

Apuntes de Modelos Matemáticos II

José A. Cañizo

Curso 2019-2020

Estos son apuntes de Modelos Matemáticos II, utilizados para el grupo de tarde del grado en Matemáticas, Universidad de Granada. Son un borrador, pueden contener errores y no contienen exactamente lo que se ha visto en clase. Los pongo aquí sólo como ayuda para la asignatura. El contenido oficial de la asignatura es siempre el que se ha visto en clase.

Últimos cambios: 21 de abril de 2020.

Índice

1. Introducción al cálculo de variaciones	4
1.1. Definiciones básicas	5
1.2. El caso de \mathbb{R}^d	6
2. Condiciones necesarias de extremo	6
2.1. Variación de una función. Condiciones necesarias de extremo	6
2.2. Ecuaciones de Euler-Lagrange	8
2.2.1. Para una variable	8
2.2.2. Para varias variables	12
2.3. Regularidad de los extremos	12
3. Condiciones suficientes de extremo: convexidad	14
3.1. El caso de \mathbb{R}^d	17
4. Extremos con restricciones	18
4.1. Extremos con restricciones en \mathbb{R}^d	18
4.2. Extremos con restricciones en dimensión infinita	19
4.3. Ecuaciones de Euler-Lagrange	21
4.3.1. Restricciones del tipo $\varphi(x, y_1(x), \dots, y_q(x)) = 0$	22
4.3.2. Restricciones del tipo isoperimétrico (es decir, de tipo integral)	23
4.3.3. Multiplicadores de Lagrange para restricciones integrales: demostración	24
4.4. Convexidad y restricciones	25

5. Algunos problemas variacionales	25
5.1. El problema isoperimétrico	25
5.2. Superficies minimales de revolución	27
5.3. La catenaria	29
5.4. La braquistócrona	30
5.5. Ejercicios	31
6. Principios variacionales en mecánica	33
7. Problemas de contorno	36
7.1. Forma autoadjunta de una ecuación de segundo orden	37
7.2. Alternativa de Fredholm	37
7.2.1. Demostración del teorema de Fredholm para condiciones de con-	
torno separadas	41
7.2.2. Demostración de Teorema de Fredholm para condiciones de con-	
torno periódicas	44
7.3. Formulación variacional de los problemas de contorno	45
7.4. Problemas de Sturm-Liouville	46
7.4.1. Demostración del Teorema de Sturm-Liouville	52
7.4.2. Relación con el cálculo de variaciones	53
7.5. Un problema de contorno de orden superior: el problema de la viga . .	57
8. Ejercicios	59
9. Series de Fourier	60
9.1. Serie de Fourier en $[-\pi, \pi]$	60
9.2. Serie de Fourier en un intervalo general	62
9.3. Series de Fourier en senos y cosenos	64
9.4. Otras representaciones de las series de Fourier	65
9.5. Series de Fourier y derivadas	66
9.5.1. Para la serie de Fourier en senos	68
9.5.2. Para la serie de Fourier en cosenos	69
9.6. Convergencia puntual y uniforme de series de Fourier	70
9.7. Ejercicios	70
10.La ecuación de ondas	71
10.1. Derivación de la ecuación	71
10.2. Propiedades básicas	74
10.3. Soluciones en \mathbb{R} : fórmula de d'Alembert	75
10.4. Separación de variables en dimensión 1	76
10.5. Separación de variables en dimensión 2	78
10.6. Problema de valores iniciales para la ecuación de ondas	79
10.7. Ejercicios	82
11.Las ecuaciones de Laplace y Poisson	82
11.1. La ecuación de Laplace	83
11.1.1. Propiedades básicas	83
11.1.2. Solución fundamental	84

11.1.3. Armónicos esféricos en \mathbb{R}^3	85
11.2. La ecuación de Poisson en \mathbb{R}^d	85
11.2.1. Solución de la ecuación de Poisson en \mathbb{R}^d	85
11.2.2. Una interpretación física: potencial gravitatorio y electrostático	87
11.3. La ecuación de Laplace en el cuadrado: separación de variables	90
11.4. Ejercicios	91
12.El Teorema de Lax-Milgram y su aplicación a las EDPs	91
12.1. El Teorema de Lax-Milgram	91
12.2. El concepto de derivada débil	93
12.3. El espacio de Sobolev $H^k(\Omega)$	97
12.4. Aplicación a la ecuación de Poisson	97
12.5. Aplicación a la ecuación $\Delta u + \lambda u = f$	99
12.6. Ejercicios	100
13.La transformada de Fourier	101
14.La ecuación del calor	104
14.1. Solución fundamental	107
14.2. Solución usando la transformada de Fourier	111
14.3. Solución en un intervalo: separación de variables	113
14.4. Solución en el cuadrado: separación de variables	114
15.La ecuación de Schrödinger	114
15.1. Ejemplos en dimensión 1	116
15.1.1. Partícula libre	116
15.1.2. Partícula en una caja	117
15.1.3. Partícula en un pozo de potencial finito	117
15.2. La partícula libre	118
15.3. El átomo de hidrógeno	118
16.Ley de acción de masas	119
17.Algunos modelos de poblaciones	120
18.Ecuaciones de reacción-difusión	120
19.Inestabilidad de Turing	120
A. Fórmulas trigonométricas útiles	124
B. Cálculo en varias variables	124
B.1. Notación para las derivadas	124
B.2. Diferenciabilidad	125
B.3. Operadores diferenciales comunes	126
B.4. Coordenadas esféricas en \mathbb{R}^3	127
B.4.1. El laplaciano en coordenadas esféricas	128
B.5. Coordenadas esféricas en \mathbb{R}^d	129
B.6. Derivación bajo la integral	130

B.7. Teorema de la divergencia y fórmulas de Green	131
C. Ecuaciones diferenciales ordinarias lineales	131
D. Ecuaciones diferenciales ordinarias especiales	134
D.1. La ecuación de Legendre	134
D.2. La ecuación de Legendre asociada	135
E. Integral de Lebesgue	135
F. Convolución y aproximación	136
G. Series de funciones	137
G.1. Modos de convergencia	137
G.2. Derivación de series de funciones término a término	139
H. Espacios de Hilbert	142
H.1. Aplicaciones bilineales y sesquilineales	142
H.2. Espacios de Hilbert	142

1. Introducción al cálculo de variaciones

Un problema común en matemáticas es el de encontrar el mínimo (o el máximo) de una cierta función definida en \mathbb{R}^n o en un subconjunto de \mathbb{R}^n . Este tipo de problema aparece cuando queremos optimizar una cantidad. Por ejemplo: ¿qué precio deben tener las entradas de un concierto para conseguir la mayor recaudación posible? De la misma forma, en muchas situaciones es útil encontrar una curva o una superficie que sea “la mejor posible” en algún sentido. ¿Qué trayectoria debe seguir un avión que vuela entre dos puntos para gastar la menor cantidad posible de combustible? En esta pregunta hay varias cosas a tener en cuenta: una, que si los dos puntos están muy alejados la curvatura de la Tierra no permite ir en línea recta. Por otra parte, es mejor volar más alto porque el aire es menos denso y se gasta menos combustible; sin embargo, *subir* gasta combustible. ¿Cuál es la altura óptima a la que debemos volar? ¿Cuál es la mejor forma de subir hasta esa altura? ¿Una subida fuerte al principio, tal vez una subida suave?

Se le da el nombre de *cálculo de variaciones* a un conjunto de técnicas para encontrar extremos de funciones (a las que llamamos *funcionales*) definidas sobre espacios de dimensión infinita (normalmente espacios de funciones como $\mathcal{C}^2[a, b]$ por ejemplo). En este sentido es una generalización de las técnicas que se usan para encontrar extremos de funciones en \mathbb{R}^d , con la diferencia de que la incógnita, el objeto óptimo que buscamos, no es un punto de \mathbb{R}^d sino una función definida en cierto conjunto. Podemos verlo como una generalización a dimensión infinita de las ideas que conocemos para encontrar extremos de funcionales en dimensión finita. Por supuesto, los problemas que aparecen son en general mucho más complejos.

Por otra parte el cálculo de variaciones es una herramienta fundamental en la formulación de las teorías actuales de la física, y da un nuevo punto de vista en varios problemas.

1.1. Definiciones básicas

Lo primero que debemos hacer es fijar el lenguaje que usamos para hablar de los extremos de una función:

Definición 1.1 (Extremos). Sea \mathcal{D} un conjunto cualquiera, y $\mathcal{F}: \mathcal{D} \rightarrow \mathbb{R}$ una función.

- Decimos que \mathcal{F} alcanza un *mínimo* en $y_* \in \mathcal{D}$, o que y_* es un *mínimo* de \mathcal{F} , o que y_* es un *minimizador* de \mathcal{F} , cuando

$$\mathcal{F}(y) \geq \mathcal{F}(y_*) \quad \text{para todo } y \in \mathcal{D}.$$

Decimos además que es un *mínimo estricto* cuando

$$\mathcal{F}(y) > \mathcal{F}(y_*) \quad \text{para todo } y \in \mathcal{D} \text{ con } y \neq y_*.$$

- Decimos que \mathcal{F} alcanza un *máximo* en $y_* \in \mathcal{D}$, o que y_* es un *máximo* de \mathcal{F} , o que y_* es un *maximizador* de \mathcal{F} , cuando

$$\mathcal{F}(y) \leq \mathcal{F}(y_*) \quad \text{para todo } y \in \mathcal{D}.$$

Decimos además que es un *máximo estricto* cuando

$$\mathcal{F}(y) < \mathcal{F}(y_*) \quad \text{para todo } y \in \mathcal{D} \text{ con } y \neq y_*.$$

Decimos que $y_* \in \mathcal{D}$ es un *extremo* de \mathcal{F} si es un máximo o un mínimo.

Si $y_* \in \mathcal{D}$ es un mínimo de \mathcal{F} , a veces decimos que $\mathcal{F}(y_*)$ es el *valor mínimo* de \mathcal{F} , y análogamente con el valor máximo. Los extremos que hemos definido son extremos *globales*, en el sentido de que la desigualdad se cumple en todo el dominio de \mathcal{F} . Si el espacio considerado está dotado de una norma (o al menos una topología) podemos definir los *extremos locales*, considerando, en lugar de todo el subconjunto \mathcal{D} , los y en un entorno de y_* .

La búsqueda de extremos de un funcional no varía si hacemos transformaciones afines del funcional:

Lema 1.2. Sea \mathcal{D} un conjunto cualquiera, y $\mathcal{F}: \mathcal{D} \rightarrow \mathbb{R}$ una función.

1. Sean $A, B \in \mathbb{R}$ con $A > 0$. Un punto $y_* \in \mathcal{D}$ minimiza a \mathcal{F} si y sólo si minimiza a $A\mathcal{F} + B$.
2. Un punto $y_* \in \mathcal{D}$ minimiza a \mathcal{F} si y sólo si maximiza a $-\mathcal{F}$.

El último punto nos dice que podemos restringir el estudio a la búsqueda de mínimos, ya que para buscar máximos no tenemos más que buscar mínimos de $-\mathcal{F}$. Haremos esto de forma sistemática, sin volver a mencionar que el caso de los máximos es completamente análogo.

1.2. El caso de \mathbb{R}^d

Comenzaremos recordando la teoría de extremos en \mathbb{R}^d , como motivación para entender las técnicas de cálculo de variaciones.

Sabemos que no todas las funciones tienen extremos. Por ejemplo, la función $f(x) = x$ definida en $\Omega = \mathbb{R}$ no tiene extremos globales ni locales. También sabemos que los valores máximos y mínimos se pueden alcanzar en más de un punto del dominio, por ejemplo la función $f(x) = x^2$ definida en $\Omega = [-1, 1]$ alcanza el máximo absoluto 1 en -1 y 1.

Para garantizar la existencia de extremos globales, recordamos el Teorema de Weierstrass.

Teorema 1.3 (Teorema de Weierstrass). *Si $\mathcal{D} \subset \mathbb{R}^n$ es compacto y $f : \mathcal{D} \rightarrow \mathbb{R}$ es continua, entonces f alcanza un valor máximo y un valor mínimo en \mathcal{D} .*

Es claro que ambas hipótesis (\mathcal{D} compacto y f continua) son necesarias (¿por qué?). También es claro que si f alcanza un máximo en un punto $x_* \in \mathcal{D}$, entonces $-f$ alcanza un mínimo en ese mismo punto $x_* \in \mathcal{D}$. Podemos, por tanto, reducirnos a dar criterios para encontrar mínimos, ya que de forma análoga encontraríamos los máximos.

Uno de los resultados más conocidos puede enunciarse así:

Teorema 1.4 (Condición necesaria de extremo en \mathbb{R}^d). *Sean $\Omega \subseteq \mathbb{R}^d$ un abierto, $f : \Omega \rightarrow \mathbb{R}$ una función de clase \mathcal{C}^1 . Si la función f alcanza un extremo en un punto $y_* \in \Omega$ entonces*

$$\nabla f(x_*) = 0. \quad (1)$$

Decimos que f tiene un *punto crítico o estacionario* en y_* si $\nabla f(y_*) = 0$. Este resultado tiene varios inconvenientes:

1. No nos da una condición suficiente para encontrar un extremo, ya que por ejemplo las funciones $f(x) = x^3$ en $\Omega = \mathbb{R}$ y $f(x, y) = x^2 - y^2$ en $\Omega = \mathbb{R}^2$ no tienen extremos en sus puntos críticos.
2. Requiere que f sea una función \mathcal{C}^1 , definida en un abierto. Por ejemplo, este resultado no nos sirve para encontrar extremos que estén en la frontera del dominio de definición, cuando este dominio no es abierto. No sirve tampoco para funciones definidas en una superficie.

2. Condiciones necesarias de extremo

2.1. Variación de una función. Condiciones necesarias de extremo

Queremos ahora encontrar un enunciado similar al teorema 1.4 en el caso de espacios de dimensión infinita. La dificultad principal que nos encontramos es que el concepto de gradiente depende de la norma que elijamos, y en un espacio de dimensión infinita no son todas equivalentes. Lo que es más importante, pedir que un funcional sea \mathcal{C}^1 es normalmente demasiado restrictivo. Para evitar esto usaremos las derivadas direccionales, definidas de la siguiente forma:

Definición 2.1 (Variación de Gâteaux). Sea X un espacio vectorial, $\mathcal{D} \subseteq X$ un subconjunto cualquiera, $\mathcal{F}: \mathcal{D} \rightarrow \mathbb{R}$ un funcional. Para $y \in \mathcal{D}$, $v \in X$ decimos que existe la derivada direccional $\delta_v \mathcal{F}(y)$ si

1. existe $\delta > 0$ tal que $y + \epsilon v \in \mathcal{D}$ para todo $\epsilon \in (-\delta, \delta)$,
2. y la función $\epsilon \mapsto \mathcal{F}(y + \epsilon v)$ es derivable en $\epsilon = 0$.

En ese caso definimos

$$\delta_v \mathcal{F}(y) := \left. \frac{d}{d\epsilon} \right|_{\epsilon=0} \mathcal{F}(y + \epsilon v) = \lim_{\epsilon \rightarrow 0} \frac{\mathcal{F}(y + \epsilon v) - \mathcal{F}(y)}{\epsilon}. \quad (2)$$

A $\delta_v \mathcal{F}(y)$ lo llamaremos *variación de Gâteaux* o *derivada direccional* de \mathcal{F} en y en la dirección de v .

Algunas de las propiedades básicas de esta derivada son las siguientes:

Lema 2.2. Sean $a, b \in \mathbb{R}$, X un espacio vectorial, \mathcal{F} y $\tilde{\mathcal{F}}$ funcionales definidos en un subconjunto $\mathcal{D} \subseteq X$, $y \in \mathcal{D}$, $v \in X$. Siempre que existan las variaciones que aparecen en las siguientes expresiones, se satisfacen las siguientes igualdades:

1. Para $v = 0$, $\delta_0 \mathcal{F}(y)$ siempre existe, y vale 0.
2. $\delta_v(a\mathcal{F} + b\tilde{\mathcal{F}})(y) = a\delta_v \mathcal{F}(y) + b\delta_v \tilde{\mathcal{F}}(y)$.
3. $\delta_{av} \mathcal{F}(y) = a\delta_v \mathcal{F}(y)$. Además, para $a \neq 0$, $\delta_{av} \mathcal{F}(y)$ está definida si y sólo si lo está $\delta_v \mathcal{F}(y)$.
4. $\delta_{-v} \mathcal{F}(y) = -\delta_v \mathcal{F}(y) = \delta_v(-\mathcal{F})(y)$. Los miembros de esta igualdad están definidos si y solo si alguno de ellos lo está.
5. Si $\mathcal{F} \in C^1(\mathbb{R}^d)$ e $y, v \in \mathbb{R}^d$, entonces $\delta_v \mathcal{F}(y) = D_v \mathcal{F}(y)$ (la derivada direccional usual). Y por tanto

$$\delta_v \mathcal{F}(y) = \nabla f(y) \cdot v.$$

Definida la variación de Gâteaux podemos dar una condición suficiente que deben cumplir los extremos de cualquier funcional:

Proposición 2.3 (Condición necesaria de extremo). Sea X un espacio vectorial, $\mathcal{D} \subseteq X$ un subconjunto, y $\mathcal{F}: \mathcal{D} \rightarrow \mathbb{R}$ una función. Si \mathcal{F} alcanza un extremo en $y_* \in \mathcal{D}$ entonces

$$\delta_v \mathcal{F}(y_*) = 0$$

para toda dirección $v \in X$ en la que exista la derivada direccional.

Esta condición es análoga a la del teorema 1.4. De la misma forma, la condición de que la variación de Gâteaux sea cero no implica que y_* sea extremo, ni determina, caso de serlo, de qué tipo sería. Es por tanto una condición necesaria, pero no suficiente.

Definición 2.4. Sea X un espacio vectorial, $\mathcal{D} \subseteq X$ un subconjunto cualquiera, $\mathcal{F}: \mathcal{D} \rightarrow \mathbb{R}$ un funcional. Llamamos *punto crítico* o *extremal* de \mathcal{F} a un punto $y_* \in \mathcal{D}$ tal que

$$\delta_v \mathcal{F}(y_*) = 0$$

para todo $v \in X$ para el que exista la derivada direccional.

Podemos por tanto enunciar la proposición 2.3 diciendo que los extremos de un funcional son siempre puntos críticos. Por supuesto, esto no contiene información para puntos en los que no existen las derivadas direccionales.

Ejemplo 2.5. Consideramos el operador que da la longitud de la gráfica de una función $y: [a, b] \rightarrow \mathbb{R}$:

$$\mathcal{F}(y) := \int_a^b \sqrt{1 + (y'(x))^2} dx,$$

definida en el dominio

$$\mathcal{D} := \{y \in \mathcal{C}^1[a, b]\}.$$

Podemos calcular la derivada direccional en la dirección de $v \in \mathcal{D}$ de la siguiente forma:

$$\begin{aligned} \delta_v \mathcal{F}(y) &= \left. \frac{d}{d\epsilon} \right|_{\epsilon=0} \mathcal{F}(y + \epsilon v) = \left. \frac{d}{d\epsilon} \right|_{\epsilon=0} \int_a^b \sqrt{1 + (y'(x) + \epsilon v'(x))^2} dx \\ &= \int_a^b \frac{y'(x)v'(x)}{\sqrt{1 + (y'(x))^2}} dx. \end{aligned}$$

Observa que esto es siempre 0 cuando y es constante.

Ejemplo 2.6. Consideramos el mismo operador que da la longitud de la gráfica de una función $y: [0, 1] \rightarrow \mathbb{R}$:

$$\mathcal{F}(y) := \int_0^1 \sqrt{1 + (y'(x))^2} dx,$$

definido ahora en el dominio

$$\mathcal{D} := \{y \in \mathcal{C}^1[0, 1] \mid y(0) = 0, y(1) = 1\}.$$

Ahora las únicas direcciones en las que podemos calcular la derivada direccional de \mathcal{F} son las funciones $v \in \mathcal{C}_c^1(0, 1)$. El resultado es el mismo que antes:

$$\delta_v \mathcal{F}(y) = \int_0^1 \frac{y'(x)v'(x)}{\sqrt{1 + (y'(x))^2}} dx.$$

Observa que esto es 0 para todo $y \in \mathcal{D}$ cuando y' es constante. ¿Puedes demostrar rigurosamente que la función $y(x) = x$ es el único mínimo de \mathcal{F} en \mathcal{D} ?

2.2. Ecuaciones de Euler-Lagrange

2.2.1. Para una variable

En muchas aplicaciones los funcionales de los que queremos buscar extremos son de la forma

$$\mathcal{F}(y) = \int_a^b F(x, y(x), y'(x)) dx, \quad (3)$$

definidos en algún subconjunto de $\mathcal{C}^k[a, b]$. Para estos funcionales la condición de la Proposición 2.3 se puede escribir de forma más útil. Esta forma, que vamos a ver en un momento, se conoce como *ecuaciones de Euler-Lagrange*.

Teorema 2.7 (Ecuaciones de Euler-Lagrange en una variable). Sea $U \subseteq \mathbb{R} \times \mathbb{R}^d \times \mathbb{R}^d$ un abierto, $F: U \rightarrow \mathbb{R}$ una función de clase \mathcal{C}^2 . Dados $a < b \in \mathbb{R}$, consideramos el funcional

$$\mathcal{F}(y) := \int_a^b F(t, y(t), y'(t)) dt$$

definido un cierto dominio \mathcal{D} tal que

$$\mathcal{D} \subseteq \{y \in \mathcal{C}^2([a, b], \mathbb{R}^d) \mid (t, y(t), y'(t)) \in U \forall t \in [a, b]\}.$$

Si existe $y \in \mathcal{D}$ tal que $\delta_v \mathcal{F}(y)$ está definida para todo $v \in \mathcal{C}_c^\infty((a, b); \mathbb{R}^d)$, y tenemos

$$\delta_v \mathcal{F}(y) = 0 \quad \text{para todo } v \in \mathcal{C}_c^\infty((a, b); \mathbb{R}^d)$$

entonces

$$\frac{d}{dt} \nabla_z F(t, y(t), y'(t)) = \nabla_y F(t, y(t), y'(t)) \quad \text{para todo } t \in (a, b). \quad (4)$$

En particular, si y es un extremo de \mathcal{F} , entonces y es solución de la ecuación (4), llamada ecuación de Euler-Lagrange.

Demostración. Sólo tenemos que escribir la derivada direccional $\delta_v \mathcal{F}(y)$:

$$\begin{aligned} \delta_v \mathcal{F}(y) &= \frac{d}{d\epsilon} \mathcal{F}(y + \epsilon v) = \frac{d}{d\epsilon} \int_a^b F(t, y(t) + \epsilon v(t), y'(t) + \epsilon v'(t)) dt \\ &= \int_a^b \nabla_y F(t, y(t), y'(t)) \cdot v(t) dt + \int_a^b \nabla_z F(t, y(t), y'(t)) \cdot v'(t) dt \\ &= \int_a^b \nabla_y F(t, y(t), y'(t)) \cdot v(t) dt - \int_a^b \frac{d}{dt} (\nabla_z F(t, y(t), y'(t))) \cdot v(t) dt \\ &= \int_a^b \left(\nabla_y F(t, y(t), y'(t)) - \frac{d}{dt} (\nabla_z F(t, y(t), y'(t))) \right) \cdot v(t) dt. \end{aligned}$$

Del siguiente lema deducimos directamente que se cumplen las ecuaciones de Euler-Lagrange. \square

Lema 2.8 (Lema fundamental del cálculo de variaciones). Sea $I = (a, b) \subseteq \mathbb{R}$ un intervalo (posiblemente con $a = -\infty$ o $b = +\infty$), $f: I \rightarrow \mathbb{R}$ una función continua. Si ocurre que

$$\int_a^b f(x) \phi(x) dx = 0 \quad \text{para toda } \phi \in \mathcal{C}_c^\infty(I)$$

entonces $f(x) = 0$ para todo $x \in I$.

Demostración. Si f no es cero en un cierto $x \in I$, usando la continuidad de f podemos encontrar $r > 0$ tal que

$$f(y) \geq \frac{f(x)}{2}, \quad y \in B_r(x).$$

Si tomamos entonces una función $\phi \in \mathcal{C}_c^\infty(I)$ que sea no negativa, positiva al menos en un punto, y con soporte en $B_r(x)$, vemos que

$$\int_I f(y) \phi(y) dy = \int_{B_r(x)} f(y) \phi(y) dy \geq \frac{f(x)}{2} \int_{B_r(x)} \phi(y) dy > 0,$$

en contradicción con las hipótesis. \square

Observación 2.9. Como consecuencia de lo anterior, el resultado también es cierto para funciones vectoriales: si $f : I \rightarrow \mathbb{R}^d$ es una función continua tal que

$$\int_a^b f(x)\psi(x) \, dx = 0 \quad \text{para toda } \psi \in \mathcal{C}_c^\infty(I, \mathbb{R}^d)$$

entonces $f(x) = 0$ para todo $x \in I$. Para verlo es suficiente con tomar $\psi := \phi e_i$, con e_i el vector i -ésimo de la base usual de \mathbb{R}^d , y $\phi \in \mathcal{C}_c^\infty(I)$ una función arbitraria. El lema 2.8 demuestra entonces que la componente i -ésima de f se anula, y esto podemos hacerlo para cada coordenada.

Definición 2.10 (Funciones extremales). Las funciones $C^2[a, b]$ que satisfacen la ecuación de Euler-Lagrange (4) se llaman *funciones estacionarias*, *funciones extremales* o *puntos críticos* del funcional \mathcal{F} .

Observación 2.11. Si consideramos un dominio \mathcal{D} sin restricciones en la frontera, es decir, $\mathcal{D} = C^1([a, b])$, entonces las funciones test, v , están también en $\mathcal{D} = C^1([a, b])$ y además de la ecuación de Euler-Lagrange (4) se deben imponer las condiciones de contorno $F_z(a, y(a), y'(a)) = F_z(b, y(b), y'(b)) = 0$ para encontrar los extremales del funcional.

Observa que las condiciones de contorno que impongamos en el dominio \mathcal{D} no importan en el Teorema 2.7: la ecuación (4) se sigue cumpliendo en cualquier caso. Sin embargo, hay veces en las que si el dominio no tiene condiciones de contorno en alguno de los extremos podemos obtener más información. Enunciamos el resultado en el caso en que las condiciones son libres en el extremo inferior; por supuesto, el resultado correspondiente es cierto para condiciones libres en el extremo superior:

Proposición 2.12 (Condiciones de contorno libres). *Con las mismas hipótesis del Teorema 2.7, supongamos que*

$$\delta_v \mathcal{F}(y) = 0 \quad \text{para todo } v \in \mathcal{C}_c^\infty([a, b]; \mathbb{R}^d)$$

(Observa que ahora el intervalo es $[a, b]$; es decir, permitimos que v tome cualquier valor en el extremo izquierdo del intervalo). Entonces, además de la conclusión del Teorema 2.7 se cumple que

$$\nabla_z F(a, y(a), y'(a)) = 0.$$

Demostración. Con el mismo cálculo que en la demostración del Teorema 2.7 llegamos a que

$$0 = \int_a^b \left(\nabla_y F(t, y, y') - \frac{d}{dt} \nabla_z F(t, y, y') \right) v \, dt - \nabla_z F(a, y(a), y'(a)) v(a)$$

para todo $v \in \mathcal{C}_c^\infty([a, b]; \mathbb{R}^d)$. Sabemos por el Teorema 2.7 que la integral tiene que ser 0, luego

$$\nabla_z F(a, y(a), y'(a)) v(a) = 0$$

para todo $v \in \mathcal{C}_c^\infty([a, b]; \mathbb{R}^d)$, lo cual implica el resultado. \square

Ejemplo 2.13. Consideramos el funcional

$$\mathcal{F}(y) := \int_a^b |y'(t)|^2 dt$$

definido en el dominio

$$\mathcal{D} := \{y \in \mathcal{C}^2([a, b], \mathbb{R}^d) \mid y(a) = y_a, y(b) = y_b\},$$

donde $y_a, y_b \in \mathbb{R}^d$ son puntos dados. La función F correspondiente a este caso es $F: \mathbb{R} \times \mathbb{R}^d \times \mathbb{R}^d \rightarrow \mathbb{R}$ dada por

$$F(t, y, z) = |z|^2,$$

y sus derivadas parciales son

$$\nabla_y F = 0, \quad \nabla_z F = 2z.$$

Por tanto, la ecuación de Euler-Lagrange es en este caso

$$0 = \frac{d}{dt}(2y'(t)) = 2y''(t),$$

es decir,

$$y''(t) = 0.$$

Ejemplo 2.14. Consideramos ahora

$$\mathcal{F}(y) := \int_a^b \left(\frac{1}{2} |y'(t)|^2 - U(y(t)) \right) dt$$

con $U: \mathbb{R}^d \rightarrow \mathbb{R}$ es una función de clase \mathcal{C}^2 , definido en el dominio

$$\mathcal{D} := \{y \in \mathcal{C}^2([a, b], \mathbb{R}^d) \mid y(a) = y_a, y(b) = y_b\},$$

donde $y_a, y_b \in \mathbb{R}^d$ son puntos dados. La función F correspondiente a este caso es $F: \mathbb{R} \times \mathbb{R}^d \times \mathbb{R}^d \rightarrow \mathbb{R}$ dada por

$$F(t, y, z) = \frac{1}{2} |z|^2 - U(y),$$

y sus derivadas parciales son

$$\nabla_y F = -\nabla U(y), \quad \nabla_z F = z.$$

Por tanto, la ecuación de Euler-Lagrange es en este caso

$$-\nabla U(y(t)) = \frac{d}{dt}(y'(t)) = y''(t),$$

es decir,

$$y''(t) = -\nabla U(y(t)).$$

Observa que esto es la ecuación de Newton para una partícula que se mueve en un potencial U .

2.2.2. Para varias variables

Observación 2.15 (Funcionales que dependen de una función de varias variables). Para los funcionales del tipo $\mathcal{F}(z) = \int_{\Omega} F(x, y, z(x, y), z_x(x, y), z_y(x, y)) \, dx \, dy$, definidos en $\mathcal{D} = \{z \in C^1(\bar{\Omega}) : z(x, y) = \gamma(x, y), \forall (x, y) \in \partial\Omega\}$, (siendo $\Omega \subseteq \mathbb{R}^2$ un dominio acotado y regular) su variación de Gâteaux es cero, si la función z satisface

$$\frac{\partial}{\partial x} F_{z_x}(x, y, z, z_x, z_y) + \frac{\partial}{\partial y} F_{z_y}(x, y, z, z_x, z_y) = F_z(x, y, z, z_x, z_y) \quad \forall x \in (a, b) \, \forall i = 1, \dots, q, \quad (5)$$

con la condición de contorno $z(x, y) = \gamma(x, y)$ para todo $(x, y) \in \partial\Omega$.

2.3. Regularidad de los extremos

Teorema 2.16. *Con las hipótesis del Teorema 2.7, consideramos el mismo funcional \mathcal{F} definido ahora en un dominio \mathcal{D} con*

$$\mathcal{D} \subseteq \{y \in C^1([a, b], \mathbb{R}^d) \mid (t, y(t), y'(t)) \in U \, \forall t \in [a, b]\}.$$

(Observa que ahora sólo requerimos que y sea C^1 .) Si existe $y \in \mathcal{D}$ tal que $\delta_v \mathcal{F}(y)$ está definida para todo $v \in C_c^\infty((a, b); \mathbb{R}^d)$, y tenemos

$$\delta_v \mathcal{F}(y) = 0 \quad \text{para todo } v \in C_c^\infty((a, b); \mathbb{R}^d)$$

entonces $t \mapsto \nabla_z F(t, y(t), y'(t)) \in C^1[a, b]$, y se cumple la ecuación de Euler-Lagrange (4).

La novedad en este resultado es que obtenemos *una regularidad adicional*: sabiendo que y es un punto crítico del funcional \mathcal{F} deducimos que cierta función de y' es de clase C^1 . Muchas veces esto es suficiente para decir que y es de hecho C^2 .

La demostración se basa en dos lemas:

Lema 2.17 (du Bois-Reymond). *Si $h \in C([a, b])$ y $\int_a^b h(x)v'(x)dx = 0$ para todo $v \in C_c^\infty(a, b)$, entonces h es constante en $[a, b]$.*

Demostración. Por continuidad, es suficiente ver que h es constante en (a, b) . Supongamos que hay dos puntos $x_1, x_2 \in (a, b)$ tales que $h(x_1) > h(x_2)$. Por continuidad, podemos encontrar $\epsilon, \delta > 0$ tal que $h(x) > h(y) + \delta$ para todo $x \in (x_1 - \epsilon, x_1 + \epsilon)$, $y \in (x_2 - \epsilon, x_2 + \epsilon)$. Tomamos una función $w \in C_c^\infty(a, b)$ con soporte contenido en $(x_1 - \epsilon, x_1 + \epsilon) \cup (x_2 - \epsilon, x_2 + \epsilon)$, tal que $w \geq 0$ en $(x_1 - \epsilon, x_1 + \epsilon)$, y de forma que

$$w(x_1 + s) = -w(x_2 + s) \quad \text{para todo } s \in (-\epsilon, \epsilon),$$

$$\int_{x_1 - \epsilon}^{x_1 + \epsilon} w(s) \, ds > 0.$$

(La función w tiene dos “bultos” simétricos: uno positivo cerca de x_1 y otro negativo cerca de x_2 ; no es difícil ver que efectivamente existe). Claramente por la primera condición, $\int_a^b w(s) \, ds = 0$, y podemos definir

$$v(x) := \int_a^x w(s) \, ds, \quad x \in [a, b].$$

Vemos que $v \in \mathcal{C}_c^\infty(a, b)$, y

$$\begin{aligned} \int_a^b h(s)v'(s) \, ds &= \int_{x_1-\epsilon}^{x_1+\epsilon} h(s)w(s) \, ds + \int_{x_2-\epsilon}^{x_2+\epsilon} h(s)w(s) \, ds \\ &= \int_{-\epsilon}^{\epsilon} h(x_1+s)w(x_1+s) \, ds + \int_{-\epsilon}^{\epsilon} h(x_2+s)w(x_2+s) \, ds \\ &= \int_{-\epsilon}^{\epsilon} (h(x_1+s) - h(x_2+s))w(x_1+s) \, ds > \delta \int_{-\epsilon}^{\epsilon} w(x_1+s) \, ds > 0, \end{aligned}$$

lo cual nos da una contradicción. \square

Gracias a este resultado se puede probar este segundo lema, que permite demostrar el teorema:

Lema 2.18. Sean $g, h \in C([a, b])$ y $\int_a^b (g(x)v(x) + h(x)v'(x))dx = 0$, $\forall v \in \mathcal{D}_0 = \{v \in \mathcal{E} : v(a) = v(b) = 0\}$, entonces $h \in C^1([a, b])$ y $h' = g$.

Demostración. Sea G una primitiva cualquiera de g . Integrando por partes tenemos que

$$0 = \int_a^b (g(x)v(x) + h(x)v'(x)) \, dx = \int_a^b (h(x) - G(x))v'(x) \, dx.$$

Por el Lema 2.17 vemos que $h(x) = G(x) + C$ para cierta constante C , luego se cumple el resultado. \square

Demostración del Teorema 2.16. Es una consecuencia directa del Lema 2.18, ya que la variación de Gâteaux del funcional (3) es

$$\delta_v \mathcal{F}(y) = \int_a^b (F_y(x, y(x), y'(x))v(x) + F_z(x, y(x), y'(x))v'(x)) \, dx, \quad (6)$$

para todo $v \in \mathcal{D}_0 = \{v \in X \mid v(a) = v(b) = 0\}$. \square

Los Lemas 2.17 y 2.18 tienen también una versión que funciona en casi todo punto, con condiciones un poco más generales:

Lema 2.19 (du Bois-Reymond, versión L^1). Si $h \in L^1[a, b]$ y $\int_a^b h(x)v'(x)dx = 0$ para todo $v \in \mathcal{C}_c^\infty(a, b)$, entonces h es constante en casi todo punto de $[a, b]$.

Lema 2.20. Sean $g, h \in L^1[a, b]$ tales que $\int_a^b (g(x)v(x) + h(x)v'(x))dx = 0$ para todo $v \in \mathcal{C}_c^\infty(a, b)$. Entonces h coincide en casi todo punto con una función \tilde{h} absolutamente continua y $\tilde{h}' = g$ en casi todo punto de $[a, b]$.

Demostración. Sea G una primitiva cualquiera de g . Igual que antes, integrando por partes tenemos que

$$0 = \int_a^b (g(x)v(x) + h(x)v'(x)) \, dx = \int_a^b (h(x) - G(x))v'(x) \, dx.$$

Por el Lema 2.19 vemos que $h(x) = G(x) + C$ para cierta constante C y casi todo $x \in [a, b]$, luego se cumple el resultado. \square

De la misma forma que antes obtenemos el siguiente teorema, con condiciones un poco más generales que las del Teorema 2.16 (y con una conclusión también un poco más débil):

Teorema 2.21. *Con las hipótesis del Teorema 2.7, consideramos el mismo funcional \mathcal{F} definido ahora en un dominio \mathcal{D} con*

$$\mathcal{D} \subseteq \{y \in \text{AC}([a, b], \mathbb{R}^d) \mid (t, y(t), y'(t)) \in U \forall t \in [a, b]\}.$$

(Observa que ahora sólo requerimos que y sea absolutamente continua.) Si existe $y \in \mathcal{D}$ tal que $\delta_v \mathcal{F}(y)$ está definida para todo $v \in \mathcal{C}_c^\infty((a, b); \mathbb{R}^d)$, y tenemos

$$\delta_v \mathcal{F}(y) = 0 \quad \text{para todo } v \in \mathcal{C}_c^\infty((a, b); \mathbb{R}^d)$$

entonces $t \mapsto \nabla_z F(t, y(t), y'(t))$ es absolutamente continua en $[a, b]$, y se cumple la ecuación de Euler-Lagrange (4) en casi todo punto de $[a, b]$.

3. Condiciones suficientes de extremo: convexidad

Para encontrar condiciones suficientes trabajaremos con funcionales convexos.

Definición 3.1 (Funcional convexo). Sea X un espacio vectorial y $\mathcal{D} \subseteq X$ un subconjunto cualquiera. Un funcional $\mathcal{F}: \mathcal{D} \rightarrow \mathbb{R}$ es *convexo* en \mathcal{D} si

$$\theta \mathcal{F}(y) + (1 - \theta) \mathcal{F}(z) \geq \mathcal{F}(\theta y + (1 - \theta)z)$$

para todo $y, z \in \mathcal{D}$ y para todo $\theta \in (0, 1)$ tal que $\theta y + (1 - \theta)z \in \mathcal{D}$. Decimos que es *estrictamente convexo* si además la desigualdad es siempre estricta.

Lema 3.2. *Sea X un espacio vectorial, $\mathcal{D} \subseteq X$ un subconjunto cualquiera, y sea $\mathcal{F}: \mathcal{D} \rightarrow \mathbb{R}$ un funcional convexo. Entonces*

$$\mathcal{F}(y + v) - \mathcal{F}(y) \geq \delta_v \mathcal{F}(y) \tag{7}$$

para todo $y \in \mathcal{D}$ y todo $v \in X$ tal que $y + v \in \mathcal{D}$ y exista $\delta_v \mathcal{F}(y)$. Si además \mathcal{F} es estrictamente convexo, la desigualdad (7) es estricta siempre que $v \neq 0$.

Demostración. Para $\epsilon > 0$,

$$y + \epsilon v = (1 - \epsilon)y + \epsilon(y + v),$$

luego para $\epsilon > 0$ suficientemente pequeño, $y + \epsilon v \in \mathcal{D}$ y tenemos

$$\mathcal{F}(y + \epsilon v) \leq (1 - \epsilon)\mathcal{F}(y) + \epsilon\mathcal{F}(y + v)$$

o equivalentemente,

$$\frac{\mathcal{F}(y + \epsilon v) - \mathcal{F}(y)}{\epsilon} \leq \mathcal{F}(y + v) - \mathcal{F}(y), \tag{8}$$

lo cual da el resultado cuando $\epsilon \rightarrow 0$. Si además \mathcal{F} es estrictamente convexa y $v \neq 0$,

$$\frac{\mathcal{F}(y + \epsilon v) - \mathcal{F}(y)}{\epsilon} < \mathcal{F}(y + v) - \mathcal{F}(y).$$

Observa que no podemos pasar directamente aquí al límite, porque las desigualdades estrictas se pierden en el límite. Pero sí podemos fijar $\epsilon > 0$ para el cual lo anterior es cierto; y entonces para cualquier $0 < \delta < 1$, usando de nuevo la desigualdad (8) para $y, \epsilon v$,

$$\frac{\mathcal{F}(y + \delta \epsilon v) - \mathcal{F}(y)}{\delta \epsilon} \leq \frac{\mathcal{F}(y + \epsilon v) - \mathcal{F}(y)}{\epsilon} < \mathcal{F}(y + v) - \mathcal{F}(y).$$

Pasando al límite $\delta \rightarrow 0$ obtenemos la desigualdad estricta. \square

Con la definición 3.1 se puede probar el siguiente resultado:

Teorema 3.3 (Condición suficiente de extremo). *Sea X un espacio vectorial y $\mathcal{D} \subseteq X$ un subconjunto cualquiera. Si $\mathcal{F}: \mathcal{D} \rightarrow \mathbb{R}$ es un funcional convexo e $y_* \in \mathcal{D}$ es tal que*

$$\delta_v \mathcal{F}(y_*) = 0 \quad \text{para todo } v \in X \text{ tal que } y_* + v \in \mathcal{D}$$

(es decir: la derivada direccional existe y vale 0), entonces \mathcal{F} alcanza un mínimo en y_ . Si además \mathcal{F} es estrictamente convexo, y_* es el único punto donde se alcanza el mínimo.*

Demostración. Dado cualquier $y \in \mathcal{D}$ llamamos $v := y - y_*$. Como \mathcal{F} es convexo,

$$\mathcal{F}(y) - \mathcal{F}(y_*) = \mathcal{F}(y_* + v) - \mathcal{F}(y_*) \geq \delta_v \mathcal{F}(y_*) = 0,$$

y por tanto \mathcal{F} alcanza un mínimo en y_* . Si \mathcal{F} es estrictamente convexo entonces la desigualdad anterior es estricta siempre que $y \neq y_*$, y el mínimo sólo puede alcanzarse en y_* . \square

La desigualdad (7) nos da también una forma de ver si una función es convexa: si se cumple siempre que $y, y + v$ estén en \mathcal{D} entonces \mathcal{F} tiene que ser convexo:

Lema 3.4. *Sea X un espacio vectorial, $\mathcal{D} \subseteq X$ un subconjunto cualquiera, $\mathcal{F}: \mathcal{D} \rightarrow \mathbb{R}$ un funcional. Supongamos que para todo $y \in \mathcal{D}$ y todo $v \in X$ tal que $y + v \in \mathcal{D}$ se cumple que existe $\delta_v \mathcal{F}(y)$, y la desigualdad (7) es cierta. Entonces \mathcal{F} es convexa. Si se cumple la desigualdad (7) estrictamente siempre que $v \neq 0$ entonces \mathcal{F} es estrictamente convexo.*

Para demostrar este lema necesitamos un resultado previo, que es esencialmente el caso de dimensión 1:

Lema 3.5. *Sea $D \subseteq \mathbb{R}$ un abierto, $f: D \rightarrow \mathbb{R}$ derivable tal que*

$$f(x + h) - f(x) \geq hf'(x) \tag{9}$$

para todo $x, h \in \mathbb{R}$ tales que $x, x + h \in D$. Entonces f es convexa en U . Si la desigualdad es estricta para todo $h \neq 0$ entonces f es estrictamente convexa.

Demostración. Si tomamos $y, z \in D$, tenemos que demostrar que

$$\theta f(y) + (1 - \theta)f(z) \geq f(\theta y + (1 - \theta)z). \tag{10}$$

Llamando $w := \theta y + (1 - \theta)z$ y usando (9) tenemos

$$\begin{aligned} \theta(z - y)f'(w) &\leq f(z) - f(w) \\ -(1 - \theta)(z - y)f'(w) &\leq f(y) - f(w), \end{aligned}$$

Multiplicando la primera desigualdad por $(1 - \theta)$, la segunda por θ , y sumando, obtenemos (10); si las desigualdades son estrictas obtenemos la desigualdad estricta en (10). \square

Observación 3.6. Es interesante destacar que en el lema anterior no es importante que sea f' la cantidad que aparece a la derecha de (9). Dicho de otra forma, el siguiente resultado es cierto, con la misma demostración: si $U \subseteq \mathbb{R}$ es un abierto, $f, \phi: U \rightarrow \mathbb{R}$ funciones cualesquiera tal que $f(x+h) - f(x) \geq h\phi(x)$ para todo $x, h \in \mathbb{R}$ tales que $x, x+h \in U$, entonces f es convexa en U . Esto permite aplicar resultados de este tipo usando por ejemplo la derivada superior (definida igual que la derivada, pero con \limsup), que siempre existe. Por supuesto, si f es derivable en x y tenemos $f(x+h) - f(x) \geq h\phi(x)$ para todo $h \in \mathbb{R}$ suficientemente pequeño, entonces necesariamente $\phi(x) = f'(x)$.

Demostración del Lema 3.4. Sean $y, z \in \mathcal{D}$, y sea $\theta \in (0, 1)$ tal que $\theta y + (1 - \theta)z \in \mathcal{D}$. Tenemos que demostrar que

$$\mathcal{F}((1 - \theta)y + \theta z) \leq (1 - \theta)\mathcal{F}(y) + \theta\mathcal{F}(z). \quad (11)$$

Si $y = z$ entonces esta desigualdad es trivial, así que suponemos que $y \neq z$. Definimos $D = \{\lambda \in \mathbb{R} \mid (1 - \lambda)y + \lambda z \in \mathcal{D}\}$. Vamos a ver que D es un abierto.

Está claro que $0, 1 \in D$. Si $\lambda \neq 0 \in D$, entonces llamemos $w := (1 - \lambda)y + \lambda z \in \mathcal{D}$. Como $w \in \mathcal{D}$, $y \in \mathcal{D}$, por hipótesis sabemos que $\delta_{y-w}\mathcal{F}(w)$ está bien definida. En particular (como $y - w \neq 0$) tenemos que $w + \epsilon(y - w) \in \mathcal{D}$ para $\epsilon \in \mathbb{R}$ con $|\epsilon|$ suficientemente pequeño; es decir,

$$(1 - \lambda(1 - \epsilon))y + \lambda(1 - \epsilon)z \in \mathcal{D}$$

para ϵ suficientemente pequeño. Es decir: hay un entorno de λ que está contenido en D . Podemos hacer el mismo razonamiento para $\lambda = 0$ observando que $y \in \mathcal{D}$, $z \in \mathcal{D}$ y por tanto $\delta_{y-z}\mathcal{F}(y)$ está bien definida, y por tanto hay un entorno de $\lambda = 0$ que está contenido en D . La conclusión es que D es un abierto de \mathbb{R} .

Ahora definimos la función

$$f: D \rightarrow \mathbb{R}, \quad f(\lambda) = \mathcal{F}((1 - \lambda)y + \lambda z).$$

Siguiendo el mismo razonamiento de arriba vemos que f es derivable, y la desigualdad (7) es exactamente la condición (9) para la función f . Por el Lema 3.5, f es convexa, luego

$$f(\lambda) \leq (1 - \lambda)f(0) + \lambda f(1)$$

para todo $\lambda \in D$. En particular, para $\lambda = \theta$ obtenemos (11). De la misma forma, si la desigualdad (7) se cumple estrictamente obtenemos la desigualdad estricta en (11). \square

Ejercicio 3.1. Sea \mathcal{D} el dominio

$$\mathcal{D} := \{y \in \mathcal{C}^2[0, 1] \mid y(0) = 0, y(1) = e^{\sqrt{2}} - e^{-\sqrt{2}}\}.$$

Consideramos el siguiente funcional \mathcal{F} , definido en \mathcal{D} :

$$\mathcal{F}(y) := \int_0^1 e^{2x} ((y'(x))^2 + (y(x))^2) dx.$$

Demuestra que el funcional \mathcal{F} alcanza un mínimo global en \mathcal{D} , y encuentra todas las funciones $y \in \mathcal{D}$ donde lo alcanza.

Solución. El funcional \mathcal{F} es convexo, como puede comprobarse directamente viendo que la función $F(x, y, z) := e^{2x}(z^2 + y^2)$ es convexa en las variables (y, z) . Por tanto, alcanza un mínimo en una función $y \in \mathcal{D}$ si y sólo si y es solución de la ecuación de Euler-Lagrange. Dicha ecuación es

$$\frac{d}{dx}(e^{2x}y'(x)) = e^{2x}y(x),$$

es decir,

$$y''(x) + 2y'(x) - y(x) = 0.$$

Esta es una ecuación ordinaria de coeficientes constantes, cuya solución general es

$$y(x) = Ae^{(-1+\sqrt{2})x} + Be^{(-1-\sqrt{2})x}.$$

Imponiendo las condiciones de contorno,

$$A + B = 0, \quad Ae^{(-1+\sqrt{2})} + Be^{(-1-\sqrt{2})} = e^{\sqrt{2}} - e^{-\sqrt{2}},$$

lo cual implica $A = e, B = -e$. El único mínimo global se alcanza por tanto en la función

$$y(x) = e^{-x-1} \left(e^{\sqrt{2}x} - e^{-\sqrt{2}x} \right) = 2e^{-x-1} \sinh(\sqrt{2}x).$$

3.1. El caso de \mathbb{R}^d

Por otro lado, si se alcanzan extremos en los puntos críticos estos serán locales y no hay en principio garantías de que sean globales. Entonces, una pregunta natural es: ¿Qué criterios podemos dar para tener garantías de que en los puntos críticos se alcanzan extremos globales?

Decimos que una función $f : \Omega \rightarrow \mathbb{R}$ es *convexa* en Ω si tiene derivadas parciales continuas en Ω y verifica: $f(x) \geq f(x_*) + \nabla f(x_*) \cdot (x - x_*) \forall x, x_* \in \Omega$. Esta condición se puede escribir de forma análoga como:

$$f(x + v) \geq f(x) + \nabla f(x) \cdot v, \quad \forall x \in \Omega \quad \text{y} \quad \forall v \in \mathbb{R}^n \text{ tal que } x + v \in \Omega.$$

¿Es la función $f(x, y) = x^2 + y^2$ convexa?

Conocidas las funciones convexas, es fácil comprobar el siguiente resultado.

Teorema 3.7. *Si $f : \Omega \rightarrow \mathbb{R}$ es una función convexa en Ω , entonces f tiene un valor mínimo en cada punto crítico de Ω .*

Decimos que una función $f : \Omega \rightarrow \mathbb{R}$ es *estrictamente convexa* en Ω si tiene derivadas parciales continuas en Ω y verifica: $f(x) \geq f(x_*) + \nabla f(x_*) \cdot (x - x_*) \forall x, x_* \in \Omega$, y la igualdad se da sólo si $x = x_*$.

Teorema 3.8. *Si $f : \Omega \rightarrow \mathbb{R}$ es una función estrictamente convexa en Ω , entonces f como máximo tiene un punto crítico y como máximo un mínimo en el interior de Ω .*

Otra forma de determinar los mínimos de f pasa por estudiar su matriz Hessiana.

Teorema 3.9. Sea $f : \Omega \rightarrow \mathbb{R}$ con derivadas parciales de segundo orden continuas. En un punto crítico $x_* \in \Omega$, f alcanza un mínimo local estricto si $\sum_{i,j=1}^n \partial_{x_i x_j} f(x_*) u_i u_j > 0$ $\forall u \in \mathbb{R}^n$ unitario. Es decir, la matriz Hessiana es definida positiva.

Análogamente, si la matriz Hessiana es definida negativa la función alcanza en ese punto crítico un máximo local. Si la matriz Hessiana es indefinida en x_* la función presenta un punto de silla.

Sabemos por teorema 1.3 que si Ω es compacto la función (continua) tiene máximo y mínimo globales. El teorema 3.9 ofrece condiciones para encontrar extremos locales en el interior de Ω , para determinar los extremos globales se tiene que analizar también lo que ocurre en su frontera. El método de los multiplicadores de Lagrange es muy útil para ese estudio.

4. Extremos con restricciones

4.1. Extremos con restricciones en \mathbb{R}^d

Queremos estudiar los extremos de una función $f : \Omega \rightarrow \mathbb{R}$, con $\Omega \subseteq \mathbb{R}^n$ un abierto, con las restricciones adicionales

$$\phi_i(x) = 0, \quad i = 1, \dots, m,$$

donde $\phi_i : \Omega \rightarrow \mathbb{R}$ son $m < n$ funciones reales dadas. Dicho de otra forma, queremos encontrar los extremos de f en el conjunto

$$S := \{x \in \Omega \mid \phi_i(x) = 0, i = 1, \dots, m\}. \quad (12)$$

Para que las restricciones $\phi_i(x) = 0$ sean “independientes” cerca de un punto $x^0 \in \mathbb{R}^n$ hace falta que determinen los valores de m coordenadas de x , fijadas las otras $n - m$. Sin pérdida de generalidad, suponemos que podemos fijar las últimas $n - m$ coordenadas y determinar las primeras m usando las condiciones $\phi_i(x) = 0, i = 1, \dots, m$, al menos cuando x está cerca de x^0 . La formalización de esto es la siguiente condición:

$$\det(\partial_{x_j} \phi_i(x^0))_{i,j=1,\dots,m} \neq 0. \quad (13)$$

El Teorema de la Función Implícita nos permite entonces encontrar un abierto $U \subseteq \mathbb{R}^{n-m}$ con $y^0 := (x_{m+1}^0, \dots, x_n^0) \in U$ y m funciones $\psi_1, \dots, \psi_m : U \rightarrow \mathbb{R}$ tales que

$$\phi_i(\psi(y), y) = 0 \quad \text{para todo } y \in U, \quad (14)$$

donde

$$y \equiv (y_1, \dots, y_{n-m}), \quad \psi(y) \equiv (\psi_1(y), \dots, \psi_m(y)),$$

y con $\psi(y^0) = (x_1^0, \dots, x_m^0)$. Es decir: en un entorno de x^0 , las m restricciones nos permiten “despejar” $n - m$ variables que quedan libres, y escribir las m variables restantes en función de ellas.

