

# Ondas

---

**Ricardo A. Depine**

Departamento de Física  
Facultad de Ciencias Exactas y Naturales  
Universidad de Buenos Aires

Versión: 5 de septiembre de 2020

Primera versión: segundo cuatrimestre de 2015

(comprobar actualizaciones en <http://bit.ly/F2Depinev02>)

© 2020 Ricardo A. Depine. All rights reserved.

# Índice general

<b>1. Introducción</b>	<b>1</b>
1.1. ¿Qué es una onda?	2
1.2. Repaso de 1 GL (grado de libertad)	2
1.3. Uso de programación y aplicaciones	3
1.4. Repaso de 1GL: ejemplo numérico	5
1.4.1. Aproximación analítica	6
1.4.2. Solución numérica	6
1.4.3. Puesta en marcha	8
1.4.4. Conservación de la energía	10
1.4.5. Péndulo con rozamiento	13
<b>2. Oscilaciones libres y modos normales</b>	<b>19</b>
2.1. Dos grados de libertad	19
2.1.1. Primer ejemplo con $N = 2$	19
2.1.2. Condiciones iniciales	22
2.1.3. Los modos normales	22
2.1.4. Modos normales y forzado estacionario	23
2.1.5. Energía cinética y potencial	24
2.1.6. El ejemplo de $N = 2$ se complica	25
2.2. Búsqueda sistemática de modos normales	26
2.2.1. Notación y $N$ grados de libertad	26
2.2.2. Busquemos un modo físicamente	27

2.2.3.	Buscamos un modo y encontramos $N$ modos . . . . .	28
2.2.4.	El movimiento más general . . . . .	29
2.2.5.	El método aplicado al ejemplo . . . . .	30
2.2.6.	Todo lo anterior en el lenguaje del álgebra lineal . . . . .	31
	Sistema desacoplado y coordenadas normales . . . . .	31
	Condiciones iniciales . . . . .	31
	Energía cinética y potencial . . . . .	32
2.3.	Matrices, autovalores y autovectores en Python . . . . .	32
2.3.1.	Visualización de resultados . . . . .	35
2.3.2.	Archivo de comandos o <i>script</i> . . . . .	39
2.4.	Batidos y pulsaciones . . . . .	41
<b>3.</b>	<b>Muchas partes móviles</b>	<b>45</b>
3.1.	Cadena con $N$ cuentas idénticas equiespaciadas . . . . .	46
3.1.1.	Ecuaciones de movimiento para oscilaciones transversales . . . . .	46
3.1.2.	Modos normales . . . . .	48
3.1.3.	Relación de dispersión . . . . .	50
3.1.4.	Modos para dos extremos fijos . . . . .	51
	Condiciones iniciales . . . . .	55
3.1.5.	Condición para extremo libre . . . . .	56
3.1.6.	Oscilaciones longitudinales . . . . .	56
3.2.	Péndulos idénticos acoplados . . . . .	56
3.3.	Aproximación continua . . . . .	59
3.3.1.	Cuerda $\implies$ ecuación de ondas clásica . . . . .	60
3.3.2.	Slinky $\implies$ ecuación de ondas clásica . . . . .	61
3.3.3.	Péndulos acoplados $\implies$ ecuación de Klein-Gordon . . . . .	63
3.4.	La ecuación de ondas . . . . .	63
3.5.	Modos normales en el medio continuo . . . . .	64
3.6.	Condiciones en los extremos . . . . .	65

3.7. Cuerda con extremos fijos . . . . .	66
3.7.1. Discretización de los modos normales . . . . .	66
3.7.2. Condiciones iniciales . . . . .	67
3.7.3. Series de Fourier . . . . .	68
Coeficientes de la serie de Fourier impar . . . . .	69
Ejemplo: diente de sierra . . . . .	71
Ejemplo: pulso rectangular . . . . .	74
Condiciones iniciales . . . . .	74
3.7.4. Evolución temporal . . . . .	75
3.8. ondas estacionarias = suma de ondas viajeras . . . . .	79
3.9. Y en otros sistemas . . . . .	81
<b>4. Ondas propagantes</b>	<b>83</b>
4.1. Ecuación clásica 1D . . . . .	83
4.1.1. Progresivas y regresivas 1D . . . . .	83
4.1.2. La solución más general 1D . . . . .	84
4.1.3. Fórmula de d'Alembert . . . . .	85
4.2. Ondas, medios y dimensiones . . . . .	87
4.2.1. La onda plana en un medio 3D . . . . .	88
La onda plana armónica en un medio 3D . . . . .	88
4.2.2. Fuentes con simetría esférica . . . . .	89
4.2.3. Fuentes con simetría cilíndrica . . . . .	90
4.3. Modulación . . . . .	91
4.3.1. Caso más sencillo . . . . .	92
4.3.2. Modulaciones reales . . . . .	93
4.3.3. Síntesis de señales . . . . .	94
Mezclando $N$ componentes armónicos . . . . .	96
Límite $N \rightarrow \infty$ , con $\Delta\omega$ constante . . . . .	96
4.3.4. Transformada de Fourier . . . . .	100

<b>5. Respuesta forzada estacionaria</b>	<b>103</b>
5.1. Forzado estacionario, cadena lineal . . . . .	104
5.1.1. Rango dispersivo . . . . .	104
5.1.2. Rango reactivo . . . . .	105
5.2. Forzado estacionario, Klein-Gordon . . . . .	106
5.2.1. Dispersivo vs reactivo . . . . .	107
5.2.2. La parte con $e^{\kappa x}$ no está permitida en sistemas infinitos . . . . .	108
5.2.3. Comentarios sobre el régimen reactivo . . . . .	109
5.3. Descripción discreta . . . . .	109
5.4. Cadena lineal diatómica . . . . .	111
5.4.1. Ecuaciones de movimiento . . . . .	112
5.4.2. Relación de dispersión . . . . .	114
5.4.3. Caso particular $f = g$ . . . . .	116
5.4.4. Modos para extremos fijos . . . . .	117
<b>6. Descripción geométrica</b>	<b>119</b>
6.1. Rayos . . . . .	120
6.2. Leyes fenomenológicas . . . . .	120
6.3. Algunas consecuencias de la ley de Ibn Sahl . . . . .	122
6.3.1. Reflexión externa e interna . . . . .	122
6.3.2. Reflexión total . . . . .	122
6.4. La DG en otros marcos teóricos . . . . .	123
6.4.1. Principio de Huygens . . . . .	124
6.4.2. Principio de Fermat . . . . .	127
6.5. Formación de imágenes . . . . .	128
6.5.1. Dioptra plana . . . . .	129
6.5.2. Dioptra esférica . . . . .	130
6.5.3. Trazado de rayos . . . . .	131
6.5.4. Lentes . . . . .	132

6.6. La DG en el marco ondulatorio . . . . .	134
6.6.1. Reflexión y transmisión de una onda plana en una superficie plana	136
6.6.2. Las condiciones de contorno imponen la cinemática de los rayos .	136
6.6.3. Conservación de componentes tangenciales de $\vec{k}$ . . . . .	138
6.6.4. La reflexión total en el marco ondulatorio . . . . .	139
<b>7. Ondas transversales</b>	<b>141</b>
7.1. Ondas planas armónicas transversales . . . . .	141
7.2. Polarización lineal . . . . .	142
7.3. Polarización elíptica . . . . .	143
7.4. Polarización circular . . . . .	144
7.5. Caso $\phi$ arbitrario . . . . .	145
7.5.1. Luz natural . . . . .	146
7.6. Polarización por absorción selectiva . . . . .	146
7.7. Polarización por reflexión . . . . .	148
7.8. Láminas retardadoras . . . . .	153
7.9. Birrefringencia en medios quirales . . . . .	153
7.10. Birrefringencia en cristales . . . . .	156
<b>8. Interferencia</b>	<b>157</b>
8.1. Fuentes puntuales . . . . .	158
8.2. Uso de fasores . . . . .	160
8.3. Sistemas de franjas . . . . .	161
8.4. Experiencia de Young . . . . .	161
8.5. Coherencia . . . . .	163
8.6. Interferómetros . . . . .	164
8.6.1. División del frente de onda . . . . .	164
8.6.2. División de amplitud . . . . .	165
Interferómetro de Michelson . . . . .	165

8.7. Interferencia en láminas . . . . .	167
8.8. Localización de franjas . . . . .	167
<b>9. Difracción</b>	<b>169</b>
9.1. Interferencia de $N$ fuentes . . . . .	169
9.1.1. Caso general . . . . .	170
9.1.2. Fuentes alineadas y equiespaciadas . . . . .	171
9.1.3. Intensidad en la zona lejana . . . . .	174
9.1.4. Intensidad en la zona cercana . . . . .	180
9.1.5. Experimentos numéricos . . . . .	182
9.1.6. Límite continuo, $N \rightarrow \infty$ , $d \rightarrow 0$ . . . . .	187
9.2. Difracción de Fresnel y de Fraunhofer . . . . .	188
9.3. Difracción de Fraunhofer . . . . .	189
9.4. Rendija de ancho $b$ . . . . .	191
9.5. Obstáculos planos con simetría de translación . . . . .	192
9.5.1. Otra vez una rendija . . . . .	194
9.5.2. $N$ rendijas idénticas equiespaciadas . . . . .	195
9.6. Obstáculos planos, caso general . . . . .	201
9.6.1. Abertura rectangular . . . . .	203
9.6.2. Abertura circular . . . . .	203
9.7. Óptica de Fourier . . . . .	203

# Capítulo 1

## Introducción

En Física 2 vamos a estudiar **ONDAS** (y en unas líneas veremos qué son). Nos proponemos **dos objetivos**:

- i) introducir **conceptos básicos de los mecanismos ondulatorios** (para esto nos basaremos esencialmente en sistemas mecánicos, cuya descripción debería resultar familiar por haber sido estudiada en Física 1);
- ii) mostrar que estos conceptos ondulatorios son comunes, y permiten la **unificación**, de una gran diversidad de fenómenos físicos, no solamente mecánicos, sino también de fenomenologías que hasta el momento no se han visto en la carrera.

Para avanzar con el punto ii) vamos a tener que proceder por **analogías** y decir cosas como “si con las ondas mecánicas pasaba esto, entonces con las ondas electromagnéticas (o acústicas, o sísmicas, o gravitatorias, o lo que sea) también tiene que pasar lo mismo”.

Como cuando se cursa Física 2 todavía no se ha visto casi nada de electromagnetismo, o de acústica, o de sísmica, es muy normal al principio sentirse un poco (o bastante) inseguros. No hay que asustarse cuando esto pasa. Para superar esta inseguridad y para poder desarrollar los objetivos del punto ii) exitosamente, iremos viendo a lo largo del curso evidencias de cuándo las analogías funcionan y de cuándo NO funcionan. No debemos perder de vista que la física es una ciencia experimental, y que las analogías sirven para “intuir” comportamientos posibles, pero sólo son válidas cuando están comprobadas **empíricamente**.

La palabra **unificación** empleada más arriba se refiere a que el estudio de las características de los fenómenos ondulatorios no depende del tipo de propiedad perturbada o del medio en cuestión. Sin embargo, no tenemos que olvidarnos que los distintos orígenes físicos confieren propiedades muy particulares a cada onda.

## 1.1. ¿Qué es una onda?

Una definición minimalista: “**Una perturbación que se propaga**”.

La *perturbación* de la que estamos hablando es de alguna propiedad física (por ejemplo, velocidad, posición, densidad, presión, campo eléctrico o campo magnético), de un sistema de partes móviles o de un medio continuo. Y si es una perturbación, se mide con respecto a un estado de referencia, que por lo general es el estado de *equilibrio*. Veremos que el sistema o el medio perturbado pueden ser de naturaleza muy diversa, como las cuentas de un collar unidas por resortes, aire, agua, un metal, etc. Incluso, el medio puede no ser un medio, como sucede cuando se propagan ondas electromagnéticas en el vacío. La perturbación se *propaga* a través del sistema, pero el sistema como un todo no se propaga. Por este motivo, conviene ampliar la definición minimalista anterior y explicitar que en una onda no hay transporte neto de materia, diciendo que una onda es **una perturbación que se propaga y que transporta energía pero no materia**. Y como hay que saber parafrasear, también podríamos decir que **llamamos onda a la propagación de una perturbación, con transporte de energía pero sin transporte de materia**.

Para que haya propagación, el medio tiene que ser extenso o tener muchas partes móviles, como una fila de fichas de dominó (con una sola ficha no se propaga nada). Por eso tenemos que aprender a tratar físicamente la perturbación del estado de equilibrio en sistemas con muchas partes. También tiene que haber un **mecanismo restitutivo**, que tienda a regresar las partes móviles a su posición de equilibrio. Sin este mecanismo restitutivo, el apartamiento del equilibrio no sería muy interesante, como sucede con muchas fichas de dominó, donde el apartamiento del equilibrio se propaga, pero el proceso termina cuando cae la última ficha y tendríamos que volver a posicionar las fichas a mano para volver a empezar.

En Física 1 se estudia como evoluciona la perturbación del estado de equilibrio de sistemas muy sencillos, en general una masa puntual o un cuerpo rígido con vínculos adecuados que sólo permitían un grado de libertad, como el péndulo o el resorte. Una de las primeras cosas que vamos a ver en Física 2 es cómo estudiar la evolución temporal de la perturbación del estado de equilibrio en sistemas *mecánicos* con muchas partes móviles (es decir, con muchos grados de libertad) y que interactúan entre sí mediante fuerzas que son las que proveen el mecanismo restitutivo. En particular también prestaremos atención al límite de  $N$  partes móviles con  $N \rightarrow \infty$ , que es la idealización de **medio continuo**.

## 1.2. Repaso de 1 GL (grado de libertad)

A continuación, una lista de cosas que en las clases teóricas voy a dar por sabidas (reparar de lo visto en F1):

1. equilibrio estable e inestable, sistemas conservativos, condiciones sobre el potencial

- en el punto de equilibrio;
2. si el equilibrio es estable la perturbación produce un movimiento limitado en el espacio;
  3. si la fuerza restitutiva es lineal (el potencial cuadrático), el movimiento es oscilatorio armónico, con una frecuencia característica  $\omega_0$  que es igual a la fuerza restitutiva por unidad de desplazamiento y por unidad de masa;
  4. si la perturbación del equilibrio estable es pequeña (hay que cuantificar “pequeña”), el movimiento es oscilatorio armónico, con una frecuencia característica  $\omega_0$  que es igual a la fuerza restitutiva por unidad de desplazamiento y por unidad de masa ( $d^2V/dx^2 > 0$ );
  5. soluciones de la ecuación diferencial del oscilador armónico amortiguado (segundo orden, lineal, homogénea y a coeficientes constantes);
  6. condiciones iniciales;
  7. amplitud, frecuencia, período, fase inicial, velocidad, aceleración, energía cinética, energía potencial, energía mecánica total, tiempo de decaimiento;
  8. cómo cambiar entre distintos tipos de representaciones:
    - $x(t) = A \cos(\omega t + \varphi_A)$
    - $x(t) = B \sin(\omega t + \varphi_B)$
    - $x(t) = C \cos(\omega t) + D \sin(\omega t)$
    - $x(t) = \Re a \exp(\pm i\omega t)$  ( $\Re$  parte real);
  9. oscilador armónico amortiguado **forzado**, fuerza impulsora armónica de frecuencia  $\omega$ , solución transitoria y estacionaria, resonancia;
  10. la solución **transitoria** representa soluciones de la ecuación diferencial homogénea y tiene información sobre las condiciones iniciales;
  11. la solución **estacionaria** es una solución particular de la ecuación diferencial no homogénea, no tiene información sobre las condiciones iniciales, tiene la misma frecuencia que la fuerza impulsora armónica;
  12. principio de superposición.

### 1.3. Uso de programación y aplicaciones

Como la Física es una ciencia experimental, es claro que la experimentación tiene una gran importancia en su enseñanza. En el plan de estudio de la Licenciatura en Física

de la UBA, la parte experimental está cubierta con al menos siete materias que se desarrollan directamente en los laboratorios de enseñanza, además de las demostraciones que se pueden hacer en clase en las materias teóricas. Sin embargo, el tiempo que tienen los estudiantes para jugar con experimentos y manejar la instrumentación es, desde limitado en los Labos hasta casi nulo en las demostraciones en el aula y en ambos casos no es posible seguir experimentando al salir de la facultad.

Este problema desaparece cuando se incorpora a la enseñanza de la física los programas y los conocimientos necesarios para poder realizar **experimentos numéricos con computadora**. Estos experimentos numéricos son habituales en todo equipo de investigación y son cada día más usados para evitar montajes experimentales muy costosos o para investigar situaciones que valga la pena someter a experimentación. En pleno siglo XXI resulta anacrónico no introducir desde un principio alguna de las herramientas numéricas más convenientes para que los estudiantes puedan seguir explorando los diferentes fenómenos físicos de manera independiente e ilustrar de manera gráfica el comportamiento de los fenómenos estudiados en las clases teóricas y de problemas. No es muy importante qué herramienta se usa, lo importante es empezar a usar alguna. Pero al principio y para entendernos todos, conviene usar una herramienta común y en el curso vamos a usar **Python**. Python es un programa libre que emplea un lenguaje de alto nivel, muy útil para cálculos numéricos y visualización de resultados. Se puede instalar y usar fácilmente tanto en Linux como en Windows como en IOS. Permite ejecutar órdenes de dos formas distintas: a) de forma interactiva (la forma que usaremos primero para ir aprendiendo sobre la marcha); y b) en forma de *script* o *archivo de órdenes* (un programa usualmente simple, que por lo regular se almacena en un archivo de texto plano). En el curso usaremos mucho una tercera manera: a través de **Cuadernos de Jupyter**, o Jupyter Notebooks, una aplicación web que permite incorporar en nuestros documentos elementos de código abierto (en nuestro caso, Python, por eso la Py the Jupyter), de texto enriquecido (con todas las cosas que necesitamos los científicos para escribir nuestras fórmulas y ecuaciones) y gráficos diversos de una forma sencilla y fluida.

Un cuaderno de Jupyter es simplemente un archivo con extensión `.ipynb` hecho por una aplicación web. Contiene muchas ‘celdas’ con código, texto con formatos muy convenientes (en particular admite el formato LaTeX, imprescindible para los físicos y muy conveniente para los foros del campus virtual), o metadatos. Jupyter interpreta estas celdas y hace que tengan una apariencia muy similar a la de un documento de texto. Las celdas que contienen código son ejecutadas, mostrando el resultado dentro de la celda. Es muy común el uso de librerías como matplotlib para mostrar gráficos, videos o imágenes que cambien o se actualicen a lo largo del tiempo y que se pueden guardar como archivos locales en cada dispositivo. En el [sitio oficial de Jupyter](#) se puede obtener información sobre diversas maneras de usar Jupyter en un dispositivo propio o en la nube. Para empezar, en el curso (empieza broma) nos vendimos a Google (termina broma) vamos a usar la aplicación **Colab**, una versión de los Cuadernos de Jupyter que tiene la ventaja de que no requiere ninguna instalación local, se ejecutan en la nube y están altamente integrados con Google Drive, lo que los hace fáciles de configurar, acceder y compartir.

Un consejo, tanto Python, como Jupyter Notebooks y Google Colab se aprenden sobre la marcha, es decir, usándolos. ¿O acaso hay gente que hace cursos de WhatsApp o de otras aplicaciones del celular? No!! Todos aprendemos usándolas. Y si nos trabamos con algo, preguntamos al que sabe, o buscamos respuestas en internet. Lo mismo pasa con estas herramientas numéricas, hay muchísima información en las redes y todas (o casi todas) las dudas están discutidas en los foros especializados.

Al trabajar con cuadernos (notebooks) Colab descubrirán que es muy fácil hacer sus propios apuntes. Y que estos apuntes, además de quedar super-prolijos y en formato digital, son INTERACTIVOS. Verán también que lo que hicieron en un cuaderno sirve también para otros, cortando y pegando, y adaptando convenientemente. Todos los físicos hemos aprendido a usar herramientas digitales así, partiendo de algo ya hecho, tratando de entenderlo y luego adaptándolo a nuestras necesidades.

## 1.4. Repaso de 1GL: ejemplo numérico

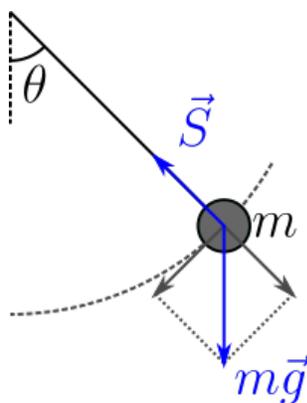


Figura 1.1: Esquema del péndulo simple

Este ejemplo está adaptado de [un cuaderno](#) publicado en [NumFys](#), dedicado al péndulo simple con y sin rozamiento. NumFys es un sitio web con recursos computacionales en línea para el uso de física computacional, con módulos de aprendizaje en Python y ejemplos que cubren temas muy variados. A continuación veremos una simplificación de este cuaderno para el caso ideal sin rozamiento. En el esquema de la fig. 1.2 se muestra el caso del **péndulo simple**, un modelo idealizado del péndulo real. Consiste en una masa puntual  $m$  suspendida por una cuerda (o varilla) no extensible y sin masa de longitud  $L$ . El péndulo solo se mueve en un plano y no hay fricción. Usamos  $\theta$  para denotar el ángulo entre el eje vertical y la cuerda, en el equilibrio  $\theta = 0$ . Como bajo estas circunstancias la posición del péndulo se puede especificar usando solamente la variable  $\theta$ , decimos que el problema tiene **un grado de libertad** (GL). Cuando se aparta la masa del equilibrio, aparece una oscilación y para estudiarla tenemos que plantear la ecuación de movimiento.

Hay dos fuerzas que actúan sobre la masa: la fuerza gravitacional  $\mathbf{F} = m\mathbf{g}$  y la ten-

sión de la cuerda  $\mathbf{S}$ . Las fuerzas se pueden descomponer en una dirección a lo largo de la cuerda (dirección radial,  $\hat{r}$ ) y una dirección a lo largo de la trayectoria de la masa puntual (dirección azimutal,  $\hat{\theta}$ ). Dado que se supone que la cuerda no se puede estirar, solo la atracción gravitacional en la dirección azimutal,  $\mathbf{F}_\theta$ , contribuirá al movimiento del péndulo. La fuerza gravitacional en la dirección radial,  $\mathbf{F}_r$ , se cancela con la tensión de la cuerda:  $\mathbf{F}_r = -\mathbf{S}$ . De la figura es claro que

$$F_\theta = -mg \sin \theta. \quad (1.1)$$

Ahora usaremos la segunda ley de Newton,  $\mathbf{F} = m\mathbf{a}$ , para encontrar una ecuación diferencial que describa el movimiento del péndulo. Sea  $t$  el tiempo. La aceleración de la masa puntual se puede escribir

$$\mathbf{a} = L \frac{d^2\theta}{dt^2} \hat{\theta} = L\ddot{\theta} \hat{\theta}, \quad (1.2)$$

y entonces el movimiento del péndulo se describe con la ecuación

$$L\ddot{\theta} = L \frac{d^2\theta}{dt^2} = -g \sin \theta. \quad (1.3)$$

### 1.4.1. Aproximación analítica

La ecuación (1.3) no se puede resolver analíticamente. Sin embargo, teniendo en cuenta que para  $\theta \ll 1$  vale que  $\sin \theta \approx \theta$ , para ángulos pequeños el movimiento se puede aproximar mediante

$$L\ddot{\theta} \approx -g\theta, \quad (1.4)$$

que tiene la sencilla solución conocida como oscilación armónica

$$\theta = \theta_0 \cos(\omega t), \quad (1.5)$$

con  $\theta_0$  la posición angular inicial a  $t = 0$  y  $\omega = \sqrt{g/L}$  la frecuencia angular. El período de la oscilación es  $T = 2\pi/\omega = 2\pi\sqrt{L/g}$ .

### 1.4.2. Solución numérica

Ahora resolveremos la ecuación (1.3) numéricamente. Para ello recurrimos al [método de Euler](#), un método de primer orden utilizado para resolver ecuaciones diferenciales ordinarias. Existen métodos muy superiores, pero el método explícito de Euler es el más simple y el más fácil de entender. Hay métodos más avanzados y precisos, como el método de cuarto orden de Runge-Kutta, pero ahora nos conformamos con el de Euler.

Necesitamos escribir la ecuación (1.3) en una forma que la computadora pueda evaluar más fácilmente. Comenzamos por desacoplar la ecuación diferencial de segundo orden en

dos ecuaciones de primer orden introduciendo la velocidad angular  $\omega = \dot{\theta}$ :

$$\begin{aligned} d\omega &= -\frac{g}{L} \sin \theta dt, \\ d\theta &= \omega dt. \end{aligned} \quad (1.6)$$

La computadora no puede manejar valores infinitesimales y, por lo tanto, necesitamos aproximarlos como finitos (pero pequeños)

$$dt \approx \Delta t, \quad d\omega \approx \Delta\omega \quad d\theta \approx \Delta\theta, \quad (1.7)$$

y cuanto menores sean estos valores, más preciso será el método. Para poner en marcha el algoritmo del método de Euler, comenzamos a  $t = 0$  sin velocidad inicial

$$\begin{aligned} \theta(0) &= \theta_0, \\ \omega(0) &= 0. \end{aligned} \quad (1.8)$$

Con los valores de  $\theta(0)$  y  $\omega(0)$ , es decir, con la información que tenemos para  $t = 0$ , la ecuación (1.6) permite evaluar las tasas de variación, o sea las derivadas  $d\theta/dt$  y  $d\omega/dt$ , de las variables de interés. Para estimar *las variaciones* que tendrían estas variables en un futuro instante muy cercano  $t = 0 + \Delta t$ , usamos las aproximaciones (1.7) y extrapolamos linealmente

$$\begin{aligned} \Delta\omega(0) &\approx -\frac{g}{L} \sin \theta(0) \Delta t \quad \longrightarrow \quad \omega(0 + \Delta t) \approx \omega(0) + \Delta\omega(0), \\ \Delta\theta(0) &\approx \omega(0) \Delta t \quad \longrightarrow \quad \theta(0 + \Delta t) \approx \theta(0) + \omega(0) \Delta t. \end{aligned} \quad (1.9)$$

Notar que este procedimiento equivale a moverse por la recta tangente a las funciones  $\theta(t)$  y  $\omega(t)$ . Y ahora, con los valores de las variables en  $t = \Delta t$ , la ecuación (1.6) permite evaluar nuevos valores de las tasas de variación de  $\theta$  y  $\omega$ . Y si volvemos a extrapolar linealmente, obtenemos estimaciones de las variables de interés en el instante  $t = 2\Delta t$

$$\begin{aligned} \Delta\omega(\Delta t) &\approx -\frac{g}{L} \sin \theta(\Delta t) \Delta t \quad \longrightarrow \quad \omega(\Delta t + \Delta t) \approx \omega(\Delta t) + \Delta\omega(\Delta t), \\ \Delta\theta(\Delta t) &\approx \omega(\Delta t) \Delta t \quad \longrightarrow \quad \theta(\Delta t + \Delta t) \approx \theta(\Delta t) + \omega(\Delta t) \Delta t. \end{aligned} \quad (1.10)$$

Repitiendo el mismo procedimiento, podemos encontrar los valores de  $\theta$  y  $\omega$  en cualquier instante  $t = j\Delta t$ , con  $j = 0, 1 \dots JMAX$ , de la siguiente manera

$$\begin{aligned} \Delta\omega|_j &\approx -\frac{g}{L} \sin \theta_j \Delta t \quad \longrightarrow \quad \omega_{j+1} \approx \omega_j + \Delta\omega|_j, \\ \Delta\theta|_j &\approx \omega_j \Delta t \quad \longrightarrow \quad \theta_{j+1} \approx \theta_j + \Delta\theta|_j. \end{aligned} \quad (1.11)$$

Así como lo planteamos, el valor de  $\Delta t$  se mantiene constante en todos los pasos, pero no habría en principio ningún inconveniente para cambiar  $\Delta t$ , adaptándolo según se esperen zonas de variaciones lentas (en este caso lo aumentaríamos, para ganar tiempo de cómputo), o zonas de variaciones rápidas (en este caso lo disminuiríamos, para ganar precisión).

### 1.4.3. Puesta en marcha

Python va a ser el idioma que vamos a usar para comunicarnos con la computadora. Si queremos que la computadora repita sin protestar los aburridos pasos del método de Euler, es muy conveniente subdividir el algoritmo en tareas que se puedan explicar de la manera mas sencilla posible. Dicho de otra manera, hay que aplicar la conocida estrategia del “divide y vencerás”.

Las tareas sencillas, una vez programadas, se conocen como funciones, tienen argumentos como entrada y dan datos en formatos muy versátiles como salida. En una función, la primera cadena literal que se encuentra es interpretada como documentación, sin importar si se usan triples comillas, un par de comillas dobles o simples o el símbolo `#` que se usa para incluir comentarios: todo lo que está después del `#` es un comentario. Es muy importante tener en cuenta que el bloque de una función debe estar separado del margen izquierdo mediante una sangría (indentación es un anglicismo pero es mas usado).

Antes de empezar se importan librerías especializadas que amplían el poder de Python, por ejemplo como primera celda vamos a poner

```
# Primera celda de mi primer cuaderno Colab
# Curso de F2, 2do cuatrimestre virtual 2020
import numpy as np          #numpy = Numerical Python
import matplotlib.pyplot as plt #graficos
plt.style.use('bmh')        #estilo
figsize = (12, 5)           #12x5 pulgadas
dpi = 600                   #Resolucion de figura
```

El procesador de textos que estoy usando (L<sup>A</sup>T<sub>E</sub>X) reconoce que estoy escribiendo código Python y colorea en azul las palabras reservadas para comandos del lenguaje.

Ahora, en otra celda introducimos los datos y codificamos las funciones en que dividimos el procedimiento descrito por (1.11)

```
# segunda celda de mi primer cuaderno Colab
# Curso de F2, 2do cuatrimestre virtual 2020
g = 9.81 # m/s^2
L = 25   # m
m = 40   # kg

def approx(t, theta0):
    """ para evaluar el resultado analitico aproximado
        """
    return theta0*np.cos(t*(g/L)**.5)

def RHS(theta, w, dt):
```

```

    """ devuelve el lado derecho de la
    ec. dif. ordinaria (1.6)
    """
    dw = -np.sin(theta)*dt*g/L
    dtheta = w*dt
    return dtheta, dw

def euler_step(theta, w, dt):
    """ avanza un paso del algoritmo de Euler """
    dtheta, dw = RHS(theta, w, dt)
    w = w + dw
    theta = theta + dtheta
    return theta, w

def euler_method(theta0, w0, dt, n):
    """ Integra Euler en 0 <= t <= T (intervalo) """
    theta = (n + 1)*[0]
    w = (n + 1)*[0]

    theta[0] = theta0
    w[0] = w0
    for i in range(n):
        theta[i + 1], w[i + 1] = \
            euler_step(theta[i], w[i], dt)

    return theta, w

```

La barra inclinada a la izquierda es una continuación explícita de línea, si consultan un manual verán que se puede obviar bajo ciertas circunstancias (acá no hace falta). Y ahora usemos el método de Euler y grafiquemos el resultado. Para comparar el resultado numérico con la aproximación analítica válida para ángulos pequeños, tomamos dos ángulos iniciales diferentes  $\theta = \{15^\circ, 60^\circ\} = \{\pi/12, \pi/3\}$

```

# tercera celda de mi primer cuaderno Colab
# Curso de F2, 2do cuatrimestre virtual 2020
theta01 = np.pi/12
theta02 = np.pi/3
T = 20
n = 10000
t = np.linspace(0, T, n + 1)
dt = T/float(n)

theta1, w1 = euler_method(theta01, 0, dt, n)

```

```
theta2, w2 = euler_method(theta02, 0, dt, n)
```

Las variables `theta1` y `w1` contienen los valores de  $\theta(t)$  y  $\dot{\theta}(t)$  obtenidos para  $\theta(0) = 15^\circ$  y discretizados en la grilla de tiempos creada con el comando `np.linspace(0, T, n + 1)`. Análogamente las variables `theta2` y `w2` contienen los valores de  $\theta(t)$  y  $\dot{\theta}(t)$  obtenidos para  $\theta(0) = 60^\circ$ , también discretizados en la misma grilla. Para obtener los gráficos de la variable angular  $\theta(t)$  en función del tiempo, usamos los siguientes comandos contenidos en una nueva celda

```
# cuarta celda de mi primer cuaderno Colab
# Curso de F2, 2do cuatrimestre virtual 2020
plt.figure(figsize=figsize, dpi=dpi)
plt.title("Angular position")
plt.plot(t, theta1, "m", label=r"$\theta_0=0f^\circ$"%(
    theta01*180/np.pi))
plt.plot(t, approx(t, theta01), "m--", label=r"Aproximado
")
plt.plot(t, theta2, "g", label=r"$\theta_0=0f^\circ$"%(
    theta02*180/np.pi))
plt.plot(t, approx(t, theta02), "g--", label=r"Aproximado
")
plt.xlabel(r"$t$, [s]")
plt.ylabel(r"$\theta(t)$, [rad]")
plt.legend()
plt.show()
```

Matplotlib es una biblioteca muy completa para hacer visualizaciones estáticas, animadas o interactivas en Python. Por ser tan versátil, hay muchísimas maneras de encarar un gráfico. Acá conviene saber solamente que el comando central es `plt.plot`, que grafica el primer argumento en las abscisas (eje horizontal) y el segundo argumento en las ordenadas (eje vertical). Se recomienda cargar a mano las celdas anteriores en un nuevo cuaderno Colab o tratar de ejecutar el cuaderno que se encuentra en [este enlace](#).

Observar que para  $\theta_0 = 15^\circ$  la aproximación  $\theta \ll 1$  es bastante buena, pero se vuelve menos precisa a medida que aumenta  $\theta_0$ .

#### 1.4.4. Conservación de la energía

Como no hemos incluido mecanismos disipativos, la energía mecánica total del sistema

$$E = U + K = mgL(1 - \cos \theta) + \frac{1}{2}mL^2\dot{\theta}^2, \quad (1.12)$$

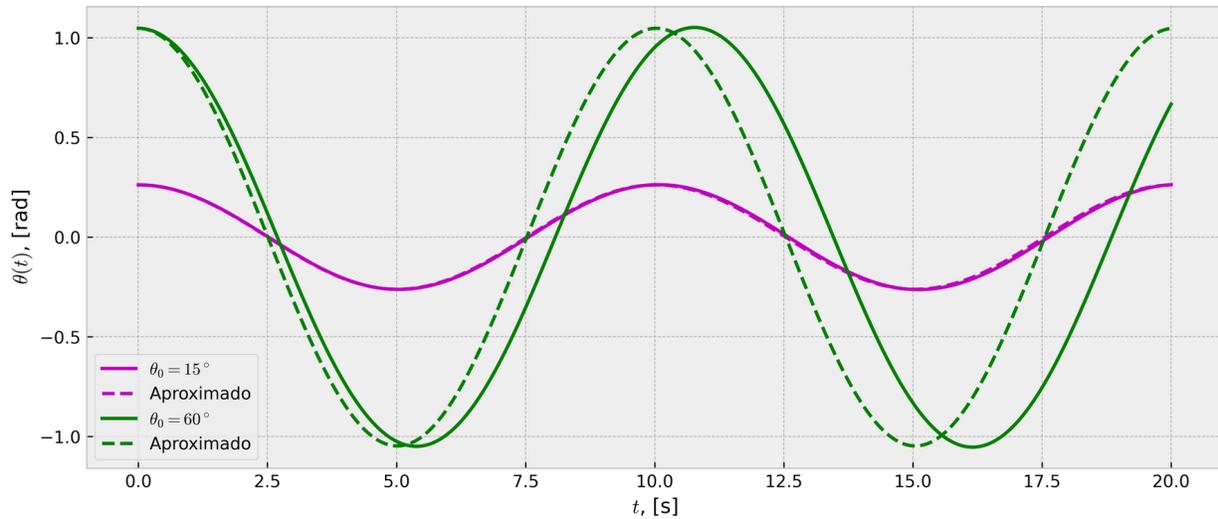


Figura 1.2: Comparación entre resultados numéricos obtenidos para  $\theta(t)$  con el método de Euler y con la aproximación analítica válida para ángulos pequeños para casos con ángulos iniciales  $\theta = 15^\circ$  y  $\theta = 60^\circ$ .

debe conservarse. Esta ley física sirve entonces como indicador de la calidad del algoritmo y de los valores asignados a parámetros numéricos, como  $dt$  (o sea,  $\Delta t$ ), que fueron elegidos bastante arbitrariamente. Para calcular la energía, introducimos una celda nueva con dos funciones Python

```
# quinta celda de mi primer cuaderno Colab
# Curso de F2, 2do cuatrimestre virtual 2020
def get_U(theta):
    """ Energia Potencial U """
    return m*g*L*(1 - np.cos(theta))

def get_K(w):
    """ Energia cinetica K """
    return 0.5*m*L**2*np.array(w)**2
```

y otra celda nueva para hacer los gráficos

```
# sexta celda de mi primer cuaderno Colab
# Curso de F2, 2do cuatrimestre virtual 2020
plt.figure(figsize=figsize, dpi=dpi)
plt.title(r"Energia mecanica, $\theta_0=%.0f^\circ$"%(
    theta0*180/np.pi))
plt.plot(t, get_U(theta2), label=r"Energia potencial")
plt.plot(t, get_K(w2), label=r"Energia cinetica")
```

```
plt.plot(t, get_U(theta2) + get_K(w2), label=r"Energía
total")
plt.xlabel(r"$t$, [s]")
plt.ylabel(r"$E$, [J]")
plt.legend(loc=0) #0 mejor, 1 arriba izquierda
from google.colab import files
plt.savefig("penduloeuler01energ.png")
files.download("penduloeuler01energ.png")
plt.show()
```

Colab trabaja en la nube de Google, pero no guarda las figuras en sus servidores, por eso es necesario importar el comando `file` desde la biblioteca especial `google.colab` y así poder guardar la figura como archivo de imagen, en este caso en formato `.png`. La

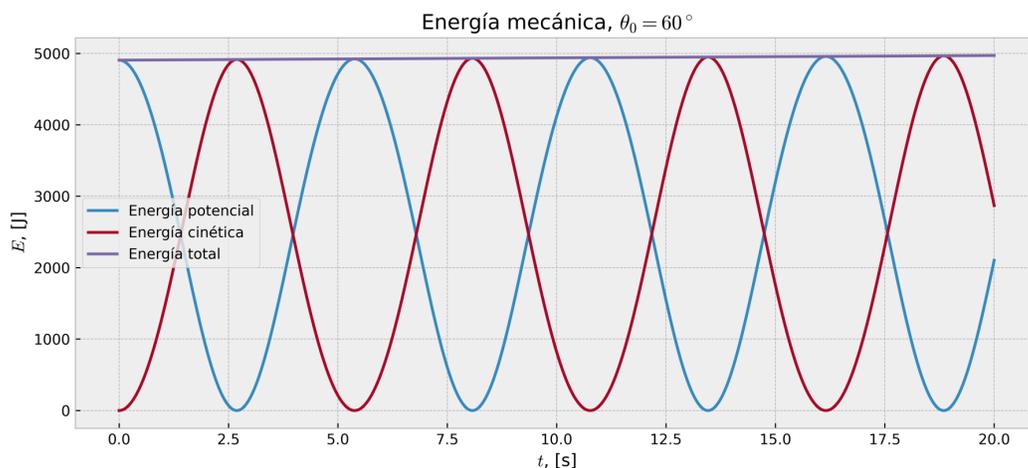


Figura 1.3: Energías  $E$ ,  $K$  y  $U$  en función del tiempo para  $\theta_0 = 60^\circ$ .

figura 1.3 muestra las curvas de energía cinética, potencial y mecánica total en función del tiempo para  $\theta_0 = 60^\circ$ . Observamos una pequeña variación de la energía mecánica que parece indicar que el valor de  $\Delta t$  sería bastante satisfactorio. Para ser un poco más precisos, evaluemos numéricamente la diferencia entre la energía mecánica total a  $t=0$  (el primer punto de la grilla temporal) y la energía mecánica total al tiempo final (el último punto de la grilla temporal) en las dos condiciones iniciales consideradas ( $\theta_0 = 15^\circ$  y  $\theta_0 = 60^\circ$ )

```
# 7ma celda de mi primer cuaderno Colab
# Curso de F2, 2do cuatrimestre virtual 2020
def get_error(theta, w):
    """ Chequea error relativo en E=cte """
    E0 = get_U(theta[0]) + get_K(w[0])
    E1 = get_U(theta[-1]) + get_K(w[-1])
```

```

    return np.abs((E0 - E1)/E0)

print("Cambio relativo en E:")
print("Theta = %.0f: %.2e"%(theta01*180/np.pi, get_error(
    theta1, w1)))
print("Theta = %.0f: %.2e"%(theta02*180/np.pi, get_error(
    theta2, w2)))

```

En ambos casos, para  $n=10000$  el error relativo en la conservación de la energía se mantiene por debajo del 2% mientras que para  $n=100000$  el error relativo en la conservación de la energía disminuye y se mantiene por debajo del 0.2%.

### 1.4.5. Péndulo con rozamiento

Debido a la fricción, los péndulos reales no oscilan eternamente, si no que se van frenando hasta detenerse. A velocidades bajas, la fuerza de fricción producida por el aire es proporcional a la velocidad,  $f = -b\dot{\theta}$  y entonces la ecuación diferencial que describe el movimiento se escribe

$$L\ddot{\theta} - \frac{b}{m}\dot{\theta} + g \sin \theta = 0. \quad (1.13)$$

Igual que en el caso anterior, la ecuación (1.13) no tiene solución analítica, excepto en el caso de pequeñas oscilaciones  $\theta \ll 1$ . Para velocidad inicial nula, la solución tiene la forma

$$\theta(t) = C e^{-t/2\tau} \cos(\omega_1 t), \quad (1.14)$$

donde  $\Gamma = b/m$ ,  $\tau = 1/\Gamma$ ,  $\omega_1^2 = \omega_0^2 - \frac{\Gamma^2}{4}$  y  $\omega_0$  es la frecuencia angular cuando no hay rozamiento. La forma aproximada de esta ecuación se obtiene con el mismo procedimiento que usamos para la ecuación (1.3)

$$\begin{aligned} \Delta\omega &= -\frac{g}{L} \sin \theta \Delta t + \frac{\Gamma}{L}\omega, \\ \Delta\theta &= \omega \Delta t. \end{aligned} \quad (1.15)$$

Comparando con la ecuación (1.9), vemos que la única diferencia es cómo se evalúa  $\Delta\omega$  en cada paso temporal y entonces podemos aprovechar todo lo programado anteriormente, cambiando solamente la definición del código Python que hace avanzar el algoritmo de Euler en cada paso temporal, o sea, solo hay que cambiar la función RHS de la siguiente manera

```

# 8va celda de mi primer cuaderno Colab
# Curso de F2, 2do cuatrimestre virtual 2020

def RHS(theta, w, dt):
    """ devuelve el lado derecho de la ec.

```

```

del oscilador armonico con rozamiento """
dw = -np.sin(theta)*dt*g/L - gamma*w*dt/L
dtheta = w*dt
return dtheta, dw

```

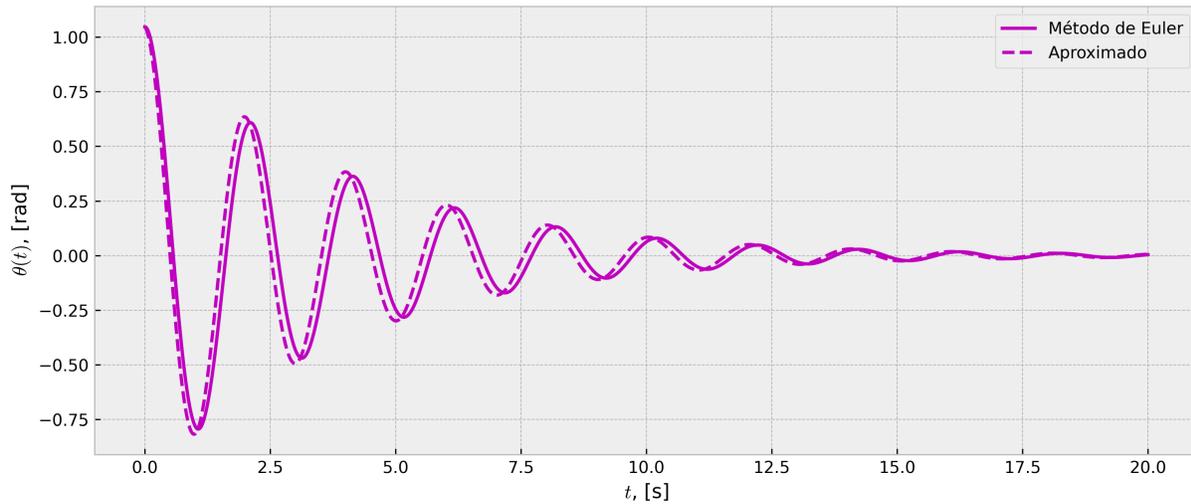


Figura 1.4: Comparación entre los resultados numéricos obtenidos para  $\theta(t)$  con el método de Euler y los resultados aproximados dados por la ecuación (1.14). Los parámetros son  $\theta_0 = 60^\circ$ ,  $\dot{\theta}_0 = 0$ ,  $m=1\text{kg}$ ,  $L = 1\text{m}$  y  $\Gamma = .5\text{m}$ .

Ahora podemos comparar los resultados obtenidos con el método de Euler con los predichos por la ecuación aproximada (1.14). A diferencia de lo obtenido en la figura 1.2, en el caso  $\theta_0 = 60^\circ$  aún se observa una coincidencia bastante aceptable (ver figura 1.4). ¿A qué se debe esta diferencia, teniendo en cuenta que la condición inicial no verifica la aproximación  $\theta \ll 1$ ? A continuación se transcribe el código usado en este ejemplo

```

# 9na celda de mi primer cuaderno Colab
# Curso de F2, 2do cuatrimestre virtual 2020

g = 9.8      # m/s^2
m = 1.      # kg -- masa
L = 1.      # m. longitud

w0 = 0      # 1/s -- velocidad angular inicial
theta0 = np.pi/3 # rad  angulo inicial

T = 20.     # s intervalo de simulacion
n = 100000 # nro de pasos
b = .5     # kg m. factor rozamiento
gammab=b/m # m

```

```

omega1=(g/L-gammab**2/4)**(1/2)
tau=1/gammab

t = np.linspace(0, T, n + 1)
theta, _ = euler_method(theta0, w0, T/float(n), n)
theta_aprox = theta0*np.exp(-t/(2*tau))*np.cos(omega1*t)

plt.figure(figsize=figsize, dpi=dpi)
plt.plot(t, theta, "m", label=r"Metodo de Euler")
plt.plot(t, theta_aprox, "m--", label=r"Aproximado")
plt.xlabel(r"$t$, [s]")
plt.ylabel(r"$\theta(t)$, [rad]")
plt.legend(loc=0) #0 mejor, 1 arriba izquierda
plt.show()

```

Exploremos ahora un caso donde la aproximación  $\theta \ll 1$  seguramente no sería válida. Llevemos la masa hasta una posición inicial  $\theta_0 = 3$ , casi coincidente con la vertical, y empujémosla con una velocidad inicial  $\dot{\theta}_0 = 10s^{-1}$ . ¿Cómo se interpreta el resultado que se muestra en la figura 1.5? Primero, notar que hay un intervalo donde el crecimiento de la posición angular con el tiempo es casi lineal, lo que indica un predominio de la energía cinética sobre la energía potencial máxima. Si no hubiera rozamiento, en estas condiciones el período del péndulo sería infinito, porque la partícula nunca quedaría atrapada en el pozo de potencial alrededor de la vertical. Pero como el trabajo realizado por la fuerza de rozamiento disminuye la energía total del sistema, llega un momento en que la energía cinética comienza a ser del mismo orden que la energía potencial máxima, hasta que la energía total toma un valor que obliga al sistema a quedarse oscilando alrededor de la posición de equilibrio estable, que en esta descripción corresponde al valor  $\theta = 6\pi$  mientras sigue perdiendo energía. Aquí el comportamiento del péndulo empieza a ser similar al visto anteriormente en la figura 1.4. Se deja como ejercicio comprobar que los tiempos característicos (período temporal y tiempo de decaimiento) en esta zona se pueden predecir con los valores que se obtienen analíticamente a partir de la ecuación aproximada (1.14) y que con estas condiciones iniciales el período tiende a infinito a medida que disminuye el valor de  $b$ .

Una manera muy recomendada de visualizar resultados de este curso es mediante animaciones. Hay varias maneras de hacerlas, acá dejo una receta que pueden adaptar para otros problemas. El archivo de video que corresponde a las condiciones de la figura 1.5 se puede ver en [este](#) link.

```

x = np.sin(theta)
y = -np.cos(theta)

from matplotlib import animation
from IPython.display import HTML

```

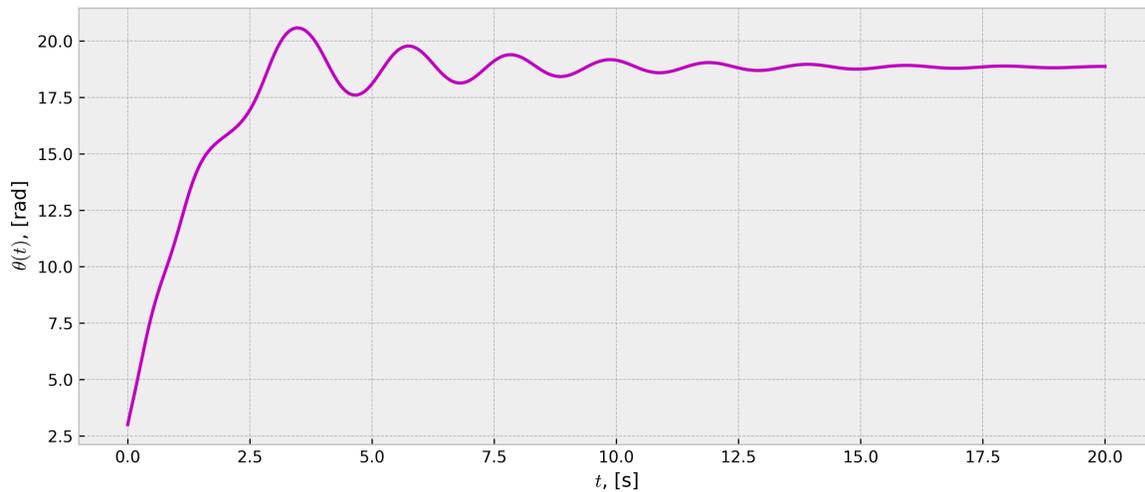


Figura 1.5: Resultados obtenidos con el método de Euler para  $\theta_0 = 3$ ,  $\dot{\theta}_0 = 10s^{-1}$  y el resto de los parámetros como en la figura 1.4.

```
FPS=30
plt.style.use('default')

# Define la figura
fig = plt.figure(figsize=(14, 14), dpi=60)
ax = plt.axes(xlim=(-1.1, 1.1), ylim=(-1.1, 1.1))
ax.set_aspect('equal')
ax.axis('off')

# Define elementos de la animacion
rod, = ax.plot([], [], color="grey", linewidth=2)
ball = plt.Circle((x[0], y[0]), 0.1, fc="grey")
ax.add_patch(ball)

# Calcula nro. de frames
framesNum = int(FPS*t[-1])

# Funcion que define animacion
def animate(j):
    i = j*int(n/framesNum)
    ball.center = (x[i], y[i])
    rod.set_data([0, x[i]], [0, y[i]])

# Crear animacion
anim = animation.FuncAnimation(fig, animate, frames=
    framesNum, interval=1000/FPS)
```

```
plt.close(anim._fig)

# Mostrar animacion
HTML(anim.to_html5_video())
```



# Capítulo 2

## Oscilaciones libres y modos normales

### 2.1. Dos grados de libertad

- El primer objetivo de este primer ejemplo es mostrar la nueva dificultad que aparece en sistemas con dos partes móviles que interactúan entre si. Es claro que llegaremos a dos ecuaciones de movimiento (una para cada parte móvil). La novedad es que la ecuación de movimiento para una parte móvil no se puede resolver sin tener antes la solución para la otra parte móvil. Es decir, las dos ecuaciones diferenciales que representan la ley de Newton para cada parte móvil están **acopladas**.
- El segundo objetivo será entonces encontrar un método que permita tratar el sistema de dos **ecuaciones diferenciales acopladas**.
- El tercer objetivo es introducir los conceptos de **coordenadas normales, modos normales y frecuencias normales**.
- Y el cuarto objetivo será poner en evidencia que **el movimiento más general** de este sistema con  $N = 2$  grados de libertad que se perturba y se deja evolucionar a partir del equilibrio es una **superposición de  $N = 2$  movimientos oscilatorios armónicos**. Luego veremos que esta característica se generaliza a cualquier sistema de  $N$  grados de libertad y con fuerzas de interacción lineales con los desplazamientos de la posición de equilibrio.

#### 2.1.1. Primer ejemplo con $N = 2$

Consideremos la configuración de la figura 2.1, donde se muestra el sistema en la posición de equilibrio (arriba) y el sistema perturbado (abajo). Las dos masas son idénticas

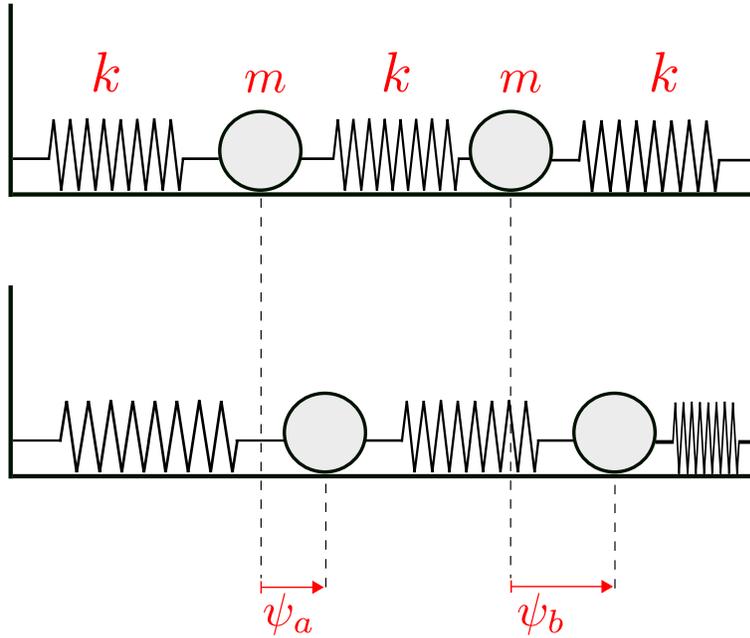


Figura 2.1: Primer ejemplo de 2GL

(masa  $m$ ) y los resortes tienen la misma constante elástica  $k$  y la misma longitud natural que en el equilibrio. No hay rozamiento y los movimientos ocurren solo en la dirección horizontal. Elegimos describir el movimiento con las coordenadas  $\psi_a(t)$  y  $\psi_b(t)$ , que dan la posición de cada parte móvil en el instante  $t$ . Las dos ecuaciones diferenciales que representan la ley de Newton para cada parte móvil son

$$\begin{aligned}\ddot{\psi}_a &= -\frac{k}{m}\psi_a + \frac{k}{m}(\psi_b - \psi_a), \\ \ddot{\psi}_b &= -\frac{k}{m}(\psi_b - \psi_a) - \frac{k}{m}\psi_b,\end{aligned}\tag{2.1}$$

cada término del miembro derecho en cada ecuación está asociado a una fuerza. El sistema se puede reescribir como

$$\begin{aligned}\ddot{\psi}_a &= -2\frac{k}{m}\psi_a + \frac{k}{m}\psi_b, \\ \ddot{\psi}_b &= \frac{k}{m}\psi_a - 2\frac{k}{m}\psi_b,\end{aligned}\tag{2.2}$$

De paso notar que, usando matrices y vectores columna, el sistema se puede representar simbólicamente como

$$\frac{d^2}{dt^2} \begin{pmatrix} \psi_a \\ \psi_b \end{pmatrix} = - \begin{pmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{pmatrix} \begin{pmatrix} \psi_a \\ \psi_b \end{pmatrix},\tag{2.3}$$

donde cambiamos la notación para las derivadas segundas y definimos  $M_{11} = 2\frac{k}{m}$ ;  $M_{12} = -\frac{k}{m}$ ;  $M_{21} = -\frac{k}{m}$  y  $M_{22} = 2\frac{k}{m}$ .

El sistema de la ecs. (2.2) es un sistema homogéneo de 2 ecuaciones diferenciales lineales, acopladas, de segundo orden que tiene como incógnitas a las funciones  $\psi_a(t)$  y  $\psi_b(t)$ . Si se definen dos nuevas incógnitas  $\psi_1(t)$  y  $\psi_2(t)$  de la forma

$$\begin{aligned}\psi_1 &= \psi_a + \psi_b, \\ \psi_2 &= \psi_a - \psi_b,\end{aligned}\tag{2.4}$$

la simetría que tiene el sistema físico (masas iguales, resortes iguales) permite ver fácilmente, simplemente sumando y restando miembro a miembro la ecs. (2.2), que las nuevas incógnitas  $\psi_1(t)$  y  $\psi_2(t)$  satisfacen las siguientes ecuaciones diferenciales **desacopladas**

$$\begin{aligned}\ddot{\psi}_1 + \frac{k}{m}\psi_1 &= 0, \\ \ddot{\psi}_2 + 3\frac{k}{m}\psi_2 &= 0.\end{aligned}\tag{2.5}$$

¡Qué bueno! Porque  $\psi_1(t)$  y  $\psi_2(t)$  no sólo satisfacen ecuaciones diferenciales desacopladas, sino que cada ecuación en el sistema (2.5) tiene la misma forma que la ecuación del oscilador armónico. ¡Genial! Porque ahora podemos usar lo que sabíamos del repaso de F1 y asegurar que las funciones incógnitas  $\psi_1(t)$  y  $\psi_2(t)$  son funciones armónicas, de la forma

$$\begin{aligned}\psi_1(t) &= A_1 \cos(\omega_1 t + \varphi_1), \\ \psi_2(t) &= A_2 \cos(\omega_2 t + \varphi_2),\end{aligned}\tag{2.6}$$

con

$$\omega_1^2 = \frac{k}{m}, \quad \omega_2^2 = 3\frac{k}{m},\tag{2.7}$$

y  $A_1$ ,  $A_2$ ,  $\varphi_1$  y  $\varphi_2$  cuatro constantes que todavía no sabemos, pero que tenemos mucha confianza que vamos a poder encontrar, porque sabemos muy bien que hasta ahora nunca usamos las posiciones y velocidades iniciales de cada parte móvil, es decir,  $\psi_a(0)$ ,  $\psi_b(0)$ ,  $\dot{\psi}_a(0)$  y  $\dot{\psi}_b(0)$ . Estos son cuatro datos físicos que seguramente tenemos que saber si sabemos cómo se perturbó el sistema. Antes de encontrar  $A_1$ ,  $A_2$ ,  $\varphi_1$  y  $\varphi_2$  en función de  $\psi_a(0)$ ,  $\psi_b(0)$ ,  $\dot{\psi}_a(0)$  y  $\dot{\psi}_b(0)$ , quedémonos pensando en el hermoso resultado que hemos obtenido: en este sistema el movimiento más general de cada parte móvil (dado por  $\psi_a(t)$  y  $\psi_b(t)$ ) es una combinación lineal de dos funciones armónicas con frecuencias características  $\omega_1$  y  $\omega_2$ . Esto es así porque la ec. (2.4) muestra que  $\psi_a(t)$  y  $\psi_b(t)$  son combinaciones lineales de  $\psi_1(t)$  y  $\psi_2(t)$  y hemos demostrado que  $\psi_1(t)$  y  $\psi_2(t)$  son funciones armónicas dadas por las ecs. (2.6).

Las nuevas funciones  $\psi_1(t)$  y  $\psi_2(t)$  que desacoplan las ecuaciones de movimiento del sistema se llaman **coordenadas normales** y las frecuencias características  $\omega_1$  y  $\omega_2$  se llaman **frecuencias normales** o también **frecuencias de resonancia** (¿por qué les habrán dado este nombre?).

### 2.1.2. Condiciones iniciales

Para encontrar  $A_1$ ,  $A_2$ ,  $\varphi_1$  y  $\varphi_2$  (ver ec. (2.6)) en función de  $\psi_a(0)$ ,  $\psi_b(0)$ ,  $\dot{\psi}_a(0)$  y  $\dot{\psi}_b(0)$  imponemos las condiciones iniciales. Según las ecs. (2.4) y (2.6)

$$\begin{aligned}\psi_1(0) &= A_1 \cos \varphi_1 = \psi_a(0) + \psi_b(0), \\ \dot{\psi}_1(0) &= -\omega_1 A_1 \sin \varphi_1 = \dot{\psi}_a(0) + \dot{\psi}_b(0), \\ \psi_2(0) &= A_2 \cos \varphi_2 = \psi_a(0) - \psi_b(0), \\ \dot{\psi}_2(0) &= -\omega_2 A_2 \sin \varphi_2 = \dot{\psi}_a(0) - \dot{\psi}_b(0).\end{aligned}\tag{2.8}$$

Elevando al cuadrado y sumando miembro a miembro, de las dos primeras ecuaciones se obtiene  $A_1^2$  y de las otras dos ecuaciones se obtiene  $A_2^2$

$$\begin{aligned}A_1^2 &= \left[ \psi_a(0) + \psi_b(0) \right]^2 + \frac{1}{\omega_1^2} \left[ \dot{\psi}_a(0) + \dot{\psi}_b(0) \right]^2, \\ A_2^2 &= \left[ \psi_a(0) - \psi_b(0) \right]^2 + \frac{1}{\omega_2^2} \left[ \dot{\psi}_a(0) - \dot{\psi}_b(0) \right]^2.\end{aligned}\tag{2.9}$$

Dividiendo miembro a miembro las dos primeras ecuaciones se obtiene  $\tan \varphi_1$  y dividiendo miembro a miembro las otras dos ecuaciones se obtiene  $\tan \varphi_2$

$$\begin{aligned}\tan \varphi_1 &= -\frac{1}{\omega_1} \frac{\dot{\psi}_a(0) + \dot{\psi}_b(0)}{\psi_a(0) + \psi_b(0)}, \\ \tan \varphi_2 &= -\frac{1}{\omega_2} \frac{\dot{\psi}_a(0) - \dot{\psi}_b(0)}{\psi_a(0) - \psi_b(0)},\end{aligned}\tag{2.10}$$

y como las ecs. (2.8) dan los signos de los senos y cosenos, se obtienen  $\varphi_1$  y  $\varphi_2$  en el intervalo  $[0, 2\pi]$ . Si las velocidades iniciales son cero,  $\varphi_1 = \varphi_2 = 0$ .

Invirtiendo la ec. (2.4) para obtener las viejas coordenadas  $\psi_a(t)$  y  $\psi_b(t)$  en función de  $\psi_1(t)$  y  $\psi_2(t)$ , vemos que cuando el sistema de la Fig. 2.1 se aparta del equilibrio, su movimiento más general resulta una superposición de dos movimientos armónicos simples

$$\begin{aligned}\psi_a(t) &= \frac{1}{2} \left[ A_1 \cos(\omega_1 t + \varphi_1) + A_2 \cos(\omega_2 t + \varphi_2) \right], \\ \psi_b(t) &= \frac{1}{2} \left[ A_1 \cos(\omega_1 t + \varphi_1) - A_2 \cos(\omega_2 t + \varphi_2) \right],\end{aligned}\tag{2.11}$$

con  $\omega_1$  y  $\omega_2$  dados por (2.7),  $A_1$  y  $A_2$  por (2.9), y  $\varphi_1$  y  $\varphi_2$  por (2.10).

### 2.1.3. Los modos normales

Hasta acá el cambio de coordenadas dado por la ec. (2.4) ha sido muy útil para pasar de un sistema diferencial acoplado para las viejas coordenadas  $\psi_a(t)$  y  $\psi_b(t)$  a un par

de ecuaciones diferenciales desacopladas para las nuevas coordenadas  $\psi_1(t)$  y  $\psi_2(t)$ . Las nuevas coordenadas realizan movimientos armónicos simples, pero según (2.11) las partes móviles no realizan un movimiento armónico simple sino una superposición de dos movimientos armónicos simples.

Como físicos curiosos no solamente es normal sino que es casi obligatorio que nos preguntemos si es posible observar el movimiento armónico simple en las partes móviles. Para que esto suceda es necesario que una de las amplitudes  $A_1$  o  $A_2$  se anule. Las ecs. (2.9) muestran que esto es posible si se elijen las condiciones iniciales en dos casos distintos

- **Modo 1:**  $A_2 = 0$ , cuando  $\psi_a(0) = \psi_b(0)$  y  $\dot{\psi}_a(0) = \dot{\psi}_b(0)$ .  
En este caso  $\psi_a(t) = \psi_b(t) = A_1 \cos(\omega_1 t + \varphi_1)$ .
- **Modo 2:**  $A_1 = 0$ , cuando  $\psi_a(0) = -\psi_b(0)$  y  $\dot{\psi}_a(0) = -\dot{\psi}_b(0)$ .  
En este caso  $\psi_a(t) = -\psi_b(t) = A_2 \cos(\omega_2 t + \varphi_2)$ .

Observar que el modo 1 describe el movimiento del centro de masa del sistema mientras que el modo 2 describe el movimiento interno, a través de la distancia relativa entre masas  $\psi_a(t) - \psi_b(t)$ . Y también observar que  $\omega_1^2$  es la fuerza restitutiva por unidad de desplazamiento y por unidad de masa sobre cada parte móvil cuando el sistema se mueve en el modo 1, con idéntica consideración para  $\omega_2^2$  (característica que excede el ejemplo particular que estamos viendo).

Vemos entonces que mediante una elección adecuada de las condiciones iniciales podemos hacer que todas las partes móviles de este sistema oscilen con un movimiento armónico simple de frecuencia  $\omega_1$  (modo 1) o de frecuencia  $\omega_2$  (modo 2). Esto quiere decir que podemos ver experimentalmente los modos normales de este sistema sin saber matemáticas o sin mencionar cambio de variables ni desacoples, simplemente jugando para ver cuándo *todas las partes móviles del sistema oscilan con la misma frecuencia y pasan por la posición de equilibrio simultáneamente*. Justamente esta parte resaltada es la definición de *modo normal* de todo sistema físico, no solamente sistemas mecánicos. Bajo ciertas circunstancias, a los sistemas físicos, cuando se los aparta del equilibrio, “les gusta” oscilar colectivamente. La cantidad de maneras (o modos) en los que les gusta oscilar depende del número de grados de libertad del sistema, un sistema de 2 GL, como el del ejemplo, tiene 2 modos y veremos que un sistema con  $N$  grados de libertad tiene  $N$  modos.

#### 2.1.4. Modos normales y forzado estacionario

¿Cómo explorar experimentalmente los modos normales de un sistema? De acuerdo con lo discutido en el párrafo anterior, el método consistiría en usar distintas condiciones iniciales para perturbar el sistema a partir de la posición de equilibrio y así, mediante prueba y error, encontrar determinadas condiciones iniciales muy especiales, para las cuales todas las partes móviles del sistema se mueven con la misma frecuencia y pasan por

la posición de equilibrio simultáneamente (en nuestro ejemplo estas condiciones iniciales especiales son justamente las que dan  $A_2 = 0$  (modo 1) y  $A_1 = 0$  (modo 2)).

Pero hay otra manera distinta que no necesita de una elección especial de condiciones iniciales y está basada en forzar el sistema con una fuerza impulsora externa armónica de frecuencia  $\omega$ . Si suponemos que hay una pequeña disipación, imprescindible para amortiguar el movimiento transitorio pero suficientemente pequeña como para que pueda despreciarse en los cálculos, llegará un momento en que el sistema se habrá olvidado de las condiciones iniciales y se moverá siguiendo a la fuerza impulsora. Debido al fenómeno de resonancia, este movimiento, que en el problema forzado corresponde a la solución estacionaria, es notablemente apreciable cuando la frecuencia de la fuerza impulsora toma un conjunto de valores que son justamente los de las frecuencia de los modos normales y es por este motivo que los modos normales reciben también el nombre de modos resonantes. Como se ve, este segundo método tiene la ventaja de no requerir condiciones iniciales especiales, un punto especialmente importante cuando se estudian ciertos sistemas físicos donde no es fácil modificar experimentalmente las condiciones iniciales.

La existencia de dos métodos experimentales para estudiar los modos normales, pone en evidencia la íntima relación que existe entre el [problema de modos](#), es decir, el problema de encontrar la evolución del sistema cuando se lo aparta del equilibrio sujeto a sus propias fuerzas restitutivas, y el [problema forzado estacionario](#), es decir el problema de encontrar cuán bien oscila el sistema a la frecuencia de la fuerza impulsora. En ambos casos, todas las partes móviles del sistema oscilan con la misma frecuencia y pasan por la posición de equilibrio simultáneamente. Sin embargo, en el primer caso, la frecuencia es una característica emergente del sistema de partes móviles, mientras que en el segundo caso, la frecuencia viene impuesta por la fuerza impulsora.

### 2.1.5. Energía cinética y potencial

En el sistema de nuestro ejemplo, la energía total se conserva, debido que no tuvimos en cuenta mecanismos disipativos, como la fricción entre las masas y el piso, o con el aire. Sin embargo, cuando el sistema se aparta del equilibrio con condiciones iniciales arbitrarias, siempre habrá un juego continuo entre la energía cinética y potencial del sistema. La energía cinética total del sistema está en las masas, que van cambiando su velocidad, mientras que la energía potencial total está en los resortes, que van cambiando su longitud. Tanto la energía cinética total

$$T(t) = \frac{1}{2}m[\dot{\psi}_a]^2 + \frac{1}{2}m[\dot{\psi}_b]^2 \quad (2.12)$$

como la energía potencial total

$$V(t) = \frac{k}{2}[\psi_a]^2 + \frac{k}{2}[\psi_b - \psi_a]^2 + \frac{k}{2}[\psi_b]^2, \quad (2.13)$$

son funciones del tiempo. Notar que mientras  $T(t)$  resulta una forma cuadrática en las velocidades de las partes móviles, sin términos cruzados,  $V(t)$  es una forma cuadrática

en los desplazamientos de las partes móviles, pero contiene términos cruzados. Usando la transformación (2.4) es fácil demostrar que

$$T(t) = \frac{m}{4}[\dot{\psi}_1]^2 + \frac{m}{4}[\dot{\psi}_2]^2, \quad (2.14)$$

y

$$V(t) = \frac{k}{4}[\psi_1]^2 + \frac{3k}{4}[\psi_2]^2, \quad (2.15)$$

es decir, que la energía cinética  $T(t)$  y la energía potencial  $V(t)$  del sistema, expresadas en términos de las coordenadas normales  $\psi_1(t)$  y  $\psi_2(t)$ , son formas cuadráticas sin términos cruzados. Esta característica subraya el rol físico que tienen las coordenadas normales: *en cuanto a la energía total del sistema se refiere, todo ocurre como si los modos del sistema se excitaran separadamente.*

Las expresiones (2.14) y (2.15) se pueden reescribir usando la dependencia temporal de las coordenadas normales dada por (2.6)

$$T(t) = \frac{m\omega_1^2}{4}A_1^2 \sin^2(\omega_1 t + \varphi_1) + \frac{m\omega_2^2}{4}A_2^2 \sin^2(\omega_2 t + \varphi_2), \quad (2.16)$$

y

$$V(t) = \frac{k}{4}A_1^2 \cos^2(\omega_1 t + \varphi_1) + \frac{3k}{4}A_2^2 \cos^2(\omega_2 t + \varphi_2). \quad (2.17)$$

Usando (2.7), la energía total del sistema  $E = T(t) + V(t)$  resulta

$$E = \frac{m\omega_1^2}{4}A_1^2 + \frac{m\omega_2^2}{4}A_2^2, \quad (2.18)$$

es decir, independiente del tiempo, de acuerdo con nuestra suposición original de excluir mecanismos disipativos. Con la ayuda de (2.9), es fácil comprobar que este valor de  $E$  coincide con la energía entregada al sistema en el instante inicial.

### 2.1.6. El ejemplo de $N = 2$ se complica

Al revisar los pasos que permitieron desacoplar las ecuaciones diferenciales, nos damos cuenta que el cambio de coordenadas (2.4) que define las nuevas coordenadas  $\psi_1(t)$  y  $\psi_2(t)$  era evidente, gracias a la simetría del sistema físico elegido. Si las masas de las partes móviles a y b no son iguales y si los tres resortes tienen distintas constantes elásticas, como se indica en el esquema de la figura 2.2, es fácil verificar que el sistema de ecuaciones de Newton se escribe ahora como

$$\begin{aligned} \ddot{\psi}_a &= -\frac{k_1+k_2}{m_a}\psi_a + \frac{k_2}{m_a}\psi_b, \\ \ddot{\psi}_b &= \frac{k_2}{m_b}\psi_a - \frac{k_2+k_3}{m_b}\psi_b, \end{aligned} \quad (2.19)$$

que, usando matrices y vectores columna, también está representado por la ec. (2.3), pero ahora  $M_{11} = \frac{k_1+k_2}{m_a}$ ;  $M_{12} = -\frac{k_2}{m_a}$ ;  $M_{21} = -\frac{k_2}{m_b}$  y  $M_{22} = \frac{k_2+k_3}{m_b}$ . El cambio de coordenadas

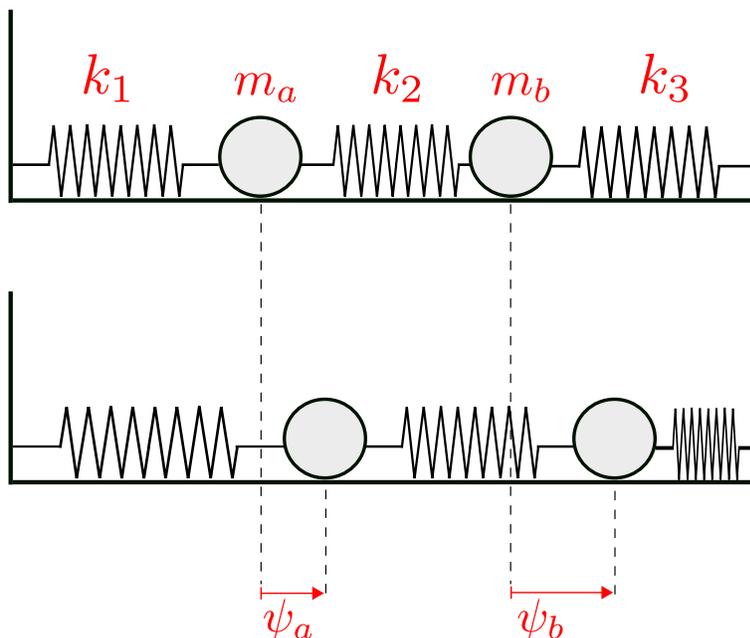


Figura 2.2: Segundo ejemplo de 2GL

que diagonaliza al sistema no es tan evidente. Y si bien es un buen ejercicio usar un poco de intuición y tratar de “adivinar” la transformación análoga a (2.4) para este nuevo problema, queda claro que es necesario un método sistemático que permita encontrar las coordenadas normales y resolver el problema en cualquier situación donde las fuerzas son lineales.

## 2.2. Búsqueda sistemática de modos normales

### 2.2.1. Notación y N grados de libertad

Ya vimos en nuestro ejemplo que las ecuaciones de movimiento (2.2) se pueden representar simbólicamente de la manera mucho más compacta (ver ec. (2.3))

$$\frac{d^2}{dt^2} \begin{pmatrix} \psi_a \\ \psi_b \end{pmatrix} = - \begin{pmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{pmatrix} \begin{pmatrix} \psi_a \\ \psi_b \end{pmatrix}, \quad (2.20)$$

usando matrices y vectores columna. Como siempre vamos a tener una ecuación diferencial para cada parte móvil (la ec. de Newton) y en el caso en que se linealicen las fuerzas que actúan sobre cada parte móvil (todo un tema, para ver en cada problema), nadie podrá dejar de admitir que en un sistema con 3 GL las ecuaciones de movimiento se tienen que

representar simbólicamente como

$$\frac{d^2}{dt^2} \begin{pmatrix} \psi_a \\ \psi_b \\ \psi_c \end{pmatrix} = - \begin{pmatrix} M_{11} & M_{12} & M_{13} \\ M_{21} & M_{22} & M_{23} \\ M_{31} & M_{32} & M_{33} \end{pmatrix} \begin{pmatrix} \psi_a \\ \psi_b \\ \psi_c \end{pmatrix}. \quad (2.21)$$

A esta altura nos tenemos que dar cuenta que etiquetar las coordenadas de cada parte móvil con letras no siempre es una buena idea. Porque si queremos seguir usando esta notación tan linda y tan compacta para un sistema con 40 partes móviles, no van a alcanzar las letras. Así que mejor le decimos chau a la notación con letras y pasamos a etiquetar las coordenadas de cada parte móvil con números naturales. Entonces, en un sistema de cuatro partes móviles (y luego de linealizar las fuerzas), tendremos que las ecuaciones de movimiento son

$$\underbrace{\frac{d^2}{dt^2} \begin{pmatrix} \psi_1 \\ \psi_2 \\ \psi_3 \\ \psi_4 \end{pmatrix}}_{\vec{\psi}} = - \underbrace{\begin{pmatrix} M_{11} & M_{12} & M_{13} & M_{14} \\ M_{21} & M_{22} & M_{23} & M_{24} \\ M_{31} & M_{32} & M_{33} & M_{34} \\ M_{41} & M_{42} & M_{43} & M_{44} \end{pmatrix}}_{\mathbb{M}} \begin{pmatrix} \psi_1 \\ \psi_2 \\ \psi_3 \\ \psi_4 \end{pmatrix}, \quad (2.22)$$

donde definimos el vector columna  $\vec{\psi}$ , la derivada segunda de este vector columna y la matriz  $\mathbb{M}$ . Lo bueno de esta notación es que si en vez de cuatro GL tenemos  $N$ , las ecuaciones de movimiento también se escribirán como

$$\frac{d^2}{dt^2} \vec{\psi} = -\mathbb{M} \vec{\psi}, \quad (2.23)$$

pero ahora el vector columna  $\vec{\psi}$  tendrá  $N$  elementos y la matriz  $\mathbb{M}$  tendrá  $N$  filas y  $N$  columnas.

Por lo general la matriz  $M$  que aparece en estos sistemas diferenciales no es diagonal, es decir las ecuaciones diferenciales están acopladas. Para buscar las coordenadas “buenas” que desacoplan el sistema, tendremos que hacer como hicimos en el ejemplo, es decir buscar combinaciones lineales de las coordenadas “malas”. Notar que no podremos seguir llamando como antes a las coordenadas “buenas”, porque ahora estamos usando las etiquetas de números naturales para identificar a las partes móviles.

### 2.2.2. Busquemos un modo físicamente

El vector columna  $\vec{\psi}(t)$  que aparece en (2.23) representa en cada instante de tiempo el estado del sistema apartado del equilibrio para el caso más general posible (es decir, condiciones iniciales arbitrarias). Supongamos que existen condiciones iniciales particulares para las cuales todas las partes móviles del sistema oscilan con la misma frecuencia y pasan simultáneamente por la posición de equilibrio. Es decir, vamos a proceder como en las demostraciones por el absurdo, suponemos que existe un modo y queremos ver si

esta suposición nos lleva a conclusiones compatibles con el resto de suposiciones. En este caso, suponer que existe un modo nos lleva a admitir que el vector columna  $\vec{\psi}(t)$  tiene que tener todos los elementos de su columna con la misma dependencia temporal, de la forma  $\cos(\omega t + \varphi)$ , todos con el mismo  $\omega$  y el mismo  $\varphi$ , aunque podrían tener distinta amplitud, por ejemplo si una parte móvil tiene mucha más masa que el resto, esperamos que se mueva menos. Esto quiere decir que si existe este modo, el vector columna  $\vec{\psi}(t)$  tiene que tener la forma

$$\vec{\psi}(t) = \begin{pmatrix} \psi_1(t) \\ \psi_2(t) \\ \vdots \\ \psi_N(t) \end{pmatrix} = \begin{pmatrix} A_1 \\ A_2 \\ \vdots \\ A_N \end{pmatrix} \cos(\omega t + \varphi) = \vec{A} \cos(\omega t + \varphi), \quad (2.24)$$

donde definimos el vector de amplitudes  $\vec{A}$ . El vector columna  $\vec{\psi}(t)$  de (2.24) describe el movimiento en un caso especial de condiciones de contorno, pero no está exceptuado de satisfacer las ecuaciones de Newton (2.23). En este caso especial es muy fácil calcular la derivada segunda del vector  $\vec{\psi}(t)$ , como se ve en (2.24) debe ser  $\frac{d^2}{dt^2} \vec{\psi} = -\omega^2 \vec{A} \cos(\omega t + \varphi)$ . Y por otro lado,  $\frac{d^2}{dt^2} \vec{\psi} = -\mathbb{M} \vec{\psi} = -\mathbb{M} \vec{A} \cos(\omega t + \varphi)$ . Con esto concluimos que

$$\mathbb{M} \vec{A} = \omega^2 \vec{A}, \quad (2.25)$$

o, dicho de otra manera: si existe un modo, entonces el vector columna de las amplitudes del modo tiene que ser **un autovector** de la matriz  $\mathbb{M}$  y la frecuencia del modo es la raíz cuadrada del **autovalor** correspondiente  $\omega^2$ .

