

Equações Diferenciais Parciais

(Versão preliminar, sujeita a mudanças)

Michael Forger

– São Paulo, 17 de agosto de 2023 –

Sumário

Prefácio	vii
Notações e convenções	ix
1 Conceitos e resultados gerais	1
1.1 Introdução	1
1.2 O teorema de Frobenius	6
1.3 O teorema de Cauchy-Kovalevski	10
1.4 Equações e operadores diferenciais lineares	13
1.5 Operadores diferenciais parciais de grau 2	18
1.6 Operadores diferenciais parciais da física	21
1.6.1 Operadores de grau 2	22
1.6.2 Operadores de grau 1	24
1.7 Condições iniciais e condições de fronteira	28
1.8 Solução formal: função de Green e núcleo	32
2 Problemas e Resultados Clássicos	37
2.1 O operador de Laplace	37
2.1.1 O princípio do máximo/mínimo	41
2.1.2 Potencial de Coulomb	42
2.1.3 Funções de Green	51
2.1.4 Teoremas da média esférica e de regularidade	58
2.1.5 Singularidades removíveis	60
2.2 O método de separação de variáveis	62
2.3 O operador de difusão	73
2.3.1 A equação do calor na barra	73

2.3.2	Solução da equação de difusão unidimensional	73
2.3.3	O princípio do máximo/mínimo	73
2.4	O operador de ondas	73
2.4.1	A equação da corda vibrante	73
2.4.2	Solução da equação de ondas unidimensional	74
2.4.3	Causalidade: domínios de dependência e de influência	76
3	Distribuições	77
3.1	Dualidade na álgebra linear	78
3.2	Espaços vetoriais topológicos e espaços localmente convexos	82
3.3	Espaços de funções teste	98
3.4	O conceito de distribuição	105
3.5	Suporte e suporte singular	114
3.6	Diferenciação	118
3.7	Multiplicação	124
3.8	Algumas classes especiais de distribuições	127
3.8.1	Distribuições de suporte compacto	130
3.8.2	Distribuições temperadas	132
3.8.3	Distribuições integráveis	134
3.9	Convergência de seqüências e séries	136
3.10	Distribuições periódicas e séries de Fourier	141
3.11	Produto tensorial	152
3.12	Convolução	158
3.13	Transformação de Fourier	164
3.14	O teorema de Payley-Wiener	171
3.15	O conjunto de frente de ondas	171
3.16	Aplicações a equações diferenciais parciais	171
A	O teorema do fluxo	173
B	Resumo do cálculo vetorial	175
B.1	Integração sobre subvariedades de \mathbb{R}^n : teoria local	175
B.2	Integração sobre subvariedades de \mathbb{R}^n : teoria global	179
B.3	Integração de campos vetoriais	183
B.4	Exemplo: Coordenadas esféricas	184
B.5	Teoremas integrais do cálculo vetorial	189
B.6	Médias esféricas	192

C	Equações diferenciais parciais clássicas na teoria dos campos	195
C.1	Eletrodinâmica: equações de Maxwell, equação de Poisson e equação de ondas	195
C.2	Hidrodinâmica: equações de balanço e equação de difusão	203
	Bibliografia	209

Prefácio

O presente livro originou-se do material apresentado em várias disciplinas sobre equações diferenciais parciais e métodos de física matemática que foram ministradas pelo autor, desde 1993, no Instituto de Matemática e Estatística da Universidade de São Paulo. Trata-se de um texto dirigido a alunos no final de um curso de Bacharelado ou no início de um curso de Mestrado em Matemática ou Matemática Aplicada, visando providenciar uma perspectiva geral da área e ao mesmo tempo uma base para estudos de assuntos mais especializados.

Tendo em vista a enorme quantidade e diversidade do conhecimento matemático acumulado desde que, em 1757, Leonhard Euler estabeleceu as equações de movimento do fluido ideal como o primeiro sistema de equações diferenciais parciais do qual se tem notícia, a escolha de conteúdo e metodologia para uma disciplina de caráter introdutório sobre a área não é uma tarefa fácil. Basicamente, podemos distinguir entre duas abordagens: uma “clássica”, na qual se usa apenas técnicas clássicas de análise, e uma mais “moderna”, que desde o início lança mão do cálculo de distribuições ou funções generalizadas.¹ A opção clássica é adotada por muitos colegas, e frequentemente por motivos predominantemente didáticos, pois pode-se pressupor que os alunos já estejam familiarizados com os fundamentos das técnicas clássicas de análise a serem empregadas, o que facilita a compreensão e o acompanhamento do conteúdo. No entanto, ela também tem as suas desvantagens. Por exemplo, as limitações técnicas inerentes a essa versão da teoria tendem a camuflar a existência de certos conceitos gerais que permitem entender a área como uma só, e não como uma coleção de subáreas com pouco ou nenhum nexos. Concretamente, podemos citar o conceito de função de Green que constitui um elo fundamental entre todas as equações diferenciais parciais – sejam elas elípticas, parabólicas ou hiperbólicas – mas que não admite uma definição geral em termos clássicos: esta ainda é possível (porém tecnicamente bem mais complicada do que no âmbito do cálculo de distribuições) para a equação de Laplace ou Poisson e também para a equação de difusão ou de calor, mas não para a equação de ondas. Como resultado, na abordagem clássica há uma tendência para uma “apresentação por amostragem”, onde métodos importantes e fenômenos típicos são estudados apenas no âmbito de problemas específicos, o que acaba dificultando a meta de fazer

¹ Os dois termos “distribuições” e “funções generalizadas” significam a mesma coisa, sendo que o primeiro foi usado pela escola francesa de L. Schwartz e a segunda pela escola russa de I. Gelfand: ambas introduziram o conceito simultaneamente e independentemente.

com que os alunos desenvolvam uma perspectiva geral da área: muitas vezes, não se passa de uma visão atomizada dos assuntos, sem compreensão mais profunda das conexões que existem entre eles.

E finalmente, mesmo de um ponto de vista puramente didático, a abordagem clássica é um pouco antiquada, excessivamente conservadora, ignorando desenvolvimentos mais modernos, mas isso não é necessariamente o que os alunos querem; estes muitas vezes preferem uma abordagem mais moderna. O sucesso didático não depende tanto da facilidade de compreensão mas, na experiência do autor, do interesse e da motivação dos alunos!

Além disso, abrir mão do moderno cálculo de distribuições que (não é exagero dizer isso) na segunda metade do século 20, provocou uma verdadeira revolução na área de equações diferenciais parciais lineares, constitui-se num retrocesso. Para alunos que querem se aprofundar na área, acarreta uma considerável perda de tempo pois serão obrigados a se familiarizar com essa técnica em outra disciplina, mais avançada, uma vez que não há como fazer matemática em pleno século 21 com as técnicas do século 19 e ignorando as do século 20. Podemos comparar a situação com a do aluno que faz primeiro uma disciplina sobre a integral de Riemann e depois outra sobre a integral de Lebesgue. Quem ainda tem tempo para isso hoje? E para os outros alunos, é pior ainda, pois eles não terão oportunidade para cursar uma disciplina mais avançada da área e assim corrigir eventuais distorções.

Por todos estes motivos, e também por uma questão de preferência pessoal, o autor optou pela abordagem moderna, usando sistematicamente o cálculo de distribuições e, mais especificamente, o método de funções de Green. Isso requer um investimento inicial maior, mas os lucros vem logo. Um deles é a clara percepção de que equações diferenciais parciais constituem uma área só, contribuindo assim à propagação do espírito segundo o qual a matemática não é conhecimento atomizado de alguns assuntos isolados e sim uma teia de conhecimento em que as coisas são fortemente interconectadas. Na opinião do autor, é muito importante que o professor de hoje consiga transmitir esta mensagem ao aluno, que será o professor de amanhã.

Enfatizamos que neste livro, tratamos do cálculo de distribuições e não da teoria de distribuições. ...

Escassez de literatura especializada em português.

Agradecer aos estudantes, entre eles: Luis Abib Finotti, Mário Otávio Salles

São Paulo, junho de 2009

Michael Forger

Notações e convenções

Neste livro, denotaremos por \mathbb{N} o semianel dos *números naturais*, com a convenção de que este contém o número 0, por \mathbb{Z} o anel dos *números inteiros* e por \mathbb{Q} , \mathbb{R} e \mathbb{C} os corpos dos *números racionais*, *reais* e *complexos*, respectivamente. Também usaremos o símbolo \mathbb{F} para denotar um corpo qualquer (de característica 0), mas neste livro, consideraremos apenas os casos em que $\mathbb{F} = \mathbb{R}$ ou $\mathbb{F} = \mathbb{C}$.

Para lidar com espaços e álgebras de funções, observamos que dado um conjunto qualquer X e um espaço vetorial W sobre um corpo \mathbb{F} , o conjunto $\text{Map}(X, W)$ de todas as aplicações f de X em W é, de modo natural, um espaço vetorial sobre \mathbb{F} , no qual a adição e a multiplicação por escalares são definidas “pontualmente”:

$$(\lambda_1 f_1 + \lambda_2 f_2)(x) = \lambda_1 f_1(x) + \lambda_2 f_2(x)$$

para $\lambda_1, \lambda_2 \in \mathbb{F}$, $f_1, f_2 \in \text{Map}(X, W)$, $x \in X$.

No caso em que W for o próprio corpo base \mathbb{F} ou, mais geralmente, uma álgebra A sobre \mathbb{F} , concluímos da mesma maneira que o espaço $\text{Map}(X, \mathbb{F})$ ou, mais geralmente, o espaço $\text{Map}(X, A)$ é, de modo natural, uma álgebra sobre \mathbb{F} , na qual a multiplicação é definida “pontualmente”:

$$(f_1 f_2)(x) = f_1(x) f_2(x)$$

para $f_1, f_2 \in \text{Map}(X, \mathbb{F})$ ou $\text{Map}(X, A)$, $x \in X$.

Finalmente, no caso em que $\mathbb{F} = \mathbb{C}$ e W for o próprio corpo base \mathbb{C} ou, mais geralmente, uma $*$ -álgebra A , concluímos novamente que o espaço $\text{Map}(X, \mathbb{C})$ ou, mais geralmente, o espaço $\text{Map}(X, A)$ é, de modo natural, uma $*$ -álgebra, na qual a conjugação é definida “pontualmente”:

$$f^*(x) = f(x)^*$$

para $f \in \text{Map}(X, \mathbb{F})$ ou $\text{Map}(X, A)$, $x \in X$.

(Resumidamente, o conceito de uma $*$ -álgebra pode ser definido da seguinte forma. Primeiro, uma *involução* em um conjunto X qualquer é uma aplicação de X em X cujo quadrado é a identidade, ou seja, uma bijeção de X que é o seu próprio inverso. Segundo, uma *conjugação* em um espaço vetorial complexo W é uma involução anti-linear $*$ em W . Terceiro, uma *conjugação* em uma álgebra complexa A é uma conjugação $*$ em A como espaço vetorial complexo que também é um antihomomorfismo, i.e., satisfaz $(a_1 a_2)^* = a_2^* a_1^*$, para todo $a_1, a_2 \in A$. Então uma $*$ -álgebra

é uma álgebra complexa associativa A munida de uma tal conjugação, sendo que no caso de uma $*$ -álgebra com unidade, exige-se ainda que $1^* = 1$. Note aqui a inversão da ordem dos fatores sob a conjugação, que se mostra essencial para que a teoria funcione bem no caso não-comutativo, permitindo que a “parte real” de A , ou seja, o subespaço real dos pontos fixos sob a conjugação, pode deixar de ser uma subálgebra: o exemplo mais elementar seria a álgebra das matrizes complexas ($n \times n$) munida da adjunção hermitiana como conjugação. Mas no contexto deste livro, encontraremos apenas $*$ -álgebras de funções que são comutativas e onde, portanto, tal fenômeno não ocorre.)

Passando à álgebra linear e multilinear, notamos que para quaisquer dois espaços vetoriais V e W sobre um corpo \mathbb{F} , as aplicações *lineares* f de V em W formam um subespaço de $\text{Map}(V, W)$ que será denotado por $L(V, W)$. No caso em que W for o próprio corpo base \mathbb{F} , falamos de *formas lineares* sobre V (em vez de aplicações lineares de V em \mathbb{F}): elas constituem o *espaço dual* $V^* = L(V, \mathbb{F})$ de V . Mais geralmente, para $p \geq 1$, as aplicações f de $V \times \dots \times V$ (p fatores) em W que são *p -multilineares*, i.e., satisfazem

$$\begin{aligned} & f(v_1, \dots, v_{i-1}, \lambda_i v_i + \lambda'_i v'_i, v_{i+1}, \dots, v_p) \\ &= \lambda_i f(v_1, \dots, v_{i-1}, v_i, v_{i+1}, \dots, v_p) + \lambda'_i f(v_1, \dots, v_{i-1}, v'_i, v_{i+1}, \dots, v_p) \\ & \text{para } i = 1, \dots, p, \lambda_i, \lambda'_i \in \mathbb{F}, v_1, \dots, v_{i-1}, v_i, v'_i, v_{i+1}, \dots, v_p \in V, \end{aligned}$$

formam um subespaço de $\text{Map}(V \times \dots \times V, W)$ que será denotado por $L^p(V, W)$.² Se W for o próprio corpo base \mathbb{F} , falamos de *formas p -multilineares* sobre V (em vez de aplicações p -multilineares de $V \times \dots \times V$ em \mathbb{F}); tais formas multilineares são também conhecidas como *tensores*. Finalmente, os símbolos $L_s^p(V, W)$ e $L_a^p(V, W)$ denotarão os subespaços de $L^p(V, W)$ de todas as aplicações p -multilineares de $V \times \dots \times V$ (p fatores) em W que são, respectivamente, *simétricas* e *antissimétricas* sob permutações arbitrárias dos seus argumentos. O primeiro destes pode também ser identificado com o espaço dos *polinômios homogêneos de grau p* sobre V a valores em W , pois por definição, um polinômio homogêneo de grau p sobre V a valores em W é uma função P sobre V a valores em W obtida a partir de uma aplicação p -multilinear simétrica f de $V \times \dots \times V$ em W por restrição à *diagonal*

$$\text{diag}(V) = \{(v, \dots, v) \mid v \in V\} \subset V \times \dots \times V,$$

i.e.,

$$P(v) = f(v, \dots, v) \quad \text{para } v \in V.$$

O fato interessante aqui é que o tensor simétrico f já é unicamente determinado pelo polinômio f e pode ser reconstruído a partir de P por uma fórmula explícita chamada “identidade de polarização”, a qual vem em várias versões, sendo uma delas a seguinte:

$$f(v_1, \dots, v_p) = \frac{1}{p!} \sum_{q=1}^p (-1)^{p-q} \sum_{\substack{1 \leq i_1 < \dots < i_q \leq p \\ \text{para } v_1, \dots, v_p \in V}} P(v_{i_1} + \dots + v_{i_q})$$

² Ao invés da expressão “ p -multilinear”, usam-se também as expressões simplificadas “ p -linear” ou, quando for claro qual deve ser o valor de p , “multilinear”.

Por exemplo, é óbvio que para $p = 1$, temos $f = P$ (polinômios homogêneos de grau 1 são funções lineares), enquanto que para $p = 2$, obtemos $f(v_1, v_2) = \frac{1}{2}(P(v_1 + v_2) - P(v_1) - P(v_2))$; para o caso geral, veja [1].

Os espaços $L(V, W)$ e $L_s^p(V, W)$ aparecem naturalmente no cálculo diferencial em espaços vetoriais: dado espaços vetoriais reais V e W de dimensão finita, um aberto Ω de V e uma aplicação

$$u : \Omega \subset V \longrightarrow W$$

de classe C^r ($r \geq 1$), a sua derivada é uma aplicação

$$Du : \Omega \subset V \longrightarrow L(V, W)$$

de classe C^{r-1} e, mais geralmente, a sua p -ésima derivada ($p \leq r$) é uma aplicação

$$D^{(p)}u : \Omega \subset V \longrightarrow L_s^p(V, W)$$

de classe C^{r-p} . Em termos de bases arbitrárias de V e de W , o valor $Du(x)$ da derivada de u em cada ponto x de Ω corresponde à matriz jacobiana de u em x , formada pelas derivadas parciais, em x , das funções que compõem a aplicação u , enquanto que o valor $D^{(p)}u(x)$ da p -ésima derivada de u em cada ponto x de Ω corresponde ao polinômio de Taylor homogêneo de grau p de u em x , cujos coeficientes são as derivadas parciais de ordem p , em x , das funções que compõem a aplicação u . Ocasionalmente, também usaremos uma “versão vetorial” de derivada parcial: supondo que $V = V_1 \times \dots \times V_r$, temos

$$Du(x)(v_1, \dots, v_r) = \sum_{i=1}^r D_i u(x)(v_i) \quad \text{para } x \in \Omega, v_1 \in V_1, \dots, v_r \in V_r$$

com

$$D_i u : \Omega \subset V_1 \times \dots \times V_r \longrightarrow L(V_i, W)$$

de classe C^{r-1} , sendo que, frequentemente, a variável que percorre Ω é da forma $x = (x_1, \dots, x_r)$ e então denotamos $D_i u$ por $D_{x_i} u$. Também sempre suporemos que Ω seja um *domínio* de V , que por definição é um subconjunto aberto e conexo de V , pois caso contrário, podemos sempre nos restringir a considerar, separadamente, cada componente conexa de Ω : isto implica que para $\dim V = 1$, Ω é um intervalo – possivelmente (semi-)infinito.

Maiores detalhes sobre o formalismo do cálculo diferencial em espaços vetoriais de dimensão finita ou espaços de Banach podem ser encontrados em livros texto clássicos tais como [Cartan], [Lang, Vol. 1, Cap. XVI] ou [Lang, Vol. 2, Cap. V].

Continuando, denotaremos, neste livro, por $C^r(\Omega)$ ($C^r(\Omega, W)$) o espaço das funções de classe C^r sobre Ω a valores em \mathbb{F} (em W), onde o símbolo r pode tomar os valores $0, 1, \dots, \infty, \omega$, sendo que a expressão “função de classe C^r ” é sinônima de “função contínua” se $r = 0$, de “função r vezes continuamente diferenciável” se r é um inteiro ≥ 1 , de “função infinitamente diferenciável” ou “suave” ou “lisa” (“smooth”, em inglês) se $r = \infty$ e de “função analítica” se $r = \omega$. (Lembramos

que uma função analítica pode ser definida como uma função suave cuja série de Taylor, em torno de cada ponto do seu domínio, converge absoluta e uniformemente para a função dada em alguma vizinhança do referido ponto.) O conceito de uma função de classe C^r sobre um subconjunto X de V que não é aberto é menos óbvio (exceto, é claro, quando $r = 0$). De fato, supondo que $r \neq 0$ e ainda que $r \neq \omega$, podemos formular (pelo menos) duas definições deste conceito que não são equivalentes. A primeira usa uma abordagem “de dentro para fora” (“outside in”, em inglês) e não requer nenhuma hipótese sobre X , estabelecendo que uma função u definida sobre X é de classe C^r se ela admitir uma extensão de classe C^r a algum aberto maior (i.e., se existirem um aberto Ω de V e uma função \tilde{u} de classe C^r sobre Ω tal que $X \subset \Omega$ e $\tilde{u}|_X = u$). A segunda pode ser caracterizada como uma abordagem “de fora para dentro” (“inside out”, em inglês) e requer que X tenha um interior X° “suficientemente grande”, no sentido de ser contido no fecho deste,

$$X \subset \overline{X^\circ}$$

(note que esta inclusão se torna uma igualdade se X for fechado), estabelecendo que u é de classe C^r sobre X se sua restrição $u|_{X^\circ}$ ao interior X° de X é de classe C^r e todas as derivadas $D^{(p)}(u|_{X^\circ})$ de ordem $p \leq r$, ou equivalentemente, todas as derivadas parciais de $u|_{X^\circ}$ de ordem $\leq r$, admitem extensões contínuas a X . (Nota-se que estas extensões são necessariamente únicas.) Será essa a definição que adotaremos no presente livro.

