

Métodos Matemáticos de la Física

Carlos B. Brizozzo

29 de junio de 2025

Índice general

1. Algebra Lineal	1
1.1. Espacios lineales	1
1.1.1. Propiedades fundamentales	1
1.1.2. Independencia lineal	2
1.1.3. Dimensión de un EV	3
1.1.4. Base de un EV	3
1.1.5. Componentes de un vector	4
1.1.6. Subespacios lineales	5
1.2. Operadores lineales	6
1.2.1. Operadores lineales	7
1.2.2. Componentes de un operador lineal	8
1.2.3. Operaciones elementales	9
1.2.4. Conmutatividad	10
1.2.5. Inversa	10
1.2.6. Funciones de operadores	11
1.3. Matrices	12
1.3.1. Propiedades fundamentales	13
1.3.2. Definiciones	14
1.3.3. Matrices notables	15
1.3.4. Operación por bloques	16
1.3.5. Funciones de matrices	17
1.4. Transformaciones de coordenadas	17
1.4.1. Matriz de transformación	18
1.4.2. Covarianza y contravarianza	18
1.4.3. Componentes de un operador	19
1.4.4. Transformaciones de semejanza	20
1.5. Formas y espacio dual	21
1.5.1. Formas	21
1.5.2. Espacio dual	21
1.5.3. Componentes de una forma	21
1.5.4. Transformaciones de coordenadas en V^*	22
1.6. Producto interno, métrica y norma	23
1.6.1. Producto interno	23
1.6.2. Métrica	24

1.6.3. Norma	24
1.6.4. De V a V^* y vuelta con el producto interno	25
1.6.5. Correspondencia geométrica	26
2. Matrices	29
2.1. Autovalores y autovectores	29
2.1.1. Autovectores a derecha	30
2.1.2. Autovectores a izquierda	33
2.1.3. Diagonalización de un operador	34
2.1.4. Operadores Hermitianos	35
2.1.5. Operadores Hermitianos con autovalores degenerados	37
2.1.6. Diagonalización simultánea de operadores Hermitianos	40
2.1.7. Operadores normales	41
2.2. Forma de Jordan	42
2.2.1. Ejemplos ilustrativos	42
2.2.2. Descomposición primaria	43
2.2.3. Forma normal de Jordan	46
2.2.4. Reducción a la forma normal de Jordan	47
2.2.5. Método abreviado para hallar la forma de Jordan	64
3. Tensores	67
3.1. Tensores	67
3.1.1. Definición de tensor	67
3.1.2. Espacio tensorial	69
3.1.3. Producto tensorial	70
3.1.4. Base y componentes	70
3.1.5. Cambio de base	72
3.1.6. Contracción de índices	72
3.1.7. Simetría	73
3.1.8. Producto exterior	75
3.1.9. Densidades tensoriales	76
3.1.10. Tensor adjunto	77
3.1.11. Ley del cociente	78
3.2. Coordenadas curvilíneas	78
3.2.1. Cambios de coordenadas locales	79
3.2.2. La base tangente	80
3.2.3. Vectores contravariantes	81
3.2.4. Vectores covariantes	82
3.2.5. La base dual	82
3.2.6. Tensores	84
3.2.7. Densidades tensoriales	85
3.2.8. Tensor métrico	85
3.2.9. Ascenso y descenso de índices	88
3.2.10. Producto escalar y norma	89
3.3. Integración en coordenadas curvilíneas	90
3.3.1. La integral de volumen	90

3.3.2. La integral de superficie	93
3.3.3. La integral de línea	98
3.3.4. Integrales de vectores y tensores	101
3.4. Derivación en coordenadas curvilíneas	102
3.4.1. La conexión afín de Levi–Civita	102
3.4.2. Derivación covariante	104
3.4.3. Gradiente	109
3.4.4. Rotor	110
3.4.5. Divergencia	111
3.4.6. Laplaciano	112
3.4.7. Componentes físicas de un vector	112
3.4.8. Cálculo en variedades	115
4. Distribuciones y espacios \mathcal{L}^2	117
4.1. Distribuciones	117
4.1.1. Funciones generalizadas	118
4.1.2. Funcionales lineales	119
4.1.3. Distribuciones	119
4.1.4. Ejemplos de distribuciones	121
4.1.5. Sucesiones de distribuciones	124
4.1.6. Distribuciones regulares y singulares	126
4.1.7. Derivación e integración	126
4.1.8. Cambio de variables	128
4.2. Espacios \mathcal{L}^2	129
4.2.1. Definición y resultados básicos	129
4.2.2. Convergencia	131
4.2.3. Funciones ortogonales	132
5. Problemas de Contorno	135
5.1. EDO lineales de segundo orden	135
5.1.1. Coeficientes analíticos	135
5.1.2. La Ecuación de Legendre	138
5.1.3. Puntos singulares regulares	141
5.1.4. La ecuación de Bessel	148
5.2. Problemas de valores iniciales	151
5.2.1. Ceros aislados	151
5.2.2. Alternancia de ceros	152
5.2.3. Reducción a la forma normal de Liouville	152
5.2.4. Soluciones oscilatorias	154
5.3. Problemas de contorno	155
5.3.1. Operadores diferenciales Hermitianos	155
5.3.2. Reducción a un operador formalmente autoadjunto	157
5.3.3. Problemas de contorno	158
5.3.4. Homogeneización de las CC	160
5.3.5. Identidad de Lagrange y fórmula de Green	160
5.4. Problemas de Sturm–Liouville	161

5.4.1.	Autovalores	163
5.4.2.	Subespacios característicos	164
5.4.3.	Ortogonalidad de las autofunciones	164
5.4.4.	Espectro de autovalores	165
5.4.5.	Compleitud de las autofunciones	166
5.5.	La notación de Dirac	166
5.5.1.	Ortogonalidad y completitud	167
5.5.2.	La delta de Dirac	168
5.5.3.	Problemas inhomogéneos	169
6.	Funciones Especiales	171
6.1.	PSL para la ecuación armónica	171
6.1.1.	Autovalores y autofunciones	171
6.1.2.	Desarrollo en autofunciones	173
6.1.3.	Funciones periódicas	174
6.1.4.	Convergencia puntual y uniforme	175
6.1.5.	Expresiones para la delta de Dirac	176
6.1.6.	Otros PSL para la ecuación armónica	178
6.2.	PSL para la ecuación de Legendre	179
6.2.1.	El PSL en $(-1, 1)$	179
6.2.2.	Fórmula de Rodrigues para $P_n(x)$	181
6.2.3.	Representaciones integrales	183
6.2.4.	Función generatriz	184
6.2.5.	Relaciones de recurrencia	185
6.2.6.	Funciones de Legendre de segundo tipo	186
6.2.7.	PSL sobre otros intervalos	187
6.2.8.	Las funciones asociadas de Legendre	188
6.2.9.	Armónicos esféricos	190
6.3.	PSL para la ecuación de Bessel	193
6.3.1.	El PSL en $(0, b)$	194
6.3.2.	Relaciones de recurrencia	197
6.3.3.	Integrales de funciones de Bessel	198
6.3.4.	Función generatriz	199
6.3.5.	Representaciones integrales	200
6.3.6.	Desarrollos asintóticos	201
6.3.7.	Otros PSL asociados a la ecuación de Bessel	201
6.3.8.	Las funciones de Bessel modificadas	202
6.3.9.	Las funciones de Bessel esféricas	204
7.	Transformadas Integrales	207
7.1.	La transformada de Fourier	208
7.1.1.	Definición de la transformada de Fourier	208
7.1.2.	Propiedades de la transformada de Fourier	211
7.1.3.	Espectro de potencia	213
7.1.4.	Ejemplos de aplicación	215
7.2.	La transformada de Laplace	220

7.2.1.	Definición de la transformada de Laplace	221
7.2.2.	Propiedades de la transformada de Laplace	222
7.2.3.	Ejemplos de aplicación	224
7.3.	Comentarios generales	228
8.	Ecuaciones Diferenciales Parciales	231
8.1.	Clasificación	231
8.1.1.	Curvas características	232
8.1.2.	La EDP hiperbólica	234
8.1.3.	La EDP parabólica	235
8.1.4.	La EDP elíptica	235
8.1.5.	La clasificación en más de dos variables	236
8.2.	Modelos físicos	237
8.2.1.	Condiciones de contorno	238
8.2.2.	La ecuación de difusión del calor	239
8.2.3.	La ecuación de la cuerda vibrante	241
8.2.4.	La cuerda infinita y las curvas características	244
9.	Método de Separación de Variables	247
9.1.	Dimensión 1 + 1	247
9.1.1.	Difusión del calor con CC homogéneas	247
9.1.2.	Difusión del calor con CC inhomogéneas	250
9.1.3.	Difusión del calor con CC periódicas en el tiempo	251
9.1.4.	Oscilaciones de una cuerda	253
9.2.	Dimensiones 2 y 2 + 1	256
9.2.1.	La ecuación de Laplace en coordenadas Cartesianas . . .	258
9.2.2.	La ecuación de Laplace en coordenadas polares	261
9.2.3.	La ecuación de Helmholtz en coordenadas Cartesianas .	263
9.2.4.	La ecuación de Helmholtz en coordenadas polares . . .	265
9.2.5.	La ecuación de Poisson en coordenadas polares	267
9.3.	Dimensiones 3 y 3 + 1	269
9.3.1.	La ecuación de Laplace en coordenadas cilíndricas . . .	269
9.3.2.	La ecuación de Helmholtz en coordenadas cilíndricas . .	271
9.3.3.	La ecuación de Laplace en coordenadas esféricas	273
9.3.4.	La ecuación de Helmholtz en coordenadas esféricas . . .	274
9.4.	Ejemplos de aplicación	276
9.4.1.	Conductor cilíndrico en un campo externo	276
9.4.2.	Radiación de una antena cilíndrica	277
9.4.3.	Difusión del calor en una placa circular	279
9.4.4.	Vibraciones de una membrana circular	283
9.4.5.	Guía de ondas acústica	284
9.4.6.	Cavidad resonante acústica	286

10. Método de Transformadas Integrales	289
10.1. Dominios semiinfinitos	289
10.1.1. Difusión del calor en una barra semiinfinita	290
10.1.2. Ondas en una cuerda semiinfinita	292
10.2. Dominios infinitos	293
10.2.1. Difusión del calor en una barra infinita	294
10.2.2. Ondas en una cuerda infinita	295
10.2.3. Potencial de una cinta plana	296
10.2.4. Condensador de cinta plana	298
11. EDP Inhomogéneas	301
11.1. La función de Green	302
11.2. Problemas en dimensión 1	305
11.2.1. Soluciones fundamentales	305
11.2.2. Funciones de Green	308
11.3. Problemas en dimensión N	311
11.3.1. Solución fundamental para la ecuación de Laplace	311
11.3.2. Solución fundamental para la ecuación de Helmholtz	313
11.3.3. Función de Green para la ecuación de Laplace	313
11.3.4. Función de Green para la ecuación de Helmholtz	316
11.4. Problemas en dimensión $N + 1$	317
11.4.1. Solución fundamental para la ecuación de difusión	318
11.4.2. Solución fundamental para la ecuación de ondas	319
11.4.3. Función de Green para la ecuación de ondas	321
11.5. Función de Green para CC inhomogéneas	323

Índice de figuras

6.1.	Fenómeno de Gibbs	177
6.2.	Polinomios de Legendre	181
6.3.	Representación Integral de P_n	183
6.4.	Funciones de Legendre de segundo tipo	187
6.5.	Funciones asociadas de Legendre P_n^1	190
6.6.	Funciones asociadas de Legendre P_n^2	191
6.7.	Armónicos esféricos	193
6.8.	Funciones de Bessel de primer tipo	195
6.9.	Funciones de Bessel de segundo tipo	196
6.10.	Funciones de Bessel modificadas de primer tipo	202
6.11.	Funciones de Bessel modificadas de segundo tipo	203
6.12.	Funciones de Bessel esféricas de primer tipo	204
6.13.	Funciones de Bessel esféricas de segundo tipo	206
7.1.	Contornos de integración para la inversión de Fourier.	219
7.2.	Camino de integración para la inversión de Laplace.	222
7.3.	Contorno de integración para la inversión del oscilador armónico.	225
7.4.	Circuito RLC serie.	226
7.5.	Contorno de integración para $s^{-1/2}$	227
8.1.	Conducción del calor en una barra.	239
8.2.	Tensiones en una cuerda.	242
8.3.	CC para una cuerda.	243
8.4.	Curvas características de la ecuación de ondas.	245
8.5.	Propagación de un pulso rectangular.	246
9.1.	Barra aislada con extremos a 0°C	248
9.2.	Cuerda pulsada.	253
9.3.	Modos de una membrana.	284

Índice de tablas

8.1. Clasificación de EDP.	232
8.2. Modelos físicos.	237
8.3. Condiciones de contorno.	238

Capítulo 1

Algebra Lineal

1.1. Espacios lineales

En Física usamos habitualmente vectores para representar magnitudes tales como posición, velocidad, fuerza, campos eléctrico y magnético, etc. Por ello es de interés formalizar y organizar algunos conceptos que quizá ya estemos habituados a usar.

Comenzaremos repasando conceptos básicos de espacios vectoriales.

1.1.1. Propiedades fundamentales

Llamaremos *espacio lineal* o *espacio vectorial* (y abreviaremos EV) a un conjunto V de objetos (llamados *vectores*), $V = \{\vec{x}, \vec{y}, \vec{z}, \dots\}$ (más un *cuerpo escalar*,¹ que denotaremos² por \mathbb{C}), que es *cerrado* bajo las siguientes operaciones:³

1. **Suma:** $\vec{a}, \vec{b} \in V \implies \vec{a} + \vec{b} = \vec{c} \in V$, que es:

- a) asociativa: $(\vec{a} + \vec{b}) + \vec{c} \equiv \vec{a} + (\vec{b} + \vec{c})$;
- b) commutativa: $\vec{a} + \vec{b} \equiv \vec{b} + \vec{a}$.

2. **Producto por un escalar:** $\vec{a} \in V, \lambda \in \mathbb{C} \implies \lambda\vec{a} \in V$, que es:

a) distributiva:

- respecto a los vectores: $\lambda(\vec{a} + \vec{b}) \equiv \lambda\vec{a} + \lambda\vec{b}$;
- respecto a los escalares: $(\lambda + \mu)\vec{a} \equiv \lambda\vec{a} + \mu\vec{a}$

¹En algunos casos basta que los escalares tengan estructura de anillo (sin inversa multiplicativa), como ser en el reticulado $V = \mathbb{Z}^n$ con “cuerpo” (anillo) \mathbb{Z} .

²Para simplificar, usaremos \mathbb{C} tanto como notación genérica para un cuerpo escalar, como para denotar el cuerpo complejo, ya que en casos particulares el contexto dejará claro si el cuerpo son los reales, los complejos, etc.

³El símbolo \equiv denotará que ambos lados de la ecuación son iguales para todo valor posible de las variables que aparezcan, es decir $x + y \equiv y + x$ será lo mismo que $x + y = y + x \forall x, y$.

b) asociativa: $\lambda(\mu\vec{a}) \equiv (\lambda\mu)\vec{a}$

Suponemos además que:

- Existe un (único) *elemento nulo* en V : $\exists! \vec{0} \in V$ t.q. $\vec{a} + \vec{0} \equiv \vec{a}$.
- Existe un (único) *elemento unidad* en \mathbb{C} : $\exists! 1 \in \mathbb{C}$ t.q. $1\vec{a} \equiv \vec{a}$.
- Para cada vector, existe un (único) *elemento inverso* (aditivo) en V : $\forall \vec{a} \in V \exists! -\vec{a} \in V$ t.q. $\vec{a} + (-\vec{a}) = \vec{0}$.

Cualquier conjunto V que posea todas estas propiedades, será para nosotros un EV.

Ejemplo 1.1.1. Los espacios tridimensionales de vectores reales, \mathbb{R}^3 , y complejos, \mathbb{C}^3 , y el reticulado Cartesiano tridimensional (red cúbica) \mathbb{Z}^3 , son EV, con cuerpos \mathbb{R} , \mathbb{C} y \mathbb{Z} , respectivamente (este último “cuerpo” es sólo un anillo: carece de inversa multiplicativa).

Ejemplo 1.1.2. Los polinomios en una variable real x de grado a lo sumo N forman un EV, $V = \{p_N(x)\}$.

Ejemplo 1.1.3. Las funciones a valores reales definidas sobre el intervalo real $[a, b]$ forman un EV, $V = \{f : [a, b] \rightarrow \mathbb{R}\}$. Sin embargo, si restringimos a las funciones a valores positivos no obtenemos un EV.

Ejercicio 1.1.1. Muestre que $\{f : [a, b] \rightarrow \mathbb{R}_+\}$ no es un EV.

1.1.2. Independencia lineal

Un conjunto cualquiera de n vectores no nulos, $\{\vec{a}_i\}_{i=1}^n$, $\vec{a}_i \in V$, $\vec{a}_i \neq \vec{0}$, $i = 1, \dots, n$ se llama *linealmente independiente* (abreviamos LI) si la única combinación lineal de ellos que se anula es la trivial, es decir si

$$\lambda_i \vec{a}_i = \vec{0} \implies \lambda_i \equiv 0, \quad \lambda_i \in \mathbb{C}, \quad i = 1, \dots, n. \quad (1.1)$$

Note el lector que de ahora en mas, salvo nota explícita en contrario, asumiremos el *convenio de suma* sobre índices repetidos: $a_i b_i := \sum_i a_i b_i$.

Si un conjunto de vectores no es LI, lo llamaremos *linealmente dependiente* (y abreviaremos LD).

Ejemplo 1.1.4. En \mathbb{R}^3 , si \hat{x} , \hat{y} y \hat{z} son los versores Cartesianos usuales, $\{\hat{x}, \hat{y}\}$, $\{\hat{x}, \hat{x} + \hat{y}\}$ y $\{\hat{x}, \hat{x} + \hat{y}, \hat{x} + \hat{y} + \hat{z}\}$ son LI, pero $\{\hat{x}, \hat{y}, \hat{x} + \hat{y}\}$ es LD.

Ejemplo 1.1.5. En $\{p_5(x)\}$, $\{1, x, x^2\}$ es LI, pero $\{x, x^2, 2x + 3x^2\}$ es LD.

Ejemplo 1.1.6. En $\{f : [0, 2\pi] \rightarrow \mathbb{R}\}$, $\{\sin x, \cos x\}$ es LI.

Ejercicio 1.1.2. Demuestre las afirmaciones precedentes.

1.1.3. Dimensión de un EV

La *dimensión* de un EV se define como el máximo número de vectores LI que podemos hallar en ese EV. Diremos entonces que la dimensión de V es n si podemos hallar un conjunto $\{\vec{x}_i\}_{i=1}^n$ que sea LI, pero no podemos construir ningún conjunto $\{\vec{x}_i\}_{i=1}^{n+1}$ que sea LI.

La dimensión de un EV V se denota por $\dim(V)$. Una definición compacta es⁴

$$\dim(V) := \max_{\mathbb{N}_0} \{n \in \mathbb{N}_0 / \exists \{\vec{x}_i\}_{i=1}^n \subset V \text{ LI}\}. \quad (1.2)$$

Nótese que la dimensión de un EV no necesita ser finita.

Ejemplo 1.1.7. $\dim(\mathbb{R}^3) = \dim(\mathbb{C}^3) = \dim(\mathbb{Z}^3) = 3$

Ejemplo 1.1.8. $\dim(\{p_N(x)\}) = N + 1$.

Ejemplo 1.1.9. $\dim(\{\sin(nx), \cos(nx), x \in [0, 2\pi], n \in \mathbb{N}\}) = \aleph_0$; el símbolo \aleph_0 , que se lee “aleph cero”, denota la cardinalidad de los naturales, también llamada “infinito numerable”.

Ejemplo 1.1.10. $\dim(\{f : [0, 1] \rightarrow \mathbb{R}\}) = \aleph_1$; el símbolo \aleph_1 , que se lee “aleph uno”, denota la cardinalidad de los reales, también llamada “infinito no numerable”.

Ejercicio 1.1.3. Demuestre las afirmaciones precedentes. Una forma de mostrar la última es considerar el conjunto de las funciones

$$f_a(x) = \begin{cases} 1 & \text{si } x < a, \\ 0 & \text{si } x \geq a, \end{cases}$$

con un parámetro real $a \in [0, 1]$, que es un subconjunto de $\{f : [0, 1] \rightarrow \mathbb{R}\}$, y probar que $f_a(x)$ y $f_{a'}(x)$ son LI si $a' \neq a$.

1.1.4. Base de un EV

Si $\dim(V) = n < \infty$, todo $\{\vec{e}_i\}_{i=1}^n \subset V$ LI se llama *base* de V . Los \vec{e}_i se llaman *vectores base*. Notemos que ninguno de ellos puede ser nulo.

Demostraremos ahora que cualquier vector puede ser escrito como una *única* combinación lineal de los vectores de una base dada.

Teorema 1.1.1. Sea $\dim(V) = n < \infty$, $\{\vec{e}_i\}_{i=1}^n$ base de V . Entonces $\forall \vec{x} \neq \vec{0} \in V$, $\exists! \{x^i\}_{i=1}^n \subset \mathbb{C}$ t.q. $\vec{x} = x^i \vec{e}_i$.

Demotración.

Existencia: Notemos $\vec{x} \neq \vec{0} \implies \{\vec{e}_i\}_{i=1}^n \cup \{\vec{x}\}$ LD. Luego existe solución no trivial para $\lambda_i \vec{e}_i + \mu \vec{x} = \vec{0}$ con $\mu \neq 0$ (porque si $\mu = 0$ la única solución sería la trivial). Sin pérdida de generalidad podemos asumir $\mu = 1$ (dividiendo todo por μ). Luego $\exists \lambda_i \neq 0$ t.q. $\vec{x} = -\lambda_i \vec{e}_i$ y basta tomar $x^i \equiv -\lambda_i$.

⁴La notación $:=$ indica que lo que está a la izquierda se *define* como lo que está a la derecha; la situación inversa se denotará por $=:$.

Unicidad: Supongamos $\vec{x} = x^i \vec{e}_i = y^i \vec{e}_i$. Luego $\vec{0} = \vec{x} - \vec{x} = x^i \vec{e}_i - y^i \vec{e}_i = (x^i - y^i) \vec{e}_i$. Pero $\{\vec{e}_i\}_{i=1}^n$ base de V , luego $x^i - y^i \equiv 0$, es decir $x^i \equiv y^i$.

□

Este resultado no se extiende directamente a EVs de dimensión infinita, sino que se reduce a la desigualdad de Cauchy–Schwartz, que será tratada en los Problemas.

Ejemplo 1.1.11. En \mathbb{R}^3 , si \hat{x} , \hat{y} y \hat{z} son los versores Cartesianos usuales, $\{\hat{x}, \hat{y}, \hat{z}\}$ es base, y $\{\hat{x}, \hat{x} + \hat{y}, \hat{x} + \hat{y} + \hat{z}\}$ también.

Ejemplo 1.1.12. En $\{p_N(x)\}$, $\{1, x, x^2, \dots, x^N\}$ es base.

Ejercicio 1.1.4. Demuestre las afirmaciones precedentes.

Ejemplo 1.1.13. En $\{f : [0, 2\pi] \rightarrow \mathbb{R}, f$ periódica y de variación acotada $\}$, $\{1\} \cup \{\sin(nx), \cos(nx), x \in [0, 2\pi], n \in \mathbb{N}\}$ es base, resultado que conocemos del desarrollo en serie de Fourier (Teorema de Dirichlet).

1.1.5. Componentes de un vector

Dado un EV V de dimensión n (finita), una base $\{\vec{e}_i\}_{i=1}^n$ y un vector $\vec{x} \neq \vec{0} \in V$, los coeficientes x^i de la (única) combinación lineal que expresa

$$\vec{x} = x^i \vec{e}_i \tag{1.3}$$

se llaman *componentes* de \vec{x} en la base $\{\vec{e}_i\}_{i=1}^n$. El Teorema 1.1.1 muestra que los \vec{e}_i , $i = 1, \dots, n$ forman un *conjunto completo*, ya que los x^i existen para *cualquier* vector $\vec{x} \neq \vec{0}$ de V . También muestra que, dada una base, las componentes de un vector en esa base son únicas, de donde es directo deducir los siguientes corolarios:

Corolario 1.1.2. *Dos vectores son iguales si y sólo si sus componentes en una dada base son iguales.*

Corolario 1.1.3. *En cualquier base, las componentes del vector nulo $\vec{x} = \vec{0}$ son $x^i \equiv 0$.*

Ejercicio 1.1.5. Demuestre los dos corolarios precedentes del Teorema 1.1.1.

Ejemplo 1.1.14. Las componentes de $a_0 + a_1 x + a_2 x^2 + \dots + a_N x^N \in \{p_N(x)\}$, en la base $\{1, x, x^2, \dots, x^N\}$, son simplemente sus coeficientes $a_0, a_1, a_2, \dots, a_N$.

Ejemplo 1.1.15. En $\{f : [0, 2\pi] \rightarrow \mathbb{R}, f$ periódica y de variación acotada $\}$, la función definida por $f(x) = x$, $-\pi < x < \pi$, $f(x + 2\pi) \equiv f(x)$ puede escribirse en la base $\{1\} \cup \{\sin(nx), \cos(nx), x \in [0, 2\pi], n \in \mathbb{N}\}$ como $f(x) = \sum_{n=1}^{\infty} (2(-1)^{n+1}/n) \sin(nx)$.

Ejercicio 1.1.6. Demuestre las afirmaciones precedentes.

1.1.6. Subespacios lineales

El tema de subespacios lineales es muy extenso, y la mayor parte de él cae fuera de los objetivos de este curso. Aquí daremos sólo unas pocas definiciones que necesitaremos más adelante.

Dado un EV V , ciertos subconjuntos (propios o no) de sus elementos pueden ser, por derecho propio, EVs. Diremos que W es un *subespacio lineal* o *subespacio vectorial* de V si:

- $\forall \vec{x} \in W, \vec{x} \in V;$
- W es un EV.

Ejemplo 1.1.16. \forall EV V , $\{\vec{0}\}$ y V son subespacios de V .

Ejemplo 1.1.17. El plano (x, y) es un subespacio de \mathbb{R}^3 , y el eje z también.

Ejemplo 1.1.18. $\{p_3(x)\}$ es un subespacio de $\{p_5(x)\}$.

Ejemplo 1.1.19. $\mathcal{C}^N[\mathbb{R}]$, el conjunto de las funciones de \mathbb{R} en \mathbb{R} continuamente diferenciables N veces, es un subespacio de $\mathcal{C}^M[\mathbb{R}]$ si $M \leq N$.

Ejercicio 1.1.7. Demuestre las afirmaciones precedentes.

Dado un conjunto de vectores $\{\vec{x}_i\}_{i=1}^k$ de V , no necesariamente LI, el conjunto de sus combinaciones lineales

$$\text{span}(\{\vec{x}_i\}_{i=1}^k) := \{\lambda_i \vec{x}_i, \lambda_i \in \mathbb{C}\} \quad (1.4)$$

es un subespacio de V , y

$$\dim(\text{span}(\{\vec{x}_i\}_{i=1}^k)) \leq k \quad (= k \text{ si } \{\vec{x}_i\} \text{ LI}). \quad (1.5)$$

Ejemplo 1.1.20. En \mathbb{R}^3 con los versores cartesianos usuales, $\text{span}(\hat{x})$ es el eje x , y $\text{span}(\{\hat{x}, \hat{y}, \hat{x} + \hat{y}\})$ es el plano (x, y) .

Ejemplo 1.1.21. En $\{p_5(x)\}$, $\text{span}(1, x, x^2) = \{p_2(x)\}$

Ejercicio 1.1.8. Demuestre la propiedad (1.5).

Dados dos subespacios U y W de un EV V , su *intersección*

$$U \cap W := \{\vec{x} \in V / \vec{x} \in U \wedge \vec{x} \in W\} \quad (1.6)$$

y su *suma vectorial*

$$U + W := \{\vec{x} + \vec{y} \in V / \vec{x} \in U \wedge \vec{y} \in W\} \quad (1.7)$$

son también subespacios de V . Dos subespacios U y W se llaman *disjuntos* si

$$U \cap W = \{\vec{0}\}, \quad (1.8)$$

es decir si no contienen vectores comunes a ambos, salvo el vector nulo.

Ejercicio 1.1.9. Muestre que dos vectores pertenecientes a sendos subespacios disjuntos son linealmente independientes.

Una noción importante es la de suma directa de subespacios. Consideremos una colección *finita* de subespacios V_i de V , $i = 1, \dots, r$, tales que

- $V_i \cap V_j = \{\vec{0}\}$ si $i \neq j$, es decir son todos disjuntos dos a dos;
- $\forall \vec{x} \in V \exists! \vec{x}_i \in V_i / \vec{x} = \sum_{i=1}^r \vec{x}_i$, es decir que todo vector de V puede descomponerse de forma *única* como suma de “contribuciones” provenientes de cada subespacio.

Entonces decimos que V es *suma directa*⁵ de sus subespacios V_i , y lo denotamos

$$V = V_1 \oplus V_2 \oplus \cdots \oplus V_r = \bigoplus_{i=1}^r V_i. \quad (1.9)$$

Es inmediato que en ese caso

$$\dim(V) = \sum_{i=1}^r \dim(V_i). \quad (1.10)$$

Ejemplo 1.1.22. \mathbb{R}^3 puede descomponerse como suma directa del plano (x, y) y el eje z , o también

$$\mathbb{R}^3 = \text{span}(\hat{x}) \oplus \text{span}(\hat{x} + \hat{y}) \oplus \text{span}(\hat{x} + \hat{y} + \hat{z}),$$

de donde podemos ver que la descomposición de un EV en suma directa de subespacios disjuntos dista de ser única.

Ejercicio 1.1.10. Demuestre la propiedad (1.10).

Otra noción de interés es la de *complemento* de un subespacio: dado $V = X \oplus Y$, diremos que X es un complemento de Y (y viceversa: Y es un complemento de X), o que X e Y son *complementarios*. El ejemplo 1.1.22 muestra claramente que dado un EV V y un subespacio X , el complemento de X no es único.

1.2. Operadores lineales

Los operadores son aplicaciones que toman elementos de un espacio V y devuelven elementos correspondientes pertenecientes a otro espacio W (en general distinto). Estudiaremos algunos casos y propiedades importantes de operadores de un EV a otro.

⁵La generalización de esta idea a espacios de dimensión infinita se llama *producto directo*.

1.2.1. Operadores lineales

Llamaremos *función vectorial* u *operador* de un EV V en otro EV W a una aplicación $\mathcal{A} : V \rightarrow W$, que a cada elemento \vec{x} de V le hace corresponder un elemento $\mathcal{A}(\vec{x})$ de W , es decir

$$V \ni \vec{x} \xrightarrow{\mathcal{A}} \mathcal{A}(\vec{x}) \in W. \quad (1.11)$$

Llamaremos *operador lineal* a un \mathcal{A} tal que

$$\mathcal{A}(\lambda\vec{x} + \mu\vec{y}) \equiv \lambda\mathcal{A}(\vec{x}) + \mu\mathcal{A}(\vec{y}), \quad (1.12)$$

es decir, tal que su acción sea distributiva respecto a la suma en V , y “transparente” a la multiplicación por un escalar.⁶

Ejemplo 1.2.1. El operador que rota (un ángulo dado) alrededor del eje z : $\mathcal{R}_z : \mathbb{R}^3 \rightarrow \mathbb{R}^3$; el operador que proyecta sobre el plano (x, y) : $\mathcal{P}_z : \mathbb{R}^3 \rightarrow \mathbb{R}^2$ (que también puede pensarse como $\mathcal{P}_z : \mathbb{R}^3 \rightarrow \mathbb{R}^3$).

Ejemplo 1.2.2. El operador que deriva respecto a x : $\partial_x : \{p_N(x)\} \rightarrow \{p_{N-1}(x)\}$; el operador que integra respecto a x : $\int dx : \{p_{N-1}(x)\} \rightarrow \{p_N(x)\}$.

Dado un operador $\mathcal{A} : V \rightarrow W$ y un vector $\vec{x} \in V$, llamaremos a $\vec{y} = \mathcal{A}\vec{x}$ la *imagen* de \vec{x} bajo \mathcal{A} , y a \vec{x} una *preimagen* de \vec{y} bajo \mathcal{A} . Para \mathcal{A} lineal, la imagen es única, pero la preimagen no necesariamente lo es.

Ejemplo 1.2.3. En \mathbb{R}^3 , la imagen de $2\hat{x} + \hat{y} - 3\hat{z}$ bajo \mathcal{P}_z es $2\hat{x} + \hat{y}$, pero cualquier vector de la forma $2\hat{x} + \hat{y} + \lambda\hat{z}$ con $\lambda \in \mathbb{C}$ es preimagen de $2\hat{x} + \hat{y}$ bajo \mathcal{P}_z .

Ejemplo 1.2.4. nx^{n-1} es la imagen de x^n bajo ∂_x , pero cualquier polinomio de la forma $a + x^n$ con $a \in \mathbb{R}$ es preimagen de nx^{n-1} bajo ∂_x .

Dado un operador lineal $\mathcal{A} : V \rightarrow W$, el conjunto de vectores de V que son llevados al vector nulo en W se conoce como el *núcleo (kernel)* de \mathcal{A} , y se denota por

$$\ker(\mathcal{A}) := \{\vec{x} \in V / \mathcal{A}\vec{x} = \vec{0} \in W\}. \quad (1.13)$$

Es fácil ver que

$$\dim(\ker(\mathcal{A})) \geq \dim(V) - \dim(W). \quad (1.14)$$

Ejemplo 1.2.5. En \mathbb{R}^3 , $\ker(\mathcal{P}_z) = \text{span}(\hat{z})$.

Ejemplo 1.2.6. En $\{p_N(x)\}$, $\ker(\partial_x) = \{p_0(x)\}$.

Ejercicio 1.2.1. Demuestre que $\ker(\mathcal{A})$ es un subespacio de V .

Ejercicio 1.2.2. Demuestre la desigualdad (1.14).

⁶En adelante, cuando un operador sea lineal omitiremos los paréntesis del argumento siempre que ello no cree ambigüedad, poniendo simplemente $\mathcal{A}\vec{x}$ en lugar de $\mathcal{A}(\vec{x})$; ya que esta notación recuerda a la de producto usual, que comparte las propiedades de linealidad.

Dado un operador lineal $\mathcal{A} : V \rightarrow V$, un subespacio W de V se llama *invariante bajo \mathcal{A}* si $\mathcal{A}\vec{x} \in W \forall \vec{x} \in W$.

Ejemplo 1.2.7. En \mathbb{R}^3 , $\text{span}(\{\hat{x}, \hat{y}\})$ y $\text{span}(\hat{z})$ son invariantes bajo \mathcal{R}_z .

Ejemplo 1.2.8. En $C^1[\mathbb{R}]$, $\text{span}(\{e^{ax}, a \in \mathbb{R}\})$ es invariante bajo ∂_x .

Ejercicio 1.2.3. Demuestre las afirmaciones precedentes.

Una clase de operadores lineales íntimamente relacionados a los subespacios de un EV V , y que usaremos mas adelante, son los llamados proyectores. Dado $V = X \oplus Y$ y $\vec{z} = \vec{x} + \vec{y}$ con $\vec{x} \in X$ e $\vec{y} \in Y$, el *proyector sobre X* (a lo largo de Y) se define por

$$\mathcal{P}_X \vec{z} = \vec{x}, \quad (1.15)$$

y es el operador que “extrae la parte de \vec{z} que está en X ”. Es inmediato que

$$\dim(\ker(\mathcal{P}_W)) = \dim(V) - \dim(W). \quad (1.16)$$

Ejercicio 1.2.4. Muestre que \mathcal{P}_X es *idempotente*: $\mathcal{P}_X^2 = \mathcal{P}_X$.

Si un dado subespacio es invariante bajo un operador lineal, entonces la acción del operador sobre los vectores de ese subespacio puede estudiarse ignorando el resto del espacio vectorial, lo que en general simplifica el problema. La formalización de esta idea lleva a definir la restricción de un operador a un subespacio: dado un operador lineal $\mathcal{A} : V \rightarrow V$ y un subespacio W invariante bajo \mathcal{A} , la *restricción de \mathcal{A} a W* se define como el operador $\mathcal{A}/W : W \rightarrow W$ que actúa sobre los elementos de W exactamente como lo haría \mathcal{A} . Debe notarse que \mathcal{A}/W actúa sobre W y *no* sobre V . En las aplicaciones que desarrollaremos más adelante, sin embargo, será mas conveniente pensar a \mathcal{A}/W como un operador sobre V , para lo cual lo redefiniremos

$$\mathcal{A}/W := \mathcal{A}\mathcal{P}_W. \quad (1.17)$$

1.2.2. Componentes de un operador lineal

Veremos ahora que un operador lineal puede descomponerse, en forma análoga a un vector, en sus componentes respecto a bases dadas.

Sea $\{\vec{e}_i\}_{i=1}^n$ base de un EV V [$\dim(V) = n$], y sea $\vec{x} = x^i \vec{e}_i \in V$.

Sea $\{\vec{f}_i\}_{i=1}^m$ base de un EV W [$\dim(W) = m$], y sea $\vec{y} = y^j \vec{f}_j \in W$.

Sea $\mathcal{A} : V \rightarrow W$ un operador lineal tal que $\vec{y} = \mathcal{A}\vec{x}$.

Por linealidad, $\mathcal{A}(\vec{x}) = \mathcal{A}(x^i \vec{e}_i) = x^i \mathcal{A}\vec{e}_i$.

Pero para cada \vec{e}_i , $\mathcal{A}\vec{e}_i$ es un vector en W ; luego $\exists!$ coeficientes A_i^j t.q. $\mathcal{A}\vec{e}_i = A_i^j \vec{f}_j$; entonces $y^j \vec{f}_j = A_i^j x^i \vec{f}_j$.

Como los \vec{f}_j forman base de W , podemos igualar las componentes: $y^j = A_i^j x^i$.

Vemos entonces que dado un operador lineal $\mathcal{A} : V \rightarrow W$, su acción sobre un vector cualquiera $\vec{x} \in V$ queda completamente determinada dando los coeficientes A_i^j , que se denominan *componentes* de \mathcal{A} en el par de bases $\{\vec{e}_i\}$, $\{\vec{f}_j\}$.

Dicho de otra forma, un operador lineal queda completamente caracterizado por su acción sobre cada uno de los vectores de una base.

Debemos notar que, así como los vectores tienen en general un significado físico independiente del sistema de coordenadas (es decir, de la base en que expresamos sus componentes), los operadores también tienen un significado físico independiente del sistema de coordenadas. Volveremos sobre esta idea en la Sección 1.4.

Ejemplo 1.2.9. Consideremos una rotación de ángulo θ (en sentido positivo) alrededor del eje z en \mathbb{R}^3 , y llamemos \mathcal{R} al correspondiente operador lineal. En este caso $V = W = \mathbb{R}^3$, y podemos tomar $\vec{f}_i \equiv \vec{e}_i$, con \vec{e}_i los versores Cartesianos usuales. Sea $\vec{x} = x^i \vec{e}_i$ un vector cualquiera en \mathbb{R}^3 ; el vector rotado será $\vec{y} = \mathcal{R}\vec{x} = x^i \mathcal{R}\vec{e}_i$, con

$$\mathcal{R}\vec{e}_i = \begin{cases} \cos \theta \vec{e}_1 + \sin \theta \vec{e}_2 & \text{si } i = 1, \\ -\sin \theta \vec{e}_1 + \cos \theta \vec{e}_2 & \text{si } i = 2, \\ \vec{e}_3 & \text{si } i = 3. \end{cases}$$

Las componentes de \mathcal{R} en el par de bases $\{\vec{e}_i\}$, $\{\vec{e}_i\}$ serán entonces los elementos R_i^j de la matriz

$$[R_i^j] = \begin{pmatrix} \cos \theta & -\sin \theta & 0 \\ \sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

y las componentes del vector rotado serán $y^j = R_i^j x^i$.

Ejemplo 1.2.10. Consideremos el operador $\partial_x : \{p_3(x)\} \rightarrow \{p_2(x)\}$; sus componentes en el par de bases $\{1, x, x^2, x^3\}$ de $\{p_3(x)\}$ y $\{1, x, x^2\}$ de $\{p_2(x)\}$ serán los elementos de la matriz

$$[\partial_i^j] = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix},$$

y el resultado de aplicarlo al elemento $ax - bx^3 \in \{p_3(x)\}$ será

$$\partial_x(ax - bx^3) = a - 3bx^2.$$

Ejercicio 1.2.5. Verifique lo afirmado en los dos ejemplos precedentes.

1.2.3. Operaciones elementales

Las operaciones elementales con operadores lineales se definen a partir de sus acciones sobre los vectores de un EV:

Igualdad: Dos operadores lineales $\mathcal{A}, \mathcal{B} : V \rightarrow W$ se dicen *iguales* si $\mathcal{A}\vec{x} \equiv \mathcal{B}\vec{x}$, y lo denotamos $\mathcal{A} = \mathcal{B}$.

Suma: Dados dos operadores lineales $\mathcal{A}, \mathcal{B} : V \rightarrow W$, su *suma* $\mathcal{C} := \mathcal{A} + \mathcal{B}$ se define como el operador lineal que realiza la asignación $\mathcal{C}\vec{x} \equiv \mathcal{A}\vec{x} + \mathcal{B}\vec{x}$.

Producto: Dados dos operadores lineales $\mathcal{A} : V \rightarrow W$, $\mathcal{B} : W \rightarrow U$, su *composición o producto* $\mathcal{C} := \mathcal{B} \circ \mathcal{A} = \mathcal{B}\mathcal{A}$ se define como el operador lineal $\mathcal{C} : V \rightarrow U$ que realiza la asignación $\mathcal{C}\vec{x} \equiv \mathcal{B}(\mathcal{A}\vec{x})$.

Producto por un escalar: Dado un operador lineal $\mathcal{A} : V \rightarrow W$, su *producto por un escalar* $\lambda \in \mathbb{C}$ se define como el operador lineal $\lambda\mathcal{A} : V \rightarrow W$ que realiza la asignación $(\lambda\mathcal{A})\vec{x} \equiv \lambda(\mathcal{A}\vec{x})$.

Identidad: El *operador identidad* $\mathcal{I} : V \rightarrow V$ se define como el que realiza la asignación $\mathcal{I}\vec{x} \equiv \vec{x}$.

Las consecuencias de estas definiciones sobre las componentes de los operadores lineales, serán tratadas en el contexto de matrices en la Sección 1.3.

1.2.4. Conmutatividad

Diremos que dos operadores lineales \mathcal{A} y \mathcal{B} *conmutan* si $\mathcal{A}\mathcal{B} = \mathcal{B}\mathcal{A}$. Para ello es necesario que ambas composiciones tengan sentido, así que es condición necesaria que $V \xrightarrow{\mathcal{A}} V \xrightarrow{\mathcal{B}} V$; pero no es condición suficiente. En general, $\mathcal{A}\mathcal{B} \neq \mathcal{B}\mathcal{A}$.

Ejercicio 1.2.6. Usando las definiciones de las operaciones elementales entre operadores lineales, muestre que $\mathcal{A}\mathcal{B} = \mathcal{B}\mathcal{A}$ requiere $V \xrightarrow{\mathcal{A}} V \xrightarrow{\mathcal{B}} V$.

Ejemplo 1.2.11. Consideremos nuevamente rotaciones en \mathbb{R}^3 . En particular, sea \mathcal{R}_x una rotación de ángulo $\pi/2$ en sentido positivo alrededor del eje x , y sea \mathcal{R}_y una rotación de ángulo $\pi/2$ en sentido positivo alrededor del eje y . Es fácil ver que $\mathcal{R}_x\hat{y} = \hat{z}$ y $\mathcal{R}_y\hat{z} = \hat{x}$, así que $\mathcal{R}_y\mathcal{R}_x\hat{y} = \hat{x}$. Por otro lado, $\mathcal{R}_y\hat{y} = \hat{y}$ y $\mathcal{R}_x\hat{y} = \hat{z}$, de modo que $\mathcal{R}_x\mathcal{R}_y\hat{y} = \hat{z}$. Como hay al menos un vector (\hat{y}) cuyas imágenes bajo $\mathcal{R}_y\mathcal{R}_x$ y bajo $\mathcal{R}_x\mathcal{R}_y$ son distintas, la definición de igualdad entre operadores nos dice que $\mathcal{R}_y\mathcal{R}_x \neq \mathcal{R}_x\mathcal{R}_y$, es decir, *no conmutan*.

Dados dos operadores lineales $\mathcal{A}, \mathcal{B} : V \rightarrow V$ definimos su *conmutador* como

$$[\mathcal{A}, \mathcal{B}] = \mathcal{A}\mathcal{B} - \mathcal{B}\mathcal{A}. \quad (1.18)$$

Es evidente que $[\mathcal{B}, \mathcal{A}] = -[\mathcal{A}, \mathcal{B}]$, $[\mathcal{A}, \mathcal{A}] = 0$ y $[\mathcal{A}, \mathcal{I}] = 0$.

1.2.5. Inversa

Un operador lineal lleva elementos de un EV a otro, o posiblemente al mismo, así que cabe preguntarse si existirá otro operador que “deshaga” lo que este hizo. Esto nos lleva a la definición de inversa de un operador. Dado $\mathcal{A} : V \rightarrow W$ y $\mathcal{B} : W \rightarrow V$ diremos que \mathcal{B} es la *inversa* de \mathcal{A} si

$$\mathcal{B}\mathcal{A} = \mathcal{I}, \quad (1.19)$$

y usualmente denotaremos $\mathcal{B} = \mathcal{A}^{-1}$. Nótese que la inversa de $\mathcal{A} : V \rightarrow W$ existe sólo si $\dim(W) \geq \dim(V)$, pero esta condición no es suficiente.

Ejemplo 1.2.12. Consideremos $\mathcal{A} : \mathbb{R}^2 \rightarrow \mathbb{R}^3$ dado por

$$\mathcal{A}\vec{e}_i = \begin{cases} \vec{e}_1 + \vec{e}_3 & \text{si } i = 1, \\ \vec{e}_2 + \vec{e}_3 & \text{si } i = 2, \end{cases}$$

y sea $\mathcal{P} : \mathbb{R}^3 \rightarrow \mathbb{R}^2$ definido como

$$\mathcal{P}\vec{e}_i = \begin{cases} \vec{e}_1 & \text{si } i = 1, \\ \vec{e}_2 & \text{si } i = 2, \\ \vec{0} & \text{si } i = 3. \end{cases}$$

Entonces es evidente que $\mathcal{P}\mathcal{A} = \mathcal{I} : \mathbb{R}^2 \rightarrow \mathbb{R}^2$, ya que

$$\mathcal{P}\mathcal{A}\vec{e}_i = \begin{cases} \mathcal{P}(\vec{e}_1 + \vec{e}_3) = \vec{e}_1 & \text{si } i = 1, \\ \mathcal{P}(\vec{e}_2 + \vec{e}_3) = \vec{e}_2 & \text{si } i = 2, \end{cases}$$

y podemos poner $\mathcal{P} = \mathcal{A}^{-1}$. Sin embargo $\mathcal{A}\mathcal{P} \neq \mathcal{I} : \mathbb{R}^3 \rightarrow \mathbb{R}^3$, ya que por ejemplo $\mathcal{A}\mathcal{P}(\vec{e}_1 + \lambda\vec{e}_3) = \mathcal{A}\vec{e}_1 = \vec{e}_1 + \vec{e}_3 \forall \lambda \in \mathbb{C}$. Es más, podemos ver que $\nexists \mathcal{B} : \mathbb{R}^2 \rightarrow \mathbb{R}^3$ tal que $\mathcal{B}\mathcal{P} = \mathcal{I}$, justamente debido a que $\mathcal{P}(\lambda\vec{e}_3) = \vec{0} \forall \lambda$: la información sobre la componente \vec{e}_3 de un vector dado, es destruida irreversiblemente por la acción de \mathcal{P} , y no habrá operador que la recupere.

Ejemplo 1.2.13. Consideremos el espacio $V = \mathcal{C}^1[\mathbb{R}]$ y los operadores derivada ∂_x e integral $\int dx$. Es fácil ver que $\partial_x \int dx = \mathcal{I}$, sin embargo $\int dx \partial_x \neq \mathcal{I}$, como se comprueba aplicándolo a la función constante.

De los ejemplos anteriores se desprende que

$$\mathcal{B}\mathcal{A} = \mathcal{I} \quad \Rightarrow \quad \mathcal{A}\mathcal{B} = \mathcal{I}. \tag{1.20}$$

También vemos que el problema para invertir un operador surge cuando el kernel del operador es no trivial, lo que nos lleva a enunciar el siguiente

Teorema 1.2.1. *Un operador lineal \mathcal{A} es invertible si $\dim(\ker(\mathcal{A})) = 0$.*

La demostración puede consultarse en cualquier buen texto de Algebra Lineal.

1.2.6. Funciones de operadores

Dado que hemos definido las operaciones elementales (suma, producto, etc.) entre operadores lineales, cabe preguntarse si podremos definir funciones más “complicadas” de un operador, y cómo. Responderemos parcialmente esta pregunta para una clase restringida de funciones.

De lo hecho hasta aquí es inmediato que, dado un operador $\mathcal{A} : V \rightarrow V$, podemos definir potencias enteras positivas arbitrarias \mathcal{A}^n del mismo.⁷ Dada entonces una función

$$f(x) = \sum_{n=0}^{\infty} a_n x^n$$

cuyo desarrollo en serie tenga un radio de convergencia $r > 0$, podemos construir

$$f(\mathcal{A}) = \sum_{n=0}^{\infty} a_n \mathcal{A}^n,$$

con la interpretación obvia

$$f(\mathcal{A})\vec{v} = \sum_{n=0}^{\infty} a_n \mathcal{A}^n \vec{v}.$$

En cuanto al “radio de convergencia” de esta serie, lo discutiremos mas adelante, una vez hayamos definido normas de operadores.

Ejercicio 1.2.7. Sea $V = \mathcal{C}^\infty[\mathbb{R}]$, y considere el operador derivada $\partial_x : V \rightarrow V$. Muestre que para $f \in V$ y $a \in \mathbb{R}$,

$$e^{a\partial_x} f(x) = f(x + a).$$

Ejercicio 1.2.8. ¿Qué sentido asignaría al operador $(1 + \partial_x)^{-1}$?

1.3. Matrices

Sea $\mathcal{A} : V \rightarrow W$ un operador lineal, y A_i^j sus componentes en el par de bases $\{\vec{e}_i\}_{i=1}^n$ de V y $\{\vec{f}_j\}_{j=1}^m$ de W . Podemos escribir estas componentes como una matriz $m \times n$ con elementos A_i^j , en la forma

$$\mathbf{A} = [A_i^j] := \begin{pmatrix} A_1^1 & A_2^1 & \dots & A_n^1 \\ A_1^2 & A_2^2 & \dots & A_n^2 \\ \vdots & \vdots & & \vdots \\ A_1^m & A_2^m & \dots & A_n^m \end{pmatrix} \quad (m \text{ filas por } n \text{ columnas}). \quad (1.21)$$

De la misma forma en que usualmente representamos a un vector por la n -upla formada por sus componentes en una base dada, representaremos en general a un operador lineal por la matriz $m \times n$ formada por sus componentes en un par de bases dadas, y nos referiremos a esa matriz como la *representación* del operador en esas bases.

⁷Por consistencia, definimos $\mathcal{A}^0 := \mathcal{I}$.

1.3.1. Propiedades fundamentales

La definición y propiedades fundamentales de las matrices y las operaciones entre ellas, pueden deducirse de aquellas de los operadores lineales que representan:

- Por la definición de **igualdad** entre operadores lineales, dos matrices serán *iguales*, $A = B$, si sus componentes o *elementos* son iguales: $A_i^j \equiv B_i^j$.
- Por la definición de **suma** entre operadores lineales, la *suma* de dos matrices A y B satisface $[A + B]_i^j \equiv A_i^j + B_i^j$; es decir, los elementos de la suma son la suma de los elementos.
- Por la definición de **producto** entre operadores lineales, el *producto* de dos matrices A y B satisface $[AB]_i^j \equiv A_k^j B_i^k$; es decir, el elemento $_{i,j}^j$ de AB es la suma de los productos de los elementos de la fila j de A por los correspondientes elementos de la columna i de B . Evidentemente, el número de *columnas* de A debe ser igual al número de *filas* de B para que el producto AB tenga sentido. Recuperamos así la conocida regla de multiplicación entre matrices.
- Por la definición de **producto por un escalar** de un operador lineal, el *producto por un escalar* λ de una matriz A satisface $[\lambda A]_i^j = \lambda A_i^j$; es decir, multiplicar a una matriz por un escalar multiplica cada uno de sus elementos por ese mismo escalar.
- Por la definición del **operador identidad** \mathcal{I} , sus componentes en cualquier base resultan $I_i^j \equiv \delta_i^j$, donde δ_i^j es la *delta de Kronecker*; por lo tanto la matriz que representa a \mathcal{I} es la *matriz identidad*

$$\mathbf{I} := \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix}.$$

La matriz identidad *siempre* es cuadrada; cuando sea necesario indicar su dimensión, lo haremos por medio de un subíndice: \mathbf{I}_n indica la matriz identidad $n \times n$.

Ejercicio 1.3.1. Demuestre las propiedades precedentes.

Ejemplo 1.3.1. Como ayuda para el ejercicio anterior, para la suma: Sean $\mathcal{A}, \mathcal{B} : V \rightarrow W$ operadores lineales, $\{\vec{e}_i\}_{i=1}^n$ base de V , $\{\vec{f}_j\}_{j=1}^m$ base de W , $\vec{x} = x^i \vec{e}_i \in V$. Sabemos que $(\mathcal{A} + \mathcal{B})\vec{x} = \mathcal{A}\vec{x} + \mathcal{B}\vec{x} \forall \vec{x} \in V$. Luego $[\mathcal{A} + \mathcal{B}]_i^j x^i \vec{f}_j = A_i^j x^i \vec{f}_j + B_i^j x^i \vec{f}_j = (A_i^j + B_i^j) x^i \vec{f}_j$. Como $\{\vec{f}_j\}$ es base, podemos igualar las componentes: $[\mathcal{A} + \mathcal{B}]_i^j x^i = (A_i^j + B_i^j) x^i \forall j$. Pero esta igualdad debe valer $\forall \vec{x}$, es decir $\forall x^i$, luego $[\mathcal{A} + \mathcal{B}]_i^j \equiv (A_i^j + B_i^j)$. Q.E.D.

Debemos destacar que, de igual modo que la composición de operadores lineales (de la que deriva sus propiedades), el producto de matrices en general *no es conmutativo*, es decir $\mathbf{AB} \neq \mathbf{BA}$. Dos matrices *comutan*, esto es $\mathbf{AB} = \mathbf{BA}$, si los correspondientes operadores lo hacen, es decir si $\mathcal{A}\mathcal{B} = \mathcal{B}\mathcal{A}$ (y por ello mismo, si dos matrices comutan, lo hacen en cualquier base).

Notemos que si $\vec{x} = x^i \vec{e}_i \in V$ [$\dim(V) = n$], podemos ordenar las componentes x^i de \vec{x} como una *matriz columna* $n \times 1$

$$\mathbf{x} = \begin{pmatrix} x^1 \\ x^2 \\ \vdots \\ x^n \end{pmatrix}.$$

Si además $\vec{y} = \mathcal{A}\vec{x} = y^j \vec{f}_j \in W$ [$\dim(W) = m$], entonces $y^j = A_{\cdot i}^j x^i$, y vemos que la notación de matriz columna es consistente con la regla de producto entre matrices:

$$\mathbf{y} = \begin{pmatrix} y^1 \\ y^2 \\ \vdots \\ y^m \end{pmatrix} = \begin{pmatrix} A_{11} & A_{12} & \dots & A_{1n} \\ A_{21} & A_{22} & \dots & A_{2n} \\ \vdots & \vdots & & \vdots \\ A_{m1} & A_{m2} & \dots & A_{mn} \end{pmatrix} \begin{pmatrix} x^1 \\ x^2 \\ \vdots \\ x^n \end{pmatrix} = \mathbf{Ax}.$$

1.3.2. Definiciones

Dada una matriz \mathbf{A} , existen diversas matrices relacionadas a ella;⁸ algunas de las mas comunes son:

Matriz	Notación	Componentes	Ejemplo
\mathbf{A}	A_i^j		$\begin{pmatrix} 1 & i \\ 1+i & 2 \end{pmatrix}$
Conjugada	\mathbf{A}^*	$[A^*]_i^j = (A_i^j)^*$	$\begin{pmatrix} 1 & -i \\ 1-i & 2 \end{pmatrix}$
Transpuesta	\mathbf{A}^T	$[A^T]_i^j = A_j^i$	$\begin{pmatrix} 1 & 1+i \\ i & 2 \end{pmatrix}$
Adjunta	\mathbf{A}^\dagger	$[A^\dagger]_i^j = (A_j^i)^*$	$\begin{pmatrix} 1 & 1-i \\ -i & 2 \end{pmatrix}$
Inversa	\mathbf{A}^{-1}	$[A^{-1}]_i^j = \frac{\text{cof } A^i_j}{\det \mathbf{A}}$	$\frac{1}{3-i} \begin{pmatrix} 2 & -i \\ -1-i & 1 \end{pmatrix}$

En particular, si \mathbf{A} es la matriz que representa a un operador lineal \mathcal{A} , entonces la *matriz inversa* \mathbf{A}^{-1} es la matriz que representa al operador inverso \mathcal{A}^{-1} ; esto puede verificarse fácilmente escribiendo ambos operadores en componentes. De la misma forma que el producto de un operador por su inversa resulta en el operador identidad, $\mathcal{A}^{-1}\mathcal{A} = \mathcal{I}$, el producto de una matriz por su inversa resulta en la matriz identidad, $\mathbf{A}^{-1}\mathbf{A} = \mathbf{I}$.

Como se recordará, la matriz inversa de \mathbf{A} existe si \mathbf{A} es *no singular*, es decir si $\det(\mathbf{A}) \neq 0$, de donde podemos escribir

$$\det(\mathbf{A}) \neq 0 \Leftrightarrow \dim(\ker(\mathcal{A})) = 0. \quad (1.22)$$

⁸Restringimos lo que sigue al caso $\dim(W) = \dim(V)$, es decir a matrices *cuadradas*.

La fórmula para los elementos de A^{-1} se justifica como sigue: Recordemos que el *cofactor* del elemento A_{ij}^j se define como

$$\text{cof } A_{ij}^j := (-1)^{i+j} \times \det[\mathbf{A} \text{ sin su fila } j \text{ y columna } i].$$

Desarrollando $\det(\mathbf{A})$ por la fila j vemos que⁹

$$\det(\mathbf{A}) = A_{ii}^j \text{ cof } A_{ii}^j,$$

luego

$$[\mathbf{AA}^{-1}]_i^j = A_k^j [\mathbf{A}^{-1}]_i^k = A_k^j \frac{\text{cof } A_k^i}{\det \mathbf{A}} = \delta_i^j \equiv [I]_i^j,$$

ya que si $i \neq j$, $A_k^j \text{ cof } A_k^i = \det[\text{matriz con filas } j \text{ e } i \text{ iguales}] = 0$, mientras que si $i = j$, $A_k^j \text{ cof } A_k^j = \det(\mathbf{A})$ (desarrollado por la fila j).

Ejercicio 1.3.2. Consideré el operador \mathcal{P}_z que proyecta de \mathbb{R}^3 al plano xy , como un operador de \mathbb{R}^3 en \mathbb{R}^3 . Escriba la matriz P_z que lo representa en la base Cartesiana usual, y verifique que no tiene inversa.

Ejercicio 1.3.3. Consideré el operador $\mathcal{R}_{z,\theta}$ en \mathbb{R}^3 , que rota un ángulo θ en sentido positivo alrededor del eje z . Escriba la matriz $R_{z,\theta}$ que lo representa en la base Cartesiana usual, y muestre que su inversa es la matriz que representa al operador $\mathcal{R}_{z,-\theta}$.

Ejercicio 1.3.4. Demuestre que:

1. $(AB)^{-1} = B^{-1}A^{-1}$.
2. $(AB)^T = B^T A^T$.
3. $\text{Tr}(AB) = \text{Tr}(BA)$.
4. $\det(AB) = \det(BA) = \det(A)\det(B)$.

1.3.3. Matrices notables

Dependiendo de sus propiedades, diversas matrices cuadradas reciben nombres especiales. Algunos de los mas usuales son:

⁹Mientras se utiliza el convenio de suma sobre índices repetidos, la aparición de índices *subrayados* significa que *no sumamos* sobre ellos; en la fórmula del texto, ponemos \underline{j} para indicar que no sumamos sobre j .

Propiedad	Definición	Ejemplo	Notas
Real	$A^* = A$	$\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$	
Simétrica	$A^T = A$	$\begin{pmatrix} i & 1 \\ 1 & 2 \end{pmatrix}$	
Antisimétrica	$A^T = -A$	$\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$	$A_{\frac{i}{2}}^i \equiv 0$
Autoadjunta	$A^\dagger = A$	$\begin{pmatrix} 1 & i \\ -i & 2 \end{pmatrix}$	$\det(A)$ real
Ortogonal	$A^{-1} = A^T$	$\begin{pmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{pmatrix}$	$\det(A) = \pm 1$
Unitaria	$A^{-1} = A^\dagger$	$\frac{1}{2} \begin{pmatrix} 1 & i \\ -i & -1 \end{pmatrix}$	$ \det(A) = 1$
Diagonal	$A_i^j = 0 \forall i \neq j$	$\begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}$	
Idempotente	$A^2 = A$	$\frac{1}{2} \begin{pmatrix} 1 & i \\ -i & 1 \end{pmatrix}$	$\det(A) = 1$
Nilpotente	$A^2 = 0$	$\begin{pmatrix} 1 & 1 \\ -1 & -1 \end{pmatrix}$	singular

Ejercicio 1.3.5. Considere las matrices de los Ejercicios 1.3.2 y 1.3.3. ¿En qué categoría de las de arriba clasificaría a cada una?

Ejercicio 1.3.6. Refiérase al Ejercicio 1.3.4. Para A y B en cada categoría de las de arriba, ¿qué puede concluir?

1.3.4. Operación por bloques

En los cálculos con matrices, muchas veces podemos simplificar las operaciones subdividiendo las matrices involucradas en *bloques* o submatrices. Por ejemplo, sean

$$A := \begin{pmatrix} A_1^1 & A_1^2 & \cdots & A_3^1 & A_4^1 & A_5^1 \\ A_1^2 & A_2^2 & \cdots & A_3^2 & A_4^2 & A_5^2 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ A_1^3 & A_2^3 & \cdots & A_3^3 & A_4^3 & A_5^3 \\ A_1^4 & A_2^4 & \cdots & A_3^4 & A_4^4 & A_5^4 \\ A_1^5 & A_2^5 & \cdots & A_3^5 & A_4^5 & A_5^5 \end{pmatrix} = \begin{pmatrix} A_1^1 & A_1^2 \\ A_2^1 & A_2^2 \end{pmatrix},$$

$$B := \begin{pmatrix} B_1^1 & B_1^2 & \cdots & B_3^1 & B_4^1 & B_5^1 \\ B_1^2 & B_2^2 & \cdots & B_3^2 & B_4^2 & B_5^2 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ B_1^3 & B_2^3 & \cdots & B_3^3 & B_4^3 & B_5^3 \\ B_1^4 & B_2^4 & \cdots & B_3^4 & B_4^4 & B_5^4 \\ B_1^5 & B_2^5 & \cdots & B_3^5 & B_4^5 & B_5^5 \end{pmatrix} = \begin{pmatrix} B_1^1 & B_1^2 \\ B_2^1 & B_2^2 \end{pmatrix},$$

con

$$A_1^1 = \begin{pmatrix} A_1^1 & A_1^2 \\ A_2^1 & A_2^2 \end{pmatrix}, \quad A_2^1 = \begin{pmatrix} A_3^1 & A_4^1 & A_5^1 \\ A_2^2 & A_4^2 & A_5^2 \end{pmatrix}, \quad \text{etc.},$$

y similarmente para B_i^j . Entonces podemos calcular el producto AB como

$$AB = \begin{pmatrix} A_1^1 & A_1^2 \\ A_2^1 & A_2^2 \end{pmatrix} \begin{pmatrix} B_1^1 & B_1^2 \\ B_2^1 & B_2^2 \end{pmatrix} = \begin{pmatrix} A_1^1 B_1^1 + A_1^2 B_2^1 & A_1^1 B_1^2 + A_1^2 B_2^2 \\ A_2^1 B_1^1 + A_2^2 B_2^1 & A_2^1 B_1^2 + A_2^2 B_2^2 \end{pmatrix}.$$

Notemos que es necesario, para proceder de esta manera, que cada uno de los productos y sumas matriciales que aparezcan tengan sentido, es decir, los bloques en que subdividimos las matrices deben tener números de filas y columnas consistentes. En particular, si la matriz es cuadrada, los bloques que yacen sobre la diagonal principal deben ser siempre cuadrados.

Así como definíamos una matriz *triangular* cuando todos los elementos por encima (o por debajo) de la diagonal principal eran nulos, diremos que una matriz es *triangular por bloques* si por ejemplo

$$A = \begin{pmatrix} A_1^1 & 0 & \dots & 0 \\ A_1^2 & A_2^2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ A_1^n & A_2^n & \dots & A_n^n \end{pmatrix}.$$

El cálculo del determinante se simplifica notablemente en este caso, ya que resulta

$$\det(A) = \prod_{j=1}^n \det(A_j^j).$$

Ejercicio 1.3.7. Sean P_z y R_z las representaciones de los operadores proyección sobre el plano xy y rotación alrededor de z , respectivamente, en la base Cartesiana usual. Calcule su producto y muestre que comutan.

1.3.5. Funciones de matrices

Retomando lo hecho en la Sección 1.2.6, vemos que si A es la representación de un operador $\mathcal{A} : V \rightarrow V$ en una dada base, deberíamos definir

$$f(A) = \sum_{n=0}^{\infty} a_n A^n,$$

donde la función f tiene un desarrollo de Taylor $f(x) = \sum_{n=1}^{\infty} a_n x^n$. Seguiremos postergando la discusión del radio de convergencia hasta que hayamos definido la norma de una matriz.

Ejercicio 1.3.8. Sea $A = \begin{pmatrix} x & a \\ 0 & x \end{pmatrix}$. Calcule e^A .

Ejercicio 1.3.9. Calcule $\exp\left(\begin{pmatrix} 0 & x \\ -x & 0 \end{pmatrix}\right)$ y $\exp\left(\begin{pmatrix} 0 & x \\ x & 0 \end{pmatrix}\right)$.

1.4. Transformaciones de coordenadas

En Física suponemos que la descripción que hagamos de un problema, las leyes físicas que formulemos, etc., deben ser independientes del sistema de coordenadas en que estemos trabajando. Por ejemplo, cuando decimos que en una onda electromagnética los vectores campo eléctrico y campo magnético son ortogonales, o cuando calculamos los niveles de energía del electrón de un átomo

de hidrógeno, no esperamos que estos resultados cambien si cambiamos de sistema de coordenadas. Sin embargo, cuando tratamos con entes como vectores (p.ej. campos eléctrico y magnético en Electromagnetismo) u operadores (p.ej. el operador energía en Mecánica Cuántica) necesitamos, por una cuestión práctica, describirlos por sus componentes en algún sistema de coordenadas dado, es decir, trabajamos con su *representación* en algún sistema de coordenadas. Y cuando cambiamos a otro sistema de coordenadas, estas componentes cambian. Es importante, entonces, saber *cómo* cambian, de modo de poder distinguir si una propiedad dada, es propiedad de los vectores u operadores, o sólo de su representación en un dado sistema de coordenadas.

Comenzaremos estudiando qué ocurre con las componentes de un vector ante un cambio de base, y veremos luego qué pasa con los operadores.

1.4.1. Matriz de transformación

Sea V un EV, $\{\vec{e}_i\}_{i=1}^n$ una base de V , y $\vec{x} = x^i \vec{e}_i \in V$. Sea $\{\vec{e}'_j\}_{j=1}^n$ otra base de V . Entonces, por ser $\{\vec{e}_i\}$ base de V , para cada $\vec{e}'_j \exists! \gamma^i_j \neq 0$ tales que

$$\vec{e}'_j = \vec{e}_i \gamma^i_j. \quad (1.23)$$

Las componentes γ^i_j de los vectores \vec{e}'_j en la base $\{\vec{e}_i\}$ son los elementos de la llamada *matriz de transformación* $\boldsymbol{\gamma}$ ($n \times n$), que representa el *cambio de base* (o transformación de coordenadas)

$$\{\vec{e}_i\}_{i=1}^n \xrightarrow{\boldsymbol{\gamma}} \{\vec{e}'_j\}_{j=1}^n.$$

Si pensamos a los \vec{e}_i y los \vec{e}'_j como columnas de las matrices ($n \times n$) E y E' respectivamente, es decir poniendo

$$E := \begin{pmatrix} \vec{e}_1 & \vec{e}_2 & \dots & \vec{e}_n \\ \downarrow & \downarrow & & \downarrow \end{pmatrix}, \quad E' := \begin{pmatrix} \vec{e}'_1 & \vec{e}'_2 & \dots & \vec{e}'_n \\ \downarrow & \downarrow & & \downarrow \end{pmatrix},$$

podemos poner (1.24) en notación matricial como

$$E' = E\boldsymbol{\gamma}. \quad (1.24)$$

Notemos que, por ser bases de V , tanto $\{\vec{e}_i\}$ como $\{\vec{e}'_j\}$ son LI. Luego la matriz $\boldsymbol{\gamma}$ es *no singular*, y por lo tanto tiene inversa $\boldsymbol{\gamma}^{-1}$, la cual evidentemente representa el cambio de base inverso

$$E = E'\boldsymbol{\gamma}^{-1}.$$

1.4.2. Covarianza y contravarianza

Siguiendo la idea de que un dado vector \vec{x} tiene un significado físico propio, independiente de la base en que lo representemos, pediremos que su desarrollo en las bases $\{\vec{e}_i\}$ y $\{\vec{e}'_j\}$ representen el mismo vector, es decir

$$\vec{x} = x^i \vec{e}_i = x'^j \vec{e}'_j.$$

Substituyendo $\vec{e}'_j = \vec{e}_i \gamma_j^i$ e igualando componentes en la base $\{\vec{e}_i\}$, obtenemos $x^i = x'^j \gamma_j^i$, y aplicando la transformación inversa resulta $x'^j = [\boldsymbol{\gamma}^{-1}]_i^j x^i$. Ordenando las componentes de \vec{x} y \vec{x}' como matrices columna, podemos poner entonces en notación matricial

$$\mathbf{x}' = \boldsymbol{\gamma}^{-1} \mathbf{x}. \quad (1.25)$$

Comparando las ecuaciones (1.24) y (1.25) vemos que los vectores de una base y las componentes de un vector en esa base, transforman de maneras exactamente opuestas. Por convención, se llama *covariante* a todo objeto que, ante cambios de base, transforme como los vectores de la base. Todo objeto que transforme a la inversa, como las componentes de un vector, se llama *contravariante*.

La notación que hemos adoptado para los índices es consistente con esta distinción: los *subíndices* corresponden a objetos covariantes, y los *superíndices* a objetos contravariantes.

Ejercicio 1.4.1. Considere la transformación

$$\vec{e}'_1 = \vec{e}_1 + \vec{e}_2, \quad \vec{e}'_2 = \vec{e}_1 - \vec{e}_2$$

y el vector $2\vec{e}_1 + 3\vec{e}_2$. Calcule la matriz de transformación y las componentes del vector en la nueva base, y verifique que se cumple (1.25).

1.4.3. Componentes de un operador

Veremos ahora que las componentes de un operador lineal en un par de bases dadas también poseen propiedades de covarianza y contravarianza bien definidas.

Sea V un EV, $\{\vec{e}_i\}_{i=1}^n \xrightarrow{\boldsymbol{\gamma}} \{\vec{e}'_j\}_{j=1}^n$ bases de V , $\boldsymbol{\gamma}$ la matriz de transformación entre ellas, y $\vec{x} = x^i \vec{e}_i = x'^j \vec{e}'_j \in V$.

Sea W otro EV, $\{\vec{f}_l\}_{l=1}^m \xrightarrow{\boldsymbol{\delta}} \{\vec{f}'_k\}_{k=1}^m$ bases de W , $\boldsymbol{\delta}$ la matriz de transformación entre ellas, e $\vec{y} = y^l \vec{f}_l = y'^k \vec{f}'_k \in W$.

Sea $\mathcal{A} : V \rightarrow W$ un operador lineal tal que $\vec{y} = \mathcal{A}\vec{x}$, de modo que $y^l = A^l_i x^i$ y $y'^k = A'^k_j x'^j$.

Las propiedades de transformación de las componentes de \vec{x} e \vec{y} nos permiten deducir las de las componentes de \mathcal{A} , dado que

$$y'^k = [\boldsymbol{\delta}^{-1}]_l^k y^l = [\boldsymbol{\delta}^{-1}]_l^k A^l_i x^i = [\boldsymbol{\delta}^{-1}]_l^k A^l_i \gamma_j^i x'^j \equiv A'^k_j x'^j \quad \forall \vec{x}.$$

Luego, igualando componentes

$$A'^k_j = [\boldsymbol{\delta}^{-1}]_l^k A^l_i \gamma_j^i,$$

o en notación matricial,

$$\mathbf{A}' = \boldsymbol{\delta}^{-1} \mathbf{A} \boldsymbol{\gamma}. \quad (1.26)$$

Vemos entonces que A'^k_j es *contravariante* en el superíndice k y *covariante* en el subíndice j , consistentemente con la notación adoptada.

1.4.4. Transformaciones de semejanza

Un caso particular muy importante de transformación de coordenadas, es aquel donde $W = V$. En ese caso $\vec{f}_i \equiv \vec{e}_i$, $\vec{f}'_k \equiv \vec{e}'_k$ y $\delta = \gamma$; luego A es $n \times n$ y

$$A' = \gamma^{-1} A \gamma, \quad A = \gamma A' \gamma^{-1}. \quad (1.27)$$

Llamaremos *transformación de semejanza* a toda transformación lineal tal que $A' = S^{-1}AS$ para todo operador lineal representado por las matrices A o A' . Evidentemente, la matriz de transformación S es no singular.

La importancia de las transformaciones de semejanza reside en que muchas propiedades de las matrices y de las operaciones entre ellas son *invariantes* bajo estas transformaciones, y serán por tanto propiedades de los operadores representados por ellas.

Ejemplo 1.4.1. $(\lambda A)' = S^{-1}(\lambda A)S = \lambda S^{-1}AS = \lambda A'$.

Ejemplo 1.4.2. $(AB)' = S^{-1}(AB)S = S^{-1}ASS^{-1}BS = A'B'$.

Ejercicio 1.4.2. Considere la transformación de semejanza $A' = S^{-1}AS$. Muestre que

1. $\det(A') = \det(A)$.
2. $\text{Tr}(A') = \text{Tr}(A)$.
3. $f(A') = S^{-1}f(A)S \forall f$ desarrollable en serie de potencias.
4. A Hermitiana $\implies A'$ Hermitiana si S unitaria.
5. A unitaria $\implies A'$ unitaria si S unitaria.
6. A ortogonal $\implies A'$ ortogonal si S ortogonal.
7. $AB = BA \implies A'B' = B'A'$.
8. $B = A^{-1} \implies B' = (A')^{-1}$.

Ejercicio 1.4.3. Considere la transformación de coordenadas del Ejercicio (1.4.1) y un operador que rota un ángulo θ en sentido positivo, como el del Ejemplo (1.2.9) pero en dos dimensiones. Calcule las componentes del operador en la nueva base.

Ejercicio 1.4.4. Sean $\vec{v} \in \mathbb{R}^2$ y $\mathcal{A} : \mathbb{R}^2 \rightarrow \mathbb{R}^2$, representados en la base Cartesiana usual por

$$v = \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \quad A = \begin{pmatrix} 1 & \frac{1}{2} \\ \frac{1}{2} & 1 \end{pmatrix},$$

y sea una nueva base definida por

$$\vec{e}'_1 = \cos \theta \vec{e}_1 + \sin \theta \vec{e}_2, \quad \vec{e}'_2 = -\sin \theta \vec{e}_1 + \cos \theta \vec{e}_2.$$

Halle la representación de $\mathcal{A}\vec{v}$ en la vieja base, halle las representaciones v' de \vec{v} y A' de \mathcal{A} en la nueva base, y verifique que $(Av)' = A'v'$.

1.5. Formas y espacio dual

Vamos a introducir ahora la importante noción de espacio dual de un EV.

1.5.1. Formas

Sea V un EV con cuerpo escalar \mathbb{C} .

Una *forma* (o 1-forma) es una aplicación $\phi : V \rightarrow \mathbb{C}$, es decir una aplicación que lleva elementos del espacio vectorial al cuerpo escalar,

$$V \ni \vec{x} \xrightarrow{\phi} \phi(\vec{x}) \in \mathbb{C}. \quad (1.28)$$

Análogamente, una *n-forma* es una aplicación $\phi : V \oplus V \oplus \cdots \oplus V \rightarrow \mathbb{C}$, es decir una aplicación que lleva n -tuplas de elementos del espacio vectorial al cuerpo escalar,

$$\underbrace{V \oplus V \oplus \cdots \oplus V}_{n \text{ veces}} \ni (\vec{x}_1, \vec{x}_2, \dots, \vec{x}_n) \xrightarrow{\phi} \phi(\vec{x}_1, \vec{x}_2, \dots, \vec{x}_n) \in \mathbb{C}. \quad (1.29)$$

Llamaremos *forma lineal* a cualquier forma que verifique

$$\phi(\lambda \vec{a} + \mu \vec{b}) \equiv \lambda \phi(\vec{a}) + \mu \phi(\vec{b}). \quad (1.30)$$

En general y salvo nota explícita en contrario, asumiremos que trabajamos con formas lineales.

1.5.2. Espacio dual

Dado un EV V con cuerpo escalar \mathbb{C} , definimos el *espacio dual* V^* de V como el espacio de todas las formas lineales definidas sobre V , es decir

$$V^* := \{\phi : V \rightarrow \mathbb{C} / \phi \text{ lineal}\}. \quad (1.31)$$

Como veremos enseguida, el dual de un EV es también un EV, y tiene su misma dimensión. Por ello, en adelante adoptaremos la notación $\tilde{\phi}$ para sus elementos.

1.5.3. Componentes de una forma

Veremos que las formas lineales, igual que los vectores, pueden representarse adecuadamente dando sus componentes en una base.

Sea $\tilde{\phi} \in V^*$, $\{\vec{e}_i\}_{i=1}^n$ base de V , $\vec{x} = x^i \vec{e}_i \in V$. Por linealidad, $\tilde{\phi}(\vec{x}) = \tilde{\phi}(\vec{e}_i)x^i$. De acuerdo a la idea general de que la acción de una aplicación lineal sobre un EV queda completamente especificada si damos su acción sobre cada vector de una base, definiremos las *componentes* de la forma $\tilde{\phi}$ como

$$\phi_i := \tilde{\phi}(\vec{e}_i), \quad (1.32)$$

y ahora podemos poner

$$\tilde{\phi}(\vec{x}) = \phi_i x^i. \quad (1.33)$$

Sin embargo, nos gustaría completar la analogía con lo realizado para vectores de un EV respecto a su desarrollo en una base dada, definiendo un desarrollo análogo para una forma. Para ello necesitamos definir una base de V^* , lo que podemos conseguir definiendo las formas \tilde{e}^i tales que

$$\tilde{e}^i(\vec{e}_j) = \delta_j^i. \quad (1.34)$$

Ahora podemos poner

$$\tilde{\phi} = \phi_i \tilde{e}^i \quad (1.35)$$

y consistentemente tendremos

$$\tilde{\phi}(\vec{x}) = \phi_i \tilde{e}^i(x^j \vec{e}_j) = \phi_i \tilde{e}^i(\vec{e}_j) x^j = \phi_i \delta_j^i x^j = \phi_i x^i.$$

La base $\{\tilde{e}^i\}_{i=1}^n$ de V^* así construida se conoce como la *base dual* de $\{\vec{e}_i\}_{i=1}^n$.

Empleando la notación matricial, vemos que podemos ordenar las componentes ϕ_i de $\tilde{\phi}$ como una *matriz fila* $1 \times n$

$$\Phi = (\phi_1, \dots, \phi_n),$$

de modo que un uso consistente de la notación nos permite expresar $\tilde{\phi}(\vec{x})$ como el producto matricial

$$\tilde{\phi}(\vec{x}) = \Phi \mathbf{x} = (\phi_1, \dots, \phi_n) \begin{pmatrix} x^1 \\ \vdots \\ x^n \end{pmatrix}.$$

1.5.4. Transformaciones de coordenadas en V^*

Las formas, al igual que los vectores y operadores, tienen “existencia” independiente de la base en la cual las representemos, pero sus componentes cambiarán si cambiamos de una base a otra. Veremos ahora cómo una transformación de coordenadas en V induce una transformación relacionada en su dual V^* .

Sea un vector $\vec{x} = x^i \vec{e}_i = x'^j \vec{e}'_j \in V$, con $\vec{e}'_j = \vec{e}_i \gamma^i{}_j$, de modo que $x'^j = [\boldsymbol{\gamma}^{-1}]^j{}_i x^i$; y sea $\tilde{\phi}(\vec{x}) = \phi_i x^i = \phi'_j x'^j$, con $\tilde{\phi} \in V^*$. Por linealidad, $\tilde{\phi}(\vec{x}) = \tilde{\phi}(x'^j \vec{e}'_j) = \tilde{\phi}(\vec{e}'_j) x'^j = \phi'_j x'^j = \phi'_j [\boldsymbol{\gamma}^{-1}]^j{}_i x^i = \phi_i x^i$; como esta relación debe valer $\forall \vec{x}$, podemos deducir que $\phi'_j [\boldsymbol{\gamma}^{-1}]^j{}_i = \phi_i$, es decir

$$\phi'_j = \phi_i \gamma^i{}_j, \quad (1.36)$$

o en notación matricial

$$\Phi' = \Phi \boldsymbol{\gamma}. \quad (1.37)$$

La ley de transformación para la base $\{\tilde{e}^i\}_{i=1}^n$ de V^* puede deducirse ahora fácilmente del hecho de que

$$\tilde{\phi} = \phi_i \tilde{e}^i = \phi'_j \tilde{e}'^j,$$

y resulta ser

$$\tilde{e}'^j = [\gamma^{-1}]^j_i \tilde{e}^i. \quad (1.38)$$

Vemos entonces que las componentes de una forma son *covariantes* y la base de V^* es *contravariante*, exactamente al revés que en V .

Ejercicio 1.5.1. Considere la transformación de coordenadas del Ejercicio (1.4.1). Escriba la base dual transformada usando $\tilde{e}'^i(\tilde{e}'_j) = \delta^i_j$ y verifique (1.38).

1.6. Producto interno, métrica y norma

Probablemente nos hayamos habituado a la definición geométrica de la norma de un vector como su “largo”, o del producto interno entre dos vectores como “el producto de las normas por el coseno del ángulo comprendido”, y en realidad así fue como estos conceptos surgieron históricamente. Sin embargo, estas definiciones son difícilmente extensibles a espacios menos “geométricos” como $\{p_N(x)\}$ o $C^\infty[\mathbb{R}]$. En esta sección veremos cómo introducir una noción mucho más general (aunque más abstracta) de producto interno y norma, que sin embargo se reducirá a la ya conocida para espacios “geométricos”. Como bonificación, obtendremos una forma de definir un mapa biyectivo entre V y V^* , es decir que podremos poner en correspondencia cada vector con una forma y viceversa.

1.6.1. Producto interno

El *producto interno* o *producto escalar* entre dos vectores \vec{a} y \vec{b} de un mismo espacio vectorial V se define como una 2-forma $\Phi : V \oplus V \rightarrow \mathbb{C}$, es decir

$$V \oplus V \ni (\vec{a}, \vec{b}) \xmapsto{\Phi} \Phi(\vec{a}, \vec{b}) \in \mathbb{C} \quad (1.39)$$

que satisface

1. $\Phi(\vec{a}, \vec{b}) = \Phi^*(\vec{b}, \vec{a});$
2. $\Phi(\vec{a}, \lambda \vec{b}) = \lambda \Phi(\vec{a}, \vec{b}); \Phi(\lambda \vec{a}, \vec{b}) = \lambda^* \Phi(\vec{a}, \vec{b});$
3. $\Phi(\vec{a}, \vec{a}) \geq 0$ (y obviamente debe ser real); $\Phi(\vec{a}, \vec{a}) = 0 \iff \vec{a} = \vec{0}.$

Cualquier 2-forma que cumpla estas condiciones será un “buen” producto interno sobre V .

Ejemplo 1.6.1. Dado el EV de las funciones a valores complejos definidas sobre el intervalo real $[a, b]$, $V = \{f : [a, b] \rightarrow \mathbb{C}\}$, y dadas $f, g \in V$, la 2-forma

$$\langle f, g \rangle = \int_a^b f^*(x)g(x) dx$$

define un producto interno sobre V .

Ejercicio 1.6.1. Demuestre la afirmación precedente.

1.6.2. Métrica

Consideremos ahora una base $\{\vec{e}_i\}_{i=1}^n$ de V y dos vectores $\vec{x} = x^i \vec{e}_i$ e $\vec{y} = y^i \vec{e}_i$. Por las propiedades de linealidad del producto interno, tendremos $\Phi(\vec{x}, \vec{y}) = x^{*i} \Phi(\vec{e}_i, \vec{e}_j) y^j$. De acuerdo a la idea de que el efecto de una aplicación queda completamente especificado dando su efecto sobre los vectores de una base, definimos

$$g_{ij} := \Phi(\vec{e}_i, \vec{e}_j). \quad (1.40)$$

Los elementos g_{ij} se conocen como *elementos de la métrica*, y podemos ordenarlos en una matriz

$$\mathbf{g} := [g_{ij}] \quad (1.41)$$

que se conoce como la *métrica*. Las propiedades del producto interno imponen que esta matriz $n \times n$ sea Hermitiana (simétrica si es real) y definida positiva.¹⁰ Dada una métrica podemos calcular el producto interno de dos vectores cualesquiera a partir simplemente de sus componentes, de acuerdo a

$$\vec{x} \cdot \vec{y} := \Phi(\vec{x}, \vec{y}) = x^{*i} g_{ij} y^j, \quad (1.42)$$

o en notación matricial

$$\vec{x} \cdot \vec{y} := \mathbf{x}^\dagger \mathbf{g} \mathbf{y}. \quad (1.43)$$

Ejercicio 1.6.2. Considere la transformación de coordenadas del Ejercicio (1.4.1). Asumiendo que $\vec{e}_i \cdot \vec{e}_j = \delta_{ij}$, calcule la métrica en la base $\{\vec{e}'_i\}$.

Ejercicio 1.6.3. Calcule la métrica en la base $\{\vec{e}'_i\}$ dada por

$$\vec{e}'_1 = \vec{e}_1, \quad \vec{e}'_2 = \vec{e}_1 + \vec{e}_2, \quad \vec{e}'_3 = \vec{e}_1 + \vec{e}_2 + \vec{e}_3.$$

1.6.3. Norma

Una vez definido un producto interno (es decir, dada una métrica) podemos usarlo para definir la *norma* de un vector como

$$\|\vec{x}\| := \sqrt{\vec{x} \cdot \vec{x}} = \sqrt{x^{*i} g_{ij} x^j}. \quad (1.44)$$

Las propiedades del producto interno garantizan que esta norma estará bien definida, aún para vectores complejos, y que sólo será nula para el vector nulo.

Es más, la norma así definida es una función *monótonamente creciente* de las normas (complejas) de cada componente de un vector. Esto permite asociarla a nuestra noción “intuitiva” de la norma de un vector como una medida de su “distancia al origen”, y permite también definir la *distancia* entre dos vectores como

$$\text{dist}(\vec{x}, \vec{y}) := \|\vec{x} - \vec{y}\|. \quad (1.45)$$

¹⁰En las teorías relativistas se utilizan métricas que no son definidas positivas, por ejemplo la métrica de Minkowski de la Relatividad Especial, que es $g_{ij} = \text{diag}(-1, 1, 1, 1)$.

Ejercicio 1.6.4. Calcule la norma de $\vec{e}_1 + 2\vec{e}_2 + 3\vec{e}_3$ en ambas bases del Ejercicio 1.6.3. ¿Coinciden? ¿Porqué?

Una vez definida la norma de un vector, podemos utilizarla para inducir una norma para operadores; la definición mas usual es

$$\|\mathcal{A}\| := \inf \{c \geq 0 : \|\mathcal{A}\vec{v}\| \leq c\|\vec{v}\| \forall \vec{v} \in V\}, \quad (1.46)$$

basada en la idea de que $\|\mathcal{A}\|$ representa el mayor valor real en el que la aplicación de \mathcal{A} puede incrementar la norma de un vector \vec{v} .¹¹ Un operador cuya norma es finita se dice *acotado*, ya que mapea subconjuntos acotados en subconjuntos acotados. Ahora podemos precisar un criterio de convergencia para funciones de operadores escritas como series de sus potencias: si la serie $f(x) = \sum_{n=0}^{\infty} a_n x^n$ tiene radio de convergencia r , la serie $f(\mathcal{A}) = \sum_{n=0}^{\infty} a_n \mathcal{A}^n$ convergerá donde $\|\mathcal{A}\| < r$.

Ejemplo 1.6.2. La serie exponencial $e^x = \sum_{n=0}^{\infty} x^n / n!$ tiene radio de convergencia infinito; luego si $V = \{f : [a, b] \rightarrow \mathbb{R} : f \in C^\infty[a, b], |f^{(n)}| \leq c < \infty \forall n \in \mathbb{N}_0\}$, la serie de $e^{\alpha \partial_x}$ convergerá para todo α finito.

1.6.4. De V a V^* y vuelta con el producto interno

Consideremos un espacio vectorial V y su dual V^* , respectivamente con bases $\{\vec{e}_i\}_{i=1}^n$ y $\{\vec{e}^i\}_{i=1}^n$, y un dado producto interno Φ que podemos especificar dando la métrica g . Recordemos que Φ es una 2-forma, es decir una aplicación que toma dos vectores de V y devuelve un escalar. Si tomamos un vector dado $\vec{x} \in V$, podemos pensar en la aplicación $\Phi(\vec{x}, \cdot)$ que obtenemos de Φ “llenando” su “primera casilla” con \vec{x} , y dejando la segunda “libre” para recibir otro vector cualquiera y y devolver un escalar. Entonces vemos que $\Phi(\vec{x}, \cdot)$ es una aplicación (lineal) de V a los escalares, y por lo tanto es una forma, es decir un elemento de V^* , que denotaremos por

$$\tilde{\phi}_{\vec{x}} := \Phi(\vec{x}, \cdot). \quad (1.47)$$

Vemos así que Φ mapea cada elemento \vec{x} de V en un elemento $\tilde{\phi}_{\vec{x}}$ de V^* .

Esto nos permite replantear el concepto de producto interno como un “producto” entre un elemento de V y uno de V^* : para tomar el producto interno $\vec{x} \cdot \vec{y}$, primero obtenemos la forma $\tilde{\phi}_{\vec{x}}$ y luego se la aplicamos a \vec{y} .

Las componentes de $\tilde{\phi}_{\vec{x}}$ pueden deducirse del hecho que $\vec{x} \cdot \vec{y} = x^{*i} g_{ij} y^j$ y que para cualquier forma $\tilde{\phi}$, $\tilde{\phi}(\vec{y}) = \phi_j y^j$, de lo que obtenemos

$$\phi_{\vec{x}j} := x^{*i} g_{ij}. \quad (1.48)$$

La métrica, entonces, lleva las componentes de un vector a las de la correspondiente forma.

¹¹Una definición alternativa será discutida en el contexto de espacios de Hilbert.

Siguiendo con la misma idea, podemos ahora deducir cuál es la base $\{\tilde{e}^j\}_{j=1}^n$ de V^* en la cual $\tilde{\phi}_{\vec{x}}$ tiene las componentes dadas, ya que por un lado debemos tener

$$\tilde{\phi}_{\vec{x}} = \phi_{\vec{x}j} \tilde{e}^j, \quad (1.49)$$

de modo que

$$\tilde{\phi}_{\vec{x}}(\vec{y}) = \phi_{\vec{x}k} \tilde{e}^k(\vec{e}_j) y^j = x^{*i} g_{ik} \tilde{e}^k(\vec{e}_j) y^j,$$

y por otro lado

$$\vec{x} \cdot \vec{y} = x^{*i} g_{ij} y^j.$$

Comparando estas expresiones, que deben valer $\forall y^j$, resulta

$$x^{*i} g_{ik} \tilde{e}^k(\vec{e}_j) = x^{*i} g_{ij},$$

que a su vez debe valer $\forall x^i$, de donde concluimos que

$$\tilde{e}^k(\vec{e}_j) = \delta_j^k,$$

es decir, $\{\tilde{e}^i\}_{i=1}^n$ es la *base dual* (o *adjunta*) de $\{\vec{e}_i\}_{i=1}^n$.

La expresión $\vec{x} \cdot \vec{y} := \vec{x}^\dagger \vec{y}$ para el producto interno de dos vectores nos provee una manera sencilla de completar estas correspondencias, interpretando el producto $\vec{x}^\dagger \vec{y}$ de dos formas alternativas:

- Teníamos las componentes de dos vectores como matrices columna x e y ; hallamos las componentes de la forma correspondiente a x como la matriz fila $x^\dagger g$, y multiplicamos por y para obtener el escalar $x^\dagger g y$.
- Teníamos las componentes de dos formas como matrices fila x^\dagger e y^\dagger ; hallamos las componentes del vector correspondiente a y^\dagger como la matriz columna $g y$, y multiplicamos por x^\dagger para obtener el escalar $x^\dagger g y$.

Notemos que con la última interpretación, ¡acabamos de definir el producto interno entre dos formas! El hecho de que la métrica, siendo definida positiva, sea no singular, y por lo tanto invertible, es lo que hace que la correspondencia entre vectores y formas así definida sea una biyección, de modo que podemos poner

$$V \ni \vec{x} \xleftrightarrow{g} \tilde{\phi}_{\vec{x}} \in V^*. \quad (1.50)$$

1.6.5. Correspondencia geométrica

Inicialmente, los vectores fueron introducidos como artificios para describir la posición de un punto en el espacio tridimensional. Tomando un punto arbitrario como origen, hacemos pasar por él tres rectas mutuamente ortogonales (en sentido geométrico), los *ejes Cartesianos*. El vector que da la posición de un punto será el segmento orientado del origen a ese punto, y sus *componentes Cartesianas* serán sus proyecciones geométricas sobre cada eje,

$$x^i := \|\vec{x}\| \cos \theta_i,$$

donde θ_i son los ángulos entre el vector y cada eje, y la norma $\|\vec{x}\|$ es su largo.

Si colocamos sobre cada eje vectores de largo 1, los *versores Cartesianos* \vec{e}_i , y definimos el producto escalar entre dos vectores \vec{x} e \vec{y} por

$$\vec{x} \cdot \vec{y} := \|\vec{x}\| \|\vec{y}\| \cos \theta,$$

donde θ es el ángulo comprendido entre ellos, las componentes Cartesianas de \vec{x} pueden escribirse $x^i = \vec{e}_i \cdot \vec{x}$, y la trigonometría nos dice que $\|\vec{x}\| = \sqrt{(x^1)^2 + (x^2)^2 + (x^3)^2} = \sqrt{\vec{x} \cdot \vec{x}}$. En particular, el producto escalar entre dos vectores geométricamente ortogonales es nulo, de modo que

$$\vec{e}_i \cdot \vec{e}_j = \delta_{ij}.$$

Comparando estas expresiones para la norma y el producto escalar (y en particular el producto escalar entre los versores) con las obtenidas previamente, vemos que resultan consistentes si definimos la *métrica Cartesiana*

$$g_{ij} := \delta_{ij}$$

Por definición (herencia de su origen geométrico), un sistema de coordenadas con esta métrica se llama *ortonormal* (los versores son *ortogonales* y de *norma unidad*). Es de destacar que, siendo la métrica la matriz identidad $\mathbf{g} = \mathbf{I}$, en coordenadas ortonormales la distinción entre objetos covariantes y contravariantes desaparece, y el mapa de V a V^* es trivial.

Capítulo 2

Matrices

2.1. Autovalores y autovectores

Consideremos el siguiente problema: dado un operador lineal $\mathcal{A} : V \rightarrow V$, ¿cuándo existen vectores en V que sean *invariantes* bajo \mathcal{A} ? Es decir, ¿bajo qué condiciones

$$\exists \vec{x} \in V / \mathcal{A}\vec{x} = \lambda\vec{x} \quad (2.1)$$

para algún $\lambda \in \mathbb{C}$?

La importancia de este problema reside en que cada vector \vec{x} que satisface (2.1) genera un subespacio de V *invariante* bajo \mathcal{A} . Si podemos dividir el espacio vectorial en una suma directa de subespacios invariantes, el estudio de las propiedades del operador se simplificará notablemente, ya que podremos limitarnos a estudiar su acción sobre los vectores de cada subespacio por separado.

Si $\mathcal{A}\vec{x} = \lambda\vec{x}$ con $\vec{x} \neq \vec{0}$, λ se llama un *autovalor* del operador \mathcal{A} , y \vec{x} se llama un *autovector* asociado (o perteneciente) a ese autovalor.

Ejemplo 2.1.1. Sea $\mathcal{R} : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ el operador que rota en un ángulo θ alrededor del eje z ; entonces cualquier vector $\vec{a} = a\hat{z}$ es un autovector de \mathcal{R} con autovalor 1: $\mathcal{R}\vec{a} = \vec{a}$.

Ejemplo 2.1.2. Sea $\partial_x : \mathcal{C}^\infty[\mathbb{R}] \rightarrow \mathcal{C}^\infty[\mathbb{R}]$; entonces e^{ax} son autovectores (llamados en este caso *autofunciones*) con autovalor a : $\partial_x e^{ax} = ae^{ax}$.

Notemos que si \vec{x} es un autovector del operador \mathcal{A} con autovalor λ , cualquier múltiplo escalar de \vec{x} también lo es. Es más, si $\vec{x}_1, \dots, \vec{x}_r$ son todos autovectores de \mathcal{A} con autovalor λ , cualquier combinación lineal de ellos también lo es.

Ejercicio 2.1.1. Demuestre las afirmaciones precedentes.

2.1.1. Autovectores a derecha

Sea $\mathcal{A} : V \rightarrow V$, y sea A su representación matricial en una base dada $\{\vec{e}_i\}$. En notación matricial, el problema (2.1) puede escribirse

$$Av = \lambda v \implies (A - \lambda I)v = 0. \quad (2.2)$$

Esta última ecuación matricial es un sistema homogéneo de ecuaciones lineales para las componentes del vector \vec{v} , y tendrá solución no trivial $\vec{v} \neq \vec{0}$ si y solo si la matriz $(A - \lambda I)$ es no invertible, o sea singular. Debemos exigir entonces que

$$\det(A - \lambda I) = 0. \quad (2.3)$$

Definimos el *polinomio característico* del operador \mathcal{A} como

$$\varphi_A(\lambda) := \det(\lambda I - A) = \sum_{j=0}^n \alpha_j \lambda^{n-j}, \quad (2.4)$$

donde $n = \dim(V)$. Este es evidentemente un polinomio en λ de grado n igual a la dimensión del espacio, que es la dimensión de la matriz A . Sus coeficientes α_j se llaman *invariantes algebraicos* (o *invariantes escalares*) del operador \mathcal{A} , por motivos que pronto se harán evidentes.

De (2.3) y (2.4) vemos que si λ es un autovalor, el polinomio característico satisface la llamada *ecuación secular*

$$\varphi_A(\lambda) = 0.$$

El Teorema Fundamental del Álgebra nos garantiza que *siempre* existen n raíces complejas $\lambda_1, \dots, \lambda_n \in \mathbb{C}$ de esta ecuación, de modo que el polinomio característico puede factorizarse, como cualquier otro polinomio, en la forma

$$\varphi_A(\lambda) = (\lambda - \lambda_1)(\lambda - \lambda_2) \dots (\lambda - \lambda_n), \quad \lambda_1, \dots, \lambda_n \in \mathbb{C}.$$

Estas raíces no son necesariamente todas distintas (pese a nuestra notación), y puede que ni siquiera estén en el cuerpo escalar asociado a nuestro espacio vectorial.

Ejemplo 2.1.3. Consideremos el operador $R_z : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ que efectúa una rotación de ángulo θ en sentido positivo alrededor del eje z . En la base Cartesiana usual, su representación matricial es

$$R_z = \begin{pmatrix} \cos \theta & -\sin \theta & 0 \\ \sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Por lo tanto

$$\varphi_{R_z}(\lambda) = (\lambda^2 - 2 \cos(\theta)\lambda + 1)(\lambda - 1)$$

y los autovalores son $e^{\pm i\theta}$ y 1; el cuerpo escalar asociado a \mathbb{R}^3 es \mathbb{R} , pero ¡dos de los autovalores son complejos!

Para λ igual a cada una de las raíces λ_i de la ecuación secular, la ecuación (2.2) tendrá (al menos) una solución no trivial $v_i \neq 0$, es decir

$$Av_i = \lambda_i v_i. \quad (2.5)$$

Por lo tanto, las raíces $\lambda_1, \dots, \lambda_n$ de la ecuación secular serán los autovalores del operador \mathcal{A} , y las correspondientes soluciones no triviales v_i de (2.5) serán las representaciones matriciales (como matrices columna), en la base $\{\vec{e}_i\}$, de los autovectores \vec{v}_i de \mathcal{A} asociados al correspondiente autovalor λ_i . El conjunto de autovalores de un operador se llama su *espectro*.

Vamos ahora a enunciar un primer resultado general, para el caso relativamente frecuente en que los autovalores son todos distintos.

Teorema 2.1.1. *Si $\varphi_{\mathcal{A}}(\lambda)$ tiene n raíces λ_i todas distintas, con $n = \dim(V)$, entonces los autovectores asociados \vec{v}_i forman una base¹ del espacio vectorial V .*

Demostración. Demostraremos la independencia lineal de $\{\vec{v}_i\}_{i=1}^n$ por inducción. Por claridad suspenderemos momentáneamente el convenio de suma sobre índices repetidos.

- $\{\vec{v}_1\}$ LI es trivial: $a^1 \vec{v}_1 = \vec{0} \implies a^1 = 0$, ya que suponemos $\vec{v}_1 \neq \vec{0}$.
- Supongamos $\{\vec{v}_i\}_{i=1}^{k-1}$ LI (para un dado $k \leq n$), y sean $a^i \in \mathbb{C}$ tales que $\sum_{i=1}^k a^i \vec{v}_i = \vec{0}$. Luego

$$\mathcal{A} \left(\sum_{i=1}^k a^i \vec{v}_i \right) = \sum_{i=1}^k a^i \mathcal{A}(\vec{v}_i) = \sum_{i=1}^k a^i \lambda_i \vec{v}_i = \vec{0}.$$

Por otro lado

$$\sum_{i=1}^k a^i \vec{v}_i = \vec{0} \implies \lambda_k \sum_{i=1}^k a^i \vec{v}_i = \vec{0}.$$

Restando miembro a miembro,

$$\sum_{i=1}^k a^i \lambda_i \vec{v}_i - \lambda_k \sum_{i=1}^k a^i \vec{v}_i = \sum_{i=1}^{k-1} a^i (\lambda_i - \lambda_k) \vec{v}_i = \vec{0}.$$

Pero $\{\vec{v}_i\}_{i=1}^{k-1}$ LI $\implies a^i (\lambda_i - \lambda_k) = 0$, $i = 1, \dots, k-1$, y por lo tanto $\lambda_i \neq \lambda_k$, $i = 1, \dots, k-1 \implies a_i = 0$, $i = 1, \dots, k-1$. Luego

$$\sum_{i=1}^k a^i \vec{v}_i = a^k \vec{v}_k = \vec{0} \implies a^k = 0, \text{ ya que suponemos } \vec{v}_k \neq \vec{0}.$$

Por lo tanto hemos probado que $a^i = 0$, $i = 1, \dots, k$, es decir

$$\{\vec{v}_i\}_{i=1}^{k-1} \text{ LI} \implies \{\vec{v}_i\}_{i=1}^k \text{ LI}.$$

¹No necesariamente ortonormal.

□

Notemos que la validez de este Teorema *no* está restringida a espacios de dimensión finita.²

Ejemplo 2.1.4. Los autovalores de la matriz $A = \begin{pmatrix} 1 & 1 \\ 0 & 2 \end{pmatrix}$ son $\lambda_1 = 1$ y $\lambda_2 = 2$, y los correspondientes autovectores son $v_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ y $v_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$; no son ortogonales ni de norma 1, pero forman base de \mathbb{R}^2 .

Ejercicio 2.1.2. Verifique las afirmaciones precedentes.

Si bien podemos definir una *norma inducida* (o subordinada) para matrices a partir de la norma (1.46) para operadores, por ejemplo como

$$\|A\| := \sup \{\|Ax\| : x \in \mathbb{C}^n, \|x\| \leq 1\} \quad (2.6)$$

con $\|x\| = \|\vec{x}\|$ o cualquier otra norma adecuada en \mathbb{C}^n , una vez introducidos los autovalores de una matriz podemos definir (con base en la norma Euclidiana $\|x\|_2$ en \mathbb{C}^n) la *norma espectral* de A como

$$\|A\|_2 := \sqrt{\lambda_{\max}(A^\dagger A)}, \quad (2.7)$$

con $\lambda_{\max}(A^\dagger A)$ el máximo autovalor de $A^\dagger A$, que es de uso mucho más frecuente. La razón para definirla así, y no en base a los autovalores de A (o A^\dagger) es que una matriz como $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ tiene sólo autovalores nulos, y sin embargo $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \neq \begin{pmatrix} 0 \\ 0 \end{pmatrix}$; en cambio $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}^\dagger \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$, con máximo autovalor 1. Ya definida la norma de una matriz, podemos precisar un criterio de convergencia para funciones de matrices escritas como series de sus potencias: si la serie $f(x) = \sum_{n=0}^{\infty} a_n x^n$ tiene radio de convergencia r , la serie $f(A) = \sum_{n=0}^{\infty} a_n A^n$ convergerá donde $\|A\| < r$.

Ejemplo 2.1.5. La serie exponencial $e^A = \sum_{n=0}^{\infty} \frac{1}{n!} A^n$ convergerá para toda A de norma finita. La serie geométrica $\frac{1}{1-A} = \sum_{n=0}^{\infty} A^n$ convergerá si $\|A\| < 1$.

Ejemplo 2.1.6. Para $A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ la serie $\sum_{n=0}^{\infty} A^n$ converge a $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} = (I - A)^{-1}$ pese a que $\|A\|_2 = 1$, lo que ocurre porque la serie se *trunca*, ya que $A^n = 0 \forall n > 2$; esto *no puede* ocurrir con series de escalares, donde si $x \neq 0$, entonces $x^n \neq 0 \forall n \in \mathbb{N}$, y muestra que con matrices, la condición de convergencia basada en norma es suficiente, pero no necesaria.

Ejercicio 2.1.3. Muestre que si las series involucradas convergen, $Av = \lambda v \implies f(A)v = f(\lambda)v$.

²Para espacios de dimensión infinita nuestra demostración sólo prueba la independencia lineal de los autovectores; el concepto de base requiere restricciones adicionales sobre los vectores “admisibles” como parte del espacio.

2.1.2. Autovectores a izquierda

Consideremos un operador $\mathcal{A} : V \rightarrow V$ tal que $\mathcal{A}\vec{v}_i = \lambda_i \vec{v}_i$, con n autovalores λ_i [$n = \dim(V)$]. Podemos expresar esta relación matricialmente como

$$\mathbf{A}\mathbf{v}_i = \lambda_i \mathbf{v}_i, \quad i = 1, \dots, n. \quad (2.8)$$

Busquemos ahora autovalores y autovectores del operador adjunto \mathcal{A}^\dagger , cuya representación matricial es la matriz adjunta \mathbf{A}^\dagger ; es decir busquemos $\mu \in \mathbb{C}$ y $\vec{u} \in V$ tales que $\mathcal{A}^\dagger \vec{u} = \mu \vec{u}$, o matricialmente

$$\mathbf{A}^\dagger \mathbf{u} = \mu \mathbf{u}, \quad \mathbf{u} \neq 0. \quad (2.9)$$

Como antes, esta ecuación puede ponerse como $(\mathbf{A}^\dagger - \mu \mathbf{I})\mathbf{u} = 0$, y la existencia de solución no trivial requiere $\det(\mu \mathbf{I} - \mathbf{A}^\dagger) = 0$. Pero

$$\det(\mu \mathbf{I} - \mathbf{A}^\dagger) = \det[(\mu^* \mathbf{I} - \mathbf{A})^\dagger] = [\det(\mu^* \mathbf{I} - \mathbf{A})]^* = \left(\sum_{j=0}^n \alpha_j \mu^{*n-j} \right)^* = (\varphi_{\mathbf{A}}(\mu^*))^*,$$

de modo que

$$\det(\mu \mathbf{I} - \mathbf{A}^\dagger) = 0 \implies \varphi_{\mathbf{A}}(\mu^*) = 0.$$

Luego μ^* debe ser igual a alguna de las raíces λ_i de $\varphi_{\mathbf{A}}(\lambda)$, es decir $\mu = \lambda_i^*$. De lo anterior podemos concluir que los autovalores μ_i de \mathbf{A}^\dagger son los conjugados de los autovalores de \mathbf{A} ,

$$\mu_i \equiv \lambda_i^*.$$

Las matrices columna \mathbf{u}_i que representan a los correspondientes autovectores \vec{u}_i son soluciones de

$$\mathbf{A}^\dagger \mathbf{u}_i = \lambda_i^* \mathbf{u}_i \quad (2.10)$$

Si los autovalores λ_i son todos distintos, los μ_i también resultan todos diferentes entre sí, y puede demostrarse que los autovectores a izquierda \vec{u}_i forman base de V , de la misma forma que se hizo para los autovectores a derecha \vec{v}_i .

El nombre de *autovectores a izquierda* se debe a que las componentes de \vec{u}_i , en notación matricial, satisfacen la ecuación

$$\mathbf{u}^{i\dagger} \mathbf{A} = \lambda_i \mathbf{u}^{i\dagger}, \quad (2.11)$$

obtenida de la anterior tomando el adjunto Hermitiano de ambos miembros, y poniendo

$$\mathbf{u}^{i\dagger} := (\mathbf{u}_i)^\dagger.$$

Notemos que esto nos permite interpretar al operador \mathcal{A} , representado por la matriz \mathbf{A} , como actuando “hacia adelante” sobre elementos de V representados como matrices columna, o actuando “hacia atrás” sobre elementos de V^* representados como matrices fila.

Ejemplo 2.1.7. Los autovectores a izquierda de la matriz del Ejemplo 2.1.4 son $\mathbf{u}_1 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$ y $\mathbf{u}_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$; tampoco son ortogonales entre sí ni tienen norma 1.

Ejercicio 2.1.4. Verifique las afirmaciones precedentes.

2.1.3. Diagonalización de un operador

Sea un operador $\mathcal{A} : V \rightarrow V$, y sea \mathbf{A} su representación matricial en una dada base $\{\vec{e}_i\}_{i=1}^n$ de V ; sean $\mathbf{A}\mathbf{v}_i = \lambda_i \underline{\mathbf{v}}_i$, $\mathbf{A}^\dagger \underline{\mathbf{u}}_i = \lambda_i^* \underline{\mathbf{u}}_i$, con n autovalores λ_i todos distintos [$n = \dim(V)$]. Entonces

$$\begin{aligned} \mathbf{A}\mathbf{v}_i &= \lambda_i \underline{\mathbf{v}}_i &\implies \underline{\mathbf{u}}^{j\dagger} \mathbf{A}\mathbf{v}_i &= \lambda_i \underline{\mathbf{u}}^{j\dagger} \underline{\mathbf{v}}_i \\ \underline{\mathbf{u}}^{j\dagger} \mathbf{A} &= \lambda_j \underline{\mathbf{u}}^{j\dagger} &\implies \underline{\mathbf{u}}^{j\dagger} \mathbf{A}\mathbf{v}_i &= \lambda_j \underline{\mathbf{u}}^{j\dagger} \underline{\mathbf{v}}_i \end{aligned}$$

Como $\lambda_i \neq \lambda_j$ si $i \neq j$, entonces comparando las dos expresiones de la derecha vemos que $\underline{\mathbf{u}}^{j\dagger} \underline{\mathbf{v}}_i = 0$ si $i \neq j$. Por otro lado, podemos multiplicar tanto a $\underline{\mathbf{u}}^{j\dagger}$ como a $\underline{\mathbf{v}}_i$ por escalares arbitrarios sin alterar nada de lo que antecede, es decir, siempre podemos *normalizarlos* de modo que

$$\lambda_i \neq \lambda_j \text{ si } i \neq j \implies \underline{\mathbf{u}}^{j\dagger} \underline{\mathbf{v}}_i = \delta_i^j. \quad (2.12)$$

Ahora bien, la matriz fila $\underline{\mathbf{u}}^{j\dagger}$ puede pensarse como la representación matricial de un elemento \tilde{u}^j de V^* , igual que la matriz columna $\underline{\mathbf{v}}_i$ es la representación matricial del elemento \tilde{v}_i de V . En estos términos, la relación anterior se escribe

$$\lambda_i \neq \lambda_j \text{ si } i \neq j \implies \tilde{u}^j(\tilde{v}_i) = \delta_i^j. \quad (2.13)$$

Esto quiere decir que las bases $\{\tilde{v}_i\}$ de V y $\{\tilde{u}^j\}$ de V^* son *bases adjuntas*. Las correspondientes matrices columna $\underline{\mathbf{v}}_i$ y fila $\underline{\mathbf{u}}^{j\dagger}$ dan sus componentes en la base $\{\vec{e}_i\}$ de V .

Si ahora queremos interpretar a $\underline{\mathbf{u}}^{j\dagger} \underline{\mathbf{v}}_i$ como el producto escalar de los autovectores \tilde{u}_j y \tilde{v}_i , debemos pensar que primero obtenemos de \tilde{u}_j la matriz columna $\underline{\mathbf{u}}_j$ de sus componentes en la base $\{\vec{e}_i\}$, luego tomamos su adjunta Hermitiana $\underline{\mathbf{u}}^{j\dagger} = (\underline{\mathbf{u}}_j)^\dagger$, y finalmente decimos que ésta es la matriz fila que representa al correspondiente elemento \tilde{u}^j de V^* . La consistencia con lo expuesto en la Sección 1.6 nos muestra entonces que debemos considerar a $\{\vec{e}_i\}$ una base ortonormal de V ; es decir, debemos asumir que la métrica, en la base $\{\vec{e}_i\}$, es la identidad.³ Con esta consideración, podremos entonces poner

$$\tilde{u}_j \cdot \tilde{v}_i = \delta_{ji} \quad (2.14)$$

y decir que los autovectores a izquierda y a derecha son dos a dos ortogonales. Nótese que no pedimos que $\tilde{v}_j \cdot \tilde{v}_i = \delta_{ji}$ ni que $\tilde{u}_j \cdot \tilde{u}_i = \delta_{ji}$, y en general *no ocurre así*.

Ejemplo 2.1.8. Los autovectores a derecha del Ejemplo 2.1.4 y a izquierda del Ejemplo 2.1.7 satisfacen $\underline{\mathbf{u}}^{j\dagger} \underline{\mathbf{v}}_i = \delta_i^j$.

Ejercicio 2.1.5. Verifique la afirmación precedente.

De lo anterior se desprende que $\underline{\mathbf{u}}^{j\dagger} \mathbf{A}\mathbf{v}_i = \underline{\mathbf{u}}^{j\dagger} \underline{\mathbf{v}}_i \lambda_i = \delta_i^j \lambda_i$ son los elementos de la matriz diagonal $\text{diag}(\lambda_1, \dots, \lambda_n)$. Sea \mathbf{U} la matriz $n \times n$ que tiene a las

³Por cierto, la métrica en la base de los $\{\vec{e}_i\}$ en general *no* será la identidad.

matrices fila $u^{j\dagger}$ por filas, y sea V la matriz $n \times n$ que tiene a las matrices columna v_i por columnas, es decir

$$U := \begin{pmatrix} u^{1\dagger} & \longrightarrow \\ \vdots & \\ u^{n\dagger} & \longrightarrow \end{pmatrix}, \quad V := \begin{pmatrix} v_1 & \cdots & v_n \\ \downarrow & & \downarrow \end{pmatrix}.$$

Luego

$$UV = \begin{pmatrix} u^{1\dagger}v_1 & \dots & u^{1\dagger}v_n \\ \vdots & & \vdots \\ u^{n\dagger}v_1 & \dots & u^{n\dagger}v_n \end{pmatrix} = I,$$

de modo que $U = V^{-1}$, y

$$UAV = V^{-1}AV = \begin{pmatrix} u^{1\dagger}Av_1 & \dots & u^{1\dagger}Av_n \\ \vdots & & \vdots \\ u^{n\dagger}Av_1 & \dots & u^{n\dagger}Av_n \end{pmatrix} = \text{diag}(\lambda_1, \dots, \lambda_n).$$

Ejemplo 2.1.9. Para la matriz de los Ejemplos 2.1.4 y 2.1.7, tenemos $U = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}$ y $V = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$, con $UV = I$ y $UAV = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}$.