### 2.2.3. Buscamos un modo y encontramos $N$ modos

La ec. (2.25) se puede reescribir como

$$(\mathbb{M} - \omega^2 \mathbb{I}) \vec{A} = 0, \quad (2.26)$$

donde  $\mathbb{I}$  es la matriz identidad. Si la matriz  $\mathbb{M} - \omega^2 \mathbb{I}$  de este sistema homogéneo tuviera determinante distinto de cero, las amplitudes incógnita  $\vec{A}$  serían cero. Y esto querría decir que la única situación donde todas las partes móviles del sistema oscilan con la misma frecuencia y pasan simultáneamente por la posición de equilibrio es cuando oscilan con amplitud cero . . . ¡es decir cuando permanecen siempre en la posición de equilibrio ( $\vec{\psi}(t) = 0$ )! No es por criticar ni porque el resultado esté mal, pero claramente este resultado no es precisamente lo que estábamos buscando, por no decir que es una trivialidad. Si queda alguna posibilidad para que exista un modo y se cumpla la ec. (2.27), entonces necesariamente debería cumplirse que

$$\det(\mathbb{M} - \omega^2 \mathbb{I}) = 0. \quad (2.27)$$

El determinante de la matriz  $\mathbb{M} - \omega^2 \mathbb{I}$  es un polinomio de grado  $N$  en la variable  $\omega^2$  y se denomina polinomio característico de  $\mathbb{M}$ . Veremos en los ejemplos que en los casos sin

disipación  $\mathbb{M}$  tiene  $N$  autovalores reales y la multiplicidad de cada uno de ellos (como raíz del polinomio característico) coincide con la dimensión del subespacio propio generado por los correspondientes autovectores. Y así llegaremos a la conclusión de que no sólo existe un modo, sino que existen  $N$  modos que representan  $N$  maneras distintas en las que todas las partes móviles del sistema oscilan con la misma frecuencia y pasan simultáneamente por su posición de equilibrio. Cada manera está caracterizada por  $\omega_i$  (la raíz cuadrada del autovalor  $\omega_i^2$ ) y por  $\vec{A}^{(i)}$  (el autovector correspondiente), que satisfacen

$$\mathbb{M} \vec{A}^{(i)} = \omega_i^2 \vec{A}^{(i)} , \quad (2.28)$$

con  $i = 1, 2, \dots, N$  y que sirven para armar el vector columna  $\vec{\psi}^{(i)}(t) = \vec{A}^{(i)} \cos(\omega_i t + \varphi_i)$  que describe el movimiento particular de cada modo.

### 2.2.4. El movimiento más general

Si cada  $\vec{\psi}^{(i)}(t)$ ,  $i = 1, 2, \dots, N$  es una solución de las ecuaciones de movimiento, toda combinación lineal de los  $\vec{\psi}^{(i)}(t)$  también será solución, es decir que a partir de las soluciones modales podemos construir nuevas soluciones donde las partes móviles del sistema ya no necesariamente oscilan con la misma frecuencia ni pasan simultáneamente por la posición de equilibrio. Escribimos esta combinación lineal como

$$\vec{\psi}(t) = \sum_{i=1}^N C_i \vec{A}^{(i)} \cos(\omega_i t + \varphi_i) , \quad (2.29)$$

donde hemos adoptado la convención usual de trabajar con autovectores de norma uno (recordar que los autovectores están definidos a menos de una constante multiplicativa). En la notación original de las coordenadas “malas”

$$\psi_j(t) = \sum_{i=1}^N C_i A_j^{(i)} \cos(\omega_i t + \varphi_i) , \quad j = 1, 2, \dots, N. \quad (2.30)$$

En las ecs. (2.29) (o (2.30)) hay  $2N$  constantes a determinar, que se obtienen a partir de las  $2N$  condiciones iniciales,  $N$  posiciones iniciales dadas por  $\vec{\psi}(0)$  y  $N$  velocidades iniciales dadas por  $\dot{\vec{\psi}}(0)$ . Si aceptamos que el conjunto de todas las soluciones modales es equivalente al conjunto de todas las soluciones del sistema diferencial desacoplado, entonces tenemos que admitir que la expresión de la ec. (2.29) representa el movimiento más general del sistema mecánico como superposición de modos normales. Pero si no se acepta que el conjunto de todas las soluciones modales es equivalente al conjunto de todas las soluciones del sistema diferencial desacoplado, entonces hay que perfeccionar los argumentos “físicos” sobre la plausibilidad de la afirmación, o pasar a hacer el tratamiento matemático riguroso que estuvimos evitando (pero que sí hicimos en el ejemplo), que es ver el tema del desacople de las ecuaciones diferenciales.

### 2.2.5. El método aplicado al ejemplo

En el ejemplo de la figura 2.1 descrito por las ecs. (2.1), la matriz  $\mathbb{M}$  es (ver ec. (2.3))

$$\mathbb{M} = \begin{pmatrix} 2\frac{k}{m} & -\frac{k}{m} \\ -\frac{k}{m} & 2\frac{k}{m} \end{pmatrix}. \quad (2.31)$$

Es fácil verificar que los autovalores de esta matriz son  $\omega_1^2 = \frac{k}{m}$  y  $\omega_2^2 = 3\frac{k}{m}$  y los autovectores correspondientes son

$$\vec{A}^{(1)} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad \vec{A}^{(2)} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \quad (2.32)$$

¿Cómo sería la matriz  $\mathbb{M}$  si en el sistema de la figura 2.1 se cambia solamente el resorte del medio, con una constante elástica  $K \neq k$ ? ¿Y los autovalores y los autovectores? Repetir para los tres resortes con constantes elásticas distintas  $K_1$ ,  $K_2$  y  $K_3$ . Y por último, repetir para constantes elásticas distintas y las dos masas  $m_1$  y  $m_2$  también distintas.

Siguiendo con el ejemplo, formemos la matriz  $\mathbb{D}$  cuyas columnas son los autovectores  $\vec{A}^{(1)}$  y  $\vec{A}^{(2)}$

$$\mathbb{D} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}. \quad (2.33)$$

Seguramente a esta altura muchos ya deben haber recordado que  $\mathbb{D}$  es una matriz de cambio de base, en este caso entre la base de coordenadas “malas” y la base de modos normales o coordenadas “buenas”, como se puede comprobar observando que el vector de desplazamientos  $\vec{\psi}(t)$  (en la base “mala”) y el vector de desplazamientos  $\vec{\Phi}(t)$  (en la base “buena”), están relacionados por

$$\vec{\psi}(t) = \mathbb{D} \vec{\Phi}(t), \quad (2.34)$$

que a menos de un factor constante, es completamente equivalente a la inversa de la transformación definida en la ecs. (2.4). Notar que  $\mathbb{D}^t \mathbb{M} \mathbb{D}$  (donde  $\mathbb{D}^t$  representa la matriz transpuesta) es una matriz diagonal. Claro, porque

$$\mathbb{D}^t \mathbb{M} \mathbb{D} = \frac{1}{2} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} 2\frac{k}{m} & -\frac{k}{m} \\ -\frac{k}{m} & 2\frac{k}{m} \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} = \begin{pmatrix} \frac{k}{m} & 0 \\ 0 & 3\frac{k}{m} \end{pmatrix}. \quad (2.35)$$

### 2.2.6. Todo lo anterior en el lenguaje del álgebra lineal

#### Sistema desacoplado y coordenadas normales

En 2.2.2 buscamos un modo imponiendo que todas las partes móviles del sistema físico realicen el mismo movimiento oscilatorio armónico (la misma frecuencia para todas las partes móviles) y pasen por su posición de equilibrio simultáneamente. La manera formal y matemáticamente correcta de ver que con esta suposición no hemos perdido soluciones y que la solución general para condiciones iniciales arbitrarias es la superposición de  $N$  movimientos de este tipo, es justamente a través de la **transformación de semejanza**  $\mathbb{D}^t \mathbb{M} \mathbb{D}$  que transforma la matriz  $\mathbb{M}$  en una matriz diagonal, siempre que las columnas de  $\mathbb{D}$  sean los autovectores de  $\mathbb{M}$  y que además la matriz  $\mathbb{M}$  satisfaga ciertas condiciones que tendremos la precaución de verificar si este tema nos pone nerviosos, pero que ahora supondremos que se cumplen. En este caso, multiplicamos miembro a miembro la ec. (2.23) por  $\mathbb{D}^t$  y observando que  $\mathbb{D}$  no depende del tiempo y que es una matriz ortogonal ( $\mathbb{D}^{-1} = \mathbb{D}^t$ ), obtenemos

$$\frac{d^2}{dt^2} \mathbb{D}^t \vec{\psi} = - [\mathbb{D}^t \mathbb{M} \mathbb{D}] \mathbb{D}^t \vec{\psi}. \quad (2.36)$$

Teniendo en cuenta que  $\mathbb{D}^t \mathbb{M} \mathbb{D}$  es diagonal, resulta entonces que el vector

$$\vec{\Phi} = \mathbb{D}^t \vec{\psi}, \quad (2.37)$$

satisface un sistema de ecuaciones diferenciales desacopladas, es decir cada componente del vector de desplazamientos  $\vec{\Phi}$  satisface una ecuación como la del oscilador armónico donde cada autovalor juega el papel de una frecuencia al cuadrado. Usando el resultado ya familiar de que la solución más general de cada una de estas  $N$  ecuaciones diferenciales desacopladas es de la forma  $C_i \cos(\omega_i t + \varphi_i)$  y volviendo a las viejas coordenadas

$$\vec{\psi} = \mathbb{D} \vec{\Phi}, \quad (2.38)$$

queda justificado el resultado de la ec. (2.30) que obtuvimos con argumentos físicos, aunque menos rigurosos desde el punto de vista matemático.

#### Condiciones iniciales

Las  $2N$  constantes  $C_i, \varphi_i, i = 1, 2, \dots, N$  en (2.29) (y en (2.30)) se obtienen a partir de las  $2N$  condiciones iniciales representadas por  $\vec{\psi}(0) = \mathbb{D} \vec{\Phi}(0)$  y  $\dot{\vec{\psi}}(0) = \mathbb{D} \dot{\vec{\Phi}}(0)$ . Pero tal como hicimos en (2.8), es más conveniente imponer las condiciones iniciales en las coordenadas normales. Esto es así porque, tal como se ve en (2.8), las dos ecuaciones que representan la condición inicial sobre la coordenada normal 1 y sobre su derivada, solamente involucran a las dos incógnitas  $C_1, \varphi_1$ , las dos ecuaciones que representen la condición inicial sobre la coordenada normal 2 y sobre su derivada, solamente involucran a las dos incógnitas  $C_2, \varphi_2$ , etc. Y en general, las dos ecuaciones que representen la

condición inicial sobre la coordenada normal  $i$ -ésima y sobre su derivada, solamente van a involucrar a las dos incógnitas  $C_i, \varphi_i$ , con el mismo valor de  $i$ , sin mezclar con otras incógnitas asociadas a otras coordenadas normales con distinto valor del índice  $i$ . De esta manera, se va a poder repetir el esquema usado para llegar a (2.8), esencialmente elevar al cuadrado y sumar miembro a miembro para obtener  $C_i$ , y el esquema usado para llegar a (2.10), dividir miembro a miembro para obtener  $\tan \varphi_i$ .

### Energía cinética y potencial

En cursos más avanzados se verá la generalización a  $N$  grados de libertad de los resultados obtenidos en la sección 2.1.5 para nuestro ejemplo particular de dos grados de libertad. El resumen es el siguiente. En las coordenadas ordinarias, la energía cinética total del sistema,  $T(t)$  es una forma cuadrática en las velocidades de las partes móviles, sin términos cruzados (el resultado que obtuvimos en (2.12)), mientras que la energía potencial total del sistema,  $V(t)$ , es una forma cuadrática en los desplazamientos de las partes móviles, que incluye términos cruzados (el resultado que obtuvimos en (2.13)). En cambio, al hacer la transformación (2.37),  $T(t)$  y  $V(t)$  resultan formas cuadráticas de las coordenadas normales  $\vec{\Phi}(t)$ , sin términos cruzados, obteniéndose así un resultado similar al que obtuvimos en (2.18) y, de esta manera, *la energía mecánica total del sistema se puede calcular sumando aportes de los distintos modos, como si hubieran sido excitados separadamente*.

## 2.3. Matrices, autovalores y autovectores en Python

Vamos ahora a cargar una matriz sencilla, como la matriz asociada al sistema de la figura 1 y dada por la ec. (2.31). Para simplificar el tema de las unidades, decidimos medir los elementos de esta matriz en unidades de  $k/m$  (una magnitud con unidades de frecuencia al cuadrado, es decir,  $s^{-2}$ ), lo que equivale a sacar factor común  $k/m$  en la ec. (2.31). La instrucción para cargar esta matriz (por ahora llamémosla matriz, aunque para Python sea solamente un "array", no vale la pena ahora detenernos en este detalle) y darle el nombre  $M$  es

```
>>> M=[[2, -1], [-1, 2]]
```

Cuando tecleamos esta instrucción y apretamos la tecla "enter", Python da señales de vida y escribe este mensaje

```
M=[[2, -1], [-1, 2]]
```

que quiere decir que entendió bien y que a partir de ahora el símbolo  $M$  representa a la matriz  $\begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}$ , que es la misma de la ec. (2.31) pero medida en unidades de  $k/m$ . Ahora

le vamos a decir a Python que calcule los autovalores y autovectores de esta matriz. Para hacerlo, vamos a cargar una extensión de Python que agrega mayor soporte para vectores y matrices y que constituye una biblioteca de funciones matemáticas de alto nivel para operar con esos vectores o matrices. La extensión se llama NumPy (python numérico para los amigos) y se carga tecleando

```
>>> import numpy
```

Aunque ahora no viene al caso, podríamos ver otras habilidades del paquete NumPy. Por ejemplo, para invocar el número  $\pi$  hay que escribir

```
>>> numpy.pi
```

y al apretar la tecla “enter” devuelve el valor 3.141592653589793. En cambio, si escribimos

```
>>> numpy.e
```

al apretar la tecla “enter” devuelve el valor 2.718281828459045.

Para comprobar con Python los resultados de las ecs. (2.32), (2.33) y (2.35), tenemos que saber que `numpy` tiene un paquete de rutinas llamado `linalg` (que como su nombre lo indica es para álgebra lineal) y este paquete tiene una función `eig` que justamente calcula autovalores y autovectores de una matriz. La función `eig` aplicada a una matriz devuelve dos objetos: una tira con los autovalores y la matriz de autovectores. A estos objetos les damos el nombre que más nos guste (para no ser rebuscados los llamamos `omega2` y `D` respectivamente) y los ponemos a la izquierda del signo igual. Del otro lado del igual invocamos a la función `numpy.linalg.eig` aplicada a la matriz que queremos diagonalizar, de esta manera

```
>>> omega2, D = numpy.linalg.eig(M)
```

y apretamos la tecla “enter”. Listo, si todo anduvo bien tenemos lo que estábamos buscando. Para comprobarlo, le decimos a Python que nos muestre `omega2`, y luego del consabido “enter” nos devuelve

```
>>> omega2
array([ 3.,  1.] )
```

que quiere decir que un autovalor es igual a 3 y el otro es igual a 1. Lo mismo que obtuvimos a mano, pero como Python es medio tonto no se da cuenta que nos hubiera gustado mucho más que el primer autovalor fuera el que vale 1 y el segundo el que vale

3. Se lo dejamos pasar, no hay que pedirle peras al Python. Análogamente, le decimos a Python que nos muestre  $D$ , y luego del “enter” nos devuelve

```
>>> D
array([[ 0.70710678,  0.70710678],
       [-0.70710678,  0.70710678]])
```

También coincide con lo que obtuvimos a mano, la primera columna corresponde al modo 2 y la segunda al modo 1.

Y ahora veamos una de las ventajas de aprender el jueguito de Python: lo que hicimos recién también sirve para cosas más complicadas. Por ejemplo, para un sistema similar al de la Fig. 1, pero con tres masas idénticas ( $m$ ) unidas por cuatro resortes idénticos ( $k$ ), la matriz  $M$  es

$$M = \begin{pmatrix} 2\frac{k}{m} & -\frac{k}{m} & 0 \\ -\frac{k}{m} & 2\frac{k}{m} & -\frac{k}{m} \\ 0 & -\frac{k}{m} & 2\frac{k}{m} \end{pmatrix}. \quad (2.39)$$

Para cargar esta matriz en Python (en unidades de  $k/m$ ) y con el mismo nombre  $M$ , se puede hacer “cut&paste” en lo que escribimos antes y con muy pocas modificaciones escribir

```
>>> M=[[2, -1, 0], [-1, 2, -1], [0, -1, 2]]
```

Lo mismo para obtener los autovalores y  $\mathbb{D}$  (la matriz de autovectores)

```
>>> omega2, D = numpy.linalg.eig(M)
>>> omega2
array([ 3.41421356,  2.          ,  0.58578644])
>>> D
array([[ -5.00000000e-01,  -7.07106781e-01,  5.00000000e
-01],
       [ 7.07106781e-01,  4.06080300e-16,  7.07106781e
-01],
       [ -5.00000000e-01,  7.07106781e-01,  5.00000000e
-01]])
```

Notar que el elemento del resultado anterior que físicamente representa el nodo del modo 2 no da cero como debería dar, si no que tiene un valor muy pequeño. Esto es así porque Python está usando precisión finita (es decir, un número finito de bits para hacer aritmética).

Consideremos otro sistema similar al de la Fig. 1 con dos masas idénticas ( $m$ ), pero ahora los resortes de los extremos tienen ambos constante elástica  $K$  distinta a la constante del resorte de acople ( $k$ ). Como cada término del miembro derecho del sistema de ecs. (2.2) está asociado a una fuerza, es fácil ver que la matriz  $\mathbb{M}$  ahora es

$$\mathbb{M} = \begin{pmatrix} \frac{k+K}{m} & -\frac{k}{m} \\ -\frac{k}{m} & \frac{k+K}{m} \end{pmatrix} = \frac{K}{m} \begin{pmatrix} 1 + \frac{k}{K} & -\frac{k}{K} \\ -\frac{k}{K} & 1 + \frac{k}{K} \end{pmatrix}, \quad (2.40)$$

donde saqué factor común  $K/m$  para facilitar el ingreso de datos en Python (en unidades de  $K/m$ ) y la interpretación de resultados numéricos. Como seguramente ya todos saben, las frecuencias de los modos normales en este caso son

$$\omega_1^2 = \frac{K}{m}, \quad \omega_2^2 = \frac{K}{m} + 2\frac{k}{m}, \quad (2.41)$$

y por la simetría del problema los modos siguen siendo iguales que en nuestro primer ejemplo. En particular, me interesa explorar con Python el caso en que las dos masas están débilmente acopladas, es decir cuando  $k \ll K$ . Probamos con el valor  $k/K = 0.1$ , cargamos la matriz (2.40) en Python (en unidades de  $K/m$ ), le ordenamos a Python que calcule la matriz diagonal de autovalores y la matriz de autovectores y obtenemos

```
>>> M=[[1.1, -0.1], [-0.1, 1.1]]      (Enter)
>>> omega2, A = numpy.linalg.eig(M)    (Enter)

>>> [A, omega2]=eig(M)                (Enter)
>>> A                                  (Enter)
array([[ 0.70710678,  0.70710678],
       [-0.70710678,  0.70710678]])
>>> omega2                             (Enter)
array([ 1.0000,  1.2000])
```

resultados que coinciden con lo que esperábamos.

En los ejemplos que hicimos hasta acá el cálculo numérico no es imprescindible, porque los autovectores y autovalores se pueden calcular analíticamente sin ninguna dificultad. Una ventaja de usar lenguajes de programación como Python es evidente cuando hay que resolver problemas con muchos grados de libertad, en los que es muy pesado, o a veces directamente imposible, calcular autovectores y autovalores de manera analítica. Por ejemplo, prueben cómo harían con un sistema similar al de la Fig. 1, pero con diez masas idénticas ( $m$ ) unidas por once resortes idénticos ( $k$ ).

### 2.3.1. Visualización de resultados

Antes de seguir, digamos que hay veces en que puede ser conveniente que Python se olvide de todas las definiciones previas. Una manera a lo bestia de hacer esto es salir del

programa, que se logra con la instrucción `Ctrl-D`, y luego volver a entrar como se explicó en la Sección 3.1. También observemos que Python/Matlab/Octave ... por lo general aceptan como variable de entrada argumentos escalares, o argumentos que son vectores filas, o vectores columnas, o matrices (el misterio por ahora de los arrays). Y que en estos casos, el resultado o salida de estas funciones es también un escalar, vector (fila o columna) o matriz, respectivamente, dependiendo de la entrada. Por ejemplo, a continuación creo una lista (como un vector) y tomo el logaritmo de la lista, que resulta ser otra lista donde cada elemento es el logaritmo (natural o decimal) del elemento correspondiente en la lista original

```
>>> import numpy
>>> list=[10,100,1000,10000]
>>> numpy.log(list)
array([ 2.30258509,  4.60517019,  6.90775528,
        9.21034037])
>>> numpy.log10(list)
array([ 1.,  2.,  3.,  4.] )
```

(se supone que saben que hubo que usar varios “enter”).

Otra situación donde también resulta extremadamente útil saber usar programas como Python es cuando se desea visualizar o ilustrar de manera gráfica el comportamiento de los fenómenos físicos estudiados. Por simplicidad seguimos con un ejemplo que ya hemos resuelto analíticamente, el de las dos masas idénticas, con los resortes de los extremos de constante  $K$  y el resorte de acople de constante  $k$ . En este caso la matriz  $M$  viene dada por la ec. (2.40). Para centrarnos en la visualización, dejamos de lado por el momento el tema de la diagonalización y usamos resultados analíticos. A pesar de que no es el mismo problema, en este caso el desplazamiento de las partes móviles viene dado por las funciones de las ecs. (2.11), con amplitudes y fases dadas por las ecs. (2.9) y (2.10) y frecuencias propias dadas por la ec. (2.41). Si además nos conformamos con limitarnos al caso en que la perturbación del equilibrio de ambas masas se hace sin velocidad inicial, las fases de las ecs. (2.9) resultan nulas y todo lo que necesitamos para visualizar el movimiento de las partes móviles son las dos posiciones iniciales, las constantes elásticas  $K$  y  $k$  y la masa  $m$ . Nos proponemos usar Python para visualizar este movimiento cuando se varían estos cinco parámetros.

Comenzamos por introducir los datos con nombres que tengan que ver con lo que estamos haciendo, en nuestro caso con las ecs. (2.11). Por ejemplo, aquí va una secuencia donde se omite que después de cada comando hay que apretar la tecla “Enter”

```
>>> psia0=1.
>>> psib0=0.
>>> A1=((psia0+psib0)**2)**(0.5)
>>> A2=((psia0-psib0)**2)**(0.5)
```

Si los desplazamientos iniciales están medidos en centímetros, los desplazamientos de las partes móviles también lo estarán. Puse raíces cuadradas ( $** (0.5)$ ) de algo al cuadrado ( $** 2$ ) como una manera de tomar el valor absoluto, en este ejemplo parece un resultado irrelevante pero no lo habría sido si hubiera dejado quieta la parte móvil *a* y movido la *b*. Para estar seguros de que Python calculó bien las amplitudes de las ecs. (2.9), podemos pedir que muestre en pantalla sus valores, de la siguiente manera

```
>>> A1
1.
>>> A2
1.
```

Ahora ingresaremos los valores de las constantes elásticas  $K$  y  $k$  y de la masa  $m$ . Elijamos los valores  $K = 1$  N/m,  $k = 0.05$  N/m y  $m = 1$  kg, las frecuencias propias estarán medidas en segundos.

```
>>> K=1.; k=0.05; m=1.
>>> w1=(K/m)**0.5
>>> w2=(K/m+2*k/m)**0.5
```

Si somos muy controladores (para programar no es un defecto, aunque en la vida real puede serlo), verificamos

```
>>> w1
1.0
>>> w2
1.0488088481701516
```

Para que Python evalúe el desplazamiento de las partes móviles, dado por las funciones de las ecs. (2.11), creamos un vector fila con los valores discretizados del tiempo a través de la función `linspace` del paquete `numpy`.

```
>>> t=numpy.linspace(0,300,1500)
```

Esta instrucción genera una secuencia equiespaciada entre un valor inicial, un valor final y con tantos elementos como indique su tercer argumento. En el ejemplo se generan 1500 valores entre 0 y 300 segundos, esos 1500 valores representan instantes de tiempo, notar que los 1500 valores están representados por el nombre de variable  $t$ , por eso es como si fuera un vector fila. Y cada elemento de este vector se podría invocar por algún índice. Python tiene la característica de que no le gusta manejar índices empezando desde 1 si no empezando desde cero, como se puede ver de la siguiente manera

```
>>> t[0]
0.0
>>> t[1499]
300.0
>>> t[30]
6.0040026684456302
>>> t[1000]
200.13342228152101
```

Y luego se generan vectores fila de 1500 elementos con los valores de los desplazamientos de cada masa de la siguiente manera

```
>>> psiat=0.5*(A1*numpy.cos(w1*t)+A2*numpy.cos(w2*t))
>>> psibt=0.5*(A1*numpy.cos(w1*t)-A2*numpy.cos(w2*t))
```

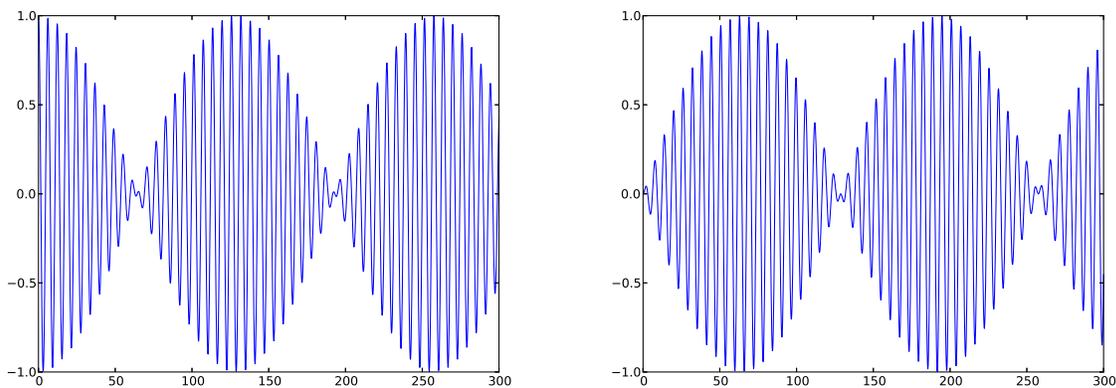


Figura 2.3: Movimiento de las partes a y b para condiciones iniciales sin velocidad inicial,  $\psi_a(0) = 1m$ ,  $\psi_b(0) = 0$ ,  $K = 1$  N/m,  $k = 0.05$  N/m y  $m = 1$  kg.

Acá `numpy.cos` recibe como argumento una lista y devuelve otra lista, funciona análogamente a lo que vimos en el ejemplo del logaritmo anteriormente. Notar que Python recuerda las definiciones anteriores sobre variables como `w1`, `w2`, `A1`, `A2`, `psia0`, `psib0`, etc. Esto siempre y cuando el gato no nos haya desconectado la computadora, o hayamos cerrado inadvertidamente la consola de trabajo u otros incidentes que suelen suceder y que se pueden evitar guardando todos los comandos en un archivo con extensión `.py`, que esencialmente sería lo que se llama un *script* o *archivo de comandos* o sea, un programa. Por ahora seguimos trabajando en la modalidad interactiva o de consola. Lo único que falta es que ahora Python muestre los gráficos del desplazamiento de las partes móviles para estas condiciones iniciales y para los resortes y masas elegidos en este cálculo. Hay muchas maneras de hacerlo, una es importar `matplotlib.pyplot` (se refiere al subconjunto `pyplot` de `matplotlib`). A esta altura las cosas empiezan a ser autoevidentes y como

no es un curso de programación, no nos ponemos obsesivos, no entramos en demasiados detalles y ejecutamos obedientemente las siguientes instrucciones

```
>>> import matplotlib.pyplot as plt
>>> fig = plt.figure()
>>> plt.plot(t, psiat)
>>> plt.draw()
>>> plt.savefig('ej1parte.a.eps')
>>> plt.savefig('ej1parte.a.jpg')
```

El comando `import` ya lo usamos, ahora tiene la novedad del `as`, que sirve para usar el nombre mas corto `plt` en vez de `matplotlib.pyplot`. El comando `plt.figure()` inicia una nueva figura, anulando toda figura anterior. Realmente no hace falta en este ejemplo particular, pero es una buena costumbre ponerlo en programas muy largos. El comando `plt.plot` grafica el segundo argumento en función del primero, tiene otras opciones (que no uso ahora) para controlar tamaño de trazo, tipo y color de línea, etc. El comando `plt.draw` vuelve a dibujar la figura en la que estamos trabajando y viene bien en modo interactivo porque permite modificar sobre la marcha. Por último, el comando `plt.savefig('ej1parte.a.eps')` lo pongo para indicar que quiero que transforme cada gráfico en una imagen en formato eps, que es las imagen que se ve en la figura 2.3a. Si preferimos el formato jpg, usamos el último comando del listado. Dejo que exploren cómo mejorar estos gráficos, poner nombre a los ejes, controlar las escalas, cambiar el color de las curvas, etc. Pregúntenle al buscador, por ejemplo yo recién puse *como graficar con python* y el tercer resultado fue este enlace <http://bit.ly/f2c3pyt>, que corresponde a un blog dedicado a enseñar y difundir programación Python.

### 2.3.2. Archivo de comandos o *script*

Si ahora quisiéramos explorar otras situaciones, deberíamos repetir el proceso explicado en 3.3.2. Los archivos de comandos o *scripts* facilitan notablemente esta tarea, la idea es que se pueden poner todos los comandos en un archivo plano de texto, con extensión `py` que se guarda en el directorio de trabajo. Por ejemplo, a la secuencia de instrucciones mostrada en la sección 3.3.2, la guardo en el archivo `Ejemplo1.py` y luego ejecuto este archivo de comandos simplemente escribiendo `python Ejemplo1.py` desde la terminal (y Enter). Como uno de los mejores métodos para aprender es haciendo, dejo el script usado para obtener la figura 2.3: i) prueben con otras condiciones de contorno, ii) generalicen para incluir el caso de velocidades iniciales no nulas, iii) grafiquen las energías cinéticas y potenciales del sistema y de cada parte móvil, y iv) generalicen al caso de masas y resortes distintos. El signo `#` se puede usar en un script para incluir comentarios: todo lo que está después del `#` es un comentario.

```
# Ejemplo de archivo de comandos en Python
%# Curso de F2, 2do cuatrimestre de 2016
```

```

# Dos masas iguales, resorte del medio distinto
# a los dos resortes de las paredes
import numpy
import matplotlib.pyplot as plt
# datos de entrada
psia0=1; psib0=0 # en cm
K=1.; k=0.05; m=1. # k en N/m, m en kg
# calculos
A1=((psia0+psib0)**2)**0.5
A2=((psia0-psib0)**2)**0.5
w1=(K/m)**(0.5)
w2=(K/m+2*k/m)**(0.5)
t=numpy.linspace(0,300,1500)
psiat=0.5*(A1*numpy.cos(w1*t)+A2*numpy.cos(w2*t))
psibt=0.5*(A1*numpy.cos(w1*t)-A2*numpy.cos(w2*t))
# graficos
fig = plt.figure()
plt.plot(t, psiat)
plt.draw()
plt.savefig('ej1partea.eps')
plt.savefig('ej1partea.jpg')
fig = plt.figure()
plt.plot(t, psibt)
plt.draw()
plt.savefig('ej1parteb.jpg')
plt.savefig('ej1parteb.eps')

```

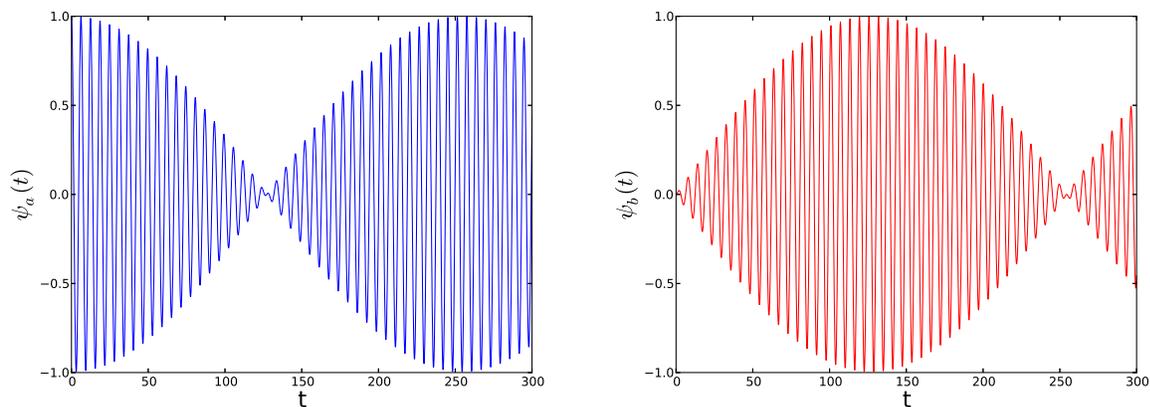


Figura 2.4: Como en la figura 2.3, excepto que el resorte que acopla tiene constante elástica  $k = 0.025$  N/m (acoplamiento más débil).

Espero que todos se den cuenta de la importancia de disponer de un script. Por ejemplo,

si quiséramos explorar el movimiento en un caso idéntico, excepto que ahora el acople es la mitad de débil que antes ( $k = 0.025$  N/m en vez de  $k = 0.05$  N/m), lo único que tendríamos que modificar es una línea y volver a ejecutar el archivo completo de instrucciones. Hagan la prueba, en este caso se obtienen los desplazamientos que se muestran en la figura 2.4. Observar que al disminuir el acomplamiento cambia el número de máximos absolutos en el intervalo de tiempo estudiado, el estudio numérico muestra que hay cosas que nos estamos perdiendo, lo que corresponde es ver si podemos profundizar el análisis, cosa que haremos a continuación.

## 2.4. Batidos y pulsaciones

Un resultado físico muy interesante obtenido en el estudio numérico ilustrado en las figuras 2.3 y 2.4 es el siguiente. En ambos casos sólo perturbamos inicialmente la parte móvil de la izquierda, pero la de la derecha no. Sin embargo, a medida que pasa el tiempo la perturbación se va transmitiendo a la parte móvil de la derecha, que se empieza a mover a partir del equilibrio mientras que la parte móvil de la izquierda empieza a disminuir la amplitud de su movimiento. Y esto de alguna manera es una perturbación que se propaga, no van a decir que no. Vemos que la energía potencial que se entregó inicialmente a la parte móvil de la izquierda, se va transfiriendo a la parte móvil de la derecha. Podríamos especular que si el último resorte estuviera atado a una tercera parte móvil (en vez de estar atado a una pared rígida) esta transferencia de energía continuaría alejándose de la parte móvil que le dió origen. Pero como está la pared, es como si rebotara y volviera de nuevo a la parte móvil de la izquierda. Lo interesante de esta manera de explicar lo que pasa en estos ejemplos es que está surgiendo el concepto de ondas que mencionamos en la primera clase, como **una perturbación que se propaga**. Es cierto que mucha propagación no puede haber en un sistema tan chico y limitado como el que estuvimos viendo, pero pronto nos convenceremos de que en sistemas menos limitados el movimiento producido por la perturbación del equilibrio de la parte móvil de un extremo se puede seguir propagando hacia otras partes móviles más lejanas, hasta el momento en que el sistema termina, en cuyo caso no tiene más remedio que rebotar y propagarse de nuevo pero en dirección contraria.

Los comportamientos graficados en las figuras 2.3 y 2.4 se puede explicar analíticamente empleando una identidad que transforma la suma de dos cosenos en un producto de otros dos cosenos. El resultado es importante en todo sistema físico donde la respuesta  $\psi(t) = \psi_1(t) + \psi_2(t)$  es la superposición de dos oscilaciones armónicas de frecuencias angulares diferentes,  $\psi_1(t) = A_1 \cos(\omega_1 t + \varphi_1)$  y  $\psi_2(t) = A_2 \cos(\omega_2 t + \varphi_2)$ . Los ejemplos anteriores de dos modos normales son un caso especial, pero algo similar ocurre en el caso de dos diapasones que suenan simultáneamente, cada uno con su propia nota, que producen en el aire que llega al tímpano diferencias de presión que son la superposición de dos movimientos oscilatorios armónicos.

Si  $A_1 = A_2$  y  $\varphi_1 = \varphi_2 = 0$ , entonces  $\psi(t) = A(\cos \omega_1 t + \cos \omega_2 t)$ . Y la suma de cosenos

entre paréntesis se puede transformar si se tiene en cuenta que

$$\cos(x + y) + \cos(x - y) = 2 \cos x \cos y. \quad (2.42)$$

Para usar esta identidad tomamos  $\omega_1 t = x + y$ ,  $\omega_2 t = x - y$ , con lo que resulta

$$x = \underbrace{\frac{\omega_1 + \omega_2}{2}}_{\omega_p} t, \quad y = \underbrace{\frac{\omega_1 - \omega_2}{2}}_{\omega_m} t. \quad (2.43)$$

De esta manera, podemos reescribir  $\psi(t)$  como

$$\psi(t) = A(\cos \omega_1 t + \cos \omega_2 t) = \underbrace{2A \cos \omega_m t}_{A_m(t)} \cos \omega_p t = A_m(t) \cos \omega_p t. \quad (2.44)$$

Debido a que el coseno que oscila a la frecuencia promedio varía en el tiempo más rápi-

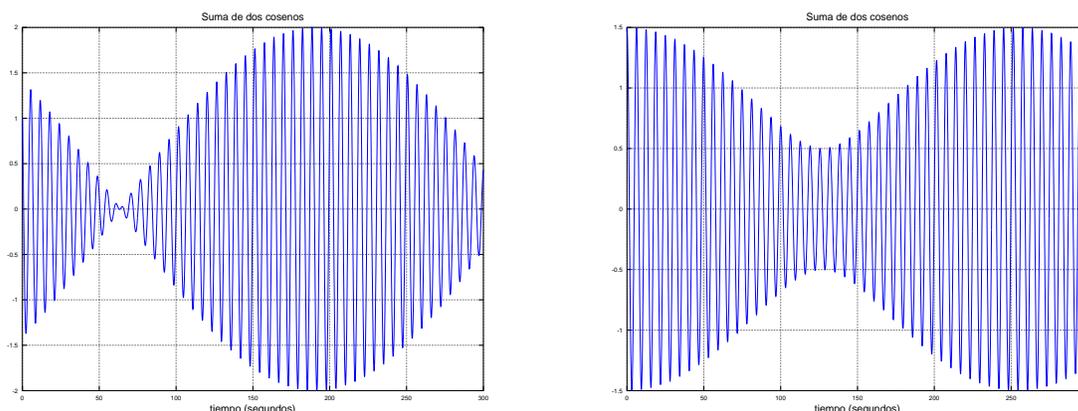


Figura 2.5: Batidos con frecuencias como en la figura 2.4, a la izquierda con amplitudes iguales y desfase de  $\pi/2$  y a la derecha con amplitudes  $A_1 = 1$ ,  $A_2 = 0.5$  y en fase. Comparar con los resultados mostrados en las figuras 2.3 y 2.4.

damente que el que oscila a la frecuencia semidiferencia (porque el período asociado al promedio es menor que el asociado a la semidiferencia), el resultado de la ec. (2.44) permite interpretar la superposición de dos oscilaciones armónicas de frecuencias angulares diferentes como un coseno a la frecuencia promedio pero no con una amplitud constante, sino con una amplitud modulada  $A_m(t)$  (llamada **amplitud de modulación**) que varía en el tiempo como otro coseno a la frecuencia semidiferencia. Notar que la  $\psi(t)$  resultante no es armónica, sino quasi-armónica, ya que la amplitud del coseno a  $\omega_p$  no es constante, y que cuanto más parecidas son las frecuencias  $\omega_1$  y  $\omega_2$  más armónica es la  $\psi(t)$  resultante. Esto se ve claramente en los gráficos de las figuras 2.3 y 2.4: el período de la modulación en la figura 2.4 es mayor que el período de la modulación en la figura 2.3, porque en el segundo caso el acoplamiento es más débil y entonces las frecuencias de resonancia de los dos modos son más parecidas.

Es interesante notar que el oído humano a veces prefiere la suma en la ec. (2.44) (por ejemplo cuando suena un do y un mi no oímos la modulación) y otras veces prefiere el producto y detecta la modulación, como en la demostración con los diapasones que hicimos en clase. Se ha comprobado que la preferencia depende de la diferencia relativa entre las frecuencias  $\omega_1$  y  $\omega_2$  (menor que  $\approx 6\%$ , modulación, mayor, acorde). Recordar que debido a que detectamos la energía contenida en la perturbación acústica, no escuchamos la frecuencia de modulación  $\omega_m$ , sino la que se llama frecuencia de pulsación  $2\omega_m$  (la velocidad de repetición fuerte-débil es la de los ciclos positivo-negativo, que tienen período mitad que los ciclos positivo-positivo). La identidad de la ec. (2.44) se usa en el primer instrumento musical eléctrico, el teremin.

Por último, explorar analíticamente y con Python el caso general en que  $\psi_1(t) = A_1 \cos(\omega_1 t + \varphi_1)$  y  $\psi_2(t) = A_2 \cos(\omega_2 t + \varphi_2)$ . En la figura 2.5 muestro un par de resultados obtenidos con este script. ¿Se puede hacer una interpretación similar a la que hicimos en el caso fácil?



# Capítulo 3

## Muchas partes móviles

Prestaremos especial atención a las *cadena lineales* de partes móviles, un tipo especial de sistemas donde las propiedades físicas permiten mostrar analíticamente muchas características de los modos propios, sin necesidad de diagonalizar de manera explícita la matriz del sistema (que para estos sistemas es tridiagonal). Conviene destacar los siguientes puntos, que muestran la importancia el estudio que haremos a continuación:

- i) bajo ciertas circunstancias, es posible explicitar soluciones para un valor arbitrario de  $N$ , mientras que en el método general no es fácil exhibir el comportamiento de los autovalores y autovectores asociados a los modos sin decir cuánto vale  $N$ ;
- ii) abre un camino para explorar el **límite continuo**, cuando hay tantas partes móviles tan juntas que no se ven sus tamaños. En este límite nos encontraremos por primera vez con la **ecuación de ondas clásica** y veremos que los modos del sistema son las **ondas estacionarias**;
- iii) gracias a la inclusión de partes móviles “virtuales”, es posible avanzar en el tratamiento analítico sin necesidad de explicitar qué pasa en el comienzo y el final de la cadena;
- históricamente fue el primer intento de estudiar las **propiedades emergentes** de un sistema físico a partir de la dinámica de sus partes. Las oscilaciones *longitudinales* de las cadenas, es decir las oscilaciones en la dirección en que se extiende el sistema en el equilibrio (el eje  $x$  en la figura 3.1) fueron estudiadas por Newton en el libro II de sus Principia (1686) en sus intentos de obtener una expresión para la velocidad del sonido. Si parece extraño que Newton usara cadenas de resortes para estudiar el sonido en el aire, más adelante nos quedará claro que era el único camino que tenía, porque para tratar el aire como medio continuo habría necesitado las ecuaciones diferenciales en derivadas parciales, que todavía no existían. Más tarde, John Bernoulli desde Basilea y Daniel Bernoulli (su hijo) desde San Petersburgo, iniciaron en 1727 una conmovedora correspondencia familiar sobre las propiedades de las oscilaciones *transversales* (es decir, en la dirección perpendicular al eje  $x$ ) de cadenas

con  $N$  cuentas equiespaciadas. Entre carta y carta demostraron que hay  $N$  modos con  $N$  frecuencias propias. Más tarde Daniel Bernoulli demostró lo que ya sabemos, que el movimiento más general está dado por la superposición de  $N$  modos y así estableció un principio de superposición que fue el detonante de lo que luego serían las **series de Fourier**. Por eso, no debe extrañar (y hasta resulta lógico cuando se lo piensa), que haya que encarar intuitivamente algunas herramientas que luego se estudiarán con más rigor en materias de matemática más avanzadas, tal como tuvieron que hacer científicos como los Bernoulli, Taylor o Euler.

### 3.1. Cadena con $N$ cuentas idénticas equiespaciadas

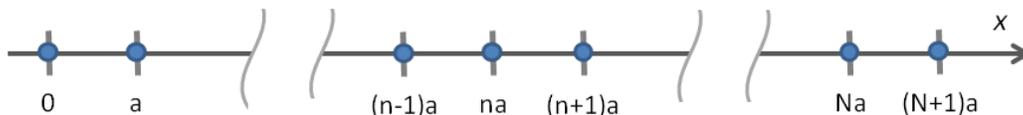


Figura 3.1:  $N$  cuentas idénticas equidistantes (separación  $a$ ) engarzadas en una cuerda elástica sin masa (o unidas por resortes idénticos). Excepto las dos masas de los extremos, el resto de las masas tienen todas dos vecinos cercanos.

Veremos un método especial para resolver el problema de modos normales de un sistema como el bosquejado en la figura 3.1:  $N$  cuentas idénticas de masa  $m$ , equidistantes (separación  $a$ ) y engarzadas en una cuerda elástica sin masa o unidas por resortes idénticos. En el equilibrio, todos los tramos de cuerda tienen igual tensión  $T_0$  y si imaginamos que en vez de cuerdas hay resortes de constante elástica  $k$  y longitud natural  $\ell_0$ , resulta  $T_0 = k(a - \ell_0)$ . La cuenta del comienzo, con  $n = 0$  en  $x = 0$ , y la cuenta del final, con  $n = N + 1$  en  $x = (N + 1)a = L$ , son cuentas virtuales a las que reservaremos un tratamiento aparte cuyos detalles dependerán de cada situación. Así por ejemplo se puede hacer que estos extremos tengan masa infinita para representar una cuerda con extremos fijos, como en una guitarra, o se pueden dejar los extremos libres como si la cuerda fuera una serpiente. Si bien elegimos estudiar apartamientos del equilibrio en la dirección perpendicular a la dirección de equilibrio (problema de oscilaciones transversales), es fácil convencerse de que el mismo tratamiento es válido para el problema de oscilaciones longitudinales.

#### 3.1.1. Ecuaciones de movimiento para oscilaciones transversales

La situación fuera del equilibrio está bosquejada en la figura 3.2. Si llamamos  $T_I$  y  $T_D$  a las tensiones en las sogas a la izquierda y derecha de la parte móvil  $n$  ( $n = 1, 2, \dots, N$ ),  $\ell_I$  y  $\ell_D$  a las longitudes de estas sogas y  $\alpha_I$  y  $\alpha_D$  a los ángulos que estas sogas forman

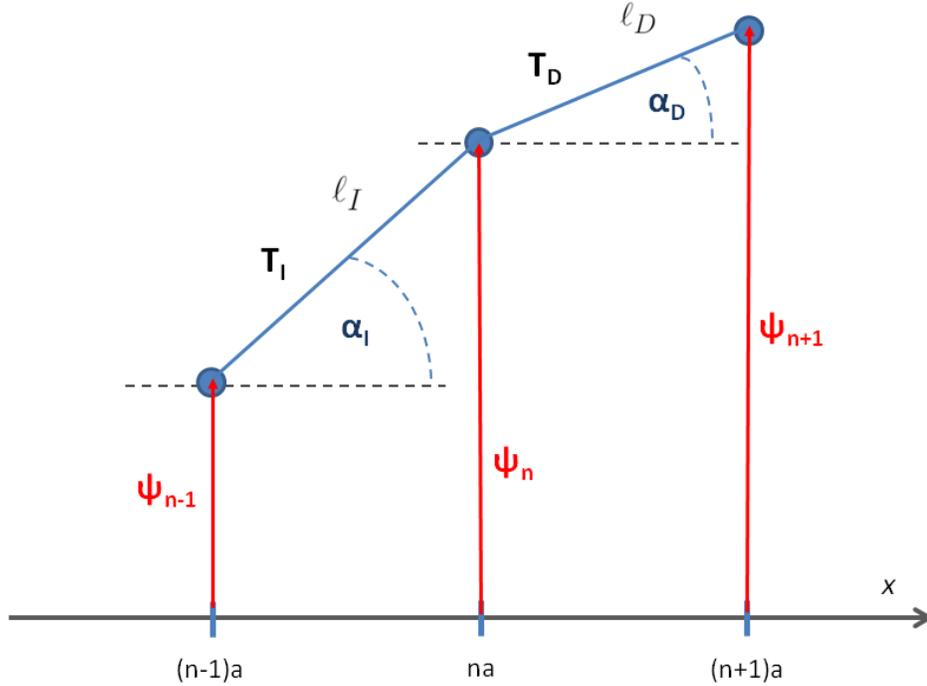


Figura 3.2: Desplazamientos transversales de la cadena con cuentas idénticas equidistantes

con la horizontal, la ecuación de movimiento de la parte móvil  $n$  resulta

$$m \ddot{\psi}_n = -T_I \sin \alpha_I + T_D \sin \alpha_D. \quad (3.1)$$

Este es un sistema de ecuaciones diferenciales *no lineales* en las funciones incógnita  $\psi_n(t)$ , porque  $T_I$  y  $T_D$

$$T_I = k(\ell_I - \ell_0), \quad T_D = k(\ell_D - \ell_0) \quad (3.2)$$

dependen de cuánto se estiraron las sogas ( $\ell_I - \ell_0$  y  $\ell_D - \ell_0$ ) y tanto estos estiramientos

$$\ell_I^2 = a^2 + (\psi_n - \psi_{n-1})^2, \quad \ell_D^2 = a^2 + (\psi_{n+1} - \psi_n)^2, \quad (3.3)$$

como los ángulos  $\alpha_I$  y  $\alpha_D$

$$\sin \alpha_I = \frac{\psi_n - \psi_{n-1}}{\ell_I}, \quad \sin \alpha_D = \frac{\psi_{n+1} - \psi_n}{\ell_D}, \quad (3.4)$$

dependen de la posición de la parte móvil  $n$  y de sus vecinas cercanas, es decir de  $\psi_{n-1}(t)$ ,  $\psi_n(t)$  y  $\psi_{n+1}(t)$ .

Queda claro entonces que el sistema (3.1) no tiene solamente primeras potencias de los desplazamientos, es decir, es no lineal. Para linealizarlo, usamos la hipótesis de pequeños apartamientos del equilibrio. En este caso aproximamos los senos por las tangentes, desarrollamos en serie de Taylor las raíces cuadradas que aparecen en (3.3) y conservamos

solamente términos lineales en los desplazamientos. Se obtiene así el siguiente sistema

$$\ddot{\psi}_n = \frac{T_0}{m} \frac{\psi_{n+1} - \psi_n}{a} - \frac{T_0}{m} \frac{\psi_n - \psi_{n-1}}{a}, \quad (3.5)$$

una forma que conviene recordar cuando encontremos que en el límite continuo la aceleración de una parte móvil es proporcional a una derivada segunda con respecto a la posición de la parte móvil dentro del sistema (etiquetada por el índice discreto  $n$ ).

Notar que si el movimiento es puramente transversal, deben estar compensadas las fuerzas restitutivas horizontales a la izquierda y a la derecha de la masa genérica  $n$ , es decir

$$T_I \cos \alpha_I = T_D \cos \alpha_D,$$

y como en la aproximación lineal los cosenos difieren de la unidad en términos cuadráticos en los desplazamientos, resulta

$$T_I \approx T_D \approx T_0. \quad (3.6)$$

Por conveniencia vamos a distinguir el caso de resortes muy estirables (*slinkies*), donde

$$\ell_0 \ll a, \quad T_I \approx T_D \approx T_0 = ka$$

del caso en que se tiene un resorte común, o una soga elástica, donde

$$\ell_I \approx \ell_D \approx a \quad \text{y} \quad T_I \approx T_D \approx T_0 = k(a - \ell_0).$$

Agrupando términos en (3.5), el sistema de ecuaciones diferenciales linealizado (y homogéneo), para toda masa con dos vecinos cercanos queda

$$\ddot{\psi}_n = \frac{T_0}{ma} (\psi_{n+1} - 2\psi_n + \psi_{n-1}). \quad (3.7)$$

Hay dos maneras equivalentes de pensar teóricamente el tema de los vecinos cercanos. La primera es imaginando que estamos considerando un sistema infinito, que no empieza ni termina nunca, de esta manera queda asegurado que cualquier parte móvil tiene dos vecinos cercanos y entonces (3.7) vale para cualquier valor de  $n$ . La segunda manera es explicitando que el sistema es finito y entonces (3.7) solo vale para  $n = 1, 2, \dots, N$ , pero no para  $n = 0$  (extremo izquierdo) ni para  $n = N + 1$  (extremo derecho).

### 3.1.2. Modos normales

En un modo normal las funciones  $\psi_n(t)$  ( $n = 1, 2, \dots, N$ ) deben representar movimientos donde todas las partes móviles oscilen con la misma frecuencia y pasen simultáneamente por la posición de equilibrio. No queda otra posibilidad que estas funciones adopten la

forma  $\psi_n(t) = A_n \cos(\omega t + \varphi)$ . Y como además tienen que ser solución del sistema (3.7), se obtiene la siguiente *relación de recurrencia*<sup>1</sup> para los  $A_n$

$$A_{n+1} - \left(2 - \frac{ma}{T_0} \omega^2\right) A_n + A_{n-1} = 0. \quad (3.8)$$

La ec. (3.8) se puede reescribir como

$$\frac{A_{n+1} + A_{n-1}}{2 A_n} = 1 - \frac{m a}{2 T_0} \omega^2. \quad (3.9)$$

Un caso fácil de resolver sería cuando se dan como datos **dos** elementos adyacentes de la sucesión  $A_n$  y **la frecuencia**  $\omega$ , porque en este caso con la ayuda de la relación de recurrencia (3.8) podrían obtenerse todos los elementos  $A_n$ . Otro caso fácil de resolver sería cuando se dan como datos **tres** elementos adyacentes, porque (3.9) daría la frecuencia  $\omega$  y también todos los elementos  $A_n$ . Por lo general no estaremos en estos casos fáciles, por ejemplo, en el caso de extremos fijos no se dan ni elementos adyacentes (se dan los elementos  $A_0$  y  $A_{N+1}$ ) ni la frecuencia. Entonces, hay que aguzar el ingenio para resolver la relación de recurrencia y encontrar la dependencia con  $n$  de los elementos  $A_n$  de la sucesión buscada. La ec. (3.9) dice que la dependencia adecuada es aquella para la cual el cociente  $\frac{A_{n+1}+A_{n-1}}{2 A_n}$  no depende de  $n$ . Podríamos ponernos a probar, o preguntarle a un matemático, que seguramente nos va a decir que la dependencia con  $n$  más general de recurrencias como la (3.8) tiene la forma

$$A_n = B \sin(n \Theta) + C \cos(n \Theta), \quad (3.10)$$

donde  $B$ ,  $C$  y  $\Theta$  son tres constantes a determinar (como en los casos fáciles mencionados anteriormente, donde también necesitábamos tres datos). Es fácil ver que la receta (3.10) realmente funciona, es decir, que el cociente  $\frac{A_{n+1}+A_{n-1}}{2 A_n}$  no depende de  $n$ . Efectivamente,

$$A_{n\pm 1} = B \sin[(n \pm 1) \Theta] + C \cos[(n \pm 1) \Theta],$$

y usando las identidades trigonométricas que dan el seno y el coseno de una suma, es fácil ver que

$$\frac{A_{n+1} + A_{n-1}}{2 A_n} = \cos \Theta, \quad (3.11)$$

es decir, que la receta (3.10) funciona independientemente de los valores de  $B$  y  $C$ . Ahora solo resta ajustar el valor de  $\Theta$  para que se cumpla la ec. (3.9). Así obtenemos la siguiente relación entre  $\Theta$ ,  $\omega$  y las magnitudes físicas relevantes del sistema

$$\cos \Theta = 1 - \frac{m a}{2 T_0} \omega^2. \quad (3.12)$$

Conviene notar que el valor de  $\Theta$  correspondiente a cada modo está asociado con un ordenamiento espacial, que se manifiesta como una periodicidad espacial observable en el

---

<sup>1</sup>Una relación de recurrencia vendría a ser la forma discreta de una ecuación diferencial.

conjunto de partes móviles. Esto es así porque la secuencia  $A_n$  de cada modo se obtiene mediante una expresión (la (3.10)) que satisface

$$A_{n+2\pi/\Theta} = A_n,$$

es decir, que se comporta periódicamente, con período  $2\pi/\Theta$ , en la variable  $n$ , lo que equivale a decir que se comporta periódicamente, con período  $2\pi a/\Theta$ , a lo largo del eje  $x$ . Vemos así que la ec. (3.12) es una relación entre el período temporal asociado con  $\omega$  y el período espacial asociado con  $\Theta$ . La relación entre el período temporal (es decir,  $\omega$ ) y el período espacial de un sistema se conoce con el nombre de *relación de dispersión*.

Técnicamente la ec. (3.12) es la *relación de dispersión para los modos transversales de una cadena con  $N$  cuentas idénticas equiespaciadas*. Notar que la forma de la relación de dispersión *siempre* viene dada por (3.12), para cualquier cadena como esta, *independientemente de las condiciones de los extremos*. Las condiciones de contorno en los extremos discretizan los valores de  $\Theta$  y de  $\omega$  permitidos, es decir, determinan las frecuencias de los modos (los valores permitidos para  $\omega$ ) y la distribución espacial de las partes móviles en cada modo (asociada al valor de  $\Theta$  que corresponde a cada valor de  $\omega$ ). Sabemos que tiene que haber  $N$  modos, aunque hasta ahora no hemos logrado explicitarlos, porque hasta ahora no hemos explicitado las condiciones en los extremos. Desde un punto de vista físico queda claro que tienen que ser las condiciones de contorno las que determinen los valores permitidos para  $\Theta$  y para  $\omega$ , porque las maneras que el sistema limitado en el espacio “prefiere” para oscilar colectivamente cuando se lo aparta del equilibrio, tienen que depender de cómo se atan (o no se atan) las fronteras del sistema (en este caso, los extremos de la cadena).

### 3.1.3. Relación de dispersión

Si reescribimos la relación (3.12) entre  $\Theta$ ,  $\omega$  y las magnitudes físicas relevantes del sistema usando una identidad trigonométrica que da el coseno en función del argumento mitad, obtenemos esta forma más linda:

$$\omega^2 = \frac{4T_0}{m a} \sin^2 \frac{\Theta}{2} \implies \omega = \underbrace{2\sqrt{\frac{T_0}{m a}}}_{\omega_{Max}} \left| \sin \frac{\Theta}{2} \right| \quad (3.13)$$

que muestra que aunque todavía no sepamos el valor de  $\Theta$ , la frecuencia  $\omega$  de un modo está comprendida en el intervalo  $[0, \omega_{Max} = 2\sqrt{\frac{T_0}{m a}}]$ . Por ejemplo, en la figura 3.3 se muestra la curva  $\omega(\Theta)$  y tres valores de frecuencia obtenidos para un sistema hipotético con tres partes móviles. Como  $\omega(\Theta)$  es periódica, a cada valor de frecuencia  $\omega_p$  sobre el tramo de curva azul (con  $\Theta_p$  en el intervalo  $0 \leq \Theta_p \leq \pi$ ) le corresponde una doble infinitud de valores de  $\Theta$  fuera de dicho intervalo: una infinitud de la forma  $\Theta_p + 2\pi q$  (en los tramos con pendiente positiva, los símbolos rojos) y otra infinitud de la forma  $-\Theta_p + 2\pi q$  (en los tramos con pendiente negativa, los símbolos verdes),  $q$  un número

entero. Es fácil ver que todos los valores de la infinitud  $\Theta_p + 2\pi q$  representan físicamente el mismo movimiento colectivo, porque según la ec. (3.10) todos tienen exactamente los mismos valores de  $A_n$ . Dicho de otra manera,  $A_n(\Theta_p + 2\pi q) = A_n(\Theta_p)$ . Los valores de la infinitud  $-\Theta_p + 2\pi q$  también representan la misma situación física, porque el lado derecho de la igualdad (3.11) toma el mismo valor tanto para  $\Theta_p$  como para  $-\Theta_p + 2\pi q$ , definiendo así la misma recurrencia, es decir, la misma sucesión de amplitudes modales  $A_n$ .

Para ver cómo en qué puntos de la curva de la figura 3.3 caen los simbolitos rojos (y los verdes), es decir, para ver cómo funciona la discretización de los valores de  $\Theta_p$  (y en consecuencia, de  $\omega_p$ ), especifiquemos las condiciones en los extremos, que es la información que hasta ahora hemos evitado explicitar. Las idealizaciones más comunes son los casos de extremo fijo y de extremo libre.

### 3.1.4. Modos para dos extremos fijos

Consideramos primero el caso de ambos extremos fijos. En este caso los extremos en  $x = 0$  y en  $x = L$ , figura 3.1, están atados a una pared, lo que equivale a decir que las masas virtuales 0 y  $N + 1$  tienen masa tan grande que no se pueden mover. Para los modos esto se traduce en que se anulen las amplitudes de las partes móviles virtuales, es decir,  $A_0 = A_{N+1} = 0$ . Y de la ec. (3.10) y para que  $A_0 = 0$  debe ser  $C = 0$ . Y si  $C = 0$ , entonces  $B \sin[(N + 1)\Theta] = 0$ , una condición que se cumple cuando  $B = 0$  (que no nos interesa, porque no representa movimiento) o cuando  $\sin[(N + 1)\Theta] = 0$ . En resumen, para que en la cadena con extremos fijos haya un modo, la cantidad  $(N + 1)\Theta$  debe ser una raíz de la función seno, es decir, que debe ser  $(N + 1)\Theta = p\pi$ , con  $p$  un número

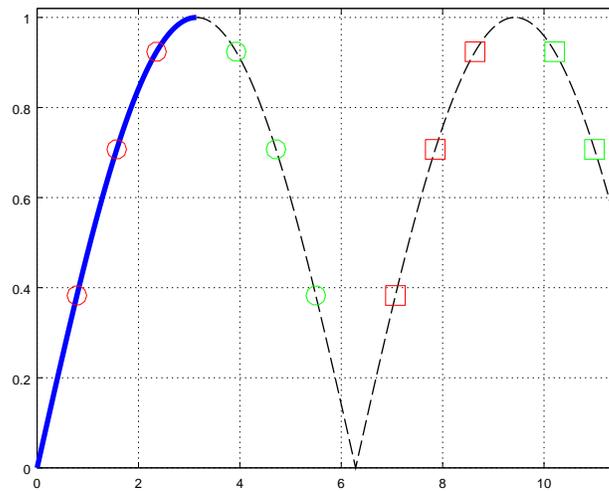


Figura 3.3: Forma de la ecuación de dispersión  $\omega(\Theta)$ , ec. (3.13),  $\Theta$  en el eje horizontal y  $\omega/\omega_{Max}$  en el eje vertical.

entero arbitrario. Los modos quedan así etiquetados por este índice  $p$ . En cada modo los desplazamientos de las partes móviles son

$$\psi_n^{(p)}(t) = A_n^{(p)} \cos(\omega_p t + \varphi_p) = \overbrace{B^{(p)} \sin \left[ \frac{p \pi}{(N+1)} n \right]}^{A_n^{(p)}} \cos(\omega_p t + \varphi_p), \quad (3.14)$$

$\Theta^{(p)}$

con

$$\omega_p = \omega_{Max} \left| \sin \frac{p \pi}{2(N+1)} \right|, \quad (3.15)$$

En el método matricial los modos se visualizan a través de dos cantidades: los autovalores, que dan el comportamiento temporal del conjunto, y los autovectores, que dan el comportamiento espacial de cada parte móvil en la oscilación colectiva. Análogamente, en el método de recurrencias que estamos estudiando, cada modo queda caracterizado por dos cantidades que podemos calcular en términos de  $p$  y de las magnitudes físicas relevantes del sistema ( $T_0$ ,  $a$ ,  $m$ ,  $N$  y  $L = (N+1)a$ ). Estas cantidades son

$$\Theta^{(p)} = \frac{p \pi}{(N+1)} = \frac{p \pi a}{L},$$

que da el *comportamiento espacial* de cada parte móvil en la oscilación colectiva, y

$$\omega_p(\Theta^{(p)}),$$

que se obtiene de la ecuación de dispersión (3.13) y que rige el *comportamiento temporal* del conjunto.

Cuando hay muchas partes móviles, puede ser conveniente etiquetar la  $n$ -ésima parte móvil con la coordenada  $x = na$  (ver figura 3.1). Esta coordenada permite indentificar a la parte móvil como aquella que en el equilibrio está ubicada sobre el eje  $x$  en la posición  $na$  respecto a un origen. En esta descripción, y entendiendo que  $x$  toma solamente valores discretos, podemos reemplazar en el argumento del primer seno en la ec. (3.14)

$$\Theta^{(p)} n \quad \text{por} \quad \underbrace{\frac{\Theta^{(p)}}{a}}_{k_p} x. \quad (3.16)$$

Este reemplazo pone en evidencia que en los modos de la ec. (3.14) la cantidad  $k_p = \Theta^{(p)}/a$  juega el papel de una frecuencia espacial y que entonces el período espacial de las funciones de  $x$  (de  $n$ ) asociadas con los modos es

$$\lambda_p = \frac{2\pi}{k_p} = \frac{2L}{p}. \quad (3.17)$$

Al período espacial asociado con movimientos armónicos colectivos de sistemas (limitados o no) se lo llama **longitud de onda** y se denota con la letra griega  $\lambda$  (lambda minúscula)

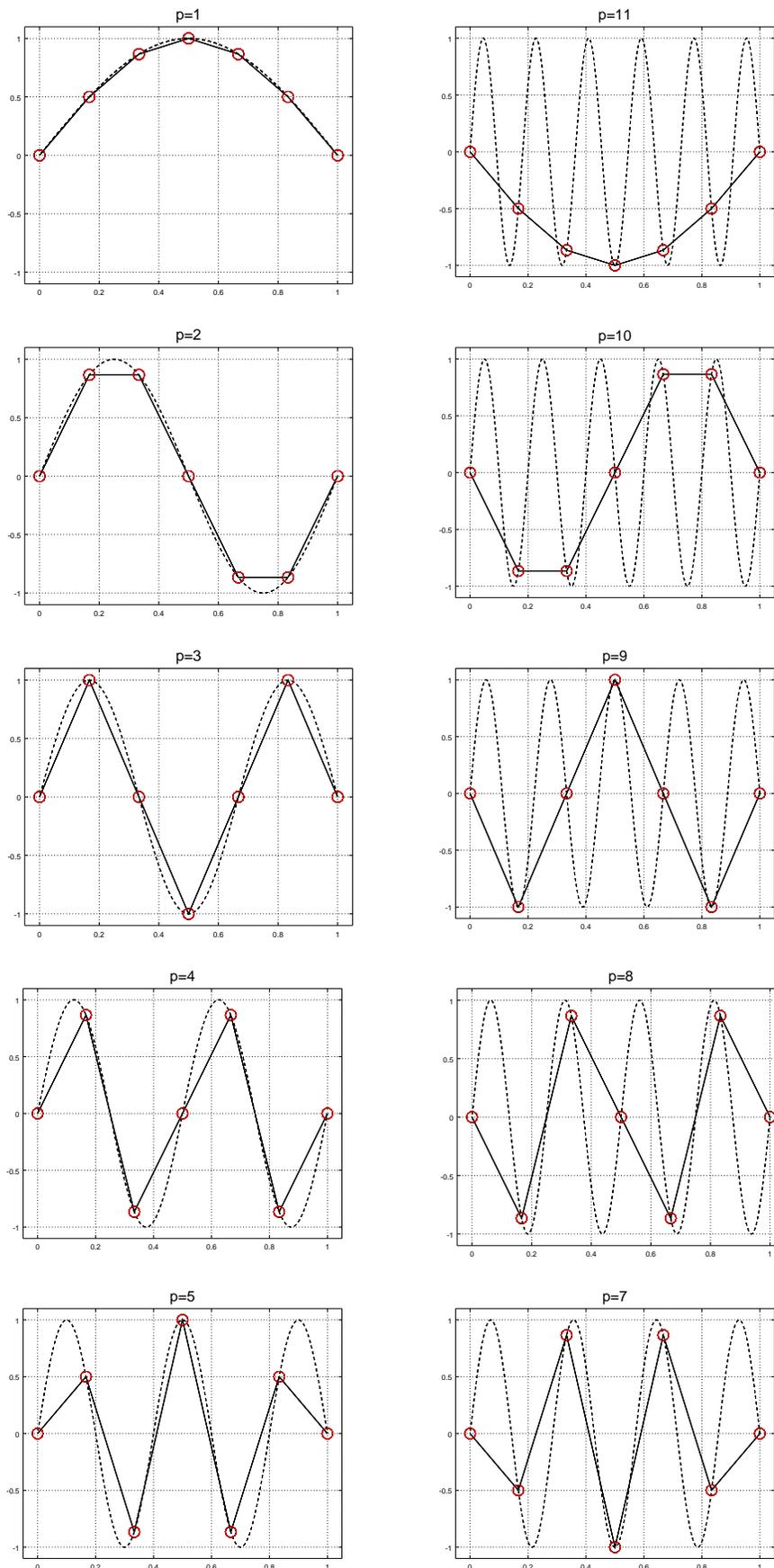


Figura 3.4: Forma de los modos para distintos valores de  $p$  en la ec. (3.14).

y a la frecuencia angular espacial asociada con este período se lo llama **número de onda** y se denota con la letra  $k$  (una elección muy desafortunada cuando hay resortes involucrados).

Notemos que en el caso de extremos fijos se verifican las propiedades que ya habíamos anunciado en la figura 3.3 para el caso general de extremos con condiciones arbitrarias. En primer lugar vemos que hay solamente  $N$  frecuencias normales en el tramo azul de la curva de la figura 3.3 y son las que tienen  $\Theta^{(p)} = p\pi/(N+1)$  con  $p = 1, 2, \dots, N-1, N$  (los circulitos rojos). El valor  $p = 0$  está excluido porque con frecuencia cero no hay movimiento colectivo y el valor  $p = N+1$  también porque para  $\Theta^{(N+1)} = \pi$ ,  $\omega_{(N+1)} = \omega_{Max}$  y tampoco hay movimiento colectivo. En segundo lugar, las otras  $N$  soluciones en el siguiente tramo de la curva de dispersión de la ec. (3.13), el tramo de trazos discontinuos y con pendiente negativa (los circulitos verdes en la figura 3.3), corresponden a  $p = N+2, N+3, \dots, 2N+1, 2N+2$  y guardan la siguiente correspondencia con las frecuencias del tramo anterior

$$\underbrace{\omega_{N+1+j}}_{\text{circulitos verdes}} = \underbrace{\omega_{N+1-j}}_{\text{circulitos rojos}}, \quad j = 1, 2, \dots, N-1, N. \quad (3.18)$$

Por lo discutido en la sección 3.1.3, las soluciones con  $p = N+2, N+3, \dots, 2N+1, 2N+2$ , describen configuraciones de modos que ya han sido tenidas en cuenta por las amplitudes relativas de las partes móviles obtenidas para las soluciones con  $p = 1, 2, \dots, N-1, N$ . Esto está ilustrado en la figura 3.4, donde se muestra la secuencia  $A_n^p$  en una cadena con  $N = 5$  cuentas y con extremos fijos. Las curvas de trazos representan las funciones  $A_n^{(p)}$  de la ec. (3.15) y solamente sirven como una guía para ubicar las partes móviles. No representan la forma de la cuerda con cuentas, porque las sogas (o los resortes) entre cuentas están tensas y entonces corresponden a tramos rectos de las figuras. En la columna de la izquierda se muestran las secuencias  $A_n^p$  para  $1 \leq p \leq N = 5$  (circulitos rojos en la figura 3.3), mientras que en la columna de la izquierda se muestran las secuencias  $A_n^p$  para  $N+1 = 6 \leq p \leq 2N = 10$  (circulitos verdes en la figura 3.3). Observar que la secuencia  $A_n$  para las soluciones con  $p = N+1+j$  (verdes), difiere en un factor constante de la secuencia  $A_n$  para soluciones con  $p = N+1-j$  (rojas), con  $j = 1, 2, \dots, N-1, N$ . Nótese que las configuraciones espaciales para los casos  $p = 0$  y  $p = 6$  no se muestran, ya que corresponden a todas las cuentas en su posición de equilibrio.

Los modos más bajos tienen  $\Theta^{(p)}$  más chico, es decir que en el dibujo de la figura 3.3 están ubicados sobre la parte de la curva de dispersión donde el seno se puede aproximar por una función lineal. Entonces, para los modos más bajos,  $\omega$  depende linealmente de  $k$ , es decir que los modos más bajos están sobre el tramo casi lineal de la parte azul de la figura 3.3. Notar que en este caso, los valores de las frecuencias de los modos son múltiplos de la frecuencia más baja o fundamental

$$\omega_p \approx \omega_{Max} \frac{p\pi}{2(N+1)} = p \sqrt{\frac{T_0 a}{m}} \frac{\pi}{L} = p\omega_1, \quad (3.19)$$

con

$$\omega_1 = \sqrt{\frac{T_0 a \pi}{m L}}.$$

Resumiendo, siempre que se satisfagan las hipótesis necesarias para que se cumpla la aproximación lineal, el movimiento más general de cada parte móvil de la cadena de masas idénticas y equiespaciadas con extremos fijos y condiciones iniciales arbitrarias, se escribe como superposición de los modos dados por la ec. (3.14)

$$\psi_n(t) = \sum_{p=1}^N B^{(p)} \sin\left[\frac{p\pi}{(N+1)}n\right] \cos(\omega_p t + \varphi_p), \quad (3.20)$$

con  $n = 1, 2, \dots, N$  (etiqueta de cada parte móvil) y  $\omega_p$  dado por la ec. (3.15).

### Condiciones iniciales

En nuestra solución general obtenida en (3.20), las  $N$  amplitudes  $B^{(p)}$  y las  $N$  fases iniciales  $\varphi_p$  son incógnitas que se obtienen a partir de las  $2N$  condiciones iniciales:  $N$  posiciones iniciales de las partes móviles

$$\psi_n(0) = \sum_{p=1}^N B^{(p)} \sin\left[\frac{p\pi}{(N+1)}n\right] \cos(\varphi_p), \quad (3.21)$$

y otras  $N$  velocidades iniciales de las partes móviles

$$\dot{\psi}_n(0) = - \sum_{p=1}^N \omega_p B^{(p)} \sin\left[\frac{p\pi}{(N+1)}n\right] \sin(\varphi_p). \quad (3.22)$$

Teniendo en cuenta que cualquier condición inicial se puede representar con un dibujito como los que están en la columna izquierda de la figura 3.4, pero con las partes móviles en posiciones bastante arbitrarias (como que las sogas/resortitos sigan siendo elásticas y que se comporten linealmente), el resultado de la ec. (3.21) es bastante sorprendente. Porque esta igualdad está diciendo que cualquier dibujito arbitrario para las posiciones iniciales de las  $N$  partes móviles equiespaciadas se puede escribir como combinación lineal de los dibujitos “normales”, es decir, los dibujitos que corresponden a los modos. Para pensar ... este resultado está relacionado con el hecho que los autovectores que generan el cambio de base son ortogonales ... a retomar más adelante <sup>2</sup>. Por ahora, nos podríamos quedar con la demostración de la siguiente identidad

$$\sum_{n=1}^N \sin\left[\frac{p\pi}{(N+1)}n\right] \sin\left[\frac{q\pi}{(N+1)}n\right] = \frac{N+1}{2} \delta_{pq}, \quad (3.23)$$

con  $\delta_{pq}$  la delta de Kronecker (igual a 1 si  $p = q$ , o igual a 0 cuando  $p \neq q$ ).

---

<sup>2</sup>Prestar atención, actualizaciones pronto

### 3.1.5. Condición para extremo libre

Supongamos que un extremo de la cadena, por ejemplo el de la derecha, ver figura 3.1, no está fijo sino que tiene un anillo de masa  $m_{(N+1)}$  que desliza sin rozamiento por una varilla vertical. Para este anillo, la ecuación de movimiento es parecida a la (3.5), pero con  $M_{N+1}$  en vez de  $m$  y sin el aporte del vecino derecho

$$M_{N+1} \ddot{\psi}_{N+1} = -T_0 \frac{\psi_{N+1} - \psi_N}{a}. \quad (3.24)$$

Vemos que esta condición de extremo es bastante general, en el sentido que incluye al caso anterior de extremo fijo, que correspondería a un anillo de masa muy grande ( $M_{N+1} \rightarrow \infty$ ). Es común idealizar un extremo libre como el límite opuesto, es decir, un anillo sin masa ( $M_{N+1} \rightarrow 0$ ). En este caso, la condición (3.24) exige

$$\psi_{N+1} = \psi_N$$

y esta igualdad provee una relación entre las constantes  $B$ ,  $C$  y  $\Theta$  que determinan las amplitudes  $A_n$  según la ec. (3.10). Otra relación entre las mismas constantes se obtiene cuando se explicita la condición sobre la masa virtual del extremo izquierdo. Procediendo de una manera completamente análoga a la empleada para extremos fijos, se obtiene ahora la discretización de los modos, que de nuevo quedan etiquetados por un índice discreto  $p$ .

### 3.1.6. Oscilaciones longitudinales

Hasta aquí hemos estudiado las oscilaciones **transversales** de la cadena con  $N$  cuentas idénticas equiespaciadas. Es fácil comprobar que las oscilaciones **longitudinales** (es decir, los apartamientos del equilibrio en la dirección que ocupa el sistema cuando está en reposo) admiten un tratamiento formalmente idéntico. Si las funciones  $\psi_n(t)$  representan ahora los apartamientos longitudinales, como en la figura 3.5, en vez de los apartamientos transversales como en la figura 3.2, las ecuaciones de movimiento también tienen la forma (3.7), con  $T_0$  la tensión de las sogas en el equilibrio (caso sogas elásticas) o  $T_0 = ka$  (caso de resortes muy estirables). Luego, los modos normales se obtienen siguiendo el tratamiento ya visto en 3.1.2, la relación de dispersión sigue siendo (3.13), las frecuencias de resonancia para extremos fijos vienen dadas por la ec. (3.15) y las coordenadas normales para los distintos modos están representadas por gráficos como los de la figura 3.4, con la única diferencia de que las funciones seno asociadas con cada modo no se visualizan tan directamente como en el caso de las oscilaciones transversales.

## 3.2. Péndulos idénticos acoplados

En las cadenas lineales vistas hasta ahora, las ecuaciones de movimiento en la aproximación lineal tenían la forma (3.5), o (3.7), tanto para oscilaciones longitudinales como

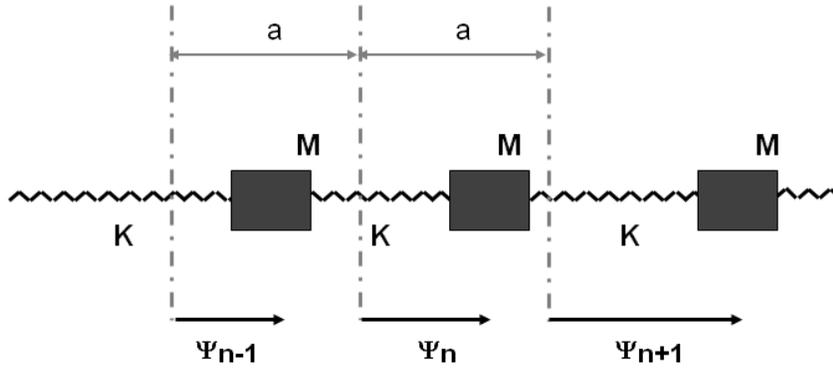


Figura 3.5: Desplazamientos longitudinales de la cadena con cuentas idénticas equidistantes

para transversales. Es claro que tal sistema de ecuaciones diferenciales acopladas conducirá siempre a relaciones de dispersión de la misma forma que (3.13). Esto quiere decir que, independientemente de las condiciones en los extremos, y tal como se representa en la figura 3.3, las frecuencias de los modos normales estarán comprendidas en el intervalo  $0 \leq \omega \leq \omega_{Max}$ . Que las frecuencias de resonancia del sistema tengan como cota mínima el valor nulo, está claramente relacionado con el hecho de que la fuerza resitutiva sobre una parte móvil, proveniente solo de los vecinos cercanos, son proporcionales a los desplazamientos relativos entre vecinos cercanos y, en consecuencia, la fuerza resitutiva es cero cuando las posiciones relativas entre partes móviles no cambian. A continuación veremos un sistema donde, además de estas características, aparece también una nueva fuerza resitutiva que depende solamente del desplazamiento de la parte móvil de su posición de equilibrio. Y entonces, la fuerza resitutiva total ya no es cero, ni siquiera en el caso en que las posiciones relativas entre partes móviles no cambien. El sistema en cuestión consiste en un conjunto de péndulos idénticos (masa  $M$ , longitud  $\ell$ ), acoplados por resortes también idénticos (constante elástica  $K$ ), esquematizado en la figura 3.6. En este caso, y de acuerdo a lo discutido anteriormente, el sistema de ecuaciones de Newton se escribe igual que el sistema (3.7), excepto por un término adicional que representa la fuerza resitutiva gravitatoria

$$\ddot{\psi}_n = -\omega_0^2 \psi_n + \frac{K}{M} (\psi_{n+1} - 2\psi_n + \psi_{n-1}), \quad (3.25)$$

con  $\omega_0^2 = g/\ell$ .