1 Conceitos e resultados gerais

1.1 Introdução

Começamos a nossa viagem pela paisagem das equações diferenciais parciais com uma simples pergunta: o que é uma equação diferencial parcial? Ou mais geralmente ainda: o que é uma equação diferencial?

De forma extremamente genérica e vaga, podemos tentar responder da seguinte forma: uma equação diferencial é uma equação que estabelece relações entre uma função ou um conjunto de funções e suas derivadas (até uma certa ordem r , digamos). No âmbito deste livro, onde não pretendemos abordar tópicos avançados tais como análise global usando equações diferenciais em variedades, tal conjunto de funções é representado por uma aplicação

$$u : \Omega \longrightarrow U \tag{1.1}$$

de classe C^r , onde Ω é um aberto de \mathbb{R}^n , U é um aberto de \mathbb{R}^N e r pode assumir os valores $1, \dots, \infty, \omega$ (excluímos o valor $r = 0$, pois sem derivadas não há equações diferenciais). Em geral, os números n e N são logicamente distintos e independentes: n é a dimensão do domínio de u , ou seja, o número de variáveis independentes, enquanto que N é a dimensão do codomínio de u , isto é, o número de variáveis dependentes. O primeiro destes números distingue entre equações diferenciais ordinárias e parciais, enquanto que o segundo distingue entre equações escalares e sistemas de equações:

- para equações diferenciais ordinárias, $n = 1$;
- para equações diferenciais parciais, $n > 1$;
- para equações diferenciais escalares, $N = 1$;
- para sistemas de equações diferenciais, $N > 1$.

No entanto, as considerações genéricas a seguir podem na sua maioria ser formuladas para qualquer valor (inteiro positivo) de n e de N , o que permite estabelecer analogias e fazer comparações úteis.

Para escrever uma equação diferencial para a aplicação u , precisamos considerar pelo menos sua primeira derivada, representada por uma aplicação

$$Du : \Omega \longrightarrow L(\mathbb{R}^n, \mathbb{R}^N) \quad (1.2)$$

de classe C^{r-1} , onde $L(\mathbb{R}^n, \mathbb{R}^N)$ denota o espaço das aplicações lineares de \mathbb{R}^n em \mathbb{R}^N . Mais geralmente, para $1 \leq p \leq r$, introduzimos sua p -ésima derivada, representada por uma aplicação

$$D^{(p)}u : \Omega \longrightarrow L_s^p(\mathbb{R}^n, \mathbb{R}^N) \quad (1.3)$$

de classe C^{r-p} , onde $L_s^p(\mathbb{R}^n, \mathbb{R}^N)$ denota o espaço das aplicações p -multilineares simétricas de $\mathbb{R}^n \times \dots \times \mathbb{R}^n$ (p fatores) em \mathbb{R}^N , cuja dimensão é dada pela fórmula

$$\dim L_s^p(\mathbb{R}^n, \mathbb{R}^N) = N \binom{n+p-1}{p}. \quad (1.4)$$

De passagem, notamos que, em cada ponto x de Ω , o polinômio de Taylor homogêneo de grau p de u em x , que corresponde à restrição à diagonal do valor $D^{(p)}u(x)$ da p -ésima derivada de u em x , pode ser escrito na forma de uma derivada comum:

$$D^{(p)}u(x)(h, \dots, h) = \left. \frac{d^p}{dt^p} u(x+th) \right|_{t=0}. \quad (1.5)$$

A seguir, denotaremos, tipicamente, pontos de Ω por x , pontos de U por y , pontos de $L(\mathbb{R}^n, \mathbb{R}^N)$ por y' e pontos de $L_s^p(\mathbb{R}^n, \mathbb{R}^N)$ por $y^{(p)}$.

Com esta notação, uma **equação diferencial** para u de **ordem** ou **grau** r sobre Ω , ou em Ω , é uma equação da forma

$$F(x, u(x), Du(x), \dots, D^{(r)}u(x)) = 0, \quad (1.6)$$

onde x percorre Ω e

$$F : \Omega \times U \times \bigoplus_{p=1}^r L_s^p(\mathbb{R}^n, \mathbb{R}^N) \longrightarrow \mathbb{R}^k \quad (1.7)$$

é uma aplicação dada. Sempre suporemos que F seja, no mínimo, de classe C^1 . Também suporemos que F dependa não-trivialmente da última variável $y^{(r)}$, pois caso contrário, estaríamos tratando de uma equação que na realidade seria de ordem $< r$; isso significa, no mínimo, que a derivada parcial de F em relação a $y^{(r)}$,

$$D_{y^{(r)}}F : \Omega \times U \times \bigoplus_{p=1}^r L_s^p(\mathbb{R}^n, \mathbb{R}^N) \longrightarrow L(L_s^r(\mathbb{R}^n, \mathbb{R}^N), \mathbb{R}^k) \quad (1.8)$$

não deve se anular. (Condições mais restritivas serão formuladas logo adiante.) A equação diferencial (1.6) é chamada **linear** se F for linear em todas as variáveis $y, y', \dots, y^{(r)}$ (sendo que $U = \mathbb{R}^N$), de modo que podemos escrever

$$F(x, y, y', \dots, y^{(r)}) = F_c(x) + F_0(x) \cdot y + F_1(x) \cdot y' + \dots + F_r(x) \cdot y^{(r)} \quad (1.9)$$

com funções $F_c : \Omega \rightarrow \mathbb{R}^k$, $F_0 : \Omega \rightarrow L(\mathbb{R}^n, \mathbb{R}^k)$, $F_1 : \Omega \rightarrow L(L(\mathbb{R}^n, \mathbb{R}^n), \mathbb{R}^k)$ e $F_p : \Omega \rightarrow L(L_s^p(\mathbb{R}^n, \mathbb{R}^n), \mathbb{R}^k)$ ($2 \leq p \leq r$); note que neste caso, a derivada parcial de F em relação a $y^{(r)}$ é exatamente F_r e é uma função apenas de x (independente de $y, y', \dots, y^{(r)}$). Mais geralmente, a equação é chamada **quasilinear** se F for linear apenas na última variável $y^{(r)}$, de modo que podemos escrever

$$F(x, y, y', \dots, y^{(r)}) = F_c(x, y, y', \dots, y^{(r-1)}) + F_l(x, y, y', \dots, y^{(r-1)}) \cdot y^{(r)} \quad (1.10)$$

com funções

$$F_c : \Omega \times U \times \bigoplus_{p=1}^{r-1} L_s^p(\mathbb{R}^n, \mathbb{R}^N) \rightarrow \mathbb{R}^k$$

e

$$F_l : \Omega \times U \times \bigoplus_{p=1}^{r-1} L_s^p(\mathbb{R}^n, \mathbb{R}^N) \rightarrow L(L_s^r(\mathbb{R}^n, \mathbb{R}^N), \mathbb{R}^k)$$

ou em outras palavras, se a derivada parcial de F em relação a $y^{(r)}$ for uma função apenas de $x, y, y', \dots, y^{(r-1)}$ (independente de $y^{(r)}$); se em particular ela for uma função apenas de x (independente de $y, y', \dots, y^{(r)}$), a equação é chamada **semi-linear**.

Para explicar melhor porque equações do tipo (1.6) são chamadas de “equações diferenciais parciais” (quando $n > 1$), é conveniente elaborar sua forma explícita em coordenadas, ou seja, em relação a uma base qualquer: isso é um exercício em álgebra (multi)linear que nos levará naturalmente ao conceito de multi-índice. Tecnicamente, a definição deste conceito é muito simples.

Definição 1.1 *Um **multi-índice** é uma n -upla $\alpha = (\alpha_1, \dots, \alpha_n)$ de inteiros não-negativos $\alpha_1, \dots, \alpha_n$, ou seja, $\alpha \in \mathbb{N}^n$. (Quando precisamos especificar n , falamos de um multi-índice em n dimensões.) A **ordem** ou o **grau** $|\alpha|$ de α é por definição a soma das suas componentes,*

$$|\alpha| = \sum_{i=1}^n \alpha_i. \quad (1.11)$$

Multi-índices proporcionam uma notação eficiente para escrever derivadas parciais de qualquer ordem,¹ baseada na abreviação

$$\partial_i \equiv \frac{\partial}{\partial x_i}, \quad \partial_\alpha \equiv \frac{\partial^{|\alpha|}}{\partial x^\alpha} = \frac{\partial^{|\alpha|}}{\partial x_1^{\alpha_1} \dots \partial x_n^{\alpha_n}}. \quad (1.12)$$

Assim, podemos pensar não apenas na primeira derivada (1.2) da aplicação (1.1) como a família $(\partial_i u)_{i=1, \dots, n}$ de todas as suas derivadas parciais de primeira ordem,

$$\partial_i u : \Omega \rightarrow \mathbb{R}^N, \quad (1.13)$$

¹ Também proporcionam uma notação eficiente para escrever polinômios em n variáveis que constitui uma generalização natural do bem conhecido procedimento em uma variável; este assunto será abordado na Seção 1.4.

mas também na p -ésima derivada (1.3) da aplicação (1.1) como a família $(\partial_\alpha u)_{|\alpha|=p}$ de todas as suas derivadas parciais de ordem p ,

$$\partial_\alpha u : \Omega \longrightarrow \mathbb{R}^N. \quad (1.14)$$

Com esta interpretação, concluímos que a equação diferencial (1.6) pode ser reescrita na forma de um **sistema de equações diferenciais**, ou mais exatamente, um sistema de k equações diferenciais para N funções de n variáveis:

$$F(x, (\partial_\alpha u(x))_{|\alpha| \leq r}) = 0. \quad (1.15)$$

Para caracterizar como a aplicação F depende das derivadas de ordem mais alta, mostrar-se-á útil considerar não apenas a derivada parcial de F “em relação a todas as derivadas parciais de ordem mais alta”, como na equação (1.8) acima, mas também cada derivada parcial de F “em relação à derivada parcial de ordem mais alta em uma determinada direção”, inclusive em função dessa direção. Para exibir o significado matematicamente preciso desta terminologia, note que, em um determinado ponto $(x_0, y_0, y'_0, \dots, y_0^{(r)})$ do domínio $\Omega \times U \times \bigoplus_{p=1}^r L_s^p(\mathbb{R}^n, \mathbb{R}^N)$ de F , a primeira é representada pela expressão

$$D_{y^{(r)}} F(x_0, y_0, y'_0, \dots, y_0^{(r)}) \in L(L_s^r(\mathbb{R}^n, \mathbb{R}^N), \mathbb{R}^k), \quad (1.16)$$

enquanto que a segunda é definida pela expressão

$$D_{y^{(r)}} F(x_0, y_0, y'_0, \dots, y_0^{(r)}) \cdot \xi^r \in L(\mathbb{R}^N, \mathbb{R}^k), \quad (1.17)$$

onde $\xi \in L(\mathbb{R}^n, \mathbb{R})$ é um vetor que pertence ao espaço dual $(\mathbb{R}^n)^*$ de \mathbb{R}^n , ou seja, uma forma linear sobre \mathbb{R}^n ,² e $\xi^r \in L_s^r(\mathbb{R}^n, \mathbb{R})$ é a forma r -multilinear simétrica sobre \mathbb{R}^n definida por $\xi^r(v_1, \dots, v_r) = \xi(v_1) \dots \xi(v_r)$, para $v_1, \dots, v_r \in \mathbb{R}^n$. Note que a expressão (1.17) resulta da expressão (1.16) por inserção, o que se torna possível lançando mão do seguinte isomorfismo canônico:

$$L(L_s^r(\mathbb{R}^n, \mathbb{R}^N), \mathbb{R}^k) \cong L(L_s^r(\mathbb{R}^n, \mathbb{R}), L(\mathbb{R}^N, \mathbb{R}^k)).$$

(Resumidamente, o método mais rápido para entender a construção deste isomorfismo usa dualidade e produtos tensoriais: para quaisquer espaços vetoriais de dimensão finita E , F e G , vale $L(E, F) \cong E^* \otimes F$ e $L_s^r(E, F) \cong L_s^r(E, \mathbb{R}) \otimes F$ e portanto

$$\begin{aligned} L(L_s^r(E, F), G) &\cong (L_s^r(E, \mathbb{R}) \otimes F)^* \otimes G \\ &\cong L_s^r(E, \mathbb{R})^* \otimes F^* \otimes G \cong L(L_s^r(E, \mathbb{R}), L(F, G)). \end{aligned}$$

Omitimos os detalhes porque não terão importância para o que segue.)

Como no caso das equações diferenciais ordinárias, podemos chamar uma equação do tipo (1.6) de **implícita** e um sistema de equações do tipo (1.15) de

² A seguir, chamaremos vetores ξ que pertencem ao espaço dual $(\mathbb{R}^n)^* \equiv L(\mathbb{R}^n, \mathbb{R})$ de \mathbb{R}^n simplesmente de covetores. Uma discussão mais detalhada de espaços duais pode ser encontrada no Capítulo 3.1.

implícito, ao contrário de uma equação **explícita** e de um sistema de equações **explícito** onde certas derivadas de ordem mais alta são colocadas em evidência. Nesta direção, há dois casos extremos para os quais existe uma teoria razoavelmente geral, apresentando um certo grau de semelhança com a teoria das equações diferenciais ordinárias: também são os únicos que parecem ter sido estudados sistematicamente na literatura.

- Colocando *todas* as derivadas parciais de ordem mais alta em evidência, obtemos o que é chamado uma equação diferencial total. Em equações deste tipo, quando consideradas como sistemas, o número k de equações é dado pela fórmula

$$k = N \binom{n+r-1}{r}. \quad (1.18)$$

Como veremos na Seção 1.2, qualquer sistema desse tipo pode ser reduzido a um sistema do mesmo tipo de primeira ordem, sob pena de aumento do número de variáveis dependentes. E para os sistemas desse tipo que são de primeira ordem, temos um teorema geral de existência e unicidade local de soluções: o *teorema de Frobenius*. A novidade é que quando $n > 1$, a existência de soluções, mesmo localmente, pode ser garantida apenas mediante uma certa condição de *integrabilidade*. Também na Seção 1.2, mostraremos como reduzir um sistema geral (implícito) a essa forma (explícita), desde que ele tenha o número correto (1.18) de equações.

- Colocando apenas *uma* das derivadas parciais de ordem mais alta em evidência, chegamos ao tipo de equação estudada no problema de Cauchy. Em equações deste tipo, quando consideradas como sistemas, o número k de equações é simplesmente

$$k = N. \quad (1.19)$$

Como veremos na Seção 1.3, para sistemas desse tipo podemos fixar “dados iniciais” ao longo de uma “hipersuperfície transversal” (em relação à direção em que atua a derivada parcial colocada em evidência) e então tentar resolver a equação resultante como se fosse uma equação diferencial ordinária na qual as demais variáveis aparecem como se fossem apenas parâmetros. E de fato, temos um teorema geral de existência e unicidade local de soluções para este problema: o *teorema de Cauchy-Kovalevski*. Infelizmente, ele só vale no contexto analítico e portanto carece de flexibilidade para se adaptar a situações realísticas. Também na Seção 1.3, mostraremos como reduzir um sistema geral (implícito) a essa forma (explícita), desde que ele tenha o número correto (1.19) de equações: este procedimento requer uma certa adaptação da notação e leva naturalmente à noção de *característica*, sendo que a redução funciona apenas quando a hipersuperfície empregada na fixação dos dados iniciais não é característica.

Como já foi mencionado acima, esses dois casos podem ser considerados extremos no que se refere ao número k de equações, pois no segundo caso k assume o seu mínimo, uma vez que se $k < N$, haveria menos equações do que funções à

disposição e portanto o sistema seria necessariamente subdeterminado, enquanto que no primeiro caso k atinge o seu máximo, pois se $k > N \binom{n+r-1}{r}$, haveria mais equações ainda do que em uma equação diferencial total, que já constitui um sistema fortemente sobredeterminado, com um amplo sistema de condições de integrabilidade, o que parece pouco razoável. Observa-se que quando $n > 1$, o número k de equações no primeiro caso é muito grande e cresce fortemente com a ordem r do sistema, o que já não ocorre no segundo caso. Tal problema não se coloca para equações diferenciais ordinárias, pois quando $n = 1$, os dois números coincidem, para qualquer valor de r . Na verdade, neste caso, ocorrem simplificações drásticas, devido ao fato de que o espaço $L(\mathbb{R}, \mathbb{R}^N)$ e, mais geralmente, para todo valor de p , o espaço $L_s^p(\mathbb{R}, \mathbb{R}^N)$ é canonicamente isomorfo ao espaço \mathbb{R}^N original, o que implica que todas as derivadas da aplicação u têm o mesmo espaço alvo que a aplicação u original.³

Tendo em vista essa discussão, chama atenção o fato de que não parece existir nenhum resultado geral sobre casos intermediários onde o número k de equações encontra-se entre o número mínimo (1.19) e o número máximo (1.18). Por exemplo, poderíamos imaginar formular, para sistemas de $N \binom{m+r-1}{r}$ equações, onde $1 < m < n$, um resultado que seria uma mistura do teorema de Cauchy-Kovalevski, com dados iniciais fixados sobre uma subvariedade “não-característica” de codimensão m , e do teorema de Frobenius, governando a “evolução transversal” m -dimensional sujeita às pertinentes condições de integrabilidade em m dimensões. Mas como não conhecemos nenhum exemplo prático de “evolução multi-temporal”, tal teorema talvez não tenha nenhuma utilidade prática.

1.2 O teorema de Frobenius

Encontrar uma forma padrão para equações diferenciais parciais explícitas nas quais todas as derivadas parciais de ordem mais alta são colocadas em evidência, sendo o número k de equações igual ao número N de funções à disposição multiplicado pelo coeficiente binomial $\binom{n+r-1}{r}$, nos leva a introduzir os seguintes conceitos. Uma **equação diferencial total** de **ordem** ou **grau** r sobre Ω , ou em Ω , é uma equação da forma

$$D^{(r)}u(x) = G(x, u(x), Du(x), \dots, D^{(r-1)}u(x)) \quad (1.20)$$

onde x percorre Ω e

$$G: \Omega \times U \times \bigoplus_{p=1}^{r-1} L_s^p(\mathbb{R}^n, \mathbb{R}^N) \longrightarrow L_s^r(\mathbb{R}^n, \mathbb{R}^N) \quad (1.21)$$

é uma aplicação dada que sempre suporemos ser, no mínimo, de classe C^1 . Esta equação pode ser reescrita na forma de um **sistema total de equações**

³ Deixamos ao leitor a verificação do fato de que o referido isomorfismo canônico pode ser obtido pela regra de levar cada aplicação linear de \mathbb{R} em \mathbb{R}^N para seu valor em $1 \in \mathbb{R}$ e, mais geralmente, cada aplicação p -multilinear simétrica de $\mathbb{R} \times \dots \times \mathbb{R}$ (p fatores) em \mathbb{R}^N para seu valor em $(1, \dots, 1) \in \mathbb{R} \times \dots \times \mathbb{R}$.

diferenciais de ordem r , que é o seguinte sistema de $N \binom{n+r-1}{r}$ equações diferenciais de ordem r para N funções de n variáveis:

$$\begin{aligned} \partial_\alpha u_a(x) &= G_{\alpha,a}(x, (\partial_\beta u_b(x))_{|\beta|<r, 1 \leq b \leq N}) \\ (\alpha \in \mathbb{N}^n, |\alpha| = r, 1 \leq a \leq N) \end{aligned} \quad (1.22)$$

Para chegar nesta forma padrão, considere a equação diferencial implícita (1.6), ou equivalentemente, o sistema de equações diferenciais implícito (1.15), com $k = N \binom{n+r-1}{r}$, e note que, neste caso, em qualquer ponto do domínio de F , a derivada parcial de F em relação a todas as derivadas parciais de ordem r , introduzida na equação (1.16), é uma matriz quadrática. Usando a hipótese de que, em um determinado ponto do domínio de F , esta seja não-singular,⁴

$$\det(D_{y^{(r)}} F(x_0, y_0, y'_0, \dots, y_0^{(r)})) \neq 0, \quad (1.23)$$

podemos usar o teorema das funções implícitas para garantir, em uma vizinhança aberta suficientemente pequena do referido ponto, a possibilidade de reduzir a equação (1.6) à forma da equação (1.20), ou equivalentemente, o sistema (1.15) à forma do sistema (1.22).