Ejercicio 2.1.6. Verifique las afirmaciones precedentes.

Podemos resumir los resultados obtenidos hasta ahora en el siguiente

Teorema 2.1.2. Sea $\mathcal{A} : V \rightarrow V$ un operador lineal sobre un espacio vectorial V de dimensión n , que tiene n autovalores diferentes $\lambda_1, \dots, \lambda_n$,⁴ sean \vec{v}_i y \vec{u}_j los autovectores a derecha y a izquierda (respectivamente) asociados a ellos; y sean \bar{u}^j los elementos de V^* asignados a \vec{u}_j a través de la métrica identidad. Entonces:

1. $\{\vec{v}_i\}_{i=1}^n$ es base de V , $\{\bar{u}^j\}_{j=1}^n$ es base de V^* , y son bases adjuntas; los \vec{v}_i y los \vec{u}_j son ortonormales dos a dos.⁵
2. Las representaciones matriciales v_i de \vec{v}_i son las columnas de una matriz V , y las representaciones matriciales $u^{j\dagger}$ de \bar{u}^j son las filas de su inversa V^{-1} , tales que la transformación de semejanza representada por V diagonaliza A : $V^{-1}AV = \text{diag}(\lambda_1, \dots, \lambda_n)$.

2.1.4. Operadores Hermitianos

Una clase de operadores muy importantes en Física son aquellos llamados Hermitianos, es decir aquellos cuya representación matricial es una matriz Hermitiana $A^\dagger = A$. En particular, en Mecánica Cuántica los operadores Hermitianos representan *observables* físicos, sus autovalores son los valores “permitidos”

⁴En este caso se dice que los autovalores son *simples*.

⁵Son ortogonales, y pueden elegirse normalizados.

de dichos observables, y sus autovectores (o autofunciones) representan los estados cuánticos correspondientes.

Supongamos que $\mathcal{A} : V \rightarrow V$ sea Hermitiano, A su representación matricial en una dada base $\{\vec{e}_i\}_{i=1}^n$ de V , y $\mathbf{Av}_i = \lambda_i \underline{v}_i$, $\mathbf{A}^\dagger \underline{u}_i = \lambda_i^* \underline{u}_i$. Entonces

1. $\mathbf{A}^\dagger = \mathbf{A} \implies \det(\lambda I - \mathbf{A}) = \det(\lambda I - \mathbf{A}^\dagger) \forall \lambda$; pero, factorizando el polinomio característico, vemos que

$$\det(\lambda I - \mathbf{A}) = (\lambda - \lambda_1) \dots (\lambda - \lambda_n),$$

$$\det(\lambda I - \mathbf{A}^\dagger) = (\lambda - \lambda_1^*) \dots (\lambda - \lambda_n^*).$$

Por lo tanto, podemos concluir que

$$\lambda_i \equiv \lambda_i^*,$$

es decir, los autovalores de un operador Hermitiano *son todos reales*.

2. Por el resultado precedente,

$$\mathbf{A}^\dagger = \mathbf{A} \implies \left\{ \begin{array}{l} \mathbf{Av}_i = \lambda_i \underline{v}_i \\ \mathbf{A}^\dagger \underline{u}_i = \lambda_i^* \underline{u}_i \end{array} \right\} \implies \underline{u}_i \equiv \underline{v}_i,$$

es decir, los autovectores a derecha e izquierda serán idénticos, $\vec{u}_i \equiv \vec{v}_i$, y $\underline{u}^{j\dagger} \equiv (\underline{v}_j)^\dagger$.

3. Asumiendo la métrica identidad en la base $\{\vec{e}_i\}_{i=1}^n$ de V , y normalizando los autovectores adecuadamente, si los autovalores λ_i son todos diferentes, entonces

$$\vec{v}_i \cdot \vec{v}_j = \delta_{ij}.$$

Con estos resultados, y los del Teorema 2.1.2, hemos demostrado el siguiente

Teorema 2.1.3. *Sea $\mathcal{A} : V \rightarrow V$ un operador lineal Hermitiano sobre un espacio vectorial V de dimensión n , con autovalores $\lambda_1, \dots, \lambda_n$; sean \vec{v}_i y \vec{u}_j los autovectores a derecha y a izquierda (respectivamente) asociados a ellos; y sean \vec{u}^j los elementos de V^* asignados a \vec{u}_j a través de la métrica identidad. Entonces:*

1. $\lambda_i^* \equiv \lambda_i$, es decir $\lambda_i \in \mathbb{R}$, $i = 1, \dots, n$, y $\vec{u}_i \equiv \vec{v}_i$.
2. Si λ_i son todos distintos, $\{\vec{v}_i\}_{i=1}^n$ es base de V , $\{\vec{v}^i\}_{i=1}^n$ es base de V^* , y son bases adjuntas y ortonormales.
3. Si λ_i son todos distintos, las representaciones matriciales \underline{v}_i de \vec{v}_i son las columnas de una matriz unitaria \mathbf{V} , y la transformación de semejanza representada por \mathbf{V} diagonaliza \mathbf{A} : $\mathbf{V}^{-1} \mathbf{A} \mathbf{V} = \mathbf{V}^\dagger \mathbf{A} \mathbf{V} = \text{diag}(\lambda_1, \dots, \lambda_n)$.

Ejemplo 2.1.10. La matriz $\mathbf{A} = \begin{pmatrix} 0 & 2i \\ -2i & 3 \end{pmatrix}$ es Hermitiana, con autovalores $\lambda_1 = -1$ y $\lambda_2 = 4$, y autovectores a derecha (normalizados) $\underline{v}_1 = \frac{1}{\sqrt{5}} \begin{pmatrix} 2 \\ i \end{pmatrix}$ y $\underline{v}_2 = \frac{1}{\sqrt{5}} \begin{pmatrix} i \\ 2 \end{pmatrix}$ mutuamente ortonormales; la transformación de semejanza viene dada por $\mathbf{V} = \frac{1}{\sqrt{5}} \begin{pmatrix} 2 & i \\ i & 2 \end{pmatrix}$ y $\mathbf{V}^\dagger = \frac{1}{\sqrt{5}} \begin{pmatrix} 2 & -i \\ -i & 2 \end{pmatrix}$, y resulta en $\mathbf{V}^\dagger \mathbf{A} \mathbf{V} = \begin{pmatrix} -1 & 0 \\ 0 & 4 \end{pmatrix}$.

Ejercicio 2.1.7. Verifique todas las afirmaciones de arriba.

2.1.5. Operadores Hermitianos con autovalores degenerados

Si los autovalores de un operador Hermitiano no son todos diferentes, nuestra demostración de que los autovectores forman base se desmorona. Veremos ahora que, aún en ese caso, podemos construir una base con los autovectores de un operador Hermitiano.

Sea $\mathcal{A} : V \rightarrow V$ Hermitiano, $\dim(V) = n$. El polinomio característico de \mathcal{A} *siempre* puede escribirse como

$$\wp_{\mathcal{A}}(\lambda) = (\lambda - \lambda_1)^{q_1}(\lambda - \lambda_2)^{q_2} \dots (\lambda - \lambda_r)^{q_r}, \quad \sum_{i=1}^r q_i = n.$$

Notemos que ahora denotamos por $\lambda_1, \lambda_2, \dots, \lambda_r$ las raíces *distintas* de $\wp_{\mathcal{A}}$. Para cada raíz λ_i , el número de veces q_i que aparece repetida se llama su *multiplicidad algebraica*. Si $q_i > 1$, el autovalor λ_i se llama *degenerado*.

Es fácil ver que $\mathcal{A}^\dagger = \mathcal{A} \implies \lambda_i^* \equiv \lambda_i \implies \lambda_i \in \mathbb{R}$, $i = 1, \dots, n$, como antes.

Para cada autovalor distinto λ_i , sabemos que $\det(\lambda_i \mathbf{I} - \mathcal{A}) = 0$. Por lo tanto, existe *al menos una* solución \vec{v}_i no trivial de $\mathcal{A}\vec{v}_i = \lambda_i \vec{v}_i$ para cada $i = 1, \dots, r$, y el correspondiente autovector a izquierda satisface $\vec{u}_i = \vec{v}_i$ como antes.

Construiremos ahora un procedimiento iterativo para obtener una base de autovectores.

1. Consideremos el autovalor λ_1 , y supongamos que hemos encontrado un autovector \vec{v}_1 asociado a él. Procedamos a normalizarlo, es decir, hagamos $\vec{v}_1^\dagger \vec{v}_1 = 1$.
2. Construyamos una base ortonormal $\{\vec{e}_i\}_{i=1}^n$ con $\mathbf{e}_1 := \vec{v}_1$, $\mathbf{e}_i^\dagger \mathbf{e}_j = \delta_{ij}$. Para ello podemos tomar cualquier conjunto LI de n vectores que contenga a \vec{v}_1 y ortonormalizarlo por el procedimiento de Gram-Schmidt.
3. Construyamos las matrices $n \times n$

$$\mathbf{V} := \begin{pmatrix} \mathbf{v}_1 & \mathbf{e}_2 & \dots & \mathbf{e}_n \\ \downarrow & \downarrow & & \downarrow \end{pmatrix}, \quad \mathbf{V}^\dagger := \begin{pmatrix} \mathbf{v}_1^\dagger & \longrightarrow \\ \mathbf{e}_2^\dagger & \longrightarrow \\ \vdots & \\ \mathbf{e}_n^\dagger & \longrightarrow \end{pmatrix},$$

que resultan no singulares por ser $\{\vec{e}_i\}_{i=1}^n$ base. Tendremos $\mathbf{V}^\dagger \mathbf{V} = \mathbf{I}$, luego \mathbf{V} es unitaria, $\mathbf{V}^{-1} = \mathbf{V}^\dagger$.

4. Construyamos

$$\mathbf{V}^\dagger \mathcal{A} \mathbf{V} = \begin{pmatrix} \mathbf{v}_1^\dagger \mathcal{A} \mathbf{v}_1 & \mathbf{v}_1^\dagger \mathcal{A} \mathbf{e}_2 & \dots & \mathbf{v}_1^\dagger \mathcal{A} \mathbf{e}_n \\ \mathbf{e}_2^\dagger \mathcal{A} \mathbf{v}_1 & \mathbf{e}_2^\dagger \mathcal{A} \mathbf{e}_2 & \dots & \mathbf{e}_2^\dagger \mathcal{A} \mathbf{e}_n \\ \vdots & \vdots & & \vdots \\ \mathbf{e}_n^\dagger \mathcal{A} \mathbf{v}_1 & \mathbf{e}_n^\dagger \mathcal{A} \mathbf{e}_2 & \dots & \mathbf{e}_n^\dagger \mathcal{A} \mathbf{e}_n \end{pmatrix} = \begin{pmatrix} \lambda_1 \mathbf{v}_1^\dagger \mathbf{v}_1 & \lambda_1 \mathbf{v}_1^\dagger \mathbf{e}_2 & \dots & \lambda_1 \mathbf{v}_1^\dagger \mathbf{e}_n \\ \lambda_1 \mathbf{e}_2^\dagger \mathbf{v}_1 & A'^2_2 & \dots & A'^2_n \\ \vdots & \vdots & & \vdots \\ \lambda_1 \mathbf{e}_n^\dagger \mathbf{v}_1 & A'^n_2 & \dots & A'^n_n \end{pmatrix},$$

donde hemos definido $A'^i{}_j := e^{i\dagger} A e_j$. Luego

$$V^\dagger A V = \left(\begin{array}{c|ccc} \lambda_1 & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & A' & \\ 0 & & & \end{array} \right),$$

con A' la matriz $(n-1) \times (n-1)$ con elementos $A'^i{}_j$.

5. Notemos que

$$\det(\lambda I - A) = \det[V^{-1}(\lambda I - A)V] = \det(\lambda I - V^{-1}AV)$$

$$= \left| \begin{array}{c|cccc} \lambda - \lambda_1 & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & \lambda I' - A' & \\ 0 & & & \end{array} \right| = (\lambda - \lambda_1) \det(\lambda I' - A'),$$

con I' la matriz identidad $(n-1) \times (n-1)$. Pero

$$\det(\lambda I - A) = (\lambda - \lambda_1)^{q_1} (\lambda - \lambda_2)^{q_2} \dots (\lambda - \lambda_r)^{q_r},$$

$$\therefore \det(\lambda I' - A') = (\lambda - \lambda_1)^{q_1-1} (\lambda - \lambda_2)^{q_2} \dots (\lambda - \lambda_r)^{q_r}.$$

6. Si $q_1 > 1$, $\det(\lambda_1 I' - A') = 0$; por lo tanto existe al menos una solución no trivial v'_2 de $A'v'_2 = \lambda_1 v'_2$. Podemos repetir entonces todo el proceso anterior, construyendo una segunda matriz unitaria

$$V' = \left(\begin{array}{c|ccccc} 1 & 0 & 0 & \cdots & 0 \\ 0 & v'_2 & e'_3 & \cdots & e'_n \\ \vdots & \downarrow & \downarrow & & \downarrow \end{array} \right)$$

tal que

$$V'^{-1} V^{-1} A V V' = \left(\begin{array}{cc|c} \lambda_1 & 0 & 0 \\ 0 & \lambda_1 & \\ \hline 0 & & A'' \end{array} \right),$$

con A'' una matriz $(n-2) \times (n-2)$.

7. Repitiendo el proceso q_1 veces, terminamos por construir una matriz unitaria $S_1 := VV'V''\dots V^{(q_1-1)}$ tal que

$$S_1 A S_1 = \left(\begin{array}{cc|c} \lambda_1 & 0 & \\ \ddots & & 0 \\ 0 & \lambda_1 & \\ \hline 0 & & A^{(q_1)} \end{array} \right).$$

El mismo procedimiento puede ahora aplicarse, por turno, a cada uno de los restantes autovalores distintos. Por lo tanto, hemos demostrado el siguiente

Teorema 2.1.4. *Sea $\mathcal{A} : V \rightarrow V$ un operador lineal Hermitiano sobre un espacio vectorial V de dimensión n , con*

$$\phi_{\mathcal{A}}(\lambda) = (\lambda - \lambda_1)^{q_1} (\lambda - \lambda_2)^{q_2} \dots (\lambda - \lambda_r)^{q_r}, \quad \sum_{i=1}^r q_i = n.$$

Entonces:

1. $\lambda_i \in \mathbb{R}$, $i = 1, \dots, r$, y a cada autovalor diferente λ_i hay asociados q_i autovectores LI.⁶
2. Los autovectores de \mathcal{A} forman una base ortonormal de V , por lo tanto \mathcal{A} es diagonalizable. En particular, $\mathbf{V} := \begin{pmatrix} v_1 & \cdots & v_n \\ \downarrow & \cdots & \downarrow \end{pmatrix}$ es unitaria, y

$$\mathbf{V}^\dagger \mathbf{A} \mathbf{V} = \text{diag}(\underbrace{\lambda_1, \dots, \lambda_1}_{q_1 \text{ veces}}, \underbrace{\lambda_2, \dots, \lambda_2}_{q_2 \text{ veces}}, \dots, \underbrace{\lambda_r, \dots, \lambda_r}_{q_r \text{ veces}}).$$

Notemos que a cada autovalor diferente λ_i hay asociado un *subespacio invariant* de \mathcal{A} que, *en este caso*, resulta ser $\ker(\lambda_i \mathbf{I} - \mathbf{A})$.

Si bien la demostración de este Teorema ha sido constructiva, el procedimiento seguido para construir la base donde la representación de \mathcal{A} es diagonal, es engorroso y nada conveniente en las aplicaciones. En la práctica, una vez que sabemos que a cada autovalor λ_i con multiplicidad q_i le corresponden q_i autovectores LI, sabemos que la solución general de

$$\mathbf{A} \mathbf{v}_i = \lambda_i \mathbf{v}_i$$

contendrá q_i parámetros libres. Haciendo sucesivamente cada uno de estos parámetros igual a 1 y los demás a cero, obtenemos q_i autovectores LI, que luego podremos ortonormalizar por el procedimiento de Gram-Schmidt. (Los autovectores correspondientes a autovalores diferentes serán automáticamente ortogonales entre sí).

Ejemplo 2.1.11. La matriz $\mathbf{A} = \begin{pmatrix} 0 & 0 & -2i \\ 0 & -1 & 0 \\ 2i & 0 & 3 \end{pmatrix}$ es Hermitiana, y tiene los autovalores $\lambda_1 = -1$ con multiplicidad 2 y $\lambda_2 = 4$ con multiplicidad 1. Los correspondientes autovectores (no normalizados) a derecha son $\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$ y $\begin{pmatrix} 2i \\ 0 \\ 1 \end{pmatrix}$ para λ_1 y $\begin{pmatrix} -i \\ 0 \\ 2 \end{pmatrix}$ para λ_2 , que evidentemente forman base de \mathbb{C}^3 . Cualquier combinación lineal de $\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$ y $\begin{pmatrix} 2i \\ 0 \\ 1 \end{pmatrix}$ es también un autovector de autovalor -1 .

Ejercicio 2.1.8. Demuestre todas las afirmaciones precedentes.

⁶En este caso se dice que los autovalores son *semisimples*.

2.1.6. Diagonalización simultánea de operadores Hermitianos

Ahora que ya sabemos que un operador Hermitiano siempre es diagonalizable, consideremos el siguiente problema: dados dos operadores lineales $\mathcal{A}, \mathcal{B} : V \rightarrow V$, ambos Hermitianos y por consiguiente diagonalizables *por separado*, ¿existirá alguna matriz V no singular tal que

$$\mathcal{A}' = V^{-1}\mathcal{A}V \quad \text{y} \quad \mathcal{B}' = V^{-1}\mathcal{B}V$$

sean *ambos* diagonales?⁷

Supongamos que tal matriz exista. Entonces, por ser diagonales.

$$\mathcal{A}'\mathcal{B}' = \mathcal{B}'\mathcal{A}',$$

y por consiguiente

$$\mathcal{A}\mathcal{B} = \mathcal{B}\mathcal{A},$$

ya que la comutatividad es propiedad de los *operadores* (véase el Ejercicio 1.4.2). Es decir, si \mathcal{A} y \mathcal{B} son diagonalizables *simultáneamente*, entonces comutan.

Supongamos ahora que \mathcal{A} y \mathcal{B} comutan.

1. Como \mathcal{A} es Hermitiana, existe alguna matriz V no singular tal que $\mathcal{A}' = V^{-1}\mathcal{A}V$ es diagonal, digamos $\mathcal{A}' = \text{diag}(a_1, \dots, a_n)$; luego $\mathcal{A}'^i_j = a_i \delta^i_j$.
2. Sea $\mathcal{B}' = V^{-1}\mathcal{B}V$; entonces como $\mathcal{A}\mathcal{B} = \mathcal{B}\mathcal{A} \implies \mathcal{A}'\mathcal{B}' = \mathcal{B}'\mathcal{A}'$, tendremos $[\mathcal{A}'\mathcal{B}']^i_j \equiv [\mathcal{B}'\mathcal{A}']^i_j$, de donde

$$\left. \begin{aligned} [\mathcal{A}'\mathcal{B}']^i_j &= A'^i_k B'^k_j = a_i \delta^i_k B'^k_j = a_i B'^i_j \\ [\mathcal{B}'\mathcal{A}']^i_j &= B'^i_l A'^l_j = B'^i_l a_l \delta^l_j = a_l B'^i_j \end{aligned} \right\} \implies a_i B'^i_j = a_l B'^i_j.$$

3. Para todo i y j tales que $a_i \neq a_j$, la ecuación anterior implica $B'^i_j = 0$. Ordenemos los elementos de \mathcal{A}' de modo que $a_1 = a_2 = \dots = a_{q_1} = \lambda_1$, el primer autovalor de \mathcal{A} ($q_1 \leq n$), y $a_{q_1+1}, \dots, a_n \neq \lambda_1$. Entonces

$$\mathcal{A}' = \begin{pmatrix} \lambda_1 I_{q_1} & 0 \\ 0 & \mathcal{A}'_2 \end{pmatrix}, \quad \mathcal{B}' = \begin{pmatrix} \mathcal{B}'_1 & 0 \\ 0 & \mathcal{B}'_2 \end{pmatrix},$$

con I_{q_1} la identidad $q_1 \times q_1$, $\mathcal{A}'_2 = \text{diag}(a_{q_1+1}, \dots, a_n)$, \mathcal{B}'_1 una matriz $q_1 \times q_1$, y \mathcal{B}'_2 una matriz $(n - q_1) \times (n - q_1)$.

⁷Este problema es de fundamental importancia en Mecánica Cuántica: que dos o más operadores Hermitianos puedan diagonalizarse simultáneamente quiere decir que podemos construir estados cuánticos con valores bien definidos de todos esos observables *a la vez*; por ejemplo, podremos tener simultáneamente valores bien definidos de la energía y del momento angular si y sólo si los correspondientes operadores pueden diagonalizarse simultáneamente.

4. \mathbf{B} Hermitiana $\implies \mathbf{B}'$ Hermitiana $\implies \exists \mathbf{W}_1$ unitaria, $q_1 \times q_1$, tal que $\mathbf{B}''_1 := \mathbf{W}_1^{-1} \mathbf{B}'_1 \mathbf{W}_1$ es *diagonal*. Definamos

$$\mathbf{W} = \begin{pmatrix} \mathbf{W}_1 & 0 \\ 0 & \mathbf{I}_{n-q_1} \end{pmatrix} \implies \mathbf{W}^{-1} = \begin{pmatrix} \mathbf{W}_1^{-1} & 0 \\ 0 & \mathbf{I}_{n-q_1} \end{pmatrix},$$

con \mathbf{I}_{n-q_1} la identidad $(n-q_1) \times (n-q_1)$. Entonces (operando por bloques)

$$\begin{aligned} \mathbf{A}'' &:= \mathbf{W}^{-1} \mathbf{A}' \mathbf{W} = \begin{pmatrix} \mathbf{W}_1^{-1} & 0 \\ 0 & \mathbf{I}_{n-q_1} \end{pmatrix} \begin{pmatrix} \lambda_1 \mathbf{I}_{q_1} & 0 \\ 0 & \mathbf{A}'_2 \end{pmatrix} \begin{pmatrix} \mathbf{W}_1 & 0 \\ 0 & \mathbf{I}_{n-q_1} \end{pmatrix} \\ &= \begin{pmatrix} \mathbf{W}_1^{-1} \lambda_1 \mathbf{I}_{q_1} \mathbf{W}_1 & 0 \\ 0 & \mathbf{I}_{n-q_1} \mathbf{A}'_2 \mathbf{I}_{n-q_1} \end{pmatrix} = \begin{pmatrix} \lambda_1 \mathbf{I}_{q_1} & 0 \\ 0 & \mathbf{A}'_2 \end{pmatrix}, \end{aligned}$$

que es una matriz diagonal, y

$$\begin{aligned} \mathbf{B}'' &:= \mathbf{W}^{-1} \mathbf{B}' \mathbf{W} = \begin{pmatrix} \mathbf{W}_1^{-1} & 0 \\ 0 & \mathbf{I}_{n-q_1} \end{pmatrix} \begin{pmatrix} \mathbf{B}'_1 & 0 \\ 0 & \mathbf{B}'_2 \end{pmatrix} \begin{pmatrix} \mathbf{W}_1 & 0 \\ 0 & \mathbf{I}_{n-q_1} \end{pmatrix} \\ &= \begin{pmatrix} \mathbf{W}_1^{-1} \mathbf{B}'_1 \mathbf{W}_1 & 0 \\ 0 & \mathbf{I}_{n-q_1} \mathbf{B}'_2 \mathbf{I}_{n-q_1} \end{pmatrix} = \begin{pmatrix} \mathbf{B}''_1 & 0 \\ 0 & \mathbf{B}''_2 \end{pmatrix}, \end{aligned}$$

con \mathbf{B}''_1 una matriz *diagonal* $q_1 \times q_1$, y $\mathbf{B}''_2 = \mathbf{B}'_2$.

5. Repitiendo este proceso para cada uno de los autovalores distintos de \mathbf{A} , terminamos construyendo una transformación de semejanza que diagonaliza simultáneamente a \mathbf{A} y \mathbf{B} .

Hemos demostrado entonces el siguiente

Teorema 2.1.5. Sean $\mathcal{A}, \mathcal{B} : V \rightarrow V$ Hermitianos. Entonces $\exists \mathbf{S}$ t.q. $\mathbf{S}^{-1} \mathcal{A} \mathbf{S}$ y $\mathbf{S}^{-1} \mathcal{B} \mathbf{S}$ son ambas diagonales, si y sólo si \mathcal{A} y \mathcal{B} comutan.

2.1.7. Operadores normales

Una clase más general de operadores, que incluye a los operadores Hermitianos como caso particular, son los operadores *normales*, que son aquellos que comutan con su adjunto Hermitiano: $\mathcal{A}\mathcal{A}^\dagger = \mathcal{A}^\dagger\mathcal{A}$.

Notemos que para toda matriz \mathbf{A} podemos escribir

$$\mathbf{A} = \mathbf{X} + i\mathbf{Y}, \text{ con } \mathbf{X} := \frac{1}{2}(\mathbf{A} + \mathbf{A}^\dagger), \mathbf{Y} := \frac{1}{2i}(\mathbf{A} - \mathbf{A}^\dagger),$$

descomposición que guarda cierta semejanza con la de un número complejo en sus partes real e imaginaria. Es inmediato de la definición que tanto \mathbf{X} como \mathbf{Y} son matrices Hermitianas, por lo que podremos diagonalizarlas simultáneamente si comutan. Pero

$$\mathbf{XY} - \mathbf{YX} = \frac{1}{4i}[(\mathbf{A} + \mathbf{A}^\dagger)(\mathbf{A} - \mathbf{A}^\dagger) - (\mathbf{A} - \mathbf{A}^\dagger)(\mathbf{A} + \mathbf{A}^\dagger)] = \frac{1}{2i}[\mathbf{A}^\dagger\mathbf{A} - \mathbf{AA}^\dagger],$$

de donde deducimos el siguiente

Teorema 2.1.6. *Todo operador lineal normal es diagonalizable.*

Debemos destacar que la conversa no es cierta: bien puede existir una transformación de semejanza que diagonalice A pero no diagonalice ni a X ni a Y por separado, como muestra el siguiente

Ejemplo 2.1.12. Consideremos la matriz $\begin{pmatrix} 1 & 1 \\ 0 & 2 \end{pmatrix}$. Es evidente que podemos diagonalizarla, ya que sus dos autovalores (1 y 2) son diferentes. Sin embargo, no es normal: $\begin{pmatrix} 1 & 1 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 1 & 2 \end{pmatrix} = \begin{pmatrix} 2 & 2 \\ 2 & 4 \end{pmatrix}$, pero $\begin{pmatrix} 1 & 0 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 5 \end{pmatrix}$.

2.2. Forma de Jordan

De lo expuesto en la Sección precedente queda claro que un operador será diagonalizable si todos sus autovalores son diferentes, o si es normal (este caso incluye a los operadores Hermitianos), o bien si, por el motivo que sea, sus autovalores son *semisimples*. La diagonalizabilidad resulta una característica altamente deseable, ya que en la base en que su representación es una matriz diagonal, la acción de un operador resulta evidente.

Es natural, entonces, preguntarnos cuál es la clase más general de operadores de los que podremos decir *a priori* que serán diagonalizables, y cuál será la forma más simple a la que podremos reducir la representación de un operador que no lo sea. Como veremos en lo que sigue, las respuestas a ambas preguntas están íntimamente relacionadas.

2.2.1. Ejemplos ilustrativos

En un sentido trivial, la respuesta a la pregunta de cuándo un operador podrá diagonalizarse es “cuando exista una base del espacio vectorial en la que su representación sea una matriz diagonal”. Sin embargo, si una matriz es diagonal, sus autovectores son simplemente los vectores de la base donde adopta esa forma (y los autovalores son obviamente los elementos de la diagonal); y si los autovectores forman base, es evidente que la representación del operador en esa base será una matriz diagonal. Pero el que los autovectores formen base es independiente de la base en que expresemos el operador, así que por ahora podemos expresar el siguiente

Lema 2.2.1. *Un operador $\mathcal{A} : V \rightarrow V$ es diagonalizable si sus autovectores forman base de V .*

Esto nos indica que, si queremos hallar operadores no diagonalizables, debemos buscar operadores cuyos autovectores *no formen base*; en ese caso, la matriz que representa al operador se dice *defectiva*. Demos algunos ejemplos:

Ejemplo 2.2.1. Consideremos un operador $\mathcal{A} : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ que en una dada base se representa por la matriz $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$. Su polinomio característico es $\varphi_A(\lambda) = (\lambda - 1)^2$, de modo que tiene un autovalor $\lambda_1 = 1$ con multiplicidad 2. Pero la ecuación de autovectores $(A - \lambda_1 I)v = 0$ tiene en este caso una sola

solución LI, ya que $(A - \lambda_1 I) = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$; poniendo $v = \begin{pmatrix} x \\ y \end{pmatrix}$ tenemos que $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \implies y = 0$, de modo que $\ker(A - \lambda_1 I) = \text{span}\left\{\begin{pmatrix} 1 \\ 0 \end{pmatrix}\right\}$.

Dicho de otro modo, el subespacio de \mathbb{R}^2 desarrollado por los autovectores (uno, en este caso) es unidimensional. Sin embargo, cabe preguntarse qué ocurre con los vectores que *no* pertenecen a este subespacio. En particular, vemos que $(A - \lambda_1 I)\begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, es decir, los vectores del subespacio complementario de $\ker(A - \lambda_1 I)$ son llevados a $\ker(A - \lambda_1 I)$ por acción de $(A - \lambda_1 I)$. Una vez allí, una segunda aplicación de $(A - \lambda_1 I)$ los enviará al vector nulo, ya que en efecto $(A - \lambda_1 I)^2 = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$, de modo que $\ker[(A - \lambda_1 I)^2] = \mathbb{R}^2$, el espacio completo.

Notemos para terminar que la ecuación $(A - \lambda_1 I)^2 v = 0$ *sí* tiene dos soluciones LI, por ejemplo $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ y $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$.

Ejemplo 2.2.2. Consideremos ahora un operador $A : \mathbb{R}^4 \rightarrow \mathbb{R}^4$, que en alguna base se representa

$$A = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 2 \end{pmatrix}$$

Su polinomio característico es $\varphi_A(\lambda) = (\lambda - 1)^2(\lambda - 2)^2$, de modo que tiene dos autovalores diferentes $\lambda_1 = 1$ y $\lambda_2 = 2$, ambos con multiplicidad 2. Tenemos

$$A - \lambda_1 I = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad A - \lambda_2 I = \begin{pmatrix} -1 & 1 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

luego

$$\ker(A - \lambda_1 I) = \text{span}\left\{\begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}\right\}, \quad \ker(A - \lambda_2 I) = \text{span}\left\{\begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}\right\}.$$

Nuevamente vemos que los autovectores no forman base, pero

$$(A - \lambda_1 I) \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad (A - \lambda_2 I) \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}.$$

Notando que

$$(A - \lambda_1 I)^2 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad (A - \lambda_2 I)^2 = \begin{pmatrix} 1 & -2 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

tenemos

$$\ker[(A - \lambda_1 I)^2] = \text{span}\left\{\begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}\right\}, \quad \ker[(A - \lambda_2 I)^2] = \text{span}\left\{\begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}\right\},$$

y vemos que son *subespacios invariantes* bajo A .

2.2.2. Descomposición primaria

De los ejemplos precedentes podemos ver aparecer un patrón, que ahora vamos a formalizar.

Supongamos que V es suma directa de cierta colección de subespacios V_i , $i = 1, \dots, r$, y denotemos por \mathcal{P}_i el proyector sobre V_i . Es inmediato que

- $\mathcal{P}_i^2 \equiv \mathcal{P}_i$;
- $\mathcal{P}_i \mathcal{P}_j = 0$ si $i \neq j$;
- $\sum_{i=1}^r \mathcal{P}_i = \mathcal{I}$.

También es inmediato que $\forall \vec{x} \in V$,

- $\vec{x}_i := \mathcal{P}_i \vec{x} \in V_i$;
- $\vec{x} = \sum_{i=1}^r \vec{x}_i$.

Las restricciones de un operador $\mathcal{A} : V \rightarrow V$ a cada subespacio V_i las denotaremos por $\mathcal{A}_i := \mathcal{A}/V_i$, y las identificaremos con los correspondientes operadores $\mathcal{AP}_i : V \rightarrow V$ (ver el comentario final de la Sección 1.2.1). De lo anterior se desprende que

$$\mathcal{A}\vec{x} \equiv \sum_{i=1}^r \mathcal{A}_i \vec{x}_i. \quad (2.15)$$

Si cada subespacio V_i es además invariante bajo \mathcal{A} , es decir si $\vec{x}_i \in V_i \implies \mathcal{A}\vec{x}_i \equiv \mathcal{A}_i \vec{x}_i \in V_i$, $i = 1, \dots, r$, diremos que \mathcal{A} es *suma directa* de sus restricciones \mathcal{A}_i a cada subespacio, y lo denominaremos

$$\mathcal{A} = \bigoplus_{i=1}^r \mathcal{A}_i. \quad (2.16)$$

Una consecuencia inmediata de lo anterior es que, si tenemos bases B_i de cada V_i y formamos una base $B = \{B_1, \dots, B_r\}$ de V colecciónándolas, entonces en la base B la matriz que representa al operador \mathcal{A} toma la forma *diagonal por bloques*

$$\mathbf{A} = \begin{pmatrix} \mathbf{A}_1 & 0 & \cdots & 0 \\ 0 & \mathbf{A}_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \mathbf{A}_r \end{pmatrix}, \quad (2.17)$$

donde cada \mathbf{A}_i es una matriz $\dim(V_i) \times \dim(V_i)$, que podemos pensar como la representación matricial de \mathcal{A}_i en la base B_i .⁸

⁸Con la definición “habitual” de \mathcal{A}/V_i , tenemos $\mathcal{A}_i : V_i \rightarrow V_i$, y las representaciones matriciales de cada \mathcal{A}_i en la correspondiente base B_i de V_i son las matrices \mathbf{A}_i ; sin embargo, si identificamos (como lo hacemos aquí) a \mathcal{A}/V_i con \mathcal{AP}_i , estos son operadores de V en V , y sus representaciones matriciales en la base B de V son

$$\tilde{\mathbf{A}}_1 = \begin{pmatrix} \mathbf{A}_1 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix}, \quad \tilde{\mathbf{A}}_2 = \begin{pmatrix} 0 & 0 & \cdots & 0 \\ 0 & \mathbf{A}_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix}, \quad \dots, \quad \tilde{\mathbf{A}}_r = \begin{pmatrix} 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \mathbf{A}_r \end{pmatrix}.$$

La ventaja de este enfoque es que podemos escribir $\mathbf{A} = \sum_{i=1}^r \tilde{\mathbf{A}}_i$.

Supongamos ahora que $\mathcal{A} : V \rightarrow V$ es un operador lineal, con $\dim(V) = n$, y que $\wp_{\mathcal{A}}(\lambda) = \prod_{i=1}^r (\lambda - \lambda_i)^{q_i}$, $\sum_{i=1}^r q_i = n$. Definimos el *subespacio característico* V_i asociado al autovalor λ_i como

$$V_i := \ker[(\mathcal{A} - \lambda_i I)^{q_i}] \quad (2.18)$$

(nótese que $\dim(V_i) = q_i$). Enunciaremos, sin demostración, el siguiente

Teorema 2.2.2 (Descomposición Primaria). *Sea $\mathcal{A} : V \rightarrow V$ un operador lineal, entonces:*

1. *V es suma directa de los subespacios característicos V_i de $\mathcal{A}: V = \bigoplus_{i=1}^r V_i$.*
2. *Cada V_i es invariante bajo \mathcal{A} , y por consiguiente $\mathcal{A} = \bigoplus_{i=1}^r \mathcal{A}/V_i$.*
3. *El polinomio característico de \mathcal{A}/V_i es $\wp_{\mathcal{A}/V_i}(\lambda) = (\lambda - \lambda_i)^{q_i}$.*
4. *Cada \mathcal{A}_i puede descomponerse como $\mathcal{A}_i = \mathcal{D}_i + \mathcal{N}_i$, donde $\mathcal{D}_i := \lambda_i \mathcal{P}_i$, que es diagonalizable, y $\mathcal{N}_i := (\mathcal{A}_i - \lambda_i I)\mathcal{P}_i$, que es nilpotente de índice $\leq q_i$.*⁹

Es sencillo demostrar que todos los operadores \mathcal{D}_i y \mathcal{N}_i commutan entre sí.

Ejercicio 2.2.1. Consideré la matriz $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ del Ejemplo 2.2.1. Muestre que $D = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ y $N = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ commutan.

Ejercicio 2.2.2. Consideré la matriz del Ejemplo 2.2.2. Construya D_1 , D_2 , N_1 y N_2 , y muestre que commutan.

Las primeras tres afirmaciones del Teorema anterior implican que, para todo operador lineal $\mathcal{A} : V \rightarrow V$, existirá una base de V donde su representación matricial adopte una forma diagonal por bloques como la mostrada en (2.17), y que cada bloque representará la restricción de \mathcal{A} al correspondiente subespacio invariante. La cuarta afirmación nos dice que, además, existirá una base donde cada bloque \mathcal{A}_i adopte la forma triangular¹⁰

$$\mathcal{A}_i = \begin{pmatrix} \lambda_i & x & x & \cdots & x & x \\ 0 & \lambda_i & x & \cdots & x & x \\ 0 & 0 & \lambda_i & \cdots & x & x \\ \vdots & & & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda_i & x \\ 0 & 0 & 0 & \cdots & 0 & \lambda_i \end{pmatrix}, \quad (2.19)$$

donde las “ x ” representan elementos posiblemente no nulos, que dan la representación matricial del nilpotente \mathcal{N}_i .

⁹Un operador \mathcal{N} se dice nilpotente de índice q si $\mathcal{N}^q = 0$ pero $\mathcal{N}^r \neq 0$ si $r < q$.

¹⁰Por convención se asume *triangular superior*.

Ejercicio 2.2.3. Considere la matriz

$$A = \frac{1}{2} \begin{pmatrix} 3 & 2 & -1 & 0 \\ 0 & 3 & 0 & -1 \\ -1 & 0 & 3 & 2 \\ 0 & -1 & 0 & 3 \end{pmatrix}.$$

Determine sus autovalores con sus multiplicidades y los correspondientes subespacios característicos, construya una base para cada uno, y construya una transformación de semejanza que reduzca A a una matriz triangular superior por bloques.

2.2.3. Forma normal de Jordan

Vamos a analizar ahora en mas detalle la estructura de cada uno de los bloques de la descomposición primaria de un operador lineal. Para ello deberemos introducir primero algo de nomenclatura.

Un operador lineal $\mathcal{A} : V \rightarrow V$ se llama *nilcíclico* de orden n [= $\dim(V)$] si existe una base $\{\vec{e}_1, \dots, \vec{e}_n\}$ de V tal que

$$\begin{aligned} \mathcal{A}\vec{e}_i &= \vec{e}_{i-1}, \quad i = n, n-1, \dots, 2, \\ \mathcal{A}\vec{e}_1 &= \vec{0}. \end{aligned}$$

Una tal base se llama *cíclica*, y en ella la representación matricial de \mathcal{A} adopta la forma

$$A = \begin{pmatrix} 0 & 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & 1 & \cdots & 0 & 0 \\ \vdots & & \ddots & \ddots & & & \vdots \\ \vdots & & & \ddots & \ddots & & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & 0 & 0 & \cdots & 0 & 0 \end{pmatrix}.$$

Notemos que para un operador nilcíclico, $\varphi_A(\lambda) = \lambda^n$, y que su único autovector, escrito en la base cíclica, es \vec{e}_1 , con autovalor cero. También se verifica que todo operador nilcíclico de orden n es un operador nilpotente de índice n .

Ejercicio 2.2.4. Demuestre las afirmaciones precedentes.

Enunciaremos ahora sin demostración el siguiente

Teorema 2.2.3. *Sea $\mathcal{N} : V \rightarrow V$ un operador nilpotente sobre un espacio V de dimensión finita n . Entonces existen subespacios V_i , $i = 1, \dots, r$, tales que*

1. *V es su suma directa: $V = \bigoplus_{i=1}^r V_i$;*
2. *las restricciones \mathcal{N}/V_i de \mathcal{N} a cada uno de ellos son operadores nilcíclicos de orden $\dim(V_i)$;*
3. *\mathcal{N} es suma directa de sus restricciones: $\mathcal{N} = \bigoplus_{i=1}^r \mathcal{N}/V_i$.*

Esto significa que, dado un operador nilpotente \mathcal{N} sobre un espacio V , existirá una base de V donde su representación matricial adopte la forma diagonal por bloques

$$\mathbf{N} = \begin{pmatrix} \mathbf{N}_1 & 0 & \cdots & 0 \\ 0 & \mathbf{N}_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \mathbf{N}_r \end{pmatrix}, \quad (2.20)$$

donde cada matriz \mathbf{N}_i representa un operador nilcíclico de orden $\dim(V_i)$ igual a la dimensión del subespacio V_i sobre el que actúa.

Aplicando el teorema precedente a la descomposición primaria de un operador $\mathcal{A} : V \rightarrow V$, escrita como en (2.17), vemos que siempre existirá alguna base de V donde cada bloque \mathbf{A}_i adopte la forma particularmente simple

$$\mathbf{A}_i = \begin{pmatrix} \mathbf{J}_{(\lambda_i)t_{i1}} & 0 & \cdots & 0 \\ 0 & \mathbf{J}_{(\lambda_i)t_{i2}} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \mathbf{J}_{(\lambda_i)t_{ir_i}} \end{pmatrix}, \quad (2.21)$$

donde cada *bloque de Jordan* $\mathbf{J}_{(\lambda_i)t_{ij}}$ será una matriz $t_{ij} \times t_{ij}$ de la forma

$$\mathbf{J}_{(\lambda_i)t_{ij}} = \begin{pmatrix} \lambda_i & 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & \lambda_i & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & \lambda_i & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \ddots & \vdots & \vdots \\ \vdots & \vdots & \vdots & & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & \lambda_i & 1 \\ 0 & 0 & 0 & 0 & \cdots & 0 & \lambda_i \end{pmatrix}, \quad (2.22)$$

con el autovalor λ_i en la diagonal principal, unos en la diagonal inmediata superior, y ceros en todas las demás posiciones. También se desprende de lo anterior que a *cada* bloque de Jordan $\mathbf{J}_{(\lambda_i)t_{ij}}$ habrá asociado un *único* autovector de \mathcal{A} de autovalor λ_i .

2.2.4. Reducción a la forma normal de Jordan

Enfrentaremos ahora el problema siguiente: dado un operador lineal $\mathcal{A} : V \rightarrow V$, y conocida su representación matricial \mathbf{A} en una dada base $\{\vec{e}_i\}_{i=1}^n$ de V , encontrar: (i) su forma normal de Jordan; (ii) la transformación de semejanza que lleva \mathbf{A} a esa forma. Daremos un método constructivo para resolver este problema, que puede entenderse como un algoritmo. El procedimiento es directo, pero largo y tedioso; el lector interesado en hallar sólo la forma de Jordan, y no la transformación que lleva a ella, puede leer sólo los comentarios al final de esta sección y continuar con la Sección 2.2.5.

Descomposición primaria

Lo primero que debemos hacer es encontrar una base de V que implemente la descomposición primaria de A . Para ello deberemos encontrar los autovalores de A y los subespacios característicos asociados a cada uno.

1. Construimos el polinomio característico

$$\varphi_A(\lambda) = \det(\lambda I - A),$$

resolvemos la ecuación secular

$$\varphi_A(\lambda) = 0$$

para encontrar los autovalores λ_i con sus respectivas multiplicidades q_i , y factorizamos el polinomio característico en la forma

$$\varphi_A(\lambda) = \prod_{i=1}^r (\lambda - \lambda_i)^{q_i}.$$

2. Para cada autovalor distinto λ_i , determinamos la matriz columna v_i , solución general de la ecuación

$$(A - \lambda_i I)^{q_i} v_i = 0.$$

Esta solución contendrá q_i parámetros libres, así que de ella podremos obtener q_i columnas v_{i1}, \dots, v_{iq_i} LI,¹¹ que desarrollan el subespacio característico V_i de A asociado al autovalor λ_i .

3. De manera automática (por construcción), todas las columnas obtenidas serán LI entre sí. Construimos entonces con las columnas v_{il} la matriz

$$T := \begin{pmatrix} v_{11} & \cdots & v_{rq_r} \\ \downarrow & & \downarrow \\ & & \end{pmatrix},$$

y calculamos su inversa T^{-1} . La transformación de semejanza definida por estas matrices llevará A a la forma diagonal por bloques

$$A' = T^{-1}AT = \begin{pmatrix} A_1 & 0 & \cdots & 0 \\ 0 & A_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & A_r \end{pmatrix}.$$

Hasta aquí, habremos realizado la *descomposición primaria* de A .¹² Debe notarse la fuerte similitud entre el procedimiento llevado a cabo aquí y el seguido usualmente para diagonalizar una matriz diagonalizable, siendo la principal diferencia que en lugar de trabajar con *autovectores*, soluciones de $(A - \lambda_i I)v_i = 0$, aquí trabajamos con las soluciones de $(A - \lambda_i I)^{q_i}v_i = 0$.

¹¹Por ejemplo, poniendo sucesivamente cada parámetro libre igual a 1 y los demás a 0.

¹²Alternativamente, en el paso (2) podemos determinar también cada matriz fila $u^{i\dagger}$, solución general de la ecuación

$$u^{i\dagger}(A - \lambda_i I)^{q_i} = 0,$$

Reducción de cada bloque a la forma normal

Pasamos ahora a trabajar en la reducción de cada bloque A_i a su forma normal de Jordan. Notemos primero que, al trabajar con cada bloque A_i , ya no necesitamos trabajar en el espacio completo V , sino sólo en el correspondiente subespacio característico V_i . Esto quiere decir que si estamos trabajando con el bloque A_i , nuestras filas y columnas sólo contendrán elementos distintos de cero en las posiciones correspondientes a ese bloque, es decir serán de la forma

$$\begin{aligned} v &= (\underbrace{0, \dots, 0}_{q_1 + \dots + q_{i-1} \text{ veces}}, \underbrace{v, \dots, v}_{q_i \text{ elementos no nulos}}, \underbrace{0, \dots, 0}_{q_{i+1} + \dots + q_r \text{ veces}})^T, \\ u^\dagger &= (\underbrace{0, \dots, 0}_{q_1 + \dots + q_{i-1} \text{ veces}}, \underbrace{u, \dots, u}_{q_i \text{ elementos no nulos}}, \underbrace{0, \dots, 0}_{q_{i+1} + \dots + q_r \text{ veces}}). \end{aligned}$$

En la práctica, entonces, para cada bloque A_i trabajaremos con columnas y filas de tamaño q_i (los únicos elementos posiblemente no nulos), y recién al final ordenaremos estos elementos en filas y columnas de tamaño n .

Como se verá a continuación, el procedimiento a seguir con cada bloque A_i consiste en ir construyendo, por un método iterativo, las bases cíclicas de cada uno de los nilcíclicos contenidos en A_i .

4. Para cada bloque A_i debemos comenzar encontrando todos los autovectores LI correspondientes al autovalor λ_i . Para ello, determinamos la matriz columna v ($\text{j}(\text{tamaño } q_i!)$), solución general de la ecuación

$$(A_i - \lambda_i I)v = 0.$$

Esta solución contendrá $m_i \leq q_i$ parámetros libres, así que de ella podemos obtener m_i columnas v_1, \dots, v_{m_i} LI,¹³ que desarrollan el subespacio $\ker(A_i - \lambda_i I)$.¹⁴ En este paso, es altamente conveniente ortonormalizar

que contendrá q_i parámetros libres, y obtener de ella q_i filas $u^{i1\dagger}, \dots, u^{iq_i\dagger}$ LI, que desarrollan el subespacio dual de V_i . Por construcción, las filas de un subespacio V_i y las columnas de cualquier otro subespacio distinto V_j , $i \neq j$, serán mutuamente ortogonales:

$$u^{il\dagger} v_{jk} = 0 \quad \forall l = 1, \dots, q_i, k = 1, \dots, q_j \text{ si } i \neq j.$$

Dentro de cada subespacio V_i , podemos aplicar a las correspondientes filas y columnas el procedimiento de Gram-Schmidt para convertirlas en mutuamente ortonormales (si ya no lo fueran). Obtenemos así un conjunto de n columnas v_{il} y otro de n filas $u^{il\dagger}$, con $i = 1, \dots, r$, $l = 1, \dots, q_i$, que son mutuamente *ortonormales*:

$$u^{il\dagger} v_{jk} = \delta^i_j \delta^l_k.$$

La matriz T se construye entonces como en el paso (3), pero su inversa se obtiene simplemente poniendo

$$T^{-1} := \begin{pmatrix} u^{11\dagger} & \longrightarrow \\ \vdots & \\ u^{rq_r\dagger} & \longrightarrow \end{pmatrix}.$$

Cuál de las dos alternativas resulta más conveniente en la práctica, dependerá de la matriz A .

¹³Por ejemplo, poniendo sucesivamente cada parámetro libre igual a 1 y los demás a 0.

¹⁴Recordemos que $\dim(\ker(A_i - \lambda_i I)) = m_i$, el número de autovectores LI asociados a λ_i .

estas columnas (p.ej. por el procedimiento de Gram-Schmidt), y en lo que sigue *asumiremos que así se hizo*, tras lo cual construimos la columna

$$\mathbf{v} = a_1 \mathbf{v}_1 + \cdots + a_{m_i} \mathbf{v}_{m_i},$$

con a_1, \dots, a_{m_i} parámetros libres y $\mathbf{v}_1, \dots, \mathbf{v}_{m_i}$ mutuamente ortonormales.

5. Buscamos la columna \mathbf{w}_1 , solución general de la ecuación

$$(\mathbf{A}_i - \lambda_i \mathbf{I})\mathbf{w}_1 = \mathbf{v}.$$

Supondremos primero que este sistema de ecuaciones *no impone ninguna condición* sobre los parámetros a_s para que exista solución. Entonces \mathbf{w}_1 contendrá $2m_i$ parámetros libres,¹⁵ ya que como toda solución general de un sistema *inhomogéneo* lleva sumada la solución general del sistema *homogéneo*; de estos parámetros, m_i serán los a_s y los otros m_i serán nuevos, llamémoslos b_s .¹⁶

Para seguir, buscamos una columna \mathbf{w}_2 , solución general de

$$(\mathbf{A}_i - \lambda_i \mathbf{I})\mathbf{w}_2 = \mathbf{w}_1,$$

e iteramos el procedimiento anterior. Otra vez aparecerán m_i nuevos parámetros, llamémoslos c_i . Es posible que a medida que avancemos algunos de los nuevos parámetros vayan quedando determinados en términos de otros o de los originales a_s , pero mientras no aparezca una condición sobre los parámetros *originales* a_s , seguimos iterando este paso.¹⁷

6. Supongamos en cambio que en una dada iteración del paso anterior determinamos (por primera vez) que el sistema

$$(\mathbf{A}_i - \lambda_i \mathbf{I})\mathbf{w}_t = \mathbf{w}_{t-1}$$

no tiene solución a menos que se cumpla cierta relación entre los parámetros *originales* a_s presentes en \mathbf{w}_{t-1} , la que adoptará la forma genérica

$$k_1 a_1 + \cdots + k_{m_i} a_{m_i} = 0,$$

con k_1, \dots, k_{m_i} dados.

¹⁵Obviamente este caso presupone $2m_i \leq q_i$.

¹⁶Estos nuevos parámetros b_s simplemente aportan la solución general del sistema homogéneo, que en principio podríamos pensar que no queremos y que deberíamos descartar poniendo $b_s \equiv 0$; porque estaríamos buscando una preimagen de $\ker(\mathbf{A}_i - \lambda_i \mathbf{I})$ bajo $(\mathbf{A}_i - \lambda_i \mathbf{I})$ que “no contenga nada” que ya estuviera en $\ker(\mathbf{A}_i - \lambda_i \mathbf{I})$ (lo que llamaríamos una preimagen “propia”). Sin embargo, pese a que estamos construyendo una base del subespacio invariante correspondiente a \mathbf{A}_i , generalmente ella no será ortogonal, de modo que “no estar” en $\ker(\mathbf{A}_i - \lambda_i \mathbf{I})$ *no implica* no tener proyección (ortogonal) sobre (o componentes en) él.

¹⁷Así, vamos generando en cada paso subespacios que son llevados cada uno al anterior [y finalmente a $\ker(\mathbf{A}_i - \lambda_i \mathbf{I})$] por sucesivas aplicaciones de $(\mathbf{A}_i - \lambda_i \mathbf{I})$.

Por una parte, esto quiere decir que existe un autovector particular en

$$\mathbf{v} = a_1 \mathbf{v}_1 + \cdots + a_{m_i} \mathbf{v}_{m_i}$$

que debemos eliminar para poder seguir adelante. Este autovector tendrá justamente la forma

$$\mathbf{v}_{10} = k_1 \mathbf{v}_1 + \cdots + k_{m_i} \mathbf{v}_{m_i},$$

y sus sucesivas preimágenes $\mathbf{v}_{11}, \mathbf{v}_{12}, \dots, \mathbf{v}_{1(t-1)}$ se obtienen análogamente, substituyendo cada a_s por el correspondiente k_s en $\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_{t-1}$, respectivamente. A éste *autovector particular* (y sus preimágenes) tendremos asociado un bloque de Jordan $J_{(\lambda_i)t}$ de tamaño $t \times t$, que actuará sobre el subespacio $V_{i1} := \text{span}\{\mathbf{v}_{10}, \mathbf{v}_{11}, \dots, \mathbf{v}_{1(t-1)}\}$ de V_i .¹⁸

Por otra parte, deberemos *imponer* la condición

$$k_1 a_1 + \cdots + k_{m_i} a_{m_i} = 0$$

a

$$\mathbf{v} = a_1 \mathbf{v}_1 + \cdots + a_{m_i} \mathbf{v}_{m_i},$$

eliminando así \mathbf{v}_{10} y obteniendo un subespacio de $\ker(\mathbf{A}_i - \lambda_i \mathbf{I})$ disjunto (es decir, complementario) de $\text{span}\{\mathbf{v}_{10}\}$. Llamaremos $\mathbf{v}', \mathbf{w}'_1, \mathbf{w}'_2, \dots, \mathbf{w}'_{t-1}$ al resultado de imponer la condición en \mathbf{v} y sus sucesivas preimágenes, todos los cuales tendrán ahora $m_i - 1$ parámetros a_s libres. Veremos que ahora la ecuación

$$(\mathbf{A}_i - \lambda_i \mathbf{I}) \mathbf{w}'_t = \mathbf{w}'_{t-1}$$

sí tiene solución. Continuaremos entonces con el proceso iterativo, pero con un parámetro original a_s menos.¹⁹

¹⁸Nótese que $\{\mathbf{v}_{10}, \mathbf{v}_{11}, \dots, \mathbf{v}_{1(t-1)}\}$ forman, por construcción, una base cíclica de V_i .

¹⁹Si hubiéramos determinado que el sistema

$$(\mathbf{A}_i - \lambda_i \mathbf{I}) \mathbf{w}_t = \mathbf{w}_{t-1}$$

no tiene solución a menos que se cumplan *dos* relaciones entre los a_s presentes en \mathbf{w}_{t-1} , de la forma

$$\begin{aligned} k_1 a_1 + \cdots + k_{m_i} a_{m_i} &= 0, \\ h_1 a_1 + \cdots + h_{m_i} a_{m_i} &= 0, \end{aligned}$$

(k_1, \dots, k_{m_i} y h_1, \dots, h_{m_i} dados), procedemos de igual forma, pero determinando ahora *dos* autovectores particulares

$$\begin{aligned} \mathbf{v}_{10} &= k_1 \mathbf{v}_1 + \cdots + k_{m_i} \mathbf{v}_{m_i}, \\ \mathbf{v}_{20} &= h_1 \mathbf{v}_1 + \cdots + h_{m_i} \mathbf{v}_{m_i}, \end{aligned}$$

cada uno con su respectivo bloque de Jordan $J_{(\lambda_i)t}$ asociado, de tamaño $t \times t$, actuando sobre sendos subespacios generados por \mathbf{v}_{10} y \mathbf{v}_{20} y sus respectivas preimágenes. Imponiendo las dos condiciones a \mathbf{v} y sus preimágenes (dejando $m_i - 2$ parámetros a_s libres en ellos), la ecuación

$$(\mathbf{A}_i - \lambda_i \mathbf{I}) \mathbf{w}'_t = \mathbf{w}'_{t-1}$$

tendrá solución, y continuaremos entonces con el proceso iterativo, con dos parámetros originales a_s menos. La generalización al caso de tres o más condiciones es directa.

A medida que procedamos, los pasos (5) y (6) se irán alternando, e irán quedando cada vez menos parámetros *originales* a_s libres, es decir sin haber sido eliminados en términos de los demás. El proceso iterativo de cada bloque A_i terminará cuando, quedando $p \geq 1$ parámetros originales libres (digamos a_1, \dots, a_p), obtengamos p condiciones

$$\begin{aligned} k_1 a_1 + \dots + k_p a_p &= 0, \\ h_1 a_1 + \dots + h_p a_p &= 0, \\ &\vdots \\ f_1 a_1 + \dots + f_p a_p &= 0 \end{aligned}$$

que impongan $a_1 = \dots = a_p = 0$. Determinaremos entonces los últimos p autovectores particulares y sus respectivas preimágenes, y los respectivos bloques de Jordan.

Una vez determinados todos los autovectores particulares y sus preimágenes, pueden quedar sin determinar parámetros b_s, c_s , etc. de las contribuciones de la solución homogénea incorporadas en cada iteración del paso 5. Procedemos entonces a “hacer limpieza” simplemente anulándolos.²⁰

Finalmente, una vez completados los pasos 4 a 6 para cada bloque A_i de la descomposición primaria, procedemos como sigue:

Construcción de la transformación de semejanza

7. En cada subespacio V_i construimos la matriz $q_i \times q_i$

$$S_i := \begin{pmatrix} v_{10} & v_{11} & \cdots & v_{1(t_1-1)} & v_{20} & v_{21} & \cdots & v_{2(t_2-1)} & \cdots \text{etc.} \cdots \\ \downarrow & \downarrow & \cdots & \downarrow & \downarrow & \downarrow & \cdots & \downarrow & \cdots \text{etc.} \cdots \end{pmatrix},$$

y su inversa S_i^{-1} ; la transformación de semejanza generada por ellas dejará el bloque A_i en la forma

$$S_i^{-1} A_i S_i = \begin{pmatrix} J_{(\lambda_i)t_1} & & & \\ & J_{(\lambda_i)t_2} & & \\ & & \ddots & \\ & & & \ddots \end{pmatrix}.$$

8. Construimos ahora las matrices $n \times n$

$$S := \begin{pmatrix} S_1 & 0 & \cdots & 0 \\ 0 & S_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & S_r \end{pmatrix}, \quad S^{-1} := \begin{pmatrix} S_1^{-1} & 0 & \cdots & 0 \\ 0 & S_2^{-1} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & S_r^{-1} \end{pmatrix}.$$

²⁰Ya que, por ejemplo, los b_s aún libres tras finalizar la iteración de los pasos 5 y 6 sólo aportan a w_1 vectores que ya están en v , a w_2 vectores que ya están en w_1 , etc.; los c_s libres sólo aportan a w_2 vectores que ya están en v , a w_3 vectores que ya están en w_1 , etc.; etc. Como se verá en los ejemplos que siguen, muchas veces podemos descartar estos parámetros tan pronto estemos seguros que ya no serán necesarios en ulteriores iteraciones.

Aplicando la transformación de semejanza generada por ellas a la descomposición primaria A' de A , tendremos

$$A'' = S^{-1}T^{-1}ATS$$

en su forma normal de Jordan.

Ejemplo 2.2.3. Presentaremos primero un ejemplo relativamente simple para fijar algunas ideas. Utilizaremos la misma numeración del texto precedente, para facilitar el seguimiento de cada paso.

Sea la matriz

$$A = \begin{pmatrix} 2 & 1 & -1 & 0 & 1 \\ 0 & 2 & 0 & 1 & 1 \\ 0 & 0 & 2 & 0 & 1 \\ 0 & 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 0 & 2 \end{pmatrix}.$$

Descomposición primaria

Un rápido examen de ésta matriz muestra que sólo tenemos un único autovector $\lambda = 2$ con multiplicidad 5, por lo que ya estamos en una base que implementa la descomposición primaria. Los pasos 1 a 3 pueden darse entonces por completados.

Reducción a forma normal

Necesitaremos

$$(A - \lambda I)x = \begin{pmatrix} 0 & 1 & -1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} = \begin{pmatrix} x_2 - x_3 + x_5 \\ x_4 + x_5 \\ x_5 \\ x_5 \\ 0 \end{pmatrix}.$$

4. Los autovectores vienen dados por

$$\begin{pmatrix} x_2 - x_3 + x_5 \\ x_4 + x_5 \\ x_5 \\ x_5 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \implies \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} = \begin{pmatrix} x_1 \\ x_2 \\ x_2 \\ 0 \\ 0 \end{pmatrix}.$$

Hay dos autovectores LI; ortonormalizándolos ponemos

$$v = a_1 \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} + a_2 \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}.$$

5. Buscamos una primera preimagen w_1 de v , solución general de

$$(A - \lambda I)w_1 = v \implies \begin{pmatrix} w_2 - w_3 + w_5 \\ w_4 + w_5 \\ w_5 \\ w_5 \\ 0 \end{pmatrix} = \begin{pmatrix} a_1 \\ a_2/\sqrt{2} \\ a_2/\sqrt{2} \\ 0 \\ 0 \end{pmatrix}.$$

La tercera y cuarta ecuaciones muestran que no habrá solución a menos que $a_2 = 0$.

6. Tenemos la condición

$$k_1 a_1 + k_2 a_2 = 0 \text{ con } k_1 = 0, k_2 = 1.$$

Por un lado, entonces, substituimos a_1 por k_1 y a_2 por k_2 en v y determinamos

$$v_{10} = k_1 \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} + k_2 \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \\ 0 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \\ 0 \end{pmatrix},$$

que tiene asociado un bloque de Jordan $J_{(2)1}$.

Por otro lado, imponemos la condición $a_2 = 0$ en v y obtenemos

$$v' = a_1 \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} + 0 \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \\ 0 \end{pmatrix} = a_1 \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$

Volvemos entonces al paso anterior.

5. Buscamos una primera preimagen w'_1 de v' , solución general de

$$(A - \lambda I)w'_1 = v' \implies \begin{pmatrix} w_2 - w_3 + w_5 \\ w_4 + w_5 \\ w_5 \\ w_5 \\ 0 \end{pmatrix} = \begin{pmatrix} a_1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \implies w'_1 = \begin{pmatrix} b_1 \\ b_2 \\ b_2 - a_1 \\ 0 \\ 0 \end{pmatrix},$$

que no impone condición alguna sobre a_1 .

Buscamos una segunda preimagen w'_2 de w'_1 , solución general de

$$(A - \lambda I)w'_2 = w'_1 \implies \begin{pmatrix} w_2 - w_3 + w_5 \\ w_4 + w_5 \\ w_5 \\ w_5 \\ 0 \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ b_2 - a_1 \\ 0 \\ 0 \end{pmatrix}.$$

Vemos que la tercera y cuarta ecuaciones imponen $b_2 = a_1$, pero no hay condición impuesta sobre a_1 . Tampoco se impone ninguna condición sobre b_1 , así que asumiremos el riesgo de descartarlo ahora, poniendo $b_1 = 0$. Recordemos que debemos substituir estos valores de b_1 y b_2 también en w'_1 , así que escribimos

$$w'_1 = \begin{pmatrix} 0 \\ a_1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad w'_2 = \begin{pmatrix} c_1 \\ c_2 \\ c_2 \\ a_1 \\ 0 \end{pmatrix}.$$

Buscamos una tercera preimagen w'_3 de w'_2 , solución general de

$$(A - \lambda I)w'_3 = w'_2 \implies \begin{pmatrix} w_2 - w_3 + w_5 \\ w_4 + w_5 \\ w_5 \\ w_5 \\ 0 \end{pmatrix} = \begin{pmatrix} c_1 \\ c_2 \\ c_2 \\ a_1 \\ 0 \end{pmatrix}.$$

Vemos que la tercera y cuarta ecuaciones imponen $c_2 = a_1$, pero no hay condición impuesta sobre a_1 . Tampoco se impone ninguna condición sobre c_1 , así que asumiremos el riesgo de descartarlo ahora, poniendo $c_1 = 0$. Substituyendo estos valores de c_1 y c_2 también en w'_2 , escribimos

$$w'_2 = \begin{pmatrix} 0 \\ a_1 \\ a_1 \\ a_1 \\ 0 \end{pmatrix}, \quad w'_3 = \begin{pmatrix} d_1 \\ d_2 \\ d_2 + a_1 \\ 0 \\ a_1 \end{pmatrix}.$$

Buscamos una cuarta preimagen w'_4 de w'_3 , solución general de

$$(A - \lambda I)w'_4 = w'_3 \implies \begin{pmatrix} w_2 - w_3 + w_5 \\ w_4 + w_5 \\ w_5 \\ w_5 \\ 0 \end{pmatrix} = \begin{pmatrix} d_1 \\ d_2 \\ d_2 + a_1 \\ 0 \\ a_1 \end{pmatrix}.$$

Vemos que la tercera y cuarta ecuaciones imponen $d_2 = -a_1$, y que no hay condición sobre d_1 , así que tomamos $d_1 = 0$ y escribimos

$$w'_3 = \begin{pmatrix} 0 \\ -a_1 \\ 0 \\ 0 \\ a_1 \end{pmatrix}.$$

Vemos también que la quinta ecuación impone la condición $a_1 = 0$.²¹

6. Tenemos la condición

$$k_1 a_1 = 0 \text{ con } k_1 = 1.$$

Por un lado, substituimos a_1 por k_1 en v' , w'_1 , w'_2 y w'_3 , y determinamos

$$v_{20} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad w'_1 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad w'_2 = \begin{pmatrix} 0 \\ 1 \\ 1 \\ 1 \\ 0 \end{pmatrix}, \quad w'_3 = \begin{pmatrix} 0 \\ -1 \\ 0 \\ 0 \\ 1 \end{pmatrix},$$

que tienen asociado un bloque de Jordan $J_{(2)4}$.

Por otro lado, imponemos la condición a_2 en v y obtenemos

$$v' = 0,$$

el subespacio nulo. Esto indica que hemos terminado con este bloque de la descomposición primaria.

Construcción de la transformación de semejanza

7. Procedemos ahora a construir las matrices

$$S = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 1/\sqrt{2} & 0 & 1 & 1 & -1 \\ 1/\sqrt{2} & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}, \quad S^{-1} = \begin{pmatrix} 0 & 0 & \sqrt{2} & -\sqrt{2} & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix},$$

y obtenemos

$$S^{-1}AS = \begin{pmatrix} 2 & 0 & 0 & 0 & 0 \\ 0 & 2 & 1 & 0 & 0 \\ 0 & 0 & 2 & 1 & 0 \\ 0 & 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 0 & 2 \end{pmatrix},$$

que ya está en forma normal de Jordan.

Ejercicio 2.2.5. Muestre que si en el ejemplo precedente, elegimos otros valores distintos para b_1 , c_1 y d_1 (que en el ejemplo fueron igualados a cero), la forma normal de Jordan resulta la misma, y por consiguiente sólo hemos obtenido *otra base cíclica diferente* para el subespacio donde actúa el bloque de Jordan $J_{(2)4}$.

²¹El lector debe convencerse de que esta condición no es “salvable”, es decir no hay valores posibles de d_1 y d_2 que permitan no anular a_1 ; es más, debe convencerse también de que, aún si no hubiéramos descartado b_1 y c_1 , la condición $a_1 = 0$ persistiría.

Ejemplo 2.2.4. Vamos a presentar ahora un ejemplo relativamente completo del uso del método para reducir una matriz a su forma normal de Jordan. De nuevo utilizaremos la misma numeración, para facilitar el seguimiento de cada paso.

Sea la matriz

$$A = \begin{pmatrix} -10 & -6 & 7 & -12 & 2 & 8 & 8 & -5 \\ -9 & -4 & 5 & -11 & 3 & 9 & 7 & -6 \\ 0 & 1 & 3 & 3 & -3 & -3 & -1 & 3 \\ 19 & 12 & -10 & 25 & -7 & -18 & -15 & 13 \\ 12 & 7 & -6 & 15 & -3 & -11 & -9 & 8 \\ -8 & -4 & 5 & -8 & 1 & 6 & 5 & -3 \\ 0 & 0 & -1 & -2 & 2 & 2 & 2 & -1 \\ -14 & -9 & 7 & -18 & 6 & 13 & 11 & -8 \end{pmatrix}.$$

Descomposición primaria

- El polinomio característico es

$$\varphi_A(\lambda) = \lambda^8 - 11\lambda^7 + 52\lambda^6 - 138\lambda^5 + 225\lambda^4 - 231\lambda^3 + 146\lambda^2 - 52\lambda + 8,$$

que factorizado queda

$$\varphi_A(\lambda) = (\lambda - 2)^3(\lambda - 1)^5.$$

Tenemos entonces

$$\begin{aligned} \lambda_1 &= 2, & q_1 &= 3, \\ \lambda_2 &= 1, & q_2 &= 5. \end{aligned}$$

- La solución general de $(A - \lambda_1 I)^3 v = 0$ resulta ser

$$\ker[(A - \lambda_1 I)^3] = \text{span} \left\{ \begin{pmatrix} 1 \\ 2 \\ 1 \\ -1 \\ -1 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \\ -1 \\ -1 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ -1 \\ 1 \\ 1 \\ 2 \\ 0 \\ 0 \\ 1 \end{pmatrix} \right\},$$

de donde podemos leer las columnas v_{11}, v_{12} y v_{13} , que son LI. Análogamente, la solución general de $(A - \lambda_2 I)^5 v = 0$ resulta ser

$$\ker[(A - \lambda_2 I)^5] = \text{span} \left\{ \begin{pmatrix} -4 \\ -1 \\ -4 \\ 2 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 3 \\ 2 \\ 0 \\ 2 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 0 \\ 2 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \\ 2 \\ 0 \\ 0 \\ 0 \\ 2 \\ 0 \end{pmatrix}, \begin{pmatrix} -4 \\ 1 \\ -4 \\ 0 \\ 0 \\ 0 \\ 0 \\ 2 \end{pmatrix} \right\},$$

de donde podemos leer las columnas $v_{21}, v_{22}, v_{23}, v_{24}$ y v_{25} , que son LI entre sí y de las anteriores.

3. Formamos la matriz

$$T = \begin{pmatrix} 1 & 1 & 0 & -4 & 0 & 2 & 2 & -4 \\ 2 & 1 & -1 & -1 & 3 & 0 & 1 & 1 \\ 1 & 0 & 1 & -4 & 2 & 2 & 2 & -4 \\ -1 & -1 & 1 & 2 & 0 & 0 & 0 & 0 \\ -1 & -1 & 2 & 0 & 2 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 2 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 2 \end{pmatrix}$$

con las columnas encontradas en el paso 2, determinamos su inversa

$$T^{-1} = \begin{pmatrix} -7 & -4 & 4 & -8 & 2 & 6 & 5 & -4 \\ -2 & -2 & 0 & -5 & 3 & 4 & 3 & -3 \\ -6 & -4 & 3 & -8 & 3 & 6 & 5 & -4 \\ -\frac{3}{2} & -1 & \frac{1}{2} & -2 & 1 & 2 & \frac{3}{2} & -\frac{3}{2} \\ \frac{3}{2} & 1 & -1 & \frac{3}{2} & 0 & -1 & -1 & \frac{1}{2} \\ 7 & 4 & -4 & 8 & -2 & -5 & -5 & 4 \\ 1 & 1 & 0 & \frac{5}{2} & -\frac{3}{2} & -2 & -1 & \frac{3}{2} \\ 3 & 2 & -\frac{3}{2} & 4 & -\frac{3}{2} & -3 & -\frac{5}{2} & \frac{5}{2} \end{pmatrix},$$

y aplicamos a A la transformación de semejanza así definida, obteniendo

$$A' = T^{-1}AT = \begin{pmatrix} 2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & -\frac{1}{2} & \frac{3}{2} & \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2 \\ 0 & 0 & 0 & 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 0 & \frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} & \frac{3}{2} \end{pmatrix}.$$

Reducción a forma normal del bloque A_1 . Trabajamos primero con

$$A_1 = \begin{pmatrix} 2 & 0 & 0 \\ 1 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}, \quad (A_1 - \lambda_1 I) = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

4. La solución general de $(A_1 - \lambda_1 I)v = 0$ resulta ser

$$\ker(A_1 - \lambda_1 I) = \text{span} \left\{ \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\}.$$

Estas columnas ya son ortonormales; ponemos entonces

$$\mathbf{v} = a_1 \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + a_2 \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

5. Buscamos una primera preimagen \mathbf{w}_1 , solución general de

$$(\mathbf{A}_1 - \lambda_1 \mathbf{I})\mathbf{w}_1 = \mathbf{v}.$$

La tercera ecuación de este sistema muestra que no habrá solución a menos que $k_1 a_1 + k_2 a_2 = 0$, con $k_1 = 0$ y $k_2 = 1$, es decir, a menos que $a_2 = 0$.

6. Substituyendo a_1 por k_1 y a_2 por k_2 en \mathbf{v} , determinamos entonces un primer autovector

$$\mathbf{v}_{10} = k_1 \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + k_2 \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix},$$

asociado a un bloque de Jordan $J_{(2)1}$, y que no tiene ninguna preimagen.

Para seguir debemos eliminar \mathbf{v}_{10} de \mathbf{v} . Para ello substituimos $a_2 = 0$ en \mathbf{v} , obteniendo

$$\mathbf{v}' = a_1 \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}.$$

5. Buscamos su primera preimagen \mathbf{w}'_1 , solución general de

$$(\mathbf{A}_1 - \lambda_1 \mathbf{I})\mathbf{w}'_1 = \mathbf{v}'.$$

Obtenemos $\mathbf{w}'_1 = (a_1, b_1, b_2)^T$; un vistazo a la iteración siguiente muestra que no se impone condición sobre los nuevos parámetros b_1 y b_2 , así que los descartamos y escribimos

$$\mathbf{w}'_1 = \begin{pmatrix} a_1 \\ 0 \\ 0 \end{pmatrix}.$$

Buscamos ahora una segunda preimagen \mathbf{w}'_2 , solución general de

$$(\mathbf{A}_1 - \lambda_1 \mathbf{I})\mathbf{w}'_2 = \mathbf{w}'_1.$$

La segunda ecuación de este sistema muestra que no habrá solución a menos que $k_1 a_1 = 0$, con $k_1 = 1$, es decir a menos que $a_1 = 0$.

6. Substituyendo a_1 por k_1 en \mathbf{v}' y su preimagen \mathbf{w}'_1 , determinamos un segundo autovector \mathbf{v}_{20} y su preimagen \mathbf{v}_{21} ,

$$\mathbf{v}_{20} = k_1 \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad \mathbf{v}_{21} = \begin{pmatrix} k_1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix},$$

asociados a un bloque de Jordan $J_{(2)2}$, y que no tienen otra preimagen.

Para seguir debemos eliminar v_{20} de v' . Para ello substituimos $a_1 = 0$ en v' , obteniendo el vector nulo, lo que indica que hemos terminado con el bloque asociado a λ_1 .

Reducción a forma normal del bloque A_2 . Pasamos ahora a trabajar con

$$A_2 = \begin{pmatrix} 1 & 1 & 0 & 0 & 1 \\ -\frac{1}{2} & \frac{3}{2} & \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} \\ 0 & 0 & 0 & 0 & 2 \\ 0 & 1 & 0 & 1 & 1 \\ \frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} & \frac{3}{2} \end{pmatrix}, \quad (A_2 - \lambda_2 I) = \begin{pmatrix} 0 & 1 & 0 & 0 & 1 \\ -\frac{1}{2} & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} \\ 0 & 0 & -1 & 0 & 2 \\ 0 & 1 & 0 & 0 & 1 \\ \frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \end{pmatrix}.$$

4. La solución general de $(A_2 - \lambda_2 I)v = 0$ resulta ser

$$\ker(A_2 - \lambda_2 I) = \text{span} \left\{ \begin{pmatrix} \frac{1}{2}\sqrt{2} \\ -\frac{1}{6}\sqrt{6} \\ \frac{1}{3}\sqrt{6} \\ \frac{1}{2}\sqrt{2} \\ \frac{1}{6}\sqrt{6} \end{pmatrix}, \begin{pmatrix} \frac{1}{2}\sqrt{2} \\ \frac{1}{6}\sqrt{6} \\ -\frac{1}{3}\sqrt{6} \\ \frac{1}{2}\sqrt{2} \\ -\frac{1}{6}\sqrt{6} \end{pmatrix} \right\}.$$

Estas columnas ya son ortonormales; ponemos entonces

$$v = a_1 \begin{pmatrix} \frac{1}{2}\sqrt{2} \\ -\frac{1}{6}\sqrt{6} \\ \frac{1}{3}\sqrt{6} \\ \frac{1}{2}\sqrt{2} \\ \frac{1}{6}\sqrt{6} \end{pmatrix} + a_2 \begin{pmatrix} \frac{1}{2}\sqrt{2} \\ \frac{1}{6}\sqrt{6} \\ -\frac{1}{3}\sqrt{6} \\ \frac{1}{2}\sqrt{2} \\ -\frac{1}{6}\sqrt{6} \end{pmatrix}.$$

5. Buscamos su primera preimagen w_1 , solución general de

$$(A_1 - \lambda_1 I)w_1 = v.$$

Obtenemos $w_1 = (b_1 + b_2 + \frac{1}{2}\sqrt{2}(a_1 + a_2), -b_1 + b_2 + \frac{1}{2}\sqrt{2}(a_1 + a_2), 2b_1 - 2b_2 + \frac{1}{3}\sqrt{6}(-a_1 + a_2), b_1 + b_2, b_1 - b_2)^T$; un vistazo a la iteración siguiente muestra que no se impondrá condición sobre los nuevos parámetros b_1 y b_2 , así que los descartamos y escribimos

$$w_1 = \begin{pmatrix} \frac{1}{2}\sqrt{2}(a_1 + a_2) \\ \frac{1}{2}\sqrt{2}(a_1 + a_2) \\ \frac{1}{3}\sqrt{6}(-a_1 + a_2) \\ 0 \\ 0 \end{pmatrix}.$$

Buscamos ahora una segunda preimagen w_2 , solución general de

$$(A_1 - \lambda_1 I)w_2 = w_1.$$

La primera y cuarta ecuaciones de este sistema muestran que no habrá solución a menos que $k_1 a_1 + k_2 a_2 = 0$, con $k_1 = 1$ y $k_2 = 1$, es decir, a menos que $a_2 = -a_1$ (la segunda y quinta ecuaciones muestran lo mismo).

6. Substituyendo a_1 por k_1 y a_2 por k_2 en \mathbf{v} y su preimagen \mathbf{w}_1 , determinamos entonces un primer autovector \mathbf{v}_{10} y su preimagen \mathbf{v}_{11} ,

$$\mathbf{v}_{10} = k_1 \begin{pmatrix} \frac{1}{2}\sqrt{2} \\ -\frac{1}{6}\sqrt{6} \\ \frac{1}{3}\sqrt{6} \\ \frac{1}{2}\sqrt{2} \\ \frac{1}{6}\sqrt{6} \end{pmatrix} + k_2 \begin{pmatrix} \frac{1}{2}\sqrt{2} \\ \frac{1}{6}\sqrt{6} \\ -\frac{1}{3}\sqrt{6} \\ \frac{1}{2}\sqrt{2} \\ -\frac{1}{6}\sqrt{6} \end{pmatrix} = \begin{pmatrix} \sqrt{2} \\ 0 \\ 0 \\ \sqrt{2} \\ 0 \end{pmatrix}, \quad \mathbf{v}_{11} = \begin{pmatrix} \frac{1}{2}\sqrt{2}(k_1 + k_2) \\ \frac{1}{2}\sqrt{2}(k_1 + k_2) \\ \frac{1}{3}\sqrt{6}(-k_1 + k_2) \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} \sqrt{2} \\ \sqrt{2} \\ 0 \\ 0 \\ 0 \end{pmatrix},$$

asociados a un bloque de Jordan $J_{(1)2}$, y que no tienen otra preimagen.

Para seguir debemos eliminar \mathbf{v}_{10} de \mathbf{v} . Para ello substituimos $a_2 = -a_1$ en \mathbf{v} , obteniendo

$$\mathbf{v}' = a_1 \begin{pmatrix} \frac{1}{2}\sqrt{2} \\ -\frac{1}{6}\sqrt{6} \\ \frac{1}{3}\sqrt{6} \\ \frac{1}{2}\sqrt{2} \\ \frac{1}{6}\sqrt{6} \end{pmatrix} - a_1 \begin{pmatrix} \frac{1}{2}\sqrt{2} \\ \frac{1}{6}\sqrt{6} \\ -\frac{1}{3}\sqrt{6} \\ \frac{1}{2}\sqrt{2} \\ -\frac{1}{6}\sqrt{6} \end{pmatrix} = \begin{pmatrix} 0 \\ -\frac{1}{3}\sqrt{6}a_1 \\ \frac{2}{3}\sqrt{6}a_1 \\ 0 \\ \frac{1}{3}\sqrt{6}a_1 \end{pmatrix}.$$

Haciendo la misma sustitución en \mathbf{w}_1 , obtenemos la primera preimagen de \mathbf{v}' ,

$$\mathbf{w}'_1 = \begin{pmatrix} 0 \\ 0 \\ -\frac{2}{3}\sqrt{6}a_1 \\ 0 \\ 0 \end{pmatrix}.$$

5. Buscamos ahora una preimagen \mathbf{w}'_2 , solución general de

$$(\mathbf{A}_1 - \lambda_1 \mathbf{I})\mathbf{w}'_2 = \mathbf{w}'_1.$$

Obtenemos $\mathbf{w}'_2 = (b_1 + b_2 + \frac{2}{3}\sqrt{6}a_1, -b_1 + b_2, 2b_1 - 2b_2 + \frac{2}{3}\sqrt{6}a_1, b_1 + b_2, b_1 - b_2)^T$; un vistazo a la iteración siguiente muestra que no se impondrá condición sobre los nuevos parámetros b_1 y b_2 , así que los descartamos y escribimos

$$\mathbf{w}'_2 = \begin{pmatrix} \frac{2}{3}\sqrt{6}a_1 \\ 0 \\ \frac{2}{3}\sqrt{6}a_1 \\ 0 \\ 0 \end{pmatrix}.$$

Buscamos la siguiente preimagen, solución general de

$$(\mathbf{A}_1 - \lambda_1 \mathbf{I})\mathbf{w}'_3 = \mathbf{w}'_2.$$

La primera y cuarta ecuaciones de este sistema muestran que no habrá solución a menos que $k_1 a_1 = 0$, con $k_1 = 1$, es decir, a menos que $a_1 = 0$.

6. Substituyendo a_1 por $k_1 = 1$ en v' y sus preimágenes w'_1 y w'_2 , respectivamente, determinamos un segundo autovector v_{20} y sus preimágenes v_{21} y v_{22} ,

$$v_{20} = \begin{pmatrix} 0 \\ -\frac{1}{3}\sqrt{6} \\ \frac{2}{3}\sqrt{6} \\ 0 \\ \frac{1}{3}\sqrt{6} \end{pmatrix}, v_{21} = \begin{pmatrix} 0 \\ 0 \\ -\frac{2}{3}\sqrt{6} \\ 0 \\ 0 \end{pmatrix}, v_{22} = \begin{pmatrix} \frac{2}{3}\sqrt{6} \\ 0 \\ \frac{2}{3}\sqrt{6} \\ 0 \\ 0 \end{pmatrix},$$

asociados a un bloque de Jordan $J_{(1)3}$, y que ya no tienen otra preimagen.

Para seguir debemos eliminar v_{20} de v' . Para ello substituimos $a_1 = 0$ en v' , obteniendo el vector nulo, lo que indica que hemos terminado con el bloque asociado a λ_2 .

Construcción de la transformación de semejanza

7. Recolectamos los autovectores y sus preimágenes asociados a λ_1 en la matriz S_1 y calculamos su inversa, obteniendo

$$S_1 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}, S_1^{-1} = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}.$$

Aplicando la transformación de semejanza así definida a A_1 , queda

$$S_1^{-1}A_1S_1 = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix}.$$

También recolectamos los autovectores y sus preimágenes asociados a λ_2 en la matriz S_2 y calculamos su inversa, obteniendo

$$S_2 = \begin{pmatrix} \sqrt{2} & \sqrt{2} & 0 & 0 & \frac{2}{3}\sqrt{6} \\ 0 & \sqrt{2} & -\frac{1}{3}\sqrt{6} & 0 & 0 \\ 0 & 0 & \frac{2}{3}\sqrt{6} & -\frac{2}{3}\sqrt{6} & \frac{2}{3}\sqrt{6} \\ \sqrt{2} & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{3}\sqrt{6} & 0 & 0 \end{pmatrix},$$

$$S_2^{-1} = \begin{pmatrix} 0 & 0 & 0 & \frac{1}{2}\sqrt{2} & 0 \\ 0 & \frac{1}{2}\sqrt{2} & 0 & 0 & \frac{1}{2}\sqrt{2} \\ 0 & 0 & 0 & 0 & \frac{1}{2}\sqrt{6} \\ \frac{1}{4}\sqrt{6} & -\frac{1}{4}\sqrt{6} & -\frac{1}{4}\sqrt{6} & -\frac{1}{4}\sqrt{6} & \frac{1}{4}\sqrt{6} \\ \frac{1}{4}\sqrt{6} & -\frac{1}{4}\sqrt{6} & 0 & -\frac{1}{4}\sqrt{6} & -\frac{1}{4}\sqrt{6} \end{pmatrix}.$$

Aplicando la transformación de semejanza así definida a A_2 , queda

$$S_2^{-1}A_2S_2 = \begin{pmatrix} 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

8. Finalmente, recolectamos las matrices S_1 y S_2 y sus inversas en las matrices

$$S = \begin{pmatrix} S_1 & 0 \\ 0 & S_2 \end{pmatrix}, \quad S^{-1} = \begin{pmatrix} S_1^{-1} & 0 \\ 0 & S_2^{-1} \end{pmatrix},$$

y aplicamos la transformación de semejanza así definida a A' , quedando

$$A'' = S^{-1}A'S = S^{-1}T^{-1}ATS = \begin{pmatrix} 2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 2 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix},$$

que está en su forma normal de Jordan. Esta última transformación es sólo formal, ya que en realidad nos limitamos a colocar los bloques $S_1^{-1}A_1S_1$ y $S_2^{-1}A_2S_2$ en la diagonal de A'' , y completar el resto con ceros.

Para finalizar, debemos destacar algunos hechos importantes que probablemente no hayan quedado explicitados.

1. Cuando diagonalizamos una matriz diagonalizable, estamos habituados a que los autovectores pueden renormalizarse arbitrariamente. Aquí eso sigue siendo cierto respecto de los autovectores, pero si multiplicamos un dado autovector por una constante, deberemos multiplicar todas sus preimágenes por *la misma* constante, ya que estas últimas son soluciones de sucesivos sistemas inhomogéneos, el primero de los cuales tiene al autovector como inhomogeneidad.
2. También cuando diagonalizamos una matriz diagonalizable, estamos habituados a que los autovectores correspondientes a un mismo autovalor pueden recombinarse (linealmente) en forma arbitraria. Para una matriz no diagonalizable esto en general *es falso*, ya que cada autovector tiene asociado un bloque de Jordan en general distinto. Sólo podremos recombinar libremente autovectores al mismo autovalor y que tengan asociados bloques de Jordan *de igual tamaño*.
3. Finalmente, también de nuestra experiencia con matrices diagonalizables puede habernos quedado la idea de que dos matrices serán semejantes si tienen el mismo espectro (o, lo que es igual, los mismos invariantes algebraicos). Para matrices defectivas (no diagonalizables) ello *no es cierto*: debe especificarse, además, la estructura de bloques de Jordan, es decir, cuántos y de qué tamaños hay asociados a cada autovalor.

2.2.5. Método abreviado para hallar la forma de Jordan

Cuando diagonalizamos una matriz diagonalizable, podemos determinar inmediatamente la forma final que obtendremos una vez determinados sus autovalores $\lambda_1, \dots, \lambda_r$ y sus multiplicidades q_1, \dots, q_r , que será²²

$$\text{diag}(\underbrace{\lambda_1, \dots, \lambda_1}_{q_1 \text{ veces}}, \underbrace{\lambda_2, \dots, \lambda_2}_{q_2 \text{ veces}}, \dots, \underbrace{\lambda_r, \dots, \lambda_r}_{q_r \text{ veces}}).$$

Cabe preguntarse, entonces, si existirá alguna forma similarmente breve de determinar, para una matriz defectiva, el número y tamaño de los bloques de Jordan sin construir explícitamente la transformación que la reduce a la forma normal. Afortunadamente, la respuesta no sólo es afirmativa, sino que existen varios procedimientos para conseguirlo; expondremos aquí uno que, en opinión del autor, es particularmente conciso y simple.

El procedimiento en cuestión parte de notar que el número y tamaño de los bloques de Jordan asociados a un autovalor λ_i de multiplicidad q_i es *independiente de la base*; por lo tanto podemos suponer que ya estamos en la base donde la matriz adopta su forma de Jordan, aunque *no sepamos cuál es*, y trabajar allí. Notemos primero que, de la Sección 2.2.3, es inmediato que el número de bloques de Jordan asociados a λ_i es igual a la cantidad de autovectores LI asociados al ese autovalor; esta cantidad s_i se conoce como la *multiplicidad geométrica* de λ_i , y es simplemente

$$s_i = \dim \ker(A - \lambda_i I),$$

que podemos determinar de inmediato contando los parámetros libres en la solución general de $(A - \lambda_i I)v = 0$. Por otro lado, la *multiplicidad algebraica* de λ_i , que podemos poner como

$$q_i = \max_{p \in \mathbb{N}} \{\dim \ker(A - \lambda_i I)^p\},$$

es la suma de los tamaños de los bloques de Jordan, igual a la dimensión del subespacio característico asociado a λ_i , y será en general mayor que s_i (si fuera igual, todos los bloques tendrían tamaño 1, y el bloque A_i de la matriz A sería diagonal). Para bloques de la descomposición primaria (es decir, subespacios característicos) de tamaños 1, 2 y 3, esta información es suficiente para determinar el número y tamaño de los posibles bloques de Jordan.

Para subespacios característicos de dimensión 4 o mayor, la clave es la estructura (2.21) del correspondiente bloque A_i de la descomposición primaria. Notemos primero que (2.21) es diagonal por bloques (de Jordan), y que

$$(A_i - \lambda_i I) = \begin{pmatrix} N_{t_{i1}} & 0 & \cdots & 0 \\ 0 & N_{t_{i2}} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & N_{t_{ir_i}} \end{pmatrix},$$

²²Esta forma es única, a menos de permutaciones de la base.

donde cada uno de los r_i sub-bloques $N_{t_{ij}}$ es un nilpotente de tamaño e índice t_{ij} . Sucesivas potencias tendrán entonces la forma

$$(A_i - \lambda_i I)^p = \begin{pmatrix} N_{t_{i1}}^p & 0 & \cdots & 0 \\ 0 & N_{t_{i2}}^p & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & N_{t_{ir_i}}^p \end{pmatrix}$$

y, a medida que p crezca, $\dim \ker(A_i - \lambda_i I)^p$ irá creciendo, cada vez en un entero igual a la cantidad de nilpotentes que aún no se hayan anulado: todo nilpotente $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ de tamaño 2 ($\dim \ker = 1$), elevado al cuadrado será $\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$ ($\dim \ker = 2$), y de allí en adelante será un bloque 2×2 de ceros que ya no cambiará; un nilpotente $\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$ ($\dim \ker = 1$), al cuadrado será $\begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$ ($\dim \ker = 2$), y del cubo en adelante será un bloque nulo ($\dim \ker = 3$); etc. De aquí es fácil ver que el número de bloques de Jordan asociados a λ_i de tamaño $\geq p$ será igual a $\dim \ker(A - \lambda_i I)^p - \dim \ker(A - \lambda_i I)^{p-1}$, y en consecuencia el número de bloques de Jordan de tamaño p será

$$n_p = 2 \dim \ker(A - \lambda_i I)^p - \dim \ker(A - \lambda_i I)^{p+1} - \dim \ker(A - \lambda_i I)^{p-1}, \quad (2.23)$$

donde $0 \leq p \leq q_i$ y se sobreentiende que $\dim \ker(A - \lambda_i I)^0 = 0$ y que $\dim \ker(A - \lambda_i I)^p = q_i$ si $p \geq q_i$.

Ejemplo 2.2.5. Sea la matriz

$$A = \begin{pmatrix} 2 & 1 & -1 & 0 & 1 \\ 0 & 2 & 0 & 1 & 1 \\ 0 & 0 & 2 & 0 & 1 \\ 0 & 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 0 & 2 \end{pmatrix}.$$

Es inmediato que tenemos un único autovalor $\lambda = 2$ con multiplicidad algebraica 5. Colocamos en una tabla las potencias p a las que elevamos $(A - \lambda I)$, la correspondiente dimensión de $\ker(A - \lambda I)^p$, y el número n_p de bloques de Jordan de tamaño p calculado según (2.23):

p	$\dim \ker(A - \lambda I)^p$	n_p
0	0	-
1	2	1
2	3	0
3	4	0
4	5	1
5	5	0
6	5	-

Esto nos muestra que la forma de Jordan de A contiene un bloque $J_{(2)1}$ de tamaño 1, y un bloque $J_{(2)4}$ de tamaño 4.

Ejercicio 2.2.6. Complete los detalles de los cálculos precedentes.

Ejercicio 2.2.7. Sea la matriz

$$A = \begin{pmatrix} -10 & -6 & 7 & -12 & 2 & 8 & 8 & -5 \\ -9 & -4 & 5 & -11 & 3 & 9 & 7 & -6 \\ 0 & 1 & 3 & 3 & -3 & -3 & -1 & 3 \\ 19 & 12 & -10 & 25 & -7 & -18 & -15 & 13 \\ 12 & 7 & -6 & 15 & -3 & -11 & -9 & 8 \\ -8 & -4 & 5 & -8 & 1 & 6 & 5 & -3 \\ 0 & 0 & -1 & -2 & 2 & 2 & 2 & -1 \\ -14 & -9 & 7 & -18 & 6 & 13 & 11 & -8 \end{pmatrix}.$$

Muestre que sus autovalores son $\lambda_1 = 2$ con $q_1 = 3$ y $\lambda_2 = 1$ con $q_2 = 5$, y encuentre su forma de Jordan.

Capítulo 3

Tensores

3.1. Tensores

Los tensores aparecen frecuentemente en Física, representando por ejemplo entes como momento cuadrupolar o gradiente del campo eléctrico, momentos de inercia, propiedades elásticas de sólidos, propiedades de transporte de fluidos, etc. En esta sección unificaremos algunos conceptos ya introducidos y daremos una formulación general de los mismos. En el camino, veremos que ya nos hemos encontrado, sin saberlo, con diversas instancias de tensores, que ahora reconoceremos como tales.

3.1.1. Definición de tensor

Ya nos hemos encontrado con el concepto de una 1-forma (lineal) $\tilde{\omega}$, que es una aplicación de un espacio vectorial V al cuerpo escalar \mathbb{C} ,

$$\tilde{\omega} : V \rightarrow \mathbb{C}, \quad V \ni \vec{u} \xrightarrow{\tilde{\omega}} \tilde{\omega}(\vec{u}) \in \mathbb{C}.$$

Menos familiar es el hecho de que podemos pensar a un vector \vec{u} como una aplicación lineal del espacio dual V^* al cuerpo escalar: si definimos

$$\vec{u}(\tilde{\omega}) := \tilde{\omega}(\vec{u}),$$

vemos que efectivamente

$$\vec{u} : V^* \rightarrow \mathbb{C}, \quad V \ni \vec{u} \xrightarrow{\tilde{\omega}} \vec{u}(\tilde{\omega}) \in \mathbb{C}.$$

Otro ente en cierto modo semejante que ya hemos encontrado es el producto interno, que definimos como una 2-forma (lineal) que lleva dos vectores de V al cuerpo escalar,

$$\Phi : V \times V \rightarrow \mathbb{C}, \quad V \times V \ni (\vec{u}, \vec{v}) \xrightarrow{\Phi} \Phi(\vec{u}, \vec{v}) \in \mathbb{C}.$$

Aquí $V \times V$ representa el *producto Cartesiano* del espacio V consigo mismo, es decir, el EV formado por todos los pares ordenados de vectores (\vec{u}, \vec{v}) , con $\vec{u}, \vec{v} \in V$.¹

Aún otro ente con características similares es un operador lineal \mathcal{A} de V en V . A primera vista esto parece contradictorio, ya que $\mathcal{A} : V \rightarrow V$, es decir no va al cuerpo escalar. Sin embargo, dado $\vec{u} \in V$, también $\mathcal{A}\vec{u} \in V$, de modo que podemos aplicarlo a una 1-forma $\tilde{\omega} \in V^*$ para obtener un escalar:

$$(\mathcal{A}\vec{u})(\tilde{\omega}) \equiv \tilde{\omega}(\mathcal{A}\vec{u}) \in \mathbb{C}.$$

Podemos entonces pensar en el operador \mathcal{A} como una aplicación lineal que toma un vector \vec{u} de V y una 1-forma $\tilde{\omega}$ de V^* y devuelve un escalar. Haciendo un ligero abuso de notación (aprovechándonos de la linealidad) desecharemos los paréntesis, y escribiremos

$$\mathcal{A} : V^* \times V \rightarrow \mathbb{C}, \quad V^* \times V \ni (\tilde{\omega}, \vec{u}) \xrightarrow{\mathcal{A}} \tilde{\omega}\mathcal{A}\vec{u} \in \mathbb{C}.$$

La notación $\tilde{\omega}\mathcal{A}\vec{u}$ es altamente sugestiva, y nos permite dar *tres* interpretaciones diferentes de la operación que indica: (i) la aplicación \mathcal{A} toma un vector \vec{u} y una forma $\tilde{\omega}$, y devuelve un escalar ($\mathcal{A} : V^* \times V \rightarrow \mathbb{C}$); (ii) el operador \mathcal{A} actúa “hacia adelante” sobre el vector \vec{u} , devolviendo otro vector sobre el que actúa $\tilde{\omega}$ para dar un escalar ($\mathcal{A} : V \rightarrow V$); y (iii) el operador \mathcal{A} actúa “hacia atrás” sobre la forma $\tilde{\omega}$, devolviendo otra forma sobre la que actúa \vec{u} para dar un escalar ($\mathcal{A} : V^* \rightarrow V^*$).

La generalización más o menos obvia de estos casos particulares es una aplicación lineal que tome un cierto número de formas de V^* y un cierto número de vectores de V y devuelva un escalar. Para formalizarla, definiremos primero

$$\Pi_r^s := \underbrace{V^* \times V^* \times \cdots \times V^*}_{r \text{ veces}} \times \underbrace{V \times V \times \cdots \times V}_{s \text{ veces}}. \quad (3.1)$$

Definiremos ahora un *tensor* T de tipo (r, s) como una aplicación

$$T : \Pi_r^s \rightarrow \mathbb{C} \quad (3.2)$$

lineal en todos sus argumentos, y escribiremos

$$T(\tilde{\omega}^1, \dots, \tilde{\omega}^r; \vec{u}_1, \dots, \vec{u}_s) = t \in \mathbb{C}, \quad \tilde{\omega}^1, \dots, \tilde{\omega}^r \in V^*, \quad \vec{u}_1, \dots, \vec{u}_s \in V.$$

El *rango* de un tensor se define como el número total de argumentos que toma la aplicación, es decir $r + s$.

Revisando los casos particulares del principio, vemos que:

- las 1-formas son tensores de tipo $(0, 1)$, rango 1;

¹En espacios de dimensión *finita*, el producto Cartesiano $V \times W$ y la suma directa $V \oplus W$ de dos espacios vectoriales V y W son equivalentes, y podemos usar indistintamente uno o la otra. Sin embargo, esto no es válido en espacios de dimensión infinita; la noción que se generaliza correctamente a este caso es la de producto Cartesiano, que es la que usaremos aquí.

- los vectores son tensores de tipo $(1, 0)$, rango 1;
- el producto interno es un tensor de tipo $(0, 2)$, rango 2;
- los operadores lineales son tensores de tipo $(1, 1)$, rango 2.

3.1.2. Espacio tensorial

Volviendo al concepto de 1-forma como aplicación lineal de V en \mathbb{C} , recordemos que lo usamos para definir el espacio dual V^* como el EV de todas las 1-formas lineales sobre V , es decir

$$V^* = \{\bar{\omega} : V \rightarrow \mathbb{C}, \bar{\omega} \text{ lineal}\}.$$

Para completar esta definición, debemos agregar al conjunto de las 1-formas las operaciones de *suma* de 1-formas y de *producto por un escalar*, con las propiedades usuales.

De modo análogo, y pensando ahora en los vectores como aplicaciones lineales de V^* en \mathbb{C} , podemos definir V como

$$V = \{\vec{u} : V^* \rightarrow \mathbb{C}, \vec{u} \text{ lineal}\},$$

agregando las operaciones de suma de vectores y de producto de un vector por un escalar, también con las propiedades usuales.

Asimismo, podemos definir el espacio consistente en todos los operadores lineales \mathcal{A} de V en V , como el conjunto

$$\{\mathcal{A} : V^* \times V \rightarrow \mathbb{C}, \mathcal{A} \text{ lineal}\},$$

agregando las operaciones de suma entre operadores lineales y de producto de un operador lineal por un escalar, y este será un espacio vectorial.

Para generalizar esta idea, debemos definir las siguientes operaciones:

- **Suma de tensores:** Dados dos tensores T y S , ambos de tipo (r, s) , su *suma* es el tensor

$$\begin{aligned} (T + S)(\bar{\omega}^1, \dots, \bar{\omega}^r; \vec{u}_1, \dots, \vec{u}_s) \\ := T(\bar{\omega}^1, \dots, \bar{\omega}^r; \vec{u}_1, \dots, \vec{u}_s) + S(\bar{\omega}^1, \dots, \bar{\omega}^r; \vec{u}_1, \dots, \vec{u}_s), \end{aligned}$$

también de tipo (r, s) .

- **Producto por un escalar:** Dado un tensor T de tipo (r, s) , su *producto por un escalar* $\lambda \in \mathbb{C}$ es el tensor

$$(\lambda T)(\bar{\omega}^1, \dots, \bar{\omega}^r; \vec{u}_1, \dots, \vec{u}_s) := \lambda T(\bar{\omega}^1, \dots, \bar{\omega}^r; \vec{u}_1, \dots, \vec{u}_s),$$

también de tipo (r, s) .

Estas operaciones tendrán las mismas propiedades (distributividad, etc.) que las definidas para vectores. Dadas estas operaciones, definimos el *espacio tensorial* V_r^s como

$$V_r^s := \{T : \Pi_r^s \rightarrow \mathbb{C}, T \text{ lineal}\}, \quad (3.3)$$

y es un espacio vectorial.

3.1.3. Producto tensorial

Un concepto útil que nos permite definir tensores sobre espacios “más grandes” a partir de otros tensores sobre espacios “más chicos” es el de producto tensorial. Dados dos tensores $T_1 \in V_{r_1}^{s_1}$ y $T_2 \in V_{r_2}^{s_2}$ en espacios tensoriales $V_{r_1}^{s_1}$ y $V_{r_2}^{s_2}$ (posiblemente distintos), su *producto tensorial* T es el nuevo tensor

$$T := T_1 \otimes T_2 \in V_{r_1+r_2}^{s_1+s_2} \quad (3.4)$$

definido por

$$\begin{aligned} T(\bar{\omega}^1, \dots, \bar{\omega}^{r_1}, \bar{\tau}^1, \dots, \bar{\tau}^{r_2}; \vec{u}_1, \dots, \vec{u}_{s_1}, \vec{v}_1, \dots, \vec{v}_{s_2}) \\ = T_1(\bar{\omega}^1, \dots, \bar{\omega}^{r_1}; \vec{u}_1, \dots, \vec{u}_{s_1}) T_2(\bar{\tau}^1, \dots, \bar{\tau}^{r_2}; \vec{v}_1, \dots, \vec{v}_{s_2}), \end{aligned} \quad (3.5)$$

donde $\bar{\omega}^1, \dots, \bar{\omega}^{r_1}, \bar{\tau}^1, \dots, \bar{\tau}^{r_2} \in V^*$ y $\vec{u}_1, \dots, \vec{u}_{s_1}, \vec{v}_1, \dots, \vec{v}_{s_2} \in V$.

Ejemplo 3.1.1. Dados dos vectores $\vec{u}, \vec{v} \in V$, definamos $T := \vec{u} \otimes \vec{v}$; entonces

$$T(\bar{\omega}, \bar{\tau}) = \vec{u}(\bar{\omega}) \vec{v}(\bar{\tau}) \quad \forall \bar{\omega}, \bar{\tau} \in V^*,$$

y vemos que T es un tensor de tipo $(2, 0)$.

Ejemplo 3.1.2. Dados un vector $\vec{u} \in V$ y una forma $\bar{\omega} \in V^*$, definamos $T := \vec{u} \otimes \bar{\omega}$; entonces

$$T(\bar{\tau}, \vec{v}) = \vec{u}(\bar{\tau}) \bar{\omega}(\vec{v}) \quad \forall \bar{\tau} \in V^*, \vec{v} \in V,$$

y vemos que T es un tensor de tipo $(1, 1)$.

3.1.4. Base y componentes

Consideremos un EV V con una base $\{\vec{e}_i\}_{i=1}^n$, y su dual V^* con la base dual $\{\bar{e}^j\}_{j=1}^n$. Recordemos cómo definimos las componentes ω_i de una forma $\bar{\omega} \in V^*$ por $\omega_i = \bar{\omega}(\vec{e}_i)$. Análogamente podríamos definir las componentes x^j de un vector $\vec{x} \in V$ por $x^j = \vec{x}(\bar{e}^j)$. También podríamos definir las componentes de un operador lineal \mathcal{A} como $A^j{}_i = \bar{e}^j \mathcal{A} \vec{e}_i$, o las componentes Φ_{ij} del producto interno Φ como $\Phi_{ij} = \Phi(\vec{e}_i, \vec{e}_j)$ (estas últimas serían los elementos de la métrica g_{ij}). Verifaremos que todas estas definiciones son consistentes con las propiedades que hemos ido hallando para las componentes de los diversos entes.

Vamos a generalizar entonces esa idea, definiendo las *componentes* $S_{j_1 \dots j_s}^{i_1 \dots i_r}$ de un tensor $S \in V_r^s$ en las bases $\{\vec{e}_i\}_{i=1}^n$ y $\{\bar{e}^j\}_{j=1}^n$, como

$$S_{j_1 \dots j_s}^{i_1 \dots i_r} := S(\bar{e}^{i_1}, \dots, \bar{e}^{i_r}; \vec{e}_{j_1}, \dots, \vec{e}_{j_s}). \quad (3.6)$$

De manera análoga a lo que ocurría para las componentes de vectores, operadores, etc., las propiedades de linealidad de los tensores harán que las componentes de una suma sean la suma de las componentes, y las componentes del producto por un escalar sean el producto del escalar por las componentes.

Consideremos ahora formas $\bar{\omega}^1 = \omega_{i_1} \bar{e}^{i_1}, \dots, \bar{\omega}^r = \omega_{i_r} \bar{e}^{i_r}$ en V^* , y vectores $\vec{u}_1 = u^{j_1} \vec{e}_{j_1}, \dots, \vec{u}_s = u^{j_s} \vec{e}_{j_s}$ en V . Por linealidad,

$$\begin{aligned} S(\bar{\omega}^1, \dots, \bar{\omega}^r; \vec{u}_1, \dots, \vec{u}_s) &= \omega_{i_1} \dots \omega_{i_r} u^{j_1} \dots u^{j_s} S(\bar{e}^{i_1}, \dots, \bar{e}^{i_r}; \vec{e}_{j_1}, \dots, \vec{e}_{j_s}) \\ &= \omega_{i_1} \dots \omega_{i_r} u^{j_1} \dots u^{j_s} S_{j_1 \dots j_s}^{i_1 \dots i_r}. \end{aligned} \quad (3.7)$$

Por otro lado, dados los productos tensoriales

$$\vec{e}_{i_1} \otimes \dots \otimes \vec{e}_{i_r} \otimes \bar{e}^{j_1} \otimes \dots \otimes \bar{e}^{j_s} \quad (3.8)$$

puede verificarse fácilmente que

$$S = \vec{e}_{i_1} \otimes \dots \otimes \vec{e}_{i_r} \otimes \bar{e}^{j_1} \otimes \dots \otimes \bar{e}^{j_s} S_{j_1 \dots j_s}^{i_1 \dots i_r} \in V_r^s, \quad (3.9)$$

es decir que S así definido es un tensor de tipo (r, s) . Esta forma de escribir un tensor es completamente análoga a la manera en que escribimos un vector $\vec{u} = u^i \vec{e}_i$ o una forma $\bar{\omega} = \omega_i \bar{e}^i$ como combinaciones lineales de los elementos de una base, donde los coeficientes de la combinación son las respectivas componentes. Puede mostrarse que los productos tensoriales (3.8) forman una *base* del espacio tensorial V_r^s .

Ejemplo 3.1.3. Sea $\mathcal{A} : V \rightarrow V$ un operador lineal sobre V , con componentes A^i_j en la base $\{\vec{e}_i\}$ de V . Entonces

$$\mathcal{A} = A^i_j \vec{e}_i \otimes \bar{e}^j,$$

con lo cual si $\vec{x} = x^i \vec{e}_i \in V$,

$$\mathcal{A}\vec{x} = A^i_j \vec{e}_i \otimes \bar{e}^j x^l \vec{e}_l = A^i_j x^l \vec{e}_i \otimes \bar{e}^j (\vec{e}_l) = A^i_j x^l \vec{e}_i \delta^j_l = (A^i_j x^j) \vec{e}_i \in V,$$

y análogamente, si $\bar{\omega} = \omega_i \bar{e}^i \in V^*$,

$$\bar{\omega}\mathcal{A} = (\omega_i A^i_j) \bar{e}^j \in V^*.$$

Ejemplo 3.1.4. Podemos pensar al operador identidad $\mathcal{I} : V \rightarrow V$ como un tensor $I \in V_1^1$ tal que

$$I(\bar{\omega}, \vec{u}) = \bar{\omega}\mathcal{I}\vec{u} \equiv \bar{\omega}\vec{u} \quad \forall \bar{\omega} \in V^*, \vec{u} \in V.$$

Entonces

$$I = \delta^i_j \vec{e}_i \otimes \bar{e}^j,$$

donde δ^i_j es la delta de Kronecker.

Ejercicio 3.1.1. Muestre que los productos tensoriales (3.8) forman una base del espacio tensorial V_r^s , y que (3.9) es un tensor de tipo (r, s) .

Ejercicio 3.1.2. Muestre que $I = \delta^i_j \vec{e}_i \otimes \bar{e}^j$.

3.1.5. Cambio de base

El comportamiento de las componentes de un tensor ante un cambio de base en los espacios V y V^* puede deducirse de la misma forma en que lo hicimos para las componentes de vectores, formas y operadores lineales. La idea central es, nuevamente, que un tensor tiene existencia independiente de las bases en que expresamos sus componentes.

Así, si $\{\vec{e}_i\}_{i=1}^n$ y $\{\vec{e}'_j\}_{j=1}^n$ son bases de V con

$$\vec{e}'_j = \vec{e}_i A^i{}_j,$$

y $\{\vec{e}^i\}_{i=1}^n$, $\{\vec{e}'^j\}_{j=1}^n$ son sus respectivas bases duales (bases de V^*), el mismo tensor S podrá expresarse en cualquiera de estas bases, como

$$S = \vec{e}_{i_1} \otimes \cdots \otimes \vec{e}_{i_r} \otimes \vec{e}^{j_1} \otimes \cdots \otimes \vec{e}^{j_s} S_{j_1 \dots j_s}^{i_1 \dots i_r} \quad (3.10)$$

$$= \vec{e}'_{k_1} \otimes \cdots \otimes \vec{e}'_{k_r} \otimes \vec{e}'^{l_1} \otimes \cdots \otimes \vec{e}'^{l_s} S_{l_1 \dots l_s}^{k_1 \dots k_r}. \quad (3.11)$$

De aquí resulta

$$S_{l_1 \dots l_s}^{k_1 \dots k_r} = [\mathbf{A}^{-1}]_{i_1}^{k_1} \dots [\mathbf{A}^{-1}]_{i_r}^{k_r} A^{j_1}_{l_1} \dots A^{j_s}_{l_s} S_{j_1 \dots j_s}^{i_1 \dots i_r}, \quad (3.12)$$

mostrando que S es s veces covariante y r veces contravariante.

Esto nos permite introducir una definición alternativa de tensor: Cualquier objeto con $r+s$ índices, cada uno que corre de 1 a $\dim(V)$, y que ante un cambio de base $\vec{e}'_j = \vec{e}_i A^i{}_j$ en V transforme de acuerdo a (3.12), lo llamaremos tensor de rango $r+s$, s veces covariante (índices abajo) y r veces contravariante (índices arriba). Esta definición práctica es la que usaremos en adelante.

Ejemplo 3.1.5. Un vector $\vec{u} = u^j \vec{e}_j$ con componentes u^j , es un tensor de rango 1 una vez contravariante, también llamado vector contravariante (o simplemente *vector*).

Ejemplo 3.1.6. Una 1-forma $\omega = \omega_i \vec{e}^i$ con componentes ω_i , es un tensor de rango 1 una vez covariante, también llamado vector covariante o *covector*.

Ejemplo 3.1.7. Un operador lineal $\mathcal{A} = A^i{}_j \vec{e}_i \otimes \vec{e}^j$ con componentes $A^i{}_j$, es un tensor de rango 2 *mixto*, una vez covariante (j) y una vez contravariante (i) .

3.1.6. Contracción de índices

Ya hemos introducido el producto tensorial, que nos permite construir un tensor de rango mayor a partir de dos o más tensores de rangos menores. La operación contraria, que nos permite obtener un tensor de rango menor a partir de otro de rango mayor, es la *contracción de índices*.

Si $S \in V_r^s$ es un tensor de tipo (r, s) con componentes $S_{j_1 \dots j_s}^{i_1 \dots i_r}$, definimos su *contraído* respecto a los índices i_n y j_m como

$$S_{j_1 \dots j_{m-1} \mathbf{k} j_{m+1} \dots j_s}^{i_1 \dots i_{n-1} \mathbf{k} i_{n+1} \dots i_r} = S_{j_1 \dots j_{m-1} \mathbf{j} m j_{m+1} \dots j_s}^{i_1 \dots i_{n-1} \mathbf{i} n i_{n+1} \dots i_r} \delta^{\mathbf{j} m}_{\mathbf{i} n}.$$

De acuerdo a la ley de transformación (3.12), éste será un tensor de tipo $(r-1, s-1)$ y rango $r+s-2$.

Ejemplo 3.1.8. Sea $T = T^i_j \vec{e}_i \otimes \vec{e}^j$ un tensor mixto de rango 2. Su *traza*, definida como

$$\text{Tr}(T) := T^i_i,$$

es un *escalar* (tensor de rango 0).

Es fundamental aclarar que el contraído de un tensor será a su vez un tensor, si y sólo si los índices que se contraen son tomados de a pares, uno covariante y el otro contravariante. Por ejemplo, si S_{ij} son las componentes de un tensor covariante de rango 2, S_{ii} no es un tensor (escalar, en este caso), ya que no resulta invariante ante cambios de base.

