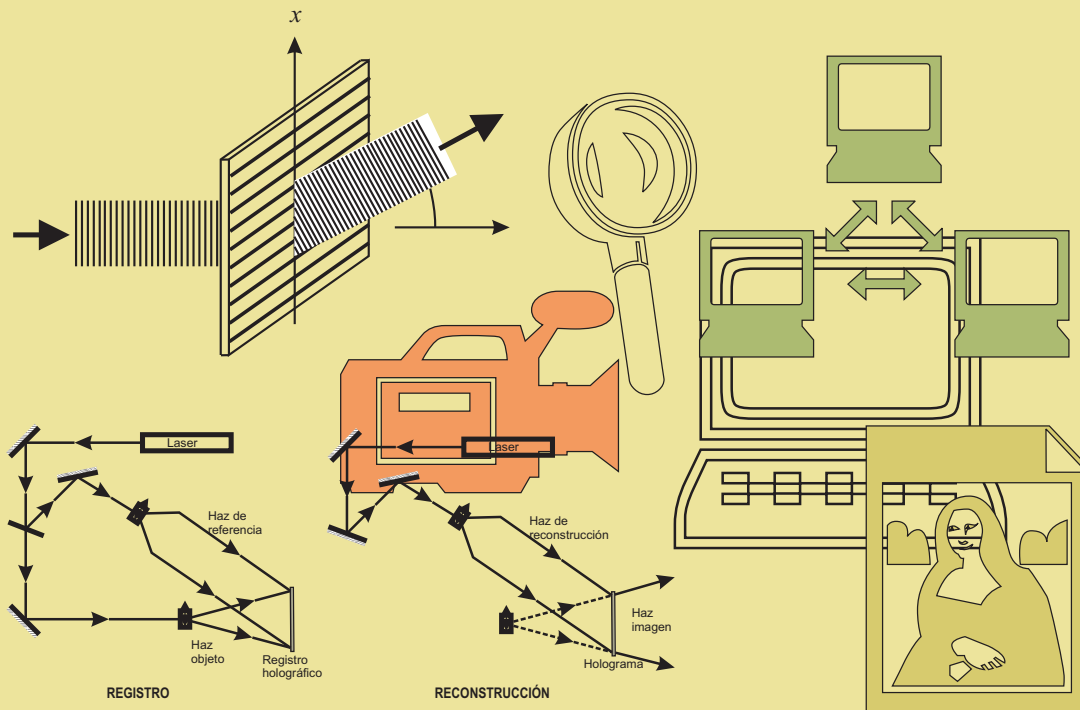


Apuntes de apoyo a la asignatura

FOTÓNICA APLICADA A LA COMPUTACIÓN Y TRANSMISIÓN DE INFORMACIÓN

Dept. de Física Aplicada 1
E.T.S. de Ingeniería Informática
UNIVERSIDAD DE SEVILLA



Francisco L. Mesa Ledesma

Copia Uso Académico

Prefacio

La presente colección de notas sobre *Fotónica Aplicada a la Computación y Transmisión de Información* (FACTI) pretende ser una ayuda al estudiante de esta asignatura cuatrimestral que se imparte en la Escuela Técnica Superior de Ingeniería Informática de la Universidad de Sevilla. Es importante resaltar que estas notas no pueden ni deben sustituir a otros textos sobre la materia mucho más elaborados.

El objetivo principal de la materia presentada es dotar al alumno de algunos de los fundamentos físicos elementales en los que se basa el funcionamiento de los dispositivos y sistemas ópticos, optoelectrónicos y fotónicos usados en Informática. Dado que parte de la tecnología actual de los computadores se basa en la Fotónica y ésta consiste básicamente en el control del flujo de los fotones, es evidente la necesidad de estudiar en primer lugar el comportamiento general de la luz y los fotones. Este estudio se realizará mediante dos temas introductorios dedicados uno a las ondas en general y otro a las ondas electromagnéticas en particular. Posteriormente, se estudiarán las bases de la Óptica paraxial y de algunos de los componentes ópticos más elementales. A continuación se abordará el estudio de la Óptica Ondulatoria y de la Óptica de Fourier, que constituyen la base de muchas de las nuevas aplicaciones tecnológicas de la Óptica. Posteriormente, y dada la importancia actual de los medios ópticos como sistemas de transmisión de información, se dedicará una última parte al estudio de la fibra óptica y de las fuentes de luz más usadas en la tecnología informática: láser y led.

FRANCISCO L. MESA LEDESMA
Sevilla, septiembre de 2000

Copia Uso Académico

Índice general

1. Nociones generales de Ondas	1
1.1. Introducción	1
1.2. Ecuación de ondas	2
1.3. Ondas armónicas	4
1.4. Energía e Intensidad de la onda	6
1.5. Interferencia de Ondas	7
1.5.1. Superposición de dos ondas armónicas	7
1.5.2. Focos incoherentes	9
1.5.3. Focos coherentes	9
1.6. Ondas estacionarias	12
1.7. Difracción	15
1.8. Grupo de Ondas	18
1.9. Ondas en tres dimensiones	23
1.9.1. Ondas Planas	23
1.9.2. Ecuación de onda tridimensional	24
1.9.3. Ondas esféricas y cilíndricas	25
1.9.4. Aproximación de Fresnel de la onda esférica: Onda Parabólica	26
1.10. Problemas propuestos	27
2. Ondas Electromagnéticas	31
2.1. Introducción	31
2.2. Ecuaciones de Maxwell	32
2.2.1. Introducción	32
2.2.2. Antecedentes	32
2.2.3. Aportaciones de Maxwell	33
2.2.4. Ecuaciones de Maxwell en forma diferencial	37
2.2.5. Ecuaciones de Maxwell en la materia	39
2.3. Ecuación de Ondas para las Ondas Electromagnéticas	40

2.3.1. Ondas electromagnéticas armónicas	41
2.4. Polarización	42
2.5. Intensidad de la onda electromagnética	44
2.6. Fuentes de las Ondas Electromagnéticas	46
2.7. Espectro electromagnético	48
2.8. Problemas propuestos	49
3. Óptica Geométrica	51
3.1. Introducción	51
3.2. Consideraciones previas	51
3.2.1. Rayo luminoso	52
3.2.2. Índice de refracción	52
3.2.3. Camino óptico	53
3.2.4. Principio de Fermat	53
3.3. Leyes de la Óptica Geométrica	54
3.4. Componentes ópticos simples	56
3.4.1. Espejos	57
3.4.2. Interfases planas	59
3.4.3. Interfases esféricas	59
3.4.4. Lentes esféricas	61
3.4.5. Guías de luz	63
3.5. Formalismo matricial de la Óptica paraxial	64
3.6. Problemas propuestos	68
4. Óptica Ondulatoria	69
4.1. Introducción	69
4.2. Relación entre rayos y frente de onda	69
4.3. Ondas paraxiales	72
4.4. Transmisión a través de elementos ópticos	73
4.4.1. Transmisión a través de una lámina transparente	73
4.4.2. Lámina transparente de espesor variable	74
4.4.3. Red de difracción	77
4.5. Problemas propuestos	79
5. Óptica de Fourier	81
5.1. Introducción	81
5.2. Aspectos preliminares	82
5.2.1. Correspondencia entre la función armónica espacial y la onda plana	83

5.2.2. Análisis espectral de frecuencias espaciales	84
5.2.3. Modulación en amplitud	85
5.2.4. Modulación en frecuencia	87
5.3. Propagación de la luz en el espacio libre	89
5.3.1. Función transferencia del espacio libre	89
5.4. Transformada de Fourier óptica	94
5.4.1. Aproximación de Fresnel	94
5.4.2. Aproximación de Fraunhofer: Transformada de Fourier en el campo lejano	95
5.4.3. Transformada de Fourier usando lentes	96
5.5. Problemas propuestos	99
6. Aplicaciones de la óptica de Fourier	101
6.1. Introducción	101
6.2. Difracción de la luz	101
6.2.1. Difracción de Fraunhofer	102
6.3. Formación de imágenes	104
6.3.1. Lente simple como sistema de formación de imágenes	105
6.4. Procesado de imágenes	106
6.4.1. Procesado óptico coherente	107
6.4.2. Procesado digital	111
6.5. Holografía	111
6.5.1. El código holográfico	113
6.5.2. Holografía fuera de eje	115
6.5.3. Registro holográfico	117
6.5.4. Aplicaciones de la holografía	117
6.6. Problemas propuestos	119
7. Fuentes de luz: led y láser	121
7.1. Introducción	121
7.2. Interacción radiación-materia	122
7.3. Características generales del láser	128
7.3.1. Cavidad Resonante	128
7.3.2. Coherencia	130
7.3.3. Longitud de Onda	131
7.3.4. Divergencia del rayo	131
7.3.5. Potencia de Salida	131
7.3.6. Modulación	132

7.4. Láseres Semiconductores	132
7.4.1. Propiedades eléctricas	132
7.4.2. Propiedades ópticas	136
7.4.3. Estructura del láser semiconductor	139
7.5. Aplicaciones del Láser	142
7.5.1. Aplicaciones del Láser de Inyección	142
7.6. Problemas propuestos	145
8. Fibra Óptica	147
8.1. Introducción	147
8.2. Guía de placas paralelas	148
8.2.1. Velocidad de grupo	150
8.2.2. Propagación multimodo	151
8.3. Guía dieléctrica plana	152
8.4. Fibra de salto de índice	154
8.4.1. Guiado de rayos	155
8.4.2. Número de modos	156
8.5. Fibra de gradiente de índice	158
8.6. Atenuación	160
8.7. Dispersión	162
8.8. Comunicaciones ópticas	164
8.9. Problemas propuestos	165
A. Análisis vectorial	167
A.1. Vectores	167
A.1.1. Notación vectorial	167
A.1.2. Suma de vectores	168
A.1.3. Producto escalar	168
A.1.4. Producto vectorial	169
A.1.5. Productos triples	170
A.1.6. Diferencial y derivada de funciones de una sola variable	170
A.1.7. Teorema fundamental del cálculo	171
A.1.8. Diferencial y derivada parcial de funciones de varias variables	171
A.1.9. Operador gradiente	172
A.1.10. Integral de camino	172
A.1.11. Teorema fundamental del gradiente	173
A.2. Integral de flujo	173
B. Funciones armónicas y Análisis fasorial	175

B.1. Funciones Armónicas	175
B.2. Análisis fasorial	175
C. Operadores divergencia y rotacional	179
C.1. Divergencia	179
C.2. Rotacional	179
D. Repaso del análisis de Fourier	181
D.1. Funciones periódicas	181
D.2. Funciones no periódicas	182
D.3. Transformada bidimensional de Fourier	183
D.3.1. Interpretación geométrica de la transformada bidimensional de Fourier	184
E. Función delta de Dirac	187
F. Introducción a los Sistemas Lineales	189

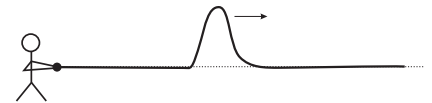
Copia Uso Académico

Tema 1

Nociones generales de Ondas

1.1. Introducción

En la Naturaleza existen muchos fenómenos físicos en los que una perturbación física viaja sin que ello lleve aparejado un desplazamiento neto de materia. Un ejemplo de esto puede ser la *ola* que se produce en el agua tras arrojar una piedra. En este fenómeno se observa el desplazamiento de una ondulación en la superficie del agua en la que las partículas individuales de agua *no viajan* sino que realizan un simple movimiento de vaivén (movimiento oscilatorio). Otro ejemplo, es la propagación del sonido, que básicamente es un desplazamiento de un cambio de presión en el aire pero sin que ello implique que las partículas de aire viajen desde el lugar donde se originó el sonido hasta el receptor; más bien cada partícula transmite su movimiento oscilatorio a la siguiente antes de volver a su posición original. Otro ejemplo bastante visual de este tipo de fenómenos se produce al agitar una cuerda por uno de sus extremos. En este caso se observaría claramente el desplazamiento de un pulso en la cuerda, siendo también evidente que cada segmento de cuerda no viaja junto a este pulso.



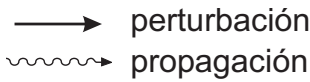
En todos los ejemplos anteriores una perturbación física se desplaza a través de un medio (agua, aire y cuerda respectivamente) sin que las partículas de este medio hayan sufrido un desplazamiento neto¹. Estos ejemplos son casos concretos de un tipo general de fenómenos físicos denominados **ondas**, las cuales pueden definirse como

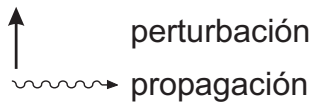
Propagación de una perturbación física sin que exista un transporte neto de materia.

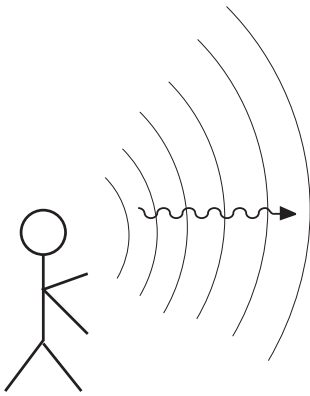
Debe notarse que la propagación de la perturbación en la onda implica el transporte de cierta energía y momento lineal. En este sentido, el comportamiento ondulatorio debe discernirse claramente del comportamiento de las *partículas*, puesto que estas últimas siempre transportan energía y momento lineal asociado a un transporte neto de materia.

Entre las posibles formas de clasificar a las ondas, a continuación se presentan dos de ellas:

¹ Debe notarse que la ausencia de un desplazamiento *neto* no implica la existencia de movimiento nulo. El movimiento oscilatorio de una partícula en torno a un punto fijo es un claro ejemplo de movimiento en el cual no existe traslación neta.







Ecuación de ondas no-dispersiva
monodimensional

■ Naturaleza física de la perturbación

- ONDAS MECÁNICAS: cuando la perturbación física es de naturaleza mecánica, por ejemplo: desplazamiento, velocidad, presión, torsión, etc.
- ONDAS ELECTROMAGNÉTICAS: cuando la perturbación es un campo electromagnético.

■ Dirección relativa de la perturbación y el desplazamiento ondulatorio

- ONDAS LONGITUDINALES: cuando la dirección de la perturbación física y de la propagación ondulatoria coinciden, por ejemplo: onda de sonido.
- ONDAS TRANSVERSALES: cuando la perturbación física se realiza en un plano transversal a la dirección de propagación de la onda; por ejemplo: el desplazamiento de un pulso en una cuerda, ondas electromagnéticas, etc.

Cuando se trata de caracterizar una onda, algunos conceptos usuales son:

- FOCO: Es el recinto donde se produce la perturbación inicial.
- SUPERFICIE/FRENTE DE ONDA: Es el lugar geométrico de los puntos en que han sido alcanzados simultáneamente por la perturbación.
- VELOCIDAD DE FASE: Velocidad con la que se propagan las superficies de onda.

Los conceptos anteriores pueden clarificarse si los concretamos en el caso de la propagación del sonido. En este caso, el *foco* sería el lugar donde se emiten los sonidos (por ejemplo la boca de alguien), la *superficie de onda* serían superficies aproximadamente esféricas centradas en el foco, y la *velocidad de fase* sería la velocidad a la que se viaja el frente de ondas, esto es, la velocidad del sonido $\sim 340\text{m/s}$.

1.2. Ecuación de ondas

Del mismo modo que existe una ecuación general que determina el momento lineal, \mathbf{p} , de una partícula (o conjunto de ellas) en función de la fuerza externa, \mathbf{F} ,

$$\mathbf{F} = \frac{d\mathbf{p}}{dt} \quad (1.1)$$

(o bien $F = md^2x/dt^2$ para el caso de movimiento monodimensional), existe también una ecuación, denominada *ecuación de ondas*, que se aplica a todos los fenómenos ondulatorios. La ecuación que describe el comportamiento ondulatorio de una perturbación física descrita matemáticamente como $u(x, t)$ que se propaga con velocidad constante v *sin distorsión* (**onda no-dispersiva**) a lo largo del eje x viene dada por

$$\frac{\partial^2 u}{\partial x^2} - \frac{1}{v^2} \frac{\partial^2 u}{\partial t^2} = 0 \quad (1.2)$$

Para mostrar que la ecuación anterior describe matemáticamente el fenómeno ondulatorio, analizaremos la propagación de un pulso en una cuerda dado que este ejemplo ofrece una imagen visual inmediata. En este caso, la perturbación $u(x, t)$ es justamente el desplazamiento vertical de cada trocito de cuerda. La forma del pulso para un instante arbitrario, que podemos tomar como $t = 0$, se muestra en la Figura 1.1(a), esto es, la forma matemática de la onda en ese instante de

tiempo viene completamente descrita por la función $u(x)$. Si tras un tiempo t , el pulso viaja sin distorsión hacia la derecha una distancia a , el perfil de la cuerda será como el mostrado en la Figura 1.1(b), pudiéndose describir matemáticamente por la función $u(x - a)$. Ahora bien, si el pulso viaja a una velocidad v , la distan-

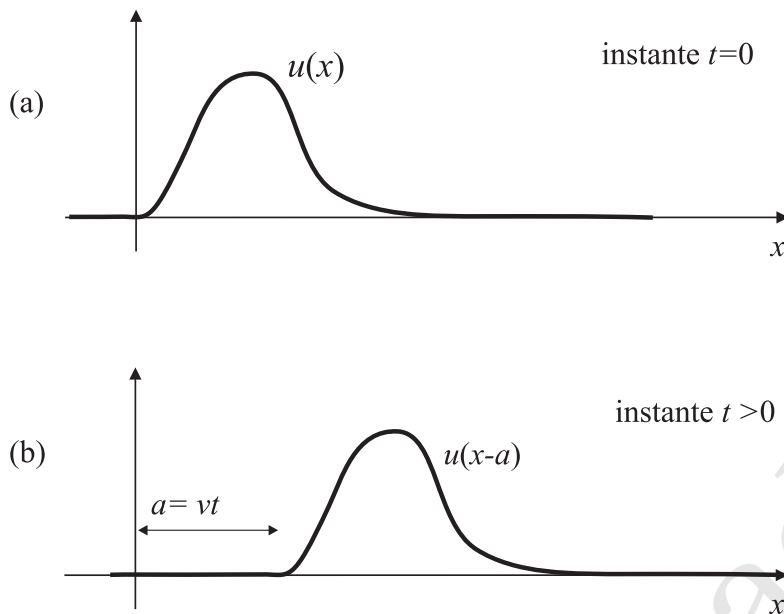


FIGURA 1.1: Evolución del pulso en una cuerda en dos instantes

cia a recorrida por el pulso puede escribirse como $a = vt$ y consecuentemente la expresión matemática de la onda en el instante t será

$$u(x, t) = f(x - vt). \quad (1.3)$$

Evidentemente, el pulso podría haber viajado igualmente hacia la izquierda, en cuyo caso, la expresión matemática de la onda viajera en la cuerda sería

$$u(x, t) = f(x + vt), \quad (1.4)$$

de modo que un movimiento ondulatorio general en la cuerda podría ser descrito por la función

$$u(x, t) = f(\chi) \text{ siendo } \chi = x \pm vt, \quad (1.5)$$

que representaría una onda que puede viajar tanto hacia la izquierda como hacia la derecha. Para encontrar la ecuación diferencial cuya solución general es una función del tipo (1.5), diferenciaremos la función $u(x, t)$ con respecto a x y a t , esto es,

$$\frac{\partial u}{\partial x} = \frac{du}{d\chi} \frac{\partial \chi}{\partial x} = u'(\chi) \quad (1.6)$$

$$\frac{\partial u}{\partial t} = \frac{du}{d\chi} \frac{\partial \chi}{\partial t} = \pm v u'(\chi). \quad (1.7)$$

Dado que las primeras derivadas no pueden relacionarse entre sí debido a la indefinición en el signo de (1.7), procedemos para obtener las derivadas segundas:

$$\frac{\partial^2 u}{\partial x^2} = \frac{\partial}{\partial x} \left[\frac{\partial u}{\partial x} \right] = \frac{d}{d\chi} [u'(\chi)] \frac{\partial \chi}{\partial x} = u''(\chi) \quad (1.8)$$

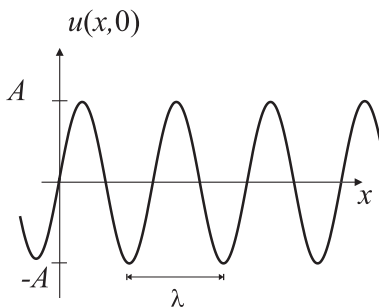
$$\frac{\partial^2 u}{\partial t^2} = \frac{\partial}{\partial t} \left[\frac{\partial u}{\partial t} \right] = \frac{d}{d\chi} [\pm v u'(\chi)] \frac{\partial \chi}{\partial t} = v^2 u''(\chi). \quad (1.9)$$

Si observamos ahora la forma de los segundos miembros de (1.8) y (1.9), podemos comprobar que al eliminar $u''(x)$ obtendríamos precisamente la ecuación general de ondas mostrada en (1.2). En consecuencia, esta ecuación diferencial en derivadas parciales tiene por soluciones a funciones del tipo (1.3) y (1.4) con la única condición de que éstas sean diferenciables hasta el segundo orden (la forma concreta de estas funciones en cada caso particular vendrá determinada por las condiciones iniciales del problema).

Una propiedad muy importante de la ecuación general de ondas es que ésta es *lineal*, lo que implica que si $u_1(x, t)$ y $u_2(x, t)$ son soluciones individuales de la ecuación de ondas, entonces la superposición de ambas, $u(x, t) = u_1(x, t) + u_2(x, t)$, también lo es. Esta propiedad de linealidad de la ecuación de ondas simplemente expresa en forma matemática el siguiente principio físico conocido como **principio de superposición de ondas**:

Principio de superposición de ondas

la perturbación ondulatoria resultante es igual a la suma de las perturbaciones coincidentes.



1.3. Ondas armónicas

Según se ha explicado en el apartado anterior, la expresión matemática general de una onda monodimensional no-dispersiva venía dada por (1.5). De entre las posibles formas matemáticas que puede tener este tipo de ondas, hay una especialmente interesante conocida como **onda armónica**. La forma de una onda armónica es una curva tipo senoidal, cuya instantánea en $t = 0$ puede venir dada por la siguiente expresión matemática:

$$u(x,0) = A \operatorname{sen} \left(\frac{2\pi}{\lambda} x \right). \quad (1.10)$$

La constante A es la **amplitud** de la onda y representa el valor máximo de la perturbación, λ es la **longitud de onda** o período espacial, esto es, la distancia en la que se repite la perturbación (por ejemplo, la distancia entre dos mínimos sucesivos). Si la onda se mueve hacia la derecha con cierta velocidad v , la función de onda en cualquier instante de tiempo t posterior vendrá dada por

$$u(x,t) = A \operatorname{sen} \left[\frac{2\pi}{\lambda} (x - vt) \right]. \quad (1.11)$$

El tiempo que tarda la onda en recorrer una longitud de onda se conoce como **período** T , por lo que

$$v = \frac{\lambda}{T} \quad \text{o} \quad \lambda = vT. \quad (1.12)$$

El período T corresponde igualmente al tiempo empleado por la perturbación en realizar una oscilación completa en un punto fijo.

Usando la definición del período, (1.11) puede escribirse como

$$u(x,t) = A \operatorname{sen} \left[2\pi \left(\frac{x}{\lambda} - \frac{t}{T} \right) \right]. \quad (1.13)$$

La expresión anterior indica claramente que la onda armónica muestra una doble periodicidad, tanto en el espacio como en el tiempo:

$$u(x,t) = u(x + n\lambda, t + mT). \quad (1.14)$$

Esta doble periodicidad es una consecuencia de la periodicidad temporal de la perturbación en el foco ($x = 0$), que se refleja en una periodicidad espacial².

La función de onda armónica puede expresarse en una forma más conveniente si se definen dos cantidades, k y ω que corresponden a la frecuencia espacial o **número de ondas** y a la frecuencia angular respectivamente, esto es,

$$k = 2\pi/\lambda \quad (1.15)$$

$$\omega = 2\pi/T. \quad (1.16)$$

Combinando las expresiones (1.15) y (1.16) junto con (1.12), obtenemos la siguiente relación para la frecuencia angular y el número de ondas de una onda armónica:

$$\omega = vk. \quad (1.17)$$

La frecuencia angular ω suele expresarse comúnmente en términos de la **frecuencia temporal**, f (siendo ésta la inversa del periodo: $f = 1/T$) mediante

$$\omega = 2\pi f. \quad (1.18)$$

La frecuencia temporal representa por tanto el número de oscilaciones realizadas por unidad de tiempo, siendo su unidad el **hertzio** (Hz).

Unidad de frecuencia: 1 hertzio (Hz)
 $\equiv \text{s}^{-1}$)

Teniendo en cuenta las definiciones dadas en (1.15) y (1.16), la función de onda armónica que viaja en el sentido positivo de las x puede escribirse como

$$u(x, t) = A \sin(kx - \omega t). \quad (1.19)$$

La expresión anterior es un caso particular de la siguiente expresión genérica usando la función coseno:

$$u(x, t) = A \cos(\omega t - kx - \varphi), \quad (1.20)$$

Expresión matemática de la onda armónica viajando en el sentido positivo de las x

donde el argumento completo del coseno se conoce como **fase** de la onda y la constante φ como **fase inicial** y se introduce para posibilitar que en $t = 0$ la perturbación en el foco ($x = 0$) pueda tomar un valor arbitrario: $u(0, 0) = A \cos \varphi$. Una onda armónica viajando en el sentido *negativo* de las x tendrá la siguiente forma general:

$$u(x, t) = A \cos(\omega t + kx - \varphi). \quad (1.21)$$

Es interesante notar que el carácter viajero de la onda en sentido positivo/negativo del eje x lo determina la desigualdad/igualdad entre los signos que acompañan a ωt y kx en la fase.

Para facilitar las operaciones con ondas armónicas, éstas suelen expresarse en forma de exponencial compleja, de manera que la onda armónica dada en (1.20) se escribirá usualmente como

$$u(x, t) = A e^{-j(kx + \varphi)} e^{j\omega t} \quad (1.22)$$

Expresión matemática compleja de la onda armónica

(ver Apéndice B para un estudio de los fasores), aunque debe considerarse que $u(x, t)$ tal como se ha expresado en (1.20) es solamente la parte real de (1.22):

$$u(x, t) = A \cos(\omega t - kx - \varphi) = \text{Re} \left(A e^{-j(kx + \varphi)} e^{j\omega t} \right). \quad (1.23)$$

No obstante, en lo que sigue del tema, cuando tratemos con ondas armónicas usaremos la notación compleja por simplicidad, debiéndose sobreentender que la onda *verdadera* es la parte real de la expresión compleja.

² De manera análoga a como un pastelero soltando pasteles cada tiempo T en un extremo de una cinta transportadora (periodicidad temporal en el foco) que se mueve con velocidad v da lugar a una periodicidad espacial en dicha cinta; esto es, los pasteles aparecen distanciados una longitud de onda.

1.4. Energía e Intensidad de la onda

La **intensidad** de una onda se define como la energía que fluye por unidad de tiempo a través de una superficie unidad situada perpendicularmente a la dirección de propagación. Si ρ_U es la densidad volumétrica de energía de la onda (esto es, la energía media por unidad de volumen contenida en el medio donde se propaga la onda) y v la velocidad de propagación de la onda, la intensidad I de la onda es

$$I = \rho_U v, \quad (1.24)$$

cuyas unidades son $(\text{ms}^{-1})(\text{Jm}^{-3}) = \text{Js}^{-1}\text{m}^{-2} = \text{Wm}^{-2}$; es decir, potencia por unidad de área.

Analicemos el caso particular en el que una perturbación ondulatoria *no-dispersiva armónica* (por ejemplo una onda sonora) se propaga a lo largo de la dirección x en un medio homogéneo, isótropo y no absorbente de densidad de masa ρ_m . En este caso, el movimiento armónico de cada partícula (y también su energía) es transmitida a las restantes partículas del medio que la rodea. La expresión matemática del desplazamiento de las partículas será

$$u(x, t) = A \cos(\omega t - kx), \quad (1.25)$$

donde la fase inicial se ha tomado como cero por simplicidad ($\varphi = 0$). Las partículas de este medio adquieren, al paso de la onda, una cierta velocidad, $\dot{u}(x, t) = \partial u / \partial t$, y por tanto una energía mecánica debida al movimiento armónico que realizan. En este tipo de movimiento, la energía total de las partículas puede igualarse a la energía cinética *máxima* que adquieren. Por tanto, la energía, ΔU , de un elemento de volumen, ΔV (de sección transversal S y anchura $\Delta l = v\Delta t$, siendo v la velocidad de propagación de la perturbación ondulatoria) puede expresarse como

$$\Delta U = \frac{1}{2} \Delta m \dot{u}_{\max}^2 = \frac{1}{2} \rho_m S v \Delta t \dot{u}_{\max}^2. \quad (1.26)$$

Para calcular la velocidad máxima, \dot{u}_{\max} , de las partículas, primero debemos obtener su velocidad derivando su desplazamiento $u(x, t)$ con respecto al tiempo, esto es,

$$\dot{u}(x, t) = \frac{\partial u(x, t)}{\partial t} = -\omega A \sin(\omega t - kx). \quad (1.27)$$

El máximo del módulo de la velocidad que adquieren las partículas será por tanto $\dot{u}_{\max} = \omega A$, por lo que la energía puede reescribirse como

$$\Delta U = \frac{1}{2} \rho_m S v \Delta t \omega^2 A^2, \quad (1.28)$$

por lo que la densidad volumétrica de energía vendrá dada por

$$\rho_U = \frac{\Delta U}{\Delta V} = \frac{1}{2} \rho_m \omega^2 A^2. \quad (1.29)$$

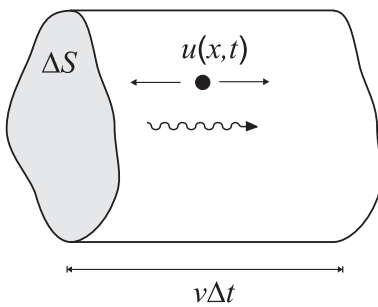
La potencia, o energía transmitida a cada capa del medio por unidad de tiempo, se obtiene a partir de

$$\text{Potencia} = \frac{\Delta U}{\Delta t} = \frac{1}{2} \rho_m S v \omega^2 A^2. \quad (1.30)$$

Dado que la intensidad, I , de la onda viajera es potencia por unidad de área, se tiene finalmente que

$$I = \frac{\text{Potencia}}{\text{Área}} = \frac{1}{2} \rho_m \omega^2 A^2 v. \quad (1.31)$$

Intensidad de la onda



Intensidad de la onda armónica
mecánica: $I \propto A^2$

Para el caso particular que estamos analizando, teniendo en cuenta (1.29), observamos que (1.31) puede también expresarse como $I = \rho v v$, en concordancia con la expresión general (1.24). Es también interesante notar en (1.31) que la intensidad de la onda armónica mecánica es proporcional al cuadrado de la frecuencia y al cuadrado de la amplitud de la perturbación.

EJEMPLO 1.1 A 1000 Hz, el umbral de audibilidad del oído humano corresponde a una intensidad de 10^{-12} W/m². ¿Cuál es el desplazamiento máximo de las moléculas de aire en este límite?

Datos: $\rho_{\text{aire}} = 1.2$ kg/m³, $v = 340$ m/s.

Dado que la intensidad de la onda sonora viene dada por

$$I = \frac{1}{2} \rho_m \omega^2 A^2 v,$$

la amplitud, A , correspondiente al desplazamiento máximo de las moléculas del medio será

$$A = \sqrt{\frac{2I}{\rho_m \omega^2 v}}.$$

Al sustituir los datos del presente problema tendremos que

$$A = \sqrt{\frac{2 \cdot 10^{-12}}{1.2 \cdot 4\pi^2 \times 10^6 \cdot 340}} \approx 1.1 \times 10^{-11} \text{ m}.$$

Este desplazamiento es minúsculo, lo que demuestra la alta sensibilidad del oído humano.

1.5. Interferencia de Ondas

Cuando dos o más ondas coinciden en el espacio en el mismo instante de tiempo se produce un fenómeno que se conoce como **interferencia**. El principio de superposición de ondas establece que cuando dos o más ondas coinciden en un punto y en un instante de tiempo, la perturbación resultante es la *suma* de las perturbaciones individuales (este principio ya fue relacionado en el Apartado 1.2 con la linealidad de la ecuación de ondas). En consecuencia, la perturbación resultante en un punto P y en un instante de tiempo t , $u(P, t)$, debido a la coincidencia de N ondas $u_i(x, t)$ se obtendrá mediante la siguiente expresión:

$$u(P, t) = \sum_{i=1}^N u_i(P, t). \quad (1.32)$$

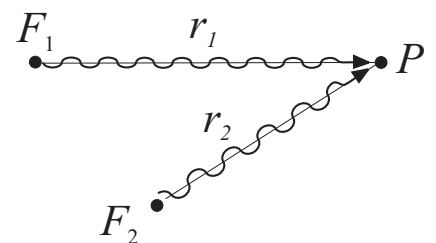
1.5.1. Superposición de dos ondas armónicas

Para estudiar los aspectos cuantitativos de la interferencia consideraremos la *superposición de dos ondas armónicas monodimensionales de la misma frecuencia pero distinta amplitud y fase inicial*,

$$u_1(r, t) = A_1 \cos(\omega t - kr - \varphi_1)$$

y

$$u_2(r, t) = A_2 \cos(\omega t - kr - \varphi_2),$$



en cierto punto P . Si r_1 y r_2 son las distancias desde los focos respectivos al punto P , la perturbación resultante vendrá dada por

$$u(P, t) = u_1(r_1, t) + u_2(r_2, t) . \quad (1.33)$$

Si usamos la notación compleja, la perturbación suma puede obtenerse a partir de

$$\begin{aligned} u(P, t) &= A_1 e^{-j(kr_1 - \omega t + \varphi_1)} + A_2 e^{-j(kr_2 - \omega t + \varphi_2)} \\ &= \left[A_1 e^{-j\varepsilon_1} + A_2 e^{-j\varepsilon_2} \right] e^{j\omega t} \\ &= A(P) e^{-j\varepsilon(P)} e^{j\omega t} , \end{aligned} \quad (1.34)$$

donde

$$\varepsilon_i = kr_i + \varphi_i \quad (1.35)$$

y $A(P)$ y $\varepsilon(P)$ son respectivamente la amplitud y la fase de la perturbación resultante en el punto P . Operando en (1.34) encontramos que

$$\begin{aligned} A(P) e^{-j\varepsilon(P)} &= A_1 e^{-j\varepsilon_1} + A_2 e^{-j\varepsilon_2} \\ &= (A_1 \cos \varepsilon_1 - j A_1 \sin \varepsilon_1) + (A_2 \cos \varepsilon_2 - j A_2 \sin \varepsilon_2) \\ &= (A_1 \cos \varepsilon_1 + A_2 \cos \varepsilon_2) - j (A_1 \sin \varepsilon_1 + A_2 \sin \varepsilon_2) , \end{aligned} \quad (1.36)$$

de donde obtenemos que la amplitud puede ser calculada como sigue:

$$\begin{aligned} A^2(P) &= A_1^2 \cos^2 \varepsilon_1 + A_2^2 \cos^2 \varepsilon_2 + 2A_1 A_2 \cos \varepsilon_1 \cos \varepsilon_2 + \\ &\quad A_1^2 \sin^2 \varepsilon_1 + A_2^2 \sin^2 \varepsilon_2 + 2A_1 A_2 \sin \varepsilon_1 \sin \varepsilon_2 \\ &= A_1^2 + A_2^2 + 2A_1 A_2 \cos(\varepsilon_1 - \varepsilon_2) , \end{aligned}$$

Amplitud de la interferencia de dos ondas armónicas de igual frecuencia

esto es,

$$A(P) = \sqrt{A_1^2 + A_2^2 + 2A_1 A_2 \cos \delta(P)} , \quad (1.37)$$

siendo

$$\begin{aligned} \delta(P) &= kr_1 - kr_2 + \varphi_1 - \varphi_2 \\ &= k\Delta r + \Delta\varphi . \end{aligned} \quad (1.38)$$

En la expresión anterior, $\delta(P)$ se denomina **diferencia de fase**, $\Delta r = r_1 - r_2$ se conoce como **diferencia de camino** entre el recorrido de las dos ondas al propagarse desde los focos respectivos hasta el punto P y $\Delta\varphi = \varphi_1 - \varphi_2$ es la **diferencia de fase inicial** entre las dos ondas. El último término de la expresión anterior,

$$2A_1 A_2 \cos \delta(P) ,$$

se denomina usualmente **término de interferencia** puesto que es el responsable de que la amplitud de la interferencia varíe al variar la diferencia de camino hasta el punto P . En concreto, si notamos que $-1 \leq \cos \delta(P) \leq 1$ encontraremos que la amplitud en un punto podrá tomar en general valores comprendidos entre

$$(A_1 - A_2) \leq A \leq (A_1 + A_2) . \quad (1.39)$$

Para obtener la intensidad resultante de la superposición de las dos ondas armónicas de igual frecuencia en el punto P debemos tener en cuenta que, según (1.31), la intensidad de las ondas armónicas depende del cuadrado de la amplitud ($I \propto A^2$). En consecuencia, a partir de (1.37), podemos deducir que la intensidad resultante será

$$I(P) = I_1 + I_2 + 2\sqrt{I_1 I_2} \cos \delta(P) . \quad (1.40)$$

1.5.2. Focos incoherentes

En el apartado anterior observamos que la amplitud resultante en el punto P oscilaba entre dos valores dependiendo del valor concreto de δ en dicho punto. No obstante, en la práctica ocurre frecuentemente que la diferencia de fase no es constante en el tiempo sino que $\delta = \delta(t)$. Esto puede ser debido a una posible variación temporal de las condiciones de emisión de los focos (usualmente en un tiempo del orden de 10^{-10} s); por ejemplo:

1. La frecuencia de los focos no es estrictamente constante sino que presenta pequeñas fluctuaciones arbitrarias que provocan que el número de ondas (y equivalentemente la longitud de onda) oscile ligeramente en torno a cierto valor promedio, \bar{k}

$$k(t) = \bar{k} + \Delta k(t).$$

2. Las fases iniciales de los dos focos presentan fluctuaciones al azar de modo que las funciones $\varphi_1(t)$ y $\varphi_2(t)$ no están correlacionadas de ninguna manera dando lugar a que la diferencia de fase inicial sea una función del tiempo,

$$\Delta\varphi = \varphi_1(t) - \varphi_2(t) = f(t),$$

que varía igualmente al azar.

Cuando nos encontramos con alguna de las condiciones anteriores decimos que los focos son **incoherentes**. Debido a esta rápida variación arbitraria en el tiempo de la diferencia de fase, el término de interferencia *se anula en promedio* durante el intervalo de observación debido a que el valor medio del coseno de un argumento que varía al azar es cero:

$$\langle \cos \delta(t) \rangle = \frac{1}{T} \int_0^T \cos \delta(t) dt = 0.$$

Esto hecho implica que la intensidad promedio en el punto P , $\langle I(P) \rangle$, venga dada por

$$\langle I(P) \rangle = I_1 + I_2 \quad \text{focos incoherentes.} \quad (1.41)$$

Notemos que en el presente caso de focos incoherentes, la anulación en promedio del término de interferencia hace que la intensidad de la perturbación NO dependa de la posición del punto de observación. Este hecho provoca que aunque podamos, en un sentido estricto, hablar de interferencia, ésta *no será observable* y usualmente diremos que “no existe interferencia”.

A menudo cuando se habla de un *único* foco también podemos decir que este foco es “incoherente”. En este caso, en realidad estamos queriendo decir que este único foco tiene cierta extensión espacial, y que las distintas partes del foco (asimilables a diversos focos puntuales) no son coherentes entre sí.

1.5.3. Focos coherentes

Cuando la fase inicial de los dos focos está completamente correlacionada, de modo que

$$\varphi_1(t) - \varphi_2(t) \neq f(t),$$

manteniendo una diferencia de fase inicial constante, se dice que los dos focos son **coherentes**. En el caso de que $\Delta\varphi = 0$, δ sólo dependerá de la diferencia de camino (en general Δr),

$$\delta = k\Delta r = 2\pi\Delta r/\lambda, \quad (1.42)$$

dando lugar así a una interferencia que sí podría ser observable.

En las circunstancias anteriores, podemos distinguir dos casos de interés, dependiendo de si $\cos \delta$ es 1 o -1 , esto es, si A adquiere su valor máximo (**interferencia constructiva**) o bien su valor mínimo (**interferencia destructiva**). Por tanto, si

$$\delta = \begin{cases} 2n\pi & \Rightarrow A = A_1 + A_2 \text{ Interferencia Constructiva} \\ (2n+1)\pi & \Rightarrow A = A_1 - A_2 \text{ Interferencia Destructiva.} \end{cases} \quad (1.43)$$

Teniendo en cuenta (1.42), la condición de interferencia constructiva o destructiva para Δr en P vendrá dada por

$$\Delta r = \begin{cases} n\lambda & \text{Interferencia Constructiva} \\ (2n+1)\frac{\lambda}{2} & \text{Interferencia Destructiva;} \end{cases} \quad (1.44)$$

es decir, si la diferencia de camino es un múltiplo *entero/semientero* de la longitud de onda, entonces tendremos interferencia constructiva/destructiva.

Desde un punto de vista práctico, una forma usual de producir focos coherentes es generar dos focos secundarios a partir de la misma fuente primaria, asegurando así que la diferencia de fase inicial en los dos focos secundarios es una constante. Uno de los primeros experimentos que mostró el fenómeno de interferencia con luz es el **experimento de la doble rendija de Young** mostrado en la Figura 1.2(a), constatando así convincentemente la naturaleza ondulatoria de la luz. En este experimento, la luz (u otra perturbación ondulatoria) proveniente de

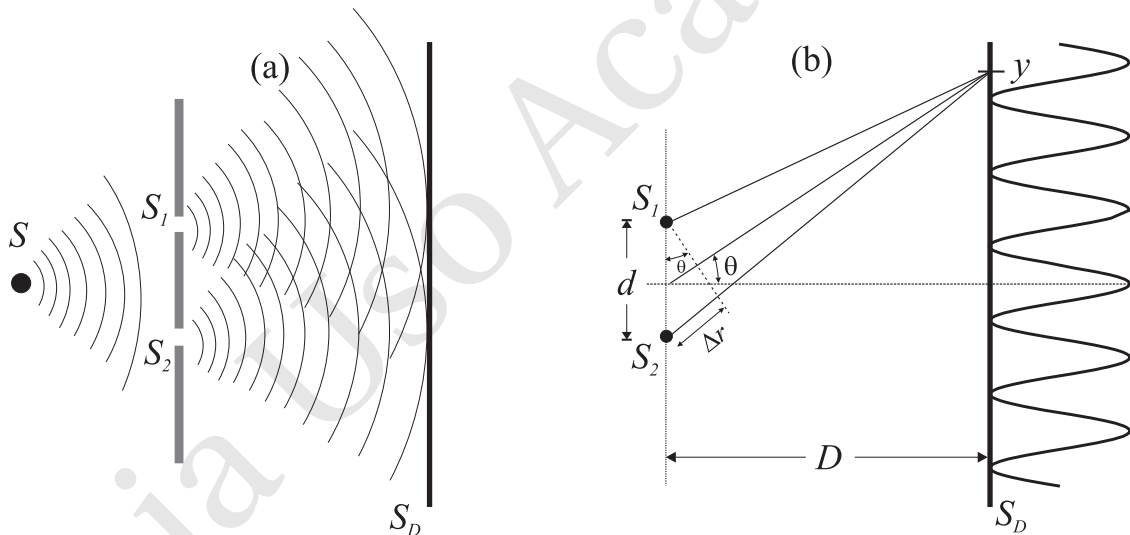


FIGURA 1.2: Experimento de la doble rendija de Young

un foco primario S se hace pasar por una pantalla en la que se han realizado dos ranuras S_1 y S_2 separadas una distancia d . Las rendijas se comportan como dos focos *coherentes* de luz cuyas ondas interfieren en el semiespacio derecho. Este fenómeno provoca un patrón de interferencias en la pantalla S_D donde aparecen regiones sombreadas (dibujadas en negro) junto a regiones más iluminadas tal y como se muestra en la Figura 1.3. En este experimento tenemos que la amplitud de las ondas que interfieren es idéntica, esto es,

$$A_1 = A_2 .$$

Si además consideramos que la pantalla S_D se coloca a una distancia $D \gg d$ de las rendijas y admitimos que $\theta \ll$ entonces, según muestra la Figura 1.2(b), encontramos que la diferencia de camino es distinta para distintas coordenadas y

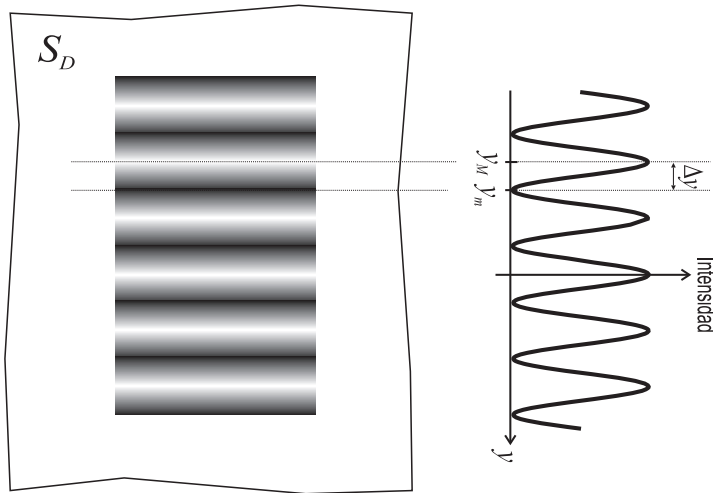


FIGURA 1.3: Patrón de interferencia resultante en el experimento de la doble rendija de Young

en la pantalla:

$$\Delta r = d \sin \theta \approx d \tan \theta \approx d \frac{y}{D}. \quad (1.45)$$

En consecuencia, el patrón de interferencia obtenido en la pantalla S_D mostrará franjas de interferencia constructiva o bien destructiva según se cumplan las siguientes condiciones:

- **Interferencia constructiva**, $y = y_M$:

$$k\Delta r = 2n\pi \Rightarrow \frac{2\pi}{\lambda} d \frac{y_M}{D} = 2n\pi, \quad (1.46)$$

de donde se deduce que las franjas $y = y_M$ de interferencia constructiva verifican

$$y_M = n \frac{d}{D} \lambda, \quad (1.47)$$

siendo la intensidad de la onda en estas franjas: $I = 4I_1$.

- **Interferencia destructiva**, $y = y_m$:

$$k\Delta r = (2n+1)\pi \Rightarrow \frac{2\pi}{\lambda} d \frac{y_m}{D} = (2n+1)\pi, \quad (1.48)$$

de donde se deduce que las franjas $y = y_m$ de interferencia destructiva verifican

$$y_m = \frac{2n+1}{2} \frac{d}{D} \lambda, \quad (1.49)$$

siendo la intensidad de la onda en estas franjas $I = 0$.

La diferencia Δy entre un máximo y mínimo consecutivo será

$$\Delta y = \frac{d}{D} \frac{\lambda}{2}. \quad (1.50)$$

Esta expresión nos proporciona adicionalmente un procedimiento muy sencillo para determinar el valor de la longitud de onda a partir de la medida de la distancia entre franjas de interferencia constructiva y destructiva.

Es interesante notar que en las franjas de interferencia constructiva se ha obtenido que la intensidad es *cuatro* veces (y no dos) el valor de la intensidad proporcionada por cada uno de los focos. Esto parece violar el principio de conservación de la energía, aunque tal hecho no se produce puesto que la energía de la onda no se distribuye homogéneamente en la pantalla S_D sino que, debido a la interferencia, existen puntos donde la energía es mayor que la suma de las energías provenientes de los focos pero también existen otros puntos donde la energía es menor (incluso cero) que la proveniente de los focos.

EJEMPLO 1.2 Un foco de luz amarilla ($\lambda = 600\text{nm}$) incide sobre dos rendijas separadas una distancia d , observándose la interferencia de la luz proveniente de estas rendijas en una pantalla situada a una distancia de 3 m. Obtener la separación d entre las rendijas para que la distancia entre máximos y mínimos consecutivos del patrón de interferencia luminoso sea mayor que 5 mm.

Según la teoría expuesta anteriormente, la distancia entre máximos y mínimos consecutivos en el experimento de Young viene dado por

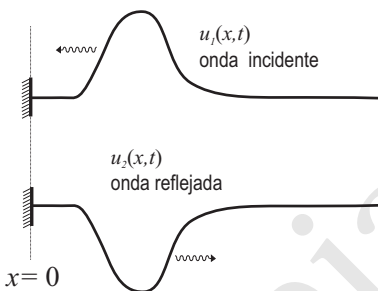
$$\Delta y > \frac{d \lambda}{D 2}.$$

Al despejar en la expresión anterior d encontramos que

$$d < \frac{D\lambda}{2\Delta y} = \frac{3 \cdot 6 \times 10^{-7}}{2 \cdot 5 \times 10^{-3}} = 1.8 \times 10^{-4} \text{m} = 180 \mu\text{m}.$$

El resultado anterior nos muestra que la separación entre rendijas debe ser muy pequeña (y aún menor si Δy se quiere mayor) por lo que en la práctica no es fácil llevar a cabo este experimento.

1.6. Ondas estacionarias



Observemos que cuando una perturbación viaja hacia la izquierda por una cuerda, al llegar al extremo, ésta se refleja de la forma mostrada en la figura adjunta.

Si la situación anterior ocurre para una perturbación armónica que viaja hacia la izquierda por una cuerda con uno de sus extremos fijos:

$$u_1(x, t) = A_1 \cos(\omega t + kx),$$

al llegar al extremo de la cuerda, $x = 0$, la onda se refleja dando lugar a una onda armónica viajando hacia la derecha:

$$u_2(x, t) = A_2 \cos(\omega t - kx).$$

Dado que las dos ondas viajeras anteriores se encuentran en una misma *región* del espacio darán lugar a un fenómeno típico de superposición o interferencia. Puesto que en el punto $x = 0$, la perturbación de la cuerda debe ser nula (puesto que, por hipótesis, es un extremo *fijo*) para cualquier instante de tiempo, tendremos que

$$\begin{aligned} u(0, t) &= A_1 \cos(\omega t) + A_2 \cos(\omega t) \\ &= (A_1 + A_2) \cos \omega t = 0, \end{aligned} \quad (1.51)$$

de donde se deduce que $A_1 = -A_2$.

Como las dos perturbaciones ondulatorias anteriores coinciden simultáneamente en la misma región, la superposición de ambas (usando notación compleja) dará lugar a la siguiente onda:

$$\begin{aligned} u(x, t) &= -A_2 e^{j(\omega t + kx)} + A_2 e^{j(\omega t - kx)} = A_2 (-e^{jkx} + e^{-jkx}) e^{j\omega t} \\ &= A \operatorname{sen}(kx) e^{j(\omega t - \pi/2)} \end{aligned} \quad (1.52)$$

(donde $A = 2A_2$ y $-j$ se ha escrito como $e^{-j\pi/2}$), cuya parte real puede finalmente escribirse como

$$u(x, t) = A \operatorname{sen} kx \operatorname{sen} \omega t. \quad (1.53)$$

Nótese que en la expresión (1.53) no aparecen explícitamente expresiones del tipo $f(\omega t \pm kx)$, lo que indica que esta perturbación no puede identificarse ya simplemente con una onda viajera, sino que constituye un nuevo tipo de onda conocido como **onda estacionaria**. En este tipo de perturbación ya no podemos decir que la energía viaja de un punto a otro sino que, como muestra la figura 1.4, esta onda

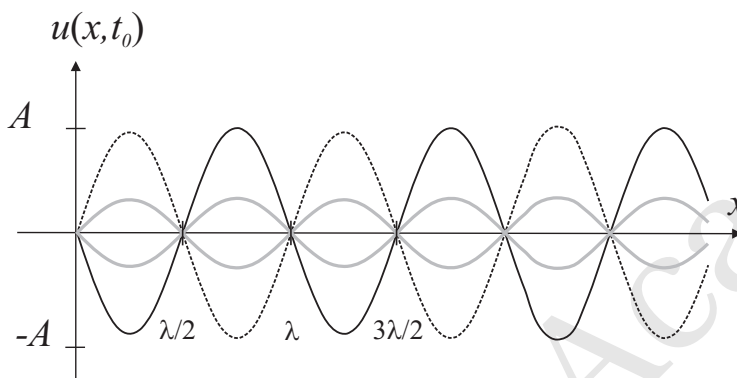


FIGURA 1.4: Instantánea de la onda estacionaria en $t = t_0$. Los nodos están separados una distancia $\lambda/2$.

estacionaria corresponde a una situación en la que cada elemento individual de la cuerda realiza un *movimiento armónico simple* cuya amplitud es una función de x , $A(x)$, pudiéndose escribir entonces que

$$u(x, t) = A(x) \operatorname{sen} \omega t, \quad (1.54)$$

siendo

$$A(x) = A \operatorname{sen} kx. \quad (1.55)$$

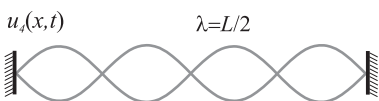
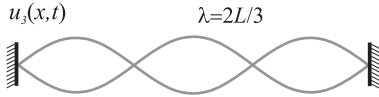
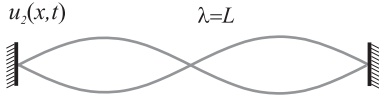
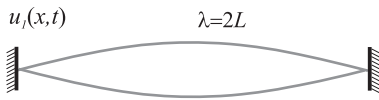
Observemos que en la situación anterior cada elemento de la cuerda es como si actuase de forma independiente, sin “transmitir su movimiento al siguiente”. De hecho podemos incluso encontrar puntos denominados **nodos** donde la perturbación es nula para todo instante de tiempo. Estos puntos son aquellos que verifican que la amplitud es nula, es decir, aquellos que satisfacen la siguiente condición:

$$kx = n\pi \Rightarrow x_{\text{nodo}} = n \frac{\lambda}{2}, \quad (1.56)$$

siendo la distancia, Δ , entre dos nodos sucesivos una semilongitud de onda ($\Delta = \lambda/2$; recuérdese que la longitud de onda está determinada por la frecuencia y la velocidad de propagación de la onda: $\lambda = v/f$).

Si ahora imponemos al problema anterior una *segunda condición* consistente en que el otro extremo de la cuerda (el punto $x = L$) también esté fijo, entonces ha de verificarse igualmente que

$$u(L, t) = 0,$$



lo cual requiere que

$$\text{sen } kL = 0 \Rightarrow kL = n\pi. \quad (1.57)$$

La condición anterior implica que tanto el número de ondas como la longitud de onda de la onda estacionaria sólo pueden tomar ciertos valores *discretos* (fenómeno conocido como *cuantización*) dados por

$$k_n = n \frac{\pi}{L} = \frac{\pi}{L}, \frac{2\pi}{L}, \frac{3\pi}{L}, \dots \quad (1.58)$$

$$\lambda_n = \frac{2L}{n} = 2L, \frac{2L}{2}, \frac{2L}{3}, \dots \quad (1.59)$$

Vemos entonces que la imposición de (1.57) ha limitado los valores de las longitudes de onda permitidas en la cuerda fija por sus dos extremos a aquéllos que cumplan la condición (1.59). De forma análoga, las frecuencias permitidas en la cuerda serán aquéllas que cumplan

$$\omega_n = vk_n = v \frac{n\pi}{L}. \quad (1.60)$$

En consecuencia podemos concluir que tanto las longitudes de onda como las frecuencias permitidas en la cuerda están **cuantizadas** y que esta cuantización es fruto de la imposición de condiciones de contorno en *ambos* extremos de la cuerda.

EJEMPLO 1.3 Una cuerda tensa de $L = 60\text{cm}$ de longitud y que tiene una masa por unidad de longitud $\rho_m = 60\text{g/m}$ es sometida a una tensión $\mathcal{T} = 900\text{N}$. Suponiendo que la cuerda se hace vibrar en su segundo armónico, ¿cuál será la frecuencia del sonido que generará?

Teniendo en cuenta que la velocidad de propagación de la onda en la cuerda viene dada por

$$v = \sqrt{\frac{\mathcal{T}}{\rho_l}},$$

en el presente caso, esto implica que la velocidad será

$$v = \sqrt{\frac{900}{6 \times 10^{-2}}} \approx 122.47\text{m/s}.$$

Dado que la cuerda vibra en su segundo armónico ($n = 2$), entonces la longitud de onda en este armónico será

$$\lambda_2 = 2L/2 = 0.6\text{m},$$

por lo que la frecuencia, f , de la perturbación será

$$f = \frac{v}{\lambda} = \frac{122.47}{0.6} \approx 204.12\text{Hz}.$$

Dado que la cuerda vibrante es el foco del sonido, la frecuencia del sonido que escucharíamos sería justamente la misma que la de vibración de la cuerda.

Finalmente observemos que en la cuerda con extremos fijos cada una de las ondas estacionarias permitidas responden a la siguiente expresión:

$$u_n(x, t) = A_n \text{sen}(k_n x) \text{sen}(\omega_n t + \varphi_n),$$

denominándose genéricamente como *armónicos*. Estos armónicos presentan la importante propiedad de que cualquier perturbación en la cuerda puede expresarse como una superposición de ellos, esto es,

$$\begin{aligned} u(x, t) &= \sum_{n=1}^{\infty} A_n \text{sen}(k_n x) \text{sen}(\omega_n t + \varphi_n) \\ &= \sum_{n=1}^{\infty} A_n \text{sen}(nk_1 x) \text{sen}(n\omega_1 t + \varphi_n), \end{aligned} \quad (1.61)$$

siendo

$$k_1 = \frac{\pi}{L}, \quad \omega_1 = v \frac{\pi}{L}$$

y A_n la amplitud del n -ésimo armónico (esta amplitud será distinta en cada caso particular). El resultado anterior puede considerarse como una conclusión particular de un teorema más general, llamado **teorema de Fourier**, que básicamente dice que una función periódica puede expresarse como la suma de senos/cosenos cuyas frecuencias son un número entero de veces la frecuencia original del problema (un tratamiento detallado de este teorema puede encontrarse en cualquier libro de Cálculo).

1.7. Difracción

Uno de los fenómenos ondulatorios más característicos es el conocido como difracción. Este fenómeno se produce cuando una onda es distorsionada en su propagación por un obstáculo, aunque también se llama difracción a la interferencia producida por *muchos* focos coherentes elementales. Desde un punto de vista físico, la difracción no se diferencia básicamente de la interferencia puesto que ambos fenómenos son fruto de la superposición de ondas. La difracción puede ser, por ejemplo, la causa de la desviación de la luz de una trayectoria recta, explicando así por qué la luz llega a puntos que, en principio, no debería alcanzar si su propagación fuese estrictamente rectilínea. Un ejemplo de difracción puede verse en la Figura 1.5(b) que muestra el patrón de sombras cuando una fuente de luz *coherente* ilumina una esquina recta. En la Fig. 1.5(a) se muestra esta misma sombra cuando no se produce difracción (por ejemplo, cuando la fuente de luz es incoherente).

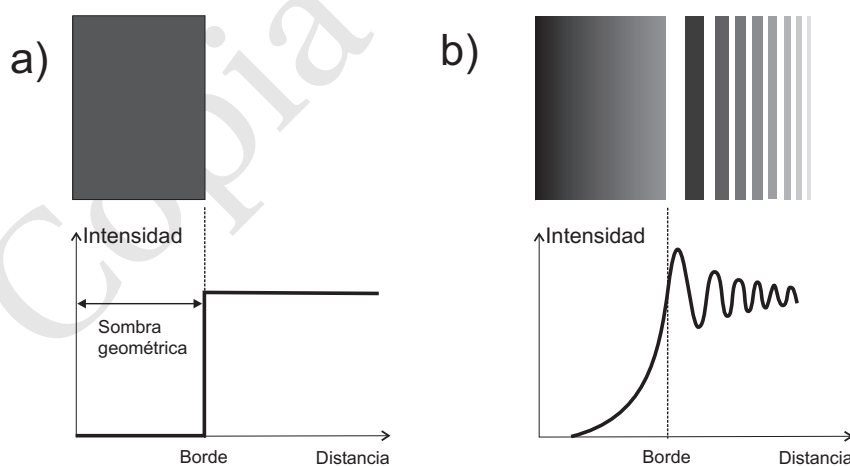
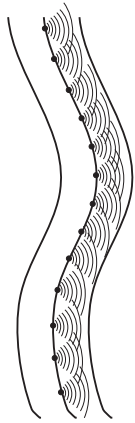


FIGURA 1.5: Sombra producida por una esquina recta iluminada por una fuente de luz cuando: (a) NO se produce difracción, (b) sí se produce difracción



En el presente estudio de la difracción, consideraremos únicamente la denominada **difracción de Fraunhofer**, que se presenta cuando las ondas incidentes pueden considerarse planas (esto es, su frente de onda es un plano) y el patrón de difracción es observado a una distancia lo suficientemente lejana como para que solo se reciban los rayos difractados *paralelamente*. Este estudio lo basaremos en el **principio de Huygens**, que explica la propagación ondulatoria suponiendo que

cada punto de un frente de ondas primario se comporta como un foco de ondas esféricas elementales secundarias que avanzan con una velocidad y frecuencia igual a la onda primaria. La posición del frente de ondas primario al cabo de un cierto tiempo es la envolvente de dichas ondas elementales.

Si siguiendo este principio, cuando un frente de onda alcanza una pantalla en la que existe una rendija de anchura b , tal y como se muestra en la Figura 1.6, sólo

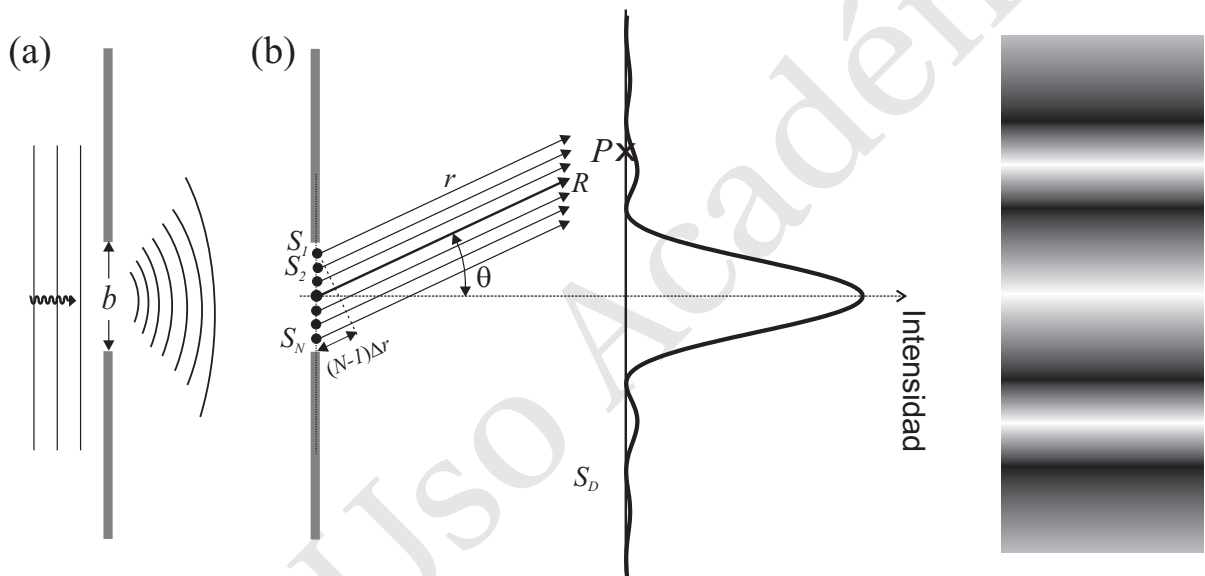


FIGURA 1.6: (a) Difracción de Fraunhofer de una rendija rectangular; (b) Cada punto de la rendija se comporta como un foco puntual emisor de ondas secundarias.

aquellos puntos del frente de ondas coincidentes con la rendija se convierten en focos emisores secundarios, de modo que la perturbación ondulatoria en cualquier punto a la derecha de la rendija puede calcularse como la superposición de las ondas originadas en cada uno de estos focos secundarios (ver Figura 1.6b).

En este sentido, y a efectos de cálculo, supondremos que existen N focos puntuales equiespaciados en la rendija. La perturbación ondulatoria, $u(x, t)$, en cierto punto P de una pantalla S_D (situada a una distancia $D \gg d$) será fruto de la *interferencia* de un gran número de fuentes equiespaciadas de igual amplitud y fase inicial, esto es,

$$u(P, t) = \sum_{n=1}^N A_0 e^{-j(kr_n - \omega t)}, \quad (1.62)$$

donde r_n es la distancia desde el foco secundario n -ésimo hasta el punto P y A_0 la amplitud constante de cada onda elemental. Notemos que, bajo la presente aproximación, todos los rayos que llegan a P se consideran paralelos. Si llamamos r a la distancia desde el foco 1 hasta P y Δr a la diferencia de camino entre la

perturbación que llega a P desde un foco y el siguiente, r_n puede escribirse como

$$r_n = r + (n-1)\Delta r.$$

La perturbación en P según (1.62) puede entonces expresarse como

$$\begin{aligned} u(P, t) &= A_0 \left[e^{-jkr} + e^{-jk(r+\Delta r)} + e^{-jk(r+2\Delta r)} + \dots \right] e^{j\omega t} \\ &= A_0 \left[1 + e^{-j\phi} + e^{-2j\phi} + \dots + e^{-j(N-1)\phi} \right] e^{-j(kr-\omega t)}, \end{aligned} \quad (1.63)$$

donde $\phi = k\Delta r$ y pudiéndose identificar la suma entre corchetes como una serie geométrica, S_g , de razón $q = e^{-j\phi}$. Dado que la suma de la siguiente serie geométrica viene dada por

$$1 + q + q^2 + \dots + q^{N-1} = \frac{1 - q^N}{1 - q},$$

el resultado de la serie geométrica en (1.63) puede expresarse como

$$\begin{aligned} S_g &= \frac{1 - e^{jN\phi}}{1 - e^{j\phi}} = \frac{e^{jN\phi/2} e^{-jN\phi/2} - e^{jN\phi/2}}{e^{j\phi/2} e^{-j\phi/2} - e^{j\phi/2}} \\ &= \frac{\text{sen}(N\phi/2)}{\text{sen}(\phi/2)} e^{j(N-1)\phi/2}, \end{aligned}$$

por lo que

$$u(P, t) = A_0 \frac{\text{sen}(kN\Delta r/2)}{\text{sen}(k\Delta r/2)} e^{-j[k(r + \frac{N-1}{2}\Delta r) - \omega t]}. \quad (1.64)$$

La expresión anterior puede reescribirse como

$$u(P, t) = A_P e^{-j(kR - \omega t)}, \quad (1.65)$$

donde

$$R = r + \frac{N-1}{2}\Delta r$$

es la distancia desde el centro de la rendija al punto P y

$$A_P = A_0 \frac{\text{sen}(kN\Delta r/2)}{\text{sen}(k\Delta r/2)}. \quad (1.66)$$

es la amplitud resultante en P . Dado que la amplitud de la onda en la pantalla varía punto a punto, también variará la intensidad, formando lo que se conoce como un **patrón de difracción**:

$$\frac{I(\theta)}{I_{\max}} = \frac{\text{sen}^2(Nk\Delta r/2)}{\text{sen}^2(k\Delta r/2)}. \quad (1.67)$$

Claramente existe un mínimo en la intensidad de la perturbación cuando $A_P \rightarrow 0$, esto es, cuando el numerador de (1.66) sea cero,

$$\text{sen}(kN\Delta r/2) = 0,$$

es decir, cuando el argumento verifica que

$$kN\Delta r/2 = m\pi. \quad (1.68)$$

Según se puede deducir de la Figura 1.6(b) (si $N \gg 1$):

$$N\Delta r \approx (N-1)\Delta r = b \text{sen} \theta,$$

por lo que la condición de mínimo (1.68) para A_P puede reescribirse como

$$\frac{2\pi}{\lambda} \frac{b \sin \theta}{2} = m\pi, \quad (1.69)$$

o equivalentemente

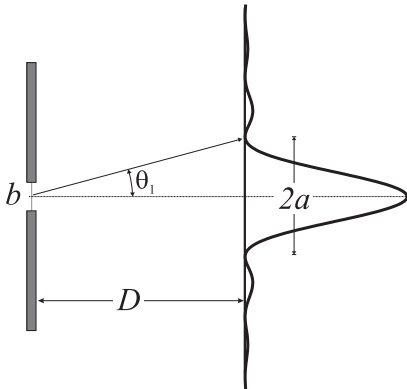
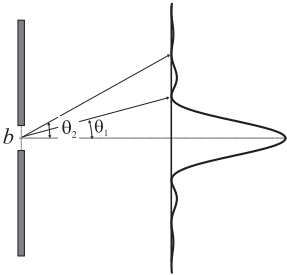
$$b \sin \theta_m = m\lambda \quad m = 1, 2, \dots \quad (1.70)$$

El primer mínimo (o mínimo de primer orden) ocurre para $m = 1$, verificándose entonces que

$$\sin \theta_1 = \frac{\lambda}{b}. \quad (1.71)$$

Puede observarse que si $\lambda \ll b$, $\theta_1 \approx 0$, por lo que apenas se observará patrón de difracción, es decir, la zona de sombra aparece bien definida tal como ocurriría si la onda se propagase en línea recta. A medida que el cociente λ/b crece, el ángulo θ_1 aumenta, haciéndose, por tanto, más evidente el fenómeno de difracción. En general, los fenómenos de difracción son más apreciables cuando las dimensiones de la rendija son del orden de la longitud de onda de la perturbación ondulatoria (no obstante, debe tenerse en cuenta que el análisis efectuado para obtener la expresión (1.70) es sólo válido si $\lambda < b$, puesto que de otro modo el seno sería mayor que uno para todo valor de m).

Condición de intensidad nula en la difracción por una rendija



EJEMPLO 1.4 Hallar la anchura de la franja central del patrón de difracción producido en una pantalla situada a una distancia de 5 m de una rendija de anchura 0.3 mm por la que se ha hecho pasar una luz laser de 600 nm.

La anchura de la franja central puede obtenerse a partir del ángulo θ_1 que nos da el primer mínimo en el patrón de difracción. Según la expresión (1.71), este ángulo viene dado por

$$\sin \theta_1 = \frac{\lambda}{b} = \frac{6 \times 10^{-7} \text{ m}}{3 \times 10^{-4} \text{ m}} = 2 \times 10^{-3}.$$

Dado que $\sin \theta_1 \ll 1$, tenemos que

$$\sin \theta_1 \approx \tan \theta_1$$

y, por tanto, la anchura de la franja central será

$$2a = 2D \tan \theta_1 \approx 2 \cdot 5 \cdot 2 \times 10^{-3} = 20 \text{ mm}.$$

1.8. Grupo de Ondas

Aunque hasta ahora hemos estado tratando con ondas armónicas, debe tenerse en cuenta que una onda armónica es una *idealización* que no puede representar, en sentido estricto, a una perturbación real. Claramente, cualquier perturbación o señal real debe *empezar* y *acabar* en determinados instantes de tiempo. La onda armónica representaría una señal que no empieza ni acaba nunca y que además tiene siempre la misma amplitud, fase y frecuencia. Una señal de este tipo, además de irreal, no transmitiría información alguna puesto que la información está necesariamente asociada a los cambios de alguna magnitud detectable (una conversación puede tener lugar porque se perciben cambios incesantes en la frecuencia y amplitud de la onda sonora).

Una onda que empieza en cierto instante de tiempo y que acaba un cierto tiempo después recibe el nombre de *pulso*. Ahora bien, este pulso no puede estar formado por la superposición de ondas armónicas de *igual frecuencia* y distinta amplitud, puesto que, como se señaló en el Apartado 1.3, la superposición

de ondas armónicas de la misma frecuencia es simplemente otra onda armónica de la misma frecuencia. El pulso puede, sin embargo, construirse sumando ondas armónicas de *distinta frecuencia* y, consecuentemente, de distinto número de ondas. En este sentido denominaremos **grupo de ondas** a un conjunto de ondas armónicas de distinta frecuencia que se propagan superpuestas en la misma dirección. Entre los distintos grupos de ondas existe uno de mucho interés práctico que se conoce como **paquete de ondas**. Un paquete de ondas está formado por la superposición de un conjunto *infinito* de ondas armónicas de amplitud y frecuencia variables con respecto al número de ondas, k , que se propagan superponiéndose a lo largo del eje x . Una de las componentes infinitesimales de este grupo de ondas puede representarse por

$$u_k(x, t) = A_k e^{-j[kx - \omega(k)t]} . \quad (1.72)$$

Si conocemos cómo varía tanto la amplitud como la frecuencia con respecto al número de ondas, podríamos formar el paquete de ondas superponiendo de forma continua ondas del tipo (1.72), es decir,

$$u(x, t) = \int_{-\infty}^{\infty} u_k(x, t) dk = \int_{-\infty}^{\infty} A(k) e^{-j[kx - \omega(k)t]} dk . \quad (1.73)$$

A continuación analizaremos la forma del paquete de ondas en función de la naturaleza dispersiva o no dispersiva del medio en el que se propaga la onda. En general diremos que un **medio no dispersivo** es aquel para el que se cumple que la función $\omega(k)$ es una función *lineal* del número de ondas:

$$\omega = vk , \quad (1.74)$$

Medio NO dispersivo: $\omega = vk$

tal como se dedujo en (1.17) a partir de la ecuación de ondas no dispersiva. Para el medio no dispersivo, al introducir (1.74) en (1.73) encontramos que

$$u(x, t) = \int_{-\infty}^{\infty} A(k) e^{-jk(x-vt)} dk , \quad (1.75)$$

esto es, la integral es una función exclusivamente de $x - vt$. En consecuencia podemos escribir que

$$u(x, t) = f(x - vt) , \quad (1.76)$$

representando así a una perturbación que se propaga *sin distorsión*. Concluimos entonces que en los medios no dispersivos el paquete de ondas se propaga sin distorsión, lo cual es imprescindible en el proceso de transmisión de información para recibir la misma "información" que se envió.

Desafortunadamente, la mayoría de los medios son **medios dispersivos** donde se observa que

- Ondas armónicas de diferente longitud de onda se propagan con diferente velocidad, esto es, $v = v(k)$.
- La forma de la onda se distorsiona cuando se propaga.