Dada uma equação diferencial total de ordem r para uma aplicação $u : \Omega \rightarrow U$, como na equação (1.20) acima, definida em termos de uma função G como na equação (1.21) acima, podemos reduzi-la a uma equação diferencial total de primeira ordem para uma aplicação $\tilde{u} : \Omega \rightarrow \tilde{U}$, ou seja, a uma equação da forma

$$D\tilde{u}(x) = \tilde{G}(x, \tilde{u}(x))$$

onde x percorre Ω e

$$\tilde{G} : \Omega \times \tilde{U} \rightarrow L(\mathbb{R}^n, \mathbb{R}^{\tilde{N}})$$

é uma aplicação dada de classe C^1 , no mínimo, pondo

$$\tilde{U} = U \times \bigoplus_{p=1}^{r-1} L_s^p(\mathbb{R}^n, \mathbb{R}^N) \subset \mathbb{R}^{\tilde{N}} = \mathbb{R}^N \oplus \bigoplus_{p=1}^{r-1} L_s^p(\mathbb{R}^n, \mathbb{R}^N)$$

com

$$\tilde{u}(x) = (u(x), Du(x), \dots, D^{(r-1)}u(x))$$

e

$$\tilde{G}(x, y, y', \dots, y^{(r-1)}) = (y', \dots, y^{(r-1)}, G(x, y, y', \dots, y^{(r-1)}))$$

onde usamos a inclusão natural $L_s^p(\mathbb{R}^n, \mathbb{R}^N) \subset L(\mathbb{R}^n, L_s^{p-1}(\mathbb{R}^n, \mathbb{R}^N))$.

Mudando de notação para fins de simplificação, consideramos no restante desta seção apenas equações diferenciais totais de primeira ordem,

$$Du(x) = G(x, u(x)) \quad (1.24)$$

⁴Note que esta é uma condição mais restritiva do que a condição anterior, formulada no contexto da equação (1.8), de que a derivada parcial na equação (1.16) não se anule.

onde x percorre Ω e

$$G : \Omega \times U \longrightarrow L(\mathbb{R}^n, \mathbb{R}^N) \quad (1.25)$$

é uma aplicação dada que sempre suporemos ser, no mínimo, de classe C^1 . Esta equação pode ser reescrita na forma de um sistema total de equações diferenciais de primeira ordem, que é o seguinte sistema de Nn equações diferenciais de primeira ordem para N funções de n variáveis:

$$\begin{aligned} \frac{\partial u_a}{\partial x_i}(x) &= G_{i,a}(x, u(x)) \\ (1 \leq i \leq n, 1 \leq a \leq N) \end{aligned} \quad (1.26)$$

Diremos que esta equação ou este sistema de equações é **integrável** se para todo ponto x de Ω e todo ponto y de U , existe uma única solução local da equação (1.24) ou do sistema de equações (1.26) que em x assume o valor y , ou seja, existem uma vizinhança aberta Ω_x de x contida em Ω e uma única aplicação $u(\cdot; x, y) : \Omega_x \longrightarrow U$ de classe C^1 , no mínimo, tal que

$$D_z u(z; x, y) = G(z, u(z; x, y)) \quad (1.27)$$

ou equivalentemente

$$\begin{aligned} \frac{\partial u_a}{\partial z_i}(z; x, y) &= G_{i,a}(z, u(z; x, y)) \\ (1 \leq i \leq n, 1 \leq a \leq N) \end{aligned} \quad (1.28)$$

e

$$u(x; x, y) = y. \quad (1.29)$$

Quando isso for o caso, podemos diferenciar a equação (1.24) para expressar a segunda derivada de cada uma dessas suas soluções u em termos das derivadas parciais

$$\begin{aligned} D_x G : \Omega \times U &\longrightarrow L^2(\mathbb{R}^n, \mathbb{R}^N) \cong L(\mathbb{R}^n, L(\mathbb{R}^n, \mathbb{R}^N)) \\ (x, y) &\longmapsto D_x G(x, y) \end{aligned}$$

e

$$\begin{aligned} D_y G : \Omega \times U &\longrightarrow L(\mathbb{R}^N, L(\mathbb{R}^n, \mathbb{R}^N)) \\ (x, y) &\longmapsto D_y G(x, y) \end{aligned}$$

sendo que $D_x G(x, y) \cdot (v_1, v_2) = (D_x G(x, y) \cdot v_1) \cdot v_2$ também é a derivada no ponto (x, y) e na direção v_1 da aplicação

$$\begin{aligned} \Omega &\longrightarrow \mathbb{R}^N \\ z &\longmapsto G(z, y) \cdot v_2 \end{aligned}$$

e $D_y G(x, y) \cdot (w, v) = (D_y G(x, y) \cdot w) \cdot v$ também é a derivada no ponto (x, y) e na direção w da aplicação

$$\begin{aligned} U &\longrightarrow \mathbb{R}^N \\ z &\longmapsto G(x, z) \cdot v \end{aligned}$$

Aplicando a regra da cadeia, obtemos

$$\begin{aligned} D^{(2)}u(x) \cdot (v_1, v_2) &= D_x G(x, u(x)) \cdot (v_1, v_2) + D_y G(x, u(x)) \cdot (Du(x) \cdot v_1, v_2) \\ &= D_x G(x, u(x)) \cdot (v_1, v_2) + D_y G(x, u(x)) \cdot (G(x, u(x)) \cdot v_1, v_2) \end{aligned}$$

ou equivalentemente, diferenciando novamente o sistema de equações (1.26),

$$\begin{aligned} \frac{\partial^2 u_a}{\partial x_i \partial x_j} &= \frac{\partial}{\partial x_i} \left(G_{j,a}(\cdot, u(\cdot)) \right) = \frac{\partial G_{j,a}}{\partial x_i}(\cdot, u(\cdot)) + \frac{\partial G_{j,a}}{\partial y_b}(\cdot, u(\cdot)) \frac{\partial u_b}{\partial x_i} \\ &= \left(\frac{\partial G_{j,a}}{\partial x_i} + \frac{\partial G_{j,a}}{\partial y_b} G_{i,b} \right) (\cdot, u(\cdot)) \end{aligned}$$

Assim, o simples fato de que a segunda derivada deve ser simétrica nos leva à **condição de integrabilidade de Frobenius**

$$\begin{aligned} D_x G(x, y) \cdot (v_1, v_2) + D_y G(x, y) \cdot (G(x, y) \cdot v_1, v_2) \\ = D_x G(x, y) \cdot (v_2, v_1) + D_y G(x, y) \cdot (G(x, y) \cdot v_2, v_1) \end{aligned} \quad (1.30)$$

ou equivalentemente,

$$\begin{aligned} \frac{\partial G_{j,a}}{\partial x_i} + \frac{\partial G_{j,a}}{\partial y_b} G_{i,b} = \frac{\partial G_{i,a}}{\partial x_j} + \frac{\partial G_{i,a}}{\partial y_b} G_{j,b} \quad . \quad (1.31) \\ (1 \leq i, j \leq n, 1 \leq a \leq N) \end{aligned}$$

Teorema 1.1 (Teorema de Frobenius): *Para que uma equação diferencial total de primeira ordem como na equação (1.24), dada por uma aplicação G de classe C^p como na equação (1.25), com $1 \leq p \leq \omega$, ou equivalentemente, um sistema total de equações diferenciais como na equação (1.26), seja integrável, é necessário e suficiente que G satisfaça à condição de integrabilidade de Frobenius (1.30), ou equivalentemente, (1.31). Neste caso, a solução da equação e a sua dependência das condições iniciais também são de classe C^p , i.e., para todo ponto x_0 de Ω e todo ponto y_0 de U existem uma vizinhança aberta $\Omega_0 \subset \Omega$ de x_0 , uma vizinhança aberta $U_0 \subset U$ de y_0 e uma única aplicação $u : \Omega_0 \times \Omega_0 \times U_0 \rightarrow U_0$ de classe C^p satisfazendo a equação diferencial (1.27), ou equivalentemente, o sistema de equações diferenciais (1.28), assim como a condição inicial (1.29), para todo $z, x \in \Omega_0, y \in U_0$.*

Para a demonstração deste teorema, veja o livro texto clássico [Dieudonné], Cap. X, Seção 9.

Um caso particularmente importante é o de uma equação diferencial total *linear* de primeira ordem,

$$Du(x) = -A(x) \cdot u(x) \quad (1.32)$$

onde x percorre Ω e

$$A : \Omega \rightarrow L(\mathbb{R}^N, L(\mathbb{R}^n, \mathbb{R}^N)) \quad (1.33)$$

é uma aplicação dada de classe C^1 , no mínimo, ou seja, temos $G(x, y) = -A(x) \cdot y$, com $U = \mathbb{R}^N$.⁵ Esta equação pode ser reescrita na forma de um sistema total linear de equações diferenciais de primeira ordem, que é o seguinte sistema linear de Nn equações diferenciais para N funções de n variáveis:

$$\frac{\partial u_a}{\partial x_i}(x) = -A_{i,ab}(x) u_b(x) \quad (1.34)$$

$$(1 \leq i \leq n, 1 \leq a \leq N)$$

Então é fácil ver que, assim como as próprias equações, as condições de integrabilidade de Frobenius também são lineares na variável y e se reduzem à condição de que a expressão definida por

$$F_{ij,ab} = \frac{\partial A_{j,ab}}{\partial x_i} - \frac{\partial A_{i,ab}}{\partial x_j} + A_{i,ac} A_{j,cb} - A_{j,ac} A_{i,cb} \quad (1.35)$$

$$(1 \leq i, j \leq n, 1 \leq a, b \leq N)$$

ou em notação matricial

$$F_{ij} = \frac{\partial A_j}{\partial x_i} - \frac{\partial A_i}{\partial x_j} + [A_i, A_j] \quad (1.36)$$

$$(1 \leq i, j \leq n)$$

se anula. Nesta forma, a condição de integrabilidade de Frobenius é bem conhecida como a **condição de curvatura zero** e aparece como tal em várias áreas da ciência, por exemplo em geometria diferencial, em teorias de calibre e na teoria de sistemas integráveis.

1.3 O teorema de Cauchy-Kovalevski

Encontrar uma forma padrão para equações diferenciais parciais explícitas nas quais apenas uma derivada parcial de ordem mais alta é colocada em evidência, sendo o número k de equações igual ao número N de funções à disposição, requer escolher as coordenadas de \mathbb{R}^n de tal modo que a derivada parcial a ser isolada seja na direção de uma das coordenadas: chamaremos esta de x_0 e as demais de x_1, \dots, x_{n-1} e usaremos letras em negrito para denotar objetos em $n-1$ dimensões: assim, o símbolo \mathbf{x} denotará um vetor em \mathbb{R}^{n-1} enquanto que o símbolo x continua denotando um vetor em \mathbb{R}^n , ou seja,

$$x = \begin{pmatrix} x_0 \\ \mathbf{x} \end{pmatrix} \quad \text{onde} \quad \mathbf{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_{n-1} \end{pmatrix},$$

⁵O sinal negativo é introduzido apenas para estabelecer coerência com algumas convenções a serem adotadas posteriormente.

e o símbolo α denotará multi-índices em \mathbb{N}^{n-1} enquanto que o símbolo a continua denotando multi-índices em \mathbb{N}^n . Com esta notação, a equação em questão pode ser escrita na forma do seguinte sistema explícito de N equações diferenciais para N funções de n variáveis:

$$\begin{aligned} \partial_0^r u_a(x) &= G_a(x, (\partial_0^q \partial_\alpha u_b(x))_{q+|\alpha| \leq r, q < r, 1 \leq b \leq N}) \\ (1 \leq a \leq N) \end{aligned} \quad (1.37)$$

Para chegar nesta forma padrão, considere a equação diferencial implícita (1.6), ou equivalentemente, o sistema de equações diferenciais implícito (1.15), com $k = N$, e note que, neste caso, em qualquer ponto do domínio $\Omega \times U \times \bigoplus_{p=1}^r L_s^p(\mathbb{R}^n, \mathbb{R}^N)$ de F , a derivada parcial de F em relação à derivada parcial de ordem r na direção de um covetor ξ qualquer, conforme definida na equação (1.17), é uma matriz quadrática.

Definição 1.2 *Para uma equação diferencial (implícita) de ordem r como na equação (1.6), dada por uma aplicação F de classe C^p como na equação (1.7), com $1 \leq p \leq \omega$, ou equivalentemente, um sistema de equações diferenciais de ordem r como na equação (1.15), um covetor ξ é dito **característico** no ponto $(x_0, y_0, y'_0, \dots, y_0^{(r)})$ do domínio de F se*

$$\det(D_{y^{(r)}} F(x_0, y_0, y'_0, \dots, y_0^{(r)}) \cdot \xi^r) = 0, \quad (1.38)$$

e um hiperplano H de \mathbb{R}^n é dito **característico** no ponto $(x_0, y_0, y'_0, \dots, y_0^{(r)})$ do domínio de F se é o núcleo de algum covetor característico neste ponto.⁶

Usando a hipótese de que, em um determinado ponto do domínio de F , ξ seja não-característico, ou seja, esta matriz seja não-singular,⁷

$$\det(D_{y^{(r)}} F(x_0, y_0, y'_0, \dots, y_0^{(r)}) \cdot \xi^r) \neq 0, \quad (1.39)$$

e efetuando uma transformação linear das antigas coordenadas para novas coordenadas x_0, x_1, \dots, x_{n-1} tais que $\xi = dx_0$, podemos usar o teorema das funções implícitas para garantir, em uma vizinhança aberta suficientemente pequena do referido ponto, a possibilidade de reduzir o sistema (1.15) à forma do sistema (1.37).

Voltando a considerar o sistema de equações diferenciais em questão na sua forma explícita (1.37), podemos e devemos complementá-lo por condições iniciais, dados por uma coleção de aplicações dadas

$$\begin{aligned} u^{(0)} : \Omega \cap H &\longrightarrow U \\ u^{(1)} : \Omega \cap H &\longrightarrow \mathbb{R}^N \quad \dots \quad u^{(r-1)} : \Omega \cap H &\longrightarrow \mathbb{R}^N \end{aligned} \quad (1.40)$$

⁶Esta definição faz sentido porque, conforme o critério (1.38), múltiplos de covetores (não-) característicos também são (não-)característicos e porque dois covetores se anulam no mesmo hiperplano se e somente se são proporcionais.

⁷Note que a existência de covetores não-característicos é uma condição mais restritiva do que a condição anterior, formulada no contexto da equação (1.8), de que a derivada parcial na equação (1.16) não se anule. (Para $N = 1$, ambas são equivalentes.)

onde H denota o hiperplano de \mathbb{R}^n definido por

$$H = \{x \in \mathbb{R}^n \mid x_0 = 0\}, \quad (1.41)$$

para chegar ao

Problema de Cauchy: Resolver o sistema explícito de equações diferenciais

$$\partial_0^r u_a(x) = G_a(x, (\partial_0^q \partial_\alpha u_b(x))_{q+|\alpha| \leq r, q < r, 1 \leq b \leq N}) \quad (1.42)$$

$$(1 \leq a \leq N)$$

sujeito às condições iniciais

$$\partial_0^q u_a(0, \mathbf{x}) = (u^{(q)})_a(\mathbf{x}) \quad (1.43)$$

$$(0 \leq q < r, 1 \leq a \leq N)$$

Notamos que juntas, as condições (1.42) e (1.43) permitem determinar *todas* as derivadas parciais da solução u sobre $\Omega \cap H$, pois para qualquer $\alpha \in \mathbb{N}^{n-1}$, as derivadas parciais da forma $\partial_0^q \partial_\alpha u$ com $q < r$ podem ser obtidas diretamente a partir da equação (1.43), aplicando o operador ∂_α aos dois lados, enquanto que as derivadas parciais da forma $\partial_0^q \partial_\alpha u$ com $q \geq r$ são calculadas recursivamente a partir dessas e usando as equações obtidas aplicando os operadores ∂_0^p com $0 \leq p \leq q - r$ à equação (1.42). Esta simples observação sugere que, no caso em que a aplicação G que define as equações diferenciais do sistema, assim como as aplicações $u^{(q)}$ que constituem as condições iniciais, forem analíticas, o problema de Cauchy acima formulado possui uma solução única, pelo menos localmente. E de fato isso é exatamente a afirmação do seguinte

Teorema 1.2 (Teorema de Cauchy-Kovalevski): *Suponha que as funções G_a e $(u^{(q)})_a$ que aparecem no problema de Cauchy constituído pelas eqs. (1.42) e (1.43) acima sejam analíticas. Então existe uma vizinhança aberta Ω_0 de $\Omega \cap H$ contida em Ω na qual o referido problema de Cauchy admite uma única solução analítica.*

Para a demonstração deste teorema, veja livros texto clássicos, tais como [Folland], Cap. 1.D, ou [John], Cap. 3.3, onde também se encontra o material necessário sobre funções analíticas (de várias variáveis reais).

Infelizmente, apesar de sua natureza geral, o teorema de Cauchy-Kovalevski é de pouca utilidade em aplicações, devido à sua natureza puramente local e, mais ainda, devido à restrição inerente ao âmbito analítico. Por exemplo, o teorema não garante existência nem unicidade de soluções globais, ou seja, geralmente temos $\Omega_0 \neq \Omega$ e não há controle sobre o grau em que Ω_0 é menor do que Ω . Também não exclui a existência de outras soluções do mesmo problema que não sejam analíticas, ou que a solução analítica admita uma extensão (ou mesmo extensões) definida(s) em um domínio maior Ω_1 , $\Omega_0 \subset \Omega_1 \subset \Omega$, que deixa(m) de ser analítica(s). E finalmente, o princípio de continuação analítica implica que soluções analíticas

são completamente determinadas, em todo o seu domínio, por condições locais em torno de um único ponto – um comportamento irreal e inaceitável, por exemplo, para soluções de equações hiperbólicas, onde tal comportamento entra em flagrante conflito com a noção de causalidade.

