

LASER Y OPTICA EN INGENIERIA

CIOP

AÑO 1989

© 1989

CIOp

• CIC

CENTRO DE INVESTIGACIONES OPTICAS

Cno. Parque Centenario entre 505 y 508 Gonnat - Tel. 84-0280 y 84-2957
Direc. Postal: C. de Correo 124, Telex 31216 CESLA, 1900 La Plata Rep. Arg.

Hecho el depósito de ley

**IMPRESO EN ARGENTINA
PRINTED IN ARGENTINA**

ISBN

PROLOGO

La presente obra es uno de los resultados de las tareas de investigación científica y tecnológica, de docencia superior y de servicios calificados a empresas públicas y privadas, realizadas por el Centro de Investigaciones Opticas (CIOp) desde su constitución, hace doce años, y que han sido volcadas durante siete años consecutivos en el Curso de Actualización y Perfeccionamiento LASER Y OPTICA EN INGENIERIA (LOI), dirigido por el Prof. Dr. Mario Gallardo.

Está dirigida particularmente a aquellos profesionales de las ramas de Ingeniería y Ciencias, empresarios y demás lectores interesados en las aplicaciones industriales de la óptica y el láser. No pretende abarcar todas ellas; sea por razones de espacio o por haberse seleccionado sólo aquellas de mayor desarrollo relativo dentro de los Programas del CIOp, tales como las referidas a macro y micrometrología, a tratamiento de materiales y a comunicaciones atmosféricas. Por otra parte, no se incluyen las numerosas aplicaciones de la óptica y el láser a otras ciencias, como biología, medicina, antropología, etc. u otras disciplinas, como las artes, el diseño y la propaganda -en las que el CIOp ha contribuido y lo sigue haciendo- por considerarlas ajenas a la finalidad arriba enunciada.

El impacto producido por los tremendos avances de la óptica y el láser en el ámbito de la ingeniería es difícil de mensurar. El CIOp trata de seguir esos avances de cerca. Por ello, la decisión de editar esta obra representa un serio compromiso; por un lado establecer una base sólida que sirva de fundamento a las aplicaciones industriales de la óptica y el láser y por el otro dar cabida a algunas aplicaciones selectas. Al revisar los textos impresos de los cursos LOI precedentes, creemos estar convencidos de haber cumplido con aquel compromiso por cuanto el que ahora sale a la luz es mucho más amplio y profundo. Sin embargo, ¡cuánto falta por hacer! ...: somos conscientes de que cumplir el compromiso implica trabajar mucho más. Invito a los lectores de esta obra a hacernos llegar sus sugerencias y requerimientos sobre los aspectos básicos y aplicados de la óptica y el láser para estudiarlos y considerarlos, en futuras ediciones.

Esta obra incluye la colaboración de muchos investigadores y profesionales. De ellos queda constancia capítulo a capítulo. Deseo dejar también constancia del trabajo esmerado de otros colaboradores, en particular de Norma E. Díaz, Marcela Tomeo, Arq. Viviana Schaposnik, Cdor. Juan N. Cap y Jorge M. Willemoës, Secretario Ejecutivo del Curso LOI.

Durante el transcurso de los siete últimos años han contribuido con su auspicio o apoyo a esta labor las siguientes instituciones:

- Secretaría de Ciencia y Técnica de la Nación (SECYT)
- Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET)

- Consejo de Rectores de Universidades Nacionales (CRUN)
- Universidad Nacional de La Plata (UNLP)
- Comisión de Investigaciones Científicas de la Provincia de Buenos Aires (CIC).

Finalmente, agradezco a la **Fundación Antorchas** el generoso apoyo brindado para lograr la presente edición.

La Plata, agosto de 1989.

Prof. Dr. Mario Garavaglia
Director
Centro de Investigaciones Opticas

CONTENIDO

-EL LASER.	
Prof. Dr. Mario Garavaglia	1
-ALGUNOS TIPOS DE LASERES.	
Prof. Dr. Héctor F. Ranea Sandoval y Dr. Lucía B. Scaffardi	91
-ESPECTROSCOPIA ATOMICA Y MOLECULAR.	
Prof. Dr. Jorge G. Reyna Almandos	129
-OPTICA FISICA. COHERENCIA. INTERFERENCIA. DIFRACCION.	
Prof. Dr. Lía M. Zerbino	151
-ANALISIS DE FOURIER.	
Prof. Lic. Beatriz N. Ruiz	199
-HOLOGRAFIA.	
Dr. Roberto D. Torroba	221
-METROLOGIA SPECKLE.	
Dr. Enrique E. Sicre	243
-CRISTALES FOTORREFRACTIVOS BSO.	
Prof. Dr. Néstor Bolognini	265
-ALINEACION, POSICIONAMIENTO, CONTROL Y MEDIDAS.	
Prof. Dr. Mario Gallardo	281
-EL LASER EN LA MEDICION DE VELOCIDADES.	
Prof. Dr. Jorge O. Tocho	311
-APLICACIONES DE LASERES EN PROCESOS QUIMICOS.	
Prof. Dr. Jorge O. Tocho	333
-EL LASER EN ESTUDIO DE LLAMAS.	
Prof. Dr. Jorge O. Tocho	341
-PROCESAMIENTO DE MATERIALES CON LASER DE CO₂.	
Dr. Eliseo Gallego Lluesma	345
-DETECTORES DE RADIACION VISIBLE E INFRARROJA.	
Ing. Aníbal P. Laquidara	451
-COMUNICACIONES ATMOSFERICAS CON LASER.	
Ing. Aníbal P. Laquidara	481
-OTRAS APLICACIONES.	
Dr. Daniel C. Schinca	497

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.

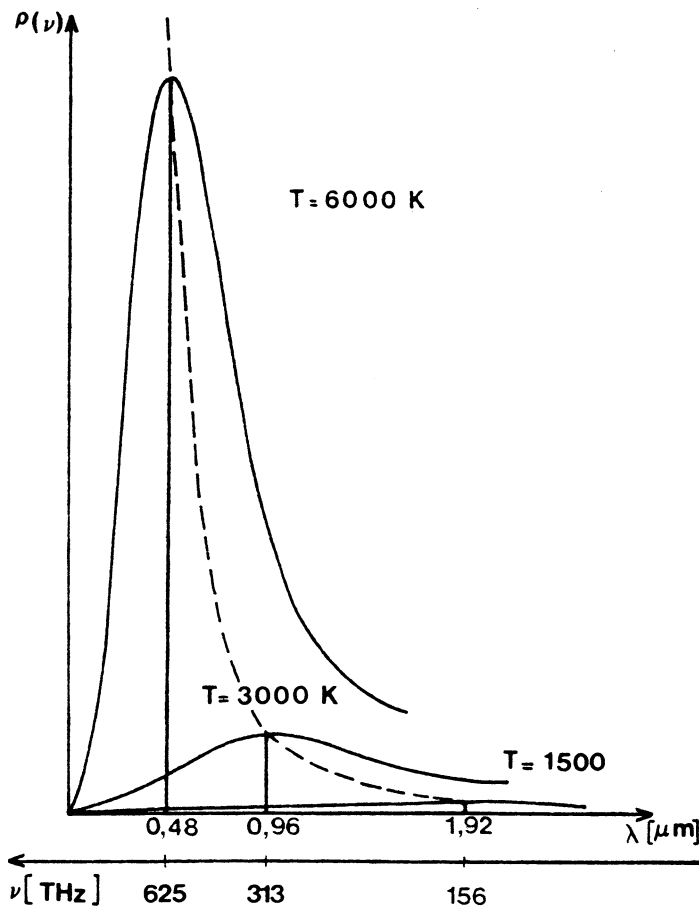


Figura 1. Densidad de radiación $\rho(\nu)$ emitida por el cuerpo negro según la temperatura T y la ley de Wien.

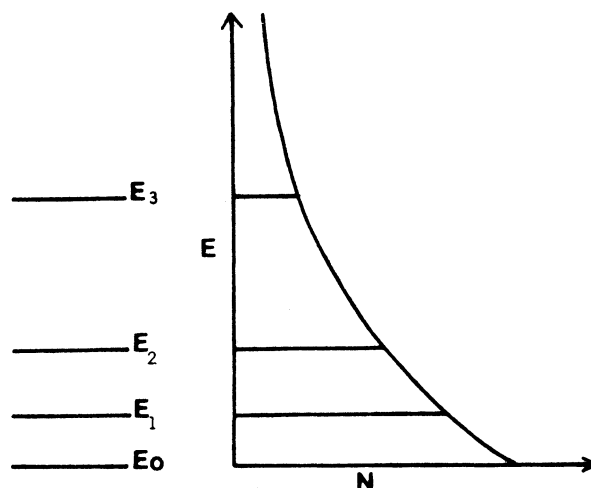


Figura 2. Ley de Boltzmann de la distribución de población.

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

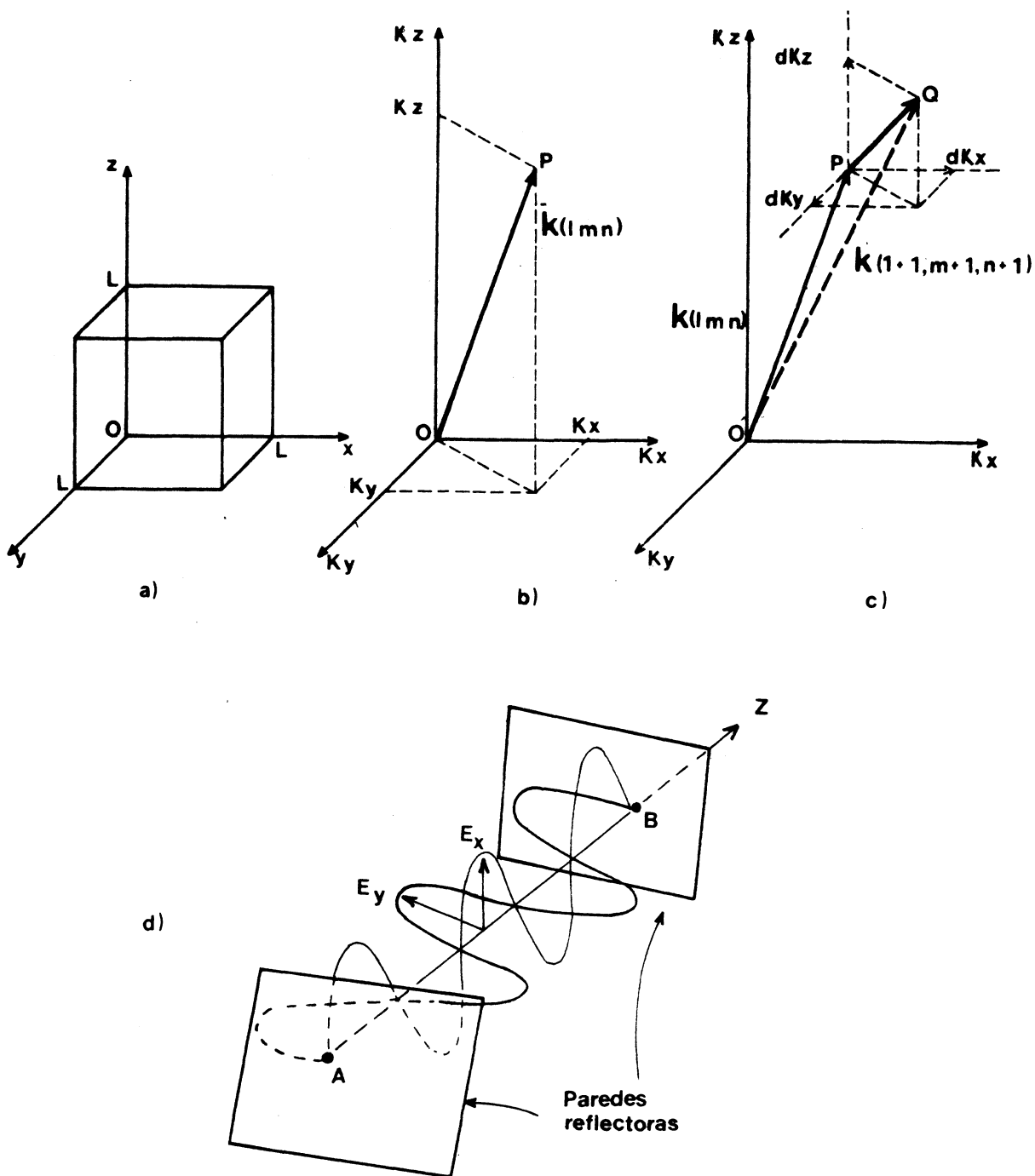


Figura 3. Vínculos entre la geometría de la cavidad y las ondas del campo de radiación. a) Cavidad cúbica de lado L . b) Vector propagación de la onda que excita al modo resonante (l, m, n) . c) Cambios en la longitud de onda de la radiación provocan la excitación del modo $(l+1, m+1, n+1)$. d) Las ondas 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:

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

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:

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

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 despreciarse 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.

3.1. Emisión estimulada y dispersión.

En el caso de la dispersión, R. Ladenburg³ y H. Kopferman⁴ hicieron uso de la expresión:

$$n^2 - 1 = \frac{e^2}{\pi m} \frac{(N_1 f_{21} - N_2 f_{12})}{\nu_{21}^2 - \nu^2}, \quad (16)$$

donde n es el índice de refracción del medio óptico en las cercanías de la transición $2 \rightarrow 1$ entre los niveles de energía E_2 y E_1 , cuyas poblaciones son, respectivamente, N_2 y N_1 . En la misma fórmula, e es la carga del electrón y m su masa, y f_{21} mide la efectividad de la transición $2 \rightarrow 1$. El coeficiente f de efectividad de la transición electrónica está íntimamente relacionado con los coeficientes de Einstein que expresan las probabilidades de la misma transición mediante los procesos de absorción, de emisión espontánea y de emisión estimulada. El coeficiente f fue introducido por los espectroscopistas antes del desarrollo de la mecánica cuántica. La importancia y trascendencia del coeficiente f aparece realmente a partir de la Regla de la Suma de los f , formulada independientemente por W. Kuhn⁵ y W. Thomas⁶. Esta regla de la suma, de origen empírico, establece que si se suman los valores de f correspondientes a todas las posibles transiciones electrónicas entre un dado nivel inicial y todos los posibles niveles finales de energía, considerando tanto las transiciones de emisión cuanto las de absorción de radiación, el resultado será igual al número total de los electrones que posee el átomo cuyo espectro se estudia. A partir de la validez de esta regla, bien establecida por la espectroscopía, W. Heisenberg⁷ en su primer trabajo sobre la me-

cánica de las matrices establece la regla de conmutación entre la coordenada y el impulso del electrón, como postulado fundamental.

La relación entre los coeficientes f y B correspondientes a la transición electrónica $2 \rightarrow 1$ es:

$$B_{21} = \frac{\pi e^2}{h \nu_{21} m} \cdot f_{21} \quad .$$

Volviendo a la expresión (16) de la dispersión, si N_2 es mayor que N_1 , el signo de la dispersión cambia y se podrá obtener dispersión negativa, concepto ignorado en la física clásica. En la ecuación (16) el término $N_2 f_{21}$ causante de la dispersión negativa, es proporcional a la población del nivel de energía superior N_2 .

El campo de radiación en el entorno de la frecuencia ν_{21} provocará una interacción con los átomos, la que puede forzar la transición $2 \rightarrow 1$, y como $E_2 > E_1$, el resultado será la emisión de radiación estimulada. Si este efecto resulta ser preponderante, debe observarse en las medidas de dispersión.

El término negativo $N_2 f_{21}$ significa pérdida de energía por la emisión estimulada que los átomos excitados ceden al campo de radiación. La frecuencia de la radiación emitida será igual a la del campo que produce la perturbación sobre los átomos excitados.

En la expresión (16) de la dispersión no aparece ningún factor relacionado con las fases de las radiaciones estimulante y estimulada. Ello equivale a afirmar que la radiación estimulada está en fase con la estimulante, o sea, que el campo de la radiación resultante será coherente. Si así no fuera, no habría contribución en la fórmula de la dispersión y su observación resultaría imposible.

Los experimentos de Ladenburg y Kopfermann, basados en la ecuación (16), consistieron esencialmente en la medida del índice de refracción de un gas en el entorno de una línea espectral. El instrumento por ellos empleado fue el interferómetro de Jamin, y el montaje del experimento se muestra en la Figura 4. La luz producida por una lámpara de filamento intensa pasaba por dos tubos de iguales dimensiones, y tras ser recombinada, entraba a un espectrógrafo. La figura de interferencia formada en la recombinación de los haces luminosos y analizada respecto a cada longitud de onda mediante el espectrógrafo, consiste en un sistema de franjas. Cualquier variación relativa de las longitudes de los caminos ópticos 1 y 2, se verá reflejada con precisión en la variación de la forma y disposición de las franjas.

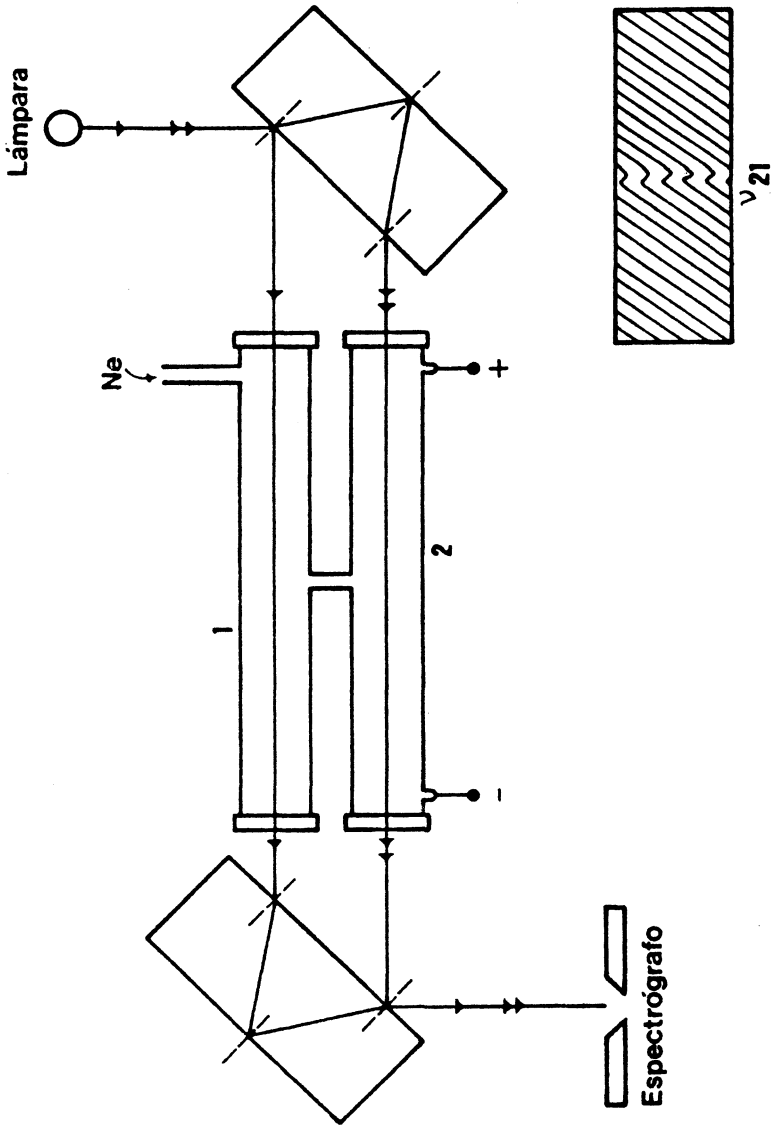


Figura 4. Experiencia de dispersión anómala para detectar la influencia de la emisión estimulada.

En este sentido, el interferómetro de Jamin actúa como un refractómetro diferencial. En los dos tubos se hizo el vacío convenientemente y se llenaron con neón a baja presión. Ambos estaban conectados, de modo que la presión del gas era la misma. Sin embargo, en uno de ellos el gas fue excitado mediante el paso de una corriente eléctrica susceptible de variación. De este modo, en los entornos de las líneas espectrales se pudo medir la diferencia de índices de refracción y obtener, a partir de la ecuación (16) el valor del factor $f_{21}(N_1-N_2)$. Si la excitación es débil, N_1 disminuirá con respecto a su valor cuando la excitación es nula, pero siempre se mantendrá $N_2 \ll N_1$, con lo cual se puede medir $f_{21} N_1$. Al incrementar la corriente de la descarga excitadora, los niveles de energía bajos se pueblan más rápidamente y crece N_2 mucho más que N_1 y, por lo tanto, decrece la dispersión en el entorno de la línea espectral ν_{21} correspondiente a la transición entre los niveles $2 \rightarrow 1$.

El resultado completo del experimento se ha representado en la Figura 5a. El producto $f_{21}(N_1-N_2)$ en función de la corriente de excitación, crece al principio y tras alcanzar un máximo decrece lentamente. La manera en que la dispersión o, lo que es lo mismo, el valor del factor $f_{21}(N_1-N_2)$ varía con la excitación, puede deberse a un aumento relativo de N_2 respecto de N_1 , o también a una disminución de N_1 . Para decidir entre estas dos posibilidades, los mencionados investigadores estudiaron la variación de la dispersión en los entornos de las cuatro líneas espectrales del neón que tienen el mismo nivel de energía terminal E_1 , como muestra la Figura 5b. En tal caso, si la variación de la dispersión se debiera a la disminución de la población del nivel 1 y no al aumento relativo de N_2 respecto de N_1 , se observaría idéntico cambio en la dispersión en todas las líneas espectrales consideradas.

Sin embargo, notaron en las cuatro líneas espectrales distinta relación funcional entre la variación de la dispersión y la excitación, aunque siempre aquella aumentaba rápidamente al comienzo y después de alcanzar el máximo decrecía lentamente, según muestra la Figura 5a.

Quedó comprobada así la existencia de la dispersión negativa y, por lo tanto, se evidenció por primera vez la emisión estimulada de radiación.

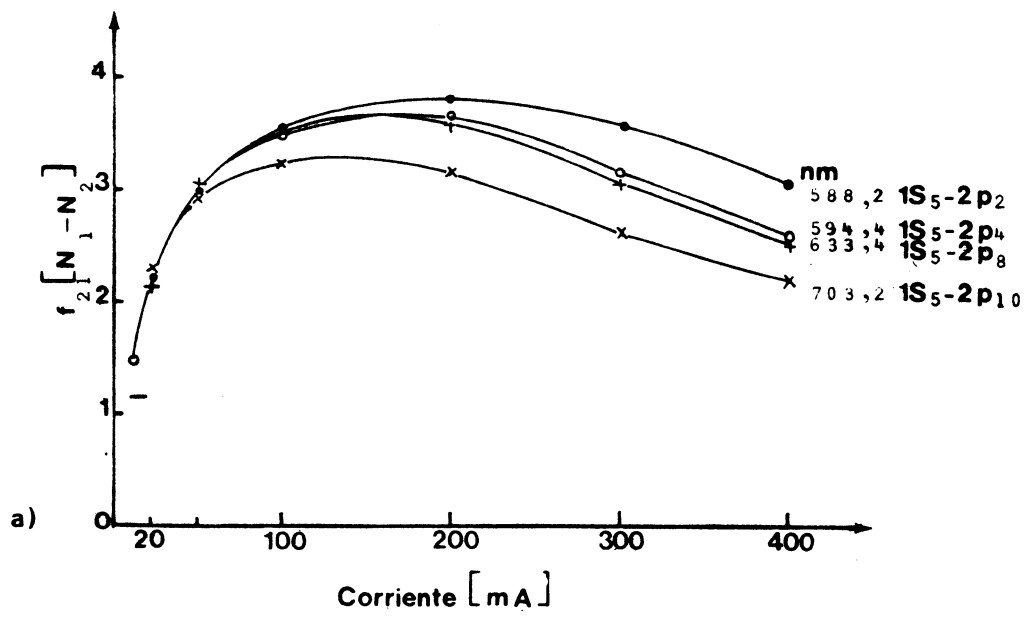


Figura 5a. Efecto de la emisión estimulada en cuatro transiciones del neón.

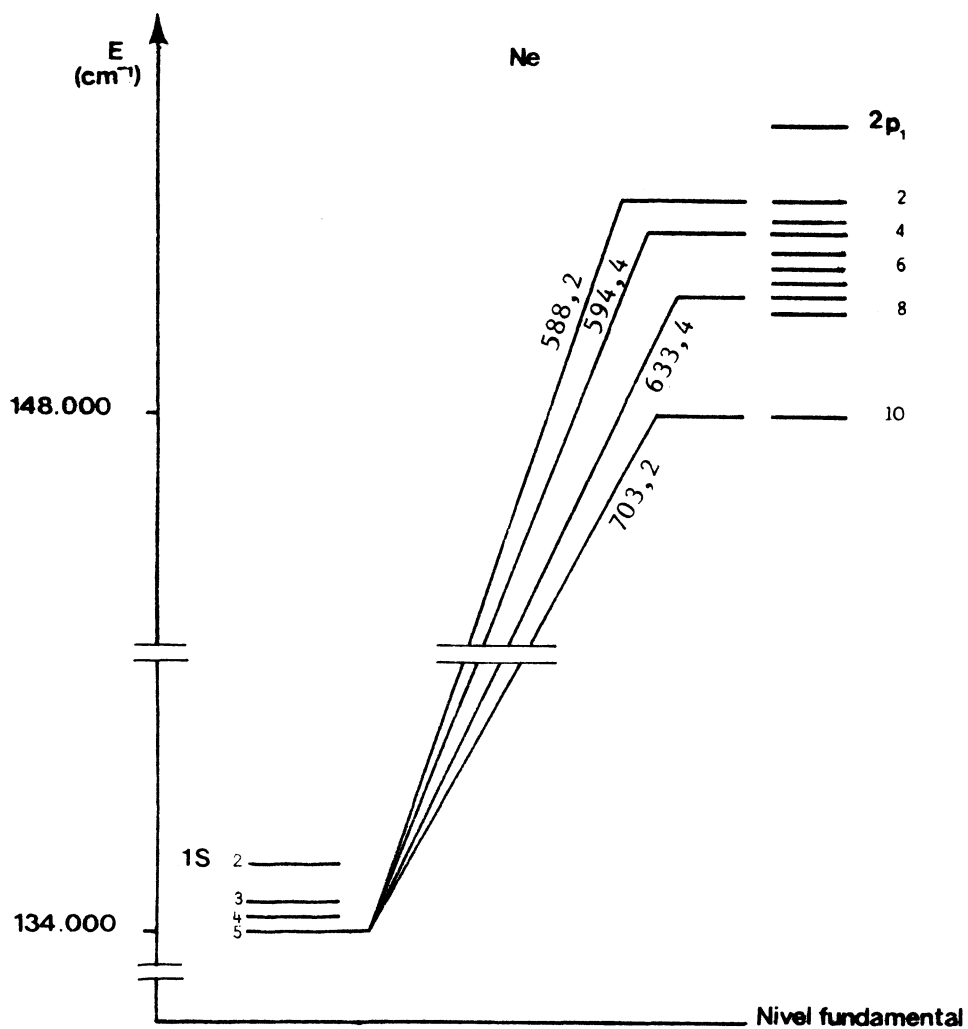


Figura 5b. Las cuatro transiciones del neón.

3.2. Emisión estimulada y absorción.

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:

$$I(x, \nu_0) = I(0, \nu_0) e^{-\alpha(\nu_0) \cdot x}, \quad (17)$$

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 $I(x, \nu_0)$.

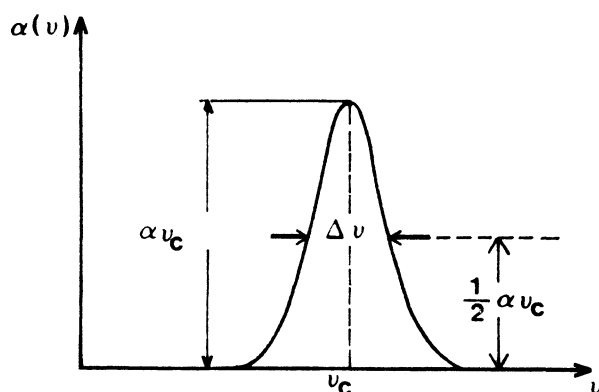


Figura 6. Absorción a través de una línea espectral de frecuencia central ν_c .

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:

$$\int \alpha(\nu) d\nu = h\nu \cdot B_{12}(N_1 - N_2) / c \quad . \quad (18)$$

La Figura 6 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:

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

De esta manera, el término de la derecha de la ecuación (18) 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:

$$\frac{g_1}{g_2} \cdot \frac{N_2}{N_1} = e^{-(E_2-E_1)/kT} \quad . \quad (20)$$

Como consecuencia, las expresiones previas de la dispersión (16) y la absorción (18) pueden completarse en caso de considerar niveles de multiplicidad mayor que la unidad. Más adelante, al tratar el láser de rubí, se discutirá cómo influye el grado de degeneración en las condiciones de excitación.

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 7 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.

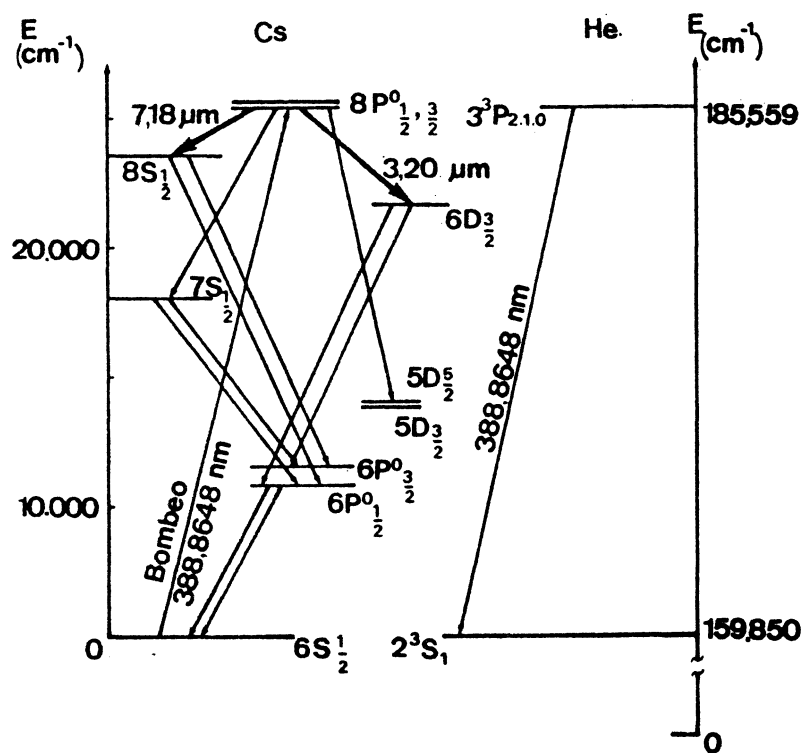


Figura 7. a) Excitación del cesio por bombeo óptico con radiación del helio.

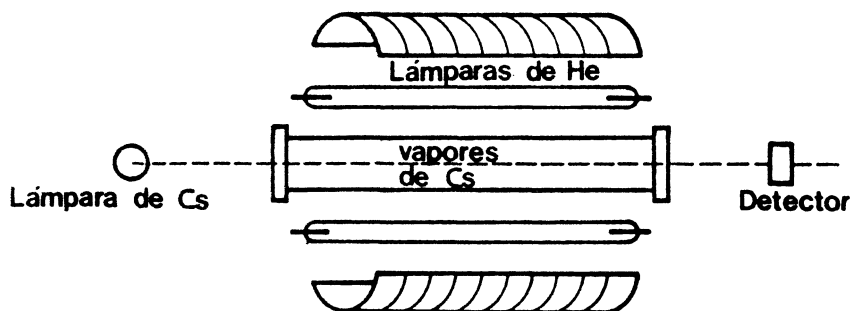


Figura 7. b) Experiencia propuesta por Fabrikant para amplificar radiación.

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 8 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.

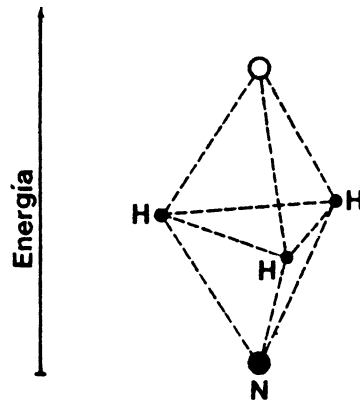


Figura 8. a) Esquema de la molécula de amoníaco.

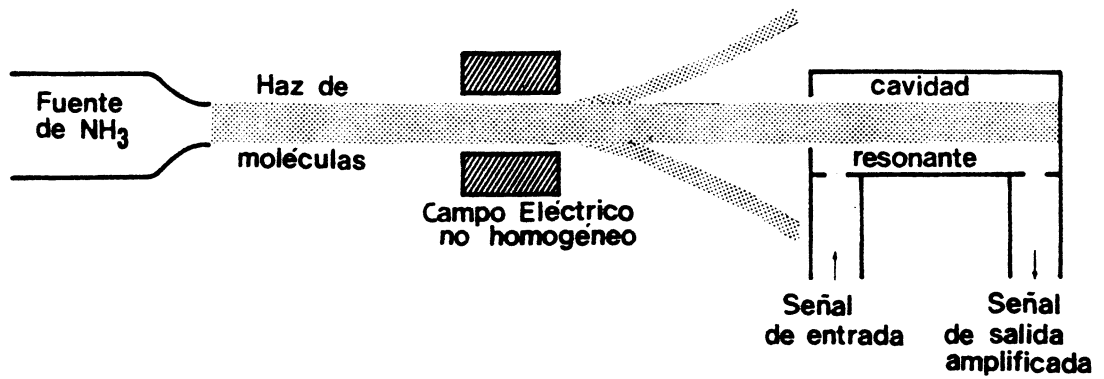


Figura 8. b) Máser de amoníaco.

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:

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

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 másers 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 másers 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 (21), es:

$$\delta\nu = 4\pi h\nu (\Delta\nu)^2 / P \quad . \quad (22)$$

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 (22) 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 9.

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.

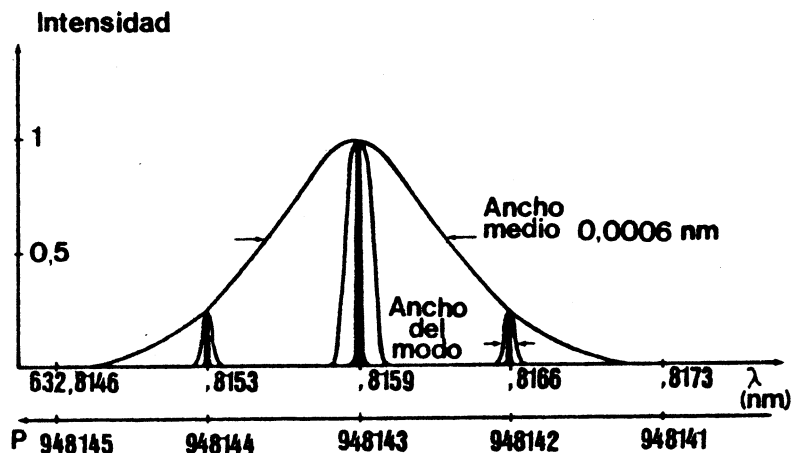


Figura 9. Modos longitudinales de resonancia bajo el perfil de línea espectral de ancho medio 0,0006 nm correspondiente a la transición de $\lambda=633$ nm del Ne. La representación de la estructura modal es la de un láser de He-Ne cuya intensidad tiene una longitud $D=30$ cm.

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° 2929922 ¹⁸ 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 logran 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 (19). 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 10 muestra el esquema de niveles del ión triple de cromo en óxido de aluminio.

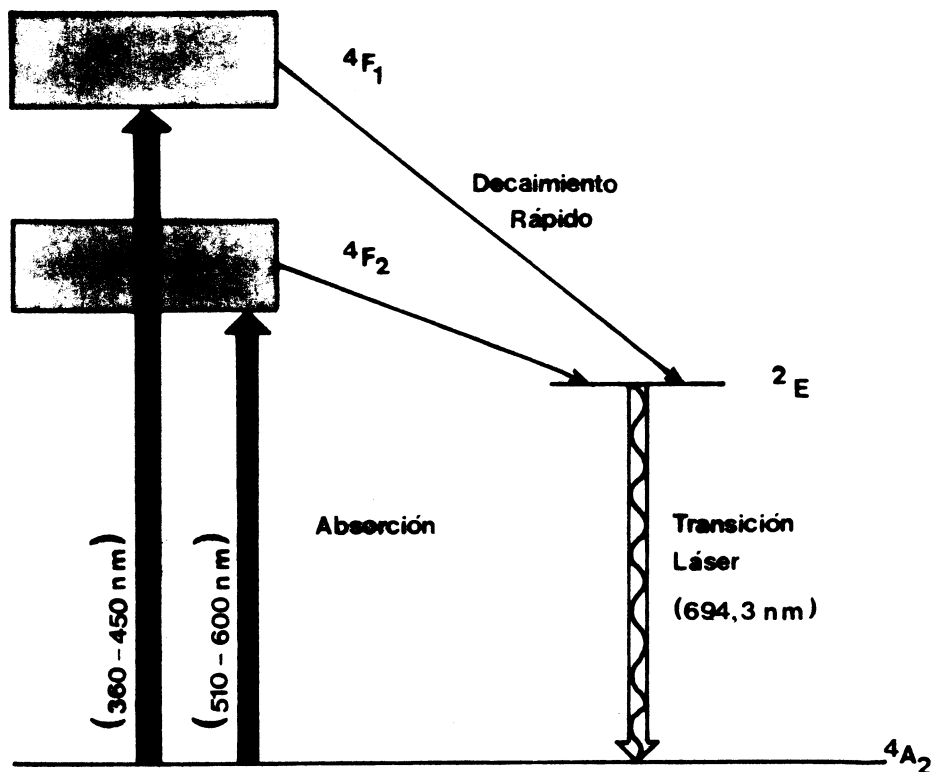


Figura 10. Esquema simplificado de los niveles de energía del ión Cr^{3+} en óxido de aluminio.

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 $1,1 \mu\text{m}$.

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.

Así pues, las técnicas de bombeo dependen de los materiales, pero las condiciones en que el bombeo debe producirse son generales y, por lo mismo, sujetas a sistematización. El análisis que sigue sólo pretende ser una introducción al tema. No obstante, por abarcar todos los elementos básicos que atañen al bombeo, se llegará a la enumeración de condiciones esenciales para el funcionamiento de todo láser. El tratamiento de las condiciones de bombeo se dividirá según los aspectos relativos al equilibrio dinámico del proceso, donde la variable tiempo no juega un papel preponderante, y a la dinámica de la inversión de población, donde la variable tiempo es capital. De todos modos, el resultado del análisis permitirá, en general, el diseño de los láseres, por cuanto se presta atención a las magnitudes físicas que caracterizan los materiales que constituyen los medios activos, así como a los aspectos instrumentales y al dimensionamiento de sus diversas partes.

9.2. Caracterización del medio material activo.

Como ya se vio, el planteo de Einstein sobre las condiciones de equilibrio dinámico entre los niveles de energía 1 y 2, se expresa por:

$$N_1 \rho(\nu_{21})B_{12} = N_2 A_{21} + N_2 \rho(\nu_{21})B_{21} \quad . \quad (23)$$

Esta igualdad indica que la energía transferida desde el nivel 1 al 2 por absorción de radiación es igual a la energía transferida desde el nivel 2 al 1 por los procesos de emisión espontánea y de emisión estimulada, cuando las poblaciones N_1 y N_2 están sujetas a la ley de Boltzmann.

Para facilitar la escritura se harán algunas concesiones.

Como los coeficientes B_{12} y B_{21} son iguales, en adelante se los designará como el coeficiente B . Además, siempre se hará referencia a las transiciones de absorción y de emisión entre los niveles 1 y 2, por lo que también se eliminarán los subíndices en la expresión de la frecuencia respectiva. De este modo ν reemplazará a ν_{21} . Por las mismas razones se indicará la probabilidad de emisión espontánea A_{21} , por A .

Designando:

$$W_i' = \rho(\nu)B \quad , \quad (24)$$

la expresión (23) resulta:

$$N_1 W_i' = N_2 A + N_2 W_i' \quad , \quad (25)$$

de modo que se puede definir W_i' como la probabilidad de transición inducida, asumiendo un papel equivalente al de A , por haber incorporado en su expresión a la densidad de radiación $\rho(\nu)$.

Teniendo en cuenta la relación mutua entre los coeficientes de Einstein A y B , expresada según la (5), se puede a su vez relacionar A y W_i' , quedando:

$$\frac{W_i'}{A} = \frac{c^3}{8\pi h \nu^3} \rho(\nu) \quad . \quad (26)$$

Las transiciones de absorción o de emisión tienen una banda de frecuencia tan estrecha que puede suponerse que el valor de la densidad de radiación $\rho(\nu)$ en toda ella apenas varía, de modo que dicho valor podría suponerse constante. Sin embargo, como el ancho de línea emitida por el láser es mucho menor que el de la línea correspondiente a la misma transición emitida espontáneamente o absorbida, será imprescindible considerar la manera de comportarse de $\rho(\nu)$ en tal intervalo de frecuencias.

En todo tratamiento teórico sobre el láser debe considerarse a la radiación emitida mediante el proceso de emisión inducida por los átomos, moléculas o iones como poseyendo una elevadísima pureza espectral. Por ello, para obtener una expresión cuantitativa de W_i' será necesario tener en cuenta el perfil de distribución de la energía emitida espontáneamente. Los caminos para introducir este perfil de línea espectral en el tratamiento teórico sobre el láser son variados y dependen básicamente de la economía de trabajo que producen en los cálculos. Uno de ellos indica que si se expresa el perfil de línea como la función $g(\nu)$, el producto $g(\nu)d\nu$ es igual a la probabilidad de que los átomos, moléculas o iones emitan o absorban, en la transición correspondiente, un fotón de energía comprendida entre $h\nu$ y $h(\nu+d\nu)$. De esta definición se obtiene inmediatamente la condición de normalización:

$$\int_0^{\infty} g(\nu) d\nu = 1 \quad , \quad (27)$$

la cual señala que hay una probabilidad total, es decir igual a uno, de que en todo el intervalo de frecuencias del espectro de radiación, desde cero hasta infinito, sea emitida o absorbida energía con una distribución espectral $g(\nu)$. Con esto, la representación gráfica de $g(\nu)$ en función de la frecuencia ν es idéntica a la que puede registrarse de la intensidad de la línea espectral emitida o absorbida en función de ν , utilizando un aparato apropiado.

El origen de la distribución de intensidad o perfil de línea espectral puede ser diverso y, en general, concurren simultáneamente varios efectos. En el caso de los gases excitados a baja presión, el perfil de línea se debe de modo principal al efecto del corrimiento Doppler. En cambio, en los sólidos, tal ensanchamiento es causado por los desplazamientos de los niveles de energía de los iones embebidos en la red cristalina desde sus posiciones de equilibrio, como consecuencia de las perturbaciones provocadas sobre ellos por la misma acción de la red. El primer mecanismo, efecto de Doppler, es la causa de aparición de una función $g(\nu)$ particular en los láseres gaseosos; mientras que el segundo, interacción de la red cristalina con los iones causantes de la absorción o emisión de radiación, resulta predominante en los láseres sólidos.

La introducción de la función perfil de línea espectral $g(\nu)$ en la expresión (26), lleva a la siguiente fórmula:

$$\frac{W_i' g(\nu) d\nu}{A g(\nu) d\nu} = \frac{c^3}{8\pi h \nu^3} \rho(\nu) \quad . \quad (28)$$

En el primer miembro, los productos $A g(\nu) d\nu$ y $W_i' g(\nu) d\nu$ son las probabilidades de que ocurra la emisión de un fotón de manera espontánea o estimulada en el intervalo de frecuencias comprendidas entre ν y $\nu + d\nu$.

Luego, la probabilidad de emisión estimulada de radiación entre ν y $\nu + d\nu$ se puede expresar por:

$$W_i' g(\nu) d\nu = \frac{c^3 A}{8\pi h \nu^3} g(\nu) \rho(\nu) d\nu \quad . \quad (29)$$

Recordando que el flujo de energía en el intervalo de frecuencia comprendido entre ν y $\nu + d\nu$ se define como:

$$I(\nu)d\nu = c \rho(\nu)d\nu \quad , \quad (30)$$

la expresión anterior se puede escribir:

$$W_i' g(\nu)d\nu = \frac{c^2 A}{8\pi h \nu^3} g(\nu) I(\nu) d\nu \quad . \quad (31)$$

Como ya se indicó, las señales luminosas emitidas por el láser deben ser consideradas como estrictamente monocromáticas, o al menos muchísimo más monocromáticas que las líneas espectrales emitidas espontáneamente. Luego, de acuerdo con las expresiones (21) y (22) encontradas por A.L. Schawlow y C.H. Townes, puede expresarse de un modo adecuado el flujo de energía que ellas transportan, adoptando el formalismo de la Delta de Dirac. Por las propiedades de la Delta, el flujo de energía de una señal de ancho espectral despreciable se puede escribir como:

$$I(\nu) = I(\nu_0) \delta(\nu_0 - \nu) \quad , \quad (32)$$

donde $I(\nu_0)$ es el flujo de energía en la frecuencia ν_0 .

Reemplazando esta expresión del flujo energético en la ecuación (31) e integrando sobre todo el intervalo de frecuencias, resulta:

$$W_i = \frac{c^2 A}{8\pi h \nu^3} g(\nu) I(\nu_0) \quad , \quad (33)$$

que es la fórmula más usada en el estudio de las condiciones bajo las cuales un medio material puede comportarse como amplificador de radiación, es decir, como un láser.

9.3. Osciladores y amplificadores de radiación.

Un láser consta de un medio material que actúa como generador y amplificador de radiación, y que ocupa una parte o todo el volumen de un resonador óptico o cavidad resonante tipo Fabry-Perot. Por lo general, en los láseres gaseosos el medio material ocupa sólo una parte del volumen definido por la cavidad resonante abierta; en cambio, el medio activo en los láseres sólidos ocupa con frecuencia todo el volumen mencionado.

La misión del resonador, como se verá más adelante en detalle, es doble. Por un lado, los espejos que constituyen el resonador favorecen la

reinyección de la luz emitida por el medio material excitado en el sentido del eje de la cavidad al mismo medio. Luego, tales espejos generan, por las leyes de la reflexión, una serie teóricamente interminable de imágenes del medio activo. En tal sentido, los espejos actúan como un multiplicador de la longitud del medio material. Por el otro, el resonador debe sustentar ondas estacionarias de luz, es decir, debe mantener una configuración estable del campo de radiación y, de preferencia, la que se corresponda con el menor número posible de modos resonantes.

Se dice que el láser entra en oscilación o comienza a oscilar cuando las pérdidas de todo el sistema son compensadas por la emisión estimulada. Este efecto consiste en la producción de ondas estacionarias en aquellas frecuencias a las cuales resuena la cavidad. Tales ondas estacionarias son generadas por las ondas de radiación emitidas por los átomos, moléculas o iones excitados que forman el medio activo, y que se propagan por sucesivas reflexiones en los espejos, atravesándolo un gran número de veces. A su vez, el láser amplifica cuando la contribución de la emisión estimulada es tan decisiva que supera las pérdidas del sistema. En esta circunstancia, a través de los espejos, que nunca reflejan toda la radiación que incide en ellos sino que transmiten una pequeñísima porción, emerge un haz luminoso estrecho, perfectamente definido y en extremo direccional.

Entre los distintos factores causantes de pérdidas en el sistema pueden enumerarse la transmisión, la absorción, la difracción y las imperfecciones de los espejos, así como la absorción, la dispersión y la difracción debidas a la falta de homogeneidad del medio activo. Desde el punto de vista matemático, todos estos factores se pueden caracterizar por un solo parámetro, al que se designa γ . Por lo dicho sobre los tipos de pérdidas que condicionan el funcionamiento del láser, debe notarse que al parámetro γ debe sumarse el efecto del coeficiente de absorción $\alpha(\nu)$ del medio material, introducido en la ecuación (17). Por ello, puede expresarse como $\alpha = \alpha(\nu)D + \gamma$, la pérdida total por pasaje de la radiación dentro de la cavidad resonante.

Tal como se definió anteriormente, para alcanzar el régimen de oscilación del láser es necesario que la energía emitida por el medio activado por acción de la emisión estimulada de radiación compense la energía perdida. Y para lograr el mantenimiento de dicho régimen deben cumplirse estrictamente las condiciones de bombeo, de modo que las poblaciones de los niveles de energía que intervienen, N_1 y N_2 , guarden una relación que asegure la inversión de población. Tales condiciones de bombeo se denominan umbral. Si el medio activo es excitado por encima del umbral, el dispositivo pasará del régimen de oscilación al de amplificación.

El parámetro γ que expresa las pérdidas permite definir el factor de calidad Q , de la cavidad resonante. Si la distancia que separa los espejos es D y la frecuencia de la radiación emitida por el medio activo que se desea amplificar es ν , resulta válida la siguiente igualdad:

$$nD/\gamma c = Q / 2\pi\nu \quad , \quad (34)$$

donde c es la velocidad de la luz en el vacío y n el índice de refracción del medio a la frecuencia ν . Dicho cociente denota que, como consecuencia de las pérdidas indicadas por γ , la onda de luz necesitará más tiempo recorrer la distancia D a través del medio activo hasta que por acción de la amplificación vuelva a tener, al menos, la misma intensidad inicial. Por eso, se suele indicar que los cocientes de la fórmula (34) son iguales a:

$$t_{\text{fotón}} = nD / \gamma c = Q / 2\pi \nu \quad , \quad (35)$$

En este tiempo $t_{\text{fotón}}$, llamado vida media de un fotón en el medio material, la intensidad decrece en un factor $1/e$. Es decir, la cantidad de fotones F en la cavidad decae según la expresión:

$$F = F_0 e^{-(\gamma c / nD)t} \quad , \quad (36)$$

siendo F_0 el número inicial de fotones.

9.4. Condición umbral e inversión crítica de población.

La relación entre las poblaciones N_2 y N_1 , desequilibrada por la emisión de radiación, debe volver, por efecto del bombeo, al valor mínimo que asegure la inversión de población.

El modo según el cual varía la relación entre N_2 y N_1 se determina analizando las variaciones temporales de la intensidad de radiación dentro de la cavidad resonante.

La variación temporal del flujo de energía o intensidad de radiación dentro de la cavidad resonante debida a la acción de la emisión estimulada, se expresa por:

$$(dI/dt)_1 = h \nu (N_2 - N_1 g_2/g_1) c W_i / V \quad , \quad (37)$$

donde $h\nu$ es la energía del fotón irradiado al producirse la transición $2 \rightarrow 1$ en un átomo. La justificación de esta expresión es la siguiente: Esta energía $h\nu$ debe multiplicarse por el número de todas las transiciones estimuladas probables, que es igual a

$$(N_2 - N_1 g_2/g_1) W_i \quad ,$$

para obtener la energía total emitida. Además, estos procesos ocurren en el volumen V ocupado por el medio activo, por lo que, al multiplicar por la velocidad de la luz c , la expresión (37), resulta ser un flujo de energía por unidad de tiempo.

Por otra parte, la variación temporal del flujo de energía por su paso dentro de la cavidad resonante debida a las pérdidas del sistema, se expresa:

$$(dI/dt)_2 = - [I(\nu_0) / t_{\text{fotón}}] \quad . \quad (38)$$

Luego, la condición de umbral, esto es, la condición para que el dispositivo entre en oscilación o para que amplifique, es la siguiente:

$$(dI/dt)_1 + (dI/dt)_2 \geq 0 \quad . \quad (39)$$

Si en la ecuación (39) el término de la izquierda es nulo, el láser entrará en oscilación y mantendrá indefinidamente ese régimen si no se producen variaciones de otros parámetros experimentales no tenidos en cuenta en la discusión. Si, en cambio, el término de la izquierda crece y llega a ser mayor que cero, el láser entrará decididamente en el régimen de amplificación.

Introduciendo en la ecuación (39) los valores respectivos y reagrupando términos, se encuentra la condición siguiente para que el láser comience a oscilar en la frecuencia central ν_c de la línea espectral correspondiente a la transición $2 \rightarrow 1$:

$$N_2 - N_1 \cdot g_2/g_1 = \frac{8\pi\nu_c^2}{Ac^3g(\nu_c) t_{\text{fotón}}} \cdot V = \Delta N_c \quad . \quad (40)$$

ΔN_c se denomina inversión crítica de población. Habitualmente se emplea en las discusiones la densidad de inversión crítica de población, esto es, el cociente $\Delta N_c/V$. La densidad de inversión crítica depende solamente del parámetro $t_{\text{fotón}}$.

La expresión (40) que señala el mínimo del valor de la inversión necesaria para provocar oscilaciones en el medio, es conocida como Condición de Schawlow-Townes.

Una vez obtenido en el medio material el grado de inversión ΔN_c , el dispositivo entrará en el régimen de oscilación. Si los mecanismos de

bombeo acrecientan el grado de inversión de manera que supere el valor ΔN_C , el dispositivo se transforma en un láser.

9.5. Inversión crítica y ensanchamientos homogéneos e inhomogéneos de las líneas espectrales.

La ecuación (40) puede escribirse particularmente para los casos de ensanchamiento de línea que generan perfiles del tipo de Lorentz o de Gauss. La Figura 11 compara ambos perfiles suponiendo que el ancho medio $\Delta \nu$ es el mismo.

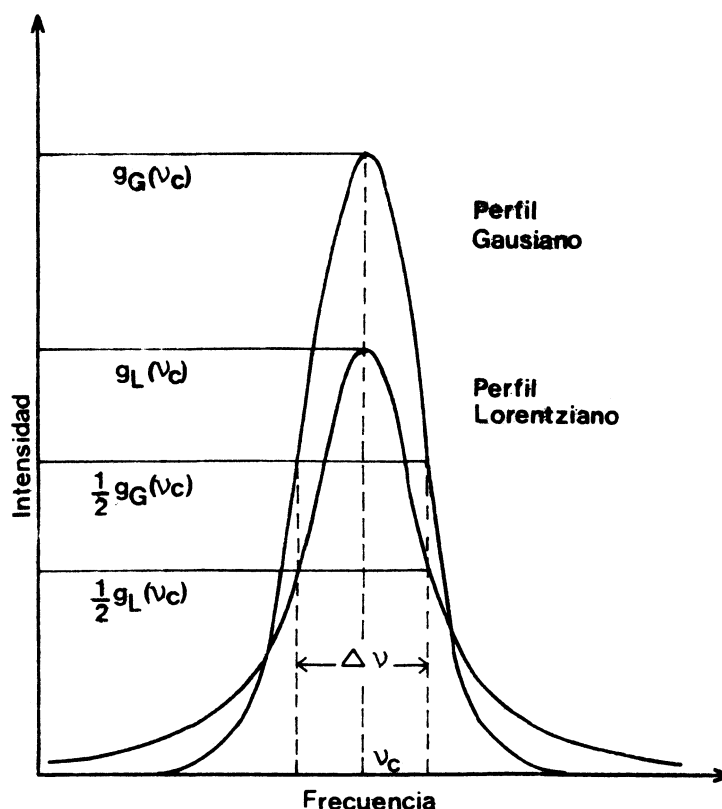


Figura 11. Comparación de los perfiles de línea espectral gaussiano y lorentziano de iguales anchos medios.

El ensanchamiento de las líneas espectrales de emisión o de absorción se produce por diversas causas. Por ejemplo, una transición entre los nive-

les de energía E_2 y E_1 de los iones inmersos en un cristal no generarán la misma frecuencia ν_{21} en las diferentes partes del cristal porque en cada una de ellas los iones estarán soportando distintas condiciones físicas - tensiones, campos eléctricos, etc. - que perturban al azar las posiciones de sus niveles 1 y 2 de energía. Otro ejemplo es el de los átomos en un gas a muy baja presión excitado por descargas eléctricas. En este caso, debido a las diferentes velocidades v de movilidad de los átomos, aparece para cada uno de ellos una variación de frecuencia $\Delta\nu$ emitida o absorbida, la que se puede determinar de acuerdo con la expresión para el efecto Doppler:

$$\Delta\nu = v/c \nu_c \quad .$$

Si la distribución de las velocidades es maxwelliana, entonces la distribución de las frecuencias alrededor de ν_c , genera un perfil de Gauss.

A las dos formas ejemplificadas se las clasifica como provocando ensanchamientos inhomogéneos de las líneas espectrales.

Otra forma que provoca ensanchamientos de las líneas espectrales pero de tipo homogéneo, se atribuye al hecho de lo limitado del tiempo medio de permanencia de los electrones en los niveles de energía excitados. Este tiempo de permanencia, llamado habitualmente vida media del nivel de energía considerado, se relaciona con la probabilidad de transición. Si la probabilidad de transición es pequeña, el tiempo de permanencia será grande y viceversa. Ahora bien, por el Principio de Incerteza enunciado por W. Heisenberg, resulta que la energía de los niveles nunca podrá determinarse con precisión y exactitud si los electrones no se establecen en ellos por un tiempo infinito. Pero, si ello ocurriera, nunca se observarían transiciones, con lo que la incongruencia sería total. De hecho, como se observan transiciones a las que se les puede determinar sus frecuencias así como sus probabilidades de transición, resulta que la incerteza en la diferencia de energía entre los dos niveles electrónicos, resulta ser igual a:

$$\Delta(E_2-E_1) \approx \frac{h}{2\pi} \left(\frac{1}{t_2} + \frac{1}{t_1} \right) \quad ,$$

donde t_1 y t_2 son las respectivas vidas medias de los niveles 1 y 2.

Si el conjunto de átomos excitados correspondiera a un gas sometido a presión alta, se producirán choques entre ellos que perturbarán la duración de la permanencia de los electrones en los niveles excitados. Tales colisiones producen una forma de ensanchamiento homogéneo de las líneas espectrales que se representa por un perfil del tipo de Lorentz.

La diferencia sustancial entre las dos formas de ensanchamiento, reside en que en el homogéneo los átomos individuales son indistinguibles, mientras que en el inhomogéneo son distinguibles. Es decir que, en este último caso, suponiendo los átomos excitados en un gas a muy baja presión se puede determinar la frecuencia emitida en una cierta transición por aquéllos que están animados con una dada velocidad, por ejemplo.

A la inversa, los átomos que emiten radiación en el seno de un gas a muy alta presión no podrá individualizárselos como para reconocer aquéllos que han sufrido un cierto tipo de colisiones de otros átomos.

En tal caso, el perfil de la línea espectral emitida espontáneamente es lorentziano y su descripción se hace mediante la fórmula siguiente:

$$g_L(\nu) = \frac{\Delta\nu}{2\pi[(\nu-\nu_c)^2 + (\frac{\Delta\nu}{2})^2]} \quad . \quad (41)$$

Si, como se supuso al establecer la expresión (40), se desea amplificar la línea espectral justamente en su frecuencia central ν_c , deberá emplearse para calcular la inversión crítica el valor de $g_L(\nu_c)$. Esta operación permite arribar al siguiente resultado:

$$g_L(\nu_c) = 2 / \pi \Delta\nu \quad ,$$

el que, reemplazado en la ecuación (40), permite obtener el valor:

$$\Delta N_c = \frac{4\pi^2 \nu_c^2 \Delta\nu}{A c^3 t_{\text{fotón}}} \cdot V \quad . \quad (42)$$

En cambio, si el perfil de línea espontáneo es gaussiano resulta

$$g_G(\nu_c) = \frac{2(\pi \ln 2)^{1/2}}{\pi \Delta\nu} \quad ,$$

con lo que la inversión crítica valdrá:

$$\Delta N_c = \frac{4\pi^2 \nu_c^2 \Delta\nu}{(\pi \ln 2)^{1/2} A c^3 t_{\text{fotón}}} \cdot V \quad . \quad (43)$$

De acuerdo con las expresiones (42) y (43), se concluye que para dos transiciones electrónicas caracterizadas por idénticas frecuencias ν_c y probabilidades de emisión espontánea A , amplificadas en dispositivos que posean igual volumen de material activo V y factor de calidad Q , aquella con ensanchamiento homogéneo y que tiene un perfil lorentziano requerirá una inversión crítica ΔN_c menor en aproximadamente 1,45 de la ensanchada en forma inhomogénea y que tiene perfil gaussiano.

9.6. 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 12 muestra un esquema general de los niveles de energía que entran en juego en cada caso.

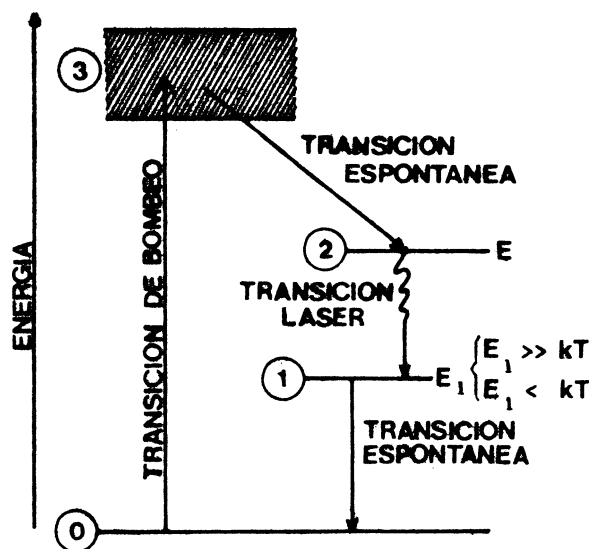


Figura 12. Esquema de tres y cuatro niveles.

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 in-

versión de población entre algunos niveles excitados y el nivel fundamental.

El láser de cuatro niveles se caracteriza por tener en su nivel terminal 1 una población N_1 en equilibrio térmico despreciable comparada con la inversión crítica ΔN_C . Por lo tanto, la condición para que se inicien las oscilaciones en el sistema será, de acuerdo con la ecuación (40), expresada por:

$$N_2 \approx \Delta N_C \quad . \quad (44)$$

Ahora bien, para que el nivel 1 se despueble lo suficiente a medida que se suceden las transiciones electrónicas desde el nivel 2, es necesario que su diferencia de energía E_1 con respecto a la del nivel fundamental sea mucho mayor que kT . En los láseres sólidos se verifica perfectamente esta condición ya que $E_1 \geq 8 kT$.

La condición recién impuesta para definir un láser de cuatro niveles garantiza que cualquier elevación de la temperatura del medio activo del láser no provoque un incremento puramente térmico de la población electrónica N_1 , de acuerdo con la ley de distribución de las poblaciones de Boltzmann.

Para mantener siempre constantemente, o al menos por breves períodos, la población del nivel 2, de modo que se cumpla la condición de inversión crítica (44), hay que proporcionar al sistema una potencia igual a:

$$P = h \nu A \Delta N_C \quad . \quad (45)$$

En esta fórmula $h\nu$ representa la energía de un fotón emitido espontáneamente en virtud de la transición $2 \rightarrow 1$. Al multiplicar $h\nu$ por ΔN_C se obtiene la energía total emitida y, a su vez, al multiplicar este producto por la probabilidad de emisión espontánea por unidad de tiempo A , resulta la potencia emitida en la transición $2 \rightarrow 1$. Por lo tanto, para conservar la población $N_2 \approx \Delta N_C$ se requerirá que el sistema reciba por lo menos una potencia igual a la emitida. Es claro que, reemplazando en la ecuación (45) la expresión $A \Delta N_C$ despejada de la (42), y suponiendo que el perfil de línea sea lorentziano, se obtiene:

$$P = \frac{4\pi^2 h \nu^3 \Delta \nu}{c^3 t_{\text{fotón}}} V \quad . \quad (46)$$

El láser de tres niveles se caracteriza, en cambio, porque el nivel de energía terminal 1 se encuentra a una altura energética E_1 con res-

pecto del nivel fundamental, que es menor que kT . Por ello, la potencia requerida para que se inicien oscilaciones en el sistema tendrá que ser mayor que la requerida por el láser de cuatro niveles en un factor $\Delta N_C / N_2$. Este factor aparece porque la población N_1 del nivel terminal de la posible transición láser se incrementa por el simple hecho de que la temperatura del medio material se aproxima al valor T o lo alcanza. Luego, para compensar la falta de eficiencia del bombeo desde el nivel fundamental hasta el nivel 2, es necesario incrementar la potencia proporcionada al sistema.

Todas las exigencias referidas hasta aquí, permiten enumerar los requerimientos que debe cumplir un medio material para que en él se produzca efecto de amplificación de radiación. Tales requerimientos son los siguientes:

- 1) De acuerdo con la expresión (46), la potencia P necesaria para el bombeo depende en grado sumo de la frecuencia ν . Tal como se afirmó antes, se confirma aquí por qué los primeros láseres que se construyeron amplificaban las señales en la región de frecuencias cercanas al infrarrojo, es decir a frecuencias bajas.
- 2) De acuerdo con la expresión (46), el ancho espectral espontáneo debe ser lo menor posible a fin de que se reduzca la potencia P necesaria para el bombeo.
- 3) El medio material debe tener un factor de pérdidas γ lo más bajo posible en la frecuencia correspondiente a la transición $2 \rightarrow 1$. Esto se justifica así: La potencia P de bombeo es inversamente proporcional a $t_{\text{fotón}}$, según la ecuación (46), y como, a su vez, $t_{\text{fotón}}$ es inversamente proporcional a γ , según la ecuación (35), resulta que, a mayor pérdida, más potencia P será necesaria para que el medio material actúe como amplificador.
- 4) La densidad de inversión crítica $\Delta N_C / V$ sólo depende del parámetro $t_{\text{fotón}}$ o, lo que es lo mismo, de las pérdidas por pasaje α . A mayor valor de inversión crítica ΔN_C , N_2 resulta mayor que $N_1 g_2 / g_1$, de acuerdo con la ecuación (40), y, por la relación expresada por la fórmula (18), el coeficiente de absorción $\alpha(\nu)$ adquiere valor negativo, con lo que el medio se convierte en un amplificador de radiación. La ecuación (17) sigue teniendo valor en estas condiciones.
- 5) Es preferible que el nivel de energía terminal 1 esté lo más alejado posible del nivel fundamental a fin de permitir la operación del dispositivo según el esquema de cuatro niveles. Esto producirá un considerable ahorro de potencia de bombeo.
- 6) Conviene que la probabilidad de transición del nivel de energía metastable 2 sea lo más cercana posible a la unidad para que la eficiencia de la conexión óptica entre el nivel 2 y el 1 sea elevada. A esta conclusión se llega a partir de la ecuación (45).
- 7) Por último, el nivel terminal 1 debe poseer una probabilidad de transi-

ción con el nivel fundamental mayor que la del nivel inicial 2 con el 1. De este modo, el nivel terminal 1 se despoblará más rápidamente de lo que se llena, y asegura así que la inversión crítica ΔN_C sea elevada.

En caso de que el nivel terminal 1 coincida con el nivel fundamental del átomo, molécula o ión, se obtiene el esquema básico de acuerdo con el cual se interpreta el funcionamiento del láser de rubí y de algunos láseres gaseosos, como se dijo ya. Estos láseres funcionan según un régimen pulsante. Es decir, la excitación que provoca el bombeo desde el nivel fundamental hasta el nivel superior 2 de la transición láser se inicia e interrumpe de modo abrupto. Así, la población almacenada en el nivel fundamental es transferida al nivel 2 en una proporción tal que inmediatamente se establece la inversión crítica ΔN_C con respecto del nivel terminal 1 o fundamental, en este caso. La eficiencia de funcionamiento mostrada por estos láseres depende de cuán abrupta sea la iniciación de la excitación, así como también de su duración. En condiciones experimentales muy especiales ha sido posible hacer que el láser de rubí funcione en régimen continuo. Una de las condiciones esenciales para conseguirlo es la refrigeración efectiva de la barra de rubí a bajísimas temperaturas.

Sólo algunos láseres que se pueden interpretar como funcionando de acuerdo con los esquemas de tres o cuatro niveles emiten radiación de esa manera dual. El láser de He-Ne, por ejemplo, puede ser excitado de modo que su emisión sea continua o intermitente.

9.7. Umbral de excitación y amplificación.

Un modo de mensurar el umbral de excitación de un medio material resulta de la discusión de la expresión de la absorción integral (18) de Füchtbauer y Ladenburg, ya que para cada transición de emisión o de absorción se puede definir su distribución espectral $g(\nu)$. En efecto, teniendo en cuenta el grado de degeneración g_i de los niveles energéticos, la expresión (18) se puede reescribir, según se señaló previamente, así:

$$\int \alpha(\nu) d\nu = \frac{h\nu}{c} B_{12} \left(\frac{g_2}{g_1} N_1 - N_2 \right) \quad . \quad (47)$$

Ahora bien, considerando la definición de $g(\nu)$ según la expresión (27) resulta que puede definirse el coeficiente de absorción de la siguiente forma

$$\alpha(\nu) = \frac{h\nu}{c} B_{12} g(\nu) \left(\frac{g_2}{g_1} N_1 - N_2 \right) \quad . \quad (48)$$

Introduciendo las relaciones (5) entre B_{12} y B_{21} en función de A_{21} , el coeficiente de absorción para una transición entre el nivel inicial 1 y el final 2, puede escribirse así:

$$\alpha(\nu) = \frac{c^2 A_{21}}{8\pi \nu^2} \cdot [g(\nu) \left(\frac{g_1}{g_2} N_1 - N_2 \right)] \quad . \quad (49)$$

Es común denominar al coeficiente de la fórmula (49):

$$\sigma = \frac{c^2 A_{21}}{8\pi \nu^2} \quad ,$$

como la sección eficaz de absorción en la transición electrónica $1 \rightarrow 2$ por átomo del medio material.

El significado físico de σ es la probabilidad de que cada átomo absorba un fotón del campo de radiación de densidad $\rho(\nu)$ en la frecuencia ν correspondiente a la transición $1 \rightarrow 2$.

Tanto la expresión (49) como las (47) y (48), se refieren al caso en que el medio material absorba de modo que el coeficiente $\alpha(\nu)$ resulta ser positivo. Pero cuando en el medio material se produce, por la acción del bombeo, la inversión de población, el término $[(g_2/g_1)N_1 - N_2]$ adquiere valores negativos, con lo que el medio activado contribuye a amplificar la radiación que lo atraviesa, en lugar de absorberla.

Las situaciones para las cuales se producen la oscilación o la amplificación en el láser se describen por la condición de umbral dada por la expresión (39) y, correlativamente, ello ocurrirá, al menos, cuando en el medio material se produzca la inversión crítica de población, según la expresión (40). Entonces la expresión del coeficiente de absorción negativo $\alpha(\nu)$, será

$$\alpha(\nu) = \sigma g(\nu) \Delta N_c \quad , \quad (50)$$

denominado habitualmente factor de amplificación o de ganancia por pasaje.

Recordando la expresión de la ley de absorción en un material:

$$I(x, \nu) = I(0, \nu) e^{-\alpha(\nu)x} \quad , \quad (17)$$

se puede considerar ahora como que en cada pasaje a través del medio

activo del láser la radiación se amplifica en una proporción igual a $e^{\alpha(\nu)D}$, donde $\alpha(\nu)$ resulta de la expresión (50) y D es la longitud del medio atravesado.

Sin embargo, como ya se mencionaron previamente formas de pérdidas que se introducen en el láser debidas a su diseño y a las características de los materiales y componentes, la expresión del factor de amplificación por pasaje (50) no siempre es aplicable.

Las pérdidas aludidas más importantes, como ya se señaló, son la transmisión, la absorción, la difracción y las imperfecciones de los espejos, así como la absorción, la dispersión y la difracción debidas a las inhomogeneidades del medio material, las que se caracterizan por el coeficiente γ .

Las primeras afectan la intensidad de la radiación una vez por cada pasaje, mientras que las restantes lo hacen aleatoriamente en cada pasaje. Todos estos factores de pérdida pueden agruparse en un solo parámetro, el que, como ya se señaló antes, al asociarlo al coeficiente $\alpha(\nu)$, permitió definir el coeficiente de pérdida total por pasaje.

Entonces, volviendo a la expresión exponencial (17), la amplificación o absorción negativa y las pérdidas por pasaje se pueden calcular empleando el factor:

$$e^{[\alpha(\nu)D - \gamma]} \quad , \quad (51)$$

cuando su exponente se anula.

Entonces, la condición umbral para que el dispositivo oscile, es decir, cuando la amplificación compensa las pérdidas, se expresa por:

$$\alpha(\nu)_m D = \gamma \quad , \quad (52)$$

donde $\alpha(\nu)_m$ es la mínima amplificación que debe obtenerse en el medio, teniendo en cuenta la distribución espectral $g(\nu)$ de la transición que desea amplificarse y la inversión crítica ΔN_c entre los niveles de energía en que ella es emitida.

Obviamente, cuando se produzca la desigualdad:

$$\alpha(\nu)D > \gamma \quad , \quad (53)$$

el dispositivo será un láser. Por otro lado, si nunca se alcanzase a cumplir la condición mínima de umbral (52) y en el medio ocurriera siempre que:

$$\alpha(\nu)D < \gamma \quad , \quad (54)$$

el dispositivo nunca permitirá poner en evidencia la emisión estimulada que ciertamente está generándose.

La Figura 13 describe el caso de un medio material que al no estar excitado absorbe la radiación que lo atraviesa (trazo continuo) y que cuando se genera en él una cierta inversión de población por el bombeo, se produce la amplificación de la radiación en el rango de frecuencias para el cual $\alpha(\nu) > \gamma$ (trazo cortado).

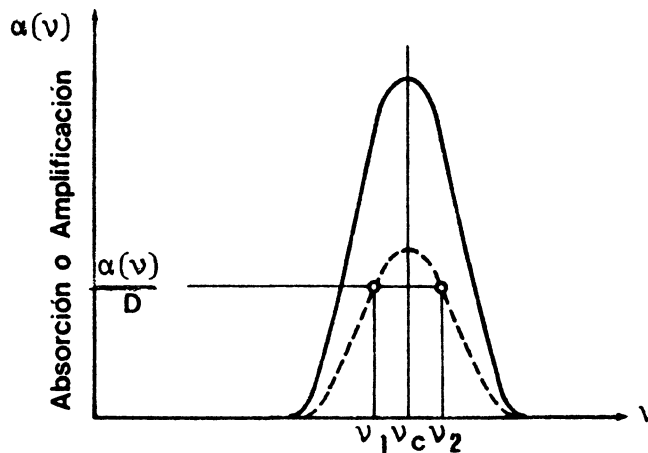


Figura 13. Perfiles de absorción y amplificación. Curva de trazo continuo: absorción del material sin excitar; curva de trazo cortado: amplificación.

Una importante conclusión a la que se arriba después de esta discusión relativa al umbral de excitación del medio activo, es que existe la posibilidad, al menos, de determinar las pérdidas que afectan al dispositivo. En efecto, a partir de las expresiones (50) y (52), se puede obtener la fórmula siguiente para la inversión crítica ΔN_C :

$$\Delta N_C = N_2 - N_1 \frac{g_2}{g_1} = \frac{\gamma}{\sigma D g(\nu)} \quad (55)$$

Esta expresión, señala con toda claridad que para mantener la inversión crítica en un nivel adecuado como para que se produzca amplificación de radiación, debe haber una relación bien controlada entre las pérdidas

γ y la longitud D del medio activo. Si las pérdidas se acrecientan y no se puede aumentar la inversión de población, será necesario alargar la longitud del dispositivo D para que siga teniendo validez la ecuación (55).

También puede elevarse la excitación del medio material para aumentar la inversión crítica ΔN_c , requerida por el dispositivo sin que sea necesario alargar su longitud D .

Más adelante se analizarán las pérdidas producidas por la reflexión y absorción en los espejos del láser, las que, por otro lado, son predominantes.

9.8. Cinética de la inversión de población.

En este párrafo se describirán las condiciones de equilibrio estacionario entre las poblaciones de los niveles de energía que interesan en la generación de la amplificación de radiación. Ello permitirá profundizar en las características del medio activo y se calculará una expresión de la potencia emitida por el láser.

El análisis que sigue se formaliza para un medio que opera según el esquema de cuatro niveles. Sin embargo, se supondrá con el objeto de simplificar, que el grupo de niveles asociados a la banda de absorción 3 de la Figura 12 actúa como simple intermediario para provocar la inversión crítica entre los niveles 2 y 1. Por ello, en la Figura 14 no se los ha graficado y se considera que la acción de bombeo deposita en el nivel 2 una población N_2 a razón de B_2 átomos por segundo.

Además, en este análisis deberá tenerse presente que los electrones pueden realizar otras transiciones desde el nivel 2 a otros niveles más bajos no graficados en la Figura 14, así como también transiciones de tipo no radiativas provocadas por la colisión entre el átomo excitado y otro átomo con diferente grado de excitación. Asimismo, deberá considerarse que el bombeo provocará también la excitación del nivel 1 a razón de B_1 átomos por segundo. Este proceso de bombeo, si bien es indeseable resulta, en la práctica, imposible de evitar.

Una manera de proceder a la simplificación de la nomenclatura y hacer menos compleja la interpretación de los procesos, es introduciendo los tiempos de permanencia de los electrones en los respectivos niveles. Estos tiempos de permanencia, entonces, deberán agrupar a las probabilidades de transición espontánea a los diferentes niveles inferiores y de transiciones no radiativas. Como ya se ha visto, el tiempo medio de permanencia del electrón o vida media del respectivo nivel de energía es una magnitud física inversa a la probabilidad de transición. De allí que, por ejemplo, el tiempo de ocupación del nivel 1 se puede expresar así:

$$1 / t_1 = A_{10} \quad ,$$

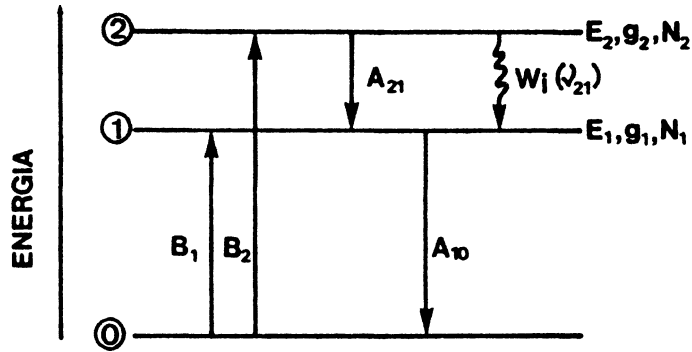


Figura 14. Sistema idealizado de niveles de energía de un medio amplificador. Con B_1 y B_2 se identifican las transiciones de absorción, mientras que con A_{10} y A_{21} se hace lo propio con las de emisión espontánea. La inversión de población ocurre entre los niveles 2 y 1, caracterizados por sus energías E_2 y E_1 , sus multiplicidades g_2 y g_1 y sus poblaciones N_2 y N_1 . La emisión láser se produce por las transiciones $W_i(\nu_{21})$.

mientras el correspondiente al nivel 2, se lo hace de la siguiente manera:

$$1/t_2 = A_{21} + \sum_{i=0}^m (1/t_{2i})_R + \sum_{i=0}^m (1/t_{2i})_{NR} \quad , \quad (56)$$

donde A_{21} es la probabilidad de transición espontánea $2 \rightarrow 1$; la primera sumatoria es la suma de las probabilidades de transición espontánea producidas entre el nivel 2 y los m niveles intermedios ubicados entre el nivel fundamental 0 y el nivel 2, exceptuando el nivel 1, y la segunda sumatoria es la suma de las probabilidades de transición no radiativa desde el nivel 2 a estos mismos niveles m .

La probabilidad W_i de transición estimulada $2 \rightarrow 1$ ya ha sido calculada anteriormente y se empleará aquí el resultado dado por la fórmula (33):

$$W_{21} = W_i = \frac{c^2 A_{21}}{8\pi h \nu_{21}^3} g(\nu) I(\nu) \quad . \quad (57)$$

Como el flujo de radiación $I(\nu)$ en la frecuencia ν_{21} afecta tan-

to a los átomos cuyo electrón excitado está en el nivel 2 cuanto a los que los tienen en el nivel 1, deberán tenerse presente las transiciones de absorción $1 \rightarrow 2$, con una probabilidad igual a:

$$W_{12} = \frac{g_2}{g_1} W_i, \quad (58)$$

donde g_1 y g_2 son las multiplicidades o factores de degeneración de los niveles 1 y 2.

Luego, se puede describir la cinética de las variaciones de las poblaciones de los niveles entre los que se produce la transición láser, teniendo en cuenta el bombeo y las transiciones espontáneas, estimuladas y no radiativas, según el sistema de ecuaciones siguientes:

$$\frac{dN_2}{dt} = B_2 - \frac{N_2}{t_2} - (N_2 - \frac{g_2}{g_1} N_1) W_i, \quad (59)$$

$$\frac{dN_1}{dt} = B_1 - \frac{N_1}{t_1} + A_{21} N_2 + \frac{N_2}{(t_{21})_N R} + (N_2 - \frac{g_2}{g_1} N_1) W_i. \quad (60)$$

La expresión (59) describe la variación temporal de la población N_2 del nivel 2 en función del bombeo B_2 que lo excita, de las transiciones de emisión radiativas y no radiativas $-N_2/t_2$ y estimulada $-N_2 W_i$ que lo despueblan y de las transiciones de absorción desde el nivel 1, $g_2 N_1 W_i / g_1$, que contribuyen a repoblarlo. Del mismo modo, la expresión (60) describe la variación temporal de la población N_1 del nivel 1 en función del bombeo B_1 que lo excita, de las transiciones espontáneas $A_{21} N_2$, no radiativas N_2/t_{21} y estimuladas $N_2 W_i$ provenientes del nivel 2 que contribuyen también a poblarlo y de todas las transiciones espontáneas y no radiativas $-N_1/t_1$ que se inician en el nivel 1, así como de las transiciones de absorción $-g_2 N_1 W_i / g_1$ con destino al nivel 2 que contribuyen a despoblarlo.

Suponiendo que todos los procesos mencionados alcanzan la condición de equilibrio estacionario, representado por la anulación de las derivadas en las expresiones (59) y (60), se puede determinar la inversión crítica de población como:

$$\Delta N_C = N_2 - \frac{g_2}{g_1} N_1 = \frac{B_2 t_2 - (B_1 + \tau B_2) t_1 \frac{g_2}{g_1}}{1 + [t_2 + (1 - \tau) t_1 \frac{g_2}{g_1}] W_i}. \quad (61)$$

En la expresión (61) se ha introducido, con el fin de simplificar su escritura, el factor τ , que es igual a:

$$\tau = t_2 \left[A_{21} + \sum_{i=0}^m (1/t_{2i})_{NR} \right] .$$

Es decir, el factor τ representa la relación entre las probabilidades de transición espontáneas $2 \rightarrow 1$ más las no radiativas y la probabilidad t_2 de las transiciones que se inician en el nivel 2.

A partir de la expresión de la inversión crítica (61) se puede calcular su valor en ausencia de emisión estimulada o sea para cuando $W_i = 0$. En este caso resulta:

$$\Delta N_0 = (N_2 - \frac{g_2}{g} N_1)_0 = B_2 t_2 - (B_1 + \tau B_2) t_1 \frac{g_2}{g_1} . \quad (62)$$

Combinando las expresiones (61) y (62), se puede expresar la inversión crítica ΔN_C así:

$$\Delta N_C = \frac{\Delta N_0}{1 + \beta W_i \frac{t_2}{\tau}} , \quad (63)$$

donde:

$$\beta = \tau \left[1 + (1 - \tau) \frac{t_1 g_2}{t_2 g_1} \right] . \quad (64)$$

La expresión (64) incluye una compleja relación entre varias probabilidades de transición y las multiplicidades de los niveles de energía de interés, la cual no importa demasiado conocer, y que, a los efectos prácticos, conviene recordar que β es una constante para cada sistema atómico.

Supóngase el caso hipotético de un medio en el que $\tau=1$ y $B_1=0$. Es decir, un medio en el que no se producen otras transiciones espontáneas desde el nivel 2 salvo la $2 \rightarrow 1$ y en el que el bombeo no contribuye a excitar el nivel 1.

Para dicho medio, la expresión (62) resulta:

$$\Delta N_0 = (N_2 - \frac{g_2}{g_1} N_1)_0 = B_2(t_2 - t_1 \frac{g_2}{g_1}) , \quad (65)$$

y, por lo tanto, se producirá en él amplificación cuando $\Delta N_0 > 0$, con lo que:

$$t_2 > t_1 \frac{g_2}{g_1} . \quad (66)$$

Obsérvese entonces que cuando $g_1 > g_2$, es posible lograr en el medio inversión crítica ΔN_0 en ausencia de emisión estimulada y, por lo mismo, obtener amplificación de radiación a pesar de que sea $t_2 < t_1$.

Esta conclusión viene a completar el enunciado del séptimo requerimiento consignado al final del parágrafo 9.6. y con respecto del cual, al no haber impuesto condiciones a la relación entre las multiplicidades g_1 y g_2 de los niveles entre los que se espera obtener inversión de población, resulta ser opuesto.

Teniendo en cuenta las expresiones (57) y (58), ocurre que si $g_1 > g_2$, la probabilidad de emisión estimulada W_{21} es mayor que la probabilidad de absorción W_{12} para la frecuencia ν_{21} . De ahí que, y tal como ya se lo había deducido previamente, se vuelve a concluir que será posible obtener emisión estimulada aún cuando $N_2 < N_1$.

En la gran mayoría de los láseres estudiados, siempre se ve satisfecha la relación $t_1 g_2 / t_2 g_1 \ll 1$. Ello significa que $\beta \approx \tau$ y, por lo tanto, la (62) queda reducida a:

$$\Delta N_C = \frac{\Delta N_0}{1 + t_2 W_i} . \quad (67)$$

9.9. Importancia de la emisión espontánea.

Conviene ahora, después de esta discusión sobre la cinética de la inversión de población, analizar el papel que juega la emisión espontánea en la frecuencia ν_{21} producida dentro del medio activo antes de que se genere suficiente inversión de población como para que el dispositivo amplifique. Se ha visto ya, cómo en el medio excitado es posible obtener una cierta inversión ΔN_0 , aún en ausencia de emisión estimulada ($W_i = 0$).

No obstante ello, la emisión espontánea emitida isotrópicamente en el medio excitado será la semilla que permitirá iniciar el proceso que conduce al establecimiento de una densidad de radiación $\rho(\nu)$ dentro del dispositivo. Parte de este campo de radiación, más precisamente aquél que se propaga según el eje de simetría elegido o el eje óptico de la cavidad resonante del dispositivo, será el responsable de introducir las primeras transiciones estimuladas. Esta primera etapa ocurre, en general, cuando el medio se encuentra por debajo del umbral. Cuando el proceso de realimentación iniciado incrementa la intensidad de la radiación de modo que la amplificación iguale a las pérdidas, se dice que el dispositivo oscila y esto ocurre cuando se encuentra justamente en el umbral. Obviamente, si el proceso de realimentación continúa contribuyendo a la amplificación de la radiación superando las pérdidas, el dispositivo se comporta francamente como un amplificador y se encuentra excitado por encima del umbral.

Debe advertirse, sin embargo, que la secuencia delineada depende fuertemente del mantenimiento del bombeo suficiente como para que el dispositivo pase de la excitación inicial - en la que la emisión espontánea es esencial - a comportarse como un amplificador - en el que la emisión estimulada resulta preponderante - pasando por la etapa intermedia del oscilador.

Si el bombeo suministrado al dispositivo nunca le hace superar la etapa de oscilación, el factor de amplificación irá disminuyendo progresivamente, hasta que la densidad de radiación $\rho(\nu)$ en el medio alcance un valor tan bajo que se anula toda posibilidad de observar la emisión estimulada. En cambio, si el bombeo le hace superar las condiciones de oscilador, el dispositivo comienza a comportarse como un láser.

En toda la discusión contenida en este párrafo nunca se hizo intervenir la variable temporal para caracterizar la excitación o bombeo. El análisis realizado no puede ir más allá del planteo de las condiciones estacionarias a partir de la anulación de las derivadas en las expresiones (59) y (60), a menos que se pueda profundizar en casos particulares cuando se conocen los valores de las probabilidades de transición y tiempo medio de ocupación de los niveles. 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 15.

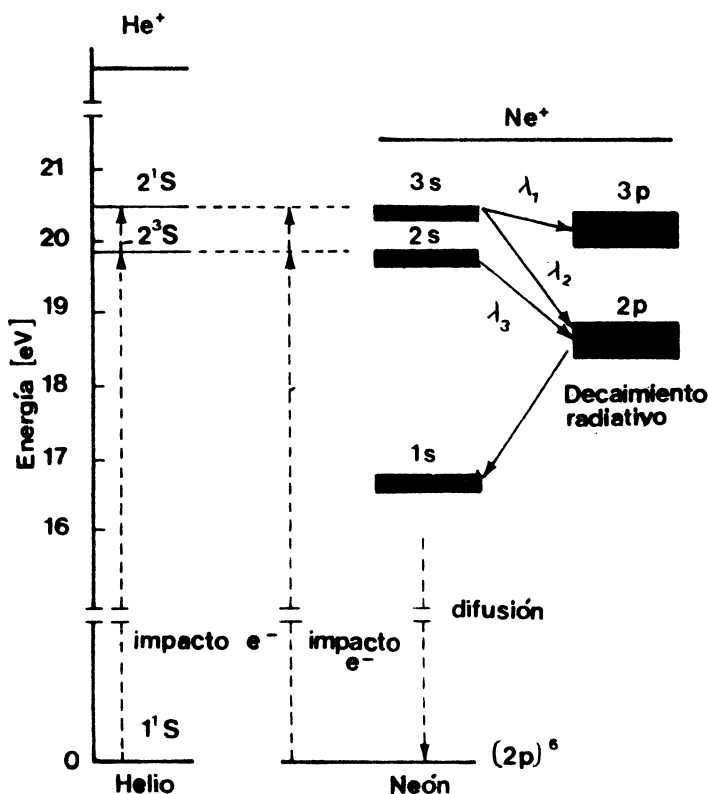


Figura 15. Esquema de los niveles de energía que intervienen en el láser de He-Ne. λ_1 y λ_3 emisión IR y λ_2 emisiones visibles.

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.

9.10. 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 16 muestra la excitación de la molécula de hidrógeno y la emisión estimulada en las bandas de Werner y de Lyman.

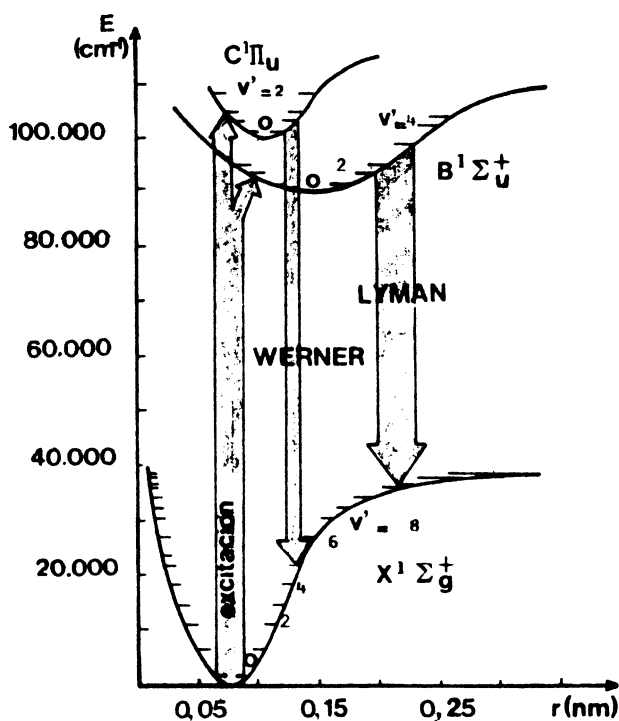


Figura 16. Excitación y emisión del láser de hidrógeno molecular.

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

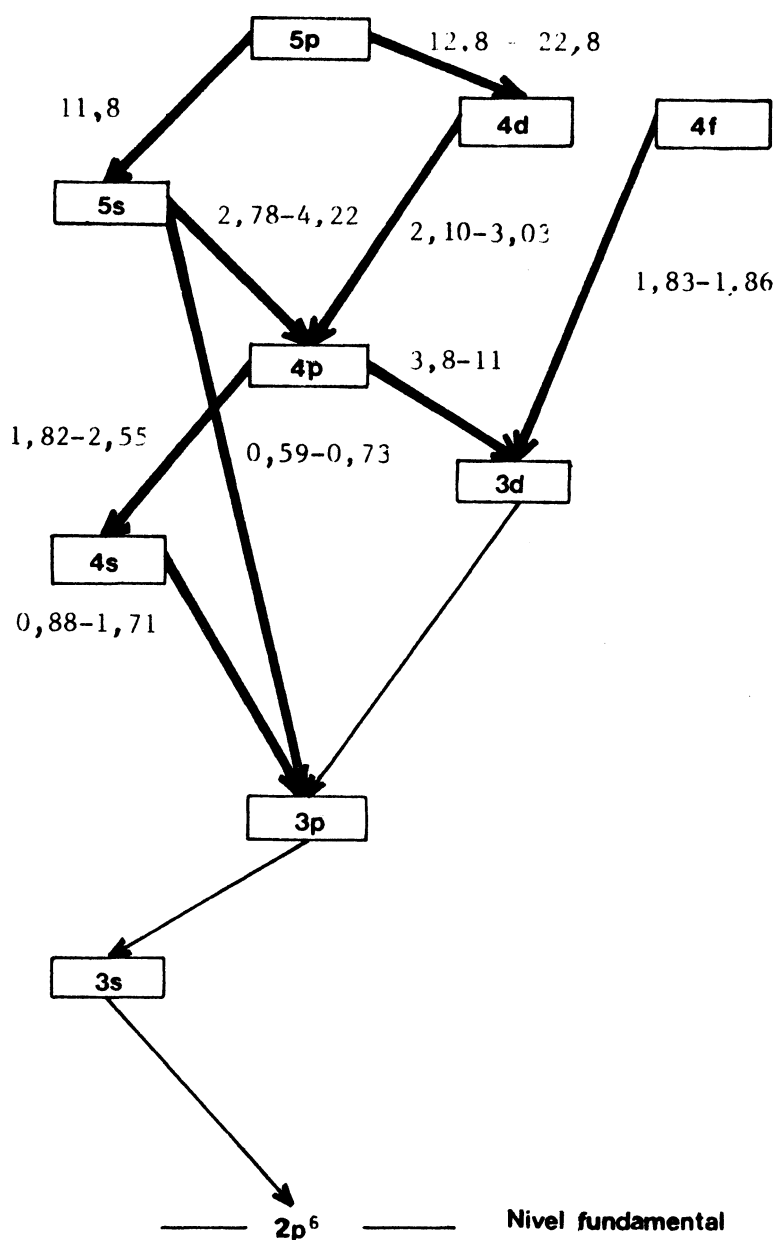


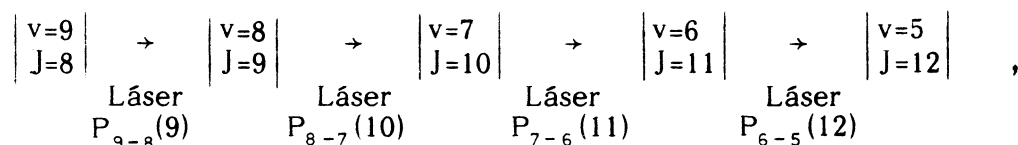
Figura 17. Transiciones en cascada entre niveles de energía del Ne I en las que se observa emisión láser ($\longrightarrow$) y emisión espontánea ($\longrightarrow$). Cada bloque de niveles de energía está indicado con su símbolo de Racah. El espaciado entre los niveles no guarda escala uniforme. Las longitudes de onda correspondientes a los grupos de transiciones están expresadas en μm .

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 H_2 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 H_2 , el H_2^- , y una vez pasado dicho tiempo, el electrón es expulsado con energía casi nula, y por lo tanto, la molécula H_2 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 17 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 CO ⁴³, 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:



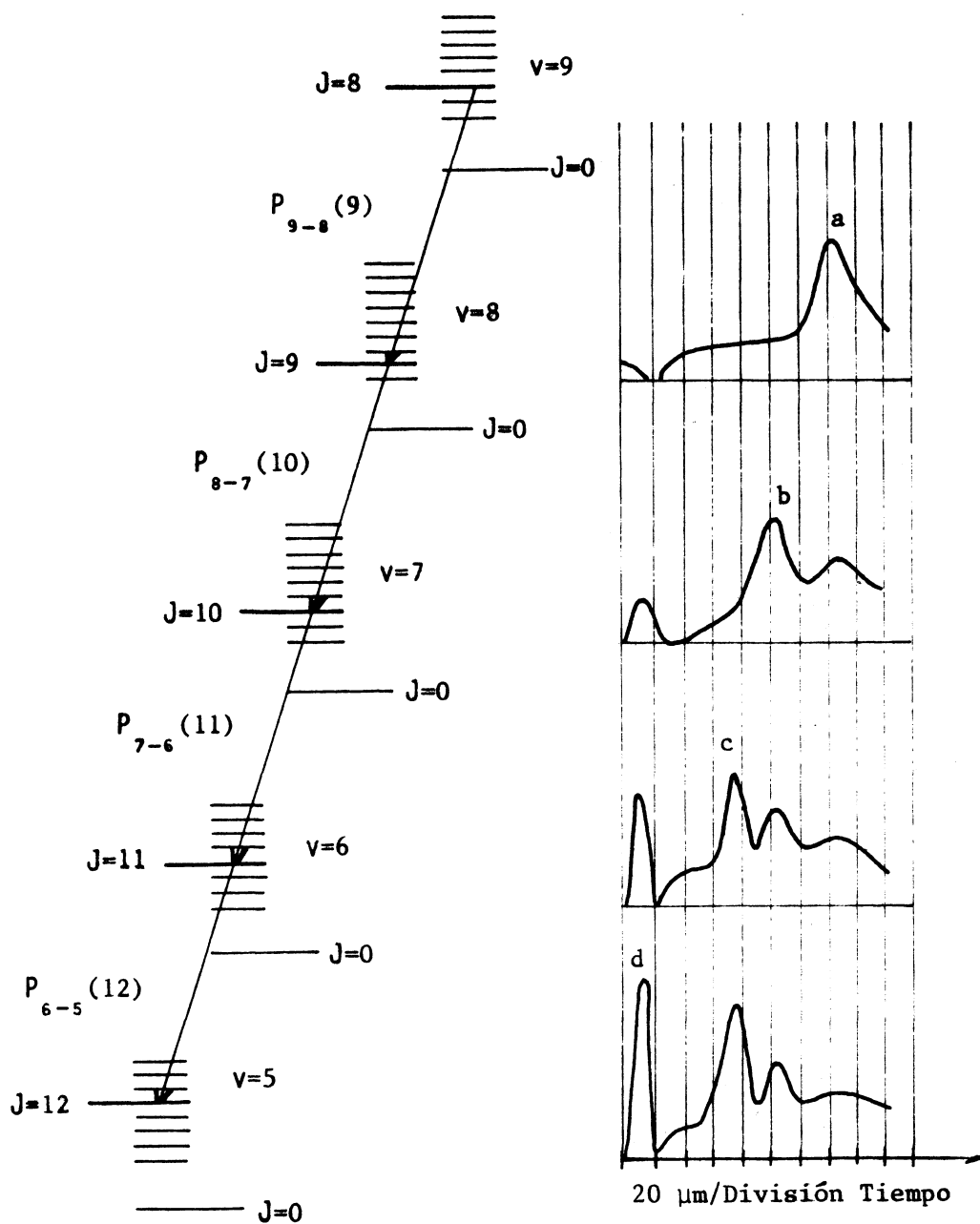


Figura 18. Cascada láser inversa entre niveles rotacionales-vibracionales del CO correspondientes a las transiciones a) $P_{6-5}(12)$ en $5,08845 \mu\text{m}$, b) $P_{7-6}(11)$ en $5,14530 \mu\text{m}$, c) $P_{8-7}(10)$ en $5,20345 \mu\text{m}$, y d) $P_{9-8}(9)$ en $5,26310 \mu\text{m}$. La posición 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 18. En la Figura 18, 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 (17) 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 (17) 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 (17) 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. Ya se hizo mención a que el coeficiente de pérdidas por pasaje γ , permite definir el factor de calidad Q de la cavidad resonante. La expresión (34) define este factor en cavidades resonantes ópticas. Q puede ser constante o variable; en este último caso, las técnicas para hacer que Q varíe desde cero hasta su valor máximo son diversas y su grado de complejidad depende del tiempo en que se desea realizar dicha variación. De todos modos, aún en los láseres cuyo funcionamiento requiere alguna técnica de Q variable, importa siempre saber cuál es el máximo valor de Q que puede lograrse al construir una cavidad resonante.

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 resuene. 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 (14) 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 9.

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:

$$\frac{L_1 L_2}{\lambda D} \geq 1 \quad , \quad (68)$$

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 (68) 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 (68) resulta:

$$N = L^2 / \lambda D \geq 1 \quad , \quad (69)$$

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

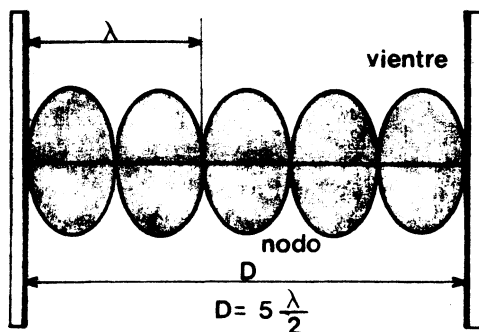


Figura 19. Onda estacionaria.

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:

$$p = 2D' / \lambda \quad , \quad (70)$$

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 19, 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:

$$D'' = \int_0^{D'} n(z) dz \quad , \quad (71)$$

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 (71), 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 (70) como sigue:

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

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

$$\Delta\lambda = \lambda^2 / 2D \quad , \quad (72)$$

o lo que es lo mismo:

$$\Delta\lambda = \lambda / p \quad . \quad (73)$$

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 (72), resulta:

$$\Delta\nu = c / 2D \quad . \quad (74)$$

La Figura 9 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 9 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 (70). 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:

$$E_{q+1} = \varepsilon E_q ,$$

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 20.

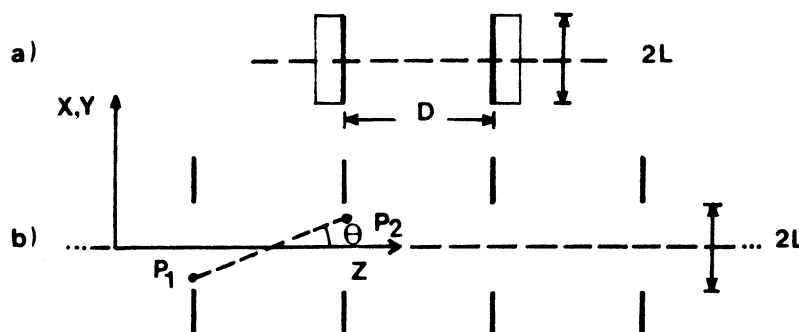


Figura 20. a) Interferómetro de Fabry-Perot de placas circulares planas y paralelas de diámetro $2L$ separadas una distancia D . b) Sucesión de pupilas circulares difractantes de diámetro $2L$ equiespaciadas a distancias D y que simulan al interferómetro de Fabry-Perot.

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 $E_i(x,y,z,t)$ se calcula según la expresión de Kirchoff:

$$E_d(x,y,z,t) = \frac{i}{2\lambda} \iint E_i(x,y,z,t) \frac{e^{-i 2\pi D/\lambda}}{D} (1+\cos\theta) dx dy \quad , \quad (75)$$

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:

$$E_{m,n} = E H_m \left[\left(\frac{2\pi}{\lambda D} \right)^{1/2} x \right] H_n \left[\left(\frac{2\pi}{\lambda D} \right)^{1/2} y \right] e^{-\pi(x^2 + y^2)/\lambda D} \quad , \quad (76)$$

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

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

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 18.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 21.a), puede obtenerse la distribución de intensidad calculando los cuadrados de E_0 , E_1 y E_2 , cuyas representaciones aparecen en la Figura 21.b).

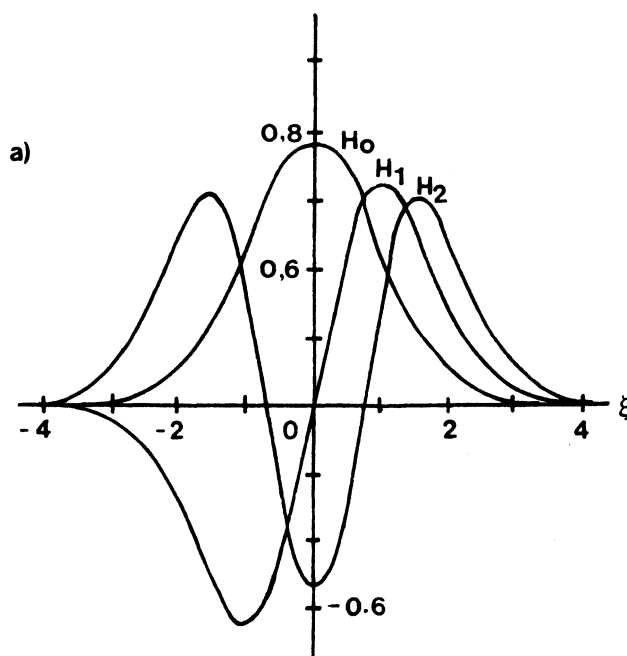


Figura 21.a) Amplitudes de la componente eléctrica del campo electromagnético para los modos $m = 0, 1$ y 2 .

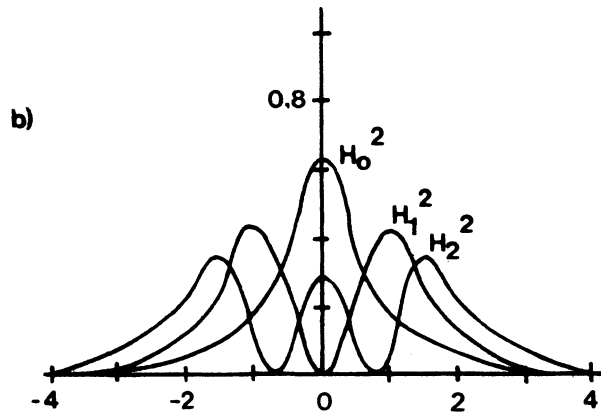


Figura 21.b) Intensidades de los modos $m = 0, 1$ y 2 .

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

$$\left(\frac{2}{\lambda}\right)^2 = \left(\frac{p}{D}\right)^2 + \left(\frac{m^2+n^2}{4L^2}\right) \quad (77)$$

Si sólo se excita el modo transversal más bajo dentro de la cavidad resultan $m = 0$ y $n = 0$, con lo que la expresión (77) se reduce a la (70) que gobierna las condiciones relativas a los modos longitudinales. En este caso, la ecuación (76) 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 (77), permite calcular la diferencia en longitud de onda entre los modos transversales excitados

en un láser:

$$\Delta\left(\frac{1}{\lambda}\right) = \frac{1}{2D} \left[\Delta p + \frac{1}{8} \frac{D\lambda}{L^2} (m \Delta m + n \Delta n) \right] \quad (78)$$

En el caso de ser $m = 0$ y $n = 0$, se obtiene la ecuación (74) para los modos longitudinales.

A pesar de todo, éstas no son las únicas conclusiones que se pueden extraer del análisis de las ecuaciones (76) y (77).

La expresión (76) 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 manojo 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 22 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 manojo 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 TEM_{00} 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.

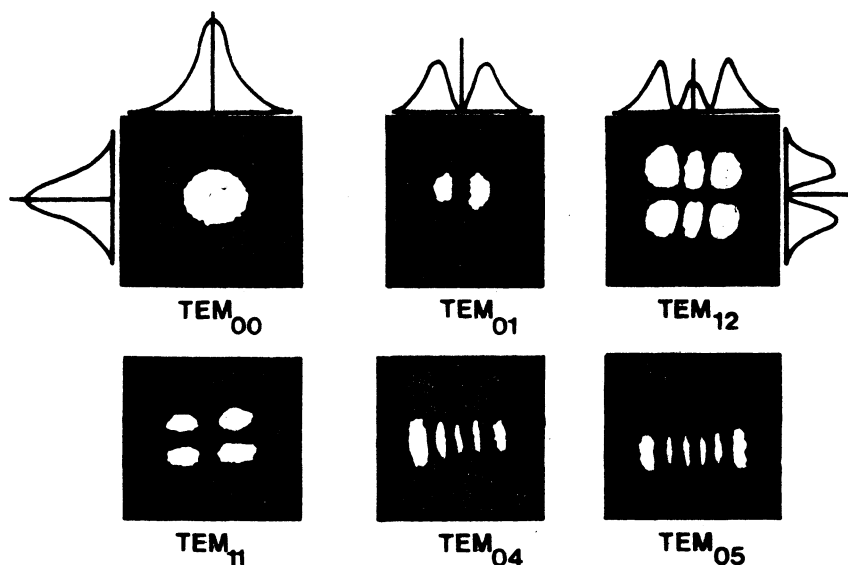


Figura 22. Fotografía de algunos modos transversales emitidos por el láser de He-Ne, $\lambda=632,8$ nm, y distribución de la intensidad de campo para aquéllos de $m = 0, 1$ y 2 , y $n = 0$ y 1 .

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 (76). 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 $z=0$:

$$W_0 = (D \lambda / \pi)^{1/2} \quad (79)$$

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

$$W = W_0 [1 + (z/D)^2]^{1/2} \quad (80)$$

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 23 ilustra convenientemente lo aquí expuesto.

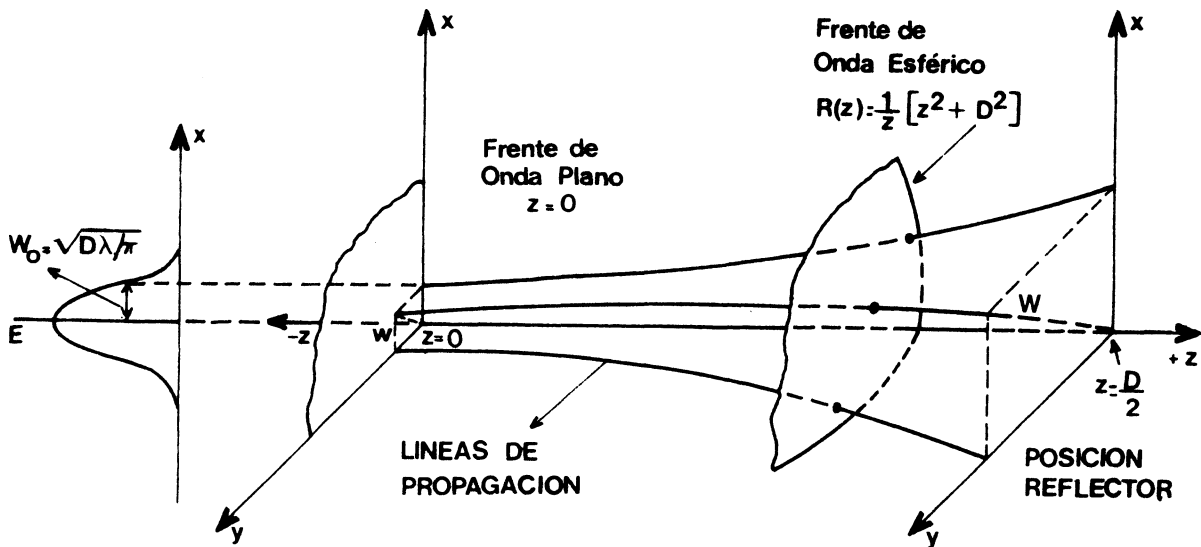


Figura 23. Parámetros de la cavidad resonante.

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:

$$R(z) = \frac{1}{z} (z^2 + D^2) \quad . \quad (81)$$

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 (81) 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:

$$0 \leq \left(1 - \frac{D}{R_1}\right) \left(1 - \frac{D}{R_2}\right) \leq 1 \quad . \quad (82)$$

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 (82). 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 24 ilustra, con varios ejemplos, las configuraciones de cavidades resonantes incluso dos de elevada pérdida o inestables porque no cumplen la ecuación (82).

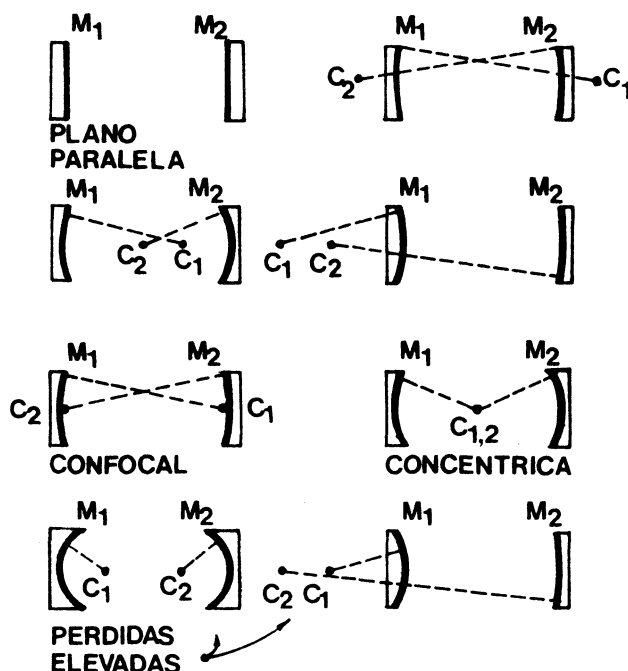


Figura 24. Diferentes configuraciones de cavidades resonantes estables e inestables, constituidas por espejos planos y esféricos.

La Figura 25 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.

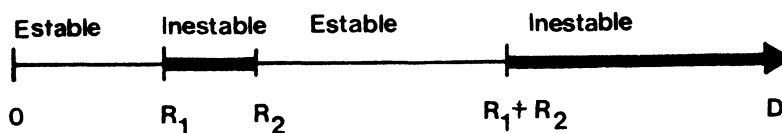


Figura 25. Caracterización de las configuraciones estables e inestables de cavidades resonantes.

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 23 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 26. La observación de la misma muestra claramente que la cavidad confocal es la mejor.

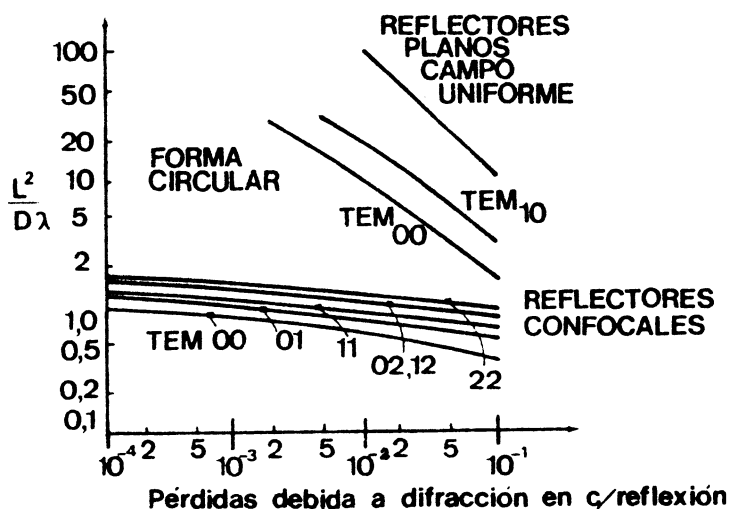


Figura 26. Pérdidas en las cavidades resonantes.

10.4. Las pérdidas de las cavidades resonantes y la potencia emitida por el láser.

En el párrafo 9.3 se definió el factor Q de calidad de la cavidad resonante y se lo relacionó con las pérdidas por pasaje, de acuerdo con la fórmula (34).

Conviene ahora circunscribir tal definición sólo a las pérdidas por reflexión en los espejos, por ser ellas las más importantes, habitualmente.

Se relacionará entonces el factor de calidad Q de la cavidad con los coeficientes de reflexión que caracterizan a los espejos. Para ello, se parte de la expresión (17):

$$I(x, \nu) = I(0, \nu) e^{-\alpha(\nu)x} ,$$

en la que se introducen las pérdidas totales por pasaje γ , de acuerdo con la expresión (51), resultando:

$$I(x, \nu) = I(0, \nu) e^{-[\alpha(\nu)x + \gamma]} . \quad (83)$$

Si la radiación recorre la cavidad de ida y vuelta entre ambos espejos, de modo que al final su intensidad resulte ser igual a la inicial, debido a que la ganancia ha compensado las pérdidas, de la expresión (83) resulta:

$$I(2D, \nu) = I(0, \nu) ,$$

con lo que:

$$e^{-[2\alpha(\nu)D + 2\gamma]} = 1 . \quad (84)$$

Ahora bien, considerando el factor de pérdida total γ como compuesto por la suma de las pérdidas γ_r por reflexión en los espejos más el resto de las pérdidas, $\gamma_p D$ a lo largo del pasaje, la (84) se puede escribir así:

$$e^{-[2\alpha(\nu)D + 2\gamma_p D]} \cdot e^{-2\gamma_r} = 1 , \quad (85)$$

de donde se deduce que:

$$\alpha(\nu)D + \gamma_p D + \gamma_r = 0 \quad . \quad (86)$$

Como la energía de la onda que se refleja dentro de la cavidad, disminuye en un pasaje completo de ida y de vuelta en un factor $r_1 r_2$, se expresan las pérdidas en las dos reflexiones sucesivas según la siguiente fórmula:

$$e^{2\gamma_r} = r_1 r_2 \quad , \quad (87)$$

con lo que la (85) queda, teniendo en cuenta las (86) y (87), así expresada:

$$e^{-\alpha(\nu)D - \gamma_p D} \cdot r_1 r_2 = 1 \quad . \quad (88)$$

Tomando logaritmos en la ecuación (88) resulta:

$$-\alpha(\nu)D - \gamma_p D + \ln r_1 r_2 = 0 \quad . \quad (89)$$

A partir de la fórmula (89) se define el valor de la ganancia $\alpha(\nu)$ como para que el dispositivo comience a oscilar. Cuando ello ocurre, se dice que el dispositivo se encuentra en las condiciones de umbral y la ganancia vale:

$$\alpha_u(\nu) = \gamma_p - \frac{1}{D} \ln r_1 r_2 = 0 \quad . \quad (90)$$

Obsérvese que los términos de la derecha de la ecuación (90) representan las pérdidas totales por pasaje, de modo que si la excitación del medio activo produce una inversión de población suficiente como para que $\alpha(\nu) > \alpha_u(\nu)$, la ganancia será mayor que las pérdidas y el dispositivo se convierte en un láser.