Para medios dispersivos, la ecuación de ondas no viene dada por (1.2) sino que en general aparecerán también derivadas de otros órdenes con respecto a x y t . Dado que en estos medios, la velocidad con que se propagan las ondas armónicas depende de la longitud de onda, encontramos que la frecuencia y el número de ondas no estarán relacionados linealmente, sino que debemos escribir ahora

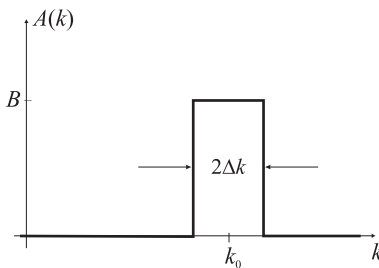
$$\omega(k) = v(k)k . \quad (1.77)$$

Una manera muy usual de reducir al máximo el inconveniente práctico de la deformación de la señal al transmitirse por un medio dispersivo consiste en la construcción de un tipo de paquete de ondas en el que la amplitud de las componentes armónicas de la onda muestre un máximo muy acusado en torno a cierto valor central k_0 y decaiga rápidamente en un entorno $\pm\Delta k$ respecto a k_0 tal que

$$\Delta k \ll k_0. \quad (1.78)$$

Así, si sumamos de forma continua ondas armónicas con un número de ondas k comprendido entre $k_0 - \Delta k \leq k \leq k_0 + \Delta k$, el paquete de ondas resultante puede expresarse como la siguiente integral:

$$u(x, t) = \int_{k_0 - \Delta k}^{k_0 + \Delta k} A(k) e^{-j[kx - \omega(k)t]} dk. \quad (1.79)$$



Si, por ejemplo, suponemos que la dependencia de la amplitud con respecto al número de ondas viene dada por

$$A(k) = \begin{cases} B & \text{si } k_0 - \Delta k \leq k \leq k_0 + \Delta k \\ 0 & \text{en otro caso,} \end{cases}$$

entonces la forma matemática final del paquete de ondas puede obtenerse integrando la siguiente expresión:

$$u(x, t) = B \int_{k_0 - \Delta k}^{k_0 + \Delta k} e^{-j[kx - \omega(k)t]} dk. \quad (1.80)$$

Debido a que $\Delta k \ll k_0$, la función $\omega(k)$ puede desarrollarse en serie de Taylor en torno a k_0 y aproximarla con los dos primeros términos de dicho desarrollo; esto es,

$$\omega(k) \approx \omega(k_0) + \left. \frac{d\omega}{dk} \right|_{k=k_0} (k - k_0). \quad (1.81)$$

Debemos notar que la validez de la aproximación anterior se deteriora cuando el medio sea muy dispersivo o si aumentamos Δk (es decir, en aquellas situaciones en las que no podamos despreciar los términos superiores del desarrollo de Taylor).

Si en (1.81) renombramos

$$\omega_0 = \omega(k_0), \quad (1.82)$$

y

$$v_g = \left. \frac{d\omega}{dk} \right|_{k=k_0}, \quad (1.83)$$

$\omega(k)$ puede entonces expresarse como

$$\omega(k) \approx \omega_0 + (k - k_0) v_g, \quad (1.84)$$

lo que permitiría reescribir

$$\begin{aligned} kx - \omega(k)t &= kx - \omega_0 t - (k - k_0) v_g t \\ &= (k_0 x - \omega_0 t) + (k - k_0)(x - v_g t). \end{aligned} \quad (1.85)$$

Introduciendo ahora (1.85) en (1.80) y realizando el cambio de variable $\alpha = k - k_0$, $u(x, t)$ puede expresarse como

$$u(x, t) = B e^{-j(k_0 x - \omega_0 t)} \int_{-\Delta k}^{\Delta k} e^{-j(x - v_g t)\alpha} d\alpha. \quad (1.86)$$

Si llevamos a cabo la integración en (1.86) encontramos que

$$\int_{-\Delta k}^{\Delta k} e^{-j(x-v_g t)\alpha} d\alpha = \frac{e^{-j(x-v_g t)\alpha}}{-j(x-v_g t)} \Big|_{-\Delta k}^{\Delta k} = \frac{2 \operatorname{sen} [\Delta k(x-v_g t)]}{x-v_g t},$$

lo que nos daría la siguiente expresión final para nuestro paquete de ondas:

$$u(x, t) = 2B \frac{\operatorname{sen} [\Delta k(x-v_g t)]}{x-v_g t} \cos(\omega_0 t - k_0 x). \quad (1.87)$$

Al analizar la forma matemática de la señal anterior, y según nos muestra la Fig. 1.7 (teniendo en cuenta que $\Delta k \ll k_0$), observamos que el paquete de ondas

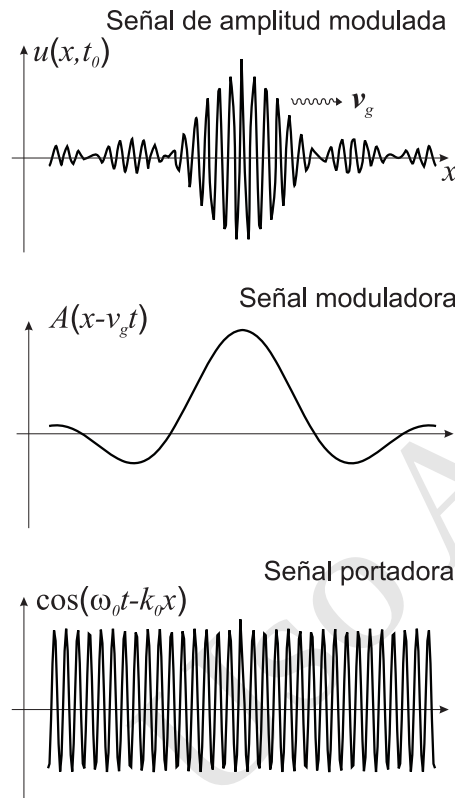


FIGURA 1.7: Descomposición de la señal de amplitud modulada en el producto de una señal moduladora por una portadora.

puede verse como una **señal portadora** armónica de frecuencia ω_0 cuya *amplitud*, $A(\cdot)$, está *modulada*, esto es,

$$u(x, t) = A(x-v_g t) \cos(\omega_0 t - k_0 x). \quad (1.88)$$

La señal resultante puede entonces considerarse como el producto de la señal portadora por la siguiente señal **señal moduladora**:

$$A(x-v_g t) = 2B \frac{\operatorname{sen} [\Delta k(x-v_g t)]}{x-v_g t}, \quad (1.89)$$

que viaja *sin distorsión* apreciable a la velocidad v_g definida en (1.83).

Debemos notar ahora que, en la práctica, lo que se detectaría comúnmente de la señal $u(x, t)$ sería la variación en su amplitud; es decir, la señal $A(x-v_g t)$. Dado

Velocidad de grupo

que esta señal viaja a velocidad v_g , esta velocidad es la que debe identificarse como la “verdadera” velocidad del grupo de ondas, por lo que se conoce como **velocidad de grupo**:

$$v_g = \left. \frac{d\omega}{dk} \right|_{k=k_0} . \quad (1.90)$$

La discusión anterior nos permite observar que una posible manera de transmitir información consistiría en la **modulación** de una señal portadora de número de ondas k_0 y frecuencia ω_0 por una señal moduladora que contiene un rango de frecuencias espaciales Δk en el entorno de k_0 y un rango de frecuencias temporales

$$\Delta\omega = \omega(k_0 + \Delta k) - \omega_0 \quad (1.91)$$

en torno a ω_0 . Notemos que si $\omega(k)$ es una función suave respecto a k , entonces encontraremos que

$$\Delta\omega \ll \omega_0 . \quad (1.92)$$

Al rango de frecuencias $\Delta\omega$ se le denomina usualmente **ancho de banda** en frecuencia de la señal. Esta magnitud es muy importante en la práctica puesto que determina la “cantidad” de información que transmite una señal; por lo tanto, interesa aumentar el ancho de banda lo máximo posible. No obstante, según nos dice la expresión (1.92), para “soportar” señales moduladoras de gran ancho de banda es necesario que la frecuencia de la señal portadora sea bastante mayor que la de la señal moduladora. El interés en obtener señales portadoras con frecuencias lo más altas posibles (y que, consecuentemente, puedan “contener” más información) es precisamente la motivación principal para usar señales luminosas como ondas portadoras. Las señales luminosas tienen una frecuencia del orden de 10^{15} Hz, por lo que permitirían ser moduladas por señales de muy alta frecuencia y con un gran contenido de información.

Es interesante resaltar que *en un medio dispersivo la velocidad de grupo no coincide con la velocidad de fase* ($v_g \neq v$) sino que usando (1.83) y (1.77), la velocidad de grupo puede escribirse en función de la longitud de onda como

$$v_g = v - \lambda \frac{dv}{d\lambda} , \quad (1.93)$$

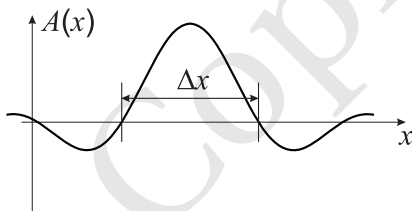
pudiendo ser esta velocidad de grupo mayor o menor que la velocidad de fase (en un medio no dispersivo $v_g = v$). Si el medio es muy dispersivo, la envolvente del grupo de ondas se distorsionará apreciablemente al propagarse. Desde un punto de vista práctico, esta distorsión se puede reducir disminuyendo el ancho de banda de la señal moduladora.

Por último es interesante mencionar que la *anchura espacial* del grupo de ondas, Δx , se define como como la distancia entre los dos primeros mínimos de amplitud para $t = 0$ (si el medio no es muy dispersivo, un instante después, la amplitud habrá viajado una distancia $v_g t$, pero la distancia entre mínimos apenas habrá variado). Los dos primeros mínimos de amplitud ocurren, según (1.87), para $x = \pm\pi/\Delta k$, por lo que

$$\Delta x = \frac{2\pi}{\Delta k}$$

y consecuentemente se llega a la siguiente relación entre la anchura espacial, Δx , y el ancho de banda en k , Δk :

$$\Delta x \Delta k = 2\pi . \quad (1.94)$$



1.9. Ondas en tres dimensiones

Debido a que, en general, el frente de ondas de una onda tridimensional se distorsiona al propagarse (es decir, la onda no es del tipo $u(x, y, z, t) = u(x - vt, y - vt, z - vt)$), la obtención de una ecuación de ondas similar a (1.2) para una perturbación ondulatoria tridimensional del tipo

$$u(\mathbf{r}, t) = u(x, y, z, t)$$

no es trivial. No obstante, existe un tipo muy importante de ondas tridimensionales en el que la propagación se realiza sin distorsión. Este tipo de ondas son las **ondas planas**.

1.9.1. Ondas Planas

Una onda plana es aquella cuyo frente de ondas es un plano, es decir, aquella onda en la que se cumple que todas las superficies que en un instante dado tienen la misma fase son un conjunto de planos. Las ondas planas son muy importantes en la práctica dado que muchos sistemas ópticos producen luz semejante a ondas planas.

La expresión matemática para un plano perpendicular a cierto vector dado \mathbf{k} y que pasa por cierto punto $\mathbf{r}_0 = (x_0, y_0, z_0)$ viene dada por

$$(\mathbf{r} - \mathbf{r}_0) \cdot \mathbf{k} = 0 \quad (1.95)$$

o equivalentemente

$$\mathbf{k} \cdot \mathbf{r} = \text{constante} = a, \quad (1.96)$$

donde \mathbf{r} es la posición de todos los puntos contenidos en dicho plano y $a = \mathbf{k} \cdot \mathbf{r}_0$.

La forma matemática de una onda plana que se propaga a velocidad v sin distorsión debe, entonces, poder escribirse, en analogía con (1.3), como

$$u(\mathbf{r}, t) = u(\hat{\mathbf{k}} \cdot \mathbf{r} - vt), \quad (1.97)$$

donde

$$\hat{\mathbf{k}} = \frac{\mathbf{k}}{|\mathbf{k}|} \quad (1.98)$$

es el vector unitario normal al plano.

Un tipo muy importante en la práctica de onda plana, es la **onda plana armónica** (en temas posteriores se verá que superponiendo muchas de estas ondas planas armónicas se puede formar cualquier tipo de onda tridimensional). Siguiendo un procedimiento análogo al llevado a cabo en el Apartado 1.3, una onda armónica plana que se repite en el espacio cada longitud de onda, λ , puede expresarse como

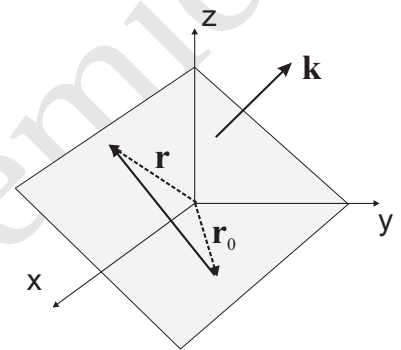
$$u(\mathbf{r}, t) = A \text{sen} \left[\frac{2\pi}{\lambda} (\hat{\mathbf{k}} \cdot \mathbf{r} - vt) \right]. \quad (1.99)$$

Definiendo el *número de ondas* de la onda plana armónica como

$$k = \frac{2\pi}{\lambda}$$

y teniendo en cuenta (1.98) y (1.17) se tiene que la onda plana armónica puede escribirse de forma general usando cosenos en vez de senos como

$$u(\mathbf{r}, t) = A \cos(\omega t - \mathbf{k} \cdot \mathbf{r} - \varphi) = \text{Re} \left\{ A e^{j(\omega t - \mathbf{k} \cdot \mathbf{r} - \varphi)} \right\}, \quad (1.100)$$



o bien simplemente como

$$u(\mathbf{r}, t) = Ae^{-j(\mathbf{k}\cdot\mathbf{r}-\omega t+\varphi)} \quad (1.101)$$

sobreentendiendo que lo que tiene sentido físico es la parte real de la anterior expresión.

Al vector \mathbf{k} se le conoce como **vector de ondas** y es un vector que tiene por módulo el número de ondas y por dirección la dirección de avance de la onda plana

$$\mathbf{k} = k_x\hat{\mathbf{x}} + k_y\hat{\mathbf{y}} + k_z\hat{\mathbf{z}} \quad (1.102)$$

$$k = \sqrt{k_x^2 + k_y^2 + k_z^2} = \frac{2\pi}{\lambda}. \quad (1.103)$$

Teniendo en cuenta (1.102), (1.101) puede reescribirse en forma compleja como (tomando $\varphi = 0$)

$$u(\mathbf{r}, t) = Ae^{-j(k_x x + k_y y + k_z z - \omega t)}. \quad (1.104)$$

1.9.2. Ecuación de onda tridimensional

Para deducir la ecuación de onda que debe satisfacer una perturbación ondulatoria tridimensional operaremos usando la onda plana armónica. Encontraremos qué ecuación diferencial verifica una solución plana armónica y extendemos esta ecuación al caso general de ondas tridimensionales.

La función $u(x, y, z, t)$ dada por (1.104) tiene las siguientes derivadas parciales segundas:

$$\frac{\partial^2 u}{\partial x^2} = -k_x^2 u \quad (1.105a)$$

$$\frac{\partial^2 u}{\partial y^2} = -k_y^2 u \quad (1.105b)$$

$$\frac{\partial^2 u}{\partial z^2} = -k_z^2 u \quad (1.105c)$$

$$\frac{\partial^2 u}{\partial t^2} = -\omega^2 u. \quad (1.105d)$$

Sumando las tres derivadas parciales espaciales y teniendo en cuenta (1.103), se obtiene que

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2} = -k^2 u. \quad (1.106)$$

Combinado esta ecuación con (1.105d) y teniendo en cuenta que $v = \omega/k$, se llega a la siguiente ecuación diferencial:

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2} = \frac{1}{v^2} \frac{\partial^2 u}{\partial t^2}. \quad (1.107)$$

La anterior ecuación puede escribirse de una forma más concisa usando el operador *laplaciano*

$$\nabla^2 \equiv \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}, \quad (1.108)$$

por lo que (1.107) puede expresarse como

$$\nabla^2 u = \frac{1}{v^2} \frac{\partial^2 u}{\partial t^2}. \quad (1.109)$$

Hasta ahora la perturbación ondulatoria se ha considerado como una magnitud escalar, pero exactamente el mismo análisis puede repetirse para cada una de las componentes de la perturbación ondulatoria en el caso de que ésta sea vectorial, en este caso, la perturbación ondulatoria se escribiría como: $\mathbf{u}(\mathbf{r}, t)$ y la ecuación de onda correspondiente sería

$$\nabla^2 \mathbf{u} - \frac{1}{v^2} \frac{\partial^2 \mathbf{u}}{\partial t^2} = 0 . \quad (1.110)$$

Una posible solución de (1.110) sería la onda plana armónica que vendría dada por

$$\mathbf{u}(\mathbf{r}, t) = \mathbf{A}e^{-j(\mathbf{k}\cdot\mathbf{r}-\omega t)} , \quad (1.111)$$

que representa una onda cuya perturbación está dirigida en una dirección constante (la marcada por \mathbf{A}) a la vez que se propaga a lo largo de la dirección \mathbf{k} . Una onda de este tipo se denomina **linealmente polarizada**. En la solución (1.111) es también interesante señalar que sus frentes de onda son planos que se extienden hasta infinito³. La orientación relativa entre la dirección de la perturbación (\mathbf{A}) y la de la propagación (\mathbf{k}) determinarán si la onda puede considerarse como longitudinal o transversal. De este modo si

$$\mathbf{k} \times \mathbf{A} = 0 \Rightarrow \text{Onda longitudinal} \quad (1.112)$$

$$\mathbf{k} \cdot \mathbf{A} = 0 \Rightarrow \text{Onda transversal.} \quad (1.113)$$

1.9.3. Ondas esféricas y cilíndricas

Otras soluciones distintas de la onda plana de bastante interés práctico son las ondas esféricas y cilíndricas.

Las **ondas esféricas** son aquellas en las que su frente de ondas tiene forma esférica. Estas ondas se originarían siempre que un foco emita radial y uniformemente en un medio isótropo, por ejemplo, la luz emitida por una estrella o el sonido emitido por un foco puntual en el aire. Este tipo de ondas puede representarse de forma genérica como

$$u(\mathbf{r}, t) = \frac{f(r \pm vt)}{r} , \quad (1.114)$$

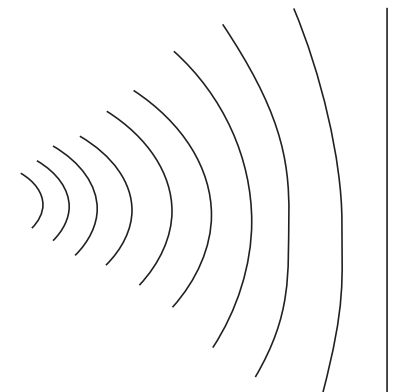
donde v es la velocidad con que se aleja la onda del origen. Un tipo particular de onda esférica es la *onda esférica armónica*,

$$u(r, t) = A(r)e^{-j(kr \pm \omega t)} = \frac{\mathcal{A}}{r} e^{-j(kr \pm \omega t)} , \quad (1.115)$$

donde la constante \mathcal{A} se conoce como *intensidad de la fuente*. Es interesante notar que para la onda esférica armónica, la amplitud de la onda $A(r)$, decrece como r^{-1} , por lo que esta onda se propaga *con distorsión*. La disminución de la amplitud de la onda está relacionada con la conservación de la energía puesto que al aumentar el área de los frentes de onda al avanzar la onda, la energía de la onda debe repartirse cada vez en más área lo que implica que $A(r)$ debe ir disminuyendo proporcionalmente. *Suficientemente lejos de la fuente, una porción del frente de ondas de una onda esférica se asemeja mucho a una porción del frente de ondas de una onda plana* y, por consiguiente en muchos casos prácticos, una onda

Ecuación diferencial de una onda vectorial tridimensional

Onda esférica armónica



³ La perturbación ocuparía todo el espacio por lo que, estrictamente, esta perturbación no describiría ninguna situación real. No obstante, la onda plana armónica será una muy buena aproximación de lo que ocurre en ciertas regiones en muchas situaciones prácticas.

luminosa emitida por un foco puntual se considerará como una onda plana si el foco puntual está suficientemente alejado.

Para las **ondas cilíndricas** (esto es, ondas cuyos frentes de onda son cilindros concéntricos) se encuentra que la forma general de una onda armónica cilíndrica que se aleja del eje z viene dada por

$$u(\mathbf{r}, t) = A(\mathbf{r})e^{-j(k\rho - \omega t)}, \quad (1.116)$$

con $\rho = (x^2 + y^2)^{1/2}$. Para distancias suficientemente lejanas del foco emisor, la perturbación anterior puede aproximarse por

Onda cilíndrica armónica

$$u(\mathbf{r}, t) = \frac{\mathcal{A}}{\sqrt{\rho}} e^{-j(k\rho - \omega t)}. \quad (1.117)$$

1.9.4. Aproximación de Fresnel de la onda esférica: Onda Parabólica

Según la expresión (1.115) para la onda esférica, la parte espacial de una onda esférica saliente viene dada por:

$$U(r) = \frac{\mathcal{A}}{r} e^{-jkr}. \quad (1.118)$$

Para puntos próximos al eje z y alejados del origen ($z \gg$) se tiene que

$$\sqrt{x^2 + y^2} \ll z.$$

Si llamamos

$$\theta^2 = \frac{x^2 + y^2}{z^2}, \quad (1.119)$$

podemos realizar el siguiente desarrollo para r :

$$r = \sqrt{x^2 + y^2 + z^2} = z\sqrt{1 + \theta^2} \approx z\left(1 + \frac{\theta^2}{2} + \frac{\theta^4}{8} + \dots\right). \quad (1.120)$$

Si realizamos ahora las siguientes aproximaciones

$$r \approx z \quad \text{para la amplitud} \quad (1.121)$$

$$r \approx z\left(1 + \frac{\theta^2}{2}\right) \quad \text{para la fase} \quad (1.122)$$

(puesto que la sensibilidad a errores en la fase es mayor que la amplitud dado que la longitud de onda puede ser muy pequeña y consiguientemente k muy grande), obtenemos la aproximación de Fresnel o de **onda parabólica** para una onda esférica:

Onda parabólica

$$U(r) \approx \frac{\mathcal{A}}{z} \exp\left(-jk\frac{x^2 + y^2}{2z}\right) \exp(-jkz). \quad (1.123)$$

Es importante señalar que la aproximación de Fresnel no es válida simplemente cuando $\theta^2 \ll 1$, sino cuando el efecto del tercer término del desarrollo (1.120) en la fase sea despreciable, es decir, cuando

$$kz\frac{\theta^4}{8} \ll \pi \quad (1.124)$$

o, equivalentemente,

$$\frac{2\pi}{\lambda} z \frac{(x^2 + y^2)^2}{8z^4} \ll \pi \Rightarrow (x^2 + y^2)^2 \ll 4\lambda z^3. \quad (1.125)$$

En otras palabras, la aproximación de Fresnel es válida para puntos cercanos al eje z y alejados una distancia $a = \sqrt{x^2 + y^2}$ del origen que verifiquen

$$\frac{N_F \theta_m^2}{4} \ll 1, \quad (1.126)$$

Condición de validez de la aproximación de Fresnel

donde

$$\theta_m = \frac{a}{z} \quad (1.127)$$

es el ángulo máximo y

$$N_F = \frac{a^2}{\lambda z} \quad (1.128)$$

se denomina *número de Fresnel*.

La onda parabólica (obtenida bajo las condiciones de la aproximación de Fresnel) puede aproximarse por una onda plana bajo la denominada **aproximación de Fraunhofer**. Esta aproximación se cumple para para puntos muy alejados del origen ($z \gg$) y cerca del eje z de modo que se verifique que

$$N_F \rightarrow 0. \quad (1.129)$$

Condición de validez de la aproximación de Fraunhofer

Bajo esta condición, la fase de la onda parabólica se aproxima a kz mientras que la amplitud permanece prácticamente constante. En este caso, tenemos que la onda esférica se puede aproximar por una onda plana del tipo

$$U(r) \simeq A \exp(-jkz). \quad (1.130)$$

EJEMPLO 1.5 Evaluar en qué puntos del eje z , en torno a un radio $a = 2.5\text{cm}$, una onda esférica de longitud de onda $\lambda = 600\text{nm}$ puede considerarse una onda parabólica y dónde una onda plana

▪ Onda parabólica:

Una onda esférica puede aproximarse por una onda parabólica en aquellos puntos que verifiquen la condición de Fresnel (1.126), o equivalentemente cuando

$$z \gg \left(\frac{a^4}{4\lambda} \right)^{1/3} = \left(\frac{(2.5 \times 10^{-2})^4}{4 \times 6 \times 10^{-7}} \right)^{1/3},$$

esto es, cuando

$$z \gg 0.5\text{m}$$

▪ Onda plana:

La condición de onda plana (1.129) se satisfecerá cuando

$$z \gg \frac{a^2}{\lambda} = \frac{(2.5 \times 10^{-2})^2}{6 \times 10^{-7}}$$

es decir, cuando

$$z \gg 1000\text{m}.$$

1.10. Problemas propuestos

1.1: El rango de frecuencias de las ondas electromagnéticas utilizadas en emisoras comerciales (radio y TV) se extiende, aproximadamente, desde 10^4 Hz hasta 10^9 Hz. Si la velocidad de propagación de las ondas electromagnéticas en vacío es $c = 3 \times 10^8$ m/s, ¿cuál es el rango de longitudes de onda de las emisoras comerciales? Repítase el problema anterior para el rango de señales acústicas audibles que se extiende, aproximadamente, desde 20 Hz hasta 20 kHz, considerando que el sonido es una onda de presión que se propaga en el aire a una velocidad aproximada de 340 m/s.

Sol.: Ondas electromagnéticas comerciales: λ entre 30.000 m y 30 cm. Ondas sonoras audibles: λ entre 17 m y 17 mm.

1.2: Demuestre explícitamente que el pulso descrito por la función

$$y(x, t) = y_0 \exp \left[- \left(\frac{x - vt}{x_0} \right)^2 \right]$$

es solución de la ecuación de onda.

1.3: El cable de un telesquí de 80 kg de masa asciende 400 m por la ladera de una montaña. Cuando el cable recibe un golpe transversal en un extremo, el pulso de retorno se detecta 12 segundos después. Calcular cuál es la velocidad de la onda y la tensión, T , en el cable. [Nota: la velocidad de la onda en la cuerda viene dada por $v = \sqrt{\mu/T}$, siendo μ la densidad lineal de masa de la cuerda y T su tensión.]

Sol. $v = 66,7 \text{ m/s}$, $T = 889 \text{ N}$.

1.4: Dos pulsos de onda que se mueven sobre una cuerda están representados por las siguientes funciones de onda:

$$y_1(x, t) = \frac{0.02}{2 + (x - 2t)^2}$$

$$y_2(x, t) = \frac{-0.02}{2 + (x + 2t)^2},$$

donde x e y están en metros y t en segundos. **a)** Dibujar por separado cada onda en función de x para $t = 0$ y describir el comportamiento de ambas al aumentar en el tiempo. **b)** Hallar la función de onda resultante para $t = 0$ y $t = 1 \text{ s}$. **d)** Dibujar la función de onda resultante en este último caso.

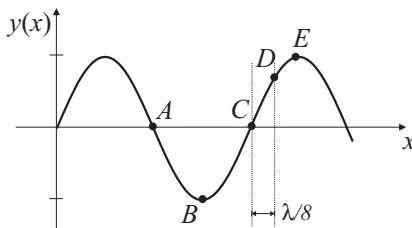
Sol. $y(x, 0) = 0$, $y(x, 1) = \frac{0.16x}{36 - 4x^2 + x^4}$

1.5: Una onda armónica se mueve a lo largo de una cuerda uniforme e infinita bajo tensión constante. La cuerda esta marcada a intervalos de 1 m. En la marca de 0 m, se observa que la cuerda alcanza su desplazamiento transversal máximo de 50 cm cada 5 s. La distancia entre máximos en un instante de tiempo cualquiera es 50 m. Encontrar la expresión de su función de onda, suponiendo que es armónica, que tiene su desplazamiento máximo en $x = 0$ cuando $t = 0$, y que está moviendo a lo largo de la cuerda de izquierda a derecha.

Sol. $y = 0.5 \cos(2\pi/50)(x - 10t) \text{ m}$.

1.6: La onda armónica en el hilo que se muestra en la figura tiene una amplitud de 25mm, una velocidad de 46m/s, una frecuencia angular de 160 rad/s y se propaga hacia $+x$. Determinar **a)** las componentes de velocidad; **b)** las componentes de aceleración de los elementos marcados por letras; **c)** la pendiente del hilo y **d)** $\partial^2 y / \partial x^2$ en los elementos marcados por letras.

Sol.: **a)** 4,0, -4, $-2\sqrt{2}, 0$ (m/s); **b)** 0,640, 0, $-320\sqrt{2}, -640$ (m/s²); **c)** -87,0,87,62,0; **d)** 0,0,3,0, -0,21, -0,3 m⁻¹.



1.7: El desplazamiento transversal de una cuerda viene dado por una onda armónica de ecuación $y(x, t) = 0.001 \sin(62.8x + 314t)$, estando x e y expresadas en metros y t en segundos. **a)** ¿En qué sentido se desplaza la onda? **b)** Hallar la longitud de onda, frecuencia y periodo de la misma. **c)** ¿Cuál es el máximo desplazamiento de un segmento cualquiera de la cuerda?

Sol.: **a)** hacia las x negativas; **b)** $\lambda = 0.1 \text{ m}$, $f = 50 \text{ Hz}$, $T = 0.02 \text{ s}$; **c)** 1 mm.

1.8: Dos ondas de la misma amplitud ($A = 46 \text{ mm}$), frecuencia, velocidad y dirección de propagación, pero desfasadas entre sí, están presentes en un hilo, siendo la amplitud resultante 31 mm. ¿Cuál es la diferencia de fase entre ambas ondas?

Sol.: 140° .

1.9: Hallar la onda resultante de la superposición de tres ondas con la misma frecuencia, longitud de onda y amplitud, que se mueven en la misma dirección y sentido y que vienen dadas por

$$y_1(x, t) = 0.05 \sin(kx - \omega t - \frac{\pi}{3})$$

$$y_2(x, t) = 0.05 \sin(kx - \omega t)$$

$$y_3(x, t) = 0.05 \sin(kx - \omega t + \frac{\pi}{3}).$$

Sol. $y(x, t) = 1.5A \sin(kx - \omega t)$.

1.10: Dos fuentes sonoras, accionadas en fase por el mismo amplificador están distantes 2m entre sí en el eje y . En un punto a distancia muy grande del eje, se oye una interferencia constructiva para el ángulo $\theta_1 = 8^\circ$ y la siguiente se oye a $\theta_1 = 16^\circ 10'$ respecto al eje x . Si la velocidad del sonido es de 340 m/s, **a)** ¿cuál es la longitud de onda de las ondas sonoras procedentes de las fuentes; **b)** ¿cuál es la frecuencia de las mismas?; **c)** ¿para qué otros ángulos se oye la interferencia constructiva?; **d)** ¿cuál es el ángulo más pequeño para el cual se contrarrestan entre sí por completo las ondas sonoras?

Sol.: **a)** 0.278 m; **b)** 1.22 kHz; **c)** $24.7^\circ, 33.8^\circ, 44.1^\circ, 56.6^\circ, 77^\circ$; **d)** 4° .

1.11: Dos ondas armónicas planas,

$$y_1(x, t) = A \cos(\omega t - kx + \phi_1)$$

$$y_2(x, t) = A \cos(\omega t + kx + \phi_2)$$

interfieren dando lugar a una onda estacionaria. Determínese: **a)** la expresión de la onda estacionaria resultante; **b)** las posiciones de los planos nodales y ventrales en dos casos: **b.1)** $\phi_1 = \phi_2 = 0$ y **b.2)** $\phi_1 = 0, \phi_2 = -\pi$.

Sol. a) $y(x, t) = y_1(x, t) + y_2(x, t) = 2A \cos[kx + (\phi_1 + \phi_2)/2] \cos[\omega t + (\phi_2 - \phi_1)/2]$;

b.1) nodos: $x = (2n + 1)\lambda/4$, vientres: $x = n\lambda/2$ (n entero);

b.2) nodos: $x = n\lambda/2$, vientres: $x = (2n + 1)\lambda/4$ (n entero).

1.12: En una columna de aire de longitud $L = 2$ m, abierta en sus dos extremos, se propagan ondas sonoras del tipo

$$u(x, t) = C \cos(kx) \cos(\omega t)$$

(los extremos de la columna representan vientres para la onda sonora). Sabiendo que la frecuencia de su n -ésima armónica es 410 Hz y la frecuencia de su $(n + 1)$ armónica es 492 Hz, determínese la velocidad del sonido en el aire bajo estas condiciones.

Sol.: $v = 328$ m/s.

1.13: Dos focos S_1 y S_2 emiten ondas esféricas de igual frecuencia:

$$y_1(r_1, t) = (A_1/r_1) \cos(\omega t - kr_1)$$

$$y_2(r_2, t) = (A_2/r_2) \cos(\omega t - kr_2)$$

respectivamente, donde r_1 y r_2 indican la distancia desde el punto considerado hasta los focos S_1 y S_2 respectivamente. Calcúlese: **a)** la amplitud de la señal resultante en cada punto; **b)** los puntos en los cuales la amplitud resultante es máxima y en los que es mínima (expresese el resultado en términos de la distancia de los puntos a los focos y de la longitud de onda).

Sol. a) $A(r_1, r_2) = \{(A_1/r_1)^2 + (A_2/r_2)^2 + 2A_1A_2 \cos[k(r_1 - r_2)]/(r_1r_2)\}^{1/2}$; **b)** puntos de amplitud máxima: $r_1 - r_2 = n\lambda$, puntos de amplitud mínima: $r_1 - r_2 = (2n + 1)\lambda/2$.

1.14: Un haz de luz monocromático que puede asimilarse a una onda plana que incide sobre una rendija larga de anchura a perpendicular a su dirección de propagación. **a)** Obtener la expresión para la anchura del máximo central de difracción en función de la longitud de onda y de la distancia, D , de la pantalla de detección a la rendija. **b)** Utilizar el resultado anterior para encontrar la anchura del máximo central cuando la luz de una lámpara de sodio ($\lambda = 590$ nm) es difractada por una rendija de anchura 0.3 mm, siendo la distancia de la pantalla a la rendija de 0.87 m.

Sol. a) Anchura $= 2\lambda D/\sqrt{a^2 - \lambda^2}$. **b)** 3.4 mm.

1.15: Dos ondas armónicas planas que se propagan en *igual* sentido interfieren dando lugar a una onda plana que también se propaga en dicho sentido. Si las ecuaciones de dichas ondas son

$$y_1(x, t) = A \sin(kx - \omega t + \phi_1)$$

$$y_2(x, t) = A \sin(kx - \omega t + \phi_2),$$

obtégase la ecuación de la onda resultante de la interferencia, indicando en qué casos la amplitud resultante es máxima y en cuáles es mínima.

Sol.: $y(x, t) = y_1(x, t) + y_2(x, t) = 2A \cos(\Delta\phi/2) \sin[kx - \omega t + (\phi_1 + \phi_2)/2]$, siendo $\Delta\phi = \phi_2 - \phi_1$; Amplitud máxima: $\Delta\phi = 2n\pi$, mínima: $\Delta\phi = (2n + 1)\pi$.

1.16: Dadas dos ondas sonoras armónicas de *distinta* frecuencia:

$$u_1(x, t) = A \cos[(\omega_0 - \Delta\omega)t - (k_0 - \Delta k)x]$$

$$u_2(x, t) = A \cos[(\omega_0 + \Delta\omega)t - (k_0 + \Delta k)x],$$

a) obtenga la expresión de la superposición de dichas ondas. **b)** Si $\Delta k \ll k_0$ y $\Delta\omega \ll \omega_0$, represente gráficamente la onda. Si las ondas se propagan en un medio dispersivo determinado por $\omega(k) = v_m \sqrt{k^2 - k_c^2}$, siendo $v_m = 200$ m/s, $\omega_0 = 2\pi \times 10^4$ rad/s, $\Delta\omega = 200\pi$ Hz y $k_c = 100$ m⁻¹, **c)** calcule los valores de la velocidad de fase, v_f , y velocidad de grupo, $v_g = (d\omega/dk)_{k=k_0}$, de la onda superposición.

Sol. a) $u(x, t) = 2A \cos(\Delta\omega t - \Delta kx) \cos(\omega_0 t - k_0x)$; **c)** $v_f \approx 190.6$ m/s, $v_g \approx 209.9$ m/s.

1.17: Determine el radio del círculo dentro del cual se puede aproximar por una onda parabólica una onda esférica de longitud de onda $\lambda = 633$ nm, cuyo foco está a 1 m de distancia. Determinar el ángulo máximo θ y el número de Fresnel N_F .

Sol.: $\theta \ll 0.04$ rad; $N_F \ll 2527$.

1.18: Una onda plana descrita por $A_1 \exp(-jkz)$ y una onda esférica aproximada por la onda parabólica $\frac{A_2}{z} e^{-jkz} \exp(-jk \frac{x^2 + y^2}{2z})$ interfieren en el plano $z = d$. Obtenga una expresión para la intensidad total $I(x, y, d)$. Verifique que el lugar geométrico de los puntos de intensidad nula es un conjunto de anillos concéntricos.

Sol.: $I = I_1 + I_2 + 2\sqrt{I_1 I_2} \cos(\frac{x^2 + y^2}{2d})$.

1.19: Dos ondas esféricas de igual intensidad I_0 , originadas en los puntos $(a, 0, 0)$ y $(-a, 0, 0)$ interfieren en el plano $z = d$. El sistema es similar al empleado en el experimento de la doble rendija de Young. Utilice la aproximación parabólica para las ondas a fin de demostrar que la intensidad medida es

$$I(x, y, d) = 2I_0 \left(1 + \cos \frac{2\pi x \theta}{\lambda} \right),$$

donde $\theta = 2a/d$ es aproximadamente el ángulo subtendido por los centros de las dos ondas en el plano de observación. El patrón de intensidad es periódico de periodo λ/θ .

Copia Uso Académico

Tema 2

Ondas Electromagnéticas

2.1. Introducción

En el tema anterior se ha discutido que la característica principal de una onda es su capacidad de transmitir una energía (junto con momento lineal/angular) sin que ello lleve aparejado un transporte neto de materia. Usualmente, las ondas consisten en la propagación de alguna perturbación física a través de algún *medio material*, por ejemplo: olas en el agua, variaciones de presión en el aire (sonido), etc. No obstante, existe un tipo de fenómeno ondulatorio que no requiere la presencia de medios materiales para su propagación (es decir, la perturbación se puede propagar en el vacío) aunque ciertamente también puede propagarse a través de medios materiales. Estas ondas son las **ondas electromagnéticas**, que consisten en la transmisión de un campo electromagnético a una velocidad $v \leq c$, siendo c la velocidad de propagación en el vacío: $c \approx 3 \times 10^8 \text{ ms}^{-1}$. Históricamente la existencia de estas ondas surgió como una consecuencia teórica posible y sorprendente de las ecuaciones de Maxwell. La combinación de las leyes de Faraday–Maxwell y Ampère–Maxwell implicaba que un campo magnético “primario” variable en el tiempo, $\mathbf{B}_1(\mathbf{r}, t)$, podía ser la fuente de un campo eléctrico variable en el tiempo, $\mathbf{E}_1(\mathbf{r}, t)$, y podía ser éste a su vez la fuente de un campo magnético “secundario” variable en el tiempo, $\mathbf{B}_2(\mathbf{r}, t)$, y así sucesivamente ...¹:

$$\mathbf{B}_1(\mathbf{r}, t) \Rightarrow \mathbf{E}_1(\mathbf{r}, t) \Rightarrow \mathbf{B}_2(\mathbf{r}, t) \Rightarrow \mathbf{E}_2(\mathbf{r}, t) \Rightarrow \mathbf{B}_3(\mathbf{r}, t) \Rightarrow \dots$$

De este modo, los campos eléctrico y magnético se generan mutuamente dando lugar a una onda electromagnética que se propaga en el espacio libre (vacío) a una velocidad $c = 1/\sqrt{\mu_0\epsilon_0} \approx 3 \times 10^8 \text{ m/s}$. Esta hipótesis teórica deducida por Maxwell fue confirmada experimentalmente en 1888 por H. Hertz.

Las ondas electromagnéticas, además de constituir uno de los fenómenos físicos más predominantes en la Naturaleza, tienen una importancia fundamental en el campo de las comunicaciones y en la Fotónica. Podría decirse que la mayoría de las comunicaciones actuales se sustentan en la transmisión de ondas electromagnéticas, ya sea a través del espacio libre (radio, televisión, teléfonos móviles,...) o bien a través de medios materiales (telefonía convencional, televisión por cable, transmisión por fibra óptica, comunicación entre ordenadores, ...). Existen muchas razones para justificar este extendido uso pero, entre otras, cabe destacar:

¹ Evidentemente si el campo *primario* fuese uno eléctrico, en vez de uno magnético, también se produciría una onda electromagnética.

- la capacidad de las ondas electromagnéticas de propagarse en el vacío;
- el desarrollo de antenas (emisoras y receptoras) eficaces que permiten la transmisión y recepción de estas ondas involucrando muy poca energía;
- la posibilidad de “guiar” estas ondas mediante diversos sistemas de guiado: línea bifilar, cable coaxial, guías de ondas metálicas, fibras ópticas, etc;
- el hecho de poder usar señales portadoras de muy alta frecuencia (hasta frecuencias ópticas) que permiten grandes anchos de banda;
- la facilidad de tratamiento de las señales electromagnéticas, por ejemplo su modulación/demodulación en fase, amplitud, frecuencia... que permite usar estas señales como soporte de información tanto analógica como digital;
- la fácil integración de los equipos de generación/recepción con la circuitería electrónica.

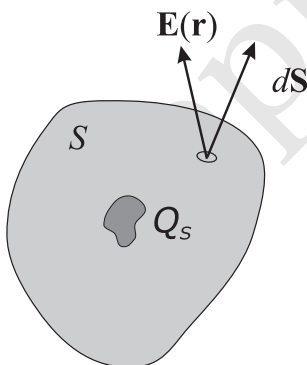
Además dado que lo que denominamos luz no es más que ondas electromagnéticas de un cierto rango de frecuencias, el estudio de la propagación de la luz y de la interacción de la luz con la materia no podría llevarse a cabo sin una comprensión previa de las propias ondas electromagnéticas.

2.2. Ecuaciones de Maxwell

Tal como se ha comentado anteriormente, la existencia de ondas electromagnéticas surgió en primer lugar como una consecuencia teórica de las ecuaciones de Maxwell. Por esta razón, a continuación dedicaremos nuestra atención a estas ecuaciones.

2.2.1. Introducción

En cursos previos de Física se han visto una serie de leyes extraídas de la experiencia y que determinaban el comportamiento del campo eléctrico y del magnético. Entre las múltiples leyes y expresiones que se han visto, podemos extraer una serie de ellas que forman la *base* del Electromagnetismo y que explican todos los fenómenos electromagnéticos. Estas leyes fueron recogidas por James C. Maxwell (1831-1879) en una labor que ha sido reconocida como una de las síntesis más fecundas de toda la historia de la Física. Las ecuaciones de Maxwell son cuatro ecuaciones diferenciales o integro-diferenciales que compendian toda la información experimental adquirida sobre los campos eléctricos y magnéticos.



2.2.2. Antecedentes

Un posible compendio de las leyes básicas que regían el comportamiento del campo eléctrico y magnético antes de la síntesis de Maxwell se muestra a continuación:

1. Ley de Gauss para el campo electrostático.

Esta ley dice que el flujo del campo electrostático, $\mathbf{E}(\mathbf{r})$, a través de una superficie cerrada, S , es igual a $1/\epsilon_0$ veces la carga *total* encerrada en el interior de dicha superficie:

$$\oint_S \mathbf{E}(\mathbf{r}) \cdot d\mathbf{S} = \frac{Q_S}{\epsilon_0}, \quad (2.1)$$

donde Q_S es la carga encerrada en el interior de la superficie S y

$$\epsilon_0 = 8.85 \times 10^{-12} \text{ C}^2/(\text{Nm}^2) \quad (2.2)$$

es una constante que se conoce como *permitividad eléctrica del vacío*.

2. Ley de Gauss para el campo magnetostático.

Dado que las líneas del campo *magnetostático*, $\mathbf{B}(\mathbf{r})$, no tienen principio ni fin, este hecho puede expresarse matemáticamente como

$$\oint_S \mathbf{B}(\mathbf{r}) \cdot d\mathbf{S} = 0. \quad (2.3)$$

Esta ley establece que el flujo del campo magnetostático a través de cualquier superficie cerrada es nulo. No es difícil imaginar que si las líneas de campo son líneas que no tienen principio ni fin, el número de líneas que entran en cualquier superficie cerrada serán las mismas que las que salen y, por tanto, el flujo total es nulo.

3. Ley de Faraday.

La ley de inducción electromagnética establece que la fuerza electromotriz (expresada comúnmente como fem, ξ) inducida en un circuito es igual a menos la variación temporal del flujo magnético, Φ_m , que atraviesa dicho circuito:

$$\xi = -\frac{d\Phi_m}{dt}. \quad (2.4)$$

En términos de los campos eléctrico y magnético, esta ley puede reescribirse como

$$\oint_{\Gamma} \mathbf{E}(\mathbf{r}, t) \cdot d\mathbf{l} = -\frac{d}{dt} \int_{S(\Gamma)} \mathbf{B}(\mathbf{r}, t) \cdot d\mathbf{S}, \quad (2.5)$$

es decir, la circulación del campo eléctrico, $\mathbf{E}(\mathbf{r}, t)$ (campo no electrostático), a lo largo del "circuito" Γ es igual a menos la variación del flujo magnético del campo $\mathbf{B}(\mathbf{r}, t)$ que atravesaba dicho circuito. Es importante notar que la ley de Faraday siempre hace referencia al medio material (conductores del circuito) como el lugar donde se genera el campo no electrostático que provoca la fem inducida.

4. Ley de Ampère para el campo magnetostático.

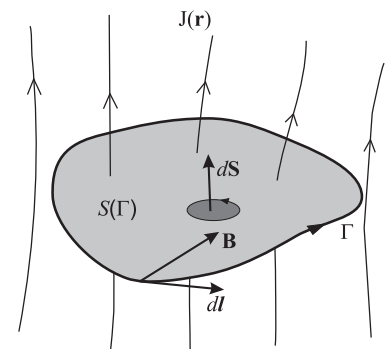
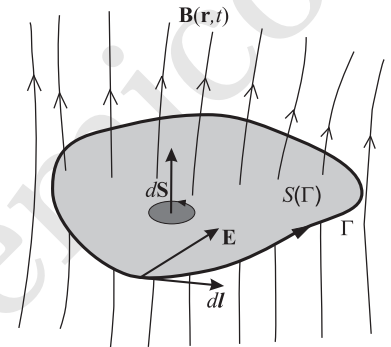
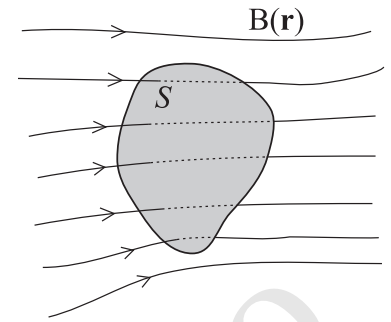
La ley básica que determina la forma del campo *magnetostático*, $\mathbf{B}(\mathbf{r})$ es la ley de Ampère, la cual establece que

$$\oint_{\Gamma} \mathbf{B}(\mathbf{r}) \cdot d\mathbf{l} = \mu_0 \int_{S(\Gamma)} \mathbf{J}(\mathbf{r}) \cdot d\mathbf{S}, \quad (2.6)$$

es decir, la circulación del campo magnetostático a través de una curva Γ es igual a μ_0 veces la intensidad de la corriente *total* que atraviesa la superficie $S(\Gamma)$. La constante

$$\mu_0 = 4\pi \times 10^{-7} \text{ N/A}^2 \quad (2.7)$$

se denomina *permeabilidad magnética del vacío*.



2.2.3. Aportaciones de Maxwell

Maxwell realiza una revisión de las leyes expuestas en el apartado anterior, extendiéndolas en general al caso de campos eléctricos y magnéticos que varían en el tiempo. Sus aportaciones pueden resumirse en lo siguiente:

Ley de Gauss para el campo eléctrico

1. Ley de Gauss para el campo eléctrico

Maxwell extendió la validez de la ley de Gauss (que en su forma inicial (2.1) sólo era aplicable a campos *electrostáticos*, $\mathbf{E} = \mathbf{E}(\mathbf{r})$) a campos eléctricos que varían en el tiempo, $\mathbf{E} = \mathbf{E}(\mathbf{r}, t)$, de modo que ésta puede escribirse como

$$\oint_S \mathbf{E}(\mathbf{r}, t) \cdot d\mathbf{S} = \frac{Q_S(t)}{\epsilon_0}, \quad (2.8)$$

donde $Q_S(t)$ es la carga *total* (que ahora puede variar en el tiempo) encerrada en el interior de la superficie S .

2. Ley de Gauss para el campo magnético.

Al igual que para la anterior ley, Maxwell extiende la validez de la ley (2.3) a campos magnéticos variables en el tiempo, $\mathbf{B} = \mathbf{B}(\mathbf{r}, t)$; esto es,

Ley de Gauss para el campo magnético

$$\oint_S \mathbf{B}(\mathbf{r}, t) \cdot d\mathbf{S} = 0. \quad (2.9)$$

Esta ley describe la observación experimental de que las líneas de campo magnético no divergen ni convergen en ningún punto del espacio; es decir, que no existen cargas magnéticas.

3. Ley de Faraday-Maxwell.

La ley de inducción electromagnética según fue establecida por Faraday estaba directamente ligada a la presencia de conductores, de modo que en la expresión (2.5), la curva Γ coincidía estrictamente con el recorrido del circuito. Maxwell notó que la identidad matemática expresada por (2.5) no tenía por qué ligarse a la existencia de conductores puesto que no hay nada en (2.5) que exija que la curva Γ deba coincidir con el recorrido del circuito. Con esta concepción en mente, la ley de Faraday-Maxwell,

Ley de Faraday-Maxwell

$$\oint_{\Gamma} \mathbf{E}(\mathbf{r}, t) \cdot d\mathbf{l} = - \int_{S(\Gamma)} \frac{\partial \mathbf{B}(\mathbf{r}, t)}{\partial t} \cdot d\mathbf{S} \quad (2.10)$$

establece que la circulación del campo eléctrico a través de *una curva arbitraria*, Γ , es igual a menos la variación del flujo magnético que atraviesa una superficie $S(\Gamma)$ cuyo contorno se apoya en Γ . Esta reinterpretación de la ley de Faraday dice mucho más que la ley original pues establece

la existencia de un campo eléctrico en cualquier punto del espacio donde exista un campo magnético variable en el tiempo.

4. Ley de Ampère-Maxwell.

La ley de Ampère tal como se escribió en (2.6) sólo era válida, en principio, para campos magnetostáticos. Para generalizar la ley de Ampère es tentador seguir el mismo procedimiento que permitió extender las leyes de Gauss (inicialmente formuladas para campos estáticos) a campos variables en el tiempo. En este sentido, podríamos realizar la siguiente extensión de la ley de Ampère:

$$\oint_{\Gamma} \mathbf{B}(\mathbf{r}, t) \cdot d\mathbf{l} \stackrel{?}{=} \mu_0 \int_{S(\Gamma)} \mathbf{J}(\mathbf{r}, t) \cdot d\mathbf{S}, \quad (2.11)$$

Para comprobar la validez de la expresión (2.11) basta considerar el proceso de carga de un conductor recorrido por una intensidad $I(t)$, donde la curva Γ rodea al conductor y la superficie $S(\Gamma)$ es tal como se muestra en la figura. Al tomar el límite cuando la curva Γ se hace cero, la superficie $S(\Gamma)$ cierra al conductor y se tiene que

$$\lim_{\Gamma \rightarrow 0} \oint_{\Gamma} \mathbf{B}(\mathbf{r}, t) \cdot d\mathbf{l} = 0, \quad (2.12)$$

puesto que el valor del campo magnético en los puntos de la curva Γ tiende a cero² en el límite $\Gamma \rightarrow 0$. Ahora bien, supuesta cierta (2.11), la expresión (2.12) también implicaría que

$$\lim_{\Gamma \rightarrow 0} \int_{S(\Gamma)} \mathbf{J}(\mathbf{r}, t) \cdot d\mathbf{S} = \oint_S \mathbf{J}(\mathbf{r}, t) \cdot d\mathbf{S} = 0, \quad (2.14)$$

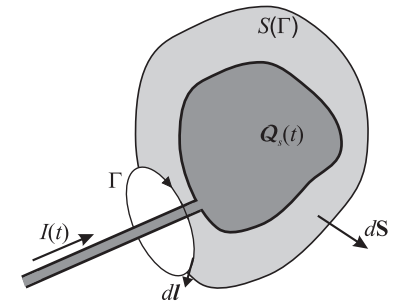
lo cual es claramente incorrecto como puede comprobarse al considerar que, efectivamente, entra una intensidad $I(t)$ en la superficie *cerrada* S ; es decir,

$$\oint_S \mathbf{J}(\mathbf{r}, t) \cdot d\mathbf{S} = I(t).$$

Si está entrando una intensidad en la superficie S , encontramos igualmente que la carga en el interior de esta superficie estaría variando en el tiempo.

Si se tiene en cuenta la ecuación de continuidad de la carga dada por

$$\oint_S \mathbf{J}(\mathbf{r}, t) \cdot d\mathbf{S} = -\frac{d}{dt} Q_S(t), \quad (2.15)$$



Ecuación de continuidad de carga

que dice que la variación de la carga encerrada en una superficie cerrada S , $Q_S(t)$, en la unidad de tiempo es igual al flujo total de densidad de corriente que atraviesa dicha superficie, observamos una clara contradicción entre lo que dice la ecuación de continuidad de la carga (2.15) y la expresión (2.14) derivada directamente de la ley de Ampère. Dado que no cabe discusión acerca de la validez de la ecuación de continuidad de la carga (ésta no es más que la expresión local del *principio de conservación de la carga*), tenemos que concluir que la ley de Ampère, tal y como se expresó en (2.6), no es válida para situaciones en las que las magnitudes implicadas varían en el tiempo. Es razonable pensar que la forma en la que esta ley debe modificarse debe ser en el sentido de hacerla compatible con la ecuación de continuidad de la carga. Así, si consideramos la expresión de la ley de Gauss dada en (2.8), la ecuación de continuidad de la carga puede reescribirse como

$$\begin{aligned} \oint_S \mathbf{J}(\mathbf{r}, t) \cdot d\mathbf{S} &= -\frac{d}{dt} \left(\oint_S \epsilon_0 \mathbf{E}(\mathbf{r}, t) \cdot d\mathbf{S} \right) \\ &= -\oint_S \epsilon_0 \frac{\partial \mathbf{E}(\mathbf{r}, t)}{\partial t} \cdot d\mathbf{S}, \end{aligned} \quad (2.16)$$

o bien

$$\oint_S \left[\mathbf{J}(\mathbf{r}, t) + \epsilon_0 \frac{\partial \mathbf{E}(\mathbf{r}, t)}{\partial t} \right] \cdot d\mathbf{S} = 0. \quad (2.17)$$

A la vista de la expresión anterior, es claro que reescribiendo la ecuación de Ampère en la forma siguiente:

$$\oint_{\Gamma} \mathbf{B}(\mathbf{r}, t) \cdot d\mathbf{l} = \mu_0 \int_{S(\Gamma)} \left[\mathbf{J}(\mathbf{r}, t) + \epsilon_0 \frac{\partial \mathbf{E}(\mathbf{r}, t)}{\partial t} \right] \cdot d\mathbf{S} \quad (2.18)$$

Ley de Ampère-Maxwell

y siguiendo el mismo procedimiento de paso al límite de la curva Γ , entonces esta ley es ya congruente con la ecuación de continuidad de la carga.

En el segundo miembro de (2.18) aparecen dos términos de densidad de corriente, a saber:

² Recuérdese que el campo magnetostático en el interior de un conductor cilíndrico rectilíneo de radio R recorrido por una intensidad I venía dado por

$$\mathbf{B}(\mathbf{r}) = \frac{\mu_0 I}{2\pi R^2} r \hat{\mathbf{r}}. \quad (2.13)$$

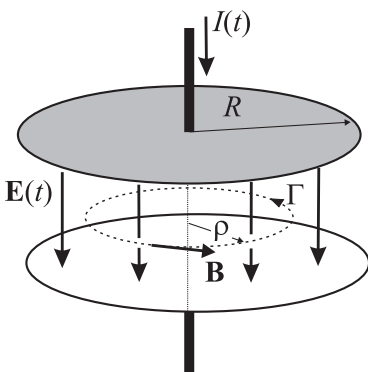
- **la densidad de corriente de cargas: \mathbf{J}_C ,** que es aquella relacionada directamente con el movimiento de las cargas; es decir, la que aparece en la ley de Ohm.
- **la densidad de corriente de desplazamiento: $\mathbf{J}_D = \epsilon_0 \frac{\partial \mathbf{E}}{\partial t}$,** que es un término de corriente que no está directamente relacionado con el movimiento de cargas sino que está asociado exclusivamente a las variaciones temporales del campo eléctrico. Su origen podemos explorarlo en el paso de la ecuación (2.15) a (2.17), donde podemos identificar este término como debido a la existencia de una carga variable en el tiempo.

Es interesante hacer notar que en el caso de que no haya corriente de cargas, la ley de Ampère–Maxwell se escribe en el vacío como

$$\oint_{\Gamma} \mathbf{B}(\mathbf{r}, t) \cdot d\mathbf{l} = \mu_0 \epsilon_0 \int_{S(\Gamma)} \frac{\partial \mathbf{E}(\mathbf{r}, t)}{\partial t} \cdot d\mathbf{S}. \quad (2.19)$$

Esta ecuación es análoga a la ecuación (2.10) y establece

la existencia de un campo magnético asociado a la existencia de una campo eléctrico variable en el tiempo.



EJEMPLO 2.1 Cálculo del campo magnético en el interior de un condensador de placas circulares de radio R alimentado por una corriente $I(t)$

El campo eléctrico en el interior de un condensador de placas paralelas de densidad superficial de carga σ viene dado por

$$\mathbf{E} = \frac{\sigma}{\epsilon_0} \hat{\mathbf{u}}$$

donde $\hat{\mathbf{u}}$ es el vector unitario que va desde la placa cargada positivamente a la cargada negativamente. Expresando ahora la densidad superficial de carga σ en función de la carga total en la placa $Q(t)$ se tiene que

$$\mathbf{E}(t) = \frac{Q(t)}{\epsilon_0 \pi R^2} \hat{\mathbf{u}},$$

y obviamente esto implica la existencia de una corriente de desplazamiento, $\mathbf{J}_D(t)$, en el interior del condensador, que viene dada por

$$\mathbf{J}_D(t) = \epsilon_0 \frac{\partial \mathbf{E}}{\partial t} = \frac{I(t)}{\pi R^2} \hat{\mathbf{u}}.$$

Aplicando ahora la ley de Ampère–Maxwell según (2.19), esto es,

$$\oint_{\Gamma} \mathbf{B}(\mathbf{r}, t) \cdot d\mathbf{l} = \mu_0 \int_{S(\Gamma)} \mathbf{J}_D \cdot d\mathbf{S} = \mu_0 I(t).$$

nos encontramos con un problema muy similar al del cálculo del campo magnetostático en el interior de un conductor cilíndrico rectilíneo, con la diferencia de que en dicho problema la corriente era de conducción. Consecuentemente usando la expresión (2.13) se llegaría a que en el interior del condensador $r \leq R$:

$$\mathbf{B}(\mathbf{r}, t) = \frac{\mu_0 I(t)}{2\pi R^2} r \hat{\boldsymbol{\tau}}$$

2.2.4. Ecuaciones de Maxwell en forma diferencial

En este apartado expresaremos las ecuaciones de Maxwell en forma diferencial en vez de en la forma integral mostrada anteriormente. Para expresar estas ecuaciones en forma diferencial (esto es, relacionando lo que pasa en cada punto e instante de tiempo con lo que pasa en puntos e instantes de tiempo muy cercanos) tenemos que hacer uso de la definición de los operadores diferenciales **divergencia** y **rotacional** discutidos brevemente en el Apéndice C.

Empecemos tratando la ley de Gauss para el campo eléctrico (2.8) teniendo en cuenta que la carga $Q_S(t)$ puede expresarse en función de la **densidad volumétrica de carga**, $\rho(\mathbf{r}, t)$, como

$$Q_S(t) = \int_V \rho(\mathbf{r}, t) dV, \quad (2.20)$$

donde la integral está extendida al volumen V encerrado por la superficie S . Haciendo uso del teorema del valor medio³, la integral volumétrica anterior puede expresarse como

$$\int_V \rho(\mathbf{r}, t) dV = \rho(\mathbf{r}_0, t) V, \quad (2.21)$$

siendo \mathbf{r}_0 cierto punto del interior del volumen.

Teniendo en cuenta lo anterior, la ley de Gauss (2.8) puede expresarse como

$$\frac{\int_{S(V)} \mathbf{E} \cdot d\mathbf{S}}{V} = \frac{\rho(\mathbf{r}_0, t)}{\epsilon_0}. \quad (2.22)$$

Al tomar el límite cuando $V \rightarrow 0$, se tiene que

$$\lim_{V \rightarrow 0} \frac{\int_{S(V)} \mathbf{E} \cdot d\mathbf{S}}{V} = \frac{1}{\epsilon_0} \lim_{V \rightarrow 0} \rho(\mathbf{r}_0, t) = \frac{\rho(\mathbf{r}, t)}{\epsilon_0}. \quad (2.23)$$

El primer miembro de la expresión anterior es justamente la definición de la divergencia de \mathbf{E} (ver Apéndice C) y el segundo miembro, al tomar el límite cuando $V \rightarrow 0$, como $\mathbf{r}_0 \in V$, el punto \mathbf{r}_0 será justamente el punto \mathbf{r} en el que colapsa el volumen V . Todo ello nos permite escribir finalmente la ley de Gauss para el campo eléctrico en forma diferencial como

$$\nabla \cdot \mathbf{E} = \frac{\rho}{\epsilon_0}. \quad (2.24)$$

Ley de Gauss para el campo eléctrico en forma diferencial

Análogamente la ley de Gauss para el campo magnético puede expresarse en forma diferencial como

$$\nabla \cdot \mathbf{B} = 0. \quad (2.25)$$

Ley de Gauss para el campo magnético en forma diferencial

Si analizamos ahora las leyes de Faraday–Maxwell (2.10) y Ampère–Maxwell (2.18), observamos que ambas tienen la forma genérica

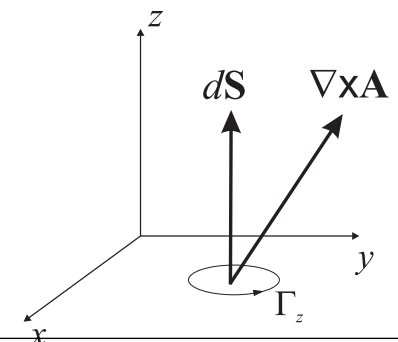
$$\oint_{\Gamma} \mathbf{A} \cdot d\mathbf{l} = \int_{S(\Gamma)} \mathbf{C} \cdot d\mathbf{S}. \quad (2.26)$$

Si particularizamos la anterior expresión para una curva situada en el plano xy encontramos que

³ Para funciones de una sola variable, el teorema del valor medio establece que

$$\exists x_0 \in (a, b) \text{ tal que } \int_a^b f(x) dx = f(x_0)(b - a).$$

Geoméricamente, este teorema dice que debe existir algún rectángulo de altura $f(x_0)$ y anchura $b - a$ que tenga el mismo área que la curva $f(x)$ y el eje x entre $x = a$ y $x = b$.



$$\oint_{\Gamma_z} \mathbf{A} \cdot d\mathbf{l} = \int_{S(\Gamma_z)} \mathbf{C} \cdot d\mathbf{S} = \int_{S(\Gamma_z)} C_z dS_z, \quad (2.27)$$

puesto que el $d\mathbf{S}$ de una superficie contenida en el plano xy será $d\mathbf{S} = dS_z \hat{\mathbf{z}}$. Usando ahora el teorema del valor medio para el segundo miembro de (2.27), podemos escribir

$$\int_{S(\Gamma_z)} C_z(\mathbf{r}, t) dS_z = C_z(\mathbf{r}_0, t) S_z \quad (2.28)$$

y, consecuentemente,

$$\frac{\oint_{\Gamma_z} \mathbf{A} \cdot d\mathbf{l}}{S_z} = C_z(\mathbf{r}_0, t). \quad (2.29)$$

Tomando ahora el límite cuando $S_z \rightarrow 0$, llegamos a que

$$\lim_{S_z \rightarrow 0} \frac{\oint_{\Gamma_z} \mathbf{A} \cdot d\mathbf{l}}{S_z} = \lim_{S_z \rightarrow 0} C_z(\mathbf{r}_0, t) = C_z(\mathbf{r}, t). \quad (2.30)$$

Teniendo en cuenta la expresión (C.5) dada en el Apéndice C para la componente z del rotacional de \mathbf{A} , (2.30) se puede reescribir como

$$(\nabla \times \mathbf{A})_z = C_z. \quad (2.31)$$

Dado que existirá una expresión análoga para las restantes componentes x e y del $\nabla \times \mathbf{A}$, podemos concluir que

$$\nabla \times \mathbf{A} = \mathbf{C}. \quad (2.32)$$

Identificando en la ecuación de Faraday-Maxwell los vectores \mathbf{A} y \mathbf{C} con \mathbf{E} y $\partial \mathbf{B} / \partial t$ respectivamente, tenemos que la forma diferencial de esta ecuación será

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}. \quad (2.33)$$

Llevando a cabo la misma identificación en la ecuación de Ampère-Maxwell, ésta se puede escribir en forma diferencial como

$$\nabla \times \mathbf{B} = \mu_0 \mathbf{J} + \mu_0 \epsilon_0 \frac{\partial \mathbf{E}}{\partial t}. \quad (2.34)$$

Las cuatro ecuaciones anteriores que relacionan el campo eléctrico, $\mathbf{E}(\mathbf{r}, t)$, y magnético, $\mathbf{B}(\mathbf{r}, t)$, con la densidad de carga total, $\rho(\mathbf{r}, t)$, y la densidad de corriente total, $\mathbf{J}(\mathbf{r}, t)$, forman el conjunto de ecuaciones de Maxwell

$$(i) \quad \nabla \cdot \mathbf{E} = \frac{1}{\epsilon_0} \rho \quad (2.35a)$$

$$(ii) \quad \nabla \cdot \mathbf{B} = 0 \quad (2.35b)$$

$$(iii) \quad \nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} \quad (2.35c)$$

$$(iv) \quad \nabla \times \mathbf{B} = \mu_0 \mathbf{J} + \mu_0 \epsilon_0 \frac{\partial \mathbf{E}}{\partial t}. \quad (2.35d)$$

Estas ecuaciones junto con la ley de Lorentz,

$$\mathbf{F} = q(\mathbf{E} + \mathbf{v} \times \mathbf{B}), \quad (2.36)$$

forman todo el cuerpo teórico de la Electrodinámica en el vacío.

Ley de Faraday-Maxwell en forma diferencial

Ley de Ampère-Maxwell en forma diferencial

Ecuaciones de Maxwell

2.2.5. Ecuaciones de Maxwell en la materia

Las ecuaciones de Maxwell (2.35a)-(2.35d) son completas y válidas en la forma en que se han presentado. No obstante, cuando existen medios materiales en nuestro problema, esta forma no es posiblemente la más "atractiva" desde un punto de vista práctico. Ello se debe a que, en presencia de medios materiales, aparecen cargas y corrientes *ligadas* a los materiales, que no tienen nada que ver con las cargas y corrientes *libres*. Con el término de cargas y corrientes libres queremos decir aquellas cargas y corrientes que también existirían en el vacío dado que son las cargas y corrientes que nosotros imponemos y que consecuentemente podemos controlar. Es por tanto muy conveniente escribir las ecuaciones de Maxwell de tal forma que aparezcan directamente estas cargas y corrientes libres. Un estudio del efecto de la presencia de medios materiales *lineales e isótropos* (que será aquí omitido) indica que los únicos cambios que hay que hacer en las ecuaciones de Maxwell (2.35a)-(2.35d) para relacionar los campos \mathbf{E} y \mathbf{B} con la densidad de carga libre, ρ_f y densidad de corriente libre \mathbf{J}_f consisten en sustituir los valores de ϵ_0 y μ_0 (permitividad dieléctrica y permeabilidad magnética del vacío) por ϵ y μ respectivamente; esto es, por los valores de la permitividad dieléctrica y permeabilidad magnética del medio material. Las ecuaciones de Maxwell para un medio material lineal e isótropo caracterizado por una permitividad dieléctrica ϵ y una permeabilidad magnética μ podrían entonces escribirse como

$$(i) \quad \nabla \cdot \mathbf{E} = \frac{1}{\epsilon} \rho_f \quad (2.37a)$$

$$(ii) \quad \nabla \cdot \mathbf{B} = 0 \quad (2.37b)$$

$$(iii) \quad \nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} \quad (2.37c)$$

$$(iv) \quad \nabla \times \mathbf{B} = \mu \mathbf{J}_f + \mu \epsilon \frac{\partial \mathbf{E}}{\partial t}, \quad (2.37d)$$

Ecuaciones de Maxwell en medios materiales

donde

$$\epsilon = \epsilon_r \epsilon_0 \quad \text{y} \quad \mu = \mu_r \mu_0, \quad (2.38)$$

siendo ϵ_r y μ_r unas constantes adimensionales llamadas respectivamente *permitividad dieléctrica relativa* y *permeabilidad magnética relativa* del medio.

Es interesante notar que si introducimos dos nuevos vectores \mathbf{D} y \mathbf{H} , definidos para medios lineales e isótropos como

$$\mathbf{D} = \epsilon \mathbf{E} \quad \text{vector desplazamiento eléctrico} \quad (2.39)$$

$$\mathbf{H} = \frac{1}{\mu} \mathbf{B} \quad \text{vector intensidad magnética,} \quad (2.40)$$

entonces las ecuaciones (2.37a)-(2.37d) pueden reescribirse como

$$(i) \quad \nabla \cdot \mathbf{D} = \rho_f \quad (2.41a)$$

$$(ii) \quad \nabla \cdot \mathbf{B} = 0 \quad (2.41b)$$

$$(iii) \quad \nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} \quad (2.41c)$$

$$(iv) \quad \nabla \times \mathbf{H} = \mathbf{J}_f + \frac{\partial \mathbf{D}}{\partial t}. \quad (2.41d)$$

Ecuaciones de Maxwell en la materia

Aunque las ecuaciones (2.41a)-(2.41d) han sido obtenidas para medios lineales e isótropos, esta es la forma usual y completamente general en la que se escriben las ecuaciones de Maxwell para cualquier medio en términos de las cargas y corrientes libres.

2.3. Ecuación de Ondas para las Ondas Electromagnéticas

Según se discutió en el Apartado 1.2, la expresión matemática de cualquier magnitud vectorial que represente a una onda debe satisfacer la ecuación de ondas (1.110). En este sentido, los campos \mathbf{E} y \mathbf{B} representarán una onda si obtenemos una ecuación de ondas para ellos.

En una región de un medio material lineal e isótropo caracterizado por ϵ y μ donde no haya cargas ($\rho_f = 0$) ni corrientes libres ($\mathbf{J}_f = 0$), las ecuaciones de Maxwell (2.37a)–(2.37d) serán

$$(i) \quad \nabla \cdot \mathbf{E} = 0 \quad (2.42a)$$

$$(ii) \quad \nabla \cdot \mathbf{B} = 0 \quad (2.42b)$$

$$(iii) \quad \nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} \quad (2.42c)$$

$$(iv) \quad \nabla \times \mathbf{B} = \mu\epsilon \frac{\partial \mathbf{E}}{\partial t} \quad (2.42d)$$

Tomando rotacional en las ecuaciones (iii) y (iv) y teniendo en cuenta la siguiente identidad vectorial:

$$\nabla \times (\nabla \times \mathbf{A}) = \nabla(\nabla \cdot \mathbf{A}) - \nabla^2 \mathbf{A}$$

(que puede verse como una aplicación de la expresión (A.21) al operador vectorial ∇ , teniendo en cuenta que $\nabla \cdot \nabla \equiv \nabla^2$) se tiene que

$$\begin{aligned} \nabla \times (\nabla \times \mathbf{E}) &= \nabla(\nabla \cdot \mathbf{E}) - \nabla^2 \mathbf{E} = \nabla \times \left(-\frac{\partial \mathbf{B}}{\partial t} \right) \\ &= -\frac{\partial}{\partial t} (\nabla \times \mathbf{B}) = -\mu\epsilon \frac{\partial^2 \mathbf{E}}{\partial t^2} \end{aligned} \quad (2.43)$$

$$\begin{aligned} \nabla \times (\nabla \times \mathbf{B}) &= \nabla(\nabla \cdot \mathbf{B}) - \nabla^2 \mathbf{B} = \nabla \times \left(\mu\epsilon \frac{\partial \mathbf{E}}{\partial t} \right) \\ &= \mu\epsilon \frac{\partial}{\partial t} (\nabla \times \mathbf{E}) = -\mu\epsilon \frac{\partial^2 \mathbf{B}}{\partial t^2} \end{aligned} \quad (2.44)$$

Considerando ahora las ecuaciones (2.42a) y (2.42b) para la divergencia de los campos eléctrico y magnético en las expresiones anteriores, llegamos a las siguientes ecuaciones para los campos $\mathbf{E}(\mathbf{r}, t)$ y $\mathbf{B}(\mathbf{r}, t)$:

$$\nabla^2 \mathbf{E} - \mu\epsilon \frac{\partial^2 \mathbf{E}}{\partial t^2} = 0 \quad (2.45)$$

$$\nabla^2 \mathbf{B} - \mu\epsilon \frac{\partial^2 \mathbf{B}}{\partial t^2} = 0. \quad (2.46)$$

Las dos ecuaciones anteriores son *ecuaciones de onda* lo que indica que, en un medio material sin fuentes, los campos \mathbf{E} y \mathbf{B} serán perturbaciones ondulatorias que viajan a una velocidad v dada por

$$v = \frac{1}{\sqrt{\mu\epsilon}} \leq c, \quad (2.47)$$

siendo c la velocidad de propagación en el espacio libre (sin medios materiales y, por tanto, $\epsilon = \epsilon_0$ y $\mu = \mu_0$) cuyo valor numérico viene dado por

$$c = \frac{1}{\sqrt{\mu_0 \epsilon_0}} = 2.99792 \times 10^8 \text{ m/s}. \quad (2.48)$$

Dado que la velocidad a la que se propaga el campo electromagnético en el vacío (obtenida de forma teórica mediante manipulaciones en las ecuaciones de Maxwell) era muy próxima a la velocidad medida experimentalmente para la luz, esta sorprendente coincidencia era una clara sugerencia de que la luz consistía en una onda electromagnética. Debe notarse que en el momento en que se dedujo teóricamente la velocidad de propagación del campo electromagnético, se admitía que la luz era una onda pero se discutía sobre la *naturaleza* de esta onda⁴.