1.4 Equações e operadores diferenciais lineares

Quando comparamos a área de equações diferenciais parciais com a de equações diferenciais ordinárias, evidencia-se uma enorme diferença entre as duas no que diz respeito ao grau de desenvolvimento da teoria geral. Por um lado, a teoria das equações diferenciais ordinárias baseia-se em alguns teoremas bastante gerais, válidos inclusive para equações e sistemas de equações não-lineares, a saber o teorema de existência e unicidade local de soluções e o teorema sobre sua dependência em relação às condições iniciais e a parâmetros.⁸ Por outro lado, na área das equações diferenciais parciais, não se conhece até o presente nenhum resultado desta abrangência, pois cada um dos dois teoremas apresentados nas duas seções anteriores tem as suas limitações que o impedem de ocupar uma posição comparável à do teorema do fluxo para equações diferenciais ordinárias: o teorema de Frobenius se aplica apenas a sistemas fortemente sobredeterminados enquanto que o teorema de Cauchy-Kovalevski é restrito ao âmbito analítico, e a grande maioria dos problemas importantes da área de equações diferenciais parciais não se enquadra em nenhuma destas hipóteses. Assim, pode-se dizer que uma teoria geral de equações diferenciais parciais que permita tratar de equações não-lineares simplesmente não existe: resultados mais profundos costumam ser restritos a equações específicas e requerem técnicas específicas para sua demonstração.

Para equações diferenciais lineares, porém, a situação é diferente, no sentido de que teoremas gerais do tipo supracitado podem ser formulados não apenas para equações diferenciais ordinárias mas também para equações diferenciais parciais. O método mais elegante para abordar essa teoria baseia-se no conceito de um *operador diferencial linear*.

Tendo em vista esta situação, vamos nos restringir, no resto deste capítulo assim como no resto deste livro, a considerar apenas equações ou sistemas de equações diferenciais lineares. Portanto, está subentendido, uma vez por todas, que de agora em diante, a expressão “operador diferencial”, sem especificação adicional, sempre significará “operador diferencial linear”, exceto quando admitirmos de maneira explícita a possibilidade de alguma não-linearidade.

De forma resumida, é fácil especificar o que se entende pela noção de uma **equação diferencial linear**: é uma equação da forma

$$Du = f, \tag{1.44}$$

onde D é um operador diferencial. Mais especificamente, a equação (1.44) é dita **homogênea** se $f = 0$ e **não-homogênea** se $f \neq 0$; ademais, costuma-se chamar

⁸Uma apresentação resumida desse conjunto de resultados, geralmente conhecido sob o termo “teorema do fluxo”, encontra-se no Apêndice A.

a função f de **fonte** ou de **termo não-homogêneo** e a função u de **solução** da referida equação, pois em geral, f é conhecida, e procura-se u . Como veremos mais adiante, a mesma terminologia se aplica quando substituirmos funções comuns por funções generalizadas, ou distribuições: neste caso, u é chamada uma **solução fraca**.

A grande vantagem de considerar equações diferenciais como sendo o resultado de aplicar operadores diferenciais a funções para gerar outras funções é que desta forma, torna-se possível separar o aspecto de propriedades específicas das funções envolvidas – questões do tipo qual seria o seu de diferenciabilidade das soluções, por exemplo – do aspecto de propriedades genuínas das próprias equações – questões do tipo se estamos tratando de uma equação elíptica, hiperbólica ou parabólica,⁹ por exemplo – propriedades que podem ser atribuídas aos operadores diferenciais que nelas aparecem. É possível até definir um operador diferencial (e portanto, uma equação diferencial) de maneira puramente formal, sem que se precise definir explicitamente qual seria o espaço de funções (ou talvez de distribuições) às quais ele deve ser aplicado, ou qual seria o espaço de funções (ou talvez de distribuições) resultando de tal aplicação. Serão tais aspectos formais de equações diferenciais – ou operadores diferenciais – que pretendemos abordar neste capítulo.

Isto posto, o problema se reduz a dar uma definição do que deve-se entender, a nível formal, pela noção de um operador diferencial (linear). A resposta pode ser resumida numa simples frase:

Definição 1.3 *Um operador diferencial formal é um polinômio em derivadas de primeira ordem. Mais explicitamente, um operador diferencial formal ordinário é um polinômio no operador básico de diferenciação comum*

$$d = \frac{d}{dx} ,$$

enquanto que um operador diferencial formal parcial em n variáveis é um polinômio nos operadores básicos de diferenciações parciais

$$\partial_1 = \frac{\partial}{\partial x_1} , \dots , \partial_n = \frac{\partial}{\partial x_n} .$$

Para esclarecer esta definição e inclusive para mostrar que ela está em sintonia com a definição dada na primeira seção deste capítulo, precisamos discutir duas questões:

- Qual é a forma mais adequada (para não dizer, canônica) para representar polinômios, principalmente em dimensão $n > 1$?
- Qual é a natureza dos coeficientes destes polinômios?

Quanto à primeira pergunta, existe uma resposta padrão que, no caso $n = 1$, é muito bem conhecida: Em dimensão $n = 1$, um polinômio é simplesmente uma

⁹Estas noções serão definidas logo em seguida.

função P sobre \mathbb{R} , ou seja, uma função P de uma variável real ξ , que pode ser escrita na forma de uma combinação linear dos monômios ξ^k , com coeficientes a_k :

$$P(\xi) = \sum_{k=0}^r a_k \xi^k . \quad (1.45)$$

Portanto, um operador diferencial ordinário formal é uma combinação linear dos operadores de diferenciação d^k/dx^k , ou seja, um operador da forma

$$P(d) = \sum_{k=0}^r a_k \frac{d^k}{dx^k} . \quad (1.46)$$

Dizemos que o polinômio $P(\xi)$ assim como o operador $P(d)$ têm grau r se $a_r \neq 0$. Em dimensão $n \geq 1$, é conveniente utilizar o conceito de multi-índice (veja a Definição 1.1). Com esta notação, um **polinômio** é simplesmente uma função P sobre \mathbb{R}^n , ou seja, uma função P de n variáveis reais ξ_1, \dots, ξ_n , que pode ser escrita na forma de uma combinação linear dos monômios

$$\xi^\alpha = \prod_{i=1}^n \xi_i^{\alpha_i} , \quad (1.47)$$

com coeficientes a_α :

$$P(\xi) = \sum_{|\alpha| \leq r} a_\alpha \xi^\alpha . \quad (1.48)$$

Portanto, um **operador diferencial formal** parcial é uma combinação linear dos operadores de diferenciação parcial já definidos na equação (1.12),

$$\partial_\alpha = \prod_{i=1}^n \partial_i^{\alpha_i} , \quad (1.49)$$

ou seja, um operador da forma

$$P(\partial) = \sum_{|\alpha| \leq r} a_\alpha \partial_\alpha . \quad (1.50)$$

Dizemos que o polinômio $P(\xi)$ assim como o operador $P(\partial)$ têm **grau** r se existe (pelo menos) um multi-índice α de grau r tal que $a_\alpha \neq 0$.

Para entender melhor a notação dos multi-índices, escrevemos explicitamente os monômios e operadores de diferenciação parcial em n variáveis de grau ≤ 2 :

$ \alpha $	α	ξ^α	∂_α
0	$(0, \dots, 0)$	1	1
1	$(0, \dots, 1, \dots, 0)$ 1 na i -ésima posição	ξ_i	$\frac{\partial}{\partial x_i}$
2	$(0, \dots, 2, \dots, 0)$ 2 na i -ésima posição	ξ_i^2	$\frac{\partial^2}{\partial x_i^2}$
2	$(0, \dots, 1, \dots, 1, \dots, 0)$ 1 na i -ésima posição e na j -ésima posição	$\xi_i \xi_j$	$\frac{\partial^2}{\partial x_i \partial x_j}$

Quanto à natureza dos coeficientes a_k nas equações (1.45), (1.46) e a_α nas equações (1.48), (1.50), existem várias opções. A primeira e mais elementar é que são simplesmente números (reais ou complexos) ou, no caso de um operador representando um sistema de k equações para N funções, matrizes ($k \times N$) (reais ou complexas). Neste caso, falamos de um operador diferencial (ou uma equação diferencial ou um sistema de equações diferenciais) *a coeficientes constantes*. A segunda, mais geral, é que são funções (a valores reais ou complexas) ou, no caso de um operador representando um sistema de k equações para N funções, matrizes ($k \times N$) de funções (a valores reais ou complexas), de uma determinada classe, definidas sobre um domínio Ω de \mathbb{R} ou de \mathbb{R}^n : a hipótese padrão é que sejam funções de classe C^∞ sobre Ω . Neste caso, falamos de um operador diferencial (ou uma equação diferencial ou um sistema de equações diferenciais) *sobre Ω , ou em Ω* . Observe que a primeira opção pode ser considerada um caso especial da segunda, já que números podem ser identificados, de maneira natural, com funções constantes, o que justifica a expressão “a coeficientes constantes”.

Para sistematizar o relacionamento entre um operador diferencial parcial e o polinômio correspondente, mesmo no caso de operadores cujos coeficientes são funções (possivelmente não-constantes), introduzimos a seguinte terminologia.

Definição 1.4 *Seja D um operador diferencial parcial de grau r (possivelmente representando um sistema) sobre um aberto Ω de \mathbb{R}^n ,*

$$D = \sum_{|\alpha| \leq r} a_\alpha \partial_\alpha, \quad (1.51)$$

*com coeficientes $a_\alpha \in C^\infty(\Omega, L(\mathbb{R}^N, \mathbb{R}^k))$. Então o **símbolo** de D é a função $\sigma_D \in C^\infty(\Omega \times \mathbb{R}^n, L(\mathbb{R}^N, \mathbb{R}^k))$ definida por*

$$\sigma_D(x, \xi) = \sum_{|\alpha| \leq r} a_\alpha(x) \xi^\alpha, \quad (1.52)$$

e o **símbolo principal** de D é a função $(\sigma_D)_p \in C^\infty(\Omega \times \mathbb{R}^n, L(\mathbb{R}^N, \mathbb{R}^k))$ definida por

$$(\sigma_D)_p(x, \xi) = \sum_{|\alpha|=r} a_\alpha(x) \xi^\alpha . \quad (1.53)$$

Enfatizamos que o símbolo e o símbolo principal são funções não de n mas de $2n$ variáveis, sendo que σ_D é um polinômio e $(\sigma_D)_p$ um polinômio homogêneo de grau r em ξ , onde r é o grau de D , ambos com coeficientes que são funções, ou matrizes de funções, de classe C^∞ em x . Também enfatizamos que é importante distinguir entre as variáveis x e ξ : longe de poderem ser identificadas, elas tem um papel dual, ou complementar, uma em relação a outra. Isso só ficará mais claro em função de desenvolvimentos posteriores (alguns fora do contexto deste livro), por exemplo no âmbito da teoria da transformação de Fourier, onde algum produto escalar entre x e ξ aparece como argumento da exponencial (veja a Seção 3.13), mas também no âmbito da mecânica quântica, onde variáveis “tipo posição” (os x) e variáveis “tipo momento” (os ξ) são complementares, no sentido de não poderem ser medidas simultaneamente com precisão arbitrária, conforme a relação de incerteza de Heisenberg.

Uma das múltiplas utilidades do símbolo principal é que permite caracterizar covetores e hiperplanos característicos.

Definição 1.5 Para um operador diferencial parcial D (possivelmente representando um sistema, com $k = N$) sobre um aberto Ω de \mathbb{R}^n , um covetor ξ é dito **característico** no ponto x_0 de Ω se o símbolo principal de D for singular no ponto (x_0, ξ) , i.e., se

$$\det((\sigma_D)_p(x_0, \xi)) = 0 , \quad (1.54)$$

e um hiperplano H de \mathbb{R}^n é dito **característico** no ponto x_0 de Ω se é o núcleo de algum covetor característico neste ponto.

A inexistência de hiperplanos característicos é típica de uma importante classe de operadores diferenciais.

Definição 1.6 Um operador diferencial parcial D (possivelmente representando um sistema, com $k = N$) sobre um aberto Ω de \mathbb{R}^n é dito **elíptico** se o seu símbolo principal for singular apenas na origem,

$$\det((\sigma_D)_p(x, \xi)) \neq 0 \quad \text{para } x \in \Omega, \xi \neq 0 . \quad (1.55)$$

Obviamente, um operador diferencial parcial D é completamente determinado por seu símbolo, e não é difícil mostrar que as definições de equação diferencial linear dadas aqui e na primeira seção deste capítulo coincidem. Também pode ser verificado sem dificuldade que as definições de covetor e hiperplano característico dadas aqui e na seção anterior deste capítulo (veja a Definição 1.2) coincidem. Deixaremos os detalhes da demonstração destas afirmações para o leitor como exercício.

Encerrando esta seção, notamos uma regra importante para calcular derivadas parciais de qualquer ordem de um produto de funções; esta regra desempenha um papel importante na manipulação de operadores diferenciais com coeficientes que não são constantes e, mais geralmente, para estabelecer estimativas em análise. Aqui, ela serve para ilustrar a grande utilidade da notação dos multi-índices.

Exercício 1.1 Usando a regra de Leibniz para o operador d/dx sobre $\Omega \subset \mathbb{R}$, mostre por indução sobre o grau k que vale a seguinte regra do produto:

$$\frac{d^k}{dx^k}(fg) = \sum_{l=0}^k \binom{k}{l} \frac{d^l f}{dx^l} \frac{d^{k-l} g}{dx^{k-l}}, \quad (1.56)$$

onde por definição,

$$\binom{k}{l} = \frac{k!}{l!(k-l)!}.$$

De modo análogo, usando a regra de Leibniz para os operadores $\partial/\partial x_i$ sobre $\Omega \subset \mathbb{R}^n$, mostre por indução sobre o grau k que vale a seguinte regra do produto: Para todo multi-índice α de grau $|\alpha| \leq k$,

$$\partial_\alpha(fg) = \sum_{\beta \leq \alpha} \binom{\alpha}{\beta} \partial_\beta f \partial_{\alpha-\beta} g, \quad (1.57)$$

onde por definição,

$$\binom{\alpha}{\beta} = \frac{\alpha!}{\beta!(\alpha-\beta)!},$$

e

$$\beta \leq \alpha \iff \beta_i \leq \alpha_i \quad \text{para } 1 \leq i \leq n,$$

$$\alpha! = \prod_{i=1}^n \alpha_i!,$$

$$\binom{\alpha}{\beta} = \prod_{i=1}^n \binom{\alpha_i}{\beta_i} = \prod_{i=1}^n \frac{\alpha_i!}{\beta_i!(\alpha_i - \beta_i)!}.$$

Sugestão: Use a identidade binomial

$$\binom{k}{l} + \binom{k}{l-1} = \binom{k+1}{l}.$$

1.5 Operadores diferenciais parciais de grau 2

Os operadores diferenciais parciais (a coeficientes constantes) de grau ≤ 2 podem ser completamente classificados. A sua classificação baseia-se na classificação das formas quadráticas em \mathbb{R}^n .

Geralmente, um polinômio de grau ≤ 2 pode ser escrito na forma

$$P(\xi) = P_2(\xi) + P_1(\xi) + P_0(\xi) = \sum_{j,k=1}^n a_{jk} \xi_j \xi_k + \sum_{j=1}^n a_j \xi_j + a_0, \quad (1.58)$$

e portanto, temos

$$P(\partial) = P_2(\partial) + P_1(\partial) + P_0(\partial) = \sum_{j,k=1}^n a_{jk} \frac{\partial}{\partial x_j} \frac{\partial}{\partial x_k} + \sum_{j=1}^n a_j \frac{\partial}{\partial x_j} + a_0, \quad (1.59)$$

onde os coeficientes a_{jk} do termo P_2 de segunda ordem formam uma matriz simétrica: $a_{jk} = a_{kj}$. (Se inicialmente esta matriz não for simétrica, podemos simplesmente substituí-la pela sua parte simétrica. Também vale mencionar que não utilizamos aqui a notação em termos de multi-índices, que é muito útil para polinômios e/ou operadores diferenciais de grau arbitrário, mas não apresenta nenhuma vantagem específica no tratamento de polinômios e/ou operadores diferenciais de grau ≤ 2 .)

Através de uma mudança apropriada de base em \mathbb{R}^n , podemos transformar o polinômio quadrático P_2 e, portanto, o polinômio P e/ou o operador diferencial $P(\partial)$, numa forma canônica: De fato, a matriz dos coeficientes a_{jk} de P_2 , sendo simétrica, pode ser diagonalizada por uma transformação linear ortogonal em \mathbb{R}^n . Posteriormente, podemos aplicar uma transformação linear reescalando cada coordenada por um fator apropriado, transformando a matriz diagonal numa matriz diagonal com elementos diagonais iguais a $+1$, 0 ou -1 . Finalmente, podemos também mudar, de maneira arbitrária, a ordem dos elementos diagonais, pois podemos efetuar uma permutação qualquer das coordenadas ξ_1, \dots, ξ_n através de uma transformação linear apropriada. (Por exemplo, a transposição das coordenadas ξ_j e ξ_k pode ser realizada por uma matriz $T(j, k)$ com os elementos $T(j, k)_{jk} = 1$, $T(j, k)_{kj} = 1$, $T(j, k)_{ll} = 1$ se $l \neq j$ e $l \neq k$, e com todos os outros elementos iguais a zero.) Isto significa que as formas quadráticas em \mathbb{R}^n podem ser completamente classificadas por dois números inteiros não-negativos p e q , a saber

$$\begin{aligned} p &= \text{número de autovalores positivos da matriz } (a_{jk}), \\ q &= \text{número de autovalores negativos da matriz } (a_{jk}), \\ r &= \text{multiplicidade do autovalor } 0 \text{ da matriz } (a_{jk}), \end{aligned} \quad (1.60)$$

onde $p+q+r = n$; a diferença $p-q$ é chamada a **assinatura** da forma quadrática,¹⁰ enquanto que r é a dimensão do seu núcleo

$$N = \left\{ \xi \in \mathbb{R}^n \mid \sum_{j,k=1}^n a_{jk} \xi_j \xi_k = 0 \text{ para todo } \eta \in \mathbb{R}^n \right\}. \quad (1.61)$$

¹⁰A afirmação de que essas são as únicas invariantes das formas quadráticas sobre \mathbb{R}^n sob transformações lineares (mudanças de base) arbitrárias, é conhecida na álgebra linear como o *teorema de inércia de Sylvester*.

Implica também que a forma padrão do polinômio P e do operador diferencial $P(\partial)$ é a seguinte:

$$P(\xi) = \left(\sum_{j=1}^p \xi_j^2 - \sum_{j=n-q+1}^n \xi_j^2 \right) + \sum_{j=1}^n a_j \xi_j + a_0, \quad (1.62)$$

$$P(\partial) = \left(\sum_{j=1}^p \frac{\partial^2}{\partial x_j^2} - \sum_{j=n-q+1}^n \frac{\partial^2}{\partial x_j^2} \right) + \sum_{j=1}^n a_j \frac{\partial}{\partial x_j} + a_0. \quad (1.63)$$

Obviamente, se escrevermos $D = P(\partial)$, então $\sigma_D(\xi) = P(\xi)$ é o símbolo de D e $(\sigma_D)_p(\xi) = P_2(\xi)$ é o símbolo principal de D . Denotando por c uma constante arbitrária, podemos considerar as equações

$$\sigma_D(\xi) = c, \quad (1.64)$$

e

$$(\sigma_D)_p(\xi) = c, \quad (1.65)$$

sendo que a última é chamada a **equação característica** do operador D .