El factor de ganancia $\alpha(\nu)$ se puede expresar, teniendo en cuenta las fórmulas (49), (62) y (67), así:

$$\alpha(\nu) = \frac{c^2 A}{8\pi \nu^2} g(\nu) \frac{\Delta N_0}{1 + W_i t_2} \quad ,$$

en la que puede denominarse:

$$\alpha_o(\nu) = \frac{c^2 A}{8\pi \nu^2} g(\nu) \Delta N_o ,$$

y, entonces, resulta:

$$\alpha(\nu) = \frac{\alpha_o(\nu)}{1 + W_i t_2} . \quad (91)$$

Ahora bien, cuando $\alpha(\nu) > \alpha_u(\nu)$, se puede calcular a partir de la expresión (91) el valor de W_i para aquellas condiciones que corresponden al funcionamiento del dispositivo por encima del umbral. En efecto, despejando W_i de la (91) e introduciendo el valor de $\alpha(\nu)$ cuando se ha sobrepasado el umbral, dado por la (90), queda:

$$W_i = \frac{1}{t_2} \left[\frac{\alpha_o(\nu)}{\alpha_u(\nu)} - 1 \right] = \frac{1}{t_2} \left[\frac{\alpha_o(\nu)}{\gamma_p - \frac{1}{D} \ln r_1 r_2} - 1 \right] \quad (92)$$

El resultado de la expresión (92) permite determinar la inversión crítica según la (49), cuando se ha sobrepasado el umbral, resultando:

$$\Delta N_u = \frac{8\pi \nu^2}{c^2 A g(\nu)} \left(\gamma_p - \frac{1}{D} \ln r_1 r_2 \right) . \quad (93)$$

Con todos estos elementos y teniendo en cuenta la expresión (49), se calcula la potencia P_1 emitida por el láser, la que fue generada dentro del volumen V_m que ocupan los modos resonantes, en función de las pérdidas debidas a las reflexiones de la luz en los espejos. Reemplazando los valores dados por las (92) y (93), la potencia P_1 resulta:

$$P_1 = \frac{8\pi \nu^3}{c^2 A t_2 g(\nu)} \frac{V_m}{D} (\gamma_p D - \ln r_1 r_2) \left[\frac{\alpha_o(\nu)}{\gamma_p D - \ln r_1 r_2} - 1 \right] . \quad (94)$$

Los coeficientes de reflexión r_1 y r_2 de ambos espejos se relacionan con las respectivas reflectividades R_1 y R_2 , según la fórmula $R_i = r_i^2$, dados que unos se aplican a las amplitudes y otros a las intensidades de la radiación. Entonces se puede simplificar la expresión (94), calculando:

$$-\ln r_1 r_2 = -\ln \sqrt{R_1 R_2} \quad . \quad (95)$$

Suponiendo que los espejos tienen iguales reflectividades $R_1 = R_2 = R$, la que a su vez es muy cercana a la unidad, y recordando que la transmitancia T de los espejos se define como $T = 1 - R$, resulta al hacer el cálculo del logaritmo por serie:

$$-\ln r_1 r_2 = -\ln (1 - T) \approx T \quad . \quad (96)$$

Luego, la (94) se escribe:

$$P_1 = \frac{8 \pi \nu^3}{c^2 A t_2 g(\nu)} \frac{V_m}{D} [\alpha_0(\nu) D - \gamma_p D - T] \quad . \quad (97)$$

El cociente V_m/D se denomina área efectiva del modo resonante láser.

Ahora bien, como sólo parte de la potencia P_1 emitida por el láser está disponible fuera de la cavidad, se define la potencia de salida útil P_s como la relación entre la potencia transmitida a través de los espejos fuera de la cavidad y la potencia perdida en el proceso de funcionamiento del láser y que, por lo tanto, no contribuye al mantenimiento de la amplificación. Dicha relación se expresa como:

$$P_s = P_1 \frac{T}{\gamma_p D + T} \quad . \quad (98)$$

La Figura 27 muestra los resultados del estudio realizado por P. Laures⁴¹ para el láser de He-Ne que emite en el rojo, $\lambda = 632,8$ nm, los que revelan claramente cómo P_s alcanza un máximo para los diferentes valores de las pérdidas, expresadas porcentualmente, y considerando una ganancia del 12%.

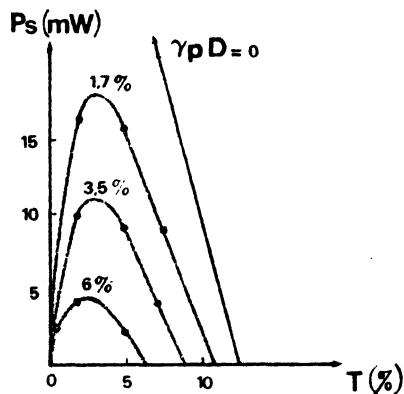


Figura 27. Potencia de salida útil P_s en función de la transmitancia T de los espejos de un láser de He-Ne que emite en $\lambda=632,8$ nm. Se supone que la ganancia del láser es $\alpha(\nu_0)D = 12\%$ para diferentes valores porcentuales de las pérdidas $\gamma_p D$.

Recordando la definición del factor Q y algunas de sus expresiones empleadas anteriormente, puede evaluárlas para algún caso particular con el propósito de ejemplificar lo ya señalado, en el sentido que para los láseres, Q adopta valores elevadísimos.

En efecto, de la (34) se despeja:

$$Q = \frac{2\pi \nu n D}{\alpha c} ,$$

en la que introduciendo el valor de α para las condiciones correspondientes al umbral de excitación, dado por la (90) y (96), queda:

$$Q = \frac{2\pi n D}{\lambda(\gamma_p + \frac{T}{D})} . \quad (99)$$

Suponiendo el caso del láser de He-Ne que emite en $\lambda=632,8$ nm, que emplea espejos de una transmitancia $T=0,02$ separados una distancia $D=30$ cm y con pérdidas γ_p prácticamente nulas, resulta $Q=10^8$.

Este valor de Q permite hacerse una idea más cabal de la cuantía

de la inversión crítica ΔN_C necesaria para que un medio excitado amplifique. Combinando las expresiones (42) o (43) con la (35) puede calcularse la inversión crítica correspondiente a perfiles de línea espectral lorentziano o gaussiano, respectivamente. Se sigue del cálculo que ΔN_C debe ser del orden de 10^8 en la región espectral de longitud de onda $\lambda=1 \mu\text{m}$.

Los espejos de revestimiento de aluminio, plata, oro o platino pueden fabricarse mediante deposición en vacío y alcanzar excelente calidad. Sin embargo, por su elevada absorción en la región visible, su empleo se limita a las regiones espectrales de longitud de onda por encima de 800 nm y por debajo de 400 nm. El valor del coeficiente de reflexión de estos espejos es del orden de 0,9. Con espejos recubiertos de multicapas de sustancias dieléctricas la situación mejora notablemente, ya que se obtienen superficies exentas de absorción, de excelente calidad óptica, de reflectividad fácilmente controlable y con valores de hasta 0,997. Estos espejos se fabrican por deposición de vapores en el vacío formando capas alternadas de materiales dieléctricos de índice de refracción distinto, uno mayor y otro menor que el índice de refracción del sustrato de vidrio o cuarzo. Las capas sucesivas tienen un espesor igual a múltiplos de $\lambda/4$. Por efecto de interferencias constructivas entre las ondas reflejadas en cada capa, el balance de la reflexión favorece la región espectral del entorno de λ . De este modo se obtienen espejos de alta selectividad: reflejan la radiación del entorno de λ y transmiten la restante. En general, se seleccionan materiales dieléctricos de muy escasa absorción en la región visible y con ellos se fabrican juegos de espejos que abarcan el intervalo entre 400 y 800 nm, en el que, como se dijo ya, resulta difícil emplear metálicos.

Por último, y a modo de resumen de todo lo descrito en relación con la emisión de los láseres y su estructura modal, debe señalarse que, dependiendo del nivel de excitación del medio material, las condiciones de umbral fijan la cantidad de modos longitudinales que quedarán efectivamente sustentados. En la Figura 28 se ha marcado la línea correspondiente a un umbral de amplificación que permite sólo mantener el modo longitudinal central de la estructura de tres que puede sustentar un láser de He-Ne de 30 cm de longitud.

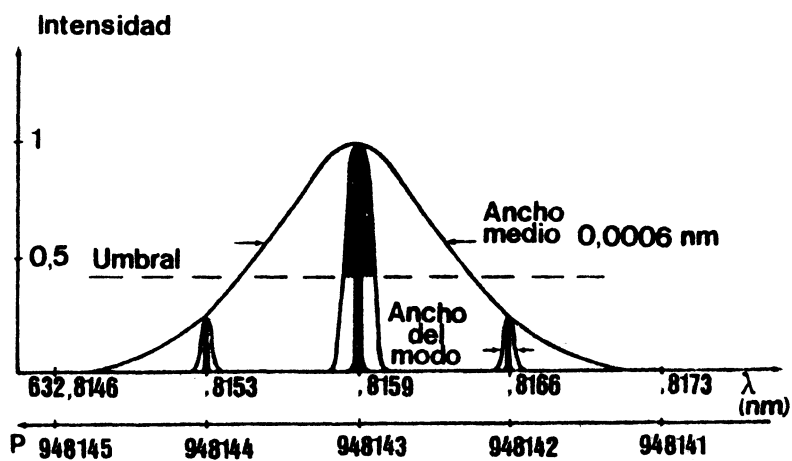


Figura 28. El modo resonante central es el único efectivamente amplificado porque su intensidad supera el umbral de la amplificación.

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.E.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 Nº 2929922 de los Estados Unidos de Norteamérica (30 de julio de 1958).
- 19 - Gould, G. - Patentes Nº 953721 a Nº 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 N° 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).

ALGUNOS TIPOS DE LASERES

H.F. Ranea Sandoval

L.B. Scaffardi

Introducción

No podemos ni siquiera intentar una enumeración exhaustiva de todos los tipos de láseres, sea clasificándolos por naturaleza, modificaciones de un mismo láser o láseres diseñados para alguna aplicación específica; resultaría tedioso aún tratándose de un tema apasionante. Tampoco podríamos analizar los láseres posibles o incluirlos en las largas listas ya mencionadas.

En la década del '60 se pretendía hacer del láser un instrumento de cualidades extraordinarias en espera de una posible aplicación: y las aplicaciones no se hicieron esperar. Se inició así una espiral siempre creciente de interacción directa entre: aplicaciones → láser adecuado → nuevas aplicaciones → láseres más perfeccionados, etc. Se puede decir que siempre se puede mejorar un láser para adecuarse mejor 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 tan dispares campos como la medicina, la ingeniería genética, la física nuclear, 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. Así, podremos ver un láser de CO_2 soldando, analizando y sintetizando sustancias químicas, generando plasmas, oficiando de bisturí óptico (debería decirse lumínico) y en otras tareas difíciles de enumerar.

Es por eso que decidimos tratar algunos tipos de láseres. Para ello, hemos recurrido a una clasificación (tan arbitraria como cualquiera otra) que refleje de alguna manera las diversas situaciones en las que vamos a encontrar trabajando a estos láseres.

El esquema que seguiremos es como sigue: en la sección 1 daremos un breve comentario sobre sistemas de excitación, o de "bombeo", tratando de poner énfasis en aquéllos empleados en los láseres que nos ocuparán más en detalle en lo que siga. Las secciones 2 y 3 estarán dedicadas, respectivamente, a láseres en los cuales la emisión sea en frecuencias fijas y a aquéllos en que la frecuencia de emisión sea variable por sintonía. Aquí debemos hacer una aclaración: muchos sistemas moleculares, atómicos o de otro tipo, son capaces de emitir varias líneas, o bandas que pueden ser seleccionadas en diversas regiones del espectro. Sin embargo, sólo algunos sistemas son capaces de variar la frecuencia de emisión (o longitud de onda de emisión) de forma que toda la energía emitida (o una muy buena parte de ella) se canalice en una longitud de onda prefijada (permítasenos, por ahora, esta simplificación de lenguaje). A estos últimos los denominamos sintonizables.

Cada una de las secciones 2 y 3 estará dividida en partes correspondientes, a saber: láseres de emisión continua, de emisión pulsada y una parte dedicada a la transformación de frecuencia por procedimientos ópticos no-lineales.

Para cada una de tales partes, haremos aún una división en láseres de potencia baja, de potencia media y de alta potencia. En este punto será necesaria otra aclaración: claramente, para láseres continuos y para láseres pulsados, los límites de tal división son necesariamente distintos y para no abundar aquí en detalles, trasladaremos la cuestión al capítulo que corresponda. Claro está, será imposible describirlos a cada uno en detalle, de modo que en ciertos casos sólo se los mencionará con alguna referencia, para derivar al lector atraído por aquél.

Una cuarta sección se ocupará brevemente de enumerar y describir en forma extremadamente breve, otros tipos de láseres que hayan quedado sin tratar.

La sección 5 tratará de resumir en forma de esquemas, todas las anteriores.

Otro de los temas que merecen un capítulo aparte, es el del desarrollo tecnológico de los láseres, en los que el acento está puesto en la durabilidad y funcionalidad de los elementos que componen al sistema láser.

En este sentido es interesante estudiar, entre otros temas, los sistemas especializados en los que sus prestaciones vienen especificadas para un restringido margen de aplicaciones, lo cual logra abaratamiento en los precios y mayor durabilidad de los componentes, a costa de una aparente pérdida de versatilidad.

1 - Procedimientos de excitación

El medio activo de un láser puede ser sólido (en realidad, iones en matrices sólidas o amorfas), líquido (moléculas en solución), o gaseoso (átomos, iones o moléculas en estado gaseoso). Para que estos sistemas puedan emitir radiación, los mismos deben ser excitados partiendo de sus niveles fundamentales, a los que llamamos nivel superior de la emisión láser y para que tal radiación pueda ser considerada emisión estimulada, debe existir al menos un desbalance neto entre las poblaciones del nivel láser superior y el nivel láser inferior, por lo que la excitación debe ser lo más selectiva que se pueda.

Existen dos métodos que se aplican en prácticamente todos los láseres conocidos; éstos son el método de excitación óptica y el de excitación por descarga eléctrica. Por lo general, el primero es especial para sólidos y líquidos, mientras que el segundo se aplica a gases y semiconductores. Existen otros tres métodos: el de excitación química, el de expansión de un gas a altas temperaturas y el de fotodisociación atómica, que no serán tratados aquí.

1 - 1. Bombeo óptico.

Cuando un fotón posee la energía suficiente como para provocar una transición óptica desde un nivel fundamental a uno excitado, en un sistema atómico o molecular, dicho sistema ha sido excitado ("bombeado") ópticamente. Si tal sistema invierte la población de un nivel respecto a un tercero, o un cuarto, el método de excitación es el adecuado para que, completadas otras operaciones, se pueda operar un láser.

El método de bombeo óptico, como ya anotáramos, es particularmente adecuado para materiales en matrices sólidas y/o líquidas, no siéndolo tanto en gases. Esto es debido a que los niveles de los iones incluidos en cristales son, como sabemos, ensanchados por el campo cristalino (caso de iones Cr^{+3} en zafiro [rubí], por ejemplo) o anchos de por sí (caso de moléculas de colorantes [dyes] en solución); este ensanchamiento da mayor probabilidad para que un fotón sea absorbido por la transición en cuestión. Veremos más adelante algunos ejemplos de esto.

Un sistema muy usado para el bombeo son las lámparas de flash, en las que un condensador cargado a tensiones que varían entre los 2 kV y los 8 kV se descarga, gracias a un disparo de alta tensión auxiliar aplicado a la lámpara que pre-ioniza el gas dentro de ella (usualmente Xe o Kr a presiones entre los 100 y los 400 Torr), y provoca que toda la carga del condensador circule a través de dicho gas, haciendo que los niveles de corriente alcancen valores del orden de 1 kA en tiempos que varían entre los 0,1 y los 10 μs (la duración está condicionada por la carga del condensador y la resistencia del plasma formado en la lámpara).

La lámpara, en estas condiciones, se puede considerar un "cuerpo negro", es decir, con una distribución espectral a la cual se puede asignar una determinada temperatura; tal distribución espectral debe ser lo más aproximada a la curva de absorción del material bajo bombeo. La radiación emitida por la lámpara debe ser canalizada al material de bombeo. Esta transferencia es altamente dependiente de la geometría en la que se dispongan lámpara y material a bombear. Por otra parte, no todos los fotones que son utilizados para excitar el material son transformados por el material para poblar el nivel láser superior (es decir, el de interés para el diseñador). De esta manera existen tres tipos de cuestiones a optimizar, que llamaremos eficiencias, a saber:

i) Eficiencia de transferencia: es decir el cociente entre la potencia que realmente es absorbida por el material a excitar (atención, no sólo por el nivel de interés) y la potencia emitida por la lámpara. Existen muchos modos de transferir lo emitido al material, cada uno mejor adaptado a ciertos láseres.

ii) Eficiencia espectral: el material absorberá ópticamente una parte de la potencia total absorbida. Este número depende del solapamiento entre las curvas de absorción del nivel de interés y la de emisión del sistema de bombeo. Para mejorar este número, 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.

iii) Eficiencia cuántica de bombeo: es decir, el 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 este número no sea la unidad se justifica debido a que existen otros canales de desexcitación del nivel al que fue bombeado el material. Sólo en muy contados casos se puede mejorar este número. Estos pueden expresarse en función de energías.

Por último digamos que, en determinadas condiciones, se requiere que la distribución de potencia sea lo más uniforme posible, para evitar tensiones térmicas o mejorar la calidad óptica del haz láser de salida.

Veamos en algunos ejemplos gráficos, de qué se tratan estos números.

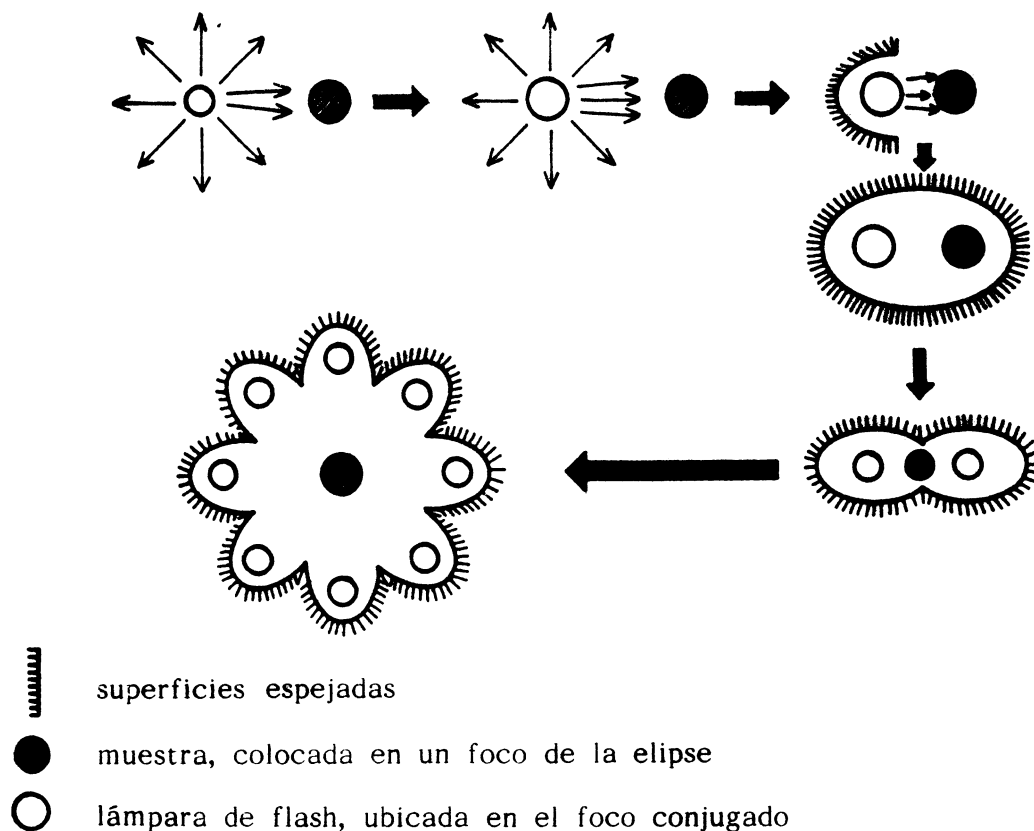


Figura 1. a) Ejemplo 1: Cómo mejorar la eficiencia de transferencia por un mejor diseño de acoplamiento radiación emitida - material a excitar.

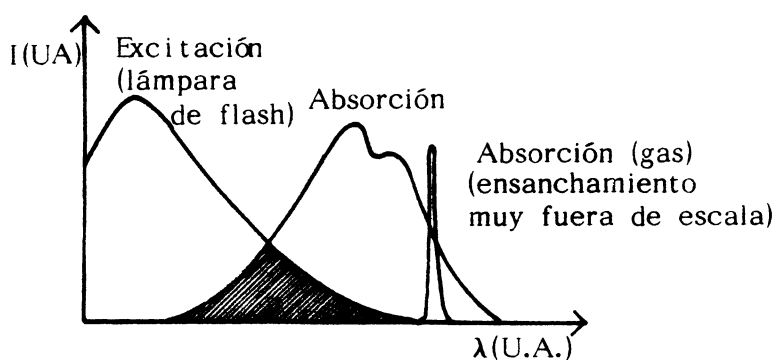


Figura 1. b) Ejemplo 2: Eficiencia de distribución espectral. Se evidencia aquí que la lámpara es muy fría para el sistema a ser excitado. Se superpone una imaginaria "curva" de absorción de un gas.

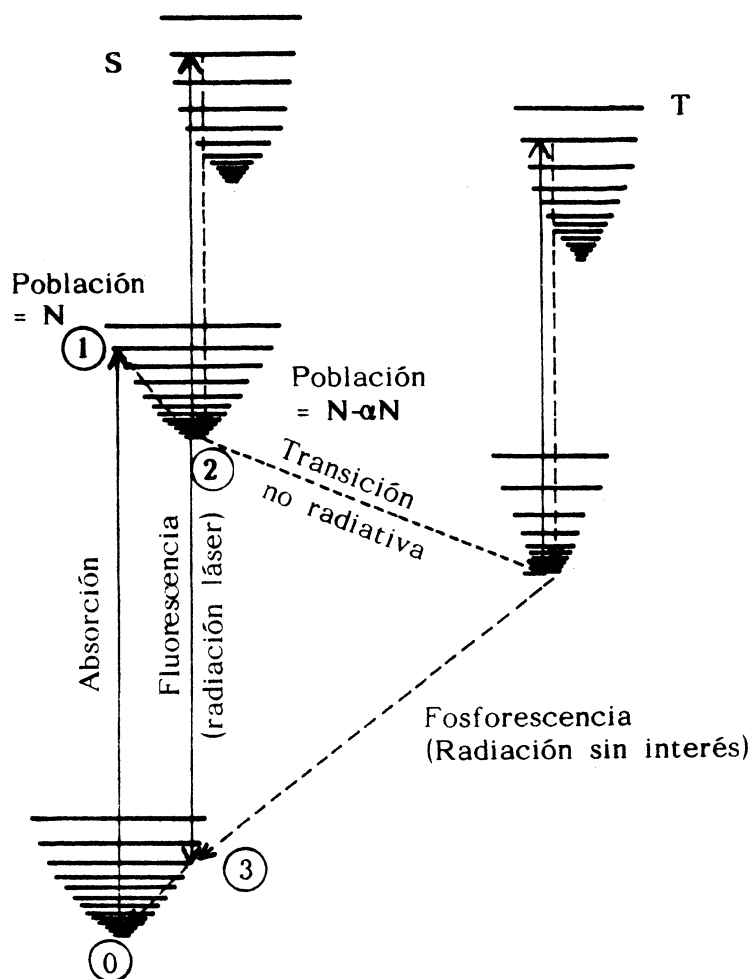


Figura 1. c) Ejemplo 3: Eficiencia cuántica. No todas las moléculas excitadas del nivel 0 al 1 terminan poblando el nivel de interés (el 2) y del 2 no todas se transfieren al 3 en forma radiativa. Los niveles (T) son metaestables por lo general, y la radiación proveniente de ellos al nivel 3 tiene una duración que puede llegar a ser del orden del segundo.

El bombeo óptico suele ser efectuado por medio de láseres que emiten en zonas espectrales de gran absorción del material. Tal el caso de los láseres de colorantes CW bombeados por un láser de Ar^+ . Otro sistema usado en láseres de estado sólido CW son las lámparas de Hg o de iodo-tungsteno; en este caso, aparte de los obviamente necesarios sistemas de refrigeración, las características generales responden a las descriptas.

1 - 2. Bombeo eléctrico.

Por descarga eléctrica en un gas entendemos la generación de una corriente en el seno del mismo, mediante la aplicación en forma continua o intermitente de una determinada tensión sobre él. La aplicación de una tensión continua sobre un tubo de descarga cargado con gas a la presión apropiada produce, si es suficientemente elevada, un fenómeno conocido como "ruptura" del gas, el que pasa a conducir con características de baja resistencia, por lo que la tensión cae abruptamente, manteniéndose en un nivel por lo general muy bajo, para estabilizar la corriente que circule en el tubo. En caso de descargas pulsadas el fenómeno es análogo, excepto que tal descarga puede diseñarse de modo de optimizarla para los fines buscados. La descarga de una lámpara de flash es precisamente un ejemplo de excitación eléctrica. Los procedimientos usados para generarlos se presentan en la Figura 2.

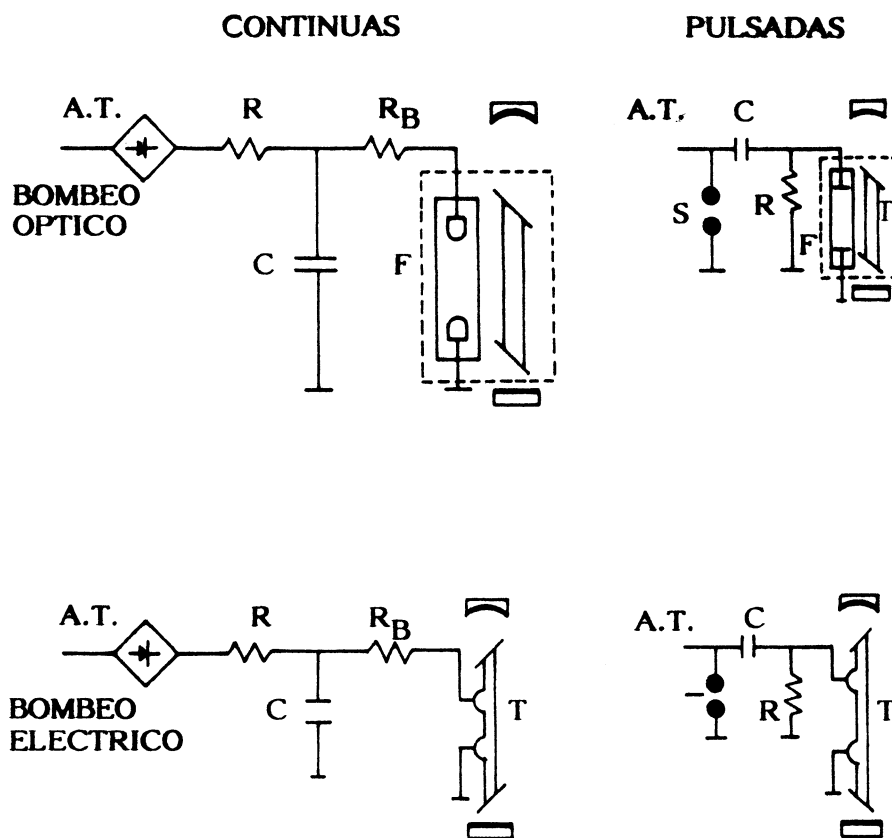
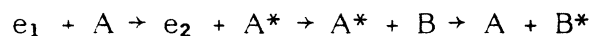
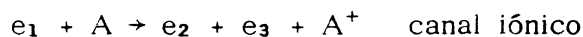


Figura 2

En el caso de descargas continuas, una alta tensión (A.T.) rectificada, descarga a través de una resistencia (R, R_B), sobre una lámpara de flash (F) o sobre un tubo de descarga (T). El condensador **C** sirve para

fijar la tensión sobre el canal de descarga. En el caso de descargas pulsadas, un condensador cargado a una tensión dada, descarga a través de una llave rápida (S) sobre una lámpara de flash o sobre el canal de descarga T.

El proceso de formación de la descarga está estrechamente relacionado con el de excitación de una especie atómica o molecular por impacto electrónico. Los electrones, acelerados por el campo eléctrico externo aplicado (E), adquieren una energía que disipan chocando con los átomos o moléculas del gas según diversos canales; los principales son: choques elásticos en los que la energía de ninguno (ni proyectil, ni blanco) resulta afectada, canal inelástico en donde se produce una excitación del sistema blanco, con la consecuente pérdida de energía de los proyectiles e ionización (que es casi un caso particular del anterior). La energía que adquieren los electrones en una descarga está relacionada con la razón E/p y la energía media de los electrones en la descarga debe optimizarse para la excitación selectiva de determinadas especies. Por otra parte, existe otro canal que denominamos de transferencia resonante de energía, en el que el electrón excita un átomo de una especie a un determinado nivel; a su vez éste, por colisiones, excita la transición buscada en el átomo B.



Resumen de procesos de excitación

2 - Láseres de frecuencia fija. 1. De emisión continua.

2.1.1. De potencia baja:

El láser de He-Ne, uno de los láseres más populares, pertenece a esta categoría. El medio activo es el gas Ne en mezcla con el He; la excitación es por descarga continua de corriente a través del gas, la que excita al He a dos niveles (el 2^3s y 2^1s) que excitan por resonancia a los niveles $2s$ y $3s$ del Ne, respectivamente; éstos se pueblan de manera que responden a las condiciones necesarias para convertirse en niveles de emisión láser. Efectivamente, existen tres transiciones clásicas, a saber, la $3s_2 \rightarrow 3p_4$ ($\lambda_1 = 3.391 \text{ nm}$), la $3s_2 \rightarrow 2p_4$ ($\lambda_2 = 632,8 \text{ nm}$) y la $2s_2 \rightarrow 2p_4$ ($\lambda_3 = 1.152 \text{ nm}$). En esta última funcionó el primer láser de

emisión continua y el primer láser gaseoso. Los sistemas comerciales más usuales estaban preparados para funcionar preferentemente en λ_2 . Para más detalle ver Figura 3.

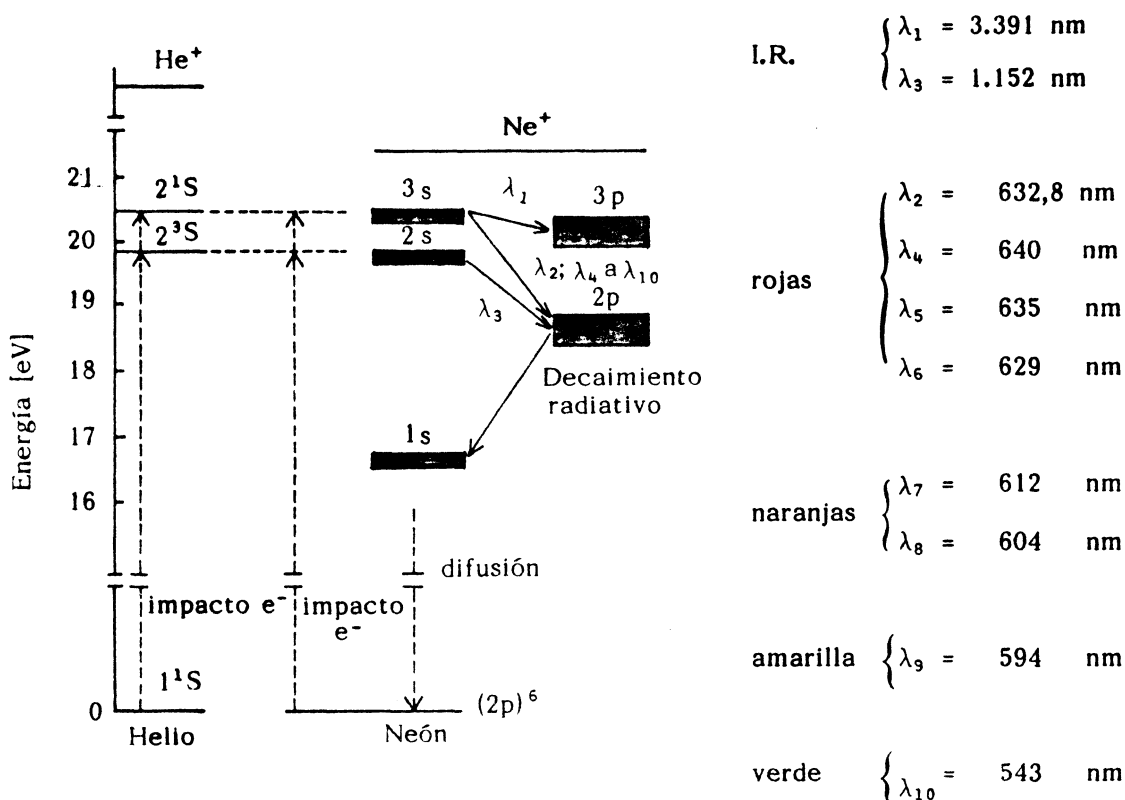


Figura 3. Esquema de niveles del Ne y del He. Se marcan las transiciones láser tratadas en el texto.

Debido al movimiento térmico de los átomos, la línea de 632,8 nm tiene un ancho establecido en el orden de los 1700 MHz, siendo los ensanchamientos de otra naturaleza muy poco importantes. Si el tubo del láser se hace de unos 30 cm de longitud, tres modos longitudinales oscilarán simultáneamente. Existen versiones comerciales con tubos de 10 cm de longitud, para lograr emisión monomodo longitudinal. Para impedir la oscilación de la línea $\lambda = 3.391 \text{ nm}$, se coloca al tubo en una zona de campo magnético inhomogéneo.

Las presiones parciales de cada gas de la mezcla, así como el diámetro interno del tubo de descarga y la densidad de corriente tienen valores óptimos, debido a los complejos procesos que intervienen en la emi-

sión. Algunos valores característicos son: presión de He: 1 Torr; presión de Ne: 0,1 Torr; diámetro interno del capilar: 1,5 mm; corriente: entre 5 y 50 mA.

Los espejos que forman la cavidad resonante son de multicapas dieléctricas (hasta 23 capas) capaces de llegar a una reflectividad del 99,9% en una longitud de onda determinada, y un espejo de menor reflectividad (aproximadamente 99%) como acoplamiento externo.

Pueden encontrarse comercialmente equipos capaces de entregar hasta 75 mW en 632,8 nm en tubos de más de 1 m de longitud, pero las potencias típicas son del orden de 5 mW. La frecuencia de oscilación puede ser estabilizada hasta 1 parte en 10^{12} .

Actualmente se pueden adquirir equipos comerciales capaces de emitir radiación láser en 13 diferentes líneas que cubren el espectro desde 611,8 nm hasta 1,52 μm , que pueden ser seleccionadas individualmente. En algunos se logran varios mW de potencia de salida en el rojo (632,8 nm), 0,5 a 1 mW en el naranja (611,8 nm o en 604 nm), varias décimas de mW en el amarillo (594 nm) o por encima de 100 μW en el verde (543 nm). Las restantes líneas pertenecen al infrarrojo (cercano a 1,1 μm).

Recientemente apareció en el mercado un láser de He-Ne de pequeño tamaño: sólo 3 cm de longitud, con emisión en 632,8 nm y distintas potencias, que van desde 0,5 a 30 mW.

Este láser puede ser considerado una fuente barata de luz de alta calidad. En ámbitos industriales y comerciales, el láser de He-Ne es ampliamente utilizado en alineación, interferometría, lectura de código de barras, "data display", shows artísticos, impresión y otras aplicaciones.

2.1.2. De media potencia.

El ejemplo más conspicuo de esta clasificación es el láser de Ar^+ . Este es un láser en el que los emisores son los iones del gas Ar, que se generan en la descarga. Para obtener la ionización de los átomos de Ar es necesario proveer de una considerable energía al gas. Esto hace que las fuentes de alimentación deban ser mucho más grandes que en nuestro ejemplo anterior. El proceso de excitación es en dos etapas, una de ionización y otra de excitación del ión, por lo que la producción de Ar^+ depende con el cuadrado de la corriente; además, la corriente debe ser elevada para que el proceso sea eficiente. Al ser mayor la corriente continua que debe circular por el tubo láser, éste sufre el calentamiento extremo, por lo que debe ser construido en materiales especiales (de óxido de berilio o segmentos de aleaciones especiales de tungsteno) y con sistemas adecuados de refrigeración. Esta puede obtenerse por circulación de agua a presión (para potencias altas) o circulación forzada de aire para potencias bajas. La corriente elevada produce el fenómeno de "gas pumping" que hace necesario un sistema de estabilización de la presión en ambos extremos del tubo, ya que se produce una migración hacia el cátodo.

La emisión de este láser comprende diversas longitudes de onda, siendo las más importantes la de $\lambda 488$ nm y la de $\lambda 514,5$ nm que se llevan cerca del 80% de la potencia total emitida por el láser. Otras líneas visibles están en 529 nm y hacia longitudes de onda menores, en 455 nm. Con óptica de cuarzo cristalino, puede conseguirse salida en la región UV en 350 nm, con potencias del orden de 3 W. Dada la gran corriente que circula por el tubo, el ensanchamiento Doppler es mucho mayor que en el caso del láser de He-Ne y no resulta posible hacerlo emitir como monomodo.

Un tubo comercial típico de óxido de berilio mide aproximadamente 60 cm de longitud, 3 mm de diámetro interno, cargado de Ar a 0,2 Torr. La corriente que circula es estabilizada y del orden de 20 A a 300 V de caída en el tubo. En este láser no existe un diámetro óptimo como en el caso del He-Ne y existen tubos de hasta 1 cm de diámetro. En el láser se utiliza un campo magnético en dirección del tubo, que disminuye la probabilidad de colisiones de electrones con el tubo, lo que significa un aumento neto de la densidad de corriente.

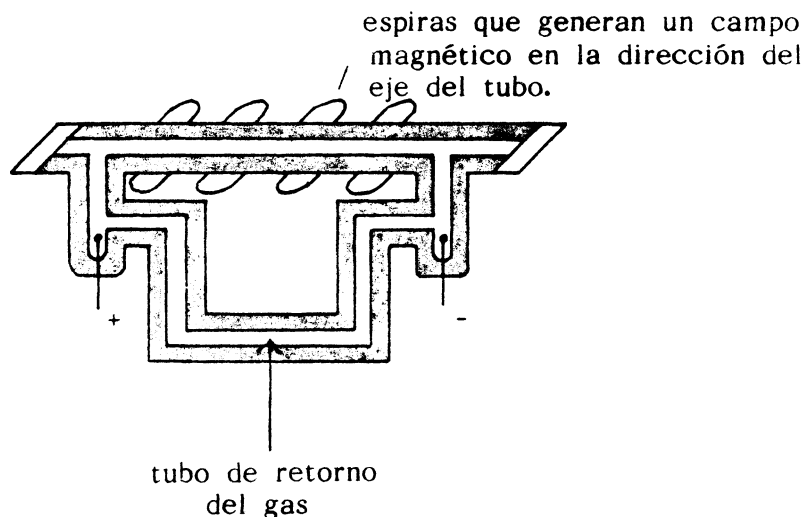


Figura 4. Diagrama esquemático de un tubo láser de Ar^+ .

Los láseres iónicos de argón comerciales poseen potencias típicas de salida en el visible por encima de los 20 W y recientemente se han obtenido potencias de 30 W. Con equipos experimentales se llega a un centenar de vatios de potencia CW.

A pesar de las grandes dificultades tecnológicas que presenta la construcción y funcionamiento confiable de éstos, se han convertido en elementos indispensables en la investigación y en campos como el de la medicina láser, sobre todo por su aplicación en oftalmología en el tratamiento de retinopatías diabéticas, en capsulotomías, en cirugía de cataratas,

etc. Recientemente se lo ha utilizado para recanalizar las arterias coronarias, vaporizando las placas que obstruyen el flujo sanguíneo normal.

El láser de argón permite el análisis automatizado del DNA. Se utiliza como fuente de luz que provoca la fluorescencia de material colorante adherido al DNA. Este instrumento representa una herramienta indispensable en investigación genética.

Una aplicación importante que se encuentra en desarrollo es la técnica de deposición química asistida por láser, para trazar líneas de interconexión entre bloques de circuitos integrados, que tiende a reducir las distancias recorridas por la señal entre "microchips", aumentando así la velocidad de procesamiento. Un argón azul enfocado en unos pocos μm incide sobre el circuito dentro de una celda de deposición de vacío. El láser causa una reacción fotoquímica sobre el hexafluoruro de tungsteno inyectado en la celda, provocando la deposición metálica buscada. Tanto éste como el láser de He-Ne llegan a una eficiencia total de $5 \times 10^{-2} \%$.

2.1.3. De alta potencia.

Los dos tipos más significativos son el láser molecular de CO_2 y el de Nd^{+3} que utiliza como emisor los iones de Nd^{+3} que son sustitutivos en un YAG. Dado que el primero es quizá el de mayor relevancia en el campo tecnológico, a la vez que el láser gaseoso de mayor eficiencia que se conoce, trataremos sólo éste. El segundo es una modificación de otro de régimen pulsado que será oportunamente descrito; algunos datos referentes a él se dan en las tablas al final del capítulo. El láser de anhídrido carbónico utiliza como medio activo tal molécula triatómica por lo general en mezcla con otros gases, normalmente N_2 y He y es, como en los dos casos anteriores, excitado por una descarga eléctrica continua.

Sin adentrarnos en la dinámica molecular, diremos que en una molécula existen transiciones entre niveles electrónicos, de energías comparables a las atómicas o iónicas ya tratadas. Pero además, por tratarse de sistemas poliatómicos, pueden existir transiciones entre niveles vibracionales y rotacionales de la molécula, que surgen del hecho de que tanto las vibraciones como las rotaciones están cuantificadas. En estos casos la energía de los fotones emitidos es mucho más pequeña. Cuando un láser opera en transición vibrorrotacional, la longitud de onda de emisión es de infrarrojo. En el caso del láser de CO_2 la longitud de onda típica es de $10,6 \mu\text{m}$; sin embargo, la emisión no se realiza como líneas únicas sino como bandas o conjuntos de líneas que corresponden a transiciones entre diversos niveles rotacionales de un mismo sistema vibracional. Puede demostrarse que en dicho conjunto de transiciones existen reglas que admiten sólo aquéllas en que el estado rotacional cambia en un cuanto en más o en menos. Esto describe al espectro como ramas (P ó R si el cambio es en +1 ó -1 respectivamente). En el caso del CO_2 la más importante es la rama P y dentro de ella las líneas de mayor intensidad son las de $10,57 \mu\text{m}$ (P18); $10,59 \mu\text{m}$ (P20) y $10,61 \mu\text{m}$ (P22), separadas ellas por

55 GHz. La excitación selectiva del CO_2 se puede lograr por dos caminos: por impacto electrónico directo o por transferencia resonante a través de colisiones molécula-molécula. En este último caso, el N_2 actúa (como el He en el caso del He-Ne) para transferir su energía en forma resonante a las moléculas del CO_2 (ver Figura 5).

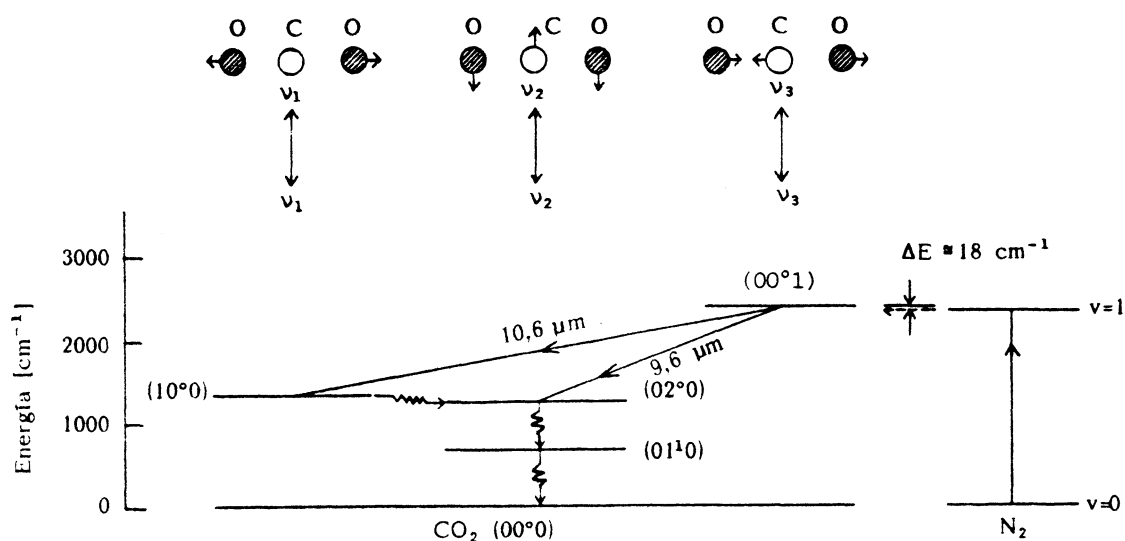


Figura 5. La molécula de CO_2 tiene tres modos vibracionales. Cada modo vibracional tiene niveles de energía característicos.

El He en este caso se usa como gas para disipar la energía térmica de las moléculas y desagotar el nivel láser inferior. Las proporciones típicas usadas en este láser son 1 Torr CO_2 , 1,5 Torr N_2 y 5 Torr He. Este láser es sintonizable en el sentido que puede operar en alguna de las longitudes de onda mencionadas, según se fijen determinados parámetros experimentales. Este ancho de línea es extremadamente menor que en el de los casos precedentes, a saber 50 MHz debido a la longitud de onda de emisión y masa molecular; así, en un tubo de 1 m, sólo puede oscilar un modo longitudinal. En los aparatos de descarga longitudinal tradicionales, para mantener baja la temperatura y eliminar los gases indeseables (como CO y O_2) formados en la descarga, se debe circular la

mezcla en modo de vaciar completamente el tubo en 0,5 s. Cuando se utilizan elementos filtrantes y catalizadores, puede recircularse la mezcla. A mayor longitud del tubo, mayor potencia. Se obtienen hasta 7 kW en tubos de 200 m y quizás se llegue a 1 MW. En tubos de aproximadamente 3 m, se obtienen 185 W en equipos adquiribles comercialmente. El problema que plantean estas potencias en tal longitud de onda es de implicancia tecnológica.

La corriente que se hace circular es de unos 36 mA (en tubos de 1 m de longitud y 15 mm de diámetro interno) a una tensión de 7 kV DC.

Los espejos usados son de sustrato metálico (acero inoxidable, Cu al berilio, molibdeno) pulido para el reflector trasero y un espejo del 20% de transmisión (para 1 m de longitud) hecho sobre sustrato de Ge con "antireflection coating" de un lado y varias capas dieléctricas del lado que debe oficiar de espejo como acoplador externo. La eficiencia alcanza el 15 - 20%, siendo sólo superada por los láseres de semiconductores.

Actualmente se ha desarrollado una técnica original conocida como "flujo axial turbulento" con la cual es posible obtener potencias continuas en el orden de los 10 kW con un diseño de tubo doble y una longitud efectiva de 4 m. Una mezcla de gases de 2,5% de CO_2 , 46% de N_2 y 51,5% de He se hace circular por los tubos a una velocidad de Mach 0,4 (170 m/s), lo que provoca turbulencias aerodinámicas controladas que se utilizan para la esabilización de la descarga eléctrica. Esta se consigue con una tensión de 27 kV y 400 mA, obteniéndose eficiencias entre 20 - 25%.

Este láser con tales potencias es de gran aplicación en distintas áreas que van desde la fusión hasta cortes y soldaduras de precisión, pasando por medicina, fotoquímica, etc.

Se ha desarrollado un proyecto de fusión por láser en Los Alamos ("Antares"), consistente en 24 haces que al dispararse simultáneamente producen una potencia total de 20 TW.

Por otro lado, hoy existen nuevos procesos que usan al láser de CO_2 para la fabricación de polvos de cerámica que se usan en la manufactura de componentes cerámicos. Las cerámicas constituyen una gran familia de materiales inorgánicos, no metálicos, con un vasto campo de aplicaciones que van desde utensilios de cocina hasta componentes para computadoras y sistemas de comunicaciones.

En la actualidad se consiguen con este láser partículas uniformes de cerámica de 0,2 μm , en contraste con las que se obtenían por métodos tradicionales de 0,2 a 1,0 μm .

2.2. De emisión pulsada.

2.2.1. De potencia baja:

En este tipo entran los láseres de semiconductores operados en régimen pulsado. El interés de estos láseres ha sido superado por el de los láseres de semiconductores de emisión continua con sistema de acoplamiento de modos, que los hace más interesantes. Dejaremos su descripción entonces para otra oportunidad.

2.2.2 De potencia media:

Si definimos media potencia aquélla que llega a los 100 kW por pulso, el representante más conspicuo de este tipo es el láser de iones de Xe.

Este láser corresponde aproximadamente al láser de Ar^+ ya descrito excepto que las transiciones utilizadas corresponden al tercer ion del Xe (Xe IV), según se ha establecido recientemente. La emisión más notable es en la zona azul-verde del espectro y comprende varias líneas.

El sistema de excitación es similar a la descarga de una lámpara de flash, excepto que la baja presión interna (~ 100 mTorr) hace innecesario el pulso de preionización.

Disparando condensadores de 100 nF cargados a 10-12 kV por un tubo de 10 mm de diámetro y 1 m de longitud puede llegarse a potencias del orden de 1 kW. Este láser puede operar a frecuencias de repetición del orden de los 100 Hz, por lo que también se hace necesario proveer un retorno a los iones que migraron al cátodo para equilibrar la presión en el tubo.

La duración del pulso luminoso en las líneas del azul-verde, en dicho sistema de excitación puede llegar a los 200 ns (en cambio, la emisión del Xe neutro (Xe I) en el infrarrojo es mucho menor para los mismos espejos que forman la cavidad). Los espejos usualmente son de multicapas dieléctricas (proveen alta reflectividad y baja absorción), pero también pueden obtenerse potencias respetables con espejos de depósito metálico (Au, Ag, Al), obviamente de menor precio).

Dada la potencia y la longitud de onda de emisión, este láser puede ser utilizado como elemento de excitación de otros láseres (v.g. los de colorantes, como veremos en una sección posterior), para el cortado de films en semiconductores, corrección de fotomáscaras, análisis de fallas de circuitos, interconexión de circuitos integrados, etc.

2.2.3. De potencia alta:

Numerosísimos son los láseres que pueden ser ubicados en este tipo. Si partimos desde potencias del orden de 1 MW en adelante, discriminándolos en orden de energía creciente, ubicaríamos el láser de nitrógeno molecular (N_2), los láseres de dímeros y halógenos-gases raros, de rubí (iones de Cr^{+3} en corindón), de CO_2 en descarga transversal a presión atmosférica (TEA) y el ya mencionado Nd^{+3} .

Describiremos sólo uno de ellos, a saber, el de Nd^{+3} , ya que, por ejemplo, el de rubí es uno de los más famosos (fue el primero en emitir láser) y es también el de mayor literatura disponible. En cuanto al de CO_2 , ya hemos descrito las características más importantes cuando correspondió a los láseres de CW y podemos decir de los TEA muchas cosas similares a las que diremos de los excímeros en capítulo posterior. Las características de su emisión se resumen en las tablas del final del capítulo. Algo similar puede decirse del de N_2 .

El láser de Nd^{+3} es un ejemplo de estado sólido. Como ya dijéramos, los responsables de la emisión son los iones de Nd incluidos como sustituyentes del Y^{+3} en granates de itrio-aluminio ($Y_3Al_2O_{15}$ (YAG) o como dopantes en vidrios especiales. Cada uno de estos láseres de Nd (YAG o vidrio) tienen características que los hacen óptimos para algunos diseños. En régimen pulsado, el láser de Nd:vidrio es más adecuado que el Nd:YAG que, a su vez, da mejor comportamiento en régimen CW, siendo la razón principal, la mala conductividad del vidrio para régimen continuo. Los niveles de energía del Nd^{+3} se encuentran ensanchados, como puede verse en la Figura 6 y el sistema es apto para ser bombeado ópticamente, tal como dijéramos en la primera parte de este capítulo.

Esquema de niveles del Nd^{+3} (Ejemplo de 4 niveles)

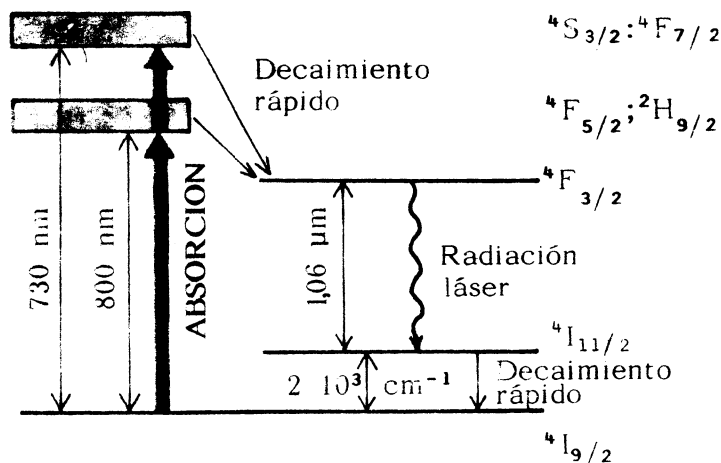


Figura 6

Además de ser un láser de buena eficiencia, la emisión láser de Nd^{+3} es de banda muy ancha (10 nm), por lo que muchos modos pueden oscilar simultáneamente y por lo tanto usarse para acoplar y producir pulsos del orden de 10^{-12} s.

Este láser ha sido utilizado en grandes cadenas de amplificación, obteniéndose en sistemas como el SHIVA (de 20 haces) hasta 2×10^{13} W o en haces de 20 cm como el ARGUS de 2 TW. En el proyecto NOVA se logró una potencia de 30 TW con 10 haces en pulsos de 3 ns. En este sistema es posible alcanzar los 100 TW acortando la duración del pulso a 0,1 ns. Todos estos sistemas fueron desarrollados con el fin de lograr fusión controlada por láser.

La longitud de onda más importante de emisión de este láser es alrededor de 1,06 μm ; las barras de material utilizadas en osciladores son de algunos centímetros de largo y 1,2 cm de diámetro en cavidades del orden de los 0,5 m.

El láser de Nd es uno de los más poderosos y desarrollados de los láseres de hoy en día. Ha desplazado por mayor eficiencia y simplicidad en la fabricación de barras, al de rubí.

En régimen pulsado y continuo es interesante la conversión de frecuencia, llegando a duplicarla y triplicarla para obtener radiación de alta potencia y energía en el verde y UV. Se lo considera un serio competidor del Ar^+ en el caso CW.

Se han desarrollado láseres de Nd:YAG Q-switched, doblados en frecuencia de 4 W y 6 W en 532 nm (verde). Aplicaciones típicas incluyen el marcado extra fino y maquinado de núcleos de ferrite para cabezas registradoras.

Los esfuerzos actuales en el láser de Nd:YAG se centran en el aumento de potencia, introducir nuevos diseños y mejorar el control electrónico para facilitar su control por medio de computadoras.

Las tendencias del mercado son hacia pulsos más cortos y más energéticos, llegándose a desarrollar un Nd:YAG Q-switched que trabaja a 1 KHz de frecuencia de repetición, dando una potencia media de 150 W.

Se ha introducido en el mercado un Nd:YAG continuo de 20 W de monomodo y hasta de 100 W multimodo.

Otro problema que actualmente se estudia es la estabilidad de la salida, que es particularmente seria en aplicaciones tales como bombeo de dyes y doblado de frecuencias.

Los láseres de muy alta energía utilizan una configuración "slab-geometry" que debido al bombeo y enfriamiento simétrico a través de las caras del slab, eliminan en primer orden los efectos de distorsión térmica sufridos por las barras cilíndricas comunes. En una variación de este diseño, en la cual hace una serie de reflexiones internas totales (zig-zag), el camino se hace lo suficientemente largo como para que opere como

oscilador. Con esta geometría se obtiene una buena calidad de haz a niveles de potencia de interés para aplicaciones industriales.

Con esta geometría se obtienen velocidades de perforación hasta 5 veces mayores que con láseres de varilla y se han cortado planchas de acero de hasta 2,5 cm de espesor.

Actualmente el uso de láseres de diodos aparece como una alternativa de las lámparas de flash en el bombeo de láseres de estado sólido (ver Figura 7).

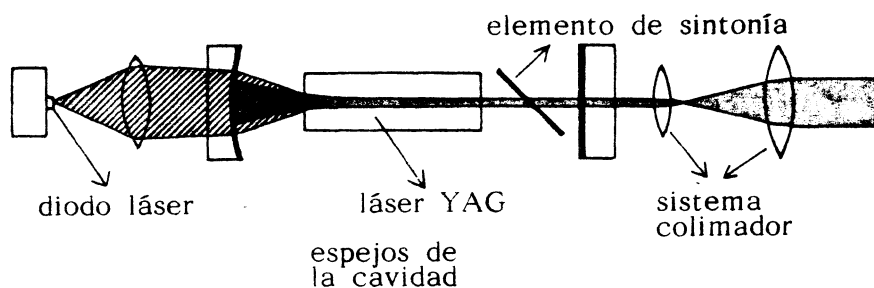


Figura 7

Las ventajas obtenidas incluyen mayor eficiencia, mayor vida útil y menor tamaño. Esta tecnología llegó a mediados de 1986, a su etapa comercial con la introducción de un Nd:YAG bombeado por diodo de 40 mW de potencia en $1,06 \mu\text{m}$ y 15 cm de longitud.

Los desarrollos en fibras ópticas proveen un medio para entregar la radiación de un Nd:YAG en niveles de energía y potencia promedio lo suficientemente altos para procesamiento de materiales, soldaduras de punto, etc.

3 - Láseres de frecuencia variable por sintonía. 1. De emisión continua.

3.1.1 De potencia baja:

En este tipo de láseres se destacan aquéllos en los que el material activo es un semiconductor. La tecnología de este tipo de láseres ha avan-

zado enormemente en los últimos años, permitiendo la aplicación de éstos en un vasto espectro de aplicaciones. Existe una gran variedad de este grupo, por lo que resulta difícil resumir todas las características de los mismos. Hay láseres de juntura (diodos) y de heterojuntura doble, cada uno con sus variedades y cambio de propiedades. En particular, es interesante notar que la longitud de onda de emisión puede ser variada por diferentes métodos, sin pérdida apreciable de potencia: así, es posible obtener emisión en bandas extremadamente angostas. Esta cualidad de sintonizabilidad es la que motiva el presente capítulo.

Para describir el funcionamiento de un láser de semiconductores, debemos hacer hincapié en la naturaleza de los semiconductores. Sabemos que en un átomo podemos describir sus niveles de energía como partícula aislada, cosa que representa bien los átomos en un gas y aproximadamente los iones en una matriz sólida, pero cuando el estado de agregación es tal que la separación interatómica es muy pequeña, se produce un fenómeno conocido como formación de bandas de energía, de modo que el estado en que pueden hallarse los electrones no es único sino que es una colección de estados con muy poca separación entre sí. Este fenómeno colectivo de formación de bandas es distintivo del estado sólido y es el responsable, por ejemplo, de la conducción eléctrica de los materiales.

El esquema de niveles de energía para un semiconductor ideal se muestra en la Figura 8a.

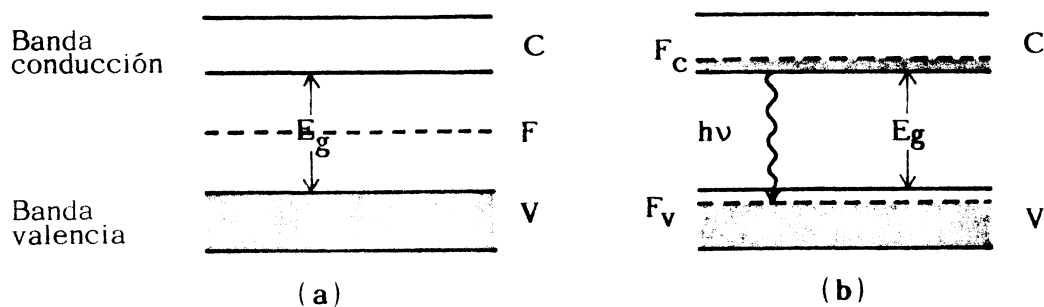


Figura 8

El espectro de niveles de energía consiste en dos bandas: son las bandas de valencia V y las bandas de conducción C , separadas por una región de energías prohibidas ("gap de energía"). Cada banda realmente consiste de un gran número de estados de energía muy cercanamente espaciados. Puede mostrarse que el número de estados en cada banda es del mismo orden que el número total de átomos en el semiconductor. F representa la energía del llamado nivel de Fermi, el que constituye el límite entre los niveles totalmente ocupados y completamente vacíos a $T = 0^\circ\text{K}$ (la probabilidad de ocupación vale 1 para energías inferiores a

F y vale 0 para energías superiores a F). Para semiconductores no degenerados, el nivel de Fermi está situado dentro del gap de energía, estando a 0°K la banda de valencia completamente llena y la de conducción completamente vacía. En estas condiciones, el semiconductor no conducirá y será por lo tanto un aislador.

Supongamos, por simplicidad, que disponemos de un semiconductor a $T = 0^\circ\text{K}$ (Figura 8a), con la banda de valencia totalmente ocupada). Supongamos ahora que de alguna manera podemos llevar electrones de la banda de valencia a la de conducción. Después de un corto tiempo ($\sim 10^{-13}$ seg) los electrones en la banda de conducción caen 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 "agujeros". Esto significa que puede conseguirse "inversión de población" entre C y V para obtener acción láser si el material es colocado en un resonador apropiado. Introduciendo el concepto del llamado "cuasi-nivel de Fermi", se puede obtener la condición de acción láser. Si bien el semiconductor como un todo no está en equilibrio térmico, podemos decir que el equilibrio es alcanzado dentro de cada banda en un tiempo muy breve. Entonces F_V y F_C son las energías de los cuasi-niveles de Fermi, que separan a $T = 0^\circ\text{K}$, la zona totalmente ocupada de la completamente vacía para cada banda. A partir de la Figura 8b puede verse que

$$h\nu < F_C - F_V$$

Para obtener acción láser debe cumplirse que

$$F_C - F_V > E_2 - E_1 = h\nu \quad ,$$

donde E_2 y E_1 son las energías de los niveles superior e inferior respectivamente.

Una forma muy simple de obtener inversión de población consiste en usar un semiconductor en la forma de diodo de juntura p-n, con regiones tipo n y p altamente dopadas ($\sim 10^{18}$ átomos donores o aceptores/ cm^3). La operación de un diodo de este tipo se muestra en la Figura 9. Como el material está altamente dopado, el nivel de Fermi F_p del semiconductor tipo p cae dentro de la banda de valencia y el F_n nivel de Fermi del semiconductor tipo n cae dentro de la banda de conducción. Cuando no hay tensión aplicada, los dos niveles de Fermi yacen sobre la misma línea horizontal (Figura 9a). En cambio, cuando se aplica una tensión V (Figura 9b), los dos niveles se separan en una cantidad dada por:

$$\Delta F = eV$$

Aplicando una tensión directa al semiconductor se produce la "inyección" en la zona de juntura, de electrones de la banda C del material n y agujeros de la banda V del material p .

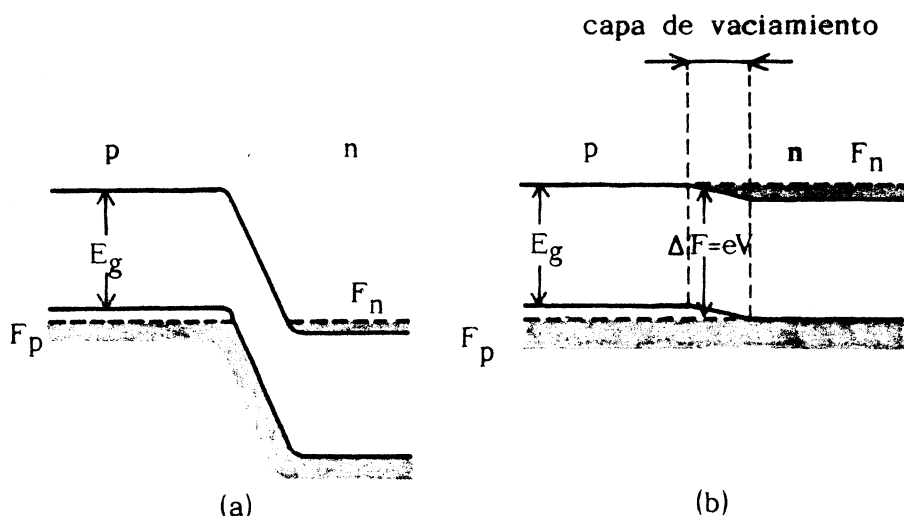


Figura 9

La inversión de población alcanzada tenderá a equilibrarse pasando por un proceso inverso que será de emisión de fotones cuya energía sea del orden del gap energético.

El bombeo puede ser también óptico (mediante lámparas de arco o láseres de longitud de onda menor que el gap), o por haces de electrones (e-beam).

Al comienzo estos láseres podían funcionar sólo a temperaturas muy bajas debido a que la corriente que debe pasar por ellos para obtener la condición de umbral depende muy fuertemente de la temperatura (en la proximidad de los 77°K la corriente (o densidad de corriente) umbral crece con T^3) llegando a hacer imposible la acción láser a temperatura ambiente en algunos casos. A 20°K se obtuvieron en junturas de AsGa, en 840 nm, potencias CW de 1 W y con los de sales de plomo, 0,35 W en operación multimodo.

En estos láseres hablaremos de dos eficiencias, primero la eficiencia cuántica que mide el porcentaje de fotones emitidos por número de electrones que atraviesen la juntura (se puede llegar hasta el 70%) y en segundo lugar, la eficiencia eléctrica, en términos de la potencia de salida en relación a la de alimentación (eléctrica); en este caso se llega a un 10%.

De los distintos tipos existentes, un tipo como el de $\text{Ga}(\text{As}_{1-x}\text{P}_x)$ puede hacerse emitir en 840 nm ($x=0$) a 640 nm ($x=1$). Para los de sales de plomo, eligiendo adecuadamente los componentes y sus proporciones, λ se puede variar entre 1,5 μm y 40 μm .

Para un compuesto dado, la emisión puede sintonizarse por cambio de temperatura, por presión hidrostática, por variación de un campo magnético externo y por sintonía de la cavidad externa. Por variación de presión, un láser de SePb a 77°K fue sintonizado desde 7,5 a 22 μm llevando la presión hasta 14 kbar. Este sistema se usa como sintonía gruesa y la variación por temperatura como sintonía fina. Por control de la cavidad externa usando una red de difracción, se obtienen líneas de hasta 0,2 Å de ancho espectral y se verifica que el ensanchamiento es prácticamente homogéneo, lo que significa en términos útiles, que sin sintonía (o sea en banda ancha) la potencia emitida es prácticamente la misma que la emitida en banda angosta.

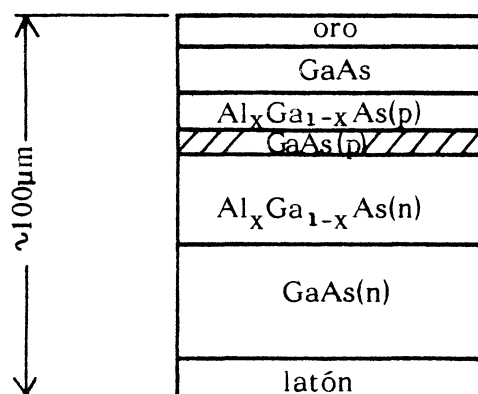


Figura 10. Diagrama esquemático de un láser de semiconductor de heterojuntura doble. La región activa aparece rayada.

En un primer tiempo las aplicaciones estaban limitadas por la desventaja de tener que refrigerar los diodos a bajas temperaturas. Este problema ya fue superado usando, en lugar de un diodo p-n, un diodo de heteroestructura doble. En este diodo hay dos junturas $\text{Al}_x\text{Ga}_{1-x}\text{As}(p)$ - GaAs y GaAs - $\text{Al}_x\text{Ga}_{1-x}\text{As}(n)$ entre diferentes materiales (Figura 10). La región activa consiste de una capa delgada de GaAs (1 μm). La operación continua se debe al efecto combinado de tres cosas: 1) el índice de refracción del $\text{Al}_x\text{Ga}_{1-x}\text{As}$ es menor que el del GaAs , asegurando con esto que la acción láser quede confinada en la capa de GaAs , que es la región de ganancia. 2) La región activa está mucho mejor confinada y es por lo tanto menor. Para la misma densidad de corriente la densidad electrónica en la región activa es incrementada, y con ello, la ganancia.

3) La capacidad de disipación de calor del diodo ha sido considerablemente mejorada. Esto se obtuvo pegando el sustrato de GaAs(n) a una placa de diamante (o latón), el cual actúa como sumidero de calor debido a su masa y conductividad térmica.

Las dimensiones típicas de un láser de semiconductores, sin contar la cavidad resonante (en caso de ser ésta externa), son de 1 mm de longitud, 100-300 μm de espesor total de material, unos 10 μm de región activa y 1 mm de ancho. Figuran por ello entre los láseres con menor región activa conocida.

En el caso de los de heterojuntura doble, la región activa es de alrededor de 1 μm . Por razones de mejor disipación del calor y mejor confinamiento de la radiación y de la región activa, éstos son usados en régimen CW a temperatura ambiente. Actualmente, con estructuras monolíticas donde los emisores forman un paquete de un solo sustrato, se alcanzan potencias de 1 W continuo a temperatura ambiente.

Puesto que los láseres de inyección sólo necesitan fuentes de polarización de 1,5 V y de unos 10-100 mA y su emisión puede ser modulada en amplitud por simple modulación de la corriente de inyección, estos láseres son de gran aplicabilidad en campos como el de las comunicaciones ópticas, debido a su simplicidad y pequeño tamaño.

Se ha desarrollado recientemente una nueva generación de láseres de diodos que pueden usarse como una fuente de bombeo casi ideal para láseres de estado sólido, los que ofrecen la ventaja de una salida con calidad espacial mejor que la de los primeros.

Los láseres de diodo han superado a las lámparas tradicionales en el bombeo de láseres de estado sólido, debido a que el tiempo de vida de aquéllos (superior a 5000 horas) es muy superior al de las lámparas (400 horas). Otra ventaja importante es que los láseres de diodo pueden ser colimados y enfocados. Por otro lado, poseen altas eficiencias y requieren poca refrigeración.

Una aplicación importante es la determinación de la temperatura y concentración de CO en una llama por espectroscopía de absorción, usando un láser de diodo sintonizable.

Se han obtenido recientemente láseres de semiconductores de óptica colimada preajustada para producir un haz paralelo con colimación menor de 0,3 mrad. Algunos modelos han desarrollado una potencia de hasta 20 mW en 820 nm.

Los láseres de semiconductor jugarán un rol más importante en los sistemas de supercomputadoras futuras que operarán a velocidades del orden del gahertz.

Recientemente se desarrolló una nueva generación de láseres de semiconductores: los de "pozo cuántico" (quantum-well lasers), que reemplazarán a los de semiconductor convencionales. Los primeros presentan una corriente umbral considerablemente menor que los segundos, evitando así

retardos en el transporte de la señal, agilizando las interconexiones. Los láseres de "pozo cuántico" consisten en una muy delgada zona activa del orden de $0,010\ \mu\text{m}$ o menos, en la cual el movimiento de los electrones está confinado a distancias extremadamente pequeñas, hecho que modifica la estructura de los niveles de energía, aumentando la ganancia óptica, respecto de los convencionales.

Actualmente se fabrican láseres de pozos cuánticos múltiples con corrientes umbrales tan bajas como 2,5 mA. Con paquetes de veinte emisores en estructura monolítica se alcanzaron potencias continuas de 1 W, y eficiencias del 30%. A fines de 1988 se comercializarán láseres de diodos visibles en 680 nm, con potencias del orden de 3 mW.

Uno de los problemas básicos con que se enfrenta la investigación en este campo, es el del diseño de cavidades acopladas (tanto longitudinal como lateralmente), con el objeto de seleccionar modos de vibración por un lado y "acoplar" en fase varios láseres, por el otro. Si bien el problema del estudio de cavidades acopladas no es nuevo, hoy día ha sido retomado con creciente interés, para aplicarse a este tipo de láseres, ya que aquellas presentan la posibilidad de operación en régimen monomodal (con la consiguiente reducción de ruido y aumento en la velocidad de modulación) y la potencial operación en frecuencia modulada.

3.1.2. De potencia media:

Para ejemplificar un tipo de láseres de régimen CW, de potencias mayores o del orden de 1 W, daremos a continuación un breve examen de los láseres de soluciones líquidas de colorantes orgánicos, los que llamaremos láseres de colorantes. En éstos, el material activo son moléculas orgánicas de gran número de átomos, cuya característica más importante es la aparición de ligaduras entre átomos de carbono dobles y simples alternadas, lo que se denomina ligadura no saturada. Estos compuestos absorben luz en forma muy eficiente en regiones espectrales desde el UV cercano al IR cercano (de 220 a 1700 nm) según su composición; la energía absorbida por éstos es reemitida posteriormente en longitudes de onda mayores en forma de fluorescencia, de corta duración, y de fosforescencia de mayor duración. Ambos procesos dan lugar a la intensa coloración que los caracteriza; sin embargo, para nuestros fines interesa estudiar sólo la fluorescencia, siendo según esta vía que se obtiene acción láser.

Las características de estas moléculas hacen que se presenten dos sistemas de niveles electrónicos (ver Figura 1c), subdivididos a su vez en un número muy grande de niveles vibrorrotacionales, a saber, el sistema de singuletes y el de tripletes. La fluorescencia se produce cuando hay una transición desde el primer nivel singulete excitado al nivel singulete fundamental; la vida media del singulete excitado es del orden de los nanosegundos; la fosforescencia aparece cuando hay transiciones del primer triplete excitado al nivel fundamental; por tratarse de una transición prohibida, la vida media radiativa del nivel triplete es de varios segundos.

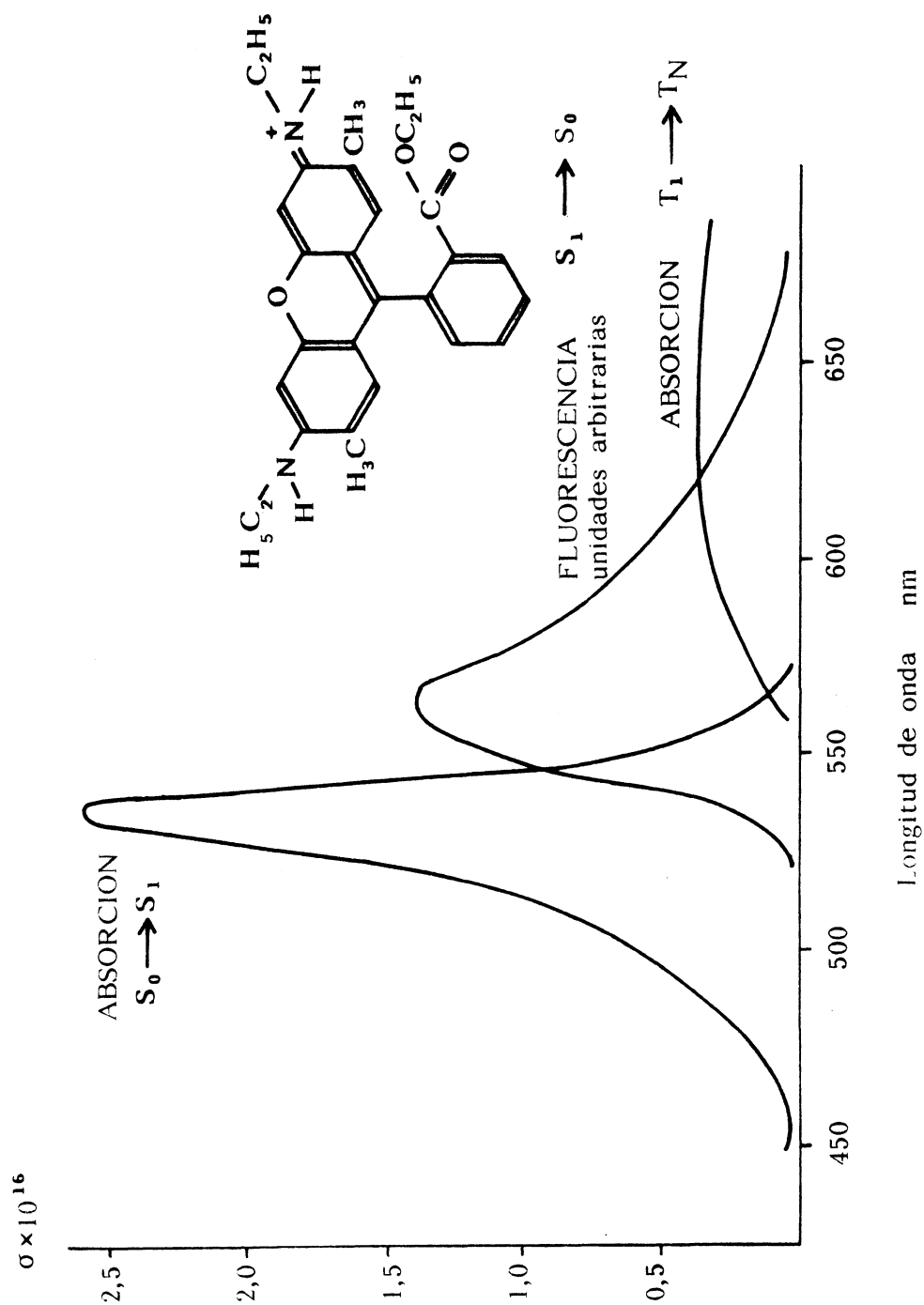


Figura 11

En el esquema de niveles que se representa en la Figura 1c, se ve que este láser es un láser de cuatro niveles y puede comprenderse el motivo por el que la radiación emitida es de menor energía que la absorbida.

Un ejemplo típico de curvas de absorción y emisión fluorescente se presenta en la Figura 11, con el caso de la rodamina 6G, colorante éste que resulta ser uno de los más eficientes y por lo tanto de los más utilizados.

La radiación emitida lo es en una banda ancha que abarca desde 530 a 650 nm. En caso de llegar a la condición umbral, esta radiación será emitida como láser en aproximadamente todo su ancho espontáneo y podrá ser sintonizada tal como dijimos de los láseres de semiconductores, por diversos métodos. Puede lograrse un corrimiento en la banda de emisión, cambiando algún sustituyente en la composición básica de la molécula, cambiando la concentración de la solución o su pH, cambiando el solvente, por variaciones de la presión externa o de la temperatura, la longitud del medio activo o por medios externos, colocando una red de difracción que seleccione la longitud de reinyección, mediante la colocación de un etalón de inclinación variable o filtros birrefringentes dentro de la cavidad resonante y por "inyección" de una longitud de onda externa en la cavidad de estos láseres. Por falta de espacio no podemos considerar ni las causas de que todas estas operaciones lleven a una sintonización de la emisión, ni todas ellas en detalle. Debe recordarse que sintonizar significa hacer que el láser emita en una banda mucho más angosta que su emisión "libre" pero sin pérdida apreciable de energía; esto se debe, como ya hiciéramos mención, al hecho de que los niveles de energía son homogéneamente ensanchados.

En el caso de los láseres que nos ocupan ahora, es decir, en régimen CW, hay que prestar mucha atención a los fenómenos que significan pérdidas en la cavidad; la utilización de cavidades especiales que puedan compensar pérdidas adicionales por astigmatismo, que luego enunciaremos y fundamentalmente problemas que surgen de efectos térmicos.

El láser CW de colorantes más utilizado es el de la cavidad plegada bombeado por un láser de Ar^+ CW con el colorante fluyendo por tobera. Este diseño consiste, como indica la Figura 12, en una cavidad de tres espejos "plegada" a un cierto ángulo, en un punto de la cual y según el ángulo de Brewster, fluye el colorante libre de celdas y paredes en forma continua. El haz de bombeo enfocado por un cuarto espejo en un punto del "jet", excitando el colorante; sin embargo, este espejo debe ser usado "inclinado" y por lo tanto aparece una deformación del foco que, de no ser compensada por la cavidad, introduciría graves pérdidas en el láser.

Con láseres de este tipo pueden obtenerse varios vatios de potencia sintonizable en 50-90 nm con anchos de línea menores que $0,2 \text{ \AA}$, es decir, que puede hacerse oscilar en unos pocos (hasta en uno) modos longitudinales de la cavidad. Bombeando el dye con un Ar^+ de modos acoplados, puede llegarse a tener pulsos extremadamente cortos, del orden de

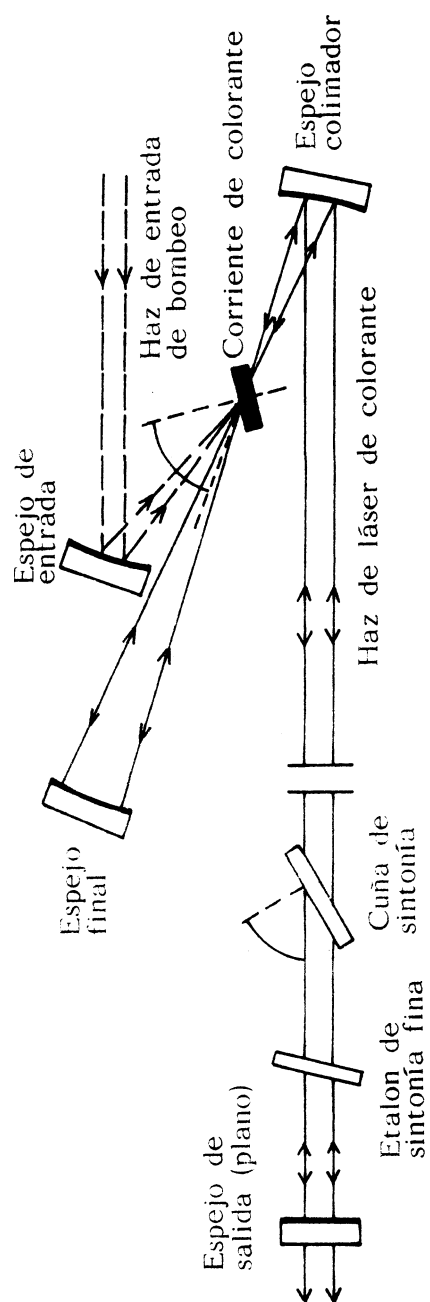


Figura 12

0,5 ps y sintonizables.

La eficiencia de estos colorantes, medida como potencia de salida a potencia de entrada (de bombeo), puede llegar hasta un 50%.

Todas las características mencionadas hacen de estos láseres instrumentos de gran utilidad en diversos campos de la tecnología y la ciencia básica.

3.1.3. De potencia alta:

Si bien no es tan notoria la sintonizabilidad como en los casos anteriormente ilustrados, el láser de CO_2 puede, como ya se dijo, ser sintonizado por medios externos, dentro de ciertos límites.

3.2. Emisión pulsada.

3.2.1. De potencia baja:

Por debajo de los 100 W de pico de potencia emitida, podemos citar como caso típico los láseres pulsados de semiconductores. Los métodos de excitación no difieren en su naturaleza de los utilizados para régimen CW, excepto que son pulsados. Naturalmente, los problemas de calentamiento pueden ser menores siempre que las frecuencias de repetición no sean altas.

Un láser de AsGa puede llegar a 3 W de potencia pico cuando se sintoniza su emisión por red a un ancho de 0,2 Å y 3,5 W en una banda de 4 nm de ancho. Puede llegarse a potencias de 20-40 W a temperatura ambiente con sistemas especiales y el intento es llegar a los 100 W de potencia pico.

3.2.2. De potencia media:

En el rango de más de 1 kW hasta 1 MW, nos ceñiremos al ejemplo de los láseres de colorantes, cuyas potencias, según sean excitados por lámparas longitudinales de flash o coaxiales o láseres de alta potencia, pueden llegar a 15 kW, 11 MW o 2 MW respectivamente, con energías del orden de los 5, 10, 0,2 J, en ese orden. Los anchos temporales de pulso están en relación con los números mencionados. A frecuencias de repetición entre 50 y 60 Hz se han obtenido más de 200 W de potencia RMS en un sistema de láser de colorante bombeado por lámparas longitudinales de flash; se logran por sintonía, líneas de 19 Hz de ancho.

Este tipo de sistema es muy usado para obtener pulsos ultracortos (~ 1 ps) sintonizables y de alta energía (~ 1 mJ aproximadamente).

Aparte de los problemas técnicos asociados con el manejo de pulsos eléctricos y ópticos de alta potencia, no existe ningún otro particular respecto a estos láseres bombeados por lámparas de flash. Los sistemas bombeados por láseres pulsados de alta potencia (por 2da. armónica del Nd, por láser de N_2 , por láseres de excímeros o láser de Xe), presentan una gran variedad de configuraciones de bombeo entre las que se destaca la configuración denominada de Hänsch, con la cual ha sido posible obtener pulsos de 50 kW sintonizables en un rango variable según la variación en la presión en un sistema de etalones, y anchos de línea tan angostos como 6×10^{-3} nm, con doble etapa amplificadora del láser de colorante.

3.2.3. De potencia alta:

Por encima del régimen de 1 MW de potencia pulsada sintonizable podemos reincidir sobre los colorantes; sin embargo, aunque el rango de sintonizabilidad sea limitado, preferimos hacer una breve mención a los láseres de excímeros, que son actualmente los representantes de una nueva generación de láseres de alta potencia.

Dichos láseres han llegado a pulsos del orden de 1 kJ de energía en el ultravioleta ($\lambda 249$ nm) y eficiencias del orden del 1%, o bien potencias medias del orden de 25 W a 100 Hz de frecuencia de repetición, con lo que resultan las fuentes de mayor energía en el UV a buenas eficiencias de conversión.

El medio activo de dichos láseres lo constituyen sistemas moleculares denominados dímeros o complejos excitados, según se trate de sistemas homo o heteronucleares, respectivamente. Estos no presentan estado fundamental ligado, por lo que puede decirse que existen sólo en estado excitado. Un esquema ilustrativo de la formación y emisión de esas "moléculas" se muestra en la Figura 13.

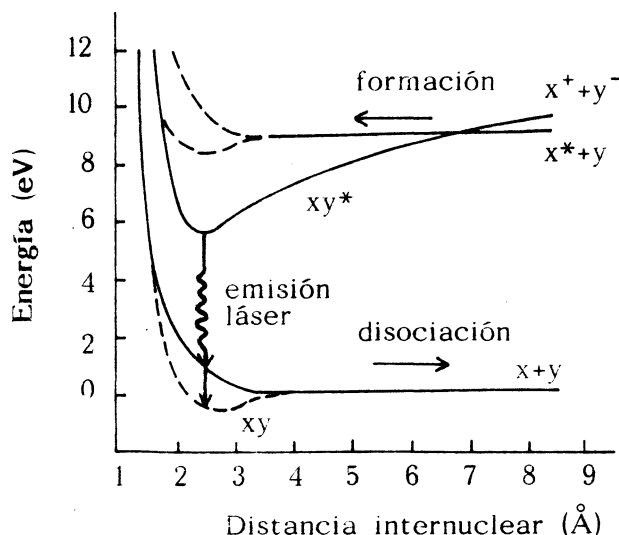


Figura 13

Las moléculas de excímeros se forman en la descarga al excitarse uno de los átomos constituyentes y atrapar al otro, cosa que en su estado fundamental resulta imposible.

Los dispositivos usados son en general celdas de alta presión en las que se provoca una descarga eléctrica, sea por bombeo por haces de electrones muy energéticos (e-beam), con los que se logran las mayores energías por pulso, o por descargas de avalancha en las que se debe desarrollar una alta potencia sobre el plasma gaseoso formado, antes de la aparición de arcos de baja impedancia, que resultan sistemas muy económicos y en los que se han alcanzado las mayores potencias medias.

Los sistemas de descarga en avalancha son comunes a muchos láseres que funcionan en régimen de presiones mayores que la atmosférica. Tales dispositivos no difieren grandemente de los ya estudiados excepto en dos aspectos: el primero, la velocidad del sistema y el segundo que, dado que las presiones a las que trabaja la celda son altas, la formación de arcos es muy probable a menos que se pre-"siembre" el gas de iones que sean capaces de homogeneizar el volumen de disparo. Esto se logra mediante una descarga auxiliar previa un tiempo antes, del orden de los 100 ns, dependiente de la presión y la tensión a la que se produce el disparo.

El sistema de disparo trabaja con sistema de duplicación de tensión, de modo que cargando el dispositivo a tensiones del orden de 25 kV, la celda puede encontrarse a ~ 50 kV durante el tiempo necesario para provocar la descarga; la corriente queda fijada por la capacidad usada, que para un láser de 100 mJ de energía es del orden de 60 nF; esto corresponde a una potencia del orden de 7 MW ya que la duración del pulso es del orden de 50 ns.

Este sistema es útil sólo en el caso de formación de excímeros de gases nobles-halógeno. En los casos gas noble-gas noble, sólo puede emplearse el bombeo por haz de electrones.

La celda de alta presión se llena con una mezcla de gases, en el caso de los halógeno-gas raro, con gas raro ($\sim 1\%$), halógeno ($\sim 0,1\%$) y un gas buffer (99%); dicha mezcla se contamina hasta impedir la acción láser luego de $\sim 10^5$ disparos.

Algunos ejemplos de excímeros gas raro-halógeno son: el ArF^* ($\lambda 193$ nm), el KrF^* ($\lambda 249$), XeCl^* ($\lambda 308$) y XeF^* ($\lambda 351$); el ArF^* es uno de los más usados en fotoquímica; el KrF^* es el más eficiente y por tanto en el que se han desarrollado mayores potencias y energía; el XeCl^* el que más tiempo permanece útil en la celda; y el XeF^* el de mayor eficiencia de bombeo de colorantes.

Estos láseres emiten en una banda homogéneamente ensanchada, por lo que puede lograrse emisión en banda angosta por medio de un sistema de red externa de difracción. Mediante este método y amplificadores adecuados se han logrado pulsos con anchos espectrales de 10^{-2} Å respecto a los anchos naturales, que son del orden de los 50 Å, con muy buena eficiencia.

Como amplificadores han sido usados para amplificar pulsos ultracortos, llevándolos a niveles altos de energía mediante un adecuado control temporal de descargas.

Estos láseres son muy útiles como fuentes de luz para espectroscopía Raman. Se los ha desarrollado últimamente de larga vida y buena confiabilidad, consiguiéndose una operación de 10^9 pulsos en algunos casos.

La aplicación tradicional de los excímeros ha sido la investigación científica y ocupa aún hoy la mayor parte del mercado, lo que incluye espectroscopía UV y bombeo de láseres de colorantes. Sin embargo, los fabricantes han comenzado a ampliar su mercado al área industrial y médica. La emisión de los excímeros en el UV, donde materiales como plásticos, tejidos biológicos y semiconductores absorben fuertemente, los hacen potencialmente útiles para estas áreas. Los pulsos de láseres de excímeros producen la ablación de regiones delgadas de material con poco daño en los alrededores. Debido a que la corta longitud de onda de la radiación puede enfocarse en regiones pequeñas, la ablación y otros efectos pueden controlarse en una escala submicrométrica, hecho que será importante en la futura integración de gran escala de circuitos electrónicos de semiconductores (VLSI).

Los excímeros pueden usarse para el grabado, deposición y dopado en films delgados. Una aplicación inmediata es la remoción del "photoresist" que cubre las marcas de alineación en las obleas de semiconductor (por fotodescomposición ablativa), mejorando así la exactitud de alineación de los equipos de fotolitografiado. El proceso está libre de daño al sustrato, ya que no hay calentamiento de éste en el proceso de ablación.

Se han desarrollado excímeros de pequeño ancho de línea (0,002 nm) y diferentes potencias, que los hace sumamente útiles en muchas aplicaciones de "patterning" de semiconductores. Estos láseres son compactos y en general poseen microprocesadores que controlan parámetros tales como energía de salida, pureza del material activo, etc. Además están contruidos en forma modular, lo que permite el fácil recambio de algunas de las partes para su mantenimiento o reparación.

Otra aplicación en desarrollo es el perforado de pequeños agujeros en fibras ópticas para una nueva técnica de acoplamiento de luz.

Aunque la investigación sobre las aplicaciones médicas se encuentra en una etapa comparativamente primitiva, se han obtenido promisorios resultados en pruebas de laboratorio sobre efecto en ojos, corazón y otros órganos.

La característica principal del excímero es su capacidad de ablación de regiones delgadas de tejido sin causar gran trauma en los alrededores. Se ha desarrollado un tipo de láser (ExciMed) conformado por una unidad completa con ruedas y un brazo articulado por donde sale el haz; tiene también la posibilidad de conexión con fibra óptica, generando pulsos de 100 mJ a una repetición de 200 Hz. Recientemente, se han logrado 10^8 pulsos con una potencia promedio de 40 W con un 20% diario de renovación de gas en un láser de KrF. En un XeCl se consiguieron 10^8 pulsos en 150 W con una sola carga de gas. Últimamente, usando un purificador de gas, se alcanzaron 10^8 pulsos en KrF con una sola carga.

4 - Otros láseres importantes

Con estos pocos ejemplos no hemos agotado ni por poco la lista de láseres existentes y además útiles.

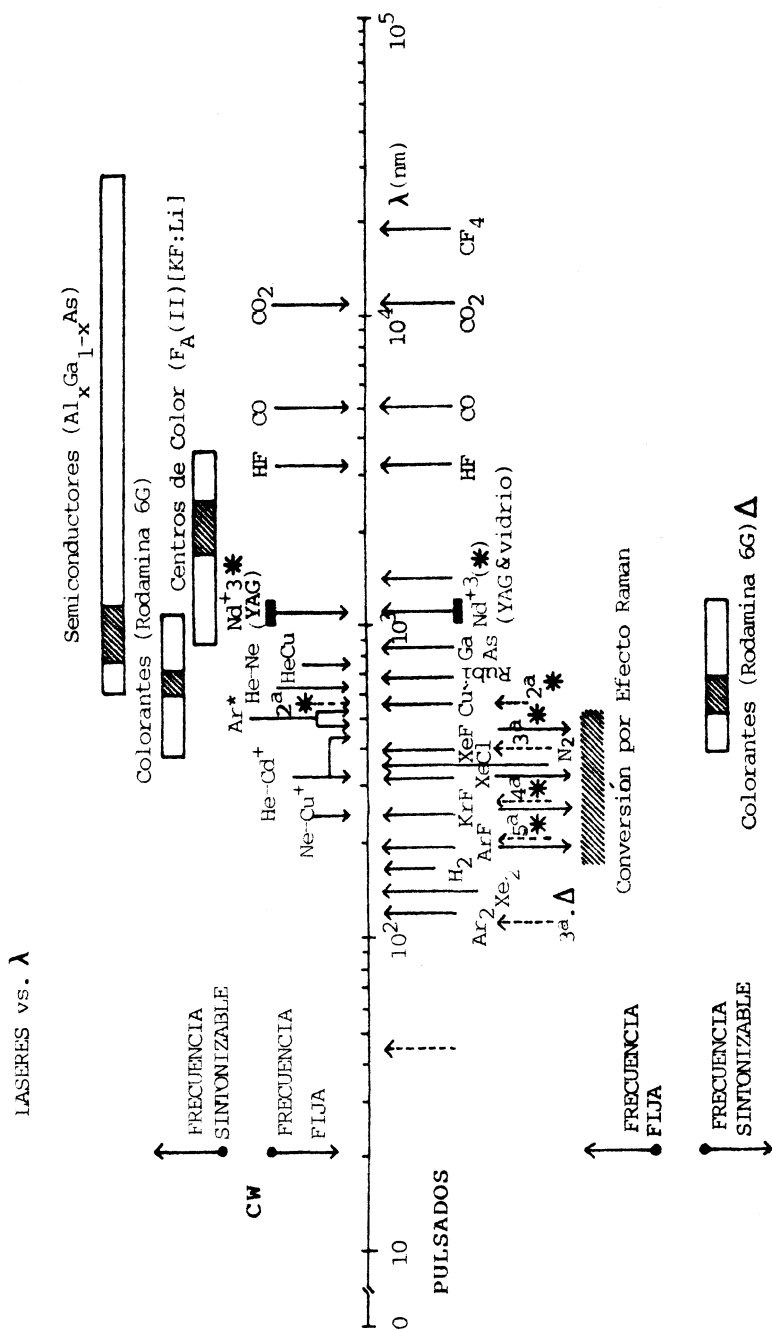
Pero ya advertimos que no es posible un muestreo más exhaustivo.

Hay láseres de electrones libres, en los que la radiación de frenado de los electrones genera acción láser. Láseres químicos, de fotodisociación, de vapores metálicos, de centros de color (F ó A-center), etc. Cada uno de ellos puede tener diversas aplicaciones de cierta importancia.

Como sabemos, cada frecuencia emitida por un láser, puede ser convertida en frecuencia doble, triple, mitad, etc., por métodos de óptica no-lineal, o también convertirse en frecuencias de efecto Raman estimulado. Estas técnicas no lineales y el uso de láseres de altísima potencia han permitido, en los últimos años, la generación de radiación coherente a longitudes de onda bajas (típicamente menores de 1000 Å). Un caso particularmente interesante son los llamados láseres de "rayos X blandos", cuyas longitudes de onda se encuentran entre los 300 y 50 Å. Es decir, que en estos láseres es necesario excitar las capas electrónicas más internas del átomo (típicamente $n = 2,3$). Ello hace que estos láseres se desarrollen principalmente al lado de equipos experimentales que desarrollan potencias elevadas, como por ejemplo los sistemas OMEGA, NOVA y PHAROS III. Cuando los diversos haces de estos láseres para fusión se enfocan sobre una pequeña muestra de material (láminas delgadas de cobre, aluminio, germanio, selenio o molibdeno), la gran irradiancia (alrededor de 10^{17} W/cm²) produce un campo eléctrico del orden del campo eléctrico atómico, lo que produce ionización multielectrónica. Se llega así a conseguir inversión de población entre niveles internos del átomo, originando emisión estimulada. En general el medio amplificador en estos experimentos provee ganancia para un solo pasaje de la radiación, pero experimentos muy recientes presentaron evidencia de un doble pasaje usando un espejo de multicapas dieléctricas de incidencia normal. Con el sistema NOVA se ha llegado a encontrar ganancia para una transición en Cu XX de alrededor de 10 Å. Actualmente la tendencia en láseres de rayos X apunta a lograr emisión estimulada en longitudes de onda cada vez más cortas y a aumentar la eficiencia sin sacrificar ganancia. Aunque estos dispositivos distan mucho de la fase comercial, es posible vislumbrar algunas aplicaciones como Holografía y Microscopía X en estudios biológicos de muestras "in vivo" basadas en el incremento del contraste de absorción entre el agua y el material orgánico.

En muchos casos pueden transformarse las salidas pulsadas o continuas en trenes de pulsos ultracortos sintonizables. Pero sobre esto no es aquí el lugar de entrar en consideraciones detalladas.

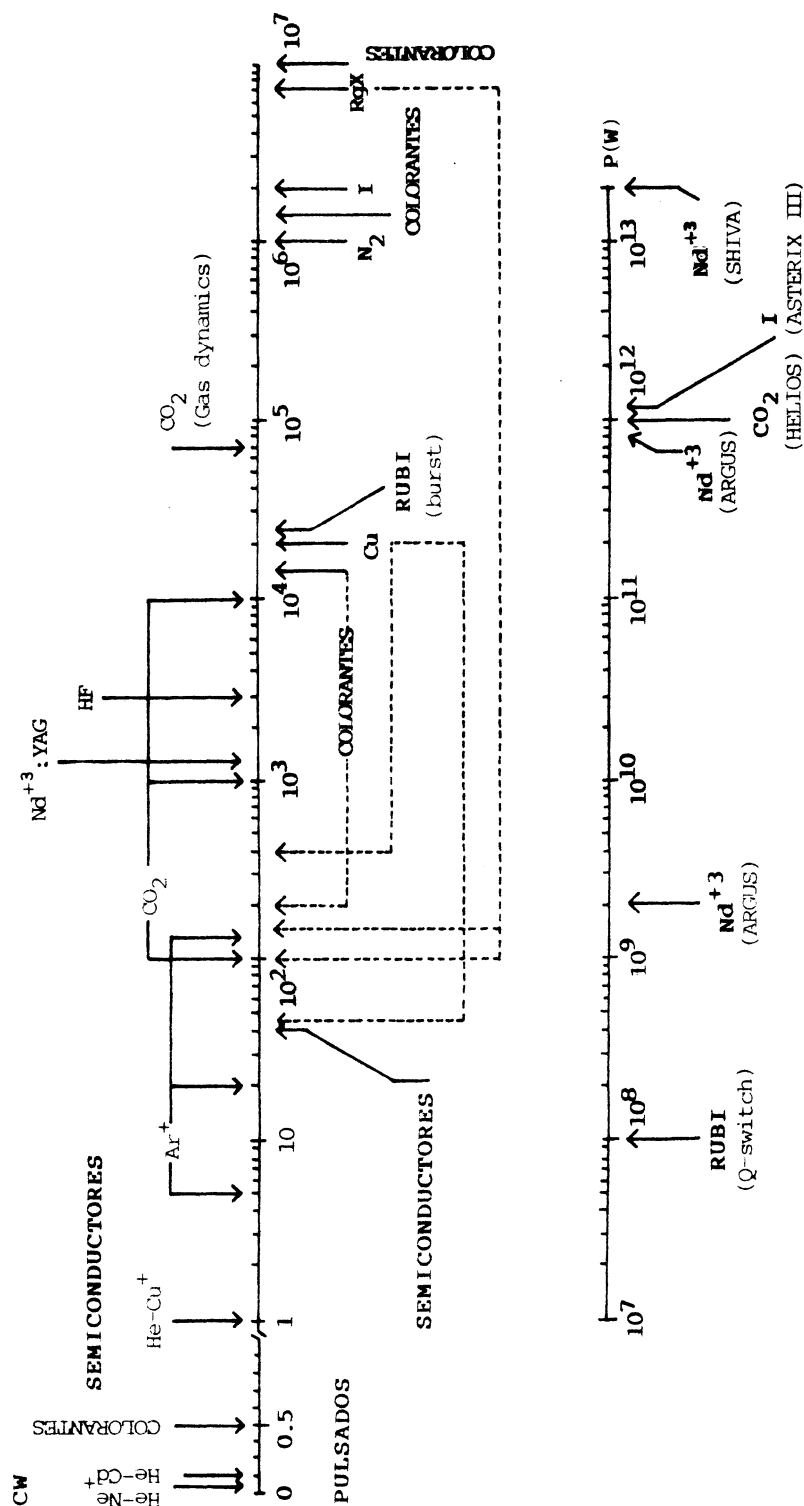
Como resumen para guía del lector, incluimos las siguientes tablas confeccionadas con datos más o menos recientes sobre los láseres citados.



- 1) Se dan los láseres más usuales en sus longitudes de onda más populares
- 2) En los casos de frecuencia sintonizable se presentan en forma global, con un ejemplo típico
- 3) * Se usa para Nd^{+3}
- 4) Δ Se usa para colorantes
- † Ancho de banda considerable (1 nm). Fuera de escala.

Tabla 1

LASERES vs. POTENCIAS PICO (Potencias RMS en líneas de trazos) ó POTENCIA CW



1) Se enumeran también con etapas amplificadoras

Tabla II

LASERES vs. ENERGIA POR PULSO

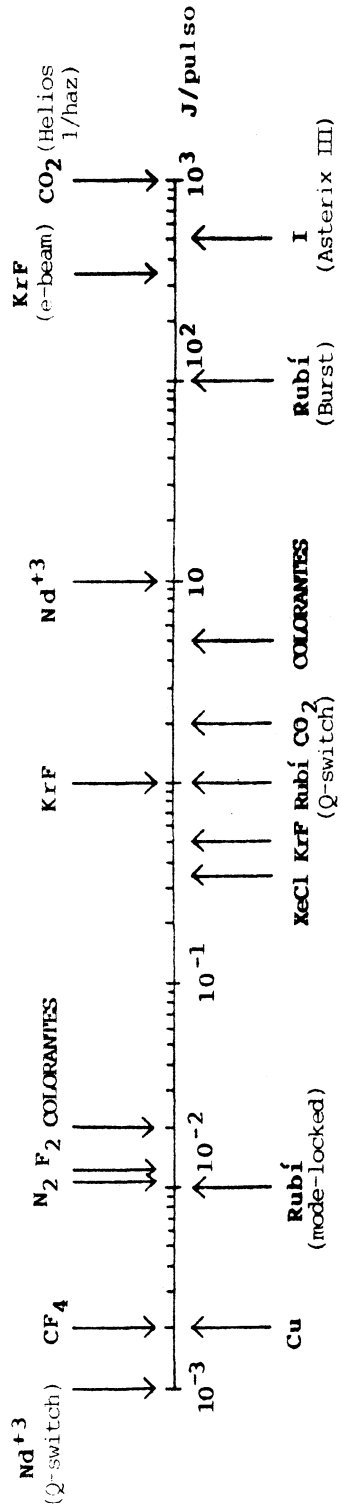


Tabla III

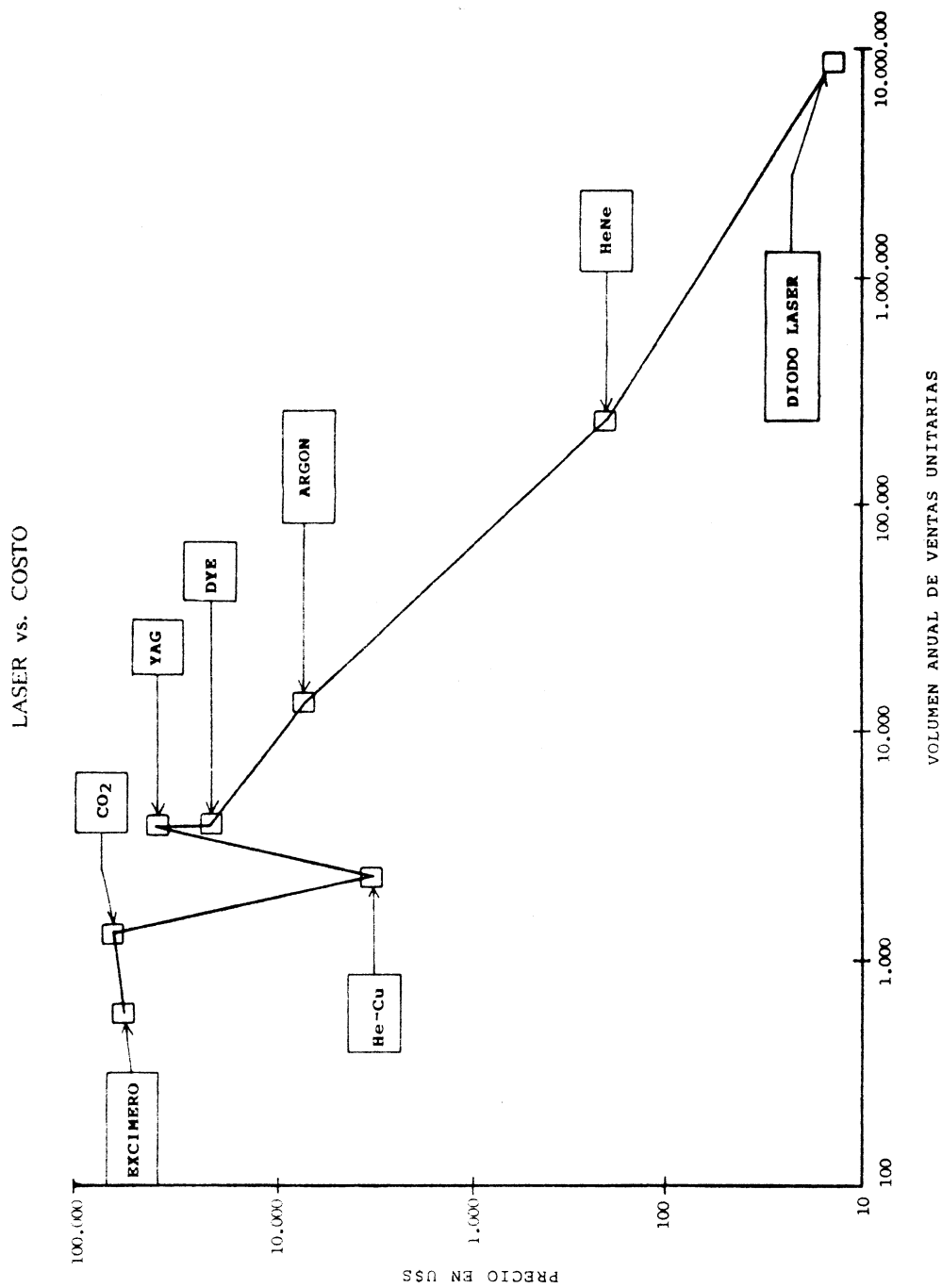


Tabla IV

Bibliografia

- "Principles of Lasers", O. Svelto, Caps. 3,5,6, Plenum Press (1979).
- "Developments of New Laser Systems", J.A. Piper, "Laser Physics". Proceedings of the 2nd. New Zealand Summer School in Laser Physics. Walls-Harvey eds., pp. 125-147, (1980).
- "Excimer Laser - a New Powerful UV Light Source", D. Basting, EPA Newsletter, pp. 64-80, April (1980).
- "Machining and Welding Applications", L. Weaver, Applications of Lasers. E. Monte Ross, Vol.1, pp. 201-238, Academic Press (1971).
- "Laser Communications", M. Ross, Applications of Lasers, Ed. M. Ross, Vol.1, pp. 239-295, Academic Press (1971).
- "High Power Lasers", A.F. Gibson, Rep. Prog. Phys., Vol. 43, pp. 3-65, (1980).
- "Principais Tipos de Lasers", R. Forneris, III Curso DFVLR:CTA. Lasers e Holografia: suas Aplicações a Ciencia e a Tecnologia. 7-18 fevereiro (1977).
- "Detetores de luz e lasers a semiconductor", C.C. Ghizoni, ibídem.
- "The Dye Laser Dye Update", S.E. Neister, Opt.Spectra, pp. 34-36 , (1980).
- "Tunable Lasers in the Visible and Near IR", H. Kogelnik, Actas del Primer Seminario Latinoamericano Sobre el Láser y sus Aplicaciones a la Física y a la Química. La Plata, 21-26 agosto (1978).
- "Tunable Diode and their Spectroscopic Applications", A. Mooradian, ibídem.
- "Laterally Coupled-cavity Semiconductor Lasers", R.J. Lang, IEEE Journal of Quantum Electronics, Vol. QE-23, nº4, April (1987).

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 Kirchhoff, 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.

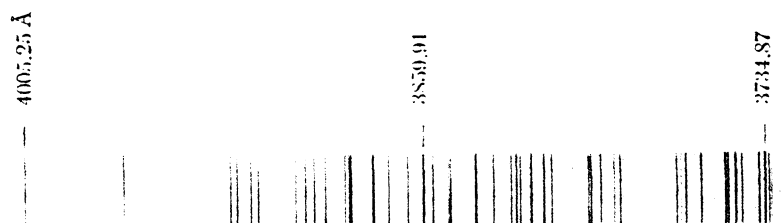


Figura 1. Parte del espectro del Fe.

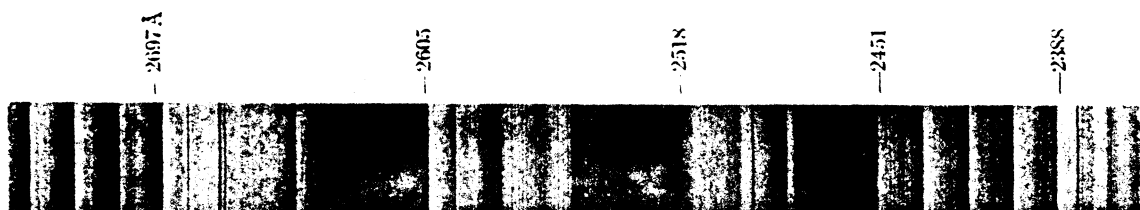


Figura 2. Espectro de banda de la molécula de PN.

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 la 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:

n_2	n_1	Serie
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.

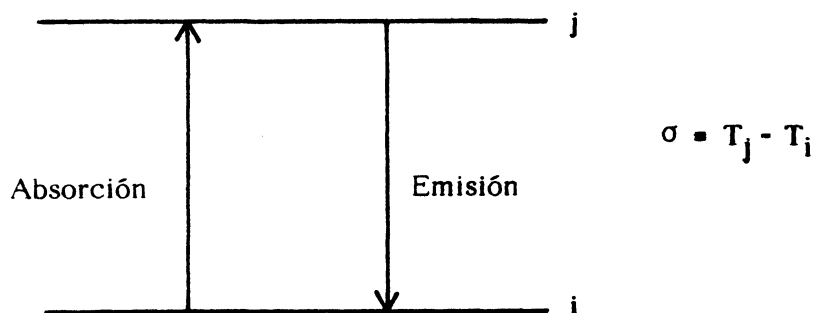


Figura 3

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 Cal.

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.

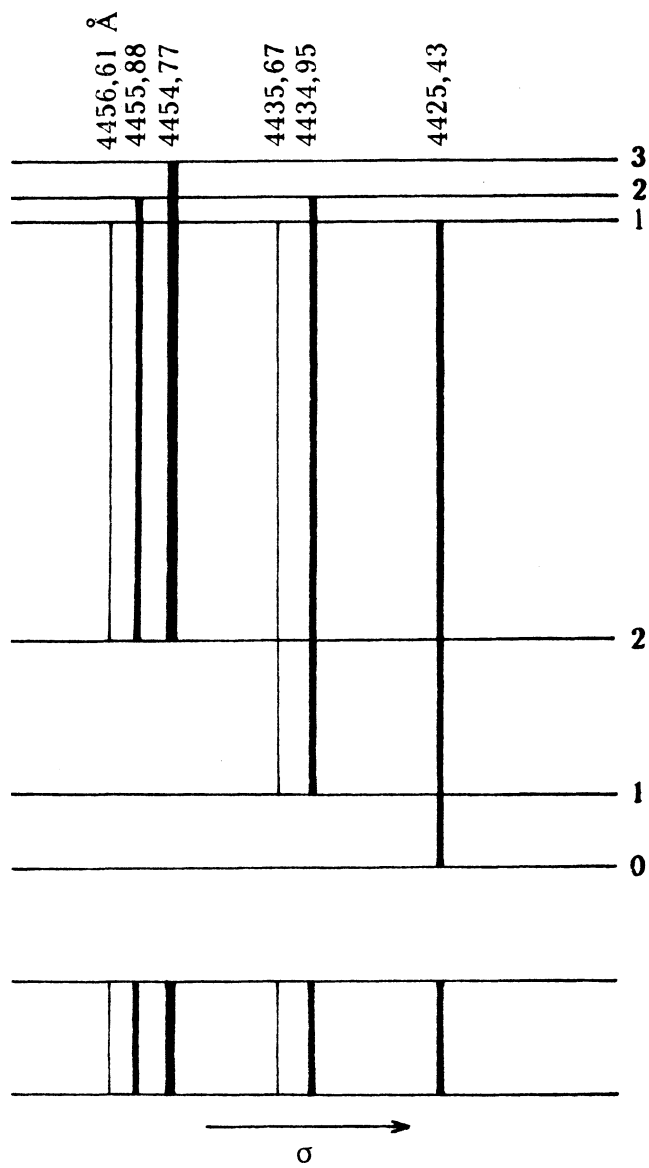


Figura 4

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

$$J = L + S \text{ } |L - S| \text{ .}$$

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

$$L = (l_1 + l_2) \text{ } |l_1 - l_2| \text{ .}$$

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

$$S = (s_1 + s_2) \text{ } |s_1 - s_2| \text{ .}$$

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, ...) .

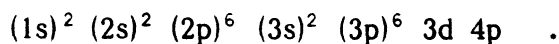
De lo expuesto surgen los siguientes términos:

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

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:



—————	$^1F_3$	40.537,86 cm^{-1}
	$^3P_2$	39.340,08
=====	$^3P_1$	39.335,32
	$^3P_0$	39.333,37
=====	$^3D_3$	38.259,10
	$^3D_2$	38.219,09
	$^3D_1$	38.192,37
—————	$^1P_1$	36.731,62
	$^1D_2$	35.835,40
=====	$^3F_4$	35.896,89
	$^3F_3$	35.818,71
	$^3F_2$	35.730,45

Figura 5

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

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

$$\Delta S = 0$$

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

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 $J=0 \nrightarrow J=0$, 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, es decir, aquellos 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 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 ,

$$l + j_c = K \quad .$$

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

$$K + S = J \quad .$$

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.

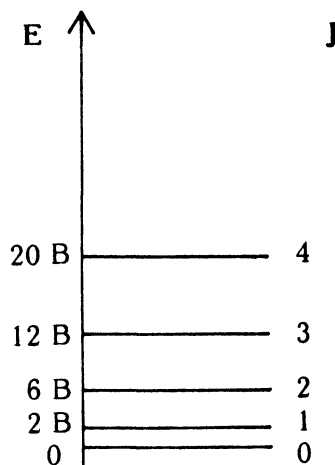


Figura 6

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 $0, 1, 2, \dots$.

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 $X, A, B, \dots$, 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 $0, 1, 2, \dots$, los que son designados por $\Sigma, \Pi, \Delta, \Phi, \dots$,

análogamente a la designación para los átomos.

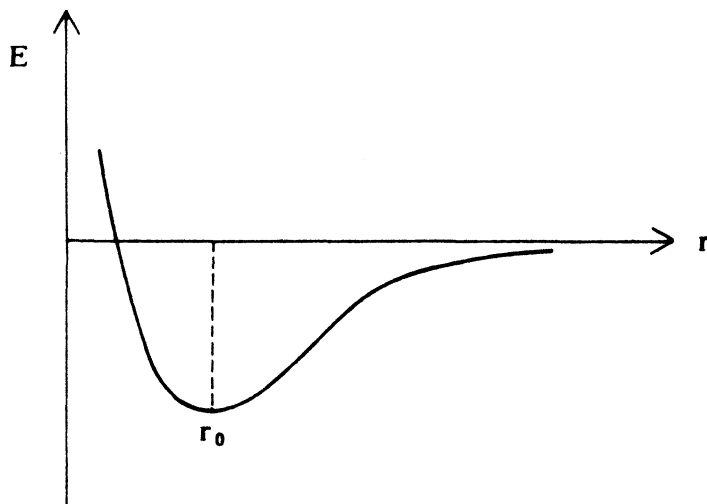


Figura 7

Espectros de moléculas diatómicas

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

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

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.

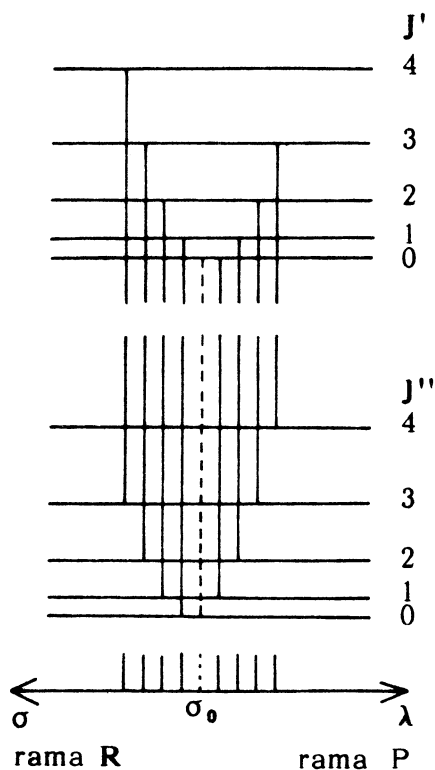


Figura 8

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.