Teorema 4.1. Sea $\Omega \subseteq \mathbb{R}^n$ abierto, $1 \leq m < n$ un entero, $f, \phi_1, \dots, \phi_m : \Omega \rightarrow \mathbb{R}$ funciones reales de clase \mathcal{C}^1 . Definimos S por (12). Si un punto $x^0 \in S$ es un extremo de f en S , y las funciones ϕ_1, \dots, ϕ_m satisfacen (13), entonces hay m números $\lambda_1, \dots, \lambda_m \in \mathbb{R}$ tales que

$$\nabla f(x^0) + \sum_{i=1}^m \lambda_i \nabla \phi_i(x^0) = 0. \quad (15)$$

Demostración. Supongamos que $x^0 \in S$ es un extremo (digamos un mínimo) de f . En particular, usando la función ψ definida como antes, escribimos

$$\tilde{\psi}: U \subseteq \mathbb{R}^{n-m} \rightarrow \mathbb{R}^n, \quad \tilde{\psi}(y) := (\psi(y), y)$$

y entonces

$$F(y) := f(\tilde{\psi}(y)) \geq f(\tilde{\psi}(y^0)) = f(x^0) \quad \text{para todo } y \in U,$$

donde llamamos $y^0 := (x_{m+1}^0, \dots, x_n^0)$. Por el Teorema 1.4, el gradiente de F con respecto a y es 0 en y^0 . Es decir,

$$\nabla f(\tilde{\psi}(y^0)) \cdot \partial_{y_i} \tilde{\psi}(y^0) = 0, \quad i = 1, \dots, n-m.$$

Observa que $\partial_{y_i} \tilde{\psi}(y^0)$ es un vector en \mathbb{R}^n , ya que $\tilde{\psi}$ es una función con n componentes. Esto quiere decir que $\nabla f(x^0)$ es perpendicular a los $n-m$ vectores $\partial_{y_i} \tilde{\psi}(y^0)$, $i = 1, \dots, n-m$. Ocurre además que los vectores

$$\{\nabla \phi_1(x^0), \dots, \nabla \phi_m(x^0), \partial_{y_1} \tilde{\psi}(y^0), \dots, \partial_{y_{n-m}} \tilde{\psi}(y^0)\}$$

forman una base de \mathbb{R}^n . El motivo es que:

1. Los vectores $\partial_{y_1} \tilde{\psi}(y^0), \dots, \partial_{y_{n-m}} \tilde{\psi}(y^0)$ son independientes, ya que $\partial_{y_i} \tilde{\psi}(y^0) = (\partial_{y_i} \psi(y^0), e_i)$, donde e_i es el vector i de la base usual de \mathbb{R}^{n-m} .
2. Los vectores $\nabla \phi_1(x^0), \dots, \nabla \phi_m(x^0)$ son independientes gracias a (13).
3. Para cualesquiera $i = 1, \dots, m$, $j = 1, \dots, n-m$, los vectores $\nabla \phi_i(x^0)$, $\partial_{y_j} \tilde{\psi}(y^0)$ son perpendiculares (lo cual se deduce de (14), ya que $\partial_{y_j} \phi_i(\tilde{\psi}(y)) = 0$).

Deducimos que $\nabla f(x^0)$ tiene que ser combinación lineal de los m vectores $\nabla \phi_i(x^0)$, $i = 1, \dots, m$. \square

Ejercicio 4.1. La condición de transversalidad para las funciones ϕ_i es importante en el Teorema 4.1. Comprueba que el mínimo de la función

$$f(x) = x_1 + x_3^2$$

en el conjunto

$$D := \{x = (x_1, x_2, x_3) \in \mathbb{R}^3 \mid x_2 - x_1^2 = 0, x_2 - x_1^3 = 0\}$$

existe y se alcanza en un sólo punto, pero no se cumple la condición (15).

4.2. Extremos con restricciones en dimensión infinita

En esta sección vamos a ver un resultado análogo al Teorema 4.1 sobre optimización con restricciones en dimensión finita. Las dificultades que aparecen al trasladarlo al caso de dimensión infinita son las siguientes:

1. En el caso de dimensión finita necesitábamos que las funciones que aparecen sea \mathcal{C}^1 . Para espacios generales esto es muy restrictivo, y de hecho el concepto de derivada depende de la norma (u otro tipo de estructura) que elijamos. Necesitamos rebajar los requisitos si queremos poder aplicar el resultado en casos útiles.

2. La condición (13) sobre la “transversalidad” de las restricciones ϕ_j estaba escrita eligiendo las primeras m coordenadas, ya que si es cierta con otras coordenadas siempre podemos reordenarlas. Ahora tenemos que traducir la misma condición a un espacio donde no tenemos un sistema de coordenadas “estándar”.

Teniendo en cuenta estas modificaciones, el siguiente resultado no es más que una versión del Teorema 4.1 en dimensión infinita:

Teorema 4.2. *Sea X un espacio vectorial, $\mathcal{D} \subseteq X$ un subconjunto, $\mathcal{F}, \Phi_1, \dots, \Phi_m: \mathcal{D} \rightarrow \mathbb{R}$ funciones. Supongamos que \mathcal{F} alcanza un extremo en $y_0 \in \mathcal{D}_L$ dentro del conjunto*

$$\mathcal{D}_L := \{y \in \mathcal{D} \mid \Phi_j(y) = 0, \ j = 1, \dots, m\}.$$

Sea $\mathcal{V} \subseteq X$ tal que $\mathcal{F}, \Phi_1, \dots, \Phi_m$ cumplen la siguiente condición de regularidad: para toda lista de $m+1$ vectores $v_1, \dots, v_{m+1} \in \mathcal{V}$ las funciones

$$\tilde{\mathcal{F}} := \mathcal{F} \circ \Psi, \quad \tilde{\Phi}_j := \Phi_j \circ \Psi \quad (j = 1, \dots, m+1)$$

están bien definidas y son de clase \mathcal{C}^1 en un entorno de $x = 0$, donde

$$\Psi: \mathbb{R}^{m+1} \rightarrow X, \quad \Psi(x) := y_0 + \sum_{i=1}^{m+1} x_i v_i. \quad (16)$$

Entonces ocurre (al menos) una de las dos opciones siguientes:

1. *Para cualesquiera $v_1, \dots, v_m \in \mathcal{V}$ se cumple que*

$$\det \left(\delta_{v_i} \Phi_j(y_0) \right)_{i,j=1,\dots,m} = 0.$$

2. *Existen constantes $\lambda_1, \dots, \lambda_m \in \mathbb{R}$ tales que*

$$\delta_v \mathcal{F}(y_0) = \sum_{j=1}^m \lambda_j \delta_v \Phi_j(y_0)$$

para todo $v \in \mathcal{V}$.

Observación 4.3. La condición de regularidad de \mathcal{F} y las funciones Φ_j es la más general que permite completar la demostración tal como está escrita, y no requiere ninguna estructura adicional de X salvo la de espacio vectorial. Si queremos una condición más familiar, es suficiente que X sea un espacio normado, $\mathcal{D} \subseteq X$ sea un abierto (en la topología de X), que las funciones $\mathcal{F}, \Phi_1, \dots, \Phi_m$ sean continuas en X , y que las derivadas $\delta_v \mathcal{F}, \delta_v \Phi_1, \dots, \delta_v \Phi_m$ estén definidas y sean continuas en un entorno de y_0 para cualquier $v \in \mathcal{V}$.

Demostración del Teorema 4.2. Si no se cumple la primera opción, tomemos vectores v_1, \dots, v_m tales que la matriz $M := (\delta_{v_i} \Phi_j(y_0))_{i,j}$ es invertible. Dado cualquier $v_{m+1} \in \mathcal{V}$ sabemos por hipótesis que hay un entorno abierto U de 0 en \mathbb{R}^{m+1} tal que la función

$$\tilde{\mathcal{F}}: U \rightarrow \mathbb{R}, \quad \tilde{\mathcal{F}}(x) := \mathcal{F}(\Psi(x))$$

está bien definida y es \mathcal{C}^1 , donde Ψ es la función (16) asociada a v_1, \dots, v_{m+1} . La función $\tilde{\mathcal{F}}$ alcanza un extremo en $x = 0$ dentro del conjunto

$$\tilde{\mathcal{D}}_L := \{x \in U \mid \tilde{\Phi}_j(x) = 0, j = 1, \dots, m\},$$

con $\tilde{\Phi}_j(x) := \Phi_j(\Psi(x))$ para $x \in U$, $j = 1, \dots, m$. Gracias a la condición de regularidad que estamos suponiendo, podemos aplicar el Teorema 4.1 sobre los multiplicadores de Lagrange en dimensión finita y deducir que existen $\lambda_1, \dots, \lambda_m$ tales que

$$\nabla \tilde{\mathcal{F}}(0) = \sum_{j=1}^m \lambda_j \nabla \tilde{\Phi}_j(0).$$

Es decir,

$$\delta_{v_i} \mathcal{F}(y_0) = \sum_{j=1}^m \lambda_j \delta_{v_i} \Phi_j(y_0), \quad i = 1, \dots, m+1. \quad (17)$$

En particular, para $i = m+1$,

$$\delta_{v_{m+1}} \mathcal{F}(y_0) = \sum_{j=1}^m \lambda_j \delta_{v_{m+1}} \Phi_j(y_0), \quad i = 1, \dots, m+1.$$

Sólo falta demostrar que $\lambda_1, \dots, \lambda_m$ son independientes de v_{m+1} ; esto es así porque (17) implica (para $i = 1, \dots, m$) que

$$(\delta_{v_i} \mathcal{F}(y_0))_{i=1, \dots, m} = M\lambda,$$

donde λ es el vector $\lambda := (\lambda_1, \dots, \lambda_m)$. Como la matriz M es invertible por hipótesis, el valor de $\lambda_1, \dots, \lambda_m$ está determinado independientemente de v_{m+1} . \square

4.3. Ecuaciones de Euler-Lagrange

Para acabar esta sección, recordamos que en los ejemplos introductorios había funcionales en los que se buscaba el máximo o el mínimo en un cierto dominio \mathcal{D} , como los vistos hasta el momento, pero añadiendo además una restricción. En el caso de las geodésicas añadíamos la restricción de que la curva estuviese en una determinada superficie y en el caso de los problemas isoperimétricos la condición adicional era que una determinada integral valiese una cantidad fija. Si recordamos el repaso hecho para \mathbb{R}^n , observamos que esta situación era la que se daba en el caso de optimización con restricciones y la resolvíamos por medio de los multiplicadores de Lagrange. Eso mismo haremos para el caso funcional.

Consideramos un funcional

$$\mathcal{F}(y_1, \dots, y_q) = \int_a^b F(x, y_1(x), \dots, y_q(x), y_1'(x), \dots, y_q'(x)) dx, \quad (18)$$

definidos en $\mathcal{D} = \{(y_1, \dots, y_q) : y_i \in C^1([a, b]), y_i(a) = y_i^0, y_i(b) = y_i^1, i = 1, \dots, q\}$ y añadimos una condición adicional, que vamos a clasificar en dos tipos.

4.3.1. Restricciones del tipo $\varphi(x, y_1(x), \dots, y_q(x)) = 0$.

En este caso restringimos el dominio \mathcal{D} a un espacio más pequeño :

$\mathcal{D}_L = \{(y_1, \dots, y_q) \in \mathcal{D} : \varphi_i(x, y_1(x), \dots, y_q(x)) = 0 \ \forall x \in [a, b], \ i = 1, \dots, m\}$ con $m < q$.

Suponiendo la regularidad que se necesite para las funciones φ_i se puede probar el siguiente teorema

Teorema 4.4. Sean $F, \phi_1, \dots, \phi_m: \mathbb{R}^d \rightarrow \mathbb{R}$ funciones de clase \mathcal{C}^2 . Consideramos el funcional

$$\mathcal{F}(y) := \int_a^b F(t, y(t), y'(t)) dt$$

en el dominio

$$\mathcal{D} := \left\{ y \in \mathcal{C}^2([a, b]) \mid \begin{array}{l} y(a) = y_a, \ y(b) = y_b, \\ \phi_j(y(t)) = 0 \text{ para } j = 1, \dots, m, \ t \in [a, b] \end{array} \right\}$$

Si $y_0 \in \mathcal{D}$ es un extremo de \mathcal{F} y suponemos además que el rango de la matriz

$$\left(\partial_{y_i} \phi_j(y_0(t)) \right)_{\substack{i=1, \dots, d \\ j=1, \dots, m}}$$

es igual a m para todo $t \in [a, b]$, entonces existen funciones $\lambda_j = \lambda_j(t)$ (llamadas multiplicadores de Lagrange), con $j = 1, \dots, m$, tales que y_0 cumple las siguientes ecuaciones de Euler-Lagrange:

$$\nabla_y F^* = \frac{d}{dt} \nabla_z F^* = 0, \quad (19)$$

donde $F^* = F + \sum_{j=1}^m \lambda_j(t) \phi_j$, y denotamos sus variables como $F^* = F^*(t, y, z)$.

Como ejemplo se puede estudiar el funcional $F(y, z) = \int_a^b \sqrt{1 + y'^2(x) + z'^2(x)} dx$ sujeto a la restricción $\varphi(x, y(x), z(x)) = 0$, que describe las líneas geodésicas, en una superficie determinada por la ecuación $\varphi(x, y(x), z(x)) = 0$.

Demostración del Teorema 4.4. (Esta demostración no ha sido hecha en clase. Consideramos sólo el caso $\varphi_i(y(x)) = 0$). Denotamos $S := \{y \in \mathbb{R}^q \mid \varphi_i(y) = 0, \ i = 1, \dots, m\}$. Supongamos que $y^0: [a, b] \rightarrow \mathcal{D}_L$ es un extremo de \mathcal{F} . Tomamos $t_0 \in (a, b)$. Como $y^0(t_0) \in S$, por el teorema de la función implícita (como el Teorema 4.1) podemos “despejar $q - m$ variables”: podemos encontrar una función $\Phi: U \subseteq \mathbb{R}^{q-m} \rightarrow \mathbb{R}^q$ tal que $0 \in U$, $\Phi(0) = y^0(t_0)$, y con $\Phi(z) \in S$ para todo $z \in U$. Sabemos también que existe $\epsilon > 0$ tal que $y^0(t) \in \Phi(U)$ para todo $t \in [t_0 - \epsilon, t_0 + \epsilon] := I$. Podemos encontrar una curva $z^0: I \rightarrow U$ tal que $\Phi(z^0(t)) = y^0(t)$ para $t \in I$. Dado que la imagen de Φ está siempre en S , la curva z^0 es un extremo del funcional

$$\mathcal{G}(z) := \int_I F\left(\Phi(z(t)), \frac{d}{dt} \Phi(z(t))\right) dt = \int_I F(\Phi(z(t)), D_z \Phi(z(t)) z'(t)) dt$$

en el conjunto

$$\mathcal{D}^\epsilon := \{z \in \mathcal{C}^2(I) \mid z(t_0 - \epsilon) = z^0(t_0 - \epsilon), \ z(t_0 + \epsilon) = z^0(t_0 + \epsilon)\}.$$

Podemos entonces escribir las ecuaciones de Euler-Lagrange para el funcional \mathcal{G} y obtener

$$\frac{d}{dt} \partial_{q_i} G(z(t), z'(t)) = \partial_{z_i} G(z(t), z'(t)), \quad i = 1, \dots, q - m,$$

donde

$$G(z, q) := F(\Phi(z), D_z \Phi(z)q), \quad z \in U, \quad q \in \mathbb{R}^{q-m}.$$

Calculando las derivadas de G obtenemos

$$\partial_{z_i} G = \nabla_y F \partial_{z_i} \Phi + \nabla_p F \partial_{z_i} (D_z \Phi(z)) p, \quad \partial_{q_i} G = \nabla_p F \partial_{z_i} \Phi,$$

así que

$$\frac{d}{dt} (\nabla_p F \partial_{z_i} \Phi) = \nabla_y F \partial_{z_i} \Phi + \nabla_p F \partial_{z_i} (D_z \Phi(z)) z'(t), \quad i = 1, \dots, q - m. \quad (20)$$

Expandiendo el primer término,

$$\frac{d}{dt} (\nabla_p F \partial_{z_i} \Phi) = \frac{d}{dt} (\nabla_p F) \partial_{z_i} \Phi + \nabla_p F \partial_{z_i} (D_z \Phi(z)) z'(t).$$

Esto nos permite cancelar un término en (20):

$$\frac{d}{dt} (\nabla_p F) \partial_{z_i} \Phi = \nabla_y F \partial_{z_i} \Phi \quad i = 1, \dots, q - m. \quad (21)$$

Con el mismo razonamiento que en el caso de los multiplicadores de Lagrange en dimensión finita (Teorema 4.1) deducimos que

$$\frac{d}{dt} \left(\nabla_p F(y(t), y'(t)) \right) - \nabla_y F(y(t), y'(t)) = \sum_{i=1}^m \lambda_i(t) \nabla \phi_i(y(t)), \quad t \in I.$$

lo cual demuestra el resultado. \square

4.3.2. Restricciones del tipo isoperimétrico (es decir, de tipo integral)

En este caso consideramos restricciones del tipo

$$\Phi_i(y) := \int_a^b F_i(x, y(x), y'(x)) dx = 0, \quad i = 1, \dots, m.$$

y restringimos el dominio \mathcal{D} a un espacio más pequeño:

$$\mathcal{D}_L = \{(y_1, \dots, y_q) \in \mathcal{D} \mid \Phi_i(y) = L_i, i = 1, \dots, m.\}$$

Teorema 4.5. Sean $F, \varphi_1, \dots, \varphi_m: \mathbb{R}^d \rightarrow \mathbb{R}$ funciones de clase \mathcal{C}^2 . Definimos

$$\Phi_j(y) := \int_a^b \varphi_j(y(t)) dt, \quad j = 1, \dots, m$$

para cualquier $y \in \mathcal{C}^2([a, b])$. Consideramos el funcional

$$\mathcal{F}(y) := \int_a^b F(t, y(t), y'(t)) dt$$

en el dominio

$$\mathcal{D} := \left\{ y \in \mathcal{C}^2([a, b]) \mid \begin{array}{l} y(a) = y_a, \quad y(b) = y_b, \\ \Phi_j(y) = L_j \text{ para } j = 1, \dots, m \end{array} \right\}$$

Si $y_0 \in \mathcal{D}$ es un extremo de \mathcal{F} , entonces se cumple (al menos) una de las dos posibilidades siguientes:

1. Para todas las listas de m funciones $v_1, \dots, v_m \in \mathcal{C}_0^2([a, b])$ ocurre que

$$\det \left(\delta_{v_k} \Phi_j(y_0) \right)_{\substack{k=1, \dots, m \\ j=1, \dots, m}} = 0$$

2. Existen números λ_j (llamados multiplicadores de Lagrange), con $j = 1, \dots, m$, tales que y_0 cumple las siguientes ecuaciones de Euler-Lagrange:

$$\nabla_y F^* = \frac{d}{dt} \nabla_z F^* = 0, \quad (22)$$

donde $F^* = F + \sum_{j=1}^m \lambda_j \varphi_j$, y denotamos sus variables como $F^* = F^*(t, y, z)$.

Observamos que en este caso los multiplicadores de Lagrange son constantes, es decir, no dependen de x como en el caso anterior.

Como ejemplo se puede estudiar el funcional $A(x, y) = \frac{1}{2} \int_0^l (x(t)y'(t) - y(t)x'(t)) dt$ sujeto a la restricción $\int_0^l \sqrt{x'^2(t) + y'^2(t)} dt = L$, con $x, y \in C^1([0, l])$ y tales que $x(0) = x(l)$ e $y(0) = y(l)$. El máximo de A , en ese espacio y con esa restricción, se alcanza en la curva cerrada que encierra más área con una longitud fija L .

Ejemplos de este tipo también son los funcionales de la forma: $\mathcal{F}(y) = \int_a^b (p(x)y'^2(x) - q(x)y^2(x)) dx$, sujetos a la restricción $\int_a^b r(x)y^2(x) dx = 1$, donde $p, q, r \in C^1([a, b])$, p y r son funciones positivas en todo $[a, b]$. Si buscamos sus extremos en $\mathcal{D} = \{y \in C^1([a, b]), y(a) = y^0, y(b) = y^1\}$, observamos que las ecuaciones de Euler-Lagrange que obtenemos son

$$(p(x)y'(x))' + q(x)y(x) + \lambda r(x)y(x) = 0, \quad (23)$$

donde $\lambda \in \mathbb{R}$ es el multiplicador de Lagrange asociado al problema. Por tanto observamos que si \mathcal{F} alcanza en y un extremo local, y debe estar en $\mathcal{D} \cap C^2([a, b])$ puesto que debe verificar la ecuación (23), junto con las condiciones de contorno $y(a) = y^0, y(b) = y^1$. La ecuación (23) junto con las condiciones de contorno $y(a) = y^0, y(b) = y^1$, es lo que se conoce como *problema de Sturm-Liouville* y será objeto de estudio más adelante.

4.3.3. Multiplicadores de Lagrange para restricciones integrales: demostración

Demostración del Teorema 4.5. Podemos aplicar directamente el Teorema 4.2 con

$$\mathcal{V} := \mathcal{C}_0^2([a, b], \mathbb{R}^d)$$

y obtener que existen $\lambda_1, \dots, \lambda_m \in \mathbb{R}$ tales que

$$\delta_v \mathcal{F}^*(y_0) = 0 \quad \text{para todo } v \in \mathcal{V},$$

donde $\mathcal{F}^*: \mathcal{D} \rightarrow \mathbb{R}$ está definida por

$$\mathcal{F}^*(y) := \mathcal{F}(y) + \sum_{j=1}^m \lambda_j \Psi_j(y) = \int_a^b F^*(t, y(t), y'(t)) dt, \quad y \in \mathcal{D}.$$

El Teorema 2.7 aplicado a \mathcal{F}^* nos dice que F^* debe satisfacer las ecuaciones de Euler-Lagrange, con lo que tenemos el resultado. \square

4.4. Convexidad y restricciones

Teorema 4.6 (Condición suficiente de extremo con restricciones). *Sea X un espacio vectorial y $\mathcal{D} \subseteq X$ un subconjunto cualquiera, $\mathcal{F}, \phi_1, \dots, \phi_m: \mathcal{D} \rightarrow \mathbb{R}$ funcionales. Sean $\lambda_1, \dots, \lambda_m \in \mathbb{R}$ tales que el funcional*

$$\mathcal{F}^* := \mathcal{F} + \sum_{j=1}^m \lambda_j \phi_j,$$

definido en \mathcal{D} , es convexo. Supongamos que $y_ \in \mathcal{D}$ es tal que*

$$\delta_v \mathcal{F}^*(y_*) = 0 \quad \text{para todo } v \in X \text{ tal que } y_* + v \in \mathcal{D}$$

(es decir: la derivada direccional existe y vale 0), y tal que

$$\phi_1(y_*) = \dots = \phi_m(y_*) = 0.$$

Entonces y_ es un minimizador de \mathcal{F} en el dominio*

$$\mathcal{D}^* := \{y \in \mathcal{D} \mid \phi_1(y) = \dots = \phi_m(y) = 0\}$$

Si además \mathcal{F}^ es estrictamente convexo, y_* es el único punto de \mathcal{D}^* donde \mathcal{F} alcanza su mínimo.*

Demostración. Por el Teorema 3.3, y_* es un minimizador de \mathcal{F}^* en \mathcal{D} , y en particular es un minimizador de \mathcal{F}^* en \mathcal{D}^* . Como \mathcal{F} coincide con \mathcal{F}^* en \mathcal{D}^* , vemos que y_* es también un minimizador de \mathcal{F} en \mathcal{D}^* . \square

5. Algunos problemas variacionales

5.1. El problema isoperimétrico

Veamos qué podemos decir de la siguiente pregunta: *¿cuál es el área máxima que podemos encerrar con una cuerda de longitud dada?* Esta pregunta se conoce como el problema isoperimétrico, porque suponemos que la longitud de la cuerda, el perímetro de la figura resultante, está fija.

Para decir algo matemáticamente sobre la pregunta tenemos que formularla de forma un poco más precisa. Una *curva* en el plano es una aplicación continua $y: [a, b] \rightarrow \mathbb{R}^2$,

donde $a < b \in \mathbb{R}$. Decimos que una curva es *cerrada* si $y(a) = y(b)$, y decimos que es *simple* si no pasa dos veces por el mismo punto, posiblemente al principio y al final: es decir, si $y(t_1) \neq y(t_2)$ para todo $t_1 \neq t_2 \in (a, b)$. De esta forma, fijando una longitud $L > 0$, podemos dar una formulación más precisa de nuestro problema: *entre todas las curvas cerradas simples de longitud L , ¿hay alguna que encierre la mayor área posible?*

Para que tenga sentido definir la longitud de la curva necesitamos que sea *rectificable*. Por ahora vamos a suponer todavía más: vamos a trabajar sólo con curvas de clase \mathcal{C}^2 , que en particular son rectificables, y que tienen siempre una longitud finita definida por

$$\mathcal{L}(y) = \int_a^b |y'(t)| dt.$$

Si la curva y está recorrida en el sentido de las agujas del reloj, no es difícil ver que el área que encierra una curva cerrada simple es

$$\mathcal{F}(y) := \int_a^b y_2(t)y_1'(t) dt. \quad (24)$$

En realidad hay algo que estamos dando por supuesto, y es el hecho de que cualquier curva cerrada simple encierra una región dada (y no, por ejemplo, dos regiones, o ninguna). Esto se conoce como Teorema de la curva de Jordan, y su demostración rigurosa para cualquier curva continua no es fácil (es algo más fácil si queremos saberlo sólo para curvas de clase \mathcal{C}^2 , pero sigue estando muy lejos de ser algo obvio). Aquí vamos a aceptar este teorema (que es cierto) para seguir con nuestro razonamiento.

Hemos llegado a la siguiente formulación de nuestra pregunta: consideramos el funcional \mathcal{F} dado en (24), definido en el dominio

$$\mathcal{D} := \{y \in \mathcal{C}^2([a, b]; \mathbb{R}^2) \mid y(a) = y(b), \mathcal{L}(y) = L, y \text{ es simple}\}.$$

La pregunta es: *¿Cuál es el máximo (si lo hay) del funcional \mathcal{F} en \mathcal{D} ?*

Una observación básica es que si el problema tiene una solución, entonces tiene muchas soluciones: cualquier reparametrización de la curva óptima es otra solución, y cualquier traslación en \mathbb{R}^2 de la curva es una solución. Esto nos dice que siempre podemos suponer que $y(a) = y(b) = 0$ (haciendo una traslación si es necesario) y que $[a, b] = [0, 1]$ (reparametrizando la curva si es necesario). De la misma forma, es fácil ver que si y es una curva óptima para el problema con longitud L dada, entonces $\tilde{y}(t) = \lambda y(t)$ es una curva óptima para el problema con longitud λL , para cualquier $\lambda > 0$ que elijamos. Así es suficiente resolver el problema para $L = 2\pi$, porque podemos obtener la solución para cualquier otro L a partir de ésta. Así que busquemos el máximo del funcional

$$\mathcal{F}(y) := \int_0^1 y_2(t)y_1'(t) dt. \quad (25)$$

en el dominio

$$\mathcal{D} := \{y \in \mathcal{C}^2([0, 1]; \mathbb{R}^2) \mid y(0) = y(1) = 0, \mathcal{L}(y) = 2\pi, y \text{ es simple}\}.$$

Escrito así, el Teorema 4.5 se aplica directamente. Podemos deducir que cualquier curva óptima $y = (y_1, y_2)$ debe cumplir

$$-y_2' = \lambda \frac{d}{dt} \frac{y_1'}{|y'|}, \quad y_1' = \lambda \frac{d}{dt} \frac{y_2'}{|y'|},$$

donde $\lambda \in \mathbb{R}$ es el multiplicador de Lagrange asociado a la única restricción que tenemos. Siempre podemos reparametrizar un máximo y de forma que $|y'(t)| = 2\pi$ para todo $t \in [0, 1]$, y gracias a la invarianza del problema la reparametrización sigue siendo un máximo. Suponiendo que es así,

$$-2\pi y'_2 = \lambda y''_1, \quad 2\pi y'_1 = \lambda y''_2.$$

Si llamamos $(p_1(t), p_2(t)) := (y'_1(t), y'_2(t))$,

$$-2\pi p_2 = \lambda p'_1, \quad 2\pi p_1 = \lambda p'_2.$$

Si $\lambda = 0$ la curva y tiene que ser constante, así que no es simple y no está en el dominio \mathcal{D} . Supongamos que $\lambda < 0$. Entonces las soluciones de este sistema lineal son

$$p_1(t) = A \cos(\omega t) + B \sin(\omega t), \quad p_2(t) = -A \sin(\omega t) + B \cos(\omega t),$$

donde $\omega := 2\pi/|\lambda| = -2\pi/\lambda$. Ahora usamos otra invarianza del problema: si y es una curva óptima, entonces una *rotación* de y en torno a $(0, 0)$ es también una curva óptima. Esto quiere decir que podemos suponer (por ejemplo) $y'(0) = (0, 2\pi)$; es decir,

$$p_1(0) = 0, \quad p_2(0) = 2\pi,$$

lo cual implica que $A = 0$, $B = 2\pi$:

$$y'_1(t) = p_1(t) = 2\pi \sin(\omega t), \quad y'_2(t) = p_2(t) = 2\pi \cos(\omega t).$$

Usando la condición $y_1(0) = y_2(0) = 0$,

$$y_1(t) = -\lambda \cos(\omega t) + \lambda, \quad y_2(t) = -\lambda \sin(\omega t).$$

Finalmente, como y tiene que cumplir $y(1) = 0$, vemos que $\cos(\omega) = 1$, $\sin(\omega) = 0$, luego $\omega = 2\pi n$ para cierto $n < 0$ entero. Como y es simple, tenemos $\omega = 2\pi$, luego $\lambda = -1$:

$$y_1(t) = 1 - \cos(2\pi t), \quad y_2(t) = \sin(2\pi t).$$

Ejercicio 5.1. Demuestra que si $\lambda > 0$ la solución correspondiente es también una circunferencia, pero descrita en el sentido contrario a las agujas del reloj.

La conclusión es que si hay una curva óptima para nuestro problema, necesariamente tiene que ser una circunferencia. El hecho de que la circunferencia es efectivamente la óptima no podemos demostrarlo con técnicas generales de cálculo de variaciones.

5.2. Superficies minimales de revolución

¿Cuál es la superficie de revolución con menor área posible? Para ser más exactos, queremos que la superficie tenga ciertos radios ya especificados en dos puntos dados. Un ejemplo práctico de esto es la película de jabón que se obtiene entre dos aros de metal, ya que la tensión de la película hace que adopte una forma lo más “estirada” posible, minimizando su superficie. Formalizando el problema, si suponemos que la sección de nuestra superficie viene descrita por la gráfica de una función $y = y(x)$, la pregunta es: *¿cuál es la menor área posible de la superficie de revolución que se obtiene al girar la*

gráfica de $y: [0, L] \rightarrow \mathbb{R}$ alrededor del eje x , suponiendo que $y(0) = y_0 > 0, y(L) = y_L > 0$?

Dada una curva y , el área de la superficie de revolución correspondiente es

$$\mathcal{F}(y) = \int_0^L y(x) \sqrt{1 + (y'(x))^2} dx,$$

de forma que nuestro problema es encontrar un mínimo de \mathcal{F} en el dominio

$$\mathcal{D} = \{y \in \mathcal{C}^2([0, L]) \mid y(0) = y_0, y(L) = y_L\},$$

con $y_0 > 0, y_L > 0$ dados. Sabemos por el Teorema 2.7 que si hay una curva y donde se alcance el mínimo debe cumplirse que

$$\frac{d}{dx} \left(\frac{yy'}{\sqrt{1 + (y')^2}} \right) = \sqrt{1 + (y')^2}.$$

Esta ecuación tiene un factor integrante: multiplicando todo por $yy'/\sqrt{1 + (y')^2}$ tenemos

$$\left(\frac{yy'}{\sqrt{1 + (y')^2}} \right)^2 = y^2 - C,$$

para cierta constante $C \in \mathbb{R}$. Expandiendo,

$$C(y')^2 = y^2 - C.$$

Cuando $C = 0$ la solución de esta ecuación es

$$y \equiv \text{const.},$$

que es por tanto una posible solución de la ecuación de Euler-Lagrange cuando $y_0 = y_L$. Cuando $C \neq 0$, tiene que ocurrir que $C > 0$ y que $y^2 \geq C$ (luego $y \geq \sqrt{C}$, ya que $y(0)$ es positiva). La forma de la ecuación sugiere el cambio

$$y = k \cosh u,$$

con $k > 0$ cierta constante, u cierta función. Esto nos da

$$Ck^2(\sinh u)^2(u')^2 = k^2(\cosh u)^2 - C,$$

que eligiendo $k^2 := C$ se simplifica a

$$ku' = 1,$$

es decir, $u(x) = u_0 + x/k$. Obtenemos entonces

$$y(x) = k \cosh \left(u_0 + \frac{x}{k} \right),$$

con $u_0 \in \mathbb{R}, k > 0$ constantes a elegir. Todas las soluciones de la ecuación de Euler-Lagrange tienen que ser de esta forma; lo que nos falta es ver si hay alguna de ellas (o más de una) que cumpla también las condiciones de frontera. En este caso son

$$k \cosh(u_0) = y_0, \quad k \cosh \left(u_0 + \frac{L}{k} \right) = y_L. \quad (26)$$

Dados $y_0 < y_L$ no es fácil determinar exactamente cuándo estas ecuaciones tienen solución para k, u_0 , y cuándo tienen más de una (aunque siempre podemos intentar una solución numérica). Pero cuando encontremos esas soluciones, ¿podemos asegurar que la y correspondiente es un mínimo del funcional \mathcal{F} ?

En algunos casos podemos ver que *no*: si L es grande y $y_0 = y_L$ (o si $y_0 \approx y_L$) es fácil encontrar superficies que tienen *menos área* que el cilindro a altura $y_0 \equiv y_L$. Esto nos dice que en esos casos la solución constante de la ecuación de Euler-Lagrange no es un mínimo del funcional. (En particular, esto quiere decir que el funcional \mathcal{F} no es convexo.) Cuando L es pequeño en comparación con y_0, y_L , estas superficies sí minimizan el área, pero no es trivial demostrarlo.

De nuevo, una solución completa al problema requiere un estudio más detallado del problema — no es suficiente con las técnicas generales que hemos visto.

Ejemplo 5.1. Supongamos que $y_0 = y_L$. Entonces las ecuaciones (26) implican

$$u_0 + \frac{L}{k} = -u_0, \quad \text{es decir,} \quad u_0 = -\frac{L}{2k},$$

luego

$$\cosh(-u_0) = -\frac{2y_0}{L}u_0.$$

Podemos ver gráficamente que esta ecuación no tiene soluciones cuando y_0/L es pequeño; tiene una solución para un cierto valor preciso de y_0/L ; y tiene dos soluciones cuando y_0/L es grande.

5.3. La catenaria

¿Cuál es la forma que tiene una cadena colgada de sus extremos? Este problema se puede resolver con mecánica de Newton: uno puede en principio resolver las fuerzas que actúan sobre cada eslabón y deducir de ahí la forma de la cadena, pero es un cálculo bastante complicado. Una forma de expresar la ley de Newton de forma más eficiente en este caso es decir que la cadena debe adoptar la forma con menor energía potencial posible. Si la altura de la cadena viene descrita por la gráfica de la función $y: [0, L] \rightarrow \mathbb{R}$, entonces su energía potencial es

$$\mathcal{F}(y) = \int_0^L y(x) \sqrt{1 + (y'(x))^2} dx,$$

salvo una constante multiplicativa que involucra a la constante gravitatoria y la densidad lineal de masa de la cadena. Observa que es el mismo funcional que en la Sección 5.2; sin embargo, ahora queremos minimizarlo en el dominio \mathcal{D} de la sección anterior, con la restricción adicional de que

$$\int_0^L \sqrt{1 + (y'(x))^2} dx = \ell, \tag{27}$$

donde ℓ es la longitud de la cadena. El Teorema 4.5 nos da entonces una condición necesaria que debe cumplir un minimizador y , en caso de que lo haya:

$$\frac{d}{dx} \left(\frac{(y + \lambda)y'}{\sqrt{1 + (y')^2}} \right) = \sqrt{1 + (y')^2},$$

donde $\lambda \in \mathbb{R}$ es el multiplicador de Lagrange. Esto quiere decir que $\tilde{y} := y + \lambda$ cumple exactamente la misma ecuación que la de la Sección 5.2. Con el mismo razonamiento deducimos que

$$y(x) = k \cosh\left(u_0 + \frac{x}{k}\right) - \lambda, \quad (28)$$

con $u_0 \in \mathbb{R}$, $k > 0$, $\lambda \in \mathbb{R}$ constantes a elegir.

Ejemplo 5.2. En el caso $y_0 = y_L$, es más sencillo trasladar la función al intervalo $[-L/2, L/2]$. Buscamos entonces una función y de la forma (28) que cumpla

$$y(-L/2) = y(L/2) = y_0.$$

Debe ocurrir entonces que

$$u_0 = 0,$$

y debemos hallar las dos constantes λ , k usando que $y(L/2) = y_0$ y que se cumple la restricción (27). Calculando la integral, ésta última es

$$2k \sinh\left(\frac{L}{2k}\right) = \ell,$$

es decir: $a = L/(2k)$ es solución de

$$\sinh(a) = \frac{\ell}{L}a.$$

Esta ecuación tiene una única solución positiva cuando $\ell > L$ (lo cual es razonable, ya que la longitud de la cuerda debe ser mayor que la distancia horizontal entre los dos postes). Esto nos permite encontrar k , aunque no podemos escribir su valor explícitamente. Usando esto en la condición $y(L/2) = y_0$,

$$y\left(\frac{L}{2}\right) = y_0 = k \cosh\left(\frac{L}{2C}\right) - \lambda,$$

lo cual nos da el valor de λ .

Para este problema sí se cumple que la ecuación de Euler-Lagrange tiene siempre exactamente una solución (con las condiciones de contorno dadas), y esta solución es el mínimo del funcional \mathcal{F} . Igual que antes, no es obvio demostrar que efectivamente el mínimo se alcanza a partir de los resultados que hemos visto (pero en este caso no es difícil extenderlos para conseguirlo).

5.4. La braquistócrona

Supongamos que un cierto peso, inicialmente parado, baja sin rozamiento por una curva predeterminada, recorriendo una distancia horizontal L . ¿Cuál es la forma de la curva que hace que se recorra esa distancia en el menor tiempo posible?

El tiempo que el peso tarda en recorrer una curva descrita por la gráfica de $y: [0, L] \rightarrow \mathbb{R}$ puede calcularse de la siguiente forma: si la altura de la curva en cierto punto horizontal x es $y(x)$, por conservación de la energía debe ocurrir que la velocidad $|v(x)|$ del objeto en ese punto es

$$\frac{1}{2}m|v(x)|^2 + mgy(x) = mgy(0),$$

donde g es la aceleración de la gravedad y m la masa del objeto. Si suponemos que $y(0) = 0$,

$$|v(x)|^2 = -2gy(x). \quad (29)$$

Por otra parte, el vector velocidad $v(x) = (v_1(x), v_2(x))$ debe ir en la dirección de $(1, y'(x))$, así que

$$v_2(x) = v_1(x)y'(x).$$

Usando esto en (29) tenemos

$$v_1(x)^2(1 + (y'(x))^2) = -2gy(x),$$

luego

$$v_1(x) = \frac{\sqrt{-2gy(x)}}{\sqrt{1 + (y'(x))^2}}.$$

Una vez que tenemos v_1 en función de la posición horizontal x , observa que la posición horizontal $x(t)$ del peso en tiempo t cumple $x'(t) = v_1(x(t))$. Resolviendo esta ecuación en variables separadas, el tiempo que se tarda en recorrer una distancia horizontal L es por tanto

$$\mathcal{F}(y) := \int_0^L \frac{1}{v_1(x)} dx = \int_0^L \frac{1}{\sqrt{2g}} \frac{\sqrt{1 + (y'(x))^2}}{\sqrt{-y(x)}} dx.$$

¿Podemos encontrar una función $y: [0, L] \rightarrow (-\infty, 0]$ que minimice este funcional y cumpla $y(0) = 0$, $y(L) = -H$? De nuevo por el Teorema 2.7, si existe una función así debe cumplir la ecuación de Euler-Lagrange correspondiente:

5.5. Ejercicios

Ejercicio 5.2. Dado el funcional $\mathcal{F}(y) = \int_0^{\frac{\pi}{2}} (y^2(t) - y'^2(t)) dx$ determina sus extremales en:

1. $\mathcal{D} = \{y \in C^2([0, \frac{\pi}{2}]) : y(0) = 0, y(\frac{\pi}{2}) = 1\}$.
2. $\mathcal{D} = C^2([0, \frac{\pi}{2}])$.
3. $\mathcal{D} = \{y \in C^2([0, \frac{\pi}{2}]) : y(\frac{\pi}{2}) = 1\}$.

Solución 5.3. Independientemente de las condiciones de contorno, la ecuación de Euler-Lagrange asociada (Teorema 2.7) es

$$y'' = -y,$$

y por tanto los extremales tienen que ser funciones de la forma

$$y(t) = A \cos t + B \sin t.$$

En el primer caso la única opción es $A = 0$, $B = 1$, luego

$$y(t) = \sin t \quad \text{en el primer caso.}$$

En el segundo caso tiene los dos extremos libres, y tenemos dos condiciones adicionales que obtenemos de la Proposición 2.12:

$$y'(0) = y'(\pi/2) = 0,$$

lo que implica que $A = B = 0$, luego

$$y(t) = 0.$$

En el tercer caso, de nuevo por la Proposición 2.12 tenemos que

$$y'(0) = 0,$$

luego $B = 0$ y tenemos una contradicción con la condición $y(\pi/2) = 1$. El funcional \mathcal{F} no tiene extremales en este caso.

Ejercicio 5.3. Minimiza el funcional $\mathcal{F}(y) = \int_0^1 y'^2(x) dx$ en $\mathcal{D} = \{y \in C^1([0, 1]) : y(0) = 0, y(1) = 1\}$ sujeto a la restricción $\int_0^1 y(x) dx = 1$.

Solución 5.4. Llamamos $\Phi(y) := \int_0^1 y(x) dx$, y sea $y \in \mathcal{D}$ un mínimo de \mathcal{F} en \mathcal{D} (si lo hay). Primero notamos que la condición 1 del Teorema 4.5 no se cumple: si tomamos cualquier $v \in \mathcal{C}_0^2([0, 1])$ cuya integral no sea 0,

$$\delta_v \Phi(y) = \int_0^1 v(x) dx \neq 0.$$

Por tanto, según el Teorema 4.5, si hay un minimizador debe cumplir la siguiente ecuación de Euler-Lagrange:

$$2y'' = \lambda,$$

con λ el multiplicador de Lagrange. Esto quiere decir que

$$y(x) = \frac{\lambda}{4}x^2 + C_1x + C_2,$$

con C_1, C_2 constantes reales. Usando las condiciones de frontera vemos que $C_2 = 0$, $\lambda/4 + C_1 = 1$. Usando la restricción vemos que $\lambda/12 + C_1/2 = 1$, y deducimos que $C_1 = 4$, $\lambda = -12$. Por tanto, si hay un minimizador éste tiene que ser

$$y(x) = -3x^2 + 4x, \quad x \in [0, 1].$$

Nos falta comprobar que es un minimizador. Observamos que la función

$$F^*(x, y, z) = z^2 + \lambda y$$

es convexa en (y, z) , y por tanto el funcional \mathcal{F}^* es convexo en \mathcal{D} (lo es para $\lambda = -12$, pero en este caso lo es para cualquier $\lambda \in \mathbb{R}$). Por el Teorema 4.6, la función $y = y(x)$ que hemos encontrado antes tiene que ser un minimizador (y es el único minimizador, ya que es el único punto crítico).

Ejercicio 5.4. ¿Es el funcional $\mathcal{F}(y) = \int_1^2 (xy'(x) + y(x)) dx$ convexo en $\mathcal{D} = \{y \in C^2([1, 2]) : y(1) = 1, y(2) = 2\}$? ¿Y estrictamente convexo? ¿Alcanza un mínimo este funcional? En caso afirmativo ¿en qué funciones lo alcanza?

Solución 5.5. La función \mathcal{F} es lineal, y por tanto es convexa pero no estrictamente convexa.

Vamos a ver lo mismo con una técnica que se puede emplear para funcionales en general: la función $F(x, y, z) := xz + y$ es convexa en (y, z) . Por tanto, \mathcal{F} es convexo. La

función F no es estrictamente convexa, así que esto no nos dice si \mathcal{F} es estrictamente convexo. Para $y, w \in \mathcal{D}$, $v \in \mathcal{C}_0^2([1, 2])$, $\theta \in \mathbb{R}$ podemos comprobar que

$$\mathcal{F}(\theta y + (1 - \theta)w) = \theta \mathcal{F}(y) + (1 - \theta)\mathcal{F}(w).$$

(Lo que estamos diciendo aquí es que \mathcal{F} es lineal, claro). Por tanto no es estrictamente convexo.

Si \mathcal{F} alcanza un mínimo en \mathcal{D} , el mínimo debe cumplir la ecuación de Euler-Lagrange correspondiente, que en este caso es trivial (obtenemos $1 = 1$). Esto quiere decir que su mínimo se alcanza en cualquier función de \mathcal{D} . Dicho de otra forma, ¡lo que ocurre es que \mathcal{F} es *constante* en \mathcal{D} !

Ejercicio 5.5. Consideramos el funcional $\mathcal{F}(y) = \int_a^b f(x)(1 + y'(x)^2) dx$, con $a < b \in \mathbb{R}$ y $f \in \mathcal{C}^1([a, b])$, definido en el conjunto $\mathcal{D} := \{y \in \mathcal{C}^1([a, b]) \mid y(a) = y(b) = 0\}$.

1. Determina la ecuación de Euler-Lagrange asociada a este funcional.
2. Demuestra que si $0 < a < b$ y $f(x) = x$ para $x \in [a, b]$ entonces el funcional F tiene un mínimo global. Calcúlalo.

Solución 5.6. La ecuación de Euler-Lagrange asociada es

$$2(f(x)y'(x))' = 0.$$

En el caso de la segunda pregunta, esta ecuación es

$$(xy')' = 0,$$

cuyas soluciones cumplen $xy' = C$ (constante), luego deben ser

$$y(x) = C \log x.$$

Para cumplir $y(a) = y(b) = 0$ la única opción es que $C = 0$. Por tanto $y(x) \equiv 0$ es el único candidato a mínimo, y se ve directamente en la expresión del funcional que efectivamente es el mínimo. (Alternativamente, puedes ver que el funcional es convexo).

6. Principios variacionales en mecánica

Un principio general muy útil en mecánica es el siguiente: denotemos por T la energía cinética de un sistema y U la potencial. Si $y = y(t)$ es la trayectoria de un sistema, entonces

$$T \equiv T(y)(t) = \frac{1}{2}m|y'(t)|^2, \quad U \equiv U(y)(t) = U(y(t)),$$

con U un potencial dado que suponemos que sólo depende de la posición $y(t)$. El principio es que *la trayectoria que describe el sistema es un punto crítico del funcional de acción*

$$\mathcal{F}(y) = \int_a^b (T(y)(t) - U(y)(t)) dt.$$

Esto es cierto incluso cuando el sistema tiene ciertas restricciones (por ejemplo, hay pesos unidos por una barra sólida o colgados de una cuerda) y expresamos T , U en función de las coordenadas generalizadas del sistema. Esto es muy práctico para calcular la EDO que describe la trayectoria de muchos sistemas. Vamos a ver algunos ejemplos de esto.

Ejemplo 6.1. Consideramos una partícula que se mueve bajo la acción de la gravedad, cerca de la superficie de la Tierra. Si la altura de la partícula en el tiempo t es $y(t)$ y su masa es m , la energía cinética es

$$T = \frac{1}{2}m|y'(t)|^2$$

y la potencial

$$U = mgy(t),$$

donde g es la constante gravitatoria.

Ejemplo 6.2. Consideramos la misma situación del ejemplo anterior, pero ahora queremos describir las tres coordenadas de posición de la partícula. Denotamos su posición por $y(t) = (y_1(t), y_2(t), y_3(t))$, donde $y_3(t)$ denota la altura. La energía cinética es

$$T = \frac{1}{2}m|y'(t)|^2,$$

donde ahora $|\cdot|$ denota el módulo, y la potencial

$$U = mgy_3(t).$$

Ejemplo 6.3. En general, si tenemos un sistema con coordenadas $y(t) \in \mathbb{R}^d$ que se mueve bajo la acción de un potencial U , la energía cinética es

$$T = \frac{1}{2}m|y'(t)|^2,$$

y la potencial

$$U = U(y(t)).$$

Ejemplo 6.4. Vamos a encontrar la ecuación que describe la trayectoria de un péndulo. Suponemos que tenemos un peso colgado de un punto fijo, sujeto por una barra de longitud r que puede girar en torno a ese punto. (Podemos pensar que es una cuerda, pero queremos dar la idea de que no se va a “doblar” aunque el péndulo tenga el peso arriba). Suponemos que la barra tiene un peso despreciable. El sistema está completamente descrito por el ángulo $\theta = \theta(t)$ de desviación de la posición de equilibrio. La posición horizontal y vertical del peso (si el origen de coordenadas está en el punto donde está “clavada” la barra) son

$$x(t) = r \sin \theta(t), \quad y(t) = -r \cos \theta(t).$$

Sus derivadas son

$$x'(t) = r\theta'(t) \cos \theta(t), \quad y'(t) = r\theta'(t) \sin \theta(t).$$

La energía cinética en este caso es

$$T = \frac{1}{2}m(|x'(t)|^2 + |y'(t)|^2) = \frac{1}{2}mr^2|\theta'(t)|^2$$

y la potencial es

$$U = mgy(t) = -mgr \cos \theta(t).$$

El funcional de acción es por tanto

$$\mathcal{F}(\theta) = \int_a^b F(y(t), y'(t)) dt = \int_a^b \left(\frac{1}{2}mr^2|\theta'(t)|^2 + mgr \cos \theta(t) \right) dt.$$

Si calculamos la ecuación de Euler-Lagrange para este funcional obtenemos

$$\frac{d}{dt}(mr^2\theta'(t)) = -mgr \sin \theta(t),$$

o lo que es lo mismo

$$\theta''(t) = -\frac{g}{r} \sin \theta(t).$$

Ejercicio 6.1. Calcula la aceleración con la que una rueda de radio r y masa m (uniformemente distribuida) desciende una pendiente de ángulo α , suponiendo que la rueda baja rodando sin deslizar por la pendiente. Compara esta aceleración con la que tendría si bajara deslizando por la misma pendiente, sin rozamiento.

Solución. Llamemos (x, y) a las coordenadas del centro de la rueda (poniendo el origen en la posición inicial de la rueda) y θ al ángulo de giro. La posición de la rueda está determinada en todo momento por su desplazamiento ℓ desde la posición inicial, ya que

$$x = \ell \cos \alpha, \quad y = \ell \sin \alpha.$$

En el caso de la rueda sin deslizar, la distancia recorrida ℓ está determinada por el ángulo θ :

$$\ell = \theta r.$$

En este caso, la energía cinética total es la energía cinética del centro de masas, más la energía de rotación:

$$T = \frac{1}{2}m((x')^2 + (y')^2) + \frac{1}{2}I(\theta')^2 = \frac{1}{2}m(\ell')^2 + \frac{1}{2r^2}I(\ell')^2 = \frac{3}{4}m(\ell')^2$$

donde $I = \frac{1}{2}mr^2$ es el momento de inercia de un disco uniforme. La energía potencial es

$$U = mgy = mg\ell \sin \alpha.$$

La lagrangiana es entonces

$$L = T - U = \frac{3}{4}m(\ell')^2 - mg\ell \sin \alpha,$$

y las ecuaciones de Euler-Lagrange son

$$\ell'' = \frac{2}{3}g \sin \alpha.$$

¡Observa que la aceleración no depende ni de la masa m ni del radio r ! En el caso con deslizamiento es fácil ver que la ecuación correspondiente es

$$\ell'' = g \sin \alpha,$$

que es mayor que en el caso sin deslizamiento.