Para analizar los modos de este sistema de péndulos acoplados, partimos de la ec. (3.25) y repetimos de manera formalmente idéntica los procedimientos realizados para analizar los modos en la eq. (3.7), es decir el caso  $\omega_0 = 0$ , página 48. Es claro que si antes llegábamos a una eq. en diferencias que se puede reescribir como

$$\omega^2 = 2\frac{K}{M} \left( 1 - \frac{A_{n+1} + A_{n-1}}{2A_n} \right), \quad (3.26)$$

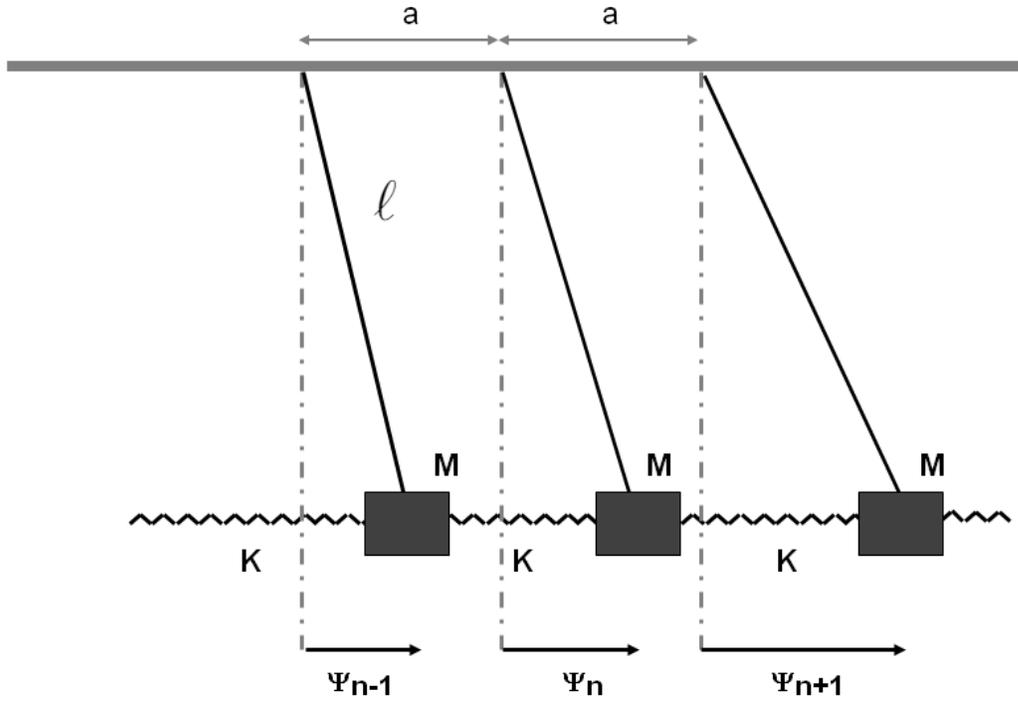


Figura 3.6: Esquema para analizar un conjunto de péndulos idénticos, acoplados mediante resortes también idénticos, un modelo de sistema donde actúan dos tipos de fuerzas restitutivas, i) las elásticas, que dependen de la posición relativa entre vecinos cercanos y ii) la gravitatoria, que no depende de la posición relativa entre vecinos cercanos, sino sólo del apartamiento local del equilibrio.

(ver (3.8)), ahora vamos a llegar a una ecuación en diferencias similar, pero con el nuevo término que corresponde a  $\omega_0 \neq 0$

$$\omega^2 = \omega_0^2 + 2 \frac{K}{M} \left( 1 - \frac{A_{n+1} + A_{n-1}}{2 A_n} \right), \quad (3.27)$$

que se puede interpretar recordando que el lado derecho de (5.12) da la fuerza restitutiva por unidad de desplazamiento y por unidad de masa cuando  $\omega_0 = 0$  (no hay gravedad o péndulos muy largos). Cuando  $\omega_0 \neq 0$ , la fuerza restitutiva aumenta de la misma manera para todas las masas y se suma a la la fuerza restitutiva por unidad de desplazamiento y por unidad de masa que ya tenían las partes móviles sin gravedad.

Y como ya sabemos que el cociente entre amplitudes del lado derecho de la ec. (5.12) se puede hacer independiente de  $n$  con sucesiones de la forma

$$A_n = B \sin(n \Theta) + C \cos(n \Theta), \quad (3.28)$$

llegamos así a una nueva relación de dispersión (ver eq. (3.13))

$$\omega^2 = \omega_0^2 + \underbrace{\frac{4K}{M}}_{\omega_{MAX}^2} \sin^2 \frac{\Theta}{2}, \quad (3.29)$$

válida para la relación entre el período temporal y el período espacial en un modo. Como sucedía en el caso  $\omega_0 = 0$ , las frecuencias de los modos quedan discretizadas cuando se imponen las condiciones en los extremos, aunque siempre estarán ubicadas sobre la curva  $\omega(\Theta)$  definida por la ec. y las situaciones de interés físico serán las que correspondan al intervalo  $0 \leq \Theta \leq \pi$  (o cualquier otro intervalo, aunque por simplicidad conviene elegir  $[0, \pi]$ ).

En el caso en que los péndulos virtuales de los extremos estén fijos a una pared, el tratamiento es completamente similar al que se hizo para la cadena con cuentas, donde  $\omega_0 = 0$ , adaptando el significado de  $\psi_n(t)$ . De esta manera, vemos que el movimiento más general de cada parte móvil del sistema esquematizado en la figura 3.6, con extremos fijos y condiciones iniciales arbitrarias, se escribe también como (3.15), es decir como superposición de modos dados por la ec. (3.14), que solo difieren en la relación de dispersión, es decir en la manera en que  $\omega_p$  se obtiene a partir de  $\Theta_p$ .

### 3.3. Aproximación continua

Veremos cómo tratar el caso límite en que el número de partes móviles del sistema es muy grande. La idea general se aplica a muchos sistemas reales y para fijar ideas podríamos pensar en las oscilaciones transversales de una cadena tensionada con  $N$  cuentas idénticas equiespaciadas, como en las figuras 3.1 o 3.2, con cuentas cada vez más chiquitas y más juntas, tan chiquitas y tan juntas que no seremos capaces de ver la separación entre cuentas a simple vista. Es decir que no podremos decir si se trata de las oscilaciones transversales de un collar con cuentas o de una cuerda. Análogamente, también podríamos pensar en las oscilaciones longitudinales de una cadena similar, como en la figura 3.5, o de sistemas con otras fuerzas restitutivas, como el sistema de péndulos acoplados de la figura 3.6. En todos estos casos esperamos que, a partir de un valor suficientemente grande de  $N$ , el número de partes móviles del sistema deje de importar y que solamente importe la inercia o la elasticidad de una cierta longitud del sistema. Es decir, esperamos que a partir de un dado valor de  $N$ , el movimiento colectivo de un metro de sistema cuando hay  $10^{23}$  partes móviles con masa  $m$ , sea casi indistinguible del movimiento colectivo de un metro de sistema cuando hay el doble de partes móviles con masa  $m/2$ , o cuando hay el cuádruple de partes móviles con masa  $m/4$ , o el óctuple con masa  $m/8$ , etc. En estos casos es conveniente cambiar la notación y dejar de etiquetar a las partes móviles con un índice discreto, porque cada vez nos interesa menos la individualidad de cada parte móvil, dicho esto sin intención de ofender a ninguna parte móvil.

Una alternativa al índice discreto para etiquetar partes móviles sería pintar el sistema de distintos colores y describir el movimiento diciendo “ahí está la parte azul, ahí está la parte verde, etc.”. No vamos a usar esta alternativa, porque no siempre se puede pintar lo que uno quiere estudiar. La alternativa que vamos a usar para etiquetar las partes móviles es relacionarlas con las posiciones que tenían en el equilibrio. Por ejemplo, para sistemas como los de las figuras 3.1, 3.5, o 3.6, la etiqueta sería la coordenada  $x$ , Entonces,

en vez de describir los desplazamientos mediante un conjunto numerable de funciones de una variable  $\psi_n(t)$ , usaremos una función de dos variables  $\psi(x, t)$ . La primera variable,  $x$ , juega el papel del índice discreto  $n$  y hace referencia a la parte móvil que estamos mirando, que es aquella que en el equilibrio estaba en la posición  $x = na$ . La segunda variable,  $t$ , sigue siendo el tiempo. En síntesis: para pasar en cada caso el sistema de ecuaciones diferenciales a la nueva notación, hay que hacer los siguientes reemplazos

$$\begin{aligned}\psi_n(t) &\longrightarrow \psi(x = na, t) \\ \psi_{n+1}(t) &\longrightarrow \psi(x + a, t) \\ \psi_{n-1}(t) &\longrightarrow \psi(x - a, t) \\ \ddot{\psi}_n(t) &\longrightarrow \partial^2 \psi(x = na, t) / \partial t^2.\end{aligned}\tag{3.30}$$

Notar que para sistemas bidimensionales, por ejemplo la tela tensa de una cama elástica, deberíamos usar las dos coordenadas cartesianas  $x$  e  $y$ , o las coordenadas polares  $\rho$  y  $\theta$ . Y para un sistema tridimensional, por ejemplo un gas en un recipiente, deberíamos usar las tres coordenadas cartesianas  $x$ ,  $y$  y  $z$ , o coordenadas esféricas.

### 3.3.1. Cuerda $\implies$ ecuación de ondas clásica

En la nueva notación y en el caso de oscilaciones transversales de la cadena con cuentas, el sistema de ecuaciones diferenciales (3.7) queda

$$\frac{\partial^2 \psi}{\partial t^2} = \frac{T_0}{m a} \left[ \psi(x + a, t) - 2\psi(x, t) + \psi(x - a, t) \right].\tag{3.31}$$

Observar que hasta acá solamente se trata de un cambio de notación, porque  $x$  toma solamente los valores discretos  $na$ . Pero si ahora pensamos en el límite físico de muchas partes móviles muy juntas, entonces podemos aproximar  $x$  por una variable continua, que es en el fondo lo que hacemos implícitamente en la ferretería cuando pedimos una cierta longitud de cable, en vez de especificar el número de partes móviles. En este caso,  $\psi(x, t)$  es una función de  $x$  que da la foto del sistema en cada instante  $t$ . Y como  $a$  es muy chiquito y si la foto no es muy “patológica”<sup>3</sup>, entonces podemos desarrollar  $\psi(x \pm a, t)$  en serie de Taylor alrededor de  $x$

$$\psi(x \pm a, t) = \psi(x, t) \pm a \frac{\partial \psi}{\partial x}(x, t) + a^2 \frac{1}{2!} \frac{\partial^2 \psi}{\partial x^2} \pm a^3 \frac{1}{3!} \frac{\partial^3 \psi}{\partial x^3} + a^4 \frac{1}{4!} \frac{\partial^4 \psi}{\partial x^4} + \dots.\tag{3.32}$$

Introduciendo estos desarrollos en (3.31), todos los términos con derivadas impares se cancelan, mientras que los términos con derivadas pares se refuerzan. Así obtenemos

$$\frac{\partial^2 \psi}{\partial t^2} = \frac{T_0}{m a} \left[ a^2 \frac{\partial^2 \psi}{\partial x^2} + a^4 \frac{2}{4!} \frac{\partial^4 \psi}{\partial x^4} + \dots \right]\tag{3.33}$$

<sup>3</sup>es decir, suponemos que en las fotos de sistemas reales las derivadas que aparecen en los desarrollos de Taylor están bien definidas

una expresión que diría que la aceleración vertical de este sistema, en el límite matemático  $a \rightarrow 0$ , es nula. Un resultado feo pero consistente con el hecho de que cuando pusimos más partes en el sistema, dejamos constante la masa de cada parte móvil. Y en el límite de  $N \rightarrow \infty$  y  $a \rightarrow 0$ , llegamos a un sistema con masa infinita, por eso la aceleración vertical resulta cero, porque no lo podemos mover.

Pero el procedimiento anterior, si bien correcto matemáticamente, es inconsistente con la motivación física inicial, donde esperábamos que a partir de un dado valor de  $N$  no importara el número de partes móviles contenidas en una longitud fija del sistema. Para que la masa del sistema con longitud  $L$  sea independiente del número de partes móviles, tenemos que pedir que en este límite la cantidad  $m/a$  permanezca constante. Vemos así que en nuestra descripción aparece naturalmente una magnitud  $\rho$

$$\rho = \lim_{a \rightarrow 0} \frac{m}{a}, \quad (3.34)$$

llamada densidad lineal de masa, o masa por unidad de longitud, que en el caso de cuentas idénticas equiespaciadas toma un valor independiente de  $x$ . Así, el lado izquierdo de (3.31) no tiende a cero cuando  $a \rightarrow 0$  y el sistema de ecuaciones para el movimiento del sistema se transforma en una proporcionalidad entre dos derivadas segundas parciales

$$\frac{\partial^2 \psi}{\partial t^2} = \frac{T_0}{\underbrace{\rho}_{v^2}} \frac{\partial^2 \psi}{\partial x^2}, \quad (3.35)$$

la derivada segunda temporal (la aceleración) y la derivada segunda espacial (relacionada con la fuerza restitutiva). Esta ecuación es la **ecuación de ondas clásicas** en su versión unidimensional para un sistema que ocupa una dirección fija en el espacio (o sea, unidimensional en cartesianas, no en esféricas). El factor de proporcionalidad entre las dos derivadas segundas parciales en (3.35)

$$v^2 = \frac{T_0}{\rho}, \quad (3.36)$$

tiene unidades de velocidad al cuadrado,  $v$  recibe el nombre de **velocidad de fase** y ya podemos intuir que es una velocidad asociada con la onda.

A primera vista es sorprendente que el conjunto de ecuaciones diferenciales ordinarias acopladas que describía la dinámica del conjunto de partes móviles del sistema, haya colapsado en esta sencilla ecuación diferencial en derivadas parciales. Aunque veremos que no tan sencilla, en particular cuando vayamos descubriendo la gran diversidad de soluciones que admite y la gran cantidad de fenómenos físicos que estas soluciones describen.

### 3.3.2. Slinky $\implies$ ecuación de ondas clásica

Como se discutió en 3.1.6, las oscilaciones longitudinales de una cadena con  $N$  cuentas idénticas equiespaciadas unidas por resortes idénticos, también tienen ecuaciones de movimiento de la misma forma que (3.7), excepto que ahora  $\psi_n(t)$  representan apartamentos

longitudinales, y que  $T_0 = ka$ , porque en vez de sogas hay resortes. La ecuación análoga a (3.7) se escribe como

$$\ddot{\psi}_n = \frac{K}{M} (\psi_{n+1} - 2\psi_n + \psi_{n-1}) . \quad (3.37)$$

Aparte de estos detalles, sigue siendo válido el cambio de notación (3.31) y también podemos repetir los desarrollos (3.32) para prepararnos para el límite  $a \rightarrow 0$ . Al reemplazar en (3.37), se siguen cancelando los términos con derivadas impares y reforzando los términos con derivadas pares y así obtenemos

$$\frac{\partial^2 \psi}{\partial t^2} = \frac{k a^2}{m} \left[ \frac{\partial^2 \psi}{\partial x^2} + a^2 \frac{2}{4!} \frac{\partial^4 \psi}{\partial x^4} + \dots \right] \quad (3.38)$$

La aceleración en el límite matemático  $a \rightarrow 0$  de nuevo parecería nula, pero en nuestro límite físico queremos que las propiedades del sistema con longitud  $L$  sean independientes del número de partes móviles. Aplicado a la masa, y al igual que en el caso de las sogas con cuentas, esto impone que la cantidad  $m/a$  permanezca constante cuando  $a \rightarrow 0$ . Y aplicado a la elasticidad equivalente, esto impone que a medida que se agregan resortes, se cambie la constante  $k$  de los mismos, de manera tal que el resorte equivalente de longitud  $L$  formado por resortes en serie con constantes  $k$ , tenga una elasticidad que sea independiente de  $N$ , es decir de  $a$ . Como la constante equivalente  $K_e$  de  $N$  resortes en serie con la misma constante  $k$  es

$$\frac{1}{K_e} = \frac{1}{k} + \frac{1}{k} + \dots + \frac{1}{k} = \frac{N}{k} = \frac{L}{ak}$$

y teniendo en cuenta que la propiedad extensiva (la que se suma cuando se aumenta el número de partes del sistema) es la inversa de las constantes de los resortes, en el límite físico tenemos que pedir que la constante inversa del resorte por unidad de longitud,

$$\frac{1}{K_e L}$$

sea una propiedad del sistema, independientemente de la longitud. Entonces, la cantidad  $ka^2/m$  que aparece en (3.38) tiende a una constante

$$\frac{ka^2}{m} = \frac{K_e L}{\rho}$$

y en el límite  $a \rightarrow 0$ , (3.38) se transforma, igual que en (3.35), en una proporcionalidad entre las dos derivadas segundas parciales

$$\frac{\partial^2 \psi}{\partial t^2} = \underbrace{\frac{K_e L}{\rho}}_{v^2} \frac{\partial^2 \psi}{\partial x^2} . \quad (3.39)$$

De nuevo encontramos la **ecuación de ondas clásicas** en su versión unidimensional para un sistema que ocupa una dirección fija en el espacio, pero ahora la velocidad de fase, el factor de proporcionalidad entre las dos derivadas segundas parciales con unidades de velocidad al cuadrado, es

$$v^2 = \frac{K_e L}{\rho} . \quad (3.40)$$

### 3.3.3. Péndulos acoplados $\implies$ ecuación de Klein-Gordon

De manera completamente análoga a lo hecho en 3.3.1 y 3.3.2, empleemos ahora la descripción continua para estudiar el caso límite de  $a \rightarrow 0$  en un sistema de péndulos idénticos acoplados como el visto en 3.2. Empezamos por cambiar la notación en el sistema de ecuaciones diferenciales, originalmente dado por (3.25). Seguimos empleando los reemplazos (3.31) y también usamos los desarrollos de Taylor (3.32) para prepararnos para el límite  $a \rightarrow 0$ . Al reemplazar en (3.25), se siguen cancelando los términos con derivadas impares y reforzando los términos con derivadas pares y luego de tener en cuenta las consideraciones sobre las propiedades intensivas de una longitud  $L$  de sistema de péndulos acoplados, obtenemos la siguiente ecuación en derivadas parciales

$$\frac{\partial^2 \psi}{\partial t^2} = -\omega_0^2 \psi + v^2 \frac{\partial^2 \psi}{\partial x^2}, \quad (3.41)$$

con  $v^2$  dado por (3.40). Esta ecuación se parece mucho a la ecuación de ondas clásica, excepto que ahora difiere en un nuevo término que contiene a la función  $\psi$  multiplicada por  $\omega_0^2$  y se conoce como la **ecuación de ondas de Klein-Gordon** unidimensional (se reduce a la ecuación de ondas clásicas (3.39) en el caso  $\omega_0 = 0$ ).

## 3.4. La ecuación de ondas

Como se discutió en el contexto de las ecs. (3.16) y (3.17), la ec. de dispersión (3.13) para el sistema discreto se puede reescribir como una relación entre  $\omega$  y la frecuencia espacial  $k = \Theta/a$ . En el límite continuo ( $a \rightarrow 0$ ) el valor del seno en (3.13) se puede aproximar por su argumento

$$\omega = 2\sqrt{\frac{T_0}{m a}} \left| \sin \frac{k a}{2} \right| \approx \sqrt{\frac{T_0 a}{m}} k, \quad (3.42)$$

y la ec. de dispersión  $\omega(k)$  para una cuerda homogénea es una función lineal, de pendiente

$$v = \sqrt{\frac{T_0}{\rho}}. \quad (3.43)$$

Esta pendiente tiene unidades de velocidad y se llama **velocidad de fase**. En el caso de la cuerda homogénea la velocidad de fase depende solamente de  $T_0$  y de  $\rho$ , pero no de la frecuencia, se dice que es un **caso no dispersivo**. En cambio, cuando la velocidad de fase depende de la frecuencia, se dice que es un **caso dispersivo**.

Ecuaciones de ondas como la (3.35), de la forma

$$\frac{\partial^2 \psi}{\partial t^2} = v^2 \frac{\partial^2 \psi}{\partial x^2}, \quad (3.44)$$

no solamente rigen la evolución del apartamiento del equilibrio de una soga elástica homogénea, sino también la perturbación de magnitudes físicas en fenomenologías muy distintas a la de la soga, como la presión en un gas (ondas acústicas, donde  $v$  depende del tipo de gas y de la presión en el equilibrio) o el campo eléctrico en el vacío (ondas electromagnéticas, donde  $v$  es una constante universal, la velocidad de la luz en el vacío).

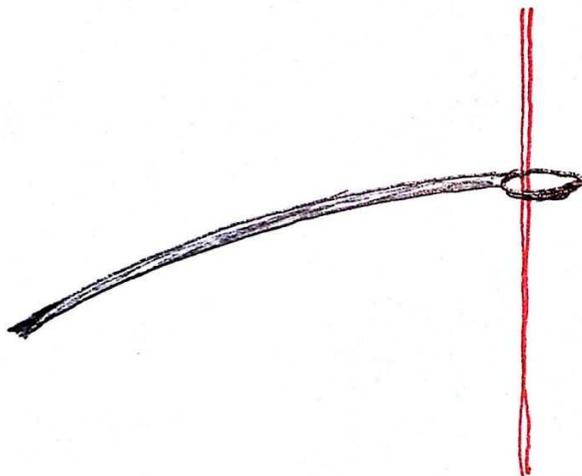


Figura 3.7: Extremo libre: la cuerda con anillo no ejerce fuerza transversal sobre la varilla  $\Rightarrow$  cuerda horizontal  $\Rightarrow$  pendiente cero.

### 3.5. Modos normales en el medio continuo

Si en vez de buscar la solución más general de la ec. (3.44) buscamos soluciones que correspondan a modos (es decir, soluciones donde las partes blablabla, blablabla, blablabla) necesariamente tenemos que restringir la búsqueda a funciones de onda de la forma

$$\psi(x, t) = A(x) \cos(\omega t + \varphi), \quad (3.45)$$

donde la función  $A(x)$  juega un papel análogo a la sucesión  $A_n$  en la versión discreta (ver 4.1.2, página 49). Los modos normales de la soga o de otros medios continuos limitados en el espacio y que responden a la ecuación de ondas clásicas se llaman *ondas estacionarias*. Para encontrar  $A(x)$ , reemplazamos (3.45) en (3.44) y queda

$$\frac{d^2 A}{dx^2} + \underbrace{\frac{\omega^2}{v^2}}_{k^2} A = 0, \quad (3.46)$$

que es como la ecuación del oscilador armónico, pero en el dominio espacial en vez del temporal. Las soluciones tienen la forma

$$A(x) = B \sin(kx) + C \cos(kx), \quad (3.47)$$

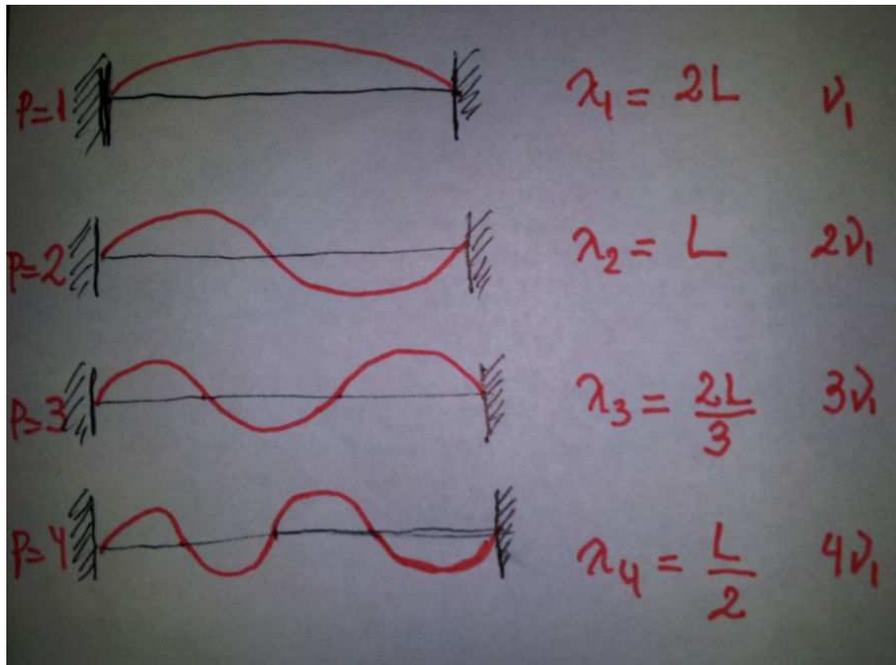


Figura 3.8: Formas de los primeros modos

donde  $B$  y  $C$  son constantes y

$$k = \frac{\omega}{v}. \quad (3.48)$$

Comparar (3.45) con la ec. (3.10) y (3.48) con (3.42). Resumen: un modo normal (u onda estacionaria) de la cuerda tiene la forma

$$\psi(x, t) = \left[ B \sin(kx) + C \cos(kx) \right] \cos(\omega t + \varphi), \quad (3.49)$$

donde determinaremos  $k$  a partir de las condiciones en los extremos ( $x = 0, L$ ) y determinaremos  $\omega(k)$  con la relación de dispersión.

### 3.6. Condiciones en los extremos

En muchos problemas ondulatorios aparecen dos condiciones de extremos que matemáticamente se manifiestan como “función igual a cero” o como “derivada igual a cero”. La condición “función igual a cero” se debe aplicar en el caso de la cuerda para un extremo fijo y es análoga a la condición usada para las amplitudes de la cadena discreta en la sección 3.1.4. Si los dos extremos de la cuerda están fijos, entonces

$$\psi(0, t) = \psi(L, t) = 0. \quad (3.50)$$

En cambio, la condición “derivada igual a cero” aparece en la cuerda homogénea en la situación de extremo libre, que podemos idealizar siguiendo los lineamientos de la sección

3.1.5, pensando que el extremo de la soga está unido a un anillo sin masa que desliza por una varilla vertical sin rozamiento (ver figura 3.9). En estas condiciones, la varilla no puede ejercer fuerza vertical sobre el sistema cuerda-anillo (porque no hay rozamiento). Entonces, el sistema cuerda-anillo no ejerce fuerza transversal sobre la varilla y la única manera para que esto suceda es que la cuerda quede horizontal, es decir, que su pendiente sea cero y así no hay proyección de la tensión en la dirección vertical. Observar que la misma condición se puede obtener a partir de la ec. (3.24) en el límite de  $a \rightarrow 0$ . Para una cuerda con dos extremos libres, se debe pedir

$$\frac{\partial\psi(0,t)}{\partial x} = \frac{\partial\psi(L,t)}{\partial x} = 0. \quad (3.51)$$

Por supuesto puede haber situaciones con un extremo fijo y el otro libre. Como en el caso discreto, las condiciones en los extremos no cambian la relación entre  $\omega$  y  $k$ , solamente seleccionan los valores de  $k$  permitidos.

## 3.7. Cuerda con extremos fijos

### 3.7.1. Discretización de los modos normales

Para una cuerda uniforme de longitud  $L$  con extremos fijos en  $x = 0$  y en  $x = L$ , se debe pedir que

$$C = 0 \quad \text{y} \quad B \sin(kL) = 0, \quad (3.52)$$

por lo tanto

$$kL = p\pi, \quad p = 1, 2, \dots, \infty \quad (3.53)$$

y se discretizan los valores de  $k$  y las frecuencias de cada modo

$$k_p = p \frac{\pi}{L}, \quad \lambda_p = \frac{2L}{p}, \quad \omega_p = p \frac{\pi}{L} \sqrt{\frac{T_0}{\rho}}, \quad p = 1, 2, \dots, \infty. \quad (3.54)$$

El movimiento más general de la cuerda continua para vibraciones transversales es una superposición de todos los modos

$$\psi(x,t) = \sum_{p=1}^{\infty} B_p \sin(k_p x) \cos(\omega_p t + \varphi_p), \quad (3.55)$$

y esperamos que esta expresión siga describiendo el comportamiento físico de un sistema discreto con muchas partes móviles siempre y cuando las condiciones iniciales exciten modos con longitudes de onda  $\lambda_p \gg a$ .

En la figura 3.8 están dibujadas las amplitudes  $A_p(x)$  (ver eq. (3.47)) con  $k_p$  y  $\omega_p = 2\pi\nu_p$  dados por las ecs. (3.54).

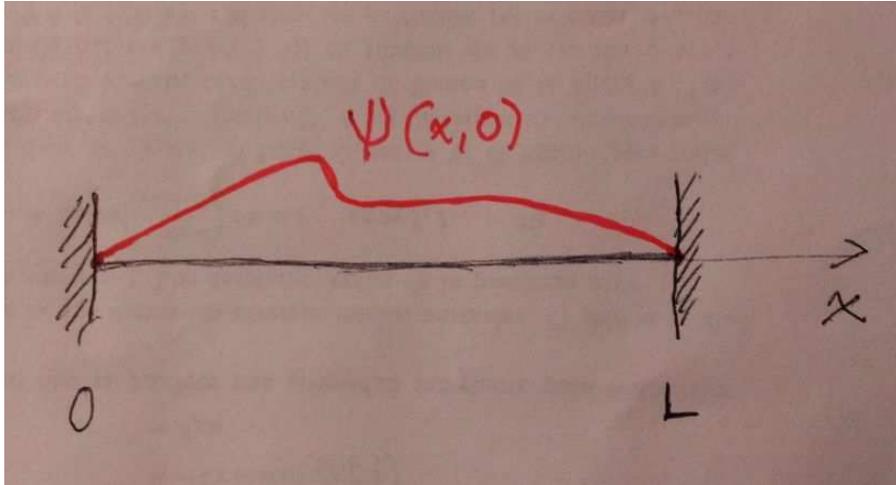


Figura 3.9: Problema de condiciones iniciales, encontrar  $B_p$  y  $\varphi_p$  en la ec. (3.54) a partir de  $\psi(x, 0)$  y de  $\frac{\partial\psi(x,0)}{\partial t} \forall x \in [0, L]$

### 3.7.2. Condiciones iniciales

Imponemos las condiciones iniciales en la expresión de la ec. (3.55)

$$\psi(x, 0) = \sum_{p=1}^{\infty} B_p \sin(k_p x) \cos \varphi_p, \quad (3.56)$$

y

$$\frac{\partial\psi(x, 0)}{\partial t} = - \sum_{p=1}^{\infty} \omega_p B_p \sin(k_p x) \sin \varphi_p. \quad (3.57)$$

Las funciones a la izquierda de las igualdades (3.56) y (3.57) son conocidas, la primera función es la foto de la forma de la sogá en el instante inicial y la otra es la velocidad inicial en cada punto de la sogá. ¿Cómo obtener la doble infinitud de amplitudes  $B_p$  y fases  $\varphi_p$ ?

Cuando los físicos queremos una solución ya mismo, pero no sabemos matemática ni tenemos tiempo de aprender, recurrimos a métodos FB (llamados así por las iniciales de Fuerza Bruta). Para nuestro problema de determinar  $B_p$  y  $\varphi_p$ , un método FB podría proceder de la siguiente manera. Primero, truncar las series y quedarse con 20 términos en cada sumatoria, 20 en (3.56) y 20 en (3.57). Se tendrían así 40 incógnitas, 20 amplitudes  $B_p$  y 20 fases  $\varphi_p$ . Y como las igualdades (3.56) y (3.57) deben cumplirse para todo  $0 < x < L$ , podría forzarse el cumplimiento de estas igualdades en tantos valores de  $x$  distintos como hiciera falta. Por ejemplo, en una grilla equiespaciada de 20 puntos  $0 < x_q < L$ , con  $q = 1, 2, \dots, 20$ . Se tendría entonces un sistema de 40 ecuaciones con 40 incógnitas. Y con suerte y con una computadora, es probable que se puedan encontrar valores de  $B_p$  y  $\varphi_p$ ,  $p = 1, 2, \dots, 20$  para que las 40 ecuaciones se cumplan lo más ajustadamente posible. Para comprobar si la propuesta FB funciona o no, el proceso se puede repetir

truncando las series con 40, en vez de 20 términos, con lo cual ahora habría 80 incógnitas, 40 amplitudes  $B_p$  y 40 fases  $\varphi_p$ , que pueden obtenerse numéricamente como antes, pero forzando el cumplimiento de las igualdades (3.56) y (3.57) en una grilla equiespaciada de 40 (en vez de 20) puntos  $0 < x_q < L$ , con  $q = 1, 2, \dots, 40$ . Si el proceso FB funciona, las 20 primeras incógnitas  $B_p$  y  $\varphi_p$  obtenidas en la segunda truncación, deberían parecerse bastante a las obtenidas en la primera truncación, además de observarse que las últimas 20 obtenidas en la segunda truncación van tomando valores menos significativos. Si esto no ocurre, se puede seguir aumentando el parámetro de truncación y ver si este proceso converge. Así tendríamos un método numérico, que habría que ver bajo que condiciones converge . . . todo esto pertenece al área de la física computacional. Afortunadamente, hay un método más fácil, el método de las series de Fourier, que permite encontrar de manera exacta las infinitas incógnitas  $B_p$  y  $\varphi_p$  que aseguran el cumplimiento de (3.56) y (3.57) en todo el intervalo  $0 < x < L$ . Veremos a continuación el método de las series de Fourier para el caso de la cuerda con extremos fijos, aunque este método se puede aplicar, con pequeñas variantes, para otras soluciones de la ecuación de onda unidimensional y para otras condiciones de extremos.

### 3.7.3. Series de Fourier

A la derecha de las igualdades (3.56) y (3.57) tenemos dos series que sabemos cuánto dan para todo valor de  $x \in [0, L]$  (sobre la soga): la serie en (3.56) tiene que dar la forma de la soga en el instante inicial y la serie en (3.57) tiene que dar la velocidad inicial de cada punto de la soga. Nuestro problema es usar esta información para obtener las amplitudes  $B_p$  y las fases  $\varphi_p$ ,  $p = 1, 2, \dots, \infty$ , porque estas cantidades son lo único que nos falta para describir completamente las vibraciones transversales de la cuerda (ver ec. (3.55)).

Algo realmente sorprendente y extrañamente antiintuitivo de la ec. (3.55) es que dice que la forma de la soga en cualquier instante de tiempo  $t_0$  (en particular también en el instante inicial, como en (3.56))

$$\psi(x, t_0) = \sum_{p=1}^{\infty} \underbrace{B_p \cos(\omega_p t_0 + \varphi_p)}_{F_p} \sin(k_p x), \quad (3.58)$$

se puede escribir como una suma infinita de funciones sinusoidales de amplitudes  $F_p$  y con frecuencias espaciales  $k_p$  que son múltiplos enteros de la frecuencia espacial fundamental  $k_1 = \pi/L$ . Puede sonar raro que una función de  $x$ , definida para  $x \in [0, L]$  y tan arbitraria como puede ser la forma de la cuerda *con extremos fijos* en un instante dado, se pueda escribir de esta manera. A Jean-Baptiste Joseph Fourier también le sonaba raro y gracias a que se salvó de ser guillotinado en la Revolución Francesa fue el primero que pudo demostrar esta propiedad en 1807. Por eso hoy decimos que en la ec. (3.58) estamos haciendo un **análisis de Fourier espacial**. Otra consecuencia sorprendente y antiintuitiva de la ec. (3.55) es que el movimiento más general de un punto arbitrario de

la cuerda (por ejemplo, el punto  $x_0 \in [0, L]$ )

$$\psi(x_0, t) = \sum_{p=1}^{\infty} \underbrace{B_p \sin(k_p x_0)}_{G_p} \cos(\omega_p t + \varphi_p), \quad (3.59)$$

siempre viene dado por una función del tiempo que se puede escribir como una suma infinita de funciones armónicas de amplitudes  $G_p$  y frecuencias  $\omega_p$  que son múltiplos enteros de una frecuencia fundamental  $\omega_1 = \pi v/L$  (ver (3.54)). Observar que todos los términos de la serie (3.55) tienen períodos que entran un número entero de veces en el período asociado con la frecuencia fundamental ( $T_p = T_1/p$ ). Esto quiere decir que la función  $\psi(x_0, t)$  es periódica, de período  $T_1$ . Y de nuevo, es raro que una función periódica de período  $T_1$ , tan arbitraria como puede ser el desplazamiento de un punto cualquiera de la cuerda, se pueda escribir como una combinación de funciones armónicas puras, con amplitudes y fases convenientes. De nuevo estamos ante un **análisis de Fourier**, pero ahora **temporal** en vez de espacial.

Por lo discutido en el párrafo anterior, queda claro que para resolver el problema de condiciones iniciales representado por las ecs. (3.56) y (3.57) hay que hacer un análisis de Fourier espacial y encontrar los coeficientes que multiplican a las funciones seno. En la ec. (3.56) estos coeficientes son directamente  $B_p \cos \varphi_p$ , mientras que en la ec. (3.57) los coeficientes son  $-\omega_p B_p \sin \varphi_p$ . Si sabemos encontrar estos coeficientes, luego podremos hallar  $B_p$  y  $\varphi_p$  tal como hicimos tantas veces (ver (2.8)), mediante el procedimiento de elevar al cuadrado y sumar miembro a miembro las ecuaciones obtenidas (para obtener  $B_p$ ) y dividir miembro a miembro (para obtener la tangente de  $\varphi_p$ ).

### Coefficientes de la serie de Fourier impar

Las ecs. (3.56) y (3.57) tienen la forma

$$F(x) = \sum_{p=1}^{\infty} F_p \sin k_p x, \quad 0 \leq x \leq L, \quad (3.60)$$

donde  $F(x)$  es dato y *satisface las condiciones de extremos fijos*, es decir,  $F(0) = F(L) = 0$ . Para encontrar  $F_p$  (los coeficientes del desarrollo en serie de Fourier de la función  $F(x)$ ), el procedimiento siempre es el mismo y consiste en multiplicar miembro a miembro la igualdad por una función y luego integrar. Si la función se elige “bien”, al integrar se logra que la serie colapse a un solo término. Las funciones que permiten obtener este efecto tan conveniente son las mismas funciones asociadas a los modos, es decir las funciones  $\sin k_p x$ ,  $p = 1, 2, \dots, \infty$ . Para ver que esto es así, calculemos las integrales

$$I_{pq} = \int_0^L \sin k_p x \sin k_q x \, dx. \quad (3.61)$$

Usando la identidad

$$2 \sin a \sin b = \cos(a - b) - \cos(a + b),$$

el integrando se puede poner como una suma de cosenos. Y como la integral de cada coseno es un seno, queda

$$I_{pq} = \frac{1}{2} \left[ \frac{\sin(k_p - k_q)x}{k_p - k_q} - \frac{\sin(k_p + k_q)x}{k_p + k_q} \right] \Big|_0^L. \quad (3.62)$$

La evaluación en  $x = 0$  da cero. Y la evaluación en  $x = L$  también, porque según (3.54),  $(k_p \pm k_q)L$  es siempre múltiplo entero de  $\pi$ . O sea, que hemos demostrado que  $I_{pq} = 0$ . ¿Seguro? ¿Y qué pasa con el denominador en (3.62) cuando  $k_p = k_q$ ? Ah, no, en ese caso resulta que  $I_{pq}$  no da cero, sino que da  $L/2$ . En definitiva, hemos demostrado que

$$I_{pq} = \int_0^L \sin k_p x \sin k_q x dx = \frac{L}{2} \delta_{pq}, \quad (3.63)$$

con  $\delta_{pq}$  la delta de Kronecker (igual a 1 si  $p = q$ , o igual a 0 cuando  $p \neq q$ ). Con este resultado encaramos el procedimiento que mencionamos después de la ec. (3.60): multiplicamos miembro a miembro la igualdad por  $\sin k_q x$  y luego integramos en el intervalo  $0 \leq x \leq L$ . En el lado derecho la integral de la suma es la suma de las integrales (excepto

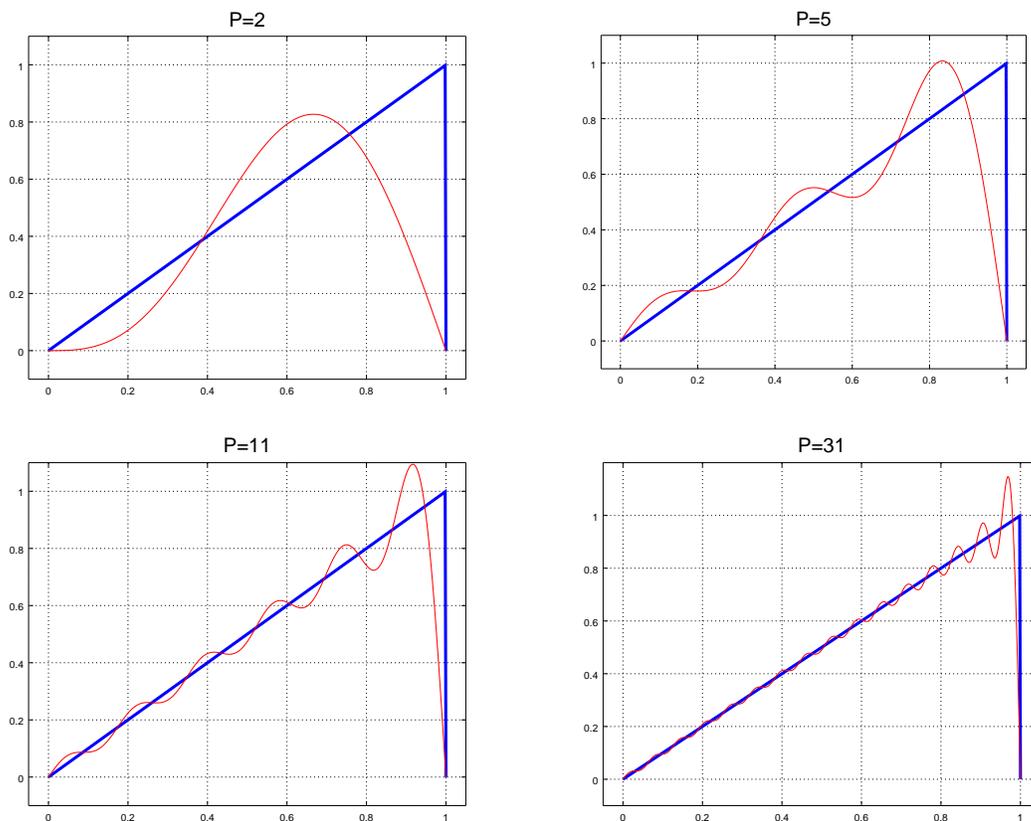


Figura 3.10: La función diente de sierra triangular definida por (3.66) (con  $r = 1$  y  $L = 1$ ) y su síntesis parcial sumando solamente los  $P$  primeros términos de su serie de Fourier (3.60)

en casos muy patológicos), y entonces queda

$$\int_0^L F(x) \sin k_q x \, dx = \sum_{p=1}^{\infty} F_p I_{pq} = \frac{L}{2} F_q, \quad (3.64)$$

es decir que

$$F_p = \frac{2}{L} \int_0^L F(x) \sin k_p x \, dx, \quad (3.65)$$

para todo valor de  $q = 1, 2, \dots, \infty$ . Y si todo converge bien, esta “receta” (3.65) nos autorizaría entonces a hacer una identificación entre la función  $F(x)$ , definida en el intervalo finito  $[0, L]$  y nula en los extremos del intervalo (el lado izquierdo de (3.60)), con la serie de funciones sinusoidales de amplitudes  $F_p$  y frecuencias espaciales múltiplos enteros de la frecuencia fundamental  $\pi/L$  (el lado derecho de (3.60)). Notar que si bien en nuestro problema físico sólo interesa el valor de  $F(x)$  en el intervalo  $0 \leq x \leq L$ , la serie a la derecha de la igualdad (3.60) puede ser evaluada fuera de este intervalo. Es claro que la serie representa una función periódica, con el período  $(2L)$  del armónico más bajo, ya que

$$\sin \left[ k_p \left( x + \frac{2\pi}{k_1} \right) \right] = \sin (k_p x + 2\pi p) = \sin k_p x.$$

¿Qué sucede en el intervalo  $[-L, 0]$  (y en todos los intervalos equivalentes, separados del intervalo  $[-L, 0]$  por múltiplos de  $2L$ ), donde la información sobre  $F(x)$  es irrelevante para el problema físico que estamos tratando? Debido a que los senos son funciones impares,  $F(-x) = -F(x)$  y entonces la serie representa una función impar. Y revisando los pasos anteriores, es fácil convencerse que bajo las mismas condiciones de convergencia, toda función periódica impar admite un desarrollo en serie de Fourier dado por (3.64) con amplitudes dadas por (3.65).

### Ejemplo: diente de sierra

Una manera infalible de poner a prueba nuestro entendimiento del significado de la igualdad (3.60), es llevarlo a la práctica numérica, por ejemplo comparando una  $F(x)$  muy sencilla, definida en el intervalo  $[0, L]$ , con distintas aproximaciones obtenidas sumando solamente un número finito de términos de la serie de Fourier que corresponde a esta misma función prolongada periódicamente por imparidad al intervalo  $[-L, 0]$ . Para experimentar numéricamente usemos la función en forma de triángulo que se muestra en color azul en la figura 3.10 y definida por

$$F(x) = \begin{cases} r x & \text{si } 0 \leq x < L, \\ 0 & \text{cuando } x = L. \end{cases} \quad (3.66)$$

Esta función tiene la ventaja de que las integrales (3.65) se obtienen analíticamente de forma muy sencilla, aunque no sería la más adecuada para representar una condición

inicial realista (recordemos que fueron las condiciones iniciales las que nos trajeron por este camino). Con cuentas muy fáciles se demuestra que

$$F_q = (-1)^{q+1} \frac{2rL}{q\pi}, \quad q = 1, 2, \dots, \infty. \quad (3.67)$$

En los gráficos de la figura 3.10 se muestran las síntesis parciales obtenidas en el intervalo  $[0, L]$ , para el caso  $r = 1$ ,  $L = 1$  y distintos truncamientos de la serie. Observar que el número de crestas y valles en las líneas rojas da directamente el número de términos empleados y que la amplitud de crestas y valles disminuye a medida que la suma de senos va aproximando mejor a la función. En la figura 3.13 se muestran más detalles sobre la convergencia del cálculo numérico y el comportamiento de la serie fuera del intervalo  $[0, L]$ . ¿No es asombroso que una función radicalmente tan diferente de una senoide se pueda sintetizar usando solamente funciones sinusoidales?

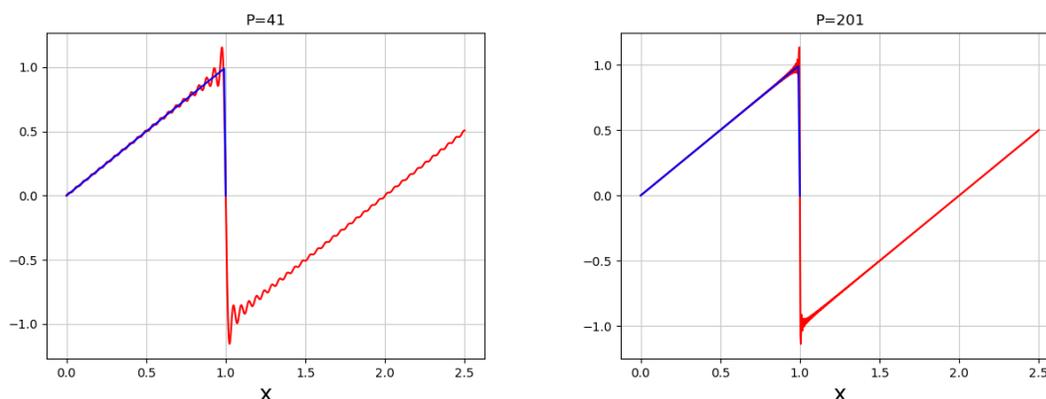


Figura 3.11: La prolongación periódica impar de la función diente de sierra triangular definida por (3.66) (con  $r = 1$  y  $L = 1$ ), sumas parciales con 41 y 201 primeros términos en (3.60).

Como las herramientas computacionales son ineludibles en la formación de todo científico, podemos experimentar numéricamente el ejemplo del diente de sierra con este sencillo script de Python, usado para obtener las figuras 3.10 y 3.13:

```
import numpy as np
import matplotlib.pyplot as plt
r=1.0 #pendiente de la recta
L=1 #longitud de la cuerda
veces=2.5 #intervalo a graficar [0,veces*L]
Pmax=32 #suma Pmax-1 senos
x=np.linspace(0.0,veces*L,1000) #grilla para serie
xsoga=np.linspace(0.0,L,100) #grilla para soga
k=np.pi/L #frecuencia espacial fundamental

Faprox=0 #inicializo aproximacion de F
F=r*xsoga #diente de sierra evaluado en xsoga
```

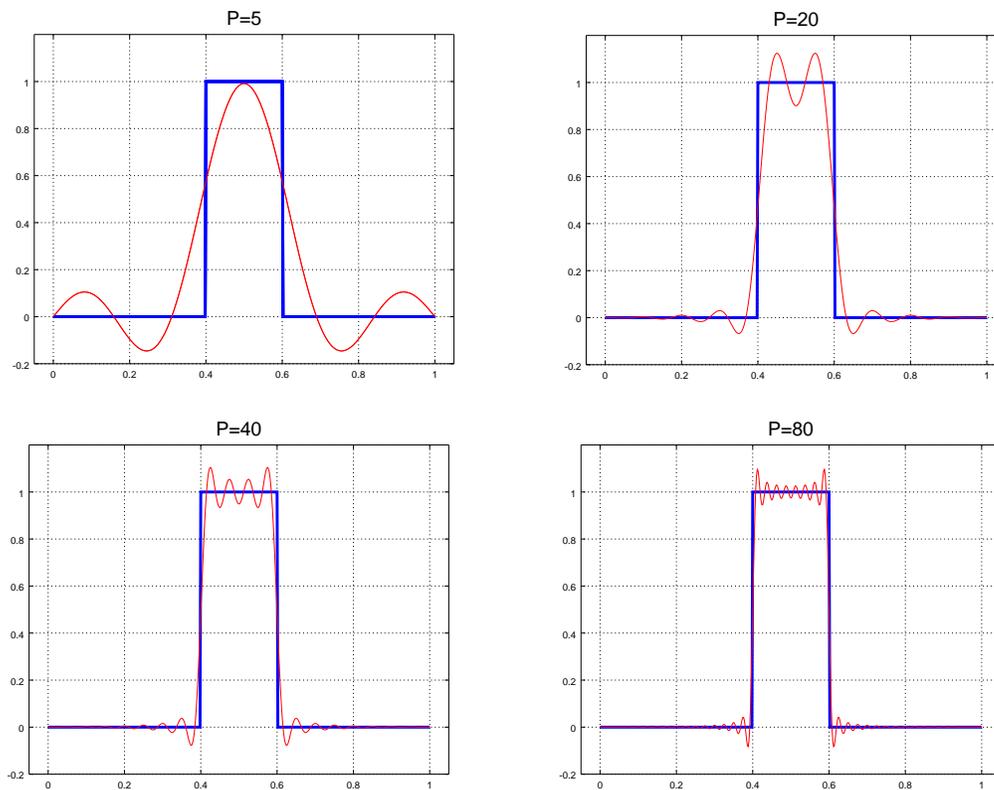


Figura 3.12: El pulso rectangular (3.68) (azul) y su síntesis parcial con  $P$  armónicos (rojo).

```

F[99]=0      #fuerzo 0 en xsoga=L
for p in range(1,Pmax):      #ciclo suma serie truncada
    Fp=(-1)**(p+1)*(2*r*L)/(p*np.pi)
    Faprox=Faprox+Fp*np.sin(p*k*x)

plt.plot(x,Faprox,c='r') #grafico F aproximada
plt.plot(xsoga,F,c='b') #grafico diente de sierra
plt.title('P=%i' %(Pmax-1)) #pongo titulo
plt.xlabel('x', fontsize=18) #etiqueto eje x
plt.grid(b=True, which='major', color='xkcd:silver', linestyle
        ='-')
#pepe=directorio de trabajo donde esta el script
plt.savefig(r'd:\pepe\fourier1.png', format='png')

```

Ejecuten el script, modifíquelo, jueguen con los datos, visualicen comportamientos físicos, en síntesis: ¡experimenten numéricamente! Un buen plan es reproducir primero los resultados de las figuras 3.10 y 3.13 y luego, modificando los scripts, probar con otras pendientes, otros truncamientos u otras funciones (cambiar la definición de  $F$ , por ejemplo una función  $F$  con dos tramos rectos).

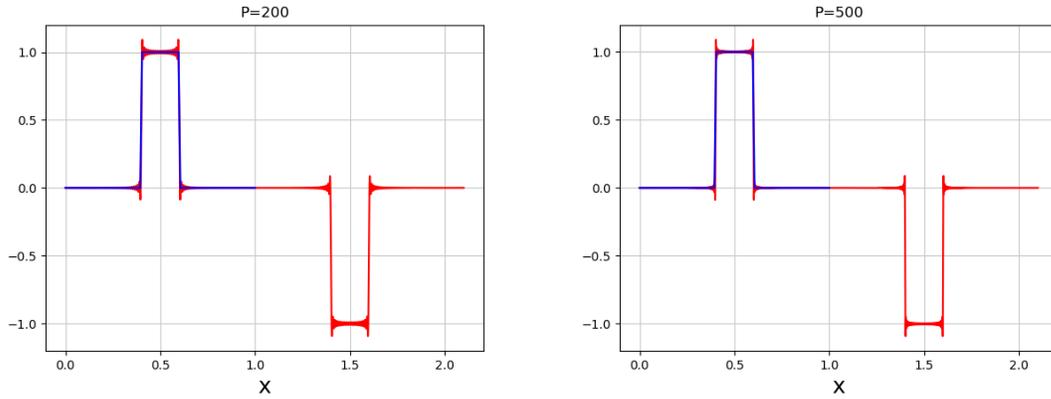


Figura 3.13: La prolongación periódica impar del pulso rectangular definido por (3.68) aproximada por las sumas parciales con 200 y 500 primeros términos en (3.60), parámetros  $a = 1$ ,  $x_0 = 0.4$ ,  $c = 0.2$ ,  $L = 1$ .

### Ejemplo: pulso rectangular

Como nuevo ejemplo, probemos con un pulso rectangular de ancho  $c$  y altura  $a$ , localizado entre  $x_0$  y  $x_0 + c$  y definida por

$$F(x) = \begin{cases} a & x \in [x_0, x_0 + c], \\ 0 & x \notin [x_0, x_0 + c]. \end{cases} \quad (3.68)$$

Ahora las amplitudes de los armónicos son

$$F_q = \frac{2a}{q\pi} \left[ \cos k_q x_0 - \cos k_q (x_0 + c) \right], \quad q = 1, 2, \dots, \infty. \quad (3.69)$$

En la figura 3.12 se muestran síntesis parciales obtenidas con distintos truncamientos en el caso  $a = 1$ ,  $x_0 = 0.4$ ,  $c = 0.2$  y  $L = 1$ . El script de Python empleado para obtener las figuras (3.10) y (3.13) <sup>4</sup> se obtuvo haciendo pequeñas modificaciones al script de la página 72, de manera de incluir la nueva función, (3.68), y los nuevos coeficientes, (3.69).

### Condiciones iniciales $\implies B_p$ y $\varphi_p$

Ahora ya estamos en condiciones de resolver completamente el problema de condiciones iniciales representado por las ecs. (3.56) y (3.57). Para eso recordemos que las dos series tienen la forma de la ec. (3.60) y que entonces sus amplitudes de Fourier  $F_p$  están determinadas por la expresión (3.65). De esta manera, resulta que

$$\begin{aligned} B_p \cos \varphi_p &= \frac{2}{L} \int_0^L \psi(x, 0) \sin k_p x \, dx, \\ -\omega_p B_p \sin \varphi_p &= \frac{2}{L} \int_0^L \frac{\partial \psi(x, 0)}{\partial t} \sin k_p x \, dx. \end{aligned} \quad (3.70)$$

<sup>4</sup>hay una copia en <http://bit.ly/Fourier0102>

Así obtenemos  $B_p$

$$B_p = \frac{2}{L} \left\{ \left[ \int_0^L \psi(x, 0) \sin k_p x \, dx \right]^2 + \frac{1}{\omega_p^2} \left[ \int_0^L \frac{\partial \psi(x, 0)}{\partial t} \sin k_p x \, dx \right]^2 \right\}^{1/2}, \quad (3.71)$$

y  $\tan \varphi_p$

$$\tan \varphi_p = -\frac{1}{\omega_p} \frac{\int_0^L \frac{\partial \psi(x, 0)}{\partial t} \sin k_p x \, dx}{\int_0^L \psi(x, 0) \sin k_p x \, dx}, \quad (3.72)$$

y como las ecs. (3.56) y (3.57) dan los signos de los senos y cosenos, las fases  $\varphi_p$  queda bien determinadas en el intervalo  $[0, 2\pi]$ .

Notar que:

- Si la velocidad inicial de todos los puntos de la sogá es cero,  $\varphi_p = 0 \, \forall p$ ;
- Si la velocidad inicial de todos los puntos de la sogá es cero y si la forma inicial tiene la forma de uno de los modos (ver figura 3.8), la sogá permanece oscilando en dicho modo;
- Si la velocidad inicial de todos los puntos de la sogá es cero y si la forma inicial tiene un **nodo** (un punto de desplazamiento nulo) que coincide con el nodo de un modo, la sogá permanece oscilando en una superposición de modos que también tengan un nodo en el mismo lugar. Ejemplo de esta situación son los “sonidos armónicos” de la guitarra y otros instrumentos de cuerda, que se obtienen cuando el ejecutante fuerza un nodo en la deformación inicial
- Análogamente, cuando la forma inicial tiene desplazamiento no nulo en el lugar que corresponde al nodo de un armónico, el efecto resultante es suprimir dicho armónico. Por este motivo, los pianos acústicos se diseñan para que los martillos que golpean las cuerdas peguen aproximadamente a una distancia que sea la séptima parte de la longitud de la cuerda. De esta manera, logran atenuar muchísimo la amplitud del séptimo armónico, cuyo sonido no es “agradable” (No entramos a discutir qué se entiende por agradable. Nos pueden decir que significa “disonante”. Pero tampoco vamos a entrar a discutir qué se entiende por disonante).

### 3.7.4. Evolución temporal

Una vez hallados  $B_p$  y  $\varphi_p$  ( $p = 1, 2, \dots, \infty$ ), el movimiento de la sogá queda determinado para todo instante. Por ejemplo, en la figura 3.14 se muestran imágenes de los primeros instantes de la evolución de la deformación para una sogá con  $L = 1$  m y condición inicial en forma de pulso rectangular, como en (3.68), con  $a = 1$  cm y con  $x_0 = 0.45$  m y  $c = 0.1$  m (pulso centrado en la mitad de la sogá). La densidad y la tensión se eligen para que la velocidad de fase sea  $v = 1$  cm/s y la velocidad inicial de todos los puntos de la sogá es nula. En este caso todos los  $\varphi_p$  son nulos y los  $B_p$  están dados por (3.69). Para seguir la deformación de la sogá tenemos

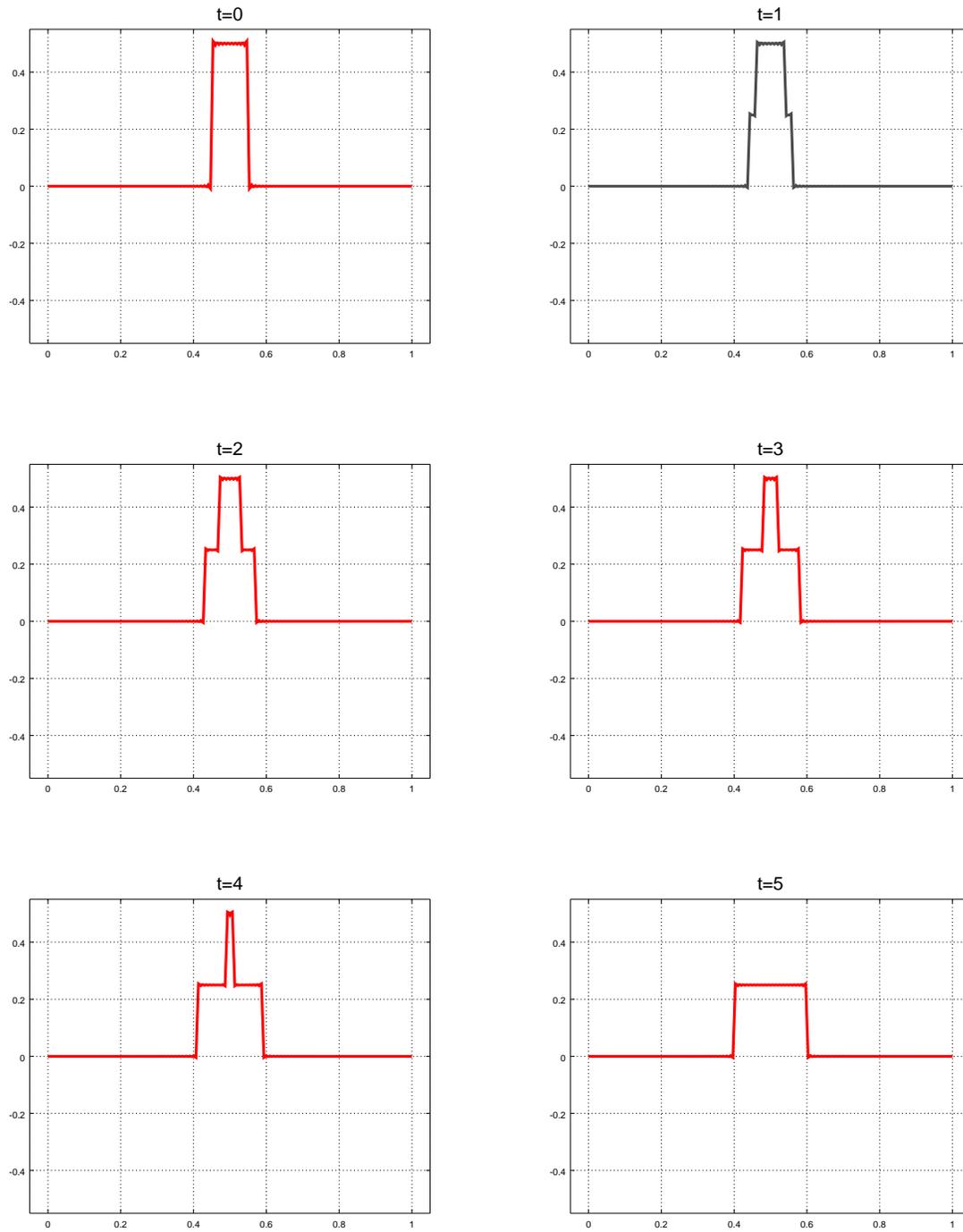


Figura 3.14:  $\psi(x,t)$  para  $t = 0s, 1s, 2s, 3s, 4s$  y  $5s$ . La deformación inicial tiene la forma de pulso rectangular indicada en (3.68), con  $x_0 = 0.45$  m,  $c = 0.1$  m,  $a = 1$  cm y  $v = 1$  cm/s. La velocidad inicial de todos los puntos de la soga es nula. Para evaluar las series se usó  $N_s = 1000$ .

que usar (3.55), truncando la serie en un número finito de términos  $N_s$ . Para evaluar esta serie en  $t \neq 0$  modificamos el script usado para evaluar la serie en  $t = 0$ .

En la figura 3.14 se observa que la deformación, localizada inicialmente alrededor de  $x_0 = 0.5$  m, se empieza a “desparramar” y se propaga hacia las zonas vecinas. En  $t = 5$  s la deformación se “ensanchó”, pero se hizo menos intensa en cada punto (bajó su amplitud).

Para  $5 s < t < 50 s$ , ver figura 3.15, la deformación se sigue extendiendo a las zonas vecinas, en la forma de dos paquetes que parecen muy independientes. Si alguien nos preguntara la razón de este comportamiento, podríamos decir que porque las cuentas dan así. ¡Pero seguramente no nos quedaríamos conformes! Sería atrayente poder explicar cosas sin tener que seguir sumando la serie en cada instante. ¿Podríamos predecir la distancia entre los centros de los dos paquetes para cada instante? Estamos en un punto crítico, porque con estos ejemplos queda claro que para

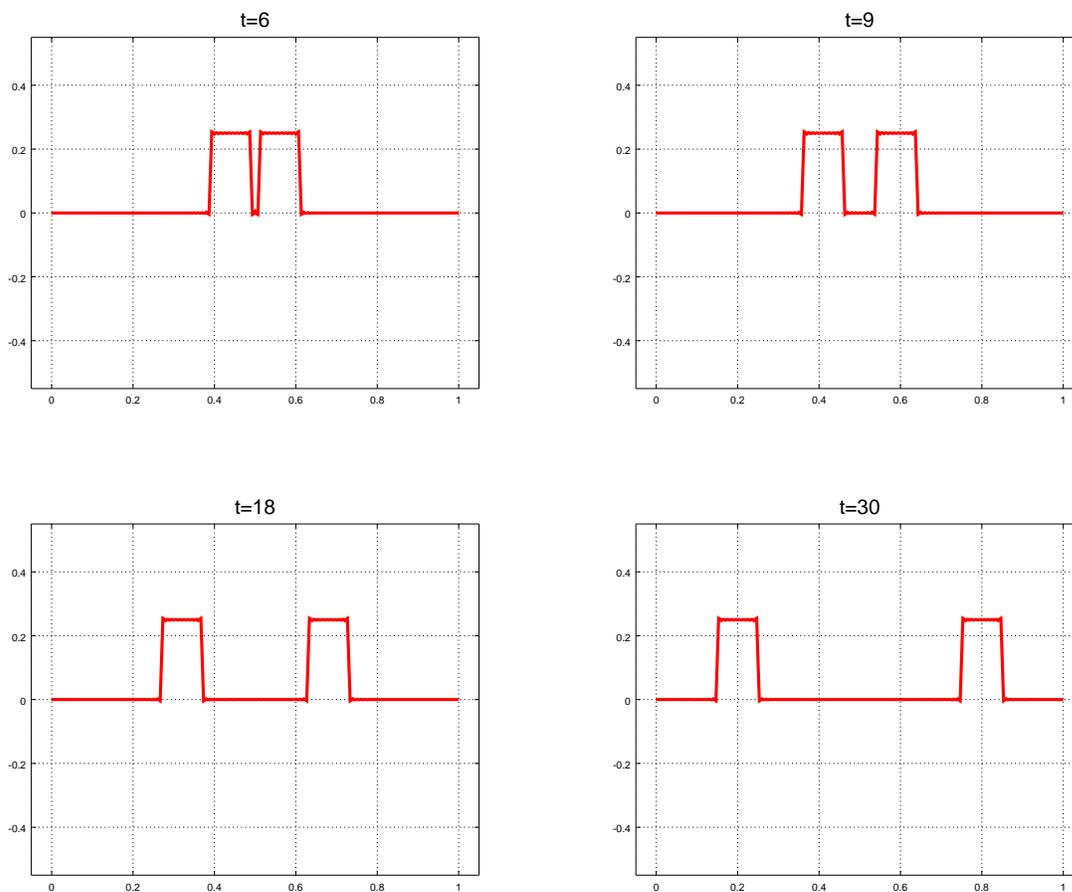


Figura 3.15:  $\psi(x, t)$  para  $5s < t < 50s$ . La deformación inicial tiene la forma de onda cuadrada indicada en la eq. (3.68), con  $x_0 = 0.45$  m,  $c = 0.1$  m,  $a = 1$  cm y  $v = 1$  cm/s. La velocidad inicial de todos los puntos de la soga es nula.

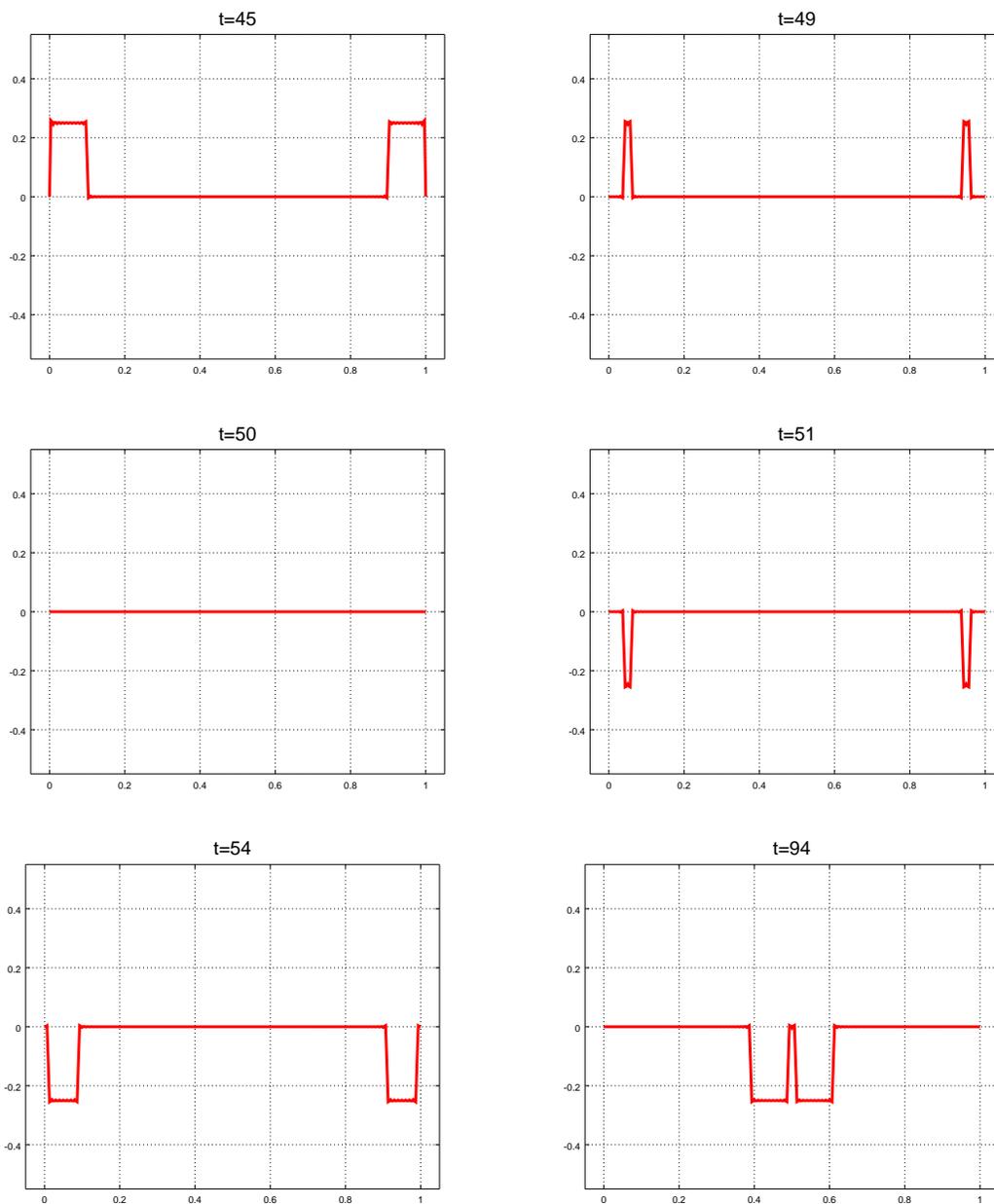


Figura 3.16:  $\psi(x, t)$  para  $45s < t < 94s$ . La deformación inicial tiene la forma de onda cuadrada indicada en la eq. (3.68), con  $x_0 = 0.45$  m,  $c = 0.1$  m,  $a = 1$  cm y  $v = 1$  cm/s. La velocidad inicial de todos los puntos de la soga es nula.

entender los comportamientos ilustrados en estas figuras, la descripción que venimos haciendo con los modos normales no parece muy adecuada. Para hablar del rasgueo de la guitarra parecía que los modos normales estaban muy bien, pero acá vemos que no explica claramente lo que sucede con estos paquetitos. Si bien las cuentas dan, nos gustaría descubrir los mecanismos físicos que nos ayuden a comprender mejor lo que estamos viendo.

Llegados a este punto, conviene reflexionar y darse cuenta que estamos cumpliendo el plan que nos habíamos propuesto inicialmente: partiendo de sistemas mecánicos con pocas partes móviles,

fuimos agregando partes móviles y en el límite continuo llegamos a una ecuación llamada “de ondas”. Y aunque todavía faltan cosas, el ejemplo que acabamos de ver muestra claramente que estamos en presencia de perturbaciones que se propagan, es decir, estamos poniendo en evidencia los comportamientos ondulatorios. La evolución de la deformación en este ejemplo se puede ver en un gif animado (fácil de hacer con Python) en este enlace <http://bit.ly/1LOZxY0>.

¿Qué sentido físico tienen en esta simulación tiempos “especiales”, como  $t = 5$  s, donde la altura del pulso disminuyó a la mitad y su ancho se duplicó respecto al pulso original? ¿O  $t = 50$  s (ver figura 3.16), donde la deformación de la sogá es nula en todo punto? ¿Estamos en condiciones de encontrar otros tiempos críticos y de relacionarlos con mecanismos sencillos, que se puedan explicar sin tener que sumar una serie? A continuación veremos que estos tiempos “especiales” están relacionados con mecanismos ondulatorios escondidos detrás de los modos normales y que estos mecanismos se pueden poner en evidencia sin necesidad de sumar las series, no solamente en el ejemplo que estamos viendo de sogá con extremos fijos, sino también en cualquier otra sistema unidimensional gobernado por la ecuación de ondas clásicas (3.35).

### 3.8. ondas estacionarias = suma de ondas viajeras

Según (3.49) (página 65), la onda estacionaria de la cuerda elástica homogénea, o de cualquier otro sistema continuo gobernado por (3.35), tiene la forma

$$\psi(x, t) = \underbrace{\left[ B \sin(kx) + C \cos(kx) \right]}_{D \sin(kx + \alpha)} \cos(\omega t + \varphi), \quad (3.73)$$

con  $k = \omega/v$ . Este tipo de solución puede representar un modo normal como en la ec. (3.49), en cuyo caso  $\omega$  queda determinada por las condiciones en los extremos. O también puede representar la respuesta forzada estacionaria de un sistema impulsado armónicamente desde alguno de sus extremos, en cuyo caso  $\omega$  es impuesta por la fuerza impulsora.

Usando identidades trigonométricas, el producto de un seno por un coseno puede transformarse en una suma de dos senos y así la onda estacionaria puede reescribirse como

$$\psi(x, t) = \frac{D}{2} \left[ \underbrace{\sin(kx - \omega t + \alpha - \varphi)}_{\text{fase1}} + \underbrace{\sin(kx + \omega t + \alpha + \varphi)}_{\text{fase2}} \right]. \quad (3.74)$$

Prestemos atención al primer seno, el que depende de la variable fase1: vale cero cuando fase1 es cero, vale 0.5 cuando fase1 es  $\pi/6$ , vale uno cuando fase1 es  $\pi/2$ , etc. Supogamos que estoy obsesionado en moverme con esta senoide siempre al lado del valor 0.5. Para asegurarme de estar en el lugar  $x$  y en el momento  $t$  adecuados lo único que tengo que pedir es que el argumento fase1 valga siempre  $\pi/6$ . Porque si me quedo parado en donde estoy voy a perder el valor 0.5, porque el argumento fase1 cambia a medida que pasa el tiempo. Dicho de otra manera, si en el instante  $t + \Delta t$  quiero seguir al lado del valor 0.5 del seno, entonces tendría que moverme a una posición  $x + \Delta x$  que me asegure de que fase1 continúa valiendo  $\pi/6$ . Es decir, que los valores  $\Delta x$  y  $\Delta t$  no son arbitrarios: tienen que preservar el valor de fase1 que me daba 0.5 para el seno, o sea

$$kx - \omega t + \alpha - \varphi = k(x + \Delta x) - \omega(t + \Delta t) + \alpha - \varphi = \pi/6,$$

$$\begin{aligned} \implies k \Delta x - \omega \Delta t &= 0, \\ \implies \frac{\Delta x}{\Delta t} &= \frac{\omega}{k} = v. \end{aligned} \quad (3.75)$$

En resumen: para seguir al valor 0.5 (o a cualquier otro valor del primer seno en el que estemos interesados), tenemos que mantener fase1 constante y para eso hay que moverse hacia mayores valores de  $x$  con una velocidad igual a la velocidad  $v$  que aparece en la ecuación de ondas. Ahora queda claro el nombre de **velocidad de fase** para  $v$  y el nombre **onda progresiva** para el primer seno en la ec. (3.74) y para cualquier otra función que dependa de la combinación  $x - vt$ . Un análisis similar para el segundo seno en (3.74) muestra que para seguir un valor determinado de la función que depende del argumento fase2, es necesario desplazarse hacia menores valores de  $x$  a medida que pasa el tiempo, es decir hay que desplazarse con velocidad  $-v$ . Por eso el nombre de **onda regresiva** para el segundo seno en la ec. (3.74) y para cualquier otra función que dependa de la combinación  $x + vt$ .

Si estamos interesados en cómo evoluciona la deformación inicial del sistema, usamos la expresión general, con modos de la forma (3.73), análoga a (3.55), pero con  $C \neq 0$ . En términos de la representación progresiva-regresiva (3.74), la evolución se escribe

$$\psi(x, t) = \frac{1}{2} \sum_{p=1}^{\infty} D_p \sin(k_p x - \omega_p t + \alpha_p - \varphi_p) + \frac{1}{2} \sum_{p=1}^{\infty} D_p \sin(k_p x + \omega_p t + \alpha_p + \varphi_p). \quad (3.76)$$

Apliquemos este resultado al ejemplo de las figuras 3.14, 3.15 y 3.16. Como en este caso la velocidad inicial de todos los puntos de la sogá es nula,  $\varphi_p = 0$  para todo valor de  $p$  y recordando que  $\omega_p = vk_p$ , resulta

$$\psi(x, t) = \frac{1}{2} \sum_{p=1}^{\infty} D_p \sin[k_p(x - vt) + \alpha_p] + \frac{1}{2} \sum_{p=1}^{\infty} D_p \sin[k_p(x + vt) + \alpha_p], \quad (3.77)$$

para todo  $0 \leq x \leq L$ . En esta representación progresiva-regresiva, la condición inicial se expresa como suma de dos series idénticas

$$\psi(x, 0) = \frac{1}{2} \sum_{p=1}^{\infty} D_p \sin[k_p x + \alpha_p] + \frac{1}{2} \sum_{p=1}^{\infty} D_p \sin[k_p x + \alpha_p], \quad (3.78)$$

de donde deducimos el resultado, bastante trivial, de que cada serie en (3.78) representa a la función  $\psi(x, 0)/2$ .

La cosa empieza a parecer menos trivial cuando usamos esta información para reinterpretar la representación (3.77) de  $\psi(x, t)$  y observamos que la primera sumatoria no es otra cosa que la deformación inicial dividida por el factor 2, pero evaluada en  $x - vt$ , es decir, desplazada  $vt$  hacia la derecha. Y que análogamente, la segunda sumatoria en (3.77) no es otra cosa que la deformación inicial dividida por el factor 2, pero ahora evaluada en  $x + vt$ , es decir, desplazada  $vt$  hacia la izquierda.

Esta separación en ondas progresivas y regresivas nos brinda otra manera de interpretar los resultados que hasta ahora obteníamos sumando ondas estacionarias, es decir, modos. Por ejemplo, consideremos el instante  $t = 5$  s, donde se ha duplicado el ancho del pulso (ver figura 3.14). Según (3.77), la deformación en este instante se puede interpretar como resultado de un

desplazamiento  $v t = 5$  cm, del pulso progresivo hacia la derecha y del pulso regresivo hacia la izquierda. Si en  $t=0$  ambos pulsos ocupaban la misma región  $0.45 \text{ cm} \leq x \leq 0.55 \text{ cm}$ , entonces en  $t = 5$  s el pulso regresivo ocupa la región  $0.4 \text{ cm} \leq x \leq 0.5 \text{ cm}$ , mientras que el pulso progresivo ocupa la región  $0.5 \text{ cm} \leq x \leq 0.6 \text{ cm}$ . Vemos así que la suma de los dos términos en (3.77) ocupan una región de longitud doble que la región ocupada inicialmente por la deformación y que la altura de la deformación resultante es la mitad de la altura de la deformación inicial.

