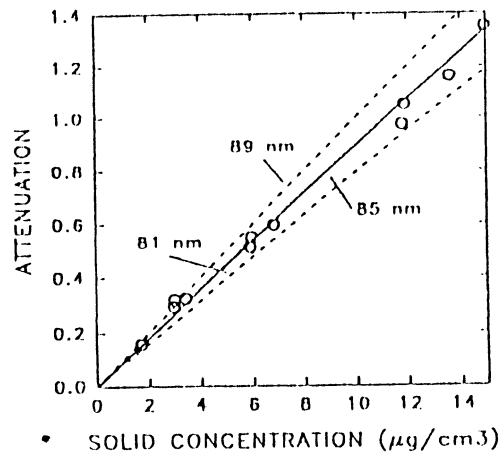


Fig. 5 Resultados con el método AT

3- Transmitancia Espectral (TS)

El método TS está basado en la medición de la intensidad de luz transmitida a través de la muestra para un determinado rango de longitudes de onda. Para implementar este método, se usó un espectrofotómetro comercial de simple haz, en la zona UV-visible (200 a 900 nm). La medición de la transmitancia sobre un rango específico de longitudes de onda fue llevado a cabo sobre una muestra de concentración conocida, y la curva resultante fue comparada con la calculada por medio de la teoría de Mic. El radio de las partículas fue determinado por el ajuste sobre todo el rango espectral completo.

Es interesante notar que, debido a que el método analiza la transmitancia sobre un rango espectral amplio, tanto el índice de refracción de las partículas como también el del solvente deben ser aproximadamente conocidos como función de la longitud de onda. El índice de refracción de las partículas puede determinarse usando un refractómetro a lámpara de sodio para varias diluciones y luego extrapolando a dilución cero. Este valor

particular a esta longitud de onda fue usado en la fórmula de dispersión para calcular los valores a otras longitudes de onda:

$$n(\lambda) = 1.59 + 15 \times 10^3 \left(\frac{1}{\lambda^2} - \frac{1}{589.32^2} \right)$$

Para el caso del solvente, en este caso agua:

$$n(\lambda) = 1.324 + 3046 / \lambda^2$$

La Fig. 6 muestra los resultados experimentales y el ajuste teórico para el método TS. La línea llena representa el mejor ajuste que corresponde a un valor del radio de 98 nm para las mismas partículas patrones de látex, mientras que las líneas de puntos muestran apartamientos de $\pm 10\%$ del dicho valor promedio. El inserto de la Fig. 6 muestra la foto obtenida con un microscopio electrónico de barrido de la muestra preparada por liofilización. Puede notarse claramente la esfericidad y la uniformidad en el tamaño de las partículas. El análisis, en este caso, dio por resultado un radio de 94 ± 5 nm.

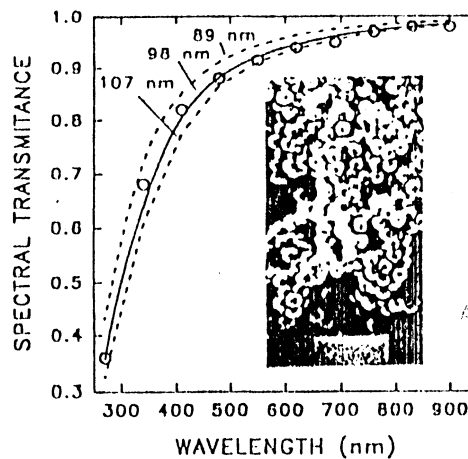


Fig. 6 Resultados del método TS

Como conclusión podemos decir que los tres métodos implementados para determinar radios promedio mostraron resultados muy satisfactorios. La medida del radio promedio era coincidente, dentro del error experimental, con el indicado por el fabricante o con el obtenido por microscopía electrónica. Los métodos de extinción han mostrado ser tan buenos como el de scattering en el caso especial donde la absorción es despreciable. Para un caso más general con materiales de absorción desconocida, el método DAS es el más apropiado.

Aunque la microscopía electrónica es la técnica más ampliamente usada para medición de radio de partículas, este método tiene algunas desventajas. Es un instrumento caro; la preparación de las muestras no siempre es posible debido a que pueden ocurrir fenómenos de coalescencia de las partículas y el procedimiento de conteo insume mucho tiempo.

El método DAS probó ser la más rápida y simple técnica de medición. Mostró ser el de menor sensibilidad a variaciones de las constantes ópticas del material. El método AT es válido sólo cuando se conocen las constantes ópticas del mismo.