7. Problemas de contorno

Un problema de contorno es una ecuación diferencial a la que imponemos condiciones que debe cumplir la solución en varios puntos. En esta sección estudiaremos problemas de contorno en ecuaciones diferenciales ordinarias. Uno de los problemas más comunes de este tipo es encontrar una solución $y : [a, b] \rightarrow \mathbb{R}$ de la ecuación

$$y'' = f(t, y, y') \quad (30)$$

en el intervalo $[a, b]$, que cumpla además

$$y(a) = y_a, \quad y(b) = y_b \quad (31)$$

para ciertos números dados y_a, y_b . Observa que estamos especificando condiciones en *dos puntos distintos* del dominio donde y está definida. Esto hace que este tipo de problemas sean distintos de los problemas de valores iniciales donde se busca una solución de (30) que cumpla

$$y(a) = y_0, \quad y'(a) = y_1 \quad (32)$$

para $y_0, y_1 \in \mathbb{R}$ dados. Hay muchos tipos de condiciones de contorno; por ejemplo, podríamos buscar soluciones que cumplan

$$y'(a) = y'(b) = 0, \quad (33)$$

o que cumplan

$$y(a) = y(b), \quad y'(a) = y'(b). \quad (34)$$

Las condiciones que involucran ecuaciones donde en cada una de ellas sólo aparece el valor de la función y sus derivadas en un sólo punto (como en (31) o (33)) se llaman *condiciones de contorno separadas*. Las condiciones (34) se llaman *condiciones de frontera periódicas*.

En general, dado que las soluciones de una ecuación de orden n forman (normalmente) una familia que depende de n parámetros, los problemas de frontera para una ecuación ordinaria de orden n suelen tener n condiciones de contorno. Esto es parecido a los problemas de valores iniciales, en los que una ecuación ordinaria de orden n requiere n condiciones en un punto para tener solución única. La situación para los problemas de contorno es más complicada que en los problemas de valores iniciales, porque los problemas de contorno no siempre tienen solución, y cuando tienen, la solución no tiene por qué ser única. Observa que para tener un problema de contorno propiamente dicho necesitamos condiciones en al menos dos puntos distintos del dominio de la solución, por lo que no suele haber problemas de contorno para ecuaciones de orden 1 (ya que el valor en un sólo punto de su dominio normalmente determina la solución completamente).

Todas estas consideraciones son ideas generales que tienen excepciones en cada caso. El propósito de esta sección es dar algunos ejemplos de problemas de contorno, explicar su motivación, y dar algunas técnicas útiles para resolverlos. Nos centraremos en problemas de contorno *lineales* de segundo orden.

Ejemplo 7.1.

$$\begin{cases} y'' + y = 0, \\ y(0) = 0, \quad y(\pi/2) = 2. \end{cases}$$

Ejemplo 7.2.

$$\begin{cases} y'' + y = 0, \\ y(0) = 0, \quad y(\pi) = 2. \end{cases}$$

Ejemplo 7.3.

$$\begin{cases} y'' + y = 0, \\ y(0) = 0, \quad y(2\pi) = 0. \end{cases}$$

Ejemplo 7.4.

$$\begin{cases} y'' + y = 0, \\ y(0) = y(2\pi), \quad y'(0) = y'(2\pi) \end{cases}$$

Ejemplo 7.5.

$$\begin{cases} y'' - y = 0, \\ y(0) = y(\pi), \quad y'(0) = y'(\pi) \end{cases}$$

7.1. Forma autoadjunta de una ecuación de segundo orden

Decimos que una ecuación lineal de segundo orden está escrita *en forma autoadjunta* cuando es del tipo

$$(p(t)y')' + q(t)y = r(t), \quad (40)$$

donde $p, q, r : [a, b] \rightarrow \mathbb{R}$ son funciones continuas, con p además de clase \mathcal{C}^1 en $[a, b]$. Decimos que la ecuación es *homogénea* cuando $r(t) = 0$ para todo $t \in [a, b]$; en otro caso decimos que es *completa*. Una ecuación de segundo orden del tipo

$$a(x)y'' + b(x)y' + c(x)y = d(x)$$

puede siempre escribirse en forma autoadjunta si a, b, c son funciones continuas con $a(x) \neq 0$ para $x \in [a, b]$.

7.2. Alternativa de Fredholm

Consideramos el problema de contorno

$$\begin{cases} (p(t)y')' + q(t)y = r(t), & (41a) \\ \alpha y(a) + \beta y'(a) = 0, & (41b) \\ \delta y(b) + \gamma y'(b) = 0, & (41c) \end{cases}$$

con condiciones de contorno separadas, o bien

$$\begin{cases} (p(t)y')' + q(t)y = r(t), & (42a) \\ y(a) = y(b), & (42b) \\ y'(a) = y'(b), & (42c) \end{cases}$$

con condiciones de contorno periódicas. En esta sección siempre suponemos que

$$\begin{aligned}
a &< b \in \mathbb{R}, \\
p, q, r &: [a, b] \rightarrow \mathbb{R} \text{ son funciones continuas,} \\
p &\text{ es una función } \mathbb{C}^1 \text{ en } [a, b], \\
p(t) &> 0 \text{ para todo } t \in [a, b]. \\
\alpha, \beta, \delta, \gamma &\in \mathbb{R}, \text{ con } \alpha, \beta \text{ no ambas nulas, } \delta, \gamma \text{ no ambas nulas.}
\end{aligned} \tag{43}$$

El problema homogéneo asociado a (41) es la misma ecuación con $r(t) \equiv 0$:

$$\begin{cases} (p(t)x')' + q(t)x = 0, & (44a) \\ \alpha x(a) + \beta x'(a) = 0, & (44b) \\ \delta x(b) + \gamma x'(b) = 0, & (44c) \end{cases}$$

De forma similar, el problema homogéneo asociado a (42) es la ecuación

$$\begin{cases} (p(t)x')' + q(t)x = 0, & (45a) \\ x(a) = x(b) & (45b) \\ x'(a) = x'(b) & (45c) \end{cases}$$

Observa que el problema homogéneo asociado (ya sea (44) o (45)) tiene siempre al menos una solución: la *solución trivial* $x(t) = 0$, $t \in [a, b]$.

A veces consideraremos la expresión

$$(p(t)x')' + q(t)x$$

como un operador diferencial que actúa sobre la función x .

Teorema 7.6 (Alternativa de Fredholm). *Consideramos el problema de contorno (41) (con condiciones de contorno separadas), asumiendo las hipótesis (43).*

1. *Si el problema homogéneo asociado (44) tiene sólo la solución trivial entonces el problema completo tiene una única solución.*
2. *Si el problema homogéneo asociado (44) tiene alguna solución distinta de la trivial, entonces el problema completo tiene solución si y sólo si*

$$\int_a^b r(t)x(t) dt = 0 \tag{46}$$

para toda solución $x : [a, b] \rightarrow \mathbb{R}$ del problema homogéneo.

Este resultado es también cierto para el problema con condiciones de contorno periódicas (42) (cuyo problema homogéneo asociado es (45)), si asumimos adicionalmente que

$$p(a) = p(b). \tag{47}$$

El teorema de la alternativa de Fredholm es en realidad mucho más general que el enunciado anterior, que es su aplicación a los problemas de contorno. El Teorema 7.6 es un caso particular de un resultado abstracto sobre operadores lineales. Podemos entender bien cómo funciona en el caso de sistemas de ecuaciones lineales en dimensión finita. Supongamos que tenemos una ecuación lineal en forma matricial

$$Ay = b, \quad (48)$$

donde A es una matriz $d \times d$, $b \in \mathbb{R}^d$ es un vector dado, y $y \in \mathbb{R}^d$ es la incógnita. Un resultado básico en álgebra lineal nos dice que:

1. Si A es invertible, entonces (48) tiene una única solución (dada por $y = A^{-1}b$).
2. Si A no es invertible, entonces (48) tiene solución si y sólo si $b \in \text{Im } A$ (la imagen de A).

Este enunciado es en realidad una versión en dimensión finita del Teorema 7.6. El punto 1, el hecho de que A sea invertible, se puede enunciar equivalentemente diciendo que el problema “homogéneo”

$$A^\top x = 0 \quad (49)$$

tiene únicamente la solución trivial $x = 0$. Por otra parte, si y es una solución del “problema completo” (48) y x es cualquier solución del “problema homogéneo” (49) entonces

$$0 = \langle A^\top x, y \rangle = \langle x, Ay \rangle = \langle x, b \rangle.$$

La afirmación recíproca es un poco más difícil de ver, pero también es cierta. Supongamos que $\langle x, b \rangle = 0$ para todo vector x que cumpla (49). Entonces,

$$b \perp \text{Ker } A^\top, \quad \text{es decir,} \quad b \in (\text{Ker } A^\top)^\perp.$$

Como $\text{Ker } A^\top = (\text{Im } A)^\perp$, tenemos

$$b \in ((\text{Im } A)^\perp)^\perp = \text{Im } A.$$

Hemos demostrado por tanto lo siguiente:

Teorema 7.7 (Alternativa de Fredholm en dimensión finita). *Sea $b \in \mathbb{R}^d$ un vector y $A \in \mathbb{R}^{d \times d}$ una matriz.*

1. *Si el problema homogéneo (49) tiene solo la solución trivial, entonces el problema completo (48) tiene una única solución.*
2. *Si el problema homogéneo (49) admite soluciones no triviales, entonces el problema completo (48) admite solución si y sólo si*

$$\langle x, b \rangle = 0$$

para toda solución x de (49).

Observa que en (49) hemos escrito la traspuesta de la matriz A , mientras que en (41a), (44a) hemos escrito el mismo operador lineal. El motivo es que el operador está escrito en forma autoadjunta, lo que hace que sea un operador simétrico (en un sentido que no hemos estudiado por ahora). Es la propiedad análoga a que la matriz A sea simétrica, de forma que $A^\top = A$.

En la demostración del Teorema 7.7 el único paso que no funciona con operadores generales es el paso donde decimos que $((\operatorname{Im} A)^\top)^\top = \operatorname{Im} A$. En un espacio de dimensión infinita esto no siempre es verdad.

Ejemplo 7.8.

$$\begin{cases} y'' - y = \cos(x), \\ y(0) = y(1) = 0. \end{cases}$$

Ejemplo 7.9.

$$\begin{cases} y'' + y = \cos(x), \\ y(0) = y(\pi) = 0. \end{cases}$$

Ejemplo 7.10.

$$\begin{cases} y'' + y = 1, \\ y(0) = y(\pi) = 0. \end{cases}$$

Ejemplo 7.11.

$$\begin{cases} y'' = f(x), \\ y'(0) = y'(1) = 0. \end{cases}$$

Ejercicio 7.1. ¿Tiene solución el siguiente problema de contorno?

$$y'' - 2y' + y = \frac{1}{1+t^2} \quad y(0) = 0, \quad y(1) = 2.$$

Si tiene solución, ¿es única?

Solución. Este problema se puede resolver explícitamente, pero para saber *si existe solución o no* necesitamos menos cálculos. Para poder aplicar la alternativa de Fredholm tenemos que transformar el problema en uno con condiciones de contorno homogéneas. Elegimos una función cualquiera que cumpla las condiciones de contorno; por ejemplo,

$$w(t) := 2t$$

y llamamos $\tilde{y} := y - w$. Está claro que $\tilde{y}(0) = \tilde{y}(1) = 0$, y además

$$\tilde{y}'' - 2\tilde{y}' + \tilde{y} = \frac{1}{1+t^2} - 4 + 2t.$$

Las soluciones del problema homogéneo correspondiente son

$$y(t) = Ae^t + Bte^t,$$

de las cuales sólo la solución trivial cumple las condiciones de contorno $y(0) = y(1) = 0$. Por tanto, el problema tiene una única solución.

Ejercicio 7.2. Demuestra que el espacio de soluciones del problema (44) (que es un espacio vectorial) tiene como mucho dimensión 1.

Ejercicio 7.3. Comprueba si las siguientes afirmaciones son correctas, y justifica brevemente tus respuestas:

1. El siguiente problema de valores en la frontera tiene una solución única en $[0, \pi]$:

$$y''(x) - y(x) = \sin x, \quad y(0) = y(\pi) = 0.$$

2. El siguiente problema de valores en la frontera tiene una solución única en $[0, \pi]$:

$$y''(x) + y(x) = \sin x, \quad y(0) = y(\pi) = 0.$$

Solución 7.12. 1. *Verdadera.* Usamos el teorema de la alternativa de Fredholm. La ecuación homogénea correspondiente es $y'' - y = 0$ con $y(0) = y(\pi) = 0$ que tiene como única solución $y \equiv 0$ (lo cual se puede comprobar directamente). Por tanto la ecuación completa tiene solución única, sea cual sea la función que ponemos a la derecha.

2. *Falsa.* De la misma forma, la ecuación homogénea correspondiente es $y'' + y = 0$ con $y(0) = y(\pi) = 0$, que tiene como soluciones $y(x) = A \sin(x)$, con $A \in \mathbb{R}$ una constante. Para $A \neq 0$, esta solución no es perpendicular a la función $\sin(x)$, luego la alternativa de Fredholm nos dice que el problema completo no tiene solución.

7.2.1. Demostración del teorema de Fredholm para condiciones de contorno separadas

Demostraremos primero el caso de condiciones de contorno separadas; el caso de condiciones de contorno periódicas se trata en la sección 7.2.2. En ambos casos usaremos la siguiente versión del teorema de integración por partes:

Lema 7.13 (Fórmula de Green para un operador de Fredholm). *Supongamos las hipótesis (43). Dada una función $u \in \mathcal{C}^2([a, b])$ definimos la función Lu como*

$$Lu = (pu')' + qu. \quad (54)$$

Entonces para cualquier función $v \in \mathcal{C}^2([a, b])$ se cumple la identidad de Lagrange

$$vLu - uLv = \frac{d}{dt}P(u, v) \quad (55)$$

donde

$$P(u, v) = p(u'v - uv').$$

Como consecuencia obtenemos la fórmula de Green

$$\int_a^b vLu - \int_a^b uLv = P(u, v) \Big|_{t=a}^{t=b}. \quad (56)$$

Demostración. La identidad (55) se comprueba directamente, y (56) se obtiene de integrar (55) entre a y b . \square

El término $P(u, v)$ que aparece en el lema anterior se llama a veces *término de frontera*. El siguiente resultado dice que se anula cuando u, v cumplen condiciones de contorno del tipo (41b)–(41c):

Lema 7.14. *Supongamos las condiciones (43). Supongamos que $u, v \in \mathcal{C}^2([a, b])$ son funciones que cumplen las condiciones de frontera (41b)–(41c); es decir, tales que*

$$\begin{cases} \alpha u(a) + \beta u'(a) = \alpha v(a) + \beta v'(a) = 0, \\ \delta u(b) + \gamma u'(b) = \delta v(b) + \gamma v'(b) = 0, \end{cases} \quad (57)$$

$$(58)$$

para ciertos $\alpha, \beta, \delta, \gamma \in \mathbb{R}$ con $|\alpha| + |\beta| > 0$, $|\delta| + |\gamma| > 0$. Entonces $P(u, v)(a) = P(u, v)(b) = 0$.

Demostración. Si llamamos

$$V_a := (\alpha, \beta), \quad V_b := (\delta, \gamma), \quad U(t) \equiv \begin{pmatrix} u(t) \\ u'(t) \end{pmatrix}, \quad V(t) \equiv \begin{pmatrix} v(t) \\ v'(t) \end{pmatrix}, \quad (59)$$

entonces las condiciones de frontera para u, v pueden escribirse como

$$V_a \cdot U(a) = V_b \cdot U(b) = V_a \cdot V(a) = V_b \cdot V(b) = 0. \quad (60)$$

Por otra parte, la expresión de $P(u, v)$ puede escribirse como

$$P(u, v) = pU \cdot J(V) \quad (61)$$

donde J es un giro de ángulo $\pi/2$:

$$J(v_1, v_2) := (v_2, -v_1) \quad \text{para } (v_1, v_2) \in \mathbb{R}^2.$$

La condición $V_a U(a) = 0$ significa que $U(a)$ es un vector perpendicular a V_a (que no es nulo por la hipótesis (43)), y lo mismo ocurre con $V(a)$. Por tanto, $U(a)$ y $V(a)$ son paralelos, así que $U(a) \cdot J(V(a)) = 0$. De (61) deducimos que $P(u, v)(a) = 0$. El razonamiento análogo funciona para $P(u, v)(b)$. \square

Corolario 7.15 (Simetría del operador diferencial). *Bajo las condiciones (43), para cualesquiera dos funciones $x, y \in \mathcal{C}^2[a, b]$ que cumplan las condiciones de contorno (41b), (41c) se cumple*

$$\int_a^b x(t) Ly(t) dt = \int_a^b (Lx(t)) y(t) dt.$$

Observación 7.16. La igualdad correspondiente es cierta (prácticamente con la misma demostración) para $x, y \in \mathcal{C}^2([a, b], \mathbb{C})$:

$$\langle x, Ly \rangle = \langle Lx, y \rangle, \quad (62)$$

donde $\langle \cdot, \cdot \rangle$ es el producto escalar definido por

$$\langle x, y \rangle := \int_a^b x(t) \overline{y(t)} dt.$$

Demostración del Teorema 7.6 para condiciones de contorno separadas. Primero traducimos el problema que queremos resolver en un lenguaje más conveniente. Para cualquier función $y \in C^1(a, b)$ escribimos

$$Ly = (py')' + qy.$$

Llamamos

$$V := \{X = (x, x') \mid Lx = 0\},$$

el conjunto de soluciones de la EDO de segundo orden $Lx = 0$, escritas como vector (x, x') . V es un espacio vectorial de dimensión 2. Sea y_p una solución particular de $Ly_p = r$, y llamamos

$$Y_p = (y_p, y'_p).$$

Sabemos que todas las soluciones de la ecuación $Ly = r$ pueden escribirse como $Y = Y_p + X$, donde $Y = (y, y')$, $X \in V$. Llamamos

$$V_a := (\alpha, \beta), \quad V_b := (\delta, \gamma),$$

y para cualquier $Y : [a, b] \rightarrow \mathbb{R}^2$ continua denotamos

$$AY := (V_a \cdot Y(a), V_b \cdot Y(b)).$$

Entonces las condiciones de contorno en (41) pueden escribirse como

$$AY = 0,$$

donde $Y = (y, y')$. Como Y es de la forma $X + Y_p$, resolver el problema de contorno (41) es encontrar $X \in V$ tal que

$$AX = -AY_p. \tag{63}$$

En particular, observamos que el operador A es un operador lineal de V a \mathbb{R}^2 (y la función Y_p no tiene por qué estar en V , claro).

Caso 1: el problema homogéneo sólo tiene la solución trivial. En el lenguaje anterior, esto quiere decir que el único $X \in V$ tal que $AX = 0$ es $X = 0$. Por tanto A , como operador lineal de V en \mathbb{R}^2 , es no singular, así que hay una única solución de (63).

Caso 2: el problema homogéneo tiene soluciones no triviales. Sea x una solución no trivial del problema de contorno homogéneo. Si el problema de contorno completo tiene una solución y , entonces usando los Lemas 7.13 y 7.14 tenemos

$$\int_a^b xr = \int_a^b xLy = \int_a^b yLx = 0.$$

Para demostrar la afirmación recíproca, supongamos que

$$\int_a^b xr = 0$$

para todo x solución del problema homogéneo. Sea x una solución no trivial; equivalentemente, sea $X \in V$, $X \neq 0$, con $AX = 0$. Como X es no trivial, sabemos que $X(a) \neq 0$, $X(b) \neq 0$. Elegimos ahora la solución de $Ly = 0$ tal que $y(0) = x(0)$, $y'(0) = x'(0)$ (es decir, $X(a) = Y(a)$). Por los Lemas 7.13 y 7.14 tenemos

$$p(a)X(a)J(Y(a)) - p(b)X(b)J(Y(b)) = \int_a^b (xLy - yLx) = \int_a^b xr = 0,$$

de forma que

$$0 = p(a)X(a)J(Y(a)) - p(b)X(b)J(Y(b)) = p(b)X(b)J(Y(b)).$$

Como $p \neq 0$, $X(b) \neq 0$, esto demuestra que $Y(b)$ es paralelo a $X(b)$. Como $V_b \cdot X(b) = 0$, deducimos que $V_b \cdot Y(b) = 0$ por tanto,

$$V_a \cdot Y(a) = 0, \quad V_b \cdot Y(b) = 0,$$

luego y es solución del problema completo. \square

7.2.2. Demostración de Teorema de Fredholm para condiciones de contorno periódicas

Lema 7.17. *Supongamos las condiciones (43). Supongamos que $u, v \in \mathcal{C}^2([a, b])$ son funciones que cumplen las condiciones de frontera (42b)–(42c); es decir, tales que*

$$\begin{cases} u(a) = u(b), & v(a) = v(b), \\ u'(a) = u'(b), & v'(a) = v'(b) \end{cases} \quad (64)$$

$$(65)$$

Supongamos también que

$$p(a) = p(b).$$

Entonces $P(u, v)(a) = P(u, v)(b) = 0$.

Demostración del Teorema 7.6 para condiciones de frontera periódicas. Reescribiendo el problema como antes, ahora definimos

$$AY := Y(a) - Y(b),$$

de forma que cumplir las condiciones de contorno es equivalente a $AY = 0$.

Caso 1: el problema homogéneo sólo tiene la solución trivial. En el lenguaje anterior, esto quiere decir que el único $X \in V$ tal que $AX = 0$ es $X = 0$. Por tanto A , como operador lineal de V en \mathbb{R}^2 , es no singular, así que hay una única solución de (63).

Caso 2 el problema homogéneo tiene soluciones no triviales. Con el mismo razonamiento que antes (y usando que $p(a) = p(b)$) vemos que si el problema completo tiene solución entonces $\int_a^b xr = 0$ para cualquier solución x del problema homogéneo.

Para demostrar la otra implicación, supongamos que

$$\int_a^b xr = 0$$

para todo x solución del problema homogéneo. Denotamos por W el espacio vectorial

$$W := \{X \in V \mid AX = 0\}.$$

(El núcleo de A en V .) Si $Y = (y, y')$ es cualquier solución de $Ly = r$, y $X = (x, x') \in W$, usando los Lemas 7.13 y 7.17 y teniendo en cuenta que $p(a) = p(b)$ tenemos

$$\begin{aligned} 0 &= \int_a^b xr = \int_a^b xLy = \int_a^b yLx + p(a)(X(b)J(Y(b)) - X(a)J(Y(a))) \\ &= p(a)X(a)(J(Y(b)) - J(Y(a))), \end{aligned}$$

es decir,

$$X(a) \cdot J(A(Y)) = 0. \quad (66)$$

Por otra parte, si $Z \in W$, con el mismo razonamiento,

$$X(a) \cdot J(A(Z)) = 0. \quad (67)$$

Distinguimos dos casos:

1. Si $W = V$ entonces $X(a)$ puede ser cualquier vector de \mathbb{R}^2 y (66) demuestra que $Y(a) = Y(b)$.
2. Si W tiene dimensión 1 entonces tomemos un generador X . La relación (67) demuestra que la imagen de A como operador lineal $A: V \rightarrow \mathbb{R}^2$ está contenida en $\langle X(a) \rangle$. Como esta imagen tiene dimensión 1, tiene que ser igual a $\langle X(a) \rangle$. Por otra parte, (66) demuestra que AY_p está también dentro de $\langle X(a) \rangle = \text{Im } A$. Por tanto, podemos resolver (63) (lo que equivale a encontrar una solución del problema completo).

□

7.3. Formulación variacional de los problemas de contorno

Si tenemos un problema de contorno en forma autoadjunta

$$(py')' + qy = r \quad (68)$$

con ciertas condiciones de contorno dadas en un intervalo $[a, b]$, siempre podemos reescribirlo como un problema variacional considerando el funcional

$$\mathcal{F}(y) := \int_a^b \left(\frac{p(t)}{2} (y'(t))^2 - \frac{q(t)}{2} y(t)^2 + r(t)y(t) \right) dt. \quad (69)$$

En este caso la función F es

$$F(t, y, z) = \frac{p}{2} z^2 - \frac{q}{2} y^2 + ry, \quad (70)$$

así que

$$\partial_y F = -qy + r, \quad \partial_z F = pz,$$

luego la ecuación de Euler-Lagrange correspondiente a (69) es exactamente (68). Usando lo que conocemos sobre los problemas de contorno de este tipo podemos demostrar el siguiente resultado:

Teorema 7.18. Sea $a < b \in \mathbb{R}$, $p \in \mathcal{C}^1[a, b]$, $q, r \in \mathcal{C}[a, b]$, y consideremos el funcional (69) definido en

$$\mathcal{D} := \{y \in \mathcal{C}^2[a, b] \mid y(a) = y(b) = 0\}.$$

Si $p > 0$ en $[a, b]$ y $q < 0$ en $[a, b]$ entonces el funcional \mathcal{F} tiene un único mínimo en \mathcal{D} , igual a la única solución del problema de contorno (68) en \mathcal{D} .

Demostración. Podemos comprobar que $F(t, y, z) := \frac{p}{2}z^2 - \frac{q}{2}y^2 + ry$ es convexa en (y, z) , así que el funcional F es convexo en \mathcal{D} . Por tanto, un punto crítico suyo tiene que ser un mínimo. La condición para ser un punto crítico es precisamente el problema (68) con condiciones de contorno $y(a) = y(b) = 0$. Para ver que dicho problema tiene solución única, vamos a ver que la única solución de la ecuación homogénea

$$(px')' + qx = 0 \quad (71)$$

con condiciones

$$x(a) = x(b) = 0 \quad (72)$$

es la trivial. Para demostrarlo, supongamos que existe una solución no trivial. La función x alcanza un máximo en $[a, b]$, y podemos suponer que su valor en el máximo es estrictamente positivo (si no, consideramos $-x$, que también es una solución y debe tener un máximo positivo entonces, ya que x no es la solución trivial). El punto t_0 donde alcanza su máximo debe estar en (a, b) (ya que $x(a) = x(b) = 0$). En ese punto t_0 debe ocurrir $x'(t_0) = 0$ y $x''(t_0) \leq 0$, luego

$$0 = p(t_0)x''(t_0) + p'(t_0)x'(t_0) + q(t_0)x(t_0) = p(t_0)x''(t_0) + q(t_0)x(t_0).$$

Pero $p(t_0)x''(t_0) \leq 0$ y $q(t_0)x(t_0) < 0$, lo cual es una contradicción.

Hemos visto que el problema homogéneo (71)–(72) tiene solo la solución trivial. Por el Teorema de Fredholm, las ecuaciones de Euler-Lagrange tienen una única solución, que tiene que ser el único mínimo de \mathcal{F} en \mathcal{D} . \square

Ejercicio 7.4. Demuestra que los siguientes problemas de contorno tienen solución única:

$$\begin{aligned} (i) \quad & y'' + y' - ty = \cos t, & y(1) = y(2) = 0 \\ (ii) \quad & y'' + y' - ty = \cos t, & y(1) = 2, y(2) = 3 \end{aligned}$$

7.4. Problemas de Sturm-Liouville

Consideramos ahora un problema de contorno homogéneo en forma autoadjunta con un parámetro λ adicional:

$$\left\{ \begin{array}{l} (p(t)y')' + q(t)y + \lambda w(t)y = 0, \\ \alpha y(a) + \beta y'(a) = 0, \\ \delta y(b) + \gamma y'(b) = 0. \end{array} \right. \quad \begin{array}{l} (73a) \\ (73b) \\ (73c) \end{array}$$

Siempre supondremos $\lambda \in \mathbb{R}$, con las mismas condiciones que en la sección anterior, junto con la condición de que w sea continua y positiva:

$$\begin{aligned}
a &< b \in \mathbb{R}, \\
p, q, w &: [a, b] \rightarrow \mathbb{R} \text{ son funciones continuas,} \\
p &\text{ es una función } \mathbb{C}^1 \text{ en } [a, b], \\
p(t) &> 0, \quad w(t) > 0 \text{ para todo } t \in [a, b]. \\
\alpha, \beta, \delta, \gamma &\in \mathbb{R}, \text{ con } \alpha, \beta \text{ no ambas nulas, } \delta, \gamma \text{ no ambas nulas.}
\end{aligned} \tag{74}$$

A un problema de este tipo cuyos parámetros cumplen estas hipótesis se le llama *problema de Sturm-Liouville regular*. La solución trivial $y(t) = 0$, $t \in [a, b]$ es siempre solución de (73). Una de las preguntas fundamentales en este tipo de problemas es saber para qué valores de $\lambda \in \mathbb{R}$ el problema (73a)–(73b)–(73c) admite soluciones no triviales.

Definición 7.19. A un valor de $\lambda \in \mathbb{C}$ para el que (73) tiene al menos una solución no trivial se le llama *valor propio* del problema (73). A una solución no trivial de (73) se le llama *función propia* asociada a λ .

Fíjate en que permitimos $\lambda \in \mathbb{C}$ en la definición anterior (y admitimos soluciones de (73) donde y tome valores complejos, $y: [a, b] \rightarrow \mathbb{C}$). Ésta es la definición estándar (el *espectro* de un operador se considera siempre en \mathbb{C}), pero no es una extensión importante porque en este caso todos los valores propios tienen que ser reales:

Lema 7.20. *Suponemos las condiciones (74), y llamamos L al operador dado por*

$$Ly := (py')' + qy.$$

Se cumple que

$$\langle y, Ly \rangle \in \mathbb{R}$$

para toda función $y: [a, b] \rightarrow \mathbb{C}$ en $L^2([a, b], \mathbb{C})$. En particular, los valores propios del problema (73) son reales.

Observación 7.21. Observa que el producto escalar que estamos usando aquí es el de $L^2([a, b], \mathbb{C})$ de funciones con valores complejos:

$$\langle x, y \rangle := \int_a^b x(t) \overline{y(t)} dt.$$

Demostración. Para cualquier $y \in L^2([a, b], \mathbb{C}; w)$,

$$\langle y, Ly \rangle = \langle Ly, y \rangle = \overline{\langle y, Ly \rangle},$$

donde la primera igualdad es por (62) y la segunda por la simetría del producto escalar complejo. Por tanto, $\langle y, Ly \rangle \in \mathbb{R}$. Con esto, si λ es un valor propio y ϕ una función propia asociada,

$$\lambda \|\phi\|_{L^2(w)}^2 = \langle L\phi, \phi \rangle \in \mathbb{R},$$

luego $\lambda \in \mathbb{R}$. □

Lema 7.22. *Suponemos las condiciones (43). Si ϕ, ψ son funciones propias del problema (73) asociadas a valores diferentes de λ entonces*

$$\int_a^b w(t) \phi(t) \psi(t) dt = 0.$$

Demostración. Llamamos L al operador dado por

$$Ly := (py')' + qy.$$

Supongamos que ϕ está asociada al valor propio λ_1 y ψ a λ_2 , con $\lambda_1 \neq \lambda_2$. Usando la fórmula de Green del Lema 7.13 tenemos

$$0 = \int_a^b (\phi L\psi - \psi L\phi) = (\lambda_1 - \lambda_2) \int_a^b w\psi\phi,$$

que implica el resultado. \square

Teorema 7.23 (Teorema de Sturm-Liouville). *Los valores propios del problema (73) son reales, simples, y forman una sucesión creciente $(\lambda_i)_{i \geq 1}$ que diverge a $+\infty$. Si llamamos ϕ_i a un generador del espacio de funciones propias asociadas a λ_i , entonces la sucesión $(\phi_i)_{i \geq 1}$ es una base ortogonal del espacio $L^2((a, b), w)$.*

Si suponemos que $p(a) = p(b)$, el resultado es también cierto también con condiciones de contorno periódicas, con la diferencia de que ahora los valores propios no tienen por qué ser simples. Si tomamos una base ortogonal de cada espacio propio del problema, su unión forma una base ortogonal de $L^2((a, b), w)$.

Ejemplo 7.24. Consideramos el problema

$$\begin{cases} y'' + \lambda y = 0, & (75a) \\ y(-\pi) = y(\pi) = 0. & (75b) \end{cases}$$

Buscamos sus valores propios (los valores de λ para los que hay soluciones no triviales).

1. Cuando $\lambda < 0$ las soluciones de la ecuación diferencial (75a) son

$$y(t) = Ae^{t\sqrt{|\lambda|}} + Be^{-t\sqrt{|\lambda|}},$$

de forma que para cumplir (75b) debe ocurrir

$$\begin{aligned} Ae^{-\pi\sqrt{|\lambda|}} + Be^{\pi\sqrt{|\lambda|}} &= 0, \\ Ae^{\pi\sqrt{|\lambda|}} + Be^{-\pi\sqrt{|\lambda|}} &= 0. \end{aligned}$$

Comprobamos que la matriz que define este sistema lineal es no singular, luego la única solución es $A = B = 0$.

2. Cuando $\lambda > 0$ las soluciones de la ecuación diferencial (75a) son

$$y(t) = A \cos(t\sqrt{\lambda}) + B \sin(-t\sqrt{\lambda}),$$

de forma que para cumplir (75b) debe ocurrir

$$\begin{aligned} A \cos(-\pi\sqrt{\lambda}) + B \sin(\pi\sqrt{\lambda}) &= 0, \\ A \cos(\pi\sqrt{\lambda}) + B \sin(-\pi\sqrt{\lambda}) &= 0. \end{aligned}$$

La matriz de coeficientes de este sistema lineal es

$$\begin{pmatrix} \cos(-\pi\sqrt{\lambda}) & \sin(\pi\sqrt{\lambda}) \\ \cos(\pi\sqrt{\lambda}) & \sin(-\pi\sqrt{\lambda}) \end{pmatrix} = \begin{pmatrix} \cos(\pi\sqrt{\lambda}) & \sin(\pi\sqrt{\lambda}) \\ \cos(\pi\sqrt{\lambda}) & -\sin(\pi\sqrt{\lambda}) \end{pmatrix},$$

con determinante $-2\cos(\pi\sqrt{\lambda})\sin(\pi\sqrt{\lambda}) = -2\sin(2\pi\sqrt{\lambda})$. Para que este determinante sea 0 debe ocurrir que

$$2\pi\sqrt{\lambda} = n\pi$$

con $n > 0$ entero; es decir,

$$\lambda = \frac{n^2}{4}.$$

Esto nos dice que para cualquier $n > 0$ entero, el número $\lambda = n^2/4$ es un valor propio de (75). Para encontrar una función propia asociada resolvemos el sistema anterior, que queda en este caso

$$\begin{aligned} A \cos\left(\frac{n\pi}{2}\right) + B \sin\left(\frac{n\pi}{2}\right) &= 0, \\ A \cos\left(\frac{n\pi}{2}\right) - B \sin\left(\frac{n\pi}{2}\right) &= 0. \end{aligned}$$

Cuando n es par obtenemos simplemente

$$A = 0$$

y cuando n es impar,

$$B = 0.$$

Por tanto, unos posibles generadores del espacio de funciones propias son

$$\begin{aligned} \sin\left(\frac{tn}{2}\right) & \quad \text{asociada a } \lambda = n^2/4 \text{ con } n \text{ par,} \\ \cos\left(\frac{tn}{2}\right) & \quad \text{asociada a } \lambda = n^2/4 \text{ con } n \text{ impar,} \end{aligned}$$

para $n \geq 1$ entero.

3. Cuando $\lambda = 0$ las soluciones de (75a) son

$$y(t) = A + Bt,$$

que podemos comprobar que sólo cumple la condición de contorno cuando $A = B = 0$.

En conclusión, los valores propios de (75) son

$$\lambda = \frac{n^2}{4}, \quad n \geq 1 \text{ entero,}$$

con funciones propias asociadas

$$\begin{aligned} \sin\left(\frac{tn}{2}\right) & \quad \text{asociada a } \lambda = n^2/4 \text{ con } n \text{ par,} \\ \cos\left(\frac{tn}{2}\right) & \quad \text{asociada a } \lambda = n^2/4 \text{ con } n \text{ impar.} \end{aligned}$$

Ejemplo 7.25. Consideramos el problema

$$\begin{cases} y'' + \lambda y = 0, & (76a) \\ y(-\pi) = y(\pi), \quad y'(-\pi) = y'(\pi), & (76b) \end{cases}$$

que ahora tiene condiciones de frontera periódicas. Buscamos sus valores propios (los valores de λ para los que hay soluciones no triviales).

1. Cuando $\lambda < 0$ las soluciones de la ecuación diferencial (76a) son

$$y(t) = Ae^{t\sqrt{|\lambda|}} + Be^{-t\sqrt{|\lambda|}},$$

de forma que para cumplir (76b) debe ocurrir

$$\begin{aligned} Ae^{-\pi\sqrt{|\lambda|}} + Be^{\pi\sqrt{|\lambda|}} &= Ae^{\pi\sqrt{|\lambda|}} + Be^{-\pi\sqrt{|\lambda|}}, \\ \sqrt{|\lambda|}Ae^{-\pi\sqrt{|\lambda|}} - \sqrt{|\lambda|}Be^{\pi\sqrt{|\lambda|}} &= \sqrt{|\lambda|}Ae^{\pi\sqrt{|\lambda|}} - \sqrt{|\lambda|}Be^{-\pi\sqrt{|\lambda|}}. \end{aligned}$$

Es decir:

$$\begin{aligned} A(e^{-\pi\sqrt{|\lambda|}} - e^{\pi\sqrt{|\lambda|}}) + B(e^{\pi\sqrt{|\lambda|}} - e^{-\pi\sqrt{|\lambda|}}) &= 0, \\ A(e^{-\pi\sqrt{|\lambda|}} - e^{\pi\sqrt{|\lambda|}}) - B(e^{\pi\sqrt{|\lambda|}} - e^{-\pi\sqrt{|\lambda|}}) &= 0. \end{aligned}$$

Comprobamos que la matriz que define este sistema lineal es no singular, luego la única solución es $A = B = 0$.

2. Cuando $\lambda > 0$ las soluciones de la ecuación diferencial (76a) son

$$y(t) = A \cos(t\sqrt{\lambda}) + B \sin(-t\sqrt{\lambda}),$$

de forma que para cumplir (76b) debe ocurrir

$$\begin{aligned} A \cos(-\pi\sqrt{\lambda}) + B \sin(\pi\sqrt{\lambda}) &= A \cos(\pi\sqrt{\lambda}) + B \sin(-\pi\sqrt{\lambda}), \\ -A \sin(-\pi\sqrt{\lambda}) + B \cos(\pi\sqrt{\lambda}) &= -A \sin(\pi\sqrt{\lambda}) + B \cos(-\pi\sqrt{\lambda}), \end{aligned}$$

es decir,

$$B \sin(\pi\sqrt{\lambda}) = 0, \quad A \sin(\pi\sqrt{\lambda}) = 0.$$

La única opción para tener alguna solución distinta de $A = B = 0$ es que $\sin(\pi\sqrt{\lambda}) = 0$, es decir,

$$\pi\sqrt{\lambda} = n\pi$$

con $n > 0$ entero; es decir,

$$\lambda = n^2.$$

Esto nos dice que para cualquier $n > 0$ entero, el número $\lambda = n^2$ es un valor propio de (76). Cualquier solución de (76a) con este valor de λ es una función propia asociada a λ , así que unos posibles generadores del espacio de funciones propias son

$$\{\sin(tn), \cos(tn)\} \quad \text{asociadas a } \lambda = n^2$$

para $n \geq 1$ entero.

3. Cuando $\lambda = 0$ las soluciones de (76a) son

$$y(t) = A + Bt,$$

que cumplen la condición de contorno cuando $B = 0$. Así que $\lambda = 0$ es también un valor propio (simple) con función propia asociada 1.

En conclusión, los valores propios de (76) son

$$\lambda = n^2, \quad n \geq 0 \text{ entero},$$

con funciones propias asociadas

$$\begin{array}{ll} 1 & \text{asociada a } \lambda = 0, \\ \{\sin(tn), \cos(tn)\} & \text{asociados a } \lambda = n^2 \text{ con } n \geq 1. \end{array}$$

Ejemplo 7.26. Los polinomios de Hermite son las soluciones no triviales del problema

$$y'' - ty' + \lambda y = 0$$

que cumplen unas condiciones de contorno un poco distintas: son aquellas que crecen de forma polinomial cuando $x \rightarrow \pm\infty$. Escrito en forma autoadjunta el problema es

$$(e^{-\frac{t^2}{2}} y')' + \lambda e^{-\frac{t^2}{2}} y = 0.$$

Ejemplo 7.27. Los polinomios de Laguerre son las soluciones no triviales del problema

$$ty'' + (1-t)y' + \lambda y = 0.$$

Ejemplo 7.28. Los polinomios de Jacobi de tipo (α, β) son las soluciones no triviales del problema

$$(1-t^2)y'' + (\beta - \alpha - (\alpha + \beta + 2)t)y' + \lambda y = 0.$$

Ejercicio 7.5. Dado $L > 0$, consideramos el siguiente problema de Sturm-Liouville:

$$\begin{cases} y'' + \lambda y = 0, & (77a) \\ y(0) = y(L) = 0. & (77b) \end{cases}$$

Demuestra que los valores propios de este problema son

$$\lambda = \frac{n^2\pi^2}{L^2}, \quad n \geq 1 \text{ entero},$$

y que los espacios de funciones propias están generados por las funciones

$$\psi_n(t) = \sin\left(\frac{n\pi t}{L}\right).$$

7.4.1. Demostración del Teorema de Sturm-Liouville

Algunas partes del Teorema 7.23 son fáciles de demostrar. Ya hemos visto que los valores propios son reales (Lema 7.20), que son simples en el caso de las condiciones de contorno separadas (Ejercicio 7.2), y que funciones propias asociadas a distintos valores propios son ortogonales (Lema 7.22). En realidad la parte más importante de la demostración es el siguiente resultado:

Lema 7.29. *Existe al menos un valor propio $\lambda \in \mathbb{R}$ del problema (73).*

Demostración. Consideramos el problema de minimizar el funcional

$$\mathcal{F}(y) := \int_a^b (p(y')^2 - qy^2) dt$$

en el dominio

$$\mathcal{D} := \{y \in \mathcal{C}^2[a, b] \mid y \text{ cumple (73b), (73c), } \|y\|_{L^2(w)} = 1\}.$$

Observa que si este problema alcanza un mínimo en $y \in \mathcal{D}$, este mínimo debe cumplir las ecuaciones de Euler-Lagrange, lo cual dice precisamente que y es una función propia de (73). Lo que vamos a demostrar es que efectivamente alcanza un mínimo.

Lo primero que notamos es que \mathcal{F} está acotado inferiormente en \mathcal{D} , ya que por la desigualdad de Poincaré tenemos

$$\begin{aligned} |\mathcal{F}(y)| &\geq \int_a^b p(y')^2 - \int_a^b |q|y^2 \geq p_{\min} \int_a^b (y')^2 - q_{\max} \int_a^b y^2 \\ &\geq (Cp_{\min} - q_{\max}) \int_a^b y^2 \geq -\frac{1}{w_{\min}} |Cp_{\min} - q_{\max}| \int_a^b wy^2 = -\frac{1}{w_{\min}} |Cp_{\min} - q_{\max}|, \end{aligned}$$

donde p_{\min} , q_{\max} , w_{\min} representan los máximos o mínimos valores en $[a, b]$ de las funciones $|p|$, $|q|$, $|w|$. Por tanto podemos considerar una sucesión minimizante $(y_n)_{n \geq 1}$; es decir, una sucesión en \mathcal{D} tal que $\mathcal{F}(y_n) \rightarrow \inf_{y \in \mathcal{D}} \mathcal{F}(y)$. Para esta sucesión, $\mathcal{F}(y_n)$ está acotado por una constante C , luego

$$p_{\min} \int_a^b (y'_n)^2 \leq \int_a^b p(y'_n)^2 \leq \mathcal{F}(y_n) + \int_a^b |q|y_n^2 \leq C + \frac{q_{\min}}{w_{\max}} \int_a^b wy_n^2 = C + \frac{q_{\min}}{w_{\max}}.$$

Esto demuestra que $\int_a^b (y'_n)^2$ está uniformemente acotado para $n \geq 1$. Como $\int_a^b y^2 \leq w_{\min}^{-1} \int_a^b wy^2 = w_{\min}^{-1}$, vemos que la sucesión (y_n) está acotada en $H^1[a, b]$. Por tanto hay una subsucesión (a la que seguimos llamando (y_n)) y una función $y_* \in H^1[a, b]$ tal que

1. (y_n) converge débilmente en H^1 a y_* ,
2. (y_n) converge en $L^2[a, b]$ a y_* .

Gracias a la convergencia en $L^2[a, b]$ vemos que $\int_a^b wy_*^2 = 1$. Usando los dos tipos de convergencia,

$$\mathcal{F}(y_*) \leq \lim_{n \rightarrow +\infty} \mathcal{F}(y_n) = \inf_{y \in \mathcal{D}} \mathcal{F}(y).$$

Lo único que nos falta es ver que $y \in \mathcal{D}$. Esto es un resultado de regularidad de las soluciones de la ecuación diferencial (73a), ya que se puede ver que y cumple la ecuación en sentido débil. Omitimos esta parte de la demostración. \square

Corolario 7.30. *El problema (73) tiene un número infinito de valores propios. Las funciones propias asociadas forman un conjunto ortogonal completo de $L^2([a, b], w)$ (el cierre del espacio que generan es $L^2([a, b], w)$).*

Demostración. Supongamos que sólo hay una cantidad finita $\lambda_1, \dots, \lambda_n$, asociados a las funciones propias ϕ_1, \dots, ϕ_n . Consideramos el espacio V generado por las funciones ϕ_1, \dots, ϕ_n . Como es un espacio cerrado, tiene un complemento ortogonal V^\perp en $L^2([a, b], w)$, que no puede ser vacío porque $L^2([a, b], w)$ tiene dimensión infinita. Es fácil ver que L se puede restringir a $X := V^\perp \cap \mathcal{D}_0$ (con \mathcal{D}_0 el conjunto de funciones \mathcal{C}^2 que cumplen las condiciones de contorno), de forma que tenemos un operador $L: X \rightarrow X$. Con un razonamiento muy parecido al del lema anterior podemos ver que L alcanza un mínimo en la esfera unidad de X , que tiene que alcanzarse en un valor propio que esté en el espacio X . Hemos encontrado por tanto una función propia nueva, lo que contradice el supuesto de que el conjunto de vectores propios está generado por ϕ_1, \dots, ϕ_n .

Podemos probar la segunda parte con un argumento parecido: si el cierre del espacio generado por ϕ_1, \dots, ϕ_n no es igual a $L^2([a, b], w)$, tiene que tener un complemento ortogonal no vacío, y por el mismo razonamiento que antes dicho complemento ortogonal tiene que contener una nueva función propia, lo cual es una contradicción. \square

Lema 7.31. *Los valores propios de (73) son aislados.*

Demostración. En caso contrario, hay una sucesión de valores propios λ_n que convergen a un punto $\lambda \in \mathbb{R}$. Tomemos funciones propias asociadas $(\phi_n)_{n \geq 1}$, normalizadas con $\|\phi_n\|_{L^2(w)} = 1$. Como son ortogonales

$$\|\phi_n - \phi_m\|_{L^2(w)} = 1 \quad \text{para todo } n \neq m. \quad (78)$$

Sin embargo, por los resultados de regularidad que hemos mencionado antes, las funciones ϕ_n están uniformemente acotadas en $H^1[a, b]$, y por tanto deben tener una parcial convergente en $L^2[a, b]$. Como las normas en $L^2[a, b]$ y $L^2([a, b], w)$ son equivalentes, esto contradice (78). \square

7.4.2. Relación con el cálculo de variaciones

Los problemas del tipo (73) aparecen de forma natural en el estudio de problemas con restricciones de tipo integral (ver Sección 4.3.2). Si consideramos el funcional

$$\mathcal{F}(y) = \int_a^b (p(t)(y'(t))^2 - q(t)y(t)^2) dt \quad (79)$$

sujeto a la restricción

$$\int_a^b w(t)y(t)^2 dt = I \quad (80)$$

(con $I > 0$ dado) entonces la ecuación de Euler-Lagrange correspondiente es

$$(py')' = -qy - \lambda wy,$$

donde $\lambda \in \mathbb{R}$ es también un parámetro por determinar. Si consideramos el funcional \mathcal{F} definido en el dominio

$$\mathcal{D} := \{y \in \mathcal{C}^2[a, b] \mid y(a) = y(b) = 0, \text{ y se cumple (80)}\} \quad (81)$$

entonces los puntos críticos del funcional \mathcal{F} en \mathcal{D} son exactamente las soluciones de un problema de Sturm-Liouville. Un resultado importante es el siguiente:

Teorema 7.32. *Suponemos las condiciones (74). El funcional (79) alcanza un mínimo en el dominio*

$$\mathcal{D} := \{y \in \mathcal{C}^1[a, b] \mid y(a) = 0, y(b) = 0, \int_a^b w(t)y(t)^2 dt = 1\}.$$

Dicho mínimo se alcanza en la primera función propia ϕ_1 (la asociada al valor propio más bajo), normalizada de forma que $\int_a^b w\phi_1^2 = 1$, del problema de Sturm-Liouville dado por

$$(py')' + qy + \lambda wy = 0, \quad y(a) = y(b) = 0. \quad (82)$$

El valor del mínimo es igual al menor valor propio de este problema.

De hecho, el mínimo del teorema anterior se alcanza en dos funciones: tanto en ϕ_1 como en $-\phi_1$. Podemos ver de la demostración que sólo se alcanza en esas dos funciones.

Demostración. Si denotamos por L el operador

$$Ly = (py')' + qy,$$

entonces podemos comprobar que para todo $y \in \mathcal{D}$ tenemos (gracias a las condiciones de contorno)

$$\mathcal{F}(y) = - \int_a^b y(t)(Ly)(t) dt. \quad (83)$$

Gracias al Teorema 7.23 sabemos que podemos encontrar una base ortonormal de $L^2((a, b), w)$ formada por funciones propias del problema (82). Denotamos esta base por $\mathcal{B} := (\phi_i)_{i \geq 0}$, con ϕ_i una función propia asociada al valor propio λ_i (ordenados de manera creciente), de forma que

$$L\phi_i = -\lambda_i w\phi_i, \quad i \geq 0.$$

Sea $y \in \mathcal{D}$ una función cualquiera. Podemos escribirla en coordenadas de la base \mathcal{B} como

$$y = \sum_{i=0}^{\infty} \mu_i \phi_i,$$

con $\mu_i \in \mathbb{R}$. Usando (83) tenemos

$$\begin{aligned} \mathcal{F}(y) &= - \int_a^b y(t)(Ly)(t) dt = - \int_a^b \left(\sum_{i=0}^{\infty} \mu_i \phi_i \right) L \left(\sum_{j=0}^{\infty} \mu_j \phi_j \right) dt \\ &= - \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \mu_i \mu_j \int_a^b \phi_i L\phi_j dt = \sum_{i=0}^{\infty} \mu_i^2 \lambda_i \int_a^b w\phi_i^2 dt = \sum_{i=0}^{\infty} \mu_i^2 \lambda_i. \end{aligned}$$

Como y tiene que satisfacer la restricción $\int_a^b w y^2 = 1$,

$$1 = \int_a^b w y^2 dt = \int_a^b w \left(\sum_{i=0}^{\infty} \mu_i \phi_i \right)^2 dt = \sum_{i=0}^{\infty} \mu_i^2 \int_a^b w \phi_i^2 dt = \sum_{i=0}^{\infty} \mu_i^2. \quad (84)$$

El menor valor posible de $\mathcal{F}(y) = \sum_{i=0}^{\infty} \mu_i^2 \lambda_i$ sujeto a la restricción (84) ocurre cuando $\mu_1 = 1$. Este valor se alcanza cuando $y = \phi_1$. \square

Podemos escribir una versión un poco más general de este resultado incluyendo otras restricciones:

Teorema 7.33. *Suponemos las condiciones (74). Denotamos por $(\lambda_i)_{i \geq 0}$ los valores propios del problema de Sturm-Liouville (82), y por $(\phi_i)_{i \geq 0}$ una base ortonormal de $L^2((a, b), w)$ formada por funciones propias del problema (de forma que ϕ_i corresponde a λ_i para cada i).*

Tomemos $N \geq 0$ un entero. El funcional (79) alcanza un mínimo en el dominio

$$\mathcal{D} := \left\{ y \in C^1[a, b] \mid y(a) = y(b) = 0, \int_a^b w y^2 = 1, \int_a^b w \phi_j y = 0 \text{ para } j = 0, \dots, N \right\}.$$

Dicho mínimo se alcanza en la función propia ϕ_{N+1} del problema de Sturm-Liouville (82), y el valor del mínimo es igual al valor propio λ_{N+1} de este problema.