La separación progresiva-regresiva de (3.76) nos provee de una explicación alternativa para los resultados de la figura 3.14, que hasta ahora solamente podíamos explicar “porque las cuentas dan así”. Veremos más adelante (4.1.3) que la separación es válida no solamente para el caso considerado en el ejemplo, de velocidad inicial nula, sino para cualquier solución de la ecuación de ondas clásica 1D.

### 3.9. Y en otros sistemas ...

Si la soga tiene otras condiciones en los extremos, pasa algo similar a lo que hemos visto, cambia la forma de los modos, pero los modos siguen siendo “ortogonales” entre sí, en el sentido de que sus integrales son deltas de Kronecker, tal como ocurre para sogas con extremos fijos, ver ec. (3.63). Esta característica se repite para otros sistemas que pueden ser muy distintos a la soga elástica, pero que también se extienden en una dirección fija del espacio, están limitados espacialmente (empiezan y terminan) y responden a la ecuación de ondas clásica unidimensional.

En la práctica se verán ejemplos no solamente de sogas con otras condiciones de contorno, sino también de otros sistemas muy distintos, como tubos con gases, que exhiben comportamientos (y admiten tratamientos) completamente análogos a los que vimos para la soga con extremos fijos. Es importante observar que en todos estos sistemas resulta que la forma de los modos permitidos solo depende de la forma de las ecuaciones de movimiento (que terminan siendo siempre la misma, la ecuación de ondas) y además los modos adoptan formas particularmente sencillas, por ejemplo, siempre aparece una relación supersencilla entre la longitud de onda del modo y la longitud total de sistema y esta relación no tiene nada que ver con magnitudes como elasticidad, inercia, densidad, temperatura, presión, etc.



# Capítulo 4

## Ondas propagantes

### 4.1. Ecuación clásica 1D

#### 4.1.1. Progresivas y regresivas 1D

Según las ecs. (3.73) y (3.74), toda onda estacionaria armónica es superposición de dos ondas viajeras, una progresiva y otra regresiva. Veremos a continuación una propiedad de alguna manera inversa, aunque bastante más general porque no está restringida a dependencias armónicas: si una magnitud física  $\psi(x, t)$  depende solamente de la combinación  $\eta_1 = x - vt$ , entonces  $\psi(x, t) = f(\eta_1)$  es solución de la ecuación de ondas (3.44). Análogamente, si la magnitud física  $\psi(x, t)$  depende solamente de la combinación  $\eta_2 = x + vt$ , entonces  $\psi(x, t) = g(\eta_2)$  también es solución de la ecuación de ondas (3.44). Dicho de otra manera, veremos que toda función de una sola variable, dos veces derivable, puede representar una posible onda clásica progresiva, cuando su argumento se evalúa en  $x - vt$ , o regresiva, cuando su argumento se evalúa en  $x + vt$  (el carácter progresivo o regresivo está referido a una dirección fija en el espacio, que sin pérdida de generalidad podemos considerar como eje  $x$ ).

Supongamos primero que  $f(\eta_1)$  es una función de la única variable  $\eta_1$ . Para fijar ideas, asociamos  $f(\eta_1)$  con su representación gráfica,  $f$  en el eje vertical versus  $\eta_1$  en el eje horizontal. Sea  $\eta_1 = x - vt$ : es claro que  $f(x - vt)$  representa una función progresiva, porque el gráfico de  $f(x - vt)$  vs  $x$  para  $t > 0$  está desplazado una distancia  $vt$  en el sentido positivo del eje  $x$  con respecto al gráfico de  $f(x - vt)$  vs  $x$  a  $t = 0$ . Notemos que esto es así siempre, independientemente de la forma que tenga  $f(\eta_1)$ , no importa si se trata de un pulso rectangular, una senoide o una función gaussiana. Y con tanta generalidad, no parecería haber argumentos físicos para asegurar que la función  $\psi(x, t)$  elegida idénticamente igual a  $f(x - vt)$  sea obligatoriamente solución de la ecuación de ondas clásica 1D (3.44). Sin embargo, lo es, como se puede ver fácilmente teniendo en cuenta que con esta elección

$$\frac{\partial^2 \psi(x, t)}{\partial x^2} = f'' \quad \text{y} \quad \frac{\partial^2 \psi(x, t)}{\partial t^2} = v^2 f'', \quad (4.1)$$

y entonces se cumple

$$\frac{\partial^2 \psi}{\partial t^2} = v^2 \frac{\partial^2 \psi}{\partial x^2}.$$

De manera completamente análoga podemos mostrar que la función  $\psi(x, t)$  elegida idénticamente igual a  $g(x + vt)$ , con  $g$  función de una única variable ( $\eta_2$ ), también satisface automáticamente la ecuación de ondas clásica 1D (3.44), puesto que

$$\frac{\partial^2 \psi(x, t)}{\partial x^2} = g'' \quad \text{y} \quad \frac{\partial^2 \psi(x, t)}{\partial t^2} = v^2 g''. \quad (4.2)$$

Ahora,  $\psi(x, t) = g(x + vt)$  representa una función regresiva, porque el gráfico de  $g(x + vt)$  vs  $x$  para  $t > 0$  está desplazado una distancia  $vt$  en el sentido negativo del eje  $x$  con respecto al gráfico de  $g(x + vt)$  vs  $x$  a  $t = 0$ .

### 4.1.2. La solución más general 1D

Hemos visto que las soluciones progresivas y regresivas son siempre solución de la ecuación de ondas clásica 1D. Y como esta ecuación es lineal, resulta que cualquier suma de funciones progresivas es también solución. Ahora vamos por más: veremos a continuación que **la solución más general** de la ec. (3.44) es **siempre** combinación lineal de una solución progresiva y de una regresiva. Con este fin, pasamos a las variables  $\eta_1$  y  $\eta_2$ , respectivamente asociadas a comportamientos progresivos y regresivos

$$\begin{aligned} \eta_1 &= x - vt \\ \eta_2 &= x + vt \end{aligned} \quad \iff \quad \begin{aligned} x &= (\eta_1 + \eta_2)/2 \\ t &= (\eta_2 - \eta_1)/(2v). \end{aligned} \quad (4.3)$$

Para ver cómo se escribe (3.44) en las nuevas variables, usamos la regla de la cadena una vez para transformar las derivadas primeras

$$\frac{\partial \psi(x, t)}{\partial x} = \frac{\partial \psi(x, t)}{\partial \eta_1} \frac{\partial \eta_1}{\partial x} + \frac{\partial \psi(x, t)}{\partial \eta_2} \frac{\partial \eta_2}{\partial x}, \quad (4.4)$$

$$\frac{\partial \psi(x, t)}{\partial t} = \frac{\partial \psi(x, t)}{\partial \eta_1} \frac{\partial \eta_1}{\partial t} + \frac{\partial \psi(x, t)}{\partial \eta_2} \frac{\partial \eta_2}{\partial t}, \quad (4.5)$$

y otra vez más para transformar las derivadas segundas

$$\frac{\partial^2 \psi(x, t)}{\partial x^2} = \frac{\partial^2 \psi(x, t)}{\partial \eta_1^2} + \frac{\partial^2 \psi(x, t)}{\partial \eta_2^2} + 2 \frac{\partial^2 \psi(x, t)}{\partial \eta_1 \partial \eta_2}, \quad (4.6)$$

$$\frac{\partial^2 \psi(x, t)}{\partial t^2} = v^2 \left[ \frac{\partial^2 \psi(x, t)}{\partial \eta_1^2} + \frac{\partial^2 \psi(x, t)}{\partial \eta_2^2} - 2 \frac{\partial^2 \psi(x, t)}{\partial \eta_1 \partial \eta_2} \right]. \quad (4.7)$$

Vemos así que en las nuevas variables  $\eta_1$  y  $\eta_2$ , (3.44) se escribe de una manera muy sencilla

$$\frac{\partial^2 \psi}{\partial \eta_2 \partial \eta_1} = 0, \quad (4.8)$$

que dice que la primera derivada parcial con respecto a  $\eta_1$  no depende de  $\eta_2$ . Por lo tanto, la primera derivada parcial con respecto a  $\eta_1$  es solamente una función  $F$  de  $\eta_1$

$$\frac{\partial \psi}{\partial \eta_1} = F(\eta_1). \quad (4.9)$$

La solución de esta ecuación diferencial en derivadas parciales es muy fácil, la función incógnita  $\psi$  es la primitiva de la función  $F(\eta_1)$  (llamémosla  $f(\eta_1)$ ), más una constante que puede depender de  $\eta_2$  (llamémosla  $g(\eta_2)$ ), es decir que hemos demostrado que la solución más general de la ec. (3.44) es

$$\psi(x, t) = f(x - vt) + g(x + vt), \quad (4.10)$$

que es lo que queríamos demostrar.

Si en un medio continuo infinito hubiera una única onda progresiva, seguiría propagándose sin modificaciones. Pero si la onda progresiva se encontrara con una discontinuidad, necesariamente debería aparecer una onda regresiva, para poder ajustar las condiciones de contorno en la discontinuidad. Por este motivo en sistemas limitados en el espacio, como en la sogá con extremos fijos, sobreviven sólo ondas estacionarias, que pueden ser interpretadas como superposición de ondas viajeras con amplitudes y periodicidades espaciales adecuadas para ajustar las condiciones en ambos extremos.

### 4.1.3. Fórmula de d'Alembert

Según (3.73) y (3.74), toda onda estacionaria es superposición de dos ondas viajeras, una progresiva y otra regresiva. Y según (4.10), toda onda unidimensional que depende de una coordenada fija en el espacio es siempre combinación lineal de dos ondas: una progresiva y otra regresiva. Si bien el resultado (4.10) es muy fuerte, porque no toda función de dos variables se escribe como suma de dos funciones de una sola variable, su demostración no dice nada sobre la forma que tienen las funciones  $f$  y  $g$ .

En la sección 3.8, cuando interpretamos en términos de ondas progresivas y regresivas la simulación de la cuerda con extremos fijos y condiciones iniciales de onda cuadrada, quedaba claro que eran las condiciones iniciales  $\psi(x, 0)$  y  $\partial\psi(x, 0)/\partial t$  las que fijaban la forma de las funciones  $f$  y  $g$ . En dicha simulación (ver figuras (3.14), (3.15) y (3.16)), la velocidad inicial de todos los puntos de la sogá era nula y con esta condición, y reescribiendo los modos como combinación de ondas progresivas y regresivas, resultó fácil encontrar que las funciones  $f$  y  $g$  eran pulsitos cuadrados, idénticos entre sí e idénticos a la forma inicial, pero con altura mitad. También vimos en la sección 3.8 que con la separación progresivas-regresivas, en vez de usar directamente los modos, era posible explicar muy satisfactoriamente la evolución del apartamiento del equilibrio en base al desplazamiento de los pulsitos. ¿Será siempre posible encontrar las funciones  $f$  y  $g$  para condiciones iniciales arbitrarias? La respuesta es positiva y veremos a continuación la demostración dada en 1747 por Jean le Rond d'Alembert, uno de los grandes genios del Iluminismo.

El procedimiento de d'Alembert parte de imponer las condiciones iniciales en la solución general dada por la ec. (4.10), primero sobre la posición inicial

$$\psi(x, 0) = f(x) + g(x), \quad (4.11)$$

y luego sobre la velocidad inicial

$$\frac{\partial\psi(x,0)}{\partial t} = v[-f'(x) + g'(x)]. \quad (4.12)$$

Integrando miembro a miembro la última ecuación, queda

$$h(x) = v[-f(x) + g(x)], \quad (4.13)$$

donde  $h(x)$  es una primitiva de la función de  $x$  que da la velocidad inicial de la soga

$$h(x) = \int \frac{\partial\psi(x,0)}{\partial t} dx + C, \quad (4.14)$$

definida a menos de una constante  $C$ . Despejando  $f(x)$  y  $g(x)$  del sistema de ecs. (4.11) y (4.14), se obtiene

$$f(x) = \frac{1}{2} \left[ \psi(x,0) - \frac{1}{v} h(x) \right], \quad (4.15)$$

$$g(x) = \frac{1}{2} \left[ \psi(x,0) + \frac{1}{v} h(x) \right]. \quad (4.16)$$

Las ecs. (4.15) y (4.16) dan la forma funcional de las funciones de una variable asociadas con las partes progresiva y regresiva de la solución más general de la ecuación de ondas. Para obtener la dependencia funcional correcta de la función de onda con las variables  $x$  y  $t$  usamos (4.10):

$$\psi(x,t) = \frac{1}{2} \left[ \psi(x-vt,0) + \psi(x+vt,0) \right] + \frac{1}{2v} \left[ h(x+vt) - h(x-vt) \right],$$

que si se usa (4.14) se puede reescribir como

$$\psi(x,t) = \frac{1}{2} \left[ \psi(x-vt,0) + \psi(x+vt,0) \right] + \frac{1}{2v} \int_{x-vt}^{x+vt} \frac{\partial\psi(x,0)}{\partial t} dx. \quad (4.17)$$

Notar que cuando la velocidad inicial de todos los puntos de la soga es cero, se anula el segundo término de (4.17) y el resultado se reduce al obtenido en (3.78).

A diferencia de la demostración en 3.8, la demostración de d'Alembert no hace referencia a dependencias temporales armónicas ni a series de Fourier. A pesar de ser descripciones equivalentes, es obvio que la descripción de la evolución de una cuerda elástica *à la d'Alembert* (ec. (4.17)) es muy diferente a la descripción *à la Fourier* (ecs. (3.55), (3.71) y (3.72)). En Física se habla de *dualidad* para describir modelos teóricos que parecen ser diferentes pero que son equivalentes, porque predicen exactamente las mismas propiedades físicas. En los ejemplos “triviales”, dos teorías parecen ser diferentes debido sólo al modo en que se presentan. Sería como si las teorías se enunciaran en español y en chino: un físico que dominara ambas lenguas podría realizar fácilmente una traducción entre una y otra lengua y demostraría así su equivalencia. En un caso trivial, con la traducción no se gana nada desde el punto de vista de la Física, un cambio de lenguaje no aporta nuevas ideas. En cambio, en los ejemplos no triviales, distintas descripciones de la misma situación física producen ideas físicas y métodos matemáticos de análisis

diferentes y complementarios. Es lo que pasa en este caso entre las descripciones de d'Alambert y de modos, cada descripción tiene su ventaja, a diferencia del caso del español y el chino, hay ideas físicas muy importantes que toman protagonismo en cada una de estas descripciones duales, tales como el concepto de frecuencia de resonancia o de  $\lambda_p$  en la descripción de modos, o el concepto de velocidad de propagación o de ondas viajeras en la descripción de d'Alambert. Mientras que la descripción de modos parece ser la más adecuada para explicar el timbre de un instrumento de cuerdas, o efectos como los sonidos armónicos de la guitarra, la descripción de d'Alambert parece ser la más adecuada para explicar la evolución de la deformación en la simulación mostrada en las figuras 3.14 y 3.15). Ambas descripciones dan el mismo resultado físico, de hecho los resultados en las figuras 3.14 y 3.15), a los que se agregan los de la figura 3.16, fueron obtenidos con la descripción de modos, pero para interpretar los “pulsitos”, necesitamos la descripción de d'Alambert. En el curso de F2 veremos otros ejemplos de descripciones duales.

Con respecto a los cuadros que se muestran en la figura 3.16, es interesante prestar atención a las ondas reflejadas en ambos extremos e interpretar la deformación de la cuerda elástica en términos de módulo y fase del coeficiente de reflexión. ¿Cómo cambiaría la evolución de la misma deformación inicial si la cuerda tuviera el extremo derecho libre?

## 4.2. Ondas, medios y dimensiones

Hasta ahora hemos visto comportamientos físicos que dependen de una dirección fija en el espacio, esencialmente porque la posición de equilibrio de los sistemas estudiados, como las sogas o las cadenas con cuentas, *es* una dirección fija en el espacio. Pero en otros sistemas donde la posición de equilibrio se describe con más dimensiones, también puede haber comportamientos ondulatorios dependientes solamente de una dirección fija en el espacio, siempre que esta simetría espacial se introduzca mediante los contornos, las deformaciones iniciales o las fuerzas impulsoras. Esto es lo que sucede por ejemplo en tubos de gas, o en canales con líquido, o en recipientes con gases, cuando las perturbaciones se hacen mediante pistones o compuertas planas. Si bien las ondas son unidimensionales, en el sentido de que se pueden describir con una sola variable espacial, es importante notar que son unidimensionales en cartesianas, no en esféricas ni en cilíndricas. Por tratarse de una dirección fija en el espacio, hemos obtenido ecuaciones como la ec. (3.44) o la ec. (??), donde en cada punto la derivada segunda temporal tiene en cuenta la aceleración de la parte móvil mientras que la derivada segunda espacial está relacionada con la fuerza restitutiva sobre dicha parte móvil producida por vecinos cercanos.

En las redes unidimensionales a lo largo del eje  $x$  la fuerza restitutiva producida por vecinos cercanos da lugar a derivadas segundas con respecto a  $x$ . Pero el número de vecinos cercanos aumenta con la dimensión del sistema. Es fácil darse cuenta que cuando las partes móviles interactúan en mallas bidimensionales, cada parte móvil no sólo tiene vecinos a lo largo de la dirección  $x$ , sino también vecinos en la dirección perpendicular, digamos  $y$ . Por este motivo, en el caso 2D la fuerza restitutiva que produce la aceleración tiene que tener dos contribuciones: i) por un lado la contribución de antes, de los vecinos cercanos en  $x$  y que da lugar a la derivadas segunda con respecto a  $x$ ; y ii) la nueva contribución de los vecinos cercanos en  $y$  que en el límite continuo dará lugar a una derivada segunda con respecto a  $y$ . Si el sistema es isótropo, es decir si  $x$  e  $y$  son indistinguibles, las dos fuerzas restitutivas se suman con el mismo factor de peso. Por este motivo, no debe sorprendernos que la ecuación de ondas para una membrana

elástica (el análogo 2D de la cuerda), tenga la forma

$$\frac{\partial^2 \psi}{\partial t^2} = v^2 \nabla^2 \psi, \quad (4.18)$$

donde  $\nabla^2$  indica la suma de las derivadas segundas espaciales  $\nabla^2 \psi = \frac{\partial^2 \psi}{\partial x^2} + \frac{\partial^2 \psi}{\partial y^2}$ . Un razonamiento similar es válido cuando las partes móviles interactúan en mallas 3D, donde de nuevo se obtiene la ec. (4.18), pero ahora  $\nabla^2 \psi = \frac{\partial^2 \psi}{\partial x^2} + \frac{\partial^2 \psi}{\partial y^2} + \frac{\partial^2 \psi}{\partial z^2}$ . Notar que si los resortes a lo largo de  $x$  tuvieran distinta elasticidad que los resortes a lo largo de  $y$  (un medio anisótropo) las dos fuerzas restitutivas se sumarían con distinto factor de peso y en tal caso no se podría formar el laplaciano a partir de las derivadas segundas.

### 4.2.1. La onda plana en un medio 3D

En un medio 3D, las soluciones de la ec. (4.18) que dependen de una dirección fija en el espacio tienen que tener la forma de la ec. (4.10), es decir que siempre son combinación lineal de dos ondas: una progresiva y otra regresiva. Para que esto sea así tenemos que poder elegir el sistema de coordenadas de manera tal que el eje  $x$  sea *esa dirección fija en el espacio*. Si no tenemos esta libertad, por ejemplo porque conviene elegir los ejes para que se adapten a los contornos, podría ser problemático. Por ejemplo, quiero mirar una perturbación que avanza por la diagonal que va del techo al piso de una habitación con forma de paralelepípedo rectangular, pero sería deseable que los planos  $x - y$ ,  $x - z$  e  $y - z$  coincidan con las paredes. Sea  $\vec{r}$  el vector posición de cada punto del medio 3D y el versor  $\hat{n}$  la dirección de esta diagonal. Estamos de acuerdo en que convendría elegir el eje  $x$  a lo largo de esta diagonal y en este caso la onda progresiva sería  $f(x - vt)$  y no tendríamos que usar notación vectorial con el vector posición  $\vec{r}$ . Pero cuando no se puede elegir el eje  $x$  a lo largo de la dirección fija en el espacio (en este ejemplo, la diagonal), no hay otro remedio que usar  $\vec{r}$ . En realidad todo se reduce al caso sencillo si notamos que los puntos con igual perturbación son planos perpendiculares al versor  $\hat{n}$ , es decir que cada plano se caracteriza porque sus puntos tienen la misma proyección del vector  $\vec{r}$  sobre  $\hat{n}$ . Esto equivale a decir que  $\hat{n} \cdot \vec{r}$  juega el papel que nos habría gustado que jugara  $x$  si se hubiera podido elegir el eje  $x$  a lo largo de la diagonal. Toda esta discusión muestra que la **onda plana** progresiva en un medio 3D se escribe como

$$\psi(\vec{r}, t) = f(\hat{n} \cdot \vec{r} - vt), \quad (4.19)$$

mientras que la regresiva

$$\psi(\vec{r}, t) = g(\hat{n} \cdot \vec{r} + vt). \quad (4.20)$$

### La onda plana armónica en un medio 3D

Si existen motivos para afirmar que la onda, además de ser plana, tiene que tener dependencia temporal armónica de frecuencia  $\omega$ , entonces se puede encontrar la forma de  $f$  y  $g$  como ya hicimos muchas veces, es decir haciendo la derivada segunda temporal, integrando la parte espacial y llegando a senos y cosenos espaciales. Otra manera es pensar esto con el resultado de la ec. (4.10). Consideremos la progresiva: decir que sólo es función de  $\hat{n} \cdot \vec{r} - vt$  es lo mismo que decir que sólo es función de  $\hat{n} \cdot \vec{r}/v - t$ , que equivale a decir que sólo es función de  $t - \hat{n} \cdot \vec{r}/v$ . Y

acá agregamos que si sabemos que debe tener dependencia temporal armónica de frecuencia  $\omega$  (por ejemplo, porque es la respuesta a un forzado estacionario) entonces debe variar como un coseno o un seno de  $\omega t$ . Pero la función  $f$  no depende de  $t$  y de  $\hat{n} \cdot \vec{r}$  por separado, sino solamente en la combinación  $t - \hat{n} \cdot \vec{r}/v$ . Entonces, resulta inevitable que toda perturbación progresiva que: i) sea solución de la ec. de ondas (4.18), ii) dependa de una dirección fija en el espacio y iii) sea armónica, se tenga que escribir, a menos de una fase constante, como

$$\psi(\vec{r}, t) = A \cos[\omega(t - \hat{n} \cdot \vec{r}/v)] = A \cos(\omega t - \vec{k} \cdot \vec{r}) = A \cos(\vec{k} \cdot \vec{r} - \omega t), \quad (4.21)$$

donde se ha definido el **vector de ondas**  $\vec{k}$

$$\vec{k} = \frac{\omega}{v} \hat{n}, \quad (4.22)$$

cuya magnitud es la del  $k = 2\pi/\lambda$  de antes, y su dirección es la de propagación. Obviamente, puede resultar conveniente usar senos en lugar de cosenos, o notación exponencial. De manera análoga se obtienen las representaciones de las ondas planas regresivas armónicas (cambia un signo).

Hemos prestado atención a las ondas planas armónicas porque las soluciones de la forma (4.21) forman base, es decir **cualquier** solución de la ecuación (4.18) se puede escribir como superposición conveniente de estas ondas, con distintos valores de  $\omega$ , distintas direcciones de  $\vec{k}$  y con el módulo del vector  $\vec{k}$  determinado por su definición de la ec. (4.22), basada en la ecuación de dispersión.

### 4.2.2. Fuentes con simetría esférica

Supongamos un medio 3D isótropo y que en este medio la magnitud física  $\psi$  satisface la ecuación de ondas clásica (4.18), con  $\nabla^2 \psi = \frac{\partial^2 \psi}{\partial x^2} + \frac{\partial^2 \psi}{\partial y^2} + \frac{\partial^2 \psi}{\partial z^2}$ . Si las condiciones iniciales o las fuentes impulsoras o forzantes tienen simetría esférica, la descripción con coordenadas cartesianas no es lo más conveniente (en realidad es horrible). Por ejemplo, cuando se informa que una explosión “se escuchó a una distancia de cien kilómetros”, se está especificando el valor de la coordenada esférica  $r$  y se da por entendido que no es necesario dar los valores de las coordenadas esféricas angulares  $\theta$  y  $\phi$ , porque los efectos de la explosión poseen simetría esférica o de revolución, es decir, que son iguales en todas las direcciones. La misma simetría aparece en el caso de fuentes pequeñas, que desde lejos se ven como puntuales. Si bien ideales, las fuentes puntuales cobran importancia cuando vale superposición, porque cualquier fuente extensa se puede analizar como una colección de fuentes puntuales.

En todos los casos mencionados, la simetría esférica de la fuente permite afirmar que la función de onda depende solamente de la coordenada  $r$ , es decir que debe ser  $\psi(r, t)$  y no  $\psi(r, \theta, \phi, t)$ . Los puntos del espacio que tienen igual perturbación, o sea, los *frentes de onda*, son esferas concéntricas. En estas condiciones el operador  $\nabla^2$  no es la derivada segunda respecto a  $r$  dos veces, si no que su forma se obtiene a partir del cambio de variables  $(x, y, z) \rightarrow (r, \theta, \phi)$ . Aplicando la regla de la cadena para pasar las derivadas con respecto a coordenadas cartesianas a derivadas con respecto a coordenadas esféricas, el laplaciano resulta

$$\frac{1}{r^2} \frac{\partial}{\partial r} \left( r^2 \frac{\partial \psi}{\partial r} \right) + \text{derivadas angulares} = \frac{1}{r} \frac{\partial^2}{\partial r^2} (r \psi) + \text{derivadas angulares}. \quad (4.23)$$

En el caso en que  $\psi(r, t)$ , las derivadas con respecto a  $\theta$  y  $\phi$  son cero y la eq. (4.18) queda

$$\frac{1}{r} \frac{\partial^2}{\partial r^2}(r \psi) = \frac{1}{v^2} \frac{\partial^2 \psi}{\partial t^2}, \quad (4.24)$$

es decir,

$$\frac{\partial^2}{\partial r^2}(r \psi) = \frac{1}{v^2} \frac{\partial^2(r \psi)}{\partial t^2}. \quad (4.25)$$

Este resultado muestra que la combinación  $r \psi$  satisface la ecuación de ondas 1D (3.44). Luego, la solución más general para  $r \psi$  viene dada por la ec. (4.10)

$$r \psi = f(r - vt) + g(r + vt), \quad (4.26)$$

y la función de onda más general para simetría esférica resulta la suma de una perturbación progresiva  $f/r$  (que se aleja del origen, explosión, o divergencia) mas una perturbación regresiva  $f/r$  (que se acerca al origen, implosión, o convergencia)

$$\psi(r, t) = \frac{f(r - vt)}{r} + \frac{g(r + vt)}{r}. \quad (4.27)$$

Para fuentes armónicas, la onda esférica progresiva que se aleja del origen resulta entonces, a menos de una fase,

$$\psi(r, t) = \frac{A}{r} \cos(kr - \omega t), \quad (4.28)$$

o en notación exponencial

$$\psi(r, t) = \frac{A}{r} \exp i(kr - \omega t). \quad (4.29)$$

Si el medio no tiene pérdidas, la energía transportada por un frente de onda particular tiene que ser constante a medida que el frente se aleja del origen. Este resultado es consistente con el hecho de que la amplitud resulta inversamente proporcional a  $r$ . Así, en cada elemento de área la densidad de energía por unidad de área es inversamente proporcional a  $r^2$ , mientras que el área total del frente de ondas aumenta como  $r^2$ . Esto explica por qué cuanto más lejos estamos del emisor, más débil se recibe la señal emitida, a menos que se agrande el área del detector, para compensar el efecto de la dependencia inversamente proporcional a  $r^2$ . Justamente “agrandar el área del detector” es lo que se hace en los radiotelescopios de observatorios astronómicos con antenas de gran tamaño, como RATAN-600 o Arecibo, con antenas de diámetro 576 m y 305 m, respectivamente.

Notar que muy lejos de las fuentes, y para detectores pequeños, un frente de onda esférico puede aproximarse por un frente de onda plano.

### 4.2.3. Fuentes con simetría cilíndrica

Si las condiciones iniciales o las fuentes impulsoras o forzantes tienen simetría cilíndrica, en vez de esférica, se debe obtener la forma del laplaciano con argumentos similares a los empleados para obtener la eq. (4.24), pero ahora se debe hacer el cambio de variables de las coordenadas cartesianas  $(x, y, z)$  a las coordenadas cilíndricas  $(r, \theta, z)$  (no es el mismo  $r$  de antes). De esta

manera, la ecuación para ondas emitidas por fuentes con simetría cilíndrica (la de un alambre infinito), resulta

$$\frac{1}{r} \frac{\partial}{\partial r} \left( r \frac{\partial \psi}{\partial r} \right) = \frac{1}{v^2} \frac{\partial^2 \psi}{\partial t^2}. \quad (4.30)$$

Notar que los puntos del espacio con igual perturbación, o sea, los frentes de onda, son cilindros concéntricos.

Las soluciones armónicas de (4.30) tienen una dependencia radial dada por las funciones de Bessel. Para lo que necesitamos en este curso, solamente diremos que el comportamiento asintótico de las soluciones armónicas lejos de las fuentes,  $r \gg \lambda$ , es

$$\psi(r, t) = \frac{A}{\sqrt{r}} \exp i(kr - \omega t). \quad (4.31)$$

De nuevo, si el medio no tiene pérdidas, la energía transportada por un frente de onda particular (en forma de cilindro) tiene que ser constante a medida que el frente se aleja del origen. Este resultado es consistente con el hecho de que la amplitud resulta inversamente proporcional a  $\sqrt{r}$ . Así, en cada elemento de área la densidad de energía por unidad de área y por unidad de tiempo (la intensidad) es inversamente proporcional a  $r$ , mientras que el área total del frente de ondas aumenta como  $r$ . Para puntos de observación muy lejanos, y para detectores pequeños, un frente de onda cilíndrico puede aproximarse por un frente de onda plano.

### 4.3. Modulación

Hasta ahora prestamos atención a oscilaciones y ondas con dependencia temporal armónica, es decir “cosenos de  $\omega t$ ” (o senos o exponenciales). ¿Por qué la fijación con esta dependencia temporal? Por lo visto anteriormente, hay solamente dos respuestas posibles (ok, hay muchas más, pero solamente dos tuvieron protagonismo en nuestro curso hasta ahora). Como ya se habrán respondido mentalmente, una de las respuestas posibles es *porque queríamos ver los modos*. Y la segunda respuesta posible es *porque estábamos interesados en la respuesta forzada estacionaria*. Si el extremo de un sistema se fuerza armónicamente,  $\psi(0, t) = A_0 \cos \omega t$ , y si la perturbación  $\psi$  es solución de la ecuación de ondas clásica 1D, entonces la propagación de la perturbación progresiva se obtiene reemplazando  $t$  por  $t - x/v(\omega)$  y así se tiene que  $\psi(x, t) = A_0 \cos(k(\omega)x - \omega t)$ . Pero con este tipo de perturbación, vista como una señal para transmitir información, no se puede transmitir nada, es como querer comunicarse cantando siempre la misma nota. Para transmitir información hay que hacer cambios de algo, ya sea cambiar la intensidad, el ritmo, la frecuencia, la fase . . . algo. Esto se llama **modular**: producir cambios que sirvan para codificar información. Si la fuerza impulsora en el extremo de un sistema que transmite ondas responde a estos cambios temporales, el sistema se puede usar para que los cambios viajen, lleguen a un receptor y así transmitir información.

A continuación veremos un tercer motivo para justificar tanto interés por las funciones armónicas. Y este tercer motivo surge de una sorprendente propiedad de una gran clase de funciones que empiezan y terminan (no como los senos y cosenos, que no empiezan ni terminan nunca). Resulta que los elementos de esta gran clase de funciones se pueden escribir como superposición continua de funciones con dependencia temporal armónica. Por este motivo, a partir de este preciso momento centraremos la atención en entender los nuevos aspectos que aparecen cuando

se superponen distintas frecuencias. El primer aspecto nuevo es incorporar que la velocidad de fase  $v(\omega)$  puede depender de la frecuencia. Si bien para el ejemplo de la cuerda o para ondas electromagnéticas en el vacío la velocidad de fase es una constante y entonces la ecuación de dispersión  $k(\omega) = \omega/v$  es una función lineal de  $\omega$ , en otros casos (como para las ondas electromagnéticas en medios materiales) la velocidad de fase  $v(\omega)$  puede depender de la frecuencia y en este caso  $k(\omega)$  no es una función lineal. Cuando la velocidad de fase depende de la frecuencia, se dice que el medio es **dispersivo** (otra acepción de dispersivo, la acepción anterior estaba en la página 107).

### 4.3.1. Caso más sencillo

El ejemplo más sencillo de superposición de funciones con dependencia temporal armónica es el de dos frecuencias distintas con la misma amplitud y fase. El resultado relevante está en la ecuación (2.44). Supongamos que la fuerza impulsora al comienzo del sistema, en vez de ser monocromática como dijimos al comienzo de 4.3, produce una deformación  $\psi(0, t)$  de la forma

$$\psi(0, t) = A_0(\cos \omega_1 t + \cos \omega_2 t) = \underbrace{2A_0 \cos \omega_m t}_{A_m(t)} \cos \omega_p t = A_m(t) \cos \omega_p t. \quad (4.32)$$

Como antes, si la perturbación  $\psi$  es solución de la ecuación de ondas clásica 1D, las perturbaciones progresivas asociadas con  $\omega_1$  y con  $\omega_2$  se obtienen reemplazando  $t$  por  $t - x/v(\omega)$ . Así se tiene que la perturbación resultante se puede escribir como

$$\psi(x, t) = A_0 \cos(k(\omega_1)x - \omega_1 t) + A_0 \cos(k(\omega_2)x - \omega_2 t) = \underbrace{2A_0 \cos(k_m x - \omega_m t)}_{A_m(x, t)} \cos(k_p x - \omega_p t), \quad (4.33)$$

con  $k(\omega_1) = \omega_1/v(\omega_1)$  y  $k(\omega_2) = \omega_2/v(\omega_2)$ . Es decir, que la perturbación resultante se obtiene como una perturbación portadora, de frecuencia

$$\omega_p = (\omega_1 + \omega_2)/2, \quad (4.34)$$

y número de onda

$$k_p = (k(\omega_1) + k(\omega_2))/2, \quad (4.35)$$

**modulada** por una perturbación moduladora, de frecuencia

$$\omega_m = (\omega_1 - \omega_2)/2, \quad (4.36)$$

y número de onda

$$k_m = (k(\omega_1) - k(\omega_2))/2. \quad (4.37)$$

Si queremos seguir al coseno que corresponde a la modulación (es decir, si queremos ir siempre con un punto característico, como la cresta, o el valle), entonces hay que mantener el argumento  $k_m x - \omega_m t$  en el mismo valor. Dicho de otra manera, para que  $A_m(x + \Delta x, t + \Delta t) = A_m(x, t)$ ,  $\Delta x$  y  $\Delta t$  tienen que estar ligados por la relación

$$\frac{\Delta x}{\Delta t} = \frac{\omega_m}{k_m} = \frac{\omega_1 - \omega_2}{k(\omega_1) - k(\omega_2)}. \quad (4.38)$$

En el límite de frecuencias muy próximas, el lado derecho de esta ecuación tiende a la derivada  $d\omega/dk$ , la *derivada* de la relación de dispersión, mientras que el lado izquierdo tiende a la velocidad necesaria para ir siempre con un punto característico de la modulación. Esta velocidad se llama **velocidad de grupo** y siempre coincide con la velocidad de propagación de la información, es decir de la energía

$$v_g = \frac{d\omega}{dk}, \quad (4.39)$$

que por lo general no coincide con la velocidad de fase, a menos que la relación de dispersión sea lineal, es decir cuando la velocidad de fase  $v$  no depende de la frecuencia.

### 4.3.2. Modulaciones reales

El caso de pulsaciones es un caso especial de fuerzas impulsoras cuya dependencia temporal puede expresarse como combinación lineal de distintas funciones armónicas,

$$\psi(0, t) = \sum_{\omega} A(\omega) \cos [\omega t + \phi(\omega)], \quad (4.40)$$

y esta combinación lineal también resulta

$$\psi(0, t) = A_{mod}(t) \cos \omega_p t, \quad (4.41)$$

producto de una perturbación portadora, a la frecuencia promedio, por una perturbación moduladora. En el caso fácil pero superidealizado de pulsaciones, la moduladora es periódica a la frecuencia semidiferencia. Pero en general, en los sistemas reales, la moduladora contiene funciones que codifican la información que se quiere transmitir, ya sea en forma digital (una secuencia de ceros y unos, como la información musical en los DVDs) o en forma analógica (como la información musical en los discos de vinilo). Para elegir la frecuencia promedio (de la portadora) se deben poner de acuerdo el emisor y el receptor, aunque en la práctica esto no es posible para ondas electromagnéticas (la base de la sociedad de la información tal como la conocemos) porque existen en todo el mundo diversos entes de control que gestionan el uso del espectro para cada aplicación. Para ondas de radio la portadora tiene la frecuencia que marca el dial,  $\nu \approx 10^3$  kHz para AM,  $\nu \approx 10^2$  MHz para FM. En cambio, las frecuencias que componen la moduladora vienen impuestas por las frecuencias características de la información que se quiere transmitir.

Tomemos el ejemplo sencillo de radio AM. La perturbación en el comienzo del sistema (es decir, las corrientes eléctricas en la antena transmisora) tienen la forma (4.41), pero la modulación  $A_{mod}(t)$ , en vez de ser una simple función armónica a la frecuencia semidiferencia, se sintetiza mediante una suma de funciones armónicas. Esta suma tiene un término constante, que representa la presión atmosférica en el estudio de radio cuando todo está en silencio, más otros términos armónicos con amplitudes  $A_{Mj}$ , frecuencias  $\omega_{Mj}$  y fases  $\phi_{Mj}$  que representan los sonidos que se desea transmitir, ya sea música, un discurso o una conversación. Resumiendo, la modulación  $A_{mod}(t)$  se expresa de la siguiente manera

$$A_{mod}(t) = A_{M0} + \sum_j A_{Mj} \cos [\omega_{Mj} t + \phi_{Mj}], \quad (4.42)$$

donde el primer subíndice es una etiqueta que dice que estamos hablando de una modulación, mientras que el segundo subíndice numera de alguna manera a las frecuencias moduladoras. Como ésta es la forma que modula a la portadora, al introducir (4.42) en (4.41), y luego de usar la identidad trigonométrica que transforma el producto de cosenos en una suma, se llega a que la perturbación en el comienzo del sistema se escribe como

$$\begin{aligned} \psi(0, t) = A_{M0} \cos \omega_p t + \frac{1}{2} \sum_j A_{Mj} \cos [(\omega_{Mj} + \omega_p)t + \phi_{Mj}] \\ + \frac{1}{2} \sum_j A_{Mj} \cos [(\omega_p - \omega_{Mj})t - \phi_{Mj}], \end{aligned} \quad (4.43)$$

La expresión (4.43) pone en evidencia que para poder irradiar ondas que tengan toda la información sobre los sonidos audibles que ocurren en el estudio de radio, entonces la “fuerza impulsora” en la antena transmisora tiene que estar sintetizada con señales armónicas con frecuencias angulares entre la menor frecuencia contenida en la segunda sumatoria (conocida como banda lateral inferior), y la mayor frecuencia contenida en la primera sumatoria (conocida como banda lateral superior), es decir

$$\omega_p - \omega_M^{Max} \leq \omega \leq \omega_p + \omega_M^{Max}, \quad (4.44)$$

donde  $\omega_M^{Max}$  es el máximo valor de todas las frecuencias de modulación que intervienen en (4.42). La diferencia entre la mayor frecuencia de la banda lateral superior y la menor frecuencia de la banda lateral inferior se conoce como *ancho de banda* y según (4.44) resulta igual a  $2\omega_M^{Max}$ . Debido a la existencia del ancho de banda, el número de canales (estaciones de radio, de televisión, etc.) en un dado rango de frecuencias está limitado. Si estamos hablando de frecuencias de audio, entre  $\nu \approx 50$  Hz y  $\nu \approx 20$  kHz, el ancho de banda debería ser de aproximadamente 40 kHz, pero debido que AM no es muy pretenciosa con la fidelidad, y teniendo en cuenta que la nota más alta del piano tiene frecuencia  $\nu \approx 4.2$  kHz, es usual tener anchos de banda del orden de 5 kHz. En cambio para televisión, las frecuencias de modulación vienen impuestas por la persistencia retiniana, el número de pixeles de cada cuadro, el número de colores . . . y es usual tener anchos de banda del orden de algunas decenas de MHz (y por este motivo, las frecuencias portadoras de las transmisiones aéreas de televisión están en el rango de las centenas de MHz).

El estudio de la dispersión en un medio es muy importante para las comunicaciones. Esto es así porque si el medio es muy dispersivo, la velocidad de fase depende fuertemente de la frecuencia, y entonces la señal emitida por la fuerza impulsora  $\psi(0, t)$  (o la antena emisora, o cualquier perturbación que se imponga en el medio para que llegue a un receptor), se va a ir degradando debido a que las distintas componentes frecuenciales viajan con distintas velocidades de fase y entonces la modulación no mantiene su forma a medida que la onda avanza. En cambio, si el medio es poco dispersivo y si el contenido de frecuencias de la señal no es muy grande, la modulación mantiene su forma y se propaga con la velocidad de grupo  $v_g$ .

### 4.3.3. Síntesis de señales

Ya vimos que cuando el extremo de un sistema que responde a la ecuación de ondas clásica 1D se fuerza armónicamente para que la perturbación inicial tenga la forma  $\psi(0, t) = A_0 \cos \omega t$ , la perturbación progresiva se propaga y penetra en el sistema. Si bien esto NO sirve como señal

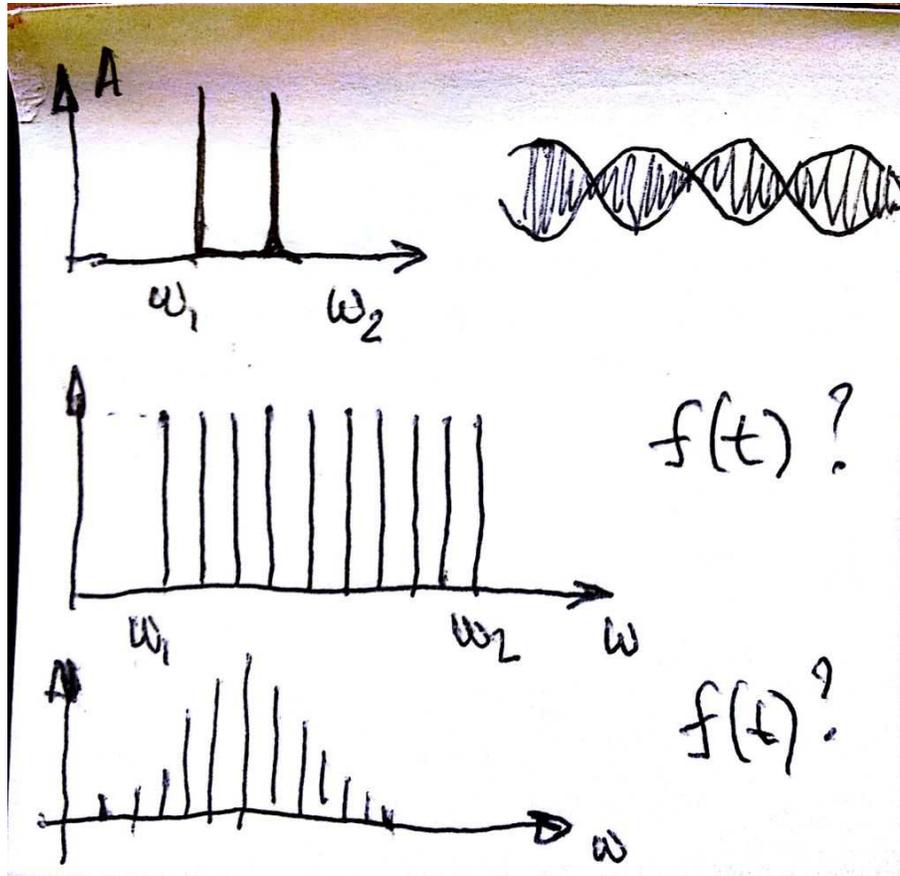


Figura 4.1: Los gráficos de la columna izquierda simbolizan cómo se elige la síntesis de funciones a partir de sumar cosenos. En la primera fila, el ejemplo de los dos cosenos, en la segunda fila, el ejemplo discutido en 4.3.3.

para transmitir información, porque una función puramente periódica no empieza ni termina nunca, Sí sirve para sintetizar otras señales que den lugar a perturbaciones progresivas finitas, es decir, con localización temporal y espacial. Esto es así porque se puede probar que muchas funciones que empiezan y terminan son sintetizables a partir de funciones armónicas con distintas frecuencias. Como la prueba rigurosa está fuera del alcance del curso, a continuación estudiaremos la **síntesis** en un caso sencillo, pero no tan supersencillo como el de dos cosenos con la misma amplitud y fase que vimos antes. Los gráficos de la columna izquierda de la Figura 4.1 representan casos de síntesis de funciones a partir de sumar cosenos. En el eje vertical se grafican las amplitudes y en el eje horizontal las frecuencias de los cosenos usados para la síntesis. En la primera fila, el ejemplo de los dos cosenos con dos frecuencias distintas con la misma amplitud y fase (como un peine con dos dientes de la misma altura), que da lugar a la conocida función del tiempo que corresponde al batido: una portadora a la frecuencia promedio, modulada por una amplitud variable en el tiempo y que viaja a la velocidad de grupo. Las funciones de las amplitudes en función de las frecuencias representadas en los gráficos de la columna izquierda de la Figura 4.1 se conocen como *espectros*. Otros ejemplos que conocemos de funciones que no empiezan ni terminan nunca es el caso de funciones periódicas, sintetizadas a partir de series de Fourier, como ya se discutió en una animación, ver enlace en <http://bit.ly/2cDIVqE>.

### Mezclando $N$ componentes armónicos

En la segunda fila de la Figura 4.1 se esquematiza el nuevo ejemplo de la clase de hoy, vamos a sintetizar una función a partir de  $N$  cosenos con la misma amplitud ( $a$ ) y con la misma fase y con frecuencias  $\omega_p$  equiespaciadas en el intervalo  $\omega_1 \leq \omega \leq \omega_2$ . El dibujito ahora es como un peine con  $N$  dientes de la misma altura y a diferencia de ejemplo con las series de Fourier, todas las frecuencias están localizadas en un intervalo finito. Primero investigaremos qué tipo de perturbación inicial  $\psi(0, t)$  se logra con esta mezcla de cosenos. Y luego, cómo se propaga esta perturbación por el sistema.

Si  $\psi(0, t)$  es la suma de estos  $N$  cosenos

$$\psi(0, t) = \sum_{n=0}^{N-1} a \cos [(\omega_1 + n \delta\omega)t], \quad (4.45)$$

con

$$\delta\omega = \frac{\omega_2 - \omega_1}{N - 1} = \frac{\Delta\omega}{N - 1}, \quad \Delta\omega = \omega_2 - \omega_1. \quad (4.46)$$

Conviene pasar a notación compleja, cada coseno es la parte real de una exponencial compleja y las amplitudes y una parte de las exponenciales sale de factor común de la sumatoria. Sobrentendiendo que luego tomaremos parte real y empleando la expresión para la suma de los  $N$  primeros términos de la sucesión geométrica, tenemos que

$$\psi(0, t) = a e^{i\omega_1 t} \sum_{n=0}^{N-1} [e^{i\delta\omega t}]^n = a e^{i\omega_1 t} \frac{e^{iN\delta\omega t} - 1}{e^{i\delta\omega t} - 1}, \quad (4.47)$$

que se puede reescribir como

$$\psi(0, t) = a e^{i[\omega_1 + (N-1)\delta\omega/2]t} \frac{\sin(\frac{N}{2}\delta\omega t)}{\sin(\frac{1}{2}\delta\omega t)} = a \frac{\sin(\frac{N}{2}\delta\omega t)}{\sin(\frac{1}{2}\delta\omega t)} \cos\omega_p t = A_{mod}(0, t) \cos\omega_p t, \quad (4.48)$$

con

$$\omega_p = \omega_1 + \frac{\Delta\omega}{2}, \quad (4.49)$$

la frecuencia promedio (la que está en el medio del peine), tal como ya habíamos anticipado en la eq. (4.41).

### Límite $N \rightarrow \infty$ , con $\Delta\omega$ constante

Si se deja fijo  $\Delta\omega$  pero se hace más pequeño el valor  $\delta\omega$ , es decir, se ponen mas dientes en el peine ( $N \rightarrow \infty$ ),  $T_{mod} \rightarrow \infty$ . Y este resultado equivale a decir que en el límite  $N \rightarrow \infty$ , la moduladora deja de ser periódica (la portadora lo sigue siendo). Es casi evidente que en este límite hemos sintetizado una función no periódica a partir de una suma de cosenos. Veremos que, además de no ser periódica, la moduladora empieza y termina, es decir que puede servir

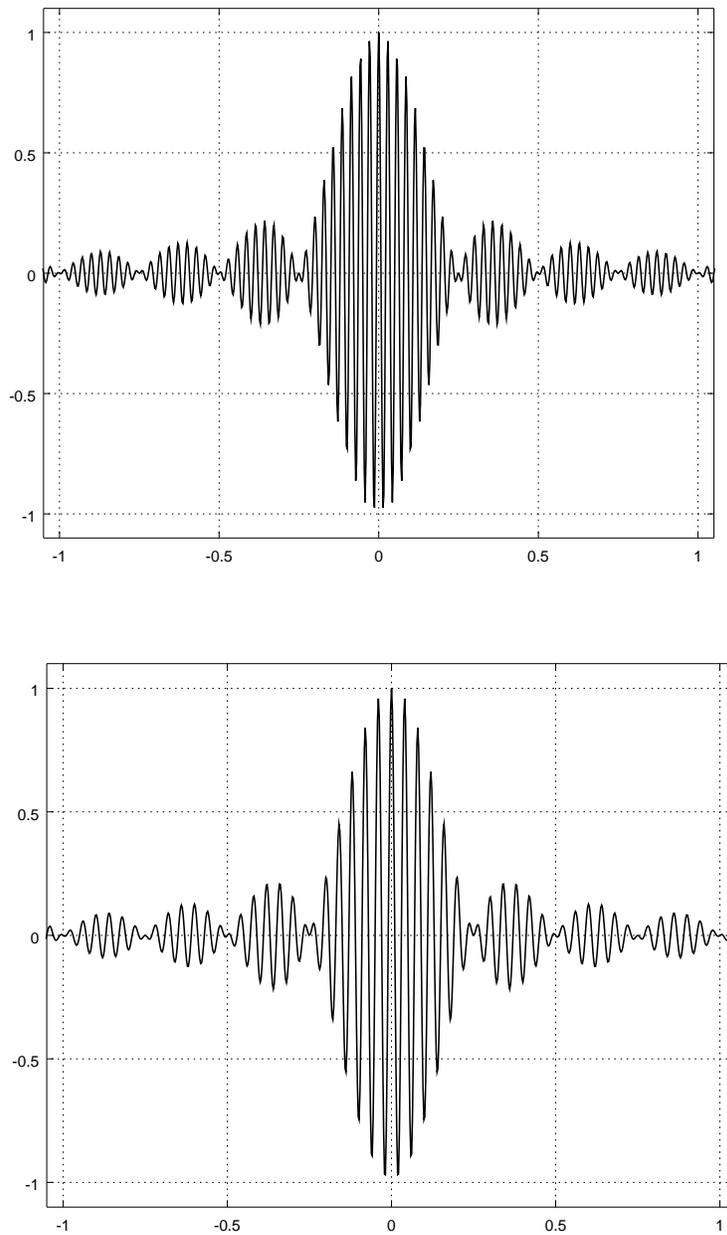


Figura 4.2:  $\psi(0, t)$  dado por la ec. (4.54) en función del tiempo,  $A_0 = 1$  y distintos valores de  $\omega_1$  y  $\omega_2$ . (a)  $\nu_1 = 33\text{Hz}$  y  $\nu_2 = 37\text{Hz}$ ; (b)  $\nu_1 = 23\text{Hz}$  y  $\nu_2 = 27\text{Hz}$ .

para representar señales.

Para que el proceso tenga sentido, al aumentar el número de dientes en el peine tenemos que disminuir la altura de los dientes (las amplitudes  $a$  de las componentes armónicas), porque no queremos que la fuerza impulsora en el extremo del sistema  $\psi(0, t)$  se haga infinita. Para que

esto no pase, pedimos que la amplitud de modulación

$$A_{mod}(0, t) = a \frac{\sin(\frac{N}{2} \delta\omega t)}{\sin(\frac{1}{2} \delta\omega t)}, \quad (4.50)$$

siempre tenga el mismo valor en el instante inicial, independientemente del valor de  $N$ . Y como  $A_{mod}(0, 0) = aN$ , esto equivale a pedir que en este límite  $aN = A_0$  cuando  $N \rightarrow \infty$  ( $A_0$  una constante). De esta manera, conviene escribir la amplitud de modulación como

$$A_{mod}(0, t) = \frac{A_0}{N} \frac{\sin(\frac{N}{2} \delta\omega t)}{\sin(\frac{1}{2} \delta\omega t)}. \quad (4.51)$$

Ahora está todo preparado para hacer el límite  $N \rightarrow \infty$ ,  $\delta\omega \rightarrow 0$ , con  $\Delta\omega$  constante. En este límite,

$$N\delta\omega \approx (N-1)\delta\omega = \Delta\omega = \omega_2 - \omega_1, \quad (4.52)$$

y la amplitud de modulación se puede escribir

$$A_{mod}(0, t) = A_0 \frac{\sin(\frac{1}{2} \Delta\omega t)}{\frac{1}{2} \Delta\omega t}, \quad (4.53)$$

y entonces la dependencia temporal de la perturbación en el comienzo del sistema resulta

$$\psi(0, t) = A_0 \frac{\sin(\frac{1}{2} \Delta\omega t)}{\frac{1}{2} \Delta\omega t} \cos \omega_p t. \quad (4.54)$$

Por otra parte, este resultado se obtuvo calculando  $\psi(0, t)$  como

$$\psi(0, t) = \lim_{N \rightarrow \infty} \sum_{n=0}^{N-1} a \cos [(\omega_1 + n \delta\omega)t] = \lim_{N \rightarrow \infty} \frac{A_0}{\Delta\omega} \sum_{n=0}^{N-1} \cos [(\omega_1 + n \delta\omega)t] \delta\omega, \quad (4.55)$$

que en el límite es igual a

$$\psi(0, t) = \frac{A_0}{\Delta\omega} \int_{\omega_1}^{\omega_2} \cos \omega t d\omega. \quad (4.56)$$

Las eqs. (4.54) y (4.56) muestran que

$$\frac{A_0}{\Delta\omega} \int_{\omega_1}^{\omega_2} \cos \omega t d\omega = A_0 \frac{\sin(\frac{1}{2} \Delta\omega t)}{\frac{1}{2} \Delta\omega t} \cos \omega_p t, \quad (4.57)$$

es decir, que sumar infinitos cosenos de la misma amplitud y la misma fase en el intervalo  $\omega_1 < \omega < \omega_2$  (lo que equivale a hacer  $\int_{\omega_1}^{\omega_2}$  de una función constante multiplicada por cosenos), da como resultado una señal portadora a la frecuencia promedio, modulada por una amplitud moduladora variable en el tiempo y que tiene la forma de la función  $\sin u/u$ , que se lee **seno cardinal de  $u$**  y que también se suele notar como  $\text{sinc}(u)$ . Es una función par, que tiende a cero cuando  $u$  crece (porque el seno está acotado) y que tiene los ceros del seno, excepto en el origen donde tiene su máximo. La variable  $u$  es

$$u = \frac{1}{2} \Delta\omega t = \pi \Delta\nu t, \quad (4.58)$$

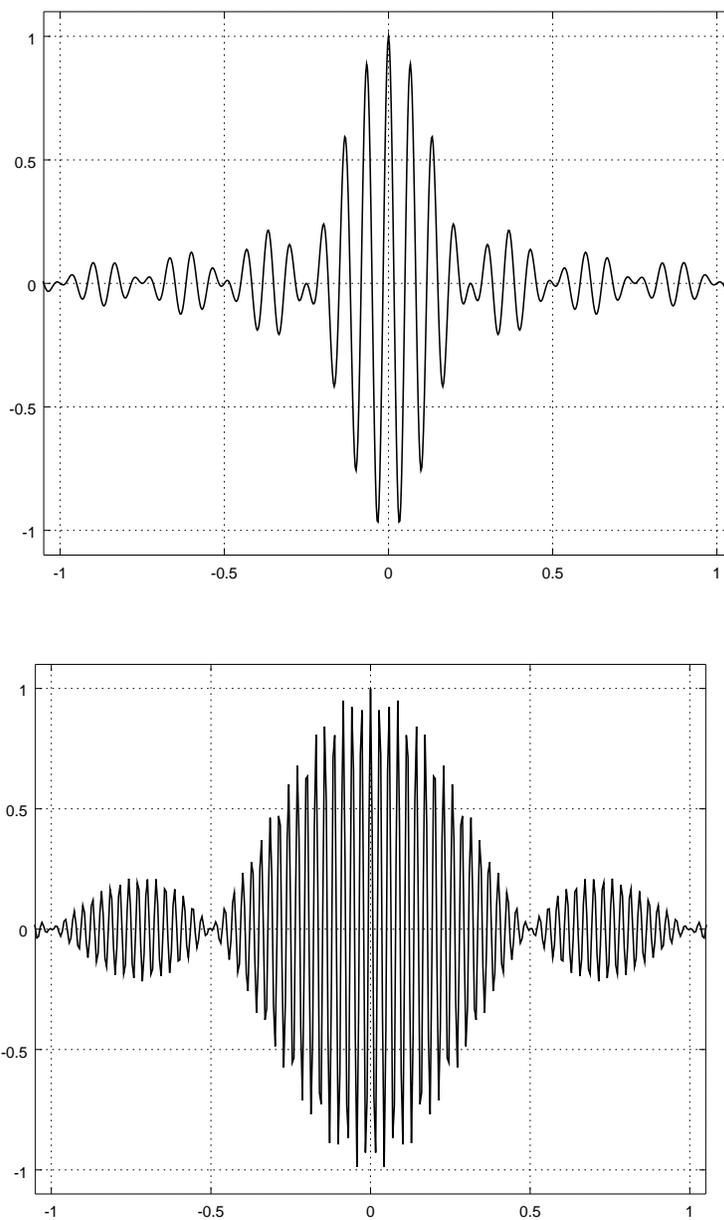


Figura 4.3:  $\psi(0, t)$  como en la Fig. 4.2, pero (a)  $\nu_1 = 13\text{Hz}$  y  $\nu_2 = 17\text{Hz}$ ; (b)  $\nu_1 = 34\text{Hz}$  y  $\nu_2 = 36\text{Hz}$ .

y la (4.54) muestra que si se deja el ancho  $\Delta\omega$  fijo pero se cambian los valores de  $\omega_1$  y  $\omega_2$ , la forma de la modulación es siempre la misma, lo que cambia es el valor de la frecuencia promedio, es decir cuántas oscilaciones de la portadora entran desde que la modulación toma su valor máximo, para  $u = 0$ , hasta que toma su primer cero, para  $u = \pi$ . Desde que la modulación toma su máximo valor hasta que se hace cero por primera vez transcurre un tiempo  $\Delta t$  que resulta

$$\Delta t = 1/\Delta\nu, \quad (4.59)$$

y que puede ser usado como medida de la localización temporal (cuánto dura) la modulación.

La función del tiempo resultante de sumar infinitos cosenos en un intervalo fijo se muestra en las figuras 4.2 y 4.3, donde se ha graficado  $\psi(0, t)$ , dado por la ec. (4.54), en función del tiempo (también se puede graficar en función de  $u$ , cuál sería la ventaja) para  $A_0 = 1$  y distintos valores de  $\omega_1$  y  $\omega_2$ . Notar que a pesar de que los gráficos de de las figura 4.2a, 4.2b y 4.3a corresponden a valores diferentes de  $\nu_1$  y  $\nu_2$ , en todos se mantuvo el mismo valor  $\Delta\nu = 4\text{Hz}$ . Y por este motivo la la modulación en estas tres curvas tiene siempre el mismo ancho característico  $\Delta t = 0.25\text{s}$ , aunque lo que cambia es la frecuencia de la portadora. En cambio, en el ejemplo de la figura 4.3b, el intervalo de frecuencias donde se hace la suma de cosenos se reduce a la mitad,  $\Delta\nu = 2\text{Hz}$ . Por eso ahora la modulación tiene un ancho característico que es el doble,  $\Delta t = 0.5\text{s}$ . Este comportamiento, anunciado por la eq. (4.59), es común a todas las funciones sintetizadas a partir de funciones armónicas. Se conoce como **principio de incerteza** y dice que la localización de la función que se sintetiza a partir de funciones armónicas (en este caso localización temporal) es inversamente proporcional a la localización de las frecuencias de los senos y cosenos usados. No es una propiedad que valga solamente para síntesis temporales, por ejemplo la misma propiedad valdría para la localización espacial de funciones que se sintetizan a partir de funciones como  $\cos kx$ , con frecuencias espaciales  $k$ . El caso límite familiar en el contexto de estos ejemplos es un coseno, que no tiene localización temporal pero está perfectamente localizado frecuencialmente. El caso límite opuesto es un impulso temporal, algo que se prende y se apaga instantáneamente, que está muy bien localizado temporalmente, pero que es de esperar que no esté bien localizado frecuencialmente. Calcular analíticamente la disminución relativa de la altura de los máximos secundarios, mas notoria cuando se observa en la energía de la onda resultante, ya que la energía va con el cuadrado de la perturbación.

Es muy fácil reproducir en Python los gráficos de las figuras 4.2 y 4.3. Hagan sus propios scripts y comparen los resultados obtenidos con los que están en éstas figuras.

#### 4.3.4. Transformada de Fourier

La superposición de cosenos que acabamos de hacer, representada por la eq. (4.56), es un caso particular de funciones del tiempo  $f(t)$  obtenidas mediante superposiciones de la forma

$$f(t) = \int_0^\infty \tilde{A}(\omega) \cos \omega t d\omega + \int_0^\infty \tilde{B}(\omega) \sin \omega t d\omega, \quad (4.60)$$

que se conoce como **integral de Fourier**. Las funciones  $\tilde{A}(\omega)$  y  $\tilde{B}(\omega)$  dan la transformada de Fourier de la función  $f(t)$ . En nuestro ejemplo  $\tilde{B}(\omega) = 0$  para todo  $\omega$  (porque no usamos funciones seno) y  $\tilde{A}(\omega)$  es una función partida con forma rectangular, que vale  $\tilde{A}(\omega) = \frac{A_0}{\Delta\omega}$  para  $\omega_1 < \omega < \omega_2$  y  $\tilde{A}(\omega) = 0$  fuera del intervalo  $\omega_1 < \omega < \omega_2$ . Lo que hicimos al construir la función  $f(t)$  se llama *síntesis* de Fourier. El proceso inverso, que consiste en descomponer una determinada  $f(t)$  en sus componentes armónicas, se llama *análisis* de Fourier. Ambos procesos, síntesis y análisis, son completamente análogos a los vistos para funciones periódicas y pueden verse como el límite para período infinito. Si se hiciera este límite a partir de las ecuaciones obtenidas para series de Fourier (ver página 69), se obtendría que las funciones  $\tilde{A}(\omega)$  y  $\tilde{B}(\omega)$  resultan

$$\tilde{A}(\omega) = \frac{1}{\pi} \int_0^\infty f(t) \cos \omega t dt, \quad (4.61)$$

$$\tilde{B}(\omega) = \frac{1}{\pi} \int_0^{\infty} f(t) \sin \omega t dt. \quad (4.62)$$

Si se prefiere usar notación exponencial, hay que escribir los senos y los cosenos en términos de las exponenciales, separar los términos con exponenciales de exponente con distinto signo, introducir amplitudes complejas  $\tilde{f}(\omega)$  definidas a partir de las amplitudes  $\tilde{A}(\omega)$  y  $\tilde{B}(\omega)$  y extendidas a frecuencias negativas. En notación compleja, la síntesis y el análisis se escriben

$$f(t) = \int_{-\infty}^{\infty} \tilde{f}(\omega) \exp(-i\omega t) d\omega. \quad (4.63)$$

$$\tilde{f}(\omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} f(t) \exp(i\omega t) dt. \quad (4.64)$$

Volviendo a la perturbación  $\psi(0, t)$  al comienzo del sistema, la expresión (4.63) es muy útil para seguir la propagación de una señal. Con  $\psi(0, t)$  como en nuestro ejemplo (eq. (4.57)), la respuesta progresiva  $\psi(x, t)$  que resulta del forzado estacionario en  $x > 0$  se obtiene reemplazando  $t$  por  $t - x/v(\omega)$  en el lado izquierdo de la eq. (4.57). En medios no dispersivos,  $v$  es una constante y la integral debe dar el mismo resultado que antes, excepto que  $t$  se reemplaza por  $t - x/v(\omega)$

$$\psi(x, t) = \frac{A_0}{\Delta\omega} \int_{\omega_1}^{\omega_2} \cos \omega (t - x/v) d\omega = A_0 \frac{\sin(\frac{1}{2} \Delta\omega (t - x/v))}{\frac{1}{2} \Delta\omega (t - x/v)} \cos \omega_p (t - x/v), \quad (4.65)$$

es decir que la forma inicial se propaga a la velocidad  $v$  sin deformarse. Pero si el medio es dispersivo,  $v(\omega)$  no es una constante y entonces el pulso inicial por lo general cambia de forma a medida que avanza. Sólo en el caso en que la función  $k(\omega)$  se puede aproximar en todo el intervalo de integración por una función lineal, la forma inicial se propaga a la velocidad  $v_g$  sin deformarse, porque en este caso la integral se puede aproximar por

$$\psi(x, t) \approx A_0 \frac{\sin(\frac{1}{2} \Delta\omega (t - x/v_g))}{\frac{1}{2} \Delta\omega (t - x/v_g)} \cos(\omega_p t - k_p x), \quad (4.66)$$

con  $k_p = k(\omega_p)$ ,  $v_g$  la velocidad de grupo definida en la ec. (4.39).



# Capítulo 5

## Respuesta forzada estacionaria

En los capítulos anteriores hemos aprendido a encontrar los modos normales de sistemas con muchas partes móviles y seguramente nos dimos cuenta que, más allá de la descripción matemática, la simetría del sistema juega un papel fundamental para descubrir los modos normales. La relación entre simetría y modos normales se puede estudiar sistemáticamente mediante la teoría de grupos, una rama de las matemáticas que trata sobre simetrías. Usando teoría de grupos (fuera del alcance de este curso) es posible encontrar los modos normales de sistemas con un gran número de partes móviles. Por ejemplo, es posible encontrar los modos normales de un cristal perfecto de silicio, uno de los materiales más importantes en la tecnología electrónica, resolviendo determinantes de  $6 \times 6$ , en vez de los determinantes monstruosos que habría que resolver solamente teniendo en cuenta el enorme número de átomos que componen un trozo macroscópico de cristal.

Otra manera de encontrar los modos normales de sistemas con muchas partes móviles es mediante el estudio de las resonancias que se observan en la respuesta forzada estacionaria. La respuesta forzada estacionaria es el movimiento que realiza el sistema cuando se lo somete a la acción de fuerzas exteriores (no solamente las fuerzas restitutivas que juegan en los modos) y cuando se han extinguido los movimientos transitorios que dependen de las condiciones iniciales. Desde el punto de vista matemático, el transitorio es la solución de ecuaciones de movimiento homogéneas, mientras que el estacionario es la solución particular cuando la fuerza impulsora es armónica.

La respuesta forzada estacionaria es notablemente intensa cuando la frecuencia de la fuerza impulsora coincide con la frecuencia de un modo normal, mientras que es poco intensa y casi despreciable en otros casos. Esto significa que las formas y las frecuencias de los modos de un sistema se revelan experimentalmente en los máximos de respuesta cuando el sistema es forzado mediante una fuerza impulsora de frecuencia variable. Y a partir de formas y frecuencias modales, es posible inferir propiedades y características físicas internas de sistemas que incluso pueden ser inaccesibles, como la atmósfera de un planeta lejano.

Notar que el movimiento colectivo de las partes móviles de un sistema en la respuesta forzada estacionaria tiene mucho en común con el movimiento colectivo de las partes móviles cuando el sistema se encuentra en un modo normal: en ambas situaciones las partes móviles oscilan con la misma frecuencia y pasan simultáneamente por la posición de equilibrio (tienen la misma fase).

Por este motivo, la descripción matemática de ambos movimientos es muy similar, aunque con varias diferencias, la más importante es que mientras que en la respuesta forzada estacionaria la frecuencia es un dato exterior conocido, impuesto por la fuerza impulsora, la frecuencia en un modo es una característica dinámica interna, impuesta por las fuerzas restitutivas. Para poner en evidencia la íntima relación que existe entre el problema de modos de cualquier sistema físico y la respuesta forzada estacionaria de dicho sistema, como así también las similitudes y diferencias entre ambos problemas, revisaremos situaciones familiares de cadenas lineales y sogas y también veremos un nuevo sistema con una nueva fuerza restitutiva.

## 5.1. Forzado estacionario, cadena lineal

Consideremos la respuesta forzada estacionaria de una cadena lineal con  $N$  cuentas idénticas equiespaciadas, como se esquematiza en la figura 3.1. Como ya estudiamos la dinámica de esta cadena para pequeños apartamientos transversales del equilibrio, sabemos que las ecuaciones linealizadas de movimiento cuando solamente actúan las fuerzas restitutivas internas vienen dadas por el sistema de ecuaciones diferenciales (3.7), con  $n = 1, 2, \dots, N$ . Si ahora forzáramos a la parte móvil 3 de este sistema mediante una fuerza impulsora armónica, entonces deberíamos retocar solamente la ecuación correspondiente a  $n = 3$ , el resto de ecuaciones quedaría sin cambios. Para no complicar demasiado, imaginemos que forzamos alguna de las partes móviles “virtuales”, por ejemplo el extremo izquierdo ( $n = 0$ ) o el extremo derecho ( $n = N$ ), así ni siquiera tenemos que retocar el sistema (3.7).

Y ahora nos damos cuenta que, si estamos interesados solamente en la respuesta forzada estacionaria, todo el trabajo ya está hecho en la sección 3.1.2. Esto es así porque en el estado forzado estacionario, al igual que en el modo normal, las funciones  $\psi_n(t)$  ( $n = 1, 2, \dots, N$ ) deben representar movimientos donde todas las partes móviles oscilan con la misma frecuencia. Mienstra que en la sección 3.1.2 esta frecuencia era la frecuencia del modo, que dependía del sistema, ahora la frecuencia es la de la fuerza impulsora, que resulta un parámetro externo. Pero más allá de esta diferencia, no queda otra posibilidad que estas funciones adopten la forma empleada en los tratamientos de la sección 3.1.2, es decir  $\psi_n(t) = A_n \cos(\omega t + \varphi)$  y por este motivo, volvemos a obtener las relaciones de recurrencia (3.8) o (3.9). Que ya sabemos que se resuelven con dependencias dadas por la ec. (3.10) y que conducen a relaciones de dispersión como (3.12) o su equivalente (3.13). Las mismas relaciones entre el período temporal asociado con  $\omega$  y la dependencia espacial asociada con  $\Theta$  que teníamos para los modos. Estas relaciones están representadas gráficamente en la figura 3.3.

### 5.1.1. Rango dispersivo

En el problema de modos, las condiciones de extremo discretizaban los valores de  $\Theta$ . Por ejemplo, en el caso de extremos fijos resultaban los valores reales  $\Theta = p\pi/(N+1)$ , con  $p$  entero y esto llevaba a valores  $0 < \omega_p < \omega_{Max}$ . En el caso de modos la existencia de una frecuencia máxima  $\omega_{Max}$  no nos había preocupado para nada. Y ahora, si la frecuencia  $\omega$  de la fuerza impulsora es menor que  $\omega_{Max}$ , tampoco nos preocupa, porque para todo valor  $\omega$  en el eje vertical de la figura 3.3, siempre será posible encontrar un valor de  $\Theta$  en el eje horizontal de

la misma figura y en virtud de la dependencia espacial dada por la ec. (3.10), este valor será interpretado físicamente como una periodicidad espacial de la respuesta forzada estacionaria de la cadena lineal, todo igual que antes, debo decir aunque parezca pesado.

Todo sistema físico donde la periodicidad temporal de la fuerza impulsora produce una periodicidad espacial de la respuesta forzada estacionaria se dice que se comporta dispersivamente. La conclusión es entonces que para fuerzas impulsoras con frecuencias en el intervalo  $0 < \omega < \omega_{Max}$ , la cadena lineal se comporta dispersivamente. Notar que la palabra *dispersivo* se usa con dos acepciones distintas: i) como en la sección 3.4, como sinónimo de que la velocidad de fase depende de la frecuencia, y ii) como ahora, como sinónimo de que una periodicidad temporal en un estímulo produce una periodicidad espacial en el comportamiento del sistema.

### 5.1.2. Rango reactivo

Pero si la frecuencia  $\omega$  de la fuerza impulsora es mayor que  $\omega_{Max}$ , entonces si que tenemos que salir de nuestra zona de confort. Que no está mal, porque la zona de confort en ciencia impide lograr grandes cosas. Así que ... bienvenidos a esta zona de incomodidad, en la que nos damos cuenta que cuando  $\omega > \omega_{Max}$  nunca será posible hallar valores de  $\Theta$  en el eje horizontal de la figura 3.3. Y como ya probamos que la dependencia espacial (3.10) funciona bien, porque según (3.11) el cociente  $\frac{A_{n+1}+A_{n-1}}{2A_n}$  no depende de  $n$ , tal como lo pide la física en la ec. (3.9), entonces tenemos que reconocer que quizás lo que no funciona es el prejuicio de pensar que, para toda frecuencia de fuerza impulsora, necesariamente tenemos que encontrar valores reales de  $\Theta$ . Que no es otra cosa que el prejuicio de pensar que los senos y cosenos están siempre acotados. Lo están para argumentos reales, y esto muestra que quizás tenemos que buscar en argumentos complejos. Por otro lado, hay un indicio muy importante que puede facilitarnos la búsqueda, y es que ya sabemos que en el límite continuo estas cosas raras no suceden, porque la relación de dispersión es lineal y  $\omega_{Max}$  convenientemente tiende a infinito y esta observación sugiere que nuestra zona de incomodidad desaparece en el límite continuo. Y ¿qué perdimos para hacer la aproximación continua? Recordemos que al hacer desarrollos de Taylor perdimos formas que no tienen bien definidas las derivadas espaciales de las ecuaciones de onda, como movimientos zig-zag.

Juntando los indicios de que los senos y cosenos de argumentos imaginarios no están acotados y de que en la aproximación continua perdimos movimientos zig-zag y observando que los senos y cosenos de argumentos complejos se entienden mejor cuando se escriben como combinación de funciones exponenciales, es entendible que para la zona  $\omega > \omega_{Max}$  haya que usar sucesiones de la forma

$$A_n = (-1)^n (b e^{-\Xi n} + c e^{\Xi n}), \quad (5.1)$$

donde  $\pm i \Xi$  juega el papel de  $\pm \Theta$  y el  $(-1)^n$  representa el zig-zag (o vaivén, para oscilaciones longitudinales) de partes móviles sucesivas. Para completar la demostración, hay que comprobar que con esta dependencia espacial, el cociente entre amplitudes  $\frac{A_{n+1}+A_{n-1}}{2A_n}$  de la ec. (3.9) ahora resulta igual a  $-\cosh \Xi$ . Y entonces, la relación de dispersión en esta zona se reescribe como

$$\omega = \omega_{Max} \cosh \frac{\Xi}{2}, \quad (5.2)$$

que siempre tiene soluciones  $\Xi$  reales para frecuencias  $\omega \geq \omega_{Max}$ . La ec. (5.1) muestra que en este rango de frecuencias, la periodicidad temporal de la fuerza impulsora NO produce una

periodicidad espacial de la respuesta forzada estacionaria, si no que produce una respuesta que es una combinación lineal de exponenciales crecientes y decrecientes y la magnitud  $a/\Xi$  representa la distancia típica en que la exponencial decreciente cambia a  $1/e$  su valor anterior (y la exponencial creciente aumenta a un valor igual a  $e$  veces su valor anterior). Veremos dentro de poco que las exponenciales crecientes no pueden aparecer en sistemas muy largos, de manera tal que si se fuerza la primera parte móvil a la izquierda, la perturbación no puede avanzar muy lejos y casi no se va a notar a varias distancias típicas.

En todo sistema físico donde la periodicidad temporal de la fuerza impulsora produce una respuesta forzada estacionaria con comportamientos espaciales exponenciales (por lo general decrecientes) se dice que el sistema se comporta reactivamente. La conclusión es entonces que para fuerzas impulsoras con frecuencias  $\omega > \omega_{Max}$ , la cadena lineal se comporta reactivamente.

Notar que un sistema no es dispersivo o reactivo, si no que se comporta de una u otra manera de acuerdo con el valor de la frecuencia de la fuerza impulsora. También notemos que en el caso de la cadena lineal, el comportamiento reactivo desaparece en la aproximación continua. En la siguiente sección veremos otros sistema de partes móviles que en la aproximación continua exhibe los dos comportamientos, dispersivo y reactivo.

## 5.2. Forzado estacionario, Klein-Gordon

En vez de buscar la solución más general de (3.41) es instructivo estudiar cómo responde el sistema cuando: i) se obliga a que el primer péndulo se mueva con un movimiento oscilatorio armónico de frecuencia  $\omega$  (con una fuerza externa, por ejemplo con un motor) y ii) se espera el tiempo suficiente como para que se apague el transitorio. Con la hipótesis ii) estamos repitiendo el esquema realizado en 5.1, de suponer que hay una pequeña disipación que no afecta el fenómeno que queremos estudiar (y que por eso despreciamos) y que esa pequeña disipación es necesaria para apagar el transitorio (y por eso no la despreciamos). Luego de pensar en situaciones prácticas en las que esta ambivalencia puede resultar una buena aproximación, retomamos la idea de la sección anterior, de aprovechar para la respuesta forzada estacionaria las cosas que sabemos de modos normales. Y así seguir los pasos que hicimos en la página 64 para estudiar los modos normales de una cuerda. De esta manera, proponemos una solución como en (3.45), reemplazamos esta expresión en (3.41) y así llegamos la siguiente ecuación diferencial ordinaria para encontrar la función  $A(x)$

$$\frac{d^2 A}{dx^2} + \underbrace{\frac{M}{Ka^2} (\omega^2 - \omega_0^2)}_{k^2} A = 0. \quad (5.3)$$

Como esta ecuación es formalmente idéntica a la ec. (3.46), su solución debe ser formalmente idéntica a la dada en la ec. (3.47)

$$A(x) = B \sin(kx) + C \cos(kx), \quad (5.4)$$

donde  $B$  y  $C$  son constantes. La **diferencia fundamental** entre el modelo de la cuerda y el modelo de los péndulos acoplados aparece en la ecuación de dispersión que relaciona  $k$  con  $\omega$ :

en el primer caso la relación de dispersión es lineal y viene dada por la ec. (3.48), mientras que en el segundo caso la relación de dispersión es

$$k = \sqrt{\frac{M}{Ka^2}} \sqrt{\omega^2 - \omega_0^2} \quad (5.5)$$

que claramente no es lineal.

### 5.2.1. Dispersivo vs reactivo

Aparte de la no linealidad, la nueva relación de dispersión muestra la existencia de dos regímenes de frecuencia donde los comportamientos físicos son muy distintos:

- **zona dispersiva**,  $\omega > \omega_0$ ,  $k$  es una cantidad real,  $A(x) = A(x + 2\pi/k)$ , la periodicidad temporal de la fuerza impulsora induce una periodicidad espacial  $\lambda = 2\pi/k$  en las amplitudes de las partes móviles.
- **zona reactiva**,  $\omega < \omega_0$ ,  $k = i\kappa$  es una cantidad imaginaria pura, con

$$\kappa = \sqrt{\frac{M}{Ka^2}} \sqrt{\omega_0^2 - \omega^2}, \quad (5.6)$$

los senos y cosenos en (5.4) pasan a ser senos y cosenos hiperbólicos y es conveniente reescribir  $A(x)$  como una combinación lineal de funciones exponenciales reales

$$e^{-\kappa x} \quad y \quad e^{\kappa x}. \quad (5.7)$$

En este caso, la periodicidad temporal de la fuerza impulsora NO induce una periodicidad espacial en las amplitudes de las partes móviles. La cantidad  $\delta = 1/\kappa$  es la distancia en la cual la exponencial decreciente cambia a  $1/e$  su valor anterior y se conoce como *distancia de penetración* (también es la distancia en la cual la exponencial creciente aumenta a un valor igual a  $e$  veces su valor anterior).