Medición de distribuciones de tamaño

En la zona entre 0.05 y 3500 μm , de acuerdo a la aplicación, es importante el conocimiento de la distribución de tamaños, sobre todo en procesos industriales relacionados con la molienda y control de calidad, investigación y desarrollo en distintas áreas como por ejemplo: cerámica, arenas, muestras de suelos, resinas, plásticos, látex, productos farmacéuticos, cementos, arcillas, cuarzo, grafito, sales orgánicas e inorgánicas; en la fabricación de toner, pigmentos, jabón en polvo, abonos, etc.; en la industria alimenticia, en la fabricación de chocolates, azúcar, café, etc.

Dentro de las técnicas que miden distribución de tamaños hablaremos brevemente de 1) Medición de luz Dispersada a Diferentes Ángulos por una Distribución de Tamaños (DASDT), 2) Conteo Óptico de Partículas aisladas o en pequeñas concentraciones (COP) y 3) Photon Correlation Spectroscopy o PCS.

1) Medición de luz dispersada a diferentes ángulos por una distribución de tamaños (DASDT)

Se registra la intensidad dispersada por un conjunto de partículas presentes en la muestra, en diferentes ángulos. Dicho de otra manera, a cada lugar físico de observación le estará llegando luz dispersada proveniente de partículas de diferentes radios. Podemos resumir esto a través de las siguientes expresiones:

$$I_i = t_{i1} w_1 + t_{i2} w_2 + \dots + t_{iN} w_N,$$

donde I_i es la intensidad correspondiente al ángulo de observación i , t_{ij} es la intensidad teórica dispersada por una partícula de radio j y w_j es el porcentaje de partículas con ese radio. Para cualquier ángulo de observación, la expresión puede condensarse en una forma matricial,

$$(I_i) = (T_{ij}) (w_j)$$

El complejo proceso que sigue ahora es encontrar cuáles son los tamaños presentes y en qué proporción w_j . Dentro de los límites de validez de esta aproximación, se puede obtener la distribución de tamaño de las partículas a partir del conocimiento del patrón de scattering siempre que se cuente con un poderoso algoritmo capaz de invertir los datos, esto es hallar $(T_{ij})^{-1}$ tal que,

$$w_j = (T_{ij})^{-1} \cdot (I_i)$$

Existen comercialmente numerosos equipos capaces de efectuar este proceso satisfactoriamente. Algunos de los que aparecen hoy en el mercado utilizan la aproximación de Fraunhofer mientras que otros incorporaron la teoría de Mie para calcular la matriz teórica (T_{ij}).

Estos sistemas presentan la ventaja de la medición simultánea de un gran número de partículas de cada uno de los tamaños intervinientes. Esto es de gran importancia en el control de procesos industriales en los que es necesario analizar conjuntos estadísticos grandes. Como desventaja podemos decir que el proceso matemático de inversión no es sencillo porque la matriz de intensidad en función del radio y del ángulo de observación, es mal condicionada, hecho que acarrea grandes oscilaciones en los valores de w_j .

2) Contador Óptico de partículas (COP)

Este método fue originariamente desarrollado para el campo de análisis de contaminación de material particulado. Las partículas en suspensiones gaseosas o líquidas fluyen a través de una zona pequeña, una angosta región laminar uniformemente iluminada, producida por un láser de diodo o un bulbo incandescente (Fig. 7). Las partículas suspendidas están suficientemente diluidas como para que el pasaje a través de la zona iluminada sea de una a la vez, evitando así coincidencias por superposición.

El paso de una partícula a través de la zona activa da origen a una señal cuya altura depende del radio medio de la partícula y del principio de detección (scattering de luz u oscurecimiento por bloqueo). El sistema de iluminación/detección en el sensor está diseñado para proveer un aumento monótono en la altura del pulso a medida que aumenta el diámetro de las partículas.