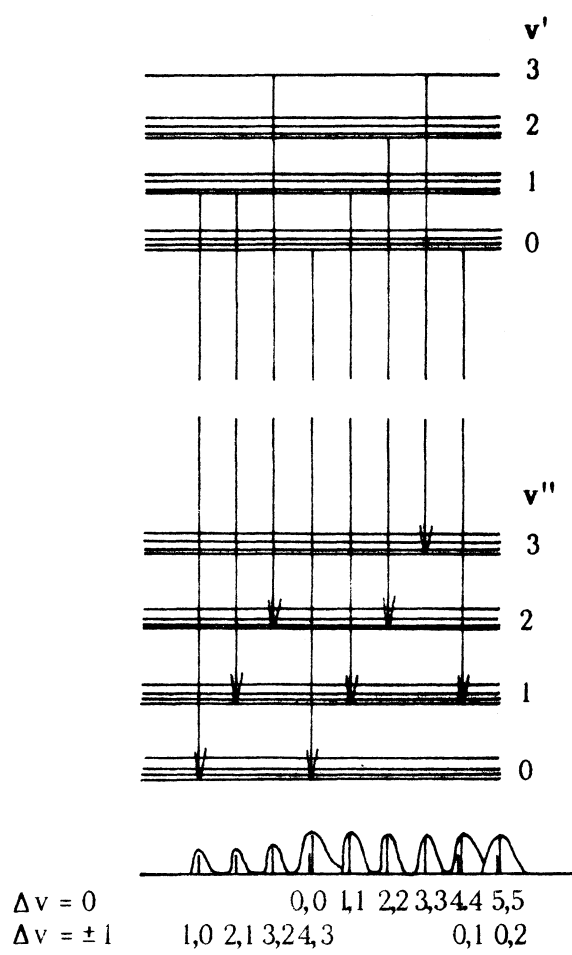


Figura 9

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.

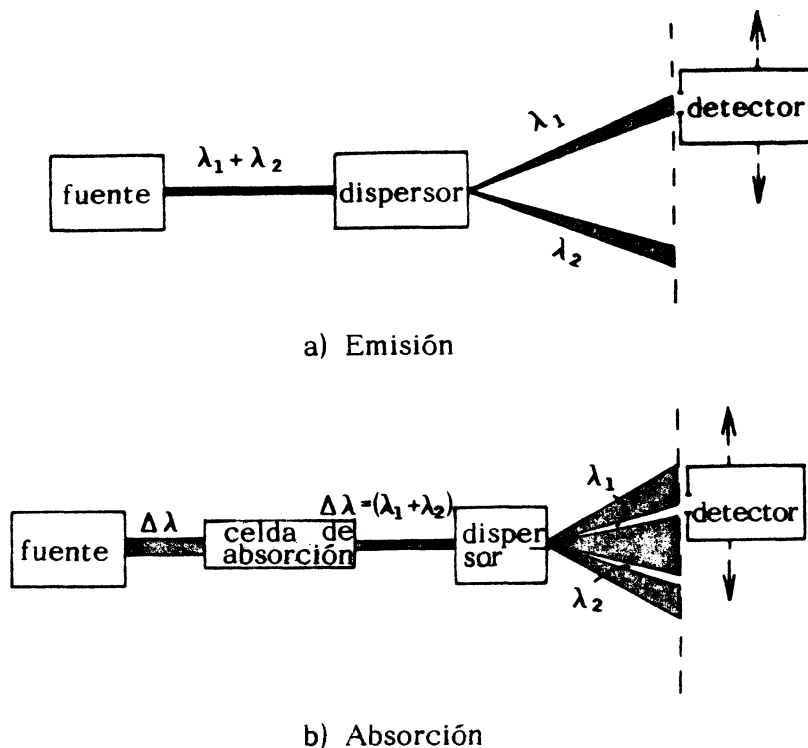


Figura 10

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.

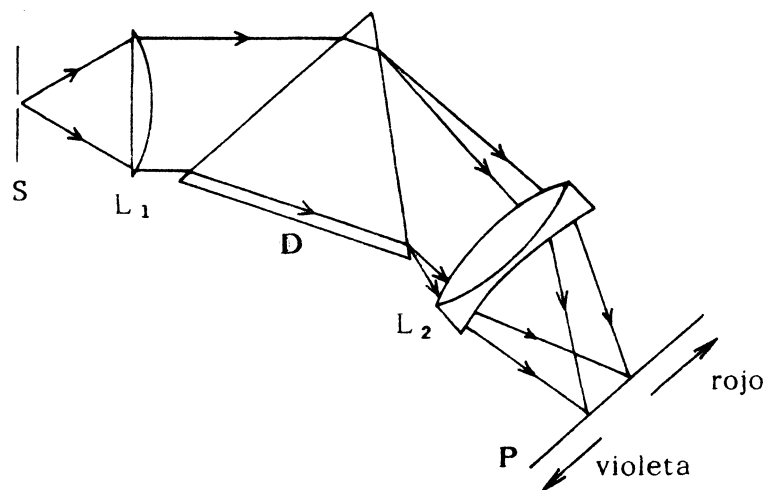
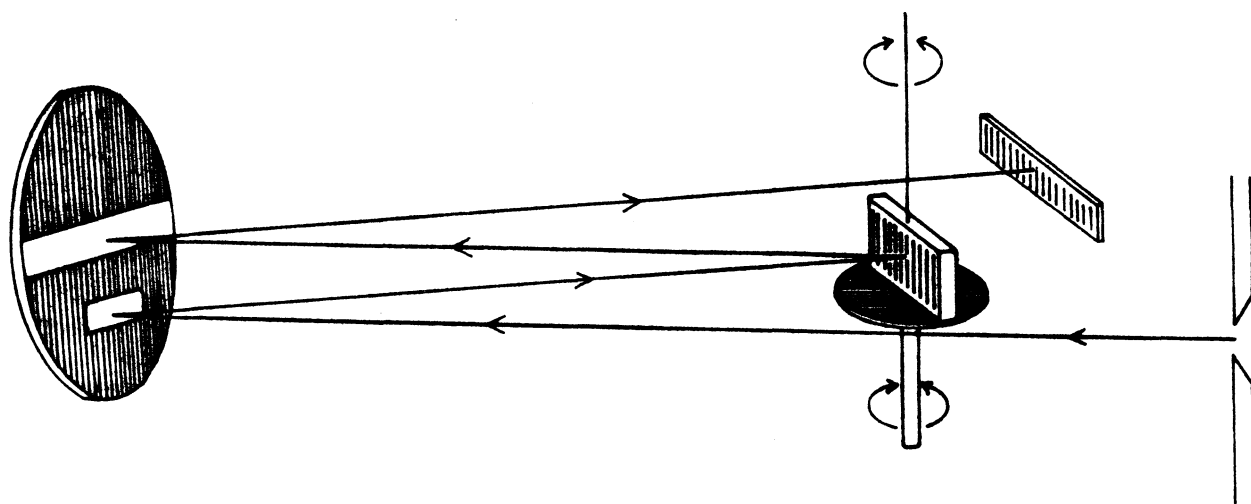
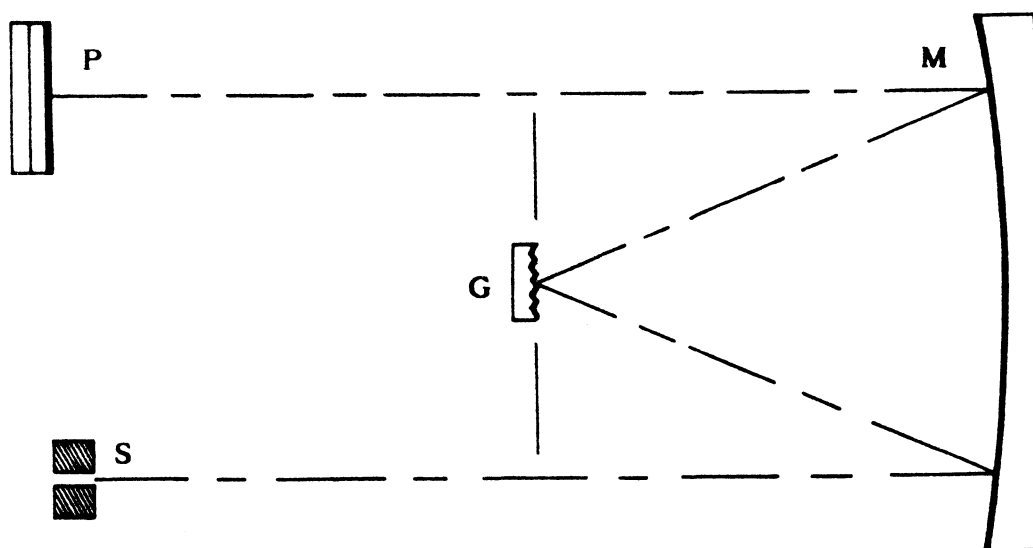


Figura 11. Componentes de un espectrógrafo de prisma convencional. S es la ranura, L_1 y L_2 las lentes colimadoras y de cámara, respectivamente, D el prisma y P el plano focal de L_2 .



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}	$1,986 \times 10^{-23}$	$1,240 \times 10^{-4}$	$1,986 \times 10^{-16}$	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).

OPTICA FISICA. 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 permiten 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 hacen 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 describir mediante una

misma ecuación matemática, representando ya sea el campo eléctrico como el magnético de 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:

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

donde $\nabla^2 E$ es el laplaciano de E y t el tiempo, y 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:

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

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:

$$U(x,y,z,t) = A \cos[2\pi(\nu t - z/\lambda) + \phi] \quad , \quad (3)$$

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 y de la distancia z sobre el eje de propagación. Representando esta función con el tiempo t (Figura 2) o la coordenada z (Figura 1) como parámetro, pueden hacerse más evidentes sus principales características:

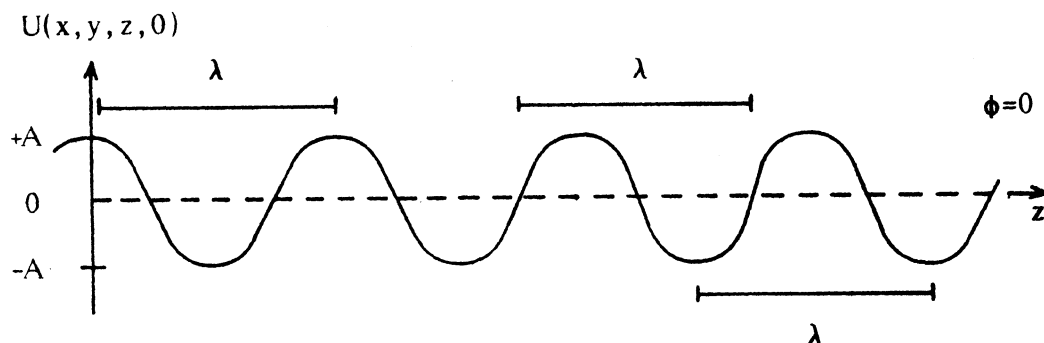


Figura 1

En la Figura 1 se ha representado $U(x, y, z, 0) = A \cos(\phi + 2\pi z/\lambda)$ pero, para simplificar, se ha puesto $\phi=0$. En el plano $z=0$ y en todos los planos paralelos a él y que están a una distancia $z=\lambda n$, 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 denomina frecuencia espacial angular, y que se mide en radianes por unidad de longitud, ese campo luminoso para $t=0$ puede escribirse:

$$U(x, y, z, 0) = A \cos[kz] \quad (4)$$

Si en cambio se considera la evolución temporal del campo luminoso U en el plano $z=0$,

$$U(x, y, 0, t) = A \cos(2\pi\nu t + \phi) \quad (5)$$

y, para $\phi=0$, se obtiene la Figura 2, donde ν es la frecuencia de oscilación de U en un punto fijo y $T = 1/\nu$ el período. Si definimos la frecuencia angular temporal $\omega = 2\pi\nu$, que se mide en radianes por segundo, ese campo luminoso en el plano $z=0$ se puede escribir:

$$U(x, y, 0, t) = A \cos \omega t \quad (6)$$

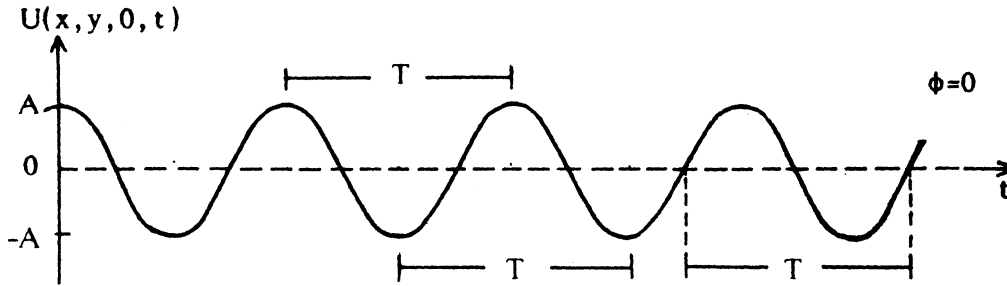


Figura 2

Con estas mismas definiciones, el campo luminoso en la ecuación (3) puede escribirse como:

$$U(x, y, z, t) = A \cos [\omega t - kz + \phi] \quad . \quad (7)$$

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 , y describe una onda plana de extensión infinita ($A(x, y) = \text{cte}$ 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 $\vec{k}$, cuya dirección y sentido son aquéllos de la propagación de la onda que queremos describir, y cuyo módulo es:

$$k = (\vec{k} \cdot \vec{k})^{1/2} = 2\pi / \lambda \quad (8)$$

y considerando además que $\vec{r}$ 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 $\vec{k}$ es:

$$U(x, y, z, t) = A \cos (\omega t - \vec{k} \cdot \vec{r} + \phi) \quad . \quad (9)$$

La frecuencia ν y la longitud de onda λ de un campo luminoso que se propaga en el vacío están ligadas por la relación:

$$\lambda \nu = c \quad ,$$

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:

$$U(x,y,z,t) = A/r \cos(\omega t - \vec{k} \cdot \vec{r}) \quad .$$

Otro de los conceptos que utilizaremos en lo que sigue es el Principio de Superposición: 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.

Muchas veces es cómodo escribir estas mismas ecuaciones en notación compleja, de la forma:

$$U(x,y,z,t) = A [\cos(\omega t - \phi) + i \sin(\omega t - \phi)] = A \exp[i(\omega t - \phi)] \quad .$$

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:

$$U = A \exp(-i\phi) \quad \text{para la onda plana} \quad ;$$

$$U = (A/r) \exp(-i\phi) \quad \text{para la onda esférica} \quad .$$

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,

$$I = U U^* .$$

Una fuente luminosa capaz de emitir vibraciones monocromáticas tales como las que hemos descripto, 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 sería una fuente luminosa circular, en cuyo interior la intensidad es constante y en cuyo exterior es cero, 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 pueden ser asimiladas 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. Para las fuentes luminosas clásicas, la longitud de coherencia es del orden de 60 cm, 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 metro-lógicas.

Superposición de ondas

Si consideramos por un momento dos fuentes luminosas puntuales que emiten vibraciones ilimitadas, monocromáticas, paralelas, podemos estudiar 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.

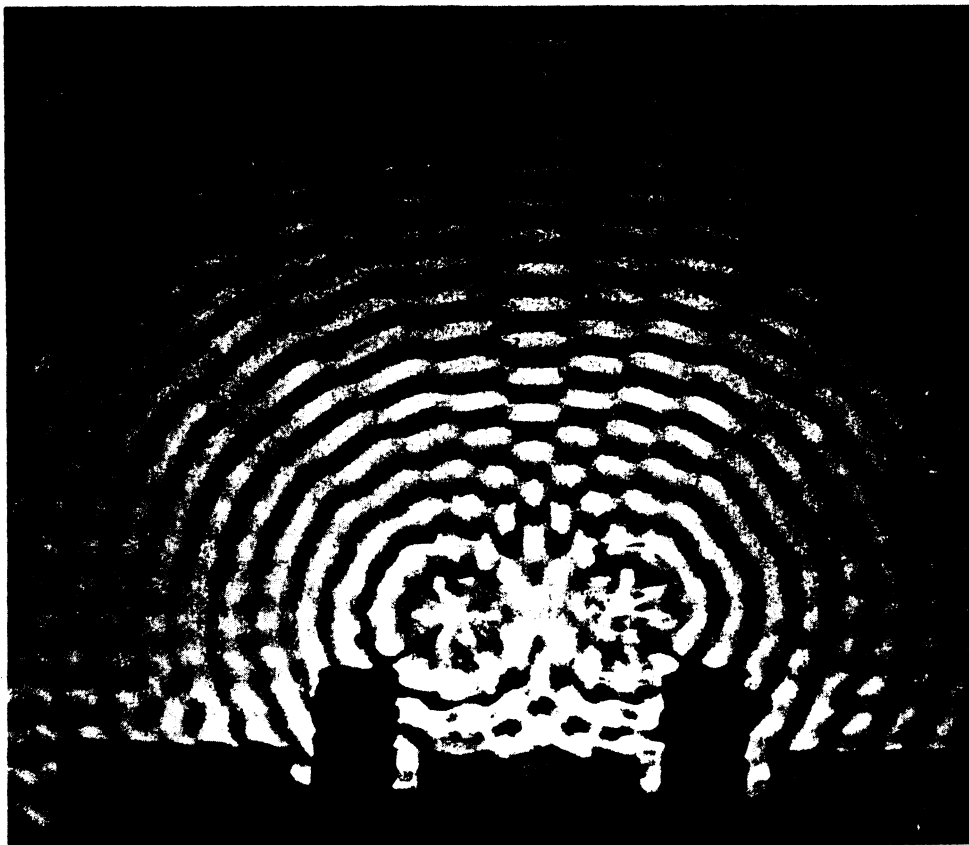


Foto 1. Superposición de ondas en agua. Las perturbaciones provienen de dos fuentes puntuales.

Interferencias

Una de las maneras de explicar el fenómeno de "speckle" es a partir del de interferencia. El primero en comprobar y estudiar este fenómeno fue Thomas Young, quien en 1801 enunció el "principio de interferencia" a partir de sus observaciones.

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 , como en el esquema de la Figura 3.

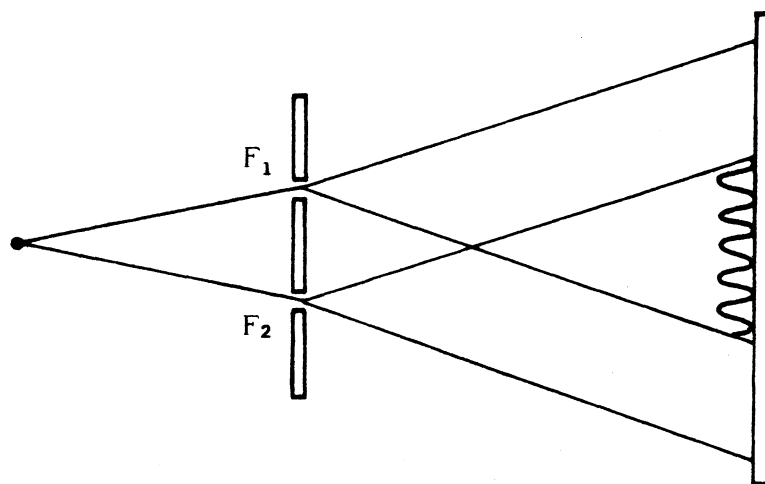


Figura 3

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.

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 de las fuentes que la producen.

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 durante 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 su coherencia espacial, en función de la visibilidad o el contraste de las franjas, definido por medio de la expresión

$$\gamma = (I_{\text{máx}} - I_{\text{mín}}) / (I_{\text{máx}} + I_{\text{mín}}) \quad .$$

Habiendo descripto 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:

$$U_1(x,y,z,t) = A \exp [(\omega t - kz + \phi_1)i] \quad .$$

$$U_2(x,y,z,t) = A \exp [(\omega t - kx \sin\theta - kz \cos\theta + \phi_2)i] \quad .$$

El campo total en el punto (x,y,z) y al tiempo t , será:

$$U(x,y,z,t) = U_1 + U_2 \quad .$$

Como ni el ojo ni los detectores más rápidos pueden seguir las variaciones del campo (del orden de los 5×10^{14} Hz), la magnitud observable es la intensidad luminosa, que se obtiene haciendo el valor medio temporal del cuadrado del campo, es decir:

$$I(x,y,z) = (1/T) \int_0^T U^2(x,y,z,t) dt$$

$$= U U^*$$

$$= 2A [1 + \cos(kx \sin \theta + (\phi_1 - \phi_2) + kz(\cos \theta - 1))] ;$$

entonces, si la pantalla está en $z=0$,

$$I(x,y,0) = 2A^2 [1 + \cos(kx \sin \theta + \Delta\phi)] ,$$

donde hemos puesto $\Delta\phi = \phi_1 - \phi_2$.

La intensidad luminosa varía como una función $(1 + \cos)$ y tiene un gráfico como el de la Figura 4.

M = máximos franjas
brillantes

m = mínimos franjas
oscuras

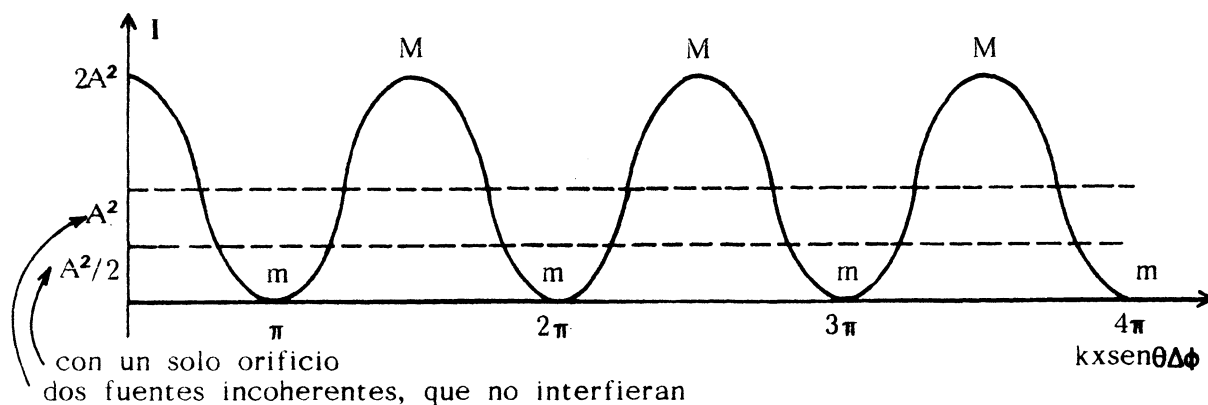


Figura 4

que consiste en una sucesión de franjas claras y oscuras. 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 $\omega t + \phi_1$ y $\omega t + \phi_2$, 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 el punto considerado y la fuente F_1 difiere de la distancia óptica entre el punto y F_2 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.

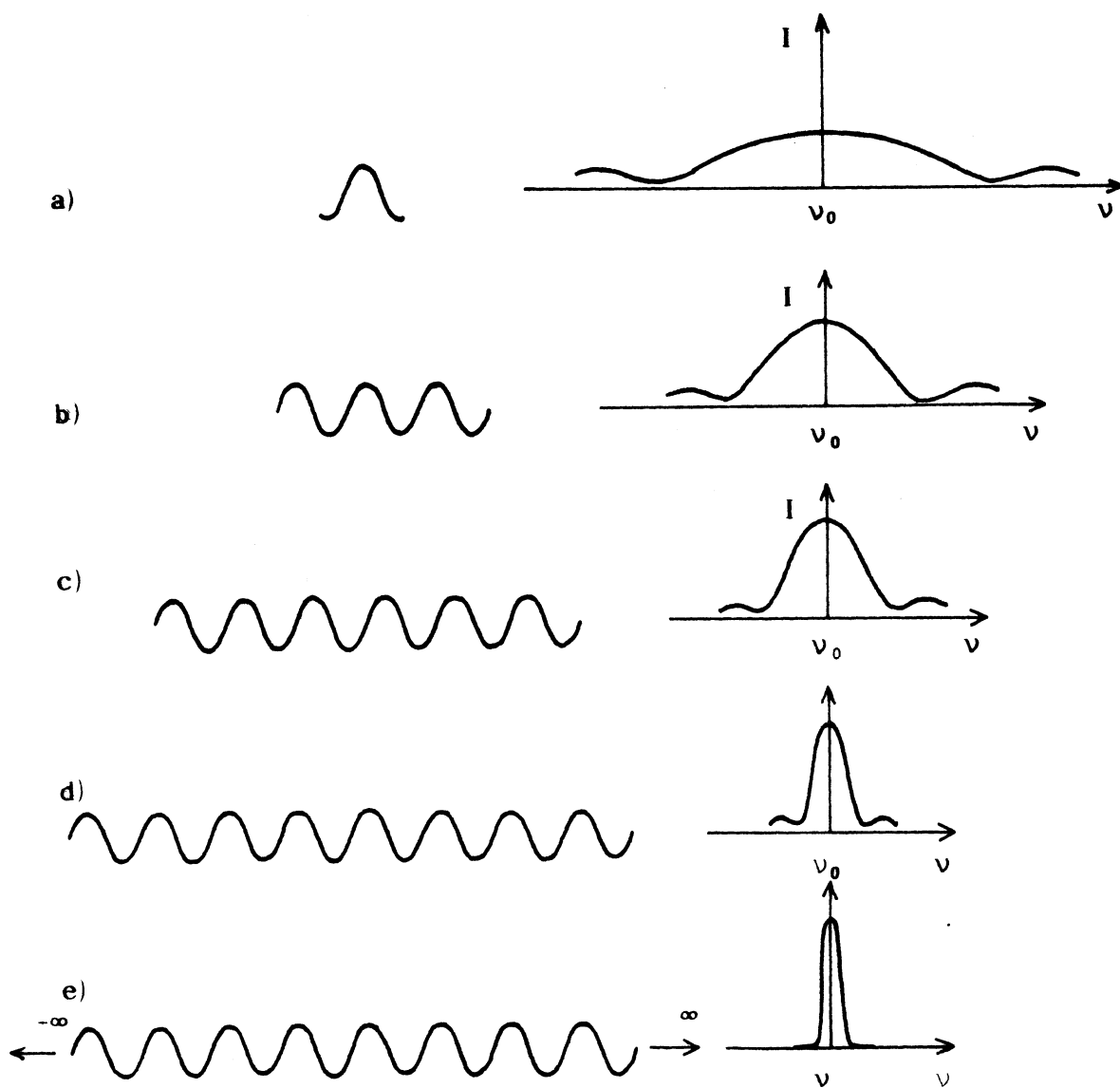


Figura 5

Las sinusoides graficadas representan trenes de onda finitos de distinta longitud, salvo 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 media 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 para 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 que difiere en $dz = cdt$.

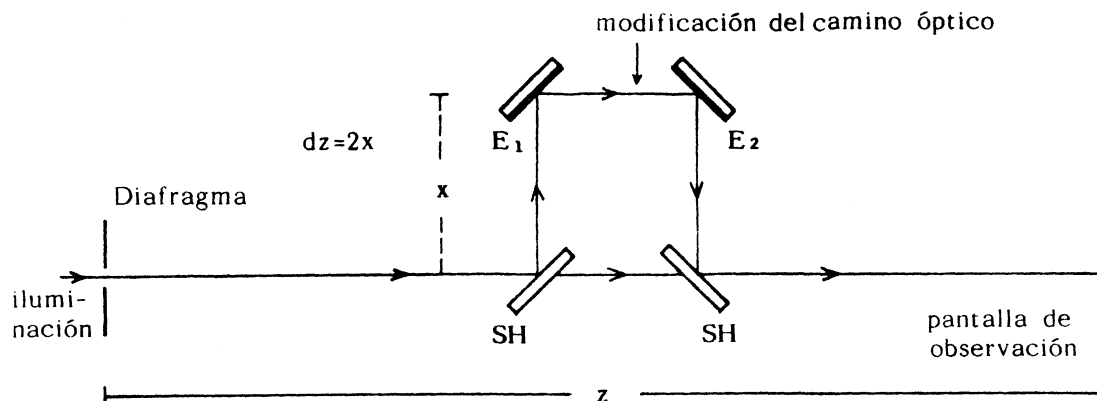


Figura 6

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:

$$\Delta\nu = c/\ell \quad \text{y} \quad \tau = 1/\Delta\nu$$

Interferómetros

En realidad, la coherencia se mide habitualmente por medio de interferómetros. Para estudiarlos, vamos a dividirlos en dos grandes grupos: aquéllos en los que la superposición es entre dos haces luminosos, y aquéllos en los que se superponen más de dos haces de luz. Como ejemplos de ellos analizaremos en particular el interferómetro de Michelson, el de Mach-Zehnder y el de Twynam-Green en el primero y el de Fabry-Perot en el segundo. 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.

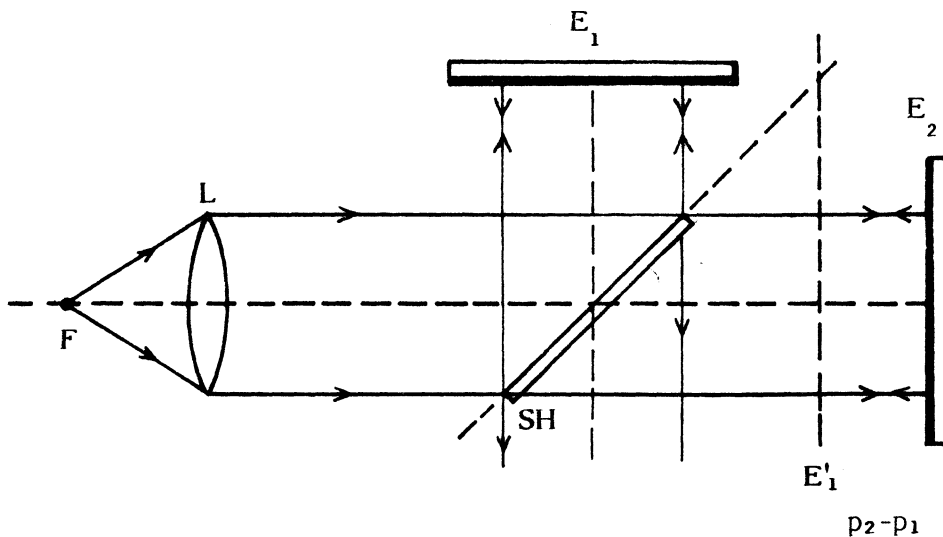


Figura 7

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 E_1 y E_2 reflejan estos haces luminosos, haciéndoles 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 $ds = 2n(p_2 - p_1)$, donde n es el índice de refracción del medio para la longitud de onda utilizada, y p_2 y p_1 las distancias recorridas por los dos haces. En la figura se indica el plano E'_1 imagen de E_1 dada por SH y la diferencia $(p_2 - p_1)$.

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 $(2m+1)\lambda/2$.

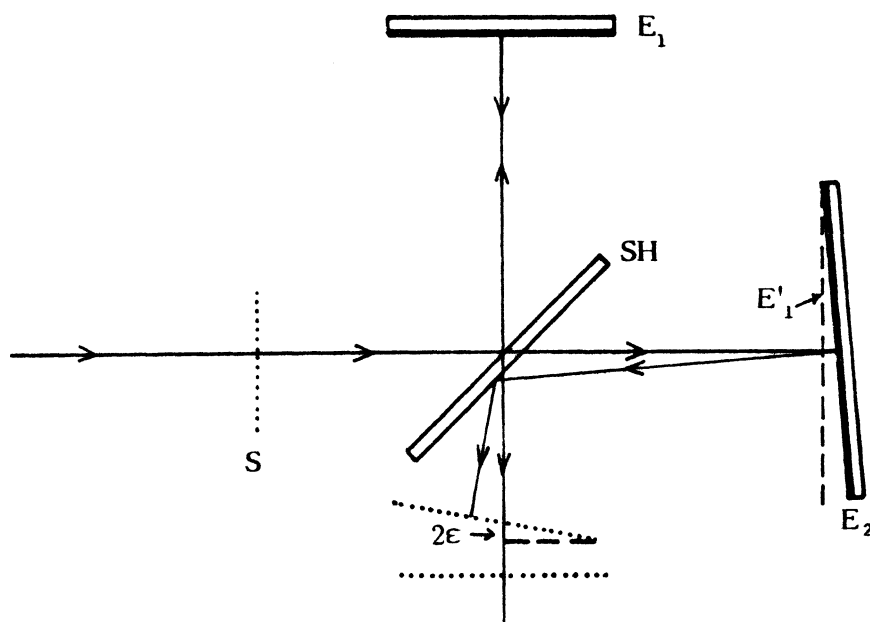


Figura 8

Si $ds=0$ y el sistema está ligeramente desalineado, como en la Figura 8, 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 la 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 paralelo a sí mismo, este desplazamiento introduce una diferencia de camino óptico δs 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-Zehnder:

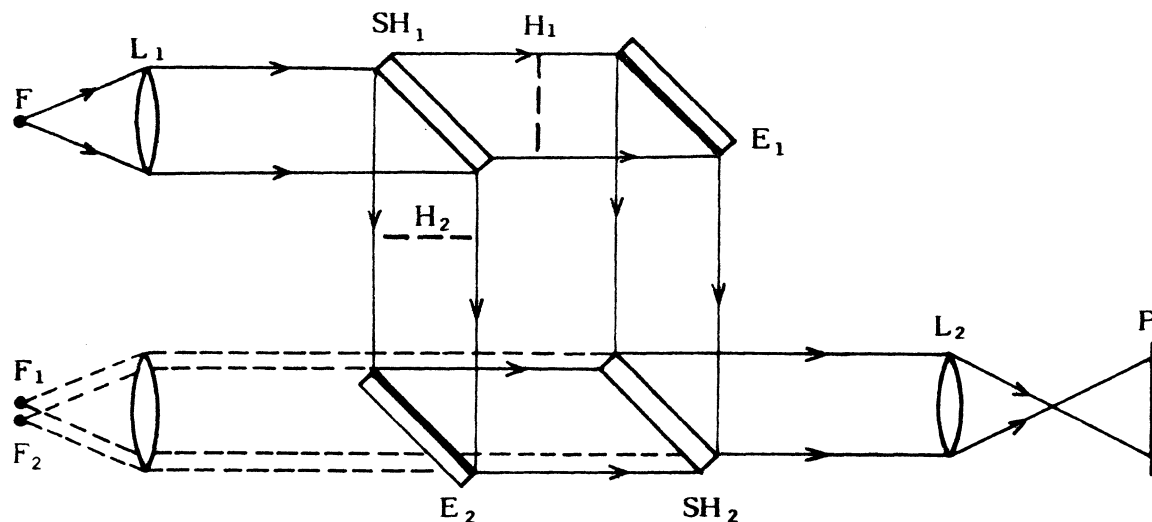


Figura 9

Es un interferómetro de dos haces, como el de Michelson, con la particularidad de 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.

Una lente colimadora L_1 envía un frente de onda plano sobre un semiespejo SH_1 , y se obtienen de esta forma dos haces planos H_1 y H_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. Ubicando adecuadamente los espejos y semiespejos 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 n es una función de la presión p , según la expresión:

$$n = 1 + 0,0003 \, p/H \quad ,$$

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.

La Figura 10 muestra un esquema de una adaptación del interferómetro de Mach-Zehnder para analizar la distribución de presión alrededor de un obstáculo O 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 A y B 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 O .

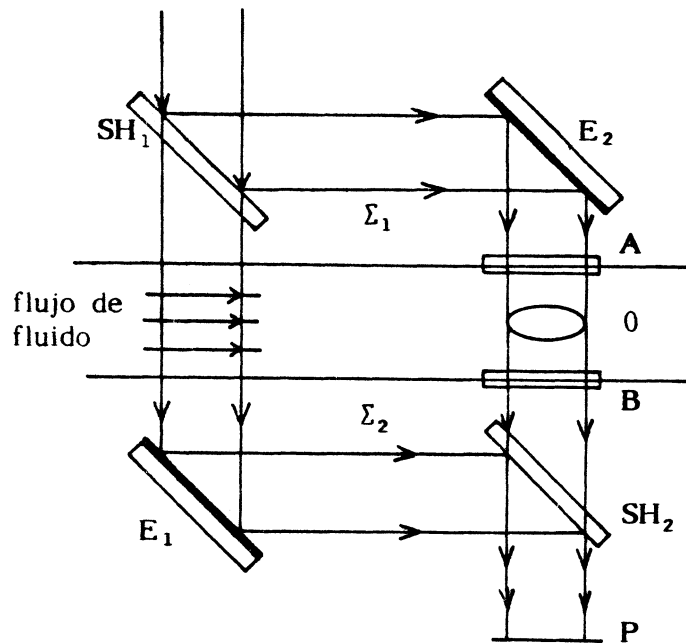


Figura 10

La sensibilidad de este instrumento, así como la de todos los interferómetros que se basan en la superposición de 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 se obtienen resultados precisos en el rango de la décima de micrometro. 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.

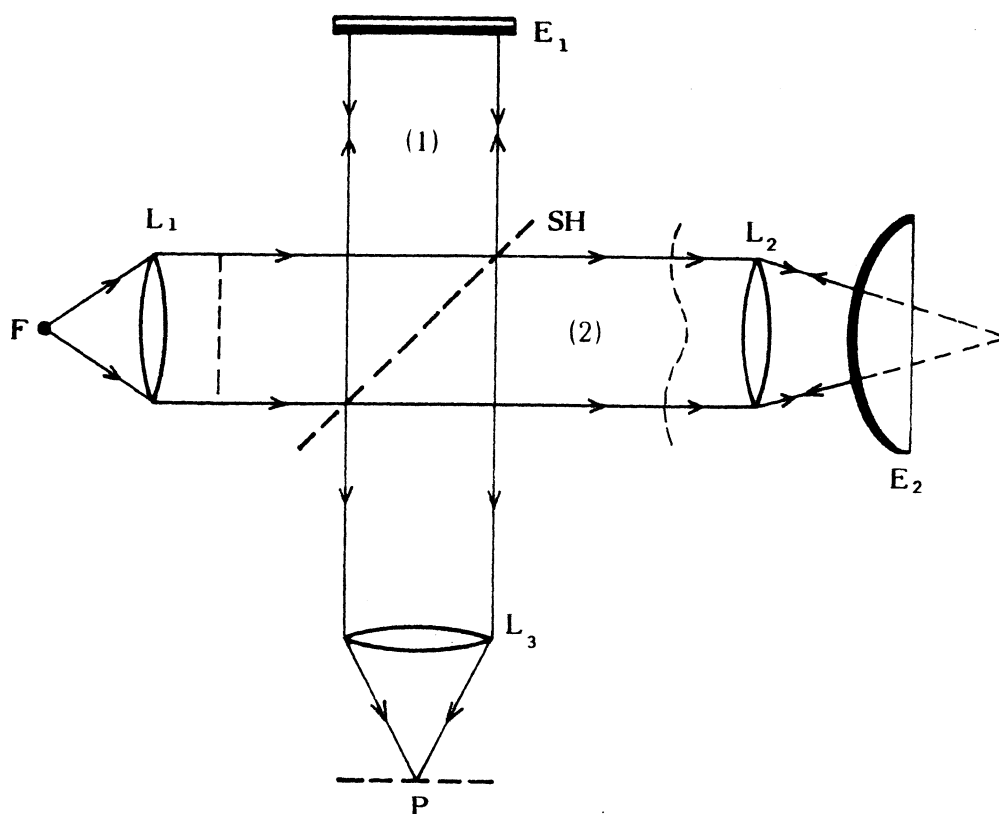


Figura 11

Como en el interferómetro de Michelson, la fuente luminosa F se coloca en el plano focal de un buen objetivo L_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 L_2 . Este elemento se ha colocado en el brazo del interferómetro y el espejo E_2 es esférico, convexo, y está colocado de manera tal que su centro de curvatura coincida con el foco de L_2 . De esta forma, si el objetivo L_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 L_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 L_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 L_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 L_2 . Esto último implica que el dispositivo tiene más sensibilidad que el interferómetro de Mach-Zehnder, pero que la interpretación de los diagramas de interferencia es más complicada.

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 semirreflectantes.

Para analizar el fenómeno, comenzaremos por estudiar las reflexiones múltiples en el interior de la cavidad formada por dos superficies planas semirreflectantes paralelas, separadas por una distancia e como la que se esquematiza en la Figura 12.

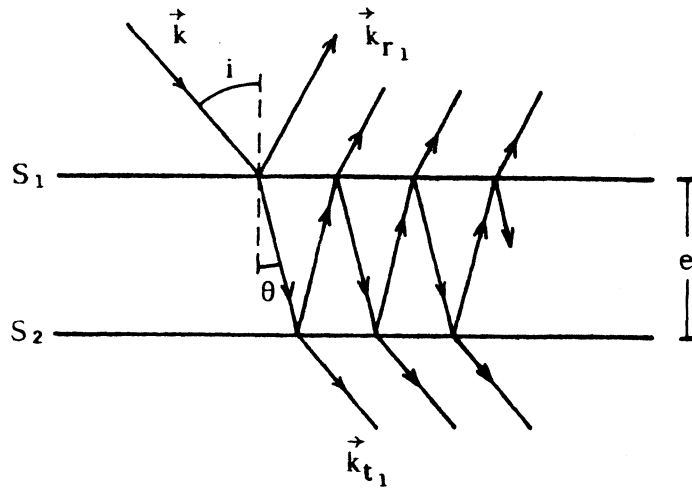


Figura 12

Sea $\vec{k}$ el vector propagación de una onda plana de amplitud A que incide bajo un ángulo i sobre la superficie S_1 , y sean $\vec{k}_{rj}$ y $\vec{k}_{tj}$ 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 :	$A r'_1$	
Para R_1 :	$A t_1 t'_1 r_2$	$\exp(-i2\phi)$
Para R_2 :	$A t_1 t'_1 r_2^2 r_1$	$\exp(-i4\phi)$
Para R_3 :	$A t_1 t'_1 r_2^3 r_1^2$	$\exp(-i6\phi)$
Para R_j :	$A t_1 t'_1 r_2^j r_1^{(j-1)}$	$\exp(-i2j\phi)$

y

Para T_1 :	$A t_1 t_2$	$\exp(-i\phi)$
Para T_2 :	$A t_1 t_2 r_2 r_1$	$\exp(-i3\phi)$

$$\text{Para } T_3: A t_1 t_2 r_2^2 - r_1^2 \exp(-i5\phi)$$

$$\text{Para } T_j: A t_1 t_2 r_2(j-1) r_1(j-1) \exp(-i(2j-1)\phi)$$

donde t_1 es el coeficiente complejo de transmisión en amplitud de la superficie S_1 para luz incidente desde el exterior;

t'_1 es el coeficiente complejo de transmisión en amplitud de la superficie S_1 para luz incidente desde el interior;

t_2 es el coeficiente complejo de transmisión en amplitud de la superficie S_2 para luz incidente desde el interior;

r' es el coeficiente complejo de reflexión en amplitud de la superficie S_1 para luz incidente desde el exterior;

r_1 es el coeficiente complejo de reflexión en amplitud de la superficie S_1 para luz incidente desde el interior;

r_2 es el coeficiente complejo de reflexión en amplitud de la superficie S_2 para luz incidente desde el interior;

ϕ es el atraso de fase introducido entre las dos superficies, dado por:

$$\phi = 2\pi n e \cos \theta / \lambda ,$$

donde θ es el ángulo de transmisión.

Recordemos que t_1 , t'_1 y t_2 son de la forma:

$$t_j = \mathcal{t} \exp(-i\psi) ,$$

donde $\mathcal{t}$ es la fracción transmitida de la amplitud, y ψ el retardo de fase introducido al atravesar la superficie, y que

$$T_j = t_j t_j^*$$

es la transmitancia de la superficie S_j definida como:

$$t_j = \text{energía transmitida} / \text{energía incidente} .$$

De modo similar, r_1 y r_2 son de la forma:

$$r_j = \mathcal{r} \exp(-i\delta) ,$$

donde $\mathcal{r}$ es la fracción reflejada de la amplitud y δ el retardo de fase introducido en la reflexión;

$$R_j = r_j r_j^*$$

es la reflectancia de la superficie S_j , definida como:

$$R_j = \text{energía reflejada} / \text{energía incidente} .$$

En la reflexión en el vidrio r es real (dado que $\delta=0$ o $\delta=\pi$), mientras que en la reflexión metálica r es complejo. Si el medio es transparente t es real, o sea $t = t^*$, y en films metálicos t es complejo.

Además, 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 puede 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:

$$U_p = \sum_{j=1}^N A t_1 t_2 r_2^{(j-1)} r_1^{(j-1)} \exp(-i(2j-1)\phi) ,$$

y si el número de reflexiones es suficientemente elevado, esta suma se convierte en una serie geométrica, cuyo resultado es:

$$U_p = \frac{A t_1 t_2}{1 - r^2 \exp(-2i\phi)} e^{-i\phi} ,$$

y la intensidad correspondiente será:

$$I_M = \frac{I_0}{1 + F \sin^2(\phi/2)} ,$$

que es la llamada fórmula de Airy, donde

$$I_0 = T^2 / (1-R)^2 \quad ;$$

$$F = 4R / (1-R)^2 \quad ;$$

y además,

$$\phi' = 2\phi + 2\delta \quad .$$

El contraste, o visibilidad de las franjas es, en estos casos:

$$v = \frac{I_{\text{máx}} - I_{\text{mín}}}{I_{\text{máx}}} = \frac{F}{1+F} = \frac{4R}{(1+R)^2} \quad .$$

La Figura 13 muestra un gráfico de los perfiles de las franjas de interferencia para distintos valores de la reflectancia R .

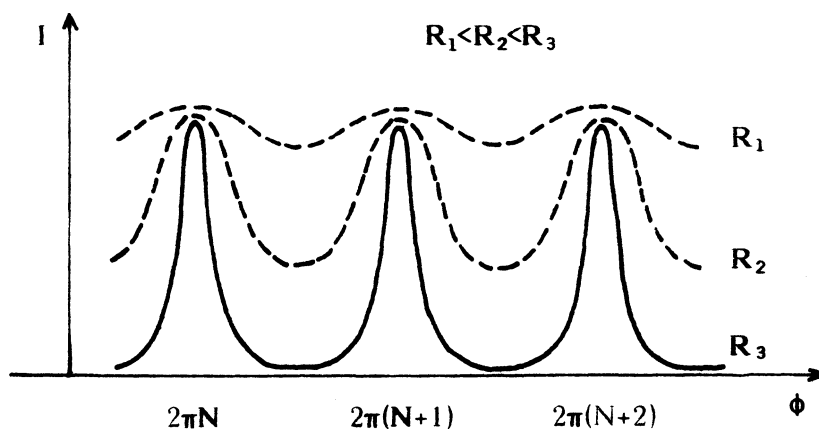


Figura 13

Si la onda que llega a la cavidad es una onda esférica, todas las ondas transmitidas 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 al esquema representado en la Figura 14.

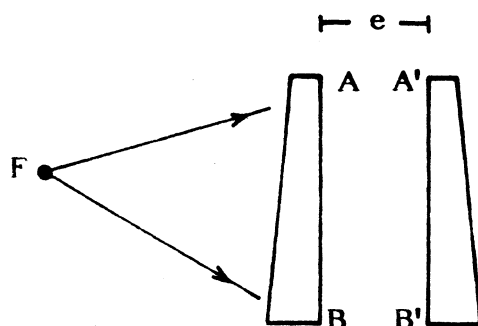


Figura 14

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 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 no será 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

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 se 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 la dirección x , podemos llamar Δx a este corrimiento o "shear" y se puede ver que si

$$s(x+\Delta x, y) - s(x, y) = m\lambda \quad ; \quad \text{con } m \in \mathbb{Z} \text{ (enteros)} \quad ,$$

donde $s(x+\Delta x, y)$ y $s(x, y)$ son los caminos ópticos recorridos por cada uno de los haces, habrá una franja de interferencia brillante.

De manera similar, si

$$s(x+\Delta x, y) - s(x, y) = (2m+1)\lambda/2 \quad : \quad \text{con } m \in \mathbb{Z} \text{ (enteros)} \quad ,$$

habrá una franja de interferencia oscura.

Estas relaciones pueden interpretarse como el promedio de la derivada del camino óptico respecto a la dirección del desplazamiento multi-

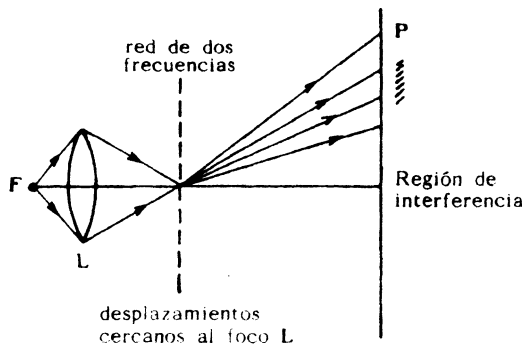
plicado por ese mismo desplazamiento, es decir, que pueden escribirse

$$\frac{\partial s(x,y)}{\partial x} \Delta x = \begin{cases} m \lambda & \text{(franja brillante)} \\ (2m+1)\lambda/2 & \text{(franja oscura)} \end{cases} .$$

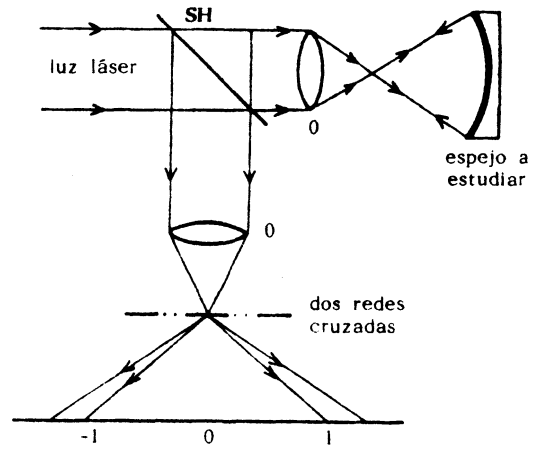
Esto pone en evidencia que los interferómetros de "shearing" lateral sólo dan información de la derivada del frente de onda en la dirección del "shear" o desplazamiento. Por lo tanto, si los frentes de onda que se desean estudiar no tienen simetría axial (en el plano xy), deberán realizarse dos interferogramas con direcciones de desplazamiento diferentes (obviamente, la interpretación será más sencilla si son ortogonales), a fin de describirlos completamente.

La Figura 15 muestra esquemas de dos ejemplos de interferómetros de "shear" lateral.

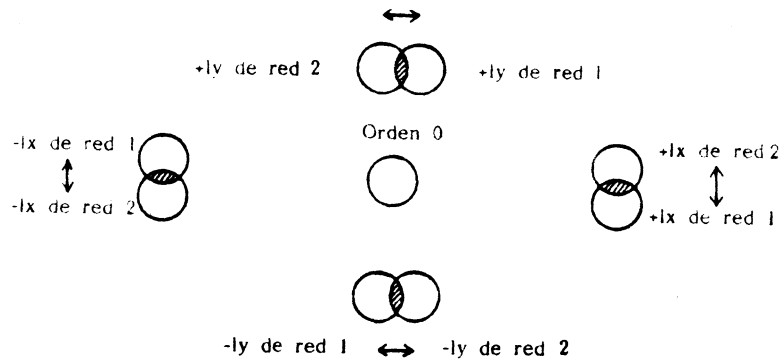
En el de la parte a), la interferencia se realiza por superposición de los frentes de onda difractados en el primer orden por dos redes paralelas de frecuencias espaciales distintas o, lo que es lo mismo, por una red que contenga dos frecuencias espaciales. Si se desea investigar el frente de onda en 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).



a)



(b)



c)

Figura 15

En el caso del "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.

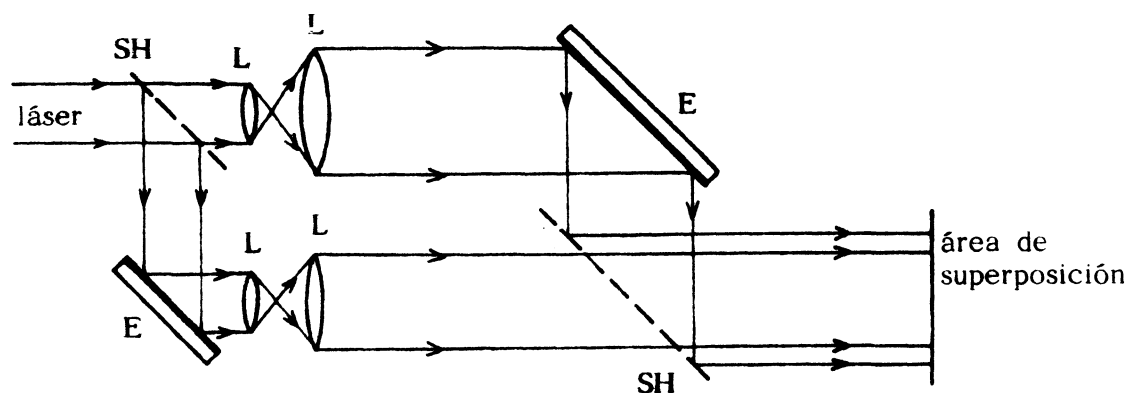


Figura 16

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 aparejados 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 semirreflectantes. Así, cuando se trabaja con iluminación láser, en lugar de semiespejos, es conveniente usar prisma-cubos o separadores "pellicle", o recubrimientos reflectantes 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 estudia para determinar la planitud o la calidad del pulido de superficies, la forma 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, y las variaciones de presión en fluidos, 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 adaptan para cada problema particular dependen tanto de las condiciones de trabajo como de las precisiones que se necesite alcanzar.

Difracción

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.

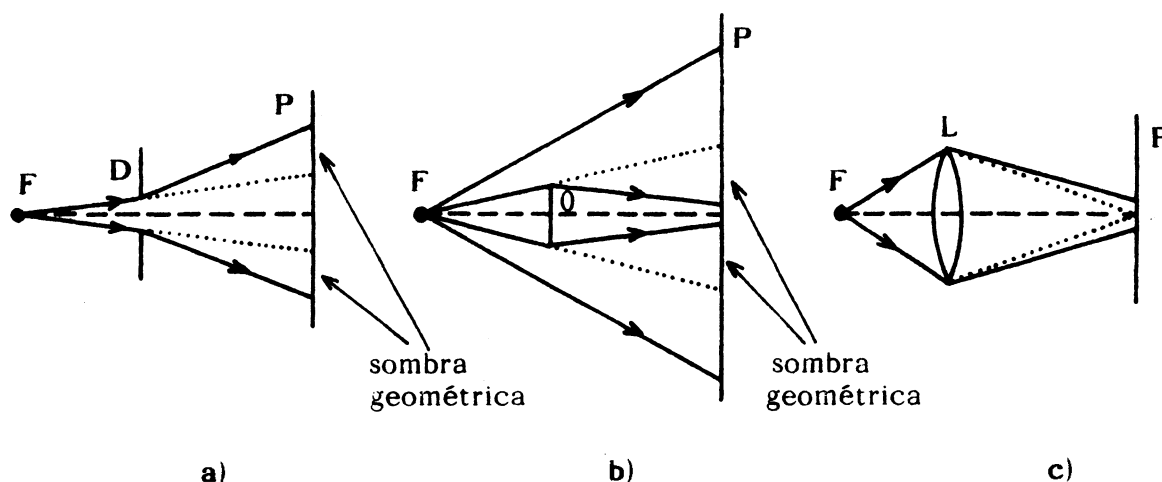


Figura 17

- a) Parte de la luz proveniente de una fuente F atraviesa un orificio practicado en el diafragma D y es recogida en una pantalla P .
- b) Parte de la luz proveniente de una fuente F es interceptada por un objeto opaco O y el resto es recogido en una pantalla P .
- c) 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.

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, 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 $t + \Delta t$ (ver Figura 18a). Este modelo de Huygens se complementa con la hipótesis de Fresnel 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 ondas 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 18b) muestra cómo el modelo Huygens-Fresnel permite explicar la difracción a través de un diafragma.

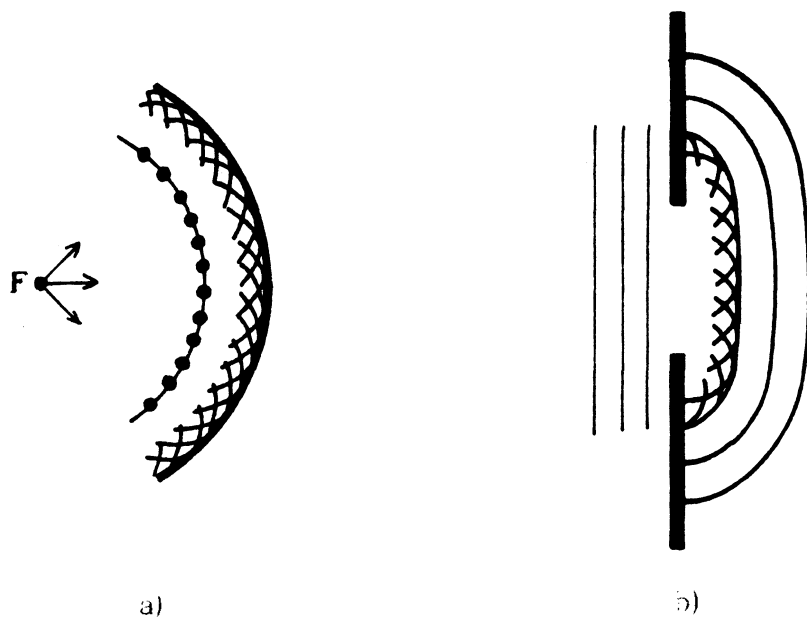


Figura 18

Formulación escalar de Fresnel-Kirchhoff. Aproximaciones:

El principio de Huygens-Fresnel es la base de la formulación matemática desarrollada por Kirchhoff, y permite calcular la amplitud del campo luminoso en el 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.

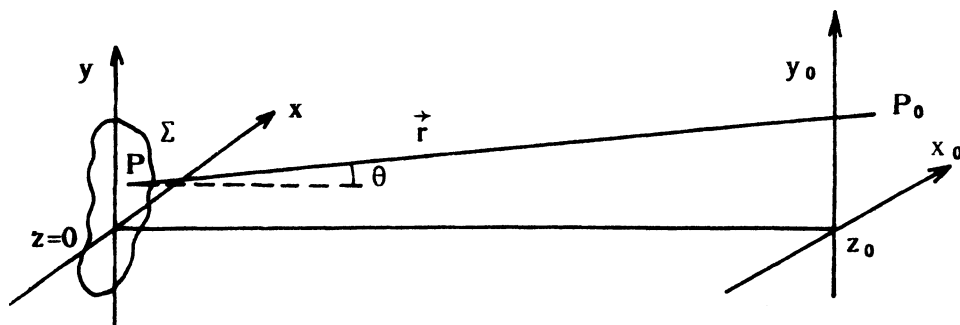


Figura 19



a)



b)

Foto 2. 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 abertura y generan a distancia un diagrama similar.

Consideremos que el diafragma colocado 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 .

Haciendo uso del principio de superposición, el campo escalar en el punto $P_0(x_0, y_0, z_0)$ se obtendrá sumando las contribuciones de cada una de las fuentes individuales P en el plano $z=0$, es decir,

$$U(x_0, y_0; z_0) = \frac{1}{2i\lambda} \iint_{\Sigma} U(x, y; 0) \frac{\exp ikr}{r} (1+\cos\theta) dx dy ,$$

donde $U(x, y; 0)$ es el valor del campo en cada punto del plano $z=0$,

$$r^2 = z_0^2 + (x_0 - x)^2 + (y_0 - y)^2 ,$$

donde θ es el ángulo entre $\vec{r} = \vec{P} \vec{P}_0$ y la normal en la dirección z , y

$$k = 2\pi/\lambda ,$$

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° ($1/i$).

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

$$U(x_0, y_0; z_0) = \frac{\exp ikz_0}{2i\lambda z_0} \iint_{\Sigma} U(x, y; 0) \exp \left\{ \frac{ik}{2z_0} [(x_0 - x)^2 + (y_0 - y)^2] \right\} dx dy ,$$

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 aún mayores, para las que es posible escribir

$$z_0 \gg k(x^2 + y^2)_{\text{máx}} / 2 ,$$

el campo se puede escribir

$$U(x_0, y_0; z_0) = C(x_0, y_0; z_0) \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} U(x, y; 0) \exp[-i2\pi(\mu x + \nu y)] dx dy .$$

donde hemos puesto

$$\mu = x_0 / \lambda z_0 \quad \text{y} \quad \nu = y_0 / \lambda z_0 ,$$

que son las frecuencias espaciales, y el factor

$$C(x_0, y_0; z_0) = (2i\lambda z_0)^{-1} \exp \left[ik \left(\frac{x_0^2 + y_0^2}{2z_0} + z_0 \right) \right] .$$

La última expresión para el campo difractado es la llamada Aproximación de Fraunhofer o de difracción en el infinito. 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 apertura sobre el plano $z=0$ no cambian el valor del campo en el plano $z=z_0$ 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

$$I = U(x_0, y_0; z_0) U^*(x_0, y_0; z_0) .$$

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,

$$U(u,v;z_0) = F\{U(x,y)\} (u,v) \quad ,$$

donde F significa transformada de Fourier, y la intensidad

$$I(u,v;z_0) = |F\{U(x,y)\} (u,v)|^2 \quad ,$$

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 un 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

$$M(x,y) = \exp \left[-\frac{ik}{2f} (x^2 + y^2) \right] \quad ,$$

donde f 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 el borde de la lente y el cambio de fase intro-

ducido 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 queda de la forma

$$U(\mu, \nu; f) = C(x_f, y_f; f) \iint_{-\infty}^{\infty} U(x, y; 0) \exp[-i2\pi(\mu x + \nu y)] dx dy ,$$

donde las frecuencias espaciales en el plano focal de la lente son

$$\mu = x_f / \lambda f \quad y \quad \nu = y_f / \lambda f .$$

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, observar una distribución en el plano focal de una lente corresponde a hacer infinita la distancia de observación de esa distribución. En la Figura 20 se esquematiza este resultado.

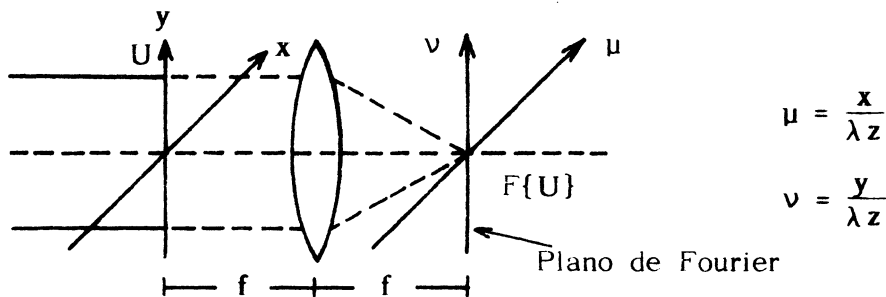


Figura 20

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.

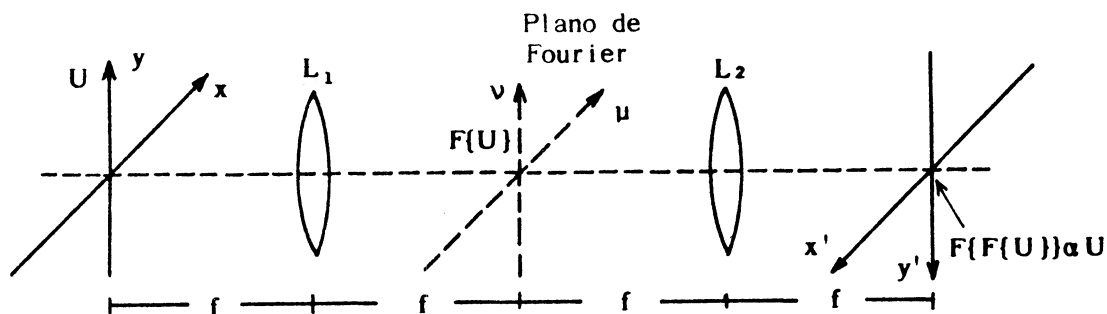


Figura 21

Estas técnicas de procesamiento óptico tienen aplicación en los campos más diversos, como la inspección de circuitos integrados (fotomáscaras), detección de defectos en objetos espacialmente periódicos, y procesamiento de imágenes de todo tipo.

Difracción de Fraunhofer de una abertura rectangular:

Si la apertura rectangular tiene dimensiones $x'y'$ y está iluminada por una onda plana, tenemos:

$$U(x,y) = \begin{cases} \text{rect}(x/x') \text{ rect}(y/y') = 1 & \text{dentro de la apertura} \\ 0 & \text{fuera de la apertura} \end{cases} ;$$

entonces el campo difractado en el infinito, o en el plano focal de una lente, puede calcularse como

$$U(u,v;f) = Q \text{ sinc}(\pi u x' / \lambda) \text{ sinc}(\pi v y' / \lambda) \quad .$$

El factor Q es proporcional a la superficie de la apertura. En general cualquiera sea la forma de la apertura, la amplitud del diagrama de interferencia en un punto es proporcional a la superficie de la apertura, y por consiguiente, la intensidad es proporcional al cuadrado de la superficie de la apertura.

En el caso de la apertura rectangular, entonces 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

$$\text{sinc}^2(\pi\mu x'/\lambda) \quad \text{en el eje } \mu$$

y

$$\text{sinc}^2(\pi\nu y'/\lambda) \quad \text{en el eje } \nu$$

El rectángulo central tiene dimensiones $(\lambda/x')(\lambda/y')$ 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 (Figura 22) de la función $\text{sinc}^2(\pi\mu x'/\lambda)$.

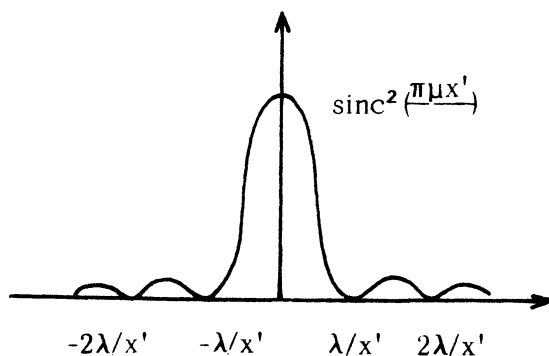


Figura 22

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 de 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 .

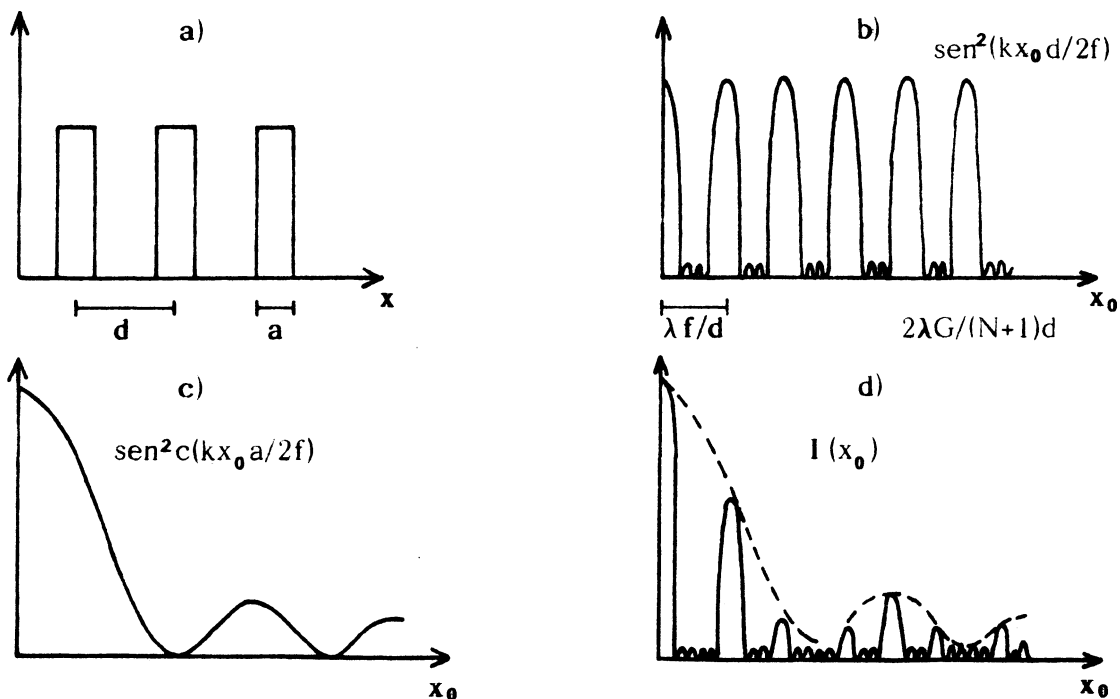


Figura 23

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 $x_0 a \pi / z_0$, 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 μm .

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.

En la tabla que se presenta en la Figura 24, 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.

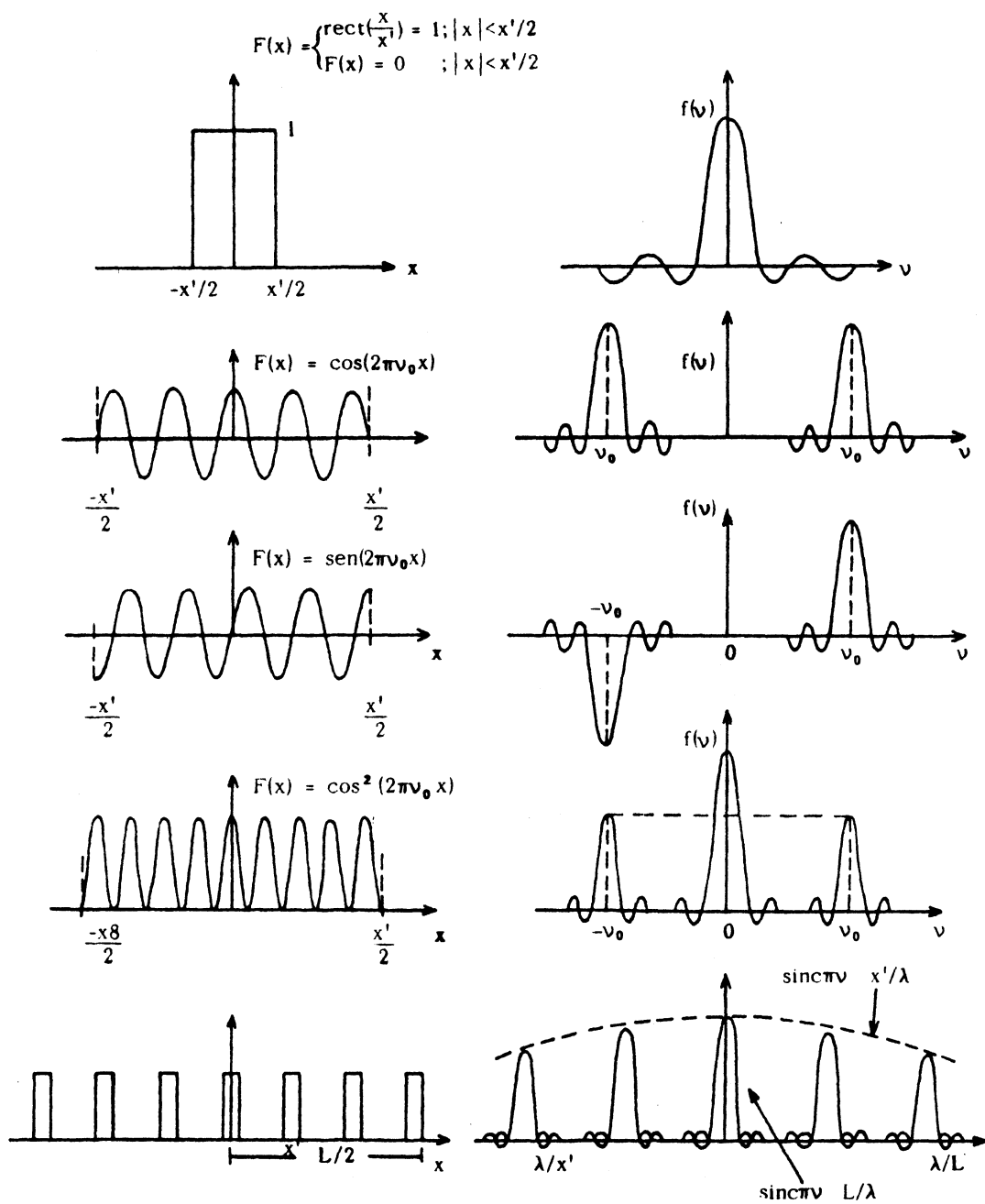


Figura 24

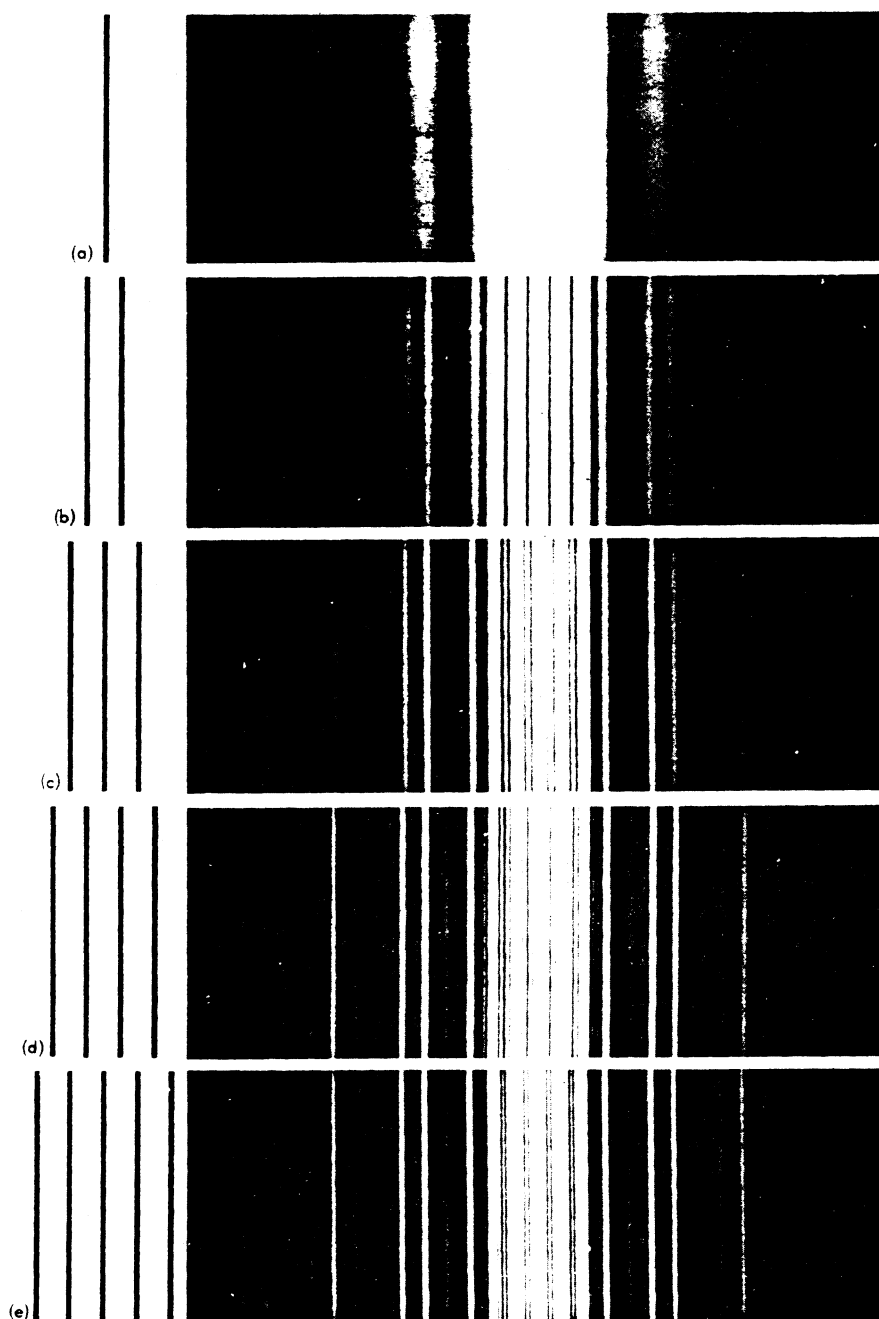


Foto 3. Distribución de intensidad de la luz difractada por la distribución de rendijas de la izquierda.

Se han representado sucesivamente una rendija unidimensional, una función coseno limitada por una rendija, una función seno y una función $\cos^2$ limitadas por la misma rendija. La función $\cos^2$ 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 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

$$I(x_0) = \frac{I_0}{(N+1)^2} \left\{ \frac{\text{sen}(N+1)(k/2f)x_0 d}{(kx_0 d/2f)} \right\}^2 \left\{ \frac{\text{sen}(kx_0 a/2f)}{(kx_0 a/2f)} \right\}^2$$

El primer factor entre llaves representa la figura de interferencia entre las $N+1$ rendijas (ver Figura 25b), 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 25c). El producto de ambos explica la distribución de intensidad del campo difractado que está representado en 25d) 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.

La información contenida en un arreglo periódico está contenida en regiones discretas y muy reducidas del plano de frecuencias. De esta manera el filtrado espacial puede resultar una herramienta 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.

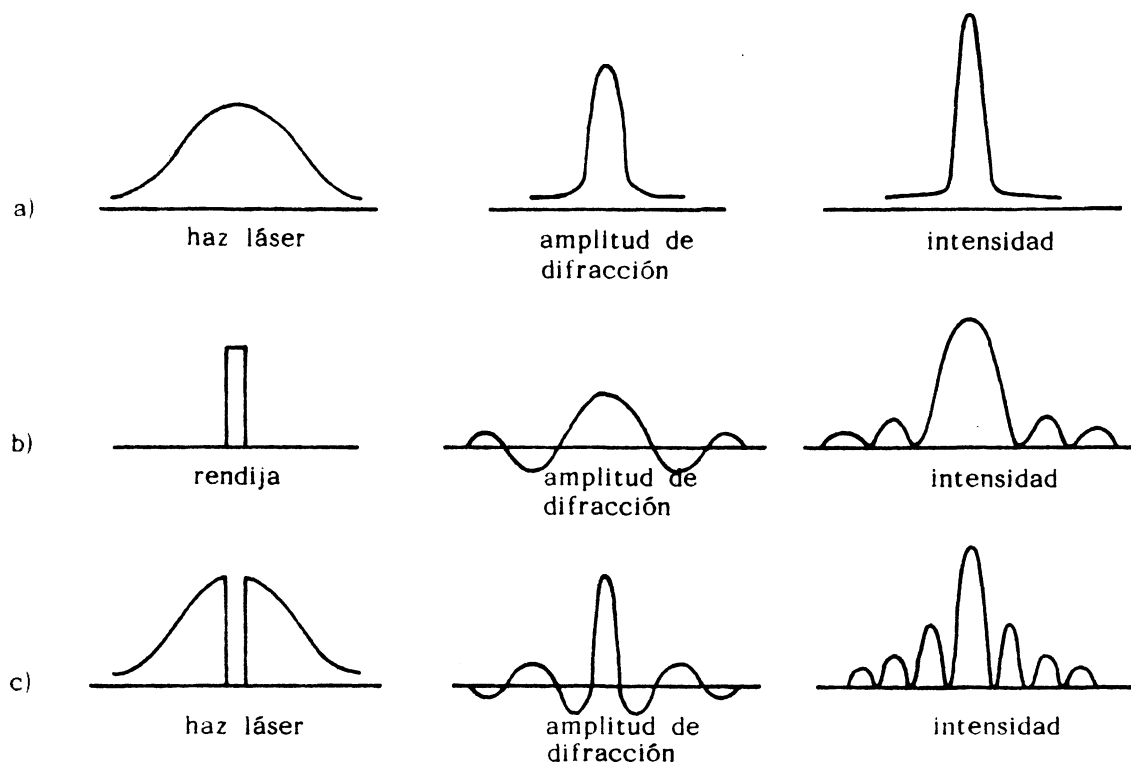


Figura 25

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", D.C. O'Shea, W.R. Callen y W.T. Rhodes, Addison-Wesley (1977).
- "Optique Formation et Traitement des Images", M. Francon, 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. Francon, Academic Press, New York (1966).
- "A Versatile 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.

ANALISIS DE FOURIER

B. Ruiz

Introducción

En este capítulo nos proponemos definir la transformada de Fourier y demostrar sus propiedades más útiles, pues se trata de una herramienta ampliamente utilizada en el estudio de circuitos eléctricos y sistemas de comunicación, en el caso de una variable independiente, y en el caso de dos variables independientes se la usa frecuentemente en los sistemas ópticos.

I - Transformada de Fourier unidimensional

Definición: Dada una función $g(t)$ de variable real t y a valores reales o complejos, se llama transformada de Fourier de $g(t)$ a la función de variable real f :

$$\int_{-\infty}^{\infty} g(t) \exp(-2\pi i f t) dt = G(f) = F\{g\}(f) \quad .$$

A $G(f)$ se la llama también espectro de Fourier o espectro de frecuencias de $g(t)$.

En forma análoga se define la transformada inversa de una función $G(f)$ como

$$F^{-1} G(t) = \int_{-\infty}^{\infty} G(f) \exp(2\pi i f t) df \quad .$$

Se prueba que si $F\{g\} = G$ entonces $F^{-1}\{G\} = g$ o en forma

equivalente

$$F^{-1}F\{g\} = g \quad , \quad F F^{-1}\{G\} = G \quad ,$$

cuando las funciones son continuas.

Antes de estudiar las propiedades de la transformada de Fourier se debe discutir su existencia.

Se conocen varios conjuntos de condiciones suficientes para la existencia de la transformada, uno de los más comunes es el siguiente:

1) g debe ser absolutamente integrable, o sea

$$\int_{-\infty}^{\infty} |g(t)| dt < \infty \quad .$$

2) g debe tener un número finito de discontinuidades y un número finito de máximos y mínimos en todo intervalo finito.

3) g no debe tener discontinuidades infinitas.

Bracewell sostiene que la cuestión de la existencia puede ser ignorada cuando la función a transformar es una descripción bastante exacta de una cantidad física.

Sin embargo, no se cumple 1) en ciertas situaciones físicas cuando hay energía infinita, por ejemplo corriente directa absolutamente estacionaria que fluye siempre; ondas armónicas, etc. No se cumple 3) cuando hay energía infinita concentrada en 0.

Para evitar estos inconvenientes se recurre a las distribuciones, lo que haremos más adelante.

En lo que sigue trataremos de hallar la transformada de Fourier de las funciones que habitualmente se utilizan en óptica.

Ejemplos:

1. La función rectángulo (función pupila), se define por medio de la expresión:

$$g(t) = \text{Rect}(t/t_0)/t_0 = \begin{cases} 1/t_0, & |t| < t_0/2 \\ 0, & |t| > t_0/2 \end{cases} .$$

La transformada de Fourier de g es:

$$G(f) = \frac{1}{t_0} \int_{-t_0/2}^{t_0/2} \exp(-2\pi i f t) dt = \frac{\exp(-2\pi i f t)}{-2\pi i f t_0} \Big|_{-t_0/2}^{t_0/2}$$

$$= \frac{\exp(-\pi i f t_0) - \exp(\pi i f t_0)}{-2\pi i f t_0} = \frac{\text{sen } \pi f t_0}{\pi f t_0} = \text{senc } \pi f t_0 .$$

Por lo tanto, la transformada de Fourier de la función pupila $g(t)$ es $G(f)$; se la conoce con el nombre de función filtrado o interpolación.

La representación gráfica de estas dos funciones se muestra en la Figura 1.

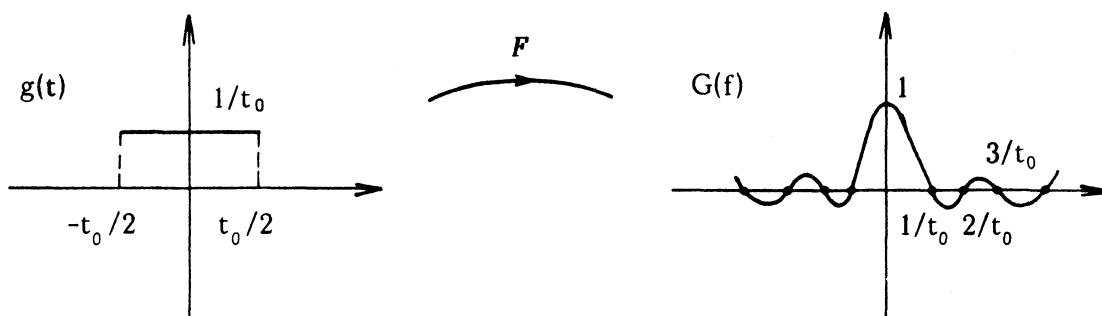


Figura 1

El par $g(t) \rightarrow G(f)$ se llama par transformado.

2. La transformada de la función gaussiana de probabilidad, definida por:
 $g(x) = C e^{-ax^2}$ con $C = \sqrt{a/\pi}$ y a constante, es

$$G(u) = C \int_{-\infty}^{\infty} e^{-ax^2} e^{-2\pi i x u} dx = C \int_{-\infty}^{\infty} e^{-(ax^2 + 2\pi i x u)} dx .$$

Completando cuadrados en el exponente:

$$ax^2 + 2\pi iux = (\sqrt{a}x + \pi iu/\sqrt{a})^2 + \pi^2 u^2/a ,$$

y haciendo $v = \sqrt{a}x + \pi iu/\sqrt{a}$, se obtiene

$$G(u) = \frac{C}{\sqrt{a}} e^{-\pi^2 u^2/a} \int_{-\infty}^{\infty} e^{-v^2} dv = \frac{e^{-\pi^2 u^2/a}}{\sqrt{\pi}} \cdot \sqrt{\pi} ,$$

$$G(u) = e^{-\pi^2 u^2/a} ,$$

que es nuevamente una función gaussiana.

La desviación estándar se define como el valor de la variable (x ó u) para la cual la exponencial toma el valor $e^{-1/2} = 0,607$. Por lo tanto,

$$\sigma_x = \frac{1}{\sqrt{2a}} \quad y \quad \sigma_u = \frac{\sqrt{a/2}}{\pi} .$$

El gráfico de la Figura 2 representa el par transformado:

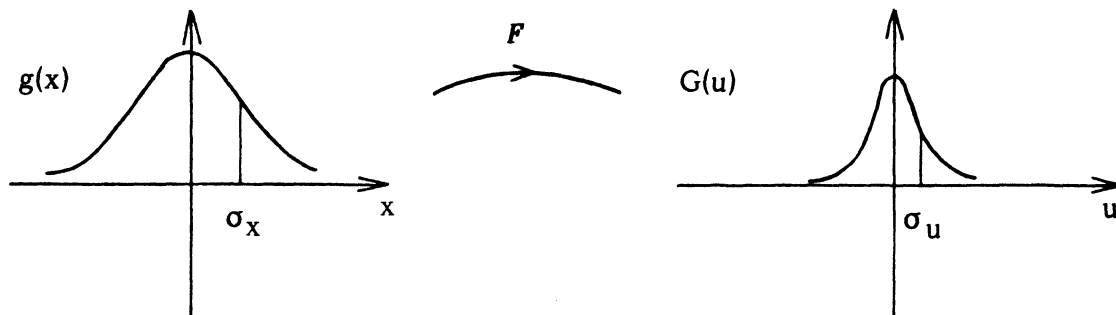


Figura 2

3. La función coseno limitada, definida por:

$$g(t) = \begin{cases} 2/t_0 \cdot \cos 2\pi f_0 t & , \quad |t| < t_0/2 \\ 0 & , \quad |t| > t_0/2 \end{cases}$$

tiene como transformada de Fourier a

$$G(f) = \frac{2}{t_0} \int_{-t_0/2}^{t_0/2} \frac{\exp(2\pi i f_0 t) + \exp(-2\pi i f_0 t)}{2} \exp(-2\pi i f t) dt$$

$$G(f) = \frac{1}{t_0} \int_{-t_0/2}^{t_0/2} \{ \exp[-2\pi i (f+f_0)t] + \exp[-2\pi i (f-f_0)t] \} dt$$

$$G(f) = \text{senc} [\pi (f+f_0) t_0] + \text{senc} [\pi (f-f_0) t_0] \quad .$$

La representación gráfica de este par transformado es el de la Figura 3.

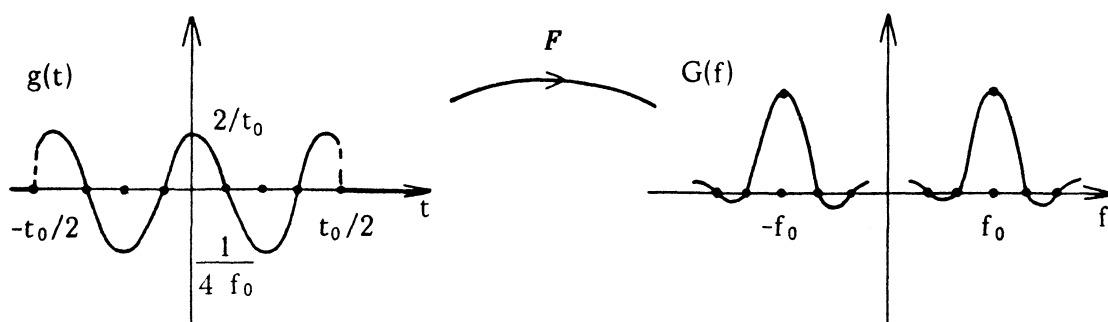


Figura 3

4. En forma análoga al ejemplo anterior, la función seno limitada, definida por

$$g(t) = \begin{cases} 2i/t_0 \operatorname{sen} 2\pi f_0 t & , \quad |t| < t_0/2 \\ 0 & , \quad |t| > t_0/2 \end{cases}$$

tiene como transformada de Fourier, a la función:

$$G(f) = \operatorname{senc} [\pi(f-f_0)t_0] - \operatorname{senc} [\pi(f+f_0)t_0] \quad .$$

Este par transformado tiene la representación gráfica ilustrada en la Figura 4.

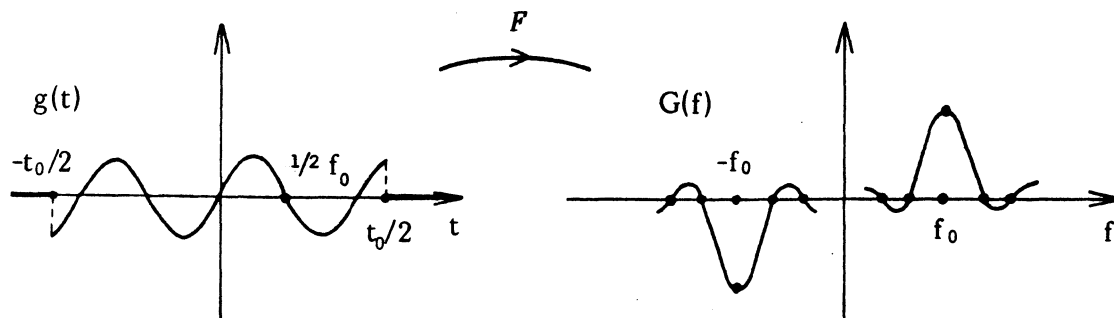


Figura 4

II - Transformada de Fourier bidimensional

Definición: La transformada de Fourier de una función real o compleja, $g(x,y)$ de dos variables reales independientes, es

$$\iint_{-\infty}^{\infty} g(x,y) \exp[-2\pi i(xu+yv)] dx dy = G(u,v) = F\{g\}(u,v) \quad .$$

Se la llama también espectro de Fourier o espectro de frecuencias de $g(x,y)$.

Se define la transformada inversa de Fourier como

$$\iint_{-\infty}^{\infty} g(u,v) \exp[2\pi i(xu+yv)] du dv = F^{-1}\{G\}(x,y) \quad .$$

Se demuestra que si $F\{g\} = G$ entonces $F^{-1}\{G\} = g$ o equivalentemente $F^{-1}F\{g\} = g$, $FF^{-1}\{G\} = G$, en el caso en que las funciones son continuas.

Las condiciones de existencia de la transformada bidimensional son análogas a las de la transformada unidimensional.

Ejemplos:

1. La función pupila rectangular se define mediante la expresión:

$$g(x,y) = 1/x_0 y_0 \cdot \text{Rect}(x/x_0) \text{Rect}(y/y_0) =$$

$$= \begin{cases} 1/x_0 y_0 & ; \quad |x| < x_0/2 \quad , \quad |y| < y_0/2 \\ 0 & ; \quad |x| > x_0/2 \quad \text{ó} \quad |y| > y_0/2 \end{cases} .$$

La transformada de Fourier de g es G , dada por:

$$G(u,v) = 1/x_0 \cdot \int_{-x_0/2}^{x_0/2} \exp(-2\pi i x u) dx \cdot 1/y_0 \cdot \int_{-y_0/2}^{y_0/2} \exp(-2\pi i y v) dy =$$

$$= \text{senc}(\pi x_0 u) \text{senc}(\pi y_0 v) \quad .$$

2. En este ejemplo nos proponemos calcular la transformada de la función pupila circular; para esto utilizamos una sucesión de cálculos de uso frecuente cuando hay simetría circular.

La función pupila circular (normalizada) se define mediante la expresión:

$$g(x,y) = 1/\pi r_0^2 \cdot \text{Circ}_{r_0}(x,y) = \begin{cases} 1/\pi r_0^2 & , \quad x^2 + y^2 < r_0^2 \\ 0 & , \quad x^2 + y^2 > r_0^2 \end{cases} = g(r)$$

recordando que $x^2 + y^2 = r^2$.

Su transformada de Fourier es

$$G(u,v) = 1/\pi r_0^2 \cdot \iint_{K_{r_0}} \exp[-2\pi i(xu+yv)] dx dy ,$$

siendo K_{r_0} el círculo de radio r_0 .

Haciendo una transformación a coordenadas polares en ambos planos (x,y) y (u,v) :

$$\begin{cases} x = r \cos \theta \\ y = r \sin \theta \end{cases} \quad \begin{cases} u = \rho \cos \phi \\ v = \rho \sin \phi \end{cases}$$

obtenemos

$$xu + yv = \rho r (\sin \phi \sin \theta + \cos \phi \cos \theta) = \rho r \cos(\phi - \theta) ,$$

de donde

$$G(u,v) = 1/\pi r_0^2 \cdot \int_0^{r_0} \int_0^{2\pi} r \exp[-2\pi i \rho r \cos(\theta - \phi)] d\theta dr .$$

Utilizando la siguiente expresión integral para la función de Bessel de primera clase y orden cero

$$J_0(z) = 1/2\pi \cdot \int_0^{2\pi} \exp[-iz \cos(\theta - \phi)] d\theta$$

obtenemos

$$G(\rho) = 2/r_0^2 \cdot \int_0^{r_0} r J_0(2\pi \rho r) dr .$$

Teniendo en cuenta que $z J_0(z) = \frac{d}{dz} [(z J_1(z))]$

$$\begin{aligned}
 G(\rho) &= \frac{2}{r_0^2} \frac{1}{(2\pi\rho)^2} \int_{r=0}^{r=r_0} 2\pi\rho r J_0(2\pi\rho r) d(2\pi\rho r) = \\
 &= \frac{2}{r_0^2} \frac{2\pi\rho r_0}{(2\pi\rho)^2} J_1(2\pi\rho r_0) = 2 \frac{J_1(2\pi\rho r_0)}{2\pi\rho r_0}
 \end{aligned}$$

Se observa que la transformada de una función $g(r)$, de simetría cilíndrica, es una función $G(\rho)$ también de simetría cilíndrica. La utilización de coordenadas apropiadas condujo a las funciones cilíndricas $J_0(z)$, $J_1(z)$.

El par transformado se puede ver de dos formas:

a) Tomando como variable independiente r , en el plano original y ρ , en el transformado, obtenemos gráficamente

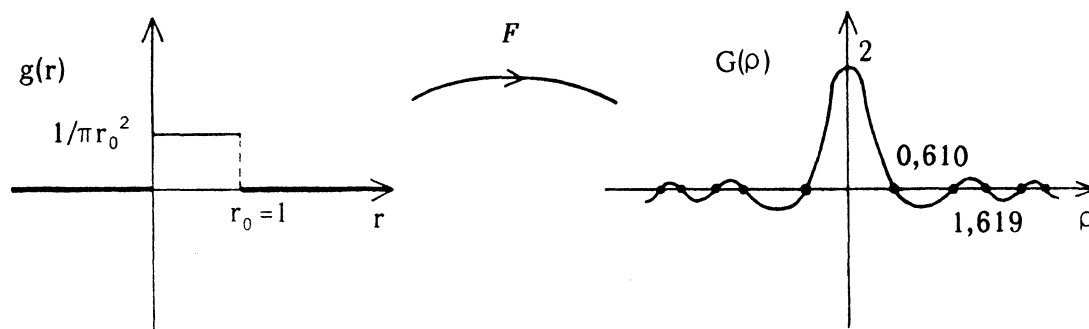


Figura 5

b) Si consideramos dos dimensiones, se evidencia la simetría circular en ambos espacios.

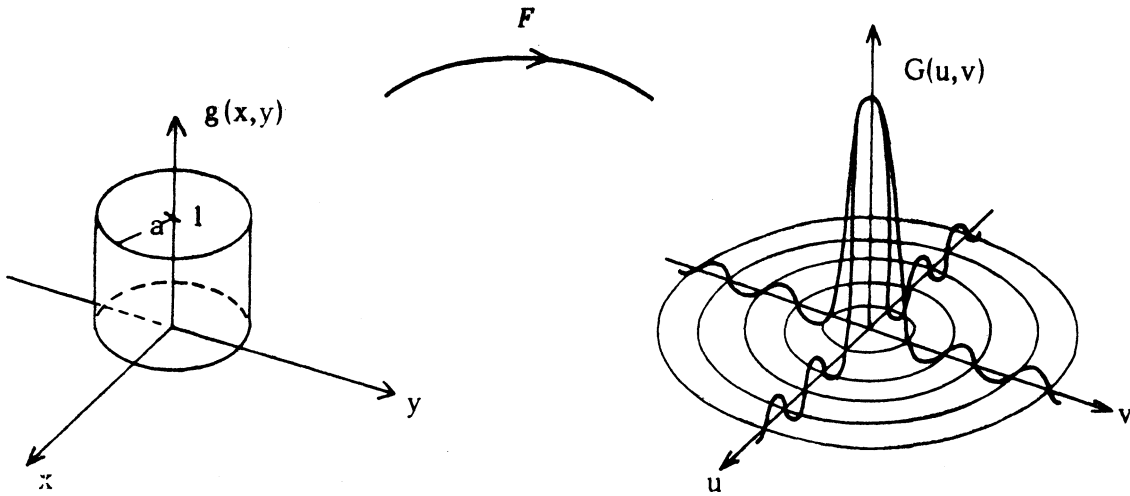


Figura 6. Función cilíndrica o tapa de sombrero y su transformada.

III. Propiedades

A continuación se demuestran las propiedades fundamentales de la transformada de Fourier bidimensional.

Eliminando una de las variables independientes se obtienen las propiedades correspondientes a la transformada unidimensional.

Teorema de linealidad

$$F\{\alpha g + \beta h\} = \alpha F\{g\} + \beta F\{h\} \quad .$$

Este teorema es consecuencia de la linealidad de las integrales que definen la transformada.

Ejemplo:

Si en el ejemplo 1 del § I tomamos $t_1 = t_0/2$, obtenemos el par transformado:

$$\frac{1}{t_1} \text{Rect} \left(\frac{t}{t_1} \right) \rightarrow \text{senc } \pi f t_1 \quad .$$

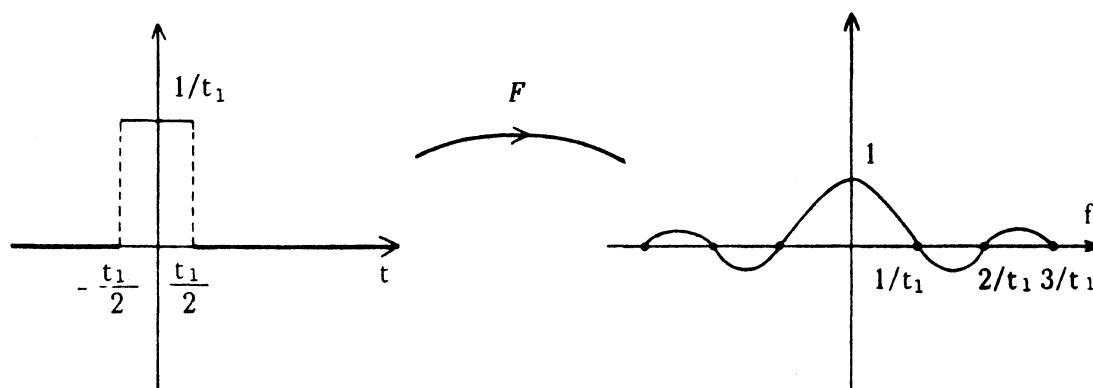


Figura 7

Si necesitamos la transformada de

$$g(t) = \frac{1}{t_0} \text{Rect}\left(\frac{t}{t_0}\right) + \frac{1}{t_1} \text{Rect}\left(\frac{t}{t_1}\right)$$

es conveniente utilizar la propiedad de linealidad para obtener

$$G(f) = \text{senc } \pi f t_0 + \text{senc } \pi f t_1 ,$$

cuya gráfica es la suma de las gráficas obtenidas.

Hemos construido así, a partir de resultados conocidos, el nuevo par transformado: $g(t) \rightarrow G(f)$, que muestran los gráficos de la Figura 8.

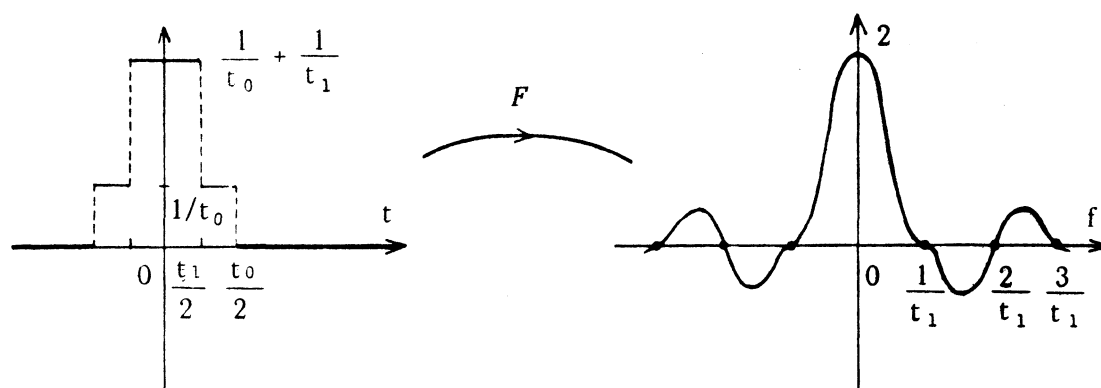


Figura 8

Teorema de dilatación

Si $F\{g\} = G$, entonces

$$F\{g(ax,by)\}(u,v) = 1/|ab| \cdot G(u/a, v/b) .$$

Demostración

$$\begin{aligned} F\{g(ax,by)\}(u,v) &= \iint_{-\infty}^{\infty} g(ax,by) \exp[-2\pi i(xu+yv)] dx dy = \\ &= \iint_{-\infty}^{\infty} g(ax,by) \exp[-2\pi i(ax \cdot u/a + by \cdot v/b)] \cdot \frac{dax}{|a|} \frac{dby}{|b|} = \\ &= 1/|ab| \cdot G(u/a, v/b) . \end{aligned}$$

Teorema de traslación

Si $F\{g\} = G$, entonces

$$a) \quad F\{g(x-a,y-b)\}(u,v) = G(u,v) \exp[-2\pi i(au+bv)] .$$

$$b) \quad F\{\exp[2\pi i(ax+by)] \cdot g(x,y)\}(u,v) = G(u-a, v-b) .$$

Demostración

$$\begin{aligned} a) \quad F\{g(x-a,y-b)\}(u,v) &= \iint_{-\infty}^{\infty} g(x-a,y-b) \exp[-2\pi i(xu+yv)] dx dy = \\ &= \iint_{-\infty}^{\infty} g(\zeta,\eta) \exp[-2\pi i(\zeta+a)u + (\eta+b)v] d\zeta d\eta = \\ &= G(u,v) \exp[-2\pi i(au+bv)] . \end{aligned}$$

$$\begin{aligned} \text{b) } F\{\exp[2\pi i(ax+by)].g(x,y)\} (u,v) &= \iint_{-\infty}^{\infty} g(x,y) \exp[-2\pi i(x(u-a)+y(v-b))] dx dy = \\ &= G(u-a, v-b) . \end{aligned}$$

Teorema de modulación

Si $F\{g\} = G$ entonces

$$\begin{aligned} F\{g(x,y) [\cos(u_0 x) \cos(v_0 y)]\} &= \frac{1}{4} [G(u + \frac{u_0}{2\pi}, v + \frac{v_0}{2\pi}) + \\ &+ G(u - \frac{u_0}{2\pi}, v - \frac{v_0}{2\pi}) + G(u - \frac{u_0}{2\pi}, v + \frac{v_0}{2\pi}) + G(u + \frac{u_0}{2\pi}, v - \frac{v_0}{2\pi})] . \end{aligned}$$

Demostración

$$\begin{aligned} F\{\frac{g(x,y)}{4} [(e^{iu_0 x} + e^{-iu_0 x})(e^{iv_0 y} + e^{-iv_0 y})]\} (u,v) &= \\ = \frac{1}{4} F\{g(x,y) e^{i(u_0 x + v_0 y)} + g(x,y) e^{-i(u_0 x + v_0 y)} + g(x,y) e^{i(u_0 x - v_0 y)} + \\ &+ g(x,y) e^{-i(u_0 x - v_0 y)}\} (u,v) = \\ = \frac{1}{4} [G(u - \frac{u_0}{2\pi}, v - \frac{v_0}{2\pi}) + G(u + \frac{u_0}{2\pi}, v + \frac{v_0}{2\pi}) + G(u - \frac{u_0}{2\pi}, v + \frac{v_0}{2\pi}) + G(u + \frac{u_0}{2\pi}, v - \frac{v_0}{2\pi})] \end{aligned}$$

Definición: Dadas dos funciones $g(x,y)$, $h(x,y)$ se llama convolución de g con h y se anota $g * h$, a la función

$$(g * h)(x,y) = \iint_{-\infty}^{\infty} g(\zeta, \eta) h(x-\zeta, y-\eta) d\zeta d\eta .$$

Se puede verificar fácilmente que $g * h = h * g$.

Ejemplo:

Trataremos de obtener gráficamente la convolución de

$$g(\zeta) = \begin{cases} \exp(-\zeta) & , \quad \zeta > 0 \\ 0 & , \quad \zeta < 0 \end{cases} \quad \text{con} \quad h(\zeta) = \begin{cases} 1 & ; \quad \zeta > 0 \\ 0 & ; \quad \zeta < 0 \end{cases}$$

$$(g * h)(x) = \int_{-\infty}^{\infty} g(\zeta) h(x - \zeta) d\zeta ;$$

gráficamente:

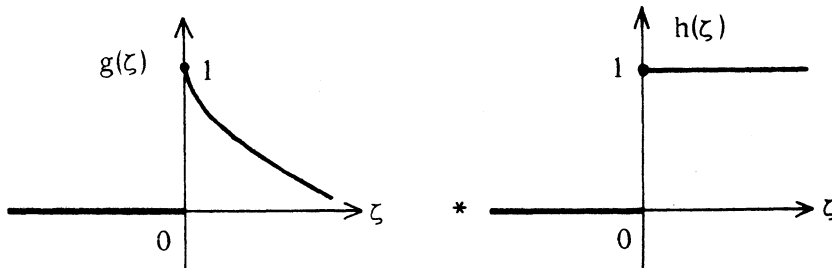


Figura 9

Esta convolución implica los cuatro pasos de las Figuras 10 y 11.

Paso 1º: reflexión

Paso 2º: traslación

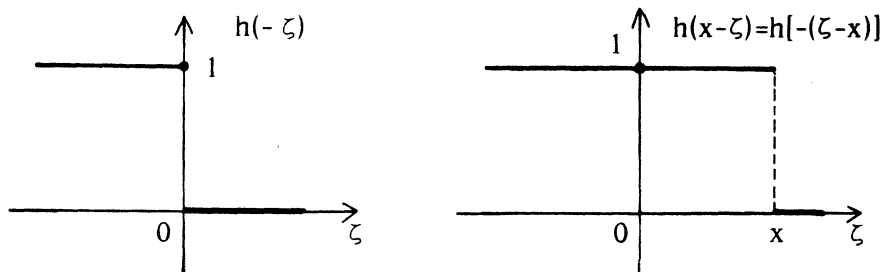
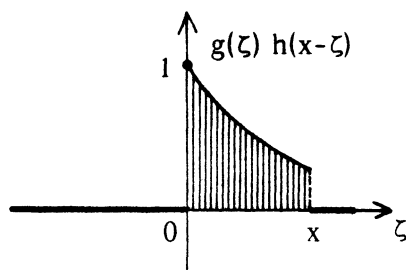


Figura 10

Paso 3º: multiplicación



Paso 4º: integración

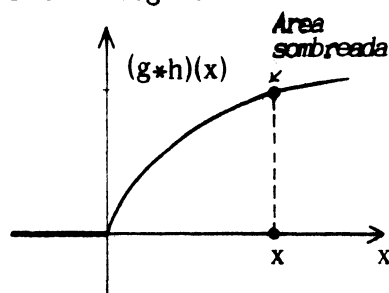


Figura 11

Teorema de convoluciónSi $F\{g\} = G$ y $F\{h\} = H$, entonces

$$F\{g*h\} = G.H.$$

Demostración

$$\begin{aligned}
 & F\left\{ \iint_{-\infty}^{\infty} g(\zeta, \eta) h(x-\zeta, y-\eta) d\zeta d\eta \right\} (u, v) = \\
 & = \iiint_{-\infty}^{\infty} g(\zeta, \eta) h(x-\zeta, y-\eta) \exp[-2\pi i(xu + yv)] d\zeta d\eta dx dy = \\
 & = \iint_{-\infty}^{\infty} g(\zeta, \eta) \left[\iint_{-\infty}^{\infty} h(x-\zeta, y-\eta) \exp[-2\pi i(xu + yv)] dx dy \right] d\zeta d\eta = \\
 & = \iint_{-\infty}^{\infty} g(\zeta, \eta) \left[\iint_{-\infty}^{\infty} h(\alpha, \beta) \exp[-2\pi i\{(\alpha+\zeta)u + (\beta+\eta)v\}] d\alpha d\beta \right] d\zeta d\eta = \\
 & = \iint_{-\infty}^{\infty} g(\zeta, \eta) \exp[-2\pi i(\zeta u + \eta v)] d\zeta d\eta \iint_{-\infty}^{\infty} h(\alpha, \beta) \exp[-2\pi i(\alpha u + \beta v)] d\alpha d\beta = \\
 & = G(u, v) H(u, v).
 \end{aligned}$$

Definición: Dadas las funciones $g(x, y)$, $h(x, y)$ se llama correlación de g con h a la función

$$\iint_{-\infty}^{\infty} g(\zeta, \eta) h^*(\zeta - x, \eta - y) d\zeta d\eta = (g * h^*)(x, y) ,$$

con $h_-(x, y) = h(-x, -y)$ y donde $*$ denota conjugación.

Teorema de autocorrelación

Si $F\{g\} = G$, entonces

$$F\left\{\iint_{-\infty}^{\infty} g(\zeta, \eta) g^*(\zeta - x, \eta - y) d\zeta d\eta\right\}(u, v) = |G(u, v)|^2 .$$

Demostración

$$\begin{aligned} F\left\{\iint_{-\infty}^{\infty} g(\zeta, \eta) g^*(\zeta - x, \eta - y) d\zeta d\eta\right\}(u, v) &= F\{g(x, y) * g^*(-x, -y)\}(u, v) = \\ &= G(u, v) F\{g^*(-x, -y)\}(u, v) = G(u, v) G^*(u, v) = |G(u, v)|^2 , \\ \text{pues} \\ \left[F\{g^*(-x, -y)\}(u, v)\right]^* &= \left[\iint_{-\infty}^{\infty} g^*(-x, -y) \exp[-2\pi i(xu + yv)] dx dy\right]^* = \\ &= \iint_{-\infty}^{\infty} g(-x, -y) \exp[2\pi i(xu + yv)] dx dy = \iint_{-\infty}^{\infty} g(\alpha, \beta) \exp[-2\pi i(\alpha u + \beta v)] d\alpha d\beta = \\ &= G(u, v) . \end{aligned}$$

IV - Distribución de Dirac en una y dos dimensiones.

Definición: La distribución de Dirac, anotada δ_{t_0} , es un funcional lineal que asigna a una función ϕ de variable t , continua en t_0 , el valor $\phi(t_0)$.

$$\delta_{t_0}(\phi) = \phi(t_0) = \int_{-\infty}^{\infty} \phi(t) \delta(t-t_0) dt = \int_{-\infty}^{\infty} \phi(t) \delta_{t_0}(t) dt \quad .$$

Las dos últimas expresiones son simbólicas y no tienen significado como integral, sólo se pretende que δ tenga propiedades formalmente iguales a las de las integrales.

En particular, se define $\delta(\phi) = \phi(0)$.

Definición: La transformada de Fourier de δ_{t_0} , es

$$F\{\delta_{t_0}\}(f) = \int_{-\infty}^{\infty} \delta_{t_0}(t) e^{-2\pi i f t} dt = \delta_{t_0}(e^{-2\pi i f t}) = e^{-2\pi i f t_0} \quad .$$

La transformada inversa de Fourier de δ_{f_0} , es:

$$F^{-1}\{\delta_{f_0}\}(t) = \delta_{f_0}(e^{2\pi i f t}) = e^{2\pi i f_0 t} \quad .$$

Resultan, por lo tanto

$$F\{\delta\}(f) = \delta(e^{-2\pi i f t}) = 1$$

$$F\{e^{2\pi i f_0 t}\} = \delta_{f_0} \quad .$$

Ejemplos:

1. Utilizando las definiciones anteriores, la transformada de Fourier de una suma de distribuciones δ es:

$$F\{\delta_{t_0} + \delta_{-t_0}\}(f) = \exp[-2\pi i f t_0] + \exp[2\pi i f t_0] = 2 \cos(2\pi f t_0) \quad ,$$

que se expresa gráficamente en la Figura 12.

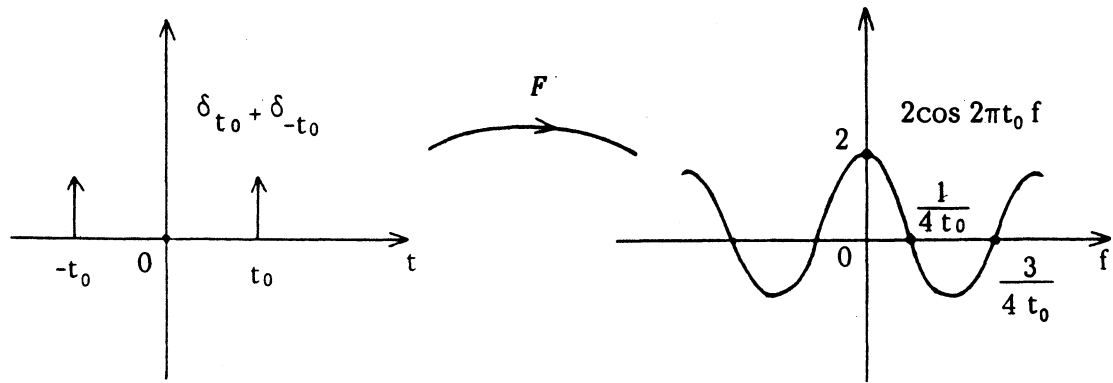


Figura 12

En el primer gráfico tenemos en cuenta que la distribución δ_{t_0} se interpreta como representando la masa 1 en el punto t_0 .

2. En forma análoga obtenemos

$$F \{ \delta_{-t_0} - \delta_{t_0} \} (f) = \exp(2\pi i f t_0) - \exp(-2\pi i f t_0) = 2i \operatorname{sen}(2\pi f t_0) .$$

Gráficamente:

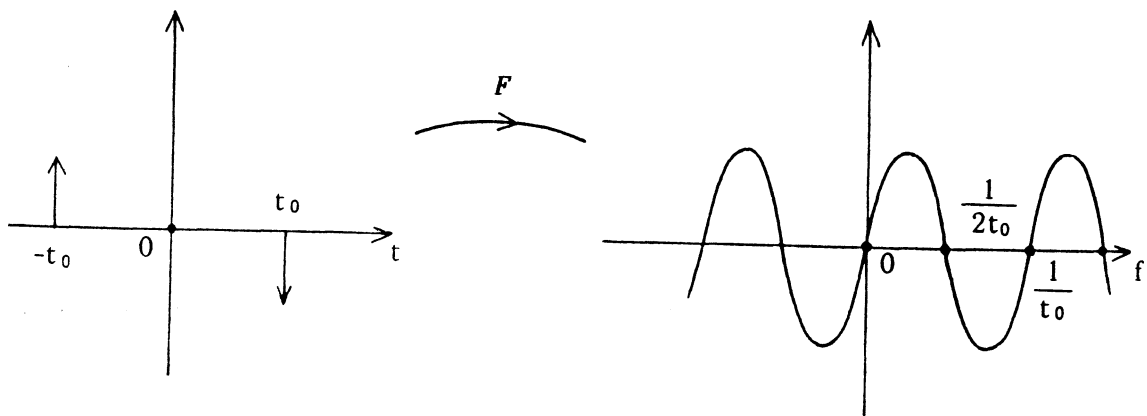


Figura 13

Definición: La convolución de una función $g(t)$ con δ_{t_0} , es

$$(g * \delta_{t_0})(t) = \int_{-\infty}^{\infty} g(t - \zeta) \delta_{t_0}(\zeta) d\zeta \quad .$$

Como corolario de la definición se deduce que

$$(g * \delta_{t_0})(t) = \delta_{t_0}[g(t - \zeta)] = g(t - t_0) = g_{t_0}(t) \quad .$$

o sea

$$g * \delta_{t_0} = g_{t_0} \quad ,$$

lo que significa que convolucionando una función $g(t)$ con una δ_{t_0} resulta la función trasladada en t_0 .

Se expresa gráficamente el resultado anterior en la Figura 14.