Conclusión: la respuesta forzada de este sistema depende dramáticamente de la frecuencia de la fuerza impulsora. Cuando la periodicidad temporal de una excitación externa (como en este caso la fuerza impulsora, pero también podrían ser los campos electromagnéticos de una antena emisora, o el voltaje alterno de 50 Hz que sale del enchufe eléctrico) induce una periodicidad espacial, se dice que el medio excitado se comporta dispersivamente. En cambio cuando la periodicidad temporal de la excitación externa induce comportamientos exponenciales, se dice que el medio excitado se comporta reactivamente. Un medio no es ni reactivo ni dispersivo, sino que su comportamiento es de un tipo u otro dependiendo de la frecuencia de la excitación externa. Si  $\omega = \omega_0$ ,  $k = 0$ ,  $\lambda = \infty$ : todos los péndulos se balancean en fase y la fuerza restauradora es solamente gravitatoria.

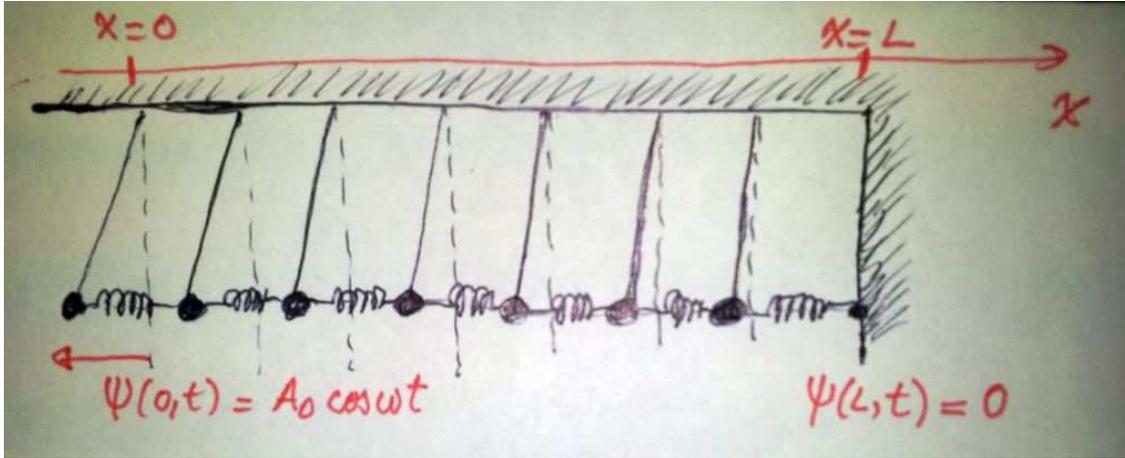


Figura 5.1: El sistema de péndulos idénticos se extiende desde  $x = 0$  hasta  $x = L$  y el último resorte está sujeto a una pared rígida.

### 5.2.2. La parte con $e^{\kappa x}$ no está permitida en sistemas infinitos

El siguiente ejemplo, esquematizado en la figura 5.1, muestra que en el régimen reactivo, cuanto más grande es el sistema, menos influye la exponencial creciente en la combinación lineal de funciones exponenciales reales. El primer péndulo es forzado con un movimiento oscilatorio armónico de frecuencia  $\omega$  y describe el movimiento  $\psi(0, t) = A_0 \cos \omega t$ . El último resorte está sujeto a una pared rígida en  $x = L$ , por lo tanto  $\psi(L, t) = 0$  siempre. Si la frecuencia de la fuerza impulsora es menor que la frecuencia característica  $\omega_0$ , en el régimen estacionario las amplitudes  $A(x)$  tienen la forma

$$A(x) = b e^{-\kappa x} + c e^{\kappa x}. \quad (5.8)$$

Las constantes  $b$  y  $c$  quedan determinadas cuando se ajusta (5.8) para que cumpla las condiciones en los dos extremos

$$\begin{aligned} b + c &= A_0, \\ b e^{-\kappa L} + c e^{\kappa L} &= 0 \implies c = -b e^{-2\kappa L} \xrightarrow{\kappa L \gg 1} 0. \end{aligned} \quad (5.9)$$

La segunda ecuación muestra que si el sistema mide varias distancias de penetración ( $L \gg \delta$ )  $\implies |c| \ll |b|$ . Para tener idea de órdenes de magnitud, si  $L/\delta = 2$ ,  $|c/b| \approx 0.14$  y si  $L/\delta = 5$ ,  $|c/b| \approx 0.007$ . Resolviendo el sistema (5.9), resulta

$$\psi(x, t) = \frac{A_0}{1 - e^{-2\kappa L}} \left[ e^{-\kappa x} - e^{-\kappa(2L-x)} \right] \cos \omega t. \quad (5.10)$$

Si  $L \gg \delta$ ,  $\psi(x, t) \approx A_0 e^{-\kappa x} \cos \omega t$  y se tiene una perturbación armónica con decaimiento espacial exponencial conocida como *onda evanescente*. En este caso resulta que en el régimen reactivo la fuerza impulsora es incapaz de mover partes móviles alejadas a distancias  $x \gg \delta$ . En este caso da lo mismo que el último resorte esté sujeto a una pared rígida o no, porque la perturbación de la fuerza impulsora no puede ser transmitida por el sistema.

### 5.2.3. Comentarios sobre el régimen reactivo

El régimen reactivo tiene aplicaciones en *filtros*: toda fuerza impulsora no armónica, Fourier mediante, se puede pensar como suma de funciones armónicas. Este sistema deja pasar solamente aquellas componentes de Fourier con frecuencias mayores que  $\omega_0$ , en cambio “rebotará” componentes de Fourier con frecuencias menores que  $\omega_0$ . Es lo que sucede en la ionósfera cuando se la intenta forzar electromagnéticamente. La ionósfera es la parte de la atmósfera terrestre ionizada permanentemente debido a la radiación solar. Y resulta que los campos electromagnéticos satisfacen en la ionósfera la ecuación de Klein-Gordon (3.41),  $\omega_0$  es la *frecuencia del plasma* y el factor que multiplica a la derivada segunda espacial es igual a la velocidad de la luz en el vacío al cuadrado. Para un plasma diurno la frecuencia del plasma es  $\nu_p \approx 10 - 30$  MHz, por eso la ionósfera se comporta reactivamente para ondas de radio AM, con frecuencias  $\nu \approx 1000$  KHz, mientras que se comporta dispersivamente para ondas de radio FM, con frecuencias  $\nu \approx 100$  MHz. ¿Cómo se comporta la ionósfera terrestre para la luz?

Supongamos que a partir de la posición  $x_1$  se produce un cambio en la longitud de los péndulos: a la izquierda de  $x_1$  con longitud  $\ell_1$  y a la derecha con longitud  $\ell_2$ . La respuesta forzada estacionaria en cada tramo dependerá del valor de la frecuencia externa con respecto a  $\omega_{01} = \sqrt{g/\ell_1}$  y  $\omega_{02} = \sqrt{g/\ell_2}$ . Si  $\ell_2 < \ell_1$ ,  $\omega_{01} < \omega_{02}$  (como en el sistema de péndulos del ejercicio 16 de la guía 1, que sería la versión discreta de lo que estamos haciendo acá). Si la frecuencia de la fuerza impulsora se elige para que  $\omega_{01} < \omega < \omega_{02}$ , la perturbación periódica causada por la fuente en el tramo izquierdo se transformará, a partir de  $x_1$ , en una perturbación con decaimiento exponencial. Como no hay disipación y debido a que las perturbaciones con decaimiento exponencial no pueden entrar mucho en la zona  $x > x_1$ , la perturbación incidente “rebotará” y como resultado se formará una onda estacionaria en la zona  $x < x_1$ , lo mismo que pasaba en un segmento finito de cuerda (independientemente de las condiciones en los extremos la energía se queda encerrada en el sistema). Comparar la solución del ejercicio 16 de la guía 1 con la solución que se obtendría en este caso en el continuo.

Si a partir de la posición  $x_2 > x_1$  la longitud de los péndulos fuera de nuevo  $\ell_1$ , la región  $x > x_2$  sería una zona con comportamiento dispersivo y en esta zona podría volver a aparecer una perturbación periódica, como la que había en el tramo izquierdo. Para que la perturbación en este tercer tramo sea apreciable se debería cumplir que  $x_2 - x_1 \lesssim \delta$ . Hay muchos fenómenos físicos en los que una perturbación se transmite a través de regiones “prohibidas” (como en el tramo entre  $x_1$  y  $x_2$ , que no puede ser traspasado si  $x_2 - x_1 \gg \delta$ ). Cuando esto ocurre se dice que hay **tuneleo** de la perturbación, un efecto que aparece en óptica, en acústica y de manera espectacular en mecánica cuántica.

## 5.3. Forzado estacionario: péndulos acoplados

Para analizar el forzado estacionario del sistema discreto, partimos de (3.25) (la que nos llevó a Klein-Gordon). Ya dijimos que la solución procede formalmente igual que la de los modos. Aprovechemos lo que se hizo en la eq. (3.7) para el caso  $\omega_0 = 0$ , página 48. Es claro que si antes

llegábamos a esta eq. en diferencias (ver (3.8))

$$\omega^2 = 2 \frac{K}{M} \left( 1 - \frac{A_{n+1} + A_{n-1}}{2A_n} \right), \quad (5.11)$$

ahora vamos a llegar a una ecuación en diferencias similar, pero con el nuevo término que corresponde a  $\omega_0 \neq 0$

$$\omega^2 = \omega_0^2 + 2 \frac{K}{M} \left( 1 - \frac{A_{n+1} + A_{n-1}}{2A_n} \right), \quad (5.12)$$

que se puede interpretar recordando que el lado derecho de (5.12) da la fuerza restitutiva por unidad de desplazamiento y por unidad de masa cuando  $\omega_0 = 0$  (no hay gravedad o péndulos muy largos). Cuando  $\omega_0 \neq 0$ , la fuerza restitutiva aumenta de la misma manera para todas las masas y se suma a la la fuerza restitutiva por unidad de desplazamiento y por unidad de masa que ya tenían las partes móviles sin gravedad.

Ya sabemos que el cociente entre amplitudes se puede hacer independiente de  $n$  con sucesiones de la forma

$$A_n = B \sin(n\Theta) + C \cos(n\Theta). \quad (5.13)$$

Así llegamos a la nueva relación de dispersión (ver eq. (3.13))

$$\omega^2 = \omega_0^2 + \underbrace{\frac{4K}{M}}_{\omega_{MAX}^2} \sin^2 \frac{\Theta}{2}, \quad (5.14)$$

válida para la relación entre período temporal y período espacial, tanto en un modo como en el caso forzado estacionario.

Cuando  $\omega = \omega_0$ ,  $\Theta = 0$ , todos los péndulos se balancean en fase (es decir, período espacial infinito, resultado ya obtenido en la aproximación continua, página 107) y la fuerza restitutiva por unidad de desplazamiento y por unidad de masa es solamente gravitatoria.

Cuando  $\omega_0^2 \leq \omega^2 \leq \omega_0^2 + \omega_{MAX}^2$  siempre es posible encontrar un valor real de  $\Theta$  que estará asociado con una periodicidad espacial. Vemos entonces que el rango dispersivo en el caso discreto tiene un límite superior

$$\sqrt{\omega_0^2 + \frac{4K}{M}} = \sqrt{\omega_0^2 + \frac{1}{a^2} \frac{4K a^2}{M}} \xrightarrow{a \rightarrow 0} \infty \quad (5.15)$$

y que este límite superior tiende a infinito cuando nos aproximamos al caso continuo, resultado que ya habíamos obtenido con Klein-Gordon.

**Rango reactivo bajo.** Sería de esperar que la zona  $\omega < \omega_0$  corresponda a un comportamiento reactivo. Y así es nomás. Una manera de comprobarlo es notar que la ec. (5.14) se satisface para valores imaginarios puros de  $\Theta = i\Xi$ , con  $\Xi$  real. En este caso los senos y cosenos

de la ec. (5.13) se transforman en senos y cosenos hiperbólicos, el cociente entre amplitudes de la ec. (5.12) resulta igual a  $\cosh \Xi$  y la relación de dispersión en esta zona se reescribe como

$$\omega^2 = \omega_0^2 - \frac{4K}{M} \sinh^2 \frac{\Xi}{2}. \quad (5.16)$$

Notar que en este rango tiene más sentido físico reescribir la ec. (5.13) en términos de combinaciones lineales de exponenciales crecientes y decrecientes del tipo  $e^{\pm \Xi n}$

$$A_n = b e^{-\Xi n} + c e^{\Xi n}. \quad (5.17)$$

Cuando  $\omega = \omega_0$  ya sabemos que todos los péndulos se balancean en fase, un resultado contenido en esta ecuación porque para esta frecuencia  $\Xi = 0$ , que corresponde a una distancia de penetración infinita. Cuando  $\omega$  disminuye desde el valor  $\omega_0$ ,  $\Xi$  aumenta, es decir que la distancia de penetración disminuye, hasta llegar a un valor máximo de  $\Xi$  cuando  $\omega \rightarrow 0$ . Nos damos cuenta que  $\Xi$  juega el papel de  $\kappa a$ .

**Rango reactivo alto.** Ya nos dimos cuenta que la zona  $\omega^2 \geq \omega_0^2 + \omega_{MAX}^2$  no corresponde a comportamientos dispersivos. Y también vimos que esta zona desaparece en el caso continuo. Nos damos cuenta que pasar algo muy parecido al caso  $\omega_0 = 0$  que vimos para las oscilaciones transversales de la cadena lineal, aunque en vez de zig-zag deberíamos decir vaivén, porque los péndulos realizan movimientos longitudinales. Es claro que las sucesiones de la forma de la ec. (5.17) no sirven, porque dan frecuencias fuera del rango que nos interesa. Pero sí a través de un  $(-1)^n$  intentamos con un vaivén entre partes móviles sucesivas

$$A_n = (-1)^n (b e^{-\Xi n} + c e^{\Xi n}), \quad (5.18)$$

es fácil comprobar que ahora el cociente entre amplitudes de la ec. (5.12) resulta igual a  $-\cosh \Xi$  y la relación de dispersión en esta zona se reescribe como

$$\omega^2 = \omega_0^2 + \frac{4K}{M} \cosh^2 \frac{\Xi}{2}. \quad (5.19)$$

que siempre tiene soluciones  $\Xi$  reales para frecuencias  $\omega \geq \sqrt{\omega_0^2 + \omega_{MAX}^2}$ .

## 5.4. Cadena lineal diatómica

En la figura 5.2 se muestra la posición de equilibrio de un sistema de cuentas que consiste en una cadena lineal de masas que interactúan mediante fuerzas elásticas con sus vecinos cercanos. Es un sistema muy parecido al considerado en la figura 3.1, con la diferencia de que ahora las cuentas no son idénticas ni están equiespaciadas. Aunque el sistema sigue siendo periódico, con período  $a$ , en este caso hay dos masas ( $m_1$  y  $m_2$ ) en cada período. Como ya sabemos que el tratamiento es el mismo, haremos simultáneamente el tratamiento matemático de modos y del forzado estacionario. Para los modos, y tal como hicimos en la sección 3.1, podemos limitar el sistema y suponer masas virtuales en  $x = 0$  y en  $x = (N+1)a = L$ . Para el forzado estacionario, y si estamos interesados en la propagación de una perturbación armónica inducida desde  $x = 0$ , podemos considerar un sistema infinito (donde solamente interesará la parte progresiva de las

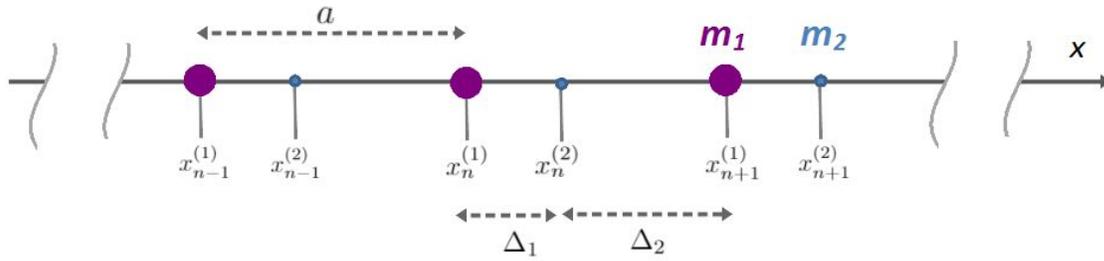


Figura 5.2: Cadena compuesta por una secuencia periódica de dos cuentas con masas  $m_1$  y  $m_2$ . Como en la figura 3.1 el período es  $a$ , pero las cuentas no están equiespaciadas

ondas) o finito (donde necesariamente habrá que considerar, como en los modos estacionarios, tanto las partes progresiva como regresivas, o las exponenciales crecientes y decrecientes, si hubiera comportamiento reactivo). Para la porción genérica  $n$ -ésima (la  $n$ -ésima *celda unitaria*), las posiciones de las masas  $m_1$  y  $m_2$  vienen dadas por  $x_n^{(1)} = na$  y  $x_{n-1}^{(2)} + \Delta_1 = x_n^{(1)} + \Delta_1$ . Notar que para  $m_1 = m_2$  y  $\Delta_1 = \Delta_2 = a/2$  el sistema se reduce al considerado en la sección 3.1. Las posiciones de las masas en las celdas unitarias adyacentes a derecha e izquierda vienen dadas por  $x_{n\pm 1}^{(1)} = x_n^{(1)} \pm a$  y  $x_{n\pm 1}^{(2)} = x_n^{(2)} \pm a$ .

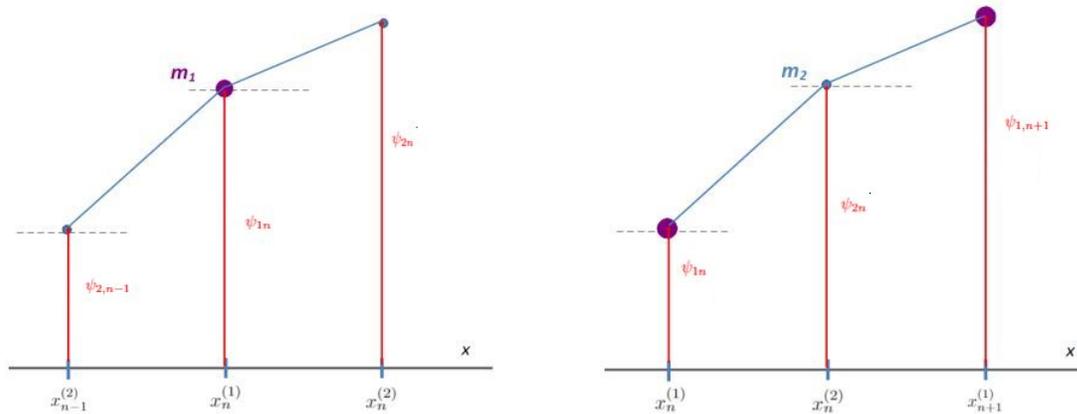


Figura 5.3: Desplazamientos transversales de la cadena periódica con cuentas de masas  $m_1$  y  $m_2$  en cada celda unitaria.

### 5.4.1. Ecuaciones de movimiento

Fuera del equilibrio, en cada celda unitaria representamos con  $\psi_{1n}(t)$  y  $\psi_{2n}(t)$  a los desplazamientos transversales de las cuentas con masas  $m_1$  y  $m_2$ . La situación está bosquejada en la figura 5.3 y es completamente similar a la considerada en la figura 3.2. Observar que tal como en el caso de la cadena homogénea, tenemos el mismo tratamiento para estudiar pequeños apartamientos transversales del equilibrio en el caso de i) cuentas engarzadas en una cuerda elástica sin masa o ii) unidas por resortes muy estirables. El sistema de ecuaciones diferenciales homogéneo linealizado para cada masa con dos vecinos cercanos tiene la misma forma que la

ec. (3.5), solo hay que tener cuidado y plantear una ecuación de Newton para las masas pares y otra para las impares, cambiando los valores de las masas y de las proyecciones de las tensiones en la dirección vertical.

Observando la figura 5.3 y repitiendo el procedimiento realizado en (3.5), obtenemos el siguiente sistema acoplado de ecuaciones diferenciales

$$\ddot{\psi}_{1n} = \frac{T_0}{m_1} \frac{\psi_{2n} - \psi_{1n}}{\Delta_1} - \frac{T_0}{m_1} \frac{\psi_{1n} - \psi_{2,n-1}}{\Delta_2}, \quad (5.20)$$

$$\ddot{\psi}_{2n} = \frac{T_0}{m_2} \frac{\psi_{1,n+1} - \psi_{2n}}{\Delta_2} - \frac{T_0}{m_2} \frac{\psi_{2n} - \psi_{1n}}{\Delta_1},$$

que se puede reescribir como

$$m_1 \ddot{\psi}_{1n} = f \psi_{2n} - (f + g) \psi_{1n} + g \psi_{2,n-1}, \quad (5.21)$$

$$m_2 \ddot{\psi}_{2n} = f \psi_{1n} - (f + g) \psi_{2n} + g \psi_{1,n+1}, \quad (5.22)$$

donde  $f = T_0/\Delta_1$  y  $g = T_0/\Delta_2$ . Tanto en modos como en forzado estacionario debe cumplirse que  $\psi_{jn}(t) = A_{jn} \cos(\omega t + \varphi)$ , con  $j=1, 2$ , o de manera más explícita

$$\psi_{1n}(t) = A_{1n} \cos(\omega t + \varphi),$$

$$\psi_{2n}(t) = A_{2n} \cos(\omega t + \varphi).$$

Introduciendo estas expresiones en el sistema (5.21)-(5.22), simplificando los cosenos y tal como ocurría en (3.8), se obtienen relaciones de recurrencia para las amplitudes de las partes móviles

$$f A_{1,2} - [(f + g) - m_1 \omega^2] A_{1,n} + g A_{2,n-1} = 0, \quad (5.23)$$

$$f A_{1,n} - [(f + g) - m_2 \omega^2] A_{1,2} + g A_{1,n+1} = 0, \quad (5.24)$$

que se pueden reescribir de manera parecida a la eq. (3.9)

$$\frac{f A_{1,2} + g A_{2,n-1}}{(f + g) A_{1,n}} = 1 - \frac{m_1}{(f + g)} \omega^2, \quad (5.25)$$

$$\frac{f A_{1,n} + g A_{1,n+1}}{(f + g) A_{1,2}} = 1 - \frac{m_2}{(f + g)} \omega^2, \quad (5.26)$$

y que también dice que la dependencia adecuada es aquella para la cual ciertos cocientes que involucran los términos de la sucesión no dependen de  $n$ . Y así de nuevo llegamos a que la solución de estas ecuaciones en recurrencias tienen que tener la forma general propuesta en (3.10), tanto para los términos pares como para los impares

$$A_{1n} = B^{(1)} \sin(n \Theta) + C^{(1)} \cos(n \Theta), \quad (5.27)$$

$$A_{2n} = B^{(2)} \sin(n \Theta) + C^{(2)} \cos(n \Theta), \quad (5.28)$$

con  $B^{(1)}$ ,  $C^{(1)}$ ,  $B^{(2)}$  y  $C^{(2)}$  cuatro constantes a determinar, a menos de una constante porque ya se puede ver que obtendremos un sistema de ecuaciones algebraicas homogéneo, y  $\Theta(\omega)$ , *la relación de dispersión del sistema diatómico*, una función a determinar con la condición que dicho sistema homogéneo tenga solución no trivial.

### 5.4.2. Relación de dispersión

Cada una de las combinaciones (5.27) y (5.28) en términos de senos y cosenos son equivalentes a una combinación lineal de exponenciales imaginarias

$$e^{in\Theta} \quad \text{y} \quad e^{-in\Theta}$$

y para simplificar las cuentas, vamos a probar primero que cada exponencial por separado resuelve las relaciones de recurrencia y conduce a las mismas relaciones de dispersión. Luego, por linealidad de las relaciones, quedarán demostradas las propuestas (5.27) y (5.28). Empecemos entonces con la primera exponencial

$$A_{1n} = u_1 e^{in\Theta} \quad \text{y} \quad A_{2n} = u_2 e^{in\Theta}. \quad (5.29)$$

Introduciendo esta propuesta en las ecs. (5.23) y (5.24), y simplificando el factor común  $e^{in\Theta}$ , llegamos a un sistema algebraico homogéneo para  $u_1$  y  $u_2$

$$[m_1 \omega^2 - (f + g)] u_1 + [f + g e^{-i\Theta}] u_2 = 0, \quad (5.30)$$

$$[f + g e^{i\Theta}] u_1 + [m_2 \omega^2 - (f + g)] u_2 = 0. \quad (5.31)$$

La relación de dispersión del sistema se obtiene pidiendo que el determinante de este sistema sea nulo, es decir

$$[m_1 \omega^2 - (f + g)][m_2 \omega^2 - (f + g)] - [f + g e^{-i\Theta}][f + g e^{i\Theta}] = 0$$

dividiendo por el producto de las masas

$$\left[\omega^2 - \frac{f + g}{m_1}\right] \left[\omega^2 - \frac{f + g}{m_2}\right] - \frac{1}{m_1 m_2} [f + g e^{-i\Theta}][f + g e^{i\Theta}] = 0$$

haciendo los productos

$$\omega^4 + \frac{(f + g)^2}{m_1 m_2} - \omega^2 (f + g) \left[\frac{1}{m_1} + \frac{1}{m_2}\right] - \frac{1}{m_1 m_2} [f^2 + g^2 + 2fg \cos \Theta] = 0$$

juntando los términos independientes de  $\omega$

$$\omega^4 - \omega^2 (f + g) \left[\frac{1}{m_1} + \frac{1}{m_2}\right] + \frac{2fg}{m_1 m_2} [1 - \cos \Theta] = 0.$$

Usando una identidad trigonométrica e introduciendo la *masa reducida*  $\mu = m_1 m_2 / (m_1 + m_2)$  queda la siguiente ecuación bicuadrática para la frecuencia  $\omega$

$$\omega^4 - \omega^2 \frac{f + g}{\mu} + \frac{4fg}{m_1 m_2} \sin^2 \frac{\Theta}{2} = 0$$

cuyas soluciones son

$$\omega^2 = \frac{f + g}{2\mu} \left[ 1 \pm \left( 1 - \frac{16fg}{(f + g)^2} \frac{\mu^2}{m_1 m_2} \sin^2 \frac{\Theta}{2} \right)^{1/2} \right]. \quad (5.32)$$

La ecuación (5.32) muestra una de las características más importantes que distingue a este sistema de los vistos hasta el momento y que consiste en que su relación de dispersión tiene dos ramas que definen dos rangos dispersivos. Una rama,  $\omega_1(\Theta)$ , corresponde al signo “-” en (5.32) y se conoce como *rama acústica*, mientras que la otra rama,  $\omega_2(\Theta)$ , corresponde al signo “+” en (5.32) y se conoce como *rama óptica*. Entre ambas ramas, y para valores de frecuencias en el intervalo comprendido en una *banda* definida por el máximo de  $\omega_1(\Theta)$  y el mínimo de  $\omega_2(\Theta)$ , aparece un rango reactivo, además del rango reactivo alto que había aparecido tanto en la cadena de partículas idénticas como en el sistema de péndulos acoplados, para frecuencias por encima de la rama más alta. Este tipo de *bandas prohibidas* (en el sentido de que las ondas con frecuencias en la banda solo pueden decaer exponencialmente a lo largo del sistema infinito) se conocen también por su nombre en inglés *band gaps*. Juegan un papel muy importante en otras fenomenologías ondulatorias análogas al ejemplo mecánico considerado y explica, por ejemplo, el comportamiento de los electrones en los semiconductores, de las ondas electromagnéticas en los cristales fotónicos o de las ondas sonoras en los cristales sónicos (o fonónicos).

También cabe destacar, como verán cuando estudien Estructura de la Materia 2, que muchos de los resultados derivados en esta sección para la cadena diatómica se pueden aplicar formalmente a cristales tridimensionales simples como los haluros alcalinos, por ejemplo el NaCl. Si en estos cristales las ondas elásticas se propagan en direcciones de alta simetría, el problema se vuelve unidimensional y puede tratarse formalmente con el modelo de la cadena diatómica. El nombre de rama acústica se debe al hecho de que en los sólidos si la longitud de onda es grande,  $2\pi/\Theta \gg 1$ , una perturbación de esta rama se propaga como una onda de sonido en un continuo elástico. Por otra parte, el nombre rama óptica proviene del hecho de que los modos con frecuencias  $\omega_2$  pueden ser excitados por campos electromagnéticos en el infrarrojo o el infrarrojo lejano. De manera general se puede decir que en una cadena unidimensional compuesta por  $J$  tipos de átomos tendremos  $J$  ramas, una acústica que pasará por el punto  $(\omega, \Theta) = (0, 0)$  y  $J - 1$  ramas ópticas.

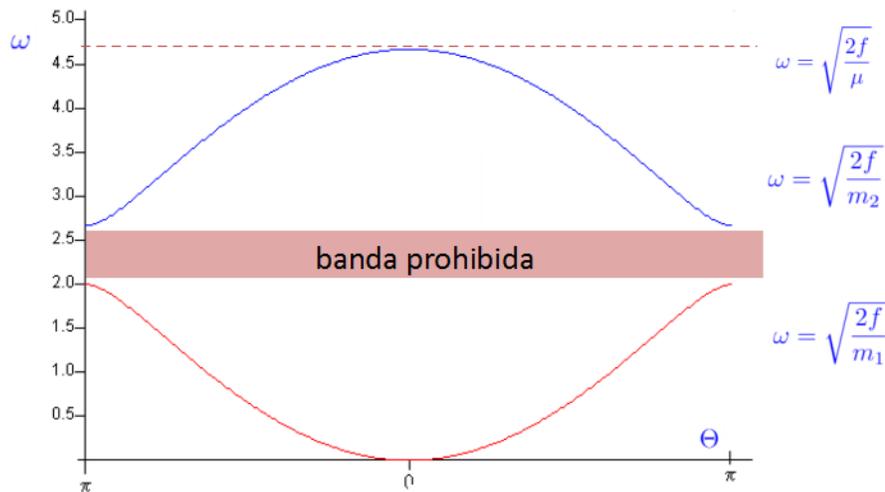


Figura 5.4: Relación de dispersión para la cadena periódica con cuentas de masas  $m_1 = 1\text{kg}$ ,  $m_2 = 0.75\text{kg}$ ,  $f = 1\text{N/m}$ .

Notar que la ec. (5.32) es invariante a intercambios  $m_1 \rightarrow m_2$  y  $f_1 \rightarrow f_2$  y que también permanece invariante cuando se cambia  $\Theta$  por  $-\Theta$ . Observamos que la función  $e^{-in\Theta}$  también

es solución de las relaciones de recurrencia, con la misma relación de dispersión y todos los resultados intermedios que hemos desarrollado para la propuesta (5.29) siguen valiendo, teniendo la precaución de hacer el reemplazo de  $\Theta$  por  $-\Theta$ .

### 5.4.3. Caso particular $f = g$

Para simplificar cuentas, consideremos el caso particular  $\Delta_1 = \Delta_2 = a/2$  (masas distintas equiespaciadas), es decir  $f = g = 2T_0/a$  (el caso de la molécula de ClNa). En este caso, las dos ramas son

$$\omega^2 = \frac{f}{\mu} \left[ 1 \pm \left( 1 - 4 \frac{\mu^2}{m_1 m_2} \sin^2 \frac{\Theta}{2} \right)^{1/2} \right]. \quad (5.33)$$

Notar que  $4m_1 m_2 \leq (m_1 + m_2)^2$ . Para  $m_1 > m_2$ , la rama acústica tiene su valor mínimo  $\omega = 0$  en  $\Theta = 0$  y su valor máximo  $\omega = (\frac{2f}{m_1})^{(1/2)}$  en  $\Theta = \pi$ . En cambio, la rama óptica tiene su valor mínimo  $\omega = (\frac{2f}{m_2})^{(1/2)}$  en  $\Theta = \pi$  y su valor máximo  $\omega = (\frac{2f}{\mu})^{(1/2)}$  en  $\Theta = 0$ . La *banda prohibida* se extiende para valores de frecuencia en la región

$$\sqrt{\frac{2f}{m_1}} \leq \omega \leq \sqrt{\frac{2f}{m_2}}$$

y tiene un ancho proporcional a  $1/\sqrt{m_2} - 1/\sqrt{m_1}$ . El gráfico de la figura muestra que no existen valores reales de  $\Theta$  para frecuencias en la banda prohibida. Al igual que en el régimen reactivo de otros sistemas, como la cadena lineal o los péndulos acoplados, en esta zona la relación de dispersión (5.33) da valores reales de  $\omega^2$  solamente cuando  $\Theta$  es un número complejo con parte imaginaria distinta de cero y esto indica que el comportamiento del sistema en esta zona es reactivo.

Notar que la relación de dispersión se obtuvo a partir del sistema de ecuaciones (5.30) y (5.31) y que este sistema se puede escribir en la forma de una ecuación de autovalores. Los autovalores son las ramas acústica y óptica y los autovectores son las amplitudes  $u_1$  y  $u_2$ . En el caso particular que estamos considerando,  $f = g = 2T_0/a$ , las ecuaciones (5.30) y (5.31) se pueden reescribir como

$$m_1 \left[ \omega^2 - \frac{2f}{m_1} \right] u_1 + f [1 + e^{-i\Theta}] u_2 = 0, \quad (5.34)$$

$$f [1 + e^{i\Theta}] u_1 + m_2 \left[ \omega^2 - \frac{2f}{m_2} \right] u_2 = 0. \quad (5.35)$$

De la ec. (5.34) se obtiene que el espacio de los autovectores está definido mediante el cociente

$$\frac{u_1}{u_2} = -\frac{f}{m_1} \frac{1 + e^{-i\Theta}}{\omega^2 - \frac{2f}{m_1}}, \quad (5.36)$$

condición que, para frecuencias que satisfacen la relación de dispersión, debe ser equivalente a la que se obtiene a partir de la ec. (5.35), ya que cuando se satisface la relación de dispersión

el sistema (5.30) y (5.31) deja de ser independiente y ambas ecuaciones deben dar los mismos resultados, aunque no lo parezca

$$\frac{u_1}{u_2} = -\frac{m_2}{f} \frac{\omega^2 - \frac{2f}{m_2}}{1 + e^{i\Theta}}, \quad (5.37)$$

En la rama acústica  $\omega^2 < 2f/m_1$ . Y cuando  $\Theta \approx 0$ , las dos masas de una celda oscilan en fase y esto se repite para todas las celdas. En cambio, en la misma rama pero cuando  $\Theta \approx \pi$ , las dos masas de una celda unitaria también oscilan en fase, pero la fase cambia de signo de una celda a la siguiente.

En la rama óptica  $\omega^2 > 2f/m_2$ . Y cuando  $\Theta \approx 0$ , ambas masas de una celda unitaria oscilan en contrafase y esto se repite para todas las celdas. En cambio, cuando  $\Theta \approx \pi$ , las dos masas de una celda unitaria también oscilan en contrafase, pero la fase cambia de signo de una celda a la siguiente. Para una cadena de átomos distintos, que las dos partículas de una misma celda oscilen en contrafase produce un momento dipolar oscilante que se puede acoplar de manera resonante con radiación electromagnética en el rango del infrarrojo y en definitiva da lugar a que el material absorba dicha radiación.

#### 5.4.4. Modos para extremos fijos

Veamos cómo usar los pasos teóricos anteriores para encontrar los modos (transversales o longitudinales) de una cadena diatómica de longitud finita con los extremos fijos. En primer lugar tomemos un número impar  $(2N + 1)$  de partes móviles, es decir que los extremos fijos, al igual que las partes móviles más próximas a los extremos, corresponderían a masas del mismo tipo (en la cadena hay N celdas elementales y media).

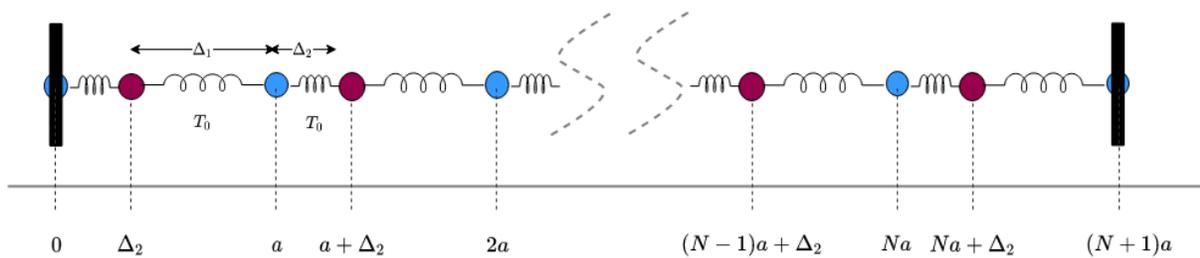


Figura 5.5: Esquema de cadena diatómica con un número impar  $(2N + 1)$  de partes móviles (N celdas elementales y media).

Las amplitudes  $A_{1,n}$  y  $A_{1,2}$  de las ecs. (5.27) y (5.28) se tienen que escribir como combinación lineal de los dos tipos de exponenciales imaginarias, positivas y negativas. Por razones matemáticas y porque físicamente los modos son ondas estacionarias, que son combinación de ondas progresivas y regresivas. Entonces, hay que modificar las expresiones (5.29) y escribir las amplitudes como combinaciones lineales de las funciones  $e^{\pm i n \Theta}$

$$A_{1,n} = u_1(\Theta) e^{in\Theta} + u_1(-\Theta) e^{-in\Theta}, \quad (5.38)$$

$$A_{2,n} = u_2(\Theta) e^{in\Theta} + u_2(-\Theta) e^{-in\Theta}, \quad (5.39)$$

donde  $u_1(\Theta)$  y  $u_2(\Theta)$  están relacionadas mediante la expresión (5.36), o su equivalente (5.37), mientras que  $u_1(-\Theta)$  y  $u_2(-\Theta)$  están relacionadas mediante las expresiones que se obtienen a partir de (5.36), o de su equivalente (5.37), pero reemplazando  $\Theta$  por  $-\Theta$ .

En la situación considerada en la figura 5.5 ambos extremos fijos son del tipo 2. Entonces debe cumplirse que

$$A_0^{(2)} = 0 \quad \text{y} \quad A_{N+1}^{(2)} = 0.$$

Luego,

$$u_2(\Theta) + u_2(-\Theta) = 0.$$

Reemplazando este resultado en la condición en el extremo derecho, resulta

$$\sin[(N+1)\Theta] = 0$$

y entonces  $(N+1)\Theta$  tiene que ser un múltiplo entero de  $\pi$ ,

$$\Theta^{(p)} = p \frac{\pi}{N+1}, \quad (5.40)$$

un resultado que si bien es completamente análogo al obtenido en la sección 3.1.4 para la cuerda con cuentas idénticas equiespaciadas, da lugar a varias diferencias notables que son consecuencia del carácter bivaluado de la relación de dispersión (5.33). En particular, para cada valor de  $\Theta^{(p)}$  dado por (5.40), ahora pueden corresponder *dos valores de frecuencia* (no uno, como en la sección 3.1.4): uno en la rama acústica y el otro en la rama óptica.

# Capítulo 6

## Descripción geométrica

En los ejemplos mecánicos vistos en los capítulos anteriores, los comportamientos ondulatorios se pudieron mostrar a partir de estudiar la dinámica impuesta en cada parte móvil por las fuerzas restitutivas del sistema. Como la dinámica de sistemas mecánicos clásicos está gobernada por las ecuaciones de Newton, la dificultad quedaba acotada a resolver matemáticamente las ecuaciones diferenciales acopladas o la ecuación de ondas resultante. Pero para estudiar fenómenos físicos tan familiares para la humanidad como el sonido o la luz (tan familiares que nacemos con detectores para ellos) el camino recorrido no ha sido éste, sino un camino inverso, que arranca de la fenomenología y con análisis, sistematización, deducción y comprobación permite llegar a leyes y teorías, aún sin conocer los principios fundamentales de la dinámica de las partes móviles.

La descripción de comportamientos ondulatorios que veremos a continuación está formulada en base a conceptos puramente geométricos que no hacen ninguna referencia a la ecuación de ondas ni a la naturaleza ondulatoria del fenómeno que se estudia (como óptica o acústica) y recibe el nombre de descripción geométrica (DG). Es una aproximación válida cuando la onda interactúa con objetos grandes comparados con la longitud de onda, es decir

$$\lambda \ll D, \tag{6.1}$$

con  $\lambda$  la longitud de onda y  $D$  el menor tamaño característico del objeto con el que interactúa la onda.

¿Cuál es el interés de la descripción geométrica? Que en su rango de validez permite describir de manera muy sencilla la *cinemática* (cómo viaja la onda) en procesos ondulatorios como la luz o el sonido. Sin embargo, no dice nada sobre la *dinámica* de la interacción onda–objeto, que requiere conocer las leyes físicas que están detrás de cada onda. En otras palabras, la descripción geométrica permite predecir qué caminos siguen las ondas que se reflejan y transmiten en un dado obstáculo –y estos caminos serían iguales para todo tipo de onda, siempre que valga la condición (6.1)– pero no dice nada sobre qué fracción de la energía incidente se va por el camino reflejado y qué fracción por el camino transmitido.

La construcción teórica de la DG sigue históricamente los avances del hombre en el entendimiento de la luz, por este motivo muchos libros la llaman *óptica geométrica*, si bien vale para cualquier tipo de onda. La luz es una onda electromagnética transversal con longitudes de onda

en el rango  $400 \text{ nm} < \lambda < 700 \text{ nm}$  ( $1 \text{ nm} = 10^{-9} \text{ m}$ ). Luego, para la mayoría de los objetos de la vida cotidiana, se satisface ampliamente la condición (6.1). En cambio, para las ondas sonoras en aire ( $\lambda \approx 1 \text{ m}$  para el  $la_{440}$ ), en la vida cotidiana por lo general no se satisface la condición (6.1) y en consecuencia no es aplicable la DG (o sea, que para que valga la DG hay que ir a otros rangos de frecuencia, es decir, de longitudes de onda).

## 6.1. Rayos

Un rayo se define como la línea que corresponde a la dirección del flujo de energía. Lo podemos imaginar como la línea recta que une la salida de un puntero laser con la manchita producida por el laser en un papel. Y se podría visualizar perfectamente si en la habitación donde se hace el experimento hay humo o polvo flotando en el aire. El concepto ideal de rayo se obtendría tratando de disminuir el tamaño de la manchita, por ejemplo limitando la apertura del puntero. Pero cuando la apertura resulta muy pequeña, la condición (6.1) deja de valer y a partir de este momento el procedimiento para visualizar el rayo ideal deja de funcionar: el tamaño de la manchita deja de disminuir y para describir lo que sucede hay que tener en cuenta el carácter ondulatorio, ya no se puede aplicar la DG y en estas circunstancias se dice que hay *difracción*.

Nos damos cuenta que los rayos, como caminos de las ondas, se corresponden con los vectores de onda  $\vec{k}$  asociados con las ondas 1D (ver pag. 88 y siguientes). Y entonces son perpendiculares a los frentes de onda: rayos paralelos se corresponden con ondas planas, rayos radiales se corresponden con ondas esféricas, etc.

## 6.2. Leyes fenomenológicas

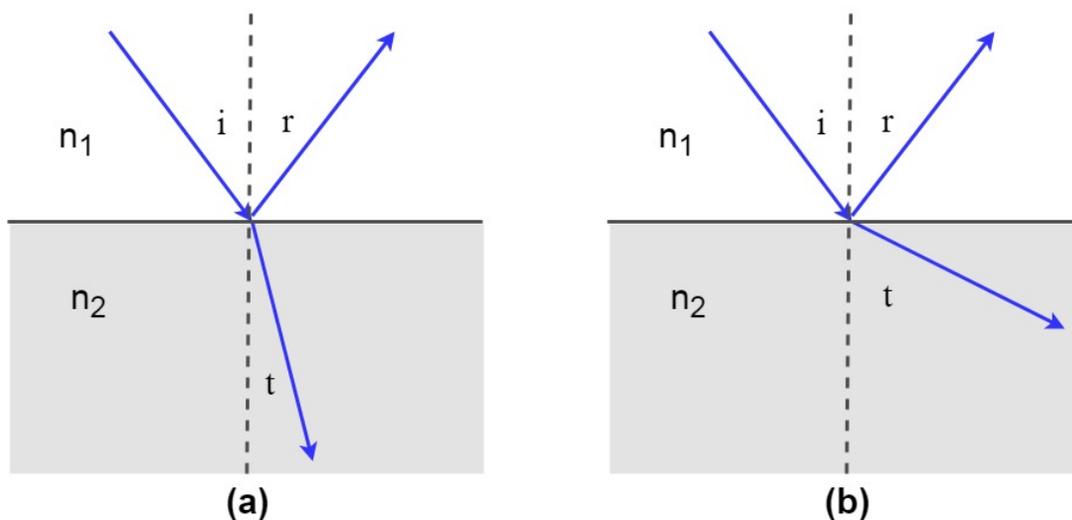


Figura 6.1: En la superficie de separación entre los medios 1 y 2, los experimentos demuestran que los rayos incidente, reflejado y refractado están contenidos en el plano determinado por el rayo incidente y la normal a la superficie (línea punteada) y las relaciones entre los ángulos  $i$ ,  $r$  y  $t$ .

Los experimentos demuestran que el comportamiento de los rayos se puede condensar en las siguientes leyes:

- en un medio homogéneo, los rayos son líneas rectas (ley cero);
- como se muestra en el esquema de la Fig. 6.1, en una superficie de separación entre dos medios homogéneos el rayo incidente da lugar a un rayo reflejado en el primer medio y a un rayo transmitido en el segundo medio;
- el rayo reflejado y el rayo transmitido están contenidos en el plano determinado por el rayo incidente y la normal a la superficie de separación. Este plano se llama *plano de incidencia*. Notar que en el caso de incidencia normal ( $i = 0$ ) el plano de incidencia no está bien definido, porque hay infinitos planos que contienen al rayo incidente y a la normal.
- $r = i$ , es decir que el ángulo  $r$  que forma el rayo reflejado con la normal es igual al ángulo  $i$  que forma el rayo incidente con la normal;
- el cociente entre el seno del ángulo  $t$  que forma el rayo transmitido con la normal y el seno del ángulo  $i$  que forma el rayo incidente con la normal sólo depende del par de medios y no del ángulo de incidencia. Esta ley, descubierta por *Ibn Sahl*, también se conoce como ley de Snell, de Descartes o ley de la refracción y se expresa

$$\frac{\sin i}{\sin t} = n_{21}. \quad (6.2)$$

La constante  $n_{21}$  se llama *índice de refracción relativo* entre el medio 2 y el medio 1. Si el medio 2 se reemplaza por un tercer medio, se verifica experimentalmente que el valor del cociente entre senos independiente del ángulo de incidencia toma ahora otro valor  $n_{31}$ . Si se vuelve a hacer el experimento en la superficie de separación entre los medios 2 y 3, ahora el cociente entre senos resulta  $n_{32} = n_{31}/n_{21}$  y esta propiedad sugiere tomar algún medio como referencia, entendiendo que las constantes involucradas en la ley de la refracción están medidas en relación a dicho medio:  $n_{21} = n_2/n_1$ . Así, la ley se escribe

$$n_1 \sin i = n_2 \sin t. \quad (6.3)$$

De la condición (6.1) se desprende que en esta aproximación todas las superficies son localmente planas en la zona donde incide el rayo, es decir que los radios de curvatura de la superficie son grandes comparados con la longitud de onda. Con experimentos caseros es posible poner en evidencia ciertos comportamientos luminosos que no responden a las leyes fenomenológicas de la DG. Por ejemplo basta con iluminar superficies, como la de un DVD o de ciertas tarjetas bancarias, que tienen detalles que el ojo humano no puede resolver. Por ejemplo, la distancia radial entre dos surcos consecutivos en la superficie de un DVD es aproximadamente 740 nm. Además, a lo largo de cada surco hay montañitas de 320 nm de ancho y como mínimo 400 nm de largo y 120 nm de alto. Estas cifras muestran que para estas superficies no se cumple la condición de validez de la DG, porque la menor distancia característica  $D$  es del mismo orden de magnitud que  $\lambda$ .

### 6.3. Algunas consecuencias de la ley de Ibn Sahl

#### 6.3.1. Reflexión externa e interna

En la figura 6.1a se muestra el esquema correspondiente al caso en que el índice de refracción del medio de incidencia es menor que el índice de refracción del medio de transmisión. En esta situación, conocida como *reflexión externa*, el rayo refractado *se acerca a la normal* y siempre hay un rayo refractado para todo valor del ángulo de incidencia.

En cambio, en la figura 6.1b se muestra el caso opuesto al anterior, cuando el índice de refracción del medio de incidencia es mayor que el índice de refracción del medio de transmisión. En esta situación, conocida como *reflexión interna*, el rayo refractado *se aleja de la normal*, tal como se pone en evidencia cuando despejamos  $\sin t$  en función de  $\sin i$ . Si tenemos en cuenta que en el intervalo  $[0, \pi/2]$ , intervalo donde están definidos los ángulos  $i$  y  $t$ , el seno es una función creciente de su argumento, resulta

$$\sin t = \frac{n_1}{n_2} \sin i \implies \begin{cases} t < i & \text{si } n_1 < n_2, \\ t > i & \text{si } n_1 > n_2. \end{cases} \quad (6.4)$$

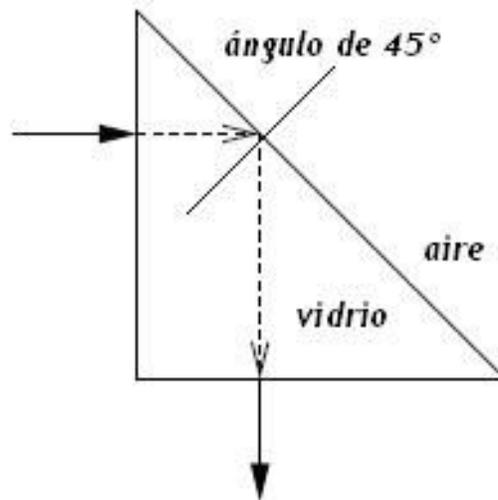


Figura 6.2: Esquema de funcionamiento del prisma de reflexión total. La sección del prisma tiene forma de triángulo rectángulo isósceles. Los rayos que inciden normalmente a uno de los catetos llegan a la hipotenusa con un ángulo mayor al que corresponde a  $i_c$  para la reflexión interna entre vidrio y aire.

#### 6.3.2. Reflexión total

Notar que en el caso  $n_1 > n_2$  existe un ángulo crítico  $i_c < \pi/2$  para el cual el rayo refractado es paralelo a la superficie  $t = \pi/2$ . Este ángulo se llama *ángulo crítico de reflexión total* y se

puede calcular a partir de la ley de Ibn Sahl

$$n_1 \sin i_c = n_2 \sin \pi/2 \quad \rightarrow \quad \sin i_c = \frac{n_2}{n_1}. \quad (6.5)$$

En este caso la ley de Ibn Sahl se puede reescribir como

$$\sin t = \frac{\sin i}{\sin i_c}. \quad (6.6)$$

Esta expresión muestra claramente que cuando

$$i > i_c \quad \rightarrow \quad \sin t > 1$$

una situación que no es nada clara en el lenguaje de rayos, porque no sabemos cómo dibujar ángulos con seno mayor que 1. Esto suena bastante parecido a lo que sucedía para comportamientos reactivos. Justamente, más adelante reexaminaremos el fenómeno con el lenguaje de las ondas y encontraremos que para  $i > i_c$ , el medio de transmisión se comporta reactivamente y entonces el rayo reflejado lleva toda la energía del rayo incidente. También encontraremos que para  $i < i_c$  el medio de transmisión se comporta dispersivamente y la energía del rayo incidente en parte se refleja y en parte se transmite, tal como sucede para todo ángulo de incidencia en el caso de reflexión externa  $n_1 < n_2$ .

Si el rayo va de vidrio ( $n_1 = 1.5$ ) hacia aire ( $n_2 = 1$ ), entonces  $\sin i_c = 1/1.5 = 2/3$  y resulta  $i_c \approx 42^\circ$ , un resultado que explica el diseño del prisma de reflexión total mostrado en la figura 6.2, donde se utiliza un prisma con forma de triángulo rectángulo isósceles para que el ángulo de incidencia en la hipotenusa resulte  $i = 45^\circ > i_c$  y así entonces el rayo transmitido en el primer cateto se refleja totalmente.

El mecanismo de reflexión total permite el confinamiento y la propagación de luz en **fibras ópticas**. Esencialmente la fibra consiste en un núcleo con un recubrimiento, con el índice de refracción del núcleo mayor que el índice del recubrimiento. De esta manera, la luz se propaga por el interior de la fibra con un ángulo de reflexión por encima del ángulo límite de reflexión total, rebotando en las superficies núcleo-recubrimiento.

## 6.4. La DG en otros marcos teóricos

Cuando estudiamos el problema de la cuerda mediante las descripciones de d'Alambert y de modos observamos que distintas descripciones de la misma situación pueden producir ideas físicas y métodos matemáticos de análisis diferentes y complementarios. Las leyes de la descripción geométrica también admiten teorías duales que pueden aplicarse como alternativa a las leyes fenomenológicas enunciadas en la página 122. Las descripciones alternativas que veremos están basadas en

- Principio de Huygens;
- Principio de Fermat;
- Marco ondulatorio riguroso.

Cada descripción tiene su ventaja, en cada una hay ideas físicas muy importantes que toman protagonismo y si estamos en el rango de validez de la DG, da lo mismo cuál se usa.

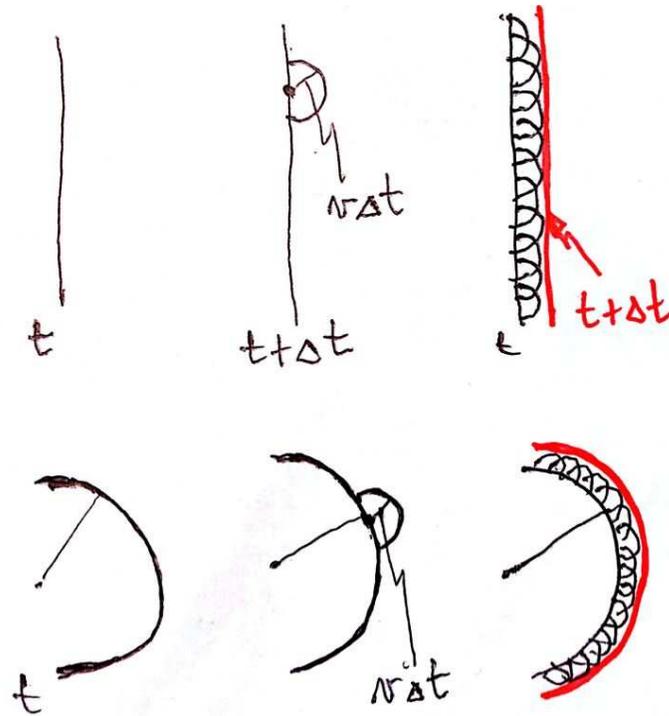


Figura 6.3: Esquema para obtener la propagación de frentes de onda planos y esféricos. Primero se muestra el frente de onda en el instante  $t$ . Luego una ondita secundaria en el instante  $t + \Delta t$ . Y finalmente la envolvente de todas las onditas secundarias en el instante  $t + \Delta t$ .

### 6.4.1. Principio de Huygens

En 1678 el físico holandés Christiaan Huygens propuso un mecanismo que permite explicar no solo la propagación de los rayos (normales a los frentes de onda) sino también las leyes fenomenológicas enunciadas en la página 122. El principio postula que todo punto de un frente de ondas actúa como fuente de onditas esféricas secundarias de igual velocidad y frecuencia que la onda original y que el frente de ondas en un instante posterior es la envolvente de todas estas onditas esféricas (onditas de Huygens). Por envolvente se entiende la superficie tangente. Para que el mecanismo funcione, Huygens supuso, sin explicaciones adicionales, que las onditas secundarias sólo viajan “hacia adelante”. El principio fue reformulado en 1815 por Fresnel para: i) hacer desaparecer las onditas que viajan “hacia atrás”, introduciendo un factor de oblicuidad que depende de la dirección de propagación; y ii) interpretar la envolvente en términos de interferencia entre las onditas secundarias.

La figura 6.3 muestra como aplicar el principio de Huygens para obtener la propagación de un frente de onda plano (arriba) y de un frente de onda esférico (abajo). En la primera columna se muestra el frente en el instante  $t$  y en la segunda columna una sola ondita secundaria en el instante  $t + \Delta t$ , notar que su frente de onda tiene un radio  $v\Delta t$ . En la tercera columna se muestra la envolvente de todas las onditas secundarias en el instante  $t + \Delta t$ . Se ha supuesto que el medio es isótropo y homogéneo, por eso todas las onditas secundarias tienen el mismo radio, si el medio no fuera homogéneo la velocidad dependería del punto, las onditas secundarias tendrían distinto radio y los frentes de onda no serían como los originales. Como los rayos son

perpendiculares a los frentes de onda, es evidente que estas construcciones equivalen a la ley cero de la página 122.

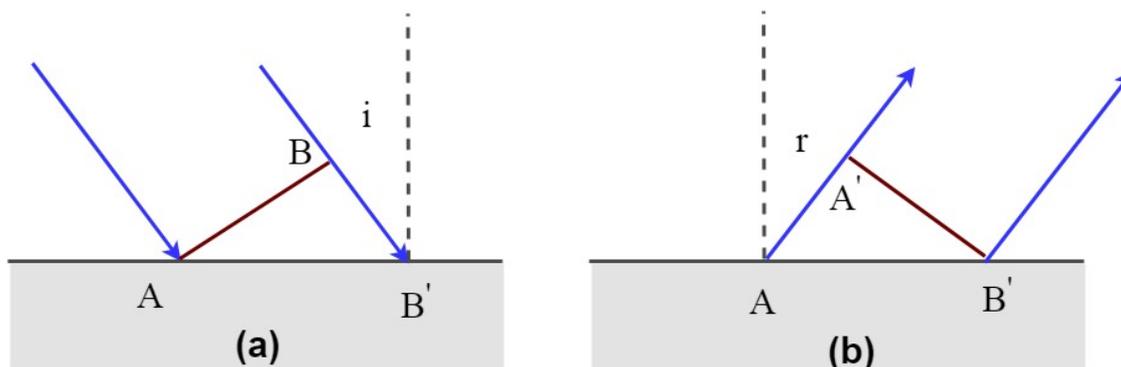


Figura 6.4: Esquema para deducir la ley de la reflexión a partir del principio de Huygens.

El Principio de Huygens también contiene la ley de la reflexión y la ley de Ibn Sahl. Para verlo, consideremos primero el esquema de la Fig. 6.4. La parte (a) de la figura representa el frente de una onda plana con rayos que forman un ángulo  $i$  con respecto a la normal a la superficie de separación. En el instante  $t$ , cuando el punto A del frente de ondas incidente toca a la superficie, genera una ondita esférica elemental reflejada que vuelve hacia el medio 1. Lo mismo hacen, en instantes posteriores, los otros puntos del frente de ondas  $\overline{AB}$ . Consideremos el instante  $t + \Delta t$ , cuando el punto B del frente de ondas incidente llega a la superficie. En ese instante, ver parte (b), la ondita esférica elemental reflejada en el punto A tiene radio

$$\overline{AA'} = v_1 \Delta t$$

mientras que la ondita esférica elemental reflejada en el punto  $B'$  recién está por empezar a emitir. Esto indica que el frente de ondas reflejado es el segmento  $\overline{A'B'}$  y que por hipótesis también  $\overline{BB'} = v_1 \Delta t$ . Como

$$\text{sen } i = \overline{BB'} / \overline{AB'} \quad \text{y} \quad \text{sen } r = \overline{AA'} / \overline{AB'}$$

entonces debe ser  $\text{sen } i = \text{sen } r$ , luego  $i = r$  y queda demostrada la ley de la reflexión.

Para deducir la ley de refracción a partir del del principio de Huygens, consideremos ahora el esquema de la Fig. 6.5 y supongamos que  $v_1$  y  $v_2$  son las velocidades de fase correspondientes a los medios de incidencia y de transmisión respectivamente. El segmento  $\overline{AB}$  representa, como antes, el frente de ondas de una onda plana con rayos que forman un ángulo  $i$  con respecto a la normal a la superficie de separación. En el instante  $t$ , cuando el punto A del frente de ondas incidente toca a la superficie, genera una ondita esférica elemental refractada que empieza a penetrar en el medio 2. Lo mismo hacen, en instantes posteriores, los otros puntos del frente de ondas  $\overline{AB}$ . Consideremos el instante  $t + \Delta t$ , cuando el punto B del frente de ondas incidente llega a la superficie, condición que impone que

$$\overline{BB'} = v_1 \Delta t.$$

En este instante, la ondita esférica elemental refractada a partir del punto A tiene radio

$$\overline{AA'} = v_2 \Delta t$$

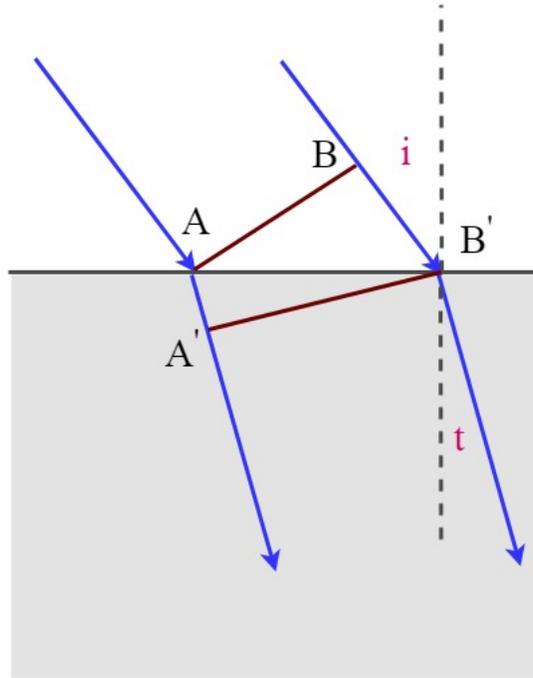


Figura 6.5: Esquema para deducir la ley de Ibn Sahl a partir del principio de Huygens.

mientras que la ondita esférica elemental refractada en el punto  $B'$  recién está por empezar a emitir (notar que en el esquema se supuso  $v_1 > v_2$ ). Esto indica que el frente de ondas refractado es el segmento  $\overline{A'B'}$  y como

$$\text{sen } i = \overline{BB'}/\overline{AB'} \quad \text{y} \quad \text{sen } t = \overline{AA'}/\overline{AB'}$$

entonces debe ser

$$\frac{\text{sen } i}{\text{sen } t} = \frac{\overline{BB'}}{\overline{AA'}} = \frac{v_1}{v_2}.$$

Esta expresión muestra que el cociente entre los senos solamente depende de los medios involucrados y no del ángulo de incidencia y equivale a la ley de Ibn Sahl escrita en la forma (6.2). El punto importante, además de entender cómo aplicar el mecanismo de Huygens cuando hay una discontinuidad, es notar que con este mecanismo Huygens permite interpretar la constante que aparece en la ley fenomenológica (6.2) como un cociente entre velocidades de fase de la onda en cada medio. Por eso la notación moderna escribe esta constante como un cociente entre índices de refracción, cantidades que para cada medio están definidas como cocientes entre la velocidad de fase en un cierto medio de referencia y la velocidad de fase en el medio en cuestión. De esta manera

$$\frac{1}{v_1} \text{sen } i = \frac{1}{v_2} \text{sen } t$$

que demuestra que, admitiendo las hipótesis de Huygens, es posible medir la velocidad relativa de la luz en dos medios distintos con sólo medir los ángulos de incidencia y refracción en una superficie de separación entre ambos medios.

Eligiendo convenientemente la forma de las superficies y los materiales, es posible diseñar elementos que introduzcan localmente desfases (retrasos o adelantos) en los frentes de on-

das, ya sea mediante reflexiones o mediante refracciones. Por ejemplo, las antenas parabólicas transforman, mediante reflexión en la superficie de un paraboloide de revolución, un frente de ondas plano en un frente de ondas que converge en el foco de la parábola. El resultado se puede demostrar aplicando la ley de la reflexión en cada punto del paraboloide y para cada rayo incidente, o estudiando los desfases de las onditas elementales de Huygens y luego calculando la envolvente.

El esquema de la figura 6.6 muestra un elemento refractivo (lente convergente) que transforma el frente de onda plano  $\overline{AB}$  de una onda incidente y lo convierte en el frente de ondas  $\widehat{A'B'}$  con forma de superficie esférica con centro en C. En el lenguaje de rayos, la lente transforma un haz de rayos paralelos en un haz convergente.

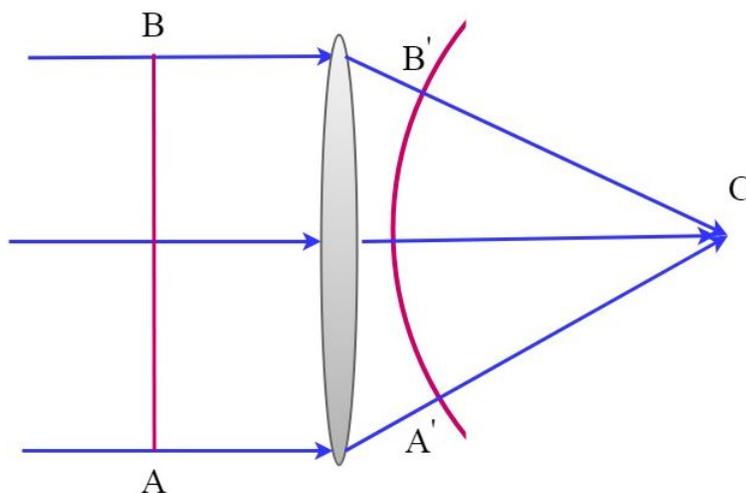


Figura 6.6: Un elemento refractivo (lente) transforma, mediante refracciones en dos superficies, un frente incidente plano en un frente de ondas esférico con centro en C (esto es, un haz de rayos paralelos en un haz convergente).

### 6.4.2. Principio de Fermat

Es un principio variacional para encontrar el camino que hace un rayo entre dos puntos (A y B). Es evidente que entre dos puntos hay infinitos caminos, algunos muy largos y otros más cortos. Matemáticamente, cada camino es una curva en el espacio 3D, que empieza en A y termina en B. ¿Cuál es el “verdadero” camino que “elige” un rayo que parte de A y llega a B y cumpliendo ciertas condiciones adicionales, como por ejemplo pasar por un espejo? Justamente el principio de Fermat da una receta para encontrar este camino.

El principio de Fermat se formula en términos de una magnitud llamada *longitud de camino óptico*, o directamente camino óptico a secas. La magnitud camino óptico está asociada con *dos puntos* (A y B) *y una trayectoria* entre ambos puntos. No tiene sentido hablar de camino óptico sin especificar dos puntos y una trayectoria.

En la figura ?? se esquematizan distintas trayectorias entre dos puntos A y B.

Definámosla. Dada la trayectoria  $j$  entre los puntos A y B, el camino óptico asociado a dicha trayectoria es la integral de camino del índice de refracción

$$[AB]_j = \int_j n(\ell) d\ell, \quad (6.7)$$

con  $d\ell$  el elemento de arco en la trayectoria  $j$ .

Si el medio es homogéneo, el índice de refracción no depende del punto de la trayectoria. Entonces, la longitud de camino óptico asociada a una trayectoria entre dos puntos en un medio homogéneo es igual al producto del índice de refracción por la longitud de la trayectoria. Si además la trayectoria es rectilínea,  $[AB]_{recta} = n \overline{AB}$ , es decir el camino óptico es el producto del índice de refracción por la distancia entre A y B.

Usando la definición de índice de refracción, se ve que la longitud de camino óptico asociada con la trayectoria  $j$  resulta proporcional al tiempo que tardaría el rayo en recorrer dicha trayectoria  $j$

$$[AB]_j = \int_j n(\ell) d\ell = c \int_j \frac{1}{v(\ell)} d\ell = c \int_j dt = c t_{AB}. \quad (6.8)$$

Y ahora que entendemos qué es el camino óptico asociado a una trayectoria entre dos puntos A y B, estamos en condiciones de formular el Principio de Fermat. Que dice así: *la trayectoria que siguen los rayos para ir de A a B es aquella para la cual el camino óptico es estacionario con respecto a variaciones de la trayectoria*. Y se debe recordar que dichas variaciones de la trayectoria deben ser compatibles con las condiciones del problema, por ejemplo en un problema de reflexión las variaciones de la trayectoria no pueden incluir caminos que dejen de pasar por el espejo.

Claramente este principio contiene la propagación rectilínea en un medio homogéneo, porque la única trayectoria con camino óptico estacionario, es decir que tenga camino óptico a primer orden igual que los caminos cercanos, es la trayectoria rectilínea, que tiene camino óptico mínimo. Máximo no hay porque es claro que el camino óptico en este caso no está acotado.

El Principio de Fermat también contiene la ley de la reflexión y la ley de Ibn Sahl . . . completar, por el momento ver <http://bit.ly/10GHKVY>.

Usando un espejo elíptico y sus deformaciones, vimos un caso donde el camino que predice el Principio de Fermat no es único, otro caso donde es único y es un máximo y por supuesto el conocido caso donde es único y es un mínimo. Un punto a tener en cuenta, hay textos que siguen diciendo que el Principio de Fermat es un principio de longitud de camino óptico mínimo.

*El principio de Fermat no puede ser la causa, pues de lo contrario estaríamos atribuyendo conocimiento a la naturaleza . . .*

Claude Clerselier (1622)

## 6.5. Formación de imágenes

Los objetos luminosos (emisores o iluminados) se comportan como una colección de fuentes puntuales y cada fuente emite rayos divergentes. Ya vimos que si una fuente puntual se coloca

en el foco de un espejo con forma de elipsoide, todos los rayos reflejados en el espejo convergen en el otro foco. En este caso se ha logrado transformar un haz de rayos divergentes con centro en un foco, en un haz de rayos convergentes con centro en el otro foco y se dice que el segundo foco es la imagen del primer foco.

Se llama sistema óptico a un conjunto de superficies reflectoras y refractoras. Si el sistema óptico transforma un haz de rayos divergentes desde el punto S en un haz de rayos convergentes en el punto S', se dice que el sistema óptico forma en el punto S' una **imagen perfecta** del punto S. Excepto la energía perdida por efectos no deseados, como reflexiones inevitables cuando se usan sistemas refractores, **toda la energía** existente en el haz que incide en el sistema óptico llega al punto S'. S y S' también reciben el nombre de **puntos conjugados**, porque de acuerdo con la reversibilidad implícita en el principio de Fermat, una fuente puntual colocada en S' debe dar una imagen perfecta en S. El principio de Fermat también asegura que todas las trayectorias posibles que unen S y S' tienen el mismo camino óptico.

Por lo general la función de un sistema óptico es recoger una porción del frente de onda incidente (un haz de rayos) para formar una imagen, es decir que el sistema óptico se diseña para que el camino óptico de todas las trayectorias posibles entre un punto emisor S y otro punto S' tenga siempre el mismo valor. Desde este punto de vista, la formación de imágenes es la manipulación controlada de los rayos (o de los frentes de onda).

Cuando a la salida del sistema óptico todos los rayos se cortan en un punto, se dice que la imagen es **real**. En cambio, cuando las prolongaciones de los rayos que salen del sistema óptico se cortan en un punto, se dice que la imagen es **virtual**.

### 6.5.1. Dioptra plana

El caso más sencillo de sistema óptico, una superficie plana que separa dos medios de índices  $n_1$  y  $n_2$ . Para estudiar la formación de imágenes en sistemas ópticos en los libros hay varias convenciones de signo, acá usamos la convención más natural en física: tomar distancias positivas hacia la derecha y negativas hacia la izquierda tanto para los objetos como para las imágenes. Suponemos que la luz viaja de izquierda a derecha. La dioptra plana está en  $x = 0$  (es el plano  $yz$ ) y la fuente puntual S tiene coordenadas  $(s, 0, 0)$  ( $s < 0$  si la luz viene de izquierda a derecha). El rayo que sale de S e incide en la dioptra en el punto  $(0, h, 0)$  forma un ángulo  $i$  con la normal y se refracta en el semiespacio  $x > 0$  formando un ángulo  $t$  con la normal. Es claro que para cada  $i$  hay un  $h$  distinto. Y el rayo que sale de S e incide en la dioptra en el punto  $(0, 0, 0)$  no se desvía cuando se refracta. Estos dos rayos no se cortan en el semiespacio  $x > 0$ , pero si en el semiespacio  $x < 0$ , en el punto  $(s', 0, 0)$ ,  $s' < 0$ . En un sistema refractivo,  $s' < 0$  equivale a que la imagen es virtual. Pero no sabemos si forma imagen. Para que esta dioptra forme imagen por transmisión, la distancia  $s' < 0$  debe ser independiente del ángulo  $i$ . Y como  $\operatorname{tg} i = h/s$  y  $\operatorname{tg} t = h/s'$ , resulta que

$$s' = \frac{\operatorname{tg} i}{\operatorname{tg} t} s = \frac{n_2 \cos t}{n_1 \cos i} s, \quad (6.9)$$

que claramente depende del ángulo  $i$ : el sistema no forma imagen perfecta. Pero si el sistema solamente recoge rayos que viajan cerca del eje, es decir, ángulos  $i \ll 1$ , entonces con muy buena aproximación se pueden poner los cosenos igual a 1, y resulta que con muy buena aproximación

se puede decir que el sistema forma imagen, ubicada a la distancia

$$s' \approx \frac{n_2}{n_1} s . \quad (6.10)$$

Esta aproximación de suponer que el sistema solamente recoge rayos que viajan cerca del eje, es decir, ángulos  $i \ll 1$ , se llama **aproximación paraxial**, o también **aproximación de primer orden**, porque equivale a aproximar la ley de Ibn Sahl por  $n_1 i \approx n_2 t$ , es decir, aproximar los senos por los ángulos. Estrictamente hablando, la fuente puntual S no produce una imagen puntual S', si no una mancha, que será más grande cuanto mayor sea la apertura de los rayos.

La formación de imágenes por reflexión se estudia de manera similar, en vez de seguir el rayo refractado en  $(0, h, 0)$  hay que seguir el rayo reflejado. Que forma con la normal el mismo ángulo que el rayo incidente, pero con la orientación contraria. La ley de la reflexión es entonces  $r = -i$ , que es como la ley de Ibn Sahl aproximada  $n_1 i \approx n_2 t$  pero con  $n_1 = -n_2$  y con  $r$  en vez de  $t$ . Y además vale en forma exacta, es decir que el espejo plano forma imagen perfecta ubicada a una distancia

$$s' = -s , \quad (6.11)$$

que se obtiene de (6.10) con  $n_1 = -n_2$ . Notar que en un sistema reflector,  $s' > 0$  indica que la imagen es virtual.

### 6.5.2. Dioptra esférica

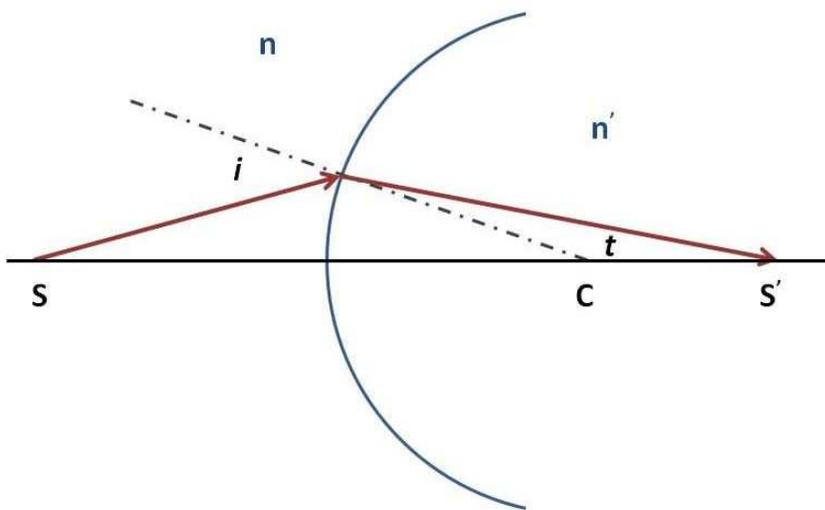


Figura 6.7: Refracción en dioptra esférica, radio  $R$  ( $R > 0$  en el esquema), C es el centro de curvatura.

Por tener curvatura constante, las dioptras esféricas son fáciles de fabricar. La mayoría de las lentes para corregir defectos de visión son esféricas y están diseñadas para trabajar en la zona paraxial. Ver esquema para la formación de imágenes por transmisión en la figura 6.7. La fuente puntual S tiene coordenadas  $(s, 0, 0)$ ,  $s < 0$  en el dibujo. El rayo que sale de S e incide

en la dioptra con un ángulo  $i$  con la normal se refracta en el semiespacio  $x > 0$  formando un ángulo  $t$  con la normal. Ahora la normal va del centro de curvatura  $C$  al lugar donde incide el rayo. El rayo que sale de  $S$  e incide en la dioptra en el punto  $(0, 0, 0)$  no se desvía cuando se refracta. En el dibujo, estos dos rayos se cortan en el semiespacio  $x > 0$  en el punto  $(s', 0, 0)$ ,  $s' > 0$ . Esto es así porque se está suponiendo que  $n' > n$  y que la dioptra es convexa (vista desde el lado que viene la luz). En cambio, con la misma curvatura, si  $n' < n$ , es claro que el rayo se habría alejado de la normal y que  $s' < 0$ . En la dioptra refractiva,  $s' > 0$  indica que la imagen es real. Pero para que esta dioptra realmente forme imagen por transmisión, se debe cumplir que esta distancia  $s' < 0$  sea independiente del ángulo  $i$ . Y resulta que esto se cumple en la aproximación paraxial. Queda como ejercicio demostrar que la siguiente ley

$$\frac{n'}{s'} - \frac{n}{s} = \frac{n' - n}{R}, \quad (6.12)$$

relaciona  $s$  y  $s'$  con todos los parámetros físicos de la dioptra. El lado derecho de esta igualdad se llama **potencia**  $\Phi$  de la dioptra.

Si la imagen de un objeto está en infinito, el objeto está en el foco objeto y su distancia a la dioptra se llama distancia focal objeto,  $s = f$ .

Se define foco imagen a la posición donde se forma la imagen de un objeto que está en infinito, su distancia a la dioptra se llama distancia focal imagen,  $s' = f'$ . De acuerdo con estas definiciones

$$\frac{n'}{s'} - \frac{n}{s} = \frac{n' - n}{R} = \frac{n'}{f'} = -\frac{n}{f} = \Phi, \quad (6.13)$$

Si  $f' > 0$  la dioptra es convergente, si  $f' < 0$  divergente. Notar que no sólo depende de que sea cóncava o convexa, también depende de la diferencia de índices de refracción.

El aumento lateral de una dioptra se define como el cociente entre la altura de imagen y objeto. Trazando el rayo que sale del extremo del objeto y pasa por el vértice de la dioptra, es fácil demostrar que el aumento lateral de la dioptra esférica vale

$$m = \frac{h'}{h} = \frac{n s'}{n' s}. \quad (6.14)$$

La formación de imágenes por espejos esféricos se estudia de manera similar a como se hizo para espejos planos: hay que seguir el rayo reflejado. Que forma con la normal un ángulo  $r = t$ , que el rayo incidente, pero con la orientación contraria. La ley de la reflexión es entonces  $r = t$ , que es como la ley de Ibn Sahl aproximada  $n_1 i \approx n_2 t$  pero con  $n_1 = -n_2$ . Vemos entonces que la ecuación para espejos esféricos resulta

$$\frac{1}{s'} + \frac{1}{s} = \frac{2}{R} = \frac{1}{f'} = \frac{1}{f}, \quad (6.15)$$

que se obtiene de (6.13) con  $n_1 = -n_2$ .

### 6.5.3. Trazado de rayos

La imagen formada por una dioptra esférica se puede localizar gráficamente de una manera sencilla utilizando las propiedades de los focos objeto e imagen y del centro de curvatura.

Teniendo en cuenta que dentro de los límites de validez de la aproximación paraxial el sistema forma imagen, cada punto de la imagen se obtendrá investigando cómo se refractan los rayos emitidos por cada punto del objeto. Es conveniente tomar los rayos emitidos por la parte superior del objeto y refractados en la dioptra. Donde estos rayos (o sus prolongaciones) se corten, allí estará la imagen, real (o virtual), de la parte superior de dicho objeto. Con la base del objeto se podría hacer lo mismo, pero no es necesario porque en la aproximación paraxial las dioptras y espejos simples transforman planos en planos <sup>1</sup>.

De los rayos emitidos por la parte superior del objeto, un rayo fácil de seguir es el paralelo a la línea central que pasa por el centro de curvatura y la base del objeto, porque este rayo se refracta pasando por el foco imagen en dioptras convergentes ( $f' > 0$ ) o con su prolongación pasando por el foco imagen en dioptras divergentes ( $f' < 0$ ). Otro rayo fácil de seguir es el que pasa, él o su prolongación, por el foco objeto. Este rayo se refracta en la dioptra saliendo paralelo a la línea central que pasa por el centro de curvatura y la base del objeto. Si bien con dos rayos es suficiente para determinar la imagen, también se podría usar un tercer rayo: el que pasa, él o su prolongación, por el centro de curvatura de la dioptra, que cuando se refracta no se desvía porque incide en la dirección de la normal local.