Ejercicio 3.1.3. Verifique que S_{ii} no es un escalar, aunque S_{ij} sea un tensor.

3.1.7. Simetría

Una propiedad importante que pueden tener los tensores, y que ya hemos conocido a través de las matrices, es la simetría o antisimetría.

Sea $S \in V_r^s$ un tensor de tipo (r, s) con componentes $S^{i_1 \dots i_r}_{j_1 \dots j_s}$. Diremos que S es *simétrico* respecto a los índices i_n, i_m si

$$S^{i_1 \dots \color{blue}{i_m} \dots \color{red}{i_n} \dots i_r}_{j_1 \dots j_s} \equiv S^{i_1 \dots \color{red}{i_n} \dots \color{blue}{i_m} \dots i_r}_{j_1 \dots j_s},$$

es decir si sus componentes son invariantes respecto del intercambio de los índices i_n, i_m . De igual modo, le llamaremos simétrico respecto a los índices j_n, j_m si

$$S^{i_1 \dots i_r}_{j_1 \dots \color{blue}{j_m} \dots \color{red}{j_n} \dots j_s} \equiv S^{i_1 \dots i_r}_{j_1 \dots \color{red}{j_n} \dots \color{blue}{j_m} \dots j_s}.$$

Si un tensor es simétrico respecto a cualquier par de índices, le llamaremos simplemente *simétrico*.

Diremos que S es *antisimétrico* respecto a los índices i_n, i_m si

$$S^{i_1 \dots \color{blue}{i_m} \dots \color{red}{i_n} \dots i_r}_{j_1 \dots j_s} \equiv -S^{i_1 \dots \color{red}{i_n} \dots \color{blue}{i_m} \dots i_r}_{j_1 \dots j_s},$$

es decir si sus componentes cambian de signo ante el intercambio de los índices i_n, i_m . De igual modo, le llamaremos antisimétrico respecto a los índices j_n, j_m si

$$S^{i_1 \dots i_r}_{j_1 \dots \color{blue}{j_m} \dots \color{red}{j_n} \dots j_s} \equiv -S^{i_1 \dots i_r}_{j_1 \dots \color{red}{j_n} \dots \color{blue}{j_m} \dots j_s}.$$

Si $S \in V_r^0$, le llamaremos *totalmente antisimétrico* si

$$S^{\Pi(i_1, \dots, i_r)} = \text{sgn}(\Pi) S^{i_1 \dots i_r},$$

donde $\Pi(i_1, \dots, i_r)$ es una permutación cualquiera del conjunto ordenado de índices $\{i_1, \dots, i_r\}$, y $\text{sgn}(\Pi)$ es su signo (+ si es una permutación par y - si es impar). De igual modo, si $S \in V_0^s$, le llamaremos *totalmente antisimétrico* si

$$S_{\Pi(j_1, \dots, j_s)} = \text{sgn}(\Pi) S_{j_1 \dots j_s}.$$

Notemos que definimos la simetría o antisimetría de un tensor solamente respecto al intercambio de pares de índices ambos covariantes o ambos contravariantes. Ello es así porque de este modo la simetría o antisimetría se preserva ante cambios de base. Si la definiéramos respecto de índices uno covariante y el otro contravariante, cambiaría con la base en que representáramos el tensor, y no podríamos considerarla una propiedad del tensor, sino sólo de su representación en alguna base particular.

Ejercicio 3.1.4. Verifique que $T_j^i = T_i^j$ no implica $T'^j{}_j = T'^i{}_i$.

Ejercicio 3.1.5. Sea $S_{i_1 \dots i_s}$ un tensor totalmente antisimétrico. Muestre que toda componente con dos o más índices repetidos *debe* ser nula. Deduzca de allí que las únicas componentes no nulas son aquellas donde $\{i_1, i_2, \dots, i_s\}$ es una permutación de $\{1, 2, \dots, s\}$, es decir, las que tienen todos sus índices diferentes. Muestre que, en consecuencia, $S_{i_1 \dots i_s}$ tiene una única *componente estricta* $s := S_{1,2,\dots,s}$, ya que toda otra componente no nula $S_{i_1 \dots i_s}$ será igual a s si $\{i_1, i_2, \dots, i_s\}$ es una permutación par de $\{1, 2, \dots, s\}$, y a $-s$ si es una permutación impar.

Ejercicio 3.1.6. Muestre que si un tensor totalmente antisimétrico tiene rango mayor que $\dim(V)$, es idénticamente nulo.

Dos operaciones relacionadas con las definiciones anteriores son la simetrización y antisimetrización de tensores. Dado un tensor $T \in V_r^0$ de tipo $(r, 0)$ con componentes $T^{i_1 \dots i_r}$, definimos su *parte simétrica* como el tensor

$$ST \in V_r^0$$

con componentes

$$T^{\{i_1 \dots i_r\}} := \frac{1}{r!} \sum_{\Pi} T^{\Pi(i_1, \dots, i_r)},$$

donde la suma corre sobre todas las permutaciones de los índices. Asimismo, definimos su *parte antisimétrica* como el tensor

$$\mathcal{A}T \in V_r^0$$

con componentes

$$T^{[i_1 \dots i_r]} := \frac{1}{r!} \sum_{\Pi} \text{sgn}(\Pi) T^{\Pi(i_1, \dots, i_r)}$$

Definiciones y notación completamente análogas valen para tensores $T \in V_0^s$ de tipo $(0, s)$.

Las operaciones de simetrización y antisimetrización pueden extenderse a conjuntos arbitrarios de índices, mientras sean todos del mismo tipo (covariantes o contravariantes). Así por ejemplo

$$T_k^{[ij]} = \frac{1}{2} [T_k^{ij} - T_k^{ji}], \quad S_{j\{kl\}}^i = \frac{1}{2} [S_{jkl}^i + S_{jlk}^i], \quad \text{etc.}$$

Ejemplo 3.1.9. Sea el tensor $T \in V_0^3$, $V = \mathbb{R}^3$ con componentes

$$T_{1ij} = \begin{pmatrix} 1 & -1 & 0 \\ 2 & 0 & -2 \\ 3 & 2 & 1 \end{pmatrix}, \quad T_{2ij} = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 0 & 0 \\ 3 & 1 & 1 \end{pmatrix}, \quad T_{3ij} = \begin{pmatrix} 0 & -1 & -2 \\ 0 & 2 & 1 \\ 0 & -3 & -3 \end{pmatrix}.$$

Tendremos

$$T_{1\{ij\}} = \begin{pmatrix} 1 & 1/2 & 3/2 \\ 1/2 & 0 & 0 \\ 3/2 & 0 & 1 \end{pmatrix}, \quad T_{2\{ij\}} = \begin{pmatrix} 1 & 1 & 3 \\ 0 & 1/2 & 1 \\ 3 & 1/2 & 1 \end{pmatrix}, \quad T_{3\{ij\}} = \begin{pmatrix} 0 & -1/2 & -1 \\ -1/2 & 2 & -1 \\ -1 & -1 & -3 \end{pmatrix}$$

y

$$T_{1[ij]} = \begin{pmatrix} 0 & -3/2 & -3/2 \\ 3/2 & 0 & -2 \\ -3/2 & 2 & 0 \end{pmatrix}, \quad T_{2[ij]} = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & -1/2 \\ 0 & 1/2 & 0 \end{pmatrix}, \quad T_{3[ij]} = \begin{pmatrix} 0 & -1/2 & -1 \\ 1 & 0 & 2 \\ 1 & -2 & 0 \end{pmatrix}.$$

Nótese que $T_{ijk} = T_{i\{jk\}} + T_{i[jk]}$.

Ejercicio 3.1.7. Complete los cálculos precedentes, y calcule $T_{\{ijk\}}$ y $T_{[ijk]}$. ¿Es cierto que $T_{ijk} = T_{\{ijk\}} + T_{[ijk]}$? ¿Por qué?

3.1.8. Producto exterior

Cuando tratamos con vectores, nos es familiar la noción de producto vectorial entre dos vectores, que es otro vector.² Un concepto relacionado a él, y hasta cierto punto similar, es el de producto exterior entre dos tensores.

Dado un tensor $S \in V_0^s$ con componentes $S_{j_1 \dots j_s}$, totalmente antisimétrico, y otro tensor $T \in V_0^t$ con componentes $T_{j_1 \dots j_t}$, también totalmente antisimétrico, su *producto exterior*

$$S \wedge T := \mathcal{A}(S \otimes T)$$

es un nuevo tensor $\in V_0^{s+t}$ totalmente antisimétrico, con componentes

$$S_{[j_1 \dots j_s} T_{l_1 \dots l_t]}.$$

Una definición completamente equivalente vale para tensores contravariantes. El producto exterior posee las siguientes propiedades:

1. $S \wedge (T_1 + T_2) = S \wedge T_1 + S \wedge T_2$.
2. $S \wedge (T \wedge R) = (S \wedge T) \wedge R = S \wedge T \wedge R$.
3. $S \wedge T = (-1)^{st} T \wedge S$.

Ejercicio 3.1.8. Demuestre las propiedades precedentes.

Ejemplo 3.1.10. El producto exterior entre dos vectores \vec{u} y \vec{v} de \mathbb{R}^3 será el tensor con componentes

$$u^{[i} v^{j]} = \frac{1}{2} [u^i v^j - u^j v^i],$$

²En realidad, como veremos en seguida, es un *pseudovector*.

que podemos representar por la matriz antisimétrica

$$\frac{1}{2} \begin{pmatrix} 0 & u^1 v^2 - u^2 v^1 & u^1 v^3 - u^3 v^1 \\ u^2 v^1 - u^1 v^2 & 0 & u^2 v^3 - u^3 v^2 \\ u^3 v^1 - u^1 v^3 & u^3 v^2 - u^2 v^3 & 0 \end{pmatrix}.$$

Notemos que $\vec{u} \wedge \vec{u} = 0$.

3.1.9. Densidades tensoriales

Existen muchos objetos con significado físico que ante cambios de base no transforman exactamente como tensores, sino de una forma estrechamente relacionada. El producto vectorial entre dos vectores es un ejemplo; las densidades de masa y carga son otro.

Todo objeto con $r+s$ índices, cada uno que corre de 1 a $\dim(V)$, y que ante un cambio de base $\vec{e}'_j = \vec{e}_i A^i_j$ en V transforme de acuerdo a

$$S'^{k_1 \dots k_r}_{l_1 \dots l_s} = \det(\mathbf{A})^p [\mathbf{A}^{-1}]^{k_1}_{i_1} \dots [\mathbf{A}^{-1}]^{k_r}_{i_r} A^{j_1}_{l_1} \dots A^{j_s}_{l_s} S^{i_1 \dots i_r}_{j_1 \dots j_s}, \quad (3.13)$$

se llama *densidad tensorial* o pseudotensor de peso p .

Ejemplo 3.1.11. Las densidades de carga eléctrica y de masa no son escalares, sino pseudoescalares de peso 1. Análogamente, el diferencial de volumen es un pseudoescalar de peso -1 . Estos resultados serán discutidos en mas detalle en las secciones siguientes.

Ejercicio 3.1.9. Demuestre que el producto tensorial de dos densidades tensoriales de pesos p_1 y p_2 , es otra densidad tensorial de peso $p_1 + p_2$.

Ejercicio 3.1.10. Demuestre que la contracción de índices no altera el peso de una densidad tensorial, sino sólo su rango.

Ejemplo 3.1.12. El símbolo de permutación de Levi–Civita se define como

$$\epsilon_{j_1 \dots j_s} = \begin{cases} 1 & \text{si } \{j_1, \dots, j_s\} \text{ es permutación par de } \{1, \dots, s\}, \\ -1 & \text{si } \{j_1, \dots, j_s\} \text{ es permutación impar de } \{1, \dots, s\}, \\ 0 & \text{para todo otro caso,} \end{cases}$$

es *totalmente antisimétrico*,³ tiene rango $s = \dim(V)$, y *por definición* sus componentes asumen *exactamente los mismos valores* en cualquier sistema de coordenadas. Está claro entonces que $\epsilon_{j_1 \dots j_s}$ no puede ser un tensor, que es el motivo por el que se lo llama *símbolo*. Alternativamente podemos definir $\epsilon^{i_1 \dots i_s}$, cuyas componentes coinciden numéricamente con las del anterior, y por las mismas razones tampoco es un tensor. Sin embargo su utilidad para realizar cálculos utilizando la notación de Einstein de suma sobre índices repetidos hace que estos

³Nótese que ello implica que tiene una *única* componente que necesitamos especificar, digamos $\epsilon_{1 \dots s} = 1$; todas las otras quedan inmediatamente especificadas por la propiedad de antisimetría total.

símbolos sean ubicuos. Por eso vale la pena mostrar cómo, con las precauciones adecuadas, $\epsilon_{j_1 \dots j_s}$ puede ser interpretado como una densidad tensorial (covariante) de peso -1 y rango $s = \dim(V)$, y $\epsilon^{i_1 \dots i_s}$ como una densidad tensorial (contravariante) de peso 1 .

Lo mostraremos explícitamente para $s = 2$. En ese caso $\epsilon_{12} = 1$, $\epsilon_{21} = -1$, y $\epsilon_{11} = \epsilon_{22} = 0$. Luego, si ante un cambio de coordenadas con matriz de transformación $A^i{}_j$ aplicamos las reglas de transformación usuales, como si ϵ_{ij} fueran las componentes de un tensor, la componente ${}_{ij}$ del objeto transformado será

$$\epsilon_{kl} A^k{}_i A^l{}_j = A^1{}_i A^2{}_j - A^2{}_i A^1{}_j = \begin{cases} \det(\mathbf{A}) & \text{si } i = 1, j = 2, \\ -\det(\mathbf{A}) & \text{si } i = 2, j = 1, \\ 0 & \text{si } i = j, \end{cases}$$

de donde es inmediato que

$$\epsilon_{ij} = \det(\mathbf{A})^{-1} A^k{}_i A^l{}_j \epsilon_{kl} = \begin{cases} 1 & \text{si } \{i, j\} \text{ es permutación par de } \{1, 2\}, \\ -1 & \text{si } \{i, j\} \text{ es permutación impar de } \{1, 2\}, \\ 0 & \text{para todo otro caso.} \end{cases}$$

Es decir, el símbolo ϵ_{ij} cumple el requerimiento de preservar sus componentes pese al cambio de coordenadas (por lo cual omitimos la “prima” en el lado izquierdo), y para hacerlo “transforma” como una densidad tensorial de peso -1 dos veces covariante.

Ejercicio 3.1.11. Demuestre que $\epsilon_{j_1 \dots j_s} = \det(\mathbf{A})^{-1} A^{i_1}_{j_1} \dots A^{i_s}_{j_s} \epsilon_{i_1 \dots i_s}$ (peso -1) y que $\epsilon^{i_1 \dots i_r} = \det(\mathbf{A}) [\mathbf{A}^{-1}]^{i_1}_{j_1} \dots [\mathbf{A}^{-1}]^{i_r}_{j_r} \epsilon^{j_1 \dots j_r}$ (peso 1).

Ejercicio 3.1.12. Use el resultado de arriba para demostrar que

$$\det(\mathbf{A}) = \epsilon_{i_1 \dots i_s} A^{i_1}_1 \dots A^{i_s}_s = A^{1}_{j_1} \dots A^{s}_{j_s} \epsilon^{j_1 \dots j_s}.$$

Muestre además que

$$\det(\mathbf{A}) = \frac{1}{s!} \epsilon_{i_1 \dots i_s} A^{i_1}_{j_1} \dots A^{i_s}_{j_s} \epsilon^{j_1 \dots j_s}.$$

3.1.10. Tensor adjunto

Un concepto íntimamente relacionado a los de densidades tensoriales y producto tensorial, y que frecuentemente aparece en conjunción con ellos, es el de tensor adjunto o dual.

Dado $\dim(V) = n$ y un tensor completamente antisimétrico $T^{i_1 \dots i_r}$ de rango $r \leq n$, definimos el tensor *adjunto*⁴ (o dual) como

$$\bar{T}_{i_1 \dots i_{n-r}} := \epsilon_{i_1 \dots i_{n-r} j_1 \dots j_r} T^{j_1 \dots j_r}, \quad (3.14)$$

⁴No confundir con la adjunta Hermitiana.

es decir, contraemos los índices de $T^{i_1 \dots i_r}$ con los últimos r índices del símbolo de Levi–Civita. El tensor adjunto es también totalmente antisimétrico, y su rango es $n - r$. Sin embargo, si $T^{i_1 \dots i_r}$ era un tensor, su adjunto será una densidad tensorial.

Definiciones completamente análogas valen para tensores covariantes e incluso para densidades tensoriales.

Ejemplo 3.1.13. El producto vectorial usual es el adjunto del producto exterior (ver Ejemplo 3.1.10) de dos vectores \vec{u} y \vec{v} de \mathbb{R}^3 ,

$$(\vec{u} \times \vec{v})_i = \epsilon_{ijk} u^{[j} v^{k]} = \begin{cases} u^2 v^3 - u^3 v^2 & \text{si } i = 1, \\ u^3 v^1 - u^1 v^3 & \text{si } i = 2, \\ u^1 v^2 - u^2 v^1 & \text{si } i = 3, \end{cases}$$

y es una densidad tensorial de peso -1 o *pseudovector* (covariante).

Ejercicio 3.1.13. Use los resultados de los Ejercicios 3.1.9, 3.1.10 y 3.1.11 para demostrar la afirmación precedente.

Ejercicio 3.1.14. Muestre que si A_{ij} es antisimétrico en los índices i y j , y S^{ij} es simétrico, entonces $A_{ij} S^{ij} = 0$. Muestre también que para todo tensor T de rango 2, $T^{ij} = T^{\{ij\}} + T^{[ij]}$ (idem con T_{ij}). Use estos resultados para mostrar que el producto vectorial entre dos vectores de \mathbb{R}^3 puede escribirse

$$\vec{u} \times \vec{v} = \epsilon_{ijk} \vec{e}^i u^j v^k = \epsilon_{ijk} u^i v^j \vec{e}^k.$$

3.1.11. Ley del cociente

Este es un procedimiento rápido y práctico para reconocer si un objeto dado es un tensor. Lo ilustraremos con un ejemplo.

Supongamos que sabemos que \vec{u} es un vector con componentes u^i , y determinamos que $T^i_j u^j$ es un vector $\forall \vec{u}$. Entonces T es un tensor, en este caso de rango 2 y tipo $(1, 1)$ o *mixto*.

Un procedimiento análogo vale para tensores de rango superior, y para densidades tensoriales.

3.2. Coordenadas curvilíneas

Hasta ahora hemos venido considerando cambios de coordenadas de la forma $\vec{e}'_j = \vec{e}_i A^i_j$ con A^i_j constantes, es decir cambios de coordenadas *globales*. Estos incluyen rotaciones, reflexiones y permutaciones de la base, y cambios entre sistemas Cartesianos no necesariamente ortogonales. Pero todos estos cambios de coordenadas comparten una característica: si los vectores de la vieja base eran paralelos a sí mismos en cada punto del espacio (es decir, si era una base Cartesiana), los de la nueva base también lo serán.

Sin embargo, en Física estamos acostumbrados a usar sistemas de coordenadas no Cartesianos, las llamadas *coordenadas curvilíneas*, que muchas veces

resultan más adecuados a la geometría del problema que estamos estudiando. Estos sistemas se caracterizan por el hecho de que los vectores base (que son tangentes a las curvas coordenadas, y en general no son ni unitarios ni ortogonales) cambian de dirección de un punto a otro del espacio. Por lo tanto, cualquier cambio de coordenadas que involucre coordenadas curvilíneas será *local*, es decir, los elementos A^i_j de la matriz de transformación serán funciones de punto.

En esta sección consideraremos las consecuencias de realizar cambios de coordenadas locales, es decir de la forma

$$\vec{e}'_j = \vec{e}_i A^i_j(\vec{x}).$$

3.2.1. Cambios de coordenadas locales

Un cambio de coordenadas cualquiera en V [$\dim(V) = n$] queda completamente caracterizado dando, por ejemplo, la expresión de las nuevas coordenadas x'^i como funciones de las viejas coordenadas x^j , es decir

$$x'^i = x'^i(x^1, \dots, x^n), \quad i = 1, \dots, n. \quad (3.15)$$

En el caso de cambios de coordenadas globales, las funciones $x'^i(x^1, \dots, x^n)$ eran lineales en todos sus argumentos, con una matriz de coeficientes no singular y constante. En el caso de cambios de coordenadas locales, consideraremos admisible cualquier forma funcional para las x'^i , siempre y cuando podamos invertir estas funciones para despejar

$$x^j = x^j(x'^1, \dots, x'^n), \quad j = 1, \dots, n. \quad (3.16)$$

La condición necesaria y suficiente para ésto es que el determinante J de la matriz Jacobiana J de la transformación no se anule,

$$J = \det(J) := \left| \frac{\partial(x^1, \dots, x^n)}{\partial(x'^1, \dots, x'^n)} \right| = \left| \frac{\partial x^i}{\partial x'^j} \right| \neq 0. \quad (3.17)$$

Si esto ocurre, diremos que la transformación es *no singular*, o que las nuevas coordenadas están “bien definidas” en términos de las viejas; en general esta no-singularidad será sólo *local* y no global.

Ejemplo 3.2.1. Sean x^1 y x^2 coordenadas Cartesianas usuales en \mathbb{R}^2 , y sean x'^1 y x'^2 coordenadas polares (en notación usual, $x^1 = x$, $x^2 = y$, $x'^1 = \rho$ y $x'^2 = \theta$). Entonces

$$x'^1 = \sqrt{(x^1)^2 + (x^2)^2}, \quad x'^2 = \arctan \frac{x^2}{x^1},$$

y

$$x^1 = x'^1 \cos x'^2, \quad x^2 = x'^1 \sin x'^2.$$

Notemos que el Jacobiano de la transformación,

$$J = \left| \frac{\partial(x^1, x^2)}{\partial(x'^1, x'^2)} \right| = \det \begin{pmatrix} \cos x'^2 & -x'^1 \sin x'^2 \\ \sin x'^2 & x'^1 \cos x'^2 \end{pmatrix} = x'^1,$$

sólo se anula para $x'^1 = 0$, hecho que conocemos como “el sistema de coordenadas polares es singular en el origen”.

Las ecuaciones (3.15) y (3.16) pueden ser altamente no lineales, lo cual generalmente dificulta trabajar con ellas. Sin embargo, la relación entre los *diferenciales* de las coordenadas viejas y nuevas,

$$dx'^i = \frac{\partial x'^i}{\partial x^j} dx^j, \quad (3.18)$$

es *siempre* lineal y homogénea, aunque los coeficientes ($\partial x'^i / \partial x^j$) puedan depender de las coordenadas, y aún si esa dependencia es no lineal. Notando que si las relaciones (3.15) fueran lineales y tuviesen la forma

$$x'^i = a^i_j x^j$$

con a^i_j constantes, la relación entre diferenciales se reduciría a

$$dx'^i = a^i_j dx^j,$$

vemos que la generalización a coordenadas curvilíneas de las definiciones de tensores, densidades, etc., dadas en la Sección 3.1, puede hacerse simplemente reemplazando la matriz de transformación A de los cambios de coordenadas globales, por la matriz Jacobiana J , es decir haciendo los reemplazos

$$[A^i_j] \rightarrow \left[\frac{\partial x^i}{\partial x'^j} \right] =: J, \quad [A^{-1}]^i_j \rightarrow \left[\frac{\partial x'^i}{\partial x^j} \right] =: J^{-1}, \quad \det(A) \rightarrow \det(J) =: J. \quad (3.19)$$

Esta identificación, junto con la linealidad de la relación entre los diferenciales, simplificarán notablemente los cálculos subsiguientes.

3.2.2. La base tangente

Una *curva coordenada* correspondiente a x'^i es una curva

$$\vec{x}(\textcolor{red}{x}^{\textcolor{red}{i}}) = x^j (x'^1 = \text{cte.}, \dots, \textcolor{red}{x}^{\textcolor{red}{i}}, \dots, x'^n = \text{cte.}) \vec{e}_j$$

obtenida al variar la coordenada x'^i manteniendo constantes todas las otras x'^k , $k \neq i$ en la expresión $\vec{x} = x^i \vec{e}_i$, con las coordenadas x^i como funciones de las x'^i como en (3.16). Las diferentes curvas coordinadas $\vec{x}(\textcolor{red}{x}^{\textcolor{red}{i}})$ están parametrizadas por los valores de las x'^k ($k \neq i$) constantes.

El vector \vec{e}'_i se *define* como el *vector tangente* a la correspondiente curva coordinada, de allí el nombre de *base tangente*. De la expresión

$$d\vec{x} = dx^i \vec{e}_i = dx'^j \vec{e}'_j$$

para un dado desplazamiento infinitesimal $d\vec{x}$, usando (3.18) es inmediato que

$$\vec{e}'_j = \frac{\partial \vec{x}}{\partial x'^j} = \frac{\partial x^i}{\partial x'^j} \vec{e}_i, \quad (3.20)$$

que es la expresión para la transformación de los vectores de la base tangente;⁵ por ello suele llamársela también *base covariante*.

Ejemplo 3.2.2. Para las coordenadas polares definidas en el Ejemplo 3.2.1, las curvas coordenadas $\vec{x}(x'^1)$ son semirrectas radiales que parten del origen, y las $\vec{x}(x'^2)$ son circunferencias concéntricas con el origen. Los correspondientes vectores base son

$$\vec{e}'_1 = \cos x'^2 \vec{e}_1 + \sin x'^2 \vec{e}_2, \quad \vec{e}'_2 = -\sin x'^1 \vec{e}_1 + \cos x'^1 \vec{e}_2.$$

Nótese que \vec{e}'_2 no es unitario, aunque en este caso \vec{e}'_1 y \vec{e}'_2 sí son ortogonales.

Ejercicio 3.2.1. Demuestre las afirmaciones precedentes.

Ejercicio 3.2.2. Determine las curvas coordenadas y los vectores base de las *coordenadas esféricas* definidas por

$$x^1 = x'^1 \sin x'^2 \cos x'^3, \quad x^2 = x'^1 \sin x'^2 \sin x'^3, \quad x^3 = x'^1 \cos x'^2,$$

donde x^1 , x^2 y x^3 son las coordenadas Cartesianas usuales. Los vectores base, ¿son ortogonales? ¿Y unitarios?

Ejercicio 3.2.3. *Idem* para las *coordenadas cilíndricas* definidas por

$$x^1 = x'^1 \cos x'^2, \quad x^2 = x'^1 \sin x'^2, \quad x^3 = x'^3.$$

3.2.3. Vectores contravariantes

Definiremos como vector *contravariante* a cualquier objeto con componentes u^i que, ante cambios de coordenadas, transformen de acuerdo a

$$u'^i = \frac{\partial x'^i}{\partial x^j} u^j, \tag{3.21}$$

que es la ley de transformación (3.18) de los diferenciales de coordenadas.

Ejemplo 3.2.3. Son vectores contravariantes: las componentes del vector desplazamiento $d\vec{x}$ (por definición); las del vector velocidad $\frac{d\vec{x}}{dt}$; etc.

Ejemplo 3.2.4. En coordenadas Cartesianas ortonormales, el vector posición en \mathbb{R}^2 se escribe

$$\vec{r} = x^1 \vec{e}_1 + x^2 \vec{e}_2.$$

Reescribiendo x^1 y x^2 en coordenadas polares tendremos

$$\vec{r} = x'^1 \cos x'^2 \vec{e}_1 + x'^1 \sin x'^2 \vec{e}_2.$$

⁵En la página 34 Blennow presenta esta ley de transformación como la primaria, y considera a (3.18) como derivada de ella. Sin embargo, su deducción de (3.20) sólo es válida si $\{\vec{e}_i\}_{i=1}^n$ es la base Cartesiana ortonormal. En cambio la deducción de (3.18) presentada aquí, y con ella la de (3.20), es válida aún si las coordenadas $\{x^i\}$ son también curvilíneas.

Comparando con la expresión para \vec{e}'_1 obtenida en el Ejemplo 3.2.2, vemos que

$$\vec{r} = x'^1 \vec{e}'_1,$$

es decir, en coordenadas polares el vector posición es puramente *radial*, y su componente “angular” es nula.

Ejercicio 3.2.4. Repita los cálculos precedentes para el vector posición en \mathbb{R}^3 en coordenadas esféricas.

3.2.4. Vectores covariantes

Sea $\varphi(\vec{x})$ una función de punto (escalar); entonces por la regla de la cadena

$$\frac{\partial \varphi}{\partial x'^i} = \frac{\partial x^j}{\partial x'^i} \frac{\partial \varphi}{\partial x^j}.$$

La variación $d\varphi$ de φ ante un desplazamiento infinitesimal $d\vec{x}$ será

$$d\varphi = \frac{\partial \varphi}{\partial x^j} dx^j,$$

y es un escalar, invariante bajo cambios de coordenadas. La ley del cociente nos lleva entonces a definir como vector *covariante* a cualquier objeto con componentes u_i que, ante cambios de coordenadas, transformen de acuerdo a

$$u'_i = \frac{\partial x^j}{\partial x'^i} u_j, \quad (3.22)$$

que es la ley de transformación (3.20) de los vectores base (inversa de la ley de transformación de los diferenciales de coordenadas).

Ejemplo 3.2.5. El gradiente de una función escalar de punto, $\nabla\varphi$, es un vector covariante con componentes⁶

$$\varphi_{,i} := \frac{\partial \varphi}{\partial x^i}.$$

Son vectores covariantes: fuerza, campo eléctrico, y todo otro que se obtenga como el gradiente de un potencial escalar.

3.2.5. La base dual

Supongamos que tenemos un vector contravariante $\vec{u} = u^i \vec{e}_i$, y un vector covariante de componentes v_i . De lo hecho hasta ahora está claro que $v_i u^i$ debe ser

⁶La derivada parcial respecto de x^i se denota por una coma seguida del *índice* de la coordenada respecto de la que derivamos, en posición de *subíndice*. Nótese que la derivación respecto de una variable con *superíndice* produce la aparición de un *subíndice*, consistentemente con el carácter escalar de $d\varphi = \varphi_{,i} dx^i$.

un *escalar*, ya que ante un cambio de base las reglas de transformación (3.21) para las componentes u^i y (3.22) para las componentes v_i cancelarán mutuamente sus efectos:

$$v'_i u'^i = v_j \frac{\partial x^j}{\partial x'^i} \frac{\partial x'^i}{\partial x^k} u^k = v_j \frac{\partial x^j}{\partial x^k} u^k = v_j u^j,$$

donde hemos usado que $\partial x^j / \partial x^k = 1$ si $j = k$ y es nula si no (véase el Ejemplo 3.2.9 mas adelante). Es natural preguntarse entonces qué significado deberíamos asignar a una expresión como

$$\vec{v} = v_i \vec{e}^i,$$

y, en particular, cómo deberíamos caracterizar la (supuesta) base $\{\vec{e}^i\}$ de los vectores covariantes. Para responder, notemos que estamos requiriendo alguna “operación” entre los vectores \vec{v} y \vec{u} , que denotaremos provisoriamente como $\vec{v}(\vec{u})$, y que debe satisfacer

$$\vec{v}(\vec{u}) = v_i u^i.$$

Comparando con la Sección 1.5, vemos que $\vec{v}(\vec{u})$ no es otra cosa que una aplicación lineal del espacio V de los vectores contravariantes al cuerpo escalar, es decir, una *forma*, y que el espacio de los vectores covariantes es V^* . Apelando a la linealidad, entonces, tendremos

$$\vec{v}(\vec{u}) = v_i \vec{e}^i (u^j \vec{e}_j) = v_i u^j \vec{e}^i (\vec{e}_j) = v_i u^i,$$

de donde es inmediato que $\{\vec{e}^i\}$ no es otra cosa que la *base dual* de $\{\vec{e}_i\}$.⁷

La regla de transformación para los vectores de la base dual es, por supuesto,

$$\vec{e}'^j = \frac{\partial x'^j}{\partial x^i} \vec{e}^i, \quad (3.23)$$

por lo que también suele llamársela *base contravariante*.

Ejemplo 3.2.6. Para las coordenadas Cartesianas ortonormales, es inmediato que $\vec{e}'^i \equiv \vec{e}_i$.

Ejemplo 3.2.7. Para las coordenadas polares del Ejemplo 3.2.2,

$$\vec{e}'^1 = \cos x'^2 \vec{e}^1 + \sin x'^2 \vec{e}^2, \quad \vec{e}'^2 = -\frac{1}{x'^1} \sin x'^2 \vec{e}^1 + \frac{1}{x'^1} \cos x'^2 \vec{e}^2.$$

Ejercicio 3.2.5. Verifique las afirmaciones precedentes.

Ejercicio 3.2.6. Construya la base dual de las coordenadas esféricas del Ejercicio 3.2.2.

⁷En este contexto lo usual es no usar la notación de “flecha invertida” \vec{v} para las formas, ya que en general quedará claro del contexto si \vec{v} es una forma (vector covariante) o un vector (contravariante). En particular, la posición del índice de los vectores base hace explícito si son vectores de la base tangente (índices abajo) o de la base dual (índices arriba).

Ejemplo 3.2.8. Consideremos una partícula en \mathbb{R}^2 que tiene una energía potencial $U(\vec{x}) = -k/x^1$ en coordenadas polares x^1, x^2 . La fuerza actuante sobre ella será

$$\vec{F} = -\nabla U = -U_{,i}\vec{e}^i = -\frac{\partial U}{\partial x^1}\vec{e}^1 - \frac{\partial U}{\partial x^2}\vec{e}^2 = -\frac{k}{(x^1)^2}\vec{e}^1.$$

Ejercicio 3.2.7. Repita el cálculo precedente en coordenadas Cartesianas orto-normales, donde $U(\vec{x}) = -k/\sqrt{(x^1)^2 + (x^2)^2}$; note la conveniencia de trabajar en coordenadas polares.

3.2.6. Tensores

Análogamente a las definiciones anteriores, un objeto con componentes $S_{j_1\dots j_s}^{i_1\dots i_r}$ que ante cambios de coordenadas transforme de acuerdo a

$$S'^{k_1\dots k_r}_{l_1\dots l_s} = \frac{\partial x'^{k_1}}{\partial x^{i_1}} \dots \frac{\partial x'^{k_r}}{\partial x^{i_r}} \frac{\partial x^{j_1}}{\partial x'^{l_1}} \dots \frac{\partial x^{j_s}}{\partial x'^{l_s}} S_{j_1\dots j_s}^{i_1\dots i_r}, \quad (3.24)$$

es un *tensor* de rango $r+s$, r veces contravariante (índices arriba) y s veces covariante (índices abajo).

Ejemplo 3.2.9. La delta de Kronecker puede definirse como

$$\delta^i_j := \frac{\partial x^i}{\partial x^j},$$

y es un tensor mixto de rango 2, ya que por la regla de la cadena

$$\delta'^i_j = \frac{\partial x'^i}{\partial x'^j} = \frac{\partial x'^i}{\partial x^k} \frac{\partial x^k}{\partial x^l} \frac{\partial x^l}{\partial x'^j} = \frac{\partial x'^i}{\partial x^k} \frac{\partial x^l}{\partial x'^j} \delta^k_l.$$

Ejemplo 3.2.10. Sean u^i las componentes de un vector contravariante, y v_j las de un vector covariante. Entonces $u^i v_j$ son las componentes de un tensor mixto de rango 2 (su producto tensorial), y el contraído $u^i v_i$ es un tensor de rango 0 (escalar), su producto interno. Notemos que ni $u^i v^i$ ni $u_i v_i$ son escalares (dependen de la base). Por ejemplo con fuerzas conservativas, la velocidad de un móvil es un vector contravariante con componentes v^i , y la fuerza aplicada es un vector covariante (ya que es el gradiente de un potencial escalar) con componentes f_i , de modo que la potencia

$$P := f_i v^i$$

es un escalar.

Ejercicio 3.2.8. Demuestre las afirmaciones precedentes.

Ejercicio 3.2.9. Muestre que el símbolo δ_{ij} , definido en ciertas coordenadas como igual a 1 si $i = j$ e igual a 0 si no, no toma los mismos valores en otras coordenadas diferentes si se lo supone un tensor de tipo $(0, 2)$.

3.2.7. Densidades tensoriales

Un objeto con componentes $S_{j_1 \dots j_s}^{i_1 \dots i_r}$ que ante cambios de coordenadas transforme de acuerdo a

$$S'_{l_1 \dots l_s}^{k_1 \dots k_r} = J^p \frac{\partial x'^{k_1}}{\partial x^{i_1}} \dots \frac{\partial x'^{k_r}}{\partial x^{i_r}} \frac{\partial x^{j_1}}{\partial x'^{l_1}} \dots \frac{\partial x^{j_s}}{\partial x'^{l_s}} S_{j_1 \dots j_s}^{i_1 \dots i_r}, \quad (3.25)$$

es una *densidad tensorial* de peso p y rango $r+s$, r veces contravariante (índices arriba) y s veces covariante (índices abajo), es decir de tipo (r, s) .

Ejemplo 3.2.11. El símbolo de permutación de Levi–Civita $\epsilon_{i_1 \dots i_n}$ es una densidad tensorial (covariante) de peso -1 , y $\epsilon^{i_1 \dots i_n}$ es una densidad tensorial (contravariante) de peso 1 .

Ejercicio 3.2.10. Demuestre las afirmaciones precedentes para $n = 2$ y $n = 3$.

Ejemplo 3.2.12. El producto vectorial $\epsilon_{ijk} u^{[j} v^{k]} = \epsilon_{ijk} u^j v^k$ entre dos vectores contravariantes en \mathbb{R}^3 es una densidad tensorial de rango 1 y peso -1 , o *pseudovector* (covariante).

Ejercicio 3.2.11. Demuestre la afirmación precedente.

Ejemplo 3.2.13. Para x^i coordenadas Cartesianas ortonormales, es válida la relación $\epsilon\text{-}\delta$:

$$\epsilon_{ijk} \epsilon_{klm} = \delta_{il} \delta_{jm} - \delta_{im} \delta_{jl}.$$

Sin embargo esto es así sólo porque en esas coordenadas los vectores covariantes y contravariantes son idénticos y las densidades tensoriales se reducen a tensores. Escrita de esta forma la relación $\epsilon\text{-}\delta$ no sería la misma en otras coordenadas, ya que el lado izquierdo de esta igualdad transformaría como un pseudotensor de peso -2 , mientras que el lado derecho lo haría como un tensor (peso 0) si lo definimos como en el Ejercicio 3.2.9, y las componentes del tensor transformado no serían las de la delta de Kronecker.

Ejercicio 3.2.12. Demuestre la relación precedente, y muestre que en coordenadas arbitrarias

$$\epsilon^{ijk} \epsilon_{klm} = \delta^i_l \delta^j_m - \delta^i_m \delta^j_l.$$

3.2.8. Tensor métrico

Ya hemos introducido anteriormente la *métrica* a través del producto interno. Ahora veremos cómo ese concepto se adapta al contexto de las coordenadas curvilíneas. La principal diferencia estriba en que ahora la métrica puede no ser constante, sino función de punto.

Sea ds el *diferencial de arco* (o diferencial de longitud), que en coordenadas Cartesianas ortonormales satisface

$$ds^2 = dx^i \delta_{ij} dx^j.$$

Como ds representa la distancia infinitesimal entre dos puntos del espacio separados por $d\vec{x} = dx^i \vec{e}_i$, debe ser un escalar; luego en un sistema de coordenadas x'^j arbitrario deberemos tener

$$ds^2 = dx'^k \frac{\partial x^i}{\partial x'^k} \delta_{ij} \frac{\partial x^j}{\partial x'^l} dx'^l = dx'^k \frac{\partial x^i}{\partial x'^k} \frac{\partial x^i}{\partial x'^l} dx'^l = dx'^k g_{kl} dx'^l.$$

El símbolo

$$g_{ij} := \frac{\partial x^k}{\partial x'^i} \frac{\partial x^k}{\partial x'^j}, \quad (3.26)$$

donde x^k son coordenadas Cartesianas ortonormales, se conoce como el *tensor métrico*, o simplemente la *métrica*, de las coordenadas curvilíneas x'^i . En notación matricial,

$$\mathbf{g} := [g_{ij}] = \mathbf{J}^T \mathbf{J}, \quad (3.27)$$

donde \mathbf{J} es la matriz Jacobiana de la transformación de Cartesianas a curvilíneas.

Ejercicio 3.2.13. Demuestre (3.27) a partir de (3.26).

Ejercicio 3.2.14. Demuestre que si \mathbf{g} es la métrica en coordenadas arbitrarias x^i , y \mathbf{g}' es la métrica en otras coordenadas arbitrarias x'^j , entonces

$$\mathbf{g}' = \mathbf{J}^T \mathbf{g} \mathbf{J}, \quad (3.28)$$

donde \mathbf{J} es la matriz Jacobiana de la transformación de x^i a x'^j .

Del ejercicio anterior es inmediato que la métrica es un tensor simétrico⁸ dos veces covariante, y que en coordenadas Cartesianas ortonormales es δ_{ij} (es decir, $\mathbf{g} = \mathbf{I}$). También es inmediato que \mathbf{g} es no singular donde \mathbf{J} tampoco lo sea, y que debe ser definida positiva si pedimos $ds^2 > 0$.

Ejercicio 3.2.15. Demuestre las afirmaciones precedentes.

En lo que sigue también serán de relevancia la *métrica inversa* \mathbf{g}^{ij} definida como

$$[g^{ij}] := \mathbf{g}^{-1}, \quad (3.29)$$

que es un tensor simétrico dos veces contravariante, y el *determinante de la métrica*

$$g := \det[g_{ij}] = \det \mathbf{g}, \quad (3.30)$$

que es un pseudoescalar de peso 2: $g' = J^2 g$.

Ejercicio 3.2.16. Demuestre las afirmaciones precedentes, y halle una expresión para \mathbf{g}^{ij} equivalente a (3.26).

Ejemplo 3.2.14. Consideremos las coordenadas polares $x'^1 = \rho$ y $x'^2 = \theta$ del Ejemplo 3.2.1, y escribamos

$$ds^2 = dx^2 + dy^2 = g'_{11} d\rho^2 + g'_{12} d\rho d\theta + g'_{21} d\theta d\rho + g'_{22} d\theta^2.$$

⁸Si \mathbf{g} es real; si es compleja, debemos pedir que sea Hermitiana.

Substituyendo $dx = d\rho \cos \theta - \rho \operatorname{sen} \theta d\theta$ y $dy = d\rho \operatorname{sen} \theta + \rho \cos \theta d\theta$, la última expresión se reduce a

$$d\rho^2 + \rho^2 d\theta^2,$$

de donde es inmediato que

$$[g'_{ij}] = \begin{pmatrix} 1 & 0 \\ 0 & \rho^2 \end{pmatrix}.$$

Alternativamente, partiendo de

$$d\vec{s} = dx'^i \vec{e}'_i$$

y escribiendo

$$ds^2 = d\vec{s} \cdot d\vec{s},$$

es inmediato que

$$g'_{ij} = \vec{e}'_i \cdot \vec{e}'_j. \quad (3.31)$$

Si la base $\{\vec{e}_i\}$ es Cartesiana ortonormal, como en este caso, basta substituir las expresiones para los \vec{e}'_i en función de los \vec{e}_i y usar que $\vec{e}_i \cdot \vec{e}_j = \delta_{ij}$.

Ejercicio 3.2.17. Calcule la métrica en coordenadas polares por este último método. Calcúlela también usando (3.27).

Ejercicio 3.2.18. Calcule la métrica en coordenadas esféricas y cilíndricas.

Ejercicio 3.2.19. Muestre que una transformación de coordenadas correspondiente a una rotación induce una transformación de semejanza para g .

Ejercicio 3.2.20. Muestre que si *definimos* el símbolo δ_{ij} en coordenadas Cartesianas ortonormales como en el Ejercicio 3.2.9, e imponemos que transforme como un tensor de tipo $(0, 2)$, entonces δ_{ij} no es otra cosa que el tensor métrico.

Existen muchos sistemas de coordenadas curvilíneas de uso común donde, notablemente, g resulta ser diagonal aunque en general J no lo sea. De (3.31) vemos que ello ocurrirá cuando los vectores tangentes curvilíneos formen una base ortogonal (aunque ésta no será, en general, ortonormal). En ese caso hablamos de un sistema de *coordenadas curvilíneas ortogonales*. Si en particular $g = I$, estamos ante un sistema de *coordenadas ortonormales*, donde $g_{ij} = \delta_{ij}$.

Ejemplo 3.2.15. Los sistemas de coordenadas polares, esféricas y cilíndricas son ortogonales, aunque en ninguno J es diagonal, y ninguno es ortonormal.

Ejercicio 3.2.21. Verifique las afirmaciones anteriores.

Para coordenadas ortogonales, los elementos diagonales de la métrica suelen escribirse como

$$g'_{ii} =: h_i^2, \quad (3.32)$$

donde h_i son llamados *factores de escala*. Si bien los factores de escala no son ni escalares ni pseudoescalares, simplifican bastante la escritura de fórmulas en coordenadas ortogonales. En particular

$$J = h_1 \cdots h_n = \prod_{i=1}^{\dim(V)} h_i$$

aunque J no sea diagonal, ya que

$$g' = h_1^2 \cdots h_n^2 = \prod_{i=1}^{\dim(V)} h_i^2.$$

Ejercicio 3.2.22. Verifique la expresión de arriba para J .

Ejercicio 3.2.23. Calcule los factores de escala, g y J para las coordenadas polares, esféricas y cilíndricas.

3.2.9. Ascenso y descenso de índices

Vimos anteriormente cómo un producto interno puede utilizarse para crear un mapa entre un EV V y su dual V^* . Utilizando la métrica podemos introducir una versión “canónica” de esa correspondencia.

Sea un vector contravariante con componentes u^i ; sus componentes covariantes se definen como

$$u_i := g_{ij} u^j.$$

Análogamente, dado un vector covariante con componentes v_i , sus componentes contravariantes se definen como

$$v^i := g^{ij} v_j.$$

En general, dado un tensor T^{ij} , definimos $T^i_j := g_{jk} T^{ik}$, $T_{ij} := g_{ik} T^{kl} g_{lj}$, etc. Decimos así que usamos la métrica para “subir” o “bajar” índices, o sea para “traducir” de componentes covariantes a contravariantes y viceversa.

En general, por ser g_{ij} y g^{ij} *tensores*, esperaríamos que la operación de subir o bajar índices no altere el peso del tensor o densidad tensorial. Sin embargo esta afirmación debe ser usada con precaución, como veremos enseguida.

Ejemplo 3.2.16. Sea v_i un vector covariante (peso 0) y $v^i = g^{ij} v_j$ su versión contravariante. Aplicando las reglas de transformación tendremos

$$v'^i = g'^{ij} v'_j = \frac{\partial x'^i}{\partial x^k} g^{kl} \frac{\partial x'^l}{\partial x^m} v_m = \frac{\partial x'^i}{\partial x^k} g_{kl} v_l = \frac{\partial x'^i}{\partial x^k} v^k,$$

que es la regla de transformación correcta para un vector contravariante (peso 0). Es decir, aquí subir los índices no alteró el peso.

Ejercicio 3.2.24. Complete los cálculos precedentes, y repítalos para la transformación de $v_i = g_{ij} v^j$.

Ejemplo 3.2.17. Consideremos las versiones covariante (peso -1) $\epsilon_{i_1 \dots i_n}$ y contravariante (peso 1) $\epsilon^{j_1 \dots j_n}$ del símbolo de Levi-Civita. Usando la métrica para subir los índices del primero, escribiríamos

$$g^{j_1 i_1} \dots g^{j_n i_n} \epsilon_{i_1 \dots i_n},$$

cuya única componente estricta es $g^{j_1 1} \dots g^{j_n n} \epsilon_{1 \dots n} = g^{-1}$. Guiándonos por el Ejemplo 3.1.12 y el Ejercicio 3.1.11, vemos que deberíamos escribir

$$\epsilon^{j_1 \dots j_n} = g g^{j_1 i_1} \dots g^{j_n i_n} \epsilon_{i_1 \dots i_n}.$$

En este sentido, la operación de subir los índices con la métrica incrementó el peso en 2 .

Ejercicio 3.2.25. Complete los cálculos precedentes y convénzase de lo que se afirma. Hágalo también bajando los índices de $\epsilon^{j_1 \dots j_n}$.

En el caso de coordenadas ortogonales, es inmediato que los ascensos y descensos de índices se reducen a

$$u_i = h_{\underline{i}} u^{\underline{i}}, \quad v^i := \frac{1}{h_{\underline{i}}} v_{\underline{i}},$$

y tendremos expresiones equivalentes para el ascenso y descenso de índices de tensores y densidades tensoriales.

Ejercicio 3.2.26. Dado un vector covariante con componentes $u_1 = a$, $u_2 = b$ y $u_3 = c$ en coordenadas esféricas, calcule sus componentes contravariantes.

Para el caso de coordenadas ortonormales, los factores de escala son todos iguales a 1 , y es inmediato que las componentes covariantes y contravariantes de cualquier vector o tensor coinciden. Esa es la situación a la que estábamos acostumbrados en coordenadas Cartesianas.

3.2.10. Producto escalar y norma

Utilizando la métrica para subir o bajar índices, podemos definir ahora el *producto escalar* entre dos vectores covariantes, o entre dos contravariantes, como

$$\vec{u} \cdot \vec{v} := u^i v_i = g^{ij} u_i v_j = g_{ij} u^i v^j. \quad (3.33)$$

La *norma* de un vector vendrá dada por

$$\|\vec{u}\|^2 := \vec{u} \cdot \vec{u} = u^i u_i = g^{ij} u_i u_j = g_{ij} u^i u^j. \quad (3.34)$$

De modo análogo, utilizando la métrica para subir y bajar índices, podemos ahora contraer pares de índices ambos covariantes o contravariantes, y obtener un tensor: por ejemplo, si T^{ij}_k es un tensor y queremos contraer sus dos superíndices, podemos bajar el segundo superíndice construyendo $T^i_{kl} = T^{ij}_k g_{jl}$, y a continuación contraer como

$$S_k = T^i_{ki} = T^{ij}_k g_{ji}$$

obteniendo un vector covariante. Etc.

Dadas (3.31) y (3.33) y la definición de base dual, de $\vec{u} = u^i \vec{e}_i$ y $\vec{v} = v_i \vec{e}^i$ es directo deducir que

$$\vec{e}_i = g_{ij} \vec{e}^j$$

y

$$\vec{e}^i = g^{ij} \vec{e}_j,$$

es decir, la métrica también nos permite traducir de la base tangente a la dual o viceversa.

En un sistema de coordenadas ortogonales las definiciones de producto escalar y norma se reducen a

$$\vec{u} \cdot \vec{v} = \frac{1}{h_i} u_i v_i = h_i u^i v^i, \quad (3.35)$$

y

$$\|\vec{u}\|^2 = \frac{1}{h_i} u_i u_i = h_i u^i u^i, \quad (3.36)$$

y la relación entre las bases tangente y dual viene dada por

$$\vec{e}_i = h_i \vec{e}^i,$$

Ejercicio 3.2.27. Calcule la norma del vector definido en el Ejercicio 3.2.26.

Finalmente, en un sistema de coordenadas ortonormales las bases tangente y dual son idénticas (y ortonormales, claro).

3.3. Integración en coordenadas curvilíneas

Como primera aplicación de lo aprendido en la sección precedente, veremos cómo el trabajar en coordenadas curvilíneas altera lo que estábamos habituados a hacer en coordenadas Cartesianas, en lo que hace al cálculo de integrales de volumen, de superficie y de línea. Por simplicidad consideraremos, mayormente, espacios tridimensionales.

3.3.1. La integral de volumen

En \mathbb{R}^3 y en coordenadas Cartesianas ortonormales estamos acostumbrados a escribir una integral de volumen como

$$\int dV = \iiint dx dy dz,$$

escribiendo el diferencial de volumen dV simplemente como $dx dy dz$. Sin embargo esta expresión es engañoso: debemos tener presente que geométricamente dV representa el volumen (infinitesimal) del paralelepípedo generado por tres vectores desplazamiento LI $dx \vec{e}_1$, $dy \vec{e}_2$ y $dz \vec{e}_3$, y bajo cambios de coordenadas transformará de manera muy diferente al producto de los diferenciales dx , dy y dz ; lo cual no debería sorprendernos, ya que $dx dy dz$ no es un escalar, que es lo que estamos acostumbrados a esperar de “ dV ”, y tampoco es un tensor.

Ejemplo 3.3.1. En las coordenadas esféricas del Ejercicio (3.2.2) tenemos

$$\begin{aligned} dx^1 &= \operatorname{sen} x'^2 \cos x'^3 dx'^1 + x'^1 \cos x'^2 \cos x'^3 dx'^2 - x'^1 \operatorname{sen} x'^2 \operatorname{sen} x'^3 dx'^3, \\ dx^2 &= \operatorname{sen} x'^2 \operatorname{sen} x'^3 dx'^1 + x'^1 \cos x'^2 \operatorname{sen} x'^3 dx'^2 + x'^1 \operatorname{sen} x'^2 \cos x'^3 dx'^3, \\ dx^3 &= \cos x'^2 dx'^1 - x'^1 \operatorname{sen} x'^2 dx'^2, \end{aligned}$$

y es evidente que $dx^1 dx^2 dx^3$ no será lo que esperaríamos, que de acuerdo al Cálculo Vectorial sería $(x'^1)^2 \operatorname{sen} x'^2 dx'^1 dx'^2 dx'^3$.

Ejercicio 3.3.1. Calcule $dx^1 dx^2 dx^3$ y verifique la afirmación precedente.

Dado que en coordenadas Cartesianas ortonormales sabemos que debemos obtener $dV = dx dy dz$, veamos cómo podemos llegar a este resultado mediante operaciones que podamos generalizar luego a coordenadas arbitrarias.

Dados tres vectores desplazamiento LI $d\vec{x}_{(1)}, d\vec{x}_{(2)}, d\vec{x}_{(3)}$ en \mathbb{R}^3 , el volumen dU_3 del paralelepípedo que los tiene como aristas es

$$dU_3 = d\vec{x}_{(1)} \cdot (d\vec{x}_{(2)} \times d\vec{x}_{(3)}) = \det \begin{pmatrix} d\vec{x}_{(1)} & d\vec{x}_{(2)} & d\vec{x}_{(3)} \\ \downarrow & \downarrow & \downarrow \end{pmatrix},$$

donde la matriz 3×3 tiene por columnas las componentes de los $d\vec{x}_{(i)}$. Reescribiendo el determinante en términos de estas componentes, es directo verificar que

$$dU_3 = \epsilon_{i_1 i_2 i_3} dx_{(1)}^{[i_1]} dx_{(2)}^{i_2} dx_{(3)}^{i_3} \equiv \overline{d\vec{x}_{(1)} \wedge d\vec{x}_{(2)} \wedge d\vec{x}_{(3)}},$$

es decir, dU_3 es el tensor adjunto del producto exterior de $d\vec{x}_{(1)}, d\vec{x}_{(2)}$ y $d\vec{x}_{(3)}$.

Este resultado puede ser generalizado inmediatamente para escribir el diferencial de “volumen” dU_n en \mathbb{R}^n generado por n vectores desplazamiento $d\vec{x}_{(1)}, \dots, d\vec{x}_{(n)}$ LI, como

$$dU_n = \overline{d\vec{x}_{(1)} \wedge \cdots \wedge d\vec{x}_{(n)}} \equiv \epsilon_{i_1 \dots i_n} dx_{(1)}^{[i_1]} \dots dx_{(n)}^{i_n}, \quad (3.37)$$

con $dx_{(k)}^l$, $l = 1, \dots, n$ las componentes contravariantes de cada vector desplazamiento $d\vec{x}_{(k)}$. Puede verificarse que, si construimos una matriz cuadrada $n \times n$ cuyas columnas sean las componentes de los $d\vec{x}_{(k)}$, la definición anterior hace que

$$dU_n = \det \begin{pmatrix} d\vec{x}_{(1)} & \cdots & d\vec{x}_{(n)} \\ \downarrow & & \downarrow \end{pmatrix}, \quad (3.38)$$

que es justamente el volumen del paralelepípedo subtendido por $d\vec{x}_{(1)}, \dots, d\vec{x}_{(n)}$. Podemos interpretar así la definición (3.37) como una “receta” algebraica para realizar esta construcción geométrica: El producto tensorial de los $d\vec{x}_{(k)}$ genera un tensor de rango n cuyas n^n componentes son todos los productos de la forma $dx_{(1)}^{i_1} \dots dx_{(n)}^{i_n}$, con $i_1, \dots, i_n = 1, \dots, n$. La antisimetrización selecciona de entre éstos los $n!$ productos que corresponden a i_1, \dots, i_n todos distintos, les asigna signos alternantes y divide por $n!$. Finalmente, la contracción con $\epsilon_{i_1 \dots i_n}$ produce el determinante en cuestión. Consistentemente, podemos tomar entonces (3.38) como una definición alternativa de dU_n , completamente equivalente a (3.37).

Ejemplo 3.3.2. Construyamos el diferencial de volumen en \mathbb{R}^2 (diferencial de área): El producto tensorial de $d\vec{x}_{(1)}$ y $d\vec{x}_{(2)}$ tiene componentes $dx_{(1)}^i dx_{(2)}^j$ con $i, j = 1, 2$, que podemos ordenar en la matriz

$$\begin{pmatrix} dx_{(1)}^1 dx_{(2)}^1 & dx_{(1)}^1 dx_{(2)}^2 \\ dx_{(1)}^2 dx_{(2)}^1 & dx_{(1)}^2 dx_{(2)}^2 \end{pmatrix}.$$

Antisimetrizando, podemos ordenar las componentes del tensor producto exterior $d\vec{x}_{(1)} \wedge d\vec{x}_{(2)}$ en la matriz

$$\frac{1}{2} \begin{pmatrix} 0 & dx_{(1)}^1 dx_{(2)}^2 - dx_{(1)}^2 dx_{(2)}^1 \\ dx_{(1)}^2 dx_{(2)}^1 - dx_{(1)}^1 dx_{(2)}^2 & 0 \end{pmatrix}.$$

Finalmente, contrayendo con el símbolo de Levi–Civita de rango 2 obtenemos

$$dU_2 = dx_{(1)}^1 dx_{(2)}^2 - dx_{(1)}^2 dx_{(2)}^1,$$

que es justamente el área del paralelogramo subtendido por $d\vec{x}_{(1)}$ y $d\vec{x}_{(2)}$.

Nótese que la expresión obtenida es justamente el determinante de

$$\begin{pmatrix} dx_{(1)}^1 & dx_{(2)}^1 \\ dx_{(1)}^2 & dx_{(2)}^2 \end{pmatrix}.$$

Nótese también que si $d\vec{x}_{(1)} = dx \vec{e}_1$ y $d\vec{x}_{(2)} = dy \vec{e}_2$, con $\{\vec{e}_1, \vec{e}_2\}$ una base ortonormal, esta expresión se reduce a $dU_2 = dx dy$.

Ejercicio 3.3.2. Construya las expresiones equivalentes para el diferencial de volumen en \mathbb{R}^3 , y muestre que si $d\vec{x}_{(i)} = dx^i \vec{e}_i$, resulta $dU_3 = dx^1 dx^2 dx^3$.

Generalizando ahora a coordenadas arbitrarias, resulta claro de lo anterior que el diferencial de volumen no será un escalar, sino una densidad escalar de peso -1 ; de allí la conocida regla de transformación

$$dV = J dV'.$$

También queda claro que, si queremos recuperar nuestro conocido diferencial de volumen en coordenadas Cartesianas ortogonales deberemos elegir para formarlo, de entre todos los posibles vectores desplazamiento $d\vec{x}_{(1)}, \dots, d\vec{x}_{(n)}$, aquellos de la forma $d\vec{x}_{(i)} = dx^i \vec{e}_i$; y en coordenadas arbitrarias, extenderemos esta elección y tomaremos los vectores desplazamiento a lo largo de los vectores de la base tangente. En consecuencia adoptamos como nuestra definición “standard”, en cualquier sistema de coordenadas,

$$dV := J \det \begin{pmatrix} d\vec{x}_{(1)} & \cdots & d\vec{x}_{(n)} \\ \downarrow & & \downarrow \end{pmatrix} \equiv \sqrt{g} \epsilon_{i_1 \dots i_n} dx_{(1)}^{[i_1} \dots dx_{(n)}^{i_n]}, \quad d\vec{x}_{(i)} = dx^i \vec{e}_i, \tag{3.39}$$

donde $J = \sqrt{g}$ es el determinante de la transformación de coordenadas Cartesianas ortogonales a las coordenadas x^i , y g el determinante de la métrica en estas coordenadas.

Ejemplo 3.3.3. Construyamos el diferencial de volumen en coordenadas polares. Poniendo como siempre $x'^1 = r$ y $x'^2 = \theta$, de (3.39) resulta

$$dV = J \det \begin{pmatrix} dr \vec{e}_1 & d\theta \vec{e}_2 \\ \downarrow & \downarrow \end{pmatrix} = J \det \begin{pmatrix} dr & 0 \\ 0 & d\theta \end{pmatrix} = r dr d\theta.$$

Ejercicio 3.3.3. Repita lo anterior y muestre que, en las coordenadas esféricas del Ejercicio 3.2.2,

$$dV = r^2 \sin \theta dr d\theta d\phi.$$

Ejemplo 3.3.4. Sea ρ la densidad de carga eléctrica; si la carga total Q contenida en un volumen dado debe ser un escalar (invariante ante transformaciones de coordenadas), es decir si pedimos

$$Q = \int \rho dV = \int \rho' dV',$$

vemos que ρ debe ser un pseudotensor de rango 0 y peso 1, es decir una densidad escalar de peso 1.

3.3.2. La integral de superficie

El razonamiento por el que construimos el diferencial de volumen también nos indica un camino posible para construir otros diferenciales multidimensionales. Por ejemplo, el diferencial de área orientado en \mathbb{R}^3 debe ser un “vector” normal a la superficie y de módulo igual al área del paralelepípedo subtendido por dos vectores desplazamiento $d\vec{x}_{(1)}$ y $d\vec{x}_{(2)}$, LI y tangentes a ella. Esto nos sugiere definirlo como el producto vectorial de dichos vectores desplazamiento,

$$dA_i = \left[\overline{d\vec{x}_{(1)} \wedge d\vec{x}_{(2)}} \right]_i = \epsilon_{ijk} dx_{(1)}^{[j]} dx_{(2)}^{k]},$$

que no es un vector sino un pseudovector (peso -1) covariante. El flujo de un vector contravariante a través de una superficie vendrá dado entonces por

$$\Phi = \int u^i dA_i,$$

que es fácil ver que será un pseudoescalar (peso -1); si Φ debiera ser un escalar, u^i debería ser una densidad vectorial de peso 1.

Ejemplo 3.3.5. Consideremos un fluido de densidad $\rho(\vec{x})$ que se mueve con velocidad $\vec{v}(\vec{x})$, ambos posiblemente dependientes de la posición. La *corriente de masa* se define como $\vec{j}(\vec{x}) = \rho(\vec{x})\vec{v}(\vec{x})$, y es una densidad (peso 1) vectorial contravariante, ya que la densidad de masa ρ es una densidad escalar (peso 1) y la velocidad \vec{v} es un vector contravariante. El flujo de masa a través de una superficie 2-dimensional S en \mathbb{R}^3 será entonces

$$\Phi = \int_S j^i dA_i = \int_S \rho v^i dA_i,$$

y será un escalar, como era de esperar.

Generalizando, si $\dim(V) = n$, el diferencial de “área” sobre una “superficie” de dimensión $n - 1$ será el pseudovector (peso -1) de componentes

$$dA_j = \left[\overline{d\vec{x}_{(1)} \wedge \cdots \wedge d\vec{x}_{(n-1)}} \right]_j = \epsilon_{j i_1 \dots i_{n-1}} dx_{(1)}^{[i_1} \dots dx_{(n-1)}^{i_{n-1}]}, \quad (3.40)$$

donde los vectores desplazamiento son LI, y nos permitirá calcular un pseudoescalar “flujo” de un vector contravariante. Análogamente, el diferencial de “área” sobre una “superficie” de dimensión $n - r$ será el pseudotensor de peso -1 y componentes

$$dA_{j_1 \dots j_r} = \left[\overline{d\vec{x}_{(1)} \wedge \cdots \wedge d\vec{x}_{(n-r)}} \right]_{j_1 \dots j_r} = \epsilon_{j_1 \dots j_r i_1 \dots i_{n-r}} dx_{(1)}^{[i_1} \dots dx_{(n-r)}^{i_{n-r}]},$$

e integrado con un tensor $u^{j_1 \dots j_r}$ nos dará un pseudoscalar de peso -1 .

El procedimiento indicado arriba adolece de dos fallas. La primera es que, a menos que la superficie sobre la que integramos sea una superficie coordenada, no podremos reproducir lo que hicimos con el diferencial de volumen al tomar los vectores desplazamiento como $d\vec{x}_{(i)} = dx^i \vec{e}_i$, ya que en general ningún vector de la base tangente será tangente a la superficie. La segunda es que, aún para una superficie coordinada, necesitamos contraer un vector o densidad vectorial con dA_i para integrar; ¿cómo, entonces, calcular por ejemplo el área de una superficie? Veremos que la remediación de la segunda falla nos llevará a remediar también la primera.

Ilustraremos el procedimiento para una superficie S de dimensión 2 definida en \mathbb{R}^3 .⁹ Para definirla introducimos dos parámetros u^1 y u^2 y definimos una función

$$\vec{s}(u^1, u^2) = s^i(u^1, u^2) \vec{e}_i$$

con \vec{e}_i los vectores de la base Cartesiana ortonormal de \mathbb{R}^3 , de modo que

$$S = \{\vec{x} = \vec{s}(u^1, u^2), (u^1, u^2) \in U \subset \mathbb{R}^2\}$$

define una *parametrización* de nuestra superficie.

Ejemplo 3.3.6. El parabolóide de revolución $z = x^2 + y^2$ puede parametrizarse como

$$\begin{aligned} s^1(u^1, u^2) &= u^1, \\ s^2(u^1, u^2) &= u^2, \\ s^3(u^1, u^2) &= (u^1)^2 + (u^2)^2, \end{aligned}$$

con $u^1, u^2 \in \mathbb{R}$.

Ejemplo 3.3.7. La esfera $x^2 + y^2 + z^2 = R^2$ puede parametrizarse como

$$\begin{aligned} s^1(u^1, u^2) &= R \sin u^1 \cos u^2, \\ s^2(u^1, u^2) &= R \sin u^1 \sin u^2, \\ s^3(u^1, u^2) &= R \cos u^1, \end{aligned}$$

con $u^1 \in [0, \pi]$ y $u^2 \in [0, 2\pi]$.

⁹Técnicamente, una *variedad*.

Es inmediato que $\partial\vec{s}/\partial u^1$ y $\partial\vec{s}/\partial u^2$ serán dos vectores tangentes a la superficie, LI si la parametrización es no singular. Recordando (3.31), definimos entonces la *métrica inducida* sobre S como

$$g_{ij} := \frac{\partial\vec{s}}{\partial u^i} \cdot \frac{\partial\vec{s}}{\partial u^j}. \quad (3.41)$$

Ejemplo 3.3.8. Para el parabolóide de revolución $z = x^2 + y^2$ tendremos

$$\vec{s}(u^1, u^2) = u^1\vec{e}_1 + u^2\vec{e}_2 + ((u^1)^2 + (u^2)^2)\vec{e}_3,$$

luego

$$\frac{\partial\vec{s}}{\partial u^1} = \vec{e}_1 + 2u^1\vec{e}_3, \quad \frac{\partial\vec{s}}{\partial u^2} = \vec{e}_2 + 2u^2\vec{e}_3$$

y usando $\vec{e}_i \cdot \vec{e}_j = \delta_{ij}$ tendremos la métrica inducida

$$[g_{ij}] = \begin{pmatrix} 1 + 4(u^1)^2 & 4u^1u^2 \\ 4u^1u^2 & 1 + 4(u^2)^2 \end{pmatrix}.$$

Ejercicio 3.3.4. Muestre que la métrica inducida sobre la esfera $x^2 + y^2 + z^2 = R^2$ es

$$[g_{ij}] = \begin{pmatrix} R^2 & 0 \\ 0 & R^2 \operatorname{sen}^2 u^1 \end{pmatrix}.$$

Una vez que tenemos la métrica inducida, podemos tratar a S como un *espacio métrico* por derecho propio, donde está claro de lo anterior que los parámetros u^1 y u^2 han pasado a jugar el papel de coordenadas curvilíneas sobre S ; y en él podemos definir un diferencial de área dA tal como definiríamos un diferencial de volumen en cualquier espacio de dimensión 2. Notemos sin embargo que los vectores desplazamiento

$$d\vec{s}_{(i)} := \frac{\partial\vec{s}}{\partial u^i} du^i$$

aunque son ambos tangentes a S , son vectores en \mathbb{R}^3 ; así que, aunque nos gustaría definir el diferencial de área como $d\vec{s}_{(1)} \times d\vec{s}_{(2)}$, éste también sería un (pseudo)vector en \mathbb{R}^3 . No obstante su *norma* es un (pseudo)escalar, y es justamente el área del paralelogramo subtendido por $d\vec{s}_{(1)}$ y $d\vec{s}_{(2)}$; la cuestión es, entonces, cómo construirla en base a cantidades *intrínsecas* a S . Para ello apelamos a la conocida identidad vectorial

$$\|d\vec{s}_{(1)} \times d\vec{s}_{(2)}\|^2 = \|d\vec{s}_{(1)}\|^2 \|d\vec{s}_{(2)}\|^2 - |d\vec{s}_{(1)} \cdot d\vec{s}_{(2)}|^2,$$

usamos (3.33) y (3.34) para reescribir esto como

$$\left(\left\| \frac{\partial\vec{s}}{\partial u^1} \right\|^2 \left\| \frac{\partial\vec{s}}{\partial u^2} \right\|^2 - \left| \frac{\partial\vec{s}}{\partial u^1} \cdot \frac{\partial\vec{s}}{\partial u^2} \right|^2 \right) (du^1 du^2)^2 = (g_{11}g_{22} - g_{12}^2)(du^1 du^2)^2,$$

y finalmente obtenemos

$$dA := \sqrt{g} du^1 du^2, \quad (3.42)$$

donde g es el determinante de la métrica inducida. Notemos la exacta correspondencia con la definición (3.39) de dV ; y, como ocurría con dV , el dA así definido no es un escalar, sino una densidad escalar.

Ejemplo 3.3.9. Para el paraboloide de revolución $z = x^2 + y^2$ tenemos

$$g = (1 + 4(u^1)^2)(1 + 4(u^2)^2) - 16(u^1 u^2)^2,$$

luego

$$dA = \sqrt{1 + 4(u^1)^2 + 4(u^2)^2} du^1 du^2.$$

Ejercicio 3.3.5. Muestre que para la esfera $x^2 + y^2 + z^2 = R^2$ se obtiene

$$dA = R^2 \sin u^1 du^1 du^2.$$

Ejemplo 3.3.10. El área del paraboloide de revolución $z = x^2 + y^2$ desde $x = y = -1$ hasta $x = y = 1$ será

$$A = \int dA = \int_{-1}^1 du^1 \int_{-1}^1 du^2 \sqrt{1 + 4(u^1)^2 + 4(u^2)^2} \simeq 7,45.$$

La integral no es trivial, y aquí se la ha aproximado numéricamente. Nótese que esta es el área de un parche de paraboloide cuya proyección sobre el plano xy es un *cuadrado*, no un círculo, y que la altura z del borde no es uniforme. Para obtener el área desde $z = 0$ hasta $z = 1$, véase el ejercicio a continuación.

Ejercicio 3.3.6. Reparametrice el paraboloide de revolución como

$$\begin{aligned} s^1(u^1, u^2) &= u^1 \cos u^2, \\ s^2(u^1, u^2) &= u^1 \sin u^2, \\ s^3(u^1, u^2) &= (u^1)^2, \end{aligned}$$

con $u^1 \in [0, \infty]$ y $u^2 \in [0, 2\pi]$, y muestre que

$$\frac{\partial \vec{s}}{\partial u^1} = \cos u^2 \vec{e}_1 + \sin u^2 \vec{e}_2 + 2u^1 \vec{e}_3, \quad \frac{\partial \vec{s}}{\partial u^2} = -u^1 \sin u^2 \vec{e}_1 + u^1 \cos u^2 \vec{e}_2,$$

$$[g_{ij}] = \begin{pmatrix} 1 + 4(u^1)^2 & 0 \\ 0 & (u^1)^2 \end{pmatrix}, \quad dA = \sqrt{1 + 4(u^1)^2} u^1 du^1 du^2,$$

y que el área del parche desde $r = 0$ hasta $r = 1$ es $(5\sqrt{5} - 1)\pi/6 \simeq 5,33$.

Ejercicio 3.3.7. Suponga que el paraboloide del ejercicio anterior tiene una densidad superficial de masa $\rho = \rho_0 z$. Calcule la masa total del paraboloide $\int \rho dA$ para $0 \leq z \leq 1$.

Ejercicio 3.3.8. Muestre que el área total de la esfera $x^2 + y^2 + z^2 = R^2$ es $A = 4\pi R^2$, y calcule su masa total si su densidad superficial de masa es $\rho \sin u^1$.

El lector habrá notado que en el camino para obtener dA hemos descartado la posibilidad de definir un diferencial de área *orientado*; sin embargo lo hicimos sólo porque deseábamos obtener un diferencial de área “escalar”. Si definimos

$$d\vec{A} := d\vec{s}_{(1)} \times d\vec{s}_{(2)},$$

éste será justamente el diferencial de área orientado que buscamos. Recordando que $\partial\vec{s}/\partial u^1$ y $\partial\vec{s}/\partial u^2$ son dos vectores tangentes a la superficie, vemos que

$$\vec{n} := \frac{\partial\vec{s}}{\partial u^1} \times \frac{\partial\vec{s}}{\partial u^2}$$

es normal a la superficie y completa una terna LI en \mathbb{R}^3 , que podemos tomar como los vectores base de un sistema de coordenadas curvilíneas; elegantemente, esto nos proporciona un sistema curvilíneo “a medida”, donde S es una “superficie coordinada”.

Ejemplo 3.3.11. En la parametrización del Ejemplo 3.3.6

$$\vec{n} = (\vec{e}_1 + 2u^1\vec{e}_3) \times (\vec{e}_2 + 2u^2\vec{e}_3) = -2u^1\vec{e}_1 - 2u^2\vec{e}_2 + \vec{e}_3,$$

donde hemos usado que los \vec{e}_i son ortonormales para calcular sus productos vectoriales. Nótese que \vec{n} apunta “hacia adentro” del paraboloide.

Ejercicio 3.3.9. Muestre que en la parametrización del Ejercicio 3.3.6

$$\vec{n} = u^1 (-2u^1 \cos u^2 \vec{e}_1 - 2u^1 \sin u^2 \vec{e}_2 + \vec{e}_3),$$

y que es paralelo al del ejemplo previo.

Ejercicio 3.3.10. Muestre que con la parametrización del Ejemplo 3.3.7, el vector normal a la esfera es

$$\vec{n} = R^2 \sin u^1 (\sin u^1 \cos u^2 \vec{e}_1 + \sin u^1 \sin u^2 \vec{e}_2 + \cos u^1 \vec{e}_3),$$

y que apunta hacia fuera de la esfera.

Notemos que en los tres casos precedentes la norma de \vec{n} es igual al coeficiente de $du^1 du^2$ en dA , es decir, es igual a \sqrt{g} con g el determinante de la métrica inducida. Un repaso de lo hecho hasta ahora nos muestra que este es un resultado general, por lo que introduciendo los vectores en \mathbb{R}^3

$$d\vec{x}_{(1)} := \frac{\partial\vec{s}}{\partial u^1} du^1, \quad d\vec{x}_{(2)} := \frac{\partial\vec{s}}{\partial u^2} du^2$$

y el vector unitario $\hat{n} := \vec{n}/\|\vec{n}\|$, dado que $\|\vec{n}\| = \sqrt{g}$ podemos escribir

$$d\vec{A} = \hat{n} \sqrt{g} du^1 du^2 = \vec{n} du^1 du^2, \quad dA_i = \epsilon_{ijk} dx_{(1)}^j dx_{(2)}^k, \quad (3.43)$$

en consonancia con nuestra propuesta inicial.

Ejercicio 3.3.11. Verifique (3.43).

Ejemplo 3.3.12. Consideremos el campo vectorial $\vec{E} = E\vec{e}_3$ en coordenadas Cartesianas ortonormales, y calculemos su flujo a través del parche del paraboloide $z = x^2 + y^2$ definido por $0 \leq z \leq 1$. Usando la parametrización del Ejemplo 3.3.6 tenemos

$$\Phi = \int_0^1 du^1 \int_0^{2\pi} du^2 u^1 (-2u^1 \cos u^2 \vec{e}_1 - 2u^1 \sin u^2 \vec{e}_2 + \vec{e}_3) \cdot (E\vec{e}_3) = \pi E,$$

que es igual al flujo a través del disco de radio unidad en el plano xy , como era de esperar.

Ejercicio 3.3.12. Calcule el flujo del mismo campo a través de la mitad superior $0 \leq u^1 \leq \pi/2$ de la esfera del Ejemplo 3.3.7.

Para finalizar, vale la pena mencionar una forma alternativa de calcular la integral de una densidad superficial (de carga, masa, etc.) ρ_s sobre una superficie en \mathbb{R}^3 definida por la ecuación $\sigma(\vec{x}) = 0$: primero convertimos ρ_s en la densidad *volumétrica*

$$\rho = \rho_s \delta(\sigma(\vec{x})),$$

donde δ es la delta de Dirac,¹⁰ y luego integramos ρ en el volumen,

$$\int \rho_s \delta(\sigma(\vec{x})) dV.$$

Ejemplo 3.3.13. Consideremos una cáscara esférica de radio r_0 con una densidad superficial de carga uniforme $\rho_s = Q/4\pi r_0^2$. En las coordenadas esféricas del Ejercicio 3.2.2, poniendo como es usual $x^1 = r$, $x^2 = \theta$ y $x^3 = \phi$ y eligiendo $d\vec{x}_{(1)} = \vec{e}_1$, $d\vec{x}_{(2)} = \vec{e}_2$ y $d\vec{x}_{(3)} = \vec{e}_3$, el diferencial de volumen puede escribirse $dV = r^2 \sin \theta dr d\theta d\phi$, y la superficie de la cáscara vendrá definida por $\sigma(\vec{x}) = r - r_0 = 0$. En este caso tendremos

$$\int \rho_s \delta(\sigma(\vec{x})) dV = \iiint \frac{Q}{4\pi r_0^2} \delta(r - r_0) r^2 \sin \theta dr d\theta d\phi = \frac{Q}{4\pi} \iint \sin \theta d\theta d\phi = Q.$$

Ejercicio 3.3.13. Verifique todas las afirmaciones precedentes.

3.3.3. La integral de línea

Usando como guía lo desarrollado en la Sección 3.3.2, vamos a investigar ahora cómo deberíamos realizar integrales a lo largo de una curva en coordenadas curvilíneas.

Una *curva* suave en \mathbb{R}^3 vendrá dada por una función de un único parámetro t ,

$$\vec{c}(t) = c^i(t) \vec{e}_i,$$

¹⁰La delta de Dirac será tratada en detalle más adelante, en el contexto de distribuciones.

con \vec{e}_i los vectores de la base Cartesiana ortonormal de \mathbb{R}^3 , de modo que

$$C = \{\vec{x} = \vec{c}(t), t \in I \subset \mathbb{R}\}$$

define una parametrización de la curva. Cuando el parámetro t varía en dt , un punto sobre la curva variará su posición \vec{x} en

$$d\vec{x} = d\vec{c}(t) = \vec{e}_i \frac{dc^i}{dt} dt,$$

de donde es inmediato que

$$\frac{d\vec{c}}{dt} = \frac{dc^i}{dt} \vec{e}_i$$

es un vector tangente a la curva; en particular, si el parámetro t representa el tiempo, será el vector velocidad.

Ejemplo 3.3.14. Una circunferencia de radio R en el plano xy puede parametrizarse como

$$c^1(t) = R \cos t, \quad c^2(t) = R \sin t, \quad c^3(t) = 0,$$

con $t \in [0, 2\pi]$. El vector tangente es

$$\frac{d\vec{c}}{dt} = -R \sin t \vec{e}_1 + R \cos t \vec{e}_2.$$

Ejercicio 3.3.14. Una *helicoide* de radio R y paso H puede parametrizarse como

$$c^1(t) = R \cos t, \quad c^2(t) = R \sin t, \quad c^3(t) = \frac{H}{2\pi} t,$$

con $t \in \mathbb{R}$. Muestre que el vector tangente es

$$\frac{d\vec{c}}{dt} = -R \sin t \vec{e}_1 + R \cos t \vec{e}_2 + \frac{H}{2\pi} \vec{e}_3.$$

Si esto es todo lo que tenemos, es decir, si simplemente parametrizamos la curva $\vec{c}(t)$ en Cartesianas en \mathbb{R}^3 , procederemos como en el análisis vectorial usual. Definiremos entonces el *largo* de la curva para $t \in [a, b]$ como

$$\ell = \int_a^b \left\| \frac{d\vec{c}}{dt} \right\| dt = \int_a^b \sqrt{\frac{d\vec{c}}{dt} \cdot \frac{d\vec{c}}{dt}} dt = \int_a^b \sqrt{\frac{dc^i}{dt} \frac{dc^i}{dt}} dt,$$

consistentemente con la definición del largo de arco. Si $\rho(t)$ es, por ejemplo, una densidad lineal de masa, la masa de la curva entre a y b será

$$m = \int_a^b \rho(t) \sqrt{\frac{dc^i}{dt} \frac{dc^i}{dt}} dt.$$

Y si $\vec{F}(t) = f_i(t) \vec{e}^i$ es una fuerza actuante sobre un punto que recorre la curva desde a hasta b , el trabajo realizado será

$$W_{ab} = \int_a^b \vec{F}(t) \cdot \frac{d\vec{c}}{dt} dt = \int_a^b f_i(t) \frac{dc^i}{dt} dt = \int_{\vec{c}(a)}^{\vec{c}(b)} f_i dc^i.$$

Ejemplo 3.3.15. El largo de la circunferencia del Ejemplo 3.3.14 es

$$\ell = \int_0^{2\pi} \sqrt{\frac{dc^i}{dt} \frac{dc^i}{dt}} dt = \int_0^{2\pi} \sqrt{R^2 \sin^2 t + R^2 \cos^2 t} dt = 2\pi R.$$

Ejercicio 3.3.15. Muestre que el largo de una vuelta ($0 \leq t \leq 2\pi$) de la helicoide del Ejercicio 3.3.14 es $\ell = 2\pi\sqrt{R^2 + H^2/4\pi^2}$, y que el trabajo de una fuerza $\vec{F}(t) = f\vec{e}_3$ a lo largo de esa misma vuelta es $W = fH$.

La traducción, relativamente trivial, de estos resultados a coordenadas curvilineas en \mathbb{R}^3 , pasa por recordar la definición (3.2.8) del diferencial de arco en curvilineas, y *reparametrizar* la curva $\vec{c}(t)$ en las coordenadas curvilineas x'^i como

$$\vec{c}(t) = c'^i(t) \vec{e}'_i = c^k(t) \frac{\partial x'^i}{\partial x^k} \frac{\partial x^j}{\partial x'^i} \vec{e}_j = c^k(t) \vec{e}_k.$$

Usando la métrica g_{ij} de las coordenadas curvilineas x'^i , tendremos entonces el largo como

$$\ell = \int_a^b \sqrt{\frac{dc'^i}{dt} g_{ij} \frac{dc'^j}{dt}} dt,$$

la integral de una densidad escalar como

$$m = \int_a^b \rho(t) \sqrt{\frac{dc'^i}{dt} g_{ij} \frac{dc'^j}{dt}} dt,$$

y la integral de línea de un campo $\vec{F}(t) = f_i(t)\vec{e}'^i = f'_j(t)\vec{e}'^j$ como

$$W_{ab} = \int_a^b f'_j(t) \frac{dc'^j}{dt} dt = \int_{\vec{c}(a)}^{\vec{c}(b)} f'_j d\vec{c}'^j.$$

Ejemplo 3.3.16. La helicoide del Ejercicio 3.3.14 puede reparametrizarse en las coordenadas cilíndricas $x'^1 = r$, $x'^2 = \theta$, $x'^3 = z$ del Ejercicio 3.2.3 como

$$c'^1(t) = R, \quad c'^2(t) = t, \quad c'^3(t) = \frac{H}{2\pi}t,$$

y el vector tangente será

$$\frac{d\vec{c}}{dt} = \vec{e}'_2 + \frac{H}{2\pi} \vec{e}'_3,$$

con

$$\vec{e}'_1 = \cos x'^2 \vec{e}_1 + \sin x'^2 \vec{e}_2, \quad \vec{e}'_2 = -x'^1 \sin x'^2 \vec{e}_1 + x'^1 \cos x'^2 \vec{e}_2, \quad \vec{e}'_3 = \vec{e}_3.$$

Ejercicio 3.3.16. Verifique las afirmaciones precedentes, y rehaga el Ejercicio 3.3.15 con esta parametrización de la helicoide. ¿Coinciden los resultados? ¿Porqué?

Resultados menos triviales se obtienen si ahora la curva $\vec{c}(t)$ está definida sobre una superficie S dada por $\vec{s}(u^1, u^2)$ como en la Sección 3.3.2, y queremos expresar todo en términos de las “coordenadas naturales” sobre S , que son los parámetros u^1 y u^2 . En ese caso primero debemos parametrizar la superficie como

$$\vec{s}(u^1, u^2) = s^i(u^1, u^2)\vec{e}_i, \quad i = 1, 2, 3$$

como antes, y luego parametrizar la curva como

$$\vec{c}(t) = u^i(t)\frac{\partial \vec{s}}{\partial u^i}, \quad i = 1, 2$$

en términos de los vectores tangentes a S . Definiendo la métrica inducida sobre S como en (3.41), es bastante directo concluir que la longitud entre $t = a$ y $t = b$ vendrá dada por

$$\ell = \int_a^b \sqrt{\frac{du^i}{dt} g_{ij} \frac{du^j}{dt}} dt$$

y la integral de una densidad escalar $\rho(t)$ por

$$m = \int_a^b \rho(t) \sqrt{\frac{du^i}{dt} g_{ij} \frac{du^j}{dt}} dt,$$

expresiones en las cuales g_{ij} es ahora la métrica inducida.

Ejercicio 3.3.17. Para el paraboloide parametrizado en coordenadas cilíndricas del Ejercicio 3.3.6, defina la “helicoide paraboloidal” parametrizada como

$$u^1(t) = \sqrt{t}, \quad u^2(t) = 2\pi t,$$

y calcule su largo para $0 \leq t \leq 1$ (puede dejar integrales indicadas, o hacerlas numéricamente).

3.3.4. Integrales de vectores y tensores

El proverbial lector atento quizá haya notado que, hasta ahora, hemos evitado definir cosas como la integral de volumen de un vector. Discutiremos brevemente porqué no lo haremos aquí.

En coordenadas Cartesianas (ortonormales o no) con base $\{\vec{e}_i\}$, definiríamos por ejemplo la integral de volumen de un campo vectorial $\vec{E} = E^i \vec{e}_i$ como

$$\int \vec{E} dV = \int (E^i \vec{e}_i) J dx^1 \cdots dx^n = \vec{e}_i \int E^i(x^1, \dots, x^n) J dx^1 \cdots dx^n,$$

donde el hecho de que los vectores base \vec{e}_i sean *independientes de la posición* nos ha permitido sacarlos fuera de la integral. Reduciríamos así la integral del vector \vec{E} a n integrales del tipo que ya hemos estudiado y obtendríamos un “ente” que transformaría, *ante cambios entre bases Cartesianas*, como un vector o

densidad vectorial (dependiendo del carácter de \vec{E}). Está claro que, en coordenadas Cartesianas, podremos hacer lo mismo con cualquier tensor o densidad tensorial.

La dificultad de extender esto a coordenadas curvilíneas es doble. Por un lado tenemos la dificultad técnica de que ahora los vectores base *varían con la posición*, así que ya no podremos extraerlos fuera de la integral.

La segunda dificultad es de un carácter más conceptual. La integral del vector sobre el volumen dado debería tener, presumiblemente, un significado físico independiente de la base; sin embargo para trabajar con ella deberemos expresarla en componentes en una base dada. Ahora bien, si los vectores base tienen diferentes direcciones y magnitudes en distintos puntos dentro del volumen en cuestión, ¿cuál de todas esas bases deberíamos usar para expresar las componentes del resultado?

La resolución de estas ambigüedades, y la cuestión misma de cuáles o cuándo pueden o no resolverse, es parte del *cálculo en variedades*, una rama de la matemática de dificultad considerablemente mayor que la asumida para este curso. Por ello aquí no profundizaremos en esta cuestión. Baste decir que cuando el objeto a integrar termine siendo un escalar o una densidad escalar, como en el resto de esta sección, estas dificultades no se presentarán, ya que dichos entes carecen de componentes.

3.4. Derivación en coordenadas curvilíneas

En una segunda aplicación de lo aprendido en la sección 3.2, veremos cómo el trabajar en coordenadas curvilíneas también altera lo que estábamos habituados a hacer en coordenadas Cartesianas, en lo que hace al cálculo de derivadas de escalares, vectores y tensores, y en particular a la forma adoptada por los operadores diferenciales más comunes: gradiente, rotor, divergencia y Laplaciano.

3.4.1. La conexión afín de Levi–Civita

Dado un escalar ϕ , hemos visto que sus derivadas $\phi_{,i}$ son las componentes de un vector covariante, ya que transforman como tales. Sin embargo, si consideramos por ejemplo un vector contravariante con componentes u^i , que sabemos que ante una transformación de coordenadas transformarán como

$$u'^i = \frac{\partial x'^i}{\partial x^k} u^k,$$

es fácil (aunque algo engoroso) ver que sus derivadas transformarán según

$$u'^{i,j} = \frac{\partial x'^i}{\partial x^k} \frac{\partial x^l}{\partial x'^j} u^k,_l + \frac{\partial^2 x'^i}{\partial x^l \partial x^k} \frac{\partial x^l}{\partial x'^j} u^k. \quad (3.44)$$

Es evidente, entonces, que las derivadas $u'^{i,j}$ *no son* las componentes de un tensor. Esto puede parecer una sorpresa desagradable, ya que frecuentemente tendremos que trabajar con las derivadas de vectores, tensores, etc. Sin embargo,

mostraremos en seguida que en realidad esto no debería sorprendernos, y cómo tratar con esta dificultad.

La complicación proviene de que, en coordenadas curvilíneas, los vectores coordinados cambian de dirección de un punto a otro; por lo tanto, si derivamos un vector escrito como $\vec{u} = u^i \vec{e}_i$, debemos derivar no sólo sus componentes u^i , sino también los vectores \vec{e}_i . Una “receta” para derivar vectores en coordenadas curvilíneas, por tanto, podría ser la siguiente: tome las componentes u^i , multiplíquelas por los vectores base \vec{e}_i y sume; derive la combinación lineal $u^i \vec{e}_i$ (*todos* los términos); y descomponga el resultado nuevamente en la base $\{\vec{e}_i\}$ para obtener las componentes de la derivada del vector. De este complicado procedimiento al menos algo quedaría claro: en general, en coordenadas curvilíneas *las componentes de la derivada no son las derivadas de las componentes*.

Evidentemente sería mucho más cómodo obtener, de una vez y para siempre, expresiones para las componentes de la derivada de un vector en términos de las componentes mismas del vector. Para ello, consideremos de nuevo la derivada de $\vec{u} = u^i \vec{e}_i$, que escribiríamos

$$\vec{u}_{,k} = u^i_{,k} \vec{e}_i + u^j \vec{e}_{j,k}. \quad (3.45)$$

Ahora bien, $\vec{e}_{j,k}$ debe poder descomponerse en la base \vec{e}_i , es decir que deben existir coeficientes Γ_{jk}^i tales que

$$\vec{e}_{j,k} = \Gamma_{jk}^i \vec{e}_i. \quad (3.46)$$

Por lo tanto, podremos poner

$$\vec{u}_{,k} = (u^i_{,k} + \Gamma_{jk}^i u^j) \vec{e}_i. \quad (3.47)$$

La expresión de arriba define lo que se conoce como una *conexión afín*, y los coeficientes Γ_{jk}^i son conocidos como los *elementos de la conexión afín*; el origen del nombre se aclara, junto con otros puntos, en la Sección 3.4.8. Aunque parezca extraño, existe más de una posible conexión afín que permite desarrollar un cálculo diferencial consistente; sin embargo en nuestro caso, el hecho de que, en el fondo, siempre definimos las coordenadas curvilíneas partiendo de un sistema Cartesiano ortonormal nos permite especificar una conexión afín particular, conocida como *conexión de Levi-Civita*. De acuerdo a lo desarrollado en la Sección 3.2, podemos contraer la expresión (3.46) con un vector \vec{e}^l de la base dual, y usar que $\vec{e}^l \vec{e}_i = \delta^l_i$ en cualquier sistema de coordenadas para obtener

$$\Gamma_{jk}^i = \vec{e}^i \vec{e}_{j,k}. \quad (3.48)$$

Si ahora usamos (3.20) para escribir $\vec{e}_j = \vec{x}_{,j}$, es inmediato que¹¹

$$\Gamma_{jk}^i = \vec{e}^i \vec{x}_{,jk} = \vec{e}^i \vec{x}_{,kj} = \Gamma_{kj}^i,$$

es decir, los elementos de la conexión afín son *simétricos* en los subíndices jk . Además, un cálculo bastante largo permite ver que Γ_{jk}^i transforma como

$$\Gamma'_{kj}^i = \frac{\partial x'^i}{\partial x^m} \frac{\partial x^l}{\partial x'^k} \frac{\partial x^h}{\partial x'^j} \Gamma_{hl}^m - \frac{\partial^2 x'^i}{\partial x^l \partial x^h} \frac{\partial x^l}{\partial x'^k} \frac{\partial x^h}{\partial x'^j}, \quad (3.49)$$

¹¹Una coma seguida de *dos* índices indica una derivada segunda: $a_{,ij} := \partial^2 a / \partial x^i \partial x^j$.

de donde resulta obvio que los elementos de la conexión afín de Levi–Civita Γ_{jk}^i no son las componentes de un tensor.

Ejemplo 3.4.1. Para las coordenadas polares $x'^1 = r$ y $x'^2 = \theta$, de los Ejemplos 3.2.2 y 3.2.7 tenemos

$$\begin{aligned}\vec{e}'_1 &= \cos \theta \vec{e}_1 + \sin \theta \vec{e}_2, & \vec{e}'_2 &= -r \sin \theta \vec{e}_1 + r \cos \theta \vec{e}_2, \\ \vec{e}'^1 &= \cos \theta \vec{e}^1 + \sin \theta \vec{e}^2, & \vec{e}'^2 &= -\frac{1}{r} \sin \theta \vec{e}^1 + \frac{1}{r} \cos \theta \vec{e}^2\end{aligned}$$

con \vec{e}_i y \vec{e}^i los vectores base Cartesianos. Entonces

$$\vec{e}'_{1,1} = 0, \quad \vec{e}'_{1,2} = -\sin \theta \vec{e}_1 + \cos \theta \vec{e}_2 = \vec{e}'_{2,1}, \quad \vec{e}'_{2,2} = -r \cos \theta \vec{e}_1 - r \sin \theta \vec{e}_2,$$

y por lo tanto

$$\Gamma_{11}^1 = \Gamma_{12}^1 = \Gamma_{21}^1 = 0, \quad \Gamma_{22}^1 = -r, \quad \Gamma_{11}^2 = \Gamma_{22}^2 = 0, \quad \Gamma_{12}^2 = \Gamma_{21}^2 = \frac{1}{r}.$$

Ejercicio 3.4.1. Halle los elementos de la conexión afín de Levi–Civita para las coordenadas cilíndricas y esféricas.

En tres dimensiones, en principio tendremos $3^3 = 27$ elementos de la conexión afín. Aún usando la simetría en los subíndices, el número de elementos independientes sólo se reduce a 18. Afortunadamente, para el frecuente caso de coordenadas *ortogonales* es posible simplificar su cálculo, como mostraremos al final de la Sección 3.4.2.

3.4.2. Derivación covariante

Volviendo a (3.47), vemos que si u^i es un vector contravariante, la cantidad

$$u^i_{;k} := u^i_{,k} + \Gamma_{jk}^i u^j \tag{3.50a}$$

es un tensor (mixto), su *derivada covariante*. De este modo, escribiremos

$$\vec{u} = u^i \vec{e}_i \implies \vec{u}_{,k} = u^i_{;k} \vec{e}_i. \tag{3.50b}$$

Ejercicio 3.4.2. Demuestre que $u^i_{;k}$ es un tensor.

Surge naturalmente la cuestión de cómo hacer lo mismo con la derivada de un vector covariante. Podríamos repetir, paso a paso, el procedimiento que llevó a (3.50a), pero es mas sencillo notar que

$$0 = \delta_{j,k}^i = (\vec{e}^i \vec{e}_j)_{,k} = \vec{e}^i_{,k} \vec{e}_j + \vec{e}^i \vec{e}_{j,k} = \Gamma_{jk}^i + \vec{e}^i_{,k} \vec{e}_j,$$

de donde es inmediata la relación “complementaria” de (3.48)

$$\Gamma_{jk}^i = -\vec{e}^i_{,k} \vec{e}_j \tag{3.51}$$

y la expresión

$$v_{i;j} = v_{i,j} - \Gamma_{ij}^k v_k \quad (3.52a)$$

para la derivada covariante de un vector covariante. De este modo, escribiremos

$$\vec{v} = v_i \vec{e}^i \implies \vec{v}_{,j} = v_{i;j} \vec{e}^i. \quad (3.52b)$$

El procedimiento anterior puede generalizarse para obtener la derivada covariante de tensores de cualquier rango y tipo. Por ejemplo, sea T un tensor con componentes T^{ij} ; expresándolo en términos de la base producto como $T = T^{ij} \vec{e}_i \otimes \vec{e}_j$ y derivando, tendremos

$$T_{,k} = T^{ij}_{,k} \vec{e}_i \otimes \vec{e}_j + T^{ij} \vec{e}_{i,k} \otimes \vec{e}_j + T^{ij} \vec{e}_i \otimes \vec{e}_{j,k},$$

y apelando a (3.48) concluiremos que

$$T^{ij}_{,k} = T^{ij}_{,k} + \Gamma_{hk}^i T^{hj} + \Gamma_{hk}^j T^{ih},$$

con lo cual

$$T = T^{ij} \vec{e}_i \otimes \vec{e}_j \implies T_{,k} = T^{ij}_{,k} \vec{e}_i \otimes \vec{e}_j.$$

De forma análoga pueden definirse las derivadas covariantes de tensores de rango arbitrario, que quedarán

$$\begin{aligned} h_{;i} &\equiv h_{,i}, \\ v_{i;j} &= v_{i,j} - \Gamma_{ij}^k v_k, & \vec{v} &= v_i \vec{e}^i \implies \vec{v}_{,j} = v_{i;j} \vec{e}^i, \\ u^i_{,j} &= u^i_{,j} + \Gamma_{kj}^i u^k, & \vec{u} &= u^i \vec{e}_i \implies \vec{u}_{,j} = u^i_{,j} \vec{e}_i, \\ t_{ij;k} &= t_{ij,k} - \Gamma_{ik}^h t_{hj} - \Gamma_{jk}^h t_{ih}, & T &= t_{ij} \vec{e}^i \otimes \vec{e}^j \implies T_{,k} = t_{ij,k} \vec{e}^i \otimes \vec{e}^j, \quad (3.53) \\ t^i_{j;k} &= t^i_{j,k} + \Gamma_{hk}^i t^h_j - \Gamma_{jk}^h t^i_h, & T &= t^i_j \vec{e}_i \otimes \vec{e}^j \implies T_{,k} = t^i_{j,k} \vec{e}_i \otimes \vec{e}^j, \\ t^{ij}_{,k} &= t^{ij}_{,k} + \Gamma_{hk}^i t^{hj} + \Gamma_{hk}^j t^{ih}, & T &= t^{ij} \vec{e}_i \otimes \vec{e}_j \implies T_{,k} = t^{ij}_{,k} \vec{e}_i \otimes \vec{e}_j, \\ &\dots & &\dots \end{aligned}$$

Ejemplo 3.4.2. Consideremos el vector contravariante $\vec{u} = u^r \vec{e}_r + u^\theta \vec{e}_\theta$ en coordenadas polares. Las componentes de $\vec{u}_{,k}$ serán

$$\begin{aligned} u^r_{,r} &= u^r_{,r} + \Gamma_{rr}^r u^r + \Gamma_{r\theta}^r u^\theta = \frac{\partial u^r}{\partial r}, \\ u^r_{,\theta} &= u^r_{,\theta} + \Gamma_{\theta r}^r u^r + \Gamma_{\theta\theta}^r u^\theta = \frac{\partial u^r}{\partial \theta} - r u^\theta, \\ u^\theta_{,r} &= u^\theta_{,r} + \Gamma_{rr}^\theta u^r + \Gamma_{r\theta}^\theta u^\theta = \frac{\partial u^\theta}{\partial r} + \frac{1}{r} u^\theta, \\ u^\theta_{,\theta} &= u^\theta_{,\theta} + \Gamma_{\theta r}^\theta u^r + \Gamma_{\theta\theta}^\theta u^\theta = \frac{\partial u^\theta}{\partial \theta} + \frac{1}{r} u^r, \end{aligned}$$

con lo cual

$$\begin{aligned} \vec{u}_{,r} &= \frac{\partial u^r}{\partial r} \vec{e}_r + \left(\frac{\partial u^\theta}{\partial r} + \frac{1}{r} u^\theta \right) \vec{e}_\theta, \\ \vec{u}_{,\theta} &= \left(\frac{\partial u^r}{\partial \theta} - r u^\theta \right) \vec{e}_r + \left(\frac{\partial u^\theta}{\partial \theta} + \frac{1}{r} u^r \right) \vec{e}_\theta. \end{aligned}$$

Ejercicio 3.4.3. Calcule \vec{u}_r y \vec{u}_θ usando (3.45) y las expresiones para la base tangente en polares del Ejemplo 3.2.2, y verifique que coinciden con las expresiones de arriba.

Ejercicio 3.4.4. Para las coordenadas esféricas $x'^1 = r$, $x'^2 = \theta$, $x'^3 = \phi$ calcule la derivada covariante de $\vec{u} = r \sin \theta \vec{e}_\phi$.

Ejercicio 3.4.5. Muestre que $\delta_{j;k}^i = 0$.

Un hecho que frecuentemente pasa desapercibido es que la derivada covariante cumple la regla de Leibniz para derivar productos de funciones, como si fuese una derivada usual: por ejemplo

$$(v_i u_j)_{;k} = (v_i u_j)_{,k} - \Gamma_{ik}^h v_h u_j - \Gamma_{jk}^h v_i u_h = v_{i;k} u_j + v_i u_{j;k} .$$

Ejercicio 3.4.6. Complete los cálculos precedentes, y muestre que también $(v^i u_j)_{;k} = v^i_{;k} u_j + v^i u_{j;k}$ y $(v^i u^j)_{;k} = v^i_{;k} u^j + v^i u^j_{;k}$. ¿Qué ocurre con $(v^i u_i)_{;k}$? ¿Porqué?

Un tensor particularmente interesante desde el punto de vista de la derivada covariante, es el tensor métrico g_{ij} . Recordando (3.31), notemos primero que

$$g_{ij,k} = (\vec{e}_i \cdot \vec{e}_j)_{,k} = \vec{e}_{i,k} \cdot \vec{e}_j + \vec{e}_i \cdot \vec{e}_{j,k} = \Gamma_{ik}^l \vec{e}_l \cdot \vec{e}_j + \vec{e}_i \cdot \vec{e}_l \Gamma_{jk}^l = \Gamma_{ik}^l g_{lj} + g_{il} \Gamma_{jk}^l , \quad (3.54)$$

de donde

$$g^{mi} g_{ij,k} = g^{mi} \Gamma_{ik}^l g_{lj} + g^{mi} g_{il} \Gamma_{jk}^l = g^{mi} \Gamma_{ik}^l g_{lj} + \delta^m_l \Gamma_{jk}^l = g^{mi} \Gamma_{ik}^l g_{lj} + \Gamma_{jk}^m ,$$

es decir

$$\Gamma_{jk}^m = g^{mi} (g_{ij,k} - \Gamma_{ik}^l g_{lj}) .$$

Pero ya vimos que $\Gamma_{jk}^m = \Gamma_{kj}^m$, de modo que podemos poner igualmente

$$\Gamma_{jk}^m = g^{mi} (g_{ik,j} - \Gamma_{ij}^l g_{lk}) .$$

Promediando las dos últimas expresiones, entonces, obtenemos

$$\Gamma_{jk}^m = \frac{1}{2} g^{mi} (g_{ij,k} + g_{ik,j} - (\Gamma_{ik}^l g_{lj} + \Gamma_{ij}^l g_{lk})) ,$$

y usando (3.54) resulta

$$\Gamma_{jk}^m = \frac{1}{2} g^{mi} (g_{ij,k} + g_{ik,j} - g_{kj,i}) , \quad (3.55)$$

que nos permite calcular los elementos de la conexión afín a partir de la métrica, y que podemos tomar como una definición de éstos alternativa a (3.48) y (3.51).¹² Lo mas notable es que si ahora calculamos la derivada covariante $g_{ij;k}$ de la métrica, usando (3.54) obtenemos

$$g_{ij;k} = g_{ij,k} - \Gamma_{ik}^h g_{hj} - \Gamma_{jk}^h g_{ih} = 0 ,$$

¹²De hecho, esta es la definición habitual de la conexión afín de Levi–Civita. Es de destacar que la derivada covariante pueda definirse, como aquí, enteramente en términos de la métrica, aunque si recordamos que la métrica justamente *mide* (permite construir longitudes, productos escalares, etc.), y la derivación involucra el límite de diferencias entre puntos cercanos, no debería extrañarnos.

que poniendo $G = g_{ij} \vec{e}^i \otimes \vec{e}^j$ dice

$$G_{,k} = 0.$$

Es decir, el tensor métrico es “constante” pese a que sus componentes g_{ij} son en general funciones de las coordenadas. Junto con la regla de Leibniz esto nos permite, por ejemplo, escribir cosas como

$$v_{i;k} = (g_{ij} v^j)_{;k} = g_{ij} v^j_{;,k},$$

es decir, extraer la métrica fuera de una derivada covariante o introducirla en ella. Dicho de otra forma, la derivación covariante conmuta con el ascenso y descenso de índices.

Ejercicio 3.4.7. Complete la demostración de la última afirmación, demostrando que también $g^{ij}_{;,k} = 0$. (Ayuda: derive $\delta^i_{,k} = g^{ij} g_{jk}$.)

Es de destacar que la derivada covariante y la derivada usual de un escalar coinciden. Sin embargo, si una cantidad H es una densidad escalar de peso p , esto ya no es así. Para entender lo que ocurre, recordemos que en ese caso $H' = J^p H$; derivando tendremos

$$\begin{aligned} H'_{,i} &= \frac{\partial H'}{\partial x^i} = \frac{\partial x^j}{\partial x'^i} \frac{\partial (J^p H)}{\partial x^j} = \frac{\partial x^j}{\partial x'^i} \left(J^p \frac{\partial H}{\partial x^j} + p J^{p-1} \frac{\partial J}{\partial x^j} H \right) \\ &= J^p \frac{\partial x^j}{\partial x'^i} (H_{,j} + p (\ln J)_{,j} H), \end{aligned}$$

de donde es inmediato que, para que las componentes de $H_{,i}$ transformen como las de una densidad tensorial de peso p , debemos definir

$$H_{,i} = H_{,i} - p (\ln J)_{,i} H.$$

Para evaluar el término $(\ln J)_{,i}$, derivemos el determinante g de la métrica: de (3.55) es fácil ver que

$$\Gamma_{jk}^k = \frac{1}{2} g^{ki} g_{ik,j} = \frac{1}{2} \delta^k_l g^{li} g_{ik,j} = \frac{1}{2} \text{Tr}(g^{-1} g_{,j}),$$

que por la fórmula de Jacobi¹³ $\det(g^{-1}) \det(g)_{,j} = \text{Tr}(g^{-1} g_{,j})$ resulta en

$$\Gamma_{jk}^k = \frac{1}{2} \frac{g_{,j}}{g} = \frac{1}{2} (\ln g)_{,j} = (\ln \sqrt{g})_{,j} = (\ln J)_{,j}.$$

¹³La fórmula de Jacobi para la derivada de una matriz es un resultado clásico, válido para cualquier matriz no singular, pero su demostración cae fuera del alcance de este curso. Sin embargo para la métrica, que es una matriz real, simétrica y definida positiva, están garantizadas la diagonalizabilidad y la positividad de sus autovalores λ_i ; la fórmula de Jacobi, siendo una relación entre determinantes y trazas, preserva su validez ante cambios de base, y ello permite construir una demostración muy breve apelando a la base donde g es diagonal, pese a que esta base cambie con el parámetro respecto del cual derivamos. Ya que $\det g = \prod_{i=1}^n \lambda_i$ y $\text{Tr } g = \sum_{i=1}^n \lambda_i$, dado cualquier parámetro t del que g dependa,

$$\partial_t \det g = \sum_{j=1}^n \prod_{i \neq j} \lambda_i \partial_t \lambda_j = \det g \sum_{j=1}^n \frac{\partial_t \lambda_i}{\lambda_i} = \det g \text{Tr}(g^{-1} \partial_t g).$$

Luego, la derivada covariante de una densidad escalar H de peso p quedará

$$H_{;i} = H_{,i} - p\Gamma_{ki}^k H. \quad (3.56)$$

El mismo procedimiento, aplicado a densidades vectoriales y tensoriales de peso p , muestra que su derivada covariante resulta la misma de un tensor (peso 0) de igual rango y tipo, más un término extra debido al carácter de densidad; reescribimos entonces (3.53) como

$$\begin{aligned} h_{;i} &= h_{,i} - p\Gamma_{ki}^k h, \\ v_{i;j} &= v_{i,j} - \Gamma_{ij}^k v_k - p\Gamma_{kj}^k v_i, \\ u^i_{\ ;j} &= u^i_{\ ,j} + \Gamma_{kj}^i u^k - p\Gamma_{kj}^k u^i, \\ t_{ij;k} &= t_{ij,k} - \Gamma_{ik}^h t_{hj} - \Gamma_{jk}^h t_{ih} - p\Gamma_{hk}^h t_{ij}, \\ t^i_{\ j;k} &= t^i_{\ j,k} + \Gamma_{hk}^i t^h_{\ j} - \Gamma_{jk}^h t^i_{\ h} - p\Gamma_{hk}^h t^i_{\ j}, \\ t^{ij}_{\ \ ;k} &= t^{ij}_{\ ,k} + \Gamma_{hk}^i t^{hj} + \Gamma_{hk}^j t^{ih} - p\Gamma_{hk}^h t^{ij}, \\ &\dots \end{aligned} \quad (3.57)$$

y estas expresiones son completamente generales, ya que para peso 0 recuperamos (3.53).

Ejemplo 3.4.3. Un ejemplo de aplicación interesante es calcular la derivada covariante del determinante de la métrica g que, recordemos, es un pseudoescalar de peso 2. De (3.57), poniendo $p = 2$ vemos que

$$g_{;i} = g_{,i} - 2\Gamma_{ki}^k g = g_{,i} - 2\frac{1}{2}\frac{g_{,i}}{g}g = g_{,i} - g_{,i} = 0,$$

es decir, el determinante de la métrica tiene derivada covariante nula; esto es consistente con que, como habíamos visto, $g_{ij;k} = 0$.

Ejercicio 3.4.8. Muestre que también $(g^{-1})_{;i} = 0$, es decir, la derivada covariante del determinante de la métrica inversa también se anula, consistentemente con $g^{ij}_{\ \ ;k} = 0$. (Ayuda: recuerde que g^{-1} es un pseudoescalar de peso -2 .)

Para el frecuente caso de coordenadas ortogonales, el cálculo de los elementos de la conexión afín puede simplificarse bastante. Por un lado, como la métrica es diagonal, (3.55) se reduce inmediatamente a

$$\Gamma_{jk}^m = \frac{1}{2}g^{mm}(g_{mj,k} + g_{mk,j} - g_{kj,m}),$$

de donde es evidente que, para que algún término del lado derecho sobreviva, debemos tener o bien $m = j$, o bien $m = k$, o bien $k = j$, o bien $m = j = k$, es decir, los únicos elementos no nulos de la conexión afín son los que tienen al menos dos índices iguales. Por otro lado, la simetría en los índices covariantes reduce aún más el número. Y por último, los lados derechos no nulos tienen

expresiones simples en términos de los factores de escala. En definitiva, sólo necesitamos calcular

$$\begin{aligned}\Gamma_{ij}^i &= (\ln h_i)_{,j}, \quad i, j = 1, 2, 3, \quad h_i \text{ dependiente de } x^j, \\ \Gamma_{jj}^i &= -\frac{(h_j^2)_{,i}}{2h_i^2}, \quad i, j = 1, 2, 3. \quad i \neq j, \quad h_j \text{ dependiente de } x^i.\end{aligned}\tag{3.58}$$

Ejercicio 3.4.9. Complete los detalles de los cálculos precedentes, y chequee (3.58) con los resultados del Ejemplo (3.4.1) y el Ejercicio (3.4.1).

Una vez que tenemos la derivada covariante, podemos introducir las definiciones correctas, y completamente generales, para los operadores diferenciales de uso más común en Física, lo que procedemos a hacer en las secciones siguientes.

3.4.3. Gradiente

Definimos el gradiente $\nabla\varphi$ de un escalar φ como el vector covariante de componentes

$$(\nabla\varphi)_i := \varphi_{,i} \equiv \varphi_{,i}.$$

Un ejemplo es el campo eléctrico, gradiente del potencial (escalar) en electrostática; otro son las fuerzas conservativas en mecánica, también gradientes de un potencial.

Si φ es un pseudoescalar de peso p , aparecen dos diferencias: por un lado, obtendremos un pseudovector; por otro lado, y más importante, la derivada covariante ya no será equivalente a la derivada usual, sino que será

$$(\nabla\varphi)_i = \varphi_{,i} = \varphi_{,i} - p\Gamma_{hi}^h\varphi.$$

Si queremos ahora definir el gradiente de un vector o tensor, o de una densidad tensorial de peso p , la extensión de las definiciones previas es obvia: simplemente, ¡tomamos su derivada covariante! Es decir, (3.57) ya nos provee (las componentes de) los gradientes de cualquier vector, tensor o densidad. Y al igual que con el gradiente de un escalar, que es un vector (covariante), el gradiente de un tensor será otro tensor de rango superior, con un (sub)índice adicional.

Ejemplo 3.4.4. El potencial electrostático de un dipolo $\vec{p} = p\hat{z}$ situado en el origen, expresado en coordenadas esféricas $x^1 = r$, $x^2 = \theta$, $x^3 = \phi$, es

$$\varphi(\vec{x}) = \frac{p}{4\pi\epsilon_0} \frac{\cos\theta}{r^2},$$

y es un escalar. Luego, su campo eléctrico $\vec{E} = E_r\vec{e}^r + E_\theta\vec{e}^\theta + E_\phi\vec{e}^\phi$ viene dado por

$$E_r = -\varphi_{,r} = \frac{p}{2\pi\epsilon_0} \frac{\cos\theta}{r^3}, \quad E_\theta = -\varphi_{,\theta} = \frac{p}{4\pi\epsilon_0} \frac{\sin\theta}{r^2}, \quad E_\phi = -\varphi_{,\phi} = 0.$$

Ejercicio 3.4.10. Complete los cálculos del ejemplo precedente.

3.4.4. Rotor

En un EV de dimensión arbitraria, el rotor de un vector covariante u_i se define como

$$r_{ij} := u_{[i;j]} = \frac{1}{2}(u_{i;j} - u_{j;i}) = \frac{1}{2}(u_{i,j} - u_{j,i}),$$

donde la antisimetrización cancela las contribuciones de la conexión afín, permitiendo escribir el resultado en términos de derivadas ordinarias. Para tomar el rotor de un vector u^i contravariante, primero bajamos el índice con la métrica y luego derivamos, es decir

$$r_{ij} := (u^k g_{k[i];j}) = \frac{1}{2}[(u^k g_{ki})_{;j} - (u^k g_{kj})_{;i}] = \frac{1}{2}[(u^k g_{ki})_{,j} - (u^k g_{kj})_{,i}].$$

En ambos casos el rotor es un tensor covariante antisimétrico de rango 2.