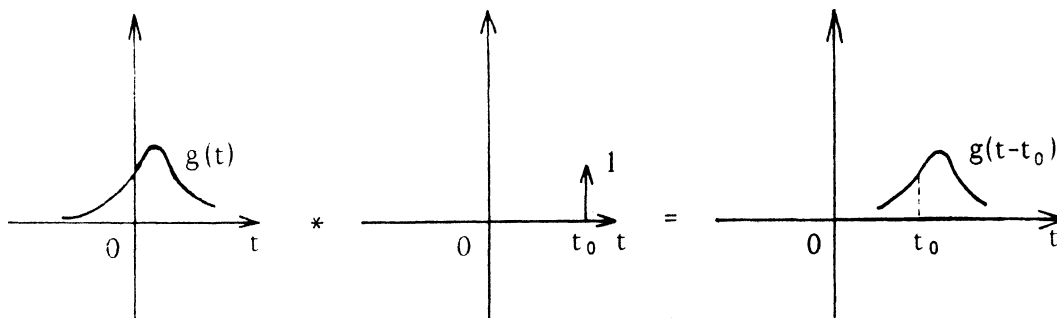


Figura 14

Definición: Se llama peine de Dirac a una serie de distribuciones de Dirac de la misma intensidad.

$$\sum_{k=-\infty}^{\infty} \delta_{k t_0} \quad , \quad \text{con} \quad \delta_0 = \delta \quad .$$

Se demuestra que la transformada de Fourier de un peine de Dirac de período t_0 es otro peine de Dirac de período $1/t_0$ y de intensidad $1/t_0$.

$$F \left\{ \sum_{k=-\infty}^{\infty} \delta_{kt_0} \right\} = \frac{1}{t_0} \sum_{k=-\infty}^{\infty} \delta_{k/t_0}$$

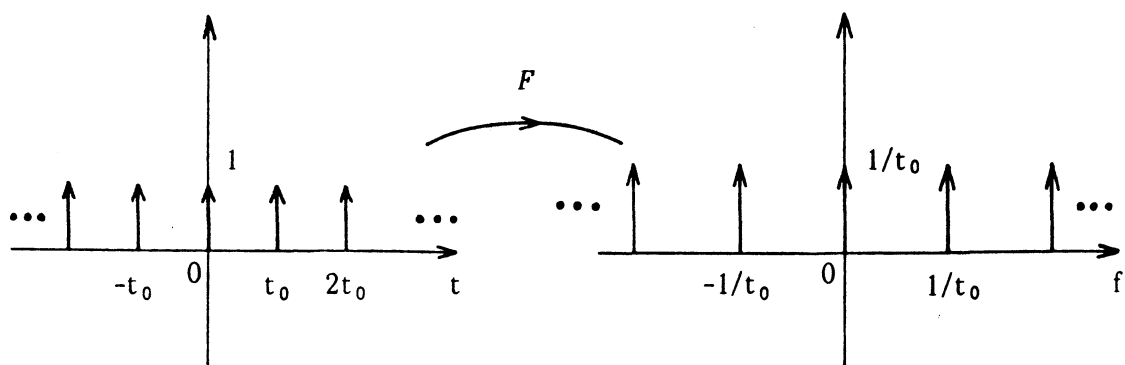


Figura 15

Asimismo se demuestra que la convolución de una función g con un peine de Dirac de período t_0 es una función periódica de período t_0 .

$$g * \sum_{k=-\infty}^{\infty} \delta_{kt_0} = \sum_{k=-\infty}^{\infty} g_{kt_0} \quad .$$

Estos resultados pueden generalizarse para dos dimensiones.

Definición:

$$\begin{aligned} \delta_{x_0, y_0}(g) &= g(x_0, y_0) = \iint_{-\infty}^{\infty} g(x, y) \delta(x - x_0, y - y_0) dx dy = \\ &= \iint_{-\infty}^{\infty} g(x, y) \delta_{x_0, y_0}(x, y) dx dy \quad . \end{aligned}$$

Propiedades

$$F \{ \delta_{x_0, y_0} \} (u, v) = \exp[-2\pi i(x_0 u + y_0 v)]$$

$$F \{ \delta \} (u, v) = 1$$

$$F \{ \exp[2\pi i(xu_0 + yv_0)] \} = \delta_{u_0, v_0}$$

$$F \left\{ \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} \delta_{mx_0} \delta_{ny_0} \right\} = 1/x_0 y_0 \cdot \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} \delta_{m/x_0} \delta_{n/y_0} \cdot$$

Nota: En algunas ocasiones se ha identificado a la función con el valor de la función en un punto, o sea, se ha puesto $g(t)$ en lugar de g , a fin de no complicar la notación.

Bibliografía

- Papoulis, A. - "The Fourier integral and its applications", McGraw-Hill, (1962).
- Bracewell, R. - "The Fourier transform and its applications", McGraw-Hill, (1965).
- Hecht, E. y Zajac, A. - "Optica", Fondo Educativo Interamericano S.A., (1977).
- Brigham, O. - "The fast Fourier transform", Prentice-Hall, (1974).
- Goodman, J.W. - "Introduction to Fourier optics", McGraw-Hill, (1968).

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.

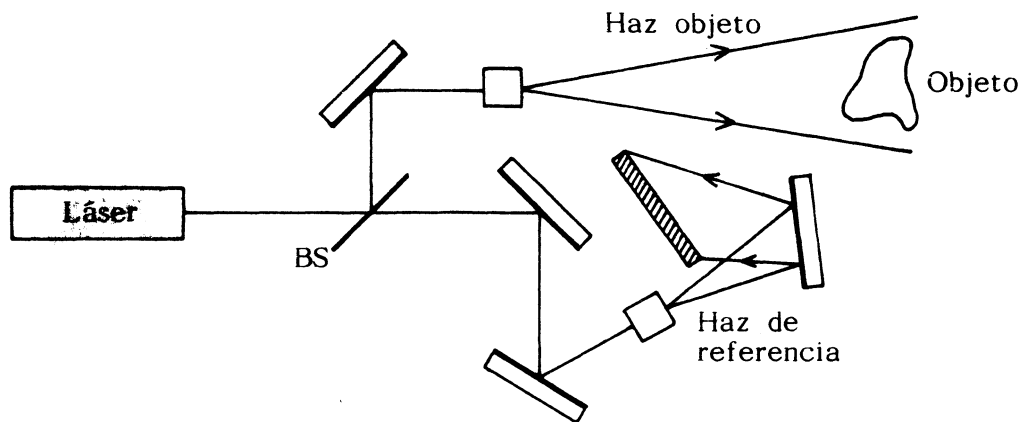


Figura 1

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.

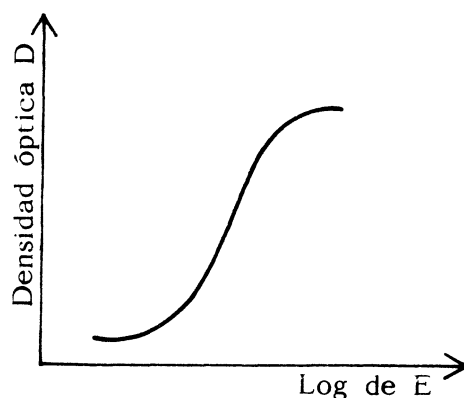


Figura 2

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

$$t \propto E^{-\gamma/2} \quad \text{ó} \quad t \propto I^{-\gamma/2} \propto |U|^{-\gamma},$$

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 .

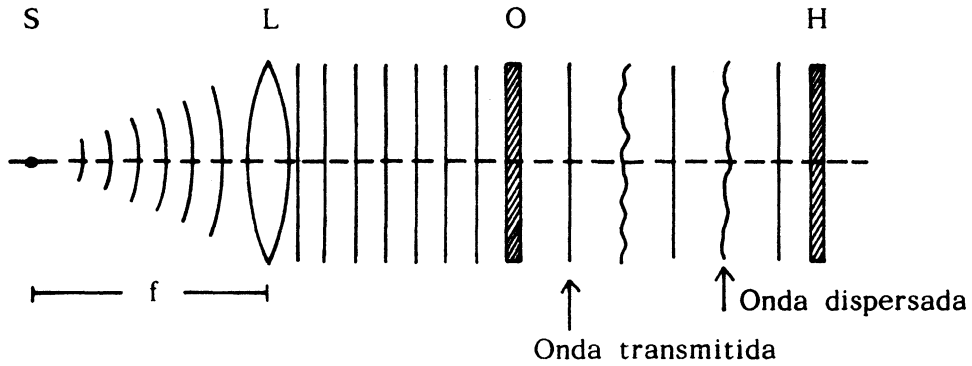


Figura 3

La placa revelada 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.

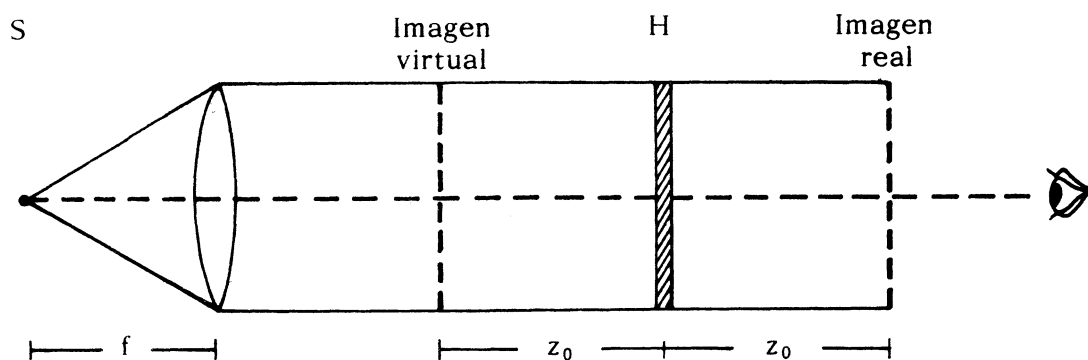


Figura 4

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 Bt_0 . 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 $\alpha = \sin \theta / \lambda$. El primer término denota a la onda plana inclinada y el segundo la luz transmitida por el objeto. La intensidad es

$$I(x,y) = A^2 + |a(x,y)|^2 + Aa(x,y)\exp(2\pi i \alpha y) + Aa^*(x,y)\exp(-2\pi i \alpha y) .$$

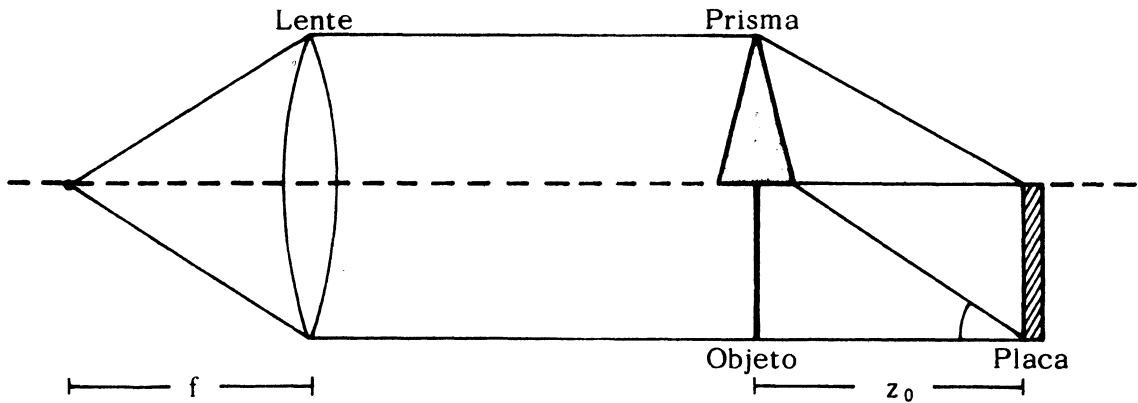


Figura 5

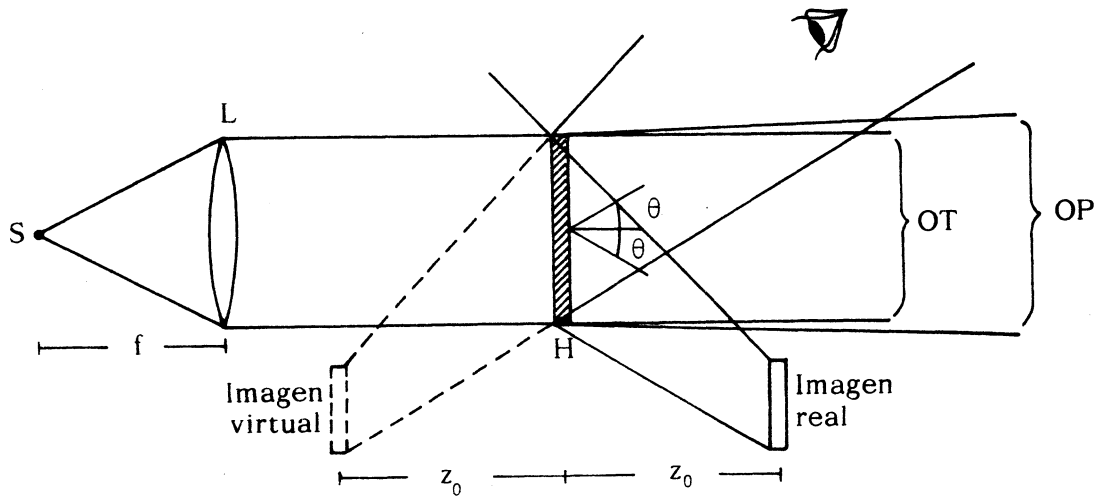
De la manera usual supondremos que el holograma revelado es iluminado por una onda plana uniforme de amplitud B , componiéndose el campo transmitido por cuatro términos, identificados como

$$U_1 = B t_b \quad , \quad U_3 = \beta B A a(x,y) \exp(2\pi i \alpha y) \quad ,$$

$$U_2 = \beta B |a(x,y)|^2, \quad U_4 = \beta B A a^*(x,y) \exp(-2\pi i \alpha y).$$

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.



OT: Onda transmitida. OP: Onda proporcional a $|Q|^2$

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 el 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

$$U_{ir} \propto A + A \exp(ik\gamma z_1 + ik\gamma z_2 \cos wt) ,$$

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 aquéllos perpendiculares a la superficie frontal de la caja. Por ello consideremos la geometría de la Figura 7, la más apta para el registro.

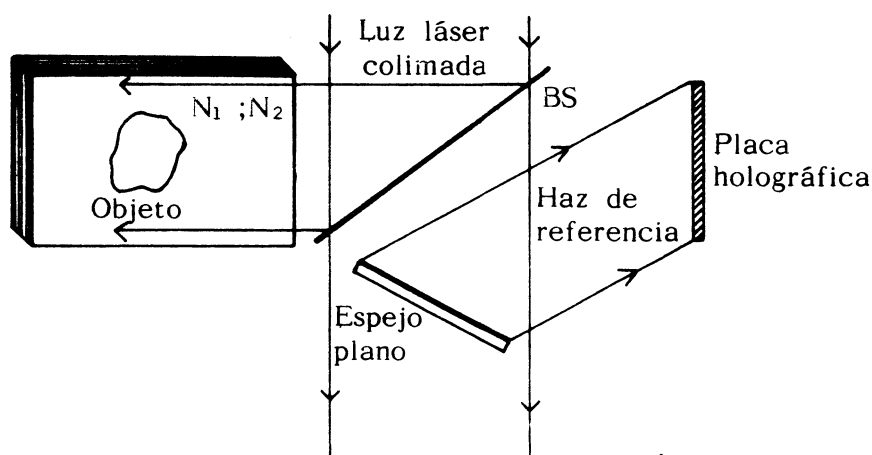


Figura 7

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 descripto 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 despla-

mientos en superficies se pueden usar cuantitativamente para medir esfuerzos. Un esfuerzo mecánico ε , resultante de aplicar una tensión, se define como el incremento fraccional, en dimensiones lineales del cuerpo tensionado, medido en una dirección particular,

$$\varepsilon = \Delta L/L .$$

Por supuesto que existe un diagrama de esfuerzos tridimensional repartido en el volumen del cuerpo, pero por supuesto un método óptico puede dar información sólo de la superficie si el cuerpo es opaco. El campo de esfuerzos se resuelve en dos direcciones ortogonales sobre la superficie. En realidad, se mide el desplazamiento ΔL en función de la longitud de onda. Se debe notar que un interferómetro convencional, tal como el Michelson, puede medir el desplazamiento de uno de sus espejos sólo en la dirección normal a la superficie del mismo. En un holograma interferométrico la situación se modifica, por el hecho que la superficie rugosa del objeto dispersa la luz en todas direcciones.

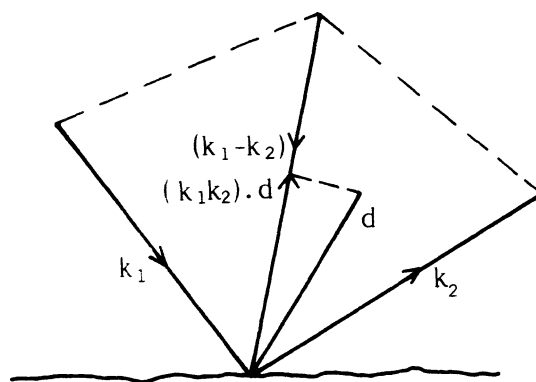


Figura 8

Si $\vec{k}_1$ es el vector unitario en la dirección de iluminación, $\vec{k}_2$ el correspondiente a la dirección de observación y $\vec{d}$ el vector desplazamiento de la superficie, el diagrama de franjas observado da sólo información sobre el desplazamiento δ , donde

$$\delta = (\vec{k}_1 - \vec{k}_2) \cdot \vec{d} .$$

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 pandeo 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 pandeo. Graficando se puede obtener la derivada segunda de z . Si se conoce la distancia z_0 de la superficie al eje neutro 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 pandeo, 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 $2 \times 7/8$ 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 despreciar 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, Nº3, 1982.
- "Incoherent Optical Techniques in Experimental Mechanics", Optical Engineering, Vol. 21, Nº4, 1982.
- "Optical Metrology", Optical Engineering, Vol. 18, Nº5, 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.

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.

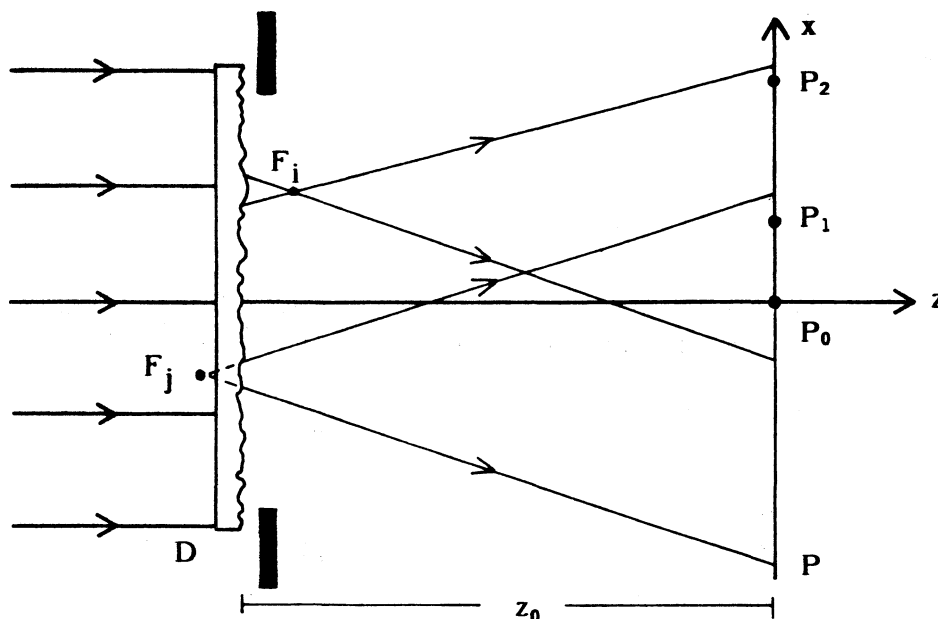


Figura 1

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 $\vec{\xi}(\xi, \eta)$. En estas condiciones, el diagrama de speckle imagen se trasladará: $\vec{d} = m\vec{s}$, 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.

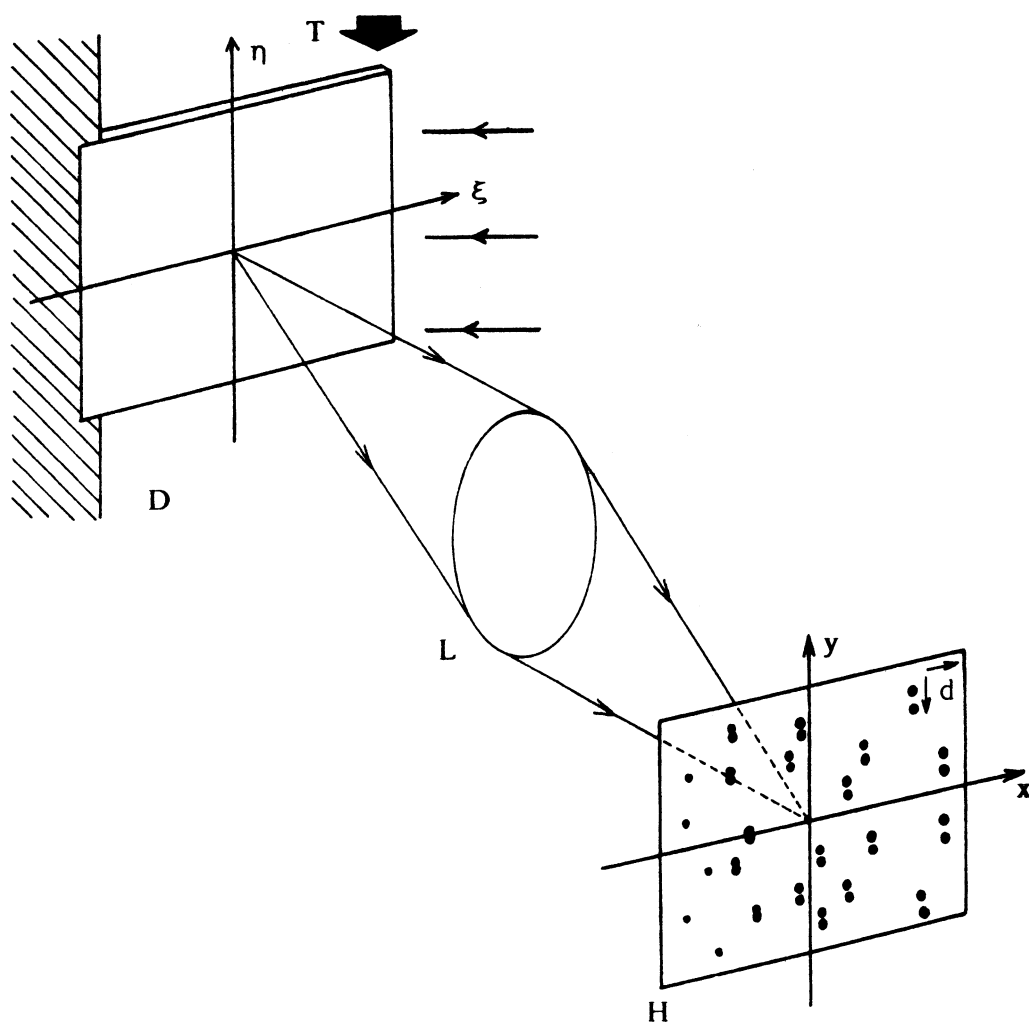


Figura 2

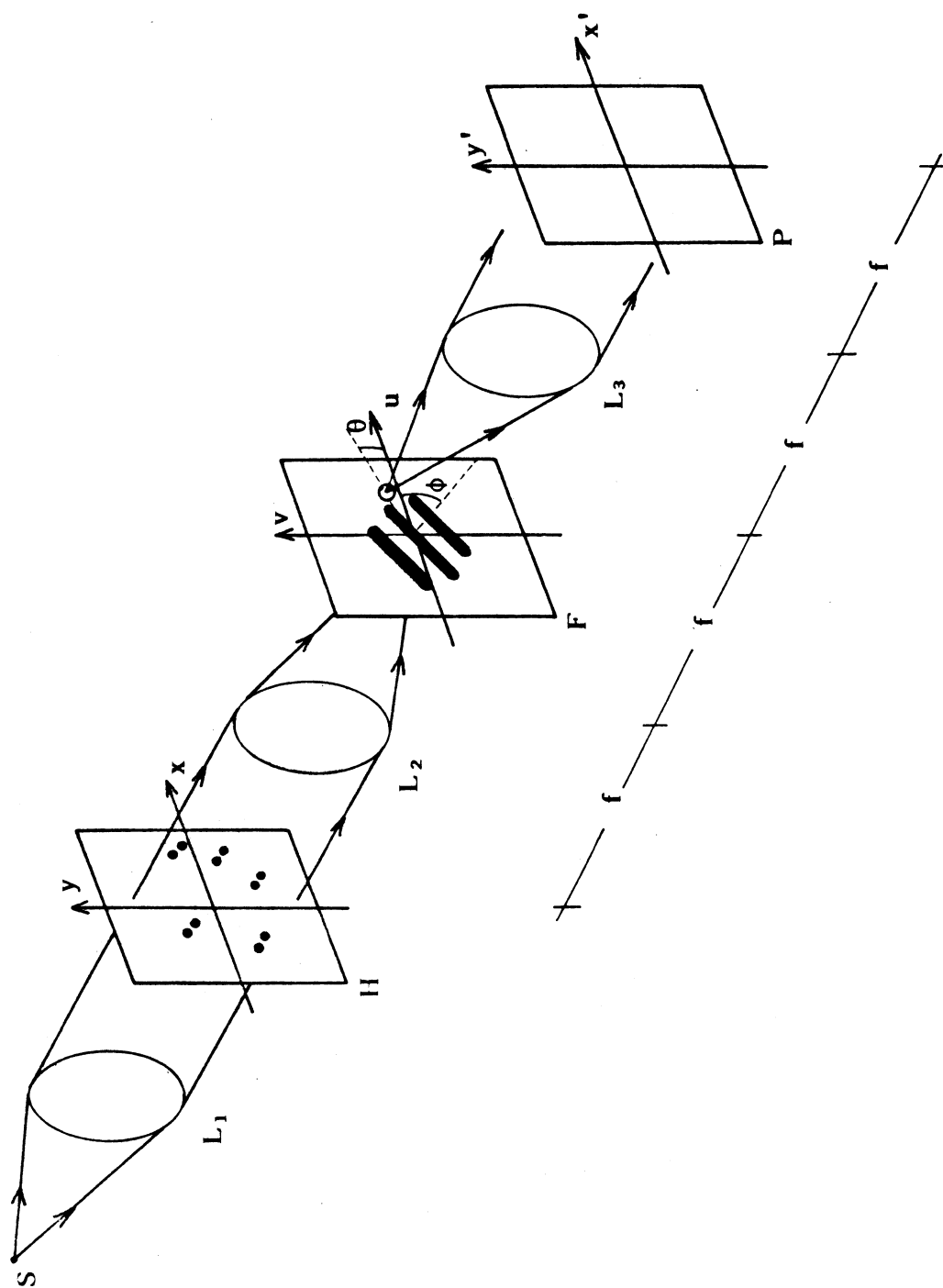


Figura 3

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:

$$u_0 \cos \phi = n\Delta \quad ; \quad n = 1,2,3,\dots \quad (3)$$

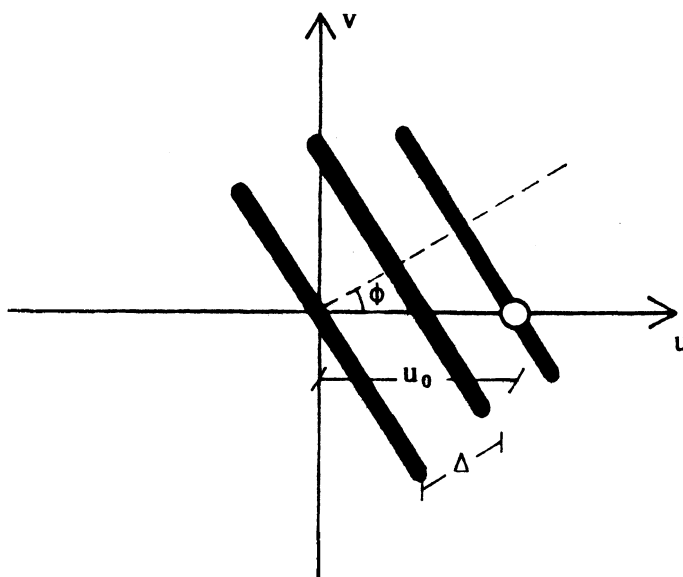


Figura 4

Teniendo en cuenta que $s_x = s \cos \phi$, la Ec.(3) se puede reescribir como:

$$s_x = n \lambda f / \mu u_0 \quad . \quad (4)$$

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:

$$s_q = n \lambda f / m q_0 \quad .$$

Por lo tanto, entre dos franjas adyacentes cualesquiera de la imagen, existirá un incremento del valor de la deformación:

$$\Delta s_q = \lambda f / m q_0 \quad .$$

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:

$$|\vec{s}(x_0/m, y_0/m)| = \lambda z / m \Delta(x_0, y_0) \quad (6a)$$

$$\hat{s}(x_0/m, y_0/m) \cdot \hat{x} = \cos[\phi(x_0, y_0)] \quad ; \quad (6b)$$

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.

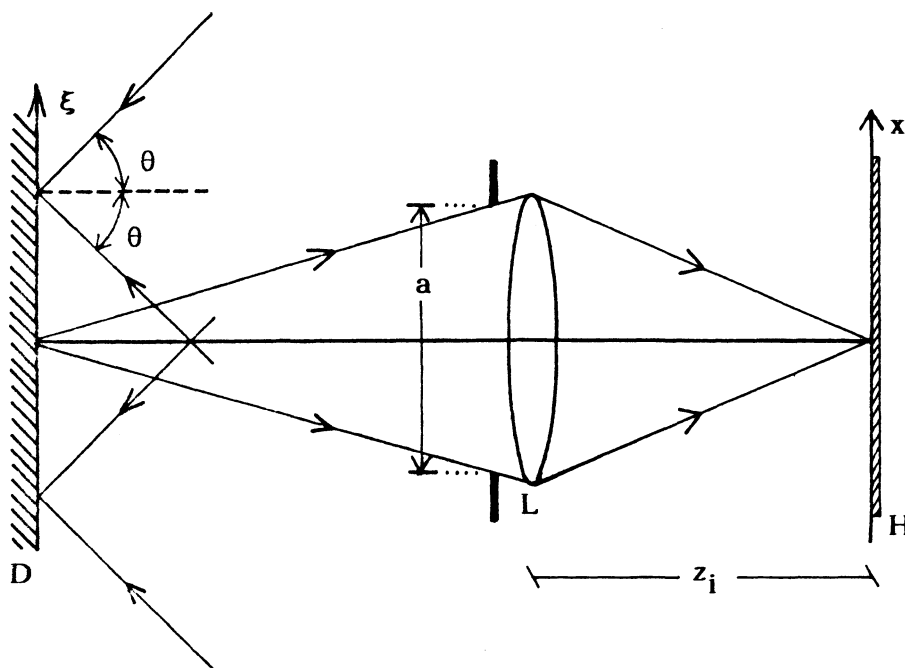


Figura 5

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).

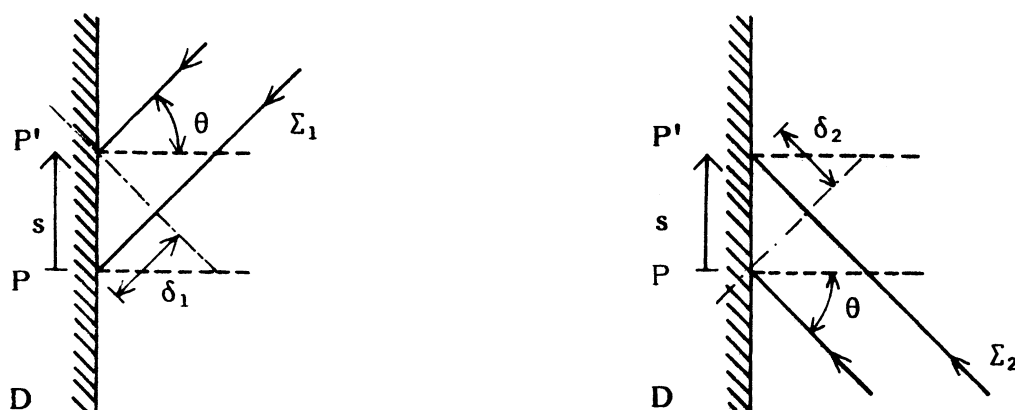


Figura 6

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_i 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: $ms < \lambda z_i/a$, o sea, que la magnitud del desplazamiento del punto imagen de P' respecto a P_i 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: $\phi_1 = -(2\pi/\lambda)s \sin \theta$, mientras que para los rayos debidos a Σ_2 el cambio de fase será: $\phi_2 = (2\pi/\lambda)s \sin \theta$. Por lo tanto, el cambio de fase relativo será:

$$\Delta\phi(\xi, \eta) = (4\pi/\lambda) \sin \theta \cdot s(\xi, \eta) \quad ; \quad (7)$$

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:

$$s = n \lambda/2 \sin \theta \quad . \quad (8)$$

De la Ec.(8) se deduce que el incremento de la deformación superficial entre dos franjas adyacentes es $\lambda/2 \sin \theta$, 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.

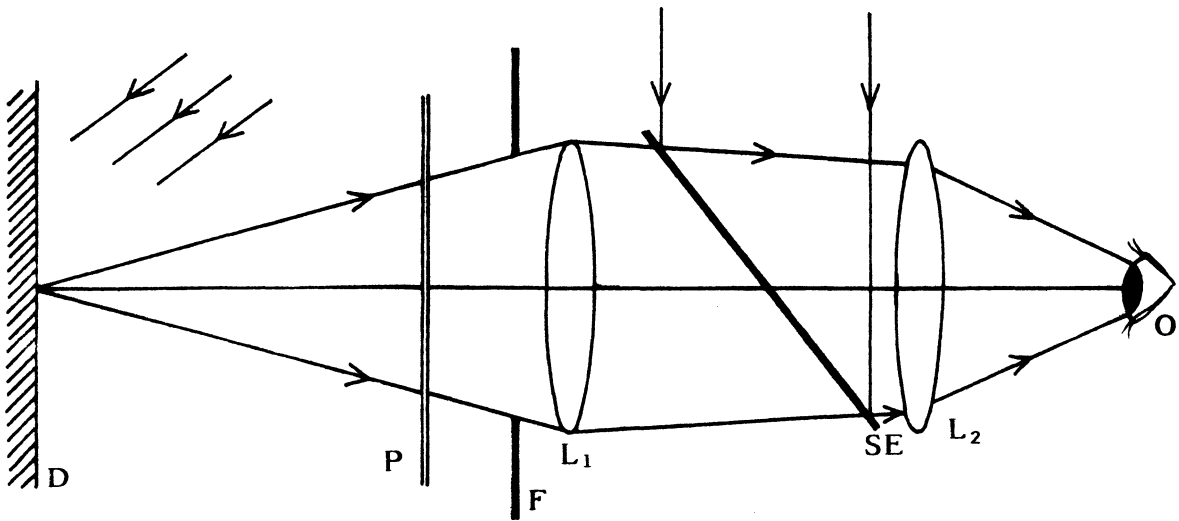


Figura 7

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.

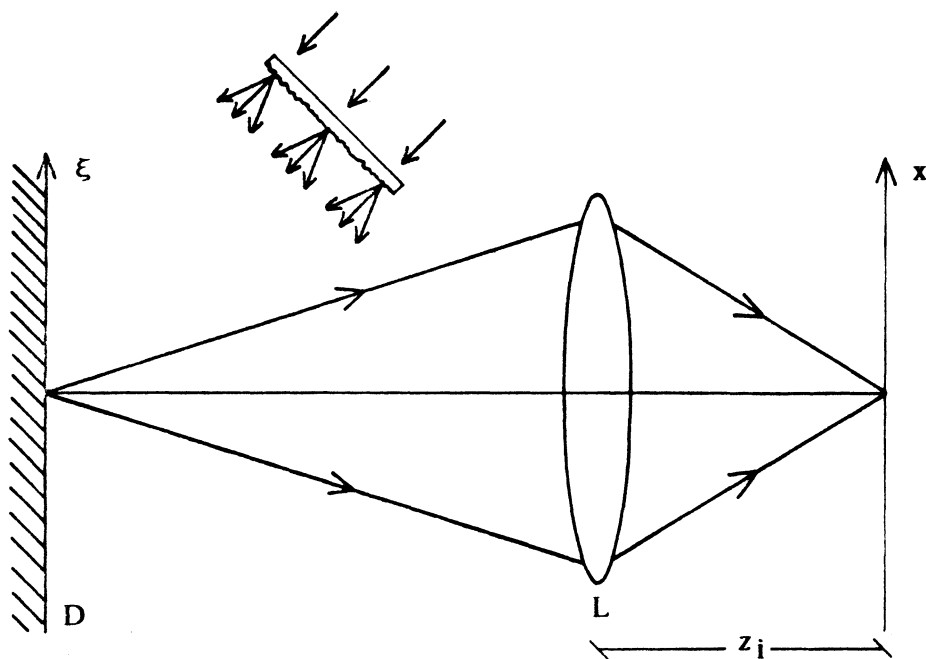


Figura 8.

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.

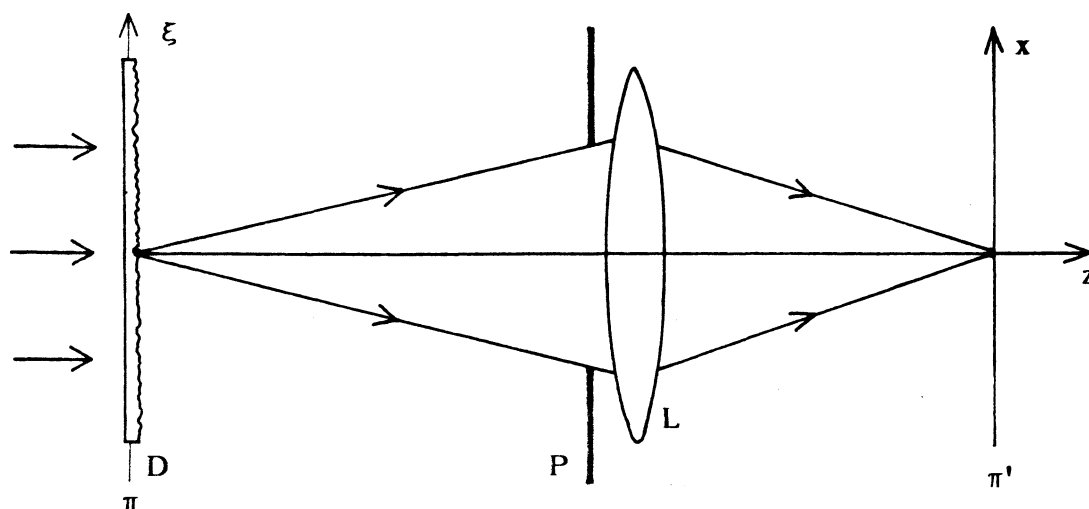


Figura 9

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:

$$U_i(x,y) = \iint_{-\infty}^{\infty} K(x,y;\xi,\eta) U_0(\xi,\eta) d\xi d\eta \quad , \quad (10)$$

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.

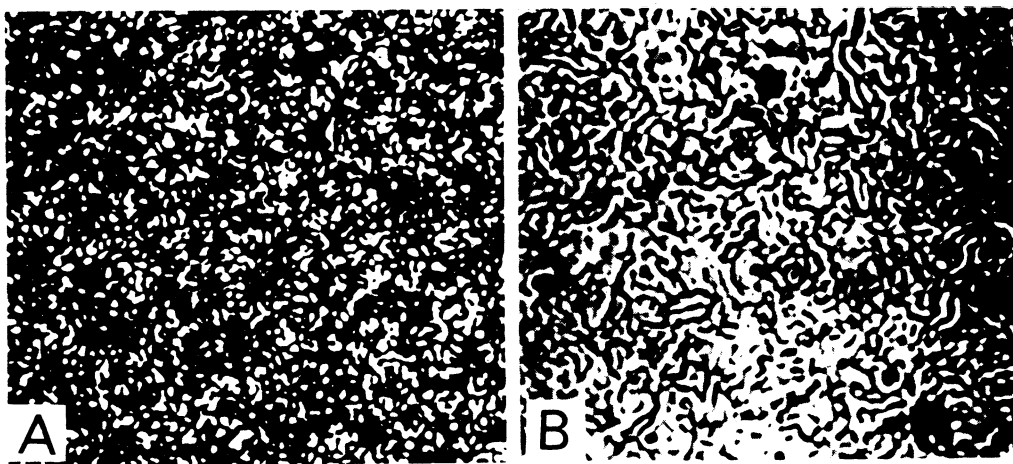
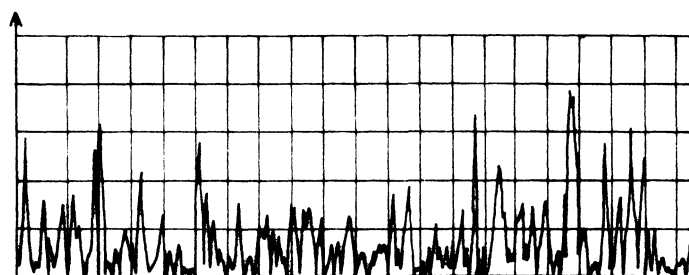
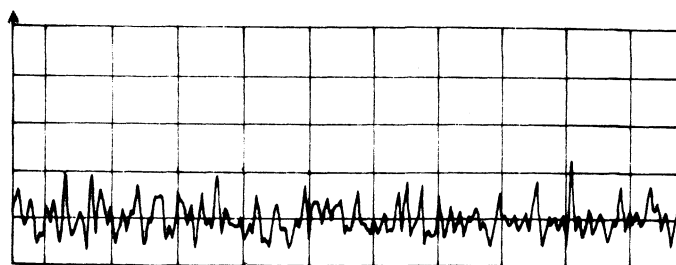


Figura 10



(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 $R_a = (\langle \Delta^2(\xi,\eta) \rangle)^{1/2}$, se obtiene una relación funcional: $C = C(R_a)$, 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:

$$\Delta\theta_2 = (\cos \theta_1 / \cos \theta_2) \Delta\theta_1 \quad . \quad (14)$$

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.

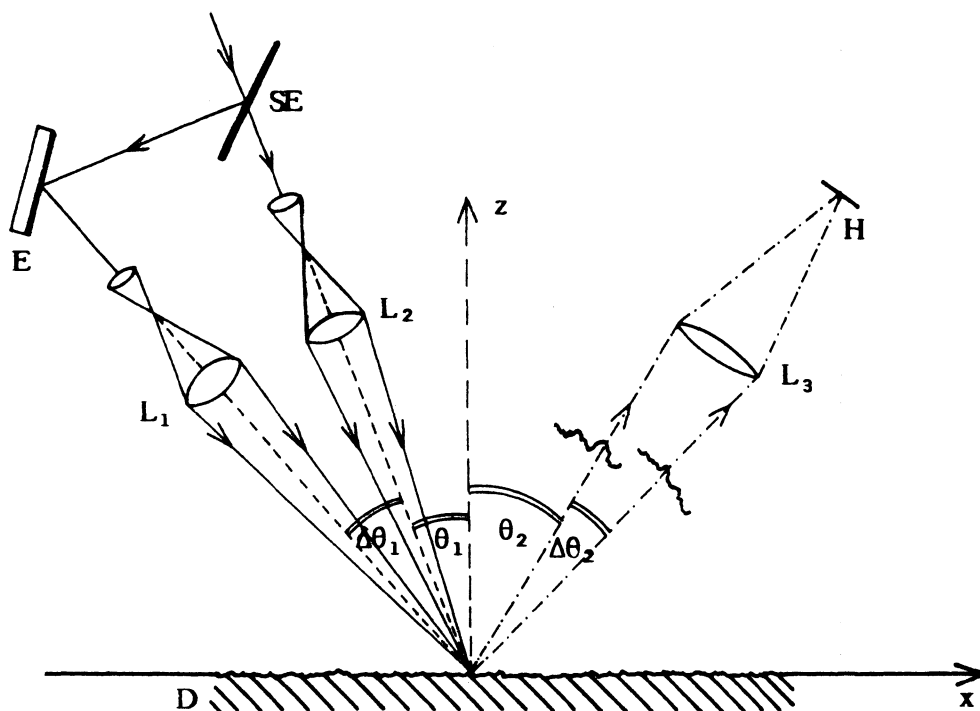


Figura 12

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 Fourier. 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:

$$C = e^{-(2\pi R_a \sin \theta_1 \Delta\theta_1 / \lambda)^2} ; \quad (15)$$

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).

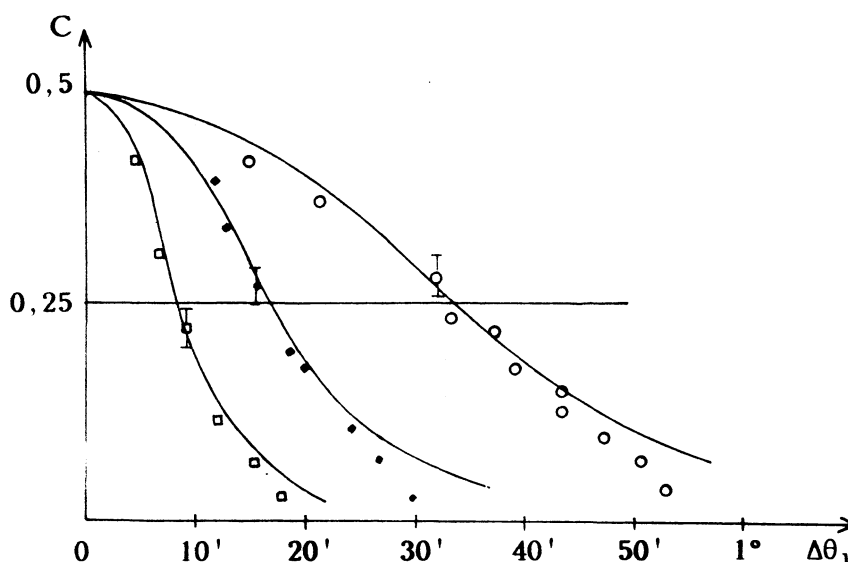


Figura 13

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.

CRISTALES FOTORREFRACTIVOS BSO

N. Bolognini

Introducción

En el año 1966 se encontró que al enfocar un haz láser (longitud de onda $\lambda = 514 \text{ nm}$) en un cristal de LiNbO_3 se provocaba un cambio en el índice de refracción del mismo. Esta inhomogeneidad producía una distorsión en el frente de onda transmitido y perturbaba la colimación del haz, entorpeciendo experiencias de óptica no-lineal en ese rango espectral. Para dar énfasis a estos aspectos negativos, se denominó a este fenómeno, "láser damage effect". Más tarde se encontró que el cristal podía retornar a su estado homogéneo original y que el fenómeno distaba mucho de ser irreversible. Por analogía con el efecto fotocromico que describe cambios en la absorción inducidos por luz, se lo denominó efecto fotorrefractivo, ya que en este caso, se producen cambios en el índice de refracción, inducidos por la luz.

El efecto fotorrefractivo ya ha sido observado en muchos materiales electro-ópticos y puede ser considerado como una propiedad general de los mismos. Entre dichos materiales, uno de los más promisorios es el monocristal de óxido de bismuto y silicio, $\text{Bi}_{12}\text{SiO}_{20}$ (BSO), por su alta sensibilidad para registrar en volumen estructuras de fase difractantes. En particular, en la configuración denominada electro-óptica transversal, permite alcanzar una resolución espacial sólo limitada por difracción. Este medio de registro fotosensible en el rango $\mu\text{W}/\text{cm}^2$ - mW/cm^2 , no requiere procesamiento de ningún tipo para acceder (leer) a la información que en él ha sido registrada pocos milisegundos antes. Así, en la práctica, funciona a tiempo real. Además estos monocristales, disponibles en dimensiones lineales de hasta 3 cm, pueden ser reciclados sin ningún tipo de fatiga. Es decir, el ciclo: registro, lectura, borrado, puede repetirse sin límites. Así puede afirmarse que los cristales fotorrefractivos son medios muy adecuados para emplearlos en técnicas de análisis no destructivo en aplicaciones industriales, ya que se adaptan para experiencias de generación de un frente de onda conjugado, procesamiento óptico de imágenes, e interferometría holográfica, con sus diversas variantes.

Efecto fotorrefractivo en el cristal BSO. Transferencia de carga bajo iluminación

El efecto de cambiar el índice de refracción de un cristal electro-óptico está ligado a la presencia en el material de centros fotoexcitables. Bajo iluminación local $I(x)$ (x coordenada espacial) se producen en el cristal fotoportadores que han sido generados a partir de la liberación de electrones de centros donores, los cuales se desplazan una distancia media L antes de ser retenidos en un nivel aceptor. La densidad de carga así originada produce un campo de carga espacial que modula por efecto electro-óptico lineal, la birrefringencia del material. El mecanismo de transferencia de carga se muestra en la Figura 1.

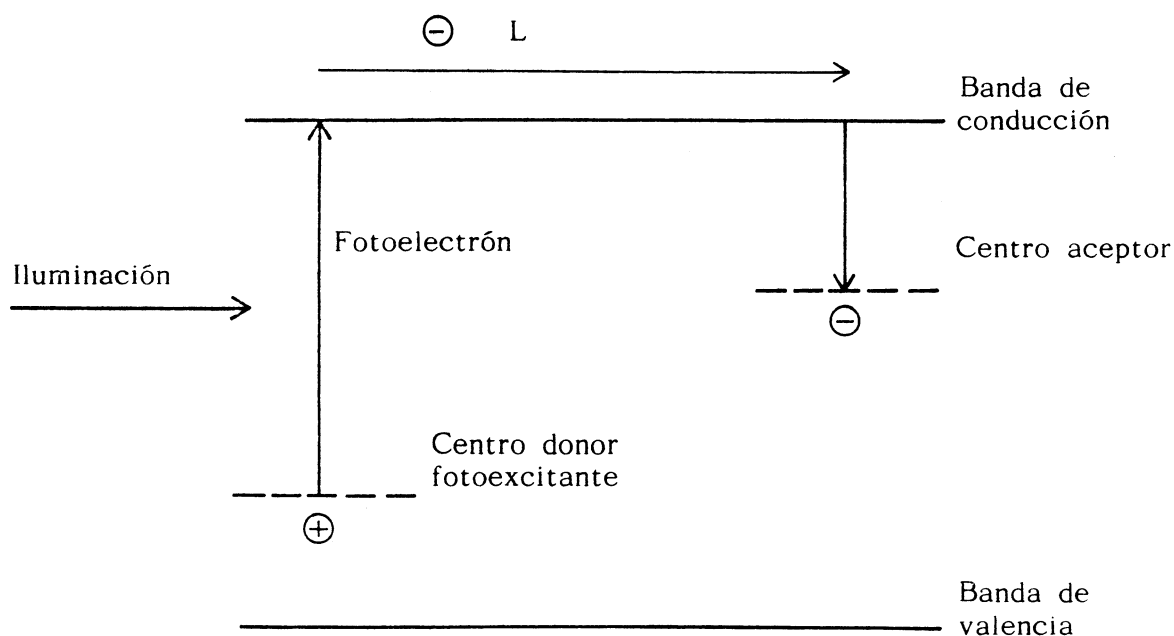


Figura 1

Los procesos físicos en juego en el cristal electro-óptico para llevar a cabo el desplazamiento de cargas fotoexcitadas, tiene lugar por la acción de dos mecanismos:

- Difusión de portadores sobre una distancia media L_d , bajo iluminación no uniforme.

- "Drift" de portadores bajo la acción de un campo eléctrico exterior sobre una distancia media L_e . En un cristal ferroeléctrico como LiNbO_3 la asimetría de la red cristalina alcanza para orientar la migración de elec-

trones, pero en cristales para-eléctricos como BSO es conveniente "orientar" los electrones mediante un campo eléctrico externo.

Según las condiciones de trabajo, uno u otro mecanismo se tornará preponderante para el transporte de carga. Los valores medios de L_d y L_e resultan:

$$L_d = C_d(\mu\tau)^{1/2} \quad ; \quad L_e = \mu\tau E_0 .$$

μ : movilidad de los fotoportadores.

τ : tiempo de tránsito entre un sitio donador y uno aceptor.

E_0 : campo eléctrico aplicado al cristal.

C_d : constante que depende de la temperatura.

Los valores típicos de L_d y L_e son del orden de 1 μm .

Registro de un campo de carga espacial

Se ilumina el cristal electro-óptico con una distribución luminosa generada por la superposición de dos ondas planas coherentes en el volumen del cristal, según muestra la Figura 2.

Resulta así una distribución espacial de intensidad luminosa de la forma:

$$I(x) = I_0(1+m \cos Kx) \quad ; \quad m = \frac{2(I_1 I_2)^{1/2}}{I_1 + I_2} \quad ; \quad K = \frac{2\pi}{\Lambda}$$

donde I_1 , I_2 son las intensidades de los haces que interfieren; m es la modulación; Λ es el espaciado de las franjas de interferencia y vale $\lambda/2 \sin(\theta/2)$ siendo θ el ángulo entre los haces y λ la longitud de onda de la luz.

La iluminación $I(x)$ en el volumen del cristal genera fotoportadores que se desplazan por difusión o "drift" desde las zonas de máxima iluminación hasta las de mínima, donde son retenidos en los sitios aceptores. Este desplazamiento conduce a un nuevo equilibrio en la distribución de carga que origina el campo de carga espacial E_{sc} , el cual modula por efecto electro-óptico lineal el índice de refracción del cristal y origina una estructura de fase difractante en volumen. Esta última no está necesariamente en fase con el sistema de franjas de interferencia. Las etapas del proceso descrito están representadas en la Figura 3.

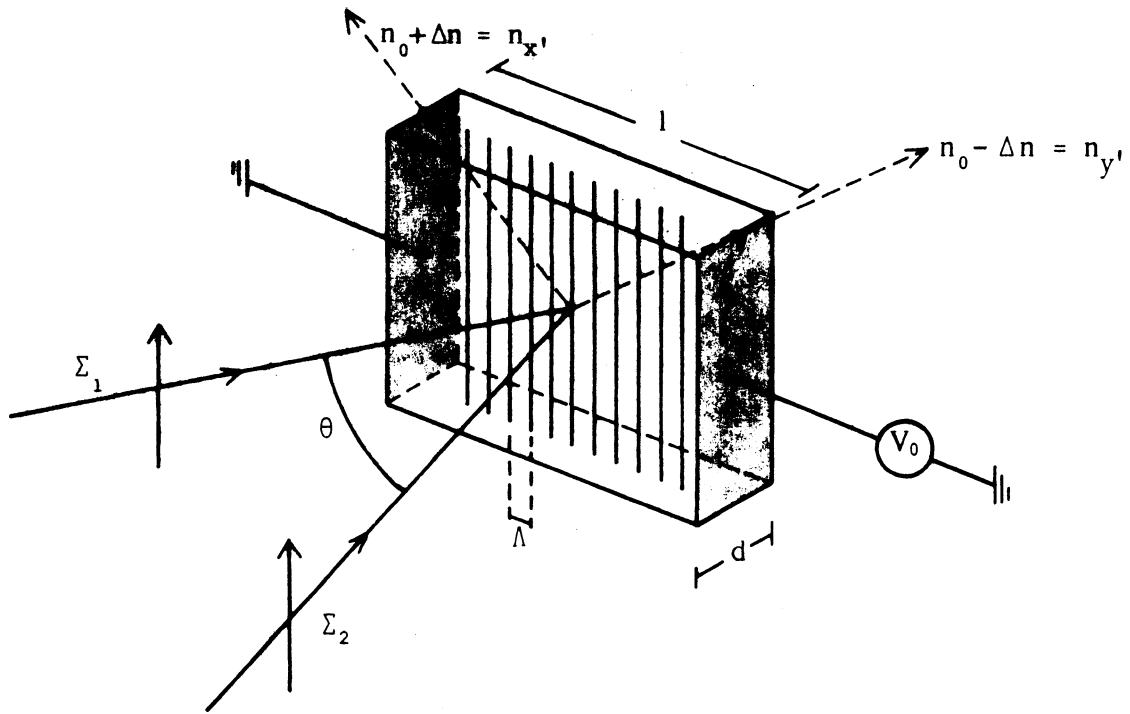


Figura 2

La iluminación $I(x)$ genera una densidad de fotoportadores por unidad de volumen, $N(x,t)$, proporcional a $I(x)$. Esta situación produce una densidad de corriente $J(x,t)$ dada por

$$J(x,t) = C_1 \frac{dN(x,t)}{dx} + C_2 N(x,t) [E_0 + E_{sc}(x,t)] .$$

C_1 y C_2 son constantes y E_0 es el campo externo aplicado.

Teniendo en cuenta la densidad de corriente que se genera en el cristal, $J(x,t)$, y a partir de la ecuación de continuidad, resulta:

$$\rho(t) = - \int_0^t \frac{dJ}{dx} dt' ;$$

t : tiempo,

$\rho(t)$: densidad de carga inducida.

la que a su vez permite calcular el campo de carga espacial:

$$E_{sc} = \frac{1}{\epsilon} \int_0^x \rho(x') dx' ;$$

No existe una expresión analítica simple de $E_{sc}(x,t)$. Sin embargo, sí se encuentra una expresión para su valor de saturación, el cual se alcanza en pocos milisegundos, dada por:

$$E_{sc} = m \left[\frac{E_0^2 + E_d^2}{E_0^2 + (E_d + E_m)^2} \right]^{1/2} E_m \cos(Kx + \psi) \quad (1)$$

$$\operatorname{tg} \psi = \frac{E_0^2 + E_d^2 + E_m E_d}{E_m E_0} ;$$

$$E_d = C_d f_x$$

$$E_m = C_m f_x^{-1}$$

$$f_x = K_x$$

donde E_m es la amplitud máxima que se podría fotoinducir en el cristal para una frecuencia espacial f_x dada, y corresponde a una separación de todas las cargas (donores y aceptores) disponibles; E_d es la amplitud de campo de carga espacial inducido por difusión; E_0 es el campo eléctrico externo aplicado; ψ representa el desfase respecto al sistema de franjas de interferencia.

Un valor típico para E_0 resulta ser alrededor de 5 kV/cm. En este caso $E_m \gg E_0$ y además $E_0 \gg E_d$ para frecuencias espaciales de hasta 2000 1n/mm. Así, la expresión para E_{sc} se reduce a:

$$E_{sc} = m C E_0 \cos Kx . \quad (2)$$

C : constante limitadora de la amplitud de E_{sc} , ligada a la densidad de centros aceptores.

El cristal naturalmente no es birrefringente. La presencia de E_{sc} induce birrefringencia. Es decir, aparecen dos modos de propagación de la luz en el cristal con distintos índices de refracción (ver Figura 2) dados por:

$$\begin{aligned} n_{x'} &= n_0 + \Delta n \\ n_{y'} &= n_0 - \Delta n \end{aligned} \quad (3)$$

donde n_0 representa el valor del índice de refracción cuando E_{sc} no ha sido inducido; Δn representa la variación inducida por la existencia de E_{sc} y puede calcularse a partir de las ecuaciones de Maxwell con el método del elipsoide de índices. Teniendo en cuenta la precisión de las medidas experimentales, la dependencia lineal de Δn con E_{sc} es suficiente. A esta aproximación se la denomina efecto Pockels o efecto electro-óptico lineal, y para ella Δn resulta:

$$\Delta n = \frac{1}{2} n_0^3 r_{41} E_{sc} . \quad (4)$$

donde r_{41} es el coeficiente fotorrefractivo del cristal BSO.

La birrefringencia que se ha inducido resulta ser entonces:

$$\begin{aligned} n_{x'} - n_{y'} &= \delta n = n_0^3 r_{41} E_{sc} = n_0^3 r_{41} m C E_0 \cos Kx \\ \delta n &= E_0 C' m \cos Kx . \end{aligned} \quad (5)$$

Así, se ha reproducido en una variación de la birrefringencia la variación de la distribución luminosa $I(x) = I_0(1 + m \cos Kx)$. Puede concluirse que, bajo ciertas condiciones, el efecto fotorrefractivo permitirá reproducir una distribución luminosa arbitraria. Como ya fue mencionado, la Figura 3 describe el fenómeno fotorrefractivo.

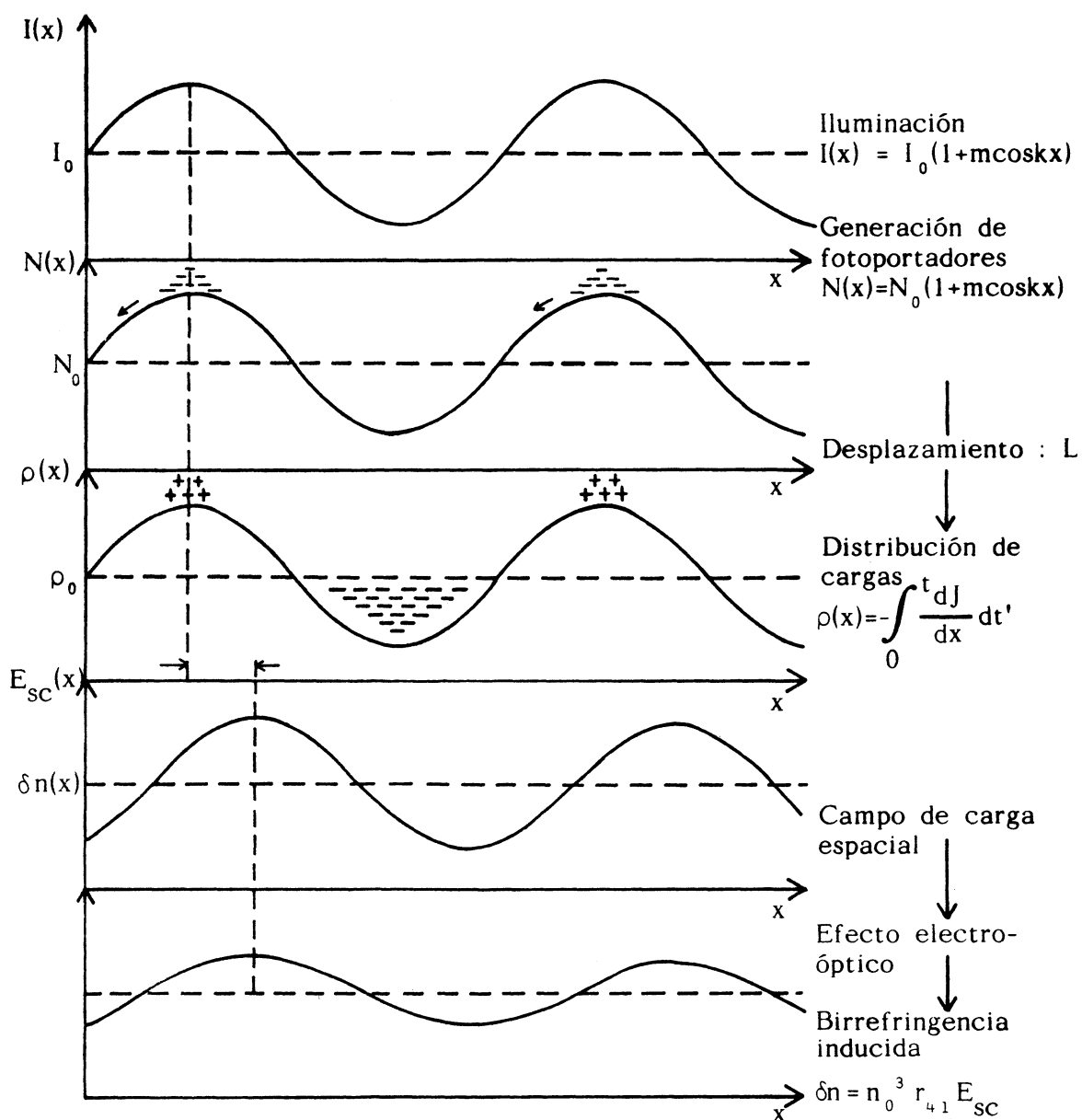


Figura 3

Holografía interferométrica con cristales BSO

El cristal BSO resulta muy adecuado para realizar una configuración óptica que permita observar en tiempo real los modos de estructuras vibrantes mediante técnicas de interferometría.

Como ha sido analizado, una distribución luminosa monocromática arbitraria (por ejemplo, $\lambda=514$ nm, láser de argón), es reproducida en el cristal a través de la variación de la birrefringencia que en él se induce. La lectura de dicho registro con la misma longitud de onda λ de registro tiene el efecto negativo de relajar la distribución espacial de carga inducida, borrando así el registro almacenado en el cristal. Esto puede evitarse si el haz de lectura tiene una longitud de onda para la cual el cristal es casi insensible.

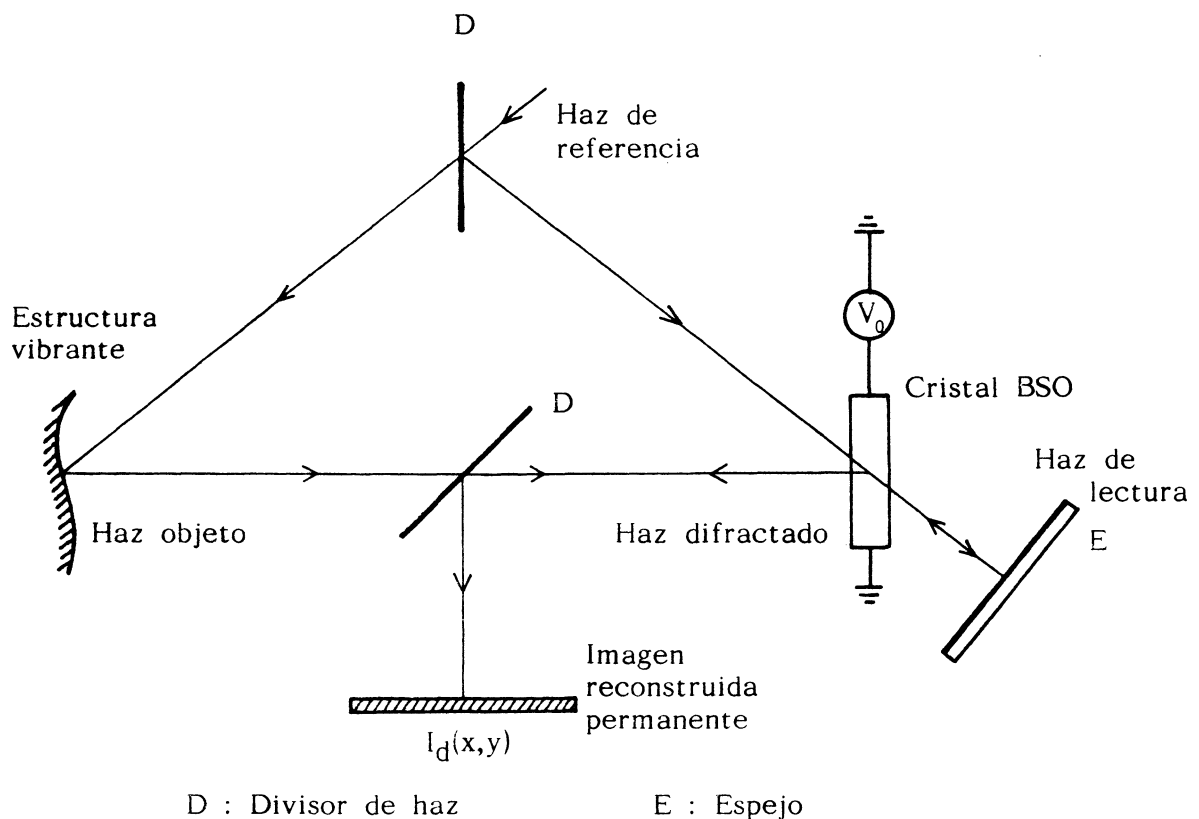


Figura 4

Por otro lado, en un holograma de volumen no es conveniente reconstruir su imagen con una longitud de onda distinta a la de registro. Esta incompatibilidad, sin embargo, puede salvarse, obteniéndose una lectura permanente, si se utiliza el arreglo experimental de la Figura 4.

Puede observarse que la parte transmitida por el cristal de la onda plana de referencia Σ_R , se emplea después de reflejarse en el espejo E para obtener una reconstrucción de la imagen real del holograma.

Otra configuración similar se muestra en la Figura 5, en caso que el objeto sea iluminado por transmisión:

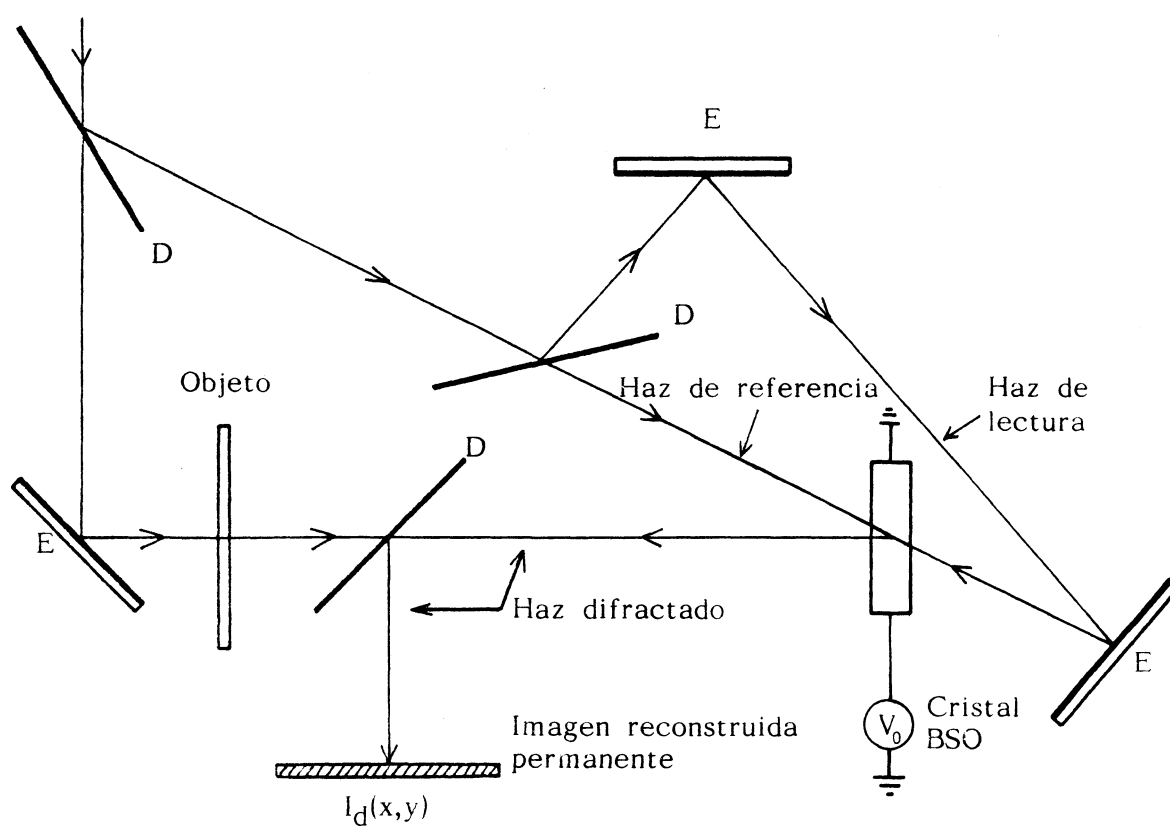


Figura 5

En este caso, la reflexión combinada de los espejos E_1 y E_2 permiten también una lectura permanente y obtener una reconstrucción de la imagen real del holograma, función que en la Figura 4 la cumple el espejo E .

La influencia del haz de lectura en la saturación de la eficiencia de difracción puede ser considerada como equivalente a una reducción en la modulación para una dada frecuencia espacial $f_x = K$.

La intensidad total incidente en el cristal para una red holográfica elemental puede ser escrita:

$$I = I_R(1-R)(1+B) + 2I_R(1-R)B^{1/2}\cos Kx + I_L(1-R) \quad (6)$$

I_R e I_L son las intensidades de los haces incidente de referencia y de lectura, respectivamente. B es el cociente entre la intensidad del objeto y del haz de referencia; R es la reflectividad del cristal. Esta ecuación supone que el haz de lectura se suma en forma incoherente. Si bien en la práctica esta suposición puede no ser válida si el láser tiene gran longitud de coherencia, el cristal no resuelve el sistema de franjas que surge de la interferencia del haz de referencia y lectura: estas franjas son paralelas al campo eléctrico externo aplicado y no generan una modulación en el campo de carga espacial E_{sc} .

La ecuación (6) puede reescribirse como

$$I = I_T \left[1 + \frac{2(I_R B)^{1/2}}{I_R(1+B) + I_L} \cos Kx \right], \quad (7)$$

con $I_T = I_R(1-R)(1+B) + I_L(1-R)$.

Según la ecuación (7) la modulación resulta ser:

$$m = \frac{2(I_R B)^{1/2}}{I_R(1+B) + I_L}.$$

El campo de carga espacial fotoinducido saturado a la frecuencia K es proporcional a m [ecuaciones (1) y (2)]. La eficiencia de difracción se define como el cociente de la intensidad incidente I_R y la intensidad difractada en el plano imagen; la ecuación de Kogelnik describe esta situación como

$$Q = \frac{I_1}{I_R} (1-R)^2 \left(\frac{\pi d \Delta n}{\lambda} \right)^2 m^2 \exp(-\alpha d) . \quad (8)$$

donde d es el espesor del cristal; α es la absorción a la longitud de onda λ ; Δn está descrito por la ecuación (4) en la cual se toma la amplitud de E_{sc} .

Según la configuración de la Figura 4:

$$I_{L1} = (1-R)^2 I_R \exp(-\alpha d) .$$

Según la configuración de la Figura 5:

$$I_{L2} = I_R .$$

Reemplazando I_{L1} y I_{L2} en la expresión (8) se obtiene un cociente:

$$Q_1/Q_2 = (1-R)^2 \left[\frac{2 + B}{1+B+(1-R)^2 \exp(-\alpha d)} \right]^2 \exp(-\alpha d) .$$

Condiciones experimentales óptimas requerirían que $B=1$, $\alpha d=1$ y $R=20\%$. En este caso $Q_1/Q_2=0,42$. Podría concluirse que la configuración de la Figura 5 es más eficiente, debido a que en la Figura 4 la onda de referencia atraviesa dos veces el cristal y aumentan las pérdidas por reflexión y absorción. Sin embargo, esta última cancela la actividad óptica que presenta el cristal y es más simple de ajustar. También se debería considerar la planitud de las caras del cristal y su paralelismo. La reflectividad puede reducirse con capas dieléctricas antirreflectantes en las caras del cristal.

Cuando el objeto vibra y es iluminado por reflexión, se obtiene una imagen difractada $I_d(x,y)$ en la cual se superponen franjas de interferencia correspondientes a la función de Bessel de orden cero, es decir

$$I_d = I_0 \left| J_0 \left(\frac{4\pi}{\lambda} \delta \right) \right|^2$$

donde δ es la amplitud local de la deformación periódica. Esto es así cuando el tiempo de registro τ_{sc} del campo de carga espacial es largo comparado con el período de la ^{sc} vibración. Para el cristal BSO, τ_{sc} es del orden de 10mseg para 15 mW/cm². Esto es conveniente para observaciones en tiempo real de estructuras vibrantes a frecuencias del orden del kHz. Frecuencias más bajas pueden examinarse reduciendo la potencia del láser.

En el caso de holografía interferométrica de doble exposición, el arreglo experimental para el análisis de deformaciones o desplazamientos (del orden de la longitud de onda λ de la luz) es semejante a los anteriores. El principio de este control no destructivo consiste en superponer y registrar en el cristal, dos frentes de onda generados por el objeto, antes y después de la deformación o desplazamiento. Así, toda diferencia entre ambos frentes registrados, múltiplo impar de $\lambda/2$, será visualizado por una franja oscura superpuesta a la imagen difractada $I_d(x,y)$.

Las etapas para este proceso son:

- registro a saturación del primer frente de onda (objeto sin deformar) durante un tiempo $T_1 = \tau_{sc}$;
- registro del segundo frente de onda durante $T_2 = \tau_{sc}/2$;
- lectura de la imagen difractada $I_d(x,y)$ modulada ^{sc} por franjas y transferencia del interferograma a un tubo vidicón (por ejemplo) con memoria.

Los dos arreglos experimentales de la Figura 4 y la Figura 5, son adecuados también y de una conveniente flexibilidad con pequeñas variantes para experiencias de procesamiento de imágenes tales como:

- rescatar una señal óptica enmascarada por algún tipo de aberración óptica ;
- mejorar la relación señal-ruido debido al speckle, inherente a cualquier sistema óptico que emplea luz de un láser ;
- resaltar los contornos de un objeto .

cabe remarcar, entonces, que para observaciones de estructuras vibrantes en tiempo real, a través de interferometría holográfica de tiempo promedio, el cristal BSO es un material óptimo, por su alta sensibilidad (comparable a las placas fotográficas) y por no requerir ningún procesamiento ni desplazamiento del cristal para realizar la lectura permanente del interferograma.

Estas características ventajosas frente a otros medios de registro, son válidas también para todas las experiencias descritas. Así, estas técnicas podrían constituirse en una herramienta muy valiosa en el área del mantenimiento predictivo, es decir, obtener información de las máquinas sin detenerlas y llegar a prevenir su desgaste sin desarmarlas.

Correlación óptica dinámica

Los procesadores ópticos coherentes tienen una gran capacidad de procesamiento en paralelo. Sin embargo, sería muy interesante contar con un medio holográfico que permitiera adaptar el filtro adecuado en el plano de Fourier del procesador en forma dinámica, es decir, poder modificarlo todas las veces que sea necesario, en tiempo real. Este requerimiento puede ser satisfecho con un cristal BSO. Esto puede ejemplificarse en la obtención de la correlación óptica a través de una configuración que utiliza la interferencia de dos frentes de ondas que han sido modulados por transmisión por dos señales ópticas. El arreglo experimental se muestra en la Figura 6.

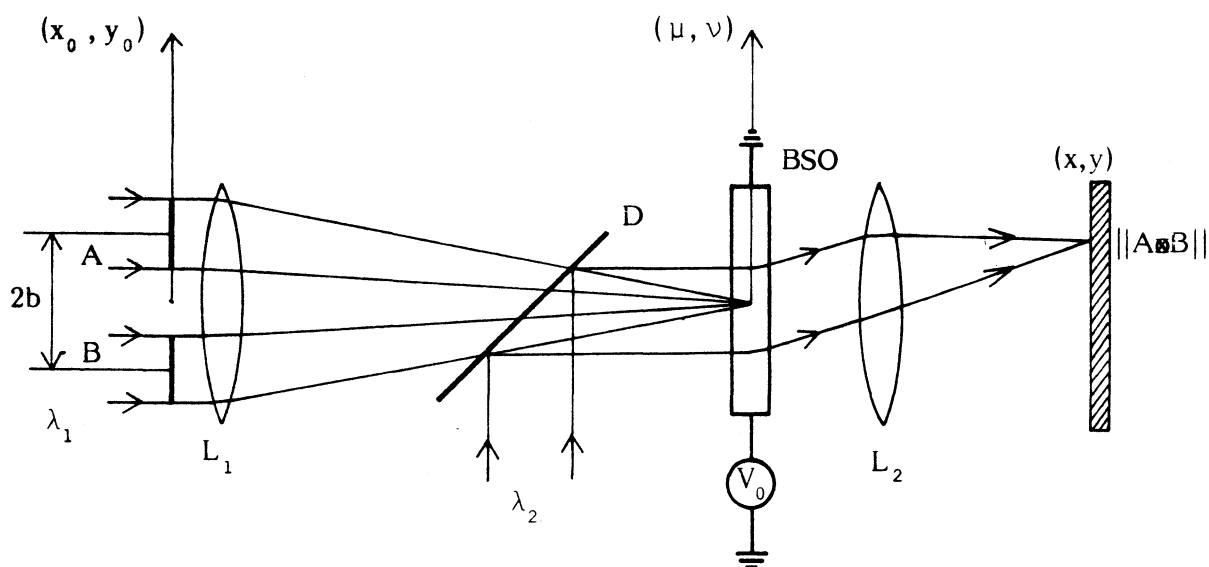


Figura 6

Las dos transparencias a ser procesadas A y B, se colocan en la pupila de la lente L_1 . De acuerdo a las propiedades de transformada de Fourier de una lente, la amplitud compleja en el plano focal de L_1 es proporcional a:

$$F\{A(x_0 - b, y_0) + B(x_0 + b, y_0)\} ;$$

F : transformada de Fourier ;

donde $\pm b$ representa la distancia entre A y B respecto al eje óptico del sistema. La correlación dinámica entre A y B se logra ubicando el cristal BSO en el plano focal de la lente transformadora L_1 . Así, la distribución luminosa en dicho plano origina una variación en el índice de refracción fotoinducido, la cual es proporcional a:

$$I(u,v) = |F\{A(x_0-b, y_0) + B(x_0+b, y_0)\}|^2 ,$$

donde $u = x/\lambda_1 f_1$, $v = y/\lambda_1 f_1$; f_1 : distancia focal de L_1 ; λ_1 longitud de onda de registro (por ejemplo $\lambda_1 = 488$ nm, línea láser de argón).

En este caso la lectura se realizará con una longitud de onda distinta para evitar efectos de disminución de la saturación en la eficiencia de difracción y obtener una buena relación señal/ruido. Entonces, al leer el holograma de volumen y de fase (variación en el índice de refracción del cristal inducida por $I(u,v)$) mediante una longitud de onda λ_2 (por ejemplo $\lambda_2 = 633$ nm), se induce en el cristal una transmitancia compleja dada por

$$t(u,v) = \exp[iCI(u,v)] .$$

Dado que la modulación de fase fotoinducida es pequeña (Δn en la ecuación (4) es del orden de 10^{-4}), $CI(u,v) \ll \pi/2$ y la transmisión del cristal se puede escribir:

$$t(u,v) = 1 + iCI(u,v) .$$

En consecuencia, la amplitud difractada en el plano focal de la lente L_2 de distancia focal f_2 es proporcional a:

$$F\{CI(u,v)\} .$$

Así, la distribución de intensidad en este plano puede ser escrita como

$$I(x,y) = |A \otimes A + B \otimes B|^2 + |A \otimes B|^2 * \delta(x+2b \frac{\lambda_2 f_2}{\lambda_1 f_1}, y) \\ + |B \otimes A|^2 * \delta(x-2b \frac{\lambda_2 f_2}{\lambda_1 f_1}, y) \quad .$$

donde $\otimes$ indica la operación de producto de correlación cruzada entre A y B y $*$ indica producto de convolución.

Entonces, si A se utiliza como señal de referencia, por ejemplo representando un signo cualquiera, cada vez que dicho signo aparece en la señal B , es detectado vía el producto de correlación en el plano focal de L_2 como un punto luminoso intenso, en la posición relativa que ocupa en B . La convolución con una función δ (delta de Dirac) indica que los picos de correlación entre la señal y la referencia se encuentran desplazados respecto al eje óptico en el factor geométrico $\pm 2b \lambda_2 f_2 / \lambda_1 f_1$. Así, los picos de correlación observados en el plano focal de L_2 revelan la presencia y posición relativa de las partes comunes de B con respecto a la referencia A .

Obsérvese que si no se contara con el cristal en el plano focal de L_1 , plano de filtrado espacial, cada vez que se quisiera comparar B con una referencia distinta A , sería necesario sintetizar en una placa fotográfica de alta resolución un holograma de la transformada de Fourier de A (filtro adaptado), procesarla y reinsertarla en el plano focal de L_1 .

Cabe mencionar que al ser un holograma de volumen el que ha sido registrado, el haz de lectura (λ_2) debe ser ajustado cuidadosamente para que los picos de correlación sean difractados en su condición óptima.

Bibliografía

- A.M. Glass, Opt. Eng., 17, 470 (1978) .
- P. Günter, Phys. Report, 93, 199 (1982) .
- N.V. Kukhtarev, V.B. Markov, S.G. Odulov, M.S. Soskin, V.L. Vinetskii, Ferroelectric 22, 949 (1979) .
- A. Marrakchi, J.P. Huignard, J.P. Herriau, Opt. Comm. 34, 15 (1980) .
- J.P. Huignard, A. Marrakchi, Opt. Comm. 38, 249 (1981) .
- Y.H. Ja, Opt. Comm. 42, 377 (1982) .
- E. Ochoa, L. Hesselink, J.W. Goodman, Appl. Opt. 24, 1826 (1985) .
- R. Rak, I. Ledoux, J.P. Huignard, Opt. Comm. 49, 302 (1984) .
- M.B. Klein, Opt. Comm. 9, 350 (1984) .

- H. Rajbenbach, J.P. Huignard, Opt. Comm. 10, 137 (1985).
- D.Z. Anderson, Opt. Lett., 11, 56 (1986).
- M. Cronin-Golomb, A. Yariv, I. Ury, Appl. Phys. Lett. 48, 1240 (1986).
- D.A. Temple, C. Warde, J. Opt. Soc. Am., B.3, 337 (1986).
- A. Smirl, G.C. Valley, R.A. Mullen, K. Bohnert, C.D. Mire, T.F. Boggess, Opt. Lett. 12, 501 (1987).
- P. Gunter y J.P. Huignard (eds.): "Photorefractive materials and their applications I". Springer-Verlag, Vol. 61 (1988).
- P. Gunter y J.P. Huignard (eds.): "Photorefractive materials and their applications II". Springer-Verlag, Vol. 62 (1988).

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 capítulo 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 capítulo, 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.

Óptica 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.

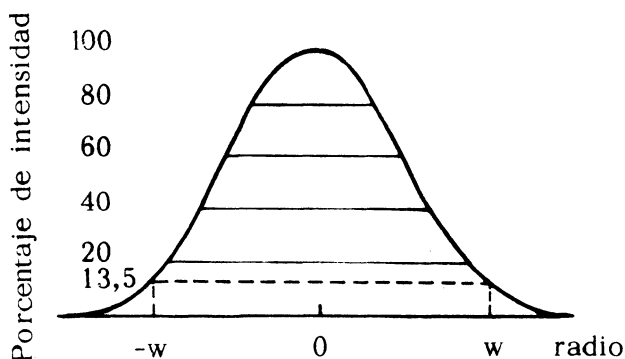


Figura 1

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 TEM₀₀ 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" gaussiano 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

$$w(z) \cong \frac{\lambda z}{\pi w_0}$$

Como z se supone mucho más grande que $\pi w^2/\lambda$, la superficie que

marca el $1/e^2$ de la intensidad tiende a un cono de radio angular (divergencia del haz gaussiano)

$$\theta = \frac{w(z)}{z} = \frac{\lambda}{\pi w_0}$$

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).

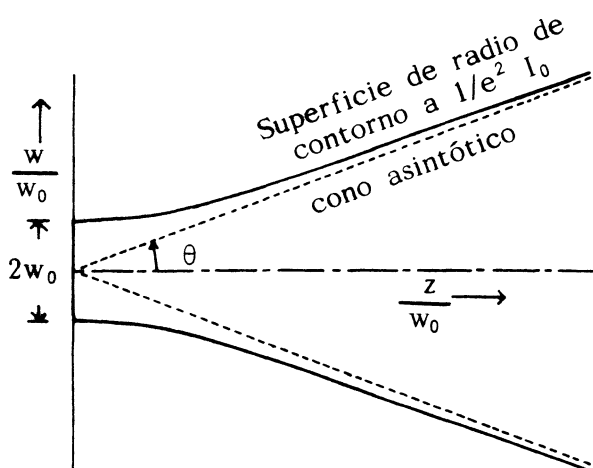


Figura 2

Observando la Guía de Optica 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

$$\theta = \frac{\lambda}{\pi w_0} = \frac{632,8 \cdot 10^6 \text{ mm}}{\pi \cdot 0,4 \text{ m m}} = 5,04 \cdot 10^{-4} \text{ rad.}$$

A la distancia de $z=100$ m

$$w(z) = z\theta = 10^5 \text{ mm } 5,04 \cdot 10^{-4} = 50,4 \text{ mm}$$

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.

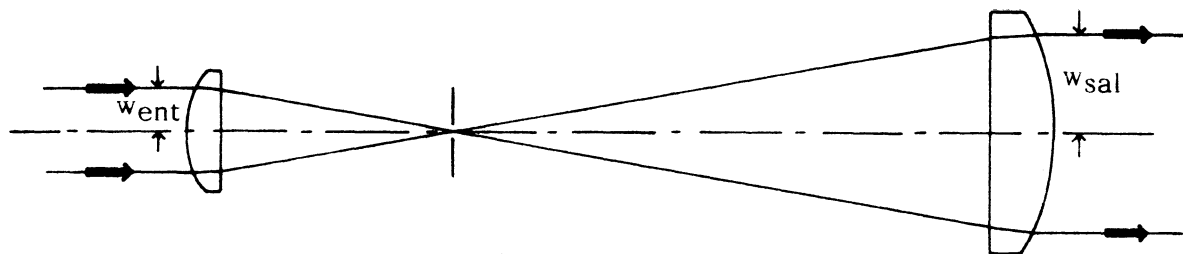


Figura 3

Supongamos que al 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 mayor.

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

$$w_0 = \frac{4}{\sqrt{2}} = 2\sqrt{2} \text{ mm}$$

$$\text{y} \quad 2z_R = \frac{2\pi w_0^2}{\lambda} = \frac{16\pi}{632,8 \cdot 10^{-6}} = 7,944 \cdot 10^4 \text{ mm} \cong 80 \text{ m}.$$

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:

$$w_0 = \frac{\lambda z}{\pi w}.$$

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 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

$$d = f + \frac{(d_0 - f)f^2}{(d_0 - f)^2 + (\pi w_0^2 / \lambda)^2}$$

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 parágrafo, 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 capítulo, 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).

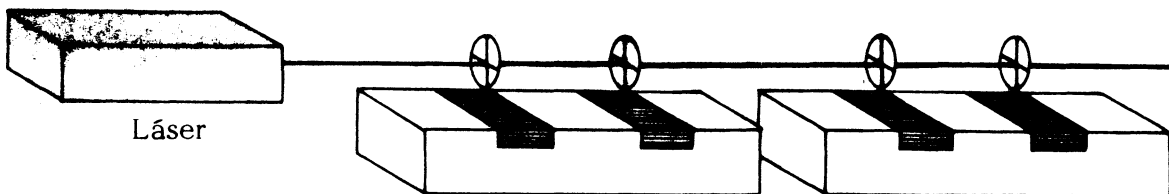


Figura 4

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 la 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.

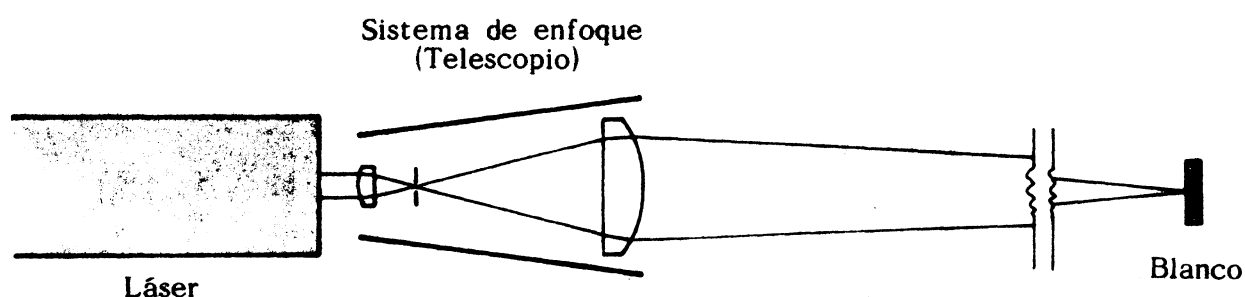


Figura 5

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 fiduciarias 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 7.

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

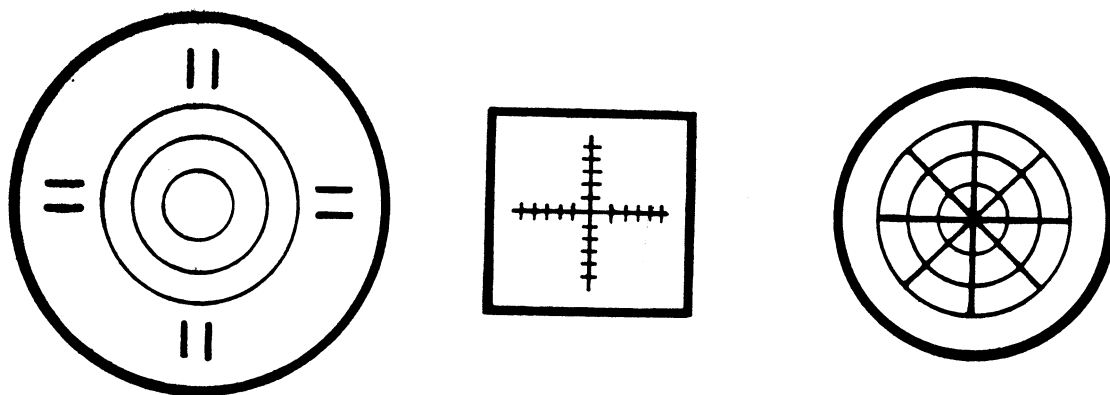


Figura 6

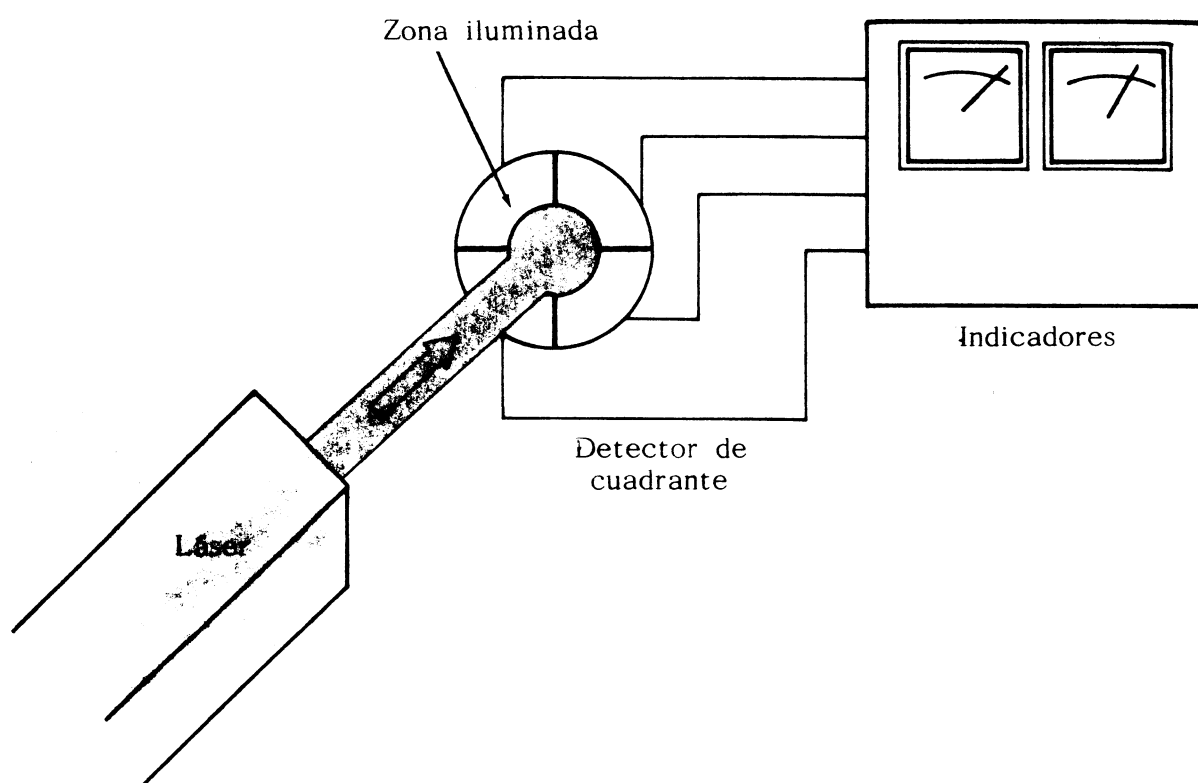


Figura 7

te no coincida con su similar del haz del láser, los indicadores, previamente calibrados, mostrarán el desplazamiento a partir de la posición de centrado.

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 μ 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 parágrafo anterior.

Esto se establece como cierto, siempre que el láser opere en el modo TEM₀₀ 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 μ 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 cen-

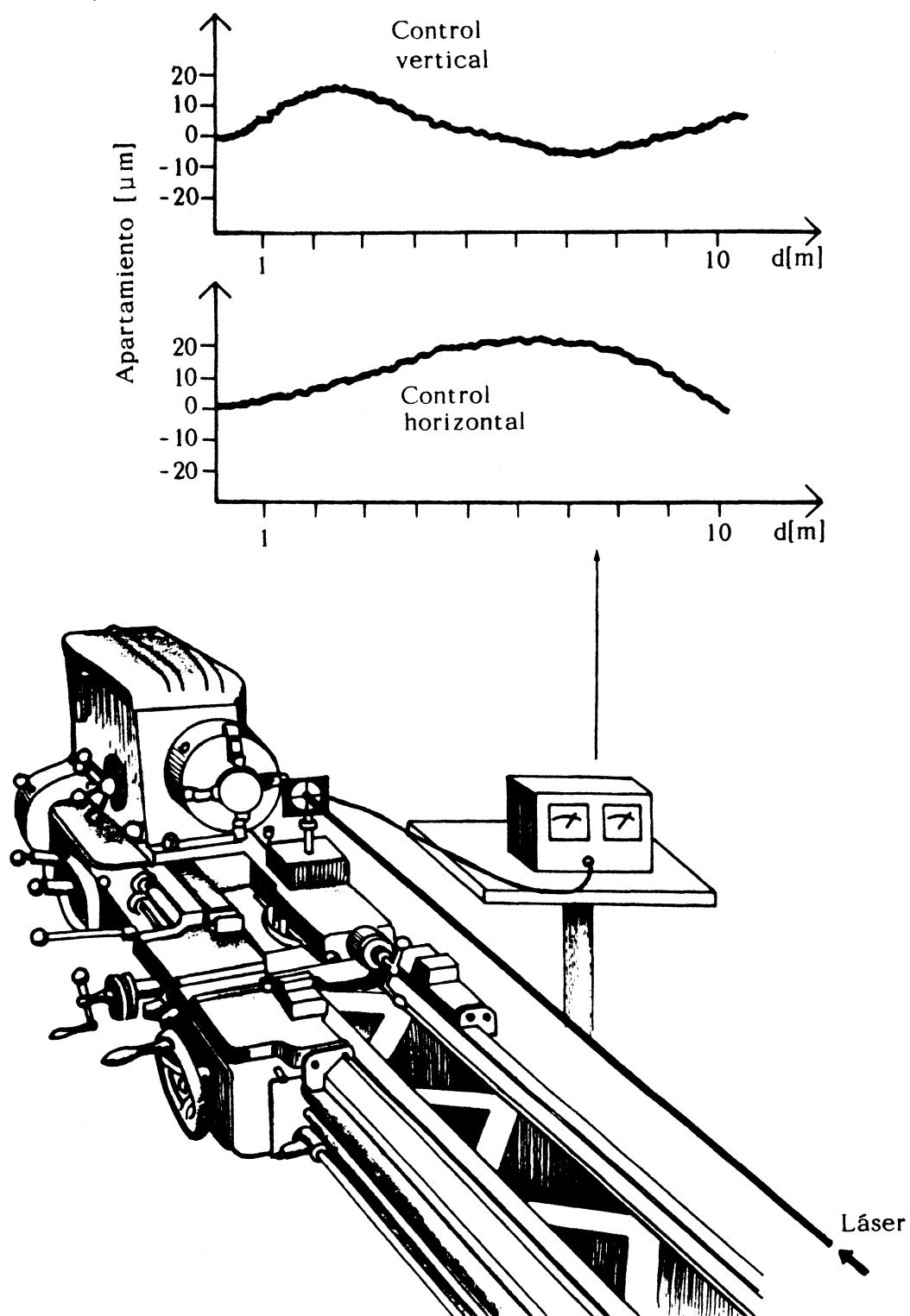


Figura 8

tros.

La calidad del sistema de arañas es determinante para la exactitud del centrado o la precisión en el control de alineación.

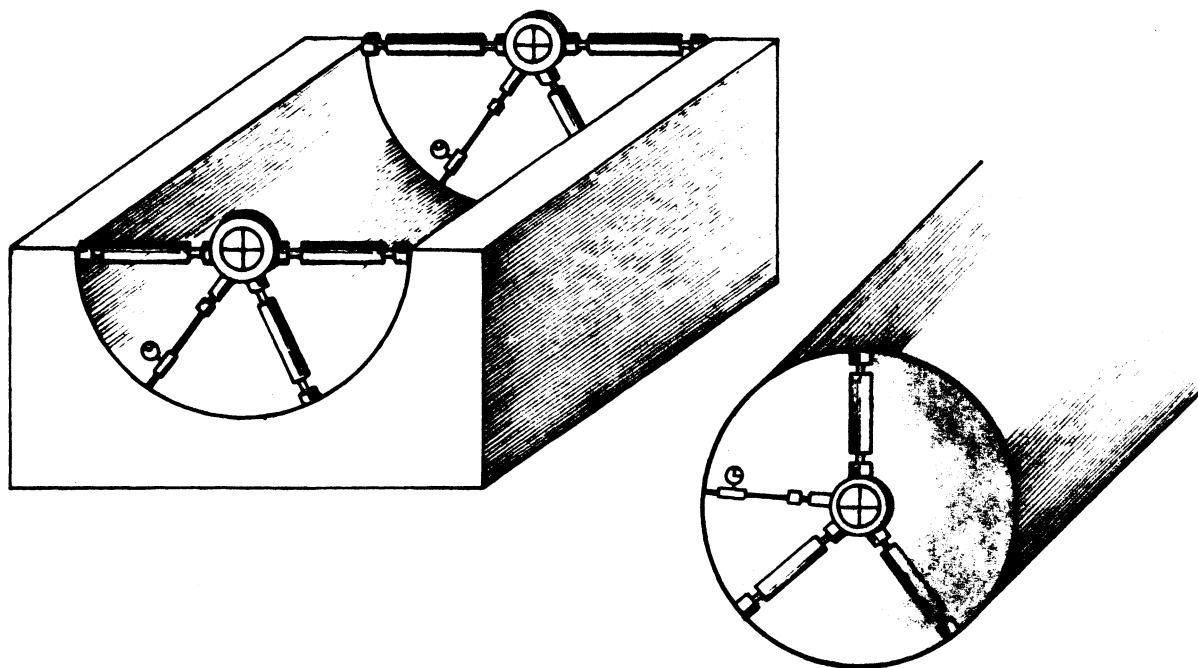


Figura 9

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 en 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 nº 19 [6].

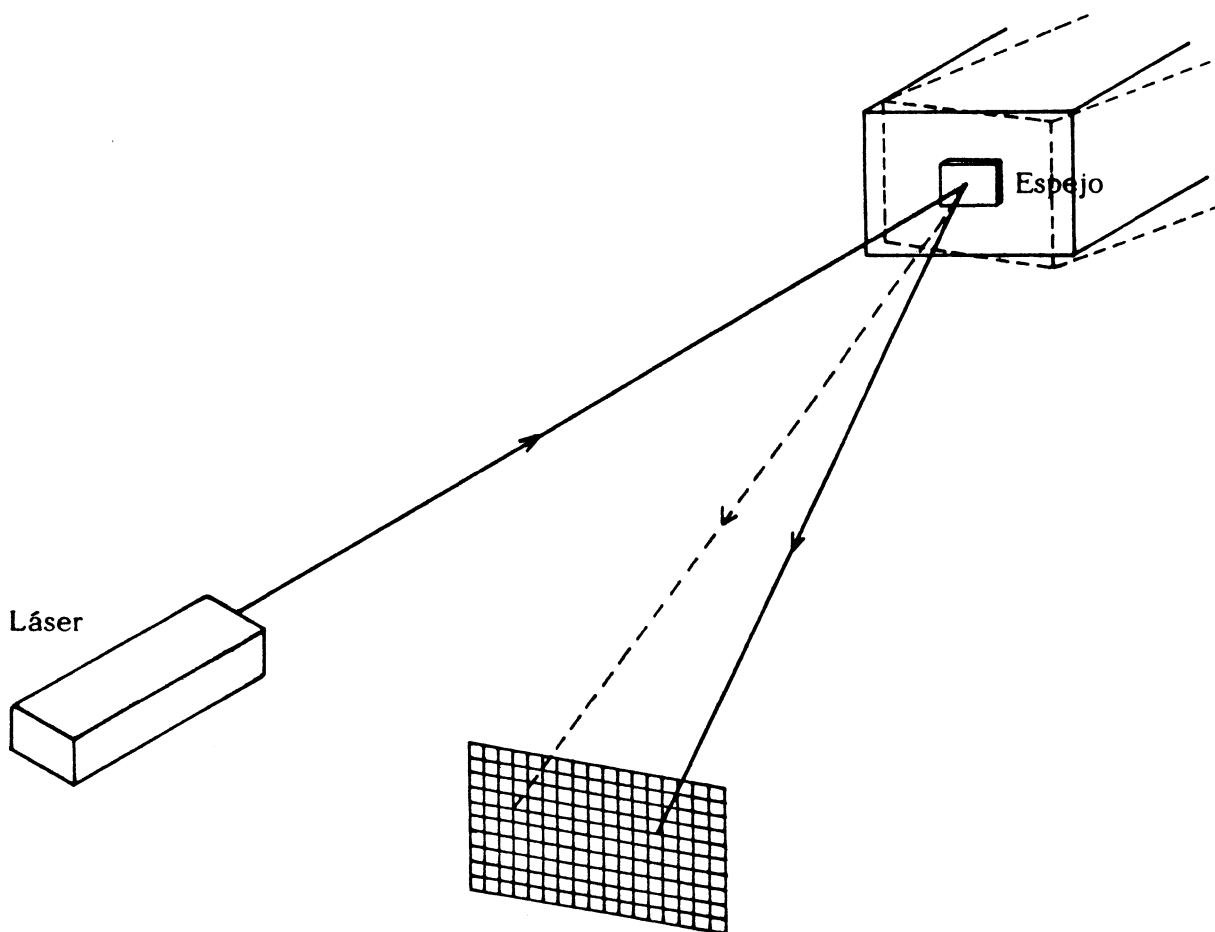


Figura 10

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 de 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 contruidos 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 parágrafos 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.

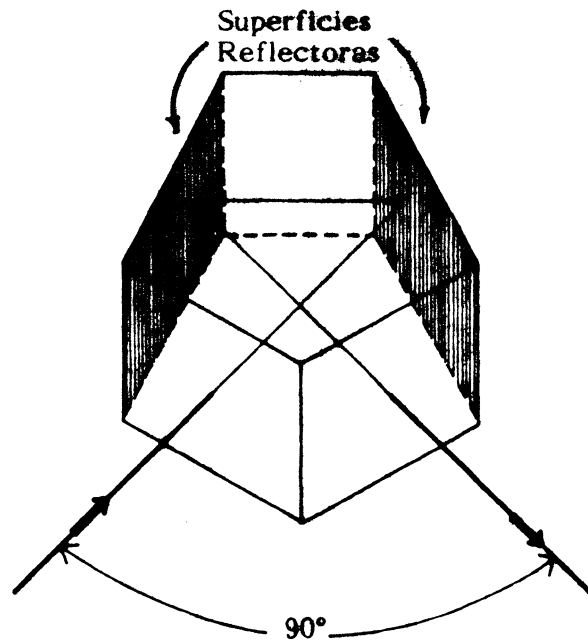


Figura 11

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 perpendicular a dicho eje. El inconveniente, para la generación del plano se concreta en la perfecta construcción del cojinete alrededor del cual gira el eje del espejo, ya que si éste no rota normal al haz incidente, la superficie es cónica.

Para resolver este problema sería necesario construir el plano sin el uso de rotaciones. Una red plana de difracción de pocas líneas por milímetro, iluminada con un haz de luz perpendicular a ella, genera un número finito de rayos que están contenidos en un plano. Si además el haz inicial está compuesto por varios colores, el pincel emergente cubre con más amplitud la superficie plana y, por lo tanto, son más los puntos accesibles del plano [7].

Otras construcciones

A partir de lo expresado en a), b) y c) y con la ayuda de la geometría elemental, se pueden generar planos perpendiculares entre sí, rectas perpendiculares a planos, rectas paralelas, etc. [3,4].

Además tomando como base un plano de luz determinado por alguno de los métodos anteriores, se pueden generar (o controlar) distintos tipos de superficies. La más elemental sería otro plano, pero conociendo la respectiva expresión matemática y disponiendo de instrumental adecuado para "marcar" los puntos en el espacio, es posible resolver distintas superficies.

Lo antedicho se traslada a la práctica con la ayuda del láser, óptica adicional y blancos adecuados (detectores de cuadrante, marcas fiducias, blancos especiales, etc.).

Las tres figuras siguientes muestran algunos arreglos experimentales para la determinación o control de planos y planos perpendiculares a rectas dadas, empleando detectores o blancos colocados sobre montajes maquinados en su base con gran precisión. Esto último determina, de manera excluyente, la exactitud en la definición o control de la superficie.

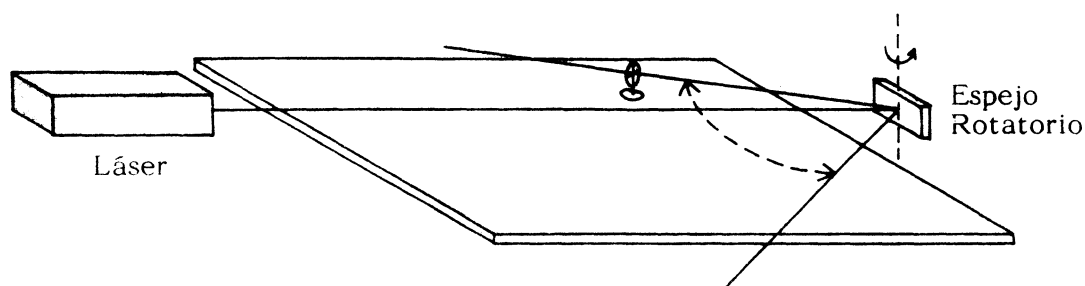


Figura 12

La Figura 12 muestra la manera de generar un plano mediante el uso de un espejo rotatorio y cómo controlar una superficie plana. La Figura 13 presenta la forma alternativa de fabricar un plano mediante el uso de una red de difracción. Por último, la Figura 14 ejemplifica lo expresado en el punto b).

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 suplemen-

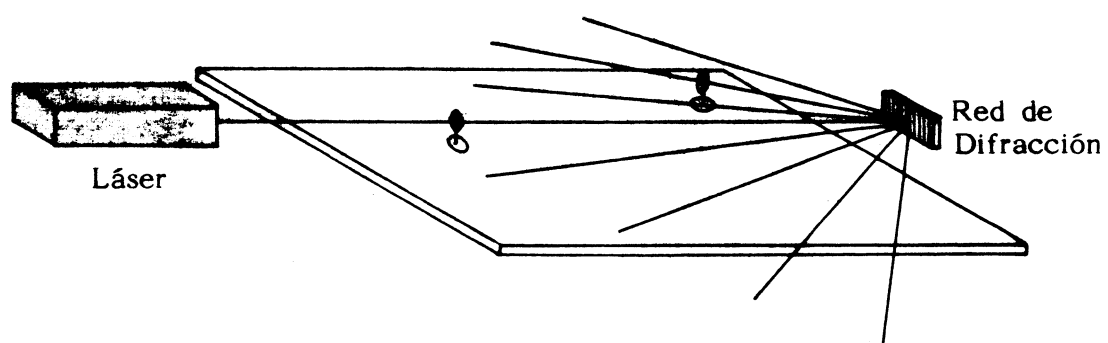


Figura 13

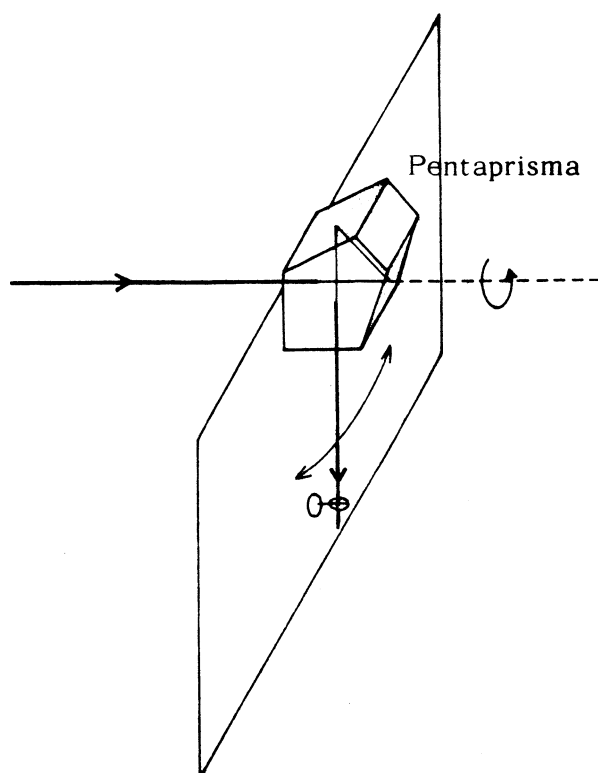


Figura 14

tarios propios o externos y en recintos que no cuentan con espacio libre.

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, yacerá en el mismo plano [8].

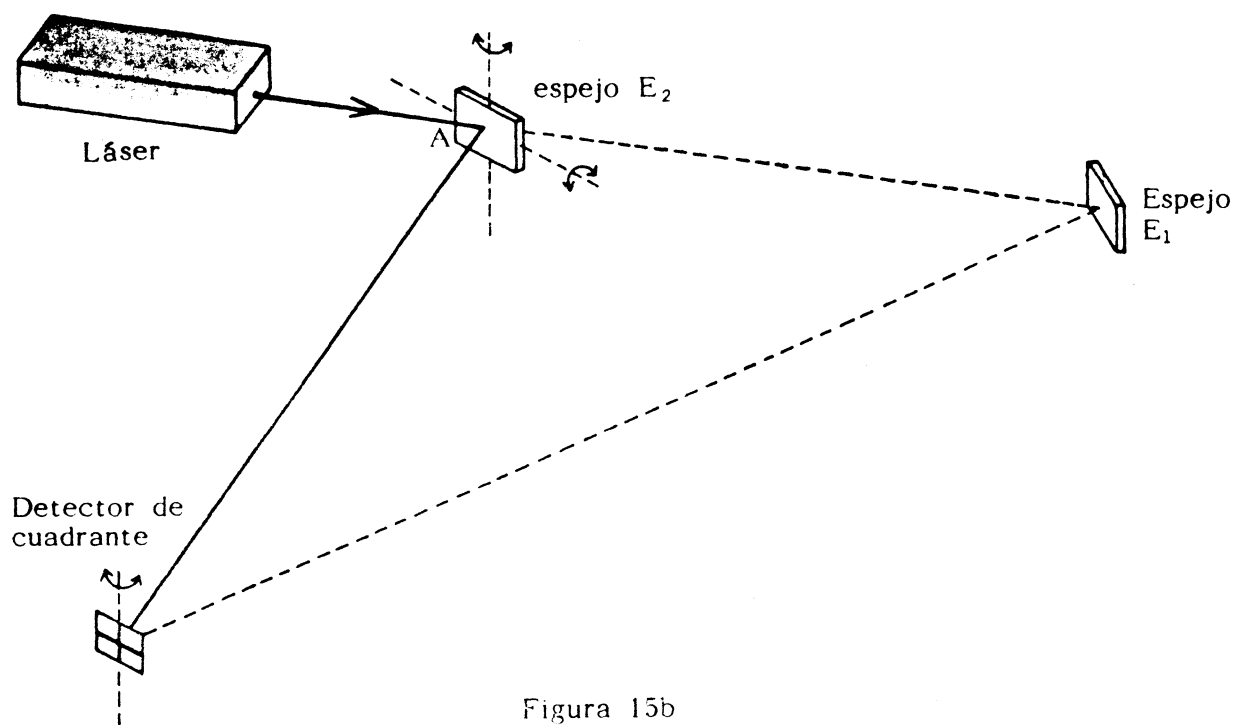
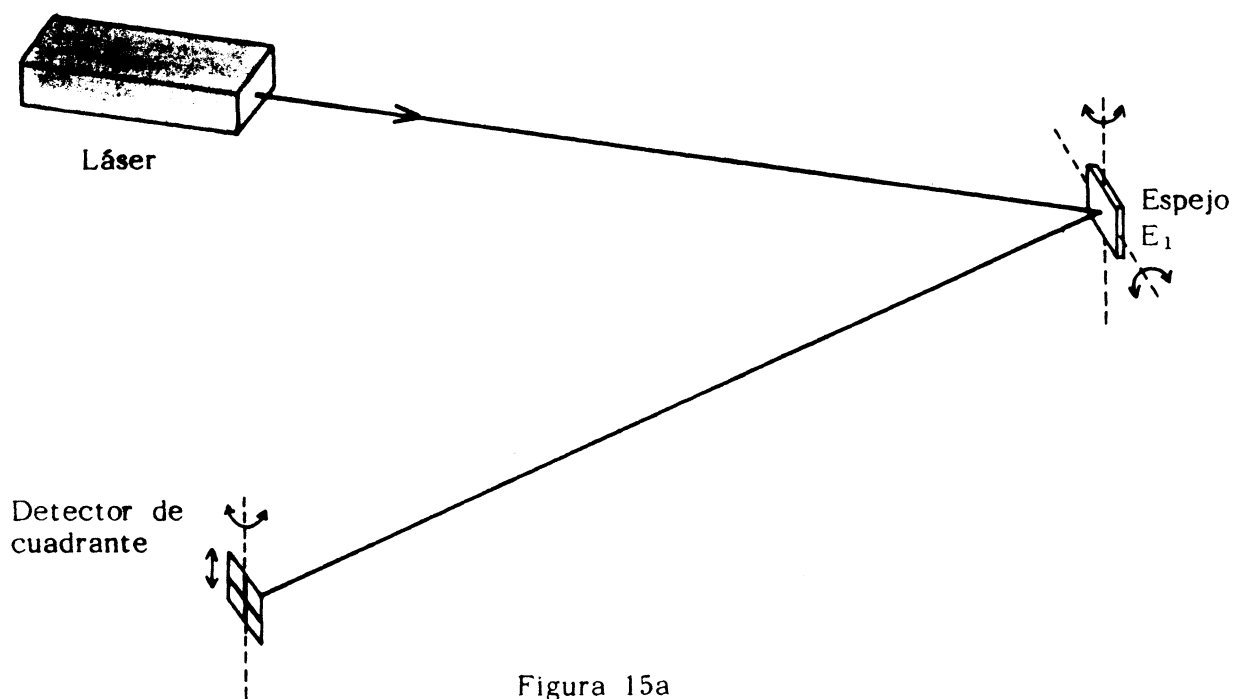
Las rectas originales se determinan de la siguiente manera: una, mediante un rayo de luz proveniente de un láser, y la otra, por su reflexión en un espejo fijo E_1 , Figura 15a. De esta manera es muy sencillo generar la parte básica del plano y lograr su ubicación más conveniente, variando a voluntad 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.