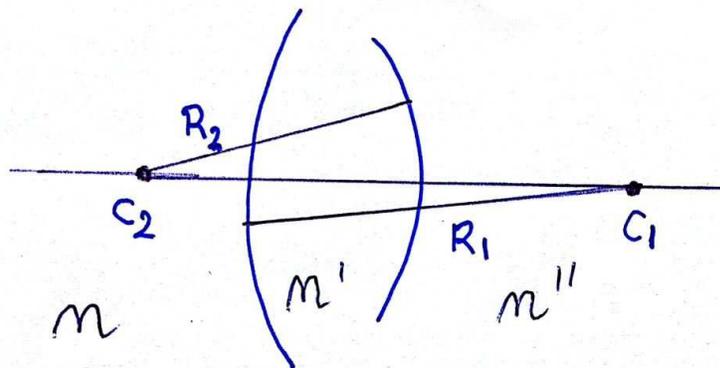


Figura 6.8: Lente gruesa: dos dioptras esféricas, radios de curvatura  $R_1$  y  $R_2$ , centros de curvatura  $C_1$  y  $C_2$ . La distancia entre dioptras, medida sobre el eje que une  $C_1$  y  $C_2$  es  $d$ .

#### 6.5.4. Lentes

Una lente no es otra cosa que dos dioptras, el esquema en la Fig. 6.8, donde se muestran dos dioptras esféricas, de radios de curvatura  $R_1$  y  $R_2$  y con centros de curvatura  $C_1$  y  $C_2$ . La distancia entre dioptras, medida sobre el eje que une  $C_1$  y  $C_2$  es  $d$ . Si la distancia  $d$  es chica, se dice que se tiene una **lente delgada**, si no, una **lente gruesa**. El análisis para la formación de imágenes es la aplicación sucesiva del análisis hecho para una dioptra. El esquema se muestra

<sup>1</sup>A un plano objeto perpendicular al eje óptico, le corresponde un plano imagen también perpendicular al eje (planos conjugados), de tal forma que todo punto objeto contenido en el plano objeto, tiene su correspondiente punto imagen contenido en el plano imagen.

en la Fig. 6.9: el objeto está situado en  $s_1$ , la imagen de este objeto producida por la primera dioptra está ubicada en  $s_1'$ . Y según (6.12)

$$\frac{n'}{s_1'} - \frac{n}{s_1} = \frac{n' - n}{R_1}. \quad (6.16)$$

Análogamente, la imagen en  $s_1'$  actúa como objeto para la segunda dioptra, que ahora produce una imagen ubicada en  $s_2'$ . Según (6.12), las distancias  $s_2$  y  $s_2'$ , medidas sobre el eje y desde la segunda dioptra, están relacionadas por

$$\frac{n''}{s_2'} - \frac{n'}{s_2} = \frac{n'' - n'}{R_2}. \quad (6.17)$$

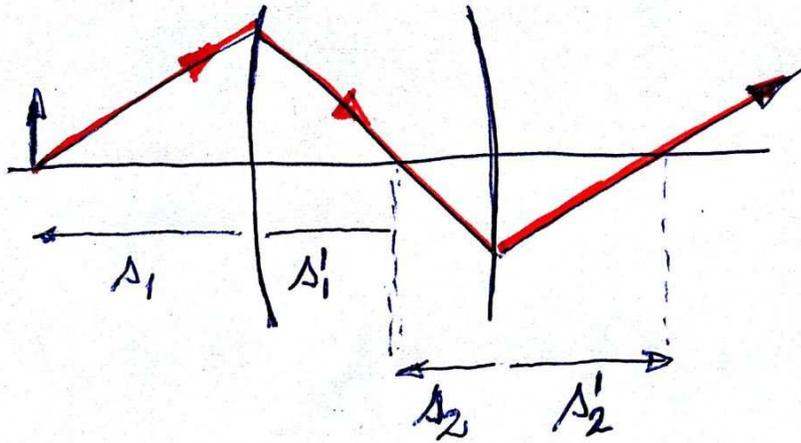


Figura 6.9: La primera dioptra produce una imagen en  $s_1'$  del objeto situado en  $s_1$  y dicha imagen actúa como objeto para la segunda dioptra.

En el esquema se ve que

$$|s_2| = d - s_1'. \quad (6.18)$$

es decir que

$$s_2 = s_1' - d. \quad (6.19)$$

y esta relación da el esquema para aplicar (6.17) a partir del resultado de (6.16) (y para seguir iterando si hubiera más dioptras).

En el caso de lente delgadas  $s_2 \approx s_1'$  y sumando miembro a miembro las ecuaciones (6.16) y (6.17) se obtiene

$$\frac{n''}{s_2'} - \frac{n}{s_1} = \frac{n' - n}{R_1} + \frac{n'' - n'}{R_2} = \Phi_1 + \Phi_2 = \frac{n''}{f'} = -\frac{n}{f}, \quad (6.20)$$

que muestra que la potencia de una lente delgada es la suma de las potencias de cada dioptra y que además define las distancias focales de la lente delgada. Si la lente es delgada no tiene

sentido seguir distinguiendo los sistemas 1 y 2, todas las distancias se miden desde el mismo lugar y por este motivo la ecuación (6.20) se reescribe como

$$\frac{n''}{s'} - \frac{n}{s} = \frac{n' - n}{R_1} + \frac{n'' - n'}{R_2}. \quad (6.21)$$

El aumento lateral de una lente, el cociente entre la altura de la imagen y la altura del objeto, se obtiene siguiendo un procedimiento análogo al empleado en clase para el aumento lateral de una dioptra. El aumento lateral de una lente resulta

$$m = \frac{h'}{h} = \frac{n s'}{n'' s}. \quad (6.22)$$

Recordar que, de acuerdo con el principio de Fermat, todas las trayectorias posibles entre un punto objeto y su imagen dada por cualquier sistema óptico, tienen el mismo camino óptico.

## 6.6. La DG en el marco ondulatorio

A continuación reinterpretaremos las leyes de la descripción geométrica desde la visión de la teoría ondulatoria. Esto nos permitirá volver a las pobres ondas que quedaron un poco abandonadas. Y relacionarlas con las leyes fenomenológicas que fuimos discutiendo en términos de conceptos puramente geométricos como *rayos* o *caminos ópticos*. Veremos que así quedarán mas claros algunos aspectos que no se terminaban de entender, como por ejemplo *qué sucede con el rayo transmitido en reflexión total cuando el ángulo de incidencia es mayor que el ángulo crítico*

Recordemos lo discutido en la página 85, cuando pensábamos mayormente en sistemas 1D como sogas o gases en tubos. En particular, habíamos dicho que una onda progresiva en un medio continuo infinito continúa propagándose sin modificaciones, a menos que se encuentre con una discontinuidad, como un obstáculo fijo o impenetrable, un cambio en la densidad o en la tensión de la soga, o un cambio en la densidad o en la presión del gas. Hemos visto que para poder ajustar las condiciones de contorno en la discontinuidad, además de la onda transmitida que representa el avance de la perturbación en el segundo medio, también aparece una onda regresiva en el primer medio, llamada onda reflejada. Las dos ondas que aparecen en este ajuste de condiciones de contorno tienen la misma frecuencia que la onda incidente y sus periodicidades espaciales  $2\pi/k$  surgen de las relaciones de dispersión  $k(\omega)$  en cada medio, es decir que el  $k$  de la onda reflejada es igual al de la onda incidente, pero diferente en general del  $k$  de la onda transmitida.

En medios 3D los mecanismos físicos son completamente análogos a los que estamos acostumbrados para medios 1D, la necesidad de ajustar condiciones de contorno en una discontinuidad produce la aparición de ondas reflejadas en el medio de incidencia y de ondas transmitidas en el segundo medio. Estas ondas tienen la misma frecuencia que la onda incidente y sus periodicidades espaciales en la dirección de propagación  $2\pi/k$  surgen de las relaciones de dispersión  $k(\omega)$  en cada medio, el  $k$  de la onda reflejada es igual que el de la onda incidente y en general diferente al  $k$  de la onda transmitida. Pero mientras que antes la discontinuidad se podía describir sólo con un valor particular de la única coordenada involucrada (por ejemplo,  $x = L$ ), en un medio

3D la discontinuidad entre dos medios homogéneos es una superficie y aún en el caso en que esta superficie fuera plana, la dirección de propagación de la onda incidente no es necesariamente perpendicular a dicha superficie. Justamente el formalismo de la descripción geométrica da la respuesta al problema cinemático de cómo orientar los vectores de onda de las ondas reflejadas y transmitidas, dada la dirección de la onda incidente y la forma de la discontinuidad.

Para encarar la interpretación de las leyes de la óptica geométrica en el marco de una descripción ondulatoria comenzaremos considerando el caso más sencillo posible de discontinuidad en un medio 3D: un plano infinito. Que tiene la ventaja de ser un obstáculo que no tiene “tamaños característicos” (esas cantidades que se comparaban con la longitud de onda para ver la validez de la descripción geométrica). En realidad un plano infinito **sí** tiene tamaños característicos: valen infinito, tanto su extensión como su curvatura. Y teniendo en cuenta que las ondas planas armónicas forman base para cualquier solución de la ecuación de ondas clásica, (sección 10.3), a continuación analizaremos como satisfacer las condiciones de contorno cuando la discontinuidad está “iluminada” por una onda plana. Veremos que todos los aspectos cinemáticos de este problema surgen solamente de admitir la existencia de condiciones de contorno, *aún sin especificar de qué condiciones de contorno estamos hablando*. Justamente por este motivo, estos comportamientos cinemáticos (las leyes de la óptica geométrica) son comunes a cualquier fenomenología ondulatoria (mecánica, acústica, electromagnética, etc.).

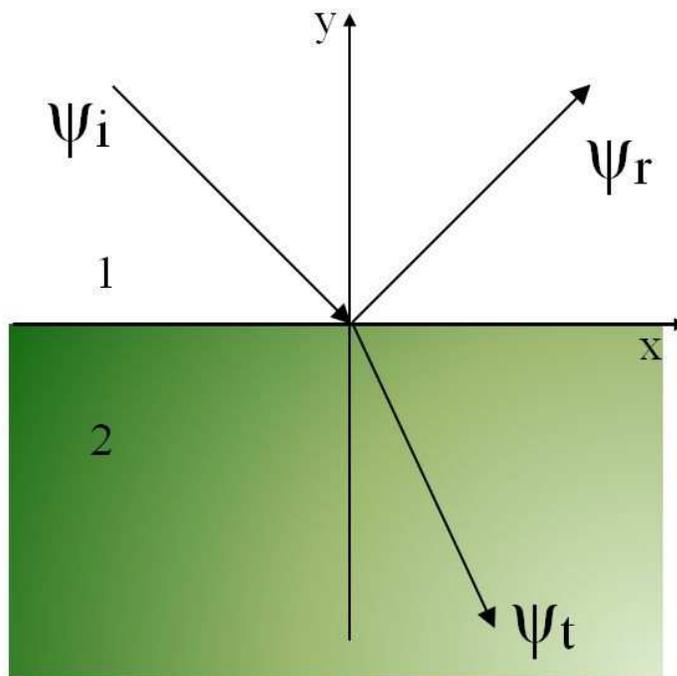


Figura 6.10: Onda plana incidente, reflejada y transmitida.

### 6.6.1. Reflexión y transmisión de una onda plana en una superficie plana

El plano  $y = 0$  separa dos medios materiales isótropos, uno en la región 1 ( $y > 0$ ) y otro en la región 2 ( $y < 0$ ). La función de onda en cada región satisface la ecuación de ondas clásica con velocidades de fase  $v_1(\omega) = c/n_1(\omega)$  y  $v_2(\omega) = c/n_2(\omega)$ ,  $n_1(\omega)$  y  $n_2(\omega)$  los índices de refracción de los medios 1 y 2. Una onda incidente monocromática plana de frecuencia  $\omega$  se propaga en la región  $y > 0$  y se encuentra con la discontinuidad en  $y = 0$ . Podríamos pensar que esta onda incidente fue originada por una fuente situada en  $y \rightarrow \infty$ , pero hay otras posibilidades. Sin pérdida de generalidad, supongamos que el plano  $x - y$  es el plano de incidencia. Según (4.21)

$$\psi_i(\vec{r}, t) = A_{0i} e^{i(\vec{k}_i \cdot \vec{r} - \omega t)} , \quad (6.23)$$

con

$$\vec{k}_i = \underbrace{\frac{\omega}{c}}_{k_1} n_1 (\sin i \hat{x} - \cos i \hat{y}) = \alpha_i \hat{x} - \beta_i \hat{y} . \quad (6.24)$$

En analogía con lo discutido para el caso 1D, intentamos ajustar las condiciones de contorno en la discontinuidad  $y = 0$  mediante una onda reflejada (en el medio 1) y una onda transmitida (en el medio 2)

$$\psi_r(\vec{r}, t) = A_{0r} e^{i(\vec{k}_r \cdot \vec{r} - \omega t)} , \quad (6.25)$$

$$\psi_t(\vec{r}, t) = A_{0t} e^{i(\vec{k}_t \cdot \vec{r} - \omega t)} . \quad (6.26)$$

Como en todo forzado estacionario de un sistema físico lineal, estas ondas tienen la misma frecuencia que la onda incidente. Y por ser soluciones de ecuaciones de ondas clásicas, las periodicidades espaciales en la dirección de propagación surgen de las relaciones de dispersión para cada medio. Como ya hemos discutido, estas relaciones fijan el módulo del vector de onda, pero no su dirección. Por eso escribimos los vectores de onda de las ondas reflejada y transmitida como

$$\vec{k}_r = \alpha_r \hat{x} + \beta_r \hat{y} + \gamma_r \hat{z} , \quad (6.27)$$

y

$$\vec{k}_t = \alpha_t \hat{x} - \beta_t \hat{y} + \gamma_t \hat{z} , \quad (6.28)$$

con

$$\alpha_r^2 + \beta_r^2 + \gamma_r^2 = \frac{\omega^2}{c^2} n_1^2 = k_1^2 , \quad (6.29)$$

$$\alpha_t^2 + \beta_t^2 + \gamma_t^2 = \frac{\omega^2}{c^2} n_2^2 = k_2^2 . \quad (6.30)$$

### 6.6.2. Las condiciones de contorno imponen la cinemática de los rayos

Aunque no sepamos en qué dirección apuntan  $\vec{k}_r$  y  $\vec{k}_t$ , las ecuaciones (6.29) y (6.30) dicen que  $\vec{k}_r$  y  $\vec{k}_t$  están en esferas de radio  $k_1$  y  $k_2$  respectivamente. La dirección surge de aceptar

que en el plano  $y = 0$  hay que satisfacer condiciones de contorno lineales. Por condiciones de contorno lineales entendemos condiciones que se puedan escribir como combinación lineal de las funciones de onda y sus derivadas a un lado y otro de la discontinuidad. Es instructivo revisar los problemas que se hicieron en los casos 1D, porque justamente éste es el tipo de condiciones de contorno que han aparecido en estos problemas, donde la física de cada situación particular condujo siempre a condiciones del tipo “continuidad de la función”, o “continuidad de la derivada de la función”, o “continuidad de la derivada de la función multiplicada por un coeficiente”. Supongamos que estamos en el caso “continuidad de la función”. Esto quiere decir que en todo instante de tiempo y para todo punto  $\vec{r}_P$  en el plano  $x - z$  ( $y = 0$ ), debe cumplirse que  $\psi_i(\vec{r}_P, t) + \psi_r(\vec{r}_P, t) = \psi_t(\vec{r}_P, t)$ . Esta condición nos lleva a pedir que

$$A_{0i} e^{i(\alpha_i x)} = -A_{0r} e^{i(\alpha_r x + \gamma_r z)} + A_{0t} e^{i(\alpha_t x + \gamma_t z)}. \quad (6.31)$$

Esto es raro ... ¿Una función de  $x$  (el lado izquierdo) igual a una función de  $x$  y de  $z$  (el lado derecho)? Sería posible (pero no demasiado interesante) si todas las amplitudes fueran cero. Si no, para que se cumpla la ecuación (6.31) debe ser

$$\gamma_t = \gamma_r = 0. \quad (6.32)$$

Y esto equivale a decir que el vector de onda reflejado  $\vec{k}_r$  y el vector transmitido  $\vec{k}_t$  están en el plano que contiene al vector de onda incidente  $\vec{k}_i$  y a la normal  $\hat{y}$ . **La ley del plano de incidencia.** Es fácil darse cuenta que habríamos llegado a la misma conclusión si en vez de considerar condiciones de contorno del tipo “continuidad de la función” hubiéramos considerado “continuidad de la función multiplicada por un coeficiente” u otro tipo de condición de contorno lineal. Vemos entonces que la ley sobre el plano de incidencia es una consecuencia de la existencia de condiciones de contorno, toda onda plana reflejada o transmitida en una superficie plana va a estar contenida en el plano de incidencia, independientemente de la naturaleza de la onda.

Admitida la igualdad (6.32), como todos los vectores están en el plano  $x - y$ , en este plano se pueden definir ángulos de reflexión y transmisión entre los vectores de onda y la normal

$$\vec{k}_r = \underbrace{\frac{\omega}{c}}_{k_1} n_1 (\sin r \hat{x} + \cos r \hat{y}) = \alpha_r \hat{x} - \beta_r \hat{y}, \quad (6.33)$$

$$\vec{k}_t = \underbrace{\frac{\omega}{c}}_{k_2} n_2 (\sin t \hat{x} - \cos t \hat{y}) = \alpha_t \hat{x} - \beta_t \hat{y}. \quad (6.34)$$

Y para resolver el problema dinámico de encontrar las amplitudes de las ondas reflejada y transmitida, habría que encontrar antes las componentes  $\alpha_r$  y  $\alpha_t$  de los vectores de onda  $\vec{k}_r$  y  $\vec{k}_t$ . Para eso, reescribimos (6.31) como

$$A_{0i} e^{i \alpha_i x} = -A_{0r} e^{i \alpha_r x} + A_{0t} e^{i \alpha_t x}. \quad (6.35)$$

El lado izquierdo de esta igualdad es una función periódica de la variable  $x$ , con período  $2\pi/\alpha_i$ . ¿Una función perfectamente periódica en todo el intervalo  $-\infty < x < \infty$  igual a la suma de dos funciones con las mismas características pero con períodos distintos? Si las amplitudes no son

cero, la única manera de que esto suceda es que los períodos sean todos iguales y para eso todos los  $\alpha$  deben ser iguales <sup>2</sup>

$$\alpha_i = \alpha_r = \alpha_t, \quad (6.36)$$

igualdad que equivale a la ley de la reflexión,  $i = r$ , y de Ibn-Sahl,  $n_1 \sin i = n_2 \sin t$ . Como antes, la misma conclusión se obtendría considerando condiciones de contorno del tipo “continuidad de la función multiplicada por un coeficiente” u otro tipo de condición de contorno lineal, en vez de la “continuidad de la función” considerada en (6.35) y por este motivo concluimos que para el caso de una onda plana incidente, las leyes de la DG son todas consecuencias necesarias de la existencia de condiciones de contorno lineales en un plano infinito, independientemente de la naturaleza (acústica, electromagnética, mecánica, etc.) de la onda.

### 6.6.3. Conservación de componentes tangenciales de $\vec{k}$

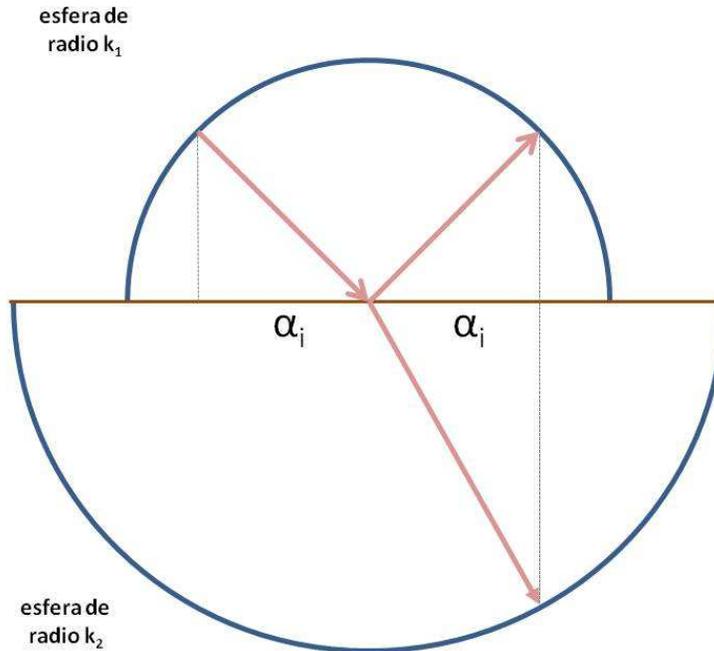


Figura 6.11: Construcción en el espacio de los vectores de onda,  $n_1 < n_2$ .

Vemos que todas las leyes de la descripción geométrica se pueden obtener como una consecuencia natural de la existencia de condiciones de contorno lineales en un plano infinito. Las ecuaciones (6.32) y (6.36) se pueden reformular como una ley de conservación de las componentes

<sup>2</sup>Otra manera de decir lo mismo es notar que el conjunto de funciones  $e^{i\alpha x}$  etiquetadas por  $\alpha$  son funciones de  $x$  linealmente independientes. Y entonces, para que se cumpla (6.35) en todo punto de la superficie de separación, solamente existen dos posibilidades: i) todos los coeficientes de la combinación lineal son cero, o ii) todos los exponentes (todos los  $\alpha$ ) son iguales.

tangenciales de los vectores de onda. En cursos superiores se verá que esta ley de conservación es una consecuencia de la simetría que tiene un plano infinito, que resulta invariante frente a traslaciones infinitesimales. No como los techos de chapa acanalada, que tienen ondulaciones periódicas y que entonces son invariantes frente a traslaciones en la dirección perpendicular a las canaletas sólo si son traslaciones finitas e iguales a la distancia entre dos canaletas, mientras que en la dirección de las canaletas si son invariantes frente a traslaciones infinitesimales. Si una de estas chapas actuara como espejo, se verían cosas raras, porque no se conservan todas las componentes tangenciales de los vectores de onda.

Trabajando en el espacio de los vectores de onda, la ley de conservación de las componentes tangenciales de los vectores de onda facilita la visualización del movimiento de los rayos reflejado y transmitido cuando cambia la inclinación del rayo incidente. Todos los vectores de onda involucrados en el ajuste de condiciones de contorno, tienen módulo impuesto por la ecuación de dispersión en cada medio. Los vectores de la onda incidente y la reflejada tienen que estar en la esfera superior (ver Fig. 6.11) y de la onda transmitida en la esfera inferior. Y como todos tienen que tener componente tangencial  $\alpha_i$ , quedan fijas las direcciones de propagación. La Fig. 6.11 muestra el caso  $n_1 < n_2$ , mientras que la Fig. 6.12 el caso  $n_1 > n_2$  justo cuando el ángulo de incidencia es igual al crítico de reflexión total, cuando la componente tangencial del vector de onda incidente es igual al radio de la esfera inferior. ¿Qué pasa cuando el ángulo de incidencia es mayor al ángulo crítico de reflexión total? ¿Cómo se dibuja en la esfera inferior un vector de onda con una componente horizontal mayor que el radio de la esfera?

#### 6.6.4. La reflexión total en el marco ondulatorio

Si  $n_1 > n_2$ ,  $\sin \theta_t > \sin \theta_i$  y entonces existe un ángulo de incidencia crítico  $\theta_{ic}$  para el cual la onda refractada se propaga en dirección paralela a la superficie

$$\sin \theta_{ic} = \frac{n_2}{n_1}. \quad (6.37)$$

En función de  $\theta_{ic}$ , la ley de Snell se escribe

$$\sin \theta_t = \frac{\sin \theta_i}{\sin \theta_{ic}}, \quad (6.38)$$

que muestra que

$$\begin{aligned} \theta_i \leq \theta_{ic} &\Rightarrow \sin \theta_t \leq 1 \Rightarrow \cos \theta_t = \sqrt{1 - \left(\frac{\sin \theta_i}{\sin \theta_{ic}}\right)^2} & (\theta_t \in \mathbf{R}), \\ \theta_i > \theta_{ic} &\Rightarrow \sin \theta_t > 1 \Rightarrow \cos \theta_t = i\sqrt{\left(\frac{\sin \theta_i}{\sin \theta_{ic}}\right)^2 - 1} & (\theta_t \notin \mathbf{R}). \end{aligned} \quad (6.39)$$

¿Un ángulo de transmisión complejo? Seguro que las descripciones anteriores no contemplaban esta novedad. Sólo el tratamiento ondulatorio muestra que para ángulos de incidencia mayores que el crítico tenemos que revisar la interpretación de  $\theta_t$  como ángulo de refracción. El significado queda en evidencia cuando investiguemos la estructura de la onda refractada en el medio 2 ( $y < 0$ )

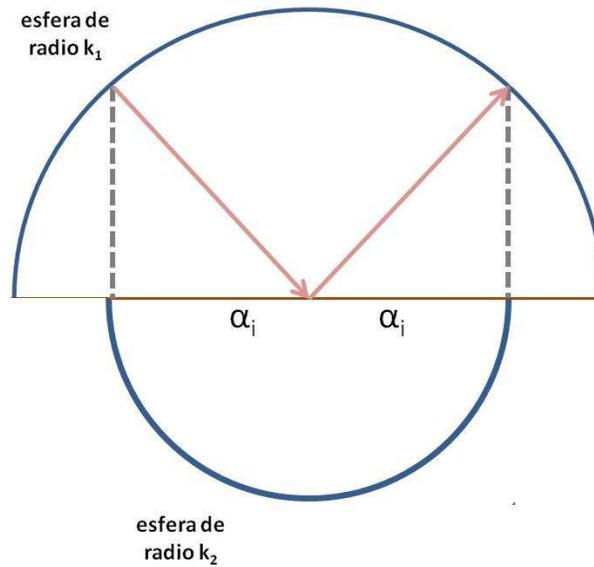


Figura 6.12: Construcción en el espacio de los vectores de onda,  $n_1 > n_2$ , ángulo crítico.

$$\begin{aligned}
 \psi_t(\vec{r}, t) &= A_{0t} e^{i\left[\frac{\omega}{c}n_2(x \sin \theta_t - y \cos \theta_t) - \omega t\right]} \\
 &= A_{0t} e^{\left(\frac{\omega}{c}n_2 \sqrt{\left(\frac{\sin \theta_i}{\sin \theta_{ic}}\right)^2 - 1}\right)y} e^{i\left[\left(\frac{\omega}{c}n_2 \sin \theta_t\right)x - \omega t\right]}.
 \end{aligned} \tag{6.40}$$

Esta expresión, válida para  $y \leq 0$ , representa una onda que se propaga a lo largo de la superficie sin atenuarse y cuya amplitud decae exponencialmente hacia el interior del medio 2 (notar que esto está relacionado con la elección de la raíz cuadrada en (6.39)). Esta onda se llama **onda evanescente** y es como las que aparecieron en el régimen reactivo de la ecuación de Klein-Gordon, página 107. La distancia de decaimiento  $\delta$  de la onda transmitida resulta

$$\frac{1}{\delta} = \frac{\omega}{c}n_2 \sqrt{\left(\frac{\sin \theta_i}{\sin \theta_{ic}}\right)^2 - 1},$$

es decir

$$\delta = \frac{\lambda_1}{2\pi} \frac{1}{\sqrt{\sin^2 \theta_i - \sin^2 \theta_{ic}}}, \tag{6.41}$$

donde  $\lambda_1$  es la longitud de onda en el medio 1. Como en el ejemplo de péndulos acoplados, este decaimiento no está asociado a ninguna pérdida, sino a un comportamiento reactivo del medio 2 para  $\theta_i > \theta_{ic}$ .

# Capítulo 7

## Ondas transversales

El objetivo de esta parte del curso es poner en evidencia fenómenos comunes a toda **onda transversal**. Fenómenos que son consecuencia del carácter **vectorial** de la perturbación. Y que hasta ahora hemos ignorado en aras de no complicar demasiado las cosas. Por este motivo, en los ejemplos previos hemos prestado atención a solamente a una de las dos componentes posibles de una onda transversal. Un ejemplo de esta afirmación se ve en la figura 3.2, allí se estudiaron las ecuaciones de movimiento de una cadena lineal para apartamientos del equilibrio solamente en el plano de la figura, pero no en el plano perpendicular. Veremos a continuación qué novedades trae considerar la presencia simultánea de las **dos** componentes posibles de una onda transversal. Ya sabemos que para ecuaciones de onda lineales basta con estudiar ondas con dependencia temporal armónica. Y dijimos, sin demostrar, que una consecuencia del análisis de Fourier espacial es que casi cualquier perturbación se pueda escribir como combinación lineal de ondas planas armónicas. Así que . . . ondas planas armónicas **transversales** pasamos a estudiar.

### 7.1. Ondas planas armónicas transversales

Si el medio que soporta la onda es isótropo y no hay direcciones “especiales”, entonces ambas componentes de la onda transversal satisfacen la misma ecuación (4.18), con el mismo valor de  $v$  para las dos componentes. Esto equivale a decir que la perturbación satisface la ecuación (4.18), pero reemplazando  $\psi$  por  $\vec{\psi}$ . En un medio anisótropo, en cambio, hay direcciones “especiales” y a la componente de  $\vec{\psi}$  en una dirección “especial” le corresponde en general una velocidad de fase distinta que a la componente de  $\vec{\psi}$  en otra dirección “especial”.

Si elegimos que la dirección de propagación de la onda coincida con el eje  $z$ , entonces la perturbación  $\vec{\psi}$  debe estar en el plano perpendicular a  $\hat{z}$ . Luego, la expresión más general para una onda plana transversal con dependencia temporal armónica se tiene que formar a partir de soluciones como la obtenida en la ecuación (4.21):

$$\vec{\psi}(z, t) = \psi_x(z, t) \hat{x} + \psi_y(z, t) \hat{y}, \quad (7.1)$$

una solución como en (4.21) para la componente  $\hat{x}$  con amplitud  $A_x$  y otra solución de la misma

forma para la componente  $\hat{y}$  con amplitud  $A_y$

$$\psi_x(z, t) = A_x \cos(kz - \omega t), \quad (7.2)$$

$$\psi_y(z, t) = A_y \cos(kz - \omega t + \phi), \quad (7.3)$$

donde  $\phi$  es el desfase entre ambas componentes. Esta forma siempre es válida con una elección adecuada del origen de  $t$  o de  $z$ , porque si en vez de la expresión anterior hubiéramos tenido

$$\psi_x(z, t) = A_x \cos(kz - \omega t + \phi_x),$$

$$\psi_y(z, t) = A_y \cos(kz - \omega t + \phi_y),$$

entonces siempre se podría definir  $t'$  de manera tal que  $-\omega t + \phi_x = -\omega t'$ . Y con el nuevo origen de tiempo el argumento del primer coseno sería igual a  $kz - \omega t'$  mientras que el argumento del segundo coseno sería igual a  $kz - \omega t' + \phi$ , con  $\phi = \phi_y - \phi_x$ . Una discusión similar vale para la elección del origen de coordenadas  $z$ .

Las eqs. (7.2) y (7.3) muestran que para definir una onda plana transversal con dependencia temporal armónica **no** son necesarios los cuatro parámetros  $A_x$ ,  $A_y$ ,  $\phi_x$  y  $\phi_y$ , si no sólo tres parámetros:  $A_x$ ,  $A_y$  y la diferencia de fase  $\phi$ . Notemos que estas ecuaciones definen dónde está la punta del vector  $\vec{\psi}(z, t)$ . Y la forma trazada por la punta del vector  $\vec{\psi}(z, t)$  es lo que se conoce como **estado de polarización** de la onda transversal. O directamente, la polarización de la onda. Como las eqs. (7.2) y (7.3) no dependen de  $z$  y de  $t$  por separado, sino de la combinación  $kz - \omega t$ , es casi equivalente (a menos de un signo) estudiar la forma trazada por la punta del vector  $\vec{\psi}$  en un lugar fijo a medida que pasa el tiempo, o la forma trazada por la punta del vector  $\vec{\psi}$  en un tiempo fijo como función de  $z$ . Veamos primero los estados de polarización que son fácilmente visualizables (y más tarde veremos que todos los casos restantes se reducen a éstos cuando se rota convenientemente el sistema de ejes  $x - y$  alrededor del eje  $z$ ).

## 7.2. Polarización lineal

Cuando  $\phi$  es un múltiplo par de  $\pi$  ( $\phi = 2m\pi$ , con  $m$  entero), las perturbaciones componentes  $\psi_x$  y  $\psi_y$  pasan por cero juntas, crecen juntas, disminuyen juntas y se hacen negativas juntas. Pensando en composición de vectores, esto quiere decir que la forma trazada por la punta del vector  $\vec{\psi}(z, t)$  en un  $z$  fijo a medida que pasa el tiempo es una recta. Y que en un dado instante de tiempo, todos los vectores  $\vec{\psi}(z, t)$  en distintos lugares están contenidos en un plano que contiene al eje  $z$  y que pasa por el primer y tercer cuadrante del plano  $x - y$ . Si necesitamos convencernos con algo más matemático, entonces dividimos miembro a miembro las ecuaciones (7.2) y (7.3) y queda que

$$\psi_y(z, t) = \frac{A_y}{A_x} \psi_x(z, t), \quad (7.4)$$

que dice que el ángulo que el vector  $\vec{\psi}(z, t)$  forma con el eje  $x$  no depende ni de  $z$  ni de  $t$ . En la visualización a  $z$  fijo diríamos que la punta del vector  $\vec{\psi}$  recorre una recta, por eso esta polarización se llama **lineal** y decimos que la onda está **linealmente polarizada**. En la visualización a  $t$  fijo diríamos que si sacamos una foto, los vectores  $\vec{\psi}$  están en un plano, por eso se dice que una onda con esta polarización está **plano polarizada**. Notar que según las

(7.2) y (7.3), la onda plana transversal con dependencia temporal armónica más general es una combinación lineal de dos ondas planas linealmente polarizadas en direcciones ortogonales.

Cuando  $\phi$  es un múltiplo impar de  $\pi$  ( $\phi = (2m + 1)\pi$  con  $m$  entero) las conclusiones son muy similares a las del párrafo anterior. Las dos perturbaciones,  $\psi_x$  y  $\psi_y$ , pasan por cero juntas, crecen en valor absoluto juntas, y en valor absoluto disminuyen juntas. Pero cuando una es positiva la otra es negativa. Pensando en composición de vectores, esto quiere decir que la forma trazada por la punta del vector  $\vec{\psi}(z, t)$  en un  $z$  fijo a medida que pasa el tiempo también es una recta. Y que en un dado instante de tiempo, todos los vectores  $\vec{\psi}(z, t)$  en distintos lugares también están contenidos en un plano que contiene al eje  $z$ . Pero ahora el plano no pasa por el primer y tercer cuadrante del plano  $x - y$ , sino por el segundo y el cuarto. Esto se ve al dividir miembro a miembro las dos ecuaciones (7.2) y (7.3), que ahora dan

$$\psi_y(z, t) = -\frac{A_y}{A_x} \psi_x(z, t) , \quad (7.5)$$

que de nuevo dice que el ángulo que el vector  $\vec{\psi}(z, t)$  forma con el eje  $x$  no depende ni de  $z$  ni de  $t$  y por eso este caso también corresponde a una **polarización lineal**.

Otra manera de analizar el estado de polarización es a través del concepto de **vector amplitud**. En los dos casos considerados, el vector  $\vec{\psi}(z, t)$  se puede escribir de la siguiente manera

$$\vec{\psi}(z, t) = (A_x \hat{x} \pm A_y \hat{y}) \cos(kz - \omega t) = \vec{A}_{\pm} \cos(kz - \omega t) , \quad (7.6)$$

que no es otra cosa que un vector cuyo módulo oscila armónicamente, tanto en el tiempo como en la dirección de propagación, pero que siempre apunta en la dirección del vector fijo  $\vec{A}_{\pm}$ . Como los vectores  $\vec{A}_+$  y  $\vec{A}_-$  son linealmente independientes, no es difícil justificar que con estos vectores se puede formar una base de perturbaciones para escribir cualquier perturbación transversal como combinación lineal de esta base.

### 7.3. Polarización elíptica

Si  $\phi$  es múltiplo impar de  $\pi/2$  y  $A_y \neq A_x$ , el coseno que aparece en  $\psi_y(z, t)$  en (7.3) se transforma, a menos de un signo, en un seno. Entonces las dos perturbaciones  $\psi_x$  y  $\psi_y$  no pasan por cero juntas: cuando una es cero la otra es máxima en valor absoluto. Es decir que hay instantes donde la punta del vector  $\vec{\psi}(z, t)$  está dirigida totalmente según un eje, mientras que en otros instantes la punta del vector  $\vec{\psi}(z, t)$  está dirigida totalmente según el eje perpendicular al primero. Estos casos corresponden a **polarización elíptica**, llamada así porque la forma trazada por la punta del vector  $\vec{\psi}(z, t)$  responde a la ecuación

$$1 = \frac{\psi_x^2(z, t)}{A_x^2} + \frac{\psi_y^2(z, t)}{A_y^2} , \quad (7.7)$$

que es una elipse con semiejes  $A_x$  y  $A_y$ . El caso particular  $A_x = A_y$  corresponde a una circunferencia.

En la visualización a  $z$  fijo la punta del vector  $\vec{\psi}$  describe una elipse y está rotando, en sentido horario o en sentido antihorario. El sentido de la rotación depende de que los desfases sean

congruentes con  $+\pi/2$  o con  $-\pi/2$ . En la visualización a  $t$  fijo diríamos que si sacamos una foto, los vectores  $\vec{\psi}$  estarán en una especie de sacacorchos elíptico, algunos como los sacacorchos de la vida real, y otros como su imagen especular, justamente dependiendo del sentido de giro.

Para definir el **vector amplitud** en el caso de elipses centradas en los ejes, conviene pasar los cosenos a la notación exponencial, es decir,

$$\begin{aligned}\psi_x(z, t) &= \Re A_x \exp(kz - \omega t), \\ \psi_y(z, t) &= \Re A_y \exp(kz - \omega t \pm \pi/2),\end{aligned}$$

y factorizando las exponenciales, se puede expresar el vector  $\vec{\psi}(z, t)$  como

$$\vec{\psi}(z, t) = \Re (A_x \hat{x} \pm i A_y \hat{y}) = \Re \vec{A}_{\pm} \exp(kz - \omega t). \quad (7.8)$$

Con la notación exponencial, donde es costumbre no escribir  $\Re$ , tenemos un **vector amplitud compleja**  $\vec{A}_{-}$  para la onda elípticamente polarizada que cuando se observa a  $z$  fijo corresponde a una onda que gira en sentido antihorario, es decir para  $\phi$  congruente a  $-\pi/2$ . Análogamente, el vector amplitud compleja  $\vec{A}_{+}$  representa una onda elípticamente polarizada que en cada  $z$  gira en sentido horario, es decir para  $\phi$  congruente a  $\pi/2$ . Como los vectores  $\vec{A}_{+}$  y  $\vec{A}_{-}$  en este caso también son linealmente independientes, no es difícil justificar que con estos vectores se puede formar una base de perturbaciones para escribir cualquier perturbación transversal como combinación lineal de ondas elípticamente polarizadas en sentidos opuestos.

## 7.4. Polarización circular

Este es un caso particular del anterior, ahora la elipse (7.7) degenera en una circunferencia y se dice que la onda transversal está **circularmente polarizada**. Notar que el valor absoluto de la perturbación  $\vec{\psi}(z, t)$  ahora permanece constante en todo lugar y tiempo, aunque su dirección rota, en un sentido o en otro dependiendo de que los desfases sean congruentes con  $+\pi/2$  o con  $-\pi/2$ . En este caso los vectores de amplitud compleja son  $\vec{A}_{\pm} = A_{\pm}(\hat{x} \pm i\hat{y})$  que también se pueden usar como elementos de una base que permite expresar cualquier perturbación transversal como combinación lineal de dos ondas circularmente polarizadas en sentidos opuestos.

Los medios formados por estructuras desordenadas son isótropos y no distinguen entre estados de polarización ortogonales. Excepto los que tienen asociado un sentido de giro, que son sensibles al sentido de giro de la polarización de las ondas electromagnéticas. En estos medios la velocidad de fase para las ondas circularmente polarizadas en un sentido es distinta a la de las que están polarizadas en sentido contrario. Si una onda linealmente polarizada atraviesa este tipo de medio, al abandonarlo habrá rotado su dirección de polarización. Para ver que esto es así, sólo hace falta escribir primero la onda linealmente polarizada en la base circularmente polarizada. Y notar que si ambas ondas circularmente polarizadas viajaran con la misma velocidad, la polarización a la salida seguiría siendo la misma que había a la entrada. Pero como cada elemento de la base circular viaja con distinta velocidad, las componentes en esta base se desfasan y esto se traduce en la rotación de la dirección de incidencia. Es el principio de funcionamiento de los polarímetros (ver video en <http://bit.ly/1PV5vvu>).

En la naturaleza abundan medios que tienen un sentido de giro asociado, son sustancias compuestas por grandes moléculas orgánicas, que tienen orientaciones espaciales en forma de hélice. El caso más conocido es el de la molécula de ácido desoxirribonucleico (ADN), ver Figura 7.1.

La mayoría de proyecciones 3D actuales usan luz circularmente polarizada en ambos sentidos para poder brindar imágenes levemente diferentes a cada ojo y así producir en el espectador la sensación estereoscópica. Si van a ver una película 3D, prueben el efecto de superponer el cristal de sus gafas correspondiente al ojo izquierdo con el correspondiente al ojo derecho de las gafas de su vecino.

## 7.5. Caso $\phi$ arbitrario

Si el valor de  $\phi$  en la página 142 no es ninguno de los considerados hasta ahora, lo que hay que hacer es eliminar el parámetro  $kz - \omega t$  de las ecuaciones paramétricas (7.2) y (7.3), para obtener una relación, ya sea explícita (como en (7.4) y en (7.5)) o implícita (como en (7.7)), entre las componentes  $\psi_x$  y  $\psi_y$ . De la primera ecuación despejamos el coseno del parámetro

$$\cos(kz - \omega t) = \psi_x(z, t) / A_x$$

Para sacar el parámetro de la segunda ecuación, desarrollamos el coseno de la segunda ecuación como

$$\cos(kz - \omega t + \phi) = \cos(kz - \omega t) \cos \phi - \sin(kz - \omega t) \sin \phi$$

Ahora reemplazamos  $\sin(kz - \omega t)$  por

$$\sqrt{1 - [\psi_x(z, t) / A_x]^2}$$

y obtenemos una ecuación implícita que parece fea o desprolija en la primera impresión. Pero eso es porque recién se despierta y necesita un poco de ayuda. Para ayudarla a mejorar su aspecto, tratamos de sacarle esa raíz cuadrada que le queda tan mal, para eso pasamos la raíz cuadrada de una lado de la igualdad y dejamos todo el resto del otro. Y (todo sea por ayudar), elevamos al cuadrado miembro a miembro. Aquí hay que aplicar eso del cuadrado del primero, mas el cuadrado del segundo, blablabla . . . y si nos tenemos fé, luego de emprolijar un poco las cosas, obtendremos la siguiente ecuación

$$\left(\frac{\psi_x}{A_x}\right)^2 + \left(\frac{\psi_y}{A_y}\right)^2 - 2 \frac{\psi_x \psi_y}{A_x A_y} \cos \phi = \sin^2 \phi \quad (7.9)$$

que hay que reconocer que mejoró muchísimo. No es otra cosa que la ecuación de una sección cónica. Y comparando el producto de los coeficientes de los términos cuadráticos (no cruzados) con el coeficiente del término cruzado, es claro que se trata de una elipse (ver <http://bit.ly/2epLUaC>). Es fácil demostrar que el eje mayor de esta elipse forma con el eje  $x$  un ángulo  $\beta$  dado por

$$\tan 2\beta = \frac{2 A_x A_y}{A_x^2 - A_y^2} \cos \phi \quad (7.10)$$

Corolario 1: en el sistema de ejes donde los ejes de la elipse están centrados, el desfase entre componentes ortogonales es  $\pm \pi/2$  y sus congruentes módulo  $2\pi$ .

Corolario 2: para describir una polarización elíptica, conviene (si se puede) girar los ejes para que coincidan con los ejes propios de la elipse, porque así los desfases serán siempre  $+\pi/2$  o  $-\pi/2$ , dependiendo del sentido de giro.

### 7.5.1. Luz natural

La radiación electromagnética (luz incluida) proveniente de fuentes térmicas se puede describir con las eqs. (7.2) y (7.3), con la diferencia de que los cosenos, en vez de no empezar ni terminar nunca, están modulados por una función que empieza y termina, con una duración  $\tau \approx 10^{-8}$ s. A estos cosenos modulados se los conoce como “trenes de onda”. Cada tren de onda difiere del anterior en el valor de la diferencia de fase  $\phi$  y el valor de  $\phi$  cambia aleatoriamente en cada intervalo  $\tau$ . Todos los valores de  $\phi$  tienen la misma probabilidad, lo que equivale a decir que todos los estados de polarización son equiprobables en cada intervalo  $\tau$  y que cambian  $\tau^{-1} \approx 10^8$  veces en un segundo. Como los detectores comunes son incapaces de resolver cien millones de cambios en un segundo, se suele decir que una onda con polarización natural está **no polarizada**. Aunque realmente lo esté, durante tiempos muy cortos. Lo más adecuado sería llamarla polarizada **al azar** (o también polarizada a la *a mí qué me importa*). Observar que una onda plana perfectamente monocromática, es decir con una frecuencia bien definida, siempre mantiene un estado de polarización bien definido. Porque si es tren, no es monocromática, no?

Si a una onda con polarización natural se le agrega o se le resta una onda polarizada, se dice que la onda tiene un estado **parcialmente polarizado** o que la onda está parcialmente polarizada.

## 7.6. Polarización por absorción selectiva

Una manera de polarizar un haz de luz natural es mediante el uso de materiales o dispositivos que absorben selectivamente una de las componentes transversales del campo eléctrico de una onda. La absorción puede estar ligada a la estructura molecular del material o a asimetrías espaciales. Un ejemplo son las rejillas de alambres conductores (Fig. 7.2) empleadas para microondas (las microondas son una forma de radiación electromagnética, como la luz visible, pero con  $\lambda$  del orden de los centímetros). La energía de la componente del campo eléctrico paralela a los alambres se “gasta” en mover los electrones libres de los alambres, mientras que la energía de la componente del campo eléctrico perpendicular a los alambres permanece casi inalterada. De esta manera, a la salida de la rejilla se tiene una onda linealmente polarizada en dirección perpendicular a los alambres.

Otro ejemplo de polarización por absorción selectiva ocurre en ciertos cristales como la turmalina, o en las láminas Polaroid, donde las estructuras moleculares juegan un papel similar al que juegan los alambres conductores en las rejillas de microondas. En los polaroids reales hay una componente que se atenúa, aunque no completamente, la lámina es más cara cuanto mejor polariza (es decir, cuanto más atenúa). Salvo que se diga lo contrario, se supone que las láminas polarizadoras son ideales, es decir que a la salida se obtiene una onda perfectamente polarizada de manera lineal.

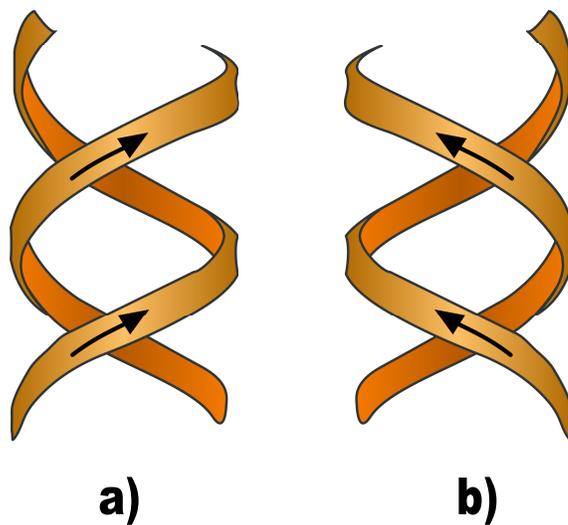


Figura 7.1: Estructuras helicoidales con sentidos de giro opuestos. Notar que la estructura de un tipo se obtiene como reflexión especular de la estructura del otro tipo (de Wikimedia Commons [https://commons.wikimedia.org/wiki/File:Levo\\_dextro.svg](https://commons.wikimedia.org/wiki/File:Levo_dextro.svg))

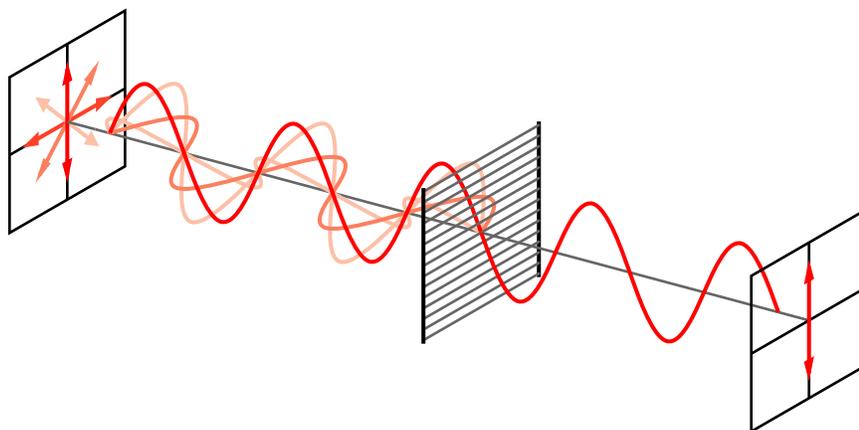


Figura 7.2: Polarizador de alambres usado para microondas (de Wikimedia Commons, archivo Wire-grid-polarizer.svg)

Si los ejes de transmisión de dos polarizadores ideales son perpendiculares, no se transmite luz. En cambio, si los ejes de transmisión de los polarizadores forman un ángulo  $\theta$ , la intensidad de la luz transmitida sigue la **ley de Malus**

$$I(\theta) = I_{max} \cos^2 \theta , \quad (7.11)$$

que surge de proyectar la amplitud de la perturbación linealmente polarizada que sale de la primera lámina en la dirección del eje de transmisión de la segunda lámina y recordar que la intensidad es proporcional al cuadrado de la amplitud.

## 7.7. Polarización por reflexión

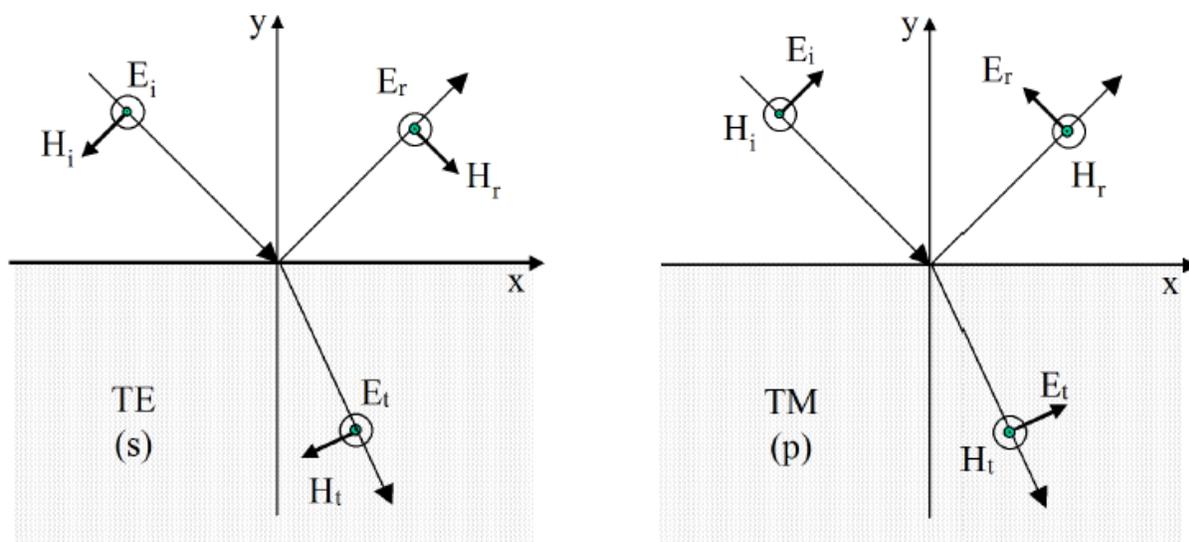


Figura 7.3: Modos de polarización en una superficie plana.  $E$  indica campo eléctrico,  $H$  indica campo magnético

Cuando una onda plana incide oblicuamente sobre una superficie plana (ver esquema de la Fig. 7.3), existe un plano especial, el plano de incidencia, donde están contenidos los rayos reflejado y transmitido. Si bien el espacio es isótropo y un plano infinito es indistinguible frente a rotaciones alrededor de un eje perpendicular, este plano de incidencia introduce una asimetría para las ondas transversales, porque la componente de la polarización incidente que está contenida en el plano de incidencia no tiene por lo general el mismo coeficiente de reflexión que la componente de la polarización incidente que es perpendicular al plano de incidencia. Una manera fácil de ver que esto es así es pensando que si la superficie reflectante fuera metálica, la componente del campo eléctrico incidente perpendicular al plano de incidencia, paralela a la superficie reflectante, es muy eficiente en mover las cargas sueltas que existen en el metal. En cambio, a la otra componente le resulta más difícil, porque no es paralela a la superficie reflectante. En fin, de todo esto mejor no decir demasiado porque a nivel F2 no se sabe nada de electromagnetismo,

que se empieza a ver en F3. Y el problema completo se resolverá recién en Física Teórica 1. Pero bueno, qué le vamos a hacer, la vida del científico es así, uno se pasa la vida completando huecos.

Lo importante es que la fracción de la potencia incidente que se lleva la onda reflejada es distinta en cada caso. Y entonces la polarización de la onda reflejada va a ser en general distinta que la polarización de la onda incidente. Los casos o “modos” de polarización, son (ver Fig. 7.3)

- polarización  $s$ , también llamada TE o transverso eléctrico, porque el campo eléctrico incidente es perpendicular al plano de incidencia. En el sistema de coordenadas de la figura es la componente  $\hat{z}$ ;
- polarización  $p$ , también llamada TM, o transverso magnético, porque el campo magnético incidente es perpendicular al plano de incidencia (y el campo eléctrico incidente está contenido en el plano de incidencia). En el sistema de coordenadas de la figura es la componente en la dirección del producto  $\hat{k}_i \times \hat{z}$ .

En la Fig. 7.4 se muestran las curvas de los coeficientes de reflexión TE y TM en función del ángulo de incidencia para una dioptra plana aire-vidrio ( $n_2 = 1.5$ ,  $n_1 = 1$ ). Estos coeficientes dan la fracción de potencia incidente que se va en la onda reflejada. Notar que ambas curvas coinciden en incidencia normal, porque en el límite de  $i \rightarrow 0$  no hay distinción entre modos TE y TM (porque no está definido el plano de incidencia). Usando las condiciones de contorno adecuadas (electromagnéticas), se puede demostrar que en incidencia normal el módulo del coeficiente de reflexión vale

$$R(i = 0) = \left| \frac{n_2 - n_1}{n_2 + n_1} \right|^2,$$

que sólo depende del índice relativo y que da lo mismo para incidencia desde el medio 1 hacia el medio 2 que para incidencia desde el medio 2 hacia el medio 1. Esta expresión muestra que la onda reflejada en la primera cara de una ventana de vidrio ( $n_2 = 1.5$ ,  $n_1 = 1$ ) se lleva un 4% de la energía incidente (y como la intensidad es proporcional al cuadrado de la amplitud, podemos decir que el módulo del campo eléctrico de la onda reflejada tiene una amplitud 0.2 veces menor que el del campo eléctrico de la onda incidente). Ambas curvas también coinciden en incidencia rasante donde la reflectividad vale 1. Este es el comportamiento usual en óptica, donde los medios son no magnéticos. La diferencia más importante entre modos TE y TM es que la reflectividad para el modo TM se anula para un ángulo de incidencia  $i_B$  conocido como **ángulo de Brewster**, mientras que la reflectividad para el modo TE nunca se anula. En el ángulo de Brewster el rayo reflejado (que en esta polarización no se refleja!) habría sido perpendicular al rayo refractado. Por eso, aplicando Ibn Sahl resulta

$$\tan i_B = \frac{n_2}{n_1}. \quad (7.12)$$

En el ángulo de Brewster, la onda con campo eléctrico en el plano de incidencia no se refleja en absoluto, sino que se refracta en su totalidad, una característica que se usa en las llamadas “ventanas Brewster”. Si la onda incidente tiene polarización natural, la luz reflejada en el ángulo de Brewster queda totalmente polarizada con campo eléctrico perpendicular al plano de incidencia. En cambio, la luz refractada queda parcialmente polarizada en dirección paralela al

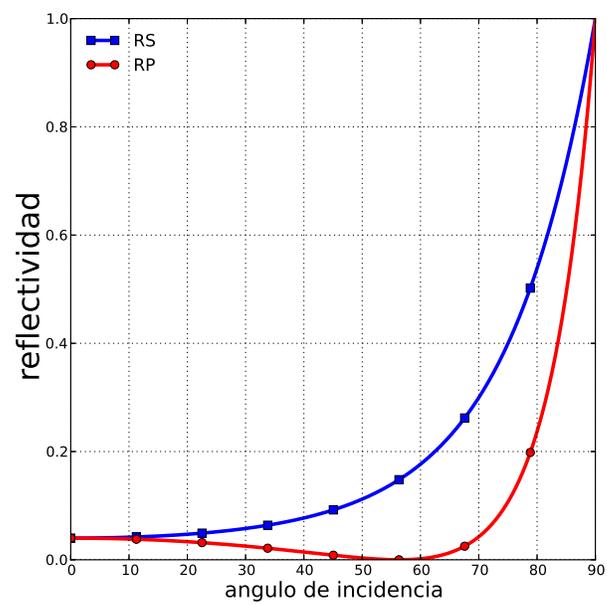


Figura 7.4: Coeficientes de reflexión en función del ángulo de incidencia para polarización  $s$  (TE) y  $p$  (TM). El medio de incidencia es aire  $n_1 = 1$ , el de transmisión es vidrio  $n_2 = 1.5$

plano de incidencia (es una mezcla de la componente TM de la onda incidente, que se refracta en su totalidad, y de una fracción de la componente TE de la onda incidente, que no se refracta en su totalidad porque una parte si se reflejó).

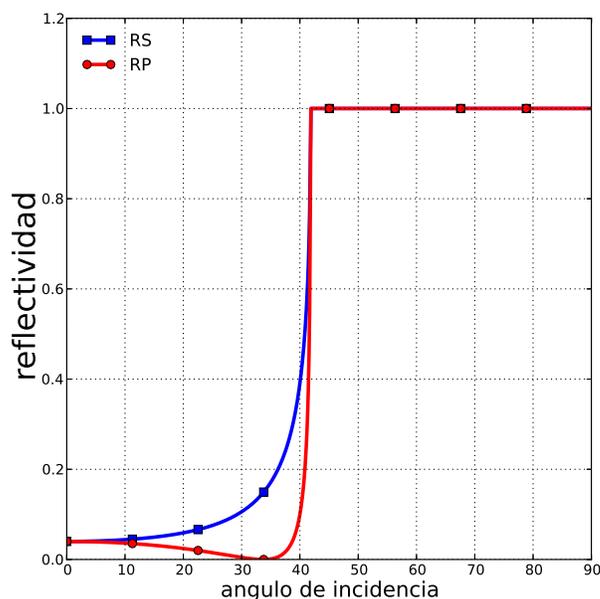


Figura 7.5: Coeficientes de reflexión en función del ángulo de incidencia para polarización  $s$  (TE) y  $p$  (TM). El medio de incidencia es es vidrio  $n_1 = 1.5$  y el de transmisión aire  $n_2 = 1$ .

Es claro que la distinta respuesta para cada componente produce el efecto de polarización por reflexión. Justamente esta es la base del diseño de filtros polarizadores en los anteojos de sol. Cuando la luz solar se refleja en una superficie horizontal, el plano de incidencia es vertical y la luz reflejada contiene preponderantemente luz polarizada en dirección horizontal. Cuando la reflexión tiene lugar en la superficie lisa de una carretera de asfalto o en la superficie de un lago, produce una reverberación indeseable que deslumbra al observador interesado en las reflexiones que provienen de otros objetos (autos, barcos). La visión mejora si se elimina esta reverberación y por eso los fabricantes de lentes colocan en posición vertical el eje de polarización del material polarizador con que están hechas las lentes de los anteojos.

La Fig. 7.5 muestra las curvas para una dioptra vidrio–aire, en vez de aire–vidrio. Notar que el resultado es cualitativamente parecido al de la Fig. 7.4, pero la reflectividad toma el valor 1 en todo el intervalo entre el ángulo crítico de reflexión total e incidencia rasante. Claramente el ángulo crítico de reflexión total tiene que ser mayor que el de Brewster (hay argumentos físicos . . . y además se ve en el resultado numérico . . . y además **se puede demostrar analíticamente** usando las expresiones conocidas para cada ángulo: hacerlo).

Sabiendo electromagnetismo se demuestra que en medios no magnéticos los coeficientes de reflexión complejos para las amplitudes vienen dados por las siguientes expresiones

$$r^s = \frac{\beta_i - \beta_t}{\beta_i + \beta_t} = r_{12}^s \quad (7.13)$$

$$r^p = \frac{\beta_i/n_1^2 - \beta_t/n_2^2}{\beta_i/n_1^2 + \beta_t/n_2^2} = r_{12}^p \quad (7.14)$$

donde se usa la notación de las ecuaciones (6.24) y (6.34) y los subíndices 1 y 2 a la derecha de las igualdades recuerdan que la reflexión ocurre cuando la onda pasa desde el medio 1 al medio 2. La reflectividad  $R(i)$  de la figura 7.4 es el módulo al cuadrado de los coeficientes  $r^s$  y  $r^p$  dados por las expresiones (7.13) y (7.14). Es más, usando los resultados de la ec. 6.39 es muy fácil comprobar que estos coeficientes son complejos cuando el ángulo de incidencia es mayor que el crítico, demostrando así lo que dije en las clases sobre los desfases introducidos en el régimen de reflexión total para cada polarización.

A continuación dejo de regalo un script de Python para calcular los coeficientes  $r^s$  y  $r^p$  y a partir de ellos la reflectividad que se muestra en las figuras 7.4 y 7.5.

```
# Coeficientes de reflexión vs ángulo de incidencia r_{12}
# Clase 21 del curso de F2
# Modos s y p pol, ver Figura 21.2 y 21.3 en apuntes
# Author R. Depine <http://users.df.uba.ar/rdep/>
#

import numpy as np
import math
import cmath
import matplotlib.pyplot as plt
pi=math.pi
Uno=1+0j

##### Entradas

ind1=1.
ind2=1.5
teta0=0.
tetaf=89.99
ihow=800

tetagrid=np.linspace(teta0,tetaf,ihow)
tetarad=tetagrid*math.pi/180

beta1=ind1*np.cos(tetarad)
beta2=np.sqrt(Uno*ind2**2-ind1**2*np.sin(tetarad)**2)
rs=(beta1-beta2)/(beta1+beta2)
rp=(beta1/ind1**2-beta2/ind2**2)/(beta1/ind1**2+beta2/ind2**2)
RS=np.absolute(rs)**2
RP=np.absolute(rp)**2

### resultados a un archivo con encabezamiento y tabla
f=open('clase21.ext','w')
print >>f, "=====
print >>f, "Script para la Figura 2 de la Clase 20 - F2 2016 2C"
print >>f, "=====
print >>f, ind1, ind2
print >>f, "ang, RS, RP"

for IJK in range(ihow):
    print >>f, "%g %g %g " %(tetagrid[IJK], RS[IJK], RP[IJK])

f.close()

### Mostramos los resultados

fig = plt.figure(figsize=(8, 8))
plt.grid()
plt.xlabel(r'ángulo de incidencia', fontsize=18)
plt.ylabel(r'reflectividad', fontsize=26)
plt.plot(tetagrid, RS, c='b', ls='-', lw=3, marker='s', markevery=100, ms=7)
plt.plot(tetagrid, RP, c='r', ls='-', lw=3, marker='o', markevery=100, ms=7)
plt.legend(['RS', 'RP'], loc='upper left', frameon=False)
plt.draw()
plt.savefig('Clase21Fig2.eps', format='eps', dpi=300)
plt.savefig('Clase21Fig2.jpg', format='jpg')
```

Para saber si es posible transformar en circularmente polarizada una onda que incide linealmente polarizada a 45 grados del plano de incidencia (en la clase anticipé que no), deberíamos poder evaluar los desfases para cada polarización en el régimen de reflexión total y a partir de ellos el desfase relativo entre las componentes s y p.

## 7.8. Láminas retardadoras

Para controlar la diferencia de fase  $\phi$  entre las dos componentes transversales de las eqs. (7.2) y (7.3), y controlar así el estado de polarización, se usan láminas fabricadas con materiales anisótropos, que tienen distinto índice de refracción para componentes de polarización paralelas o perpendiculares a un eje preferencial del material, asociado con su estructura cristalina. Si la onda descrita por las eqs. (7.2) y (7.3) incide sobre una lámina de este tipo y orientada para que sus ejes preferenciales coincidan con los ejes x e y, para describir la propagación de la onda dentro de la lámina hay que tener en cuenta que la velocidad de propagación para la componente x no es igual a la velocidad de propagación para la componente y. En consecuencia, si el número de onda  $k = \omega/c$  era el mismo para describir la propagación en el vacío de las dos componentes de la onda transversal, dentro de la lámina hay dos números de onda, uno para cada componente

$$k_x = \frac{\omega}{c}n_x, \quad k_y = \frac{\omega}{c}n_y. \quad (7.15)$$

De esta manera, mientras que en el vacío las dos componentes mantienen su diferencia de fase (en cada z mantienen su estado de polarización), en la lámina las dos componentes se van desfasando a medida que avanzan. En particular, a la salida de la lámina, luego de atravesar un espesor  $d$ , su diferencia de fase será la que tenían a la entrada ( $\phi$ ), más un desfase adicional, igual a

$$\Delta\phi = (k_y - k_x)d = \frac{\omega}{c}(n_y - n_x)d = \frac{2\pi}{\lambda_0}(n_y - n_x)d. \quad (7.16)$$

Si se elige el espesor de la lámina para que  $\Delta\phi$  sea congruente a  $\pm\pi/2$  (un cuarto de toda una vuelta), se dice que se tiene una **lámina de cuarto de onda**. Una lámina de este tipo transforma una onda linealmente polarizada a 45° de los ejes x-y en una onda circularmente polarizada y una onda elípticamente polarizada con ejes principales x-y en una onda linealmente polarizada.

Si se elige el espesor de la lámina para que  $\Delta\phi$  sea congruente a  $\pm\pi$  (media vuelta), se dice que se tiene una **lámina de media onda**. Esta lámina se puede usar para cambiar el sentido de giro de una onda elípticamente polarizada o para rotar el plano de polarización de una onda linealmente polarizada.

## 7.9. Birrefringencia en medios quirales

Así como la estructura interna de los medios anisótropos da como resultado la existencia de distintos índices de refracción para componentes de polarización en la dirección de ciertos

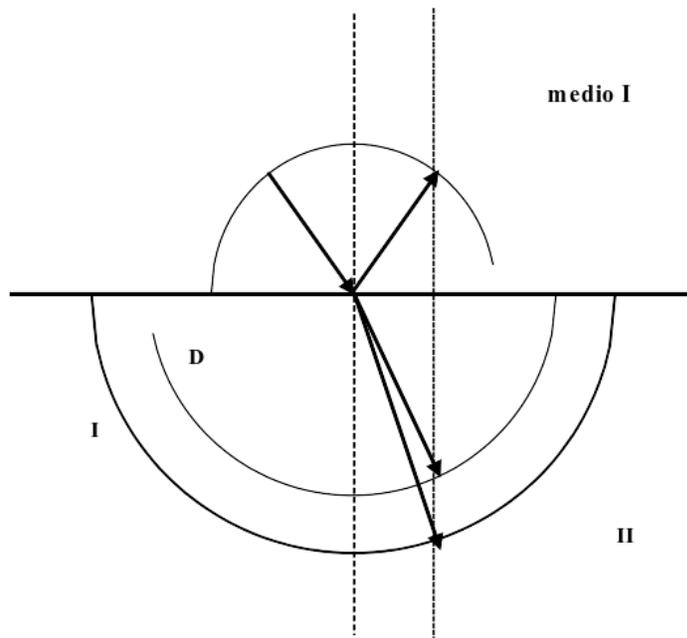


Figura 7.6: Birrefringencia en superficie aquiral-quiral, los dos índices de refracción de las ondas circularmente polarizadas izquierda y derecha son mayores que el índice de refracción del medio de incidencia.

ejes preferenciales, en la naturaleza existen medios cuya estructura interna tiene un sentido de giro privilegiado y esto produce que las ondas polarizadas circularmente con un sentido de giro tengan velocidades de propagación distintas que las que tienen sentido de giro opuesto. Este tipo de medios son de interés en la física de polímeros, en la industria farmacéutica o en química biológica, ya que las grandes moléculas orgánicas tienen generalmente asociado un sentido de giro.

Que las ondas polarizadas circularmente con sentidos de giro opuestos tengan distinta velocidad de propagación es una propiedad común a todo medio constituido microscópicamente por objetos que no tienen simetría de reflexión. Los medios formados por hélices izquierdas o derechas son un ejemplo, pero lo mismo ocurre con cualquier otro medio formado por colecciones de objetos tales que cada objeto difiere de su imagen especular. La palabra **quiral** tiene raíz griega, y proviene de **mano** (la mano izquierda no coincide con su imagen especular, que es la mano derecha). Este fenómeno es la base de lo que se conoce como actividad óptica: una onda linealmente polarizada entra en un medio quiral, se propaga, y al salir del mismo se comprueba que ha rotado su dirección de vibración (ver video en <http://bit.ly/1PV5vvu>).

Mediante un análisis similar al que hicimos en 6.6.3, ver figuras 6.11 y 6.12, es fácil ver que en la superficie plana entre un medio común (aquiral) y otro quiral, una onda incidente cuya dirección de propagación forma un ángulo no nulo con la normal a la superficie, da origen, por lo general, a dos ondas transmitidas que viajan en direcciones distintas. La existencia de dos ondas transmitidas para una misma onda incidente se conoce como **birrefringencia**. Cuando el medio tiene un sentido de giro derecho, la onda circularmente polarizada derecha se propaga

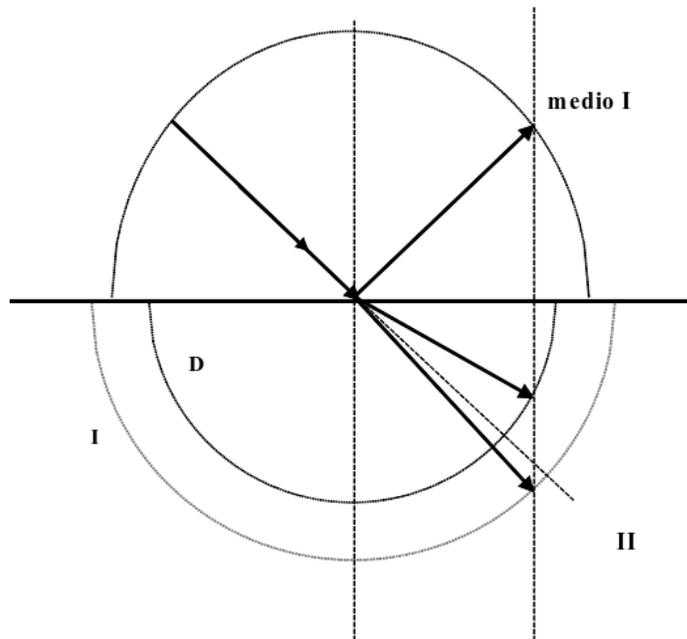


Figura 7.7: Birrefringencia en superficie aquiral-quiral, el valor del índice de refracción del medio de incidencia es intermedio entre los dos índices de refracción de las ondas circularmente polarizadas.

más rápido que la izquierda:  $v_D > v_I$ , en consecuencia  $k^{(I)} > k^{(D)}$ . Inversamente, cuando el medio tiene un sentido de giro izquierdo, la onda circularmente polarizada izquierda se propaga más rápido que la derecha:  $v_I > v_D$ , en consecuencia  $k^{(D)} > k^{(I)}$ . El punto importante es darse cuenta que en el medio quiral puede haber dos ondas circularmente polarizadas con sentidos de rotación opuestos. Esto se indica en las figuras mediante las esferas D e I. Las direcciones de propagación de estas dos ondas transmitidas quedan determinadas como en el caso de medios comunes, solamente pidiendo la conservación de las componentes tangenciales de los vectores de onda.

Dependiendo del tipo de medios, pueden darse situaciones en las que:

- la onda incidente da lugar a dos ondas transmitidas; una circular izquierda y otra circular derecha y las direcciones de ambas (respecto de la dirección incidente) se acercan a la normal. En este caso, que es el mostrado en la figura anterior, siempre hay dos ondas transmitidas para todo ángulo de incidencia.
- la onda incidente da lugar a dos ondas transmitidas; pero con respecto a la dirección incidente la onda circularmente polarizada izquierda se acerca a la normal, mientras que la onda circularmente polarizada derecha se aleja de la normal. En este caso, existe un ángulo de incidencia crítico a partir del cual la onda circularmente polarizada derecha se transforma en evanescente (pero la izquierda nunca se hace evanescente).
- la onda incidente da lugar a dos ondas transmitidas y tanto la circularmente polarizada

izquierda como la derecha se alejan de la normal con respecto a la dirección incidente. En este caso, existe un ángulo crítico de incidencia a partir del cual la onda circularmente polarizada derecha se transforma en evanescente (y la izquierda no evanescente) y otro ángulo crítico mayor a partir del cual ambas ondas transmitidas se transforman en evanescentes (reflexión total).

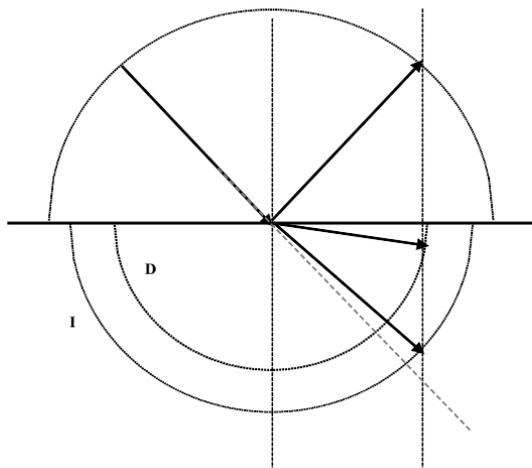


Figura 7.8: Birrefringencia en superficie aquiral-quiral, el valor del índice de refracción del medio de incidencia es mayor que los dos índices de refracción de las ondas circularmente polarizadas.

## 7.10. Birrefringencia en cristales

En la clase vimos que la existencia de dos ondas transmitidas para una misma onda incidente también ocurre en la superficie plana entre un medio isótropo común y un medio anisótropo como el que se usa para construir láminas retardadoras. Las láminas retardadoras se usan en incidencia normal y por eso la birrefringencia no es evidente. Pero cuando la incidencia no es normal, se observan dos rayos transmitidos dentro del cristal y además las ondas asociadas a estos rayos están linealmente polarizadas en direcciones ortogonales, una propiedad que se usa para construir dispositivos polarizadores (ver, por ejemplo prismas de Nicol y de Wollaston, sección 8.4 del libro de Hecht y Zajac).

### Para pensar

¿Cómo se debe proceder experimentalmente para determinar unívocamente el estado de polarización de una muestra incógnita?

# Capítulo 8

## Interferencia

La ecuación de ondas clásica es lineal, entonces vale el principio de superposición. Esto quiere decir que si en el punto P estamos observando la perturbación producida por dos fuentes  $S_1$  y  $S_2$  (ver Fig. 8.1, la perturbación medida tiene que ser igual a la suma de la perturbación producida cuando la fuente  $S_1$  estaba encendida y la fuente  $S_2$  apagada, más la perturbación producida cuando la fuente  $S_2$  estaba encendida y la fuente  $S_1$  apagada. La superposición de dos o más ondas en una región del espacio da lugar a una onda resultante que es la suma de las perturbaciones individuales. Cuando esto sucede se dice que hay **interferencia**. Desde este punto de vista, siempre se produce interferencia entre dos o más ondas regidas por ecuaciones de onda lineales. Pero con luz natural no es fácil que esta interferencia pueda ser detectada por los sentidos humanos, probablemente por eso se tardó tanto tiempo en evidenciar experimentalmente el carácter ondulatorio de la luz.

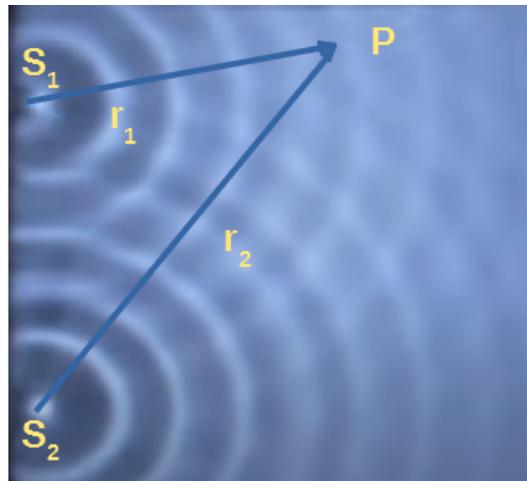


Figura 8.1: Esquema conceptual de la superposición con dos fuentes puntuales.

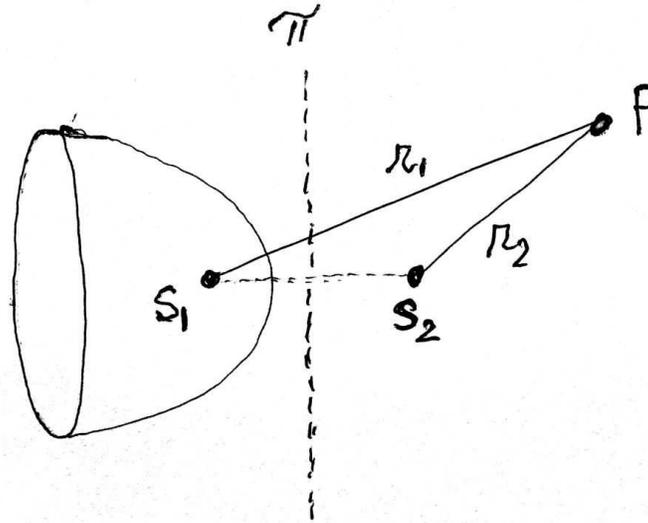


Figura 8.2: Las perturbaciones de dos fuentes puntuales interfieren en el punto P.

## 8.1. Fuentes puntuales

Empezamos con el caso más sencillo, analicemos la interferencia entre las perturbaciones producidas por dos fuentes puntuales monocromáticas de frecuencia  $\omega$ . Cada fuente produce una onda esférica progresiva, como se vió en la eq. (4.28). Al punto P llegan dos perturbaciones de la forma

$$\begin{aligned}\psi_1(r_1, t) &= a_1 \cos(\omega t - kr_1 + \delta_1) = a_1 \cos(\omega t + \Delta_1), \\ \psi_2(r_2, t) &= a_2 \cos(\omega t - kr_2 + \delta_2) = a_2 \cos(\omega t + \Delta_2),\end{aligned}$$

donde  $\Delta_1 = \delta_1 - kr_1$ ,  $\Delta_2 = \delta_2 - kr_2$ ,  $k = \omega/v$ ,  $r_1$  y  $r_2$  se miden desde cada fuente al punto P,  $a_1$  y  $a_2$  dependen de las intensidades de cada fuente y de  $r_1$  y de  $r_2$  y  $\delta_1$  y  $\delta_2$  son fases iniciales que se introducen para tener en cuenta que las fuentes se pueden encender en tiempos distintos y en consecuencia que pueden estar desfasadas.