2.3.1. Ondas electromagnéticas armónicas

Ya se señaló en el Apartado 1.3 que una solución particularmente importante de la ecuación de ondas era la solución armónica. En concreto, una importante solución de (2.45) es un campo eléctrico del tipo

$$\mathbf{E}(\mathbf{r}, t) = \mathbf{E}_0 e^{j(\omega t - \mathbf{k} \cdot \mathbf{r})}, \quad (2.49)$$

que representa una onda plana armónica polarizada linealmente que se propaga en la dirección del vector de ondas \mathbf{k} —tal como la mostrada en (1.111). El campo magnético asociado a este campo eléctrico en la onda electromagnética puede calcularse a partir de (2.42c),

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t},$$

teniendo en cuenta que, en el presente caso, la aplicación del operador vectorial nabla al campo (2.49) es equivalente a la aplicación del vector $-j\mathbf{k}$, de modo que

$$\nabla \times \mathbf{E} = -j\mathbf{k} \times \mathbf{E},$$

y por tanto (2.42c) queda como

$$-j\mathbf{k} \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}. \quad (2.50)$$

Integrando la ecuación anterior obtenemos que

$$\begin{aligned} \mathbf{B}(\mathbf{r}, t) &= \int j\mathbf{k} \times \mathbf{E}_0 e^{j(\omega t - \mathbf{k} \cdot \mathbf{r})} dt = \frac{\mathbf{k}}{\omega} \times \mathbf{E}_0 e^{j(\omega t - \mathbf{k} \cdot \mathbf{r})} \\ &= \mathbf{B}_0 e^{j(\omega t - \mathbf{k} \cdot \mathbf{r})}, \end{aligned} \quad (2.51)$$

de donde se deduce que

$$\frac{\mathbf{k}}{\omega} \times \mathbf{E} = \mathbf{B}. \quad (2.52)$$

Si se multiplica vectorialmente ambos miembros de (2.52) por \mathbf{k} , teniendo en cuenta la identidad (A.21), encontramos

$$\begin{aligned} \mathbf{k} \times (\mathbf{k} \times \mathbf{E}) &= \omega \mathbf{k} \times \mathbf{B} \\ &= (\mathbf{k} \cdot \mathbf{E})\mathbf{k} - (\mathbf{k} \cdot \mathbf{k})\mathbf{E} = -k^2 \mathbf{E}, \end{aligned}$$

puesto que al aplicar (2.42a) al caso de la onda armónica plana tenemos que

$$\mathbf{k} \cdot \mathbf{E} = 0 \quad (2.53)$$

y, por tanto, puede escribirse

$$\mathbf{E} = -\frac{\omega}{k^2} \mathbf{k} \times \mathbf{B}. \quad (2.54)$$

⁴ Se postulaba, por ejemplo, que la luz, en analogía con las ondas mecánicas, era una vibración de las partículas de un medio que "impregnaba" todo el universo denominado *éter*.

Teniendo ahora en cuenta que el módulo del vector velocidad de fase, \mathbf{v} , de la onda plana armónica viene dada por $v = \omega/k$ y que la dirección de propagación de la onda plana armónica (y por tanto la de avance del frente de onda) viene determinada por el vector de onda \mathbf{k} , el vector velocidad de fase puede expresarse como

$$\mathbf{v} = \frac{\omega}{k^2} \mathbf{k} = \frac{\omega}{k} \hat{\mathbf{k}}.$$

En consecuencia, la expresión (2.54) puede también reescribirse como

$$\mathbf{E} = \mathbf{B} \times \mathbf{v}, \quad (2.55)$$

lo que implica que

- Los campos eléctrico y magnético de una onda plana armónica están en fase.
- Las amplitudes de los campos eléctrico y magnético de una onda plana armónica están relacionados según

$$E_0 = B_0 v. \quad (2.56)$$

- \mathbf{E} , \mathbf{B} y \mathbf{v} forman un *triedro rectángulo* (esto es, cada uno de estos vectores es perpendicular a los otros dos). Esta última conclusión se deduce al considerar (2.55) junto con (2.53). Dado que el plano de vibración de los vectores \mathbf{E} y \mathbf{B} es perpendicular a la dirección de propagación de la onda, podemos concluir que *la onda electromagnética plana armónica es transversal* (ver Figura 2.1).

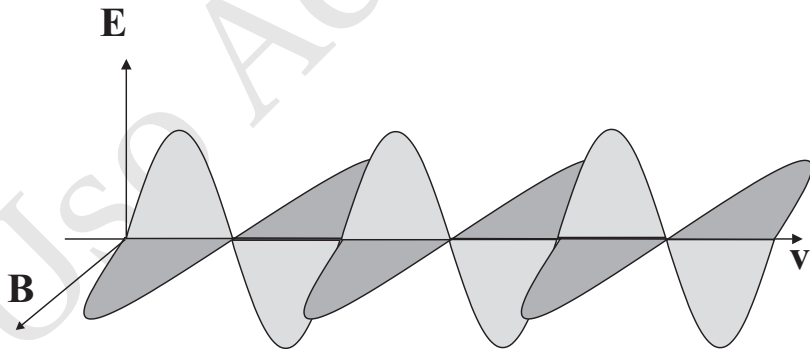
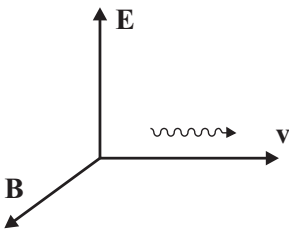


FIGURA 2.1: Representación de una onda plana armónica polarizada linealmente que se propaga en la dirección marcada por \mathbf{k} . El dibujo representa la onda en algún instante de tiempo

2.4. Polarización

Hasta ahora hemos tratado únicamente ondas electromagnéticas polarizadas linealmente, es decir, ondas en las que la orientación del campo eléctrico/magnético es constante aunque su magnitud y signo varíen con el tiempo:

$$\mathbf{E} = E_0 \cos(\omega t - \mathbf{k} \cdot \mathbf{r}).$$

Esta situación es un caso particular de lo que le ocurre en general a la dirección del campo eléctrico en una onda plana armónica. Para analizar un caso más general, consideremos el campo eléctrico producido por la siguiente superposición

de dos ondas planas de la misma frecuencia, \mathbf{E}_1 y \mathbf{E}_2 , polarizadas linealmente y ortogonales entre sí que se propagan a lo largo de la dirección z :

$$\mathbf{E}_1(z, t) = E_{0x} \cos(\omega t - kz) \hat{\mathbf{x}} \quad (2.57)$$

$$\mathbf{E}_2(z, t) = E_{0y} \cos(\omega t - kz + \varphi) \hat{\mathbf{y}}. \quad (2.58)$$

El campo resultante, $\mathbf{E}(z, t)$, será obviamente la suma de los campos anteriores,

$$\mathbf{E}(z, t) = E_{0x} \cos(\omega t - kz) \hat{\mathbf{x}} + E_{0y} \cos(\omega t - kz + \varphi) \hat{\mathbf{y}}. \quad (2.59)$$

En función del valor del ángulo de desfase φ y de la relación entre las amplitudes podemos encontrar diversas situaciones:

■ **Polarización lineal:**

Si el ángulo de desfase es nulo o un múltiplo entero de 2π ,

$$\varphi = 2n\pi \quad (n \in \mathbb{Z}), \quad (2.60)$$

entonces las dos ondas están en fase y el campo resultante vendrá dado por

$$\mathbf{E}(z, t) = (E_{0x} \hat{\mathbf{x}} + E_{0y} \hat{\mathbf{y}}) \cos(\omega t - kz). \quad (2.61)$$

La onda resultante tiene una amplitud, \mathbf{E}_0 , constante igual a

$$\mathbf{E}_0 = (E_{0x} \hat{\mathbf{x}} + E_{0y} \hat{\mathbf{y}})$$

y, por consiguiente, es también una onda polarizada linealmente. Esta situación también se da si φ es un múltiplo entero de π , en cuyo caso las ondas se dice que están en *contrafase*.

■ **Polarización circular:**

Este caso se da cuando la amplitud de ambas ondas es idéntica y el ángulo de desfase es

$$\varphi = \frac{\pi}{2} + 2m\pi \quad \text{con } m = 0, \pm 1, \pm 2, \dots, \quad (2.62)$$

en cuyo caso

$$\mathbf{E}_1(z, t) = E_0 \cos(\omega t - kz) \hat{\mathbf{x}} \quad (2.63)$$

$$\mathbf{E}_2(z, t) = -E_0 \sin(\omega t - kz) \hat{\mathbf{y}}. \quad (2.64)$$

La onda resultante queda entonces

$$\mathbf{E}(z, t) = E_0 [\cos(\omega t - kz) \hat{\mathbf{x}} - \sin(\omega t - kz) \hat{\mathbf{y}}]. \quad (2.65)$$

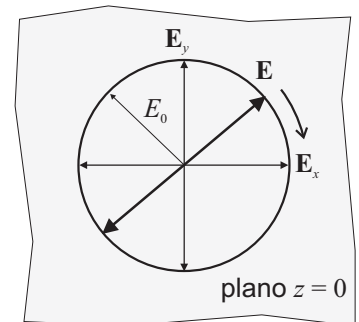
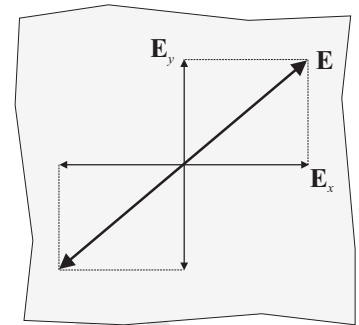
Obsérvese que el módulo del campo eléctrico es constante, $|\mathbf{E}(z, t)| = E_0$, pero su dirección es variable en el tiempo. Para un punto concreto del espacio (tomemos $z = 0$ por sencillez) la perturbación eléctrica toma el siguiente valor:

$$\mathbf{E}(0, t) = E_0 [\cos(\omega t) \hat{\mathbf{x}} - \sin(\omega t) \hat{\mathbf{y}}]. \quad (2.66)$$

Observando la evolución de este vector en el tiempo se ve que éste rota en sentido horario recorriendo su extremo un círculo y, por este motivo, se dice que la onda (2.65) tiene *polarización circular derecha*. Análogamente una onda como

$$\mathbf{E}(z, t) = E_0 [\cos(\omega t - kz) \hat{\mathbf{x}} + \sin(\omega t - kz) \hat{\mathbf{y}}] \quad (2.67)$$

se dice que presenta *polarización circular izquierda* puesto que la amplitud de esta onda recorre un círculo en sentido antihorario.



Un resultado interesante es que una onda linealmente polarizada puede construirse superponiendo dos ondas con polarización circular opuesta de la misma amplitud. Este hecho puede comprobarse sumando las ondas (2.65) y (2.67) para obtener

$$\mathbf{E}(z, t) = 2E_0 \cos(\omega t - kz) \hat{\mathbf{x}}, \quad (2.68)$$

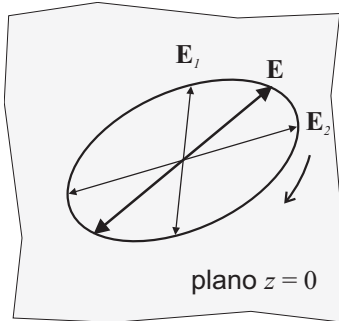
que es evidentemente una onda polarizada linealmente.

■ Polarización elíptica:

Los dos casos anteriores pueden considerarse casos particulares del caso más general de polarización elíptica. Dado que para un punto z fijo, el campo eléctrico puede expresarse en general como

$$\mathbf{E}(z, t) = E_{0x} \cos(\omega t - kz) \hat{\mathbf{x}} - E_{0y} \sin(\omega t - kz + \varphi) \hat{\mathbf{y}}, \quad (2.69)$$

podemos comprobar que el extremo del anterior vector campo eléctrico recorre una elipse situada en el plano x, y . La orientación y tamaño de esta elipse depende del valor del desfase φ y de la relación E_{0x}/E_{0y} .



2.5. Intensidad de la onda electromagnética

En el Apartado 1.24 se definió la intensidad instantánea, I_{inst} , de una onda como la energía que fluye por unidad de tiempo a través de una superficie perpendicular a la dirección de propagación —ver expresión (1.24). Por tanto, para calcular la intensidad de la onda electromagnética debemos obtener en primer lugar la densidad volumétrica de energía asociada con esta onda. La energía del campo eléctrico y magnético en un medio lineal e isótropo vienen dadas respectivamente por

$$u_E = \frac{1}{2} \epsilon E^2 \quad \text{densidad de energía eléctrica} \quad (2.70)$$

$$u_B = \frac{B^2}{2\mu} \quad \text{densidad de energía magnética}, \quad (2.71)$$

por lo que la intensidad instantánea de la onda electromagnética vendrá dada por

$$I_{\text{inst}} = (u_E + u_B) v. \quad (2.72)$$

Para ondas planas armónicas, encontramos que la relación entre los módulos de los campos eléctrico y magnético de la onda electromagnética verificaban que $E = vB$. Esto nos permite escribir la densidad de energía almacenada en el campo magnético de estas ondas como

$$u_B = \frac{B^2}{2\mu} = \frac{E^2}{2\mu v^2} = \frac{1}{2} \epsilon E^2, \quad (2.73)$$

donde se ha tenido en cuenta que $v^2 = 1/\mu\epsilon$. Se ha obtenido, por tanto, que para una onda plana electromagnética armónica, la densidad de energía almacenada en el campo magnético es idéntica a la almacenada en el campo eléctrico, esto es,

$$u_E = u_B. \quad (2.74)$$

La anterior expresión nos permite escribir la densidad de energía de la onda electromagnética, u_{EB} , como

$$u_{EB} = u_E + u_B = \frac{1}{2} \epsilon E^2 + \frac{1}{2} \epsilon E^2 = \epsilon E^2 = \frac{B^2}{\mu} = \frac{EB}{\mu v}, \quad (2.75)$$

Intensidad instantánea de una onda electromagnética

y, consecuentemente, podemos expresar la intensidad instantánea de dicha onda como

$$I_{\text{inst}} = u_{EB} v = v \epsilon E^2 = v \frac{B^2}{\mu} = \frac{EB}{\mu} = EH. \quad (2.76)$$

En el medio lineal e isótropo, la energía de la onda viaja en la dirección de propagación de la onda; esto es, en una dirección perpendicular tanto a \mathbf{E} como \mathbf{B} para una onda plana armónica. Por otra parte, para este tipo de ondas, la intensidad de la onda se puede expresar, según (2.76), en función de los módulos de los campos eléctrico y magnético de la onda. Todo ello sugiere la introducción de un vector \mathbf{S} , denominado **vector de Poynting**, que determine energéticamente a la onda electromagnética y que, por tanto, tenga por dirección la dirección de propagación de la energía y por módulo el valor de la intensidad instantánea de la onda electromagnética. A la vista de las expresiones anteriores, para una onda plana armónica, este vector vendría dado por el siguiente producto vectorial:

$$\mathbf{S}(\mathbf{r}, t) = \mathbf{E} \times \mathbf{H}. \quad (2.77)$$

Vector de Poynting

Aunque la expresión anterior del vector de Poynting se ha obtenido para el caso concreto de una onda plana armónica en un medio lineal e isótropo, cálculos más elaborados mostrarían que la expresión (2.77) tiene validez general para cualquier tipo de onda electromagnética y medio material.

Para la onda plana armónica polarizada linealmente, el vector de Poynting viene dado por

$$\mathbf{S}(\mathbf{r}, t) = v \epsilon E_0^2 \cos^2(\omega t - \mathbf{k} \cdot \mathbf{r}) \hat{\mathbf{k}} = v u_{EB} \hat{\mathbf{k}}, \quad (2.78)$$

y consecuentemente la intensidad instantánea de esta onda es

$$I_{\text{inst}}(\mathbf{r}, t) = v \epsilon E_0^2 \cos^2(\omega t - \mathbf{k} \cdot \mathbf{r}). \quad (2.79)$$

Hay que tener en cuenta que los valores instantáneos de la intensidad (o bien de la potencia) no tienen mucho interés práctico dado que usualmente las ondas electromagnéticas varían muy rápidamente (del orden de 10^{15} veces en un segundo para la luz). Es, por tanto, más significativo obtener el promedio de la intensidad, I_{med} , en un periodo de tiempo, para lo cual debemos promediar temporalmente (2.79):

$$\begin{aligned} I_{\text{med}} &= \langle I_{\text{inst}}(\mathbf{r}, t) \rangle = v \epsilon E_0^2 \frac{1}{T} \int_0^T \cos^2(\omega t - \mathbf{k} \cdot \mathbf{r}) dt \\ &= \frac{v \epsilon E_0^2}{2} = \frac{1}{2} \frac{E_0 B_0}{\mu} = \frac{1}{2} E_0 H_0. \end{aligned} \quad (2.80)$$

Intensidad promedio de una onda electromagnética armónica

EJEMPLO 2.2 Sabiendo que la amplitud del campo eléctrico de la radiación solar que llega a la superficie terrestre es de aproximadamente $E_0 = 850$ V/m, calcule la potencia total que incidiría sobre una azotea de 100 m².

Para calcular la potencia promedio que incide en una superficie S debemos primero obtener el valor de la intensidad promedio, I_{med} , de la onda. En este caso dado que conocemos el valor de la amplitud del campo eléctrico, esta intensidad vendrá dada por

$$I_{\text{med}} = \frac{1}{2} c \epsilon_0 E_0^2 = \frac{3 \times 10^8 \cdot 8.85 \times 10^{-12}}{2} (850)^2 \approx 959 \frac{\text{W}}{\text{m}^2}.$$

Una vez calculada la intensidad promedio, la potencia promedio, P_{med} , que incide sobre la superficie será simplemente

$$P_{\text{med}} = I_{\text{med}} S = 959 \cdot 100 \approx 9.6 \times 10^4 \text{ W}.$$

Aunque esta potencia es realmente alta, debe tenerse en cuenta que está distribuida en una área grande y que su aprovechamiento total es imposible. De hecho con placas solares típicas se podría transformar en potencia eléctrica aproximadamente el 10% de la radiación solar, debiéndose tener en cuenta además que los datos dados en el problema se refieren a las horas de iluminación de días soleados.

2.6. Fuentes de las Ondas Electromagnéticas

Hasta ahora hemos estado estudiando algunas características de las ondas electromagnéticas pero todavía no sabemos dónde y cómo se originan estas ondas. Dado que las ondas electromagnéticas son simplemente campos eléctricos y magnéticos oscilantes y las fuentes de estos campos son las cargas estáticas y/o en movimiento, es razonable suponer que serán igualmente las cargas las fuentes de las ondas electromagnéticas. No obstante, debemos notar que estamos hablando de las fuentes de los campos "primarios" puesto que, como se discutió anteriormente en el Apartado 2.1, una vez que se han generado estos campos primarios, son precisamente los propios campos los responsables de la generación de los campos subsiguientes. No obstante, para que los campos primarios generen otros campos, éstos debían ser campos variables en el tiempo por lo que *ni cargas estáticas ni cargas en movimiento uniforme pueden producir ondas electromagnéticas*⁵. Consecuentemente serán las cargas eléctricas en el único estado de movimiento no considerado anteriormente, esto es, cargas aceleradas las que originen los campos primarios y, por tanto, podemos concluir que

las cargas eléctricas aceleradas son las fuentes de las ondas electromagnéticas.

En concreto, una partícula con carga q que se mueve en el espacio libre con velocidad $\mathbf{v}(t)$ y aceleración $\dot{\mathbf{v}}(t)$ produce unos campos de radiación dados por⁶

$$\mathbf{E} = \frac{q}{4\pi\epsilon_0 c^2 r} [\hat{\mathbf{r}} \times (\hat{\mathbf{r}} \times \dot{\mathbf{v}})] \quad (2.81)$$

$$\mathbf{B} = \frac{1}{c} \hat{\mathbf{r}} \times \mathbf{E}. \quad (2.82)$$

El modulo del campo eléctrico puede escribirse como

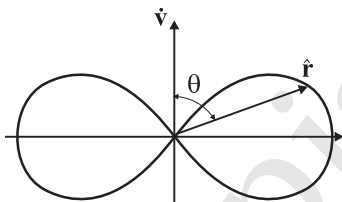
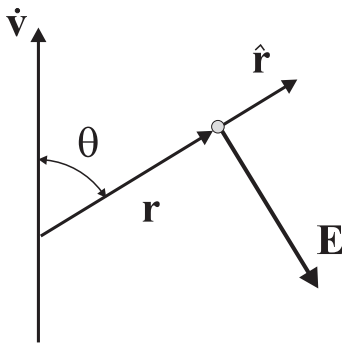
$$E = \frac{q|\dot{\mathbf{v}}|}{4\pi\epsilon_0 c^2 r} \sin\theta, \quad (2.83)$$

donde θ es el ángulo entre los vectores $\hat{\mathbf{r}}$ y $\dot{\mathbf{v}}$. Los campos dados anteriormente están *retardados* lo que quiere decir que \mathbf{v} debe ser evaluada un tiempo $t = r/c$ anterior a la medida de \mathbf{E} y \mathbf{B} . Juntos los campos dados por (2.81) y (2.82) forman una radiación que emite energía predominantemente en el plano normal a la dirección de la aceleración de la carga. Según (2.77), el vector de Poynting para la situación analizada vendrá dado por

$$\mathbf{S} = \frac{q^2 |\dot{\mathbf{v}}|^2}{16\pi^2 \epsilon_0 c^3 r^2} \sin^2\theta \hat{\mathbf{r}}. \quad (2.84)$$

⁵ Recordemos que las cargas estáticas son las fuentes de campos eléctricos estáticos y las cargas en movimiento uniforme (esto es, las corrientes eléctricas) son las fuentes de los campos magnéticos estáticos.

⁶ Se llaman campos de radiación a aquéllos que muestran una dependencia del tipo r^{-1} . Esto asegura que la intensidad de la onda decae como r^{-2} y consecuentemente es capaz de transportar energía hasta el infinito. La carga acelerada produciría también otros campos que no son de radiación y que no serán considerados aquí.



La expresión anterior muestra claramente que existe un máximo de radiación emitida en la dirección normal al vector aceleración.

Uno de los sistemas más elementales en los que encontramos una carga permanentemente acelerada y, por tanto, radiando es una *carga oscilando armónicamente*. Esta situación se conoce como **dipolo oscilante** y simula muy aproximadamente lo que ocurre en muchas antenas o bien en átomos y moléculas. Si la carga está oscilando a una frecuencia ω a lo largo del eje z , su posición viene dada por

$$z(t) = A \cos(\omega t) \quad (2.85)$$

y, por tanto, su aceleración será

$$\dot{\mathbf{v}} = -A\omega^2 \cos(\omega t) \hat{\mathbf{z}}. \quad (2.86)$$

De acuerdo con las expresiones (2.81)–(2.82), los campos radiados por este dipolo oscilante son

$$\mathbf{E} = -\frac{qA\omega^2}{4\pi\epsilon_0 c^2 r} [\hat{\mathbf{r}} \times (\hat{\mathbf{r}} \times \hat{\mathbf{z}})] \quad (2.87)$$

$$\mathbf{B} = \frac{1}{c} \hat{\mathbf{r}} \times \mathbf{E}. \quad (2.88)$$

La potencia instantánea, $P(t)$, radiada por el dipolo oscilante puede calcularse integrando el vector de Poynting en una superficie esférica de radio r , obteniéndose entonces que

$$P(t) = \frac{p_0^2 \omega^4}{6\pi\epsilon_0 c^3} \cos^2(\omega t), \quad (2.89)$$

donde $p_0 = qA$. La potencia promedio radiada por el dipolo oscilante vendrá finalmente dada por

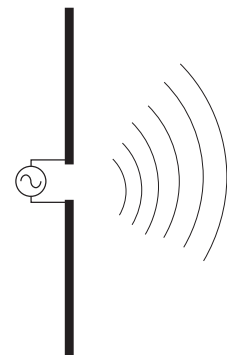
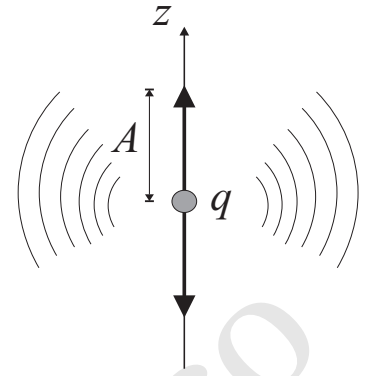
$$P_{\text{med}} = \langle P \rangle = \frac{p_0^2 \omega^4}{12\pi\epsilon_0 c^3}, \quad (2.90)$$

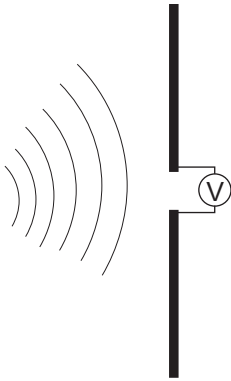
teniendo en cuenta que el promedio de $\cos^2(\omega t)$ es $1/2$. Uno de los aspectos más destacados de la expresión anterior es la fuerte dependencia de la potencia promedio con la frecuencia,

$$P_{\text{med}} \propto \omega^4, \quad (2.91)$$

lo que explicaría, por ejemplo, el color azul del cielo. Para ello debemos considerar que las moléculas de la atmósfera actúan como dipolos oscilantes que *reirradian* la energía recibida del sol. Debido a la forma de la expresión (2.91), las componentes azules del espectro visible (es decir, aquéllas de mayor frecuencia) son potenciadas a expensas de otras componentes de menor frecuencia, resultando todo este proceso en el característico color azul del cielo.

La discusión anterior ha mostrado que cargas eléctricas oscilando a una determinada frecuencia ω son focos de ondas electromagnéticas de esa misma frecuencia, y con una longitud de onda en el espacio libre dada por: $\lambda = 2\pi c/\omega$. Normalmente la oscilación de una única carga produce una onda cuya intensidad es prácticamente indetectable, por ello las ondas electromagnéticas suelen originarse en la práctica cuando un número importante de cargas están oscilando conjuntamente. Este hecho se produce, por ejemplo, en las **antenas** emisoras. Un tipo básico de estas antenas lo constituyen dos varillas conductoras alimentadas mediante un generador de corriente alterna. El generador de corriente alterna provoca que los electrones de las varillas conductoras viajen desde un extremo a otro de las varillas realizando un movimiento oscilatorio determinado por la frecuencia del generador. Este tipo de antenas es el comúnmente usado para generar ondas de radio y TV ($\text{MHz} \lesssim f \lesssim \text{GHz}$). Las ondas de la luz visible, que oscilan a





una $f \sim 10^{15}$ Hz, son originadas por el movimiento de las cargas atómicas mientras que radiaciones de mayor frecuencia son debidas a las rápidas oscilaciones electrónicas y nucleares.

El mismo mecanismo que justifica que los electrones (cargas) en movimiento en un conductor originan ondas electromagnéticas, esto es, forman una *antena emisora*, también explica por qué este mismo dispositivo (sin el generador) sería una *antena receptora*. Los campos eléctricos que llegan a la antena ejercen una fuerza sobre las cargas móviles del conductor (electrones) que las hacen oscilar a la misma frecuencia que la onda electromagnética incidente. Claramente, el movimiento de estas cargas, que simplemente sigue el patrón de la radiación incidente, produce una corriente eléctrica oscilante que puede ser detectada por algún dispositivo adecuado. De esta manera el patrón de variación temporal que se produjo en el generador de la antena emisora es ahora “recogido” en el detector de la antena receptora⁷. Mediante el procedimiento anterior se ha transmitido información desde un sitio a otro del espacio usando como intermediaria a la onda electromagnética. Esta manera de transmitir información es muy eficaz ya que pone en juego muy poca energía y permite transmitir información entre puntos muy lejanos entre sí (incluyendo comunicaciones con satélites y vehículos espaciales).

2.7. Espectro electromagnético

Uno de los aspectos más interesantes de las ondas electromagnéticas es que distintos fenómenos ondulatorios aparentemente inconexos como la luz, las ondas de radio, las microondas, los rayos X, los rayos gamma y otras radiaciones son todos ellos ondas electromagnéticas de distinta frecuencia y longitud de onda. Todos los fenómenos anteriores son simplemente campos eléctricos y magnéticos oscilantes a determinada frecuencia. En el espacio libre, la relación entre la frecuencia f y la longitud de onda λ viene, evidentemente, dada por

$$f = \frac{c}{\lambda}. \quad (2.92)$$

El conjunto de todas las radiaciones electromagnéticas se conoce como **espectro electromagnético**, distinguiéndose en él las distintas denominaciones que toman las ondas electromagnéticas en función de la frecuencia, tal como muestra la Figura 2.2.

Según se ha estudiado en el Tema 1, las propiedades de las ondas vienen determinadas fundamentalmente por su frecuencia y longitud de onda. En el Apartado 1.7 se vio, por ejemplo, que los fenómenos de difracción dependían básicamente de la relación entre la longitud de onda y el tamaño físico de los objetos donde se producía la difracción. Esto justificaba que los efectos de difracción de la luz fuesen difíciles de percibir debido a la corta longitud de onda de la luz visible ($400 \lesssim \lambda(\text{nm}) \lesssim 700$) y que, por tanto, la luz pueda ser considerada en muchas situaciones prácticas como un *rayo*. La misma explicación sirve para entender por qué grandes obstáculos como edificios o montes no afectan drásticamente a la propagación de ondas de radio largas ($10^7 \lesssim \lambda(\text{m}) \lesssim 10^2$). La interacción de la onda electromagnética con la materia también depende básicamente de la longitud de la onda y así, la pequeña longitud de onda de los rayos X ($10^{-12} \lesssim \lambda(\text{m}) \lesssim 10^{-8}$)

⁷ Los electrones de la antena receptora se mueven tal como lo hacían los electrones de la antena emisora, sólo que cierto intervalo de tiempo después; justamente el necesario para que la onda recorra la distancia entre las dos antenas.

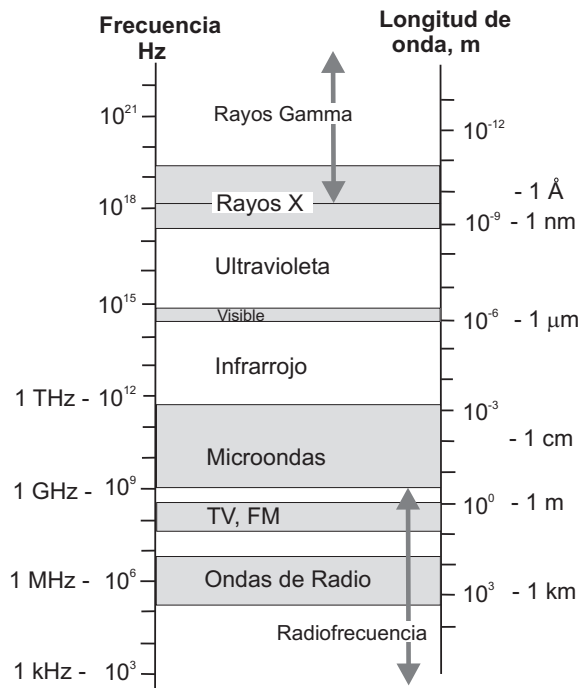


FIGURA 2.2: Espectro electromagnético

es la que explica por qué estos rayos pueden atravesar fácilmente muchos materiales que son opacos para radiaciones de mayor longitud de onda. Igualmente, al coincidir la longitud de onda de las ondas generadas en los hornos de microondas ($\lambda \sim 15\text{cm}$) con el espectro de absorción de las moléculas de agua se explica que esta radiación sea fuertemente absorbida por las moléculas de agua que contienen los alimentos y, consecuentemente, éstos se calienten.

2.8. Problemas propuestos

2.1: Demostrar por sustitución directa que la siguiente expresión:

$$E_y(x, t) = E_0 \sin(kx - \omega t) = E_0 \sin k(x - ct),$$

donde $c = \omega/k$, satisface la ecuación (2.45).

2.2: Hallar la longitud de onda de **a)** una onda de radio de AM típica con una frecuencia de 100 kHz, **b)** una onda de radio de FM típica de 100 MHz; **c)** la frecuencia de una microonda de 3 cm y **d)** la frecuencia de unos rayos X con una longitud de onda de 0,1 nm.

Sol. **a)** $\lambda = 300\text{m}$; **b)** $\lambda = 300\text{m}$; **c)** $f = 10\text{GHz}$; **d)** $f = 3 \times 10^{18}\text{Hz}$;

2.3: Una onda electromagnética (OEM) plana se propaga en el vacío. Sabiendo que su frecuencia es de 98.4 MHz y su amplitud de campo eléctrico es de 20 mV/m, calcúlese: **a)** la amplitud del campo magnético; **b)** la intensidad de onda (potencia media por unidad de área).

Sol. **a)** $B_0 = 0.66 \times 10^{-10}\text{T}$; **b)** $I = 0.53 \mu\text{W}/\text{m}^2$.

2.4: Una OEM plana se propaga a lo largo del eje X con una longitud de onda de 3 cm, transportando una potencia media por unidad de área de $6 \mu\text{W}/\text{m}^2$. Determinense las expresiones completas de los campos \mathbf{E} y \mathbf{B} sabiendo que el campo eléctrico está dirigido según el eje Y .

Sol. $\mathbf{E}(x, t) = 67.26 \times 10^{-3} \cos(2\pi \times 10^{10} t - 200\pi x/3 + \phi) \hat{y} \text{ V/m}$,
 $\mathbf{B}(x, t) = 22.42 \times 10^{-11} \cos(2\pi \times 10^{10} t - 200\pi x/3 + \phi) \hat{z} \text{ T}$.

2.5: Cierta pulso de campo electromagnético puede asimilarse a una onda plana cuyos campos son $\mathbf{E}(x, t) = E_0 e^{-(x-ct)^2} \hat{y}$ (V/m) y $\mathbf{B}(x, t) = B_0 e^{-(x-ct)^2} \hat{z}$ (T). Demostrar que ambos campos verifican la ecuación de onda y obtener la relación entre E_0 y B_0 sabiendo que de acuerdo con la ley de Faraday

debe cumplirse que $\partial E_y(x, t)/\partial x = -\partial B_z(x, t)/\partial t$.

Sol.: $E_0 = cB_0$.

2.6: La antena de un receptor radioeléctrico es equivalente a una barra conductora de 2 m de altura y está orientada paralelamente al campo eléctrico de la OEM que se desea sintonizar. Si la tensión eficaz entre los extremos de la antena al recibir la onda es de 4 mV, determinense las amplitudes de los campos eléctrico y magnético de la onda sintonizada, así como la potencia media por unidad de área transportada por la onda.

Sol.: $E_e = 2 \times 10^{-3}$ V/m; $B_e = 0.666 \times 10^{-11}$ T, $I = 10^{-8}$ W/m².

2.7: En la superficie de la Tierra, el flujo solar medio aproximado es de 0,75 kW/m². Se desea diseñar un sistema de conversión de energía solar a eléctrica para que proporcione una potencia eléctrica de 25 kW que permita cubrir las necesidades de una casa. Si el sistema tiene una eficacia del 30%, ¿cuál será el área necesaria de los colectores solares, supuestos que son absorbentes perfectos?

Sol. 111 m².

2.8: Un pulso de láser tiene una energía de 20 J y un radio de haz de 2 mm. La duración del pulso es de 10 ns y la densidad de energía es constante dentro del pulso. **a)** ¿Cuál es la longitud espacial del pulso? **b)** ¿Cuál es la densidad de energía dentro del mismo? **c)** Hallar los valores de la amplitud de los campos eléctrico y magnético.

Sol.: **a)** 3 m; **b)** 5.31×10^5 J/m³; **c)** $E_0 = 3.46 \times 10^8$ V/m, $B_0 = 1.15$ T.

2.9: El campo eléctrico de una onda electromagnética oscila en la dirección z, viniendo su vector de Poynting dado por

$$\mathbf{S}(x, t) = -(100 \text{ W/m}^2) \cos^2[10x + (3 \times 10^9)t] \hat{\mathbf{x}},$$

donde x está en metros y t en segundos. **a)** ¿En qué dirección se propaga la onda? **b)** Calcular la longitud de onda y la frecuencia. **c)** Hallar los campos eléctrico y magnético. **d)** Comprobar que la densidad de energía eléctrica y magnética son idénticas.

Sol.: **a)** sentido negativo de x ; **b)** $\lambda = 0.620$ m, $f = 4.77 \times 10^8$ Hz;

c) $\mathbf{E} = (194 \text{ V/m}) \cos[10x + (3 \times 10^9)t] \hat{\mathbf{z}}$, $\mathbf{B} = (0.647 \times 10^{-6} \text{ T}) \cos[10x + (3 \times 10^9)t] \hat{\mathbf{y}}$.

2.10: El campo eléctrico de una onda electromagnética armónica plana tiene la expresión $\mathbf{E}(z, t) = 3 \times 10^{-3} \cos(kz - 2\pi \times 10^8 t) \hat{\mathbf{y}}$ (V/m). Determinense: **a)** la longitud de onda, frecuencia, periodo y número de onda; **b)** el campo magnético, \mathbf{B} , así como el vector de Poynting, \mathbf{S} , y la intensidad de onda, I .

Sol.: **a)** $\lambda = 3$ m, $f = 100$ MHz, $T = 10$ ns, $k = 2\pi/3 \text{ m}^{-1}$;

b) $\mathbf{B}(z, t) = -0.01 \cos(2\pi z/3 - 2\pi \times 10^8 t) \hat{\mathbf{x}}$ nT,

$\mathbf{S}(z, t) = 0.0239 \cos^2(2\pi z/3 - 2\pi \times 10^8 t) \hat{\mathbf{z}}$ $\mu\text{W/m}^2$, $I = \langle S \rangle = 0.01195 \mu\text{W/m}^2$.

2.11: Una OEM armónica plana de longitud de onda $\lambda = 6$ m se propaga en el sentido negativo del eje de las X siendo su campo magnético $\mathbf{B}(x, t) = 2 \times 10^{-10} \cos(\omega t + kx + \pi/4) \hat{\mathbf{y}}$ T. Determinense: **a)** el número de ondas, la frecuencia y el periodo de la onda; **b)** las expresiones del campo eléctrico, \mathbf{E} , y del vector de Poynting, \mathbf{S} , así como la intensidad de onda, I .

Sol.: **a)** $k = \pi/3 \text{ m}^{-1}$, $f = 50$ MHz, $T = 20$ ns;

b) $\mathbf{E}(x, t) = 60 \times 10^{-3} \cos(\pi \times 10^8 t + kx + \pi/4) \hat{\mathbf{z}}$ V/m,

$\mathbf{S}(x, t) = -(30/\pi) \cos^2(\pi \times 10^8 t + kx + \pi/4) \hat{\mathbf{x}}$ $\mu\text{W/m}^2$, $I = \langle S \rangle = 15/\pi \mu\text{W/m}^2$.

Constantes: $c = 3 \times 10^8$ m/s, $\mu_0 = 4\pi \times 10^{-7}$ H/m, $\epsilon_0 = 8.854 \times 10^{-12}$ F/m.

Tema 3

Óptica Geométrica

3.1. Introducción

En el tema anterior se ha discutido que lo que se conoce usualmente como *luz* es una radiación electromagnética de longitud de onda comprendida entre 400 y 700 nm. En este sentido, el estudio de la propagación de la luz en los diversos medios y la interacción de ésta con la materia debería estudiarse como una particularización del análisis de ondas electromagnéticas. No obstante, cuando se estudia la *propagación* de ondas electromagnéticas, se encuentra que esta propagación se rige por simples *leyes geométricas* que, en general, no dependen de la longitud de onda y de la intensidad de la radiación. (Esto no sucede cuando lo que se estudia es la interacción de la radiación electromagnética con la materia). El hecho anterior permite estudiar la propagación de la radiación luminosa mediante el concepto de *rayos* que se propagan en línea recta en medios uniformes y que están relacionados por las leyes de la reflexión y refracción de Snell en las superficies de separación. Para que esta consideración básica de la Óptica Geométrica sea válida, es fundamental que las dimensiones de los objetos involucrados sean mucho mayores que la longitud de onda de la luz¹, lo cual se cumple generalmente para la mayoría de los objetos que usamos. Bajo estas suposiciones, la Óptica Geométrica responderá de una forma muy satisfactoria a la mayoría de las cuestiones relacionadas con los instrumentos ópticos, proporcionando además una teoría mucho más simple que el tratamiento directo de ondas electromagnéticas.

En consecuencia, podemos concluir que la Óptica Geométrica tratará sobre las cuestiones relacionadas con la propagación de la luz, determinando las trayectorias de la energía radiante a través de los distintos medios materiales o bien deduciendo la disposición adecuada de los medios para que la luz siga una trayectoria determinada.

3.2. Consideraciones previas

Para determinar las leyes geométricas básicas que sigue la radiación luminosa al propagarse debemos hacer una serie de consideraciones que se exponen a continuación.

¹ Este hecho permite ignorar los fenómenos de difracción que, como ya se ha discutido, son significativos únicamente cuando la dimensión de los objetos es del orden de la longitud de onda

3.2.1. Rayo luminoso

La radiación luminosa al propagarse sigue unas determinadas trayectorias que se denominan *rayos de luz*. Evidentemente, el rayo luminoso es una abstracción geométrica de una situación física que se daría cuando de la radiación total emitida por un emisor sólo consideramos aquella parte que atraviesa un pequeño orificio situado en una pantalla opaca.

3.2.2. Índice de refracción

En el tema anterior se discutió que la velocidad de la luz, v , en un medio material caracterizado por una permitividad eléctrica relativa ϵ_r y una permeabilidad magnética relativa μ_r venía dado por

$$v = \frac{1}{\sqrt{\epsilon_0 \epsilon_r \mu_0 \mu_r}} = \frac{c}{\sqrt{\epsilon_r \mu_r}} \quad (3.1)$$

donde $c = 1/\sqrt{\epsilon_0 \mu_0} \approx 3 \times 10^8$ m/s era la velocidad de la luz en el espacio libre.

Desde el punto de vista de la óptica, un medio material quedará caracterizado por el *índice de refracción*, n , que es la relación entre la velocidad de la luz en el espacio libre y en el medio material. Este índice vendrá dado, en virtud de (3.1), por

Índice de refracción

$$n = \frac{c}{v} = \sqrt{\epsilon_r \mu_r} > 1. \quad (3.2)$$

En función de las características del índice de refracción, los medios pueden clasificarse como siguen:

- MEDIOS HOMOGÉNEOS E ISÓTROPOS.

En estos medios, el índice de refracción es el mismo en todo el medio: $n \neq n(\mathbf{r})$ (homogéneo) y además no depende de la dirección del espacio (isótropo). Ejemplos de este tipo de medios son el aire, el vidrio, el agua, etc.

- MEDIOS HETEROGÉNEOS.

Cuando el índice de refracción varía de punto a punto del medio: $n = n(\mathbf{r})$ se dice que el medio es heterogéneo o también medio GRIN (de GRadiente de Índice). Esto le ocurre, por ejemplo, a la atmósfera, ya que al ser distinta su composición con respecto a la altura, tendrá un índice de refracción variable con ésta.

- MEDIOS ANISÓTROPOS.

Estos medios presentan la característica de que la velocidad de luz es diferente para distintas direcciones del espacio, provocando así que el índice de refracción dependa de la dirección del espacio. En este caso, el índice de refracción no puede venir definido por un escalar sino que ha de definirse como un tensor. Esta situación se encuentra, por ejemplo, en todos aquellos cristales que no cristalizan en el sistema cúbico (como el zafiro).

Es interesante señalar también que, en general, el índice de refracción en los medios materiales depende de la longitud de onda: $n = n(\lambda)$, en cuyo caso, se dice que existe **dispersión cromática**. La dispersión cromática proviene del hecho señalado en el Apartado 1.8 acerca de que la velocidad de fase depende del número de ondas. A estos medios se les denominaba en general medios dispersivos y ahora, con respecto a la luz, diremos que presentan dispersión cromática. El único medio que no presenta estrictamente dispersión cromática es el vacío. No obstante, los medios gaseosos de baja densidad apenas presentan dispersión apreciable por lo que el aire se considerará en general un medio no dispersivo.

3.2.3. Camino óptico

Si en un medio homogéneo e isótropo de índice de refracción n , la luz recorre un trayecto de longitud s , el *camino óptico*, \mathcal{S} , recorrido por la luz se define como

$$\mathcal{S} = ns. \quad (3.3)$$

Si el rayo de luz se propaga a través de N medios homogéneos, cada uno de ellos caracterizado por un índice de refracción n_i y en cada uno de ellos recorre un trayecto de longitud s_i , entonces el camino óptico viene dado por

$$\mathcal{S} = \sum_{i=1}^N n_i s_i. \quad (3.4)$$

En el caso de que el rayo luminoso viaje desde un punto A hasta un punto B en un medio GRIN, el camino óptico puede entonces definirse como

$$\mathcal{S} = \int_A^B n ds. \quad (3.5)$$

Teniendo en cuenta la definición de n dada en (3.2), (3.5) puede también reescribirse como

$$\mathcal{S} = c \int_A^B \frac{ds}{v} = c \int_{t_A}^{t_B} dt = ct, \quad (3.6)$$

lo que nos dice que el camino óptico es c veces el tiempo que tarda el rayo en viajar desde A hasta B .

3.2.4. Principio de Fermat

Toda la Óptica Geométrica puede deducirse del principio de Fermat, que en su forma más general puede enunciarse como sigue:

El camino óptico de la trayectoria real seguida por la luz en su propagación desde A hasta B es estacionario con respecto a variaciones de ese camino.

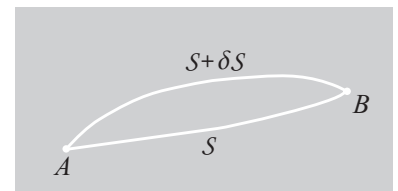
El hecho de que el camino óptico de la trayectoria *real* sea estacionario quiere decir que este camino óptico puede ser un punto de inflexión de tangente horizontal, un máximo o un mínimo con respecto a otros caminos ópticos correspondientes a otras trayectorias que unan A con B . En forma matemática, este principio puede expresarse como

$$\delta \mathcal{S} = \delta \int_A^B n ds = 0. \quad (3.7)$$

En muchas situaciones prácticas, el carácter estacionario viene dado por la condición de mínimo y así el principio de Fermat muchas veces se enuncia de la siguiente manera:

el rayo luminoso al ir desde un punto a otro sigue aquella trayectoria cuyo camino óptico sea menor.

Esta versión del principio de Fermat implica que la trayectoria seguida en la realidad por la luz será aquella que tienda a minimizar, en la medida de lo posible, el



recorrido por las zonas donde el índice de refracción sea mayor. Dado que en el caso de medios gaseosos, el índice de refracción más alto corresponde a las regiones más densas, la interpretación anterior (ver Fig. 3.1) permite explicar por qué podemos “ver” el Sol aún después de situarse bajo la línea del horizonte. En la Fig. 3.1 se observa cómo el rayo minimiza su recorrido en las regiones infe-

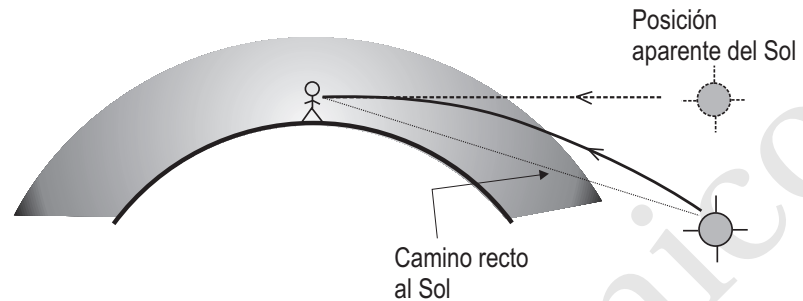


FIGURA 3.1: Posición aparente del sol en su puesta debido a la trayectoria curvada de los rayos solares.

riores, más densas, de la atmósfera. Esta misma explicación permite igualmente entender por qué a veces una carretera vista desde un ángulo rasante refleja los alrededores y se comporta entonces como un espejo (al estar más caliente la región cercana a la superficie, ésta es menos densa y, en consecuencia, el rayo se curva hacia esta región).

3.3. Leyes de la Óptica Geométrica

A partir del principio de Fermat deduciremos a continuación las leyes básicas de la Óptica Geométrica. Para ello aplicaremos este principio a un rayo de luz que saliendo de A en un medio de índice n_1 llega a la superficie de separación S con un segundo medio de índice n_2 exactamente en el punto I , donde se refracta para llegar en el segundo medio al punto B . Según el esquema mostrado en la Figura 3.2 el camino óptico, \mathcal{S} , viene dado por

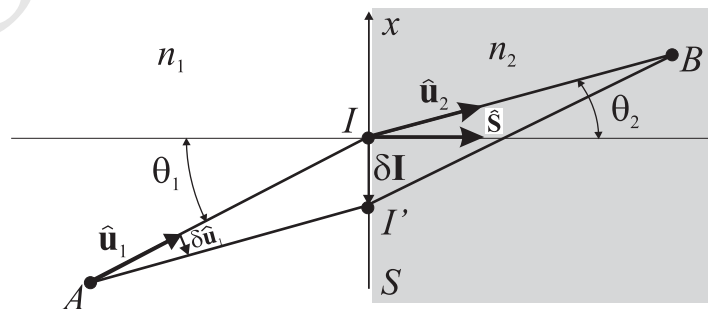


FIGURA 3.2: Aplicación del principio de Fermat a la refracción en una superficie de separación de dos medios

$$\begin{aligned}\mathcal{S} &= n_1 \overline{AI} + n_2 \overline{IB} \\ &= n_1 \hat{\mathbf{u}}_1 \cdot (\mathbf{I} - \mathbf{A}) + n_2 \hat{\mathbf{u}}_2 \cdot (\mathbf{B} - \mathbf{I})\end{aligned}\quad (3.8)$$

donde \overline{AI} representa la distancia entre el punto A y el punto I , los vectores $\hat{\mathbf{u}}_i$ son los vectores unitarios en la dirección del rayo en cada medio y \mathbf{A} , \mathbf{B} e \mathbf{I} son

los vectores de posición de los puntos correspondientes. Según el principio de Fermat, una pequeña variación del camino óptico debe ser nula, $\delta\mathcal{S} = 0$, esto es,

$$\begin{aligned}\delta\mathcal{S} &= n_1\delta\hat{\mathbf{u}}_1 \cdot (\mathbf{I} - \mathbf{A}) + n_1\hat{\mathbf{u}}_1 \cdot (\delta\mathbf{I} - \delta\mathbf{A}) + \\ &\quad n_2\delta\hat{\mathbf{u}}_2 \cdot (\mathbf{B} - \mathbf{I}) + n_2\hat{\mathbf{u}}_2 \cdot (\delta\mathbf{B} - \delta\mathbf{I}) \\ &= (n_1\hat{\mathbf{u}}_1 - n_2\hat{\mathbf{u}}_2) \cdot \delta\mathbf{I} = 0,\end{aligned}\quad (3.9)$$

donde se ha usado el hecho de que $\delta\hat{\mathbf{u}}_i$ es perpendicular a las trayectorias de los rayos y que $\delta\mathbf{A} = \delta\mathbf{B} = 0$ puesto que los puntos inicial y final no varían. La expresión (3.9) nos dice entonces que el vector $(n_1\hat{\mathbf{u}}_1 - n_2\hat{\mathbf{u}}_2)$ debe ser perpendicular a $\delta\mathbf{I}$ y puesto que está contenido en la superficie de separación S , tenemos que

$$n_1\hat{\mathbf{u}}_1 - n_2\hat{\mathbf{u}}_2 = q\hat{\mathbf{S}}, \quad (3.10)$$

donde $q\hat{\mathbf{S}}$ es cierto vector normal a la superficie S en el punto I . A partir de la expresión (3.10) podemos deducir lo siguiente:

- *El rayo incidente, el refractado y la normal a la superficie están en el mismo plano (plano de incidencia).*

- **Ley de Snell para la refracción.**

Multiplicando escalarmente (3.10) por el vector unitario en la dirección x , encontramos que

$$(n_1\hat{\mathbf{u}}_1 - n_2\hat{\mathbf{u}}_2) \cdot \hat{\mathbf{x}} = q\hat{\mathbf{S}} \cdot \hat{\mathbf{x}}. \quad (3.11)$$

Puesto que el segundo miembro es igual a cero ($\hat{\mathbf{x}} \perp \hat{\mathbf{S}}$), (3.11) establece que

$$n_1\hat{\mathbf{u}}_1 \cdot \hat{\mathbf{x}} = n_2\hat{\mathbf{u}}_2 \cdot \hat{\mathbf{x}}, \quad (3.12)$$

o equivalentemente,

$$n_1 \sin\theta_1 = n_2 \sin\theta_2. \quad (3.13)$$

Ley de Snell

La expresión (3.13) se conoce como ley de Snell para la refracción.

- **Ley de la reflexión.**

Si la superficie de separación de los medios se comportase igualmente como una superficie reflectora, de modo que el rayo 2 se reflejase en el medio 1, entonces la aplicación del procedimiento anterior nos diría (dado que $n_1 = n_2$) que

$$\theta_1 = \theta_2 \quad (3.14)$$

o lo que es lo mismo: el ángulo que forma el rayo reflejado con la normal a la superficie es idéntico al formado por el rayo incidente.

El ángulo que forma el rayo reflejado con la normal a la superficie es idéntico al formado por el rayo incidente.

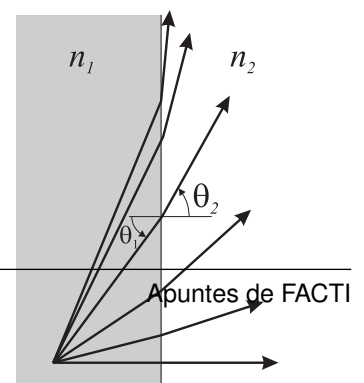
- **Las trayectorias de los rayos luminosos son reversibles.**

Si el rayo de luz viajase en sentido contrario al supuesto inicialmente, las leyes anteriores quedarían invariantes puesto que los únicos cambios estarían en el signo de $\hat{\mathbf{u}}_1$, $\hat{\mathbf{u}}_2$ y $\hat{\mathbf{S}}$, lo cual no tendría ningún efecto en las leyes de reflexión y refracción.

Reflexión total

Una consecuencia adicional e interesante que podemos deducir de la ley de refracción (3.13) es que si

$$n_1 > n_2 \Rightarrow \theta_2 > \theta_1,$$



pudiéndose llegar a la situación en la que $\theta_2 = \pi/2$. Este caso se dará cuando el rayo en el medio 1 incide en la superficie de separación con un ángulo $\hat{\theta}_1$ determinado por

$$n_1 \operatorname{sen} \hat{\theta}_1 = n_2, \quad (3.15)$$

esto es, cuando

$$\hat{\theta}_1 = \operatorname{arcsen} \left(\frac{n_2}{n_1} \right). \quad (3.16)$$

Para ángulos de incidencia $\theta_1 \geq \hat{\theta}_1$ *no existe refracción* y, por tanto, decimos que estamos en una situación de reflexión total.

3.4. Componentes ópticos simples

En este apartado estudiaremos el comportamiento óptico, desde el punto de vista de la Óptica Geométrica, de algunos de los componentes ópticos más utilizados en la práctica.

Un **sistema óptico** puede definirse como un conjunto de superficies que separan medios distintos. Aquí nos centraremos en particular en algunos sistemas ópticos sencillos que están formados por superficies esféricas cuyos centros están alineados y que se denominan **sistemas ópticos centrados**. Supondremos además que estos sistemas ópticos son **astigmáticos**, esto es, que para cada cono de rayos que parte de cierto punto del sistema existe un punto del sistema al cual llega este cono de rayos.

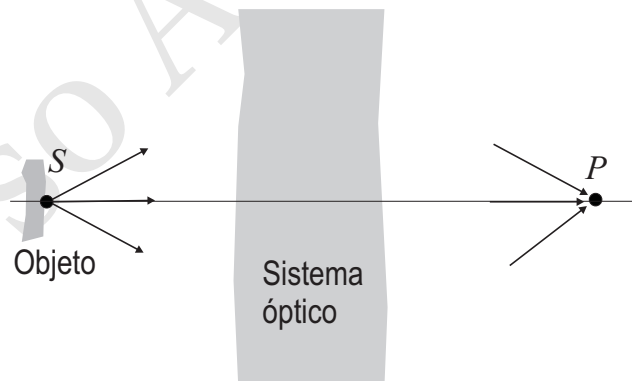


FIGURA 3.3: Esquema de un sistema óptico centrado astigmático

Aplicaremos ahora la teoría de rayos a estos sistemas ópticos astigmáticos centrados bajo la **aproximación paraxial**. Esta aproximación asume que, para un sistema centrado de eje óptico z , todos los rayos son casi paralelos al eje óptico. Esto implica que al ser los ángulos tan pequeños, los senos y tangentes de dichos ángulos pueden sustituirse directamente por los ángulos, esto es,

Aproximación paraxial

$$\theta \ll \Rightarrow \operatorname{sen} \theta \approx \theta, \quad \tan \theta \approx \theta.$$

Teniendo en cuenta las anteriores consideraciones, a continuación estudiaremos algunos componentes ópticos simples: espejos, interfases de separación, lentes y guías de ondas.

3.4.1. Espejos

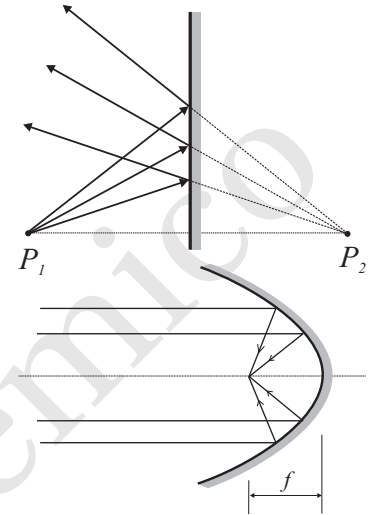
Los espejos se comportan como superficies reflectoras perfectas, cumpliéndose la ley (3.14).

Espejo plano

En un espejo plano, los rayos se reflejan de modo que los rayos reflejados que provienen de un punto P_1 parece como si vinieran de un punto P_2 (*punto imagen*) situado detrás del espejo. En este sentido, los objetos reflejados en este espejo aparecen como si estuvieran situados detrás del espejo plano.

Espejo parabólico

Este espejo, formado por una superficie parabólica de revolución, tiene la importante propiedad de que todos los rayos incidentes *paralelos* al eje del espejo se reflejan en un punto llamado **foco**. Estos espejos se usan a menudo en los telescopios para recoger todos los rayos paralelos incidentes. También pueden usarse para emitir rayos paralelos provenientes de un punto emisor.



Espejo elíptico

En este espejo, todos los rayos provenientes de uno de los focos de la elipse se encuentran, tras reflejarse, en el otro foco.

Espejo esférico

Los espejos esféricos son más fáciles de fabricar que los parabólicos y los elípticos pero no presentan la propiedad de enfoque de los anteriores. En concreto, los rayos que inciden paralelamente al eje del espejo no se encuentran en un solo punto del eje sino sobre una curva que se denomina *caústica*.

Aunque, en general, este espejo no enfoca en un solo punto, los rayos *paraxiales* sí que son enfocados en un punto F situado a una distancia $(-R)/2$ del centro del espejo C . Por convenio, tomaremos R negativo para espejos cóncavos y R positivo para espejos convexos.

Para verificar la anterior propiedad, consideremos la Figura 3.4, donde se ha tomado el siguiente convenio para los ángulos: *ángulo positivo para rayos que ascienden desde el eje y ángulo negativo para rayos que descienden hacia el eje*.

En la Fig. 3.4 podemos observar que

$$\theta_1 = \theta_0 - \theta \quad (3.17)$$

$$(-\theta_2) = \theta_0 + \theta, \quad (3.18)$$

de donde se deduce que

$$(-\theta_2) + \theta_1 = 2\theta_0 \quad (3.19)$$

y, como bajo la aproximación paraxial,

$$\theta_0 = \frac{y}{(-R)}, \quad (3.20)$$

encontramos que

$$(-\theta_2) + \theta_1 = \frac{2y}{(-R)}. \quad (3.21)$$

Convenio de signos:

$R < 0$ espejo cóncavo

$R > 0$ espejo convexo

$\theta > 0$ rayo ascendente desde eje

$\theta < 0$ rayo que desciende al eje.



$$\alpha = \pi/2 - \theta_0$$

$$\alpha + \theta = \pi/2 - \theta_1$$

$$\theta_1 = \theta_0 - \theta$$

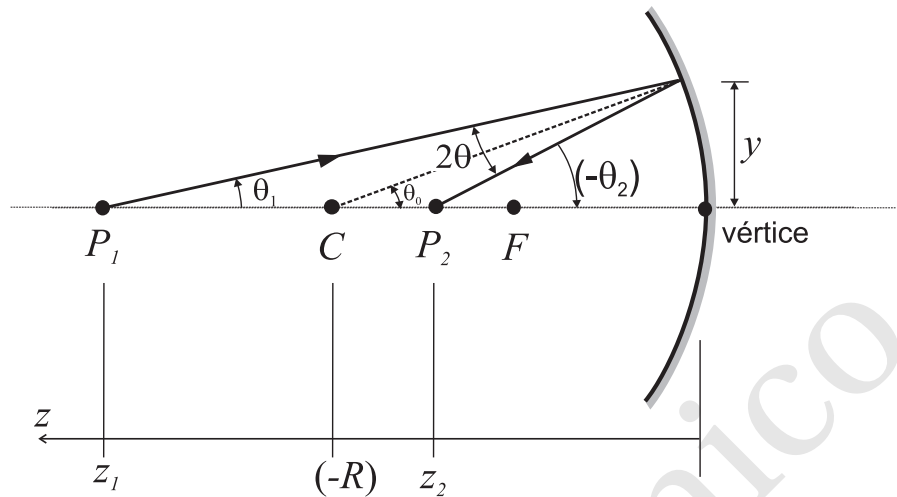


FIGURA 3.4: Reflexión de rayos paraxiales en un espejo esférico **cóncavo** de radio $(-R)$.

Análogamente, la aproximación paraxial permite escribir que

$$\theta_1 = \frac{y}{z_1} \quad \text{y} \quad (-\theta_2) = \frac{y}{z_2}$$

y, consecuentemente, expresar (3.21) como

$$\frac{1}{z_1} + \frac{1}{z_2} = \frac{1}{f}, \quad (3.22)$$

siendo f la *distancia focal* de valor

$$f = \frac{(-R)}{2}. \quad (3.23)$$

En la expresión (3.22) puede verse que todos los rayos que entren paralelos al eje del espejo ($z_1 \rightarrow \infty$) convergerán en el punto F mostrado en la Figura 3.4 determinado por $z = f$, demostrándose así la propiedad de focalización de los espejos esféricos bajo aproximación paraxial.

EJEMPLO 3.1 Delante de un espejo cóncavo de 50 cm de distancia focal y a 25 cm de su vértice se encuentra un objeto cuya altura, perpendicular al eje, es de 1 cm. Calcular la posición y el tamaño de la imagen suponiendo que los rayos incidentes corresponden a la zona paraxial. Repetir el problema considerando que el espejo es convexo.

■ **Espejo cóncavo.**

La posición de la imagen se puede calcular a partir de la siguiente ecuación para espejos esféricos

$$\frac{1}{z_1} + \frac{1}{z_2} = \frac{1}{f}$$

donde f es la distancia focal del espejo ($f = 50$ cm), z_1 es la posición del objeto (25 cm) y z_2 es la posición de la imagen del objeto. De la ecuación anterior obtenemos $z_2 = -25$ cm. La imagen es virtual, aparece “detrás” del espejo. El aumento lateral del espejo es

$$m = -\frac{z_2}{z_1} = -(-25)/25 = 1,$$

por lo que la imagen tendrá 1 cm de tamaño.

- **Espejo convexo.**

En este caso, la distancia focal es $f = -50$ cm, por lo que, realizando un cálculo análogo al del apartado anterior, obtenemos $z_2 = -50/3$ cm. La imagen es virtual y de $4/3$ cm de tamaño, puesto que el aumento lateral es

$$m = -\frac{z_2}{z_1} = 50/75 = 2/3.$$

3.4.2. Interfases planas

Si una interfase plana separa dos medios de índice de refracción n_1 y n_2 , la relación entre los ángulos incidente y refractado viene dada por la ley de Snell (3.13). Esta ley dice que según sea la relación n_1/n_2 podemos distinguir dos casos:

- **Refracción externa** $n_1 < n_2$.

Dado que $n_1 < n_2$ tenemos que $\theta_2 < \theta_1$ y por tanto los rayos refractados se acercan a la normal a la superficie, tal como muestra la Figura 3.5(a).

- **Refracción interna** $n_1 > n_2$.

Según la Figura 3.5(b), si $n_1 > n_2$ tenemos que $\theta_2 > \theta_1$ y por tanto los rayos refractados se alejan de la normal a la superficie.

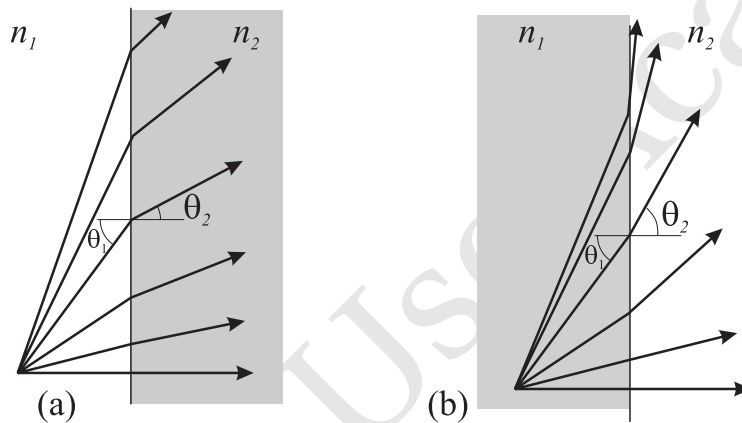


FIGURA 3.5: Refracción externa e interna.

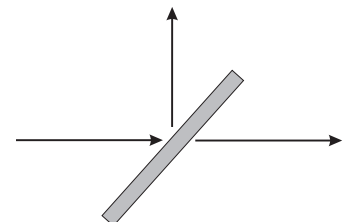
La ley de Snell, bajo la aproximación paraxial, puede reescribirse como

$$n_1\theta_1 = n_2\theta_2. \quad (3.24)$$

Ley de Snell bajo aproximación paraxial

Divisores de haz

Son componentes ópticos que dividen el haz incidente en un haz reflejado y un haz refractado. Estos componentes se construyen a menudo depositando una fina película metálica semitransparente sobre un sustrato dieléctrico.



3.4.3. Interfases esféricas

Para estudiar la refracción en una superficie esférica usaremos la ley de Snell aplicándola al sistema mostrado en la Figura 3.6. En esta figura se cumple que

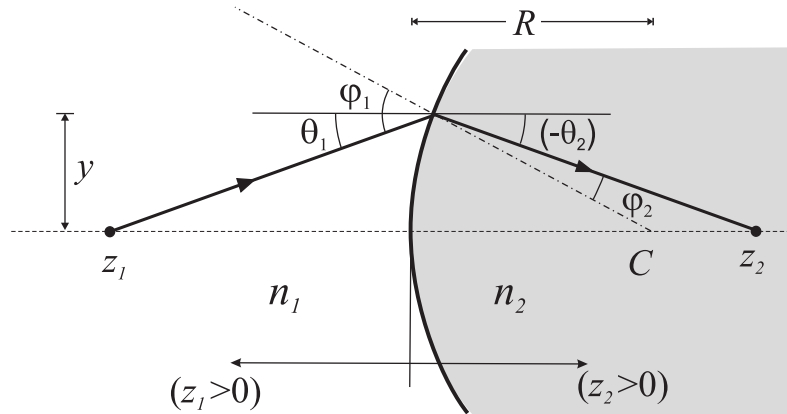


FIGURA 3.6: Refracción de rayos paraxiales en una interfase esférica **convexa** de radio $R > 0$.

$$n_1 \varphi_1 = n_2 \varphi_2 \quad (3.25)$$

y teniendo en cuenta que

$$\theta_1 = \frac{y}{z_1} \quad \varphi_1 - \theta_1 = \frac{y}{R} \quad (3.26)$$

$$(-\theta_2) = \frac{y}{z_2} \quad \varphi_2 + (-\theta_2) = \frac{y}{R}, \quad (3.27)$$

encontramos la siguiente relación:

$$n_1 \left(\frac{y}{R} + \theta_1 \right) = n_2 \left(\frac{y}{R} + \theta_2 \right), \quad (3.28)$$

que puede reescribirse como

$$n_1 \left(\frac{y}{R} + \frac{y}{z_1} \right) = n_2 \left(\frac{y}{R} + \frac{y}{z_2} \right). \quad (3.29)$$

Reagrupando términos y simplificando, la ecuación anterior puede expresarse como

$$\frac{n_1}{z_1} + \frac{n_2}{z_2} = \frac{n_2 - n_1}{R}. \quad (3.30)$$

Es también interesante obtener el ángulo refractado θ_2 en función del ángulo incidente θ_1 y la coordenada y . Utilizando (3.25) y dado que

$$\begin{aligned} (-\theta_2) &= \frac{y}{R} - \varphi_2 \\ &= \frac{y}{R} - \frac{n_1}{n_2} \varphi_1 \\ &= \frac{y}{R} - \frac{n_1}{n_2} \theta_1 - \frac{n_1 y}{n_2 R} \\ &= -\frac{n_1}{n_2} \theta_1 + \left(1 - \frac{n_1}{n_2} \right) \frac{y}{R}, \end{aligned} \quad (3.31)$$

obtenemos finalmente que

$$\theta_2 = \frac{n_1}{n_2} \theta_1 - \frac{n_2 - n_1}{n_2} \frac{y}{R}. \quad (3.32)$$

3.4.4. Lentes esféricas

Una lente esférica es un trozo de material de índice de refracción n limitado por dos superficies esféricas de radios R_1 y $(-R_2)$. Si la distancia Δ es pequeña en comparación con el radio de las esferas, entonces despreciaremos este espesor y diremos entonces que la lente es una **lente delgada** (la lente se comportaría entonces como cierta interfase de separación).

Para la lente delgada de la Figura 3.7 se cumple, de acuerdo con la expresión

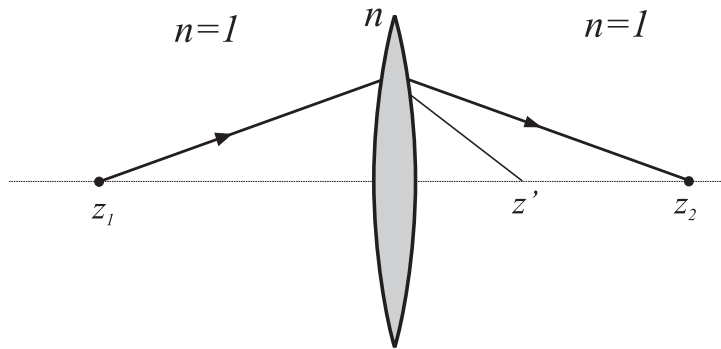


FIGURA 3.7: Refracción en una lente esférica delgada.

(3.30) al aplicarla reiteradamente a cada una de las dos superficies esféricas, que

$$\frac{1}{z_1} + \frac{n}{z'} = \frac{n-1}{R_1} \tag{3.33}$$

$$\frac{n}{(-z')} + \frac{1}{z_2} = \frac{1-n}{R_2} \tag{3.34}$$

Despejando n/z' en (3.33) y sustituyendo en (3.34) se llega a

$$-\frac{n-1}{R_1} + \frac{1}{z_1} + \frac{1}{z_2} = \frac{1-n}{R_2} \tag{3.35}$$

expresión que puede reescribirse como

$$\frac{1}{z_1} + \frac{1}{z_2} = \frac{1}{f} \tag{3.36}$$

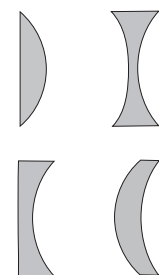
Ecuación de la lente delgada

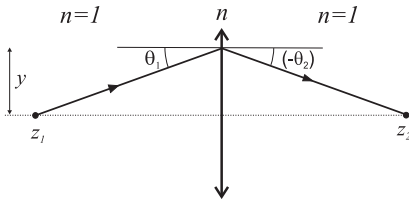
que es la *ecuación de la lente delgada* donde la distancia focal, f , de la lente esférica viene dada por

$$\frac{1}{f} = (n-1) \left(\frac{1}{R_1} - \frac{1}{R_2} \right) \tag{3.37}$$

Distancia focal de una lente esférica

(Recordemos que según el convenio establecido $(-R_2)$ será un número positivo en nuestra lente. Notemos además que la distancia focal no varía si invertimos la lente). En la expresión anterior, la distancia focal f de la lente determina la posición del foco, esto es, el punto donde convergen los rayos que entran paralelos al eje de la lente. Es interesante señalar que la expresión (3.36) junto con (3.37) sigue siendo válida para otro tipo de lente distinto al mostrado en la Figura 3.7. La lente mostrada en esta figura pertenece al tipo conocido como *convexa, convergente o positiva*, que son más gruesas en el centro. En concreto es una lente biconvexa. Otro tipo de lente convergente son las lentes plano convexas, formadas por una superficie plana y otra convexa. El tipo de lentes que son más delgadas por el centro se conocen como lentes *cóncavas, divergentes o negativas*. Claramente el tipo de lente se manifiesta en la ecuación (3.36) a través del valor de la distancia focal,





reflejándose este hecho en los distintos signos (o bien $R \rightarrow \infty$ para la superficie plana) que toman los radios en (3.37).

Si queremos obtener el ángulo de refracción θ_2 en la figura adjunta, podemos aplicar reiteradamente la expresión (3.32) a cada una de las superficies esféricas de la lente, esto es,

$$\theta' = \frac{1}{n}\theta_1 - \frac{y}{R_1} \frac{n-1}{n} \quad (3.38)$$

$$\theta_2 = n\theta' - (1-n) \frac{y}{R_2}. \quad (3.39)$$

Sustituyendo θ' en (3.39) se obtiene que

$$\begin{aligned} \theta_2 &= n \left(\frac{1}{n}\theta_1 - \frac{y}{R_1} \frac{n-1}{n} \right) + (n-1) \frac{y}{R_2} \\ &= \theta_1 - y \frac{n-1}{R_1} + y \frac{n-1}{R_2} \end{aligned} \quad (3.40)$$

y por tanto, usando (3.37), obtenemos finalmente que

$$\theta_2 = \theta_1 - \frac{y}{f}. \quad (3.41)$$

Formación de imágenes finitas

Una cuestión importante respecto a las lentes es la formación de imágenes de objetos finitos. En concreto, y con referencia a la Fig. 3.8, buscamos la relación que existe entre las alturas de un objeto lineal y de su imagen; esto es, la relación

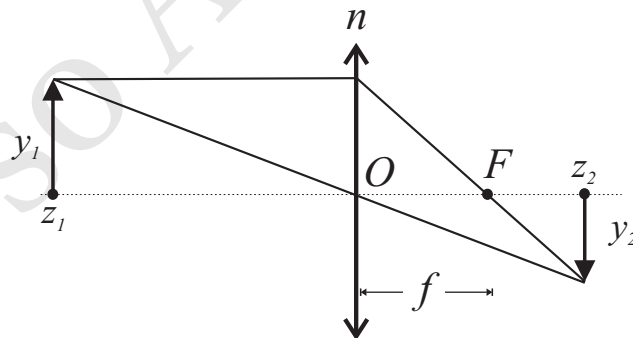


FIGURA 3.8: Formación de imágenes con una lente de distancia focal f .

entre y_2 e y_1 . Dado que la lente considerada es un sistema astigmático, dos rayos cualesquiera son suficientes para fijar la posición de un punto que no está en el eje. Una posible manera de hacer esto consiste en considerar un rayo que entra paralelo al eje de la lente y otro que pasa por el centro de la lente. Debido a la propiedad focal de la lente, el rayo paralelo pasará por el punto F situado una distancia f de la lente. Por otra parte, el rayo que incide en el centro de la lente no sufre desviación debido a que sufre dos refracciones en superficies que son localmente planas y *paralelas*. Según el razonamiento anterior encontramos que

$$\frac{y_1}{z_1} = \frac{(-y_2)}{z_2} \quad (3.42)$$

y por tanto

$$y_2 = -\frac{z_2}{z_1} y_1, \quad (3.43)$$

de donde encontramos que el **aumento lateral o transversal**, M_T , viene dado por

$$M_T \equiv \frac{y_2}{y_1} = -\frac{z_2}{z_1} \quad (3.44)$$

Es importante notar que cualquier punto del plano $z = z_1$ tiene su imagen en el plano $z = z_2$, por lo que ambos planos se denominan **planos conjugados**.