Demostración. La demostración es exactamente igual que antes: si escribimos una función $y \in \mathcal{D}$ en la base $(\phi_i)_{i \geq 0}$, las restricciones implican que

$$y = \sum_{i=N+1}^{\infty} \mu_i \phi_i,$$

con $\mu_i \in \mathbb{R}$ tales que $\sum_{i=N+1}^{\infty} \mu_i^2 = 1$. Observa que los términos desde $i = 0$ hasta $i = N$ desaparecen debido a las condiciones de ortogonalidad $\int_a^b w \phi_j y = 0$ para $j = 0, \dots, N$. Al igual que antes tenemos

$$\mathcal{F}(y) = \sum_{i=N+1}^{\infty} \mu_i^2 \lambda_i.$$

El mínimo de esta expresión sujeto a la restricción $\sum_{i=N+1}^{\infty} \mu_i^2 = 1$ se alcanza cuando $\mu_{N+1} = 1$, $\mu_j = 0$ para $j > N + 1$. Esto corresponde a la función ϕ_{N+1} . \square

Ejercicio 7.6. Calcula el mínimo del funcional

$$\mathcal{F}(y) = \int_0^{\pi} (y'(t))^2$$

en el dominio

$$\mathcal{D} = \left\{ y \in C^1[0, \pi] \mid y(0) = y(\pi) = 0, \int_0^{\pi} (y(t))^2 dt = 1 \right\}.$$

Calcula también su mínimo (si tiene) en los dominios

$$\begin{aligned} \mathcal{D}_2 &= \left\{ y \in C^1[0, \pi] \mid y(0) = y(\pi) = 0, \int_0^{\pi} (y(t))^2 dt = 1, \int_0^{\pi} y(t) \sin t dt = 0 \right\}, \\ \mathcal{D}_3 &= \left\{ y \in C^1[0, \pi] \mid y(0) = y(\pi) = 0, \int_0^{\pi} (y(t))^2 dt = 1, \int_0^{\pi} y(t) dt = 0 \right\}. \end{aligned}$$

Ejercicio 7.7. Calcula el mínimo del funcional

$$\mathcal{F}(y) = \int_1^e t(y'(t))^2$$

en el dominio

$$\mathcal{D} = \left\{ y \in C^2[1, e] \mid y(1) = y(e) = 0, \int_1^e \frac{y^2}{t} dt = 1, \int_1^e \frac{y}{t} \sin(\pi \log t) dt = 0 \right\}.$$

Solución. Para saber si es un problema al que le podamos aplicar el Teorema 7.33 tenemos que calcular la primera función propia del problema de Sturm-Liouville

$$(t(y'))' + \frac{\lambda}{t}y = 0,$$

es decir,

$$t^2y'' + ty' + \lambda y = 0. \quad (85)$$

Esta es una ecuación de tipo Euler, que puede resolverse con el cambio $s = \log t$. Si escribimos $z = z(s) = y(t)$ entonces

$$y' = \frac{1}{t}z', \quad y'' = -\frac{1}{t^2}z' + \frac{1}{t^2}z'',$$

luego la ecuación (85) queda

$$z'' + \lambda z = 0,$$

con las condiciones de contorno

$$z(0) = z(1) = 0.$$

Sabemos que los valores propios de esta ecuación son positivos (como en ejemplos anteriores), y para $\lambda > 0$ tenemos

$$z(t) = A \cos(t\sqrt{\lambda}) + B \sin(t\sqrt{\lambda}),$$

que con las condiciones de contorno implica que

$$\sin(\sqrt{\lambda}) = 0,$$

es decir,

$$\lambda = n^2\pi^2, \quad n \geq 1 \text{ entero.}$$

El primer valor propio es $\lambda_1 = \pi^2$, y sus funciones propias asociadas son

$$\phi(s) = B \sin(\pi s).$$

Deshaciendo el cambio $s = \log t$, la primera función propia del problema (85) es

$$\psi(t) = B \sin(\lambda \log t).$$

Por tanto este problema es justo del tipo de los que aparecen en el Teorema 7.33, y por tanto el mínimo del funcional se alcanza en la segunda función propia ψ_2 , normalizada tal que

$$1 = \int_1^e \frac{1}{t} (\psi_2(t))^2 dt,$$

y el valor del mínimo es el segundo valor propio, $4\pi^2$.

Ejercicio 7.8. Sea \mathcal{D} el dominio

$$\mathcal{D} := \{y \in C^2[0, \pi] \mid y(0) = y(\pi) = 0, \int_0^\pi (y(x))^2 dx = 5\}.$$

Consideramos el siguiente funcional \mathcal{F} , definido en \mathcal{D} :

$$\mathcal{F}(y) := \int_0^\pi ((y'(x))^2 + 3(y(x))^2) dx.$$

Demuestra que el funcional \mathcal{F} alcanza un mínimo global en \mathcal{D} , y encuentra todas las funciones $y \in \mathcal{D}$ donde lo alcanza.

Solución. Para este ejercicio podemos aplicar directamente el Teorema 7.32 sobre la interpretación variacional de los problemas de Sturm-Liouville. La ecuación de Sturm-Liouville asociada es

$$y'' - 3y + \frac{\lambda}{5}y = y'' + \left(\frac{\lambda}{5} - 3\right)y = 0,$$

con condición de frontera $y(0) = y(\pi) = 0$. Si llamamos $\mu = \lambda/5 - 3$, sabemos que las únicas soluciones no triviales (funciones propias) se dan para $\mu = n^2$, con $n \geq 1$ un entero, y son múltiplos de $y_n(x) = \sin(nx)$. Estos valores de μ corresponden a

$$\lambda = 5(n^2 + 3),$$

que son los valores propios del problema de Sturm-Liouville. El valor del mínimo es por tanto el valor del menor valor propio, obtenido para $n = 1$: $\lambda = 20$. Este valor se alcanza en

$$y(x) = A \sin(x),$$

con la constante A elegida de tal forma que se cumpla la restricción

$$\int_0^\pi (y(x))^2 dx = 5,$$

es decir,

$$5 = A^2 \int_0^\pi (\sin(x))^2 dx = A^2 \frac{\pi}{2}, \quad \text{luego } A = \pm \sqrt{\frac{10}{\pi}}.$$

Por tanto el mínimo se alcanza en dos funciones:

$$y(x) = \pm \sqrt{\frac{10}{\pi}} \sin x, \quad x \in [0, \pi].$$

7.5. Un problema de contorno de orden superior: el problema de la viga

Supongamos que tenemos una viga sujeta por sus dos extremos a dos paredes, y que está soportando una cierta carga a lo largo de su longitud. ¿Cuál es la forma que adopta la viga bajo el peso?

La deformación de la viga depende del material del que esté hecha, de la forma exacta en que esté sujeta en sus extremos, y por supuesto del reparto del peso. Supongamos que la distancia entre las dos paredes es L , y sea $y : [0, L] \rightarrow \mathbb{R}$ la curva que describe la viga. Supongamos que los dos extremos de la viga están a la misma altura (que suponemos igual a 0, ajustando el eje vertical), de forma que

$$y(0) = y(L) = 0. \tag{86}$$

El peso que soporta la viga lo describimos por medio de una *densidad de carga* $p = p(x)$ tal que $\int_a^b p(x) dx$ es el peso que soporta la viga entre los puntos $a, b \in [0, L]$. La forma que adopta la viga es aquella *que minimice la energía total*. La energía total tiene dos partes: una elástica que tiene que ver con la deformación de la viga a través de su *curvatura*:

$$E_e := M \int_0^L |y''(x)|^2 dx.$$

La constante M se llama *módulo de elasticidad a flexión* de la viga. Por otra parte, la *energía potencial* viene dada por la altura a la que está la viga y la densidad de carga p :

$$E_p := g \int_0^L p(x)y(x) dx,$$

donde g es la aceleración de la gravedad. De ahora en adelante ponemos las constantes $M, g = 1$ cambiando las unidades, ya que el problema matemático es el mismo. Así que estamos buscando el mínimo del funcional

$$F[y] := \int_0^L |y''(x)|^2 dx + \int_0^L p(x)y(x) dx$$

entre las funciones y para las cuales tenga sentido, y que cumplan la condición de contorno (86). De hecho, nos faltan otras condiciones de contorno que especifican exactamente *cómo* está sujeta la viga a la pared. La viga podría estar *empotrada* en la pared de forma que entre de forma horizontal, con lo que tendríamos

$$y'(0) = y'(L) = 0, \quad (87)$$

o podría estar simplemente apoyada en un pivote (por ejemplo, si fuera una estantería con libros en lugar de una viga). En ese caso tendríamos

$$y''(0) = y''(L) = 0. \quad (88)$$

Podría ocurrir también que uno de los extremos estuviera libre, por ejemplo el extremo en $x = L$, en cuyo caso deberíamos imponer

$$y''(L) = y'''(L) = 0. \quad (89)$$

Supongamos por ahora (87). Estamos por tanto buscando el mínimo del funcional F definido sobre el conjunto

$$\mathcal{D} := \{y \in \mathcal{C}^4[0, L] \mid \text{se cumplen (86) y (87)}\}. \quad (90)$$

Para hacer el cálculo más fácil, supongamos que y es una función de \mathcal{D} donde se alcanza el mínimo del funcional F , y que además $y \in \mathcal{C}^4[0, L]$. No sabemos directamente cuáles son las ecuaciones de Euler-Lagrange en este caso porque el funcional depende también de y'' , pero podemos encontrar una condición necesaria que y debe cumplir con un razonamiento muy parecido. Sea φ una función \mathcal{C}^∞ y con soporte compacto en $(0, L)$. La variación de Gâteaux de F en la dirección de φ es

$$\delta_\varphi F[y] = \int_0^L (2y''(x)\varphi''(x) + p(x)\varphi(x)) dx,$$

que integrando por partes da

$$\delta_\varphi F[y] = \int_0^L (2y^{(4)}(x)\varphi(x) + p(x)\varphi(x)) dx = \int_0^L (2y^{(4)}(x) + p(x))\varphi(x) dx.$$

Si esto es cierto para toda φ tenemos

$$2y^{(4)}(x) + p(x) = 0 \quad \text{para } x \in [0, L]. \quad (91)$$

Lo que obtenemos es que y debe cumplir un problema de contorno de cuarto orden, con las condiciones de contorno (86), (87).

Ejemplo 7.34. Calculemos la forma que adopta una estantería bajo un peso uniformemente repartido ($p(x) \equiv 1$), con extremos que están simplemente apoyados sobre un soporte. Un mínimo de la energía F (si existe) debe cumplir la ecuación (91), que en este caso es

$$y^{(4)}(x) = -\frac{1}{2} \quad \text{para } x \in [0, L]. \quad (92)$$

con las condiciones de contorno de tipo apoyo:

$$y(0) = y(L) = 0, \quad y''(0) = y''(L) = 0. \quad (93)$$

Integrando la ecuación (92) con respecto a x vemos que

$$y'''(x) = C_1 - \frac{x}{2}.$$

Para cierta constante $C_1 \in \mathbb{R}$. Integrando de nuevo,

$$y''(x) = C_2 + C_1x - \frac{x^2}{4}.$$

La condición de contorno $y''(0) = 0$ implica $C_2 = 0$, y la condición $y''(L) = 0$ implica

$$C_1L - \frac{L^2}{4} = 0, \quad \text{luego} \quad C_1 = \frac{L}{4}. \quad (94)$$

Integrando de nuevo,

$$y'(x) = C_3 + \frac{C_1}{2}x^2 - \frac{x^3}{12},$$

y finalmente

$$y(x) = C_4 + C_3x + \frac{C_1}{6}x^3 - \frac{x^4}{48}. \quad (95)$$

La condición $y(0) = 0$ da $C_4 = 0$, y con $y(L) = 0$ obtenemos

$$C_3L + \frac{C_1}{6}L^3 - \frac{L^4}{48}, \quad \text{luego} \quad C_3 = \frac{L^4}{48} - \frac{L^4}{24} = -\frac{L^4}{48}. \quad (96)$$

La solución final es por tanto

$$y(x) = -\frac{L^4}{48}x + \frac{L}{24}x^3 - \frac{1}{48}x^4. \quad (97)$$

8. Ejercicios

Ejercicio 8.1. Calcula los valores propios y las funciones propias del siguiente Problema de Sturm Liouville en el intervalo $[0, \pi]$:

$$y'' + \lambda y = 0, \quad y'(0) = y'(\pi) = 0. \quad (98)$$

Ejercicio 8.2. Sea $L > 0$. Consideramos el funcional

$$F[y] := \int_0^L (|y''(x)|^2 + |y'(x)|^2) dx + \int_0^L y(x) dx,$$

definido en el dominio

$$D := \{y \in \mathcal{C}^4([0, L]) \mid y(0) = y'(0) = y''(L) = y'''(L) = 0\}. \quad (99)$$

1. Encuentra las ecuaciones de Euler-Lagrange asociadas a este funcional.
2. Encuentra una función que sea un punto crítico de este funcional (dicha función es de hecho un mínimo, como demostraremos más adelante en la Sección 12 — ver Ejercicio 12.9).

9. Series de Fourier

9.1. Serie de Fourier en $[-\pi, \pi]$

Definición 9.1 (Serie de Fourier). Dada una función integrable $f : [-\pi, \pi] \rightarrow \mathbb{R}$ su *serie de Fourier* es la expresión

$$\frac{a_0}{2} + \sum_{n \geq 1} a_n \cos(nx) + \sum_{n \geq 1} b_n \sin(nx), \quad (100)$$

donde los coeficientes $\{a_n\}_{n \geq 0}$, $\{b_n\}_{n \geq 1}$ vienen dados por

$$\begin{aligned} a_n &:= \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos(nx) \, dx, & (n \geq 0), \\ b_n &:= \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \sin(nx) \, dx, & (n \geq 1), \end{aligned} \quad (101)$$

La lista de funciones que aparecen en la serie (100),

$$\{1\} \cup \{\cos(nx), \sin(nx) \mid n \geq 1 \text{ entero}\},$$

se conoce como *base de Fourier en $[-\pi, \pi]$* . Son las que obtuvimos como funciones propias de un problema de Sturm-Liouville en el Ejercicio 7.25. Las funciones $\cos(nx)$, $\sin(nx)$ corresponden al mismo valor propio, y las hemos elegido de forma que son ortogonales con el producto escalar usual en $L^2([-\pi, \pi])$:

$$\int_{-\pi}^{\pi} \cos(nx) \sin(nx) \, dx = 0, \quad (n \geq 1) \quad (102)$$

como puede verse por ejemplo con el cambio de variable $x \mapsto -x$. Las funciones correspondientes a distintos valores de n son también ortogonales, como sabemos, ya que son funciones propias correspondientes a valores propios distintos de un problema de Sturm-Liouville. Es decir,

$$\begin{aligned} \int_{-\pi}^{\pi} \cos(nx) \cos(mx) \, dx &= \int_{-\pi}^{\pi} \sin(nx) \sin(mx) \, dx = \int_{-\pi}^{\pi} \cos(nx) \sin(mx) \, dx = 0, \\ &\text{para } n, m \geq 1, \quad n \neq m. \end{aligned} \quad (103)$$

También podemos demostrarlo directamente usando por ejemplo las expresiones

$$\cos(nx) \cos(mx) = \frac{1}{2} (\cos((n-m)x) + \cos((n+m)x)), \quad (104)$$

$$\sin(nx) \sin(mx) = \frac{1}{2} (\cos((n-m)x) - \cos((n+m)x)), \quad (105)$$

$$\sin(nx) \cos(mx) = \frac{1}{2} (\sin((n+m)x) + \sin((n-m)x)). \quad (106)$$

La función constante $1/2$ es también ortogonal a todas las demás. Por otra parte,

$$\int_{-\pi}^{\pi} (\cos(nx))^2 dx = \int_{-\pi}^{\pi} (\sin(nx))^2 dx = \pi, \quad (107)$$

$$\int_{-\pi}^{\pi} \left(\frac{1}{2}\right)^2 dx = \frac{\pi}{2}. \quad (108)$$

(como puede verse por ejemplo usando (104)). El Teorema 7.23 nos dice que las funciones $\{1/2, \cos(nx), \sin(nx) \mid n \geq 1\}$ forman una base de $L^2([-\pi, \pi])$ en el siguiente sentido:

Teorema 9.2. *Sea $f \in L^2([-\pi, \pi])$. Entonces la serie (100), con coeficientes dados por (101), converge en $L^2([-\pi, \pi])$ a la función f :*

$$\frac{a_0}{2} + \sum_{n=1}^m a_n \cos(nx) + \sum_{n=1}^m b_n \sin(nx) \xrightarrow{m \rightarrow +\infty} f \quad \text{en } L^2([-\pi, \pi]). \quad (109)$$

A la inversa, si (109) se cumple para ciertos coeficientes $\{a_i\}_{i \geq 0}$, $\{b_i\}_{i \geq 1}$ entonces a_i , b_i tienen que ser iguales a los definidos en (101).

Lema 9.3 (Compleitud de la serie de Fourier). *Sea $f: [-\pi, \pi] \rightarrow \mathbb{R}$ una función integrable. Si se cumple que*

$$\begin{aligned} \int_{-\pi}^{\pi} f(x) \sin(nx) dx &= 0 & n \geq 1 \text{ entero,} \\ \int_{-\pi}^{\pi} f(x) \cos(nx) dx &= 0 & n \geq 0 \text{ entero,} \end{aligned}$$

entonces $f = 0$ en casi todo punto de $[-\pi, \pi]$.

Idea de la demostración. Podemos escribir una aproximación de la función δ como

$$\varphi_k(x) := A_k \left(\frac{1 + \cos x}{2} \right)^k,$$

que es también una combinación lineal finita de funciones que aparecen en la base de Fourier. Como consecuencia, podemos ver que

$$\varphi_k * f = 0 \quad \text{en } [-\pi, \pi],$$

porque $\varphi(x - y)$, para x fijo, también se escribe como combinación finita de elementos de la base de Fourier. En el límite (como convergencia en $L^1[-\pi, \pi]$) obtenemos que

$$f = \delta * f = 0$$

en casi todo punto de $[-\pi, \pi]$.

(Hay muchas aproximaciones posibles para la función δ como combinación de elementos de la base de Fourier, y cualquiera de ellas da una demostración parecida de este resultado.) \square

Teorema 9.4 (Identidad de Parseval). Sea $f \in L^1([-\pi, \pi])$ y $\{a_i\}_{i \geq 0}$, $\{b_i\}_{i \geq 1}$ sus coeficientes de Fourier dados por (101). Entonces $f \in L^2([-\pi, \pi])$ si y sólo si $\sum_{n=1}^{\infty} (|a_n|^2 + |b_n|^2) < \infty$, y en ese caso

$$\frac{|a_0|^2}{2} + \sum_{n=1}^{\infty} (|a_n|^2 + |b_n|^2) = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x)^2 dx. \quad (110)$$

Idea de la demostración. Si $f \in L^2[-\pi, \pi]$, entonces sabemos por (109) que la serie de Fourier de f converge a f en L^2 . En particular, la norma de las sumas parciales de la serie de Fourier de f converge a la norma de f , lo que demuestra (110).

Por otra parte, si $\sum_n (|a_n|^2 + |b_n|^2)$ converge, entonces la serie de Fourier de f es de Cauchy. Como $L^2[-\pi, \pi]$ es completo, vemos que la serie de Fourier de f converge a cierta función $g \in L^2[-\pi, \pi]$, que debe cumplir

$$\begin{aligned} \int_{-\pi}^{\pi} f(x) \sin(nx) dx &= \int_{-\pi}^{\pi} g(x) \sin(nx) dx, & n \geq 1, \\ \int_{-\pi}^{\pi} f(x) \cos(nx) dx &= \int_{-\pi}^{\pi} g(x) \cos(nx) dx, & n \geq 0. \end{aligned}$$

Se puede ver que esto último implica que $f - g = 0$ en casi todo punto de $[-\pi, \pi]$, por el Lema 9.3. Por tanto, $f \in L^2[-\pi, \pi]$. \square

Ejercicio 9.1. Calcula los coeficientes de la serie de Fourier de la función $f(x) = x$ en el intervalo $[-\pi, \pi]$.

Ejercicio 9.2. ¿Son correctas las siguientes afirmaciones?

1. Existe una función en $L^2([-\pi, \pi])$ cuyos coeficientes de Fourier son $a_n = 0$ ($n \geq 0$), $b_n = 1/\sqrt{n}$ ($n \geq 1$).
2. Existe una función en $L^2([-\pi, \pi])$ cuyos coeficientes de Fourier son $a_n = 0$ ($n \geq 0$), $b_n = 1/n$ ($n \geq 1$).

Solución 9.5. 1. *Falsa.* La sucesión $1/\sqrt{n}$ no es de cuadrado sumable, y por tanto el Teorema 9.4 nos dice que no hay ninguna función en $L^2([-\pi, \pi])$ con dichos coeficientes.

2. *Verdadera.* La función es

$$f(x) = \sum_{n=1}^{\infty} \frac{1}{n} \sin(nx),$$

donde el límite se entiende en el sentido de $L^2([-\pi, \pi])$. Sabemos que esto define una función en $L^2([-\pi, \pi])$ por el Teorema 9.4, ya que la sucesión $1/n$ es de cuadrado sumable.

9.2. Serie de Fourier en un intervalo general

La serie de Fourier de la Definición 9.1 se aplica a funciones definidas en el intervalo $[-\pi, \pi]$. Si queremos representar de la misma forma una función definida en otro

intervalo $[a, b]$ es fácil trasladar y reescalar las funciones básicas que usamos a ese otro intervalo: las funciones

$$\frac{1}{2}, \quad \cos\left(\frac{n\pi(2x - b - a)}{b - a}\right), \quad \sin\left(\frac{n\pi(2x - b - a)}{b - a}\right), \quad (n \geq 1) \quad (111)$$

resultan de componer las funciones que hemos usado en (100) con la transformación afín

$$\phi : [a, b] \rightarrow [-\pi, \pi] \quad (112)$$

$$\phi(x) := \frac{\pi(2x - b - a)}{b - a}, \quad x \in [a, b], \quad (113)$$

que lleva el intervalo $[a, b]$ en el $[-\pi, \pi]$. La serie de Fourier en el intervalo $[a, b]$ es entonces

$$\frac{a_0}{2} + \sum_{n \geq 1} a_n \cos\left(\frac{n\pi(2x - b - a)}{b - a}\right) + \sum_{n \geq 1} b_n \sin\left(\frac{n\pi(2x - b - a)}{b - a}\right). \quad (114)$$

Es fácil ver que los coeficientes tienen que ser

$$\begin{aligned} a_n &:= \frac{2}{b - a} \int_a^b f(x) \cos\left(\frac{n\pi(2x - b - a)}{b - a}\right) dx, & (n \geq 0), \\ b_n &:= \frac{2}{b - a} \int_a^b f(x) \sin\left(\frac{n\pi(2x - b - a)}{b - a}\right) dx, & (n \geq 1). \end{aligned} \quad (115)$$

Ejemplo 9.6. El caso en que el intervalo sea $[0, L]$ la serie correspondiente es

$$\frac{a_0}{2} + \sum_{n \geq 1} a_n \cos\left(\frac{2\pi nx}{L} - n\pi\right) + \sum_{n \geq 1} b_n \sin\left(\frac{2\pi nx}{L} - n\pi\right) \quad (116)$$

con coeficientes dados por

$$\begin{aligned} a_n &:= \frac{2}{b - a} \int_a^b f(x) \cos\left(\frac{2\pi nx}{L} - n\pi\right) dx, & (n \geq 0), \\ b_n &:= \frac{2}{b - a} \int_a^b f(x) \sin\left(\frac{2\pi nx}{L} - n\pi\right) dx, & (n \geq 1). \end{aligned} \quad (117)$$

Observa que

$$\cos\left(\frac{2\pi nx}{L} - n\pi\right) = \begin{cases} -\cos\left(\frac{2\pi nx}{L}\right), & n \text{ impar} \\ \cos\left(\frac{2\pi nx}{L}\right), & n \text{ par} \end{cases} \quad (118)$$

y la igualdad correspondiente es también cierta para el seno, así que para simplificar se suele escribir (116) en la forma

$$\frac{\tilde{a}_0}{2} + \sum_{n \geq 1} \tilde{a}_n \cos\left(\frac{2\pi nx}{L}\right) + \sum_{n \geq 1} \tilde{b}_n \sin\left(\frac{2\pi nx}{L}\right) \quad (119)$$

con coeficientes dados por

$$\begin{aligned} \tilde{a}_n &:= \frac{2}{L} \int_a^b f(x) \cos\left(\frac{2\pi nx}{L}\right) dx, & (n \geq 0), \\ \tilde{b}_n &:= \frac{2}{L} \int_a^b f(x) \sin\left(\frac{2\pi nx}{L}\right) dx, & (n \geq 1). \end{aligned} \quad (120)$$

El resultado de la suma es el mismo, ya que lo único que hemos hecho es cambiar el signo de las funciones que usamos y de sus coeficientes.

9.3. Series de Fourier en senos y cosenos

Podemos también representar cualquier función en $[0, \pi]$ como una suma *sólo de senos*. La *serie de Fourier en senos* de una función integrable $f : [0, \pi] \rightarrow \mathbb{R}$ es la serie

$$\sum_{n \geq 1} b_n \sin(nx) \quad (121)$$

con coeficientes

$$b_n := \frac{2}{\pi} \int_0^\pi f(x) \sin(nx) dx. \quad (122)$$

Observa que ahora el intervalo que estamos usando es $[0, \pi]$, no $[-\pi, \pi]$. Por supuesto, podemos hacer lo mismo en cualquier otro intervalo usando el mismo cambio de variable que en la sección 9.2. Puedes ver que la serie (121) es la misma que la serie de Fourier en $[-\pi, \pi]$ de la *extensión impar* de f :

$$\tilde{f}(x) := \operatorname{sgn}(x) f(|x|), \quad x \in [-\pi, \pi], \quad (123)$$

donde $\operatorname{sgn}(x)$ es la función signo,

$$\operatorname{sgn}(x) := \begin{cases} -1 & \text{si } x < 0 \\ 0 & \text{si } x = 0 \\ 1 & \text{si } x > 0. \end{cases} \quad (124)$$

Si calculamos la serie de Fourier usual de \tilde{f} en $[-\pi, \pi]$ los coeficientes de a_n se anulan para todo $n \geq 0$, así que sólo quedan los términos en b_n , que son

$$b_n = \frac{1}{\pi} \int_{-\pi}^\pi \tilde{f}(x) \sin(nx) dx = \frac{2}{\pi} \int_0^\pi f(x) \sin(nx) dx, \quad (125)$$

lo mismo que en (122). Esto demuestra que se cumple un teorema muy parecido al Teorema 9.2:

Teorema 9.7 (Convergencia de la serie de Fourier en senos). *Sea $f \in L^2([0, \pi])$. Entonces la serie (121), con coeficientes dados por (122), converge en $L^2([0, \pi])$ a la función f :*

$$\sum_{n=1}^m b_n \sin(nx) \xrightarrow{m \rightarrow +\infty} f \quad \text{en } L^2([0, \pi]). \quad (126)$$

A la inversa, si (126) se cumple para ciertos coeficientes $\{b_i\}_{i \geq 1}$ entonces los b_i tienen que ser iguales a los definidos en (122).

De forma parecida podemos definir la *serie de Fourier en cosenos* de cualquier función $f : [0, \pi] \rightarrow \mathbb{R}$ integrable: es la serie

$$\frac{a_0}{2} + \sum_{n \geq 1} a_n \cos(nx) \quad (127)$$

con coeficientes

$$a_n := \frac{2}{\pi} \int_0^\pi f(x) \cos(nx) dx. \quad (128)$$

La serie (121) es la misma que la serie de Fourier en $[-\pi, \pi]$ de la *extensión par* de f :

$$\tilde{f}(x) := f(|x|), \quad x \in [-\pi, \pi]. \quad (129)$$

Y por supuesto tenemos también un teorema de convergencia:

Teorema 9.8 (Convergencia de la serie de Fourier en cosenos). Sea $f \in L^2([0, \pi])$. Entonces la serie (127), con coeficientes dados por (128), converge en $L^2([0, \pi])$ a la función f :

$$\frac{a_0}{2} + \sum_{n=1}^m a_n \cos(nx) \xrightarrow{m \rightarrow +\infty} f \quad \text{en } L^2([0, \pi]). \quad (130)$$

A la inversa, si (130) se cumple para ciertos coeficientes $\{a_i\}_{i \geq 0}$ entonces los a_i tienen que ser iguales a los definidos en (128).

9.4. Otras representaciones de las series de Fourier

A veces es más cómodo escribir la serie de Fourier (100) de otra forma. Usando que

$$\cos(nx) = \frac{e^{inx} + e^{-inx}}{2}, \quad \sin(nx) = \frac{e^{inx} - e^{-inx}}{2i},$$

la serie (100) se puede poner como

$$\begin{aligned} f(x) &= \frac{a_0}{2} + \sum_{n \geq 1} a_n \cos(nx) + \sum_{n \geq 1} b_n \sin(nx) \\ &= \frac{a_0}{2} + \sum_{n \geq 1} a_n \frac{e^{inx} + e^{-inx}}{2} + \sum_{n \geq 1} b_n \frac{e^{inx} - e^{-inx}}{2i} \\ &= \frac{a_0}{2} + \sum_{n \geq 1} e^{inx} \frac{1}{2} (a_n - ib_n) + \sum_{n \geq 1} e^{-inx} \frac{1}{2} (a_n + ib_n) \\ &= \frac{a_0}{2} + \sum_{n \geq 1} e^{inx} \frac{1}{2} (a_n - ib_n) + \sum_{n \leq -1} e^{inx} \frac{1}{2} (a_{-n} + ib_{-n}) = \sum_{n \in \mathbb{Z}} \hat{f}_n e^{inx}, \end{aligned}$$

donde

$$\hat{f}_n := \begin{cases} \frac{1}{2}(a_{-n} + ib_{-n}) & n < 0, \\ \frac{1}{2}a_0, & n = 0, \\ \frac{1}{2}(a_n - ib_n) & n > 0. \end{cases}$$

Es decir: dada una función $f: [-\pi, \pi] \rightarrow \mathbb{R}$, podemos escribirla como

$$f(x) = \sum_{n \in \mathbb{Z}} \hat{f}_n e^{inx}, \quad (131)$$

donde los coeficientes \hat{f}_n están definidos por

$$\hat{f}_n = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{-inx} f(x) dx. \quad (132)$$

La igualdad en (131) es cierta en el sentido de la convergencia en $L^2([-\pi, \pi])$:

$$\sum_{n=-m}^m \hat{f}_n e^{inx} \xrightarrow{m \rightarrow +\infty} f \quad \text{en } L^2([-\pi, \pi]).$$

Observa que la suma desde $-m$ hasta m es una función real, ya que los términos para n y $-n$ son uno conjugado del otro, y la parte imaginaria se cancela siempre. En realidad, la serie (131) se puede escribir también para funciones complejas $f: [-\pi, \pi] \rightarrow \mathbb{C}$, y entonces la convergencia es cierta en el espacio $L^2([-\pi, \pi], \mathbb{C})$.

Teorema 9.9 (Identidad de Parseval, segunda representación). Sea $f \in L^1([-\pi, \pi])$ cualquiera. La suma $\sum_{n \in \mathbb{Z}} |\hat{f}_n|^2$ es finita si y sólo si $f \in L^2([-\pi, \pi])$, y en ese caso se tiene

$$\sum_{n \in \mathbb{Z}} |\hat{f}_n|^2 = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x)^2 dx. \quad (133)$$

9.5. Series de Fourier y derivadas

Lema 9.10 (Serie de Fourier de la derivada de una función). Sea $f : [-\pi, \pi] \rightarrow \mathbb{R}$ una función absolutamente continua en $[-\pi, \pi]$, con $f(-\pi) = f(\pi)$. Se cumple que

$$\widehat{(f')} = in \hat{f}_n \quad \text{para todo } n \in \mathbb{Z}. \quad (134)$$

Demostración. Dado que f es absolutamente continua, su derivada existe en casi todo punto y $f' \in L^1([-\pi, \pi])$. Esto nos permite integrar por partes en la expresión de los coeficientes de f' :

$$\widehat{(f')} = \frac{1}{2\pi} \int_{-\pi}^{\pi} f'(x) e^{-inx} dx = \frac{in}{2\pi} \int_{-\pi}^{\pi} f(x) e^{-inx} dx = in \hat{f}_n,$$

donde el término de frontera desaparece debido a la condición $f(-\pi) = f(\pi)$. \square

Definición 9.11 (El espacio de Sobolev H^k). Sea $a < b \in \mathbb{R}$. Para $k \geq 1$ entero definimos el espacio $H^k([a, b])$ como

$$H^k([a, b]) := \left\{ f \in \mathcal{C}^{k-1}([a, b]) \mid \begin{array}{l} f^{(k-1)} \text{ es absolutamente continua} \\ \text{y } f^{(k)} \text{ está en } L^2([a, b]) \end{array} \right\}. \quad (135)$$

Definimos el espacio $H_{\text{per}}^k([a, b])$ como

$$H_{\text{per}}^k([a, b]) := \left\{ f \in H^k([a, b]) \mid f^{(j)}(a) = f^{(j)}(b), j = 1, \dots, k-1 \right\}. \quad (136)$$

Para $k = 0$ denotamos

$$H^0([a, b]) = H_{\text{per}}^0([a, b]) = L^2([a, b]).$$

El espacio $H_{\text{per}}^k([a, b])$ es simplemente el espacio $H^k([a, b])$ con condiciones periódicas en el borde.

Teorema 9.12 (Diferenciabilidad de una serie de Fourier). Sea $k \geq 0$ y $f : [-\pi, \pi] \rightarrow \mathbb{R}$ una función integrable. Son equivalentes:

1. Existe una función $g \in H_{\text{per}}^k([-\pi, \pi])$ tal que $f = g$ en casi todo punto de $[-\pi, \pi]$.
2. Se cumple que

$$\sum_{n \in \mathbb{Z}} n^{2k} |\hat{f}_n|^2 < +\infty. \quad (137)$$

En caso de que se cumpla alguna de las equivalencias anteriores, para $0 \leq j \leq k-1$ la serie de Fourier de $f^{(j)}$ converge uniformemente en $[-\pi, \pi]$, y la serie de Fourier de f se puede derivar término a término k veces en $[-\pi, \pi]$ (la derivación término a término es cierta en todo $x \in [-\pi, \pi]$ para las $k-1$ primeras derivadas; es cierta en casi todo $x \in [-\pi, \pi]$ para la derivada de orden k).

Demostración. Es fácil completar la demostración por inducción si tenemos el caso $k = 0, 1$. Para $k = 0$, el enunciado es exactamente el del Teorema 9.9. Para $k = 1$, supongamos primero que $f = g$ en casi todo punto, con $g \in H_{\text{per}}^1([-\pi, \pi])$. Por supuesto, los coeficientes de Fourier de f son los mismos que los de g . El Lema 9.10 nos dice que los coeficientes de Fourier de g' son inf_n , y como $g' \in L^2([-\pi, \pi])$ sabemos que

$$\sum_{n \in \mathbb{Z}} n^2 |\hat{f}_n|^2 < +\infty.$$

Para la otra implicación, supongamos que esta suma es finita. Esto implica que la serie de funciones

$$\sum_{n \in \mathbb{Z}} inf_n e^{inx}$$

converge en $L^2([-\pi, \pi])$ a cierta función $h \in L^2([-\pi, \pi])$, y en particular en $L^1([-\pi, \pi])$. El Teorema G.8 sobre derivación de series de funciones dice entonces que la serie

$$\sum_{n \in \mathbb{Z}} \hat{f}_n e^{inx}$$

converge uniformemente a una función g absolutamente continua cuya derivada (en casi todo punto) es igual a h . Como $h \in L^2([-\pi, \pi])$, tenemos que $g \in H^1([-\pi, \pi])$. Como $f(x) = \sum_{n \in \mathbb{Z}} \hat{f}_n e^{inx}$ en casi todo $x \in [-\pi, \pi]$, tenemos que $f = g$ en casi todo punto de $[-\pi, \pi]$. Sólo nos queda por demostrar que g cumple $g(-\pi) = g(\pi)$, lo cual es consecuencia de

$$g(-\pi) = \sum_{n \in \mathbb{Z}} \hat{f}_n e^{-in\pi} = \sum_{n \in \mathbb{Z}} \hat{f}_n e^{in\pi} = g(\pi). \quad \square$$

Teorema 9.13 (Diferenciabilidad de una serie de Fourier y sumabilidad de los coeficientes). *Sea $f : [-\pi, \pi] \rightarrow \mathbb{R}$ una función integrable y $k \geq 0$.*

1. *Supongamos que*

$$\sum_{n \in \mathbb{Z}} n^k |\hat{f}_n| < +\infty. \quad (138)$$

Entonces f es igual en casi todo punto a una función $g \in \mathcal{C}^k([-\pi, \pi])$ con $g^{(j)}(-\pi) = g^{(j)}(\pi)$ para todo $j = 0, \dots, k$.

2. *Supongamos que $f \in H^{k+1}([-\pi, \pi])$, y que $f^{(j)}(-\pi) = f^{(j)}(\pi)$ para todo $j = 0, \dots, k$. Entonces*

$$\sum_{n \in \mathbb{Z}} n^k |\hat{f}_n| < +\infty. \quad (139)$$

Demostración. Punto 1. Es suficiente probarlo para $k = 0, 1$, ya que a partir de esto es fácil obtener el caso general por inducción. Para $k = 0$, el hecho de que $\sum_{n \in \mathbb{Z}} |\hat{f}_n| < +\infty$ implica que la serie de Fourier de f converge uniforme-absolutamente a una función g , que en particular debe ser continua (ya que es límite uniforme de funciones continuas). Como el límite de la serie de Fourier de f es igual a f en el sentido de $L^2([-\pi, \pi])$, tenemos que $f = g$ en casi todo punto. Observa también que $g(-\pi) = g(\pi)$ por definición de la serie de Fourier.

Para el caso $k = 1$, consideramos la serie de Fourier con coeficientes $in\hat{f}_n$,

$$\sum_{n \in \mathbb{Z}} in\hat{f}_n e^{inx}. \quad (140)$$

Esta serie debería ser la serie de Fourier de f' , pero no podemos asegurarlo porque todavía no sabemos que f sea derivable. Lo que sí sabemos por hipótesis es que esta serie converge uniforme-absolutamente, ya que

$$\sum_{n \geq k} |in\hat{f}_n e^{inx}| = \sum_{n \geq k} n|\hat{f}_n| \rightarrow 0 \quad \text{cuando } k \rightarrow +\infty.$$

Por el Teorema G.7 de derivación de series de funciones, la serie de Fourier de f converge a una función \mathcal{C}^1 en $[-\pi, \pi]$ (a la cual llamamos g) y su derivada viene dada por el límite de la serie (140). Por el mismo motivo que antes, $f = g$ en casi todo punto de $[-\pi, \pi]$. La convergencia puntual de la serie de f y la de f' en $[-\pi, \pi]$ demuestra que $g(\pi) = g(-\pi)$, $g'(\pi) = g'(-\pi)$, ya que $e^{in\pi} = e^{-in\pi}$ para todo n .

Punto 2. Si $f \in H^{k+1}([-\pi, \pi])$ el Teorema 9.12 implica que

$$\sum_{n \in \mathbb{Z}} n^{2(k+1)} |\hat{f}_n|^2 < +\infty. \quad (141)$$

Usando la desigualdad de Cauchy-Schwarz,

$$\sum_{n \in \mathbb{Z}} |n^k \hat{f}_n| \leq \left(\sum_{n \in \mathbb{Z}} n^{2k} (|n| + 1)^2 |\hat{f}_n|^2 \right)^{\frac{1}{2}} \left(\sum_{n \in \mathbb{Z}} (|n| + 1)^{-2} \right)^{\frac{1}{2}}. \quad (142)$$

Los dos últimos paréntesis están acotados: el último es una serie sumable y para el primero podemos usar que

$$n^{2k} (|n| + 1)^2 = n^{2(k+1)} + 2|n|^{2k+1} + n^{2k} \leq 4n^{2(k+1)}, \quad (143)$$

con lo que (141) demuestra que la serie del primer paréntesis también es sumable. \square

Ejercicio 9.3. Completa el paso de inducción en la demostración del resultado anterior.

9.5.1. Para la serie de Fourier en senos

Lema 9.14. Sea $k \geq 1$, $f \in H^k([0, \pi])$. Definimos la extensión antisimétrica de f a $[-\pi, \pi]$ como

$$\tilde{f}(x) = \begin{cases} f(x) & \text{si } x \in [0, \pi], \\ -f(-x) & \text{si } x \in [-\pi, 0]. \end{cases} \quad (144)$$

Entonces $\tilde{f} \in H^k([-\pi, \pi])$ si y sólo si

$$f^{(j)}(0) = 0 \text{ para todo } j \text{ par con } 0 \leq j \leq k-1. \quad (145)$$

Demostración. Notamos que para $j = 1, \dots, k-1$,

$$\tilde{f}_+^{(j)}(0) = f_+^{(j)}(0), \quad \tilde{f}_-^{(j)}(0) = (-1)^{j+1} f_+^{(j)}(0). \quad (146)$$

Esto demuestra que

$$\tilde{f}_+^{(j)}(0) = \tilde{f}_-^{(j)}(0), \quad j = 0, \dots, k-1, \quad (147)$$

si y sólo si (145) se cumple. Esto es una condición necesaria para que $\tilde{f} \in H^k([-\pi, \pi])$. Por otra parte, si (145) se cumple deducimos que $\tilde{f} \in \mathcal{C}^{k-1}([-\pi, \pi])$. Como $f^{(k-1)}$ es absolutamente continua en $[0, \pi]$ deducimos que $\tilde{f}^{(k-1)}$ lo es en $[-\pi, \pi]$, y como $f \in L^2([0, \pi])$ vemos que $\tilde{f} \in L^2([-\pi, \pi])$. \square

Teorema 9.15 (Diferenciabilidad de la serie de Fourier en senos y decaimiento de los coeficientes). *Sea $k \geq 1$, $f : [0, \pi] \rightarrow \mathbb{R}$ una función integrable y $\{b_n\}_{n \geq 1}$ sus coeficientes de Fourier en senos. Entonces son equivalentes:*

1. $f \in H^k([0, \pi])$ y $f^{(j)}(0) = f^{(j)}(\pi) = 0$ para todo j par con $0 \leq j \leq k-1$.
2. Se cumple que

$$\sum_{n=1}^{\infty} n^{2k} |b_n|^2 < +\infty. \quad (148)$$

Teorema 9.16 (Diferenciabilidad de la serie de Fourier en senos y sumabilidad de los coeficientes). *Sea $k \geq 1$, $f : [0, \pi] \rightarrow \mathbb{R}$ una función integrable y $\{b_n\}_{n \geq 1}$ sus coeficientes de Fourier en senos.*

1. Supongamos que

$$\sum_{n=1}^{\infty} n^k |b_n| < +\infty. \quad (149)$$

Entonces $f = g$ en casi todo $x \in [0, \pi]$ para cierta función $g \in \mathcal{C}^k([0, \pi])$ con $g^{(j)}(0) = g^{(j)}(\pi) = 0$ para todo j par con $0 \leq j \leq k$.

2. Supongamos que $f \in H^{k+1}([0, \pi])$ para cierto $k \geq 1$, y que $f^{(j)}(0) = f^{(j)}(\pi) = 0$ para todo j par con $0 \leq j \leq k$. Entonces

$$\sum_{n=1}^{\infty} n^k |b_n| < +\infty. \quad (150)$$

9.5.2. Para la serie de Fourier en cosenos

Lema 9.17. *Sea $k \geq 1$, $f \in H^k([0, \pi])$. Definimos la extensión simétrica de f a $[-\pi, \pi]$ como*

$$\tilde{f}(x) = f(|x|), \quad x \in [-\pi, \pi]. \quad (151)$$

Entonces $\tilde{f} \in H^k([-\pi, \pi])$ si y sólo si

$$f^{(j)}(0) = 0 \text{ para todo } j \text{ impar con } 1 \leq j \leq k-1. \quad (152)$$

Demostración. Notamos que para $j = 1, \dots, k-1$,

$$\tilde{f}_+^{(j)}(0) = f_+^{(j)}(0), \quad \tilde{f}_-^{(j)}(0) = (-1)^j f_+^{(j)}(0). \quad (153)$$

Esto demuestra que

$$\tilde{f}_+^{(j)}(0) = \tilde{f}_-^{(j)}(0), \quad j = 1, \dots, k-1, \quad (154)$$

si y sólo si (152) se cumple. Esto es una condición necesaria para que $\tilde{f} \in H^k([-\pi, \pi])$. Por otra parte, si (152) se cumple deducimos que $\tilde{f} \in \mathcal{C}^{k-1}([-\pi, \pi])$. Como $f^{(k-1)}$ es absolutamente continua en $[0, \pi]$ deducimos que \tilde{f} lo es en $[-\pi, \pi]$, y como $f \in L^2([0, \pi])$ vemos que $\tilde{f} \in L^2([-\pi, \pi])$. \square

Teorema 9.18 (Diferenciabilidad de la serie de Fourier en cosenos y decaimiento de los coeficientes). *Sea $k \geq 1$ y $f : [0, \pi] \rightarrow \mathbb{R}$ una función integrable y $\{a_n\}_{n \geq 0}$ sus coeficientes de Fourier en cosenos. Entonces son equivalentes:*

1. $f \in H^k([0, \pi])$ y $f^{(j)}(0) = f^{(j)}(\pi) = 0$ para todo j impar con $0 \leq j \leq k-1$.
2. Se cumple que

$$\sum_{n=0}^{\infty} n^{2k} |a_n|^2 < +\infty. \quad (155)$$

9.6. Convergencia puntual y uniforme de series de Fourier

Teorema 9.19. *Sea $f \in H^1([-\pi, \pi])$ tal que $f(-\pi) = f(\pi)$. Entonces la serie de Fourier de f converge uniforme-absolutamente a f en $[-\pi, \pi]$.*

Demostración. En realidad esto ya lo mencionamos el Teorema 9.13, salvo el hecho de que la convergencia es uniforme-absoluta. Las colas de la serie de Fourier de f en $x \in [-\pi, \pi]$ pueden acotarse usando la desigualdad de Cauchy-Schwarz: para $m \geq 1$,

$$\sum_{|n| \geq m} |\hat{f}_n| |e^{inx}| \leq \left(\sum_{|n| \geq m} n^2 |\hat{f}_n|^2 \right)^{\frac{1}{2}} \left(\sum_{|n| \geq m} n^{-2} \right)^{\frac{1}{2}}. \quad (156)$$

El último paréntesis es una serie convergente, y el primer paréntesis converge a cero cuando $m \rightarrow +\infty$ debido al Teorema 9.12. \square

Teorema 9.20 (Convergencia de la serie de Fourier en cosenos). *Sea $f \in H^1([0, \pi])$. Entonces la serie de Fourier en cosenos de f converge uniforme-absolutamente a f en $[0, \pi]$.*

Teorema 9.21 (Convergencia de la serie de Fourier en senos). *Sea $f \in H^1([0, \pi])$ una función tal que $f(0) = f(\pi) = 0$. Entonces la serie de Fourier en senos de f converge uniforme-absolutamente a f en $[0, \pi]$.*

9.7. Ejercicios

Ejercicio 9.4. Calcula la serie de Fourier en $[-\pi, \pi]$ de la función $f : [-\pi, \pi] \rightarrow \mathbb{R}$ en cada uno de los siguientes casos:

1. $f(x) = x$, $x \in [-\pi, \pi]$.
2. $f(x) = |x|$, $x \in [-\pi, \pi]$.
3. $f(x) = -1$ para $-\pi \leq x \leq 0$, $f(x) = 1$ para $0 < x \leq \pi$.

Ejercicio 9.5. Consideramos la función $f : [0, \pi] \rightarrow \mathbb{R}$, $f(x) = x$ para $x \in [0, \pi]$.

1. Calcula su serie de Fourier en senos.
2. Calcula su serie de Fourier en cosenos.

Ejercicio 9.6. Calcula la serie de Fourier de la función $f(x) = x^2$, definida en $[-\pi, \pi]$.

1. Demuestra que la serie de Fourier de f converge uniformemente a f en $[-\pi, \pi]$.
2. Evaluando la serie para $x = \pi$ calcula el valor de

$$\sum_{n=1}^{\infty} \frac{1}{n^2}. \quad (157)$$

Dicho valor es la función zeta de Riemann evaluada en $z = 2$, normalmente denotado por $\zeta(2)$.

10. La ecuación de ondas

10.1. Derivación de la ecuación

La ecuación de ondas es una ecuación en derivadas parciales que sirve como modelo para el comportamiento de ondas de muchos tipos. Se aplica por ejemplo a la propagación de ondas de sonido en el aire, de vibraciones en una cuerda tensa o de ondas electromagnéticas, y es una ecuación básica que forma parte de muchos fenómenos más complicados. El caso que vamos a estudiar en esta sección es uno de los más sencillos: las vibraciones en una cuerda.

Supongamos que tenemos una cuerda sujeta por sus dos extremos a dos puntos fijos que están separados por una distancia L . Suponemos que la cuerda está tensa, de forma que cuando está en equilibrio describe una recta entre sus dos extremos. Puedes imaginar por ejemplo la cuerda de una guitarra. En cierto momento separamos la cuerda de su posición de equilibrio y la soltamos (por ejemplo, punteando la guitarra), y queremos saber cuál es el movimiento que describe. Si describimos la cuerda según la mecánica clásica (que es suficiente para lo que vamos a hacer) entonces su movimiento viene determinado por las leyes de Newton. La dificultad es que el número de partículas que conforman la cuerda es muy grande y no podemos calcular las fuerzas que actúan sobre cada una de ellas y escribir una ecuación diferencial para cada una. Aquí es muy útil el principio variacional en mecánica que estudiamos en la sección 6: las leyes de Newton son equivalentes a las ecuaciones de Euler-Lagrange de un problema variacional asociado. De acuerdo con este principio, para calcular las ecuaciones del movimiento tenemos que calcular la energía potencial U y la energía cinética T de la cuerda, y luego escribir las ecuaciones de Euler-Lagrange correspondientes al funcional $L = T - U$.

Empecemos por describir la posición de la cuerda como una función $y : [0, +\infty) \times [0, L] \rightarrow \mathbb{R}$, de manera que la forma de la cuerda en el momento t es la gráfica de la función $x \mapsto y(t, x)$, con t fijo. Visto de otra forma, la trayectoria que describe la partícula x viene dada por la curva $t \mapsto y(t, x)$, con x fijo. Elegimos las coordenadas de forma que los extremos estén fijados en $(0, 0)$ y $(L, 0)$, así que

$$y(t, 0) = y(t, L) = 0.$$

La energía cinética es fácil de escribir: es la suma de la energía cinética de cada partícula:

$$T = \frac{m}{2} \int_0^L (\partial_t y(t, x))^2 dx, \quad (158)$$

donde m es la densidad de masa de la cuerda (en kilos por metro, por ejemplo). Para la energía potencial vamos a despreciar la gravedad (para el ejemplo de la guitarra, las cuerdas vibran igual independientemente de si la guitarra está de pie o tumbada). Sólo vamos a tener en cuenta la energía potencial elástica de la cuerda, que suponemos que es el efecto más importante. La energía potencial depende de cómo de estirada está la cuerda. Si la cuerda estuviese recta, entonces la energía potencial sería

$$U = \frac{1}{2}KL^2, \quad (159)$$

que es la expresión de la energía potencial de un muelle estirado a una longitud L con constante de restitución K . Pero la cuerda no está recta, así que debemos aplicar esto localmente en cada sección pequeña de la cuerda. La distancia ℓ entre la partícula que está en $(x, y(x))$ y la que está en $(x + h, y(x + h))$ es

$$\ell^2 = (y(t, x + h) - y(t, x))^2 + h^2, \quad (160)$$

o

$$\frac{\ell^2}{h^2} = \left(\frac{y(t, x + h) - y(t, x)}{h} \right)^2 + 1 \approx (\partial_x y(t, x))^2 + 1. \quad (161)$$

Si consideramos el segmento de cuerda entre x y $x + h$ su energía es por lo tanto proporcional a

$$\frac{\ell^2}{h^2} = (\partial_x y(t, x))^2 + 1, \quad (162)$$

y la energía elástica total es proporcional a

$$\int_0^L ((\partial_x y(t, x))^2 + 1) dx. \quad (163)$$

Como podemos elegir el punto cero de la energía potencial (podemos añadirle o quitarle una constante si queremos), usaremos

$$U := \frac{k}{2} \int_0^L (\partial_x y(t, x))^2 dx, \quad (164)$$

donde k es la *densidad de energía elástica* de la cuerda (por ejemplo, en julios por metro). De (158) y (164) obtenemos al final que

$$T - U = \frac{m}{2} \int_0^L (\partial_t y(t, x))^2 dx - \frac{k}{2} \int_0^L (\partial_x y(t, x))^2 dx. \quad (165)$$

Tenemos entonces un funcional que depende de una función $y : [0, +\infty) \times [0, L]$. Podemos escribir las ecuaciones de Euler-Lagrange correspondientes como vimos en la sección 1. Escribimos el funcional como

$$\mathcal{F}[y] = T - U = \int_0^L F(t, x, y(t, x), \partial_t y(t, x), \partial_x y(t, x)) dt \quad (166)$$

donde F es la función

$$F(t, x, y, z, w) = \frac{m}{2}z^2 - \frac{k}{2}w^2.$$

Sabemos que las ecuaciones de Euler-Lagrange son

$$\partial_y F = \operatorname{div}_{(t,x)} (\nabla_{(z,w)} F), \quad (167)$$

donde las derivadas de F están siempre evaluadas en $(t, x, y(t, x), \partial_t y(t, x), \partial_x y(t, x))$. Si escribimos la ecuación completamente, esto es

$$\begin{aligned} \partial_y F(t, x, y(t, x), \partial_t y(t, x), \partial_x y(t, x)) \\ = \operatorname{div}_{(t,x)} (\nabla_{(z,w)} F(t, x, y(t, x), \partial_t y(t, x), \partial_x y(t, x))) . \end{aligned} \quad (168)$$

Vamos a calcular esto: para empezar, $\partial_y F = 0$, así que el miembro izquierdo de (167) es 0. Para el miembro derecho necesitamos

$$\partial_z F = mz, \quad \partial_w F = -kw, \quad (169)$$

luego

$$\nabla_{(z,w)} F = (mz, -kw) \quad (170)$$

y evaluando en el punto adecuado,

$$\nabla_{(z,w)} F(t, x, y(t, x), \partial_t y(t, x), \partial_x y(t, x)) = (m\partial_t y(t, x), -k\partial_x y(t, x)). \quad (171)$$

Calculamos su divergencia:

$$\begin{aligned} \operatorname{div}_{(t,x)} (\nabla_{(z,w)} F(t, x, y(t, x), \partial_t y(t, x), \partial_x y(t, x))) \\ = \operatorname{div}_{(t,x)} ((m\partial_t y(t, x), -k\partial_x y(t, x))) = m\partial_{tt} y(t, x) - k\partial_{xx} y(t, x). \end{aligned} \quad (172)$$

Por tanto, las ecuaciones de Euler-Lagrange asociadas al funcional $\mathcal{F} = T - U$ son

$$m\partial_{tt} y(t, x) = k\partial_{xx} y(t, x), \quad (173)$$

o escritas de forma más compacta,

$$m\partial_{tt} y = k\partial_{xx} y. \quad (174)$$

La ecuación en derivadas parciales (174) se conoce como *ecuación de ondas*. En esta sección queremos ver si es posible resolverla y obtener información sobre el comportamiento de la cuerda.