La perturbación total en el punto P es

$$\psi_P(t) = \psi_1(r_1, t) + \psi_2(r_2, t)$$

y tiene que tener la forma

$$\psi_P(t) = a \cos(\omega t + \Delta). \quad (8.1)$$

Para ver que esto es así basta con encontrar la amplitud  $a$  y la fase inicial  $\Delta$  de la onda resultante en función de amplitudes  $a_1$  y  $a_2$  y fases iniciales  $\Delta_1$  y  $\Delta_2$  de las ondas individuales. Notemos por un lado que  $\psi_1(r_1, t) + \psi_2(r_2, t)$  se puede escribir como la suma de dos términos: uno que varía como  $\cos \omega t$  y otro que varía como  $\sin \omega t$ . Por otro lado,  $\psi(r, t)$  dado por la ec. (8.1) también se puede escribir de la misma manera. Entonces, para que la superposición valga en

todo instante, el coeficiente del término que varía como  $\cos \omega t$  en  $\psi_1(r_1, t) + \psi_2(r_2, t)$  tiene que ser igual al coeficiente del término que varía como  $\cos \omega t$  en  $\psi(r, t)$ . Y lo mismo tiene que pasar para los coeficientes de los términos que varían como  $\sin \omega t$ . Al igualar estos coeficientes se obtienen las siguientes condiciones

$$\begin{aligned} a \cos \Delta &= a_1 \cos \Delta_1 + a_2 \cos \Delta_2, \\ a \sin \Delta &= a_1 \sin \Delta_1 + a_2 \sin \Delta_2, \end{aligned}$$

un sistema de ecuaciones con el que estamos familiarizados, hay que elevar al cuadrado miembro a miembro cada ecuación y luego sumar para obtener

$$a^2 = a_1^2 + a_2^2 + 2a_1 a_2 \cos \phi, \quad (8.2)$$

donde la diferencia de fase  $\phi$  entre las dos perturbaciones en el punto P

$$\phi = \Delta_2 - \Delta_1 = \delta_2 - \delta_1 + k(r_1 - r_2), \quad (8.3)$$

tiene dos aportes: uno  $(\delta_2 - \delta_1)$  proviene de las fases iniciales de las fuentes, y el otro  $(k(r_1 - r_2))$  proviene de la diferencia de camino óptico (porque  $k$  es proporcional al índice de refracción) entre los caminos recorridos por las dos perturbaciones. Análogamente, se podría hallar  $\Delta$  en términos de  $a_1$ ,  $a_2$ ,  $\Delta_1$  y  $\Delta_2$ , pero por el momento no nos hace falta.

Como en todo fenómeno ondulatorio, la intensidad en el punto P,  $I_P$ , es proporcional a  $a^2$ , el cuadrado de la amplitud de la perturbación resultante en dicho punto. En cambio, cuando solo está encendida la fuente 1, la intensidad  $I_1$  que se mide en el punto P es proporcional a  $a_1^2$  y cuando solo está encendida la fuente 2, la intensidad  $I_2$  en el punto P es proporcional a  $a_2^2$ . Entonces, en términos de las intensidades, la ec. (8.2) se puede reescribir como

$$I_P = I_1 + I_2 + 2\sqrt{I_1 I_2} \cos \phi \quad (8.4)$$

que dice que la intensidad en el punto P cuando están encendidas las dos fuentes simultáneamente **no es la suma de las intensidades** producidas por cada fuente por separado. Esto es así porque, además de la suma de intensidades individuales (proporcionales a  $a_1^2$  y a  $a_2^2$ ), hay un tercer término que puede ser positivo o negativo dependiendo del valor de  $\phi$ . Este tercer término se llama **término de interferencia**. O sea, que la intensidad en el punto P tiene máximos y mínimos dependiendo del valor de la diferencia de fase  $\phi$  entre las dos perturbaciones en dicho punto. Cuando  $\phi$  es congruente con  $2\pi$  el término de interferencia toma su máximo valor y entonces la intensidad en el punto P toma el máximo valor posible, proporcional a

$$a_{max}^2 = a_1^2 + a_2^2 + 2a_1 a_2 = (a_1 + a_2)^2. \quad (8.5)$$

En este caso se dice que la interferencia es **constructiva**. En cambio, cuando  $\phi$  es congruente con  $\pi$ , el término de interferencia toma su mínimo valor, la intensidad en el punto P toma el mínimo valor posible, proporcional a

$$a_{min}^2 = a_1^2 + a_2^2 - 2a_1 a_2 = (a_1 - a_2)^2, \quad (8.6)$$

y se dice que la interferencia es **destructiva**. Notar que en el caso  $a_1 = a_2$ , los mínimos valen cero y podemos decir que “luz + luz = oscuridad”. Suena raro, porque no es usual observar que luz + luz da oscuridad. Y esto pasa porque estamos habituados a un mundo donde las

fuentes de luz son fuentes térmicas, como las que mencionamos cuando hablamos de luz natural en la página 146. Y entonces, en vez de producir dos perturbaciones  $\psi_1(z, t)$  y  $\psi_2(z, t)$  como supusimos, en forma de coseno con fases  $\Delta_1$  y  $\Delta_2$  constantes en el tiempo, cada fuente térmica produce “trenes de onda”, es decir, cosenos modulados por una función con una duración  $\tau$ , que para la luz suele ser  $\approx 10^{-8}$ s. Por este motivo, la diferencia de fase  $\Delta_2 - \Delta_1$  cambia aleatoriamente en cada intervalo  $\tau$  y entonces hay  $\tau^{-1} \approx 10^8$  cambios de intensidad en un segundo, algo imposible de detectar con detectores comunes o a simple vista. Por este motivo se dice que las fuentes térmicas son **incoherentes**, en el sentido de que no pueden exhibir fenómenos de interferencia como los predichos por las ecs. (8.5) y (8.6). Es claro que esto es una convención impuesta por la resolución de los aparatos a los que estamos acostumbrados, porque si pudiéramos resolver tiempos muy pequeños, estaríamos mas familiarizados con las variaciones de intensidad producidas por interferencia.

## 8.2. Uso de fasores

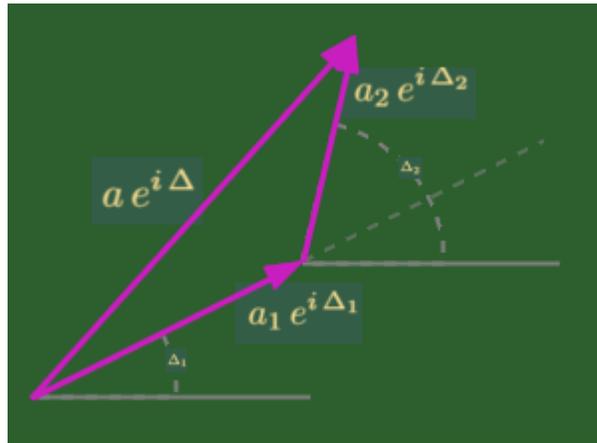


Figura 8.3: La suma de fasores es equivalente a la suma de dos perturbaciones armónicas.

Una vez entendido el procedimiento para encontrar la resultante de la suma de las dos perturbaciones armónicas usadas en (8.1), veamos cómo se puede simplificar el álgebra, pagando el precio de introducir notación compleja. Escribiendo los cosenos como la parte real de exponenciales imaginarias, obtenemos

$$\begin{aligned}\psi_1(r_1, t) &= a_1 \cos(\omega t + \Delta_1) = \text{Re } a_1 e^{i(\omega t + \Delta_1)}, \\ \psi_1(r_2, t) &= a_2 \cos(\omega t + \Delta_2) = \text{Re } a_2 e^{i(\omega t + \Delta_2)}.\end{aligned}$$

Con esta notación, y omitiendo  $\text{Re}$  (pero recordando que habrá que tomar la parte real de los resultados que obtengamos), la suma que hicimos en notación trigonométrica se puede expresar como

$$\psi_P(t) = \sum_{n=1}^2 a_n e^{i(\omega t + \Delta_n)} = e^{i\omega t} \sum_{n=1}^2 a_n e^{i\Delta_n} = a e^{i\Delta} e^{i\omega t}. \quad (8.7)$$

Observando que el factor  $e^{-i\omega t}$  está en todas partes, lo simplificamos y vemos que se puede interpretar la suma de oscilaciones armónicas como suma de dos números complejos, cada uno

con su módulo y su fase, y que el resultado es otro número complejo, con módulo  $a$  y fase  $\Delta$ . Y usando la representación de los números complejos como vectores del plano (el plano complejo), tenemos que el esquema de la sección anterior es completamente equivalente al esquema mostrado en la figura 8.3. Así vemos que la amplitud resultante dada por (8.2) no es otra cosa que la aplicación del teorema del coseno a la resultante que se obtiene de sumar los dos vectores del plano complejo. Y que este resultado no depende de la orientación de los vectores 1 y 2, solamente depende de su orientación relativa que viene dada por la diferencia  $\Delta_2 - \Delta_1$  entre los ángulos de cada vector con el eje horizontal, que no es otra cosa que la diferencia de fase  $\phi$  en (8.3).

### 8.3. Sistemas de franjas

La distribución espacial de máximos y mínimos de intensidad producida por la interferencia de distintas perturbaciones recibe el nombre de sistema de franjas. La visibilidad de los sistemas de franjas depende de la diferencia que haya entre las intensidades máxima y mínima. Se suele cuantificar la visibilidad  $V$  o contraste de un sistema de franjas mediante la cantidad

$$V = \frac{I_{max} - I_{min}}{I_{max} + I_{min}}, \quad (8.8)$$

que varía entre 1 (máximo contraste, cuando la intensidad mínima es nula, es decir cuando  $a_1 = a_2$ ) y cero (mínimo contraste, cuando la intensidad máxima y la mínima son iguales).

¿Cuál es el sistema de franjas asociado a dos fuentes puntuales coherentes? Para visualizar un sistema de franjas es conveniente determinar el lugar geométrico de los puntos con igual intensidad. Si las distancias de observación  $r_1$  y  $r_2$  son grandes comparadas con la longitud de onda, entonces la variación espacial más importante de  $a^2$  es a través del coseno (ver (8.2)) y no a través de las amplitudes  $a_1$  y  $a_2$ . Entonces, el lugar geométrico de los puntos con igual intensidad es el de los puntos con igual valor de  $\phi$ . Y según (8.3), esto corresponde a superficies con el mismo valor de  $r_2 - r_1$ . Si las fuentes emiten en fase, los puntos de intensidad máxima están en el plano  $r_2 = r_1$ , que es perpendicular al segmento que une  $S_1$  con  $S_2$  (el plano  $\Pi$  en la figura 8.2). En este plano las dos perturbaciones llegan completamente en fase luego de haber recorrido el mismo camino óptico. Otros puntos de intensidad máxima están en los hiperboloides (de dos hojas) de revolución  $r_2 - r_1 = \pm\lambda$ , con focos en  $S_1$  y  $S_2$ , donde las dos perturbaciones llegan desfasadas en  $2\pi$ . En general, el lugar geométrico de los puntos con intensidad constante es una familia de hiperboloides confocales, con focos en  $S_1$  y  $S_2$  (una de las hojas se indica en la figura 8.2). Si el sistema de franjas se observa en un plano perpendicular al segmento  $\overline{S_1S_2}$ , se verán círculos concéntricos (anillos) con intensidad máxima y anillos con intensidad mínima. Y si el sistema de franjas se observa en un plano que contiene al segmento  $\overline{S_1S_2}$  (el eje del hiperboloide), se verán hipérbolas.

### 8.4. Experiencia de Young

Uno de los primeros experimentos cuantitativos encaminados a poner de manifiesto la interferencia de la luz proveniente de dos fuentes fue realizado en 1800 por el científico inglés

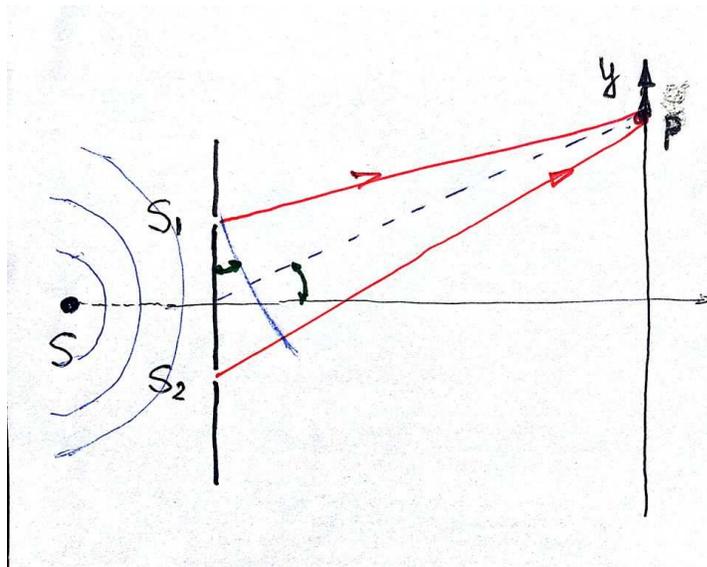


Figura 8.4: Experimento de Young.

Thomas Young. En la figura se esquematiza cómo hizo Young para obtener dos fuentes secundarias coherentes  $S_1$  y  $S_2$  a partir de una fuente de luz natural. Como las perturbaciones de diferentes partes de una fuente extensa ordinaria no están sincronizadas, se elige sólo una región de la fuente con un orificio estrecho  $S$ , de aproximadamente  $1\mu\text{m}$  de ancho, en una pantalla opaca. Este orificio se comporta como una fuente puntual idealizada (en las versiones actuales de este experimento se usa un laser como fuente y este orificio no hace falta). La “fuente”  $S$  ilumina una pantalla con orificios  $S_1$  y  $S_2$ , separados entre sí una distancia  $d$ . A partir de  $S$  se propagan frentes de onda que recorren distancias iguales y que llegan a  $S_1$  y  $S_2$  en fase, por consiguiente se puede considerar que  $S_1$  y  $S_2$  son fuentes puntuales que están en fase, es decir que son fuentes coherentes. A partir de acá el análisis corresponde al de cualquier par de fuentes puntuales coherentes. En todo punto iluminado simultáneamente por las perturbaciones que salen de  $S_1$  y  $S_2$  se observará la interferencia de la luz, según como se interactúe con el sistema de hiperboloides asociado con  $S_1$  y  $S_2$ . Si esta interacción se hace con una placa fotográfica o con una pantalla ubicada a una distancia  $D$  y paralela a la pantalla con orificios, se visualizará la intersección con el sistema de hiperboloides, es decir trozos de hipérbolas.

Para simplificar el análisis del experimento de Young, se suele suponer que la distancia  $D$  es grande comparada con la distancia  $d$  entre  $S_1$  y  $S_2$  (en un experimento podría ser  $d$  aproximadamente unos pocos milímetros y  $D$  un metro o más). En este caso, la diferencia de camino  $\overline{S_2P} - \overline{S_1P}$  satisface aproximadamente la relación

$$\frac{y}{D} \approx \frac{\overline{S_2P} - \overline{S_1P}}{d}, \quad (8.9)$$

Si suponemos que la “fuente”  $S$  equidista de los orificios  $S_1$  y  $S_2$ , entonces  $\delta_2 = \delta_1$  y la diferencia de fase de la ec. (8.3) tiene solo el aporte de la diferencia de camino  $\overline{S_2P} - \overline{S_1P}$ . Cuando esta cantidad es igual a un número entero de longitudes de onda, la intensidad en el punto  $P$  alcanza su valor máximo posible, mientras que si esta cantidad es igual a un número impar de

semilongitudes de onda, la intensidad en el punto P alcanza su mínimo valor posible. De estas consideraciones surge que

$$y_{max} \approx m \frac{\lambda D}{d}, \quad (8.10)$$

con  $m$  un número entero. Las perturbaciones que llegan a  $y = 0$  interfieren constructivamente para todo valor de  $\lambda$ , entonces el máximo central no está coloreado si se hace el experimento con luz blanca. En cambio, las posiciones  $y_{max}$  donde hay interferencia constructiva para  $m \neq 0$  dependen de  $\lambda$ , es decir que la posición del máximo para  $m = 1$  del color azul no es la misma que la posición del máximo para  $m = 1$  del color rojo y en consecuencia el sistema de franjas se ve coloreado cuando se usa luz blanca.

Si los orificios  $S_1$  y  $S_2$  tienen el mismo tamaño, podemos suponer que las amplitudes  $a_1$  y  $a_2$  son aproximadamente iguales en la zona central de la pantalla de observación (son exactamente iguales en  $y = 0$ ) y la expresión (8.2) para  $a^2$ , proporcional a la intensidad resultante, queda

$$a^2 = 2 a_1^2 (1 + \cos \phi) = 4 a_1^2 \cos^2 \frac{\phi}{2}, \quad (8.11)$$

y como sobre la pantalla

$$\phi = k (\overline{S_2 P} - \overline{S_1 P}) = \frac{k d}{D} y, \quad (8.12)$$

concluimos que la intensidad en la pantalla varía con  $y$  como un  $\cos^2$ . Si sobre la pantalla se coloca una película fotográfica, la transparencia del negativo varía con  $y$  como un  $\cos^2$ . Y si se coloca un material como las fotoresinas usadas en holografía, cuya solubilidad en un revelador es proporcional a la intensidad luminosa recibida, se obtiene una superficie rugosa con forma de seno (este dispositivo se llama red holográfica).

## 8.5. Coherencia

La falta de coherencia de las fuentes luminosas que nos rodean habitualmente es la gran dificultad para observar franjas de interferencia en la vida cotidiana. Ya vimos que esta dificultad fue resuelta por Young mediante el dispositivo esquematizado en la Fig. 8.4. El dispositivo de Young divide el frente de onda generado por la fuente térmica, con fase inicial que cambia muy rápidamente, en dos frentes de onda que pasan a ser coherentes, porque su diferencia de fase permanece constante en el tiempo. Esta solución resuelve el problema de la coherencia temporal, porque ahora las dos porciones en que se dividió el frente de onda original pueden interferir y dar máximos y mínimos de intensidad, como si provinieran de dos fuentes coherentes. Pero no tenemos que perder de vista que los trenes de onda siguen teniendo una longitud finita y que la fase inicial en las dos porciones en que se dividió cada tren siguen cambiando rápidamente. El truco usado por Young consigue que en ambas porciones la fase inicial cambie de la misma manera y así la cantidad  $\delta_1 - \delta_2$  (página 158) se mantiene constante. Esta afirmación es claramente cierta para el punto de la pantalla de observación que equidista de  $S_1$  y  $S_2$  y también para cualquier otro punto P en el que la diferencia de camino  $r_1 - r_2$  no sea comparable con la longitud de un tren de onda. Pero si el punto P se aleja mucho del centro de la pantalla, o si la fuente primaria  $S$  se coloca de manera asimétrica con respecto a  $S_1$  y  $S_2$ , podría ocurrir

que en el punto P se superpongan perturbaciones provenientes de dos trenes de onda diferentes. Y si pasa esto, el truco de Young deja de funcionar, porque la fase  $\phi$  de la ec. (8.3) pasa a cambiar muy rápidamente, porque lo hace la diferencia de fase  $\delta_2 - \delta_1$ . Por este motivo, para que el truco de Young funcione, las diferencias de camino acumuladas desde la fuente al punto de observación deben ser menores que la longitud de coherencia

$$\overline{SS_2} + \overline{S_2P} - \overline{SS_1} - \overline{S_1P} < v\tau. \quad (8.13)$$

De esta manera, la coherencia temporal de la fuente se traduce en una condición espacial, asociada a la duración limitada de los trenes de onda. En la práctica aparece otra dificultad, relacionada con otro tema: el tamaño finito de la fuente. Esto sucede porque las perturbaciones de diferentes partes de una fuente extensa ordinaria no están sincronizadas. En este caso se habla de coherencia espacial, que es una condición distinta a la expresada en la ec. (8.13). En el caso de fuentes extensas, los orificios  $S_1$  y  $S_2$  deben estar iluminados por frentes de onda sincronizados, no por frentes de onda provenientes de puntos desincronizados de la fuente extensa (por este motivo Young usó como fuente  $S$  un orificio muy pequeño, si hubiera usado el Sol como fuente, habría fracasado como fracasó Grimaldi que en 1665 había intentado algo similar).

## 8.6. Interferómetros

Los interferómetros son dispositivos para ver sistemas de franjas, o sea, para controlar problemas de coherencia. El interferómetro de Young es el padre de todos los interferómetros.

### 8.6.1. División del frente de onda

Las mismas consideraciones matemáticas y físicas que discutimos en el experimento de Young, se aplican directamente en otras configuraciones donde el frente de ondas proveniente de una fuente puntual primaria se divide en dos frentes de onda que interfieren entre si y que parecen provenir de dos fuentes puntuales secundarias. Por eso este tipo de interferómetros se llaman *de división del frente de onda*. Hay muchos ejemplos, como el espejo doble de Fresnel, biprisma de Fresnel, o espejo de Lloyd. muy bien descriptos en la última parte de la sección 9.3 del libro de Hecht, *Óptica* (páginas 397-399 de la traducción al español, misma numeración de páginas en la versión en inglés, cuarta edición).

Entre los interferómetros de división del frente de onda merece destacarse, por razones históricas, el espejo doble de Fresnel, que jugó un papel muy importante para invalidar las críticas que Young recibió de sus contemporáneos. Cuando Young publicó su trabajo en 1804, la comprensión más ampliamente aceptada de la naturaleza de la luz era la teoría corpuscular de Newton. En esta visión, la luz era una corriente de partículas que podrían agitar el éter e interactuar con objetos materiales a través de fuerzas atractivas y repulsivas. En ese momento, la mayoría de las personas interesadas en óptica seguía a Newton sin discusiones, excepto algunos visionarios como Young en Inglaterra, y Arago y Fresnel en Francia. Para ellos la luz era una onda elástica en el éter. Hoy podríamos pensar que el experimento de Young es una prueba irrefutable de

la naturaleza ondulatoria de la luz. Pero ese no fue el caso, no es fácil cambiar los paradigmas, especialmente cuando son sostenidos por autoridades como Newton. Los seguidores de la teoría corpuscular argumentaban que al pasar por las dos rendijas estrechas, las partículas de luz habían interactuado de manera mecánica con el material de los bordes de las rendijas . . . . O sea, no se querían convencer. El experimento del doble espejo de Fresnel en 1816 tenía la gran virtud de acabar con las rendijas por completo, y así proporciona más evidencia de que las franjas se producen, no por la acción de los bordes sino por la superposición de dos haces de luz . Con este experimento y otros trabajos teóricos y experimentales de Fresnel, la teoría ondulatoria de la luz gradualmente ganó preeminencia, y hacia 1830 comenzó a aceptarse que la teoría ondulatoria era más poderosa, porque explicaba más fenómenos observados, que la corpuscular.

### 8.6.2. División de amplitud

En los *interferómetros de división de amplitud*, el frente de ondas proveniente de una fuente puntual primaria se divide, mediante reflexiones y transmisiones parciales (no totales como en el espejo de Lloyd o en los espejos dobles de Fresnel) en láminas u otros objetos semitransparentes. Así se logran dos frentes de onda que por lo general tienen distinta amplitud, pero que son coherente entre sí debido a que provienen del mismo frente de ondas y por eso mantienen su diferencia de fase constante en el tiempo, siempre que no recorran caminos que difieran entre sí en cantidades mayores que la longitud de coherencia de la fuente utilizada. Un ejemplo de interferómetro por división de amplitud es el Interferómetro de Michelson.

#### Interferómetro de Michelson

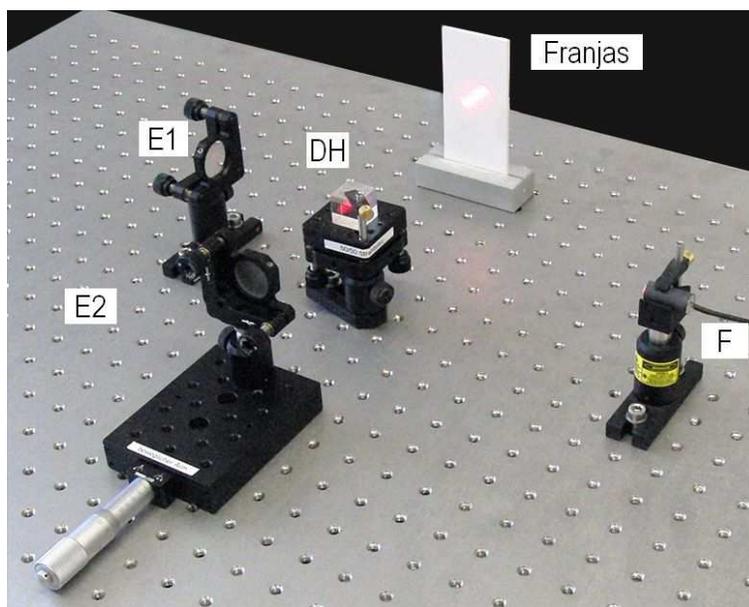


Figura 8.5: Interferómetro de Michelson (de Wikimedia Commons, archivo Aufbau-Michelson-Interferometer.jpg).

Además de sus múltiples aplicaciones actuales, este dispositivo experimental tiene importancia histórica ya que hace un siglo aportó uno de los puntales experimentales clave de la teoría de la relatividad. Hoy se continúa usando para mediciones muy precisas de la longitud de onda y en metrología óptica en general, cuando se necesita medir distancias muy pequeñas. Con el objetivo de detectar ondas gravitacionales, en el observatorio de LIGO se ha mejorado la sensibilidad hasta un nivel sin precedentes. LIGO, siglas de Laser Interferometer Gravitational-Wave Observatory, es decir, Observatorio de ondas gravitacionales por interferometría láser, utiliza largos interferómetros de tipo Michelson a fin de detectar ondas gravitacionales provenientes de agujeros negros, estrellas de neutrones, supernovas o remanentes del Big Bang.

Al igual que el interferómetro de Young, el interferómetro de Michelson se puede usar con una fuente de luz monocromática. La onda proveniente de esta fuente se divide en dos ondas que siguen caminos diferentes. Pero mientras que en el interferómetro de Young la división se hace enviando parte de la luz a través de una ranura y parte a través de otra; en el interferómetro de Michelson se emplea un dispositivo llamado divisor de haz, por eso pertenece a la familia de interferómetros por división de amplitud. En ambos casos hay interferencia cuando se combinan de nuevo las dos perturbaciones luminosas.

En la figura 8.5 se ve una foto del montaje en un banco óptico del interferómetro de Michelson. La luz proveniente de la fuente monocromática  $F$  incide en el divisor de haz  $DH$ , que es una placa de vidrio con un delgado recubrimiento de plata. Parte de la luz atraviesa  $DH$  y se refleja en el espejo  $E1$  y la parte restante se refleja en la superficie plateada y se dirige hacia el espejo  $E2$ . La parte que se refleja en el espejo  $E1$  regresa a través de  $DH$  y se refleja en la superficie plateada de  $DH$  hacia el espacio donde se coloca la pantalla. La parte que se refleja en el espejo  $E2$  regresa, atraviesa  $DH$  y se superpone con la otra parte en el espacio donde se coloca la pantalla. En clase vimos que las perturbaciones que llegan a esta zona parecen provenir de dos fuentes secundarias (virtuales) que están detrás de  $E2$ , en el eje que une  $E2$  con la pantalla de observación. Estas fuentes secundarias definen el eje del sistema de hiperboloides, así que si está todo bien alineado (es decir,  $E1$  y  $E2$  a  $90^\circ$  entre sí y a  $45^\circ$  de la cara plateada de  $DH$ ) en la pantalla se observan circunferencias concéntricas, tal como se muestra en el video <http://bit.ly/1k7XEho>.

El aparato de la figura está montado en su totalidad sobre una estructura antivibratoria (las mesas ópticas modernas lo son, Michelson llegó a usar mesas flotantes en mercurio porque los carruajes que pasaban por la calle hacían mover las franjas). La posición del espejo  $E2$  se puede ajustar con un tornillo micrométrico muy fino y preciso (se observa en la foto). Cuando se desplaza el espejo  $E2$  hacia atrás o hacia adelante una distancia  $\lambda/2$ , la diferencia de caminos entre las perturbaciones que pasaron por  $E1$  y las que pasaron por  $E2$  cambia en  $\lambda$  y cada franja se desplaza a la izquierda o a la derecha una distancia igual a la separación entre franjas, como puede verse en el video <http://bit.ly/1k7XEho>. Si se mueve el espejo una distancia  $d$  tal que la franja de orden  $M \pm m$  esté en el lugar en que antes estaba la franja de orden  $M$ , entonces

$$d = m \frac{\lambda}{2}, \quad (8.14)$$

Si  $m$  es de varios miles, la distancia  $d$  es suficientemente grande como para medirla con buena exactitud y así se obtiene un valor muy exacto para la longitud de onda. Por otra parte, si  $\lambda$  ya se conoce, el método sirve para medir distancias del orden de una longitud de onda de la luz empleada con relativa facilidad.

## 8.7. Interferencia en láminas

También vimos en clase cómo analizar la interferencia producida con láminas y una fuente puntual, de nuevo todo se reducía al análisis anterior, o sea, a descubrir dónde están las fuentes secundarias  $S_1$  y  $S_2$  y tener en cuenta los desfases adicionales que se producen durante la reflexión. Si la onda incide desde un medio menos denso hacia otro más denso, la onda reflejada se desfasa en  $\pi$ , en cambio si la onda incide desde un medio más denso hacia otro menos denso, la onda reflejada no se desfasa. Las ondas transmitidas no se desfasan en ningún caso. En el caso de observar por reflexión, la lámina cumple la función de dividir la amplitud de la onda incidente de manera que se puede imaginar que las perturbaciones en el semiespacio superior provienen de dos fuentes coherentes virtuales colocadas detrás de la lámina, que son las dos imágenes de S formadas por: i) reflexión en la primera cara y ii) por transmisión en la primera cara, reflexión en la segunda cara y de nuevo transmisión en la primera cara.

## 8.8. Localización de franjas

Por todo lo visto anteriormente, es claro que cuando en la región donde coinciden las perturbaciones producidas por dos fuentes puntuales coherentes (por lo general, las fuentes secundarias creadas mediante un interferómetro), el sistema de franjas de interferencia producido siempre está asociado con los hiperboloides de revolución o con parte de estos hiperboloides. Estos hiperboloides llenan todo el espacio donde hay superposición, por eso se dice que el sistema de franjas es **no localizado**, o también “localizado en todo el espacio”. “Con fuente puntual, franjas no localizadas” sería la regla. Una fuente extensa es una colección de fuentes puntuales. Pero como por lo general estas fuentes puntuales no están sincronizadas, entre ellas *no interfieren*, es decir que cada fuente puntual de una fuente extensa es incoherente con respecto a las otras. Y haciendo las reservas del caso sobre todo lo que está detrás de las expresiones *no interfieren* y *son incoherentes*.

Si la fuente extensa se usa en un interferómetro, cada punto dará lugar a un par de fuentes secundarias con su sistema de hiperboloides. Pero a menos que los máximos y mínimos de todos los hiperboloides coincidan, las franjas asociadas con determinado punto de la fuente extensa van a borrar el sistema de franjas asociado con los puntos vecinos. Por este motivo, cuando hay fuente extensa, las franjas, si se ven, están **localizadas**: sólo se ven en aquellos lugares donde coinciden los máximos y mínimos de todos los sistemas de hiperboloides, independientemente del punto de la fuente extensa en el que se originaron. Todo problema de localización de franjas de interferencia se resuelve analizando si existe un lugar donde esto ocurre. Sólo se observarán franjas claras cuando las diferencias de fase entre pares de haces recibidos en un punto dado procedentes de partes diferentes del emisor sean iguales o casi iguales. En general, esta condición sólo se cumplirá cuando el punto de observación está en una cierta superficie, es decir, sólo se verán si el ojo del observador (o su instrumento de observación) se enfoca sobre dicha superficie.

Los ejemplos de localización más accesibles y fáciles de discutir en un curso introductorio de ondas son las franjas producidas en láminas o películas transparentes. En clase empezamos

analizando las diferencias de camino y los desfases en reflexión necesarios para entender las **franjas de igual inclinación** que se observan con fuente puntual y una lente muy grande (texto asociado a la figura 9.17 del libro de Hecht). Una vez entendido este fenómeno, pasamos a ver lo que pasa cuando la lente (o el instrumento de observación) tiene pequeña abertura (texto asociado a figura 9.18 del libro de Hecht). Y así finalmente llegar al caso de fuente extensa (explicaciones de figuras 9.19, 9.20 y 9.21 del libro de Hecht), donde la localización se da en el infinito, porque es allí donde coincide la posición de los máximos y mínimos de todos los hiperboloides. También discutimos de manera muy general las llamadas franjas de igual espesor (explicaciones de figuras 9.22-9.26 del libro de Hecht).

# Capítulo 9

## Difracción

Si se acepta la validez del principio de Huygens, se debe reconocer que, cuando una onda interactúa con un obstáculo, la difracción siempre está presente. Sin embargo, y tal como vimos en el capítulo 6, la difracción aparece de manera significativa sólo cuando no se cumple la condición (6.1).

Recordemos que el principio de Huygens postula que todo punto de un frente de ondas primario actúa como fuente puntual que emite ondas esféricas secundarias de igual velocidad y frecuencia que la onda original. Y que el frente de ondas en todo instante es la envolvente de estas onditas esféricas (también llamadas onditas de Huygens) emitidas en un instante anterior. Por envolvente, Christiaan Huygens entendía *la superficie tangente*, un concepto fácil de visualizar geoméricamente, pero quizás poco adecuado tanto para interpretar físicamente como para obtener resultados cuantitativos. A continuación, para poner en marcha el principio de Huygens y para cuantificar los apartamientos de la descripción geométrica, seguiremos la interpretación de Fresnel, el primero en imaginar que dicha envolvente no era otra cosa que la interferencia de todas las onditas secundarias (u onditas de Huygens). O sea: el mecanimso de Huygens según Fresnel consiste en sumar las onditas de muchas fuentes puntuales. Entonces, para poner en marcha el mecanimso de Huygens-Fresnel, y aprovechando lo aprendido en el capítulo anterior sobre la suma de ondas provenientes de dos fuentes puntuales (ver página 158), ahora generalizaremos el análisis para considerar la suma de muchas perturbaciones. Empecemos por el caso de  $N$  fuentes puntuales coherentes, es decir, de la misma frecuencia y que mantienen su fases iniciales constantes en el tiempo.

### 9.1. Interferencia de $N$ fuentes

En la página 96 sintetizamos una función del tiempo a partir de  $N$  cosenos temporales con la misma amplitud ( $a$ ), la misma fase y frecuencias equiespaciadas en el intervalo  $\Delta\omega = \omega_2 - \omega_1$ . El espaciado era  $\delta\omega = \Delta\omega/(N - 1)$ . El resultado de esta síntesis se obtuvo con el procedimiento esquematizado en las ecuaciones (4.45), (4.47) y (4.48). Ahora queremos estudiar el efecto que producen las  $N$  perturbaciones armónicas que llegan desde las  $N$  fuentes a un punto P genérico. Si bien este problema parece completamente diferente del problema de sintetizar una señal con

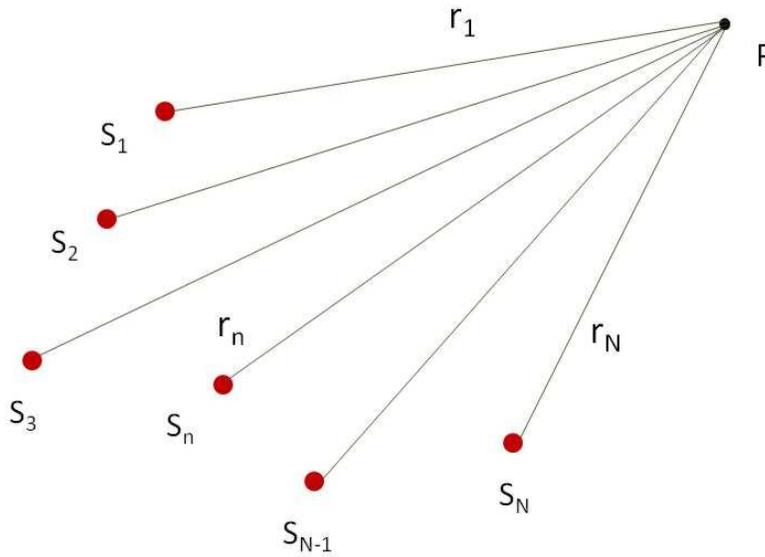


Figura 9.1: Las perturbaciones de  $N$  fuentes puntuales interfieren en el punto  $P$ .

$N$  funciones armónicas de distintas frecuencias, veremos que ambos problemas tienen muchísimo en común.

### 9.1.1. Caso general

En la figura 9.1 se esquematiza una situación general. Notar que más allá de la alineación o el equiespaciado de las fuentes, la linealidad de la ecuación de ondas nos conduce a una suma como en la página 158, pero con  $N$  términos, en vez de solamente dos. La situación es completamente análoga: cada fuente produce una onda esférica progresiva y al punto  $P$  llegan  $N$  perturbaciones de la forma

$$\psi_n(r_n, t) = a_n \cos(\omega t - kr_n + \delta_n) = a_n \cos(\omega t + \Delta_n), \quad (9.1)$$

donde  $\Delta_n = \delta_n - kr_n$ ,  $k = \omega/v$ ,  $r_n$  se mide desde cada fuente al punto  $P$ ,  $a_n$  depende de la intensidad de cada fuente y de  $r_n$ , mientras que  $\delta_n$  es la fase inicial que se introduce para tener en cuenta que las fuentes se pueden encender en tiempos distintos y, en consecuencia, que pueden estar desfasadas. La perturbación total en el punto  $P$  es

$$\psi_P(t) = \sum_{n=1}^N a_n \cos(\omega t + \Delta_n), \quad (9.2)$$

y esta suma tiene que tener la forma

$$\psi_P(t) = A \cos(\omega t + \text{fase}). \quad (9.3)$$

Tenemos que encontrar una expresión para la amplitud  $A$  de la onda resultante en función de amplitudes  $a_n$  y fases iniciales  $\Delta_n$  ( $n = 1, 2, \dots, N$ ) de las ondas individuales y la expresión

buscada se tiene que reducir a (8.2) en el caso  $N = 2$ . Como en el caso  $N = 2$ , la fase de la resultante no es interesante en experimentos donde solamente se mida la intensidad de la perturbación.

Tal como hicimos en 4.3.3 pasamos a notación compleja y sobreentendiendo “parte real” (que tomaremos al finalizar los cálculos), podemos escribir

$$\psi_P(t) = \sum_{n=1}^N a_n e^{i(\omega t + \Delta_n)} = e^{i\omega t} \sum_{n=1}^N a_n e^{i\Delta_n}. \quad (9.4)$$

Comparando la expresión (9.4) con la ec. (4.47) obtenida en la sección 4.3.3, notamos un gran parecido entre ambas, excepto por: i) la presencia de las amplitudes  $a_n$  dentro de la sumatoria en (9.4) (que para el caso considerado en (4.47) son iguales y salen de factor común) y ii) el argumento de las exponenciales complejas, que en (9.4) es  $\Delta_n = \delta_n - kr_n$ , un valor bastante arbitrario porque aún no se ha dicho nada de las fases iniciales  $\delta_n$  ni de cuál es la dependencia con  $n$  de las distancias  $r_n$ , (mientras que los  $\delta_n$  para la (4.47) son proporcionales a  $n$ ). Con esto queda bastante en evidencia lo que ya se había anunciado, que el problema de hallar la perturbación resultante de sumar  $N$  perturbaciones armónicas provenientes de  $N$  fuentes tiene un gran parecido formal con el problema de sintetizar una señal con  $N$  funciones armónicas de distintas frecuencias. La identidad formal de los dos problemas mencionados aparece justamente cuando se agregan hipótesis adicionales que permiten eliminar las diferencias mencionadas en i) y ii).

### 9.1.2. Fuentes alineadas y equiespaciadas

La dependencia de  $a_n$  con  $r_n$  se puede despreciar si las fuentes están localizadas en una región limitada del espacio y realizamos la medición en un punto P muy lejos de todas las fuentes. Para poner un caso concreto, consideremos que las fuentes están alineadas, como se muestra en la figura 9.2, y que están todas en fase ( $\delta_n$  independiente de  $n$ , valor que podemos tomar cero con una elección conveniente del origen de tiempos y entonces  $\Delta_n = -kr_n$ ). Con estas hipótesis, la suma de amplitudes complejas de la ec. (9.4) queda

$$\sum_{n=1}^N a_n e^{i\Delta_n} = \sum_{n=1}^N a_n e^{-ikr_n} = e^{-ikr_1} \sum_{n=1}^N a_n e^{-ik(r_n - r_1)} \quad (9.5)$$

donde el último paso equivale a tomar  $r_1$  como referencia, sacar factor común fuera de la sumatoria y medir todos los caminos en las exponenciales con respecto a  $r_1$ .

Para evaluar las diferencias  $r_n - r_1$ , consideremos que el plano del dibujo es el plano  $x = 0$ , que las fuentes están ubicadas a lo largo del eje y en las posiciones  $y_n$  y que el plano de observación X-Y dista una distancia  $Z$  del plano x-y de las fuentes. Entonces

$$r_n = \sqrt{Z^2 + (Y - y_n)^2} \approx Z + \frac{(Y - y_n)^2}{2Z}$$

$$r_1 = \sqrt{Z^2 + (Y - y_1)^2} \approx Z + \frac{(Y - y_1)^2}{2Z}$$

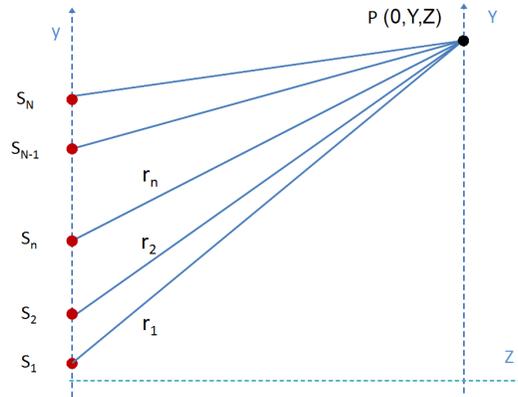


Figura 9.2: Las perturbaciones de  $N$  fuentes puntuales colineales (no necesariamente equiespaciadas) interfieren en un punto  $P$  alejado de la región donde están localizadas las fuentes.

y restando y quedándonos con el primer orden en  $Y/Z$ , nos queda

$$r_n - r_1 \approx (y_1 - y_n) \frac{Y}{Z} \quad (9.6)$$

que coincide con la diferencia de camino obtenida en la aproximación paraxial para el experimento de Young (la distancia entre fuentes por el seno, o la tangente, de un ángulo). Esta expresión también muestra que una variación de camino óptico de media longitud de onda importa mucho en la fase, porque está pesada por  $k$  y entonces las perturbaciones individuales “tiran” juntas (si los caminos son iguales) o “tiran” en direcciones opuestas (si los caminos difieren en media longitud de onda). En cambio, esa misma variación de camino no modifica demasiado las amplitudes, porque hace que una amplitud pase de un valor inversamente proporcional a  $1/r_1$  a otro valor inversamente proporcional a  $1/(r_1 + \lambda/2)$ , ambos muy parecidos entre sí cuando  $r_1$  es mucho mayor que la longitud de onda  $\lambda$ . Como valores típicos que ilustran esta aproximación para luz visible, podríamos tomar  $r_1 \approx 1m$  y  $\lambda \approx 0.5 \cdot 10^{-6} m$ .

La ec. (9.6) muestra que las fases de las exponenciales dentro de la sumatoria que representa la amplitud compleja resultante (ver ec. (9.5)) es siempre proporcional a la distancia entre la fuente 1 y la  $n$ . Entonces, si las fuentes están equiespaciadas, separadas todas la misma distancia  $d$ , las fases están equiespaciadas (como en el problema de suma de frecuencias), porque los caminos entre las perturbaciones provenientes de dos fuentes consecutivas cualesquiera, siempre son aproximadamente iguales a  $\Delta = d \sin \theta$  (Figura 9.3), con  $\theta$  la coordenada angular del punto lejano  $P$  medida desde la normal a la línea que pasa por todas las fuentes (eje  $Z$ ).

El caso límite en que el punto  $P$  está en el infinito, equivale a observar en el plano focal de una lente. Y en este caso, a la interferencia en  $P$  solamente contribuyen aquellas perturbaciones asociadas a la dirección de propagación (a los rayos diríamos en la DG) con inclinación  $\theta$ .

En resumen, en la configuración de la Figura 9.3, los caminos al punto de observación desde dos fuentes consecutivas difieren en el valor  $\Delta = d \sin \theta$  y esta diferencia es independiente de la etiqueta que numera a las fuentes. De esta manera,  $r_2 = r_1 - d \sin \theta$ ,  $r_3 = r_1 - 2d \sin \theta$ ,  $\dots$ ,

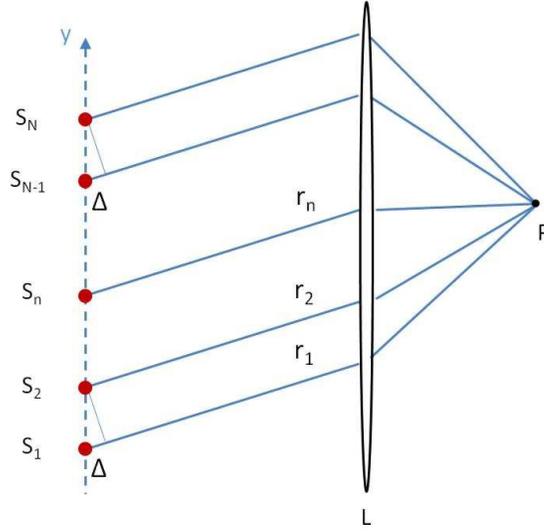


Figura 9.3: Las perturbaciones de N fuentes puntuales colineales equiespaciadas interfieren en el punto P en el plano focal de la lente L.

$r_n = r_1 - (n - 1) d \sin \theta$ . Luego, la expresión (9.4) para  $\psi_P(t)$  se escribe

$$\psi_P(t) = a e^{i\omega t} \sum_{n=1}^N e^{i\Delta_n} = e^{i\omega t} e^{-i k r_1} a \sum_{n=1}^N e^{i k (n-1) d \sin \theta}, \quad (9.7)$$

que en la representación compleja es una oscilación armónica con amplitud compleja igual al factor que multiplica a  $e^{i\omega t}$ . Este factor contiene una sumatoria entre 1 y N que mediante un cambio de índice se puede escribir como una sumatoria entre 0 y N-1 con la misma forma que la ec. (4.47) de la sección 4.3.3

$$\sum_{n=0}^{N-1} \left[ e^{i k d \sin \theta} \right]^n, \quad (9.8)$$

donde  $k d \sin \theta$  juega el mismo papel que  $\delta\omega t$  en (4.47). De esta manera, es claro que la sumatoria se puede escribir como

$$\frac{e^{i N k d \sin \theta} - 1}{e^{i k d \sin \theta} - 1} = e^{i[(N-1) k d \sin \theta]/2} \frac{\sin(\frac{N}{2} k d \sin \theta)}{\sin(\frac{1}{2} k d \sin \theta)}. \quad (9.9)$$

Los términos  $e^{-i k r_1}$  en (9.7) y  $e^{i[(N-1) k d \sin \theta]/2}$  en (9.9) tienen módulo unitario y por eso contribuyen solamente a la fase de la amplitud compleja de  $\psi_P(t)$ . Luego, la verdadera amplitud  $A_P$  (real) de la oscilación armónica resultante en P es

$$A_P = a \frac{\sin(\frac{N}{2} k d \sin \theta)}{\sin(\frac{1}{2} k d \sin \theta)} = a \frac{\sin N \alpha}{\sin \alpha}, \quad (9.10)$$

donde por brevedad definimos

$$\alpha = \frac{k}{2} d \sin \theta. \quad (9.11)$$

La intensidad  $I_P$  de la perturbación armónica resultante en P, proporcional al cuadrado de la amplitud  $A_P$ , queda

$$I_P = I \frac{\sin^2 N \alpha}{\sin^2 \alpha}, \quad (9.12)$$

donde  $I$ , proporcional al cuadrado de la amplitud  $a$ , representa la intensidad en P cuando está encendida una sola de las fuentes individuales.

Del resultado (8.5) obtenido en el caso  $N = 2$  sabemos que cuando las fuentes interfieren constructivamente, la amplitud resultante es la suma de las amplitudes individuales ( $2a$ ) y así la intensidad resultante es  $4 = N^2$  veces la intensidad que se obtendría con una única fuente. Y antes de seguir con las cuentas, podemos convencernos que, bajo las hipótesis que hemos explicitado, tiene que pasar también algo similar en el caso  $N > 2$ . Porque si todas las perturbaciones están en fase (si todas “tiran” para el mismo lado), la amplitud de la perturbación armónica resultante en P debe ser  $N$  veces la amplitud de una perturbación individual ( $Na$ ) y así la intensidad resultante debe ser  $N^2$  veces la intensidad de una única fuente ( $N^2 I$ ). Veamos a continuación que la (9.12) predice justamente este resultado.

### 9.1.3. Intensidad en la zona lejana

La ec. (9.12) describe la variación de la intensidad  $I_P$  en términos de la variable  $\alpha$ , que contiene a las variables físicas  $\lambda = 2\pi/k(\omega)$ ,  $d$ ,  $N$  y  $\theta$ .

Con numerador y denominador acotados, los máximos de  $I_P$  ocurren cuando se hace cero el denominador, es decir cuando

$$\alpha = m \pi, \quad \text{o sea} \quad \Delta = d \sin \theta = m \lambda, \quad (9.13)$$

con  $m$  un número entero. Esta condición confirma que la interferencia constructiva (máxima intensidad posible) aparece cuando todos los caminos desde las fuentes al punto de observación difieren en múltiplos de una longitud de onda. Cuando se cumple la condición (9.13) también se anula el numerador de (9.12) y al resolver la indeterminación se obtiene que la máxima intensidad es  $N^2 I$ , un resultado que también habíamos anticipado.

En cambio, la interferencia destructiva (mínima intensidad posible) ocurre cuando se anula el numerador de (9.12), pero no su denominador, es decir cuando

$$\sin N \alpha = 0 \quad \text{y} \quad \sin \alpha \neq 0 \quad (9.14)$$

o sea

$$N \alpha = j \pi, \quad \text{pero} \quad \alpha \neq l \pi \quad (9.15)$$

con  $m$  y  $l$  enteros, que en términos de  $\theta$  se escribe

$$N \Delta = N d \sin \theta = j \lambda, \quad (9.16)$$

con  $j$  un entero distinto de cero y que no sea múltiplo de  $N$ , para que no se anule el denominador de la ec. (9.12). Como indica claramente la ec. (8.6) para en el caso  $N = 2$ , notemos al pasar que el hecho de que en este caso sea cero la mínima intensidad posible es una consecuencia de

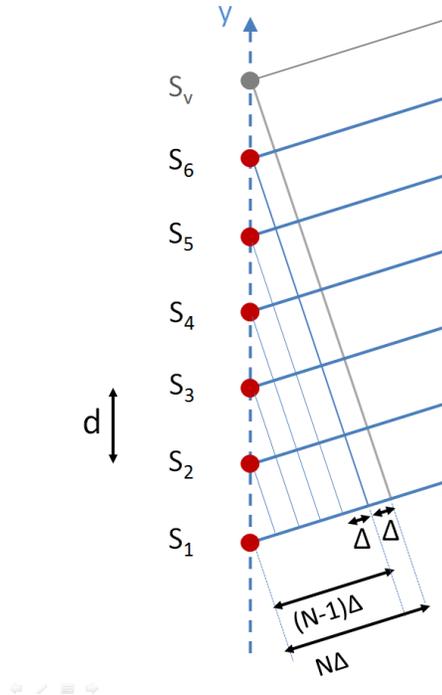


Figura 9.4: Cuando se cumple la condición (9.16) para  $j = 1$ , la primera fuente ( $n = 1$ ) estaría en fase con una supuesta fuente virtual ( $n = N + 1 = 7$ ) alineada y con el mismo espaciamiento, es decir que estaría en contrafase con la fuente en el centro ( $n = 4$ ) del arreglo de  $N + 1$  fuentes. Y lo mismo pasaría para las fuentes con  $n = 2$  y  $n = 5$  y así sucesivamente.

la suposición que se hizo al comienzo de la subsección 9.1.2 sobre la igualdad de las amplitudes de todas las perturbaciones en P.

Es instructivo interpretar físicamente la condición (9.16). Con este fin consideremos el esquema de la Figura 9.4 y observemos que según la ec. (9.7),  $(N - 1)\Delta$  es la diferencia de camino entre las perturbaciones que llegan a P desde la primera y la última fuente. Es decir que el lado izquierdo de la condición (9.16) sería la diferencia de camino entre las perturbaciones que llegan a P desde la primera fuente y una eventual fuente virtual alineada con las fuentes reales y separada la misma distancia  $d$  de la última fuente. Y esto es equivalente a decir que la primera fuente estaría en contrafase con la fuente en el centro ( $n = 4$ ) del arreglo de  $N + 1$  fuentes y que toda fuente en la primera mitad del arreglo estaría en contrafase con otra fuente en la segunda mitad del arreglo. Un argumento similar vale para otros valores de  $j$ . En general, siempre que se cumpla la condición (9.16), es posible formar pares en contrafase y estos pares, que claramente interfieren destructivamente, producen ceros de intensidad.

También es instructivo interpretar las condiciones de máximos y mínimos de intensidad en el marco de la suma de números complejos  $a_n e^{i\Delta_n}$  que aparece del lado derecho de la eq. (9.4). En el plano complejo estos números son vectores de módulo  $a_n$  e inclinación  $\Delta_n$  con respecto al eje real. Si todas las inclinaciones  $\Delta_n$  son iguales o difieren en múltiplos de  $2\pi$ , la resultante de la sumatoria es un vector de amplitud igual a la suma de las amplitudes, aunque sean distintas (interferencia constructiva, condición (9.13)). En cambio, si se tienen pares de vectores de igual módulo y orientaciones opuestas (con  $\Delta_n$ 's que difieren en  $\pi$ ), la resultante de la sumatoria es

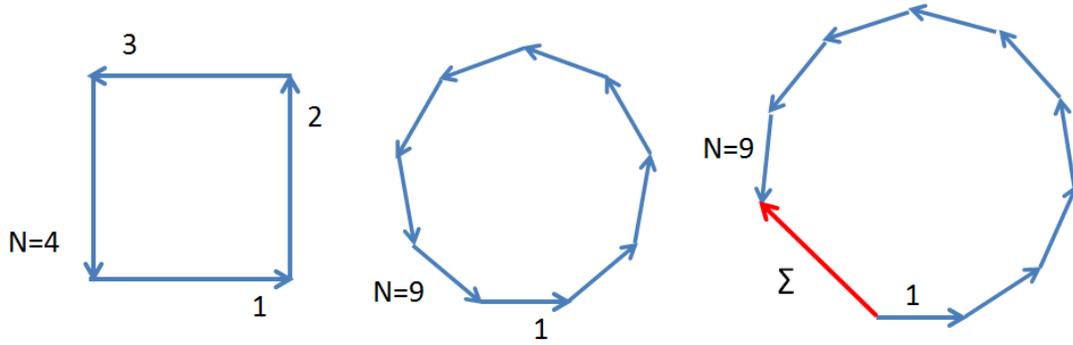


Figura 9.5: La suma de amplitudes complejas  $a_n e^{i\Delta_n}$  de la eq. (9.4) como suma de vectores en el plano complejo cuando todos los vectores tienen el mismo módulo y los valores  $\Delta_n$  difieren en el mismo valor  $2\alpha$ . Cuando el valor  $2\alpha$  está dado por la condición de mínimo (9.15), la resultante es cero, como se ilustra para el caso  $j = 1$ , con  $N = 4$  (izquierda) y  $N = 9$  (centro). A la derecha se muestra la resultante (vector rojo) para  $N = 9$  y un valor  $2\alpha$  que no satisface ni la condición de mínimo (9.15) ni la de máximo (9.13) ( $2\alpha = 33^\circ$ ).

un vector de amplitud nula (caso de interferencia destructiva, del que la condición (9.13) es un caso particular cuando todas las amplitudes son iguales y  $N$  es par). Si todas las amplitudes son iguales y la inclinación relativa entre dos vectores consecutivos se mantiene constante, situación que es equivalente a la considerada en la ec. (9.7) ( $\Delta$  independiente de  $n$ ), es fácil ver que siempre que  $\Delta$  satisfaga la condición (9.16), los vectores puestos uno a continuación del otro determinan en el plano complejo un polígono regular donde la punta del último vector coincide con el comienzo del primero, *cualquiera sea el valor de  $N$  (par o impar)*, o en otras palabras, que la resultante de la suma de vectores es nula. Esto es así porque según la ec. (9.11),  $2\alpha$  es la inclinación relativa entre vectores complejos asociados con fuentes consecutivas. Y según (9.15), en la condición de mínimo esta inclinación vale  $j 2\pi/N$  ( $j$  distinto de cero y no múltiplo de  $N$ ), es decir, justo la inclinación necesaria como para que la flecha del último vector toque el comienzo del primero, tal como se muestra en la figura 9.5 para  $j = 1$  y en los casos  $N = 4$  y  $N = 8$ .

En resumen, la ec. (9.12) muestra que la intensidad  $I_P$  en el plano focal de la lente tiene máximos cuando  $\alpha = \dots -2\pi, -\pi, 0, \pi, 2\pi \dots$  y la ec. (9.15) muestra que entre dos máximos consecutivos  $I_P$  se hace cero  $N - 1$  veces, un resultado que está de acuerdo con lo que habíamos obtenido para el interferómetro de Young ( $N = 2$ ), donde entre dos máximos de interferencia constructiva encontramos solamente un lugar donde la intensidad se anula. Por ejemplo entre el máximo en  $\alpha = 0$  y el máximo en  $\alpha = \pi$ , la intensidad vale cero cuando  $j$  en (9.15) vale  $1, 2 \dots N - 1$ , es decir, cuando  $\alpha = \pi/N, 2\pi/N \dots (N - 1)\pi/N$ , o sea cuando  $d \sin \theta = \lambda/N, 2\lambda/N \dots (N - 1)\lambda/N$ . Lo mismo ocurre entre el máximo en  $\alpha = -\pi$  y el máximo en  $\alpha = 0$ , donde la intensidad vale cero  $N - 1$  veces cuando  $j$  en (9.15) vale  $-(N - 1), -(N - 2), \dots, -2, -1$ , es decir, cuando  $\alpha = -(N - 1)\pi/N, -(N - 2)\pi/N, \dots, -2\pi/N, -\pi/N$  y mismo ocurre entre cualquier otro par de máximos con valores consecutivos de  $m$  en la condición (9.13).

La dependencia de la intensidad  $I_P$  con la variable  $\alpha$  está ilustrada en la figura 9.6 para el caso  $N = 5$ ,  $I = 1$  y cuatro valores del cociente  $d/\lambda = 1.2, 1.5, 2$  y  $4$ . Este cociente es el parámetro adimensional característico del fenómeno de interferencia que estamos estudiando. El resultado

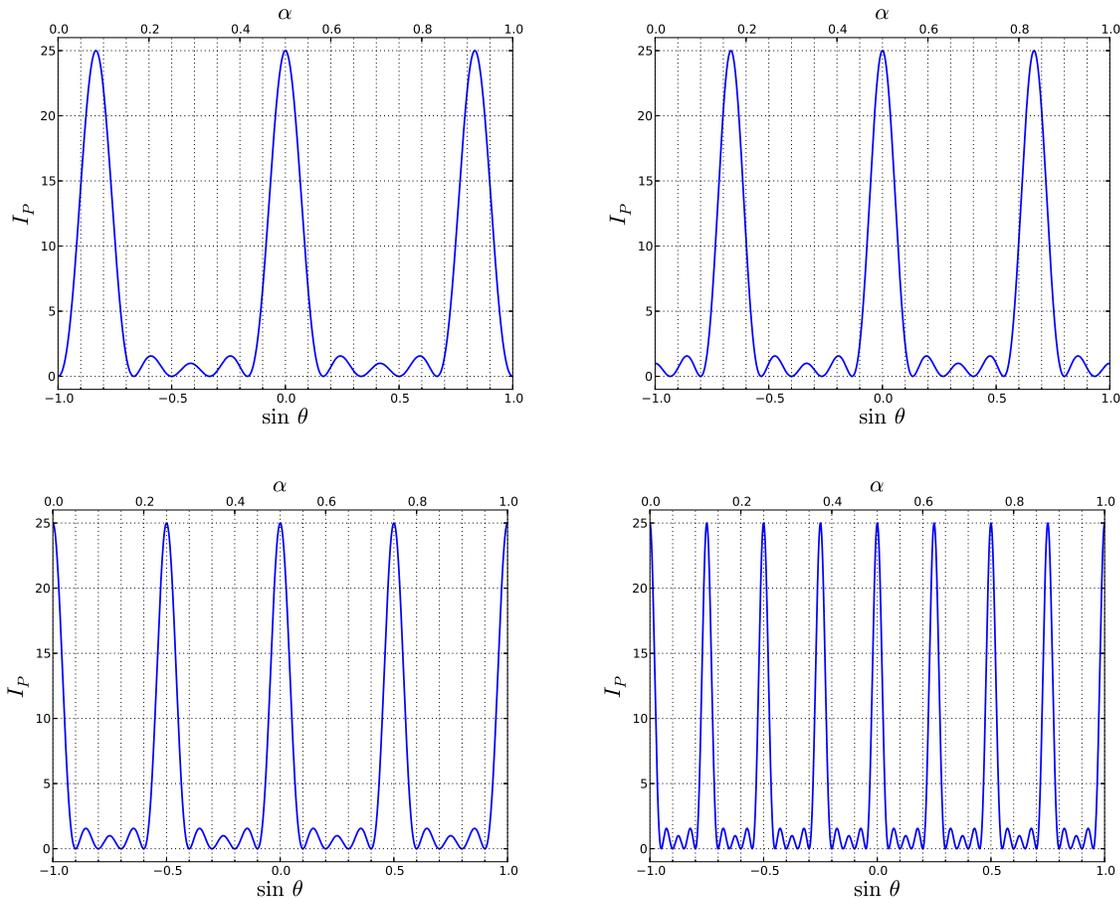


Figura 9.6: Intensidad  $I_P$  dada por la ec. (9.12) como función de  $\alpha$  (eje superior) y de  $\sin \theta$  (eje inferior) para  $N = 5$  fuentes idénticas equiespaciadas y para distintos valores del cociente  $d/\lambda$ : 1.2 (arriba izquierda), 1.5 (arriba derecha), 2 (abajo izquierda) y 4 (abajo derecha).

de la interferencia en este problema no depende de los valores de  $\lambda$  o  $d$  por separado ni tampoco de la naturaleza del fenómeno. Que los ejemplos de interferencia se hayan discutido mayormente en el contexto de las ondas luminosas no debería hacernos olvidar que, con las mismas hipótesis que hemos venido usando, se obtienen los mismos efectos de interferencia con otras ondas, como sonido audible, ultrasonido, ondas en la superficie libre de un fluido (olas) o con cualquier otro tipo de onda.

Como la intensidad producida en P por una única fuente se tomó arbitrariamente igual a uno, todos los máximos principales de la figura 9.6 alcanzan el mismo valor ( $N^2 = 25$ ). Estos máximos son efecto de la interferencia constructiva, tienen un origen físico completamente análogo a los llamados órdenes del experimento de Young y cada uno tiene asociado un valor distinto del entero  $m$  de la ec. (9.13). Igual que en el experimento de Young, se los denomina *órdenes de interferencia*.

Notar que los órdenes de interferencia están equiespaciados en  $\alpha$  (y en  $\sin \theta$ ), pero no están equiespaciados en el plano focal de la lente. Es decir, no están equiespaciados en  $\theta$  ni en las distancias medidas en el plano focal. Una vez fijado el valor de  $d$ , el equiespaciado en las variables

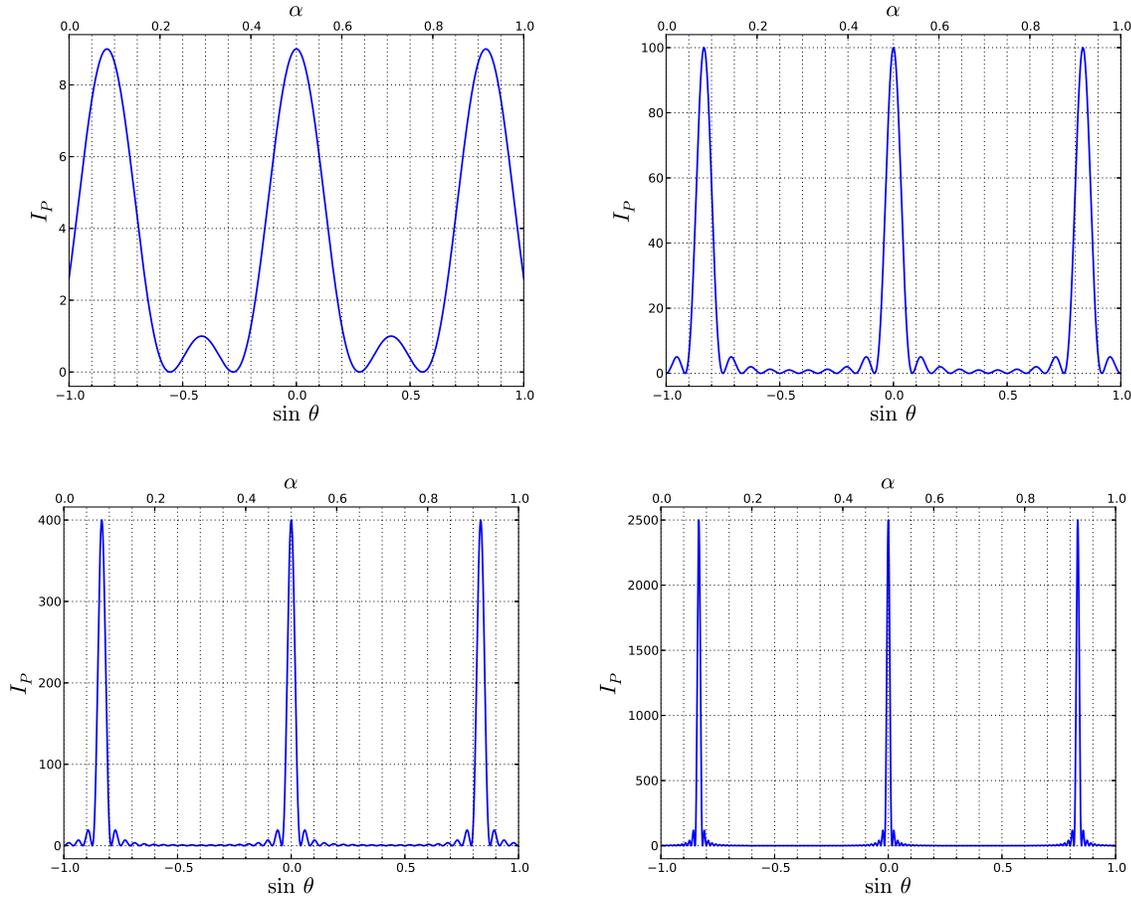


Figura 9.7: Intensidad  $I_P$  dada por la ec. (9.12) como función de  $\alpha$  (eje superior) y de  $\sin \theta$  (eje inferior) para  $d/\lambda = 1.2$  y distintos valores de  $N$ : 3 (arriba izquierda), 10 (arriba derecha), 20 (abajo izquierda) y 50 (abajo derecha).

$\alpha$  y  $\sin \theta$  depende del valor de la longitud de onda: a menor longitud de onda (mayor valor de  $d/\lambda$ ), menor es el equiespaciado entre órdenes. Y debido al hecho físico de que solamente tenemos acceso a ángulos  $\theta$  reales entre  $-\pi/2$  y  $\pi/2$ , la variación de  $\sin \theta$  está acotada como así también los valores permitidos del entero  $m$

$$-\frac{d}{\lambda} \leq m \leq \frac{d}{\lambda} \quad (9.17)$$

Vemos que el número total de órdenes resulta dos veces la parte entera del parámetro adimensional característico  $d/\lambda$ , y es menor cuanto mayor es la longitud de onda. Físicamente, la explicación es que si la longitud de onda es muy grande comparada con  $d$ , por más que variemos la coordenada angular que mide la inclinación  $\theta$  del punto lejano  $P$  (ver figura 9.3), los caminos ópticos provenientes de cada fuente en el punto  $P$  siempre serán muy parecidos, sus variaciones serán siempre muy pequeñas comparadas con  $\lambda$  y así las variaciones de intensidad resultarán casi inapreciables. En cambio, si la longitud de onda es muy pequeña comparada con  $d$ , una pequeña variación de la coordenada angular  $\theta$  puede producir variaciones de los caminos ópticos del orden de varias longitudes de onda, una condición necesaria para que al mover el detector en el plano focal de la lente, la intensidad pueda variar un número apreciable de veces entre sus valores máximo y mínimo y así tener un número apreciable de franjas. Tanto la discusión

```

DifraccionNequi1.py (~/.Dropbox/Ubuntu/F2) - gedit
Open Save Undo
DifraccionNequi1.py x
#####
# Python para interferencia de N fuentes
#           id'enticas equiespaciadas
# Curso de F2, 1er cuatrimestre de 2017
# ver http://users.df.uba.ar/rdep/
#####

import numpy as np
import matplotlib.pyplot as plt

#           # Datos entrada
N=50        # Nro. de fuentes
dl=1.2     # d/lambda
npuntos=20000 # puntos en curva

pi=np.pi
sint=np.linspace(-1.,1.,npuntos) # grilla del seno
alfa=pi*dl*sint
inteP=(np.sin(N*alfa)/np.sin(alfa))**2

fig = plt.figure()
ax1 = fig.add_subplot(111) # para tener 2 ejes
ax1.plot(sint, inteP, c='b', linewidth=1.7, ls='-')
ax1.set_xlim(-1, 1)
ax1.set_xlabel(r"$\sin \theta$", fontsize=22) # etiqueta eje x1
ax1.set_ylabel(r"$I_P$", fontsize=22) # etiqueta eje y
ax1.grid(True, which="both")
ax1.minorticks_on() # subgrilla
ax1.set_ylim(-100, 2600.) # cambiar con N

ax2 = ax1.twinx()
ax2.set_xlabel(r"$\alpha$", fontsize=22)

plt.draw()
plt.savefig('C9IP2.eps') # figura en eps
plt.savefig('C9IP2.jpg') # figura en jpg

Python Tab Width: 8 Ln 5, Col 7 INS

```

Figura 9.8: Comandos Python para obtener figuras 9.6 y 9.7.

anterior como la expresión (9.17) muestran que  $N$  fuentes puntuales en fase y “muy juntas” ( $d/\lambda < 1$ ) se comportan en la zona lejana casi como una única fuente puntual, en el sentido que no producen órdenes de interferencia ya que no existen variaciones apreciables de camino óptico en puntos de observación lejanos. Estas consideraciones pueden usarse para establecer un criterio que involucre  $d$ ,  $\lambda$  y la posición del punto de observación y que sirva para establecer cuándo una fuente extensa puede ser considerada puntual.

En la figura 9.7 exploramos cómo cambian las curvas de  $I_P$  predichas por la ec. (9.12) cuando se aumenta el número de fuentes. El valor  $d/\lambda = 1.2$  se mantiene constante y se muestran los casos  $N = 3, 10, 20$  y  $50$ . Como la intensidad  $I = 1$  de las fuentes individuales se deja constante para todas las curvas, la intensidad máxima en  $P$  aumenta con el cuadrado del número de fuentes. Se observa que a medida que aumenta  $N$ :

- i) se hace mas pequeña la distancia angular entre cada orden de interferencia y su dos mínimos adyacentes a izquierda y derecha, tal como lo predice la ec. (9.15). Esto equivale a un afinamiento muy notable de los picos de intensidad correspondientes a los órdenes de interferencia, que pasan a ocupar regiones cada vez limitadas del espacio y a estar cada vez más definidas. En particular, notar que para  $N = 2$  podríamos decir que las colitas de los órdenes  $\pm 1$  se van más allá del horizonte, que puede sonar casi esotérico pero no.
- ii) disminuye la intensidad relativa de los  $N - 2$  máximos secundarios que hay entre dos órdenes consecutivos. Si bien nunca calculamos explícitamente la posición angular ni la intensidad de estos máximos secundarios, al hacer la cuenta se obtiene que no todos tienen la misma intensidad y que están aproximadamente a mitad de camino entre dos ceros consecutivos, como se puede comprobar muy fácilmente con un sencillo código Python como el de la Figura 9.8.
- iii) la función  $I_P(\alpha)$  tiende a un peine de deltas, los dientes son los órdenes, cada vez mas delgados y el fondo de máximos secundarios se hace cada vez menos significativo. Excepto por el distinto significado de las variables físicas involucradas, este peine es completamente análogo al esquematizado en la segunda fila de la Figura 4.1.

#### 9.1.4. Intensidad en la zona cercana

La figura 9.9 se corresponde con el esquema de la figura 9.3, pero con la diferencia de que el punto de observación, en vez de estar muy alejado de las  $N$  fuentes colineales, es un punto “**cercano**”. Entonces, ni las amplitudes ni las fases relativas de las perturbaciones individuales satisfacen las hipótesis que usamos para obtener la expresión (9.10). Por este motivo, no nos queda otro remedio que volver al punto de partida, que es el principio de superposición expresado en la suma de la ec. (9.2).

Este problema es muy importante, no solamente porque da lugar a muchísimas aplicaciones tecnológicas con fenómenos ondulatorios sino también porque la envolvente de las onditas secundarias postuladas por Huygens no es otra cosa que la *interferencia* de muchas onditas secundarias, tal como propuso Fresnel. El mensaje es el siguiente: si estamos familiarizados con expresiones del tipo (9.2), seremos capaces de entender mejor el principio de Huygens y ponerlo

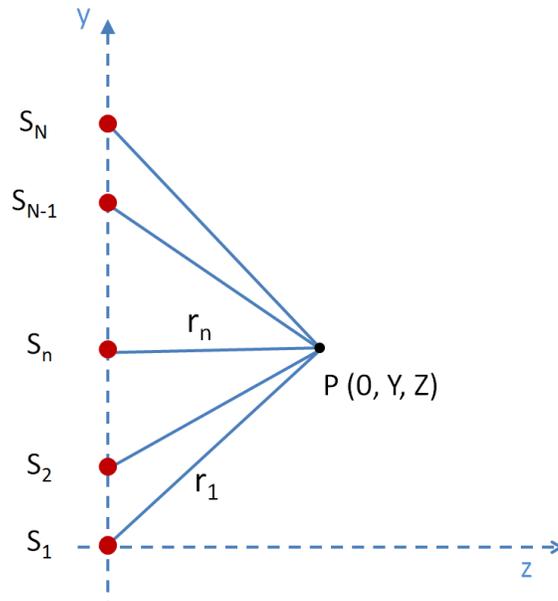


Figura 9.9: Las perturbaciones de  $N$  fuentes puntuales colineales equiespaciadas interfieren en un punto  $P$  “cercano” en el plano  $x = 0$  que contiene a las fuentes.

en marcha para obtener una herramienta matemática que permita predecir resultados cuantitativos mucho más específicos que los que se obtienen solamente con el concepto geométrico de “envolvente”.

Es conveniente mencionar que el interés en la interferencia con muchas fuentes (como así también el interés en la difracción y en los apartamientos de la descripción geométrica) no está limitado a problemas donde el punto de observación está lejos de las fuentes. Si bien el estudio sistemático para puntos cercanos está fuera del alcance de un curso introductorio de ondas, es muy instructivo evaluar la suma involucrada en la expresión (9.4) empleando el método FB (de fuerza bruta), que bien usado y con la ayuda de una computadora o cualquier dispositivo con microprocesador programable (computadora de escritorio, notebook, tablet, celular) es hoy en día una herramienta tan válida e imprescindible como lo fue la regla de cálculo o las tablas de funciones especiales en otras épocas.

Consideremos que las  $N$  fuentes puntuales están ubicadas a lo largo del eje  $y$  con coordenadas cartesianas  $x_n = 0$ ,  $y_n = (n - 1)d$ ,  $z_n = 0$  y que situamos el punto  $P$  “cercano” en el plano  $x = 0$  que contiene a las fuentes, con coordenadas cartesianas  $(0, Y, Z)$ . Las fases  $\Delta_n$  que intervienen en (9.4) siguen siendo  $\Delta_n = \delta_n - kr_n$  y si mantenemos la hipótesis de que las fuentes están todas en fase,  $\delta_n = 0$ . Pero ahora dos valores de  $r_n$  consecutivos no difieren en un valor constante, como pasaba en el caso de puntos lejanos (recordar figura 9.3). Esto es así porque la nueva expresión para  $r_n$  es

$$r_n = \sqrt{Z^2 + (Y - y_n)^2}, \quad (9.18)$$

en vez de la dada unas líneas antes de la ec. 9.7. Luego, la sumatoria en la ec. (9.4) queda

$$\sum_{n=1}^N a_n e^{i\Delta_n} = \sum_{n=1}^N a_n e^{ikr_n} = e^{ikr_1} \sum_{n=1}^N a_n e^{-ik(r_1-r_n)} \quad (9.19)$$

donde el argumento de la exponencial en la última sumatoria no depende linealmente de la etiqueta que identifica a las fuentes, como ocurría en el caso considerado en la sección 9.1.3. Si bien el cambio impide encontrar una expresión cerrada para la amplitud de la perturbación resultante en P, todo físico que se precie debe ser capaz de analizar el problema con herramientas computacionales.

### 9.1.5. Experimentos numéricos

La solución numérica de problemas físicos sin solución analítica explícita fue históricamente la primera aplicación científica de las primeras computadoras. Hoy en día constituye un área multidisciplinaria con identidad propia donde no solo se superponen física, matemáticas y computación (tres carreras de nuestra Facultad de Ciencias Exactas y Naturales de la Universidad de Buenos Aires) sino que también actúa de puente entre dichas disciplinas.

Por otro lado, desarrollar una herramienta numérica para resolver un problema físico y asegurarse de que la herramienta está funcionando correctamente es un excelente ejercicio para poner a prueba nuestra comprensión de muchos detalles que sin este ejercicio hubiera pasado desapercibidos.

A continuación concentrémonos en la simulación numérica del problema planteado en la eq. (9.19). Un brevísimo código Python podría ser el siguiente

```
#####
# Campo Cercano
# N fuentes id'enticas equiespaciadas
#
# Curso de F2, 1er cuatrimestre de 2017
# ver http://users.df.uba.ar/rdep/>
#####

import numpy as np
import matplotlib.pyplot as plt
import cmath; import math

#
# Datos entrada
#
N=3 # Nro. de fuentes
dl=1.6 # d/lambda
# # delimita zona
xini=-.1+1.*10.**(-5)
xfin=2.5 # en unidades de d
nrox=150
yini=-3.6+1.*10.**(-5) # en unidades de d
yfin=6.8 # en unidades de d
nroy=500

af=1. # amplitud fuentes
pi=np.pi
ceror=0.; ceroi=0.; cero=ceror+1j*ceroi

f=open('NfuentesCercano.dat','w')
print >>f, "x x psi2 logpsi2"
print >>f, "Nro. fuentes = %g" %N
print >>f, "d/lambda = %g" %dl
print >>f, "xini = %g, xfin= %g, nrox= %g " %(xini, xfin, nrox)
print >>f, "yini = %g, yfin= %g, nroy= %g " %(yini, yfin, nroy)

print >>f, "
print >>f, "x y psi2 logpsi2"
```

```

x1 = dl*np.linspace(xini, xfin, nrox)
y1 = dl*np.linspace(yini, yfin, nroy)

yf = dl*np.arange(N)          # yj de fuentes en x=0
                               # medidas en lambdas

for ii in range(nrox):
    for jj in range(nroy):
        aux1=cero
        r1=(x1[ii]**2+y1[jj]**2)**0.5
        for jf in range(N):
            rj=(x1[ii]**2+(y1[jj]-yf[jf])**2)**0.5
            deltarj=rj-r1
            aux1=aux1+af*cmath.exp(1j*2*pi*deltarj)/rj
        psi2=abs(aux1**2)
        print >>f, "%g %g %g" %(x1[ii],y1[jj],psi2)

```

donde los datos de entrada incluyen el número de fuentes  $N$ , el cociente  $d/\lambda$  y las coordenadas del punto de observación. Para explorar cómo varía espacialmente la intensidad de la perturbación resultante, el punto P se mueve en una grilla rectangular determinada por las variables de entrada  $xini$ ,  $xfin$  y  $nrox$  (grilla  $x1$  de puntos en la dirección  $Z$ ) e  $yini$ ,  $yfin$  y  $nroy$  (grilla  $y1$  de puntos en la dirección  $Y$ ). Se eligió dar las longitudes medidas en unidades de  $d$ , construir las grillas en estas unidades y luego multiplicar por  $dl$ . Luego, se usan dos ciclos para recorrer cada punto de la grilla y un tercer ciclo anidado para hacer la sumatoria indicada en (9.19). Al terminar la suma se calcula el valor absoluto al cuadrado, se escriben las coordenadas y la intensidad en un archivo y luego se pueden graficar los datos de este archivo con el método preferido. El gráfico se podría haber hecho en el mismo script con comandos Python, pero quizás es preferible tener los datos en un archivo y luego jugar con comandos Python o usar un graficador externo, como Gnuplot.