En un espacio tridimensional, podemos obtener un vector (que es la definición de rotor a la que estamos habituados) tomando el tensor adjunto de r_{ij} ; tendríamos así

$$(\text{rot } \vec{u})^k = (\nabla \times \vec{u})^k := r_{ij}\epsilon^{jik} \equiv u_{i,j}\epsilon^{jik}.$$

Sin embargo esta definición,¹⁴ que es la usual y correcta en coordenadas Cartesianas ortogonales, no toma en cuenta las propiedades de transformación del símbolo de Levi–Civita si la reproducimos tal cual en coordenadas curvilíneas. Un vistazo al Ejemplo 3.1.12 nos muestra que la definición correcta es

$$\nabla \times \vec{u} := \sqrt{g^{-1}}\epsilon^{jik}u_{i,j}\vec{e}_k,$$

válida en cualquier sistema de coordenadas, y donde ϵ^{jik} es el *símbolo* de Levi–Civita; notemos que el rotor resulta ser una densidad vectorial. Un ejemplo es el campo magnético, rotor de un potencial vectorial.

Ejemplo 3.4.5. El potencial vector de un dipolo magnético ubicado en el origen se escribe

$$\vec{A} = \frac{\mu_0}{4\pi} \frac{\vec{m} \times \vec{x}}{\|\vec{x}\|^3},$$

donde \vec{m} es el momento magnético del dipolo y μ_0 es la permeabilidad magnética del vacío. Suponiendo sin pérdida de generalidad que $\vec{m} = m\hat{z}$, en coordenadas Cartesianas ortonormales tenemos

$$\vec{A} = \frac{\mu_0 m}{4\pi} \frac{x\hat{y} - y\hat{x}}{(x^2 + y^2 + z^2)^{3/2}},$$

donde \hat{x} , \hat{y} y \hat{z} son los versores Cartesianos; en coordenadas esféricas, esta expresión se convierte en

$$\vec{A} = \frac{\mu_0 m}{4\pi} \frac{1}{r^3} \vec{e}_\phi = \frac{\mu_0 m}{4\pi} \frac{\sin^2 \theta}{r} \vec{e}_\phi.$$

¹⁴Esta es la definición que aparece en el texto de Santaló.

Luego, el campo magnético del dipolo queda

$$\vec{B} = \nabla \times \vec{A} = \frac{\mu_0 m}{4\pi} \frac{1}{r^3} \left(2 \cos \theta \vec{e}_r + \frac{\sin \theta}{r} \vec{e}_\theta \right).$$

Ejercicio 3.4.11. Complete los cálculos precedentes. También calcule \vec{B} tomando el rotor de \vec{A} en Cartesianas, convierta a esféricas, y compare.

Ejercicio 3.4.12. Para el campo \vec{B} del último ejemplo, muestre que $\nabla \times \vec{B} = \vec{0}$.

3.4.5. Divergencia

Definimos la divergencia de un vector contravariante u^i como

$$\operatorname{div} \vec{u} = \nabla \cdot \vec{u} := u^i_{;i},$$

que es un escalar. Otra forma de escribirla es explicitando la derivada covariante, como

$$\nabla \cdot \vec{u} = u^i_{,i} + \Gamma^i_{ik} u^k = u^i_{,i} + (\ln \sqrt{g})_{,k} u^k = u^i_{,i} + \frac{(\sqrt{g})_{,i}}{\sqrt{g}} u^i = \frac{1}{\sqrt{g}} (\sqrt{g} u^i)_{,i}.$$

Para un vector covariante u_i , primero subimos el índice con la métrica (pasándolo a un vector contravariante) y después derivamos,

$$\operatorname{div} \vec{u} = \nabla \cdot \vec{u} := (g^{ij} u_j)_{,i} = \frac{1}{\sqrt{g}} (\sqrt{g} g^{ij} u_j)_{,i},$$

obteniendo también un escalar; la última expresión se obtiene, como para un vector contravariante, explicitando la derivada covariante, y usando la relación

$$g^{ij} \Gamma^k_{ij} = -\frac{(\sqrt{g} g^{kl})_{,l}}{\sqrt{g}}.$$

Ejercicio 3.4.13. Verifique la relación anterior. (Advertencia: este ejercicio es largo.) (Ayuda: recuerde que $g_{,i}/g = g^{jk} g_{jk,i}$; derive $\delta^i_j = g^{ik} g_{kj}$ para obtener una relación entre $g^{ij}_{,k}$ y $g_{ij,k}$ y use (3.55) para expresar Γ^k_{ij} .)

Ejemplo 3.4.6. La divergencia del potencial vector \vec{A} del Ejemplo 3.4.5 es

$$\nabla \cdot \vec{A} = \frac{1}{\sqrt{g}} (\sqrt{g} A^i)_{,i} = \frac{1}{r^2 \sin \theta} (r^2 \sin \theta A^\phi)_{,\phi} = 0.$$

Ejercicio 3.4.14. Complete los cálculos precedentes. Calcule también $\nabla \cdot \vec{B}$.

Podemos extender la definición de la divergencia a tensores de cualquier rango; sin embargo, a diferencia de lo que ocurre con un vector, existe una posible ambigüedad: con cuál de los índices del tensor contraemos el único índice del operador ∇ . Por ejemplo, para un tensor T contravariante de rango 2 podemos

definir $T^{ij}_{;i}$ o $T^{ij}_{;j}$ y, en general, serán objetos (vectores contravariantes) diferentes; la excepción son los tensores simétricos: si $T^{ij} = T^{ji}$ es evidente que ambas definiciones coinciden, y podremos escribir $\nabla \cdot T$ sin ambigüedad.

Por último, notemos que la divergencia implica una *contracción de índices*, por lo que aplicada a un tensor de rango k resulta en uno de rango $k - 1$, con un índice contravariante menos; en consecuencia, no tiene sentido aplicada a un escalar.

3.4.6. Laplaciano

Definimos el Laplaciano de un tensor como la divergencia de su gradiente. Aplicado a un escalar φ , el gradiente produce un vector covariante $\varphi_{;i}$; luego, para tomar su divergencia debemos subir su índice con la métrica formando $g^{ji}\varphi_{;i}$ y recién entonces contraer el nuevo superíndice con una segunda derivada covariante; el resultado es de nuevo un escalar,

$$\nabla^2\varphi = \nabla \cdot \nabla\varphi := (g^{ji}\varphi_{;i})_{;j} \equiv g^{ji}\varphi_{;ij},$$

donde la última expresión se obtiene extrayendo la métrica fuera de la derivada covariante en virtud de $g^{ij}_{;k} = 0$. Explicitando la derivada covariante como hicimos con la divergencia, obtenemos la expresión equivalente

$$\nabla^2\varphi = \frac{1}{\sqrt{g}}(\sqrt{g}g^{ij}\varphi_{,j})_{,i}.$$

Ejemplo 3.4.7. El Laplaciano del potencial dipolar electrostático del Ejemplo 3.4.4 es

$$\nabla^2\varphi = \frac{1}{r^2 \operatorname{sen} \theta} ((r^2 \operatorname{sen} \theta g^{rr}\varphi_{,r})_{,r} + (r^2 \operatorname{sen} \theta g^{\theta\theta}\varphi_{,\theta})_{,\theta} + (r^2 \operatorname{sen} \theta g^{\phi\phi}\varphi_{,\phi})_{,\phi}) = 0.$$

Ejercicio 3.4.15. Complete los cálculos precedentes, y chequéelos calculando la divergencia del campo eléctrico obtenido en el Ejemplo 3.4.4.

Debemos notar que el Laplaciano es, en sí mismo y por definición, un *operador escalar*, que no altera el carácter tensorial (rango y tipo) del objeto al que lo apliquemos. En consecuencia podemos aplicarlo a un tensor sin la ambigüedad mencionada en el caso de la divergencia, ya que el índice de la segunda derivada covariante (la divergencia) se contraerá *siempre* con el índice introducido por la primera (el gradiente). Escribimos entonces

$$\nabla^2 t^{k_1 \dots k_r}_{h_1 \dots h_s} := (g^{ji}t^{k_1 \dots k_r}_{h_1 \dots h_s;i})_{;j} \equiv g^{ji}t^{k_1 \dots k_r}_{h_1 \dots h_s;ij}.$$

3.4.7. Componentes físicas de un vector

Quizá el lector se descubra algo confuso al advertir que, una vez que podemos subir y bajar índices con la métrica, un vector contravariante como ser la velocidad, con componentes v^i , puede ser convertido en un vector covariante

con componentes $v_i = g_{ij}v^j$. Es evidente que, salvo en coordenadas Cartesianas ortonormales, donde $g_{ij} \equiv \delta_{ij}$, los valores numéricos de v^i y v_i pueden ser muy diferentes. Si ahora se le pide que *mida* la componente de la velocidad del móvil en la dirección de \vec{e}_i , ¿qué contestaría?

Ejemplo 3.4.8. Supongamos que, en coordenadas polares $x^1 = r$ y $x^2 = \theta$, la velocidad de un móvil viene dada por $\vec{v} = \dot{r}\vec{e}_r + \dot{\theta}\vec{e}_\theta$. Como en coordenadas Cartesianas, hallaremos que \dot{r} viene expresada en, digamos, m/s; pero $\dot{\theta}$ vendrá dada en radianes/s ($^{\circ}/s$); ¿cómo podemos sumarlas? Lo que ocurre es que en coordenadas polares, la información sobre *longitudes* viene codificada en el vector tangente \vec{e}_θ , que en términos de los vectores base (versores) Cartesianos se escribe $-\dot{r} \sin \theta \vec{e}_x + \dot{r} \cos \theta \vec{e}_y$ y cuya norma es $\|\vec{e}_\theta\| = r$, con dimensiones de longitud. En cambio $\|\vec{e}_r\| = 1$, que es adimensional. El resultado es que tanto $\dot{r}\vec{e}_r$ como $\dot{\theta}\vec{e}_\theta$ tienen dimensiones de velocidad. Si queremos codificar la información de longitud en las *componentes*, de modo que ambas tengan las mismas dimensiones, deberemos definir los *versores* $\hat{e}_r := \vec{e}_r/\|\vec{e}_r\| = \vec{e}_r$ y $\hat{e}_\theta := \vec{e}_\theta/\|\vec{e}_\theta\| = \vec{e}_\theta/r$, y redefinir las componentes como $v^r := \dot{r}\|\vec{e}_r\| = \dot{r}$ y $v^\theta := \dot{\theta}\|\vec{e}_\theta\| = r\dot{\theta}$. Notemos que la base formada por \vec{e}_r y \vec{e}_θ era sólo ortogonal, mientras que la base formada por \hat{e}_r y \hat{e}_θ es *ortonormal*, y podemos considerarla como una base Cartesiana ortonormal (local).

El ejemplo anterior muestra que la aparente ambigüedad proviene del hecho que ni las componentes covariantes ni las contravariantes de un vector son lo que usualmente identificamos como sus componentes *geométricas* o *físicas*; estas últimas son las proyecciones (geométricas) del vector sobre las direcciones tangentes a las curvas coordinadas, que son las direcciones de los vectores de la base tangente \vec{e}_i , y son componentes en la base de los *versores* $\hat{e}_i := \vec{e}_i/\|\vec{e}_i\|$. Podemos aprovechar esta idea si notamos que la norma de un vector, que usualmente identificamos como su “largo” (el módulo de la velocidad, por ejemplo), *sí* tiene un valor independiente de si es covariante o contravariante:

$$\|\vec{v}\| = \sqrt{\vec{v} \cdot \vec{v}} = \sqrt{v^i v_i} = \sqrt{g^{ij} v_i v_j} = \sqrt{g_{ij} v^i v^j}.$$

Una breve reflexión nos muestra que lo que llamariamos componentes físicas de un vector vienen dadas por

$$U^i := \sqrt{g_{ii}} u^i, \quad V_i := \sqrt{g^{ii}} v_i.$$

Estas cantidades *no son* componentes de vectores covariantes ni contravariantes.

Ejemplo 3.4.9. Dado que en coordenadas polares $g = \begin{pmatrix} 1 & 0 \\ 0 & r^2 \end{pmatrix}$, vemos que las componentes físicas del vector velocidad $\vec{v} = \dot{r}\vec{e}_r + \dot{\theta}\vec{e}_\theta$ serán $V^r = \dot{r}$ y $V^\theta = r\dot{\theta}$, consistentemente con el Ejemplo previo.

En términos de las componentes físicas, las expresiones de los operadores diferenciales usuales quedan:

Gradiente:

$$\nabla \Phi := \sqrt{g^{ii}} \Phi_{,i} \hat{e}_i.$$

Rotor:

$$\nabla \times \vec{V} := \frac{\sqrt{g^{kk}}}{\sqrt{g}} \left(\frac{V_i}{\sqrt{g^{ii}}} \right)_{,j} \epsilon^{jik} \hat{e}_k.$$

Divergencia:

$$\nabla \cdot \vec{U} := \frac{1}{\sqrt{g}} \left(\frac{\sqrt{g}}{\sqrt{g^{ii}}} U^i \right)_{,i}.$$

Laplaciano:

$$\nabla^2 \Phi := \frac{1}{\sqrt{g}} (\sqrt{g} g^{ij} \Phi_{,j})_{,i}.$$

Nótese que en estas expresiones los versores \hat{e}_i tampoco son covariantes ni contravariantes; de hecho, $\hat{e}^i \equiv \hat{e}_i$, por lo que podemos usar indistintamente unos u otros.

Estas expresiones pueden escribirse de forma más sencilla (o al menos, más habitual) para el caso de coordenadas ortogonales, usando el hecho de que los factores de escala son justamente $h_i = \sqrt{g_{ii}} = 1/\sqrt{g^{ii}}$, que $g_{ij} = \delta_{ij} h_i^2$ y $g^{ij} = \delta_{ij}/h_i^2$, y que las direcciones de los vectores de la base tangente \vec{e}_i , de la base dual \vec{e}^i y de los versores \hat{e}_i coinciden, con lo cual los versores forman efectivamente una *base Cartesiana ortonormal local* (una copia rotada de la base Cartesiana ortonormal, cuya orientación depende del punto). Obtenemos así

$$\nabla \varphi = \frac{1}{h_i} \varphi_{,i} \hat{e}_i, \quad (3.59a)$$

$$\nabla \times \vec{U} = \epsilon_{ijk} \frac{h_i}{J} (h_k U_k)_{,j} \hat{e}_i, \quad (3.59b)$$

$$\nabla \cdot \vec{U} = \frac{1}{J} \left(\frac{J}{h_i} U_i \right)_{,i}, \quad (3.59c)$$

$$\nabla^2 \varphi = \frac{1}{J} \left(\frac{J}{h_i^2} \varphi_{,i} \right)_{,i}, \quad (3.59d)$$

donde $J = h_1 h_2 h_3$ es el determinante Jacobiano del cambio de Cartesianas a curvilíneas. Nótese nuevamente que aquí los índices *no* indican covarianza o contravarianza, si bien seguimos sumando sobre índices repetidos.

Ejercicio 3.4.16. Usando (3.59), calcule explícitamente el gradiente, el rotor, la divergencia y el Laplaciano en coordenadas cilíndricas y esféricas.

Ejercicio 3.4.17. Usando (3.59), repita los cálculos de los Ejemplos 3.4.4, 3.4.5, 3.4.6 y 3.4.7, y de los Ejercicios 3.4.10, 3.4.11, 3.4.12, 3.4.14 y 3.4.15.

3.4.8. Cálculo en variedades

Capítulo 4

Distribuciones y espacios \mathcal{L}^2

En este capítulo expondremos, tan sucintamente como sea posible, la teoría matemática de los espacios de funciones que encontraremos del Capítulo 5 en adelante. Veremos que muchas de ellas no son, en mayor o menor medida, lo que identificamos como “funciones” en cursos elementales de análisis; sin embargo estas extensiones del concepto de función son indispensables para el tratamiento adecuado de los espacios de soluciones de ecuaciones diferenciales ordinarias y parciales.

4.1. Distribuciones

En la primera mitad del Siglo XX, el físico británico Paul Adrien Maurice Dirac¹ introdujo una “función” $\delta(x)$ que se suponía debía satisfacer

$$\delta(x) = \begin{cases} 0, & x \neq 0, \\ \infty (?), & x = 0, \end{cases} \quad \text{“} \int_{-\infty}^{\infty} f(x)\delta(x) dx = f(0) \text{”}$$

y tener otras propiedades igualmente bizarras. Dirac llamó a estas entidades “funciones impropias” y afirmaba que, aunque no siempre podamos asignarles valores bien definidos, cuando aparecían bajo un signo integral el resultado *sí* era bien definido.

El problema con expresiones como éstas es que no sabemos cómo definir “”, ya que el límite de la suma de Riemann no está bien definido: depende de cómo lo tomemos. Sin embargo, la observación de Dirac sobre el comportamiento de δ bajo un signo integral fué la base de la teoría rigurosa desarrollada por el matemático francés Laurent Schwartz en 1950, que describiremos sucintamente en esta sección.

¹Uno de los padres de la Mecánica Cuántica, a quien sus colegas apodaban jocosamente PAM.

4.1.1. Funciones generalizadas

La idea central de la teoría de las distribuciones, también llamadas funciones generalizadas, es la siguiente: para una función $f(x)$, si tomamos como argumento un número x , la función devuelve otro número; para una distribución f , definimos un funcional lineal que, cuando tomamos como argumento una función $\varphi(x)$ (de cierta clase especial conocida como funciones de prueba) devuelve un número. No podemos tomar un x_0 y hacer que f le asigne un número, pero podemos hacer que f asigne un número a una función de prueba φ que tenga un pico muy alto en x_0 y se anule fuera de un intervalo muy estrecho que lo contenga, y esto basta para definir completamente f .

Funciones de prueba

Definimos una *función de prueba (test function)* como una $\varphi : \mathbb{R}^n \rightarrow \mathbb{C}$ tal que

1. $\varphi(\vec{x}) \in \mathcal{C}^\infty[\mathbb{R}^n]$.
2. Dos alternativas:
 - a) clase $\mathcal{C}_0^\infty[\mathbb{R}^n]$ (de *soporte acotado*), que significa que $\forall \varphi \exists a < \infty$ tal que $\varphi(\vec{x}) = 0$ si $\|\vec{x}\| \geq a$;
 - b) clase $\mathcal{S}[\mathbb{R}^n]$ (de *Schwartz*), que significa que $x_1^{p_1} \cdots x_n^{p_n} \varphi(\vec{x}) \xrightarrow[\|\vec{x}\| \rightarrow \infty]{} 0$
 $\forall p_1, \dots, p_n \in \mathbb{N}$.

Ejemplo 4.1.1.

$$\mathcal{C}_0^\infty[\mathbb{R}] \ni \varphi(x) = \begin{cases} e^{-\frac{1}{1-x^2}}, & |x| < 1, \\ 0, & |x| \geq 1. \end{cases}$$

Ejercicio 4.1.1. Muestre que todas las derivadas de la φ definida arriba son continuas para todo x .

Ejemplo 4.1.2.

$$\mathcal{S}[\mathbb{R}] \ni \varphi(x) = e^{-x^2}.$$

Ejemplo 4.1.3.

$$\mathcal{S}[\mathbb{R}^3] \ni \varphi(x) = e^{-x^2-y^2-z^2}.$$

Propiedades básicas

Tanto $\mathcal{C}_0^\infty[\mathbb{R}^n]$ como $\mathcal{S}[\mathbb{R}^n]$ son *espacios lineales*, en principio con cuerpo \mathbb{R} o quizás \mathbb{C} .

Si $\varphi(\vec{x}) \in \mathcal{C}_0^\infty[\mathbb{R}^n]$ y $f(\vec{x}) \in \mathcal{C}^\infty[\mathbb{R}^n]$, entonces $f(\vec{x})\varphi(\vec{x}) \in \mathcal{C}_0^\infty[\mathbb{R}^n]$.

Si $\varphi(\vec{x}) \in \mathcal{S}[\mathbb{R}^n]$ y $f(\vec{x}) \in \{p_N(\vec{x})\}$, entonces $f(\vec{x})\varphi(\vec{x}) \in \mathcal{S}[\mathbb{R}^n]$.

Si $\varphi(\vec{x})$ está en $\mathcal{C}_0^\infty[\mathbb{R}^n]$ o $\mathcal{S}[\mathbb{R}^n]$, entonces todas sus derivadas parciales de orden finito también lo están.

Ejercicio 4.1.2. Demuestre las propiedades anteriores.

Ejercicio 4.1.3. Muestre que $\mathcal{C}_0^\infty[\mathbb{R}^n]$ es un subespacio de $\mathcal{S}[\mathbb{R}^n]$.

4.1.2. Funcionales lineales

Dado un espacio lineal $V = \{\varphi : \mathbb{R}^n \rightarrow \mathbb{C}\}$, un *funcional* es cualquier aplicación F de V al cuerpo escalar:

$$V \ni \varphi \xrightarrow{F} F[\varphi] \in \mathbb{C}. \quad (4.1)$$

En otras palabras, un funcional es una *forma* sobre V .

Definición

Diremos que un funcional es *lineal* si

$$F[a\varphi + b\psi] \equiv aF[\varphi] + bF[\psi],$$

donde a y b son escalares; es decir, es una *forma lineal*.

Ejemplo 4.1.4. Dada $f \in \mathcal{C}^\infty(\mathbb{R})$ y $\varphi \in \mathcal{C}_0^\infty(\mathbb{R})$,

$$F[\varphi] = \int_{-\infty}^{\infty} f(x)\varphi(x) dx$$

es un funcional lineal de $\mathcal{C}_0^\infty(\mathbb{R})$ en \mathbb{C} .

Dada $f \in \mathcal{C}^\infty(\mathbb{R}^n)$ y $\varphi \in \mathcal{C}_0^\infty(\mathbb{R}^n)$ o $\mathcal{S}(\mathbb{R}^n)$, definiremos el funcional lineal asociado a f como

$$F[\varphi] = \langle f, \varphi \rangle := \int_{-\infty}^{\infty} \cdots \int f(\vec{x})\varphi(\vec{x}) d^n x. \quad (4.2)$$

Propiedades fundamentales

El funcional lineal asociado a f satisface las siguientes propiedades:

1. Es *bilineal*: $\langle af + bg, \varphi \rangle \equiv a\langle f, \varphi \rangle + b\langle g, \varphi \rangle$.
2. $\langle \psi f, \varphi \rangle \equiv \langle f, \psi \varphi \rangle$.
3. $\langle f^*, \varphi \rangle \equiv \langle f, \varphi^* \rangle^*$.

Diremos que el funcional $\langle f, \varphi \rangle$ es *real* si devuelve un número real para toda φ real.

4.1.3. Distribuciones

Definimos una *distribución* $f(\vec{x})$ como un funcional lineal

$$\langle f, \cdot \rangle : \mathcal{C}_0^\infty(\mathbb{R}^n) \rightarrow \mathbb{C} \quad (4.3a)$$

o

$$\langle f, \cdot \rangle : \mathcal{S}(\mathbb{R}^n) \rightarrow \mathbb{C} \quad (4.3b)$$

con una condición de continuidad que discutiremos enseguida.

Debe advertirse que la notación $f(\vec{x})$ es sólo eso, y tiene por único objetivo explicitar la variable independiente de las funciones de prueba; no implica para nada que podamos tratar a $f(\vec{x})$ como una función de la variable \vec{x} y, de hecho, en general no lo será: justamente, la definición de la distribución $f(\vec{x})$ no pasa por poder asignar valores a f para cualquier valor dado de \vec{x} , sino por poder asignar valores a $\langle f, \varphi \rangle$ para cualquier función de prueba φ .

Convergencia de Schwartz

Consideremos una sucesión de funciones de prueba φ_j que converja a otra ψ ; queremos definir una noción de *continuidad* para $\langle f, \cdot \rangle$, exigiendo la continuidad del funcional respecto de la noción de convergencia que elijamos para φ ; es decir, requerimos

$$\langle f, \varphi_j \rangle \rightarrow \langle f, \psi \rangle \quad \text{siempre que } \varphi_j \rightarrow \psi. \quad (4.4)$$

Cuanto mas “fuerte” el modo de convergencia de φ_j a ψ , menores serán las restricciones impuestas a la clase de funcionales por la condición de continuidad.

El modo de convergencia más fuerte, denotado por \mathcal{D} , es la *continuidad de Schwartz*. Dadas $\psi, \varphi_j \in \mathcal{C}_0^\infty(\mathbb{R}^n)$, diremos que

$$\varphi_j \xrightarrow{\mathcal{D}} \psi \quad (4.5)$$

siempre que:

1. $\exists B \subset \mathbb{R}^n$ acotado, t.q. $\varphi_j(\vec{x}) = 0 \forall \vec{x} \notin B, \forall j$; es decir, B contiene los soportes de todas las φ_j .
2. $\varphi_j(\vec{x}) \xrightarrow{j \rightarrow \infty} \psi(\vec{x})$ uniformemente en $\vec{x} \forall \vec{x} \in B$.
3. Toda derivada parcial de φ_j converge de igual modo a la correspondiente derivada parcial de ψ .

En ese caso, el requerimiento (4.4) hace del funcional $\langle f, \cdot \rangle$ una *distribución* (de Schwartz) sobre \mathbb{R}^n .

Modos de convergencia más débiles para φ_j llevan a clases más “leves” de distribuciones. En particular, dadas $\psi, \varphi_j \in \mathcal{C}_0^\infty(\mathbb{R})$ o $\psi \in \mathcal{S}(\mathbb{R})$, decimos que

$$\varphi_j \xrightarrow{\mathcal{S}} \psi \quad (4.6a)$$

siempre que, para todo p y para todo k ,

$$\sup_{(x)} \left| x^p \left(\varphi_j^{(k)}(x) - \psi^{(k)}(x) \right) \right| \xrightarrow{j \rightarrow \infty} 0. \quad (4.6b)$$

En ese caso, el requerimiento (4.4) hace del funcional $\langle f, \cdot \rangle$ una *distribución temperada* sobre \mathbb{R} ; la generalización a distribuciones temperadas sobre \mathbb{R}^n es obvia.

Debe notarse que, para funciones en \mathcal{C}_0^∞ , $\varphi_j \xrightarrow{\mathcal{D}} \psi$ implica $\varphi_j \xrightarrow{\mathcal{S}} \psi$; en consecuencia, si f es una distribución temperada, la restricción de $\langle f, \cdot \rangle$ a \mathcal{C}_0^∞ es una distribución de Schwartz. Las distribuciones temperadas cobrarán relevancia en el Capítulo 7, en el contexto de Transformadas de Fourier.

Yendo al extremo, un modo de convergencia aún más débil, la *convergencia en norma*, lleva a los elementos de los espacios \mathcal{L}^2 con los que trataremos en la Sección 4.2; como distribuciones, éstos son tan “leves” que a menudo los llamamos simplemente “funciones”.

4.1.4. Ejemplos de distribuciones

Vamos a dar ahora varios ejemplos de distribuciones de uso corriente en Física. Notemos que, una vez definido el funcional lineal $\langle f, \cdot \rangle$, somos libres de escribir, si conviene, cosas como

$$\langle f, \varphi \rangle = \int_{-\infty}^{\infty} f(x) \varphi(x) dx,$$

sin que ello implique que el lado derecho esté bien definido como límite de una suma de Riemann; simplemente, una vez que sabemos que f es una distribución y φ una función de prueba, *interpretamos* el lado derecho como el funcional del lado izquierdo, del cual *sí* pedimos las propiedades que sean deseables.

La delta de Dirac

Ahora ya podemos definir la delta de Dirac $\delta(x)$ con cierta confianza de obtener resultados consistentes. Para ello definimos un funcional $\langle \delta, \cdot \rangle$ con la propiedad

$$\langle \delta, \phi \rangle = \phi(0) \quad \forall \phi \in \mathcal{C}_0^\infty[\mathbb{R}] \text{ o } \mathcal{S}[\mathbb{R}]. \quad (4.7)$$

Escribiendo, con un ligero abuso de notación,

$$\langle \delta(x), \phi(x) \rangle = \int_{-\infty}^{\infty} \delta(x) \phi(x) dx,$$

vemos que, *si* $\delta(x)$ fuera una función, deberíamos tener

$$\langle \delta(x-a), \phi(x) \rangle = \int_{-\infty}^{\infty} \delta(x-a) \phi(x) dx = \int_{-\infty}^{\infty} \delta(x) \phi(x+a) dx = \langle \delta(x), \phi(x+a) \rangle,$$

donde hemos usado las reglas habituales de cambio de variable de integración. Vemos entonces que es conveniente exigir la propiedad adicional

$$\langle \delta(x-a), \phi(x) \rangle = \phi(a) \quad \forall \phi \in \mathcal{C}_0^\infty[\mathbb{R}] \text{ o } \mathcal{S}[\mathbb{R}], \quad (4.8)$$

donde, por supuesto, el soporte de ϕ debe incluir el punto a . Para evitar el abuso de notación, podemos tomar esto como la definición de un funcional $\langle \delta_a, \cdot \rangle$ con la propiedad $\langle \delta_a, \phi \rangle = \phi(a)$, y adoptar la convención de que $\langle \delta, \cdot \rangle = \langle \delta_0, \cdot \rangle$.

El mismo procedimiento nos permite definir cosas como $\delta(ax + b)$: de nuevo, cambiando variables bajo la integral como si δ fuera una función, vemos que es conveniente definir

$$\langle \delta(ax + b), \phi(x) \rangle = \frac{1}{a} \phi(-b/a) \quad \forall \phi \in \mathcal{C}_0^\infty[\mathbb{R}] \text{ o } \mathcal{S}[\mathbb{R}]. \quad (4.9)$$

Un resultado importante, que usaremos en la Sección 5.1.3 y más adelante, es que $(x - x_0)\delta(x - x_0)$ es la *distribución nula*: dada cualquier función de prueba φ , tendremos

$$\langle (x - x_0)\delta(x - x_0), \varphi(x) \rangle = \langle \delta(x - x_0), (x - x_0)\varphi(x) \rangle = 0 \varphi(x_0) = 0.$$

Derivadas de la delta de Dirac

Ahora queremos definir una distribución $\delta'(x)$. Debe remarcarse que, en principio, podríamos definirla como quisiéramos; sin embargo, usando como guía la expresión

$$\langle \delta', \phi \rangle = \int_{-\infty}^{\infty} \delta'(x)\phi(x) dx$$

vemos que, si $\delta(x)$ fuese una función diferenciable, podríamos integrar por partes para obtener

$$\langle \delta', \phi \rangle = \delta(x)\phi(x)|_{-\infty}^{\infty} - \int_{-\infty}^{\infty} \delta(x)\phi'(x) dx,$$

y el término integrado por partes se anularía por ser ϕ de soporte acotado o rápidamente decreciente con $|x|$. Esto sugiere fuertemente definir $\langle \delta', \cdot \rangle$ como

$$\langle \delta', \phi \rangle := -\langle \delta, \phi' \rangle = -\phi'(0) \quad \forall \phi \in \mathcal{C}_0^\infty[\mathbb{R}] \text{ o } \mathcal{S}[\mathbb{R}]. \quad (4.10)$$

El mismo razonamiento nos lleva a definir la derivada n -ésima $\delta^{(n)}$ como

$$\langle \delta^{(n)}, \phi \rangle := (-1)^n \langle \delta, \phi^{(n)} \rangle = (-1)^n \phi^{(n)}(0) \quad \forall \phi \in \mathcal{C}_0^\infty[\mathbb{R}] \text{ o } \mathcal{S}[\mathbb{R}]. \quad (4.11)$$

Con esto ya podemos definir cosas como

$$\langle a\delta + b\delta', \phi \rangle = a\phi(0) - b\phi'(0),$$

etc.

La delta de Dirac multidimensional

Consideremos una función de prueba $\varphi(\vec{x}) = \varphi(x, y, z)$ en \mathbb{R}^3 , donde x, y, z son coordenadas Cartesianas ortonormales. Si convenimos que

$$\langle \delta(z), \varphi(x, y, z) \rangle = \varphi(x, y, 0),$$

podemos definir

$$\langle \delta(\vec{x}), \varphi(\vec{x}) \rangle = \langle \delta(x), \langle \delta(y), \langle \delta(z), \varphi(x, y, z) \rangle \rangle \rangle,$$

y escribir sin ambigüedad

$$\langle \delta(\vec{x}), \varphi(\vec{x}) \rangle = \varphi(\vec{0}) \quad \forall \varphi \in \mathcal{C}_0^\infty[\mathbb{R}^3] \text{ o } \mathcal{S}[\mathbb{R}^3]. \quad (4.12)$$

El mismo razonamiento de antes nos sugiere escribir, consistentemente,

$$\delta(\vec{x}) = \delta(x)\delta(y)\delta(z). \quad (4.13)$$

Análogamente, definiremos

$$\langle \delta(\vec{x} - \vec{x}_0), \varphi(\vec{x}) \rangle = \varphi(\vec{x}_0) \quad \forall \varphi \in \mathcal{C}_0^\infty[\mathbb{R}^3] \text{ o } \mathcal{S}[\mathbb{R}^3], \quad (4.14)$$

y

$$\langle \partial_x^n \partial_y^m \partial_z^l \delta(\vec{x}), \varphi(\vec{x}) \rangle = (-1)^{n+m+l} \partial_x^n \partial_y^m \partial_z^l \varphi(\vec{x}_0) \quad \forall \varphi \in \mathcal{C}_0^\infty[\mathbb{R}^3] \text{ o } \mathcal{S}[\mathbb{R}^3], \quad (4.15)$$

etc.

Distribuciones de carga

Dada una distribución volumétrica de carga eléctrica, la densidad de carga $\rho(\vec{x})$ es una función ordinaria,² sin embargo, para cargas superficiales, lineales o puntuales la densidad volumétrica diverge en la superficie, curva o punto en cuestión, y debemos definirla como una distribución.

Ejemplo 4.1.5. Consideremos una carga puntual q en el origen de \mathbb{R}^3 , inmersa en un potencial electrostático $\varphi(\vec{x})$ que sea una función de prueba, es decir $\varphi \in \mathcal{C}_0^\infty[\mathbb{R}^3]$ o $\mathcal{S}[\mathbb{R}^3]$. Su energía potencial electrostática U , que debe ser un escalar, será por definición $q\varphi(\vec{0})$. Notando que si $\rho(\vec{x})$ fuese una función (es decir, si tuviésemos una distribución volumétrica de carga), la energía se calcularía como

$$U = \int \rho(\vec{x})\phi(\vec{x}) d^3x,$$

vemos que debemos definir

$$\rho(\vec{x}) = q\delta(\vec{x}), \quad (4.16)$$

de modo de tener

$$U = \langle \rho(\vec{x}), \varphi(\vec{x}) \rangle = \int q\delta(\vec{x})\phi(\vec{x}) d^3x = q\varphi(\vec{0}).$$

Ejemplo 4.1.6. Consideremos un dipolo eléctrico de momento \vec{p} en el origen de \mathbb{R}^3 , inmerso en un potencial $\varphi(\vec{x})$; su energía potencial será, como sabemos,

$$U = -\vec{p} \cdot \nabla \varphi(\vec{0}).$$

El mismo razonamiento del ejemplo anterior, entonces, nos lleva a definir

$$\rho(\vec{x}) = \vec{p} \cdot \nabla \delta(\vec{x}). \quad (4.17)$$

²Una densidad escalar en el sentido del Capítulo 3.

Ejemplo 4.1.7. Sea ahora un cuadrupolo Q^{ij} en el origen de \mathbb{R}^3 , como siempre en un potencial $\varphi(\vec{x})$; su energía potencial será, en Cartesianas,

$$U = Q^{ij} \partial_i \partial_j \varphi(\vec{0})$$

y, por el mismo razonamiento de antes, definiremos

$$\rho(\vec{x}) = Q^{ij} \partial_i \partial_j \delta(\vec{x}). \quad (4.18a)$$

Para definir ρ en coordenadas arbitrarias, debemos recordar que U debe ser un *escalar*, y por lo tanto escribiríamos

$$U = Q^{ij} \varphi_{;ij}(\vec{0}) = Q^{ij} \left(\varphi_{,ij}(\vec{0}) - \Gamma_{ij}^k \varphi_{,k}(\vec{0}) \right).$$

Pondremos entonces

$$\rho(\vec{x}) = Q^{ij} (\partial_i \partial_j \delta(\vec{x}) - \Gamma_{ij}^k \partial_k \delta(\vec{x})). \quad (4.18b)$$

Ejemplo 4.1.8. Consideremos una superficie S suave y cerrada en \mathbb{R}^3 , y una *densidad superficial de carga* $\rho_s(\vec{x})$, que es una función continua de \vec{x} sobre S . Como siempre tomamos un potencial electrostático $\varphi(\vec{x})$ que sea una función de prueba, y queremos que la energía potencial electrostática *total*

$$U = \int_S \rho_s(\vec{x}) \varphi(\vec{x}) dA$$

sea un escalar, donde la integral se extiende a $\vec{x} \in S$ y dA es el diferencial de área escalar sobre S . Vemos que

$$\langle \rho, \varphi \rangle = \int_S \rho_s(\vec{x}) \varphi(\vec{x}) dA \quad \forall \varphi \in \mathcal{C}_0^\infty[\mathbb{R}^3] \text{ o } \mathcal{S}[\mathbb{R}^3] \quad (4.19)$$

define una distribución $\rho(\vec{x})$ sobre \mathbb{R}^3 , que se comporta como una delta de Dirac definida en cada punto de S .

Ejemplo 4.1.9. Con la misma superficie S del ejemplo previo, consideremos una densidad superficial de momentos dipolares $\rho_p(\vec{x})$ en la dirección de \hat{n} , la normal exterior unitaria a S . Entonces

$$\langle \rho, \varphi \rangle = \int_S \rho_p(\vec{x}) \hat{n} \cdot \nabla \varphi(\vec{x}) dA \quad \forall \varphi \in \mathcal{C}_0^\infty[\mathbb{R}^3] \text{ o } \mathcal{S}[\mathbb{R}^3] \quad (4.20)$$

define una distribución $\rho(\vec{x})$ sobre \mathbb{R}^3 que se comporta, en cada punto de S , como la derivada de la delta de Dirac en la dirección de \hat{n} .

4.1.5. Sucesiones de distribuciones

Si $f(x)$ es una función en el sentido usual, por ejemplo $f \in \mathbb{C}^\infty[\mathbb{R}]$, está claro que

$$\langle f, \varphi \rangle = \int_{-\infty}^{\infty} f(x) \varphi(x) dx$$

será una distribución si $\varphi \in \mathcal{C}_0^\infty[\mathbb{R}]$, y también si $\varphi \in \mathcal{S}[\mathbb{R}]$ siempre que f crezca a lo sumo polinomialmente para $|x| \rightarrow \infty$. Como en ese caso podemos aplicar las reglas del cálculo usual a la integral, surge naturalmente la cuestión de si podremos definir, por ejemplo, $\langle \delta, \cdot \rangle$ como el límite de una sucesión de funcionales de este tipo, ya que las propiedades a exigir de, digamos, $\langle \delta', \cdot \rangle$ surgirían “por herencia”, sin necesidad de imponerlas por separado.

Formalmente diremos que si $\{f_n\}$ es una sucesión de funciones tal que el límite de $\int_{-\infty}^{\infty} f_n(x)\varphi(x) dx$ existe para toda función de prueba φ , entonces la sucesión de distribuciones $\langle f_n, \varphi \rangle$ converge a la distribución

$$\langle f, \varphi \rangle := \lim_{n \rightarrow \infty} \int_{-\infty}^{\infty} f_n(x)\varphi(x) dx. \quad (4.21)$$

Más generalmente, si $\{f_n\}$ es una sucesión de distribuciones, y f es otra distribución tal que $\langle f_n, \varphi \rangle \rightarrow \langle f, \varphi \rangle$ para toda función de prueba φ , entonces diremos que $f_n \rightarrow f$.

Ejemplo 4.1.10. Consideremos una familia de funciones $f_\nu(x)$ de soporte acotado (pero no \mathcal{C}_0^∞) definida como

$$f_\nu(x) = \begin{cases} \nu, & |x| \leq 1/2\nu, \\ 0, & |x| > 1/2\nu. \end{cases}$$

Entonces es inmediato que

$$\lim_{\nu \rightarrow 0} \langle f_\nu, \varphi \rangle = \lim_{\nu \rightarrow 0} \int_{-\infty}^{\infty} f_\nu(x)\varphi(x) dx = \lim_{\nu \rightarrow 0} \nu \int_{-1/2\nu}^{1/2\nu} \varphi(x) dx = \varphi(0)$$

para toda φ en \mathcal{C}_0^∞ o \mathcal{S} , con lo cual tenemos

$$f_\nu \xrightarrow{\nu \rightarrow 0} \delta.$$

Ejercicio 4.1.4. Considere la sucesión de funciones en $\mathcal{C}^\infty[\mathbb{R}]$

$$f_n = \frac{n}{\sqrt{\pi}} e^{-n^2 x^2}.$$

Muestre que

$$f_n \xrightarrow{n \rightarrow \infty} \delta.$$

¿A qué converge f'_n ? ¿Y $f_n^{(k)}$? (Ayuda: note que toda función de prueba en $\mathcal{C}_0^\infty[\mathbb{R}]$ o $\mathcal{S}[\mathbb{R}]$ es entera, y por consiguiente tiene un desarrollo de Taylor con radio de convergencia infinito.)

Ejercicio 4.1.5. Muestre que

$$\frac{1 - \cos nx}{n\pi x^2} \xrightarrow{n \rightarrow \infty} \delta.$$

4.1.6. Distribuciones regulares y singulares

Decimos que una distribución $\langle f, \cdot \rangle$ es *regular* si f es una función localmente integrable.³ En la práctica podemos reducir este requerimiento a:

1. $f(x) \in \mathcal{C}[\mathbb{R}]$ salvo a lo sumo en un número finito de discontinuidades finitas.
2. $f(x)$ es integrable, es decir, $\langle f, \varphi \rangle = \int_{-\infty}^{\infty} f(x)\varphi(x) dx$ existe y es finita para toda función de prueba φ .

Toda distribución $\langle f, \cdot \rangle$ con $f(x)$ singular como función, se dice *singular*; por ejemplo δ , δ' , etc. son distribuciones singulares, aunque pueden obtenerse como límite de sucesiones de distribuciones regulares. La extensión de estas definiciones a \mathbb{R}^n es directa.

Ejemplo 4.1.11. Consideremos la función $f(x) = 1/x$, singular en $x = 0$, y definamos

$$f_\nu(x) = \begin{cases} 1/x, & |x| > \nu, \\ 0, & |x| \leq \nu, \end{cases}$$

que es una sucesión de funciones regulares que converge a f para $\nu \rightarrow 0$, *como funciones*. Definamos las distribuciones regulares $\langle f_\nu, \cdot \rangle$ como

$$\langle f_\nu, \varphi \rangle = \int_{-\infty}^{\infty} f_\nu(x)\varphi(x) dx.$$

Luego

$$\langle f, \varphi \rangle := \lim_{\nu \rightarrow 0} \langle f_\nu, \varphi \rangle = \lim_{\nu \rightarrow 0} \int_{|x|>\nu} \frac{\varphi(x)}{x} dx = \text{f} \frac{\varphi(x)}{x} dx,$$

donde f indica el valor principal de Cauchy. Esto *define* la distribución singular $\langle 1/x, \cdot \rangle$.

El resultado es completamente general, y podemos enunciarlo como el siguiente

Teorema 4.1.1. *Toda distribución f es el límite de una sucesión f_n de distribuciones regulares.*⁴

4.1.7. Derivación e integración

Derivación

Dada una distribución f sobre \mathbb{R} , su derivada f' se *define* como

$$\langle f', \varphi \rangle := -\langle f, \varphi' \rangle \quad \forall \varphi \in \mathcal{C}_0^\infty[\mathbb{R}] \text{ o } \mathcal{S}[\mathbb{R}]. \quad (4.22)$$

³Una función $f(\vec{x})$ es localmente integrable si $\int_M f(\vec{x}) d^n x$ es finita para todo $M \subset \mathbb{R}^n$ compacto.

⁴En otras palabras, el conjunto de distribuciones regulares es *denso* en el conjunto de distribuciones.

Notemos que si φ es una función de prueba, φ' también lo es. Si f es una función, la definición es simplemente el resultado usual de integrar por partes y descartar las contribuciones de superficie.

Ejemplo 4.1.12. Sea $\Theta(x)$ la función escalón de Heaviside, definida como

$$\Theta(x) := \begin{cases} 1, & x \geq 0, \\ 0, & x < 0. \end{cases}$$

Como Θ es una función continua salvo por un número finito de discontinuidades finitas, es localmente integrable, así que podemos poner

$$\langle \Theta, \varphi' \rangle = \int_{-\infty}^{\infty} \Theta(x) \varphi'(x) dx = \int_0^{\infty} \varphi'(x) dx = \varphi(x)|_0^{\infty} = -\varphi(0),$$

donde hemos usado que φ es una función de prueba para poner $\varphi(\infty) = 0$. Se deduce entonces que

$$\langle \Theta', \varphi \rangle = -\langle \Theta, \varphi' \rangle = \varphi(0),$$

y por lo tanto

$$\Theta' = \delta.$$

Ejercicio 4.1.6. Dada la función $f(x) = |x|$ definida sobre \mathbb{R} , halle las distribuciones f' y f'' .

Integración

Puede mostrarse que dada una distribución cualquiera g sobre \mathbb{R} , existe otra distribución f llamada su *primitiva*, tal que $f' = g$, y que f es única a menos de una constante aditiva.

Para verlo, dada una función de prueba $\psi(x)$ definamos

$$\varphi(x) = - \int_{-\infty}^x \psi(x') dx',$$

que es \mathcal{C}^∞ y satisface $\varphi' = -\psi$. Si $\int_{-\infty}^{\infty} \psi(x') dx' = 0$, entonces $\varphi(x)$ también es una función de prueba de la misma clase que $\psi(x)$, y simplemente definimos

$$\langle f, \psi \rangle = \langle g, \varphi \rangle.$$

Caso contrario, elegimos una función de prueba ψ_1 tal que $\int_{-\infty}^{\infty} \psi_1(x') dx' = 1$ y una constante arbitraria c , y definimos

$$\langle f, \psi_1 \rangle = c.$$

Notando que cualquier función de prueba ψ puede escribirse como

$$\psi = \psi_0 + a\psi_1, \quad \int_{-\infty}^{\infty} \psi_0(x') dx' = 0, \quad a = \int_{-\infty}^{\infty} \psi(x') dx',$$

vemos que

$$\begin{aligned}\langle f, \psi \rangle &= \langle f, \psi_0 \rangle + \langle f, a\psi_1 \rangle \\ &= \langle f, \psi_0 \rangle + a \langle f, \psi_1 \rangle \\ &= \langle f, \psi_0 \rangle + \int_{-\infty}^{\infty} \psi(x') dx' c \\ &= \langle f, \psi_0 \rangle + \langle c, \psi \rangle,\end{aligned}$$

donde el último término representa la función aditiva constante c .

4.1.8. Cambio de variables

Consideremos una función $f(\vec{x}) \in \mathcal{C}^\infty[\mathbb{R}^n]$ y una función de prueba $\varphi(\vec{x})$ en $\mathcal{C}_0^\infty[\mathbb{R}^n]$ o $\mathcal{S}[\mathbb{R}^n]$; bajo una transformación de variables $\vec{x} = \vec{x}(\vec{y})$ con matriz Jacobiana $J = \partial(\vec{x})/\partial(\vec{y})$, sabemos por la Sección 3.3 que tendremos

$$\langle f, \varphi \rangle = \int f(\vec{x}) \varphi(\vec{x}) d^n x = \int f(\vec{x}(\vec{y})) \varphi(\vec{x}(\vec{y})) J d^n y.$$

Luego, si queremos escribir

$$\langle f, \varphi \rangle = \langle g, \psi \rangle = \int g(\vec{y}) \psi(\vec{y}) d^n y$$

manteniendo la (muy razonable) definición

$$\psi(\vec{y}) = \varphi(\vec{x}(\vec{y}))$$

para la nueva función de prueba $\psi(\vec{y})$, vemos que debemos definir la nueva función $g(\vec{y})$ como

$$g(\vec{y}) = J f(\vec{x}(\vec{y})).$$

Si ahora tomamos una sucesión $f_n \rightarrow f$ de funciones $\mathcal{C}^\infty[\mathbb{R}^n]$ que converja a una distribución f , y la correspondiente sucesión $g_n \rightarrow g$, la regla de transformación para distribuciones que nos permite escribir

$$\langle f, \varphi \rangle = \langle g, \psi \rangle \tag{4.23a}$$

será

$$g(\vec{y}) = J f(\vec{x}(\vec{y})), \quad \psi(\vec{y}) = \varphi(\vec{x}(\vec{y})). \tag{4.23b}$$

Vemos entonces que, en el contexto del Capítulo 3, las distribuciones transforman como *densidades escalares* de peso 1.

Ejemplo 4.1.13. Consideremos una carga eléctrica puntual q colocada en las coordenadas Cartesianas ortonormales $(x, y, z) = (1, 1, 1)$, es decir, en $\vec{x}_0 = \hat{x} + \hat{y} + \hat{z}$; en coordenadas esféricas su posición será $r = \sqrt{3}$, $\theta = \pi/4$, $\phi = \pi/4$. La carga total calculada en Cartesianas será, muy razonablemente,

$$Q = \langle q \delta(\vec{x} - \vec{x}_0), 1 \rangle = \iiint_{-\infty}^{\infty} q \delta(x - 1) \delta(y - 1) \delta(z - 1) dx dy dz = q.$$

La misma carga, calculada en esféricas, será igualmente

$$\begin{aligned} Q &= \left\langle q\delta(r - \sqrt{3})\delta(\theta - \pi/4)\delta(\phi - \pi/4), 1 \right\rangle \\ &= \int_0^\infty r^2 dr \int_0^\pi \sin \theta d\theta \int_0^{2\pi} d\phi q \frac{\delta(r - \sqrt{3})\delta(\theta - \pi/4)\delta(\phi - \pi/4)}{r^2 \sin \theta} \\ &= q. \end{aligned}$$

Nótese que, en ausencia del factor J^{-1} en la distribución δ en esféricas, el resultado habría sido $3q/\sqrt{2}$.

4.2. Espacios \mathcal{L}^2

Los espacios \mathcal{L}^2 son probablemente el ejemplo más importante, al menos en Física, de los denominados *espacios de Hilbert*, y generalizan a espacios de dimensión infinita muchos de los conceptos habituales en espacios Euclidianos de dimensión finita. Los espacios de Hilbert resultan de definir una *topología*, es decir una noción de “distancia” entre elementos de un espacio abstracto, de funciones en nuestro caso. Para ello debemos definir una *norma* en base a un *producto interno*, que en espacios \mathcal{L}^2 se basará en integrales de las funciones involucradas. A este respecto la integral de Riemann se mostró insuficiente, y el área realmente “despegó” con la introducción de la integral de Lebesgue, que permite la integración de una clase mucho más amplia de funciones, y de la que proviene la notación “ \mathcal{L} ”; el “2” proviene de requerir que el cuadrado del módulo de una función sea integrable en el sentido de Lebesgue. Pese a lo dicho, al nivel de este curso raramente (o nunca) necesitaremos la integral de Lebesgue, y salvo nota explícita en contrario todas las integrales que aparezcan podrán (y deberán) interpretarse como de Riemann.

4.2.1. Definición y resultados básicos

Consideremos el espacio lineal de funciones de un intervalo real $[a, b]$ a valores complejos; no exigiremos por ahora condiciones de continuidad para esas funciones, sólo de integrabilidad. Para f, g en ese espacio, definimos un *producto interno* como

$$\langle f, g \rangle := \int_a^b f^*(x)g(x) dx, \quad (4.24)$$

que, como puede verificarse fácilmente, cumple todos los requerimientos de la Sección 1.6.1. En base a él definimos la *norma* de una función f como

$$\|f\| := \sqrt{\langle f, f \rangle} = \sqrt{\int_a^b |f(x)|^2 dx}. \quad (4.25)$$

Definiremos ahora

$$\mathcal{L}^2[a, b] := \{f : [a, b] \rightarrow \mathbb{C} / \|f\| < \infty\}. \quad (4.26)$$

Notemos que $\mathcal{L}^2[a, b]$, $\mathcal{L}^2(a, b]$, $\mathcal{L}^2[a, b)$ y $\mathcal{L}^2(a, b)$ son todos el mismo conjunto, por definición.

Es bastante directo demostrar la *desigualdad de Cauchy–Schwartz*

$$|\langle f, g \rangle| \leq \|f\| \|g\|, \quad (4.27)$$

y la *desigualdad triangular*

$$\|f + g\| \leq \|f\| + \|g\|. \quad (4.28)$$

Notemos que la desigualdad triangular garantiza que \mathcal{L}^2 es un espacio lineal.

La noción de *distancia* $\|f - g\|$ entre dos funciones nos permite definir la *igualdad en \mathcal{L}^2* :

$$f \stackrel{\mathcal{L}^2}{=} g \iff \|f - g\| = 0, \quad (4.29)$$

y, cuando el significado quede claro del contexto, pondremos simplemente $f = g$. Debe notarse que

$$f \stackrel{\mathcal{L}^2}{=} g \Rightarrow f(x) = g(x) \forall x \in [a, b],$$

ya que dos funciones pueden, por ejemplo, ser iguales salvo en un conjunto numerable de puntos, y sin embargo ser iguales en \mathcal{L}^2 . La noción de igualdad en \mathcal{L}^2 hace que en realidad, cuando hablamos de una $f \in \mathcal{L}^2$, no nos estemos refiriendo a *una* función $f(x)$ en el sentido usual, sino a una *clase de equivalencia*: la de aquellas funciones cuya norma es $\|f\|$. Debe notarse que lo mismo se aplica a las distribuciones definidas en la Sección 4.1.

Dos funciones $f, g \in \mathcal{L}^2[a, b]$ se dicen *ortogonales en $\mathcal{L}^2[a, b]$* si $\langle f, g \rangle = 0$.

Ejemplo 4.2.1. Las funciones $\{1, \cos(nx), \operatorname{sen}(nx), n = 1, 2, \dots\}$ son ortogonales dos a dos en $\mathcal{L}^2[-\pi, \pi]$.

Ejercicio 4.2.1. Demuestre la afirmación precedente.

Dada una función $\rho(x)$, real y positiva en (a, b) , definimos el producto interno respecto a la *función de peso* ρ como

$$\langle f, g \rangle_\rho := \int_a^b f^*(x)g(x) \rho(x) dx, \quad (4.30)$$

y la correspondiente norma como

$$\|f\|_\rho := \sqrt{\langle f, f \rangle_\rho}. \quad (4.31)$$

El producto interno $\langle f, g \rangle_\rho$ también satisface todos los requerimientos de la Sección 1.6.1, y cumple las desigualdades de Cauchy–Schwartz y triangular.

Ejercicio 4.2.2. Demuestre la última afirmación.

Denotaremos el correspondiente espacio lineal como

$$\mathcal{L}_\rho^2[a, b] := \{f : [a, b] \rightarrow \mathbb{C} / \|f\|_\rho < \infty\}. \quad (4.32)$$

Notemos que, desde el punto de vista del producto interno, $\rho(x)$ juega el papel de una métrica diagonal.

4.2.2. Convergencia

Comencemos recordando algunas nociones de convergencia ya conocidas. Dada una sucesión de funciones $f_n : [a, b] \rightarrow \mathbb{C}$, la noción más básica es la de *convergencia puntual*:

$$f_n \rightarrow f \text{ si } \lim_{n \rightarrow \infty} f_n(x) = f(x) \forall x \in [a, b]. \quad (4.33)$$

Una noción “más fuerte” de convergencia es la de *convergencia uniforme*: decimos que

$$f_n \xrightarrow{u} f \text{ si } \forall \epsilon > 0 \exists N / n \geq N \implies |f_n(x) - f(x)| < \epsilon \forall x \in [a, b], \quad (4.34)$$

es decir, si podemos acotar la diferencia $|f_n(x) - f(x)|$ por una cantidad *independiente de x*. La convergencia uniforme implica la puntual, pero no al revés. Recordemos que la convergencia de *series* de funciones puede entenderse como la convergencia de la sucesión de sumas parciales, $\sum_{n=0}^{\infty} f_n(x) = f(x)$ si $\sum_{n=0}^N f_n(x) \rightarrow f(x)$. Recordemos también que la convergencia uniforme es la que garantiza que si $f_n \rightarrow f$, entonces f_n continua implica f continua, $f'_n \rightarrow f'$ y $\int f_n \rightarrow \int f$, y es la que permite derivar o integrar series término a término.

Introduciremos ahora la noción de *convergencia en \mathcal{L}^2* : dada una sucesión de funciones $f_n \in \mathcal{L}^2(a, b)$, diremos que

$$f_n \xrightarrow{\mathcal{L}^2} f \text{ si } \lim_{n \rightarrow \infty} \|f_n - f\| = 0, \quad (4.35)$$

y diremos que f es el *límite en \mathcal{L}^2* de la sucesión de las f_n .

Ejemplo 4.2.2. Puntualmente

$$x^n \rightarrow \begin{cases} 0, & 0 \leq x < 1 \\ 1, & x = 1. \end{cases}$$

Pero

$$\|x^n - 0\| = \sqrt{\int_0^1 x^{2n} dx} = \sqrt{\frac{1}{2n+1}} \rightarrow 0,$$

así que

$$x^n \xrightarrow{\mathcal{L}^2} 0.$$

Ejemplo 4.2.3. Sea

$$f_n(x) = \begin{cases} 0, & x = 0, \\ n, & 0 < x \leq 1/n, \\ 0, & 1/n < x \leq 1. \end{cases}$$

Puntualmente

$$f_n \rightarrow 0.$$

Pero

$$\|x^n - 0\| = \sqrt{\int_0^1 f_n^2(x) dx} = \sqrt{\int_0^{1/n} n^2 dx} = \sqrt{n} \not\rightarrow 0,$$

así que f_n no converge a 0 en \mathcal{L}^2 .

Los ejemplos de arriba muestran claramente que la convergencia puntual no implica convergencia en \mathcal{L}^2 ni a la inversa. Pero puede demostrarse que si una sucesión f_n converge tanto puntualmente como en \mathcal{L}^2 , lo hace *al mismo límite*; y también puede demostrarse que la convergencia uniforme implica *ambas*.

Otra noción importante del análisis clásico que conviene recordar, es la de una *sucesión de Cauchy*. La extenderemos a \mathcal{L}^2 diciendo que una sucesión de funciones $f_n \in \mathcal{L}^2$ es de Cauchy si

$$\forall \epsilon > 0 \exists N / m, n \geq N \implies \|f_n - f_m\| < \epsilon. \quad (4.36)$$

Es claro que toda sucesión convergente en \mathcal{L}^2 es de Cauchy, ya que si $f_n \xrightarrow{\mathcal{L}^2} f$, por la desigualdad triangular

$$\|f_n - f_m\| \leq \|f_n - f\| + \|f_m - f\|,$$

y podemos hacer el lado derecho tan pequeño como queramos tomando m y n suficientemente grandes. La conversa se enuncia en el siguiente

Teorema 4.2.1 (Completitud de \mathcal{L}^2). *Para toda sucesión de Cauchy $f_n \in \mathcal{L}^2$ existe una función $f \in \mathcal{L}^2$ tal que $f_n \xrightarrow{\mathcal{L}^2} f$.*

Ejercicio 4.2.3. Muestre que la serie $\sum_{n=1}^{\infty} \sin(nx)/n$ converge en $\mathcal{L}^2(-\pi, \pi)$. (Ayuda: muestre que la sucesión de sumas parciales es de Cauchy.)

Otro resultado relacionado es el siguiente

Teorema 4.2.2 (\mathcal{C} denso en \mathcal{L}^2). *Para toda $f \in \mathcal{L}^2(a, b)$ y todo $\epsilon > 0$, existe una función $g \in \mathcal{C}[a, b]$ tal que $\|f - g\| < \epsilon$.*

Este resultado es, en cierto modo, similar a la densidad de \mathbb{Q} en \mathbb{R} , y tiene consecuencias equivalentes para la convergencia de sucesiones: toda función $f \in \mathcal{L}^2(a, b)$, continua o no, es el límite de alguna sucesión de funciones $f_n \in \mathcal{C}[a, b]$ (de hecho, de infinitas sucesiones diferentes); nótese la correspondencia con el Teorema 4.1.1, que afirma que toda distribución, regular o singular, es el límite de alguna sucesión de distribuciones regulares.

4.2.3. Funciones ortogonales

Consideremos una sucesión $\{\varphi_1, \varphi_2, \varphi_3, \dots\}$ de funciones en \mathcal{L}^2 , todas mutuamente ortogonales, es decir $\langle \varphi_i, \varphi_j \rangle = 0$ si $i \neq j$, y una función arbitraria $f \in \mathcal{L}^2$. Supongamos que queremos hallar la “mejor aproximación” a f en el sentido que

$$\left\| f - \sum_{k=1}^n \alpha_k \varphi_k \right\|$$

sea mínima. Es bastante directo mostrar que este mínimo se alcanza para

$$\alpha_k = \frac{\langle \varphi_k, f \rangle}{\langle \varphi_k, \varphi_k \rangle},$$

y que su valor viene dado por

$$\left\| f - \sum_{k=1}^n \frac{\langle \varphi_k, f \rangle}{\langle \varphi_k, \varphi_k \rangle} \varphi_k \right\|^2 = \|f\|^2 - \sum_{k=1}^n \frac{|\langle \varphi_k, f \rangle|^2}{\langle \varphi_k, \varphi_k \rangle} \geq 0,$$

lo que da la relación

$$\sum_{k=1}^n \frac{|\langle \varphi_k, f \rangle|^2}{\langle \varphi_k, \varphi_k \rangle} \leq \|f\|^2.$$

Como esta relación vale para todo n , debe valer para $n \rightarrow \infty$, lo que nos da la *desigualdad de Bessel*

$$\sum_{k=1}^{\infty} \frac{|\langle \varphi_k, f \rangle|^2}{\langle \varphi_k, \varphi_k \rangle} \leq \|f\|^2. \quad (4.37)$$

Es inmediato que la desigualdad de Bessel se vuelve una igualdad si, y sólo si,

$$f \stackrel{\mathcal{L}^2}{=} \sum_{k=1}^{\infty} \frac{\langle \varphi_k, f \rangle}{\langle \varphi_k, \varphi_k \rangle} \varphi_k.$$

Cuando esto ocurre, decimos que el conjunto ortogonal $\{\varphi_k\}_{k=1}^{\infty}$ es *completo en \mathcal{L}^2* , y forma una *base* para \mathcal{L}^2 ; como este espacio tiene dimensión infinita, toda base debe constar de un número infinito de elementos.

Cuando la desigualdad de Bessel se vuelve una igualdad, la relación resultante

$$\|f\|^2 = \sum_{k=1}^{\infty} \frac{|\langle \varphi_k, f \rangle|^2}{\langle \varphi_k, \varphi_k \rangle} \quad (4.38)$$

se conoce como *relación de Parseval* o *relación de completitud*, resultado también expresado como el

Teorema 4.2.3 (Parseval). *Un conjunto ortogonal $\{\varphi_k\}_{k=1}^{\infty}$ es completo en \mathcal{L}^2 si, y sólo si, satisface la relación de Parseval (4.38) para toda $f \in \mathcal{L}^2$.*

Podemos interpretar este resultado como que, en el sentido de la norma, no hay nada en f que no esté representado por sus “proyecciones” $\langle f, \varphi_k \rangle / \langle \varphi_k, \varphi_k \rangle$ sobre cada elemento de la base. A este respecto, notemos que en principio tendríamos $\dim\{f : [a, b] \rightarrow \mathbb{C}\} = \aleph_1$; la exigencia de que $\|f\| < \infty$ junto con la igualdad en \mathcal{L}^2 definida en el sentido de la norma, es lo que nos permite representar *todos y cada uno* de los elementos de $\mathcal{L}^2(a, b)$ como combinaciones lineales de elementos de un conjunto cuya cardinalidad es \aleph_0 .

Capítulo 5

Problemas de Contorno

5.1. EDO lineales de segundo orden

Caveat lector: Esta sección tiene por único objetivo refrescar conocimientos que el lector necesitará en las secciones subsiguientes, en particular la solución por series de potencias de EDOs con coeficientes analíticos, y por series de potencias generalizadas de EDOs con puntos singulares regulares; ello incluye los ejemplos clásicos de las ecuaciones de Legendre y de Bessel, respectivamente, y el enunciado formal del método de Frobenius. El lector familiarizado con estos temas puede pasar directamente a la Sección 5.2.

5.1.1. Coeficientes analíticos

Recordemos que una función $f(x)$ se llama *analítica* en un dominio $|x - x_0| < \rho$ si existe un desarrollo en serie de potencias

$$f(x) = \sum_{n=0}^{\infty} a_n (x - x_0)^n \quad (5.1)$$

que converge absolutamente en $|x - x_0| < \rho$.

Consideremos una EDO de la forma

$$y'' + P_1(x)y' + P_2(x)y = 0, \quad x \in I := (x_0 - \rho, x_0 + \rho) \subset \mathbb{R}, \quad (5.2)$$

con P_1 y P_2 analíticos en I , es decir, I no contiene ningún punto singular de la EDO. La analiticidad de los coeficientes implica la existencia de desarrollos en serie

$$P_1(x) = \sum_{n=0}^{\infty} b_n (x - x_0)^n, \quad P_2(x) = \sum_{n=0}^{\infty} c_n (x - x_0)^n \quad (5.3)$$

absolutamente convergentes en I , y cuyos coeficientes b_n y c_n daremos por conocidos. Es natural entonces buscar una solución de (5.2) que también sea analítica en I .

Propodremos entonces para $y(x)$ una serie de potencias de $(x - x_0)$,

$$y = \sum_{n=0}^{\infty} a_n (x - x_0)^n, \quad (5.4)$$

cuyos coeficientes a_n deberemos determinar. Asumiendo momentáneamente la convergencia absoluta de esta serie en I , que en seguida demostraremos, podemos derivarla término a término para obtener

$$y' = \sum_{n=1}^{\infty} a_n n (x - x_0)^{n-1} = \sum_{n=0}^{\infty} (n+1) a_{n+1} (x - x_0)^n, \quad (5.5a)$$

$$y'' = \sum_{n=2}^{\infty} a_n n(n-1) (x - x_0)^{n-2} = \sum_{n=0}^{\infty} (n+1)(n+2) a_{n+2} (x - x_0)^n. \quad (5.5b)$$

(5.5c)

Multiplicando y' por P_1 e y por P_2 tenemos

$$P_1(x)y' = \sum_{n=0}^{\infty} \left(\sum_{m=0}^n (m+1) a_{m+1} b_{n-m} \right) (x - x_0)^n, \quad (5.6a)$$

$$P_2(x)y = \sum_{n=0}^{\infty} \left(\sum_{m=0}^n a_m c_{n-m} \right) (x - x_0)^n, \quad (5.6b)$$

donde hemos reescrito las sumas dobles usando

$$\sum_{p=0}^{\infty} \alpha_p x^p \sum_{q=0}^{\infty} \beta_q x^q = \sum_{p=0}^{\infty} \sum_{q=0}^{\infty} \alpha_p \beta_q x^{p+q} = \sum_{n=0}^{\infty} \sum_{m=0}^n \alpha_n \beta_{n-m} x^n. \quad (5.7)$$

Substituyendo (5.21) y (5.4) en (5.2) tendremos

$$\sum_{n=0}^{\infty} \left[(n+2)(n+1)a_{n+2} + \sum_{m=0}^n ((m+1)a_{m+1}b_{n-m} + a_m c_{n-m}) \right] (x - x_0)^n = 0. \quad (5.8)$$

Notando que esta serie se obtuvo por derivación y producto de series absolutamente convergentes en I , vemos que también debe ser absolutamente convergente en I . Por lo tanto, para anularse idénticamente en I , el coeficiente de cada potencia $(x - x_0)^n$ debe ser nulo, es decir

$$(n+2)(n+1)a_{n+2} + \sum_{m=0}^n ((m+1)a_{m+1}b_{n-m} + a_m c_{n-m}) = 0, \quad n \geq 0. \quad (5.9)$$

Esta es una *relación de recurrencia* para los coeficientes a_n , que permite calcular a_{n+2} si conocemos a_0, a_1, \dots, a_{n+1} y b_m y c_m . Eligiendo arbitrariamente a_0 y a_1 y calculando a_2, a_3, \dots mediante (5.9), obtenemos una serie de la forma (5.4) que satisface la EDO (5.2).

Vamos a mostrar ahora que (5.4) converge absolutamente en I . Sea $x_1 \neq x_0$, $x_1 \in I$, y sea $t := |x_1 - x_0|$. Dado que P_1 y P_2 convergen absolutamente en I , sus términos deben ser acotados; sean entonces $M_1, M_2 \in \mathbb{R}$ tales que

$$|b_m(x - x_o)^m| = |b_m|t^m \leq M_1, \quad |c_m(x - x_o)^m| = |c_m|t^m \leq M_2, \quad (5.10)$$

con $M_1 > 0$, $M_2 > 0$, y sea $M := \max(M_1, tM_2)$, de modo que

$$|b_m| \leq \frac{M}{t^m}, \quad |c_m| \leq \frac{M}{t^{m+1}}. \quad (5.10')$$

Substituyendo (5.10') en (5.9) y tomando valor absoluto vemos que

$$\begin{aligned} (n+2)(n+1)|a_{n+2}| &= \left| \sum_{m=0}^n ((m+1)a_{m+1}b_{n-m} + a_m c_{n-m}) \right| \\ &\leq \sum_{m=0}^n ((m+1)|a_{m+1}| |b_{n-m}| + |a_m| |c_{n-m}|) \\ &\leq \sum_{m=0}^n \left((m+1)|a_{m+1}| \frac{M}{t^{n-m}} + |a_m| \frac{M}{t^{n-m+1}} \right) \\ &= \frac{M}{t^{n+1}} \left(\sum_{m=0}^n (m+1)|a_{m+1}|t^{m+1} + \sum_{m=0}^n |a_m|t^m \right). \end{aligned} \quad (5.11)$$

Ahora bién,

$$\begin{aligned} \sum_{m=0}^n |a_m|t^m &= \sum_{m=-1}^{n-1} |a_{m+1}|t^{m+1} = \sum_{m=0}^n |a_{m+1}|t^{m+1} + |a_0| - |a_{n+1}|t^{n+1} \\ &\leq \sum_{m=0}^n |a_{m+1}|t^{m+1} + |a_0|. \end{aligned} \quad (5.12)$$

Luego

$$\begin{aligned} (n+2)(n+1)|a_{n+2}| &\leq \frac{M}{t^{n+1}} \left(\sum_{m=0}^n (m+2)|a_{m+1}|t^{m+1} + |a_0| \right) \\ &= \frac{M}{t^{n+1}} \sum_{m=-1}^n (m+2)|a_{m+1}|t^{m+1} \\ &= \frac{M}{t^{n+1}} \sum_{m=0}^{n+1} (m+1)|a_m|t^m, \quad n \geq 0. \end{aligned} \quad (5.13)$$

Sean ahora $A_0 := |a_0|$, $A_1 := |a_1|$, y

$$(n+2)(n+1)A_{n+2} = \frac{M}{t^{n+1}} \sum_{m=0}^{n+1} (m+1)|a_m|t^m, \quad (5.13')$$

de modo que $|a_n| \leq A_n \forall n \geq 0$; entonces

$$\sum_{n=0}^{\infty} a_n(x - x_0)^n \leq \sum_{n=0}^{\infty} |a_n| |x - x_0|^n \leq \sum_{n=0}^{\infty} A_n t^n. \quad (5.14)$$

Substituyendo n por $(n - 1)$ en $(5.13')$ y multiplicando por t^{-1} tendremos

$$(n + 1)nt^{-1}A_{n+1} = \frac{M}{t^{n+1}} \sum_{m=0}^n (m + 1)A_m t^m, \quad n \geq 1, \quad (5.13'')$$

y restando $(5.13'')$ de $(5.13')$

$$(n + 2)(n + 1)A_{n+2} - (n + 1)nt^{-1}A_{n+1} = M(n + 2)A_{n+1} \quad (5.15)$$

Luego

$$A_{n+2} = A_{n+1} \frac{M(n + 2) + t^{-1}(n + 1)n}{(n + 2)(n + 1)} \quad (5.16)$$

y

$$\frac{A_{n+2}|x - x_0|^{n+2}}{A_{n+1}|x - x_0|^{n+1}} = \frac{M(n + 2) + t^{-1}(n + 1)n}{(n + 2)(n + 1)} |x - x_0| \xrightarrow{n \rightarrow \infty} |x - x_0|t^{-1}. \quad (5.17)$$

Aplicando el criterio del cociente por paso al límite, vemos entonces que $\sum_{n=0}^{\infty} A_n |x - x_0|^n$ converge si $|x - x_0| < t = |x_1 - x_0|$, y por (5.14) , $\sum_{n=0}^{\infty} a_n(x - x_0)^n$ converge si $|x - x_0| < |x_1 - x_0|$. Como x_1 era un punto arbitrario de $I = (x_0 - \rho, x_0 + \rho)$, concluimos que

$$\sum_{n=0}^{\infty} a_n(x - x_0)^n \text{ converge absolutamente } \forall x \in I. \quad (5.18)$$

Finalmente, notando de (5.4) que $y(x_0) = a_0$, y de $(5.5a)$ que $y'(x_0) = a_1$, vemos que la *base canónica* del espacio de soluciones de la EDO (5.2) es la formada por ϕ_1 y ϕ_2 definidas por

$$\left. \begin{aligned} \phi_1(x) &= \sum_{n=0}^{\infty} a_n(x - x_0)^n, \quad a_0 = 1, a_1 = 0, \\ \phi_2(x) &= \sum_{n=0}^{\infty} a_n(x - x_0)^n, \quad a_0 = 0, a_1 = 1, \end{aligned} \right\} a_2, a_3, \dots \text{ según } (5.9). \quad (5.19)$$

5.1.2. La Ecuación de Legendre

En su forma más usual, la Ecuación de Legendre se escribe

$$(1 - x^2)y'' - 2xy' + \alpha(\alpha + 1)y = 0, \quad \alpha \in \mathbb{R}, \quad x \in (-1, 1). \quad (5.20)$$

Si $x \neq \pm 1$ podemos dividir por $(1 - x^2)$ y poner

$$y'' + P_1(x)y' + P_2(x)y = 0, \quad (5.20')$$

con

$$P_1(x) = -\frac{2x}{1-x^2} = -2 \sum_{n=0}^{\infty} x^{2n+1}, \quad (5.21a)$$

$$P_2(x) = \frac{\alpha(\alpha+1)}{1-x^2} = \alpha(\alpha+1) \sum_{n=0}^{\infty} x^{2n}, \quad (5.21b)$$

analíticas en $|x| < 1$. Los puntos $x = \pm 1$ son *puntos singulares* de la ecuación.

Buscaremos una solución en serie de potencias de la forma

$$y = \sum_{n=0}^{\infty} a_n x^n, \quad x \in (-1, 1). \quad (5.22)$$

En lugar de utilizar la relación de recurrencia genérica (5.9), substituiremos (5.22) en (5.20). Derivando (5.22) tenemos

$$y' = \sum_{n=1}^{\infty} a_n n x^{n-1}, \quad y'' = \sum_{n=2}^{\infty} a_n n(n-1) x^{n-2}. \quad (5.23)$$

Luego

$$2xy' = \sum_{n=1}^{\infty} 2na_n x^n = \sum_{n=0}^{\infty} a_n x^n, \quad (5.24a)$$

$$\begin{aligned} (1-x^2)y'' &= \sum_{n=2}^{\infty} n(n-1)a_n x^{n-2} - \sum_{n=2}^{\infty} n(n-1)a_n x^n \\ &= \sum_{n=0}^{\infty} (n+1)(n+2)a_{n+2} x^n - \sum_{n=0}^{\infty} n(n-1)a_n x^n \\ &= \sum_{n=0}^{\infty} ((n+1)(n+2)a_{n+2} - n(n-1)a_n) x^n. \end{aligned} \quad (5.24b)$$

Substituyendo (5.22) y (5.24) en (5.20) resulta

$$\sum_{n=0}^{\infty} [(n+1)(n+2)a_{n+2} - n(n-1)a_n - 2na_n + \alpha(\alpha+1)a_n] x^n = 0, \quad (5.25)$$

e igualando a cero el coeficiente de x^n ,

$$(n+1)(n+2)a_{n+2} + ((\alpha-n)(\alpha+n+1))a_n = 0, \quad (5.26)$$

obtenemos la relación de recurrencia

$$a_{n+2} = -\frac{(\alpha-n)(\alpha+n+1)}{(n+1)(n+2)} a_n. \quad (5.27)$$

Notemos que (5.27) sólo relaciona a_{n+2} con a_n , sin involucrar a_{n+1} . De este modo, tenemos dos relaciones de recurrencia separadas, una para los coeficientes de índice par y otra para los de índice impar. Dado a_0 obtendremos

$$\begin{aligned} a_2 &= -\frac{\alpha \cdot (\alpha + 1)}{1 \cdot 2} a_0, \\ a_4 &= (-1)^2 \frac{\alpha(\alpha - 2) \cdot (\alpha + 1)(\alpha + 3)}{1 \cdot 2 \cdot 3 \cdot 4} a_0, \\ &\vdots \\ a_{2n} &= (-1)^n \frac{\alpha(\alpha - 2) \dots (\alpha - 2n + 2) \cdot (\alpha + 1)(\alpha + 3) \dots (\alpha + 2n - 1)}{(2n)!} a_0, \end{aligned} \tag{5.28a}$$

y dado a_1 tendremos

$$\begin{aligned} a_3 &= -\frac{(\alpha - 1) \cdot (\alpha + 2)}{2 \cdot 3} a_1, \\ a_5 &= (-1)^2 \frac{(\alpha - 1)(\alpha - 3) \cdot (\alpha + 2)(\alpha + 4)}{2 \cdot 3 \cdot 4 \cdot 5} a_1, \\ &\vdots \\ a_{2n+1} &= (-1)^n \frac{(\alpha - 1)(\alpha - 3) \dots (\alpha - 2n + 1) \cdot (\alpha + 2)(\alpha + 4) \dots (\alpha + 2n)}{(2n + 1)!} a_1. \end{aligned} \tag{5.28b}$$

La solución general de (5.20) queda entonces

$$y(x) = a_0 \phi_1^\alpha(x) + a_1 \phi_2^\alpha(x), \tag{5.29}$$

con

$$\phi_1^\alpha(x) := 1 + \sum_{n=1}^{\infty} (-1)^n \frac{\alpha(\alpha - 2) \dots (\alpha - 2n + 2) \cdot (\alpha + 1)(\alpha + 3) \dots (\alpha + 2n - 1)}{(2n)!} x^{2n}, \tag{5.30a}$$

$$\phi_2^\alpha(x) := x + \sum_{n=1}^{\infty} (-1)^n \frac{(\alpha - 1)(\alpha - 3) \dots (\alpha - 2n + 1) \cdot (\alpha + 2)(\alpha + 4) \dots (\alpha + 2n)}{(2n + 1)!} x^{2n+1}, \tag{5.30b}$$

que forman la *base canónica* del espacio de soluciones de (5.20), y convergen absolutamente para $|x| < 1$. Notemos que ϕ_1^α es una función par de x , y ϕ_2^α impar.

Puede verse fácilmente que, cuando α es un entero ≥ 0 , una de las dos series se reduce a un polinomio. Por ejemplo, supongamos que $\alpha = 2m$, $m = 0, 1, 2, \dots$. De la relación de recurrencia (5.27) es inmediato que $a_{2m+2} = 0$, y que todos

los coeficientes de índice par que le siguen, también se anularán. Dado que en este caso

$$\begin{aligned}\alpha(\alpha - 2) \dots (\alpha - 2n + 2) &= 2m(2m - 2) \dots (2m - 2n + 2) \\ &= \frac{2^nm!}{(m-n)!}\end{aligned}\quad (5.31a)$$

y

$$\begin{aligned}(\alpha + 1)(\alpha + 3) \dots (\alpha + 2n - 1) &= (2m + 1)(2m + 3) \dots (2m + 2n - 1) \\ &= \frac{(2m + 2n)!}{(2m)!} \frac{m!}{2^n(m+n)!},\end{aligned}\quad (5.31b)$$

resulta

$$\phi_1^{2m} = 1 + \sum_{n=1}^m (-1)^n \frac{(m!)^2}{2m!} \frac{(2m+2n)!}{(m-n)!(m+n)!(2n)!} x^{2n}. \quad (5.32)$$

Supongamos en cambio que $\alpha = 2m + 1$, $m = 0, 1, 2, \dots$. De la relación de recurrencia (5.27) es inmediato que $a_{2m+3} = 0$, y que todos los coeficientes de índice impar que le siguen, también se anularán. Las expresiones para $(\alpha - 1)(\alpha - 3) \dots (\alpha - 2n + 1)$ y $(\alpha + 2)(\alpha + 4) \dots (\alpha + 2n)$ dan ahora resultados totalmente análogos a (5.31), y resulta

$$\phi_2^{2m+1} = x + \sum_{n=1}^m (-1)^n \frac{(m!)^2}{(2m+1)!} \frac{(2m+2n+1)!}{(m-n)!(m+n)!(2n+1)!} x^{2n+1}. \quad (5.33)$$

Debe notarse que, en ambos casos, la *otra* solución sigue siendo una serie infinita.

Por convención, se define el *polinomio de Legendre de orden n* como proporcional a la solución cuya serie se trunca para $\alpha = n$, de acuerdo a

$$P_n(x) := \frac{1}{2^n} \sum_{r=0}^{[n/2]} \frac{(-1)^r (2n-2r)!}{r!(n-r)!(n-2r)!} x^{n-2r}, \quad (5.34)$$

donde $[n/2]$ es la parte entera de $n/2$ ($n/2$ si n es par, $(n-1)/2$ si n es impar). Generalmente, la *otra* solución de la base, cuya serie no se trunca, se denota por $Q_n(x)$, con un factor de normalización también convencional. Estas funciones serán estudiadas en detalle más adelante, en el Capítulo 6, dedicado a Funciones Especiales. Aquí nos limitaremos a notar que $P_n(x)$, justamente por ser un polinomio, es analítica en toda la recta real (y en realidad en todo el plano complejo), incluso en los puntos singulares $x = \pm 1$ de (5.20).

5.1.3. Puntos singulares regulares

Diremos que una EDO de la forma

$$y'' + P_1(x)y' + P_2(x)y = 0, \quad x \in I := (x_0 - \rho, x_0 + \rho) \subset \mathbb{R}, \quad (5.35)$$

tiene un *punto singular regular* en $x_0 \in I$ si $P_1(x)$ y $P_2(x)$ no son analíticas en x_0 , pero $(x - x_0)P_1(x)$ y $(x - x_0)^2P_2(x)$ son analíticas en I . En ese caso, frecuentemente reescribiremos (5.35) como¹

$$(x - x_0)^2y'' + (x - x_0)\tilde{P}_1(x)y' + \tilde{P}_2(x)y = 0, \quad (5.35')$$

donde

$$\tilde{P}_1(x) = (x - x_0)P_1(x) = \sum_{n=0}^{\infty} b_n(x - x_0)^n, \quad (5.36a)$$

$$\tilde{P}_2(x) = (x - x_0)^2P_2(x) = \sum_{n=0}^{\infty} c_n(x - x_0)^n, \quad (5.36b)$$

y ambas series son absolutamente convergentes en I por hipótesis.

Por analogía con el caso de una EDO con coeficientes analíticos, nos gustaría proponer una solución de (5.35') en serie de potencias de $(x - x_0)$, pero ello implicaría asumir una solución analítica en x_0 , lo que no está garantizado ya que es un punto singular. La salida más simple es multiplicar esta serie por un factor de la forma $|x - x_0|^r$ con r (por ahora) indeterminado, que permita dar cuenta de un posible comportamiento singular de la solución en x_0 . Proponemos entonces para $y(x)$ la *serie de potencias generalizada*²

$$y = |x - x_0|^r \sum_{n=0}^{\infty} a_n(x - x_0)^n, \quad r \in \mathbb{C}, \quad a_0 \neq 0. \quad (5.37)$$

Suponiendo convergencia de esta serie en I , podemos derivar término a término y obtener³

$$y' = |x - x_0|^{r-1} \operatorname{sgn}(x - x_0) \sum_{n=0}^{\infty} (r+n)a_n(x - x_0)^n, \quad (5.38a)$$

$$y'' = |x - x_0|^{r-2} \sum_{n=0}^{\infty} (r+n)(r+n-1)a_n(x - x_0)^n, \quad (5.38b)$$

¹La multiplicación por $(x - x_0)^2$ podría introducir soluciones espúreas, que veremos en seguida cómo evitar.

²La elección de un prefactor de la forma $|x - x_0|^r$ se debe a que r puede no ser entero y, en principio, estamos buscando soluciones reales; si admitimos soluciones complejas (y generalizamos el intervalo $I \subset \mathbb{R}$ a un disco en el plano complejo), podemos cambiarlo en todo lo que sigue por $(x - x_0)^r$, y tratar con las posibles singularidades esenciales en x_0 y los correspondientes cortes ramales. Por otro lado, asumir $a_0 \neq 0$ no implica pérdida de generalidad, ya que si $a_0 = 0$ simplemente estaríamos “desplazando” la serie, reemplazando r por $r+1$ y a_n por a_{n+1} .

³El lector atento notará que $|x - x_0|$ no es dos veces diferenciable en x_0 en el sentido usual ya que, como vimos en el Capítulo 4 en el contexto de distribuciones, $\partial_x|x - x_0| = 2\Theta(x - x_0) - 1 = \operatorname{sgn}(x - x_0)$ es discontinua allí. Sin embargo $\partial_x^2|x - x_0| = 2\delta(x - x_0)$. En principio, entonces, y'' debería incluir una contribución aditiva proporcional a la delta de Dirac $\delta(x - x_0)$. Pero como vimos en la Sección 4.1.4, $(x - x_0)\delta(x - x_0) \equiv 0$, de modo que esta contribución “desaparece” de (5.35'). Contribuciones de este tipo corresponden a soluciones espúreas introducidas al multiplicar (5.35) por $(x - x_0)^2$ para obtener (5.35'), que justamente no debemos incluir.

ya que $|x - x_0| \operatorname{sgn}(x - x_0) = (x - x_0)$. Multiplicando y'' por $(x - x_0)^2$, y' por $(x - x_0)\tilde{P}_1(x)$ e y por $\tilde{P}_2(x)$ tenemos

$$(x - x_0)^2 y'' = |x - x_0|^r \sum_{n=0}^{\infty} (r+n)(r+n-1)a_n(x - x_0)^n, \quad (5.39a)$$

$$(x - x_0)\tilde{P}_1(x)y' = |x - x_0|^r \sum_{n=0}^{\infty} \left(\sum_{m=0}^n (r+m)a_m b_{n-m} \right) (x - x_0)^n, \quad (5.39b)$$

$$\tilde{P}_2(x)y = |x - x_0|^r \sum_{n=0}^{\infty} \left(\sum_{m=0}^n a_m c_{n-m} \right) (x - x_0)^n, \quad (5.39c)$$

donde hemos reescrito las sumas dobles usando (5.7).

Substituyendo (5.39) en (5.35') e igualando a cero el coeficiente de $(x - x_0)^n$ obtenemos

$$(r+n)(r+n-1)a_n + \sum_{m=0}^n ((r+m)b_{n-m} + c_{n-m}) a_m = 0, \quad n = 0, 1, 2, \dots, \quad (5.40)$$

y agrupando todos los términos que contienen a_n para cada n ,⁴

$$[(r+n)(r+n-1) + (r+n)b_0 + c_0]a_n + \sum_{m=0}^{n-1} [(r+m)b_{n-m} + c_{n-m}]a_m = 0. \quad (5.41)$$

Recordando que asumimos $a_0 \neq 0$, vemos de (5.41) para $n = 0$ que debemos tener

$$q(r) := r(r-1) + rb_0 + c_0 = r(r-1) + r\tilde{P}_1(x_0) + \tilde{P}_2(x_0) = 0, \quad (5.42)$$

donde $q(r)$ es el llamado *polinomio indicial* de (5.35'). Luego, para que exista una solución de la forma (5.37), r debe ser una raíz de $q(r)$.

Por otro lado, para $n > 0$ (5.41) puede escribirse

$$q(r+n)a_n + \sum_{m=0}^{n-1} [(r+m)b_{n-m} + c_{n-m}]a_m = 0, \quad n = 1, 2, \dots \quad (5.43)$$

La suma del segundo término es una combinación lineal de a_0, \dots, a_{n-1} con coeficientes que son funciones de r (y de los coeficientes conocidos b_m y c_m), de modo que podemos escribir a_n como una función de r ,

$$a_n(r) = -\frac{\sum_{m=0}^{n-1} [(r+m)b_{n-m} + c_{n-m}]a_m}{q(r+n)}, \quad n = 1, 2, \dots \quad (5.44)$$

Esta expresión se conoce como la *relación de recurrencia* para los coeficientes a_n , ya que permite determinar sucesivamente cada uno en función de los anteriores, siempre que $q(r+n) \neq 0$.

⁴La suma del segundo término debe considerarse nula para $n = 0$.

Supongamos ahora que r_1 y r_2 son las raíces de $q(r)$, y ordenémoslas de modo que $\operatorname{Re} r_1 \geq \operatorname{Re} r_2$; entonces $\operatorname{Re}(r_1 + n) > \operatorname{Re} r_2$, $n = 1, 2, \dots$, y $q(r_1 + n) \neq 0 \forall n > 0$, de modo que $a_k(r_1)$ está bien definido $\forall k \geq 0$. Eligiendo $a_0(r_1) \neq 0$ [por ejemplo $a_0(r_1) = 1$], vemos que

$$\phi_1(x) := |x - x_0|^{r_1} \sum_{n=0}^{\infty} a_n(r_1)(x - x_0)^n \quad (5.45)$$

es una solución de $(5.35')$, siempre que converja, en I .

Si $r_2 = r_1$, $a_k(r_2) \equiv a_k(r_1)$ y no podemos usarlos para obtener una segunda solución. Por otro lado, si $r_2 \neq r_1$ tendremos $\operatorname{Re} r_2 < \operatorname{Re} r_1$, pero ello implicará $q(r_2 + n) \neq 0$, $n = 1, 2, \dots$, si y sólo si $r_2 + n \neq r_1 \forall n = 1, 2, \dots$; es decir, $a_k(r_2)$ estará bien definido $\forall k > 0$ si y sólo si $r_2 - r_1$ no es un entero positivo, y en este caso $a_k(r_2) \not\equiv a_k(r_1)$ por (5.44) . En definitiva, si (y sólo si) $r_1 - r_2 \neq 0, 1, 2, \dots$,

$$\phi_2(x) := |x - x_0|^{r_2} \sum_{n=0}^{\infty} a_n(r_2)(x - x_0)^n \quad (5.45')$$

es otra solución de $(5.35')$, siempre que converja, en I .

La convergencia de (5.45) puede demostrarse en forma análoga a como se hizo en la Sección precedente para las soluciones de la EDO lineal con coeficientes analíticos. Asumiremos que $r_1 - r_2 \neq 0, 1, 2, \dots$. Dado que $q(r) = (r - r_1)(r - r_2)$, entonces $\forall k \in \mathbb{N}$

$$q(r_1 + k) = k(k + r_1 - r_2) \quad \text{y} \quad q(r_2 + k) = k(k + r_2 - r_1), \quad (5.46)$$

luego

$$|q(r_1 + k)| \geq k(k - |r_1 - r_2|) \quad \text{y} \quad |q(r_2 + k)| \geq k(k - |r_2 - r_1|). \quad (5.47)$$

Sea σ un real cualquiera tal que $0 < \sigma < \rho$; la convergencia absoluta de (5.36) implica que sus términos son acotados, es decir $\exists M > 0$ tal que

$$|b_j|\sigma^j \leq M, \quad |c_j|\sigma^j \leq M, \quad j = 0, 1, 2, \dots \quad (5.48)$$

Substituyendo (5.47) y (5.48) en (5.43) ,

$$k(k - |r_1 - r_2|) |a_k(r_1)| \leq M \sum_{j=0}^{k-1} (j + 1 + |r_1|) \sigma^{j-k} |a_j(r_1)|, \quad k = 1, 2, \dots \quad (5.49)$$

Sea N el entero que satisface $N - 1 \leq |r_1 - r_2| < N$, y sean $\alpha_0, \alpha_1, \dots$ definidos por

$$\alpha_0 = a_0(r_1) = 1, \quad (5.50a)$$

$$\alpha_k = |a_k(r_1)|, \quad k = 1, 2, \dots, N - 1, \quad (5.50b)$$

$$k(k - |r_1 - r_2|) \alpha_k = M \sum_{j=0}^{k-1} (j + 1 + |r_1|) \sigma^{j-k} \alpha_j, \quad k = N, N + 1, \dots \quad (5.50c)$$

Luego, comparando (5.50c) con (5.49) vemos que

$$|a_k(r_1)| \leq \alpha_k, \quad k = 1, 2, \dots \quad (5.51)$$

Ahora, reemplazando k por $k + 1$ en (5.50c) resulta

$$\begin{aligned} (k+1)(k+1-|r_1-r_2|)\alpha_{k+1} &= M \sum_{j=0}^k (j+1+|r_1|)\sigma^{j-k-1}\alpha_j \\ &= \frac{M}{\sigma} \sum_{j=0}^k (j+1+|r_1|)\sigma^{j-k}\alpha_j \\ &= \frac{1}{\sigma} [k(k-|r_1-r_2|) + M(k+1+|r_1|)]\alpha_k \end{aligned} \quad (5.52)$$

para $k \geq N$, luego

$$\left| \frac{\alpha_{k+1}(x-x_0)^{k+1}}{\alpha_k(x-x_0)^k} \right| = \frac{k(k-|r_1-r_2|) + M(k+1+|r_1|)}{\sigma(k+1)(k+1-|r_1-r_2|)} |x-x_0| \xrightarrow[k \rightarrow \infty]{} \frac{|x-x_0|}{\sigma}. \quad (5.53)$$

Aplicando el criterio del cociente por paso al límite, vemos que $\sum_{k=0}^{\infty} \alpha_k |x-x_0|^k$ converge para $|x-x_0| < \sigma$. Como $\sigma < \rho$ era arbitrario, entonces

$$\sum_{k=0}^{\infty} a_k(r_1)(x-x_0)^k \leq \sum_{k=0}^{\infty} |a_k(r_1)| |x-x_0|^k \leq \sum_{k=0}^{\infty} \alpha_k |x-x_0|^k < \infty \text{ si } |x-x_0| < \rho, \quad (5.54)$$

es decir, la serie (5.45) para ϕ_1 converge en I si $r_1 - r_2 \neq 0, 1, 2, \dots$. Cambiando r_1 por r_2 , un argumento completamente análogo nos lleva a concluir que la serie (5.45') para ϕ_2 también converge en I si $r_1 - r_2 \neq 0, 1, 2, \dots$

La independencia lineal de las dos soluciones ϕ_1 y ϕ_2 puede deducirse inmediatamente de (5.45) y (5.45') notando que, si $r_1 - r_2 \neq 0, 1, 2, \dots$,

$$\frac{\phi_2(x)}{\phi_1(x)} \xrightarrow{x \rightarrow x_0} |x-x_0|^{r_2-r_1} \neq \text{cte.} \quad (5.55)$$

y recordando que dos funciones cualesquiera son LD si y sólo si son proporcionales entre sí.

Necesitamos ahora encontrar una segunda solución de (5.35') LI de ϕ_1 para el caso en que $r_1 - r_2 = 0, 1, 2, \dots$. Para ello definamos

$$\Phi(x, r) := |x-x_0|^r \sum_{k=0}^{\infty} a_k(r) (x-x_0)^k, \quad |x-x_0| < \rho, \quad (5.56)$$

donde los coeficientes $a_k(r)$ vienen dados por la relación de recurrencia (5.44) para algún $a_0(r) \neq 0$ dado. La convergencia absoluta de esta serie en I puede demostrarse igual que acabamos de hacerlo para ϕ_1 . Definamos también el operador diferencial lineal

$$L := (x-x_0)^2 \partial_x^2 + (x-x_0) \tilde{P}_1(x) \partial_x + \tilde{P}_2(x). \quad (5.57)$$

Cuando apliquemos este operador a Φ , la relación (5.44) hará que los coeficientes de todas las potencias de $(x - x_0)$ se anulen, con excepción del de $(x - x_0)^0$, de modo que tendremos

$$L(\Phi) = a_0(r)q(r)|x - x_0|^r. \quad (5.58)$$

Para el caso $r_1 = r_2$, tenemos $q(r_1) = q'(r_1) = 0$. Derivando (5.58) respecto de r tendremos

$$\frac{\partial}{\partial r}L(\Phi) = L\left(\frac{\partial\Phi}{\partial r}\right) = [a_0(r)q'(r) + (a'_0(r) + a_0(r)\ln|x - x_0|)q(r)]|x - x_0|^r, \quad (5.59)$$

y poniendo $r = r_1$ vemos que $L\left(\frac{\partial\Phi}{\partial r}\right)|_{r=r_1} = 0$, así que $\frac{\partial\Phi}{\partial r}(x, r_1)$ es otra solución de (5.35') en I . Derivando (5.56) respecto de r y tomando $a_0(r) = 1$, ponemos entonces

$$\begin{aligned} \phi_2(x) := \frac{\partial\Phi}{\partial r}(x, r_1) &= |x - x_0|^{r_1} \sum_{k=0}^{\infty} (a'_k(r_1) + \ln|x - x_0|a_k(r_1))(x - x_0)^k \\ &= |x - x_0|^{r_1} \sum_{k=0}^{\infty} a'_k(r_1)(x - x_0)^k + \ln|x - x_0| \phi_1(x), \end{aligned} \quad (5.60)$$

con ϕ_1 dada por (5.45) con $a_0(r_1) = 1$. Esta solución converge donde lo hace Φ , es decir converge absolutamente en I , y al no ser proporcional a ϕ_1 es LI de ella. Notemos que si tomamos $a_0(r_1) = 1$, entonces $a'_0(r_1) = 0$.

Para el caso $r_1 = r_2 + m$ con $m \in \mathbb{N}$, $a_m(r_2)$ no está definida según (5.44) porque el denominador se anula. Sin embargo, como $q(r) = (r - r_1)(r - r_2)$, entonces $q(r+m) = (r - r_2)(r + m - r_2)$. El factor que se anula en el denominador de (5.44) cuando $n = m$ es justamente $(r - r_2)$, de modo que definiendo $a_0(r) = (r - r_2)$ todos los $a_k(r)$ con $k = 1, 2, \dots, m - 1$ contienen este factor en el numerador, y $a_m(r)$ ya no diverge en $r = r_2$, ya que este factor cancela al $(r - r_2)$ del denominador de (5.44). Para $k > m$ este factor ya no aparece en $a_k(r)$, así que estos coeficientes estarán bien definidos. Podremos definir entonces

$$\Psi(x, r) := |x - x_0|^r \sum_{k=0}^{\infty} a_k(r)(x - x_0)^k, \quad a_0(r) = r - r_2, \quad (5.61)$$

que converge absolutamente en I , y es solución de (5.35') para $r = r_2$. Pero $a_0(r) = (r - r_2)$ impone $a_0(r_2) = a_1(r_2) = \dots = a_{m-1}(r_2) = 0$, así que la serie de $\Psi(x, r_2)$ recién comienza con el término $(x - x_0)^m$, y tendremos

$$\Psi(x, r_2) = \text{sgn}(x - x_0)^m |x - x_0|^{r_2+m} \sum_{k=0}^{\infty} a_{k+m}(r)(x - x_0)^k \propto \phi_1(x), \quad (5.62)$$

ya que $r_2 + m = r_1$. Para obtener una segunda solución LI de ϕ_1 , notamos que Ψ también satisface (5.58); derivando respecto de r con $a_0(r) = r - r_2$ tendremos

$$\frac{\partial}{\partial r}L(\Psi) = L\left(\frac{\partial\Psi}{\partial r}\right) = [q(r) + (r - r_2)(q'(r) + \ln|x - x_0|q(r))]|x - x_0|^r, \quad (5.63)$$

y haciendo $r = r_2$ vemos que $L\left(\frac{\partial\Psi}{\partial r}\right)|_{r=r_2} = 0$, así que $\frac{\partial\Psi}{\partial r}(x, r_1)$ es otra solución de (5.35') en I . Derivando (5.61) respecto de r tomando $a_0(r) = r - r_2$, y haciendo luego $r = r_2$, ponemos entonces

$$\phi_2(x) := \frac{\partial\Psi}{\partial r}(x, r_2) = |x - x_0|^{r_2} \sum_{k=0}^{\infty} (a'_k(r_2) + \ln|x - x_0| a_k(r_2)) (x - x_0)^k, \quad (5.64)$$

y recordando $a_0(r_2) = a_1(r_2) = \dots = a_{m-1}(r_2) = 0$, $a_k(r_2) \propto a_m(r_2)$ para $k > m$ y $r_1 = r_2 + m$, tendremos

$$\begin{aligned} \phi_2(x) &= |x - x_0|^{r_2} \left[\sum_{k=0}^{\infty} a'_k(r_2) (x - x_0)^k + \ln|x - x_0| \sum_{k=0}^{\infty} a_{k+m}(r_2) (x - x_0)^{k+m} \right] \\ &= |x - x_0|^{r_2} \sum_{k=0}^{\infty} a'_k(r_2) (x - x_0)^k + a_m(r_2) \ln|x - x_0| \phi_1(x), \end{aligned} \quad (5.65)$$

con ϕ_1 dada por (5.45) con $a_0(r_1) = 1$. Notemos que $a_m(r_2)$ *puede* ser nula. Esta solución converge donde lo hace Ψ , es decir converge absolutamente en I , y al no ser proporcional a ϕ_1 es LI de ella.

Los resultados obtenidos en esta sección pueden resumirse en el siguiente

Teorema 5.1.1 (Frobenius). *Dada la EDO*

$$(x - x_0)^2 y'' + (x - x_0) \tilde{P}_1(x) y' + \tilde{P}_2(x) y = 0, \quad (5.66)$$

con \tilde{P}_1 y \tilde{P}_2 analíticas en $I = (x_0 - \rho, x_0 + \rho)$, y dadas las raíces r_1 y r_2 de su polinomio indicial

$$q(r) = r(r - 1) + \tilde{P}_1(x_0) r + \tilde{P}_2(x_0) \quad (5.67)$$

con $r_1 \geq r_2$, entonces una solución de la EDO se escribe

$$\phi_1(x) = |x - x_0|^{r_1} \sum_{n=0}^{\infty} a_n (x - x_0)^n, \quad a_0 = 1. \quad (5.68a)$$

Si $r_1 - r_2 \notin \mathbb{N} \cup \{0\}$, otra solución LI se escribe

$$\phi_2(x) = |x - x_0|^{r_2} \sum_{n=0}^{\infty} b_n (x - x_0)^n, \quad b_0 = 1. \quad (5.68b)$$

Si en cambio $r_1 - r_2 = m \in \mathbb{N} \cup \{0\}$, otra solución LI se escribe

$$\phi_2(x) = |x - x_0|^{r_2} \sum_{n=0}^{\infty} b_n (x - x_0)^n + c \ln|x - x_0| \phi_1(x), \quad (5.68c)$$

donde si $m = 0$, $b_0 = 0$, $b_1 = 1$ y $c \neq 0$, mientras que si $m \neq 0$, $b_0 = 1$ y c puede ser nula. Todas las series convergen absolutamente en I , y son soluciones de la EDO en I .

Notemos que las raíces del polinomio indicial *pueden* ser complejas; pero en ese caso, como (5.67) tiene coeficientes *reales*, tendremos $r_2 = r_1^*$ y *no pueden ser iguales ni diferir en un entero*, ya que $r_1 - r_2 = 2i \operatorname{Im} r_1$. El método aún se aplica, y las dos soluciones LI tienen la forma (5.68a) y (5.68b).

5.1.4. La ecuación de Bessel

La Ecuación de Bessel se escribe generalmente en la forma

$$x^2y'' + xy' + (x^2 - \alpha^2)y = 0, \quad \operatorname{Re}(\alpha) \geq 0, \quad x \in (-\infty, \infty). \quad (5.69)$$

Es inmediato que (5.69) es de la forma (5.66), con $x_0 = 0$, $\tilde{P}_1(x) = 1$, $\tilde{P}_2(x) = x^2 - \alpha^2$, de modo que $x = 0$ es un punto singular regular. El polinomio indicial de (5.69) es

$$q(r) = r(r-1) + \tilde{P}_1(x_0)r + \tilde{P}_2(x_0) = r(r-1) + r - \alpha^2 = r^2 - \alpha^2, \quad (5.70)$$

de modo que sus raíces son

$$r_1 = \alpha, \quad r_2 = -\alpha. \quad (5.71)$$

Busquemos una solución de la forma

$$y = |x|^\alpha \sum_{n=0}^{\infty} a_n x^n, \quad a_0 \neq 0. \quad (5.72)$$

Notando que $b_0 = 1$, $b_n = 0$, $n \geq 1$, y que $c_0 = -\alpha^2$, $c_1 = 0$, $c_2 = 1$, $c_n = 0$, $n \geq 3$, y usando $q(r_1 + n) = (\alpha + n)^2 - \alpha^2$, la relación de recurrencia (5.43) implica

$$[(\alpha + 1)^2 - \alpha^2]a_1 = 0, \quad (5.73a)$$

$$[(\alpha + n)^2 - \alpha^2]a_n = -a_{n-2}, \quad n = 2, 3, \dots \quad (5.73b)$$

Como $\operatorname{Re}(\alpha) \geq 0$, (5.73a) impone $a_1 = 0$. Luego por (5.73b) tendremos $a_{2n+1} = 0 \forall n \geq 0$. Reescribiendo (5.73b) en la forma

$$a_n = -\frac{a_{n-2}}{n(n+2\alpha)} \quad (5.74)$$

vemos que

$$\begin{aligned} a_2 &= \frac{-a_0}{2^2(1+\alpha)}, \\ a_4 &= \frac{(-1)^2 a_0}{2^4 2!(1+\alpha)(2+\alpha)}, \\ a_6 &= \frac{(-1)^3 a_0}{2^6 3!(1+\alpha)(2+\alpha)(3+\alpha)}, \text{ etc.} \end{aligned}$$

Usando

$$(1 + \alpha)(2 + \alpha)(3 + \alpha) \dots (n + \alpha) = \frac{\Gamma(n + 1 + \alpha)}{\Gamma(1 + \alpha)},$$

tendremos una primera solución de (5.69) en la forma (5.68a),

$$\phi_1(x) = a_0|x|^\alpha \left(1 + \sum_{n=1}^{\infty} \frac{(-1)^n \Gamma(1 + \alpha) x^{2n}}{2^{2n} n! \Gamma(n + 1 + \alpha)} \right). \quad (5.75)$$

Tomando por convención $a_0 = 1/(2^\alpha \Gamma(1 + \alpha))$ obtenemos la *función de Bessel de primera especie y orden α* , definida $\forall x > 0$ como

$$J_\alpha(x) := \left(\frac{x}{2}\right)^\alpha \sum_{n=0}^{\infty} \frac{(-1)^n}{n! \Gamma(n + 1 + \alpha)} \left(\frac{x}{2}\right)^{2n}. \quad (5.76)$$

Si $2\alpha \neq 0, 1, 2, \dots$, podemos obtener una segunda solución de (5.69) LI de J_α en la forma (5.68b), simplemente cambiando α por $-\alpha$ en (5.73)–(5.76) para obtener

$$J_{-\alpha}(x) := \left(\frac{x}{2}\right)^{-\alpha} \sum_{n=0}^{\infty} \frac{(-1)^n}{n! \Gamma(n + 1 - \alpha)} \left(\frac{x}{2}\right)^{2n}. \quad (5.77)$$

Debemos notar que $J_{-\alpha}$ está bien definida (para $x > 0$) aún si $2\alpha = 0, 1, 2, \dots$. Sin embargo, si $\alpha = m \in \mathbb{N}$, como $\Gamma(s)$ diverge $\forall s$ entero ≤ 0 , los coeficientes de la serie de $J_{-\alpha}$ son cero para $n < m$ (y $\therefore J_{-\alpha}$ está bien definida $\forall x$), pero no será LI de J_α , sino proporcional a ella; lo mismo ocurre, obviamente, si $\alpha = 0$. En cambio, si $\alpha = \frac{1}{2}, \frac{3}{2}, \frac{5}{2}, \dots$ la aplicación del Teorema 5.1.1 nos lleva de nuevo a (5.77).

Si $\alpha = 0$, buscaremos una segunda solución de (5.69) LI de J_α en la forma (5.68c). En este caso tendremos

$$J_0(x) := \sum_{n=0}^{\infty} \frac{(-1)^n}{(n!)^2} \left(\frac{x}{2}\right)^{2n}, \quad (5.76')$$

ya que (5.76) es válida $\forall \alpha$. La forma más sencilla de hallar ϕ_2 es partir de (5.68c), escribiéndola en la forma

$$\phi_2(x) = \sum_{n=0}^{\infty} b_n x^n + \ln(x) J_0(x) \quad (5.78)$$

(donde hemos absorbido la constante c en la definición de los b_n), y substituir en (5.69), que para $\alpha = 0$ toma la forma

$$L(y) := x^2 y'' + x y' + x^2 y = 0. \quad (5.69')$$

Derivando (5.78) tenemos

$$\phi_2'(x) = \sum_{n=1}^{\infty} n b_n x^{n-1} + \frac{J_0(x)}{x} + \ln(x) J_0'(x), \quad (5.79a)$$

$$\phi_2''(x) = \sum_{n=2}^{\infty} n(n-1) b_n x^{n-2} - \frac{J_0(x)}{x^2} + \frac{2}{x} J_0'(x) + \ln(x) J_0''(x), \quad (5.79b)$$

de modo que substituyendo en (5.69') y usando $L(J_0) = 0$ y $b_0 = 0$ (ver Teorema 5.1.1), resulta

$$\begin{aligned}
 L(\phi_2) &= x^2\phi_2'' + x\phi_2' + x^2\phi_2 \\
 &= \left(\sum_{n=2}^{\infty} n(n-1)b_nx^n - J_0 + 2xJ_0' + x^2\ln(x)J_0'' \right) \\
 &\quad + \left(\sum_{n=1}^{\infty} nb_nx^n + J_0 + x\ln(x)J_0' \right) + \left(\sum_{n=0}^{\infty} b_nx^{n+2} + x^2\ln(x)J_0 \right) \\
 &= \sum_{n=2}^{\infty} (n(n-1)b_n + nb_n + b_{n-2})x^n + b_1x + 2xJ_0' + \ln(x)L(J_0) \\
 &= b_1x + 2^2b_2x^2 + \sum_{n=3}^{\infty} (n^2b_n + b_{n-2})x^n + 2 \sum_{m=1}^{\infty} \frac{(-1)^m 2mx^{2m}}{2^{2m}(m!)^2} = 0.
 \end{aligned} \tag{5.80}$$

Luego

$$b_1 = 0, \quad 2^2b_2 - 1 = 0, \quad 3^2b_3 + b_1 = 0, \quad \text{etc.} \tag{5.81}$$

De aquí es fácil verificar que $b_{2n+1} = 0$, $n = 1, 2, \dots$, y obtener para los coeficientes de índice par la recurrencia

$$(2m)^2b_{2m} + b_{2m-2} = \frac{(-1)^{m+1}m}{2^{2m-2}(m!)^2}, \quad m = 2, 3, \dots \tag{5.82}$$

Explícitamente, los primeros coeficientes son

$$b_2 = \frac{1}{2^2}, \quad b_4 = -\frac{1}{2^24^2} \left(1 + \frac{1}{2} \right), \quad b_6 = \frac{1}{2^24^26^2} \left(1 + \frac{1}{2} + \frac{1}{3} \right), \quad \text{etc.} \tag{5.83}$$

Así, la segunda solución de (5.69'), LI de J_0 , queda expresada

$$K_0(x) := - \sum_{m=1}^{\infty} \frac{(-1)^m}{((2m)!!)^2} \left(1 + \frac{1}{2} + \cdots + \frac{1}{m} \right) \left(\frac{x}{2} \right)^{2m} + \ln(x)J_0(x), \tag{5.77'}$$

que se conoce como *función de Bessel de segunda especie y orden 0*.

Para el caso $\alpha = m \in \mathbb{N}$, damos sin demostración el resultado de sustituir (5.68c) en (5.69),

$$\begin{aligned}
 K_m(x) &:= \ln(x)J_m(x) - \frac{1}{2} \left(\frac{x}{2} \right)^{-m} \sum_{n=0}^{m-1} \frac{(m-n-1)!}{n!} \left(\frac{x}{2} \right)^{2n} \\
 &\quad + \frac{1}{2} \left(\frac{x}{2} \right)^m \sum_{n=0}^{\infty} \frac{(-1)^n}{n!(n+m)!} (h_n + h_{n+m}) \left(\frac{x}{2} \right)^{2n},
 \end{aligned} \tag{5.77''}$$

donde hemos definido

$$h_n := \begin{cases} 0 & n = 0 \\ 1 + \frac{1}{2} + \cdots + \frac{1}{n} & n = 1, 2, \dots \end{cases} \tag{5.84}$$

Junto con J_m obtenida de (5.76) para $\alpha = m$, $(5.77'')$ es una segunda solución LI de (5.69) para $\alpha = m$.

Todas estas funciones serán estudiadas con mayor detalle en el Capítulo 6, dedicado a Funciones Especiales.

5.2. Problemas de valores iniciales

Consideremos una EDO de segundo orden

$$y'' + r(x)y' + s(x)y = f(x), \quad x \in I \subset \mathbb{R}, \quad (5.85)$$

donde r , s y f son funciones complejas sobre el intervalo real I , que supondremos que no contiene ningún punto singular de la ecuación. Una *solución* de (5.85) es una función $y \in \mathcal{C}^2(I)$ tal que (5.85) se satisface idénticamente. Recordemos (sin demostración) el

Teorema 5.2.1 (Existencia y unicidad de soluciones). *Si las funciones r , s y f son continuas en I , y x_0 es un punto en I , entonces dados dos números cualesquiera φ_0 y φ'_0 existe una única solución $\varphi(x)$ de (5.85) tal que*

$$\varphi(x_0) = \varphi_0, \quad \varphi'(x_0) = \varphi'_0. \quad (5.86)$$

Las condiciones (5.86) se denominan *condiciones iniciales*, que abreviaremos CI. El sistema formado por las ecuaciones (5.85) y (5.86) es llamado un *problema de valores iniciales*. Recordemos que si $\varphi_0 = 0$ y $\varphi'_0 = 0$ entonces la (única) solución es la *trivial*, $y \equiv 0$ en I . Recordemos también que la solución general de la ecuación homogénea (es decir con $f \equiv 0$ en I) es una combinación lineal de dos soluciones LI de la misma ecuación, y que la solución general de (5.85) se obtiene sumando a la solución general de la ecuación homogénea una solución particular cualquiera de la inhomogénea.

De lo anterior es evidente que determinar soluciones de la ecuación homogénea es fundamental. Sin embargo muchas veces no será posible encontrar soluciones de la EDO (5.85) homogénea en forma cerrada, y las soluciones en series de la Sección 5.1 pueden ser difíciles de construir, y aún más de interpretar. Afortunadamente existen muchos resultados generales que pueden demostrarse con sólo conocer el comportamiento de las funciones r y s , aún sin especificar las CI. Listaremos algunos de ellos en el resto de esta sección.

5.2.1. Ceros aislados

Una función $f : I \rightarrow \mathbb{C}$ se dice que tiene un *cero aislado* en x_0 si $f(x_0) = 0$ y existe una vecindad U de x_0 tal que $f(x) \neq 0$ para todo $x \in I \cap U \setminus \{x_0\}$. Tenemos

Teorema 5.2.2. *Si y es una solución no trivial de la Ecuación (5.85) homogénea, entonces todo cero de y en I es aislado.*

Demostración. Supongamos que x_0 es un cero de y . Si $y'(x_0) = 0$, entonces $y \equiv 0$ por el Teorema 5.2.1. Si $y'(x_0) \neq 0$, como y' es continua en I , existe una vecindad U de x_0 tal que $y' \neq 0$ en $U \cap I$. Por lo tanto y es estrictamente creciente o estrictamente decreciente en $U \cap I$. \square

5.2.2. Alternancia de ceros

Vamos a demostrar ahora un resultado de gran importancia en el estudio de soluciones de (5.85).