Definição 1.7 *Um operador diferencial parcial D de grau 2 a coeficientes constantes é dito*

- a) **elíptico** se o seu símbolo principal for uma forma quadrática definida (positiva ou negativa), isto é, se $p = n$, $r = 0$, $q = 0$ ou $p = 0$, $r = 0$, $q = n$;
- b) **hiperbólico** se o seu símbolo principal for uma forma quadrática não-degenerada de assinatura $\pm(n-2)$, isto é, se $p = n-1$, $r = 0$, $q = 1$ ou $p = 1$, $r = 0$, $q = n-1$;
- c) **parabólico** se o seu símbolo principal for uma forma quadrática degenerada com núcleo N uni-dimensional e definida (positiva ou negativa) sobre qualquer complemento direto de N , isto é, se $p = n-1$, $r = 1$, $q = 0$ ou $p = 0$, $r = 1$, $q = n-1$, e se o vetor a formado pelos coeficientes dos termos de primeira ordem não pertencer ao hiperplano de \mathbb{R}^n aniquilado pelos covetores em N , ou seja,

$$\xi \in N \setminus \{0\} \implies \sum_{j=1}^n \xi_j a_j \neq 0. \quad (1.66)$$

Nota-se que essa definição de um operador elíptico é um caso especial da anterior, sendo que a Definição 1.6 é muito mais geral, pois também se aplica a operadores com coeficientes que não são constantes, que têm grau arbitrário e que podem representar sistemas. Infelizmente, não existe até o presente nenhuma definição tão abrangente de que seria um operador hiperbólico ou parabólico. Nota-se também a condição técnica adicional (1.66) imposta na definição de um operador parabólico, que não pode ser formulada exclusivamente em termos do símbolo principal e que

serve para garantir que o operador em questão não seja meramente uma família a um parâmetro de operadores elípticos em $n-1$ dimensões.

Para explicar a origem dos termos “elíptico”, “hiperbólico” e “parabólico”, observamos que a equação (1.64) define uma família de hiper-superfícies em \mathbb{R}^n (parametrizadas por c), cuja forma providencia a motivação desejada:¹¹

Exemplo 1.1 Em $n=2$ dimensões, temos as seguintes alternativas:

- a) $p = 2, r = 0, q = 0$ ou $p = 0, r = 0, q = 2$: Usando a equação (1.62) e absorvendo várias constantes numa redefinição da constante c , a equação (1.64) assume a forma

$$(\xi_1 + \frac{1}{2}a_1)^2 + (\xi_2 + \frac{1}{2}a_2)^2 = c \quad (c > 0),$$

que é a equação de uma elipse (no caso, um círculo, devido à normalização das coordenadas ξ_1 e ξ_2).

- b) $p = 1, r = 0, q = 1$: Usando a equação (1.62) e absorvendo várias constantes numa redefinição da constante c , a equação (1.64) assume a forma

$$(\xi_1 + \frac{1}{2}a_1)^2 - (\xi_2 - \frac{1}{2}a_2)^2 = c \quad (c \in \mathbb{R}),$$

que é a equação de uma hipérbole (no caso, uma hipérbole equilátera, devido à normalização das coordenadas ξ_1 e ξ_2), que para $c=0$ degenera nas duas retas diagonais assíntotas.¹²

- c) $p = 1, r = 1, q = 0$ ou $p = 0, r = 1, q = 1$: Usando a equação (1.62) e absorvendo várias constantes numa redefinição da constante c , a equação (1.64) assume a forma

$$(\xi_1 + \frac{1}{2}a_1)^2 + a_2\xi_2 = c \quad (c \in \mathbb{R}),$$

com $a_2 \neq 0$, ou

$$a_1\xi_1 - (\xi_2 - \frac{1}{2}a_2)^2 = c \quad (c \in \mathbb{R}),$$

com $a_1 \neq 0$, que é a equação de uma parábola.

1.6 Operadores diferenciais parciais da física

Nesta seção, apresentamos uma lista dos operadores diferenciais parciais mais importantes que aparecem na física. Todos eles são operadores a coeficientes constantes de grau ≤ 2 : é um fato que a natureza parece ter uma forte preferência por equações diferenciais de ordem ≤ 2 , mas ninguém sabe o porquê.

¹¹Para podermos incluir o caso parabólico, estamos obrigados a considerar a equação (1.64), ao invés da equação característica (1.65).

¹²O sinal de c adquire importância em dimensões maiores do que 2, por exemplo já em dimensão $n=3$, onde a equação (1.65) descreve um hiperbolóide de uma folha quando $p = n-1, q = 1$ e $c > 0$ ou $p = 1, q = n-1$ e $c < 0$, um hiperbolóide de duas folhas quando $p = n-1, q = 1$ e $c < 0$ ou $p = 1, q = n-1$ e $c > 0$, e um cone duplo quando $c = 0$.

1.6.1 Operadores de grau 2

1. Operador de Laplace (ou laplaciano)

O operador de Laplace, ou laplaciano, em \mathbb{R}^n é definido por

$$\Delta = \sum_{j=1}^n \frac{\partial^2}{\partial x_j^2}. \quad (1.67)$$

Portanto, o seu símbolo, igual ao seu símbolo principal, é a forma quadrática definida por

$$\sigma_{\Delta}(\xi) = \sum_{j=1}^n \xi_j^2. \quad (1.68)$$

(Escrevendo o operador Δ na forma $P(\partial)$ e utilizando a representação (1.48) de P , precisamos tomar

$$a_{\alpha} = 1 \quad \text{se } \alpha = (2, 0, \dots, 0) \text{ ou } \dots \text{ ou } (0, 0, \dots, 2),$$

e $a_{\alpha} = 0$ para todos os outros valores de α .)

O operador de Laplace é o exemplo padrão de operador elíptico.

Definição 1.8 *A equação homogênea associada ao operador de Laplace se chama **equação de Laplace** e a equação não-homogênea **equação de Poisson**. As soluções u da equação de Laplace $\Delta u = 0$ se chamam **funções harmônicas**.*

Por muitos autores, o termo “função harmônica” é estendido às soluções u da equação de Laplace modificada $\Delta u = \lambda u$, com $\lambda \in \mathbb{R}$, ou seja, a todas as autofunções do operador de Laplace (com autovalor $\lambda \in \mathbb{R}$ arbitrário), e nós seguiremos esta convenção aqui.

2. Operador de difusão (ou transporte)

O operador de difusão, ou transporte, em \mathbb{R}^{n+1} é definido por

$$\frac{\partial}{\partial t} - \lambda \sum_{j=1}^n \frac{\partial^2}{\partial x_j^2}, \quad (1.69)$$

onde λ é uma constante positiva – a constante de difusão, que descreve a taxa de difusão ou transporte dos fenômenos descritos por esta equação. (Na termodinâmica, por exemplo, pode ser a condutibilidade térmica, na química, a solubilidade de uma substância em um solvente, etc.) Portanto, o seu símbolo é

$$\sigma(\omega, \xi) = \omega - \lambda \sum_{j=1}^n \xi_j^2, \quad (1.70)$$

enquanto que o seu símbolo principal é a forma quadrática definida por

$$\sigma(\omega, \xi) = -\lambda \sum_{j=1}^n \xi_j^2. \quad (1.71)$$

(Escrevendo o operador em questão na forma $P(\partial)$ e modificando a equação (1.48) para escrever o polinômio P como função de uma variável ω (substituindo $\partial/\partial t$) e n variáveis ξ_j (substituindo $\partial/\partial x_j$), conforme

$$P(\omega, \xi) = \sum_{|\alpha| \leq 2} a_\alpha(\omega, \xi)^\alpha = \sum_{|\alpha| \leq 2} a_\alpha \omega^{\alpha_0} \xi_1^{\alpha_1} \dots \xi_n^{\alpha_n}$$

com multi-índices $\alpha = (\alpha_0, \alpha_1, \dots, \alpha_n) \in \mathbb{N}^{n+1}$, precisamos tomar

$$\begin{aligned} a_\alpha &= 1 \quad \text{se } \alpha = (1, 0, \dots, 0), \\ a_\alpha &= -\lambda \quad \text{se } \alpha = (0, 2, \dots, 0) \text{ ou } \dots \text{ ou } (0, 0, \dots, 2), \end{aligned}$$

e $a_\alpha = 0$ para todos os outros valores de α .)

O operador de difusão é o exemplo padrão de operador parabólico.

3. Operador de ondas (ou d'alembertiano)

O operador de ondas, ou d'alembertiano, em \mathbb{R}^{n+1} é definido por

$$\square = \frac{1}{c^2} \frac{\partial^2}{\partial t^2} - \sum_{j=1}^n \frac{\partial^2}{\partial x_j^2}, \quad (1.72)$$

onde c é uma constante positiva com a dimensão de uma velocidade – a velocidade de propagação dos fenômenos (ondas, por exemplo) descritos por esta equação. (Na acústica, será a velocidade do som, no eletromagnetismo, a velocidade da luz, etc..) Portanto, o seu símbolo, igual ao seu símbolo principal, é a forma quadrática definida por

$$\sigma_\square(\omega, \xi) = \frac{\omega^2}{c^2} - \sum_{j=1}^n \xi_j^2. \quad (1.73)$$

(Escrevendo o operador \square na forma $P(\partial)$ e modificando a equação (1.48) para escrever o polinômio P como função de uma variável ω (substituindo $\partial/\partial t$) e n variáveis ξ_j (substituindo $\partial/\partial x_j$), conforme

$$P(\omega, \xi) = \sum_{|\alpha| \leq 2} a_\alpha(\omega, \xi)^\alpha = \sum_{|\alpha| \leq 2} a_\alpha \omega^{\alpha_0} \xi_1^{\alpha_1} \dots \xi_n^{\alpha_n}$$

com multi-índices $\alpha = (\alpha_0, \alpha_1, \dots, \alpha_n) \in \mathbb{N}^{n+1}$, precisamos tomar

$$\begin{aligned} a_\alpha &= 1/c^2 \quad \text{se } \alpha = (2, 0, \dots, 0), \\ a_\alpha &= -1 \quad \text{se } \alpha = (0, 2, \dots, 0) \text{ ou } \dots \text{ ou } (0, 0, \dots, 2), \end{aligned}$$

e $a_\alpha = 0$ para todos os outros valores de α .)

O operador de ondas é o exemplo padrão de operador hiperbólico.

4. Operador de Schrödinger

O operador de Schrödinger em \mathbb{R}^{n+1} é definido por

$$i \frac{\partial}{\partial t} - \lambda \sum_{j=1}^n \frac{\partial^2}{\partial x_j^2}, \quad (1.74)$$

onde λ é uma constante real. Formalmente, o operador de Schrödinger é uma variante do operador de difusão, obtida pela aparentemente inócua substituição do fator 1 pelo fator i . A teoria matemática da equação de Schrödinger e da equação de difusão, porém, assim como as propriedades das suas soluções e a sua interpretação física, apresentam diferenças drásticas: as analogias se restringem a aspectos puramente formais.

5. Operador de Helmholtz ou Klein-Gordon

O operador de Helmholtz ou operador de Klein-Gordon em \mathbb{R}^{n+1} é definido por

$$\square - \frac{1}{l^2}, \quad (1.75)$$

onde \square é o operador de ondas e l é uma constante positiva com a dimensão de um comprimento – a distância típica de alcance dos fenômenos (forças nucleares, por exemplo) descritos por esta equação.

1.6.2 Operadores de grau 1

6. Operadores de Cauchy-Riemann

No plano complexo $\mathbb{C} = \mathbb{R}^2$, introduzimos coordenadas complexas

$$z = x + iy, \quad \bar{z} = x - iy,$$

e consideramos os operadores $\partial/\partial z$ e $\partial/\partial \bar{z}$ de Cauchy-Riemann, definidos por

$$\frac{\partial}{\partial z} = \frac{1}{2} \left(\frac{\partial}{\partial x} - i \frac{\partial}{\partial y} \right), \quad \frac{\partial}{\partial \bar{z}} = \frac{1}{2} \left(\frac{\partial}{\partial x} + i \frac{\partial}{\partial y} \right). \quad (1.76)$$

Os símbolos destes operadores são

$$\sigma_{\partial/\partial z}(\zeta, \bar{\zeta}) = \bar{\zeta}/2, \quad \sigma_{\partial/\partial \bar{z}}(\zeta, \bar{\zeta}) = \zeta/2. \quad (1.77)$$

Estes dois operadores podem ser combinados, de maneira natural, para formar um operador diferencial matricial, isto é, um sistema de operadores diferenciais:

$$\partial = 2 \begin{pmatrix} 0 & \partial/\partial z \\ \partial/\partial \bar{z} & 0 \end{pmatrix}. \quad (1.78)$$

Esse é um exemplo típico de um sistema elíptico de operadores diferenciais. Também é interessante observar que

$$\partial^2 = \Delta,$$

(onde seguimos a convenção usual de suprimir a matriz 1), ou explicitamente,

$$\Delta = 4 \frac{\partial}{\partial z} \frac{\partial}{\partial \bar{z}} = 4 \frac{\partial}{\partial \bar{z}} \frac{\partial}{\partial z} ,$$

isto é, o operador (1.78) é uma “raiz” do operador de Laplace.¹³ Finalmente, notamos que as soluções da equação homogênea $\partial u/\partial \bar{z} = 0$ respectivamente $\partial v/\partial z = 0$ têm um papel fundamental na teoria das funções de uma variável complexa: são as funções analíticas (ou holomorfas) respectivamente anti-analíticas (ou anti-holomorfas).

7. Operadores associados ao cone de luz

No espaço de Minkowski bidimensional \mathbb{R}^2 , introduzimos as tão-chamadas coordenadas do cone de luz ou coordenadas características

$$x^\pm = ct \pm x ,$$

e consideramos os operadores ∂_+ e ∂_- associados ao cone de luz, definidos por

$$\partial_\pm \equiv \frac{\partial}{\partial x^\pm} = \frac{1}{2} \left(\frac{1}{c} \frac{\partial}{\partial t} \pm \frac{\partial}{\partial x} \right) , \quad (1.79)$$

Os símbolos destes operadores são

$$\sigma_{\partial_\pm}(\xi^+, \xi^-) = \xi^\pm/2 = \frac{1}{2} \left(\frac{\omega}{c} \pm \xi \right) . \quad (1.80)$$

Como no caso anterior, estes dois operadores podem ser combinados, de maneira natural, para formar um operador diferencial matricial, isto é, um sistema de operadores diferenciais:

$$\not\partial = 2 \begin{pmatrix} 0 & \partial_+ \\ \partial_- & 0 \end{pmatrix} , \quad (1.81)$$

Apesar de não termos definido aqui a noção de um sistema hiperbólico de operadores diferenciais, mencionamos que esse seria um exemplo típico de um tal sistema. Também é interessante observar que

$$\not\partial^2 = \square ,$$

(onde seguimos a convenção usual de suprimir a matriz 1), ou explicitamente,

$$\square = 4 \partial_+ \partial_- = 4 \partial_- \partial_+ ,$$

isto é, o operador (1.81) é uma “raiz” do operador de ondas.¹³ Esta fatorização fornece a base para a solução da equação de ondas em uma dimensão espacial, que será discutida no próximo capítulo.

¹³O preço a ser pago para poder definir a “raiz” do operador de Laplace ou de ondas é justamente que tal raiz deixa de ser um operador escalar: uma equação de segunda ordem corresponde a um sistema de equações de primeira ordem.

8. Operador de Dirac

Generalizando as construções dos dois itens anteriores de duas para n variáveis, podemos definir o operador de Dirac em \mathbb{R}^n como um operador diferencial de grau 1 cujos coeficientes $\gamma_1, \dots, \gamma_n$ são matrizes:

$$\not\partial = \sum_{j=1}^n \gamma_j \frac{\partial}{\partial x_j}. \quad (1.82)$$

O símbolo de um tal operador é

$$\sigma_{\not\partial}(\xi) = \sum_{j=1}^n \gamma_j \xi_j. \quad (1.83)$$

Calculamos a diferença entre o quadrado do operador $\not\partial$ e o operador

$$\mathcal{D} = \sum_{j,k=1}^n \eta_{jk} \frac{\partial}{\partial x_j} \frac{\partial}{\partial x_k}, \quad (1.84)$$

onde η é a matriz diagonal $\text{diag}(+1, \dots, +1, -1, \dots, -1)$ (com $+1$ aparecendo p vezes e -1 aparecendo q vezes, $p + q = n$), ou seja

$$\mathcal{D} = \sum_{j=1}^p \frac{\partial^2}{\partial x_j^2} - \sum_{j=p+1}^{p+q} \frac{\partial^2}{\partial x_j^2}, \quad (1.85)$$

o que inclui o operador de Laplace ($p = n, q = 0$ ou $p = 0, q = n$) assim como o operador de ondas ($p = n - 1, q = 1$ ou $p = 1, q = n - 1$):

$$\begin{aligned} \not\partial^2 - \mathcal{D} &= \sum_{j,k=1}^n \gamma_j \gamma_k \frac{\partial}{\partial x_j} \frac{\partial}{\partial x_k} - \sum_{j,k=1}^n \eta_{jk} \frac{\partial}{\partial x_j} \frac{\partial}{\partial x_k} \\ &= \sum_{j,k=1}^n \left(\frac{1}{2} (\gamma_j \gamma_k + \gamma_k \gamma_j) - \eta_{jk} \right) \frac{\partial}{\partial x_j} \frac{\partial}{\partial x_k}. \end{aligned}$$

Portanto, para termos

$$\not\partial^2 = \mathcal{D}, \quad (1.86)$$

as matrizes $\gamma_1, \dots, \gamma_n$ devem satisfazer a relação

$$\gamma_j \gamma_k + \gamma_k \gamma_j = 2 \eta_{jk}. \quad (1.87)$$

É possível mostrar que $\gamma_1, \dots, \gamma_n$, quando realizadas de maneira irredutível, devem ser matrizes $(2^{n/2} \times 2^{n/2})$ quando n é par, e $(2^{(n-1)/2} \times 2^{(n-1)/2})$ quando n é ímpar. A álgebra gerada por estas matrizes se chama *álgebra de Clifford* sobre \mathbb{R}^n , relativa à forma quadrática η .

Para $n=2$ e $n=3$, é fácil dar uma representação explícita desta construção, utilizando as matrizes de Pauli

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad (1.88)$$

que satisfazem as relações

$$\sigma_1^2 = 1, \quad \sigma_2^2 = 1, \quad \sigma_3^2 = 1, \quad (1.89)$$

$$\begin{aligned} \sigma_1\sigma_2 &= i\sigma_3 = -\sigma_2\sigma_1, \\ \sigma_2\sigma_3 &= i\sigma_1 = -\sigma_3\sigma_2, \end{aligned} \quad (1.90)$$

$$\sigma_3\sigma_1 = i\sigma_2 = -\sigma_1\sigma_3.$$

Em particular, para $n=2$, podemos por $\gamma_1 = \sigma_1$ e $\gamma_2 = \sigma_2$ ou $\gamma_1 = \sigma_1$ e $\gamma_2 = i\sigma_2$, para concluir que o operador de Dirac (1.82) se torna idêntico, respectivamente, ao operador (1.78) ou ao operador (1.81).