Podemos hacer un razonamiento muy parecido en dimensiones mayores. Por ejemplo, en dimensión 2 consideramos una membrana sujeta en los bordes de una región acotada $\Omega \subseteq \mathbb{R}^d$ (imaginamos la membrana de un tambor), descrita por una función $y: [0, +\infty) \times \bar{\Omega} \rightarrow \mathbb{R}$ de forma que $y(t, x)$ es la altura sobre la posición $x \in \bar{\Omega}$ en tiempo t . La energía elástica de un trozo pequeño de la membrana es proporcional a su área, y haciendo un cálculo como el de antes no es difícil llegar a que y debe cumplir

$$m\partial_{tt} y = k\Delta y, \quad (175)$$

donde $\Delta y = \sum_{i=1}^d \partial_{x_i}^2 y$ es el Laplaciano de y .

10.2. Propiedades básicas

Vamos a estudiar con más detenimiento la ecuación en derivadas parciales

$$\partial_t^2 y = \Delta y, \quad (176)$$

donde la incógnita $y = y(t, x)$ es una función que depende de $t \in I$, $x \in \Omega$, con $I \subseteq \mathbb{R}$ un intervalo abierto y Ω un dominio de \mathbb{R}^d . Normalmente decimos que t representa el tiempo, y que x representa el espacio. Decimos que la ecuación está planteada en Ω cuando queremos destacar el dominio Ω , y decimos que está planteada en dimensión d para destacar que $\Omega \subseteq \mathbb{R}^d$ (por ejemplo, en la sección anterior obtuvimos la ecuación (174), que está planteada en dimensión 1).

Las condiciones naturales a imponer son las *condiciones iniciales*

$$y(0, x) = g(x), \quad \partial_t y(0, x) = h(x), \quad x \in \Omega, \quad (177)$$

donde $g, h: \Omega \rightarrow \mathbb{R}$ son funciones dadas. Normalmente imponemos también la *condición de frontera*

$$y(t, x) = 0, \quad t \in I, \quad x \in \partial\Omega. \quad (178)$$

Concepto de solución Aunque el concepto de solución de la ecuación es bastante obvio por ahora, es bueno enunciarlo explícitamente. Para las EDPs en general hay varios tipos de soluciones (más adelante veremos por ejemplo *soluciones débiles*) y es importante destacar que por ahora estamos trabajando sólo con *soluciones clásicas*:

Definición 10.1. Sea $I \subseteq \mathbb{R}$ un intervalo abierto, $\Omega \subseteq \mathbb{R}^d$ un abierto. Una *solución* (o *solución clásica*) en $I \times \Omega$ de la ecuación de ondas (176) es una función $y: I \times \Omega \rightarrow \mathbb{R}$ de clase \mathcal{C}^2 en $I \times \Omega$, tal que

$$\partial_t^2 y(t, x) = \Delta y(t, x) \quad \text{para todo } (t, x) \in I \times \Omega.$$

Decimos que una solución y de (176) cumple la condición inicial (177) cuando $0 \in \bar{I}$, las funciones $y, \partial_t y$ tienen una extensión continua a $(I \cup \{0\}) \times \Omega$ (a las que llamamos también $y, \partial_t y$), y dicha extensión continua cumple (177).

Decimos que una solución y de (176) cumple la condición de frontera (178) cuando y tiene una extensión continua a $I \times \bar{\Omega}$ (a la que llamamos también y), y dicha extensión continua cumple (178).

A veces consideramos la ecuación $m\partial_t^2 y = k\Delta y$ con ciertas constantes $m, k > 0$, o ecuaciones similares. La definición de solución es la misma en cada caso, con las modificaciones evidentes.

Principio de superposición Una de las características más importantes de la ecuación (176) es que es una ecuación en derivadas parciales *lineal*. Esto quiere decir que si y_1, y_2 son soluciones de (176), entonces λy_1 es también una solución para cualquier $\lambda \in \mathbb{R}$, y la suma $y_1 + y_2$ es también una solución.

Soluciones particulares Hay muchas soluciones sencillas de la ecuación (176). Si $I \subseteq \mathbb{R}$ es cualquier intervalo abierto y $\Omega \subseteq \mathbb{R}^d$ es un abierto entonces cualquier función afín en (t, x) definida en $I \times \Omega$ es solución.

Reversibilidad La ecuación (176) es *reversible en tiempo*: si y es una solución en $I \times \Omega$, entonces

$$z(t, x) := y(-t, x)$$

es una solución en $(-I) \times \Omega$. Si pensamos en la solución y como en una “película” que muestra el comportamiento de la onda, el cambio $t \mapsto -t$ corresponde a “reproducir la película hacia atrás”. El hecho de que también es solución significa que el proceso que vemos al invertir el tiempo es también un proceso físico posible. Más adelante veremos ecuaciones donde esto no es cierto, como la ecuación del calor.

Traslación en tiempo La ecuación de ondas (176) es *autónoma*: no tiene coeficientes que dependan de t . Gracias a eso es fácil ver que si y es una solución en $I \times \Omega$, entonces para cualquier $\tau \in \mathbb{R}$ la función

$$z(t, x) := y(t - \tau, x)$$

es solución en $(I + \tau) \times \Omega$. Esta solución es una traslación en tiempo de y .

Cambio de escala Si y es una solución de (176) en $I \times \Omega$, entonces

$$z(t, x) := y(\lambda t, \mu x)$$

es una solución de la ecuación

$$\mu^2 \partial_t^2 z = \lambda^2 \Delta z$$

en $(\frac{1}{\lambda}I) \times (\frac{1}{\mu}\Omega)$. Este tipo de cambio de variable se llama *cambio de escala*. Observa en particular que para encontrar soluciones de la ecuación

$$m \partial_t^2 y = k \Delta y$$

con $m, k > 0$ constantes, es suficiente hacerlo para $m = k = 1$, ya que siempre podemos reducirnos a este caso con un cambio de escala.

10.3. Soluciones en \mathbb{R} : fórmula de d'Alembert

Vamos a buscar soluciones de

$$\left\{ \begin{array}{ll} \partial_t^2 y = \partial_x^2 y, & (t, x) \in \mathbb{R}^2, \\ y(0, x) = g(x), & x \in \mathbb{R}, \\ \partial_t y(0, x) = h(x), & x \in \mathbb{R} \end{array} \right. \quad \begin{array}{l} (179a) \\ (179b) \\ (179c) \end{array}$$

en $\mathbb{R} \times \mathbb{R}$, donde $f, g: \mathbb{R} \rightarrow \mathbb{R}$ son funciones continuas dadas. Para esto hay una idea que consiste en escribir

$$\partial_t^2 y - \partial_x^2 y = (\partial_t + \partial_x)(\partial_t - \partial_x)y = (\partial_t - \partial_x)(\partial_t + \partial_x)y.$$

Esto nos sugiere que las funciones que cumplan

$$(\partial_t + \partial_x)y = 0 \quad \text{o bien} \quad (\partial_t - \partial_x)y = 0$$

son también soluciones de la ecuación de ondas. Dichas funciones son fáciles de encontrar: son aquellas de la forma

$$y(t, x) = F(x - t) \quad \text{o bien} \quad y(t, x) = G(x + t),$$

donde F, G son funciones que tomamos de clase \mathcal{C}^2 (para asegurarnos de que y es una solución de la ecuación de ondas). Estas soluciones son *ondas viajeras*: son funciones que tienen siempre la misma forma, para t fijo, y que se mueven trasladándose hacia la izquierda o hacia la derecha. Como la ecuación de ondas es lineal, la función

$$y(t, x) = F(x - t) + G(x + t)$$

es también una solución. Para resolver (179) nos falta encontrar F, G adecuadas de forma que se cumplan (179b) y (179c). Es decir,

$$F(x) + G(x) = g(x), \quad -F'(x) + G'(x) = h(x), \quad x \in \mathbb{R}.$$

Integrando la segunda ecuación obtenemos

$$G(x) - F(x) = H(x),$$

donde H es una primitiva de h . Junto con la primera ecuación,

$$G(x) = \frac{1}{2}(g(x) + H(x)), \quad F(x) = \frac{1}{2}(g(x) - H(x)).$$

Por tanto,

$$\begin{aligned} y(t, x) = F(x - t) + G(x + t) &= \frac{1}{2}(g(x - t) + g(x + t)) + \frac{1}{2}(H(x + t) - H(x - t)) \\ &= \frac{1}{2}(g(x - t) + g(x + t)) + \frac{1}{2} \int_{x-t}^{x+t} h(\tau) \, d\tau. \end{aligned}$$

Esta última expresión se llama *fórmula de d'Alembert*. Puedes comprobar el siguiente resultado:

Proposición 10.2 (Fórmula de d'Alembert para la ecuación de ondas). *Sean $g, h: \mathbb{R} \rightarrow \mathbb{R}$ con g de clase \mathcal{C}^2 , h de clase \mathcal{C}^1 . La función $y: \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$ definida por*

$$y(t, x) = \frac{1}{2}(g(x - t) + g(x + t)) + \frac{1}{2} \int_{x-t}^{x+t} h(\tau) \, d\tau, \quad (t, x) \in \mathbb{R} \times \mathbb{R}$$

es una solución de (179).

10.4. Separación de variables en dimensión 1

Se conoce como *método de separación de variables* a la idea de buscar soluciones particulares de una ecuación en derivadas parciales que sean de la forma

$$y(t, x) = \phi(t)\psi(x). \tag{180}$$

Intentemos buscar soluciones de esta forma para la ecuación de ondas (174), suponiendo como antes que los extremos están fijos:

$$y(t, 0) = y(t, L) = 0. \quad (181)$$

Si sustituimos (180) en (174) obtenemos

$$m\phi''(t)\psi(x) = k\phi(t)\psi''(x) \quad (182)$$

así que (al menos en los puntos (t, x) donde $\phi(t) \neq 0$, $\psi(x) \neq 0$)

$$m \frac{\phi''(t)}{\phi(t)} = k \frac{\psi''(x)}{\psi(x)}. \quad (183)$$

Los dos miembros de esta ecuación dependen de variables distintas, así que tienen que ser constantes. Encontramos entonces que

$$m\phi''(t) = -\lambda\phi(t), \quad k\psi''(x) = -\lambda\psi(x) \quad (184)$$

para cierta constante λ que no conocemos (el signo negativo es arbitrario; está para simplificar un poco los cálculos que vamos a hacer después). Por otra parte la condición de contorno (181) implica que

$$\psi(0) = \psi(L) = 0. \quad (185)$$

Vemos entonces que ψ satisface un problema de contorno de segundo orden. Por supuesto, estamos interesados en las soluciones no triviales de la ecuación de ondas, así que sólo nos interesan las soluciones ψ que no sean siempre 0. Este problema es precisamente un problema de Sturm-Liouville; es uno de los casos más sencillos y puede resolverse explícitamente. Las únicas soluciones no triviales posibles son

$$\psi_n(x) = C \sin\left(\frac{\pi nx}{L}\right), \quad (186)$$

para cualquier $n \geq 1$ dado, asociadas respectivamente a

$$\lambda = \frac{k\pi^2 n^2}{L^2}. \quad (187)$$

Para estos valores de λ las soluciones ϕ que satisfacen (184) son

$$\phi_n(t) = A_n \cos\left(\sqrt{\frac{k}{m}} \frac{\pi nt}{L}\right) + B_n \sin\left(\sqrt{\frac{k}{m}} \frac{\pi nt}{L}\right), \quad (188)$$

donde A_n , B_n son constantes reales libres. Resumiendo lo anterior, hemos encontrado la siguiente familia de soluciones de (174):

$$y(t, x) = \sin\left(\frac{\pi nx}{L}\right) \left(A_n \cos\left(\sqrt{\frac{k}{m}} \frac{\pi nt}{L}\right) + B_n \sin\left(\sqrt{\frac{k}{m}} \frac{\pi nt}{L}\right) \right). \quad (189)$$

Estas soluciones se llaman a veces *modos de vibración* de una cuerda. Observa que la frecuencia (en tiempo) de estas soluciones es $n\omega$, donde

$$\omega := \frac{\sqrt{k}}{2L\sqrt{m}}.$$

En particular, todas las frecuencias son un múltiplo entero de la *frecuencia fundamental* ω .

10.5. Separación de variables en dimensión 2

Podemos intentar la misma idea para la ecuación (175) en dimensión $d \geq 2$. Suponiendo que $y(t, x) = \phi(t)\psi(x)$ (ahora con $x \in \Omega$) obtenemos

$$m\phi''(t) = -\lambda\phi(t), \quad k\Delta\psi = -\lambda\psi.$$

La segunda ecuación es un tipo de “problema de Sturm-Liouville para una ecuación en derivadas parciales”. No tenemos solución explícita para un dominio general Ω pero sigue siendo cierto un resultado análogo al Teorema 7.23: el conjunto de $\lambda \in \mathbb{R}$ para los que existe una solución de $k\Delta\psi = -\lambda\psi$ (los valores propios) forman una sucesión discreta que tiende a $+\infty$, y si tomamos una base de cada espacio propio obtenemos una base de Hilbert de $L^2(\Omega)$.

Un caso en el que sí podemos encontrar la solución es aquel en el que Ω es un cuadrado en \mathbb{R}^2 . Si tomamos

$$\Omega := (0, L) \times (0, L)$$

podemos buscar soluciones explícitas en variables separadas de la ecuación

$$\begin{cases} \partial_t^2 y = \Delta y, & (t, x) \in \mathbb{R} \times (0, L)^2, \\ y(t, x_1, 0) = y(t, x_1, L) = 0, & x_1 \in [0, L], \\ y(t, 0, x_2) = y(t, L, x_2) = 0, & x_2 \in [0, L]. \end{cases} \quad \begin{matrix} (190a) \\ (190b) \\ (190c) \end{matrix}$$

Buscamos soluciones particulares de la forma

$$y(t, x_1, x_2) = \phi(t)\psi_1(x_1)\psi_2(x_2).$$

La ecuación (190a) se traduce en

$$\phi''(t)\psi_1(x_1)\psi_2(x_2) = \phi(t)\psi_1''(x_1)\psi_2(x_2) + \phi(t)\psi_1(x_1)\psi_2''(x_2).$$

En los puntos donde no se anulan $\phi(t), \psi_1(x_1), \psi_2(x_2)$ tenemos

$$\frac{\phi''(t)}{\phi(t)} = \frac{\psi_1''(x_1)}{\psi_1(x_1)} + \frac{\psi_2''(x_2)}{\psi_2(x_2)}.$$

Esto nos dice que deben existir constantes $\lambda, \mu \in \mathbb{R}$ tales que

$$\begin{aligned} \phi''(t) &= -\lambda\phi(t), \\ \psi_1''(x_1) &= -\mu\psi_1(x_1), \\ \psi_2''(x_2) &= -(\lambda - \mu)\psi_2(x_2). \end{aligned}$$

Las condiciones de frontera (190b)–(190c) se traducen en

$$\psi_1(0) = \psi_1(L) = \psi_2(0) = \psi_2(L) = 0,$$

y por lo tanto las funciones ψ_1, ψ_2 son solución de un problema de Sturm-Liouville, precisamente el problema del ejercicio 7.5. Sabemos entonces que las únicas posibilidades son

$$\mu = \frac{n^2\pi^2}{L^2}, \quad \lambda - \mu = \frac{m^2\pi^2}{L^2}$$

para enteros $m, n \geq 1$. Es decir,

$$\mu = \frac{n^2 \pi^2}{L^2}, \quad \lambda = \frac{(m^2 + n^2) \pi^2}{L^2}.$$

Las funciones propias correspondientes son (salvo una constante multiplicativa)

$$\begin{aligned} \psi_1(x_1) &= \text{sen} \left(\frac{n\pi x_1}{L} \right), \\ \psi_2(x_2) &= \text{sen} \left(\frac{m\pi x_2}{L} \right). \end{aligned}$$

Y la función ϕ correspondiente a m, n es

$$\phi(t) = \text{sen} \left(\frac{\pi t \sqrt{m^2 + n^2}}{L} \right).$$

Las soluciones finales que obtenemos son entonces

$$y(t, x_1, x_2) = \text{sen} \left(\frac{\pi t \sqrt{m^2 + n^2}}{L} \right) \text{sen} \left(\frac{n\pi x_1}{L} \right) \text{sen} \left(\frac{m\pi x_2}{L} \right).$$

Observa que la frecuencia en tiempo de la función anterior es

$$\frac{\sqrt{m^2 + n^2}}{2L}.$$

La frecuencia fundamental (la más baja) es $(L\sqrt{2})^{-1}$, y no es cierto que las otras frecuencias (los *armónicos*) sean múltiplos enteros suyos.

Ejercicio 10.1. Dados $L_1, L_2 > 0$, encuentra las soluciones en variables separadas del problema

$$\begin{cases} \partial_t^2 y = \Delta y, & (t, x_1, x_2) \in \mathbb{R} \times (0, L_1) \times (0, L_2), \\ y(t, x_1, 0) = y(t, x_1, L_2) = 0, & x_1 \in [0, L_1], \\ y(t, 0, x_2) = y(t, L_1, x_2) = 0, & x_2 \in [0, L_2], \end{cases}$$

definidas en $\mathbb{R} \times [0, L_1] \times [0, L_2]$.

10.6. Problema de valores iniciales para la ecuación de ondas

Hemos encontrado algunas soluciones de la ecuación de ondas, pero el problema que nos gustaría resolver en general es el siguiente: si conocemos el estado de la cuerda en cierto momento, ¿podemos predecir cuál será en momentos posteriores? Esto se conoce como *problema de valores iniciales* o *problema de Cauchy*. Dado que estamos calculando la evolución en tiempo de un sistema físico que sigue las leyes de Newton es razonable pensar que necesitamos conocer la posición y velocidad iniciales de cada una de las partículas que lo forman. Con la interpretación que hacíamos al principio, necesitamos conocer la posición inicial de cada partícula,

$$f(x) := y(0, x), \quad x \in [0, L], \quad (191)$$

y la velocidad inicial de cada una,

$$g(x) := \partial_t y(0, x), \quad x \in [0, L]. \quad (192)$$

Dadas estas dos funciones f , g , ¿podemos encontrar una solución de la ecuación de ondas que cumpla (191)–(192)? Ésta es la pregunta que intentamos contestar ahora.

Las soluciones particulares (189) que hemos encontrado cumplen

$$y(0, x) = A_n \sin\left(\frac{\pi n x}{L}\right), \quad (193)$$

$$\partial_t y_n(0, x) = \sqrt{\frac{k}{m}} \frac{\pi n}{L} B_n \sin\left(\frac{\pi n x}{L}\right). \quad (194)$$

Usando el principio de superposición podemos obtener muchas otras soluciones a partir de éstas: cualquier combinación lineal de ellas es también una solución. Pongamos nombre a algunas de estas soluciones:

$$y_n(t, x) := \sin\left(\frac{\pi n x}{L}\right) \cos\left(\sqrt{\frac{k}{m}} \frac{\pi n t}{L}\right), \quad (195)$$

$$z_n(t, x) := \sin\left(\frac{\pi n x}{L}\right) \sin\left(\sqrt{\frac{k}{m}} \frac{\pi n t}{L}\right). \quad (196)$$

Si consideramos por ejemplo

$$y(t, x) := \sum_{n=1}^k A_n y_n(t, x) + \sum_{n=1}^k B_n z_n(t, x) \quad (197)$$

entonces y es una solución que cumple

$$y(0, x) = \sum_{n=1}^k A_n \sin\left(\frac{\pi n x}{L}\right), \quad (198)$$

$$\partial_t y(0, x) = \sqrt{\frac{k}{m}} \sum_{n=1}^k \frac{\pi n}{L} B_n \sin\left(\frac{\pi n x}{L}\right). \quad (199)$$

Sabemos que podemos aproximar cualquier función por medio de sumas como las anteriores: lo que necesitamos es encontrar la *serie de Fourier en senos* de las funciones f , g . Tenemos también que tener cuidado al pasar al límite, ya que en realidad necesitamos considerar la suma de una cantidad infinita de soluciones, y demostrar que es también una solución. Esto es lo que hacemos en el siguiente resultado:

Teorema 10.3. *Sea $L > 0$, $f \in H^3([0, L])$, $g \in H^2([0, L])$ tales que $f(0) = f(L) = f''(0) = f''(L) = g(0) = g(L) = 0$, y sean b_n , c_n los coeficientes de sus series de Fourier en senos:*

$$f = \sum_{n=1}^{\infty} b_n \sin\left(\frac{\pi n x}{L}\right), \quad g = \sum_{n=1}^{\infty} c_n \sin\left(\frac{\pi n x}{L}\right). \quad (200)$$

(La igualdad se entiende como límite en $L^2([0, L])$.) Entonces el problema

$$\begin{cases} \partial_{tt}y = \partial_{xx}y, & (t, x) \in (0, +\infty) \times (0, L) \\ y(0, x) = f(x), & x \in (0, L) \\ \partial_t y(0, x) = g(x), & x \in (0, L) \\ y(t, 0) = y(t, L) = 0, & t \in [0, +\infty) \end{cases} \quad \begin{matrix} (201) \\ (202) \\ (203) \\ (204) \end{matrix}$$

tiene una solución dada por

$$y(t, x) := \sum_{n=1}^{\infty} b_n y_n(t, x) + \sum_{n=1}^{\infty} \frac{c_n L}{\pi n} z_n(t, x), \quad (t, x) \in \mathbb{R} \times [0, L]. \quad (205)$$

(Las funciones y_n, z_n se definieron en (195), (196).)

Demostración. Lo primero que debemos hacer es probar que la serie (205) converge puntualmente, y que su límite y es una función de clase \mathcal{C}^2 en $[0, +\infty) \times [0, L]$. Para esto usamos el Corolario G.10 sobre derivación de series de funciones: comprobamos que

$$|b_n y_n(t, x)| \leq |b_n|, \quad \left| \frac{c_n L}{\pi n} z_n(t, x) \right| \leq C_1 \frac{|c_n|}{n}, \quad n \geq 1. \quad (206)$$

Gracias al Teorema 9.16 sabemos que

$$\sum_{n=1}^{\infty} n^2 |b_n| < +\infty, \quad \sum_{n=1}^{\infty} n |c_n| < +\infty, \quad (207)$$

lo cual junto con (206) implica que la serie que define y es uniformemente convergente en $\mathbb{R} \times \mathbb{R}$ (aunque en la definición de y sólo se menciona el conjunto $\mathbb{R} \times [0, L]$). De forma parecida podemos acotar las derivadas de cada uno de los términos de (205):

$$|b_n \partial_x y_n(t, x)| \leq \frac{\pi n}{L} |b_n|, \quad \left| \frac{c_n L}{\pi n} \partial_x z_n(t, x) \right| \leq \sqrt{\frac{k}{m}} |c_n|, \quad (208)$$

lo cual demuestra de nuevo usando (207) que la serie de la derivada ∂_x término a término converge uniformemente en $\mathbb{R} \times \mathbb{R}$. De forma parecida obtenemos que para cierta $C_2 > 0$ independiente de n ,

$$|b_n \partial^\alpha y_n(t, x)| \leq C_2 n^2 |b_n|, \quad n \geq 1, \quad (209)$$

$$\left| \frac{c_n L}{\pi n} \partial^\alpha z_n(t, x) \right| \leq C_2 n |c_n|, \quad (210)$$

para toda derivada parcial ∂^α de orden menor o igual que 2. Deducimos usando (207) que la serie de las derivadas término a término convergen uniformemente, para todas las derivadas parciales de orden menor o igual que 2. El Corolario G.10 nos asegura entonces que y es una función de clase \mathcal{C}^2 en $\mathbb{R} \times \mathbb{R}$, y en particular en $\mathbb{R} \times [0, L]$. El Corolario G.10 también asegura que podemos calcular las derivadas parciales de y derivando la serie término a término, por lo que y es una solución de la ecuación en derivadas parciales. La función y cumple la condición inicial en todo punto por el Teorema 9.6, y puede comprobarse directamente que cumple la condición $y(t, 0) = y(t, L) = 0$ para todo $t \in \mathbb{R}$. \square

10.7. Ejercicios

Ejercicio 10.2. Si $u = u(t, x)$ es una solución de la ecuación de ondas definida en $\mathbb{R} \times \mathbb{R}$ que además cumple que

$$u(t, x) = -u(t, -x), \quad t \in \mathbb{R}, x \geq 0, \quad (211)$$

demuestra que u restringida al conjunto $\mathbb{R} \times [0, +\infty)$ es solución del problema

$$\begin{cases} \partial_t^2 u = \partial_x^2 u & \text{en } \mathbb{R} \times (0, +\infty) \\ u(t, 0) = 0 & \text{para } t \in \mathbb{R}. \end{cases} \quad (212)$$

Ejercicio 10.3. Sea u es una solución de la ecuación de ondas en $[0, +\infty) \times \mathbb{R}$ con datos iniciales f, g :

$$\begin{cases} \partial_t^2 u = \partial_x^2 u & \text{en } (0, +\infty) \times \mathbb{R} \\ u(0, x) = f(x) & \text{para } x \in \mathbb{R}. \\ \partial_t u(0, x) = g(x) & \text{para } x \in \mathbb{R}. \end{cases} \quad (213)$$

Supongamos que f, g son funciones en $\mathcal{C}^2(\mathbb{R})$ con soporte contenido en $[-1, 1]$. Demuestra que para cada $t \geq 0$ la función $x \mapsto u(t, x)$ tiene soporte contenido en $[-1-t, 1+t]$.

Ejercicio 10.4 (Equipartición de la energía). Sea $u = u(t, x)$ una solución de la ecuación de ondas en $[0, +\infty) \times \mathbb{R}$ con datos iniciales f, g :

$$\begin{cases} \partial_t^2 u = \partial_x^2 u & \text{en } (0, +\infty) \times \mathbb{R} \\ u(0, x) = f(x) & \text{para } x \in \mathbb{R}. \\ \partial_t u(0, x) = g(x) & \text{para } x \in \mathbb{R}. \end{cases} \quad (214)$$

Supongamos que f, g son funciones \mathcal{C}^2 con soporte compacto definidas en \mathbb{R} . Definimos la energía cinética $T = T(t)$ y la potencial $U = U(t)$ como al principio de la sección 10. Demuestra que

1. $T(t) + U(t)$ es una constante independiente de t , para todo $t \geq 0$,
2. $T(t) = U(t)$ para todo t suficientemente grande.

Ejercicio 10.5. Estamos intentando construir un instrumento de cuerda con cuerdas de un cierto material, con una tensión fija. Ajustamos la tensión de forma que una cuerda de 50 centímetros de longitud produce un sonido que fijamos como la nota La. ¿Qué longitud deben tener otras seis cuerdas para completar toda la escala desde Do hasta Si?

11. Las ecuaciones de Laplace y Poisson

La ecuación de Laplace es la ecuación en derivadas parciales

$$\Delta u = 0, \quad (215)$$

normalmente planteada en un abierto $\Omega \subseteq \mathbb{R}^d$. En muchas situaciones es natural imponer alguna condición en la frontera del abierto Ω (a no ser que $\Omega = \mathbb{R}^d$). Una de las más comunes son las condiciones de frontera *de tipo Dirichlet*,

$$u|_{\partial\Omega} = g, \quad (216)$$

donde g es una función dada, definida en $\partial\Omega$. La correspondiente ecuación con una función f a la derecha,

$$\Delta u = f \quad (217)$$

se llama *ecuación de Poisson*. A la función f la llamamos a veces *término de fuente*. A continuación precisamos lo que entendemos por solución de este problema:

Definición 11.1. Sea $\Omega \subseteq \mathbb{R}^d$ un abierto, $g: \partial\Omega \rightarrow \mathbb{R}$ y $f: \Omega \rightarrow \mathbb{R}$ funciones continuas. Una *solución clásica* (o simplemente solución) de la ecuación de Poisson

$$\begin{cases} \Delta u = f & \text{en } \Omega, \\ u = g & \text{en } \partial\Omega \end{cases} \quad (218)$$

es una función $u \in \mathcal{C}^2(\Omega) \cap \mathcal{C}(\overline{\Omega})$ que satisface (218).

En particular, cuando $\Omega = \mathbb{R}^d$ no hay frontera y una solución es simplemente una función en $\mathcal{C}^2(\mathbb{R}^d)$ que cumple $\Delta u = f$ en \mathbb{R}^d .

Forma variacional de la ecuación de Laplace Consideramos el problema de minimizar el funcional

$$\mathcal{F}(u) := \int_{\Omega} |\nabla u(x)|^2 dx,$$

definido en el dominio

$$D := \{u \in \mathcal{C}^1(\Omega) \cap \mathcal{C}(\overline{\Omega}) \mid u|_{\partial\Omega} = g\}.$$

Podemos escribir las ecuaciones de Euler-Lagrange correspondientes, que en este caso son

$$2 \operatorname{div}_x(\nabla u(x)) = 0 \quad \text{en } \Omega,$$

es decir,

$$\Delta u(x) = 0 \quad \text{en } \Omega.$$

Los puntos críticos de este funcional satisfacen precisamente la ecuación de Laplace.

11.1. La ecuación de Laplace

11.1.1. Propiedades básicas

Linealidad. Invarianza por movimientos rígidos. Soluciones particulares.

Proposición 11.2. Sea $\Omega \subseteq \mathbb{C}$ un abierto. Si $f: \Omega \rightarrow \mathbb{C}$ es una función holomorfa en Ω entonces sus partes real e imaginaria $u := \Re(f)$, $v := \Im(f)$ cumplen

$$\Delta u = \Delta v = 0 \quad \text{en } \Omega. \quad (219)$$

11.1.2. Solución fundamental

Como la ecuación de Laplace es simétrica por rotaciones, tiene sentido buscar una solución que tenga el mismo tipo de simetría. Si suponemos que u es de la forma

$$u(x) = \psi(|x|)$$

podemos encontrar la ecuación que debe cumplir ψ . Tenemos, denotando $|x| = r$, y para $i = 1, \dots, d$,

$$\begin{aligned}\partial_{x_i} u(x) &= \psi'(r) \frac{x_i}{|x|}, \\ \partial_{x_i}^2 u(x) &= \psi''(r) \frac{x_i^2}{|x|^2} + \psi'(r) \frac{1}{|x|} - \psi'(r) \frac{x_i^2}{|x|^3}.\end{aligned}$$

Sumando desde $i = 1$ hasta $i = d$,

$$\Delta u(x) = \psi''(r) + \frac{d-1}{r} \psi'(r).$$

La ecuación de Laplace implica por tanto que

$$\psi''(r) + \frac{d-1}{r} \psi'(r) = 0,$$

que podemos resolver escribiendo $\phi := \psi'$ para obtener

$$\phi(r) = C \exp(-(d-1) \log r) = Cr^{-(d-1)},$$

y finalmente

$$\psi(r) = Kr^{-(d-2)}, \quad u(x) = \begin{cases} K|x|^{-(d-2)} & \text{para } d \neq 2, \\ K \log |x| & \text{para } d = 2, \end{cases}$$

donde $K \in \mathbb{R}$ es una constante. La solución que hemos encontrado es diferenciable en todo \mathbb{R}^d menos en $x = 0$. En realidad, el mismo razonamiento nos sirve para ver que no hay otras soluciones con simetría rotacional: todas tienen una singularidad en $x = 0$ (salvo la solución trivial $u = 0$). Se suele hacer la siguiente elección de la constante K , por el motivo que vamos a ver en la siguiente sección:

Definición 11.3. Sea $d \geq 1$ un entero. La función

$$\Psi: \mathbb{R}^d \setminus \{0\} \rightarrow \mathbb{R}, \quad \Psi(x) = \begin{cases} \frac{1}{(d-2)\omega_d} |x|^{-(d-2)} & \text{para } d \neq 2, \\ -\frac{1}{2\pi} \log |x| & \text{para } d = 2, \end{cases}$$

(donde ω_d denota la superficie de la esfera unidad en \mathbb{R}^d) se llama *solución fundamental de la ecuación de la Laplace*. A veces denotamos la constante de la expresión anterior por K_d ,

$$K_d = \begin{cases} \frac{1}{(d-2)\omega_d} & \text{para } d \neq 2, \\ -\frac{1}{2\pi} & \text{para } d = 2. \end{cases}$$

11.1.3. Armónicos esféricos en \mathbb{R}^3

En lugar de buscar soluciones invariantes por rotaciones, podemos buscar soluciones de la ecuación de Laplace en variables separadas, considerando coordenadas radiales. El caso más conocido es el de \mathbb{R}^3 , donde podemos considerar las coordenadas esféricas (r, θ, φ) (Sección B.4), y buscar una función armónica de la forma

$$u(x) = R(r)\Theta(\theta)\Phi(\varphi). \quad (220)$$

El caso de la sección anterior, donde hemos encontrado la solución fundamental, corresponde a tomar $\Theta = \Phi = 1$; hemos visto que en ese caso las únicas soluciones son $R(r) = C_1 + C_2/r$, y de entre ellas las únicas armónicas en todo \mathbb{R}^3 son las constantes (el caso $C_2 = 0$).

El laplaciano en coordenadas esféricas se escribe (ver Sección B.4.1) como

$$\Delta u = \frac{1}{r^2} \partial_r (r^2 \partial_r u) + \frac{1}{r^2 (\sin \varphi)^2} \partial_\theta^2 u + \frac{1}{r^2 \sin \varphi} \partial_\varphi (\sin \varphi \partial_\varphi u).$$

Si u es una función armónica en \mathbb{R}^3 , usando (220) tenemos que

$$0 = \Delta u = \Theta \Phi \frac{1}{r^2} \partial_r (r^2 R') + R \Phi \frac{1}{r^2 (\sin \varphi)^2} \Theta'' + R \Theta \frac{1}{r^2 \sin \varphi} \partial_\varphi (\sin \varphi \Phi').$$

Multiplicando por $r^2/(R\Theta\Phi)$ tenemos

$$0 = \Delta u = \frac{1}{R} \partial_r (r^2 R') + \frac{1}{\Theta (\sin \varphi)^2} \Theta'' + \frac{1}{\Phi \sin \varphi} \partial_\varphi (\sin \varphi \Phi').$$

Por tanto, para cierto $\lambda \in \mathbb{R}$,

$$(r^2 R')' = \lambda R, \quad (221)$$

$$\frac{1}{\Theta (\sin \varphi)^2} \Theta'' + \frac{1}{\Phi \sin \varphi} \partial_\varphi (\sin \varphi \Phi') = \lambda. \quad (222)$$

Las funciones $Y(\theta, \varphi) = \Theta(\theta)\Phi(\varphi)$ que resuelven la ecuación (222) se llaman *armónicos esféricos*.

11.2. La ecuación de Poisson en \mathbb{R}^d

11.2.1. Solución de la ecuación de Poisson en \mathbb{R}^d

Usando la solución fundamental de la ecuación de Laplace podemos encontrar una solución del problema de Poisson

$$\Delta u = f \quad \text{en } \mathbb{R}^d,$$

con f una función dada. La solución fundamental Ψ cumple $\Delta \Psi = 0$ en $\mathbb{R}^d \setminus \{0\}$. La idea es que, para $y \in \mathbb{R}^d$ dado, las funciones

$$\Psi(\cdot - y): \mathbb{R}^d \setminus \{y\} \rightarrow \mathbb{R}, \quad \Psi(x - y) = K_d |x - y|^{-(d-2)}$$

son también soluciones de $\Delta u = 0$ en $\mathbb{R}^d \setminus \{y\}$. Como la ecuación de Laplace es lineal, uno esperaría que la “combinación lineal” de soluciones dada por

$$u(x) = \int_{\mathbb{R}^d} f(y) \Psi(x - y) dy = \int_{\mathbb{R}^d} f(x - y) \Psi(y) dy \quad (223)$$

sea también una solución de la ecuación de Laplace. Sin embargo, la singularidad de Ψ hace que en realidad esta función u tenga un laplaciano que no es 0, y que de hecho es igual a una constante por f . Vamos a ver esto. Supongamos que f es una función de clase \mathcal{C}^2 y de soporte compacto en \mathbb{R}^d . Para calcular el laplaciano de la función u definida por la primera integral en (223) tenemos el problema de que no podemos derivar bajo la integral, ya que la singularidad de Ψ hace que no se cumplan las condiciones que necesitamos. Usamos para eso la segunda integral:

$$\Delta u(x) = \int_{\mathbb{R}^d} \Delta f(x - y) \Psi(y) dy. \quad (224)$$

Queríamos usar la fórmula de Green para pasar el Laplaciano a la función Ψ , pero no podemos hacerlo directamente, de nuevo porque $\Delta \Psi(y)$ no es integrable cerca de $y = 0$. Para evitar eso partimos la integral en dos partes:

$$\begin{aligned} \Delta u(x) &= \int_{\mathbb{R}^d} \Delta f(x - y) \Psi(y) dy \\ &= \int_{\mathbb{R}^d \setminus B_\epsilon} \Delta f(x - y) \Psi(y) dy + \int_{B_\epsilon} \Delta f(x - y) \Psi(y) dy = I_1 + I_2. \end{aligned}$$

La segunda integral tiende a 0 cuando $\epsilon \rightarrow 0$, ya que

$$|I_2| \leq \|\Delta f\|_\infty \int_{B_\epsilon} |\Psi(y)| dy \xrightarrow{\epsilon \rightarrow 0} 0.$$

En cuanto a la primera, podemos usar la fórmula de Green y obtener

$$I_1 = \int_{\mathbb{R}^d \setminus B_\epsilon} \nabla f(x - y) \nabla \Psi(y) dy - \int_{S_\epsilon} \Psi(y) \nabla f(x - y) \cdot \eta(y) dS(y) = J_1 + J_2,$$

donde $\eta(y)$ es el normal *interior* a B_ϵ (el exterior a $\mathbb{R}^d \setminus B_\epsilon$). (Observa que el signo ha cambiado porque $\Delta f(x - y) = (\operatorname{div}(\nabla f))(x - y) = -\operatorname{div}_y(\nabla f(x - y))$.) De nuevo, J_2 tiende a 0 cuando $\epsilon \rightarrow 0$:

$$\begin{aligned} \left| \int_{S_\epsilon} \Psi(y) \nabla f(x - y) \cdot \eta(y) dS(y) \right| &\leq \|\nabla f\|_\infty \int_{S_\epsilon} |\Psi(y)| dS(y) \\ &= K_d \|\nabla f\|_\infty \epsilon^{-(d-2)} \int_{S_\epsilon} dS(y) = K_d |S_1| \|\nabla f\|_\infty \epsilon^{-(d-2)} \epsilon^{d-1} = \text{const.} \cdot \epsilon. \end{aligned}$$

Para J_1 podemos de nuevo usar la fórmula de Green:

$$\begin{aligned} J_1 &= \int_{\mathbb{R}^d \setminus B_\epsilon} \nabla f(x - y) \nabla \Psi(y) dy = - \int_{\mathbb{R}^d \setminus B_\epsilon} \nabla_y (f(x - y)) \nabla \Psi(y) dy \\ &= \int_{\mathbb{R}^d \setminus B_\epsilon} f(x - y) \Delta \Psi(y) dy - \int_{S_\epsilon} f(x - y) \nabla \Psi(y) \cdot \eta(y) dS(y) \\ &= - \int_{S_\epsilon} f(x - y) \nabla \Psi(y) \cdot \eta(y) dS(y), \end{aligned}$$

donde la otra integral desaparece debido a que $\Delta\Psi(y) = 0$ para $y \in \mathbb{R}^d \setminus B_\epsilon$. Para calcular esta última integral usamos que

$$\eta(y) = -\frac{y}{|y|}, \quad \nabla\Psi(y) = \begin{cases} -K_d(d-2)|y|^{-d}y & (d \neq 2), \\ -K_d\frac{y}{|y|^2} & (d = 2), \end{cases}$$

luego

$$\nabla\Psi(y) \cdot \eta(y) = \begin{cases} K_d(d-2)|y|^{-(d-1)} & (d \neq 2), \\ \frac{K_d}{|y|} & (d = 2), \end{cases}$$

y tenemos, cuando $d \neq 2$,

$$\begin{aligned} - \int_{S_\epsilon} f(x-y) \nabla\Psi(y) \cdot \eta(y) \, dS(y) &= -K_d(d-2) \int_{S_\epsilon} f(x-y) |y|^{-(d-1)} \, dS(y) \\ &= -K_d(d-2) \epsilon^{-(d-1)} \int_{S_\epsilon} f(x-y) \, dS(y) \\ &= -K_d(d-2) |S_1| \int_{S_\epsilon} f(x-y) \, dS(y) \xrightarrow{\epsilon \rightarrow 0} -K_d(d-2) |S_1| f(x) = -f(x). \end{aligned}$$

Y cuando $d = 2$,

$$\begin{aligned} - \int_{S_\epsilon} f(x-y) \nabla\Psi(y) \cdot \eta(y) \, dS(y) &= -K_d \int_{S_\epsilon} f(x-y) |y|^{-1} \, dS(y) \\ &= -K_d \epsilon^{-1} \int_{S_\epsilon} f(x-y) \, dS(y) \\ &= -K_d |S_1| \int_{S_\epsilon} f(x-y) \, dS(y) \xrightarrow{\epsilon \rightarrow 0} -K_d |S_1| f(x) = -f(x). \end{aligned}$$

Tomando el límite cuando $\epsilon \rightarrow 0$ finalmente obtenemos que

$$-\Delta u(x) = f(x) \quad \text{para todo } x \in \mathbb{R}^d.$$

Acabamos de demostrar entonces lo siguiente:

Teorema 11.4. *Sea $f \in \mathcal{C}(\mathbb{R}^d)$ con soporte compacto. Entonces la función*

$$u(x) = \int_{\mathbb{R}^d} f(y) \Psi(x-y) \, dy \tag{225}$$

es de clase \mathcal{C}^2 en \mathbb{R}^d y cumple

$$-\Delta u(x) = f(x) \quad \text{para todo } x \in \mathbb{R}^d.$$

11.2.2. Una interpretación física: potencial gravitatorio y electrostático

Sabemos que el potencial gravitatorio causado por una masa M en el punto $x_0 \in \mathbb{R}^3$ es

$$V(x) = -\frac{MG}{|x - x_0|},$$

donde $G = 6.67408 \times 10^{-11} \text{ m}^3 \cdot \text{kg}^{-1} \cdot \text{s}^{-2}$ es la constante gravitatoria. A partir de esto, el campo creado en el punto $x \in \mathbb{R}^3$ es

$$E(x) = -\nabla V(x) = -\frac{MG(x-x_0)}{|x-x_0|^3} = \frac{MG}{|x-x_0|^2} \sigma,$$

donde $\sigma = \frac{x_0-x}{|x-x_0|}$ es el vector unitario en la dirección de la línea desde x hasta x_0 . Esta última expresión es tal vez la forma más familiar, donde se ve la dependencia del inverso del cuadrado de la distancia entre los puntos. Por supuesto, la *fuerza* sobre un cuerpo con masa m situado en $x \in \mathbb{R}^d$ es

$$F(x) = -m\nabla V(x) = \frac{mMG}{|x-x_0|^2} \sigma.$$

Ahora, en lugar de una masa (ideal) concentrada en un punto $x_0 \in \mathbb{R}^d$, supongamos que tenemos una distribución de masa dada por una función $\rho: \mathbb{R}^d \rightarrow [0, +\infty)$, de forma que

$$\int_{\Omega} \rho(x) \, dx \equiv \text{masa en la región } \Omega \subseteq \mathbb{R}^d.$$

El potencial creado por la distribución de masa ρ es el que resulta de *sumar* las contribuciones de cada uno de sus puntos:

$$V(x) = - \int_{\mathbb{R}^d} \frac{G}{|x-y|} \rho(y) \, dy = - \int_{\mathbb{R}^d} \frac{G}{|y|} \rho(x-y) \, dy.$$

Esto es igual que la expresión (225), salvo un factor constante:

$$V(x) = -\frac{G}{K_3} \int_{\mathbb{R}^d} \Psi(x-y) \rho(y) \, dy. \quad (226)$$

Teniendo en cuenta que $K_3^{-1} = 4\pi$ (el área de la esfera unidad) por el Teorema 11.4 tenemos (suponiendo que ρ es continua y de soporte compacto) que

$$\Delta V(x) = 4\pi G \rho(x).$$

Es decir: el potencial gravitatorio generado por una distribución de masa ρ es una solución de la ecuación de Poisson con término independiente $4\pi G \rho$.

Por supuesto, el mismo razonamiento funciona también para las fuerzas electrostáticas porque su forma es la misma. El potencial eléctrico causado por una carga Q en el punto $x_0 \in \mathbb{R}^3$ es

$$V(x) = \frac{1}{4\pi\epsilon_0} \frac{Q}{|x-x_0|},$$

donde $\epsilon_0 = 8.8542 \times 10^{-12} \text{ F m}^{-1} = 8.8542 \times 10^{-12} \text{ A}^2 \text{ s}^4 \text{ kg}^{-1} \text{ m}^{-3}$ es una constante conocida como *permitividad del vacío*. La ecuación es la misma que la del potencial gravitatorio, con una constante distinta delante, y con el signo contrario. Siguiendo el mismo razonamiento que antes, si $\rho = \rho(x)$ representa una densidad de carga y V es el potencial producido por ella, deducimos que

$$V(x) = \frac{1}{4\pi\epsilon_0} \int_{\mathbb{R}^d} \frac{1}{|x-y|} \rho(y) \, dy = \frac{1}{\epsilon_0} \int_{\mathbb{R}^d} \Psi(x-y) \rho(y) \, dy,$$

luego

$$\Delta V(x) = -\frac{1}{\epsilon_0} \rho(x).$$

Ejemplo 11.5. Vamos a calcular el campo gravitatorio V causado por la Tierra, suponiendo por el momento que es una esfera perfecta y homogénea de radio R y densidad (constante) ρ , centrada en 0. De acuerdo con (226), deberíamos calcular la integral

$$V(x) = -\rho \int_{B_R} \frac{G}{|x-y|} dy,$$

pero esta integral no es fácil de calcular directamente. En su lugar, sabemos que V tiene que ser una función radialmente simétrica y que cumple

$$\Delta V(x) = \begin{cases} 4\pi G\rho & \text{para } |x| < R, \\ 0 & \text{para } |x| > R. \end{cases}$$

Es fácil intuir una solución para este problema, ya que las funciones cuadráticas tienen laplaciano constante, y la solución fundamental que conocemos tiene laplaciano nulo. Podemos poner

$$V(x) = \begin{cases} \frac{2}{3}\pi G\rho|x|^2 + C_1 & \text{para } |x| \leq R, \\ C_2 + \frac{C_3}{|x|} & \text{para } |x| > R. \end{cases}$$

Podemos ver que la única constante que puede cumplir esto es $C_3 = -\rho G$, ya que de otra forma el comportamiento cuando $x \rightarrow +\infty$ de V no sería el correcto. Las constantes C_1 y C_2 pueden encontrarse ya que el valor de V es fácil de calcular en $x = 0$, y sabemos que V debe ser continuo en $|x| = R$.

Ejemplo 11.6. ¿Cuál es el potencial eléctrico en el interior de un sólido conductor cargado? La diferencia con el ejemplo anterior es que antes suponíamos que la densidad de masa era conocida, pero ahora no sabemos la densidad de carga. Como suponemos que el sólido es conductor, los electrones pueden moverse libremente, y adoptarán alguna posición de equilibrio que nos da la densidad de carga. El principio físico básico para esto es que adoptan las posiciones que den energía mínima. La energía total asociada a una densidad de carga ρ es

$$E[\rho] := \frac{1}{2} \int_{\mathbb{R}^3} \rho(x)V(x) dx = \frac{1}{8\pi\epsilon_0} \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} \rho(x)\rho(y) \frac{1}{|x-y|} dx dy.$$

Lo que buscamos es una distribución ρ tal que

$$\int_{\mathbb{R}^3} \rho(x) dx = Q,$$

y tal que $E[\rho]$ sea la mínima posible. Usando una variación admisible h con $\int_{\mathbb{R}^3} h = 0$ obtenemos que

$$0 = \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} h(x)\rho(y) \frac{1}{|x-y|} dx dy = 4\pi\epsilon_0 \int_{\mathbb{R}^3} h(x)V(x) dx.$$

De esto deducimos que $V(x)$ debe ser constante en el soporte de ρ (ya que h tiene que tener soporte contenido en el soporte de ρ). Pero además, $\Delta V = -(1/\epsilon_0)\rho(x)$ para x en el interior de Ω , luego ρ debe ser 0 en el interior de Ω .

11.3. La ecuación de Laplace en el cuadrado: separación de variables

Al igual que hicimos para la ecuación de ondas, podemos encontrar soluciones particulares en variables separadas de la ecuación de Laplace, y después usarlas para encontrar una solución general del problema de frontera asociado. Fijemos $L > 0$, la longitud del lado del cuadrado. Buscamos soluciones de la ecuación

$$\Delta u = 0 \quad \text{en } (0, L) \times (0, L) \quad (227)$$

que sean de la forma

$$u(x, y) = \phi(x)\psi(y).$$

Sustituyendo en (227) tenemos

$$\phi''(x)\psi(y) + \phi(x)\psi''(y) = 0.$$

De la misma forma que en el caso de la ecuación de ondas, deducimos que debe haber una constante $\lambda \in \mathbb{R}$ tal que

$$\phi''(x) = \lambda\phi(x), \quad \psi''(y) = -\lambda\psi(y), \quad (228)$$

al menos en los puntos (x, y) donde $\phi(x) \neq 0$, $\psi(y) \neq 0$. Vamos a suponer, para empezar, que la solución u vale 0 en tres lados del cuadrado:

$$u(x, 0) = u(x, L) = u(0, y) = 0, \quad x, y \in [0, L].$$

esto implica que

$$\psi(0) = \psi(L) = 0, \quad \phi(0) = 0.$$

Junto con la ecuación (228) obtenemos un problema de Sturm-Liouville para ψ , que sabemos que tiene como únicas soluciones

$$\psi(y) = A \operatorname{sen} \left(\frac{n\pi y}{L} \right),$$

para $A \in \mathbb{R}$, $n \geq 1$ un entero, que corresponden a

$$\lambda = \frac{n^2\pi^2}{L^2}.$$

En cuanto a ϕ , para respetar la condición de contorno $\phi(0) = 0$ debe ser

$$\phi(x) = B \operatorname{senh} \left(\frac{n\pi x}{L} \right).$$

Por tanto (quitando las constantes), hemos encontrado las siguientes soluciones en variable separadas, para $n \geq 1$ entero:

$$u_n(x, y) = \phi(x)\psi(y) = \operatorname{senh} \left(\frac{n\pi x}{L} \right) \operatorname{sen} \left(\frac{n\pi y}{L} \right).$$

En el borde $\{x = L\}$ esta solución vale

$$u_n(L, y) = \phi(L)\psi(y) = A \operatorname{senh}(n\pi) \operatorname{sen} \left(\frac{n\pi y}{L} \right).$$

Si consideramos una combinación lineal de estas soluciones (formal por ahora),

$$u(x, y) = \sum_{n=1}^{\infty} A_n u_n(x, y),$$

evaluando en (L, y) obtenemos

$$u(L, y) = \sum_{n=1}^{\infty} A_n \sinh(n\pi) \sin\left(\frac{n\pi y}{L}\right).$$

Si queremos que esto sea igual a g entonces debemos elegir los coeficientes A_n de forma que

$$A_n \sinh(n\pi) = b_n,$$

con b_n los coeficientes de la serie de Fourier en senos de g . Esto puede hacerse de forma rigurosa:

Teorema 11.7. Sea $L > 0$, $g: [0, L] \rightarrow \mathbb{R}$ una función en $H^3([0, L])$ con $g(0) = g(L) = g''(0) = g''(L) = 0$. La ecuación de Laplace

$$\begin{cases} \Delta u = 0 & \text{en } (0, L) \times (0, L), \\ u(x, 0) = u(x, L) = u(0, y) = 0, & x, y \in [0, L], \\ u(L, y) = g(y) & y \in [0, L] \end{cases} \quad (229)$$

tiene una solución dada por

$$u(x, y) = \sum_{n=1}^{\infty} \frac{b_n}{\sinh(n\pi)} \sinh\left(\frac{n\pi x}{L}\right) \sin\left(\frac{n\pi y}{L}\right), \quad (230)$$

donde b_n son los coeficientes de la serie de Fourier en senos de g en $[0, L]$.

Demostración. Gracias a la regularidad de g el Teorema 9.16 nos dice que

$$\sum_{n=1}^{\infty} n^2 |b_n| < +\infty.$$

Esto es suficiente para ver que la serie (230) converge uniformemente, se puede derivar dos veces término a término, y resuelve la ecuación. \square

11.4. Ejercicios

12. El Teorema de Lax-Milgram y su aplicación a las EDPs

12.1. El Teorema de Lax-Milgram

Teorema 12.1 (Lax-Milgram). Sea H un espacio de Hilbert real, $a: H \times H \rightarrow \mathbb{R}$ una aplicación bilineal continua y $\ell: H \rightarrow \mathbb{R}$ una aplicación lineal continua. Supongamos que existe $\lambda > 0$ tal que

$$a(v, v) \geq \lambda \|v\|^2 \quad \text{para todo } v \in H. \quad (231)$$

Entonces existe un único $u \in H$ tal que

$$a(u, v) = \ell(v) \quad \text{para todo } v \in H. \quad (232)$$

Si además a es simétrica, este $u \in H$ es también el único punto crítico del funcional $F : H \rightarrow \mathbb{R}$ definido por

$$F[u] = \frac{1}{2}a(u, u) - \ell(u), \quad (233)$$

y el mínimo del funcional F se alcanza en u .