Teorema 5.2.3 (de separación de Sturm). *Si y_1 e y_2 son dos soluciones de (5.85) homogénea, y son LI en I , entonces los ceros de y_1 y los de y_2 son diferentes y se alternan, es decir, y_1 tiene exactamente un cero entre dos ceros sucesivos de y_2 y viceversa.*

Demostración. Como y_1 e y_2 son LI en I , su Wronskiano

$$W(y_1, y_2)(x) = y_1(x)y'_2(x) - y_2(x)y'_1(x)$$

no se anula en I , y por lo tanto tiene el mismo signo en todo I . Notemos primero que y_1 e y_2 no pueden tener ningún cero en común, porque W se anularía allí. Supongamos ahora que x_1 y x_2 son dos ceros sucesivos de y_2 ; entonces

$$\begin{aligned} W(x_1) &= y_1(x_1)y'_2(x_1) \neq 0, \\ W(x_2) &= y_1(x_2)y'_2(x_2) \neq 0, \end{aligned}$$

y $y_1(x_1)$, $y_1(x_2)$, $y'_1(x_1)$ e $y'_1(x_2)$ deben ser todos no nulos. Como y'_2 es continua en I , existe una vecindad U_1 de x_1 donde y'_2 no cambia de signo, y otra vecindad U_2 de x_2 donde y'_2 tampoco cambia de signo; pero el signo de y'_2 en ambas vecindades no puede ser el mismo, ya que si y_2 es creciente en una debe ser decreciente en la otra, dado que es continua y por hipótesis no tiene otros ceros entre x_1 y x_2 . Luego, para que $W(x_1)$ y $W(x_2)$ tengan el mismo signo, $y_1(x_1)$ e $y_1(x_2)$ deben tener también signos opuestos, y como y_1 es continua en I , debe tener al menos un cero entre x_1 y x_2 ; pero no puede tener más de uno, porque si x_3 y x_4 fuesen dos ceros de y_1 , ambos entre x_1 y x_2 , el mismo argumento implicaría que y_2 debe tener al menos un cero entre x_3 y x_4 , contradiciendo la hipótesis de que x_1 y x_2 son ceros consecutivos de y_2 . \square

De aquí se desprende otro resultado fundamental:

Corolario 5.2.4. *Si dos soluciones de (5.85) homogénea tienen un cero en común en I , entonces son LD en I .*

5.2.3. Reducción a la forma normal de Liouville

La Ecuación (5.85) homogénea puede reducirse a una sin el término y' ; el procedimiento fué introducido por Liouville, y se lo suele llamar *transformación de Liouville*.

Definamos

$$y(x) = u(x)v(x),$$

substituyamos en (5.85) con $f = 0$, y agrupemos u y sus derivadas; obtendremos

$$vu'' + (2v' + rv)u' + (v'' + rv' + sv)u = 0.$$

Ahora elijamos

$$2v' + rv = 0,$$

con lo cual

$$v(x) = \exp\left(-\frac{1}{2} \int^x r(x') dx'\right).$$

Finalmente, definamos

$$\rho(x) = s(x) - \frac{1}{4}r^2(x) - \frac{1}{2}r'(x).$$

Entonces

$$u'' + \rho(x)u = 0. \quad (5.87)$$

Esta última ecuación se dice que está en la *forma normal de Liouville*. Notemos que v , siendo una función exponencial, no se anula en \mathbb{R} , con lo cual los ceros de u coinciden con los de y ; la forma normal permite entonces un estudio más sencillo de los ceros de y . Enunciaremos como ejemplo dos resultados muy potentes que pueden obtenerse fácilmente de (5.87).

Teorema 5.2.5 (de comparación de Sturm). *Sean φ y ψ soluciones no triviales de*

$$\begin{aligned} y'' + \rho_1(x)y &= 0, & x \in I, \\ y'' + \rho_2(x)y &= 0, & x \in I, \end{aligned}$$

respectivamente, y supongamos $\rho_1(x) \geq \rho_2(x) \forall x \in I$. Entonces φ tiene al menos un cero entre cada par de ceros consecutivos de ψ , a menos que $\rho_1(x) \equiv \rho_2(x)$, en cuyo caso φ es un múltiplo constante de ψ .

Demuestração. Notemos primero que tanto φ como ψ pueden asumirse positivas en (x_1, x_2) ; si no lo son, invertimos su signo. Sean x_1 y x_2 dos ceros consecutivos de ψ en I , y supongamos que φ no tiene ceros en (x_1, x_2) . Como ψ' es continua tenemos $\psi'(x_1) \geq 0$ y $\psi'(x_2) \leq 0$, de modo que

$$\begin{aligned} W(\varphi, \psi)(x_1) &= \varphi(x_1)\psi'(x_1) \geq 0, \\ W(\varphi, \psi)(x_2) &= \varphi(x_2)\psi'(x_2) \leq 0 \end{aligned}$$

Pero

$$W'(\varphi, \psi)(x) = \varphi(x)\psi''(x) - \varphi''(x)\psi(x) = [\rho_1(x) - \rho_2(x)]\varphi(x)\psi(x) \geq 0$$

en (x_1, x_2) y por lo tanto W es creciente, lo que es una contradicción, a menos que $\rho_1(x) - \rho_2(x) \equiv 0$ y en consecuencia $W(x) \equiv 0$, en cuyo caso φ y ψ son LD. \square

Corolario 5.2.6. *Sea ψ una solución no trivial de (5.87) en I . Si $\rho(x) \leq 0$ entonces ψ tiene a lo sumo un cero en I .*

Demostración. Supongamos que ψ tiene más de un cero en I , y digamos que x_1 y x_2 son dos ceros consecutivos; entonces toda solución no trivial φ de $y'' = 0$ debería tener al menos un cero en (x_1, x_2) ; sin embargo $\varphi(x) \equiv 1$ es una de estas soluciones, y no se anula, lo que es una contradicción. \square

5.2.4. Soluciones oscilatorias

Una solución no trivial de (5.87) se dice *oscilatoria* si tiene un número infinito de ceros, y por el Teorema 5.2.5, que lo sea o no dependerá del comportamiento de $\rho(x)$. Si $\rho(x) \leq 0$, por el Corolario 5.2.6 no podemos tener soluciones oscilatorias; pero si existe alguna constante $k > 0$ tal que

$$\rho(x) > k^2 > 0,$$

toda solución de (5.87) tendrá una cantidad infinita de ceros distribuidos entre los ceros de cualquier solución de $y'' + k^2y = 0$, como ser $a \sin(x - b)$, cuyos ceros son

$$x_n = b + \frac{n\pi}{k}, \quad n \in \mathbb{Z}.$$

Luego todo intervalo I de longitud π/k contendrá al menos un cero de cualquier solución de (5.87), y a medida que k crezca, la distancia entre ceros sucesivos disminuirá. También podemos concluir que si una solución de (5.87), o de (5.85) homogénea, es oscilatoria, entonces *todas* lo serán.

Ejemplo 5.2.1. La ecuación de Bessel (5.69) puede escribirse en la forma (5.85) como

$$y'' + \frac{1}{x}y' + \left(1 - \frac{\nu^2}{x^2}\right)y = 0,$$

y reducirse a la forma (5.87) poniendo $u = \sqrt{xy}$, lo que da

$$u'' + \left(1 + \frac{1 - 4\nu^2}{4x^2}\right)u = 0.$$

Comparándola con $u'' + u = 0$ vemos que

$$\rho(x) = 1 + \frac{1 - 4\nu^2}{4x^2} = \begin{cases} \geq 1, & 0 \leq \nu \leq 1/2, \\ < 1, & \nu > 1/2, \end{cases}$$

y aplicando el Teorema 5.2.5 concluimos que:

1. Si $0 \leq \nu \leq 1/2$, en cualquier subintervalo de $(0, \infty)$ de largo π toda solución no trivial tiene al menos un cero.
2. Si $\nu > 1/2$, en cualquier subintervalo de $(0, \infty)$ de largo π toda solución no trivial tiene a lo sumo un cero.

3. Si $\nu = 1/2$, la distancia entre ceros sucesivos de cualquier solución no trivial es exactamente π .

Aplicando el Corolario 5.2.6 también concluimos que para $\nu > 1/2$, el intervalo $(0, \sqrt{\nu^2 - 1/4})$ contiene a lo sumo un cero de cualquier solución no trivial.

Ejercicio 5.2.1. Considere la ecuación de movimiento para un oscilador armónico amortiguado y sin forzamiento,

$$\ddot{x} + 2\gamma\dot{x} + \omega^2 x = 0,$$

donde $\omega > 0$ es la frecuencia natural del oscilador y $2\gamma > 0$ es la constante de fricción. Reduzca la ecuación a la forma normal de Liouville y muestre que todas las soluciones no triviales decaen exponencialmente con el tiempo. Halle la relación entre γ y ω para que existan soluciones oscillatorias.

5.3. Problemas de contorno

En la Sección 5.2 hemos considerado las EDO lineales de segundo orden como *problemas de valores iniciales*. Sin embargo muchos problemas de la Física llevan a considerar otra clase de problemas diferente, aunque relacionada a la anterior: los *problemas de contorno*. En un problema de contorno también tratamos con una EDO como (5.85), pero buscamos soluciones que satisfagan condiciones como $y(a) = y_a$ e $y(b) = y_b$. Una de las principales diferencias que encontraremos será que, dada una EDO, aunque cualquier problema de valores iniciales con $y(a)$ e $y'(a)$ especificadas tenga solución, la solución del problema de contorno con $y(a)$ e $y(b)$ dadas bien puede no existir. En esta sección expondremos algunos resultados generales para problemas de contorno.

5.3.1. Operadores diferenciales Hermitianos

Como dijimos, para determinar la solución general de una EDO lineal necesitamos determinar la solución general de su versión homogénea. Comencemos entonces considerando una EDO de la forma

$$p(x)y'' + q(x)y' + r(x)y = 0, \quad x \in (a, b) \subseteq \mathbb{R}, \quad (5.88)$$

donde, para empezar, pediremos que $y \in C^2(a, b)$.

Nuestro problema puede reescribirse ahora como

$$Ly = 0,$$

donde el *operador diferencial lineal* L se define como

$$L = p(x)\frac{d^2}{dx^2} + q(x)\frac{d}{dx} + r(x),$$

y actúa sobre elementos del espacio lineal $\mathcal{C}^2(a, b)$. Resolver (5.88) se traduce así en determinar $\ker L$, y resolver la ecuación inhomogénea

$$Ly = f$$

se traduce en determinar una inversa de L . A este respecto vale la pena recordar que si podemos de alguna forma “diagonalizar” L , esta tarea se simplifica enormemente; y para ello debemos encontrar “autovalores” λ y “autovectores” ϕ tales que

$$L\phi = \lambda\phi.$$

De la Sección 2.1.4 recordaremos que estas tareas resultaban particularmente simples para operadores (matrices, en aquel contexto) *Hermitianos*, para los cuales los autovectores formaban una base ortogonal del espacio en la cual la representación del operador era diagonal; y, según vimos en la Sección 4.2, justamente es en espacios \mathcal{L}^2 donde tenemos nociones bien definidas de ortogonalidad y producto interno. Vamos entonces a ceñirnos a ecuaciones de la forma (5.88) donde el operador L actúe sobre espacios $\mathcal{L}^2(a, b) \cap \mathcal{C}^2(a, b)$ y sea Hermitiano,⁵ en particular

$$L : \mathcal{L}^2(a, b) \cap \mathcal{C}^2(a, b) \rightarrow \mathcal{L}^2(a, b).$$

Debemos entonces, primeramente, determinar qué entenderemos cuando digamos que L es Hermitiano; en particular, debemos definir L^\dagger , y para ello asumiremos por el momento que $p, q, r \in \mathcal{C}^2(a, b)$; nuestra definición, copiada literalmente de las Secciones 1.3.3 y 2.1.4, y consistentemente con lo discutido en la Sección 4.2, será

$$\langle L^\dagger g, f \rangle = \langle g, Lf \rangle \quad \forall f, g \in \mathcal{L}^2(a, b) \cap \mathcal{C}^2(a, b). \quad (5.89)$$

De la definición (4.24) del producto interno y realizando sucesivas integraciones por partes, vemos que

$$\begin{aligned} \langle g, Lf \rangle &= \int_a^b g^*(pf'' + qf' + rf) dx \\ &= [p(f'g^* - fg'^*) + (q - p')fg^*]_a^b + \langle (p^*g)'' - (q^*g)' + r^*g, f \rangle, \end{aligned}$$

de donde

$$\langle g, Lf \rangle = \langle L^\dagger g, f \rangle + [p(f'g^* - fg'^*) + (q - p')fg^*]_a^b \quad (5.90a)$$

con

$$L^\dagger g = (p^*g)'' - (q^*g)' + r^*g = p^*g'' + (2p'^* - q^*)g' + (p''^* - q'^* + r^*)g. \quad (5.90b)$$

Notemos que el lado derecho de (5.90b) está bien definido con sólo exigir $p \in \mathcal{C}^2(a, b)$, $q \in \mathcal{C}^1(a, b)$ y $r \in \mathcal{C}(a, b)$, por lo que ya podemos relajar la exigencia inicial a esta.

⁵La exigencia de Hermiticidad *puede* relajarse, y desarrollaríamos una teoría más general; sin embargo ello resulta innecesario para los problemas que trataremos en este curso, y oscurecería los puntos relevantes. La exigencia de que L actúe sobre $\mathcal{L}^2 \cap \mathcal{C}^2$ también puede relajarse en circunstancias que indicaremos en cada caso.

Ejercicio 5.3.1. Complete los cálculos precedentes.

El operador

$$L^\dagger = p^* \frac{d^2}{dx^2} + (2p'^* - q^*) \frac{d}{dx} + (p''^* - q'^* + r^*)$$

se conoce como el *adjunto formal* de L . El operador L se dice *formalmente autoadjunto* cuando $L^\dagger = L$, es decir cuando

$$p^* = p, \quad 2p'^* - q^* = q, \quad p''^* - q'^* + r^* = r,$$

condiciones que se satisfacen si, y sólo si, p, q y r son reales y $p' = q$.

Ejercicio 5.3.2. Verifique la afirmación precedente.

Si L es formalmente autoadjunto, entonces

$$Lf = pf'' + p'f' + rf = (pf')' + rf,$$

es decir

$$L = \frac{d}{dx} \left(p(x) \frac{d}{dx} \right) + r(x). \quad (5.91)$$

En ese caso (5.90a) se reduce a

$$\langle Lg, f \rangle = \langle g, Lf \rangle + [p(f'g^* - fg'^*)]_a^b, \quad (5.92)$$

por lo que L será *autoadjunto* (o *Hermitiano*) si es formalmente autoadjunto y

$$[p(f'g^* - fg'^*)]_a^b = 0 \quad \forall f, g \in \mathcal{L}^2(a, b) \cap \mathcal{C}^2(a, b). \quad (5.93)$$

Notemos que en este caso basta pedir $p \in \mathcal{C}^1(a, b)$ y $r \in \mathcal{C}(a, b)$, con lo que podemos relajar aún más la exigencia inicial.

5.3.2. Reducción a un operador formalmente autoadjunto

Consideremos una EDO lineal de segundo orden de la forma

$$\tilde{L}(y) = a_0(x)y'' + a_1(x)y' + a_2(x)y = f(x), \quad x \in (a, b), \quad (5.94)$$

con $a_0 \in \mathcal{C}^1[a, b]$ que no se anula en $[a, b]$, y $a_1, a_2 \in \mathcal{C}[a, b]$; podemos suponer sin pérdida de generalidad (SPG) que $a_0(x) > 0$ en $[a, b]$, y esto garantiza que no hay puntos críticos de (5.94) en $[a, b]$. Si $a_1 = a'_0$ entonces \tilde{L} es formalmente autoadjunto; si no, podemos multiplicarlo por un factor integrante $\rho(x)$ para tener

$$L = \rho \tilde{L} = \rho a_0 \frac{d^2}{dx^2} + \rho a_1 \frac{d}{dx} + \rho a_2,$$

que será formalmente autoadjunto si

$$\rho a_1 = (\rho a_0)' = \rho' a_0 + \rho a'_0.$$

La solución general de esta EDO de primer orden es

$$\rho(x) = \frac{c}{a_0(x)} \exp \left(\int^x \frac{a_1(x')}{a_0(x')} dx' \right), \quad (5.95)$$

con c una constante. Notemos que $\rho \in C^1[a, b]$ y es estrictamente positiva (eli-giendo $c > 0$) en $[a, b]$, y se reduce a la constante c si $a_1 = a'_0$. La ecuación (5.94) queda

$$L(y) = \rho f(x), \quad x \in (a, b), \quad (5.96)$$

donde L adopta la forma (5.91) con $p = \rho a_0$ y $r = \rho a_2$. Notemos que $\rho > 0$ en $[a, b]$ garantiza que no introdujimos soluciones espúreas al multiplicar por $\rho(x)$; también garantiza que tanto a como b sean puntos regulares de (5.96), y que ésta no tenga puntos singulares en (a, b) .

En caso de que, digamos, $x = a$ sea un punto singular *regular*, en general aún podremos efectuar la reducción: por ejemplo la ecuación de Bessel (5.69), que tiene un punto singular regular en $x = 0$, puede reescribirse como

$$xy'' + y' + \frac{x^2 - \alpha^2}{x}y = 0, \quad 0 < x < \infty,$$

cuyo operador tiene la forma (5.91) con $p(x) = x$, poniendo $\rho(x) = 1/x$. Lo mismo ocurre en caso de que tanto a como b sean puntos singulares regulares: por ejemplo la ecuación de Legendre (5.20)

$$(1 - x^2)y'' - 2xy' + \alpha(\alpha + 1)y = 0, \quad -1 < x < 1,$$

ya tiene su operador en la forma (5.91) con $p(x) = 1 - x^2$. Debe notarse que en ambos casos p se anula en los puntos singulares, y este comportamiento es genérico, ya que $p(x)$ es el coeficiente de d^2/dx^2 en (5.91).

5.3.3. Problemas de contorno

La posibilidad de reducir EDOs lineales de segundo orden a otras con un operador formalmente autoadjunto hace que muchos problemas físicos lleven a considerar EDO de la forma

$$L(y) = \frac{d}{dx} \left[p(x) \frac{dy}{dx} \right] - q(x)y = f(x), \quad -\infty \leq a < x < b \leq \infty \quad (5.97a)$$

junto con *condiciones de contorno* (que abreviaremos CC), en general dos para una EDO de segundo orden. Si a y b son *finitos*, y $p(a) \neq 0$ y $p(b) \neq 0$ de modo que $x = a$ y $x = b$ son *puntos regulares* de (5.97a), las CC más generales serán de la forma

$$\begin{aligned} \alpha_1 y'(a) + \beta_1 y(a) + \gamma_1 y'(b) + \delta_1 y(b) &= u_1, \\ \alpha_2 y'(a) + \beta_2 y(a) + \gamma_2 y'(b) + \delta_2 y(b) &= u_2, \end{aligned} \quad (5.97b)$$

donde algunos de los coeficientes *reales* α_i , β_i , γ_i y δ_i pueden ser nulos, pero los vectores $(\alpha_1, \beta_1, \gamma_1, \delta_1)$ y $(\alpha_2, \beta_2, \gamma_2, \delta_2)$ deben ser LI; si en cambio $x = a$,

$x = b$ o ambos son *puntos singulares regulares* de (5.97a), en general las CC adoptarán respectivamente la forma

$$\lim_{x \downarrow a} |y(x)| < \infty, \quad \alpha y'(b) + \beta y(b) = u, \quad a \text{ singular}, \quad b \text{ regular}, \quad (5.97c)$$

$$\alpha y'(a) + \beta y(a) = u, \quad \lim_{x \uparrow b} |y(x)| < \infty, \quad a \text{ regular}, \quad b \text{ singular}, \quad (5.97d)$$

$$\lim_{x \downarrow a} |y(x)| < \infty, \quad \lim_{x \uparrow b} |y(x)| < \infty, \quad a \text{ singular}, \quad b \text{ singular}, \quad (5.97e)$$

con α y β coeficientes reales, no ambos nulos. Si $a = -\infty$ y/o $b = \infty$ podemos tener versiones de las CC anteriores expresadas como límites (por ejemplo $\lim_{x \rightarrow \infty} y(x) = c$, $\lim_{x \rightarrow \infty} |y(x)| < \infty$, $\lim_{x \rightarrow -\infty} y'(x) = \lim_{x \rightarrow \infty} y'(x)$) o como *comportamientos asintóticos* requeridos de la solución (por ejemplo $y(x) \underset{x \rightarrow \infty}{\asymp} \sin(x)/x$). Un *problema de contorno* queda definido dando la EDO, el dominio de interés y las CC, *todo junto*.

Las CC pueden ser *separadas*, donde una sólo involucra a $y(a)$ y/o $y'(a)$ y la otra sólo involucra a $y(b)$ y/o $y'(b)$, o *no separadas*, donde al menos una CC involucra valores de y y/o y' tanto en $x = a$ como en $x = b$. Las más comunes, y de hecho las únicas que trataremos en este curso, son:

CC de Dirichlet: separada, $y(a) = u_1$ y/o $y(b) = u_2$.

CC de Neumann: separada, $y'(a) = u_1$ y/o $y'(b) = u_2$.

CC de Robin: separada, $y'(a) + c_1 y(a) = u_1$ y/o $y'(b) + c_2 y(b) = u_2$.

CC periódicas: no separada, $y(b) = y(a)$ e $y'(b) = y'(a)$.

CC de función finita: separada, $\lim_{x \downarrow a} |y(x)| < \infty$ y/o $\lim_{x \uparrow b} |y(x)| < \infty$,

o sus correspondientes versiones asintóticas. Las CC se clasifican además como *homogéneas* si el correspondiente $u_i = 0$, e *inhomogéneas* si el correspondiente $u_i \neq 0$; las CC periódicas son, por definición, homogéneas; las CC de función finita deben entenderse formalmente, por lo general, como homogéneas. Nótese que para CC separadas, podemos tener un tipo de CC (Dirichlet, Neumann, Robin o de función finita) en $x = a$ y otro diferente en $x = b$; si ello ocurre, decimos que las CC son *mixtas*.

A su vez, los problemas de contorno se clasifican como

Homogéneo: $f(x) \equiv 0$ en $[a, b]$ y CC homogéneas.

Inhomogéneo: $f(x) \not\equiv 0$ en $[a, b]$ y/o al menos una CC inhomogénea.

Ejercicio 5.3.3. Muestre que las CC de Dirichlet, de Neumann y de Robin son diferentes casos particulares de la CC separada genérica

$$\frac{d \ln y}{dx}(a) = c_1, \quad \frac{d \ln y}{dx}(b) = c_2, \quad 0 \leq c_1, c_2 \leq \infty.$$

Debe destacarse que la variedad de posibilidades para la EDO, el intervalo de interés (a, b) y las CC imposibilitan una exposición sucinta de la teoría, la cual por otra parte es aún hoy un área de investigación activa en matemática. Con esta salvedad en mente, concluiremos esta sección presentando un par resultados relativamente generales para problemas de contorno.

5.3.4. Homogeneización de las CC

El problema de contorno inhomogéneo más general con CC separadas sobre un intervalo finito,

$$\begin{cases} L(y) = f(x), & a < x < b, \\ \alpha_1 y'(a) + \beta_1 y(a) = u_1, \\ \alpha_2 y'(b) + \beta_2 y(b) = u_2, \end{cases} \quad (5.98)$$

siempre puede ser transformado en uno con CC homogéneas. Para ello basta con hallar una función cualquiera $g(x)$ que cumpla las CC inhomogéneas

$$\begin{cases} \alpha_1 g'(a) + \beta_1 g(a) = u_1, \\ \alpha_2 g'(b) + \beta_2 g(b) = u_2 \end{cases} \quad (5.99)$$

pero aparte de ello sea arbitraria para $a < x < b$. Definiendo la nueva variable dependiente \tilde{y} como

$$\tilde{y}(x) := y(x) - g(x), \quad (5.100)$$

es inmediato que

$$\begin{cases} L(\tilde{y}) = f(x) - L(g)(x), & a < x < b, \\ \alpha_1 \tilde{y}'(a) + \beta_1 \tilde{y}(a) = 0, \\ \alpha_2 \tilde{y}'(b) + \beta_2 \tilde{y}(b) = 0. \end{cases} \quad (5.101)$$

Por supuesto, si es posible usar la considerable arbitrariedad en la elección de g para simplificar la nueva inhomogeneidad $f - L(g)$, tanto mejor.

La extensión de esta técnica a intervalos infinitos con las correspondientes CC dadas en forma asintótica, es inmediata.

5.3.5. Identidad de Lagrange y fórmula de Green

Consideremos dos funciones $y(x)$ y $z(x)$ que satisfacen las EDO inhomogéneas

$$L(y) = f(x), \quad a < x < b, \quad (5.102a)$$

$$L(z) = g(x), \quad a < x < b. \quad (5.102b)$$

Multiplicando la primera ecuación por z y la segunda por y y restando tenemos

$$z(p y')' - y(p z')' = fz - gy, \quad (5.103)$$

y dado que $[p(zy' - yz')]' = z(py')' - y(pz')'$, obtenemos la *identidad de Lagrange*

$$\frac{d}{dx} \left[p(x) \left(z \frac{dy}{dx} - y \frac{dz}{dx} \right) \right] = f(x)z(x) - g(x)y(x). \quad (5.104)$$

Integrando ahora desde a hasta b resulta la *fórmula de Green*

$$\int_a^b [zL(y) - yL(z)] dx = \left[p(x) \left(z \frac{dy}{dx} - y \frac{dz}{dx} \right) \right]_a^b = \int_a^b [f(x)z(x) - g(x)y(x)] dx. \quad (5.105)$$

Debemos notar que, si las ecuaciones (5.102) corresponden a un mismo problema de contorno con CC homogéneas, es decir si y y z satisfacen *las mismas CC* homogéneas, el segundo término en la fórmula de Green se anula. Notemos que la fórmula de Green sigue siendo válida para intervalos infinitos, a condición de que las integrales involucradas converjan.

Ejercicio 5.3.4. Demuestre la afirmación precedente.

Si ϕ_1 y ϕ_2 son dos soluciones de $L(y) = 0$, la identidad de Lagrange (5.104) implica

$$p(x) \left(\phi_1 \frac{d\phi_2}{dx} - \phi_2 \frac{d\phi_1}{dx} \right) = c \text{ (constante)}, \quad (5.106)$$

de donde

$$W(\phi_1, \phi_2)(x) = \frac{c}{p(x)}. \quad (5.107)$$

Si ϕ_1 y ϕ_2 son LI en $[a, b]$ la constante c será no nula, y si son LD tendremos $c = 0$. Nótese que lo anterior implica que, dadas dos soluciones LI de (5.97a), en cada punto singular, donde p debe anularse, al menos una de ellas será singular.

5.4. Problemas de Sturm–Liouville

Denotemos por L el operador formalmente autoadjunto definido por

$$L(y) = \frac{d}{dx} \left(p(x) \frac{dy}{dx} \right) - q(x)y \quad (5.108)$$

donde $p \in \mathcal{C}^1(a, b)$ y $q \in \mathcal{C}^0(a, b)$ (reales), y $p(x) > 0 \forall x \in (a, b)$. El extremo a del intervalo puede ser $-\infty$ y el extremo b puede ser ∞ . Recordemos que la positividad de p implica que (a, b) no contiene *ningún* punto singular de L , aunque tanto a como b *pueden* ser puntos singulares, en cuyo caso p se anula allí; en lo que sigue, asumiremos que si a y/o b son puntos singulares, son *regulares* en el sentido de la Sección 5.1.3.

Denotemos por $\mathcal{C}_H^2(a, b) \subset \mathcal{C}^2(a, b)$ el conjunto de funciones $\mathcal{C}^2(a, b)$ que satisfacen CC tales que se verifica (5.93), es decir

$$[p(f^*g' - f'^*g)]_a^b = 0 \quad \forall f, g \in \mathcal{C}_H^2(a, b). \quad (5.109)$$

Tales CC se dicen *Hermitianas* (o *autoadjuntas*) para L , ya que hacen de L un operador Hermitiano. Notemos que $\mathcal{C}_H^2(a, b)$ es un espacio lineal. Notemos también que, usando que tanto f como g satisfacen las mismas CC, puede mostrarse que las CC listadas en la Sección 5.3.3 son Hermitianas si son homogéneas, con pocas exigencias adicionales (o ninguna):

CC separadas: Si en $x = a$ las CC son de Dirichlet, entonces $f(a) = g(a) = 0$.

Si son de Neumann, entonces $f'(a) = g'(a) = 0$. Y si son de Robin, entonces $f'(a) = -cf(a)$ y $g'(a) = -cg(a)$ con la misma constante real c , con lo cual $f^*(a)g'(a) = f'^*(a)g(a)$. El mismo razonamiento muestra la anulación de $f^*g' - f'^*g$ en $x = b$.

CC periódicas: Es inmediato que

$$f^*(b)g'(b) - f'^*(b)g(b) = f^*(a)g'(a) - f'^*(a)g(a),$$

de modo que

$$[p(f^*g' - f'^*g)]_a^b = 0$$

a condición de que $p(b) = p(a)$.

CC de función finita: Supongamos que $x = a$ es un punto singular regular de L . Aplicando a (5.66) la reducción de la Sección 5.3.2, es fácil concluir que

$$p(x) \underset{x \downarrow a}{\asymp} (x - a).$$

Por otra parte, f y g son al menos $C^1(a, b)$ y $\mathcal{C}[a, b]$; si son funciones reales, el Teorema del Valor Medio (TVM) garantiza que

$$\exists c \in (a, b) / f(x) - f(a) = (x - a)f'(c)$$

y equivalentemente para g , con lo cual

$$p(x)f'(x) \underset{x \downarrow a}{\asymp} f(x) - f(a) \xrightarrow{x \downarrow a} 0,$$

y por lo tanto $p f' g \rightarrow 0$ ya que $\lim_{x \downarrow a} |g(x)| < \infty$; lo mismo vale para $p f g'$. El mismo razonamiento puede aplicarse si $x = b$ es un punto singular regular. El resultado puede extenderse a funciones f, g complejas holomorfas.⁶

⁶Las posibles generalizaciones del TVM a funciones complejas son aún hoy materia de investigación activa, pero la extensión del resultado a funciones complejas holomorfas puede hacerse prescindiendo de ellas: basta aplicar el TVM a $\operatorname{Re} f(ta + (1-t)b)$ e $\operatorname{Im} f(ta + (1-t)b)$ con $t \in [0, 1]$, ya que aunque usualmente si f es holomorfa $\operatorname{Re} f$ no es diferenciable (o si lo es, f es constante), en este caso estamos considerando $\operatorname{Re} f(\gamma(t))$ con γ una parametrización del intervalo real $(a, b) \subset \mathbb{C}$, y usando la continuidad de $\operatorname{Re} f$ puede mostrarse que el límite

$$\lim_{\Delta t \rightarrow 0} \frac{\operatorname{Re} f(\gamma(t + \Delta t)) - \operatorname{Re} f(\gamma(t))}{\Delta t}$$

existe; lo mismo ocurre con $\operatorname{Im} f$.

Se conoce como *problemas de Sturm-Liouville* (abreviaremos PSL) a los problemas de contorno de la forma

$$L(y) + \lambda\rho(x)y = 0, \quad a < x < b, \quad (5.110)$$

con L formalmente autoadjunto, $\rho \in \mathcal{C}[a, b]$ positiva en (a, b) y CC Hermitianas para L . Si el intervalo (a, b) es finito y $[a, b]$ no contiene puntos singulares de L , el PSL se dice *regular*; caso contrario, se lo llama *singular*. Toda constante λ para la que exista solución no trivial se llama un *autovalor* del problema, y la correspondiente solución es una *autofunción* correspondiente a ese autovalor. El factor $\rho(x)$ se incluye a fin de tener en cuenta los casos donde para hacer a L formalmente autoadjunto se ha debido recurrir a la reducción de la Sección 5.3.2.

Los autovalores y autofunciones de un PSL poseen una cantidad de propiedades útiles, algunas de las cuales listaremos a continuación, aunque no las demostraremos todas. Debemos destacar que la teoría conserva cierto grado de sencillez sólo para problemas regulares con CC separadas; sin embargo, la mayoría de los problemas de interés físico son singulares, hecho escasamente destacado en la literatura. Afortunadamente la mayoría de los resultados importantes pueden extenderse a estos casos, y así lo indicaremos donde sea relevante.

5.4.1. Autovalores

Los autovalores de un PSL son reales. La demostración de este resultado es exactamente la misma que la dada en el Teorema 2.1.3 para autovalores de una matriz Hermitiana, pero la repetiremos en este contexto con propósitos ilustrativos.

Demostración. Supongamos

$$Lf + \lambda\rho f = 0$$

con f no trivial. Entonces

$$\lambda\|f\|_\rho^2 = \lambda\langle f, f \rangle_\rho = \langle f, \lambda f \rangle_\rho = -\langle f, Lf \rangle.$$

Como L es Hermitiano,

$$-\langle f, Lf \rangle = -\langle Lf, f \rangle = -\langle \lambda f, f \rangle_\rho = \lambda^* \langle f, f \rangle_\rho = \lambda^* \|f\|_\rho^2.$$

Luego

$$(\lambda - \lambda^*)\|f\|_\rho^2 = 0,$$

y como $\|f\|_\rho > 0$ por ser f no trivial,

$$\lambda = \lambda^*.$$

□

El resultado es válido para cualquier PSL, regular o singular, a condición de que el producto interno y la norma estén bien definidos.

5.4.2. Subespacios característicos

Dado un autovalor λ de un PSL, éste tiene asociadas a lo sumo dos autofunciones LI, es decir, $\dim \ker(L + \lambda\rho) \leq 2$; si las CC son separadas, cada autovalor tiene asociada una única autofunción, es decir, $\dim \ker(L + \lambda\rho) = 1$. Lo demostraremos para intervalos *finitos*.

Demostración. El operador diferencial $L + \lambda\rho$ es lineal y de segundo orden, por lo que el problema de valores iniciales

$$(L + \lambda\rho)y = 0, \quad y(a) = y_a, \quad y'(a) = y'_a$$

tendrá dos soluciones no triviales LI, y no más. De ellas al menos una satisface las CC en $x = b$, porque λ es un autovalor; si ambas lo hacen, el espacio de soluciones es un subespacio 2-dimensional de $C^2(a, b)$.

Una CC separada en $x = a$ de la forma $\alpha y'(a) + \beta y(a) = 0$ con α y β no ambos nulos, selecciona un subespacio 1-dimensional $\alpha y'_a + \beta y_a = 0$ del espacio 2-dimensional de las CI y_a e y'_a , y por lo tanto una única autofunción.

Una CC separada de la forma $\lim_{x \downarrow a} |g(x)| < \infty$ en un punto singular regular de L selecciona sólo las soluciones finitas en $x = a$, que por el Teorema 5.1.1 serán al menos una y a lo sumo dos; pero por (5.107) vemos que $p(a) = 0$ implica que W diverge en a , así que en general sólo una solución será finita en a .

Finalmente, una CC periódica (*¡no separada!*) $y(a) = y(b)$, $y'(a) = y'(b)$ en general no restringe la elección de las CI y_a e y'_a y deja dos autofunciones LI. \square

Si $a = -\infty$ y/o $b = \infty$, el problema de valores iniciales

$$(L + \lambda\rho)y = 0, \quad y(c) = y_c, \quad y'(c) = y'_c, \quad a < c < b$$

tendrá dos soluciones no triviales LI, y no más, y de ellas al menos una satisface las CC en $x = a$ y $x = b$ ya que λ es un autovalor; las CC asintóticas pueden ser lo suficientemente exigentes para reducir este número a una, pero pueden no serlo. El resultado $\dim \ker(L + \lambda\rho) \leq 2$ se extiende entonces trivialmente a estos casos, pero la reducción a $\dim \ker(L + \lambda\rho) = 1$ no puede garantizarse en general.

5.4.3. Ortogonalidad de las autofunciones

Las autofunciones de un PSL correspondientes a autovalores diferentes, son ortogonales en $\mathcal{L}_\rho^2(a, b)$ (*i.e.*, ortogonales con peso ρ); para autovalores degenerados, las autofunciones LI correspondientes a un mismo autovalor pueden orthogonalizarse entre sí. Para autovalores diferentes, la demostración es exactamente la misma que la dada en el Teorema 2.1.3 para autovectores de una matriz Hermitiana, pero la repetiremos en este contexto con propósitos ilustrativos.

Demostración. Supongamos $\mu \neq \lambda$ y

$$Lf + \lambda\rho f = 0, \quad Lg + \mu\rho g = 0$$

con f, g no triviales. Entonces, por ser L Hermitiano y λ, μ reales,

$$\lambda \langle g, f \rangle_\rho = -\langle g, Lf \rangle = -\langle Lg, f \rangle = \mu \langle g, f \rangle_\rho,$$

de donde

$$(\lambda - \mu) \langle g, f \rangle_\rho = 0$$

Pero $\lambda - \mu \neq 0$, luego

$$\langle g, f \rangle_\rho = 0.$$

Si y_1 e y_2 son dos autofunciones LI correspondientes a un dado autovalor λ , cualquier combinación lineal de ellas también lo será. Podemos obtener entonces dos autofunciones ortogonales a partir de y_1 e y_2 por cualquier procedimiento de ortogonalización, por ejemplo el de Gram-Schmidt. \square

El resultado es válido para cualquier PSL, regular o singular, a condición de que el producto interno y la norma estén bien definidos.

5.4.4. Espectro de autovalores

Este es uno de los puntos donde la variedad de dominios, CC y comportamientos de p y q hacen imposible listar todos los resultados conocidos, y aún más demostrarlos. Enunciaremos sin demostración los más importantes, aunque intentando en lo posible delimitar su validez.⁷

El espectro de un PSL es generalmente *acotado por debajo*, es decir, existe un $\lambda_{\inf} > -\infty$ tal que $\lambda_{\inf} \leq \lambda$ para todo autovalor λ . Una excepción es $(a, b) = (0, \infty)$ con la función $\rho(x)$ de la EDO reducida a su forma normal de Liouville (5.87) que satisface $\rho(x) \xrightarrow{x \rightarrow \infty} -\infty$.

El espectro puede contener una parte *discreta* y una *continua*. El caso más común donde se presentan ambos es cuando $(a, b) = (0, \infty)$ y en (5.87) tenemos $\rho(x) \xrightarrow{x \rightarrow \infty} 0$; en ese caso habrá un espectro discreto de autovalores $\lambda \in (\lambda_{\inf}, 0)$ y uno continuo con $\lambda \in (0, \infty)$; un ejemplo es el espectro del PSL radial para el átomo de hidrógeno en Mecánica Cuántica. Si en cambio $(a, b) = (0, \infty)$ o $(-\infty, \infty)$ y en (5.87) tenemos $\rho(x) \xrightarrow{|x| \rightarrow \infty} \infty$, el espectro es puramente discreto; un ejemplo es el espectro del PSL para el oscilador armónico en Mecánica Cuántica.

Para PSL con (a, b) finito y CC separadas, *regulares* o *singulares* con a y/o b puntos singulares regulares de L , el espectro es puramente discreto, con

$$-\infty < \lambda_1 < \lambda_2 < \cdots < \lambda_n < \cdots, \quad \lambda_n \xrightarrow{n \rightarrow \infty} \infty.$$

⁷El lector interesado puede consultar varios textos excelentes, entre ellos, en orden de creciente dificultad técnica:

- M. A. Al-Gwaiz, *Sturm-Liouville Theory and its Applications*. Springer Undergraduate Mathematics Series, 2008.
 R. B Guenther y J. W. Lee, *Sturm-Liouville Problems. Theory and Numerical Implementation*. CRC Press, 2018.
 E. C. Titchmarsh, *Eigenfunction Expansions associated with Second-order Differential Equations*. Part One, 2nd. ed. Oxford University Press, 1962.

Si las CC son periódicas (no separadas) el resultado aún es válido pero recordando que cada λ_i tiene multiplicidad 2 (aunque en general λ_1 tiene multiplicidad 1).

5.4.5. Completitud de las autofunciones

Las autofunciones $\varphi_\lambda(x)$ correspondientes a los diferentes autovalores λ del PSL

$$(L + \lambda\rho)\varphi_\lambda = 0, \quad \text{CC Hermitianas en } a \text{ y } b,$$

forman un *conjunto completo en $\mathcal{L}_\rho^2(a, b)$* según se lo definió en la Sección 4.2.3, es decir tendremos

$$f \stackrel{\mathcal{L}_\rho^2}{=} \sum_{\lambda} \frac{\langle \varphi_\lambda, f \rangle_\rho}{\langle \varphi_\lambda, \varphi_\lambda \rangle_\rho} \varphi_\lambda$$

para toda $f \in \mathcal{L}_\rho^2(a, b)$. Notemos que si existen autovalores degenerados, en la suma de arriba deberemos incluir *las dos* autofunciones correspondientes, adecuadamente ortogonalizadas.

Este resultado es válido para todos los casos donde el espectro es discreto, y puede extenderse a la mayoría de los casos donde el espectro es mixto o continuo, al menos si está acotado por debajo. En ese caso además de la suma tendremos una integral sobre λ para los autovalores del espectro continuo.

Una demostración rigurosa de estos resultados pasa por la definición de una *función de Green* para el PSL; sin embargo, si bien definiremos (y usaremos) funciones de Green en el Capítulo 11, la demostración es larga y técnica, por lo cual la omitiremos; el lector interesado puede hallarla en cualquiera de los textos citados en la sección anterior.

5.5. La notación de Dirac

El mismo P.A.M. Dirac que introdujo la distribución δ , introdujo también una notación muy conveniente para la clase de problemas que venimos tratando. Dirac notó que el producto interno

$$\langle g, f \rangle = \int g^*(x)f(x) dx$$

era similar al de dos vectores \vec{u} y \vec{v} escrito como

$$\langle \vec{u}, \vec{v} \rangle = \sum_i u_i^* v_i$$

donde u_i y v_i son las componentes de \vec{u} y \vec{v} en una base Cartesiana, con la diferencia que, para funciones, el “índice de componente” es una variable continua x . Ahora bien, para vectores existe una forma más compacta de escribir $\langle \vec{u}, \vec{v} \rangle$ que nos ahorra escribir explícitamente la suma: simplemente escribimos ambos vectores como matrices columna

$$\mathbf{u} = \begin{pmatrix} u_1 \\ u_2 \\ \vdots \end{pmatrix}, \quad \mathbf{v} = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \end{pmatrix}$$

y ponemos

$$\langle \vec{u}, \vec{v} \rangle = u^\dagger v.$$

La regla de multiplicación matricial se encarga del resto, y el producto interno queda indicado simplemente por la *yuxtaposición* de u^\dagger y v .

Para hacer lo mismo con funciones, Dirac reescribió el corchete (*bracket*) $\langle g, f \rangle$ como $\langle g|f \rangle$ y lo interpretó como la *yuxtaposición* del *bra* $\langle g|$ y el *ket* $|f\rangle$, “simplificando” la doble barra central resultante:

$$\langle g| |f\rangle = \langle g|f \rangle = \langle g, f \rangle.$$

Así, la integral está *implícita*, igual que lo está la suma al multiplicar matrices por simple *yuxtaposición*.

En esta notación, representaremos a una función $f(x)$ de algún espacio V por el ket $|f\rangle$, a una función $g(x)$ de su dual V^* por el bra $\langle g|$, y, aprovechando la Hermiticidad de L , pondremos

$$\langle g, Lf \rangle = \langle Lg, f \rangle = \langle g|L|f\rangle,$$

donde podemos pensar a L actuando “hacia adelante” sobre $|f\rangle$ o “hacia atrás” sobre $\langle g|$, de manera completamente análoga a como escribiríamos

$$u^\dagger Lv$$

con matrices.

Si nuestro producto interno está definido con una función de peso $\rho(x)$, simplemente lo asumiremos definiendo

$$\langle g, f \rangle_\rho = \langle g|f \rangle = \int dx \rho(x) g^*(x) f(x)$$

y recuperando la definición de $\langle g, f \rangle$ cuando $\rho(x) \equiv 1$. En general el contexto dejará claro cual es la función de peso que deberemos usar al escribir integrales.

5.5.1. Ortogonalidad y completitud

Usando la notación de Dirac, la ortogonalidad de las autofunciones de un PSL se expresa como

$$\langle \varphi_\lambda | \varphi_\mu \rangle = \delta_{\lambda\mu} \langle \varphi_\mu | \varphi_\mu \rangle,$$

y su completitud como

$$|f\rangle = \sum_\lambda \frac{\langle \varphi_\lambda | f \rangle}{\langle \varphi_\lambda | \varphi_\lambda \rangle} |\varphi_\lambda \rangle.$$

Definiendo las *autofunciones normalizadas*

$$\psi_\lambda(x) = \frac{\varphi_\lambda(x)}{\|\varphi_\lambda(x)\|_\rho},$$

la relación de ortogonalidad ahora será una de *ortonormalidad*,

$$\langle \psi_\lambda | \psi_\mu \rangle = \delta_{\lambda\mu},$$

y la de completitud se escribirá más sencillamente, como

$$|f\rangle = \sum_{\lambda} \langle \psi_\lambda | f \rangle |\psi_\lambda\rangle.$$

5.5.2. La delta de Dirac

Vamos a obtener ahora un primer resultado cuya deducción es enormemente simplificada por la notación de Dirac. Reescribiendo la relación de completitud como

$$|f\rangle = \sum_{\lambda} |\psi_\lambda\rangle \langle \psi_\lambda | f \rangle = \left(\sum_{\lambda} |\psi_\lambda\rangle \langle \psi_\lambda| \right) |f\rangle, \quad (5.111)$$

notamos varias cosas: primero, que un símbolo como $|\psi_\lambda\rangle \langle \psi_\lambda|$ debe interpretarse como un *operador*, ya que $(|\psi_\lambda\rangle \langle \psi_\lambda|) |f\rangle \in V$, $\langle g| (|\psi_\lambda\rangle \langle \psi_\lambda|) \in V^*$, y

$$\langle g| (|\psi_\lambda\rangle \langle \psi_\lambda|) |f\rangle = \langle g| \psi_\lambda \rangle \langle \psi_\lambda | f \rangle$$

será un escalar; segundo, que este operador en particular,

$$P_\lambda = |\psi_\lambda\rangle \langle \psi_\lambda|,$$

será el *proyector* (ortogonal) sobre $|\psi_\lambda\rangle$; y tercero, que el operador

$$I = \sum_{\lambda} |\psi_\lambda\rangle \langle \psi_\lambda|$$

es efectivamente el *operador identidad*, ya que $I |f\rangle \equiv |f\rangle$. La completitud de las autofunciones, entonces, significa que si sumamos todas las proyecciones de cualquier función sobre cada autofunción, reobtenemos la función original; nada se pierde por el camino (la igualdad, por supuesto, debe entenderse en \mathcal{L}^2).

Ahora, “retraduciendo” (5.111) a nuestra vieja notación, obtendremos un resultado fundamental: tendremos

$$f(x) = \sum_{\lambda} \psi_\lambda(x) \int dx' \rho(x') \psi_\lambda(x') f(x') = \int \left(\sum_{\lambda} \psi_\lambda(x) \rho(x') \psi_\lambda(x') \right) f(x') dx',$$

de lo cual es inmediato que

$$\sum_{\lambda} \psi_\lambda(x) \rho(x') \psi_\lambda(x') = \delta(x - x'). \quad (5.112)$$

En muchos contextos esta es tomada como la *definición* de la relación de completitud.

5.5.3. Problemas inhomogéneos

Finalmente, veremos cómo la notación de Dirac nos permite resolver con sencillez problemas de contorno *inhomogéneos*. Supongamos que queremos resolver

$$(L + \rho\mu)y = \rho f, \quad a < x < b,$$

donde f es una inhomogeneidad dada, y ya hemos hallado los autovalores λ y las autofunciones normalizadas $\psi_\lambda(x)$ del correspondiente PSL (homogéneo)

$$(L + \rho\lambda)y = 0, \quad a < x < b.$$

Reescribiendo el problema como

$$(L + \rho\mu)|y\rangle = \rho|f\rangle,$$

usando la completitud para desarrollar

$$|y\rangle = \sum_{\lambda} |\psi_{\lambda}\rangle \langle \psi_{\lambda}|y\rangle, \quad |f\rangle = \sum_{\lambda} |\psi_{\lambda}\rangle \langle \psi_{\lambda}|f\rangle$$

y recordando que $L|\psi_{\lambda}\rangle = -\rho\lambda|\psi_{\lambda}\rangle$, tendremos

$$\sum_{\lambda} (\mu - \lambda) |\psi_{\lambda}\rangle \langle \psi_{\lambda}|y\rangle = \sum_{\lambda} |\psi_{\lambda}\rangle \langle \psi_{\lambda}|f\rangle.$$

Multiplicando ahora a izquierda por $\langle \psi_{\lambda'}|$ y usando la ortonormalidad $\langle \psi_{\lambda'}|\psi_{\lambda}\rangle = \delta_{\lambda'\lambda}$, resulta

$$(\mu - \lambda) \langle \psi_{\lambda}|y\rangle = \langle \psi_{\lambda}|f\rangle,$$

de donde

$$\langle \psi_{\lambda}|y\rangle = \frac{\langle \psi_{\lambda}|f\rangle}{\mu - \lambda},$$

y substituyendo en el desarrollo para $|y\rangle$ obtenemos

$$|y\rangle = \sum_{\lambda} |\psi_{\lambda}\rangle \frac{\langle \psi_{\lambda}|f\rangle}{\mu - \lambda} = \left(\sum_{\lambda} \frac{|\psi_{\lambda}\rangle \langle \psi_{\lambda}|}{\mu - \lambda} \right) |f\rangle.$$

Vemos así que el operador

$$\sum_{\lambda} \frac{|\psi_{\lambda}\rangle \langle \psi_{\lambda}|}{\mu - \lambda}$$

actúa como una *inversa formal* del operador $L + \rho\mu$. Notemos que, al igual que en el problema $(L - \mu I)v = u$ no existe solución si μ coincide con algún autovalor λ de L y u tiene proyección no trivial sobre $\ker(L - \lambda I)$, aquí no existirá solución si μ coincide con uno de los autovalores λ y $\langle \psi_{\lambda}|f\rangle \neq 0$.

Capítulo 6

Funciones Especiales

Caveat lector (otra vez). El tema de las llamadas *funciones especiales* abarca una variedad enorme de funciones que han demostrado su utilidad en la Física; en este capítulo trataremos sólo unas pocas de aquellas relacionadas a soluciones de problemas de Sturm–Liouville de ocurrencia común. El lector interesado puede consultar, por ejemplo, la *NIST Digital Library of Mathematical Functions* (<https://dlmf.nist.gov/>) para darse una idea de la vastedad del tema, y aprender más sobre el mismo.

6.1. PSL para la ecuación armónica

El primer caso que consideraremos es el de la *ecuación armónica*, que en el caso de CC periódicas nos llevará al desarrollo de funciones en *serie de Fourier*.

Consideraremos el problema de contorno

$$y'' + \lambda y = 0, \quad 0 < x < \ell, \tag{6.1a}$$

$$y(0) = y(\ell), \quad y'(0) = y'(\ell). \tag{6.1b}$$

El operador $L = \partial_x^2$ es formalmente autoadjunto, ya que tiene la forma (5.108) con $p(x) = 1$ y $q(x) = 0$. Las CC son Hermitianas para L , que resulta así Hermítano. Comparando con (5.110) vemos que (6.1) es un PSL regular con función de peso $\rho(x) = 1$.

6.1.1. Autovalores y autofunciones

Para $\lambda < 0$ la solución general de (6.1a) es

$$y(x) = a \cosh(\sqrt{-\lambda} x) + b \operatorname{senh}(\sqrt{-\lambda} x).$$

Las CC (6.1b) imponen que

$$\begin{aligned} a &= a \cosh(\sqrt{-\lambda} \ell) + b \operatorname{senh}(\sqrt{-\lambda} \ell), \\ b &= a \operatorname{senh}(\sqrt{-\lambda} \ell) + b \cosh(\sqrt{-\lambda} \ell). \end{aligned}$$

Este sistema de ecuaciones para a y b tiene solución no trivial si, y sólo si,

$$\det \begin{pmatrix} \cosh(\sqrt{-\lambda}\ell) - 1 & \operatorname{senh}(\sqrt{-\lambda}\ell) \\ \operatorname{senh}(\sqrt{-\lambda}\ell) & \cosh(\sqrt{-\lambda}\ell) - 1 \end{pmatrix} = 2(1 - \cosh(\sqrt{-\lambda}\ell)) = 0,$$

que no tiene solución para $\lambda < 0$. Luego en este caso sólo tenemos la solución trivial.

Para $\lambda = 0$ la solución general es

$$y(x) = ax + b,$$

e imponer las CC requiere $a = 0$; la solución es entonces la función constante, que tomaremos como $y(x) = 1$.

Para $\lambda > 0$ es conveniente definir $\lambda = k^2$; la solución general es

$$y(x) = a \cos(kx) + b \operatorname{sen}(kx).$$

Las CC imponen que

$$a = a \cos(k\ell) + b \operatorname{sen}(k\ell),$$

$$b = -a \operatorname{sen}(k\ell) + b \cos(k\ell).$$

Este sistema de ecuaciones para a y b tiene solución no trivial si, y sólo si,

$$\det \begin{pmatrix} \cos(k\ell) - 1 & \operatorname{sen}(k\ell) \\ -\operatorname{sen}(k\ell) & \cos(k\ell) - 1 \end{pmatrix} = 2(1 - \cos(k\ell)) = 0,$$

es decir si

$$k\ell = 2n\pi, \quad n \in \mathbb{N}.$$

En ese caso la matriz de coeficientes $\begin{pmatrix} \cos(k\ell) & \operatorname{sen}(k\ell) \\ \operatorname{sen}(k\ell) & \cos(k\ell) \end{pmatrix}$ se reduce a la identidad, y cualquier par de constantes a, b es solución; esto ocurre porque el determinante tiene una *raíz doble* para $k\ell = 2n\pi$, e implica que a cada autovalor le corresponden *dos* autofunciones LI.

Resumiendo, los autovalores λ_n y las autofunciones $y_n(x)$ del PSL (6.1) son

$$\begin{aligned} \lambda_0 &= 0 \longleftrightarrow y_0(x) = 1, \\ \lambda_n &= k_n^2 = \left(\frac{2n\pi}{\ell}\right)^2 \longleftrightarrow y_n(x) = \begin{cases} \cos(k_n x) \\ \operatorname{sen}(k_n x) \end{cases}, \quad n \in \mathbb{N}. \end{aligned} \tag{6.2a}$$

La notación $\begin{cases} \cos(k_n x) \\ \operatorname{sen}(k_n x) \end{cases}$ se entenderá, de ahora en adelante, como “cualquier combinación lineal de $\cos(k_n x)$ y $\operatorname{sen}(k_n x)$ ”.

Un conjunto alternativo de autofunciones puede obtenerse si admitimos como soluciones funciones complejas. En ese caso tendremos

$$\lambda_n = k_n^2 = \left(\frac{2n\pi}{\ell}\right)^2 \longleftrightarrow y_n(x) = e^{ik_n x}, \quad n \in \mathbb{Z}. \tag{6.2b}$$

Los autovalores λ_n son, por supuesto, los mismos, y a cada uno (salvo λ_0) le corresponden las dos autofunciones $e^{\pm 2in\pi z/\ell}$.

Ejercicio 6.1.1. Muestre que las autofunciones (6.2b) pueden obtenerse de las autofunciones (6.2a) y viceversa.

6.1.2. Desarrollo en autofunciones

De acuerdo a lo visto en la Sección 5.4, las autofunciones (6.2) del PSL (6.1) forman un conjunto ortogonal y completo en $\mathcal{L}^2(0, \ell)$. Ello quiere decir que, dada cualquier función $f \in \mathcal{L}^2(0, \ell)$, tendremos

$$f(x) = a_0 + \sum_{n=1}^{\infty} [a_n \cos(k_n x) + b_n \sin(k_n x)] = \sum_{n=-\infty}^{\infty} c_n e^{ik_n x}. \quad (6.3)$$

Estas expresiones se conocen como el *desarrollo en serie de Fourier* real y compleja, respectivamente, de la función f . La igualdad, así como la convergencia de la serie, deben entenderse, por supuesto, en \mathcal{L}^2 .

Para obtener los coeficientes $a_n, b_n \in \mathbb{R}$ o $c_n \in \mathbb{C}$, apelamos a la *ortogonalidad* de las autofunciones discutida en la Sección 5.4. Tendremos

$$\begin{aligned} a_0 &= \frac{\langle 1, f \rangle}{\langle 1, 1 \rangle}, & a_n &= \frac{\langle \cos(k_n x), f \rangle}{\langle \cos(k_n x), \cos(k_n x) \rangle}, & b_n &= \frac{\langle \sin(k_n x), f \rangle}{\langle \sin(k_n x), \sin(k_n x) \rangle}, \\ c_n &= \frac{\langle e^{ik_n x}, f \rangle}{\langle e^{ik_n x}, e^{ik_n x} \rangle}. \end{aligned} \quad (6.4)$$

En cualquier aplicación resultará conveniente tener calculadas de antemano las *integrales de normalización*

$$\begin{aligned} \langle 1, 1 \rangle &= \int_0^\ell dx = \ell, \\ \langle \cos(k_n x), \cos(k_n x) \rangle &= \int_0^\ell \cos^2(k_n x) dx = \frac{\ell}{2}, \\ \langle \sin(k_n x), \sin(k_n x) \rangle &= \int_0^\ell \sin^2(k_n x) dx = \frac{\ell}{2}, \\ \langle e^{ik_n x}, e^{ik_n x} \rangle &= \int_0^\ell e^{-ik_n x} e^{ik_n x} dx = \int_0^\ell dx = \ell. \end{aligned} \quad (6.5)$$

Ejercicio 6.1.2. Verifique los cálculos precedentes.

Ejemplo 6.1.1. Hallemos el desarrollo en serie de Fourier de la función $f(x) = x$ en el intervalo $(0, \ell)$. Tendremos

$$\begin{aligned} a_0 &= \frac{1}{\ell} \langle 1, f \rangle = \frac{1}{\ell} \int_0^\ell x dx = \frac{\ell}{2}, \\ a_n &= \frac{2}{\ell} \langle \cos(k_n x), f \rangle = \frac{2}{\ell} \int_0^\ell x \cos(k_n x) dx = 0, \\ b_n &= \frac{2}{\ell} \langle \sin(k_n x), f \rangle = \frac{2}{\ell} \int_0^\ell x \sin(k_n x) dx = -\frac{2}{k_n}, \end{aligned}$$

con lo cual

$$x = \frac{\ell}{2} - \sum_{n=1}^{\infty} \frac{2}{k_n} \sin\left(\frac{2n\pi}{\ell} x\right).$$

Ejercicio 6.1.3. Complete los cálculos precedentes; halle también el desarrollo en serie de Fourier compleja de $f(x) = x$, $x \in (0, \ell)$.

6.1.3. Funciones periódicas

Diremos que una función $f(x)$ es *periódica en x con periodo p* si

$$f(x + p) = f(x) \quad \forall x \in \mathbb{R}. \quad (6.6)$$

Notemos que una función periódica de periodo p también es periódica de periodos $2p$, $3p$, etc. Suele definirse el *periodo minimal* como

$$\inf \{ p \in \mathbb{R}_+ \mid f(x + p) = f(x) \quad \forall x \in \mathbb{R} \}. \quad (6.7)$$

Nótese que esta definición permite considerar a una función constante como periódica de periodo arbitrario y periodo minimal 0.

Si bien la expresión (6.3) ha sido obtenida para funciones en $\mathcal{L}^2(0, \ell)$, es evidente que si extendemos el dominio *de las series* a $(-\infty, \infty)$ obtendremos una representación de la función

$$g(x) = \begin{cases} f(x), & 0 \leq x < \ell, \\ f(x \bmod \ell), & x \notin [0, \ell), \end{cases} \quad (6.8)$$

periódica de periodo ℓ . Viceversa, toda función periódica de periodo ℓ podrá representarse por las series (6.3), aunque evidentemente no estará en \mathcal{L}^2 . Debe notarse que al extender de esta manera el dominio de $f(x)$, la extensión $g(x)$ puede presentar discontinuidades en $x = n\ell$, $n \in \mathbb{Z}$; la función $f(x) = x$ del Ejemplo 6.1.1 es un buen ejemplo. Debe notarse también que, como

$$\int_0^\ell h(x) dx = \int_a^{\ell+a} h(x) dx$$

para toda función h periódica de periodo ℓ , no solo los desarrollos (6.3) para g permanecerán inalterados al cambiar el intervalo de $(0, \ell)$ a $(a, \ell + a)$, sino que las integrales de normalización (6.5) también serán las mismas.

Un caso particularmente útil es cuando trasladamos el intervalo para hacerlo simétrico alrededor del origen, es decir, cambiamos a $(-\ell/2, \ell/2)$. Notando que las autofunciones (6.2a) de índice par son pares en x , y las de índice impar son impares, vemos que la serie de Fourier real (6.3) de una función par en $(-\ell/2, \ell/2)$ sólo contendrá autofunciones pares (la constante y cosenos), y la de una función impar sólo contendrá autofunciones impares (senos).

Ejercicio 6.1.4. Demuestre la afirmación precedente.

Ejemplo 6.1.2. La función $f(x) = x^2$, $x \in (-\ell/2, \ell/2)$ tiene un desarrollo en serie de cosenos (y quizás de la función constante); la función $f(x) = x$, $x \in (-\ell/2, \ell/2)$ tiene un desarrollo en serie de senos.

Ejercicio 6.1.5. Halle los desarrollos y demuestre las afirmaciones precedentes.

6.1.4. Convergencia puntual y uniforme

Está claro que la convergencia de las series (6.3) debe entenderse en \mathcal{L}^2 . Esto tiene la ventaja de permitirnos expresar de esa forma cualquier función localmente integrable sobre un intervalo finito, lo cual, como mencionamos en la Sección (4.1.6), en este contexto significa que la función tiene a lo sumo una cantidad finita de discontinuidades finitas en el intervalo en cuestión. Sin embargo es lícito preguntarse a qué valor convergen *puntualmente* las series (6.3) en una discontinuidad, si es que lo hacen, y bajo qué condiciones su convergencia será *uniforme* en algún intervalo.

Una función $f(x)$ se dice *continua a trozos* en $[a, b]$ si

1. $f(x)$ es continua en $[a, b]$ excepto en, a lo sumo, un conjunto finito de puntos $\{x_i\}_{i=1}^k$;
2. en cada punto de discontinuidad los límites a derecha e izquierda

$$f(x_{i-}) = \lim_{x \uparrow x_i} f(x), \quad f(x_{i+}) = \lim_{x \downarrow x_i} f(x)$$

existen y son finitos.

Se dice que f tiene una *discontinuidad de salto* en cada uno de esos puntos, con un salto igual a $f(x_{i+}) - f(x_{i-})$; en los extremos a, b del intervalo la continuidad y los límites se sobreentienden unilateralmente.

Respecto de la convergencia puntual y uniforme, enunciaremos sin demostración los siguientes resultados:

Teorema 6.1.1 (Convergencia puntual). *1. La serie de Fourier (6.3) converge a $f(x)$ puntualmente en $(0, \ell)$ si $f(x)$ es continua y $f'(x)$ es continua a trozos en $[0, \ell]$.*

2. Si $f(x)$ y $f'(x)$ son continuas a trozos en $[0, \ell]$, la serie de Fourier (6.3) converge en cada punto de $(0, \ell)$ a

$$\frac{1}{2} [f(x_+) + f(x_-)].$$

Teorema 6.1.2 (Convergencia uniforme). *La serie de Fourier (6.3) converge a $f(x)$ uniformemente en $[0, \ell]$ si*

1. $f(x)$ es continua y $f'(x)$ es continua a trozos en $[0, \ell]$;
2. $f(x)$ es periódica de periodo ℓ .

Notemos que los teoremas anteriores se extienden inmediatamente a cualquier intervalo en que convirtamos $[0, \ell]$. Notemos también que si $f(\ell) \neq f(0)$ el comportamiento de la serie será el mismo que en una discontinuidad de salto en $(0, \ell)$, como puede verse fácilmente desplazando el intervalo a uno que contenga a 0.

Ejemplo 6.1.3. La serie de Fourier de la función

$$f(x) = \begin{cases} -1, & -\frac{\ell}{2} \leq x < 0, \\ 1, & 0 \leq x < \frac{\ell}{2}, \end{cases} \quad (6.9)$$

converge a -1 en $(-\frac{\ell}{2}, 0)$, a 1 en $(0, \frac{\ell}{2})$, y a 0 en $0, -\frac{\ell}{2}$ y $\frac{\ell}{2}$.

Ejercicio 6.1.6. Calcule la serie de Fourier del ejemplo de arriba.

Que la convergencia es puntual pero no uniforme en intervalos que contengan discontinuidades de $f(x)$ queda claramente ilustrado en la Figura 6.1. Allí se muestra el desarrollo en serie de Fourier de la función (6.9), conocida como *onda cuadrada*, con la suma truncada a 5, 25 y 125 términos. Queda de manifiesto que la convergencia es tanto más “lenta” cuanto más cerca de las discontinuidades estemos.

También queda de manifiesto el llamado *fenómeno de Gibbs*. Este consiste en que el primer máximo y mínimo de la serie truncada a cada lado de las discontinuidades, lejos de desaparecer a medida que agregamos términos, tiende a una altura constante de aproximadamente un 8,9 % de la amplitud de la discontinuidad, en nuestro caso $\pm 1,089$. Sin embargo el máximo o mínimo correspondiente se acercan indefinidamente a la discontinuidad al ir agregando términos, por lo que la serie efectivamente converge a $f(x)$ para todo punto donde esta es continua.

Debe destacarse que los resultados presentados en esta sección, en particular los Teoremas 6.1.1 y 6.1.2 y el fenómeno de Gibbs, no son exclusivos del desarrollo en serie de Fourier, sino que aparecen en cualquier desarrollo en autofunciones de un PSL, como los que trataremos en las Secciones 6.2 y 6.3.

6.1.5. Expresiones para la delta de Dirac

La completitud en $\mathcal{L}^2(0, \ell)$ de las autofunciones del PSL (6.1) nos permite obtener representaciones de la delta de Dirac que son de gran utilidad en las aplicaciones.

De (5.112), usando las autofunciones complejas (6.2b) obtenemos

$$\delta(x - x') = \frac{1}{\ell} \sum_{n=-\infty}^{\infty} e^{2n\pi i(x-x')/\ell}. \quad (6.10a)$$

En particular, usando las autofunciones complejas para el intervalo $(0, 2\pi)$ o $(-\pi, \pi)$ tenemos

$$\delta(x - x') = \frac{1}{2\pi} \sum_{n=-\infty}^{\infty} e^{in(x-x')}. \quad (6.10b)$$

Sin embargo, debemos tener cuidado al interpretar estas expresiones: o bien agregamos las restricciones $x, x' \in (0, \ell)$ a la primera y $x, x' \in (0, 2\pi)$ o $x, x' \in (-\pi, \pi)$ a la segunda, o bien interpretamos el lado izquierdo como un *tren de deltas* repetidas periódicamente, con periodo ℓ en el primer caso y 2π en el segundo.

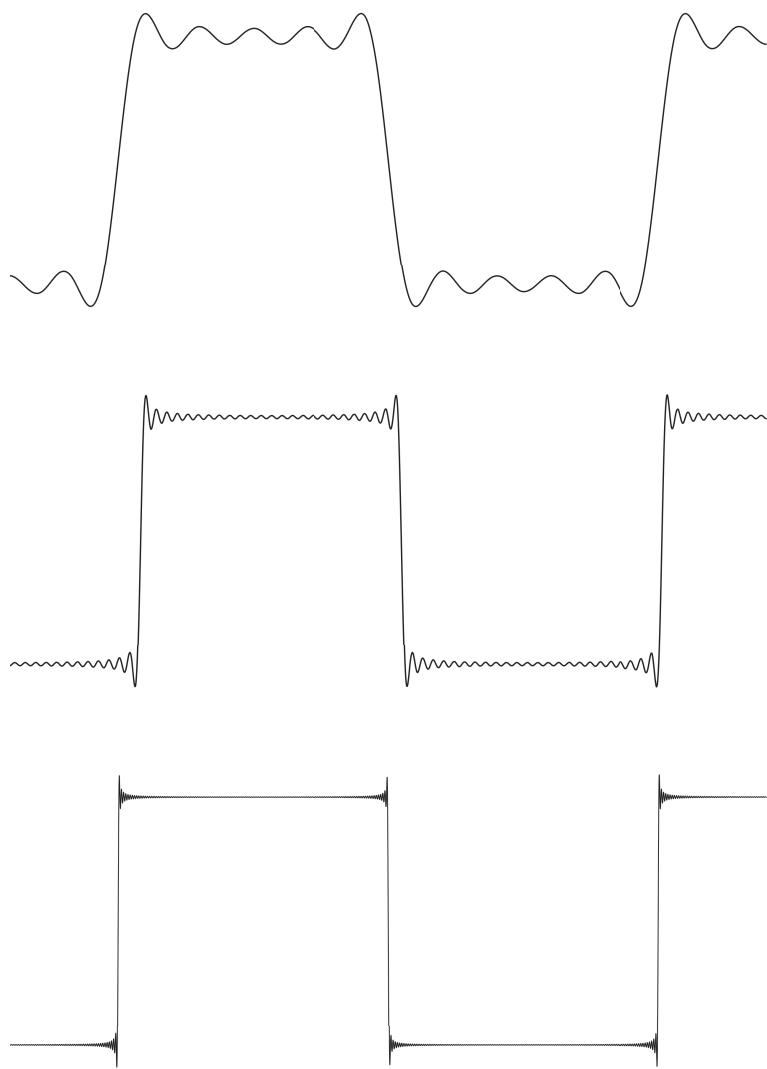


Figura 6.1: serie de Fourier de la función (6.9) truncada a 5 términos (arriba), 25 términos (centro) y 125 términos (abajo). Imágenes por Johannes Rössel, dominio público.

6.1.6. Otros PSL para la ecuación armónica

Existen, por supuesto, más PSL asociados a la ecuación armónica que (6.1). Daremos un vistazo a un par de ellos, breve dado que las técnicas serán las mismas que ya hemos aplicado, y las conclusiones, similares.

Consideremos el problema de contorno

$$y'' + \lambda y = 0, \quad 0 < x < \ell, \quad (6.11a)$$

$$y(0) = y(\ell) = 0. \quad (6.11b)$$

La ecuación y el intervalo son los mismos de (6.1a), pero ahora las CC son de Dirichlet, que siguen siendo Hermitianas para $L = \partial_x^2$. Los autovalores y autofunciones serán

$$\lambda_n = k_n^2 = \left(\frac{n\pi}{\ell}\right)^2 \longleftrightarrow y_n(x) = \sin(k_n x), \quad n \in \mathbb{N}, \quad (6.12)$$

y el desarrollo para una $f \in \mathcal{L}^2(0, \ell)$ será una *serie de Fourier de senos*

$$f(x) = \sum_{n=1}^{\infty} b_n \sin(k_n x), \quad (6.13)$$

con

$$b_n = \frac{\langle \sin(k_n x), f \rangle}{\langle \sin(k_n x), \sin(k_n x) \rangle} = \frac{2}{\ell} \int_0^{\ell} \sin\left(\frac{n\pi}{\ell} x\right) f(x) dx. \quad (6.14)$$

Análogamente, consideremos el problema de contorno

$$y'' + \lambda y = 0, \quad 0 < x < \ell, \quad (6.15a)$$

$$y'(0) = y'(\ell) = 0. \quad (6.15b)$$

La ecuación y el intervalo son de nuevo los mismos, pero con CC de Neumann, que también son Hermitianas para $L = \partial_x^2$. Los autovalores y autofunciones serán

$$\lambda_n = k_n^2 = \left(\frac{n\pi}{\ell}\right)^2 \longleftrightarrow y_n(x) = \cos(k_n x), \quad n \in \mathbb{N}_0, \quad (6.16)$$

y el desarrollo para una $f \in \mathcal{L}^2(0, \ell)$ será una *serie de Fourier de cosenos*

$$f(x) = \sum_{n=0}^{\infty} a_n \cos(k_n x), \quad (6.17)$$

con

$$a_n = \frac{\langle \cos(k_n x), f \rangle}{\langle \cos(k_n x), \cos(k_n x) \rangle} = \begin{cases} \frac{1}{\ell} \int_0^{\ell} f(x) dx, & n = 0, \\ \frac{2}{\ell} \int_0^{\ell} \cos\left(\frac{n\pi}{\ell} x\right) f(x) dx, & n > 0. \end{cases} \quad (6.18)$$

Otros PSL con CC de Robin o mixtas, pueden tratarse de manera similar.

6.2. PSL para la ecuación de Legendre

En esta sección consideraremos dos familias de funciones especiales de muy frecuente aparición en Física, los polinomios de Legendre y los polinomios asociados de Legendre, junto con los problemas de Sturm–Liouville de los que provienen y que simplificarán derivar muchas de sus propiedades. Su ubicuidad se debe a que, como veremos en el Capítulo 9, son autofunciones de la parte angular del operador ∇^2 en coordenadas esféricas.

6.2.1. El PSL en $(-1, 1)$

La ecuación diferencial de Legendre (5.20) aparece habitualmente en el contexto del PSL

$$(1 - x^2)y'' - 2xy' + \alpha(\alpha + 1)y = 0, \quad \alpha \in \mathbb{R}, \quad x \in (-1, 1), \quad (6.19a)$$

$$\lim_{x \downarrow -1} |y(x)| < \infty, \quad \lim_{x \uparrow 1} |y(x)| < \infty. \quad (6.19b)$$

El operador diferencial es formalmente autoadjunto, es decir de la forma (5.108) con $p(x) = 1 - x^2$ y $q(x) = 0$. El PSL es singular, ya que p se anula en -1 y 1 , que son puntos singulares regulares de (6.19a); las CC (6.19b) son Hermitianas para L . Los autovalores tienen la forma $\lambda = \alpha(\alpha + 1)$, y la función de peso es $\rho(x) = 1$.

En la Sección 5.1.2 vimos que la EDO (6.19a) tiene dos soluciones LI en forma de series de potencias enteras de x , dadas por (5.30), de la forma

$$\phi_1^\alpha(x) = 1 + \sum_{n=1}^{\infty} a_{2n} x^{2n}, \quad \phi_2^\alpha(x) = x + \sum_{n=1}^{\infty} a_{2n+1} x^{2n+1}, \quad (6.20)$$

con lo cual ϕ_1^α es par en x y ϕ_2^α es impar. También vimos que si, y sólo si, $\alpha = k$ con k un entero no negativo, la solución par (si k es par) o la impar (si k es impar) se trunca a un polinomio de orden k , mientras que la otra sigue siendo una serie infinita.

Supongamos que este no es el caso, de modo que ninguna de las dos soluciones se trunca a un polinomio. De la relación de recurrencia (5.27) para los coeficientes de las series vemos que, para los cocientes de términos sucesivos de la serie par,

$$q_n(x) = \left| \frac{a_{2n+2} x^{2n+2}}{a_{2n} x^{2n}} \right| = \left| \frac{(2n - \alpha)(2n + \alpha + 1)}{(2n + 1)(2n + 2)} x^2 \right|.$$

Evaluando en $x = \pm 1$, entonces,

$$q_n(\pm 1) = \left| \frac{(1 - \frac{\alpha}{2n})(1 + \frac{\alpha+1}{2n})}{(1 + \frac{1}{2n})(1 + \frac{1}{n})} \right| \underset{n \rightarrow \infty}{\asymp} 1 - \frac{1}{n} + \mathcal{O}\left(\frac{1}{n^2}\right),$$

y aplicando el criterio de Gauss, vemos que la serie par diverge en $x = \pm 1$. Notemos que para $|x| < 1$ la serie puede ser acotada por una serie geométrica en x^2 , y por lo tanto converge (absolutamente). El mismo resultado se obtiene para la serie impar.

Ejercicio 6.2.1. Verifique las afirmaciones precedentes.

De lo anterior es inmediato que ninguna solución de (6.19a) satisface las CC (6.19b) a menos que

$$\alpha = n, \quad n \in \mathbb{N}_0. \quad (6.21)$$

En ese caso la solución en serie con la paridad de n se trunca a un polinomio de grado n , y por lo tanto es analítica en toda la recta real (y en realidad en todo el plano complejo), incluso en los puntos singulares $x = \pm 1$ de (6.19a).

Aunque ya lo enunciamos en la Sección 5.1.2, por conveniencia repetiremos aquí lo siguiente: Por convención, se define el *polinomio de Legendre de orden n* como proporcional a la solución cuya serie se trunca para $\alpha = n$, de acuerdo a

$$P_n(x) := \frac{1}{2^n} \sum_{r=0}^{[n/2]} \frac{(-1)^r (2n - 2r)!}{r!(n-r)!(n-2r)!} x^{n-2r}, \quad (6.22)$$

donde $[n/2]$ es la parte entera de $n/2$ ($n/2$ si n es par, $(n-1)/2$ si n es impar). Generalmente, la *otra* solución, cuya serie no se trunca, se denota por $Q_n(x)$, con un factor de normalización también convencional, y es *singular* en $x = \pm 1$.

Los autovalores y autofunciones del PSL (6.19) son entonces

$$\lambda_n = n(n+1) \longleftrightarrow y_n(x) = P_n(x), \quad (6.23)$$

y los polinomios $P_n(x)$ forman un conjunto ortogonal y completo en $\mathcal{L}^2(-1, 1)$ con peso 1 y CC (6.19b). La normalización en (6.22) se elige convencionalmente para obtener

$$P_n(1) = 1, \quad P_n(-1) = (-1)^n. \quad (6.24)$$

Las integrales de normalización resultan entonces

$$\langle P_n, P_n \rangle = \int_{-1}^1 P_n^2(x) dx = \frac{2}{2n+1}. \quad (6.25)$$

La completitud de los polinomios de Legendre en $\mathcal{L}^2(-1, 1)$ hace que cualquier función $f \in \mathcal{L}^2(-1, 1)$ pueda desarrollarse como

$$f(x) = \sum_{n=0}^{\infty} a_n P_n(x), \quad (6.26)$$

y la normalización hace que los coeficientes a_n del desarrollo sean

$$a_n = \frac{\langle P_n, f \rangle}{\langle P_n, P_n \rangle} = \frac{2n+1}{2} \int_{-1}^1 P_n(x) f(x) dx. \quad (6.27)$$

Asimismo, de acuerdo a (5.112) tandremos la representación

$$\delta(x - x') = \sum_{n=0}^{\infty} P_n(x) \frac{2n+1}{2} P_n(x'), \quad -1 < x, x' < 1. \quad (6.28)$$

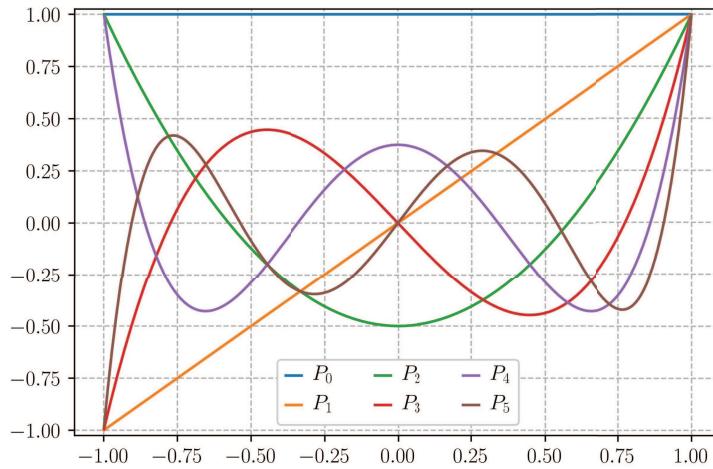


Figura 6.2: Los primeros seis polinomios de Legendre. Imagen de Geek3, bajo licencia Creative Commons Attribution-Share Alike 3.0 Unported.

Los primeros seis polinomios de Legendre se muestran en la Figura 6.2, y se escriben

$$\begin{aligned} P_0(x) &= 1, & P_3(x) &= \frac{1}{2}(5x^3 - 3x), \\ P_1(x) &= x, & P_4(x) &= \frac{1}{8}(35x^4 - 30x^2 + 3), \\ P_2(x) &= \frac{1}{2}(3x^2 - 1), & P_5(x) &= \frac{1}{8}(63x^5 - 70x^3 + 15x). \end{aligned} \quad (6.29)$$

6.2.2. Fórmula de Rodrigues para $P_n(x)$

La *fórmula de Rodrigues*, introducida por el matemático y banquero francés Olindo Rodrigues en 1816,¹ permite escribir los miembros de una familia de polinomios ortogonales como derivadas (de orden creciente) de una función dada. Más específicamente, enunciaremos sin demostración el siguiente²

Teorema 6.2.1 (Fórmula de Rodrigues). *Sean $B(x)$ y $A(x)$ polinomios en x de grados a lo sumo 2 y 1 respectivamente, y $\{f_n(x)\}_{n=0}^{\infty}$ polinomios de grado n*

¹E independientemente por el matemático británico Sir James Ivory en 1824 y por el matemático alemán Carl Gustav Jacobi en 1827.

²La formulación de este teorema es reconocidamente poco rigurosa. El lector interesado en versiones más “pulidas” (y demostraciones) puede consultar https://en.wikipedia.org/wiki/Rodrigues%27_formula https://www.physics.rutgers.edu/~shapiro/464/lects/rodrOrthPoly_2.pdf

que satisfacen

$$B(x)f_n'' + A(x)f_n' + \left[-\frac{1}{2}n(n-1)B'' - nA' \right] f_n = 0, \quad x \in (a, b).$$

Entonces

$$f_n(x) = \frac{c_n}{\rho(x)} \frac{d^n}{dx^n} [\rho(x)B^n(x)], \quad \rho(x) = \frac{1}{B(x)} \exp \left(\int^x \frac{A(x')}{B(x')} dx' \right),$$

con las constantes c_n elegidas para imponer la normalización deseada, y los polinomios son ortogonales en (a, b) con peso ρ ,

$$\langle f_n, f_m \rangle_\rho = k_n \delta_{nm}.$$

En el caso de los polinomios de Legendre, de (6.19a) tenemos $B(x) = 1 - x^2$ y $A(x) = B'(x) = -2x$, de modo que $\rho(x) = 1$ y obtenemos

$$P_n(x) = \frac{(-1)^n}{2^n n!} \frac{d^n}{dx^n} [(1 - x^2)^n]. \quad (6.30)$$

Notemos que esta expresión para P_n puede verificarse de manera directa, aunque laboriosa, por sustitución en la EDO (6.19a).

Ejercicio 6.2.2. Demuestre la afirmación precedente.

La fórmula de Rodrigues (6.30) nos permite verificar la expresión (6.25) para la integral de normalización de P_n . Tendremos

$$\|P_n\|^2 = \int_{-1}^1 P_n^2(x) dx = \frac{1}{2^{2n}(n!)^2} \int_{-1}^1 \left(\frac{d^n}{dx^n} [(x^2 - 1)^n] \right)^2 dx.$$

Integrando por partes n veces obtendremos

$$\|P_n\|^2 = \frac{(-1)^n}{2^{2n}(n!)^2} \int_{-1}^1 \left(\frac{d^{2n}}{dx^{2n}} [(x^2 - 1)^n] \right) (x^2 - 1)^n dx.$$

Notando que

$$\frac{d^{2n}}{dx^{2n}} [(x^2 - 1)^n] = (2n)!,$$

resulta

$$\|P_n\|^2 = \frac{(2n)!}{2^{2n}(n!)^2} \int_{-1}^1 (1 - x^2)^n dx = \frac{2}{2n + 1}.$$

Ejercicio 6.2.3. Complete los cálculos precedentes.

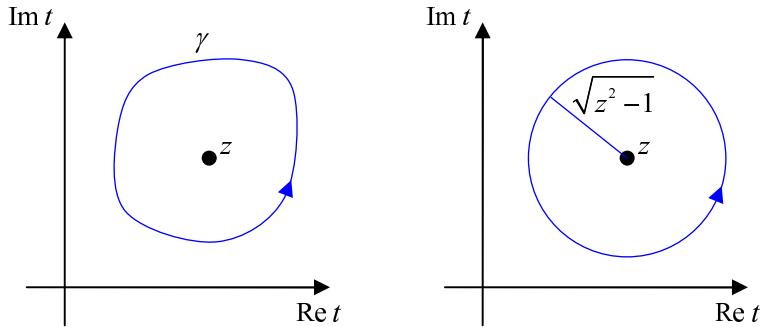


Figura 6.3: a la izquierda, circuito de integración para la representación integral de Schläfli (6.31); a la derecha, circuito para la representación integral de Laplace (6.32).

6.2.3. Representaciones integrales

A partir de la fórmula de Rodrigues (6.30) podemos obtener algunas representaciones integrales útiles para los Polinomios de Legendre.

De la fórmula de Cauchy

$$\frac{d^n}{dz^n} f(z) = \frac{n!}{2\pi i} \oint \frac{f(t) dt}{(t-z)^{n+1}},$$

donde el camino de integración γ envuelve z una vez en sentido positivo como se muestra en la Figura 6.3, y $f(z)$ es analítica en el interior de γ , la sustitución de (6.30) lleva a la *representación integral de Schläfli*

$$P_n(z) = \frac{1}{2^n} \frac{1}{2\pi i} \oint \frac{(t^2 - 1)^n}{(t-z)^{n+1}} dt. \quad (6.31)$$

Si ahora en (6.31) elegimos como contorno de integración la circunferencia centrada en z de radio $\sqrt{z^2 - 1}$ mostrada en la Figura 6.3, y la parametrizamos por

$$t = z + \sqrt{z^2 - 1} e^{i\phi}, \quad dt = i(t-z)d\phi, \quad 0 \leq \phi < 2\pi,$$

obtenemos la *representación integral de Laplace*

$$P_n(z) = \frac{1}{\pi} \int_0^\pi (z + \sqrt{z^2 - 1} \cos \phi)^n d\phi. \quad (6.32)$$

Esta última representación nos permite verificar fácilmente que de hecho la normalización en (6.30) es correcta: poniendo $z = \pm 1$ en (6.32) es inmediato que $P_n(1) = 1$ y $P_n(-1) = (-1)^n$.

6.2.4. Función generatriz

Definamos una función $F(h, z)$ como la serie

$$F(h, z) := \sum_{n=0}^{\infty} h^n P_n(z), \quad |h| < 1. \quad (6.33)$$

Usando la representación integral de Laplace (6.32) tendremos

$$\begin{aligned} F(h, z) &= \frac{1}{\pi} \int_0^\pi \sum_{n=0}^{\infty} h^n (z + \sqrt{z^2 - 1} \cos \phi)^n d\phi \\ &= \frac{1}{\pi} \int_0^\pi \frac{d\phi}{(1 - hz) - h\sqrt{z^2 - 1} \cos \phi}, \end{aligned}$$

de donde

$$F(h, z) = \frac{1}{\sqrt{1 - 2hz + h^2}}. \quad (6.34)$$

La función $F(h, z)$ se conoce como *función generatriz* de los polinomios de Legendre.

Ejercicio 6.2.4. Complete los cálculos anteriores.

Como una primera aplicación de la función generatriz, consideremos la distancia inversa entre dos vectores \vec{x} y \vec{x}' en \mathbb{R}^3 , de normas r y r' que subtienden un ángulo θ ,

$$\frac{1}{\|\vec{x} - \vec{x}'\|} = \frac{1}{\sqrt{r^2 + r'^2 - 2rr' \cos \theta}}.$$

Supongamos que $r' < r$, luego

$$\frac{1}{\|\vec{x} - \vec{x}'\|} = \frac{1}{r} \frac{1}{\sqrt{1 + \left(\frac{r'}{r}\right)^2 - 2\left(\frac{r'}{r}\right) \cos \theta}} = \frac{1}{r} F\left(\frac{r'}{r}, \cos \theta\right),$$

es decir

$$\frac{1}{\|\vec{x} - \vec{x}'\|} = \sum_{n=0}^{\infty} \frac{r'^n}{r^{n+1}} P_n(\cos \theta).$$

La expresión para $r < r'$ es la misma con r y r' intercambiados, así que para el caso general tendremos

$$\frac{1}{\|\vec{x} - \vec{x}'\|} = \sum_{n=0}^{\infty} \frac{r_-^n}{r_+^{n+1}} P_n(\cos \theta), \quad (6.35)$$

donde $r_- = \min(r, r')$ y $r_+ = \max(r, r')$.

Como una segunda aplicación, volveremos a calcular $P_n(1)$ y calcularemos $P_n(0)$. Evaluando (6.33) y (6.34) en $z = 1$ tenemos

$$F(h, 1) = \sum_{n=0}^{\infty} h^n P_n(1) = \frac{1}{1-h},$$

de donde es inmediato que $P_n(1) = 1$. Haciendo lo mismo en $z = 0$ tenemos

$$F(h, 0) = \sum_{n=0}^{\infty} h^n P_n(0) = \frac{1}{\sqrt{1+h^2}},$$

que desarrollando el último término en potencias de h lleva a

$$P_n(0) = \begin{cases} 0, & n \text{ impar}, \\ \frac{(n-1)!!(-1)^{n/2}}{2^{n/2} (\frac{n}{2})!}, & n \text{ par}. \end{cases} \quad (6.36)$$

Ejercicio 6.2.5. Complete los cálculos anteriores.

6.2.5. Relaciones de recurrencia

La función generatriz permite derivar fácilmente relaciones entre polinomios de Legendre de diferentes órdenes y/o sus derivadas, conocidas como *relaciones de recurrencia*. Derivando (6.34) respecto de h tenemos

$$\frac{\partial F}{\partial h} = \frac{z-h}{(1-2hz+h^2)^{3/2}},$$

de donde

$$(1-2hz+h^2) \frac{\partial F}{\partial h} = (z-h)F.$$

Substituyendo (6.33) e igualando potencias de h , entonces,

$$(n+1)P_{n+1}(z) - (2n+1)zP_n(z) + nP_{n-1}(z) = 0, \quad n \geq 1. \quad (6.37a)$$

Esta relación de recurrencia nos permite, por ejemplo, construir todos los polinomios de Legendre a partir de $P_0(x) = 1$ y $P_1(x) = x$.

Si en cambio derivamos (6.34) respecto de z , tenemos

$$(1-2hz+h^2) \frac{\partial F}{\partial z} = hF,$$

y substituyendo (6.33) e igualando potencias de h resulta

$$P'_{n+1}(z) - 2zP'_n(z) + P'_{n-1}(z) = P_n(z), \quad n \geq 1. \quad (6.37b)$$

A partir de (6.37a) y (6.37b), o combinando F y sus derivadas parciales, podemos obtener una cantidad de relaciones de recurrencia adicionales, como ser

$$P'_{n+1}(z) = (n+1)P_n(z) + zP'_n(z), \quad n \geq 0, \quad (6.37c)$$

$$(2n+1)P_n(z) = P'_{n+1}(z) - P'_{n-1}(z), \quad n \geq 1, \quad (6.37d)$$

$$nP_n(z) = zP'_n(z) - P'_{n-1}(z), \quad n \geq 1, \quad (6.37e)$$

$$(1-z^2)P'_n(z) = nP_{n-1}(z) - nzP_n(z), \quad n \geq 1, \quad (6.37f)$$

etc.

Ejercicio 6.2.6. Demuestre las relaciones de recurrencia (6.37a) a (6.37f).

6.2.6. Funciones de Legendre de segundo tipo

La segunda solución $Q_n(x)$ de la EDO (6.19a), que junto con $P_n(x)$ forma una base del espacio de soluciones para $\lambda = n$ pero no satisface las CC (6.19b), se conoce como *función de Legendre de segundo tipo*. Puede obtenerse de las soluciones en series de potencias generalizadas (5.30), y se expresa convencionalmente como³

$$Q_n(x) = \frac{1}{2}P_n(x) \ln\left(\frac{1+x}{1-x}\right) - \sum_{k=1}^n \frac{1}{k} P_{k-1}(x) P_{n-k}(x). \quad (6.38)$$

Una expresión alternativa se obtiene proponiendo como segunda solución de (6.19a)

$$y(x) = v(x)P_n(x),$$

donde $v(x)$ es una nueva función incógnita. Substituyendo en (6.19a) con $\alpha = n$, usando que $P_n(x)$ es solución y colectando derivadas de v , tenemos

$$(1-x^2)P_n v'' + 2[(1-x^2)P'_n - xP_n]v' = 0,$$

que reordenamos como

$$\frac{v''}{v'} = -2\frac{P'_n}{P_n} + \frac{2x}{1-x^2}.$$

Notando que esto equivale a

$$(\ln v')' = -2(\ln P_n)' - (\ln(1-x^2))',$$

dos integraciones elementales nos llevan a

$$v(x) = c_1 \int \frac{dx}{(1-x^2)P_n^2(x)} + c_2,$$

donde c_1 y c_2 son dos constantes de integración arbitrarias. La elección convencional define

$$Q_n(x) = P_n(x) \int \frac{dx}{(1-x^2)P_n^2(x)}. \quad (6.39)$$

Cualquiera de las definiciones anteriores nos da las mismas expresiones para Q_0 y Q_1 ; a partir de ellas, y usando la relación de recurrencia (6.37a), que es válida también para Q_n , obtenemos la recurrencia

$$Q_n(x) = \begin{cases} \frac{1}{2} \ln\left(\frac{1+x}{1-x}\right), & n = 0, \\ P_1(x)Q_0(x) - 1, & n = 1, \\ \frac{2n-1}{n}xQ_{n-1}(x) - \frac{n-1}{n}Q_{n-2}(x), & n \geq 2. \end{cases} \quad (6.40)$$

³Existen muchas definiciones alternativas de Q_n . La mostrada aquí está tomada de <https://dlmf.nist.gov/14.7>, que es una referencia standard.

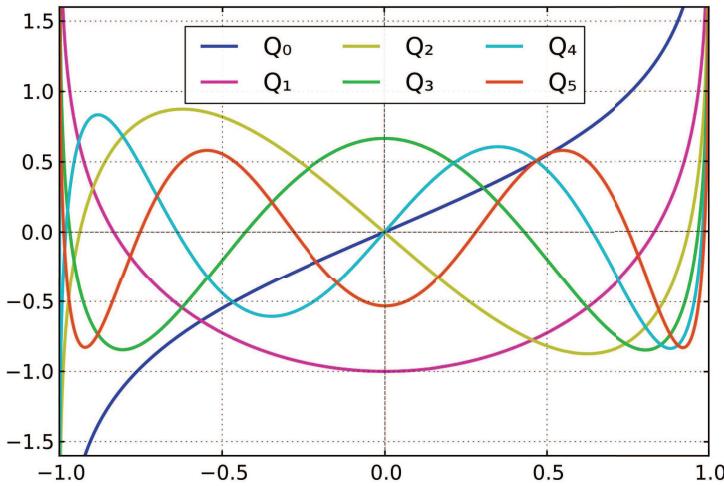


Figura 6.4: Las primeras seis funciones de Legendre de segundo tipo. Imagen de Geek3, bajo licencia GNU Free Documentation License.

Las primeras funciones de Legendre de segundo tipo se escriben

$$\begin{aligned} Q_0(x) &= \frac{1}{2} \ln \left(\frac{1+x}{1-x} \right), & Q_2(x) &= \frac{3x^2-1}{4} \ln \left(\frac{1+x}{1-x} \right) - \frac{3x^2}{2}, \\ Q_1(x) &= \frac{x}{2} \ln \left(\frac{1+x}{1-x} \right) - 1, & Q_3(x) &= \frac{5x^3-3x}{4} \ln \left(\frac{1+x}{1-x} \right) - \frac{5x^2}{2} + \frac{2}{3}, \end{aligned} \quad (6.41)$$

y se muestran en la Figura 6.4.

6.2.7. PSL sobre otros intervalos

Las funciones de Legendre de segundo tipo cobran relevancia cuando necesitamos resolver un PSL con la EDO (6.19a) pero sobre un intervalo $(a, b) \subset (-1, 1)$, es decir que no incluya los puntos singulares, por ejemplo

$$(1-x^2)y'' - 2xy' + \alpha(\alpha+1)y = 0, \quad -1 < a < x < b < 1, \quad (6.42a)$$

$$y(a) = y(b) = 0, \quad (6.42b)$$

con CC de Dirichlet. En este caso el PSL es *regular*, y α no tiene porqué ser un entero.

La solución general de la ODE (6.42a) es ahora

$$y(x) = c_1 P_\alpha(x) + c_2 Q_\alpha(x), \quad (6.43)$$

donde P_α y Q_α son dos soluciones LI obtenidas de las series (5.30), y en general ninguna se trunca a un polinomio (y por lo tanto ambas serán singulares en $x = \pm 1$, lo cual es irrelevante ya que nuestro intervalo no incluye esos puntos).

En un caso como este los autovalores *también* quedan determinados por las CC

$$\begin{aligned} c_1 P_\alpha(a) + c_2 Q_\alpha(a) &= 0, \\ c_1 P_\alpha(b) + c_2 Q_\alpha(b) &= 0. \end{aligned}$$

La primera ecuación, por sí misma, admite infinitas soluciones (c_1, c_2) proporcionales entre sí; seleccionando $c_1 = cQ_\alpha(a)$ y $c_2 = -cP_\alpha(a)$ con c una nueva constante, se satisface idénticamente, y la segunda ecuación queda

$$c(Q_\alpha(a)P_\alpha(b) - P_\alpha(a)Q_\alpha(b)) = 0.$$

Ésta tendrá solución no trivial $c \neq 0$ si, y sólo si,

$$Q_\alpha(a)P_\alpha(b) - P_\alpha(a)Q_\alpha(b) = 0,$$

que es la *ecuación de autovalores* que determina α (numéricamente, por lo general).

Una vez determinados así los posibles α_n , y con ellos los autovalores

$$\lambda_n = \alpha_n(\alpha_n + 1) \quad (6.44)$$

del PSL (6.42), las autofunciones serán

$$y_n(x) = Q_{\alpha_n}(a)P_{\alpha_n}(x) - P_{\alpha_n}(a)Q_{\alpha_n}(x), \quad (6.45)$$

que por construcción satisfacen *ambas* CC (6.42b), y formarán un conjunto ortogonal y completo en $L^2(a, b)$ (con peso 1). Debe notarse que las relaciones de recurrencia (6.37), y en general cualquier otro resultado que pueda deducirse de la propiedad de y_n de ser soluciones de la EDO (6.42a), mantendrán su validez (cambiando n por α_n).

El mismo procedimiento aquí delineado puede aplicarse, *mutatis mutandis*, a problemas donde las CC sean de Neumann, de Robin o mixtas.

6.2.8. Las funciones asociadas de Legendre

En muchos problemas provenientes de la Física, en particular al resolver ecuaciones diferenciales parciales que involucren el operador ∇^2 por separación de variables en coordenadas esféricas, nos encontramos con las llamadas *funciones asociadas de Legendre*. Estas son soluciones de la EDO

$$(1 - x^2)y'' - 2xy' + \left[\nu(\nu + 1) - \frac{\mu^2}{1 - x^2} \right] y = 0. \quad (6.46)$$

Notemos que el operador diferencial es formalmente autoadjunto, es decir de la forma (5.108), con $p(x) = 1 - x^2$ y $q(x) = \mu^2/(1 - x^2)$, lo que introduce una

singularidad adicional; sin embargo los puntos singulares $x = \pm 1$ siguen siendo *regulares* de acuerdo a la definición de la Sección 5.1.3. Las dos soluciones LI standard se conocen como *funciones asociadas de Legendre* de primer y segundo tipo, y se denotan respectivamente por $P_\nu^\mu(x)$ y $Q_\nu^\mu(x)$, donde ν es llamado el *grado* y μ el *orden*. Pueden obtenerse como desarrollos en series de potencias alrededor del punto regular $x = 0$ en forma completamente análoga a como se obtienen P_α y Q_α y, en general, para ν y μ arbitrarios ambas resultan en series infinitas y son singulares en los puntos singulares $x = \pm 1$. Sólo para $\nu = n$, un entero ≥ 0 , y $|\mu| \leq n$, las soluciones P_n^μ son finitas en $x = \pm 1$.

El caso de más común aparición es el PSL

$$(1-x^2)y'' - 2xy' + \left[\nu(\nu+1) - \frac{\mu^2}{1-x^2} \right] y = 0, \quad x \in (-1, 1), \quad (6.47a)$$

$$\lim_{x \downarrow -1} |y(x)| < \infty, \quad \lim_{x \uparrow 1} |y(x)| < \infty, \quad (6.47b)$$

con CC Hermitianas para L ; los autovalores tienen la forma $\lambda = \nu(\nu+1)$, y la función de peso es $\rho(x) = 1$. Las CC de función finita imponen, en ese caso, que $\nu = n \in \mathbb{N}_0$, y que las autofunciones se restrinjan a $P_n^\mu(x)$ con $|\mu| \leq n$. Frecuentemente $\mu = m$, otro entero, y en ese caso las CC de función finita imponen $-n \leq m \leq n$. En ese caso tendremos

$$\lambda_n = n(n+1) \longleftrightarrow y_n(x) = P_n^m(x), \quad n = m, m+1, m+2, \dots, \quad (6.48)$$

y las autofunciones P_n^m para un m dado serán ortogonales y completas en $L^2(-1, 1)$. Las autofunciones P_n^m suelen llamarse polinomios asociados de Legendre aunque, como veremos enseguida, sólo son polinomios para m par.

Debemos notar que para $m = 0$ la EDO (6.47a) se reduce a la ecuación de Legendre (6.19a), por lo que tendremos $P_n^0(x) = P_n(x)$, los polinomios de Legendre; los primeros P_n^0 son entonces los mostrados en la Figura 6.2, y sus expresiones son las dadas en (6.29). Para $m = 1$ los primeros P_n^1 se muestran en la Figura 6.5, y se escriben

$$\begin{aligned} P_1^1(x) &= (1-x^2)^{1/2}, & P_3^1(x) &= \frac{3}{2}(5x^2-1)(1-x^2)^{1/2}, \\ P_2^1(x) &= 3x(1-x^2)^{1/2} & P_4^1(x) &= \frac{5}{2}(7x^3-3x)(1-x^2)^{1/2}. \end{aligned} \quad (6.49)$$

Para $m = 2$ los primeros P_n^2 se muestran en la Figura 6.6, y se escriben

$$\begin{aligned} P_2^2(x) &= 3(1-x^2), & P_4^2(x) &= \frac{15}{2}(7x^2-1)(1-x^2), \\ P_3^2(x) &= 15x(1-x^2), & P_5^2(x) &= \frac{105}{2}(3x^3-x)(1-x^2). \end{aligned} \quad (6.50)$$

Para $m \geq 0$ tenemos⁴

$$P_n^m(x) = (1-x^2)^{m/2} \frac{d^m P_n(x)}{dx^m}, \quad (6.51)$$

⁴Algunos autores introducen un factor $(-1)^m$ en el lado derecho.

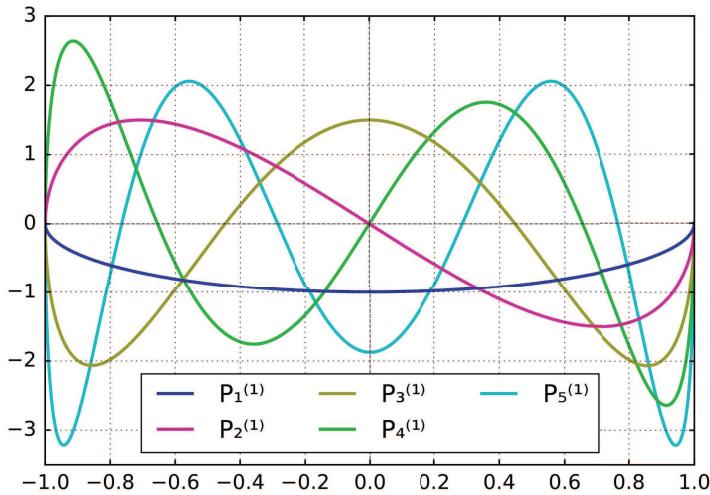


Figura 6.5: Las primeras cinco funciones asociadas de Legendre P_n^1 . Imagen de Geek3, bajo licencia Creative Commons Attribution-Share Alike 4.0 International.

que puede verificarse derivando m veces la EDO (6.19a) para los polinomios de Legendre. De aquí es inmediato que $P_n^m(x) = 0$ si $m > n$, y que en general P_n^m será el producto de $(1 - x^2)^{m/2}$ por un polinomio de grado $n - m$, lo que se aprecia cualitativamente en las Figuras 6.5 y 6.6.

La completitud de $\{P_n^m\}_{n=m}^\infty$ en $\mathcal{L}^2(-1, 1)$ nos permite escribir

$$f(x) = \sum_{n=m}^{\infty} a_n P_n^m(x) \quad (6.52)$$

para toda $f \in \mathcal{L}^2(-1, 1)$, y por ortogonalidad

$$a_n = \frac{\langle P_n^m, f \rangle}{\langle P_n^m, P_n^m \rangle} \quad (6.53)$$

con

$$\langle P_n^m, P_n^m \rangle = \int_{-1}^1 (P_n^m(x))^2 dx = \frac{2(n+m)!}{(2n+1)(n-m)!}. \quad (6.54)$$

6.2.9. Armónicos esféricos

De acuerdo a lo visto en la Sección 6.1, las funciones armónicas $\{e^{im\phi}\}_{m=-\infty}^\infty$, soluciones de la ecuación armónica

$$\frac{d^2u}{d\phi^2} + m^2 u = 0 \quad (6.55)$$

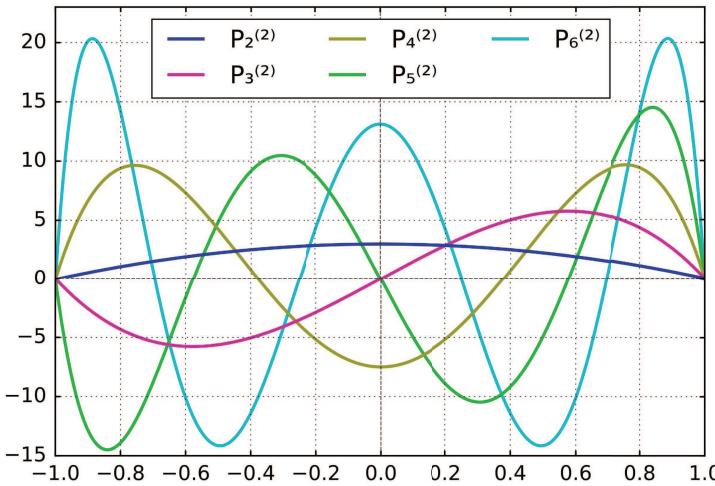


Figura 6.6: Las primeras cinco funciones asociadas de Legendre P_n^2 . Imagen de Geek3, bajo licencia Creative Commons Attribution-Share Alike 4.0 International.

con CC “naturales” (es decir, periódicas) en $[0, 2\pi]$, pueden pensarse como un conjunto ortogonal y completo para funciones definidas sobre la *1-esfera* S^1 (la circunferencia unidad), formando una base para desarrollar funciones en $\mathcal{L}^2[S^1]$. Adelantándonos a los problemas que trataremos en el Capítulo 9, presentaremos su extensión “natural” a una base de soluciones de la ecuación armónica

$$\nabla^2 y + \lambda y = 0 \quad (6.56)$$

con CC naturales sobre la *2-esfera* S^2 (la esfera unidad), que pensaremos parametrizada por los ángulos esféricos usuales $\theta \in [0, \pi]$ y $\phi \in [0, 2\pi]$.

Si proponemos

$$y(\theta, \phi) = \sum_{m=-\infty}^{\infty} y_m(\theta) e^{im\phi}$$

y usamos

$$\nabla^2 = \frac{\partial^2}{\partial \theta^2} + \cot \theta \frac{\partial}{\partial \theta} + \csc^2 \theta \frac{\partial^2}{\partial \phi^2}$$

y la ortogonalidad de $e^{im\phi}$, obtenemos

$$\frac{d^2 y_m}{d\theta^2} + \cot \theta \frac{dy_m}{d\theta} + \left[\lambda - \frac{m^2}{\sin^2 \theta} \right] y_m = 0,$$

que con el cambio de variables

$$x = \cos \theta$$

se reduce a

$$(1-x^2)y_m'' - 2xy_m' + \left[\lambda - \frac{m^2}{1-x^2} \right] y_m = 0, \quad x \in (-1, 1). \quad (6.57a)$$

Como aplicamos CC naturales, es decir de función finita en $x = \pm 1$ (*i.e.* en $\theta = 0$ y $\theta = \pi$), y (6.57a) es invariante ante un cambio de signo en m , los autovalores y autofunciones serán

$$\lambda_n = n(n+1) \longleftrightarrow y_{mn}(x) = P_n^{|m|}(x), \quad n = |m|, |m|+1, |m|+2, \dots, \quad (6.57b)$$

y, recordando que $P_n^{|m|}(x) = 0$ si $|m| > n$ y reordenando la suma, la solución general de (6.56) se escribirá

$$y(\theta, \phi) = \sum_{n=0}^{\infty} \sum_{m=-n}^n c_{nm} P_n^{|m|}(\cos \theta) e^{im\phi}, \quad (6.58)$$

con c_{nm} constantes (en general complejas).

Usando ahora la ortogonalidad de $\{e^{im\phi}\}$ en $(0, 2\pi)$ y de $P_n^m(x)$ en $(-1, 1)$, y reescribiendo (6.54) como

$$\langle P_n^m, P_n^m \rangle = \int_0^\pi (P_n^{|m|}(\cos \theta))^2 \sin \theta d\theta = \frac{2(n+|m|)!}{(2n+1)(n-|m|)!}, \quad (6.59)$$

definimos los *armónicos esféricos* $Y_{nm}(\theta, \phi)$ como⁵

$$Y_{nm}(\theta, \phi) = \sqrt{\frac{2n+1}{4\pi} \frac{(n-|m|)!}{(n+|m|)!}} P_n^{|m|}(\cos \theta) e^{im\phi}. \quad (6.60)$$

Estas son las *autofunciones normalizadas* del operador ∇^2 en S^2 , que satisfacen

$$\nabla^2 Y_{nm} + n(n+1)Y_{nm} = 0 \quad (6.61)$$

y son ortonormales y completas en $\mathcal{L}^2[S^2]$, con

$$\langle Y_{nm}, Y_{n'm'} \rangle = \int_0^\pi \sin \theta d\theta \int_0^{2\pi} d\phi Y_{nm}^*(\theta, \phi) Y_{n'm'}(\theta, \phi) = \delta_{nn'} \delta_{mm'} \quad (6.62)$$

⁵Es tradicional en Física extender la definición de P_n^m a $m < 0$ usando (6.47b) junto con la fórmula de Rodrigues (6.30), como

$$P_n^{-m}(x) = \frac{1}{2^n n!} (1-x^2)^{m/2} \frac{d^{n+m}}{dx^{n+m}} (x^2 - 1)^n,$$

con lo cual se tiene

$$P_n^{-m}(x) = \frac{(n-m)!}{(n+m)!} P_n^m(x), \quad m \geq 0$$

y la definición de los armónicos esféricos queda

$$Y_{nm}(\theta, \phi) = \sqrt{\frac{2n+1}{4\pi} \frac{(n-m)!}{(n+m)!}} P_n^m(\cos \theta) e^{im\phi},$$

que por supuesto es completamente equivalente a (6.60). Simplemente se debe cuidar la consistencia al decidir usar una u otra.

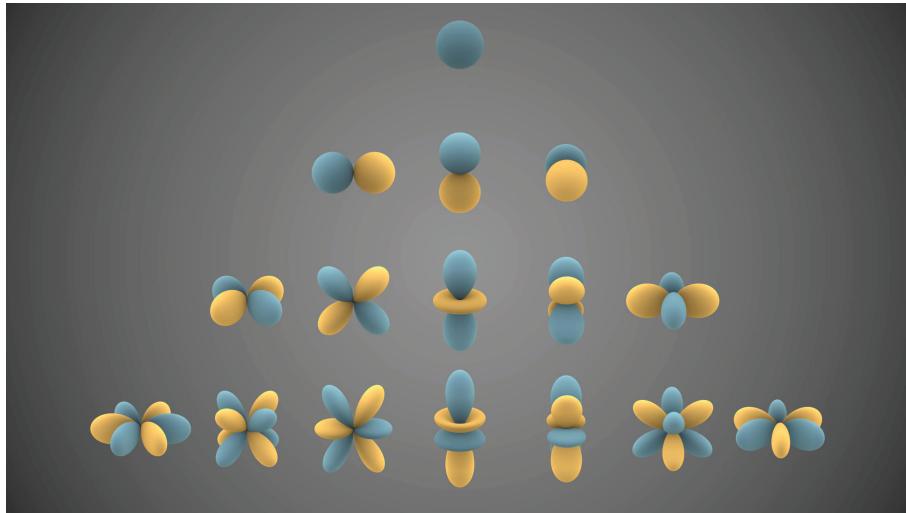


Figura 6.7: Los primeros armónicos esféricos $Y_{nm}(\theta, \phi)$. La columna central contiene los Y_{n0} con n creciente hacia abajo desde 0; a la izquierda, $\operatorname{Re} Y_{nm}$ con m creciente hacia la izquierda; a la derecha, $-\operatorname{Im} Y_{nm}$ con m creciente hacia la derecha. La distancia de cada superficie al origen representa la magnitud de la función, en azul para valores positivos y en amarillo para valores negativos. El eje x apunta hacia la izquierda, el eje y hacia el lector, inclinado un poco hacia abajo, y el eje z hacia arriba, un poco inclinado hacia el lector. Imagen de Inigo.quilez, bajo licencia Creative Commons Attribution-Share Alike 3.0 Unported.

y

$$\delta(\cos \theta - \cos \theta')\delta(\phi - \phi') = \sum_{n=0}^{\infty} \sum_{m=-n}^n Y_{nm}^*(\theta, \phi) Y_{nm}(\theta', \phi'). \quad (6.63)$$

Los armónicos esféricos son así la extensión natural a $\mathcal{L}^2[S^2]$ de las funciones armónicas $\{e^{im\phi}\}$, que son ortogonales y completas en $\mathcal{L}^2[S^1]$.

6.3. PSL para la ecuación de Bessel

Los problemas de Sturm–Liouville que involucran la ecuación de Bessel (5.69) presentan usualmente una diferencia importante con los tratados en la Sección 6.2. Estos PSL surgen cuando aplicamos el procedimiento de separación de variables que estudiaremos en el Capítulo 9 a ciertas EDP que involucran el operador ∇^2 en coordenadas cilíndricas y esféricas, y en ellos el parámetro α de (5.69), que es el *índice* de las autofunciones, está usualmente fijo de antemano. Este parámetro, entonces, no tendrá el papel de *autovalor* del PSL, y

la ortogonalidad de las autofunciones no se definirá entre autofunciones de diferentes índices como en la Sección 6.2, sino respecto de otro parámetro, como veremos enseguida; en este aspecto el problema es similar al de la Sección 6.1.

La ecuación a la que arribaremos usualmente tiene la forma

$$x^2y'' + xy + (k^2x^2 - \nu^2)y = 0 \quad (6.64)$$

con ν fijo y $\lambda = k^2$ a determinar, y puede obtenerse de (5.69) mediante las sustituciones

$$y(x) \longrightarrow y(kx), \quad x \longrightarrow kx. \quad (6.65)$$

Ejercicio 6.3.1. Verifique la afirmación precedente.

Ayuda: cambie x por λx en (5.69) e introduzca $z(x) = y(kx)$.

6.3.1. El PSL en $(0, b)$

La ecuación diferencial de Bessel (6.64) aparece habitualmente en el contexto de problemas de Sturm–Liouville, una de cuyas versiones más corrientes es

$$xy'' + y' + \left(k^2x - \frac{\nu^2}{x} \right) y = 0, \quad \nu \in \mathbb{R}, \quad x \in (0, b), \quad (6.66a)$$

$$\lim_{x \downarrow 0} |y(x)| < \infty, \quad y(b) = 0. \quad (6.66b)$$

El operador diferencial es formalmente autoadjunto, es decir de la forma (5.108) con $p(x) = x$ y $q(x) = \nu^2/x$. El PSL es singular, ya que p se anula en 0, que es un punto singular regular de (6.66a); las CC (6.66b) son Hermitianas para L . Los autovalores tienen la forma $\lambda = k^2$, y la función de peso es $\rho(x) = x$.⁶

En la Sección 5.1.4 vimos que la EDO (6.66a) tiene dos soluciones en forma de series de potencias generalizadas de x dadas por (5.76) y (5.77), que mediante las sustituciones (6.65) aquí se expresarán como

$$J_\nu(kx) = \left(\frac{kx}{2} \right)^\nu \sum_{n=0}^{\infty} \frac{(-1)^n}{n! \Gamma(n+1+\nu)} \left(\frac{kx}{2} \right)^{2n}, \quad (6.67a)$$

$$J_{-\nu}(kx) = \left(\frac{kx}{2} \right)^{-\nu} \sum_{n=0}^{\infty} \frac{(-1)^n}{n! \Gamma(n+1-\nu)} \left(\frac{kx}{2} \right)^{2n}, \quad (6.67b)$$

con J_ν regular en $x = 0$, y $J_{-\nu}$ singular, si $\nu \geq 0$.⁷ También vimos que si $\nu \neq 0, 1, 2, \dots$ estas dos soluciones son LI, pero son LD si no. En este último caso una segunda solución LI de J_ν podía obtenerse como (5.77') o (5.77''). Ahora bien, mas allá de la complejidad de estas últimas expresiones, el sólo hecho de tener que cambiar la definición de la segunda solución LI cada vez que ν pase

⁶Notemos que de acuerdo a (5.110), la función de peso $\rho(x)$ es el coeficiente del autovalor λ (k^2 en este caso) una vez reducido el operador a uno formalmente autoadjunto.