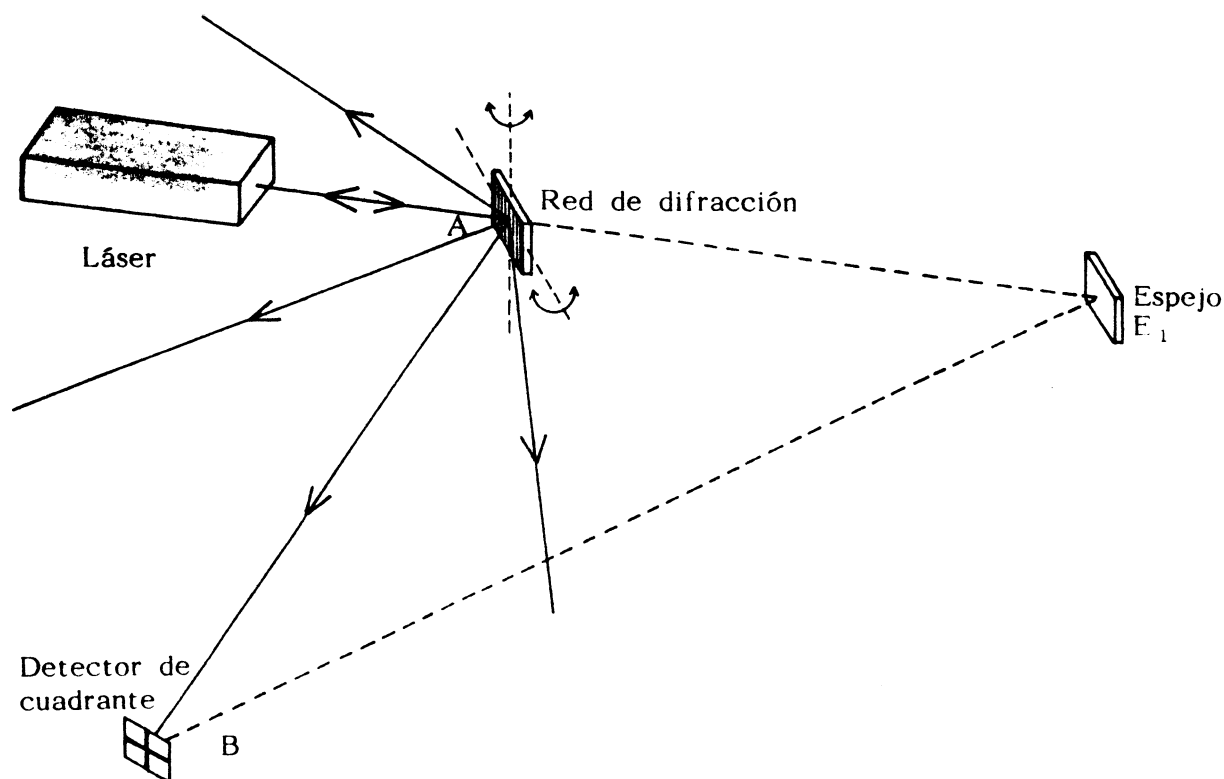


Figura 16

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 bañera 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.

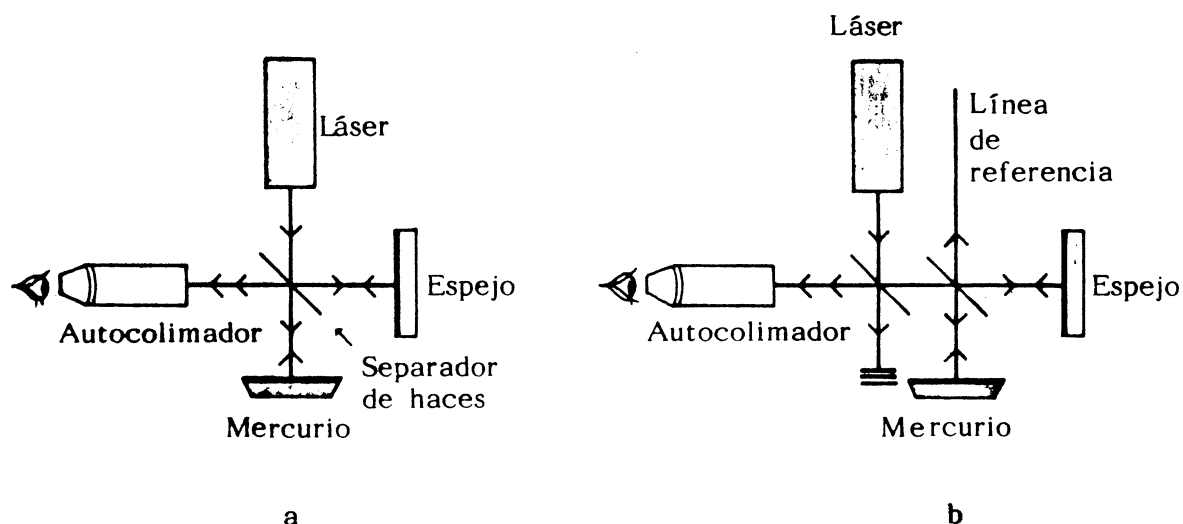


Figura 17

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 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.

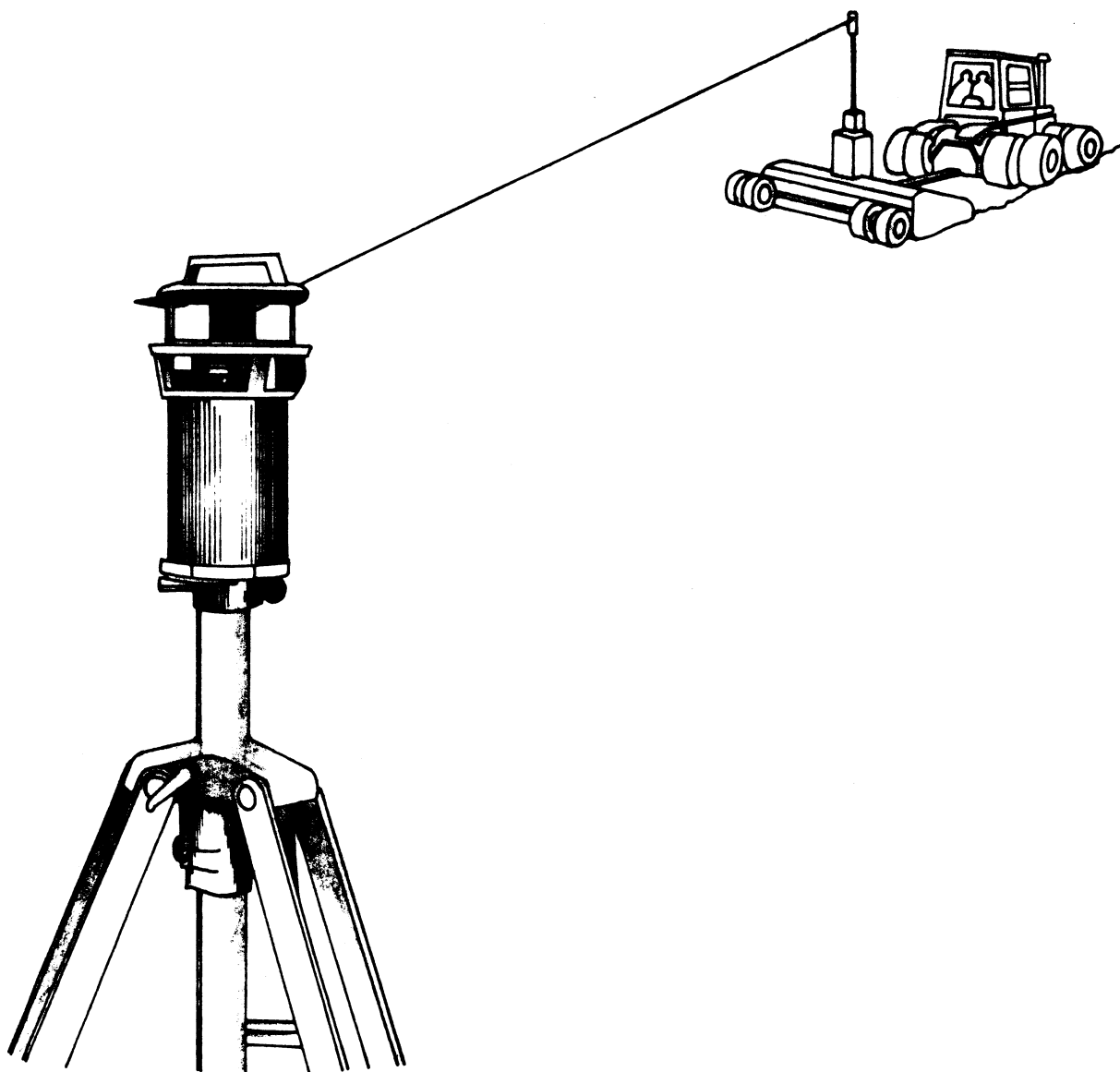


Figura 18

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 rígidamente 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 el 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 del 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.

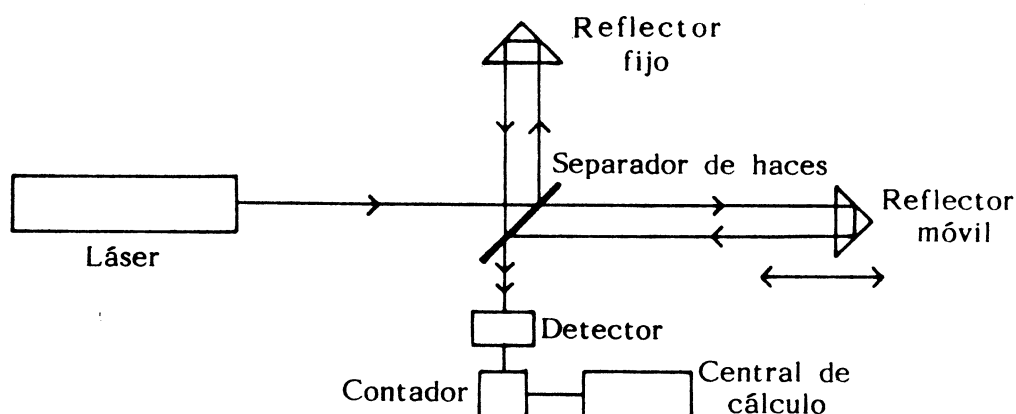


Figura 19

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 coincida 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 período, 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.

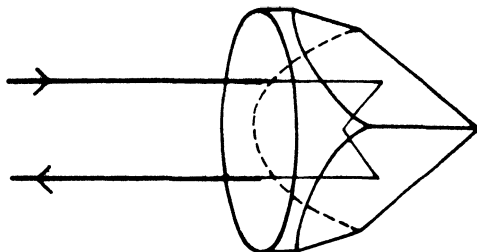


Figura 20

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.

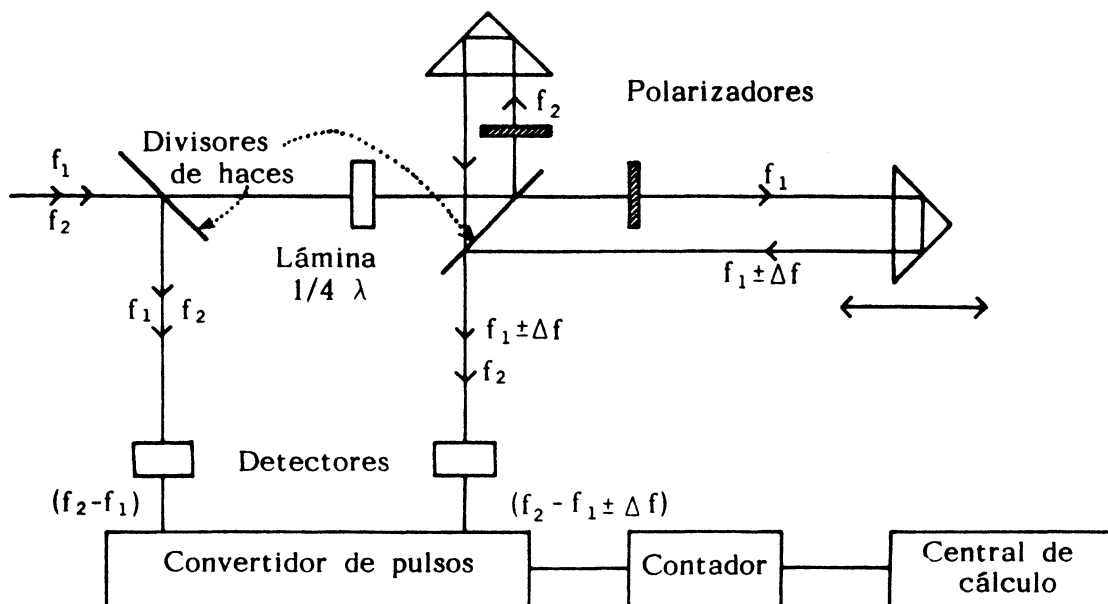


Figura 21

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 retrorreflector.

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 a 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 del 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 sinnúmero 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 Pcia. de Bs. As., 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).

EL LASER EN LA MEDICION DE VELOCIDADES

J. O. Tocho

Introducción

Desde su desarrollo en la década del sesenta, el láser constituye una parte indispensable de los diversos métodos ópticos utilizados en la determinación de velocidades. Estos métodos permiten realizar mediciones de alta localización explorando zonas inaccesibles para otras técnicas, cubrir rangos muy grandes de velocidades y fundamentalmente son no invasivos de manera que operan sin distorsionar al sistema en estudio. Los métodos ópticos basados en el empleo de láseres permiten medir velocidades desde pocos milímetros por segundo, como por ejemplo las encontradas en la circulación en vasos sanguíneos, hasta mediciones supersónicas de algunos miles de metros por segundo, en túneles de ensayos aerodinámicos. El uso de estas técnicas es común en el estudio de sistemas de combustión, en modelos hidráulicos, en el ensayo de turbomáquinas y en el control de trenes de laminado o producción de papel.

El empleo de los métodos ópticos en la determinación de velocidades no es nuevo. Desde principios de este siglo se han venido determinando velocidades de cuerpos astronómicos mediante el empleo del efecto Döppler. Este fenómeno, estudiado por primera vez por el físico austríaco C. J. Döppler en conexión con las ondas sonoras, permite relacionar la frecuencia de la luz con la velocidad relativa entre la fuente y el observador. Mientras que en el sonido, el cambio de la frecuencia se manifiesta en un cambio del tono o de altura, la modificación de la frecuencia luminosa se advierte como un cambio en el color de la luz. Si se analiza el espectro de un cuerpo que se aleja del observador las frecuencias disminuyen y las líneas espectrales que éste emite se corren hacia el rojo. Si el cuerpo se acerca al observador el corrimiento es hacia el azul. Estos desplazamientos son fácilmente perceptibles cuando la fuente luminosa se mueve con una velocidad comparable con la velocidad de la luz ($c = 3 \cdot 10^8$ m/s). Por esta razón el fenómeno se observó por primera vez en los espectros de la luz que llega a la Tierra desde estrellas y galaxias que se mueven a velocidades enormes. Todas las galaxias lejanas muestran un corrimiento hacia el rojo, indicando que se alejan de la Tierra y dando una prueba

inequívoca de la expansión del universo proclamada por diversos modelos cosmológicos. En la Figura 1a se muestran algunos ejemplos de los desplazamientos en las líneas de absorción del calcio presente en la mayoría de las atmósferas estelares en función de la distancia entre la galaxia y la Tierra. Otros ejemplos de las aplicaciones astronómicas del efecto Döppler se muestran en la Figura 1b y 1c. Se han podido separar así estrellas binarias que aún utilizando el telescopio de mayor resolución aparecen como una sola y determinar también la rotación del Sol alrededor de su eje.

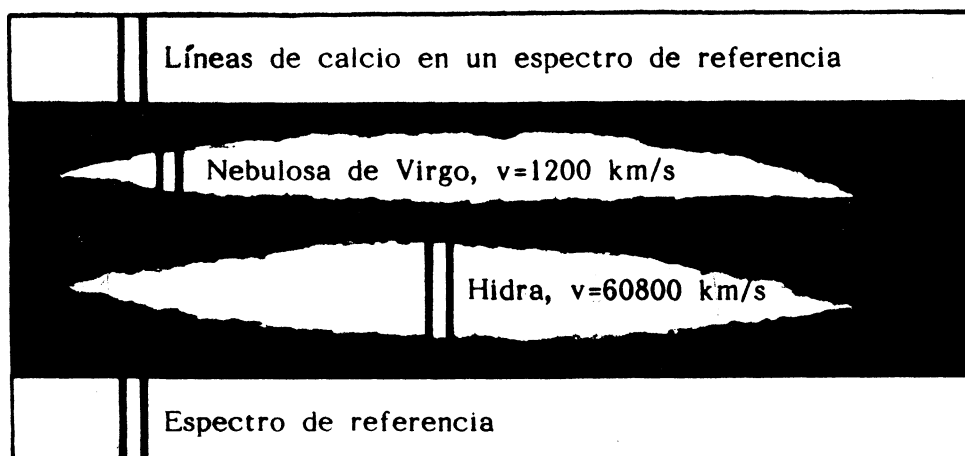


Figura 1a.

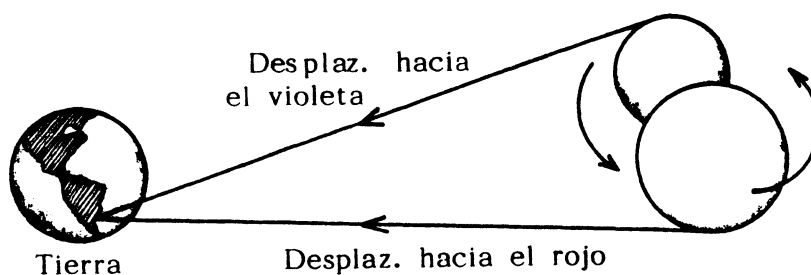


Figura 1b.

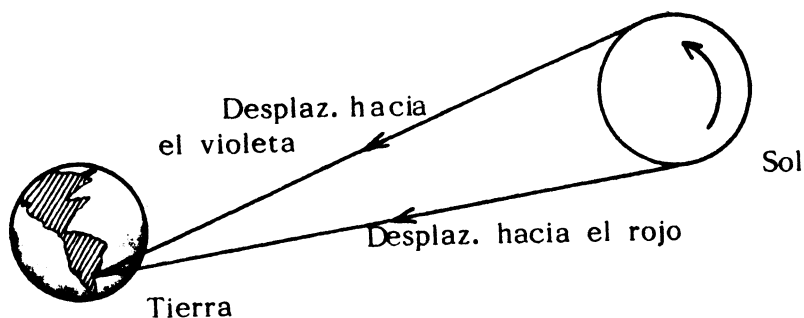


Figura 1c.

Con el advenimiento del láser, una fuente continua de radiación electromagnética coherente en las frecuencias ópticas, las mediciones pudieron trasladarse a escala terrestre. Yeh y Cummins (1), fueron los primeros en demostrar tal posibilidad al construir el primer velocímetro láser por efecto Döppler (LDV) en 1964 y determinar el perfil de velocidades para el flujo laminar en un tubo circular. El método consistía en mezclar coherentemente (heterodinar) la luz dispersada por pequeñas partículas agregadas al fluido con un haz de referencia tomado del mismo láser que iluminaba las partículas. Este éxito inició una intensa serie de trabajos de investigación para desarrollar y aplicar esta técnica en el campo de la medición de velocidades.

A diferencia de las aplicaciones astronómicas, donde es la propia fuente emisora la que se mueve con respecto al observador, las aplicaciones que nos ocupan aquí, implican la iluminación con un láser de pequeñas partículas suspendidas en el fluido en estudio o de pequeñas irregularidades superficiales en el caso de sólidos. son estos entes secundarios los que reemiten la luz hacia el detector (observador) quien producirá una señal proporcional a la velocidad. En el caso de utilizarse el efecto Döppler el corrimiento de frecuencia con respecto a la fuente es doble, primero las partículas en movimiento reciben una frecuencia diferente a la del emisor y segundo las partículas se convierten en fuentes en movimiento con respecto al observador produciendo un nuevo corrimiento.

Mientras los sólidos y el agua, aún el agua de una red potable, contienen un número suficiente de centros dispersores como para ser usados directamente, las mediciones en gases exigen normalmente el agregado de partículas especiales (sembrado). Se utilizan a tal fin sales, humo y aerosoles y en el caso de fluidos, esferitas de poliestireno, gotas de aceite o de agua.

Es gracias al fenómeno de la dispersión que un haz de luz intenso es claramente visible al atravesar una habitación medianamente oscurecida donde existe una elevada concentración de polvo en suspensión o de humo de cigarrillo. Si bien los aspectos cuantitativos de la dispersión dependen de manera complicada del tamaño de las partículas, de su índice de refracción y de la longitud de onda de la luz de iluminación, pueden enunciarse algunas consideraciones generales. La dispersión es más pronunciada hacia adelante, esto es en la dirección de la iluminación. Es más asimétrica cuando las partículas son grandes comparadas con la longitud de onda de la luz. El estado de polarización de la luz dispersada es en general diferente al de la luz incidente. En la Figura 2 puede verse un diagrama polar de la intensidad de la luz dispersada en función de la relación entre el tamaño de las partículas y de la longitud de onda de la luz.

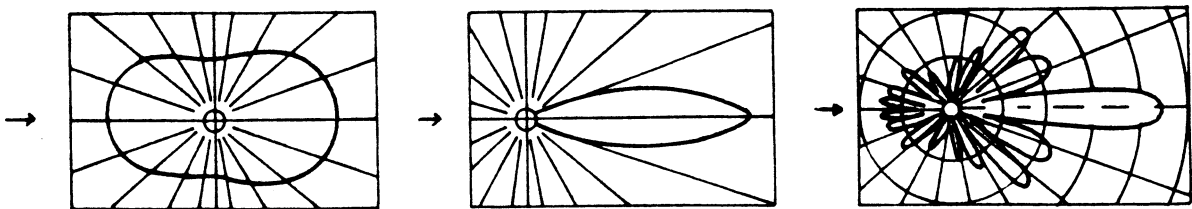


Figura 2.

Mediciones por tiempo de tránsito

El método más simple para determinar velocidades es el velocímetro láser de dos focos (L2F) desarrollado independientemente por Schodl (2) y Jading (3). La idea básica es medir el tiempo de vuelo entre dos zonas iluminadas, de las partículas suspendidas en el fluido o de las irregularidades en un sólido. Este método se conoce también como velocimetría láser de tiempo de tránsito (LTA, de laser transit anemometry) (4). Las zonas iluminadas están a una distancia calibrada y deben ser lo más localizadas posible, por lo que el uso de un láser resulta muy conveniente ya que puede enfocarse el haz que produce en diámetros menores que $100\ \mu\text{m}$.

Cuando una partícula pasa a través de un foco, genera un pulso de disparo y cuando la "misma" partícula llega al segundo foco, se detiene la

cuenta. Conociendo la separación de los focos y el tiempo de vuelo, se determina fácilmente la velocidad. En detrimento de su simpleza, este método choca con algunas dificultades. La probabilidad que una partícula cruce sucesivamente los dos focos es escasa, no mas del 30% para un flujo de aire y por lo tanto el sistema debe distinguir pulsos incorrectos generados cuando la cuenta se detiene por una partícula espúrea que atraviesa el segundo foco sin pasar por el primero (Figura 3).

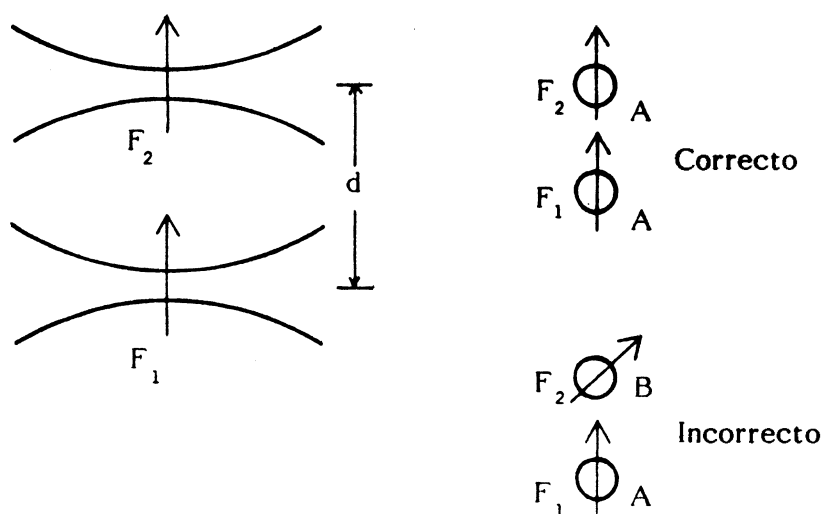


Figura 3.

La precisión puede aumentarse considerablemente si se amplía el número de marcas que deben atravesar las partículas y si se emplean técnicas de correlación digital para identificar los pares "start-stop" de las señales correctas. El primer proceso se logra cruzando dos haces coherentes en la zona a medir y produciendo numerosas franjas de interferencia separadas por una distancia que depende del ángulo de cruce y de la longitud de onda del láser. La correlación compara una forma de onda de la señal del detector con ella misma convenientemente retrasada; el máximo de esta comparación se logra cuando el retraso de la señal es igual al período de la misma. De la determinación de este período y de las constantes que gobiernan la separación entre las franjas se determina la velocidad (5). En estas condiciones el método se torna completamente equivalente al método Döpler heterodino simétrico que será analizado más adelante.

Las mediciones en sólidos no necesitan en general de estas sofisticaciones; es más, en algunos casos de movimientos unidimensionales puede

prescindirse del empleo de láseres e iluminar la zona de la medición con una fuente convencional enfocada por una lente o bien utilizar diodos emisores de luz (LED).

El efecto Döppler

Consideremos una fuente o emisor que emite una onda de frecuencia ν . Con respecto a un sistema de referencia fijo al observador, la fuente se mueve con velocidad $\vec{v}$. Si la observación se realiza en la dirección $\vec{r}/r$, el observador estacionario registrará una frecuencia ν' diferente a la emitida por la fuente. Esto se debe a que las ondas emitidas estarán concentradas en una distancia $(c - \vec{v} \cdot \vec{r}/r) \cdot t$, mientras que para una fuente estacionaria la distancia equivalente está dada por $c \cdot t$ (Figura 4).

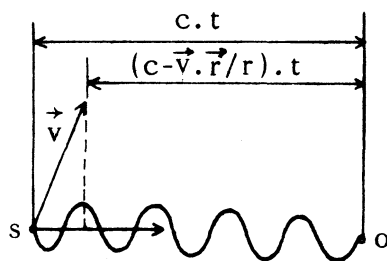


Figura 4.

De la constancia de la velocidad de la luz resulta:

$$c = \lambda \nu = \lambda' \nu' ,$$

y de la Figura 4, si T es el periodo de la fuente:

$$\lambda' = \frac{(c - \vec{v} \cdot \vec{r}/r) \cdot t}{t/T} = \frac{c - \vec{v} \cdot \vec{r}/r}{\nu} ,$$

entonces la frecuencia registrada por el observador es:

$$\nu' = \frac{c}{\lambda'} = \frac{c\nu}{c - \frac{\vec{v} \cdot \vec{r}}{r}} = \frac{\nu}{1 - (v/c)\cos\theta},$$

donde v es el módulo de la velocidad y θ es el ángulo entre $\vec{v}$ y $\vec{r}$. Si se linealiza la ecuación restringiéndose al caso en que $v \ll c$, resulta:

$$\nu' = \nu(1 + (v/c)\cos\theta). \quad (1)$$

Si es el observador el que se mueve con respecto a la fuente, un análisis similar muestra que la frecuencia ν' registrada por el observador es:

$$\nu'' = \nu(1 - (v/c)\cos\theta). \quad (2)$$

Estos cambios de frecuencia son normalmente tan pequeños que impiden su medición directa (el caso astronómico ya ha sido considerado y nos restringimos ahora a mediciones terrestres). La forma práctica de utilizar este efecto es superponer sobre un detector que responda a la intensidad luminosa (fotodiodo o fotomultiplicador) dos ondas de frecuencias ν_a y ν_b corridas por efecto Doppler con respecto a una frecuencia ν original. Llegan entonces al fotoreceptor dos señales de amplitud E_{oa} y E_{ob} con frecuencias ν_a y ν_b :

$$E_a = E_{oa} \exp[i(2\pi\nu_a t)],$$

$$E_b = E_{ob} \exp[i(2\pi\nu_b t)],$$

y la intensidad medida por el detector será:

$$I \propto |E_a + E_b|^2 = (E_a + E_b) \cdot (E_a^* + E_b^*),$$

si las ondas con coherentes, resulta

$$I = I_a + I_b + 2\sqrt{I_a I_b} \cos[2\pi(\nu_a - \nu_b)t] , \quad (3)$$

y la señal del receptor tiene una modulación con frecuencia ν_D igual al batido de las dos ondas:

$$\nu_D = \nu_a - \nu_b . \quad (4)$$

Métodos Doppler en la determinación de velocidades

Luego de la demostración experimental de la factibilidad de medir velocidades por efecto Doppler utilizando láseres (1) se idearon distintas configuraciones prácticas para materializar las ideas descriptas en el punto anterior. Los sistemas básicos son tres:

- a) haz de referencia (Figura 5),
- b) heterodino diferencial (Figura 6) y
- c) heterodino simétrico (Figura 7).

En todos estos sistemas los componentes básicos son los mismos: láser, divisor de haz, lentes de enfoque y de recolección, apertura, fotodetector y sistema de procesamiento de la señal.

Haz de referencia: en este arreglo se hace que uno de los haces que se intersectan incida directamente sobre el fotodetector. Para un buen funcionamiento, en este caso se requiere una elevada concentración de partículas. Se puede conseguir un pequeño volumen de medición, lo que facilita la medición de gradientes acentuados de velocidad pero, se requiere una óptica compleja y de difícil alineación. Deriva este método del trabajo de Yeh y Cummins quienes tomaban la referencia directamente desde el láser.

Heterodino diferencial, doble haz ó método de las franjas: aquí se intersectan dos haces de igual intensidad y se recoge solamente la luz dispersada, siendo interceptados los haces láser directos. Es posible trabajar con bajas concentraciones de partículas. Su ventaja principal radica en la gran flexibilidad, ya que el volumen de medición puede regularse para que esté adaptado al tamaño de las partículas y al gradiente de velocidades. La óptica es relativamente sencilla y su alineación es fácil.

Heterodino simétrico: en este caso se emplea un único haz cuya luz es dispersada por las partículas. En la recolección, se intercepta el haz directo, recogiendo solamente la luz dispersada. Este arreglo puede funcionar con baja concentración de partículas. La óptica requerida es simple y la alineación es relativamente fácil y el ensanchamiento en el espectro de la señal Doppler que surge del hecho de tener que utilizar aperturas necesariamente finitas, es relativamente menor que en el caso de doble haz.

Sin embargo la zona de medición no es de fácil observación ni resulta tan bien definida como en el caso anterior por lo que es un método poco usado en equipos comerciales.

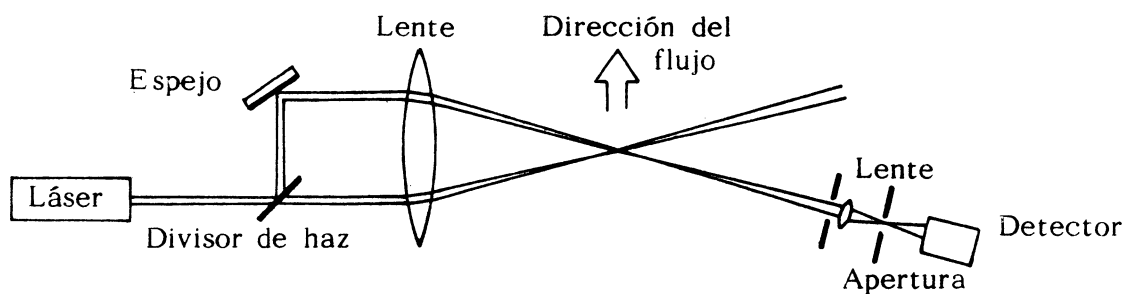


Figura 5.

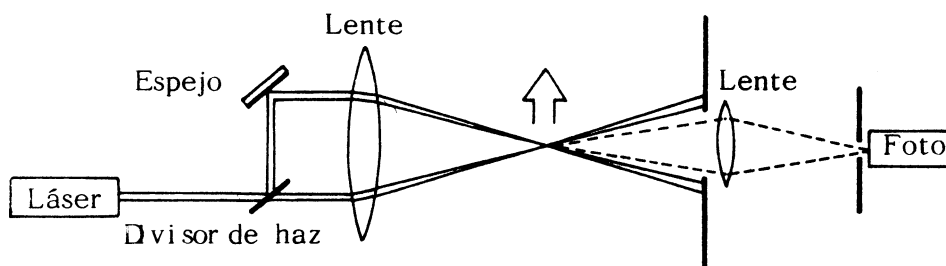


Figura 6.

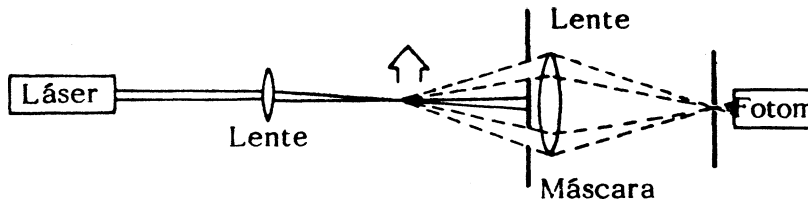


Figura 7.

Analicemos con detalle el origen de la señal en el método heterodino diferencial, los resultados son válidos para los otros dos casos. Consideremos dos haces coherentes que se intersectan en la zona de medición formando un ángulo 2α (Figura 8). Las partículas en suspensión en el fluido o las irregularidades en un sólido actúan primero como receptores. Para el haz 1 la frecuencia recibida por la partícula es ν_1 y para el haz 2 es ν_2 . Por aplicación de la ecuación (2) resulta:

$$\nu_1 - \nu = v(v/c)\sin \alpha ,$$

$$\nu_2 - \nu = v(v/c)\sin \alpha ,$$

En una segunda etapa los receptores se convierten en emisores. En la dirección de la bisectriz del ángulo que forman los haces incidentes, esto es a 90° con respecto a la velocidad v considerada, la ecuación (1) indica que no hay corrimiento de frecuencia entre la emitida y la registrada por el receptor.

Llegan entonces al fotoreceptor las ondas dispersadas del haz 1 y del haz 2. Estas dos ondas (E_1 y E_2) tienen amplitud E_{01} y E_{02} y sus frecuencias ν_1 y ν_2 están dadas por las relaciones anteriores.

Estas ondas de frecuencia similar se "baten" en el fotoreceptor que produce una señal proporcional a la intensidad promediada sobre varios ciclos. Sobre el receptor E_1 y E_2 se pueden escribir:

$$E_1 = E_{01} \exp[i(2\pi\nu_1 t)] ,$$

$$E_2 = E_{02} \exp[i(2\pi\nu_2 t)] ,$$

de manera que la intensidad $I \propto |E_1 + E_2|^2 = (E_1 + E_2)(E_1^* + E_2^*)$, resulta de acuerdo con la ecuación (3):

$$I = I_1 + I_2 + 2\sqrt{I_1 I_2} \cos(2\pi(\nu_1 - \nu_2)t) ,$$

donde I_1 es la intensidad que llega al fotoreceptor si se bloquea el haz 2 e I_2 es la intensidad detectada si se intercepta el haz 1. La señal proporcional a la intensidad tiene el aspecto de la Figura 9.

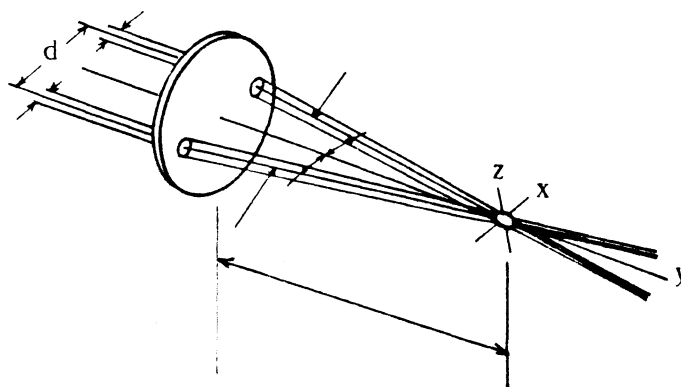


Figura 8.

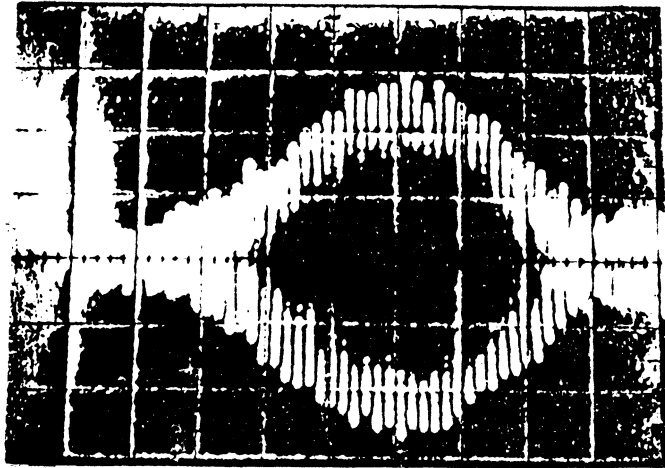


Figura 9.

Mientras que la envolvente de esta señal es función del tamaño de los emisores que pasen por la zona de intersección, la frecuencia de la pulsación es directamente proporcional a la velocidad y su determinación es el eje central de todos los métodos basados en el efecto Doppler. Si llamamos ν_D a esta frecuencia, tenemos:

$$\nu_D = \nu_1 - \nu_2 = 2\nu(v/c)\sin \alpha ,$$

a partir de la medición de esta frecuencia se calcula la velocidad buscada:

$$v = \frac{c\nu_D}{\nu 2 \sin \alpha} = \frac{\lambda}{2 \sin \alpha} \cdot \nu_D \quad (5)$$

Para una disposición mas o menos convencional que utiliza un láser de He-Ne ($\lambda=632,8$ nm), haces separados por 35 mm y enfocados por una lente de 250 mm de distancia focal, la frecuencia Doppler resulta 221 KHz para $v = 1\text{m/s}$.

Se debe recalcar que v es la componente de la velocidad $\vec{v}$ en el plano de los haces de iluminación y que es perpendicular a la dirección de observación. La extensión del método a la determinación de la otra componente en el plano perpendicular a la observación se puede encarar de dos maneras diferentes. En el caso en que las mediciones se puedan hacer secuencialmente la solución consiste en que gire 90° el plano de los haces de iluminación y realizar otra medición. Si las mediciones deben ser simultáneas el módulo de recepción debe duplicarse y estar preparado para distinguir de alguna manera las señales de una componente y de la otra. Normalmente se utiliza la polarización o el color (láser de argón $\lambda=488$ nm, $\lambda=514$ nm) para individualizar a las componentes. La determinación simultánea de las tres componentes de la velocidad, si bien es posible, es mucho más complicada y no será tratada aquí (6).

El modelo de franjas

Existe una forma alternativa de analizar el funcionamiento del método heterodino simétrico que permite visualizar el efecto de la concentración y del tamaño de los emisores sobre la calidad de la señal. Puede pensarse que los emisores al pasar por la zona de intersección de los dos haces atraviesan zonas alternadas de luz y sombra que surgen de la interferencia de los haces coherentes de iluminación. La separación entre estas franjas, df , resulta igual a $\lambda/2\sin\alpha$ y el tiempo que tarda la partícula en atravesarla es igual a df/v , de modo que habrá una señal con un periodo $T = df/v = \lambda/2v\sin\alpha$ y la frecuencia $\nu=1/T$ será igual a la ya calculada $\nu = \nu_D = (2v/\lambda)\sin\alpha$.

Para haces con una distribución de intensidad gaussiana la zona de intersección tiene forma de elipsoide de revolución cuyo contorno se define de manera que la intensidad sea $1/e^2$ de la intensidad en el centro (Figura 10). El semieje mayor de dicho elipsoide está en la dirección de observación. Puede calcularse el largo de la zona de medición, l_m , como:

$$l_m = 4\lambda f / \pi D \sin\alpha ,$$

donde D es el diámetro de los haces de iluminación. En la dirección transversal, el diámetro del elipsoide resulta:

$$d_m = 4\lambda f / \pi D \cos\alpha = l_m \cdot \operatorname{tg}\alpha .$$

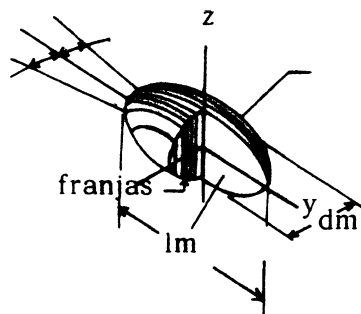


Figura 10.

El número de franjas contenidas en el elipsoide es entonces $N_f = dm/(\lambda/2\sin\alpha) = 4d/\pi D$, donde d es la separación de los haces de iluminación antes del enfoque. Retomando la discusión del ejemplo práctico que estamos desarrollando y considerando un láser de He-Ne de 5 mW con un diámetro de haz de 0,8 mm (Spectra Physics modelo 120), resulta:

$$l_m = 3,6 \text{ mm} \quad , \quad d_m = 0,25 \text{ mm} \quad , \quad d_f = 4,5 \text{ m} \quad \text{y} \quad N_f = 56 \text{ .}$$

Para reducir el tamaño de la zona de medición, alcanzando así mayor resolución espacial, es necesario aumentar el diámetro D de los haces de iluminación trabajando con expansores tipo telescopios invertidos y aumentar el ángulo de convergencia α empleando lentes de mayor diámetro y menor distancia focal. La solución es costosa y se puede decir que el límite práctico para un sistema comercial es 1 mm para la longitud de la zona de medición. Este es el factor que limita la resolución espacial del método.

Si las partículas suspendidas en el fluido o las irregularidades en un sólido tienen un tamaño mayor que la separación entre las franjas, o su concentración hace que estadísticamente haya más de una por vez dentro del elipsoide, la señal pierde nitidez. El efecto del tamaño de las partículas es producir una menor modulación de la señal (Figura 11a). La señal que proviene de dos o más partículas simultáneas adolece de los mismos problemas, puede no tener una frecuencia única y normalmente invalida la medición (Figura 11b).

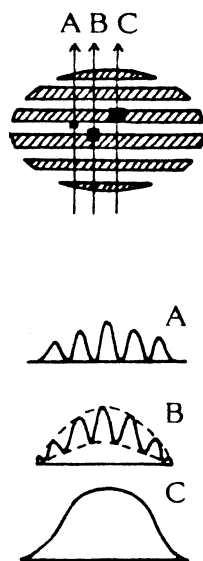


Figura 11a.

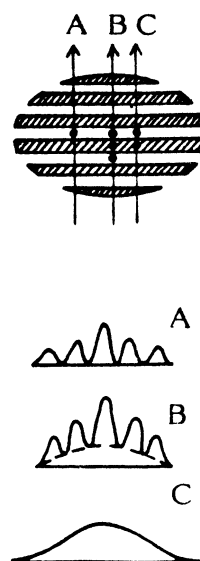


Figura 11b.

En los sistemas que utilizan agua potable es normalmente innecesario el agregado de partículas. Este fluido contiene suficiente concentración de partículas con tamaño comprendido entre $0,3$ y $3 \mu\text{m}$ como para trabajar con láseres con menos de 5 mW de potencia sin recurrir al sembrado de sustancias especiales. En otros líquidos o en aire el sembrado es normalmente inevitable. En el ejemplo que se está utilizando en esta descripción, la separación entre franjas es del orden de $5 \mu\text{m}$, de manera que las señales deben ser de buena calidad aún empleando agua sin sembrar. Como los valores típicos de las partículas suspendidas en el agua potable cambian de un lugar a otro, es conveniente comenzar cualquier estudio de velocimetría Doppler, caracterizando a las partículas, tanto en tamaño como en concentración.

Procesamiento de la señal

Es claro que en cualquiera de los métodos descriptos la información está contenida en el periodo o frecuencia de la señal y no en la amplitud

de la misma. Esta última responde al tamaño de las partículas o irregularidades que atraviesan la zona iluminada. Generalmente el procesamiento de la señal comienza en un filtro pasa banda centrado en la frecuencia media de la señal y con un ancho de banda compatible con la turbulencia o variaciones esperadas en la velocidad (Figura 12). Posteriormente un detector de nivel elimina las señales muy pequeñas y las muy grandes. Las primeras tienen relación señal-ruido insuficiente y las segundas provienen de partículas muy grandes y tienen las dificultades mencionadas en el punto anterior. El paso siguiente es eliminar completamente la dependencia con la amplitud y retener sólo la frecuencia o periodo. Para ello se genera una onda rectangular de altura constante que comienza y termina en cada paso por cero de la señal original. Para reducir al mínimo el desfase entre la señal analógica y la digital se utiliza un circuito tipo Schmitt con tiempo de respuesta mínimo.

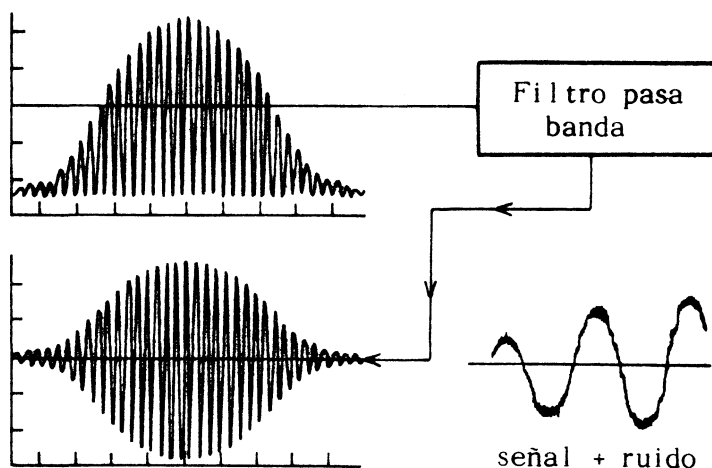


Figura 12.

La última etapa de procesamiento de la señal es precisamente la determinación de su frecuencia. Si el procesador es del tipo "correlador", se toman muestras de la señal cada un tiempo establecido y se realiza el proceso matemático de correlación. A tal fin se suman los valores de las memorias separadas por un número constante de intervalos de muestreo, luego se varía el número de intervalos y se encuentra el máximo de dicha suma. Se determina de esta manera el tiempo que tarda la señal en replicarse como el tiempo necesario para maximizar la correlación. Una vez calcu-

lado el periodo de la señal se determina su frecuencia y puede evaluarse la velocidad.

Si el procesador es del tipo "contador" existen diversos métodos para determinar el periodo de la señal. En uno de ellos se fija en N el número de pulsos que se desean considerar y se mide el tiempo τ_N equivalente a N periodos comparando con un reloj de cristal. En un equipo comercial se mide también el periodo correspondiente a $N/2$ pulsos y si la diferencia entre las dos medidas es superior a cierto porcentaje se desecha la medida.

El error en la medición de la frecuencia Doppler se reduce midiendo tantos pulsos como sea posible. La fuente principal de error es el ruido superpuesto a la señal de frecuencia ν_D y que conduce a una indeterminación ε en la medición del periodo. La frecuencia medida es entonces para N ciclos:

$$\gamma = \frac{N}{\tau_N \pm \varepsilon} = \frac{N}{\tau_N} \left(\frac{1}{1 \pm \varepsilon/\tau_N} \right) = \frac{N}{\tau_N} \left(1 \mp \frac{\varepsilon}{\tau_N} \right) = \nu_D \pm \delta \nu$$

y el error relativo $\delta \nu / \nu$ resulta:

$$\frac{\delta \nu}{\nu} = \frac{N \varepsilon / \tau_N^2}{N / \tau_N} = \frac{\varepsilon}{\tau_N} = \frac{\varepsilon}{N \tau_1},$$

donde τ_1 es el periodo de la señal. Se ve claramente la conveniencia de medir un número grande de pulsos. El límite para la precisión con que se puede medir el tiempo está dado por la frecuencia del reloj con que se compara. El error con que se mide el tiempo coincide con un pulso o ciclo del reloj; entonces $\varepsilon_c = 1/f_c$ y el error relativo mínimo en la frecuencia es:

$$\frac{\delta \nu}{\nu} = \frac{1}{N \tau_1 f_c},$$

para señales de buena calidad donde $\varepsilon \approx \varepsilon_c$.

Este razonamiento es válido hasta que el número de pulsos N es alrededor de 4/5 del número total de franjas N_f . Mas allá, la calidad de

la señal se deteriora porque las franjas de los bordes del elipsoide tienen poca modulación y pueden mostrar cierta distorsión si la alineación de la óptica no es perfecta. La lente de recolección y el diafragma frente al fotodetector deben diseñarse para seleccionar la luz dispersada en la región central del elipsoide.

Con un reloj de 10 MHz y la disposición que hemos estado discutiendo ($v=221$ KHz para $v=1$ m/s) el error en la medición de la velocidad es del orden del 2% para un ciclo y 0,05% para $N=45$, el número máximo de pulsos.

Si se tiene en cuenta el error introducido por la medición del ángulo α se tiene:

$$\frac{\Delta v}{v} = \frac{\delta \alpha}{\tan \alpha} = 0,4\% ,$$

donde se ha tomado una indeterminación de 1 minuto de arco en el ángulo α y $\tan \alpha = 17,5\text{mm}/250\text{mm} = 0,07$ para haces separados por 35 mm antes de ser enfocados por una lente de 250 mm de distancia focal. Los equipos comerciales especifican precisiones del orden del 0,5% para mediciones en fluidos y hasta el 0,1% en equipos diseñados especialmente para ser utilizados en sólidos con movimiento unidimensional, como en los trenes de laminación.

Aplicaciones

Las mediciones que exigen la instrumentación más simple son las mediciones unidimensionales en sólidos y las de fluidos en condiciones de flujo laminar. El perfil de velocidad en un tubo se ha medido y comparado satisfactoriamente con la solución exacta de la ecuación de Navier-Stokes con una precisión de 0,2%. Controlando el volumen de medición hasta la zona de $10 \mu\text{m}$ utilizando microscopios invertidos se han medido velocidades de ~ 1 mm/s en capilares de $200 \mu\text{m}$ de diámetro. Estos trabajos tienen como objetivo llegar a realizar mediciones en vivo del flujo sanguíneo.

La velocidad de respuesta de los velocímetros láser ha permitido realizar mediciones en fluidos con turbulencias y estudiar las correlaciones espaciales y temporales. Estas mediciones fueron realizadas en ductos, jets libres y coaxiales, conjuntos pistón-cilindro, etc.

Con un procesamiento de señal suficientemente rápido es posible realizar mediciones en flujos supersónicos. La dificultad más importante

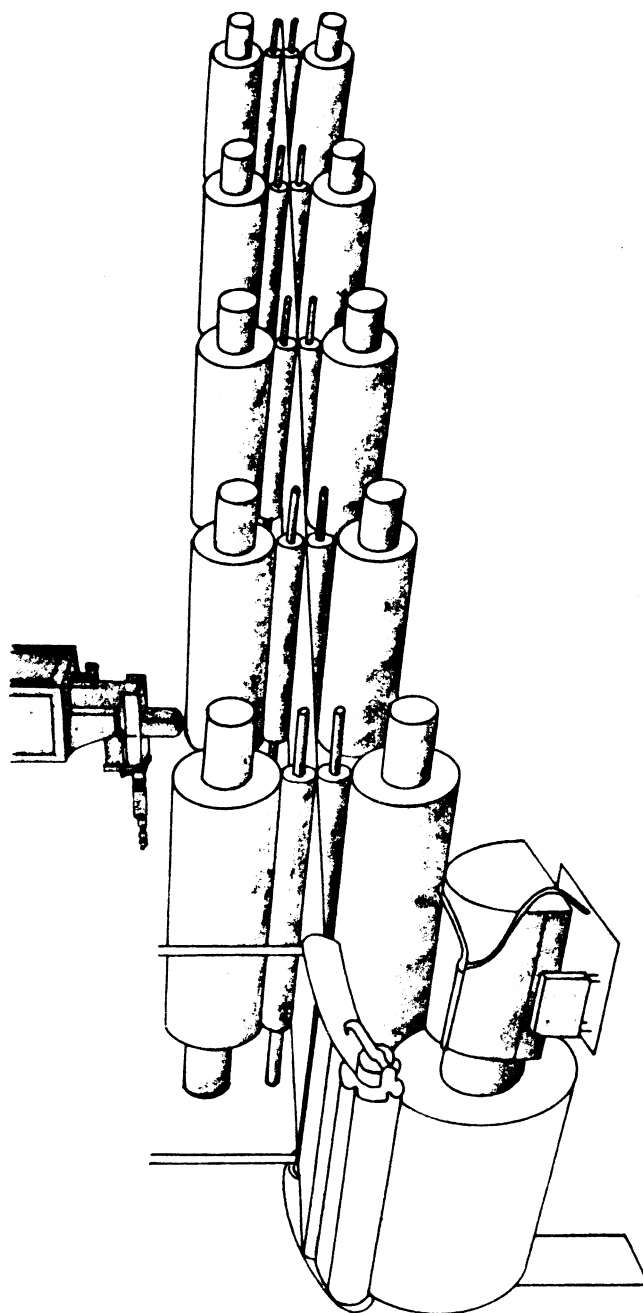


Figura 13.

proviene de la necesidad de contar con partículas suficientemente pequeñas para que puedan seguir el flujo sin distorsión y aún dispersen la cantidad de luz requerida. Generalmente se utilizan en este campo láseres de argón con hasta 2 W de potencia y procesadores tipo correlador que operan a altas frecuencias con una cantidad mínima de luz.

El estudio de las turbomáquinas constituye otro campo donde se han empleado satisfactoriamente los velocímetros láser. Se han realizado mediciones en compresores, en motores de aviación y en rotores de helicópteros. En algunos casos los resultados se han comparado con los obtenidos mediante anemómetros de hilo caliente y tubos pitot, brindando valores para las velocidades medias dentro del 2 al 3% unas de otras. Esto justifica el uso de los velocímetros láser donde los métodos convencionales no pueden utilizarse.

Otro campo donde la aplicación de los velocímetros láser es cada vez más frecuente es en el estudio de los sistemas de combustión y en los dispositivos de procesamiento térmico a elevadas temperaturas.

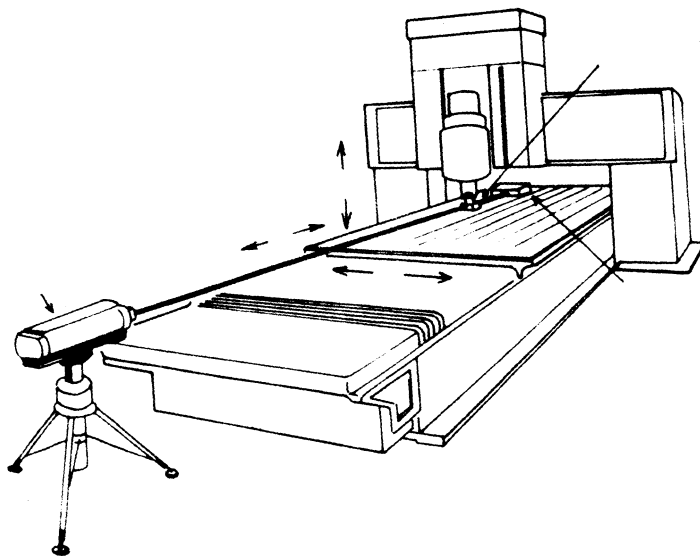


Figura 14.

En los sistemas que operan en una sola dimensión en sólidos la precisión en la medición de la velocidad es suficiente para alcanzar buenas mediciones de distancia por integración. En su versión más simplificada estos sistemas se están utilizando para controlar la longitud de productos siderúrgicos elaborados en continuo (Figura 13). En la Figura 14 se representa un sistema de elevada complejidad desarrollado por la empresa Hewlett-Packard para comandar el funcionamiento de una máquina herramienta. Este equipo combina el efecto Doppler con la interferometría para alcanzar una elevada precisión.

Bibliografía

- 1 - Yeh, Y. y Cummins, H.Z., Appl. Phys. Letters. 4, 176, 1964.
- 2 - Schodl, R., ASME paper 74-GT-157, 1974.
- 3 - Jading, L., Danish Atomic Energy Commission, Report M-1368, 1971.
- 4 - Flow Lines, The Fluid Mechanics Magazine from TSI Inc., p.4, fall 1986.
- 5 - Pike, E.R., J. Physics D. Appl. Physics, 5, Letters to the Editor, 1972.
- 6 - "Principles and practice of laser-Doppler anemometry", F. Durst, A. Melling y J.H. Whitelaw, 2nd. ed., Academic Press, London (1981).

APLICACIONES DE LASERES EN PROCESOS QUIMICOS

J. O. Tocho

Introducción

La utilización del láser en química provee medios para iniciar y controlar selectivamente reacciones químicas. Estas reacciones responden a una variedad de estímulos tales como temperatura, presión, concentraciones relativas, excitación vibracional y electrónica de las distintas especies. Utilizando láseres, los químicos pueden modificar algunos de estos parámetros de manera altamente selectiva para alcanzar un mejor control sobre las reacciones químicas.

En este capítulo nos concentraremos en la descripción de aquellas aplicaciones del láser que conduzcan al mejor manejo de reacciones químicas. Queda relegada para otro momento la discusión del uso de la espectroscopía láser en el diagnóstico de reacciones, el análisis químico y el control de procesos. Es en este área donde quedan pocas dudas de que el uso de las técnicas láser poseen una capacidad única en lo que respecta a sensibilidad y resolución espacial y temporal.

El programa de este capítulo es el siguiente: veremos en principio qué tipos de reacciones y procesos químicos pueden atacarse mediante el uso de láseres y vislumbraremos qué resultados novedosos pueden esperarse. Discutiremos posteriormente de manera somera la factibilidad de estas técnicas en el contexto de la industria química actual. En base a este análisis veremos que sólo un pequeño número de procesos quedan como candidatos para su comercialización actual o futura. Por último, estudiaremos con un poco más de detalle algunas reacciones que ya han sido probadas como exitosas.

Fotoquímica láser

Las reacciones activadas por absorción de radiación incluyen como primera etapa la producción de especies vibracionales excitadas, electrónicamente excitadas o bien de especies iónicas o radicales libres. Para producir excitaciones vibracionales normalmente se utilizan láseres infrarrojos. La excitación se produce por la absorción de un solo fotón IR, la absorción secuenciada de varios fotones IR o bien la absorción simultánea de muchos fotones IR coherentes (absorción multifotónica). Todos estos procesos producen estados vibracionales excitados en el estado electrónico fundamental. El uso de luz láser visible o ultravioleta (UV) produce normalmente moléculas excitadas electrónicamente que posteriormente pueden reaccionar químicamente. Por último, radiación láser intensa, normalmente pulsada, ya sea IR, visible o UV puede utilizarse para producir especies iónicas o radicales libres, ya sea por absorción normal o multifotónica, o bien por rotura dieléctrica por campo.

Los procesos de excitación electrónica e ionización con radiación visible o UV son ampliamente conocidos y si bien existen numerosas vías de explotarlos con láseres para alcanzar resultados químicos novedosos, el mecanismo de excitación no difiere del que se obtiene con lámparas comunes y no será discutido en detalle. Sí mencionaremos en la última parte varios procesos que utilizando láseres hacen uso de estos mecanismos. Por el otro lado, el uso de fuentes IR en química es muy reciente y sólo fue posible luego del desarrollo de láseres IR de alta potencia.

Una molécula sometida a un pulso intenso de radiación IR puede absorber muchos fotones y excitarse vibracionalmente y sin embargo, permanecer "fría" en términos rotacionales y traslacionales. Esta es una diferencia fundamental con respecto al calentamiento convencional donde la energía se distribuye uniformemente. Una evidencia de esta diferencia es el hecho de que moléculas excitadas con láser pueden acumular energía, en exceso con respecto a la energía de reacción. Por ejemplo, una molécula con energía de activación de 50 Kcal/mol puede adquirir tanto como 70-100 Kcal/mol de energía vibracional antes que la velocidad de reacción pueda competir con la velocidad de excitación. Este proceso de calentamiento instantáneo resulta normalmente en un comportamiento químico completamente diferente.

Volviendo a las generalidades relacionadas con el uso de láseres en fotoquímica, debemos mencionar la cuestión de la selectividad modal. Para un químico, la mayor utilidad de un láser estaría dada por la posibilidad de irradiar una determinada ligadura o grupo funcional de una molécula compleja e inducir la reacción en ese sitio específico. Por supuesto que el problema es competir con los mecanismos por los cuales la molécula redistribuye la energía absorbida antes de reaccionar. Con la presente generación de láseres pulsados, que emiten pulsos de más de 1 ns, no

hay muchas pruebas de este fenómeno de selectividad modal. El desarrollo de láseres con pulsos de picosegundos de duración puede convertir este sueño en realidad.

No debe confundirse selectividad modal con selectividad molecular, el fenómeno por el cual la fotoquímica puede inducir la reacción de una determinada molécula en una mezcla de especies, si la primera absorbe radiación de la longitud de onda utilizada y las demás son transparentes. Este grado de selectividad puede llegar a un extremo tal que mediante el empleo de láseres pueden sintonizarse reacciones de determinadas especies isotópicas.

La economía de la química con láseres

La practicabilidad de la puesta en marcha de sistemas de producción que empleen láseres está determinada en gran medida por el impacto que estas técnicas producen en el precio por mol de producto. Para analizar este punto nos valdremos del estudio realizado por Fischer [Angew. Chem. Int. Ed. (Eng) 17 (1978), 16].

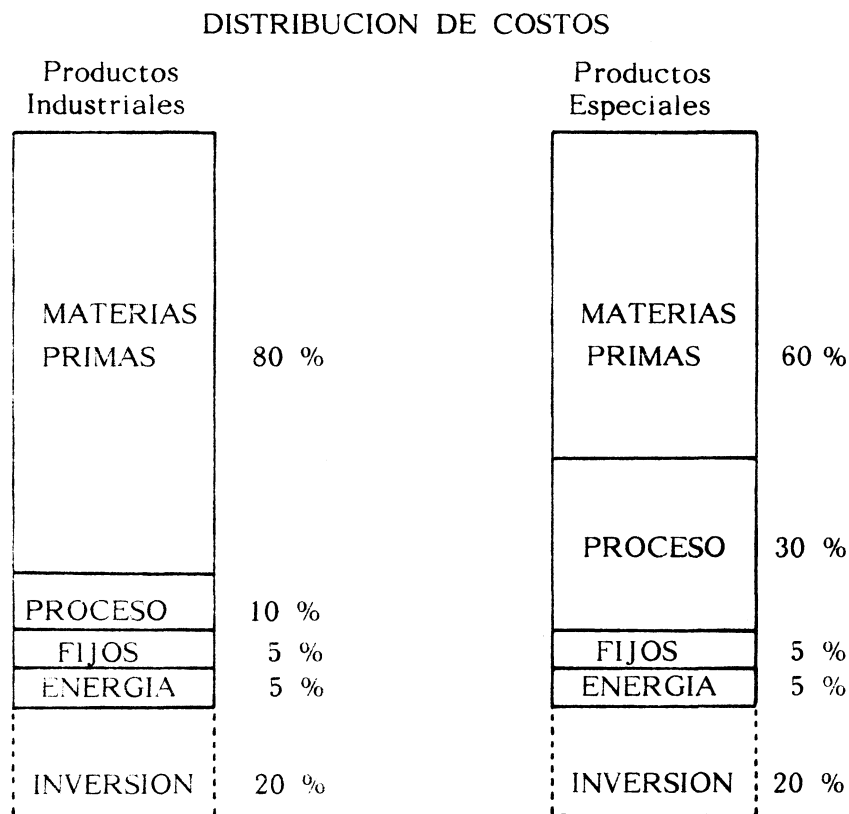


Figura 1

En el gráfico de la Figura 1 puede verse la distribución de costos para una variedad de procesos químicos englobados en dos grupos: productos industriales y productos especiales. Para los primeros, más del 80% del costo total reside en las materias primas y para el caso de los productos especiales, principalmente farmacéuticos, sólo el 30% de su alto costo reside en costos de producción.

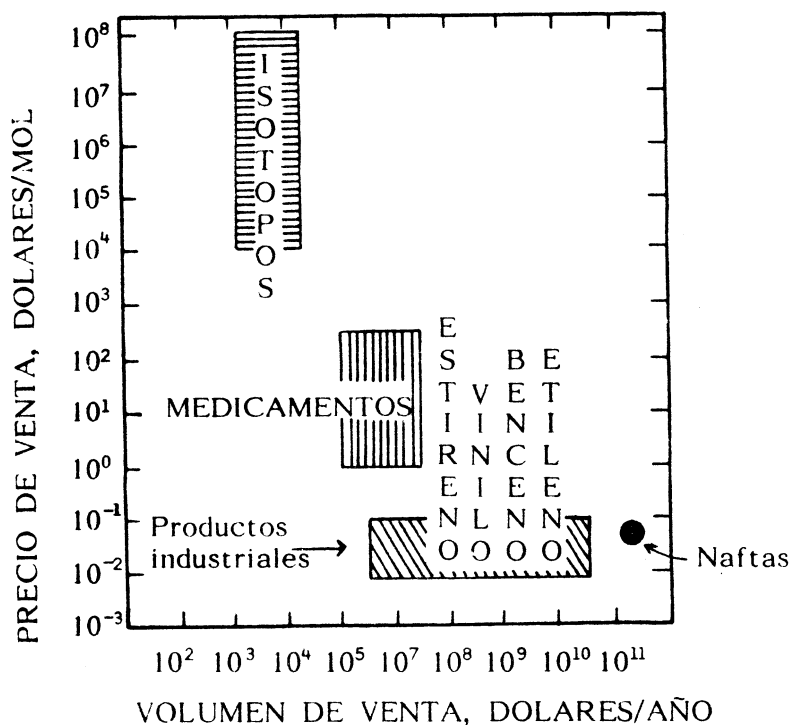


Figura 2

En la Figura 2 se muestra el precio de venta por mol de una selección de productos en función de su volumen de venta anual en Estados Unidos de Norteamérica. Puede verse que el precio de venta de la mayoría de los productos de gran volumen de producción está por debajo de \$ 1 por mol.

Por último llegamos a la Figura 3, que resulta clave para completar el análisis de la economía involucrada en esta tecnología. Aquí el costo del proceso láser por mol se grafica en función del rendimiento cuántico (moléculas producidas por fotón). La curva con una eficiencia del 2% corresponde al láser de KrF (\$ 300/W por año) y la de 6% al láser de CO_2 (\$ 100/W por año). Se calcula un uso de dos años para todos los láseres

utilizados.

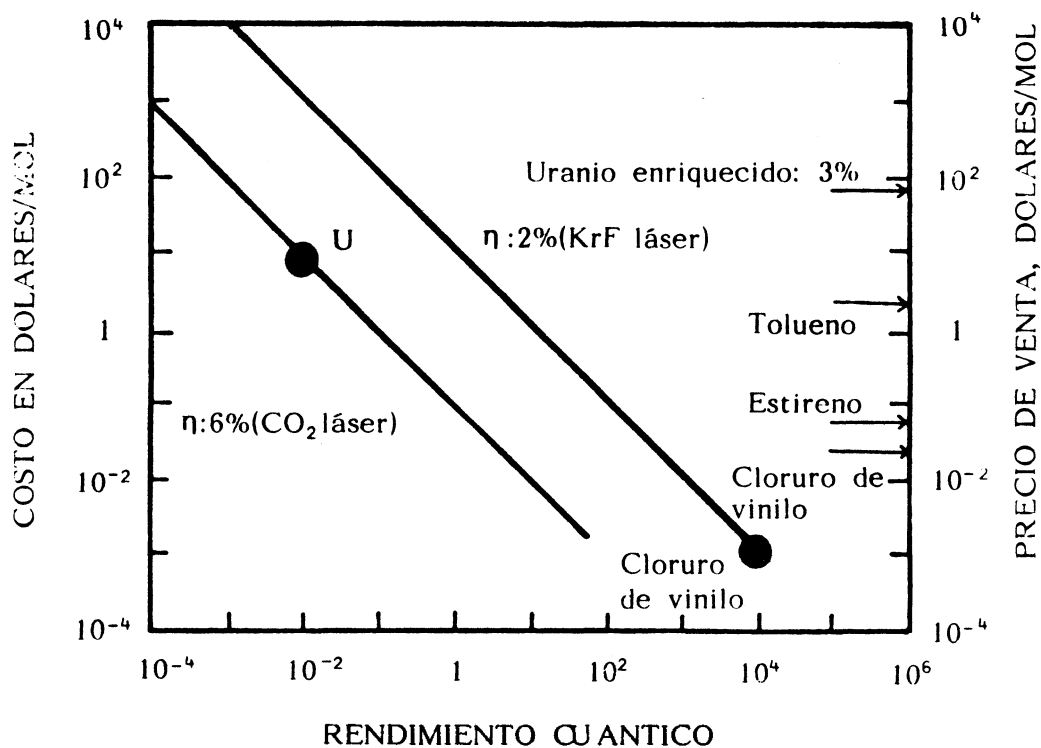


Figura 3

Resulta inmediato inferir qué procesos con bajo rendimiento cuántico tal como absorción multifotónica IR son solamente útiles cuando el proceso químico influya en gran medida en el costo del producto, tal como sucede en isótopos y materiales ultrapuros. En el otro extremo pueden verse qué procesos con altísimo rendimiento cuántico son los únicos capaces de tener algún impacto en la producción de materiales de alto consumo. En base a estas figuras es fácil comprender porqué el optimismo inicial que caracterizó este tema dio lugar a una posición más cautelosa. De todas maneras, no es todavía el momento para juzgar la química láser sólo por sus objetivos tecnológicos o económicos. Todavía se deben realizar investigaciones básicas para explorar nuevas posibilidades y desafíos.

En lo que sigue describiremos una serie de procesos ya comerciales o en vías de comercialización, que enfatizan ciertas áreas de aplicación. Un solo proceso, la separación isotópica, acaparó prácticamente todo el trabajo de la última década. Si bien en principio esto puede considerarse como una traba en el desarrollo más orgánico del tema, el volumen de información que será puesta en consideración de la comunidad científica cuando se levanten las condiciones de clasificación será tan enorme que el tiempo perdido será recuperado rápidamente.

Separación isotópica

Como ya hemos dicho, los procesos que agregan un alto costo al producto final son los únicos candidatos a ser reemplazados por procedimientos que empleen láseres de IR. La separación de isótopos de U es sin duda el tema que ha captado la mayor atención. Por lo menos tres métodos avanzados de separación utilizando láseres han sido desarrollados en los EE.UU. y uno de ellos, el propuesto por el Lawrence Livermore National Laboratory, ha sido apoyado por el gobierno para una posterior industrialización. Lamentablemente poco es lo que se conoce de estos programas y por esta razón mencionaremos solamente algunos principios básicos obtenibles en la literatura, que hacen a la separación de isótopos de U. En general, se utilizan compuestos volátiles de óxido de uranio como $\text{UO}_2(\text{h-facac})_2\text{B}$, donde h-facac es un ligando orgánico como el trimetil fosfato. Los modos vibracionales O-U-O en el compuesto volátil caen dentro del rango de frecuencias del láser de CO_2 . Esta molécula tiene un corrimiento isotópico de $0,7\text{ cm}^{-1}$ por lo que las secciones de absorción a una dada frecuencia son distintas para U^{235} y U^{238} . Para la línea P(4) del láser de CO_2 se excita preferentemente el U^{235} en un factor del orden del 30%. Posteriormente a la excitación, el complejo se disocia, liberando la base B. El proceso posterior de separación es convencional. En resumen, mediante excitación selectiva se obtiene un producto descompuesto enriquecido en U^{235} , mientras que el complejo inalterado se enriquece en U^{238} . La ventaja de un proceso multifotónico reside en el hecho de su capacidad multiplicadora del efecto selectivo. Cada vez que se absorbe un fotón la selectividad hace que en mayor proporción se excite una determinada especie.

El mercado de isótopos distintos al U es muy pequeño, y todavía no se vislumbran muchas posibilidades de extender estas técnicas a otros elementos. Pero, sin ninguna duda, el desarrollo de estas técnicas ayudará a la introducción del láser en la ingeniería de procesos químicos.

Purificación

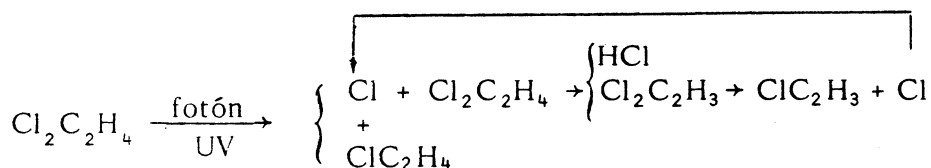
Como la separación isotópica, la purificación mediante láseres apunta a la producción de materiales con alto valor agregado. Como la purificación incluye la separación de distintas especies químicas, puede esperarse que se alcancen más fácilmente altas selectividades. Un objetivo de los sistemas de purificación con láser es que éste actúe sólo sobre las impurezas, dejando el material mayoritario inalterado, evitando así el manejo de grandes cantidades de material. Varios esquemas de purificación utilizando láseres IR y UV han sido propuestos y las estimaciones económicas resultan alentadoras especialmente en materiales utilizados en electrónica.

Chen y Borzileri [IEEE J.Q.E. 16 (1980), 1229] han demostrado la purificación de gas sintético (CO y H_2) utilizando como materia prima en diversos procesos de obtención de combustibles sintéticos. La impureza es aquí el H_2S que en proporciones mayores a 1 ppm envenena los catalizadores. La eficiente remoción del H_2S se basa en la alta absorción del H_2S relativa a los otros componentes del gas sintético cuando la mezcla se ilumina con láseres UV. Un esquema similar fue utilizado por Clark y Anderson [Appl. Phys. Lett. 32 (1978), 46] para purificar silano, un material clave en la fabricación de semiconductores de alta calidad. Luego de 6000 pulsos de un láser de ArF el contenido de fosfina y arsina se reduce a menos de 4 ppm y una posterior irradiación permite alcanzar valores de impurezas menores a 0,5 ppm.

Desde un punto de vista completamente diferente, el uso de láseres reduce el costo de procesos de purificación en la fabricación de vitamina D. El método fotoquímico convencional produce un producto con sólo 20% de vitamina D que exige un costoso proceso de purificación posterior. El uso de un láser UV con una longitud de onda optimizada produce un producto con más del 80% de vitamina, que hace innecesaria la purificación [J. Amer. Chem. Soc. 103 (1981) 6781].

Reacciones en cadena

Muchos procesos utilizados en la producción de materiales de alto consumo, como cloruro de vinilo, estireno o etileno, involucran mecanismos de reacciones con radicales libres. En los procesos térmicos normales se necesita calor para mantener la reacción en cadena, que es normalmente endotérmica. En muchos casos es posible reducir la temperatura del proceso si los radicales se generan fotoquímicamente, disminuyendo considerablemente las reacciones secundarias. Este concepto fue demostrado por primera vez por Wolfrum et al [Patente alemana 293853 (1981)], quienes obtuvieron resultados realmente interesantes en la producción de monómero de cloruro de vinilo a partir del dicloroetano. Mientras que el proceso convencional a 500°C y 10 atmósferas de presión tiene un rendimiento del 60% y una selectividad del 85%, el uso del láser de XeCl permitió disminuir la temperatura a 350°C y, manteniendo el mismo rendimiento por paso, incrementar la selectividad a más del 99,9%. El rendimiento cuántico de esta reacción es extremadamente alto, más de 10^4 moléculas del monómero producidas por fotón absorbido. Esta técnica está siendo instrumentada en Alemania en escala semiindustrial. El mecanismo de la reacción es el siguiente:



EL LASER EN ESTUDIOS DE LLAMAS

J. O. Tocho

Introducción

Numerosos métodos de diagnóstico basados en el uso del láser han sido desarrollados para estudiar procesos de combustión. Las medidas obtenidas para las concentraciones de distintas especies involucradas y las temperaturas de las llamas muestran excelentes acuerdos con cálculos teóricos, lo que hace suponer que estas técnicas representan una contribución importantísima para el mejor conocimiento de los fenómenos de combustión.

La llama comprende una interacción realmente compleja de numerosas reacciones químicas, movimiento de fluidos, transferencia de calor y difusión de especies atómicas y moleculares. Sólo recientemente se dispone de modelos matemáticos para describir y predecir el curso de estos efectos y el uso de tecnología láser brinda nuevos métodos para medir precisamente parámetros como concentraciones, velocidades de fluidos y tamaño de partículas en llamas, que son inestimables o valiosos para la comprobación de dichos modelos y teorías.

La naturaleza no perturbadora de las técnicas láser es una de las mayores ventajas sobre otras tecnologías. Se eliminará así la alteración que provocaba anteriormente la inclusión de diversas sondas como termocuplas o tubos de muestreo de gases, en las zonas bajo estudio. Dichas sondas alteran apreciablemente el flujo y/o la química de algunas llamas. Otra propiedad importante es que estas técnicas permiten alcanzar una elevada resolución espacial y realizar mediciones casi puntuadas, ya sea focalizando el haz del láser o cruzando dos o más haces focalizados para definir un pequeño volumen de interacción. Alternativamente, otro tipo de experimentos láser hacen uso de haces expandidos para iluminar uniformemente grandes zonas y seguir la dinámica macroscópica de la llama. Tales métodos incluyen la holografía, fotografía schlieren y la interferometría pulsada, las que serán estudiadas en otro capítulo de este curso.

Es nuestro propósito aquí compendiar brevemente diversas técnicas espectroscópicas capaces de medir temperaturas y concentraciones. Espectroscopías pulsadas permiten alcanzar no sólo resolución espacial sino que también proveen una descripción dinámica con elevada resolución temporal.

En todas las técnicas espectroscópicas se analizan las frecuencias (o longitudes de onda) de la luz dispersada, absorbida o emitida cuando la zona bajo estudio se ilumina con uno o más láseres. Este espectro está determinado por el o los láseres utilizados y de manera característica, como lo es la firma o la impresión digital de un ser humano, por las especies atómicas o moleculares presentes. La intensidad de la radiación estudiada está relacionada con la concentración de dichas especies y el estudio de la intensidad relativa de distintas frecuencias que provienen de la interacción con distintos estados cuánticos de una misma especie permite inferir la temperatura a que ésta se encuentra. Si la especie molecular no se encuentra en equilibrio térmico, pueden determinarse sus temperaturas vibracionales, rotacionales y traslacionales.

Scattering Raman espontáneo y estimulado

Una de las primeras espectroscopías usadas en el diagnóstico de la combustión fue la dispersión Raman espontánea (SRS). En un experimento típico de SRS un láser de frecuencia fija se focaliza en la región de interés y esta región se utiliza como objeto de un sistema óptico que está enfocado sobre la ranura de un espectrómetro. La luz Raman dispersada se analiza con un fotomultiplicador o un OMA (optical multichannel analyzer) en frecuencias diferentes a la del láser incidente. La aparición de distintas frecuencias corridas hacia uno y otro lado (frecuencias stokes y antistokes) de la frecuencia de excitación puede extenderse en base a un modelo clásico de scattering inelástico predicho por primera vez en 1923 por Swekal y observado algunos años más tarde por Raman y Krishman en líquidos y por Landsberg y Mandelstam en cuarzo. Según este modelo, cuando un haz luminoso monocromático atraviesa una colección de átomos o moléculas induce dipolos oscilantes que reirradian en la frecuencia incidente y en varias frecuencias laterales. Estas bandas laterales están separadas de la incidente por frecuencias electrónicas, vibracionales o rotacionales que son características de la especie que dispersa la radiación incidente. Sin embargo, este modelo clásico no es capaz de explicar ciertos aspectos, como intensidades y reglas de selección, que sólo son predichos por una teoría mecano-cuántica de la interacción radiación-materia.

Las especies más extensamente estudiadas por este método son N_2 , O_2 , y oxígeno atómico. Una de las formas más utilizadas para determinar temperatura en llamas es el estudio de la forma de la rama vibracional Q del espectro Raman del N_2 . Han habido numerosos estudios del SRS

del N_2 para determinar la distribución de temperatura de diversas llamas premezcladas, como $C_3H_8 - O_2 -$ aire, $CH_4 -$ aire, $NH_3 - O_2$ y de H_2 difundiéndose en aire. Otra utilidad del SRS es la capacidad de detectar átomos libres. Fue observado recientemente el SRS del oxígeno atómico formado en una llama premezclada de $H_2 - O_2$ donde el O fue detectado en una fracción no mayor del 1%.

Sin embargo, el SRS tiene severas limitaciones en el estudio de llamas. Una seria dificultad constituye la señal relativamente pequeña del SRS que hace que cualquier luminosidad proveniente de la llama se presente como una fuente de ruido intenso. Este efecto hace imposible el uso de los láseres continuos en SRS, ya que son necesarios láseres pulsados de alta potencia para producir señales SRS comparables a la luminosidad espúrea. Pero otro importante efecto ocurre entonces, ya que pequeñas partículas comúnmente en las llamas pueden llegar a la incandescencia al absorber estos intensos pulsos láser y emitir radiación de cuerpo negro que enmascara completamente la señal SRS.

Muchos de los problemas de SRS pueden resolverse mediante las espectroscopías Raman coherentes o estimuladas. Aquí, si bien el equipamiento es mucho más complicado y la interpretación de los resultados un poco más compleja, la ventaja reside en la coherencia de la señal emitida (direccional) y la mayor señal, lo que en definitiva incrementa enormemente la relación señal-ruido con respecto al SRS. En estudios de combustión, la espectroscopía Raman coherente más utilizada es la antistokes o CARS (Coherent Anti-stokes Raman Scattering). El CARS se realiza con dos láseres de frecuencias ω_1 y ω_2 de manera que la diferencia $2\omega_1 - \omega_2 = \omega_4$ resuena con alguna frecuencia Raman característica del medio que está siendo estudiado. Usualmente los experimentos de CARS se realizan con tres haces de excitación (dos de la misma frecuencia) y la dirección de propagación de la señal está dada por la conservación del momento o condición de ajuste de fase, lo que hace que las propagaciones generalmente no tengan que ser colineales ni aún estar en el mismo plano.

Teets y Bechtel han determinado oxígeno atómico por primera vez mediante CARS y el experimento mostró una resolución diez veces mayor que la obtenida mediante SRS. Otras especies, como CO en llamas premezcladas de $CH_4 - CO -$ aire, O_2 y N_2 así también como temperaturas en varios tipos de llamas han sido determinadas mediante CARS. Si uno tiene que monitorear especies muy diluidas, como ser algunas partes por millón de volumen, deben seleccionarse otros métodos, ya que ni SRS ni CARS proveen la sensibilidad necesaria. Esos tipos de espectroscopías serán mencionados brevemente en lo que sigue.

Un método de elevada sensibilidad, utilizado especialmente para detectar radicales inestables en llamas, es la fluorescencia inducida por láser (LIF). Este método ha permitido detectar especies en cantidades del orden de los 10^4 átomos/cm³ o sea concentraciones 10^{-17} M. La técnica de LIF es extremadamente simple en principio pero resulta esencial para llegar a determinar concentraciones como las mencionadas, el uso de láseres específicos sintonizables y con reducidísimo ancho de línea. El fun-

damento de LIF consiste en iluminar la zona de la llama de interés con un láser sintonizable, excitar niveles discretos y característicos de la especie relevante y recibir la fluorescencia inducida con un fotomultiplicador, previo paso a través de ciertos filtros interferenciales que seleccionan la zona espectral. Resulta obvio que la intensidad de la fluorescencia estará directamente relacionada a la sensibilidad final del método. Esa intensidad está gobernada por la eficiencia en la excitación de los niveles superiores fluorescentes. De ahí la importancia en la sintonizabilidad del láser. Luego de producido el enganche o resonancia entre el láser y la especie, es importante que toda la intensidad del láser se concentre en un reducido ancho espectral alrededor de la frecuencia de resonancia de la especie, ya que todo lo que quede fuera del ancho propio de la transición será perdido, y peor aún, provocará fluorescencia de otros niveles de esa especie o aún de otras especies.

Los láseres que cumplen las condiciones impuestas por esta técnica y por lo tanto son más comúnmente utilizados en LIF son los láseres de colorantes en la región ultravioleta-visible y los láseres de semiconductores en el infrarrojo. Además de las mencionadas, una gran variedad de técnicas espectroscópicas han sido utilizadas en estudios de combustión, entre ellas, dispersión Rayleigh, espectroscopía optogalvánica y optoacústica y fluorescencia por absorción multifotónica.

Sólo nos detendremos brevemente para mencionar el principio del efecto optogalvánico. Esta técnica se basa en la medida de la corriente eléctrica que circula entre dos electrodos colocados en una porción de la llama cuando ésta se ilumina con un láser sintonizable. La corriente se produce cuando la radiación alcanza a ionizar una especie atómica o molecular presente en la llama o más indirectamente, mediante una ionización por choque luego de una previa excitación óptica. Este método es capaz de alcanzar sensibilidades comparables a las del método LIF y presenta la ventaja sobre otras espectroscopías más convencionales en que la detección es directamente electrónica, por lo que se evitan interferencias provenientes de radiación de fondo o dispersada, comunes a los otros métodos. En el caso particular de las llamas, una desventaja es el hecho de tener que colocar los electrodos que miden la corriente inducida en la cercanía de la llama, alterando la estructura de ésta.

Como conclusión final diremos que los distintos métodos ópticos discutidos aquí encuentran cada vez más cabida en el estudio de procesos de combustión. Las especies determinadas en llamas incluyen: CH_4 , C_3H_8 , NH_3 , NO , O_2 , O , N_2 , H_2O , CO , CO_2 , H_2 y HO , en concentraciones desde 100 ppm hasta pocas partes por millón. La mayoría de los experimentos han sido desarrollados en llamas controladas de laboratorio y algunos pocos sobre motores de combustión interna y turbinas.

PROCESAMIENTO DE MATERIALES CON LASER DE CO₂

E. Gallego Lluesma

1. Introducción

Una descripción de la Física del láser ha sido dada por el Profesor Garavaglia [1], por lo que no será repetida aquí. Los conceptos básicos serán dados por comprendidos, para facilitar la escritura de la temática específica del presente capítulo. Algunos tipos de láseres han sido descritos, en particular, más detalladamente en este libro [2], por lo que, en lo que respecta al láser de CO₂ sólo se proveerán detalles adicionales a los descritos allí. Se recomienda en consecuencia, la lectura previa de estas dos referencias.

El láser de CO₂ ha sido desde su invención por Patel [3], uno de los láseres que más fuera aplicado industrialmente. Desde 1964 hasta mediados de la década del setenta, las potencias obtenibles en 10,6 micrones de longitud de onda, en el lejano infrarrojo, fueron aumentando en forma paulatina. Las aplicaciones pioneras fueron mayormente espectroscópicas y de investigación básica en materiales. A tal punto, que en el primer libro publicado sobre los láseres, el Dr. Bela Lengyel [4] en 1966, incluye al láser de CO₂ en un pequeño párrafo, en "otros láseres". Informes más completos de la excitación del CO₂ fueron publicados por Patel [5,6] en relación a las descargas pulsadas. En 1964 Legay y Legay-Sommaire [7] informaron que podría conseguirse una generación de luz coherente mucho más eficiente, basada en la transferencia de la energía vibracional del N₂ al CO₂. Un láser basado en este principio fue construido por Patel ese mismo año [8]. Otro avance significativo ocurrió cuando Patel descubrió que el agregado de He permitía obtener potencias de emisión mayores que 100 vatios. En poco más de un año se pasó de láseres de CO₂ emitiendo con 1 mW, a láseres que emitían 100 W. Las eficiencias de conversión electrón-fotón pasaron en el mismo período de 10⁻⁶ al 6 por ciento. En 1969, Deutsch y colaboradores [9] y Lavarini y colaboradores [10], afianzaron la tecnología de flujo axial rápido, que permitió la obtención de 1,4 kW por metro de medio activo pa-

ra una presión de 70 Torr y cavidad de un metro. En 1970, Beaulieu [11] informó sobre la emisión a presión atmosférica y flujo transversal, marcando una nueva generación de láseres de potencia. Sin embargo en la última década, mediante el avance en técnicas electrónicas para la producción de las descargas de excitación y diversas geometrías de flujo de gas, se ha hecho posible un aumento considerable desde el punto de vista industrial de las potencias utilizables, aumentando además la fiabilidad de los sistemas, lo que hizo que fueran disponibles comercialmente un sinnúmero de opciones. Tradicionalmente, en la tecnología de flujo axial se obtenían potencias del orden de 50 vatios de radiación láser por metro activo. Con la introducción de los láseres de flujo axial rápido, en condiciones de excitación semejantes, se obtuvieron potencias de 1000 vatios por metro activo en forma comercial, con gran estabilidad de modos. La más reciente configuración de flujo transversal puede proveer hasta 10 kilovatios de potencia por metro activo.

La óptica de infrarrojo mejoró también notablemente merced a los esfuerzos realizados por los proveedores de los programas de fusión nuclear, para la concreción de la obtención de altas densidades de potencia focalizadas.

No haremos un intento por repetir en este capítulo conceptos ya vertidos en los anteriores, pero apuntaremos en lo factible detalles importantes de los cuales dependen las prestaciones de los procesos que describiremos. Nuestro trabajo no intenta ser exhaustivo sino demostrativo de las aplicaciones más desarrolladas y rentables industrialmente. Compararemos en lo posible las tecnologías empleando láser con las tecnologías convencionales o de punta que compiten con ellas, donde de ser procedente haremos referencia a nuestro entorno tecnológico, y a los nichos tecnológicos que ocupan actualmente.

Declaramos que, cualquier intento de ser objetivo puede quedar trunco debido a que el autor se declara entusiasta de los láseres y sus aplicaciones.

2. Sistemas de Procesamiento

Desde el punto de vista industrial, un láser no es aplicable si no está incorporado a un sistema que permita su uso específico. Los sistemas son bastante diferentes según el uso final requerido. Estos sistemas suelen compararse por el inversor en forma muy conservadora. Sus referencias más importantes son cuántos se han instalado con anterioridad al que él pudiera comprar y cuáles son las prestaciones, gastos de funcionamiento, también llamados operativos, y cuál será el régimen de amortización del capital invertido en función de su expectativa de mercado de aplicabilidad.

Un sistema dedicado al corte de metales difiere muchísimo de uno dedicado a la deposición química de vapores o al maquinado de cerámicas. Pero sin entrar en detalles específicos podríamos describir lo que sería un sistema de investigación y desarrollo en el procesamiento de materiales, lo suficientemente versátil como para aplicaciones muy diversas. Lógicamente esta versatilidad lo haría de baja prestación en alguna de las aplicaciones.

El sistema de procesamiento láser, incluye además de este último, una mesa computada de alta precisión. Sensores y sus correspondientes circuitos de lectura por medio de una interfaz a la computadora del sistema. Posicionamientos típicos son del orden de 3 a 15 micrones por metro de desplazamiento, según la configuración mecánica del sistema. Las reubicaciones, esto es, el posicionamiento en un punto de coordenadas luego de haberse desplazado de él un largo trecho, tienen precisiones entre 10 y 50 micrones. Estas precisiones son necesarias para sacar ventaja del tamaño del punto focal del láser en su límite de difracción [12]. Este diámetro límite mínimo puede ser de unos 50 micrones, dependiendo de la óptica y de los modos de oscilación del láser, como veremos más adelante. La mesa posee generalmente dos movimientos coordinados x e y , la coordenada z es por lo general destinada a la focalización del haz. La lente de salida posee por lo general un sistema de seguimiento de aproximación de la superficie, siendo los más comunes ópticos, capacitivos o mecánicos.

En la Figura 1 podemos ver la descripción realizada.

La repetibilidad del procesamiento está condicionada por las propiedades ópticas de la superficie a tratar, específicas para el acoplamiento luz-material. En general los metales son muy reflectores de la radiación del infrarrojo lejano (más de 7 micrones de longitud de onda). Estas Propiedades Ópticas serán detalladas en la sección 6.

Otro parámetro importante es el tiempo de interacción de la radiación con el material. Los láseres poseen capacidad de descargar muchos joules de energía en tiempos muy cortos; al mismo tiempo poseen capacidad de focalización con altas densidades de potencia (vatios por centímetro cuadrado) a lo que hay que adicionar altas irradiancias (joules por cm^2) y la flexibilidad del manejo de la radiación como si fuera luz (modulación, reflexión, etc.).

El último tópico a tocar de un sistema de procesamiento láser es el software. Este es de dos clases, el primero general del tipo CAD/CAM y el segundo específico que es capaz del control en tiempo real del procesamiento con interfaz a la máquina. Este tipo de soft es en general del tipo conversacional y los Controles numéricos computados (CNC) no requieren que el operador sea un experto en soft. El soft es en general auto-instructivo. Posee una rutina de auto diagnóstico y un archivo de piezas y procesos ya realizados que pueden evocarse. Puede controlar el seguimiento de formas intrincadas, controlar el flujo de gases (del medio activo

PROCESOS

Corte
Perforado
Endurecimiento
Enchapado
Soldadura
LCVD

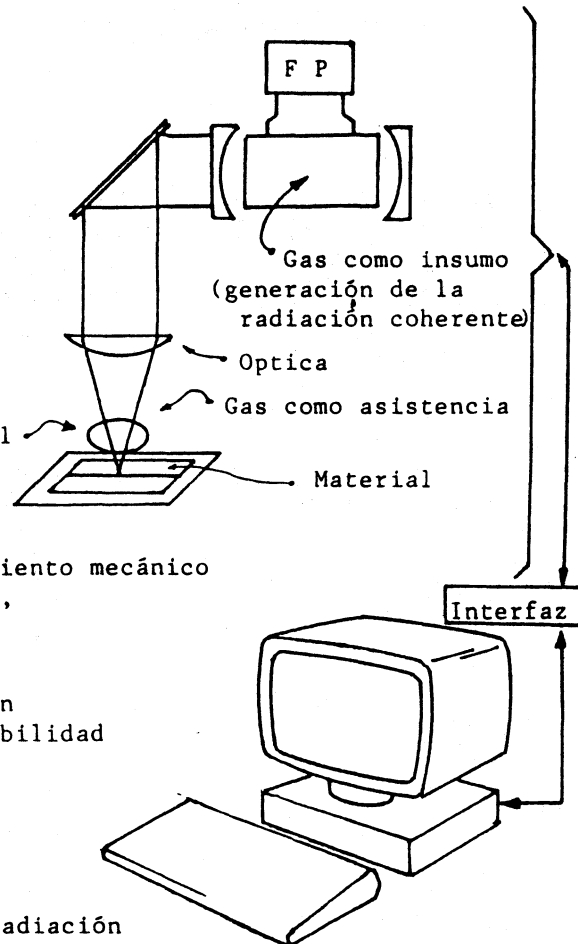
Mecanismo de
acoplamiento
radiación-material

Posicionamiento mecánico
x,y,z, ,

{
Ubicación
avance y
reubicación
con repetibilidad

Software

Formas
Intensidad de radiación
Flujo de gases
Regímenes de operación del láser
Velocidad de la mesa
Focalización
Posicionamiento
CAD/CAM
Modelado



F P: Fuente de Potencia

Figura 1: Sistema de procesamiento láser para desarrollo y producción mediante CAD/CAM.

y de aporte al proceso que se realiza), escoger el modo de operación (continuo, pulsado, híbrido), seleccionar la velocidad más adecuada de la mesa, focalizar el haz, localizar una coordenada de referencia y reposicionar. Los sistemas más sofisticados toman datos del proceso que se realiza para su posterior modelado y/o simulación. Los sistemas más convencionales, de menos precisión y dedicados al corte solamente de laminados, poseen lectores pantográficos. En este último caso, la precisión de corte es aproximadamente de 0,5 mm. Este tipo de sistemas requieren del manejo de planos de papel y poseen un lector óptico bidimensional que desplaza el pico del láser sincrónicamente a la del lector.

Los nuevos sistemas con CAD/CAM, suelen ser híbridos, de tal suerte que desplazan la óptica del láser en una coordenada y la mesa en otra coordenada perpendicular, siendo el eje z únicamente de enfoque. Esta configuración posibilita una gran rapidez. La Figura 2 nos muestra los diferentes montajes posibles. El último de ellos es un montaje de 5 ejes, instrumentando la posibilidad de procesamiento espacial. En la Figura 3 se muestra cómo es el sistema óptico-mecánico para este sistema robotizado.

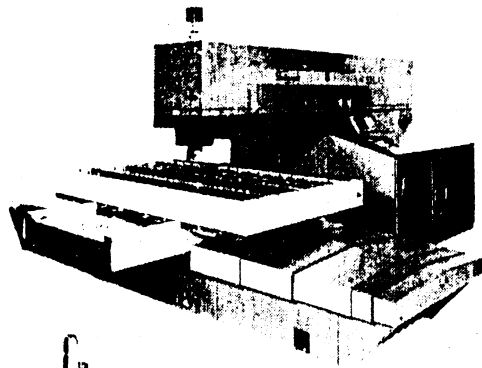
3. Modos del Láser

Hemos visto que la amplificación de luz dentro de la cavidad, se realiza merced a las múltiples reflexiones que ocurren en la cavidad. Esta cavidad para espejos planos, se comporta entonces como un Fabry-Perot, que ya hemos estudiado. Después de un número grande de reflexiones, considerando una onda plana generada en su interior cuya amplitud es A_0 , el campo total dentro de la cavidad es:

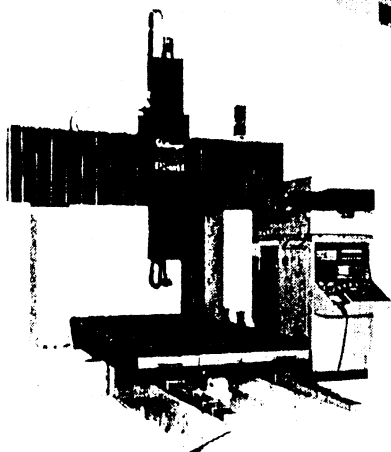
$$E = A_0 (1 + r^2 e^{-i\phi} + r^4 e^{-2i\phi} + \dots) \quad (1)$$

para una onda que viaja de izquierda a derecha, r es el coeficiente de reflexión en ambos espejos y ϕ es el cambio de fase asociado con un viaje de ida y vuelta en la cavidad ($= 2kd$). Se asume que las ondas viajan perpendiculares a los espejos. Al calcular la suma de la serie obtenida obtenemos que la intensidad I dentro de la cavidad es:

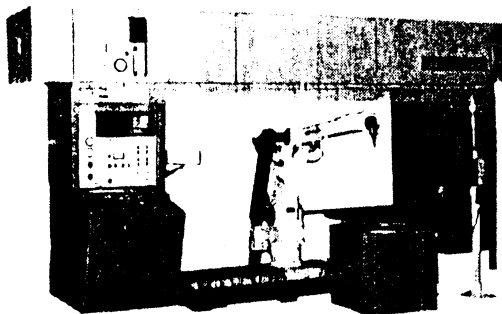
$$I = \frac{I_0}{1 + F \sin^2 \left(\frac{\phi}{2} \right)} \quad (2)$$



Sistema de tres ejes de alta precisión. CNC con capacidad de programación fuera del sistema. Gran variedad de opciones y tamaños de mesas.



Sistema de 5 ejes con CNC con capacidad de generar contornos sofisticados. Sistema híbrido, mesa sólo en x, torre sólo en y. Cabeza articulada con más grados de libertad.



Sistema robotizado multiarticulado con control del haz y gran flexibilidad.

Figura 2: Sistemas.

donde $F = 4R/(1-R)^2$ es el término que habíamos denominado "finess". R es la reflectancia. Esta expresión encontrada es la transmitancia de un interferómetro de F-P. I tiene el valor I_0 cuando $m\lambda = 2d$, la condición para que exista interferencia constructiva. Para todos los otros valores de λ , la intensidad dentro de la cavidad es pequeña, indicando que la onda 0 es absorbida o transmitida por los espejos, así que la oscilación no tiene lugar. Es conveniente considerar que el espejo de salida posee una reflectancia efectiva R_{ef} igual a R solamente cuando $m\lambda = 2d$. Por lo tanto, escribimos:

$$R_{ef} = \frac{R}{1 + F \sin^2 \left(\frac{\phi}{2} \right)} \quad (3)$$

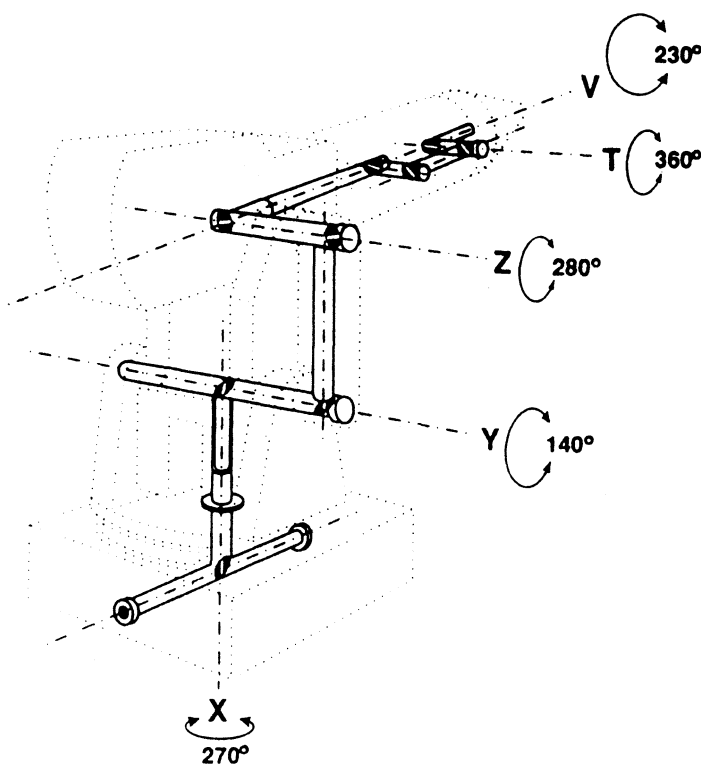


Figura 3: Sistema óptico-mecánico de un brazo articulado, robotizado de cinco ejes.

En la mayoría de los láseres, uno de los espejos posee reflectancia del 100%. Debido a que R_{ef} tiene la misma forma funcional que la transmitancia de un interferómetro de F-P, podemos adoptar todos los resultados que hemos encontrado en su discusión. Específicamente encontramos que el interferómetro transmite una serie entera de longitudes de onda separada por una cantidad igual al rango espectral libre:

$$\Delta\lambda = \lambda^2/2d, \text{ o en términos de frecuencia: } \Delta\nu = c/2d.$$

En el láser, entonces, la reflectancia efectiva es igual a R sólo a frecuencias discretas separadas por ν . La Figura 4 nos ilustra el efecto de la cavidad sobre la emisión láser. La curva superior de la figura ilustra la ganancia G como función de la frecuencia. La curva punteada es la ganancia umbral, para las cuales G^2R excede de la unidad. La curva central re-

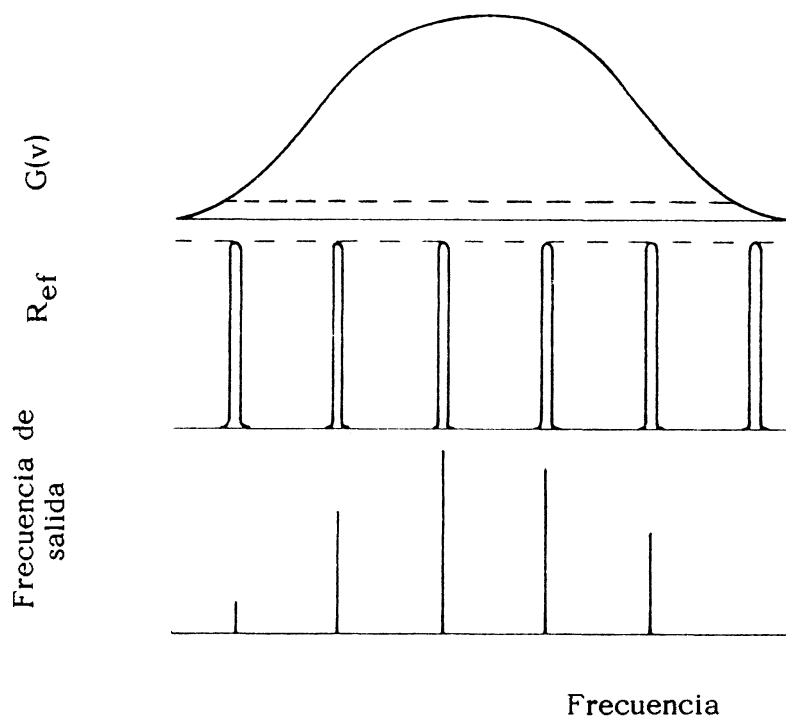


Figura 4: Oscilación en varios modos transversales.

presenta R_{ef} en función de la frecuencia. R_{ef} es grande solamente (en este ejemplo) en cinco puntos de intervalo de frecuencias dentro de los cuales G es grande. En la parte inferior se muestra la emisión láser solamente a estas cinco frecuencias discretas. Cada una de estas frecuencias es conocida como modo espectral o más comúnmente modo axial. Estos modos son en la práctica mucho más estrechos en frecuencia de lo que podría sugerir la curva de R_{ef} . El ancho de la línea espectral de emisión está dado por el gran número de reflexiones que tiene lugar antes que la ganancia supere el umbral de la emisión láser. Además, la amplificación se hace selectiva sobre la frecuencia central más que sobre los extremos del perfil que está amplificando, lo que ayuda al angostamiento de la línea. Cuando la cavidad posee ganancia, el número de idas y vueltas de los frentes de onda es muy grande y esto resulta en un espectro muy selectivo de frecuencias y con gran brillo.

También existen otros modos, denominados modos transversales. Estos se entienden mejor cuando consideramos una cavidad del tipo confocal. Esta es una cavidad que posee dos espejos idénticos con un punto focal común. Los modos transversales son los que se obtienen a partir de la distribución del campo eléctrico asociado con cualquier camino geométrico que permita realizar a la onda electromagnética un viaje de ida y regreso. Las cavidades que no permiten que exista un camino cerrado son denominadas resonadores inestables. El modo más simple que puede establecerse en una cavidad confocal es el que se forma a lo largo del eje axial del resonador. Este se denomina TEM_{00} (Transversal Electro Magnético 00). La Figura 5 nos muestra la cavidad confocal. El modo 00 es el que se conforma cuando con una geometría igual a la de la parte superior de la figura, se tiene una simetría axial y además no se dan las condiciones para la realimentación de otro modo. La distribución de intensidad de los modos puede apreciarse en la Figura 6. La difracción también contribuye para que la distribución de intensidades tengan las formas registradas en la figura. La salida de luz del láser en este modo de oscilación posee un frente de onda esférico con una distribución de intensidad gaussiana. El ancho del haz está expresado como el radio w_0 al cual la intensidad del haz cae a $1/e^2$ de su valor central máximo. El modo de oscilación transversal siguiente más simple que el 00, es el 01, sin geometría axial, como puede verse en la parte inferior de la Figura 5. Este modo solamente puede oscilar si el medio activo y los espejos poseen el radio suficientemente grande como para establecer el camino geométrico. Cuando se da la simetría axial se obtiene el modo 01* que posee forma de rosquilla. Esta figura registrada en la Figura 6, la que está un poco saturada, no permite ver la intensidad menor que posee en el centro. Los modos superiores al 00 requieren de un camino geométrico mayor y con un número mayor de reflexiones. Como norma diremos que se designan los modos con el número de mínimos que se encuentran en la figura de intensidad cuando el haz es registrado primero en forma horizontal (primer número) y luego verticalmente (segundo número). En la práctica, los modos más elevados poseen mayores pérdidas (debido a la difracción) en

comparación al modo 00. Si un láser oscila a un modo elevado, también lo hace en los órdenes menores; esto es, si un láser oscila en el modo 01 también lo hará en el modo 00. Estos láseres llamados multimodales poseen una gran potencia, y esto se realiza con una gran eficiencia, ya que el número de electrones requeridos por la fuente de potencia eléctrica para cada fotón emitido por el láser se minimiza, en la emisión multimodal. Sin embargo, la emisión monomodal es muchas veces deseada para trabajos de precisión donde se requiere minimizar el tamaño del punto focal; esto se logra prácticamente cambiando espejos o colocando diafragmas dentro de la cavidad (cuando es posible), de tal forma que sólo se permite la oscilación del modo 00.

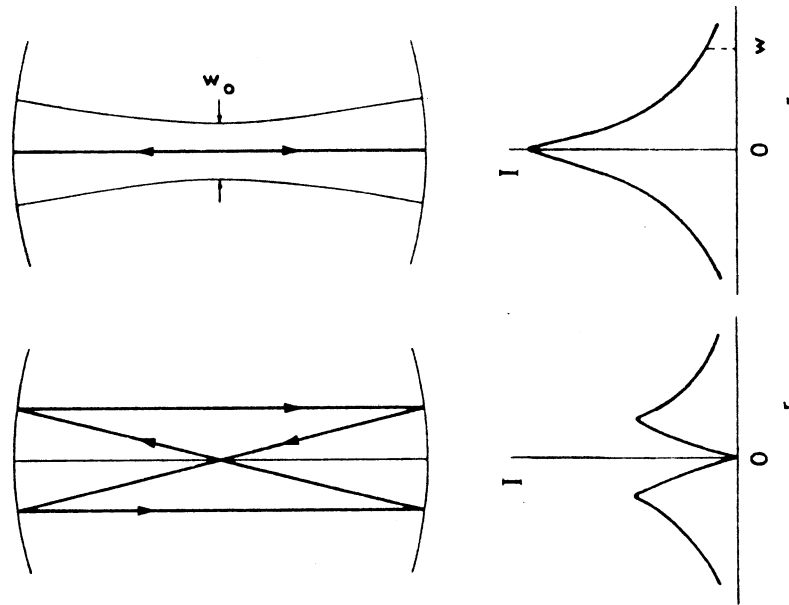


Figura 5: Modos transversales en una cavidad confocal.

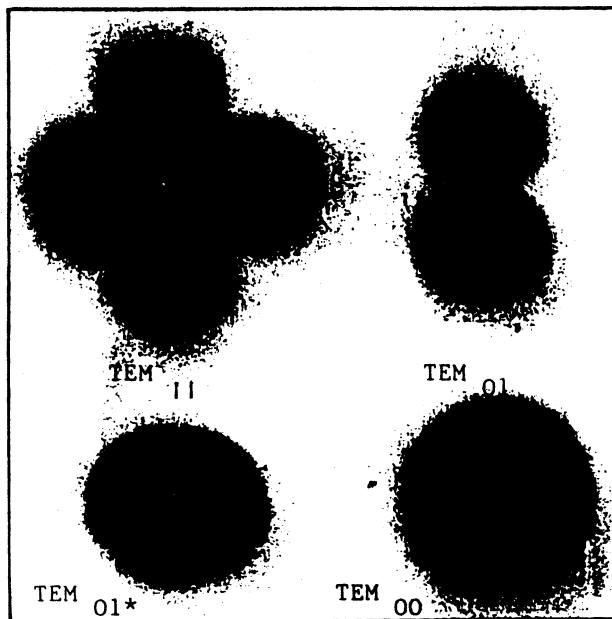


Figura 6: Intensidad registrada en placa fotográfica de los modos transversales de un láser. TEM_{00} , TEM_{01} , TEM_{11} y TEM_{01*}

Para ilustrar este último caso, la Figura 7 muestra los perfiles del frente de onda de un láser de 6 kilovatios de potencia, que ofrece la posibilidad de operación en los modos 00 , 01^* y multimodo. En el modo 00 ; se obtiene una potencia de 4 kilovatios con un diámetro del haz de 12,5 mm; en el modo 01^* , se obtiene una potencia de 5 kilovatios con un diámetro del haz de 14 mm y finalmente para la operación multimodal, una potencia de 6 kilovatios con un diámetro del haz de 16,5 mm. El modo 00 es ideal para el corte metalúrgico, el modo 01^* , el más indicado para la soldadura y el multimodal para los tratamientos térmicos. Se deduce de la práctica diaria que es necesario estar instrumentado para poder realizar un diagnóstico del perfil de onda del haz del láser, para poder controlar el procesamiento y el alineamiento del sistema. La Figura 8 nos muestra dos perfiles calculados y graficados tridimensionalmente del modo 00 y de un modo 20^* puros. La Figura 9 nos muestra un perfil de distribución de intensidad de un láser de 1,5 kW de potencia, al cual se han efectuado dos cortes para una mejor evaluación de la técnica de diagnóstico. La Figura 10 nos muestra una me-

	Modo	Máximo Nivel de potencia (kW)	Diámetro del haz de salida (mm)
	TEM ₀₀	4.0	12.5
	TEM ₀₁ *	5.0	14.0
	Multimodo	6.0	16.5

Figura 7: Prestación de un láser de potencia de acuerdo al modo de operación. Válido para un modelo específico. El diámetro del haz de salida está medido a la potencia equivalente de los límites $1/e^2$ del modo TEM₀₀, que contienen el 85% de la potencia disponible.

dida real de un haz focalizado en el modo 10*, mostrando claramente la coexistencia del modo 00 en el centro del modo 10*. Son tres los parámetros que deben controlarse con el fin de obtenerse un procesamiento reproducible:

- 1) la distribución de intensidad del frente de onda del haz;
- 2) el radio del foco utilizado; y
- 3) la divergencia del haz en el foco con su mapa de potencias [13].

Volveremos sobre este tema más adelante.

Hemos notado que cuando el láser oscila en el modo 00, la distribución de intensidad posee un perfil gaussiano. Los modos más elevados también poseen un perfil gaussiano pero multiplicados por polinomios de Hermite. El tratamiento de la difracción de una onda plana por una abertura circular conduce a una expresión para la divergencia angular del haz igual a $1,22 \lambda/D$. Pero este resultado no es satisfactorio para haces de perfil no uniforme, como es el caso de los perfiles gaussianos. La distribución del campo eléctrico para cualquier cavidad de espejos curvos posee una parte circular más angosta. En una cavidad simétrica, esta

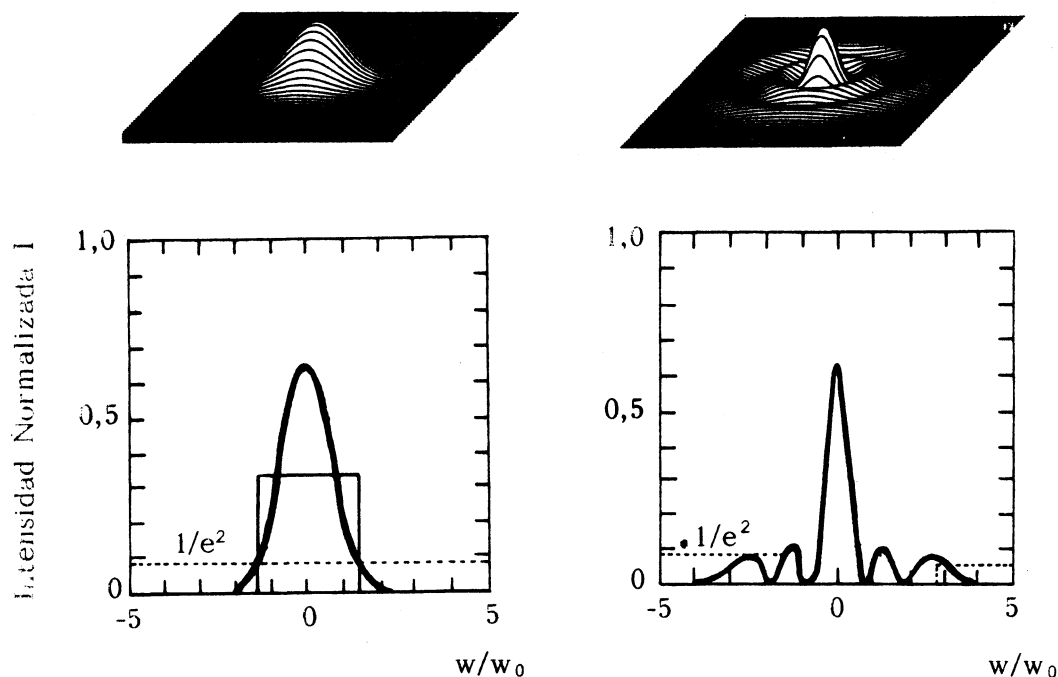


Figura 8: Definición del radio focal para dos modos diferentes. Estas definiciones carecen de practicidad.

parte está localizada en el centro de la cavidad como muestra la Figura 5. La distribución de intensidad en el plano del angostamiento es gaussiana para el modo 00, dada por:

$$I(r) = \exp(-r^2/w_0^2) \quad ,$$

donde r es la distancia desde el centro del haz. Por conveniencia la intensidad se normaliza a la unidad del centro del haz. Cuando $r = w_0$ la intensidad es igual a $1/e^2$ veces la intensidad del centro de la cavidad. Los órdenes más elevados son caracterizados por la misma distribución gaussiana multiplicada por el polinomio de Hermite apropiado para cada caso. El haz se propaga dentro y fuera de la cavidad conservando su perfil gaussiano. A una distancia z del lugar del angostamiento la intensidad de la distribución está dada por la ecuación precedente, pero reemplazando

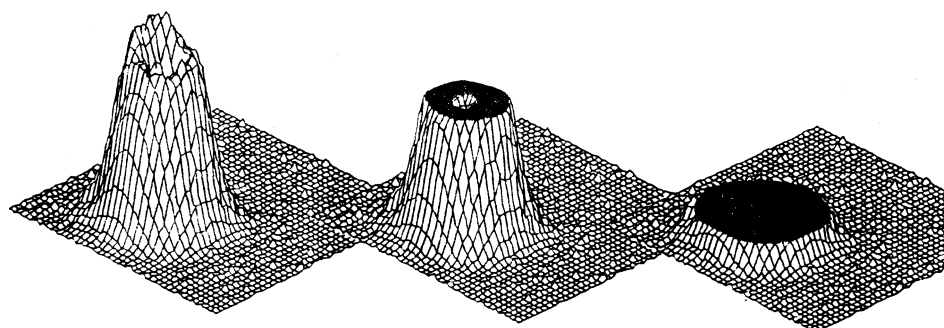


Figura 9: Distribución de la densidad de potencia de un láser que está focalizado. Los cortes transversales ilustran el carácter del interior de la distribución.

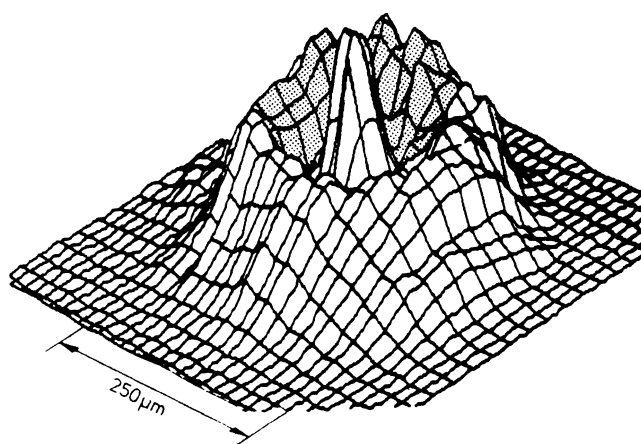


Figura 10: Resultado típico de la medida de la densidad de potencia de un láser focalizado de 2 kW operando en forma continua en el modo TEM_{10} . Gráfico obtenido mediante 3321 pixels (41 x 81).

zando w_0 por $w(z)$, donde:

$$w(z) = w_0 [1 + (\lambda z / \pi w_0^2)^2]^{1/2} .$$

A largas distancias z de la región angosta, el término del paréntesis es grande comparado con la unidad. Por lo cual, a distancias largas el haz gaussiano diverge con un ángulo θ , donde:

$$\theta = \lambda / \pi w_0 .$$

Esta es la divergencia de campo lejano para haces gaussianos. Si el haz se focaliza con una lente cuyo diámetro es al menos $2w_0$, el radio focal es $\lambda f' / \pi w_0$, el que es algo menor que el que correspondería a un radio de una función de Airy, igual a $0,61 \lambda f' / w_0$. La figura de difracción no es un disco de Airy, pero tiene una distribución gaussiana sin máximos secundarios (a menos que el diafragma de la lente recorte una parte significativa del haz incidente sobre ella). La radiación converge hacia la parte más angosta del haz y luego diverge; luego en este punto el frente de onda debe ser plano. A una distancia z de este punto de referencia, el radio de curvatura $R(z)$ del frente de onda es:

$$R(z) = z [1 + (\pi w_0^2 / \lambda z)^2]$$

donde se aprecia que solamente a grandes distancias el haz adquiere el radio de curvatura igual a z . En una cavidad confocal donde los espejos están separados por una distancia d , w_0 está dado por:

$$w_0 = (\lambda d / 2\pi)^{1/2}$$

donde el punto $z=0$ corresponde al centro de la cavidad.