Os operadores de Schrödinger, de Helmholtz ou Klein-Gordon e de Dirac ocupam uma posição central na física quântica, o que constitui um assunto que não vamos e nem podemos discutir aqui. Por outro lado, cada um dos três primeiros operadores tem um papel fundamental na física clássica, onde aparecem em diversas áreas:

- A equação de ondas

$$\square u = \left(\frac{1}{c^2} \frac{\partial^2}{\partial t^2} - \Delta \right) u = f \quad (1.91)$$

descreve *processos oscilatórios* ou *ondulatórios*. Na mecânica dos meios contínuos, por exemplo, a sua versão unidimensional

$$\frac{1}{c^2} \frac{\partial^2 u}{\partial t^2} - \frac{\partial^2 u}{\partial x^2} = f, \quad (1.92)$$

governa as vibrações de uma corda e a sua versão bidimensional

$$\frac{1}{c^2} \frac{\partial^2 u}{\partial t^2} - \frac{\partial^2 u}{\partial x^2} - \frac{\partial^2 u}{\partial y^2} = f, \quad (1.93)$$

as vibrações de uma membrana, sendo u a elongação transversal da corda ou membrana, enquanto que f representa uma (possível) força externa. Na sua versão tridimensional

$$\frac{1}{c^2} \frac{\partial^2 u}{\partial t^2} - \frac{\partial^2 u}{\partial x^2} - \frac{\partial^2 u}{\partial y^2} - \frac{\partial^2 u}{\partial z^2} = f, \quad (1.94)$$

ela descreve a propagação de ondas acústicas, sendo u a densidade de massa ou a pressão no meio de propagação (mais exatamente, o desvio desta quantidade em relação ao seu valor médio), enquanto que, novamente, f representa uma (possível) força externa. Da mesma forma, a equação (1.94) rege a propagação de ondas eletromagnéticas no espaço, sendo agora u o potencial escalar ou uma componente do potencial vetorial (no calibre de Lorentz), enquanto que f representa a densidade de carga ou uma componente da densidade de corrente.

- A equação de difusão

$$\left(\frac{\partial}{\partial t} - \lambda \Delta\right) u = f \quad (1.95)$$

descreve *processos dissipativos* ou *compensativos* que tendem para algum equilíbrio quando $t \rightarrow \infty$. Ela aparece naturalmente na hidrodinâmica, principalmente na sua interface com a termodinâmica.

- Finalmente, a equação de Laplace ou Poisson

$$\Delta u = f \quad (1.96)$$

descreve *situações de equilíbrio* já estabelecido, normalmente como resultado de algum processo dissipativo já concluído. Por exemplo, a equação (1.96) determina qual é o campo eletrostático gerado por uma distribuição estática de cargas no espaço, sendo u o potencial escalar, enquanto que f representa a densidade de carga.

Para maiores detalhes, veja o Apêndice C, onde discutimos o contexto em que as três equações diferenciais parciais clássicas surgem na física – mais exatamente, na teoria dos campos, que constitui o arcabouço formal para várias áreas da física teórica, entre elas o eletromagnetismo e a dinâmica dos fluidos. A título de exemplo, mostramos como a equação de Laplace ou Poisson e a equação de ondas podem ser deduzidas a partir das leis fundamentais da eletrodinâmica, isto é, a partir das equações de Maxwell. Também discutimos, no contexto da hidrodinâmica, um dos conceitos centrais da teoria dos campos: a noção de equações de balanço e, em particular, de leis de conservação, e mostramos como estas levam, através de uma aproximação simples, à equação de difusão.

Observamos que a equação de ondas e a equação de difusão são exemplos de uma **equação de evolução**, enquanto que a equação de Laplace ou Poisson é uma **equação estática**: não há nela nenhum aspecto de evolução temporal. Essa distinção pode ser generalizada no sentido de que equações hiperbólicas ou parabólicas são equações de evolução, enquanto que equações elípticas são estáticas.

A diferença essencial entre a equação de ondas e a equação de difusão encontra-se na ordem da derivada temporal. A equação de ondas contém uma segunda derivada em relação ao tempo, e portanto, é invariante sob a inversão temporal, isto é, sob a transformação $t \rightarrow -t$. De fato, se u é uma solução da equação $\square u = f$, podemos por $v(t, \mathbf{x}) = u(-t, \mathbf{x})$ e $g(t, \mathbf{x}) = f(-t, \mathbf{x})$ para concluir que v é uma solução da equação $\square v = g$. O mesmo já não ocorre no caso da equação de difusão. Portanto, a primeira equação descreve processos *reversíveis* e a segunda, processos *irreversíveis*.

1.7 Condições iniciais e condições de fronteira

É um fenômeno geral na teoria das equações diferenciais que além da equação propriamente dita, precisa-se especificar condições suplementares para garantir

que o problema admita uma solução única. Isso já ocorre no caso das equações diferenciais ordinárias, cujas soluções são unicamente determinadas pelas condições iniciais impostas. Para a grande maioria das equações diferenciais parciais – mais exatamente para todas que podem ser classificadas como equações de evolução – a situação é completamente análoga, levando à formulação do

Problema de Cauchy: Estudar as soluções de uma equação de evolução em função de condições iniciais dadas.

A natureza específica das condições iniciais a serem impostas depende exclusivamente da ordem da derivada temporal no operador:

- *Problema de Cauchy para a equação de difusão:*

Dada uma função $a(x)$, determinar $u_a(t, x)$, solução da equação (1.95), tal que

$$u_a(0, x) = a(x). \quad (1.97)$$

- *Problema de Cauchy para a equação de ondas:*

Dadas duas funções $a(x)$ e $b(x)$, determinar $u_{a,b}(t, x)$, solução da equação (1.91), tal que

$$u_{a,b}(0, x) = a(x), \quad \frac{\partial u_{a,b}}{\partial t}(0, x) = b(x). \quad (1.98)$$

Por outro lado, deve-se ressaltar que para equações estáticas tais como a equação de Laplace ou Poisson ou qualquer outra equação elíptica, não faz muito sentido falar em problema de Cauchy ou em condições iniciais, uma vez que nesta situação não existe nenhuma noção natural de “variável tipo tempo” nem de “hipersuperfície de Cauchy”: devido à invariância do operador de Laplace sob rotações em \mathbb{R}^n , não há nenhuma direção preferida que possa caracterizar estas noções. Portanto, neste caso, o problema de Cauchy torna-se, na melhor das hipóteses, altamente artificial.

Uma outra classe de condições suplementares, que é típica das equações diferenciais parciais (de qualquer natureza), possuindo um análogo apenas rudimentar na teoria das equações diferenciais ordinárias, são as condições de fronteira. Quando consideramos um operador diferencial parcial – mesmo um a coeficientes constantes – sobre um domínio Ω com fronteira $\partial\Omega$, as condições de fronteira prescreverão o comportamento das soluções procuradas sobre a “parte espacial” de $\partial\Omega$. No caso de uma equação estática onde não há noção de tempo, esta “parte espacial” abrange toda a fronteira $\partial\Omega$. Para equações de evolução (em $n + 1$ dimensões) e se o domínio Ω for da forma $I \times \Omega_0$ onde I é um intervalo aberto em \mathbb{R} e Ω_0 é um domínio em \mathbb{R}^n , a “parte espacial” da fronteira $\partial\Omega$ é $I \times \partial\Omega_0$.

Um aspecto que merece destaque é o fato de que quando o domínio Ω não for limitado, o conceito de “condições de fronteira” deve ser entendido num sentido amplo: mais exatamente, ele inclui condições sobre o comportamento assintótico das soluções (tipicamente condições de decaimento) na medida em que nos aproximamos da “parte da fronteira que se encontra no infinito”.

No caso da equação de Laplace ou Poisson (1.96), por exemplo, temos (no mínimo) duas alternativas naturais para formular condições de fronteira. De fato, dado um domínio Ω em \mathbb{R}^n com fronteira $\partial\Omega$, podemos procurar as soluções u da equação de Poisson (1.96) em Ω , onde a fonte f é uma função dada sobre Ω , que resolvem o

- *Problema de Dirichlet:*

$$u|_{\partial\Omega} = g_D, \quad (1.99)$$

onde g_D é uma função dada sobre $\partial\Omega$,

ou o

- *Problema de Neumann:*

$$\frac{\partial u}{\partial n} \equiv \mathbf{n} \cdot \nabla u|_{\partial\Omega} = g_N, \quad (1.100)$$

onde g_N é uma função dada sobre $\partial\Omega$, sendo que $\partial/\partial n$ denota a derivada normal à fronteira.

Estes mesmos dois tipos de problema também aparecem, de maneira natural, em equações de evolução como a equação de difusão ou a equação de ondas: dado um domínio Ω_0 em \mathbb{R}^n com fronteira $\partial\Omega_0$ e um intervalo I em \mathbb{R} , procuramos para ambas as equações as suas soluções que resolvem o

- *Problema de Dirichlet:*

$$u|_{I \times \partial\Omega_0} = g_D, \quad (1.101)$$

onde g_D é uma função dada sobre $I \times \partial\Omega_0$,

ou o

- *Problema de Neumann:*

$$\frac{\partial u}{\partial n} \equiv \mathbf{n} \cdot \nabla u|_{\partial\Omega} = g_N, \quad (1.102)$$

onde g_N é uma função dada sobre $I \times \partial\Omega_0$, sendo que $\partial/\partial n$ denota a derivada normal à fronteira.

Mais concretamente, considere o exemplo da equação do calor em uma barra finita (e suficientemente fina para que possamos negligenciar a sua extensão nas direções transversais). Matematicamente, trata-se da equação de difusão unidimensional,

$$\frac{\partial u}{\partial t} - \lambda \frac{\partial^2 u}{\partial x^2} = f \quad (1.103)$$

sobre um intervalo em \mathbb{R} , digamos $\Omega_0 =]a, b[$, na qual, fisicamente, a variável dependente u representa a temperatura, considerada como função do tempo t e da posição x .¹³ Então temos os seguintes dois tipos de condições de fronteira:

- Problema de Dirichlet:

$$u(t, a) = g_a(t) \quad , \quad u(t, b) = g_b(t) \quad , \quad (1.104)$$

onde g_a e g_b são funções dadas de t . O caso mais simples em que g_a e g_b não dependem de t corresponde a um problema de condução do calor em uma barra finita cujas duas extremidades estão em contato com banhos térmicos com temperaturas dadas por g_a e g_b .

- Problema de Neumann:

$$\frac{\partial u}{\partial x}(t, a) = g_a(t) \quad , \quad \frac{\partial u}{\partial x}(t, b) = g_b(t) \quad , \quad (1.105)$$

onde g_a e g_b são funções dadas de t . O caso mais simples $g_a = g_b = 0$ corresponde a um problema de condução do calor em uma barra finita cujas duas extremidades estão termicamente isoladas.

De modo análogo, para a equação da corda vibrante (1.92) sobre um intervalo em \mathbb{R} ou a equação da membrana vibrante (1.93) sobre um domínio retangular ou circular em \mathbb{R}^2 , as condições de fronteira no problema de Dirichlet são as seguintes:

- Equação da corda vibrante, $\Omega_0 =]a, b[$:

$$u(t, a) = g_a(t) \quad , \quad u(t, b) = g_b(t) \quad , \quad (1.106)$$

onde g_a e g_b são funções dadas de t . A escolha mais simples $g_a = g_b = 0$ corresponde à corda de extremos fixos.

- Equação da membrana vibrante retangular, $\Omega_0 =]a, b[\times]c, d[$:

$$\begin{aligned} u(t, a, y) &= g_a(t, y) \quad , \quad u(t, b, y) = g_b(t, y) \quad , \\ u(t, x, c) &= g_c(t, x) \quad , \quad u(t, x, d) = g_d(t, x) \quad , \end{aligned} \quad (1.107)$$

onde g_a e g_b são funções dadas de t e de $y \in [c, d]$, enquanto que g_c e g_d são funções dadas de t e de $x \in [a, b]$, com

$$g_a(t, c) = g_c(t, a) \quad , \quad g_a(t, d) = g_d(t, a) \quad , \quad g_b(t, c) = g_c(t, b) \quad , \quad g_b(t, d) = g_d(t, b) \quad .$$

A escolha mais simples $g_a = g_b = g_c = g_d = 0$ corresponde à membrana de bordo fixo (trampolim).

¹³Na verdade, tal interpretação é apenas uma aproximação, pois a rigor, a variável u neste contexto representa a energia térmica e não a temperatura: em sistemas termodinâmicos, a equação do calor descreve o transporte, por difusão, da energia térmica. Assim, identificar u com a temperatura do sistema se justifica na medida em que a energia térmica é uma função linear da temperatura, o que ocorre em muitos sistemas, sendo que o coeficiente de proporcionalidade recebe o nome de “calor específico”. Porém, este procedimento perde sua validade quando aparecem grandes variações de temperatura (seja no tempo, seja no espaço) e se torna completamente inadequado perto de um ponto de transição de fase, onde a energia térmica, como função da temperatura, deixa de ser contínua.

- Equação da membrana vibrante circular, $\Omega_0 = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 < R^2\}$:

$$u(t, x, y) = g(t, \varphi) \quad \text{quando} \quad x^2 + y^2 = R^2, \quad \begin{matrix} x = R \cos \varphi \\ y = R \sin \varphi \end{matrix}. \quad (1.108)$$

onde g é uma função dada de t e da variável angular φ usada para parametrizar o círculo. A escolha mais simples $g = 0$ corresponde à membrana de bordo fixo (tambor).

1.8 Solução formal: função de Green e núcleo

Encerrando este capítulo introdutório e, em particular, nossa discussão de aspectos formais de operadores e equações diferenciais parciais, e com a intenção de motivar a utilidade do cálculo de distribuições a ser desenvolvido no Capítulo 3 para a resolução de problemas na área, apresentamos nesta seção a noção de função de Green, ou solução fundamental, de um operador diferencial parcial. Também mostramos como resolver duas das mais importantes equações diferenciais parciais em $\mathbb{R} \times \mathbb{R}^n$, a equação de difusão e a equação de ondas, através do conhecimento de uma solução específica da equação homogênea, chamada núcleo. Os argumentos a serem utilizados são puramente formais no sentido de que as funções envolvidas não são funções bem comportadas e que as operações às quais serão sujeitas são aplicadas segundo as regras usuais do cálculo clássico, sem que as hipóteses do cálculo clássico sejam satisfeitas. A meta principal do cálculo de distribuições será fornecer uma base matemática rigorosa para essas manipulações formais.

Começamos por introduzir a famosa *função delta de Dirac* que, por definição, é uma “função” δ sobre \mathbb{R}^n com as seguintes propriedades:

$$\delta(x) = 0 \quad \text{para} \quad x \neq 0, \quad (1.109)$$

$$\int d^n x \delta(x) = 1. \quad (1.110)$$

Mais geralmente, exige-se que para uma certa classe de funções φ sobre \mathbb{R}^n (como, por exemplo, a classe de todas as funções contínuas sobre \mathbb{R}^n), valha

$$\int d^n x \delta(x) \varphi(x) = \varphi(0). \quad (1.111)$$

Aplicando uma das propriedades fundamentais da integral, a saber sua invariância por translações, obtemos que para qualquer vetor $a \in \mathbb{R}^n$,

$$\int d^n x \delta(x - a) \varphi(x) = \varphi(a). \quad (1.112)$$

Da mesma forma, podemos efetuar uma mudança linear de variáveis na integral para concluir que para qualquer matriz A inversível,

$$\delta(Ax) = |\det A|^{-1} \delta(x). \quad (1.113)$$

Em particular, a função delta de Dirac é par: $\delta(-x) = \delta(x)$. Finalmente, quando f pertence a uma certa classe de funções sobre \mathbb{R}^n (como, por exemplo, a classe de todas as funções contínuas sobre \mathbb{R}^n), vale a regra

$$f(x) \delta(x) = f(0) \delta(x) . \quad (1.114)$$

Outra função importante é a *função de Heaviside* θ sobre \mathbb{R} , definida por

$$\theta(t) = \begin{cases} 1 & \text{para } t > 0 \\ 0 & \text{para } t < 0 \end{cases} \quad (1.115)$$

(O valor de θ em 0 não é definido: pode ser 0, 1, 1/2 ou qualquer outro valor.) Finalmente, como veremos mais adiante, a função delta de Dirac (sobre \mathbb{R}) é simplesmente a derivada da função de Heaviside:

$$\theta'(t) = \delta(t) . \quad (1.116)$$

Repetimos que todas estas considerações, pelo menos por enquanto, não passam de um conjunto de regras formais. Em primeiro lugar, a função delta de Dirac não existe como uma função no sentido usual, pois pelas propriedades da integral, a condição (1.109) implica que as integrais aparecendo nas equações (1.110)-(1.112) não poderiam ser diferentes de 0. Além disto, a função de Heaviside obviamente não é diferenciável no sentido usual. Portanto, deve-se entender que as integrais nas equações (1.110)-(1.112) não são integrais clássicas e a derivada na equação (1.116) não é uma derivada clássica. Mesmo assim, vamos utilizar a regra de Leibniz para a diferenciação de um produto e a regra de diferenciação sob o sinal de integração, entre outras, como se não precisássemos nos importar com tais escrúpulos!

Isso posto, podemos definir a noção de *função de Green* ou *solução fundamental* de um operador diferencial parcial $P(\partial)$: é uma “função” G tal que

$$P(\partial)G = \delta . \quad (1.117)$$

A importância deste conceito provém do fato de que a partir de uma solução G da equação não-homogênea com a função delta de Dirac como fonte, é fácil deduzir uma solução u_f da equação não-homogênea

$$P(\partial)u_f = f , \quad (1.118)$$

para uma grande classe de funções fonte F sobre \mathbb{R}^n (como, por exemplo a classe de todas as funções contínuas sobre \mathbb{R}^n de suporte compacto): basta definir u_f como sendo a *convolução* de f e G :

$$u_f(x) = \int d^n y f(y) G(x - y) . \quad (1.119)$$

De fato, um cálculo formal simples mostra que a expressão em (1.119) satisfaz a

equação diferencial (1.118):

$$\begin{aligned}
 (P(\partial)u_f)(x) &= \int d^n y f(y) P(\partial_x) G(x-y) \\
 &= \int d^n y f(y) P(\partial_z) G(z) \Big|_{z=x-y} \\
 &= \int d^n y f(y) \delta(x-y) = f(x) .
 \end{aligned}$$

Portanto, a solução geral de uma equação diferencial parcial se reduz à solução de dois problemas mais simples: a determinação de uma função de Green e a solução geral da respectiva equação homogênea.

Para certas equações de evolução como a equação de difusão e a equação de ondas, estes dois problemas podem ser reduzidos ainda mais: basta achar uma única solução específica da equação homogênea, o núcleo.