EJEMPLO 3.2 Una lente de vidrio tiene 20 cm de distancia focal en el aire. Calcular su distancia focal en el agua. El índice de refracción del vidrio es 3/2; el del agua es 4/3.

Suponiendo que se trata de una lente delgada, su comportamiento queda descrito por la siguiente ecuación:

$$(n-1)\left(\frac{1}{R_1} - \frac{1}{R_2}\right) = \frac{1}{f}$$

donde f es la distancia focal de la lente, R_1 y R_2 son los radios de sus dos caras, y n es el índice de refracción relativo del vidrio de la lente respecto al medio en el que se encuentra. Cuando la lente se encuentra en el aire, suponiendo que el índice de refracción del aire es 1, la anterior ecuación puede escribirse del siguiente modo:

$$(3/2-1)\left(\frac{1}{R_1} - \frac{1}{R_2}\right) = \frac{1}{f_1}$$

Cuando la lente esté en el agua, la ecuación tendrá la siguiente forma:

$$\left(\frac{3/2}{4/3}-1\right)\left(\frac{1}{R_1} - \frac{1}{R_2}\right) = \frac{1}{f_2}$$

Combinando las dos expresiones anteriores, tenemos

$$f_1(3/2-1) = f_2\left(\frac{3/2}{4/3}-1\right),$$

de donde obtenemos el siguiente valor para f_2 , distancia focal de la lente en el agua:

$$f_2 = 20 \frac{(3/2-1)}{9/8-1} = 80 \text{ cm}$$

Por tanto, la distancia focal de la lente en el agua, con índice de refracción 4/3, es cuatro veces mayor que en el aire.

3.4.5. Guías de luz

Tal como muestra la figura adjunta, la luz puede ser guiada desde un sitio hasta otro por medio de lentes o espejos. Dado que en las lentes existe inevitablemente reflexión y en los espejos refracción, las pérdidas luminosas a medida que la luz se propaga en uno de estos sistemas son muy importantes. Una manera muy eficaz de evitar estas pérdidas consistiría en guiar la luz haciendo uso de la propiedad de reflexión interna total que se produce en la interfase entre dos medios de distintos índices de refracción. Tal como muestra la Figura 3.9, este mecanismo

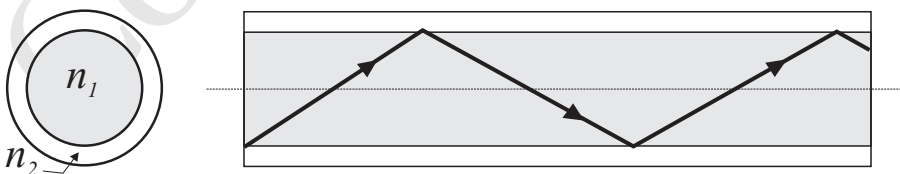
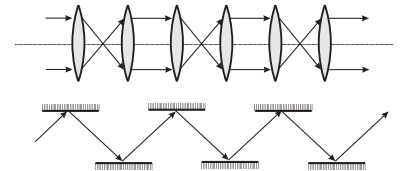


FIGURA 3.9: Mecanismo de guiado de luz en una fibra óptica ($n_1 > n_2$).

de guiado es justamente el que se usa en las fibras ópticas que tanta importancia tienen en las comunicaciones ópticas actuales.



3.5. Formalismo matricial de la Óptica paraxial

Anteriormente se han obtenido algunas ecuaciones que rigen el comportamiento de los rayos luminosos en algunos componentes simples dotados de simetría axial bajo la aproximación paraxial. Si se quiere extender el procedimiento anterior y obtener las ecuaciones correspondientes para sistemas más complicados (por ejemplo, aquéllos compuestos de varios componentes simples), encontramos que esta extensión es una tarea bastante laboriosa y complicada. Afortunadamente, podemos aplicar un formalismo matricial muy adecuado al estudio de sistemas complicados basado en la *linealidad* de las ecuaciones obtenidas para los componentes simples. La aplicación de este formalismo matricial nos permitirá analizar sistemas bastante complicados mediante una técnica formalmente muy sencilla y especialmente adaptada para el cálculo numérico por ordenador. Este método asignará una matriz característica a cada uno de los componentes simples. Es conveniente recordar que este formalismo sólo será aplicable a *sistemas ópticos centrados bajo aproximación paraxial*.

En el presente tratamiento, un rayo vendrá determinado por su posición y su ángulo (y, θ) respecto al eje óptico. Tal como muestra la Figura 3.10, el efecto del sistema óptico consistirá en modificar estas variables cuando el rayo se propaga a través del sistema. Bajo la aproximación paraxial encontraremos que la relación

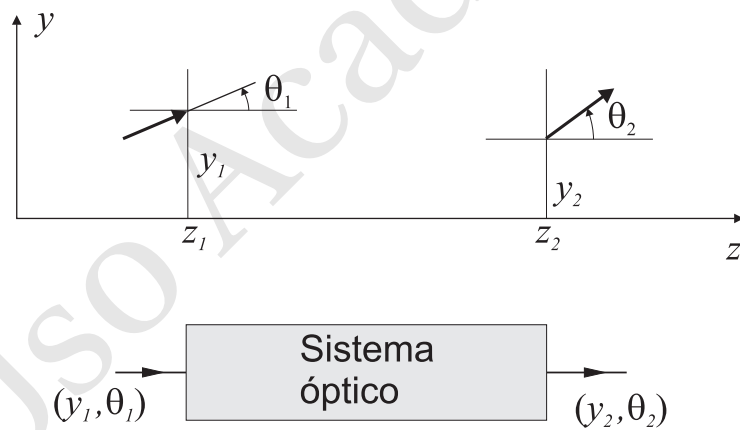


FIGURA 3.10: Efecto sobre el rayo de un sistema óptico general.

lineal entre el ángulo y la posición del rayo a la salida y entrada del sistema óptico viene dada por

$$y_2 = Ay_1 + B\theta_1 \quad (3.45a)$$

$$\theta_2 = Cy_1 + D\theta_1, \quad (3.45b)$$

que puede reescribirse en forma matricial como

$$\begin{bmatrix} y_2 \\ \theta_2 \end{bmatrix} = \underbrace{\begin{bmatrix} A & B \\ C & D \end{bmatrix}}_{[M]} \begin{bmatrix} y_1 \\ \theta_1 \end{bmatrix}, \quad (3.46)$$

siendo $[M]$ la **matriz de transferencia** del sistema.

A continuación deduciremos la forma de algunas matrices de transferencia características de ciertos componentes ópticos simples.

Propagación

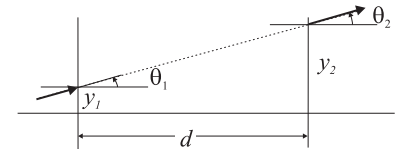
Un rayo propagándose en un medio homogéneo se propaga en línea recta, por lo que

$$y_2 = y_1 + d\theta_1 \quad (3.47a)$$

$$\theta_2 = \theta_1 \quad (3.47b)$$

y consecuentemente la *matriz de traslación* sería

$$[M] = \begin{bmatrix} 1 & d \\ 0 & 1 \end{bmatrix}. \quad (3.48)$$



Refracción en una superficie plana

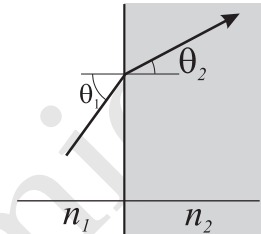
Una aplicación directa de la ley de Snell a la interfase plana dice que

$$y_2 = y_1 \quad (3.49a)$$

$$n_2\theta_2 = n_1\theta_1 \quad (3.49b)$$

por lo que la matriz $[M]$ toma la siguiente forma:

$$[M] = \begin{bmatrix} 1 & 0 \\ 0 & \frac{n_1}{n_2} \end{bmatrix}. \quad (3.50)$$



Refracción en una superficie esférica

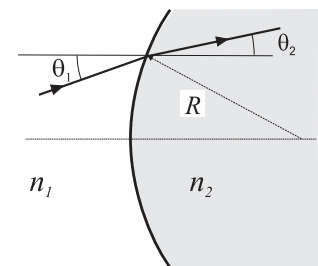
La aplicación de la ley de Snell a una superficie esférica bajo la aproximación paraxial decía, según vimos en (3.32), que

$$y_2 = y_1 \quad (3.51a)$$

$$\theta_2 = \frac{n_1}{n_2}\theta_1 - \frac{n_2 - n_1}{n_2} \frac{y_1}{R} \quad (3.51b)$$

y consecuentemente:

$$[M] = \begin{bmatrix} 1 & 0 \\ -\frac{(n_2 - n_1)}{n_2 R} & \frac{n_1}{n_2} \end{bmatrix}. \quad (3.52)$$



$R > 0$ convexa, $R < 0$ cóncava

Transmisión a través de una lente delgada

Según la expresión (3.41), en una lente delgada bajo aproximación paraxial tenemos que

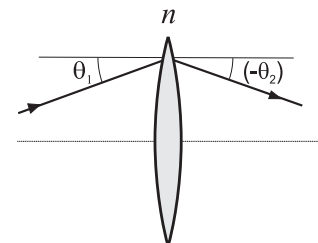
$$y_2 = y_1 \quad (3.53a)$$

$$\theta_2 = \theta_1 - \frac{y}{f}. \quad (3.53b)$$

La *matriz de transmisión* de este componente será entonces

$$[M] = \begin{bmatrix} 1 & 0 \\ -\frac{1}{f} & 1 \end{bmatrix}, \quad (3.54)$$

donde $f > 0$ si la lente es convexa y $f < 0$ si la lente es cóncava.



Reflexión en un espejo plano

Según la ley de reflexión,

$$y_2 = y_1 \quad (3.55a)$$

$$\theta_2 = \theta_1 \quad (3.55b)$$

por lo que la *matriz de reflexión* será

$$[M] = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}. \quad (3.56)$$

Reflexión en un espejo esférico

La expresión (3.21) dice que

$$y_2 = y_1 \quad (3.57a)$$

$$\theta_2 = \theta_1 + \frac{2y_1}{R}, \quad (3.57b)$$

viniendo entonces la matriz de reflexión dada por

$$[M] = \begin{bmatrix} 1 & 0 \\ \frac{2}{R} & 1 \end{bmatrix}, \quad (3.58)$$

con $R > 0$ espejo convexo y $R < 0$ espejo cóncavo.

Matriz transferencia de un sistema compuesto

Usando las matrices de transferencia de los componentes simples anteriores es posible calcular la matriz de transferencia de un sistema compuesto. Para ello, si tenemos una sucesión encadenada de N componentes simples caracterizados por $[M_1], [M_2], \dots, [M_N]$, entonces la matriz $[M]$ de transferencia del sistema compuesto vendrá dada por el siguiente producto de matrices:

$$[M] = [M_N] \cdots [M_2][M_1], \quad (3.59)$$

donde debe notarse que el orden de multiplicación empieza por el componente situado justo a la salida del sistema para acabar por el componente situado a la entrada.

EJEMPLO 3.3 Cálculo de la matriz de transferencia de una lente gruesa.

Según la figura adjunta, una lente gruesa puede considerarse como la serie de tres componentes ópticos simples:

1. Refracción en una superficie esférica convexa de radio R_1 , siendo su matriz de transferencia

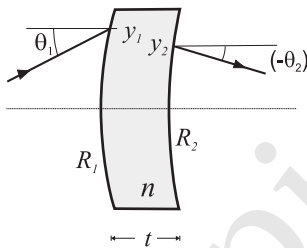
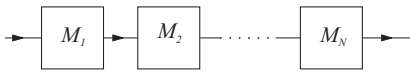
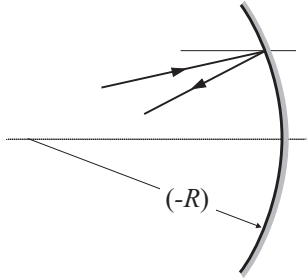
$$[M_1] = \begin{bmatrix} 1 & 0 \\ \frac{1-n}{nR_1} & \frac{1}{n} \end{bmatrix}$$

2. Traslación en un medio de espesor t e índice n , siendo

$$[M_2] = \begin{bmatrix} 1 & t \\ 0 & 1 \end{bmatrix}$$

3. Refracción en una superficie esférica convexa de radio R_2 , con

$$[M_3] = \begin{bmatrix} 1 & 0 \\ \frac{n-1}{R_2} & n \end{bmatrix}.$$



En consecuencia el efecto del sistema compuesto puede describirse mediante el siguiente producto de matrices:

$$\begin{bmatrix} y_2 \\ \theta_2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ \frac{n-1}{R_2} & n \end{bmatrix} \begin{bmatrix} 1 & t \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ \frac{1-n}{nR_1} & \frac{1}{n} \end{bmatrix} \begin{bmatrix} y_1 \\ \theta_1 \end{bmatrix}.$$

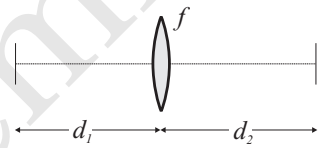
La matriz de transferencia del sistema será entonces

$$[M] = \begin{bmatrix} 1 + t \frac{1-n}{nR_1} & \frac{t}{n} \\ (n-1) \left(\frac{1}{R_2} - \frac{1}{R_1} \right) - t \frac{(n-1)^2}{nR_1 R_2} & 1 + t \frac{n-1}{nR_2} \end{bmatrix}$$

Tomando límite en la expresión anterior para $t \rightarrow 0$ recuperamos la expresión para la lente delgada dada en (3.54).

EJEMPLO 3.4 Cálculo de la matriz de transferencia para el sistema de la figura adjunta

En este caso y dado que el sistema está compuesto de una traslación de distancia d_1 , una lente delgada y una última traslación de distancia d_2 , la matriz de transferencia será el siguiente producto:



$$\begin{aligned} [M] &= \begin{bmatrix} 1 & d_2 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ -\frac{1}{f} & 1 \end{bmatrix} \begin{bmatrix} 1 & d_1 \\ 0 & 1 \end{bmatrix} \\ &= \begin{bmatrix} 1 - \frac{d_2}{f} & d_1 + d_2 - \frac{d_1 d_1}{f} \\ -\frac{1}{f} & 1 - \frac{d_1}{f} \end{bmatrix}. \end{aligned}$$

Si queremos que cualquier objeto situado en el plano de entrada de un sistema óptico tenga su correspondiente imagen en el plano de salida (es decir, si queremos sistemas astigmáticos donde los planos de entrada y salida sean conjugados), debemos exigir que todos los rayos que salgan de algún punto del plano de entrada confluyan en algún otro punto del plano de salida. Esta condición se cumple si todos los rayos que lleguen a y_2 son independientes del ángulo θ_1 con el cual partieron. Dado que y_2 no debe depender de θ_1 , esto implica que en la matriz $[M]$, $M_{12} \equiv B = 0$. En este caso encontramos que

$$A = \frac{y_2}{y_1}. \quad (3.60)$$

Para el sistema óptico analizado en el ejemplo anterior se tiene que la condición $B = 0$ implica que

$$d_1 + d_2 - \frac{d_1 d_1}{f} = 0,$$

o equivalentemente

$$\frac{1}{f} = \frac{1}{d_1} + \frac{1}{d_2}.$$

Podemos comprobar que esta condición coincide con la ecuación de la lente delgada, expresión (3.36), obtenida anteriormente.

3.6. Problemas propuestos

3.1: El índice de refracción de una determinada luz de longitud de onda en el vacío $\lambda_0 = 589 \text{ nm}$ es de 2.417 para el diamante y 1.923 para el zirconio. Calcular la relación de sus longitudes de onda en el diamante y en el zirconio.

Sol: $\lambda_z/\lambda_d = 1.257$

3.2: Un pequeño cuerpo luminoso está en el fondo de un estanque lleno de agua ($n = 4/3$) a 1 m de profundidad, y emite rayos luminosos en todas las direcciones. En la superficie del agua se forma un círculo luminoso por los rayos que se refractan al pasar al aire. Fuera de este círculo los rayos se reflejan totalmente en la superficie del agua. Hallar el radio R del círculo.

Sol: $R = 1.13 \text{ m}$.

3.3: Hallar el radio de curvatura de un espejo esférico convexo que proporciona una imagen virtual de tamaño la mitad del objeto, si éste se halla a 20 cm del espejo.

Sol: $R = 40 \text{ cm}$.

3.4: Un objeto situado a 300 cm de un espejo esférico cóncavo tiene una imagen real a 150 cm del espejo. Hallar en qué posición el objeto y la imagen de él coinciden.

Sol: 200 cm a la izquierda del espejo.

3.5: Calcular las distancias focales de las siguientes lentes delgadas, cuyo radio de curvatura es siempre de 40 mm y que están fabricadas de un vidrio de $n = 1.5$: **a)** una lente biconvexa; **b)** una lente bicóncava; **c)** una lente plano convexa; **d)** una lente plano cóncava. Calcular la situación de la imagen producida por la primera lente de un objeto real situado en el eje principal y a 20 cm delante de la lente.

Sol: a) 4 cm; b) -4 cm; c) 8 cm; d) -8 cm;

3.6: Se tiene una lente biconvexa de vidrio ($n = 1.5$) de 40 cm de distancia focal, el radio de una de las caras es 6 cm. Determinar: **a)** el radio de la otra cara; **b)** delante de ella, a 50 cm, se coloca un objeto de 3 cm de altura, determinar la posición de la imagen; **c)** calcular el tamaño de la imagen anterior y el aumento; **d)** yuxtapuesta con la anterior se coloca una lente divergente del mismo vidrio, de distancia focal 25 cm. ¿Cuál es la distancia focal del sistema?

Sol: a) 8.6 cm ; b) 2 m ; c) $m = -4$; d) $f = 67 \text{ cm}$.

3.7: Un objeto de 2 mm de altura está situado a 90 cm a la izquierda de una lente delgada divergente de 30 cm de distancia focal. A continuación de la lente divergente se dispone una lente delgada convergente de 20 cm de distancia focal. **a)** Determinar cuál debe ser la distancia entre las dos lentes para que la imagen definitiva del objeto anterior sea real y esté situada a 30 cm a la derecha de la lente convergente. **b)** Determinar la distancia focal de una lente única que produzca el mismo efecto.

Sol: a) 37.5 cm ; b) 12.63 cm.

3.8: Un sistema óptico está formado por dos lentes delgadas, la primera convergente de distancia focal $f_1 = 10 \text{ cm}$ y la segunda divergente de distancia focal $f_2 = -20 \text{ cm}$, separadas una distancia de 60 cm. Obtener la matriz de transferencia del sistema.

Sol:(distancias expresadas en metros)

$$M = \begin{bmatrix} -5 & 0.6 \\ -35 & 4 \end{bmatrix}$$

Tema 4

Óptica Ondulatoria

4.1. Introducción

La Óptica Geométrica, discutida en el tema anterior, trataba la propagación de la luz a partir de una serie de leyes puramente geométricas. Esto era posible gracias a la introducción del concepto de *rayo de luz*. Como veremos, la aproximación de rayo óptico se fundamenta básicamente en que la dimensión de los objetos es mucho mayor que la longitud de onda de la radiación luminosa, lo cual permite de algún modo ignorar la “longitud de onda” de la luz. Esta aproximación implicaba consecuentemente desprestigiar el carácter ondulatorio de la luz, dando así lugar a una teoría en la que la luz siempre se propagaba en línea recta en medios homogéneos (esto se manifestaba, por ejemplo, en la teoría de sombras geométricas). No obstante existen multitud de situaciones prácticas en las que no se puede ignorar el carácter ondulatorio de la luz (es decir, los efectos de interferencia, difracción,...) por lo que la teoría puramente geométrica del tema anterior debe complementarse adecuadamente e incorporar los aspectos ondulatorios de la radiación luminosa y sus fenómenos asociados.

Aunque en el presente tema introduciremos el carácter ondulatorio de la luz, la teoría que presentaremos no será todavía completamente rigurosa ya que consideraremos la radiación luminosa como un fenómeno puramente *escalar* en vez de vectorial (como obviamente es el carácter de la onda electromagnética). En ciertas situaciones, el carácter vectorial de la radiación luminosa no puede ser obviado (por ejemplo, cuando tratemos cuestiones energéticas, fenómenos de polarización, propagación en medios anisótropos o situaciones en las que el acoplo entre el campo eléctrico y el magnético no pueda ser ignorado), pero en multitud de otras situaciones prácticas, el considerar la radiación como una onda escalar nos permitirá disponer de una teoría más completa que la Óptica Geométrica pero sin la complicación matemática que impondría la inclusión del carácter vectorial de la luz. A esta teoría intermedia entre la Óptica Geométrica y el Electromagnetismo la denominaremos **Óptica Ondulatoria Escalar** y será el objeto del presente tema.

4.2. Relación entre rayos y frente de onda

Bajo la aproximación ondulatoria escalar, la radiación luminosa **monocromática** será considerada una *onda escalar armónica*. Usando notación compleja, esta

onda puede expresarse, en general, como

$$u(\mathbf{r}, t) = U(\mathbf{r}) \exp(j\omega t) \quad (4.1a)$$

$$= A(\mathbf{r}) \exp[j\phi(\mathbf{r})] \exp(j\omega t), \quad (4.1b)$$

donde $U(\mathbf{r})$ representa la parte espacial de la onda monocromática, denominada usualmente **amplitud compleja** de la onda y $A(\mathbf{r})$ y $\phi(\mathbf{r})$ son funciones *reales* que representan la amplitud y la fase de la onda respectivamente.

Si aplicamos la ecuación de onda (1.110) a la anterior expresión, encontramos la siguiente ecuación para $U(\mathbf{r})$:

Ecuación de Helmholtz

$$(\nabla^2 + k^2)U(\mathbf{r}) = 0, \quad (4.2)$$

que se conoce como *ecuación de Helmholtz*, siendo $k = \omega/v$ el número de ondas.

La amplitud compleja suele escribirse como

$$U(\mathbf{r}) = A(\mathbf{r}) \exp[-jk_0 \mathcal{S}(\mathbf{r})], \quad (4.3)$$

donde k_0 es el número de ondas del espacio libre ($k = nk_0$) y $\mathcal{S}(\mathbf{r})$ es una función real escalar llamada **eikonal** que define la fase de la onda¹, de modo que los frentes de onda (para un t fijo, lugar geométrico de los puntos con idéntica fase) vienen determinados por la ecuación

$$\mathcal{S}(\mathbf{r}) = \text{cte}. \quad (4.4)$$

La ecuación anterior también nos dice que la dirección normal a los frentes de onda en cada punto viene dada por el gradiente de la eikonal,

$$\nabla \mathcal{S}(\mathbf{r}) \equiv \text{normal al frente de onda en } \mathbf{r}. \quad (4.5)$$

Al sustituir la expresión (4.3) en la ecuación de Helmholtz (teniendo en cuenta que $\nabla^2 U(\mathbf{r}) = \nabla \cdot \{\nabla U(\mathbf{r})\}$), obtenemos la siguiente ecuación para las funciones reales $A(\mathbf{r})$ y $\mathcal{S}(\mathbf{r})$:

$$\{k_0^2 [n^2 - |\nabla \mathcal{S}|^2] A + \nabla^2 A\} - j \{k_0 [2\nabla \mathcal{S} \cdot \nabla A + A \nabla^2 \mathcal{S}]\} = 0. \quad (4.6)$$

La anterior ecuación debe anularse por separado tanto para la parte real como la imaginaria. Igualando a cero la parte real de (4.6) encontramos que

$$|\nabla \mathcal{S}|^2 = n^2 + \frac{\nabla^2 A}{k_0^2 A}. \quad (4.7)$$

Si la onda monocromática se propaga en un medio de índice de refracción $n(\mathbf{r})$, tal que las variaciones del índice de refracción son muy pequeñas en una escala del orden de la longitud de onda, encontraremos entonces que la amplitud de la onda variará muy lentamente con respecto a su fase, es decir tendremos que

$$\nabla^2 A \ll k_0^2 A$$

y, por tanto, este término puede ser despreciado en el límite $\lambda_0 \rightarrow 0$. En consecuencia, la ecuación (4.7) puede reescribirse como

Ecuación de la Eikonal

$$|\nabla \mathcal{S}(\mathbf{r})|^2 \approx n^2(\mathbf{r}) \quad (\lambda_0 \rightarrow 0), \quad (4.8)$$

¹ Nótese que para una onda plana uniforme en un medio homogéneo, la eikonal sería $\mathcal{S}(\mathbf{r}) = n\hat{\mathbf{k}} \cdot \mathbf{r}$

que se conoce como **ecuación de la eikonal** y que constituye la base de la Óptica Geométrica² (esto es, cuando se ignora el carácter ondulatorio de la radiación luminosa)

Un estudio más riguroso de una radiación luminosa del tipo (4.1a), que tuviese en cuenta el carácter vectorial electromagnético, nos diría que el promedio del vector de Poynting, $\langle \mathbf{S} \rangle$, de esta radiación puede expresarse como

$$\langle \mathbf{S} \rangle \propto \langle u_E \rangle \nabla \mathcal{S}(\mathbf{r}) \quad (4.9)$$

(siendo $\langle u_E \rangle$ el promedio de energía eléctrica) es decir, la dirección de propagación de la energía, que en un medio isótropo coincide con la dirección de propagación de los frentes de onda, está determinada por el gradiente de la eikonal, $\nabla \mathcal{S}(\mathbf{r})$.

Teniendo en cuenta que la dirección de propagación de la energía, $\nabla \mathcal{S}(\mathbf{r})$, define los rayos ópticos y que ésta a su vez, según (4.5), define la normal a los frentes de onda, podemos concluir que

los rayos de la Óptica Geométrica son las trayectorias ortogonales a los frentes de onda definidos por $\mathcal{S}(\mathbf{r}) = cte$

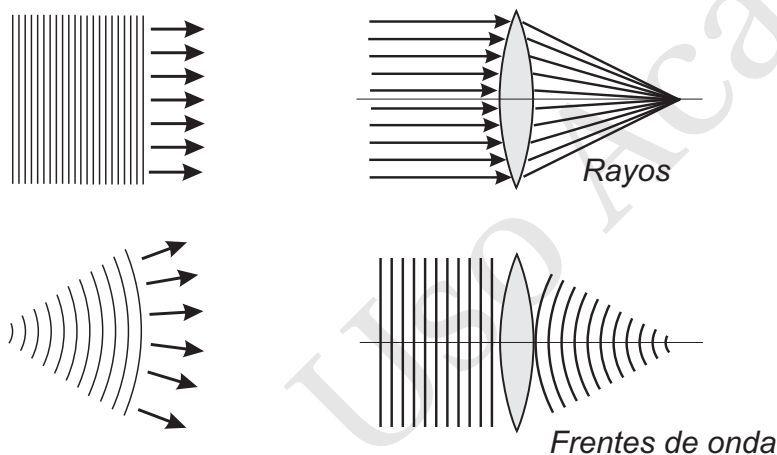


FIGURA 4.1: Relación entre los frentes de onda y sus rayos ópticos asociados

Desde este punto de vista, podemos considerar que el efecto de la lente es “distorsionar” los frentes de onda que le llegan. La Óptica Geométrica puede, por tanto, considerarse como una teoría que determina los efectos aproximados de los componentes ópticos sobre las normales al frente de onda.

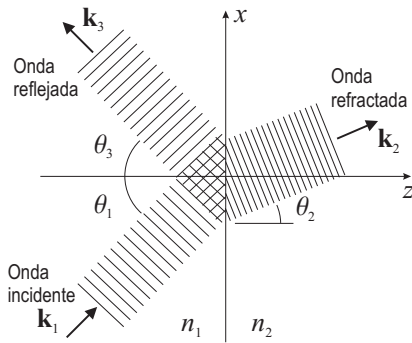
² De hecho, el principio de Fermat puede deducirse de esta ecuación, y viceversa. Si tenemos en cuenta que la ecuación (4.8) nos dice que

$$\frac{\nabla \mathcal{S}}{n} = \hat{\mathbf{s}},$$

donde $\hat{\mathbf{s}}$ es un vector normal en la dirección del gradiente de la eikonal, y que $d\mathcal{S} = \nabla \mathcal{S} \cdot d\mathbf{r}$, entonces encontramos que

$$[\mathcal{S}]_B - [\mathcal{S}]_A = \int_A^B d\mathcal{S} = \int_A^B \nabla \mathcal{S} \cdot d\mathbf{r} = \int_A^B n \hat{\mathbf{s}} \cdot d\mathbf{r} = \int_A^B n ds,$$

esto es, el *camino óptico* entre estos dos puntos es la diferencia entre el valor de las eikonales en A y B.



EJEMPLO 4.1 Reflexión y refracción de una onda plana en una interfase dieléctrica plana

Si una onda plana de vector de ondas

$$\mathbf{k}_1 = (n_1 k_0 \sin \theta_1, 0, n_1 k_0 \cos \theta_1)$$

incide sobre la superficie plana de separación de dos medios de índice de refracción n_1 y n_2 , se producirá una onda plana reflejada con vector de ondas \mathbf{k}_3 y una onda plana refractada con vector de ondas \mathbf{k}_2 . Claramente debe cumplirse que $k_3 = k_1 = n_1 k_0$ y que $k_2 = n_2 k_0$. Del dibujo de la figura observamos entonces que

$$\mathbf{k}_2 = (n_2 k_0 \sin \theta_2, 0, n_2 k_0 \cos \theta_2)$$

$$\mathbf{k}_3 = (n_1 k_0 \sin \theta_3, 0, n_1 k_0 \cos \theta_3).$$

En la superficie de separación ($z = 0$), la dependencia espacial de las fases de las tres ondas deben coincidir (puesto que esta superficie forma parte simultáneamente a las tres ondas), por lo que

$$\mathbf{k}_1 \cdot \mathbf{r} = \mathbf{k}_2 \cdot \mathbf{r} = \mathbf{k}_3 \cdot \mathbf{r} \quad \text{para todo } \mathbf{r} = (x, y, 0)$$

Teniendo en cuenta la forma de los vectores de onda y tomando $\mathbf{r} = \hat{x}$, obtenemos las siguientes identidades:

$$\theta_1 = \theta_3$$

$$n_1 \sin \theta_1 = n_2 \sin \theta_2,$$

esto es, las leyes de la reflexión y refracción que fueron obtenidas anteriormente a partir del principio de Fermat.

4.3. Ondas paraxiales

En el apartado 3.4 se introdujo la *aproximación paraxial* y se señaló que las situaciones en las cuales se iba a aplicar el formalismo de la óptica geométrica se restringiría a aquellos casos que satisficieran dicha aproximación. En el presente tema, también estaremos especialmente interesados en aquellas situaciones que verifiquen la aproximación paraxial (debido a que cubren muchas situaciones de interés práctico y a su mayor simplicidad matemática). Puesto que ahora estamos considerando el carácter ondulatorio de la radiación luminosa, y teniendo en cuenta la relación entre rayos ópticos y frentes de onda descrita anteriormente, definiremos la *onda paraxial* como aquella cuyo frente de ondas sea normal a un rayo paraxial.

Dado que la onda paraxial debe propagarse en una dirección que forme un ángulo muy pequeño con el eje óptico, una posible forma general de expresar una onda paraxial será

$$U(\mathbf{r}) = A(\mathbf{r}) \exp(-jkz), \quad (4.10)$$

debiéndose cumplir que la amplitud de esta onda, $A(\mathbf{r})$, varíe muy poco en una distancia del orden de la longitud de onda $\lambda = 2\pi/k$. Bajo esta suposición, *la onda paraxial puede considerarse localmente como una onda plana de amplitud débilmente modulada*.

Es interesante señalar que la aproximación paraxial de una onda esférica es justamente la *onda parabólica* descrita en el Apartado 1.9.4. Podemos comprobar

Expresión matemática de una onda paraxial

que bajo las condiciones de la aproximación de Fresnel descritas allí, la amplitud de (1.123),

$$A(\mathbf{r}) = \frac{\mathcal{A}}{z} \exp\left(-jk \frac{x^2 + y^2}{2z}\right),$$

varía muy poco en una distancia del orden de $2\pi/k$.

4.4. Transmisión a través de elementos ópticos

De los diversos fenómenos que podemos analizar siguiendo la óptica ondulatoria escalar nos centraremos de momento en la **transmisión** de la radiación luminosa a través de diversos componentes ópticos. Los fenómenos de reflexión asociados serán ignorados en el presente tratamiento (de hecho, la óptica ondulatoria escalar no puede dar cuenta apropiadamente de estos fenómenos dado que obvia las *condiciones de salto* que debe verificar el campo electromagnético en las fronteras de separación).

4.4.1. Transmisión a través de una lámina transparente

Consideremos una onda plana cuya parte espacial viene dada por $U(x, y, z)$ y que se propaga en el espacio libre. Esta onda incide en una lámina *transparente* de espesor d e índice de refracción n , situada entre los planos $z = 0$ y $z = d$. Dado que ignoramos los efectos de reflexión, la onda debe ser continua en las superficies planas exteriores de la lámina. La relación entre el valor de la onda luminosa en ambas superficies se denomina **transmitancia en amplitud** de la lámina:

$$t(x, y) = \frac{U(x, y, d)}{U(x, y, 0)}. \quad (4.11)$$

■ Incidencia normal.

Si la onda incide normal a la superficie de la lámina, entonces dentro de ésta, la onda se propagará igualmente en la dirección normal como una onda plana cuyo número de ondas será nk_0 . En cualquier punto z del interior, la onda vendrá dada por

$$U(x, y, z) = A \exp(-jnk_0 z), \quad (4.12)$$

luego en la superficie exterior ($z = d$), tendremos que

$$U(x, y, d) = A \exp(-jnk_0 d) = U(x, y, 0) \exp(-jnk_0 d), \quad (4.13)$$

por lo que

$$t(x, y) = \frac{U(x, y, d)}{U(x, y, 0)} = \exp(-jnk_0 d), \quad (4.14)$$

es decir, la lámina simplemente introduce un desplazamiento de fase $\varphi = nk_0 d = 2\pi nd / \lambda_0$.

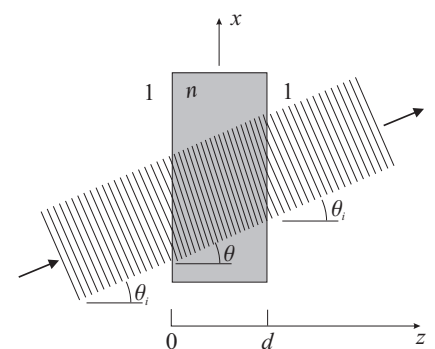
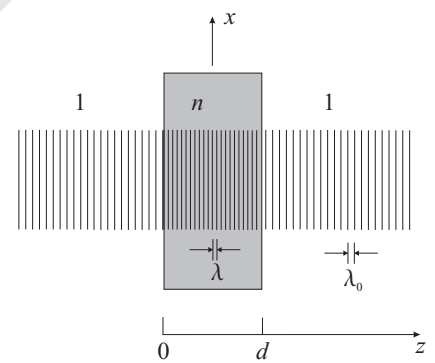
■ Incidencia oblicua.

Si la onda viaja en el interior de la lámina formando un ángulo θ con el eje z , entonces la onda dentro de la lámina podrá escribirse como

$$U(x, y, z) = A \exp(-j\mathbf{k} \cdot \mathbf{r}), \quad (4.15)$$

siendo

$$\mathbf{k} = nk_0 (\cos\theta \hat{\mathbf{z}} + \sin\theta \hat{\mathbf{x}}),$$



por lo que $U(x, y, d)$ puede expresarse como

$$U(x, y, d) = U(x, y, 0) \exp[-jnk_0(d \cos \theta + x \operatorname{sen} \theta)] . \quad (4.16)$$

En consecuencia, la transmitancia en amplitud vendrá dada por

$$t(x, y) = \frac{U(x, y, d)}{U(x, y, 0)} = \exp[-jnk_0(d \cos \theta + x \operatorname{sen} \theta)] . \quad (4.17)$$

Bajo aproximación paraxial, el ángulo de incidencia, θ_i , es muy pequeño, por lo que $\theta \approx \theta_i/n$ lo será igualmente. En este caso, y supuesto un espesor de la lámina pequeño, tendremos que

$$k_0 d \cos \theta \approx k_0 d \quad k_0 x \operatorname{sen} \theta \approx 0 ,$$

por lo que la transmitancia de la lámina en incidencia oblicua es aproximadamente la misma que la del caso de incidencia normal.

4.4.2. Lámina transparente de espesor variable

En el caso de que la lámina transparente sea de espesor variable hay que considerar que el desfase que introduce la lámina es ahora dependiente de la posición. En concreto, para incidencia normal, entre $z = 0$ y $z = d_0$, la fase de la onda se debe a los siguientes tres desfases parciales:

1. Región de aire entre $z = 0$ y $z = z_1$.
El desfase en esta región será $\varphi_1 = k_0 z_1$,
2. Región de material entre $z = z_1$ y $z = z_2$.
En el interior de la lámina, el desfase viene dado por $\varphi_2 = nk_0 d$ pues $d = z_2 - z_1$,
3. Región de aire entre $z = z_2$ y $z = d_0$. En esta última región de aire, el desfase es $\varphi_3 = k_0(d_0 - z_2)$,

por lo que el desfase total será

$$\varphi = k_0(d_0 - d) + nk_0 d$$

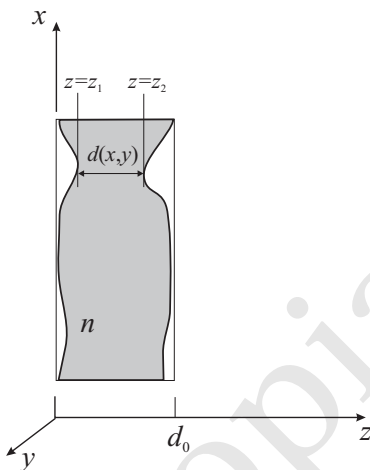
y la transmitancia de la lámina será, por tanto,

$$t(x, y) = \exp[-jk_0(d_0 - d)] \exp(-jnk_0 d) \quad (4.18)$$

o equivalentemente,

$$t(x, y) = h_0 \exp[-j(n-1)k_0 d(x, y)] , \quad (4.19)$$

siendo h_0 un factor de fase constante de valor $h_0 = \exp(-jk_0 d_0)$.



Transmitancia de una lámina de espesor variable

EJEMPLO 4.2 Obtener el efecto de un prisma delgado de ángulo $\alpha \ll$ sobre una onda plana que incide normalmente sobre éste.

En la figura adjunta tenemos que

$$d(x, y) = x \tan \alpha \approx \alpha x,$$

por lo que transmitancia del prisma, según (4.19), viene dada por

$$t(x, y) \approx h_0 \exp[-j(n-1)k_0 \alpha x].$$

El efecto de este prisma sobre una onda incidente del tipo

$$U_i = A \exp(-jkz) \Big|_{z=0}$$

provoca que

$$\begin{aligned} U_t(x, y, z) &= U_t(x, y, d) \exp(-jkz) = t(x, y) U_i(x, y, 0) \exp(-jkz) \\ &= h_0 A \exp\{-jk_0[(n-1)\alpha x + z]\}, \end{aligned}$$

que puede reescribirse como

$$U_t = h_0 A \exp(-j\mathbf{k} \cdot \mathbf{r}),$$

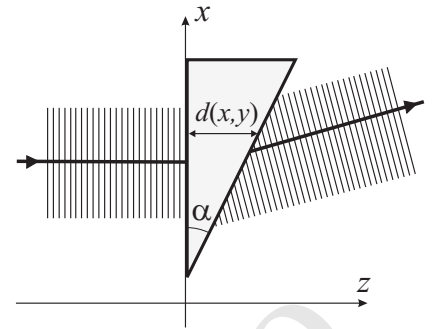
siendo \mathbf{k} el vector de ondas de esta onda plana oblicua,

$$\mathbf{k} = k_x \hat{\mathbf{x}} + k_z \hat{\mathbf{z}} = k_0(n-1)\alpha \hat{\mathbf{x}} + k_0 \hat{\mathbf{z}}.$$

El efecto del prisma es, por tanto, deflectar la onda plana respecto al eje z un ángulo θ dado aproximadamente por

$$\theta \approx (n-1)\alpha.$$

(Verificar este mismo resultado usando óptica geométrica).



Transmitancia en amplitud de una lente delgada.

En primer lugar trataremos la **incidencia normal** en una lente plano convexa como la mostrada en la figura adjunta. Lo único que debemos hacer es aplicar (4.19) y, por tanto, únicamente debemos obtener la expresión para $d(x, y)$. Para hacer esto consideremos, según el dibujo, que

$$\begin{aligned} d &= d_0 - \overline{PQ} = d_0 - \{R - \overline{QC}\} \\ &= d_0 - \{R - [R^2 - (x^2 + y^2)]^{1/2}\}. \end{aligned}$$

Ahora bien, para puntos (x, y) que verifiquen $x^2 + y^2 \ll R^2$, tenemos que

$$\sqrt{R^2 - (x^2 + y^2)} = R \left(1 - \frac{x^2 + y^2}{R^2}\right)^{1/2} \approx R \left(1 - \frac{x^2 + y^2}{2R^2}\right),$$

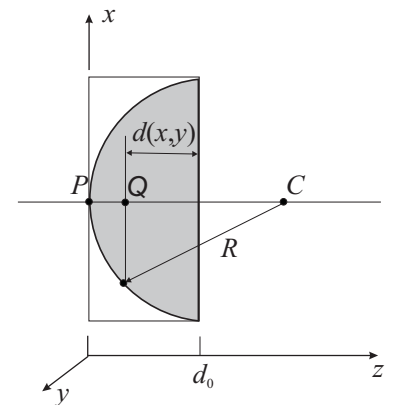
luego

$$d(x, y) \approx d_0 - \frac{x^2 + y^2}{2R}. \tag{4.20}$$

Al sustituir en (4.19) obtenemos finalmente que

$$t(x, y) \approx h_0 \exp\left(jk_0 \frac{x^2 + y^2}{2f}\right), \tag{4.21}$$

siendo $f = R/(n-1)$ la distancia focal de la lente y h_0 un factor de fase constante dado por $h_0 = \exp(-jnk_0 d_0)$. Aunque la expresión (4.21) ha sido obtenida



Transmitancia de una lente delgada

para una lente plano convexa, ésta es válida para cualquier tipo de lente delgada simplemente sustituyendo el valor de la distancia focal, f , por el de la lente correspondiente.

Podemos comprobar que el efecto de la lente delgada sobre una onda plana incidente,

$$U_i(x, y, z) = A \exp(-jk_0 z), \quad (4.22)$$

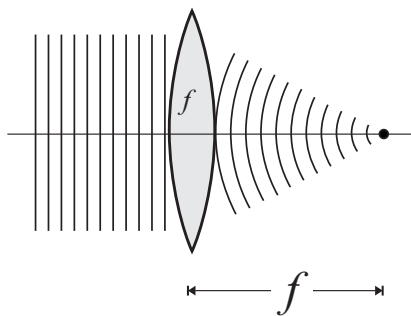
sería “transformarla” en la siguiente onda:

$$U_t(x, y, z) = t(x, y)U_i(x, y, 0) \exp(-jk_0 z) \quad (4.23)$$

$$= Ah_0 \exp\left(jk_0 \frac{x^2 + y^2}{2f}\right) \exp(-jk_0 z), \quad (4.24)$$

que, de acuerdo con (4.10), es una onda parabólica cuyo centro está en f . En consecuencia, el efecto de la lente de distancia focal f será

introducir un factor de fase que modifica los frentes de onda de modo que transforma la onda plana en una onda parabólica centrada a una distancia f de la lente.



Si una onda plana paraxial incide **oblicuamente** sobre la lente delgada con ángulos θ_x y θ_y , entonces, bajo la aproximación de *lente delgada*, el efecto de la lente es simplemente focalizar la onda plana en un punto del plano focal de la lente dado por

$$x = f \tan \theta_x \approx f \theta_x \quad (4.25a)$$

$$y = f \tan \theta_y \approx f \theta_y, \quad (4.25b)$$

tal como muestra la Figura 4.2.

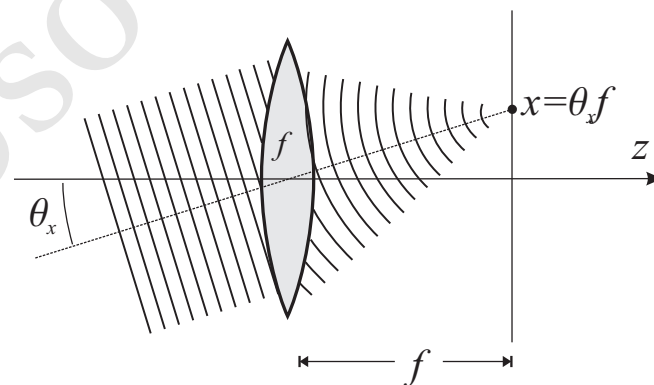


FIGURA 4.2: Efecto de una lente delgada sobre una onda plana que incide oblicuamente a la lente.

Este efecto proviene del hecho de que la incidencia oblicua paraxial se manifiesta en un cambio en la expresión de la onda incidente (4.22), que ahora será

$$U_i(x, y, 0) = A \exp[-jk_0(x\theta_x + y\theta_y)]. \quad (4.26)$$

Al introducir esta nueva expresión en (4.23) obtendremos que

$$U_t(x, y, z) \approx Ah_0 \exp\left(jk_0 \frac{x^2 + y^2}{2f}\right) \exp[-jk_0(x\theta_x + y\theta_y + z)], \quad (4.27)$$

que puede reescribirse como

$$U_t(x, y, z) \approx Ah'_0 \exp \left[jk_0 \frac{(x - f\theta_x)^2 + (y - f\theta_y)^2}{2f} \right] \exp(-jk_0z), \quad (4.28)$$

donde h'_0 es un factor de fase constante de valor

$$h'_0 = h_0 \exp \left[-jk_0(f/2)(\theta_x^2 + \theta_y^2) \right].$$

Al comparar (4.28) con (4.24) observamos que el efecto de la incidencia oblicua se manifiesta básicamente en la modificación del punto donde focaliza la onda parabólica.

4.4.3. Red de difracción

Se conoce como *red de difracción* al componente óptico que modula de una forma periódica la amplitud y/o fase de la onda incidente. El análisis de este elemento mediante óptica ondulatoria nos dirá que el efecto de la red de difracción sobre el frente de ondas será “dividir” la onda plana incidente en múltiples ondas planas que viajan en distintas direcciones (fenómenos que ciertamente no puede explicarse de ninguna manera en el marco de la óptica geométrica).

Consideremos la red de difracción de fase mostrada en la Fig. 4.3 fabricada

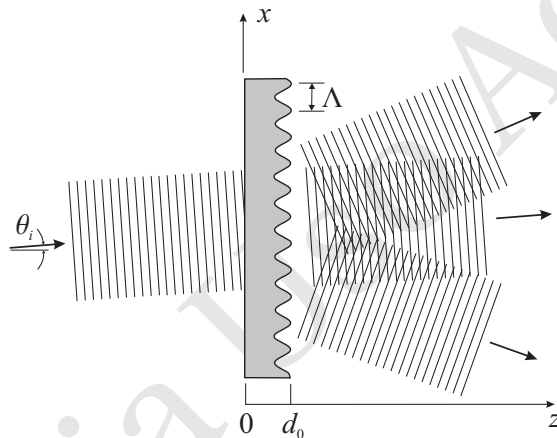


FIGURA 4.3: Efecto de una red de difracción de fase sobre una onda plana que incide oblicuamente. Puede observarse que la red de difracción divide la onda incidente en múltiples ondas planas viajando en distintas direcciones.

con una lámina transparente situada en el plano $z = 0$ y cuyo espesor varía periódicamente en la dirección x con periodo Λ . En este caso,

$$d(x, y) = \frac{1}{2}d_0 \left[1 + \cos\left(\frac{2\pi x}{\Lambda}\right) \right], \quad (4.29)$$

y, en consecuencia, la transmitancia de la red de difracción puede escribirse como

$$t(x, y) = h_0 \exp \left[-j\frac{1}{2}(n-1)k_0d_0 \cos\left(\frac{2\pi x}{\Lambda}\right) \right], \quad (4.30)$$

siendo $h_0 = \exp[-j\frac{1}{2}(n+1)k_0d_0]$.

Una onda paraxial que incide sobre la red de difracción con un ángulo θ_i ($\sin\theta_i \approx \theta_i$) y de longitud de onda $\lambda_0 \ll \Lambda$ se transforma por efecto de ésta en

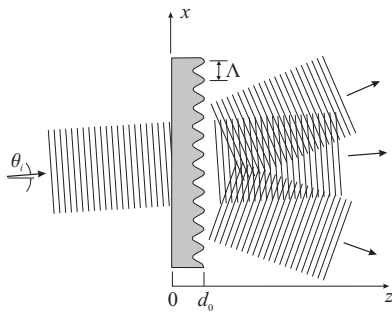
$$U_t(x, z) = t(x)U_i(x, z). \quad (4.31)$$

Antes de realizar la anterior multiplicación, debemos tener en cuenta que la transmitancia $t(x, y)$ es una función *periódica* en x de periodo Λ y, por tanto, puede expresarse como la siguiente serie de Fourier (ver Apéndice D.1):

$$t(x) = \sum_{q=-\infty}^{\infty} B_q \exp\left(-jq \frac{2\pi x}{\Lambda}\right). \quad (4.32)$$

Sustituyendo la expresión anterior en (4.31), obtenemos que la onda a la salida de la red de difracción puede expresarse como

$$\begin{aligned} U_t(x, z) &\propto \left\{ \sum_{q=-\infty}^{\infty} B_q \exp\left(-jq \frac{2\pi x}{\Lambda}\right) \right\} \exp(-jk_0 x \sin\theta_i) \exp(-jk_0 z \cos\theta_i) \\ &\propto \left\{ \sum_{q=-\infty}^{\infty} B_q \exp\left[-jk_0 x \left(\sin\theta_i + q \frac{\lambda_0}{\Lambda}\right)\right] \right\} \exp(-jk_0 z \cos\theta_i) \\ &\propto \sum_{q=-\infty}^{\infty} B_q \exp(-j\mathbf{k}_q \cdot \mathbf{r}), \end{aligned} \quad (4.33)$$



esto es, una superposición de ondas planas con vector de ondas \mathbf{k}_q dado por

$$\mathbf{k}_q = k_0 \left[\left(\sin\theta_i + q \frac{\lambda_0}{\Lambda} \right) \hat{\mathbf{x}} + \cos\theta_i \hat{\mathbf{z}} \right]. \quad (4.34)$$

Bajo la aproximación paraxial, cada una de estas ondas planas sale de la red de difracción formando un ángulo θ_q dado por

$$\theta_q = \theta_i + q \frac{\lambda_0}{\Lambda}, \quad (4.35)$$

donde $q = 0, \pm 1, \pm 2, \dots$, se denomina *orden de difracción*.

EJEMPLO 4.3 Una red de difracción por transmisión tiene 100 líneas/mm, 5 cm de tamaño y se ilumina normalmente con un haz de luz monocromática de 450 nm. Calcular: **a)** el periodo de la red, **b)** el máximo orden de difracción posible, q , así como su dispersión angular ($\mathcal{D} = \theta_q/\lambda$) y el poder de resolución ($R = qN$, N número total de líneas) para ese orden, y **c)** el ángulo de difracción del máximo de cuarto orden.

- a)** Dada la frecuencia de la red del enunciado, podemos obtener el periodo a partir de la siguiente expresión:

$$\Lambda = \frac{1}{\nu} = 1 \cdot 10^{-5} \text{ m/línea}$$

- b)** Aplicando la aproximación paraxial, el ángulo de desviación de las ondas de salida de orden q viene dado por

$$\theta_q = q \frac{\lambda}{\Lambda} = 0.045q.$$

El máximo ángulo que tiene sentido es $\pm\pi/2$, es decir, q debe estar comprendido entre -34 y 34. Tomando $q = 34$, la dispersión angular de la red es

$$\mathcal{D} = \frac{\theta_q}{\lambda} = \frac{\theta_q d}{\lambda d} = \frac{q}{\Lambda \cos\theta_q} = \frac{34}{4.08 \times 10^{-7}} = 83.36 \times 10^6 \text{ líneas/m.}$$

El poder de resolución de la red de difracción para el orden 34 es $R = 34N$, siendo $N = 1000 \times 5$, por lo que $R = 170000$.

- c) Para el cuarto orden, teniendo en cuenta que la incidencia es normal, el ángulo de difracción es

$$\theta_q = \arcsen\left(q \frac{\lambda}{\Lambda}\right) = 0.181 \text{ rad}$$

resultado similar al que se obtiene aplicando bajo la aproximación paraxial:

$$\theta_q = q \frac{\lambda}{\Lambda} = 0.18 \text{ rad}.$$

4.5. Problemas propuestos

4.1: Sabiendo que la transmitancia en amplitud de una lente planoconvexa de distancia focal f viene dada por $t(x, y) \approx h_0 \exp[jk_0(x^2 + y^2)/2f]$, demostrar que la transmitancia en amplitud de una lente biconvexa viene dada por la misma expresión si

$$\frac{1}{f} = (n-1) \left(\frac{1}{R_1} - \frac{1}{R_2} \right).$$

4.2: Demostrar que una lente delgada de distancia focal f convierte una onda parabólica centrada en el punto P_1 , con $z = z_1$, en una onda parabólica centrada en P_2 , con $z = z_2$, verificándose que

$$\frac{1}{z_1} + \frac{1}{z_2} = \frac{1}{f}.$$

4.3: Luz de color verde de longitud de onda 514 nm incide normalmente sobre una red de difracción por transmisión de 400 líneas/mm y 2 cm de tamaño. Calcular: **a)** el periodo de la red, y **b)** la desviación angular del máximo de tercer orden, así como la dispersión D y el poder de resolución R para ese orden.

Sol: **a)** $\Lambda = 2.5 \times 10^{-6}$ m/linea;

b) $\theta_q = 0.665$ rad, poder de resolución es 24000.

4.4: Una red de difracción tiene 315 líneas/mm, y luz blanca incide normalmente sobre la red. ¿En qué longitudes de onda en el espectro visible se puede observar una difracción de quinto orden?

Sol: Bajo aproximación paraxial todas las longitudes de onda del espectro visible dan lugar a armónicos difractados de quinto orden. Sin aproximación paraxial sólo se observará difracción de quinto orden para longitudes de onda inferiores a 634 nm, esto es, no habrá difracción de quinto orden para el color rojo.

4.5: Las longitudes de onda del espectro visible están comprendidas aproximadamente entre 400 y 700 nm. Hállese la amplitud angular del espectro visible de primer orden producido por una red de difracción plana que tiene 6000 líneas/cm, cuando la incidencia de la luz sobre la red es normal.

Sol: el espectro visible de primer orden es dispersado un ángulo de 0.180 rad

4.6: Demostrar que, para cualquier red de difracción, el violeta del espectro de tercer orden se solapa con el rojo del espectro de segundo orden.

4.7: Las amplitudes complejas de una onda monocromática de longitud de onda λ en los planos $z = 0$ y $z = d$ son $f(x, y)$ y $g(x, y)$, respectivamente. Si $f(x, y)$ es una función con simetría circular y frecuencia espacial máxima de 200 líneas/mm, determinar el ángulo del cono dentro del cual están confinadas las direcciones de onda. Suponer $\lambda = 633$ nm.

Sol: $\theta_{\max} = 0.13$ rad.

Copia Uso Académico

Tema 5

Óptica de Fourier

5.1. Introducción

La Óptica de Fourier constituye aquella parte de la óptica que trata la luz a partir de un análisis de Fourier bidimensional y la teoría de sistemas lineales. El análisis de Fourier (ver Apéndice D) es una herramienta matemática ampliamente usada en Física que consiste básicamente en expresar las funciones que representan a las magnitudes físicas como una combinación lineal (serie o integral) de funciones elementales de tipo armónico. La ventaja de aplicar este análisis a los sistemas ópticos es que, supuesto que el sistema óptico sea un Sistema Lineal Invariante (SLI, ver Apéndice F), conociendo la respuesta del sistema a una función armónica podemos fácilmente determinar la respuesta del sistema a una entrada arbitraria mediante el uso del análisis de Fourier a la entrada y aplicando superposición a la salida.

Una función de dos variables, $f(x, y)$, que represente la distribución luminosa en el plano (x, y) de una onda luminosa puede expresarse, mediante el análisis de Fourier, como la siguiente superposición lineal continua (una integral) de funciones armónicas elementales:

$$f(x, y) = \iint_{-\infty}^{\infty} F(v_x, v_y) \exp[-j2\pi(v_x x + v_y y)] dv_x dv_y \quad (5.1)$$

siendo

$$F(v_x, v_y) = \iint_{-\infty}^{\infty} f(x, y) \exp[j2\pi(v_x x + v_y y)] dx dy, \quad (5.2)$$

donde v_x y v_y son las *frecuencias espaciales* en las direcciones x e y respectivamente. En este sentido, la función

$$F(v_x, v_y) \exp[-j2\pi(v_x x + v_y y)] \quad (5.3)$$

es la “pieza base” sobre la que se construye la teoría de la óptica de Fourier. Como muestra la Figura 5.1, superponiendo funciones armónicas del tipo (5.3) podemos construir cualquier imagen.

Desde un punto de vista físico, la transformada de Fourier bidimensional (5.1) permite interpretar el efecto de un sistema óptico sobre una onda luminosa arbitraria en términos de los efectos particulares a ondas planas del tipo

$$\exp[-j(k_x x + k_y y + k_z z)].$$

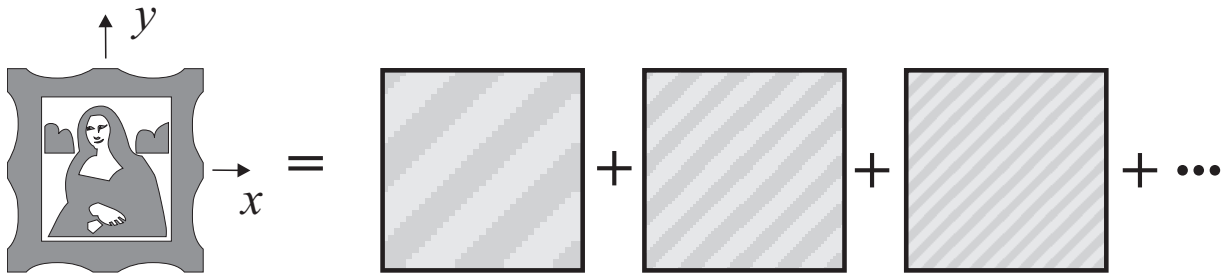


FIGURA 5.1: Una imagen arbitraria puede construirse superponiendo armónicos bidimensionales

Nótese que la anterior onda plana en $z = 0$ es justamente una función armónica bidimensional del tipo (5.3) sin más que identificar $v_x = k_x/2\pi$ y $v_y = k_y/2\pi$.

Dado que la superposición de funciones armónicas produce cualquier función arbitraria, esto implica que una onda arbitraria puede igualmente construirse superponiendo ondas planas. En consecuencia, estudiando el efecto de un sistema óptico lineal sobre una onda plana y luego superponiendo estos efectos tendremos una forma general de describir el sistema.

La interpretación anterior (o equivalentemente, el análisis de Fourier) sugiere tratar el efecto de un sistema óptico mediante la teoría de sistemas lineales. Tal como muestra la Figura 5.2, las amplitudes complejas en dos planos normales al

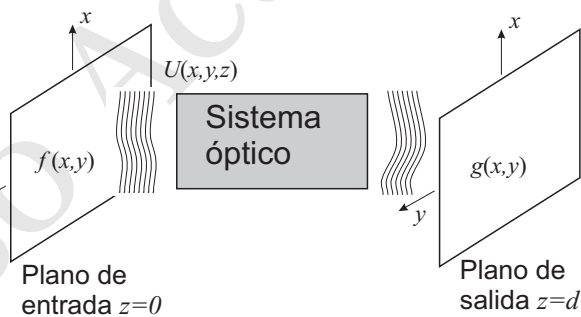


FIGURA 5.2: Efecto de un sistema óptico sobre una onda luminosa.

eje óptico (z) se considerarán como la *entrada* y la *salida* de un sistema lineal,

$$\begin{aligned} f(x, y) &= U(x, y, 0) \\ g(x, y) &= U(x, y, d) \end{aligned} \quad (5.4)$$

Dicho sistema lineal (esto es, el sistema óptico) quedará entonces caracterizado mediante su respuesta impulsiva o su **función de transferencia** (ver Apéndice F).

5.2. Aspectos preliminares

A continuación veremos algunos aspectos importantes previos al estudio específico de los sistemas ópticos mediante la transformada de Fourier y la teoría de sistemas lineales.

5.2.1. Correspondencia entre la función armónica espacial y la onda plana

Consideremos la siguiente onda plana de longitud de onda λ propagándose en el espacio libre:

$$U(x, y, z) = A \exp[-j(k_x x + k_y y + k_z z)], \quad (5.5)$$

cuyo vector de ondas es

$$\mathbf{k} = k_x \hat{\mathbf{x}} + k_y \hat{\mathbf{y}} + k_z \hat{\mathbf{z}} \quad (5.6)$$

y su número de ondas, $k = |\mathbf{k}|$, viene dado por

$$k = \sqrt{k_x^2 + k_y^2 + k_z^2} = \frac{2\pi}{\lambda}. \quad (5.7)$$

Para esta onda plana, si θ_x , θ_y y θ_z son los ángulos que forma el vector \mathbf{k} con los planos yz , xz y xy respectivamente, tenemos que

$$\text{sen } \theta_x = \frac{k_x}{k} \quad (5.8a)$$

$$\text{sen } \theta_y = \frac{k_y}{k} \quad (5.8b)$$

$$\text{sen } \theta_z = \sqrt{1 - \text{sen}^2 \theta_x - \text{sen}^2 \theta_y}. \quad (5.8c)$$

Observemos también que la amplitud compleja de la onda (5.5) en el plano $z = 0$ viene dada por la siguiente función armónica bidimensional:

$$f(x, y) \equiv U(x, y, 0) = A \exp[-j2\pi(v_x x + v_y y)], \quad (5.9)$$

con frecuencias espaciales

$$v_x = \frac{k_x}{2\pi} \quad (5.10a)$$

$$v_y = \frac{k_y}{2\pi}. \quad (5.10b)$$

Teniendo en cuenta las expresiones anteriores, (5.8a)–(5.8b) pueden reescribirse como

$$\text{sen } \theta_x = \lambda v_x \quad (5.11a)$$

$$\text{sen } \theta_y = \lambda v_y, \quad (5.11b)$$

o, equivalentemente, llamando $\Lambda_x = 1/v_x$ y $\Lambda_y = 1/v_y$ a los periodos espaciales en las direcciones x y y respectivamente, como

$$\text{sen } \theta_x = \frac{\lambda}{\Lambda_x} \quad (5.12a)$$

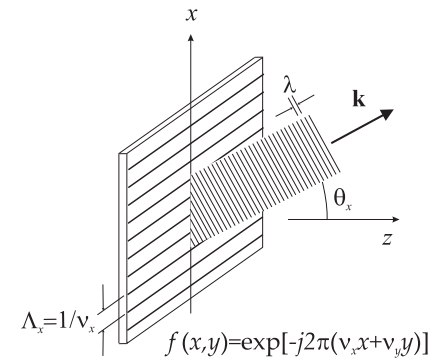
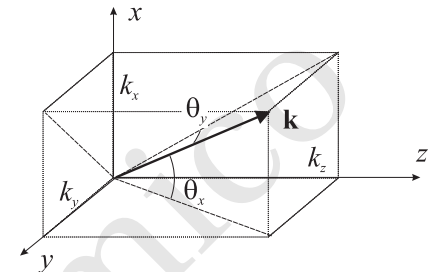
$$\text{sen } \theta_y = \frac{\lambda}{\Lambda_y}. \quad (5.12b)$$

Obsérvese que si la onda plana es paraxial (esto es, si $k_x \ll k$ y $k_y \ll k$), dado que los ángulos θ_x y θ_y son muy pequeños, podremos escribir que

$$\theta_x = \lambda v_x \quad (5.13a)$$

$$\theta_y = \lambda v_y, \quad (5.13b)$$

lo que implica que



bajo la aproximación paraxial, los ángulos de inclinación θ_x y θ_y del vector de onda son directamente proporcionales a las frecuencias espaciales de su correspondiente función armónica.

Hemos visto que para una onda plana $U(x, y, z)$ de longitud de onda λ podemos formar una función armónica asociada $f(x, y)$. Similarmente, dada una función armónica $f(x, y)$ podemos construir su onda plana asociada de longitud de onda λ , $U(x, y, z)$, de la siguiente manera:

$$U(x, y, z) = f(x, y) \exp(-jk_z z), \quad (5.14)$$

donde k_z debe satisfacer de acuerdo con (5.7) que

$$k_z = \sqrt{k^2 - k_x^2 - k_y^2}, \quad k = \frac{2\pi}{\lambda}. \quad (5.15)$$

Observemos que para que la expresión (5.14) represente realmente una onda plana *propagativa* debe cumplirse que

$$k_x^2 + k_y^2 < k^2, \quad (5.16)$$

puesto que esta condición asegura que k_z sea un número real, dando así lugar a una onda que se *propaga* en la dirección z (si k_z fuese un número imaginario la onda se *atenuaría* en la dirección z). Nótese además que de (5.11a)–(5.11b) se deduce que los ángulos θ_x y θ_y estarán definidos únicamente si se verifica que

$$\lambda v_x < 1 \quad (\Rightarrow \lambda < \Lambda_x) \quad (5.17a)$$

$$\lambda v_y < 1 \quad (\Rightarrow \lambda < \Lambda_y). \quad (5.17b)$$

5.2.2. Análisis espectral de frecuencias espaciales

Si una onda plana, $U(x, y, z)$, de longitud de onda λ ($k = 2\pi/\lambda$) y amplitud unidad viaja a lo largo de la dirección z ,

$$U(x, y, z) = \exp(-jkz), \quad (5.18)$$

y atraviesa un elemento óptico de espesor nulo situado en $z = 0$ (ver Fig. 5.3), cuya transmitancia compleja viene dada por la siguiente función armónica:

$$f_a(x, y) = \exp[-j2\pi(v_x x + v_y y)], \quad (5.19)$$

entonces la amplitud de la onda plana incidente es modulada por esta función armónica, de modo que en $z = 0^+$ (de acuerdo con (4.11)) se cumple que

$$U(x, y, 0^+) = f_a(x, y)U(x, y, 0^-) \equiv f_a(x, y). \quad (5.20)$$

La onda plana incidente normal se ha convertido en una onda plana cuyo vector de onda presenta unos ángulos de inclinación θ_x y θ_y dados por

$$\theta_x = \arcsen(\lambda v_x) \quad (5.21a)$$

$$\theta_y = \arcsen(\lambda v_y). \quad (5.21b)$$

Nótese que el efecto de elemento óptico de transmitancia (5.19) es similar al efecto de un prisma (ver Ejemplo 4.2) cuando “desviaba” la onda plana incidente, aunque el ángulo de desviación del prisma no dependía directamente de la longitud de onda incidente (sólo lo haría si $n = n(\lambda)$).

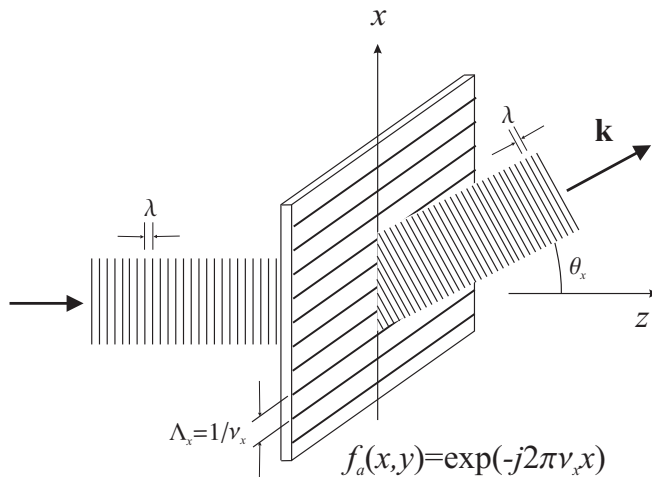


FIGURA 5.3: Efecto de un elemento óptico de transmitancia $f_a(x, y) = \exp(-j2\pi\nu_x x)$ sobre una onda plana que incide normalmente. La onda plana resultante es deflexada un ángulo $\theta_x = \arcsen(\lambda\nu_x)$.

Si la transmitancia del elemento óptico fuese ahora una función arbitraria, $f(x, y)$, que es la suma de un conjunto de funciones armónicas de diferentes frecuencias espaciales, entonces, la onda transmitida podría igualmente considerarse como la suma de un conjunto de ondas planas desviadas en diferentes direcciones, siendo la amplitud de cada una de estas ondas planas proporcional a la amplitud de cada armónico de $f(x, y)$. En general como $f(x, y)$ puede representarse como la siguiente superposición integral de funciones armónicas:

$$f(x, y) = \iint_{-\infty}^{\infty} F(\nu_x, \nu_y) \exp[-j2\pi(\nu_x x + \nu_y y)] d\nu_x d\nu_y, \quad (5.22)$$

la onda transmitida, $U(x, y, z)$, puede expresarse como la siguiente superposición de ondas planas:

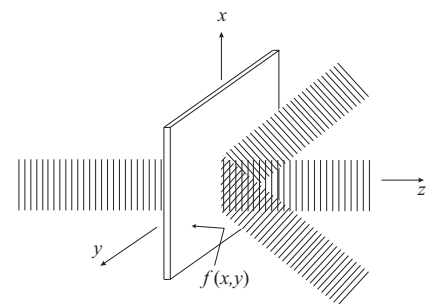
$$U(x, y, z) = \iint_{-\infty}^{\infty} F(\nu_x, \nu_y) \exp[-j2\pi(\nu_x x + \nu_y y)] \exp(-jk_z z) d\nu_x d\nu_y, \quad (5.23)$$

donde

$$k_z = \sqrt{k^2 - k_x^2 - k_y^2} = 2\pi\sqrt{\frac{1}{\lambda^2} - \nu_x^2 - \nu_y^2}. \quad (5.24)$$

Dado que, según el análisis de Fourier, cualquier función arbitraria puede ser expresada como una superposición integral del tipo (5.22), la luz transmitida a través de un elemento óptico delgado de transmitancia arbitraria puede igualmente escribirse como una superposición de ondas planas; aunque, en rigor, deberíamos decir como una superposición de ondas planas *propagativas* (k_z real), es decir, aquellas que verifiquen

$$\nu_x^2 + \nu_y^2 < \frac{1}{\lambda^2}. \quad (5.25)$$



5.2.3. Modulación en amplitud

Consideremos una transparencia cuya transmitancia compleja, $f_0(x, y)$, tiene una transformada de Fourier, $F_0(\nu_x, \nu_y)$, que está extendida unos

intervalos Δv_x y Δv_y con respecto a $(v_x, v_y) = (0, 0)$. Si esta transparencia es iluminada por una onda plana, de acuerdo a las ecuaciones (5.21a)–(5.21b), esta onda plana será deflectada en unos rangos angulares, $\Delta\theta_x$ y $\Delta\theta_y$, con respecto al eje z dados por

$$\Delta\theta_x = \arcsen(\lambda\Delta v_x) \quad (5.26a)$$

$$\Delta\theta_y = \arcsen(\lambda\Delta v_y). \quad (5.26b)$$

Consideremos ahora una segunda transparencia de transmitancia compleja

$$f(x, y) = f_0(x, y) \exp[-j2\pi(v_{x0}x + v_{y0}y)], \quad (5.27)$$

tal que la función $f_0(x, y)$ varíe muy poco con respecto a la función armónica $\exp[-j2\pi(v_{x0}x + v_{y0}y)]$, es decir,

$$\Delta v_x \ll v_{x0} \quad (5.28a)$$

$$\Delta v_y \ll v_{y0}. \quad (5.28b)$$

Bajo estas condiciones, la función $f(x, y)$ puede considerarse como una función de *amplitud modulada* cuyas frecuencias espaciales son v_{x0} y v_{y0} y que está suavemente modulada por $f_0(x, y)$.

Usando la propiedad (D.20) de la transformada de Fourier, encontramos que

$$F(v_x, v_y) = F_0(v_x - v_{x0}, v_y - v_{y0}). \quad (5.29)$$

Tal como muestra la Figura 5.4, si la transparencia de transmitancia $f_0(x, y)$ de-

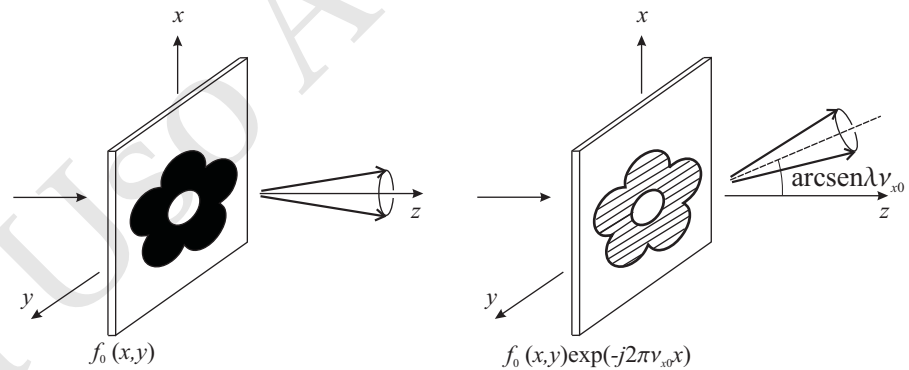


FIGURA 5.4: Deflexión de la onda plana por las transparencias de transmitancia $f_0(x, y)$ y $f(x, y) = f_0(x, y) \exp[-j2\pi(v_{x0}x + v_{y0}y)]$.

flectaba la onda plana en unos rangos angulares $\Delta\theta_x$ y $\Delta\theta_y$ con respecto al eje z , la transparencia de transmitancia dada por (5.27), deflectará la onda plana con los mismos rangos angulares pero centrados éstos ahora con respecto a los ángulos

$$\theta_{x0} = \arcsen(\lambda v_{x0}) \quad (5.30a)$$

$$\theta_{y0} = \arcsen(\lambda v_{y0}). \quad (5.30b)$$

El efecto de la transparencia de transmitancia $f(x, y)$ sería completamente equivalente al de la transparencia de transmitancia $f_0(x, y)$ en contacto con una red de difracción de transmitancia $\exp[-j2\pi(v_{x0}x + v_{y0}y)]$.

La idea anterior puede usarse para realizar un **multiplexado de frecuencias espaciales**. Dos imágenes distintas, $f_1(x, y)$ y $f_2(x, y)$, pueden registrarse en la

misma transparencia si se registran en la forma

$$f(x, y) = f_1(x, y) \exp[-j2\pi(v_{x1}x + v_{y1}y)] + f_2(x, y) \exp[-j2\pi(v_{x2}x + v_{y2}y)], \quad (5.31)$$

puesto que serán deflecionadas en ángulos distintos (esto es, se separarán espacialmente) si la transparencia es iluminada por una onda plana. Este artificio es ampliamente usado en holografía (Apartado 6.5) donde, a menudo, se desea separar dos imágenes que han sido registradas en la misma transparencia.