Cuando la cavidad no es confocal, es costumbre definir los parámetros de estabilidad g_1 y g_2 por las ecuaciones:

$$g_1 = 1 - (d/R_1) \quad y \quad g_2 = 1 - (d/R_2)$$

donde R_1 y R_2 son los radios de curvatura de los espejos.

Para encontrar el lugar del angostamiento, razonamos que el radio de curvatura del frente de onda debe coincidir con el de los espejos cuando la onda llega a ellos. Si esto no fuera así, la distribución del campo eléctrico no sería estable. En términos de nuestra anterior descripción, los rayos no seguirían un camino cerrado. Entonces conocemos el radio de curvatura del frente de onda en estos dos puntos, que llamaremos z_1 y z_2 . Haciendo $R(z_1) = R_1$ y $R(z_2) = R_2$, podemos resolver para encontrar la parte angosta. Esto nos da:

$$z_2 = g_2 (1 - g_1) d / g_1 + g_2 - 2g_1g_2$$

en términos de los parámetros de estabilidad. Similarmente se encuentra que:

$$w_0 = (\lambda d / \pi)^{1/2} \cdot [g_1g_2(1 - g_1g_2)/g_1 + g_2 - 2g_1g_2]^{1/4}.$$

Podemos encontrar a partir de los resultados de la teoría de haces gaussianos el radio $R(z)$ o el radio focal en cualquier punto z . En particular, en el espejo 1:

$$w(z_1) = (\lambda d / \pi)^{1/2} \cdot [g_2/g_1(1 - g_1g_2)]^{1/4}$$

y en el espejo 2:

$$w(z_2) = (g_1/g_2)^{1/2} \cdot w(z_1)$$

En las expresiones de $w(z_1)$ y $w(z_2)$ se obtendrá una raíz negativa a menos que:

$$(1 - g_1g_2) > 0 \quad \text{o sea} \quad g_1 \cdot g_2 < 1$$

de esta forma no se obtendrán tamaños focales infinitos o imaginarios. Las cavidades focales para las cuales $g_1 \cdot g_2$ excede de la unidad son inestables.

bles. Por lo tanto:

$$0 < g_1 \cdot g_2 = (1 - d/R_1)(1 - d/R_2) < 1 \quad .$$

El caso límite es el $g_1 \cdot g_2 = 1$, que es una hipérbola, como se muestra en la Figura 11; los resonadores estables están entre los dos brazos de la hipérbola y los inestables fuera de estos dos brazos. Los resonadores que son menos sensibles a cambios dentro de la cavidad son los que pertenecen a la hipérbola $g_1 \cdot g_2 = 1/2$.

Hemos seguido mayormente el libro de B. Lengyel [14] y el de M. Young [15] en esta descripción. Para una mayor ilustración sobre los modos consultar los trabajos de Fox y Li [16], Kogelnik y Li [17] y G. Toraldo di Francia [18].

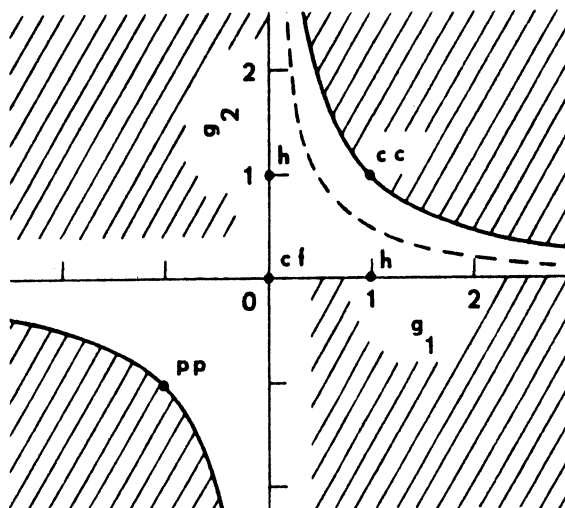


Figura 11: Gráfico de estabilidad de las diferentes cavidades resonantes de láseres.

4. Enfoque

El láser emite sus fotones coherentes, conformando un haz de baja divergencia, típicamente del orden de 1 a 2 milirradiantes para los láseres de potencia; generalmente esta salida es vertical u horizontal, dependiendo esto de la ubicación del resonador en el sistema. Lo más corriente es que el resonador esté horizontal, por lo tanto hace falta colocar un espejo que reflecte el haz para colocarlo en la posición vertical, que es la posición de trabajo natural, ya que los materiales se trabajan con sus mayores dimensiones en forma horizontal. Luego entonces del espejo se coloca una lente de enfocado, con un dispositivo que permite la refrigeración de la lente y el cambio de la lente por otras de diferentes focales. La tubuladura que mantiene este dispositivo posee un dispositivo de enfoque grueso y fino. Además contiene entradas para gases auxiliares para el procesamiento y para el aire limpio de manutención de la óptica. La Figura 12 muestra este montaje con todas estas partes detalladas.

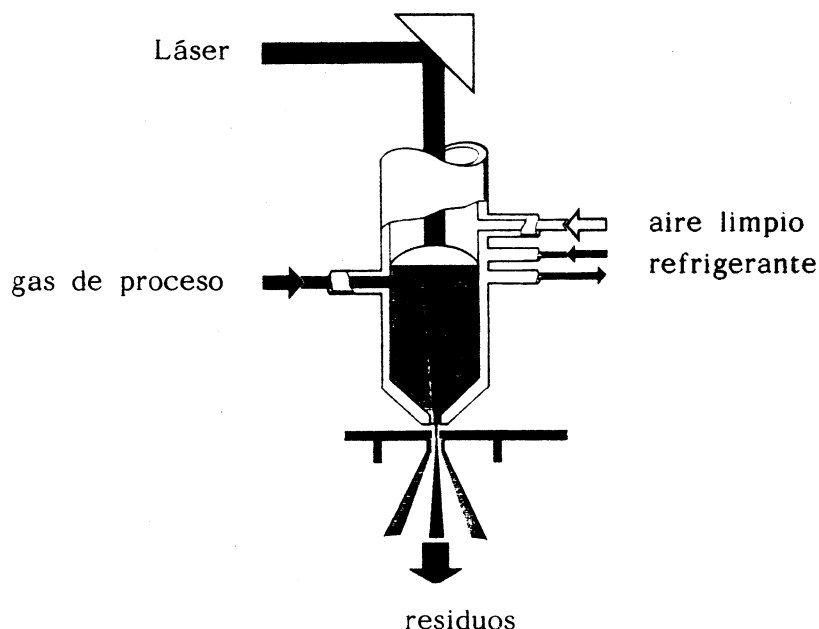
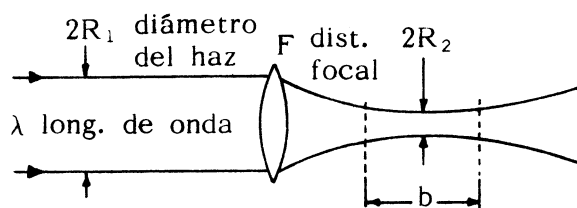


Figura 12: Montaje de la lente de salida de un láser de potencia con refrigeración, gases auxiliares y ajustes de enfoque.

La lente de enfoque es uno de los condicionamientos del tamaño mínimo de foco obtenible; igualmente ocurre con la profundidad de ese foco, ambas dimensiones gobernadas por la difracción. La Figura 13 nos muestra las dimensiones, denominadas b para el largo del foco que concentra hasta el 80% de la energía en cualquiera de sus planos y $2R_2$ el diámetro mínimo obtenible, a partir del diámetro $2R_1$ del haz que llega a la óptica de enfoque, la que tiene a su vez un diámetro mayor (en un 25% por lo menos) que el haz. Estas dimensiones están gobernadas por la difracción [19]:

$$2R_2 = 4rF/nR_1 \quad \text{y} \quad b = (2r/n) (F/R_1)^2$$

La versatilidad en el cambio de ópticas y la facilidad de limpieza de sus partes es esencial para el trabajo con emisiones de gran energía, ya que las impurezas suelen absorber energía y hacer explotar los elementos ópticos por dilatación localizada.



$$\text{Mínimo diámetro de enfoque} \quad 2R_2 = \frac{4\lambda F}{\pi R_1}$$

$$\text{Profundidad del foco} \quad b = \frac{2\lambda}{\pi} \left(\frac{F}{R_1} \right)^2$$

Figura 13: Dimensiones del volumen focal condicionadas por la difracción.

5. Régímenes de operación

La energía de los láseres industriales operados en forma continua, es de gran estabilidad. Hoy en día, variaciones de potencias menores del 1% , en períodos prolongados, son comunes en las especificaciones de los fabricantes. Para lo cual suele exigirse una variación máxima del 15% a la tensión de alimentación. Muchos de los sistemas pueden trabajar en régimen pulsado hasta frecuencias de 15 kHz.

El pulso puede ser conformado temporalmente, para redistribuir la energía disponible para conseguir densidades temporales de energía que permitan la obtención de grandes potencias. Esto es, la misma energía disponible, depositarla en un tiempo mucho más corto. Se puede además conformar la forma del pulso para lograr los efectos necesarios. Con la conformación de pulso corto se puede obtener hasta unas ocho veces la potencia máxima obtenible, durante un tiempo corto. El pulso corto es importante para el corte de metales. Otra forma de conformar los pulsos es la de pulso largo, en el que se obtiene un pico de potencia unas cuatro o cinco veces la obtenible en forma continua y luego el decaimiento del pulso por algunas unidades de tiempo adicionales. Esta conformación de pulso es importante para la soldadura láser, pues mantiene la temperatura después de la potencia pico entregada para mantener la reacción. La Figura 14 muestra las potencias de pico y la evolución temporal

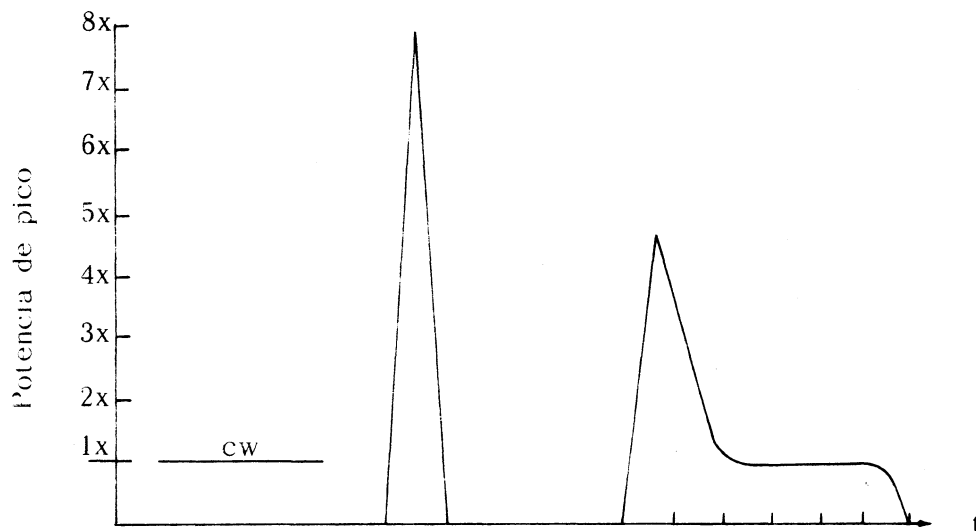


Figura 14: Conformación temporal de la emisión del láser de CO_2 . Régimen continuo (CW), pulsado corto y largo.

de los mismos. La repetición de estos pulsos a frecuencias elevadas limita la potencia máxima entregada en el pico. En general, a mayor repetición se obtiene menor potencia de pico. El mayor desempeño pulsado para un instrumento cuya prestación se grafica en la Figura 15, que en forma continua provee 220 vatios, es de 1,4 kilovatios a 100 Hz, con un ancho de pulso de 0,1 milisegundo. Otra forma de conformación de pulso es la obturación. Por este método se obtienen rampas y decaimientos u operaciones a diferentes potencias, en tiempos que son del orden

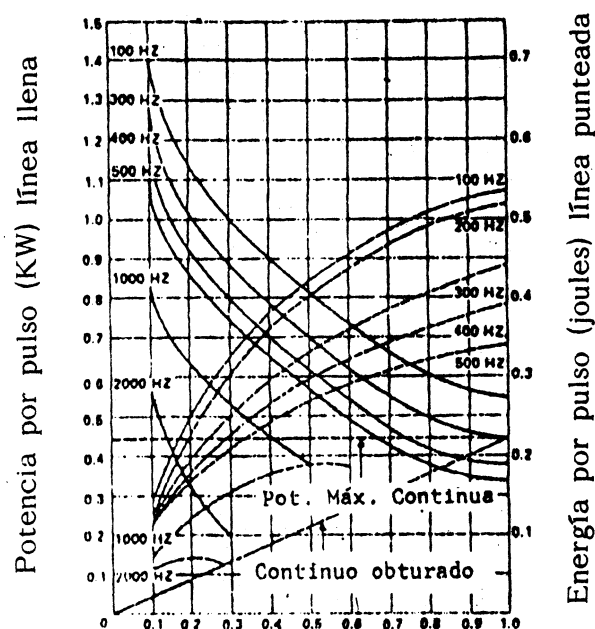


Figura 15: Limitaciones en la prestación de láser de CO_2 en regímenes continuo y pulsado en función del ancho de pulso y frecuencia de repetición de éstos.

mínimo de décimas de segundo hasta segundos y cuyas potencias de operación máxima no superan las obtenibles en régimen continuo. La Figura 16 muestra esta forma de conformación.

Por último el pulso denominado hiperpulso, que es un modo que permite la operación pulsada por encima de una energía entregada en forma continua. La Figura 17 nos muestra las diferentes formas en las que se entrega la potencia en las formas descriptas.

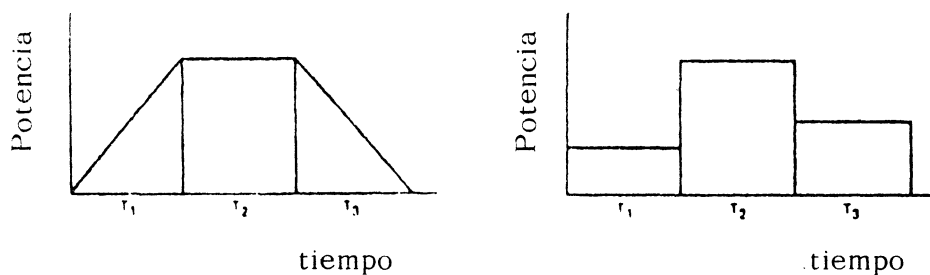


Figura 16: Conformación de pulsos que exigen baja velocidad de ajuste electrónico (del orden de déimas de segundo) incluidas en la denominación de obturación.

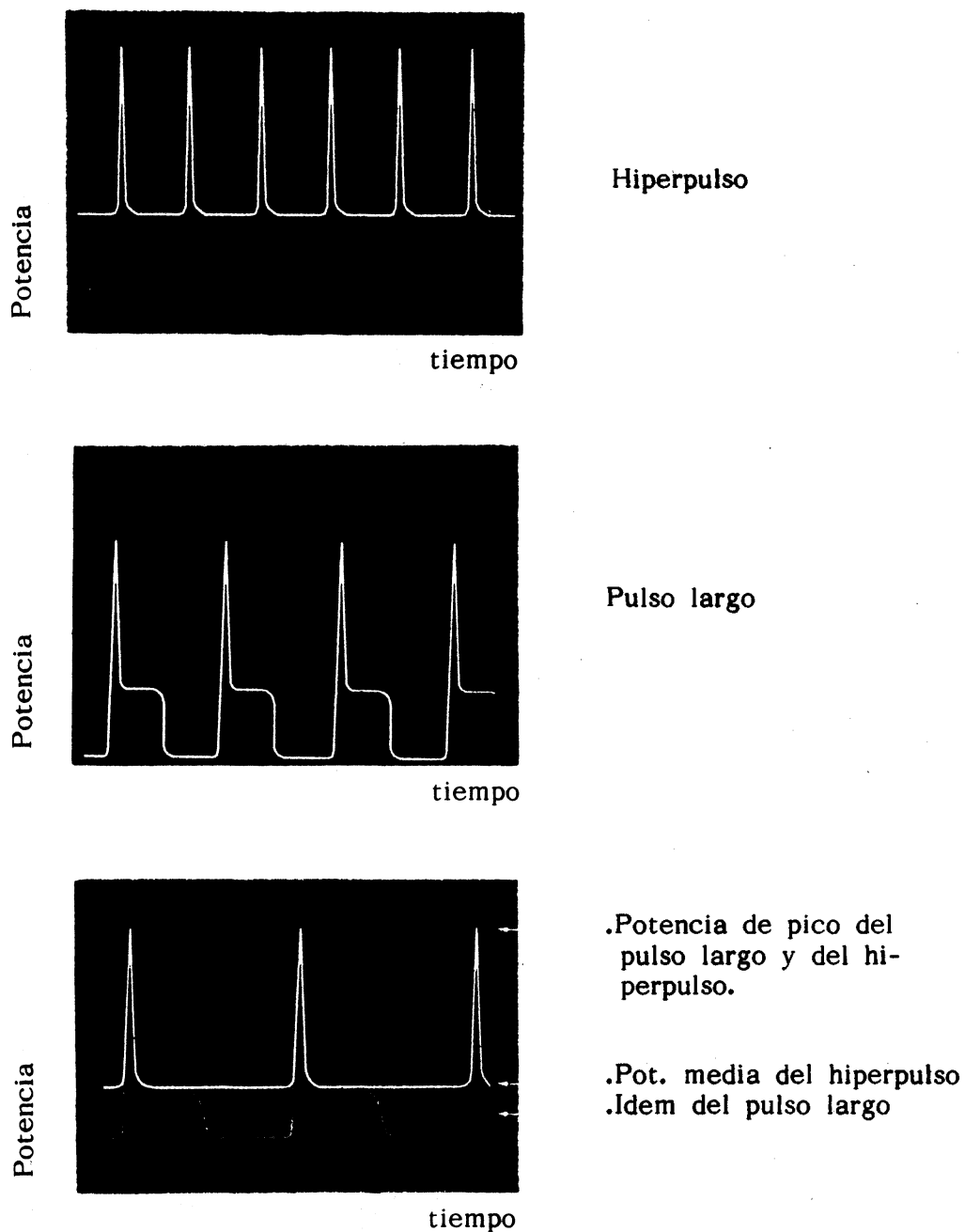


Figura 17: Entregas de potencia en los diversos regímenes de pulsación del láser. Nótese que la potencia pico del hiperpulso y del pulso largo son idénticas, pero difieren en la potencia media entregada. El pulso largo también se denomina superpulso.

6. Propiedades Ópticas de los Materiales

Para poder causar algún efecto, la luz del láser debe ser absorbida por el material que desea procesarse. Aunque esta afirmación parezca trivial, la absorción es a menudo la parte más crítica en el procesamiento láser. El proceso de absorción puede verse él mismo como una fuente secundaria de energía dentro del material. Esta fuente, aunque generada por el haz incidente, tiende a desarrollar una dinámica propia y puede en algunos casos desviarse de las leyes de la óptica ordinaria. Esta "fuente secundaria" es la que determina con su comportamiento, qué le sucede al material irradiado. Detallaremos a continuación las propiedades ópticas que determinan las propiedades de absorción de la luz, así como el comportamiento de los materiales bajo radiación intensa.

Propagación de una onda plana:

La forma más simple de luz, es una onda plana monocromática y polarizada linealmente. Esta será suficiente para describir un haz láser real. El Campo eléctrico de una onda que se propaga en un medio homogéneo no absorbente, puede ser representado como:

$$\vec{E} = \vec{E}_0 e^{i [2\pi z/\lambda - \omega t]} , \quad (1)$$

donde z es la coordenada a lo largo de la cual se propaga la luz, ω es la frecuencia angular y λ (lambda) es la longitud de onda. Estas están relacionadas entre sí con la velocidad de fase c/n_1 , donde c es la velocidad de la luz y n_1 es el índice de refracción del medio, por:

$$\lambda = \frac{2\pi}{\omega} \frac{c}{n_1} . \quad (2)$$

Una expresión análoga a la provista para el campo eléctrico, puede darse para el campo magnético H . Las amplitudes de campo magnético y eléctrico están relacionadas por:

$$\vec{H}_0 = \vec{E}_0 n_1 \epsilon c , \quad (3)$$

donde E_0 es la constante dieléctrica en el vacío. En promedio, pueden el campo magnético y eléctrico portar la misma cantidad de energía. Sin embargo, la fuerza ejercida por la onda electromagnética sobre un electrón:

$$\vec{f} = -e \left[\vec{E} + \frac{n_1}{c} \vec{v} \wedge \vec{H} \right] \quad , \quad (4)$$

posee un término de contribución debido al campo magnético que es menor a aquella producida por el campo eléctrico por un factor v/c donde v es la velocidad del electrón, y que es por lo tanto no significativa. Es el término $-eE$ en la última ecuación, la que produce todo el fenómeno que discutiremos.

El flujo de energía por unidad de área de la onda se denomina irradiancia y está dado por:

$$I = |\vec{E} \times \vec{H}| = n_1 \epsilon_0 c E_0^2 \quad . \quad (5)$$

Una irradiancia I corresponde a un flujo de $I/\hbar\omega$ fotones de energía $\hbar\omega$. La luz interacciona con la materia con un número entero de fotones. Aunque la Mecánica Cuántica es requerida para entender los procesos microscópicos de la interacción, como los flujos fotónicos son muy intensos en los láseres, es decir, se trata con muchísimos fotones, la mecánica clásica resulta suficiente para describir los fenómenos de interacción haz-sólido.

El concepto de haz implica que la irradiancia es máxima cerca del eje óptico de propagación y cae lateralmente. La distribución lateral más común es la gaussiana cilindro-simétrica, dada por:

$$I(r) = I_0 e^{-r^2/\omega^2} \quad , \quad (6)$$

donde I_0 es la irradiancia en el eje ($r = 0$), y ω describe el radio del haz. La potencia total del haz está entonces dada por:

$$P = \pi \omega^2 I_0 \quad . \quad (7)$$

La distribución gaussiana se aplica estrictamente sólo a los haces emitidos por láseres que operan en el modo fundamental TEM_{00} del resonador, de acuerdo a lo ya visto con anterioridad en la sección 3. En lo que sigue, supondremos que es siempre éste el caso que nos ocupa. El frente de onda de un haz gaussiano que se propaga libremente puede ser tomado como plano. En primera aproximación, donde los efectos de la difracción son ignorados, esto es aún válido para un haz enfocado en la vecindad del punto focal.

En un medio absorbente el índice de refracción real n_1 debe ser reemplazado por un índice de refracción complejo $n = n_1 + i n_2$. El significado de n_2 (llamado coeficiente de extinción) se hace evidente cuando en la ecuación (2) se introduce $\underline{n}$ y se reemplaza en la ecuación (1). El campo eléctrico, después de propagarse una distancia z , ahora decrece en un factor $\exp(\omega n_2 z/c)$, indicando qué parte de la energía luminosa es absorbida. El coeficiente de absorción para la irradiancia de la ecuación (5) es:

$$\alpha = -1/I \frac{dI}{dz} = 2\omega n_2/c = 4\pi n_2/\lambda \quad . \quad (8)$$

El inverso de α se denomina longitud de absorción.

En medios inhomogéneos el índice de refracción varía en el espacio; en algunos cristales, éste depende aún de la dirección de propagación. Las variaciones espaciales del índice de refracción deforman el frente de onda, curvan la trayectoria y causan ondas secundarias que se separan de las primarias, como se discute en la doble refracción.

Recordaremos la Fórmula de Fresnel que da cuenta de la reflexión de una onda que incide normalmente sobre una superficie plana de un sólido con índice de refracción $\underline{n}$. La proporción R de irradiancia reflejada está dada por:

$$R = [(n - 1) / (n + 1)]^2 \quad . \quad (9)$$

La reflectancia y el coeficiente de absorción determinan la cantidad de energía del haz que será absorbida dentro del material. La densidad depositada a la profundidad z del material por un haz perpendicular incidente de irradiancia I está descripta por:

$$J_a(z) = I (1 - R) [1 - e^{-\int_0^z \alpha(z') dz'}] \quad . \quad (10)$$

Esta expresión representa nuestra "fuente secundaria". La integral en la función exponencial está referida al camino óptico recorrido dentro del material entre 0 y z. En los materiales opacos ($z \gg 1/\alpha$) la fracción de energía absorbida es determinada por la cantidad $(1 - R)$ solamente, también conocida como absorbancia.

Propiedades Macroscópicas de los Materiales

La relación entre el índice de refracción y las propiedades de los medios de propagación está dada formalmente por las Ecuaciones de Maxwell. Estas, para materiales isótropos no magnéticos, con constante dieléctrica ϵ y conductividad σ son:

$$\nabla \cdot \vec{E} = 0 \quad (11)$$

$$\nabla \cdot \vec{H} = 0 \quad (12)$$

$$\nabla \times \vec{E} = - \frac{1}{\epsilon_0 c^2} \frac{\partial \vec{H}}{\partial t} \quad (13)$$

$$\nabla \times \vec{H} = \epsilon \epsilon_0 \frac{\partial \vec{E}}{\partial t} + \sigma \vec{E} \quad (14)$$

donde:

$$\nabla \cdot \equiv \text{divergencia}$$

$$\nabla \times \equiv \text{rotor}$$

$$\nabla \equiv \text{gradiente}$$

$$\nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \quad .$$

. Tomando el rotor de (13) y teniendo en cuenta la relación:

$$\nabla \times (\nabla \times \vec{E}) \equiv \nabla(\nabla \cdot \vec{E}) - \nabla^2 \vec{E} \quad ;$$

reemplazando estas expresiones en la (14) se obtiene la ecuación de onda:

$$\nabla^2 \vec{E} = \frac{\epsilon}{c^2} \frac{\partial^2 \vec{E}}{\partial t^2} + \frac{\sigma}{\epsilon_0 c^2} \frac{\partial \vec{E}}{\partial t} \quad . \quad (15)$$

La inserción de la expresión de la onda plana para el campo eléctrico y el uso de la relación (2) con el índice de refracción complejo, nos provee la relación de este último con la constante dieléctrica ϵ y la conductividad σ , por:

$$n^2 = \epsilon + i \sigma / \epsilon_0 \omega \equiv \epsilon = \epsilon_1 + i \epsilon_2 \quad . \quad (16)$$

La cantidad ϵ es definida como la función dieléctrica compleja, función que puede ser interpretada como una respuesta generalizada del material. Sus partes real e imaginaria están relacionadas con las del índice de refracción por medio de:

$$\epsilon_1 = n_1^2 - n_2^2 \quad \epsilon_2 = 2 n_1 n_2 \quad , \quad (17)$$

$$n_1^2 = \frac{|\epsilon| + \epsilon_1}{2} \quad n_2^2 = \frac{|\epsilon| - \epsilon_1}{2} \quad . \quad (18)$$

La función dieléctrica describe la respuesta del material a la irradiación electromagnética débil. Depende de la frecuencia de la luz; la forma como el medio responde está condicionada por la estructura microscópica del material. Veremos brevemente las principales características de

los no-metales (aislantes y semiconductores) y de los metales.

No-metales

Aislantes y semiconductores en la ausencia de excitación poseen solamente electrones ligados, y son esencialmente transparentes, excepto en la vecindad de lo que denominamos resonancias. En el modelo clásico de Lorentz, el electrón está representado como un oscilador armónico forzado por la fuerza oscilante del campo eléctrico de la onda. Los electrones oscilantes producen una polarización macroscópica del material que se superpone sobre el campo eléctrico de la onda electromagnética incidente. La función dieléctrica es simplemente la relación del campo eléctrico total (onda más polarización) en relación al campo eléctrico solamente. En la mecánica cuántica, una resonancia corresponde a una transición de un electrón entre dos estados de energía, la diferencia de energía ΔE , la cual determina una frecuencia de resonancia igual a $\omega_0 = \Delta E/\hbar$. Los tratamientos de la mecánica cuántica y clásica conducen a expresiones casi idénticas para la función dieléctrica. Para un no-metal con N_e electrones ligados que posee una resonancia sola, tenemos:

$$\epsilon = 1 + \frac{N_e e^2}{m_e \epsilon_0} f_{\text{osc}} (\omega^2 - \omega_0^2 + i \Gamma \omega) / (\omega^2 - \omega_0^2)^2 - \Gamma^2 \omega^2, \quad (19)$$

donde f_{osc} , es la "oscillator strength"; es una medida de la probabilidad de la transición, mientras que la constante de amortiguamiento Γ describe el ancho de la resonancia que resulta del ancho finito de los estados energéticos inicial y final de las transiciones. Las consecuencias ópticas de una resonancia están ilustradas en la Figura 18. Esta muestra la función dieléctrica compleja, junto a las cantidades derivadas, índice de refracción complejo $\underline{n}$, R la reflectividad y α el coeficiente de absorción, como funciones de la frecuencia de la luz, obtenidas de las expresiones 8, 9, 17-19. Nótese que un pico en el coeficiente de absorción debido a una resonancia viene siempre acompañado por un pico en la reflectancia.

Los materiales reales suelen tener muchas resonancias, y entonces en la expresión de la función dieléctrica el término de la resonancia de (19) es reemplazada por una suma de varios términos que describe cada una de ellas. Las resonancias más importantes provienen de transiciones de electrones en la banda de valencia a la banda de conducción, denominadas transiciones interbanda. Para inducir una transición interbanda, el

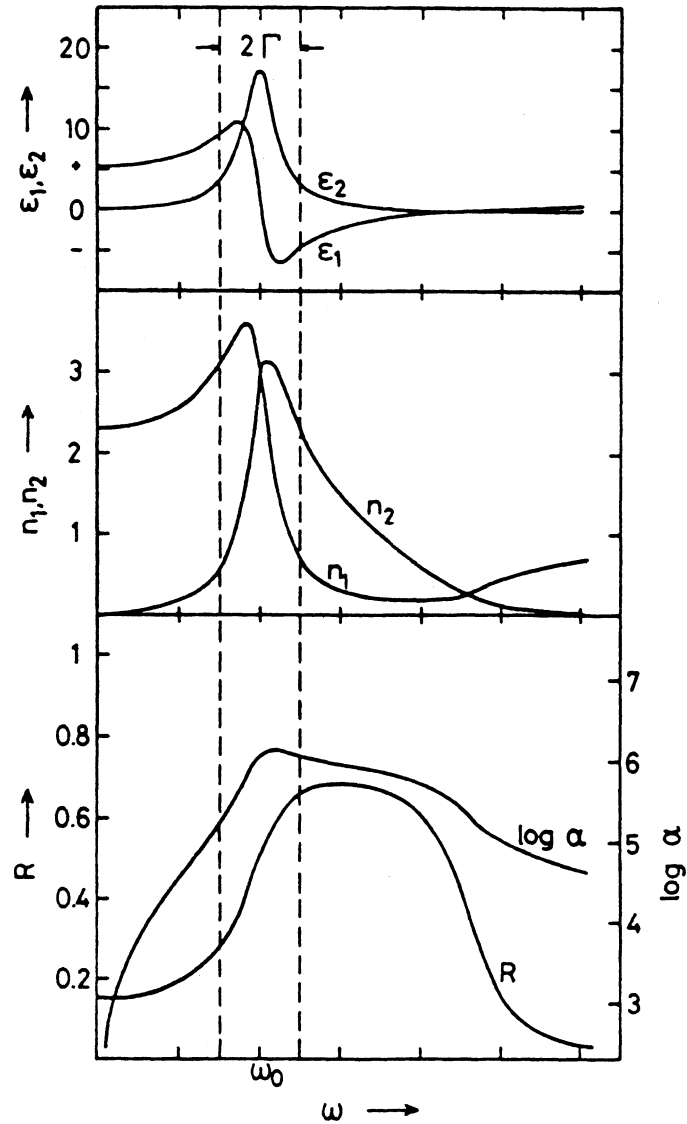


Figura 18: La Función Dieléctrica en función de la frecuencia. Índice de refracción. Reflectancia de Fresnel y el Coeficiente de absorción para un medio con una sola frecuencia de resonancia ω_0 . Cálculo para $h\omega_0 = 4$ eV, $h\nu = 1$ eV, $N = 5 \cdot 10^{22} \text{ cm}^{-3}$.

fotón incidente sobre el material debe poseer una energía por lo menos igual a la energía de la banda prohibida E_g . Los portadores libres (electrones y lagunas) creados en pares en transiciones interbanda pueden en caso de haber un número suficiente, influenciar la respuesta óptica del material. Los aislantes poseen bandas prohibidas con energías correspondientes a longitudes de onda, cuyas frecuencias están en el ultravioleta de vacío. Bajo condiciones de iluminación de luz diurna, los aislantes poseen concentración de portadores muy por debajo del número que pudiera llegar a alterar la respuesta pronosticada por la función dieléctrica. Los semiconductores poseen bandas prohibidas de energía que corresponden a longitudes de onda del espectro visible o infrarrojo. Los portadores libres generados óptica o térmicamente, contribuyen de una medida mesurable a la reflectancia del tipo metálica que muchos semiconductores poseen en la región visible del espectro.

La mayoría de los no-metales, en adición de la contribución a la función dieléctrica de las transiciones electrónicas interbanda, poseen acoplamiento resonante con fonones ópticos de alta frecuencia localizados en la región del infrarrojo cercano del espectro. El acoplamiento fonónico puede ser descrito en términos de resonancia similares a los de la ecuación (19), pero con masas y constantes de amortiguamiento características de las vibraciones de la red, en vez de vibraciones electrónicas.

Las Figuras 19 y 20 muestran el coeficiente de absorción y la reflectancia en función de la longitud de onda para dos no-metales representativos, el cuarzo con $E_g = 6,9$ eV de energía, correspondiente a una longitud de onda de 180 nm y el silicio monocristalino, con $E_g = 1,1$ eV, correspondiente a una longitud de onda de 1,13 μm . Ambos materiales exhiben las mismas características esenciales, picos de absorción alrededor de 10 μm debido al acoplamiento fonónico, absorción muy débil en las frecuencias intermedias y un incremento de absorción muy fuerte cerca de la energía que corresponde a E_g .

Con relación a la Figura 20 debe notarse que el coeficiente de absorción del Si se incrementa suavemente cuando la longitud de onda se aproxima a 1,13 μm ($h\nu = 1,1$ eV), pero lo hace en forma abrupta al aproximarse a 0,360 μm ($h\nu = 3,4$ eV). Esto se debe a que el Si posee dos tipos de transiciones interbanda. La primera es una transición directa, de los estados de energía de la banda de valencia a la banda de conducción con el mismo vector de onda. Tales transiciones son sólo posibles en el Si para los fotones que exceden la energía del gap (o banda prohibida) de 3,4 eV. La segunda clase de transiciones interbanda, involucra a estados en las bandas de valencia y de conducción de diferente vector de onda. Esta diferencia de vector de onda debe ser provista por un fonón, ya que el fotón provee un vector momento muy pequeño. Debido a que el gap fundamental a 1,1 eV es un gap indirecto, permite solamente transiciones electrónicas asistidas por fonones. Las probabilidades de estas transiciones dependen del estado de ocupación fonónica y son relativamente pequeñas y dependientes de la temperatura. Otros materiales

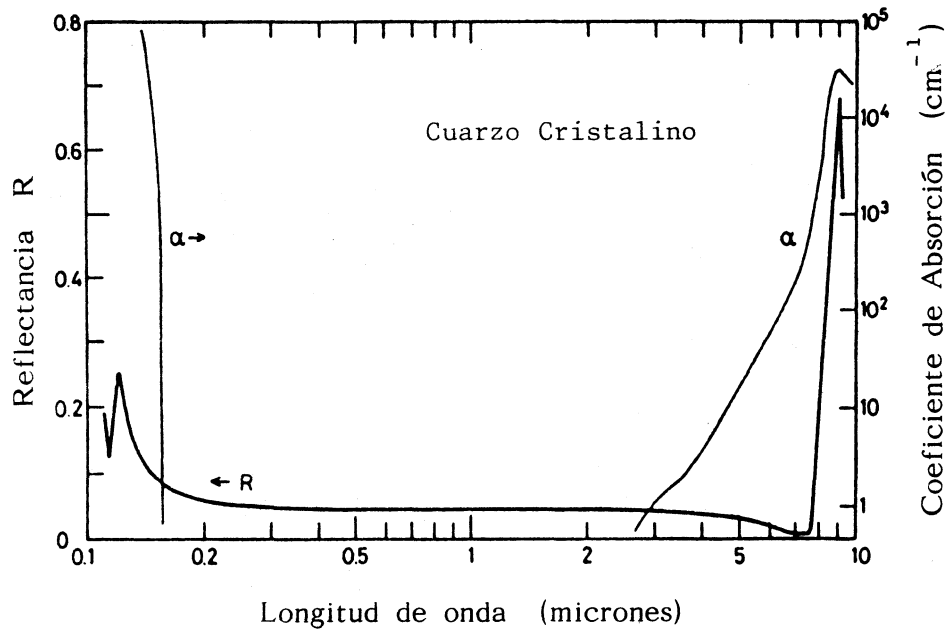


Figura 19: Reflectancia y Coeficiente de Absorción en función de la Longitud de onda para el Cuarzo cristalino.

de gap indirecto son el germanio (Ge), diamante (C), carburo de silicio (SiC), fosfuro de galio (GaP), fosfuro de aluminio (AlP) y el arsenuro de aluminio (AlAs), mientras que materiales como selenio (Se), telurio (Te), óxido de cinc (ZnO), así como la mayoría de los compuestos binarios II-VI y IV-VI tienen gaps fundamentales directos. Las aleaciones ternarias y cuaternarias (que involucran a tres o cuatro tipos de átomos diferentes) de los materiales III-V y II-VI realizados en forma no estequiométrica poseen en gran parte gaps directos y regiones de gaps indirectos, los que se utilizan para aplicaciones específicas.

El modelo simple que se ha descripto hasta aquí, se aplica solamente para materiales homogéneos. En no-metales inhomogéneos en una escala de una longitud de onda o más, la apariencia óptica es modificada por la dispersión de la luz en las rugosidades de contorno o inclusiones. Esto resulta en fuertes absorciones en materiales que deberían ser intrínsecamente transparentes, por ejemplo el caso de las cerámicas. El efecto puede

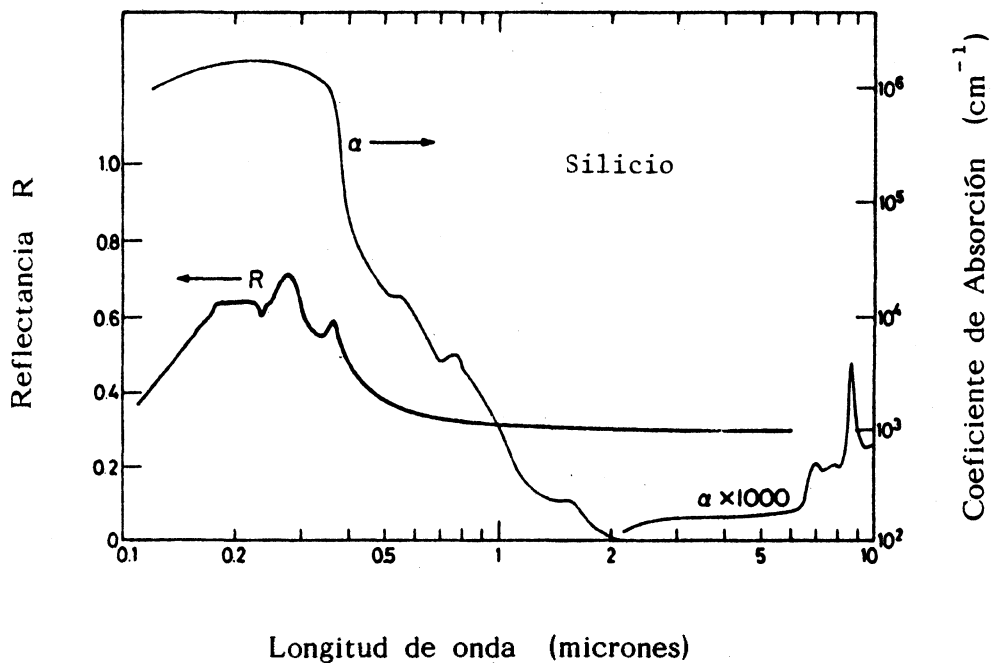


Figura 20: Reflectancia y Coeficiente de Absorción en función de la longitud de onda para el Silicio cristalino.

ser racionalizado en términos de una efectiva multiplicación del recorrido de la luz dentro del material, causado por un sinnúmero de eventos de dispersión luminosa. Inclusiones en número pequeño, pueden ser tratadas en términos de una función dieléctrica promedio (o pseudo función dieléctrica). Una expresión usada a menudo para describir una distribución aleatoria de inclusiones (metálicas o no-metálicas) con la función dieléctrica ϵ_i en una matriz con la función dieléctrica ϵ_m fue derivada por Maxwell-Garnett para explicar los colores de ciertas suspensiones y vidrios de colores.

La función dieléctrica promedio ϵ es:

$$\epsilon - \epsilon_m / \epsilon + 2\epsilon_m = V_f (\epsilon_i - \epsilon_m) / \epsilon_i + 2\epsilon_m, \quad (20)$$

donde $V_f \ll 1$ denota el volumen fraccional ocupado por las inclusiones. Esta última ecuación predice resonancias que conducen a picos de absorción que no están presentes en ninguno de los constituyentes del sistema en forma individual aislada, de huésped o inclusión.

Metales

Las Propiedades Ópticas de los metales están dominadas por los electrones de conducción. El gas de electrones es degenerado; sólo los electrones en estados de energía cercanos al nivel de Fermi, denominados electrones libres, contribuyen a las Propiedades Ópticas. No existen frecuencias de resonancia para los electrones libres, la única interacción con la red cristalina es por medio de colisiones. La función dieléctrica de un metal de electrones libres puede ser obtenida de la ecuación (19), simplemente reemplazando la constante de amortiguamiento por el tiempo inverso de colisión $1/\tau_e$ y colocando la frecuencia de resonancia igual a cero y la f_{osc} igual a la unidad. El resultado de esto nos provee de la ecuación:

$$\epsilon = 1 + \omega_p^2 (-\tau_e^2 + i \tau_e/\omega) (1 + \omega^2 \tau_e^2) \quad , \quad (21)$$

donde:

$$\omega_p = \sqrt{N_e e^2 / m_e \epsilon_0} \quad , \quad (22)$$

es la frecuencia de plasma. A $\omega = \omega_p$ (que está localizada en la región de ultravioleta de vacío para la mayoría de los metales) ambas ϵ_1 y n_1 son cero. La variación de ϵ y las cantidades asociadas se muestran en la Figura 21 en función de la frecuencia de la luz.

La frecuencia de plasma muestra dos regiones bien diferenciadas para las Propiedades Ópticas:

- 1) Grandes R y α para $\omega < \omega_p$
- 2) Pequeñas R y α para $\omega > \omega_p$.

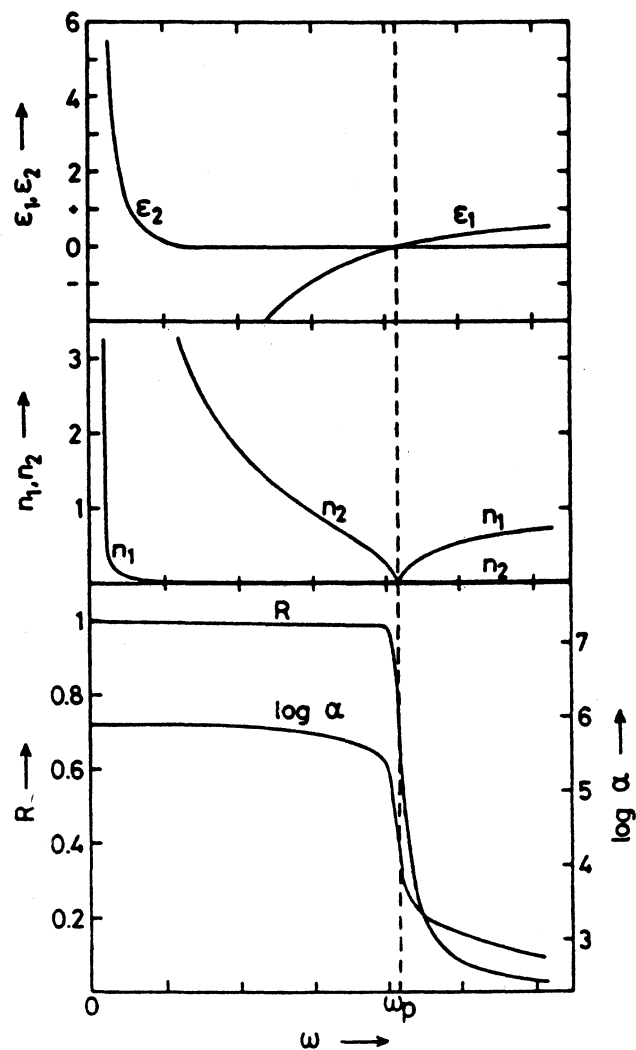


Figura 21: Dependencia de la Función Dieléctrica con la frecuencia. Índice de Refracción, Reflectancia de Fresnel y Coeficiente de absorción para un metal de electrones libres calculados para $\hbar\omega_p = 8.3$ eV, correspondiente a $N_e = 5 \cdot 10^{22} \text{ cm}^{-3}$ y $\hbar/\tau_e = 0.02$ eV.

Las Propiedades Ópticas de un metal de electrones libres para $\omega < \omega_p$ están relacionadas con su conductividad de corriente continua σ_0 . Esta está dada por la expresión de Drude:

$$\sigma_0 = N_e e^2 \tau_e / m_e = \omega_p^2 \epsilon_0 \tau_e \quad . \quad (23)$$

Aproximaciones útiles de los parámetros ópticos pueden ser obtenidos. Cuando $\omega \ll 1/\tau_e$ (región del lejano infrarrojo) la ecuación (21) nos da:

$$\epsilon_1 \approx -\sigma_0 \tau_e / \epsilon_0 \quad y \quad \epsilon_2 \approx \sigma_0 / \epsilon_0 \omega \therefore n_1 \approx n_2 \approx (\sigma_0 / 2 \epsilon_0 \omega)^{1/2} \quad , \quad (24)$$

de las que se obtiene, con ayuda de las (8) y (9):

$$1 - R \approx (8 \epsilon_0 \omega / \sigma_0)^{1/2} \quad y \quad \alpha \approx (2 \omega \sigma_0 / \epsilon_0 c^2)^{1/2} \quad .$$

Para la región donde $1/\tau_e \ll \omega \ll \omega_p$ (regiones del infrarrojo cercano y visible del espectro para la mayoría de los metales), tenemos que:

$$n_1 \approx \omega_p / 2 \omega^2 \tau_e \approx 0 \quad y \quad n_2 \approx \omega_p / \omega \quad , \quad (25)$$

de las cuales deducimos:

$$1 - R \approx 2 / \omega_p \tau_e = 2 \epsilon_0 \omega_p / \sigma_0 \quad , \quad (26)$$

y

$$\alpha \approx 2 \omega_p / c \quad . \quad (27)$$

En los metales reales, el comportamiento de un electrón libre es modificado por un número de efectos secundarios, en particular por transiciones interbanda. Como ilustración de esto mostramos en la Figura 22 los parámetros de acoplamiento para el aluminio y el oro. Las curvas de ambos metales muestran oscilaciones causadas por las transiciones interbanda, tales como las que se aprecian en $0,83 \mu\text{m}$ (energía de $1,5 \text{ eV}$) en el aluminio. Una comparación con la Figura 21, indica que la frecuencia de plasma para el Al está en la región ultravioleta de vacío, mientras que la del Au está en la región visible. Esta frecuencia de plasma anómala de baja frecuencia del Au (la cual es causante del color amarillo de este metal) es debido a las transiciones desde los estados de la banda d que se realizan a una energía de fotón de 2 eV . El corrimiento de ω_p es entendida de la inspección de las Figuras 18-21; los electrones d hacen una contribución a la función dieléctrica estando ligados y la contribución a ϵ_1 es positiva, el punto $\epsilon_1 = 0$ que define ω_p es corrido a una frecuencia menor. Comportamiento similar se encuentra en el Cu y la Ag.

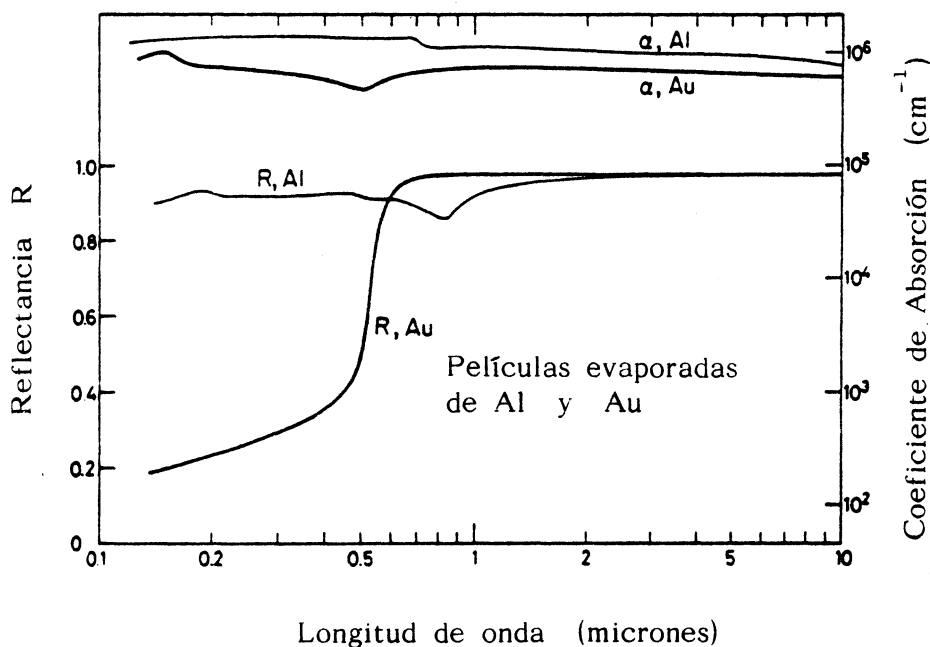


Figura 22: Reflectancia y Coeficiente de Absorción en función de la longitud de onda del aluminio y el oro.

Las Propiedades Ópticas mostradas en la Figura 21 son propiedades de volumen. En la práctica, las superficies metálicas muestran usualmente reflectancias menores que aquéllas de volumen, debido principalmente a la contaminación superficial (adsorbidos, capas de óxido, etc.) o a defectos macroscópicos. Esto justifica la gran dispersión de valores encontrados para la reflectividad de los metales publicados en la literatura. Otros desvíos del comportamiento de volumen pueden manifestarse debido a efectos intrínsecos de las superficies, como pueden ser la aparición de plasmones superficiales o la dispersión difusa de electrones, efectos estos importantes en películas delgadas. Las Propiedades Ópticas son completamente dominadas por efectos de superficie cuando las dimensiones de la película en cuestión se hace comparable al valor de la longitud de absorción, como es el caso de películas metálicas preparadas por evaporación o pulverizado (sputtering) o en estructuras agregadas consistentes en partículas metálicas pequeñas aisladas en una matriz soporte.

Con esto terminamos la descripción de las Propiedades Ópticas lineales de los materiales. Para mayores detalles sugerimos la consulta de los textos de Von Allmen [20] y Wooten [21].

Propiedades Ópticas Modificadas

Las Propiedades Ópticas Fundamentales descritas hasta ahora representan la respuesta de los materiales a la luz que es lo suficientemente débil para no perturbar los estados de los electrones y átomos en forma significativa. La potente irradiación láser, sin embargo, altera las propiedades ópticas de muchos materiales y a menudo en forma drástica. El acoplamiento ya no es caracterizado apropiadamente por la función dieléctrica estática, sino que la función se vuelve un proceso dinámico. La cantidad de luz absorbida por el material puede resultar entonces menor o mayor de la esperada a partir de las Propiedades Ópticas Fundamentales.

Los cambios inducidos en las Propiedades Ópticas por haces de luz intensos pueden ser atribuidos a uno de los siguientes tres mecanismos. Estos son aproximadamente en secuencia de irradiancia en aumento:

- 1) Producción de calor y cambios resultantes en la densidad o en las características electrónicas del material. Estos efectos incluyen el autoenfocado térmico en medios transparentes, así como fenómenos de escape térmico de portadores en semiconductores y metales.
- 2) Generación Óptica de portadores libres por transiciones interbanda o ionización por impacto en semiconductores y aislantes. Como resultante de esto, los coeficientes de absorción son incrementados dramáticamente, llegando incluso al daño explosivo del material.
- 3) Distorsión no lineal de los orbitales electrónicos o de la molécula

entera por el campo eléctrico del haz intenso de luz. Fenómenos ópticos no lineales como el autoenfocado y la absorción multifotónica son debidos a campos eléctricos intensos.

Además de estos efectos intrínsecos descriptos, el acoplamiento de la luz con el material tiende a ser afectado por las modificaciones en las formas del material, usualmente provocados por la fusión o evaporación.

7. Diagramas de Fases de Sistemas de Materiales

7.1. Generalidades

La mayoría de los sólidos poseen una estructura microscópica que muestra detalles que dependen del estado de agregación de los átomos y moléculas que los componen, así como de la distribución espacial de estos agregados. Estos estados de agregación se denominan fases.

Algunos cristales pueden mostrar formas externas que ponen en evidencia planos cristalinos de clivaje, defectos de crecimiento, defectos puntuales o lineales, o ser impecables como en el caso de ser monocristales (como en las gemas).

Los sólidos policristalinos poseen gránulos con bordes que separan cristales de diferentes orientaciones y aún de diferentes fases. En ambos casos, monocristales o policristales, y éstos de una sola fase o de múltiples fases; la forma de una fase y su distribución son dos de los más importantes parámetros estructurales. Ambos pueden ser cambiados por tratamientos térmicos o por deformación, o por una combinación de los dos. Este control sobre el arreglo de las fases en un material y cómo se modifican con diversos procesamientos, correlacionados con las propiedades de los diferentes estados de agregación, hace posible que puedan modificarse las propiedades de acuerdo al uso final, optimizando el control de propiedades.

Las fases de un material pueden entonces definirse, en términos microestructurales, como las regiones que difieren unas de las otras por su composición, estructura o ambas. El estado de agregación (sólido, líquido o vapor) puede en todos los casos, no ser suficiente para caracterizar estructuralmente un material. Un elemento sólido, compuesto o aleación, puede existir en cualquiera de varias fases con diferentes estructuras cristalinas. Un material de dos componentes en su estado de fusión, puede consistir de dos fases líquidas con diferentes composiciones.

Para describir detalladamente la estructura de un material dibujamos mapas que muestran estas fases presentes y sus cantidades relativas de sus materiales que lo componen, en función de la temperatura, presión o com-

posición. Para un material en el cual cada fase está en equilibrio con su contorno, el mapa así construido se denomina diagrama de equilibrio; si no, se denomina diagrama de fases. Adecuadamente interpretado, un diagrama de equilibrio muestra el número de fases posibles, sus composiciones relativas en función de la temperatura o presión. Muchos de los materiales de aplicación en la ingeniería existen en estado metaestable o fuera de equilibrio y cualquier cambio espontáneo se producirá buscando el equilibrio. Mucha información útil relacionada con los cambios de fase puede ser extraída de los diagramas de equilibrio apropiados. Los diagramas de equilibrio de más de un componente pueden ser clasificados como binarios, ternarios, cuaternarios, etc., dependiendo del número de componentes puros que lo integran. Los diagramas binarios son los más extensamente utilizados y son a menudo clasificados de acuerdo al número de transformaciones invariantes de fase que poseen.

Gibbs encontró a partir de consideraciones termodinámicas, la relación entre el número de fases (P) que pueden coexistir en equilibrio en un dado sistema, el número mínimo de componentes (C) que pueden ser utilizados para formar el sistema y los grados de libertad (F), siendo estos temperatura, presión, composición, los que pueden ser cambiados sin cambiar el número de fases en equilibrio. La Regla de Gibbs de las Fases es:

$$P + F = C + 2$$

Hay $P(C - 1)$ composiciones variables para cualquier sistema, ya que la concentración de cada componente puede ser especificado, menos uno de ellos, para definir la composición de cada fase, y hay P fases.

Nos acordamos de haber estudiado el caso del diagrama temperatura-presión de un elemento, donde las coordenadas (P , T) definen si el elemento está en estado sólido, líquido o vapor, o en alguna línea común a dos de estos estados: vapor-sólido, vapor-líquido o líquido-sólido, pudiendo inclusive estar en el punto triple de coexistencia de las tres fases.

7.2. Diagrama de Fases de equilibrio del Sistema Hierro-Carburo de Hierro (Fe_3C)

Este diagrama, puede construirse para el hierro. Encontramos, sin embargo que el Fe posee tres estados de cristalización: Fe delta (cristalino del tipo cúbico de cuerpo centrado o ccc), Fe gamma o austenita (cristalino del tipo cúbico de faz centrada o cfc) y Fe alfa o ferrita (cristalino del tipo ccc). En la estructura cfc los átomos están menos

compactados que en la ccc y el carbono posee mayor facilidad para difundirse en la primera. En las aleaciones de hierro y carbono, cuando el C excede los límites de solubilidad, se forma una segunda fase, la que está mayormente constituida por carburo de hierro (Fe_3C) o cementita. Forma celdas unitarias de 12 átomos, 8 de Fe y 4 de C, de forma ortorrómbica. La cementita es mucho más dura que la ferrita y la austenita. La presencia de cementita hace las aleaciones poco dúctiles. Para más detalles referimos al lector a los libros de Thornton y Colangelo [22] y de Van Vlack [23]. (El hierro y el acero son los materiales que más nos interesan para el tema de procesamiento láser, en lo que sigue solamente nos ocuparemos de ellos).

La Figura 23 nos muestra esquemáticamente la curva de enfriamiento que presentan las transformaciones polimórficas y magnética. Esta última es la temperatura de Curie, a los 780 C. A los 1539 C, el líquido coexiste con el material sólido a la misma temperatura hasta la fusión to-

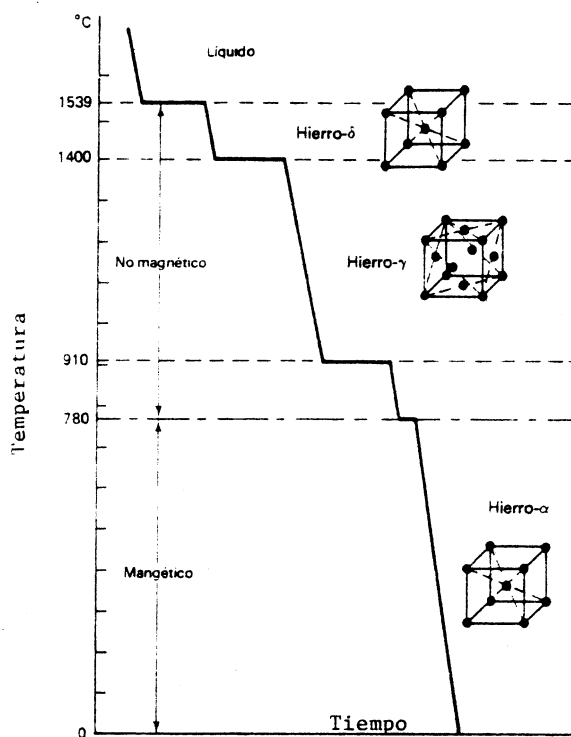


Figura 23

tal de éste. A los 1400 C ocurre la transformación sólido-sólido de Fe delta en Fe gamma y a 910 C otra transformación de la misma especie para convertir el Fe gamma en Fe alfa (ferrita).

El diagrama de equilibrio del hierro-carburo de hierro o el diagrama del hierro-carbono, como es más conocido, se muestra en la Figura 24. Este diagrama es de gran utilidad, ya que las impurezas presentes en bajas proporciones, encontradas industrialmente, de Mg , Si , S , P y otros elementos, no lo afectan significativamente. El diagrama de la Figura 24 muestra solamente la parte importante desde el punto de vista comercial e industrial.

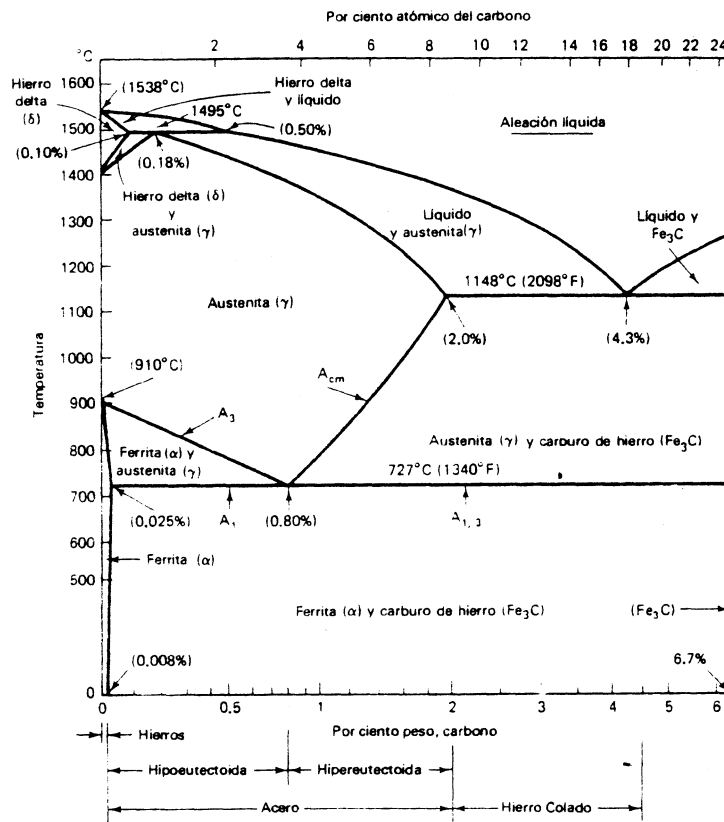
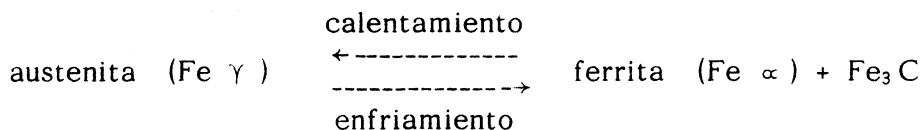


Figura 24

Nuestro interés en el conocimiento de este diagrama radica en que los procesos más usuales de corte y endurecimiento, están asociados también a la realización de una transformación cuyo resultado final resultará ser función de la composición de los aceros utilizados.

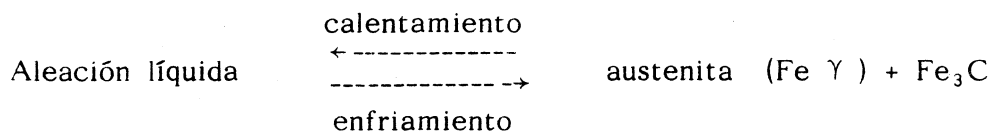
Este diagrama muestra que el hierro (hierro forjado) es un material que casi no posee carbono. Posee menos de 0,008% de carbono. Sin embargo, este material posee (agregado en forma intencional) un porcentaje pequeño de escoria (material silícico no-metálico). El hierro forjado presenta una buena resistencia a la corrosión y a las fallas por fatiga. Al no contener elementos de aleación, es muy dúctil y por lo tanto muy apto a la fabricación de alambres, clavos, pernos, caños, tubos y otras piezas que no requieran de alta dureza o resistencia.

El acero puede contener carbono aleado hasta proporciones del 2% . Esta proporción es la cantidad máxima de carbono soluble en el acero en estado sólido. Esta categoría de aleaciones ferrosas se clasifica en hipoeutectoides (que contienen menos de 0,8% de C) y en hipereutectoides (que contienen más del 0,8% de C). El término eutectoide se refiere a la reacción del estado sólido que tiene lugar a una composición y temperatura específicas, esto es, para el caso del acero, en condiciones de equilibrio a 723 C con una composición de 0,8% de C . Si en este punto se aplica la regla de las Fases de Gibbs, hay cero grado de libertad. Esto significa que la reacción eutectoide es invariante. Esta reacción puede esquematizarse:



El producto de esta transformación de estado sólido es una mezcla íntima de ferrita (alfa) y de carburo de hierro (Fe_3C), conocida como perlita.

La reacción eutéctica en el diagrama Fe-carburo de Fe , tiene lugar en la región del hierro colado (denominación para la aleación con más del 2% de carbono) a una temperatura de 1148 C y 4,3% de C. Esta puede esquematizarse como:



La metalografía nos revela cómo son las composiciones de estas aleaciones cuando realizamos enfriamientos lo suficientemente lentos, tales como los que se muestran en la Figura 25, para un par de elementos A y B que forman una aleación eutéctica, que son completamente solubles en estado líquido y completamente insolubles en estado sólido (veremos después qué validez tiene esto para el sistema Fe-Fe₃C).

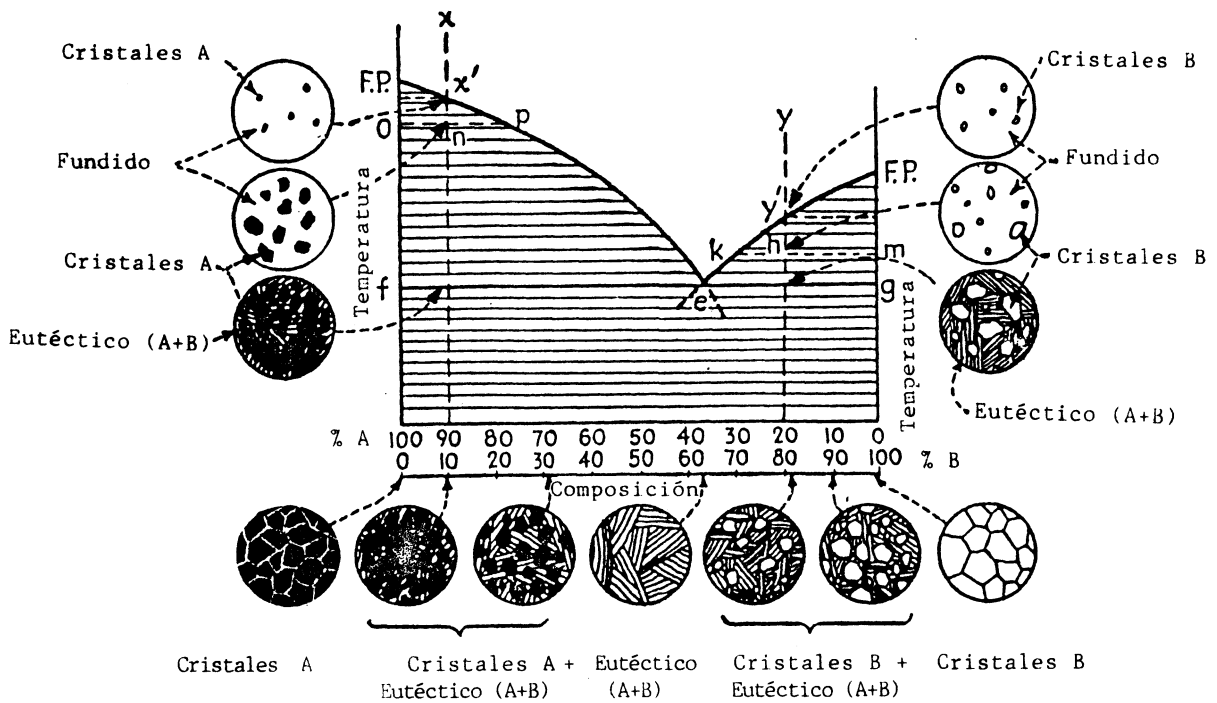


Figura 25

Para apreciar esta microestructura es necesario tener la ayuda de un microscopio, aunque algunas veces la macroestructura es perceptible a simple vista. Esta figura nos muestra una aleación x que tiene ahora un comportamiento predecible en su enfriamiento lento. Se muestra en primera instancia cristales A que comienzan a nuclearse al bajar las temperaturas por debajo de la línea liquidus (ésta marca las temperaturas por encima de las cuales cualquier composición líquido-sólida se licúa totalmente). La curva liquidus se conforma para este caso en particular por la Ley de Raoult que dice "La temperatura de solidificación de una sustancia pura disminuye con la adición de una segunda sustancia, siempre que ésta sea soluble con la primera en estado líquido e insoluble cuando se solidifica, siendo la disminución proporcional al peso molecular del soluto". En nuestra figura se muestra que al continuar el enfriamiento para la composición x , los cristales que se formaron en la primera instancia de la nucleación, llamados cristales primarios han crecido de tamaño y ahora en tercera instancia han aparecido otros cristales menores de A y B que los rodean (estos han nucleado más tarde y por ende han tenido menor tiempo para crecer) formando una matriz eutéctica alternada finalmente dividida de cristales A y B . Las proporciones en cada instante pueden ser predichas por la regla de la palanca (ver más adelante). En la parte inferior del gráfico se muestra cómo esta composición conlleva a diferentes estructuras para varias aleaciones. Nótese que la proporción de cristales primarios A se incrementa de cero en la composición eutéctica hasta el 100% para A puro, mientras que para B se incrementa de cero para la aleación eutéctica hasta llegar al 100% para una aleación de B puro.

La línea horizontal fg a lo largo del diagrama se denomina solidus pues marca para cada composición la solidificación completa para todas las aleaciones a esta temperatura.

Las aleaciones que solidifican de esta forma son muy raras, ya que la nucleación de cristallitos tendría comportamiento similar a la de los metales puros. La mayoría de las aleaciones completamente solubles depositan cristales que muestran un grado mesurado de diferencias en composición dependientes de la solución donde crecen (cuya composición es temporalmente dependiente). La curva solidus en estos casos no es ya ni recta ni horizontal.

La solidificación en una aleación a partir del estado líquido cuando existe completa solubilidad, se realiza por dos mecanismos:

- 1) Formación de cristales en el fundido.
- 2) Homogeneización de la composición en varias partes de cada cristal:
 - a) por difusión entre núcleo y cáscara.
 - b) por difusión entre núcleo y fundido.

El proceso 1) es mucho más rápido que el proceso 2), como es

el caso usual en la industria metalúrgica y no existe equilibrio en ningún momento entre los cristales y el fundido; así la solidificación de las aleaciones no seguirá la línea solidus del diagrama exactamente. Esto es importante y lo describiremos. La solidificación de aleaciones completamente solubles se relata de acuerdo a la Figura 26. Un fundido de composición C se enfría hasta cortar la línea liquidus AB, donde de acuerdo a la regla 1 enunciada, precipita cristales de composición D. Estos se muestran en la Figura 27 a. Estos cristales precipitados son más ricos en A que el fundido mismo, ya que con su formación dejan al fundido más rico en B. Ellos no son, sin embargo, cristales de A puro, ya que al disminuir la temperatura al punto E, el fundido comienza a precipitar cristales de composición F, en forma separada a los D ya formados o bien rodeando estos últimos, como se muestra en la Figura 27 b. Para que la composición de la fase cristalina sea aquélla del punto F del diagrama, algunos átomos de B deben difundirse hacia el centro de los cristales ricos en A desde el líquido, obviamente por una difusión entre el material D y el F y esto resulta en una composición intermedia entre D y F y no en la composición F. Si la difusión tiene tiempo de producirse, esto es, si el enfriamiento se hace lentamente, el baño se mueve bajando a lo largo de la línea solidus en ambas temperatura y composición conforme la aleación se enfría. Si por el contrario, el enfriamiento se hace tan rápido que la difusión no tiene lugar, los cristales a la tem-

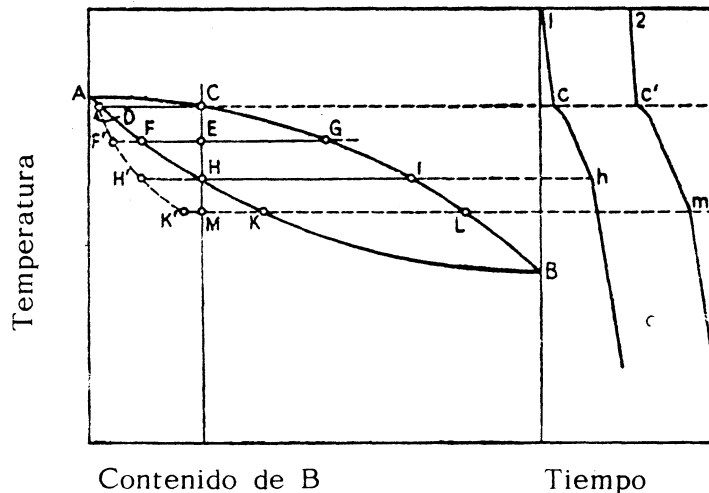


Figura 26

peratura E no serán homogéneos de composición F, sino que contendrán núcleos centrales de composición D rodeados por el material de composición F y que tendrán una composición media F'. A la temperatura H, la composición media de la fase cristalina no es H sino H'; esto significa que la fase cristalina es constantemente más rica en A de lo que sería en condiciones de equilibrio.

Examinemos ahora el caso de solubilidad parcial en estado sólido. Haremos referencia a la Figura 28. Las líneas f-F.P., g-F.P. en el sistema eutéctico que se describe en esta figura, encierran conjuntamente con sus respectivas liquidus las regiones de coexistencia de dos fases: fundido + cristales alfa y fundido + cristales beta. Ver Figura 28 a. Los líquidos en estas aleaciones precipitan cristales que nunca son puro A o puro B, sino que siempre contienen un poco del otro componente en solución sólida, lo que es similar al caso ya considerado de aleaciones completamente solubles en estado sólido. Sin embargo, a la temperatura eutéctica e, las liquidus se intersectan y el líquido precipita alternativamente no cristales de puro A y B como en un simple eutéctico de Raoult, sino cristales de A conteniendo f% de B en ellos y cristales de B conteniendo g% de A. Veremos cómo se calcula la cantidad de matriz eutéctica después de completada la solidificación. Por ejemplo, la aleación z' en la Figura 28 c a la temperatura fne contendría $f_n/f_e \times 100$ de líquido (de composición e) y $n_e/f_e \times 100$ de cristales (de composición f). A temperaturas por encima del eutéctico se comportan simplemente como completamente solubles, mostrando la estructura nucleada si se enfrían rápidamente. Por debajo de la temperatura de la línea del eutéctico feg, un fenómeno familiar tiene lugar. A medida que la temperatura

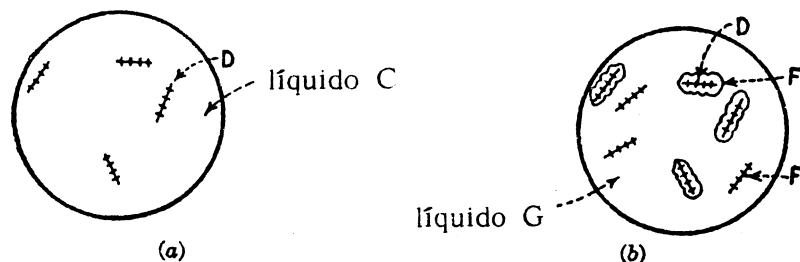


Figura 27

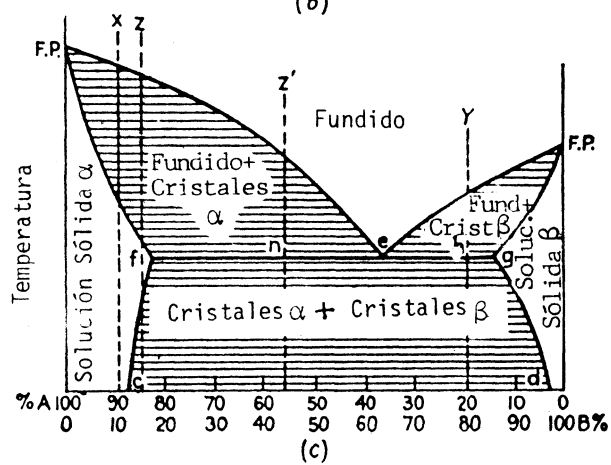
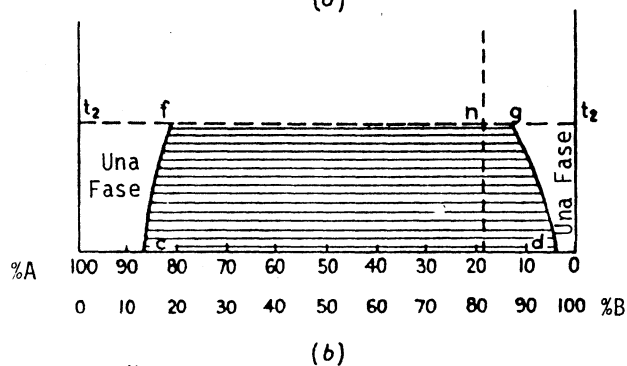
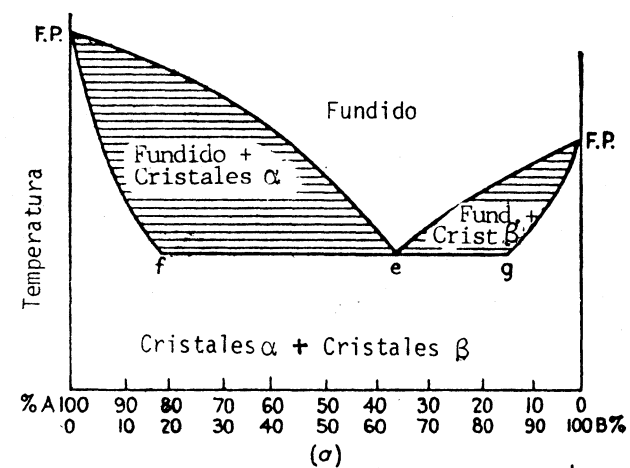


Figura 28

decrece, también hay un decrecimiento de la solubilidad de B en cristales de A y un decrecimiento de la solubilidad de B en cristales de A, hecho que puede verse en la Figura 28 b, donde se aprecian las regiones de existencia de una sola fase y la de coexistencia de dos fases.

Este diagrama se ha compuesto pues de tres casos simples ya conocidos: la reacción eutéctica, parte de los diagramas de solubilidad completa y la solubilidad parcial, lo que resulta fácil en su análisis global a partir de los casos simples, de cada una de las regiones expuestas.

La estructura de las aleaciones parcialmente solubles muestran un núcleo rodeado de otro compuesto en enfriado rápido y estructura de una fase, homogeneizada, luego de completada la difusión. Esto es válido para una aleación x de la Figura 28 c. La aleación z mostraría una estructura homogénea similar a la x (después de completado el homogeneizado), excepto que, dado que aquélla cae en la zona del campo de dos fases fgdc a temperatura ambiente, los granos homogéneos que debieran haberse formado se han vuelto supersaturados con respecto a los átomos de B a la temperatura por debajo de f y los ha precipitado. Los granos homogéneos contendrían así un fino precipitado de cristales beta, de composición d, que aparecen como segregados en los contornos de los granos y en los planos cristalográficos de los cristales huéspedes de la solución alfa sólida. La separación de este precipitado beta del huésped, en este caso cristalitas de solución sólida alfa, tiene lugar como un mecanismo similar al de la formación de una nueva fase polimórfica en un metal puro. La composición final de los cristales alfa cuando alcanzan la temperatura ambiente serán los de c. Esta estructura se llama de Widmannstaetten.

Los aceros muestran este tipo de solubilidad decrecientes. En la Figura 24 se muestra la solubilidad decreciente de la austenita (Fe gamma) en la ferrita (Fe alfa), por debajo de 910 C, marcados por la línea A₃. También la línea casi vertical a la izquierda del diagrama entre 723 C y 0 C muestra la solubilidad decreciente del carbono en la ferrita cuando disminuye la temperatura.

Volviendo a la Figura 28, las aleaciones z' e y contendrán una matriz de cristales de c y de d, la aleación z' mostrará cristales primarios de composición d. Estos cristalitas no serán homogéneos, pero de igual modo que los cristalitas alfa de la aleación z, tendrá un fino precipitado de uno en otro en sus bordes de grano y dentro de los mismos cristales, debido al decrecimiento de la solubilidad de A y B conforme la temperatura disminuye desde g hasta d.

En régimen de enfriamiento lento y en completo equilibrio, la aleación z no cruza la línea eutéctica, y el último líquido antes de la solidificación no alcanzará el punto e; por lo tanto, no debe ser esperada la formación de una estructura eutéctica en los bordes de grano. Sin embargo, en régimen de enfriamiento rápido, la fase cristalina se vuelve muy rica en el metal A (debido a la falta de difusión) y resulta que el líqui-

do se vuelve poseedor de una cantidad de B mayor de la que indica el diagrama, ya que este diagrama sólo vale para las condiciones de equilibrio. A velocidad rápida de enfriamiento, la difusión es prácticamente cero y la fase líquida alcanza el punto e. A este punto, está saturada en B y cristales de la solución beta sólida se separarán alternativamente con alfa. Un eutéctico es obtenido en la aleación z. El mismo resultado puede ser obtenido para la aleación x si es rápidamente enfriado. Un tratamiento de homogeneización puede hacer desaparecer el eutéctico por difusión de los átomos B en los granos con núcleos ricos en A con granos primarios alfa. Si la aleación z se enfría rápidamente desde la temperatura f_{ne} , cruzará la línea fc sin precipitar beta. A temperatura ambiente, es obtenida una solución sólida supersaturada que contiene f por ciento de B.

Lo que acabamos de describir sucede merced a que la velocidad de difusión atómica varía como $\exp(-C/T)$, donde C es aproximadamente una constante y T es la temperatura absoluta. Se entiende entonces que las transformaciones de fase que ocurren por difusión a menudo pueden ser completamente o casi completamente inhibidas, enfriando el sólido rápidamente a una temperatura baja. De esta forma se toma ventaja práctica para confeccionar las aleaciones endurecidas por precipitación.

Cuando una transformación invariante que normalmente ocurre en estado sólido es inhibida por enfriamiento rápido, puede formarse una nueva fase, completamente nueva, fuera de equilibrio. Este es el caso de los aceros de bajo tenor de carbono sin otros elementos adicionales de aleación. Las fases de equilibrio en estas condiciones a temperatura ambiente serían hierro y grafito. En la práctica, el grafito nunca se forma; lo hace sin embargo la cementita o carburo de hierro (Fe_3C), que es el elemento estable fuera de equilibrio. La única transformación invariante que tiene lugar enteramente en estado sólido es la eutectoide. Debido a que la difusión de estado sólido es muy lenta, ésta puede ser completamente inhibida si se enfría lo suficientemente rápido el acero desde una temperatura superior a $723\text{ }^{\circ}\text{C}$.

La Figura 29 nos muestra el diagrama de equilibrio de aspecto típico de A puro y A_xB_y con (a) solubilidad completa en estado sólido, (b) insolubilidad completa en estado sólido, y (c) solubilidad parcial de uno en otro en estado sólido.

Para entender las estructuras que se forman fuera de equilibrio a altas velocidades de enfriamiento y que resultan de la inhibición de la descomposición del eutectoide del acero, discutiremos primero las estructuras que se forman a bajas velocidades. Ya hemos anticipado que el carbono es mucho más soluble en la fase de austenita que en las fases de ferrita y delta, debido a que los espacios intersticiales son mayores en la estructura cristalina cúbica de cara centrada (ccc). Debido a esto, un acero que contenga más del 0,02% de C, en un enfriamiento lento desde la fase de austenita, precipitará en exceso de C para formar Fe_3C . Los

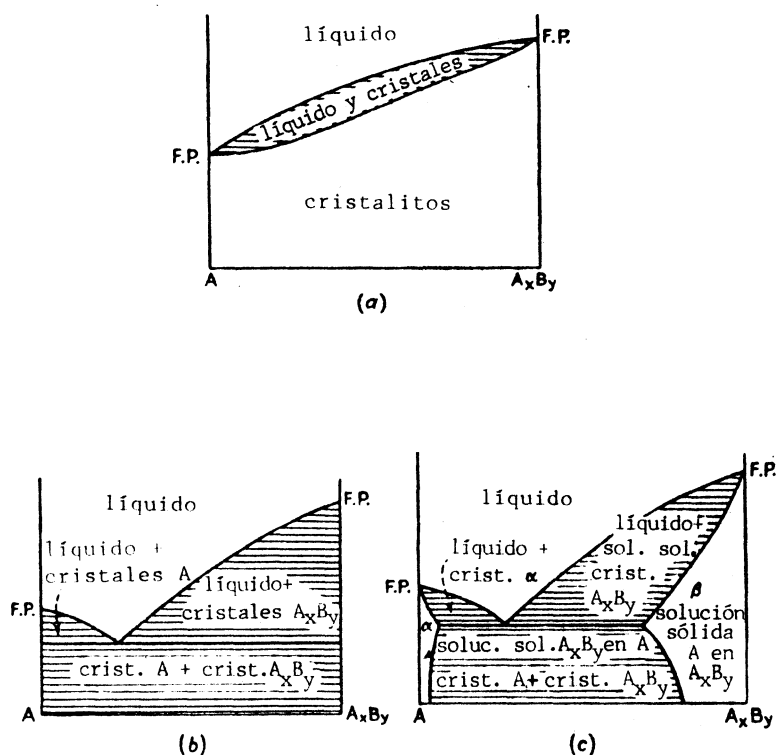


Figura 29

cambios progresivos en la microestructura de un acero con 0,4% de C cuando se enfría lentamente están ilustrados esquemáticamente en la Figura 30. Primero los granos de la solución sólida de alfa nuclean en los bordes de granos de la austenita, hasta que a una temperatura un poco por encima de 723 C, la estructura consiste de gamma (0,8% C) y alfa (0,02% C). A esta temperatura, el peso fraccional de austenita es determinado por la regla de la palanca:

$$f(\text{gamma}) = 0,4 - 0,02 / 0,8 - 0,02 = 0,49$$

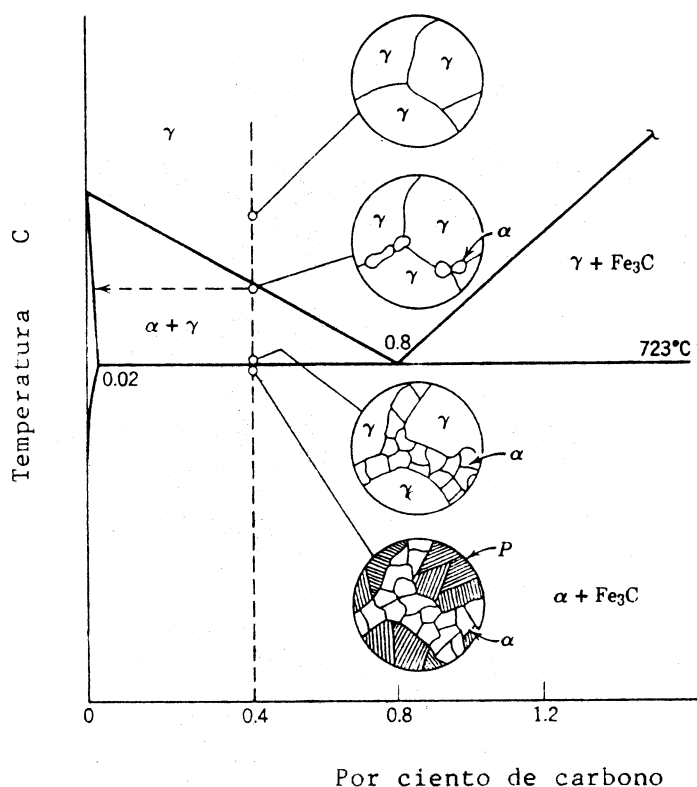


Figura 30

Cuando la aleación es enfriada por debajo de 723 C, esta austenita se transforma en una fina y laminar mezcla de alfa y Fe₃C . La mezcla es distinguida fácilmente al microscopio y se denomina perlita. El comportamiento es el mismo para un acero al 1% de C , excepto que hay una parte fraccional de austenita mayor cerca y por encima de 723 C ,

$$F(\text{gamma}) = \frac{6,67 - 1,0}{6,67 - 0,8} = 0,97 \quad ,$$

y la fase proeutectoide es el Fe₃C en vez de alfa. Microestructuras como éstas, en las cuales se observan fases a bajas temperaturas, ferrita y

cementita, son las que resultan del normalizado, o lento enfriamiento del acero. Las composiciones de las fases y sus componentes relativas pueden ser determinadas del diagrama de fases Fe-Fe₃C.

Si el acero es enfriado rápidamente ("quenching"), dos microestructuras son formadas en vez de la perlita, al inhibirse la transformación eutectoide. La bainita, que puede formarse cuando la austenita se enfría súbitamente a temperaturas entre 200 y 400 C y mantenidas a estas temperaturas por un instante. La bainita es una dispersión submicroscópica de carburos en una matriz alfa filamentar que contiene más del 0,02% de C. La martensita puede formarse cuando esa transformación se realiza a temperaturas aún menores. Es una fase extremadamente dura y frágil en la cual el C está atrapado en una solución sólida supersaturada. El exceso de C distorsiona la estructura tetragonal de cuerpo centrado (tcc) y la cantidad de distorsión medida por la relación media de c/a de las aristas de la celda tetragonal es aproximadamente una medida del C que contiene atrapado.

La transformación de austenita en perlita y bainita es en ambos casos dependiente del tiempo y de la temperatura; así también la obtención de los proeutectoides ferrita y cementita, pues ocurren por nucleación y crecimiento del grano. No ocurre lo mismo con la transformación de austenita en martensita, la que es casi independiente del tiempo. Ocurre por un mecanismo de cizallamiento similar al de la torsión mecánica. La proporción de martensita a lograrse depende casi exclusivamente de la temperatura más baja a lograrse. Cuanto más baja, mayor cantidad de martensita.

Diagramas de transformación isotérmica:

El diagrama usual no representa el curso de la descomposición de las fases dependientes del tiempo, como los que acabamos de describir. Este tipo de proceso puede ser descrito por un gráfico T-T-T (Tiempo - Temperatura - Transformación). Este tipo de gráfico se confecciona después de un considerable esfuerzo metalográfico, donde se determinan un sinnúmero de situaciones que conlleva a conclusiones sobre la obtención de estructuras deseadas. El procedimiento consiste (para el caso de los aceros) en austenizar el material un número de muestras finas y luego sumergirlas en un baño a una determinada temperatura por debajo de la eutéctica. Las muestras son retiradas del baño isotérmico, luego que comienza la descomposición y son colocados a temperatura ambiente o más baja. La extensión de la transformación es ponderada metalográficamente, como se muestra en la Figura 31, donde las primeras muestras fueron enfriadas súbitamente a una temperatura lo suficientemente alta como para formar perlita; a temperaturas más bajas podría haber formado bainita. Con este proceso repetidamente, se obtiene una familia de cur-

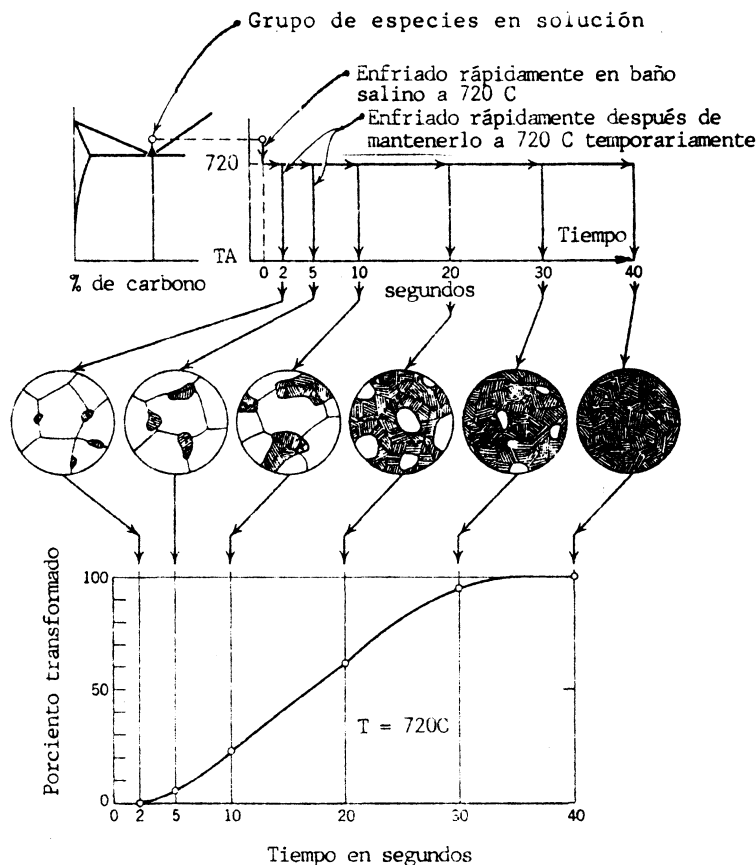


Figura 31

vas, las que llevadas al gráfico T-T-T, permite decir el porcentaje de austenita transformada. La Figura 32 a, muestra una curva T-T-T para un acero eutectoide (0,8% C); las curvas "comienzo" y "final" describen el comienzo y el final de la descomposición austenítica. Las microestructuras que se obtienen son resultado de la temperatura a las que se enfrían isotérmicamente. A temperaturas un poco por debajo de la eutectoide, la velocidad de difusión es alta y la de nucleación baja; pocos núcleos se forman, pero crecen rápido y producen una perlita de grano grande con estructura laminar muy espaciada. A temperaturas inferiores (unos 200 C), la velocidad de nucleación es mucho mayor y la difusión más lenta; así resulta una perlita fina. Unos cientos de grados por debajo, la bainita es el producto de la transformación. A temperaturas aún más bajas, donde la curva de enfriamiento se intersecta con M_s (comienzo de la forma-

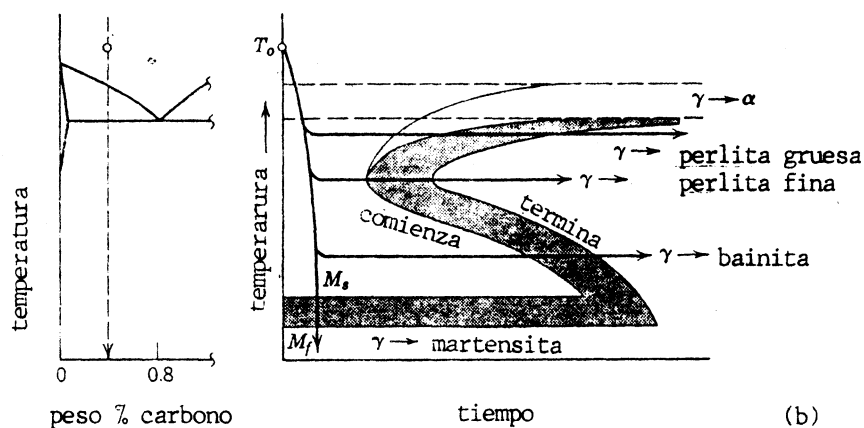
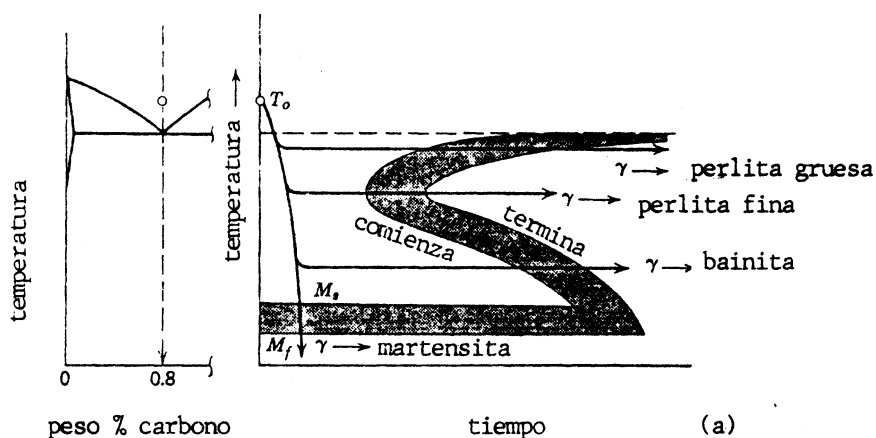


Figura 32

ción de martensita) la martensita se forma. Si la curva de enfriamiento no cruza M_f (final de la transformación martensítica), alguna austenita quedará sin transformar. La curva T-T-T posee una "nariz" o tiempo mínimo, por debajo del cual ninguna transformación comienza. Los aceros aleados poseen una segunda nariz más baja ("chin") en el rango donde se

forma la bainita. Si se desea la producción de martensita, el acero debe ser enfriado rápidamente sin cruzar la nariz o el chin. El mínimo intervalo temporal antes de comenzar la transformación determina la severidad del enfriamiento necesario. A menudo otros elementos son añadidos a la aleación para aumentar este tiempo de acceso al endurecimiento. Esto persigue el propósito de endurecer piezas grandes sin necesidad de someter las piezas a una eventual fractura superficial debida al gradiente térmico elevado que requeriría el material sin el compuesto de aleación de endurecimiento.

En el diagrama T-T-T de la Figura 32 a, se representa un acero eutectoide. Si el porcentaje de C es menor (hipoeutectoide) o mayor (hipereutectoide) a la composición eutectoide, ocurre la precipitación de una fase proeutectoide. Si el acero hipoeutectoide de la Figura 32 b, se enfría lentamente desde T_0 , la ferrita proeutectoide se forma primero y la composición de la austenita cambia a 0,8% de C a la temperatura eutectoide. La formación de la ferrita proeutectoide (o Fe_3C en el caso de un acero hipereutético) es un proceso de nucleación y crecimiento que puede ser suprimido. Si el acero es rápidamente enfriado desde T_0 a una temperatura debajo de la "nariz" de la curva de transformación y luego se permite la transformación isotérmica, como se indica en el T-T-T de la Figura 32 b, la ferrita proeutectoide no se forma, produciéndose una estructura enteramente bainítica.

De expreso no prestaremos atención a otras partes del diagrama de equilibrio del $Fe-Fe_3C$, ya que las mayores aplicaciones de los láseres para el corte se centran alrededor de los laminados de hierros y aceros.

8. Procesamiento de hierros y aceros

La mayor actividad metalúrgica es la producción de los hierros y aceros laminados. Los laminados se producen después de haberse fabricado las aleaciones deseadas, pudiendo ser realizados en frío o en caliente. La realización de los laminados permite, en sus múltiples espesores, recortar piezas para el posterior procesamiento, estampado, soldado, etc., que aproxima el material ferroso a la concreción del producto. El corte de metales a partir de las planchas de dimensiones estandarizadas, es una actividad de producción y servicios muy importante por su volumen, en el ciclo productivo de los hierros y aceros. Los laminados se producen en chapas de espesores desde décimas de milímetro hasta los quinientos milímetros en forma estándar y de espesores mayores bajo pedido. En lo que sigue detallaremos el corte, perforado, endurecimiento, enchapado, aleación superficial, soldadura y la elaboración de matricería con láseres de CO_2 .

Para obtener una interpretación física de la intensidad de un haz láser, es interesante asociarla a una "temperatura equivalente" de una fuente térmica de energía. Recordemos que la relación de Stefan-Boltzmann para la energía total irradiada por unidad de área por una fuente térmica es:

$$E_T = \epsilon \sigma (T - T_0)^4 ,$$

donde σ = constante de Stefan-Boltzmann, ϵ = emisividad, T = temperatura absoluta de la fuente emisora y T_0 = temperatura absoluta del ambiente.

Como referencia, notemos que una densidad de potencia de $10 \exp 6$ Watt/cm² corresponde a una fuente térmica radiante a 20.500 K.

Comparando al láser con dispositivos convencionales, el láser provee mucha más energía concentrada en un punto, lo que facilita la fusión y vaporización de cualquier material conocido.

En la Figura 33 vemos el espectro de interacción láser-material que se expande sobre muchos órdenes de magnitud de densidad de potencia [8.0, 8.1]. Esta figura provee junto al espectro de densidades de potencias los tiempos de interacción con el material para diferentes tipos de procesamiento. En lo que sigue, lo explicaremos con más detalle. En presentaciones a congresos realizadas con colaboradores, se ha hecho difusión de estas técnicas en el país, con anterioridad [8.1].

8.1. Corte de hierros y aceros

El corte a guillotina mecánica fue por mucho tiempo, para laminados hasta de 25 mm de espesor, el método más utilizado para el corte metálico. Se cortaron por corte oxiacetilénico (también llamado oxicorte) en forma manual durante muchos años todos los laminados ferrosos. El corte pantográfico, donde se lee en forma óptica un plano y se corta simultáneamente con el oxicorte en múltiples picos las piezas leídas, se difundió ampliamente en los últimos quince años. Otros métodos de corte son el mecanizado químico, para espesores menores al mm, el plasma, para metales no ferrosos y aceros inoxidables y el láser. La Figura 34 nos muestra las precisiones obtenidas en el corte de metales en función de los espesores de laminados. El láser tiene su nicho tecnológico en los espesores por debajo del centímetro, pudiendo realizar cortes de precisión, como veremos, en espesores desde fracciones de mm hasta cuatro mm. Al hacer esta aseveración, lo hacemos teniendo en cuenta que el noventa por ciento de los sistemas comercializados se restringen a estas dimensiones.

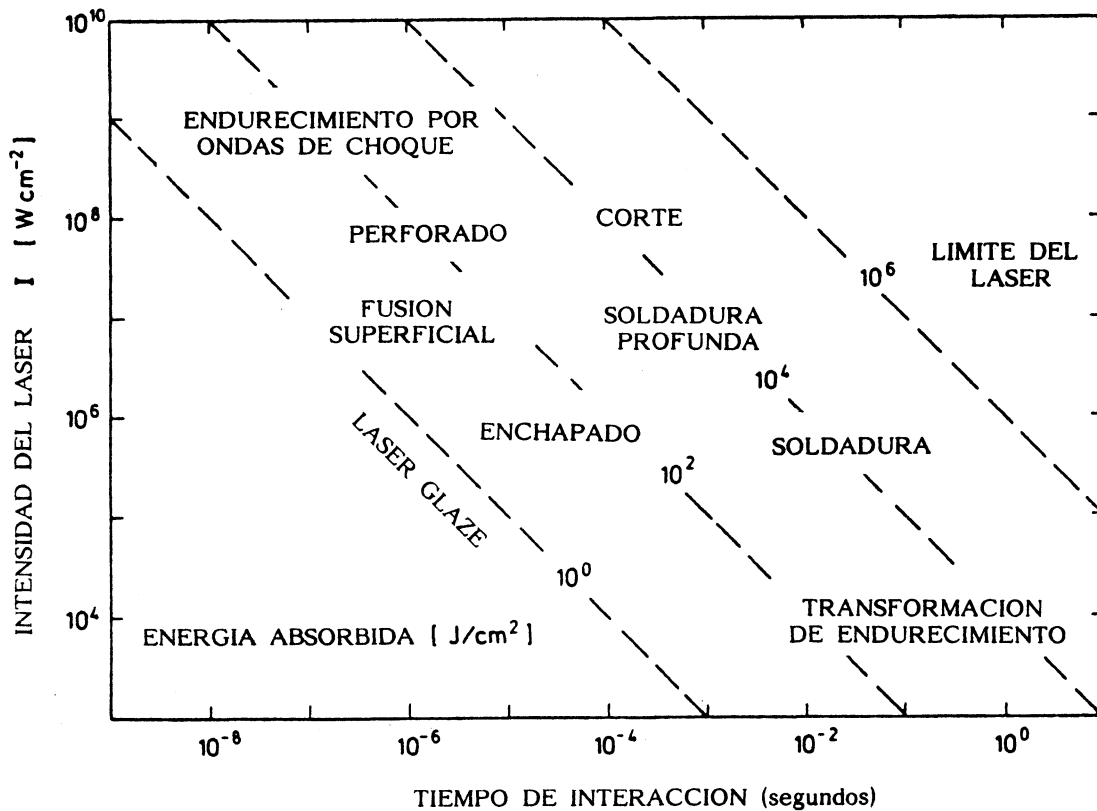


Figura 33

Si bien se fabrican sistemas mucho mayores, el uso es muy restringido y para aplicaciones especiales, con alimentadores de material diseñados para cada propósito.

Los sistemas de corte en que nos centraremos, son los sistemas híbridos de corte y los XYZ. Estos sistemas son de gran precisión. Las mesas de trabajo para los primeros, pueden llegar a 2 x 4 metros, mientras que los segundos raramente superan el metro cuadrado de superficie y poseen velocidades de trabajo limitadas. Precisiones típicas son del orden de los 100 micrones en la ubicación de un punto de coordenadas (x, y) y en la repetibilidad al encontrarlo nuevamente, luego de haberse desplazado a un extremo del plano de trabajo al otro en diagonal. Los sistemas híbridos desplazan la cabeza del láser (o un sistema de espejo) en una coordenada y la mesa en dirección perpendicular a la primera. Es-

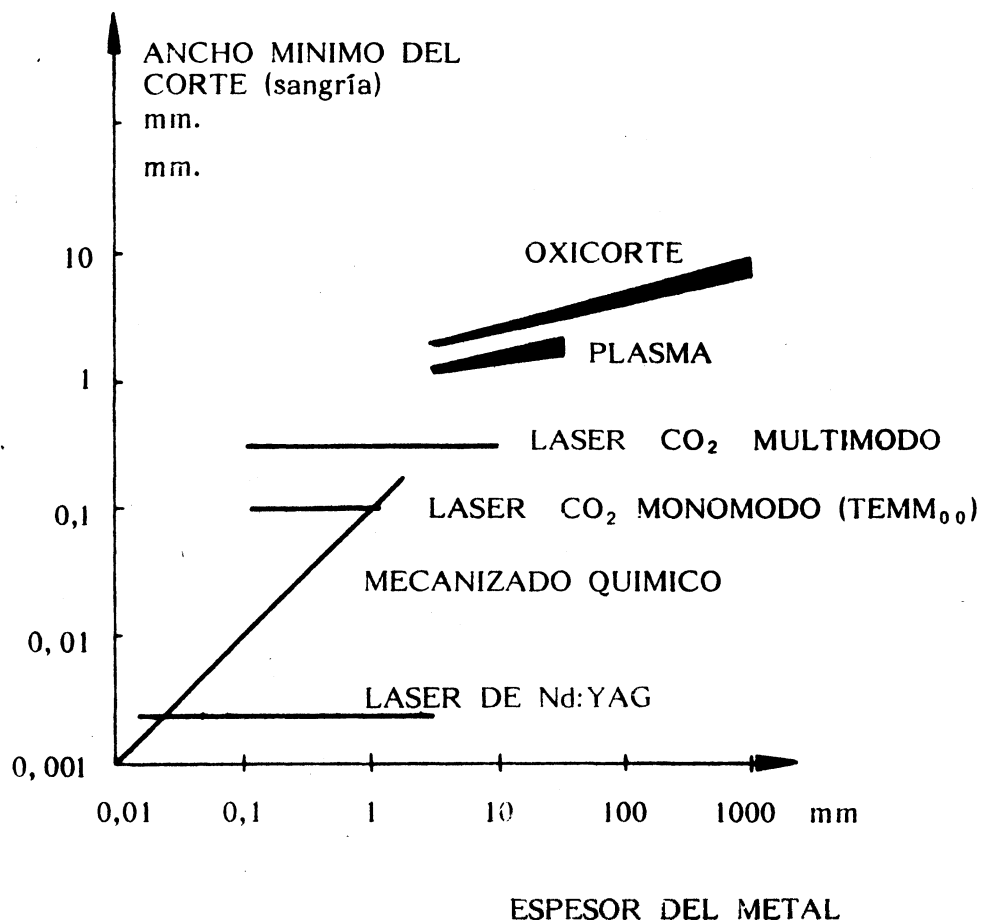


Figura 34: Nichos tecnológicos para el corte.

te sistema se concibió como alternativo a los sistemas pantográficos que desplazaban la cabeza del láser en ambas direcciones. Ellos poseían gran inercia y tenían ciertas restricciones para realizar cortes de piezas que tuvieran ángulos rectos. Los sistemas híbridos surgen como alternativa al sistema XYZ convencional, donde se consigue un sistema muy versátil, rápido, pero de menor precisión. Por último, los sistemas robotizados de 5 o más ejes, que son destinados al procesamiento (mayormente corte, perforación, endurecimiento y soldadura) de piezas de volumen.

Restringiremos nuestro estudio a un sistema específico híbrido, para corte de laminados. En particular tomaremos como ejemplo de ilustración la confección de dos piezas relacionadas con la fabricación de un motor

de corriente continua (motor DC) y de una chapa sostén para una aplicación industrial del motor. El caso que se estudia, es la generación de un prototipo de este motor. Habitualmente estos motores se realizan en laminados que luego se apilan para conformar las piezas del rotor y del estator.

Tradicionalmente, este trabajo se realizaba con la tecnología de mecanizado químico (chemical milling o etching), debiéndose realizar aproximaciones sucesivas para conseguir las prestaciones deseadas del mismo o para lograr el aprovechamiento de alguna característica del material a utilizarse. Los problemas de este método son varios. En primer lugar debemos apuntar que los cambios de geometrías en las piezas a ensayarse, exigen nuevos dibujos y nuevas impresiones de películas fotosensibles o resinas, que se imprimen sobre el laminado para su posterior ataque químico. La precisión de esta tecnología es proporcional al espesor de la pieza que se corta. Por encima de unos 300 micrones comienza el "undercutting" o ataque químico periférico. Normalmente estas tareas son de baja convergencia hacia la solución, debido a que integra diversos equipos de trabajo, normalmente no bajo el mismo techo.

Los nuevos sistemas láser han ganado en rapidez y precisión de acuerdo a lo descrito anteriormente, pero con la inclusión de un sistema de diseño asistido por computadora para el dibujo de partes y su posterior manufactura (CAD/CAM), se ha ganado en versatilidad y rapidez en la concreción de proyectos especiales. También es el caso de bajas producciones que encuentran en este sistema una respuesta casi inmediata, quedando el tiempo de las correcciones limitado al procesamiento de algunas pocas instrucciones sobre el diseño original, en contraposición con un nuevo dibujo realizado manualmente. Los nuevos sistemas no requieren siquiera la confección del dibujo sobre papel, ya que almacenados en medio magnético, son trasladados al control numérico computado (CNC) que gobierna la manufactura por medio de una interfaz y se ejecuta sin lector físico alguno. La lectura de planos fue uno de los factores limitantes de los sistemas de la generación anterior.

Debemos a priori explicar cómo influye en la calidad del corte láser el régimen pulsado de operación.

Las precisiones requeridas demandan de la manufactura, lo siguiente:

- 1) Precisión de la posición entre agujero y agujero.
- 2) Corte del agujero sin excentricidad.
- 3) Superficie cortada sin asperezas que requieran una posterior operación.
- 4) Poder cortar ángulos agudos.
- 5) Separación uniforme de corte.
- 6) Corte sin inclinación.
- 7) Rebarbas inexistentes o mínimas y fácilmente removibles.
- 8) Control de ZAC.

El modo continuo de operación del láser es utilizado principalmente para el corte de alta velocidad que no requiere grandes precisiones.

El modo pulsado es efectivo para cortar, cumpliendo con los requerimientos 1 a 8 detallados anteriormente y con precisión. Es interesante mostrar cómo al variar la potencia de pico (P_p) manteniendo la potencia media (P_a) constante, se observa un control de la ZAC. La Figura 35 muestra el corte de una pieza de 5 mm de acero de alto tenor de carbono SK3. La figura muestra que los cortes de evaluación de la

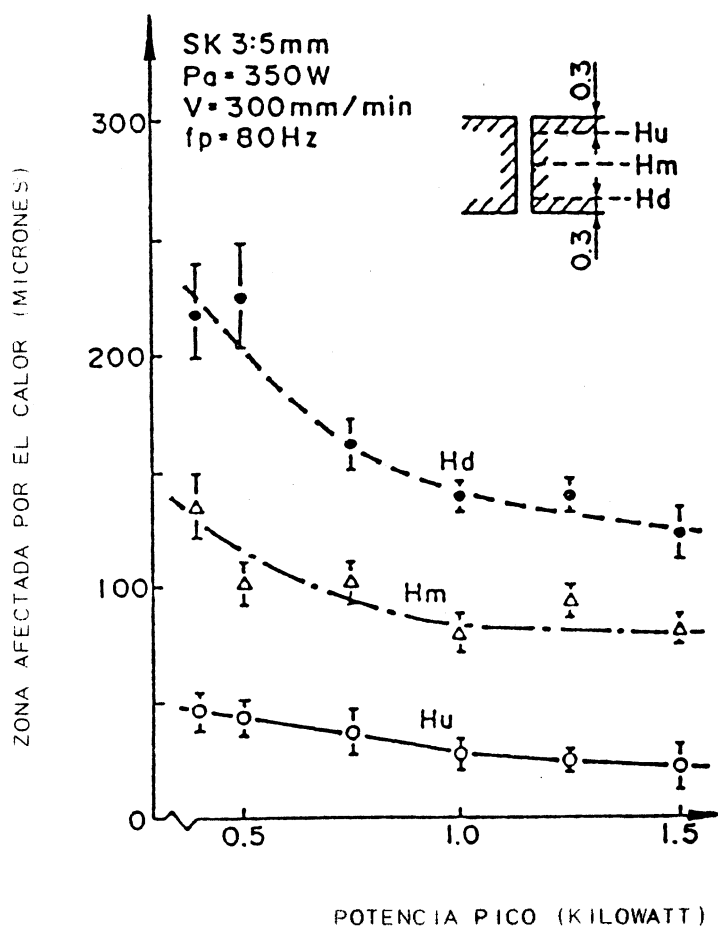


Figura 35: Relación entre la potencia pico del pulso y la profundidad de la zona afectada por el calor. M. Kanaoka, Comunicación personal, Porto, Portugal, 1988.

ZAC incrementa su penetración desde arriba hacia abajo. Las medidas se han realizado a 0,3 mm de la superficie superior (Hu), al medio de la chapa (Hm) y a 0,3 mm del borde inferior (Hd). Estos resultados muestran que pasado un cierto valor de potencia pico, la influencia sobre la penetración de la ZAC se vuelve constante, y que en su parte media (Hm) es casi constante y de alrededor de 50 micrones.

La influencia de la frecuencia de la pulsación del láser sobre la penetración de la ZAC, se muestra en la Figura 36. Puede apreciarse

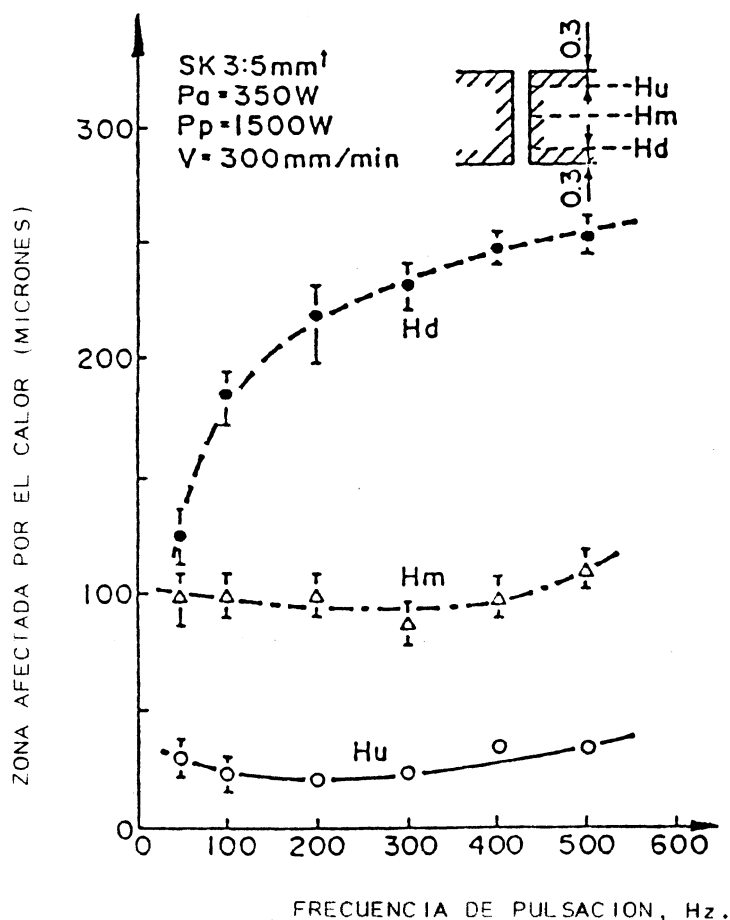


Figura 36: Relación entre frecuencia de pulso y profundidad de la zona afectada por el calor. M. Kanaoka, Comunicación personal, Porto, Portugal, 1988.

que con el aumento de la frecuencia, las profundidades de las ZAC arriba (Hu) y del medio (Hm), no varían mucho. No sucede así con la parte inferior (Hd) que se muestra, dependiente de la frecuencia hasta lograr una saturación alrededor de los 500 Hz. La calidad del corte, por su terminación, se muestra también dependiente de la frecuencia, habiéndose demostrado con la práctica que con una Pp del orden de tres veces la potencia continua necesaria para el corte, la calidad de la superficie posee una aspereza que no incomoda al tacto, lo que se acepta como para que no sea necesaria otra operación de terminación. Estas figuras fueron suministradas por M. Kanaoka [8.2].

Los ajustes manuales a ser efectuados para el procesamiento correcto son entonces:

- 1) Potencia del láser.
- 2) Operación continua o pulsada.
- 3) Gas de aporte y presión del mismo.
- 4) Ajuste del pico de salida a la focal del láser, sobre el material.
- 5) Geometría de enfoque y óptica.
- 6) Velocidad de corte.

Polarización del haz del láser.

Otro parámetro que afecta a la calidad del corte pero en menor grado, es la polarización del haz del láser. Este efecto fue descubierto por Olsen [8.3], quien puso de manifiesto que el efecto era más notable para el caso de los metales que el de los dieléctricos, pues aquéllos poseen una gran reflectividad para la radiación incidente en forma normal. Hasta el trabajo de Olsen, la polarización del haz era despreciada; su consideración ha hecho posible el corte con gran precisión.