Consideremos un sistema de 3 fuentes equiespaciadas y usemos el archivo de comandos mostrado arriba para explorar la distribución de intensidad en las distintas zonas. Los resultados se visualizan mejor en un mapa de colores con escala logarítmica para las intensidades, tal como se muestra en la figura 9.10 para la región indicada en el script de arriba, comprendida entre  $0.12 \leq Z/\lambda \leq 4$  (muy cercana). Lo primero que nos llama la atención es que en esta zona la perturbación resultante tiene una distribución espacial muy irregular, nada parecida a la que anuncia la ec. (9.12). Mientras la ec. (9.12) para campo lejano predice diagramas de interferencia que solamente dependen de las coordenadas angulares, no radiales, del punto de observación, el diagrama de intensidad obtenido en la figura 9.10 varía violentamente en la dirección radial, con máximos de gran intensidad (puntos calientes, rojos) muy cerca de las fuentes (en  $y_n/\lambda = 0, 1.6$  y  $3.2$ ), junto a otras zonas menos intensas (amarillas, casi 1 decibel menos) y otras zonas donde la intensidad sigue bajando un decibel (verde, verde azulado) hasta llegar al azul oscuro que es una zona con intensidad casi cien mil veces menor que las zonas de intensidad máxima (zonas frías en un horno de microondas, o silenciosas en el lenguaje acústico).

Para sistemas de fuentes con tamaño típico menor que media longitud de onda, la zona cercana suele ser muy chica, menor que una longitud de onda. En cambio, para sistemas de fuentes de tamaño típico mayor que media longitud de onda, como en este caso, donde el tamaño típico es  $D = (N - 1)d = 3.2\lambda$ , el límite entre zona cercana y lejana viene dado aproximadamente por la *distancia de Fraunhofer*  $D_F$

$$\frac{D_F}{\lambda} = 2\left(\frac{D}{\lambda}\right)^2, \quad (9.20)$$

donde  $D$  es la mayor distancia característica del sistema de fuentes. Que el valor  $D_F/\lambda$  en este caso sea  $\approx 20$  está de acuerdo con el hecho de que ni en la figura 9.10 ni en la figura 9.11 (para una región ampliada, comprendida entre  $1.6 \leq Z/\lambda \leq 11.2$ ) podemos ver claramente la

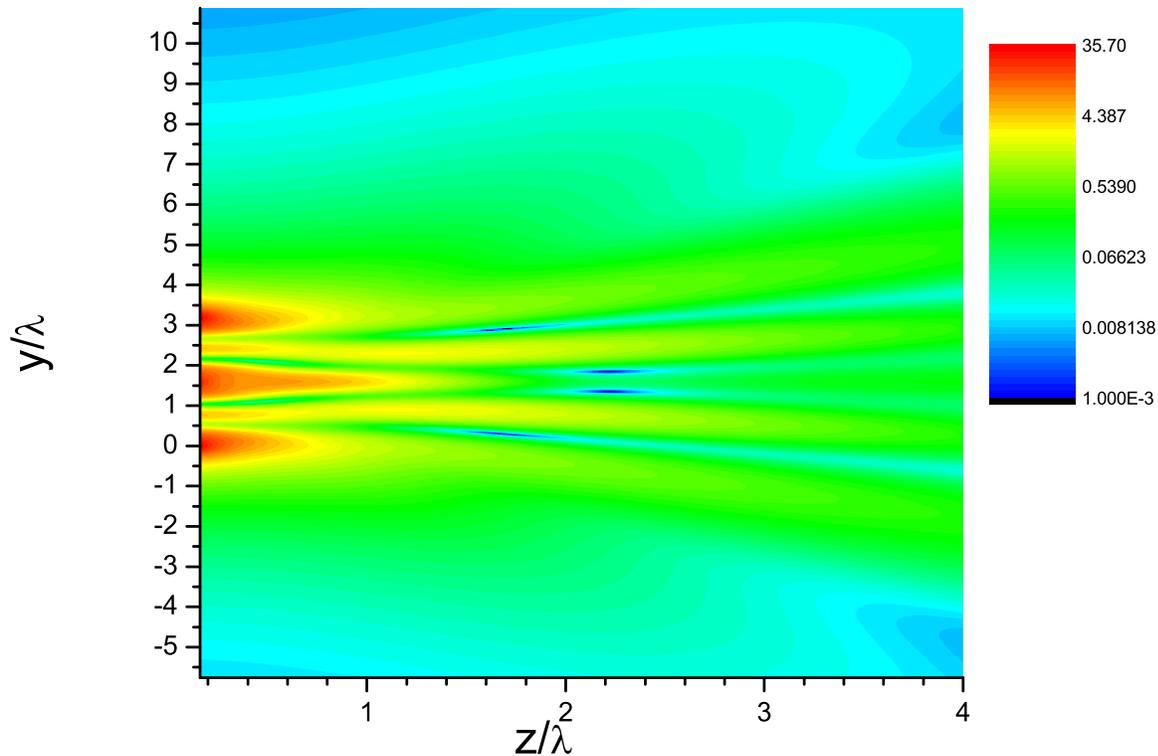


Figura 9.10: mapa de colores en escala logarítmica para la intensidad en la zona  $0.12 \leq Z/\lambda \leq 4$  producida por  $N = 3$  fuentes idénticas, colocadas en  $y_1 = 0$ ,  $y_2 = 1.6\lambda$  e  $y_3 = 3.2\lambda$ .

estructura ilustrada en el primer panel de la figura 9.7.

Sin embargo, en las figuras 9.10 y 9.11, especialmente a partir de distancias de algunas longitudes de onda, ya se empieza a vislumbrar una direccionalidad y las ganas de formar tres órdenes de interferencia, tal como se espera para la zona lejana. Esta tendencia se acentúa cuando el punto de observación se sigue alejando de las fuentes, tal como se muestra en los dos paneles de la figura 9.12.

No se puede negar que gracias a la ayuda del corto y sencillo script de la página 182 hemos sido capaces de revelar interesantes comportamientos físicos contenidos, pero bastante ocultos, en la expresión (9.4), comportamientos que habría sido muy difícil obtener sin ayuda computacional. Por este motivo, podríamos considerar que la física computacional es una rama intermedia entre la física teórica y la física experimental, es decir, un camino que complementa la teoría y el experimento. Es como tener un laboratorio particular, donde tenemos que entender muy bien los mecanismos y leyes físicas del fenómeno en estudio para poder instruir correctamente a nuestro ayudante computacional y que así nos entregue resultados confiables de cierto experimento. Los resultados del experimento ya están volcados en las figuras, ahora “midamos” nuevos observables. Por ejemplo, midamos en el panel superior de la figura 9.12 el ángulo que forma el “haz” cuyo centro parece pasar por el punto de coordenadas  $Y/\lambda = 16.5$  y  $Z/\lambda = 19.2$ . Si medimos “a ojo” (se podría mejorar inspeccionando la tabla con la que se hizo el gráfico) este ángulo desde el centro del sistema de fuentes, su tangente resulta igual al cociente entre  $14.9 (= 16.5 - 1.6)$  y  $19.2$ , valor que corresponde aproximadamente a un ángulo de  $37.81^\circ$ . De igual manera midamos

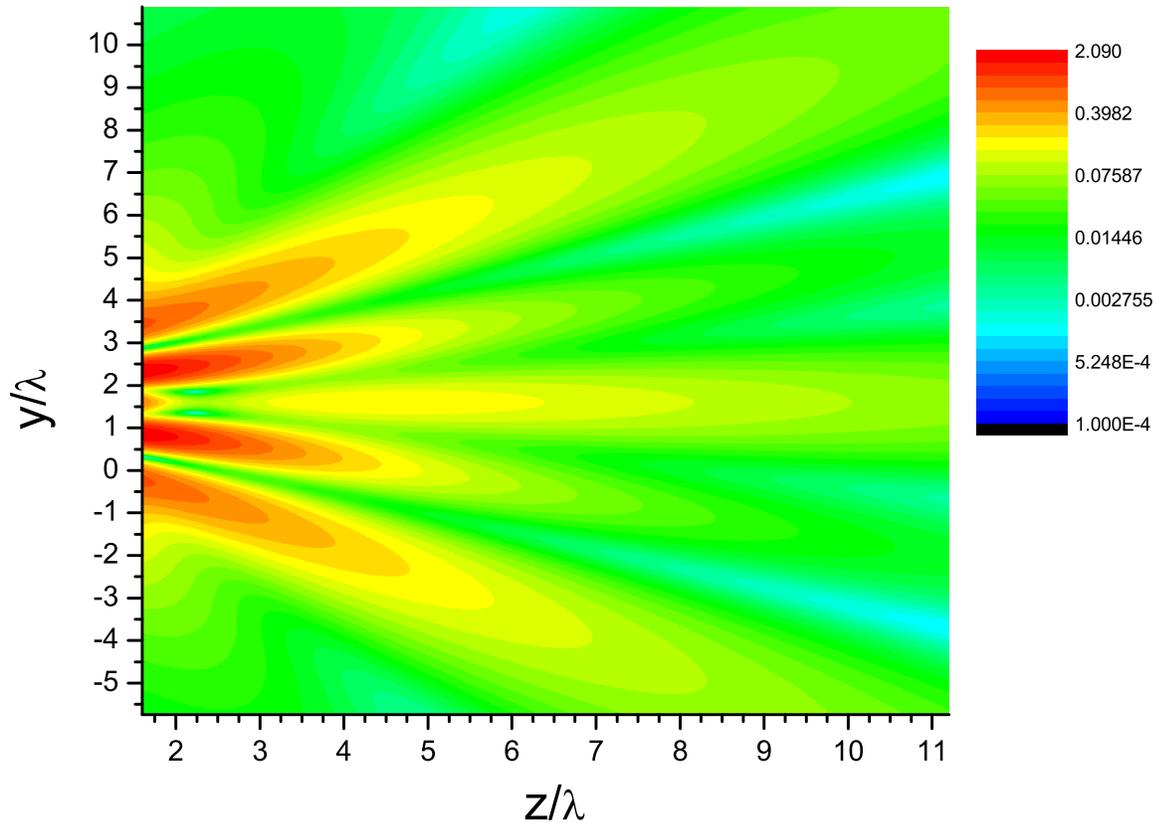


Figura 9.11: como la figura 9.10, pero para la zona  $1.6 \leq Z/\lambda \leq 11.2$ .

también en el panel superior el ángulo que forma el “haz” cuyo centro parece pasar por el punto de coordenadas  $Y/\lambda = -13.5$  y  $Z/\lambda = 19.2$ . Medido desde el centro del sistema de fuentes, su tangente resulta igual al cociente entre  $15.1 (= 13.5 + 1.6)$  y  $19.2$ , valor que ahora corresponde aproximadamente a un ángulo de  $-38.18^\circ$ . Como medimos  $Y/\lambda =$  “a ojo”, con un error de digamos media rayita, salvo el signo ambos valores son iguales dentro del error, una simetría esperada en los resultados de campo lejano. Más aún, la ec. (9.13) para campo lejano predice que los ángulos con la horizontal  $\theta_{\pm 1}$  formados por los órdenes de interferencia  $\pm 1$  satisfagan  $\sin \theta_{\pm 1} = \pm \lambda/d = \pm 1/1.6$ , esto es,  $\theta_{\pm 1} \approx \pm 38.68^\circ$ . Notar que hay otro “haz” mas intenso cuyo centro pasa muy cerca del punto con coordenadas  $Y/\lambda = 1.6$  y  $Z/\lambda = 19.2$ , o sea, que es casi paralelo a la línea horizontal que pasa por el centro del sistema de fuentes y entonces corresponde al orden de interferencia de la ec. (9.13) con  $m = 0$ . Las conclusiones son similares cuando se “miden” ángulos de haces en el panel inferior de la figura 9.12, donde además se puede apreciar un único máximo secundario entre dos órdenes de interferencia consecutivos, en coincidencia con la intensidad mostrada para el caso  $N = 3$  en la figura reffig:IPNequi2.

Realmente es notable la coincidencia entre los valores del “experimento numérico” y los valores teóricos correspondientes al caso límite en que el punto de observación está infinitamente alejado de las fuentes, especialmente si tenemos en cuenta todas las aproximaciones que hicimos para deducir la ec. (9.12), incluida la presencia de una lente.

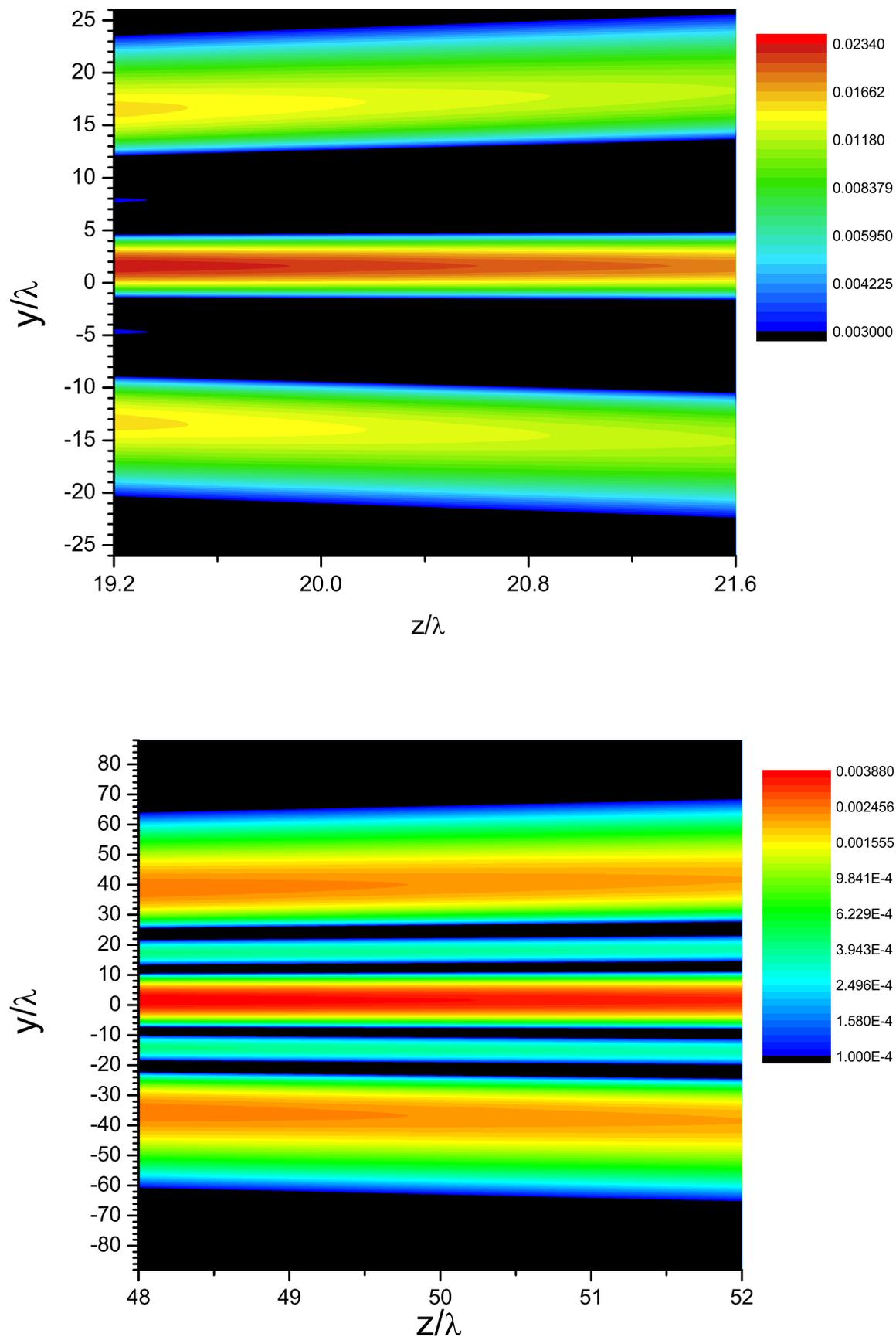


Figura 9.12: como en las figuras 9.10 y 9.11, pero para zonas más alejadas de las fuentes: a) en la zona  $19.2 \leq Z/\lambda \leq 21.6$  (panel superior) y b)  $48 \leq Z/\lambda \leq 52$  (panel inferior). Observar la estructura de haces bien direccionales, en las direcciones predichas por la ec. (9.12) y la correlación con los resultados mostrados en el primer panel de la figura 9.7

### 9.1.6. Límite continuo, $N \rightarrow \infty$ , $d \rightarrow 0$

Estudiemos un caso muy similar al de la figura 9.3, con la diferencia de que ahora aumentamos indefinidamente el número  $N$  de fuentes ( $N \rightarrow \infty$ ) mientras mantenemos constante la distancia  $b = (N - 1)d$  entre la primera y la última fuente. Claramente las fuentes están más juntas y entonces  $d \rightarrow 0$ . Y para que la intensidad resultante en situaciones de interferencia constructiva, que ya sabemos que crece con  $N^2$ , no diverja (y para que no salga demasiado caro pagar la energía consumida por las fuentes), a medida que aumentamos  $N$  vamos a ir disminuyendo la intensidad de cada fuente, de manera tal que la intensidad para interferencia constructiva sea independiente del número de fuentes. Es claro que esto significa que en la ec. (9.12) el límite que interesa es mantener constante el producto  $N^2 I$ , llamémoslo  $I_0$ . Y como escribimos en el renglón siguiente a la ec. (9.12), esto equivale a que las amplitudes individuales de cada fuente sean inversamente proporcionales a  $N$ .

Con estas consideraciones, procedemos a tomar el límite en la ec. (9.12). Según (9.11),  $\alpha \rightarrow 0$ , pero no  $N\alpha$ , ya que  $Nd \approx (N - 1)d = b$ . Entonces, el argumento del seno en el numerador de la ec. (9.12) se mantiene acotado e igual a

$$\lim_{N \rightarrow \infty} N\alpha \rightarrow \beta = \frac{k}{2} b \sin \theta = \pi \frac{b}{\lambda} \sin \theta, \quad (9.21)$$

mientras que el argumento del seno en el denominador de la ec. (9.12) tiende a cero y se puede aproximar por su argumento. Y al reemplazar  $I$  por  $I_0/N^2$ , en el denominador aparece la variable  $\beta^2$ , y así resulta que la intensidad producida por una distribución continua de fuentes puntuales que ocupan un segmento de longitud  $b$  cuando el punto de observación está infinitamente lejos, viene dada por la expresión

$$I_P = I_0 \left( \frac{\sin \beta}{\beta} \right)^2, \quad (9.22)$$

donde aparece la ya conocida función seno cardinal. Esta función ya apareció en las ecuaciones (4.54) y (4.57), en un problema que parece completamente distinto. Vemos que el problema de sumar infinitos cosenos de la misma amplitud y la misma fase con frecuencias distribuidas (uniformemente porque las amplitudes son iguales) en el intervalo  $\omega_1 \leq \omega \leq \omega_2$  da una amplitud moduladora que tiene la forma de la función  $\sin u/u$ , la misma dependencia que según (9.22) tiene ahora la amplitud de la perturbación que resulta de hacer interferir infinitas fuentes monocromáticas idénticas uniformemente distribuidas en el intervalo  $0 \leq y \leq b$ . La función seno cardinal, máxima para  $\beta = 0$  y ceros adyacentes para  $\beta = \pm \pi$  (es decir,  $\sin \theta = \pm \lambda/b$ ) está representada en las figuras 4.2 y 4.3 para distintos valores de la diferencia de frecuencias  $\Delta\nu$ , la intensidad es el cuadrado de esta función. Notar que de acuerdo a lo discutido para un número finito de fuentes,  $2\beta$  es la diferencia de fase asociada a los caminos provenientes de ambos extremos del sistema de fuentes. Luego, en los ceros adyacentes al máximo principal,  $2\beta = \pm 2\pi$  y el camino desde el extremo superior del sistema de fuentes estaría en contrafase con el camino desde el centro del sistema y todo camino desde un punto en la mitad superior del sistema estaría en contrafase con otro camino desde un punto en la mitad inferior del sistema.

Si comparamos el argumento  $u$  del seno cardinal en el problema del capítulo 4, dado por la ec. (4.58), con el argumento  $\beta$  del seno cardinal en la ec (9.22), ver lado derecho de (9.21), vemos que  $\Delta\nu$  se corresponde con  $b/\lambda$ , mientras que el tiempo  $t$  se corresponde con  $\sin \theta$ . Y de acuerdo con lo discutido en la ec. (4.59), la posición angular del primer cero puede ser usada como

medida de la localización. Antes era localización temporal, ahora es localización espacial, es decir localización en la variable  $\sin \theta$ : cuánto se extienden en el plano focal de la lente las manchas de interferencia producidas por el sistema de fuentes. La localización temporal de la moduladora era inversamente proporcional al contenido frecuencial  $\Delta\nu$  (cuanto más monocromática, menos localizada), la localización angular del diagrama de interferencia es inversamente proporcional al ancho  $b$  del sistema de fuentes (cuanto mayor es  $b$ , más se parece el sistema de fuentes a un sistema de onditas de Huygens sin limitación en la dirección  $y$ , es decir a un sistema de rayos paralelos no obstruidos, que según la DG debería dar en el foco de la lente solamente una mancha (imagen) puntual).

Es claro que con estas herramientas de interferencia podemos pasar a estudiar difracción, simplemente como una aplicación de la interferencia de las infinitas ondas secundarias de Huygens.

## 9.2. Difracción de Fresnel y de Fraunhofer

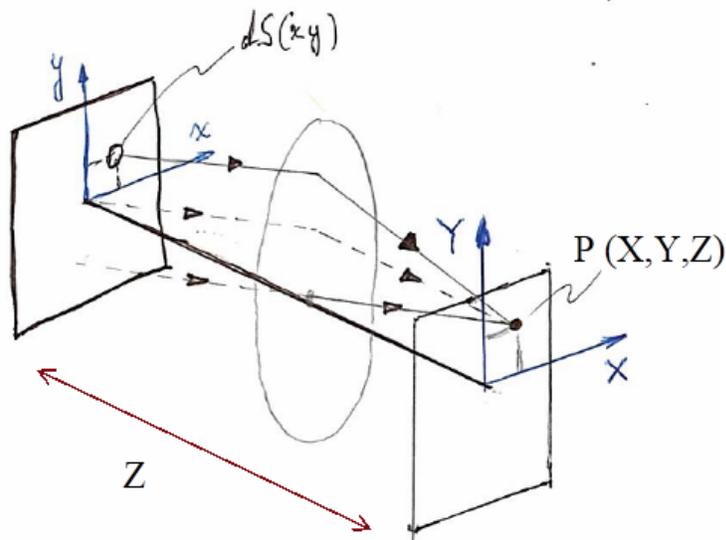


Figura 9.13: Un frente incidente plano ilumina un obstáculo colocado en el plano  $x-y$ . Cada elemento de superficie  $dS$  asociado con parte de un agujero del obstáculo está identificado por coordenadas genéricas  $(x, y)$  e irradia hacia adelante para dar lugar a la perturbación resultante en  $P(X, Y, Z)$ .

La difracción, o desviación de una onda de la propagación rectilínea predicha por la DG, ocurre siempre que la onda interactúa con un obstáculo y así la envolvente de las onditas de Huygens, que mantenía la forma del frente de ondas original antes de interactuar con el obstáculo, cambia su forma. Los frentes de onda planos dejan de ser planos y aparecen curvaturas como resultado de efectos de borde. Así los rayos perpendiculares a los frentes dejan de ser todos paralelos y los rayos divergen, o los frentes de onda esféricos, que seguían siendo esféricos, empiezan a cambiar

de forma también como resultado de efectos de borde, así aparecen rayos que dejan de tener todos el mismo centro. Cuando la onda interactúa con objetos grandes comparados con la longitud de onda, es decir, cuando se cumple la relación (6.1), los apartamientos de la propagación rectilínea son tan pequeños que se pueden ignorar. En cambio, cuando la onda interactúa con objetos de tamaños comparables con la longitud de onda, hay que investigar la envolvente de Huygens, o lo que es equivalente. hay que calcular la interferencia que producen en cada punto de observación las partes no obstruidas del frente de ondas que se propagaron más allá del obstáculo. Como se ve, no hay una diferencia física significativa entre interferencia y difracción. Y lo que sigue tiene que resultar muy familiar, ya que estamos estudiando interferencia desde el capítulo anterior.

Por razones fenomenológicas y metodológicas, el estudio de la difracción se divide en dos partes que se atacan por separado. En la **difracción de Fraunhofer**, tanto la fuente como el punto de observación están muy lejos del obstáculo difractor, idealmente ambos en infinito. La situación es muy parecida a la que vimos en la sección 9.1.3, si la fuente está en infinito, sobre el objeto incide una onda plana y todas las fuentes de Huygens están entonces en fase, resultando además que las diferencias de camino desde las distintas fuentes hasta el punto de observación  $P$  dependen linealmente de las etiquetas o coordenadas que identifican a las fuentes. En este caso, al alejarnos del objeto, y tal como vimos en los paneles de la figura 9.12, los diagramas de difracción cambian de tamaño, pero no de forma. En cambio, en la **difracción de Fresnel** la fuente o el punto de observación están “cerca” del obstáculo difractor y la situación es muy parecida a la que vimos en la sección 9.1.4, donde las diferencias de camino desde las distintas fuentes hasta el punto de observación *no dependen linealmente* de las etiquetas o coordenadas que identifican a las fuentes, tal como es evidente a partir de la ec. (9.18). En estos casos la distribución de intensidad de la perturbación resultante puede variar muy bruscamente cuando se cambia la posición del punto de observación, tal como vimos en las figuras 9.10 y 9.11.

Si bien estamos en condiciones de explorar los efectos de la difracción de Fresnel, partiendo de expresiones como (9.2) y usando procedimientos numéricos como los vistos en la sección 9.1.4, parece adecuado encarar primero el caso asintótico de Fraunhofer y dejar el estudio de la difracción de Fresnel para cursos avanzados.

### 9.3. Difracción de Fraunhofer

Por sus múltiples aplicaciones en ciencia y tecnología y por ser menos demandante matemáticamente, nos limitaremos a estudiar la difracción de Fraunhofer en el caso de *obstáculos planos*. La situación se esquematiza en la figura 9.13. La onda incidente proviene de una fuente puntual que, o está muy lejos o se coloca en el foco objeto de una lente (no dibujada). De esta manera, el frente de ondas incidente es plano y si el plano del frente de ondas es además paralelo al plano  $x - y$ , que es el plano que contiene al obstáculo, entonces, todos los elementos de superficie  $dS$  asociados con partes no obstruidas del obstáculo, identificados por sus coordenadas genéricas  $(x, y)$ , son iluminados al mismo tiempo. Y así resulta que están en fase *todas las fuentes secundarias*, que según el principio de Huygens iradian hacia adelante para dar lugar a la perturbación en los puntos de observación  $P(X, Y, Z)$  más allá del obstáculo.

Notar que la situación es completamente análoga a la considerada en la figura 9.3 y que nos había conducido a expresiones como (9.4), (9.7) y (9.8), con dos diferencias importantes:

- i) son nulos todos los  $\delta_n$  que aparecen en el argumento  $\Delta_n = \delta_n - kr_n$  de las exponenciales complejas de la eq. (9.4), o sea  $\Delta_n = -kr_n$ ;
- ii) la suma de perturbaciones se hace sobre una colección de fuentes que no están necesariamente alineadas y equiespaciadas, sino que están distribuidas de manera continua en las partes no obstruidas del plano  $x - y$  (las coordenadas  $x$  e  $y$  juegan el papel de la etiqueta  $n$ ).

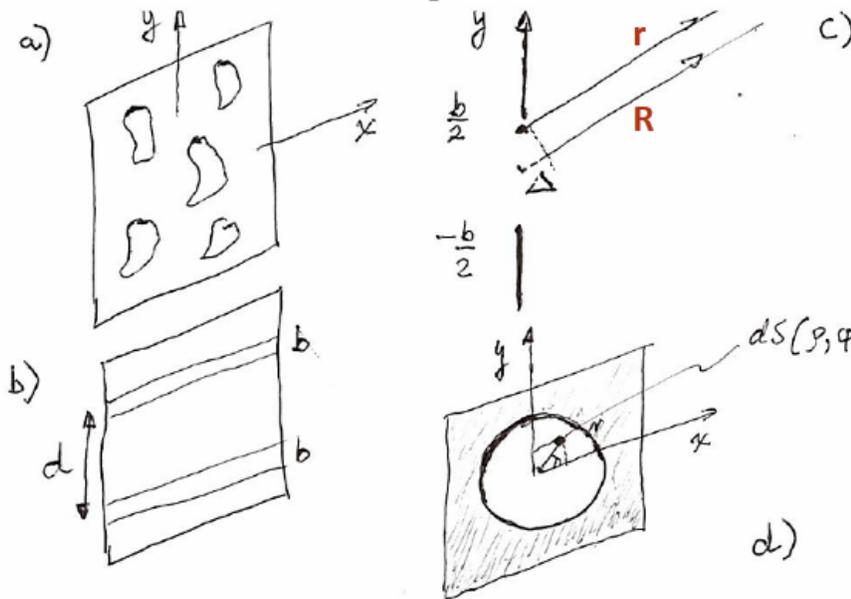


Figura 9.14: Algunos obstáculos planos: a) chapa con perforaciones de la misma o distinta forma; b) interferómetro de Young con dos rendijas lineales del mismo ancho  $b$  separadas una distancia  $d$ ; c) esquema para rendija lineal de ancho  $b$ ; d) disco opaco, o abertura circular en pantalla opaca, según el sombreado represente opacidades o transparencias.)

La figura 9.14 da algunos ejemplos de la diversidad contenida en la familia de obstáculos planos: una chapa con una o varias perforaciones de la misma o distinta forma (figura 9.14a), un interferómetro de Young con dos rendijas lineales del mismo ancho  $b$  separadas una distancia  $d$  (figura 9.14b) y un disco opaco o una abertura circular en una pantalla opaca (figura 9.14d, según la parte sombreada se considere opaca o transparente).

Volviendo al esquema de la figura 9.13, observamos que la perturbación en cada punto de observación  $P(X, Y, Z)$  tendrá aportes de todas las fuentes elementales situadas en  $(x, y, 0)$ . Cuanto más lejos del obstáculo se ubique el punto de observación  $P$ , más paralelos serán los rayos asociados a las perturbaciones provenientes de las fuentes elementales. En el caso límite de llevar el punto de observación al infinito, o lo que es lo mismo, de poner la lente y observar en su plano focal, solamente llegan a  $P$  las perturbaciones asociadas a rayos estrictamente paralelos y existe una relación unívoca entre las coordenadas  $X - Y$  del punto  $P$  y la inclinación de estos rayos paralelos provenientes de fuentes elementales no obstruidas, es decir, que a cada  $P(X, Y)$

solo llegan aquellas perturbaciones provenientes de distintos elementos de área  $dS$  que tengan la inclinación adecuada para converger en  $P$ .

## 9.4. Rendija de ancho $b$

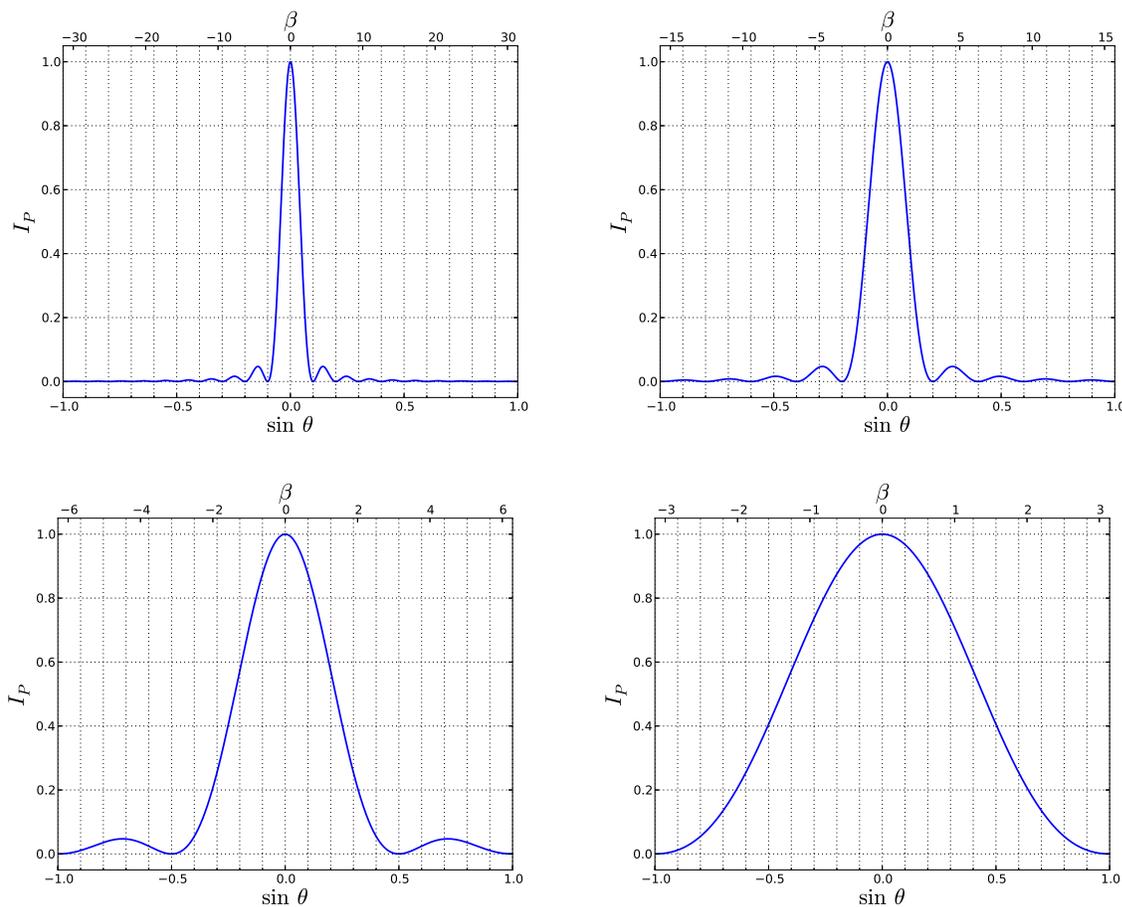


Figura 9.15: Intensidad  $I_P$  dada por la ec. (9.22) como función de  $\beta$  (eje superior) y de  $\sin \theta$  (eje inferior) para  $b/\lambda = 10$  (arriba izquierda), 5 (arriba derecha), 2 (abajo izquierda) y 1 (abajo derecha).

El caso más sencillo de obstáculo plano es el de una rendija lineal infinitamente larga orientada en la dirección  $x$  y de ancho  $b$  en la dirección  $y$ , que sería la idealización de una abertura rectangular de largo  $L$ , con  $L \gg b$ , para la que no esperamos difracción asociada con las limitaciones a lo largo de  $x$ , pero si esperamos difracción asociada con el ancho  $b$ , si  $b \approx \lambda$ . En el marco del principio de Huygens, estamos en presencia de esencialmente el mismo problema que ya consideramos en 9.1.6, aunque dada la simetría de este problema, en vez de sumar perturbaciones provenientes de fuentes puntuales podríamos simplificar las cuentas sumando directamente perturbaciones provenientes de fuentes lineales, es decir dividir la rendija en tiras de largo  $L$  y ancho  $dy$ . Así, cada fuente elemental ubicada entre  $y$  e  $y + dy$  emite una onda cilíndrica con amplitud proporcional al elemento de área  $dS = Ldy$  e inversamente proporcional a la potencia  $1/2$  de la distancia entre fuente y punto de observación (ver (4.31)), siempre y

cuando  $-b/2 \leq y \leq b/2$ , es decir, las partes obstruidas no contribuyen a la perturbación en P (ver figura 9.14c). Salvo este detalle de la dependencia de las amplitudes con la distancia entre fuente y punto de observación, para puntos lejanos sigue siendo válido considerar que las amplitudes de las perturbaciones individuales son iguales y sacarlas el factor común  $a$  de la suma en (9.7). Con esta discusión, no queda más remedio que aceptar que una rendija de ancho  $b$  produce un diagrama de difracción de Fraunhofer con las mismas características que el diagrama de interferencia de un sistema continuo de fuentes uniformemente distribuido en un segmento de longitud  $b$  y que en consecuencia la intensidad en la pantalla de observación responde a la ec. (9.22).

Podemos calcular fácilmente la distribución de intensidad para valores distintos de  $b/\lambda$  adaptando levemente el script de la página 179. Los resultados obtenidos para  $I_P/I_0$  se muestran en la figura 9.15 para  $b/\lambda = 10, 5, 2$  y 1. Además del máximo principal que corresponde a  $\beta = 0$ , hay máximos secundarios, aproximadamente a mitad de camino entre dos ceros consecutivos. Entonces, la intensidad relativa de los máximos secundarios cerca de  $\beta \approx \pm 3\pi/2$  es  $I/I_0 \approx 5\%$ , la de los máximos secundarios cerca de  $\beta \approx \pm 5\pi/2$  es  $I/I_0 \approx 2\%$ , etc. La intensidad de los máximos secundarios sigue cayendo rápidamente a medida que nos alejamos del centro de la pantalla y por este motivo normalmente no es fácil apreciarlos.

Notar que la posición angular del primer cero,  $\sin \theta = \pm \lambda/b$ , mide la localización en la variable espacial  $\sin \theta$  y cuantifica la difracción en el plano focal de la lente. Una onda plana no obstruida debería focalizarse como una imagen puntual en el foco de la lente, pero en vez de una imagen puntual se tiene una mancha de semiancho  $\lambda/b$  en la variable  $\sin \theta$ , que es muy parecido al semiancho angular si  $\theta \ll 1$ , como es razonable que pase en todo sistema óptico donde la difracción se considere un fenómeno indeseable (por ejemplo, en microscopios y telescopios).

## 9.5. Obstáculos planos con simetría de translación

Es instructivo volver a obtener los resultados de la sección anterior sin necesidad de pasar por el caso discreto, es decir, directamente a partir del principio de Huygens. Cuando el plano del frente de ondas incidente es paralelo al plano  $x - y$ , las fuentes secundarias asociadas con partes no obstruidas en  $dS$  y etiquetadas por  $(x, y)$  irradian en fase y se superponen en el punto de observación  $P(X, Y, Z)$ .

Si el obstáculo tiene simetría de translación en una dirección, digamos  $x$ , conviene considerar que las fuentes secundarias son tiras de largo  $L$  y ancho  $dy$  y que cada una emite una onda cilíndrica con amplitud proporcional al elemento de área  $dS = Ldy$  e inversamente proporcional a la potencia  $1/2$  de la distancia entre fuente y punto de observación. Y si en el campo lejano despreciamos la variación de la amplitud con la distancia, la amplitud de la perturbación en el punto  $P(X, Y, Z)$  proveniente de la fuentecita ubicada entre  $y$  e  $y + dy$  resulta

$$\frac{\mathcal{E}_0}{\sqrt{R}} L dy, \quad (9.23)$$

donde escribimos  $R$ , el camino al punto P desde el origen del sistema de coordenadas  $x - y$  que describe al obstáculo, en lugar de  $r$ , el camino al punto P desde la fuente elemental que estamos considerando, ver figura 9.14c. A esta altura tiene que quedar claro que la combinación (9.23)

juega en la superposición de Huygens el mismo papel que juega la amplitud  $a$  en la ecuación (9.7). De la figura 9.14c sale que  $r = R - y \sin \theta$  y entonces la expresión para la contribución elemental  $d\psi_P(t)$  resulta

$$d\psi_P(t) = \frac{\mathcal{E}_0 L}{\sqrt{R}} e^{i\omega t} e^{-ikR} e^{iky \sin \theta} dy, \quad (9.24)$$

que juega ahora el mismo papel que antes jugaban los términos de la sumatoria en (9.7).

En total paralelismo con (9.7), para obtener la perturbación resultante en el punto  $P$  solamente falta sumar sobre todas las fuentes, que equivale a integrar sobre todas las regiones del plano  $x - y$  no obstruidas, que para la simetría considerada quiere decir sumar sobre todas las tiras paralelas a  $x$ , es decir, integrar en la variable  $y$ . Tanto lío para decir *integrar en la variable  $y$* , pero decirlo tan resumido, sin saber decir todo lo anterior, es hacer matemática, no física. De esta manera, tenemos que

$$\psi_P(t) = \frac{\mathcal{E}_0 L}{\sqrt{R}} e^{i(\omega t - kR)} \int_{\text{PNO}} e^{iky \sin \theta} dy, \quad (9.25)$$

donde PNO quiere decir “partes no obstruidas”. Que no será una notación elegante, pero que se entiende muy bien. Notar que la integral es proporcional a la amplitud compleja de la oscilación armónica resultante en P, o sea que la intensidad en P será proporcional al valor absoluto al cuadrado de la integral en (9.25). Y también notar que la no elegancia de la notación está relacionada con nuestras pretensiones de generalidad, ya que (9.25) vale para cualquier distribución de rendijas paralelas en una pantalla opaca, por ejemplo también valdría para el sistema de dos rendijas en la figura 9.14b.

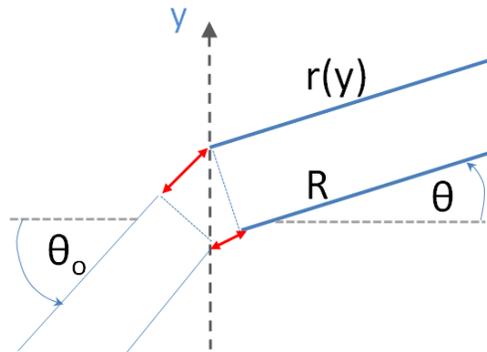


Figura 9.16: Cuando el plano del frente de ondas incidente no es paralelo al plano  $x - y$  del obstáculo, las fuentes elementales ubicadas entre  $y$  e  $y + dy$  **no** se encienden al mismo tiempo. Para un ángulo  $\theta_0 > 0$  (como en el dibujo), la fuente con etiqueta  $y > 0$  se enciende más tarde que la fuente con etiqueta  $y = 0$  (la que tiene un camino  $R$  hasta P) y el retardo entre ambas es proporcional a  $y$ .

Cuando el plano del frente de ondas incidente no es paralelo al plano  $x - y$  del obstáculo, las fuentes elementales ubicadas entre  $y$  e  $y + dy$  asociadas con partes no obstruidas **no** se encienden al mismo tiempo, es decir, que no están en fase. Como se observa en la figura 9.16, si la dirección de propagación del frente de ondas incidente forma un ángulo  $\theta_0 > 0$  con el eje  $x$ , la

fuente con etiqueta  $y > 0$  se enciende más tarde que la fuente con etiqueta  $y = 0$ . Es claro que el retardo es  $y \sin \theta_0$  (proporcional a  $y$  con su signo: las fuentes con  $y < 0$  están adelantadas, es decir tienen un retardo negativo). Y entonces, la cantidad  $\delta(y)$ , análoga a los  $\delta_n$  del argumento  $\Delta_n$  de las exponenciales complejas de la eq. (9.4), es  $\delta(y) = -ky \sin \theta_0$  ya que para nosotros (ver eq. (9.4)), una fase inicial positiva es un adelanto. Luego, la expresión para la contribución elemental  $d\psi_P(t)$  resulta

$$d\psi_P(t) = \frac{\mathcal{E}_0 L}{\sqrt{R}} e^{i\omega t} e^{-ikR} e^{iky(\sin \theta - \sin \theta_0)} dy, \quad (9.26)$$

y la perturbación total

$$\psi_P(t) = \frac{\mathcal{E}_0 L}{\sqrt{R}} e^{i(\omega t - kR)} \int_{\text{PNO}} e^{iky(\sin \theta - \sin \theta_0)} dy. \quad (9.27)$$

Vemos que la integral en la ec. (9.27) es formalmente idéntica a la de la integral en la ec. (9.25) y que se pasa de (9.25) a (9.27) simplemente reemplazando el parámetro  $\sin \theta$  en (9.25) por  $\sin \theta - \sin \theta_0$ . Esto significa que, independientemente de la forma y distribución de las ranuras en el plano  $x - y$ , la intensidad como función de la variable  $\sin \theta$  obtenida en el caso de incidencia inclinada es la misma función de  $\sin \theta$  obtenida en el caso de incidencia paralela al eje  $x$ , pero desplazada del origen en el valor  $\sin \theta_0$ . Cuidado, esto no significa que la distribución de intensidad obtenida en la pantalla de observación en el caso de incidencia inclinada sea simplemente una translación de la distribución de intensidad obtenida en la pantalla de observación cuando la incidencia es paralela al eje  $x$ .

### 9.5.1. Otra vez una rendija

Una vez familiarizados con la idea de Fresnel, de que la envolvente del principio de Huygens no es otra cosa que la interferencia entre onditas elementales, resulta instructivo verificar que la distribución de intensidad obtenida a partir de la ec. (9.25) para la difracción de Fraunhofer producida por una rendija de ancho  $b$  iluminada por un frente de ondas incidente paralelo al plano  $x - y$  es idéntica a la distribución de intensidad que resulta de la interferencia entre  $N$  fuentes idénticas, colineales, equiespaciadas y en fase, en el límite en que la distancia entre fuentes tiende a cero y el número de fuentes tiende a infinito manteniendo el sistema de fuentes en un segmento de longitud  $b$ .

Para una rendija de ancho  $b$  es muy fácil escribir los límites en la (9.25), porque las únicas partes no obstruidas son las que están en el intervalo  $-b/2 \leq y \leq b/2$ . Luego,

$$\psi_P(t) = \frac{\mathcal{E}_0 L}{\sqrt{R}} e^{i(\omega t - kR)} \int_{-b/2}^{b/2} e^{iky \sin \theta} dy. \quad (9.28)$$

La integral de una exponencial que depende linealmente de la variable de integración es la misma exponencial multiplicada por una constante y la diferencia de esta exponencial evaluada en límites con signo cambiado es proporcional a la función  $\sin \beta$  que aparece en la ec. (9.22). La cantidad  $\beta = \frac{k}{2} b \sin \theta$  se forma en el denominador con las constantes que aparecieron al hacer la integral. Y así verificamos que el principio de Huygens predice que la amplitud compleja difractada por una rendija de ancho  $b$  está en total acuerdo con el resultado de la ec. (9.22),

obtenido como un límite especial de la interferencia de  $N$  fuentes lineales equiespaciadas y uniformemente distribuídas en un segmento de longitud  $b$ . Cuando el frente plano incidente no es paralelo al plano de la rendija, las cuentas son iguales porque la exponencial sigue dependiendo linealmente de la variable de integración y todo lo dicho luego de la ec. (9.28) sigue valiendo, con la única diferencia de que ahora

$$\beta = \frac{k}{2} b (\sin \theta - \sin \theta_0), \quad (9.29)$$

un resultado que está de acuerdo con los desplazamientos que habíamos anticipado a partir de la ec. (9.27) para el caso general de cualquier obstáculo plano con simetría de translación.

Es interesante observar que el tamaño de la rendija se percibe en el campo lejano a través del tamaño angular del máximo central (figura 9.15). En cambio, muy cerca de la rendija predomina la influencia directa de las fuentes (secundarias), con una zona muy iluminada de ancho  $b$  y una zona de sombras, con intensidades mucho menores, tal como comprobamos en la simulación de campo cercano de la figura 9.10. Vemos entonces que en la zona de Fraunhofer, donde predomina la estructura de haces bien direccionales, la pantalla de observación tiene que estar iluminada esencialmente en una zona que corresponda al máximo central, mientras que en la zona cercana la pantalla de observación tiene que estar iluminada en una región del mismo orden que el ancho  $b$ . El máximo central se extiende en un rango angular del orden de  $\lambda/b$  (en realidad éste sería el rango del seno, pero nos estamos preparando para hacer una estimación para establecer un criterio, y en estas cosas todo vale, así que aproximamos nomás el seno y la tangente por el ángulo). Luego, en una pantalla colocada a una distancia  $Z$  del obstáculo, el máximo central ocupa una zona con dimensiones del orden de  $Z \lambda/b$  (arco, distancia por ángulo). Y entonces diremos que la zona donde predomina la direccionalidad de la difracción (la zona de Fraunhofer) es donde se cumple

$$Z \lambda/b \gg b, \quad (9.30)$$

es decir

$$Z/\lambda \gg (b/\lambda)^2, \quad (9.31)$$

que coincide con la definición de la distancia de Fraunhofer dada en la ec. (9.20) (excepto por un factor 2 ... y no es momento ahora de hacernos problema por un factor 2 en un criterio tan aproximado).

### 9.5.2. $N$ rendijas idénticas equiespaciadas

Un caso de sumo interés es el de una pantalla plana con  $N$  rendijas idénticas equiespaciadas. Si las rendijas tienen ancho  $b$ , están separadas una distancia  $d$  y están iluminadas por una onda plana incidente que viaja en la dirección  $x$ , entonces la integral sobre partes no obstruidas que aparece en la expresión (9.25) se escribe como una suma con  $N$  términos

$$\int_{\text{PNO}} = \int_{-b/2}^{b/2} + \int_{-b/2+d}^{b/2+d} + \dots + \int_{-b/2+(N-1)d}^{b/2+(N-1)d} . \quad (9.32)$$

Con el cambio de variable  $y' = y - d$ , la segunda integral resulta igual a la primera multiplicada por el factor  $e^{ikd \sin \theta}$ , con el cambio de variable  $y' = y - 2d$ , la tercera integral resulta igual a

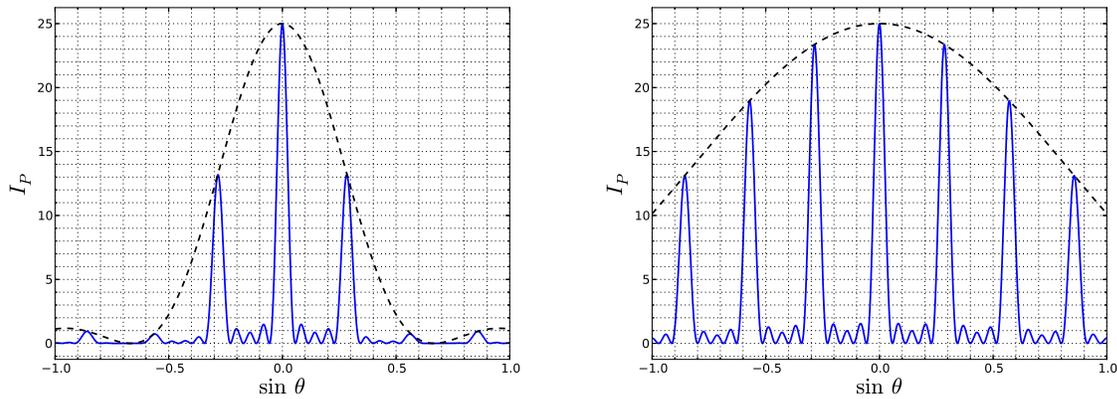


Figura 9.17: Intensidad  $I_P$  dada por la ec. (9.34) como función de  $\sin \theta$  para  $N = 5$  rendijas idénticas equiespaciadas con  $d/\lambda = 3.5$  y  $b/\lambda = 1.5$  (izquierda) y  $b/\lambda = 0.5$  (derecha). La línea punteada corresponde al término de difracción.

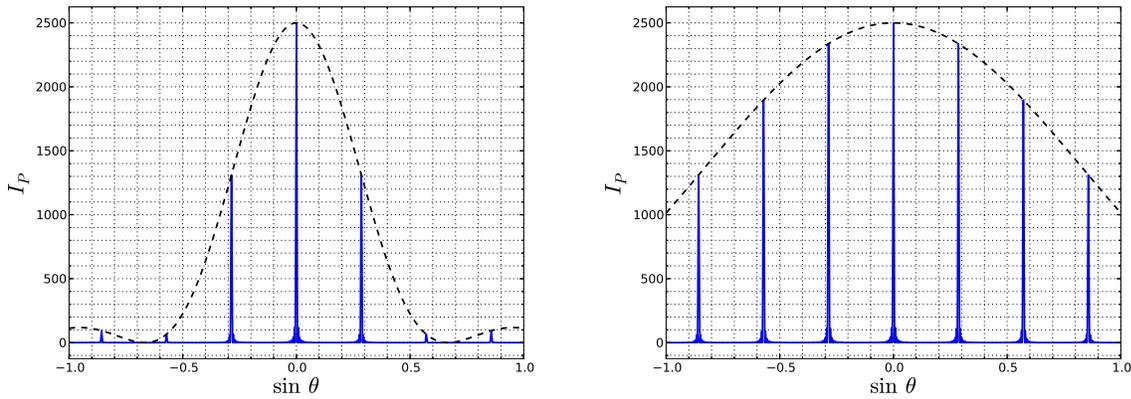


Figura 9.18: Como en la figura 9.17, pero para  $N = 50$ .

la primera multiplicada por el factor  $e^{i2kd \sin \theta}$ . Y repitiendo un procedimiento similar para las integrales restantes, la integral entre  $-b/2$  y  $b/2$  es un factor común en todos los términos y entonces

$$\begin{aligned} \int_{\text{PNO}} &= \left[ 1 + e^{ikd \sin \theta} + e^{i2kd \sin \theta} + \dots + e^{i(N-1)kd \sin \theta} \right] \int_{-b/2}^{b/2} \\ &= \sum_{n=0}^{N-1} \left[ e^{ikd \sin \theta} \right]^n \int_{-b/2}^{b/2}. \end{aligned} \quad (9.33)$$

Hemos obtenido que el lado izquierdo de la ec. (9.33), proporcional a la amplitud compleja en el punto P, es igual al producto de dos términos:

- i) una sumatoria idéntica a la que se obtuvo en (9.8) para calcular la amplitud compleja resultante de la **interferencia** lejana de  $N$  fuentes equiespaciadas a lo largo del eje  $y$ . Según la ec. (9.10) esta sumatoria terminaba siendo proporcional a  $\frac{\sin N\alpha}{\sin \alpha}$ .

- ii) una integral idéntica a la que se obtuvo en la ec. (9.28) para calcular la **difracción** por una única rendija. Según la ec. (9.22) esta integral terminaba siendo proporcional a  $\frac{\sin \beta}{\beta}$ .

Así obtenemos la intensidad en la pantalla de observación

$$I_P = I_0 \left( \frac{\sin N \alpha}{\sin \alpha} \right)^2 \left( \frac{\sin \beta}{\beta} \right)^2, \quad (9.34)$$

con  $\alpha$  y  $\beta$  funciones de  $\sin \theta$  (a cada P le corresponde un  $\theta$ ) dadas por las ecs. (9.11) y (9.21) respectivamente. La expresión (9.34) también vale cuando las rendijas están iluminadas a lo largo de una dirección inclinada un ángulo  $\theta_0$  con el eje  $x$  (figura 9.16), pero en este caso

$$\alpha = \frac{k}{2} d (\sin \theta - \sin \theta_0), \quad (9.35)$$

y  $\beta$  viene dado por (9.29).

La validez de la **factorización** de  $I_P$  como producto de un término llamado *de interferencia*, asociado con la interferencia lejana de  $N$  fuentes puntuales idénticas colocadas en el centro de cada ranura, y otro término llamado *de difracción*, representativo de la difracción de una sola ranura, excede este ejemplo particular. Es fácil darse cuenta que la factorización también vale si las ranuras idénticas no están equiespaciadas. En este caso no cambia la forma del llamado término de difracción, que sigue siendo  $(\sin \beta)^2/\beta^2$ , pero si cambia la forma del llamado término de interferencia, que no viene más dado por  $(\sin N \alpha)^2/(\sin \alpha)^2$ , sino que se obtiene a partir de la primera sumatoria en la ec. (9.7). Usando transformadas de Fourier y el llamado teorema de convolución, es fácil demostrar que la amplitud de Fraunhofer en la pantalla de observación producida por un obstáculo plano con  $N$  aberturas idénticas (no necesariamente con simetría de translación) siempre se factoriza como producto de dos términos: uno asociado con **la interferencia** lejana de  $N$  fuentes puntuales idénticas colocadas en el centro de cada abertura y otro representativo de **la difracción por una sola abertura**.

La distribución de intensidad (9.34) se muestra en la figura 9.17 junto con la modulación asociada al término de difracción (línea punteada) para  $N = 5$  rendijas idénticas equiespaciadas,  $d/\lambda = 3.5$ ,  $b/\lambda = 1.5$  (izquierda) y  $b/\lambda = 0.5$  (derecha). Observar que el término de difracción puede modificar notablemente el diagrama de intensidad y eventualmente provocar la desaparición de máximos de interferencia.

Al aumentar el número de rendijas, los órdenes de interferencia se afinan angularmente, tal como se aprecia en la figura 9.18, correspondiente a los mismos parámetros que en la figura 9.17 pero para  $N = 50$ . Este afinamiento angular se aprovecha en instrumentos llamados *redes de difracción* para separar líneas espectrales de fuentes con longitudes de onda muy parecidas. Las redes de difracción tienen muchas ventajas respecto a otros elementos dispersivos basados en prismas de vidrio. Cuando una fuente que contiene dos frecuencias muy parecidas, por ejemplo una fuente estelar, ilumina la red de difracción, cada frecuencia da lugar a un diagrama como los de las figuras 9.17 y 9.18. La situación es equivalente a tener dos fuentes lejanas con distintas frecuencias (y por lo general, con distinta intensidad en cada frecuencia). La intensidad resultante es entonces la suma de las intensidades individuales correspondientes a cada fuente, una con longitud de onda  $\lambda$  y la otra con longitud de onda  $\lambda + \Delta\lambda$ . En la figura 9.19a los diagramas asociados con ambas longitudes de onda se superponen de manera tal que la curva de intensidad

correspondiente a la suma (curva negra) de curvas de intensidad individuales ((curvas coloreadas) no exhibe ninguna característica que permita asociar posiciones de máximos con longitudes de onda: hay un máximo para el orden cero (este orden siempre es no dispersivo, porque tiene origen físico en la igualdad de caminos geométricos) y hay otros dos máximos para los órdenes  $\pm 1$ . Se dice que esta red no resuelve el valor de  $\Delta\lambda$  en este rango espectral. Como muestra la figura 9.19b esto es así porque la curva negra tiene un máximo que enmascara los máximos de las curvas coloreadas. Si los máximos de las curvas coloreadas estuvieran más alejadas entre sí (es decir, para mayor  $\Delta\lambda$ ), o si las curvas coloreadas fueran más angostas (es decir, para  $N$  más grande), entonces la curva envolvente negra debería revelar la existencia de los máximos individuales. En la figura 9.20 se muestra los diagramas obtenidos para parámetros físicos iguales a los de la figura 9.19, excepto que ahora el número de rendijas se decuplicó ( $N = 50$ ). Cerca de los picos individuales correspondientes a los órdenes  $\pm 1$  de  $\lambda$  y de  $\lambda + \Delta\lambda$ , no hay casi superposición y la curva negra tiene esencialmente la individualidad de las curvas coloreadas, salvo en la zona de máximos secundarios, que aquí es irrelevante desde el punto de vista espectroscópico por su baja intensidad. En este caso se dice que para esta  $\lambda$  la red resuelve la separación  $\Delta\lambda$ .

Para todo espectroscopista resulta relevante preguntarse cuál es el límite entre las situaciones mostradas en las figuras 9.19b (la red no resuelve  $\Delta\lambda$ ) y 9.20b (la red resuelve ampliamente  $\Delta\lambda$ ). Un criterio, con su cuota de arbitrariedad como todos los criterios, se conoce como *criterio de Rayleigh* porque fue propuesto por el físico inglés John William Strutt (Lord Rayleigh para los amigos). Rayleigh (Lord Rayleigh) hizo notar que si el máximo del orden  $m$  de la curva de  $\lambda + \Delta\lambda$  (la curva roja en la figura 9.19b) coincide con el primer cero adyacente al máximo del orden  $m$  de la curva de  $\lambda$  (la curva azul en la figura 9.19b), entonces la envolvente (la curva negra) tiene un mínimo local, aproximadamente a mitad de camino entre los dos máximos y con el contraste suficiente como para que los picos individuales puedan dar el presente. Y propuso que se tome esta distancia como indicador de la capacidad del instrumento para separar longitudes de onda (o lo que es lo mismo, frecuencias) cercanas.

Se conoce como *poder resolvente* de una red de difracción al cociente  $\lambda/\Delta\lambda$ , donde  $\Delta\lambda$  es la menor diferencia entre longitudes de onda parecidas que se pueden resolver. De acuerdo con el criterio de Rayleigh esta menor diferencia entre longitudes de onda parecidas se calcula usando que el máximo del orden  $m$  de la curva de  $\lambda$  está en un ángulo  $\theta$  que satisface la condición (9.13)

$$d \sin \theta = m \lambda,$$

mientras que al máximo de orden  $m$  de la curva para  $\Delta\lambda$  le corresponde un valor mayor de  $\sin \theta$  (es decir un valor mayor de  $\theta$ ), que también satisface la condición (9.13)

$$d [\sin \theta + \Delta(\sin \theta)] = m (\lambda + \Delta\lambda).$$

Entonces

$$d \Delta(\sin \theta) = m \Delta\lambda. \quad (9.36)$$

Por otra parte, y de acuerdo a lo discutido en la página 176, la distancia entre la posición del máximo de orden  $m$  y la posición de su primer cero adyacente, medida en términos de la variable  $\sin \theta$  es

$$d \Delta(\sin \theta) = \lambda/N. \quad (9.37)$$

Y según el criterio de Rayleigh, la situación límite de resolución ocurre cuando la distancia entre el máximo de orden  $m$  de las curvas de  $\lambda$  y de  $\lambda + \Delta\lambda$  (lado izquierdo de la penúltima ecuación)

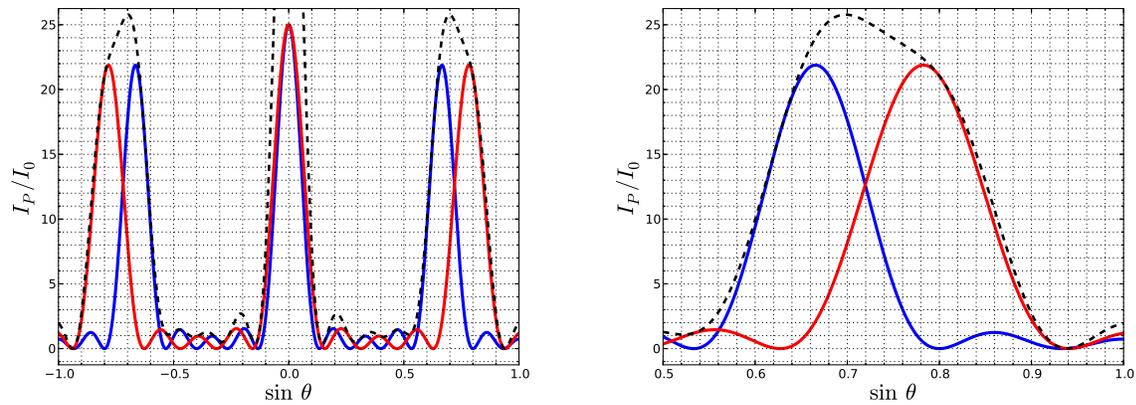


Figura 9.19: Propiedades dispersivas de  $N = 5$  rendijas idénticas equiespaciadas con  $d = 1 \mu\text{m}$  y  $b = 0.3 \mu\text{m}$ . La línea negra punteada representa el diagrama de difracción de Fraunhofer de una fuente que emite con longitudes de onda  $\lambda = 1 \mu\text{m}$  y  $\lambda + \Delta\lambda = 1.176 \mu\text{m}$ , la línea azul (roja) corresponde a una fuente monocromática de  $\lambda$  ( $\lambda + \Delta\lambda$ ). En el panel de la derecha se muestra ampliada la zona del orden de interferencia  $+1$ .

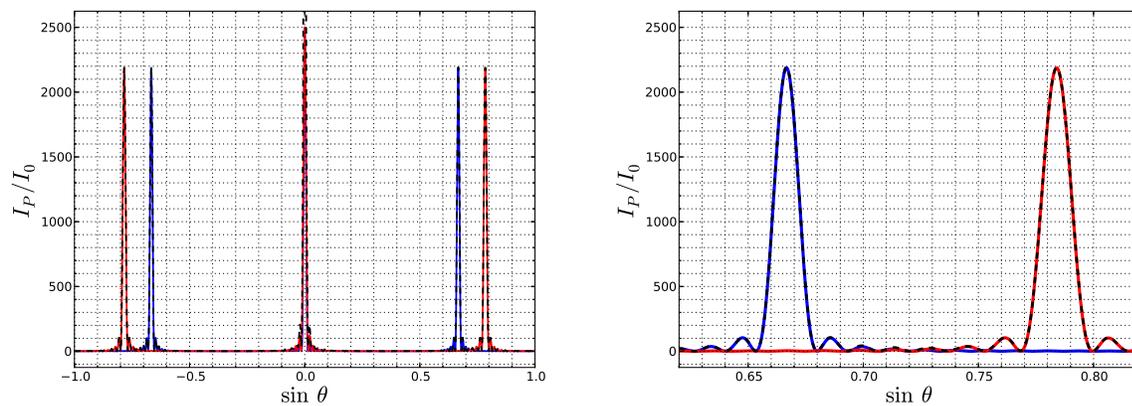


Figura 9.20: Como en la figura 9.19, pero para  $N = 50$ .

resulta igual a la distancia entre el máximo del orden  $m$  y su primer cero adyacente en la curva correspondiente a  $\lambda$  (lado izquierdo de ecuación previa), es decir  $m \Delta\lambda = \lambda/N$ , de donde se obtiene la siguiente expresión

$$\frac{\lambda}{\Delta\lambda} = m N, \quad (9.38)$$

que dice que el poder resolvente de una red de difracción depende del número  $N$  de rendijas (“líneas” en la jerga espectroscópica) y del orden espectral  $m$ . Notar que implícitamente también depende de  $d$  a través de la relación (9.17), porque  $|m| \geq 1$ . Cuando el lado derecho de la ec.

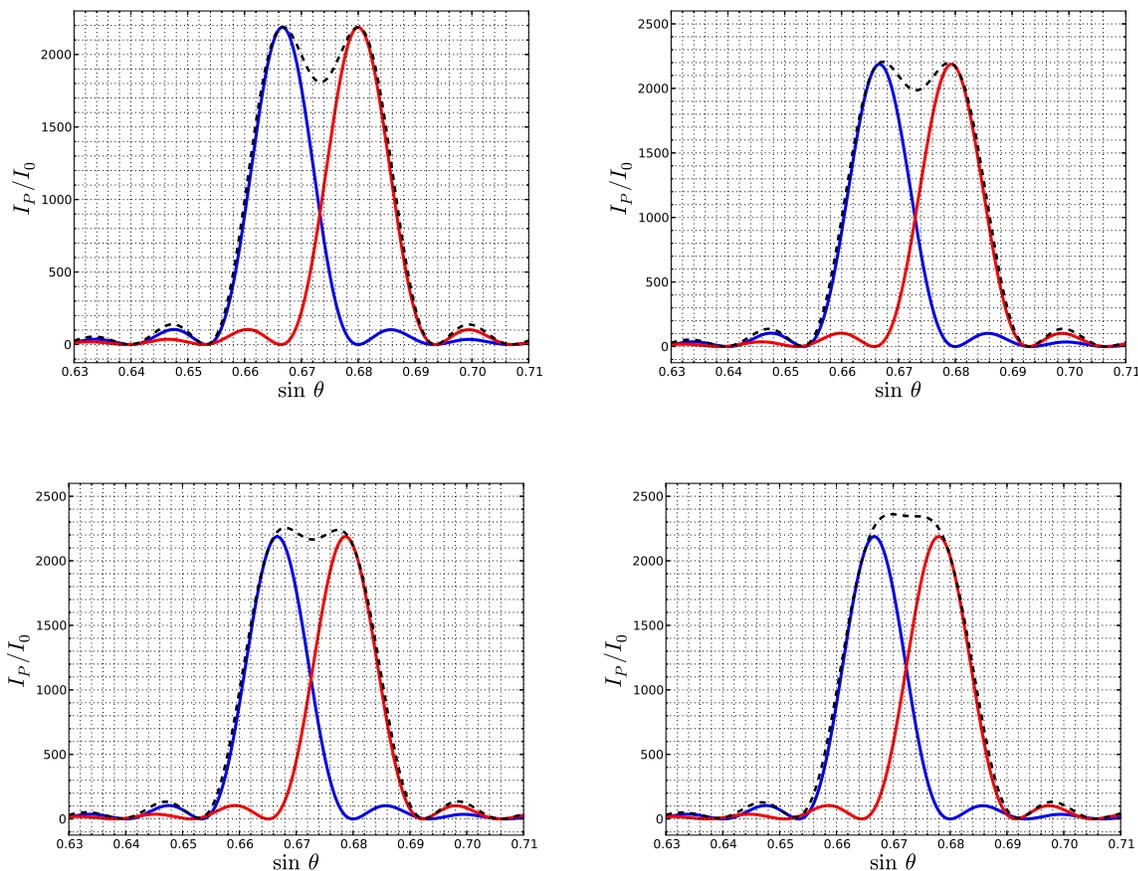


Figura 9.21: Distribución de intensidad cerca del orden  $m = +1$  para  $N = 50$  rendijas idénticas equiespaciadas con  $d = 1 \mu\text{m}$  y  $b = 0.3 \mu\text{m}$ . La línea negra punteada representa la intensidad de una fuente que emite con longitudes de onda  $\lambda = 1 \mu\text{m}$  y  $\lambda + \Delta\lambda$  para distintos valores de  $\Delta\lambda$ . La línea azul (roja) corresponde a una fuente monocromática de  $\lambda$  ( $\lambda + \Delta\lambda$ ). i) arriba izquierda:  $\Delta\lambda$  es igual al límite justo calculado con el criterio de Rayleigh, ec. (9.38); arriba derecha:  $\Delta\lambda$  es igual al 95 % del valor del límite justo; abajo izquierda:  $\Delta\lambda$  es igual al 90 % del valor del límite justo; abajo derecha:  $\Delta\lambda$  es igual al 85 % del valor del límite justo.

(9.36) es justo igual al lado derecho de la ec. (9.37), se dice que  $\Delta\lambda$  es el *límite justo de resolución* para esta  $\lambda$ . La definición del límite justo de resolución es un poco arbitraria, como se aprecia en las simulaciones mostradas en la figura (9.21).

En la práctica, hay una tensión entre la tendencia a usar órdenes de difracción altos, para aumentar el poder resolvente, y la tendencia a no tener demasiados órdenes, para no redistribuir

la intensidad recibida. En ciertas aplicaciones la intensidad de la señal que llega al instrumento espectroscópico puede ser muy pequeña (en estrellas lejanas, por ejemplo) y repartirla en órdenes de difracción que no serán utilizados sería “malgastar” la intensidad del haz incidente. También hay un problema relacionado con la existencia de la campana de difracción, que por esas casualidades de la naturaleza tiene su máximo justamente en el orden  $m = 0$ , que no exhibe propiedades dispersivas. Para maximizar la fracción de la intensidad incidente que llega al orden que tiene interés espectroscópico, se han desarrollado diversas técnicas que permiten desplazar el máximo de la campana de difracción de la posición del máximo de orden  $m = 0$ . Lograr este efecto implica modificar el sistema de ranuras de manera tal que el efecto de interferencia en el punto P de las onditas secundarias que salen de una sola ranura cambie sustancialmente, pero sin que se modifique la interferencia en el punto P entre las  $N$  perturbaciones que provienen de cada ranura. Una manera de lograr el efecto, conocido como *efecto de blaze*, es cambiando la transmisión de las rendijas, por ejemplo colocando prismas.

## 9.6. Obstáculos planos, caso general

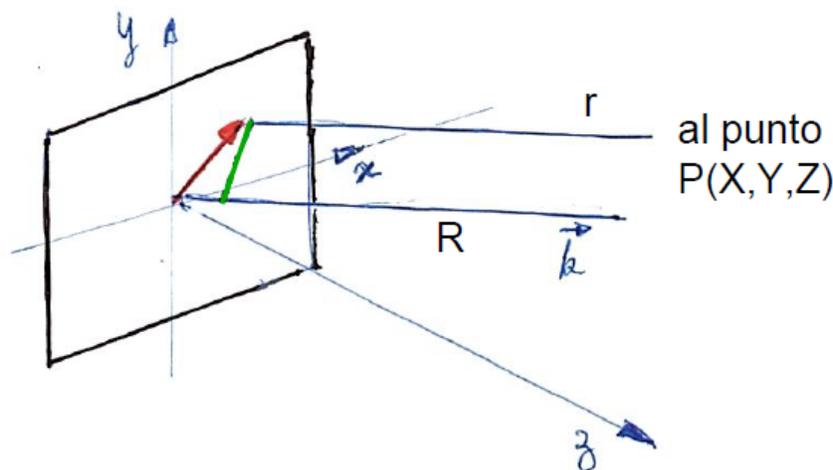


Figura 9.22: el vector  $x\hat{x} + y\hat{y}$  apunta desde el origen del obstáculo al diferencial de superficie  $dS$  ubicado entre  $x$  y  $x + dx$  y entre  $y$  y  $y + dy$ . La distancia  $r$  entre  $dS$  y P difiere de la distancia  $R$  entre el origen y P en la proyección del vector  $x\hat{x} + y\hat{y}$  a lo largo del versor unitario  $\hat{k}$  paralelo al vector de onda representativo de las ondas planas que llegan a P

Consideremos ahora el caso general de obstáculos planos esquematizado en la figura 9.13 y generalicemos el procedimiento de la sección 9.5, que fue hecho con la hipótesis de obstáculos planos con simetría de translación. En vez de tiras, las fuentes elementales son elementos de área  $dS = dx dy$ . Y en vez emitir una onda cilíndrica, cada elemento de área emite una onda esférica que, como antes, tiene amplitud proporcional al elemento de área. Sin embargo, en vez de ser a la inversamente proporcional a la potencia  $1/2$  de la distancia  $r$  entre fuente y punto de observación, la amplitud es ahora inversamente proporcional a  $r$ .

Para campo lejano seguimos despreciando la variación de la amplitud con  $r$  ( $r \approx R$ ) y entonces

la contribución elemental  $d\psi_P(t)$  a la perturbación en el punto  $P(X, Y, Z)$ , proveniente de la fuentecita ubicada entre  $x$  y  $x + dx$  y entre  $y$  e  $y + dy$ , resulta

$$d\psi_P(t) = \frac{\mathcal{E}_0}{R} e^{i\omega t} e^{-ikr} dx dy, \quad (9.39)$$

donde  $R$  representa el camino al punto  $P(X, Y, Z)$  desde el origen del sistema de coordenadas  $x - y$  del obstáculo. En la figura 9.22 se muestran detalles de estos caminos cerca del obstáculo.

Si el punto  $P$  está en el foco de una lente, los caminos a lo largo de  $r$  y de  $R$  son paralelos, es decir tienen la misma inclinación y representan rayos perpendiculares a los frentes de onda que llegan a la lente para converger en  $P$ . En otras palabras, estos rayos están en la dirección del versor unitario  $\hat{k}$  paralelo al vector de onda  $\vec{k}$  representativo de las ondas planas que llegan a  $P$ . Para referir la distancia  $r$ , entre cada fuente elemental  $dS(x, y)$  y el punto de observación, al camino  $R$  desde el origen (donde puede que haya o no fuente elemental, todo depende de si el origen no está o está obstruido), notemos que los trayectos por estos caminos son iguales desde cualquier plano perpendicular a  $\hat{k}$ . Luego,  $r$  y  $R$  difieren en la proyección sobre  $\hat{k}$  del vector  $x\hat{x} + y\hat{y}$  que apunta a  $dS$  desde el origen y entonces el argumento de la exponencial espacial en 9.39 se escribe

$$kr = kR - k \left[ \hat{k} \cdot (x\hat{x} + y\hat{y}) \right] = kR - k_x x - k_y y, \quad (9.40)$$

con  $k_x = \vec{k} \cdot \hat{x}$  y  $k_y = \vec{k} \cdot \hat{y}$ . Así, la contribución elemental (9.39) se puede reescribir como

$$d\psi_P(t) = \frac{\mathcal{E}_0}{R} e^{i(\omega t - kR)} e^{i(k_x x + k_y y)} dx dy, \quad (9.41)$$

una cantidad que juega el mismo papel que jugaban los términos de la sumatoria en (9.7).

Observar que la fase de la contribución elemental (9.41) es otra vez una función lineal de las coordenadas que etiquetan a la fuente elemental en  $dS(x, y)$ . Sumando sobre todas las fuentes, es decir, integrando sobre todas las regiones del plano  $x - y$  no obstruidas, se obtiene finalmente la siguiente expresión

$$\psi_P(t) = \frac{\mathcal{E}_0}{R} e^{i(\omega t - kR)} \int_{\text{PNO}} e^{i(k_x x + k_y y)} dx dy, \quad (9.42)$$

que representa la perturbación total en el punto  $P$  debida a la difracción por un obstáculo plano caracterizado por partes que son completamente transparentes o completamente opacas (es decir, no hay partes con zonas intermedias, más o menos transparentes).

Las cantidades  $\hat{k} \cdot \hat{x} = k_x/k$  y  $\hat{k} \cdot \hat{y} = k_y/k$  en (9.42) definen la inclinación de la dirección de observación a lo largo de los ejes  $x$  e  $y$  respectivamente. Representan la dirección en la que debería apuntar un telescopio para enfocar al punto  $P$  muy lejano y juegan el rol análogo al de la cantidad  $\sin \theta$  en la ec. (9.25). Las componentes de  $\hat{k}$  se pueden escribir en términos de los ángulos de coordenadas esféricas  $\vartheta$  (el ángulo polar entre  $\hat{k}$  y el eje  $z$ ) y  $\varphi$  (el ángulo azimutal, entre la proyección de  $\hat{k}$  en el plano  $x - y$  y el eje  $x$ ) y estos ángulos se expresan en función de las coordenadas  $X, Y$  y  $Z$  del punto  $P$ . En el caso de usar una lente de distancia focal  $f$ ,  $\tan \vartheta = \sqrt{X^2 + Y^2}/f$  y  $\tan \varphi = Y/X$ . Es instructivo seguir estos pasos para confirmar que a cada par  $X - Y$  (a cada punto  $P$  de la pantalla de observación) le corresponde un par  $\vartheta - \varphi$  (una inclinación del rayo que pasa por el centro de la lente y no se desvía) y a cada par  $\vartheta - \varphi$

le corresponde un par  $k_x - k_y$ . Dicho de otra manera: de todas las perturbaciones emitidas por las onditas de Huygens, solamente se detectan en el punto  $P$  aquellas que tienen la inclinación correcta, descartando todas las demás. Y eligiendo la forma de las aberturas en el plano  $x - y$  se controla la superposición de componentes con la misma inclinación.

Cuando el plano del frente de ondas incidente no es paralelo al plano  $x - y$ , las fuentes elementales no se encienden al mismo tiempo. Si a la dirección de propagación del frente de ondas incidente le corresponde el vector de onda  $\vec{k}_0$ , es fácil repetir los pasos que dimos en el caso de rendijas y obtener que

$$\psi_P(t) = \frac{\mathcal{E}_0}{R} e^{i(\omega t - k R)} \int_{\text{PNO}} e^{i[(k_x - k_{0x})x + (k_y - k_{0y})y]} dx dy, \quad (9.43)$$

### 9.6.1. Abertura rectangular

Para una abertura rectangular de lados  $a$  (según  $x$ ) y  $b$  (según  $y$ ) centrada en el origen, los límites en las (9.42) y (9.43) son  $-a/2 \leq x \leq a/2$  y  $-b/2 \leq y \leq b/2$ . Luego, como el integrando se factoriza en una función exclusivamente de  $x$  por una función exclusivamente de  $y$ , se obtiene un producto de dos integrales con la misma forma que en la ec. (9.28). De esta manera, la amplitud compleja es proporcional a un producto de senos cardinales, uno correspondiente a la integral en  $x$  y otro a la integral en  $y$ . La intensidad en el punto  $P$  queda

$$I_P = I_0 \left( \frac{\sin \beta_x}{\beta_x} \right)^2 \left( \frac{\sin \beta_y}{\beta_y} \right)^2, \quad (9.44)$$

con

$$\beta_x = \frac{1}{2}(k_x - k_{0x}) a = \pi \frac{a}{\lambda} \left[ (\hat{k} \cdot \hat{x}) - (\hat{k}_0 \cdot \hat{x}) \right], \quad (9.45)$$

$$\beta_y = \frac{1}{2}(k_y - k_{0y}) b = \pi \frac{b}{\lambda} \left[ (\hat{k} \cdot \hat{y}) - (\hat{k}_0 \cdot \hat{y}) \right]. \quad (9.46)$$

Suponiendo  $\hat{k}_0 = \hat{k}_0 = 0$  (frente de ondas incidente paralelo al plano  $x - y$ ), la figura de difracción de la abertura rectangular tiene un máximo principal con intensidad  $I_0$  en  $\beta_x = \beta_y = 0$ , acompañado de un doble juego de máximos secundarios y que están limitados por ceros en inclinaciones definidas por  $\beta_x = n\pi$  y  $\beta_y = m\pi$ , con  $n$  y  $m$  enteros distintos de cero. El rango de inclinaciones horizontales y verticales abarcadas por el máximo principal es  $\Delta(\hat{k} \cdot \hat{x}) = 2\lambda/a$  y  $\Delta(\hat{k} \cdot \hat{y}) = 2\lambda/b$  respectivamente.

### 9.6.2. Abertura circular

a ser actualizado próximamente

## 9.7. Óptica de Fourier

a ser actualizado próximamente