⁷Nótese que para $\nu > 0$ no entero, $J_{-\nu}(kz)$ tendrá un corte ramal que parte del origen del plano complejo z ; es habitual considerar este corte ramal como el eje real negativo $\text{Re } z < 0$.

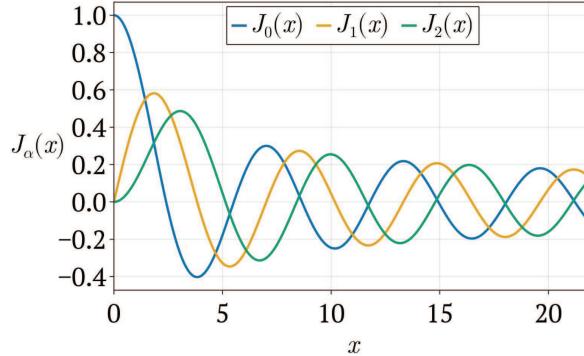


Figura 6.8: Las primeras tres funciones de Bessel de primer tipo $J_n(x)$. Imagen de Aspir8, bajo licencia Creative Commons CC0 1.0 Universal Public Domain Dedication.

por un valor entero, es incómodo en la práctica. Por ello es costumbre definir una segunda solución Y_ν por

$$Y_\nu(kx) = \frac{\cos(\nu\pi)J_\nu(kx) - J_{-\nu}(kx)}{\sin(\nu\pi)}, \quad (6.67c)$$

que será LI de J_ν para todo ν : para $\nu \notin \mathbb{N}_0$ es combinación lineal de J_ν y $J_{-\nu}$, con $J_{-\nu}$ LI de J_ν ; para $\nu = 0, 1, 2, \dots$ debe evaluarse resolviendo el límite indeterminado $\frac{0}{0}$ por la regla de L'Hôpital derivando respecto de ν , que es el mismo procedimiento implementado en la Sección 5.1.3 para tratar con estos casos, derivando respecto de las raíces del polinomio indicial.⁸ Las soluciones J_ν se conocen como *funciones de Bessel de primer tipo de orden ν* , y serán finitas en $x = 0$ para $\nu \geq 0$; las soluciones Y_ν se conocen como *funciones de Bessel de segundo tipo*, y en general divergen en $x = 0$. Las funciones de Bessel de primer tipo para $\nu = 0, 1, 2$ se muestran en la Figura 6.8; las correspondientes de segundo tipo se muestran en la Figura 6.9. Ambas figuras ilustran claramente lo deducido acerca de las soluciones de (6.66a) en el Ejemplo 5.2.1: todas las soluciones son *oscilatorias*.

⁸Para interpretar mejor lo que estamos haciendo aquí, consideremos la base de soluciones $\{\cos(\omega x), \sin(\omega x)\}$ de la ecuación $y'' + \omega^2 y = 0$, cuya solución general escribimos como $a \cos(\omega x) + b \sin(\omega x)$. Para $\omega \rightarrow 0$ esta debería reducirse a $a + bt$, la solución general de $y'' = 0$, pero no lo hace: se reduce a la constante a . Lo que ocurre es que para $\omega \rightarrow 0$ la base “colapsa”, como queda de manifiesto si calculamos el determinante Wronskiano de $\cos(\omega x)$ y $\sin(\omega x)$, que es igual a ω . Para que esto no ocurra debemos elegir una base de soluciones que sean LI aún cuando $\omega \rightarrow 0$, por ejemplo $\{\cos(\omega x), \sin(\omega x)/\omega\}$, cuyo determinante Wronskiano es 1 para todo ω , y donde $\sin(\omega x)/\omega \rightarrow x$ cuando $\omega \rightarrow 0$. Aquí hacemos algo similar, ya que

$$W[J_\nu(z), J_{-\nu}(z)] = -\frac{2 \sin(\nu\pi)}{\pi z}, \quad W[J_\nu(z), Y_\nu(z)] = \frac{2}{\pi z}.$$

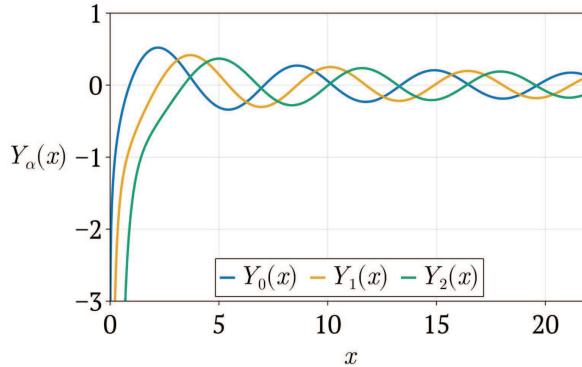


Figura 6.9: Las primeras tres funciones de Bessel de segundo tipo $Y_n(x)$. Imagen de Aspir8, bajo licencia Creative Commons CC0 1.0 Universal Public Domain Dedication.

Es inmediato que cualquier combinación lineal de J_ν e Y_ν será también solución de la EDO (6.66a) y, como ésta es a coeficientes reales, cualquier combinación lineal compleja de soluciones reales también será solución. Es habitual definir las *funciones de Hankel* (o funciones de Bessel de tercer tipo) como

$$H_\nu^{(1)}(kz) := J_\nu(kz) + iY_\nu(kz), \quad H_\nu^{(2)}(kz) := J_\nu(kz) - iY_\nu(kz), \quad (6.67d)$$

que son dos soluciones LI complejas (mutuamente conjugadas) de (6.66a).

Queda todavía la tarea de imponer las CC (6.66b). De lo anterior es claro que la CC de función finita en $x = 0$ se satisface si, y sólo si, descartamos las funciones de segundo tipo Y_ν (y las funciones de Hankel). Debemos entonces tener, por la CC en $x = b$,

$$J_\nu(kb) = 0. \quad (6.68)$$

Sabemos que J_ν tendrá infinitos ceros simples sucesivos (por ser oscilatoria), digamos

$$J_\nu(\kappa_{\nu n}) = 0, \quad n = 1, 2, 3, \dots \quad (6.69)$$

Luego cada vez que $kb = \kappa_{\nu n}$ tendremos una autofunción de (6.66a), con lo que los autovalores y autofunciones quedan⁹

$$\lambda_n = k_n^2 = \left(\frac{\kappa_{\nu n}}{b} \right)^2 \longleftrightarrow y_n(x) = J_\nu(k_n x) = J_\nu \left(\frac{\kappa_{\nu n}}{b} x \right) \quad (6.70)$$

Estas autofunciones forman un conjunto ortogonal y completo en $\mathcal{L}_\rho^2(0, b)$, con peso $\rho(x) = x$. Las integrales de normalización son

$$\left\langle J_\nu \left(\frac{\kappa_{\nu n}}{b} x \right), J_\nu \left(\frac{\kappa_{\nu n}}{b} x \right) \right\rangle = \int_0^b J_\nu^2 \left(\frac{\kappa_{\nu n}}{b} x \right) x dx = \frac{b^2}{2} [J'_\nu(\kappa_{\nu n})]^2 \quad (6.71)$$

⁹Nótese la similitud con el procedimiento seguido para obtener (6.2a), con $\kappa_{\nu n}/b$ jugando el papel de $k_n = 2n\pi/\ell$ y $J_\nu(\kappa_{\nu n}x/b)$ en lugar de $\cos(k_n x)$ o $\sin(k_n x)$.

(recordemos que J_ν y J'_ν no tienen ceros comunes por el Teorema 5.2.2). Toda f en $\mathcal{L}^2(0, b)$ puede escribirse como

$$f(x) = \sum_{n=1}^{\infty} a_n J_\nu \left(\frac{\kappa_{\nu n}}{b} x \right) \quad (6.72)$$

con

$$a_n = \frac{2}{b^2 [J'_\nu(\kappa_{\nu n})]^2} \int_0^b J_\nu \left(\frac{\kappa_{\nu n}}{b} x \right) f(x) x dx, \quad (6.73)$$

y de (5.112) tenemos

$$\delta(x - x') = \sum_{n=1}^{\infty} J_\nu \left(\frac{\kappa_{\nu n}}{b} x \right) \frac{2x'}{b^2 [J'_\nu(\kappa_{\nu n})]^2} J_\nu \left(\frac{\kappa_{\nu n}}{b} x' \right), \quad x, x' \in (0, b). \quad (6.74)$$

6.3.2. Relaciones de recurrencia

Las relaciones de recurrencia básicas para J_ν y $J_{-\nu}$ pueden obtenerse directamente de sus desarrollos en serie de potencias generalizadas de x . Sin embargo antes de proceder es conveniente notar que, por ser Y_ν una combinación lineal de J_ν y $J_{-\nu}$, también valdrán para ella,¹⁰ y por consiguiente para $H_\nu^{(1)}$ y $H_\nu^{(2)}$. Es conveniente también notar aquí que las series para J_ν y $J_{-\nu}$ pueden escribirse como una y la misma serie,

$$J_\nu(kx) = \sum_{n=0}^{\infty} \frac{(-1)^n}{n! \Gamma(n+1+\nu)} \left(\frac{kx}{2} \right)^{2n+\nu},$$

donde ahora ν puede ser cualquier real, positivo o negativo.¹¹

Denotando por C_ν cualquiera de J_ν , Y_ν , $H_\nu^{(1)}$ o $H_\nu^{(2)}$, entonces, es directo mostrar a partir de la serie para J_ν que

$$C_{\nu-1}(z) + C_{\nu+1}(z) = \frac{2\nu}{z} C_\nu(z), \quad (6.75a)$$

$$C_{\nu-1}(z) - C_{\nu+1}(z) = 2C'_\nu(z). \quad (6.75b)$$

Ejercicio 6.3.2. Demuestre las fórmulas de recurrencia anteriores.

Sumando y restando las relaciones de arriba, obtenemos

$$C_{\nu-1}(z) = \frac{\nu}{z} C_\nu(z) + C'_\nu(z), \quad (6.75c)$$

$$C_{\nu+1}(z) = \frac{\nu}{z} C_\nu(z) - C'_\nu(z). \quad (6.75d)$$

¹⁰La convergencia absoluta y uniforme de las series de J_ν y $J_{-\nu}$ para $x > 0$, junto con la continuidad de las soluciones de (6.66a) respecto de ν , garantizan la validez de esta afirmación aún cuando ν tiende a un entero.

¹¹Nótese que si $\nu \rightarrow -m$ con m natural, todos los términos de la serie con $n < m$ se anularán, ya que $1/\Gamma(k) = 0$ para todo entero negativo k . La consecuencia es la conocida relación $J_{-m}(x) = (-1)^m J_m(x)$ para $m \in \mathbb{N}$.

De aquí obtenemos también los casos particulares

$$\begin{aligned} J'_0(z) &= -J_1(z), & Y'_0(z) &= -Y_1(z), \\ H_0^{(1)\prime}(z) &= -H_1^{(1)}(z), & H_0^{(2)\prime}(z) &= -H^{(2)}(z). \end{aligned} \quad (6.76)$$

6.3.3. Integrales de funciones de Bessel

De las relaciones de recurrencia podemos obtener relaciones útiles a la hora de integrar funciones de Bessel. Por ejemplo de $C'_0(z) = -C_1(z)$ es inmediato que

$$\int C_1(x) dx = -C_0(x). \quad (6.77a)$$

Otro ejemplo es, usando (6.75c),

$$[x^\nu C_\nu(x)]' = x^\nu \left[C'_\nu(x) + \frac{\nu}{x} C_\nu(x) \right] = x^\nu C_{\nu-1}(x),$$

de donde obtenemos

$$\int x^\nu C_{\nu-1}(x) dx = x^\nu C_\nu(x). \quad (6.77b)$$

Y aún otro es, usando (6.75d),

$$[x^{-\nu} C_\nu(x)]' = x^{-\nu} \left[C'_\nu(x) - \frac{\nu}{x} C_\nu(x) \right] = -x^{-\nu} C_{\nu+1}(x),$$

de donde obtenemos

$$\int x^{-\nu} C_{\nu+1}(x) dx = -x^{-\nu} C_\nu(x). \quad (6.77c)$$

Otra integral útil puede obtenerse directamente de la ecuación de Bessel. Integrando una vez por partes,

$$\int C_\nu^2(x) x dx = \frac{1}{2} x^2 C_\nu^2(x) - \int C_\nu(x) C'_\nu(x) x^2 dx.$$

Pero de (6.64)

$$x^2 C_\nu(x) = \nu^2 C_\nu(x) - x C'_\nu(x) - x^2 C''_\nu(x),$$

luego

$$\int C_\nu^2(x) x dx = \frac{1}{2} x^2 C_\nu^2(x) - \int C'_\nu(x) [\nu^2 C_\nu(x) - x C'_\nu(x) - x^2 C''_\nu(x)] dx,$$

de donde finalmente

$$\int C_\nu^2(kx) x dx = \frac{1}{2} \left(x^2 - \frac{\nu^2}{k^2} \right) C_\nu^2(kx) + \frac{1}{2} x^2 [C'_\nu(kx)]^2, \quad (6.78)$$

que es la integral de normalización para funciones de Bessel.

Ejercicio 6.3.3. Complete los cálculos precedentes y verifique (6.71).

6.3.4. Función generatriz

Las relaciones de recurrencia también facilitan la construcción de una función generatriz para las funciones de Bessel de orden entero,

$$F(z, h) = \sum_{n=-\infty}^{\infty} h^n C_n(z). \quad (6.79)$$

Multiplicando (6.75a) por h^n y sumando sobre n obtenemos

$$\left(h + \frac{1}{h} \right) F(z, h) = \frac{2h}{z} \frac{\partial F(z, h)}{\partial h},$$

e integrando esta EDP resulta

$$F(z, h) = \phi(z) \exp \left[\frac{z}{2} \left(h - \frac{1}{h} \right) \right],$$

con $\phi(z)$ a determinar.

La elección de la función ϕ nos permite construir funciones generatrices para diferentes funciones de Bessel. Por ejemplo, eligiendo $\phi(z) = 1$ obtenemos

$$F(z, h) = \sum_{n=-\infty}^{\infty} h^n J_n(z) = \exp \left[\frac{z}{2} \left(h - \frac{1}{h} \right) \right], \quad (6.80)$$

la función generatriz para las funciones de Besel de primer tipo. Este resultado puede verificarse de manera directa escribiendo

$$F(z, h) = e^{zh/2} e^{-z/2h} = \sum_{r=0}^{\infty} \left(\frac{z}{2} \right)^r \frac{h^r}{r!} \sum_{s=0}^{\infty} (-1)^s \left(\frac{z}{2} \right)^s \frac{h^{-s}}{s!}$$

y reordenando la doble suma para obtener

$$F(z, h) = \sum_{n=-\infty}^{\infty} h^n \sum_{s=0}^{\infty} \frac{(-1)^s}{s!(s+n)!} \left(\frac{z}{2} \right)^{2s+n},$$

donde el coeficiente de h^n es $J_n(z)$ según (6.67a).

Ejercicio 6.3.4. Complete la demostración precedente.

Ejemplo 6.3.1. Como ejemplo de la utilidad de la función generatriz (6.80), tenemos

$$\begin{aligned} \sum_n J_n(x+y) h^n &= \exp \left[\frac{x+y}{2} \left(h - \frac{1}{h} \right) \right] \\ &= \exp \left[\frac{x}{2} \left(h - \frac{1}{h} \right) \right] \exp \left[\frac{y}{2} \left(h - \frac{1}{h} \right) \right] = \sum_r J_r(x) h^r \sum_s J_s(y) h^s, \end{aligned}$$

de donde

$$J_n(x+y) = \sum_m J_m(x) J_{n-m}(y). \quad (6.81)$$

Ejercicio 6.3.5. Complete los cálculos anteriores.

Si en (6.80) reemplazamos h por $e^{i\theta}$, obtenemos la serie de Fourier

$$e^{iz \operatorname{sen} \theta} = \sum_{n=-\infty}^{\infty} J_n(z) e^{in\theta}, \quad (6.82a)$$

y cambiando θ por $\theta + \frac{\pi}{2}$ tenemos

$$e^{iz \cos \theta} = \sum_{n=-\infty}^{\infty} i^n J_n(z) e^{in\theta}. \quad (6.82b)$$

De estas expresiones, separando las partes real e imaginaria, obtenemos las series de Fourier

$$\cos(z \operatorname{sen} \theta) = J_0(z) + 2 \sum_{n=1}^{\infty} J_{2n}(z) \cos(2n\theta), \quad (6.82c)$$

$$\operatorname{sen}(z \operatorname{sen} \theta) = 2 \sum_{n=0}^{\infty} J_{2n+1}(z) \operatorname{sen}((2n+1)\theta), \quad (6.82d)$$

$$\cos(z \cos \theta) = J_0(z) + 2 \sum_{n=1}^{\infty} (-1)^n J_{2n}(z) \cos(2n\theta), \quad (6.82e)$$

$$\operatorname{sen}(z \cos \theta) = -2 \sum_{n=0}^{\infty} (-1)^n J_{2n+1}(z) \operatorname{sen}((2n+1)\theta). \quad (6.82f)$$

Las expresiones (6.82) se conocen como *desarrollos de Jacobi–Anger*.

6.3.5. Representaciones integrales

De la función generatriz (6.80), usando la fórmula de Cauchy es inmediato obtener la *representación integral de Schläfli*

$$J_n(z) = \frac{1}{2\pi i} \oint \frac{\exp\left[\frac{z}{2}\left(t - \frac{1}{t}\right)\right]}{t^{n+1}} dt, \quad (6.83)$$

donde el contorno de integración envuelve el origen una vez en sentido positivo.

Parametrizando el camino de integración de (6.83) como $t = e^{i\theta}$ obtenemos la *integral de Bessel*

$$J_n(z) = \frac{1}{2\pi} \int_0^{2\pi} e^{i(z \operatorname{sen} \theta - n\theta)} d\theta = \frac{1}{\pi} \int_0^\pi \cos(n\theta - z \operatorname{sen} \theta) d\theta. \quad (6.84)$$

Ejercicio 6.3.6. Complete los cálculos anteriores.

6.3.6. Desarrollos asintóticos

Las expresiones asintóticas de las funciones de Bessel para $z \rightarrow 0$ y $z \rightarrow \infty$ son relevantes en muchas aplicaciones prácticas, y se obtienen como el primer término del desarrollo en serie alrededor de $z = 0$ o del desarrollo en serie asintótica para $z \rightarrow \infty$, respectivamente. Daremos sin demostración los principales resultados.

Para $z \rightarrow 0$ (con $|\arg z| < \pi$ si ν no es entero) tenemos

$$J_{\pm n}(z) \underset{z \rightarrow 0}{\asymp} (\pm 1)^n \frac{(z/2)^n}{n!}, \quad n = 0, 1, 2, \dots \quad (6.85a)$$

$$J_\nu(z) \underset{z \rightarrow 0}{\asymp} \frac{(z/2)^\nu}{\Gamma(\nu + 1)}, \quad \nu \neq \pm 1, \pm 2, \dots \quad (6.85b)$$

$$Y_0(z), -iH_0^{(1)}(z), iH_0^{(2)}(z) \underset{z \rightarrow 0}{\asymp} \frac{2}{\pi} \ln(z), \quad (6.85c)$$

$$Y_\nu(z), -iH_\nu^{(1)}(z), iH_\nu^{(2)}(z) \underset{z \rightarrow 0}{\asymp} \frac{1}{\pi} \Gamma(\nu)(z/2)^{-\nu}, \quad \nu > 0. \quad (6.85d)$$

Para $z \rightarrow \infty$ con $|\arg z| < \pi$ y para todo ν tenemos

$$J_\nu(z) \underset{z \rightarrow \infty}{\asymp} \sqrt{\frac{2}{\pi z}} \cos\left(z - \nu \frac{\pi}{2} - \frac{\pi}{4}\right), \quad (6.86a)$$

$$Y_\nu(z) \underset{z \rightarrow \infty}{\asymp} \sqrt{\frac{2}{\pi z}} \sin\left(z - \nu \frac{\pi}{2} - \frac{\pi}{4}\right), \quad (6.86b)$$

$$H_\nu^{(1)}(z) \underset{z \rightarrow \infty}{\asymp} \sqrt{\frac{2}{\pi z}} \exp\left[i\left(z - \nu \frac{\pi}{2} - \frac{\pi}{4}\right)\right], \quad (6.86c)$$

$$H_\nu^{(2)}(z) \underset{z \rightarrow \infty}{\asymp} \sqrt{\frac{2}{\pi z}} \exp\left[-i\left(z - \nu \frac{\pi}{2} - \frac{\pi}{4}\right)\right]. \quad (6.86d)$$

6.3.7. Otros PSL asociados a la ecuación de Bessel

Las funciones de Bessel de segundo tipo Y_ν cobran relevancia si debemos resolver un PSL donde $x = 0$ no pertenece al intervalo de interés. Un ejemplo típico es

$$xy'' + y' + \left(k^2x - \frac{\nu^2}{x}\right)y = 0, \quad \nu \in \mathbb{R}, \quad x \in (a, b), \quad (6.87a)$$

$$y(a) = y(b) = 0 \quad (6.87b)$$

con $0 < a < b < \infty$. En ese caso la solución general de (6.87a) en (a, b) es alguna combinación lineal de $J_\nu(kx)$ e $Y_\nu(kx)$; el procedimiento para ajustar las CC (6.87b) es completamente equivalente al seguido en la Sección 6.2.7 y nos lleva a construir autofunciones de la forma

$$y_n(x) = Y_\nu(k_n a) J_\nu(k_n x) - J_\nu(k_n a) Y_\nu(k_n x), \quad n = 1, 2, 3, \dots,$$

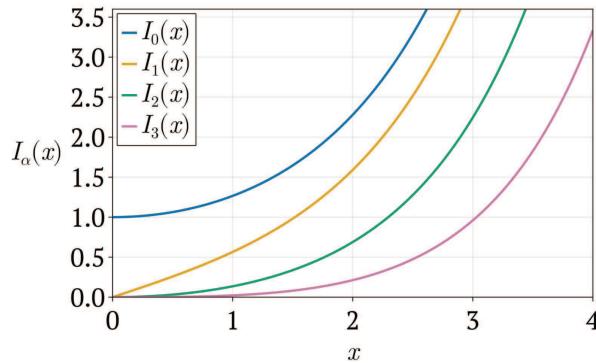


Figura 6.10: Las primeras tres funciones de Bessel modificadas de primer tipo I_n . Imagen de Aspir8, bajo licencia Creative Commons CC0 1.0 Universal Public Domain Dedication.

con los autovalores k_n^2 determinados por las soluciones de la ecuación

$$Y_\nu(k_n a) J_\nu(k_n b) - J_\nu(k_n a) Y_\nu(k_n b) = 0.$$

Como siempre, las autofunciones $\{y_n\}$ forman un conjunto ortogonal (con peso $\rho(x) = x$) y completo en $L^2_\rho(a, b)$, y la integral de normalización puede calcularse usando (6.78). Las relaciones de recurrencia (6.75) y las integrales (6.77) siguen siendo válidas.

Las funciones de Hankel $H_\nu^{(1)}$ y $H_\nu^{(2)}$ cobran relevancia en problemas donde $b \rightarrow \infty$ y especificamos condiciones asintóticas, por ejemplo (como veremos en el Capítulo 9) si debemos ajustar CC oscilatorias en el tiempo en $x = a > 0$ y condiciones de onda saliente para $x \rightarrow \infty$.

6.3.8. Las funciones de Bessel modificadas

La ecuación diferencial de Bessel (6.64) aparece también habitualmente en la forma

$$xy'' + y' - \left(k^2 x + \frac{\nu^2}{x}\right)y = 0, \quad \nu \in \mathbb{R}, \quad x \in (a, b). \quad (6.88)$$

Las dos soluciones LI de esta EDO son conocidas como las *funciones de Bessel modificadas* de primer y segundo tipo de orden ν , respectivamente definidas como

$$I_\nu(kx) = \sum_{n=0}^{\infty} \frac{1}{n! \Gamma(n+1+\nu)} \left(\frac{kx}{2}\right)^{2n+\nu}, \quad (6.89a)$$

$$K_\nu(kx) = \frac{\pi}{2} \frac{I_{-\nu}(kx) - I_\nu(kx)}{\sin(\nu\pi)}. \quad (6.89b)$$

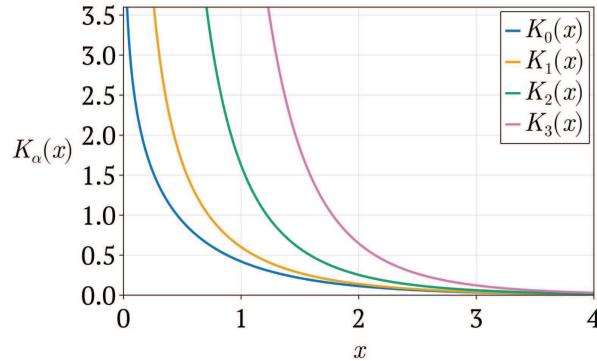


Figura 6.11: Las primeras tres funciones de Bessel modificadas de segundo tipo K_n . Imagen de Aspir8, bajo licencia Creative Commons CC0 1.0 Universal Public Domain Dedication.

Las primeras funciones modificadas de primer tipo se muestran en la Figura 6.10, y las correspondientes de segundo tipo en la Figura 6.11; allí puede apreciarse que estas funciones *no* son oscilatorias, a diferencia de J_ν e Y_ν y de acuerdo a lo previsto por el Corolario 5.2.6.

Las funciones I_ν y K_ν tiene una relación directa con J_ν , Y_ν , $H_\nu^{(1)}$ y $H_\nu^{(2)}$,

$$I_\nu(z) = i^{-\nu} J_\nu(iz), \quad K_\nu(z) = \frac{\pi}{2} i^{1+\nu} H_\nu^{(1)}(iz), \quad -\pi < \arg z < \pi/2, \quad (6.90a)$$

$$I_\nu(z) = i^\nu J_\nu(-iz), \quad K_\nu(z) = -\frac{\pi}{2} i^{1-\nu} H_\nu^{(2)}(-iz), \quad -\pi/2 < \arg z < \pi, \quad (6.90b)$$

como puede verificarse comparando sus definiciones; por este motivo suele llamarlas *funciones de Bessel de argumento imaginario*.

Estas relaciones permiten obtener con facilidad las relaciones de recurrencia a partir de (6.75), por ejemplo

$$Z_\nu(z) = \frac{z}{2\nu} [Z_{\nu-1}(z) - Z_{\nu+1}(z)], \quad (6.91a)$$

$$Z'_\nu(z) = \frac{1}{2} [Z_{\nu-1}(z) + Z_{\nu+1}(z)] \quad (6.91b)$$

(con Z_ν cualquiera de I_ν o K_ν), la función generatriz

$$F(z, h) = \sum_{n=-\infty}^{\infty} h^n I_n(z) = \exp \left[\frac{z}{2} \left(h + \frac{1}{h} \right) \right], \quad (6.92)$$

la representación integral

$$I_n(z) = \frac{1}{\pi} \int_0^\pi e^{z \cos \theta} \cos(n\theta) d\theta, \quad (6.93)$$

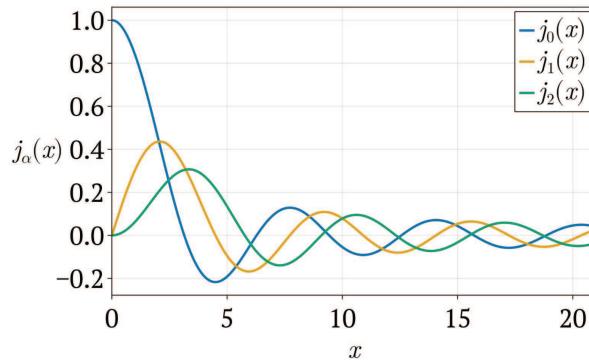


Figura 6.12: Las primeras tres funciones de Bessel esféricas de primer tipo j_n . Imagen de Aspir8, bajo licencia Creative Commons CC0 1.0 Universal Public Domain Dedication.

o los desarrollos asintóticos para $z \rightarrow 0$ (con $|\arg z| < \pi$ si ν no es entero)

$$I_{\pm n}(z) \underset{z \rightarrow 0}{\asymp} \frac{(z/2)^n}{n!}, \quad n = 0, 1, 2, \dots \quad (6.94a)$$

$$I_\nu(z) \underset{z \rightarrow 0}{\asymp} \frac{(z/2)^\nu}{\Gamma(\nu + 1)}, \quad \nu \neq \pm 1, \pm 2, \dots \quad (6.94b)$$

$$K_0(z) \underset{z \rightarrow 0}{\asymp} \ln(z/2), \quad (6.94c)$$

$$K_\nu(z) \underset{z \rightarrow 0}{\asymp} \frac{1}{2}\Gamma(\nu)(z/2)^{-\nu}, \quad \nu > 0, \quad (6.94d)$$

y para $z \rightarrow \infty$ con $|\arg z| < \pi$ y para todo ν

$$I_\nu(z) \underset{z \rightarrow \infty}{\asymp} \frac{1}{\sqrt{2\pi}} z^{-1/2} e^z, \quad (6.95a)$$

$$K_\nu(z) \underset{z \rightarrow \infty}{\asymp} \frac{1}{\sqrt{2\pi}} z^{-1/2} e^{-z}. \quad (6.95b)$$

6.3.9. Las funciones de Bessel esféricas

Las funciones de Bessel de orden *entero* suelen llamarse *funciones cilíndricas* debido a que aparecen asociadas a la separación de EDP que contienen ∇^2 en coordenadas cilíndricas. Su equivalente para la separación en coordenadas esféricas son las funciones de Bessel de orden *semientero*, conocidas como *funciones de Bessel esféricas*. Si bien desde un punto de vista puramente matemático pueden tratarse simplemente como otros casos particulares de las funciones de Bessel de orden arbitrario, el uso les ha asignado una notación particular que emplea letras *minúsculas* donde la notación general emplea mayúsculas.

Las funciones de primer y segundo tipo se definen como

$$j_n(x) = \sqrt{\frac{\pi}{2x}} J_{n+\frac{1}{2}}(x), \quad (6.96a)$$

$$y_n(x) = \sqrt{\frac{\pi}{2x}} Y_{n+\frac{1}{2}}(x), \quad (6.96b)$$

y las primeras de ambas se muestran, respectivamente, en las Figuras 6.12 y 6.13. De su definición es directo ver que

$$j_n(x) = (-1)^n y_{-n-1}(x), \quad y_n(x) = (-1)^{n+1} j_{-n-1}(x). \quad (6.97)$$

Un hecho notable es que las funciones de Bessel esféricas pueden expresarse de manera sencilla en términos de funciones elementales, por ejemplo mediante las *fórmulas de Rayleigh*

$$j_n(x) = (-x)^n \left(\frac{1}{x} \frac{d}{dx} \right)^n \frac{\sin x}{x}, \quad y_n(x) = -(-x)^n \left(\frac{1}{x} \frac{d}{dx} \right)^n \frac{\cos x}{x}, \quad (6.98)$$

con los casos particulares

$$j_0(x) = \frac{\sin x}{x}, \quad y_0(x) = -\frac{\cos x}{x}. \quad (6.99)$$

También sus funciones generatrices tienen expresiones particularmente simples válidas para $2|h| < |z|$,

$$\frac{1}{z} \cos(\sqrt{z^2 - zh} z) = \frac{\cos z}{z} + \sum_{n=1}^{\infty} \frac{h^n}{n!} j_{n-1}(z), \quad (6.100a)$$

$$\frac{1}{z} \sin(\sqrt{z^2 - zh} z) = \frac{\sin z}{z} + \sum_{n=1}^{\infty} \frac{h^n}{n!} y_{n-1}(z), \quad (6.100b)$$

y tenemos las útiles representaciones integrales

$$j_n(z) = \frac{z^n}{2^{n+1} n!} \int_0^\pi \cos(z \cos \theta) \sin^{2n+1} \theta d\theta \quad (6.101a)$$

$$= \frac{(-1)^n}{2} \int_0^\pi e^{iz \cos \theta} P_n(\cos \theta) \sin \theta d\theta, \quad (6.101b)$$

donde P_n son los polinomios de Legendre.

Las funciones de Hankel esféricas se definen, como es de esperar, por

$$h_n^{(1)}(x) = j_n(x) + iy_n(x), \quad (6.102a)$$

$$h_n^{(2)}(x) = j_n(x) - iy_n(x), \quad (6.102b)$$

de donde es inmediato que

$$h_0^{(1)}(x) = \frac{e^{ix}}{ix}, \quad h_0^{(2)}(x) = \frac{e^{-ix}}{-ix} \quad (6.103)$$

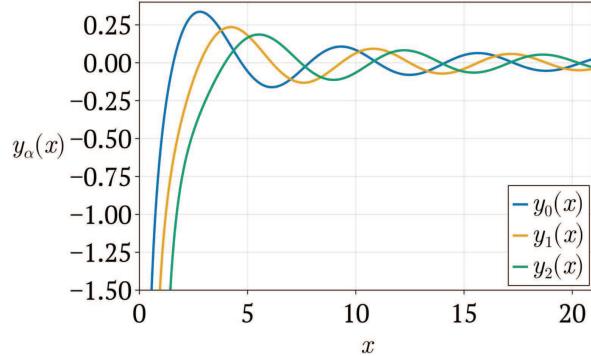


Figura 6.13: Las primeras tres funciones de Bessel esféricas de segundo tipo y_n . Imagen de Aspir8, bajo licencia Creative Commons CC0 1.0 Universal Public Domain Dedication.

y que, para $2|h| < |z|$,

$$\frac{1}{z} e^{i\sqrt{z^2-zh}z} = \frac{e^{iz}}{z} + \sum_{n=1}^{\infty} \frac{h^n}{n!} h_{n-1}^{(1)}(z), \quad (6.104a)$$

$$\frac{1}{z} e^{-i\sqrt{z^2-zh}z} = \frac{e^{-iz}}{z} + \sum_{n=1}^{\infty} \frac{h^n}{n!} h_{n-1}^{(2)}(z). \quad (6.104b)$$

Finalmente, las funciones de Bessel esféricas modificadas se definen como

$$i_n(x) = \sqrt{\frac{\pi}{2x}} I_{n+\frac{1}{2}}(x), \quad (6.105a)$$

$$k_n(x) = \sqrt{\frac{\pi}{2x}} K_{n+\frac{1}{2}}(x), \quad (6.105b)$$

con las fórmulas de Rayleigh

$$i_n(x) = x^n \left(\frac{1}{x} \frac{d}{dx} \right)^n \frac{\sinh x}{x}, \quad k_n(x) = (-x)^n \left(\frac{1}{x} \frac{d}{dx} \right)^n \frac{e^{-x}}{x} \quad (6.106)$$

y los casos particulares

$$i_0(x) = \frac{\sinh x}{x}, \quad k_0(x) = -\frac{e^{-x}}{x}. \quad (6.107)$$

Capítulo 7

Transformadas Integrales

En los Capítulos 5 y 6 hemos visto cómo podemos obtener representaciones de funciones f en \mathcal{L}^2 sobre intervalos dados, como combinaciones lineales de autofunciones de un PSL sobre ese intervalo. En todos los casos tratados hasta ahora, la naturaleza del PSL ha sido tal que los autovalores han formado un conjunto discreto (infinito numerable), por lo que estas representaciones adoptaron la forma de *series*.

Cabe preguntarse qué ocurre cuando, por el motivo que sea, el conjunto de autovalores forma un conjunto *continuo*. En concreto, dado un problema de autovalores (aunque no necesariamente de Sturm–Liouville)

$$Ly_\lambda + \lambda\rho y_\lambda = 0, \quad x \in I \subset \mathbb{R},$$

donde y_λ es la autofunción de autovalor λ y estos pueden adoptar valores en todo un intervalo J (finito o no) de \mathbb{R} , ¿existirá una representación equivalente para funciones en $\mathcal{L}^2[I]$? Y en caso afirmativo, ¿qué forma adoptará? La intuición parece indicar que deberíamos esperar algo como

$$f(x) = \int_J \phi_\lambda y_\lambda(x) d\lambda$$

con

$$\phi_\lambda = \frac{\langle y_\lambda, f \rangle_\rho}{\langle y_\lambda, y_\lambda \rangle_\rho},$$

pero evidentemente faltan muchos detalles.

Gran parte de estos problemas cae dentro del área conocida como *transformadas integrales*, mapas lineales que llevan funciones $f(t)$ a otras $\mathcal{T}f(s)$ mediante

$$\mathcal{T}f(s) = \int_{t_1}^{t_2} f(t)K(t, s) dt,$$

donde $K(t, s)$ se conoce como el *kernel* o *núcleo* de la transformación. El área abarca mucho más de lo que este curso pretende (o puede) cubrir. Trataremos

aquí con unas pocas transformadas integrales de mucha utilidad en problemas típicos de la Física, sin pretender excesivo rigor ni mucho menos desarrollar una teoría general.

7.1. La transformada de Fourier

Comenzaremos considerando una función $f \in L^2(-\ell/2, \ell/2)$, que como vimos en la Sección 6.1 puede desarrollarse en las autofunciones (6.2b) del PSL (6.1). Reescribiendo (6.3) y (6.4) como

$$f(x) = \sum_{n=-\infty}^{\infty} a_n e^{ik_n x}, \quad a_n = \frac{1}{\ell} \int_{-\ell/2}^{\ell/2} e^{-ik_n x} f(x) dx, \quad k_n = \frac{2n\pi}{\ell},$$

queremos ver qué ocurre si $\ell \rightarrow \infty$, ya que justamente en ese caso la distancia $2\pi/\ell$ entre k_n sucesivos tenderá a cero, y los autovalores $\lambda_n = k_n^2$ pasarán a formar un continuo.

Definiendo

$$\tilde{f}(k_n) := \ell a_n, \quad \Delta k := \frac{2\pi}{\ell},$$

tendremos

$$f(x) = \frac{1}{2\pi} \sum_{n=-\infty}^{\infty} \tilde{f}(k_n) e^{ik_n x} \Delta k, \quad \tilde{f}(k_n) = \int_{-\ell/2}^{\ell/2} e^{-ik_n x} f(x) dx.$$

En el límite en que $\ell \rightarrow \infty$, la variable discreta k_n pasa a ser una variable continua k , $\Delta k \rightarrow 0$, y esperaríamos tener

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \tilde{f}(k) e^{ikx} dk, \quad \tilde{f}(k) = \int_{-\infty}^{\infty} f(x) e^{-ikx} dx,$$

suponiendo que las integrales tengan sentido.

Por supuesto, lo anterior no es (ni pretende ser) un argumento riguroso, sino simplemente una motivación para las definiciones que introduciremos a continuación.

7.1.1. Definición de la transformada de Fourier

En base a lo discutido arriba, vamos a definir

$$\mathcal{F}_{x \rightarrow k}[f(x)] = \tilde{f}(k) := \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(x) e^{-ikx} dx, \quad (7.1a)$$

$$\mathcal{F}^{-1}_{k \rightarrow x}[\tilde{f}(k)] = f(x) := \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \tilde{f}(k) e^{ikx} dk, \quad (7.1b)$$

donde $\tilde{f}(k)$ se conoce como la *transformada de Fourier* de $f(x)$, y $f(x)$ se dice la *antitransformada de Fourier* de $\tilde{f}(k)$. Consideraremos estas definiciones válidas

para todo par de *funciones* f y \tilde{f} para las cuales las correspondientes integrales (improperas) estén bien definidas. Esto incluye funciones en $\mathcal{L}^1(-\infty, \infty)$ (funciones de módulo integrable, es decir $\int_{-\infty}^{\infty} |f(s)| ds < \infty$) y en $\mathcal{L}^2(-\infty, \infty)$. La elección de “repartir” simétricamente el factor $1/2\pi$ es la más común en Física, y ahorra problemas en el uso de las fórmulas; definiciones alternativas usuales son

$$\tilde{f}(k) = \int_{-\infty}^{\infty} f(x) e^{-ikx} dx, \quad f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \tilde{f}(k) e^{ikx} dk$$

y

$$\tilde{f}(k) = \int_{-\infty}^{\infty} f(x) e^{-2\pi ikx} dx, \quad f(x) = \int_{-\infty}^{\infty} \tilde{f}(k) e^{2\pi ikx} dk.$$

Ejemplo 7.1.1. La transformada de Fourier de la función

$$f(x) = \text{rect}(x) = \begin{cases} 1, & |x| \leq \frac{1}{2}, \\ 0, & |x| > \frac{1}{2}, \end{cases}$$

es la función

$$\tilde{f}(k) = \frac{1}{\sqrt{2\pi}} \text{sinc}(k/2) = \frac{1}{\sqrt{2\pi}} \frac{\sin(k/2)}{k/2}.$$

Ejercicio 7.1.1. Complete los cálculos del ejemplo de arriba.

Ahora bien, en las aplicaciones muchas veces querremos transformar o antitransformar funciones que *no* están en $\mathcal{L}^1(-\infty, \infty)$ o $\mathcal{L}^2(-\infty, \infty)$, o que ni siquiera son funciones, sino *distribuciones*. Por ejemplo, ¿cómo definiríamos la transformada de la delta de Dirac?

Para extender las definiciones, comencemos por notar que toda función de prueba de Schwartz $\varphi(x) \in \mathcal{S}[\mathbb{R}]$, de acuerdo a la definición de la Sección 4.1.1, tiene una transformada de Fourier $\tilde{\varphi}(k)$ bien definida según (7.1),

$$\tilde{\varphi}(k) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \varphi(x) e^{-ikx} dx.$$

Asimismo, para todo entero $n \geq 0$, la integral

$$\frac{d^n \tilde{\varphi}(k)}{dk^n} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} (-ix)^n \varphi(x) e^{-ikx} dx$$

también existe y es finita, ya que $\varphi(x) \in \mathcal{S}[\mathbb{R}]$; es decir, $\tilde{\varphi}(k) \in \mathcal{C}^\infty$. Además, acotando el valor absoluto de la última integral es fácil verificar que para todo entero $p \geq 0$

$$|k^p \tilde{\varphi}^{(n)}(k)| < \infty$$

cuando $|k| \rightarrow \infty$.

Ejercicio 7.1.2. Demuestre la afirmación precedente.

Concluimos así que la transformada de Fourier de una función de prueba de Schwartz es también una función de prueba de Schwartz y, evidentemente, lo mismo vale para la antittransformada, lo que podemos escribir como

$$\mathcal{F}, \mathcal{F}^{-1} : \mathcal{S}[\mathbb{R}] \rightarrow \mathcal{S}[\mathbb{R}],$$

y estos mapas son evidentemente lineales.

Ahora consideremos una función $f(x) \in \mathcal{L}^2[\mathbb{R}] \cap \mathcal{C}^\infty[\mathbb{R}]$ o $\mathcal{L}^1[\mathbb{R}] \cap \mathcal{C}^\infty[\mathbb{R}]$, cuya transformada de Fourier $\tilde{f}(k)$ exista según la definición (7.1), y definamos de la manera usual las distribuciones correspondientes mediante los funcionales $\langle f, \varphi \rangle$ y $\langle \tilde{f}, \varphi \rangle$, con $\varphi \in \mathcal{S}[\mathbb{R}]$; luego

$$\begin{aligned}\langle \tilde{f}, \varphi \rangle &= \int_{-\infty}^{\infty} \tilde{f}(s) \varphi(s) ds = \int_{-\infty}^{\infty} \tilde{f}(s) \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \tilde{\varphi}(t) e^{ist} dt ds \\ &= \int_{-\infty}^{\infty} \tilde{\varphi}(t) \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \tilde{f}(s) e^{ist} ds dt = \int_{-\infty}^{\infty} f(t) \tilde{\varphi}(t) dt \\ &= \langle f, \tilde{\varphi} \rangle\end{aligned}$$

para toda distribución regular.

Finalmente, dada una distribución temperada f según la definimos en la Sección 4.1.3, sea regular o no, el Teorema 4.1.1 garantiza que existe una sucesión de distribuciones regulares f_n que converge a f , y para cada una de las f_n tendremos $\langle \tilde{f}_n, \varphi \rangle = \langle f_n, \tilde{\varphi} \rangle$; por lo tanto definimos, *para cualquier distribución temperada f* , su transformada de Fourier \tilde{f} por

$$\langle \tilde{f}, \varphi \rangle := \langle f, \tilde{\varphi} \rangle, \quad (7.1c)$$

definición que generaliza (7.1).¹

Ejemplo 7.1.2. La delta de Dirac es una distribución temperada, como queda de manifiesto en el Ejercicio 4.1.4; luego si

$$f(x) = \delta(x - x_0),$$

entonces

$$\langle \tilde{f}, \varphi \rangle = \langle f, \tilde{\varphi} \rangle = \tilde{\varphi}(x_0) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \varphi(y) e^{-iyx_0} dy = \left\langle \frac{1}{\sqrt{2\pi}} e^{-ikx_0}, \varphi(k) \right\rangle,$$

de donde

$$\tilde{f}(x) = \mathcal{F}_{x \rightarrow k} [\delta(x - x_0)] = \frac{1}{\sqrt{2\pi}} e^{-ikx_0}.$$

Como caso particular, tendremos

$$\mathcal{F}_{x \rightarrow k} [\delta(x)] = \frac{1}{\sqrt{2\pi}},$$

¹El lector interesado en más detalles puede consultar el excelente pero avanzado texto de R. D. Richtmyer, *Principles of advanced Mathematical Physics*, Vol. 1. Springer, 1978.

con la notable implicación de que

$$\mathcal{F}_{k \rightarrow x}^{-1} \left[\frac{1}{\sqrt{2\pi}} \right] = \delta(x).$$

Esto nos proporciona otra representación de la delta de Dirac:

$$\delta(x - x_0) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{ik(x-x_0)} dk. \quad (7.2)$$

Ejercicio 7.1.3. Muestre que

$$\mathcal{F}_{x \rightarrow k}[\delta^{(n)}(x - x_0)] = \frac{(ik)^n}{\sqrt{2\pi}} e^{-ikx_0}, \quad \mathcal{F}_{k \rightarrow x}^{-1}[\delta^{(n)}(k - k_0)] = \frac{(-ix)^n}{\sqrt{2\pi}} e^{ik_0 x}.$$

Todo lo dicho hasta aquí puede generalizarse de manera directa a más de una dimensión; por ejemplo, las definiciones (7.1a) y (7.1b) pueden generalizarse a tres dimensiones como

$$\mathcal{F}_{\vec{x} \rightarrow \vec{k}}[f(\vec{x})] = \tilde{f}(\vec{k}) := \frac{1}{(2\pi)^{3/2}} \iiint_{-\infty}^{\infty} f(\vec{x}) e^{-i\vec{k} \cdot \vec{x}} d^3x, \quad (7.3a)$$

$$\mathcal{F}_{\vec{k} \rightarrow \vec{x}}^{-1}[\tilde{f}(\vec{k})] = f(\vec{x}) := \frac{1}{(2\pi)^{3/2}} \iiint_{-\infty}^{\infty} \tilde{f}(\vec{k}) e^{i\vec{k} \cdot \vec{x}} d^3k, \quad (7.3b)$$

y lo mismo con (7.1c). Ello no quiere decir que las coordenadas Cartesianas sean necesariamente las más ventajosas para realizar las integrales: las simetrías de f y \tilde{f} pueden hacer preferible el uso de coordenadas curvilíneas, como veremos en el ejemplo de aplicación 7.1.6.

7.1.2. Propiedades de la transformada de Fourier

A partir de las definiciones (7.1) podemos deducir una serie de propiedades útiles de la transformada de Fourier, algunas de las cuales presentaremos a continuación. La generalización a transformadas multidimensionales es directa, y se deja como ejercicio para el lector.

Linealidad.

La transformada y antitransformada de Fourier son, evidentemente, lineales:

$$\mathcal{F}[\alpha f + \beta g] = \alpha \mathcal{F}[f] + \beta \mathcal{F}[g], \quad (7.4a)$$

$$\mathcal{F}^{-1}[\alpha f + \beta g] = \alpha \mathcal{F}^{-1}[f] + \beta \mathcal{F}^{-1}[g]. \quad (7.4b)$$

Derivación.

A partir de las definiciones (7.1) es inmediato mostrar que, para cualquier f en \mathcal{L}^1 , \mathcal{L}^2 o distribución temperada,

$$\mathcal{F}_{x \rightarrow k}[f^{(n)}(x)] = (ik)^n \mathcal{F}_{x \rightarrow k}[f(x)]. \quad (7.5)$$

Ejercicio 7.1.4. Demuestre la afirmación precedente.

Integración.

Sea

$$f(x) = g'(x),$$

luego

$$\tilde{f}(k) = ik \tilde{g}(k).$$

Pero

$$g(x) = \int f(x) dx - \frac{c}{\sqrt{2\pi}},$$

donde elegimos escribir la constante de integración como $-c/\sqrt{2\pi}$. Entonces

$$\mathcal{F}_{x \rightarrow k} \left[\int f(x) dx \right] = \frac{1}{ik} \mathcal{F}_{x \rightarrow k}[f(x)] + c\delta(k), \quad (7.6)$$

con c una constante arbitraria. Otra forma de obtener este resultado es recordando que, como mostramos en la Sección 4.1.4, $k\delta(k)$ es la distribución nula.

Ejercicio 7.1.5. Complete los cálculos que llevan al resultado anterior.

Traslación en x .

De las definiciones (7.1) tenemos

$$\mathcal{F}_{x \rightarrow k}[f(x+a)] = e^{ika} \mathcal{F}_{x \rightarrow k}[f(x)]. \quad (7.7)$$

Ejercicio 7.1.6. Demuestre la afirmación precedente.

Traslación en k .

De las definiciones (7.1) también resulta

$$\mathcal{F}_{x \rightarrow k}[e^{i\alpha x} f(x)] = \mathcal{F}_{x \rightarrow k}[f(x)](k - \alpha) = \tilde{f}(k - \alpha). \quad (7.8)$$

Ejercicio 7.1.7. Demuestre el resultado precedente.

Escaleo

Otro resultado inmediato de las definiciones (7.1) es

$$\mathcal{F}_{x \rightarrow k}[f(ax)] = \frac{1}{|a|} \mathcal{F}_{x \rightarrow k}[f(x)](k/a) = \frac{1}{|a|} \tilde{f}\left(\frac{k}{a}\right). \quad (7.9)$$

Ejercicio 7.1.8. Demuestre el resultado precedente.

Multiplicación por potencias.

De las definiciones (7.1) tendremos

$$\mathcal{F}_{x \rightarrow k}[x^n f(x)] = i^n \frac{\partial^n}{\partial k^n} \mathcal{F}_{x \rightarrow k}[f(x)]. \quad (7.10)$$

Ejercicio 7.1.9. Demuestre el resultado anterior.

Convolución

Dadas dos funciones $f_1(x)$ y $f_2(x)$, definimos su *convolución* $f_1 * f_2$ como

$$(f_1 * f_2)(x) := \int_{-\infty}^{\infty} f_1(x') f_2(x - x') dx'. \quad (7.11)$$

De las definiciones (7.1) es fácil mostrar que

$$\mathcal{F}_{x \rightarrow k}[(f_1 * f_2)(x)] = \sqrt{2\pi} \tilde{f}_1(k) \tilde{f}_2(k) \quad (7.12)$$

y, análogamente,

$$\mathcal{F}_{x \rightarrow k}[f_1(x) f_2(x)] = \frac{1}{\sqrt{2\pi}} (\tilde{f}_1 * \tilde{f}_2)(k). \quad (7.13)$$

Ejercicio 7.1.10. Demuestre los resultados de arriba.

7.1.3. Espectro de potencia

Supongamos que $f(t)$ es una *señal* (por ejemplo, el campo eléctrico de una onda, la amplitud de oscilación de un oscilador, la corriente en un circuito) y que, como función de t , es \mathcal{L}^2 . La *energía* contenida en la señal será proporcional a $\int_{-\infty}^{\infty} |f(t)|^2 dt$ (por ejemplo, si $f(t)$ es la corriente a través de una resistencia R , la potencia disipada será $R|f(t)|^2$).

Usando (7.1b), la energía total (de comienzo a fin) de la señal será entonces proporcional a

$$\begin{aligned} E &:= \int_{-\infty}^{\infty} |f(t)|^2 dt = \int_{-\infty}^{\infty} \left| \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \tilde{f}(\omega) e^{i\omega x} d\omega \right|^2 dt \\ &= \frac{1}{2\pi} \int_{-\infty}^{\infty} \left(\int_{-\infty}^{\infty} \tilde{f}(\omega) e^{i\omega x} d\omega \right)^* \left(\int_{-\infty}^{\infty} \tilde{f}(\omega') e^{i\omega' x} d\omega' \right) dt \\ &= \int_{-\infty}^{\infty} d\omega \tilde{f}^*(\omega) \int_{-\infty}^{\infty} d\omega' \tilde{f}(\omega') \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-i(\omega-\omega')t} dt, \end{aligned}$$

de donde, usando (7.2), es inmediato que

$$\int_{-\infty}^{\infty} |f(t)|^2 dt = \int_{-\infty}^{\infty} |\tilde{f}(\omega)|^2 d\omega. \quad (7.14a)$$

Esta expresión es la versión “continua” de la relación de Parseval (4.38), y en el presente contexto se la conoce como *Teorema de Plancherel* (o también como Teorema de Parseval).

Ejercicio 7.1.11. Usando el argumento de arriba, muestre que

$$\int_{-\infty}^{\infty} g^*(t)f(t) dt = \int_{-\infty}^{\infty} \tilde{g}^*(\omega)\tilde{f}(\omega) d\omega \quad (7.14b)$$

para $f, g \in \mathcal{L}^2(-\infty, \infty)$ (es decir, $\langle g, f \rangle \equiv \langle \tilde{g}, \tilde{f} \rangle$).

En (7.14a) podemos ver que, al igual que la energía de la señal contenida en el intervalo temporal $[t, t+dt]$ es $|f(t)|^2 dt$, la energía contenida en el intervalo de frecuencias $[\omega, \omega+d\omega]$ es $|\tilde{f}(\omega)|^2 d\omega$, lo que nos lleva a definir la *densidad espectral de energía*

$$\bar{S}(\omega) := |\tilde{f}(\omega)|^2 \equiv \mathcal{F}_{t \rightarrow \omega}[(f^* * f)(t)]. \quad (7.15a)$$

Si recordamos la extensión (7.1c) de la definición de \mathcal{F} a distribuciones, podríamos pensar que ambos resultados (7.14) deberían extenderse tal cual a distribuciones temperadas; y en el caso en que $f(t)$ es una distribución temperada y $g(t)$ es una función de prueba de Schwartz, efectivamente (7.14b) aún sería válido. Sin embargo, aún así es de esperar que la energía *total* de, por ejemplo, $f(t) = \cos(\omega_0 t)$, diverja. En casos como ese tiene más sentido definir una *potencia media* de $f(t)$ como

$$P := \lim_{T \rightarrow \infty} \frac{1}{T} \int_{t_0 - \frac{T}{2}}^{t_0 + \frac{T}{2}} |f(t)|^2 dt,$$

donde t_0 es, en principio, arbitrario, y lo asumiremos nulo.² La formalización de esta idea pasa por definir

$$f_T(t) := f(t) \operatorname{rect}\left(\frac{t}{T}\right),$$

usando la función rect definida en el ejemplo 7.1.1, lo que trunca la señal al intervalo $(-\frac{T}{2}, \frac{T}{2})$ y permite escribir³

$$P := \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-\infty}^{\infty} |f_T(t)|^2 dt.$$

²Esta es en realidad una *hipótesis de estacionariedad* de la señal $f(t)$, en el sentido que el límite se asume independiente de t_0 .

³Este procedimiento, conocido como *windowing*, es común en el análisis de señales; puede realizarse con otras “ventanas” diferentes de la “ventana rectangular” rect.

Notando que

$$\mathcal{F}_{t \rightarrow \omega}[f_T(t)] = \tilde{x}(\omega) * \frac{T}{\sqrt{2\pi}} \operatorname{sinc}\left(\frac{\omega T}{2}\right),$$

un cálculo directo pero bastante largo y engorroso permite mostrar que, en este caso, debemos definir una *densidad espectral de potencia* como

$$S(\omega) = \lim_{T \rightarrow \infty} \frac{1}{T} |\tilde{f}_T(\omega)|^2 \equiv \lim_{T \rightarrow \infty} \frac{1}{T} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \mathcal{F}_{t \rightarrow \omega}[(f_T^* * f_T)(t)] dt \equiv \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \mathcal{F}_{\tau \rightarrow \omega}[\kappa(\tau)] d\tau, \quad (7.15b)$$

que es la *potencia* de la señal contenida en el intervalo de frecuencias $[\omega, \omega + d\omega]$, y donde

$$\kappa(\tau) := \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-\infty}^{\infty} f_T^*(t - \tau) f_T(t) dt \equiv \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} f^*(t - \tau) f(t) dt \quad (7.16)$$

es la *autocorrelación* de la señal $f(t)$. Si bien las expresiones intermedias en (7.15b) son correctas, pueden ser complicadas o incluso difíciles de interpretar en algunos casos (véase el ejemplo de aplicación 7.1.3 más adelante); en ellos el cálculo de $S(\omega)$ como transformada de Fourier de $\kappa(\tau)$ es más claro y rápido.

7.1.4. Ejemplos de aplicación

En esta sección vamos a desarrollar unos pocos ejemplos que ilustren la aplicación de la transformada de Fourier a problemas sencillos. Otros ejemplos más complicados serán desarrollados en el Capítulo 10, donde aplicaremos transformadas integrales a la resolución de ecuaciones diferenciales parciales.

Ejemplo 7.1.3. Consideremos la ecuación de movimiento de un oscilador armónico libre y sin amortiguación,

$$\ddot{x} + \omega_0^2 x = 0.$$

Aplicando la transformada de Fourier $\mathcal{F}_{t \rightarrow \omega}$, obtenemos

$$(-\omega^2 + \omega_0^2) \tilde{x}(\omega) = 0.$$

La solución de esta ecuación algebraica, limitada a *funciones*, sería $\tilde{x}(\omega) \equiv 0$, lo que evidentemente es erróneo ya que sabemos que la ecuación original tiene soluciones no triviales osculatorias. Sin embargo, si factorizamos el coeficiente del lado izquierdo para obtener

$$(\omega + \omega_0)(\omega - \omega_0) \tilde{x}(\omega) = 0$$

y recordamos que $x\delta(x)$ es la distribución nula, vemos que si admitimos *distribuciones* como soluciones tendremos

$$\tilde{x}(\omega) = a\delta(\omega + \omega_0) + b\delta(\omega - \omega_0).$$

Tomando ahora la antittransformada $\mathcal{F}_{\omega \rightarrow t}^{-1}$ resulta

$$x(t) = ae^{-i\omega_0 t} + be^{i\omega_0 t},$$

que podemos reescribir como

$$x(t) = A \sin(\omega_0 t) + B \cos(\omega_0 t).$$

Debe advertirse que el hecho de que la solución en sentido distribucional es indispensable para recuperar correctamente soluciones de una ecuación *homogénea*, es muchas veces omitido en la literatura.

Vamos a calcular ahora la densidad espectral de potencia de $x(t)$. Suponiendo por simplicidad que $B = 0$, notemos que en ese caso

$$\tilde{x}(\omega) = \frac{A\sqrt{2\pi}}{2i} [\delta(\omega - \omega_0) - \delta(\omega + \omega_0)],$$

con lo cual realizar el cálculo mediante las expresiones intermedias de (7.15b) sería complicado. Comencemos entonces por calcular la autocorrelación

$$\begin{aligned} \kappa(\tau) &= A^2 \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T/2}^{T/2} \sin(\omega_0 t - \omega_0 \tau) \sin(\omega_0 t) dt \\ &= \frac{A^2}{4\omega_0} \lim_{T \rightarrow \infty} \frac{1}{T} [\sin(\omega_0 \tau - 2\omega_0 t) + 2\omega_0 t \cos(\omega_0 \tau)]_{-T/2}^{T/2} \\ &= \frac{A^2}{2} \cos(\omega_0 \tau). \end{aligned}$$

Luego

$$S(\omega) = \frac{A^2}{2} \frac{\delta(\omega - \omega_0) + \delta(\omega + \omega_0)}{2}.$$

Notemos que la densidad de potencia tiene dos picos simétricos en $\omega = \pm\omega_0$, cada uno con la mitad de la potencia media $A^2/2$ de la señal sinusoidal $A \sin(\omega_0 t)$, de modo que la potencia total resulta $\int_{-\infty}^{\infty} S(\omega) d\omega = A^2/2$, como era de esperar.

Ejemplo 7.1.4. Consideremos un oscilador armónico forzado y (sub)amortiguado, descripto por la ecuación

$$\ddot{x} + 2\gamma\dot{x} + \omega_0^2 x = a \sin(\omega_f t),$$

donde $x(t)$ es la amplitud de oscilación, $2\gamma > 0$ es la constante de amortiguación, $\omega_0 > \gamma$ es la *frecuencia natural* del oscilador, a es la amplitud del forzamiento y ω_f es su frecuencia. Aplicando la transformada de Fourier $\mathcal{F}_{t \rightarrow \omega}$ a ambos miembros de la ecuación, obtenemos

$$(-\omega^2 + 2i\omega\gamma + \omega_0^2) \tilde{x}(\omega) = \frac{a\sqrt{2\pi}}{2i} [\delta(\omega - \omega_f) - \delta(\omega + \omega_f)].$$

Suponiendo que nos interesa el *estado de régimen* $x_r(t)$ del oscilador, es decir la solución de la ecuación a tiempos muy posteriores al inicial, las soluciones de la

ecuación homogénea, que decaen exponencialmente en el tiempo como $e^{-\gamma t}$, ya habrán desaparecido, independientemente de sus amplitudes iniciales; ignoramos entonces las posibles soluciones de

$$(-\omega^2 + 2i\omega\gamma + \omega_0^2) \tilde{x}(\omega) = 0,$$

que serían

$$\tilde{x}_h(\omega) = \delta \left(i\gamma \pm \sqrt{\omega_0^2 - \gamma^2} \right),$$

y simplemente despejamos $\tilde{x}_r(\omega)$ como

$$\tilde{x}_r(\omega) = \frac{a\sqrt{2\pi}}{2i} \frac{\delta(\omega - \omega_f) - \delta(\omega + \omega_f)}{\omega_0^2 - \omega^2 + 2i\omega\gamma}.$$

Tomando ahora la antitransformada $\mathcal{F}^{-1}_{\omega \rightarrow t}$ de ambos miembros, tenemos

$$x_r(t) = a \frac{(\omega_0^2 - \omega_f^2) \sin(\omega_f t) - 2\gamma\omega_f \cos(\omega_f t)}{(\omega_0^2 - \omega_f^2)^2 + 4\gamma^2\omega_f^2}.$$

Ejercicio 7.1.12. Complete los cálculos del ejemplo precedente, y muestre que

$$S(\omega) = \frac{a^2}{2} \frac{1}{(\omega_0^2 - \omega_f^2) + 2\gamma\omega_f} \frac{\delta(\omega - \omega_f) + \delta(\omega + \omega_f)}{2}.$$

Ejemplo 7.1.5. Consideremos un “paquete de ondas Gaussiano”

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma^2} \exp\left(-\frac{(x-x_0)^2}{2\sigma^2}\right) e^{ik_0x},$$

es decir, una “onda plana” e^{ik_0x} de amplitud modulada por una envolvente Gaussiana (normalizada) de semiancho σ centrada en x_0 . Su transformada de Fourier será

$$\tilde{f}(k) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{(k-k_0)^2}{2/\sigma^2}\right) e^{-i(k-k_0)x_0},$$

es decir, otro “paquete Gaussiano”: una “onda plana” $e^{-i(k-k_0)x_0}$ de amplitud modulada por una Gaussiana (pero *no* normalizada) de semiancho $1/\sigma$. Notemos que los anchos de los paquetes en coordenadas x y en número de onda k son *inversos*: un paquete muy localizado en posición (σ pequeño) tendrá una gran dispersión en número de onda ($1/\sigma$ grande), y viceversa. En Mecánica Cuántica este hecho cobra una importancia fundamental como el *principio de incertidumbre*: un estado con posición x bien definida tendrá un momento $p = \hbar k$ muy mal definido y viceversa, con $\Delta x \Delta p \sim \hbar$.

Ejercicio 7.1.13. Complete los cálculos del ejemplo precedente.

Ejercicio 7.1.14. Halle la densidad espectral de energía de la señal

$$f(t) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(t-t_0)^2}{2\sigma^2}\right) e^{i\omega_0 t},$$

y muestre que es independiente de t_0 .

Ejemplo 7.1.6. En Mecánica Cuántica, la *función de onda* $f(\vec{x})$ de una partícula es una función compleja tal que $|f(\vec{x})|^2$ es la densidad de probabilidad para la posición \vec{x} de la misma. Una particularmente simple es

$$f(\vec{x}) = \left(\frac{2}{\pi a^2}\right)^{3/4} e^{-r^2/a^2} = N e^{-r^2/a^2},$$

que corresponde a una densidad de probabilidad Gaussiana normalizada en tres dimensiones. De acuerdo a (7.3a) su transformada de Fourier es

$$\tilde{f}(\vec{k}) = \frac{1}{(2\pi)^{3/2}} \iiint_{-\infty}^{\infty} f(\vec{x}) e^{-i\vec{k}\cdot\vec{x}} d^3x.$$

Notemos que, cualquiera sea \vec{k} , siempre podemos reorientar nuestro sistema de coordenadas xyz para que el eje z coincida con la dirección de \vec{k} , sin afectar el resultado de la integración. Una vez hecho esto, tendremos

$$\begin{aligned} \tilde{f}(\vec{k}) &= \frac{1}{(2\pi)^{3/2}} \int_0^\infty dr r^2 \int_0^\pi d\theta \sin\theta \int_0^{2\pi} d\phi f(\vec{x}) e^{-ikr \cos\theta} \\ &= \frac{N}{2\pi} \int_0^\infty dr r^2 e^{-r^2/a^2} \int_0^\pi e^{-ikr \cos\theta} \sin\theta d\theta. \end{aligned}$$

Cambiando variables a $u = \cos\theta$, el resto de las integraciones es bastante directo y lleva a

$$\tilde{f}(\vec{k}) = (2\pi a^2)^{3/4} e^{-k^2 a^2/4}.$$

Nótese que, como en una dimensión, la transformada de Fourier de una Gaussiana en \vec{x} es una Gaussiana en \vec{k} , y sus anchos son inversamente proporcionales.

Ejercicio 7.1.15. Complete los cálculos del ejemplo anterior.

Ejemplo 7.1.7. En muchas ocasiones deberemos invertir, como parte de la solución de un problema, transformadas de Fourier de funciones con denominadores polinomiales. Un ejemplo ilustrativo es la inversión de

$$\tilde{f}(\omega) = \frac{a}{\omega^2 + a^2} = \frac{a}{(\omega + ia)(\omega - ia)}, \quad a, b > 0.$$

Dada (7.1b), es evidente que la forma más económica de evaluar la integral de inversión es por el método de los residuos, como

$$\begin{aligned} f(t) &= \frac{1}{\sqrt{2\pi}} \lim_{R \rightarrow \infty} \left(\int_{-R}^R \tilde{f}(\omega) e^{i\omega t} d\omega + \int_{\gamma_\pm} \tilde{f}(z) e^{izt} dz \right) \\ &= \frac{2\pi i}{\sqrt{2\pi}} \operatorname{Res}_{z=\pm ia} \left\{ \tilde{f}(z) e^{izt} \right\}, \end{aligned}$$

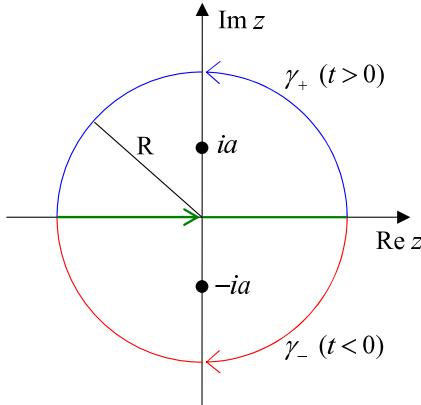


Figura 7.1: Contornos de integración para la inversión de $\tilde{f}(t)$. En verde la integral real; en rojo el semicírculo γ_- para $t < 0$; en azul el semicírculo γ_+ para $t > 0$.

donde γ_+ es un semicírculo de radio R en el semiplano superior $\text{Im } z > 0$ y γ_- es un semicírculo de radio R en el semiplano inferior, como se muestra en la Figura 7.1. La elección de cuál semicírculo usar para cerrar el contorno debe hacerse en base a la convergencia de $\int_{\gamma_\pm} \tilde{f}(z)e^{izt} dz$: dado que $\tilde{f}(z)$ es analítica aparte de sus polos simples, y que parametrizando con $z = Re^{i\theta}$ tendremos

$$e^{izt} = \exp(-Rt \sin \theta + iRt \cos \theta),$$

vemos que debemos elegir el rango de θ de modo que

$$t \sin \theta > 0,$$

ya que con esa elección

$$\lim_{R \rightarrow \infty} \int_{\gamma_\pm} \tilde{f}(z)e^{izt} dz = 0.$$

Luego, para $t < 0$ elegimos $-\pi < \theta < 0$ (cerramos en el semiplano inferior) de modo que $\sin \theta < 0$, y tenemos

$$f(t) = -\frac{2\pi i}{\sqrt{2\pi}} \underset{z=-ia}{\text{Res}} \left\{ \tilde{f}(z)e^{izt} \right\} = -\frac{2\pi i}{\sqrt{2\pi}} \frac{ae^{at}}{-2ia}, \quad t < 0$$

donde el signo negativo se debe a que recorremos el contorno de integración en sentido negativo (horario); mientras que para $t > 0$ elegimos $0 < \theta < \pi$ (cerramos en el semiplano superior) de modo que $\sin \theta > 0$, y tenemos

$$f(t) = \frac{2\pi i}{\sqrt{2\pi}} \underset{z=ia}{\text{Res}} \left\{ \tilde{f}(z)e^{izt} \right\} = \frac{2\pi i}{\sqrt{2\pi}} \frac{ae^{-at}}{2ia}, \quad t > 0.$$

Juntando los dos resultados, tenemos

$$f(t) = \sqrt{\frac{\pi}{2}} e^{-at},$$

válida para todo t .

7.2. La transformada de Laplace

En la Sección 7.1.4 hemos visto la utilidad de la transformada de Fourier para resolver ecuaciones diferenciales ordinarias. Sin embargo también han quedado de manifiesto ciertas limitaciones en este aspecto; en particular, las ecuaciones a resolver deben estar definidas sobre un dominio *infinito*. Si la variable independiente representa una coordenada (o varias), esto no resulta molesto, ya que podemos imponer las CC necesarias; pero si representa el tiempo, esto nos limita a la búsqueda de *soluciones de régimen*, que no admiten CI, como en los ejemplos de aplicación 7.1.3 y 7.1.4. ¿Cómo aplicar entonces la transformada a la resolución de EDOs con CI a un tiempo finito?

Una posible solución es restringir el dominio de las funciones a tiempos mayores que un dado tiempo inicial t_0 , que siempre podemos elegir como cero; por ejemplo si nos interesa $f(t)$ para $t > 0$ podemos buscar la transformada de Fourier de $\Theta(t)f(t)$, con Θ la función escalón de Heaviside, con la esperanza de aplicar el resultado (7.13) y escribir $\mathcal{F}_{t \rightarrow \omega}[\Theta(t)f(t)]$ como la convolución de $\tilde{\Theta}(\omega)$ con $\tilde{f}(\omega)$. Este atractivo enfoque adolece, sin embargo, de un defecto: ¡la transformada de Fourier de $\Theta(t)$ sólo está definida en sentido distribucional! Para verlo, recordemos del Ejemplo 4.1.12 que $\Theta'(t) = \delta(t)$; luego, usando (7.6),

$$\tilde{\Theta}(\omega) = \frac{1}{\sqrt{2\pi i\omega}} + c\delta(\omega),$$

con c una constante que debe elegirse para que la antittransformada de $\tilde{\Theta}(\omega)$ sea efectivamente $\Theta(t)$, es decir para que

$$\Theta(t) = \frac{1}{2\pi i} \int_{-\infty}^{\infty} \frac{e^{i\omega t}}{\omega} d\omega + \frac{c}{\sqrt{2\pi}}.$$

Ahora bien, es directo verificar que

$$\frac{1}{2\pi i} \int_{-\infty}^{\infty} \frac{e^{i\omega t}}{\omega} d\omega = \frac{1}{\pi} \int_0^{\infty} \frac{\sin(\omega t)}{\omega} d\omega \times \begin{cases} 1, & t > 0, \\ -1, & t < 0, \end{cases}$$

y que

$$\int_0^{\infty} \frac{\sin \omega}{\omega} d\omega = \frac{\pi}{2},$$

de modo que debemos tomar $c = \sqrt{2\pi}/2$ y, en definitiva,

$$\tilde{\Theta}(\omega) = \frac{1}{\sqrt{2\pi}} \left[\frac{1}{i\omega} + \pi\delta(\omega) \right], \quad (7.17)$$

que es una distribución singular, donde el término $1/i\omega$ deberá ser tratado como un valor principal de Cauchy cuando aparezca en integrales, como vimos en el Ejemplo 4.1.11. Si bien podemos seguir por este camino, es claro que no es exactamente lo que estamos buscando.

Pese a lo anterior, la idea de multiplicar $f(t)$ por $\Theta(t)$ sigue siendo atractiva; es válido entonces preguntarse qué ocurre si además multiplicamos, por ejemplo, por e^{-ct} con $c > 0$ (un *factor de convergencia*), ya que en ese caso la integral

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \Theta(t) e^{-ct} e^{-i\omega t} dt = \frac{1}{\sqrt{2\pi}} \int_0^{\infty} e^{-(c+i\omega)t} dt = \frac{1}{\sqrt{2\pi}} \frac{1}{c+i\omega}$$

está bien definida, y

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \Theta(t) e^{-ct} f(t) e^{-i\omega t} dt \quad (7.18)$$

convergerá incluso si $f(t) \underset{t \rightarrow \infty}{\asymp} e^{\alpha t}$ con $\alpha > 0$, siempre que $c > \alpha$.⁴ Nuestro par transformada-antitransformada se escribiría entonces

$$\begin{aligned} g(\omega) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(t) e^{-ct} \Theta(t) e^{-i\omega t} dt, \\ f(t) e^{-ct} \Theta(t) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} g(\omega) e^{i\omega t} d\omega. \end{aligned}$$

7.2.1. Definición de la transformada de Laplace

Vamos a definir primero una nueva variable,

$$s := c + i\omega.$$

En términos de s ,

$$\sqrt{2\pi} \int_{-\infty}^{\infty} \Theta(t) e^{-ct} f(t) e^{-i\omega t} dt = \int_0^{\infty} f(t) e^{-ct} e^{-i\omega t} dt = \int_0^{\infty} f(t) e^{-st} dt = \bar{f}(s).$$

Invirtiendo la transformada de Fourier,

$$\Theta(t) e^{-ct} f(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \bar{f}(s) e^{i\omega t} d\omega = \frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} \bar{f}(s) e^{st} e^{-ct} ds,$$

donde la integración se realiza a lo largo de la recta $\text{Re } s = c$ en el plano complejo s , como se muestra en la Figura 7.2. Notando que el factor e^{-ct} puede extraerse de la última integral y simplificarse con el del lado izquierdo, definimos

$$\mathcal{L}_{t \rightarrow s}[f(t)] = \tilde{f}(s) := \int_0^{\infty} f(t) e^{-st} dt, \quad (7.19a)$$

$$\mathcal{L}_{s \rightarrow t}^{-1}[\tilde{f}(s)] = f(t) \Theta(t) := \frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} \tilde{f}(s) e^{st} ds, \quad (7.19b)$$

⁴La divergencia *polinomial* de $f(t)$ para $t \rightarrow \infty$ queda cubierta si entendemos a f como una distribución temperada; la divergencia exponencial, no. La multiplicación por el factor de convergencia hace tratables funciones peor comportadas que un polinomio.

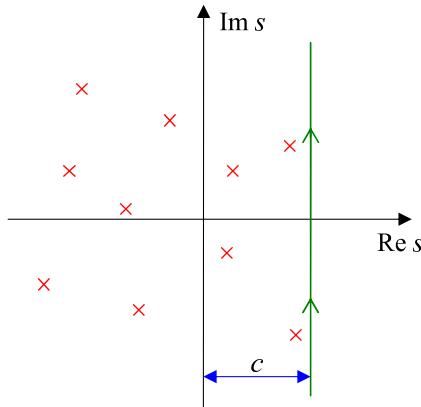


Figura 7.2: En verde, el camino de integración para la inversión de la transformada de Laplace según (7.19b). En rojo los polos de $\tilde{f}(s)$, todos los cuales deben tener parte real $< c$.

donde $\tilde{f}(s)$ se conoce como la *transformada de Laplace* de $f(t)$, y $f(t)$ se dice la *antitransformada de Laplace* de $\tilde{f}(s)$. Notemos que, dado que

$$\int_0^T e^{at} e^{-st} dt = \frac{1}{a-s} (e^{(a-s)T} - 1),$$

la convergencia de (7.19a) requiere que $\operatorname{Re} s > \operatorname{Re} a$, en cuyo caso tendremos

$$\int_0^\infty e^{at} e^{-st} dt = \frac{1}{s-a}.$$

El camino de integración en (7.19b), entonces, debe yacer *a la derecha de todos los polos de $\tilde{f}(s)$* , según lo muestra la Figura 7.2.

Consideraremos estas definiciones válidas para toda función $f(t)$ cuyo crecimiento para $t \rightarrow \infty$ pueda acotarse por una exponencial, es decir tal que

$$\exists M < \infty / \lim_{t \rightarrow \infty} |f(t)e^{-Mt}| = 0,$$

y en ese caso tendremos $c \geq M$ en (7.19b). Esto incluye funciones en $\mathcal{L}^1(0, \infty)$ y en $\mathcal{L}^2(0, \infty)$; la extensión de estas definiciones a distribuciones temperadas es inmediata, siguiendo la línea de lo realizado para arribar a (7.1c). La elección de dónde colocar el factor $1/2\pi$ en (7.19) es la seguida siempre en la literatura.

7.2.2. Propiedades de la transformada de Laplace

A partir de las definiciones (7.19) podemos deducir una serie de propiedades útiles de la transformada de Laplace, algunas de las cuales presentaremos a continuación.

Linealidad.

La transformada y antitransformada de Laplace son, evidentemente, lineales:

$$\mathcal{L}[\alpha f + \beta g] = \alpha \mathcal{L}[f] + \beta \mathcal{L}[g], \quad (7.20a)$$

$$\mathcal{L}^{-1}[\alpha f + \beta g] = \alpha \mathcal{L}^{-1}[f] + \beta \mathcal{L}^{-1}[g]. \quad (7.20b)$$

Derivación.

A partir de la definición (7.19a) es inmediato mostrar, integrando por partes, que para cualquier función f de crecimiento acotado por una exponencial,

$$\begin{aligned} {}_{t \rightarrow s} \mathcal{L}[f'(t)] &= s {}_{t \rightarrow s} \mathcal{L}[f(t)] - f(0), \\ {}_{t \rightarrow s} \mathcal{L}[f''(t)] &= s^2 {}_{t \rightarrow s} \mathcal{L}[f(t)] - sf(0) - f'(0), \\ {}_{t \rightarrow s} \mathcal{L}[f^{(n)}(t)] &= s^n {}_{t \rightarrow s} \mathcal{L}[f(t)] - s^{n-1}f(0) - s^{n-2}f'(0) - \cdots - f^{(n-1)}(0), \end{aligned} \quad (7.21)$$

donde $f(0)$ debe entenderse como $f(0_+) = \lim_{t \downarrow 0} f(t)$. La extensión a distribuciones es inmediata.

Ejercicio 7.2.1. Complete los cálculos precedentes.

Integración.

De la definición (7.19a) es inmediato mostrar que

$${}_{t \rightarrow s} \mathcal{L} \left[\int_0^t f(t') dt' \right] = \frac{1}{s} {}_{t \rightarrow s} \mathcal{L}[f(s)]. \quad (7.22)$$

Ejercicio 7.2.2. Complete los cálculos que llevan al resultado anterior.

Traslación en t .

De la definición (7.19), notando que la integral no incluye valores de $f(t)$ para $t < 0$, tenemos

$${}_{t \rightarrow s} \mathcal{L}[f(t+a)] = \begin{cases} e^{as} \left({}_{t \rightarrow s} \mathcal{L}[f(t)] - \int_0^a f(t)e^{-st} dt \right), & a > 0, \\ e^{as} {}_{t \rightarrow s} \mathcal{L}[f(t)], & a < 0. \end{cases} \quad (7.23)$$

Ejercicio 7.2.3. Verifique el resultado precedente.

Traslación en s .

De la definición (7.19) también resulta

$${}_{t \rightarrow s} \mathcal{L}[e^{\alpha t} f(t)] = {}_{t \rightarrow s} \mathcal{L}[f(t)](s - \alpha) = \tilde{f}(s - \alpha). \quad (7.24)$$

Ejercicio 7.2.4. Demuestre el resultado precedente.

Escaleo

Otro resultado inmediato de la definición (7.19) es

$$\mathcal{L}_{t \rightarrow s}[f(at)] = \frac{1}{a} \mathcal{L}_{t \rightarrow s}[f(s)](s/a) = \frac{1}{a} \tilde{f}\left(\frac{s}{a}\right), \quad a > 0. \quad (7.25)$$

Ejercicio 7.2.5. Demuestre el resultado precedente.

Multiplicación por potencias.

De la definición (7.19) tendremos

$$\mathcal{L}_{t \rightarrow s}[t^n f(t)] = (-1)^n \frac{d^n}{ds^n} \mathcal{L}_{t \rightarrow s}[f(t)]. \quad (7.26)$$

Ejercicio 7.2.6. Demuestre el resultado anterior.

Convolución

Dadas dos funciones $f_1(t)$ y $f_2(t)$, definimos su *convolución* $f_1 * f_2$ como

$$(f_1 * f_2)(t) := \int_0^t f_1(t') f_2(t - t') dt'. \quad (7.27)$$

De la definición (7.19) es fácil mostrar que

$$\mathcal{L}_{t \rightarrow s}[(f_1 * f_2)(t)] = \tilde{f}_1(s) \tilde{f}_2(s). \quad (7.28)$$

También puede mostrarse que, si al antitransformar $\tilde{f}_1(s)$ debemos tomar $\operatorname{Re} s > c_1$, y al antitransformar $\tilde{f}_2(s)$ debemos tomar $\operatorname{Re} s > c_2$, entonces

$$\mathcal{L}_{t \rightarrow s}[f_1(t) f_2(t)] = \frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} \tilde{f}_1(s') \tilde{f}_2(s - s') ds', \quad \operatorname{Re} s - c_2 > c > c_1. \quad (7.29)$$

7.2.3. Ejemplos de aplicación

En esta sección vamos a desarrollar unos pocos ejemplos que ilustren la aplicación de la transformada de Laplace a problemas sencillos. Otros ejemplos más complicados serán desarrollados en el Capítulo 10, donde aplicaremos transformadas integrales a la resolución de ecuaciones diferenciales parciales.

Ejemplo 7.2.1. Consideremos un oscilador armónico inicialmente en reposo en su punto de equilibrio, al que se le aplica a partir de $t = 0$ una fuerza constante e igual a la unidad; el problema se escribe

$$\begin{aligned} \ddot{x} + \omega^2 x &= 1, \quad 0 < t < \infty, \\ x(0) &= x'(0) = 0. \end{aligned}$$

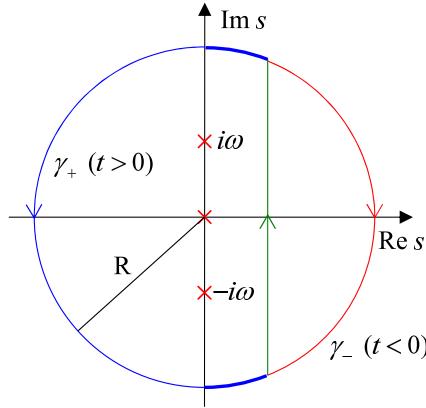


Figura 7.3: Contorno de integración para la inversión de la transformada de Laplace en el Ejemplo 7.2.1. Las cruces rojas indican los polos de $\tilde{x}(s)$; en verde la recta $\text{Re } s = c > 0$; en rojo el arco γ_- para $t < 0$; en azul el arco γ_+ para $t > 0$, y en trazo grueso los arcos cuya contribución se anula para $R \rightarrow \infty$.

Aplicando la transformada de Laplace y usando las CI, obtenemos

$$(s^2 + \omega^2)\tilde{x}(s) = \frac{1}{s},$$

luego

$$\tilde{x}(s) = \frac{1}{s(s + i\omega)(s - i\omega)}.$$

Para completar la solución del problema debemos antitransformar, y para ello lo más práctico es apelar al teorema de los residuos. Para eso necesitamos cerrar el camino de integración, lo que haremos con un arco de circunferencia de radio $R \rightarrow \infty$ como se muestra en la Figura 7.3. Tendremos entonces

$$x(t) = \lim_{R \rightarrow \infty} \frac{1}{2\pi i} \left(\int_{c-iR}^{c+iR} \frac{e^{st} ds}{s(s + i\omega)(s - i\omega)} + \int_{\gamma_\pm} \frac{e^{st} ds}{s(s + i\omega)(s - i\omega)} \right)$$

donde deberemos elegir cuál arco de circunferencia usar y, dada la posición de los tres polos simples del integrando, basta tomar $c > 0$ pero tan pequeño como queramos. Notando que debemos tener $x(t) = 0$ para $t < 0$, vemos que en ese caso debemos cerrar el contorno de integración usando el arco γ_- de la derecha, con lo cual el contorno no encierra ningún polo y la integral se anula; mientras que para $t > 0$ deberemos usar el arco γ_+ de la izquierda, donde es fácil comprobar que la integral sobre γ_+ converge a cero para $R \rightarrow \infty$. Tendremos así

$$x(t) = \frac{1}{\omega^2} - \frac{e^{i\omega t}}{2\omega^2} - \frac{e^{-i\omega t}}{2\omega^2} = \frac{1 - \cos(\omega t)}{\omega^2}.$$

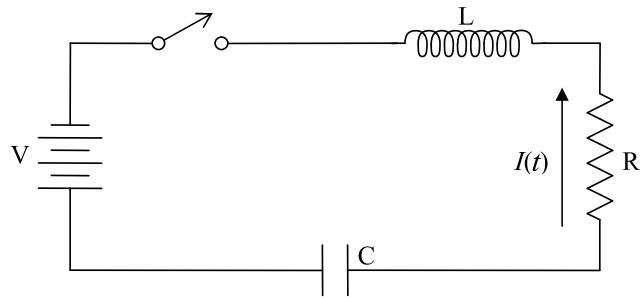


Figura 7.4: El circuito RLC serie del Ejemplo 7.2.2. Todas las corrientes y cargas son nulas para $t < 0$. La llave se cierra en $t = 0$.

Ejercicio 7.2.7. Complete los cálculos del ejemplo previo.

Aprovechando el ejemplo anterior, notemos que si hubiéramos tenido polos con parte real positiva el argumento usado para descartar la integral sobre γ_+ no habría perdido validez: aún si $c > 0$ y finito, los arcos “problemáticos” de γ_+ , marcados con línea gruesa en la Figura 7.3, cuya contribución podría diverger cuando $R \rightarrow \infty$, tienen una extensión angular que tiende a cero en ese límite, y es fácil verificar que sus contribuciones a la integral sobre γ_+ tienden a cero.

Ejercicio 7.2.8. Verifique lo afirmado arriba.

Ejemplo 7.2.2. Consideremos el circuito RLC serie mostrado en la Figura 7.4. Todas las corrientes y cargas son nulas para $t < 0$, y la llave se cierra en $t = 0$. En términos de la corriente $I(t)$, el problema se escribe

$$\begin{aligned} RI + L \frac{dI}{dt} + \frac{1}{C} \left(Q(0) + \int_0^t I(t') dt' \right) &= V, \quad 0 < t < \infty \\ I(0) = Q(0) &= 0. \end{aligned}$$

Para obtener una ecuación diferencial para $I(t)$ deberíamos derivar una vez respecto de t ; sin embargo aquí esto no será necesario: aplicando la transformada de Laplace a la EDO y usando las CI y los resultados (7.21) y (7.22) tenemos

$$R\tilde{I}(s) + L[s\tilde{I}(s) - I(0)] + \frac{1}{C} \frac{\tilde{I}(s)}{s} = \frac{V}{s},$$

de donde

$$\tilde{I}(s) = \frac{V}{L} \frac{1}{(s+a)^2 + b^2},$$

con

$$a = \frac{R}{2L}, \quad b = \sqrt{\frac{1}{LC} + \frac{R^2}{4L^2}}.$$

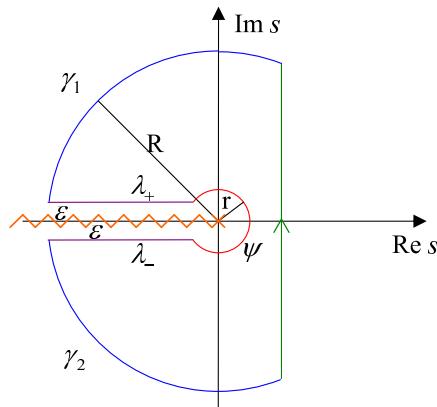


Figura 7.5: Contorno de integración para la inversión de $1/\sqrt{s}$. En verde la recta $\text{Re } s = c > 0$; en azul los arcos γ_1 y γ_2 de radio R ; en violeta los tramos rectos λ_{\pm} a cada lado del corte ramal y a distancia ϵ de él; en rojo el arco ψ de radio r rodeando la singularidad en el origen; en naranja el corte ramal a lo largo del eje real negativo.

Antitrasnformando por integración en el plano complejo s como en el Ejemplo 7.2.1 , o simplemente notando que

$$\mathcal{L}_{t \rightarrow s} [e^{-at} \sin(bt)] = \frac{b}{(s+a)^2 + b^2},$$

obtenemos

$$I(t) = \frac{V}{L} \frac{e^{-at} \sin(bt)}{b}.$$

Esto ilustra también la conveniencia de tener calculadas de antemano las transformadas de Laplace de ciertas funciones de aparición frecuente, o de contar con una tabla de transformadas de Laplace.⁵

Ejercicio 7.2.9. Complete los cálculos del ejemplo anterior.

Ejemplo 7.2.3. En más de una ocasión deberemos antitransformar una función con singularidades esenciales y los consiguientes cortes ramales. Un ejemplo que resulta ilustrativo es

$$\tilde{x}(s) = \frac{1}{\sqrt{s}},$$

⁵Algunos lugares donde encontrarlas son
https://en.wikipedia.org/wiki/List_of_Laplace_transforms
https://tutorial.math.lamar.edu/classes/de/laplace_table.aspx
<https://stanford.edu/~boyd/ee102/laplace-table.pdf>
 Dos buenos calculadores online son
<https://www.wolframalpha.com/input/?i=laplace+transform>
<https://www.symbolab.com/solver/laplace-calculator>

donde la singularidad está en el origen y colocamos el corte ramal sobre el eje real negativo. En ese caso deberemos deformar el arco γ_+ usado en el Ejemplo 7.2.1 para evitar el corte ramal. Como se muestra en la Figura 7.5 , esto se consigue interrumpiendo el arco γ_+ cerca de $\text{Im } s = 0$ para formar un arco superior γ_1 y otro inferior γ_2 , y empalmando estos con dos tramos horizontales λ_{\pm} con $\text{Im } s = \pm\epsilon \rightarrow 0$ y un tramo circular ψ de radio $r \rightarrow 0$ alrededor del origen. En este caso el integrando no tendrá polos dentro del contorno de integración, así que tendremos

$$0 = \left[\lim_{R \rightarrow \infty} \left(\int_{c-iR}^{c+iR} + \int_{\gamma_1} + \int_{\gamma_2} \right) + \lim_{\epsilon \rightarrow 0} \left(\int_{\lambda_+} + \int_{\lambda_-} \right) + \lim_{r \rightarrow 0} \int_{\psi} \right] \frac{e^{st}}{\sqrt{s}} ds.$$

Es bastante directo verificar que \int_{γ_1} e \int_{γ_2} convergen a cero para $R \rightarrow \infty$, como en el Ejemplo 7.2.1; también es directo verificar que \int_{ψ} converge a cero cuando $r \rightarrow 0$; nos queda entonces⁶

$$\begin{aligned} \mathcal{L}_s^{-1}[\tilde{x}(s)] &= \frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} \frac{e^{st}}{\sqrt{s}} ds \\ &= -\frac{1}{2\pi i} \left(\int_{-\infty}^0 |s|^{-1/2} (e^{i\pi})^{-1/2} e^{st} ds + \int_0^{-\infty} |s|^{-1/2} (e^{-i\pi})^{-1/2} e^{st} ds \right) \\ &= \frac{1}{\pi} \int_0^{\infty} s^{-1/2} e^{-st} ds = \frac{1}{\sqrt{\pi t}}. \end{aligned}$$

Ejercicio 7.2.10. Complete los cálculos del último ejemplo.

7.3. Comentarios generales

Repasando los ejemplos de aplicación, podemos destacar un par de características comunes de las transformadas de Fourier y de Laplace. En primer lugar, dado que el kernel de ambas transformadas es autofunción del operador derivada (e^{-ikx} o e^{-iwt} en el caso de Fourier y e^{-st} o e^{-sx} en el de Laplace), la aplicación de la transformada convierte una EDO a coeficientes constantes en una ecuación puramente *algebraica* para la transformada de la función incógnita, que puede resolverse por métodos elementales. En segundo lugar, una vez hallada la transformada de la solución, la inversión o antitransformada se reduce a una integración que, en general, puede realizarse en el plano complejo aplicando el teorema de los residuos, lo que la simplifica notablemente.

Por otra parte, hemos visto que ambas transformadas presentan una diferencia importante: La transformada de Fourier se adapta bien a la resolución de problemas sobre dominios *infinitos*, en particular a hallar soluciones *de régimen* que no necesitan ni admiten CI a tiempo finito. La transformada de Laplace,

⁶Este cálculo puede verse paso a paso en <https://www.youtube.com/watch?v=yY39Xz70Lo>

en cambio, se adapta bien a la resolución de problemas sobre dominios *semi-infinitos*, e incorpora de manera automática las correspondientes CI o CC del problema.

La transformación de una ecuación diferencial en una algebraica es un rasgo distintivo de las transformadas integrales, y no es para nada casual: Sin prender rigurosidad, dado un problema de autovalores

$$LK(x, \lambda) + \lambda K(x, \lambda) = 0, \quad a < x < b,$$

donde L es un operador diferencial (en x) lineal y $K(x, \lambda)$ son las autofunciones de autovalor λ , y éstos forman un conjunto *continuo*

$$\alpha < \lambda < \beta,$$

podemos definir una transformada integral “a medida” mediante el par

$$\begin{aligned} f(x) &= \int_{\alpha}^{\beta} K(x, \lambda) \tilde{f}(\lambda) d\lambda, \\ \tilde{f}(\lambda) &= \int_a^b f(x) K(x, \lambda) dx. \end{aligned}$$

Si ahora queremos resolver el problema

$$Lf(x) + \mu f(x) = g(x), \quad a < x < b,$$

la aplicación de esta transformada lo traducirá en la ecuación algebraica

$$(\mu - \lambda) \tilde{f}(\lambda) = \tilde{g}(\lambda).$$

Esto se debe a que, por la ortogonalidad y completitud asumidas para las autofunciones $K(x, \lambda)$, tendremos⁷

$$\begin{aligned} \int_{\alpha}^{\beta} K(x, \lambda) K(x', \lambda) d\lambda &= \delta(x - x'), \quad a < x, x' < b, \\ \int_a^b K(x, \lambda) K(x, \lambda') dx &= \delta(\lambda - \lambda'), \quad \alpha < \lambda, \lambda' < \beta. \end{aligned}$$

Un ejemplo de lo dicho es la *transformada de Hankel*⁸

$$\begin{aligned} f(x) &= \int_0^{\infty} \tilde{f}_{\nu}(k) J_{\nu}(kx) k dk, \\ \tilde{f}_{\nu}(k) &= \int_0^{\infty} f(x) J_{\nu}(kx) x dx, \end{aligned}$$

definida para $\nu > -\frac{1}{2}$, que mapea

$$\frac{d^2 f}{dx^2} + \frac{1}{x} \frac{df}{dx} - \frac{\nu^2}{x^2} f \quad \longleftrightarrow \quad -k^2 \tilde{f}_{\nu}(k).$$

⁷El lector interesado puede consultar https://en.wikipedia.org/wiki/Integral_transform y las referencias allí citadas.

⁸El lector interesado puede consultar https://en.wikipedia.org/wiki/Hankel_transform.

Capítulo 8

Ecuaciones Diferenciales Parciales

En este capítulo comenzaremos el estudio de las *Ecuaciones Diferenciales Parciales*, o ecuaciones en derivadas parciales, que abreviaremos EDP. Si bien a partir de los capítulos anteriores tenemos ya una “caja de herramientas” bastante completa para proceder a su resolución, es necesario previamente estudiar, al menos someramente, su clasificación y sus posibles CI y CC, presentar algunas definiciones básicas y discutir ciertas cuestiones generales. Ello será el objetivo del presente capítulo.

La forma más general de una EDP es una relación funcional

$$F(x, y, \dots; u; u_x, u_y, \dots; u_{xx}, u_{xy}, \dots; \dots) = 0$$

entre una función u , la *variable dependiente*, las coordenadas o *variables independientes* x, y , etc. de las que depende, y sus *derivadas parciales*

$$u_x = \frac{\partial u}{\partial x}, \quad u_y = \frac{\partial u}{\partial y}, \dots, \quad u_{xx} = \frac{\partial^2 u}{\partial x^2}, \quad u_{xy} = \frac{\partial^2 u}{\partial x \partial y}, \dots, \quad \text{etc.}$$

respecto a ellas. Llamaremos a la EDP *homogénea* si $u \equiv 0$ es solución. En general restringiremos nuestro estudio a EDP *lineales*, donde F es una función lineal de u y sus *derivadas*, aunque no necesariamente de las variables independientes; también nos limitaremos generalmente a EDP de *segundo orden*, donde a lo sumo intervienen derivadas segundas de u .

8.1. Clasificación

Por concisión comenzaremos con la clasificación de EDP de la forma

$$A(x, y)u_{xx} + 2B(x, y)u_{xy} + C(x, y)u_{yy} + F(x, y, u, u_x, u_y) = 0 \quad (8.1a)$$

Tipo	D	Sistema	Ejemplo
Hiperbólico	> 0	ondas	$U_{tt} - c^2 \nabla^2 U = 0$
Parabólico	0	calor, difusión sistemas cuánticos	$U_t - \kappa \nabla^2 U = 0$ $i\hbar\psi_t + \frac{\hbar^2}{2m} \nabla^2 \psi + V(\vec{x}, t)\psi = 0$
Elíptico	< 0	fenómenos estacionarios	Laplace: $\nabla^2 \phi = 0$ Helmholtz: $\nabla^2 U + k^2 U = 0$ Poisson: $\nabla^2 \phi = \rho$

Tabla 8.1: Clasificación básica de las EDP lineales de segundo orden.

en sólo dos variables independientes, ya que la fenomenología que encontraremos se extiende directamente a casos con más variables. Nótese que a este fin, en principio sólo suponemos linealidad respecto de las derivadas de orden superior, y F podría ser una función no lineal de sus argumentos.

La clasificación se basará en el signo del *discriminante*

$$D(x, y) := B^2(x, y) - A(x, y)C(x, y), \quad (8.1b)$$

y la nomenclatura y ejemplos notables se muestran en la Tabla 8.1. Es claro que una misma EDP puede tener diferentes tipos en diferentes regiones del plano xy ; cuando ello ocurre, la resolución de la EDP suele complicarse notablemente.

8.1.1. Curvas características

Como paso intermedio y necesario para reducir las EDP de cada tipo a formas más tratables, introduciremos el concepto de curvas características.

Reescribamos por conveniencia las ecuaciones (8.1) como

$$Au_{xx} + 2Bu_{xy} + Cu_{yy} + F(x, y, u, u_x, u_y) = 0, \quad (8.2a)$$

$$D = B^2 - AC. \quad (8.2b)$$

Vamos a introducir una transformación de coordenadas

$$\xi = \phi(x, y), \quad \eta = \psi(x, y), \quad (8.3)$$

la cual supondremos invertible en el dominio de interés, es decir

$$J = \left| \frac{\partial(\xi, \eta)}{\partial(x, y)} \right| \neq 0. \quad (8.4)$$

Tendremos

$$\begin{aligned} u_x &= u_\xi \xi_x + u_\eta \eta_x, \\ u_y &= u_\xi \xi_y + u_\eta \eta_y, \\ u_{xx} &= u_\xi \xi_x^2 + 2u_\xi \xi_x \eta_x + u_\eta \eta_x^2 + u_\xi \xi_{xx} + u_\eta \eta_{xx}, \\ u_{xy} &= u_\xi \xi_x \xi_y + u_\xi (\xi_x \eta_y + \xi_y \eta_x) + u_\eta \eta_x \eta_y + u_\xi \xi_{xy} + u_\eta \eta_{xy}, \\ u_{yy} &= u_\xi \xi_y^2 + 2u_\xi \xi_y \eta_y + u_\eta \eta_y^2 + u_\xi \xi_{yy} + u_\eta \eta_{yy}. \end{aligned}$$

De aquí es inmediato que (8.2a) se reescribe

$$\bar{A}u_{\xi\xi} + 2\bar{B}u_{\xi\eta} + \bar{C}u_{\eta\eta} + \bar{F}(\xi, \eta, u, u_\xi, u_\eta) = 0 \quad (8.5a)$$

con

$$\begin{aligned} \bar{A} &= A\xi_x^2 + 2B\xi_x\xi_y + C\xi_y^2, \\ \bar{B} &= A\xi_x\eta_x + B(\xi_x\eta_y + \xi_y\eta_x) + C\xi_y\eta_y, \\ \bar{C} &= A\eta_x^2 + 2B\eta_x\eta_y + C\eta_y^2, \end{aligned}$$

lo que hace que

$$\bar{D} = \bar{B}^2 - \bar{A}\bar{C} = J^2(B^2 - AC) = J^2D, \quad (8.5b)$$

es decir, mientras la transformación (8.3) no sea singular, no cambia el signo del discriminante ni, por consiguiente, el tipo de la EDP: (8.2a) y (8.5a) tienen el mismo tipo. Debe notarse además que si $F(x, y, u, u_x, u_y)$ es una función *lineal* de u, u_x y u_y , entonces $\bar{F}(\xi, \eta, u, u_\xi, u_\eta)$ será una función lineal de u, u_ξ y u_η , es decir, la transformación (8.3) lleva EDP lineales en EDP lineales.

Ejercicio 8.1.1. Complete los cálculos y verifique las afirmaciones anteriores.

Ahora vamos a aprovechar la invariancia del tipo bajo cambios de coordenadas para reducir las EDP de cada tipo a formas más simples. Primeramente enunciaremos el siguiente¹

Teorema 8.1.1. *La función $z = \varphi(x, y)$ es solución de la EDP*

$$Az_x^2 + 2Bz_x z_y + Cz_y^2 = 0 \quad (8.6a)$$

si, y sólo si, las curvas $\varphi(x, y) = c$ con c constante son soluciones de la EDO

$$Ady^2 - 2Bdxdy + Cdx^2 = 0. \quad (8.6b)$$

A continuación aplicaremos este teorema a los diferentes tipos de EDP, restringiéndolas a lo que se conoce como su *forma canónica*. Las curvas $\varphi(x, y) = c$,

¹La demostración puede consultarse en el texto de A. Tijonov y A. Samarsky, *Ecuaciones de la Física Matemática*, Editorial Mir, Moscú, 1972.

llamadas *curvas características* de la EDP, jugarán un papel fundamental en el proceso; estas serán soluciones de EDOs de la forma

$$\frac{dy}{dx} = \frac{B \pm \sqrt{B^2 - AC}}{A} = \frac{B \pm \sqrt{D}}{A}, \quad (8.7)$$

obtenidas dividiendo la *ecuación característica* (8.6b) por dx^2 . Veremos que las curvas características proveen un sistema de coordenadas (al menos local) donde la EDP adopta su expresión más sencilla, es decir su forma canónica.

8.1.2. La EDP hiperbólica

Para una EDP hiperbólica tenemos $D > 0$, y una EDO de la forma (8.7) tendrá dos familias de soluciones reales $\phi(x, y) = c$ y $\psi(x, y) = c'$, con

$$\left. \frac{dy}{dx} \right|_{\phi(x,y)=c} = \frac{B(x, y) + \sqrt{D(x, y)}}{A(x, y)}, \quad (8.8a)$$

$$\left. \frac{dy}{dx} \right|_{\psi(x,y)=c'} = \frac{B(x, y) - \sqrt{D(x, y)}}{A(x, y)}. \quad (8.8b)$$

Luego, eligiendo $\xi = \phi(x, y)$ tendremos, por el Teorema 8.1.1, que

$$\bar{A} = A\xi_x^2 + 2B\xi_x\xi_y + C\xi_y^2 = 0, \quad (8.9a)$$

y eligiendo $\eta = \psi(x, y)$ tendremos

$$\bar{C} = A\eta_x^2 + 2B\eta_x\eta_y + C\eta_y^2 = 0. \quad (8.9b)$$

Con esta elección de las nuevas coordenadas (ξ, η) , entonces, la EDP (8.2a) se reduce a la forma

$$2\bar{B}(\xi, \eta)u_{\xi\eta} + \bar{F}(\xi, \eta, u, u_\xi, u_\eta) = 0,$$

y como $\bar{D} = \bar{B}^2 > 0$, podemos dividir por $2\bar{B}(\xi, \eta)$ y llevarla a la *primera forma canónica* de las EDP hiperbólicas

$$u_{\xi\eta} + \bar{G}(\xi, \eta, u, u_\xi, u_\eta) = 0, \quad (8.10a)$$

donde

$$\bar{G}(\xi, \eta, u, u_\xi, u_\eta) = \frac{\bar{F}(\xi, \eta, u, u_\xi, u_\eta)}{2\bar{B}(\xi, \eta)}.$$

Alternativamente, podemos definir otras nuevas coordenadas $\alpha = \frac{\xi+\eta}{2}$ y $\beta = \frac{\xi-\eta}{2}$, y obtendremos la *segunda forma canónica* de las EDP hiperbólicas

$$u_{\alpha\alpha} - u_{\beta\beta} + \bar{H}(\alpha, \beta, u, u_\alpha, u_\beta) = 0, \quad (8.10b)$$

donde

$$\bar{H}(\alpha, \beta, u, u_\alpha, u_\beta) = 4\bar{G}(\xi, \eta, u, u_\xi, u_\eta).$$

8.1.3. La EDP parabólica

Para una EDP parabólica tenemos $D = 0$, y una EDO de la forma (8.7) tendrá una única familia de soluciones reales $\phi(x, y) = c$, con

$$\frac{dy}{dx} \Big|_{\phi(x,y)=c} = \frac{B(x, y)}{A(x, y)}. \quad (8.11)$$

Luego, eligiendo $\xi = \phi(x, y)$ tendremos, por el Teorema 8.1.1, que

$$\bar{A} = A\xi_x^2 + 2B\xi_x\xi_y + C\xi_y^2 = 0, \quad (8.12a)$$

y como $\bar{D} = \bar{B}^2 - \bar{A}\bar{C} = \bar{B}^2 = J^2 D = 0$, también tendremos

$$\bar{B} = 0. \quad (8.12b)$$

Para la segunda coordenada nos limitamos a elegir $\eta = \psi(x, y)$ con ψ cualquier función independiente de ϕ , es decir tal que se cumpla (8.4).

Con esta elección de las nuevas coordenadas (ξ, η) , entonces, la EDP (8.2a) se reduce a la forma

$$\bar{C}(\xi, \eta)u_{\eta\eta} + \bar{F}(\xi, \eta, u, u_\xi, u_\eta) = 0,$$

y como $\bar{C} \neq 0$ porque $\psi(x, y)$ no es solución de (8.11), podemos dividir por $\bar{C}(\xi, \eta)$ y llevarla a la *forma canónica* de las EDP parabólicas

$$u_{\eta\eta} + \bar{G}(\xi, \eta, u, u_\xi, u_\eta) = 0, \quad (8.13)$$

donde

$$\bar{G}(\xi, \eta, u, u_\xi, u_\eta) = \frac{\bar{F}(\xi, \eta, u, u_\xi, u_\eta)}{\bar{C}(\xi, \eta)}.$$

8.1.4. La EDP elíptica

Para una EDP elíptica tenemos $D < 0$, y una EDO de la forma (8.7) tendrá dos familias de soluciones complejas conjugadas $\phi(x, y) = c$ y $\phi^*(x, y) = c'$, con

$$\frac{dy}{dx} \Big|_{\phi(x,y)=c} = \frac{B(x, y) + i\sqrt{-D(x, y)}}{A(x, y)}, \quad (8.14a)$$

$$\frac{dy}{dx} \Big|_{\phi^*(x,y)=c'} = \frac{B(x, y) - i\sqrt{-D(x, y)}}{A(x, y)}. \quad (8.14b)$$

Luego, eligiendo $\xi = \phi(x, y)$ y $\eta = \phi^*(x, y)$, usando el Teorema 8.1.1 e introduciendo las coordenadas reales $\alpha = \frac{\xi+\eta}{2}$ y $\beta = \frac{\xi-\eta}{2i}$ tendremos, análogamente a la deducción de (8.10b), la *forma canónica* de las EDP elípticas

$$u_{\alpha\alpha} + u_{\beta\beta} + \bar{H}(\alpha, \beta, u, u_\alpha, u_\beta) = 0. \quad (8.15)$$

8.1.5. La clasificación en más de dos variables

Una EDP de segundo orden en n variables independientes x_1, \dots, x_n , lineal en las derivadas segundas, puede escribirse como

$$A_{ij}(x_1, \dots, x_n)u_{x_i x_j} + F(x_1, \dots, x_n; u; u_{x_1}, \dots, u_{x_n}) = 0. \quad (8.16)$$

Notemos que el operador diferencial $A_{ij}\partial_{x_i}\partial_{x_j}$ es una forma cuadrática en las derivadas parciales ∂_{x_i} y, como tal, *en un punto dado* (pero arbitrario) puede asumirse simétrica, es decir $A_{ij} = A_{ji}$.

Consideremos la introducción de nuevas coordenadas ξ_1, \dots, ξ_n que, al menos en un entorno del punto de interés, satisfagan

$$J = \left| \frac{\partial(\xi_1, \dots, \xi_n)}{\partial(x_1, \dots, x_n)} \right| \neq 0. \quad (8.17)$$

En particular, *en un punto dado* (arbitrario, pero el mismo de antes) podemos considerar una transformación *lineal*

$$\xi_i = a_{ij}x_j \quad (8.18)$$

con a_{ij} constantes, con lo cual en ese punto

$$\partial_{x_i} = a_{ki}\partial_{\xi_k}. \quad (8.19)$$

Bajo esta transformación tendremos, entonces,

$$A_{ij}\partial_{x_i}\partial_{x_j} = a_{ki}a_{lj}A_{ij}\partial_{\xi_k}\partial_{\xi_l} = \bar{A}_{kl}\partial_{\xi_k}\partial_{\xi_l}. \quad (8.20)$$

Ahora bien, la relación

$$\bar{A}_{kl} = a_{ki}a_{lj}A_{ij} \quad (8.21)$$

puede reescribirse matricialmente como

$$\bar{\mathbf{A}} = \mathbf{a}\mathbf{A}\mathbf{a}^T. \quad (8.22)$$

Eligiendo las nuevas coordenadas de modo que $\mathbf{a}^T = \mathbf{a}^{-1}$, que siempre es posible porque asumimos $J \neq 0$, esta no es sino una transformación de semejanza ortogonal aplicada a la matriz simétrica \mathbf{A} , y podemos elegirla para que además la diagonalice.

En definitiva, en un punto cualquiera de \mathbb{R}^n siempre podemos reducir la EDP inicial a la forma

$$\bar{A}_{ii}u_{\xi_i\xi_i} + \bar{F}(\xi_1, \dots, \xi_n; u; u_{\xi_1}, \dots, u_{\xi_n}) = 0, \quad (8.23)$$

con \bar{A}_{ij} diagonal. La clasificación de las EDP se basa, en este caso, en el signo de los autovalores de \mathbf{A} , que serán los elementos diagonales de $\bar{\mathbf{A}}$ en cada punto, y que pueden reducirse a ± 1 y 0 mediante un reescalamiento adicional de las coordenadas. Los tipos más comunes son los siguientes:

Ente matemático	Modela a:
Ecuación diferencial parcial	Dinámica interna del sistema físico
Condiciones iniciales	Preparación del sistema para el experimento
Condiciones de contorno	Influencia del entorno, que no forma parte del sistema

Tabla 8.2: Partes de un sistema físico o experimento modelada por cada parte de un problema matemático EDP+CC+CI.

Hiperbólico: Un coeficiente tiene signo opuesto a los demás, y ninguno es nulo; la ecuación se escribe

$$-u_{\xi_1 \xi_1} + u_{\xi_2 \xi_2} + \cdots + u_{\xi_n \xi_n} + \bar{F} = 0. \quad (8.24a)$$

Parabólico: Uno o más coeficientes se anulan y los demás tienen todos el mismo signo; la ecuación se escribe

$$u_{\xi_1 \xi_1} + u_{\xi_2 \xi_2} + \cdots + u_{\xi_m \xi_m} + \bar{F} = 0, \quad 1 < m < n. \quad (8.24b)$$

(Recuérdese que \bar{F} en general todavía depende de $u_{\xi_{m+1}}, \dots, u_{\xi_n}$.)

Elíptico: Todos los coeficientes tienen el mismo signo y ninguno es nulo; la ecuación se escribe

$$u_{\xi_1 \xi_1} + u_{\xi_2 \xi_2} + \cdots + u_{\xi_n \xi_n} + \bar{F} = 0. \quad (8.24c)$$

Ultrahiperbólico: Algunos coeficientes tienen un signo y los demás el opuesto, y ninguno es nulo; la ecuación se escribe

$$-u_{\xi_1 \xi_1} - \cdots - u_{\xi_m \xi_m} + u_{\xi_{m+1} \xi_{m+1}} + \cdots + u_{\xi_n \xi_n} + \bar{F} = 0, \quad 1 < m < n. \quad (8.24d)$$

Por supuesto, el tipo puede ser diferente en diferentes puntos de \mathbb{R}^n , si uno o más autovalores de A cambian de signo entre ellos.

8.2. Modelos físicos

Como veremos, los problemas que involucran EDPs sólo estarán completos si además especificamos sus CC y sus CI. Como tal, un conjunto EDP+CC+CI aspirará a ser un *modelo* de un sistema o proceso físico, y cada una de sus componentes tendrá un significado razonablemente preciso, sobre todo si pensamos en el problema como un modelo de un *experimento*. Esta relación se esquematiza en la Tabla 8.2.

CC espacio-temporales	Contorno	EDP Hiperb.	EDP Elíp.	EDP Parab.
Dirichlet	abierto	indefinida	indefinida	única
Neumann	cerrado	indefinida	única	sobredet.
Robin				
Cauchy	abierto	única	sobredet.	sobredet.
	cerrado	sobredet.	sobredet.	sobredet.

Tabla 8.3: Condiciones de contorno espacio-temporales (CC+CI) para diferentes tipos de EDP, y el carácter de la solución resultante en cada caso.