Normalmente el ángulo de incidencia del láser se considera ser cercano a cero grado, esto es, de incidencia normal. Sin embargo, esto no es siempre cierto. En el proceso de corte, perforado y soldadura, el ángulo inicial es cero, pero cuando se forma la pileta de fundido o el frente de fusión donde el haz del láser hace contacto con el material, el ángulo es cercano a los 90°, es decir, es cercano a la incidencia tangencial, como lo muestra la Figura 37. La reflectividad de la superficie de un material irradiado con láser puede ser calculada de las fórmulas de Fresnel:

$$R_p = \tan^2 [\theta_1 - \theta_2] / \tan^2 [\theta_1 + \theta_2]$$

$$R_s = \sin^2 [\theta_1 - \theta_2] / \sin^2 [\theta_1 + \theta_2] \quad ,$$

donde R_p es la reflectividad para la luz polarizada en un plano paralelo al plano de incidencia y R_s es la reflectividad para la luz polarizada perpendicular al plano de incidencia. Los ángulos θ_1 y θ_2 son los ángulos de incidencia y el ángulo de refracción respectivamente, donde conociendo uno, puede saberse el otro utilizando la fórmula de Snell:

$$\text{sen } \theta_1 = n \text{ sen } \theta_2 \quad ,$$

donde n es el índice de refracción complejo del material, de acuerdo a lo expuesto con anterioridad. Las propiedades ópticas del material pueden ser expresadas en términos de la función dieléctrica compleja:

$$\epsilon = \epsilon_1 - i \epsilon_2 \quad \text{y}$$

$$n = [\epsilon]^{1/2} \quad .$$

La función dieléctrica varía cuando la superficie es calentada por el láser. Es conocido que el acoplamiento del haz con el material aumenta

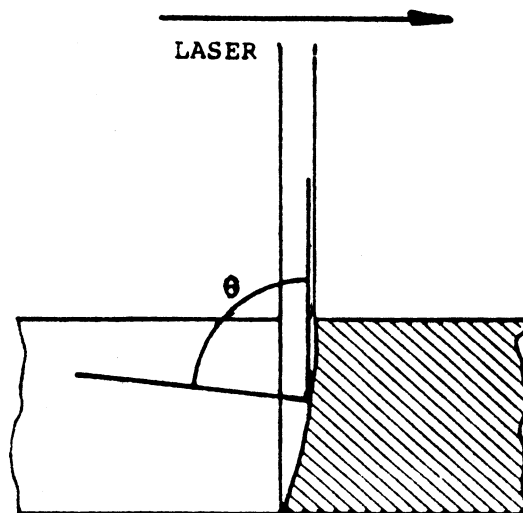


Figura 37

mucho cuando la temperatura de la superficie se acerca a la temperatura de fusión del material. La Figura 38 a, muestra la reflectividad R_0 , calculada en función de ϵ . Medidas experimentales demuestran que los valores de ϵ para el acero a $10,6\mu$ está en la región de $\epsilon_1 = -1000$ y $\epsilon_2 = 1000$. Usando las anteriores ecuaciones, se calcula R_p y R_s para una superficie de acero en función del ángulo de incidencia, que se muestra en la Figura 38 b. Para metales irradiados con el láser de CO_2 , el acoplamiento tendrá coeficientes A_p y A_s iguales a:

$$A_p = 1 - R_p \quad \text{y} \quad A_s = 1 - R_s,$$

lo cual significa que toda la luz que pase la superficie será absorbida.

Aunque la exacta variación de la ϵ no puede ser determinada en función de la temperatura, sí puede ser estimada la relación A_p/A_s . En la Figura 39 la relación A_p/A_s se muestra en función del ángulo de incidencia para diferentes valores de R_0 , donde $\epsilon_1 = -\epsilon_2$. Estas curvas muestran que la relación A_p/A_s para un dado ángulo de incidencia decrece cuando la reflectividad normal decrece, esto es, cuando la temperatura aumenta.

Es dificultoso determinar exactamente, aún en forma teórica, la relación A_p/A_s para el corte láser, donde la radiación es absorbida parcialmente por el material fundido, vapor y plasma y parcialmente por el material sólido. Sin embargo, la absorción del láser bajo ángulos cercanos a la tangencial en metales es dependiente del ángulo entre el plano de polarización y el plano de incidencia. La evaluación de la calidad del corte bajo estas circunstancias puede verse esquematizado en la Figura 40.

Estudio del corte de precisión de una pieza

La Figura 41, muestra el perfil del rotor de un motor DC (también podría ser de un motor del tipo paso a paso) cortado sobre un laminado y mostrado por transiluminación. La separación entre la pieza cortada y el remanente metálico posee una media de 40 micrones, motivo por el cual se lo muestra transiluminado y con la pieza sostenida y fijada con cinta adhesiva plástica transparente.

Aquí se presentan dos casos de trabajo: 1) La elaboración de prototipos y 2) Pequeñas producciones. En el primer caso, se realiza el trabajo sin prestar atención a los tiempos, ya que esta pieza cortada cuidadosamente insume su mayor tiempo en la búsqueda del material adecuado, su

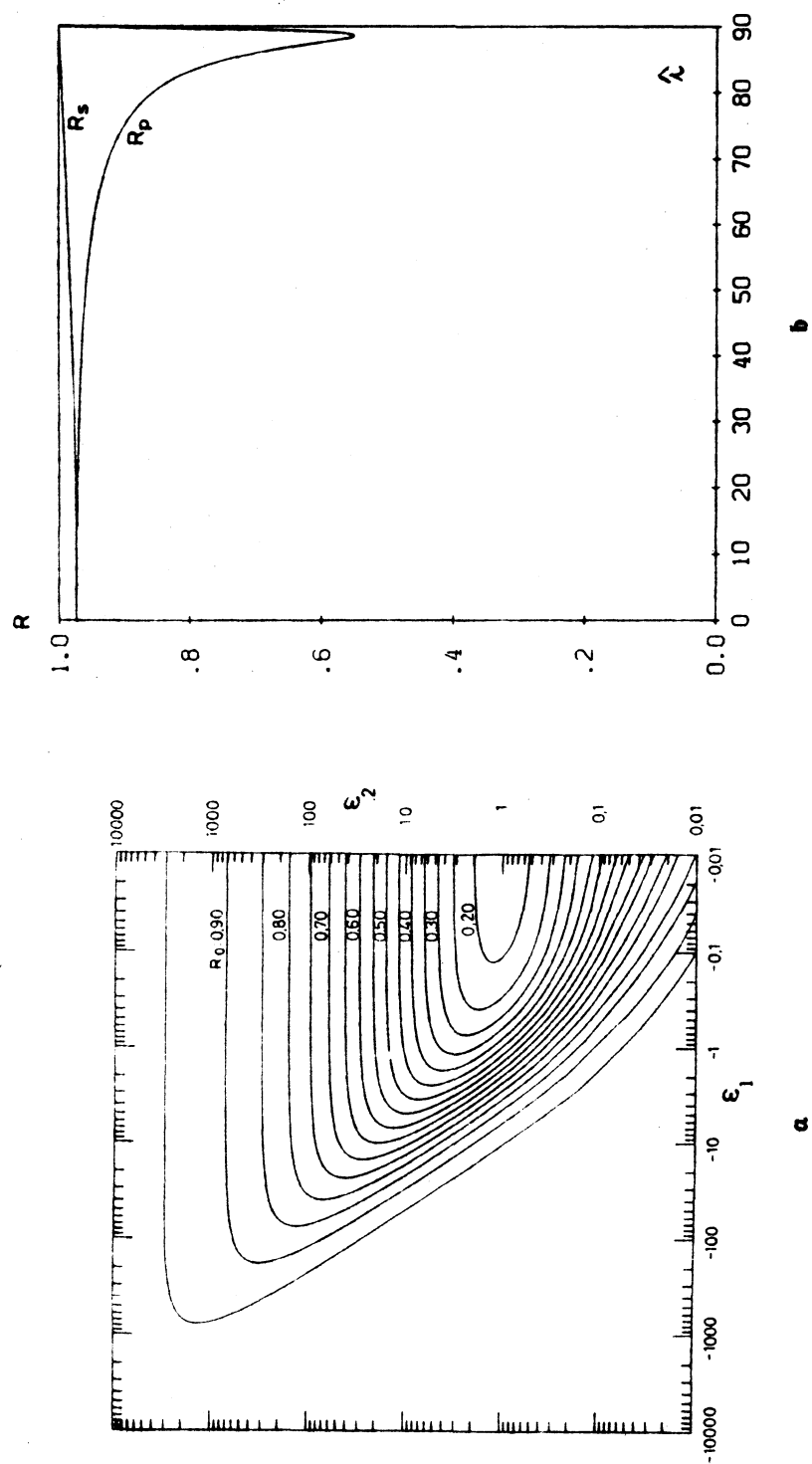


Figura 38

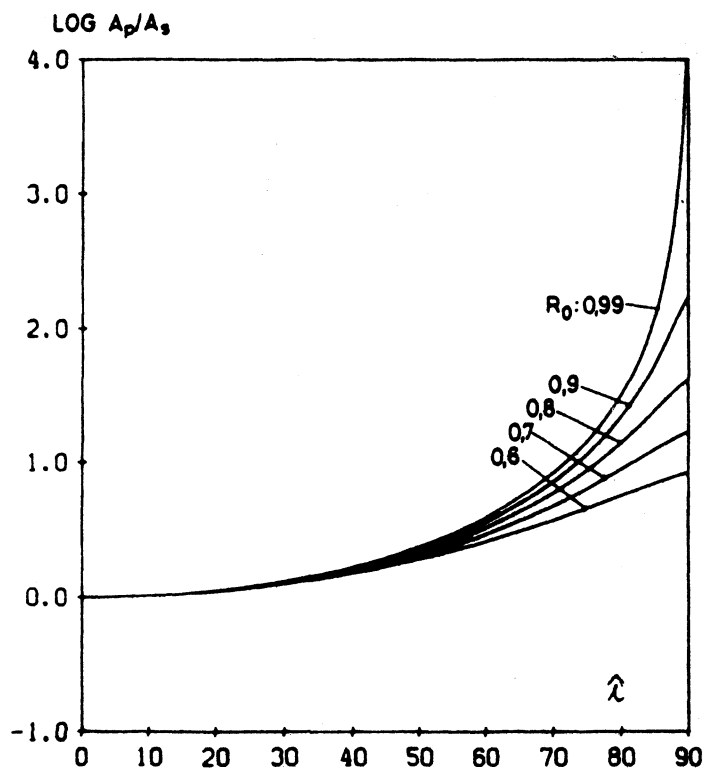
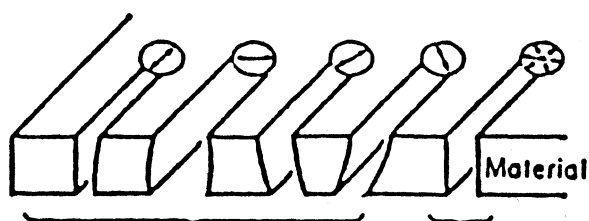


Figura 39



(a) Haz polarizado lineal
o elípticamente

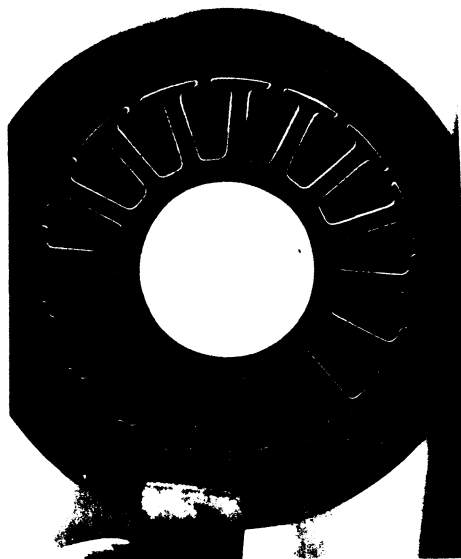
(b) Haz polarizado
circularmente

Figura 40

colocación, verificación del software, etc. El segundo caso es el que nos interesa, pues es el más importante económicamente. Es el que permite a un usuario o fabricante la manufactura en forma limitada a medida que se necesita, evitando la acumulación de stocks innecesarios. A esto se suma la posibilidad de indicar retoques en los diseños, sin necesidad de desplazarse y que los dibujos de las piezas puedan ser guardados en archivo magnético.

Método experimental

La experiencia industrial, publicada por diversos proveedores de maquinarias, demuestra, de acuerdo a lo adelantado anteriormente, que el régimen pulsado es importante para la repetibilidad del procesamiento. Desafortunadamente, los tres sistemas láser instalados en la Argentina no pudieron ser utilizados para el presente estudio. El sistema del CIOp,



a)



b)

Figura 41

si bien es sumamente versátil, no tiene la potencia suficiente para cortar chapa de espesor superior al medio milímetro. El presente estudio se realizó en los Estados Unidos y se repitió con varios proveedores para la verificación de conclusiones.

La Figura 41 a, muestra la pieza de evaluación. La Figura 41 b muestra los puntos claves en la determinación de las medidas, para conseguir la repetibilidad. Estos puntos fueron escogidos por tratarse de curvas pronunciadas, que deben realizarse de acuerdo a la especificación a diferentes velocidades. Al variar la velocidad, necesariamente estamos variando la potencia utilizada, con el fin de que la energía depositada por unidad de superficie (la focalización en este caso se deja constante, garantizando un área de irradiación constante) y por unidad de tiempo, sea la misma. Este parámetro se denomina fluencia y la repetibilidad hace posible con sentido la comparación que haremos de los diferentes intentos. Se midieron con la máquina de medir del CIOp los puntos A y B indicados en la Figura 41 b y además todos los diámetros (pata a pata) del rotor, desde el punto central del pequeño arco de cada pata del rotor. La primera observación realizada es que las máquinas de la última generación que poseen régimen pulsado no poseen un método automático de cambiar velocidad o potencia pico dejando constante la potencia media y la fluencia. Con lo que se obliga al operador a realizar por el método de la prueba y el error cuáles son los parámetros más apropiados, los que suelen ser más apropiados para una situación geométrica en particular, no generalizables en lo absoluto, si queremos conservar el valor de error en las medidas por debajo de ciertos límites impuestos por la necesidad del corte. Si bien esto es un punto que puede resolverse fácilmente, no es ciertamente algo que podamos pasar al operador de la máquina para hacer esto en forma repetitiva. Lo más lógico es hacer el cálculo y guardarlo conjuntamente con el programa de la pieza en particular para ser suministrado al CNC.

Medidas de la precisión del corte

Lamentablemente, los proveedores no abren al comprador el paquete de software, por lo que los datos que hemos acumulado, en diversas medidas, aparecen ponderados por los caprichos del sistema utilizado con una dispersión mayor de los 50 micrones requeridos a los puntos escogidos.

Los datos suministrados por M. Kanaoka permiten evaluar que el endurecimiento de la pieza cortada depende en sus partes superior, inferior y media, tanto de la potencia de pico utilizada, como de la frecuencia de la pulsación.

La Figura 42 muestra los puntos señalados del rotor del motor que



a)



b)



c)

Figura 42

fueron verificados, tal como son apreciados desde el ocular de la máquina de medir. Allí puede verse el efecto de la pulsación, que muestra en esta microescala como si hubiera sido cortado con un micro-sacabocados. En la Figura 43 mostramos cómo la potencia pico al aumentar produce menor rebarba en el corte, fundamental para un corte nítido sin segunda operación [8.4].

De lo antes expuesto deducimos que, independientemente de que las



a)



b)

Figura 43

piezas posean partes intrincadas, que son a su vez de baja disipación y que por ende deberán ser cortadas sin "quemarlas", estas piezas también poseen partes rectas y curvas que no ofrecen dificultad alguna, por lo que la programación a disponerse debe tener la posibilidad de programar el corte a diferentes velocidades para diferentes partes de las piezas a confeccionar, so pena en caso de poseerse esta facilidad de recargar temporalmente la manufactura o lo que es lo mismo decir que: con la mayoría de los sistemas a los que no puede accederse a este tipo de control, el precio de la pieza va a quedar condicionado por la parte que exija mayor cuidado y no por el número de partes intrincadas que ella posea. Las piezas pueden poseer pequeños agujeros, como el mostrado en la Figura 43 donde puede ser importante el endurecimiento del mismo de acuerdo a la función. La Figura 44 muestra pequeñas piezas con corte de calidad con una magnificación de diez veces su tamaño natural.

Velocidades de corte

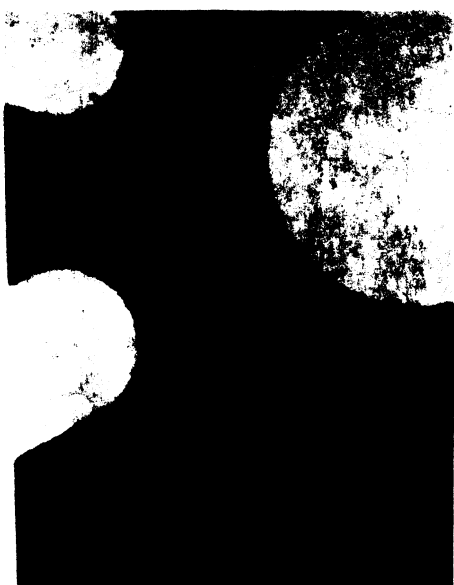
Como hemos visto al estudiar las propiedades ópticas de los metales, la reflectividad depende de la frecuencia de plasma, lográndose una mayor penetración en el titanio y los materiales ferrosos. El Apéndice I muestra una clasificación de la capacidad de procesamiento de diversos materiales. Para el caso específico del corte, la Figura 45 muestra cómo el régimen continuo de operación la eficiencia del corte, medida por la velocidad del mismo, depende del espesor del material. En la figura se muestra para el Ti, aceros de bajo tenor de carbono y acero inoxidable (304), con el láser operando a 200 W.

Para los sistemas de última generación, con un kilovatio de potencia, pueden obtenerse los 14 metros por minuto para espesor de 1 mm de acero de bajo tenor de carbono.

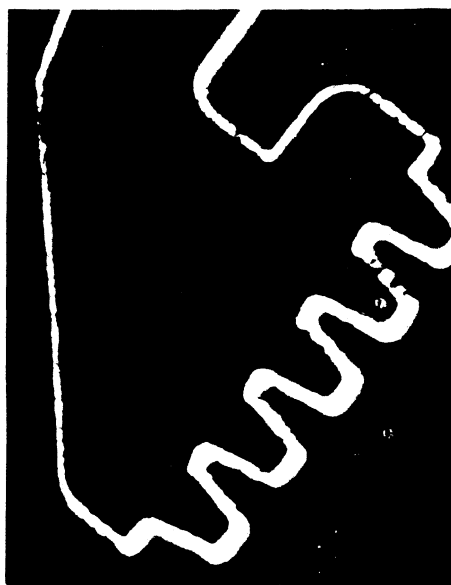
La velocidad de corte en función del espesor del material, depende no sólo de la reflectividad del material a una dada temperatura, sino además de la penetración del haz del láser en el material en cuestión. Veremos a tratar el tema de la profundidad con mayor detalle, al exponer la soldadura-láser, pues es en ese caso más crítica que en el del corte de metales.

8.2. Perforado

El uso de los láseres para aplicaciones de perforado cubre las expectativas de las tecnologías de punta. Se debe esto a que pueden aplicarse



a)



b)



c)



d)

Figura 44

con solvencia especificaciones del agujero a realizarse tales como: diámetro, profundidad, conicidad, refundido (metal vaporizado que no es expulsado y resolidifica alrededor y dentro del agujero), microfisuras, ángulo de la perforación en relación a la normal de la superficie de la pieza.

Las variables del láser que producen efecto sobre uno o más de estos requerimientos son el pulso, la energía, la duración, el diámetro del haz en la lente (apertura) y la calidad del haz. Todas las ventajas apuntadas ya para el corte, son en principio extrapolables aquí para el perfo-

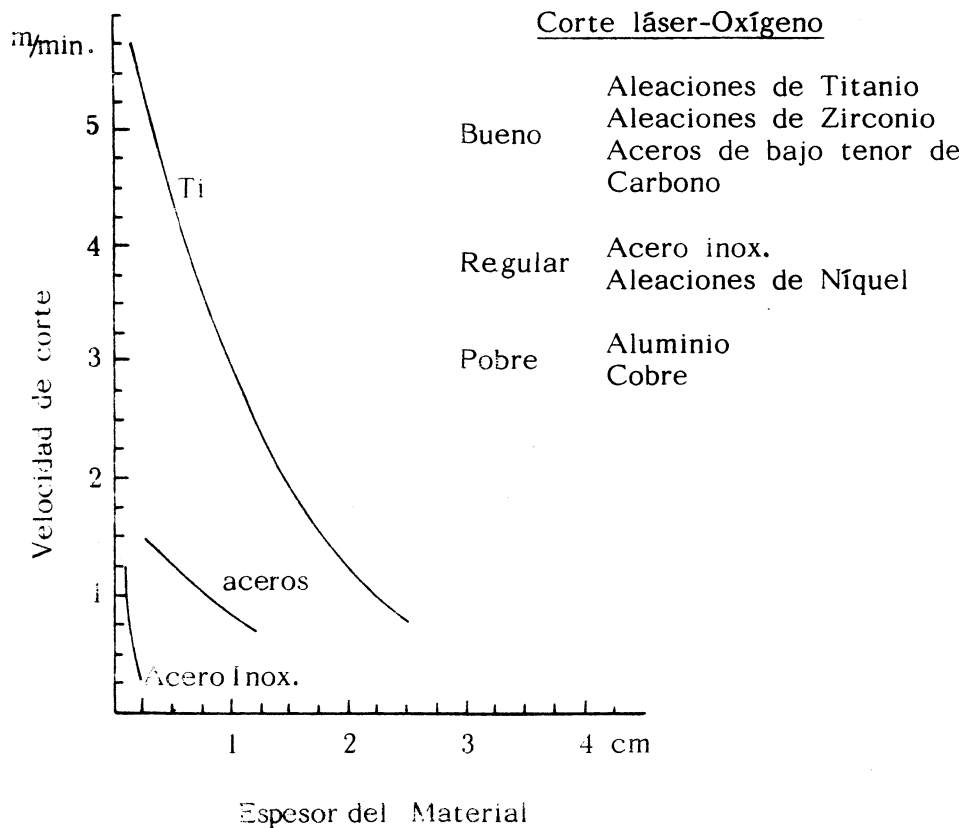


Figura 45

rado. Deberíamos añadir que las perforaciones en ángulo se realizan con gran rapidez con el láser y se cumplen las especificaciones con gran precisión, pues no existe el desplazamiento lateral de la herramienta, defecto típico en estos casos cuando se utiliza una broca. Esta ventaja se hace más notable cuanto menor es el diámetro a perforar y mayor la inclinación requerida. El límite práctico de agujero pequeño a realizarse con el láser de CO_2 está limitado por la difracción y el modo de oscilación de operación del láser.

8.3. Endurecimiento superficial de aceros

El endurecimiento del acero o del hierro por calentamiento al rojo y enfriado rápido ha sido conocido por siglos. Como ya hemos explicado al estudiar las transformaciones de fase, el acero en su forma de austenita, se transforma a una estructura dura (martensita) si el enfriamiento se hace lo suficientemente rápido como para pasar la "curva de Bain" o "codo de Bain" del diagrama T-T-T. Esta transformación de endurecimiento es extensivamente utilizada en la industria para mejorar las características de desgaste, resistencia a la erosión y durabilidad de las partes. En la Figura 24 la temperatura de austenización es aquella (para una aleación dada) que está por encima de la línea A_3 , donde la austenita es una fase estable.

A diferencia de las aleaciones que se pueden endurecer por precipitación, las cuales permanecen "blandas" después del enfriado rápido y que endurecen cuando la precipitación comienza, los aceros son endurecidos por enfriado rápido y "ablandados" por precipitación. La razón de esta diferencia es la transformación alotrópica del acero que conduce a la formación de martensita. Esta última puede ser "ablandada" por precipitación de carbono como carburo de Fe; el proceso es denominado templado. El templado de baja temperatura produce una estructura de dos fases finamente dispersas de carburo en martensita de bajo tenor de carbono. A temperaturas más elevadas, el contenido de carbono de la martensita es reducido aún más, hasta que a una temperatura un poco por debajo de la eutéctica, el templado resulta en una microestructura de pequeñas partículas de carburo en una matriz de ferrita; esta estructura es denominada esferoidita.

Variaciones de estos ciclos de enfriado rápido y templado pueden también ser utilizados. El austemplado se realiza enfriando el acero hasta justo encima de la temperatura donde comienza la formación de martensita, M_s , y después mantenerlo isotérmicamente hasta la formación de bainita. El martemplado se consigue enfriando rápidamente hasta M_s , y luego que el equilibrio térmico se ha obtenido en el material y un enfriado rápido nuevamente hasta el M_f para la obtención de martensita.

El ausformado se consigue con el enfriamiento rápido hasta justo encima de M_s y después trabajar la aleación hasta distorsionar la austenita y finalmente enfriamiento rápido hasta la formación de martensita. Estos dos últimos procesos habitualmente son seguidos de un templeado.

Estos tratamientos descriptos, en forma convencional, son realizados en volumen, por lo que la fragilidad de las piezas tratadas térmicamente sin templeado hace peligrar las funciones de trabajo de las mismas, ante la eventual fractura. Esto fue solucionado con la introducción del endurecimiento superficial por medio de radiofrecuencias, tecnología con la cual compite el endurecimiento con láser, debido a que la mayoría de los cortes con láser son tratados térmicamente con el mismo, in situ, cuando es requerido, sin la necesidad de hacer la operación en otro lugar, a posteriori, evitando traslados y demoras, procesándose en forma simultánea varios procesos.

El endurecimiento con láser es limitado por la difusión térmica del material. Así que, si la densidad de potencia del láser es incrementada, la profundidad del endurecimiento disminuye debido al aumento de temperatura más rápido de la superficie. En el otro extremo, bajas densidades de potencia y larga exposición del material producen una difusión térmica sustancial hacia el interior del material. Es decir, se termostatiza. Bajo estas condiciones, cuando el haz del láser pasa, el gradiente térmico no es suficiente para proveer un enfriamiento rápido de la superficie.

En los aceros se utiliza fosfato de manganeso para prevenir la oxidación superficial. Pinturas con pigmentos de óxidos son buenos absorbentes. Una película negra, por ejemplo, es un buen absorbedor. Para ilustrar el hecho de que la apariencia visual no necesariamente identifica las propiedades del infrarrojo, también una película blanca es absorbente e igualmente efectiva.

Cuando se utilizan películas para mejorar la absorción, debe notarse que se crea al mismo tiempo una barrera térmica que afecta las condiciones de enfriamiento. Así, la repetibilidad del procesamiento es extremadamente dependiente de la producción de películas reproducibles. En este aspecto, el láser está en desventaja con el método de haz de electrones (electron-beam), que no es reflejado por las superficies metálicas.

Veremos el endurecimiento superficial del acero SAE 1050 con láser. Los aceros más importantes a ser procesados son aquéllos que luego están sujetos a fricciones a altas temperaturas. Anticipamos la exposición, diciendo que el endurecimiento con láser de metales, usualmente provee la máxima dureza para la mayoría de ellos. Una ventaja adicional consiste en que por este método no se distorsionan las superficies ni se modifican sus medidas, por lo que no se requiere de maquinado post-tratamiento, como sucede con otros tratamientos, por ejemplo la aplicación de cromoduro industrial.

La Figura 46 es un diagrama T-T-T para el acero 1050. Este

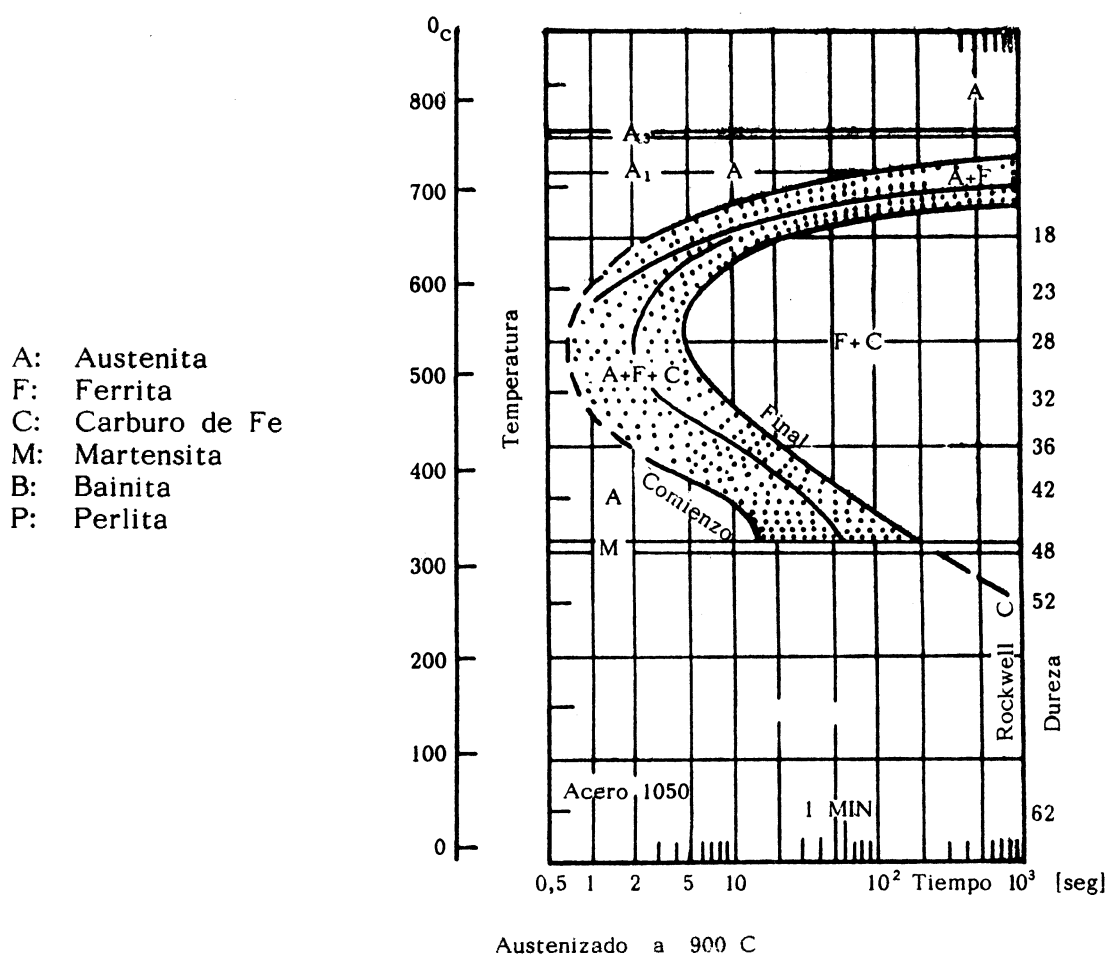


Figura 46

tipo de acero altamente reflectivo a temperatura ambiente requiere del tratamiento superficial para mejorar la absorción. Usualmente se utiliza fosfato de manganeso o una pintura uniforme a base de carbón.

La transformación de endurecimiento en este tipo de material, es altamente dependiente del contenido de carbono en el metal (martensita). Los gráficos de la Figura 47 ilustran la dureza en grados Rockwell C para diferentes porcentajes obtenidos de martensita para diferentes contenidos de carbono.

Veamos ahora el endurecimiento del acero SAE 1045. Esencialmente, el problema del tratamiento de este acero es que a altas temperaturas, la formación de austenita, así como la redistribución de carbono,

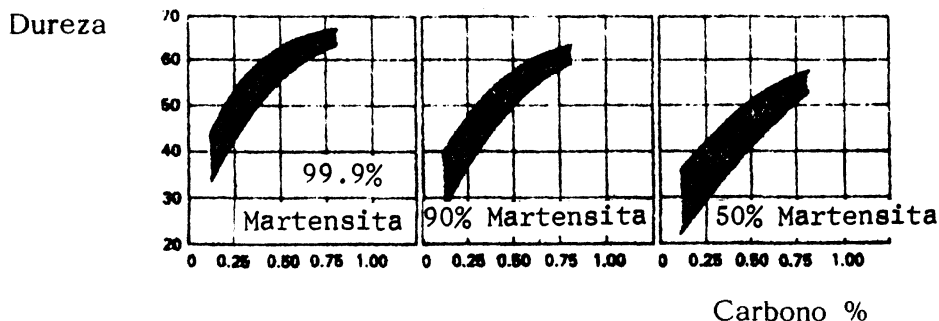


Figura 47

necesario para formar la solución sólida Fe-C , lleva un tiempo pequeño pero finito. La técnica de endurecimiento se realiza exponiendo la superficie del metal al haz del láser fijo o viceversa, de tal modo que la velocidad con que se expone el metal esté relacionada con los tiempos convenientes para este procesamiento.

Cuando la capa de acero a tratarse es calentada por el láser, la austenita comienza a formarse a la temperatura T_1 . A medida que más material se transforma en austenita, el material alcanza la temperatura T_3 , a la cual el proceso está completado. A una temperatura entre T_1 y T_2 sólo una fracción del material sufrirá la transformación, independientemente del tiempo que la pieza de trabajo sea mantenida a esa temperatura. Esto es, entre las temperaturas T_1 y T_3 , dos fases están en equilibrio y la proporción entre las dos fases depende de la temperatura. Para un acero al carbono con 0,4 de C , la temperatura T_1 es de 727 C y la T_3 de 800 C . Si la superficie del acero es calentada rápidamente con una potencia de 2250 W/cm² y el rayo es retirado para permitir el autotemplado por conducción, el material se endurece por la formación de martensita. Como ilustra la Figura 48, todo el material que ha superado la temperatura T_3 se endurecerá completamente. El material que no superó tal temperatura mostrará progresivamente, de la superficie a su interior, menos dureza, hasta el punto que alcanzó la temperatura T_1 . En este punto no hubo endurecimiento, pues no se alcanzó a formar la austenita. Así que tendremos una zona superficial de dureza y una zona de transición, cuyo ancho depende, para un metal dado, de la potencia por unidad de área irradiada y de la velocidad de ba-

ruido de exposición del material. En estos procesos, la temperatura de transición de fase, el calor adicional para la transformación, la conductividad térmica, capacidad calorífica y la difusividad, entran en juego para un análisis cuantitativo del problema.

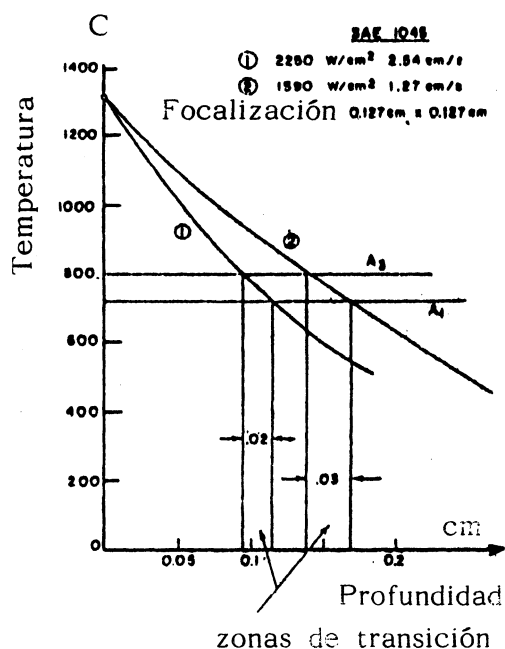


Figura 48

¿Qué es lo que hace del tratamiento térmico con láser algo atractivo desde el punto de vista de la producción? Pues simplemente la habilidad de producir el tratamiento en forma repetitiva, precisamente, en forma localizada, en espesores delgados con una estructura de grano muy

fina y con demanda mínima de energía.

Los tratamientos térmicos con láser fueron descritos por primera vez por L.I. Mirkin y N.E. Pilipetskii [8.5] en la Unión Soviética en 1966. El primer intento fabril fue realizado por el Centro de Investigaciones de la General Motors Company de Warren, estado de Michigan en los Estados Unidos, en 1973. Estos estudios terminaron en la aplicación de láseres para el endurecimiento en producción de engranajes de transmisión en 1975 y en el endurecimiento de camisas de cilindro de automóviles y locomotoras en 1978, donde la compañía afirma que esas partes poseen, producidas de esta forma, una vida útil mayor del 30% . El primer proveedor independiente del servicio de tratamiento térmico con láser data de 1983 en el estado de Kentucky [8.6]. Las industrias automotrices del Japón han ensayado con éxito el sinterizado en caliente de engranajes y el posterior endurecimiento con láseres; si bien esto se admite como una realidad, no existen en nuestro conocimiento publicaciones que lo confirmen [8.7]. Se asegura [8.8] que solamente unos 40 sistemas láser están funcionando en los Estados Unidos para tratamientos térmicos, representando el 0,5% del total de las instalaciones con el mismo propósito, de todas las tecnologías juntas. El método del haz electrónico (HE) (electron beam) posee el mismo porcentaje del mercado, contando este último sin embargo con la desventaja de la operación en vacío, lo que representa una pérdida de tiempo en el ciclo de producción. Sin embargo el HE posee la ventaja de la soldadura en grandes espesores, motivo por el cual muchas de las instalaciones son utilizadas en forma simultánea o alternativamente para endurecer metales. Esto lo veremos en la sección siguiente con mayor detalle.

Para finalizar este tema, digamos algo específico de la metalurgia del endurecimiento con láser. Por lo expuesto, hemos visto que no hay diferencias fundamentales entre el endurecimiento tradicional, es decir, el endurecimiento en volumen y el endurecimiento superficial, ya que ocurren las mismas reacciones metalúrgicas. Sin embargo, en el endurecimiento superficial el tiempo del ciclo de calentamiento se realiza en un tiempo mucho más corto que en el endurecimiento tradicional. Esto es particularmente así para el caso del endurecimiento con láser, donde el calentamiento hasta las temperaturas de austenizado ocurren en un segundo o aún en fracciones de segundo. De hecho, en este caso, el período de calentamiento es menor que el período de enfriamiento.

La respuesta de los materiales ferrosos al enfriamiento rápido desde la región de austenita ha sido estudiada con gran detalle por muchas décadas y está muy bien entendida. No puede decirse lo mismo, sin embargo, del período de calentamiento del proceso. Los problemas asociados con calentamientos muy rápidos, es que la formación de austenita, así como la redistribución del carbono, necesaria para la formación de una solución homogénea sólida de $Fe\gamma-C$, requiere un tiempo pequeño pero finito. La cinética de estos procesos a velocidades muy elevadas

de calentamiento es todavía incierta. La reacción básica que ocurre en el calentamiento es la transformación de $Fe\alpha$ en $Fe\gamma$. Esto ocurre por nucleación y crecimiento de una nueva fase en la matriz de la vieja fase. En el proceso de calentamiento convencional (lento), esto comienza a 723 C , como fue explicado anteriormente en la línea A_1 y en un acero al carbono se completará en la línea A_3 . Sin embargo, cuando el calentamiento es muy veloz, el sistema está muy lejos del equilibrio y la línea A_3 tiende a desplazarse a temperaturas más elevadas. Así puede darse el caso que el nivel de temperatura proyectado no resulte lograrse aún para comenzar la austenización. Por eso los parámetros para el endurecimiento láser se diseñan para lograr temperaturas pico muy por encima de los valores convencionales para lograr la austenización previa al enfriamiento. La estructura del hierro y el acero a temperatura ambiente, en condiciones de equilibrio, contiene carbono en forma de carburo de hierro o grafito como una fase separada. Para lograr el endurecimiento total al enfriar rápidamente, este carbono debe disolverse totalmente en la fase de austenita. Para esto, el carbono se debe redistribuir por difusión a las áreas de ferrita, libres de carbono. Siendo este proceso dependiente del tiempo, aún a las temperaturas elevadas utilizadas en el endurecimiento-láser, bajo ciertas condiciones, puede el procesamiento ser tan rápido que no permita una difusión completa. Esta inconveniencia, obviamente, es más importante cuando la distribución de carbono inicialmente encontrada en el material no es uniforme, como en los casos de las estructuras perlíticas gruesas, estructuras con cementita proeutectoide, materiales globulizados y aceros fundidos con alto contenido de ferrita libre. La Figura 49 muestra la estructura de un acero fundido dúctil después del procesamiento láser. La extensión de la difusión de carbono que ocurrió durante el procesamiento está notablemente marcada por la región de martensita rodeando los nódulos de grafito.

Las temperaturas elevadas requeridas por el endurecimiento láser, combinadas con la necesidad de permitir un tiempo suficiente encima de la línea A_3 para formar una austenita homogénea resulta en altos picos de temperatura. En tales circunstancias puede haber una tendencia a formar una capa fina (de 5 a 20 micrones de espesor) superficial, de una fase que aún no ha sido categorizada precisamente. Austenita retenida ha sido encontrada en materiales endurecidos con láser y detectada por medio de análisis por Rayos X; lo que no está bajo estudio es cómo los parámetros de procesamiento afectan a la austenita retenida.

El límite superior de la transformación de estado sólido para el endurecimiento, lo da ciertamente el punto de fusión del material, ya que esto es indeseable para este procesamiento. Esto es fácil en piezas planas, pero no es lo mismo cuando se trata de piezas en volumen con diferentes formas intrincadas que pueden poseer diferentes disipaciones. Otro inconveniente asociado a este procesamiento es la necesidad de solapar diferentes "pasadas" del haz del láser debido al carácter secuencial del procedimiento. Debido al flujo lateral del calor del área de irradia-



Figura 49

ción efectuada por el láser, es inevitable el templado de las áreas previamente endurecidas por una pasada anterior adyacente. Este efecto es mayor cuando menor es la velocidad de procesamiento. Estos inconvenientes pueden obviarse mediante la elección de la óptica adecuada o la forma de efectuar el procesamiento, de forma tal que la textura final no condicione la funcionalidad de la pieza tratada.

Los parámetros que deben fijarse para el procesamiento de endurecimiento láser son:

1) Densidades de potencia del orden entre 500 y 5000 W/cm². Esto corresponde a tiempos de interacción entre 0,1 y 10 segundos.

2) Materiales muy endurecibles deben ser procesados a bajas densidades de potencia y relativamente largos tiempos de exposición, mientras que materiales poco endurecibles deben ser procesados a altas densidades de potencia y cortos tiempos de exposición.

3) La temperatura máxima de la superficie es proporcional a la raíz cuadrada de la velocidad de exposición, así que al duplicar la potencia se requiere el cuadruplicar la velocidad de exposición para obtener temperaturas máximas equivalentes en la superficie.

4) El aumento de la densidad de potencia resulta en menor ingreso de energía en el volumen del sólido para la misma temperatura de la superficie.

5) El acero, cuando está normalizado, templado o con estructuras globulizadas, acero con cementita proeutéctica, hierros fundidos con carburos aleados en forma estable, requieren tiempos de exposición mayores que los aceros que han sido endurecidos y templados.

6) Las piezas pequeñas requerirán mayores densidades de potencia y menores tiempos de exposición que las piezas mayores, salvo que se utilice un disipador acoplado a ellas.

A modo de ilustración podemos ver para el acero 1035 cómo con irradiancias de 0 a 8000 Watt/cm² y teniendo como parámetro a la temperatura de la superficie, ver Figura 50, cómo los diferentes tiempos de interacción producen calentamiento del material por encima de la temperatura de austenizado a diferentes profundidades. A mayor exceso térmico de la línea A₃ mayor capacidad de austenización y de homogeneización por difusión. El endurecimiento superficial irá disminuyendo hacia el interior del material [8.8-9].

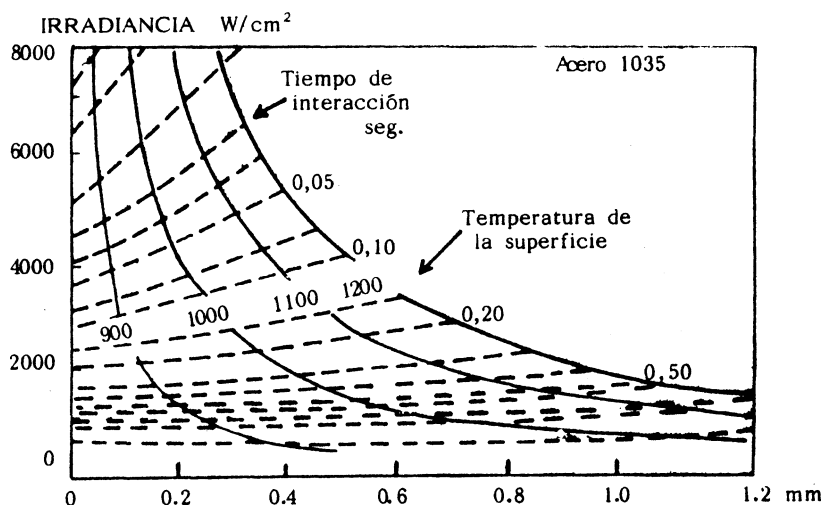


Figura 50: Zona calentada por encima de la temperatura de austenizado.

8.4. Soldadura

Los láseres han sido utilizados desde sus comienzos como útiles herramientas para soldar. Hasta el año 1970, las soldaduras fueron restringidas a materiales relativamente delgados, debido a la limitada potencia disponible en forma continua, de unos pocos vatios. La aparición de láseres de CO_2 de muchos kilovatios de potencia, ha incrementado las posibilidades y el interés en la soldadura láser (SL), lo que ha hecho que estas máquinas pasaran de curiosidades de laboratorio al uso diario industrial.

La capacidad de generar grandes densidades de potencia (mayores que 10^6 Watt/cm^2) es el factor primordial de aplicación para la SL. Esta capacidad se ha visto incrementada desde la publicación de C.O. Brown [8.10] y la de C.O. Brown y J.W. Davis [8.11], en 1970. El primer informe con una evaluación profesional de envergadura, en el área de soldadura, aparece en el Informe de la United Technologies, escrito por C.M. Banas en 1977 [8.12]. El mismo autor compara las técnicas de haz electrónico (HE o "electron beam"), láser y arco en [8.13]. Un documento lleno de rico contenido es la tesis de J. Mazumder, realizada en la Universidad de Londres [8.14] dirigida por el Dr. Ing. William Steen, un innovador en varios campos del procesamiento con láseres [8.15]. Posteriores informes de Banas, Mazumder y Arata, este último de la Universidad de Kyoto, han presentado líneas claras en la sedimentación tecnológica en soldaduras y la operación económica con láseres. Resumiremos brevemente sus resultados [8.16-20], ya que esta literatura, en general, no es de fácil acceso en nuestras bibliotecas. Las referencias provistas no pretenden ser exhaustivas, pero sí muy representativas.

Resumiremos las ventajas de la soldadura con láser:

- 1) La luz no posee inercia, transporta energía sin aporte de masa, por lo que permite alta velocidad de proceso, donde el arranque y detención instantáneo del mismo son posibles.
- 2) La focalización del láser posee la fuente térmica de mayor concentración conocida por tecnología alguna.
- 3) La soldadura puede realizarse a presión atmosférica, compitiendo así con la de HE.
- 4) Materiales de alto valor tecnológico como Ti, cuarzo, etc. pueden soldarse.
- 5) Las piezas a soldarse no necesitan estar firmemente sujetas.
- 6) No hacen falta electrodos o adición de materiales.
- 7) Pueden hacerse soldaduras muy estrechas.
- 8) Se pueden realizar soldaduras muy precisas.

9) Pueden realizarse soldaduras de baja contaminación.

10) Las adyacencias de la soldadura (o corte) son poco afectadas por elevación térmica.

11) Pueden realizarse formas intrincadas de soldaduras (o corte) utilizando técnicas de reflexión de luz.

12) Pueden realizarse técnicas de tiempo compartido del haz del láser en los procesos.

13) Pueden soldarse piezas de materiales disímiles o de masas muy diferentes, sin distorsión de formas. Esta última propiedad es sólo viable con láseres.

14) La SL es muy propicia para la robotización del proceso.

La máxima penetración en SL informada hasta 1977, fue la realizada a 5 cm de espesor en placas de acero inoxidable, utilizando un láser de CO_2 de 77 kW [8.12]. Sin embargo, hoy en día se obtienen soldaduras de ese espesor con potencias más bajas, debido a la preparación de las superficies y a las técnicas de regímenes continuo y pulsado en forma simultánea. Con la técnica de HE se obtienen soldaduras más profundas, pero requieren que el trabajo sea realizado al vacío, lo que crea un serio inconveniente desde el punto de vista industrial.

El mecanismo de la penetración de la energía en la SL es muy similar al que se encuentra en el método de HE, esto es, la transferencia energética vía "key hole" [8.21]. Este efecto se produce cuando el haz posee suficiente densidad de potencia como para producir la vaporización del material y la presión producida por el vapor en el cráter, causa el desplazamiento hacia arriba del metal fundido bordeando las paredes del agujero. Este agujero actúa como un cuerpo negro radiante y ayuda a la absorción de la radiación del láser, así como a la distribución del calor del material.

En las soldaduras convencionales, el proceso de soldadura se realiza por la energía depositada en la superficie de las piezas a soldar y ésta es transportada al interior de las mismas por conducción.

Una vez formado el "key hole", los procesos de conducción son secundarios, ya que la fusión del material produce el elevamiento de la absorción en forma considerable. La elevación paulatina de la densidad de potencia conduce inevitablemente al corte. En este control el auxilio de los gases de aporte juega un papel esencial que explicaremos después.

La Figura 51 ilustra el fenómeno y correlaciona el efecto de incrementar la densidad de potencia. La línea a trazos indica que la absorción del láser de Nd:YAG que emite en 1,06 micrones de longitud de onda (infrarrojo cercano) es menos reflejada que la del CO_2 . A pesar de esta ventaja aparente, los láseres de CO_2 son más económicos hasta el momento por la diversidad de tareas que pueden realizar por el aporte total

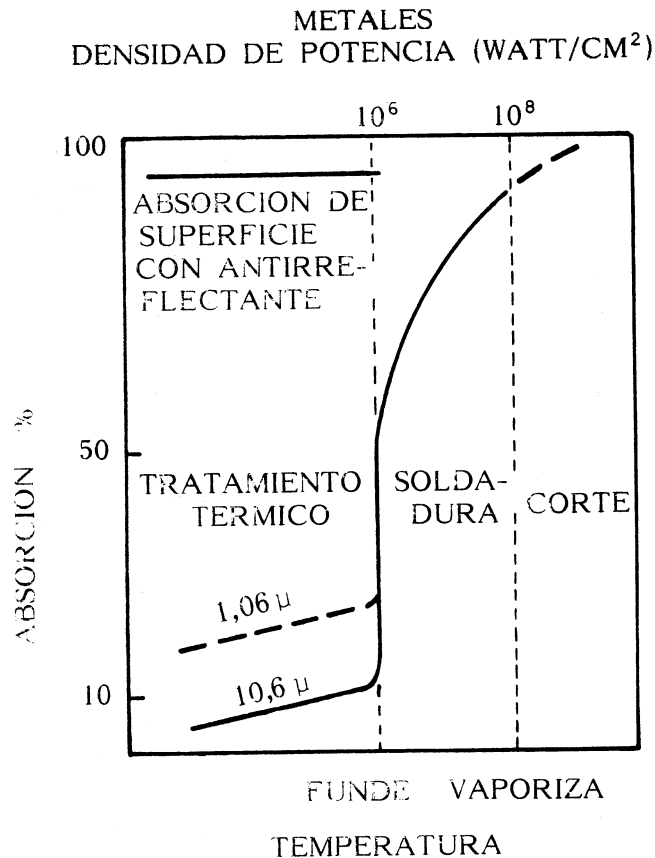


Figura 51: Absorción de luz infrarroja vs. densidad de potencia.

de energía que pueden proveer al procesamiento. Por debajo de 10^6 Watt/cm², la superficie sin tratamiento del metal absorbe sólo el 10% de la energía del láser (o menos) a la longitud de onda de 10,6 micrones. Este nivel de potencia es, sin embargo, suficiente para tratar superficies de metales tratadas con una película antirreflectora que favorece fuertemente la absorción. Entre 10^6 y 10^8 Watt/cm² de potencia por unidad de superficie, el metal se encuentra en estado líquido, absorbiendo entre el 60 y el 80% de la energía dirigida a la piletta de fundido. Por encima de estos valores, el rayo vaporiza el metal y corta.

Las condiciones de energía y flujo de material durante la SL han sido investigadas por Klemens [8.22]. El "key hole" es formado sólo si el haz del láser posee suficiente densidad de potencia. Esto se obtiene a potencias de operación continua superiores al kilovatio de energía, como puede apreciarse en la Figura 52. Este mecanismo marca los dos modos de soldadura posibles: el limitado por la conducción para potencias menores al kilovatio y el de penetración profunda vía "key hole".

Aunque existe una considerable literatura con información teórica y empírica describiendo las prestaciones de los diferentes tipos de láseres de diversos laboratorios, prestadores de servicios y de los mismos fabricantes, muchas veces las comparaciones y las estadísticas no son viables. Las variaciones en la operación que ya hemos explicado al tratar los modos de los láseres, son de gran significación al escoger un proveedor para tareas específicas. Por ejemplo, un láser de CO_2 de 1 kW de potencia con tecnología de flujo axial puede proveer mayor penetración que un idéntico láser de flujo transversal. Esto resulta de las pequeñas variacio-

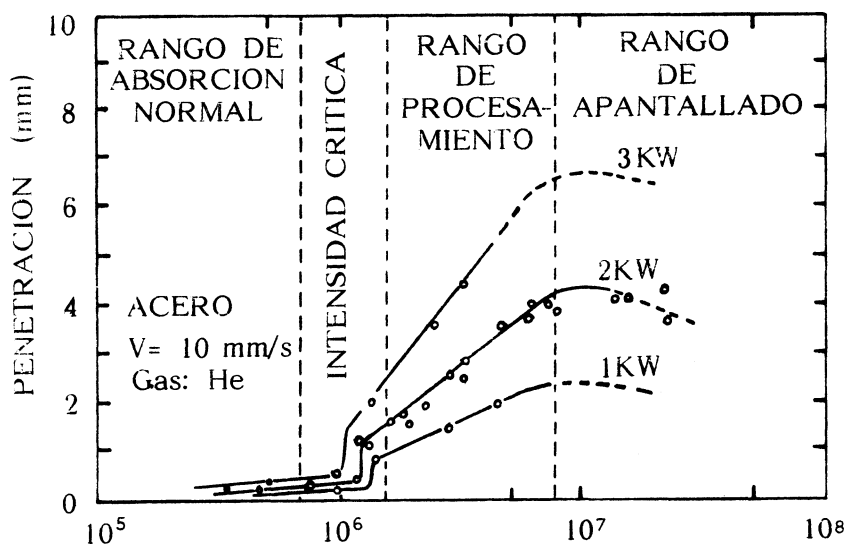


Figura 52: Acoplamiento con la muestra en función de la energía incidente, definiendo rangos.

nes que ocurren en la divergencia del haz láser o a la vibración de la óptica debido a la microfónica inherente a esta última tecnología.

Durante la soldadura, la cavidad del "key hole" está rodeada de líquido, el que a su vez está rodeado de material sólido. El flujo de líquido y la tensión superficial tienden a cerrar la cavidad, mientras que el vapor que es continuamente generado tiende a mantener la cavidad. Hay un continuo flujo de material hacia afuera de la cavidad en el punto donde incide el láser. Para un haz en movimiento, ver Figura 53, el "key hole" adquiere un estado estacionario. Esto es, la cavidad y la piletta de fundido asociadas al haz del láser se mueven en la dirección del avance del haz, mientras que el material perdido por la vaporización se muestra como una depresión en la parte solidificada, como porosidad o como una deformación hacia el volumen del material o posiblemente una combinación de ambas.

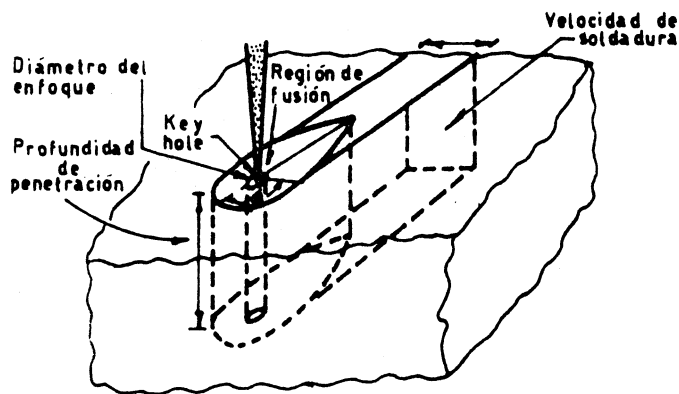


Figura 53

El requerimiento de que suficiente vapor sea producido para mantener el estado estacionario conduce a una velocidad de avance mínima para un estado estacionario dado. Mientras la cavidad se mueve a través del sólido y el material líquido a una velocidad determinada por la velo-

cidad del haz, el material debe ser movido continuamente desde una región enfrentada por el haz hacia una región dejada atrás. El trabajo experimental de Sickman [8.23] y de Arata [8.24] lo confirman.

Variables para la soldadura láser

Las variables independientes de la SL más importantes resultan ser:

- 1) Potencia del haz incidente.
- 2) Diámetro del haz.
- 3) Absortividad (cuidado, no confundir con la absorción, que es una propiedad del material que no depende del estado de su superficie, ¡ésta sí depende!).
- 4) Velocidad transversal a la superficie de trabajo.

Las variables dependientes de la SL son:

- 1) Penetración del haz.
- 2) Microestructura y propiedades metalúrgicas de las soldaduras.

De acuerdo al trabajo de Banas [8.12] para el acero inox. 304, la máxima penetración es proporcional a la potencia elevada a la 0,7, por lo tanto:

$$\text{penetración} \propto [\text{potencia}]^{0,7} .$$

El diámetro del haz para perfiles gaussianos se representa para valores de intensidad que poseen el $1/e^2$ del máximo. Este diámetro contiene el 80% de la energía del haz. Si se toma el diámetro considerando sólo hasta el valor $1/e$ del máximo, sólo el 60% de la energía será considerada.

La absorción infrarroja por los metales depende fuertemente de la absorción fotónica por los electrones libres, de acuerdo a lo estudiado en las Propiedades Ópticas de los Materiales. Arata demostró [8.24] que para metales pulidos la absortividad A de varios materiales es proporcional a la raíz cuadrada de la resistividad eléctrica r_e .

$$A = 112,2 [r_e]^{1/2} .$$

Cuando una radiación incide sobre un cuerpo, algo de ella es reflejada, algo es absorbida y alguna parte es transmitida. Así, las relaciones de cada una de ellas en relación a la potencia incidente debe ser la unidad [8.25], por lo tanto:

$$\rho + \alpha + \tau = 1 ,$$

lo que puede ser expresado en términos de flujos electromagnéticos:

$$\Phi_{\text{ref}} + \Phi_{\text{abs}} + \Phi_{\text{trans}} = \Phi_{\text{incidente}}$$

donde:

$$\rho = \Phi_{\text{ref}} / \Phi_{\text{inc}} = \text{reflectividad}$$

$$\alpha = \Phi_{\text{abs}} / \Phi_{\text{inc}} = \text{absortividad}$$

$$\tau = \Phi_{\text{trans}} / \Phi_{\text{inc}} = \text{transmisividad} ;$$

esto es también válido para cada longitud de onda:

$$\rho(\lambda) + \alpha(\lambda) + \tau(\lambda) = 1$$

Una relación dependiente de la temperatura entre la resistividad eléctrica y la emisividad del metal, fue derivada por Bramson [8.26]. Con esta fórmula puede calcularse teóricamente la absortividad a partir de la resistividad eléctrica, ya que absortividad y emisividad son equivalentes, y sólo referidas a un flujo entrante y saliente a través de una superficie. Este cálculo sólo será válido para metales calentados en el vacío, sin capa de óxido en su superficie. La presencia de esta última aumenta la absortividad. La relación es:

$$\epsilon\lambda(T) = 0,365 [r_e(T)/\lambda]^{1/2} - 0,667 [r_e(T)/\lambda] + 0,006 [r_e(T)/\lambda]^{3/2} ,$$

donde $r_e(T)$ es la resistividad eléctrica a la temperatura absoluta T expresada en Ohm-cm . $C_r(T)$ es la emisividad del sustrato a la tempera-

tura T , de la radiación de longitud de onda r .

Valores típicos de absorptividad a 300 K :

aleación Ti - 6Al - 4V : 0,15

Al, Ag y Cu : entre 0,2 y 0,3

acero 304, Fe y Zr : entre 0,12 y 0,15

Esto significa pues que las reflexiones son las pérdidas más importantes. La aplicación de polvo absorbente o anodizar la superficie a soldar son muy efectivos para minimizar la reflexión [8.24].

La absorptividad puede ser incrementada aportando un gas reactivo. El uso de un gas inerte ayuda a evitar la oxidación en la soldadura. Sin embargo, añadiendo un 10% de oxígeno al gas de aislación inerte, puede incrementarse un 100% la penetración de la soldadura. El flujo de gas en cambio, no tiene efecto significativo sobre la profundidad de soldadura. Jorgensen [8.27] en cambio, relacionó el decrecimiento de la reflectividad con la pequeña adición de oxígeno.

Una comparación de la SL con la soldadura por haces electrónicos (SHE) es mostrada en la Figura 54. La parte a ilustra cómo la presión ambiental afecta la penetración de la soldadura y la parte b muestra cómo la velocidad de soldadura afecta la penetración de la SL a baja presión. Estas figuras son del trabajo de C.O. Brown y C.M. Banas [8.28]. En ese mismo trabajo se remarcaban las diferencias entre ambas tecnologías de soldadura, que no son tan marcadas como se esperaba. La SHE es más penetrante, lo cual puede asociarse a su "longitud de onda" equivalente mucho más corta. En ese artículo los autores concluyen que las placas gruesas pueden cortarse en el vacío mediante la SHE y que la SL tomará el mercado de placas hasta 3 cm por la facilidad de hacerlo a presión atmosférica. En verdad, hoy en día la SL puede efectuarse a presión atmosférica hasta con chapas de 5 cm de espesor, gracias al avance de esta tecnología, habiéndose además adoptado ampliamente en todo Occidente en la industria de producción y servicios, mientras que la SHE está mayormente en los laboratorios de investigación y servicios en esferas de gobierno.

Una comparación de la SL con la soldadura de arco de tungsteno en gas (GTA) fue realizada por Baeslack III y Banas en soldaduras de aleaciones de titanio. En el trabajo los autores concluyen que la SL exhibe una mayor ductibilidad con resistencias iguales o mayores que con la GTA, con igual resultado con tratamiento térmico post soldadura.

Una de las propiedades interesantes de la SL , es la posibilidad de soldar materiales con coeficientes de dilatación muy diferentes. Esta posi-

bilidad, debida a la descarga pulsada con gran energía prácticamente congela las propiedades de transmisión de calor de ambas partes a ser soldadas. Esto resulta de interés para la manufactura de circuitos híbridos, donde cerámicas y metales y/o vidrios son soldados.

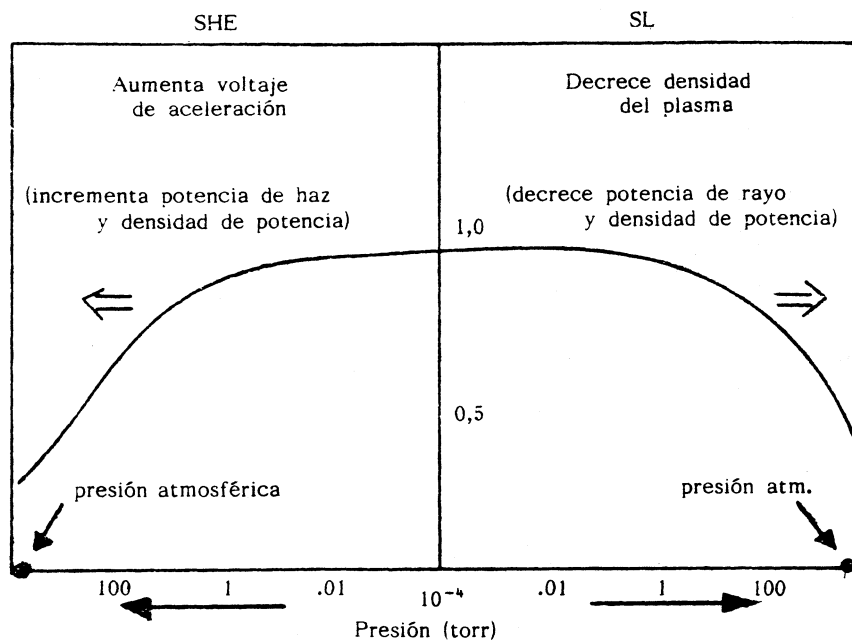


Figura 54: Comparación de la profundidad de penetración con la presión entre las soldaduras por haz electrónico y láser.

Gases de aporte en el corte y la soldadura láser

Varios tipos de gases son utilizados en el procesamiento con láseres. Para el corte de metales, el oxígeno es utilizado para aumentar la efectividad del mismo, consiguiéndose cortes en mayores espesores y el aumento de velocidad del proceso. Este favorece la reacción exotérmica que tiene lugar en el corte de metales. Cuando los espesores a cortarse no son muy gruesos (1 a 2 mm), el aporte de aire puede ser utilizado con la misma finalidad. El aporte de gases inertes como el argón y el nitrógeno en el corte produce bordes menos ásperos y elimina la necesidad de una segunda operación de acabado de las piezas. Estos gases que se alimentan coaxialmente con el haz del láser protegen la óptica de focalización de los humos de reacción. El gas dirigido de esa forma actúa como una cortina de flujo laminar que hace menos sensitivo el procesamiento al habitat de trabajo. Para cortes de alta velocidad, estos gases se emplean a presiones de hasta 4 bar con flujos de 20 litros/min. A esta presión la velocidad con que el gas emerge del pico es de alrededor de 1,8 Mach. Para materiales más espesos que requieren velocidades menores de maquinado, se utilizan flujos más pequeños. (Figura 55).

Los gases inertes utilizados como pantalla para las soldaduras ofrecen la posibilidad de obtener diferentes penetraciones de soldadura. A bajas potencias el argón es suficiente para apantallar el trabajo. El argón (Ar) es un gas mucho más denso que el aire y ofrece una simple barrera para el aire en la región de soldadura. Cuando se utiliza el láser en forma pulsada a alta potencia, el Ar se ioniza (pierde uno o dos electrones) y produce un plasma que absorbe la energía incidente del láser. En ese caso, un gas como el helio (He), que tiene un potencial de ionización (cantidad de energía necesaria para arrancarle un electrón) mucho mayor, resulta más interesante. Sin embargo, el He es mucho más liviano que el aire y entonces resulta difícil contenerlo en la región de soldadura, así que a menudo se lo mezcla con un 10% de un gas más denso, Ar nuevamente, para apantallar con mayor eficiencia. La absorción de energía durante la soldadura puede ser incrementada adicionando pequeñas cantidades de gas reactivo, como el O_2 , al gas de apantallado.

La Figura 55, nos muestra cómo depende la penetración de la soldadura para un láser operando a 1 y 2 kW de potencia continua en función de la velocidad de soldadura. Se muestra el efecto de utilizar He, N_2 o Ar como gas de aporte. La forma de la piletta de fundido condicionará (como puede correlacionarse de la experiencia) las propiedades mecánicas a posteriori. Obsérvese que hasta llegar a la potencia de 1 kW no hay diferencia significativa con los tres gases utilizados. Dependiendo del tipo de gas, la presión y el flujo, el producto de la intensidad por el radio focal (que es un parámetro práctico de procesamiento) y la velocidad de soldadura, el efecto del gas de asistencia varía. Pero a velocidades menores, el efecto es más notable. Para un conjunto de parámetros,

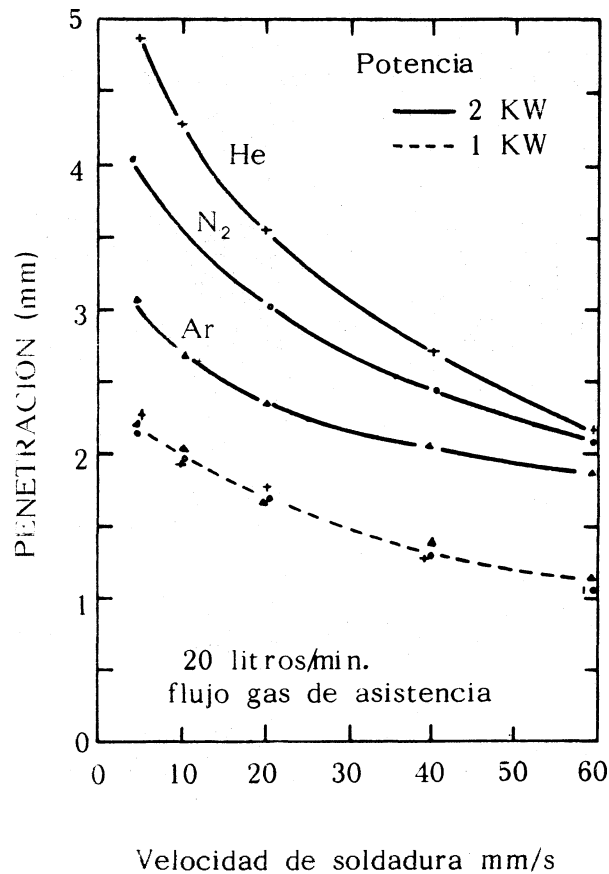


Figura 55

la máxima penetración se obtiene con el He, pero si la intensidad se aumenta sistemáticamente, puede suceder que el plasma formado, flote por encima de la superficie y sin contacto. En este caso el proceso de soldadura se interrumpe. A velocidades elevadas casi es indistinto qué gas utilizar, pero el encendido de la descarga luminosa se produce a menores intensidades con el Ar. El tener un sistema de gran velocidad de procesamiento ayuda pues a la economía del funcionamiento, no sólo por la velocidad, sino porque favorece el uso de gases de apantallamiento más baratos (el He es el más caro) [8.29].

La Figura 52 nos muestra cómo el gas de asistencia que protege de la oxidación a la soldadura, termaliza el plasma formado, definiendo, al aumentar la intensidad, una zona de procesamiento y una zona de aislación [8.29].

La Figura 56 muestra cómo la profundidad de soldadura es capaz de producirse con diferentes tecnologías y cómo varían en función de las velocidades de soldadura [8.30].

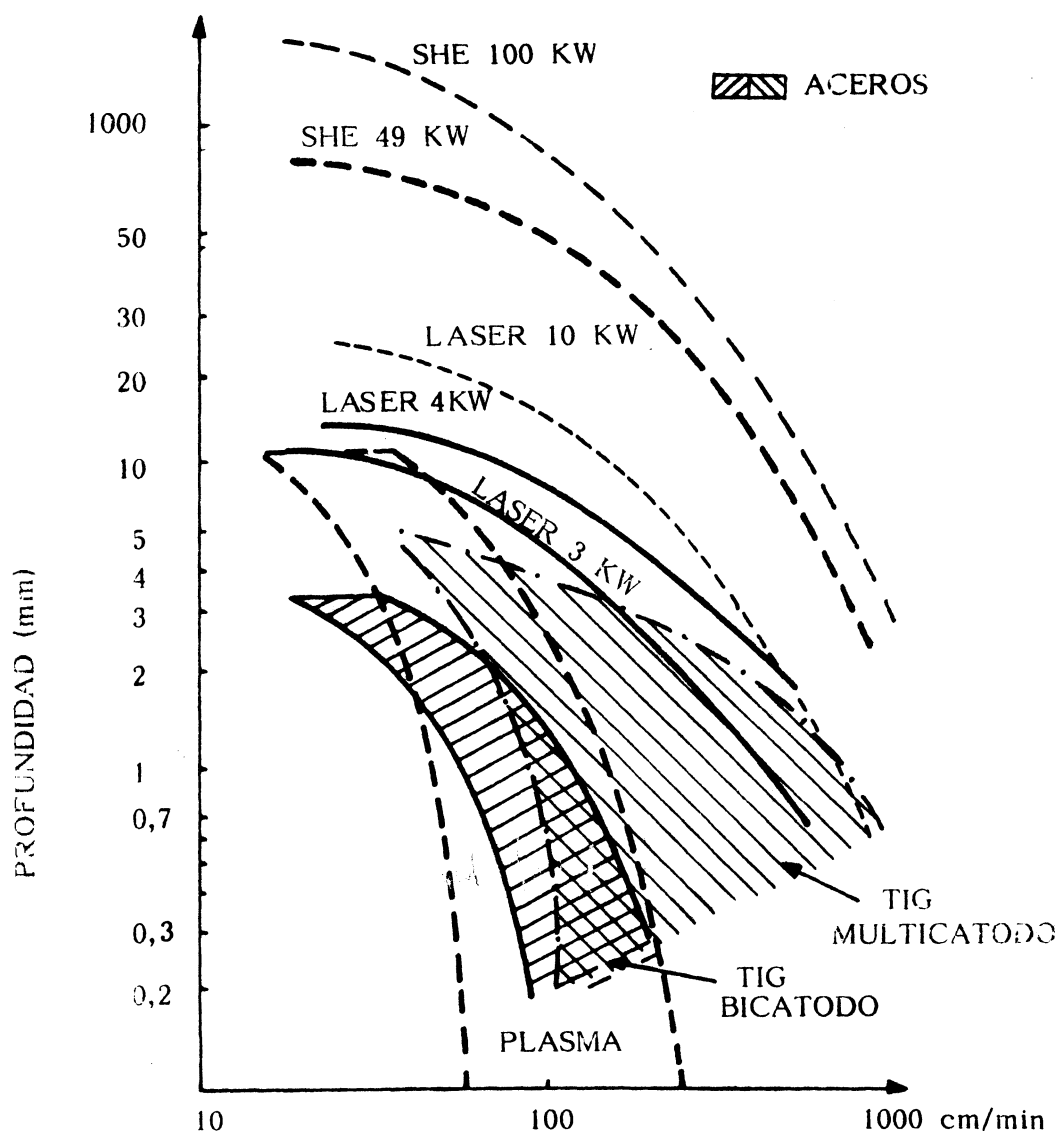


Figura 56

Por último, para los grandes sistemas láser, de más de 10 kW de potencia, entre los cuales la soldadura es el objetivo para grandes espesores utilizados en la industria naval, gasoductos, automotores, etc., generalmente se publican costos poco creíbles, sobre todo teniendo en cuenta los valores de amortización que ellos poseen. La Figura 57 ofrece una comparación de costos de soldadura para tres sistemas de diferente potencia máxima de operación. El gráfico muestra cómo es dependiente el costo de la potencia máxima de operación del láser. Para mayores potencias instaladas, se hace más onerosa la operación a bajas potencias, pero sin embargo es más económica la operación a potencias elevadas, estando estas últimas además capacitadas para soldar mayores espesores de metal [8.29].

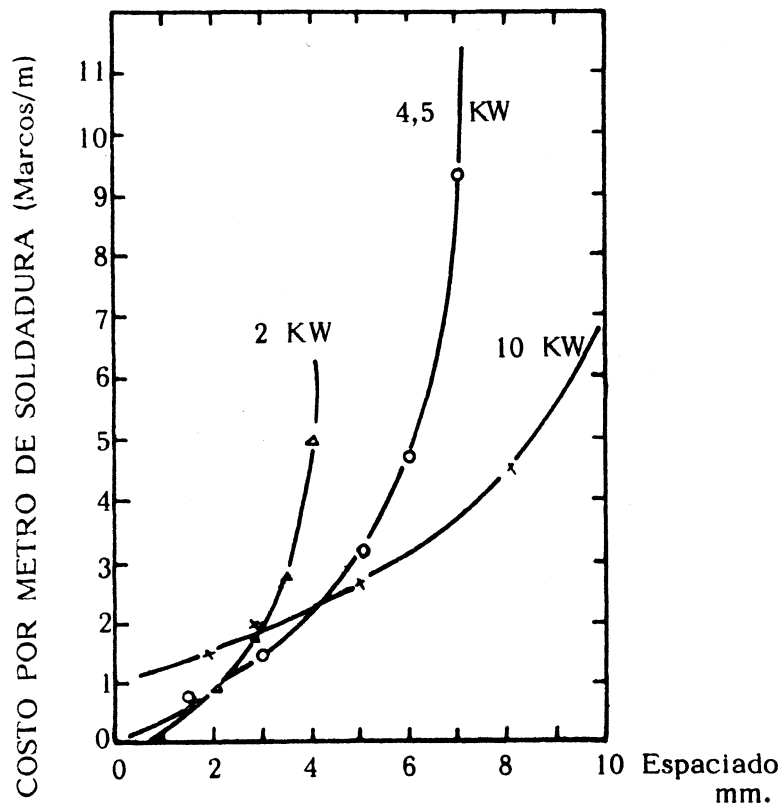


Figura 57

8.5. Otros procesamientos y otros metales

Hay otros procesamientos que se realizan con el láser de CO_2 , de los que no daremos muchos detalles, pero sí describiremos brevemente.

La aleación superficial es un método utilizado para realizar aleaciones en la superficie de laminados y piezas, destinado a proveer mejores condiciones contra la abrasión, corrosión y desgaste en sus superficies. Por este método, productos que utilizan materiales de aleación que son caros o resultan estratégicos o de importación, se ven reducidos en sus costos sin que sean afectadas las funciones para los cuales son empleados. Resulta interesante destacar que muchas de las aleaciones que se obtienen, no son posibles de obtener de otra forma, ya que gracias al congelamiento rápido, se consiguen por la aleación superficial con láser, aleaciones que contienen metales no solubles en otros. Estas películas, con espesores que dependen de la naturaleza de las mismas y de sus constituyentes, pueden variar entre 5 y 200 micrones. La manera de fabricarlas puede ser en forma de particulado diminuto que se adiciona en forma de spray con el gas de aporte o bien por evaporación o pulverizado de una película sobre la superficie a procesar, sobre la cual se realiza la posterior aleación.

La fusión superficial con láser es esencialmente el mismo procesamiento que el descrito antes, con la salvedad de que no posee material de aporte. Se obtienen algunos resultados muy interesantes, inclusive con los mismos aceros inoxidables, los cuales, sometidos luego de este procesamiento a pruebas de corrosión (picado, específicamente) la corriente galvánica disminuye en las pruebas voltamperométricas, dos órdenes de magnitud. Estos procedimientos pueden ser muy importantes en el futuro en el manejo de materiales químicos muy corrosivos.

Con el avance de esta tecnología, se consigue hoy cortar con precisión metales no ferrosos, que antes era difícil de hacerlo o la calidad del corte resultaba muy disminuida en relación a la conseguible en hierros y aceros. Si bien es cierto que los espesores de corte con calidad aceptables en los no ferrosos, suelen ser un 30% menos del espesor conseguible con calidad equivalente en los ferrosos. Vale la pena destacar que se producen dos fenómenos en relación a esto desde el punto de vista comercial. El corte con calidad de los NF hace posible producciones pequeñas para las Pequeñas y Medianas Empresas (PYME's), que antes hacían sus chasis y plantillas metálicas en forma artesanal; y que el corte de laminados con CAD/CAM ha hecho posible reemplazar a muchas piezas antes realizadas en aluminio, pues se reducen los stocks al haber respuestas inmediatas de producción, resultando irrelevante el tratamiento posterior como pintura o anodizado, etc., que el laminado pueda requerir adicionalmente en relación al aluminio.

Los no ferrosos más utilizados son el aluminio, latón y cobre y en

los tres se obtienen muy buenos resultados.

También se procesan industrialmente en forma masiva plásticos y materiales orgánicos, así como también se hace deposición química de vapores (LCVD) en las industrias de los dispositivos semiconductores, pero éstos escapan al objetivo que nos hemos fijado aquí.

Otros láseres se utilizan también industrialmente. El Nd:YAG es el que sigue en la escala de valores industriales; se calcula que la relación de instalaciones de 1 en 60 en relación al láser de CO_2 . El láser de excímero igualmente se está comenzando a utilizar industrialmente. Otro láser que promete una faja importante de aplicaciones es el de monóxido de carbono que resumiría dos cualidades importantes del Nd:YAG y del CO_2 . Estas son la capacidad de ser guiado por fibras ópticas y mejores propiedades ópticas con los ferrosos del primero, y la potencia y energía disponibles del segundo.

8.6. Elaboración de matrices

Cuando vimos el corte de precisión que efectúa el láser en el estado actual de la tecnología, vimos que pueden concretarse los cortes de gran precisión, como puede apreciarse en diversos ejemplos que mostramos en la Figura 44, todas las figuras allí mostradas tienen un aumento de 10 veces.

En la pieza cortada y medida, la dispersión de las medidas realizadas en los puntos críticos fue superior a los 75 micrones. Si bien debemos destacar que este requerimiento es excelente para la casi totalidad de las necesidades industriales, no lo es si pensamos en la manufactura de matricería de corte de bajo precio y rápida entrega que podría lograrse por este método. Obteniéndose un corte por debajo de los 50 micrones en el error de coordenadas (lo cual depende de la calidad del sistema) y con control local de endurecimiento, pueden construirse matrices de corte que no requieran de las piezas a confeccionarse con ellas una tolerancia de $\pm 0,1$ mm en sus medidas. Estas tolerancias son suficientes para la mayoría de las necesidades de nuestro mercado y de la región.

APENDICE 1

Propiedades Ópticas relevantes en el Procesamiento Láser (CO_2) de varios materiales.

METALES (baja absorción)	de absorción insuficiente	Oro Plata Cobre Aluminio Latón
		<u>Con bajo punto de fusión</u> Estaño, 232 C Plomo, 327 C Níquel, 1453 C Hierro, 1530 C
	de absorción suficiente	<u>Con alto punto de fusión</u> Titanio, 1668 C Zirconio, 1852 C Cromo, 1875 C Molibdeno, 2610 C Tungsteno, 3410 C
NO METALES (alta absorción)	Orgánicos	<u>Combustibilidad baja</u> Acrílicos PTFE Polietileno Polipropileno Policarbonatos Gomas naturales y sintéticas
		<u>Combustibilidad alta</u> Cueros Maderas Lana Algodón Laminados PVC
	Inorgánicos	<u>Con alta expansión térmica</u> Vidrios
		<u>Con baja expansión térmica</u> Cerámicas Cuarzo Alúminas Loza Amianto Mica Piedras naturales

APENDICE 2

Composición Química de las Chapas de Acero Inoxidable (ASTM A-240)

AISI	302	304	304L	310S	316	316L	321	409	410	410S	430
máx%C	,15	,08	,03	,08	,08	,03	,08	,08	,15	,08	,12
% Cr	17-19	18-20	18-20	24-26	16-18	16-18	17-19	10.5- -11,75	11,50 - 13,50		16-18
% Ni	8-10	8-10.5	8-12	19-22	10-14	10-14	9-12	,50	,75	,60	,75
% M	-	-	-	-	2-3	2-3	-	-	-	-	-
máx%N	,1	,1	,1	-	,1	,1	,7	,75	-	-	-

Agradecimientos

Deseo expresar aquí mi reconocimiento a la Comisión de Investigaciones Científicas de la Provincia de Buenos Aires (CIC) por haber manifestado esta temática de interés y por el continuo apoyo recibido. A la Secretaría de Ciencia y Tecnología de la Nación por el subsidio recibido dentro del Programa de Componentes Electrónicos.

La firma **OXICORTE S.A.** patrocinó estudios de soldadura, fusión superficial de metales, endurecimiento de aceros y la deposición química de vapores con láseres en el **CIOP** con fondos de su Proyecto de Centro de Procesamiento Láser por lo que agradezco la confianza depositada y el continuo aliento a la actividad de desarrollo industrial del Programa de Propiedades Ópticas de los Materiales del **CIOP**.

Mi agradecimiento al **Dr. Flemming Olsen** a quien debo orientación y madurez en este tema y al **Prof. William Steen** por gratos momentos compartidos los cuales han sido sumamente instructivos para mí.

Al **Prof. José Manuel Orza**, por su siempre buena voluntad y generosidad para soportar mis requerimientos de literatura y opiniones.

Bibliografía

- 1.1 Garavaglia, Mario; "El láser", Capítulo 1, este libro.
- 1.2 Ranea Sandoval, Héctor F. y Scaffardi, Lucía B.; "Algunos tipos de Láseres", Capítulo 2, este libro.
- 1.3 Patel, C. Kumar N.; Phys. Rev. Lett. 12, 588 (1964).
- 1.4 Lengyel, Bela; "Introduction to Laser Physics", Wiley, N.Y., 1966.
- 1.5 Patel, C. Kumar N.; Phys. Rev. Lett. 13, 617 (1964).
- 1.6 Patel, C. Kumar N.; Phys. Rev. 136, A1187 (1964).
- 1.7 Legay, F. and Lagay-Sommaire, N.; C.R. Acad. Sci. 259B, 99 (1964).
- 1.8 Patel, C. Kumar N.; Appl. Phys. Lett. 7, 15 (1965).
- 1.9 Labarini, B., Bettini, J.P., Brunet, H. and Michon, M.; C.R. Acad. Sci. 269, 1301 (1969).
- 1.10 Deutsch, T.F., Horrigan, F.A. and Rudko, R.I., Appl. Phys. Lett. 15, 88 (1969).
- 1.11 Beaulieu, A.J.; Appl. Phys. Lett. 16, 504 (1970).

- 2.1 Born, M. and Wolf, E.; "Principles of Optics, Electromagnetic Theory of Propagation, Interference and Diffraction of Light", Fifth Edition, Pergamon Press, Oxford, 1975.

- 3.1 Beyer, E., Herziger, G., Kramer, R., and Loosen, P.; "A Diagnostic System for Measurement of the Focused Beam Diameter of High Power CO₂ Lasers", The Changing Frontiers of Laser Materials Processing, Ed. by Banas, C.M. and Whitney, G.L., IFS (Publications) Ltd., Reino Unido, 1987.
- 3.2 Lengyel, B., "Lasers", Second Edition, Wiley-Interscience, New York, 1971.
- 3.3 Young, M., "Optics and Lasers, An Engineering Physics Approach", Springer-Verlag, N.Y., 1977.
- 3.4 Fox, A.G. and Li, T., "Resonant Modes in a Maser Interferometer", Bell System Tech. J., 40, 453 (1961).
- 3.5 Kogelnik, H. and Li, T., "Modes in Optical Resonators", Lasers Ed. by Levine, K.E., Marcel Dekker, New York, 1966.

3.6 Toraldo di Francia, T., "Optical Resonators", Opt. Acta, 13, 323 (1966).

4.1 Ver 2.1

6.1 Von Allmen, M.; "Laser-Beam Interaction with Materials, Physical Principles and Applications", Springer Verlag, Berlin, 1986.

6.2 Wooten, F.; "Optical Properties of Solids", Academic Press, New York, 1972.

7.1 Thornton, P.A. y Colangelo, V.J.; "Ciencia de Materiales para Ingeniería", Prentice Hall Hispanoamericana, S.A., México 1987.

7.2 Van Vlat, L.H.; "Elements of Materials Science and Engineering", 4th. Edition, Addison Wesley Pub. Co., Reading, Massachusetts, 1980.

Bibliografía general para este tema:

7.3 Moffatt, W.G., Pearsall, G.W. and Wulff, J., "The Structure and Properties of Materials, vol. 1 Structure", Wiley and Sons, Inc., New York 1964.

7.4 Doan, G.E.; "The Principles of Physical Metallurgy", 3rd. Edition, McGraw-Hill-Kogakusha, Tokyo 1953.

8.0 Sona, A.; "Applied Laser Tooling", Ed. by Soares, O.D.D. and Perez-Amor, M., Martinus Nijhoff Pub. Co., Dordrech, 1987.

8.1 a) Gallego Lluesma, E., "Láser y Optica en Ingeniería", Centro de Investigaciones Opticas, cursos 1983-88.

b) Pomarico, J.E. y Gallego Lluesma, E., "Termografía Infrarroja de los Procesos Metalúrgicos con Láser", 73ava. Reunión de la Asociación Física Argentina (AFA), Mar del Plata, 16-21 de octubre de 1988.

c) Pomarico, J.E. y Gallego Lluesma, E., "Procesamiento Superficial de Metales con Láseres de Dioxido de Carbono", 73ava. AFA.

d) Gallego Lluesma, E., "Corte de Metales de Alta Calidad con Láser de CO₂", 73ava. AFA.

- 8.2 Kanaoka, M.; Comunicación personal, Porto, Portugal, 1988.
- 8.3 Olsen, F.; Comunicación personal, Porto, Portugal, 1988 y "Cutting with Polarized Beams", DVS-berichte 63, 197, 1980.
- 8.4 Andrés, J.B., Gardella, H. y Gallego Lluesma, E.; "Evaluación del corte de Precisión con Láser de CO₂, Efectuado Mediante Diseño Asistido por Computadora y Control Numérico", Jornadas Metalúrgicas 1989 de la Sociedad Argentina de Metales (SAM), Universidad de Morón, Mayo 8-12, 1989, Morón, Pcia. de Buenos Aires.
- 8.5 Mirkin, L.I. and Pilipetskii; "The Hardening of Steel by the Action of an Optical Laser Beam", Metallovd. i Termich. Obrab. Materialov, 4, 70 (1966).
- 8.6 Creal, R.; "Laser vs. Furnaces: Light Makes Right", Heat Treating, 23, March 1985.
- 8.7 Comunicaciones personales de integrantes japoneses a la Conferencia "Lasers in Industry", Porto, Portugal, junio 1988.
- 8.8 Belforte, D.A., "Laser Heat Treating", Industrial Laser Review, dec. 1988.
- 8.9 Sandven, O.; "Laser Surface Transformation Hardening" en Metal Handbook, 9ne. Edition, American Society for Metals (ASM), vol. 4 "Heat Treating", Ed. by Alban, L.E., Eastday, J.R. and Zamborsky, D.S., ASM, Metals Park, Ohio 1988.
- 8.10 Brown, C.O., Appl. Phys. Lett. 17, 9 (1970).
- 8.11 Brown, C.O. and Davis, J.M., "Electrical Discharge Convection Lasers", Proc. IEEE Intl. Electronic Devices Meeting, Washington, DC. 1970.
- 8.12 Banas, C.M., "Laser Welding to 100 kW", United Technologies Research Center, Report # R76-912260-2, Feb 1977.
- 8.13 Banas, C.M., "Electron Beam, Laser Beam and Plasma Arc Welding Studies", Contract NASA 1-12565, March 1974.
- 8.14 Mazumder, J., Ph.D. Thesis, London University, 1978.
- 8.15 Steen, W.M., fundador del Laser Processing Laboratory, Imperial College, Londres. Primer laboratorio con un láser industrial en un ambiente universitario.
- 8.16 Mazumder, J. and Steen, W.M., TMS-AIME Fall Meeting, paper F-79-17, Set 1979.
- 8.17 Banas, C.M., Report NASA, CR-132386, March 1975.
- 8.18 Banas, C.M., United Technologies Research Center, Report R75-412260-1, July 1975.
- 8.19 Mazumder, J. and Steen, W.M., J. Appl. Phys., Feb. 1980, pág. 944.

- 8.20 Chande, T. and Mazumder, J., "Lasers in Metallurgy", AIME, Annual Meeting, Chicago, Feb 1981. Ed. Mukhery K. and Mazumder, J., AIME 1982.
- 8.21 Duley, W.W., "CO₂ Lasers, Effects and Applications", Academic Press, New York 1976.
- 8.22 Klemens, P.G., J. Appl. Phys. 47, 2165 (1976).
- 8.23 Sickman, J.G. and Morijn, R., Philips Res. Rep. 23, 367 y Philips Res. Rep. 23, 375 (1968).
- 8.24 Arata, Y. and Miyamoto, I., Technocrat. 11, 5, 33 (1978) y Laser Focus, March 1977.
- 8.25 "The Infrared Book" ed. by Wolfe, W. and Zissis G.J., IRIA Center 1978.
- 8.26 Bramson, M.A., "Infrared Radiation: A Handbook of Applications", Plenum Press, New York 1968.
- 8.27 Jorgensen, M., Metal Construction 12, 2, 88 (1980).
- 8.28 Brown, C.O. and Banas, C.M., Welding, J., July 1986, pág. 48.
- 8.29 Rosen, H.-G., Proc. of the International Congress on Applications of Lasers and Electro-Optics, ICALEO 86, Ed. by Banas, C.M. and Whitney, G.L., IFS Pub. Ltd., Reino Unido 1987.
- 8.30 Ebbro, J., Pieklo, C.A., Berreta, J.C. y Gallego Lluesma, E., "Procesos Metalúrgicos Asistidos por Gases", Congreso Argentino de Soldadura, Instituto Argentino de Siderurgia, Auditorio de SOMISA, Buenos Aires, 1988.

DETECTORES DE RADIACION VISIBLE E INFRARROJA

A. P. Laquidara

Introducción

El detector ideal de fotones es aquél que captura y entrega a un instrumento de medida una señal eléctrica sin ambigüedades, contando todos los eventos.

En la realidad, estos detectores no existen porque factores internos y externos a los dispositivos compiten con el fotón para dar una medida errónea.

Si la intensidad de la radiación es débil o la longitud de onda es mayor de algunos pocos micrometros ($\mu\text{m} = 10^{-6} \text{ m}$), los problemas son más complicados aún para detectar un fotón o haz de fotones.

Los primeros detectores de fotones fueron el ojo y la placa fotográfica, pero éstos no sirven para medir radiación de baja intensidad.

Los dispositivos fotoeléctricos dan una salida eléctrica (electrones o electrones-huecos), excitados por la absorción de fotones.

Cada detector tiene un rango óptimo de longitud de onda a la que es sensible. Los detectores fotoemisores de vacío en general responden a la zona ultravioleta y visible del espectro. Los semiconductores son los únicos que funcionan cuando la longitud de onda es poco mayor de $1 \mu\text{m}$. Los térmicos responden indiscriminadamente a todas las longitudes de onda: desde los rayos X a las microondas, pero son muy poco sensibles y no sirven para detectar pocos fotones. Básicamente hay 3 tipos de detectores:

TIPOS	{	FOTOEMISIVOS	{	FOTODIODO DE VACIO
				FOTOMULTIPLICADOR
	{	SEMICONDUCTORES	{	FOTORESISTORES
				FOTODIODOS
	{	TERMICOS		TERMISTORES
				TERMOPILAS
				PIROELECTRICOS

Fotoemisivos

Consisten en una superficie expuesta a la radiación, revestida con un material fotosensible que forma el fotocátodo. Los fotones que chocan con estas superficies y son absorbidos excitan algunos electrones y aquéllos que alcanzan la superficie en contacto con el vacío son extraídos por un campo eléctrico acelerador.

La relación entre los electrones emitidos y fotones absorbidos se define como eficiencia cuántica η . Todos los esfuerzos realizados para lograr detectar señales débiles se centran en conseguir materiales del fotocátodo que absorban radiación y emitan más electrones. La eficiencia cuántica depende de la longitud de onda de la radiación y va desde un máximo del 90% para el ultravioleta hasta unas décimas por ciento para el infrarrojo cercano (hasta 1,2 μm).

Pasando ese valor, los fotoemisivos no funcionan y la detección se realiza con los semiconductores.

La teoría de la detección establece que si una radiación de potencia constante P y frecuencia ν incide sobre un detector ideal de fotones, se producirán electrones y éstos generarán una corriente de valor

$$i = \frac{\eta e P}{h\nu} = R_i P \quad ;$$

siendo R_i la responsividad del detector (A/W), η la eficiencia cuántica, e la carga del electrón, y $h\nu$ la energía del fotón.

Decir ideal aquí significa que la radiación incidente sobre el detector es toda la emitida por la fuente y sólo ésa. Como el número de electrones producidos por unidad de tiempo no es constante, por la naturaleza aleatoria de la luz, existe una fuente de ruido inevitable y mínimo para todo detector, llamado ruido de fotón o ruido cuántico, debido a la fluctuación en el número de eventos.

Este ruido se expresa como una media cuadrática o eficaz:

$$\overline{i_N^2} = 2e i B \quad ;$$

i es la corriente generada por la radiación y B es el ancho de banda del sistema al que se acopla el detector, por ejemplo un amplificador.

El ruido cuántico propio del detector y el ruido térmico o de Johnson del amplificador, son factores limitantes de la sensibilidad de cualquier detector de fotones.

La corriente de ruido térmico generada en el amplificador es función de la resistencia eléctrica R que presenta el detector,

$$\overline{i_N^2} = \frac{4KTB}{R} .$$

Todo detector fotoemisor tiene una fuente de ruido propia debida a la emisión termoiónica de electrones.

Estos electrones emitidos a cualquier temperatura superior a 0° K constituyen la llamada corriente oscura, es decir una señal de salida en ausencia de radiación incidente. Esta corriente depende fuertemente de la temperatura, por lo que enfriando el fotocátodo puede reducirse varios órdenes de magnitud, hasta límites casi despreciables.

La corriente oscura tiene una densidad dada por

$$J_d = 120 T^2 \cdot e^{-(h\nu_c/KT)} \text{ (A/cm}^2\text{)} ,$$

donde ν_c es la frecuencia de corte detectada por el fotocátodo.

El ruido producido por la corriente oscura vale

$$\overline{i_N^2} = 2e i_d B ,$$

con i_d : corriente oscura.

Una importante limitación más a la detección es la indeseada radiación de fondo: cósmica, solar y de los cuerpos que rodean a la fuente de radiación medida, como también a los que se interponen en el camino al detector.

Aunque la fuente de radiación de fondo sea perfectamente constante en el tiempo, produce un ruido cuántico que compite con la corriente de los fotones que se quieren observar.

La forma de reducir su influencia en la medida, es reducir el ángulo sólido de observación del detector sobre la fuente de interés y enfriar todos los elementos ópticos que debe atravesar la señal.

La sensibilidad de la mayoría de los detectores contempla en su definición un factor debido a la radiación de fondo que se supone emitida por un cuerpo negro a temperatura de 295° K .

En resumen, existen tres categorías de ruido que limitan la detección:

- 1 - Ruido de Fotones o Cuántico
- 2 - Ruido de Amplificador
- 3 - Ruido de fondo .

Magnitudes que definen a un detector.

NEP (Noise Equivalent Power): Potencia equivalente al ruido.

Es la potencia de entrada igual al ruido propio del detector. Por ejemplo, para calcular el NEP, si la corriente de ruido está compuesta por el ruido cuántico y el debido a una radiación de fondo, la corriente producida es:

$$i = \frac{e\eta(P_S + P_B)}{h\nu} .$$

Esta corriente generará un ruido de valor:

$$\overline{i_N^2} = \frac{2e^2\eta(P_S + P_B)B}{h\nu}$$

La relación señal-ruido es $S/N = i_S^2 / \overline{i_N^2}$; $i_S = \eta e P_S / h\nu$.

$$S/N = \frac{\eta P_S^2}{2h\nu B(P_S + P_B)} .$$

Cuando esta relación es igual a 1, la potencia de señal iguala al NEP,

$$NEP = \sqrt{\frac{2h\nu B(P_S + P_B)}{\eta}} . \quad (P_S < P_B)$$

Esto vale para la detección limitada por ruido de fondo.

Detectividad específica D^* .

La inversa del NEP da un valor que crece con la sensibilidad del detector y se denomina detectividad D .

Como ésta disminuye con la superficie del detector y para referirla a un ancho de banda normalizado en 1 Hz, se define la detectividad específica D^* :

$$D^* = \frac{\sqrt{AB}}{NEP} \quad \frac{\text{cm}\sqrt{\text{Hz}}}{\text{W}}$$

Para un fotodetector ideal en presencia de ruido de fondo, se calcula la D^* en función de la longitud de onda y da el siguiente gráfico:

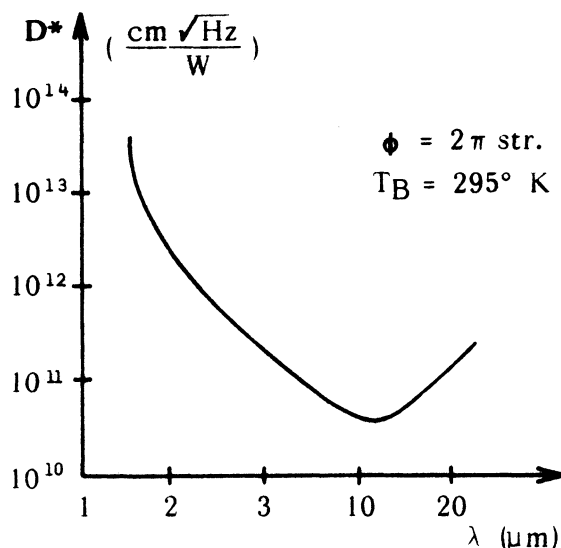


Figura 1

Todo detector real tendrá una detectividad específica inferior a la graficada.

Esta detectividad, para todo dispositivo limitado por ruido de fondo, es el valor de la máxima sensibilidad alcanzable.

Los detectores fotoemisivos más importantes son:

- 1 - El fotodiodo de vacío
- 2 - El fotomultiplicador .

1 - Fotodiodo de vacío.

Consiste en un fotocátodo y un ánodo polarizados con aproximadamente 2000 V.

Se emplea para detectar señales relativamente intensas y tiene tiempos de respuesta de menos de 1 ns (10^{-9} s).

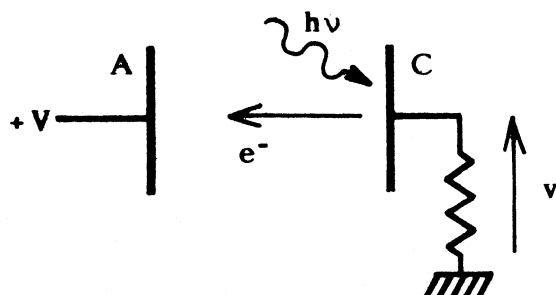


Figura 2

2 - Fotomultiplicadores.

Consisten en un fotocátodo y una cadena multiplicadora de electrones que consigue corrientes apreciables a partir de señales luminosas extremadamente débiles.

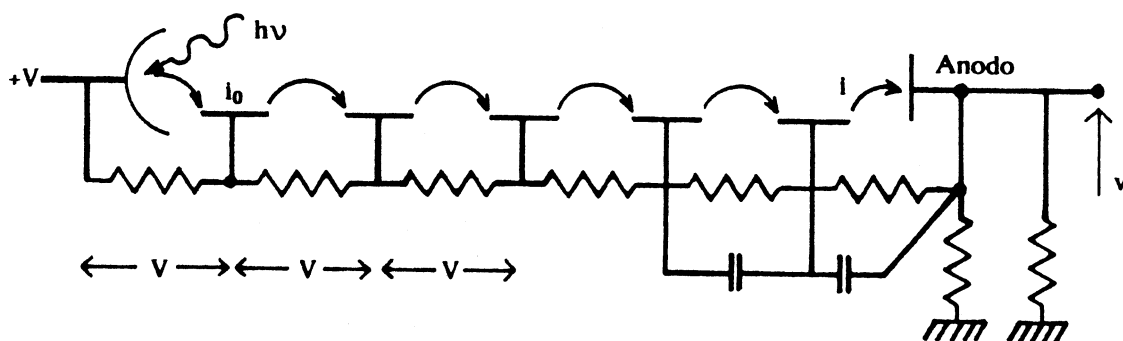


Figura 3

Los electrones emitidos por el fotocátodo son acelerados por el potencial V y al chocar contra el primer dinodo producen m electrones secundarios que a su vez siguen multiplicándose hasta producir una corriente varios órdenes de magnitud mayor que la i_0 .

La ganancia de corriente que se puede lograr con 10 o 12 etapas es de 10^6 o más.

Los fotocátodos de los fotodiodos de vacío y fotomultiplicadores tienen designaciones específicas según los elementos de los que están compuestos y de la respuesta espectral.

Los modernos fotocátodos de elementos de los grupos III y V de la tabla periódica tienen una alta eficiencia cuántica y una respuesta espectral casi plana entre el ultravioleta y el infrarrojo cercano ($< 1,2 \mu\text{m}$).

En cuanto a la respuesta temporal a los pulsos luminosos, el tiempo de crecimiento de la salida depende de la dispersión de e^- que se produce en el proceso de multiplicación.

Para lograr una buena coherencia entre los electrones multiplicados, la geometría de los dinodos se diseña de modo de producir el enfoque de los electrones y recorridos iguales.

En la estructura llamada de enfoque lineal se obtienen tiempos de crecimiento de hasta 2 ns ($t_r = 2 \cdot 10^{-9} \text{s}$). Ver Figura 4.

En estructuras más complejas del tipo de campo cruzado y de placas microcanales, se consiguen tiempos de crecimiento de 0,12 ns ($t_r = 0,12 \cdot 10^{-9} \text{s}$), con ganancia de 10^5 a 10^7 . Ver Figuras 5 y 6.

Cantidades que caracterizan a un fotomultiplicador.

- 1) Respuesta espectral: Depende de los materiales del fotocátodo. Se designa por una S-1 , S-4 , S-13 , etc.
- 2) Sensibilidad del cátodo: Es la corriente generada en el cátodo medida empleando como fuente luminosa una lámpara de filamento de tungsteno $T = 2856^\circ \text{K}$. Se expresa en $\mu\text{A/lm}$.
- 3) Sensibilidad del ánodo: Es la corriente leída en el ánodo y es igual a la corriente del fotocátodo multiplicada por la ganancia o amplificación del multiplicador de electrones. Ver Figura 7.
Se indica para distintas líneas espectrales, y se expresa en A/W .
- 4) Ganancia: Es el valor de la amplificación de electrones entre el fotocátodo y el ánodo. Depende del número de etapas, de la tensión y de los materiales que recubren los dinodos. Oscila entre 10^3 y 10^{10} .
- 5) Corriente oscura: Es el valor de la corriente leída en el ánodo con la máxima tensión aplicada y sin radiación incidente. Depende de las fugas de aislación en la base del tubo fotomultiplicador, de la emisión termoiónica y de la ionización de gases residuales. Puede reducirse enormemente enfriando el cátodo con hielo seco, agua fría o termoelectricamente.
- 6) Ruido equivalente a la entrada (ENI): Es el valor de la potencia incidente que produce una corriente de salida igual a la corriente de ruido

(relación $S/N = 1$). Es una expresión similar al NEP, que es específico de los semiconductores. Las fuentes de ruido consideradas al definir el ENI son: la corriente oscura y el ruido térmico si fuese conocida la resistencia de carga.

$$\text{ENI} = \frac{1}{S_A} \left(\sqrt{2e I_d A B} + \sqrt{\frac{4KT B}{R}} \right) .$$

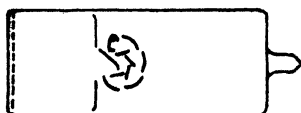
Si sólo se considera la corriente oscura como fuente de ruido,

$$\text{ENI} = \sqrt{\frac{2e I_d A B}{S_A}} .$$

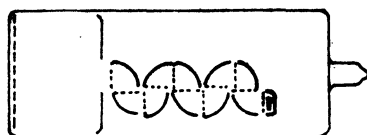
TIPOS DE FOTOMÚLTIPlicADORES



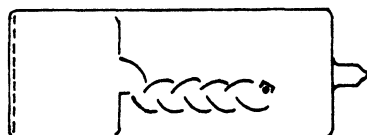
a) Caja Circular
(Entrada lateral)



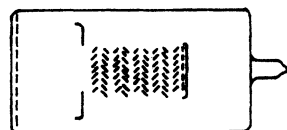
b) Caja Circular
(Entrada Frontal)



c) Caja y Grilla



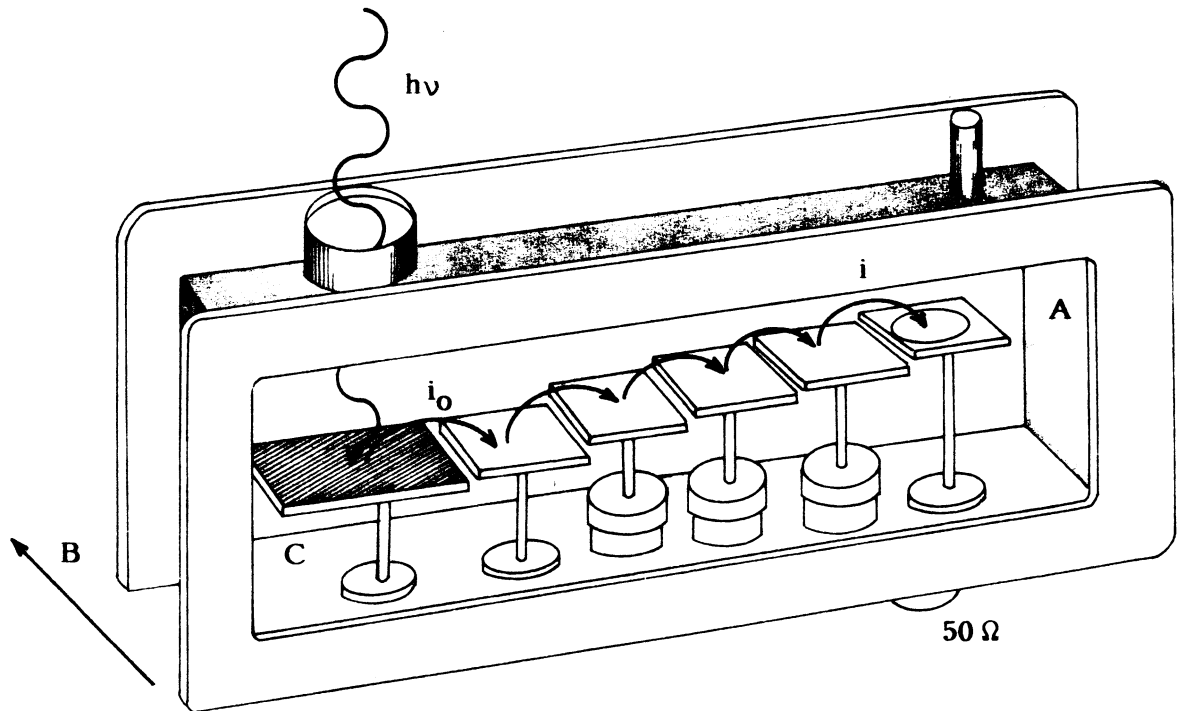
d) Enfoque Lineal



e) Ciego Veneciano

Figura 4

FOTOMULTIPLICADOR CAMPO CRUZADO ESTÁTICO



Detecta pulsos de 10^{-12} W con $\tau_r \approx 120$ pseg. (3 GHz)

Fotocátodo III-V

Dinodos GaP . Cuenta Foton.

Ganancia: 10^3 a 10^5 .

Figura 5

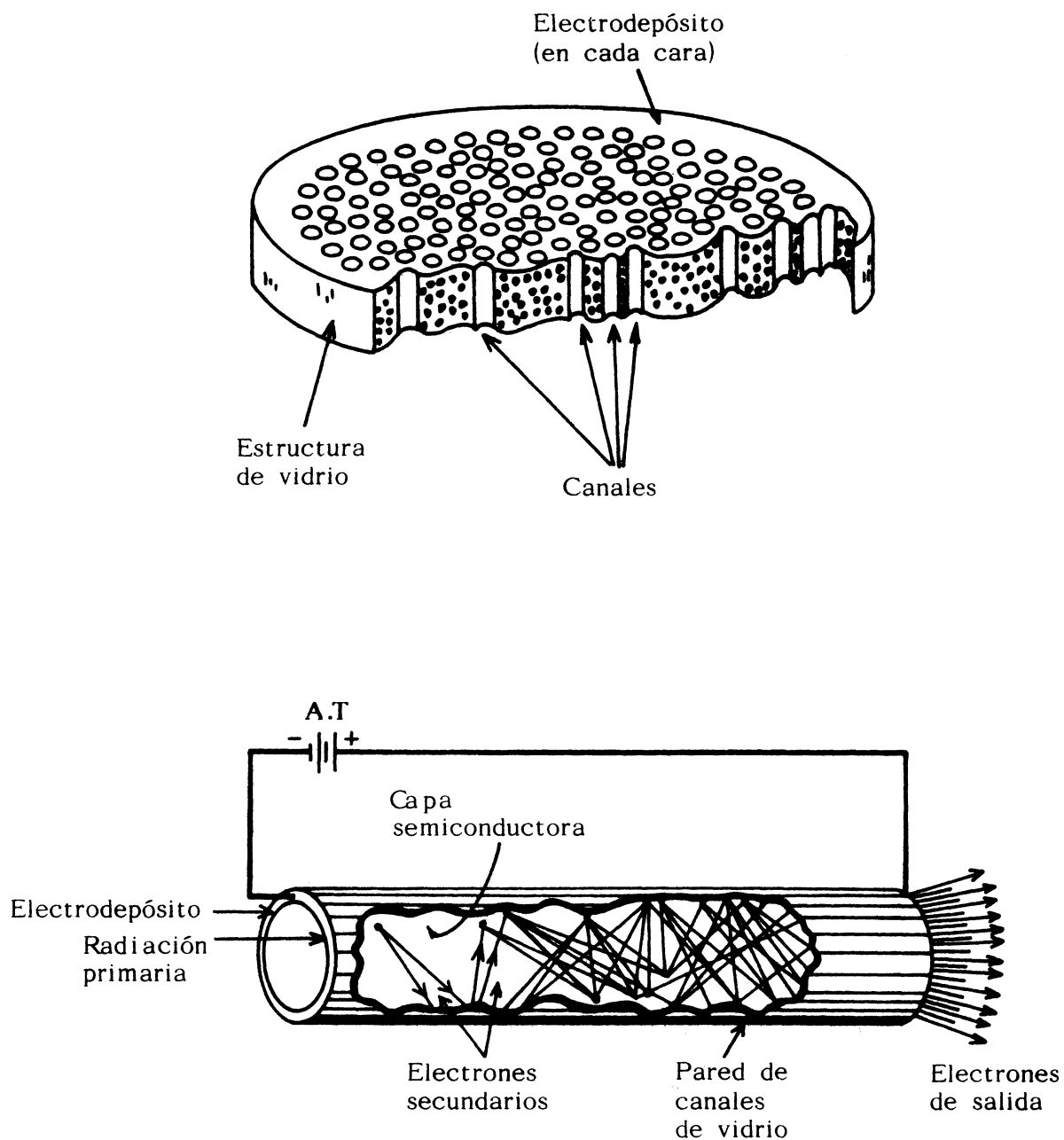


Figura 6

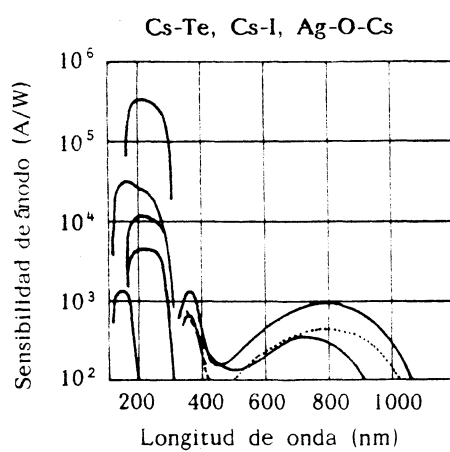
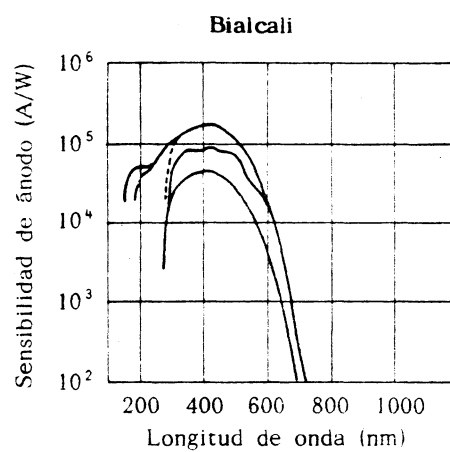
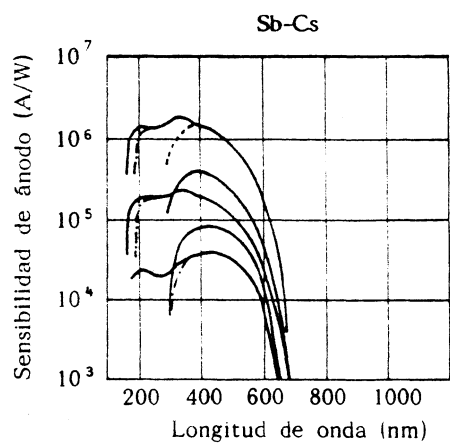


Figura 7.a

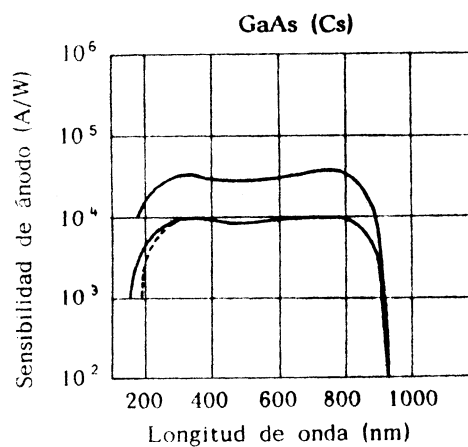
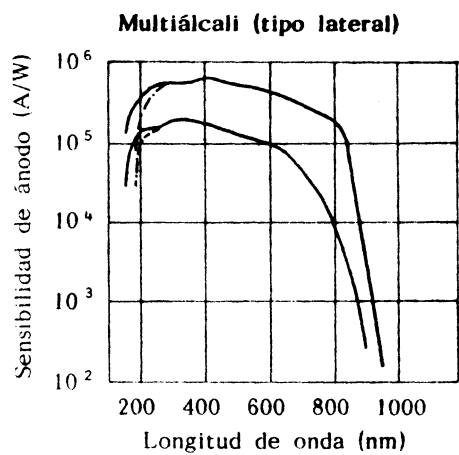
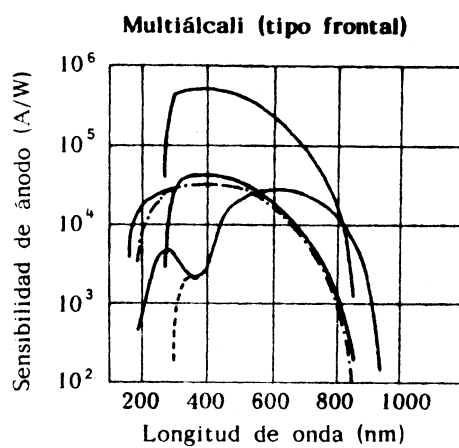


Figura 7.b

$$I_{dA} = G \cdot i_{d \text{ cátodo}}$$

$$S_A = G \cdot S \text{ cátodo} \quad .$$

La corriente total de ruido en un fotomultiplicador es la suma del ruido cuántico, de fondo y de corriente oscura.

$$\overline{i_N^2} = \frac{2\eta e^2 G B r}{h} (P_S + P_B + \frac{h\nu}{\eta e} I_{dc}) \quad .$$

r : fluctuación del factor de multiplicación (>1).
 $r = m / m-1$.

m : ganancia de extracción.

N : número de dinodos.

G : $m^2 N$

I_{dc} : corriente oscura en el cátodo.

Fotodetectores semiconductores

Son los únicos dispositivos capaces de absorber radiación de más de $1 \mu m$ de longitud de onda para producir una corriente de salida. Los de tipo extrínseco producen una sola clase de portadores cuando son excitados algunos de sus niveles de impurezas donoras o aceptoras.

La longitud de onda más alta corresponde a algunas decenas de meV. Un ejemplo es el Ge con impurezas de Au, Cu y Ag y la λ_c es de cientos de μm .

Los del tipo intrínseco producen electrones y huecos cuando absorben radiación de la longitud de onda correspondiente a la banda prohibida.

La longitud de onda más alta corresponde a una energía de aproximadamente 1 meV. Ejemplos de semiconductores intrínsecos son el Si, el Ge y el SPb.

Podemos distinguir dos tipos de detectores semiconductores:

- Fotoconductores
- Fotodiodos semiconductores

Fotoconductores

La radiación controla la conductividad del detector, produciendo una variación en su corriente o en la tensión aplicada sobre él. La corriente generada por la radiación incidente está formada por impulsos cuya carga no es igual a la de un electrón y tampoco es constante.