O *núcleo* da equação de difusão é, por definição, a solução D da equação homogênea

$$\left(\frac{\partial}{\partial t} - \lambda \Delta\right) D(t, x) = 0 , \quad (1.120)$$

sujeita à condição inicial

$$D(0, x) = \delta(x) . \quad (1.121)$$

Então a solução geral da equação homogênea, isto é, a solução u_a da equação (1.95) com $f = 0$ sujeita à condição inicial (1.97), é dada por

$$u_a(t, x) = \int d^n y a(y) D(t, x-y) . \quad (1.122)$$

De fato, um cálculo formal simples mostra que a expressão em (1.122) satisfaz a equação diferencial (1.95),

$$\begin{aligned}
 &\left(\frac{\partial}{\partial t} - \lambda \Delta\right) u_a(t, x) \\
 &= \int d^n y a(y) \left(\frac{\partial}{\partial t} - \lambda \sum_{i=1}^n \frac{\partial^2}{\partial z_i^2}\right) D(t, z) \Big|_{z=x-y} \\
 &= 0 ,
 \end{aligned}$$

assim como a condição inicial (1.97),

$$\begin{aligned}
 u_a(0, x) &= \int d^n y a(y) D(0, x-y) \\
 &= \int d^n y a(y) \delta(x-y) \\
 &= a(x) .
 \end{aligned}$$

Além disto,

$$G(t, x) = \theta(t) D(t, x) \quad (1.123)$$

é uma função de Green para o operador de difusão, pois

$$\begin{aligned} \left(\frac{\partial}{\partial t} - \lambda \Delta \right) G(t, x) &= \theta'(t) D(t, x) + \theta(t) \left(\frac{\partial}{\partial t} - \lambda \Delta \right) D(t, x) \\ &= \delta(t) D(t, x) = \delta(t) D(0, x) \\ &= \delta(t) \delta(x) = \delta(t, x) . \end{aligned}$$

De maneira análoga, o *núcleo* da equação de ondas é, por definição, a solução D da equação homogênea

$$\square D = \left(\frac{1}{c^2} \frac{\partial^2}{\partial t^2} - \Delta \right) D = 0 , \quad (1.124)$$

com as condições iniciais

$$D(0, x) = 0 , \quad \frac{\partial D}{\partial t}(0, x) = \delta(x) . \quad (1.125)$$

Então a solução geral da equação homogênea, isto é, a solução $u_{a,b}$ da equação (1.91) com $f = 0$ sujeita à condição inicial (1.98), é dada por

$$u_{a,b}(t, x) = \int d^n y \left(a(y) \frac{\partial D}{\partial t}(t, x - y) + b(y) D(t, x - y) \right) . \quad (1.126)$$

De fato, um cálculo formal simples mostra que a expressão em (1.124) satisfaz a equação diferencial (1.91),

$$\begin{aligned} &\left(\frac{1}{c^2} \frac{\partial^2}{\partial t^2} - \Delta \right) u_{a,b}(t, x) \\ &= \int d^n y a(y) \left(\frac{1}{c^2} \frac{\partial^2}{\partial t^2} - \sum_{i=1}^n \frac{\partial^2}{\partial z_i^2} \right) \frac{\partial D}{\partial t}(t, z) \Big|_{z=x-y} \\ &\quad + \int d^n y b(y) \left(\frac{1}{c^2} \frac{\partial^2}{\partial t^2} - \sum_{i=1}^n \frac{\partial^2}{\partial z_i^2} \right) D(t, z) \Big|_{z=x-y} \\ &= \int d^n y a(y) \frac{\partial}{\partial t} \left(\frac{1}{c^2} \frac{\partial^2}{\partial t^2} - \sum_{i=1}^n \frac{\partial^2}{\partial z_i^2} \right) D(t, z) \Big|_{z=x-y} \\ &\quad + \int d^n y b(y) \left(\frac{1}{c^2} \frac{\partial^2}{\partial t^2} - \sum_{i=1}^n \frac{\partial^2}{\partial z_i^2} \right) D(t, z) \Big|_{z=x-y} \\ &= 0 , \end{aligned}$$

assim como as condições iniciais (1.98),

$$\begin{aligned} u_{a,b}(0, x) &= \int d^n y \left(a(y) \frac{\partial D}{\partial t}(0, x - y) + b(y) D(0, x - y) \right) \\ &= \int d^n y a(y) \delta(x - y) \\ &= a(x) , \end{aligned}$$

$$\begin{aligned} \frac{\partial u_{a,b}}{\partial t}(0, x) &= \int d^n y \left(a(y) \frac{\partial^2 D}{\partial t^2}(0, x - y) + b(y) \frac{\partial D}{\partial t}(0, x - y) \right) \\ &= \int d^n y \left(c^2 a(y) \sum_{i=1}^n \frac{\partial^2 D}{\partial z_i^2}(0, z) \Big|_{z=x-y} + b(y) \delta(x - y) \right) \\ &= b(x) . \end{aligned}$$

Além disto,

$$G(t, x) = c^2 \theta(t) D(t, x) \tag{1.127}$$

é uma função de Green para o operador de ondas, pois

$$\theta'(t) D(t, x) = D(t, x) \delta(t) = D(0, x) \delta(t) = 0$$

e portanto

$$\begin{aligned} \left(\frac{1}{c^2} \frac{\partial^2}{\partial t^2} - \Delta \right) G(t, x) &= \frac{\partial}{\partial t} \left(\theta'(t) D(t, x) + \theta(t) \frac{\partial D}{\partial t}(t, x) \right) - c^2 \theta(t) \Delta D(t, x) \\ &= \theta'(t) \frac{\partial D}{\partial t}(t, x) + \theta(t) \left(\frac{\partial^2}{\partial t^2} - c^2 \Delta \right) D(t, x) \\ &= \delta(t) \delta(x) = \delta(t, x) . \end{aligned}$$

2 Problemas e Resultados Clássicos

Neste capítulo, apresentamos alguns problemas e resultados clássicos referentes aos três operadores diferenciais parciais clássicos, que são o operador de Laplace, o operador de difusão e o operador de ondas.

2.1 O operador de Laplace

Nesta primeira seção, discutimos várias propriedades proeminentes do operador de Laplace. Após considerações gerais acerca da questão de unicidade e existência de soluções dos problemas de Dirichlet e de Neumann, apresentamos vários teoremas e conceitos centrais da teoria, entre eles o princípio do máximo/mínimo, a fórmula de Poisson para a solução da equação não-homogênea em \mathbb{R}^n , a noção de função de Green para um domínio, a terceira identidade de Green, o teorema da média esférica para funções harmônicas, o teorema de regularidade e o teorema sobre singularidades removíveis de funções harmônicas.

Inicialmente, lembramos que, de acordo com as notações e convenções adotadas neste livro, um domínio em \mathbb{R}^n é, por definição, um subconjunto aberto e conexo de \mathbb{R}^n , com fecho denotado por $\bar{\Omega}$ e fronteira denotada por $\partial\Omega = \bar{\Omega} \setminus \Omega$. No que segue, consideraremos, além do próprio \mathbb{R}^n , tão somente domínios limitados, sendo que problemas em domínios não limitados são geralmente estudados considerando estes como limites de domínios limitados. Uma hipótese mais restritiva, a ser imposta na medida em que se torne necessária ou conveniente, é que a fronteira $\partial\Omega$ de Ω seja uma hipersuperfície de classe C^1 em \mathbb{R}^n , sempre orientada de tal forma que o campo unitário \mathbf{n} normal a $\partial\Omega$ aponta para fora de Ω . Lembramos também (veja Capítulo 1.7) que dado um domínio limitado Ω em \mathbb{R}^n , o problema de Dirichlet em Ω consiste em procurar as soluções das equações

$$\Delta u = f \quad \text{e} \quad u|_{\partial\Omega} = g_D, \quad (2.1)$$

onde f é uma função dada sobre Ω e g_D é uma função dada sobre $\partial\Omega$, enquanto que dado um domínio limitado Ω em \mathbb{R}^n cuja fronteira $\partial\Omega$ é uma hipersuperfície de

classe C^1 em \mathbb{R}^n , o problema de Neumann em Ω consiste em procurar as soluções das equações

$$\Delta u = f \quad \text{e} \quad \frac{\partial u}{\partial n} \equiv \mathbf{n} \cdot \nabla u|_{\partial\Omega} = g_N, \quad (2.2)$$

onde f é uma função dada sobre Ω e g_N é uma função dada sobre $\partial\Omega$, sendo que $\partial/\partial n$ denota a derivada normal à fronteira.

Nota-se que, com estes dados, os dois problemas ainda não são bem postos, pois não especificamos as propriedades a serem exigidas das funções f e g_D ou g_N , nem as propriedades das soluções u que consideraremos admissíveis. Condições mínimas no contexto clássico, em termos das notações e convenções adotadas neste livro, são as seguintes:¹

- Condições mínimas gerais de continuidade:

$$f \in C(\Omega) \quad , \quad g_D, g_N \in C(\partial\Omega) \quad , \quad (2.3)$$

i.e., a fonte f é uma função contínua sobre Ω e as funções g_D e g_N são contínuas sobre $\partial\Omega$;

- Condições mínimas para o problema de Dirichlet:

$$u \in C^2(\Omega) \cap C(\bar{\Omega}) \quad , \quad (2.4)$$

i.e., u é, no mínimo, uma função de classe C^2 sobre Ω que admite uma extensão contínua a $\bar{\Omega}$.

- Condições mínimas para o problema de Neumann:

$$u \in C^2(\Omega) \cap C^1(\bar{\Omega}) \quad , \quad (2.5)$$

i.e., u é, no mínimo, uma função de classe C^2 sobre Ω tal que u e ∇u admitem extensões contínuas a $\bar{\Omega}$.

Hipóteses mais restritivas poderão ser introduzidas na medida em que se tornam necessárias, principalmente para a formulação de critérios que garantam unicidade e/ou existência de soluções. Tipicamente, podemos supor que u e/ou f pertençam a um dos espaços $C^k(\bar{\Omega})$, para valores apropriados de k .

Como observação preliminar e elementar, notamos que – independentemente da questão das condições corretas de regularidade a serem impostas sobre as funções envolvidas – o problema de Neumann difere do problema de Dirichlet em dois aspectos. Primeiro, é óbvio que uma solução u do problema de Neumann será, no máximo, unicamente determinada a menos de uma constante aditiva. Segundo, o teorema da divergência (Teorema B.3) implica que uma solução do problema de

¹Excluímos aqui condições que exigem tão somente a existência (e possivelmente a continuidade) de algumas derivadas parciais mas não de outras (da mesma ordem), por exemplo das segundas derivadas parciais puras mas não necessariamente das segundas derivadas mistas: consideramos tais condições como altamente artificiais pois são incompatíveis até com o simples critério de invariância rotacional.

Neumann só pode existir quando a integral da função f sobre o domínio Ω coincide com a integral da função g_N sobre a hipersuperfície $\partial\Omega$:

$$\int_{\Omega} d^n x f(x) = \int_{\partial\Omega} d\sigma(x) g_N(x). \quad (2.6)$$

Com essa ressalva, podemos em ambos os casos dar uma resposta positiva à questão de unicidade, sob hipóteses razoavelmente gerais. Um argumento simples para estabelecer unicidade da solução para o problema de Dirichlet e unicidade da solução a menos de uma constante aditiva para o problema de Neumann, em qualquer domínio limitado Ω em \mathbb{R}^n cuja fronteira $\partial\Omega$ é uma hipersuperfície de classe C^1 em \mathbb{R}^n , e mediante a hipótese adicional de que a solução pertença ao espaço $C^2(\overline{\Omega})$, baseia-se na primeira identidade de Green: De fato, a diferença $u = u_1 - u_2$ de duas soluções $u_1, u_2 \in C^2(\overline{\Omega})$ do mesmo problema é uma função $u \in C^2(\overline{\Omega})$ harmônica sobre Ω tal que u (problema de Dirichlet), ou $\partial u / \partial n$ (problema de Neumann) se anula sobre a fronteira $\partial\Omega$ de Ω . Portanto, podemos aplicar a primeira identidade de Green (B.44) (Teorema B.5) com $\varphi = u = \psi$ para concluir que

$$\int_{\Omega} d^n x (\nabla u)^2 = \int_{\partial\Omega} d\sigma \cdot u \nabla u - \int_{\Omega} d^n x u \Delta u$$

se anula; portanto, ∇u se anula identicamente em Ω . Logo, u é constante em $\overline{\Omega}$.

Observamos, porém, que no caso do problema de Dirichlet, a afirmação de unicidade vale sob hipóteses consideravelmente mais gerais:

Teorema 2.1 (Unicidade da Solução Clássica do Problema de Dirichlet): *Seja Ω um domínio limitado em \mathbb{R}^n . Então, o problema de Dirichlet em Ω sempre admite no máximo uma solução, isto é, para toda função $f \in C(\Omega)$ e toda função $g_D \in C(\partial\Omega)$, existe no máximo uma função $u \in C^2(\Omega) \cap C(\overline{\Omega})$ que satisfaz a equação (2.1).*

DEMONSTRAÇÃO: A diferença de duas soluções do problema descrito no teorema é uma função em $C^2(\Omega) \cap C(\overline{\Omega})$ que é harmônica sobre Ω e se anula sobre $\partial\Omega$; portanto, ela se anula sobre $\overline{\Omega}$, devido ao princípio do máximo/mínimo (veja o Teorema 2.3 abaixo). □

Ainda no caso do problema de Dirichlet, a questão da existência da solução, sob hipóteses tão gerais como as acima mencionadas, é bem mais difícil. A estratégia usualmente adotada consiste em dividir o problema em duas partes: a) estabelecer existência de alguma solução da equação não-homogênea em \mathbb{R}^n , sem levar em consideração qualquer tipo de condição de fronteira, e b) estabelecer existência da solução da equação homogênea no domínio Ω , para condições de fronteira gerais.

Quanto ao item a), apresentaremos logo em seguida (veja o Teorema 2.5 abaixo) uma fórmula devida a Poisson que, para fontes f de classe C^1 em \mathbb{R}^n com suporte compacto, fornece soluções u de classe C^2 em \mathbb{R}^n da equação de Poisson $\Delta u = f$. Quanto ao item b), temos o seguinte

Teorema 2.2 (*Existência da Solução Clássica do Problema de Dirichlet Homogêneo*): Seja Ω um domínio limitado em \mathbb{R}^n cuja fronteira $\partial\Omega$ é uma hipersuperfície de classe C^1 em \mathbb{R}^n . Então o problema de Dirichlet homogêneo em Ω sempre admite uma solução, isto é, para toda função $g_D \in C(\partial\Omega)$, existe uma função $u \in C^2(\Omega) \cap C(\bar{\Omega})$ que satisfaz a equação (2.1) (com $f = 0$).

Neste teorema, a hipótese de regularidade sobre a fronteira de Ω é essencial. De fato, existência da solução não pode ser garantida para domínios Ω arbitrários, como mostra o seguinte

Exemplo 2.1 (*Contra-exemplo de Zaremba*): Seja \dot{B}_a^n a bola aberta de raio a em \mathbb{R}^n em torno da origem, com a origem removida:

$$\dot{B}_a^n = \{x \in \mathbb{R}^n / 0 < |x| < a\}.$$

O seu fecho é a bola fechada \bar{B}_a^n de raio a e portanto, a sua fronteira é a união disjunta de duas componentes conexas: a esfera de raio a e a origem:

$$\partial\dot{B}_a^n = S_a^{n-1} \cup \{0\}.$$

Seja g_D a função contínua sobre $\partial\dot{B}_a^n$ dada por

$$g_D(x) = c_1 \quad \text{para } x \in S_a^{n-1} \quad \text{e} \quad g_D(0) = c_0,$$

onde c_0 e c_1 são constantes. Então, se $c_0 \neq c_1$, não existe nenhuma solução do problema de Dirichlet homogêneo em \dot{B}_a^n satisfazendo estas condições de fronteira.

A demonstração desta afirmação será apresentada mais adiante, como conseqüência do teorema sobre singularidades removíveis de funções harmônicas (Teorema 2.9 abaixo).

Observa-se que o contra-exemplo de Zaremba não contradiz a afirmação de existência do Teorema 2.2, pois a componente conexa $\{0\}$ de $\partial\dot{B}_a^n$ não é uma hipersuperfície, sendo que ela forma uma subvariedade de dimensão 0 e não de dimensão $n - 1$.

Tendo em vista esta situação, é conveniente introduzir a seguinte terminologia:

Definição 2.1 Um domínio limitado Ω em \mathbb{R}^n é chamado um **domínio de Dirichlet** se o problema de Dirichlet homogêneo em Ω sempre admite uma solução, ou seja, se para toda função $g_D \in C(\partial\Omega)$, existe uma função $u \in C^2(\Omega) \cap C(\bar{\Omega})$ que satisfaz a equação (2.1) (com $f = 0$). (Pelo Teorema 2.1, esta solução u é necessariamente única.)

Nestes termos, a afirmação de existência do Teorema 2.2 garante que um domínio limitado Ω cuja fronteira $\partial\Omega$ é uma hipersuperfície de classe C^1 em \mathbb{R}^n é um domínio de Dirichlet. Em particular, uma bola aberta é um domínio de Dirichlet, enquanto que – segundo o contra-exemplo de Zaremba – uma bola aberta cujo centro foi removido não é.

A identificação e caracterização completa dos domínios de Dirichlet constitui um problema difícil. É um tema central da “teoria do potencial”, intensamente estudada nos últimos dois séculos por um grande número de matemáticos. Foram desenvolvidas várias técnicas de abordagem, entre elas o método das funções de Green que, no caso de um domínio limitado cuja fronteira é uma hipersuperfície de classe C^1 , permite reduzir a resolução do problema de fronteira geral à resolução de um único problema desta natureza, para uma escolha específica da fonte e da condição de fronteira. Apresentaremos a seguir alguns aspectos selecionados deste método, sem pretensão de completude.

2.1.1 O princípio do máximo/mínimo

Um dos teoremas mais importantes sobre funções harmônicas é o seguinte:

Teorema 2.3 (*Princípio do máximo/mínimo para funções harmônicas*): *Sejam Ω um domínio limitado em \mathbb{R}^n e $u \in C^2(\Omega) \cap C(\overline{\Omega})$ uma função harmônica a valores reais. Então u assume seu máximo e seu mínimo na fronteira $\partial\Omega$ de Ω .*

Mais geralmente, temos

Teorema 2.4 (*Princípio do máximo/mínimo para soluções da equação de Poisson*): *Sejam Ω um domínio limitado em \mathbb{R}^n e $u \in C^2(\Omega) \cap C(\overline{\Omega})$ uma função a valores reais. Então*

- se $\Delta u \geq 0$ em Ω , u assume seu máximo na fronteira $\partial\Omega$ de Ω ,
- se $\Delta u \leq 0$ em Ω , u assume seu mínimo na fronteira $\partial\Omega$ de Ω .

Observa-se que segundo as hipóteses, a função u , sendo contínua sobre o compacto $\overline{\Omega}$, assume seu máximo e seu mínimo em $\overline{\Omega}$. Observa-se também que o princípio do máximo e o princípio do mínimo são equivalentes, pela simples substituição $u \rightarrow -u$.

DEMONSTRAÇÃO: Suponhamos inicialmente que $\Delta u > 0$ em Ω . Neste caso, podemos afirmar que u assume seu máximo *apenas* na fronteira $\partial\Omega$ de Ω e não no interior. De fato, se $x_0 \in \Omega$ fosse um máximo (ou mesmo um máximo local) de u , teríamos

$$\frac{\partial u}{\partial x_i}(x_0) = 0 \quad \text{e} \quad \frac{\partial^2 u}{\partial x_i^2}(x_0) \leq 0 \quad (1 \leq i \leq n)$$

e portanto, $(\Delta u)(x_0) \leq 0$, contrariamente à hipótese.