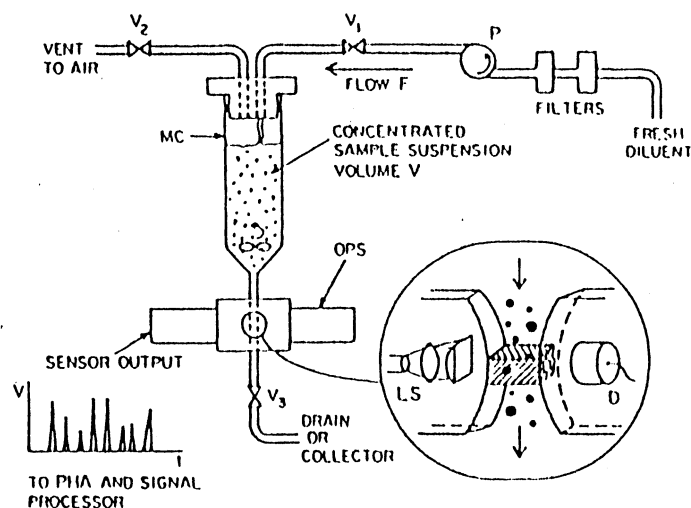


Fig.7 Contador óptico de partículas (COP)

La distribución de tamaño de las partículas se construye de a una por vez, comparando la altura del pulso detectado con una curva de calibración standard, obtenida a partir de un conjunto de partículas uniformes de diámetros conocidos.

El método COP es indispensable en el campo de análisis de contaminantes, como así también en la industria farmacéutica, en estudios de fluidos hidráulicos, etc.

Estos sistemas presentan la ventaja de que permiten la reconstrucción de distribuciones de tamaños complejas y su precisión no está limitada por el proceso posterior de análisis de datos pero también muestran una limitación: se necesita una alta velocidad de conteo para alcanzar suficiente precisión estadística en la determinación de la distribución de tamaños en tiempos razonables.

3- Photon Correlation Spectroscopy (P.C.S)

Dentro del rango comprendido entre 5-7 nm hasta 2-3 μm , la técnica más difundida es la de **Photon Correlation Spectroscopy (P.C.S) o Light Scattering Quasi-elástico**. Una suspensión relativamente diluida de partículas es iluminada por un haz láser, resultando luz dispersada a partir de cada partícula. La información sobre las dimensiones de las mismas viene dada por el análisis de las fluctuaciones temporales de la intensidad luminosa difusa.

Tales fluctuaciones provienen generalmente de la interferencia de la luz difusa de cada partícula individual que se mueve de manera Browniana en el interior del fluido en el cual está suspendida. El análisis temporal de las fluctuaciones mediante técnicas de autocorrelación consiste en determinar el coeficiente de difusión traslacional D_T de las partículas.

En el caso de partículas esféricas de diámetro d , el coeficiente de difusión traslacional viene dado por la ecuación de Stokes-Einstein $D_T = K_B T / 3\pi\eta d$, donde K_B es la constante de Boltzmann, T es la temperatura y η la viscosidad del fluido en el cual está suspendida la partícula. Por lo tanto, conociendo T y η , es posible determinar d sin conocer la densidad o el índice de refracción de la partícula.

Esta técnica presenta varias ventajas: es rápida y de fácil operación. No produce perturbación de la suspensión de la muestra y además es independiente de una calibración previa. La distribución de tamaño de partículas es independiente de la composición y de la concentración (siempre que la suspensión esté dentro de los límites de scattering simple e independiente).

Esta técnica se utiliza en sistemas comerciales de interés como látex y otros polímeros sintéticos, cerámicos, colorantes, pigmentos, sílicas coloidales, emulsiones aceite-agua, liposomas, etc. Por debajo de 5 nm se estudian sistemas de mucho interés como micelas surfactantes, (iónicas y no iónicas), microemulsiones de agua en aceite, sílicas coloidales ultrafinas, proteínas y otras moléculas de interés biológico.

Bibliografía

- 1- C. F. Bohren and D.R. Huffman, *Absorption and Scattering of Light by Small Particles*, John Wiley & Sons, New York (1983)
- 2- H. C. van der Hulst, *Light Scattering by Small Particles*, Dover, New York (1981)
- 3- L. Yebrin, J. Tocho and L. Scaffardi, "Determinación de tamaño de partículas poliméricas por distintas técnicas ópticas," *Anales Asociación Física Argentina*, vol. 4, 132 (1992)
- 4- L. Scaffardi, L. Yebrin, J. Tocho, and C. Cantera, "Medición del radio de partículas por scattering óptico", *Anales Asociación Física Argentina*, vol. 5, 262-266 (1993)
- 5- L. B. Scaffardi, J. O. Tocho, L. L. Yebrin and C. S. Cantera, "Sizing particles used in the leather industry by light scattering", in *Optical Engineering*, vol. 35, No 1, 52 (1996)