5.2.4. Modulación en frecuencia

Consideremos ahora una transparencia formada por diferentes regiones de modo que la transmitancia de cada una de ellas es una función armónica de alguna frecuencia espacial determinada (ver Figura 5.5). Si las dimensiones de cada

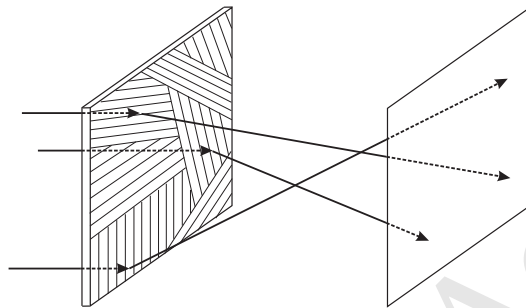


FIGURA 5.5: Deflexión de la luz por una transparencia constituida por diferentes regiones espaciales armónicas.

región son mucho mayores que el periodo espacial de la función armónica definida en ella, entonces

cada región armónica actúa como una red de difracción desviando la luz incidente en diferentes direcciones.

En consecuencia, distintas porciones del frente de onda incidente son desviadas en direcciones distintas. Esta idea puede usarse para crear mapas de interconexión óptica controlando las direcciones en que son deflecionadas las distintas porciones del frente de onda.

Si las distintas regiones espaciales armónicas son *infinitesimales*, entonces podemos considerar que la transmitancia de cierta transparencia viene dada por

$$f(x, y) = \exp[-j2\pi\phi(x, y)], \quad (5.32)$$

donde $\phi(x, y)$ es una función continua de las variables espaciales x e y que varía suavemente en distancias del orden de la longitud de onda de la luz incidente. Llevando a cabo un desarrollo de Taylor de esta función en torno a un punto arbitrario (x_0, y_0) tenemos que

$$\begin{aligned} \phi(x, y) &= \phi(x_0, y_0) + (x - x_0) \left. \frac{\partial \phi}{\partial x} \right|_{(x_0, y_0)} + (y - y_0) \left. \frac{\partial \phi}{\partial y} \right|_{(x_0, y_0)} + \dots \\ &= \phi(x_0, y_0) + (x - x_0) v_{x0} + (y - y_0) v_{y0} + \dots, \end{aligned} \quad (5.33)$$

donde v_{x0} y v_{y0} son las frecuencias espaciales *locales* en el entorno del punto (x_0, y_0) definidas como

$$v_{x0} = \left. \frac{\partial \phi}{\partial x} \right|_{(x_0, y_0)}, \quad v_{y0} = \left. \frac{\partial \phi}{\partial y} \right|_{(x_0, y_0)}. \quad (5.34)$$

Teniendo en cuenta el desarrollo (5.33), la transmitancia $f(x, y)$ se comporta localmente en el entorno de (x_0, y_0) como

$$f(x, y) \approx f_0 \exp[-j2\pi(v_{x0}x + v_{y0}y)], \quad (5.35)$$

siendo f_0 un factor de fase local de valor

$$f_0 = \exp\{-j2\pi[\phi(x_0, y_0) - (v_{x0}x_0 + v_{y0}y_0)]\}.$$

En analogía con la transparencia de transmitancia (5.19), la dada por (5.35) desviará la porción del frente de ondas que incide en la posición (x_0, y_0) hacia una dirección determinada por los ángulos

$$\theta_x = \arcsen \left[\lambda \left. \frac{\partial \phi}{\partial x} \right|_{(x_0, y_0)} \right] \quad (5.36a)$$

$$\theta_y = \arcsen \left[\lambda \left. \frac{\partial \phi}{\partial y} \right|_{(x_0, y_0)} \right], \quad (5.36b)$$

que claramente son *dependientes de la posición*.

EJEMPLO 5.1 Encontrar los ángulos de desviación producidos por una transparencia cuya transmitancia viene dada por

$$f(x, y) = \exp\left(j\pi \frac{x^2}{\lambda f}\right).$$

De acuerdo a (5.36a), la función de fase de la transmitancia anterior es

$$\phi(x, y) = -\frac{x^2}{2\lambda f},$$

por lo que el ángulo de desviación del frente de ondas que llega a un punto (x, y) arbitrario viene dado por

$$\theta_x = \arcsen\left(\lambda \frac{\partial \phi}{\partial x}\right) = \arcsen\left(-\frac{x}{f}\right) \quad (5.37a)$$

$$\theta_y = 0. \quad (5.37b)$$

Si además se cumple que $|x/f| \ll 1$, entonces

$$\theta_x \approx -\frac{x}{f}, \quad (5.38)$$

es decir, el ángulo de desviación es proporcional a la distancia x . Esta idea se puede usar para controlar el ángulo de desviación de un haz de luz que incide sobre la transparencia y construir un *scanner*. Si como muestra la Figura 5.6, una transparencia con transmitancia dada por (5.1) se mueve con velocidad v_x , entonces dado que $x = v_x t$, el ángulo de desviación del haz vendrá dado por

$$\theta_x \approx -\frac{v_x}{f} t.$$

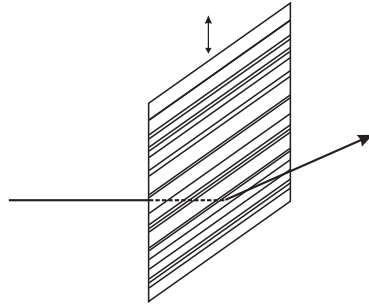


FIGURA 5.6: Una transparencia de frecuencia espacial modulada puede usarse para controlar la desviación de un haz luminoso.

La modulación en frecuencia puede usarse igualmente para simular el efecto de lentes. Si la transmitancia de una lente es $t(x, y)$, construyendo una transparencia cuya transmitancia compleja, $f(x, y)$, sea justamente la transmitancia de la lente, dicha transparencia se comportará igual que la lente. Por ejemplo, una transparencia de transmitancia

$$f(x, y) = \exp\left(j\pi \frac{x^2 + y^2}{\lambda f}\right) \quad (5.39)$$

actuaría exactamente igual que una lente delgada esférica de distancia focal f ya que según se deduce de la expresión (4.21), el efecto de la lente es simplemente introducir un desfase dependiente de la posición, o equivalentemente, desviar los rayos paralelos incidentes hasta la posición del foco.

5.3. Propagación de la luz en el espacio libre

El primer sistema óptico que vamos a estudiar es precisamente el “espacio libre”. Veremos cómo la propagación de la luz en el espacio libre puede concebirse como el efecto de cierto sistema óptico dotado de su correspondiente función de transferencia y respuesta impulsiva.

5.3.1. Función transferencia del espacio libre

Siguiendo la teoría de sistemas lineales, nos planteamos a continuación determinar la amplitud compleja de una onda $U(x, y, z)$ que viaja en el espacio libre desde el plano $z = 0$ (plano de entrada) hasta el plano $z = d$ (plano de salida). Según muestra la Figura 5.7, la onda en $z = 0$ será considerada la entrada al sistema óptico mientras que la onda en $z = d$ será la salida del sistema,

$$\begin{aligned} f(x, y) &= U(x, y, 0) \\ g(x, y) &= U(x, y, d). \end{aligned} \quad (5.40)$$

Evidentemente el espacio libre se comporta como un sistema lineal invariante (es claramente lineal e invariante por naturaleza) por lo que este sistema quedará completamente caracterizado por su función de transferencia $H(\nu_x, \nu_y)$, como se discute en el Apéndice F.

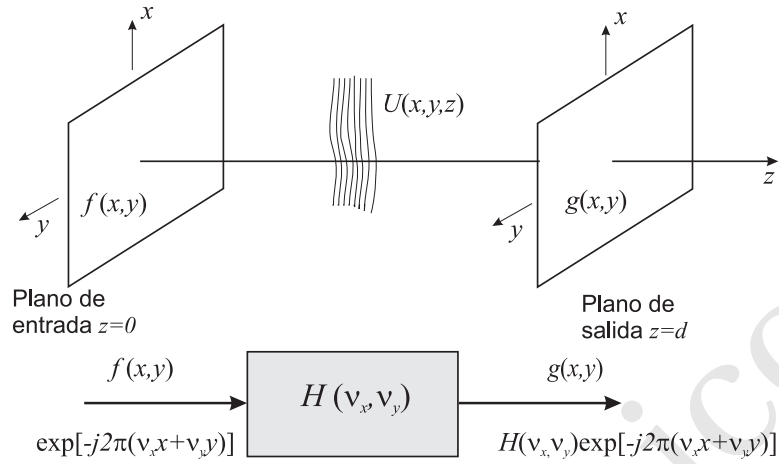


FIGURA 5.7: Propagación de la luz en el espacio libre considerada desde el punto de vista de los sistemas lineales.

Dado que la función de transferencia es el factor por el que una señal armónica de entrada es multiplicada para dar la señal armónica de salida, consideraremos como señal de entrada a

$$f(x, y) = A \exp[-j2\pi(v_x x + v_y y)] . \tag{5.41}$$

Tal como se explicó en el Apartado 5.2.1, a esta señal armónica se le puede asociar una onda plana de longitud de onda λ ,

$$U(x, y, z) = A \exp[-j(k_x x + k_y y + k_z z)] , \tag{5.42}$$

donde

$$k_x = 2\pi v_x \tag{5.43}$$

$$k_y = 2\pi v_y \tag{5.44}$$

$$k_z = \sqrt{k^2 - k_x^2 - k_y^2} = 2\pi \sqrt{\frac{1}{\lambda^2} - v_x^2 - v_y^2} . \tag{5.45}$$

La función a la salida del sistema (es decir, la onda plana tras viajar una distancia d a lo largo del eje óptico) es

$$g(x, y) = A \exp[-j(k_x x + k_y y + k_z d)] , \tag{5.46}$$

por lo que

$$H(v_x, v_y) = \frac{g(x, y)}{f(x, y)} = \exp(-j k_z d) , \tag{5.47}$$

es decir, la función de transferencia del espacio libre será

Función de transferencia del espacio libre

$$H(v_x, v_y) = \exp\left(-j2\pi \sqrt{\frac{1}{\lambda^2} - v_x^2 - v_y^2} d\right) . \tag{5.48}$$

Tal como se discutió en el Apéndice F, una vez que conocemos la función de transferencia del sistema, $H(v_x, v_y)$ (en este caso, espacio libre), podemos construir la respuesta, $g(x, y)$, a cualquier función arbitraria de entrada

$$f(x, y) = \iint_{-\infty}^{\infty} F(v_x, v_y) \exp[-j2\pi(v_x x + v_y y)] dv_x dv_y . \tag{5.49}$$

Para ello necesitamos realizar los siguientes pasos:

1. Calcular la transformada de Fourier de la función entrada,

$$F(v_x, v_y) = \iint_{-\infty}^{\infty} f(x, y) \exp[j2\pi(v_x x + v_y y)] dx dy. \quad (5.50)$$

2. El producto $H(v_x, v_y)F(v_x, v_y)$ proporciona las amplitudes complejas de cada función armónica elemental (onda plana) en el plano de salida.

3. Obtención de la función de salida superponiendo cada una de las funciones armónicas elementales con su correspondiente amplitud:

$$g(x, y) = \iint_{-\infty}^{\infty} F(v_x, v_y) \exp\left(-j2\pi\sqrt{\frac{1}{\lambda^2} - v_x^2 - v_y^2} d\right) \times \exp[-j2\pi(v_x x + v_y y)] dv_x dv_y. \quad (5.51)$$

En la expresión anterior es interesante distinguir dos casos en función de los valores de las frecuencias espaciales del integrando:

- $v_x^2 + v_y^2 \leq 1/\lambda^2$

Para funciones armónicas cuyas frecuencias verifiquen la condición anterior, el efecto del sistema (esto es, el espacio libre) sobre dicha función es producirle un *desplazamiento de fase* a medida que se propaga. La amplitud no se ve alterada.

- $v_x^2 + v_y^2 > 1/\lambda^2$

En este caso, la función de transferencia del espacio libre puede escribirse como

$$H(v_x, v_y) = \exp\left(-2\pi\sqrt{v_x^2 + v_y^2 - \frac{1}{\lambda^2}} d\right), \quad (5.52)$$

siendo ahora una función real que provoca una *atenuación* de la función armónica de entrada. Si $d \gg \lambda$, el factor de atenuación es muy intenso y no llegará señal a la salida.

El factor $1/\lambda$ representa la **frecuencia espacial de corte** del espacio libre, de modo que características de la función de entrada $f(x, y)$ que tengan frecuencias espaciales mayores que $1/\lambda$ (equivalentes a detalles espaciales en $f(x, y)$ menores que λ) no pueden ser transmitidas distancias $d \gg \lambda$ mediante una onda luminosa de longitud de onda λ .

Aproximación de Fresnel

Aunque la expresión (5.51) proporciona un método sistemático y riguroso para obtener el efecto de la propagación en el espacio libre de una onda luminosa, la forma del integrando de (5.51) es excesivamente complicada para poder obtener expresiones cerradas en la mayoría de los casos. Usualmente hemos de recurrir a ciertas aproximaciones que sí nos permitirán deducir expresiones simplificadas y más manejables.

Si la función de entrada, $f(x, y)$, tiene una característica espectral tal que

$$v_x^2 + v_y^2 \ll 1/\lambda^2, \quad (5.53)$$

entonces es posible simplificar la expresión (5.48) de la función de transferencia del espacio libre.

La expresión (5.53) es generalmente compatible con la condición de rayos paraxiales, es decir, ángulos θ_x y θ_y suficientemente pequeños de modo que encontráramos (ver expresiones (5.12a) y (5.12b)) que

$$\theta_x \approx \lambda v_x \quad , \quad \theta_y \approx \lambda v_y .$$

Denominando

$$\theta^2 = \theta_x^2 + \theta_y^2 ,$$

podemos hacer el siguiente desarrollo para pequeños valores de θ^2 :

Aproximación de Fresnel

$$2\pi\sqrt{\frac{1}{\lambda^2} - v_x^2 - v_y^2} d = 2\pi\frac{d}{\lambda}\sqrt{1 - \theta^2} \approx 2\pi\frac{d}{\lambda}\left(1 - \frac{\theta^2}{2}\right) , \quad (5.54)$$

Función de transferencia del espacio libre bajo la aproximación de Fresnel

por lo que la función de transferencia del vacío puede escribirse como

$$H(v_x, v_y) = H_0 \exp[j\pi\lambda d(v_x^2 + v_y^2)] , \quad (5.55)$$

siendo H_0 un factor de fase constante de valor

$$H_0 = \exp(-jkd) . \quad (5.56)$$

Tal como se discutió en el Apartado 1.9.4, la condición de validez de la aproximación de Fresnel deberá expresarse en el presente caso como

$$\frac{\theta^4 d}{4\lambda} \ll 1 . \quad (5.57)$$

Si a representa la distancia radial mayor en el plano de salida y $\theta_m = a/d$ es el ángulo mayor con respecto al eje z , entonces la condición (5.57) puede escribirse como

$$\frac{N_F \theta_m^2}{4} \ll 4 \quad ; \quad N_F = \frac{a^2}{\lambda d} . \quad (5.58)$$

Teniendo en cuenta la aproximación de Fresnel (rayos paraxiales), si la señal de entrada es $f(x, y)$, la señal a la salida puede expresarse como

$$g(x, y) \approx H_0 \iint_{-\infty}^{\infty} F(v_x, v_y) \exp[j\pi\lambda d(v_x^2 + v_y^2)] \times \exp[-j2\pi(v_x x + v_y y)] dv_x dv_y . \quad (5.59)$$

Respuesta impulsiva del espacio libre

Una vez conocida la función de transferencia, $H(v_x, v_y)$, de un sistema, podemos encontrar la función respuesta impulsiva, $h(x, y)$, obteniendo su transformada de Fourier inversa (ver expresión (F.13)). Para el caso del espacio libre bajo la aproximación de Fresnel, tendremos entonces que

$$\begin{aligned} h(x, y) &= \iint_{-\infty}^{\infty} H(v_x, v_y) \exp[-j2\pi(v_x x + v_y y)] dv_x dv_y \\ &= \iint_{-\infty}^{\infty} H_0 \exp[j\pi\lambda d(v_x^2 + v_y^2)] \exp[-j2\pi(v_x x + v_y y)] dv_x dv_y . \end{aligned} \quad (5.60)$$

Operando en (5.60) obtendríamos que

$$h(x, y) = h_0 \exp\left(-jk \frac{x^2 + y^2}{2d}\right), \quad (5.61)$$

Función respuesta impulsiva del espacio libre bajo la aproximación de Fresnel

siendo h_0 una constante de valor

$$h_0 = \frac{j}{\lambda d} \exp(-jkd). \quad (5.62)$$

Observemos que la respuesta impulsiva del espacio libre (5.61) tiene la forma de una onda parabólica centrada en el origen del plano de entrada. En este sentido, y tal como muestra la Fig.5.8, podemos decir que cada punto del plano de

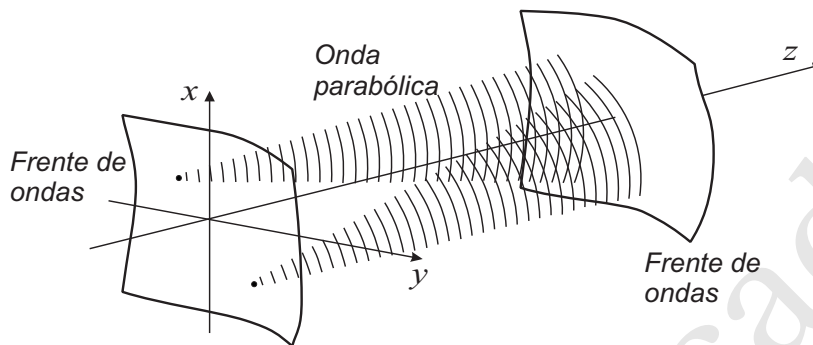


FIGURA 5.8: Construcción de la señal en el plano de salida mediante la convolución de ondas parabólicas.

entrada genera una onda parabólica que se superpone a las demás en el plano de salida (esto proporciona una justificación del principio de Huygens). La imagen en el plano de salida puede entonces expresarse mediante la siguiente integral de convolución:

$$g(x, y) = \iint_{-\infty}^{\infty} f(\xi, \eta) h(x - \xi, y - \eta) d\xi d\eta, \quad (5.63)$$

que bajo la aproximación de Fresnel puede expresarse como

$$g(x, y) = h_0 \iint_{-\infty}^{\infty} f(\xi, \eta) \exp\left[-j\pi \frac{(x - \xi)^2 + (y - \eta)^2}{\lambda d}\right] d\xi d\eta. \quad (5.64)$$

Observemos que bajo la aproximación de Fresnel hemos encontrado las dos siguientes maneras de determinar la amplitud compleja $g(x, y)$ en el plano de salida:

1. *En el dominio de la frecuencia.*

Mediante la descomposición en ondas planas de la señal de entrada y su relación con funciones armónicas de frecuencias ν_x y ν_y , la amplitud compleja a la salida venía dada por la expresión (5.59).

2. *En el dominio espacial.*

Cuando la onda de entrada es descompuesta en ondas parabólicas elementales obtenemos la amplitud compleja en el plano de salida a partir de la expresión (5.64).

5.4. Transformada de Fourier óptica

En los anteriores apartados hemos visto que la propagación de la luz en el espacio libre podía describirse convenientemente haciendo uso del análisis de Fourier. La amplitud compleja de una onda monocromática de longitud de onda λ en el plano de entrada, $z = 0$, dada por la función $f(x, y)$ puede descomponerse en funciones armónicas de diferentes frecuencias espaciales. A cada una de estas funciones armónicas le podemos asociar una onda plana viajando en una dirección fijada por los ángulos θ_x y θ_y con una amplitud dada por $F(v_x, v_y)$ —esto es, la transformada de Fourier de $f(x, y)$. Tal como se discutió en el Apartado 5.2.2, esto sugiere que la luz puede actuar como un mecanismo que realiza transformadas de Fourier, para lo cual tendríamos que tener una transparencia de transmitancia en amplitud $f(x, y)$ que sea iluminada con una onda plana de amplitud unidad.

En general, todas las ondas planas se superponen entre sí en el plano de salida, por lo que si queremos obtener la transformada de Fourier debemos encontrar un método para separar las distintas ondas planas. Comprobaremos que para distancias suficientemente alejadas del plano de salida, sólo una onda plana contribuye a la amplitud de cada punto en el plano de salida. En la práctica, esto se consigue mediante lentes que enfoquen cada onda plana en un punto distinto del plano focal.

5.4.1. Aproximación de Fresnel

Operando en la expresión (5.64), que nos daba la amplitud compleja $g(x, y)$ en el plano de salida a partir de la respuesta impulsiva del espacio libre bajo la aproximación de Fresnel, encontraremos una forma alternativa de expresar $g(x, y)$ relacionada con la transformada de Fourier de cierta función.

En el integrando (5.64) notemos que las variables (ξ, η) hacen referencia al *plano de entrada* mientras que las variables (x, y) se refieren al *plano de salida*. Teniendo ahora en cuenta que

$$\exp\left[-j\pi \frac{(x-\xi)^2 + (y-\eta)^2}{\lambda d}\right] = \exp\left(-j\pi \frac{x^2 + y^2}{\lambda d}\right) \exp\left(-j\pi \frac{\xi^2 + \eta^2}{\lambda d}\right) \exp\left(j2\pi \frac{x\xi + y\eta}{\lambda d}\right). \quad (5.65)$$

la integral de convolución (5.64) puede reescribirse como

$$g(x, y) = h_0 \exp\left(-j\pi \frac{x^2 + y^2}{\lambda d}\right) \times \iint_{-\infty}^{\infty} \left\{ f(\xi, \eta) \exp\left(-j\pi \frac{\xi^2 + \eta^2}{\lambda d}\right) \right\} \exp\left(j2\pi \frac{x\xi + y\eta}{\lambda d}\right) d\xi d\eta, \quad (5.66)$$

donde la integral doble en la expresión anterior puede identificarse como la transformada de Fourier de la función

$$\tilde{f}(\xi, \eta) = f(\xi, \eta) \exp\left(-j\pi \frac{\xi^2 + \eta^2}{\lambda d}\right) \quad (5.67)$$

(esto es, la función f multiplicada por un factor cuadrático de fase) con respecto

a las variables:¹

$$v_x = \frac{x}{\lambda d}, \quad v_y = \frac{y}{\lambda d}. \quad (5.68)$$

En consecuencia, la expresión (5.66) puede finalmente expresarse como

$$g(x, y) = h_0 \exp\left(-j\pi \frac{x^2 + y^2}{\lambda d}\right) \tilde{F}\left(\frac{x}{\lambda d}, \frac{y}{\lambda d}\right). \quad (5.69)$$

Imagen en el plano de salida bajo la aproximación de Fresnel

Bajo la aproximación de Fresnel, la amplitud compleja en el plano de salida se ha podido expresar en términos de la transformada de Fourier de la función $\tilde{f}(x, y)$.

5.4.2. Aproximación de Fraunhofer: Transformada de Fourier en el campo lejano

A continuación comprobaremos que a distancias suficientemente alejadas del plano de entrada, la única onda plana que contribuye a la amplitud compleja en el punto (x, y) es aquella cuya dirección está determinada por los ángulos $\theta_x \approx x/d$ y $\theta_y \approx y/d$.

En el supuesto de que el dominio de la función f en el plano de entrada esté confinado a un pequeño círculo de radio b , es decir,

$$\xi^2 + \eta^2 \leq b^2, \quad (5.70)$$

de modo que se cumpla la condición de Fraunhofer (1.129) en el plano de entrada (esto es, que el número de Fresnel, N_F^{ent} , en el plano de entrada sea muy pequeño),

$$\frac{\xi^2 + \eta^2}{\lambda d} \leq \frac{b^2}{\lambda d} \equiv N_F^{\text{ent}} \rightarrow 0, \quad (5.71)$$

tenemos que el factor de fase cuadrático

$$\exp\left(-j\pi \frac{\xi^2 + \eta^2}{\lambda d}\right) \rightarrow 1$$

en los puntos donde $f(\xi, \eta) \neq 0$ y, por tanto, la función en el plano de salida puede escribirse como

$$g(x, y) = h_0 \exp\left(-j\pi \frac{x^2 + y^2}{\lambda d}\right) \iint_{-\infty}^{\infty} f(\xi, \eta) \exp\left(j2\pi \frac{x\xi + y\eta}{\lambda d}\right) d\xi d\eta, \quad (5.72)$$

o, equivalentemente,

$$g(x, y) = h_0 \exp\left(-j\pi \frac{x^2 + y^2}{\lambda d}\right) F\left(\frac{x}{\lambda d}, \frac{y}{\lambda d}\right). \quad (5.73)$$

Si, en el análisis de la imagen en el plano de salida, nos limitamos ahora a puntos dentro de un círculo de radio a centrado alrededor del eje z de modo que el número de Fresnel en el plano de salida es muy reducido,

$$\frac{x^2 + y^2}{\lambda d} \leq \frac{a^2}{\lambda d} \equiv N_F^{\text{sal}} \rightarrow 0, \quad (5.74)$$

entonces, podemos expresar la imagen $g(x, y)$ en el plano de salida del sistema óptico como

Imagen en el plano de salida bajo la aproximación de Fraunhofer

¹ Notemos que la transformada de Fourier, $Q(\beta_x, \beta_y)$, de la función $q(\xi, \eta)$ viene dada por

$$Q(\beta_x, \beta_y) = \iint_{-\infty}^{\infty} q(\xi, \eta) \exp[j2\pi(\beta_x \xi + \beta_y \eta)] d\xi d\eta,$$

por lo que identificando $q(\xi, \eta)$ con $\tilde{f}(\xi, \eta)$ y las variables $\beta_x = x/\lambda d$, $\beta_y = y/\lambda d$ en (5.66) obtenemos la equivalencia señalada.

$$g(x, y) = h_0 F\left(\frac{x}{\lambda d}, \frac{y}{\lambda d}\right). \quad (5.75)$$

Hemos obtenido, por tanto, que

Para distancias suficientemente alejadas del plano de entrada ($z = 0$), la amplitud compleja, $g(x, y)$, de una onda de longitud de onda λ en el plano de salida $z = d$ es proporcional a la transformada de Fourier $F(v_x, v_y)$ de la amplitud compleja $f(x, y)$ en el plano de entrada evaluada en las frecuencias espaciales

$$v_x = \frac{x}{\lambda d}, \quad v_y = \frac{y}{\lambda d}.$$

Desde un punto de vista físico podemos interpretar el resultado anterior observando que en el punto (x, y) del plano de salida, todas las ondas planas que no satisfacen la condición señalada, interfieren destructivamente.

5.4.3. Transformada de Fourier usando lentes

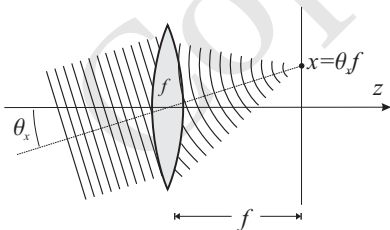
La obtención de la transformada de Fourier mediante el procedimiento expuesto en el apartado anterior está obviamente sujeto a que se verifiquen las condiciones impuestas por (5.71) y (5.74). Según vimos en el Ejemplo 1.5, las condiciones prácticas en las que se verifica la aproximación de Fraunhofer exigen comúnmente una gran distancia entre el plano de salida y el plano de entrada (si a o b son del orden del centímetro, entonces d debe ser del orden de cientos de metros). Puesto que estas condiciones son a menudo difíciles de obtener en la práctica, es muy conveniente que busquemos otras posibilidades para realizar la transformada de Fourier. Tal como veremos a continuación, esta posibilidad nos la proporcionará el simple uso de una lente.

Dado que bajo la aproximación de Fresnel, una lente delgada de distancia focal f transforma una onda plana incidente en una onda parabólica que focaliza en cierto punto del plano focal de la lente (ver expresiones (4.25a)–(4.25b) y su correspondiente discusión), es fácil entender que las ondas planas cuya superposición constituye una onda cualquiera pueden, en consecuencia, separarse haciendo uso de una lente de distancia focal f .

Si una onda plana incide *paraxialmente* sobre la lente delgada formando unos ángulos θ_x y θ_y pequeños, la onda parabólica resultante focaliza en un punto del plano focal de la lente de coordenadas

$$x \approx f\theta_x, \quad y \approx f\theta_y. \quad (5.76)$$

Usando esta propiedad, y tal como muestra la Figura 5.9, podemos observar que si sobre la lente incide una onda arbitraria, la lente separará las distintas ondas planas constituyentes de la onda incidente.



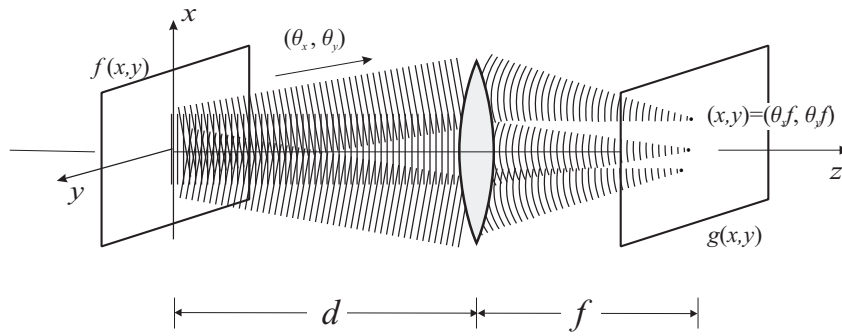


FIGURA 5.9: La lente focaliza las distintas ondas planas componentes en distintos puntos del plano focal.

Consideremos una onda plana viajando a lo largo del eje óptico,

$$U(x, y, z) = A \exp(-jkz),$$

que incide en el sistema óptico mostrada en la figura adjunta, donde $t_A(x, y)$ y $t_L(x, y)$ son las transmitancias de la transparencia y de la lente respectivamente. Si A es la amplitud compleja de la onda que incide sobre la transparencia, entonces tras la transparencia, la amplitud de la onda será

$$f(x, y) = A t_A(x, y), \tag{5.77}$$

cuya transformada de Fourier puede expresarse como

$$F(v_x, v_y) = \mathcal{F}\{f(x, y)\}.$$

Si tenemos en cuenta la expresión (5.55) que nos da la función de transferencia del espacio libre bajo la aproximación de Fresnel, entonces, en el plano *anterior* a la lente, tendremos que la transformada de Fourier de la onda viene dada por

$$Q_1(v_x, v_y) = F(v_x, v_y) \exp\left[j\pi\lambda d(v_x^2 + v_y^2)\right], \tag{5.78}$$

(no se ha considerado el factor de fase H_0) y, consecuentemente, la correspondiente amplitud compleja $q_1(x, y)$ como la siguiente transformada de Fourier inversa:

$$q_1(x, y) = \mathcal{F}^{-1}\{Q_1(v_x, v_y)\}. \tag{5.79}$$

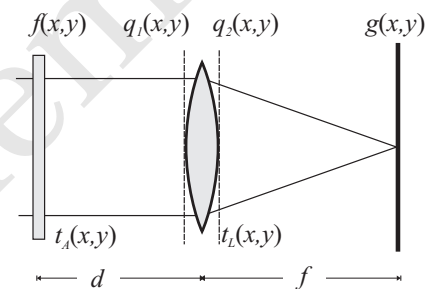
El efecto de la lente sobre la onda incidente puede ahora determinarse mediante la transmitancia de la lente (ver expresión (4.21)), de modo que en el plano *posterior* a la lente (que se caracterizará mediante las variables ξ y η), la amplitud compleja de la onda resultante, $q_2(\xi, \eta)$, vendrá dada por

$$q_2(\xi, \eta) = q_1(\xi, \eta) \exp\left[j\frac{k}{2f}(\xi^2 + \eta^2)\right], \tag{5.80}$$

donde tampoco se ha considerado factor de fase $h_0 = \exp(-jkf)$.

Finalmente en el plano focal de la lente (distante una distancia $d = f$ de la lente), la amplitud compleja $g(x, y)$ puede obtenerse usando la expresión general (5.66):

$$g(x, y) = \frac{\exp\left(-j\pi\frac{x^2+y^2}{\lambda f}\right)}{\lambda f} \times \iint_{-\infty}^{\infty} \left\{ q_2(\eta, \xi) \exp\left(-j\pi\frac{\xi^2 + \eta^2}{\lambda f}\right) \right\} \exp\left(j2\pi\frac{x\xi + y\eta}{\lambda f}\right) d\xi d\eta. \tag{5.81}$$



(En la expresión anterior no se ha considerado el factor de fase que aparece en (5.62) y que afecta a la constante h_0). Si tenemos en cuenta que el factor de fase cuadrático de $q_2(\eta, \xi)$ (ver expresión (5.80)) se cancela con el que aparece explícitamente en el integrando anterior, la amplitud compleja en el plano focal puede escribirse como

Distribución luminosa en el plano focal de una lente de distancia focal f

$$g(x, y) = \frac{\exp\left(-j\pi \frac{x^2+y^2}{\lambda f}\right)}{\lambda f} \iint_{-\infty}^{\infty} q_1(\eta, \xi) \exp\left(j2\pi \frac{x\xi + y\eta}{\lambda f}\right) d\xi d\eta$$

$$= \frac{\exp\left(-j\pi \frac{x^2+y^2}{\lambda f}\right)}{\lambda f} Q_1\left(\frac{x}{\lambda f}, \frac{y}{\lambda f}\right). \quad (5.82)$$

Teniendo ahora en cuenta la relación (5.78) entre Q_1 y F , (5.82) puede reescribirse como

$$g(x, y) = \frac{\exp\left(-j\pi \frac{x^2+y^2}{\lambda f}\right)}{\lambda f} F\left(\frac{x}{\lambda f}, \frac{y}{\lambda f}\right) \exp\left[j\pi \lambda d \left(\frac{x^2+y^2}{\lambda^2 f^2}\right)\right]$$

$$= \frac{\exp\left[j\pi \lambda \frac{(x^2+y^2)(d-f)}{\lambda^2 f^2}\right]}{\lambda f} F\left(\frac{x}{\lambda f}, \frac{y}{\lambda f}\right). \quad (5.83)$$

Notemos que el factor de fase que acompaña a F desaparece si $d = f$, por lo que si la transparencia hubiese estado situada justamente en el plano focal objeto de la lente, entonces

$$g(x, y) = F\left(\frac{x}{\lambda f}, \frac{y}{\lambda f}\right). \quad (5.84)$$

En consecuencia,

la amplitud compleja de la luz en el punto (x, y) del plano focal imagen de una lente de distancia focal f es proporcional a la transformada de Fourier de la amplitud compleja en el plano focal objeto evaluada en las frecuencias espaciales

$$v_x = \frac{x}{\lambda f} \quad v_y = \frac{y}{\lambda f}.$$

Podemos observar que el factor de fase antes mencionado (que sólo se anula si $d = f$) no afecta si estamos interesados en la *intensidad* de la luz en el plano focal imagen,

$$I(x, y) = |g(x, y)|^2,$$

puesto que en este caso

Intensidad óptica en el plano focal imagen de una lente

$$I(x, y) = \frac{1}{(\lambda f)^2} \left| F\left(\frac{x}{\lambda f}, \frac{y}{\lambda f}\right) \right|^2 \quad (5.85)$$

y, por consiguiente, podemos concluir que

la intensidad óptica en el plano focal imagen es proporcional al cuadrado del módulo de la transformada de Fourier de la amplitud compleja de la onda en el plano de entrada, evaluada en las frecuencias espaciales

$$v_x = \frac{x}{\lambda f} \quad v_y = \frac{y}{\lambda f}.$$

5.5. Problemas propuestos

5.1: Demostrar que la aproximación de Fraunhofer es más restrictiva que la de Fresnel, tomando $\lambda = 0.5 \mu\text{m}$, asumiendo que los puntos del objeto, (x, y) , están contenidos en un círculo de radio $b = 1 \text{ cm}$, y determinando el rango de distancias d para las cuales se pueden aplicar ambas aproximaciones.

Sol: Aprox. de Fresnel $d \geq \sqrt[3]{\frac{1}{2}} \approx 0.794 \text{ m}$.

Aprox. de Fraunhofer $\frac{b^2}{\lambda d} \leq 10^{-2} \Leftrightarrow d \geq 200 \text{ m}$.

5.2: Las amplitudes complejas de una onda monocromática de longitud de onda $\lambda = 633 \text{ nm}$ en los planos $z = 0$ y $z = d$ son $f(x, y)$ y $g(x, y)$, respectivamente. Si $f(x, y)$ es una función con simetría circular y frecuencia espacial máxima de 200 líneas/mm, determinar el ángulo del cono dentro del cual están confinadas las direcciones de onda.

Sol: $\theta_{\max} = \arcsen(0.2\lambda) = 0.13 \text{ rad}$.

5.3: Una transparencia de transmitancia en amplitud $t(x, y) = \exp[-j2\pi\phi(x)]$ se ilumina con una onda plana uniforme de longitud de onda $\lambda = 1 \mu\text{m}$. La luz transmitida se enfoca con una lente adyacente de distancia focal $f = 100 \text{ cm}$. **a)** ¿Cuál debe ser $\phi(x)$ para que el rayo que incide sobre la transparencia en la posición x sea deflektado y enfocado en una posición $x' = 1/x$? **b)** Si se elimina la lente, ¿cómo habría que modificar $\phi(x)$ para que el sistema realizase la misma función?

Ayuda: Tenga en cuenta los contenidos del Apartado 5.2.4 y el hecho de que una lente no desvía los rayos que atraviesan su parte central.

Sol: **a)** $\phi(x) = 10^6 \ln x$; **b)** El sistema se comportaría de igual manera.

Copia Uso Académico

Tema 6

Aplicaciones de la óptica de Fourier

6.1. Introducción

Una vez que se han visto los fundamentos de la óptica de Fourier en el tema anterior, el presente tema estará dedicado a la aplicación de esta técnica al estudio de cuatro fenómenos ópticos de mucha importancia práctica: difracción, formación de imágenes, procesado de imágenes y holografía. Veremos que la óptica de Fourier nos proporciona un marco muy adecuado para el estudio los fenómenos anteriores.

6.2. Difracción de la luz

Aunque el fenómeno de difracción ondulatoria se ha discutido de forma básica en el Apartado 1.7, a continuación trataremos este importante fenómeno desde un punto de vista más riguroso haciendo uso de la teoría desarrollada en el Tema 5.

Recordemos que lo que se conoce como **patrón de difracción** luminoso es la distribución de la intensidad de una onda luminosa que ha sido distorsionada en su propagación en el espacio libre por un obstáculo (por ejemplo: una abertura/-rendija en una pantalla opaca). Debido a la naturaleza ondulatoria de la luz, el patrón de difracción puede ser muy parecido o muy distinto a la sombra del obstáculo, dependiendo este hecho de la distancia entre el obstáculo y el plano de observación, la longitud de onda de la onda luminosa y las dimensiones del obstáculo. En general, el patrón de difracción difiere apreciablemente de la sombra del obstáculo si el tamaño de éste es del orden de la longitud de onda. Para la luz, este hecho requiere que los obstáculos sean muy pequeños por lo que usualmente no se observa el fenómeno de difracción en la vida diaria (además del hecho de que se necesitan fuentes de luz monocromáticas y coherentes).

La formulación teórica más elemental de la difracción luminosa se basa en la hipótesis de que la onda luminosa escalar no sufre cambios apreciables al atravesar el plano del obstáculo, excepto que la amplitud compleja de la onda se hace nula en la parte posterior de las zonas opacas del obstáculo (esta hipótesis es sólo una aproximación que da resultados razonablemente buenos en multitud de

situaciones prácticas).

Si en la Figura 6.1, la transmitancia en amplitud del obstáculo, también cono-

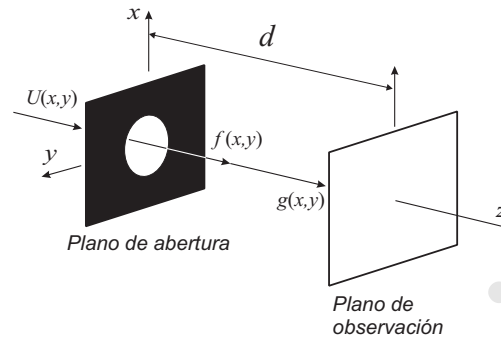


FIGURA 6.1: Tras el obstáculo de transmitancia $p(x, y)$, la onda de amplitud compleja $U(x, y)$ se transforma en $f(x, y) = U(x, y)p(x, y)$

cida como **función abertura**, viene dada por $p(x, y)$, siendo

$$p(x, y) = \begin{cases} 1 & \text{zonas libres} \\ 0 & \text{zonas opacas,} \end{cases} \quad (6.1)$$

y $U(x, y)$ representa la amplitud compleja de la onda incidente en el plano del obstáculo ($z = 0$), entonces la amplitud compleja justamente tras el obstáculo será

$$f(x, y) = U(x, y)p(x, y). \quad (6.2)$$

La amplitud compleja, $g(x, y)$, en el plano de observación $z = d$ puede calcularse usando haciendo uso de las técnicas de la transformada de Fourier óptica desarrolladas en el Apartado 5.4. El patrón de difracción, $I(x, y)$, será simplemente

$$I(x, y) = |g(x, y)|^2. \quad (6.3)$$

Según la aproximación usada para calcular $g(x, y)$ se distingue entre difracción de Fraunhofer o bien difracción de Fresnel. Debido a su mayor simplicidad, nosotros estudiaremos únicamente la difracción de Fraunhofer.

6.2.1. Difracción de Fraunhofer

Las condiciones que hacían válida la aproximación de Fraunhofer podían resumirse en que los números de Fresnel en los planos de entrada y salida fuesen muy pequeños. En concreto, si el obstáculo es una abertura en una pantalla opaca cuya dimensión radial máxima es b , entonces debemos exigir que

Condición de validez de la aproximación de Fraunhofer

$$N_F^{\text{ent}} \rightarrow 0 \Rightarrow \frac{b^2}{\lambda d} \ll 1. \quad (6.4)$$

Suponiendo que la onda incidente es una onda plana de intensidad I_i que se propaga en la dirección z , de modo que

$$U(x, y) = I_i^{1/2}, \quad (6.5)$$

entonces la amplitud compleja de la onda justo detrás del plano del obstáculo será

$$f(x, y) = I_i^{1/2} p(x, y). \quad (6.6)$$

Haciendo uso de la expresión (5.75) podemos escribir que la amplitud compleja, $g(x, y)$, en el plano de observación viene dada por

$$g(x, y) = I_i^{1/2} h_0 P\left(\frac{x}{\lambda d}, \frac{y}{\lambda d}\right); \quad h_0 = \frac{j}{\lambda d} \exp(-jkd), \quad (6.7)$$

donde

$$P(v_x, v_y) = \iint_{-\infty}^{\infty} p(x, y) \exp[j2\pi(v_x x + v_y y)] dv_x dv_y \quad (6.8)$$

es la transformada de Fourier de la función abertura.

El patrón de difracción será, por tanto,

$$I(x, y) = \frac{I_i}{(\lambda d)^2} \left| P\left(\frac{x}{\lambda d}, \frac{y}{\lambda d}\right) \right|^2. \quad (6.9)$$

Patrón de difracción de Fraunhofer de una abertura

La expresión anterior sigue siendo válida incluso en situaciones en las que $N_F^{\text{sal}} \rightarrow 0$ puesto que, según nos muestra la expresión (5.73), el factor de fase dependiente de x e y desaparecería al tomar módulo. Por consiguiente podemos concluir que

el patrón de difracción de Fraunhofer en un punto (x, y) del plano de observación es proporcional al módulo al cuadrado de la transformada de Fourier de la función abertura, evaluada en las frecuencias espaciales

$$v_x = \frac{x}{\lambda d} \quad y \quad v_y = \frac{y}{\lambda d}.$$

EJEMPLO 6.1 Obtener el patrón de difracción de Fraunhofer producido por una abertura rectangular de dimensiones D_x y D_y .

La abertura rectangular está caracterizada por la siguiente función abertura:

$$p(x, y) = \text{rect}\left(\frac{x}{D_x}\right) \text{rect}\left(\frac{y}{D_y}\right),$$

y dado que su patrón de difracción vendrá dado por la expresión (6.9), debemos obtener la transformada de Fourier de la anterior función abertura.

Antes de aplicar (6.9), tengamos en cuenta que

$$\mathcal{F}\{\text{rect}(x)\text{rect}(y)\} = \text{senc}(\pi v_x) \text{senc}(\pi v_y)$$

(siendo $\text{senc}(x) = \frac{\text{sen } x}{x}$) y que

$$\mathcal{F}\{f(ax, by)\} = \frac{1}{ab} F\left(\frac{v_x}{a}, \frac{v_y}{b}\right),$$

entonces

$$\begin{aligned} \mathcal{F}\{p(x, y)\} &= P(v_x, v_y) \\ &= D_x D_y \text{senc}(\pi D_x v_x) \text{senc}(\pi D_y v_y), \end{aligned}$$

por lo que

$$P\left(\frac{x}{\lambda d}, \frac{y}{\lambda d}\right) = D_x D_y \text{senc}\left(\frac{\pi D_x}{\lambda d} x\right) \text{senc}\left(\frac{\pi D_y}{\lambda d} y\right)$$

y consecuentemente

$$I(x, y) = I_0 \text{senc}^2\left(\frac{\pi D_x}{\lambda d} x\right) \text{senc}^2\left(\frac{\pi D_y}{\lambda d} y\right),$$

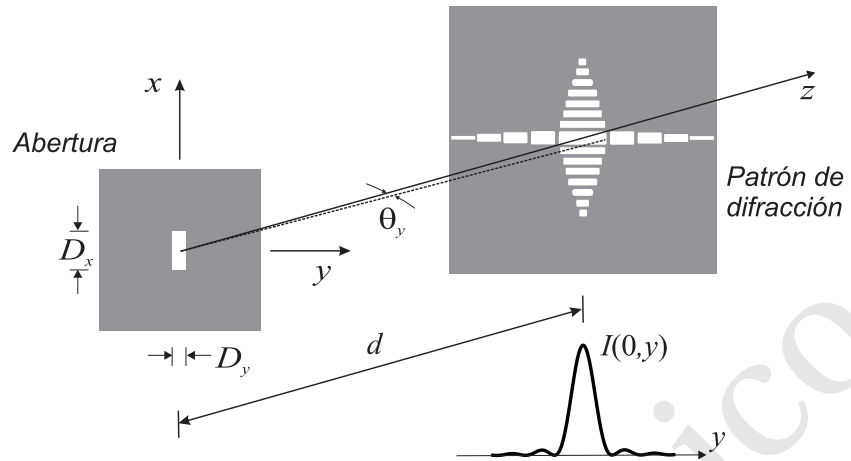


FIGURA 6.2: Difracción de Fraunhofer de una abertura rectangular.

donde

$$I_0 = \frac{D_x^2 D_y^2}{(\lambda d)^2} I_i.$$

Esta función presenta mínimos de intensidad en la dirección x en aquellos puntos x_{\min} que verifiquen

$$\frac{\pi D_x}{\lambda d} x_{\min} = n\pi \Rightarrow x_{\min} = n\lambda \frac{d}{D_x} \quad (n \neq 0).$$

Nótese la concordancia, bajo la presente aproximación paraxial, de este resultado con el obtenido anteriormente en (1.70).

En el ejemplo anterior es importante notar que no podríamos analizar la difracción por una rendija en la que $D_y \rightarrow \infty$ siguiendo la expresión (6.9) puesto que dicha situación no cumpliría la condición (6.4) de validez de la aproximación de Fraunhofer. Por la misma razón, no tendría mucho sentido aplicar la aproximación de Fraunhofer para obtener la difracción producida por un obstáculo en vez de una abertura. Por consiguiente debemos tener en cuenta que la expresión (6.9) no puede aplicarse al estudio de la difracción por obstáculos (en este último caso estaríamos en las condiciones de *difracción de Fresnel* que no serán tratados aquí).

6.3. Formación de imágenes

Bajo las aproximaciones consideradas en el Tema 3, vimos que los sistemas ópticos tratados siempre formaban imágenes ideales de los objetos en cierto plano (aparte de la posible magnificación). No obstante, debemos tener en cuenta que algunos fenómenos ineludibles como el enfoque imperfecto y la difracción provocan que los sistemas ópticos reales nunca sean formadores de imágenes perfectos, es decir, nunca obtendremos una réplica fiel del objeto en el plano imagen. En el presente Apartado, trataremos brevemente estos fenómenos y veremos cómo es posible minimizarlos.

6.3.1. Lente simple como sistema de formación de imágenes

En general, la propiedad más conocida de una lente es su habilidad para formar imágenes. Situando un objeto iluminado en frente de una lente (plano objeto), bajo condiciones apropiadas, aparece un segundo plano con una distribución de intensidad luminosa muy cercana a la del objeto. Esta distribución de intensidad luminosa es lo que se conoce como *imagen* del objeto.

A continuación estudiaremos las condiciones bajo las cuales la amplitud compleja, $g(x, y)$, en el plano de salida puede decirse razonablemente que es una imagen de la distribución $f(x, y)$ en el plano objeto; es decir, cuando $g(x, y) \cong f(x, y)$. Debido al carácter lineal del sistema podemos escribir que

$$g(x, y) = \iint_{-\infty}^{\infty} f(\xi, \eta) h(x, y; \xi, \eta) d\xi d\eta. \quad (6.10)$$

El sistema óptico producirá "imágenes de calidad" cuando $g(x, y)$ sea muy parecida a $f(x, y)$. Para que esto ocurra debemos tener una respuesta impulsiva del sistema, $h(x, y; \xi, \eta)$, que se parezca mucho a una función delta,

$$h(x, y; \xi, \eta) \approx K\delta(x \pm M\xi, y \pm M\eta) \quad (6.11)$$

(M es la magnificación del sistema). La condición anterior es equivalente a que la imagen de un punto sea muy aproximadamente otro punto.

Para obtener la respuesta impulsiva del sistema procederemos a calcular la imagen, $h(x, y)$, en el plano de salida correspondiente a un punto del plano objeto centrado en $(x, y) = (0, 0)$, tal como muestra la Figura 6.3. Este punto es el

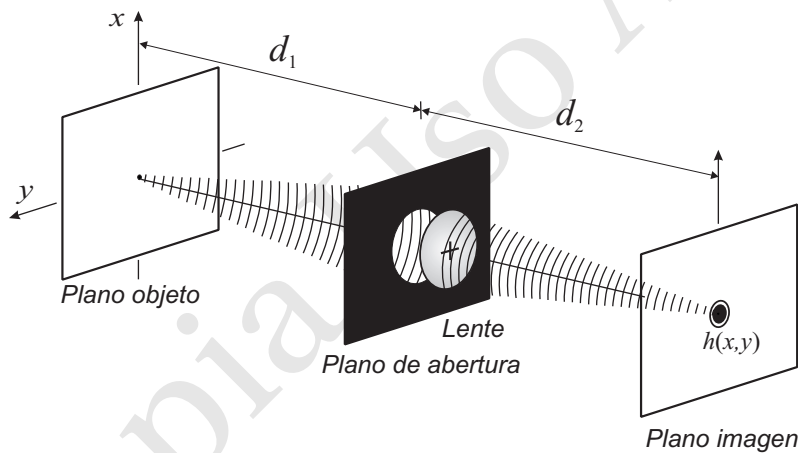
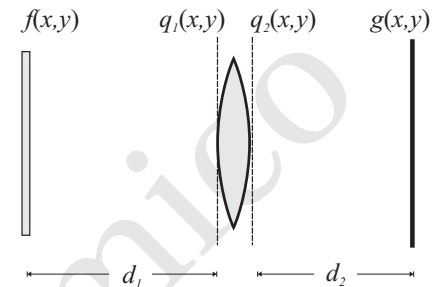


FIGURA 6.3: Sistema de imágenes formada por una lente simple

origen de una onda esférica que, bajo aproximación paraxial, puede expresarse en el plano anterior de la lente como

$$q_1(x, y) = h_1 \exp\left(-jk \frac{x^2 + y^2}{2d_1}\right) \quad h_1 = \frac{j}{\lambda d_1} \exp(-jk d_1). \quad (6.12)$$

Después de pasar por la apertura y por la lente, la amplitud compleja será

$$q_2(x, y) = q_1(x, y) \exp\left(jk \frac{x^2 + y^2}{2f}\right) p(x, y), \quad (6.13)$$

donde $p(x, y)$ es la función abertura, que aparece bien porque haya una abertura física o bien debido al *tamaño finito* de la lente.

Después de propagarse una distancia d_2 en el espacio libre, la aplicación de (5.64) nos dice que la imagen del punto vendrá dada por

$$h(x, y) = h_2 \iint_{-\infty}^{\infty} q_2(u, v) \exp\left[-j\pi \frac{(x-u)^2 + (y-v)^2}{\lambda d_2}\right] du dv. \quad (6.14)$$

Al sustituir (6.13) en (6.14) y llevando a cabo un procedimiento similar al mostrado en el Apartado 5.4.1 encontramos que $h(x, y)$ puede expresarse como

$$h(x, y) = h_1 h_2 \exp\left(-j\pi \frac{x^2 + y^2}{\lambda d_2}\right) P_1\left(\frac{x}{\lambda d_2}, \frac{y}{\lambda d_2}\right), \quad (6.15)$$

donde $P_1(v_x, v_y)$ es la transformada de Fourier de la *función abertura modificada*,

$$p_1(x, y) = p(x, y) \exp\left(-j\pi \epsilon \frac{x^2 + y^2}{\lambda}\right), \quad (6.16)$$

siendo

Error de desenfoque

$$\epsilon = \frac{1}{d_2} + \frac{1}{d_1} - \frac{1}{f} \quad (6.17)$$

un parámetro conocido como *error de desenfoque*. Notemos que el error de desenfoque se hace nulo cuando se verifica la ecuación de la lente delgada (3.36).

En el supuesto de que la imagen se extienda sobre una región pequeña del plano imagen, podemos despreciar el factor de fase que aparece en (6.15), dando lugar a que

$$h(x, y) \propto P_1\left(\frac{x}{\lambda d_2}, \frac{y}{\lambda d_2}\right). \quad (6.18)$$

Incluso admitiendo que el error de desenfoque sea nulo ($\epsilon = 0$), en cuyo caso

$$h(x, y) \propto P\left(\frac{x}{\lambda d_2}, \frac{y}{\lambda d_2}\right), \quad (6.19)$$

observamos que siempre aparecerá un efecto de difracción debido al tamaño finito de la lente por lo que la imagen de cada punto objeto no será otro punto sino más bien una "pequeña mancha". Tal como nos dice (6.19), la imagen de cada punto será únicamente un punto (es decir, $P(x/\lambda d_2, y/\lambda d_2) \rightarrow \delta(x, y)$) si la lente se extendiese hasta infinito o bien si se toma el límite cuando $\lambda \rightarrow 0$ (es decir, bajo la aproximación que nos llevaba a la Óptica Geométrica). Por este motivo, las lentes se consideran sistemas ideales formadores de imágenes en Óptica Geométrica.

Sistema óptico formado por la lente simple es un sistema lineal pero NO un sistema lineal invariante

Es interesante observar que el sistema óptico formado por la lente simple no es un sistema lineal invariante (aunque sí es lineal) puesto que un desplazamiento Δ en un punto del plano objeto no da lugar a un simple desplazamiento de su imagen en el plano imagen. Debido a la abertura de la lente, la imagen difractada de un punto situado en el centro del plano objeto se deforma, además de desplazarse, cuando el punto objeto se mueve. Dado que el sistema formador de imágenes compuesto por una lente simple no es lineal invariante, no tiene mucho sentido definir la función de transferencia para este sistema dado que no será útil.

6.4. Procesado de imágenes

Una de las ramas donde más aplicación ha tenido recientemente la óptica de Fourier es justamente en el procesado de imágenes. Usualmente los sistemas ópticos nos proporcionan imágenes inadecuadas para su utilización debido a que

no contienen la información que deseamos o bien porque han sufrido cierto tipo de degradación en su proceso de registro. El objetivo del *procesado o tratamiento de imágenes* es precisamente modificar *a posteriori*, según nuestra conveniencia, las imágenes mediante ciertas transformaciones. En general distinguiremos dos tipos de procesado de imágenes: procesado óptico coherente y procesado digital de imágenes.

6.4.1. Procesado óptico coherente

En este apartado estudiaremos el procesado de imágenes que se realiza mediante el **procesador óptico coherente** que se muestra en la Fig. 6.4 formado por dos lentes de distancia focal f .

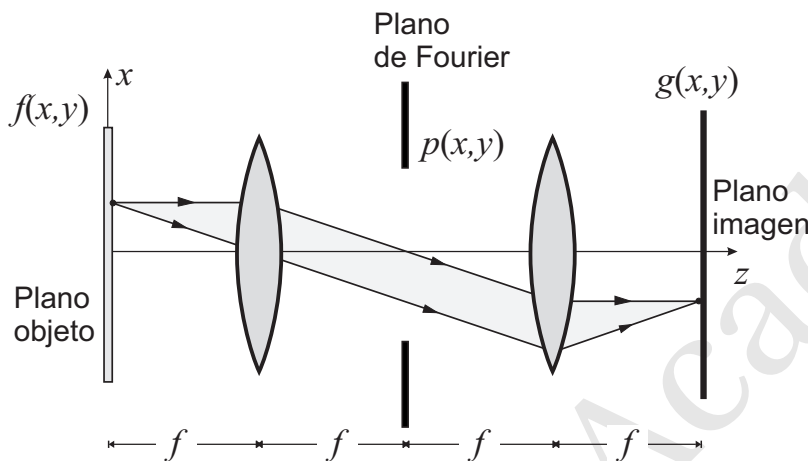


FIGURA 6.4: Sistema 4- f (o procesador óptico coherente) para producir una transformada de Fourier seguida de una transformada inversa, de modo que la imagen es una réplica del objeto.

Según vimos en el Apartado 5.4.3, si una distribución luminosa $f(x, y)$ está situada en el plano focal objeto de la lente, entonces la distribución luminosa, $g(x, y)$, el plano focal imagen de la lente reproduce la transformada de Fourier del objeto, es decir,

$$g(x, y) = F\left(\frac{x}{\lambda f}, \frac{y}{\lambda f}\right). \quad (5.84)$$

En este sentido, el plano focal imagen de la primera lente será denominado *plano de Fourier*. Dado que este plano de Fourier está situado justamente en el plano focal objeto de la segunda lente, el plano focal imagen de esta lente reproducirá de nuevo la transformada (que en este caso equivaldría a una transformada inversa de Fourier), dando así lugar a una reproducción perfecta del objeto en ausencia de máscara. Ahora bien, si situamos una transparencia adecuada en el plano de Fourier, podremos llevar a cabo ciertas “operaciones” sobre la transformada de Fourier, de modo que la transformada inversa de ésta dé lugar a una imagen “procesada” convenientemente (ver Fig. 6.5).

Es importante que consideremos la importancia de una iluminación *coherente* en este proceso. Para ello debemos recordar que la formación de la transformada de Fourier en el plano focal de la primera lente es equivalente a un proceso de difracción en esta lente, por lo que es necesario el uso de luz coherente para que el “patrón de difracción” sea estable en el tiempo. El uso de luz incoherente provocaría que los “términos de interferencia” se perdieran (ver Apartado 1.5.2 y

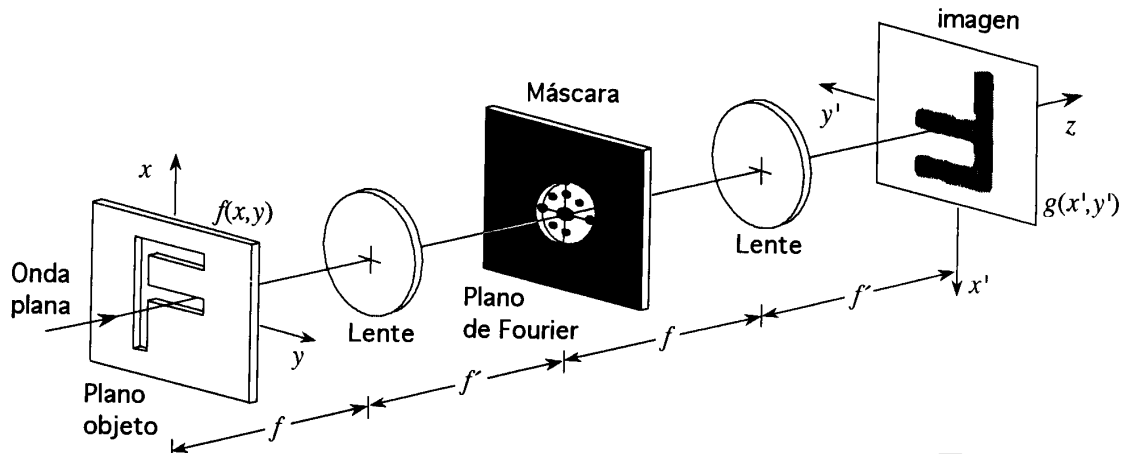


FIGURA 6.5: Sistema 4- f cuya transparencia es una máscara que nos permite seleccionar las frecuencias que se desean bloquear o transmitir.

recuérdese que la difracción era esencialmente un fenómeno de interferencia) y que consecuentemente no obtengamos el patrón de difracción deseado.

Según la expresión (5.84), la componente de Fourier de $f(x, y)$ asociada a la frecuencia espacial (v_x, v_y) corresponde a una localización espacial (x, y) en el plano de Fourier dada por

$$x = \lambda f v_x \quad (6.20a)$$

$$y = \lambda f v_y. \quad (6.20b)$$

En consecuencia, para implementar un filtro cuya función de transferencia sea $H(v_x, v_y)$, debemos colocar en el plano de Fourier una máscara cuya transmitancia en amplitud, $p(x, y)$, sea proporcional a la función de transferencia deseada. En consecuencia la función de transferencia asociada a una máscara de transmitancia $p(x, y)$ viene dada por

$$H(v_x, v_y) = p(\lambda f v_x, \lambda f v_y), \quad (6.21)$$

donde hemos ignorado el factor de fase $j \exp(-j2kf)$. En el sistema 4- f se cumplirá entonces que la relación entre las transformadas de Fourier de la imagen, $G(v_x, v_y)$, y del objeto, $F(v_x, v_y)$, estarán relacionadas mediante la siguiente expresión:

$$G(v_x, v_y) = H(v_x, v_y)F(v_x, v_y) \quad (6.22)$$

y, por tanto, la distribución luminosa a la salida se obtendrá como la transformada inversa de la expresión anterior

$$g(x, y) = \iint_{-\infty}^{\infty} F(v_x, v_y)H(v_x, v_y) \exp[-j2\pi(v_x x + v_y y)] dv_x dv_y. \quad (6.23)$$

La correspondiente función respuesta impulsiva, $h(x, y)$, del sistema vendrá dada por la transformada inversa de $H(v_x, v_y)$, es decir,

$$h(x, y) = \frac{1}{(\lambda f)^2} P\left(\frac{x}{\lambda f}, \frac{y}{\lambda f}\right), \quad (6.24)$$

donde $P(v_x, v_y)$ es la transformada de Fourier de la transmitancia, $p(x, y)$, de la máscara.

Función de transferencia del sistema 4- f con una máscara de transmitancia $p(x, y)$

Filtrado de frecuencias espaciales

Como un posible ejemplo de las operaciones de procesado óptico coherente analizaremos a continuación el filtrado de frecuencias espaciales. Según vemos en la Fig. 6.5, usando una transparencia (en este caso llamada *máscara*) que bloquee completamente el paso de ciertas frecuencias conseguiremos un filtrado selectivo de ciertas frecuencias, de modo que la imagen del objeto carecerá de dichas frecuencias “bloqueadas”.

Algunos de los filtros espaciales más frecuentemente usados son

- **Filtro pasa-baja**

Tal como muestra la Figura 6.6, este filtro tiene una función de transmisión

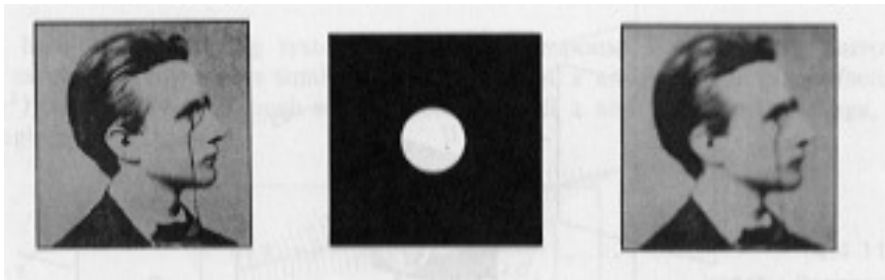


FIGURA 6.6: Filtro pasa-baja.

dada por

$$H(v_x, v_y) = \begin{cases} 1 & \text{si } v_x^2 + v_y^2 < v_c^2 \\ 0 & \text{en otro caso,} \end{cases} \quad (6.25)$$

por lo que bloquea las frecuencias más altas que v_c , que se conoce como *frecuencia de corte*. Dado que los detalles finos del objeto están relacionados con las frecuencias más altas, este filtro produce un suavizado de la imagen. El filtro se implementa mediante una máscara con una abertura circular de radio R , cuya relación con la frecuencia de corte viene dada por

$$R = \lambda f v_c. \quad (6.26)$$

Para una longitud de onda dada, la frecuencia de corte será $v_c = R/\lambda f$ y consecuentemente el tamaño del detalle más fino que discernirá este filtro será de $\lambda f/R$. Si, por ejemplo, $R = 2\text{cm}$, $\lambda = 1\mu\text{m}$ y $f = 200\text{cm}$, el detalle más fino que aparecería en la imagen sería de 0.1 mm .

- **Filtro pasa-alta**

La función de transmisión de este filtro es

$$H(v_x, v_y) = \begin{cases} 1 & \text{si } v_x^2 + v_y^2 > v_c^2 \\ 0 & \text{en otro caso,} \end{cases} \quad (6.27)$$

por lo que este filtro es el complementario del filtro anterior dado que bloquea las frecuencias bajas y deja pasar las altas. Tal como muestra la Figura 6.7, este filtro se construye con una máscara consistente en un círculo opaco central. Dado que la respuesta del filtro consistirá en potenciar las partes de la imagen relacionadas con detalles finos, este filtro será útil para realzar los bordes de las imágenes.

- **Filtro pasa-vertical**

La Figura 6.8, muestra que este filtro deja pasar las frecuencias verticales blo-

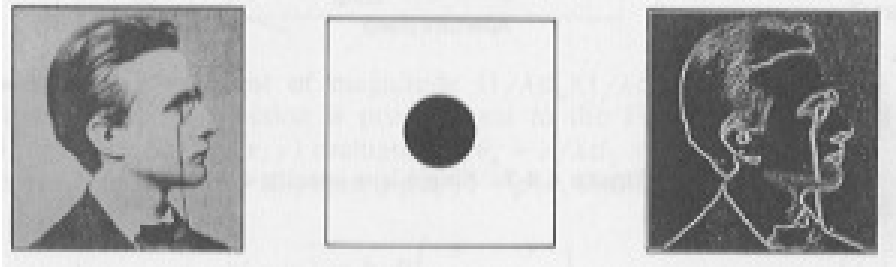


FIGURA 6.7: Filtro pasa-alta.

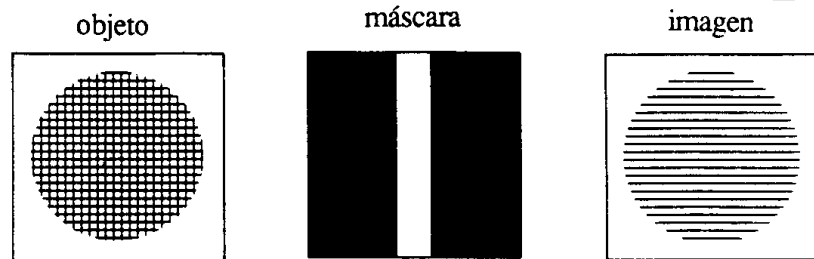


FIGURA 6.8: Filtro pasa-vertical.

queando las horizontales. Por ello, sólo las variaciones espaciales en la dirección vertical x son transmitidas. Si la anchura de la rendija vertical es D_y , la frecuencia horizontal más alta que pasa es la de valor

$$v_y|_{\max} = \frac{D_y/2}{\lambda f}. \quad (6.28)$$

■ Filtro pasa-horizontal

Al contrario que el anterior filtro, este filtro bloquea las frecuencias verticales y transmite las horizontales. Por tanto, como muestra la Figura 6.9, sólo



FIGURA 6.9: Filtro pasa-horizontal.

las variaciones horizontales (en la dirección y) aparecerán en la imagen, bloqueando las verticales de frecuencia mayor que

$$v_x|_{\max} = \frac{D_x/2}{\lambda f}. \quad (6.29)$$

En la práctica también se usan otro tipo de filtros, como los filtros en amplitud que nos permitirían, por ejemplo, aumentar el contraste de las imágenes; los filtros de fase usados, por ejemplo, en el microscopio de contraste de fase; los filtros complejos, usados en la restauración de imágenes, etc.

6.4.2. Procesado digital

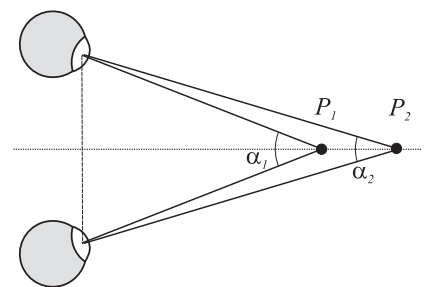
Actualmente un medio muy usual de registrar imágenes consiste en discretizar la imagen original respecto a sus coordenadas espaciales, de modo que se crea una matriz cuyos índices de fila y columna indican la posición del punto en la imagen y su valor nos proporciona el grado de intensidad de la luz (si tuviésemos imágenes en color necesitaríamos tres valores de intensidad, uno para cada tipo de color básico). Cada uno de los elementos de la matriz anterior se denominan elementos de imagen o **pixels**. En la práctica, el proceso anterior lo realiza un *digitalizador*, que en las cámaras digitales fotográficas y de video consiste usualmente en una retícula de $N \times M$ detectores semiconductores fotosensibles (comúnmente CCDs o *charge-coupled devices*) de dimensiones aproximadas 8.8×6.6 mm.

Si la imagen digitalizada anterior se almacena en forma binaria ¹, un ordenador podría llevar a cabo cualquier tipo de operación compleja sobre la imagen, tal como transformadas de Fourier, filtrados espaciales, aumento de contraste, etc. Aunque estas operaciones requieren un gran esfuerzo computacional, el procesado de imágenes por ordenador tiene la gran ventaja de que puede aplicarse indistintamente a imágenes formadas tanto con luz coherente como incoherente (ciertamente el valor binario asociado al pixel que esté procesando el ordenador no contiene este tipo de información). Dado que muchas de las imágenes de trabajo son obtenidas con fuentes de luz incoherente (luz solar, luz de focos, bombillas,...), el procesado digital de imágenes por ordenador presenta un campo de aplicación muy extenso. Además, el ordenador puede realizar ciertas operaciones sobre las imágenes que son muy difíciles de implementar en la práctica con un procesador óptico coherente (por ejemplo, diversas convoluciones bidimensionales que buscan efectos específicos). El uso del ordenador para tratar imágenes constituye un amplio y prometedor campo de estudio donde se combinan frecuentemente las técnicas de la óptica de Fourier con nuevos recursos.

6.5. Holografía

Aunque usualmente no repararemos en ello, debemos considerar el hecho de que los sistemas comunes de registro de imágenes (ojo, película fotográfica, matriz de CCDs, etc) son únicamente sensibles a la cantidad de energía que llega a ellos pero NO a cuánto camino ha recorrido esa energía; es decir, sólo son sensibles a la intensidad de la onda luminosa pero no a su fase. Este hecho podemos igualmente expresarlo diciendo que recogen la información de la intensidad de la onda pero no del frente de ondas. Debido a ello, cuando observamos este registro de la onda no obtenemos una imagen "real" del objeto original sino más bien una imagen bidimensional (2D) sin profundidad (la profundidad aparente es simplemente fruto de la perspectiva).

Si, por ejemplo, analizamos la supuesta visión 3D que nos proporcionan los ojos, debemos notar que el ojo percibe la profundidad (es decir, que un punto está más lejano que otro) comparando los ángulos que forman entre sí los ejes visuales de los dos ojos cuando estos son dirigidos al objeto. En este sentido, los puntos P_1 y P_2 se observan a distinta distancia si la diferencia entre los ángulos α_1 y α_2 es mayor que cierto límite (como promedio, en el ojo humano, $\alpha_1 - \alpha_2$ debe ser mayor que 30°). Esta posibilidad de visión estereoscópica o visión 3D sólo ocurre cuando lo que llega al ojo es directamente el frente de onda que procede del



¹ En el caso de que tengamos, por ejemplo, 780×480 detectores, una imagen en blanco y negro con $128 (= 2^7)$ niveles de gris requeriría 2.62 Mbits de memoria, $2620800 = 780 \cdot 480 \cdot 7$.

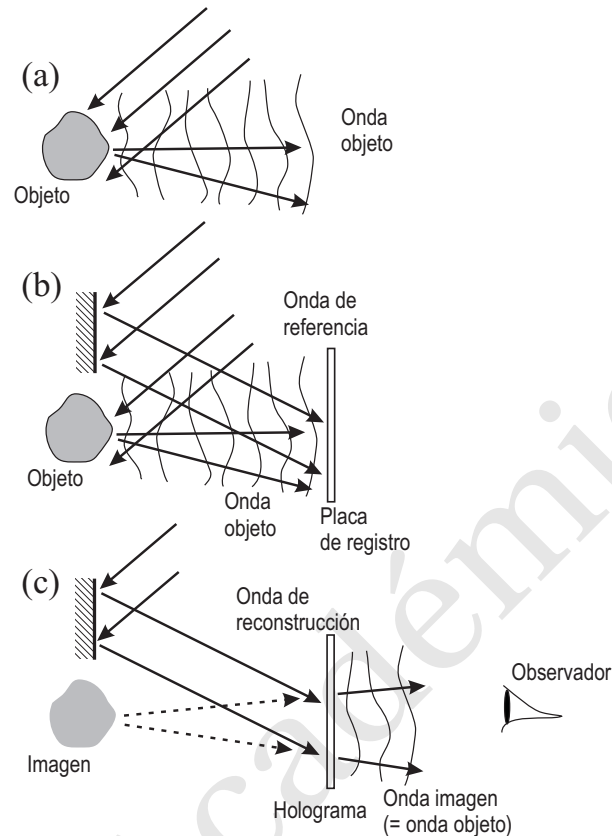


FIGURA 6.10: Esquema básico de la holografía: (a) proceso "natural" de visión estereoscópica, (b) registro de la onda objeto, (c) reconstrucción de la onda objeto.

objeto. Claramente, con la luz que procede de una fotografía no se puede llevar a cabo el anterior efecto estereoscópico.

La holografía surge como un posible método para superar los inconvenientes de los métodos tradicionales de registro luminoso. En consecuencia, la holografía pretende registrar tanto la intensidad como la fase de la onda luminosa; es decir, la propia onda objeto. Podría decirse que la holografía es un método para "congelar" la onda que procede del objeto y posteriormente "ponerla otra vez en marcha". Denominaremos **holograma** al soporte físico del registro de la onda objeto.

El proceso holográfico se muestra en la Figura 6.10. En la parte (a) vemos el proceso "natural" de visión estereoscópica que pretendemos reproducir. En la etapa de registro (b) se hace *interferir* la onda proveniente del objeto con una onda de referencia, almacenándose este patrón de interferencia en el holograma. En la etapa de reconstrucción (c), el holograma es iluminado por una onda idéntica a la onda de referencia, recuperando de esta manera la onda objeto original. Básicamente podemos decir que el registro del holograma tiene lugar mediante un proceso de interferencia mientras que la reconstrucción de la onda se basa en un proceso de difracción.

Desde un punto de vista práctico, la elaboración de un holograma debe tener en cuenta los dos siguientes puntos básicos:

1. Fuentes de luz coherente.

Dado que en el holograma debe registrarse un patrón de interferencia, es primordial que se satisfagan las condiciones para que la interferencia tenga lugar.