8.2.1. Condiciones de contorno

En principio, podríamos pensar que dada una EDP en n variables independientes, bastaría con definir un dominio cualquiera $\Omega \subset \mathbb{R}^n$ y especificar CC de algún modo “adecuadas” sobre su frontera $\partial\Omega$, de modo que la solución exista y sea única. Sin embargo, aunque esto puede ser matemáticamente correcto, al menos para algunos tipos de EDPs, en Física el tiempo juega un papel marcadamente diferente al de las coordenadas espaciales: en general, si pensamos en un experimento, tenemos la imagen de un sistema que es preparado en un estado inicial específico, y a partir de cierto instante es dejado “libre” de evolucionar de acuerdo a su propia dinámica.² De allí la distinción entre CI y CC “propiamente dichas”, que se refieren a las coordenadas *espaciales*, reflejada en la Tabla 8.2.

Si nuestra EDP involucra el tiempo como una de las variables independientes, entonces, las CI deberán corresponder al *orden más alto* de la derivada temporal que aparezca: Si la EDP es de primer orden en t , deberemos en general especificar CI de Dirichlet, Neumann o Robin, como lo haríamos para una EDO de primer orden en t . Si es de segundo orden, deberemos especificar por separado el valor de la función incógnita u y de su derivada temporal u_t en el instante inicial, como lo haríamos para una EDO; estas se conocen como condiciones *de Cauchy*.

En cuanto a las CC *espaciales* a especificar en los límites espaciales del sistema modelado, en principio pueden ser de cualquiera de los cuatro tipos, dependiendo del sistema y de si el contorno espacial es abierto o cerrado; en el caso de más de una dimensión espacial, estas se definen, para x en la frontera del dominio espacial y \hat{n} la normal saliente en la frontera, como:

Dirichlet: Especificamos $u(\vec{x})$.

Neumann: Especificamos $\hat{n} \cdot \nabla u(\vec{x})$.

Robin: Especificamos $\alpha\hat{n} \cdot \nabla u(\vec{x}) + \beta u(\vec{x})$.

Cauchy: Especificamos $u(\vec{x})$ y $\hat{n} \cdot \nabla u(\vec{x})$.

²Aún si este “instante inicial” es diferente en distintos puntos del espacio.

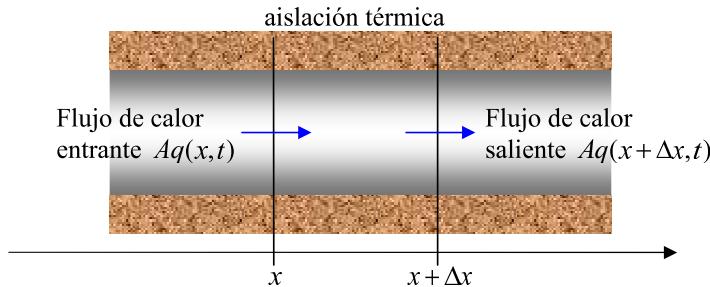


Figura 8.1: Conducción del calor en una barra delgada lateralmente aislada.

La Tabla 8.3 muestra un resumen de los resultados de aplicar diferentes CC y CI a EDPs de cada tipo, y allí ambas se agrupan como “CC espacio-temporales”.

Discutiremos a continuación cómo se deducen las EDP para dos de los fenómenos más frecuentemente modelados: difusión del calor y ondas elásticas.

8.2.2. La ecuación de difusión del calor

La difusión o transporte de calor en un medio material es uno de los problemas más antiguos de la Física, a cuyo estudio contribuyeron Newton y Fourier entre muchos otros; de hecho, es el problema que impulsó a Fourier a desarrollar su método de resolución por series trigonométricas. Por sencillez discutiremos primero el problema en una dimensión espacial.

Consideremos una barra sólida delgada que se extiende a lo largo del eje x , aislada lateralmente. La delgadez de la barra puede cuantificarse estableciendo que el área A de su sección transversal, que asumiremos uniforme (es decir independiente de x) es despreciable frente al cuadrado de su longitud L ; junto con la aislación lateral esto nos permite afirmar que, con excelente aproximación, la temperatura de la barra será uniforme en cualquier sección transversal (es decir, a x constante), de modo que la temperatura u dependerá solamente de la posición x y el tiempo t , y el flujo de calor tendrá la dirección de x .

Consideremos ahora el balance energético de una “rodaja” de la barra entre x y $x + \Delta x$, como se muestra en la Figura 8.1. El balance de energía interna de este elemento de volumen $A\Delta x$ contendrá tres términos:

En primer lugar, el flujo neto de calor entrante al elemento de volumen,

$$Q(x, t) = Aq(x, t) - Aq(x + \Delta x, t), \quad (8.25)$$

donde $q(x, t)$ es el flujo de calor por unidad de área, que por la ley de enfriamiento de Fourier es

$$q(x, t) = -K(x) \frac{\partial u(x, t)}{\partial x}, \quad (8.26)$$

con $K(x)$ la *conductividad térmica* del material, que admitiremos puede ser función de la posición.

En segundo lugar, la tasa de generación de energía en el elemento de volumen,

$$G(x, t) = A\Delta x g(x, t), \quad (8.27)$$

donde $g(x, t)$ es la tasa de generación de energía por unidad de volumen. Esta puede deberse a efecto Joule, decaimiento radiactivo, etc.

En tercer lugar, la tasa de aumento de energía interna del elemento de volumen debida al cambio de su temperatura con el tiempo,

$$\dot{U}(x, t) = A\Delta x \rho(x) C(x) \frac{\partial u(x, t)}{\partial t}, \quad (8.28)$$

donde $\rho(x)$ es la densidad de masa y $C(x)$ la capacidad calorífica del material, y ambas pueden depender de la posición x .

Ahora, el balance de energía del elemento de volumen se expresa

$$\dot{U}(x, t) = Q(x, t) + G(x, t). \quad (8.29)$$

Substituyendo las expresiones para cada contribución tenemos

$$A\Delta x \rho(x) C(x) \frac{\partial u(x, t)}{\partial t} = Aq(x, t) - Aq(x + \Delta x, t) + A\Delta x g(x, t), \quad (8.30)$$

que simplificando A , dividiendo por Δx y tomando el límite $\Delta x \rightarrow 0$ queda

$$\rho(x) C(x) \frac{\partial u(x, t)}{\partial t} = -\frac{\partial q(x, t)}{\partial x} + g(x, t), \quad (8.31)$$

y substituyendo la expresión para $q(x, t)$ resulta

$$\rho C \frac{\partial u}{\partial t} - \frac{\partial}{\partial x} \left(K \frac{\partial u}{\partial x} \right) = g. \quad (8.32)$$

Esta es la forma más general de la *ecuación de difusión del calor* en una dimensión. En el caso (relativamente común) donde $g = 0$ y ρ , C y K son todos constantes (que corresponde a un material homogéneo), se reduce a

$$\frac{\partial u}{\partial t} - \kappa \frac{\partial^2 u}{\partial x^2} = 0, \quad (8.33)$$

con

$$\kappa = \frac{K}{\rho C} \quad (8.34)$$

la *difusividad térmica*.

A continuación debemos determinar las CC que corresponde imponer en cada extremo de la barra, que supondremos que abarca el intervalo (a, b) de la recta real. La condición más general es la de *contacto térmico imperfecto* con otro cuerpo, a temperaturas $u_a(t)$ y $u_b(t)$. En el extremo en a , la ley de enfriamiento de Newton nos dice que el flujo *entrante* de calor $Q_a(t)$ viene dado por

$$Q_a(t) = Ah_a(u_a(t) - u(a, t)), \quad (8.35)$$

donde h_a es el *coeficiente de transferencia térmica* entre el extremo $x = a$ de la barra y el cuerpo en contacto con ella. Este flujo debe ser igual al previsto por la ley de enfriamiento de Fourier, que es $Aq(a, t)$, es decir

$$Ah_a(u_a(t) - u(a, t)) = -AK(a)u_x(a, t). \quad (8.36)$$

Luego

$$K(a)u_x(a, t) - h_a u(a, t) = -h_a u_a(t), \quad (8.37)$$

con lo que tenemos una CC de Robin inhomogénea. La CC en $x = b$ resulta similar, pero teniendo en cuenta que ahora $Q(b, t) = Q_b(t)$ representa un flujo de calor *saliente* de la barra, se escribirá

$$K(b)u_x(b, t) + h_b u(b, t) = h_b u_b(t), \quad (8.38)$$

Los casos límite más comunes son $h_a \ll K(a)$, que corresponde a un extremo térmicamente *aislado* y se reduce a las CC de Neumann

$$u_x(a, t) = 0, \quad (8.39)$$

y $h_a \gg K(a)$, que corresponde a *contacto térmico perfecto* y se reduce a las CC de Dirichlet

$$u(a, t) = u_a(t), \quad (8.40)$$

con expresiones equivalentes en $x = b$.

La generalización de la ecuación de difusión a más dimensiones es directa, y se obtiene

$$\rho C \frac{\partial u}{\partial t} - \nabla \cdot (K \nabla u) = g, \quad (8.41)$$

que para $g \equiv 0$ y ρ, C y K independientes de la posición se reduce a

$$\frac{\partial u}{\partial t} - \kappa \nabla^2 u = 0. \quad (8.42)$$

Las CC genéricas son las de Robin

$$K(\vec{x})\hat{n} \cdot \nabla u(\vec{x}, t) + h(\vec{x})u(\vec{x}, t) = h(\vec{x})u_e(\vec{x}, t), \quad \vec{x} \in \partial\Omega, \quad (8.43)$$

donde \hat{n} es la normal saliente en la frontera $\partial\Omega$ del dominio $\Omega \subset \mathbb{R}^n$ donde vale la EDP y u_e es la temperatura del medio externo, y pueden reducirse a CC de Dirichlet o de Neumann en toda la frontera o partes de ella.

Ejercicio 8.2.1. Verifique las expresiones para la EDP y las CC dadas arriba.

8.2.3. La ecuación de la cuerda vibrante

Las vibraciones de una cuerda tensa son otro problema “clásico” cuya descripción completa debió esperar al método de Fourier, y fué el prototipo de los sistemas donde aparece la ecuación de ondas en su forma más “pura”, al menos hasta la llegada del Electromagnetismo.

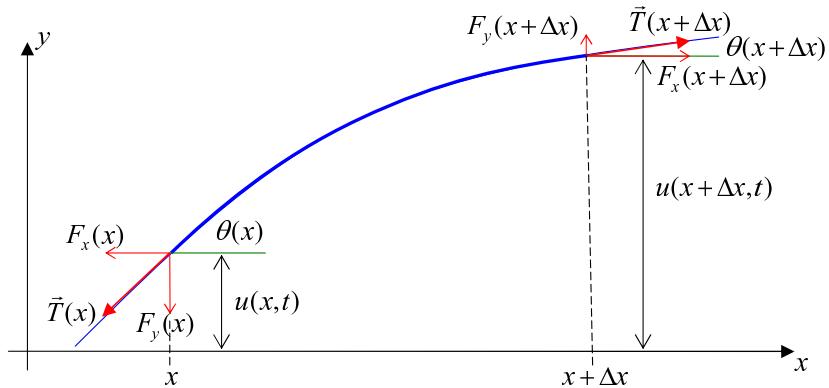


Figura 8.2: Un segmento de cuerda delgada y las fuerzas aplicadas.

Las principales hipótesis simplificadorias, respecto a las características de la cuerda, son: primera, que la cuerda es delgada, es decir, puede asimilarse a una curva matemática; segunda, que es infinitamente flexible, es decir, las fuerzas generadas por el mero hecho de curvarla son despreciables; tercera, que es inextensible, es decir que el cambio de largo respecto a su longitud “en reposo” debido a sus oscilaciones, es también despreciable.

Para la versión más sencilla del problema, las hipótesis simplificadorias respecto a su dinámica son: primera, que en su estado de reposo la cuerda es recta, digamos horizontal; segunda, que las oscilaciones tienen lugar en un plano, digamos vertical; tercera, que las oscilaciones son “pequeñas”, lo que no sólo significa que los desplazamientos respecto al estado de reposo lo son, sino que en toda su longitud la pendiente de la cuerda también es pequeña (no hay grandes curvaturas, la cuerda es casi horizontal todo a lo largo), lo que veremos que es fundamental aunque pocas veces se explica; cuarta, que la tensión a lo largo de la cuerda es prácticamente constante; y quinta, que no hay otras fuerzas exteriores aplicadas salvo las que fijan la posición de sus extremos.

La posición o forma de la cuerda queda entonces descripta dando la curva $y = u(x, t)$ que la representa en el plano en que oscila, donde en reposo la cuerda coincide con el eje x . Un segmento de la cuerda entre x y $x + \Delta x$ estará sujeto sólo a las fuerzas aplicadas en sus extremos por el resto de la cuerda debido a su tensión T , que serán tangentes a la cuerda en los puntos correspondientes, como se muestra en la Figura 8.2. Notemos primero que, dado que la tensión es prácticamente constante a lo largo de la cuerda, y los ángulos $\theta(x)$ y $\theta(x + \Delta x)$ son muy pequeños, las fuerzas *horizontales* $F_x(x) = -T \cos \theta(x)$ y $F_x(x + \Delta x) = T \cos \theta(x + \Delta x)$ tendrán la misma magnitud y se cancelarán; por consiguiente *no hay fuerzas netas horizontales* sobre ningún segmento de la cuerda, y suponiendo inicialmente velocidades horizontales nulas, *cada punto de la cuerda se mueve solamente en vertical*.

La aceleración vertical del segmento en cuestión vendrá dada por la segunda

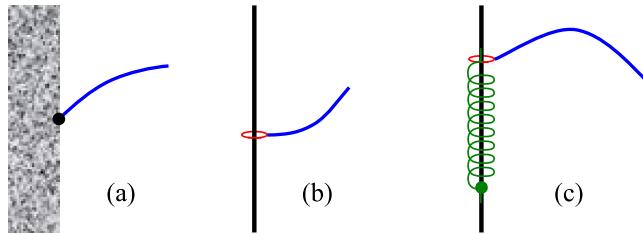


Figura 8.3: Condiciones de contorno para la cuerda vibrante: (a) extremo fijo; (b) extremo libre; (c) enlace elástico.

ley de Newton, como

$$\rho(x)\Delta x \frac{\partial^2 u(x, t)}{\partial t^2} = F_y(x + \Delta x) + F_y(x), \quad (8.44)$$

donde $\rho(x)$ es la densidad lineal de masa de la cuerda, que puede depender de la posición, y $\rho(x)\Delta x$ es la masa del segmento. Pero dada la pequeñez de los ángulos,

$$\begin{aligned} F_y(x) &= -T \operatorname{sen} \theta(x) \simeq -T \tan \theta(x) = -Tu_x(x, t), \\ F_y(x + \Delta x) &= T \operatorname{sen} \theta(x + \Delta x) \simeq T \tan \theta(x + \Delta x) = Tu_x(x + \Delta x, t). \end{aligned}$$

Substituyendo en la ecuación anterior, dividiendo por Δx y tomando el límite $\Delta x \rightarrow 0$, entonces,

$$\rho(x)u_{tt}(x, t) - Tu_{xx}(x, t) = 0, \quad (8.45)$$

que es la forma general de la *ecuación de ondas* en una dimensión. Si la densidad lineal de masa es constante, esta se simplifica a

$$u_{tt}(x, t) - c^2 u_{xx}(x, t) = 0, \quad (8.46)$$

donde

$$c = \sqrt{\frac{T}{\rho}} \quad (8.47)$$

es la *velocidad de propagación de las ondas*.

Ahora debemos determinar las CC a aplicar en cada extremo. Suponiendo que la cuerda abarca el intervalo (a, b) , tenemos tres casos principales:

Extremo fijo: El extremo $x = a$ de la cuerda está firmemente unido a un objeto esencialmente inmóvil, por ejemplo una pared, como se ilustra en la Figura 8.3.a. La CC es evidentemente de Dirichlet,

$$u(a, t) = 0. \quad (8.48)$$

Extremo libre: El extremo $x = a$ de la cuerda puede moverse libremente *en la dirección vertical*, no así en la horizontal, de modo que *no hay fuerzas verticales* aplicadas al extremo; por ejemplo, está unido a un anillo (sin masa) que desliza sin fricción sobre un eje vertical, como se ilustra en la Figura 8.3.b. Dado que la fuerza vertical aplicada por la cuerda al anillo también debe ser nula, y es proporcional a la pendiente de la cuerda, la CC es de Neumann,

$$u_x(a, t) = 0. \quad (8.49)$$

Enlace elástico: Similar al caso anterior, pero ahora el anillo está unido a un punto fijo del eje por un resorte de constante elástica k y longitud natural nula, como muestra la Figura 8.3.c. La fuerza vertical neta aplicada al anillo por la cuerda y el resorte debe anularse (porque la masa del anillo es nula), de modo que la CC es de Robin,

$$Tu_x(a, t) - ku(a, t) = 0. \quad (8.50)$$

Condiciones equivalentes pueden formularse en el extremo $x = b$.

La generalización de la ecuación de ondas a una *membrana* bidimensional es directa, y se obtiene

$$\rho(\vec{x})u_{tt}(\vec{x}, t) - T\nabla^2 u(\vec{x}, t) = 0, \quad (8.51)$$

donde ahora $\rho(\vec{x})$ es una densidad superficial de masa y \vec{x} es un vector en el plano donde yace la membrana cuando está en reposo. La generalización de las CC de borde fijo (Dirichlet) y de borde libre (Neumann) es inmediata, y la CC de enlace elástico se escribe, en el caso más general,

$$T\hat{n} \cdot \nabla u(\vec{x}, t) - k(\vec{x})u(\vec{x}, t) = 0, \quad \vec{x} \in \partial\Omega, \quad (8.52)$$

donde \hat{n} es la normal saliente en la frontera $\partial\Omega$ del dominio $\Omega \subset \mathbb{R}^2$ donde vale la EDP. Las CC pueden ser de distintos tipos en partes diferentes de la frontera.

Ejercicio 8.2.2. Verifique las expresiones para la EDP y las CC dadas arriba.

8.2.4. La cuerda infinita y las curvas características

Como ejemplo del papel que juegan las curvas características de una EDP, vamos a estudiar el problema de las oscilaciones transversales de una cuerda homogénea infinita. La EDP adopta la forma (8.46), y de acuerdo a (8.6b), las curvas características serán soluciones de la EDO

$$dx^2 - c^2 dt^2 = 0.$$

Factorizándola como

$$(dx + cdt)(dx - cdt) = 0,$$

es inmediato que tendremos las soluciones

$$\phi(x, t) = x - ct, \quad \psi(x, t) = x + ct.$$

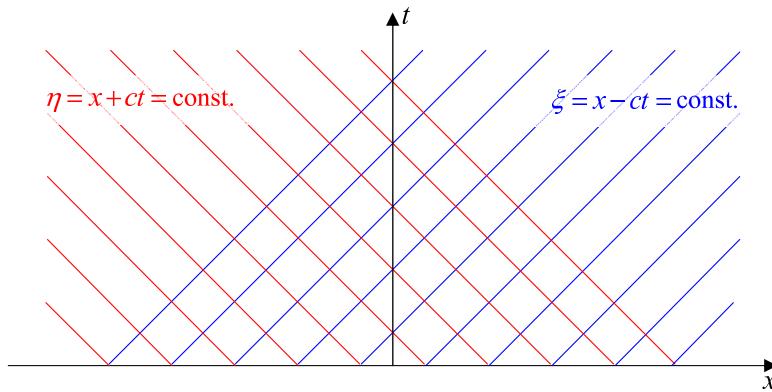


Figura 8.4: Curvas características de la ecuación de ondas 1D.

Las curvas características $\phi(x, t) = \text{constante}$ y $\psi(x, t) = \text{constante}$ y las nuevas coordenadas $\xi = \phi(x, t)$ y $\eta = \psi(x, t)$ se muestran en la Figura 8.4.

La EDP original, puesta en la forma (8.2a), es

$$u_{xx} - \frac{1}{c^2} u_{tt} = Au_{xx} + 2Bu_{xt} + Cu_{tt} = 0,$$

luego $A = 1$, $B = 0$ y $C = -1/c^2$. Definiendo las nuevas coordenadas $\xi = \phi(x, t)$ y $\eta = \psi(x, t)$ según (8.3), entonces, de acuerdo a (8.5a) la EDP quedará en la forma

$$4u_{\xi\eta} = 0.$$

La solución general es inmediata, y se escribe

$$u(\xi, \eta) = f(\xi) + g(\eta),$$

donde f y g son dos funciones C^2 arbitrarias. Esta es conocida como la *solución de D'Alembert*. Volviendo a las variables originales, se expresa

$$u(x, t) = f(x - ct) + g(x + ct), \quad (8.53)$$

y consiste en dos perturbaciones con la *forma de onda* f y g que se mueven respectivamente hacia la derecha e izquierda a velocidad c .

Ejercicio 8.2.3. Verifique todos los cálculos que llevan a (8.53).

Las dos funciones f y g de la solución general quedan especificadas dando CI de Cauchy, por ejemplo en $t = 0$:

$$\begin{aligned} f(x) + g(x) &= h(x), \\ -cf'(x) + cg'(x) &= v(x). \end{aligned}$$

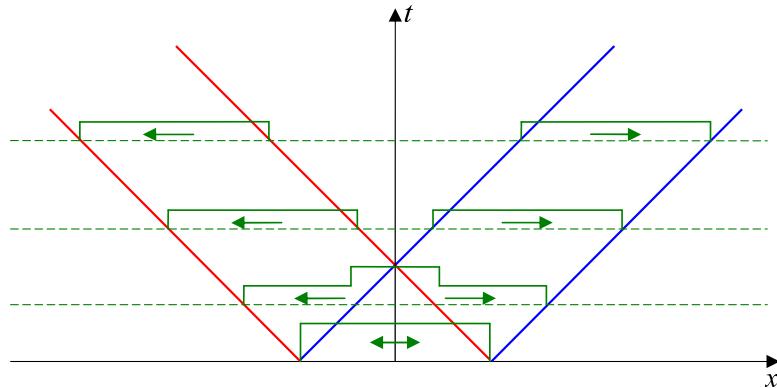


Figura 8.5: Propagación de un pulso rectangular en la ecuación de ondas 1D.

Ejercicio 8.2.4. Muestre que la solución de D'Alembert con las CC de arriba queda

$$u(x, t) = \frac{1}{2} [h(x - ct) + h(x + ct)] + \frac{1}{2c} \int_{x-ct}^{x+ct} v(x') dx'. \quad (8.54)$$

Un ejemplo de la “propagación” de las soluciones se muestra en la Figura 8.5; las CI elegidas son

$$\begin{aligned} f(x) + g(x) &= \text{rect}(x), \\ -cf'(x) + cg'(x) &= 0. \end{aligned}$$

Allí vemos que el perfil rectangular inicial se va separando en dos idénticos de la mitad de su altura, que propagan hacia la derecha y la izquierda a velocidad c . Esto ilustra cómo, en la ecuación de ondas, la información contenida en los datos iniciales *se propaga a lo largo de las curvas características*; esto se debe a que, por construcción, una solución $u = f(x, t)$ de la EDP tiene un valor *constante* a lo largo de una curva característica.

Ejercicio 8.2.5. Suponga que la cuerda, en lugar de ser infinita, se extiende para $x > 0$ y está sujeta a CC de Dirichlet $u(0, t) = 0$ (imagine que la cuerda está fija a una pared allí). Considere un pulso rectangular que avanza hacia la pared. Muestre que el pulso *se refleja* en la pared y luego se aleja de ella, pero con u cambiada por $-u$, es decir, invertido verticalmente (la reflexión *invierte la fase* del pulso).

Ayuda: imponga la CC mediante el método de las imágenes, asimilando el problema al de una cuerda infinita donde dos pulsos antisimétricos respecto al origen avanzan inicialmente uno hacia el otro, de modo que automáticamente $u(0, t)$ se anula para todo t .

Capítulo 9

Método de Separación de Variables

El método de separación de variables es probablemente el más usado en la resolución de EDPs, debido tanto a su simplicidad como a su amplia aplicabilidad. La esencia del método consiste en la descomposición de la función incógnita, que depende de varias variables independientes, en varias funciones dependientes cada una de una única variable independiente, para cada una de las cuales se obtiene un problema de Sturm–Liouville.

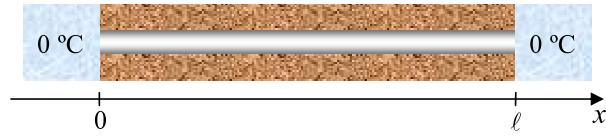
Más allá de lo dicho, no intentaremos desarrollar una teoría general del método *in abstracto*. Por experiencia, resulta mucho más productivo (y llevadero) presentarlo a través de su aplicación a los problemas de más frecuente aparición en Física, de donde irán surgiendo naturalmente sus características destacables.

9.1. Dimensión $1 + 1$

Los problemas más sencillos a tratar involucran dos variables independientes, el tiempo y una coordenada espacial; de allí el mote “ $1 + 1$ ”. Trataremos principalmente con la ecuación de difusión del calor y la ecuación de ondas, como han sido presentadas en la Sección 8.2; su simplicidad permitirá una introducción igualmente simple del método, sin oscurecerla con complejidades técnicas que aparecerán en dimensiones superiores.

9.1.1. Difusión del calor con CC homogéneas

El primer problema a tratar, extremadamente sencillo, será el de difusión del calor en una barra delgada y homogénea aislada lateralmente, como la tratada en la Sección 8.2.2. Supondremos que la barra yace a lo largo del eje x y se extiende desde $x = 0$ a $x = \ell$, como se muestra en la Figura 9.1. Para mayor simplicidad, supondremos además que los extremos de la barra están en contacto térmico

Figura 9.1: Barra aislada con extremos a 0°C .

perfecto con sendos reservorios de calor a 0°C ; de este modo las CC en $x = 0$ y $x = \ell$ resultan de Dirichlet y homogéneas. Finalmente, asumiremos que en el instante inicial $t = 0$ la temperatura $u(x, 0)$ de la barra es una función lineal de la posición, que va de 0 en $x = 0$ a α en $x = \ell$.

Como buena costumbre adoptaremos una formulación ordenada del problema matemático resultante, como sigue:

$$u_t - \kappa u_{xx} = 0, \quad 0 < x < \ell, \quad 0 < t < \infty, \quad (9.1a)$$

$$u(0, t) = u(\ell, t) = 0, \quad (9.1b)$$

$$u(x, 0) = \frac{\alpha}{\ell}x, \quad (9.1c)$$

es decir, listamos las partes del problema, primero la EDP con su dominio de validez, luego las CC y finalmente las CI si las hay.

Consideremos la EDP (9.1a); vamos a buscar soluciones de la forma

$$u(x, t) = U(x)T(t), \quad (9.2)$$

que sean productos de una función de x y una de t . Substituyendo en (9.1a) tenemos

$$UT' - \kappa U''T = 0,$$

que dividiendo por UT se escribe

$$\frac{U''}{U} = \frac{1}{\kappa} \frac{T'}{T}.$$

La idea central del método de separación de variables es notar que en esta última ecuación el lado izquierdo es función sólo de x mientras que el lado derecho es función sólo de t y que, por lo tanto, la igualdad de ambos *para todo* x y *para todo* t sólo es posible si *ambos lados son iguales a una y la misma constante*. Llamando a esa *constante de separación* $-k^2$ para futura comodidad, nuestro problema se escribe

$$\frac{U''}{U} = \frac{1}{\kappa} \frac{T'}{T} = -k^2,$$

lo que lleva inmediatamente a las *EDO separadas*

$$U'' + k^2U = 0, \quad (9.3a)$$

$$T' + \kappa k^2T = 0. \quad (9.3b)$$

Las soluciones generales de las EDOs (9.3) son

$$U(x) = a \cos(kx) + b \sin(kx), \quad (9.4a)$$

$$T(t) = ce^{-\kappa k^2 t}, \quad (9.4b)$$

con lo que, absorbiendo la constante c en las constantes a y b , la solución general de (9.1a) de la forma separada (9.2) se escribe

$$u(x, t) = U(x)T(t) = (a \cos(kx) + b \sin(kx))e^{-\kappa k^2 t}. \quad (9.5)$$

El siguiente paso es imponer las CC (9.1b) a la solución (9.5). Es inmediato que $T(t)$ no se anulará, así que debemos imponer

$$U(0) = U(\ell) = 0, \quad (9.6)$$

con lo que tendremos $a = 0$ y $\sin(k\ell) = 0$. $U(x)$ deberá ser entonces alguna de las *autofunciones espaciales*

$$U_n(x) = \sin(k_n x), \quad k_n = \frac{n\pi}{\ell}, \quad n \in \mathbb{N}, \quad (9.7)$$

a cada una de las cuales le corresponderá una autofunción

$$u_n(x, t) = \sin(k_n x)e^{-\kappa k_n^2 t} \quad (9.8)$$

del problema formado por la EDO (9.1a) más las CC (9.1b). La solución general de este problema será, entonces, una combinación lineal de ellas,

$$u(x, t) = \sum_{n=1}^{\infty} b_n u_n(x, t) = \sum_{n=1}^{\infty} b_n \sin(k_n x)e^{-\kappa k_n^2 t}. \quad (9.9)$$

Finalmente, debemos imponer las CI (9.1c), haciendo

$$u(x, 0) = \sum_{n=1}^{\infty} b_n \sin(k_n x) = \frac{\alpha}{\ell} x. \quad (9.10)$$

Notando que las autofunciones espaciales (9.7) son las autofunciones del PSL

$$U'' + k^2 U = 0, \quad 0 < x < \ell, \quad U(0) = U(\ell) = 0, \quad (9.11)$$

y por lo tanto ortogonales y completas en $L^2(0, \ell)$, vemos que

$$b_n = \frac{\langle U_n, \frac{\alpha}{\ell} x \rangle}{\langle U_n, U_n \rangle} = \frac{(-1)^n 2\alpha}{n\pi}. \quad (9.12)$$

Entonces, la solución del problema (9.1) se escribe

$$u(x, t) = \frac{2\alpha}{\pi} \sum_{n=1}^{\infty} \frac{(-1)^n}{n} \sin(k_n x)e^{-\kappa k_n^2 t}. \quad (9.13)$$

Vale la pena hacer algunos comentarios acerca de la solución obtenida y el proceso para obtenerla. En primer lugar, hay que notar que la serie (9.13) converge puntualmente en $(0, \ell)$, pero esa convergencia es muy lenta, y no es ni uniforme ni absoluta. Tenemos garantizada, sin embargo, la convergencia en $\mathcal{L}^2(0, \ell)$.

En segundo lugar, el proceso para obtener la solución (9.13) podría haberse abreviado en cierta medida: Dado que las CC (9.1b) son homogéneas, es evidente que al separar $u(x, t)$ como $U(x)T(t)$ deben ser satisfechas por $U(x)$, ya que forzar a $T(t)$ a satisfacerlas nos llevaría de inmediato a la solución trivial. Podríamos entonces habernos ahorrado escribir (9.5) y haber resuelto primero el PSL (9.11), pasando a escribir directamente (9.9). Esta clase de “atajos” son ventajosos, y los tomaremos en adelante.

9.1.2. Difusión del calor con CC inhomogéneas

Resolveremos ahora un problema semejante al anterior, pero donde las CC son inhomogéneas, a fin de ilustrar el uso del procedimiento de homogeneización de las CC discutido brevemente en la Sección 5.3.4. Concretamente,

$$u_t - \kappa u_{xx} = 0, \quad 0 < x < \ell, \quad 0 < t < \infty, \quad (9.14a)$$

$$u(0, t) = 0, \quad u(\ell, t) = \alpha, \quad (9.14b)$$

$$u(x, 0) = 0. \quad (9.14c)$$

De acuerdo a lo descripto en la Sección 5.3.4, vamos a descomponer la función incógnita u como

$$u(x, t) = v(x, t) + w(x, t), \quad (9.15)$$

donde w satisface las CC inhomogéneas (9.14b), es decir

$$w(0, t) = 0, \quad w(\ell, t) = \alpha.$$

Aquí es donde una elección juiciosa de la función auxiliar w resulta crucial para simplificar el problema. Notando de la Sección 9.1.1 que la solución de un problema de difusión (es decir, de una EDP parabólica) con CC independientes del tiempo tenderá para $t \rightarrow \infty$ a un *estado estacionario*, es decir a una solución independiente de t , buscaremos que w sea una solución de este tipo, es decir, que sea solución del *problema estacionario*

$$w_{xx} = 0, \quad 0 < x < \ell, \quad (9.16a)$$

$$w(0) = 0, \quad w(\ell) = \alpha. \quad (9.16b)$$

La solución general de (9.16a) es una función lineal de x , y las CC (9.16b) nos llevan a

$$w(x) = \frac{\alpha}{\ell}x. \quad (9.17)$$

La ventaja de haber elegido así la función w queda de manifiesto si advertimos que, por construcción,

$$w_t - \kappa w_{xx} \equiv 0,$$

ya que si ahora substituimos (9.15) en el problema original (9.14), obtenemos para v el problema

$$v_t - \kappa v_{xx} = 0, \quad 0 < x < \ell, \quad 0 < t < \infty, \quad (9.18a)$$

$$v(0, t) = v(\ell, t) = 0, \quad (9.18b)$$

$$v(x, 0) = -w(x), \quad (9.18c)$$

donde tanto las CC (9.18b) como la EDP (9.18a) son homogéneas, y la única inhomogeneidad está en las CI (9.18c). Y en vista de (9.17), vemos que el problema para v es idéntico al problema (9.1) cambiando α por $-\alpha$, por lo que usando (9.13) podemos escribir inmediatamente

$$v(x, t) = -\frac{2\alpha}{\pi} \sum_{n=1}^{\infty} \frac{(-1)^n}{n} \operatorname{sen}(k_n x) e^{-\kappa k_n^2 t}. \quad (9.19)$$

Lo único que resta por hacer, entonces, para obtener la solución completa del problema (9.14), es substituir (9.17) y (9.19) en (9.15), lo que resulta en

$$u(x, t) = \frac{\alpha}{\ell} x - \frac{2\alpha}{\pi} \sum_{n=1}^{\infty} \frac{(-1)^n}{n} \operatorname{sen}(k_n x) e^{-\kappa k_n^2 t}. \quad (9.20)$$

Debemos enfatizar que, siempre que la EDP sea parabólica y las CC sean inhomogéneas pero independientes del tiempo, podremos obtener una simplificación equivalente del problema usando para la función auxiliar w la solución del correspondiente problema estacionario con las mismas CC inhomogéneas. La única diferencia con lo realizado aquí será que, en general, no tendremos la solución del problema estacionario calculada de antemano, y deberemos calcularla como parte del proceso.

9.1.3. Difusión del calor con CC periódicas en el tiempo

Hemos visto en la sección anterior que, dada una EDP parabólica y CC inhomogéneas pero constantes, la solución del problema tiende, a tiempos largos, a un estado estacionario. Cabe preguntarse, entonces, qué ocurrirá en un problema similar pero donde las CC inhomogéneas sean además funciones del tiempo. Para una dependencia arbitraria con el tiempo la respuesta deberá aguardar al Capítulo 10. Aquí nos limitaremos a mostrar cómo proceder si las CC son funciones *periódicas* del tiempo.

El problema tendrá la forma

$$u_t - \kappa u_{xx} = 0, \quad 0 < x < \ell, \quad -\infty < t < \infty, \quad (9.21a)$$

$$u(0, t) = 0, \quad u(\ell, t) = \alpha(t), \quad (9.21b)$$

$$u(x, -\infty) = \dots, \quad (9.21c)$$

donde la función $\alpha(t)$ es periódica de periodo T y, por consiguiente, puede expresarse como una serie de Fourier compleja

$$\alpha(t) = \sum_{n=-\infty}^{\infty} \alpha_n e^{i\omega_n t}, \quad \omega_n = \frac{2n\pi}{T}. \quad (9.22)$$

El lector habrá notado un par de peculiaridades en la formulación del problema (9.21): por un lado, en (9.21a) hemos extendido el intervalo temporal a $(-\infty, \infty)$; por otro, en (9.21c) hemos asumido que las CI se especificaron en $t = -\infty$, y estamos omitiendo su expresión. La justificación es la siguiente: Sabemos que la solución del problema puede escribirse como la suma de dos funciones, la solución u_h del problema homogéneo (EDP y CC homogéneas) y una solución u_r del problema con CC periódicas en t . De éstas, la que ajusta las CI es u_h ; pero como ya hemos visto, la solución del problema homogéneo decrece con el tiempo, de modo que mucho después de imponer las CI será despreciable. Asumir que las CI se impusieron “hace mucho tiempo”, es decir en $t = -\infty$, nos deja entonces sólo con la *solución de régimen* u_r , que no necesita ni admite CI, y que será la que nos interesa determinar.

Una simplificación adicional proviene de la periodicidad de $\alpha(t)$: para cada componente $e^{i\omega_n t}$ de (9.22) podemos resolver el correspondiente problema

$$u_{n,t} - \kappa u_{n,xx} = 0, \quad 0 < x < \ell, \quad -\infty < t < \infty, \quad (9.23a)$$

$$u_n(0, t) = 0, \quad u_n(\ell, t) = e^{i\omega_n t} \quad (9.23b)$$

para después obtener la solución de (9.21) apelando al *principio de superposición lineal*, como

$$u(x, t) = \sum_{n=-\infty}^{\infty} \alpha_n u_n(x, t). \quad (9.24)$$

Finalmente, aún otra simplificación proviene del hecho de que $e^{i\omega_n t}$ es una *autofunción del operador* ∂_t : lo aprovecharemos para proponer una forma específica para la separación de variables, como

$$u_n(x, t) = U_n(x)T_n(t) = U_n(x)e^{i\omega_n t}. \quad (9.25)$$

Substituyendo en la EDP (9.23) obtenemos

$$i\omega_n U_n(x)e^{i\omega_n t} - \kappa U_n''(x)e^{i\omega_n t} = 0,$$

de donde resulta para $U(x)$ el problema

$$U_n'' - \frac{i\omega_n}{\kappa} U_n = 0, \quad U_n(0) = 0, \quad U_n(\ell) = 1. \quad (9.26)$$

Para resolver este último problema, proponemos soluciones de la forma $e^{k_n x}$ y obtenemos

$$k_n^2 = \frac{i\omega_n}{\kappa},$$

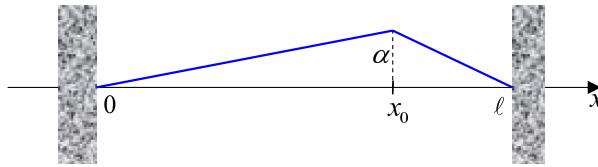
es decir

$$k_n = \frac{1+i}{\sqrt{2}} \sqrt{\frac{\omega_n}{\kappa}}.$$

Luego¹

$$U_n(x) = ae^{k_n x} + be^{-k_n x} = A \cosh(k_n x) + B \sinh(k_n x).$$

¹Nótese que $e^{k_n x}$ y $e^{-k_n x}$ no son complejas conjugadas una de la otra, de modo que las funciones $U_n(x)$ son *complejas*; esto era de esperar, ya que la CC en $x = \ell$ es compleja.

Figura 9.2: Deformación inicial de una cuerda tensa pulsada en el punto x_0 .

Las CC imponen

$$A = 0, \quad B \operatorname{senh}(k_n \ell) = 1,$$

de modo que

$$U_n(x) = \frac{\operatorname{senh}(k_n x)}{\operatorname{senh}(k_n \ell)}, \quad k_n = \frac{1+i}{\sqrt{2}} \sqrt{\frac{\omega_n}{\kappa}}. \quad (9.27)$$

Volviendo atrás, tendremos

$$u_n(x, t) = \frac{\operatorname{senh}(k_n x)}{\operatorname{senh}(k_n \ell)} e^{i \omega_n t} \quad (9.28)$$

y, finalmente,

$$u(x, t) = \sum_{n=-\infty}^{\infty} \alpha_n \frac{\operatorname{senh}(k_n x)}{\operatorname{senh}(k_n \ell)} e^{i \omega_n t}. \quad (9.29)$$

Ejercicio 9.1.1. Halle las soluciones *reales* del problema (9.21) si las CC son

$$u(0, t) = 0, \quad u(\ell, t) = A \cos(\omega t).$$

9.1.4. Oscilaciones de una cuerda

Vamos a ejemplificar ahora la separación de variables para la ecuación de ondas unidimensional, resolviendo el problema de las oscilaciones de una cuerda tensa de longitud ℓ fija en sus extremos, *pulsada* inicialmente en un punto $x_0 \in (0, \ell)$;² esta CI quiere decir que la cuerda es deformada inicialmente como se muestra en la Figura 9.2, y soltada desde el reposo.

El problema tendrá la forma

$$u_{tt} - c^2 u_{xx} = 0, \quad 0 < x < \ell, \quad 0 < t < \infty, \quad (9.30a)$$

$$u(0, t) = 0, \quad u(\ell, t) = 0, \quad (9.30b)$$

$$u(x, 0) = f(x), \quad u_t(x, 0) = 0, \quad (9.30c)$$

con la deformación inicial

$$f(x) = \begin{cases} \frac{\alpha}{x_0} x, & 0 < x < x_0, \\ \frac{\alpha}{\ell - x_0} (\ell - x), & x_0 < x < \ell. \end{cases} \quad (9.31)$$

²Un típico instrumento musical de cuerda pulsada es la guitarra.

Proponiendo como antes

$$u(x, t) = U(x)T(t)$$

y substituyendo en la EDP (9.30a), obtenemos

$$\frac{T''}{T} = c^2 \frac{U''}{U} = -\omega^2,$$

donde por conveniencia futura hemos llamado $-\omega^2$ a la constante de separación. Obtenemos así las EDO separadas

$$U'' + k^2 U = 0, \quad k = \frac{\omega}{c}, \quad (9.32a)$$

$$T'' + \omega^2 T = 0. \quad (9.32b)$$

El PSL espacial queda

$$U'' + k^2 U = 0, \quad 0 < x < \ell, \quad U(0) = U(\ell) = 0, \quad (9.33)$$

y, como ya sabemos, las autofunciones espaciales serán

$$U_n(x) = \sin(k_n x), \quad k_n = \frac{n\pi}{\ell}, \quad n \in \mathbb{N}. \quad (9.34)$$

Notemos que esto ya determina las correspondientes frecuencias $\omega_n = ck_n$, de modo que las soluciones de (9.32b) serán

$$T_n(t) = A \cos(\omega_n t) + B \sin(\omega_n t) \quad (9.35)$$

y la solución general de la EDP (9.30a) con las CC (9.30b) se escribirá

$$u(x, t) = \sum_{n=1}^{\infty} \sin(k_n x) (A_n \cos(\omega_n t) + B_n \sin(\omega_n t)). \quad (9.36)$$

Ahora debemos imponer las CI (9.30c), lo que lleva a

$$u(x, 0) = \sum_{n=1}^{\infty} A_n \sin(k_n x) = f(x),$$

$$u_t(x, 0) = \sum_{n=1}^{\infty} B_n \omega_n \sin(k_n x) = 0.$$

Usando la ortogonalidad y completitud de $U_n(x)$ en $L^2(0, \ell)$, entonces, resulta

$$A_n = \frac{2}{\ell} \langle \sin(k_n x), f(x) \rangle, \quad B_n = 0. \quad (9.37)$$

Ejercicio 9.1.2. Complete el cálculo de los coeficientes A_n , y estudie qué modos espaciales de oscilación $\sin(k_n x)$ se excitan y cuáles no, para $x_0 = \ell/2$.

Si en lugar de ser pulsada en x_0 la cuerda es *percutida* en el mismo lugar,³ es decir, estando tensa y en reposo se le comunica una velocidad inicial v_0 en una región muy pequeña alrededor de x_0 , el único cambio en el problema (9.30) es que las CI pasan a escribirse en la forma idealizada

$$u(x, 0) = 0, \quad u_t(x, 0) = v_0 \delta(x - x_0). \quad (9.38)$$

Imponiendo estas nuevas CI a la solución (9.36) obtenemos ahora

$$\begin{aligned} u(x, 0) &= \sum_{n=1}^{\infty} A_n \sin(k_n x) = 0, \\ u_t(x, 0) &= \sum_{n=1}^{\infty} B_n \omega_n \sin(k_n x) = v_0 \delta(x - x_0), . \end{aligned}$$

de donde la ortogonalidad y completitud de $U_n(x)$ lleva a

$$A_n = 0, \quad B_n = \frac{2v_0}{\ell \omega_n} \sin(k_n x_0), \quad (9.39)$$

y finalmente a

$$u(x, t) = \frac{2v_0}{\ell} \sum_{n=1}^{\infty} \frac{\sin(k_n x_0)}{\omega_n} \sin(k_n x) \sin(\omega_n t). \quad (9.40)$$

Notemos que si la cuerda es percutida en un punto x_0 donde el modo espacial n -ésimo $\sin(k_n x)$ tiene un *nodo*, ese modo no se excitará, ya que tendremos $\sin(k_n x_0) = 0$. Asimismo, no todos los modos que se excitan lo hacen con la misma amplitud de oscilación, ya que esta será proporcional a $\sin(k_n x_0)/\omega_n$; ello permite, eligiendo adecuadamente x_0 , variar el *timbre* del sonido emitido por la cuerda al vibrar, ya que este queda determinado por la amplitud (relativa) de las oscilaciones a cada frecuencia. Notemos también otra fenomenología interesante del problema de vibraciones de una cuerda tensa: todas las frecuencias de oscilación ω_n son múltiplos enteros de la frecuencia más baja, llamada *fundamental*, ω_1 ; se dice que son sus *armónicos*, porque la superposición de sonidos cuyas frecuencias relativas son racionales (pequeños) suena armónica al oído; este es también el origen del nombre “ecuación armónica” para una EDO de la forma $y'' + \omega^2 y = 0$.

Como una ligera extensión del problema que acabamos de resolver, consideraremos brevemente las oscilaciones de una cuerda tensa en un *medio viscoso* que frena progresivamente su movimiento. En ese caso la EDP (9.30) adquiere un término disipativo proporcional a la velocidad u_t , y pasa a escribirse

$$u_{tt} + 2\gamma u_t - c^2 u_{xx} = 0, \quad (9.41)$$

donde $\gamma > 0$ es la constante de disipación.⁴ Es inmediato que al separar variables, el PSL espacial (9.33) no se verá afectado, y los modos espaciales seguirán siendo

³Un típico instrumento musical de cuerda percutida es el piano.

⁴Nótese que la EDP sigue siendo hiperbólica.

$\sin(k_n x)$ con $k_n = n\pi/\ell$; pero la EDO (9.32b) asociada al n -ésimo modo espacial pasará a escribirse

$$T'' + 2\gamma T' + \omega^2 T = 0, \quad (9.42)$$

y sus soluciones serán ahora de la forma

$$e^{-\gamma t} e^{\pm i\sqrt{c^2 k_n^2 - \gamma^2}}. \quad (9.43)$$

Ejercicio 9.1.3. Complete los cálculos de arriba y muestre que, para fricción grande, algunos modos de oscilación pueden estar sobreamortiguados. (*¿Cuáles?*)

El lector debe estar advertido de que el modelo de disipación empleado en (9.41) es inadecuado en muchos casos: si la viscosidad del fluido es baja y/o la velocidad de las oscilaciones grande, el flujo alrededor de la cuerda será *turbulento* en lugar de laminar, y la disipación será proporcional al *cuadrado* de la velocidad; ello hace que los modos de frecuencias más altas sean los más amortiguados, al revés que aquí.

9.2. Dimensiones 2 y 2 + 1

Pasamos ahora a estudiar problemas con dos dimensiones espaciales, y el tiempo; incluiremos también problemas en dos dimensiones espaciales donde el tiempo no aparece. En este caso es productiva una breve discusión teórica previa.

Como hemos visto en las Secciones 8.2.2 y 8.2.3, las ecuaciones de difusión y de ondas, en su forma más simple, se escribirán

$$u_t - \kappa \nabla^2 u = 0, \quad (9.44a)$$

$$u_{tt} - c^2 \nabla^2 u = 0. \quad (9.44b)$$

Separando las partes espacial y temporal como

$$u(\vec{x}, t) = U(\vec{x})T(t) \quad (9.45)$$

tendremos entonces

$$\frac{1}{\kappa} \frac{T'}{T} = \frac{\nabla^2 U}{U} = -k^2, \quad (9.46a)$$

$$\frac{1}{c^2} \frac{T''}{T} = \frac{\nabla^2 U}{U} = -k^2, \quad (9.46b)$$

donde la notación $-k^2$ para la constante de separación se elige por conveniencia y no implica que k^2 sea positivo o siquiera real. Las EDOs para las partes temporales se escribirán entonces

$$T' + \kappa k^2 T = 0, \quad (9.47a)$$

$$T'' + c^2 k^2 T = 0, \quad (9.47b)$$

pero las EDPs para las partes espaciales tendrán la misma forma en ambos casos,

$$\nabla^2 U + k^2 U = 0. \quad (9.48)$$

Esta es conocida como la *ecuación de Helmholtz*, y si k es real, es una ecuación elíptica.

Las otras PDE elípticas que encontraremos frecuentemente son la *ecuación de Poisson*

$$\nabla^2 U = \rho, \quad (9.49)$$

y su versión homogénea, la *ecuación de Laplace*

$$\nabla^2 U = 0. \quad (9.50)$$

Ambas son comunes en problemas de electrostática, por ejemplo, donde U es un potencial escalar.

Vamos a estudiar ahora la separación de variables en las ecuaciones de Laplace y Helmholtz, en coordenadas Cartesianas y polares. La ecuación de Poisson reaparecerá en el Capítulo 11 en el marco más general de la resolución de EDP inhomogéneas, pero introduciremos aquí su tratamiento por el método de separación de variables. Antes de proceder, es conveniente establecer algunos resultados generales acerca de las soluciones de la ecuación de Laplace. Recorremos primero el siguiente

Teorema 9.2.1 (de la divergencia). *Sea $\Omega \subset \mathbb{R}^n$ compacto con frontera $\partial\Omega$ suave a trozos, y $\vec{\psi}(\vec{x})$ un campo vectorial $C^1[\Omega]$. Entonces*

$$\int_{\Omega} \nabla \cdot \vec{\psi} dV = \oint_{\partial\Omega} \vec{\psi} \cdot \hat{n} dS,$$

donde dV es el diferencial de volumen en Ω , y \hat{n} es la normal unitaria saliente y dS el diferencial de área en $\partial\Omega$.

La demostración puede consultarse en cualquier texto de Análisis Vectorial. Aplicaremos ahora el teorema de la divergencia para obtener los resultados mencionados.

Teorema 9.2.2. *El problema*

$$\begin{aligned} \nabla^2 U &= 0, & \vec{x} \in \Omega, \\ \hat{n} \cdot \nabla U &= f(\vec{x}), & \vec{x} \in \partial\Omega \end{aligned}$$

tiene solución sólo si

$$\oint_{\partial\Omega} f(\vec{x}) dS = 0.$$

Ejercicio 9.2.1. Demuestre el Teorema anterior. Ayuda: invoque el teorema de la divergencia con $\vec{\psi} = \nabla u$.

Teorema 9.2.3. Si

$$\begin{aligned}\nabla^2 U &= 0, \quad \vec{x} \in \Omega, \\ U(\vec{x}) &= 0, \quad \vec{x} \in \partial\Omega,\end{aligned}$$

entonces

$$U(\vec{x}) = 0, \quad \vec{x} \in \Omega.$$

Ejercicio 9.2.2. Demuestre el Teorema anterior. Ayuda: invoque el teorema de la divergencia con $\vec{\psi} = u\nabla u$ y use que $\nabla \cdot (u\nabla u) = \nabla u \cdot \nabla u + u\nabla^2 u$.

Corolario 9.2.4. La solución de la ecuación de Laplace con CC de Dirichlet sobre $\partial\Omega$ es única, y con CC de Neumann es única a menos de una constante aditiva.

Ejercicio 9.2.3. Demuestre el corolario anterior. Ayuda: considere dos soluciones u y v con idénticas CC, y estudie el problema para $w = u - v$.

9.2.1. La ecuación de Laplace en coordenadas Cartesianas

Consideremos el problema

$$\nabla^2 u = 0, \quad 0 < x < a, \quad 0 < y < b, \quad (9.51)$$

y busquemos soluciones de la forma

$$u(x, t) = X(x)Y(y). \quad (9.52)$$

Substituyendo tendremos

$$\frac{X''}{X} + \frac{Y''}{Y} = 0, \quad (9.53)$$

de donde es inmediato que X''/X e Y''/Y deben ser ambas constantes, con suma nula; sin embargo esto deja a nuestro arbitrio el signo de cada constante: sólo impone que sean opuestos.

Elijamos primero esta constante nula. En ese caso tendremos

$$X'' = 0, \quad Y'' = 0, \quad (9.54)$$

lo que nos lleva a soluciones lineales

$$X(x) = A + Bx, \quad Y(y) = C + Dy \quad (9.55)$$

y podemos, renombrando las constantes, escribir una solución

$$u_0(x, y) = \alpha + \beta x + \gamma y + \delta xy. \quad (9.56)$$

Elijamos ahora

$$X'' + k^2 X = 0, \quad Y'' - k^2 Y = 0 \quad (9.57)$$

con k real. Tendremos

$$X(x) = A \cos(kx) + B \sin(kx), \quad Y(y) = C \cosh(ky) + D \sinh(ky). \quad (9.58)$$

Elijamos finalmente

$$X'' - k^2 X = 0, \quad Y'' + k^2 Y = 0 \quad (9.59)$$

con k nuevamente real. Tendremos

$$X(x) = A \cosh(kx) + B \sinh(kx), \quad Y(y) = C \cos(ky) + D \sin(ky). \quad (9.60)$$

Veamos ahora cómo imponer CC a soluciones de este tipo. Consideremos primero el problema

$$\nabla^2 u = 0, \quad 0 < x < a, \quad 0 < y < b, \quad (9.61a)$$

$$u(0, y) = u(a, y) = u(x, 0) = 0, \quad u(x, b) = 1. \quad (9.61b)$$

Las CC homogéneas en $x = 0$ y $x = a$ deben ser satisfechas por $X(x)$, lo que nos dejará

$$X_n(x) = \sin(k_n x), \quad k_n = \frac{n\pi}{a}, \quad n \in \mathbb{N}. \quad (9.62)$$

Igualmente, la CC homogénea en $y = 0$ debe ser satisfecha por $Y(x)$, y notando que los posibles valores de k ya están fijados y deben ser alguno de los k_n , tendremos

$$Y_n(y) = \sinh(k_n y). \quad (9.63)$$

Hasta ahora nuestra solución se escribe

$$u(x, y) = \sum_{n=1}^{\infty} c_n \sin(k_n x) \sinh(k_n y), \quad (9.64)$$

y recién ahora vamos a imponerle las CC inhomogéneas en $y = b$, que implican

$$\sum_{n=1}^{\infty} c_n \sin(k_n x) \sinh(k_n b) = 1, \quad 0 < x < a. \quad (9.65)$$

La ortogonalidad y completitud de $\{\sin(k_n x)\}_{n=1}^{\infty}$ en $L^2(0, a)$ implica entonces que

$$c_n \sinh(k_n b) = \frac{\langle \sin(k_n x), 1 \rangle}{\langle \sin(k_n x), \sin(k_n x) \rangle} = \frac{2}{n\pi} (1 - (-1)^n), \quad (9.66)$$

y la solución queda finalmente

$$u(x, y) = \frac{2}{\pi} \sum_{n=1}^{\infty} \frac{1 - (-1)^n}{n} \sin(k_n x) \frac{\sinh(k_n y)}{\sinh(k_n b)}. \quad (9.67)$$

Notemos que si la CC en $y = b$ hubiese sido $u(x, b) = s(x)$, habríamos obtenido

$$c_n \sinh(k_n b) = \frac{2}{a} \langle \sin(k_n x), s(x) \rangle, \quad (9.68)$$

y la solución sería

$$u(x, y) = \frac{2}{a} \sum_{n=1}^{\infty} \langle \operatorname{sen}(k_n x), s(x) \rangle \operatorname{sen}(k_n x) \frac{\operatorname{senh}(k_n y)}{\operatorname{senh}(k_n b)}. \quad (9.69)$$

Consideremos ahora el problema con CC

$$u(0, y) = u(a, y) = u(x, b) = 0, \quad u(x, 0) = r(x). \quad (9.70)$$

Podríamos repetir el procedimiento anterior, pero es más sencillo notar que la solución debe ser la misma de recién, pero con $r(x)$ en el lugar de $s(x)$ y cambiando y por $b - y$, es decir

$$u(x, y) = \frac{2}{a} \sum_{n=1}^{\infty} \langle \operatorname{sen}(k_n x), r(x) \rangle \operatorname{sen}(k_n x) \frac{\operatorname{senh}(k_n(b-y))}{\operatorname{senh}(k_n b)}. \quad (9.71)$$

De manera completamente análoga podemos construir la solución del problema con CC

$$u(x, 0) = u(x, b) = u(0, y) = 0, \quad u(a, y) = q(y), \quad (9.72)$$

que será

$$u(x, y) = \frac{2}{b} \sum_{n=1}^{\infty} \langle \operatorname{sen}(k'_n y), q(y) \rangle \operatorname{sen}(k'_n y) \frac{\operatorname{senh}(k'_n x)}{\operatorname{senh}(k'_n a)}, \quad (9.73)$$

donde

$$k'_n = \frac{n\pi}{b}, \quad (9.74)$$

y la del problema con CC

$$u(x, 0) = u(x, b) = u(a, y) = 0, \quad u(0, y) = p(y), \quad (9.75)$$

que será

$$u(x, y) = \frac{2}{b} \sum_{n=1}^{\infty} \langle \operatorname{sen}(k'_n y), p(y) \rangle \operatorname{sen}(k'_n y) \frac{\operatorname{senh}(k'_n(a-x))}{\operatorname{senh}(k'_n a)}. \quad (9.76)$$

Y es inmediato que ahora podemos construir la solución del problema con CC

$$u(0, y) = p(y), \quad u(a, y) = q(y), \quad u(x, 0) = r(x), \quad u(x, b) = s(x) \quad (9.77)$$

simplemente sumando las cuatro.

Este último procedimiento es completamente general: cuando tengamos un problema con CC inhomogéneas sobre más de una curva o superficie frontera, podemos aplicar la linealidad de la EDP y las CC para dividirlo en subproblemas, cada uno con CC inhomogéneas sobre una única curva o superficie frontera y homogéneas sobre las demás, para después sumar las respectivas soluciones.

Ejercicio 9.2.4. Halle el estado estacionario $u(\vec{x})$ del problema de difusión del calor

$$\begin{aligned} u_t - \kappa \nabla^2 u &= 0, & 0 < x < a, \quad 0 < y < a, \\ u(0, y) &= a - y, \quad u(x, 0) = a - x, \quad u(x, a) = u(a, y) = 0. \end{aligned}$$

La extensión de lo realizado a CC de Neumann, de Robin y mixtas se deja como ejercicio para el lector.

Ejercicio 9.2.5. Halle el estado estacionario $u(\vec{x})$ del problema de difusión del calor

$$\begin{aligned} u_t - \kappa \nabla^2 u &= 0, & 0 < x < a, \quad 0 < y < a, \\ u(0, y) &= u(a, y) = 0, \quad u(x, 0) = 1 \quad u_x(x, a) = 0. \end{aligned}$$

9.2.2. La ecuación de Laplace en coordenadas polares

Consideremos el problema

$$\nabla^2 u = 0, \quad 0 \leq a < r < b, \quad 0 \leq \theta < \alpha, \quad (9.78)$$

donde r y θ son coordenadas polares planas, y busquemos soluciones de la forma

$$u(r, \theta) = R(r)Q(\theta). \quad (9.79)$$

Recordando que en coordenadas polares

$$\nabla^2 = \frac{\partial^2}{\partial r^2} + \frac{1}{r} \frac{\partial}{\partial r} + \frac{1}{r^2} \frac{\partial^2}{\partial \theta^2},$$

tendremos

$$R''Q + \frac{1}{r}R'Q + \frac{1}{r^2}RQ'' = 0,$$

que multiplicando por $r^2/(RQ)$ queda

$$\frac{r^2 R''}{R} + \frac{r R'}{R} = -\frac{Q''}{Q} = \nu^2,$$

donde hemos elegido llamar ν^2 a la constante de separación sin que ello implique, como siempre, que sea positiva o real. Obtenemos así las EDOs separadas

$$Q'' + \nu^2 Q = 0, \quad (9.80a)$$

$$r^2 R'' + r R' - \nu^2 R = 0. \quad (9.80b)$$

Vemos que (9.80a) es una ecuación armónica, y su solución general será⁵

$$Q(\theta) = \left\{ \begin{array}{l} \cos(\nu\theta) \\ \sin(\nu\theta) \end{array} \right\} = \{e^{\pm i\nu\theta}\}$$

⁵Recordemos que

$$\left\{ \begin{array}{l} \cos(\nu\theta) \\ \sin(\nu\theta) \end{array} \right\}$$

significa “cualquier combinación lineal de $\cos(\nu\theta)$ y $\sin(\nu\theta)$ ”.

si $\nu \neq 0$, y

$$Q(\theta) = \begin{cases} 1 \\ \theta \end{cases}$$

si $\nu = 0$. Por otro lado, para $\nu \neq 0$ (9.80b) es una *ecuación de Euler*, cuya solución general es

$$R(r) = \begin{cases} r^\nu \\ r^{-\nu} \end{cases} = \{r^{\pm\nu}\},$$

mientras que para $\nu = 0$ es

$$R(r) = \begin{cases} 1 \\ \ln r \end{cases}.$$

Las soluciones de (9.78) tienen entonces la forma

$$u(r, \theta) = \begin{cases} r^\nu \\ r^{-\nu} \end{cases} \begin{cases} \cos(\nu\theta) \\ \sin(\nu\theta) \end{cases} (+) \begin{cases} 1 \\ \ln r \end{cases} \begin{cases} 1 \\ \theta \end{cases}. \quad (9.81)$$

Ahora debemos imponer CC a estas soluciones. Un caso típico es el problema

$$\nabla^2 u = 0, \quad 0 < r < b, \quad 0 \leq \theta < 2\pi, \quad (9.82a)$$

$$\lim_{r \downarrow 0} |u(r, \theta)| < \infty, \quad u(b, \theta) = f(\theta), \quad (9.82b)$$

$$u(r, 0) = u(r, 2\pi), \quad u_\theta(r, 0) = u_\theta(r, 2\pi).$$

Para $Q(\theta)$ tendremos el PSL

$$Q'' + \nu^2 Q = 0, \quad (9.83a)$$

$$Q(0) = Q(2\pi), \quad Q'(0) = Q'(2\pi). \quad (9.83b)$$

Las CC (9.83b) imponen $\nu = n \in \mathbb{Z}$, como ya vimos en la Sección 6.1; luego $Q(\theta)$ deberá ser alguna de

$$Q_n(\theta) = \begin{cases} \cos(n\theta) \\ \sin(n\theta) \end{cases}, \quad n \in \mathbb{N}_0. \quad (9.84)$$

Por otro lado, la CC (9.82b) de función finita debe ser satisfecha por la correspondiente $R_n(r)$, lo que implica que debemos descartar r^{-n} y $\ln r$. Tenemos entonces

$$u_n(r, \theta) = r^n (A_n \cos(n\theta) + B_n \sin(n\theta)), \quad (9.85)$$

y la solución deberá tener la forma⁶

$$u(r, \theta) = \sum_{n=0}^{\infty} r^n (A_n \cos(n\theta) + B_n \sin(n\theta)). \quad (9.86)$$

⁶Para una notación más concisa incluimos el término $n = 0$ en el caso general $n > 0$, lo que es posible en este caso pero *no siempre*.

Recién ahora, una vez ajustadas todas las CC (9.82b) *homogéneas*, vamos a imponer la CC inhomogénea pidiendo que

$$u(b, \theta) = \sum_{n=0}^{\infty} b^n (A_n \cos(n\theta) + B_n \sin(n\theta)) = f(\theta). \quad (9.87)$$

La ortogonalidad y completitud de las autofunciones angulares en $\mathcal{L}^2(0, 2\pi)$ nos permite hallar las constantes A_n y B_n a partir de esta ecuación. Por simplicidad vamos a suponer que $f(\theta) = \sin(\theta)$; luego por ortogonalidad es inmediato que

$$A_n \equiv 0, \quad B_n = \begin{cases} 1/b, & n = 1, \\ 0, & n \neq 1, \end{cases}$$

y la solución quedará

$$u(r, \theta) = \frac{r}{b} \sin(\theta). \quad (9.88)$$

Ejercicio 9.2.6. Halle la solución del problema anterior si $f(\theta) = \cos^2(\theta)$, y si

$$f(\theta) = \begin{cases} 1, & 0 < \theta < \pi, \\ -1, & \pi < \theta < 2\pi. \end{cases}$$

9.2.3. La ecuación de Helmholtz en coordenadas Cartesianas

Como vimos al principio de la Sección 9.2, la ecuación de Helmholtz (9.48) suele aparecer como la EDP satisfecha por la parte espacial de una función incógnita, luego de la separación inicial entre ésta y la parte temporal; por ello mismo, dependiendo del problema, la constante de separación $-k^2$ puede no ser ya completamente arbitraria cuando planteemos la ecuación de Helmholtz. Recordemos que, por ser en general una EDP elíptica, no necesitará ni admitirá condiciones iniciales, sólo CC.

Vamos a ilustrar su tratamiento en coordenadas Cartesianas mediante un ejemplo típico, el problema

$$\nabla^2 U + k^2 U = 0, \quad 0 < x < a, \quad 0 < y < b, \quad (9.89a)$$

$$U(0, y) = U(a, y) = U(x, 0) = U(x, b) = 0. \quad (9.89b)$$

Proponiendo

$$U(x, y) = X(x)Y(y)$$

y substituyendo en la EDP, tendremos

$$\frac{X''}{X} + \frac{Y''}{Y} + k^2 = 0,$$

de donde el argumento usual muestra que tanto X''/X como Y''/Y deben ser constantes cuya suma sea $-k^2$. En este caso resulta ventajoso introducir dos

constants auxiliares k_x y k_y y escribir

$$\frac{X''}{X} = -k_x^2, \quad \frac{Y''}{Y} = -k_y^2, \quad k_x^2 + k_y^2 = k^2.$$

Tenemos así los PSL separados

$$X'' + k_x^2 X = 0, \quad 0 < x < a, \quad (9.90a)$$

$$X(0) = X(a) = 0 \quad (9.90b)$$

y

$$Y'' + k_y^2 Y = 0, \quad 0 < y < b, \quad (9.91a)$$

$$Y(0) = Y(b) = 0. \quad (9.91b)$$

Ambos PSL son ecuaciones armónicas, cuyas soluciones ya conocemos, así que podemos escribir directamente

$$X(x) = \sum_{n=1}^{\infty} a_n \operatorname{sen}(k_{xn}x), \quad k_{xn} = \frac{n\pi}{a}, \quad (9.92a)$$

$$Y(y) = \sum_{m=1}^{\infty} b_m \operatorname{sen}(k_{ym}y), \quad k_{ym} = \frac{m\pi}{b} \quad (9.92b)$$

o, más compactamente,

$$U(x, y) = \sum_{n,m=1}^{\infty} c_{nm} U_{nm}(x, y) \quad (9.93)$$

con

$$U_{nm}(x, y) = \operatorname{sen}(k_{xn}x) \operatorname{sen}(k_{ym}y) \quad (9.94)$$

las *autofunciones espaciales*, también llamadas *modos espaciales*. Estas son autofunciones del operador Laplaciano, ya que

$$\nabla^2 U_{nm} = -(k_{xn}^2 + k_{ym}^2) U_{nm} = -k_{nm}^2 U_{nm}. \quad (9.95)$$

Notemos que si, por ejemplo, $b = a$, algunos de los autovalores k_{nm}^2 de ∇^2 serán *degenerados*, ya que tendremos $k_{nm}^2 = k_{mn}^2$ pero las autofunciones U_{nm} y U_{mn} serán diferentes; en general, puede mostrarse que siempre que $\frac{b}{a} \in \mathbb{Q}$ existirán autovalores degenerados.

Ejercicio 9.2.7. Resuelva la ecuación de Helmholtz en el rectángulo $[0, a] \times [0, b]$ con las CC

$$U(0, y) = U(a, y) = U_y(x, 0) = U_y(x, b) = 0.$$

¿Existen autovalores degenerados?

9.2.4. La ecuación de Helmholtz en coordenadas polares

El procedimiento para separar la ecuación de Helmholtz en coordenadas polares es muy similar al seguido en coordenadas Cartesianas. Nuevamente lo ilustramos con un ejemplo típico, en este caso el problema

$$\nabla^2 U + k^2 U = 0, \quad 0 < r < b, \quad 0 \leq \theta < 2\pi, \quad (9.96a)$$

$$\begin{aligned} \lim_{r \downarrow 0} |u(r, \theta)| &< \infty, & u(b, \theta) &= 0, \\ u(r, 0) &= u(r, 2\pi), & u_\theta(r, 0) &= u_\theta(r, 2\pi). \end{aligned} \quad (9.96b)$$

Proponiendo

$$U(r, \theta) = R(r)Q(\theta)$$

y substituyendo en la EDP, tendremos, siguiendo el mismo procedimiento de la Sección (9.2.2),

$$\frac{r^2 R''}{R} + \frac{r R'}{R} + k^2 r^2 = -\frac{Q''}{Q} = \nu^2,$$

lo que nos lleva, con las CC (9.96b), a los PSL separados

$$r R'' + R' + \left(k^2 r - \frac{\nu^2}{r} \right) R = 0, \quad 0 < r < b, \quad (9.97a)$$

$$\lim_{r \downarrow 0} |R(r)| < \infty, \quad R(b) = 0 \quad (9.97b)$$

y

$$Q'' + \nu^2 Q = 0, \quad 0 \leq \theta < 2\pi, \quad (9.98a)$$

$$Q(0) = Q(2\pi), \quad Q_\theta(0) = Q_\theta(2\pi). \quad (9.98b)$$

Es más conveniente resolver primero el PSL (9.98), ya que ello fijará los posibles valores de ν . La EDO (9.98a) es una ecuación armónica, y dadas las CC periódicas (9.98b) tendremos $\nu = n$ entero y

$$Q(\theta) = \sum_{n=0}^{\infty} (A_n \cos(n\theta) + B_n \sin(n\theta)). \quad (9.99)$$

El mismo $\nu = n$ entero es el que aparece ahora en la EDO (9.97a), de modo que sus soluciones para cada n serán

$$R_n(r) = C_n J_n(kr) + D_n Y_n(kr).$$

Imponiendo las CC (9.97b), la de función finita en $r = 0$ requiere $D_n = 0$; por lo tanto en $r = b$ tendremos

$$J_n(kb) = 0$$

y en consecuencia

$$R_n(r) = \sum_{m=1}^{\infty} C_m J_n \left(\frac{\kappa_{nm} r}{b} \right) = \sum_{m=1}^{\infty} C_m J_n(k_{nm} r), \quad k_{nm} = \frac{\kappa_{nm}}{b}, \quad (9.100)$$

donde κ_{nm} es el m -ésimo cero de J_n ,

$$J_n(\kappa_{nm}) = 0, \quad m = 1, 2, 3, \dots$$

Juntando (9.99) y (9.100) y redefiniendo las constantes de las respectivas combinaciones lineales, entonces, arribamos a la solución general del problema (9.96),

$$U(r, \theta) = \sum_{n=0}^{\infty} \sum_{m=1}^{\infty} J_n(k_{nm}r) (A_{nm} \cos(n\theta) + B_{nm} \sin(n\theta)). \quad (9.101)$$

Vemos que los *modos espaciales*

$$U_{nm}^c(r, \theta) = J_n(k_{nm}r) \cos(n\theta), \quad m = 1, 2, \dots, \quad n = 0, 1, 2, \dots, \quad (9.102a)$$

$$U_{nm}^s(r, \theta) = J_n(k_{nm}r) \sin(n\theta), \quad m = 1, 2, \dots, \quad n = 1, 2, \dots, \quad (9.102b)$$

son efectivamente autofunciones del operador Laplaciano de autovalor k_{nm} , ya que, por ejemplo,

$$\begin{aligned} \nabla^2 U_{nm}^c &= \left(\frac{\partial^2}{\partial r^2} + \frac{1}{r} \frac{\partial}{\partial r} + \frac{1}{r^2} \frac{\partial^2}{\partial \theta^2} \right) J_n(k_{nm}r) \cos(n\theta) \\ &= \frac{1}{r} \left(r J_n''(k_{nm}r) + J_n'(k_{nm}r) - \frac{n^2}{r} J_n(k_{nm}r) \right) \cos(n\theta), \end{aligned}$$

y usando que $J_n(k_{nm}r)$ es solución de (9.97a), resulta

$$\nabla^2 U_{nm}^c = \frac{1}{r} (k_{nm}^2 r J_n(k_{nm}r)) \cos(n\theta) = -k_{nm}^2 U_{nm}^c. \quad (9.103a)$$

Un cálculo similar muestra que también

$$\nabla^2 U_{nm}^s = -k_{nm}^2 U_{nm}^s. \quad (9.103b)$$

¿Qué ocurre si cambiamos las CC (9.96b) por otras diferentes? Bien, en esencia nada, sólo detalles. Supongamos por ejemplo que en (9.96) el intervalo radial fuese ahora $a < r < b$ con $U(a, \theta) = 0$; entonces en vez del PSL (9.97) tendríamos el PSL (6.87), y en lugar de $J_n(k_{nm}r)$ tendríamos

$$R_{nm}(r) = Y_n(k_{nm}a) J_n(k_{nm}r) - J_n(k_{nm}a) Y_n(k_{nm}r), \quad m = 1, 2, 3, \dots,$$

con los autovalores k_{nm}^2 determinados por las soluciones de la ecuación

$$Y_n(k_{nm}a) J_n(k_{nm}b) - J_n(k_{nm}a) Y_n(k_{nm}b) = 0.$$

Y si en (9.96) el intervalo angular se cambiase a $0 < \theta < \alpha < 2\pi$ con $U(r, 0) = U(r, \alpha) = 0$, en lugar de $\cos(n\theta)$ y $\sin(n\theta)$ tendríamos

$$\sin(\nu_n \theta), \quad \nu_n = \frac{n\pi}{\alpha}, \quad n = 1, 2, 3, \dots,$$

y además

$$R_{nm}(r) = Y_{\nu_n}(k_{nm}a)J_{\nu_n}(k_{nm}r) - J_{\nu_n}(k_{nm}a)Y_{\nu_n}(k_{nm}r), \quad m = 1, 2, 3, \dots,$$

con los autovalores k_{nm}^2 determinados ahora por las soluciones de la ecuación

$$Y_{\nu_n}(k_{nm}a)J_{\nu_n}(k_{nm}b) - J_{\nu_n}(k_{nm}a)Y_{\nu_n}(k_{nm}b) = 0.$$

Etc.

9.2.5. La ecuación de Poisson en coordenadas polares

A diferencia de las ecuaciones de Laplace y Helmholtz, la ecuación de Poisson es *inhomogénea*, lo que *a priori* parecería impedir la utilización del método de separación de variables. Sin embargo, si notamos que la versión homogénea de la ecuación de Poisson no es más que la ecuación de Laplace, y recordamos que por linealidad la solución general de una ecuación inhomogénea

$$Lu = \rho$$

es la solución general de la homogénea

$$Lu_h = 0$$

mas una solución particular cualquiera de la inhomogénea

$$Lu_p = \rho,$$

tenemos un principio de solución. La técnica consistirá en separar el problema inhomogéneo para u con dadas CC en la superposición de: (a) un problema con EDP inhomogénea cuya solución u_p buscaremos desarrollándola, al igual que la inhomogeneidad, en autofunciones del operador, pero con CC no especificadas, mas (b) un problema con EDP homogénea para u_h con las CC adecuadas para que $u_h + u_p$ cumpla con las CC de u . Ilustraremos esta técnica con un ejemplo en coordenadas polares; la aplicación en coordenadas Cartesianas se deja como ejercicio para el lector.

Consideremos el problema

$$\nabla^2 u = \frac{r^2}{b^4} \operatorname{sen}(\theta), \quad 0 < r < b, \quad 0 \leq \theta < 2\pi, \quad (9.104a)$$

$$\begin{aligned} \lim_{r \downarrow 0} |u(r, \theta)| &< \infty, & u(b, \theta) &= \cos(\theta), \\ u(r, 0) &= u(r, 2\pi), & u_\theta(r, 0) &= u_\theta(r, 2\pi). \end{aligned} \quad (9.104b)$$

Vamos a descomponerlo en dos subproblemas, escribiendo $u = u_h + u_p$.

El primer subproblema será

$$\nabla^2 u_p = \frac{r^2}{b^4} \operatorname{sen}(\theta), \quad 0 < r < b, \quad 0 \leq \theta < 2\pi, \quad (9.105)$$

donde en principio no inponemos CC. En la Sección 9.2.2 vimos que las *autofunciones angulares* del operador ∇^2 son

$$\begin{Bmatrix} \cos(\nu\theta) \\ \sin(\nu\theta) \end{Bmatrix}$$

y, dado que la inhomogeneidad de (9.105) contiene sólo valores enteros de ν , asumiremos $\nu = n$ entero; y como sólo contiene *senos*, asumiremos que sólo tendremos $\sin(n\theta)$. También vimos que las *autofunciones radiales* del operador ∇^2 son

$$\begin{Bmatrix} r^\mu \\ r^{-\mu} \end{Bmatrix}$$

y, como la inhomogeneidad de (9.105) contiene sólo valores enteros positivos de μ , asumiremos $\mu = m$, entero positivo (o cero), y que sólo tendremos r^m . Sin embargo, a diferencia de lo que hicimos en la Sección 9.2.2, aquí *no exigiremos que $m = n$* , ya que *no queremos obtener autofunciones de ∇^2 de autovalor nulo*. En cambio propondremos para u_p un desarrollo de la forma

$$u_p(r, \theta) = \left(\sum_{m=0}^{\infty} A_m r^m \right) \left(\sum_{n=1}^{\infty} B_n \sin(n\theta) \right) = \sum_{m=0}^{\infty} \sum_{n=1}^{\infty} C_{mn} r^m \sin(n\theta).$$

Es fácil verificar que

$$\nabla^2 r^m \sin(n\theta) = \left(\frac{\partial^2}{\partial r^2} + \frac{1}{r} \frac{\partial}{\partial r} + \frac{1}{r^2} \frac{\partial^2}{\partial \theta^2} \right) r^m \sin(n\theta) = (m^2 - n^2) r^{m-2} \sin(n\theta),$$

de donde, reemplazando en (9.105), tendremos

$$\sum_{m=0}^{\infty} \sum_{n=1}^{\infty} C_{mn} (m^2 - n^2) r^{m-2} \sin(n\theta) = \frac{r^2}{b^4} \sin(\theta).$$

Finalmente, usando la ortogonalidad y completitud de las autofunciones radiales y angulares en los respectivos dominios, podemos igualar coeficientes de $r^m \sin(n\theta)$ entre ambos miembros, lo que nos deja

$$C_{41} = \frac{1}{15b^4}, \quad C_{mn} = 0, \quad m \neq 4, \quad n \neq 1$$

y

$$u_p(r, \theta) = \frac{1}{15b^4} r^4 \sin(\theta). \quad (9.106)$$

Con lo anterior, el segundo subproblema, para u_h , queda

$$\nabla^2 u_h = 0, \quad 0 < r < b, \quad 0 \leq \theta < 2\pi, \quad (9.107a)$$

$$\lim_{r \downarrow 0} |u_h(r, \theta)| < \infty, \quad u_h(b, \theta) = \cos(\theta) - \frac{1}{15} \sin(\theta), \quad (9.107b)$$

$$u_h(r, 0) = u_h(r, 2\pi), \quad u_{h,\theta}(r, 0) = u_{h,\theta}(r, 2\pi),$$

que ya sabemos resolver por las técnicas de la Sección 9.2.2.

Ejercicio 9.2.8. Complete los cálculos y halle la solución del problema (9.104).

9.3. Dimensiones 3 y 3 + 1

Estudiaremos ahora problemas con tres dimensiones espaciales más el tiempo, y problemas en tres dimensiones espaciales donde el tiempo no aparece. Cabe destacar que la separación de la parte temporal de las ecuaciones de ondas y difusión procederá exactamente como se discutió al principio de la Sección 9.2 y llevará a la ecuación de Helmholtz, por lo que no la repetiremos aquí. También debemos destacar que la separación de las ecuaciones de Laplace y Helmholtz en coordenadas Cartesianas sigue el mismo procedimiento de las Secciones 9.2.1 y 9.2.3 con modificaciones triviales, por lo que omitiremos su discusión en bien de la brevedad. Finalmente, y de nuevo por concisión, muchos detalles del proceso de separación de variables serán omitidos, a menos que presenten diferencias significativas con lo ya visto en las Secciones 9.1 y 9.2.

9.3.1. La ecuación de Laplace en coordenadas cilíndricas

Consideremos el problema

$$\nabla^2 u = 0, \quad 0 \leq a < r < b, \quad 0 \leq \theta < \alpha, \quad 0 < z < \ell, \quad (9.108)$$

donde r , θ y z son coordenadas cilíndricas, y busquemos soluciones de la forma

$$u(r, \theta, z) = R(r)Q(\theta)Z(z). \quad (9.109)$$

Recordando que en coordenadas cilíndricas

$$\nabla^2 = \frac{\partial^2}{\partial r^2} + \frac{1}{r} \frac{\partial}{\partial r} + \frac{1}{r^2} \frac{\partial^2}{\partial \theta^2} + \frac{\partial^2}{\partial z^2}, \quad (9.110)$$

tendremos una primera separación para Z ,

$$\frac{R''}{R} + \frac{1}{r} \frac{R'}{R} + \frac{1}{r^2} \frac{Q''}{Q} = -\frac{Z''}{Z} = \pm \nu^2,$$

y una segunda separación para R y Q ,

$$r^2 \frac{R''}{R} + r \frac{R'}{R} \mp r^2 \nu^2 = -\frac{Q''}{Q} = \pm \mu^2,$$

donde la elección de los signos de $\pm \nu^2$ y $\pm \mu^2$, que *suponemos reales*, deberá realizarse en base a las CC (homogéneas o inhomogéneas) que esperemos tener que imponer en diferentes superficies coordinadas.

Por ejemplo, eligiendo $-\nu^2$ y $+\mu^2$ obtenemos las EDOs separadas

$$Z'' - \nu^2 Z = 0, \quad (9.111a)$$

$$Q'' + \mu^2 Q = 0, \quad (9.111b)$$

$$r^2 R'' + r R' + (\nu^2 r^2 - \mu^2) R = 0. \quad (9.111c)$$

Dado que (9.111a) y (9.111b) son ecuaciones armónicas, y (9.111c) es la ecuación de Bessel (6.66a), las correspondientes soluciones serán

$$Z(z) = \{e^{\pm\nu z}\} = \begin{cases} \cosh(\nu z) \\ \operatorname{senh}(\nu z) \end{cases}, \quad (9.112\text{a})$$

$$Q(\theta) = \{e^{\pm i\mu\theta}\} = \begin{cases} \cos(\mu\theta) \\ \operatorname{sen}(\mu\theta) \end{cases}, \quad (9.112\text{b})$$

$$R(r) = \begin{cases} J_\mu(\nu r) \\ Y_\mu(\nu r) \end{cases}. \quad (9.112\text{c})$$

El carácter oscilatorio de $R(r)$ y $Q(\theta)$, y no oscilatorio de $Z(z)$, hará que estas soluciones se adapten bien a CC homogéneas para valores dados de r y θ , e inhomogéneas para valores dados de z .

En cambio, eligiendo $+\nu^2$ y $+\mu^2$ obtenemos las EDOs separadas

$$Z'' + \nu^2 Z = 0, \quad (9.113\text{a})$$

$$Q'' + \mu^2 Q = 0, \quad (9.113\text{b})$$

$$r^2 R'' + r R' - (\nu^2 r^2 + \mu^2) R = 0. \quad (9.113\text{c})$$

Dado que (9.113a) y (9.113b) son de nuevo ecuaciones armónicas, pero (9.113c) es ahora la ecuación de Bessel modificada (6.88), las correspondientes soluciones serán

$$Z(z) = \{e^{\pm i\nu z}\} = \begin{cases} \cos(\nu z) \\ \operatorname{sen}(\nu z) \end{cases}, \quad (9.114\text{a})$$

$$Q(\theta) = \{e^{\pm i\mu\theta}\} = \begin{cases} \cos(\mu\theta) \\ \operatorname{sen}(\mu\theta) \end{cases}, \quad (9.114\text{b})$$

$$R(r) = \begin{cases} I_\mu(\nu r) \\ K_\mu(\nu r) \end{cases}. \quad (9.114\text{c})$$

El carácter oscilatorio de $Q(\theta)$ y $Z(z)$, y no oscilatorio de $R(r)$, hará que estas soluciones se adapten bien a CC homogéneas para valores dados de θ y z , e inhomogéneas para valores dados de r .

Basados en lo anterior, si ahora planeamos imponer CC homogéneas para valores dados de r y z , e inhomogéneas para valores dados de θ , deberíamos elegir $+\nu^2$ y $-\mu^2$, lo que nos llevaría a las EDOs separadas

$$Z'' + \nu^2 Z = 0, \quad (9.115\text{a})$$

$$Q'' - \mu^2 Q = 0, \quad (9.115\text{b})$$

$$r^2 R'' + r R' - (\nu^2 r^2 - \mu^2) R = 0. \quad (9.115\text{c})$$

Vemos que (9.115a) y (9.115b) son otra vez ecuaciones armónicas, pero (9.115c) es la ecuación de Bessel modificada de índice *imaginario*,⁷ así que las corres-

⁷Véase por ejemplo <https://dlmf.nist.gov/10.45> y referencias citadas allí.

pondientes soluciones serán

$$Z(z) = \{e^{\pm i\nu z}\} = \begin{cases} \cos(\nu z) \\ \sin(\nu z) \end{cases}, \quad (9.116a)$$

$$Q(\theta) = \{e^{\pm \mu\theta}\} = \begin{cases} \cosh(\mu\theta) \\ \operatorname{senh}(\mu\theta) \end{cases}, \quad (9.116b)$$

$$R(r) = \begin{cases} I_{i\mu}(\nu r) \\ K_{i\mu}(\nu r) \end{cases}. \quad (9.116c)$$

Las funciones de Bessel de orden imaginario se conocen desde hace casi siglo y medio,⁸ pero hasta tiempos bastante recientes eran consideradas de interés mayormente académico. El autor confiesa que comparte la probable incredulidad del lector ante la prácticamente nula mención a estas funciones en la literatura de, por ejemplo, problemas de electrostática, y en general en los textos de métodos matemáticos de la Física, pese a que ya a mediados del siglo pasado fueron usadas por los matemáticos soviéticos M. Kantorovich y N. Lebedev precisamente para hallar soluciones de la ecuación de Laplace por separación de variables.⁹ Recientemente estas funciones han concitado interés en el contexto de la dinámica de partículas cargadas en cuadrupolos de isotrayectoria.¹⁰

Cabe agregar el mismo *caveat* que para la separación en dos dimensiones: si bien no lo implementamos por concisión, el lector *siempre* debe tener presente la posibilidad de que alguna de las constantes de separación se anule, con el consiguiente cambio de carácter de las EDO separadas y sus soluciones.

9.3.2. La ecuación de Helmholtz en coordenadas cilíndricas

Podríamos pensar que la separación de la ecuación de Helmholtz en coordenadas cilíndricas debería ser muy semejante a la de la ecuación de Laplace; sin embargo, veremos que la presencia del término $k^2 u$ introduce diferencias considerables. El problema se escribirá

$$\nabla^2 u + k^2 u = 0, \quad 0 \leq a < r < b, \quad 0 \leq \theta < \alpha, \quad 0 < z < \ell, \quad (9.117)$$

y buscaremos soluciones de la forma

$$u(r, \theta, z) = R(r)Q(\theta)Z(z). \quad (9.118)$$

⁸M. Bôcher, *On Some Applications of Bessel's Functions with Pure Imaginary Index*, Annals of Mathematics **6**(6) (May, 1892), pp. 137–160. <https://www.jstor.org/stable/1967335>

⁹M. Kantorovich y N. Lebedev, en *Selected Problems of Mathematical Theory of Electromagnetic Phenomena*, editado por G. G. Abramovitch. Publicación de la Academia de Ciencias, Moscú-Leningrado, 1948 (en ruso).

¹⁰A. A. Matyshev and E. Fohtung, *On the Computation and Applications of Bessel Functions with Pure Imaginary Indices*, <https://arxiv.org/abs/0910.0365>

Matyshev, A. A., Fohtung, E. B. (2014), *On the computation and applications of bessel functions with pure imaginary indices (orders) in physics and specifically in corpuscular optics* **24**(1). <https://cyberleninka.ru/article/n/on-the-computation-and-applications-of-bessel-functions-with-pure-imaginary-indices-orders-in-physics-and-specifically-in-corpuscular>

Nuevamente recordando la expresión (9.110) del Laplaciano en coordenadas cilíndricas, tendremos

$$\frac{R''}{R} + \frac{1}{r} \frac{R'}{R} + \frac{1}{r^2} \frac{Q''}{Q} + \frac{Z''}{Z} + k^2 = 0,$$

con lo que una primera separación para Z nos dará

$$\frac{R''}{R} + \frac{1}{r} \frac{R'}{R} + \frac{1}{r^2} \frac{Q''}{Q} + k^2 = -\frac{Z''}{Z} = \pm \nu^2,$$

y una segunda separación para R y Q dejará

$$r^2 \frac{R''}{R} + r \frac{R'}{R} + (k^2 \mp \nu^2)r^2 = -\frac{Q''}{Q} = \pm \mu^2.$$

donde la elección de los signos de $\pm \nu^2$ y $\pm \mu^2$, que seguimos suponiendo *reales*, deberá realizarse nuevamente en base a las CC (homogéneas o inhomogéneas) que esperemos tener que imponer en diferentes superficies coordinadas.

A diferencia de la ecuación de Laplace, que por el Teorema 9.2.3 sólo tiene solución trivial para CC de Dirichlet homogéneas, y por el Corolario 9.2.4 sólo constante para CC de Neumann homogéneas, la ecuación de Helmholtz tiene en general soluciones no triviales ni constantes aún para CC homogéneas. Para obtenerlas debemos elegir, en principio, los signos $+\nu^2$ y $+\mu^2$, con lo que obtenemos las EDO separadas

$$Z'' + \nu^2 Z = 0, \quad (9.119a)$$

$$Q'' + \mu^2 Q = 0, \quad (9.119b)$$

$$r^2 R'' + r R' + [(k^2 - \nu^2)r^2 - \mu^2]R = 0. \quad (9.119c)$$

Dado que (9.119a) y (9.119b) son ecuaciones armónicas, y (9.119c) es la ecuación de Bessel, las correspondientes soluciones serán

$$Z(z) = \{e^{\pm i\nu z}\} = \begin{cases} \cos(\nu z) \\ \sin(\nu z) \end{cases}, \quad (9.120a)$$

$$Q(\theta) = \{e^{\pm i\mu\theta}\} = \begin{cases} \cos(\mu\theta) \\ \sin(\mu\theta) \end{cases}, \quad (9.120b)$$

$$R(r) = \left\{ J_\mu \left(\sqrt{k^2 - \nu^2} r \right) \right\}. \quad (9.120c)$$

Notemos sin embargo el factor $\sqrt{k^2 - \nu^2}$ en el argumento de J_μ e Y_μ ; este hará que si $\nu^2 > k^2$ el argumento se vuelva imaginario puro, con lo que las funciones radiales pasarán a ser en efecto I_μ y K_μ , perdiendo así su carácter oscilatorio. Este efecto es relevante, entre otros casos, en la propagación de ondas en una *guía de ondas*, de lo cual daremos un ejemplo en la Sección 9.4.

Otras elecciones de signos para ν^2 y μ^2 pueden tratarse sobre las mismas líneas de la Sección 9.3.1, y se dejan como ejercicio para el lector.

9.3.3. La ecuación de Laplace en coordenadas esféricas

Consideremos el problema

$$\nabla^2 u = 0, \quad 0 \leq a < r < b, \quad 0 \leq \alpha < \theta < \beta \leq \pi, \quad 0 < \phi < \gamma \leq 2\pi, \quad (9.121)$$

donde r , θ y ϕ son coordenadas esféricas, y busquemos soluciones de la forma

$$u(r, \theta, z) = R(r)Q(\theta)F(\phi). \quad (9.122)$$

Recordando que en coordenadas esféricas

$$\nabla^2 = \frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial}{\partial r} \right) + \frac{1}{r^2 \sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial}{\partial \theta} \right) + \frac{1}{r^2 \sin^2 \theta} \frac{\partial^2}{\partial \phi^2}, \quad (9.123)$$

aplicándolo a RQF y dividiendo por RQF tenemos

$$\frac{1}{Rr^2} \frac{d}{dr} \left(r^2 \frac{dR}{dr} \right) + \frac{1}{Qr^2 \sin \theta} \frac{d}{d\theta} \left(\sin \theta \frac{dQ}{d\theta} \right) + \frac{1}{Fr^2 \sin^2 \theta} \frac{d^2 F}{d\phi^2} = 0.$$

Multiplicando por $r^2 \sin^2 \theta$ tenemos una primera separación para F ,

$$\sin^2 \theta \frac{1}{R} \frac{d}{dr} \left(r^2 \frac{dR}{dr} \right) + \sin \theta \frac{1}{Q} \frac{d}{d\theta} \left(\sin \theta \frac{dQ}{d\theta} \right) = -\frac{F''}{F} = \mu^2,$$

y volviendo a dividir por $\sin^2 \theta$ obtenemos una segunda separación para R y Q ,

$$\frac{1}{R} \frac{d}{dr} \left(r^2 \frac{dR}{dr} \right) = -\frac{1}{Q \sin \theta} \frac{d}{d\theta} \left(\sin \theta \frac{dQ}{d\theta} \right) + \frac{\mu^2}{\sin^2 \theta} = \nu(\nu + 1).$$

Por simplicidad, esta vez hemos omitido explicitar un posible doble signo de μ como $\pm \mu^2$, pero debemos recordar que, en principio, μ puede ser imaginario; consideraciones equivalentes valen para ν . Como siempre, estas elecciones deberán realizarse en base a las CC (homogéneas o inhomogéneas) que esperemos tener que imponer en diferentes superficies coordinadas.

Antes de escribir las EDO separadas que obtenemos, es mejor reescribir la EDO para $Q(\theta)$, que queda

$$\frac{d^2 Q}{d\theta^2} + \frac{\cos \theta}{\sin \theta} \frac{dQ}{d\theta} + \left[\nu(\nu + 1) - \frac{\mu^2}{\sin^2 \theta} \right] Q = 0,$$

en una forma más habitual y fácilmente reconocible, haciendo el cambio de variable independiente

$$x = \cos \theta, \quad \sin^2 \theta = 1 - x^2$$

y notando que

$$\frac{dQ}{d\theta} = -\sin \theta \frac{dQ}{dx}, \quad \frac{d^2 Q}{d\theta^2} = -\cos \theta \frac{dQ}{dx} + \sin^2 \theta \frac{d^2 Q}{dx^2}.$$

Obtenemos así las EDO separadas

$$F'' + \mu^2 F = 0, \quad (9.124a)$$

$$(1 - x^2)Q'' - 2xQ' + \left[\nu(\nu + 1) - \frac{\mu^2}{1 - x^2} \right] Q = 0, \quad (9.124b)$$

$$r^2 R'' + 2rR' - \nu(\nu + 1)R = 0. \quad (9.124c)$$

Vemos que (9.124a) es una ecuación armónica, (9.124b) es la ecuación asociada de Legendre (6.46), y (9.124c) es una ecuación de Euler, de modo que las correspondientes soluciones serán

$$F(\phi) = \{e^{\pm i\mu\phi}\} = \begin{cases} \cos(\mu\phi) \\ \sin(\mu\phi) \end{cases}, \quad (9.125a)$$

$$Q(\theta) = \begin{cases} P_\nu^\mu(\cos\theta) \\ Q_\nu^\mu(\cos\theta) \end{cases}, \quad (9.125b)$$

$$R(r) = \begin{cases} r^\nu \\ r^{-(\nu+1)} \end{cases}. \quad (9.125c)$$

Notemos que en el caso relativamente frecuente en que Q_ν^μ están ausentes y ν y μ son enteros, la solución puede reescribirse

$$R(r)Q(\theta)F(\phi) = \begin{cases} r^n \\ r^{-(n+1)} \end{cases} \{Y_{nm}(\theta, \phi)\}$$

en términos de los armónicos esféricos (6.60).

9.3.4. La ecuación de Helmholtz en coordenadas esféricas

Al igual que en coordenadas cilíndricas, en coordenadas esféricas la presencia del término $-k^2 u$ introduce diferencias importantes en la separación de la ecuación de Helmholtz respecto de la ecuación de Laplace. Consideraremos el problema

$$\nabla^2 u + k^2 u = 0, \quad 0 \leq a < r < b, \quad 0 \leq \alpha < \theta < \beta \leq \pi, \quad 0 < \phi < \gamma \leq 2\pi, \quad (9.126)$$

y buscaremos soluciones de la forma

$$u(r, \theta, z) = R(r)Q(\theta)F(\phi). \quad (9.127)$$

Aplicando el Laplaciano (9.123) a $u = RQF$ y dividiendo por RQF tenemos ahora

$$\frac{1}{Rr^2} \frac{d}{dr} \left(r^2 \frac{dR}{dr} \right) + \frac{1}{Qr^2 \sin\theta} \frac{d}{d\theta} \left(\sin\theta \frac{dQ}{d\theta} \right) + \frac{1}{Fr^2 \sin^2\theta} \frac{d^2F}{d\phi^2} + k^2 = 0.$$

Multiplicando por $r^2 \sin^2\theta$ tenemos una primera separación para F ,

$$\sin^2\theta \frac{1}{R} \frac{d}{dr} \left(r^2 \frac{dR}{dr} \right) + \sin\theta \frac{1}{Q} \frac{d}{d\theta} \left(\sin\theta \frac{dQ}{d\theta} \right) + k^2 r^2 \sin^2\theta = -\frac{F''}{F} = \mu^2,$$

y volviendo a dividir por $\sin^2 \theta$ obtenemos una segunda separación para R y Q ,

$$\frac{1}{R} \frac{d}{dr} \left(r^2 \frac{dR}{dr} \right) + k^2 r^2 = -\frac{1}{Q \sin \theta} \frac{d}{d\theta} \left(\sin \theta \frac{dQ}{d\theta} \right) + \frac{\mu^2}{\sin^2 \theta} = \nu(\nu + 1).$$

Nuevamente, por simplicidad hemos omitido explicitar un posible doble signo de μ como $\pm \mu^2$, pero como antes debemos recordar que, en principio, μ puede ser imaginario, y que consideraciones equivalentes valen para ν . Como siempre, estas elecciones deberán realizarse en base a las CC (homogéneas o inhomogéneas) que esperemos tener que imponer en diferentes superficies coordenadas.

Es fácil ver que las EDO separadas para $F(\phi)$ y $Q(\theta)$ serán las mismas que en la Sección 9.3.3. Sin embargo esta vez es la EDO para $R(r)$, que queda

$$\frac{d}{dr} \left(r^2 \frac{dR}{dr} \right) + [k^2 r^2 - \nu(\nu + 1)] R = 0,$$

la que necesita cierta reelaboración para llevarla a una forma conocida. Definiendo

$$v(r) = \sqrt{r} R(r),$$

es directo ver que

$$\frac{dR}{dr} = \frac{d}{dr} \left(\frac{v}{\sqrt{r}} \right) = \frac{1}{\sqrt{r}} \frac{dv}{dr} - \frac{1}{2} \frac{v}{r^{3/2}}$$

y

$$\frac{d}{dr} \left(r^2 \frac{dR}{dr} \right) = r^{3/2} \frac{d^2 v}{dr^2} + r^{1/2} \frac{dv}{dr} - \frac{1}{4} r^{-1/2} v,$$

lo que convierte a la EDO para R en la ecuación de Bessel

$$\frac{d^2 v}{dr^2} + \frac{1}{r} \frac{dv}{dr} + \left[k^2 - \frac{(\nu + \frac{1}{2})^2}{r^2} \right] v = 0,$$

con soluciones

$$v(r) = \begin{cases} J_{\nu+\frac{1}{2}}(kr) \\ Y_{\nu+\frac{1}{2}}(kr) \end{cases}.$$

Obtenemos así las EDO separadas

$$F'' + \mu^2 F = 0, \quad (9.128a)$$

$$(1 - x^2) Q'' - 2x Q' + \left[\nu(\nu + 1) - \frac{\mu^2}{1 - x^2} \right] Q = 0, \quad (9.128b)$$

$$r^2 (\sqrt{r} R)'' + r(\sqrt{r} R)' + [k^2 r^2 - (\nu + \frac{1}{2})^2] (\sqrt{r} R) = 0. \quad (9.128c)$$

Como antes, (9.128a) es una ecuación armónica y (9.128b) es la ecuación asociada de Legendre, pero ahora (9.128c) es una ecuación de Bessel para $\sqrt{r} R(r)$;

recordando la definición (6.96) de las funciones de Bessel esféricas j_ν e y_ν , las correspondientes soluciones serán

$$Z(z) = \{e^{\pm i\mu\phi}\} = \begin{cases} \cos(\mu\phi) \\ \sin(\mu\phi) \end{cases}, \quad (9.129a)$$

$$Q(\theta) = \begin{cases} P_\nu^\mu(\cos\theta) \\ Q_\nu^\mu(\cos\theta) \end{cases}, \quad (9.129b)$$

$$R(r) = \begin{cases} j_\nu(kr) \\ y_\nu(kr) \end{cases}. \quad (9.129c)$$

Nuevamente, si Q_ν^μ están ausentes y ν y μ son enteros, la solución puede reescribirse

$$R(r)Q(\theta)F(\phi) = \begin{cases} j_n(kr) \\ y_n(kr) \end{cases} \{Y_{nm}(\theta, \phi)\}$$

en términos de los armónicos esféricos (6.60).

9.4. Ejemplos de aplicación

En esta sección presentaremos algunos ejemplos ilustrativos de la aplicación del método de separación de variables a problemas físicos que llevan a ecuaciones diferenciales parciales. Los ejemplos no han sido seleccionados en base a criterios de sencillez, sino que apuntan a mostrar el uso del método en casos “realistas” y resaltar su potencia.

9.4.1. Conducto cilíndrico en un campo externo

Este es un problema clásico de la electrostática. Consideremos un campo eléctrico uniforme $\vec{E} = -E\hat{x}$ que llena todo el espacio; se introduce en este campo un cilindro conductor de radio a puesto a tierra, con su eje a lo largo de \hat{z} , y tan largo que puede considerarse infinito. El objetivo es encontrar el potencial electrostático en el exterior del cilindro, considerando que muy lejos de este en la dirección radial el campo eléctrico permanece inalterado. Este será un típico problema con CC asintóticas, y es conveniente, antes siquiera de plantearlo, calcular el potencial electrostático del campo *antes de introducir el cilindro*, que será, omitiendo constantes por comodidad,

$$\phi = Ex = Er \cos\theta, \quad (9.130)$$

con r y θ coordenadas cilíndricas.

El problema matemático es entonces

$$\nabla^2 u = 0, \quad a < r < \infty, \quad 0 \leq \theta < 2\pi, \quad -\infty < z < \infty, \quad (9.131a)$$

$$u(a, \theta, z) = 0, \quad u(r, \theta, z) \underset{r \rightarrow \infty}{\asymp} Er \cos\theta, \quad (9.131b)$$

$$u(r, 0, z) = u(r, 2\pi, z), \quad u_\theta(r, 0, z) = u_\theta(r, 2\pi, z).$$

Notemos que el problema es estático, y por lo tanto no incluye CI. Notemos también que la invariancia traslacional en z del problema hace que este sea efectivamente bidimensional. Buscaremos entonces una solución $u(r, \theta)$ independiente de z , trabajando en coordenadas polares r, θ .

De acuerdo a (9.81), tendremos

$$u(r, \theta) = \begin{Bmatrix} r^\nu \\ r^{-\nu} \end{Bmatrix} \begin{Bmatrix} \cos(\nu\theta) \\ \sin(\nu\theta) \end{Bmatrix} (+) \begin{Bmatrix} 1 \\ \ln r \end{Bmatrix} \begin{Bmatrix} 1 \\ \theta \end{Bmatrix}.$$

Las CC periódicas en θ imponen $\nu = n$, un entero ≥ 0 , y la ausencia del término lineal en θ . La CC asintótica, por su parte, sugiere fuertemente conservar sólo los términos con $n = 1$, y sólo los que contienen el coseno. Proponemos entonces

$$u(r, \theta) = \left(Ar + \frac{B}{r} \right) \cos \theta.$$

La CC en $r = a$ impone

$$Aa + \frac{B}{a} = 0,$$

es decir

$$B = -a^2 A$$

y por consiguiente

$$u(r, \theta) = A \left(r - \frac{a^2}{r} \right) \cos \theta.$$

Finalmente, imponiendo la CC asintótica para $r \rightarrow \infty$ vemos que

$$A = E,$$

de modo que la solución completa del problema (9.131) es

$$u(r, \theta, z) = E \left(r - \frac{a^2}{r} \right) \cos \theta.$$

9.4.2. Radiación de una antena cilíndrica

Este es otro problema clásico, esta vez de la electrodinámica, aunque en esta versión sólo trataremos con *ondas escalares*. Consideraremos un conductor cilíndrico de radio a , infinitamente largo, con su eje en la dirección de z . El conductor se halla dividido longitudinalmente en dos mitades, $-\pi < \theta < 0$ y $0 < \theta < \pi$, forzadas a mantener potenciales electrostáticos iguales y opuestos $\pm V_0 \cos(\omega t)$. El objetivo es hallar el potencial eléctrico $u(r, \theta, z, t)$ en el exterior del cilindro en el estado de régimen, considerando que lejos de él en la dirección radial se satisface la condición de *onda saliente* $u \propto e^{\pm i(kr - \omega t)}$. El problema no sólo involucra así CC asintóticas, sino además CC dependientes del tiempo.

El problema matemático se escribe

$$u_{tt} - c^2 \nabla^2 u = 0, \quad a < r < \infty, \quad -\pi \leq \theta < \pi, \quad -\infty < z < \infty, \quad (9.132a)$$

$$\begin{aligned} u(a, \theta, z, t) &= \begin{cases} -V_0 \cos(\omega t), & -\pi < \theta < 0, \\ V_0 \cos(\omega t), & 0 < \theta < \pi, \end{cases} \\ u(r, \theta, z, t) &\underset{r \rightarrow \infty}{\asymp} \{f(r, \theta, z)e^{\pm i(kr - \omega t)}\}, \\ u(r, -\pi, z, t) &= u(r, \pi, z, t), \quad u_\theta(r, -\pi, z, t) = u_\theta(r, \pi, z, t). \end{aligned} \quad (9.132b)$$

Notemos que el problema no especifica CI, ya que estamos interesados sólo en el estado de régimen. Notemos además que la CC asintótica de onda saliente es, por ahora, bastante inespecífica, y simplemente requiere el comportamiento asintótico $u \propto e^{\pm i(kr - \omega t)}$ multiplicado por una función $f(r, \theta, z)$ que presumiblemente tenderá a cero para $r \rightarrow \infty$, aunque aún no sepamos en detalle cómo. Finalmente, notemos también que, al igual que en la Sección 9.4.1, la invariancia traslacional en z del problema hará que este sea efectivamente bidimensional, es decir dimensión $2 + 1$; nos limitaremos entonces a buscar una solución $u(r, \theta, t)$ independiente de z .

Comencemos por partir el problema (9.132) en dos idénticos con la única diferencia de que la CC en $r = a$ de cada uno contendrá $e^{i\omega t}$ en un caso y $e^{-i\omega t}$ en el otro; ello simplificará los cálculos restantes, y podremos reobtener el problema original simplemente sumándolos (y dividiendo por 2); trataremos primero la CC $e^{i\omega t}$. A continuación, como en la Sección 9.1.3, propongamos una primera separación en la forma

$$u(r, \theta, t) = U(r, \theta)e^{i\omega t}.$$

Substituyendo en la EDP (9.132a) obtenemos

$$\nabla^2 U + k^2 U = 0, \quad k^2 = \frac{\omega^2}{c^2},$$

es decir, una ecuación de Helmholtz donde, como habíamos anticipado que podría ocurrir, el parámetro k está fijado de antemano.

La separación de $U(r, \theta)$ en una función $R(r)$ y una $Q(\theta)$ procede exactamente como en la Sección 9.2.4, y en principio tendremos

$$U(r, \theta) = \begin{Bmatrix} J_\nu(kr) \\ Y_\nu(kr) \end{Bmatrix} \begin{Bmatrix} \cos(\nu\theta) \\ \sin(\nu\theta) \end{Bmatrix} = \begin{Bmatrix} H_\nu^{(1)}(kr) \\ H_\nu^{(2)}(kr) \end{Bmatrix} \begin{Bmatrix} \cos(\nu\theta) \\ \sin(\nu\theta) \end{Bmatrix}.$$

La CC periódica en θ impone $\nu = n$, un entero ≥ 0 ; es más, como la CC en $r = a$ deberá ser ajustada seleccionando los coeficientes de $\cos(n\theta)$ y $\sin(n\theta)$, y esta CC es una función *impar* de θ , vemos que podremos descartar los cosenos y que tendremos $n > 0$.

Por otro lado, la CC asintótica es tratada más sencillamente en términos de las funciones de Hankel: de (6.86) vemos que corresponderá elegir $H_n^{(2)}(kr)$,

que para $r \rightarrow \infty$ se comportan como e^{-ikr}/\sqrt{r} , en lugar de $H_n^{(1)}(kr)$, que se comportan como e^{ikr}/\sqrt{r} . Tenemos entonces, hasta ahora,

$$u(r, \theta, t) = \sum_{n=1}^{\infty} A_n H_n^{(2)}(kr) \sin(n\theta) e^{i\omega t}.$$

Aún falta imponer la CC en $r = a$, que se expresa

$$\sum_{n=1}^{\infty} A_n H_n^{(2)}(ka) \sin(n\theta) e^{i\omega t} = V_0 e^{i\omega t} \begin{cases} -1, & -\pi < \theta < 0, \\ 1, & 0 < \theta < \pi. \end{cases}$$

El factor $e^{i\omega t}$ se cancelará (lo que era precisamente el propósito de la primera separación), y si recordamos que ya conocemos el desarrollo de la onda cuadrada en serie de senos,

$$\frac{4}{\pi} \sum_{n=1,3,5,\dots}^{\infty} \frac{\sin(n\theta)}{n} = \begin{cases} -1, & -\pi < \theta < 0, \\ 1, & 0 < \theta < \pi, \end{cases}$$

vemos inmediatamente que debemos tener

$$A_n = \begin{cases} 0, & n \text{ par}, \\ \frac{4V_0}{n\pi H_n^{(2)}(ka)}, & n \text{ impar}. \end{cases}$$

La solución del problema con CC en $r = a$ proporcional a $e^{i\omega t}$, entonces, queda

$$u(r, \theta, t) = \frac{4V_0}{\pi} \sum_{n=1,3,5,\dots}^{\infty} \frac{H_n^{(2)}(kr)}{H_n^{(2)}(ka)} \frac{\sin(n\theta)}{n} e^{i\omega t}.$$

Para el segundo subproblema, con CC en $r = a$ proporcional a $e^{-i\omega t}$, es inmediato que sólo necesitamos cambiar $H_n^{(2)}(kr)$ por $H_n^{(1)}(kr)$. La solución completa del problema (9.132) se escribe, entonces,

$$u(r, \theta, t) = \frac{2V_0}{\pi} \sum_{n=1,3,5,\dots}^{\infty} \left[\frac{H_n^{(2)}(kr)}{H_n^{(2)}(ka)} e^{i\omega t} + \frac{H_n^{(1)}(kr)}{H_n^{(1)}(ka)} e^{-i\omega t} \right] \frac{\sin(n\theta)}{n}.$$

Notemos que, dado que $H_n^{(2)}(kr)$ y $H_n^{(1)}(kr)$ son complejas conjugadas una de la otra, la solución es efectivamente real. Esta solución puede reescribirse explícitamente en términos de funciones reales (J_n e Y_n , y senos y cosenos de ωt), lo que se deja como ejercicio para el lector.

9.4.3. Difusión del calor en una placa circular

Consideremos una placa plana circular y homogénea de radio a y difusividad térmica κ , inicialmente a temperatura nula. En el instante inicial la mitad

superior de la placa es puesta en contacto térmico perfecto con un reservorio a temperatura T_0 , mientras que su mitad inferior se pone en contacto térmico perfecto con un reservorio a temperatura $-T_0$. El objetivo es hallar la distribución de temperatura $u(r, \theta, t)$ en el interior de la placa.

El problema matemático se escribe

$$u_t - \kappa \nabla^2 u = 0, \quad 0 \leq r < a, \quad -\pi < \theta < \pi, \quad 0 < t < \infty, \quad (9.133a)$$

$$\lim_{r \downarrow 0} |u(r, \theta, t)| < \infty, \quad u(a, \theta, t) = \begin{cases} -T_0, & -\pi < \theta < 0, \\ T_0, & 0 < \theta < \pi, \end{cases} \quad (9.133b)$$

$$\begin{aligned} u(r, -\pi, t) &= u(r, \pi, t), & u_\theta(r, -\pi, t) &= u_\theta(r, \pi, t), \\ u(r, \theta, 0) &= 0. \end{aligned} \quad (9.133c)$$

Es conveniente separar

$$u(r, \theta, t) = v(r, \theta, t) + w(r, \theta),$$

donde w es la solución de estado estacionario con las mismas CC.