Observación 12.2. El teorema anterior puede enunciarse también para un espacio vectorial complejo sustituyendo la aplicación bilineal a por una sesquilineal y la condición (237) por

$$\Re(a(v, v)) \geq \lambda \|v\|^2 \quad \text{para todo } v \in H. \quad (234)$$

Ejercicio 12.1. Supongamos que a cumple las condiciones del Teorema 12.1, pero no es simétrica. Consideramos su *simetrización*

$$a_s(u, v) := \frac{1}{2}(a(u, v) + a(v, u)), \quad u, v \in H.$$

Demuestra que un punto crítico del funcional $F : H \rightarrow \mathbb{R}$ dado por (233) debe cumplir

$$a_s(u, v) = \ell(v) \quad \text{para todo } v \in H.$$

Como consecuencia del Teorema 12.1, demuestra que existe un único punto crítico del funcional F , y que en él se alcanza el mínimo de F .

Demostración del Teorema 12.1. Por el teorema de representación de Riesz H.9, existe un único $f \in H$ tal que

$$\ell(v) = \langle f, v \rangle \quad \text{para todo } v \in H.$$

Por otra parte, para cada $v \in H$ la aplicación de H en \mathbb{R} dada por $w \mapsto a(v, w)$ es lineal y continua. De nuevo por el teorema de representación de Riesz, existe un único $A(v) \in H$ tal que

$$a(v, w) = \langle A(v), w \rangle.$$

Esto define una aplicación $A : H \rightarrow H$ que es simplemente otra forma de representar la aplicación bilineal $a : H \times H \rightarrow \mathbb{R}$. Se puede comprobar que $A : H \rightarrow H$ es lineal y continua: para todo $v \in H$ tenemos

$$\|A(v)\|^2 = \langle A(v), A(v) \rangle = a(v, A(v)) \leq C\|v\|\|A(v)\|, \quad (235)$$

luego

$$\|A(v)\| \leq C\|v\| \quad \text{para todo } v \in H. \quad (236)$$

En cuanto a la coercividad, en términos de A se traduce en

$$\langle v, A(v) \rangle = \langle A(v), v \rangle \geq \lambda \|v\|^2 \quad \text{para todo } v \in H. \quad (237)$$

Nuestro objetivo, escrito en términos de A y f , es encontrar $u \in H$ tal que $\langle A(u), w \rangle = \langle f, w \rangle$ para todo $w \in H$, o lo que es lo mismo, tal que

$$A(u) = f.$$

Esto se puede conseguir siempre que $A : H \rightarrow H$ sea una aplicación invertible. Para ver que efectivamente A es invertible, elegimos $r > 0$ y calculamos

$$\|(rA - I)(v)\|^2 = \langle rA(v) - v, rA(v) - v \rangle = r^2\|A(v)\|^2 + \|v\|^2 - r\langle A(v), v \rangle - r\langle v, A(v) \rangle, \quad (238)$$

para cualquier $v \in H$. Usando (236) y (237) tenemos

$$\|(rA - I)(v)\|^2 \leq r^2 C\|v\|^2 + \|v\|^2 - 2r\lambda\|v\|^2 = (Cr^2 - 2r\lambda + 1)\|v\|^2. \quad (239)$$

Podemos elegir $r > 0$ suficientemente pequeño de forma que el coeficiente de $\|v\|^2$ sea menor que 1. Esto demuestra que $\|rA - I\| < 1$, luego rA es invertible, o lo que es lo mismo, A es invertible. Esto demuestra que existe un único $u \in H$ que cumple (232).

Vamos a demostrar la parte sobre el funcional F en el caso de a simétrica. Por definición, un punto crítico u de F es aquel que cumple que para todo $v \in H$

$$\begin{aligned} 0 &= \left. \frac{d}{dt} \right|_{t=0} \left(\frac{1}{2}a(u + tv, u + tv) - \ell(u + tv) \right) = \frac{1}{2}a(u, v) + \frac{1}{2}a(v, u) - \ell(v) \\ &= a(u, v) - \ell(v), \end{aligned} \quad (240)$$

es decir, es una solución de (232). Esto prueba que existe un punto crítico y es único. En ese punto crítico u tenemos que

$$\begin{aligned} F[u + h] &= \frac{1}{2}a(u + h, u + h) - \ell(u + h) \\ &= \frac{1}{2}a(u, u) + a(u, h) + \frac{1}{2}a(h, h) - \ell(u) - \ell(h) \\ &= F[u] + a(u, h) + \frac{1}{2}a(h, h) - \ell(h) = F[u] + \frac{1}{2}a(h, h) \geq F[u]. \end{aligned} \quad (241)$$

Esto demuestra que u es de hecho un mínimo de F . \square

12.2. El concepto de derivada débil

La derivada débil de una función es una generalización de la idea de derivada que puede aplicarse a algunas funciones localmente integrables que no tienen derivada en sentido clásico. La idea más importante es que la derivada débil de una función es otra que hace que siga siendo cierto el teorema de integración por partes. Esta idea es en realidad una versión un poco primitiva de la derivación en teoría de distribuciones, que no vamos a tratar en este curso.

Supongamos que $u : \Omega \rightarrow \mathbb{R}$ es una función derivable (en sentido de Fréchet) definida en un abierto $\Omega \subseteq \mathbb{R}^d$ con frontera regular (al menos de clase \mathcal{C}^1). Si $\phi : \Omega \rightarrow \mathbb{R}^d$ es una función derivable de soporte compacto en Ω , usando la fórmula de Green (Corolario B.5) es fácil ver que

$$\int_{\mathbb{R}^d} u(x) \partial_{x_n} \phi(x) \, dx = - \int_{\mathbb{R}^d} \phi(x) \partial_{x_n} u(x) \, dx,$$

para $n = 1, \dots, d$. Esta fórmula es precisamente la que vamos a tomar como *definición* del concepto de derivada débil:

Definición 12.3. Sea $\Omega \subseteq \mathbb{R}^d$ un abierto, $u : \Omega \rightarrow \mathbb{R}$ una función localmente integrable. Para $n = 1, \dots, d$, decimos que una función g localmente integrable en Ω es la derivada débil de u con respecto a x_n cuando

$$\int_{\Omega} u \partial_{x_n} \phi = - \int_{\Omega} g \phi \quad (242)$$

para toda función $\phi \in \mathcal{C}_c^\infty(\Omega)$. Cuando esto ocurre decimos que g es la derivada débil de u respecto a x_n , y escribimos

$$\partial_{x_n} u = g.$$

Lema 12.4. Sean $\Omega \subseteq \mathbb{R}^d$ un abierto, $u : \Omega \rightarrow \mathbb{R}$ una función localmente integrable, $n = 1, \dots, d$.

1. La derivada débil de u con respecto a x_n , si existe, es única salvo redefinición en conjuntos de medida nula.
2. Si una función tiene todas sus derivadas débiles igual a 0, entonces es constante en casi todo punto.
3. Si $g = \partial_{x_n} u$ (en sentido débil) y $\tilde{g} : \Omega \rightarrow \mathbb{R}$ es igual a g en casi todo punto de Ω entonces \tilde{g} es también una derivada débil de u respecto a x_n .
4. Si $u \in \mathcal{C}^1(\Omega)$ entonces su derivada débil existe y coincide en casi todo punto con su derivada parcial usual.

Demostración. La demostración es directa usando el lema fundamental del cálculo de variaciones. \square

Lema 12.5. La derivada débil es lineal. Es decir, si $\Omega \subseteq \mathbb{R}^d$ es un abierto, $n \in \{1, \dots, d\}$, y $u, v : \Omega \rightarrow \mathbb{R}$ son funciones localmente integrables cuyas derivadas débiles con respecto a x_n existen entonces

$$\partial_{x_n}(u + v) = \partial_{x_n} u + \partial_{x_n} v \quad \text{en casi todo punto de } \Omega.$$

En el caso de \mathbb{R} se puede dar una caracterización sencilla del concepto de derivada débil:

Proposición 12.6. Sea $I \subseteq \mathbb{R}^d$ un abierto, $u : I \rightarrow \mathbb{R}$ una función localmente integrable. Son equivalentes:

1. Existe la derivada débil de u .
2. Hay una función v , localmente absolutamente continua en I que es igual a u en casi todo punto de I .

Si ocurre lo anterior entonces v' (definida en casi todo punto) es una derivada débil de u .

Demostración. Si existe una derivada débil de u , a la que llamamos g , tenemos

$$\int_I u \phi' = - \int_I g \phi \quad \text{para todo } \phi \in \mathcal{C}_c^\infty(I).$$

Por otra parte, como g es localmente integrable, el teorema de integración por partes para la integral de Lebesgue dice que

$$-\int_I g\phi = \int_I G\phi',$$

donde G es una primitiva cualquiera de G , que podemos elegir de forma que coincida con u en al menos un punto de I . Entonces

$$\int_I u\phi' = \int_I G\phi' \quad \text{para todo } \phi \in \mathcal{C}_c^\infty(I),$$

lo cual implica que u es igual en casi todo punto a G más una constante, que debe ser 0 ya que u, G coinciden en al menos un punto. Esto demuestra el punto 2 con $v = G$.

Para la implicación inversa, usando de nuevo el teorema de integración por partes para la integral de Lebesgue,

$$\int_I u\phi' = \int_I v\phi' = -\int_I v'\phi \quad \text{para todo } \phi \in \mathcal{C}_c^\infty(I),$$

luego v' es una derivada débil de u . □

En dimensión mayor que 1 no hay una caracterización tan sencilla para la derivada débil. Por ejemplo, el siguiente ejercicio muestra funciones que no son continuas en \mathbb{R}^d , que son derivables en sentido débil:

Ejercicio 12.2. Demuestra que la función $u : \mathbb{R}^d \rightarrow \mathbb{R}$ definida por

$$u(x) = |x|^k, \quad x \neq 0,$$

(y con $u(0) = 0$, por ejemplo) es derivable en sentido débil (con respecto a cualquier componente) si y sólo si $k > 1 - d$, o bien $k = 0$. Encuentra su derivada débil en este caso.

Solución. Para que tenga sentido el concepto de derivada débil, la función u tiene que ser localmente integrable; esto sólo es cierto si $k > -d$, así que suponemos siempre $k > -d$ (si no, por definición u no tiene ninguna derivada débil).

El caso $k = 0$ es especial, porque entonces u es constante y tiene derivada débil siempre. Así que supondremos también que $k \neq 0$.

Supongamos que u tiene derivada débil g con respecto a x_1 por ejemplo (el razonamiento es análogo con respecto a las demás coordenadas). Entonces para toda función $\varphi \in \mathcal{C}_c^\infty(\mathbb{R}^d)$ debe cumplirse

$$\int_{\mathbb{R}^d} |x|^k \partial_{x_1} \varphi(x) dx = - \int_{\mathbb{R}^d} g(x) \varphi(x) dx.$$

En particular, se cumple para toda $\varphi \in \mathcal{C}_c^\infty(\mathbb{R}^d \setminus \{0\})$. En ese caso, como u es una función \mathcal{C}^1 en $\mathbb{R}^d \setminus \{0\}$,

$$\int_{\mathbb{R}^d} |x|^k \partial_{x_1} \varphi(x) dx = - \int_{\mathbb{R}^d} \partial_{x_1} (|x|^k) \varphi(x) dx = - \int_{\mathbb{R}^d} k x_1 |x|^{k-2} \varphi(x) dx.$$

De las últimas dos ecuaciones vemos que

$$\int_{\mathbb{R}^d} g(x) \varphi(x) dx = \int_{\mathbb{R}^d} kx_1 |x|^{k-2} \varphi(x) dx$$

para toda $\varphi \in \mathcal{C}_c^\infty(\mathbb{R}^d \setminus \{0\})$. Por el lema fundamental,

$$g(x) = kx_1 |x|^{k-2} \quad \text{p.c.t. } x \in \mathbb{R}^d \setminus \{0\},$$

o lo que es lo mismo, $g(x) = kx_1 |x|^{k-2}$ en casi todo $x \in \mathbb{R}^d$. Así que si u tiene derivada débil respecto de x_1 , esa derivada débil debe ser igual a $kx_1 |x|^{k-2}$ en casi todo punto.

Por definición, esta función g que hemos encontrado sólo puede ser una derivada débil cuando sea localmente integrable. Vamos a ver cuándo ocurre esto: como $|x_1| \leq |x|$, en cualquier bola B_R centrada en 0 y de radio $R > 0$ tenemos

$$\int_{B_R} |g(x)| dx \leq |k| \int_{B_R} |x|^{k-1} dx,$$

que es finito si y sólo si $k > 1 - d$, luego g es localmente integrable en este rango. Cuando $k \leq 1 - d$ podemos ver que de hecho no es localmente integrable, porque:

$$\int_{B_R} |x_1| |x|^{k-2} dx \geq \int_{B_R} |x_1|^2 |x|^{k-3} dx = \frac{1}{d} \int_{B_R} |x|^{k-1} dx,$$

donde el último paso es debido a que la integral no cambia si cambiamos x_1 por cualquier otra coordenada. Por lo tanto, hemos demostrado que para $k \leq 1 - d$, u no tiene ninguna derivada débil (con la excepción de $k = 0$ en dimensión 1, que sí tiene).

Por otra parte, cuando $k > 1 - d$, demostramos al principio que para toda $\varphi \in \mathcal{C}_c^\infty(\mathbb{R}^d \setminus \{0\})$,

$$\int_{\mathbb{R}^d} |x|^k \partial_{x_1} \varphi(x) dx = - \int_{\mathbb{R}^d} kx_1 |x|^{k-2} \varphi(x) dx. \quad (243)$$

Como en el caso $k > 1 - d$ sabemos que la función $kx_1 |x|^{k-2}$ es localmente integrable, podemos aproximar cualquier función $\varphi \in \mathcal{C}_c^\infty(\mathbb{R}^d)$ por una sucesión creciente de funciones $\varphi_n \in \mathcal{C}_c^\infty(\mathbb{R}^d \setminus \{0\})$. El teorema de la convergencia dominada nos permite pasar al límite en n en la ecuación (243) para φ_n , y ver que se cumple para toda $\varphi \in \mathcal{C}_c^\infty(\mathbb{R}^d)$. Esto es la definición de derivada débil, y demuestra que en este caso

$$\partial_{x_1}(|x|^k) = kx_1 |x|^{k-2}$$

en sentido débil.

Ejercicio 12.3. Sea $\Omega := (0, 1) \times (0, 1) \subseteq \mathbb{R}^2$ y $f: \Omega \rightarrow \mathbb{R}$ definida por

$$f(x, y) := \begin{cases} f_1(x) & 0 < x < 1/2, \\ f_2(x) & 1/2 \leq x < 1, \end{cases}$$

donde $f_1 \in \mathcal{C}^1([0, 1/2])$, $f_2 \in \mathcal{C}^1([1/2, 1])$.

1. Demuestra que existe $\partial_y f$ en sentido débil, y es igual a 0.
2. Demuestra que si f no es continua (es decir, si $f_1(1/2) \neq f_2(1/2)$) entonces no existe $\partial_x f$ en sentido débil.
3. Suponiendo que f es continua (es decir, si $f_1(1/2) = f_2(1/2)$), demuestra que existe $\partial_x f$ en sentido débil, y calcúlala.

12.3. El espacio de Sobolev $H^k(\Omega)$

Definición 12.7. Sea $\Omega \subseteq \mathbb{R}^d$ abierto, $k \geq 1$ entero. Definimos

$$H^k(\Omega) := \left\{ f \in L^2(\Omega) \left| \begin{array}{l} \text{para todo multiíndice } \alpha \text{ con } |\alpha| \leq k, \\ \partial_\alpha f \text{ existe (en sentido débil)} \\ \text{y está en } L^2(\Omega) \text{ para } n = 1, \dots, d. \end{array} \right. \right\} \quad (244)$$

En $H^k(\Omega)$ definimos un producto escalar

$$\langle u, v \rangle_{H^k(\Omega)} = \sum_{|\alpha| \leq k} \int_{\Omega} \partial_\alpha u \partial_\alpha v \quad \text{para } u, v \in H^k(\Omega), \quad (245)$$

donde la suma es en todos los d -multiíndices de orden menor o igual que k . En particular, para $k = 1$ tenemos

$$\langle u, v \rangle_{H^1(\Omega)} = \int_{\Omega} uv + \int_{\Omega} \nabla u \nabla v \quad \text{para } u, v \in H^1(\Omega). \quad (246)$$

Lema 12.8. Para $k \geq 1$ entero, la expresión (245) define un producto escalar en $H^k(\Omega)$, y $H^k(\Omega)$ es completo con la métrica inducida por $\langle \cdot, \cdot \rangle_{H^k(\Omega)}$ (es decir, $H^k(\Omega)$ es un espacio de Hilbert con el producto escalar definido por (245)).

La norma asociada al producto escalar (245) es

$$\|u\|_{H^k(\Omega)} = \left(\sum_{|\alpha| \leq k} \int_{\Omega} (\partial_\alpha u)^2 \right)^{\frac{1}{2}}, \quad u \in H^k(\Omega).$$

En particular, la asociada al producto escalar (246) es

$$\|u\|_{H^1(\Omega)} = \left(\int_{\Omega} u^2 + \int_{\Omega} |\nabla u|^2 \right)^{\frac{1}{2}}, \quad u \in H^1(\Omega). \quad (247)$$

Definición 12.9. Sea $\Omega \subseteq \mathbb{R}^d$ abierto. Definimos $H_0^k(\Omega)$ como la completación en la norma de $H^k(\Omega)$ del subespacio $C_c^\infty(\Omega) \subseteq H^k(\Omega)$.

Ejercicio 12.4. Sea $\Omega \subseteq \mathbb{R}^d$ un abierto, $k \geq 1$, $i \in \{1, \dots, d\}$. Demuestra que

1. $u \in H^k(\Omega) \implies \partial_{x_i} u \in H^{k-1}(\Omega)$.
2. $u \in H_0^k(\Omega) \implies \partial_{x_i} u \in H_0^{k-1}(\Omega)$.

12.4. Aplicación a la ecuación de Poisson

Definición 12.10 (Solución débil de la ecuación de Poisson). Sea $\Omega \subseteq \mathbb{R}^d$ un abierto no vacío, $f \in L^2(\Omega)$. Decimos que $u \in H_0^1(\Omega)$ es una *solución débil* del problema

$$\begin{cases} \Delta u = f & \text{en } \Omega, \\ u = 0 & \text{en } \partial\Omega \end{cases} \quad (248)$$

cuando

$$\int_{\Omega} \nabla u \nabla v = - \int_{\Omega} f v \quad (249)$$

para todo $v \in H_0^1(\Omega)$.

Lema 12.11 (Una solución débil regular es solución clásica). *Supongamos que $u : \bar{\Omega} \rightarrow \mathbb{R}$ es una solución débil de la ecuación de Poisson en un abierto $\Omega \subseteq \mathbb{R}^d$ (se entiende que es solución débil cuando la restringimos a Ω). Si $u \in C^2(\bar{\Omega})$ entonces u es también una solución clásica.*

Ejercicio 12.5. Demuestra el lema anterior.

Teorema 12.12 (Desigualdad de Poincaré en H_0^1). *Sea $\Omega \subseteq \mathbb{R}^d$ un abierto acotado y no vacío. Existe una constante $\lambda_\Omega > 0$ tal que*

$$\lambda_\Omega \|u\|_{L^2(\Omega)} \leq \|\nabla u\|_{L^2(\Omega)} \quad \text{para todo } u \in H_0^1(\Omega). \quad (250)$$

Ejercicio 12.6. Sea $\Omega \subseteq \mathbb{R}^d$ un abierto. Para $u \in H_0^1(\Omega)$ definimos

$$\|u\|_*^2 := \int_{\Omega} |\nabla u|^2.$$

Demuestra que $\|\cdot\|_*$ define una norma en $H_0^1(\Omega)$ que es equivalente a $\|\cdot\|_{H^1(\Omega)}$.

Ejercicio 12.7. Demuestra la desigualdad de Poincaré cuando $\Omega = (a, b) \subseteq \mathbb{R}$ es un intervalo abierto y acotado.

Teorema 12.13 (Existencia de solución débil de la ecuación de Poisson). *Sea $\Omega \subseteq \mathbb{R}^d$ abierto y acotado, $f \in H^1(\Omega)$. Existe una única solución débil $u \in H_0^1(\Omega)$ de la ecuación de Poisson*

$$\Delta u = f. \quad (251)$$

Dicha solución débil es el único punto crítico del funcional

$$F : H_0^1(\Omega) \rightarrow \mathbb{R}, \quad (252)$$

$$F[u] := \frac{1}{2} \int_{\Omega} |\nabla u|^2 - \int_{\Omega} f u, \quad u \in H_0^1(\Omega), \quad (253)$$

y el mínimo de F se alcanza en u .

Demostración. Consideramos la aplicación bilineal

$$a(u, v) := \int_{\Omega} \nabla u \nabla v, \quad (254)$$

definida para todo $u, v \in H_0^1(\Omega)$, y la aplicación lineal

$$\ell : H \rightarrow \mathbb{R}, \quad \ell(v) := \int_{\Omega} f v.$$

Estas aplicaciones satisfacen las hipótesis del teorema de Lax-Milgram en el espacio $H_0^1(\Omega)$, luego sabemos que existe un único $u \in H_0^1(\Omega)$ que es solución débil de la ecuación. \square

El siguiente resultado es la justificación real del concepto de solución débil: una vez que hemos encontrado una solución débil es posible demostrar que es una solución regular, suponiendo que el dominio y los datos del problema tienen regularidad suficiente. La demostración de este resultado es larga y no vamos a verla:

Lema 12.14 (Regularidad de soluciones débiles). *Supongamos que $u : \Omega \rightarrow \mathbb{R}$ es una solución débil de la ecuación de Poisson en un abierto $\Omega \subseteq \mathbb{R}^d$ con frontera \mathcal{C}^2 , y con dato $f \in \mathcal{C}(\Omega)$. Entonces*

1. $u \in \mathcal{C}^2(\Omega)$ (mejor dicho, existe $v \in \mathcal{C}^2$ tal que $u = v$ p.c.t. $x \in \Omega$),
2. u puede extenderse de forma única a una función $\tilde{u} \in \mathcal{C}^2(\overline{\Omega})$,
3. y \tilde{u} es una solución clásica de la ecuación de Poisson.

12.5. Aplicación a la ecuación $\Delta u + \lambda u = f$

Podemos usar la misma estrategia para estudiar el siguiente problema de frontera: para $\Omega \subseteq \mathbb{R}^d$ abierto, $\lambda \in \mathbb{R}$ y $f \in L^2(\Omega)$ consideramos:

$$\begin{cases} \Delta u + \lambda u = f & \text{en } \Omega, \\ u = 0 & \text{en } \partial\Omega \end{cases} \quad (255)$$

La ecuación $\Delta u + \lambda u = 0$ se conoce como *ecuación de Helmholtz*. El problema (255) es por tanto una versión no homogénea de la ecuación de Helmholtz.

El concepto de solución débil de (255) se define de forma análoga al de la ecuación de Poisson:

Definición 12.15 (Solución débil de (255)). Sea $\Omega \subseteq \mathbb{R}^d$ un abierto no vacío, $f \in L^2(\Omega)$. Decimos que $u \in H_0^1(\Omega)$ es una *solución débil* del problema (255) cuando

$$\int_{\Omega} \nabla u \nabla v - \lambda \int_{\Omega} uv = - \int_{\Omega} f v \quad \text{para todo } v \in H_0^1(\Omega). \quad (256)$$

Teorema 12.16 (Existencia de solución débil de (255)). Sea $\Omega \subseteq \mathbb{R}^d$ un abierto no vacío, $f \in L^2(\Omega)$, y λ_{Ω} la constante de Poincaré de Ω . Para todo $c < \lambda_{\Omega}$ hay una única solución débil del problema (255) en $H_0^1(\Omega)$.

Notamos que no estamos diciendo nada cuando $c \geq \delta$. Lo que ocurre en realidad es que este problema se comporta como los problemas de contorno de la sección 7.4 (de hecho, la versión en dimensión 1 de (255) cuando Ω es un intervalo I es un problema de Sturm-Liouville): hay una sucesión $\{\lambda_n\}_{n \geq 0}$ que tiende a infinito tal que el problema (255) siempre tiene solución para $\lambda \notin \{\lambda_n\}_{n \geq 0}$. Estos λ_n son los *valores propios* del laplaciano en el dominio Ω ; para $\lambda = \lambda_n$ el problema (255) se cumple una versión de la alternativa de Fredholm: el problema tiene solución si y sólo si f es perpendicular a cualquier solución de la ecuación homogénea correspondiente (la ecuación de Helmholtz con $\lambda = \lambda_n$).

Demostración del Teorema 12.16. Definimos la forma bilineal $a : H_0^1(\Omega) \times H_0^1(\Omega) \rightarrow \mathbb{R}$,

$$a(u, v) := \int_{\mathbb{R}^d} \nabla u \nabla v - \lambda \int_{\mathbb{R}^d} uv, \quad u, v \in H_0^1(\Omega).$$

Se ve fácilmente que esta aplicación es bilineal y continua; para aplicar el teorema de Lax-Milgram tenemos que demostrar que es también coerciva:

$$a(u, u) = \int_{\mathbb{R}^d} |\nabla u|^2 - \lambda \int_{\mathbb{R}^d} u^2 \geq (\lambda_{\Omega} - \lambda) \int_{\mathbb{R}^d} u^2.$$

El coeficiente $(\lambda_{\Omega} - \lambda)$ es positivo cuando $\lambda < \lambda_{\Omega}$, y entonces se puede completar el razonamiento igual que en el Teorema 12.13. \square

12.6. Ejercicios

Ejercicio 12.8. Sea $L > 0$ y $p: [0, L] \rightarrow \mathbb{R}$ una función continua. Consideramos el funcional asociado al problema de la viga de la sección 7.5:

$$F[y] := \int_0^L |y''(x)|^2 dx + \int_0^L p(x)y(x) dx, \quad (257)$$

definido en el dominio

$$D := \{y \in \mathcal{C}^4([0, L]) \mid y(0) = y'(0) = y(L) = y'(L) = 0\}. \quad (258)$$

Demuestra que este funcional tiene un único mínimo en D .

Ejercicio 12.9. Demuestra que el funcional del Ejercicio 8.2 tiene un único mínimo en el dominio en el que está definido.

Ejercicio 12.10. Demuestra que el funcional

$$\mathcal{F}(u) := \int_0^1 (u'(x))^2 dx - \int_0^1 (u(x) - \sin(x))^2 dx,$$

definido para $u \in H_0^1(0, 1)$, tiene un mínimo y lo alcanza en una única función de $H_0^1(0, 1)$. (Sugerencia: puedes usar que la constante de Poincaré del intervalo $[0, 1]$ es π^2 ; es decir, $\pi^2 \int_0^1 u^2 \leq \int_0^1 (u')^2$ para toda $u \in H_0^1(0, 1)$.)

Solución. Expandimos \mathcal{F} como

$$\mathcal{F}(u) := \int_0^1 (u'(x))^2 dx - \int_0^1 u(x)^2 dx + 2 \int_0^1 u(x) \sin(x) dx - \int_0^1 (\sin(x))^2 dx.$$

El último término es constante, luego consideramos

$$\tilde{\mathcal{F}}(u) := \int_0^1 (u'(x))^2 dx - \int_0^1 u(x)^2 dx + 2 \int_0^1 u(x) \sin(x) dx,$$

que alcanza su mínimo si y sólo si lo alcanza \mathcal{F} , y en las mismas funciones. Para aplicar el Teorema de Lax-Milgram podemos definir

$$a(u, v) := 2 \int_0^1 u'(x)v'(x) dx - 2 \int_0^1 u(x)v(x) dx, \quad \ell(v) := -2 \int_0^1 u(x) \sin(x) dx,$$

para $u, v \in H_0^1(0, 1)$. Comprobamos fácilmente que a es bilineal y continua, y que ℓ es bilineal y continua. Para ver que a es coerciva usamos la desigualdad de Poincaré

$$\int_0^1 u(x)^2 dx \leq \frac{1}{\pi^2} \int_0^1 (u'(x))^2 dx,$$

con lo cual

$$\begin{aligned} a(v, v) &= 2 \int_0^1 (v'(x))^2 dx - 2 \int_0^1 v(x)^2 dx \\ &\geq 2 \int_0^1 (v'(x))^2 dx - \frac{2}{\pi^2} \int_0^1 (v'(x))^2 dx = 2 \left(1 - \frac{1}{\pi^2}\right) \int_0^1 (v'(x))^2 dx, \end{aligned}$$

para todo $v \in H_0^1(0, 1)$. Usando esta desigualdad y de nuevo la desigualdad de Poincaré,

$$\|v\|_{H^1}^2 = \int_0^1 (v')^2 + \int_0^1 v^2 \leq (1 + \frac{1}{\pi^2}) \int_0^1 (v')^2 \leq \frac{1 + \frac{1}{\pi^2}}{1 - \frac{1}{\pi^2}} a(v, v) = \frac{\pi^2 + 1}{\pi^2 - 1} a(v, v).$$

Esto demuestra la coercividad. El teorema de Lax-Milgram nos dice entonces que existe un único mínimo del funcional \mathcal{F} . (Aunque no es parte del ejercicio, este caso la función donde se alcanza el mínimo se puede calcular. ¿Puedes calcularla?)

13. La transformada de Fourier

La transformada de Fourier es una herramienta fundamental en matemáticas, en parte por lo útil que resulta para “traducir” unos problemas a otros. Nosotros la usaremos principalmente para resolver algunas ecuaciones en derivadas parciales (como ya hicimos con las series de Fourier de la sección 9).

En esta sección los espacios $L^p(\mathbb{R}^d)$ se refieren siempre a espacios de funciones con valores complejos (es decir, $L^p(\mathbb{R}^d) \equiv L^p(\mathbb{R}^d; \mathbb{C})$).

Definición 13.1 (Transformada de Fourier de una función integrable). La *transformada de Fourier* de una función $f \in L^1(\mathbb{R}^d)$ es la función $\mathcal{F}(f): \mathbb{R}^d \rightarrow \mathbb{C}$ dada por

$$\mathcal{F}(f)(\xi) := (2\pi)^{-\frac{d}{2}} \int_{\mathbb{R}^d} f(x) e^{-ix \cdot \xi} dx, \quad \xi \in \mathbb{R}^d.$$

La *transformada de Fourier inversa* de una función $f \in L^1(\mathbb{R}^d)$ es la función $\mathcal{F}^{-1}(f): \mathbb{R}^d \rightarrow \mathbb{C}$ dada por

$$\mathcal{F}^{-1}(f)(\xi) := (2\pi)^{-\frac{d}{2}} \int_{\mathbb{R}^d} f(x) e^{ix \cdot \xi} dx, \quad \xi \in \mathbb{R}^d.$$

A veces escribimos

$$\hat{f}(\xi) \equiv \mathcal{F}(f)(\xi), \quad \check{f}(\xi) \equiv \mathcal{F}^{-1}(f)(\xi).$$

En el caso de dimensión 1, para una función $f \in L^1(\mathbb{R})$ tenemos

$$\hat{f}(\xi) = \mathcal{F}(f)(\xi) := \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} f(x) e^{-ix \cdot \xi} dx, \quad \xi \in \mathbb{R}.$$

Fíjate en que esta expresión se parece a la representación (132) de las series de Fourier, con la diferencia de que ahora tenemos una integral en lugar de una suma (y el coeficiente ha cambiado).

Una propiedad importante es que tanto la transformada de Fourier como la transformada inversa son *lineales*, lo cual se ve directamente a partir de la definición: si $\lambda, \mu \in \mathbb{R}$ y $f, g \in L^1(\mathbb{R}^d)$,

$$\mathcal{F}(\lambda f + \mu g) = \lambda \mathcal{F}(f) + \mu \mathcal{F}(g), \quad \mathcal{F}^{-1}(\lambda f + \mu g) = \lambda \mathcal{F}^{-1}(f) + \mu \mathcal{F}^{-1}(g).$$

Ejemplo 13.2. Es fácil por ejemplo calcular la transformada de la función característica de un intervalo de la forma $[-R, R]$:

$$\sqrt{2\pi} \mathcal{F}(\mathbb{1}_{[-R, R]})(\xi) = \int_{-R}^R e^{-ix \cdot \xi} dx = -\frac{1}{i\xi} (e^{-iR\xi} - e^{iR\xi}) = \frac{2 \operatorname{sen}(R\xi)}{\xi},$$

luego para $\xi \in \mathbb{R}$ tenemos

$$\mathcal{F}(\mathbb{1}_{[-R,R]})(\xi) = \frac{\sqrt{2} \operatorname{sen}(R\xi)}{\sqrt{\pi}\xi}.$$

Ejemplo 13.3. También podemos calcular la transformada de Fourier de una Gaussiana. Tomamos por ejemplo

$$f(x) := e^{-\frac{|x|^2}{2}}, \quad x \in \mathbb{R}^d.$$

Y calculamos

$$(2\pi)^{\frac{d}{2}} \mathcal{F}(f)(\xi) = \int_{\mathbb{R}^d} e^{-\frac{|x|^2}{2} - ix\xi} dx = e^{-\frac{1}{2}|\xi|^2} \int_{\mathbb{R}^d} e^{-\frac{1}{2}|x+i\xi|^2} dx.$$

Esta última integral se puede calcular explícitamente, ya que el caso de dimensión 1 da

$$\int_{-\infty}^{\infty} e^{-\frac{1}{2}|x+i\xi|^2} dx = \int_{-\infty}^{\infty} e^{-\frac{1}{2}x^2} dx = \sqrt{2\pi},$$

donde el cambio en la integral se puede ver deformando continuamente la integral a lo largo de una curva compleja. En dimensión d tenemos

$$\int_{\mathbb{R}^d} e^{-\frac{1}{2}|x+i\xi|^2} dx = \prod_{j=1}^d \int_{-\infty}^{\infty} e^{-\frac{1}{2}|x_j+i\xi_j|^2} dx_j = (2\pi)^{\frac{d}{2}}.$$

Por tanto,

$$\mathcal{F}(e^{-\frac{1}{2}|x|^2}) = e^{-\frac{1}{2}|\xi|^2}. \quad (259)$$

Lema 13.4 (Continuidad $L^1 - L^\infty$). Para toda $f \in L^1(\mathbb{R}^d)$ se cumple

$$|\mathcal{F}(f)(\xi)| \leq \|f\|_1, \quad \xi \in \mathbb{R}^d.$$

Proposición 13.5 (Propiedades de la transformada de Fourier). Sea $d \geq 1$ un entero.

Traslaciones Para $f \in L^1(\mathbb{R}^d)$, $v \in \mathbb{R}^d$ se cumple

$$\mathcal{F}[f(x-v)](\xi) = e^{-iv\xi} \mathcal{F}(f)(\xi), \quad \xi \in \mathbb{R}^d.$$

Cambio de escala Para $f \in L^1(\mathbb{R}^d)$,

$$\mathcal{F}(f(\lambda x))(\xi) = \frac{1}{\lambda^d} \mathcal{F}(f)\left(\frac{\xi}{\lambda}\right), \quad x \in \mathbb{R}^d. \quad (260)$$

Derivadas Si $f \in L^1(\mathbb{R}^d)$ y $\partial_{x_j} f \in L^1(\mathbb{R}^d)$,

$$\mathcal{F}(\partial_{x_j} f)(\xi) = i\xi_j \mathcal{F}(f)(\xi), \quad \xi \in \mathbb{R}^d, \quad j = 1, \dots, d.$$

Convolución Si $f, g \in L^1(\mathbb{R}^d)$,

$$\mathcal{F}(f * g) = (2\pi)^{\frac{d}{2}} \mathcal{F}(f) \mathcal{F}(g). \quad (261)$$

Demostración. Estas propiedades son fáciles de demostrar, y dejamos la demostración como ejercicio. Probamos por ejemplo la propiedad de la derivada. Si $f \in \mathcal{C}_c^1(\mathbb{R}^d)$ usando integración por partes tenemos que

$$\begin{aligned}\mathcal{F}(\partial_{x_i} f)(\xi) &= \mathcal{F}(\partial_{x_i} f)(\xi) = (2\pi)^{-\frac{d}{2}} \int_{\mathbb{R}^d} \partial_{x_i} f(x) e^{-ix \cdot \xi} dx \\ &= (2\pi)^{-\frac{d}{2}} \int_{\mathbb{R}^d} i\xi_i f(x) e^{-ix \cdot \xi} dx = i\xi_i \mathcal{F}(f)(\xi).\end{aligned}$$

En general, si $f \in L^1(\mathbb{R}^d)$, $\partial_{x_j} f \in L^1(\mathbb{R}^d)$, podemos aproximar f por una sucesión de funciones $f_n \in \mathcal{C}_c^1(\mathbb{R}^d)$ de forma que $f_n \rightarrow f$, $\partial_{x_j} f_n \rightarrow \partial_{x_j} f$ en $L^1(\mathbb{R}^d)$. Usando la continuidad de \mathcal{F} en L^1 (Lema 13.4) podemos demostrar el resultado por aproximación. \square

Ejemplo 13.6. Usando (259) y la propiedad de la escala es fácil encontrar la transformada de Fourier de cualquier Gaussiana:

$$\mathcal{F}(e^{-a|x|^2}) = (2a)^{-\frac{d}{2}} e^{-\frac{|\xi|^2}{4a}}. \quad (262)$$

Lema 13.7 (Isometría en L^2). Para toda $f \in L^1(\mathbb{R}^d) \cap L^2(\mathbb{R}^d)$ se cumple que $\mathcal{F}(f) \in L^2(\mathbb{R}^d)$ y

$$\|\mathcal{F}(f)\|_2 = \|f\|_2.$$

Demostración. De forma no rigurosa, la cuenta que queríamos hacer es la siguiente:

$$\begin{aligned}\|\mathcal{F}(f)\|_2^2 &= \int_{\mathbb{R}^d} |\mathcal{F}(f)(\xi)|^2 d\xi = (2\pi)^{-d} \int_{\mathbb{R}^d} \left| \int_{\mathbb{R}^d} f(x) e^{-ix \cdot \xi} dx \right|^2 d\xi \\ &= (2\pi)^{-d} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} f(x) \overline{f(y)} e^{-ix \cdot \xi} e^{iy \cdot \xi} dx dy d\xi \\ &= (2\pi)^{-d} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} f(x) \overline{f(y)} \left(\int_{\mathbb{R}^d} e^{-ix \cdot \xi} e^{iy \cdot \xi} d\xi \right) dx dy,\end{aligned}$$

pero no podemos hacerlo directamente porque la función $\xi \mapsto e^{-ix \cdot \xi} e^{iy \cdot \xi}$ no es integrable en el sentido de Lebesgue (para ningunos $x, y \in \mathbb{R}^d$), y el último cambio de orden en la integral no es válido. (El truco que hemos usado para quitar el cuadrado añadiendo una integral más se conoce como “desdoblamiento de variables”.) En realidad nos gustaría decir que para $y \in \mathbb{R}^d$ fijo,

$$\int_{\mathbb{R}^d} e^{-ix \cdot \xi} e^{iy \cdot \xi} d\xi = \delta_{x-y},$$

con lo cual (si esto tuviera sentido) obtendríamos lo que queremos.

Para hacer una demostración rigurosa podemos *truncar* la integral. Elegimos una función $\psi: \mathbb{R}^d \rightarrow \mathbb{R}$ que sea continua, integrable, con $\psi(0) = 1$, simétrica y radialmente decreciente, y escribimos

$$\|\mathcal{F}(f)\|_2^2 = \lim_{\epsilon \rightarrow 0} \int_{\mathbb{R}^d} \psi(\epsilon \xi) |\mathcal{F}(f)(\xi)|^2 d\xi.$$

Observa que el límite es cierto por el teorema de la convergencia monótona (es el motivo por el que hemos puesto las condiciones sobre ψ). Con el mismo cálculo de antes,

$$\begin{aligned} \int_{\mathbb{R}^d} \psi(\epsilon\xi) |\mathcal{F}(f)(\xi)|^2 d\xi &= (2\pi)^{-d} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \psi(\epsilon\xi) f(x) \overline{f(y)} e^{-ix \cdot \xi} e^{iy \cdot \xi} dx dy d\xi \\ &= (2\pi)^{-d} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} f(x) \overline{f(y)} \left(\int_{\mathbb{R}^d} \psi(\epsilon\xi) e^{-i(x-y) \cdot \xi} d\xi \right) dx dy \\ &= (2\pi)^{-d} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} f(x) \overline{f(y)} \epsilon^{-d} \mathcal{F}(\psi)\left(\frac{x-y}{\epsilon}\right) dx dy = (2\pi)^{-d} \int_{\mathbb{R}^d} f(x) \Phi_\epsilon(x) dx, \end{aligned}$$

donde hemos usado la propiedad de cambio de escala (260) y hemos definido

$$\Phi_\epsilon(x) := \int_{\mathbb{R}^d} \overline{f(y)} \epsilon^{-d} \mathcal{F}(\psi)\left(\frac{x-y}{\epsilon}\right) dy.$$

En general, si la función ψ que elegimos es razonable, la función Φ_ϵ converge a \overline{f} en $L^2(\mathbb{R}^d)$ cuando $\epsilon \rightarrow 0$. Tenemos un caso especialmente sencillo si elegimos

$$\psi(\xi) := e^{-\frac{|\xi|^2}{2}},$$

con lo cual

$$\Phi_\epsilon(x) := \int_{\mathbb{R}^d} \overline{f(y)} \epsilon^{-d} e^{-\frac{|x-y|^2}{2\epsilon^2}} dy.$$

□

Lema 13.8 (Inversa de la transformada de Fourier en L^2). Si $f \in L^1(\mathbb{R}^d) \cap L^2(\mathbb{R}^d)$,

$$\mathcal{F}^{-1}(\mathcal{F}(f)) = f.$$

Proposición 13.9 (Propiedades de la transformada de Fourier II). Para $f, g \in L^1(\mathbb{R}^d) \cap L^2(\mathbb{R}^d)$,

Producto

$$\mathcal{F}(fg) = (2\pi)^{-\frac{d}{2}} \mathcal{F}(f) * \mathcal{F}(g).$$

Demostración. La fórmula (261) también es cierta para la inversa \mathcal{F}^{-1} . Usando esto y el hecho de que \mathcal{F}^{-1} es la inversa de \mathcal{F} ,

$$\mathcal{F}^{-1}(\mathcal{F}(f) * \mathcal{F}(g)) = (2\pi)^{\frac{d}{2}} \mathcal{F}^{-1}(\mathcal{F}(f)) \mathcal{F}^{-1}(\mathcal{F}(g)) = (2\pi)^{\frac{d}{2}} fg.$$

Aplicando \mathcal{F} a los dos lados obtenemos el resultado.

□

14. La ecuación del calor

En esta sección vamos a estudiar una ecuación en derivadas parciales que sirve como modelo para varios fenómenos:

1. La difusión del calor en una habitación (en un gas) o en un sólido.
2. La difusión de una sustancia en un líquido en el que puede disolverse (por ejemplo, tinta en agua).

3. En ecología, la expansión de una especie en un ecosistema.

El comportamiento que tienen en común estos modelos se suele llamar *difusión*, y representa la expansión de una cierta sustancia a causa de que las partículas que la forman se mueven aleatoriamente (el calor se transmite a causa de los movimientos aleatorios de las moléculas; una sustancia se disuelve en otra por el mismo motivo; el movimiento de cada individuo de una especie tiene un componente aleatorio). Los principios básicos de los que partimos son los siguientes:

1. La masa total de la sustancia que consideramos se conserva.
2. La sustancia fluye de las regiones con más concentración a las regiones con menos concentración.

Vamos a representar por $u = u(t, x)$ la densidad de la sustancia en la posición $x \in \mathbb{R}^d$, en tiempo $t \geq 0$. Fíjate en que estamos eligiendo representar sólo la *densidad* de la sustancia (de la tinta en agua): u nos dice qué cantidad de tinta hay por centímetro cúbico en cada punto, por ejemplo. En particular, estamos renunciando a modelar *la posición de cada partícula*, porque hay demasiadas y el problema se vuelve demasiado complicado.

El primer punto nos dice que la variación de la masa en una cierta región $A \subseteq \mathbb{R}^d$ tiene que ser igual al flujo en su frontera:

$$\frac{d}{dt} \int_A u(t, x) dx = - \int_{\partial A} F(t, x) \cdot N(x) dS(x),$$

donde $N(x)$ es el normal exterior a A y $F(t, x)$ representa el flujo en un punto x . Por el punto 2 es razonable suponer que el flujo es proporcional a $-\nabla_x u$, ya que $-\nabla_x u$ representa la dirección más rápida para ir de mayor a menos concentración:

$$F(t, x) = -D\nabla u(t, x). \quad (263)$$

Supondremos por ahora que $D \in \mathbb{R}$ es una constante. Esto nos dice que

$$\begin{aligned} \frac{d}{dt} \int_A u(t, x) dx &= \int_{\partial A} D\nabla u(t, x) \cdot N(x) dS(x) \\ &= \int_A D \operatorname{div}(\nabla_x u(t, x)) dx = D \int_A \Delta_x u(t, x) dx. \end{aligned}$$

Si esto es cierto para toda región A vemos que

$$\partial_t u(t, x) = D\Delta_x u(t, x) \quad (264)$$

para todo $t \geq 0$, $x \in \mathbb{R}^d$. Esta ecuación se conoce como *ecuación del calor* o *ecuación de difusión*. Es una ecuación lineal que cumple el principio de superposición: si u, v son soluciones entonces las funciones $u + v$ y λu también lo son, para cualquier $\lambda \in \mathbb{R}$. Por supuesto, el concepto de solución (clásica) es el siguiente:

Definición 14.1. Sea $I \subseteq \mathbb{R}$ un intervalo abierto no vacío, $\Omega \subseteq \mathbb{R}^d$ un abierto no vacío. Decimos que $u: I \times \Omega \rightarrow \mathbb{R}$ es una *solución de la ecuación del calor en $I \times \Omega$* cuando

1. $\partial_t u$ existe en todo punto de $I \times \Omega$ y es continua en $I \times \Omega$,
2. $\partial_{x_n} u$, $\partial_{x_m} \partial_{x_n} u$ existen en todo punto de $I \times \Omega$, para todo $n, m = 1, \dots, d$, y son funciones continuas en $I \times \Omega$,

y u satisface (264) en todo punto $(t, x) \in I \times \Omega$.

Para acortar, a veces nos referimos a las condiciones 1 y 2 diciendo que u es \mathcal{C}^1 en t , \mathcal{C}^2 en x .

El primero de los supuestos que hemos hecho para obtener esta ecuación es que *la masa total debe conservarse*. La propiedad rigurosa que refleja esto es la siguiente:

Proposición 14.2 (Conservación de la masa). *Sea $I \subseteq \mathbb{R}$ un intervalo abierto no vacío, y sea u una solución de la ecuación de calor en $I \times \mathbb{R}^d$ tal que*

1. *Para cada $t \in I$ fijo la función $x \mapsto u(t, x)$ es integrable en \mathbb{R}^d*
2. *Existe una función integrable $M: \mathbb{R}^d \rightarrow \mathbb{R}$ tal que*

$$|\partial_t u(t, x)| \leq M(x) \quad \text{para todo } (t, x) \in I \times \mathbb{R}^d.$$

3. *Para cada $t \in I$ fijo la función $x \mapsto \nabla u(t, x)$ cumple que*

$$\lim_{R \rightarrow \infty} R^{d-1} \sup_{|x|=R} |\nabla u(t, x)| = 0.$$

Entonces

$$\int_{\mathbb{R}^d} u(t, x) \, dx = \int_{\mathbb{R}^d} u(s, x) \, dx \quad \text{para todo } s, t \in I. \quad (265)$$

Este resultado es muy incompleto y lo damos sólo como ejemplo de cómo se puede hacer un razonamiento riguroso. El motivo principal por el que es incompleto es que por ahora no sabemos si las soluciones de la ecuación del calor cumplen las condiciones del resultado (o si *alguna* lo cumple). Más adelante veremos que la conservación de la masa es cierta bajo condiciones más sencillas.

Demostración. Las condiciones 1 y 2 nos permiten derivar bajo la integral y obtener que

$$\frac{d}{dt} \int_{\mathbb{R}^d} u(t, x) \, dx = \int_{\mathbb{R}^d} \partial_t u(t, x) \, dx = \int_{\mathbb{R}^d} \Delta u(t, x) \, dx.$$

Por otra parte, la condición 3 nos permite decir, usando la fórmula de Green, que

$$\begin{aligned} \left| \int_{\mathbb{R}^d} \Delta u(t, x) \, dx \right| &= \left| \lim_{R \rightarrow +\infty} \int_{B_R} \Delta u(t, x) \, dx \right| = \left| \lim_{R \rightarrow +\infty} \int_{\partial B_R} \nabla u(t, x) \cdot N(x) \, dS(x) \right| \leq \\ &\lim_{R \rightarrow +\infty} \int_{\partial B_R} |\nabla u(t, x)| \, dS(x) \leq \omega_{d-1} \lim_{R \rightarrow +\infty} R^{d-1} \sup_{|x|=R} |\nabla u(t, x)| = 0. \end{aligned}$$

Por tanto $t \mapsto \int_{\mathbb{R}^d} u(t, x) \, dx$ es una función constante. \square

Teniendo en cuenta la derivación que hemos hecho de la ecuación parece razonable pensar que la ecuación del calor debe admitir una condición inicial: si sabemos la distribución de densidad de la sustancia en un cierto momento deberíamos poder conocerla en momentos posteriores. Esto es de hecho así: veremos más adelante que (bajo ciertas condiciones) la ecuación del calor tiene una única solución para una condición inicial dada. El *problema de valores iniciales* o *problema de Cauchy* correspondiente es

$$\begin{cases} \partial_t u = D\Delta u & \text{en } I \times \Omega, \\ u = g & \text{en } t = 0. \end{cases} \quad (266)$$

Observa que este problema en general requiere *una única condición en $t = 0$* , a diferencia de la ecuación de ondas. La ecuación de ondas es una consecuencia de la ley de Newton (y así fue como la dedujimos), y por tanto requiere la posición y velocidad iniciales como datos del problema. La difusión es un fenómeno distinto que tiene que ver con el comportamiento aleatorio de las partículas, y requiere sólo conocer sus posiciones como dato inicial.

Definición 14.3. Sea $I = [0, b) \subseteq \mathbb{R}$ un intervalo no vacío (con b posiblemente igual a ∞), $\Omega \subseteq \mathbb{R}^d$ un abierto no vacío, y $g : \Omega \rightarrow \mathbb{R}$ una función. Decimos que $u : I \times \Omega \rightarrow \mathbb{R}$ es una solución del problema (266) cuando

1. u es continua en $I \times \Omega$,
2. u es una solución de la ecuación del calor en $(a, b) \times \Omega$ (es decir, $u = u(t, x)$ es \mathcal{C}^1 en t , \mathcal{C}^2 en x , y es solución de la ecuación del calor en $(a, b) \times \Omega$),
3. y se cumple que $u(0, x) = g(x)$ para todo $x \in \Omega$.

14.1. Solución fundamental

Al igual que hicimos en el caso de la ecuación de ondas y la de Laplace, empezamos por buscar algunas soluciones particulares. La ecuación puede plantearse en un dominio cuadrado (en lugar de todo \mathbb{R}^d), y podemos buscar soluciones en variables separadas, pero en lugar de eso vamos a empezar por buscar soluciones particulares de una forma específica, motivada por el siguiente resultado:

Lema 14.4. Sea u una solución de la ecuación del calor en $(0, +\infty) \times \mathbb{R}^d$.

1. Para $\mu > 0$ fijo la función $v : (0, +\infty) \times \mathbb{R}^d \rightarrow \mathbb{R}$ definida por

$$v(t, x) := \mu u(t, x), \quad (t, x) \in (0, +\infty) \times \mathbb{R}^d,$$

es también una solución en $(0, +\infty) \times \mathbb{R}^d$.

2. Para $\lambda > 0$ fijo la función $v : (0, +\infty) \times \mathbb{R}^d \rightarrow \mathbb{R}$ definida por

$$v(t, x) := u(\lambda^2 t, \lambda x), \quad (t, x) \in (0, +\infty) \times \mathbb{R}^d,$$

es también una solución en $(0, +\infty) \times \mathbb{R}^d$.

3. Si $R: \mathbb{R}^d \rightarrow \mathbb{R}^d$ es un movimiento rígido entonces la función

$$v(t, x) := u(t, R(x)), \quad (t, x) \in (0, \infty) \times \mathbb{R}^d,$$

es también una solución.