Esto requiere, como se discutió en el Apartado 1.5, que las ondas que interfieren provengan de *focos coherentes*. En el Tema 7 veremos que las características de alta coherencia de la luz láser hacen a ésta una candidata ideal para la Holografía.

2. Registros con alto poder de resolución.

El poder de resolución del material de registro debe ser muy elevado para poder recoger los detalles del patrón de interferencia que suelen ser muy finos (del orden de $1 \mu\text{m}$). Esto ha requerido el desarrollo de emulsiones fotográficas muy sensibles.

6.5.1. El código holográfico

Consideremos una onda plana oblicua de longitud de onda λ que se propaga formando un ángulo θ con el eje z y que en el plano $z = 0$ tiene la siguiente forma:

$$U_0(x, y) = \exp(-jk \text{sen}\theta x) \quad k = \frac{2\pi}{\lambda}. \quad (6.30)$$

Esta onda podría registrarse en una transparencia cuya transmitancia *compleja* fuese

$$t(x, y) = \exp(-jk \text{sen}\theta x). \quad (6.31)$$

Tal como se ha explicado anteriormente, al iluminar esta transparencia con una onda plana del tipo $\exp(-jkz)$, ésta se deflectaría un ángulo θ de modo que se habría conseguido reconstruir la onda original. (En el caso de una onda arbitraria, se podría seguir el mismo procedimiento para todas las ondas planas constitutivas y de este modo se tendría el registro adecuado).

Aunque el procedimiento anterior es perfectamente legítimo desde un punto de vista teórico, el problema práctico que presenta es que *no es posible realizar una transparencia que registre la información de la fase de la onda*. Tal como se ha señalado anteriormente, la transparencia sólo es sensible a la intensidad de la onda, por lo que, en el caso de la onda (6.30), dicha transparencia simplemente registraría intensidad unidad en toda su extensión. Puesto que la transparencia sólo registra intensidad y nosotros queremos también registrar información de la fase, una posible solución sería registrar un patrón de interferencia o **diagrama interferencial**. El patrón de interferencia es la distribución de intensidad en un plano de dos ondas que interfieren, dependiendo esta distribución de intensidad de la diferencia de fase de las dos ondas.

Codificación de la información

Consideremos un caso más general en el que la forma de la **onda objeto** en el plano $z = 0$ sea

$$U_0(x, y) = \sqrt{I_0} \exp[-j\phi_0(x, y)]. \quad (6.32)$$

Si esta onda interfiere en $z = 0$ con una **onda de referencia**,

$$U_r(x, y) = \sqrt{I_r} \exp[-j\phi_r(x, y)], \quad (6.33)$$

el diagrama interferencial que ambas producen tendrá una distribución de intensidad, $I(x, y)$, dada por

$$\begin{aligned} I(x, y) &= |U_r + U_0|^2 = |U_0|^2 + |U_r|^2 + U_0 U_r^* + U_0^* U_r \\ &= I_0 + I_r + 2\sqrt{I_0 I_r} \cos[\phi_0(x, y) - \phi_r(x, y)]. \end{aligned} \quad (6.34)$$

Si esta intensidad se registra fotográficamente, se podrá realizar una transparencia (holograma) cuya transmitancia $t(x, y)$ sea

$$t(x, y) \propto I_0 + I_r + 2\sqrt{I_0 I_r} \cos[\phi_0(x, y) - \phi_r(x, y)]. \quad (6.35)$$

Notemos que de esta manera hemos podido registrar en la transparencia información relacionada con la fase, $\phi_0(x, y)$, de la onda objeto.

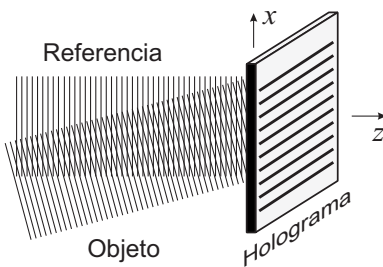
Decodificación de la información

Para reconstruir la onda objeto original necesitamos “decodificar” la información registrada en el holograma. Esto puede llevarse a cabo si el holograma es ahora iluminado con una onda igual a la onda de referencia, de modo que en $z = 0$ la onda resultante vendrá dada por

$$\begin{aligned} U(x, y) &= t(x, y)U_r(x, y) \\ &\propto U_r I_0 + U_r I_r + I_r U_0 + U_r^2 U_0^*. \end{aligned} \quad (6.36)$$

Notemos que el tercer miembro de la expresión anterior es la onda original, U_0 , multiplicada por la intensidad de la onda de referencia. En el caso de que esta intensidad no dependa de x e y , este término constituye justamente la onda reconstruida deseada. No obstante, al iluminar el holograma, este término irá siempre acompañado de los otros tres, por lo que para aislarlo debemos separarlo convenientemente.

EJEMPLO 6.2 Holograma de una onda plana oblicua.



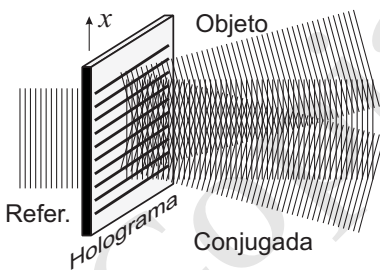
Si la onda de referencia, $U_r(x, y)$, es una onda plana que viaja a lo largo del eje z y la onda objeto, $U_0(x, y)$, es una onda plana oblicua que forma un ángulo θ con el eje z , entonces en el plano $z = 0$, estas ondas pueden escribirse como

$$\begin{aligned} U_r(x, y) &= \sqrt{I_r} \\ U_0(x, y) &= \sqrt{I_0} \exp(-jk \operatorname{sen} \theta x). \end{aligned}$$

Al iluminar el holograma formado por el diagrama interferencial de las dos ondas anteriores obtendremos, según (6.36), que

$$U(x, y) \propto I_r + I_0 + \sqrt{I_r I_0} \exp(-jk \operatorname{sen} \theta x) + \sqrt{I_r I_0} \exp(jk \operatorname{sen} \theta x).$$

Los dos primeros términos son constantes y corresponden a una onda que se propaga en la dirección z . El tercer término es proporcional a la onda original mientras que el cuarto término es una conjugada a la anterior que se propaga formando un ángulo $-\theta$ con el eje z .



El proceso descrito en el ejemplo anterior podría también interpretarse como que la transparencia de transmitancia

$$t(x, y) = I_0 + I_r + 2\sqrt{I_0 I_r} \cos(k \operatorname{sen} \theta x)$$

(esto es, el patrón de interferencia (6.35) correspondiente al ejemplo 6.2) es equivalente al elemento óptico “prisma de ángulo θ ”. En este sentido tenemos entonces que la anterior transparencia es un “holograma” del prisma de ángulo θ para la longitud de onda $2\pi/k$. Esto nos sugiere que diferentes hologramas podrían sustituir a diversos elementos ópticos que tradicionalmente se han realizado de otra manera (vidrios tallados convenientemente, superficies reflectoras, etc).

6.5.2. Holografía fuera de eje

El ejemplo anterior de la onda plana oblicua sugiere que una forma de separar las cuatro componentes que aparecen en (6.36) es asegurar que estas componentes viajen en direcciones distintas; esto es, que estén multiplexadas en distintas frecuencias espaciales (ver discusión al final del Apartado 5.2.3). Tal como muestra la Figura 6.11, este multiplexado en frecuencias espaciales puede conseguirse haciendo que las ondas objeto y referencia lleguen al holograma desde direcciones bien separadas.

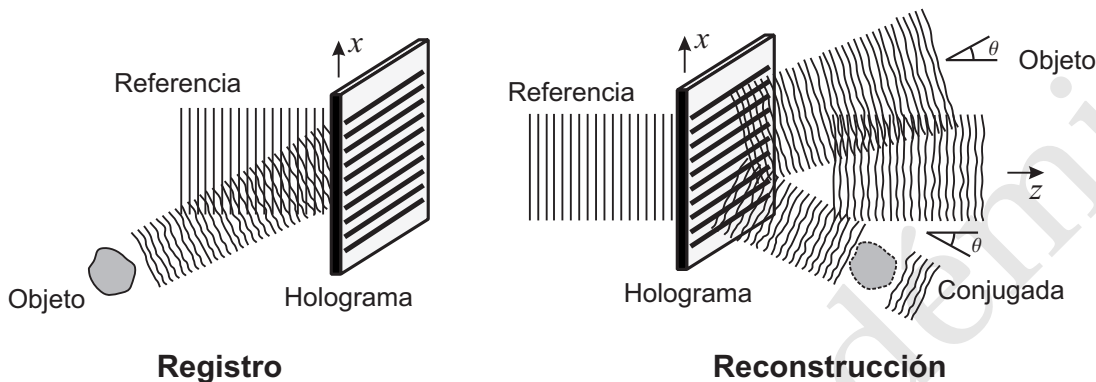


FIGURA 6.11: Esquema básico de la holografía fuera de eje.

Si la amplitud compleja de la onda objeto tiene la siguiente forma:

$$U_0(x, y) = f(x, y) \exp(-jk \operatorname{sen} \theta x), \quad (6.37)$$

es decir, la onda tiene una amplitud $f(x, y)$ modulada por un factor de fase igual al que le introduciría un prisma con ángulo de desviación θ . Supuesto que $f(x, y)$ varíe muy poco en una distancia $2\pi/k \operatorname{sen} \theta$, podemos entonces considerar que la onda objeto está básicamente dirigida según el ángulo θ .

Si la onda de referencia es una onda plana que se propaga en la dirección z , la onda resultante al iluminar el correspondiente holograma será

$$U(x, y) \propto I_r + |f(x, y)|^2 + \sqrt{I_r} f(x, y) \exp(-jk \operatorname{sen} \theta x) + \sqrt{I_r} f^*(x, y) \exp(+jk \operatorname{sen} \theta x). \quad (6.38)$$

En esta onda, al igual que ocurría en el ejemplo 6.2, el tercer término es una réplica de la onda objeto original propagándose según el ángulo θ . El cuarto término es la onda conjugada que se propaga según el ángulo $-\theta$. El primer y segundo término representan onda propagándose según el eje z .

EJEMPLO 6.3 Considerar un holograma registrado con una onda de referencia que es una onda esférica centrada en el punto $(0, 0, -d)$ y determinar el diagrama de interferencia almacenado así como las características de la onda reconstruida cuando la onda objeto es

- una onda plana que se propaga en una dirección que forma un ángulo θ , con el eje z ;
- una onda esférica centrada en el punto $(-d, 0, -d)$.

Aproximar las ondas esféricas por ondas parabólicas.

Considerando que la placa de registro está situada en el plano $z = 0$, la expresión de la onda de referencia en el plano de registro es

$$U_r(x, y) = \sqrt{I_r} e^{-j\phi_r(x, y)} = \frac{A_r}{\sqrt{d^2 + x^2 + y^2}} e^{-jk\sqrt{d^2 + x^2 + y^2}},$$

donde A_r es la intensidad de la fuente de la onda esférica. Bajo la aproximación de Fresnel, obtenemos la siguiente onda parabólica:

$$U_r(x, y) = \frac{A_r}{d} e^{-jk\frac{x^2 + y^2}{2d}}.$$

En general, si la onda objeto tiene la forma

$$U_0(x, y) = \sqrt{I_0} e^{-j\phi_0(x, y)},$$

el diagrama de interferencia que se almacena en la placa viene dado por la siguiente expresión:

$$I(x, y) = I_0 + I_r + 2\sqrt{I_0 I_r} \cos[\phi_0(x, y) - \phi_r(x, y)].$$

Una vez obtenido el holograma, al iluminarlo con la onda de referencia, se produce una onda definida por

$$U(x, y) \propto U_r I_0 + U_r I_r + I_r U_0 + U_r^2 U_0^*.$$

El tercer término de la expresión anterior corresponde a una onda proporcional a la onda objeto, mientras que el primero y el segundo conforman una onda que se propaga del mismo modo que U_r , y el cuarto está asociado a una onda proporcional a la onda objeto que se propaga con ángulo complementario al de ésta.

- En el primer caso, la expresión de la onda objeto es

$$U_0 = \sqrt{I_0} e^{-jk \text{sen} \theta x},$$

por lo que el diagrama de interferencia, suponiendo que se utiliza una fuente de luz monocromática y coherente, es

$$\begin{aligned} I(x, y) &= I_0 + I_r + 2\sqrt{I_0 I_r} \cos\left(k \text{sen} \theta x - k \frac{x^2 + y^2}{2d}\right) \\ &= I_0 + \frac{A_r^2}{d^2} + 2\sqrt{I_0 \frac{A_r^2}{d^2}} \cos\left[k\left(\text{sen} \theta x - \frac{x^2 + y^2}{2d}\right)\right]. \end{aligned}$$

Al iluminar el holograma con la onda de referencia, se obtendrá como resultado la siguiente composición de ondas:

$$\begin{aligned} U(x, y) &\propto U_r I_0 + U_r I_r + I_r U_0 + U_r^2 U_0^* \\ &= \frac{A_r}{d} e^{-jk\frac{x^2 + y^2}{2d}} (I_0 + I_r) + I_r \sqrt{I_0} e^{-jk \text{sen} \theta x} + \frac{A_r^2}{d^2} e^{-2jk\frac{x^2 + y^2}{2d}} \sqrt{I_0} e^{jk \text{sen} \theta x}. \end{aligned}$$

- Si la onda objeto es una onda esférica centrada en $(-d, 0, -d)$, en el plano $z = 0$, tenemos

$$U_0(x, y) = \frac{A_0}{\sqrt{(x+d)^2 + d^2 + y^2}} e^{-jk\sqrt{(x+d)^2 + d^2 + y^2}}.$$

Aproximando esta onda por una onda parabólica, obtenemos

$$U_0(x, y) = \frac{A_0}{d} e^{-jk\frac{(x+d)^2 + y^2}{2d}}.$$

El diagrama de interferencia registrado es, por tanto,

$$I(x, y) = I_0 + I_r + 2\sqrt{I_0 I_r} \cos\left(k \frac{(x+d)^2 + y^2}{2d} - k \frac{x^2 + y^2}{2d}\right) \\ = \frac{A_r^2}{d^2} + \frac{A_0^2}{d^2} + 2 \frac{A_0 A_r}{d^2} \cos(kx + \varphi),$$

siendo $\varphi = \pi d/\lambda$.

Si el diagrama anterior se ilumina con la onda de referencia, se obtiene como resultado la composición de cuatro términos descrita por la siguiente expresión:

$$\frac{A_r}{2d} e^{-jk \frac{x^2 + y^2}{2d}} \left(\frac{A_0^2}{4d^2} + \frac{A_r^2}{4d^2} \right) + \frac{A_r^2}{\sqrt{4d^2}} \frac{A_0}{2d} e^{-jk \frac{(x+d)^2 + y^2}{2d}} + \frac{A_r^2}{d^2} \frac{A_0}{2d} e^{j(kx + \varphi)},$$

donde el tercer sumando es proporcional a la onda objeto original.

6.5.3. Registro holográfico

Dado que el holograma debe registrar *diferencias de fase*,

$$\arg(U_r) - \arg(U_0),$$

es condición indispensable que la luz usada tanto en el proceso de registro como en el de reconstrucción sea una *fente monocromática con fluctuaciones de fase mínimas*. Tal como se discutió en el Apartado 1.5, la existencia de *interferencias* (fruto de la diferencias de fase) estaba ligada a la existencia de fuentes de luz *coherentes* para conseguir así que la diferencia de fase sea “permanente”. En este sentido, una fuente de luz usual (como la proveniente de una bombilla) no es válida dado que no es coherente. Tal como discutiremos en el Tema 7, una de las características fundamentales de la luz **láser** es precisamente su coherencia, lo que hará que esta fuente de luz sea un elemento básico en el proceso holográfico.

6.5.4. Aplicaciones de la holografía

Además de su uso como procedimiento para registrar y visualizar imágenes 3-D, la holografía tiene un amplísimo campo de aplicaciones en la óptica actual. En concreto, una de las aplicaciones más importantes de la Holografía es el diseño y fabricación de **elementos ópticos holográficos**. Estos elementos ópticos son algunos de los componentes típicos de un sistema óptico (lentes, espejos, prismas, redes de difracción,...) que tradicionalmente se realizaban usando como material base el vidrio y que, aplicando técnicas holográficas, se pueden también implementar sobre cualquier soporte capaz de registrar la estructura de franjas de interferencia deseada (por ejemplo, un acetato, una placa fotográfica, etc). Los elementos ópticos convencionales basaban su funcionamiento en las leyes de la reflexión y/o refracción mientras que el comportamiento de los elementos ópticos holográficos esta gobernado básicamente por fenómenos de difracción.

En general en los elementos ópticos holográficos, el objeto del que se hace un holograma es un punto. Dentro de la aproximación paraxial, este holograma funcionará igualmente para objetos y sistemas ópticos que verifiquen la anterior aproximación (al igual que ocurre con los sistemas ópticos convencionales). Algunos de los elemento ópticos holográficos más relevantes son:

■ Lente holográfica

Al igual que ocurre en una lente convencional y como muestra la Fig. 6.12, una lente holográfica transformará una onda incidente (en nuestro caso, lo que hemos llamado onda de reconstrucción) en una onda imagen (o equivalentemente, la onda objeto reconstruida)

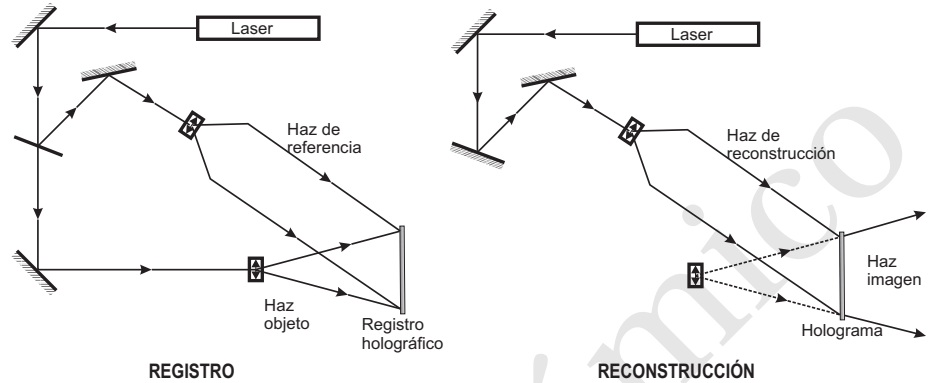


FIGURA 6.12: Esquema básico del proceso de registro y reconstrucción en una lente holográfica.

Operando de forma análoga a como hicimos en los apartados anteriores obtendríamos la siguiente ecuación para una lente holográfica

$$\frac{1}{z_c} + \frac{1}{z_i} = \frac{1}{f}. \quad (6.39)$$

Esta ecuación es formalmente idéntica a la ecuación de una lente (3.36), donde ahora debemos identificar z_c con la posición del origen de la onda esférica de reconstrucción, z_i la posición del punto imagen y f la distancia focal de la lente, que viene dada por

$$\frac{1}{f} = \frac{\lambda_c}{\lambda_r} \left(\frac{1}{z_0} - \frac{1}{z_r} \right), \quad (6.40)$$

donde λ_c y λ_r son las longitudes de onda de las ondas de reconstrucción y de referencia respectivamente y z_0 y z_r las posiciones del origen de las ondas objeto y de referencia. Podemos observar la fuerte dependencia que presenta la distancia focal de la lente con respecto a la longitud de onda (al contrario de lo que ocurre en las lentes convencionales). Algunas otras características importantes de las lentes holográficas son

- Muy fáciles de construir. El registro en un acetato es más fácil que el pulido de un vidrio.
- Muy poco pesadas. Esto puede ser muy importante a la hora de construir lentes de gran diámetro o cuando las lentes deben incorporarse en sistemas móviles diminutos.
- Se pueden realizar varios elementos en la misma placa.

■ Espejos holográficos

Usando una técnica similar a la de registro y uso de lentes holográficas podemos igualmente construir espejos holográficos.

■ Acopladores de fibras ópticas

Mediante un sistema adecuado de lentes (o espejos) holográficas podemos acoplar el haz entre dos fibras ópticas que no están alineadas.

- **Scanners holográficos**

Según vimos en el Ejemplo 5.1, el uso de una transparencia con cierta transmitancia compleja podía ser usado como un scanner. Usando técnicas holográficas se puede conseguir este mismo efecto.

- **Interconectores en óptica integrada**

El uso de lentes y otros elementos ópticos holográficos nos permitirían guiar los haces luminosos dentro de los sustratos soporte de los dispositivos realizados mediante óptica integrada (de manera similar a los interconectores usados en electrónica integrada).

- **Memorias holográficas**

La gran capacidad de almacenamiento de un registro holográfico (es una red bi o tridimensional de “puntos” de información) puede ser usado para realizar memorias ópticas para un ordenador.

- **Visores holográficos**

Mediante el uso de espejos holográficos podemos situar una imagen sobre el visor de un determinado casco de modo que simultáneamente se vea el exterior y los datos proporcionados por la imagen holográfica.

6.6. Problemas propuestos

6.1: Una rendija de anchura $187 \mu\text{m}$ se ilumina con un haz paralelo de luz monocromática de longitud de onda 694 nm . Si se sitúa una pantalla de observación a una distancia de 105 cm de la rendija, calcular: **a)** la anchura de la franja central brillante y de los máximos secundarios del diagrama de difracción; **b)** la distancia entre el primer y octavo mínimos.

Sol: a): 7.79 mm , 3.89 mm ; b): 27.23 mm .

6.2: Una onda plana monocromática de longitud de onda 694 nm incide perpendicularmente sobre una pantalla opaca que tiene una abertura rectangular. Se observa el diagrama de difracción en el plano focal de una lente convergente de 185 cm de focal situada directamente detrás de la abertura, obteniéndose que las dimensiones en las direcciones de los ejes x e y del rectángulo formado por las líneas oscuras que rodean al máximo central son 9.5 mm y 3.8 mm , respectivamente. Calcular los lados de la abertura rectangular.

Sol: $D_x = 0.15 \text{ mm}$, $D_y = 0.36 \text{ mm}$.

6.3: Rayos paralelos de luz verde de mercurio, cuya longitud de onda es 560 nm , pasan a través de una rendija de 0.4 mm de anchura, situada delante de una lente convergente de 40 cm de focal imagen, ¿Cuál es la distancia entre el máximo central y el primer mínimo sobre una pantalla colocada en el plano focal de la lente?

Sol: 0.56 mm .

6.4: Un holograma se registra con luz proveniente de un láser de He-Ne ($\lambda_0 = 633 \text{ nm}$) mediante la interferencia de dos ondas planas de modo que la normal a la placa holográfica corresponde a la bisectriz del ángulo formado por las dos ondas. Bajo estas condiciones obtener el mínimo ángulo que deben formar las ondas objeto y referencia para que la frecuencia espacial en el holograma no supere las 200 líneas/mm .

Sol: $\theta \leq 0.0633 \text{ rad}$.

6.5: Una lente holográfica se registra con luz ultravioleta de longitud de onda 337 nm y se reconstruye con luz de un láser de He-Ne de longitud de onda 633 nm . Si las ondas de referencia y de reconstrucción son dos ondas planas que inciden de la misma manera sobre la lente, obtener el aumento producido por la lente sobre el objeto original.

Sol: $m = 1.88z_i / z_0$.

Copia Uso Académico

Tema 7

Fuentes de luz: led y láser

7.1. Introducción

Hasta ahora hemos estado estudiando las ondas luminosas y algunas de sus posibles aplicaciones. En todo el estudio anterior ha quedado de manifiesto la importancia del carácter monocromático y coherente de la radiación luminosa. No obstante, nada ha sido dicho sobre cuáles son las fuentes de las ondas monocromáticas y coherentes. Este tema estará dedicado a este estudio y veremos que las fuentes más usadas en las tecnologías de la Informática son los LED y los diodos láser.

El *láser*, junto con el satélite, los reactores nucleares, el computador y los circuitos integrados forman parte de los símbolos de lo que se ha dado en llamar “alta tecnología”. La historia del láser nace con los primeros estudios de Einstein (1917) acerca de la *radiación estimulada* dentro del marco de la interacción radiación-materia. Estas primeras consideraciones teóricas no fueron suficientemente explotadas hasta la realización por C.H. Townesen 1957 del *máser*. Este dispositivo inicial, cuyo nombre proviene de los acrónimos de *Microwave Amplification of Stimulated Emission Radiation*, consistía en un haz de moléculas de amoníaco que tras ser excitado por un campo externo se dividía en dos subhaces: uno consistente de moléculas en el estado base y otro cuyas moléculas estaban todas excitadas. El subhaz excitado se hacía entrar en una cavidad resonante cuyo diseño permitía absorber la energía en exceso de las moléculas excitadas en forma de radiación electromagnética de 24 GHz (microondas). Esta radiación tenía dos importantes propiedades: a) presentaba una anchura de banda muy estrecha y b) era un haz coherente. Este montaje fue el primer dispositivo artificial que produjo una radiación estimulada y dio paso así a una amplia investigación para conseguir radiaciones basadas en emisiones estimuladas con distintas longitudes de onda. En este esfuerzo, fue T. Maiman en 1960 quién primero presentó un dispositivo que emitía radiación por emisión estimulada en el rango de frecuencias ópticas, este dispositivo se denominó *láser* siguiendo el acrónimo de *Light Amplification of Stimulated Emission Radiation*. A partir de este primer láser, numerosos laboratorios del mundo iniciaron una intensa investigación que culminó en el desarrollo de muchos tipos distintos de láser: láser de gas, láser de estado sólido, láser de inyección, etc... y cubriendo un amplia gama de longitudes de onda, desde el infrarrojo hasta el ultravioleta (e incluso rayos X). En los primeros años de la década de los sesenta, el desarrollo del láser no fue acompañado de unas aplicaciones tecnológicas inmediatas y así jocosamente se denominada al láser “una solución en busca de un problema”. Actualmente las aplicaciones del láser son inmensas

y en continua expansión, formando ya parte de la “tecnología de uso diario” en lectores de *compact disc*, impresoras láser, memorias para ordenadores magneto-ópticas, lectores de códigos de barras, etc....

Entre las características comunes de todos los láseres cabe señalar

- **Longitud de onda:** la radiación emitida suele presentar un ancho de banda muy estrecho.
- **Potencia de salida:** existe un amplio margen de valores respecto a la potencia de salida. Aunque la mayoría de los láseres comerciales son de baja potencia (menores que una décima de vatio), existen algunos láseres que pueden emitir potencias mayores que el millar de vatios en régimen continuo.
- **Salida continua o en pulsos.** La radiación emitida puede ser en pulsos de diversa anchura (tan pequeña como 10^{-15} s) o en régimen continuo.
- **Divergencia y direccionalidad del haz.** Una de las principales características de los láseres es que su emisión presenta mucha directividad de modo que a veces aparece como un “rayo”.
- **Coherencia.** Debido a la naturaleza de la emisión estimulada en la que se basa el láser, la radiación presenta comúnmente altos grados de coherencia.
- **Eficiencia:** La mayoría de los láseres presentan muy baja eficiencia ($< 0.1\%$), aunque algunos láseres (láseres de inyección) pueden llegar al 30%.

7.2. Interacción radiación-materia

Como ya se ha señalado, las ideas originales sobre el *láser* parten del estudio de la interacción radiación materia. Por consiguiente, se analizarán brevemente las características más importantes de esta interacción. El análisis de esta interacción se realizará dentro del marco de la física cuántica y así consideraremos que la radiación electromagnética (luz) está constituida por cuantos de energía (y momento) localizados denominados **fotones** y que juegan el papel de “partículas”. Dado que la radiación estará formada por un gran número de fotones, usaremos métodos estadísticos para determinar sus propiedades más destacadas. En este sentido, una variable que jugará un papel destacado será $n(E)$, esto es, el número de partículas por unidad de volumen cuyas energías tienen valores comprendidos entre E y $E + dE$. Esta variable puede escribirse en términos de la función de distribución, $f(E)$, y de la función densidad de estados, $g(E)$, como

$$n(E) = g(E)f(E).$$

La función de distribución de **partículas clásicas** se determina con base en la distinguibilidad, la no restricción del número de partículas en cualquier estado y el principio de balance detallado, dando como resultado la conocida distribución de Boltzmann:

$$f_{\text{Boltz}}(E) = A \exp\left(-\frac{E}{kT}\right), \quad (7.1)$$

donde k es la constante de Boltzmann ($k = 1.381 \times 10^{-23}$ J/K) y T la temperatura absoluta (es decir, la temperatura expresada en grados Kelvin).

Una de las características más destacadas de las **partículas cuánticas** (fotones, electrones, protones, neutrinos,...) es que presentan un comportamiento estadístico muy distinto al predecible para una sistema de partículas clásicas. Este

comportamiento peculiar proviene del fenómeno de **indistinguibilidad** intrínseca de las partículas cuánticas, que a su vez descansa en la naturaleza dual onda-corpúsculo que presentan estos entes. Además de esta indistinguibilidad cuántica, aparece una propiedad microscópica muy importante, determinada por el **principio de exclusión de Pauli**, que afecta de forma esencial al comportamiento estadístico de ciertas partículas cuánticas. Este principio restringe el número de partículas que pueden estar en el mismo estado cuántico. Las partículas que siguen dicho principio de exclusión se denominan **fermiones** y aquellas que no lo siguen **bosones**. Dadas las peculiaridades distintas de cada tipo de partículas, cada una de ellas vendrá caracterizada por una función de distribución diferente, en concreto,

$$f(E) = \begin{cases} \text{[función de Bose]} \frac{1}{A \exp\left(\frac{E}{kT}\right) - 1} & \text{para bosones} \\ \text{[función de Fermi]} \frac{1}{A \exp\left(\frac{E}{kT}\right) + 1} & \text{para fermiones} \end{cases} \quad (7.2)$$

donde el parámetro A vendrá determinado por el número total de partículas.

Se encuentra que los fotones (cuantos de radiación electromagnética de energía $E = h\nu$, siendo $h = 6.626 \times 10^{-34}$ J s la constante de Planck y ν la frecuencia de la radiación) son bosones y que por tanto su comportamiento estadístico vendrá caracterizado por la función de distribución de Bose. Este comportamiento puede interpretarse para el caso de los fotones diciendo que "la probabilidad de que un átomo emita un fotón en un estado final determinado, aumenta en un factor $(n + 1)$ si ya hay n fotones en el mismo estado". A continuación, analizaremos las consecuencias físicas de este hecho con respecto a la interacción de la luz con los átomos. Para ello, supongamos una situación en la cual tenemos n fotones idénticos –igual frecuencia, dirección y polarización– contenidos dentro de una caja, y también un átomo de la caja que puede emitir otro fotón en el mismo estado. La probabilidad de que emita un fotón es entonces

$$(n + 1)|a|^2, \quad (7.3)$$

siendo $|a|$ la amplitud de la probabilidad cuando no existen otros fotones. Por otra parte, la probabilidad de que el átomo absorba un fotón será

$$n|a|^2, \quad (7.4)$$

dado que la probabilidad de esta absorción es la misma que la que aparece cuando al existir $(n - 1)$ fotones, el átomo emite otro más en la misma situación. La ecuación (7.4) dice que la probabilidad de que un átomo absorba un fotón y haga una transición a un estado de energía más alto es proporcional a la intensidad de la luz presente. Por otra parte, la expresión (7.3) sugiere (tal como lo consideró Einstein en 1917) que la transición energética atómica desde un estado excitado a otro más bajo con la emisión de un fotón consta de dos partes. Una consiste en la probabilidad de que el átomo realice una **transición espontánea** de valor $|a|^2$ y la segunda en la probabilidad de que realice una **transición inducida o estimulada** de valor $n|a|^2$ que es proporcional a la intensidad de la luz (esto es, al número de fotones presentes). Nótese que una de las características más destacadas de la emisión estimulada es el carácter **coherente** de ésta. Esta propiedad proviene del carácter bosónico de los fotones emitidos en presencia de otros fotones; recuérdese que el fotón será emitido muy probablemente en el mismo estado que los ya presentes. Se encuentra además que los coeficientes de absorción y de emisión inducida son los mismos y están relacionados con la probabilidad de emisión espontánea. Podemos, por tanto, representar gráficamente la interacción elemental de un fotón con el átomo según se representa en la Fig. 7.1.

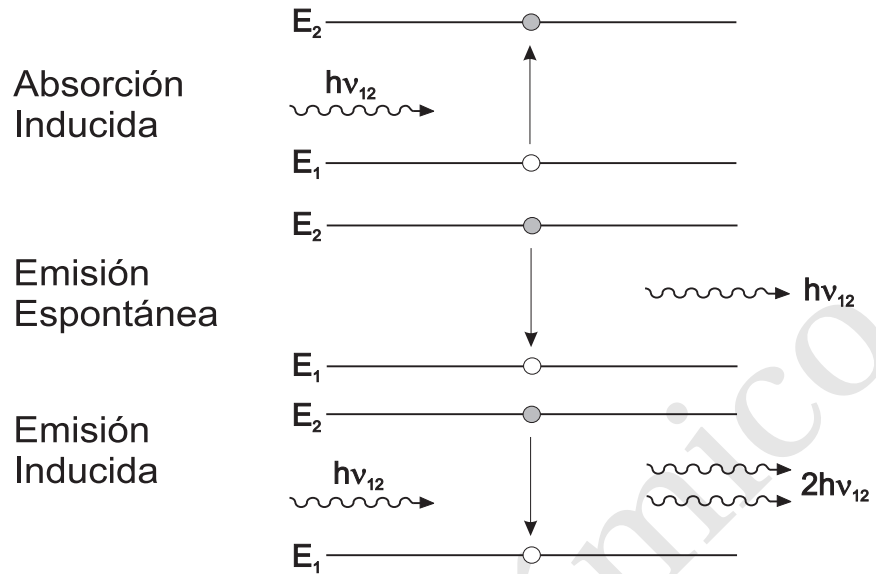


FIGURA 7.1: Interacción radiación-materia

Las anteriores leyes para fotones pueden aplicarse al cálculo de la densidad de energía electromagnética, $u(\nu)$, en el estudio de la radiación del cuerpo negro en equilibrio térmico. Denominaremos $R_{1 \rightarrow 2}$ a la probabilidad por unidad de tiempo para una transición del estado 1 al estado 2 en el átomo y se escribirá como

$$R_{1 \rightarrow 2} = B_{12} u(\nu), \quad (7.5)$$

siendo B_{12} es la probabilidad de absorción por unidad de tiempo y por unidad de intensidad de radiación. Análogamente $R_{2 \rightarrow 1}$ será la probabilidad por unidad de tiempo de que un átomo en el estado 2 realice una transición al estado 1. Dado que esta transición puede considerarse como fruto de dos procesos, escribiremos

$$R_{2 \rightarrow 1} = A_{21} + B_{21} u(\nu), \quad (7.6)$$

donde A_{21} es la probabilidad por unidad de tiempo de transición espontánea y B_{21} la probabilidad de emisión estimulada por unidad de tiempo y por unidad de intensidad de radiación.

Si hay N_2 átomos en el nivel E_2 y N_1 en el nivel E_1 en equilibrio térmico a una temperatura T con el campo de radiación de densidad de energía $u(\nu)$, entonces la razón total de absorción para el sistema, $\mathcal{R}_{1 \rightarrow 2}$ será

$$\mathcal{R}_{1 \rightarrow 2} = N_1 R_{1 \rightarrow 2} \quad (7.7)$$

y la razón de emisión total, $\mathcal{R}_{2 \rightarrow 1}$,

$$\mathcal{R}_{2 \rightarrow 1} = N_2 R_{2 \rightarrow 1} \quad (7.8)$$

deben ser iguales, esto es,

$$N_1 R_{1 \rightarrow 2} = N_2 R_{2 \rightarrow 1}, \quad (7.9)$$

que puede escribirse como

$$N_1 B_{12} u(\nu) = N_2 [A_{21} + B_{21} u(\nu)] \quad (7.10)$$

y resolviendo para la densidad de energía, se obtiene

$$u(\nu) = \frac{A_{21}/B_{12}}{(N_1/N_2)(B_{12}/B_{21}) - 1}. \quad (7.11)$$

Si los átomos están en equilibrio térmico y se les considera como partículas clásicas¹, se tiene

$$\frac{N_1}{N_2} = \exp\left(\frac{E_2 - E_1}{kT}\right) = \exp\left(\frac{h\nu}{kT}\right) \quad (7.12)$$

y por tanto

$$u(\nu) = \frac{A_{21}/B_{21}}{(B_{12}/B_{21}) \exp\left(\frac{h\nu}{kT}\right) - 1}, \quad (7.13)$$

expresión que tiene la misma forma que la ley de Planck del cuerpo negro para la densidad de energía de la radiación en equilibrio con la materia, si se hace

$$\frac{A_{21}}{B_{21}} = \frac{8\pi\nu^3}{c^3} \quad (7.14)$$

y

$$\frac{B_{12}}{B_{21}} = 1. \quad (7.15)$$

La ecuación (7.15) manifiesta la igualdad entre los coeficientes de emisión estimulada y absorción estimulada tal como ya se expuso en la discusión de la sección anterior. Por otra parte, dada la dependencia con ν^3 que presenta la expresión (7.14), esto implica que a medida que aumenta la diferencia energética entre los dos estados implicados, más probable será la emisión espontánea con respecto a la emisión estimulada. También puede obtenerse de (7.10), teniendo en cuenta (7.12), el cociente entre la probabilidad de emisión espontánea y la probabilidad de emisión estimulada, cuando la materia está en equilibrio térmico con la radiación:

$$\frac{\text{probabilidad de emisión espontánea}}{\text{probabilidad de emisión estimulada}} = \frac{A_{21}}{B_{21}u(\nu)} = \exp\left(\frac{h\nu}{kT}\right) - 1. \quad (7.16)$$

En consecuencia si $h\nu \gg kT$ (esto es, en transiciones electrónicas en átomos y moléculas, y radiativas en núcleos), la radiación emitida es predominantemente espontánea, pudiendo ignorarse la fracción de emisión estimulada. No obstante, si $h\nu \ll kT$ (por ejemplo en ciertas transiciones entre estados vibracionales o rotacionales en átomos y moléculas que corresponden a la región de microondas del espectro) la emisión estimulada (coherente) puede ser importante.

Tal como se ha estado discutiendo, cuando la materia y la radiación están en equilibrio térmico, no existe absorción o emisión neta dado que el número total por unidad de tiempo de transiciones por emisión y por absorción es el mismo. En el caso general de interacción de la radiación con la materia sin que haya necesariamente equilibrio térmico, se tiene que

$$\frac{\text{tasa de emisión}}{\text{tasa de absorción}} = \frac{[A_{21} + B_{21}u(\nu)]N_2}{B_{12}u(\nu)N_1} = \left(1 + \frac{A_{21}}{B_{21}u(\nu)}\right) \frac{N_2}{N_1}. \quad (7.17)$$

Si estamos en situaciones en las que la diferencia energética $E_2 - E_1 \ll kT$, de la expresión (7.16), se tiene que el segundo término del paréntesis es despreciable frente a la unidad y por tanto se puede escribir

$$\frac{\text{tasa de emisión}}{\text{tasa de absorción}} \approx \frac{N_2}{N_1}. \quad (7.18)$$

Si el sistema está en equilibrio térmico, donde el factor de Boltzmann es aplicable, N_2 es menor que N_1 y la tasa de emisión también es menor que la de absorción

¹ Esto es una suposición razonable en la mayoría de los casos, ya que debido al tamaño típico de las distancias interatómicas, las funciones de onda de cada uno de los átomos individuales no solapan apreciablemente entre sí

(ver Fig. 7.2a). Sin embargo, en condiciones fuera de equilibrio cualquier relación es posible. Así, si por algún procedimiento se invierte la población relativa de los niveles excitado y fundamental, de modo que $N_2 > N_1$, la tasa de emisión será mayor que la de absorción. Esto quiere decir que la radiación aplicada de frecuencia $\nu = (E_2 - E_1)/h$ será **amplificada** en intensidad por el proceso de interacción, ya que emerge más radiación que la que entra (ver Fig. 7.2b). Este proceso reducirá

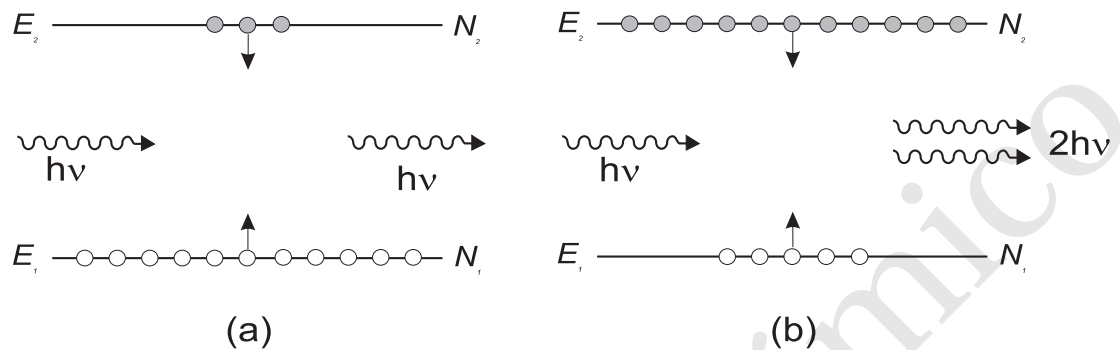


FIGURA 7.2: Principio básico del funcionamiento del láser y del máser

la población del nivel más alto hasta que se reestablezca el equilibrio térmico. Por lo tanto para sostener el proceso, es necesario utilizar algún método que mantenga la **inversión de población** de los estados. El mecanismo de excitación para producir la inversión de población, conocido frecuentemente como **bombeo**, no puede ser térmico, ya que en el equilibrio térmico siempre existirán más átomos en niveles bajos de energía que en los altos. Hay, por tanto, que excitar de forma selectiva, siendo las dos formas más comunes de excitación: luz y electricidad.

A menudo, los estados excitados tienen tiempos de vida media muy cortos (del orden de 10^{-9} s) y se desprenden de su energía mediante emisiones espontáneas². Debido al principio de incertidumbre, $\Delta E \delta t \sim h$, un tiempo de vida medio muy corto implica una gran anchura energética. Debe tenerse en cuenta que una situación estacionaria de inversión de población implica que el nivel láser superior debe poblarse con más rapidez que la razón de decaimiento de la población. Por tanto, para mantener la inversión de población, se necesitarían estados excitados con tiempos de vida media largos (y a su vez poca anchura energética). Estos estados se denominan **metaestables** y presentan un tiempo de vida media del orden de 10^{-3} s (este tiempo de vida media "corto" se debe usualmente a que presentan poca probabilidad de emisión en la forma dominante de radiación de dipolo eléctrico y alta probabilidad en formas menos dominantes de dipolo magnético o cuadripolo eléctrico). Debe tenerse en cuenta que una situación estacionaria de inversión de población implica que el nivel láser superior debe poblarse con más rapidez que la razón de decaimiento de la población.

Los dos primeros esquemas que se usaron para producir láser, presentaban la siguiente disposición:

■ Láser de 3 niveles:

Según se muestra en la Fig. 7.3, el nivel fundamental de energía es a su vez el

² Cuando el átomo es una mezcla de los dos estados cuánticos, su distribución de carga oscila precisamente a la frecuencia del fotón emitido en la transición entre los dos niveles. El cálculo cuántico del ritmo de emisión de energía del dipolo eléctrico oscilante determina igualmente la razón de transición atómica, esto es, la probabilidad por segundo de que el átomo emita espontáneamente. Esta razón de transición atómica viene determinada por los *elementos de matriz del momento eléctrico dipolar tomado entre los estados inicial y final* y, por consiguiente, los tiempos de vida media vienen igualmente caracterizados por estos elementos de matriz.

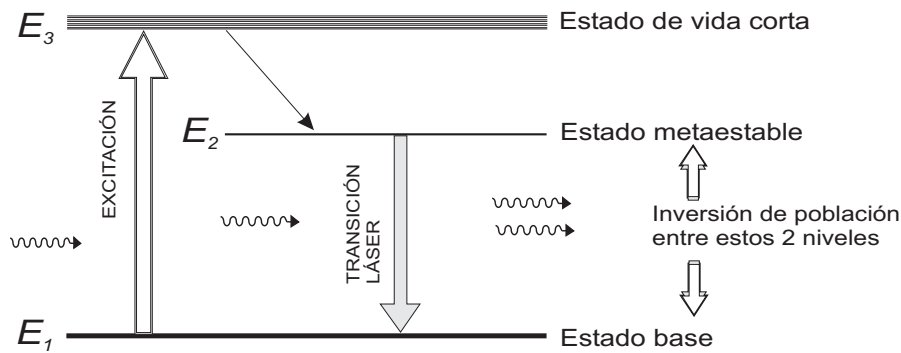


FIGURA 7.3: Niveles de energía relevantes en un láser de 3 niveles

nivel láser inferior, el nivel energético superior es un estado excitado de vida media muy corta y el nivel energético inferior es un estado metaestable que actúa como nivel láser superior. Para producir la inversión de población, muchos átomos deben ser excitados al nivel superior, requiriendo esta situación mucha energía. Además, parte de la radiación láser puede ser absorbida para excitar alguno de los numerosos átomos que permanece en el estado fundamental. Este es básicamente el esquema del primer láser (láser de rubí) que realizó T. Maiman en 1960. Normalmente los láseres que presentan este esquema energético operan en modo pulsado.

■ Láser de 4 niveles:

La Fig. 7.3, muestra cómo el esquema de cuatro niveles puede mejorar algu-

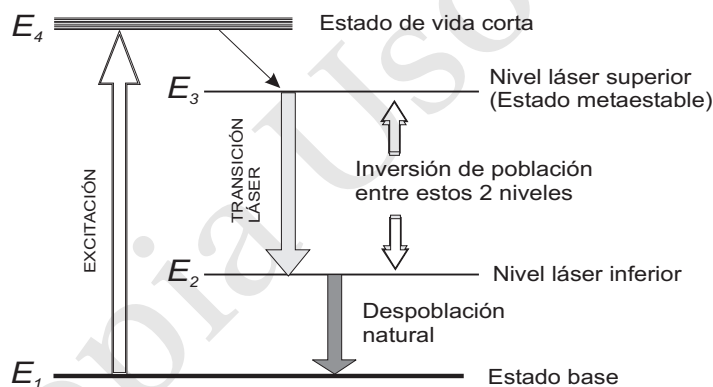


FIGURA 7.4: Niveles de energía relevantes en un láser de 4 niveles

nos de los inconvenientes de los láseres de 3 niveles. En particular, nótese que se puede obtener una acción láser efectiva con sólo un 2% de la población en el nivel láser superior y un 1% en el nivel láser inferior. Además, si el nivel láser inferior presentara una vida media menor que el nivel láser superior, esto significaría un constante vaciado del nivel láser inferior que ayuda a mantener la inversión de población. Por otra parte, la emisión estimulada no puede ser absorbida por los átomos que están en el estado base. Este esquema energético es el que presenta, por ejemplo, el láser de gas de He-Ne.

7.3. Características generales del láser

Cabe destacarse que la razón histórica por la cual desde los primeros estudios de Einstein acerca de la emisión estimulada hasta la realización del primer láser pasaran 40 años, fue la suposición mayoritaria acerca de la extrema dificultad de invertir la población; de hecho se pensaba que la materia en forma natural sólo podía estar en equilibrio térmico. En estas condiciones de equilibrio térmico, la mayor energía luminosa la proporcionaría el cuerpo negro ideal. Sólo después de la construcción del primer máser en 1957, se reconoció la existencia de máseres naturales en el espacio constituidos por nubes gaseosas cercanas a estrellas muy calientes. La luz de la estrella excita moléculas del gas a niveles altos de energía que luego decaen a estados metaestables que, si coinciden con niveles inferiores de menor vida media, pueden dar lugar a situaciones de inversión de población y por tanto acción máser. No obstante, dado que los máseres cósmicos no producen rayos, su radiación se identifica con la acción máser debido a su carácter muy monocromático, coincidiendo su longitud de onda con las transiciones de ciertas moléculas (un gas caliente usual emitiría un espectro continuo).

7.3.1. Cavity Resonante

Según acabamos de señalar, una situación en la que exista inversión de población no basta para producir haces muy directivos, ya que cada ente excitado emite en direcciones distintas y no correlacionadas. Para extraer energía eficientemente de un medio con inversión de población y producir un haz láser, se necesita una **cavidad resonante** que ayude a formar/amplificar la emisión estimulada por realimentación (reflejando parte de la emisión en el medio láser). Un ejemplo de cavidad resonante se muestra en la Fig. 7.5.

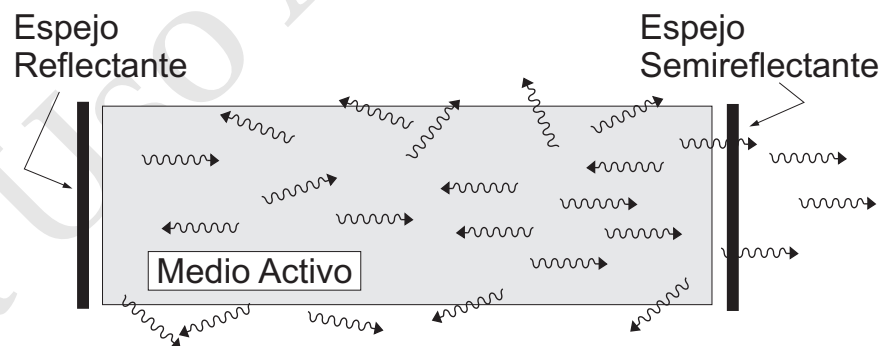


FIGURA 7.5: Cavidad resonante en un medio láser.

Se denomina **ganancia** a la cantidad de emisión estimulada que un fotón puede generar cuando viaja una cierta longitud. Si la ganancia es muy alta se puede prescindir del resonador. En caso de que la ganancia no sea tan alta, el efecto de los espejos hace que se estimule más radiación en la dirección perpendicular a los espejos y si uno de los espejos es semitransparente se llega a un equilibrio entre la potencia que sale y la que está resonando. En un láser funcionando en modo continuo, debe cumplirse que

$$\text{Amplificación} = \text{Pérdidas} + \text{Potencia emitida} .$$

En principio, aquellas radiaciones cuya longitud de onda sea un divisor del

doble de la longitud del resonador,

$$N\lambda = 2L, \quad (7.19)$$

serán fuertemente amplificadas. Aunque esta condición parece muy restrictiva, hay que considerar que la acción láser presentará una cierta **anchura de línea**, $\Delta\lambda$, debido a

- La curva de ganancia del medio activo presenta una cierta anchura espectral, Λ , centrada en torno a la longitud de onda central, λ (ver Fig. 7.6);
- Las diferentes longitudes de onda que pueden resonar en la cavidad están muy próximas unas de otras puesto que el sistema medio activo–resonador amplificará aquellas longitudes de onda separadas una distancia

$$\lambda_F = \frac{\lambda^2}{2L}, \quad (7.20)$$

o, en términos de la frecuencia, aquellas que estén separadas

$$\nu_F = \frac{c}{2L}. \quad (7.21)$$

(La expresión (7.20) se obtiene a partir de (7.21) considerando que en general $\Delta\lambda = (\lambda^2/c)\Delta\nu$). Según se muestra en la Fig. 7.6, varias resonancias de la cavidad pueden ajustarse dentro de la anchura espectral de un láser típico.

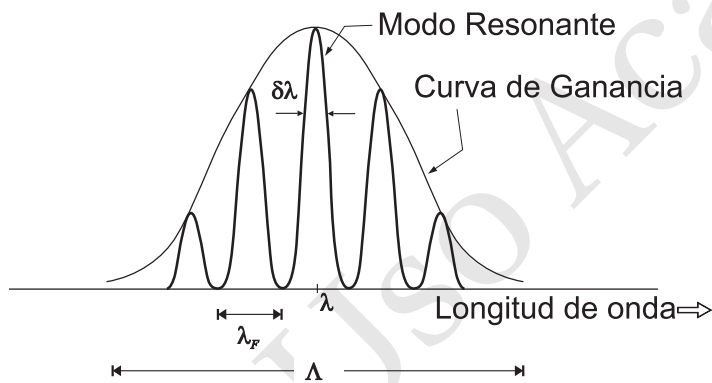


FIGURA 7.6: Diversas resonancias en el margen de la anchura de banda del láser.

Esto implica que el haz láser producido por una cavidad simple no es puramente monocromático. El diseño de la cavidad y de la óptica de salida determinará por tanto las características del rayo láser (anchura de banda, $\Delta\nu$, tamaño, divergencia, distribución de luz, etc...). Cabe señalar que un resonador de espejos planos no es muy conveniente ya que un posible pequeño desalineamiento entre sus espejos redundaría en unas pérdidas de potencia importantes. El uso de espejos curvos ayudaría a mantener la luz en el medio láser, constituyendo así un resonador estable. La óptica de la cavidad resonante del láser puede ayudar también a seleccionar una determinada longitud de onda. Para ellos puede usarse una óptica que presente menos pérdidas en la longitud de onda deseada o bien puede usarse una rejilla de difracción en el espejo (ver Fig. 7.7).

La anchura de línea del láser dependerá, por tanto, de cuántos modos resonantes sean finalmente amplificados por el sistema láser. Si sólo se emite un modo resonante, entonces $\Delta\lambda = \delta\lambda$; si se amplifican dos modos resonantes, $\Delta\lambda \approx \lambda_F$; y si se amplifican muchos modos resonantes, entonces $\Delta\lambda \approx \Lambda$.

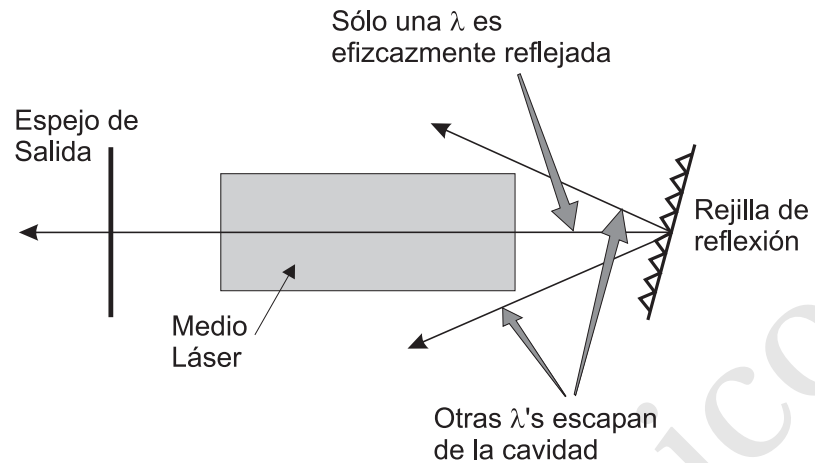


FIGURA 7.7: Disposición para sintonizar una longitud de onda.

7.3.2. Coherencia

Para que exista coherencia, deben darse las dos siguientes condiciones:

1. Las ondas deben empezar con la misma fase en la misma posición.
2. Las longitudes de onda de estas ondas deben ser la misma.

No obstante, la luz láser no presente una coherencia perfecta debido a

- No todos los fotones del rayo láser provienen del mismo fotón original, de modo que no todos empiezan en fase.
- El principio de incertidumbre causa pequeñas variaciones en las longitudes de onda de los fotones emitidos de forma que al acumularse pueden llegar a ser significativas.

En la práctica, los haces más coherentes provienen de láseres que trabajan en modo continuo y con pequeña ganancia.

Podemos distinguir dos tipos de coherencia:

- **Coherencia Temporal:** mide el tiempo durante el que las ondas luminosas permanecen en fase cuando viajan. Se denomina **longitud de coherencia**, L_c , a aquella longitud en la que luz presenta coherencia temporal, y puede expresarse como

$$L_c = \frac{c}{2\Delta\nu} = \frac{\lambda^2}{2\Delta\lambda}, \quad (7.22)$$

dependiendo fuertemente de la anchura de banda. Valores típicos de esta longitud para diferentes luces son: Bombilla común ~ 700 nm; diodo láser ~ 0.3 mm; láser de He-Ne ~ 10 cm y láser He-Ne sintonizado ~ 100 m.

- **Coherencia Espacial:** mide el área sobre la cual la luz es coherente (no está relacionada con la coherencia temporal). En un láser que emite un único modo transversal³, la coherencia espacial viene determinada por el diámetro del haz.

³ Los haces pueden oscilar en diversos modos transversales, que se manifiestan en diferentes patrones de la intensidad transversal del haz.

No siempre la coherencia es conveniente, puesto que a veces se forman patrones de interferencia indeseables debido a la presencia de minúsculas partículas presentes en el medio en el que se propaga el haz (estos patrones se manifestarían como ruido de fondo).

7.3.3. Longitud de Onda

Como ya se señaló, en la práctica la luz láser no es estrictamente monocromática sino que presenta usualmente una cierta **anchura de línea** $\Delta\lambda$. Esto se debe principalmente a la anchura energética de los niveles entre los que se produce la transición láser. Además de esta razón, también contribuyen a la anchura de banda de la curva de ganancia:

- El principio de Incertidumbre. Una transición está definida por una función que presenta un máximo pronunciado en cierta longitud de onda, pero esta función es continua.
- Movimiento inherente (por ejemplo, térmico) de los átomos y moléculas que produce efecto Doppler. Este efecto provoca típicamente una $\Delta\nu/\nu \sim 1/10^6$.
- Interacción de los átomos que participan en la acción láser con otros átomos presentes (por ejemplo, choques).

No obstante, a pesar de la anchura de la curva de ganancia, un buen diseño de la cavidad resonante puede producir una muy pequeña anchura de banda en la radiación emitida.

7.3.4. Divergencia del rayo

La divergencia del rayo láser depende principalmente del diseño del resonador, el tamaño de la apertura de salida y los efectos de difracción. Una divergencia típica es del orden de 1 miliradian (los diodos láser presentarán divergencias mucho mayores).

La difracción producida por el orificio de salida del haz determina un límite teórico mínimo para la divergencia del rayo (esto es, $\theta = \lambda/D$).

7.3.5. Potencia de Salida

Desde el punto de vista del usuario, la potencia de salida del láser es una de las características más importantes. No obstante, esta potencia de salida es únicamente una fracción de la potencia de entrada, quedando la potencia de salida determinada por lo que se conoce como **eficiencia** del láser. Esta eficiencia suele ser muy baja ya que intervienen muchos procesos y cada uno de ellos aporta un cierto grado de eficiencia que al multiplicarse va disminuyendo. Algunos puntos que reducen la eficiencia pueden ser:

- la potencia suministrada por la fuente de alimentación (usualmente potencia eléctrica) debe ser frecuentemente transformada en energía luminosa apta para producir la inversión de población.
- La potencia de excitación no es usada totalmente para promover átomos al estado láser superior; parte de esta potencia puede ser absorbida por el medio y

disipada posteriormente en forma de calor. En el caso de excitación luminosa, del espectro total sólo será aprovechada la parte de energía que corresponda al rango de longitudes de onda involucradas en el proceso de bombeo.

- En el láser de 4 niveles, la excitación inicial requiere más energía de la que luego entrará en juego en la acción láser.
- En el láser de 3 niveles, el proceso de inversión de población requiere al menos promover el 50 % de la población total. Toda esta energía no puede usarse en la acción láser y esto se manifiesta en la presencia de un umbral de energía a partir del cual ya se produce la acción láser (este hecho será significativo en el diodo láser)

7.3.6. Modulación

Muchas aplicaciones de láser requieren cambiar o modular la potencia del haz, bien para hacerla variar en el tiempo o simplemente para “encenderla” o “apagarla”. Los mecanismos más típicos para producir esta modulación son:

- Modulación mecánica: con diafragmas o ruedas dentadas giratorias. No son muy rápidas y no pueden modular la intensidad.
- Modulación directa: controlando los sistemas de bombeo. Salvo en los semiconductores, este sistema no será muy efectivo ya que existirá un tiempo de retraso importante.
- Modulación electro-óptica: consiste en un girador de la polarización que puede ser controlado por una fuente de tensión externa.

7.4. Láseres Semiconductores

La idea inicial de realizar un láser usando materiales semiconductores parece que fue propuesta de forma esquemática por J. von Neumann en 1953. Posteriormente Basov en 1961 llevó a cabo un desarrollo teórico más completo que finalmente redundó en la realización del primer láser semiconductor funcionando a 77 K por R.N. Hall en 1962. El primer láser semiconductor a temperatura ambiente fue realizado en el año 1970 por varios grupos rusos y en los ATT Bell Laboratories.

El láser semiconductor suele denominarse **diodo láser** debido a que se basa en las propiedades físicas de la unión de dos semiconductores dopados tipo p y n , al igual que los diodos de rectificación eléctrica. También se conocen estos láseres con el nombre de **láseres de inyección** debido a que el procedimiento de bombeo más usado consiste en la inyección de portadores de carga en la unión pn .

7.4.1. Propiedades eléctricas

A continuación se comentarán brevemente algunas de las propiedades más relevantes de los semiconductores en cuanto a su relación con el láser de inyección:

- Las propiedades eléctricas/ópticas de los semiconductores están relacionadas con la distribución energética de los electrones que son colectivamente compartidos por todos los iones del cristal.

► El esquema de los niveles de energía de un semiconductor intrínseco ideal se muestra en la Fig. 7.8. La densidad de estados en la Banda de Valencia (BV) y en la

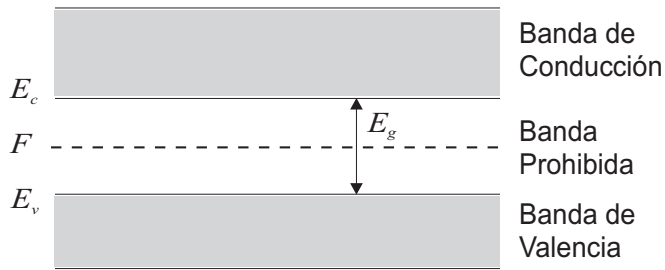


FIGURA 7.8: Esquema energético de bandas de un semiconductor intrínseco ideal.

Banda de Conducción (BC) es una función de la energía, E y cada estado puede ser ocupado como máximo por dos electrones (considerando los dos posibles estados de spin).

► Dado que los electrones son **fermiones**, la probabilidad de ocupación de un estado viene dada por la función de distribución de Fermi-Dirac —expresión (7.2) con $A = 1$:

$$f(E) = \frac{1}{\exp\left(\frac{E-F}{kT}\right) + 1} \quad (7.23)$$

El parámetro F se conoce como **energía de Fermi** del semiconductor y puede interpretarse como el nivel más energético más alto posible que pueden tener los electrones a $T = 0\text{K}$. A $T > 0\text{K}$. La conducción eléctrica es fruto del movimiento de electrones en las bandas de conducción y valencia. El movimiento global de los electrones en la BV puede ser formalmente descrito como el movimiento de unas pseudopartículas de masa y carga positivas, llamadas **huecos**.

► Aunque en la Fig. 7.8 se ha obviado, el diagrama de bandas de energía es una representación de la función $E = E(\mathbf{p})$, donde \mathbf{p} es el “pseudomomento” del electrón que está relacionado con su vector de onda mecánico-cuántico⁴ $\mathbf{k} = \mathbf{p}/\hbar$.

Una representación más realista de la estructura de bandas de energía de un semiconductor puede verse en la Fig. 7.9, donde se muestran los dos casos usuales de semiconductor directo (por ejemplo: GaAs) y semiconductores indirectos (Si, Ge...).

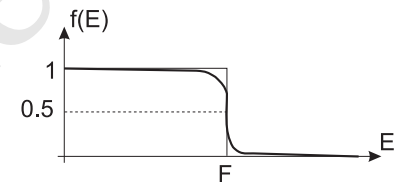
En un semiconductor directo, un electrón puede realizar una transición energética desde el valor más bajo de la BC al valor más alto de la BV sin un cambio significativo en su momento. Por el contrario, en un semiconductor intrínseco, la anterior transición requeriría un cambio importante de momento. En una estructura periódica de período espacial a (típicamente $a \sim 1\text{Å}$), el mayor cambio de momento permitido sería

$$p_{\max} \sim \frac{h}{2a}.$$

Por otra parte, el momento de un fotón luminoso ($\lambda_f \sim 10^4 \text{Å}$), $p_f = h/\lambda_f$ y, por tanto,

$$\frac{p_f}{p_{\max}} \sim \frac{2a}{\lambda_f} (= 10^{-4}).$$

⁴ Debe tenerse en cuenta que, en un sentido estricto, \mathbf{k} representa el vector de ondas característico del grupo de ondas que describe el comportamiento del electrón en el semiconductor.



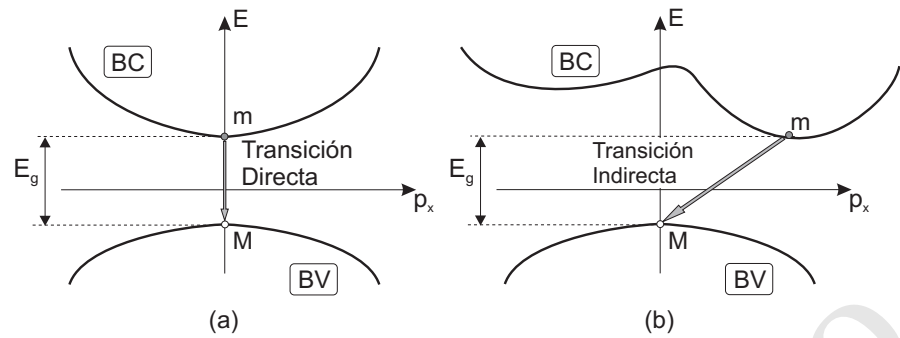


FIGURA 7.9: Esquema energético de bandas realista de (a) un semiconductor directo y (b) un semiconductor indirecto.

Podemos ver entonces que las diferencias de momento requeridas en una transición indirecta no pueden ser satisfechas únicamente por la participación de un fotón óptico y así se requiere también la participación de otro tipo de partícula que proporcione la cantidad de momento requerido. Esta partícula es usualmente un *fonón*, que básicamente representa una “vibración” de la red cristalina cuyo efecto principal, en el presente proceso, es la disipación de calor. Por otra parte, en el caso de una transición directa, los pequeños cambios de momento requeridos en el proceso pueden ahora ser satisfechos únicamente por un fotón.

► En el caso de un semiconductor dopado, el esquema de niveles energéticos se ve modificado ligeramente. La estructura general de bandas de energía no se modifica apreciablemente pero aparecen nuevos niveles energéticos discretos que corresponden bien a las impurezas donadoras (en el entorno del borde inferior de BC) o a las impurezasceptoras (en el entorno del borde superior de BV). Tal como muestra la Fig. 7.10, esto se refleja básicamente en el desplazamiento del nivel de Fermi hacia los extremos de las bandas. El tipo de dopado predominante sirve para clasificar los semiconductores en tipo *p* y tipo *n*.

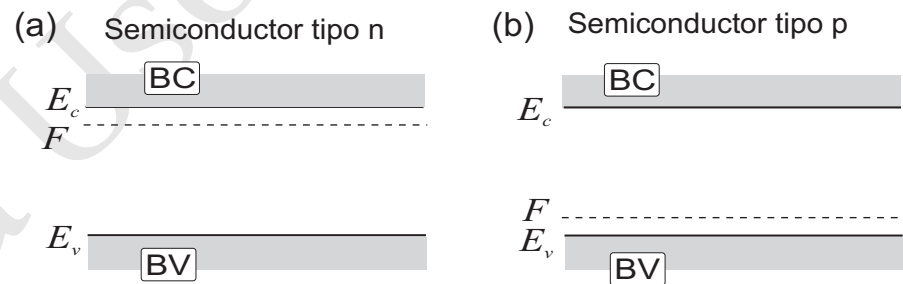


FIGURA 7.10: (a): Semiconductor tipo *n*; (b) semiconductor tipo *p*.

Las más interesantes aplicaciones de los semiconductores surgen cuando se dopan regiones distintas con distintos tipos de impurezas, formándose así una **región de transición** que se denomina unión *pn*. En esta región de transición se igualan los niveles de Fermi de ambas partes formándose un potencial de contacto, según se muestra en la Fig. 7.11(a). Si la anterior unión *pn* se somete ahora a una fuente de tensión externa de modo que eleva el potencial del lado *p* con respecto al del lado *n*, lo que se denomina **polarización directa**, el esquema energético resultante puede verse en la Fig. 7.11(b). En esta situación, los electrones del lado *n* se desplazarán hacia el lado *p* y análogamente huecos del lado *p* viajarán hacia el lado *n*. En consecuencia, electrones y huecos se encontrarán en la misma región espacial, siendo posible su **recombinación** mediante las transicio-

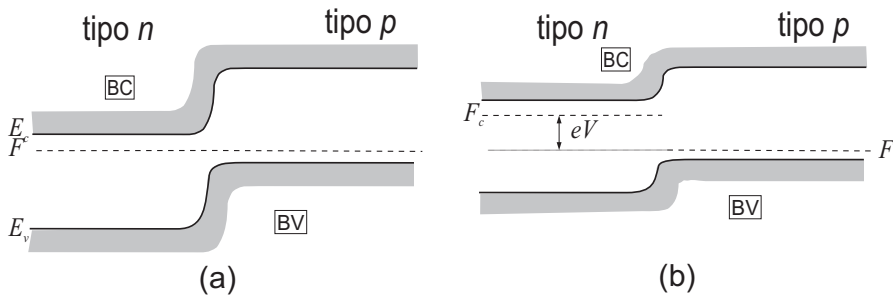


FIGURA 7.11: Esquema energético de la región de transición de una unión pn (a): en equilibrio térmico y (b): sometida a una fuente de tensión externa que la polariza directamente.

nes directas/indirectas que antes se mencionaron (ver Fig. 7.12).

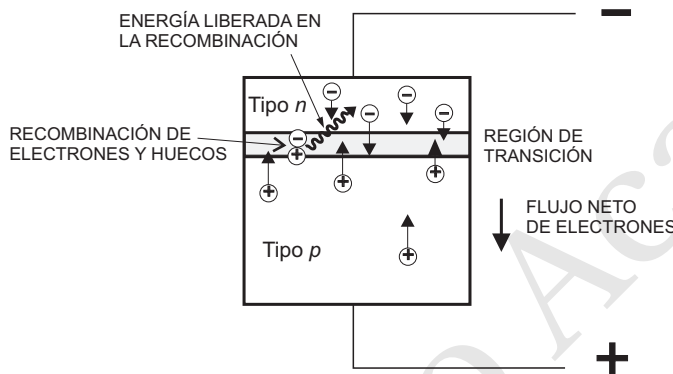


FIGURA 7.12: Flujo de portadores en un diodo semiconductor polarizado directamente.

Si en estas recombinaciones predomina la emisión espontánea, se tiene un LED (light emitting diode) y si por el contrario, predominan las emisiones estimuladas, se produce una acción láser.

Si en los semiconductores de la situación anterior se incrementan mucho los niveles de dopaje, el efecto de las impurezas añadidas podría llegar a distorsionar la estructura de bandas energéticas del semiconductor. En el presente caso consideraremos que un incremento apreciable de las impurezas donadoras únicamente provoca que el nivel de Fermi del semiconductor aparezca en el interior de la Banda de Conducción. En este caso el semiconductor se denomina **degenerado**. La Fig. 7.13(a) muestra la situación energética a la que se llega en la región de transición de una unión pn degenerada en equilibrio térmico. Si a esta unión se le aplica una polarización directa, V , tal que $eV > E_g$, se llega a la situación mostrada en la Fig. 7.13(b). Es importante notar que aunque estamos en una situación de no equilibrio térmico, lejos de la región de transición y dentro de cada banda, se establece rápidamente un cuasi-equilibrio térmico (esto es, la dinámica de los electrones fuera de la región de transición no es influenciada significativamente por el campo externo aplicado). Esto posibilita la descripción de los electrones en las bandas mediante la distribución de Fermi-Dirac y se puede, por tanto, seguir usando el parámetro *Energía de Fermi* como característico de esa banda.

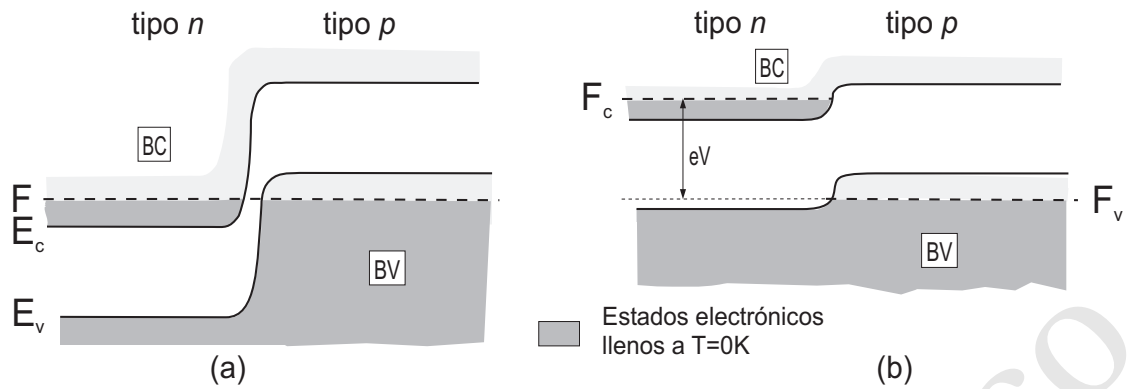


FIGURA 7.13: Esquema energético de la región de transición de una unión pn degenerada (a): en equilibrio térmico y (b) polarizada directamente.

7.4.2. Propiedades ópticas

La interacción de la radiación con los semiconductores está fundamentalmente determinada por la estructura energética que posean. Así, en un semiconductor homogéneo

- La emisión se producirá en la forma de fotones de energía igual a la anchura energética de la banda prohibida, $h\nu = E_g$, dado que, dentro de cada banda, los electrones se mueven muy rápidamente hacia los niveles más bajos/altos disponibles en la banda de conducción/valencia. Esta emisión de fotones es fruto del llamado proceso de recombinación de pares electron-hueco.
- La absorción de radiación puede producirse en un rango mucho más ancho de frecuencias, denominándose a este proceso generación de un par electrón-hueco.

La recombinación de portadores más probable en un semiconductor que forma una unión pn se produce cuando electrones y huecos se encuentran en la misma región espacial. Según se verá más adelante, para densidades de corriente moderadas, la radiación emitida tiene características de emisión espontánea y únicamente a partir de cierta intensidad umbral, I_U , se producirá emisión estimulada.

Dado que los procesos de emisión y absorción mediante recombinación o generación de pares electron-hueco pueden considerarse como procesos de interacción luz-materia descritos en la sección 7.2, para que exista acción láser la cuestión es

¿Bajo qué condiciones la razón de emisión estimulada supera a la absorción en una unión pn en polarización directa?

Dado que los procesos de interacción radiación-semiconductor ocurrirán predominantemente en la región de transición, los procesos involucrados se suponen asociados a transiciones entre un grupo de niveles del borde inferior de la Banda de Conducción y un grupo de niveles del borde superior de la Banda de Valencia (ver Fig. 7.14).