Si sobre una pastilla semiconductor sometida a una tensión incide radiación P se produce una corriente

$$i = \frac{\eta e}{h\nu} P g \quad .$$

El factor g llamado ganancia fotoconductiva puede ser mayor o menor que la unidad y depende de la relación entre el tiempo de vida de los portadores fotoinducidos y el tiempo de tránsito de los mismos entre los contactos eléctricos del detector. Es decir que por cada portador generado pueden entregarse al circuito g electrones en promedio,

$$g = \frac{\tau}{l/v} \quad .$$

v : velocidad de tránsito por campo.

l : separación entre contactos.

τ : tiempo de vida del portador.

La ganancia g puede valer hasta 10.

La eficiencia cuántica empleada en la fórmula de la corriente se define:

$$\eta = (1-r)\alpha \int_0^t e^{-\alpha x} dx \quad .$$

r : reflectividad de la superficie.

α : coeficiente de absorción de cada λ

t : espesor de la pastilla

x : dirección de incidencia de la radiación

Debido a la impredicción en el ancho de los pulsos fotoinducidos, por la fluctuación en el tiempo de recombinación, el ruido generado dobla al ruido cuántico de un detector fotoemisor y la carga es $2 g e$.

$$\overline{i_N^2} = 4 g e B i$$

se llama ruido de generación-recombinación.

Del análisis de la expresión completa de la relación señal-ruido en un fotoconductor que incluya todas las fuentes de ruido:

- cuántico
- de fondo
- de corriente oscura
- de amplificador ,

se deduce que la limitación a la sensibilidad no es el ruido cuántico debido a la resistencia no infinita del detector fotoconductor. Por lo tanto, la limitación estará dada por:

- fondo
- corriente oscura
- amplificador.

Considerando la limitación por amplificador, si se quiere tener una cierta frecuencia de corte f_c que estará fijada por la constante electrónica RC con

$$R = R_{\text{Det}} // R_{\text{Entr. Amplif.}}$$

$$R_{\text{Det}} = \Delta V / \Delta i \quad \text{del detector}$$

y

C: capacidad del dispositivo + la de entrada al amplificador.

Deberá fijarse una R que cumpla:

$$R = \frac{1}{2\pi f_c C}$$

y por consiguiente la detectividad estará fijada por esta condición, porque

$$NEP = \frac{2h\nu}{g\eta} \sqrt{\frac{KTB}{R}} \quad \text{y} \quad D^* = \frac{\sqrt{AB}}{NEP} .$$

$$D^* = K_1 \sqrt{1/fc} .$$

Por lo tanto, elevadas detectividades no son alcanzables cuando se desea operar a frecuencias altas.

Para el caso de fotoconductores intrínsecos, la detectividad está asociada con el valor de la resistencia medida sin radiación o resistencia oscura y, considerando la corriente oscura como principal fuente de ruido, la detectividad resulta:

$$D^* = \frac{\eta e}{h\nu} \sqrt{\frac{\mu\tau}{2e}} R_D \qquad D^* = K_2 \sqrt{R_D}$$

$\mu\tau$: movilidad por tiempo de vida de los principales portadores.

R_D : es constante para cada material y temperatura.

Esta ley sigue el comportamiento de los de los detectores limitados por ruido de fondo.

Cantidades que caracterizan a un detector fotoconductor.

- Temperatura 295° K, 193° K, 77° K.
- Detectividad (para una cierta λ pico y frecuencia de chopper).
- Responsividad (para la máxima tensión y en cortocircuito).
- Resistencia dinámica R_D .
- Constante de tiempo $\tau = 1/Wc$ (Wc : -3 db).

Fotodiodos semiconductores

Polarizados inversamente, como fotoconductivos, producen una corriente proporcional a la radiación incidente en las proximidades de la juntura.

Sin polarización, como fotovoltaicos, generan una tensión en sentido directo proporcional a la potencia irradiante. Ver Figura 8.

Como fotoconductores se los polariza con una tensión inversa, lo que produce una corriente inversa de saturación llamada corriente oscura.

El ruido de un fotodiodo semiconductor es similar al de uno de vacío, con la diferencia que en el primero puede circular corriente en ambos sentidos.

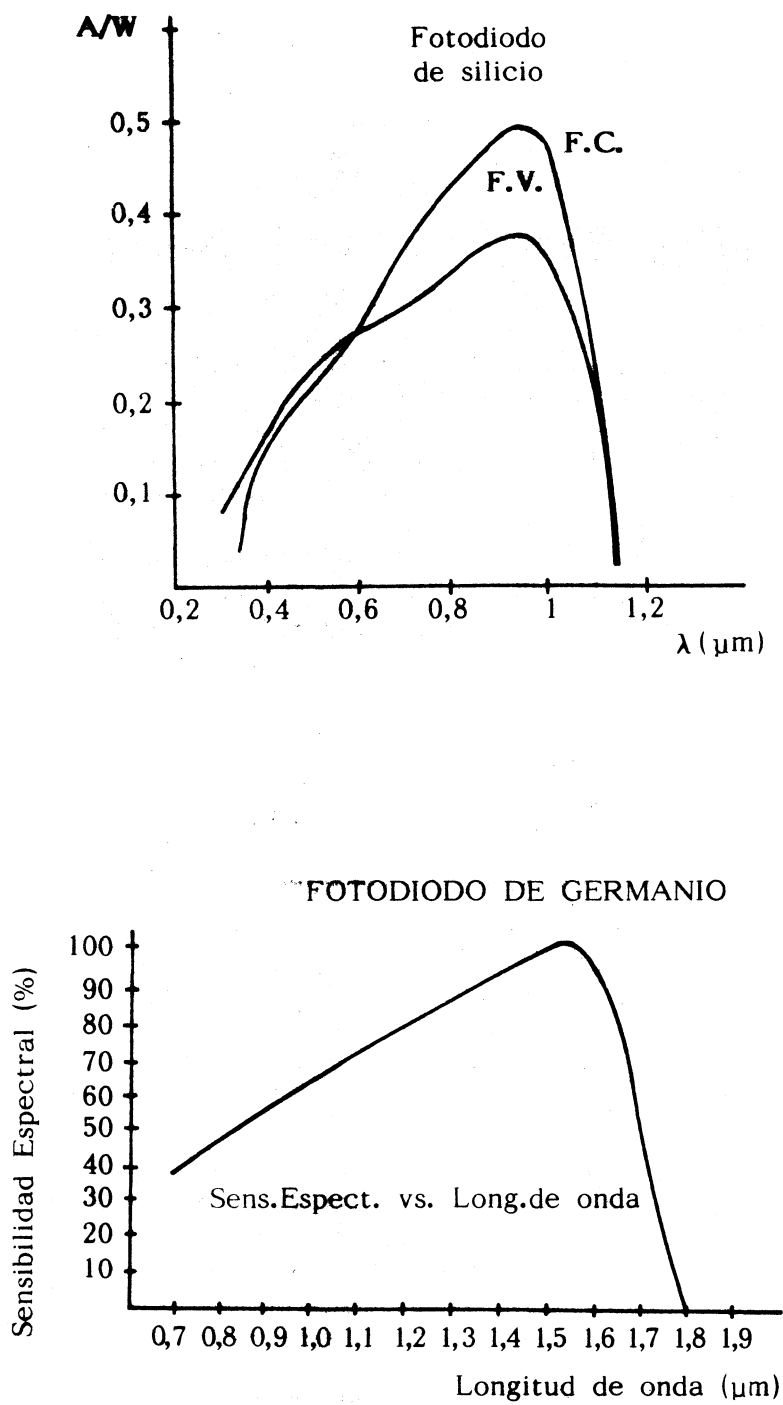


Figura 8

Si se contempla el ruido generado por las siguientes fuentes:

- señal más fondo
- corriente oscura
- polarización directa o nula,

la corriente total de ruido resulta, por tratarse de fuentes estadísticamente independientes:

$$\overline{i_N^2} = 2e [i_D e^{eV_A/KT} + i_D + \frac{\eta e}{h\nu} (P_S + P_B)] B \quad .$$

Además, la juntura presenta una resistencia dinámica con polarización nula $R = dV/dt$ que produce un ruido térmico

$$\overline{i_N^2} = \frac{4KTB}{R} \quad .$$

Desarrollando la expresión de la relación S/N que contempla todas las fuentes de ruido para polarización inversa, incluso la de la resistencia, resulta una fórmula idéntica a la del fotodiodo de vacío, con un ruido total:

$$\overline{i_N^2} = 2e [i_D + \frac{\eta e(P_S + P_B)}{h\nu}] B + \frac{4KTB}{R} \quad .$$

Como fotovoltaicos, sin polarización, la radiación que incide en la juntura genera portadores que se dirigen a la zona de carga espacial, produciendo una corriente inversa. Con el diodo acoplado a una resistencia infinita, la corriente es compensada por una acumulación de cargas de signo contrario, lo que produce una polarización directa del diodo.

La sensibilidad no es mejor que el caso fotoconductor, pero se usa cuando una polarización inversa produciría una elevada corriente oscura en un fotodiodo de baja resistencia inversa.

La corriente directa es $e^{eV/KT}$ veces superior a la de saturación y de sentido contrario y compensa a la corriente fotoinducida.

$$\frac{\eta e}{h\nu} (P_S + P_B) = i_D (e^{eV/KT} - 1) \quad .$$

La tensión inducida vale

$$V = \frac{KT}{e} \ln\left(\frac{i_B + i_S + i_D}{i_D}\right) \quad \frac{KT}{e} = 26 \text{ mV a } 25^\circ \text{ C.}$$

Debido a la fluctuación de la corriente en ambos sentidos, la corriente de ruido cuántico del fotovoltaico es doble a la del fotoconductor.

$$\overline{i_N^2} = 4eB (i_S + i_B + i_D) \quad .$$

La relación señal-ruido difiere de la del diodo polarizado inversamente en que, debido a la doble fluctuación de la corriente de ruido, la relación S/N es la mitad de la del fotoconductor. La resistencia dinámica del fotodiodo fotoconductor se obtiene de:

$$\frac{1}{R_D} = \frac{1}{dV} [i_D (e^{eV/KT} - 1)] = \frac{e}{KT} (i_S + i_B + i_D) \quad ,$$

y queda conectada en paralelo con la carga R_L .

Ejemplos de fotodiodos son:

Si λ_c aprox. $1,2 \mu\text{m}$

Ge λ_c aprox. $1,8 \mu\text{m}$

InSb λ_c aprox. $5 \mu\text{m}$

Pb - Sn - Te

Pb - Sn - Se más de $30 \mu\text{m}$

Hg - Cd - Te

} Trimetálicos

Cantidades que caracterizan a un fotodiodo

- Detectividad específica D^* (para la λ pico, frecuencia de troceado de la radiación y ancho de banda).
- Respuesta espectral (λ de corte máx y mín).
- Corriente oscura (para una V inversa y temperatura específica).
- Corrientes máximas en régimen pulsado y continuo.
- Sensibilidad y responsividad $\eta e/h\nu$ para una fuente de λ determinada

y tensión dadas.

Circuitos para fotoconductores y fotovoltáicos ver Figura 9.

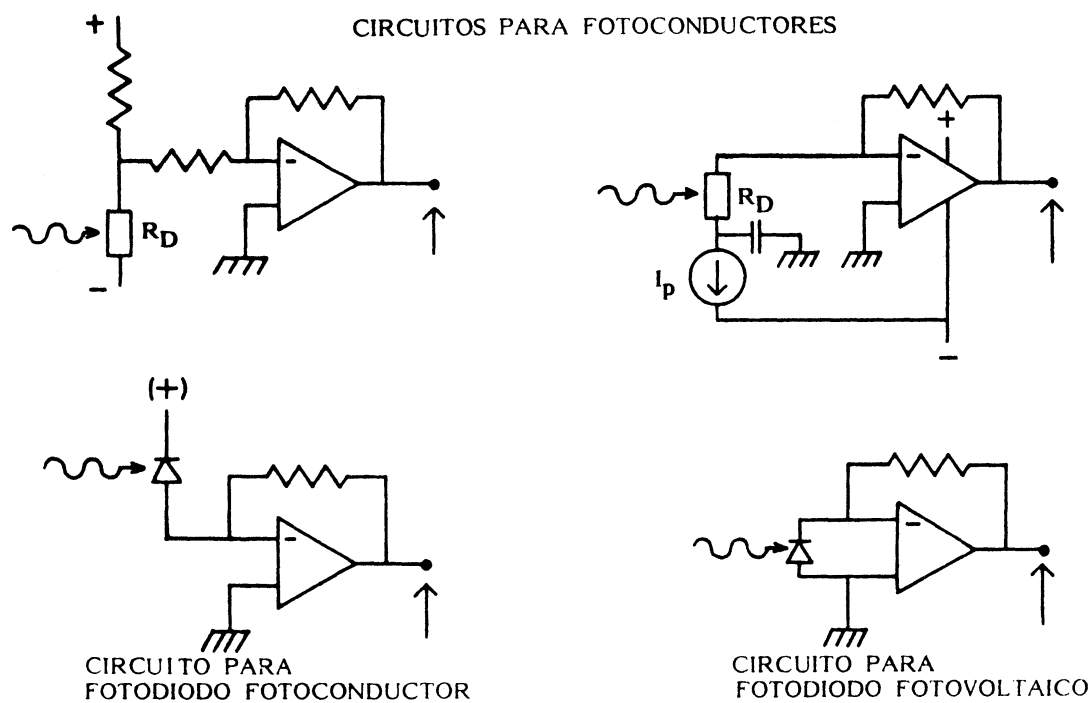


Figura 9

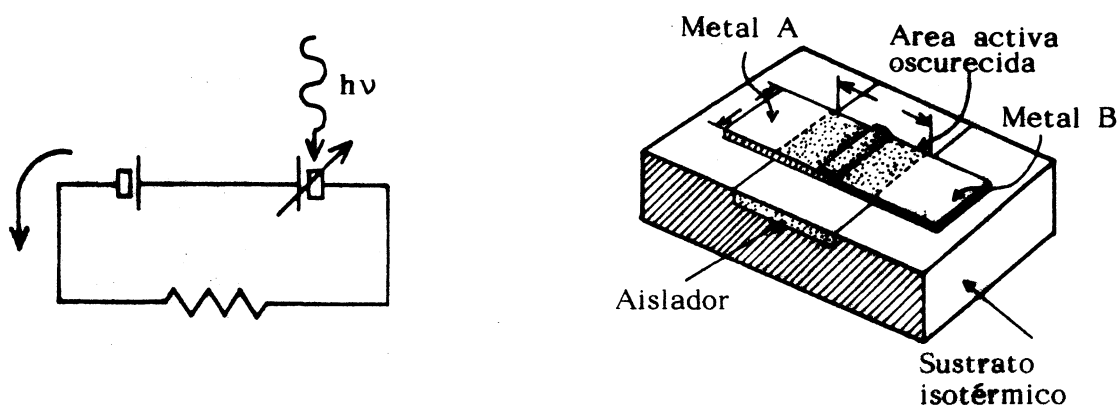


Figura 10

Fotodiodos de avalancha

Cuando los fotodiodos se polarizan con una tensión inversa elevada que produce un campo eléctrico de aproximadamente 10^5 V/cm, los portadores que atraviesan la juntura, a gran velocidad ionizan nuevos átomos generándose un proceso de avalancha y multiplicación de la corriente de salida. Es el proceso análogo al de un fotomultiplicador pero en semiconductores.

Los electrones y huecos generan a su paso nuevos pares de portadores y contribuyen a incrementar la corriente inversa total. Se define un factor de multiplicación M para el diodo como relación de las corrientes multiplicada a la fotoinducida.

Eligiendo adecuadamente los materiales de los fotodiodos se obtienen factores de multiplicación de hasta $M=100$ y tiempos de crecimiento muy cortos, hasta de $\tau_r < 0,1$ ns, pero el ruido producido es superior al del fotodiodo común debido al proceso de la multiplicación por colisiones, $M = I/I(0)$. La corriente eficaz de ruido resulta

$$\overline{i_N^2} = 2eM^3 \ln(O) B \quad , \quad \text{con } \ln(O)$$

la corriente de electrones antes de la multiplicación.

Si los coeficientes de ionización se eligen diferentes para los dos materiales p y n , puede obtenerse un parecido casi total a la corriente de ruido de un fotomultiplicador:

$$\overline{i_N^2} = 2e M^{2,2} \ln(O) B \quad .$$

Se han desarrollado fotodiodos de avalancha para uso en fibras ópticas y en pulsos ultracortos, de:

Si	hasta	$\lambda = 1,3 \mu\text{m}$
AsGa	hasta	$\lambda = 0,9 \mu\text{m}$
InSb	hasta	$\lambda = 5 \mu\text{m}$

Detectores térmicos

Los detectores térmicos convierten la energía radiante en un cambio de temperatura del dispositivo y esto produce, a su vez, la modificación de sus propiedades eléctricas: resistividad, tensión o corriente inducida.

Los principales detectores térmicos son:

- Bolómetro termistor
- Termopila
- Piroeléctrico .

El bolómetro funciona como fotoconductor y su resistencia es modulada por la temperatura. Los materiales que lo forman, por ejemplo delgadas películas semiconductoras de óxidos de Mn, Ni, CO, muestran un elevado coeficiente negativo de temperatura (NTC).

Se conectan de a pares montados sobre un mismo disparador para compensar la T_{amb} y se hace incidir la radiación sobre uno de ellos:

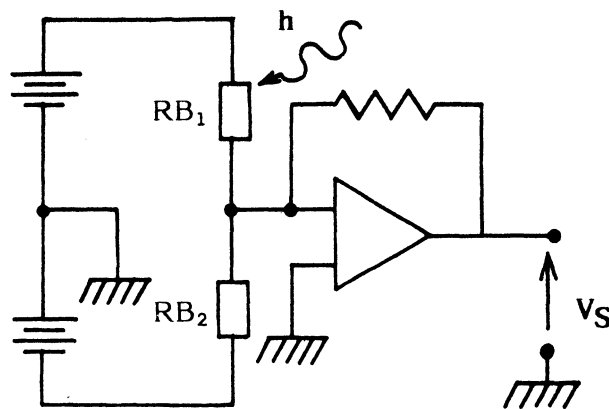


Figura 11

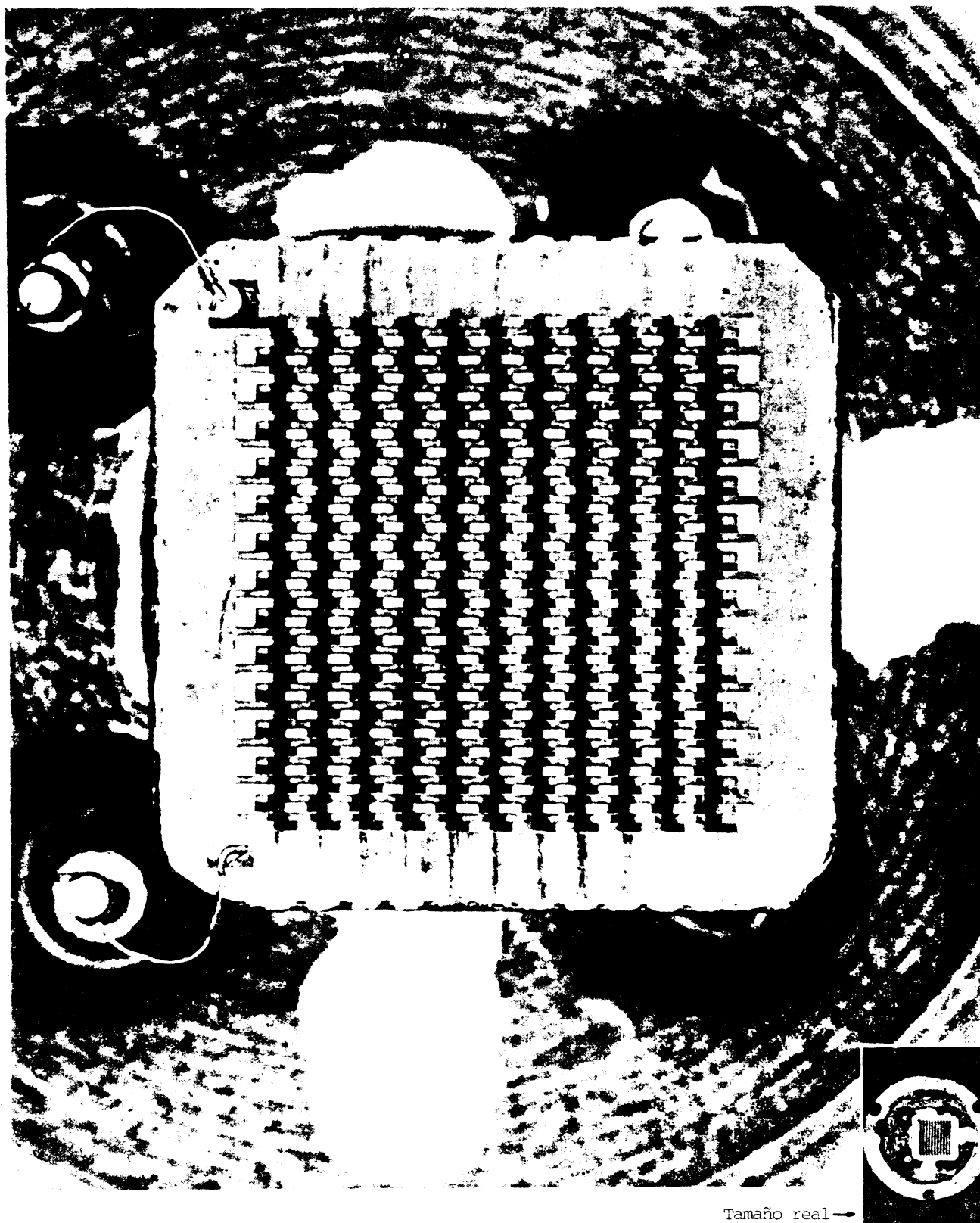


Figura 12

El desbalance producido en RB_1 por la radiación es amplificado.

Cantidades que caracterizan a un termistor:

- Constante de tiempo $T = 1/Wc$ ($Wc = 1/R_{det} \cdot C_{circ}$)
- Tensión de polarización
- Resistencia eléctrica a $T_a = 25^\circ C$
- Responsividad V/W para una tensión dada $R_v(V/W)$
- NEP (Pot. equiv. al ruido) $\sim 10^{-11} W/\sqrt{Hz}$.

Termopilas.

Están formadas por la conexión en serie de gran número de termocuplas elementales que funcionan de a pares, estando una de ellas tratada para absorber radiación y la siguiente como referencia. La tensión generada es independiente de la longitud de onda y se construyen con bismuto y antimonio en películas delgadas. Las termocuplas que absorben radiación son pintadas de negro. Ver Figuras 11 y 12.

Cantidades que caracterizan las termopilas:

- Número de junturas: 20 - 300
- Temperatura: $-65^\circ C$ a $+125^\circ C$
- Resistencia: 200 - 5000 Ω
- Constante de tiempo: 200 ms - 15 ms
- NEP: $1 \cdot 10^{-9} W / \sqrt{Hz}$
- D^* : $1,7 \cdot 10^{-9} cm \sqrt{Hz}/W$.

Detectores piroeléctricos.

El material de los detectores piroeléctricos tiene una polarización eléctrica espontánea orientada a lo largo de un eje polar del cristal, cerámica o plástico.

La radiación absorbida se convierte en calor que altera el espaciado atómico del material por expansión térmica y esto cambia la polarización, generando una variación de la carga eléctrica superficial en la dirección perpendicular al eje polar.

Si se colocan electrodos sobre esas superficies, se puede medir una tensión proporcional al cambio de temperatura del material.

Algunos de los materiales usados son:

- TGS : sulfato de triglicina

PVF₂: fluoruro de polivinildeno

SBN : niobato de Sr y Ba ,

y, más modernamente, el tantalato de litio, LiTaO₃ .

El detector piroeléctrico es sensible sólo a la radiación pulsada o chopeada, porque sólo responde a las variaciones de temperatura; por consiguiente, no responde a la radiación constante de fondo $I = p(T) A \, dT/dt$. (I : corriente piroeléctrica, A : área, T : temperatura).

Comparado con el detector de SPb, la detectividad del piroeléctrico es dos órdenes de magnitud inferior, pero como todos los detectores térmicos tienen la ventaja de absorber radiación de cualquier longitud de onda, desde los rayos X hasta las microondas.

Son superiores a los demás detectores térmicos en cuanto a su respuesta a las altas frecuencias o radiación pulsada a varios cientos de MHz. Cuando se lo emplea como medidor de energía, almacena la carga total generada por una señal transitoria integrándola en el tiempo. Adaptado a una línea de transmisión de 50Ω es capaz de reproducir señales pulsadas con tiempos de crecimiento de menos de 0,5 ns ($\tau_r < 0,5 \text{ ns}$).

Las frecuencias límites de utilización están determinadas por los siguientes factores:

- Baja frecuencia: constante de tiempo térmica τ_T .
- Alta frecuencia: constante de tiempo eléctrica τ_E .

El detector piroeléctrico y su etapa amplificadora asociada pueden representarse por una resistencia en paralelo con un capacitor que forman la constante de tiempo eléctrica $\tau_E = R_E C_E$.

También, desde el punto de vista térmico, el detector piroeléctrico es un dispositivo que tiene capacidad térmica C_T (almacena calor) y resistencia térmica R_T (al flujo del calor), las que resultan en una constante de tiempo térmica

$$\tau_T = R_T C_T \quad .$$

La respuesta temporal del detector piroeléctrico está determinada por estas dos constantes y vale:

$$R_v(t) = \alpha R_i R_E (1 - e^{-t/R_E C_E}) e^{-t/R_T C_T} \quad ;$$

$R_v(t)$ es la responsividad de tensión en función del tiempo de excitación en forma de escalón de radiación (V/W) .

R_i : responsividad de corriente (A/W)

R_E : resistencia eléctrica (Ω)

C_E : capacidad eléctrica (pF)

R_T : resistencia térmica ($^{\circ}\text{C}/\text{W}$)

C_T : capacidad calórica (Joule/ $^{\circ}\text{C}$)

Ver Figuras 13, 14 y 15, Detectores de Infrarrojo.

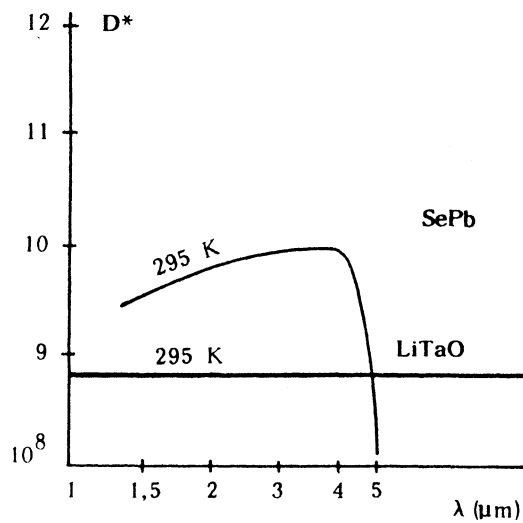
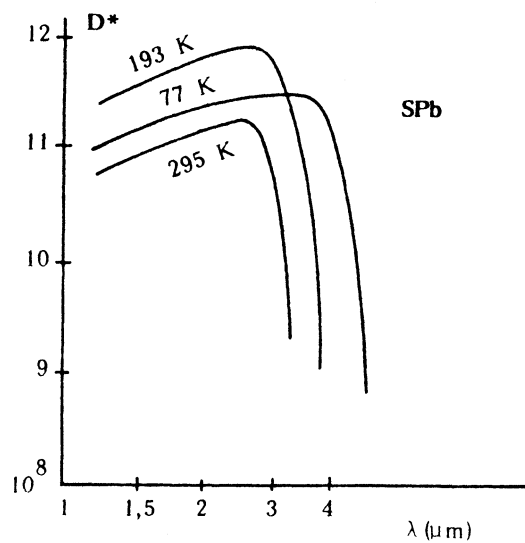


Figura 13

DETECTORES PARA INFRARROJO

TIPO	Temp. °K	Refrigerante	Ventana	λ pico	VALORES TÍPICOS			
					D^* pico 10^{10} cm ² /Hz/W	R_i pico (A/W)	R_D (Ω)	τ
PbS	295	-	-	2,4	8 - 15	0,5 - 4	0,2 M - 1M	0,1 - 0,5 ms
	193	Freón T.elect.	Si(2,7 μ)	2,8	40 - 60	1 - 10	0,5 M - 5M	2 - 5 ms
			Zafiro	2,4				
	77	Nitrógeno	Si(2,7 μ)	2,8	15 - 25	0,1 - 1	1 M - 5M	2 - 5 ms
			Zafiro	3,2				
PbSe	295	-	-	3,8	0,1 - 0,5	0,004 - 0,03	0,5 M - 5M	1 - 3 μ s
	193	Freón T.elect.	Si(4,5 μ)	4,8	1,5 - 3,0	0,01 - 0,05	10 M - 50M	10 - 40 μ s
			Zafiro	4,8				
	77	Nitrógeno	Si(4,5 μ)	5,0	1,5 - 3,0	0,04 - 0,4	0,3 M - 5 M	15 - 50 μ s
			Zafiro	5,0				
InSb	77	Nitrógeno	Zafiro	5,0	8,0 - 20,0	1,5 - 2,0	0,2 M - 1 M	20 - 200 ns
			Si(4,5 μ)	5,0				
Ge: Au	77	Nitrógeno	Irtran 2	5,0	0,15 - 0,7	0,01 - 0,60	20 K - 10 M	1 - 100 ns
Ge: Hg	4	Helio	Irtran 2 o KRS-5	10,4	1,0 - 2,0	0,03 - 20,0	20 K - 5 M	0,3 - 100 ns
HgCdTe (FC)	77	Nitrógeno	Irtran 2	12,0	0,5 - 4,0	3,0 - 30,0	10 - 100	0,2 - 2,0 μ s
	77	Nitrógeno	Irtran 4	15,0				
HgCdTe (FC)	193	Freón 13 T.E.C.	Zafiro	4,6	3,0 - 6,0	3,0 - 30,0	100 - 1000	0,02 - 10 μ s
	295	-	Zafiro	3,9	0,05 - 0,10	0,7 - 2,5	50 - 500	50 - 100 ns
HgCdTe (FV)	77	Nitrógeno	Irtran 2	10,6	1,0 - 4,0	5,0 - 6,0	50 - 200	50 - 100 ns
Ge: Cd	4	Helio	KRS-5	19,0	2,0 - 3,0	0,5 - 3,0	20 K - 200 K	10 - 100 ns
Ge: Cu	4	Helio	KRS-5	21,0	1,3 - 3,0	0,1 - 0,3	20 K - 5 M	0,3 - 100 ns
Ge: Zn	4	Helio	KRS-5	37,0	1,0 - 2,0	0,1 - 0,5	20 K - 200 K	10 - 100 ns

Figura 14

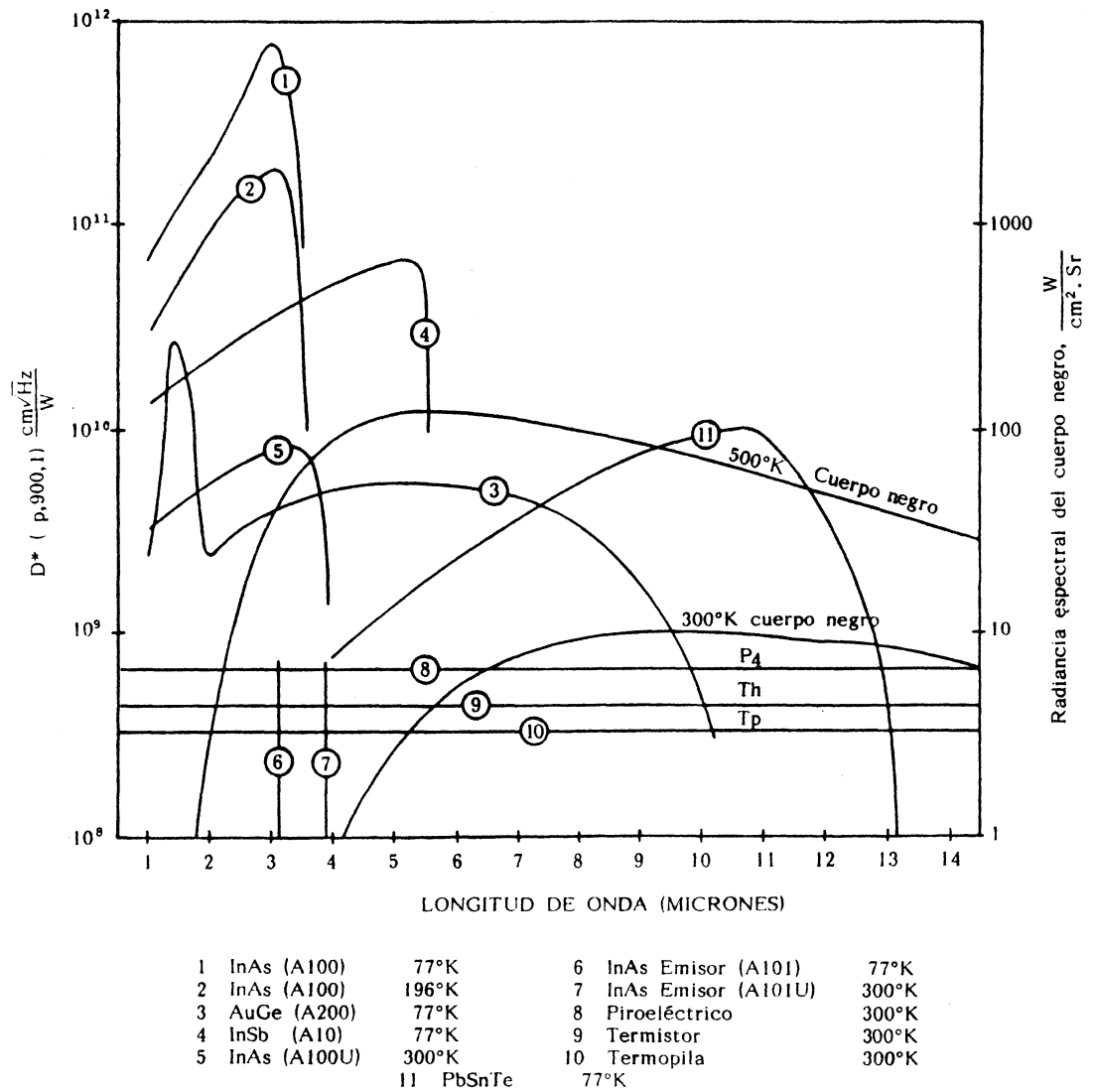


Figura 15

Bibliografia

- R.H. Kingston, "Detection of Optical and Infrared Radiation", Springer-Verlag, Optical Sciences 10.
- EG&G Application Notes D3011C-1 y D3000C-2.
- Thermal and Photon IR Detectors, Barnes 2-351.
- Hamamatsu Photomultiplier Tubes Cat. Nº SC-1-3.
- Optical Radiation Measurements. James Rockwell. Selected Reprints from EOSD Magazine.
- Barnes Handbook of IR Measurement. Octubre 1978.
- Infrared Detectors. Hudson y Hudson Benchmark Papers in Optics.

COMUNICACIONES ATMOSFERICAS CON LASER

A. P. Laquidara

Introducción

La aparición del láser en 1960 permitió disponer de una fuente de radiación luminosa de características adecuadas para la transmisión de información a distancia.

La baja divergencia de un haz láser muy intenso y el enorme ancho de banda disponible, atrajeron el interés de los ingenieros ocupados en la comunicación de gran caudal de información. Un canal de radio de microondas en la frecuencia de, por ejemplo, 4,7 GHz ($4,7 \cdot 10^9$ Hz), es capaz de conducir alrededor de 1000 conversaciones telefónicas simultáneas. Si comparamos esta frecuencia portadora con la del láser de He-Ne, $\nu_p = 470.000$ GHz ($4,7 \cdot 10^{14}$ Hz), la extrapolación llevaría a una cifra de 100 millones de conversaciones o 60.000 canales de televisión simultáneos.

Además de las barreras tecnológicas que habría que superar para lograr esos valores, carece de sentido por ahora intentar la transmisión de tal enorme cantidad de información por un solo canal.

A pesar de las barreras, el desarrollo de nuevos materiales está convirtiendo a la luz en el medio más conveniente para la transmisión de información. El reemplazo del electrón por el fotón posibilita el procesamiento de la información a mayor velocidad y con un considerable ahorro de espacio y energía.

Las fibras ópticas están reemplazando a los conductores de cobre con las ventajas mencionadas y a causa de su baja atenuación, se ha logrado transmitir información a alta velocidad sobre distancias de más de 100 Km sin intercalar estaciones repetidoras inevitables en los cables eléctricos de igual longitud.

Si comparamos las transmisiones de radio de microondas por el espacio con las ópticas, las primeras aventajan al láser dentro de la atmósfera, debido al efecto dispersivo que las gotas de agua y la turbulencia producen sobre el haz láser. Pero fuera de la atmósfera, las emisiones

láser pueden ser detectadas aunque su potencia sea muy inferior a las equivalentes de radio (naves y satélites espaciales).

Para imprimir a un haz luminoso un mensaje, es necesario modularlo. El límite a la capacidad de un canal óptico lo fija la velocidad del modulador para producir sobre el rayo un cambio de intensidad, frecuencia o fase, de acuerdo con la información a transmitir.

La radiación láser puede polarizarse y mediante un par de polarizadores y un modulador electro-óptico se puede modular en amplitud un haz continuo en forma analógica o digital.

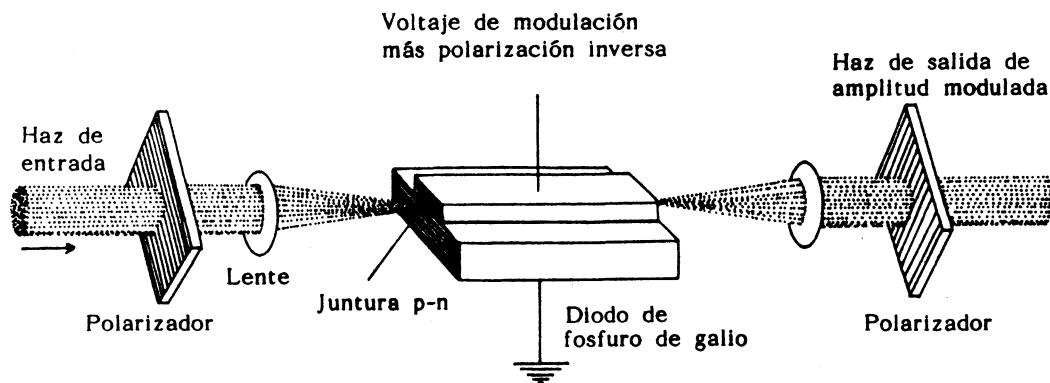


Figura 1

También se modulan los láseres de semiconductores en forma pulsada, inyectando corriente por intervalos muy breves en la juntura.

Los diodos láser son una clase especial de diodos emisores de luz o LEDs.

En diodos con junturas formadas por elementos como el arsénico y galio (AsGa), el pasaje de corriente en forma directa produce la emisión de radiación en el plano de juntura. Cuando esta juntura, que en general es de construcción laminar, está terminada en caras reflectantes que reinyectan la luz hacia el interior, se obtiene una cavidad resonante que facilita la amplificación de determinadas bandas de la radiación, convirtiéndose este diodo en un láser.

Las características de la radiación emitida por un láser y un LED contruidos con iguales materiales, difieren en la intensidad, el ancho espectral de la luz y la longitud de onda.

Cuando la densidad de corriente inyectada en la juntura de un diodo láser es inferior a un valor llamado umbral, la radiación es como la de un LED: incoherente y débil. Pero al superar ese valor, la radiación crece en intensidad muy rápidamente y se vuelve monocromática.

El desarrollo de láseres de semiconductor de estructura multicapa de elementos como el As, Ga, Al, In, P, permitió reducir la corriente de excitación a algunos miliamperes en diodos para fibras ópticas y decenas de amperes en diodos para comunicaciones atmosféricas.

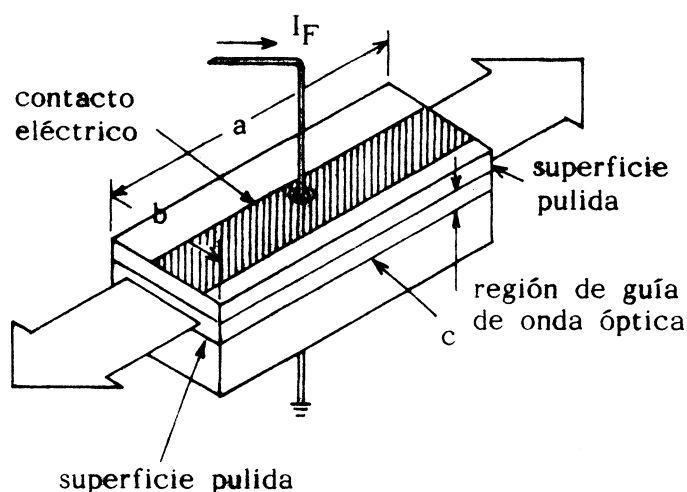


Figura 2. Estructura típica de un diodo láser de inyección. a: longitud de la cavidad. b: ancho de la emisión. c: juntura PN.

La longitud de onda y la potencia emitida por un diodo láser son fuertemente dependientes de la temperatura. En los siguientes gráficos se observa la influencia de los cambios de temperatura de la juntura de un diodo láser pulsado sobre la potencia pico y la corriente umbral.

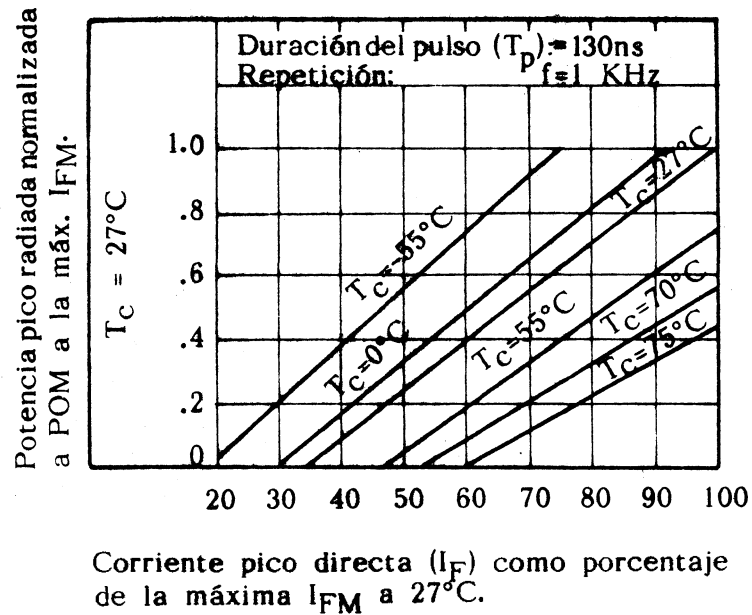
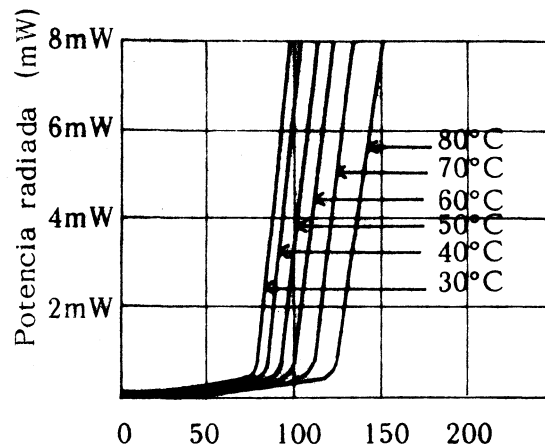


Figura 3



Potencia de salida vs. corriente directa a varias temperaturas.

Figura 4

En los siguientes gráficos se observa el corrimiento de la longitud de onda y la reducción de la potencia pico emitida al aumentar la frecuencia de repetición de los pulsos o ciclo de operación.

$$d = t_p / T$$

t_p : ancho del pulso de excitación.

T : período de la repetición de pulsos.

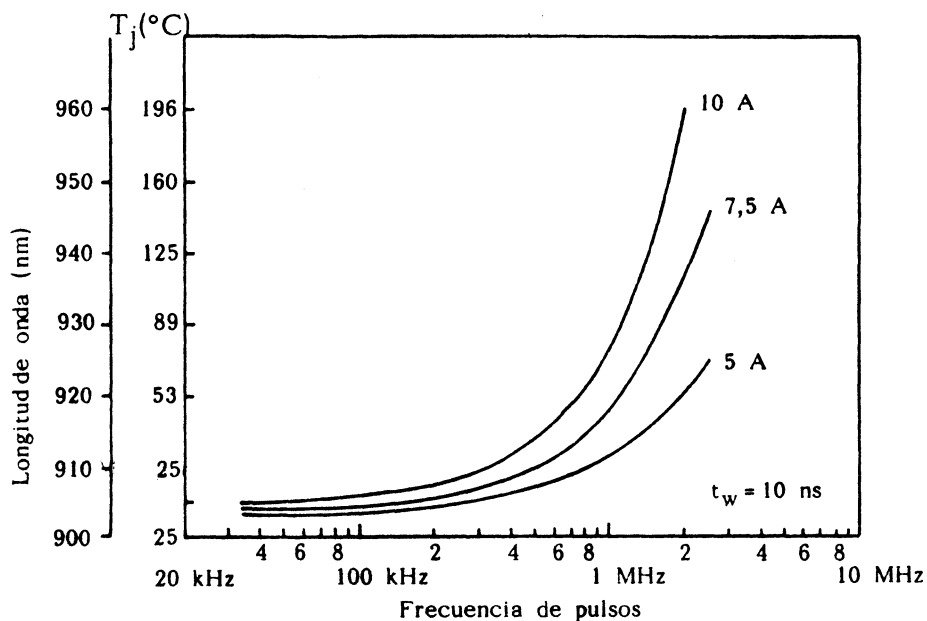


Figura 5. Temperatura de juntura del láser vs. frecuencia de pulsos.

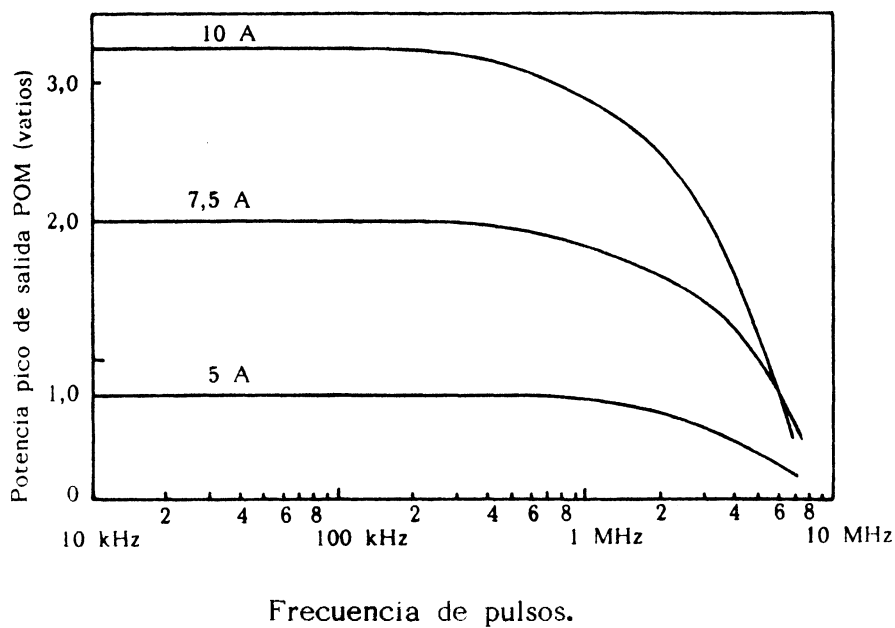


Figura 6. Potencia pico de salida vs. frecuencia de pulsos.

La detección de la luz emitida por un diodo láser se realiza por medio de fotodiodos de silicio que pueden conectarse según dos configuraciones básicas:

Fotodiodo Fotovoltaico: La radiación incidente sobre el fotodiodo genera una tensión en sus extremos proporcional a la potencia incidente. Esta configuración es apta para la detección de señales pulsadas de no menos de algunos microsegundos de duración.

Fotodiodo Fotoconductor: La corriente generada en el fotodiodo proporcional a la potencia incidente puede ser convertida en tensión por medio de un amplificador operacional. Esta conexión permite detectar pulsos de hasta algunos nanosegundos de duración.

Los tipos de detectores más empleados en las comunicaciones por diodos láser son los fotodiodos de silicio del tipo PIN y Avalancha.

El fotodiodo PIN consta de tres capas semiconductoras: P, N y de alta resistividad: I.

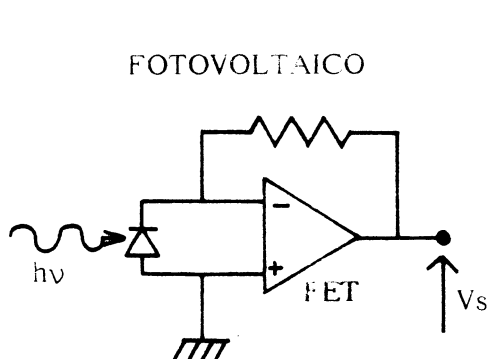


Figura 7

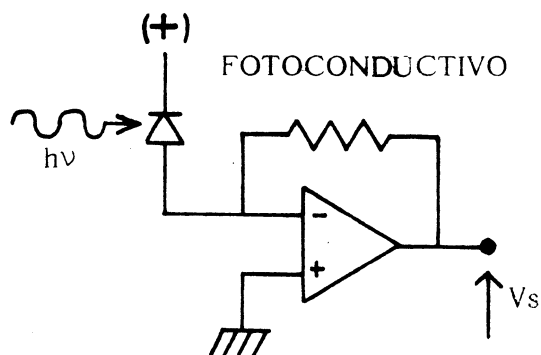


Figura 8

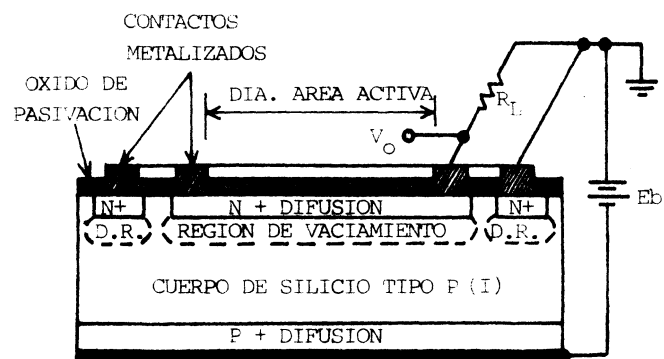
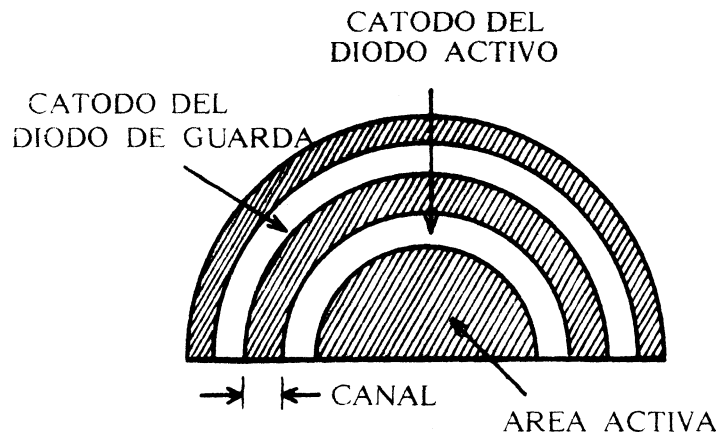


Figura 9

Al soportar tensión inversa la capa I queda desprovista de portadores y al recibir radiación de la longitud de onda adecuada (aproximadamente $\lambda=1 \mu\text{m}$) se inducen e^- y h^+ que viajan a las capas N y P respectivamente, generándose una corriente inversa proporcional a la potencia recibida.

La corriente generada por la señal y la radiación de fondo captada simultáneamente vale

$$i = \frac{\eta e}{h\nu} (P_s + P_B) \quad ,$$

donde:

- η : eficiencia cuántica
- e : carga del electrón
- h : constante de Planck
- ν : frecuencia de la luz
- P_s : potencia de señal
- P_B : potencia de fondo.

Esta corriente genera en el fotodiodo un ruido debido al flujo discontinuo de la radiación, llamado ruido cuántico o de fotón.

$$\overline{i_N^2} = 2 e i \Delta f n \quad ,$$

siendo $\Delta f n$ el ancho de banda del ruido.

El amplificador al cual se conecta el fotodiodo, produce un ruido propio generado térmicamente en su resistencia de entrada:

$$\overline{i_T^2} = \frac{4KT\Delta f n}{R_e} \quad ,$$

con:

- K : constante de Boltzman
- T : temperatura absoluta de la resistencia
- R_e : resistencia de entrada.

Considerando el conjunto formado por fotodiodo y amplificador, la

relación señal-ruido total se obtiene como

$$S/N = \frac{i_s^2}{i_N^2 + i_T^2} .$$

Cuando esta relación vale 1, la potencia obtenida de ella es el NEP o potencia equivalente al ruido. La potencia de señal recibida debe ser superior al NEP para obtener buena detectividad.

El fotodiodo de avalancha tiene una estructura similar al diodo PIN, pero es operado a una tensión inversa más elevada que establece un campo eléctrico intenso en la juntura. Los portadores fotoinducidos en el material I son acelerados hacia la juntura donde al colisionar ionizan más átomos, multiplicando la corriente inicial por un factor M . Esta multiplicación interna mejora la relación señal-ruido del detector y contribuye a incrementar el valor de la relación S/N del conjunto detector-amplificador. El factor de multiplicación vale en general $M = 100$ y el ruido cuántico del fotodiodo de avalancha es:

$$\overline{i_N^2} = 2 e \cdot M^{2,3} i \Delta f n .$$

La ganancia óptima se obtiene cuando el ruido propio del detector iguala al del amplificador.

$$M_{opt}^{2,3} \cdot 2 e i \Delta f n = \frac{4KT \Delta f n}{R_e}$$

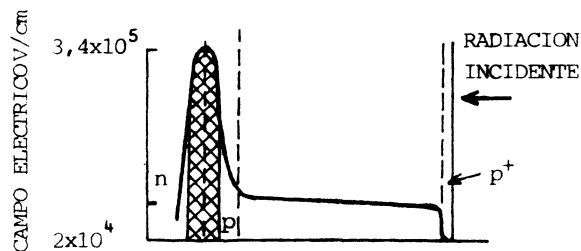
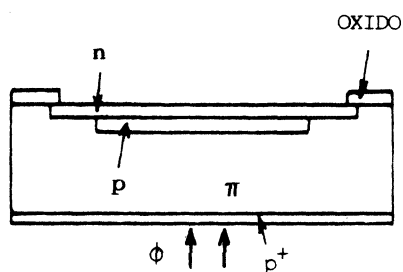


Figura 10

Comunicaciones e instrumentos ópticos

Para dirigir el haz luminoso emitido por un diodo láser hacia un punto distante es necesario el empleo de un adecuado sistema óptico transmisor.

El diodo láser emite un haz que diverge con un ángulo total de entre 10° y 30° . La energía contenida en este ángulo sólido debe enfocarse en el blanco distante del receptor por medio de, por ejemplo, una lente convergente. La lente hará que los rayos divergentes se propaguen paralelos y en forma rectilínea. Estos rayos se propagarían de esta forma indefinidamente en el vacío si la lente fuera libre de aberraciones, y las dimensiones de la fuente emisora fueran nulas.

Las dimensiones reales de la fuente emisora, las lentes y la atmósfera hacen que el límite práctico alcanzable sea el de un ángulo de aproximadamente 1 miliradian.

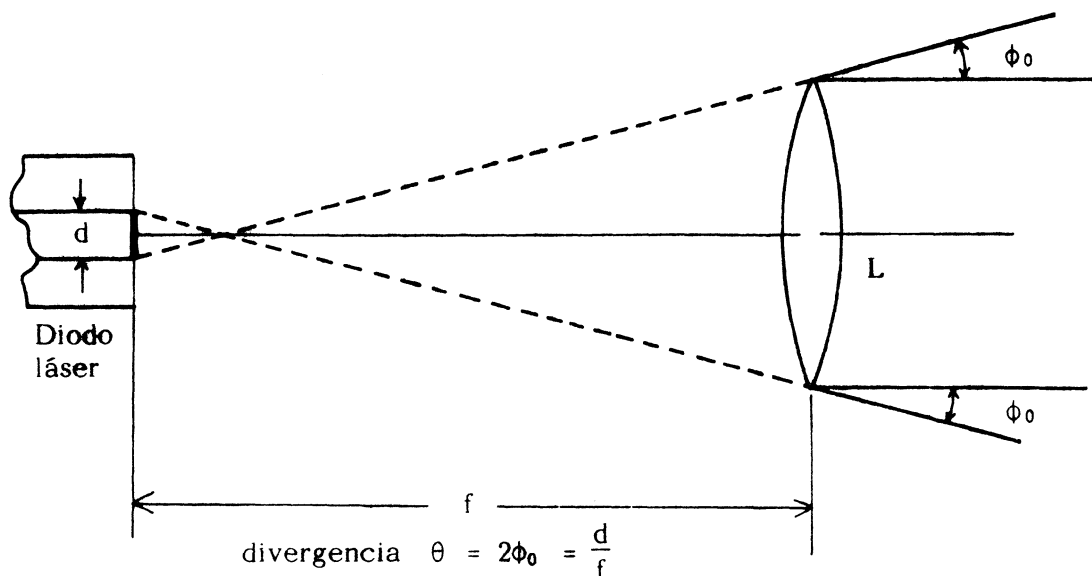


Figura 11

Para coleccionar la energía en el receptor también se emplean sistemas ópticos convergentes que enfocan la radiación sobre la superficie del fotodiodo detector.

Las dimensiones del detector dan por resultado un campo visual de recepción de aproximadamente 5 miliradianes.

La atmósfera: Canal de comunicación

La señal emitida por un diodo láser, al atravesar la atmósfera en dirección horizontal, sufre atenuación debido a tres factores principales:

1. Absorción por moléculas de H_2O , CO_2 y O_3 .

Las moléculas de estos gases absorben la radiación láser y producen una atenuación considerable a ciertas longitudes de onda. Las longitudes de onda empleadas en comunicaciones aéreas son aquéllas para las cuales existen "ventanas atmosféricas" de máxima transmitancia.

La radiación de los diodos láser coincide con una de estas ventanas atmosféricas, cuya transmitancia es del 80% al 90%.

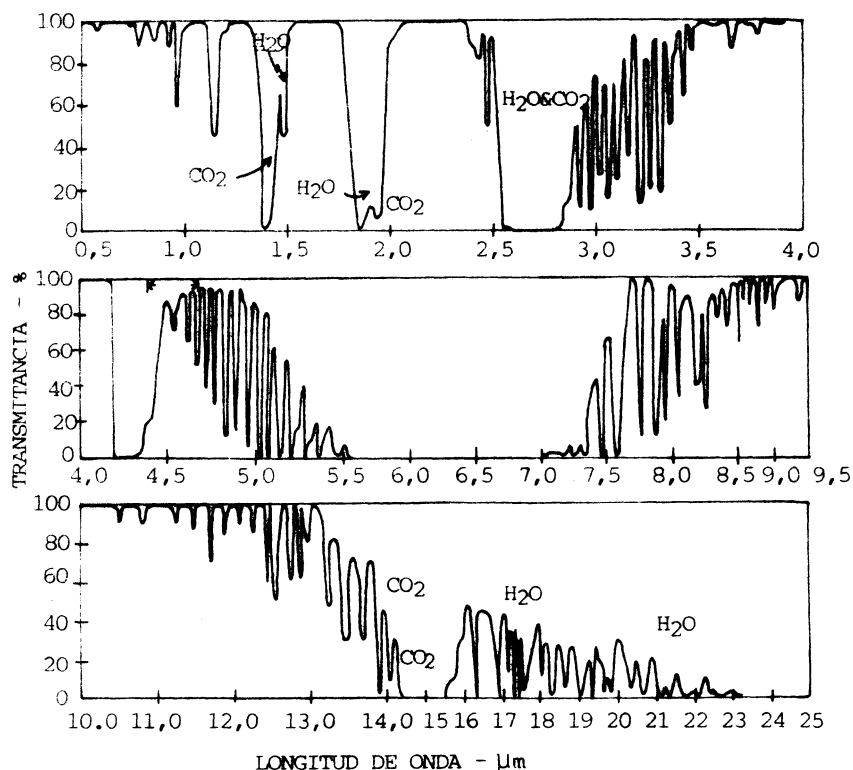


Figura 12. Transmitancia de 1000 pies de camino aéreo horizontal a nivel del mar conteniendo 5,7 mm de agua precipitable a 79° F.

2. Dispersión de la radiación en partículas.

La radiación es dispersada cada vez que encuentra en su camino partículas cuyas dimensiones se aproximan a la longitud de onda emitida. La transmitancia de la atmósfera afectada por dispersión vale:

$$T = e^{-\sigma R} \quad ; \quad \sigma = \sigma_{0_3} \quad \sigma_{Ry} \quad \sigma_{Mie} \quad ,$$

donde σ coeficiente de extinción; R distancia ; σ_{Ry} absorción Rayleigh por N_2 y O_2 ; σ_{Mie} dispersión en polvo y aerosoles.

El término dominante en la dispersión de la radiación de diodos láser es el de Mie,

$$\sigma_{Mie} = (3,91/Rv) \cdot (\lambda_c/0,55)^{-0,585} \cdot Rv^{1/3} \quad ,$$

con σ (1/Km); Rv (Km); y λ_c (μm) . Rv es la distancia de visibilidad atmosférica registrada en el momento de la transmisión.

Cuando llueve en forma continuada, el coeficiente de extinción debe incrementarse en un término,

$$\sigma_{lluvia} = 0,248 \cdot r^{0,67} \quad ; \quad r \text{ (mm/hora)} \quad .$$

Existe un límite máximo de distancia alcanzable con diodos láser en la atmósfera, de aproximadamente 60 Km en condiciones ideales de visibilidad y para una tasa de error de bits o $BER = 10^{-6}$ en transmisión por pulsos codificados.

3. Turbulencia.

Los efectos principales del calentamiento de las masas de aire atmosférico a diferentes horas del día son:

- 3.1. Rotación del haz: La variación del índice de refracción del aire por el desplazamiento de masas frías y calientes produce cambios en la dirección del rayo emitido que puede alejarse del blanco receptor interrumpiendo la comunicación.
- 3.2. Centelleo: Los cambios incesantes en la fase y amplitud de los rayos emitidos producen distintos recorridos que al converger sobre el receptor interfieren, atenuando fuertemente la señal. La pérdida de señal puede ser de hasta 60 db y la frecuencia del centelleo oscila entre 1 y 100 Hz.

Seguridad con láseres

El uso de equipos transmisores a diodos láser implica el riesgo de que el operador pueda sufrir daños irreparables en la visión. La radiación pulsada o continua de diodos láser es absorbida por la retina y cuando es focalizada en puntos pequeños produce la destrucción del tejido, llegando a producir hasta la pérdida de la visión. Estos graves efectos deben evitarse conociendo y divulgando el peligro que representa el manejo de emisores láser y debe evitarse dirigir la radiación hacia los ojos, ya sea en forma directa o reflejada.

Muchas instituciones internacionales dictan las regulaciones que deben respetarse al fabricar, instalar y operar láseres.

Los diodos láser para comunicaciones aéreas más empleados son los de AsGa y AsGaAl. Para ellos existen los siguientes límites de potencia o energía que el ojo puede aceptar sin dañarse:

AsGaAl	$\lambda = 850 \text{ nm}$	$\phi \leq 76 \text{ } \mu\text{W}$
Pulsado t_p	$= 100 \text{ ns, } 10 \text{ Hz} - 100 \text{ KHz}$	$Q_p \leq 0,4 \text{ } \mu\text{Joule}$
AsGa	$\lambda = 904 \text{ nm}$	$\phi \leq 97,5 \text{ } \mu\text{W}$
Pulsado t_p	$= 100 \text{ ns, } 10 \text{ Hz} - 100 \text{ Hz}$	$Q_p \leq 0,5 \text{ } \mu\text{Joule}$
Continuo:		
	Potencia media $\leq 6,4 \text{ W/m}^2$	
	Por 10 s $P \leq 64 \text{ W/m}^2$	.

Los siguientes gráficos representan la distancia nominal de daño ocular (NOHD) tabulada para distintos tipos de operación, longitud de onda y divergencia de haz (ver Figura 13).

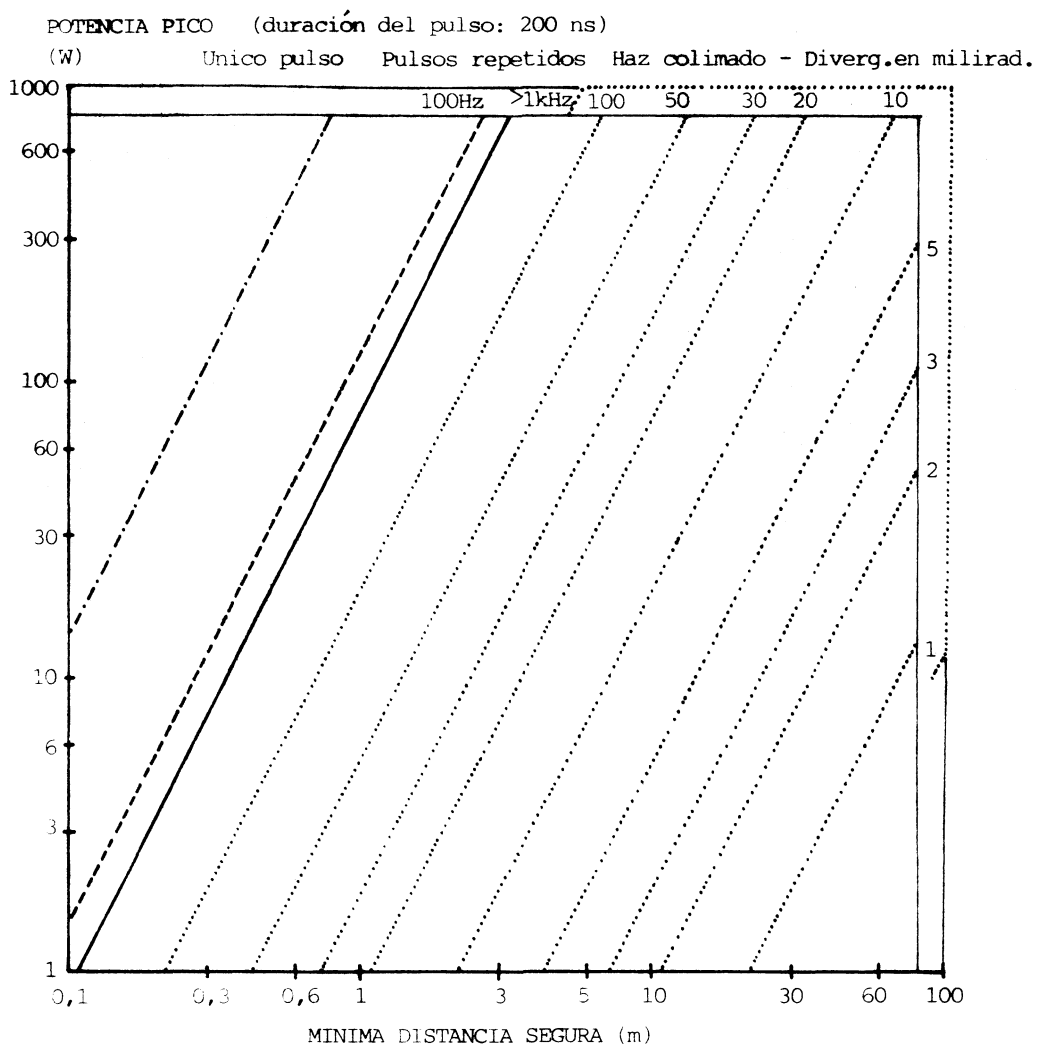


Figura 13. a)

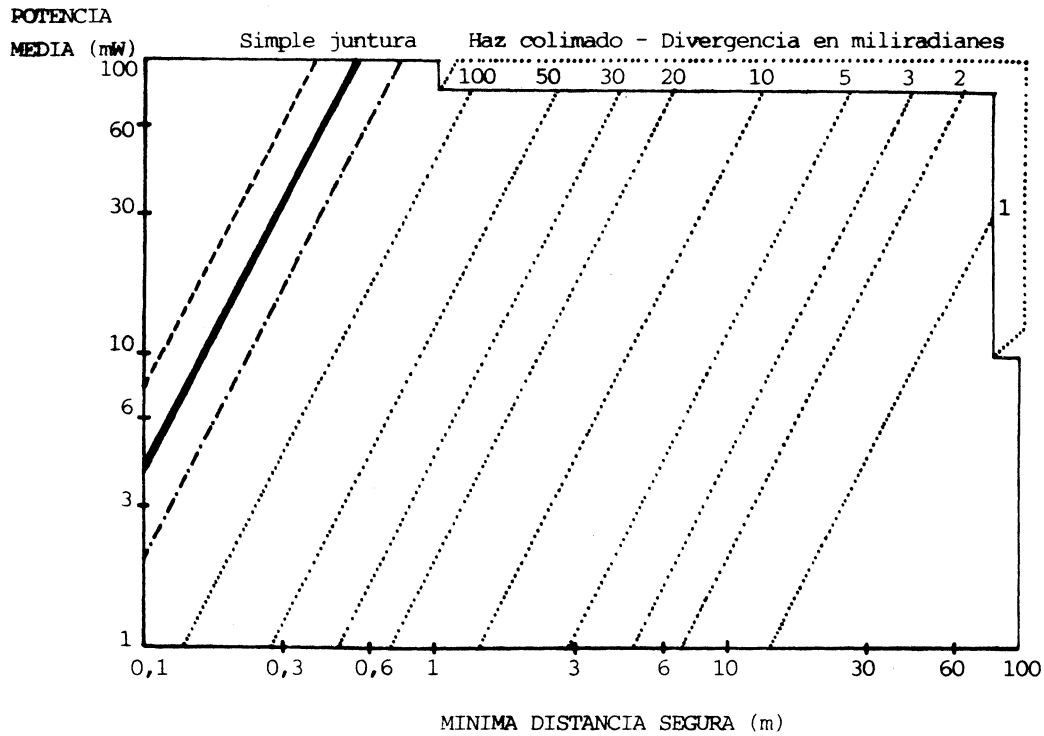


Figura 13. b)

Bibliografía

- Practical Introduction to Optical Communications, Forrest Mims III.
- Solid State Emitters. RCA SSE-100.
- RCA Electro-Optics Handbook EOH-11.
- EG&G Application Note D3000C-2 Silicon Photodiode Application Notes.
- D. Sliney y M. Wolbarsht: "Safety with Lasers and other Optical Sources" Plenum Press, New York (1980).

OTRAS APLICACIONES

D. C. Schinca

Introducción

Las aplicaciones efectivas y posibles del láser son lo suficientemente extensas y crecientes como para abarcar algunos volúmenes. No es, obviamente, la intención de este capítulo alcanzar semejante objetivo. Se han elegido entonces algunas de las aplicaciones más conspicuas, tanto de las que podríamos llamar tradicionales como de las de reciente instrumentación.

Se describen así aplicaciones del láser al procesamiento de materiales, al maquinado de piezas, a la industria electrónica, al rastreo de objetos y al sondeo ambiental. También se describen algunas aplicaciones en ensayos no-destructivos, como el control de calidad y la detección de fallas en diversas piezas. Finalmente, se resumen usos recientes del láser, tales como en impresoras, lectores de código de barras, y algunos usos en experiencias de laboratorio para el estudio de propiedades de semiconductores, materia condensada y plasmas.

Maquinado

Cuando se habla de los láseres aplicados a la industria se quiere significar, generalmente, la utilización de los mismos en el procesamiento de materiales y en el maquinado de piezas. En este rubro, los tres láseres más utilizados son, en orden de importancia, el de CO_2 , el Nd:YAG y el de excímeros. El láser de CO_2 fue el primero en ser utilizado en el área industrial por su potencia y versatilidad. Como uno de los capítulos de este libro trata el tema en forma detallada, no es necesario explayarnos aquí sobre lo mismo. Vamos, sí, a reseñar la utilización del Nd:YAG y del excímero, este último en franco ascenso en cuanto al número y variedad de aplicaciones.

El láser de Nd:YAG es, luego del láser de CO₂, el más extensivamente utilizado en el área industrial. A diferencia del segundo, éste sólo opera en forma pulsada (repetición hasta 500 Hz), lo que limita su velocidad en el corte de piezas. El esquema de un sistema así puede verse en la Figura 1.

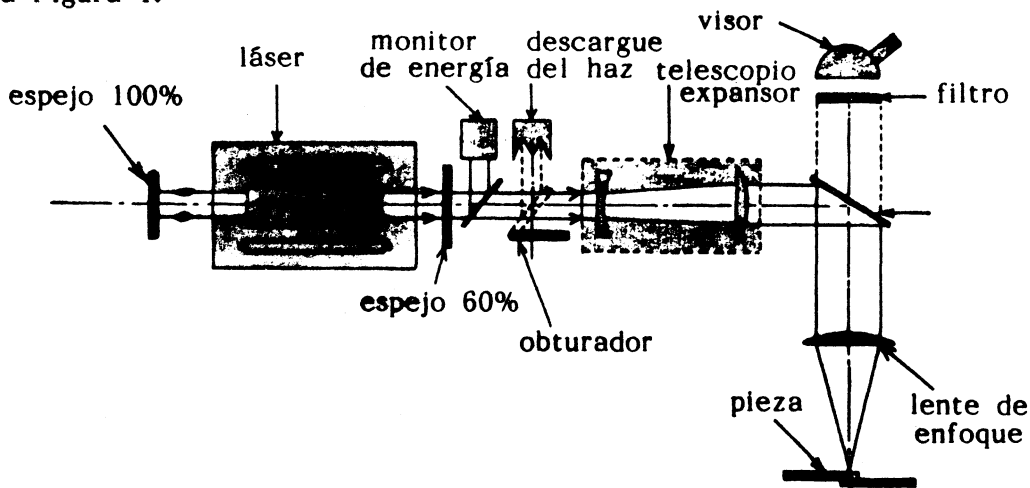


Figura 1. Esquema de un sistema láser de Nd:YAG tipo industrial.¹

Su operación pulsada le confiere excelentes características en la perforación de pequeños agujeros. El diámetro de éstos es controlado variando tanto la potencia del láser como la fineza del enfoque del haz sobre la muestra. El diámetro mínimo está determinado por la longitud de onda del láser, la óptica de enfoque y el material. Un Nd:YAG puede producir agujeros en láminas delgadas, de un diámetro 10 veces menor que los que puede hacer un CO₂, esto es, aproximadamente 5 μm .

En lo que a soldaduras concierne, el Nd:YAG utiliza los mismos mecanismos básicos que el CO₂. Aunque, como dijimos, el primero sólo opera en forma pulsada, posee ventajas específicas sobre otros sistemas. Debido a su corta longitud de onda (con respecto a la del CO₂), su energía es absorbida más rápidamente, lo que permite reducir notablemente la potencia umbral para una soldadura dada. Esto permite usar menor energía por soldadura y un mejor control del calor incidente, lo que resulta particularmente útil cuando se trabaja con juntas pequeñas de bordes delgados. Un ejemplo lo constituye la soldadura de un cabezal de níquel cobreado a una pieza cilíndrica de acero inoxidable. El cabezal puede tener elementos que no deben ser térmicamente perturbados como, por ejemplo, terminales tipo "feed-through" selladas a vidrio. Una soldadura de este tipo puede conseguirse fácilmente con un Nd:YAG de 250 W a una velocidad de 80 cm/min. Los materiales altamente reflectores resultan campo propicio para la soldadura con este láser: una potencia de 400 W puede soldar oro, plata, cobre o aluminio, permitiendo un muy buen control del

acabado de la misma. Además, es posible transmitir el haz por fibras ópticas (cosa que no se puede hacer con el láser de CO_2 , debido a su longitud de onda) y soldar varios puntos al mismo tiempo, lo que obviamente aumenta el número de piezas soldadas por unidad de tiempo y aumenta la eficiencia del proceso.¹

La industria joyera también se beneficia con el uso del Nd:YAG, por ejemplo en la soldadura de uniones de oro en diversas joyas. El alambre de oro usado para tal fin tiene un baño de aceite que se debe quitar para mejorar las características de la soldadura. Esto se logra gracias a la posibilidad que brinda este láser de controlar la forma del pulso: la primera parte del mismo se regula a baja potencia para vaporizar el aceite del alambre, mientras que la segunda parte se regula a alta potencia para entregar el suficiente calor como para fundir el oro y formar la soldadura; seguido de una entrega de energía en forma de rampa para crear un buen acabado final. En general, la mayoría de los metales que pueden soldarse por cualquier otro método, también pueden hacerlo por láser.

Con respecto a los láseres de excímeros, están encontrando un uso cada vez mayor en el procesamiento de materiales, debido a la particular forma en que la intensa radiación UV emitida por los mismos interactúa con la materia. Estos láseres producen pulsos luminosos de corta duración (10 ns), alta potencia pico (10 MW) y alta potencia promedio (100 W a 600 Hz de repetición), hecho éste que lo ha propulsado del área de investigación básica al de producción. El proceso de interacción de este láser se denomina ablación (o fotoablación), y desde su descubrimiento en 1982, ha causado una aceleración en la tecnología del mismo. Los polímeros (cadenas moleculares con 1.000 a 100.000 átomos de C, H, O y N) y muchos sólidos inorgánicos absorben fuertemente radiación en el UV lejano. Cuando tales materiales se irradian con un excímero, la energía se absorbe en una capa cuyo espesor está entre 0,1 y 0,5 μm . Cuando los fotones de la radiación UV de alta intensidad del excímero son absorbidos por estas moléculas, la energía de aquéllos rompe las ligaduras intermoleculares, dejando a los fragmentos con la energía suficiente como para ser expelidos violentamente por la brusca expansión creada por la excitación. Estos fragmentos, expulsados a velocidades supersónicas, van acompañados por una pluma luminiscente fácilmente visible en la oscuridad. A este proceso de rompimiento se lo llama ablación. Debido a que no es un mecanismo de interacción de tipo térmico, no posee efectos secundarios como fusión, carbonización o vaporización del material circundante a la zona de interacción, lo que resulta en un altísimo grado de precisión en el procesamiento de materiales y una excelente definición de bordes. Un ejemplo de esto puede verse en la Figura 2.

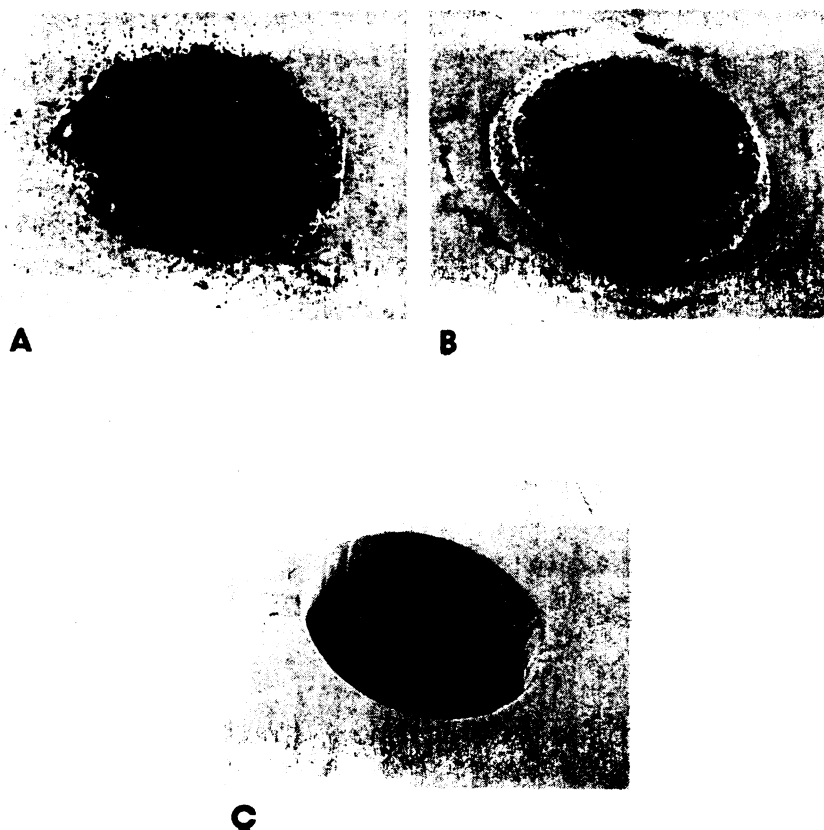


Figura 2. Orificios en poliamida, por: a) Nd:YAG; b) CO₂; c) excímero.²

En la figura se observan tres agujeros de 300 μm de diámetro en una hoja de poliamida de 75 μm de espesor usando los tres láseres mencionados. Para el caso del de excímeros, se ve que el borde está perfectamente definido y no hay deterioros del material circundante, como en los otros dos casos. Características similares se ejemplifican en la Figura 3, que muestra fotografías tomadas con microscopía electrónica, de muescas y agujeros hechos por un láser de excímeros en un cabello humano de 100 μm de espesor.

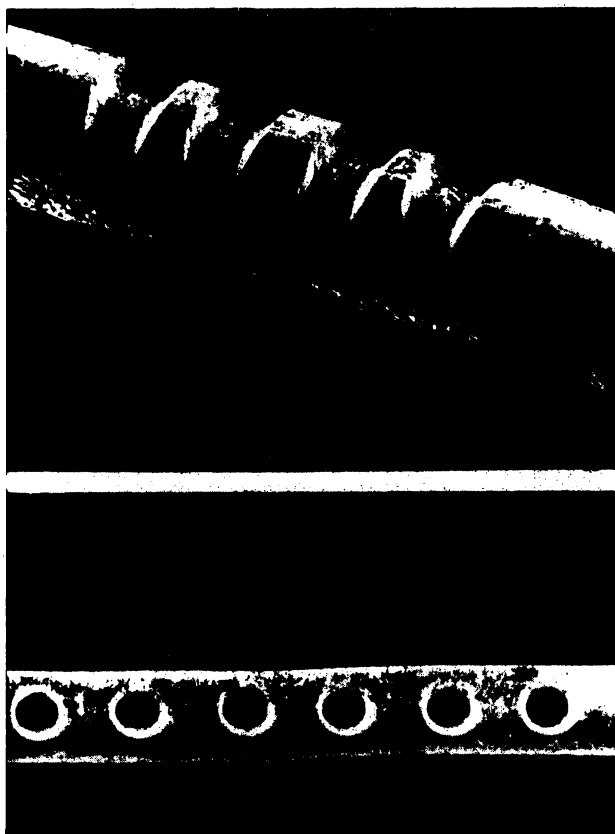


Figura 3. Muescas hechas con excímero en un cabello humano.²

En el caso convencional del láser de CO_2 o Nd:YAG, el haz es generalmente enfocado en un pequeño punto y se lo mueve respecto a la pieza a maquinar. En el caso del excímero, una máscara metálica que reproduce el patrón a procesar es iluminada por el láser y luego enfocada sobre la pieza. Como el patrón de la máscara puede ser cualquiera, es posible procesar un diseño complicado en forma instantánea, tal como se puede ver esquemáticamente en la Figura 4.

En el caso de patrones repetitivos, esta técnica unida a la capacidad de procesamiento en paralelo, permiten aumentar notablemente el número de piezas procesadas por unidad de tiempo. Otra obvia aplicación es en el procesamiento de circuitos impresos, donde estos láseres entran en acción en la etapa de remoción de la resina en áreas estrictamente limitadas.

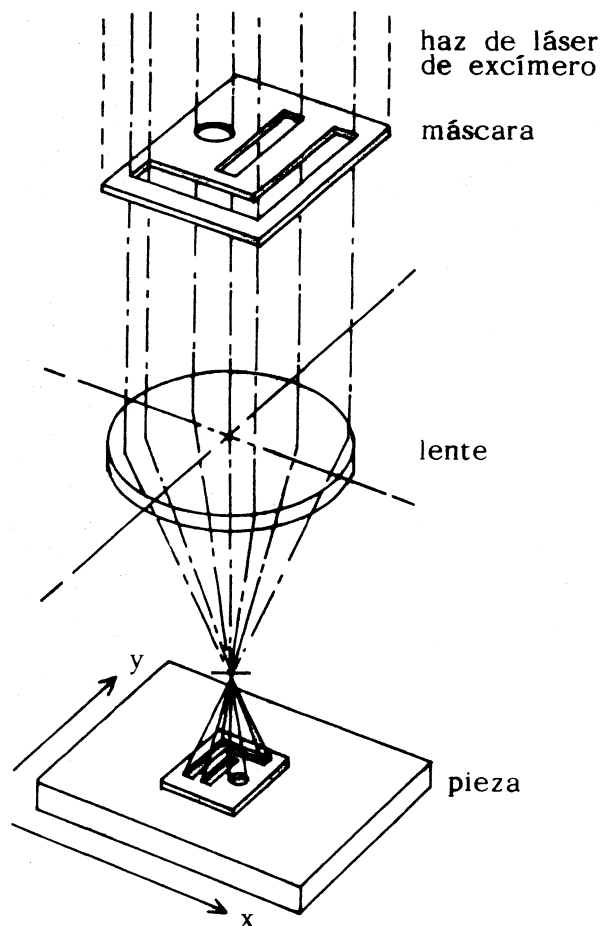


Figura 4. Arreglo típico usado en el procesamiento con láser de excímero.²

A pesar de que el uso de los láseres de excímeros en forma industrial está recién en sus comienzos, son ya especialmente recomendados cuando hay que procesar materiales orgánicos con alto grado de precisión y poco daño de las áreas circundantes.

Sondeo ambiental y detección de objetos

Como toda radiación electromagnética, el láser sufre reflexión y absorción al llegar a la superficie de un dado material. Este hecho permite utilizarlo en el monitoreo de objetos lejanos de variado tamaño, que va desde un avión hasta un conjunto de moléculas. La utilización de radiación electromagnética para estos fines no es nueva, ya que el uso de ondas de radio y microondas para el seguimiento de objetos en movimiento y determinación de su posición (RADAR) data de casi medio siglo. Sin embargo, las características del láser lo hacen apto para estos fines, particularmente en determinados casos de baja relación señal/ruido. El sistema utilizado recibe el nombre de LIDAR (por "Light Detection and Ranging") y su principio es muy similar al del radar, con las diferencias propias de tener que emitir y detectar un haz láser. Con el lidar se pueden hacer básicamente dos clases de monitoreo: de objetos macroscópicos ("tracking" o seguimiento) y de objetos microscópicos, tales como partículas, moléculas, etc. ("remote sensing" o sondeo a distancia). Vamos pues a describir brevemente cada una de estas aplicaciones.

El sistema de seguimiento de objetos macroscópicos con láser opera, como ya dijimos, sobre los mismos principios básicos que el radar. Generalmente, estos sistemas lidar emiten un pulso de radiación láser y miden el tiempo que le toma llegar al objeto en cuestión y reflejarse en el mismo. Algunos sistemas operan en modo continuo, en cuyo caso la radiación debe ser modulada. El lidar tiene ventajas y desventajas sobre el radar. Entre las primeras se encuentran la direccionalidad y corta longitud de onda, que permiten un mejor manejo de la dirección hacia el blanco y una menor probabilidad de interferencia con otra clase de radiación electromagnética, siempre respecto a las microondas. También permite el uso de retrorreflectores (tipo ojo de gato) de tamaño reducido, fáciles de incorporar al objeto a seguir y que proveen un haz de retorno bien colimado, lo que aumenta la relación señal/ruido, reforzando así la eficiencia de detección. Entre las desventajas, se encuentra la imposibilidad de operar el sistema con niebla, lluvia o nieve debido a la atenuación y dispersión de las cortas longitudes de onda bajo dichas condiciones. También, cuando se debe monitorear un objeto iluminado por el sol, el "ruido" de fondo aumenta considerablemente y debe ser filtrado, lo que trae aparejado problemas de resolución y baja relación señal/ruido. Un esquema típico de estos sistemas lidar puede verse en la Figura 5.

El láser más frecuentemente usado en el lidar es el Nd:YAG, generalmente en modo Q-switch y a veces doblado en frecuencia (es decir, que la longitud de onda de salida es la mitad de la longitud de onda de entrada). También se utiliza el láser de CO_2 y, con menos frecuencia, el de argón continuo.

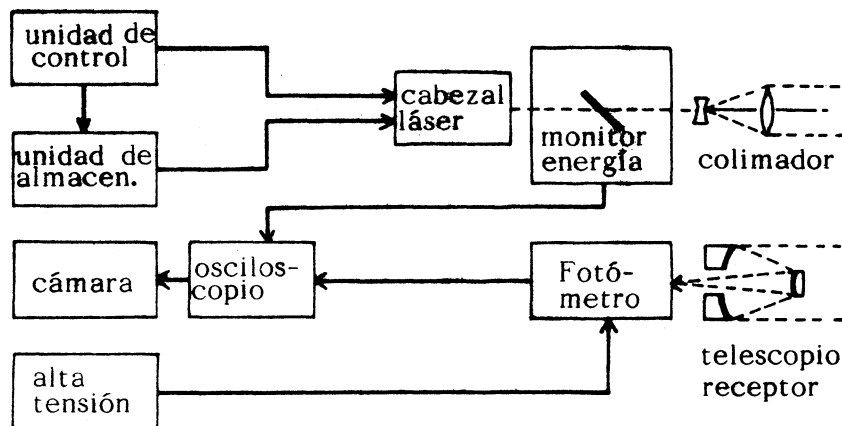


Figura 5. Diagrama en bloques de un sistema lidar pulsado.³

En la mayoría de los sistemas, el haz láser no se transmite directamente al objeto, sino que primero se expande con un sistema óptico como el que muestra la Figura 6.

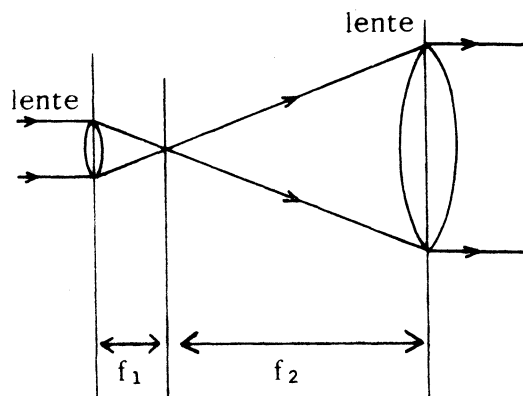


Figura 6. Esquema de un sistema óptico expansor.

La primera lente enfoca el haz láser divergente en un cierto punto y una segunda lente, con foco en ese punto, recibe el haz y lo transmite en forma paralela (colimada). Esto se hace para reducir aún más la pequeña divergencia natural que tiene el láser, la que puede molestar en ciertas aplicaciones.

Existen distintas formas de mantener el haz que sale del colimador

dirigido hacia el objeto. La más simple es la guía óptica, en la que el operador mantiene visualmente al objeto dentro de la retícula montada en el sistema telescópico de transmisión de la Figura 6. Si la trayectoria del objeto es conocida, el seguimiento puede hacerse automáticamente vía sistema computarizado.

Para la detección del haz reflejado, el sistema de recepción debe estar dirigido al objeto. Usualmente consiste en un telescopio que colecta la mayor cantidad posible de luz reflejada, la que es filtrada tanto espacial como temporalmente para disminuir el "ruido" de fondo y pasa luego al detector que convierte la señal óptica en eléctrica. Esta es procesada convenientemente en forma electrónica u observada directamente en un osciloscopio o cámara fotográfica. El detector puede ser un fotodiodo de avalancha o bien un fotomultiplicador, donde los fotones incidentes generan electrones que son acelerados por un cierto potencial, produciendo así una corriente eléctrica que resulta ser proporcional a la intensidad de luz incidente. En la elección del detector correspondiente a determinada aplicación, es conveniente conocer la llamada eficiencia cuántica del detector, que mide la cantidad de electrones emitidos por fotón incidente. Las eficiencias cuánticas de los electrones más comúnmente usados se muestran en la Figura 7.

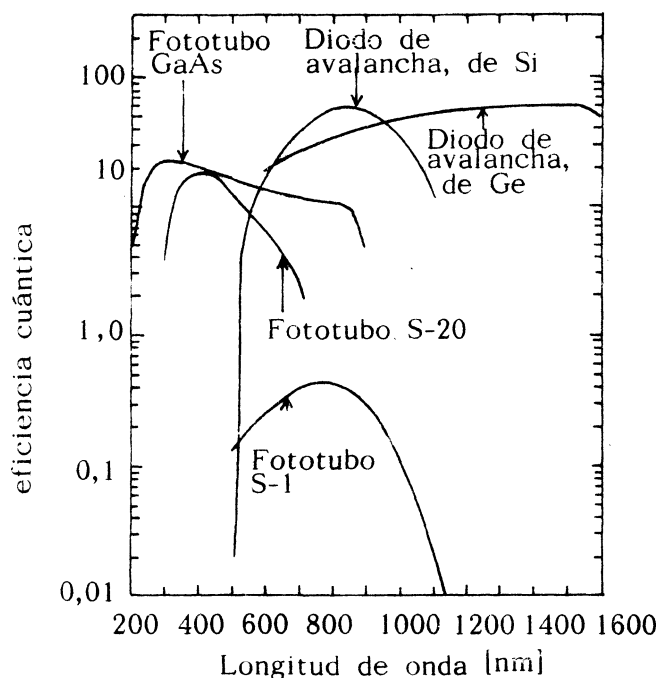


Figura 7. Diversas eficiencias cuánticas en función de la longitud de onda.

Cuando la señal de retorno es muy baja (unos pocos fotones), resulta importante usar detectores de gran eficiencia cuántica.

Evidentemente, la mayor importancia y desarrollo de estos sistemas de seguimiento se encuentran en el área militar y aeroespacial. Un ejemplo es el lidar que posee la NASA para el control del encuentro de naves espaciales en órbita. El sistema trabaja con arreglos de 100 láseres de semiconductor en ambos vehículos (llamados "vehículo blanco" y "vehículo rastreador"). El rastreador puede localizar el blanco cuando están en órbita casi coplanares y a no más de 120 Km de separación. Una vez localizado, el rastreador se dirige hacia él y es maniobrado apropiadamente hasta que se produce el acoplamiento. El sistema en el rastreador usa dos telescopios coaxiales, de distinto f /número, que permiten dos rangos de seguimiento: de 3 a 120 Km y de 0 a 3 Km, con una resolución de 10 m y 1 cm respectivamente.³

Otro ejemplo de sistema de seguimiento es el fabricado por la firma Sylvania para la serie de cohetes Saturno V, también de la NASA. Aquí el láser usado es de tipo continuo (el primero fue un He-Ne) y permite seguir la trayectoria del cohete desde su lanzamiento hasta una altura de 10 Km, con sólo 50 mW de potencia láser. Al nivel de lanzamiento, el sistema puede determinar la posición del cohete con una precisión del orden del cm, aventajando al sistema convencional de radar de microondas, cuya operatividad se ve seriamente degradada en esas circunstancias debido al ruido de fondo a nivel de la superficie. Actualmente una firma canadiense ha desarrollado un batímetro lidar de barrido, que puede medir la profundidad del agua marina con una capacidad de penetración de $40 \text{ m} \pm 0,3 \text{ m}$ en aguas claras. El sistema, que se lleva a bordo de un avión, emite simultáneamente un corto pulso láser infrarrojo y otro en la región verde del espectro, dirigidos ambos hacia la superficie del agua. El primero es reflejado por la misma, mientras que el segundo penetra en el agua y es retrodispersado desde el lecho marino. Ambos pulsos de retorno son detectados en el avión por un sistema receptor que utiliza el tiempo transcurrido entre ambos pulsos para determinar la profundidad. El láser utilizado es un Nd:YAG en modo Q-switch y con un cristal doblador de frecuencia (transforma la longitud de onda en su mitad), que emite pulsos en 1.064 nm y 532 nm simultáneamente. La frecuencia de repetición de los pulsos es de 20 Hz, lo que determina el ritmo de adquisición de datos del sistema.⁴ Un resultado típico de un barrido cercano a la costa se ve en la Figura 8.

El pulso marcador de la superficie del agua es el primero de cada trazo; el segundo pico es la retrodispersión del volumen de agua, y el tercer pico es la señal de retorno desde el fondo. Este sistema también promete ser valioso en la detección de objetos sumergidos, perfiles de terrenos, medición de elevaciones, etc.

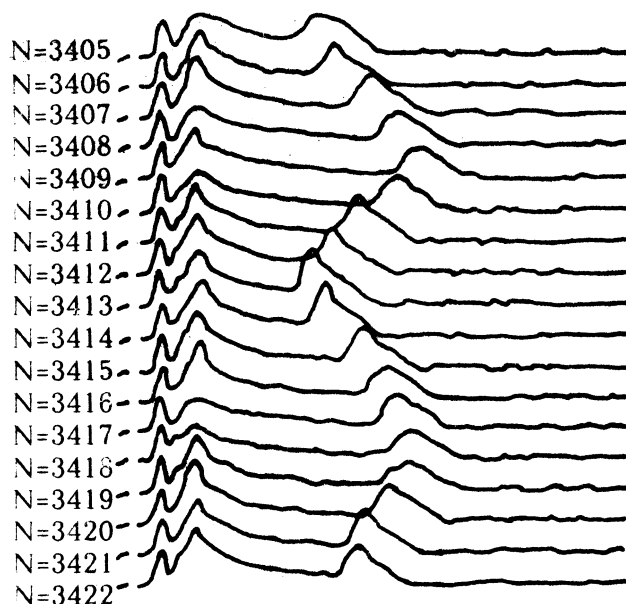


Figura 8. Muestra de un barrido con batímetro láser.

Otra faceta de la aplicación de los láseres en el rastreo de objetos es la llamada detección a distancia o sondeo remoto (monitoreo ambiental). En estos casos, se envía un pulso láser hacia determinada región de la atmósfera y se estudia el pulso de retorno (de la misma longitud de onda que el emitido), que puede provenir de objetos tales como nubes, frentes de tormenta, gotas de lluvia, humo, etc. La luz dispersada por moléculas en la atmósfera puede ser detectada a distancias de varios kilómetros con dispositivos de tamaño moderado. Ejemplos típicos de pulsos láser retrodispersados pueden verse en la Figura 9, donde en el eje vertical se representa la corriente de salida de un fotodetector lineal y en el eje horizontal se barre en tiempo (lo que significa que, por calibración, se representa la distancia). La curva superior muestra la señal que se obtendría de una atmósfera "puramente molecular" (es decir, que sólo contiene gases puros como N_2 , O_2 , A y CO_2); en cambio, la curva inferior representa la señal de una atmósfera "real", con su siempre presente concentración de aerosoles, que producen una tenue niebla visible sobre el horizonte. La parte recta de la curva B-C indica la existencia de una capa de niebla a una altura entre H_B y H_C .³

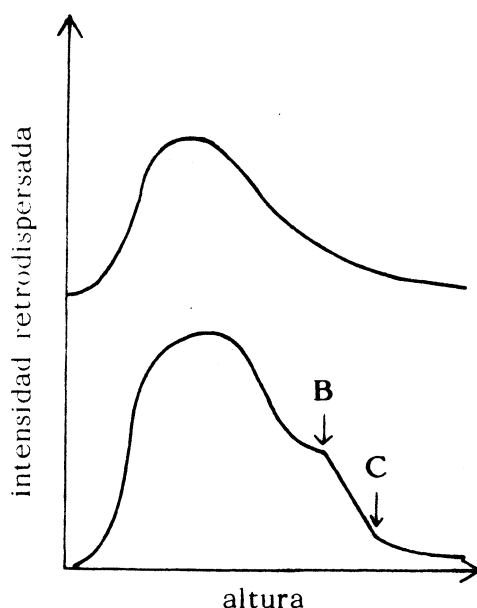


Figura 9. Trazos de lidar en monitoreo ambiental.

Como el tamaño de las partículas de aerosoles (que oscila entre 1 nm y 1 μm) es función de la humedad relativa, es posible monitorear la altura de las capas húmedas a través del estudio del diámetro de las partículas de aerosoles. También es posible utilizar técnicas de lidar combinadas con otras más tradicionales, tales como perfiles de temperatura o concentraciones de especies moleculares hechas en globo, para estudiar la dinámica de corrientes atmosféricas o de difusión, de particular importancia en el control de la polución ambiental en áreas urbanas.

Otro aspecto del monitoreo ambiental es el sondeo remoto de determinadas especies moleculares. Aquí, se emite un pulso láser de alta potencia hacia la atmósfera, el que es absorbido por determinadas moléculas. Estas a su vez reemiten radiación (fluorescencia), en otra longitud de onda, la que es recibida por un detector apropiado. La medida de la intensidad de dicha fluorescencia permite conocer características tales como tipo de especie molecular presente, concentración, etc. Para que la absorción por parte de la especie molecular a sondear tenga lugar, es necesario ajustar la longitud de onda del láser incidente, para que coincida con una determinada transición de la molécula. Para estas aplicaciones es conveniente usar láseres de colorante, que se pueden sintonizar (variar su longitud de onda sin pérdida apreciable de potencia), sobre grandes rangos del espectro visible. El rango de sintonía de los láseres de colorantes hacia el infrarrojo puede extenderse, utilizando láseres de estado sólido sintonizables entre 800 y 1500 nm, de reciente fabricación.

Con estos láseres de ancho de banda angosto, es posible medir perfiles de líneas de absorción, los que a su vez proveen información sobre presión atmosférica y temperatura de la región sondeada.

Este sistema de monitoreo es usado también para el control de contaminación ambiental en áreas fabriles, detección de concentraciones peligrosas de gases potencialmente dañinos en áreas mineras o monitoreo a tiempo real de parámetros de interés en una línea de producción. Un ejemplo de aplicación es el sistema desarrollado en Italia para investigar la contaminación de derivados de hidrocarburos en el mar. En este caso, un pulso láser enviado desde un avión excita las moléculas de aceite depositado en el mar, y luego se detecta su fluorescencia y su tiempo de decaimiento. Este fluorosensor emplea un láser de Nd:YAG en modo Q-switch con triplicado de frecuencia (355 nm) para excitación del aceite, un telescopio de barrido para coleccionar la fluorescencia y una cámara de barrido ultrarrápido conectada a una computadora para la reducción de los datos. En las Figuras 10 y 11 pueden verse las características espectrales de emisión de algunos derivados y los correspondientes tiempos de decaimiento, respectivamente.

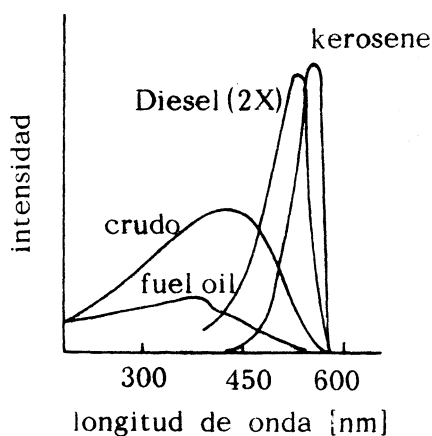


Figura 10. Fluorescencia de diversos derivados de petróleo.²

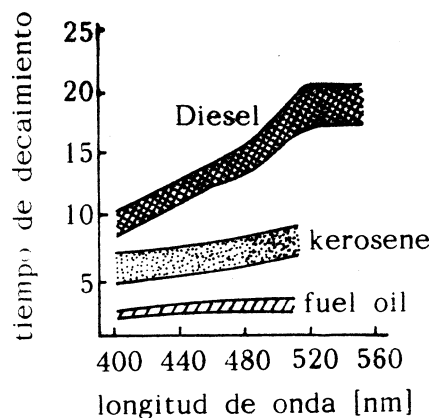


Figura 11. Tiempos de decaimiento fluorescente de algunos derivados de petróleo.

El análisis de estos dos conjuntos de datos suministra información sobre la compleja composición de las manchas de derivados de petróleo en el mar.

Próximamente, Japón pondrá en funcionamiento un equipo lidar de absorción diferencial, como parte de un programa internacional para el estudio de la influencia de gases fluorocarbonados en la capa de ozono de la atmósfera. El sistema utiliza un láser de excímero que emite en tres líneas de absorción del ozono en el UV. Como la absorción disminuye con la longitud de onda, el monitoreo de la señal de retorno para esas tres líneas permite un cálculo preciso del perfil del ozono con la altitud.⁵

Interferometría holográfica. Ensayos no destructivos.

Dentro del marco de ensayo de materiales y control de calidad de productos, las técnicas ópticas juegan un rol importante por su carácter no destructivo y no invasivo. Dentro de éstas, la más conspicua es la interferometría holográfica. En esta técnica, se registran sobre una placa fotográfica un holograma de determinado objeto. Si sobre la misma placa se realiza una segunda exposición del objeto, al que ahora se lo ha desplazado o deformado de alguna manera, se generará un patrón de franjas de interferencia que están relacionadas con el desplazamiento, a través de la ecuación:

$$\vec{k} \cdot \vec{r} = m \lambda \quad ,$$

donde $\vec{k}$ es el vector propagación; $\vec{r}$ el vector desplazamiento; m un número natural; y λ la longitud de onda.

Por ser la relación anterior de naturaleza vectorial, la interpretación exacta del patrón de franjas resulta complicado, aunque permite obtener muy buenos resultados cualitativos en primera instancia.



Figura 12. Interferograma de un neumático con "soplado".

El método tiene aplicaciones en diversas áreas. Un ejemplo típico lo constituye la detección de fallas ("soplado"), en las capas componentes de una cubierta de automóvil, como se ve en la Figura 12. Aquí se realiza una doble exposición, la primera con el neumático a presión atmosférica y la segunda a la presión de inflado normal. De existir una falla, ésta se revela por una alta densidad de franjas de interferencia concéntricas alrededor de la misma.³

El mismo sistema se utiliza para la detección de fallas en cilindros comerciales de gas a alta presión y aún en estructuras soldadas, tal como el rotor de helicóptero que se muestra en la Figura 13.

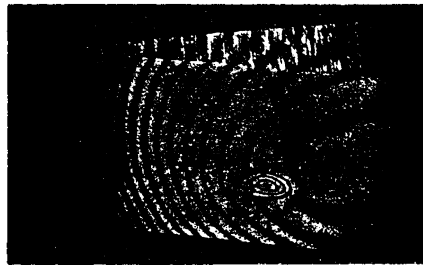


Figura 13. Detección de falla en rotor de helicóptero por interferencia holográfica.⁶

También es posible utilizar métodos interferométricos para el control de pérdidas en dispositivos herméticamente sellados. Un ejemplo puede verse en la Figura 14, donde se muestran cuatro baterías de marcapasos en etapa de control.

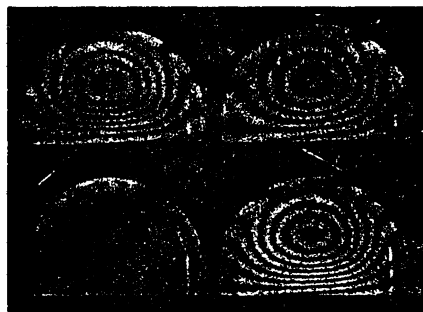


Figura 14. La ausencia de franjas indica una batería defectuosa.⁶

En este caso, las baterías se colocan en una cámara de vacío y se forma un holograma de referencia. Si se cambia la presión en la cámara, las baterías en buen estado se deformarán y producirán franjas de interferencia al tomar un segundo holograma sobre la misma placa. En cambio, la batería que tenga una fuga no sufrirá deformación y no presentará franjas. Así, es posible probar varios dispositivos a la vez y aumentar la eficiencia del proceso de control.

La industria electrónica se beneficia también con esta técnica. Es posible, como se ve en la Figura 15, utilizar un interferograma para poner de manifiesto las diversas tensiones que sufre una placa de circuito impreso, alimentada, a causa de gradientes de temperatura indebidos, en sus componentes.

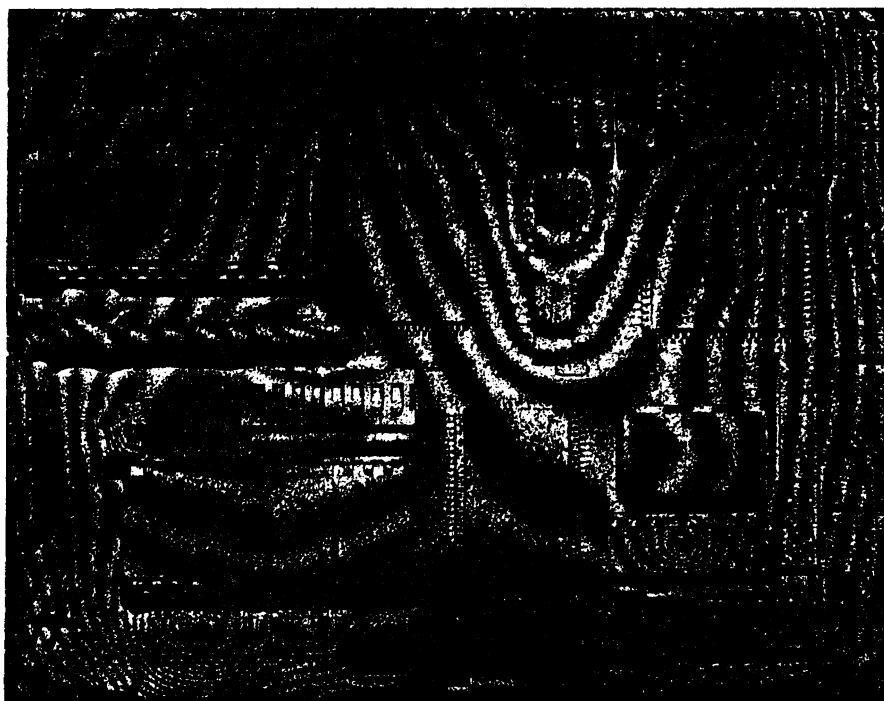


Figura 15. Interferograma mostrando las tensiones en placa de circuito impreso.⁶

Una de las ventajas del método de interferometría holográfica sobre los procedimientos más comunes (rayos X o ultrasonido) es que resulta suficiente tener acceso a la superficie de la muestra para realizar el correspondiente ensayo y visualizar los posibles defectos que yacen dentro del material. También posee la ventaja de poderse observar las franjas de interferencia a tiempo real, lo que lo hace versátil y relativamente com-

pacto como para permitir su uso en control de calidad en líneas de producción.

Registro y lectura con láser.

El láser ha encontrado también un lugar importante en el procesamiento de la información, tanto en la etapa de almacenamiento como en la de lectura de los datos compilados. Esto forma parte de un proceso general conocido como almacenamiento óptico de la información. El conjunto de datos a manipular es variable, siendo el más común el procesamiento de imágenes que proceden vía electrónica a partir de una amplia gama de fuentes, tales como cintas, discos compactos, discos rígidos, etc.

Uno de los últimos dispositivos en esta materia se conoce como impresora láser, cuyo esquema básico de funcionamiento puede verse en la Figura 16.

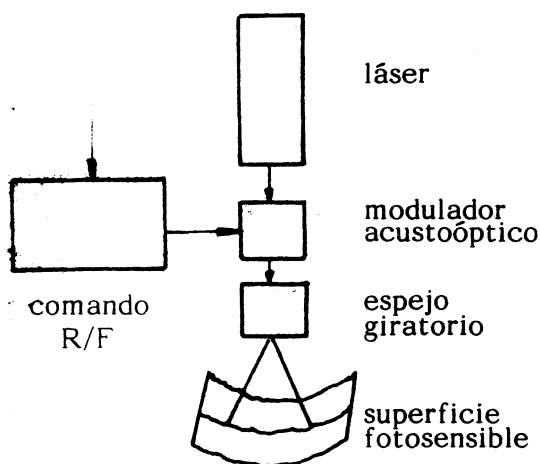


Figura 16. Esquema en bloques de una impresora láser.

La señal generada por una de las fuentes anteriormente mencionadas se usa para modular la intensidad de un láser (típicamente, un He-Ne de 2 a 5 mW), siendo los intervalos de encendido del mismo, a través de un modulador acustoóptico. Este haz así modulado contiene información que se correlaciona con el pulso electrónico de entrada y es desviado

por un espejo giratorio hacia una superficie fotosensible donde queda almacenada la información en forma de distribución de intensidad. Al revelar el film fotosensible, dicha distribución puede "leerse" con otro láser.

La información procesada puede corresponder a páginas impresas o imágenes fotográficas. En este último caso, resulta importante que el sistema sea capaz de generar diferentes tonos de grises sobre el papel o film. Estos tonos pueden lograrse por dos métodos. Cuando los grises se forman aplicando tóner de una densidad fija sobre el papel, los matices se logran variando el tamaño del punto creado por el láser: puntos grandes y poco distanciados producen los matices más oscuros, mientras que los puntos pequeños y más espaciados dan lugar a los tonos más claros. Cuando se usan materiales fotosensibles de tono continuo, los matices se logran variando la exposición de los "pixels" que conforman la imagen: a mayor exposición, más oscura será la imagen después del revelado. Este último método difiere de los sistemas digitales en los cuales el haz es interrumpido a determinados intervalos, escribiendo o no un pixel en la forma binaria habitual.⁷

Recientemente, una firma británica ha desarrollado un medio de registro basado en un film de poliéster de unos 25 μm de espesor, similar al film usado para cintas magnéticas. Está cubierto por un polímero de colorante capaz de absorber el infrarrojo y todo está envuelto por un recubrimiento protector que permite manipularlo como si fuese papel.

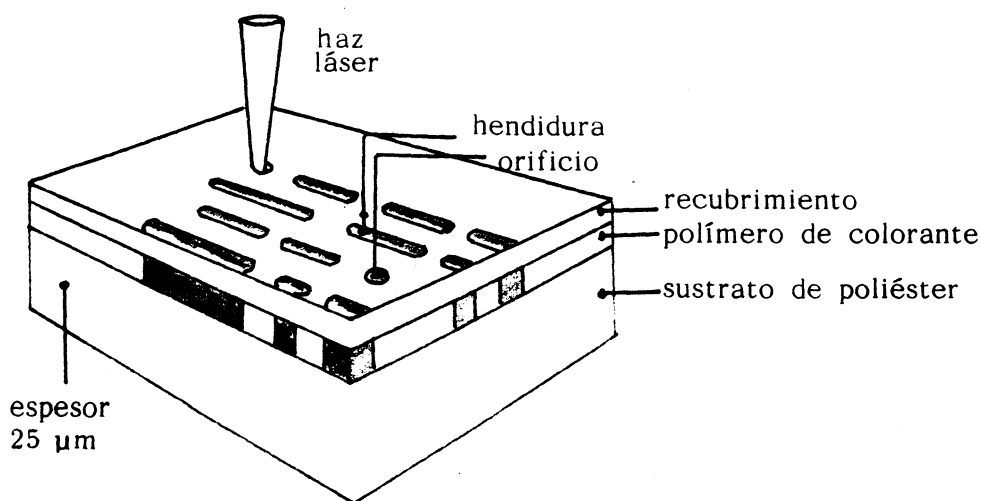


Figura 17. Esquema de un papel digital.⁸

Para escribir sobre el film, se utiliza un láser de estado sólido que emite en 780 nm u 830 nm, enfocado en diámetros de 1 μ m sobre el polímero, creando así diminutas hendiduras que tienen una menor reflexión que las áreas circundantes, tal como puede verse en la Figura 17.

Para la etapa de lectura se usa el mismo láser a menor potencia. Se puede conseguir una frecuencia de datos de 10 MHz y se trabaja actualmente para aumentar la densidad de almacenamiento de datos.

Otra aplicación del láser en el procesamiento de datos es el conocido uso del mismo como lector de código de barras. El conjunto de barras de diferentes anchos y espaciados, que aparecen en los productos que se adquieren en el mercado, llevan codificada una variada información, tal como país de origen, tipo de producto, características de peso neto o volumen, precio, etc. El dispositivo lector del código posee un láser de baja potencia (típicamente un He-Ne de 2 mW), cuyo haz pasa a través de un disco rotante con lentes holográficas que permiten deflectar el haz en diferentes ángulos y así "envolver" casi toda la periferia del envase en cuestión. Esto hace que el haz llegue efectivamente al código de barras, que devuelve luz reflejada conteniendo información (de carácter binario), sobre la distribución de las mismas. La reflexión así modulada es recibida por un detector que convierte la señal óptica en eléctrica y la transmite a una computadora que se encarga de llevar el control del movimientos de los diversos productos (total de venta por cliente, "stock", aviso de necesidad de reposición, ventas diarias totales, etc.).

El láser en el laboratorio.

Hasta aquí no hemos agotado ni por poco las aplicaciones efectivas o posibles del láser en temas relacionados con la ingeniería. Existe aún un gran conjunto de aplicaciones en el que el láser se usa, no ya como parte integrante de un determinado dispositivo, sino como herramienta fundamental en experiencias básicas de laboratorio. En general, dichas experiencias estudian propiedades de materiales a través del análisis de la respuesta de la materia al interactuar con radiación láser. Esta respuesta puede tomar diversas formas: puede ser la luz reemitida por el material (fluorescencia), de la que se puede analizar su distribución en longitud de onda (espectroscopía), razón de polarización, etc.; puede ser, en el caso de sólidos y materia condensada, la onda acústica generada por una rápida liberación de calor debida a la absorción de un pulso láser (efecto fotoacústico); o bien, en descargas gaseosas a baja presión, una variación en la tensión de la misma al hacer incidir sobre el gas radiación láser de determinada longitud de onda (efecto optogalvánico).^{9,10}

Con estas técnicas es posible estudiar propiedades tales como efectos de combustión de gases, termometría en jets supersónicos, medición instantánea de campos de temperatura a presión atmosférica, diagnóstico

de llamas y plasmas, estudios de fenómenos de superficies y films delgados, análisis de trazas de elementos o impurezas en gases, estudios de propiedades térmicas y ópticas en semiconductores y cristales de interés tecnológico, estudios de fenómenos electroquímicos en interfaces líquido-sólido, determinación de tamaño de partículas, etc.

Referencias

- 1 - Photonics Spectra, Noviembre 1987.
- 2 - Laser Focus - Electro Optics, Junio 1987.
- 3 - "Laser Applications", Vol II. Ed.: Monte Ross. Academic Press (1974).
- 4 - Laser Focus - Electro Optics, Febrero 1987.
- 5 - Laser Focus - Electro Optics, Mayo 1987.
- 6 - Photonics Spectra, Diciembre 1987.
- 7 - Photonics Spectra, Abril 1987.
- 8 - Photonics Spectra, Abril 1988.
- 9 - "Laser Spectroscopy VIII". Eds.: W. Persson y S. Svanberg, Springer-Verlag (1987).
- 10- "Photoacustical and Photothermal Phenomena". Eds.: P. Hess y J. Pelzl , Springer-Verlag (1987).

Impreso en
COPIAS 55 de DANIEL CABRERA
calle 55 N° 609, La Plata,
en el mes de agosto de 1989