El problema para w se escribe

$$\nabla^2 w = 0, \quad 0 \leq r < a, \quad 0 \leq \theta < 2\pi, \quad (9.134a)$$

$$\lim_{r \downarrow 0} |w(r, \theta)| < \infty, \quad w(a, \theta) = \begin{cases} -T_0, & -\pi < \theta < 0, \\ T_0, & 0 < \theta < \pi, \end{cases} \quad (9.134b)$$

$$w(r, -\pi) = w(r, \pi), \quad w_\theta(r, -\pi) = w_\theta(r, \pi).$$

De acuerdo a (9.81) tendremos

$$w(r, \theta) = \left\{ \begin{array}{l} r^\nu \\ r^{-\nu} \end{array} \right\} \left\{ \begin{array}{l} \cos(\nu\theta) \\ \sin(\nu\theta) \end{array} \right\} (+) \left\{ \begin{array}{l} 1 \\ \ln r \end{array} \right\} \left\{ \begin{array}{l} 1 \\ \theta \end{array} \right\}.$$

Las CC periódicas en $-\pi < \theta < \pi$ imponen $\nu = n$, un entero ≥ 0 , y la ausencia del término lineal en θ , y las CC de función finita en $r = 0$ eliminan las potencias negativas r^{-n} y el término $\ln r$, de modo que tendremos

$$w(r, \theta) = A_0 + \sum_{n=1}^{\infty} r^n (A_n \cos(n\theta) + B_n \sin(n\theta)).$$

La CC restante, en $r = a$, se escribe

$$A_0 + \sum_{n=1}^{\infty} a^n (A_n \cos(n\theta) + B_n \sin(n\theta)) = \begin{cases} -T_0, & -\pi < \theta < 0, \\ T_0, & 0 < \theta < \pi, \end{cases}$$

y corresponde nuevamente al desarrollo en serie de Fourier (6.9) de la onda cuadrada; tenemos entonces

$$w(a, \theta) = \frac{4T_0}{\pi} \sum_{n=1,3,5,\dots}^{\infty} \frac{\sin(n\theta)}{n}.$$

Vemos que

$$A_n = 0, \quad B_n = \begin{cases} 0, & n \text{ par}, \\ \frac{4T_0}{n\pi a^n}, & n \text{ impar}, \end{cases}$$

de modo que

$$w(r, \theta) = \frac{4T_0}{\pi} \sum_{n=1,3,5,\dots}^{\infty} \left(\frac{r}{a}\right)^n \frac{\sin(n\theta)}{n}.$$

Dado que w es solución de la EDP (9.133a), el problema para v queda

$$v_t - \kappa \nabla^2 v = 0, \quad 0 \leq r < a, \quad 0 \leq \theta < 2\pi, \quad 0 < t < \infty, \quad (9.135a)$$

$$\lim_{r \downarrow 0} |v(r, \theta, t)| < \infty, \quad v(r, \theta, t) = 0, \quad (9.135b)$$

$$v(r, -\pi, t) = v(r, \pi, t), \quad v_\theta(r, -\pi, t) = v_\theta(r, \pi, t),$$

$$v(r, \theta, 0) = -w(r, \theta). \quad (9.135c)$$

Usando (9.47a), (9.99) y (9.100), tenemos

$$v(r, \theta, t) = \begin{Bmatrix} J_\nu(kr) \\ Y_\nu(kr) \end{Bmatrix} \begin{Bmatrix} \cos(\nu\theta) \\ \sin(\nu\theta) \end{Bmatrix} e^{-\kappa k^2 t}.$$

Las CC de periodicidad en $-\pi < \theta < \pi$ imponen $n = \nu$, un entero ≥ 0 , y las CC de función finita en $r = 0$ excluyen Y_ν ; luego

$$v(r, \theta, t) = \sum_{n=0}^{\infty} J_n(kr) (A_n \cos(n\theta) + B_n \sin(n\theta)) e^{-\kappa k^2 t}.$$

La CC homogénea de Dirichlet en $r = a$ se escribe ahora

$$v(a, \theta, t) = \sum_{n=0}^{\infty} J_n(ka) (A_n \cos(n\theta) + B_n \sin(n\theta)) e^{-\kappa k^2 t} = 0,$$

de donde

$$J_n(ka) = 0.$$

Luego k debe adoptar los valores

$$k_{nm} = \frac{\kappa_{nm}}{a}, \quad m = 1, 2, 3, \dots$$

con κ_{nm} el m -ésimo cero de J_n : $J_n(\kappa_{nm}) = 0$. Tenemos así

$$v(r, \theta, t) = \sum_{n=0}^{\infty} \sum_{m=1}^{\infty} J_n(k_{nm}r) (A_{nm} \cos(n\theta) + B_{nm} \sin(n\theta)) e^{-\kappa k_{nm}^2 t}.$$

Imponiendo ahora la CI (9.135c) resulta

$$\sum_{n=0}^{\infty} \sum_{m=1}^{\infty} J_n(k_{nm}r) (A_{nm} \cos(n\theta) + B_{nm} \sin(n\theta)) = -\frac{4T_0}{\pi} \sum_{n=1,3,5,\dots}^{\infty} \left(\frac{r}{a}\right)^n \frac{\sin(n\theta)}{n}.$$

Usando la ortogonalidad de los senos y cosenos en $\mathcal{L}^2(-\pi, \pi)$, es inmediato que $A_{nm} \equiv 0$ y $B_{nm} = 0$ para n par, y que tendremos

$$\sum_{m=1}^{\infty} B_{nm} J_n(k_{nm}r) = -\frac{4T_0}{n\pi} \left(\frac{r}{a}\right)^n, \quad n = 1, 3, 5, \dots,$$

con lo cual, apelando a la ortogonalidad de $J_n(k_{nm}r)$ en $\mathcal{L}_r^2(0, a)$,

$$B_{nm} \int_0^a J_n^2(k_{nm}r) r dr = -\frac{4T_0}{n\pi a^n} \int_0^a r^{n+1} J_n(k_{nm}r) dr, \quad m = 1, 2, 3, \dots$$

Usando la integral de normalización (6.71)

$$\int_0^a J_n^2(k_{nm}r) r dr = \frac{a^2}{2} [J'_n(\kappa_{nm})]^2,$$

y usando (6.77b)

$$\int_0^a r^{n+1} J_n(k_{nm}r) dr = \frac{1}{k_{nm}^{n+2}} \int_0^{k_{nm}a} x^{n+1} J_n(x) dx = \left[x^{n+1} J_{n+1}(x) \right]_0^{k_{nm}a},$$

que recordando que $J_n(0) = 0$ para $n > 0$ queda

$$\int_0^a r^{n+1} J_n(k_{nm}r) dr = k_{nm}^{n+1} a^{n+1} J_{n+1}(k_{nm}a) = \kappa_{nm}^{n+1} J_{n+1}(\kappa_{nm}).$$

Resulta entonces

$$B_{nm} = -\frac{8T_0 \kappa_{nm}^{n+1}}{\pi a^{n+2}} \frac{J_{n+1}(\kappa_{nm})}{[J'_n(\kappa_{nm})]^2}, \quad n = 1, 3, 5, \dots, \quad m = 1, 2, 3, \dots,$$

y en definitiva

$$v(r, \theta, t) = -\frac{8T_0}{\pi a} \sum_{n=1,3,5,\dots}^{\infty} \sum_{m=1}^{\infty} \frac{k_{nm}^{n+1}}{n} \frac{J_{n+1}(\kappa_{nm})}{[J'_n(\kappa_{nm})]^2} J_n(k_{nm}r) \sin(n\theta) e^{-\kappa k_{nm}^2 t}.$$

Finalmente, la solución completa del problema (9.133) se escribe

$$u(r, \theta, t) = \frac{4T_0}{\pi} \sum_{n=1,3,5,\dots}^{\infty} \frac{\sin(n\theta)}{n} \left[\left(\frac{r}{a}\right)^n - \frac{2}{a} \sum_{m=1}^{\infty} k_{nm}^{n+1} \frac{J_{n+1}(\kappa_{nm})}{[J'_n(\kappa_{nm})]^2} J_n(k_{nm}r) e^{-\kappa k_{nm}^2 t} \right].$$

Nótese la corrección física del resultado obtenido: para $t \rightarrow \infty$ todas las exponentiales $e^{-\kappa k_{nm}^2 t}$ tienden a cero, y se recupera la solución estacionaria $w(r, \theta)$; mientras que, por construcción, a $t = 0$ el segundo término en el corchete se reduce al desarrollo de $(r/a)^n$ en serie de $J_n(k_{nm}r)$, de modo que se recupera la condición inicial nula.

9.4.4. Vibraciones de una membrana circular

Este problema es un clásico de la literatura, y vale la pena analizarlo, ya que se lo encuentra a menudo en la vida cotidiana en la forma de *parche de tambor*, o incluso de cama elástica. El sistema (idealizado) consiste de una membrana circular tensa, fija en su borde a un aro rígido y plano de radio a , por lo que las CC serán allí homogéneas de Dirichlet. Despreciaremos la influencia de la gravedad, y querremos estudiar los posibles modos de oscilación transversales de la membrana.

El problema matemático se escribe

$$u_{tt} - c^2 \nabla^2 u = 0, \quad 0 \leq r < a, \quad 0 \leq \theta < 2\pi, \quad (9.136a)$$

$$\begin{aligned} \lim_{r \downarrow 0} |u(r, \theta, t)| &< \infty, \quad u(a, \theta, t) = 0, \\ u(r, 0, t) &= u(r, 2\pi, t), \quad u_\theta(r, 0, t) = u_\theta(r, 2\pi, t), \end{aligned} \quad (9.136b)$$

donde por el momento no especificaremos condiciones iniciales. De acuerdo a (9.97) y (9.98), tendremos

$$u(r, \theta, t) = \left\{ \begin{array}{l} J_\nu(kr) \\ Y_\nu(kr) \end{array} \right\} \left\{ \begin{array}{l} \cos(\nu\theta) \\ \sin(\nu\theta) \end{array} \right\} \left\{ \begin{array}{l} \cos(\omega t) \\ \sin(\omega t) \end{array} \right\},$$

con

$$\omega = ck.$$

La CC periódica en θ impone $\nu = n$, un entero, y la CC de función finita en $r = 0$ excluye los Y_n . La CC de Dirichlet en $r = a$ impone entonces

$$J_n(ka) = 0,$$

con lo que k debe ser uno de

$$k_{nm} = \frac{\kappa_{nm}}{a}, \quad m = 1, 2, 3, \dots,$$

con κ_{nm} los sucesivos ceros de J_n , es decir $J_n(\kappa_{nm}) = 0$, y el correspondiente ω es $\omega_{nm} = ck_{nm}$. La solución es entonces

$$u(r, \theta, t) = \sum_{n=0}^{\infty} \sum_{m=1}^{\infty} c_{nm} J_n(k_{nm} r) \left\{ \begin{array}{l} \cos(n\theta) \\ \sin(n\theta) \end{array} \right\} \left\{ \begin{array}{l} \cos(\omega_{nm} t) \\ \sin(\omega_{nm} t) \end{array} \right\}.$$

Los *modos espaciales* de oscilación de la membrana,

$$U_{nm}(r, \theta) = J_n(k_{nm} r) \left\{ \begin{array}{l} \cos(n\theta) \\ \sin(n\theta) \end{array} \right\},$$

se muestran esquemáticamente en la Figura 9.3 para los primeros valores de n y m . Notemos que salvo para $n = 0$, para cada $n = 1, 2, \dots$ tenemos *dos* modos espaciales, correspondientes al coseno y seno de $n\theta$. Los modos con $n = 0$ son los únicos que tienen simetría de rotación; los demás modos carecen de ella,

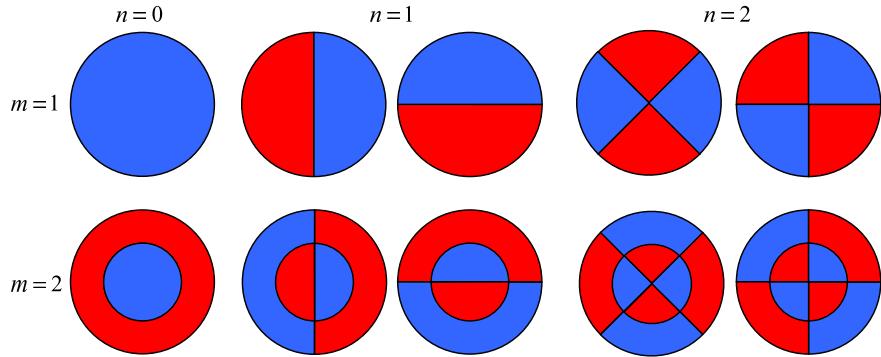


Figura 9.3: Los primeros modos espaciales de una membrana circular. En azul los desplazamientos positivos, en rojo los negativos. Las líneas donde los colores cambian son nodos del modo correspondiente. Para $n = 1$ y $n = 2$ se muestran los modos correspondientes a $\cos(n\theta)$ a la izquierda y a $\sin(n\theta)$ a la derecha.

aunque el problema (9.136) en sí es invariante bajo rotaciones. Este es un ejemplo de *rompimiento de simetría*, que ocurre cuando las soluciones de un problema tienen menos simetrías que el problema mismo.

En la Figura 9.3 las líneas donde los colores cambian son nodos del modo correspondiente, lugares donde $U_{nm} = 0$ y por consiguiente la membrana está *quieta* a todo tiempo si ese modo es el único que se excita. Notemos que si excitamos por ejemplo ambos modos con $m = n = 1$, que tienen la misma frecuencia de oscilación ω_{11} pero están desfasados angularmente en $\pi/2$, y lo hacemos con la misma amplitud pero con un desfasaje temporal también de $\pi/2$, tendremos

$$\begin{aligned} u(r, \theta, t) &= J_1(k_{11}r) [\cos(\theta) \cos(\omega_{11}t) \pm \sin(\theta) \sin(\omega_{11}t)] \\ &= J_1(k_{11}r) \cos(\theta \mp \omega_{11}t), \end{aligned}$$

que corresponde al mismo modo espacial *rotando* a derecha o izquierda con velocidad angular ω_{11} . Este fenómeno es equivalente a la descomposición de una onda estacionaria en una cuerda en ondas viajeras a derecha e izquierda, o viceversa. También es el mismo fenómeno detrás de la descomposición de una onda luminosa circularmente polarizada en ondas con polarización lineal y viceversa.

En <https://www.falstad.com/circosc/index.html> se provee una applet Java muy sencilla de usar que permite visualizar los modos de oscilación de una membrana circular.

9.4.5. Guía de ondas acústica

Para pequeños apartamientos de la presión atmosférica P_0 , supuesta constante, la presión del aire puede escribirse $P(\vec{x}, t) = P_0 + u(\vec{x}, t)$, y puede de-

mostrarse que la *sobrepresión* $u(\vec{x}, t)$ satisface la ecuación de ondas escalares. Puesto que la aceleración de un dado volumen (pequeño) de aire es proporcional al gradiente ∇u de la sobrepresión, y la aceleración normal a una superficie rígida debe ser nula, las CC en las superficies rígidas deben ser homogéneas de Neumann. Supongamos que no nos interesa estudiar este problema en toda su generalidad, sino sólo estudiar la propagación de ondas acústicas a lo largo de un cilindro (supuesto infinito) hueco y rígido de radio interior a . Si el eje del cilindro coincide con \hat{z} , entonces, buscaremos soluciones de la ecuación de ondas para u que sean proporcionales a, digamos, $e^{\pm i(k_z z - \omega t)}$, lo que nos garantiza tener ondas no atenuadas que propagan en la dirección positiva del eje z .

El problema matemático tendrá la forma

$$u_{tt} - c^2 \nabla^2 u = 0, \quad 0 \leq r < a, \quad 0 \leq \theta < 2\pi, \quad -\infty < z < \infty, \quad (9.137a)$$

$$\begin{aligned} \lim_{r \downarrow 0} |u(r, \theta, z, t)| &< \infty, & u_r(a, \theta, z, t) &= 0, \\ u(r, 0, z, t) &= u(r, 2\pi, z, t), & u_\theta(r, 0, z, t) &= u_\theta(r, 2\pi, z, t). \end{aligned} \quad (9.137b)$$

Notemos que no estamos especificando CC en z ni CI; en lugar de ello proponemos una separación en la forma

$$u(r, \theta, z, t) = R(r)Q(\theta)e^{i(k_z z - \omega t)},$$

con lo cual, de acuerdo a (9.120), tendremos

$$u(r, \theta, z, t) = \begin{cases} J_\mu\left(\sqrt{\omega^2/c^2 - k_z^2} r\right) & \left\{ \begin{array}{l} \cos(\mu\theta) \\ \sin(\mu\theta) \end{array} \right\} e^{i(k_z z - \omega t)}. \\ Y_\mu\left(\sqrt{\omega^2/c^2 - k_z^2} r\right) & \end{cases}$$

La CC de función finita en $r = 0$ excluye los Y_μ , y la CC periódica en θ impone $\mu = m$, un entero. Por su parte, la CC de Neumann en $r = a$ requiere

$$J'_m\left(\sqrt{\omega^2/c^2 - k_z^2} a\right) = 0,$$

lo que determina los posibles valores de k_z como

$$k_{zmn} = \sqrt{\frac{\omega^2}{c^2} - \frac{\alpha_{mn}^2}{a^2}},$$

donde α_{mn} son los ceros de J'_m ,

$$J'_m(\alpha_{mn}) = 0, \quad n = 1, 2, 3, \dots$$

La interpretación física de este resultado es la siguiente: Tendremos una superposición de ondas viajeras que propagan en la dirección \hat{z} a la velocidad del sonido c , cada una con número de onda k_{zmn} y un “perfil” o *modo transversal* dado por

$$U_{mn}(r, \theta) = J_m\left(\sqrt{\frac{\omega^2}{c^2} - k_{zmn}^2} r\right) \left\{ \begin{array}{l} \cos(m\theta) \\ \sin(m\theta) \end{array} \right\} = J_m\left(\frac{\alpha_{mn}}{a} r\right) \left\{ \begin{array}{l} \cos(m\theta) \\ \sin(m\theta) \end{array} \right\},$$

de por sí muy semejantes a los modos espaciales de oscilación de una membrana circular vistos en la Sección 9.4.4. Sin embargo, dado que α_{mn} son crecientes con n , sólo propagarán en la dirección z un *número finito* de modos, aquellos para los cuales

$$\alpha_{mn} < a \frac{\omega}{c}$$

y por consiguiente k_{zmn} resulte *real*. Los modos con n mayor que este valor crítico n_c (que dependerá de m) tendrán k_{zmn} *imaginario*, y en lugar de propagar como ondas, su amplitud, si es finita por ejemplo en $z = 0$, *decaerá exponencialmente* con z .¹¹ Si el cilindro es estrecho (a pequeño) sólo frecuencias lo bastante altas (ω grande) propagarán como ondas; este es el origen de la perceptible distorsión tonal de los sonidos propagados a lo largo de un tubo. Debemos notar que las CC de Neumann introducen una peculiaridad en esta fenomenología: como $J'_0(0) = 0$, el *modo uniforme* (o fundamental)

$$u_{00}(r, \theta, z, t) = e^{i(k_z z - \omega t)} = \exp\left(i\frac{\omega}{c}(z - ct)\right),$$

correspondiente a una onda plana (es decir, sin dependencia de r ni θ) *siempre propagará* como onda.

Una fenomenología similar se observa en la transmisión de ondas electromagnéticas a lo largo de tubos conductores o dieléctricos, aunque en ese caso el carácter vectorial de las ondas complica los detalles.

9.4.6. Cavidad resonante acústica

Si tenemos una esfera rígida hueca de radio interior a llena, por ejemplo, de aire, podemos tener ondas acústicas estacionarias si la disipación es despreciable. Estas serán ondas escalares de sobrepresión, como las consideradas en la Sección 9.4.5, y las CC serán homogéneas de Neumann. El problema que vamos a tratar es entonces un análogo en dimensión 3+1 de los tratados en las Secciones 9.1.4 (oscilaciones de una cuerda, dimensión 1+1) y 9.4.4 (vibraciones de una membrana circular, dimensión 2+1). El objetivo no será una solución completa del problema, sino sólo obtener los posibles modos espaciales de oscilaciones en la cavidad.

El problema matemático se escribe

$$u_{tt} - c^2 \nabla^2 u = 0, \quad 0 \leq r < a, \quad 0 < \theta < \pi, \quad 0 \leq \phi < 2\pi, \quad (9.138a)$$

$$\begin{aligned} \lim_{r \downarrow 0} |u(r, \theta, \phi, t)| &< \infty, & u_r(a, \theta, \phi, t) &= 0, \\ \lim_{\theta \downarrow 0} |u(r, \theta, \phi, t)| &< \infty, & \lim_{\theta \uparrow \pi} |u(r, \theta, \phi, t)| &< \infty, \\ u(r, \theta, 0, t) &= u(r, \theta, 2\pi, t), & u_\phi(r, \theta, 0, t) &= u_\phi(r, \theta, 2\pi, t), \end{aligned} \quad (9.138b)$$

¹¹En otro contexto estas se conocen como *ondas evanescentes*.

donde por el momento no especificaremos condiciones iniciales. De acuerdo a (9.129), tendremos

$$u(r, \theta, \phi, t) = \begin{Bmatrix} j_\nu(kr) \\ y_\nu(kr) \end{Bmatrix} \begin{Bmatrix} P_\nu^\mu(\cos \theta) \\ Q_\nu^\mu(\cos \theta) \end{Bmatrix} \begin{Bmatrix} \cos(\mu\phi) \\ \sin(\mu\phi) \end{Bmatrix} \begin{Bmatrix} \cos(\omega t) \\ \sin(\omega t) \end{Bmatrix},$$

con

$$\omega = ck.$$

La CC de periodicidad en ϕ impone $\mu = m$, un entero ≥ 0 ; las CC de función finita en $\theta = 0$ y $\theta = \pi$ excluyen las Q_ν^μ , e imponen $\nu = n$, otro entero ≥ 0 ; las CC de función finita en $r = 0$ excluyen las y_ν ; y la CC de Neumann en $r = a$ impone

$$j'_n(ka) = 0,$$

con lo que tenemos

$$k = k_{n\ell} = \frac{\kappa_{n\ell}}{a}, \quad \omega = \omega_{n\ell} = ck_{n\ell}, \quad \ell = 1, 2, 3, \dots,$$

con $\kappa_{n\ell}$ el ℓ -ésimo cero de j'_n . La solución de (9.138) se escribe entonces

$$u(r, \theta, \phi, t) = \sum_{n=0}^{\infty} \sum_{m=-n}^n \sum_{\ell=1}^{\infty} c_{nml} j_n(k_{n\ell} r) Y_n^m(\theta, \phi) \begin{Bmatrix} \cos(\omega_{n\ell} t) \\ \sin(\omega_{n\ell} t) \end{Bmatrix},$$

donde las constantes c_{nml} deberían determinarse a partir de las CI, si las impusiéramos, y los modos espaciales de las ondas en la cavidad serán

$$U_{nml}(r, \theta, \phi) = j_n(k_{n\ell} r) Y_n^m(\theta, \phi).$$

Si calculamos la frecuencia de oscilación más baja para una cavidad de radio $a = 1$ m llena de aire a presión normal y 20°C , y por lo tanto con $c = 343 \text{ ms}^{-1}$, tendremos

$$\nu_{01} = \frac{\omega_{01}}{2\pi} = \frac{\kappa_{01}c}{2\pi a} \simeq 245 \text{ Hz},$$

donde hemos usado que $\kappa_{01} \simeq 4,4934$. Notemos de paso que de acuerdo a (6.99), los ceros de j'_0 son las soluciones de $\tan x = x$.

Vale la pena notar que, salvo el modo fundamental U_{001} , todos los demás modos son *degenerados*: todos los modos U_{nml} con n y ℓ dados tendrán la misma frecuencia de oscilación $\omega_{n\ell}$, y la degeneración será $2n + 1$.

Capítulo 10

Método de Transformadas Integrales

El método de transformadas integrales resulta extremadamente útil en la resolución de algunas EDPs definidas sobre dominios *infinitos* o *semiinfinitos*, aunque muchas veces estas puedan tratarse igualmente por el método de separación de variables, como se hizo en las Secciones 9.4.1, 9.4.2 y, especialmente, 9.4.5. Dada una EDP cuyo dominio sea infinito o semiinfinito en alguna coordenada (espacial o temporal), el uso de una transformada integral apropiada en esa coordenada reducirá la cantidad de variables independientes de la EDP transformada. Sin embargo, por eso mismo pocas veces se lo aplica de manera aislada, siendo en general usado en conjunción con el método de separación de variables. En consecuencia la teoría es más dispersa y menos organizada que en el caso de este último. La exposición será entonces más fructífera y llevadera a través de ejemplos seleccionados.

La utilidad del método dependerá en gran medida de la dificultad de calcular las antitransformadas necesarias. Por ello debemos destacar que, cuando ambos métodos sean aplicables, el lector deberá usar su buen juicio a la hora de decidir entre ellos.

10.1. Dominios semiinfinitos

Comenzaremos estudiando problemas en los que el dominio de alguna coordenada sea semiinfinito. Ello incluye naturalmente cualquier EDP dependiente del tiempo y con CI a un tiempo dado, lo que no implica que automáticamente debamos elegir la variable temporal como candidata a una transformación; sin embargo, veremos que este será generalmente el caso.

10.1.1. Difusión del calor en una barra semiinfinita

Consideremos una barra delgada y homogénea, que se extiende desde $x = 0$ en adelante y está aislada lateralmente, y supongamos que inicialmente se halla a temperatura nula. En el instante $t = 0$ su extremo se pone en contacto térmico perfecto con un reservorio a temperatura $T_0 > 0$. El objetivo es hallar la temperatura $u(x, t)$ de la barra para todo $t > 0$ y $x > 0$, bajo la razonable suposición de que esta será finita para todo x y t , y en particular para $x \rightarrow \infty$.

El problema matemático se escribe

$$u_t - \kappa u_{xx} = 0, \quad 0 < x < \infty, \quad 0 < t < \infty, \quad (10.1a)$$

$$u(0, t) = T_0, \quad \lim_{x \rightarrow \infty} |u(x, t)| < \infty, \quad (10.1b)$$

$$u(x, 0) = 0. \quad (10.1c)$$

En este caso ambos dominios (en x y en t) son semiinfinitos. Podríamos estar tentados a tomar transformada de Laplace en x ; sin embargo por (7.21) para ello deberíamos especificar CC en $x = 0$ para u y para u_x , y sólo tenemos especificada la CC para u ; la CC de función finita para $x \rightarrow \infty$ no resuelve este problema. Optamos entonces por tomar transformada de Laplace en t .

Poniendo

$$\tilde{u}(x, s) = \mathcal{L}_{t \rightarrow s}[u(x, t)]$$

y usando (7.21) tendremos el problema

$$s\tilde{u} - \kappa \tilde{u}_{xx} = 0, \quad 0 < x < \infty, \quad (10.2a)$$

$$\tilde{u}(0, s) = \frac{T_0}{s}, \quad \lim_{x \rightarrow \infty} |\tilde{u}(x, s)| < \infty. \quad (10.2b)$$

Notemos tres cosas: Primera, que la EDP ahora es una EDO. Segunda, que hemos tomado también la transformada de Laplace de las CC (10.1b), que resultan en las CC (10.2b). Y tercera, que la CC de función finita para u se ha transformado en otra del mismo tipo para \tilde{u} .

La EDO (10.2a) tiene como solución general

$$\tilde{u}(x, s) = A \exp\left(\sqrt{\frac{s}{\kappa}}x\right) + B \exp\left(-\sqrt{\frac{s}{\kappa}}x\right).$$

La CC (10.2b) de función finita impone $A = 0$; la CC en $x = 0$ impone $B = T_0/s$. Luego

$$\tilde{u}(x, s) = \frac{T_0}{s} \exp\left(-\sqrt{\frac{s}{\kappa}}x\right).$$

De tablas, o usando una calculadora *online*,¹ vemos que

$$\mathcal{L}_{t \rightarrow s}^{-1}\left[\frac{1}{s}e^{-a\sqrt{s}}\right] = \text{erfc}\left(\frac{a}{2\sqrt{t}}\right)$$

¹En este caso, <https://www.wolframalpha.com/input/?i=inverse+Laplace+transform>

de modo que, finalmente, la solución del problema (10.1) resulta

$$u(x, t) = T_0 \operatorname{erfc} \left(\frac{x}{2\sqrt{\kappa t}} \right),$$

donde erfc es la *función error complementaria*²

$$\operatorname{erfc}(z) = \frac{2}{\sqrt{\pi}} \int_z^\infty e^{-y^2} dy.$$

Una variación simple del problema anterior es cambiar la CC en $x = 0$ a una dependiente del tiempo, en principio de manera arbitraria, es decir

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

donde el único requerimiento sobre f es que su transformada de Laplace exista. Es sencillo ver que en ese caso la CC transformada será

$$\tilde{u}(0, s) = \tilde{f}(s),$$

con lo cual ahora tendremos

$$\tilde{u}(x, s) = \tilde{f}(s) \exp \left(-\sqrt{\frac{s}{\kappa}} x \right).$$

Podríamos pensar que no podemos avanzar más a menos que especifiquemos f o dejemos indicada la antitransformada. Sin embargo la propiedad (7.28) nos permite un paso más: notando que

$$\mathcal{L}_{s \rightarrow t}^{-1} \left[e^{-a\sqrt{s}} \right] = \frac{a}{2\sqrt{\pi}} \frac{e^{-a^2/4t}}{t^{3/2}},$$

vemos que podemos escribir

$$u(x, t) = \frac{x}{2\sqrt{\pi\kappa}} \int_0^t f(t') \frac{e^{-x^2/4\kappa(t-t')}}{(t-t')^{3/2}} dt',$$

un resultado en *forma cerrada* en términos de la función original $f(t)$. Esta clase de resultados es típica del método de transformadas integrales.

Aún otra variación del problema es introducir, además, una CI inhomogénea, digamos

$$u(x, 0) = g(x),$$

²La función error complementaria satisface $\operatorname{erfc}(z) = 1 - \operatorname{erf}(z)$, con

$$\operatorname{erf}(z) = \frac{2}{\sqrt{\pi}} \int_0^z e^{-y^2} dy$$

la *función error*. Una aproximación sencilla a esta última es

$$\operatorname{erf}(z) \simeq \tanh \left(\frac{2}{\sqrt{\pi}} \left(z + \frac{11}{123} z^3 \right) \right),$$

con un error menor que 0,000358 para todo x .

Usando (7.21) tendremos ahora la EDO transformada *inhomogénea*

$$s\tilde{u} - \kappa\tilde{u}_{xx} = g(x).$$

Si la EDP original es también inhomogénea, por ejemplo

$$u_t - \kappa u_{xx} = h(x, t),$$

la EDO transformada quedará

$$s\tilde{u} - \kappa\tilde{u}_{xx} = g(x) + \tilde{h}(x, s).$$

Es decir, tanto la inhomogeneidad de las CI como de la EDP original resultan en una inhomogeneidad de la EDO transformada; en este aspecto, tener una, la otra o ambas no hace diferencia en cuanto a la dificultad del problema. La EDO transformada inhomogénea, por su parte, puede resolverse por el consabido método de sumar una solución particular de la ecuación inhomogénea a la solución general de la homogénea, o por los métodos generales que discutiremos en el Capítulo 11.

De este sencillo problema derivamos dos lecciones generales. La primera es que, debido a la propiedad (7.21), sólo tendrá sentido tomar transformada de Laplace de variables para las cuales, en la frontera de su dominio semiinfinito, el problema especifique tantas CC o CI como el grado de la EDP en esa variable. Dado que las EDP que venimos tratando sólo admiten este tipo de CC para la variable temporal, ya que imponerlas a las coordenadas espaciales llevaría a problemas sobredeterminados según la Tabla 8.3, vemos que en general la utilidad de la transformada de Laplace se restringirá a eliminar la variable temporal para obtener una EDP o EDO sólo en coordenadas espaciales.

La segunda es que aún problemas sencillos como los tratados hasta aquí llevan al cálculo de antittransformadas que, de realizarse por integración en el plano complejo, pueden ser de considerable extensión y complejidad. La utilidad práctica del método depende entonces, en buena medida, de contar con tablas extensivas de pares transformada-antittransformada, o la ayuda de calculadoras *online* como la utilizada más arriba; preferentemente ambas.

10.1.2. Ondas en una cuerda semiinfinita

Consideraremos una cuerda delgada e inextensible, que se extiende desde $x = 0$ en adelante, y supongamos que inicialmente tiene un perfil de deformación y de velocidad nulos, es decir, está quieta y sin deformar; en el instante inicial $t = 0$ su extremo en $x = 0$ comienza a ser agitado sinusoidalmente en el tiempo con una amplitud constante a y frecuencia ω . El objetivo es hallar la deformación $u(x, t)$ de la cuerda para todo $x > 0$ y $t > 0$, bajo la suposición de que la deformación permanece nula en el infinito (a tiempos finitos).

El problema matemático se escribe

$$u_{tt} - c^2 u_{xx} = 0, \quad 0 < x < \infty, \quad 0 < t < \infty, \quad (10.3a)$$

$$u(0, t) = a \operatorname{sen}(\omega t), \quad \lim_{x \rightarrow \infty} |u(x, t)| = 0, \quad (10.3b)$$

$$u(x, 0) = 0, \quad u_t(x, 0) = 0. \quad (10.3c)$$

Basándonos en lo visto en la sección 10.1.1, vamos a tomar transformada de Laplace en t . Poniendo

$$\tilde{u}(x, s) = \mathcal{L}_{t \rightarrow s}[u(x, t)]$$

y usando (7.21) tendremos el problema

$$s^2 \tilde{u} - c^2 \tilde{u}_{xx} = 0, \quad 0 < x < \infty, \quad (10.4a)$$

$$\tilde{u}(0, s) = \frac{a\omega}{\omega^2 + s^2}, \quad \lim_{x \rightarrow \infty} |\tilde{u}(x, s)| < 0. \quad (10.4b)$$

La EDO (10.4a) tiene como solución general

$$\tilde{u}(x, s) = A \exp\left(\frac{s}{c}x\right) + B \exp\left(-\frac{s}{c}x\right).$$

La CC (10.4b) de función finita impone $A = 0$; la CC en $x = 0$ impone

$$B = \frac{a\omega}{\omega^2 + s^2}.$$

Luego

$$\tilde{u}(x, s) = \frac{a\omega}{\omega^2 + s^2} \exp\left(-\frac{s}{c}x\right).$$

Finalmente, antitrasnformando tenemos

$$u(x, t) = \Theta\left(t - \frac{x}{c}\right) a \operatorname{sen}\left(\omega\left(t - \frac{x}{c}\right)\right),$$

mostrando que la perturbación sinusoidal se propaga hacia la derecha con velocidad c . Notemos que, aunque el resultado era esperable, hubiera sido difícil (si no imposible) llegar a él por otros métodos.

Ejercicio 10.1.1. Complete los cálculos anteriores.

Ejercicio 10.1.2. Resuelva el mismo problema, pero donde la perturbación en el extremo $x = 0$ de la soga cesa después de n períodos.

Ejercicio 10.1.3. Resuelva el mismo problema, pero donde la amplitud de la perturbación en el extremo $x = 0$ de la soga decae exponencialmente como $e^{-t/\tau}$.

10.2. Dominios infinitos

Pasaremos a estudiar ahora problemas en los que el dominio de alguna coordenada sea infinito. Ello incluye naturalmente cualquier EDP dependiente del tiempo para la cual busquemos soluciones de régimen, lo cual nuevamente no implica que automáticamente debamos elegir la variable temporal como candidata a una transformación.

10.2.1. Difusión del calor en una barra infinita

Consideremos una barra delgada y homogénea, de longitud infinita y aislada lateralmente, y supongamos que inicialmente se halla a temperatura nula. En el instante $t = 0$ se libera un pulso de calor de magnitud Q_0 en el punto x_0 . El objetivo es hallar la temperatura $u(x, t)$ de la barra para todo x y $t > 0$, bajo la suposición de que esta será finita para todo x y t .

El problema matemático se escribe

$$u_t - \kappa u_{xx} = 0, \quad -\infty < x < \infty, \quad 0 < t < \infty, \quad (10.5a)$$

$$\lim_{x \rightarrow -\infty} |u(x, t)| < \infty, \quad \lim_{x \rightarrow \infty} |u(x, t)| < \infty, \quad (10.5b)$$

$$u(x, 0) = Q_0 \delta(x - x_0). \quad (10.5c)$$

Podríamos estar tentados a tomar transformada de Laplace en t ; sin embargo, como vimos en la Sección 10.1.1, ello nos llevaría a una EDO inhomogénea en x , lo que por ahora preferimos evitar. Optamos entonces por tomar transformada de Fourier en x .

Poniendo

$$\tilde{u}(k, t) = \mathcal{F}_{x \rightarrow k}[u(x, t)]$$

y usando (7.5), resulta el problema transformado

$$\tilde{u}_t + \kappa k^2 \tilde{u} = 0, \quad 0 < t < \infty, \quad (10.6a)$$

$$\tilde{u}(k, 0) = \frac{Q_0}{\sqrt{2\pi}} e^{-ikx_0}. \quad (10.6b)$$

La solución general de la EDO (10.6a) se escribe

$$\tilde{u}(k, t) = A e^{-\kappa k^2 t},$$

y la CC (10.6b) impone

$$A = \frac{Q_0}{\sqrt{2\pi}} e^{-ikx_0},$$

de modo que la solución del problema (10.6) queda

$$\tilde{u}(k, t) = \frac{Q_0}{\sqrt{2\pi}} e^{-(\kappa k^2 t + ikx_0)}.$$

Antitransformando obtenemos

$$u(x, t) = \frac{Q_0}{\sqrt{4\pi\kappa t}} e^{-(x-x_0)^2/4\kappa t},$$

una Gaussiana centrada en x_0 cuyo ancho crece como $2\sqrt{\kappa t}$.

Si en lugar del pulso inicial (10.5c) hubiéramos tenido una distribución inicial de temperatura

$$u(x, 0) = f(x),$$

la solución del problema (10.6) habría quedado

$$\tilde{u}(k, t) = \tilde{f}(k) e^{-\kappa k^2 t}.$$

Aún sin especificar f , podemos obtener en ese caso una solución en forma cerrada: antitransformando y usando (7.12) resulta

$$u(x, t) = \frac{1}{\sqrt{4\pi\kappa t}} \int_{-\infty}^{\infty} e^{-(x-x')^2/4\kappa t} f(x') dx'.$$

Ejercicio 10.2.1. Consideré una barra homogénea infinita aislada lateralmente, con una distribución inicial de temperatura $-T_0$ para $x < 0$ y T_0 para $x > 0$. Halle la temperatura $u(x, t)$ de la barra para todo x y $t > 0$. Ayuda: recuerde la discusión sobre la transformada de Fourier de la función escalón de Heaviside, al principio de la Sección 7.2.

Ejercicio 10.2.2. Una barra semiinfinita, aislada lateralmente, se halla inicialmente a temperatura nula, y su extremo $x = 0$ también se encuentra aislado. En el instante inicial $t = 0$ se inyecta un pulso puntual de calor de magnitud Q_0 en el punto $x_0 > 0$. Encuentre la temperatura de la barra para $x > 0$ y $t > 0$. Ayuda: use el método de las imágenes para reducir este problema al de una barra infinita.

10.2.2. Ondas en una cuerda infinita

Consideremos una cuerda delgada e inextensible de longitud infinita, que inicialmente tiene un perfil de deformación $h(x)$ y un perfil de velocidad $v(x)$, ambos localizados. El objetivo es hallar la deformación $u(x, t)$ de la cuerda para todo x y $t > 0$, bajo la suposición de que la deformación permanece nula en el infinito (a tiempos finitos).

El problema matemático se escribe

$$u_{tt} - c^2 u_{xx} = 0, \quad -\infty < x < \infty, \quad 0 < t < \infty, \quad (10.7a)$$

$$\lim_{x \rightarrow -\infty} |u(x, t)| = 0, \quad \lim_{x \rightarrow \infty} |u(x, t)| = 0, \quad (10.7b)$$

$$u(x, 0) = h(x), \quad u_t(x, 0) = v(x). \quad (10.7c)$$

La naturaleza infinita del dominio en x , junto con las CC correspondientes, parecen sugerir la conveniencia de tomar transformada de Fourier. Poniendo

$$\tilde{u}(k, t) = \mathcal{F}_{x \rightarrow k}[u(x, t)]$$

y usando (7.5), tendremos el problema transformado

$$\tilde{u}_{tt} + c^2 k^2 \tilde{u} = 0, \quad 0 < t < \infty, \quad (10.8a)$$

$$\tilde{u}(k, 0) = \tilde{h}(k), \quad \tilde{u}_t(k, 0) = \tilde{v}(k). \quad (10.8b)$$

La solución general de la EDO (10.8a) es

$$\tilde{u}(k, t) = Ae^{ikct} + Be^{-ikct}.$$

Las CI (10.8b) imponen

$$A + B = \tilde{h}(k), \quad ikc(A - B) = \tilde{v}(k),$$

de donde

$$A = \frac{1}{2} \left(\tilde{h}(k) + \frac{\tilde{v}(k)}{ikc} \right), \quad B = \frac{1}{2} \left(\tilde{h}(k) - \frac{\tilde{v}(k)}{ikc} \right).$$

La solución del problema (10.8) resulta entonces

$$\begin{aligned} \tilde{u}(k, t) &= \tilde{h}(k) \cos(kct) + \tilde{v}(k) \frac{\sin(kct)}{kc} \\ &= \tilde{h}(k) \cos(kct) + \tilde{v}(k) t \operatorname{sinc}(kct) \end{aligned}$$

con lo cual antitransformando y usando (7.12) tenemos

$$\begin{aligned} u(x, t) &= \frac{1}{2} \int_{-\infty}^{\infty} (\delta(x - x' - ct) + \delta(x - x' + ct)) h(x') dx' \\ &\quad + \frac{1}{2c} \int_{-\infty}^{\infty} \operatorname{rect}\left(\frac{x - x'}{2ct}\right) v(x') dx', \end{aligned}$$

que es exactamente la solución de D'Alembert (8.54).

Ejercicio 10.2.3. Complete los cálculos y verifique la afirmación de arriba.

10.2.3. Potencial de una cinta plana

Este problema es un clásico de la electrostática, de difícil resolución a menos que acudamos al método de transformadas integrales. Supongamos que el plano xy es mantenido a potencial nulo, excepto por una cinta o franja de ancho $2a$, que se extiende desde $x = -a$ hasta $x = a$ y es infinita en la dirección de y , la que es mantenida a potencial V_0 . El objetivo es hallar el potencial electrostático $u(x, y, z)$ para todo el espacio, bajo la suposición de que este se anula para $z \rightarrow \pm\infty$.

Notemos que el problema tiene simetría de reflexión respecto del plano xy ; resolveremos primero el problema en el semiespacio $z > 0$. Notemos también que el problema tiene simetría de traslación en y , por lo que u dependerá sólo de x y z , es decir, podemos considerar un problema *plano*. Este problema matemático se escribe

$$\nabla^2 u = 0, \quad -\infty < x < \infty, \quad 0 < z < \infty, \quad (10.9a)$$

$$u(x, 0) = V_0 \operatorname{rect}\left(\frac{x}{2a}\right), \quad \lim_{z \rightarrow \infty} u(x, z) = 0. \quad (10.9b)$$

Vamos a tomar transformada de Fourier en x , con

$$\tilde{u}(k, z) = \mathcal{F}_{x \rightarrow k}[u(x, z)].$$

Usando (7.5), tendremos el problema transformado

$$\tilde{u}_{zz} - k^2 \tilde{u} = 0, \quad 0 < z < \infty, \quad (10.10a)$$

$$\tilde{u}(k, 0) = V_0 \frac{2a}{\sqrt{2\pi}} \operatorname{sinc}(2ak), \quad \lim_{z \rightarrow \infty} u(k, z) = 0. \quad (10.10b)$$

La solución general de la EDO (10.10a) es

$$\tilde{u}(k, z) = Ae^{kz} + Be^{-kz}.$$

Para $k > 0$, la CC (10.10b) asintótica impone $A = 0$, con lo cual la CC en $z = 0$ determina B ; para $k < 0$, tendremos en cambio $B = 0$ y A determinado por la CC en $z = 0$. La solución del problema (10.10) es entonces

$$\tilde{u}(k, z) = V_0 \frac{2a}{\sqrt{2\pi}} \operatorname{sinc}(2ak) e^{-z|k|}.$$

Notando que

$$\mathcal{F}_{k \rightarrow x}^{-1}[e^{-z|k|}] = \sqrt{\frac{2}{\pi}} \frac{z}{x^2 + z^2}$$

y usando (7.12), tendremos

$$u(x, z) = \frac{V_0}{\pi} \int_{x-a}^{x+a} \frac{z}{x'^2 + z^2} dx' = \frac{V_0}{\pi} \left[\arctan\left(\frac{a+x}{z}\right) + \arctan\left(\frac{a-x}{z}\right) \right].$$

Para obtener la solución para $z < 0$ simplemente debemos recordar la paridad del problema en z . Tendremos así que la solución del problema (10.9) extendido a todo el espacio es

$$u(x, z) = \frac{V_0}{\pi} \int_{x-a}^{x+a} \frac{z}{x'^2 + z^2} dx' = \frac{V_0}{\pi} \left[\arctan\left(\frac{a+x}{|z|}\right) + \arctan\left(\frac{a-x}{|z|}\right) \right].$$

Ejercicio 10.2.4. El plano $z = 0$ es mantenido a potencial nulo, excepto por un rectángulo $-a < x < a$, $-b < y < b$ mantenido a potencial V_0 . Muestre que el potencial electrostático $u(x, y, z)$ en todo el espacio es

$$\frac{V_0}{\pi^2} \left[\arctan\left(\frac{a+x}{|z|}\right) + \arctan\left(\frac{a-x}{|z|}\right) \right] \left[\arctan\left(\frac{b+y}{|z|}\right) + \arctan\left(\frac{b-y}{|z|}\right) \right].$$

Ayuda: tome transformada de Fourier en x y en y .

10.2.4. Condensador de cinta plana

Otro clásico problema de la electrostática es una modificación del problema de la Sección 10.2.3. Consideremos el mismo plano a potencial nulo con la cinta a potencial V_0 , pero ahora colocado en $z = \ell$, y un segundo plano a potencial nulo en $z = 0$, formando un condensador. El objetivo es encontrar el potencial electrostático $u(x, y, z)$ en el espacio entre ambos planos.

Como en la Sección 10.2.3, la invariancia traslacional en y nos lleva a plantear un problema bidimensional,

$$\nabla^2 u = 0, \quad -\infty < x < \infty, \quad 0 < z < \ell, \quad (10.11a)$$

$$u(x, 0) = 0, \quad u(x, \ell) = V_0 \operatorname{rect}\left(\frac{x}{2a}\right). \quad (10.11b)$$

Tomando transformada de Fourier en x , el problema transformado queda

$$\tilde{u}_{zz} - k^2 \tilde{u} = 0, \quad 0 < z < \ell, \quad (10.12a)$$

$$\tilde{u}(k, 0) = 0, \quad \tilde{u}(k, \ell) = V_0 \frac{2a}{\sqrt{2\pi}} \operatorname{sinc}(2ak). \quad (10.12b)$$

Es más conveniente esta vez escribir la solución general de la EDO (10.12a) como

$$\tilde{u}(k, z) = A \cosh(kz) + B \operatorname{senh}(kz).$$

La CC (10.12b) en $x = 0$ impone $A = 0$, con lo que imponiendo la CC en $x = \ell$ tenemos

$$\tilde{u}(k, z) = V_0 \frac{2a}{\sqrt{2\pi}} \operatorname{sinc}(2ak) \frac{\operatorname{senh}(kz)}{\operatorname{senh}(k\ell)}$$

como solución del problema (10.12).

Aquí el proceso parece estancarse, ya que la mayoría de las tablas no contienen la antitransformada de $\operatorname{senh}(kz)/\operatorname{senh}(k\ell)$, y las calculadoras *online* no la calculan. Sin embargo podemos ver que no es difícil calcularla por integración en el plano complejo. Tendremos

$$\mathcal{F}_x^{-1} \left[\frac{\operatorname{senh}(kz)}{\operatorname{senh}(k\ell)} \right] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \frac{\operatorname{senh}(kz)}{\operatorname{senh}(k\ell)} e^{ikx} dk,$$

y vemos que el integrando tiene polos simples en $k\ell = in\pi$, con n entero no nulo; la singularidad en $k = 0$ es removible, ya que

$$\lim_{k \rightarrow 0} \frac{\operatorname{senh}(kz)}{\operatorname{senh}(k\ell)} = \frac{z}{\ell}.$$

Los correspondientes residuos son³

$$\operatorname{Res}_{k=in\pi/\ell} \left\{ \frac{\operatorname{senh}(kz)e^{ikx}}{\operatorname{senh}(k\ell)} \right\} = \lim_{k \rightarrow in\pi/\ell} \frac{\operatorname{senh}(kz)e^{ikx}}{\ell \operatorname{cosh}(k\ell)} = \frac{i \operatorname{sen}(n\pi z/\ell) e^{-n\pi x/\ell}}{(-1)^n \ell}.$$

³Apelando al teorema del Análisis Complejo que dice que si $\psi(z)$ es regular en $z = a$ y $\varphi(z)$ tiene un cero simple en $z = a$, entonces

$$\operatorname{Res}_{z=a} \left\{ \frac{\psi(z)}{\varphi(z)} \right\} = \lim_{z \rightarrow a} \frac{\psi(z)}{\varphi'(z)}.$$

Procediendo como en el Ejemplo 7.1.7, vemos que deberemos cerrar el camino de integración con un semicírculo de radio $\rightarrow \infty$ en el semiplano inferior k cuando $x < 0$, y en el semiplano superior cuando $x > 0$. Para $x > 0$ tendremos entonces

$$\begin{aligned}\mathcal{F}^{-1}_{k \rightarrow x} \left[\frac{\operatorname{senh}(kz)}{\operatorname{senh}(k\ell)} \right] &= \frac{1}{\sqrt{2\pi}} 2\pi i \sum_{n=1}^{\infty} \operatorname{Res}_{k=i n \pi/\ell} \left\{ \frac{\operatorname{senh}(kz) e^{ikx}}{\operatorname{senh}(k\ell)} \right\} \\ &= \frac{\sqrt{2\pi}}{\ell} \sum_{n=1}^{\infty} (-1)^{n+1} \operatorname{sen} \left(\frac{n\pi z}{\ell} \right) e^{-n\pi x/\ell}.\end{aligned}$$

Para $x < 0$ el resultado será el mismo cambiando simplemente x por $-x$, ya que el cambio de signo debido al sentido (ahora negativo) del circuito de integración se cancelará con el cambio de signo de n dentro de los senos. El resultado para todo x es entonces

$$\begin{aligned}\mathcal{F}^{-1}_{k \rightarrow x} \left[\frac{\operatorname{senh}(kz)}{\operatorname{senh}(k\ell)} \right] &= \frac{\sqrt{2\pi}}{\ell} \sum_{n=1}^{\infty} (-1)^{n+1} \operatorname{sen} \left(\frac{n\pi z}{\ell} \right) e^{-n\pi|x|/\ell} \\ &= \frac{\sqrt{2\pi}}{2\ell} \frac{\operatorname{sen}(\pi z/\ell)}{\cosh(\pi|x|/\ell) + \cos(\pi z/\ell)},\end{aligned}$$

donde hemos reescrito $\operatorname{sen}(n\pi z/\ell)$ como combinación de exponenciales y hemos sumado las series geométricas resultantes.

Para completar la solución, procedemos como en la Sección 10.2.3. Usando (7.12) tendremos

$$\begin{aligned}u(x, z) &= \frac{V_0}{2\ell} \int_{x-a}^{x+a} \frac{\operatorname{sen}(\pi z/\ell)}{\cosh(\pi x'/\ell) + \cos(\pi z/\ell)} dx' \\ &= \frac{V_0}{2\pi} \arctan \left(\frac{\cos(\pi z/\ell) - 1}{\operatorname{sen}(\pi z/\ell)} \tanh(\pi x'/2\ell) \right) \Big|_{x'=x-a}^{x'=x+a},\end{aligned}$$

donde hemos usado la paridad del coseno hiperbólico para descartar las barras de valor absoluto.

Ejercicio 10.2.5. Verifique el cálculo precedente, y use las identidades

$$\begin{aligned}\operatorname{arctan}(x) - \operatorname{arctan}(y) &= \operatorname{arctan} \left(\frac{x-y}{1-xy} \right), \\ \tanh(x+y) - \tanh(x-y) &= \frac{2 \operatorname{senh}(2y)}{\cosh(2x) + \cosh(2y)}, \\ \tanh(x+y) \tanh(x-y) &= \frac{\cosh(2x) - \cosh(2y)}{\cosh(2x) + \cosh(2y)}\end{aligned}$$

e identidades trigonométricas para simplificar el resultado.

Capítulo 11

EDP Inhomogéneas

Hasta ahora hemos tratado mayormente con EDOs y EDPs (lineales) homogéneas. Las EDOs y los problemas de Sturm–Liouville del Capítulo 5, y las EDPs de Laplace y Helmholtz del Capítulo 9, son todos problemas homogéneos. Las funciones especiales del Capítulo 6 son soluciones de PSL, por definición homogéneos. El método de separación de variables del Capítulo 9 se desarrolló para problemas homogéneos. Incluso en el tratamiento general de las EDPs y la deducción de las ecuaciones de difusión y de ondas en el Capítulo 8 asumimos generalmente la homogeneidad.

Esta insistencia en la homogeneidad no ha sido casual; por el contrario, nos ha proporcionado una gran ventaja a la hora de analizar, por ejemplo, problemas de contorno. Un problema de contorno homogéneo genérico tiene la forma

$$Lu = 0, \quad \vec{x} \in \Omega, \quad (11.1a)$$

$$\alpha \hat{n} \cdot \nabla u + \beta u = 0, \quad \vec{x} \in \partial\Omega, \quad (11.1b)$$

e incluso si las CC son periódicas, seguimos teniendo un problema homogéneo. Poco importa a este respecto si el operador diferencial lo escribimos simplemente como L o como $L + \lambda\rho$; lo importante es que sea lineal, y que ni la ecuación diferencial ni las CC involucren términos donde la función incógnita no aparezca. La ventaja de estos problemas consiste en que *sus soluciones forman un espacio lineal*. Ello *no* ocurre si admitimos inhomogeneidades en la ecuación o las CC: la suma de dos soluciones, aún con una misma inhomogeneidad, *no es una solución del mismo problema*, ya que *las inhomogeneidades se suman*; el espacio de soluciones de un problema inhomogéneo es un *espacio afín*, como queda de manifiesto si notamos que la solución trivial (que juega el papel de elemento nulo) *no* es solución de un problema inhomogéneo.

Pese a las ventajas de la homogeneidad, muchos, si no la mayoría de los problemas provenientes de la Física involucran inhomogeneidades; recordemos sin ir más lejos que, como vimos en la Sección 9.2, la única solución del problema de Laplace homogéneo es la trivial. En el Capítulo 9 hemos tratado repetidamente con problemas donde la EDP era homogénea pero las CC eran inhomogéneas:

en la Sección 9.1.2 y 9.4.3 lo hemos hecho buscando una solución particular de la EDP que satisfaga las CC inhomogéneas y sumándola a la solución general con CC homogéneas; en las Secciones 9.1.3, 9.4.2 y 9.4.5 lo hemos hecho proponiendo formas *ad hoc* de la solución al separar variables; el método más general ha sido elegir adecuadamente (los signos de) las constantes de separación de la EDP para construir soluciones que permitan ajustar CC inhomogéneas, en general de a una curva o superficie frontera por vez, como en las Secciones 9.2.1, 9.2.2, 9.3.1 y 9.4.1. Sólo en una ocasión, en la Sección 9.2.5, hemos intentado adaptar el método de separación de variables a un problema intrínsecamente inhomogéneo, aunque allí lo hicimos de manera poco sistemática. El método de transformadas integrales del Capítulo 10, por su parte, es relativamente inmune a la presencia de inhomogeneidades, sea en la ecuación o las CC, pero deja bastante que desear en cuanto a generalidad y sistematicidad.

En vista de lo anterior, entonces, cabe preguntarse si existirá alguna forma general y sistemática de tratar problemas inhomogéneos, en particular en la ecuación, equivalente en ese aspecto al método de separación de variables. La respuesta, afirmativa, surge del trabajo del molinero y físico-matemático George Green.¹ Probablemente el aspecto más notable es que, en líneas generales, podemos construir soluciones genéricas de problemas inhomogéneos a partir de las de problemas homogéneos; lo único que necesitamos del problema es su *linealidad*.

11.1. La función de Green

Consideremos una ecuación diferencial con un operador lineal L ,

$$Lu(\vec{r}) = f(\vec{r}),$$

y desentendámonos por el momento de las CC (o CI). Si bien como dijimos las soluciones de esta ecuación para diferentes inhomogeneidades f forman un espacio afín, no uno vectorial, la linealidad les confiere una propiedad que podemos explotar: si tenemos una serie de problemas

$$Lu_1(\vec{r}) = f_1(\vec{r}),$$

$$Lu_2(\vec{r}) = f_2(\vec{r}),$$

$$Lu_3(\vec{r}) = f_3(\vec{r}),$$

etc.

¹Un rasgo biográfico notable de George Green es su carácter casi completamente autodicta. Su trabajo *An Essay on the Application of Mathematical Analysis to the Theories of Electricity and Magnetism* (1828) introdujo entre otros conceptos la idea de funciones potencial, el Teorema de Green y las Funciones de Green. Formuló la primera teoría matemática de la electricidad y el magnetismo, sobre la que se basaron Maxwell y Thomson, y su trabajo en la teoría de potenciales estuvo a la par del de Gauss. Otro rasgo notable es que todo esto lo consiguió ya en su madurez (y poco antes de su muerte), tras haberse encargado por muchos años del molino de grano de la familia, lo que hizo decir a Julian Schwinger “Los jóvenes físicos teóricos de hace una o dos generaciones suscribían la creencia de que si no habías hecho algo importante antes de los 30 años, nunca lo harías. Obviamente no estaban familiarizados con la historia de George Green, el molinero de Nottingham.”

todos con *el mismo operador* pero diferentes inhomogeneidades f_i y soluciones u_i , entonces

$$L \left(\sum_i \alpha_i u_i(\vec{r}) \right) = \sum_i \alpha_i f_i(\vec{r}).$$

Notemos que no estamos limitados a combinaciones lineales discretas: si

$$Lu(\vec{r}; \vec{k}) = f(\vec{r}; \vec{k})$$

donde \vec{k} es un parámetro continuo y L actúa sólo sobre \vec{r} y es independiente de \vec{k} , también podemos escribir

$$L \left(\int \alpha(\vec{k}) u(\vec{r}; \vec{k}) d^n k \right) = \int \alpha(\vec{k}) f(\vec{r}; \vec{k}) d^n k,$$

donde $d^n k$ es el diferencial de volumen en el espacio n -dimensional de los parámetros \vec{k} .

Ahora bien, toda inhomogeneidad “razonable” $f(\vec{r})$ puede escribirse como

$$f(\vec{r}) = \int \delta(\vec{r} - \vec{r}') f(\vec{r}') d^n r',$$

que no es sino una combinación lineal de “inhomogeneidades elementales” $\delta(\vec{r} - \vec{r}')$, donde los “coeficientes” de la combinación lineal son $f(\vec{r}')$. Supongamos que hemos conseguido encontrar una $G(\vec{r}, \vec{r}')$, solución de

$$L_{\vec{r}} G(\vec{r}, \vec{r}') = \delta(\vec{r} - \vec{r}'), \quad (11.2)$$

donde hemos agregado a L el subíndice \vec{r} para recordar que actúa sobre \vec{r} y no sobre \vec{r}' (evidentemente G será en general una distribución, y la ecuación deberá entenderse en sentido distribucional). Entonces

$$\begin{aligned} L_{\vec{r}} \int G(\vec{r}, \vec{r}') f(\vec{r}') d^n r' &= \int L_{\vec{r}} G(\vec{r}, \vec{r}') f(\vec{r}') d^n r' \\ &= \int \delta(\vec{r} - \vec{r}') f(\vec{r}') d^n r' = f(\vec{r}), \end{aligned}$$

es decir,

$$u(\vec{r}) = \int G(\vec{r}, \vec{r}') f(\vec{r}') d^n r' \quad (11.3a)$$

será solución de

$$L_{\vec{r}} u(\vec{r}) = f(\vec{r}), \quad (11.3b)$$

y el resultado ¡es válido para cualquier inhomogeneidad $f(\vec{r})$!

Las ecuaciones (11.3) muestran claramente que, si podemos construir tal distribución G , el operador integral

$$\int d^n r' G(\vec{r}, \vec{r}') \quad (11.4)$$

actuará como una *inversa* del operador diferencial $L_{\vec{r}}$. La distribución G solución del problema (11.2), donde *no especificamos CC*, se conoce como *solución fundamental* de dicho problema. Debemos notar que como tal *no es única*, ya que podemos sumarle cualquier solución u_h del correspondiente problema homogéneo

$$L_{\vec{r}}u_h = 0,$$

es decir, cualquier función $u_h(\vec{r}) \in \ker L_{\vec{r}}$, y por linealidad tendremos

$$L_{\vec{r}}(G(\vec{r}, \vec{r}') + u_h(\vec{r})) = L_{\vec{r}}G(\vec{r}, \vec{r}') + L_{\vec{r}}u_h(\vec{r}) = L_{\vec{r}}G(\vec{r}, \vec{r}') = \delta(\vec{r} - \vec{r}').$$

También debemos notar que la solución de (11.3b) adolece de la misma falta de unicidad, y por el mismo motivo. Es claro que esto ocurrirá siempre que $\ker L_{\vec{r}} \neq \emptyset$ y por lo tanto, estrictamente hablando, $L_{\vec{r}}$ no sea invertible, y muestra que en ese caso el operador integral (11.4) será una *pseudoinversa* del operador diferencial $L_{\vec{r}}$.²

Supongamos que ahora queremos resolver un problema como (11.3b) pero *con CC*, digamos

$$L_{\vec{r}}u(\vec{r}) = f(\vec{r}), \quad \vec{r} \in \Omega, \quad (11.5a)$$

$$\alpha(\vec{r})\hat{n} \cdot \nabla u(\vec{r}) + \beta(\vec{r})u(\vec{r}) = 0, \quad \vec{r} \in \partial\Omega. \quad (11.5b)$$

Notemos que estamos asumiendo que las CC son *homogéneas*, lo cual podemos hacer sin perder generalidad, ya que si no lo fueran, como vimos en la Sección 5.3.4 siempre podemos sumar a u una función h que cumpla las CC inhomogéneas, y en el peor de los casos ello simplemente sumará un término $-L_{\vec{r}}h$ a la inhomogeneidad f de la EDP, dejándonos un problema con las mismas CC pero homogéneas. Siguiendo el mismo razonamiento de antes, si ahora conseguimos encontrar una solución del problema

$$L_{\vec{r}}G(\vec{r}, \vec{r}') = \delta(\vec{r} - \vec{r}'), \quad \vec{r}, \vec{r}' \in \Omega, \quad (11.6a)$$

$$\alpha(\vec{r})\hat{n} \cdot \nabla_{\vec{r}}G(\vec{r}, \vec{r}') + \beta(\vec{r})G(\vec{r}, \vec{r}') = 0, \quad \vec{r} \in \partial\Omega, \vec{r}' \in \Omega, \quad (11.6b)$$

que también será una *distribución*, es fácil ver que una función $u(\vec{r})$ definida como

$$u(\vec{r}) = \int_{\Omega} G(\vec{r}, \vec{r}')f(\vec{r}') d^n r' \quad (11.7)$$

no sólo será solución de la EDP (11.5a), sino que también satisfará las CC (11.5b), ya que

$$\begin{aligned} & [\alpha(\vec{r})\hat{n} \cdot \nabla + \beta(\vec{r})] \int_{\Omega} G(\vec{r}, \vec{r}')f(\vec{r}') d^n r' \\ &= \int_{\Omega} [\alpha(\vec{r})\hat{n} \cdot \nabla + \beta(\vec{r})]G(\vec{r}, \vec{r}')f(\vec{r}') d^n r' = 0. \end{aligned}$$

²Las soluciones fundamentales se conocen también como “soluciones de espacio libre”, ya que en principio no establecemos un dominio en cuya frontera debamos aplicar CC, sino que este es todo el espacio; por lo mismo es común en la literatura tratarlas (incluso explícitamente) como “soluciones sin CC”, lo cual en opinión del autor es incorrecto: las soluciones fundamentales *satisfacen CC en el infinito*, bien definidas aunque muchas veces no las elijamos explícitamente.

La distribución G solución del problema (11.6) que provee una solución del problema (11.5) en la forma (11.3a), se conoce como la *función de Green* de este último, y podemos pensar en ella como una solución fundamental a la que le hemos impuesto CC.

Notemos que para lo anterior resulta indispensable que las CC (11.5b) y (11.6b) sean *homogéneas*: si en el lado derecho de (11.6b) tuviésemos $g(\vec{r}, \vec{r}')$ en lugar de cero, habríamos obtenido

$$[\alpha(\vec{r})\hat{n} \cdot \nabla + \beta(\vec{r})] \int_{\Omega} G(\vec{r}, \vec{r}') f(\vec{r}') d^n r' = \int_{\Omega} g(\vec{r}, \vec{r}') f(\vec{r}') d^n r', \quad (11.8)$$

y no está claro cómo conseguir que el lado derecho de esta ecuación reproduzca una posible inhomogeneidad $g(\vec{r})$ en la CC (11.5b), que deberíamos poder elegir *independientemente* de la inhomogeneidad $f(\vec{r})$ en la EDP (11.5a). Aunque esto puede en principio conseguirse, el procedimiento a seguir depende de la forma detallada del operador diferencial $L_{\vec{r}}$; y dado que siempre podemos trasladar la inhomogeneidad de las CC a la EDP por el procedimiento de la Sección 5.3.4, en general este será el camino preferido a seguir. Discutiremos sin embargo, más adelante, el tratamiento del problema para las EDPs más comunes que involucren el operador ∇^2 .

Pasaremos ahora a estudiar cómo construir tanto una solución fundamental como una función de Green para diversos problemas.

11.2. Problemas en dimensión 1

Para adquirir una mejor idea de cómo construir G y qué propiedades esperar de ella, estudiaremos primero problemas unidimensionales. Notemos que estos problemas serán ecuaciones diferenciales *ordinarias*, no parciales.

11.2.1. Soluciones fundamentales

Por una solución fundamental de una EDO entendemos una solución de

$$L_x G(x, x') = \delta(x - x')$$

válida para $-\infty < x < \infty$, con L_x un operador diferencial que actúa sobre la variable x . Si la EDO es a coeficientes constantes, la transformada de Fourier nos provee un camino sencillo a la solución. Como ejemplo busquemos la solución fundamental del problema

$$(\partial_x^2 - a^2)G(x, x') = \delta(x - x'). \quad (11.9)$$

Definiendo

$$\mathcal{F}_{x \rightarrow k}[G(x, x')] = \tilde{G}(k, x')$$

tendremos

$$-(a^2 + k^2)\tilde{G}(k, x') = \frac{1}{\sqrt{2\pi}} e^{-ikx'},$$

de donde

$$\tilde{G}(k, x') = -\frac{1}{\sqrt{2\pi}} \frac{e^{-ikx'}}{a^2 + k^2}$$

y antitrasformando

$$G(x, x') = -\frac{1}{2a} e^{-a|x-x'|}.$$

Con esto ya podemos escribir la solución de

$$(\partial_x^2 - a^2)u(x) = f(x)$$

con una inhomogeneidad f arbitraria, que será

$$u(x) = - \int_{-\infty}^{\infty} \frac{1}{2a} e^{-a|x-x'|} f(x') dx'.$$

Notemos que, implícitamente, hemos impuesto a G ciertas CC en $x \rightarrow \pm\infty$: las necesarias para que su transformada de Fourier exista. Notemos también un rasgo característico de G : su simetría respecto a sus argumentos x y x' .

Pese al éxito obtenido, este método no funcionará si los coeficientes del operador no son constantes. Será provechoso, entonces, usar el mismo ejemplo extremadamente simple para desarrollar un método que pueda tratar con esos casos. Notemos que la EDO (11.9) es *homogénea* para $x \neq x'$. Deberíamos poder, entonces, obtener una solución fundamental *empalmando* en $x = x'$ dos soluciones de la EDO homogénea, válidas para $x < x'$ una y para $x > x'$ la otra, como

$$G(x, x') = \begin{cases} u_-(x), & x < x', \\ u_+(x), & x > x', \end{cases}$$

con

$$\begin{aligned} (\partial_x^2 - a^2)u_-(x) &= 0, & -\infty < x < x', \\ (\partial_x^2 - a^2)u_+(x) &= 0, & x' < x < \infty. \end{aligned}$$

Tendremos

$$\begin{aligned} u_-(x) &= A_- e^{ax} + B_- e^{-ax}, \\ u_+(x) &= A_+ e^{ax} + B_+ e^{-ax}, \end{aligned}$$

y ahora sí deberemos explicitar CC para cada una para $x \rightarrow \pm\infty$, que serán obviamente de función finita e impondrán $B_- = A_+ = 0$, dejando

$$\begin{aligned} u_-(x) &= A_- e^{ax}, \\ u_+(x) &= B_+ e^{-ax}. \end{aligned}$$

Nos quedan por determinar las constantes A_- y B_+ , que justamente nos permitirán empalmar ambas funciones en $x = x'$ para obtener G . De la EDO (11.9) es inmediato que G debe ser continua y con derivada discontinua en $x = x'$,

para que la derivada segunda genere una delta. La segunda condición es mejor tratada integrando (11.9) entre $x' - \epsilon$ y $x' + \epsilon$: tendremos

$$\begin{aligned} & \int_{x'-\epsilon}^{x'+\epsilon} (\partial_x^2 - a^2) G(x, x') dx \\ &= G_x(x' + \epsilon, x') - G_x(x' - \epsilon, x') - \int_{x'-\epsilon}^{x'+\epsilon} a^2 G(x, x') dx = 1. \end{aligned}$$

Tomando $\epsilon \rightarrow 0$, la continuidad de G hará que la segunda integral se anule, y podremos escribir las *condiciones de empalme* como

$$\begin{aligned} u_+(x') - u_-(x') &= 0, \\ u'_+(x') - u'_-(x') &= 1, \end{aligned}$$

es decir

$$\begin{aligned} B_+ e^{-ax'} - A_- e^{ax'} &= 0, \\ B_+ e^{-ax'} + A_- e^{ax'} &= -\frac{1}{a}, \end{aligned}$$

de donde

$$B_+ = -\frac{1}{2a} e^{-ax'}, \quad A_- = -\frac{1}{2a} e^{ax'}.$$

Luego tendremos

$$G(x, x') = -\frac{1}{2a} \begin{cases} e^{a(x-x')}, & x < x', \\ e^{a(x'-x)}, & x > x', \end{cases}$$

es decir

$$G(x, x') = -\frac{1}{2a} e^{-a|x-x'|}.$$

Obviamente hemos recuperado el mismo resultado obtenido antes. Pero lo importante es que es evidente que el *método de empalme* puede tratar con EDOs a coeficientes variables; la única condición es que tengamos la solución general de la ecuación homogénea, la que llegado el caso podemos obtener por los métodos generales de la Sección 5.1. Así, supongamos que tenemos

$$(p(x)\partial_x^2 + q(x)\partial_x + r(x))G(x, x') = \delta(x - x'), \quad (11.10)$$

con p , q y r continuas y $p \neq 0$,³ y que la solución general de la ecuación homogénea sea

$$u(x) = Au_1(x) + Bu_2(x).$$

Definiendo u_- y u_+ como recién tendremos cuatro constantes indeterminadas, y dos CC en $x \rightarrow \pm\infty$ y dos condiciones de empalme. La única diferencia será que

³Si $p(x) = 0$ para algún x finito, tendremos allí un punto singular de la EDO, y el dominio no podrá ser infinito.

al empalmar las derivadas conviene dividir (11.10) por $p(x)$, y obtendremos las condiciones de empalme en la forma

$$\begin{aligned} u_+(x') - u_-(x') &= 0, \\ u'_+(x') - u'_-(x') &= \frac{1}{p(x')}. \end{aligned}$$

11.2.2. Funciones de Green

Por una función de Green de una EDO entendemos una solución de

$$\begin{aligned} L_x G(x, x') &= \delta(x - x'), \quad a < x < b, \\ \text{CC homogéneas en } a \text{ y } b. \end{aligned}$$

Nuevamente vamos a usar un ejemplo simple para presentar los métodos de solución. Vamos a buscar la función de Green del problema

$$(\partial_x^2 + k^2)G(x, x') = \delta(x - x'), \quad 0 < x, x' < a, \quad (11.11a)$$

$$G(0, x') = G(a, x') = 0, \quad 0 < x' < a. \quad (11.11b)$$

Comenzaremos por el método de empalme. La solución general de la EDO homogénea es

$$u(x) = A \cos(kx) + B \sin(kx),$$

de modo que propondremos

$$G(x, x') = \begin{cases} u_-(x), & 0 < x < x', \\ u_+(x), & x' < x < a, \end{cases}$$

donde conviene reescribir la solución general para $x > x'$, dejando

$$\begin{aligned} u_-(x) &= A_- \cos(kx) + B_- \sin(kx), \\ u_+(x) &= A_+ \cos(k(a - x)) + B_+ \sin(k(a - x)). \end{aligned}$$

La CC homogénea en $x = 0$ impone $A_- = 0$ y la CC homogénea en $x = a$ impone $A_+ = 0$, lo que deja

$$\begin{aligned} u_-(x) &= B_- \sin(kx), \\ u_+(x) &= B_+ \sin(k(a - x)). \end{aligned}$$

Las condiciones de empalme quedan

$$\begin{aligned} u_+(x') - u_-(x') &= 0, \\ u'_+(x') - u'_-(x') &= 1, \end{aligned}$$

es decir

$$\begin{aligned} B_+ \sin(k(a - x')) - B_- \sin(kx') &= 0, \\ B_+ \cos(k(a - x')) + B_- \cos(kx') &= -\frac{1}{k}, \end{aligned}$$

con solución

$$B_+ = -\frac{\operatorname{sen}(kx')}{k \operatorname{sen}(ka)}, \quad B_- = -\frac{\operatorname{sen}(k(a-x'))}{k \operatorname{sen}(ka)}.$$

Obtenemos entonces

$$G(x, x') = -\frac{1}{k \operatorname{sen}(ka)} \begin{cases} \operatorname{sen}(kx) \operatorname{sen}(k(a-x')), & 0 < x < x', \\ \operatorname{sen}(kx') \operatorname{sen}(k(a-x)), & x' < x < a, \end{cases}$$

que puede escribirse más brevemente como

$$G(x, x') = -\frac{\operatorname{sen}(kx_<) \operatorname{sen}(k(a-x_>))}{k \operatorname{sen}(ka)}$$

con $x_< = \min(x, x')$ y $x_> = \max(x, x')$. Notemos nuevamente la simetría de G ante el intercambio de x y x' .

Como el dominio ahora es finito, la transformada de Fourier no nos ayudará a construir G por un método alternativo. Sin embargo disponemos de algo estrechamente relacionado e igualmente poderoso: podemos desarrollar G en autofunciones del problema homogéneo. Las autofunciones normalizadas del problema

$$(\partial_x^2 + k^2)u(x) = 0, \quad 0 < x < a, \tag{11.12a}$$

$$u(0) = u(a) = 0 \tag{11.12b}$$

ya sabemos que son

$$\psi_n(x) = \sqrt{\frac{2}{a}} \operatorname{sen}\left(\frac{n\pi}{a}x\right), \quad k_n = \frac{n\pi}{a}, \quad n \in \mathbb{N}.$$

Aquí la notación de Dirac nos ahorrará tiempo y esfuerzo. Ya vimos en la Sección 5.5.3 que si tenemos el problema inhomogéneo

$$(L + \rho\mu)y = \rho f,$$

donde f es una inhomogeneidad dada, y ya hemos hallado los autovalores λ y las autofunciones normalizadas $\psi_\lambda(x)$ del correspondiente PSL (homogéneo)

$$(L + \rho\lambda)y = 0,$$

reescribiendo el problema como

$$(L + \rho\mu)|y\rangle = \rho|f\rangle,$$

obtenemos

$$|y\rangle = \sum_{\lambda} \frac{|\psi_{\lambda}\rangle \langle \psi_{\lambda}|}{\mu - \lambda} |f\rangle.$$

Aplicando este resultado a nuestro caso, queremos la solución del problema

$$(\partial_x^2 + k^2)u(x) = f(x), \quad 0 < x < a, \quad (11.13a)$$

$$u(0) = u(a) = 0. \quad (11.13b)$$

Vemos que $L = \partial_x^2$, $\rho = 1$, $\mu = k^2$, $\lambda = k_n^2$, así que

$$|u\rangle = \sum_{n=1}^{\infty} \frac{|\psi_n\rangle \langle \psi_n|}{k^2 - k_n^2} |f\rangle.$$

Retraducido a la notación con integrales este resultado se expresa

$$u(x) = \int_0^a \sum_{n=1}^{\infty} \frac{\psi_n(x)\psi_n(x')}{k^2 - k_n^2} f(x') dx', \quad (11.14)$$

de donde es inmediato que la función de Green se escribe

$$G(x, x') = \sum_{n=1}^{\infty} \frac{\psi_n(x)\psi_n(x')}{k^2 - k_n^2} = \frac{2}{a} \sum_{n=1}^{\infty} \frac{\sin(k_n x) \sin(k_n x')}{k^2 - k_n^2}. \quad (11.15)$$

Vemos entonces que el *método de desarrollo en autofunciones*, potenciado por el uso de la notación de Dirac, es potente y rápido para la construcción de funciones de Green. Mejor aún, en seguida veremos que es fácilmente extensible a problemas multidimensionales. Notemos además que, nuevamente, G es simétrica ante el intercambio de x y x' , y esta vez el método de desarrollo en autofunciones nos muestra claramente el origen esta propiedad; también nos muestra que, si las autofunciones son a valores complejos, deberemos tener $G(x, x') = G^*(x', x)$.

Antes de pasar al estudio de problemas multidimensionales, cabe preguntarse si es posible obtener funciones de Green para problemas de *valores iniciales*. La respuesta es afirmativa, y la presentaremos mediante un ejemplo simple. Consideraremos un oscilador armónico inicialmente en reposo en el origen, sobre el que actúa desde $t = 0$ en adelante una fuerza dependiente del tiempo. En las variables adecuadas el problema se escribirá

$$(\partial_t^2 + \omega^2)u(t) = f(t), \quad 0 < t < \infty, \quad (11.16a)$$

$$u(0) = u_t(0) = 0. \quad (11.16b)$$

Este problema es ideal para el uso de la transformada de Laplace. Definiendo

$$\tilde{u}(s) = \mathcal{L}_{t \rightarrow s}[u(t)]$$

tendremos

$$\tilde{u}(s) = \frac{\tilde{f}(s)}{s^2 + \omega^2},$$

y antitransformando queda

$$u(t) = \frac{1}{\omega} \int_0^t \sin(\omega(t-t')) f(t') dt'.$$

Como debemos tener

$$u(t) = \int_0^\infty G(t, t') f(t') dt',$$

es inmediato que

$$G(t, t') = \Theta(t - t') \frac{1}{\omega} \sin(\omega(t - t')).$$

Notemos que en este caso G no es simétrica ante el intercambio de sus argumentos. Esto es consistente con la *causalidad* que deberíamos esperar del resultado: sólo los valores de f a tiempos $t' < t$ pueden afectar el valor de u al tiempo t .

11.3. Problemas en dimensión N

Pasaremos ahora al estudio de soluciones fundamentales y funciones de Green para problemas en varias dimensiones espaciales. Los problemas que incluyan el tiempo serán tratados en la Sección 11.4.

11.3.1. Solución fundamental para la ecuación de Laplace

Recordemos que la ecuación de Laplace aparece en relación a estados estacionarios de problemas de difusión del calor y en problemas de electrostática. Su solución fundamental vendrá dada por

$$\nabla_{\vec{r}}^2 G(\vec{r}, \vec{r}') = \delta(\vec{r} - \vec{r}'), \quad (11.17)$$

donde el lado derecho representará una *fuente puntual* unitaria en \vec{r}' : una carga puntual, una fuente puntual de calor, etc.

Comenzaremos suponiendo $\vec{r}, \vec{r}' \in \mathbb{R}^3$. Notemos que la ecuación (11.17) posee simetría translacional, por lo que podemos limitarnos a buscar soluciones de

$$\nabla_{\vec{r}}^2 g(\vec{r}) = \delta(\vec{r}).$$

Para $\vec{r} \neq \vec{0}$ esta ecuación es homogénea, por lo que su solución podrá escribirse, de acuerdo a lo visto en la Sección 9.3.3, como⁴

$$g(\vec{r}) = \left\{ \frac{r^n}{r^{-(n+1)}} \right\} \{Y_{nm}(\theta, \phi)\}.$$

Pero además la ecuación para g posee simetría esférica, así que debemos tener $n = m = 0$ y g queda

$$g(\vec{r}) = A + \frac{B}{r}.$$

Evidentemente, la constante A fija el valor de g para $r \rightarrow \infty$; la elección usual es tomarla nula (lo contrario sólo agrega una solución de la ecuación homogénea

⁴Los índices son enteros y los Q_n^m están ausentes porque el rango de ϕ es $[0, 2\pi]$ con CC periódicas, y el rango de θ es $[0, \pi]$ con CC de función finita sobre el eje z (fuera del origen).

$\nabla^2 g = 0$). Para fijar la constante B acudimos al teorema de la divergencia (Teorema 9.2.1), que en este caso dice que

$$\int_{\Omega} \nabla^2 g(\vec{r}) dV = \int_{\partial\Omega} \hat{n} \cdot \nabla g(\vec{r}) dS.$$

Integrando sobre una esfera de radio R centrada en el origen, resulta

$$1 = \int_{\Omega} \delta(\vec{r}) dV = 4\pi R^2 \left(-\frac{B}{R^2} \right),$$

con lo que $B = -1/4\pi$ y

$$g(\vec{r}) = -\frac{1}{4\pi r}.$$

Pasando de vuelta ahora de \vec{r} a $\vec{r} - \vec{r}'$, tenemos finalmente

$$G(\vec{r}, \vec{r}') = -\frac{1}{4\pi \|\vec{r} - \vec{r}'\|}.$$

Para $\vec{r}, \vec{r}' \in \mathbb{R}^2$ la línea de argumentación es la misma pero, siguiendo la Sección 9.2.2, llegaremos a

$$g(\vec{r}) = \begin{Bmatrix} r^\nu \\ r^{-\nu} \end{Bmatrix} \begin{Bmatrix} \cos(\nu\theta) \\ \sin(\nu\theta) \end{Bmatrix} (+) \begin{Bmatrix} 1 \\ \ln r \end{Bmatrix} \begin{Bmatrix} 1 \\ \theta \end{Bmatrix}$$

que imponiendo CC periódicas en $\theta \in [0, 2\pi]$ y simetría rotacional, se reduce a

$$g(\vec{r}) = A + B \ln r.$$

La constante A podría ser tomada nula de nuevo; pero en este caso no hay elección posible de B (ni de A) que haga que la solución tienda a cero para $r \rightarrow \infty$. Para determinar B recurrimos nuevamente al teorema de la divergencia. El resultado final es

$$g(\vec{r}) = A + \frac{1}{2\pi} \ln r,$$

que pasando nuevamente de \vec{r} a $\vec{r} - \vec{r}'$ nos deja

$$G(\vec{r}, \vec{r}') = A + \frac{1}{2\pi} \ln \|\vec{r} - \vec{r}'\|.$$

Ejercicio 11.3.1. Complete el cálculo anterior.

Ejercicio 11.3.2. Verifique que, en una dimensión, la solución fundamental de

$$\partial_x^2 G(x, x') = \delta(x - x')$$

es

$$G(x, x') = \frac{1}{2} |x - x'|.$$

11.3.2. Solución fundamental para la ecuación de Helmholtz

La solución fundamental para la ecuación de Helmholtz

$$(\nabla_{\vec{r}}^2 + k^2)G(\vec{r}, \vec{r}') = \delta(\vec{r} - \vec{r}')$$

puede construirse siguiendo el mismo procedimiento que para la ecuación de Laplace, y se dejará como ejercicio para el lector.

Ejercicio 11.3.3. Muestre que para $\vec{r}, \vec{r}' \in \mathbb{R}^3$ la solución fundamental de la ecuación de Helmholtz es

$$G(\vec{r}, \vec{r}') = -\frac{k}{4\pi}y_0(k\|\vec{r} - \vec{r}'\|) = -\frac{\cos(k\|\vec{r} - \vec{r}'\|)}{4\pi\|\vec{r} - \vec{r}'\|}.$$

Ejercicio 11.3.4. Muestre que para $\vec{r}, \vec{r}' \in \mathbb{R}^2$ la solución fundamental de la ecuación de Helmholtz es

$$G(\vec{r}, \vec{r}') = \frac{1}{4}Y_0(k\|\vec{r} - \vec{r}'\|).$$

Ayuda: use (6.85).

11.3.3. Función de Green para la ecuación de Laplace

Indudablemente el método más rápido para obtener funciones de Green es el de desarrollo en autofunciones. Daremos algunos ejemplos para CC de Dirichlet, en diferentes dimensiones y sistemas de coordenadas.

Supongamos que estamos en \mathbb{R}^2 y queremos la función de Green para la ecuación de Laplace en el rectángulo $0 < x < a, 0 < y < b$. El problema se escribe

$$\nabla_{\vec{r}}^2 G(\vec{r}, \vec{r}') = \delta(\vec{r} - \vec{r}'), \quad 0 < x, x' < a, \quad 0 < y, y' < b, \quad (11.18a)$$

$$G(\vec{r}, \vec{r}') = 0, \quad \vec{r} = (0, y), (a, y), (x, 0), (x, b). \quad (11.18b)$$

Aquí parecemos encontrar un problema insuperable: porque la solución del correspondiente problema homogéneo

$$\begin{aligned} \nabla^2 u(\vec{r}) &= 0, & 0 < x < a, \quad 0 < y < b, \\ u(0, y) &= u(a, y) = u(x, 0) = u(x, b) = 0 \end{aligned}$$

es idénticamente nula. Sin embargo, recordemos que

$$\varphi_n(x) = \sqrt{\frac{2}{a}} \sin(k_n x), \quad k_n = \frac{n\pi}{a}, \quad n \in \mathbb{N}$$

son un conjunto ortonormal y completo en $L^2(0, a)$ de autofunciones de ∂_x^2 , ya que son las autofunciones del PSL

$$(\partial_x^2 + k^2)u(x) = 0, \quad 0 < x < a, \quad u(0) = u(a) = 0.$$

Análogamente,

$$\phi_m(y) = \sqrt{\frac{2}{b}} \sin(k_m y), \quad k_m = \frac{m\pi}{b}, \quad m \in \mathbb{N}$$

son un conjunto ortonormal y completo en $\mathcal{L}^2(0, b)$ de autofunciones de ∂_y^2 , ya que son las autofunciones del PSL

$$(\partial_y^2 + k^2)u(y) = 0, \quad 0 < y < b, \quad u(0) = u(b) = 0.$$

Luego

$$\psi_{nm}(x, y) = \varphi_n(x)\phi_m(y)$$

serán un conjunto ortonormal y completo en $\mathcal{L}^2[(0, a) \times (0, b)]$ de autofunciones del problema

$$(\nabla^2 + k^2)u(\vec{r}) = 0, \quad \vec{r} \in (0, a) \times (0, b), \quad u(\vec{r}) = 0, \quad \vec{r} = (0, y), (a, y), (x, 0), (x, b)$$

con $k_{nm}^2 = k_n^2 + k_m^2$, y según vimos en la Sección 5.5.2, la completitud implica que

$$\delta(\vec{r} - \vec{r}') = \delta(x - x')\delta(y - y') = \sum_{n,m=1}^{\infty} \psi_{nm}(\vec{r})\psi_{nm}(\vec{r}'),$$

es decir

$$I = \sum_{n,m=1}^{\infty} |\psi_{nm}\rangle \langle \psi_{nm}|.$$

Apelando nuevamente a lo desarrollado en la Sección 5.5.3, entonces, tendremos

$$G = - \sum_{n,m=1}^{\infty} \frac{|\psi_{nm}\rangle \langle \psi_{nm}|}{k_{nm}^2},$$

es decir

$$\begin{aligned} G(\vec{r}, \vec{r}') &= - \sum_{n,m=1}^{\infty} \frac{\psi_{nm}(\vec{r})\psi_{nm}(\vec{r}')}{k_n^2 + k_m^2} \\ &= - \sum_{n,m=1}^{\infty} \frac{\sin(k_n x) \sin(k_m y) \sin(k_n x') \sin(k_m y')}{k_n^2 + k_m^2}. \end{aligned}$$

Este procedimiento formaliza y sistematiza lo que hicimos, trabajosamente, en la Sección 9.2.5.

Antes de aplicar el método de desarrollo en autofunciones a la ecuación de Laplace en otros dominios, veamos qué ocurre si intentamos obtener la función de Green de este mismo problema usando la solución fundamental hallada en la Sección 11.3.1. Para ello debemos proponer una solución del problema (11.18) en la forma

$$G(\vec{r}, \vec{r}') = G_0(\vec{r}, \vec{r}') + u(\vec{r}),$$

donde G_0 es la solución fundamental y u la solución del problema

$$\begin{aligned}\nabla^2 u(\vec{r}) &= 0, \quad 0 < x < a, \quad 0 < y < b, \\ u(0, y) &= -G_0(0, y; x', y'), \quad u(a, y) = -G_0(a, y; x', y'), \\ u(x, 0) &= -G_0(x, 0; x', y'), \quad u(x, b) = -G_0(x, b; x', y').\end{aligned}$$

Dado que

$$G_0(x, y; x', y') = A + \frac{1}{2\pi} \ln \sqrt{(x - x')^2 + (y - y')^2},$$

comparando con el ejemplo de la Sección 9.2.1 vemos que deberemos obtener, por ejemplo, los coeficientes $c_n(x', y')$ del desarrollo

$$A + \frac{1}{4\pi} \ln((x - x')^2 + (b - y')^2) = \sum_{n=1}^{\infty} c_n(x', y') \operatorname{sen}(k_n x) \operatorname{senh}(k_n b)$$

como funciones de x' e y' , y otros equivalentes en los otros tres lados del dominio, lo que no es precisamente sencillo, pero es factible.⁵ Notablemente, el método de empalme puede ser más simple de implementar.⁶

Una vez dilucidada la mecánica del método de desarrollo en autofunciones, es sencillo implementarlo en otros dominios y sistemas de coordenadas. Por ejemplo para el dominio $0 < r < a$ en coordenadas polares r, θ tendremos

$$\nabla_r^2 G(\vec{r}, \vec{r}') = \delta(\vec{r} - \vec{r}'), \quad 0 < r, r' < a, \quad 0 \leq \theta, \theta' < 2\pi, \quad (11.19a)$$

$$G(\vec{r}, \vec{r}') = 0, \quad r = a, \quad 0 < r' < a, \quad 0 \leq \theta, \theta' < 2\pi. \quad (11.19b)$$

Las autofunciones normalizadas de ∇^2 en coordenadas polares planas se escriben

$$\psi_{nm} = \frac{1}{\sqrt{\pi} a J'_{|n|}(\kappa_{|n|m})} J_{|n|} \left(\kappa_{|n|m} \frac{r}{a} \right) e^{in\theta}, \quad n \in \mathbb{Z}, \quad m \in \mathbb{N},$$

y sus autovalores son

$$k_{nm}^2 = \left(\frac{\kappa_{|n|m}}{a} \right)^2,$$

donde κ_{nm} es la m -ésima raíz de J_n ,

$$J_n(\kappa_{nm}) = 0, \quad n \in \mathbb{N}_0, \quad m \in \mathbb{N}.$$

Luego tendremos

$$G = - \sum_{n=-\infty}^{\infty} \sum_{m=1}^{\infty} \frac{|\psi_{nm}\rangle \langle \psi_{nm}|}{k_{nm}^2},$$

⁵La convergencia de las series obtenidas por el método de desarrollo puede ser lenta, sobre todo si \vec{r}' está cerca de la frontera del recinto, lo que lleva a la acumulación de errores numéricos; en esos casos, el “método de splitting” que acabamos de presentar puede resultar más conveniente, pese a su dificultad analítica; véase por ejemplo <http://dx.doi.org/10.4279/PIP.070015>

⁶El lector interesado puede consultar el texto de J. Mathews y R. L. Walker, *Mathematical Methods of Physics*, Addison-Wesley, 1971.

es decir

$$G(r, \theta; r', \theta') = - \sum_{n=-\infty}^{\infty} \sum_{m=1}^{\infty} \frac{J_{|n|}(\kappa_{|n|m} \frac{r}{a}) J_{|n|}(\kappa_{|n|m} \frac{r'}{a}) e^{in(\theta-\theta')}}{k_{nm}^2 \pi a^2 [J'_{|n|}(\kappa_{|n|m})]^2}.$$

Notando que los coeficientes de $e^{in(\theta-\theta')}$ son pares en n , esto puede reescribirse

$$G(r, \theta; r', \theta') = - \sum_{n=0}^{\infty} \varepsilon_n \sum_{m=1}^{\infty} \frac{J_n(\kappa_{nm} \frac{r}{a}) J_n(\kappa_{nm} \frac{r'}{a}) \cos(n(\theta-\theta'))}{\pi \kappa_{nm}^2 [J'_n(\kappa_{nm})]^2},$$

donde

$$\varepsilon_n = \begin{cases} 1, & n = 0, \\ 2, & n > 0. \end{cases}$$

Las funciones de Green para la ecuación de Laplace en coordenadas cilíndricas y esféricas en \mathbb{R}^3 pueden obtenerse de manera similar.

Ejercicio 11.3.5. Muestre que la solución del problema

$$\begin{aligned} \nabla_{\vec{r}}^2 G(\vec{r}, \vec{r}') &= \delta(\vec{r} - \vec{r}'), & 0 < r, r' < a, & 0 < \theta, \theta' < \pi, & 0 \leq \phi, \phi' < 2\pi, \\ G(\vec{r}, \vec{r}') &= 0, & r = a, & 0 < r' < a, & 0 < \theta, \theta' < \pi, & 0 \leq \phi, \phi' < 2\pi, \end{aligned}$$

puede escribirse

$$G(r, \theta, \phi; r', \theta', \phi') = - \sum_{n=0}^{\infty} \sum_{m=-n}^n \sum_{\ell=1}^{\infty} \frac{j_n(\kappa_{n\ell} \frac{r}{a}) j_n(\kappa_{n\ell} \frac{r'}{a}) Y_{nm}(\theta, \phi) Y_{nm}^*(\theta', \phi')}{\kappa_{n\ell}^2 \frac{a}{2} j_{n+1}^2(\kappa_{n\ell})}.$$

Ejercicio 11.3.6. Muestre que la solución del problema

$$\begin{aligned} \nabla_{\vec{r}}^2 G(\vec{r}, \vec{r}') &= \delta(\vec{r} - \vec{r}'), & 0 < r, r' < a, & 0 \leq \theta, \theta' < 2\pi, & 0 < z, z' < h, \\ G(\vec{r}, \vec{r}') &= 0, & r = a \text{ o } z = 0, h, \end{aligned}$$

puede escribirse

$$\begin{aligned} G(r, \theta, z; r', \theta', z') &= - \sum_{m=0}^{\infty} \varepsilon_m \sum_{\ell=1}^{\infty} \sum_{n=1}^{\infty} \frac{J_m(\kappa_{m\ell} \frac{r}{a}) J_m(\kappa_{m\ell} \frac{r'}{a}) \cos(m(\theta-\theta')) \sin(n\pi \frac{z}{h}) \sin(n\pi \frac{z'}{h})}{\left[\frac{\kappa_{m\ell}^2}{a^2} + \frac{n^2 \pi^2}{h^2} \right] \frac{\pi h a^2}{2} J_m'^2(\kappa_{m\ell})}. \end{aligned}$$

Discuta porqué no aparece aquí una frecuencia o longitud de onda de corte, como sí aparece para la guía de ondas de la Sección 9.4.5.

11.3.4. Función de Green para la ecuación de Helmholtz

Las ventajas del método de desarrollo en autofunciones y de la notación de Dirac se ponen de manifiesto cuando pretendemos calcular funciones de Green para la ecuación de Helmholtz. Porque, repasando lo hecho en la Sección 5.5.3

y al comienzo de la Sección 11.3.3, es inmediato que si tenemos problemas de Helmholtz y de Laplace en un mismo dominio y con idénticas CC, tendremos

$$\nabla^2 G = \delta(\vec{r} - \vec{r}') \quad \longrightarrow \quad G = \sum_{\lambda} \frac{|\psi_{\lambda}\rangle \langle \psi_{\lambda}|}{-\lambda}, \quad (11.20a)$$

$$(\nabla^2 + \mu)G = \delta(\vec{r} - \vec{r}') \quad \longrightarrow \quad G = \sum_{\lambda} \frac{|\psi_{\lambda}\rangle \langle \psi_{\lambda}|}{\mu - \lambda}. \quad (11.20b)$$

Podemos construir, entonces, la función de Green de una ecuación de Helmholtz a partir de aquella de la correspondiente ecuación de Laplace con las mismas CC, simplemente reemplazando los autovalores en el denominador.

Como ejemplo, la función de Green del problema

$$(\nabla_{\vec{r}}^2 + k^2)G(\vec{r}, \vec{r}') = \delta(\vec{r} - \vec{r}'), \quad 0 < x, x' < a, \quad 0 < y, y' < b, \quad (11.21a)$$

$$G(\vec{r}, \vec{r}') = 0, \quad \vec{r} = (0, y), (a, y), (x, 0), (x, b) \quad (11.21b)$$

será

$$G(\vec{r}, \vec{r}') = \sum_{n,m=1}^{\infty} \frac{\operatorname{sen}(k_n x) \operatorname{sen}(k_m y) \operatorname{sen}(k_n x') \operatorname{sen}(k_m y')}{k^2 - (k_n^2 + k_m^2)},$$

mientras que la del problema

$$(\nabla_{\vec{r}}^2 + k^2)G(\vec{r}, \vec{r}') = \delta(\vec{r} - \vec{r}'), \quad 0 < r, r' < a, \quad 0 \leq \theta, \theta' < 2\pi, \quad (11.22a)$$

$$G(\vec{r}, \vec{r}') = 0, \quad r = a, \quad 0 < r' < a, \quad 0 \leq \theta, \theta' < 2\pi \quad (11.22b)$$

será

$$\sum_{n=0}^{\infty} \varepsilon_n \sum_{m=1}^{\infty} \frac{J_n(\kappa_{nm} \frac{r}{a}) J_n(\kappa_{nm} \frac{r'}{a}) \cos(n(\theta - \theta'))}{[k^2 - \frac{\kappa_{nm}^2}{a^2}] \pi a^2 [J'_n(\kappa_{nm})]^2}.$$

Ejercicio 11.3.7. Escriba las funciones de Green para los problemas de Helmholtz correspondientes a los de Laplace de los Ejercicios 11.3.5 y 11.3.6.

11.4. Problemas en dimensión $N + 1$

En principio podríamos pensar que la extensión de los resultados de la Sección 11.3 a problemas que incluyan el tiempo debería ser directa; y mayormente tendríamos razón. Sin embargo debemos tener presente que, físicamente, esperamos que la evolución temporal de un sistema respete la causalidad, y esto puede imponer ciertas restricciones a las soluciones fundamentales y funciones de Green que construyamos. En particular, muchas veces esto se manifestará en la imposición de condiciones iniciales a cierto tiempo, aunque sean implícitas. Reconocidamente, esto vuelve algo borrosa la distinción entre soluciones fundamentales y funciones de Green; pese a ello intentaremos mantenerla, dentro de lo razonable.

11.4.1. Solución fundamental para la ecuación de difusión

En principio, en la Sección 10.2.1 ya hemos obtenido la solución fundamental para la ecuación de difusión. Adaptando mínimamente lo hecho allí, vimos que la solución del problema

$$\begin{aligned} u_t - \kappa u_{xx} &= 0, & -\infty < x < \infty, & t_0 < t < \infty, \\ u(x, t_0) &= \delta(x - x_0), \end{aligned}$$

correspondiente a la aplicación de un pulso unitario de calor en x_0 al tiempo t_0 , es

$$u(x, t) = \frac{e^{-(x-x_0)^2/4\kappa(t-t_0)}}{\sqrt{4\pi\kappa(t-t_0)}}.$$

Notemos que $u(x, t)$ satisface la ecuación de difusión *homogénea* para todo $t > t_0$; por otro lado, físicamente esperamos que la solución del problema *antes* de aplicar el pulso, es decir para $t < t_0$, sea la trivial. Es inmediato, entonces, que deberíamos poder extender la validez de la solución a todo t , anterior o posterior, simplemente multiplicándola por una función escalón; definimos entonces

$$G(x, t; x', t') = \Theta(t - t') \frac{e^{-(x-x')^2/4\kappa(t-t')}}{\sqrt{4\pi\kappa(t-t')}},$$

y es fácil verificar que, efectivamente,

$$(\partial_t - \kappa \partial_x^2) G(x, t; x', t') = \delta(x - x') \delta(t - t'),$$

resultado en el cual la presencia de la función escalón resulta fundamental.

Ejercicio 11.4.1. Complete los detalles del cálculo precedente.

El artificio utilizado aquí para incorporar la causalidad y aún así extender una solución con CI a una válida para todo tiempo, es completamente *genérico*. Notando que la extensión del problema a más dimensiones espaciales,

$$\begin{aligned} u_t - \kappa \nabla^2 u &= 0, & \vec{r} \in \mathbb{R}^n, & t_0 < t < \infty, \\ u(\vec{r}, t') &= \delta(\vec{r} - \vec{r}'), \end{aligned}$$

tiene una solución igualmente sencilla, concluimos que la solución de

$$(\partial_t - \kappa \nabla_{\vec{r}}^2) G(\vec{r}, t; \vec{r}', t') = \delta(\vec{r} - \vec{r}') \delta(t - t')$$

será

$$G(\vec{r}, t; \vec{r}', t') = \Theta(t - t') \frac{e^{-\|\vec{r}-\vec{r}'\|^2/4\kappa(t-t')}}{(4\pi\kappa(t-t'))^{n/2}}.$$

La solución del problema

$$\begin{aligned} u_t - \kappa \nabla^2 u &= f(\vec{r}, t), & \vec{r} \in \mathbb{R}^n, & t_0 < t < \infty, \\ u(\vec{r}, t_0) &= g(\vec{r}), \end{aligned}$$

se escribe entonces

$$u(\vec{r}, t) = \int dV' G(\vec{r}, t; \vec{r}', t_0) g(\vec{r}') + \int_{t_0}^t dt' \int dV' G(\vec{r}, t; \vec{r}', t') f(\vec{r}, t').$$

Ejercicio 11.4.2. Demuestre las afirmaciones anteriores.

Ejercicio 11.4.3. Un medio infinito tridimensional de difusividad térmica κ se halla inicialmente a temperatura homogénea T_0 . En el instante $t = 0$ se activa una fuente puntual de calor $q\delta(\vec{r})$ en el origen, y permanece activada. Halle la temperatura para todo \vec{r} y t .

11.4.2. Solución fundamental para la ecuación de ondas

Al igual que con la ecuación de difusión, ya hemos hallado la solución fundamental para la ecuación de ondas unidimensional: la solución de D'Alembert (8.54) del problema

$$\begin{aligned} u_{tt} - c^2 u_{xx} &= 0, \quad -\infty < x < \infty, \quad 0 < t < \infty, \\ u(x, 0) &= h(x), \quad u_t(x, 0) = v(x), \end{aligned}$$

que según lo desarrollado en la Sección 10.2.2 puede escribirse

$$\begin{aligned} u(x, t) &= \frac{1}{2} \int_{-\infty}^{\infty} (\delta(x - x' - ct) + \delta(x - x' + ct)) h(x') dx' \\ &\quad + \frac{1}{2c} \int_{-\infty}^{\infty} \text{rect}\left(\frac{x - x'}{2ct}\right) v(x') dx', \end{aligned}$$

Procediendo como en la Sección 10.2.2,⁷ es bastante directo mostrar que la solución fundamental *causal* del problema

$$(\partial_t^2 - c^2 \partial_x^2) G(x, t; x', t') = \delta(x - x') \delta(t - t')$$

debe escribirse

$$G(x, t; x', t') = \frac{1}{2c} \Theta(t - t') [\Theta((x - x') + c(t - t')) - \Theta((x - x') - c(t - t'))],$$

y la solución *causal* del problema

$$\begin{aligned} u_{tt} - c^2 u_{xx} &= f(x, t), \quad -\infty < x < \infty, \quad t_0 < t < \infty, \\ u(x, t_0) &= h(x), \quad u_t(x, t_0) = v(x) \end{aligned}$$

se escribe

$$\begin{aligned} u(x, t) &= \int_{-\infty}^{\infty} [G(x, t; x', t_0) v(x') + G_t(x, t; x', t_0) h(x')] dx' \\ &\quad + \int_{t_0}^t dt' \int_{-\infty}^{\infty} dx' G(x, t; x', t') f(x', t'), \end{aligned}$$

⁷El lector puede encontrar este desarrollo en el texto de M. Blennow *Mathematical methods for physics and engineering*, CRC Press (2018).

donde

$$G_t(x, t; x', t') = \frac{1}{2} \Theta(t - t') [\delta((x - x') + c(t - t')) + \delta((x - x') - c(t - t'))].$$

Ejercicio 11.4.4. Complete los cálculos precedentes, y verifique que para $f(x, t) = 0$ se recupera la solución de D'Alembert.

Para $\vec{r}, \vec{r}' \in \mathbb{R}^3$, la solución fundamental de la ecuación de ondas satisface

$$\begin{aligned} (\partial_t^2 - c^2 \nabla_{\vec{r}}^2) G(\vec{r}, t; \vec{r}', t') &= \delta(\vec{r} - \vec{r}') \delta(t - t'), \\ G(\vec{r}, t; \vec{r}', t') &= 0, \quad t < t', \end{aligned}$$

donde la única CC que exigimos es la de causalidad. Es inmediato que el problema tiene simetría de traslación espacial y temporal, así que buscaremos la solución de

$$\begin{aligned} (\partial_t^2 - c^2 \nabla^2) g(\vec{r}, t) &= \delta(\vec{r}) \delta(t), \\ g(\vec{r}, t) &= 0, \quad t < 0. \end{aligned}$$

Tomando transformada de Fourier y definiendo

$$\tilde{g}(\vec{k}, t) = \mathcal{F}_{\vec{r} \rightarrow \vec{k}} [g(\vec{r}, t)]$$

tendremos

$$(\partial_t^2 + c^2 k^2) \tilde{g}(\vec{k}, t) = \frac{\delta(t)}{(2\pi)^{3/2}},$$

donde $k = \|\vec{k}\|$. Comparando con el problema (11.16) y su solución, vemos que

$$\tilde{g}(\vec{k}, t) = \frac{\Theta(t)}{(2\pi)^{3/2}} \frac{\operatorname{sen}(kct)}{kc},$$

así que

$$g(\vec{r}, t) = \frac{\Theta(t)}{(2\pi)^3} \iiint \frac{\operatorname{sen}(kct)}{kc} e^{i\vec{k} \cdot \vec{r}} d^3 k.$$

La integración puede realizarse orientando el sistema de coordenadas en el espacio \vec{k} como queramos; lo más sencillo es elegir el eje \hat{k}_z a lo largo de \vec{r} , de modo que $\vec{k} \cdot \vec{r} = kr \cos \theta$ con $r = \|\vec{r}\|$ y θ el ángulo polar de un sistema de coordenadas esféricas (k, θ, ϕ) en el espacio \vec{k} ; la integral sobre ϕ es trivial, y obtenemos

$$\begin{aligned} g(\vec{r}, t) &= \frac{\Theta(t)}{(2\pi)^2 c} \int_0^\infty k dk \operatorname{sen}(kct) \int_0^\pi \operatorname{sen} \theta d\theta e^{ikr \cos \theta} \\ &= -\frac{\Theta(t)}{2\pi c r} (\delta(r + ct) - \delta(r - ct)). \end{aligned}$$

Notando que $r > 0$ y $t > 0$, vemos que la primera delta no contribuirá, por lo que finalmente obtenemos

$$G(\vec{r}, t; \vec{r}', t') = \frac{1}{2\pi c \|\vec{r} - \vec{r}'\|} \Theta(t - t') \delta(\|\vec{r} - \vec{r}'\| - c(t - t')).$$

Ejercicio 11.4.5. Complete los cálculos precedentes.

Ejercicio 11.4.6. Considere un helicóptero en vuelo estacionario a una altura H sobre el origen de coordenadas, con el suelo como plano xy . El helicóptero emite sonido en múltiples frecuencias; si aislamos una, tenemos una ecuación de ondas para la sobrepresión atmosférica $u(\vec{r}, t)$ con una fuente puntual en $(0, 0, H)$ emitiendo con frecuencia ω y condiciones de contorno homogéneas de Neumann en el plano xy . Encuentre u como función de x, y, z, t ($z > 0$). Para una dada distancia horizontal al origen, halle la altura $h_0(\omega)$ a la cual se encuentra el primer mínimo de intensidad para una dada frecuencia ω , debido a la interferencia destructiva entre los frentes de onda directo y reflejado. Muestre que la interferencia hace que para pequeñas alturas las frecuencias graves se refuercen y las agudas se amortigüen. Ayuda: utilice el método de las imágenes para imponer las CC; recuerde que estas son de Neumann.

Ejercicio 11.4.7. Muestre que la solución fundamental para la ecuación de ondas bidimensional es

$$G(\vec{r}, t; \vec{r}', t') = \frac{1}{2\pi c} \frac{\Theta(c(t - t') - \|\vec{r} - \vec{r}'\|)}{\sqrt{c^2(t - t')^2 - (\vec{r} - \vec{r}')^2}}.$$

11.4.3. Función de Green para la ecuación de ondas

Si ahora queremos resolver el problema

$$(\partial_t^2 - c^2 \nabla_{\vec{r}}^2) G(\vec{r}, t; \vec{r}', t') = \delta(\vec{r} - \vec{r}') \delta(t - t'), \quad \vec{r}, \vec{r}' \in \Omega,$$

CC homogéneas en $\partial\Omega$,

$$G(\vec{r}, t; \vec{r}', t') = 0, \quad t < t'$$

como lo hicimos en la Sección 11.4.2, inmediatamente tropezamos con una dificultad: ya no disponemos de la transformada de Fourier, porque el problema está circunscripto a un recinto finito Ω con CC en su frontera $\partial\Omega$. Sin embargo aún podemos usar la transformada de Laplace.

Consideremos el problema

$$(\partial_t^2 - c^2 \nabla_{\vec{r}}^2) u(\vec{r}, \vec{r}'; t) = \delta(\vec{r} - \vec{r}') f(t), \quad \vec{r}, \vec{r}' \in \Omega,$$

CC homogéneas en $\partial\Omega$,

$$u(\vec{r}, \vec{r}'; 0) = u_t(\vec{r}, \vec{r}'; 0) = 0.$$

Definiendo

$$\tilde{u}(\vec{r}, \vec{r}'; s) = \mathcal{L}_{t \rightarrow s}[u(\vec{r}, \vec{r}'; t)]$$

tendremos

$$\left(\nabla_{\vec{r}}^2 - \frac{s^2}{c^2} \right) \tilde{u} = -\frac{\tilde{f}}{c^2} \delta(\vec{r} - \vec{r}')$$

en el mismo recinto y con las mismas CC en su frontera; y este es un problema que ya sabemos resolver, en particular por el método de desarrollo en autofunciones. Porque si tenemos

$$(\nabla^2 + \mu^2)g = \delta(\vec{r} - \vec{r}') \longrightarrow g = \sum_{\lambda} \frac{|\psi_{\lambda}\rangle \langle \psi_{\lambda}|}{\mu^2 - \lambda^2}$$

con $|\psi_{\lambda}\rangle$ autofunciones de ∇^2 de autovalor λ^2 , es decir soluciones de

$$(\nabla^2 + \lambda^2)g = 0$$

en el mismo recinto y con las mismas CC, entonces tendremos

$$\tilde{u} = \sum_{\lambda} \frac{\tilde{f}(s)}{s^2 + c^2\lambda^2} |\psi_{\lambda}\rangle \langle \psi_{\lambda}|,$$

y antitransformando

$$u(\vec{r}, \vec{r}'; t) = \sum_{\lambda} |\psi_{\lambda}\rangle \langle \psi_{\lambda}| \frac{1}{c\lambda} \int_0^t \sin(c\lambda(t-t')) f(t') dt'.$$

Comparando con la resolución del problema (11.16), entonces,

$$u(\vec{r}, \vec{r}'; t) = \sum_{\lambda} \frac{|\psi_{\lambda}\rangle \langle \psi_{\lambda}|}{c\lambda} \int_0^{\infty} \Theta(t-t') \sin(c\lambda(t-t')) f(t') dt',$$

de donde es inmediato que

$$G(\vec{r}, t; \vec{r}', t') = \Theta(t-t') \sum_{\lambda} \frac{|\psi_{\lambda}\rangle \langle \psi_{\lambda}|}{c\lambda} \sin(c\lambda(t-t')).$$

Caveat lector: el que antecede es un desarrollo propio que el autor no ha conseguido encontrar en la literatura, y como tal debe ser sometido a una mirada saludablemente crítica.

Como ejemplo de aplicación del desarrollo que acabamos de hacer, consideremos la ecuación de ondas escalares en el interior de una esfera rígida de radio a , con las consabidas CC de Neumann para el campo de sobrepresiones. El problema para la función de Green se escribirá

$$(\partial_t^2 - c^2 \nabla_{\vec{r}}^2) G(\vec{r}, t; \vec{r}', t') = \delta(\vec{r} - \vec{r}') \delta(t - t'), \quad 0 < r, r' < a,$$

$$G_r(a, \theta, \phi, t; \vec{r}', t') = 0,$$

$$G(\vec{r}, t; \vec{r}', t') = 0, \quad t < t'.$$

De la función de Green para el problema de Helmholtz en la misma geometría (véase el Ejercicio 11.3.7), que se escribe

$$\sum_{n=0}^{\infty} \sum_{m=-n}^n \sum_{\ell=1}^{\infty} \frac{j_n(\kappa_{n\ell} \frac{r}{a}) j_n(\kappa_{n\ell} \frac{r'}{a}) Y_{nm}(\theta, \phi) Y_{nm}^*(\theta', \phi')}{[k^2 - \frac{\kappa_{n\ell}^2}{a^2}] \frac{a^3}{2} j_{n+1}^2(\kappa_{n\ell})},$$

podemos deducir que la función de Green para la ecuación de ondas se escribirá

$$G(\vec{r}, t; \vec{r}', t') = \Theta(t-t') \sum_{n=0}^{\infty} \sum_{m=-n}^n \sum_{\ell=1}^{\infty} \frac{j_n(\kappa_{n\ell} \frac{r}{a}) j_n(\kappa_{n\ell} \frac{r'}{a}) Y_{nm}(\theta, \phi) Y_{nm}^*(\theta', \phi') \sin(c \frac{\kappa_{n\ell}}{a} (t-t'))}{c \kappa_{n\ell} \frac{a^2}{2} j_{n+1}^2(\kappa_{n\ell})}.$$

La construcción de la función de Green para la ecuación de ondas en otras geometrías se deja como ejercicio para el lector, así como la construcción de la función de Green para la ecuación de difusión, que procede a lo largo de las mismas líneas.

11.5. Función de Green para CC inhomogéneas

Hasta ahora, y por muy buenas razones discutidas al final de la Sección 11.1, hemos evitado construir funciones de Green para problemas con inhomogeneidades en las CC. Puede resultar sorprendente, por lo tanto, que la misma función de Green que construimos para un problema con CC homogéneas (la cual satisface esas mismas CC *homogéneas*), pueda usarse para resolver el mismo problema con CC *inhomogéneas*. Veamos cómo.

Si tenemos un operador L autoadjunto y funciones u y v definidas sobre un recinto Ω con CC homogéneas en $\partial\Omega$, tendremos

$$\int_{\Omega} (Lu)v \, dV - \int_{\Omega} (Lv)u \, dV = 0.$$

Pero si u y v satisfacen CC inhomogéneas sobre $\partial\Omega$, tendremos

$$\int_{\Omega} (Lu)v \, dV - \int_{\Omega} (Lv)u \, dV = \text{términos de frontera.}$$

Una expresión de este tipo se conoce como *fórmula de Green*, y su forma exacta depende de L y de Ω , y del tipo de inhomogeneidad de las CC. Pero aún sin tener más detalles, la fórmula de Green nos permite deducir cómo proceder. Porque, supongamos que queremos resolver

$$Lu(\vec{r}) = f(\vec{r}) + \text{CC inhomogéneas.}$$

Poniendo $v(\vec{r}) = G(\vec{r}, \vec{r}')$ en la fórmula de Green, tendremos

$$\int_{\Omega} (LG)u \, dV - \int_{\Omega} (Lu)G \, dV = \text{términos de frontera.}$$

Pero como $LG = \delta(\vec{r} - \vec{r}')$ y $Lu = f$, tendremos

$$u(\vec{r}') = \int_{\Omega} G(\vec{r}, \vec{r}') f(\vec{r}) \, dV + \text{términos de frontera.}$$

La cuestión clave es cómo tratar los términos de frontera. Lo mostraremos con un ejemplo.

Supongamos que nuestro problema es

$$\begin{aligned}\nabla^2 u(\vec{r}) &= f(\vec{r}), \quad \vec{r} \in \Omega, \\ u(\vec{r}) &= h(\vec{r}), \quad \vec{r} \in \partial\Omega.\end{aligned}$$

En ese caso, la fórmula de Green es simplemente la *identidad* de Green

$$\int_{\Omega} (u \nabla^2 v - v \nabla^2 u) dV = \oint_{\partial\Omega} (u \nabla v - v \nabla u) \cdot \hat{n} dS.$$

Tomemos como función de Green la solución de

$$\begin{aligned}\nabla^2 G(\vec{r}, \vec{r}') &= \delta(\vec{r} - \vec{r}'), \quad \vec{r}, \vec{r}' \in \Omega, \\ G(\vec{r}, \vec{r}') &= 0, \quad \vec{r} \in \partial\Omega\end{aligned}$$

y pongamos $v(\vec{r}) = G(\vec{r}, \vec{r}')$ en la identidad de Green; tendremos

$$u(\vec{r}') = \int_{\Omega} G(\vec{r}, \vec{r}') f(\vec{r}) dV + \oint_{\partial\Omega} h(\vec{r}) \nabla_{\vec{r}} G(\vec{r}, \vec{r}') \cdot \hat{n} dS.$$

Notemos que *la misma* función de Green, que satisface CC homogéneas, nos permite tener en cuenta el efecto de las inhomogeneidades en la EDP y en las CC. Además, si como ocurre casi siempre nuestra función de Green es simétrica ante el intercambio de sus argumentos, podemos reescribir este resultado como

$$u(\vec{r}) = \int_{\Omega} G(\vec{r}, \vec{r}') f(\vec{r}') dV' + \oint_{\partial\Omega} \hat{n}(\vec{r}') \cdot \nabla_{\vec{r}'} G(\vec{r}, \vec{r}') h(\vec{r}') dS'.$$

Un procedimiento similar, que se deja al lector, permite hacer algo equivalente para las funciones de Green de otros problemas.

Ejercicio 11.5.1. El plano xy se halla a potencial nulo, excepto por un cuadrado de lado a que se halla a potencial V_0 . Halle el potencial electrostático en todo el semiespacio $z > 0$ usando como función de Green la solución fundamental de la ecuación de Laplace, e integrando sobre la frontera $z = 0$.