Las probabilidades de emisión y absorción estarán relacionadas con la densidad de la radiación presente, $u(\nu)$, y la probabilidad de ocupación de los niveles

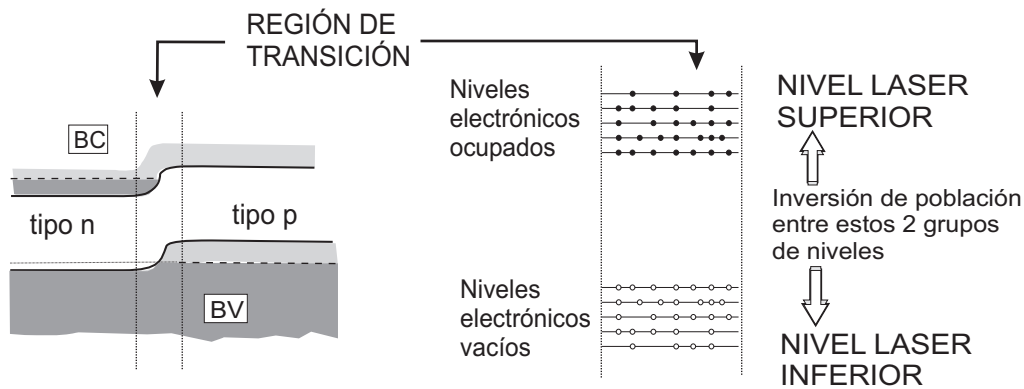


FIGURA 7.14: Niveles involucrados en la acción láser de una región de transición de una unión pn degenerada.

de la banda de conducción, $f_c(E)$ y en la banda de valencia, $f_v(E)$. Así, se tiene que

- *Emisión estimulada:* proporcional al número de transiciones posibles entre los estados superiores llenos y los estados inferiores vacíos.
- *Absorción estimulada:* proporcional al número de transiciones posibles entre estados inferiores llenos y estados superiores vacíos.

La probabilidad de que un estado con energía E_2 en la Banda de Conducción esté ocupado y que un estado con energía E_1 en la Banda de Valencia esté vacío es

$$f_c(E_2) [1 - f_v(E_1)] \quad \text{o bien} \quad f_c(1 - f_v)$$

y por tanto el ritmo de transiciones $E_2 \rightarrow E_1$, según (7.7), será proporcional a $B_{21}u(\nu)f_c(1 - f_v)$. Análogamente, el ritmo de transiciones $E_1 \rightarrow E_2$ será proporcional a $B_{12}u(\nu)f_v(1 - f_c)$.

Para que el ritmo de emisión estimulada supere a la absorción deberá ocurrir entonces que $f_c(1 - f_v) > f_v(1 - f_c)$, que puede expresarse simplemente como

$$f_c(E_2) > f_v(E_1), \quad (7.24)$$

esto es, la *probabilidad de ocupación* en el nivel de energía E_2 en BC debe ser mayor que la del estado E_1 en BV.

Dado que el tiempo de vida medio de los portadores es mucho mayor que el tiempo en que los electrones/huecos en ambas bandas alcanzan el equilibrio térmico, se supone que la probabilidad de ocupación de los niveles energéticos correspondiente a los electrones o huecos inyectados en la región de transición viene descrita por la estadística de Fermi-Dirac, por lo que, de acuerdo a (7.23),

$$f_c^{-1} = 1 + \exp\left(\frac{E_2 - F_c}{kT}\right) \quad \text{y} \quad f_v^{-1} = 1 + \exp\left(\frac{E_1 - F_v}{kT}\right).$$

A la vista de las anteriores expresiones, la condición (7.24) implica que

$$\exp\left(\frac{E_2 - F_c}{kT}\right) < \exp\left(\frac{E_1 - F_v}{kT}\right), \quad (7.25)$$

que puede reducirse a

$$F_c - F_v > E_2 - E_1. \quad (7.26)$$

La condición de inversión de población se da, por tanto, para aquellos niveles E_2 de BC y niveles E_1 de BV que satisfagan la condición (7.26), esto es, los niveles de BC por debajo de F_c y los niveles de BV por encima de F_v . Puesto que $E_2 - E_1$ es necesariamente mayor que la anchura de la banda energética, E_g , se puede concluir finalmente que la condición para que exista inversión de población en la región de transición de una unión pn directamente polarizada es

Condición de inversión de población en un diodo de inyección

$$F_c - F_v > E_g \quad (7.27)$$

De todo lo expuesto anteriormente podemos concluir que la acción láser requiere:

1. *El semiconductor debe estar degenerado* (pues de otra manera, la diferencia entre los niveles de Fermi de la zona p y la zona n nunca superaría la anchura energética de la banda prohibida).
2. *La tensión de polarización debe exceder el valor E_g/e* ($e = 1.6 \times 10^{-19}$ C es el módulo de la carga del electrón).
3. *El semiconductor debe ser directo* (para que las posibles recombinaciones electrón-hueco se realicen con la mayor probabilidad de emisión de fotones).

A pesar de cumplir las condiciones anteriores, una seria limitación para la acción láser en los semiconductores es la **absorción de portadores libres**. Esto se debe a que la estructura de bandas del semiconductor permite que un electrón de la Banda de Conducción pueda interactuar con uno de los fotones emitidos de tal forma que el electrón pase a uno de los niveles superiores y vacíos de dicha banda. Este exceso de energía será posteriormente liberado en forma de calor, dándose de esta manera transiciones energéticas de valor $h\nu$ que pueden reducir considerablemente la densidad de emisión estimulada y por tanto dificultar la realimentación de la acción láser.

Según se indicó anteriormente, cuando la corriente de alimentación del diodo láser es pequeña, éste genera emisión espontánea, según el mismo proceso que la emisión de un LED. Sin embargo, al aumentar la corriente de alimentación (o análogamente la tensión de polarización hasta conseguir que se cumpla (7.27)), se alcanza un nivel umbral donde se invierte la población de electrones y comienza la acción láser. Este fenómeno puede observarse en la Fig. 7.15 donde puede

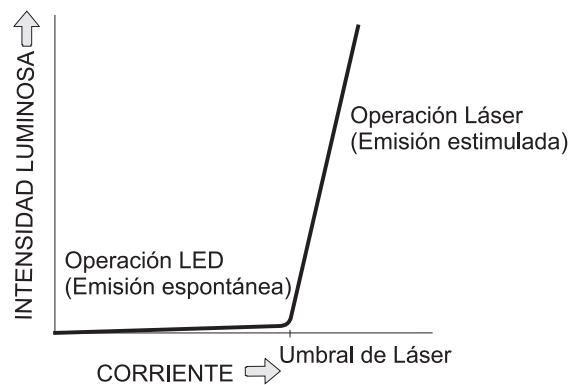


FIGURA 7.15: Potencia de salida de un láser semiconductor en función de la corriente de alimentación.

verse como por debajo de la corriente umbral, la intensidad luminosa proporcionada por el diodo es muy pequeña, disipándose la mayor parte de la potencia de

entrada en calor. A partir del umbral de intensidad, aumenta considerablemente la eficiencia óptica del diodo, aunque debe tenerse en cuenta que si la intensidad umbral es alta, esto significa una considerable pérdida de potencia en calor para mantener la acción láser. Las investigaciones actuales en los diodos láser van encaminadas a reducir esta corriente umbral y así asegurar funcionamientos muy duraderos a temperatura ambiente.

7.4.3. Estructura del láser semiconductor

Como ya se indicó en la sección 7.3, una acción láser eficiente requiere el uso de resonadores ópticos. En el caso del diodo láser horizontal y tal como se muestra en la Fig. 7.16, la realimentación que concentra la emisión estimulada en el plano

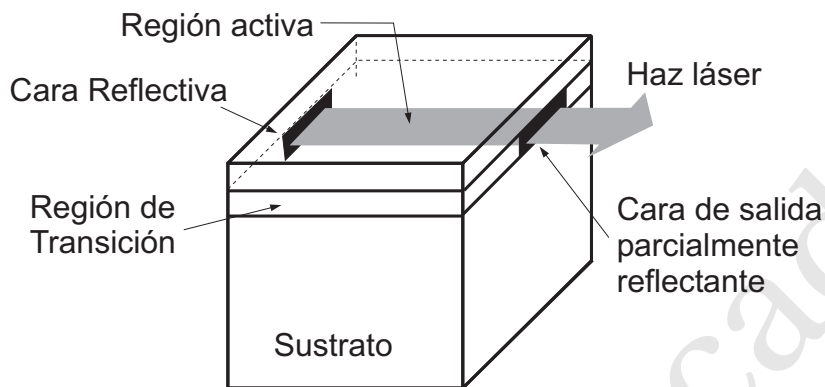


FIGURA 7.16: Las regiones reflectantes producen acción láser en el plano de la unión en un diodo láser

de la unión, proviene de los bordes pulidos del cristal semiconductor que actúan como espejos debido al alto índice de refracción de los semiconductores ($\sim 3 - 4$), que hace que estas caras pulidas reflejen aproximadamente el 30% de la luz incidente.

Una de las características más determinantes en las propiedades del láser semiconductor es la estructura de las capas adyacentes a la capa activa (región de transición donde se realiza la acción láser). En los primeros diodos láser, todas las capas estaban construidas sobre la base del mismo material y por tanto se denominan **homoestructuras**. Posteriores desarrollos llevaron a la realización de láseres semiconductores en los que las capas adyacentes estaban construidas con distintos materiales, denominándose **heteroestructuras**. Algunas características relevantes de ambas son:

- **Homoestructuras.**

No proporcionan un confinamiento óptico óptimo ya que los índices de refracción de las diversas capas eran muy parecidos; ver Fig. 7.17(a). Esto exigía unas densidades de corrientes umbrales muy altas, $I_U/\text{cm}^2 \sim 50000 \text{ A}/\text{cm}^2$, que daban lugar a muchas pérdidas por calor (lo cual exigía trabajar a temperaturas de Helio líquido $T \sim 77 \text{ K}$) y poca vida media de los dispositivos.

- **Heteroestructuras.**

Las capas presentan distintos índices de refracción, de forma que, según se muestra en la Fig. 7.17(b), esto crea un efecto guía de onda (similar al de la fibra óptica) que confina la luz en la capa activa. Además, si las capas adyacentes poseen bandas energéticas prohibidas de mayor anchura se reducirá

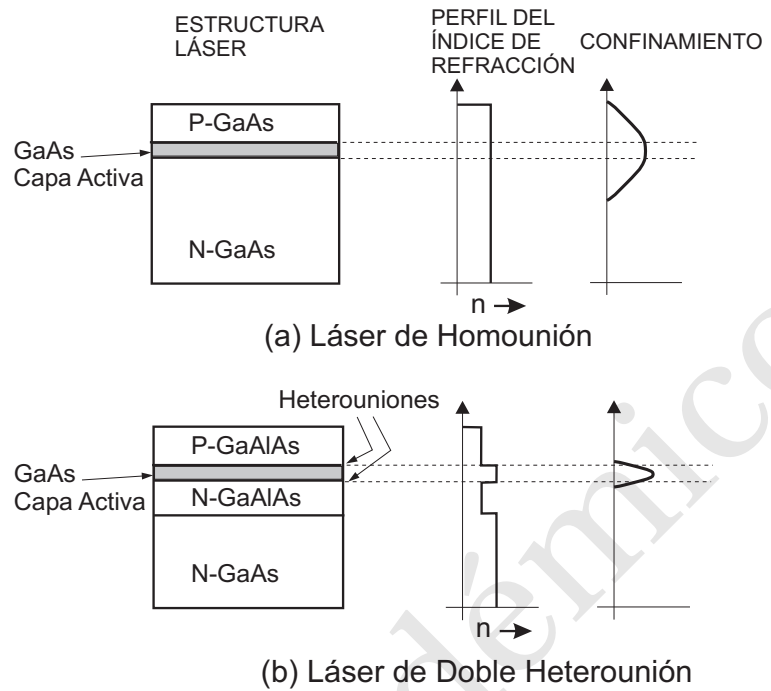


FIGURA 7.17: Láser de homounión y heterounión, señalando el perfil del índice de refracción y el confinamiento luminoso.

sustancialmente la absorción por portadores libres. Este tipo de láseres puede ya funcionar en modo continuo a temperatura ambiente al ser $I_U/cm^2 \sim 2000 - 4000 A/cm^2$, $I_U \sim 1.2 A$.

El esquema con el que se realizan actualmente la mayoría de los diodos láser comerciales consiste en una modificación de la heteroestructura anterior que se muestra en la Fig. 7.18. Esta estructura proporciona confinamiento lateral para la luz, lo que reduce la corriente umbral hasta $I_U \sim 60 mA$.

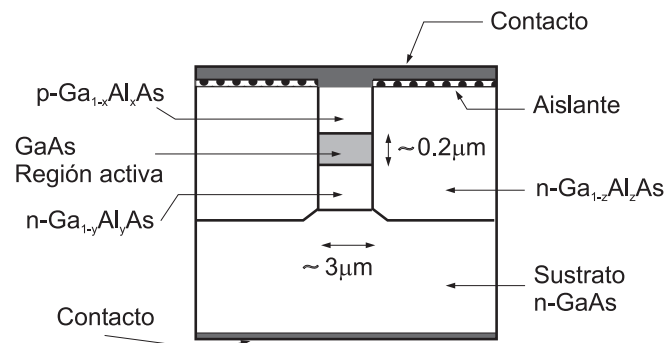


FIGURA 7.18: Heteroestructura típica de los láseres semiconductores comerciales

Usando esta estructura, es posible fabricar sobre el mismo sustrato semiconductor, muchos láseres paralelos tipo-tiras y así conseguir potencias de hasta 10 vatios en modo continuo.

Pozos cuánticos

Un tipo de confinamiento distinto basado en fenómenos cuánticos se ha conseguido mediante el uso de capas de material muy finas (~ 10 nm.) que se comportan respecto a los electrones como un **pozo cuántico**. La idea básica del pozo cuántico se consigue al intercalar una capa muy fina con una banda prohibida de anchura pequeña entre capas gruesas con anchuras de banda prohibida mayores. Los electrones que pasan por esta configuración de bandas pueden ser *capturados* en el pozo cuántico. Estos electrones capturados tienen suficiente energía para permanecer en la Banda de Conducción del material que forma la fina capa de pequeña anchura de banda prohibida pequeña, pero no para entrar en la capa de mayor anchura de banda prohibida.

Si los pozos cuánticos se colocan en una unión *pn* de un diodo láser, permiten concentrar en capas muy finas los electrones y huecos, consiguiendo una eficaz recombinación de éstos que disminuye considerablemente la intensidad umbral. Además, dado que las capas de confinamiento tienen diferentes índices de refracción, también posibilitan el confinamiento de luz.

Longitud de onda

En general, los medios activos de los diodos láser presentan altas ganancias, de forma que no se requieren muchas oscilaciones en la cavidad resonante para producir acción láser. Dado que típicamente la región que actúa de cavidad tiene una longitud de $\sim 300 - 500 \mu\text{m}$., los modos resonantes de esta cavidad están más separados que en otros láseres (cuyas cavidades resonantes son al menos de $\sim 10-30$ cm.) y como resultado, la mayoría de los diodos láser normalmente oscilan a una sola frecuencia *al mismo tiempo*. No obstante, debido a que la anchura de la curva de ganancia es grande, el diodo láser puede emitir distintas longitudes de onda (*saltando* de una λ resonante a otra) dando lugar a un funcionamiento bastante inestable. Este fenómeno puede solucionarse mediante el uso de dispositivos añadidos de sintonización y estabilización.

Divergencia del haz

Otra característica importante del diodo láser es la inusual divergencia del haz emitido. La geometría de la pequeña cavidad junto con la minúscula área de emisión (generalmente una pequeña rendija del orden de los μm^2) se combinan para producir un haz no circular con un ángulo de divergencia típicamente del orden de 40 grados. Se encuentra por tanto que los rayos de un diodo láser divergen más rápidamente que los de una buena lámpara de flash. Afortunadamente, el uso de una óptica adecuada permite una buena focalización del haz y así, por ejemplo, los diodos láser se usan como fuente de luz de los *punteros láser* o de lectores de CD-ROM y DVD.

Modulación

Según puede observarse en la Fig. 7.15, la salida óptica del diodo láser depende linealmente de la corriente que circula a través de éste. Esta simple dependencia lineal hace que la modulación del haz pueda conseguirse modulando directamente la corriente de alimentación, constituyendo este hecho una de las principales ventajas de los láseres de inyección. Dado que la respuesta de la salida

óptica a las variaciones de la corriente de alimentación es muy rápida, se consiguen anchuras de banda típicas en la modulación del orden de los Gigahertzios.

7.5. Aplicaciones del Láser

Las posibles aplicaciones del láser pueden dividirse en varias categorías. Una de las más usuales consiste en distinguir la acción de los láseres de baja potencia de aquellos de alta potencia. En la presente sección se expondrán únicamente algunas de las aplicaciones de los láseres de baja potencia. Generalmente, se denomina baja potencia a cantidades menores que un vatio, bajo la suposición de que el haz láser no afecta destructivamente a los objetos que ilumina. Es interesante señalar que en muchas de las posibles aplicaciones de láser, éste se emplea básicamente como una *bombilla de muy alta eficacia*. Algunas de las características diferenciadoras de los láseres de baja potencia respecto a otras posibles fuentes de luz son:

- Láseres producen luz *bien controlada* que puede ser focalizada con precisión sobre una minúscula región de un objeto. Este hecho tiene multitud de aplicaciones, por ejemplo en *contadores, lectores de códigos de barras, lectores de CD, discos magneto-ópticos, etc.*
- Al focalizarse la luz sobre una pequeña región, se pueden conseguir altas densidades de potencia, que pueden usarse para un calentamiento selectivo de pequeñas regiones.
- Debido a la alta directividad de los rayos láser, éstos pueden definir líneas rectas. Esta propiedad se usa para el diseño de *posicionadores y punteros*.
- La luz láser posee normalmente un estrecho rango de longitudes de onda, siendo este hecho altamente aprovechado en *espectroscopía*.
- La alta coherencia que presentan muchos láseres es aprovechada para la *holografía* y fundamentalmente para *metrología*.
- Los láseres pueden generar pulsos extremadamente cortos, siendo esto muy útil para la *transmisión de datos*.
- Láseres de He-Ne y de inyección producen potencias muy constantes en modo continuo durante miles de horas de uso.
- Los diodos láser son muy compactos y baratos.
- Dado que los diodos láser pueden ser modulados directamente a altas velocidades mediante cambios en la corriente de alimentación, son muy usados en la *transmisión de datos*.

7.5.1. Aplicaciones del Láser de Inyección

Debido a las peculiaridades del láser de inyección (tamaño reducido, bajo coste, modulación directa...), éste se ha convertido con diferencia en el láser más popular, siendo actualmente usado en millones de lectores de discos ópticos (en equipos de música, computadores, vídeos,...) y en la mayoría de los sistemas de alimentación de fibras ópticas para transmisión de datos.

Discos ópticos; Lectura de datos para computadores

El funcionamiento básico de lectura de un CD-ROM se muestra en la Fig. 7.19, donde puede verse como la luz de un diodo láser es focalizada sobre una disminu-

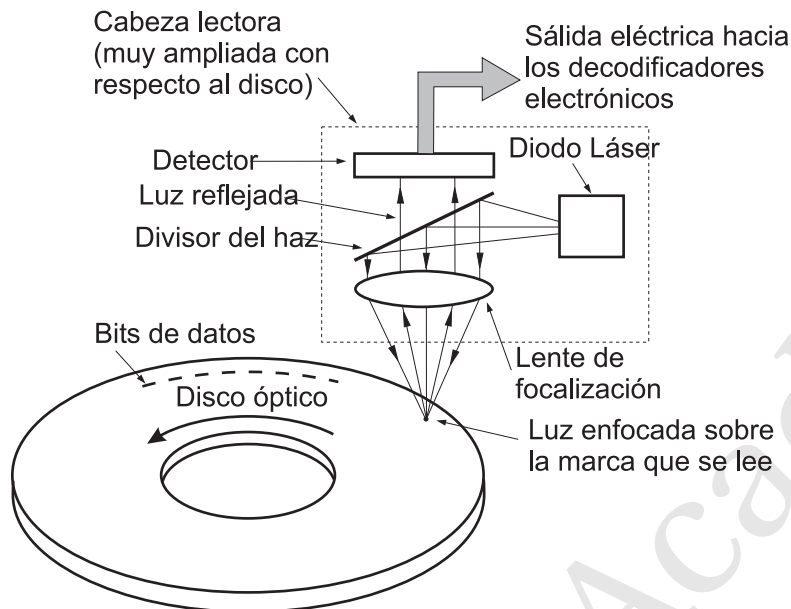


FIGURA 7.19: Sistema de lectura de un CD-ROM.

ta región de un disco que gira rápidamente. La superficie del disco está cubierta con minúsculas partículas, con diferente reflectividad que el fondo del disco, de acuerdo a un cierto patrón de bits de datos. La luz láser reflejada por el disco es enfocada sobre un detector que genera una señal eléctrica en forma de una serie de pulsos correspondientes al patrón de datos registrado en el disco.

Debido a que el haz del diodo láser puede ser enfocado sobre una región que cubre aproximadamente un diámetro de una longitud de onda, esto significa que el uso de diodos láser de $\lambda = 780 \text{ nm}$ permite que cada una de las marcas del disco debe cubrir aproximadamente $1 \mu\text{m}^2$. De este modo, un CD-ROM de 12 cm de diámetro permite almacenar $\sim 600 \text{ MBytes}$, incrementándose varios cientos de veces la densidad de almacenamiento de datos con respecto a los anteriores discos magnéticos flexibles. La focalización a distinta profundidad permite igualmente el uso de discos ópticos *multicapas*, dando así lugar a un incremento adicional muy importante de la cantidad de información que se puede almacenar.

Dado que los diodos láser son muy pequeños y pueden ser montados en la cabeza lectora, esto permite un acceso a la información muy rápido (si se dispone de un software de control adecuado) con independencia de donde esté localizada en el disco. Por otra parte, los diodos láser requieren poca potencia para su funcionamiento, siendo además esta potencia del mismo orden que la que se usa en los dispositivos electrónicos, aumentando así la compatibilidad entre la parte electrónica y la óptica.

Discos magneto-ópticos; Almacenamiento y Lectura de datos para computadores

El mayor inconveniente de los CD-ROM proviene de su carácter de *memoria de sólo lectura*. Las intensas investigaciones realizadas para conseguir discos de alta densidad de almacenamiento de datos y con posibilidad de ser regrabados por el usuario han resultado en el desarrollo del *disco magneto-óptico*. La diferencia fundamental de este disco con respecto al CD-ROM consiste en una cubierta de un material parcialmente transparente con propiedades ferrimagnéticas. Este material manifiesta una magnetización neta a temperatura ambiente que puede ser modificada por encima de cierta temperatura (temperatura de Neel) cuando está en la fase paramagnética. Tal como se muestra en la Fig. 7.20, el aumento de temperatura local de una pequeña región del disco se consigue mediante el haz

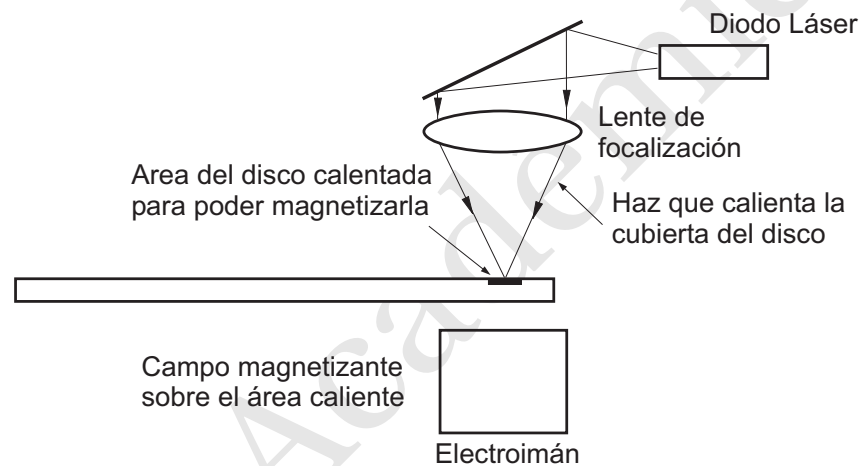


FIGURA 7.20: Esquema de grabación de un disco magneto-óptico.

focalizado de un diodo láser. Una vez que se ha alcanzado la temperatura adecuada, la pequeña región es magnetizada en la dirección impuesta por un electroimán. Nótese que la grabación de datos es puramente magnética, pero debido a que la magnetización se realiza en dirección normal al disco y que se pueden magnetizar regiones muy localizadas, se consiguen unas densidades de almacenamiento de datos muy altas.

La lectura del disco magneto-óptico se realiza mediante un procedimiento similar al del CD-ROM, aunque en este caso se usa el hecho de que la refracción de un haz polarizado en un material magnetizado produce un haz reflejado cuya polarización sufre una variación que depende de la dirección de magnetización del material refractante. Detectando, por tanto, los estados de polarización del haz reflejado por la cubierta magnetizada del disco y registrando estos estados en pulsos eléctricos se consigue recobrar el patrón de datos grabado previamente de forma magnética.

Comunicaciones por fibra óptica

Actualmente las fibras ópticas se han convertido en el soporte físico dominante en las telecomunicaciones de larga distancia *punto a punto*. Los LEDs pueden excitar los sistemas de fibra óptica cuando la transmisión se realiza a corta distancia (por ejemplo, en un mismo edificio) pero se necesitan fuentes más monocromáticas para las comunicaciones de larga distancia, usándose para este cometido como fuentes estándar los diodos láser, debido a

- Diodo láser emite en un área de unos pocos μm^2 ; lo cual es ideal para acoplarse bien con el diminuto núcleo de la fibra óptica ($\sim 10\mu\text{m}$)
- Es muy compacto y opera con los mismos niveles de tensión e intensidad que los dispositivos electrónicos convencionales.

El esquema básico de transmisión se muestra en la Fig. 7.21. La señal que al-

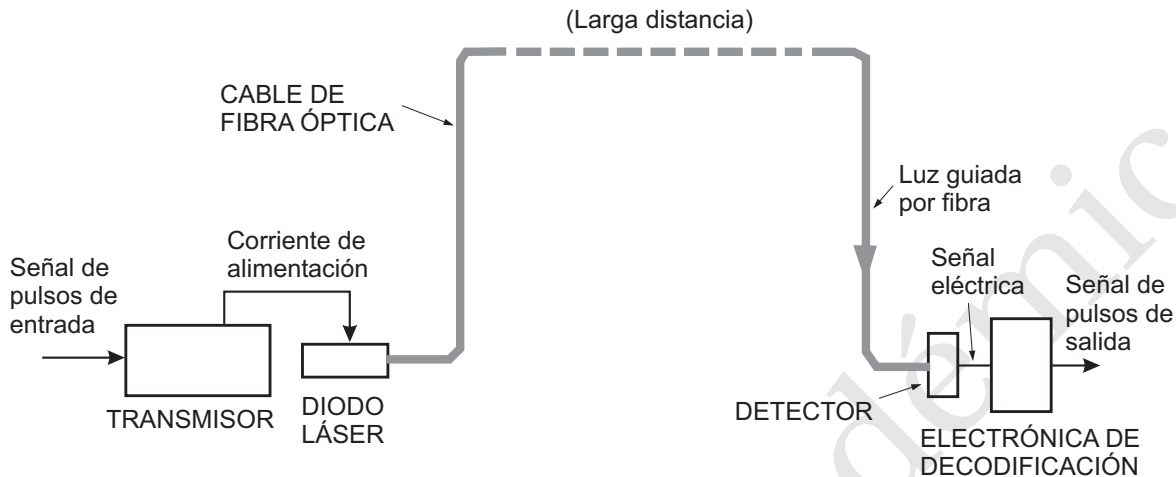


FIGURA 7.21: Sistema de fibra óptica para larga distancia

canza el transmisor genera una corriente que pasa a través de un láser semiconductor. La señal es una serie de pulsos eléctricos que genera una correspondiente serie de pulsos luminosos en el diodo láser. El láser emite los pulsos directamente al núcleo de la fibra óptica, siendo transmitida por ésta (hasta $\sim 10^4$ km) a un receptor lejano. Allí, un detector de luz convierte los pulsos luminosos en pulsos eléctricos.

7.6. Problemas propuestos

7.1: Para un sistema en equilibrio térmico a temperatura ambiente ($T \approx 300\text{K}$), calcule la relación, R , entre la tasa de emisión estimulada y la espontánea para una radiación de longitud de onda $\lambda = 600\text{nm}$. Determinar la tendencia de esta tasa con respecto a la frecuencia y a la temperatura.

Sol. $R \rightarrow 0$.

7.2: Describa las ventajas de un sistema láser de cuatro niveles respecto a un sistema láser de tres niveles.

7.3: Un láser de He-Ne tiene una cavidad resonante de 1.5m , $\lambda = 633\text{nm}$ y una anchura espectral de ganancia de 1500MHz . Calcule la separación en frecuencia, ν_F , entre dos modos adyacentes y determine el número, N , de modos resonantes que serán amplificados. Explique por qué la longitud de coherencia está en este caso determinada por la anchura espectral de ganancia.

Sol. $\nu_F = 100\text{MHz}$; $N = 15$.

7.4: Calcule la longitud aproximada, L , de la cavidad resonante de un láser de He-Ne con $\Delta\nu = 500\text{MHz}$ que oscila en dos modos.

Sol. $L = 30\text{cm}$

7.5: Calcule la longitud de coherencia de un láser semiconductor que emite en una frecuencia central de 500nm y cuya cavidad resonante es de 1mm .

Sol. $L_c \approx 1\text{mm}$.

7.6: Explique por qué la acción láser ocurre en un láser de inyección únicamente cuando la polarización externa $V > E_g/e$.

Copia Uso Académico

Tema 8

Fibra Óptica

8.1. Introducción

En los primeros temas hemos estado estudiando las bases y algunas posibles aplicaciones de la Fotónica, habiendo dedicado el último tema a las fuentes de luz. Sin duda, en el ámbito de la Informática una de las aplicaciones más importantes de la Fotónica es la transmisión de información a través de fibras ópticas. Básicamente, la información que se desea transmitir se codifica en una onda luminosa (ya sea en su amplitud, fase, frecuencia,...) cuyo espectro de frecuencias cae en el rango del infrarrojo. Una de las principales ventajas del uso de la fibra óptica estaría en la gran cantidad de información que es capaz de transmitir debido a la elevada frecuencia de las ondas portadoras. Este hecho haría que el ancho de banda de la fibra óptica fuese, en principio, bastante más elevado que el de otros sistemas alternativos de transmisión como cables bifilares, coaxiales, guías de ondas,...

La fibra óptica empezó a usarse comercialmente como sistema de transmisión a partir de 1970, aunque inicialmente su uso fue bastante limitado, debido principalmente a las enormes pérdidas que presentaba. No obstante, los recientes avances en la tecnología de los materiales han permitido el desarrollo de fibras ópticas con muy bajas pérdidas ($\lesssim 0.1$ dB/km) lo cual ha provocado que la fibra óptica sea actualmente uno de los sistemas de transmisión más usados en comunicaciones a larga distancia (cables submarinos con amplificadores ópticos pueden operar actualmente a velocidades de 5 gigabits por segundo) y posiblemente también lo sean en el futuro en todo tipo de comunicaciones.

Algunas de las ventajas principales de las fibras ópticas son

- *Alta capacidad de transmisión* debido al gran ancho de banda disponible a frecuencias ópticas.
- *Inmunidad a interferencias electromagnéticas*. Dado que las fibras no están compuestas de conductores metálicos sino por dieléctricos, éstas no presentan los problemas de interferencia típicos de los sistemas eléctricos.
- *Posibilidad de miniaturización*. Las fibras ópticas usuales son del tamaño de un pelo humano, de modo que un haz de fibras ópticas del tamaño de un cable coaxial corriente ($\varnothing \sim 1$ cm) estaría constituido por miles de fibras.

8.2. Guía de placas paralelas

Antes de iniciar el estudio de la fibra óptica empezaremos estudiando un modelo de la transmisión de energía luminosa en un tipo de guía de ondas muy elemental: la guía de placas paralelas. Este modelo, que consiste en dos placas paralelas de un material conductor ideal ($\sigma \rightarrow \infty$), nos dará una primera idea cualitativa de algunos de los procesos físicos más importantes asociados con la transmisión de la luz en una fibra.

Tal como muestra la Fig. 8.1, consideremos lo que sucede cuando una onda

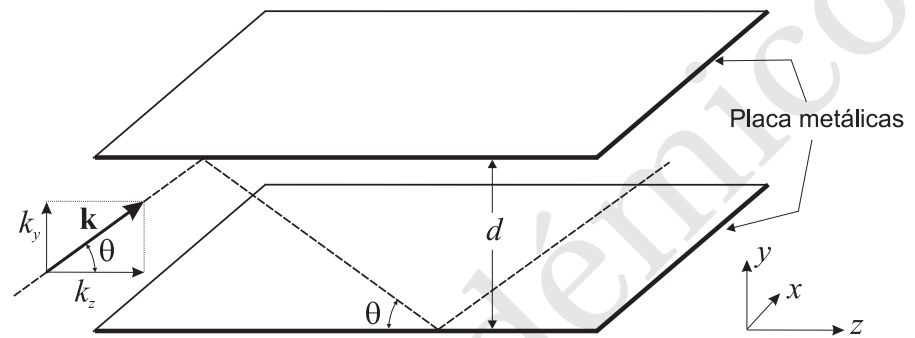


FIGURA 8.1: Propagación a lo largo de una guía de placas paralelas. En las placas metálicas, la onda sufre una reflexión total.

plana transversal electromagnética caracterizada por un vector de ondas

$$\mathbf{k} = k_y \hat{\mathbf{y}} + k_z \hat{\mathbf{z}} \quad (8.1)$$

(en general denominaremos $k_z \equiv \beta$) y *polarizada linealmente a lo largo del eje x* se propaga en el interior de la guía de placas paralelas de anchura d . Cada placa metálica actúa como un *espejo* donde se produce una reflexión total de la onda¹, provocando así múltiples reflexiones en las placas. Para que exista una onda con la orientación dada anteriormente y que ésta sea compatible con las condiciones de campo nulo impuestas por las placas metálicas debemos exigir que la componente vertical de \mathbf{k} (esto es, k_y) sea tal que dé lugar a *nodos* en $y = 0$ e $y = d$. Esta situación es similar a la estudiada en el Apartado 1.6, por lo que podemos deducir que debe verificarse que

$$k_{y,m} = m \frac{\pi}{d}, \quad m = 1, 2, 3, \dots, \quad (8.2)$$

dando así lugar al siguiente campo eléctrico:

$$\mathbf{E}_m(x, y, z) = E_{0m} \sin\left(\frac{m\pi}{d} y\right) \exp(-j\beta_m z) \hat{\mathbf{x}}. \quad (8.3)$$

Podemos observar que este campo eléctrico es justamente la combinación de una *onda estacionaria* en la dirección vertical junto con una *onda viajera* (o propagativa) en la dirección longitudinal. Los distintos campos del tipo (8.3) para cada valor de m se denominan **modos**. En general denominaremos

modos a aquellas distribuciones de campo cuyo perfil transversal y polarización se mantiene cuando se propaga.

¹ En los conductores metálicos perfectos ($\sigma \rightarrow \infty$), las componentes del campo eléctrico paralelas a ellos deben ser nulas pues de otra manera un campo finito provocaría una densidad de corriente infinita.

Modo m -ésimo de una guía de placas paralelas

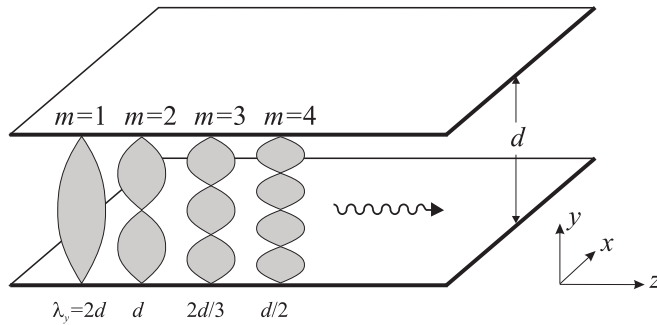


FIGURA 8.2: Ondas estacionarias en la dirección vertical (y) correspondientes a los distintos modos de propagación ($\lambda_y = 2\pi/k_y$).

Dado que para cada modo m debe cumplirse que

$$k^2 = k_{y,m}^2 + \beta_m^2, \quad (8.4)$$

podemos expresar la constante de propagación en la dirección longitudinal, β_m , como

$$\beta_m^2 = k^2 \cos^2 \theta_m = k^2 (1 - \sin^2 \theta_m), \quad (8.5)$$

siendo θ_m el ángulo mostrado en la figura adjunta, que a su vez verificará que

$$\sin \theta_m = \frac{k_{y,m}}{k} = \frac{m\pi/d}{2\pi/\lambda} = m \frac{\lambda}{2d}. \quad (8.6)$$

Observemos que a medida que aumenta el orden del modo, su correspondiente ángulo (θ_m) se hace mayor, lo que indica (según la Fig. 8.1) que los rayos se van haciendo cada vez más verticales.

Claramente para que exista propagación a lo largo de la guía debe satisfacerse que $\beta_m^2 \geq 0$, para lo cual, según (8.5), debe satisfacerse que

$$\sin \theta_m \leq 1,$$

o, equivalentemente,

$$m \frac{\lambda}{2d} \leq 1. \quad (8.7)$$

De la expresión anterior podemos extraer dos importantes conclusiones:

1. Existencia de una frecuencia de corte.

Dado que $m \geq 1$, la longitud de onda máxima, λ_{\max} , de una onda propagativa será justamente

$$\lambda_{\max} = 2d, \quad (8.8)$$

lo cual quiere decir (teniendo en cuenta que $v = c/\lambda$) que existe una frecuencia mínima o **frecuencia de corte**,

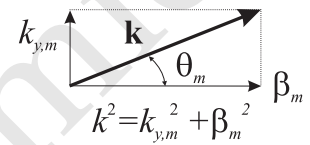
$$v_{\text{corte}} = \frac{c}{2d}, \quad (8.9)$$

por debajo de la cual no existe propagación en la guía.

2. Número máximo de modos.

Para una determinada longitud de onda, $\lambda < \lambda_{\max}$, la expresión (8.7) impone una restricción al número de modos que pueden propagarse por la guía. En concreto, el número máximo de modos, M , vendrá dado por

$$M = \text{Ent} \left(\frac{2d}{\lambda} \right), \quad (8.10)$$



Frecuencia de corte en una guía de placas paralelas

donde $\text{Ent}(x)$ es una función que nos da la parte entera del número real x [$\text{Ent}(0.345) = 0$, $\text{Ent}(8.94) = 8$].

EJEMPLO 8.1 Para una guía de placas paralelas con $d = 1$ cm, determine (a) la frecuencia de corte y (b) el número máximo de modos que puede propagarse a una frecuencia de 610 GHz.

(a): La longitud de onda máxima de una posible onda propagativa en la presente guía de placas paralelas será

$$\lambda_{\max} = 2 \times 1 = 2 \text{ cm},$$

por lo que la frecuencia de corte vendrá dada por

$$v_{\text{corte}} = \frac{c}{\lambda_{\max}} = \frac{3 \times 10^8}{2 \times 10^{-2}} = 1.5 \times 10^{10} \text{ Hz}.$$

Hemos obtenido, por tanto, que la guía de placas paralelas no transmite por debajo de una frecuencia de 15 GHz.

(b): A $v = 610$ GHz, la longitud de onda es

$$\lambda = \frac{3 \times 10^8}{6.1 \times 10^{11}} = 4.91 \times 10^{-4} \text{ m},$$

lo que implica que se propagarán como máximo M modos, siendo

$$M = \text{Parte entera} \left(\frac{2d}{\lambda} \right) = \text{Parte entera} \left(\frac{2 \times 10^{-2}}{4.91 \times 10^{-4}} \right) = 40.$$

8.2.1. Velocidad de grupo

De acuerdo con la discusión en el Apartado 1.8, debemos notar que, usualmente, las señales luminosas que se propagan por la guía no serán puramente monocromáticas sino que presentan cierta anchura de banda en frecuencia, $\Delta\omega$, o, equivalente, en número de ondas, $\Delta\beta$. La superposición de todas las componentes anteriores daba lugar a un *grupo de ondas* al que le podemos definir una velocidad de grupo, v_m , asociada a cada uno de los m modos definidos en (8.3). A partir de la expresión (1.90) encontraremos entonces que la velocidad de grupo asociada al modo m -ésimo vendrá dada por

$$v_m = \frac{d\omega}{d\beta_m} \quad (8.11)$$

(debemos entender que las constantes de propagación de las componentes del grupo de ondas están centradas alrededor de β_m). Haciendo uso de la expresión (8.4), también conocida como **relación de dispersión**, que puede escribirse como

$$\omega^2 = c^2 \beta_m^2 - c^2 \frac{m^2 \pi^2}{d^2}, \quad (8.12)$$

al operar para obtener (8.11) encontramos que

$$2\omega \frac{d\omega}{d\beta_m} = 2c^2 \beta_m \Rightarrow \frac{d\omega}{d\beta_m} = \frac{c}{\omega} c \beta_m.$$

Teniendo en cuenta que $\beta_m = k \cos \theta_m$ y que $k = \omega/c$, la velocidad de grupo puede finalmente escribirse como

$$v_m = c \cos \theta_m. \quad (8.13)$$

Esta expresión nos indica que la velocidad de grupo asociada a cada modo es distinta, siendo menor a medida que θ_m crece (esto es, para modos de mayor índice m).

8.2.2. Propagación multimodo

En las discusiones anteriores podría dar la impresión de que sólo aquellas distribuciones luminosas que sean modos pueden propagarse por la guía de placas paralelas. No obstante, en la práctica encontramos que cualquier distribución luminosa arbitraria cuyo campo eléctrico se anule en las placas se propaga a través de la guía. En general, esta distribución luminosa no mantendrá constante su perfil transversal sino que éste puede distorsionarse al propagarse. Este hecho puede entenderse si consideramos que cualquier distribución luminosa polarizada a lo largo de la dirección x que se propague en la guía tendrá un campo eléctrico, $E_x(y, z)$, que puede expresarse como la siguiente superposición de modos:

$$E_x(y, z) = \sum_{m=1}^M E_{0m} \sin\left(\frac{m\pi}{d} y\right) \exp(-j\beta_m z) \quad (8.14)$$

(ver Fig. 8.3). Podemos interpretar este hecho diciendo que la potencia óptica se divide en diferentes grupos asociados a los distintos modos propagativos. Cada uno de estos grupos de ondas se propaga *sin distorsión* a lo largo de la guía con su correspondiente velocidad de grupo (8.13), pero no así su superposición. A este fenómeno lo denominaremos **dispersión modal**.

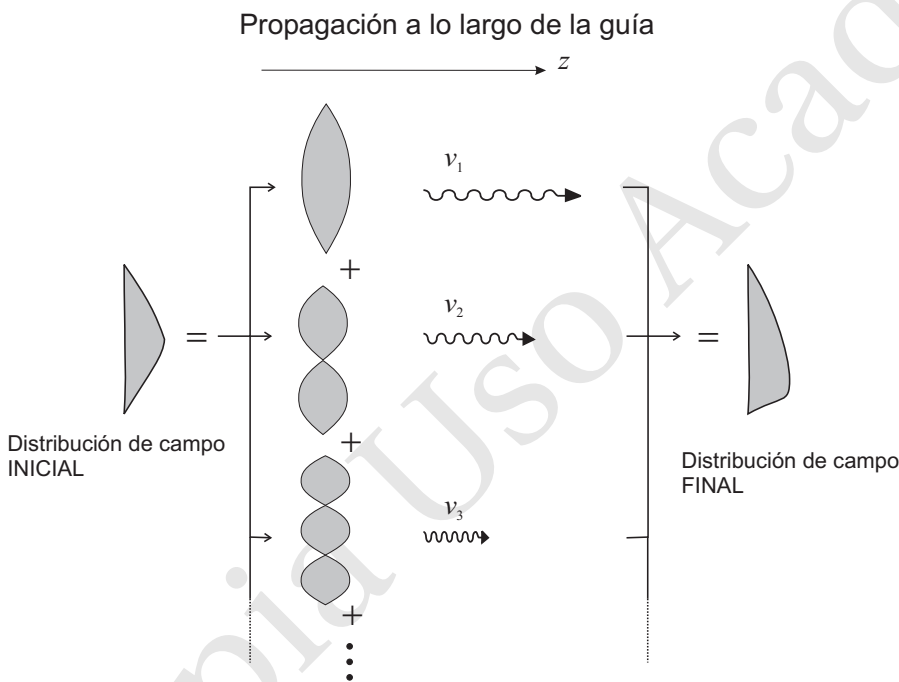


FIGURA 8.3: Debido a la distinta velocidad de cada modo, la distribución inicial de campo se distorsiona cuando ésta viaja a lo largo de la guía.

Si todos los modos viajasen con la misma velocidad de grupo, encontraríamos que la señal a la salida es exactamente la misma que la señal a la entrada. No obstante si cada modo viaja a una velocidad distinta (como ocurre en el presente caso), entonces la señal se irá distorsionando al propagarse y puede ocurrir que obtengamos a la salida del sistema de transmisión una señal muy distinta de la que teníamos a la entrada. Una posibilidad para evitar la dispersión modal es claramente trabajar a una longitud de onda tal que por la guía sólo se propague un único modo; en cuyo caso diremos que tenemos una guía **monomodo**. En el presente caso la guía será monomodo para aquellos valores de frecuencia com-

prendidos entre la frecuencia de corte del primer y el segundo mode, esto es, si

$$\frac{c}{2d} < \nu < \frac{c}{d}. \tag{8.15}$$

En el caso de la guía analizada en el Ejemplo 8.1 encontramos que dicha guía es monomodo en el rango $15 < \nu(\text{GHz}) < 30$.

8.3. Guía dieléctrica plana

En el Apartado anterior hemos visto que el guiado de la luz en la guía de placas paralelas se producía debido a un efecto de reflexión en las placas metálicas (espejos). Basado en este mismo principio básico podemos pensar en un sistema de transmisión efectivo para la luz basado en el fenómeno de **reflexión interna total**. Este sistema de transmisión podría ser la guía dieléctrica plana mostrada en la Fig. 8.4.

Tal como vimos en el Apartado 3.3, la ley de refracción de Snell ($n_1 \text{ sen } \phi_1 = n_2 \text{ sen } \phi_2$) nos dice que existirá reflexión interna total, si $n_1 > n_2$, para aquellos ángulos $\phi_1 > \phi_c$, siendo

$$\text{sen } \phi_c = \frac{n_2}{n_1}.$$

Con respecto a la Fig. 8.4, podemos observar que una onda plana caracteriza-

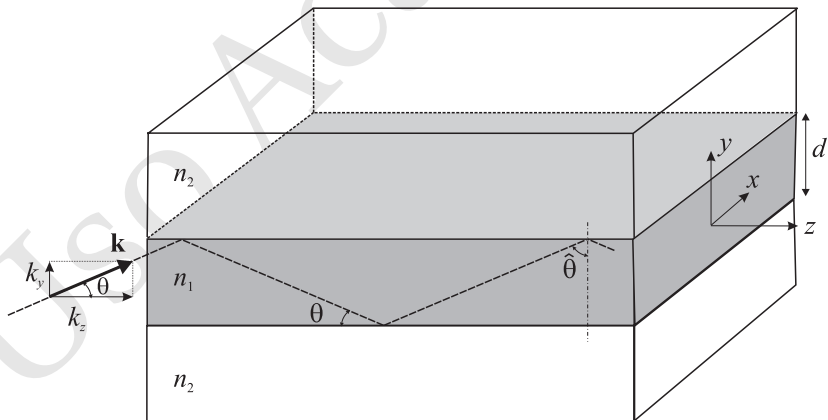
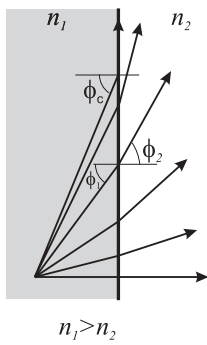


FIGURA 8.4: Propagación a lo largo de una guía dieléctrica plana. En las interfaces entre los medios 1 y 2 se produce una reflexión interna total.

da por un vector de ondas

$$\mathbf{k} = n_1 k_0 \cos \theta \hat{\mathbf{y}} + n_1 k_0 \text{sen } \theta \hat{\mathbf{z}} \tag{8.16}$$

($k_0 = 2\pi/\lambda_0$ es el número de ondas del vacío) y polarizada a lo largo de la dirección x (al igual que en el Apartado anterior) sufre reflexión interna total, y por consiguiente, se propagará en la guía dieléctrica plana siempre que

$$\theta < \hat{\theta}_c, \tag{8.17}$$

donde $\hat{\theta}_c$ es el ángulo complementario de θ_c ($\hat{\theta}_c = \pi/2 - \theta_c$), que viene dado por

$$\hat{\theta}_c = \text{arc cos} \left(\frac{n_2}{n_1} \right). \tag{8.18}$$

Análogamente a lo que ocurría en la guía de placas paralelas, sólo existe un número discreto y finito de valores de $k_y = n_1 k_0 \sin \theta$ para los cuales la orientación del vector de ondas dado da lugar a distribuciones de luz que se propagan sin distorsión a lo largo de la guía; esto es, a modos. Cada uno de estos modos está caracterizado por cierto ángulo θ_m , aunque desafortunadamente (y al contrario de lo que vimos en (8.6)) no existe en el presente caso una expresión cerrada para este ángulo. No obstante, este hecho no impide que podamos establecer una analogía total entre los fenómenos de propagación en la guía dieléctrica y los que ocurrían en la guía de placas paralelas. La diferencia principal con el caso anterior estará en el distinto valor de los ángulos θ_m (comprendidos ahora entre 0 y $\hat{\theta}_c$), que dará lugar a las siguientes constantes de propagación:

$$\beta_m = n_1 k_0 \cos \theta_m \quad (8.19)$$

y a unas distribuciones de campo eléctrico correspondientes a cada modo que *no se anulan* en las interfases dieléctricas.

La forma de los modos con polarización a lo largo del eje x de la guía dieléctrica serán similares a los de la guía de placas paralelas aunque debemos considerar que ahora los modos no serán ondas estacionarias en la dirección vertical sino más bien ondas “cuasiestacionarias” que se propagan en la dirección z con cierta constante de propagación β_m . La razón de ello estriba en el hecho de que los modos de la guía dieléctrica plana no tienen por qué presentar *nodos* en $y = 0$ e $y = d$ (es decir, en las interfases entre capas dieléctricas) sino más bien un campo que decaiga hasta cero en las proximidades de las interfases. Otra diferencia importante con el caso de la guía de placas metálicas paralelas es que *no existe frecuencia de corte en la guía dieléctrica plana*; es decir, existe un modo caracterizado por $m = 0$ que puede propagarse desde frecuencia cero y que se denomina **modo fundamental**. Este hecho está igualmente relacionado con la no existencia de nodos en las interfases, lo cual nos libera de la exigencia de que la longitud de onda de la onda cuasiestacionaria en la dirección vertical tenga que ser aproximadamente menor que la anchura d de la capa de dieléctrico con $n = n_1$.

No existe frecuencia de corte en la guía dieléctrica plana

Número máximo de modos

Un estudio que no será aquí desarrollado nos proporciona la siguiente expresión para el número máximo de modos, M , que puede propagarse por una guía dieléctrica de anchura d :

$$M = \text{Ent} \left(\frac{\text{sen } \hat{\theta}_c}{\lambda / 2d} \right) + 1. \quad (8.20)$$

Teniendo en cuenta que

$$\text{sen } \hat{\theta}_c = \sqrt{1 - \left(\frac{n_2}{n_1} \right)^2} = \frac{\sqrt{n_1^2 - n_2^2}}{n_1},$$

podemos reescribir (8.20) como

$$M = \text{Ent} \left(\frac{2d}{\lambda_0} \text{AN} \right) + 1, \quad (8.21)$$

Número máximo de modos en una guía dieléctrica plana

donde el coeficiente AN se denomina **apertura numérica** y se define como

$$\text{AN} = \sqrt{n_1^2 - n_2^2}. \quad (8.22)$$

Apertura Numérica

Si las características de la guía dieléctrica y la frecuencia de trabajo son tales que

$$\frac{\lambda}{2d} > \text{sen } \hat{\theta}_c$$

o, equivalentemente,

$$\frac{2d}{\lambda_0} AN < 1, \quad (8.23)$$

entonces estamos en una condición en la que $M = 1$; es decir, se propagará un único modo por la guía. Esta guía se denomina **guía monomodo**.

EJEMPLO 8.2 Calcule la frecuencia máxima de trabajo para que una guía dieléctrica plana con $n_1 = 1.47$, $n_2 = 1.46$ y $d = 100 \mu\text{m}$ sea una guía monomodo.

Según la expresión (8.23) y teniendo en cuenta que $\lambda_0 = c/\nu$, la frecuencia máxima, ν_{max} , para la que la guía es monomodo será

$$\frac{2d\nu_{\text{max}}}{c} AN = 1,$$

esto es,

$$\nu_{\text{max}} = \frac{c}{2dAN}.$$

Dado que para el presente caso tenemos que

$$AN = \sqrt{n_1^2 - n_2^2} = 0.171,$$

obtenemos finalmente que

$$\nu_{\text{max}} = \frac{3 \times 10^8}{2 \cdot 100 \times 10^{-6} \cdot 0.171} = 8.77 \times 10^{12} \text{ Hz} = 8.77 \text{ THz}.$$

En consecuencia, por debajo de 8.77 THz, solo se propaga un modo por esta guía.

8.4. Fibra de salto de índice

Una fibra óptica de salto de índice es una guía dieléctrica *cilíndrica* caracterizada por los índices de refracción de su núcleo y cubierta, n_1 y n_2 , de radios a y b respectivamente (ver Fig. 8.5). Usualmente las fibras ópticas se fabrican con

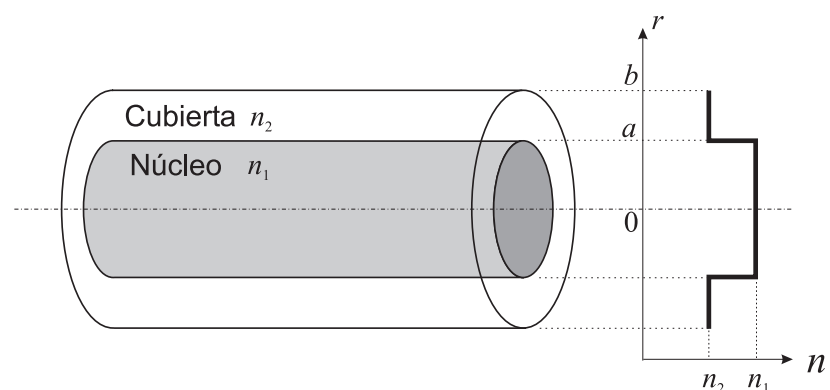


FIGURA 8.5: Geometría de una fibra de salto de índice.

vidrios de óxido de silicio (SiO_2) muy puros cuyo índice de refracción suele estar

entre 1.44 y 1.46. Las diferencias en el índice de refracción del núcleo y cubierta se consiguen mediante la adición de pequeñas concentraciones de impurezas. De este modo se consiguen pequeñas diferencias entre índices de refracción, por lo que definiremos el siguiente parámetro:

$$\Delta = \frac{n_1 - n_2}{n_1}, \quad (8.24)$$

que, en la práctica, suele tener valores comprendidos entre 0.001 y 0.002 (esta pequeña variación es suficiente para conseguir reflexión interna total). Tamaños típicos de los diámetros de los núcleos y cubiertas, $2a/2b$, son (expresados en μm) 8/125, 50/125, 62.5/125, 85/125 y 100/140.

8.4.1. Guiado de rayos

Al igual que en la guía dieléctrica plana, el guiado de la luz en la fibra óptica se realiza mediante el mecanismo de reflexión interna total. Si nos fijamos en un plano que contenga al eje de la fibra óptica (ver Fig.8.6), observamos que son guiados aquellos rayos cuyo ángulo θ verifique que $\theta < \hat{\theta}_c$, siendo $\hat{\theta}_c = \pi/2 - \theta_c$ el

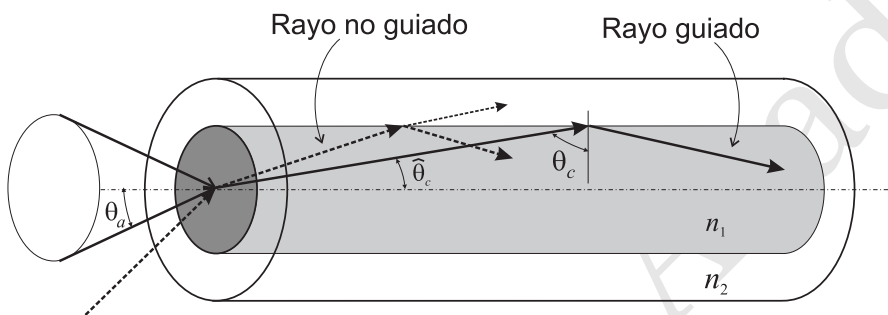


FIGURA 8.6: Los rayos situados dentro del cono de ángulo θ_a (cono de captación) serán guiados por la fibra óptica.

complementario del ángulo crítico, θ_c , para el que se verifica

$$\theta_c = \arcsen\left(\frac{n_2}{n_1}\right). \quad (8.25)$$

Si queremos obtener el ángulo máximo, θ_a , de un rayo que viniendo desde el exterior ($n = 1$) es guiado por la fibra (esto es, el ángulo del *cono de captación*), consideremos que en la interfase aire-fibra tendremos que

$$\text{sen } \theta_a = n_1 \text{sen } \hat{\theta}_c, \quad (8.26)$$

o, equivalentemente (considerando que $\text{sen } \hat{\theta}_c = \text{cos } \theta_c$),

$$\text{sen } \theta_a = n_1 \text{cos } \theta_c = n_1 \sqrt{1 - \text{sen } \theta_c^2}. \quad (8.27)$$

Teniendo en cuenta (8.25), la expresión anterior puede reescribirse como

$$\text{sen } \theta_a = n_1 \sqrt{1 - \frac{n_2^2}{n_1^2}} = \sqrt{n_1^2 - n_2^2}, \quad (8.28)$$

lo que nos dice, considerando la definición de apertura numérica (AN) dada en (8.22), que

Ángulo del cono de captación

$$\theta_a = \arcsen(\text{AN}) . \quad (8.29)$$

Dado que usualmente n_1 es un valor superior pero muy próximo a n_2 ($n_1 + n_2 \approx 2n_1$), encontramos que la apertura numérica de la fibra puede también escribirse como

$$\text{AN} = \sqrt{n_1^2 - n_2^2} = \sqrt{n_1 \Delta (n_1 + n_2)} \approx \sqrt{2n_1^2 \Delta} , \quad (8.30)$$

por lo que obtenemos finalmente que

$$\text{AN} \approx n_1 \sqrt{2\Delta} . \quad (8.31)$$

Claramente, aperturas numéricas grandes indican un gran cono de captación en la fibra considerada y viceversa.

EJEMPLO 8.3 Calcule la apertura numérica y el ángulo máximo de captación para una fibra óptica con $n_1 = 1.47$ y $n_2 = 1.46$.

Para el presente caso encontraremos que

$$\Delta = \frac{n_1 - n_2}{n_1} = \frac{0.01}{1.47} = 0.0068 ,$$

por lo que la apertura numérica será

$$\text{AN} = \text{sen } \theta_a \approx n_1 \sqrt{2\Delta} \approx 1.47 \sqrt{2 \times 0.0068} = 0.1715 .$$

El ángulo máximo del rayo incidente que será guiado por esta fibra será

$$\theta_a = \arcsen(\text{AN}) = \arcsen(0.1715) = 9.87^\circ .$$

Observemos que diferencias mayores entre los índices de refracción del núcleo y la cubierta incrementarían el ángulo de captación.

8.4.2. Número de modos

Aunque el procedimiento básico de guiado en la fibra óptica de salto de índice es el explicado en el apartado anterior, los detalles relativos al número y características de los distintos modos de propagación sólo se obtendrían tras la elaborada resolución de las ecuaciones de Maxwell. Dado que este cálculo queda fuera del objetivo y alcance del presente tema, a continuación expondremos simplemente algunos de los resultados más relevantes que se obtienen.

Debido a que en la mayoría de las fibras se cumple que $n_1 \approx n_2$ (o, equivalentemente, que $\Delta \ll 1$) tenemos que los rayos guiados son paraxiales; esto es, son aproximadamente paralelos al eje de la fibra. En estas situaciones encontramos que las ondas luminosas guiadas sin distorsión son muy aproximadamente *ondas transversales* que pueden presentarse en dos estados de polarización lineal independientes y que forman un conjunto de modos caracterizados por el par de números naturales l y m denotados usualmente como modos LP_{lm} (l y m pueden identificarse con el número de oscilaciones de la onda estacionaria en la dirección radial y azimutal respectivamente y LP proviene de "Linearly Polarized"). El modo fundamental de la fibra (es decir, aquel que se propaga desde frecuencia cero) es el LP_{01} .

En el estudio del número de modos que se propagan a una determinada frecuencia (o longitud de onda) aparece un parámetro muy importante conocido como **parámetro de la fibra** o parámetro V que viene dado por

Apertura Numérica de una fibra óptica de salto de índice

Parámetro de la fibra

$$V = 2\pi \frac{a}{\lambda_0} AN . \quad (8.32)$$

Este parámetro nos determina, por ejemplo, la siguiente condición de **fibra monomodo**:

$$V < 2.405 , \quad (8.33)$$

Condición de fibra monomodo

que nos da aquella condición para la que empezaría a propagarse un segundo modo (el LP₁₁) además del fundamental. La expresión (8.33) nos permite deducir fácilmente la frecuencia, ν_{11} , a partir de la cual empezaría a propagarse el modo LP₁₁. Esta frecuencia verificará

$$2\pi \frac{a\nu_{11}}{c} AN = 2.405 ,$$

por lo que

$$\nu_{11} = 2.405 \frac{c}{2\pi aAN} . \quad (8.34)$$

Frecuencia de corte modo LP₁₁

La expresión anterior nos indica claramente que ν_{11} es inversamente proporcional al radio del núcleo de la fibra y al valor de la apertura numérica. En consecuencia, núcleos pequeños y aperturas numéricas bajas aumentan considerablemente el rango de frecuencias en el que la fibra es monomodo. En general puede decirse que existen multitud de ventajas en el uso de fibras monomodo ya que, como veremos más adelante, reducen la dispersión y presentan bajos valores de atenuación.

El análisis de los resultados de las ecuaciones de Maxwell nos permite también encontrar una expresión bastante aproximada para calcular el número de modos que se propagan por fibras de salto de índice con un parámetro V alto. Bajo este último supuesto, encontraríamos que el número máximo de modos, M , viene dado por la siguiente expresión:

$$M \approx \frac{4}{\pi^2} V^2 . \quad (8.35)$$

Número de modos propagativos en una fibra de salto de índice con $V \gg 1$

EJEMPLO 8.4 Calcule el rango en frecuencias para el que la fibra del Ejemplo 8.3 con $a = 100 \mu\text{m}$ es monomodo. Determine asimismo la frecuencia para la cual se propagarían por la fibra al menos 400 modos.

Aplicando la expresión (8.34), encontramos que

$$\nu_{11} = 2.405 \frac{3 \times 10^8}{2\pi \times 100 \times 10^{-6} \times 0.1715} = 6.69 \times 10^{12} \text{ Hz} .$$

Por tanto para frecuencias $\nu \lesssim 6.7 \text{ THz}$ la fibra es monomodo.

Si queremos que por la fibra se propagan 400 modos, al aplicar (8.35) obtendremos que

$$400 \approx \frac{4}{\pi^2} V^2 \Rightarrow V \approx \frac{\pi}{2} \sqrt{400} \approx 31.41 ,$$

que corresponde a una frecuencia de

$$\begin{aligned} \nu &= V \frac{c}{2\pi aAN} \\ &= 31.41 \frac{3 \times 10^8}{2\pi \times 100 \times 10^{-6} \times 0.1715} = 8.74 \times 10^{13} \text{ Hz} , \end{aligned}$$

por lo que a una frecuencia de aproximadamente 87.4 THz tendríamos al menos 400 modos propagándose.

Para fibras con $V \gg 1$ podemos obtener igualmente una expresión aproximada para las constantes de propagación, β_{lm} , de los distintos modos LP_{lm} . Bajo el supuesto de $\Delta \ll 1$ encontramos

Constantes de propagación
($V \gg 1$)

$$\beta_{lm} \approx n_1 k_0 \left[1 - \frac{(l+2m)^2}{M} \Delta \right] \quad (8.36)$$

$$l = 0, 1, \dots, \sqrt{M}$$

$$m = 1, 2, \dots, (\sqrt{M} - 1)/2 .$$

Para la velocidad de grupo asociada a los distintos modos, $v_{lm} = d\omega/d\beta_{lm}$, podemos encontrar la siguiente expresión aproximada:

Velocidad de grupo ($V \gg 1$)

$$v_{lm} \approx c_1 \left[1 - \frac{(l+2m)^2}{M} \Delta \right], \quad (8.37)$$

siendo $c_1 = c/n_1$.

Observamos que existe *dispersión modal* dado que la velocidad de grupo asociada con cada modo es distinta. Podemos comprobar que para $M \gg 1$, la variación de las velocidades de grupo del modo más rápido ($l+2m=2$) y el más lento ($l+2m=\sqrt{M}$) viene dada aproximadamente por

$$v_{\max} - v_{\min} \approx c_1 \Delta . \quad (8.38)$$

Esto implica que si quisiéramos fibras con valores altos de AN (es decir, según (8.29), amplios conos de captación de luz), esto también llevaría aparejado diferencias importantes entre las velocidades de grupo de los distintos modos que se propagan por la guía.

8.5. Fibra de gradiente de índice

Una posible manera de obtener simultáneamente altas capacidades de captación luminosa junto con una pequeña dispersión modal la proporciona el uso de la fibra de gradiente de índice. Tal como muestra la Fig. 8.7, el índice de refracción

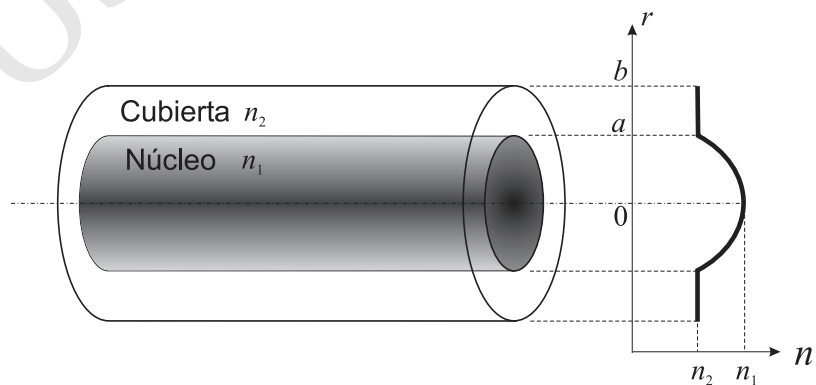


FIGURA 8.7: Geometría de una fibra de gradiente de índice.

del núcleo presenta una variación radial, $n_1 = n_1(r)$, siendo máximo en el eje para tomar el valor del índice de refracción de la cubierta en $r = a$. La existencia de un índice de propagación inhomogéneo provoca que la propagación de los rayos en el núcleo no sea en línea recta sino que éstos mostrarán una trayectoria curvada (similar a la estudiada en el Apartado 3.2.4). La comparación de estas trayectorias con las que ocurren en fibras de salto de índice monomodo y multimodo se

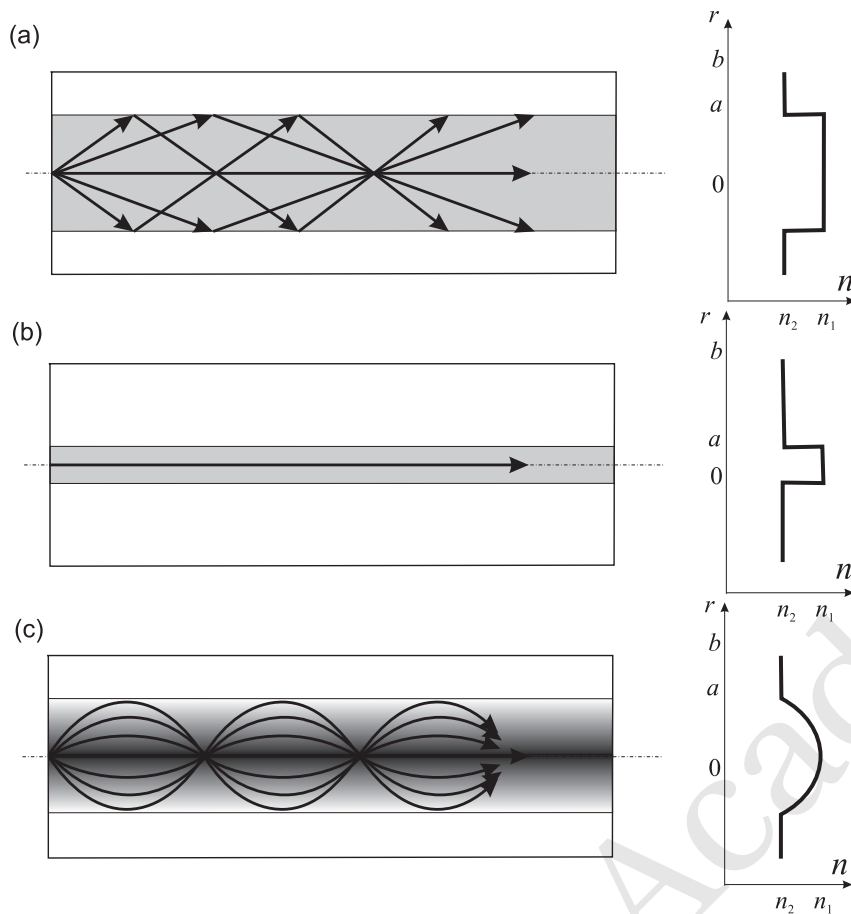


FIGURA 8.8: Trayectorias de los rayos en (a) fibra salto de índice multimodo, (b) fibra salto de índice monomodo, y (c) fibra gradiente de índice.

muestran en la Fig. 8.8. Debido a la especial característica de las trayectorias de los rayos en la fibra de gradiente de índice, las velocidades de grupo de los diferentes modos son más similares, mejorando así considerablemente el efecto de dispersión modal. En concreto, podemos encontrar que las diferencias entre la velocidad del modo más rápido respecto al más lento pueden alcanzar ahora el siguiente valor:

$$v_{\max} - v_{\min} \approx c_1 \frac{\Delta^2}{2}, \quad (8.39)$$

en vez de $c_1 \Delta$ tal como encontrábamos para las fibras de salto de índice. En consecuencia, se reduce la diferencia relativa en un factor $\Delta/2$.

De manera aproximada podemos encontrar igualmente que el número de modos en la fibra de gradiente de índice viene dado por

$$M \approx \frac{V^2}{4}. \quad (8.40)$$

Número máximo de modos en una fibra de gradiente de índice ($V \gg 1$)

Observamos que para el mismo valor de V se propagarían más modos por esta guía.

8.6. Atenuación

Cuando la luz viaja en las fibras, debido al carácter no ideal de los dieléctricos empleados en su fabricación, se observa una pérdida continua de potencia luminosa. A medida que la luz viaja en la fibra encontramos que la potencia luminosa a una distancia z del origen viene dada por

$$P(z) = P(0)e^{-\alpha z}, \quad (8.41)$$

es decir, existe una pérdida exponencial de potencia determinada por el parámetro α , conocido como **coeficiente de atenuación**. El factor de transmisión de potencia,

$$\mathcal{T}(z) = \frac{P(z)}{P(0)} = e^{-\alpha z}, \quad (8.42)$$

nos proporciona una medida del porcentaje de potencia pérdida por transmisión a lo largo de la guía. Para una fibra de longitud L , tomando logaritmos decimales en (8.42) y multiplicando por 10 (para obtener los valores en decibelios, dB), obtenemos que

$$10 \log_{10} \mathcal{T}(L) = 10(-\alpha L) \log_{10} e,$$

que puede reescribirse como

$$10\alpha \log_{10} e \equiv \hat{\alpha} = \frac{1}{L} 10 \log_{10} \frac{1}{\mathcal{T}(L)}. \quad (8.43)$$

Coeficiente de atenuación p.u.l.

Encontramos entonces que $\alpha \approx 0.23\hat{\alpha}$, siendo $\hat{\alpha}$ el coeficiente de atenuación en decibelios por unidad de longitud y que suele expresarse usualmente en dB/km (supuesto que L se exprese en km). Observemos que un valor de atenuación de 3 dB/km nos dice que la potencia luminosa se reduce a la mitad cada kilómetro; $\hat{\alpha} = 10, 20, 30, \dots$ dB/km nos indica que la potencia se reduce en un factor 10, 100, 1000... cada kilómetro. Cuando se empezaron a usar las fibras ópticas (~1970), las pérdidas eran tan importantes ($\hat{\alpha} > 20$ dB/km) que uso estaba restringido a comunicaciones a muy corta distancia. Afortunadamente, los desarrollos en la técnica de fabricación de materiales redujeron drásticamente las pérdidas, consiguiéndose actualmente valores de $\hat{\alpha} < 0.2$ dB/km, que ya sí permiten su uso en comunicaciones de larga distancia (por ejemplo, en los cables submarinos).

Las causas más importantes de la pérdida de energía luminosa en la fibra óptica son: absorción, difusión de Rayleigh y efectos extrínsecos. A continuación analizaremos cada causa por separado.

Absorción

Como cualquier otro material, el vidrio de silicio fundido del que suelen estar hechas las fibras ópticas, presentan ciertas bandas de absorción a determinadas longitudes de onda. En particular, el SiO₂ presenta una banda de absorción en el infrarrojo mediano (debido a transiciones energéticas relacionadas con vibraciones térmicas) y otra banda de absorción en el ultravioleta (debido a transiciones relacionadas con estados energéticos electrónicos y moleculares). El continuo avance en los materiales usados está permitiendo que las pérdidas debido a estas bandas de absorción sean cada vez menos significativas.

Difusión de Rayleigh

En cualquier material dieléctrico, el índice de refracción presenta pequeñas inhomogeneidades distribuidas al azar debido a irregularidades microscópicas.

Estas irregularidades microscópicas se comportan como centros difusores discretos cuando sobre este medio continuo incide una onda luminosa. Este fenómeno es análogo al que sucede cuando la luz atraviesa la atmósfera debido a la presencia de moléculas (centros difusores discretos) distribuidas al azar. En el Apartado 2.6 vimos que un dipolo (o un conjunto discreto de ellos, como las moléculas de la atmósfera) reirradiaba la potencia recibida según (2.91). Esta ley puede igualmente escribirse como

$$P(\lambda) \propto \lambda^{-4}. \quad (8.44)$$

En el caso de la fibra óptica, este fenómeno conocido como *difusión de Rayleigh*, da lugar a que la luz que atraviesa la fibra sea difundida por las inhomogeneidades en todas direcciones, afectando esta difusión mayormente a las longitudes de onda más cortas (ultravioleta). Todo este proceso es equivalente a que una parte muy significativa de la energía luminosa guiada por la fibra (aquella de menor longitud de onda) se “pierde” por difusión y no continúa siendo guiada por la fibra. En la práctica, este fenómeno provoca una de las mayores restricciones al guiado “ideal” de luz en la fibra, de modo que sólo para aquellas longitudes de onda para las que la difusión de Rayleigh no sea importante (alrededor del infrarrojo) es factible la transmisión de luz por la fibra. Este fenómeno es además prácticamente imposible de evitar usando “mejores” materiales, ya que está relacionado con el carácter “discreto” del material aunque éste sea modelado teóricamente como un medio dieléctrico continuo.