Cuando las soluciones de una ecuación tienen una propiedad de invarianza frente a una familia de transformaciones, una pregunta natural es si existe alguna solución que sea invariante frente a esas mismas transformaciones o alguna subfamilia de ellas. En este caso nos preguntamos si existe alguna solución u de la ecuación del calor que cumpla

$$u(t, x) = u(\lambda^2 t, \lambda x).$$

para todo $(t, x) \in (0, +\infty) \times \mathbb{R}^d$, $\lambda > 0$. Eligiendo $\lambda = 1/\sqrt{t}$ vemos que

$$u(t, x) = u\left(1, \frac{x}{\sqrt{t}}\right) =: v\left(\frac{x}{\sqrt{t}}\right).$$

Una solución de esta forma (que de hecho existe) no puede satisfacer la conservación de la masa; para arreglar esto buscamos una solución de la forma

$$u(t, x) = t^{-d/2} v\left(\frac{x}{\sqrt{t}}\right).$$

Podemos también imponer que la solución sea simétrica por rotaciones; es decir, que sea de la forma

$$u(t, x) = t^{-d/2} v\left(\frac{x}{\sqrt{t}}\right) = t^{-d/2} \phi\left(\frac{|x|}{\sqrt{t}}\right).$$

Usando esta forma en la ecuación (264) y llamando $y := x/\sqrt{t}$ vemos que v tiene que cumplir

$$-\frac{d}{2} t^{-\frac{d}{2}-1} v(y) - t^{-\frac{d}{2}} \nabla v(y) \frac{x}{2t^{3/2}} = t^{-\frac{d}{2}-1} \Delta v(y)$$

o, escrito en función de y ,

$$-\frac{d}{2} t^{-\frac{d}{2}-1} v(y) - t^{-\frac{d}{2}-1} \nabla v(y) \frac{y}{2} = t^{-\frac{d}{2}-1} \Delta v(y)$$

Podemos quitar el factor común $t^{-\frac{d}{2}-1}$ y obtener

$$-\frac{d}{2} v(y) - \nabla v(y) \frac{y}{2} = \Delta v(y). \quad (267)$$

Ahora, si llamamos

$$r := \frac{|x|}{\sqrt{t}} = |y|$$

tenemos que

$$\nabla v(y) = \frac{y}{|y|} \phi'(r), \quad \Delta v(y) = \phi''(r) + \frac{d-1}{r} \phi'(r).$$

con lo que de (267) obtenemos

$$-\frac{d}{2} \phi(r) - \frac{r}{2} \phi'(r) = \phi''(r) + \frac{d-1}{r} \phi'(r)$$

Esto es una EDO lineal de segundo orden para ϕ . En principio no parece fácil de resolver pero al multiplicar por r^{d-1} se convierte en

$$(r^{d-1}\phi')' + \frac{1}{2}(r^d\phi)' = 0,$$

que ahora sí puede resolverse. Esta ecuación implica que

$$r^{d-1}\phi' + \frac{1}{2}r^d\phi = C$$

para cierta constante $C \in \mathbb{R}$. Podemos resolver esta ecuación explícitamente y vemos que la única solución acotada se tiene para $C = 0$, y es la siguiente:

$$\phi(r) = Ae^{-\frac{r^2}{4}}$$

para cierta constante $A \in \mathbb{R}$, es decir,

$$u(t, x) = \frac{A}{t^{d/2}} e^{-\frac{|x|^2}{4t}}.$$

Haciendo una elección particular de la constante definimos lo siguiente:

Definición 14.5. La función $\Phi: (0, +\infty) \times \mathbb{R}^d \rightarrow \mathbb{R}$ definida por

$$\Phi(t, x) := \frac{1}{(4\pi t)^{d/2}} e^{-\frac{|x|^2}{4t}}, \quad t > 0, x \in \mathbb{R}^d. \quad (268)$$

se llama *solución fundamental de la ecuación del calor en \mathbb{R}^d* .

Ejercicio 14.1. 1. Comprueba que la función Φ es efectivamente una solución de la ecuación del calor (en $(0, \infty) \times \mathbb{R}^d$).

2. Comprueba que

$$\int_{\mathbb{R}^d} \Phi(t, x) dx = 1 \quad \text{para todo } t > 0. \quad (269)$$

Teorema 14.6. Sea $g: \mathbb{R}^d \rightarrow \mathbb{R}$ una función continua y acotada (i.e., para cierta $C > 0$ se tiene que $|g(x)| \leq C$ para todo $x \in \mathbb{R}^d$). La función

$$u(t, x) := \int_{\mathbb{R}^d} \Phi(t, x - y)g(y) dy, \quad (t, x) \in (0, +\infty) \times \mathbb{R}^d \quad (270)$$

es una solución de la ecuación del calor en $(0, +\infty) \times \mathbb{R}^d$ que cumple

$$\lim_{(t,x) \rightarrow (0,x_0)} u(t, x) = g(x_0) \quad \text{para todo } x_0 \in \mathbb{R}^d. \quad (271)$$

Es decir: u puede extenderse de forma única a una función definida en $[0, +\infty) \times \mathbb{R}^d$, y dicha extensión es una solución del problema (270) (con $D = 1$).

Demostración. Para ver que u es una solución de la ecuación del calor en $(0, +\infty) \times \mathbb{R}^d$ podemos aplicar directamente el teorema de derivación bajo la integral en cualquier conjunto de la forma $(a, b) \times B_R$, donde $0 < a < b$ y B_R es la bola de centro 0 y radio R en \mathbb{R}^d .

La parte crucial de la demostración es ver que se cumple (271). Tomemos $x_0 \in \mathbb{R}^d$. Para $\epsilon > 0$ dado, elijamos $\delta > 0$ tal que

$$|g(x) - g(x_0)| \leq \frac{\epsilon}{2} \quad \text{siempre que } |x - x_0| < \delta.$$

(Podemos hacer esto porque g es continua.) Entonces, usando (269) tenemos que

$$\begin{aligned} |u(t, x) - g(x_0)| &= \left| \int_{\mathbb{R}^d} \Phi(t, x - y) g(y) \, dy - g(x_0) \right| \\ &= \left| \int_{\mathbb{R}^d} \Phi(t, x - y) (g(y) - g(x_0)) \, dy \right| \leq \int_{\mathbb{R}^d} \Phi(t, x - y) |g(y) - g(x_0)| \, dy \end{aligned}$$

para todo $x \in \mathbb{R}^d$ y todo $t > 0$. Partimos esta integral en dos trozos:

$$\begin{aligned} \int_{\mathbb{R}^d} \Phi(t, x - y) |g(y) - g(x_0)| \, dy &= \int_{B_\delta(x_0)} \Phi(t, x - y) |g(y) - g(x_0)| \, dy \\ &\quad + \int_{\mathbb{R}^d \setminus B_\delta(x_0)} \Phi(t, x - y) |g(y) - g(x_0)| \, dy =: I_1 + I_2. \end{aligned}$$

Para la primera integral tenemos $|g(y) - g(x_0)| \leq \epsilon$, luego

$$I_1 \leq \epsilon \int_{\mathbb{R}^d} \Phi(t, x - y) \, dy = \epsilon.$$

Para la segunda, vamos a suponer que $|x - x_0| < \delta/2$. Entonces para $y \in \mathbb{R}^d \setminus B_\delta(x_0)$ tenemos $|x - x_0| < |y - x_0|/2$, luego

$$|x - y| \geq |y - x_0| - |x_0 - x| \geq \frac{|y - x_0|}{2},$$

lo que implica

$$\Phi(t, x - y) = \frac{1}{(4\pi t)^{d/2}} e^{-\frac{|x-y|^2}{4t}} \leq \frac{1}{(4\pi t)^{d/2}} e^{-\frac{|y-x_0|^2}{16t}}.$$

Entonces para $|x - x_0| < \delta/2$ se cumple

$$\begin{aligned} I_2 &= \int_{\mathbb{R}^d \setminus B_\delta(x_0)} \Phi(t, x - y) |g(y) - g(x_0)| \, dy \leq \frac{2\|g\|_\infty}{(4\pi t)^{d/2}} \int_{\mathbb{R}^d \setminus B_\delta(x_0)} e^{-\frac{|y-x_0|^2}{16t}} \, dy \\ &= \frac{2\|g\|_\infty}{(4\pi t)^{d/2}} \int_{\mathbb{R}^d \setminus B_\delta(0)} e^{-\frac{|y|^2}{16t}} \, dy. \end{aligned}$$

Se puede comprobar que

$$\mu_\delta(t) := \frac{1}{t^{d/2}} \int_{\mathbb{R}^d \setminus B_\delta(0)} e^{-\frac{|y|^2}{16t}} \, dy \rightarrow 0 \quad \text{cuando } t \rightarrow 0. \quad (272)$$

Hemos demostrado finalmente que para cada $\epsilon > 0$ fijo existe $\delta > 0$ tal que

$$|g(x) - g(x_0)| \leq \epsilon + C\mu_\delta(t) \quad \text{siempre que } |x - x_0| < \delta/2,$$

para cierta constante $C > 0$. Esto prueba (271) y termina la demostración \square

Ejercicio 14.2. 1. Escribe completamente el argumento del principio de la demostración anterior para probar que u es solución de la ecuación del calor en $(0, +\infty) \times \mathbb{R}^d$, comprobando las condiciones del teorema de derivación bajo la integral.

2. Demuestra el límite (272).

Ejercicio 14.3. Supongamos que, además de las condiciones del Teorema 14.6, la función g está en $L^1(\mathbb{R}^d)$. Demuestra rigurosamente que la solución dada en el Teorema 14.6 conserva la masa.

14.2. Solución usando la transformada de Fourier

Consideramos de nuevo la ecuación del calor en \mathbb{R}^d con condición inicial dada:

$$\begin{cases} \partial_t u = \Delta u & \text{en } (0, +\infty) \times \mathbb{R}^d, \\ u = g & \text{en } t = 0. \end{cases} \quad (273)$$

Una forma de resolverla es calcular su transformada de Fourier *en la variable x* . Supongamos que hay una solución u tal que $x \mapsto u(t, x)$ es suficientemente regular (por ejemplo, está en $L^1(\mathbb{R}^d) \cap L^2(\mathbb{R}^d)$, y es una función $\mathcal{C}^1(\mathbb{R}^d)$ cuyas derivadas están también en $L^1(\mathbb{R}^d) \cap L^2(\mathbb{R}^d)$). Usando las propiedades de la transformada de Fourier de la sección 13 tenemos

$$\partial_t \hat{f}(t, \xi) = -|\xi|^2 \hat{f}(t, \xi), \quad t \geq 0, \xi \in \mathbb{R}^d.$$

Para cada $\xi \in \mathbb{R}^d$ fijo esto es una ecuación diferencial *ordinaria*. Su solución es

$$\hat{f}(t, \xi) = \hat{f}(0, \xi) e^{-t|\xi|^2} = \hat{g}(\xi) e^{-t|\xi|^2}.$$

Podemos usar ahora la propiedad del producto de la transformada inversa para decir que

$$\begin{aligned} f(t, x) &= \mathcal{F}^{-1}(\hat{g}(\xi) e^{-t|\xi|^2}) = (2\pi)^{-\frac{d}{2}} g(x) * \mathcal{F}^{-1}(e^{-t|\xi|^2}) \\ &= (4\pi t)^{-\frac{d}{2}} g(x) * e^{\frac{|x|^2}{4t}} = (4\pi t)^{-\frac{d}{2}} \int_{\mathbb{R}^d} g(y) e^{\frac{|x-y|^2}{4t}} dy. \end{aligned}$$

Observa que esta fórmula es por supuesto la misma que obtuvimos en la sección anterior (Teorema 14.6). Falta demostrar que es efectivamente una solución (suponiendo que g es suficientemente regular), lo cual ya hicimos en el Teorema 14.6.

De forma muy parecida, la transformada de Fourier puede usarse para resolver explícitamente algunas ecuaciones diferenciales con coeficientes constantes.

Ejercicio 14.4. Encuentra la solución (clásica y que conserve la masa) de la ecuación

$$\partial_t u = \Delta u - u,$$

donde $u = u(t, x)$ con $t \geq 0$, $x \in \mathbb{R}$, con condición inicial

$$u(0, x) = g(x) := e^{-x^2/2}, \quad x \in \mathbb{R}^d.$$

(Sugerencia: el ejercicio se puede resolver usando la transformada de Fourier).

Solución. Si aplicamos la transformada de Fourier en la variable x obtenemos

$$\partial_t \hat{u}(t, \xi) = -(\xi^2 + 1)\hat{u}(t, \xi),$$

lo cual implica que

$$\hat{u}(t, \xi) = \hat{u}(0, \xi)e^{-(\xi^2+1)t} = \hat{u}(0, \xi)e^{-\xi^2 t}e^{-t}.$$

Sabemos que $\hat{u}(0, \xi) = \hat{g}(\xi) = e^{-\xi^2/2}$, luego

$$\hat{u}(t, \xi) = e^{-\xi^2(t+\frac{1}{2})}e^{-t}.$$

Tomando la transformada de Fourier inversa y recordando la transformada de una Gaussiana,

$$u(t, x) = e^{-t}(2(t + \frac{1}{2}))^{-\frac{d}{2}}e^{-\frac{x^2}{4t+2}} = (2t+1)^{-\frac{d}{2}}e^{-\frac{x^2}{4t+2}}e^{-t}.$$

Un método alternativo para encontrar la misma solución es el siguiente: si hacemos el cambio

$$v(t, x) := e^t u(t, x)$$

vemos que $\partial_t v = \Delta v$. Es decir, v es solución de la ecuación del calor en \mathbb{R}^d . Podemos entonces usar la fórmula (270) y obtener

$$v(t, x) = (4\pi t)^{-\frac{d}{2}} \int_{\mathbb{R}^d} e^{-\frac{|y|^2}{2}} e^{-\frac{|x-y|^2}{4t}} dy. \quad (274)$$

Esta última integral se puede calcular explícitamente: en general, para cualesquiera $a, b > 0$, completando cuadrados vemos que

$$a|y|^2 + b|x-y|^2 = \left| y\sqrt{a+b} - \frac{b}{\sqrt{a+b}} \right|^2 + \frac{ab}{a+b}|x|^2.$$

Usando esto podemos calcular la convolución de dos Gaussianas:

$$\begin{aligned} \int_{\mathbb{R}^d} e^{-a|y|^2} e^{-b|x-y|^2} dy &= e^{-\frac{ab}{a+b}|x|^2} \int_{\mathbb{R}^d} e^{-\left| y\sqrt{a+b} - \frac{b}{\sqrt{a+b}} \right|^2} dy \\ &= e^{-\frac{ab}{a+b}|x|^2} \int_{\mathbb{R}^d} e^{-|y\sqrt{a+b}|^2} dy = (a+b)^{-\frac{d}{2}} e^{-\frac{ab}{a+b}|x|^2} \int_{\mathbb{R}^d} e^{-|z|^2} dz \\ &= \pi^{\frac{d}{2}} (a+b)^{-\frac{d}{2}} e^{-\frac{ab}{a+b}|x|^2}, \end{aligned}$$

donde hemos usado el cambio $z = y\sqrt{a+b}$. Si usamos esto con $a := \frac{1}{2}$, $b := \frac{1}{4t}$ obtenemos

$$\int_{\mathbb{R}^d} e^{-\frac{|y|^2}{2}} e^{-\frac{|x-y|^2}{4t}} dy = \pi^{\frac{d}{2}} \left(\frac{2t+1}{4t} \right)^{-\frac{d}{2}} e^{-\frac{|x|^2}{4t+2}}.$$

Usando esto en (274),

$$v(t, x) = (4\pi t)^{-\frac{d}{2}} \pi^{\frac{d}{2}} \left(\frac{2t+1}{4t} \right)^{-\frac{d}{2}} e^{-\frac{|x|^2}{4t+2}} = (2t+1)^{-\frac{d}{2}} e^{-\frac{|x|^2}{4t+2}}.$$

Deshaciendo el cambio $u = e^{-t}v$ obtenemos la misma fórmula que antes.

14.3. Solución en un intervalo: separación de variables

Ahora consideramos la ecuación del calor en el intervalo $[0, L]$, con condiciones de frontera tipo Dirichlet. Llamando $\Omega := (0, L)$, queremos resolver la ecuación

$$\begin{cases} \partial_t u = \Delta u & \text{en } (0, +\infty) \times \Omega, \\ u = g & \text{en } t = 0, \\ u = 0 & \text{en } \partial\Omega, \end{cases} \quad (275)$$

Denotamos las variables de u como $u = u(t, x)$, con $t \geq 0$, $x \in [0, L]$. Buscamos primero soluciones en variables separadas *sin tener en cuenta la condición inicial*; esto es, soluciones del tipo

$$u(t, x) = \varphi(t)\psi(x).$$

La ecuación en derivadas parciales queda entonces

$$\varphi'(t)\psi(x) = \varphi(t)\psi''(x),$$

lo cual sugiere que debe existir una constante $\lambda \in \mathbb{R}$ tal que

$$\varphi'(t) = -\lambda\varphi(t), \quad \psi''(x) = -\lambda\psi(x),$$

y la condición de contorno implica que

$$\psi(0) = \psi(L) = 0.$$

La ecuación para ψ es un problema de Sturm-Liouville que hemos resuelto varias veces, y sabemos que sus soluciones no triviales son un espacio vectorial generado por

$$\psi_n(x) = \sin\left(\frac{\pi nx}{L}\right),$$

donde $n \geq 1$ es un entero y ψ_n correspondiente al valor

$$\lambda = \lambda_n = \frac{\pi^2 n^2}{L^2}.$$

Vemos entonces que las soluciones correspondientes para φ son

$$\varphi_n(t) = e^{-\lambda_n t} \varphi(0).$$

En resumen, hemos encontrado las siguientes soluciones en variables separadas de (275):

$$u_n(t, x) = \exp\left(-\frac{\pi^2 n^2 t}{L^2}\right) \sin\left(\frac{\pi nx}{L}\right), \quad t \geq 0, x \in [0, L].$$

Todas ellas cumplen además la condición de contorno de tipo Dirichlet. Ahora queremos encontrar una combinación lineal de las u_n que cumpla además la condición inicial: si escribimos

$$u = \sum_{n=1}^{\infty} A_n u_n,$$

entonces

$$u(0, x) = \sum_{n=1}^{\infty} A_n \sin\left(\frac{\pi nx}{L}\right),$$

lo cual nos dice que debemos tomar como A_n los coeficientes de la serie de Fourier en senos de g .

Teorema 14.7. Sea $L > 0$, $g \in H^3([0, L])$ tal que $g(0) = g(L) = g'(0) = g'(L)$, y sean b_n , los coeficientes de su serie de Fourier en senos:

$$g = \sum_{n=1}^{\infty} b_n \operatorname{sen} \left(\frac{\pi n x}{L} \right), \quad (276)$$

(La igualdad se entiende como límite en $L^2([0, L])$.) Entonces el problema

$$\begin{cases} \partial_t u = \Delta u & \text{en } (0, +\infty) \times \Omega, \\ u = g & \text{en } t = 0, \\ u = 0 & \text{en } \partial\Omega, \end{cases} \quad (277)$$

tiene una solución dada por

$$u(t, x) := \sum_{n=1}^{\infty} b_n \exp \left(-\frac{\pi^2 n^2 t}{L^2} \right) \operatorname{sen} \left(\frac{\pi n x}{L} \right), \quad t \geq 0, x \in [0, L]. \quad (278)$$

14.4. Solución en el cuadrado: separación de variables

Podemos hacer lo mismo que en la sección anterior para la ecuación del calor en el cuadrado $[0, L] \times [0, L]$, con condiciones de frontera tipo Dirichlet. Llamando $\Omega := (0, L) \times (0, L)$, queremos resolver la ecuación

$$\begin{cases} \partial_t u = \Delta u & \text{en } (0, +\infty) \times \Omega, \\ u = g & \text{en } t = 0, \\ u = 0 & \text{en } \partial\Omega, \end{cases} \quad (279)$$

Denotamos las variables de u como $u = u(t, x, y)$, con $t \geq 0$, $x, y \in [0, L]$. Buscamos primero soluciones en variables separadas, esto es, soluciones del tipo

$$u(t, x, y) = \varphi(t) \psi_1(x) \psi_2(y)$$

15. La ecuación de Schrödinger

Una de las ecuaciones fundamentales en mecánica cuántica es la *ecuación de Schrödinger*, que modela el comportamiento de una partícula en un campo eléctrico:

$$i\hbar \partial_t \Psi = -\frac{\hbar^2}{2m} \Delta \Psi + V \Psi, \quad (280)$$

donde

- t representa el tiempo, $x \in \mathbb{R}^3$ representa el espacio.
- $\Psi = \Psi(t, x) \in \mathbb{C}$ es la llamada *función de onda*. Una de las diferencias importantes con otras EDPs que hemos visto es que aquí $\Psi(t, x)$ puede tomar valores complejos.
- $V = V(x)$ es el potencial eléctrico.
- $\hbar = \frac{h}{2\pi} = 1.05 \text{ J s}$ es la *constante de Planck reducida*; $h = 6.63 \times 10^{-34} \text{ J s}$ es la constante de Planck.

- m es la masa de la partícula.

La interpretación de la función de onda Ψ es la siguiente: si medimos la posición de la partícula en cierto momento t , entonces la distribución que da la probabilidad de encontrarla en un punto $x \in \mathbb{R}^3$ es $|\Psi(t, x)|^2$. De hecho, la distribución de probabilidad del resultado de cualquier medida que hagamos sobre la partícula puede obtenerse a partir de Ψ .

Esta ecuación no puede deducirse a partir de principios fundamentales, y es uno de los supuestos básicos de la teoría cuántica. Hay formas más generales de esta ecuación que también se conocen como “ecuación de Schrödinger” en varios contextos, pero ésta es posiblemente el caso más extendido. A partir de esta ecuación pueden deducirse gran cantidad de propiedades de sistemas cuánticos sencillos, y en particular las propiedades del átomo de hidrógeno. Hay versiones más complejas de la ecuación que modelan átomos más complicados con varios electrones, pero no vamos a tratarlas aquí. En esta sección vamos a ver cómo usar esta ecuación para deducir algunas propiedades cuánticas sencillas.

Una primera observación es que esta ecuación puede obtenerse como ecuación de Euler-Lagrange del siguiente funcional:

$$\mathcal{F}(\Psi) := \int_{t_0}^{t_1} \int_{\mathbb{R}^3} \left(\frac{i\hbar}{2} (\Psi^* \partial_t \Psi - \Psi \partial_t \Psi^*) - \frac{\hbar^2}{2m} |\nabla \Psi|^2 - V |\Psi|^2 \right) dx dt.$$

Una de las ideas que llevó a la primera deducción de la ecuación de Schrödinger fue la siguiente analogía: de la misma forma que la óptica geométrica no es una descripción exacta del comportamiento de la luz, la mecánica de Newton tampoco lo es, y requiere una modificación del mismo tipo. La óptica geométrica se deduce del principio de Fermat, pero una teoría más realista viene dada por la ecuación de ondas. De la misma forma, la teoría de Newton se deduce del principio de acción estacionaria, y una teoría más realista debe obtenerse a partir de un principio comparable para una ecuación de ondas. Esto es por supuesto un razonamiento no riguroso, pero tiene una importancia histórica porque fue la idea que llevó al descubrimiento de (280).

Matemáticamente, la primera técnica que podemos intentar para resolver (280) es la separación de variables. Si buscamos una solución del tipo

$$\Psi(t, x) = f(t)\Phi(x),$$

entonces vemos que

$$i\hbar \frac{\partial_t f(t)}{f(t)} = -\frac{\hbar^2}{2m} \frac{\Delta \Phi(x)}{\Phi(x)} + V(x),$$

(siempre que $f(t) \neq 0$, $\Phi(x) \neq 0$), lo cual sugiere que debe ocurrir

$$i\hbar \partial_t f(t) = E f(t), \tag{281}$$

$$-\frac{\hbar^2}{2m} \Delta \Phi(x) + V(x)\Phi(x) = E\Phi(x) \tag{282}$$

para cierta constante E . La primera ecuación tiene la solución inmediata

$$f(t) = \exp\left(\frac{-iEt}{\hbar}\right),$$

y la segunda es una ecuación muy conocida que se llama *ecuación de Schrödinger estacionaria* (y a veces se le llama también ecuación de Schrödinger simplemente).

Adimensionalización. Podemos elegir las unidades de forma conveniente para simplificar la ecuación un poco. Se suelen elegir las unidades de manera que $\hbar = m = e = 1$ (donde $-e$ es la carga del electrón); las unidades resultantes se llaman *unidades atómicas*. En estas unidades la ecuación queda

$$i\partial_t\Psi = -\frac{1}{2}\Delta\Psi + V\Psi, \quad (283)$$

y la ecuación estacionaria (282) es ahora

$$-\frac{1}{2}\Delta\Phi + V\Phi = E\Phi. \quad (284)$$

Una vez hechos los cálculos de esta forma es fácil traducir los resultados a las dimensiones usuales, de forma que trabajaremos siempre con estas expresiones.

Propiedades básicas

1. Linealidad - principio de superposición.
2. Reversibilidad en tiempo (frente a conjugación).
3. Conservación de $\int |\psi|^2$.

15.1. Ejemplos en dimensión 1

Para entender el comportamiento de la ecuación (280) podemos estudiar la misma ecuación, pero en dimensión 1, lo cual es una especie de “modelo de juguete” con el que los cálculos son más fáciles. Vamos a considerar la ecuación estacionaria:

$$-\frac{1}{2}\Phi'' + V\Phi = E\Phi, \quad (285)$$

donde $\Phi: \mathbb{R} \rightarrow \mathbb{C}$ es la incógnita.

15.1.1. Partícula libre

El ejemplo más sencillo es el caso con $V = 0$,

$$-\frac{1}{2}\Phi'' = E\Phi. \quad (286)$$

Normalmente buscamos soluciones tales que $|\Phi|^2$ sea integrable con integral 1. En este caso no hay soluciones así; las soluciones acotadas son

$$\Phi(x) = A \operatorname{sen}(x\sqrt{2E}) + B \cos(x\sqrt{2E}).$$

15.1.2. Partícula en una caja

Supongamos que tenemos una partícula confinada a una “caja” de longitud L . Un modelo para esto es suponer que el potencial es $V = 0$ en el intervalo $[-L/2, L/2]$ y $V = +\infty$ fuera de $[-L/2, L/2]$, de forma que ninguna partícula clásica (con ninguna energía) puede salir de esa región. En este caso buscamos una solución $\Phi: [-L/2, L/2] \rightarrow \mathbb{R}$ que cumpla

$$-\frac{1}{2}\Phi'' = E\Phi, \quad x \in [-L/2, L/2], \quad (287)$$

con la condición

$$\Phi(-L/2) = \Phi(L/2) = 0.$$

Las soluciones de esta ecuación se pueden encontrar de forma directa. Son

$$\begin{aligned} \Phi(x) &= A \cos\left(\frac{n\pi x}{L}\right), & \text{para } n > 0 \text{ impar,} \\ \Phi(x) &= B \sin\left(\frac{n\pi x}{L}\right), & \text{para } n > 0 \text{ par,} \end{aligned}$$

que corresponden a los valores

$$E = \frac{n^2\pi^2}{2L^2}.$$

15.1.3. Partícula en un pozo de potencial finito

Ahora consideramos

$$V(x) = \begin{cases} v, & |x| > L/2, \\ 0, & |x| \leq L/2, \end{cases}$$

y queremos resolver la ecuación

$$-\frac{1}{2}\Phi'' + V\Phi = E\Phi, \quad x \in \mathbb{R}. \quad (288)$$

Es decir:

$$\begin{aligned} -\frac{1}{2}\Phi'' &= (E - v)\Phi, & |x| > L/2, \\ -\frac{1}{2}\Phi'' &= E\Phi, & |x| \leq L/2. \end{aligned}$$

Vamos a buscar por el momento solamente soluciones tales que $|\Phi|^2$ sea integrable en \mathbb{R} , que corresponden a estados confinados de la partícula, y que sean derivables en sentido débil al menos dos veces, de forma que podamos decir que se cumple la ecuación en sentido débil. Para esta última condición de regularidad, en la práctica lo que haremos será exigir que la solución sea de clase \mathcal{C}^1 .

La primera observación es que cada espacio propio de este problema tiene que tener dimensión 1, ya que el espacio de soluciones de una EDO de segundo orden tiene dimensión 2; si imponemos la condición de que $\lim_{x \rightarrow -\infty} \Phi(x) = 0$ el espacio resultante tiene dimensión 1 (¿puedes hacer este argumento de forma rigurosa?).

La segunda observación que hacemos es que las funciones propias tienen que ser simétricas o antisimétricas, por el siguiente motivo: si Φ es una función propia que no es antisimétrica, entonces

$$\tilde{\Phi}(x) := \Phi(x) + \Phi(-x)$$

es una función propia asociada al mismo valor propio, que además es simétrica. Como los espacios propios tienen dimensión 1, deducimos que Φ tiene que ser simétrica.

Esto simplifica mucho nuestro problema. Buscamos primero las funciones propias simétricas. Para que sean de cuadrado integrable debe ocurrir que $E - v < 0$ y

$$\Phi(x) = Ae^{-\alpha|x|}, \quad |x| \geq L/2,$$

con $\alpha := \sqrt{-2(E - v)}$. Dentro del intervalo $[-L/2, L/2]$ debe ocurrir que $E > 0$ y

$$\Phi(x) = B \cos(\omega x),$$

donde $\omega := \sqrt{2E}$. (No hay funciones propias con $E \leq 0$, ¿puedes demostrarlo?). Como la función Φ debe ser \mathcal{C}^1 , deben cumplirse las condiciones

$$Ae^{-\alpha L/2} = B \cos(\omega L/2), \quad -A\alpha e^{-\alpha L/2} = -B\omega \sin(\omega L/2).$$

Estas ecuaciones tienen solución no trivial cuando

$$0 = \det \begin{pmatrix} e^{-\alpha L/2} & \cos(\omega L/2) \\ -\alpha e^{-\alpha L/2} & -\omega \sin(\omega L/2) \end{pmatrix} = -e^{-\alpha L/2} (\omega \sin(\omega L/2) - \alpha \cos(\omega L/2)).$$

Es decir, cuando

$$\tan(\omega L/2) = \frac{\alpha}{\omega},$$

o escrito de otra forma,

$$\tan \left(\frac{L\sqrt{2E}}{2} \right) = \frac{\sqrt{-2(E - v)}}{\sqrt{2E}},$$

Esta ecuación tiene un número finito de soluciones, que no podemos escribir de forma analítica.

15.2. La partícula libre

- * Ondas planas
- * Solución fundamental
- * Solución con condición inicial dada

15.3. El átomo de hidrógeno

Para un átomo de hidrógeno el potencial que consideramos es el generado por el protón del núcleo:

$$V(x) = -\frac{Ze^2}{4\pi\epsilon_0|x|} = -\frac{Z}{|x|},$$

donde la expresión simplificada es debido al uso de unidades atómicas. La correspondiente ecuación de Schrödinger estacionaria es entonces

$$-\frac{1}{2}\Delta\Phi(x) - \frac{Z}{|x|}\Phi(x) = E\Phi(x). \quad (289)$$

Para resolverla vamos a buscar soluciones en variables separadas, pero escribiendo Φ en coordenadas esféricas. Las coordenadas esféricas estándar (r, θ, φ) (ver Sección B.4) vienen dadas por

$$\begin{aligned}x &= r \sin \varphi \cos \theta \\y &= r \sin \varphi \sin \theta \\z &= r \cos \varphi\end{aligned}$$

y el operador laplaciano puede expresarse en coordenadas esféricas como

$$\Delta \Phi = \frac{1}{r^2} \partial_r (r^2 \partial_r \Phi) + \frac{1}{r^2 \sin \varphi} \partial_\varphi (\sin \varphi \partial_\varphi \Phi) + \frac{1}{r^2 (\sin \varphi)^2} \partial_\theta^2 \Phi.$$

Si buscamos una solución de la forma

$$\Phi(x) = R(r)Y(\varphi, \theta)$$

entonces debe ocurrir que

$$-\frac{1}{2} \partial_r (r^2 \partial_r R) - r Z R = E r^2 R - \lambda R, \quad (290)$$

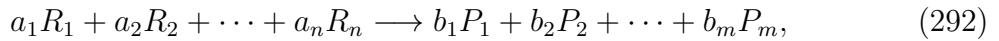
$$\frac{1}{\sin \varphi} \partial_\varphi (\sin \varphi \partial_\varphi Y) + \frac{1}{(\sin \varphi)^2} \partial_\theta^2 Y = \lambda Y \quad (291)$$

para cierto $\lambda \in \mathbb{R}$. Empecemos por estudiar la ecuación (290). Podemos simplificarla un poco como

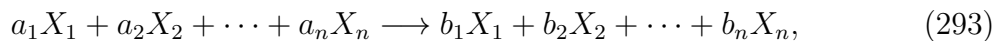
$$-\frac{1}{2} R'' - \frac{1}{r} R' - \frac{1}{r} Z R = E R - \frac{\lambda}{r^2} R.$$

16. Ley de acción de masas

La ley de acción de masas es un principio general que debe cumplir la velocidad a la que ocurre una reacción química. Es una que se cumple con gran aproximación en muchos contextos, pero no es una ley física exacta. Una reacción química es un proceso en el que intervienen ciertas sustancias llamadas *reactivos* o *reactantes*, que se transforman en otras sustancias llamadas *productos*. Si denotamos las sustancias que actúan como reactivos por R_1, R_2, \dots, R_n y las sustancias que se obtienen como productos por P_1, P_2, \dots, P_m , normalmente denotamos la reacción química por



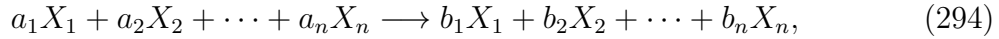
donde los números $a_1, \dots, a_n, b_1, \dots, b_m$ son enteros no negativos llamados *coeficientes estequiométricos*, que representan el número de moléculas de cada reactivo y cada producto que intervienen en la reacción. Por supuesto, algunos de los productos pueden ser las mismas sustancias que algunos de los reactivos, así que podemos escribir una reacción general como



donde varios de los coeficientes pueden ser 0.

Ley de acción de masas: *La velocidad a la que ocurre una reacción química es proporcional a cada uno de los reactivos que intervienen en la reacción.*

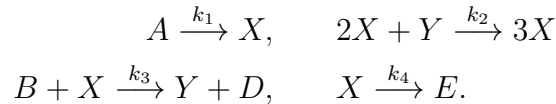
Si denotamos las concentraciones de cada una de las sustancias X_i por x_i , la ley de acción de masas nos dice que la velocidad de la reacción (293) es proporcional a $x_1^{a_1} x_2^{a_2} \dots x_n^{a_n}$. La constante de proporcionalidad se llama *constante de la reacción*. A veces, si queremos especificar la constante de reacción, escribimos la reacción química como



donde $\lambda > 0$ es la constante de reacción. En general, debido a la reacción (294), la derivada en tiempo de la concentración x_i de la sustancia X_i es

$$\frac{d}{dt} x_i = \lambda (b_i - a_i) x_1^{a_1} x_2^{a_2} \dots x_n^{a_n}.$$

Ejercicio 16.1. Consideramos el siguiente sistema de reacciones químicas, con constantes de reacción $k_1, k_2, k_3, k_4 > 0$:



1. Usando la ley de acción de masas, escribe la ecuación diferencial ordinaria que satisfacen las concentraciones x, y de las especies X, Y .
2. Suponiendo que las concentraciones de A y B son constantes (e ignorando la concentración de D y E), encuentra todos los posibles valores de equilibrio de x, y .

Solución. 1. Las ecuaciones obtenidas de la ley de acción de masas son

$$\begin{aligned} x' &= k_1 a + k_2 x^2 y - k_3 b x - k_4 x, \\ y' &= -k_2 x^2 y + k_3 b x. \end{aligned}$$

2. Poniendo $x' = y' = 0$ en las ecuaciones anteriores obtenemos

$$\begin{aligned} 0 &= k_1 a + k_2 x^2 y - k_3 b x - k_4 x, \\ 0 &= -k_2 x^2 y + k_3 b x, \end{aligned}$$

lo cual implica $x = k_1 a / k_4$, $y = k_3 b / (k_2 x) = k_3 b / (k_2 k_1 a)$. El equilibrio es único.

17. Algunos modelos de poblaciones

18. Ecuaciones de reacción-difusión

19. Inestabilidad de Turing

Si tenemos unas cuantas sustancias químicas que reaccionan según una lista determinada de reacciones, sabemos cuál es la evolución en tiempo de las concentraciones

gracias a la ley de acción de masas. Normalmente (al menos cuando hay balance detallado) los equilibrios de este sistema son estables. Si consideramos el mismo conjunto de reacciones con difusión en espacio, entonces las mismas concentraciones constantes en espacio son también un equilibrio. En principio es posible que al añadir la difusión haya otros equilibrios nuevos, no constantes en espacio.

Se conoce como *inestabilidad de Turing* al siguiente fenómeno: hay reacciones químicas que tienen un equilibrio estable, pero que se vuelve inestable cuando tenemos en cuenta la difusión. Es un fenómeno que va en contra de la intuición porque uno tiende a pensar que añadir difusión a una ecuación en todo caso podría tener el efecto de *favorecer* la estabilidad, no evitarla. En esta sección vamos a ver alguna evidencia de que esto es así, y a dar condiciones que nos digan cuándo puede pasar.

Consideramos un dominio abierto y acotado $\Omega \subseteq \mathbb{R}^d$ y dos sustancias A, B en Ω con concentraciones $u = u(t, x)$, $v = v(t, x)$, dependientes del tiempo $t \geq 0$ y la posición $x \in \Omega$. Sus concentraciones evolucionan según el siguiente sistema de reacción-difusión:

$$\left. \begin{aligned} \partial_t u &= D_1 \Delta u + f_1(u, v) \\ \partial_t v &= D_2 \Delta v + f_2(u, v) \end{aligned} \right\} \quad (295)$$

con condición de contorno de tipo Neumann,

$$\nabla u \cdot N = \nabla v \cdot N = 0 \quad \text{en } \partial\Omega. \quad (296)$$

Si las concentraciones son homogéneas en espacio (es decir, u, v no dependen de x) entonces obtenemos un sistema de ecuaciones ordinarias,

$$\left. \begin{aligned} \frac{d}{dt} u &= f_1(u, v) \\ \frac{d}{dt} v &= f_2(u, v). \end{aligned} \right\} \quad (297)$$

Supongamos que $u = u_\infty, v = v_\infty$ es un equilibrio de (297) (y por tanto $u(t, x) \equiv u_\infty, v(t, x) \equiv v_\infty$ es un equilibrio de (295), homogéneo en espacio). *Nos preguntamos qué condiciones deberían cumplir f_1, f_2, D_1, D_2 para que este equilibrio sea estable para la ecuación (297) e inestable para la ecuación (295)–(296).*

Estudio de estabilidad de la ecuación sin difusión. La forma estándar de estudiar la estabilidad del sistema (297) es considerar su linealización en torno al equilibrio (u_∞, v_∞) : si definimos las fluctuaciones U, V alrededor del equilibrio por

$$u = u_\infty + U, \quad v = v_\infty + V, \quad (298)$$

entonces U, V satisfacen aproximadamente (a primer orden) la ecuación lineal

$$\left. \begin{aligned} \frac{d}{dt} U &= a_{11}U + a_{12}V \\ \frac{d}{dt} V &= a_{21}U + a_{22}V, \end{aligned} \right\} \quad (299)$$

donde la matriz $A = (a_{ij})_{i,j=1,2}$ es la matriz jacobiana en (u_∞, v_∞) de la función que define la ecuación (297):

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = \begin{pmatrix} \nabla f_1(u_\infty, v_\infty) \\ \nabla f_2(u_\infty, v_\infty) \end{pmatrix}. \quad (300)$$

No es difícil ver lo siguiente:

Lema 19.1. Consideramos una matriz $A = (a_{i,j})_{i,j=1,2}$ de tamaño 2×2 .

1. Todos los valores propios de A tienen parte real negativa si y sólo si

$$\operatorname{tr} A < 0, \quad \det A > 0. \quad (301)$$

2. Existe un valor propio de A con parte real positiva si y sólo si

$$\operatorname{tr} A > 0 \quad \text{o bien} \quad \det A < 0.$$

Demostración. La demostración es directa: los valores propios de A son

$$\frac{1}{2} \left(\operatorname{tr} A \pm \sqrt{(\operatorname{tr} A)^2 - 4 \det A} \right).$$

□

En el caso 1 sabemos que el equilibrio (u_∞, v_∞) es estable para (297); en el caso 2 sabemos que es inestable. En el caso complementario (cuando alguno de los valores propios tiene parte real 0) no podemos decir nada de la ecuación no lineal (297) mirando solamente su linealización (299).

Como estamos buscando condiciones que nos aseguren la existencia de una inestabilidad de Turing, a partir de ahora supondremos que se cumple (301).

Estudio de estabilidad de la ecuación con difusión en dimensión 1. Podemos estudiar ahora la linealización de la EDP (295) en torno al equilibrio homogéneo (u_∞, v_∞) : de nuevo definiendo las perturbaciones U, V como en (298), en la aproximación de primer orden tenemos

$$\left. \begin{aligned} \partial_t U &= D_1 \Delta U + a_{11} U + a_{12} V, \\ \partial_t V &= D_2 \Delta V + a_{21} U + a_{22} V, \end{aligned} \right\} \quad (302)$$

siempre con la condición de contorno del tipo (296). Como primer paso consideremos el caso en dimensión 1, de forma que Ω es un intervalo abierto y acotado de \mathbb{R} .

$$\left. \begin{aligned} \partial_t U &= D_1 \partial_x^2 U + a_{11} U + a_{12} V, \\ \partial_t V &= D_2 \partial_x^2 V + a_{21} U + a_{22} V. \end{aligned} \right\} \quad (303)$$

Podemos buscar soluciones particulares del tipo

$$u(t, x) = \alpha_1 e^{\sigma t} \cos(\omega x), \quad v(t, x) = \alpha_2 e^{\sigma t} \cos(\omega x), \quad (304)$$

para ciertos $\sigma \in \mathbb{C}$, $\omega \geq 0$. Para cumplir las condiciones de contorno es necesario que

$$\omega = \frac{n\pi}{L} \quad \text{para cierto } n \geq 0 \text{ entero.}$$

El caso $n = 0$ corresponde a las soluciones de (303) homogéneas en espacio. Si encontramos alguna de esta forma con $\sigma > 0$ entonces el sistema (303) es inestable, y la ecuación no lineal (295) es linealmente inestable en torno a (u_∞, v_∞) . (La pregunta de si la ecuación no lineal (295) es estable o no es más difícil de justificar rigurosamente,

pero en general esperamos que el equilibrio (u_∞, v_∞) de (295) sea inestable siempre que es linealmente inestable.)

Usando la forma (304) en (303) y cancelando $e^{\sigma t} \cos(\omega x)$ obtenemos

$$\left. \begin{aligned} \sigma \alpha_1 &= -D_1 \omega^2 \alpha_1 + a_{11} \alpha_1 + a_{12} \alpha_2, \\ \sigma \alpha_2 &= -D_2 \omega^2 \alpha_2 + a_{21} \alpha_1 + a_{22} \alpha_2. \end{aligned} \right\}$$

Es decir: σ es un valor propio de la matriz

$$A_\omega := \begin{pmatrix} a_{11} - D_1 \omega^2 & a_{12} \\ a_{21} & a_{22} - D_2 \omega^2 \end{pmatrix}.$$

De la discusión anterior conocemos una condición sencilla para que esta matriz tenga un valor propio con parte real positiva: tiene que ocurrir que su traza sea positiva, o bien que su determinante sea negativo. Su traza no puede ser positiva porque es más pequeña que $\text{tr } A$, que estamos suponiendo que es negativa (por (301)). Así que la única posibilidad es que su determinante sea negativo:

$$(a_{11} - D_1 \omega^2)(a_{22} - D_2 \omega^2) - a_{12} a_{21} < 0,$$

que reorganizando los términos es

$$D_1 D_2 \omega^4 - (D_1 a_{22} + D_2 a_{11}) \omega^2 + \det A < 0. \quad (305)$$

Esto es un polinomio cuadrático en ω^2 , con primer y último coeficiente positivo. Para que pueda ser negativo en algún valor de ω^2 (que es siempre ≥ 0) tiene que ocurrir al menos que

$$D_1 a_{22} + D_2 a_{11} > 0. \quad (306)$$

Aquí vemos que para cumplir esto las difusiones D_1 y D_2 no pueden ser iguales, ya que $a_{11} + a_{22} < 0$. Además tiene que ocurrir que el polinomio (305) efectivamente tenga valores negativos; es decir, que su discriminante sea positivo:

$$\det A < \frac{(D_1 a_{22} + D_2 a_{11})^2}{4D_1 D_2}. \quad (307)$$

Por último, tiene que haber un valor admisible de ω^2 que esté dentro del intervalo donde el polinomio (305) es negativo: tiene que existir un entero $n \geq 1$ tal que

$$\begin{aligned} D_1 a_{22} + D_2 a_{11} - \sqrt{(D_1 a_{22} + D_2 a_{11})^2 - 4D_1 D_2 \det A} \\ < \frac{n^2 \pi^2}{L^2} < D_1 a_{22} + D_2 a_{11} + \sqrt{(D_1 a_{22} + D_2 a_{11})^2 - 4D_1 D_2 \det A}. \end{aligned} \quad (308)$$

Observamos que si se cumple (308) para algún $n \geq 1$, entonces también se cumple (306), así que podemos eliminar (306) de nuestra lista de condiciones. Finalmente hemos encontrado un conjunto de condiciones que nos aseguran que hay una inestabilidad de Turing, al menos en el sentido linealizado:

Teorema 19.2. *Consideramos el sistema de reacción-difusión*

$$\left. \begin{aligned} \partial_t u &= D_1 \Delta u + f_1(u, v) \\ \partial_t v &= D_2 \Delta v + f_2(u, v) \end{aligned} \right\} \quad (309)$$

planteado para funciones $u = u(t, x)$, $v = v(t, x)$ definidas en $[0, +\infty) \times [0, L]$, con $L, D_1, D_2 > 0$ dados, $f_1, f_2 \in \mathcal{C}^1(\mathbb{R}^2)$. Supongamos que u_∞, v_∞ son tales que

$$f_1(u_\infty, v_\infty) = 0 = f_2(u_\infty, v_\infty),$$

y llamemos $(a_{11}, a_{12}) := \nabla f_1(u_\infty, v_\infty)$, $(a_{21}, a_{22}) := \nabla f_2(u_\infty, v_\infty)$. Supongamos que se cumplen las siguientes condiciones:

$$\begin{aligned} a_{11} + a_{22} &< 0, & a_{11}a_{22} - a_{12}a_{21} &> 0, \\ \det A &< \frac{(D_1a_{22} + D_2a_{11})^2}{4D_1D_2}, \\ \text{existe } n \in \mathbb{N} &\text{ tal que se cumple (308).} \end{aligned}$$

Entonces el equilibrio (u_∞, v_∞) es estable para perturbaciones homogéneas del sistema (309), pero linealmente inestable.

El resultado anterior nos dice cuáles son las soluciones del sistema linealizado (303) que divergen: son de la forma (304), con $\omega = n\pi/L$, y σ dado por

$$\sigma = \frac{1}{2} \left(\operatorname{tr} A_\omega + \sqrt{(\operatorname{tr} A_\omega)^2 - 4 \det A_\omega} \right).$$

Esto nos sugiere la forma que toman las soluciones inestables inicialmente: esperamos que sea parecida a la función (304) correspondiente al mayor valor posible de ω (que es la solución que diverge más rápido).

A. Fórmulas trigonométricas útiles

Seno y coseno de la suma de ángulos

$$\begin{aligned} \operatorname{sen}(\alpha + \beta) &= \operatorname{sen} \alpha \cos \beta + \cos \alpha \operatorname{sen} \beta \\ \cos(\alpha + \beta) &= \cos \alpha \cos \beta - \operatorname{sen} \alpha \operatorname{sen} \beta \end{aligned}$$

Productos de senos y cosenos

$$\begin{aligned} \cos(\alpha) \cos(\beta) &= \frac{1}{2} (\cos(\alpha - \beta) + \cos(\alpha + \beta)), \\ \sin(\alpha) \sin(\beta) &= \frac{1}{2} (\sin(\alpha - \beta) - \cos(\alpha + \beta)), \\ \operatorname{sen}(\alpha) \cos(\beta) &= \frac{1}{2} (\sin(\alpha + \beta) + \sin(\alpha - \beta)). \end{aligned}$$

B. Cálculo en varias variables

B.1. Notación para las derivadas

Si $f : D \rightarrow \mathbb{R}$ es una función definida en un abierto no vacío $D \subseteq \mathbb{R}^d$, para denotar sus derivadas parciales normalmente elegimos un nombre “estándar” para las variables de f . Por ejemplo, si llamamos $x = (x_1, \dots, x_n)$ a un punto genérico de D , entonces

escribimos la derivada parcial de f con respecto a la variable n -ésima en el punto x como

$$\partial_{x_n} f(x) := \lim_{h \rightarrow 0} \frac{f(x + h e_n) - f(x)}{h}, \quad (310)$$

cuando dicho límite existe, donde $1 \leq n \leq d$ y donde e_n es el vector n -ésimo de la base usual de \mathbb{R}^d (el vector que tiene 1 en la componente n y 0 en las demás). A veces las variables de la función f tienen un nombre específico y entonces escribimos las derivadas parciales usando ese nombre. Por ejemplo, si escribimos $f = f(t, x)$ (con $t \in \mathbb{R}$, $x \in \mathbb{R}$), entonces sus derivadas parciales son

$$\partial_t f, \partial_x f.$$

Para las derivadas parciales de orden superior usamos normalmente la notación con exponente:

$$\partial_{x_n}^2 f, \partial_{x_n}^3 f, \dots \quad (311)$$

o, para las parciales cruzadas,

$$\partial_{x_n} \partial_{x_m} f \quad (312)$$

con $n, m \in \{1, \dots, d\}$. A veces por brevedad usamos

$$\partial_{x_n x_m} f \quad (313)$$

en lugar de (312).

También usamos a veces la notación con multiíndices para las derivadas parciales. Un *multiíndice* de dimensión d es un vector $\alpha = (\alpha_1, \dots, \alpha_d)$ donde cada una de las componentes α_n es un entero mayor o igual que 0. El *orden* de un multiíndice α de dimensión d es

$$|\alpha| := \alpha_1 + \dots + \alpha_d.$$

Escribimos $\partial^\alpha f$ para denotar

$$\partial^\alpha f := \partial_{x_1}^{\alpha_1} \dots \partial_{x_d}^{\alpha_d} f.$$

B.2. Diferenciabilidad

En conjuntos abiertos la siguiente definición de regularidad \mathcal{C}^k es estándar:

Definición B.1. Sea $D \subseteq \mathbb{R}^d$ un abierto no vacío, $f : D \rightarrow \mathbb{R}$ una función y $k \geq 0$ un entero. Decimos que f es de clase \mathcal{C}^k en D (escrito a veces $f \in \mathcal{C}^k(D)$) cuando f es continua en D y todas las derivadas parciales de f de orden menor o igual que k existen y son continuas en D .

Observa que la definición anterior se hace sólo en conjuntos abiertos D . En un conjunto no abierto en general no sabemos definir las derivadas parciales de f y no hay una definición estándar de “clase \mathcal{C}^k ”. Sin embargo, hay una extensión sencilla de la definición anterior para conjuntos que son un abierto, más algunos de los puntos de su frontera:

Definición B.2. Sea $D \subseteq \mathbb{R}^d$ un abierto no vacío, C un conjunto tal que $D \subseteq C \subseteq \overline{D}$, $f : C \rightarrow \mathbb{R}$ una función y $k \geq 0$ un entero. Decimos que f es de clase \mathcal{C}^k en C (escrito a veces $f \in \mathcal{C}^k(C)$) cuando f es continua en C y todas las derivadas parciales de f de orden menor o igual que k existen en D , son continuas en D , y tienen una extensión continua a C .

Ejercicio B.1. Sean $U, D \subseteq \mathbb{R}^d$ abiertos con $U \subseteq D \subseteq \overline{U}$. Si $f \in \mathcal{C}^k(U)$ y sus derivadas parciales de orden menor o igual que k tienen una extensión continua a D , entonces $f \in \mathcal{C}^k(D)$.

Ejercicio B.2. Demuestra que la definición B.2 depende sólo del conjunto C y no de cuál sea el abierto D . Es decir: si D_1, D_2 son abiertos no vacíos tales que

$$D_1 \subseteq C \subseteq \overline{D_1}, \quad D_2 \subseteq C \subseteq \overline{D_2},$$

entonces las siguientes afirmaciones son equivalentes:

1. f es continua en C y todas las derivadas parciales de f de orden menor o igual que k existen en D_1 , son continuas en D_1 , y tienen una extensión continua a C .
2. f es continua en C y todas las derivadas parciales de f de orden menor o igual que k existen en D_2 , son continuas en D_2 , y tienen una extensión continua a C .

La definición B.2 incluye conjuntos cerrados de la forma $[a, b]$ con $a < b$ o semicerrados de la forma $(a, b]$ o $[a, b)$ (siempre con $a < b$), y en general cualquier conjunto $A \subseteq \mathbb{R}^d$ tal que $A \subseteq \text{int}(\overline{A})$.

Cuando tenemos una función $f \in \mathcal{C}^k(C)$, con C un conjunto como el de la definición B.2, y x es un punto de ∂C , estrictamente no podemos escribir $\partial^\alpha f(x)$ (con α un multiíndice de orden $\leq k$), ya que la derivada parcial no tiene por qué estar definida en un punto de la frontera. En un punto de la frontera escribiremos $\partial^\alpha f(x)$ para referirnos al valor de la extensión continua de $\partial^\alpha f$ en x . Esto es un valor bien definido por lo siguiente:

Ejercicio B.3. Sea $D \subseteq \mathbb{R}^d$ un abierto no vacío, $f : D \rightarrow \mathbb{R}$ es una función continua y $C \subseteq \mathbb{R}^d$ un conjunto tal que $D \subseteq C \subseteq \overline{D}$. Demuestra que si existe una extensión continua de f definida en C , dicha extensión es única.