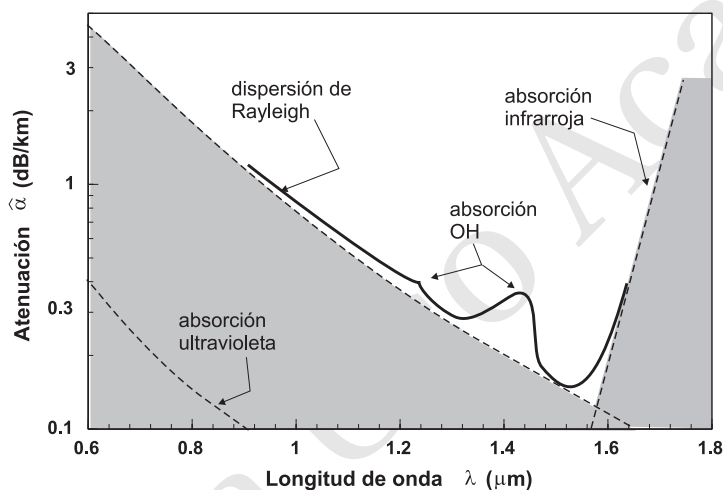


FIGURA 8.9: Dependencia del coeficiente de atenuación de una fibra óptica de SiO_2 con respecto a la longitud de onda. Obsérvese un mínimo local a $1.3 \mu\text{m}$ ($\hat{\alpha} \approx 0.3 \text{ dB/km}$) y un mínimo absoluto a $1.55 \mu\text{m}$ ($\hat{\alpha} \approx 0.16 \text{ dB/km}$)

Efectos extrínsecos

Además de los anteriores efectos intrínsecos, existen otros factores extrínsecos que también afectan a las pérdidas en la fibra. Entre ellos podríamos destacar los siguientes:

- *Presencia de impurezas OH.*

La presencia de este tipo de impurezas es muy difícil de eliminar en la práctica, provocando su presencia ciertas bandas de absorción que se evitan en las comunicaciones ópticas.

- *Curvas en la fibra.*

La existencia de pequeñas variaciones al azar en la geometría de la fibra y la presencia de tramos curvados de fibra da origen a que parte de la energía luminosa se escape de la fibra, dando así lugar a pérdidas de potencia.

8.7. Dispersión

Cuando un pequeño pulso luminoso se propaga a través de una fibra óptica se produce un “esparcimiento” del pulso en el tiempo, de modo que la “anchura temporal” del pulso aumenta distorsionando así la forma inicial del pulso. En este caso decimos que el pulso se ha *dispersado*. Este fenómeno limita considerablemente la capacidad de transmisión del canal puesto que hay que respetar una cierta “distancia temporal” entre pulsos. En la práctica este fenómeno junto con la atenuación limita muy seriamente la gran anchura de banda ideal de la fibra óptica ($\sim 10^{14}$ Hz), haciendo que las fibras ópticas actuales presenten anchos de banda únicamente del orden de una decena de veces superiores a los de los cables coaxiales.

Comúnmente reconocemos cuatro causas principales de dispersión en fibras: dispersión en el material, dispersión intermodal, dispersión de guía de ondas y dispersión no-lineal. A continuación analizaremos brevemente las anteriores causas.

Dispersión en el material

Debido al hecho de que el índice de refracción de los materiales presenta una cierta dependencia con la longitud de onda, $n = n(\lambda)$, teníamos según (1.93) que

$$v_g = v - \lambda \frac{dv}{d\lambda}.$$

Dado que $v = c/n$ encontramos que

$$v_g = \frac{c}{n} - \lambda \frac{d(c/n)}{d\lambda} = \frac{c}{n} - \lambda \left(-\frac{c}{n^2} \frac{dn}{d\lambda} \right),$$

o bien

$$v_g = \frac{c}{n} \left(1 + \frac{\lambda}{n} \frac{dn}{d\lambda} \right). \quad (8.45)$$

Dado que el pulso es un paquete de ondas, la expresión anterior nos indica que las diversas componentes del espectro con diferente longitud de onda viajarán a distinta velocidad de grupo, provocando así que el pulso se “ensanche”. Para ver esto mejor, reescribamos (8.45) como

$$v_g = \frac{c}{n_g}, \quad (8.46)$$

donde n_g es el *índice de refracción de grupo*, que se define como

$$n_g = \frac{n}{1 + \frac{\lambda}{n} \frac{dn}{d\lambda}} \approx n - \lambda \frac{dn}{d\lambda}, \quad (8.47)$$

usando la aproximación $1/(1+x) \approx 1-x$ para x pequeños. A partir de las expresiones anteriores podemos deducir que un pulso atravesaría la fibra de longitud L en un tiempo

$$\tau(\lambda) = \frac{Ln_g}{c} = \frac{L}{c} \left(n - \lambda \frac{dn}{d\lambda} \right), \quad (8.48)$$

donde $\tau(\lambda)$ es el tiempo de tránsito o **retraso de grupo**. Si usamos una fuente de luz con un ancho de línea $\Delta\lambda$ para producir el pulso que se propaga en la fibra, entonces la variación del retraso de grupo, $\Delta\tau$, fruto de la dispersión vendrá dada por

$$\Delta\tau = \left| \frac{d\tau}{d\lambda} \right| \Delta\lambda, \quad (8.49)$$

de donde obtenemos que

$$\frac{\Delta\tau}{\tau} = \frac{\lambda}{n} \frac{d^2 n}{d\lambda^2} \Delta\lambda, \quad (8.50)$$

debiéndose evaluar la derivada en la longitud de onda central del paquete.

Dispersión modal

Esta dispersión sucede en una fibra *multimodo* debido a la propagación simultánea de diversos modos con diferente velocidad de grupo. Si v_{\min} y v_{\max} son las velocidades de grupo mínimas y máximas, una estimación razonable del ensanchamiento del pulso puede ser

$$\sigma_T = \frac{1}{2} \left(\frac{L}{v_{\min}} - \frac{L}{v_{\max}} \right).$$

Dado que los valores de v_{\min} y v_{\max} han sido tratados para el caso de fibras de salto de índice y de gradiente de índice, obtenemos que

$$\sigma_T \approx \begin{cases} \frac{L}{c_1} \frac{\Delta}{2} & \text{fibra multimodo salto de índice,} \\ \frac{L}{c_1} \frac{\Delta^2}{4} & \text{fibra multimodo gradiente de índice.} \end{cases} \quad (8.51)$$

Claramente el ensanchamiento del pulso es menor en la fibra multimodo de gradiente de índice, lo que constituye una de las ventajas más importantes de este tipo de fibras.

Dispersión de guía de ondas

Incluso en ausencia de dispersión en el material, la velocidad de grupo de un modo concreto muestra una dependencia con la longitud de onda que provoca el consiguiente ensanchamiento temporal del pulso. Si en el caso de una fibra de salto de índice tenemos en cuenta la expresión (8.37) junto con (8.35) y consideramos la dependencia de M con la longitud de onda ($M \gg 1$), obtendríamos que

$$v_{lm} \approx c_1 \left[1 - \frac{(l+2m)^2 \Delta}{16a^2 \Delta N^2} \lambda_0^2 \right]. \quad (8.52)$$

Esta tipo de dispersión es especialmente importante en fibras monomodo dado que en estas fibras no existe dispersión modal y la dispersión por el material puede reducirse considerablemente.

Dispersión no-lineal

A medida que la intensidad luminosa crece aparecen fenómenos no lineales que provocan una alta dependencia de las propiedades del índice de refracción con la longitud de onda.

8.8. Comunicaciones ópticas

Hoy en día el uso más común de las fibras ópticas se encuentra en las comunicaciones a larga distancia. El proceso de comunicación mediante fibra óptica requiere evidentemente un primer paso que es la codificación de la información en una onda luminosa. Para este proceso hacen falta dos componentes: la fuente de luz (LED y láser tal como se discutió en el Apartado 7.5.1) y el correspondiente dispositivo de modulación. Posteriormente la señal luminosa debe introducirse en la fibra óptica para su consiguiente transmisión. Finalmente la información inscrita en la señal luminosa debe demodularse. Ilustremos este proceso con la transmisión de un mensaje telefónico digital (ver Fig. 7.21):

- La señal eléctrica de voltaje que sale del micrófono se muestrea a una frecuencia doble de la máxima que se quiere transmitir (de acuerdo al *teorema de Nyquist*) para asegurar que no se pierde información.
- El valor numérico de cada uno de los muestreos de la señal de voltaje se convierte en una secuencia en código binario de ocho dígitos, que a su vez se transmite mediante pulsos de luz. Estos pulsos de luz pueden obtenerse modulando convenientemente un diodo láser (ver Apartado 7.4.3). Dado que el tiempo entre muestreos es mucho mayor que la anchura de los pulsos binarios, podemos insertar varias conversaciones telefónicas en la misma fibra (*multiplexado en el tiempo*).
- Al final de la fibra, los pulsos luminosos se convierten de nuevo en pulsos eléctricos mediante un fotodiodo detector y posteriormente se separan las diferentes conversaciones (*demultiplexado*) para finalmente reconstruir la señal analógica de voltaje que se lleva al receptor telefónico.

Evidentemente el esquema anterior también sería válido para la transmisión de datos, imágenes de TV, etc...

Es interesante señalar que en la red telefónica se ha convenido en fijar la frecuencia máxima transmitida en 4 kHz, lo que hace que la frecuencia de muestreo deba ser de 8 kHz. Dado que cada muestreo requiere 8 bits, un línea telefónica supone la transmisión de 64 kbit/s. Los enlaces telefónicos de mayor orden (por ejemplo, los usados para conectar entre sí centrales de distintas ciudades) usados en Europa contienen 7620 líneas, lo que requiere un canal con una capacidad de transmisión de aproximadamente 500 Mbit/s. Esta capacidad de transmisión es soportada por un cable coaxial (cuya capacidad es de aproximadamente 1 Gbit/s) o por una fibra óptica. Mientras que es muy difícil superar la actual capacidad de transmisión de un cable coaxial, en principio podríamos esperar que la de la fibra óptica pudiera llegar hasta 10^{14} bit/s (usando una luz de $\lambda = 1 \mu\text{m}$); esto es, unas cien mil veces más que el cable coaxial. No obstante, esta suposición es bastante optimista ya que los fotodetectores actuales no son capaces de discernir variaciones menores de 10^{-10} s, lo que sitúa el límite efectivo de transmisión de un sistema óptico en 10 Gbit/s. Incluso esta cota es todavía optimista puesto que no se han considerado los efectos de dispersión que provocan el ensanchamiento del pulso y que, por tanto, reduce el número de pulsos que se pueden propagar por segundo. En este sentido, sólo si usamos fibras monomodo (que son las que presentan menor dispersión) alimentadas por diodos láser muy estables podríamos alcanzar en la práctica capacidades de transmisión de 10 Gbit/s. No obstante, dado el pequeño grosor de las fibras ópticas (del orden de un cabello humano), podemos introducir cientos de ellas en un tubo del grosor de un cable coaxial estándar. En consecuencia, para el mismo "grosor de cable" tenemos que el canal con fibras

ópticas incrementa la capacidad del mejor coaxial en un factor de 1000. Si además consideramos la posibilidad del *multiplexado en longitud de onda* (transmisión simultánea de varias señales a distintas longitudes de onda), tenemos que los sistemas ópticos de comunicaciones superan en varios millares de veces la capacidad de transmisión de los mejores sistemas eléctricos.

Otras ventajas adicionales de las fibras ópticas son

- Bajos niveles de pérdidas, lo que permite disminuir el número de estaciones amplificadoras y situarlas más distantes entre sí.
- Posibilidad de añadir impurezas a las fibras de modo que en ellas se produzca un efecto de amplificación de luz similar al que se produce en un medio láser activo. Esto permitiría reducir aún más el número de estaciones amplificadoras, además de la idoneidad que se obtiene al integrar el amplificador en el propio medio de transmisión.
- Insesibilidad respecto a las perturbaciones electromagnéticas. Debido a la ausencia de conductores metálicos, las fibras ópticas son prácticamente inmunes a las interferencias eléctricas y/o magnéticas.
- Ventajas mecánicas: las fibras ópticas son flexibles y pesan poco.

8.9. Problemas propuestos

8.1: Determinar la apertura numérica de una fibra de salto de índice sabiendo que el núcleo tiene un índice $n_1 = 1.62$ y el revestimiento $n_2 = 1.52$. ¿Cuál es el ángulo de aceptación cuando está sumergida en el aire? ¿Y si está sumergida en agua ($n = 1.33$)?

Sol. $AN = 0.314$, $\theta_a = 0.421$ rad.

8.2: Comparar las aperturas numéricas de una fibra de salto de índice con $n_1 = 1.45$ y $\Delta = 0.01$ y una fibra de gradiente de índice con $n_1 = 1.45$, $\Delta = 0.01$.

8.3: Una fibra de salto de índice tiene un radio $a = 5 \mu\text{m}$, índice de refracción en el núcleo $n_1 = 1.45$, e incremento de índice de refracción $\Delta = 0.002$. Determinar la longitud de onda menor, λ_c , para la cual la fibra actuaría como una guía de ondas monomodo.

Sol. $\lambda_c = 1.198 \mu\text{m}$.

8.4: Una fibra de salto de índice con índices de refracción $n_1 = 1.444$ y $n_2 = 1.443$ opera a una longitud de onda $\lambda_0 = 1.55 \mu\text{m}$. Determinar el radio del núcleo para el cual el parámetro V de la fibra vale 10.

Sol. $a = 45.91 \mu\text{m}$.

8.5: Calcular el tiempo de retraso, $\tau = L/v_g$, entre un rayo axial y otro que entra con el ángulo de aceptación en una fibra de salto de índice, de 1 km de longitud, sumergida en aire, sabiendo que los índices de refracción del núcleo y del revestimiento son $n_1 = 1.500$ y $n_2 = 1.485$.

Sol. $\tau_m = 50$ ns.

8.6: La dispersión material de una fibra monomodo viene descrita por

$$n(\lambda) = 1.45084 - 0.00334\lambda^2 + 0.00292/\lambda^2.$$

Si la longitud de la fibra es de 3 km y está acoplada a un LED de longitud de onda $1.3 \mu\text{m}$ y de anchura de línea $\Delta\lambda = 100$ nm, calcular la variación del retraso de grupo resultante de la dispersión material y de la dispersión de guía de onda.

Sol. $\tau_{\text{mat}} \approx 0.66$ ns, $\tau_{\text{gdo}} \approx 5$ ns.

8.7: Una fibra óptica multimodo de salto de índice tiene un núcleo de $100 \mu\text{m}$ de diámetro, una diferencia relativa de índices $\Delta = 1\%$, y se utiliza para transmitir luz monocromática de 800 nm. Suponiendo que el índice de refracción del núcleo es $n_1 = 1.46$, calcular la frecuencia normalizada de la fibra, V , el número de modos guiados y la apertura numérica si la fibra está sumergida en aire.

Sol. $AN = 0.206$, $V = 161.79$, $M \approx 13088$.

8.8: El núcleo de una fibra de gradiente de índice tiene un diámetro de $90 \mu\text{m}$ y un índice de refracción de perfil parabólico con un valor en el eje $n_1 = 1.46$. Si la apertura numérica de la fibra es 0.19 y la fibra opera con una longitud de onda de $1.3 \mu\text{m}$, calcular la frecuencia normalizada de la fibra, V , y el número de modos guiados.

Sol. $V = 82.65$, $M = 1708$.

8.9: Un LED típico que opera a $1.3 \mu\text{m}$ y presenta una anchura de línea de 100nm ilumina una fibra óptica "ideal". Calcule el ancho de banda en frecuencia $\Delta\nu$ correspondiente a la anterior anchura de línea. ¿Cuántos canales de vídeo podrían transmitirse por esta fibra si cada canal de video ocupa 10MHz ? ¿Cuántos canales de audio separados cada 10kHz podrían transmitirse?

Sol. $\Delta\nu \approx 1.8 \times 10^7 \text{MHz}$, $\sim 2 \times 10^6$ canales de vídeo, $\sim 2 \times 10^9$ canales de audio.

Copia Uso Académico

Apéndice A

Análisis vectorial

A.1. Vectores

En la naturaleza existen magnitudes físicas que están completamente determinadas por su valor y sus unidades. De forma genérica puede decirse que estas magnitudes son *escalares*. Ejemplos de estas magnitudes son la masa, la distancia, la temperatura, etc. Por el contrario, existen otras magnitudes que además de su valor y unidades están “dotadas” de una propiedad adicional: su dirección. Este tipo de magnitudes se conocen con el nombre de magnitudes *vectoriales* e incluyen a magnitudes tales como la posición, la velocidad, la fuerza, el campo eléctrico, etc. Para expresar las magnitudes vectoriales se hace uso de los *vectores* y por tanto se hace imprescindible el álgebra de vectores.

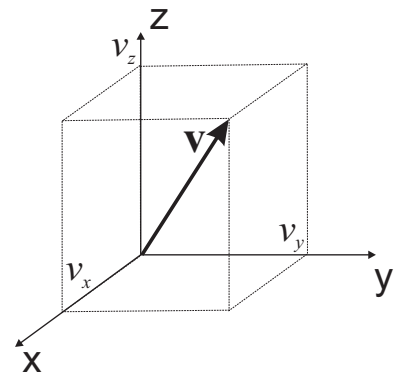
A.1.1. Notación vectorial

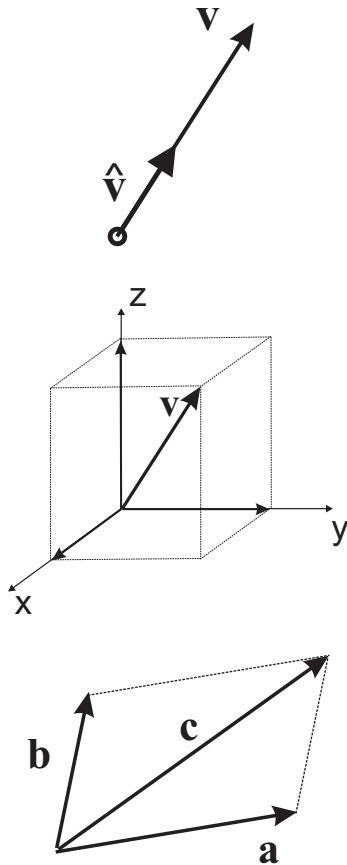
Las magnitudes vectoriales se denotarán de forma comprimida mediante letras minúsculas o mayúsculas en **negrita**, \mathbf{v} , \mathbf{V} , o bien con una flecha/raya encima de dichas letras, \vec{v} , \vec{V} . Para especificar dichos vectores se usan frecuentemente varios tipos de notación.

- Mediante una terna de números que son las componentes del vector en los ejes cartesianos x, y, z , esto es,

$$\mathbf{v} = (v_x, v_y, v_z). \quad (\text{A.1})$$

Geoméricamente, las componentes del vector son las proyecciones de este vector en los ejes cartesianos.





- El vector \mathbf{v} puede también expresarse en función de su módulo y de su **vector unitario**. El módulo del vector \mathbf{v} suele denotarse como v o bien $|\mathbf{v}|$ y viene dado según el teorema de Pitágoras por

$$|\mathbf{v}| \equiv v = \sqrt{v_x^2 + v_y^2 + v_z^2}. \quad (\text{A.2})$$

El vector unitario asociado con el vector \mathbf{v} se define como aquel vector de módulo unidad que tiene la misma dirección y sentido que \mathbf{v} . Dicho vector se denotará de forma genérica como $\hat{\mathbf{v}}$, esto es,

$$\hat{\mathbf{v}} = \frac{\mathbf{v}}{v} = \frac{(v_x, v_y, v_z)}{\sqrt{v_x^2 + v_y^2 + v_z^2}}. \quad (\text{A.3})$$

Obviamente el vector \mathbf{v} puede escribirse como: $\mathbf{v} = v\hat{\mathbf{v}}$.

- Expresando el vector como suma de las componentes del vector por los vectores unitarios a lo largo de los ejes coordenados. Los vectores unitarios a lo largo de los ejes x, y, z se denotaran como $\hat{\mathbf{x}}, \hat{\mathbf{y}}, \hat{\mathbf{z}}$ respectivamente. Otras notaciones frecuentes para estos vectores unitarios son $\mathbf{i}, \mathbf{j}, \mathbf{k}$ o bien $\mathbf{e}_x, \mathbf{e}_y, \mathbf{e}_z$. Usando esta notación, el vector \mathbf{v} se escribirá como:

$$\mathbf{v} = v_x\hat{\mathbf{x}} + v_y\hat{\mathbf{y}} + v_z\hat{\mathbf{z}}. \quad (\text{A.4})$$

A.1.2. Suma de vectores

La suma de vectores se realiza sumando sus componentes. De este modo si

$$\mathbf{a} = a_x\hat{\mathbf{x}} + a_y\hat{\mathbf{y}} + a_z\hat{\mathbf{z}}$$

$$\mathbf{b} = b_x\hat{\mathbf{x}} + b_y\hat{\mathbf{y}} + b_z\hat{\mathbf{z}},$$

el vector \mathbf{c} suma de los dos anteriores será por tanto:

$$\begin{aligned} \mathbf{c} &= \mathbf{a} + \mathbf{b} \\ &= (a_x + b_x)\hat{\mathbf{x}} + (a_y + b_y)\hat{\mathbf{y}} + (a_z + b_z)\hat{\mathbf{z}}. \end{aligned} \quad (\text{A.5})$$

A.1.3. Producto escalar

El producto escalar de dos vectores \mathbf{a} y \mathbf{b} , denotado como $\mathbf{a} \cdot \mathbf{b}$ es un *escalar* fruto de la siguiente operación:

$$\mathbf{a} \cdot \mathbf{b} = a_x b_x + a_y b_y + a_z b_z \quad (\text{A.6})$$

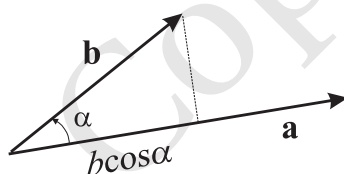
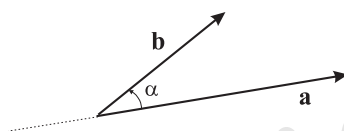
$$= ab \cos \alpha, \quad (\text{A.7})$$

siendo α el ángulo formado por los dos vectores (es independiente si este ángulo se mide en dirección horaria o antihoraria ya que $\cos(\pi - \alpha) = -\cos \alpha$). El producto escalar $\mathbf{a} \cdot \mathbf{b}$ puede interpretarse geoméricamente como la proyección de uno de los vectores sobre el otro (salvo factores numéricos). Este hecho se manifiesta claramente en el producto escalar de \mathbf{a} por uno de los vectores unitarios según los ejes coordenados, esto es,

$$\mathbf{a} \cdot \hat{\mathbf{x}} = a_x,$$

donde se ve claramente que $\mathbf{a} \cdot \hat{\mathbf{x}}$ es justamente la proyección de \mathbf{a} sobre el eje x .

Algunas de las propiedades del producto escalar son:



- El producto escalar es conmutativo:

$$\mathbf{a} \cdot \mathbf{b} = \mathbf{b} \cdot \mathbf{a}. \quad (\text{A.8})$$

- El producto escalar es distributivo respecto a la suma de vectores:

$$\mathbf{a} \cdot (\mathbf{b} + \mathbf{c}) = \mathbf{a} \cdot \mathbf{b} + \mathbf{a} \cdot \mathbf{c}. \quad (\text{A.9})$$

- El producto escalar de dos vectores perpendiculares es nulo:

$$\mathbf{a} \cdot \mathbf{b} = 0 \Rightarrow \mathbf{a} \perp \mathbf{b}. \quad (\text{A.10})$$

- El producto escalar de un vector por sí mismo es igual al cuadrado del módulo de dicho vector:

$$\mathbf{a} \cdot \mathbf{a} = a^2. \quad (\text{A.11})$$

A.1.4. Producto vectorial

El producto vectorial de dos vectores \mathbf{a} y \mathbf{b} , denotado como $\mathbf{a} \times \mathbf{b}$, es un *vector* definido como

$$\mathbf{a} \times \mathbf{b} = ab \sin \alpha \hat{\mathbf{n}}, \quad (\text{A.12})$$

siendo α el ángulo *más pequeño* formado por los dos vectores y $\hat{\mathbf{n}}$ el vector unitario normal exterior al plano que contiene a los vectores \mathbf{a} y \mathbf{b} . Puesto que el plano tiene dos normales (cada una con distinto sentido), el vector $\hat{\mathbf{n}}$ que aparece en (A.12) siempre se refiere a la normal que apunta según la *regla de la mano derecha*. Esta regla dice que usando la mano derecha y apuntando el dedo índice en la dirección de \mathbf{a} y el dedo corazón en la de \mathbf{b} , el dedo pulgar indicará la dirección de $\hat{\mathbf{n}}$ ¹. Geométricamente, el módulo del producto vectorial, $|\mathbf{a} \times \mathbf{b}|$, es igual al área del paralelogramo generado por los vectores \mathbf{a} y \mathbf{b} .

A partir de la definición del producto vectorial (A.12) pueden deducirse las siguientes propiedades:

- El producto vectorial es anticonmutativo:

$$\mathbf{a} \times \mathbf{b} = -\mathbf{b} \times \mathbf{a}. \quad (\text{A.13})$$

- El producto vectorial es distributivo respecto a la suma de vectores:

$$\mathbf{a} \times (\mathbf{b} + \mathbf{c}) = \mathbf{a} \times \mathbf{b} + \mathbf{a} \times \mathbf{c}. \quad (\text{A.14})$$

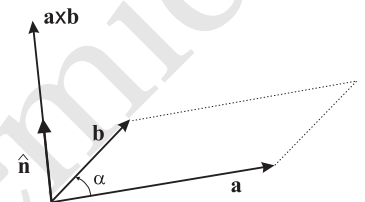
- El producto vectorial de dos vectores paralelos es nulo:

$$\mathbf{a} \times \mathbf{b} = 0 \Rightarrow \mathbf{a} \parallel \mathbf{b}. \quad (\text{A.15})$$

- Multiplicación por un escalar α :

$$\alpha(\mathbf{a} \times \mathbf{b}) = \alpha\mathbf{a} \times \mathbf{b} = \mathbf{a} \times \alpha\mathbf{b}. \quad (\text{A.16})$$

¹ Esta regla también se conoce a veces como *regla del tornillo* cuando dice que considerando el giro que va desde \mathbf{a} hasta \mathbf{b} por el camino más corto, si este giro se aplica a un tornillo, el sentido de avance o retroceso del tornillo indica hacia donde se dirige la normal.



Teniendo en cuenta la definición (A.12) y las propiedades (A.13)–(A.15), el producto vectorial de \mathbf{a} por \mathbf{b} puede obtenerse como

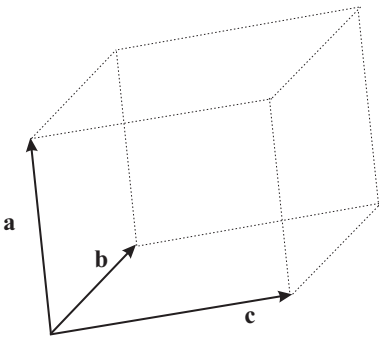
$$\begin{aligned}\mathbf{a} \times \mathbf{b} &= (a_x \hat{\mathbf{x}} + a_y \hat{\mathbf{y}} + a_z \hat{\mathbf{z}}) \times (b_x \hat{\mathbf{x}} + b_y \hat{\mathbf{y}} + b_z \hat{\mathbf{z}}) = \\ &= (a_y b_z - a_z b_y) \hat{\mathbf{x}} + (a_z b_x - a_x b_z) \hat{\mathbf{y}} + (a_x b_y - a_y b_x) \hat{\mathbf{z}}.\end{aligned}\quad (\text{A.17})$$

Usando la definición del determinante, la expresión anterior puede escribirse como

$$\mathbf{a} \times \mathbf{b} = \begin{vmatrix} \hat{\mathbf{x}} & \hat{\mathbf{y}} & \hat{\mathbf{z}} \\ a_x & a_y & a_z \\ b_x & b_y & b_z \end{vmatrix}.\quad (\text{A.18})$$

A.1.5. Productos triples

Dado que el producto vectorial de dos vectores es otro vector, este vector puede a su vez multiplicarse escalar o vectorialmente para formar lo que se conoce como productos triples.



- **Producto triple escalar:** $\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c})$.

Desde un punto de vista geométrico, este producto triple escalar puede interpretarse como el volumen del paralelepípedo generado por los tres vectores \mathbf{a} , \mathbf{b} y \mathbf{c} dado que según la figura adjunta $|\mathbf{b} \times \mathbf{c}|$ es el área de la base y $|a \cos \alpha|$ es la altura (α es el ángulo entre \mathbf{a} y $\mathbf{b} \times \mathbf{c}$). Usando esta interpretación geométrica es fácil deducir que

$$\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c}) = \mathbf{b} \cdot (\mathbf{c} \times \mathbf{a}) = \mathbf{c} \cdot (\mathbf{a} \times \mathbf{b}).\quad (\text{A.19})$$

Es interesante notar que en la expresión anterior se ha preservado el “orden alfabético”.

El producto triple escalar puede también obtenerse a partir del siguiente determinante:

$$\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c}) = \begin{vmatrix} a_x & a_y & a_z \\ b_x & b_y & b_z \\ c_x & c_y & c_z \end{vmatrix}.\quad (\text{A.20})$$

- **Producto triple vectorial:** $\mathbf{a} \times (\mathbf{b} \times \mathbf{c})$.

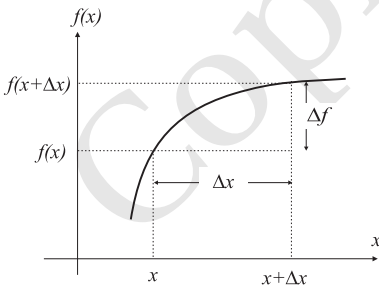
Este producto triple vectorial puede también obtenerse como

$$\mathbf{a} \times (\mathbf{b} \times \mathbf{c}) = \mathbf{b}(\mathbf{a} \cdot \mathbf{c}) - \mathbf{c}(\mathbf{a} \cdot \mathbf{b}).\quad (\text{A.21})$$

Nótese que el vector

$$(\mathbf{a} \times \mathbf{b}) \times \mathbf{c} = -\mathbf{c} \times (\mathbf{a} \times \mathbf{b}) = -\mathbf{a}(\mathbf{b} \cdot \mathbf{c}) + \mathbf{b}(\mathbf{a} \cdot \mathbf{c})\quad (\text{A.22})$$

es un vector completamente diferente al definido en la expresión (A.21).



A.1.6. Diferencial y derivada de funciones de una sola variable

Dada una función de una sola variable $f = f(x)$, se define la derivada de la función $f(x)$ con respecto a x como

$$\frac{df(x)}{dx} = \lim_{\Delta x \rightarrow 0} \frac{f(x + \Delta x) - f(x)}{\Delta x} = \lim_{\Delta x \rightarrow 0} \frac{\Delta f}{\Delta x}\quad (\text{A.23})$$

y expresa geoméricamente el valor de la pendiente de la tangente a la curva $f(x)$ en el punto x . El concepto de **diferencial** de $f(x)$, denotado genéricamente como df , expresa la variación infinitesimal de la función $f(x)$ entre x y $x + dx$, esto es,

$$df(x) = f(x + dx) - f(x). \quad (\text{A.24})$$

Desde un punto de vista matemático, este diferencial viene dado por el siguiente producto:

$$df(x) = \left(\frac{df}{dx} \right) dx. \quad (\text{A.25})$$

Debe notarse que df/dx **no** expresa un cociente entre df y dx sino que por el contrario debe entenderse como la acción del *operador* d/dx sobre la función $f(x)$. Este hecho se pone de manifiesto con otras notaciones que prefieren expresar la derivada de la función $f(x)$ con respecto a x como $D_x f(x)$, donde D_x es precisamente el operador derivada.

A.1.7. Teorema fundamental del cálculo

El teorema fundamental del cálculo establece la siguiente relación entre las operaciones de integración y diferenciación de la función $f(x)$:

$$\int_a^b \left(\frac{df}{dx} \right) dx = f(b) - f(a). \quad (\text{A.26})$$

Es posible “deducir” la expresión anterior teniendo en cuenta que $df(x) = \left(\frac{df}{dx} \right) dx$ y por tanto

$$\int_a^b df(x) = f(b) - f(a). \quad (\text{A.27})$$

A.1.8. Diferencial y derivada parcial de funciones de varias variables

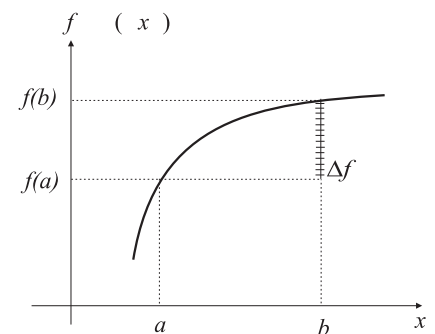
Es muy frecuente que en la naturaleza las magnitudes dependan de más de una variable, así la temperatura de una habitación depende de la posición del punto donde se mide, esto es, de las *tres* coordenadas espaciales del punto. Este hecho se manifiesta matemáticamente diciendo que la temperatura es función de x, y y z y se denota como $T = T(x, y, z)$.

Similarmente al concepto de derivada introducido en la sección anterior para funciones de una sola variable, puede ahora definirse el concepto de *derivada parcial*. Esta derivada hace referencia a la variación de cierta función con respecto a *una sola* de las variables cuando las demás permanecen constantes. Así, se define la derivada parcial de la función $f(x, y, z)$ con respecto a x como

$$\frac{\partial f}{\partial x} = \lim_{\Delta x \rightarrow 0} \frac{f(x + \Delta x, y, z) - f(x, y, z)}{\Delta x} \quad (\text{A.28})$$

y análogamente para las restantes variables. A partir del concepto de derivada parcial, puede deducirse que una variación infinitesimal de la función $f(x, y, z)$ cuando dicha función varía entre los puntos x y $x + dx$ podrá expresarse como:

$$df|_x = \left(\frac{\partial f}{\partial x} \right) dx. \quad (\text{A.29})$$



La variación infinitesimal de la función $f(x, y, z)$ cuando ésta varía entre los puntos (x, y, z) y $(x+dx, y+dy, z+dz)$ podría obtenerse, por tanto, sumando las variaciones parciales a lo largo de cada una de las coordenadas. De este modo, puede escribirse que

$$df = \left(\frac{\partial f}{\partial x}\right) dx + \left(\frac{\partial f}{\partial y}\right) dy + \left(\frac{\partial f}{\partial z}\right) dz. \quad (\text{A.30})$$

A.1.9. Operador gradiente

Es interesante notar en la expresión (A.30) que el diferencial de la función $f(x, y, z)$, df , puede expresarse como el siguiente producto escalar:

$$df = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z}\right) \cdot (dx, dy, dz). \quad (\text{A.31})$$

Definiendo el operador vectorial ∇ como

$$\nabla \equiv \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}\right) \quad (\text{A.32})$$

Operador ∇

$$\equiv \hat{x} \frac{\partial}{\partial x} + \hat{y} \frac{\partial}{\partial y} + \hat{z} \frac{\partial}{\partial z}, \quad (\text{A.33})$$

al aplicarlo a la función $f(x, y, z)$ se obtiene el **gradiente** de f , ∇f , que es evidentemente una magnitud *vectorial*:

$$\nabla f(x, y, z) = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z}\right) \quad (\text{A.34})$$

Definición de gradiente de f

$$= \frac{\partial f}{\partial x} \hat{x} + \frac{\partial f}{\partial y} \hat{y} + \frac{\partial f}{\partial z} \hat{z}. \quad (\text{A.35})$$

Esta definición permite escribir el diferencial de la función f como el siguiente producto escalar:

$$df = \nabla f \cdot d\mathbf{r}, \quad (\text{A.36})$$

donde $d\mathbf{r} = (dx, dy, dz)$.

Usando la definición de producto escalar, df también puede escribirse como

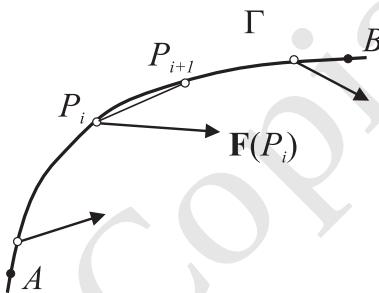
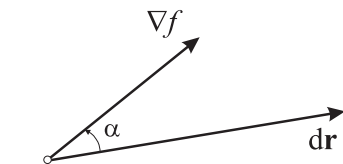
$$df = |\nabla f| |d\mathbf{r}| \cos \alpha, \quad (\text{A.37})$$

lo que permite deducir que la máxima variación de la función $f(x, y, z)$ se produce cuando $\alpha = 0$, esto es, cuando $d\mathbf{r}$ es paralelo al gradiente de f , ∇f . Consecuentemente, la dirección del vector ∇f marca la dirección de máxima variación de la función en el punto (x, y, z) .

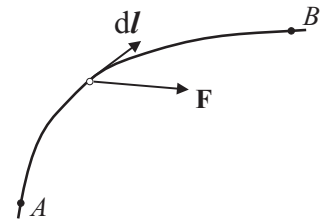
A.1.10. Integral de camino

Dado un campo vectorial \mathbf{F} (esto es, una magnitud *vectorial* cuyas componentes dependen de la posición espacial), la integral de camino de \mathbf{F} entre dos puntos A y B a lo largo de la curva Γ se define como la siguiente integral:

$$\begin{aligned} \mathcal{C}_A^B &= \int_{A, \Gamma}^B \mathbf{F} \cdot d\mathbf{l} \\ &= \lim_{P_{i+1} \rightarrow P_i(\Gamma)} \sum_i \mathbf{F}(P_i) \cdot \overline{P_i P_{i+1}}. \end{aligned} \quad (\text{A.38})$$



La integral anterior puede interpretarse como la superposición infinitesimal del producto escalar $\mathbf{F} \cdot d\mathbf{l}$ para cada elemento diferencial de la curva Γ entre los puntos A y B (El vector $d\mathbf{l}$ es un vector que tiene por módulo la longitud de un elemento diferencial de la curva y por dirección la de la tangente a la curva en dicho punto). Las integrales de camino son muy usuales en Física, definiendo, por ejemplo, el trabajo que realiza cierta fuerza entre dos puntos a través de cierta trayectoria. En general, *la integral de camino depende del camino que se elija para ir desde A hasta B .*



Algunas de las propiedades más importantes de las integrales de camino son:

- $\int_{A,\Gamma}^B \mathbf{F} \cdot d\mathbf{l} = - \int_{B,\Gamma}^A \mathbf{F} \cdot d\mathbf{l} .$
- Si A' es un punto intermedio de la curva Γ entre A y B , se tiene que $\int_{A,\Gamma}^B \mathbf{F} \cdot d\mathbf{l} = \int_{A,\Gamma}^{A'} \mathbf{F} \cdot d\mathbf{l} + \int_{A',\Gamma}^B \mathbf{F} \cdot d\mathbf{l} .$

A.1.11. Teorema fundamental del gradiente

De forma similar a como se hizo para funciones de una sola variable en (A.26), se verifica que

$$\int_A^B \nabla f \cdot d\mathbf{l} = f(B) - f(A) , \tag{A.39}$$

donde la integral en la expresión anterior es una *integral de camino*.

La expresión (A.39) puede “justificarse” considerando la definición del diferencial de f dada por (A.36). A partir de esta definición, la integral en (A.39) puede verse como una superposición infinitesimal de las variaciones de la función entre los puntos A y B , y esto es precisamente $f(B) - f(A)$.

Dos importantes corolarios se pueden extraer de la expresión (A.39)

1.

$$\int_{A,\Gamma}^B \nabla f \cdot d\mathbf{l} = \int_{A,\gamma}^B \nabla f \cdot d\mathbf{l} , \tag{A.40}$$

esto es, $\int_A^B \nabla f \cdot d\mathbf{l}$ es *independiente del camino* tomado entre los puntos A y B . Debe notarse que, en general, la integral de camino $\int_{A,\Gamma}^B \mathbf{F} \cdot d\mathbf{l}$ sí depende del camino (considérese, por ejemplo, el trabajo realizado por un coche para desplazarse entre dos puntos siguiendo distintas carreteras).

2.

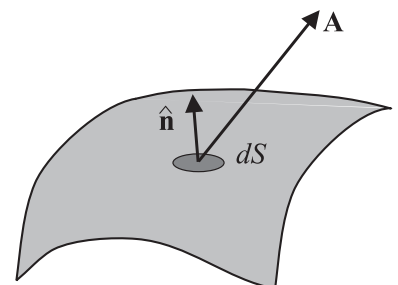
$$\oint_{\Gamma} \nabla f \cdot d\mathbf{l} = 0 . \tag{A.41}$$

La integral de camino anterior a través de cualquier curva cerrada, Γ , es nula.

A.2. Integral de flujo

Una integral muy útil que aparece en Física es la integral de flujo. El flujo de un campo vectorial \mathbf{A} a través de una superficie S se define como la siguiente integral de superficie:

$$\Phi = \int_S \mathbf{A} \cdot d\mathbf{S} , \tag{A.42}$$



donde S es una superficie arbitraria y $d\mathbf{S}$ es el vector *diferencial de superficie*, definido como

$$d\mathbf{S} = dS\hat{\mathbf{n}}, \quad (\text{A.43})$$

que tiene por módulo el área del elemento diferencial y por dirección y sentido el del vector unitario normal exterior a la superficie. Por ejemplo, para el caso del plano $z = \text{Cte}$, el diferencial de superficie será $d\mathbf{S} = dx dy \hat{\mathbf{z}}$.

Copia Uso Académico

Apéndice B

Funciones armónicas y Análisis fasorial

B.1. Funciones Armónicas

Una función armónica $f(t)$ es aquella que varía en el tiempo de la forma genérica:

$$f(t) = A \cos(\omega t + \varphi), \quad (\text{B.1})$$

donde A es la **amplitud**, ω la frecuencia angular y φ el desfase. La amplitud, A , determina el rango de variación de la señal, esto es

$$-A \leq f(t) \leq A.$$

La frecuencia angular está relacionada con la frecuencia f a través de

$$\omega = 2\pi f = \frac{2\pi}{T}, \quad (\text{B.2})$$

donde T el **periodo** de la señal, esto es, aquel valor tal que $f(t) = f(t + T)$. El desfase φ determina el origen del tiempo, esto es, cuál es el valor de la función en el instante $t = 0$:

$$f(0) = A \cos \varphi.$$

Es interesante recordar algunas relaciones trigonométricas de las funciones seno o coseno, a saber:

$$\operatorname{sen}(a \pm b) = \operatorname{sen} a \cos b \pm \cos a \operatorname{sen} b \quad (\text{B.3})$$

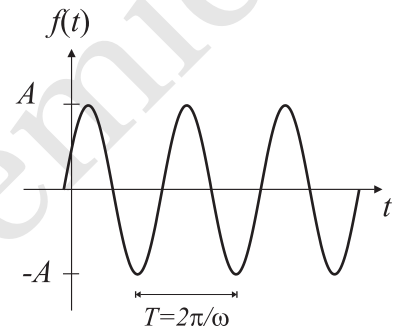
$$\operatorname{cos}(a \pm b) = \operatorname{cos} a \cos b \mp \operatorname{sen} a \operatorname{sen} b, \quad (\text{B.4})$$

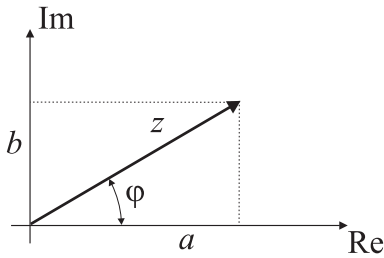
de donde puede deducirse, por ejemplo, que

$$\operatorname{cos}(\omega t + \varphi - \pi/2) = \operatorname{sen}(\omega t + \varphi). \quad (\text{B.5})$$

B.2. Análisis fasorial

En la resolución de ecuaciones de segundo grado, es frecuente encontrarse con soluciones que implican tomar la raíz cuadrada de un negativo, por ejemplo $\sqrt{-9}$. No obstante, es fácil notar que no existe ningún número *real* (esto es, que





pertenezca al conjunto \mathbb{R}) tal que su cuadrado sea -9 . Para solucionar esta cuestión se introducen los números *imaginarios*, que pueden formarse a partir de la definición de la **unidad imaginaria**, j :

$$j = \sqrt{-1}, \quad (\text{B.6})$$

de modo que

$$\sqrt{-9} = \sqrt{-1 \times 9} = \sqrt{-1} \times \sqrt{9} = j3.$$

Los números que tienen tanto parte real como imaginaria se conocen como **números complejos** y pueden definirse como

$$z = a + jb, \quad (\text{B.7})$$

donde $a = \text{Re}(z)$ se dice que es la parte real de z y $b = \text{Im}(z)$ la parte imaginaria de z .

Usualmente los números complejos se representan en un plano de modo que sobre el eje vertical se sitúa el eje imaginario y sobre el eje horizontal el eje real. De este modo, el número z queda caracterizado por un punto (como se muestra en la figura adjunta) que está a una distancia $|z|$ dada por

$$|z| = \sqrt{a^2 + b^2}, \quad (\text{B.8})$$

que se conoce como **módulo** de z , y con un ángulo φ medido en sentido antihorario a partir del eje real dado por

$$\varphi = \arctan\left(\frac{b}{a}\right), \quad (\text{B.9})$$

que se denomina **argumento** de z .

Es fácil observar en la figura que z puede escribirse como

$$z = |z|(\cos \varphi + j \sen \varphi),$$

y dado que la *identidad de Euler* dice que

$$e^{j\varphi} = \cos \varphi + j \sen \varphi, \quad (\text{B.10})$$

se tiene que el número complejo z puede reescribirse como

$$z = |z|e^{j\varphi}. \quad (\text{B.11})$$

Teniendo en cuenta la identidad (B.10), es fácil notar que la función armónica

$$f(t) = A \cos(\omega t + \varphi)$$

puede escribirse como

$$f(t) = \text{Re}\left(Ae^{j(\omega t + \varphi)}\right) = \text{Re}\left(Ae^{j\varphi}e^{j\omega t}\right). \quad (\text{B.12})$$

Si ahora definimos el **fasor**, \tilde{f} , de la función $f(t)$ como

$$\tilde{f} = Ae^{j\varphi}, \quad (\text{B.13})$$

se tiene que

$$f(t) = \text{Re}\left(\tilde{f}e^{j\omega t}\right). \quad (\text{B.14})$$

La identidad (B.14) permite establecer una relación biunívoca entre las funciones armónicas y sus fasores asociados, de modo que a toda función armónica se le puede asociar un fasor, esto es,

$$f(t) \leftrightarrow \tilde{f}. \quad (\text{B.15})$$

Siguiendo las propiedades más elementales del cálculo de números complejos pueden demostrarse fácilmente las siguientes propiedades:

$$f_1(t) + f_2(t) \leftrightarrow \tilde{f}_1 + \tilde{f}_2 \quad (\text{B.16})$$

$$\alpha f(t) \leftrightarrow \alpha \tilde{f}, \quad (\text{B.17})$$

siendo $f_i(t) = A_i \cos(\omega t + \varphi_i)$ y α un número real. Una propiedad adicional de fundamental importancia práctica es

$$\frac{df(t)}{dt} \leftrightarrow j\omega \tilde{f}. \quad (\text{B.18})$$

Esta última propiedad puede deducirse como sigue:

$$\begin{aligned} \frac{df(t)}{dt} &= -\omega A \sin(\omega t + \varphi) = -\omega A \cos(\omega t + \varphi - \pi/2) \\ &= \text{Re} \left(-\omega A e^{j(\omega t + \varphi - \pi/2)} \right) = \text{Re} \left(-\omega A e^{j\varphi} e^{-j\pi/2} e^{j\omega t} \right) \\ &= \text{Re} \left(j\omega A e^{j\varphi} e^{j\omega t} \right) = \text{Re} \left(j\omega \tilde{f} e^{j\omega t} \right), \end{aligned}$$

de donde se deduce que el fasor asociado a df/dt es precisamente $j\omega \tilde{f}$.

Copia Uso Académico

Apéndice C

Operadores divergencia y rotacional

C.1. Divergencia

Dado un campo vectorial $\mathbf{A}(\mathbf{r})$, si tomamos el producto escalar del operador ∇ (A.32) por este campo, podemos formar la siguiente magnitud *escalar*, $\nabla \cdot \mathbf{A}$:

$$\begin{aligned}\nabla \cdot \mathbf{A} &= \left(\hat{\mathbf{x}} \frac{\partial}{\partial x} + \hat{\mathbf{y}} \frac{\partial}{\partial y} + \hat{\mathbf{z}} \frac{\partial}{\partial z} \right) \cdot (\hat{\mathbf{x}} A_x + \hat{\mathbf{y}} A_y + \hat{\mathbf{z}} A_z) \\ &= \frac{\partial A_x}{\partial x} + \frac{\partial A_y}{\partial y} + \frac{\partial A_z}{\partial z},\end{aligned}\tag{C.1}$$

que se conoce como **divergencia** de \mathbf{A} . Resulta que esta magnitud escalar expresa la *densidad volumétrica de flujo* del campo vectorial \mathbf{A} y, por tanto, $\nabla \cdot \mathbf{A}$ puede definirse alternativamente como

$$\nabla \cdot \mathbf{A} = \lim_{V \rightarrow 0} \frac{\int_{S(V)} \mathbf{A} \cdot d\mathbf{S}}{V}.\tag{C.2}$$

Divergencia de \mathbf{A}

A la vista de (C.2), la operación $\nabla \cdot \mathbf{A}$ nos dice si hay fuentes ($\nabla \cdot \mathbf{A} > 0$) o sumideros ($\nabla \cdot \mathbf{A} < 0$) del campo $\mathbf{A}(\mathbf{r})$ en el punto \mathbf{r} , esto es, si hay líneas de campo que empiezan o acaban en dicho punto. Obviamente, si no hay líneas de campo que empiezan o acaban en el punto \mathbf{r} entonces $\nabla \cdot \mathbf{A} = 0$. Un campo para el que se cumpla que

$$\nabla \cdot \mathbf{A} = 0 \quad \forall \mathbf{r},\tag{C.3}$$

indica que las líneas de este campo no empiezan ni acaban en ningún punto del espacio. Un campo con estas condiciones se denomina **solenoidal**.

C.2. Rotacional

Análogamente a como en el apartado anterior se formó el producto escalar del operador ∇ por el campo vectorial \mathbf{A} , ahora formaremos el siguiente producto

vectorial $\nabla \times \mathbf{A}$:

$$\begin{aligned} \nabla \times \mathbf{A} &= \begin{vmatrix} \hat{\mathbf{x}} & \hat{\mathbf{y}} & \hat{\mathbf{z}} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ A_x & A_y & A_z \end{vmatrix} \\ &= \left(\frac{\partial A_z}{\partial y} - \frac{\partial A_y}{\partial z} \right) \hat{\mathbf{x}} + \left(\frac{\partial A_x}{\partial z} - \frac{\partial A_z}{\partial x} \right) \hat{\mathbf{y}} + \left(\frac{\partial A_y}{\partial x} - \frac{\partial A_x}{\partial y} \right) \hat{\mathbf{z}}, \end{aligned} \quad (\text{C.4})$$

que se denomina rotacional de \mathbf{A} . Evidentemente, la magnitud anterior, como todo producto vectorial, es un *vector*. Cada una de las componentes de este vector, $\nabla \times \mathbf{A}$, mide la *densidad superficial de circulación* en superficies normales a dichas componentes. Así, la componente z de $\nabla \times \mathbf{A}$, puede expresarse como

Componente z del rotacional de \mathbf{A}

$$(\nabla \times \mathbf{A})_z = \lim_{S_z \rightarrow 0} \frac{\oint_{\Gamma_z} \mathbf{A} \cdot d\mathbf{l}}{S_z}. \quad (\text{C.5})$$

Desde un punto de vista geométrico, la anterior definición sugiere que $\nabla \times \mathbf{A}$ es una medida del número de “remolinos” del campo \mathbf{A} en el punto \mathbf{r} . Un campo que no tenga “remolinos” en ningún punto de espacio, es decir,

$$\nabla \times \mathbf{A} = \mathbf{0} \quad \forall \mathbf{r}, \quad (\text{C.6})$$

se dice que es un campo **irrotacional**.

Apéndice D

Repaso del análisis de Fourier

D.1. Funciones periódicas

La teoría de Fourier dice que una función periódica, $f(x)$, de periodo λ ,

$$f(x) = f(x + \lambda),$$

puede expresarse como una serie de senos y cosenos de la siguiente manera:

$$f(x) = a_0 + \sum_{n=1}^{\infty} a_n \cos n \frac{2\pi}{\lambda} x + \sum_{n=1}^{\infty} b_n \operatorname{sen} n \frac{2\pi}{\lambda} x, \quad (\text{D.1})$$

donde los coeficientes (o pesos) de las senos/cosenos vienen dados por

$$a_n = \frac{2}{\lambda} \int_0^{\lambda} f(x) \cos n \frac{2\pi}{\lambda} x \, dx \quad (\text{D.2})$$

$$b_n = \frac{2}{\lambda} \int_0^{\lambda} f(x) \operatorname{sen} n \frac{2\pi}{\lambda} x \, dx. \quad (\text{D.3})$$

El desarrollo (D.1) de la función $f(x)$ puede igualmente expresarse como la siguiente serie de exponenciales complejas

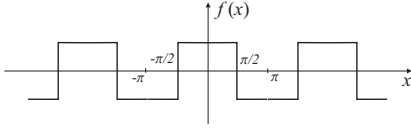
$$f(x) = \sum_{n=-\infty}^{\infty} A_n \exp\left(-jn \frac{2\pi}{\lambda} x\right), \quad (\text{D.4})$$

donde podemos identificar al comparar (D.4) con (D.1)

$$\left. \begin{array}{l} A_{-n} + A_n = a_n \\ j(A_{-n} - A_n) = b_n \end{array} \right\} \Rightarrow \begin{array}{l} A_{-n} = \frac{1}{2}(a_n - jb_n) \\ A_n = \frac{1}{2}(a_n + jb_n) \end{array}$$

y, por consiguiente, escribir los coeficientes de las exponenciales como

$$A_n = \frac{1}{\lambda} \int_0^{\lambda} f(x) \exp\left(jn \frac{2\pi}{\lambda} x\right) \, dx. \quad (\text{D.5})$$



EJEMPLO D.1 Encontrar el desarrollo de Fourier de una onda cuadrada par de periodo 2π

Dado que $\lambda = 2\pi$, tenemos que

$$\begin{aligned} A_n &= \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) e^{jn x} dx \\ &= \frac{1}{2\pi} \left[- \int_{-\pi}^{-\pi/2} e^{jn x} dx + \int_{-\pi/2}^{\pi/2} e^{jn x} dx - \int_{\pi/2}^{\pi} e^{jn x} dx \right] \\ &= \frac{1 - e^{jn\pi}}{n\pi} \operatorname{sen} \frac{n\pi}{2}. \end{aligned}$$

Podemos observar que estos coeficientes van decreciendo y representan los "peso" de las distintas exponenciales en la serie para formar la onda cuadrada, esto es,

$$f(x) = \sum_{n=-\infty}^{\infty} \frac{1 - e^{jn\pi}}{n\pi} \operatorname{sen} \frac{n\pi}{2} \exp(-jn x).$$

D.2. Funciones no periódicas

El análisis anterior puede extenderse a funciones no periódicas haciendo, por ejemplo, tender el periodo de la función $f(x)$ a infinito $\lambda \rightarrow \infty$. De este modo y tras tomar los oportunos límites, para una función $f(x)$ de cuadrado integrable, esto es,

$$\int_{-\infty}^{\infty} |f(x)|^2 dx < \infty, \quad (\text{D.6})$$

puede escribirse que¹

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} F(k) \exp(-jkx) dk \quad (\text{D.7a})$$

$$F(k) = \int_{-\infty}^{\infty} f(x) \exp(jkx) dx, \quad (\text{D.7b})$$

o equivalentemente

$$f(x) = \int_{-\infty}^{\infty} F(\nu) \exp(-j2\pi\nu x) d\nu \quad (\text{D.8a})$$

$$F(\nu) = \int_{-\infty}^{\infty} f(x) \exp(j2\pi\nu x) dx, \quad (\text{D.8b})$$

donde la variable de Fourier, ν , puede identificarse como una *frecuencia espacial*. La función $F(\nu)$ dada por (D.8b) se conoce como **transformada de Fourier** o también *espectro en frecuencia* de la función $f(x)$ y se denota a veces de forma simbólica como

$$F(\nu) = \mathcal{F}\{f(x)\}. \quad (\text{D.9})$$

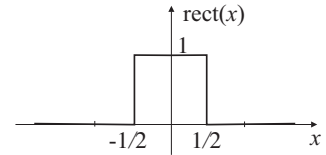
El análisis de Fourier permite interpretar la expresión (D.7a) como una descomposición de la función $f(x)$ en términos de una combinación lineal (en este caso *continua* por ser una integral) de funciones elementales del tipo $\exp(-j2\pi\nu x)$. En consecuencia, la transformada de Fourier $F(\nu)$ puede considerarse como un "factor de peso" que debe aplicarse a cada una de las funciones elementales de frecuencia espacial ν para sintetizar la función $f(x)$.

¹ En general, las funciones matemáticas que representan *magnitudes físicas* son todas susceptibles de ser tratadas mediante el análisis de Fourier.

EJEMPLO D.2 Transformada de Fourier de una función rectángulo $\text{rect}(x)$

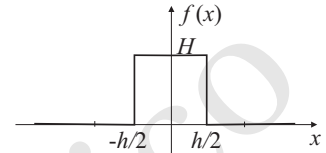
Si se define la función $\text{rect}(x)$ como

$$\text{rect}(x) = \begin{cases} 1 & \text{si } |x| < 1/2 \\ 1/2 & \text{si } |x| = 1/2, \\ 0 & \text{si } |x| > 1/2 \end{cases} \quad (\text{D.10})$$



la transformada de Fourier, $F(v)$, de la función

$$f(x) = H \text{rect}\left(\frac{x}{h}\right)$$



mostrada en la figura viene dada por

$$\begin{aligned} F(v) &= \int_{-\infty}^{\infty} f(x)e^{j2\pi vx} dx = \int_{-h/2}^{h/2} He^{j2\pi vx} dx = H \frac{e^{j2\pi vx}}{j2\pi v} \Big|_{-h/2}^{h/2} \\ &= H \frac{e^{j\pi vh} - e^{-j\pi vh}}{j2\pi vh} = Hh \frac{\text{sen } \pi vh}{\pi vh} = Hh \text{senc}(\pi vh), \end{aligned}$$

donde la función $\text{senc}(x)$ se define como

$$\text{senc}(x) = \frac{\text{sen } x}{x}. \quad (\text{D.11})$$

Es interesante observar que cuanto más ancha es la función rectángulo (h más grande), más estrecha es su correspondiente transformada de Fourier (y viceversa). Esto sugiere que funciones cuya localización espacial es muy reducida tienen un espectro en frecuencia muy amplio. De forma genérica puede decirse que

variaciones muy rápidas de la función $f(x)$ implican un amplio espectro de frecuencia para su correspondiente transformada de Fourier.

Algunas de las propiedades más útiles de la transformada de Fourier son:

$$\mathcal{F}\{af(x) + bf(x)\} = aF(v) + bG(v) \quad (\text{D.12})$$

$$\mathcal{F}\{f(x \pm a)\} = \exp(\mp j2\pi va)F(v) \quad (\text{D.13})$$

$$\mathcal{F}\{f(ax)\} = \frac{1}{a}F\left(\frac{v}{a}\right) \quad (\text{D.14})$$

$$\mathcal{F}\{f^*(x)\} = F^*(-v) \quad (\text{D.15})$$

$$\mathcal{F}\{f \otimes g\} = F(v)G(v). \quad (\text{D.16})$$

En la última expresión $f \otimes g$ representa el **producto de convolución** de la función $f(x)$ por la función $g(x)$ y viene dado por

$$f \otimes g = \int_{-\infty}^{\infty} f(\xi)g(x - \xi) d\xi. \quad (\text{D.17})$$

D.3. Transformada bidimensional de Fourier

Anteriormente se ha aplicado el análisis de Fourier a funciones de una sola variable. Este análisis puede igualmente ser aplicado a funciones de más de una

variable. En concreto, para el caso de dos variables tenemos que para una función $f(x, y)$ de cuadrado integrable podemos definir el siguiente par de transformadas:

$$f(x, y) = \iint_{-\infty}^{\infty} F(v_x, v_y) \exp[-j2\pi(v_x x + v_y y)] dv_x dv_y \quad (\text{D.18a})$$

Definición del par de transformadas bidimensionales de Fourier

$$F(v_x, v_y) = \iint_{-\infty}^{\infty} f(x, y) \exp[j2\pi(v_x x + v_y y)] dx dy, \quad (\text{D.18b})$$

donde (D.18a) puede interpretarse como la descomposición de la función $f(x, y)$ en términos de una combinación lineal continua de funciones elementales del tipo $\exp[-j2\pi(v_x x + v_y y)]$, donde v_x y v_y representan frecuencias espaciales con respecto a x e y respectivamente. Claramente, $F(v_x, v_y)$ representa el “factor peso” de cada función elemental de frecuencias v_x y v_y .

EJEMPLO D.3 Transformada de Fourier de la función “rectángulo” en dos dimensiones

Para una función del tipo

$$f(x, y) = \text{rect}\left(\frac{x}{a}\right) \text{rect}\left(\frac{y}{b}\right),$$

la transformada de Fourier correspondiente será

$$\begin{aligned} F(v_x, v_y) &= \iint_{-\infty}^{\infty} \text{rect}\left(\frac{x}{a}\right) \text{rect}\left(\frac{y}{b}\right) \exp[j2\pi(v_x x + v_y y)] dx dy \\ &= \left\{ \int_{-\infty}^{\infty} \text{rect}\left(\frac{x}{a}\right) \exp(j2\pi v_x x) dx \right\} \left\{ \int_{-\infty}^{\infty} \text{rect}\left(\frac{y}{b}\right) \exp(j2\pi v_y y) dy \right\} \\ &= ab \text{senc}(\pi v_x a) \text{senc}(\pi v_y b). \end{aligned}$$

Claramente cuanto más localizada espacialmente esté la función $f(x, y)$, más amplio será su espectro en frecuencias.

Algunas de las propiedades más significativas de la transformada de Fourier bidimensional son:

$$\mathcal{F}\{af(x, y) + bf(x, y)\} = aF(v_x, v_y) + bG(v_x, v_y) \quad (\text{D.19})$$

$$\mathcal{F}\{f(x-a, y-b)\} = \exp[j2\pi(v_x a + v_y b)]F(v_x, v_y) \quad (\text{D.20})$$

$$\mathcal{F}\{f(ax, by)\} = \frac{1}{ab} F\left(\frac{v_x}{a}, \frac{v_y}{b}\right) \quad (\text{D.21})$$

$$\mathcal{F}\{f^*(x, y)\} = F^*(-v_x, -v_y) \quad (\text{D.22})$$

$$\mathcal{F}\{f \otimes g\} = F(v_x, v_y)G(v_x, v_y), \quad (\text{D.23})$$

donde

$$f \otimes g = \iint_{-\infty}^{\infty} f(\xi, \eta) g(x-\xi, y-\eta) d\xi d\eta.$$

D.3.1. Interpretación geométrica de la transformada bidimensional de Fourier

Paralela a la interpretación *matemática* que se ha dado para

$$f(x, y) = \iint_{-\infty}^{\infty} F(v_x, v_y) \exp[-j2\pi(v_x x + v_y y)] dv_x dv_y \quad (\text{D.24})$$

como una combinación lineal (en este caso, una integral) de funciones elementales de la forma $\exp[-j2\pi(v_x x + v_y y)]$, cada una de ellas "pesadas" por el factor $F(v_x, v_y)$, podemos dar una interpretación más "física".

En este sentido, podemos considerar que la expresión integral (D.24) es una combinación lineal de **ondas planas** del tipo

$$F(v_x, v_y) \exp[-j(k_x x + k_y y)] , \quad (\text{D.25})$$

cuyo vector de onda \mathbf{k} es

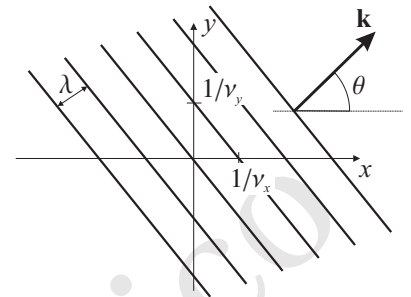
$$\mathbf{k} = k_x \hat{\mathbf{x}} + k_y \hat{\mathbf{y}} ; \quad k_x = 2\pi v_x , \quad k_y = 2\pi v_y . \quad (\text{D.26})$$

Tal como muestra la figura adjunta, cada "onda plana elemental" es una onda plana cuyo vector de ondas \mathbf{k} está dirigido en el plano (x, y) formando un ángulo θ con respecto al eje x dado por

$$\theta = \arctan\left(\frac{v_y}{v_x}\right) . \quad (\text{D.27})$$

El periodo espacial, λ , de esta onda (esto es, la distancia entre dos frentes de onda de igual fase) viene dado por

$$\lambda = \frac{2\pi}{|\mathbf{k}|} = \frac{1}{\sqrt{v_x^2 + v_y^2}} . \quad (\text{D.28})$$



Copia Uso Académico

Apéndice E

Función delta de Dirac

La función delta de Dirac monodimensional no es propiamente una *función*, sino que es una entidad más general denominada, *funcional* o distribución. La diferencia de una funcional con respecto a una función es que una función hace corresponder números con números mientras que una funcional hace corresponder funciones con números. Un sencillo ejemplo de funcional es la siguiente integral definida:

$$\int_{-\infty}^{\infty} f(\xi) d\xi,$$

que hace corresponder cualquier función $f(x)$ con el valor numérico de su área respecto al eje x .

En este sentido, la característica definitoria de la función delta (la seguiremos llamando “función” debido a su extendido uso) es la conocida como **propiedad de muestreo** bajo el signo de integral, esto es,

$$\int_{-\infty}^{\infty} f(\xi) \delta(a - \xi) d\xi = f(a). \quad (\text{E.1})$$

Esta propiedad es la que define la funcional que conocemos como función delta. Aunque la integral anterior es una representación conveniente para esta correspondencia, dicha integral debiera ser interpretada como el siguiente límite de un conjunto de integrales:

$$\int_{-\infty}^{\infty} f(\xi) \delta(a - \xi) d\xi = \lim_{N \rightarrow \infty} \int_{-\infty}^{\infty} f(\xi) g_N(a - \xi) d\xi, \quad (\text{E.2})$$

donde g_N es una secuencia de funciones que en límite $N \rightarrow \infty$ exhibe la propiedad de muestreo requerida. Estas funciones deben todas tener área unidad y deben hacerse “en algún sentido” más estrechas a medida que N crece.

A pesar de no ser formalmente correcto, es usual representar la función delta como el límite de la secuencia de funciones $g_N(x)$,

$$\delta(x) = \lim_{N \rightarrow \infty} g_N(x), \quad (\text{E.3})$$

aunque debe entenderse siempre la anterior “identidad” en el sentido dado en la expresión (E.2). Teniendo esto en cuenta, puede comprobarse, por ejemplo, que

$$\delta(x) = \lim_{N \rightarrow \infty} N \exp(-N^2 \pi x^2) \quad (\text{E.4})$$

$$\delta(x) = \lim_{N \rightarrow \infty} N \text{rect}(Nx) \quad (\text{E.5})$$

$$\delta(x) = \lim_{N \rightarrow \infty} N \text{senc}(Nx). \quad (\text{E.6})$$

En teoría de circuitos, la función delta se usa frecuentemente para representar un pulso instantáneo de corriente o voltaje. En óptica, el concepto equivalente es una *fente puntual de luz*, o un "pulso espacial" de volumen unidad. En óptica, por tanto, la función delta que nos interesa es bidimensional, $\delta(x, y)$, cuya definición es una sencilla extensión del caso monodimensional:

$$\iint_{-\infty}^{\infty} f(\xi, \eta) \delta(a - \xi, b - \eta) d\xi d\eta = f(a, b) . \quad (\text{E.7})$$

Algunas posibles definiciones de funciones delta son

$$\delta(x, y) = \lim_{N \rightarrow \infty} N^2 \exp[-N^2 \pi(x^2 + y^2)] \quad (\text{E.8})$$

$$\delta(x, y) = \lim_{N \rightarrow \infty} N^2 \text{rect}(Nx) \text{rect}(Ny) \quad (\text{E.9})$$

$$\delta(x, y) = \lim_{N \rightarrow \infty} N^2 \text{senc}(Nx) \text{senc}(Ny) . \quad (\text{E.10})$$

Apéndice F

Introducción a los Sistemas Lineales

Una forma matemática muy conveniente de expresar la acción de un sistema físico es mediante el concepto de *operador matemático*. De este modo, si la entrada de cierto sistema viene descrita por $f(x, y)$, la salida de este sistema, $g(x, y)$, puede expresarse como

$$g(x, y) = \mathcal{S}\{f(x, y)\} \quad (\text{E.1})$$

Un sistema se dice que es **lineal** si cumple la siguiente propiedad:

$$\mathcal{S}\{af_1(x, y) + bf_2(x, y)\} = a\mathcal{S}\{f_1(x, y)\} + b\mathcal{S}\{f_2(x, y)\}, \quad (\text{E.2})$$

esto es, si efecto del sistema sobre una suma de entradas viene dado por la suma de los efectos sobre cada una de las entradas individuales.

La gran ventaja matemática que proporcionan los sistemas lineales es que permiten expresar la respuesta de un sistema a una entrada arbitraria en función de las respuestas a ciertas funciones elementales en las que se pueda descomponer la entrada.

Una posible forma de descomponer la entrada en funciones elementales hace uso de la propiedad de muestreo de la función delta de Dirac. De este modo, la función $f(x, y)$ puede expresarse como

$$f(x, y) = \iint_{-\infty}^{\infty} f(\xi, \eta) \delta(x - \xi, y - \eta) d\xi d\eta, \quad (\text{E.3})$$

es decir, como una combinación lineal continua de funciones δ pesadas y desplazadas.

La salida de este sistema lineal, $g(x, y)$, podrá entonces obtenerse como

$$\begin{aligned} g(x, y) &= \mathcal{S}\left\{ \iint_{-\infty}^{\infty} f(\xi, \eta) \delta(x - \xi, y - \eta) d\xi d\eta \right\} \\ &= \iint_{-\infty}^{\infty} f(\xi, \eta) \mathcal{S}\{\delta(x - \xi, y - \eta)\} d\xi d\eta. \end{aligned} \quad (\text{E.4})$$



Si denominamos **respuesta impulsiva** del sistema a

$$h(x, y; \xi, \eta) = \mathcal{S}\{\delta(x - \xi, y - \eta)\}, \quad (\text{E.5})$$

entonces la salida del sistema puede expresarse como la siguiente *integral de superposición*:

$$g(x, y) = \iint_{-\infty}^{\infty} f(\xi, \eta) h(x, y; \xi, \eta) d\xi d\eta. \quad (\text{E.6})$$

Podemos ver que la respuesta impulsiva del sistema es simplemente la respuesta del sistema a una entrada tipo delta de Dirac. Para el caso de un sistema óptico, el resultado (E.6) puede reinterpretarse diciendo que

el efecto de los elementos ópticos (lentes, filtros,...) queda completamente descrito especificando las imágenes (tanto la amplitud como la fase) correspondientes a fuentes puntuales localizadas en el campo objeto.

Sistemas ópticos lineales invariantes

Condición que debe satisfacer un sistema lineal invariante

Un sistema óptico lineal es **invariante espacial** si su respuesta impulsiva depende únicamente de las distancias entre los puntos fuentes objeto y los puntos imágenes respuesta, esto es, si

$$h(x, y; \xi, \eta) = h(x - \xi, y - \eta). \quad (\text{E.7})$$

La condición anterior implica que la imagen de un punto fuente objeto cambia únicamente en localización, y no en forma funcional, cuando el punto fuente explora el campo objeto (en otras palabras, si el objeto global cambia de posición en el plano (x, y) de entrada, la imagen de este objeto en el plano de salida seguiría teniendo la misma "forma").

Para Sistemas Lineales Invariantes (SLI), la integral de superposición toma la siguiente forma:

$$g(x, y) = \iint_{-\infty}^{\infty} f(\xi, \eta) h(x - \xi, y - \eta) d\xi d\eta, \quad (\text{E.8})$$

que podemos observar que es una *convolución bidimensional*, que suele escribirse como

$$g = f \otimes h. \quad (\text{E.9})$$

Teniendo en cuenta la propiedad (D.23), la transformada de Fourier de la imagen respuesta vendrá dada por

$$G(v_x, v_y) = H(v_x, v_y)F(v_x, v_y), \quad (\text{E.10})$$

donde la función $H(v_x, v_y)$ se conoce como **función de transferencia** del sistema y puede obtenerse a partir de

$$H(v_x, v_y) = \frac{G(v_x, v_y)}{F(v_x, v_y)} = \iint_{-\infty}^{\infty} h(x, y) \exp[j2\pi(v_x x + v_y y)] dv_x dv_y. \quad (\text{E.11})$$

Nótese que la función de transferencia es la transformada de Fourier de la respuesta impulsiva, esto es,

$$H(v_x, v_y) = \mathcal{F}\{h(x, y)\} \quad (\text{E.12})$$

y, análogamente, la función respuesta impulsiva puede obtenerse como la transformada de Fourier inversa de la función de transferencia,

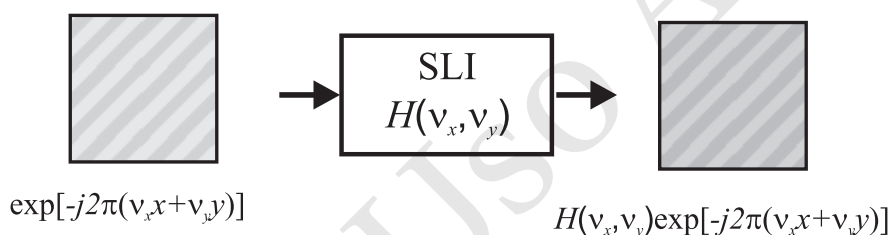
$$h(x, y) = \iint_{-\infty}^{\infty} H(v_x, v_y) \exp[-j2\pi(v_x x + v_y y)] dv_x dv_y. \quad (\text{E.13})$$

Aunque la función de transferencia puede siempre “calcularse” a partir de (E.12), debe tenerse en cuenta que esta función es solamente útil en sistemas lineales invariantes.

A partir de la función de transferencia, la respuesta del SLI puede obtenerse como

$$g(x, y) = \iint_{-\infty}^{\infty} H(v_x, v_y) F(v_x, v_y) \exp[-j2\pi(v_x x + v_y y)] dv_x dv_y. \quad (\text{E.14})$$

Las anteriores relaciones sugieren que para un SLI la entrada puede ser descompuesta en funciones elementales más convenientes que las funciones delta de Dirac de (E.3), en concreto en funciones exponenciales complejas. La razón de ello es que estas funciones mantienen su forma bajo el efecto del SLI (es decir, son autofunciones del sistema).



Una función armónica $\exp[-j2\pi(v_x x + v_y y)]$ como entrada a un SLI produce una respuesta armónica de la *misma* frecuencia, afectada simplemente por un factor de amplitud y fase. El efecto del SLI sobre la entrada armónica es simplemente cambiar la amplitud y fase de la entrada, reflejándose este hecho en el factor multiplicativo complejo $H(v_x, v_y)$.