B.3. Operadores diferenciales comunes

Gradiente Si $f : \mathbb{R}^d \rightarrow \mathbb{R}$ es una función de clase \mathcal{C}^1 entonces

$$\nabla f(x) := (\partial_{x_1} f(x), \dots, \partial_{x_d} f(x)) \quad \text{para } x = (x_1, \dots, x_d) \in \mathbb{R}^d.$$

Divergencia. Si $X : \mathbb{R}^d \rightarrow \mathbb{R}^d$ es un campo de clase \mathcal{C}^1 , cuyas coordenadas denotamos como $X = (X_1, \dots, X_d)$, entonces

$$\text{div } X(x) := \partial_{x_1} X_1(x) + \dots + \partial_{x_d} X_d(x) \quad \text{para } x = (x_1, \dots, x_d) \in \mathbb{R}^d.$$

Laplaciano Si $f : \mathbb{R}^d \rightarrow \mathbb{R}$ es una función de clase \mathcal{C}^2 entonces

$$\Delta f(x) := \partial_{x_1}^2 f(x) + \dots + \partial_{x_d}^2 f(x) \quad \text{para } x = (x_1, \dots, x_d) \in \mathbb{R}^d.$$

B.4. Coordenadas esféricas en \mathbb{R}^3

A veces, sobre todo en situaciones con simetría esférica, es conveniente describir los puntos (o vectores) de \mathbb{R}^3 usando el *radio* $r \in [0, +\infty)$ (la distancia al origen), el *azimut* $\theta \in [0, 2\pi)$ (el ángulo en radianes que forma el eje x con la proyección del vector en el plano xy) y el *ángulo polar* $\varphi \in [0, \pi]$ (el ángulo en radianes que forma el vector con el eje z). Estas coordenadas se conocen como *coordenadas esféricas*, y una notación usual es (r, θ, φ) para (radio, azimut, ángulo polar). En física es común intercambiar el significado de θ, φ ; nosotros usaremos siempre θ para el azimut y φ para el ángulo polar.

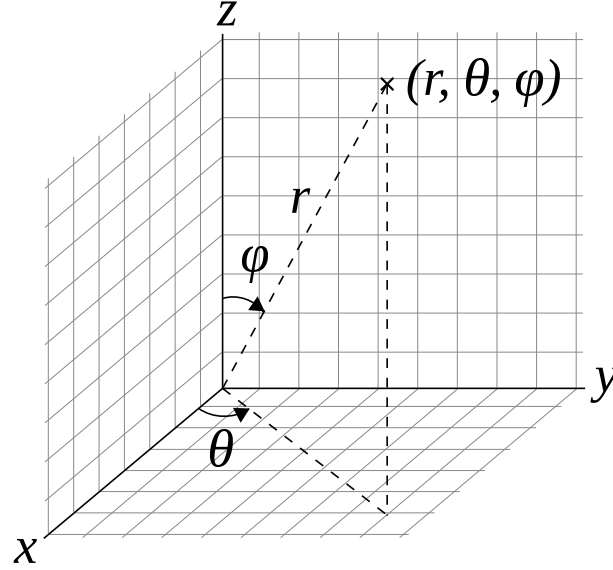


Figura 1: Representación gráfica de las coordenadas esféricas, indicando el radio r , el azimut θ , y el ángulo polar φ . (Tomada de [Wikimedia Commons](#) por Dmcq [CC BY-SA 3.0])

En la descripción de coordenadas terrestres estas coordenadas son esencialmente el radio, la longitud, y la latitud, con la diferencia de que la latitud se mide desde el ecuador, pero el ángulo polar se mide desde el Polo Norte (eje z).

La relación entre las coordenadas cartesianas usuales (x, y, z) y las esféricas es

$$\left. \begin{aligned} x &= r \sin \varphi \cos \theta \\ y &= r \sin \varphi \sin \theta \\ z &= r \cos \varphi \end{aligned} \right\} \text{ con inversa } \left\{ \begin{aligned} r &= \sqrt{x^2 + y^2 + z^2} \\ \theta &= \arctan \frac{y}{x} \\ \varphi &= \arccos \frac{z}{\sqrt{x^2 + y^2 + z^2}} \end{aligned} \right.$$

Siempre usaremos el rango

$$0 \leq r < +\infty, \quad 0 \leq \theta < 2\pi, \quad 0 \leq \varphi \leq \pi.$$

Con más precisión, el cambio entre las coordenadas esféricas y las usuales se puede describir como la aplicación

$$\begin{aligned} X: [0, +\infty) \times [0, 2\pi) \times [0, \pi] &\rightarrow \mathbb{R}^3 \\ X(r, \theta, \varphi) &= (r \sin \varphi \cos \theta, r \sin \varphi \sin \theta, r \cos \varphi). \end{aligned}$$

Para los puntos del eje z el azimut θ no tiene sentido, de forma que el punto de coordenadas $(0, 0, z)$ no tiene una única descripción en coordenadas esféricas. Si eliminamos el eje z , entonces la aplicación

$$X: [0, +\infty) \times [0, 2\pi) \times (0, \pi) \rightarrow \mathbb{R}^3 \setminus \{(0, 0, z) \mid z \in \mathbb{R}\}$$

$$X(r, \theta, \varphi) = (r \sen \varphi \cos \theta, r \sen \varphi \sen \theta, r \cos \varphi)$$

es biyectiva, con inversa

$$\Psi: \mathbb{R}^3 \setminus \{(0, 0, z) \mid z \in \mathbb{R}\} \rightarrow [0, +\infty) \times [0, 2\pi) \times (0, \pi)$$

$$\Psi(x, y, z) = \left(\sqrt{x^2 + y^2 + z^2}, \arctan \frac{y}{x}, \arccos \frac{z}{\sqrt{x^2 + y^2 + z^2}} \right),$$

con la convención de que $\arctan \frac{y}{0} = \pi/2$ para $y > 0$, $\arctan \frac{y}{0} = 3\pi/2$ para $y < 0$.

Jacobiano Para usar este cambio de variables en integrales necesitamos su jacobiano, que puede calcularse de la siguiente forma:

$$\text{Jac } X(r, \theta, \varphi) = \det(\partial_r X, \partial_\theta X, \partial_\varphi X).$$

Teniendo en cuenta que los vectores

$$\begin{aligned} \partial_r X &= (\sen \varphi \cos \theta, \sen \varphi \sen \theta, \cos \varphi), \\ \partial_\theta X &= (-r \sen \varphi \sen \theta, r \sen \varphi \cos \theta, 0), \\ \partial_\varphi X &= (r \cos \varphi \cos \theta, r \cos \varphi \sen \theta, -r \sen \varphi) \end{aligned}$$

son ortogonales entre sí, tenemos que

$$\text{Jac } X(r, \theta, \varphi) = \det(\partial_r X, \partial_\theta X, \partial_\varphi X) = |\partial_r X| |\partial_\theta X| |\partial_\varphi X| = r^2 \sen \varphi.$$

B.4.1. El laplaciano en coordenadas esféricas

Dada una función $f = f(r, \theta, \varphi)$ en coordenadas esféricas, nos gustaría expresar su laplaciano en función de sus derivadas con respecto a (r, θ, φ) . El resultado final al que vamos a llegar es que

$$\Delta f = \frac{1}{r^2} \partial_r (r^2 \partial_r f) + \frac{1}{r^2 (\sen \varphi)^2} \partial_\theta^2 f + \frac{1}{r^2 \sen \varphi} \partial_\varphi (\sen \varphi \partial_\varphi f). \quad (314)$$

Esto se puede probar mediante una aplicación directa de la regla de la cadena, calculando $\partial_x, \partial_y, \partial_z$ en función de $\partial_r, \partial_\theta, \partial_\varphi$, y a partir de esto calculando $\partial_x^2, \partial_y^2, \partial_z^2$. Es un cálculo horrible que al final da una fórmula bastante sencilla, lo que sugiere que hay una manera más fácil de hacerlo. Una idea que simplifica mucho las cuentas es expresar el gradiente y la divergencia en una base distinta, mejor adaptada a las coordenadas esféricas. Si escribimos las coordenadas como

$$X(r, \theta, \varphi) := (r \sen \varphi \cos \theta, r \sen \varphi \sen \theta, r \cos \varphi)$$

y llamamos

$$\begin{aligned} v_r &:= \frac{1}{|\partial_r X|} \partial_r X = (\sin \varphi \cos \theta, \sin \varphi \sin \theta, \cos \varphi), \\ v_\theta &:= \frac{1}{|\partial_\theta X|} \partial_\theta X = (-\sin \theta, \cos \theta, 0), \\ v_\varphi &:= \frac{1}{|\partial_\varphi X|} \partial_\varphi X = (\cos \varphi \cos \theta, \cos \varphi \sin \theta, -\sin \varphi), \end{aligned}$$

vemos que $(v_r, v_\theta, v_\varphi)$ es una base ortonormal de \mathbb{R}^3 . Tenemos por tanto

$$\nabla = v_r(v_r \cdot \nabla) + v_\theta(v_\theta \cdot \nabla) + v_\varphi(v_\varphi \cdot \nabla).$$

Por la regla de la cadena, tenemos también que

$$\begin{aligned} \partial_r f &= \nabla f \cdot \partial_r X = v_r \cdot \nabla f, \\ \partial_\theta f &= \nabla f \cdot \partial_\theta X = r \sin \varphi v_\theta \cdot \nabla f, \\ \partial_\varphi f &= \nabla f \cdot \partial_\varphi X = r v_\varphi \cdot \nabla f, \end{aligned}$$

luego

$$\nabla = v_r \partial_r + \frac{1}{r \sin \varphi} v_\theta \partial_\theta + \frac{1}{r} v_\varphi \partial_\varphi.$$

Para calcular $\Delta = \nabla \cdot \nabla$ podemos usar que los vectores $(v_r, v_\theta, v_\varphi)$ son ortogonales y que

$$\begin{aligned} v_r \partial_r v_r &= 0, & v_\theta \partial_\theta v_r &= \sin \varphi & v_\varphi \partial_\varphi v_r &= 1, \\ v_r \partial_r v_\theta &= 0, & v_\theta \partial_\theta v_\theta &= 0 & v_\varphi \partial_\varphi v_\theta &= 0, \\ v_r \partial_r v_\varphi &= 0, & v_\theta \partial_\theta v_\varphi &= \cos \varphi, & v_\varphi \partial_\varphi v_\varphi &= 0. \end{aligned}$$

Usando esto,

$$\begin{aligned} \Delta &= \nabla \cdot \nabla = (v_r \partial_r + \frac{1}{r \sin \varphi} v_\theta \partial_\theta + \frac{1}{r} v_\varphi \partial_\varphi) \cdot (v_r \partial_r + \frac{1}{r \sin \varphi} v_\theta \partial_\theta + \frac{1}{r} v_\varphi \partial_\varphi) \\ &= \partial_r^2 + \frac{1}{r} \partial_r + \frac{1}{r^2 (\sin \varphi)^2} \partial_\theta^2 + \frac{1}{r} \partial_r + \frac{1}{r^2} \partial_\varphi^2 + \frac{\cos \varphi}{r^2 \sin \varphi} \partial_\varphi, \end{aligned}$$

que es lo mismo que (314).

B.5. Coordenadas esféricas en \mathbb{R}^d

Hay una extensión de las coordenadas esféricas a \mathbb{R}^d , que se puede obtener teniendo que cuenta que si fijamos la última coordenada x_d , el resto de coordenadas están en una esfera de dimensión $d-2$, de forma que podemos definir las coordenadas recursivamente. Por ejemplo en \mathbb{R}^4 , dado un vector $x = (x_1, \dots, x_4) \in \mathbb{R}^4$, si φ_1 es el ángulo que forma x con el eje x_1 , esto resulta en

$$x_1 = r \cos \varphi_1,$$

donde r sigue denotando el radio (el módulo del vector x). La longitud de la proyección de x en el plano $\{x_1 = 0\}$ es

$$r' = r \sin \varphi_1.$$

Usando las coordenadas en dimensión 3 con r' en lugar de r (e identificando por convenio $(x_2, x_3, x_4) \equiv (z, x, y)$, $(\varphi, \theta) \equiv (\varphi_2, \varphi_3)$) tenemos

$$\begin{aligned}x_1 &= r \cos \varphi_1, \\x_2 &= r \sen \varphi_1 \cos \varphi_2, \\x_3 &= r \sen \varphi_1 \sen \varphi_2 \cos \varphi_3, \\x_4 &= r \sen \varphi_1 \sen \varphi_2 \sen \varphi_3,\end{aligned}$$

donde $\varphi_1, \varphi_2 \in [0, \pi]$ y $\varphi_3 \in [0, 2\pi)$. La generalización a \mathbb{R}^d que se obtiene recursivamente es

$$\begin{aligned}x_1 &= r \cos \varphi_1, \\x_i &= r \sen \varphi_1 \cdots \sen \varphi_{i-1} \cos \varphi_i, \quad i = 2, \dots, d-1, \\x_d &= r \sen \varphi_1 \cdots \sen \varphi_{d-1},\end{aligned}$$

donde $\varphi_1, \varphi_2, \dots, \varphi_{d-2} \in [0, \pi]$ y $\varphi_{d-1} \in [0, 2\pi)$. El jacobiano de esta transformación puede calcularse de forma parecida al caso de dimensión 3, ya que los vectores

$$\partial_r x, \partial_{\varphi_i} x, \quad i = 1, \dots, d-1$$

son ortogonales, con norma

$$|\partial_r x| = 1, \quad |\partial_{\varphi_i} x| = r \sen \varphi_1 \cdots \sen \varphi_{i-1}, \quad i = 1, \dots, d-1,$$

cuyo producto es igual al jacobiano:

$$\text{Jac } x(r, \varphi_1, \dots, \varphi_{d-1}) = r^{d-1} \prod_{i=1}^{d-2} \sen(\varphi_i)^{d-1-i}.$$

B.6. Derivación bajo la integral

Teorema B.3 (Derivación bajo la integral). *Sea $\Omega \subseteq \mathbb{R}^d$ un abierto no vacío, $I \subseteq \mathbb{R}^d$ un intervalo abierto no vacío, $f: I \times \Omega \rightarrow \mathbb{R}$ una función (que denotamos $f = f(t, x)$) tal que*

1. *Para cada $t \in I$, la función $x \mapsto f(t, x)$ es integrable en Ω .*
2. *Para cada $x \in \Omega$ la función $t \mapsto f(t, x)$ es una función \mathcal{C}^1 en I .*
3. *Existe una función $M: \Omega \rightarrow [0, +\infty)$ tal que*

$$|\partial_t f(t, x)| \leq M(x), \quad x \in \Omega, \quad t \in I, \quad (315)$$

y

$$\int_{\Omega} M(x) \, dx < +\infty.$$

Entonces existe la derivada parcial $\partial_t \int_{\Omega} f(t, x) \, dx$ para todo $t \in I$, la función $\partial_t f(t, x)$ es integrable en Ω para todo $t \in I$ fijo, y se cumple que

$$\partial_t \int_{\Omega} f(t, x) \, dx = \int_{\Omega} \partial_t f(t, x) \, dx \quad \text{para todo } t \in I. \quad (316)$$

B.7. Teorema de la divergencia y fórmulas de Green

Teorema B.4. Sea $d \geq 2$ un entero, $\Omega \subseteq \mathbb{R}^d$, un abierto acotado con frontera de clase \mathcal{C}^2 , $X: \overline{\Omega} \rightarrow \mathbb{R}^d$ un campo vectorial de clase \mathcal{C}^1 . Entonces

$$\int_{\Omega} \operatorname{div}(X(y)) \, dy = \int_{\partial\Omega} X(y) \cdot \eta(y) \, dS(y),$$

donde $\eta(y)$ denota el normal exterior a $\partial\Omega$ en el punto y .

Corolario B.5. Sea $\Omega \subseteq \mathbb{R}^d$ un abierto acotado con frontera de clase \mathcal{C}^2 , $X: \overline{\Omega} \rightarrow \mathbb{R}^d$ un campo vectorial de clase \mathcal{C}^1 , $f: \overline{\Omega} \rightarrow \mathbb{R}$ de clase \mathcal{C}^1 . Se cumple que

$$\int_{\Omega} \operatorname{div}(X(y))f(y) \, dy = - \int_{\Omega} X(y) \cdot \nabla f(y) \, dy + \int_{\partial\Omega} f(y)X(y) \cdot \eta(y) \, dy,$$

donde $\eta(y)$ denota el normal exterior a $\partial\Omega$ en el punto y .

Corolario B.6. Sea $\Omega \subseteq \mathbb{R}^d$ un abierto acotado con frontera de clase \mathcal{C}^2 , $f, g: \overline{\Omega} \rightarrow \mathbb{R}$ de clase \mathcal{C}^1 . Se cumple que

$$\int_{\Omega} \nabla g(y)f(y) \, dy = - \int_{\Omega} g(y)\nabla f(y) \, dy + \int_{\partial\Omega} f(y)g(y)\eta(y) \, dy,$$

donde $\eta(y)$ denota el normal exterior a $\partial\Omega$ en el punto y .

Si $f, g: \overline{\Omega} \rightarrow \mathbb{R}$ son de clase \mathcal{C}^1 y tales que $|\nabla g|f$ es integrable en \mathbb{R}^d y $|x|^{d-1}f(x)g(x) \rightarrow 0$ cuando $x \rightarrow +\infty$, entonces $|\nabla f|g$ es integrable en \mathbb{R}^d y

$$\int_{\mathbb{R}^d} \nabla g(y)f(y) \, dy = - \int_{\mathbb{R}^d} g(y)\nabla f(y) \, dy.$$

C. Ecuaciones diferenciales ordinarias lineales

Una ecuación diferencial ordinaria lineal es una ecuación del tipo

$$y^{(n)} = a_0(t)y + \cdots + a_{n-1}(t)y^{(n-1)} + b(t), \quad (317)$$

donde $a_0, \dots, a_{n-1}, b: I \rightarrow \mathbb{R}$ son funciones definidas en un intervalo $I \subseteq \mathbb{R}$. Normalmente consideramos I un intervalo abierto y suponemos que a_0, \dots, a_{n-1}, b son continuas en I . En el caso de una ecuación de primer orden la expresión general es

$$y' = a(t)y + b(t). \quad (318)$$

Por otra parte, un *sistema* lineal de ecuaciones diferenciales ordinarias de orden 1 es una ecuación del tipo

$$x' = A(t)x + b(t), \quad (319)$$

donde ahora la incógnita x es un vector dependiente del tiempo $x: I \rightarrow \mathbb{R}^d$, A es una matriz $A: I \rightarrow \mathcal{M}_{d \times d}$ y b es un vector $b: I \rightarrow \mathbb{R}^d$. Por supuesto, en el caso de dimensión $d = 1$ la ecuación (319) se reduce a (318).

La ecuación de orden n (317) puede reescribirse como un sistema de ecuaciones de primer orden: es equivalente a las n ecuaciones

$$Y'_1 = Y_2, \quad (320)$$

$$\vdots \quad (321)$$

$$Y'_{n-1} = Y_n, \quad (322)$$

$$Y'_n = a_0(t)Y_1 + \cdots + a_{n-1}(t)Y_n + b(t), \quad (323)$$

donde $Y_i \equiv y^{(i-1)}$. Esto puede escribirse en forma matricial como

$$Y' = A(t)Y + B(t), \quad (324)$$

donde

$$Y \equiv (y, y', \dots, y^{(n-1)}), \quad (325)$$

la matriz A es

$$A(t) := \begin{pmatrix} 0 & 1 & 0 & \cdots & \cdots & 0 \\ 0 & 0 & 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & 1 & 0 \\ 0 & 0 & \cdots & \cdots & 0 & 1 \\ a_0(t) & a_1(t) & \cdots & \cdots & \cdots & a_{n-1}(t) \end{pmatrix} \quad (326)$$

y el vector B es

$$B(t) := \begin{pmatrix} 0 \\ \vdots \\ 0 \\ b(t) \end{pmatrix}. \quad (327)$$

La ecuación de orden n (317) es completamente equivalente al sistema de n ecuaciones de orden 1 (324). Esta última forma resulta útil porque la teoría para sistemas de ecuaciones de primer orden suele ser más cómoda de escribir.

Solución de la ecuación lineal homogénea Un sistema lineal homogéneo de ecuaciones lineales puede resolverse explícitamente en algunos casos:

Supongamos que $A(t)$, $A(s)$ conmutan para todo $t, s \in I$. Entonces el sistema de ecuaciones de primer orden

$$x' = A(t)x, \quad (328)$$

(con la misma notación que antes: $A(t)$ es una matriz) tiene como solución general

$$x(t) = e^{\int_{t_0}^t A(s) ds} C, \quad (329)$$

donde $t_0 \in I$ es un punto cualquiera, y $C \in \mathbb{R}^d$ es un vector (escrito como columna).

La exponencial que aparece en esta expresión es la exponencial de una matriz, que puede definirse de varias formas análogas a la exponencial de números reales. Una de

las posibles definiciones es la siguiente: si (ϕ_1, \dots, ϕ_n) es un sistema fundamental de soluciones de (328) tal que

$$\phi_i(t_0) = e_i, \quad i = 1, \dots, d,$$

con e_i el vector i de la base usual (que tiene un 1 en la posición i , y 0 en las demás), entonces

$$e^{\int_{t_0}^t A(s) ds} = \begin{pmatrix} | & \cdots & | \\ \phi_1(t) & \cdots & \phi_n(t) \\ | & \cdots & | \end{pmatrix},$$

la matriz que resulta de poner los vectores ϕ_1, \dots, ϕ_n como columnas. Hay métodos para calcular la exponencial de una matriz usando su forma de Jordan, así que la fórmula (333) da de hecho un método para el cálculo de las soluciones de la ecuación (328). Observa que en el caso de una ecuación con coeficientes constantes tenemos la matriz A no depende de t y la condición de conmutatividad se cumple trivialmente (así que la fórmula (333) es siempre válida para una ecuación con coeficientes constantes).

En particular esto nos permite resolver la ecuación diferencial de orden n

$$y^{(n)} = a_0(t)y + \cdots + a_{n-1}(t)y^{(n-1)} + b(t), \quad (330)$$

usando el sistema equivalente (324). La solución general es

$$Y(t) = e^{\int_{t_0}^t A(s) ds} C, \quad (331)$$

donde $Y(t) = (y(t), y'(t), \dots, y^{(n-1)}(t))$ y $A(t)$ es la matriz (326).

Solución de la ecuación lineal completa En los casos en los que podemos escribir explícitamente la solución del sistema lineal homogéneo (328) podemos obtener la del sistema completo:

Consideramos sistema completo de ecuaciones de primer orden

$$x' = A(t)x + b(t), \quad (332)$$

(con la misma notación que antes: $A(t)$ es una matriz $d \times d$, $b(t) \in \mathbb{R}^d$ un vector). Supongamos que $\{\phi_1, \dots, \phi_d\}$ es un sistema fundamental de soluciones de la ecuación homogénea (328) que cumple

$$\phi_i(t_0) = e_i, \quad i = 1, \dots, d,$$

con e_i el vector i de la base usual (que tiene un 1 en la posición i , y 0 en las demás). Definimos

$$M(t) := \begin{pmatrix} | & \cdots & | \\ \phi_1(t) & \cdots & \phi_d(t) \\ | & \cdots & | \end{pmatrix},$$

de forma que $M(t_0) = I_d$, la matriz identidad $d \times d$. Entonces la solución general de la ecuación (332) es

$$x(t) = M(t)C + M(t) \int_{t_0}^t M(s)^{-1} b(s) ds, \quad (333)$$

donde $t_0 \in I$ es un punto cualquiera, y $C \in \mathbb{R}^d$ es un vector (escrito como columna).

Cuando $A(t)$, $A(s)$ conmutan para todo $t, s \in I$ (en particular, cuando A es constante en tiempo) la solución (333) puede escribirse como

$$x(t) = e^{\int_{t_0}^t A(\tau) d\tau} C + \int_{t_0}^t e^{\int_s^t A(\tau) d\tau} b(s) ds. \quad (334)$$

De nuevo, esta fórmula da también la solución de la ecuación lineal completa de orden n usando el cambio (325)–(327).

D. Ecuaciones diferenciales ordinarias especiales

D.1. La ecuación de Legendre

La ecuación diferencial de Legendre es la siguiente ecuación diferencial ordinaria:

$$((1 - t^2)x')' = -\lambda x, \quad (335)$$

escrita aquí en forma autoadjunta (ver Sección 7.1). Como es corriente para ecuaciones con coeficientes polinómicos, podemos calcular sus soluciones escribiéndolas como series de potencias: suponiendo que

$$x(t) = \sum_{i=0}^{\infty} a_i t^i$$

tenemos

$$\sum_{i=2}^{\infty} i(i-1)a_i t^{i-2} - \sum_{i=1}^{\infty} i(i+1)a_i t^i = -\lambda \sum_{i=0}^{\infty} a_i t^i,$$

es decir,

$$\sum_{i=0}^{\infty} ((i+2)(i+1)a_{i+2} - i(i+1)a_i + \lambda a_i) t^i = 0.$$

Esto implica que cada término debe ser 0, luego

$$(i+2)(i+1)a_{i+2} = (i(i+1) - \lambda)a_i, \quad i \geq 0.$$

Esto nos permite encontrar una solución recursivamente para los coeficientes a_i : dados a_0, a_1 ,

$$a_i = a_0 \prod_{\substack{j=0 \\ j \text{ par}}}^{i/2} \frac{j(j+1) - \lambda}{(j+1)(j+2)}, \quad i \geq 2 \text{ par}, \quad (336)$$

$$a_i = a_1 \prod_{\substack{j=1 \\ j \text{ impar}}}^{(i-1)/2} \frac{j(j+1) - \lambda}{(j+1)(j+2)}, \quad i \geq 3 \text{ impar}. \quad (337)$$

Cuando $\lambda = \ell(\ell+1)$ para algún $\ell \geq 0$ entonces una de las soluciones es un polinomio, ya que uno de los productos anteriores (el de los términos pares o el de los impares, según ℓ sea par o impar) es cero a partir de un cierto índice. En los casos en que todos los coeficientes son no nulos, vemos que

$$\frac{a_i}{a_{i+2}} = \frac{(m+1)(m+2)}{m(m+1) - \lambda} \rightarrow 1 \quad \text{cuando } i \rightarrow +\infty,$$

con

$$m = \begin{cases} i/2 & \text{si } i \text{ par,} \\ (i-1)/2 & \text{si } i \text{ impar.} \end{cases}$$

Esto demuestra que la serie que define x tiene radio de convergencia 1. Es fácil entonces demostrar lo siguiente:

Teorema D.1. *Dado $\lambda \in \mathbb{R}$, la solución de la ecuación de Legendre*

$$((1-t^2)x')' = -\lambda x,$$

con condición inicial $x(0) = a_0$, $x'(0) = a_1$ es

$$x(t) = \sum_{i=0}^{\infty} a_i t^i, \quad t \in (-1, 1),$$

donde los coeficientes a_i , $i \geq 2$ vienen dados por (336)-(337).

Los *polinomios de Legendre* son las soluciones polinómicas de la ecuación anterior.

D.2. La ecuación de Legendre asociada

La ecuación de Legendre asociada es la siguiente ecuación de segundo orden, con parámetros $\lambda, m \in \mathbb{R}$:

$$((1-t^2)x')' = \lambda x + \frac{m^2}{1-t^2}x. \quad (338)$$

Es una ecuación de segundo orden con coeficientes no constantes. Si la escribimos en forma normal (de forma que el coeficiente de x'' sea 1), sus coeficientes tienen singularidades en $t = \pm 1$, de forma que buscamos soluciones en el intervalo $[-1, 1]$.

* Soluciones en series de potencias.

* Otras soluciones?

Si x es una solución de la ecuación de Legendre (D.1), derivando m veces obtenemos

$$((1-t^2)x')^{(m+1)} = -\lambda x^{(m)}$$

esto es,

$$(1-t^2)x^{(m+2)} - 2(m+1)tx^{(m+1)} - 8(m+1)mx^{(m)} = \lambda x^{(m)}.$$

Llamando $y = x^{(m)}$,

$$(1-t^2)y'' - 2(m+1)ty' - 8(m+1)my = \lambda y,$$

E. Integral de Lebesgue

Esencialmente, la integral de Lebesgue es una extensión del concepto de integral que tiene mejores propiedades de paso al límite. Suponiendo que conocemos la integral de Riemann en \mathbb{R}^d , la forma más corta de definirla es tal vez como completación de la norma L^1 sobre funciones continuas de soporte compacto. Para una función $f \in \mathcal{C}_c(\mathbb{R}^d)$ podemos definir la integral de Riemann en el sentido usual, y definimos su norma L^1 como

$$\|f\|_1 := \int_{\mathbb{R}^d} |f(x)| dx, \quad (f \in \mathcal{C}_c(\mathbb{R}^d)).$$

Definición E.1 (Conjuntos de medida nula). Un conjunto $A \subseteq \mathbb{R}^d$ es *de medida nula* cuando para todo $\epsilon > 0$ existen sucesiones $x_n \in \mathbb{R}^d$, $r_n > 0$ tales que

$$A \subseteq \bigcup_{i=1}^{\infty} B_{r_n}(x_n),$$

$$\sum_{i=1}^{\infty} |B_{r_n}(x_n)| \leq \epsilon.$$

Si $\Omega \subseteq \mathbb{R}^d$ es un conjunto cualquier y una propiedad referente a puntos de Ω es cierta para todo $x \in \Omega$ salvo en un conjunto de medida nula, decimos que la propiedad es cierta *para casi todo punto de Ω* .

Definición E.2. (Integral de Lebesgue) Sea $f: \mathbb{R}^d \rightarrow \mathbb{R}$ una función. Decimos que f es *integrable* (en el sentido de Lebesgue) cuando existe una sucesión $(f_n)_{n \geq 1}$ de funciones en $\mathcal{C}_c(\mathbb{R}^d)$ tales que

$$f_n(x) \rightarrow f(x) \quad \text{p.c.t. } x \in \mathbb{R}^d$$

y (f_n) es una sucesión de Cauchy en la norma L^1 . En este caso, definimos la integral de f como

$$\int_{\mathbb{R}^d} f(x) \, dx := \lim_{n \rightarrow +\infty} \int_{\mathbb{R}^d} f_n(x) \, dx$$

para cualquier sucesión f_n que cumpla lo anterior.

La definición anterior no está completa: requiere comprobar que el límite que define la integral de f no depende de la sucesión que elijamos para aproximar f . Escribimos esto como un lema que no vamos a demostrar:

Lema E.3. Si f es una función integrable en el sentido de la definición E.2, entonces el límite de $\int_{\mathbb{R}^d} f_n$ cuando $n \rightarrow +\infty$ es el mismo para cualquier sucesión f_n en $\mathcal{C}_c(\mathbb{R}^d)$ que sea una sucesión de Cauchy en la norma L^1 , y tal que $f_n(x) \rightarrow f(x)$ para casi todo $x \in \mathbb{R}^d$.

La definición anterior no es la usual en los libros que explican la integral de Lebesgue. En su lugar, se suele desarrollar primero el concepto de la medida de Lebesgue en \mathbb{R}^d , y después se define la integral por aproximación con funciones simples. La definición usual es mucho más larga que la Definición E.2, pero por supuesto hay una pega: no conozco ninguna demostración del Lema E.3 que no requiera esencialmente los mismos argumentos que se usan para dar la definición usual. Demostrar el Lema E.3 requiere “cuantificar” de alguna forma el conjunto de puntos donde f_n está cerca de f , y la idea más natural para eso es seguramente usar el concepto de medida de Lebesgue. Dar una definición completa de la integral de Lebesgue no es algo obvio, pero la definición anterior es correcta y sirve perfectamente para trabajar con ella.

F. Convolución y aproximación

Definición F.1 (Convolución). Dadas dos funciones $f, g \in L^1(\mathbb{R}^d)$, definimos su *convolución* $f * g$ como una función que cumpla

$$f * g(x) = \int_{\mathbb{R}^d} f(y)g(x - y) \, dy, \quad \text{p.c.t. } x \in \mathbb{R}^d. \quad (339)$$

Lo anterior define $f * g$ en casi todo punto de \mathbb{R}^d . Podemos comprobar que la definición es correcta porque la función

$$(x, y) \mapsto f(y)g(x - y), \quad x, y \in \mathbb{R}^d$$

es integrable en \mathbb{R}^d , y el teorema de Fubini asegura entonces que la integral (339) está definida para casi todo $x \in \mathbb{R}^d$.

Teorema F.2 (Aproximación por convolución en L^p). *Sea $\varphi: \mathbb{R}^d \rightarrow \mathbb{R}$ una función integrable con $\int_{\mathbb{R}^d} \varphi = 1$, y definimos φ_ϵ como*

$$\varphi_\epsilon(x) = \frac{1}{\epsilon^d} \varphi\left(\frac{x}{\epsilon}\right), \quad x \in \mathbb{R}^d.$$

*Entonces $f * \varphi_\epsilon \rightarrow f$ en $L^p(\mathbb{R}^d)$ cuando $\epsilon \rightarrow 0$.*

G. Series de funciones

G.1. Modos de convergencia

Para una serie numérica tenemos los siguientes conceptos de convergencia:

Definición G.1. Sea $(a_n)_{n \geq 1}$ una sucesión de números complejos.

Convergencia usual. Decimos que la serie $\sum_{n \geq 1} a_n$ converge a $L \in \mathbb{C}$ cuando

$$\left| \sum_{n=1}^m a_n - L \right| \xrightarrow{m \rightarrow +\infty} 0.$$

Convergencia absoluta. Decimos que la serie $\sum_{n \geq 1} a_n$ converge absolutamente cuando

$$\sum_{n=m}^{\infty} |a_n| \xrightarrow{m \rightarrow +\infty} 0.$$

Equivalentemente, la serie converge absolutamente cuando

$$\sum_{n=1}^{\infty} |a_n| < +\infty,$$

pero la primera definición se presta a una analogía más directa con el caso de series de funciones. Para una serie de funciones se aplican los mismos conceptos, y sus versiones uniformes:

Definición G.2. Sea $I \subseteq \mathbb{R}$ un subconjunto no vacío cualquiera, $(f_n)_{n \geq 1}$ una sucesión de funciones con $f_n: I \rightarrow \mathbb{C}$ para cada $n \geq 1$.

Convergencia puntual. Decimos que la serie $\sum_{n \geq 1} f_n$ converge *puntualmente* a una función $f: I \rightarrow \mathbb{C}$ cuando

$$\left| \sum_{n=1}^m f_n(x) - f(x) \right| \xrightarrow{m \rightarrow +\infty} 0 \quad \text{para todo } x \in I.$$

Convergencia absoluta. Decimos que la serie $\sum_{n \geq 1} f_n$ converge absolutamente en I cuando

$$\sum_{n=m}^{\infty} |f_n(x)| \xrightarrow{m \rightarrow +\infty} 0 \quad \text{para todo } x \in I.$$

Convergencia uniforme. Decimos que la serie $\sum_{n \geq 1} f_n$ converge uniformemente en I a una función $f: I \rightarrow \mathbb{C}$ cuando

$$\sup_{x \in I} \left| \sum_{n=1}^m f_n(x) - f(x) \right| \xrightarrow{m \rightarrow +\infty} 0.$$

Convergencia uniforme-absoluta. Decimos que la serie $\sum_{n \geq 1} f_n$ converge uniforme-absolutamente en I cuando

$$\sup_{x \in I} \sum_{n=m}^{\infty} |f_n(x)| \xrightarrow{m \rightarrow +\infty} 0. \quad (340)$$

Lema G.3. Existe la siguiente relación entre los conceptos anteriores:

$$c. \text{ uniforme-absoluta} \Rightarrow \begin{cases} c. \text{ uniforme} \Rightarrow c. \text{ puntual} \\ c. \text{ absoluta} \Rightarrow c. \text{ puntual} \end{cases}$$

Demostración. Vamos a demostrar sólo que la convergencia uniforme-absoluta implica tanto la uniforme como la absoluta, ya que las otras implicaciones son más conocidas. Si una serie $\sum_{n \geq 1} f_n$ converge uniforme-absolutamente es obvio que converge absolutamente, y podemos definir

$$f(x) := \sum_{n=1}^{\infty} f_n(x), \quad x \in I.$$

Entonces

$$\sum_{n=m}^{\infty} |f_n(x)| \geq \left| \sum_{n=m}^{\infty} f_n(x) \right| = \left| f(x) - \sum_{n=1}^m f_n(x) \right|,$$

y tomando el supremo en x en ambos lados tenemos el resultado. \square

Observación G.4. Hay series de funciones que convergen uniformemente y absolutamente, pero no uniforme-absolutamente. Un ejemplo es el siguiente: si $a_n := (-1)^n n^{-1}$ es la serie armónica, definimos $f_n: (0, 1) \rightarrow \mathbb{R}$ como

$$f_n(x) = \begin{cases} a_n & \text{si } x \leq \frac{1}{n}, \\ 0 & \text{si } x > \frac{1}{n}. \end{cases}$$

Dicho de otra forma: en el intervalo $[\frac{1}{k+1}, \frac{1}{k})$ la serie f_n es una serie de funciones constantes, cuyos k primeros términos los de la serie armónica, y que son 0 a partir del término $k+1$. Según nos acercamos a $x = 0$, la serie contiene cada vez más términos de

la serie armónica. La serie $\sum_{n \geq 1} f_n$ converge absolutamente porque en cada x es una serie con un número finito de términos. Converge también uniformemente a la función

$$f(x) = \sum_{n=1}^{x^{-1}} a_n,$$

porque

$$\left| \sum_{n=1}^m f_n(x) - f(x) \right| = \left| \sum_{n=m}^{x^{-1}} a_n \right| \leq \frac{1}{m} \quad \text{para todo } x \in (0, 1),$$

donde se sobreentiende que la última suma es 0 cuando $m < 1/x$. Sin embargo, para $m \geq 1$ tenemos

$$\sum_{n=m}^{\infty} |f_n(x)| = \sum_{n=m}^{x^{-1}} |a_n| = \sum_{n=m}^{x^{-1}} \frac{1}{n},$$

Esto no es una función acotada para $x \in (0, 1)$.

En el caso de funciones, tenemos también el concepto de convergencia en un espacio funcional. En particular, podemos considerar la convergencia en $L^p(I)$:

Definición G.5. Sea $I \subseteq \mathbb{R}$ un intervalo, $p \in [1, +\infty]$, $(f_n)_{n \geq 1}$ una sucesión de funciones con $f_n: I \rightarrow \mathbb{C}$ para cada $n \geq 1$. Decimos que la serie $\sum_{n \geq 1} f_n$ converge en $L^p(I)$ a una función $f \in L^p(I)$ cuando

$$\int_I \left| \sum_{n=1}^m f_n(x) - f(x) \right|^p dx \xrightarrow{m \rightarrow +\infty} 0.$$

Teorema G.6 (Criterio de Weierstrass). Sea $I \subseteq \mathbb{R}$ un subconjunto no vacío cualquiera, $(f_n)_{n \geq 1}$ una sucesión de funciones con $f_n: I \rightarrow \mathbb{C}$ para cada $n \geq 1$. Supongamos que

$$|f_n(x)| \leq a_n \quad \text{para todo } x \in I,$$

con $(a_n)_{n \geq 1}$ una sucesión de números no negativos. Si

$$\sum_{n=1}^{\infty} a_n < +\infty,$$

entonces la serie $\sum_{n \geq 1} f_n$ converge uniforme-absolutamente en I .

Demostración. La demostración es directa a partir de la definición de convergencia en (340). \square

G.2. Derivación de series de funciones término a término

Teorema G.7 (Derivación de una serie de funciones). Sea $I \subseteq \mathbb{R}$ un intervalo acotado, $f_n: I \rightarrow \mathbb{R}$ funciones para $n \in \mathbb{N}$. Supongamos que

1. $f_n \in \mathcal{C}^1(I)$ para todo $n \in \mathbb{N}$.
2. $\sum_{n \in \mathbb{N}} f'_n$ converge uniformemente en I .

3. $\sum_{n \in \mathbb{N}} f_n$ converge al menos en un punto de I .

Entonces la serie $\sum_{n \in \mathbb{N}} f_n$ converge uniformemente en I a una función $\mathcal{C}^1(I)$ y

$$\left(\sum_{n \in \mathbb{N}} f_n \right)'(x) = \sum_{n \in \mathbb{N}} f'_n(x) \quad \text{para todo } x \in I. \quad (341)$$

Demostración. Sea g la función a la que converge la serie $\sum_{n \geq 1} f'_n$. Tenemos

$$\varepsilon_m := \sup_{x \in I} \left| \sum_{n=1}^m f'_n(x) - g(x) \right| \xrightarrow{m \rightarrow +\infty} 0.$$

Sea $x_0 \in I$ un punto tal que converge la serie $\sum_{n \geq 1} f_n(x_0)$, y sea $|I|$ la longitud del intervalo I . Si G es una primitiva cualquiera de g en I , integrando desde x_0 hasta cualquier $x \in I$ obtenemos que

$$\begin{aligned} |I| \varepsilon_m &\geq \int_{x_0}^x \left| \sum_{n=1}^m f'_n(y) - g(y) \right| dy \geq \left| \int_{x_0}^x \left(\sum_{n=1}^m f'_n(y) - g(y) \right) dy \right| \\ &= \left| \sum_{n=1}^m (f_n(x) - f_n(x_0)) - G(x) + G(x_0) \right| \\ &\geq \left| \sum_{n=1}^m f_n(x) - G(x) \right| - \left| \sum_{n=1}^m f_n(x_0) - G(x_0) \right|. \end{aligned}$$

Si elegimos la primitiva G de forma que $G(x_0) = \sum_{n=1}^{\infty} f_n(x_0)$ entonces el último término tiende a 0 y, como $\varepsilon_m \rightarrow 0$, deducimos que

$$\sum_{n \geq 1} f_n(x) \rightarrow G(x) \quad \text{uniformemente en } I$$

Como $G' = \sum_{n=1}^{\infty} f'_n$, obtenemos la conclusión del resultado. \square

Ejercicio G.1. Observa que el resultado anterior requiere que el intervalo I sea *acotado*. Encuentra un ejemplo que demuestre que no podemos quitar esta condición.

Teorema G.8 (Derivación de una serie de funciones, versión p.c.t.). Sea $I \subseteq \mathbb{R}$ un intervalo cualquiera, $f_n : I \rightarrow \mathbb{R}$ funciones para $n \in \mathbb{N}$. Supongamos que

1. f_n es absolutamente continua en I para todo $n \in \mathbb{N}$.
2. $\sum_{n \in \mathbb{N}} f'_n$ converge en $L^1(I)$.
3. $\sum_{n \in \mathbb{N}} f_n$ converge al menos en un punto de I .

Entonces la serie $\sum_{n \in \mathbb{N}} f_n$ converge uniformemente en I a una función absolutamente continua, y

$$\left(\sum_{n \in \mathbb{N}} f_n \right)'(x) = \sum_{n \in \mathbb{N}} f'_n(x) \quad \text{p.c.t. } x \in I. \quad (342)$$

Demostración. Sea g una función a la que converge la serie $\sum_{n \geq 1} f'_n$ (el límite g es único salvo conjuntos de medida nula). Tenemos

$$\varepsilon_m := \int_I \left| \sum_{n=1}^m f'_n(x) - g(x) \right| dx \xrightarrow{m \rightarrow +\infty} 0.$$

Sea $x_0 \in I$ un punto tal que converge la serie $\sum_{n \geq 1} f_n(x_0)$. Si G es una primitiva cualquiera de g en I , integrando desde x_0 hasta cualquier $x \in I$ obtenemos que

$$\begin{aligned} \varepsilon_m &\geq \int_{x_0}^x \left| \sum_{n=1}^m f'_n(y) - g(y) \right| dy \geq \left| \int_{x_0}^x \left(\sum_{n=1}^m f'_n(y) - g(y) \right) dy \right| \\ &= \left| \sum_{n=1}^m (f_n(x) - f_n(x_0)) - G(x) + G(x_0) \right| \\ &\geq \left| \sum_{n=1}^m f_n(x) - G(x) \right| - \left| \sum_{n=1}^m f_n(x_0) - G(x_0) \right|. \end{aligned}$$

Si elegimos la primitiva G de forma que $G(x_0) = \sum_{n=1}^{\infty} f_n(x_0)$ entonces el último término tiende a 0 y, como $\varepsilon_m \rightarrow 0$, deducimos que

$$\sum_{n \geq 1} f_n(x) \rightarrow G(x) \quad \text{uniformemente en } I$$

Como $G' = \sum_{n=1}^{\infty} f'_n$ en casi todo $x \in I$, obtenemos la conclusión del resultado. \square

Teorema G.9 (Derivación de una serie de funciones en varias variables). *Sea $D \subseteq \mathbb{R}^d$ un abierto, $f_n : D \rightarrow \mathbb{R}$ funciones para $n \in \mathbb{N}$. Supongamos que*

1. $\partial_1 f_n$ existe y es continua en D para todo $n \in \mathbb{N}$.
2. $\sum_{n \in \mathbb{N}} \partial_1 f_n$ converge uniformemente en D .
3. $\sum_{n \in \mathbb{N}} f_n$ converge puntualmente a una función f en D .

Entonces existe la derivada parcial $\partial_1 f$, y

$$\partial_1 f(x) = \sum_{n \in \mathbb{N}} \partial_1 f_n(x) \quad \text{para todo } x \in D. \quad (343)$$

Corolario G.10. *Sea k un entero ≥ 1 , $D \subseteq \mathbb{R}^d$ un abierto, $f_n : D \rightarrow \mathbb{R}$ funciones para $n \in \mathbb{N}$. Supongamos que*

1. $f_n \in \mathcal{C}^k(\mathbb{R}^d)$ para todo $n \in \mathbb{N}$.
2. $\sum_{n \in \mathbb{N}} \partial^\alpha f_n$ converge uniformemente en D , para todo multiíndice α de orden $\leq k$.
3. $\sum_{n \in \mathbb{N}} f_n$ converge puntualmente a una función f en D .

Entonces $f \in \mathcal{C}^k(D)$ y para todo multiíndice α de orden $\leq k$ se cumple que

$$\partial^\alpha f(x) = \sum_{n \in \mathbb{N}} \partial^\alpha f_n(x) \quad \text{para todo } x \in D. \quad (344)$$

H. Espacios de Hilbert

H.1. Aplicaciones bilineales y sesquilineales

Definición H.1. Sea H un espacio de Hilbert (sobre $K = \mathbb{R}$ o \mathbb{C}). Una *aplicación bilineal* o *forma bilineal* es una aplicación $a : H \times H \rightarrow K$ que cumple

$$a(\lambda u + \mu v, w) = \lambda a(u, w) + \mu a(v, w), \quad (345)$$

$$a(u, \lambda v + \mu w) = \lambda a(u, v) + \mu a(u, w), \quad (346)$$

para todo $u, v, w \in H$ y todo $\lambda, \mu \in K$.

Definición H.2. Sea H un espacio de Hilbert complejo. Una *aplicación sesquilineal* o *forma sesquilineal* es una aplicación $a : H \times H \rightarrow \mathbb{C}$ que cumple

$$a(\lambda u + \mu v, w) = \bar{\lambda} a(u, w) + \bar{\mu} a(v, w), \quad (347)$$

$$a(u, \lambda v + \mu w) = \lambda a(u, v) + \mu a(u, w), \quad (348)$$

para todo $u, v, w \in H$ y todo $\lambda, \mu \in \mathbb{C}$.

Lema H.3. Sea H un espacio de Hilbert y $a : H \times H \rightarrow K$ una forma bilineal o sesquilineal. Son equivalentes:

1. La aplicación a es continua (en la norma de H).
2. Existe $C > 0$ tal que

$$|a(u, v)| \leq C \|u\| \|v\| \quad \text{para todo } u, v \in H. \quad (349)$$

H.2. Espacios de Hilbert

Definición H.4 (Producto escalar). Sea V un espacio vectorial sobre $K = \mathbb{R}$ o bien $K = \mathbb{C}$. Un producto escalar en V es una aplicación $\langle \cdot, \cdot \rangle : V \times V \rightarrow K$ que cumple lo siguiente para todo $u, v, w \in V$, $\lambda \in K$:

$$\langle u + v, w \rangle = \langle u, w \rangle + \langle v, w \rangle \quad (350)$$

$$\langle u, v + w \rangle = \langle u, v \rangle + \langle u, w \rangle \quad (351)$$

$$\langle \lambda u, v \rangle = \bar{\lambda} \langle u, v \rangle, \quad \langle u, \lambda v \rangle = \lambda \langle u, v \rangle \quad (352)$$

$$\langle u, v \rangle = \overline{\langle v, u \rangle} \quad (353)$$

$$\langle u, u \rangle \geq 0 \quad (354)$$

$$\langle u, u \rangle = 0 \text{ si y sólo si } u = 0. \quad (355)$$

La tres primeras condiciones dicen que la aplicación $\langle \cdot, \cdot \rangle$ debe ser bilineal (o sesquilineal en el caso complejo). La cuarta condición es que el producto debe ser simétrico (o *hermítico* en el caso complejo). Las dos últimas condiciones dicen que el producto $\langle \cdot, \cdot \rangle$ es estrictamente definido positivo.

Definición H.5 (Espacio de Hilbert). Sea $K = \mathbb{R}$ o bien $K = \mathbb{C}$. Un espacio de Hilbert es un espacio vectorial H sobre el cuerpo K con un producto escalar $\langle \cdot, \cdot \rangle : H \times H \rightarrow K$, tal que H es completo (con la métrica inducida por el producto escalar). Decimos que H es real en el caso $K = \mathbb{R}$, y complejo en el caso $K = \mathbb{C}$.

Definición H.6 (Base ortonormal). Sea H un espacio de Hilbert sobre $K = \mathbb{R}$ o \mathbb{C} . Una *base ortonormal* es una familia de vectores $(v_j)_{j \in \Lambda}$ tales que

$$\begin{aligned}\langle v_i, v_j \rangle &= 0 \quad \text{para todo } i \neq j \in \Lambda \\ \langle v_i, v_i \rangle &= 1 \quad \text{para todo } i \in \Lambda\end{aligned}$$

El subespacio generado por $(v_j)_{j \in \Lambda}$ es denso en H .

Observación H.7. Una base ortonormal puede ser finita (si el espacio H tiene dimensión finita), numerable o no numerable. El conjunto Λ donde tomamos los índices puede tener cualquier cardinalidad.

En un espacio de dimensión infinita, una base ortonormal no es una base en el sentido usual (no todo elemento de H puede escribirse como combinación lineal *finita* de elementos de una base ortonormal). Para distinguirlas, las bases en el sentido usual se llaman *bases de Hamel*. En un espacio de Banach general, una sucesión que permite escribir los vectores del espacio de forma única como combinación lineal *infinita* de sus elementos se llama *base de Schauder*.

Lema H.8. Sea H un espacio de Hilbert real o complejo con una base numerable $(v_n)_{n \geq 0}$. Entonces, para cualquier $v \in H$ tenemos

$$v = \sum_{n=1}^{\infty} \langle v, v_n \rangle v_n,$$

donde la convergencia se entiende en el sentido de la norma de H . Además,

$$\|v\|^2 = \sum_{n=1}^{\infty} |\langle v, v_n \rangle|^2$$

Esta última igualdad se llama identidad de Parseval.

Teorema H.9 (Riesz). Sea H un espacio de Hilbert (sobre $K = \mathbb{R}$ o \mathbb{C}) y $L : H \rightarrow K$ una aplicación lineal continua. Existe un único $u \in H$ tal que

$$Av = \langle u, v \rangle \quad \text{para todo } v \in H. \quad (356)$$

Corolario H.10. Si u, v son elementos de un espacio de Hilbert tales que

$$\langle u, w \rangle = \langle v, w \rangle \quad \text{para todo } w \in H \quad (357)$$

entonces $u = v$.

Teorema H.11 (Punto fijo de Banach). Sea X un espacio de Banach y $\Phi : X \rightarrow X$ una aplicación contractiva, es decir, tal que existe $0 \leq L < 1$ tal que

$$\|\Phi(u) - \Phi(v)\| \leq L\|u - v\| \quad \text{para todo } u, v \in X. \quad (358)$$

Entonces existe un único $x \in X$ tal que $\Phi(x) = x$.

Corolario H.12. Sea X un espacio de Banach y $T : X \rightarrow X$ una aplicación lineal continua tal que $L := \|T - I\| < 1$. Entonces T es biyectiva y su inversa es continua.

Demostración. Para ver que es biyectiva, veamos que dado $y \in X$ cualquiera existe un único $x \in X$ con $Tx = y$. Estamos buscando $x \in X$ tal que

$$-Tx + x + y = x,$$

y es fácil ver que la aplicación $\Phi: X \rightarrow X$ dada por $\Phi(x) := -Tx + x + y$ es contractiva: $\|\Phi(u) - \Phi(v)\| = \|(T - I)(u - v)\| \leq L\|u - v\|$. Por tanto tiene un único punto fijo x , lo cual demuestra que T es biyectiva. La inversa de T es continua porque $\|T - I\| < 1$ implica que para todo $x \in X$

$$L\|x\| \geq \|Tx - x\| \geq \|x\| - \|Tx\|,$$

luego

$$\|Tx\| \geq (1 - L)\|x\| \quad \text{para todo } x \in X,$$

o equivalentemente,

$$\|y\| \geq (1 - L)\|T^{-1}y\| \quad \text{para todo } y \in X.$$

□