Para tratar do caso geral quando $\Delta u \geq 0$ em Ω , escolhemos um ponto de referência $x_0 \in \Omega$ qualquer, pomos

$$R = \sup_{x \in \overline{\Omega}} |x - x_0|$$

e introduzimos, para todo $\epsilon > 0$, a função $u_\epsilon \in C^2(\Omega) \cap C(\bar{\Omega})$ dada por

$$u_\epsilon(x) = u(x) + \epsilon |x - x_0|^2 \quad \text{para } x \in \bar{\Omega} .$$

Então

$$\Delta u_\epsilon = \Delta u + 2n\epsilon > 0 \quad \text{em } \Omega$$

e portanto, pelo resultado anterior,

$$\sup_{x \in \bar{\Omega}} u_\epsilon(x) = \sup_{x \in \partial\Omega} u_\epsilon(x) .$$

Logo, para qualquer $z \in \bar{\Omega}$,

$$\begin{aligned} u(z) &\leq u_\epsilon(z) \leq \sup_{x \in \bar{\Omega}} u_\epsilon(x) = \sup_{x \in \partial\Omega} u_\epsilon(x) \\ &\leq \sup_{x \in \partial\Omega} u(x) + \epsilon \sup_{x \in \partial\Omega} |x - x_0|^2 \\ &\leq \sup_{x \in \partial\Omega} u(x) + \epsilon R^2 . \end{aligned}$$

Tomando o limite $\epsilon \rightarrow 0$, chegamos ao resultado desejado. \square

2.1.2 Potencial de Coulomb

Inicialmente, queremos determinar a solução geral u da equação de Laplace no domínio $\mathbb{R}^n \setminus \{0\}$ que apresenta simetria esférica, ou seja, que satisfaz

$$u(x) = u(r) \quad \text{onde} \quad r^2 \equiv |x|^2 = \sum_{i=1}^n x_i^2 . \quad (2.7)$$

Para tais funções u , temos

$$\begin{aligned} \frac{\partial u}{\partial x_i}(x) &= u'(r) \frac{x_i}{r} , \\ \frac{\partial^2 u}{\partial x_i^2}(x) &= u''(r) \frac{x_i^2}{r^2} + u'(r) \frac{r^2 - x_i^2}{r^3} , \end{aligned}$$

e portanto

$$(\Delta u)(x) = u''(r) + \frac{n-1}{r} u'(r) . \quad (2.8)$$

Logo, a equação de Laplace se reduz à seguinte equação diferencial ordinária na variável radial r :

$$u''(r) + \frac{n-1}{r} u'(r) = 0 .$$

A solução geral é

$$u(r) = \begin{cases} -\frac{c}{(n-2)r^{n-2}} + c' & (n > 2) \\ c \ln r + c' & (n = 2) , \\ cr + c' & (n = 1) \end{cases}$$

onde c e c' são constantes. Logo,

$$\frac{\partial u}{\partial x_i}(x) = \frac{c}{r^{n-1}} \frac{x_i}{r} ,$$

e usando a relação (B.24), obtemos

$$\int_{S_a^{n-1}} d\sigma \cdot \nabla u = c \operatorname{vol}(S^{n-1}) ,$$

independentemente de a .

Definição 2.2 *O potencial de Coulomb é a função G_0 de classe C^∞ sobre $\mathbb{R}^n \setminus \{0\}$ que a) é harmônica, b) é invariante sob o grupo de rotações em \mathbb{R}^n e c) satisfaz as condições de normalização*

$$\int_{S_a^{n-1}} d\sigma \cdot \nabla G_0 = 1 \quad \text{para todo } a > 0 \quad (2.9)$$

e

$$\begin{aligned} \lim_{r \rightarrow \infty} G_0(r) &= 0 & (n > 2) \\ G_0(r_0) &= 0 & (n = 2) , \\ G_0(0) &= 0 & (n = 1) \end{aligned} \quad (2.10)$$

onde r_0 é uma constante. Explicitamente,

$$G_0(x) = \begin{cases} -\frac{1}{(n-2) \operatorname{vol}(S^{n-1})} \frac{1}{r^{n-2}} & (n > 2) \\ \frac{1}{2\pi} \ln \frac{r}{r_0} & (n = 2) . \\ \frac{1}{2} r & (n = 1) \end{cases} \quad (2.11)$$

Diferenciando, obtemos

$$\frac{\partial G_0}{\partial x_i}(x) = \frac{1}{\operatorname{vol}(S^{n-1})} \frac{x_i}{r^n} , \quad (2.12)$$

$$\frac{\partial^2 G_0}{\partial x_i \partial x_j}(x) = \frac{1}{\operatorname{vol}(S^{n-1})} \frac{r^2 \delta_{ij} - n x_i x_j}{r^{n+2}} . \quad (2.13)$$

Para uso posterior, notamos que G_0 e suas derivadas parciais são singulares na origem, mas que G_0 e suas primeiras derivadas parciais ainda são integráveis sobre compactos de \mathbb{R}^n que contêm a origem, enquanto que suas derivadas parciais de ordem ≥ 2 não o são.² De fato, para a bola B_a^n de raio a em torno da origem, com $0 < a < r_0$ no caso $n = 2$, temos

$$\int_{B_a^n} d^n x |G_0(x)| = \text{vol}(S^{n-1}) \int_0^a dr r^{n-1} |G_0(r)| = g_n(a), \quad (2.14)$$

onde

$$g_n(a) = \begin{cases} \frac{1}{2(n-2)} a^2 & (n > 2) \\ \frac{1}{4} \left(a^2 \left| \ln \frac{a^2}{r_0^2} \right| + r_0^2 - |r_0^2 - a^2| \right) & (n = 2) \\ \frac{1}{2} a^2 & (n = 1) \end{cases}, \quad (2.15)$$

enquanto que

$$\int_{B_a^n} d^n x |\nabla G_0(x)| = a. \quad (2.16)$$

Este caráter singular da função G_0 é a origem de muitas dificuldades técnicas que, no entanto, podem ser superadas pela idéia de introduzir uma *regularização* de G_0 , isto é, uma família de funções $G_{0,\epsilon}$ de classe C^∞ sobre \mathbb{R}^n que diferem de G_0 tão somente numa bola de raio ϵ em torno da singularidade e tais que

$$G_0(x) = \lim_{\epsilon \rightarrow 0} G_{0,\epsilon}(x), \quad (2.17)$$

pelo menos pontualmente (para todo $x \neq 0$). (Haverá também convergência em outros sentidos, mas não precisaremos deste fato aqui.) O método padrão de efetuar uma tal regularização é através de multiplicação por uma função reguladora. Para tanto, introduzimos primeiro uma função χ de classe C^∞ sobre \mathbb{R} a valores reais tal que

$$0 \leq \chi \leq 1, \quad \text{supp } \chi \subset [-1, 1], \quad \chi \equiv 1 \quad \text{sobre } [-\frac{1}{2}, \frac{1}{2}]. \quad (2.18)$$

(A existência de uma tal função será demonstrada no Capítulo 3.3.) Definimos agora, para todo $\epsilon > 0$,

$$\chi_\epsilon(x) = \chi(r/\epsilon) \quad \text{para } x \in \mathbb{R}^n. \quad (2.19)$$

Destarte, χ_ϵ é uma função de classe C^∞ sobre \mathbb{R}^n a valores reais tal que

$$0 \leq \chi_\epsilon \leq 1, \quad \text{supp } \chi_\epsilon \subset \bar{B}_\epsilon^n, \quad \chi_\epsilon \equiv 1 \quad \text{sobre } \bar{B}_{\epsilon/2}^n. \quad (2.20)$$

²O caso $n = 1$ é excepcional no sentido de que as singularidades de G_0 e suas derivadas são mais amenas: na origem, G_0 é contínua e G'_0 apresenta apenas uma descontinuidade finita.

Também temos

$$\partial_i \chi_\epsilon(x) = \frac{\chi'(r/\epsilon)}{\epsilon} \frac{x_i}{r}, \quad (2.21)$$

$$\partial_i \partial_j \chi_\epsilon(x) = \frac{\chi''(r/\epsilon)}{\epsilon^2} \frac{x_i x_j}{r^2} + \frac{\chi'(r/\epsilon)}{\epsilon} \frac{r^2 \delta_{ij} - x_i x_j}{r^3}, \quad (2.22)$$

e portanto

$$|\partial_i \chi_\epsilon(x)| \leq \frac{c_1}{\epsilon}, \quad |\partial_i \partial_j \chi_\epsilon(x)| \leq \frac{c_2}{\epsilon^2} + \frac{c_1}{\epsilon r}, \quad (2.23)$$

onde

$$c_1 = \sup_{t \in \mathbb{R}} |\chi'(t)|, \quad c_2 = \sup_{t \in \mathbb{R}} |\chi''(t)|. \quad (2.24)$$

Para uso posterior, definimos ainda

$$l_n = \text{vol}(S^{n-1}) \int_0^\infty dr r^{n-1} \chi(r). \quad (2.25)$$

Então a regularização $G_{0,\epsilon}$ de G_0 é definida por

$$G_{0,\epsilon}(x) = G_0(x) (1 - \chi_\epsilon(x)). \quad (2.26)$$

Isso posto, podemos formular o primeiro teorema central que, dentro de um contexto clássico, consubstancia a idéia de resolver uma equação diferencial parcial não-homogênea através de uma convolução entre a fonte e uma função de Green ou solução fundamental – uma idéia que já foi apresentada, em nível formal, no Capítulo 1.8 (veja as equações (1.117)-(1.119)):

Teorema 2.5 (Fórmula de Poisson): *Para toda função $f \in C^r(\mathbb{R}^n)$ ($r \geq 1$) com suporte compacto, a integral*

$$u(x) = \int d^n y f(y) G_0(x-y) \quad (2.27)$$

fornece uma solução $u \in C^{r+1}(\mathbb{R}^n)$ da equação de Poisson $\Delta u = f$ em \mathbb{R}^n .

Observa-se que, formalmente, diferenciando a equação (2.27) sob a integral leva à fórmula

$$\frac{\partial u}{\partial x_i}(x) = \int d^n y f(y) \partial_i G_0(x-y) = \frac{1}{\text{vol}(S^{n-1})} \int d^n y f(y) \frac{x_i - y_i}{|x-y|^n}. \quad (2.28)$$

DEMONSTRAÇÃO: Pretendemos demonstrar várias afirmações deste teorema sob hipóteses mais gerais, que são as seguintes:

- a) A função f é de suporte compacto, mensurável e limitada.
- b) A função f é de suporte compacto, diferenciável e com gradiente ∇f mensurável e limitado.

Primeiro, definimos

$$u_i(x) = \int d^n y f(y) \partial_i G_0(x-y) \quad (2.29)$$

e

$$u_{ij}(x) = \int d^n y \partial_i f(y) \partial_j G_0(x-y) \quad (2.30)$$

e observamos que, embora a função $G_0(x-y)$ e suas primeiras derivadas parciais $\partial_i G_0(x-y)$ (consideradas como funções de y , para x fixo) apresentarem singularidades no ponto $y=x$, as integrais nas equações (2.27) e (2.29) existem se f satisfaz a hipótese a). De fato, pomos

$$\|f\| = \sup_{y \in \mathbb{R}^n} |f(y)| = \sup_{y \in \text{supp} f} |f(y)|$$

e, para qualquer subconjunto compacto K de \mathbb{R}^n ,

$$R_K = \sup_{x \in K, y \in \text{supp} f} d(x, y).$$

Então, para todo $x \in K$, temos $\text{supp} f \subset B_{R_K}(x)$, de modo que podemos utilizar as equações (2.14) e (2.16) para deduzir as estimativas

$$|u(x)| \leq \int_{B_{R_K}(x)} d^n y |f(y)| |G_0(x-y)| \leq g_n(R_K) \|f\|,$$

onde a função g_n é dada pela equação (2.15), e

$$|u_i(x)| \leq \int_{B_{R_K}(x)} d^n y |f(y)| |\partial_i G_0(x-y)| \leq R_K \|f\|.$$

De forma análoga, prova-se que a integral na equação (2.30) existe se f satisfaz a hipótese b).

Para prosseguir, precisaremos da regularização $G_{0,\epsilon}$ de G_0 introduzida acima, que também fornecerá uma regularização u_ϵ de u , definida por

$$u_\epsilon(x) = \int d^n y f(y) G_{0,\epsilon}(x-y). \quad (2.31)$$

Tendo em vista que a função $(x, y) \mapsto G_{0,\epsilon}(x-y)$ é, por construção, de classe C^∞ sobre $\mathbb{R}^n \times \mathbb{R}^n$, a função u_ϵ será de classe C^∞ sobre \mathbb{R}^n , desde que f satisfaça a hipótese a), e

$$\partial_i u_\epsilon(x) = \int d^n y f(y) \partial_i G_{0,\epsilon}(x-y), \quad (2.32)$$

$$\partial_i \partial_j u_\epsilon(x) = \int d^n y f(y) \partial_i \partial_j G_{0,\epsilon}(x-y), \quad (2.33)$$

etc.. Quando f satisfaz a hipótese b), podemos integrar a última equação por partes, obtendo

$$\partial_i \partial_j u_\epsilon(x) = \int d^n y \partial_i f(y) \partial_j G_{0,\epsilon}(x-y). \quad (2.34)$$

Por outro lado, utilizando as estimativas

$$\begin{aligned} |u_\epsilon(x) - u(x)| &\leq \int d^n y |f(y)| |G_0(x-y)| |\chi_\epsilon(x-y)| \\ &\leq \int_{B_\epsilon(x)} d^n y |f(y)| |G_0(x-y)| \leq g_n(\epsilon) \|f\|, \\ |\partial_i u_\epsilon(x) - u_i(x)| &\leq \int d^n y |f(y)| \left(|\partial_i G_0(x-y)| |\chi_\epsilon(x-y)| \right. \\ &\quad \left. + |G_0(x-y)| |\partial_i \chi_\epsilon(x-y)| \right) \\ &\leq \int_{B_\epsilon(x)} d^n y |f(y)| |\partial_i G_0(x-y)| \\ &\quad + \frac{c_1}{\epsilon} \int_{B_\epsilon(x)} d^n y |f(y)| |G_0(x-y)| \\ &\leq \left(\epsilon + c_1 \frac{g_n(\epsilon)}{\epsilon} \right) \|f\|, \end{aligned}$$

que decorrem das equações (2.14)-(2.16) e das estimativas (2.23), (2.24), em conjunto com o fato de que $g_n(\epsilon)/\epsilon \rightarrow 0$ quando $\epsilon \rightarrow 0$, concluímos que

$$\begin{aligned} u(x) &= \lim_{\epsilon \rightarrow 0} u_\epsilon(x), \\ u_i(x) &= \lim_{\epsilon \rightarrow 0} \partial_i u_\epsilon(x), \end{aligned}$$

uniformemente para $x \in \mathbb{R}^n$, desde que f satisfaça a hipótese a). Logo, nestas condições, u é de classe C^1 e $\partial_i u = u_i$. Quando f satisfaz a hipótese b), podemos concluir, de forma análoga, que

$$|\partial_i \partial_j u_\epsilon(x) - u_{ij}(x)| \leq \left(\epsilon + c_1 \frac{g_n(\epsilon)}{\epsilon} \right) \|\partial_i f\|,$$

mostrando que

$$u_{ij}(x) = \lim_{\epsilon \rightarrow 0} \partial_i \partial_j u_\epsilon(x),$$

uniformemente para $x \in \mathbb{R}^n$. Logo, nestas condições, u é de classe C^2 e $\partial_i \partial_j u = u_{ij}$. Em particular, isso ocorre quando f é de classe C^1 e de suporte compacto. De modo análogo, prova-se que quando f é de classe C^r ($r \geq 1$) e de suporte compacto, então u é de classe C^{r+1} e, para $1 \leq j \leq n$ e todo multi-índice $\alpha \in \mathbb{N}^n$ com $|\alpha| \leq r$, vale $\partial_\alpha \partial_j u = u_{\alpha,j}$, onde

$$u_{\alpha,j}(x) = \int d^n y \partial_\alpha f(y) \partial_j G_0(x-y).$$

Finalmente, queremos mostrar que $\Delta u = f$. Temos

$$\Delta u(x) = \lim_{\epsilon \rightarrow 0} \Delta u_\epsilon(x),$$

uniformemente para $x \in K$, onde K é qualquer subconjunto compacto de \mathbb{R}^n , e

$$\Delta u_\epsilon(x) = \int d^n y f(y) \Delta G_{0,\epsilon}(x-y).$$

Pondo $z = x - y$ e observando que

$$(1 - \chi_\epsilon)(z) = 0 \quad \text{se } |z| \leq \epsilon/2,$$

$$\Delta(G_0(1 - \chi_\epsilon))(z) = \Delta G_0(z) = 0 \quad \text{se } |z| \geq \epsilon,$$

concluimos que

$$\Delta u_\epsilon(x) = - \int_{\epsilon/2 \leq |z| \leq \epsilon} d^n z f(x-z) \Delta(G_0 \chi_\epsilon)(z).$$

Como as funções $G_0(z)$ e $\chi_\epsilon(z)$ dependem apenas do módulo $r = |z|$ de z , podemos aplicar a fórmula (2.8) para escrever

$$\Delta G_0(z) = G_0''(r) + \frac{n-1}{r} G_0'(r) = 0$$

e

$$\begin{aligned} \Delta(G_0 \chi_\epsilon)(z) &= (G_0 \chi_\epsilon)''(r) + \frac{n-1}{r} (G_0 \chi_\epsilon)'(r) \\ &= G_0(r) \chi_\epsilon''(r) + \left(2G_0'(r) + \frac{n-1}{r} G_0(r)\right) \chi_\epsilon'(r). \\ &= G_0(r) \frac{\chi''(r/\epsilon)}{\epsilon^2} + \left(2G_0'(r) + \frac{n-1}{r} G_0(r)\right) \frac{\chi'(r/\epsilon)}{\epsilon}. \end{aligned}$$

Substituindo a forma explícita de G_0 (o que requer distinguir os casos $n \neq 2$ e $n = 2$), obtemos

$$\Delta(G_0 \chi_\epsilon)(z) = \frac{1}{\text{vol}(S^{n-1})} \left\{ \frac{n-3}{n-2} \frac{\chi'(r/\epsilon)}{\epsilon r^{n-1}} - \frac{1}{n-2} \frac{\chi''(r/\epsilon)}{\epsilon^2 r^{n-2}} \right\}$$

para $n \neq 2$ (isto é, para $n > 2$ assim como para $n = 1$) e

$$\Delta(G_0 \chi_\epsilon)(z) = \frac{1}{2\pi} \left\{ \left(2 + \ln \frac{r}{r_0}\right) \frac{\chi'(r/\epsilon)}{\epsilon r} + \ln \frac{r}{r_0} \frac{\chi''(r/\epsilon)}{\epsilon^2} \right\}$$

para $n = 2$, o que permite chegar às seguintes conclusões:

- A integral de $\Delta(G_0\chi_\epsilon)(z)$ sobre a casca $\{z \in \mathbb{R}^n / \epsilon/2 \leq |z| \leq \epsilon\}$ vale -1 .

De fato, as fórmulas anteriores podem ser reescritas para mostrar que $\Delta(G_0\chi_\epsilon)(z)$, como função da variável radial $r = |z|$, é uma derivada total, que é facilmente integrada: Temos

$$\Delta(G_0\chi_\epsilon)(z) = \frac{1}{\text{vol}(S^{n-1})} \frac{1}{r^{n-1}} \frac{d}{dr} \left(\chi(r/\epsilon) - \frac{1}{n-2} \frac{r}{\epsilon} \chi'(r/\epsilon) \right)$$

para $n \neq 2$ (isto é, para $n > 2$ assim como para $n = 1$) e

$$\Delta(G_0\chi_\epsilon)(z) = \frac{1}{2\pi} \frac{1}{r} \frac{d}{dr} \left(\chi(r/\epsilon) + \ln \frac{r}{r_0} \frac{r}{\epsilon} \chi'(r/\epsilon) \right)$$

para $n = 2$, o que implica

$$\begin{aligned} \int_{\epsilon/2 \leq |z| \leq \epsilon} d^n z \Delta(G_0\chi_\epsilon)(z) &= \left(\chi(r/\epsilon) - \frac{1}{n-2} r \frac{\chi'(r/\epsilon)}{\epsilon} \right) \Big|_{\epsilon/2}^{\epsilon} \\ &= \left(\chi(1) - \chi(\frac{1}{2}) \right) - \frac{1}{n-2} \left(\chi'(1) - \frac{1}{2} \chi'(\frac{1}{2}) \right) \\ &= -1 \end{aligned}$$

para $n \neq 2$ (isto é, para $n > 2$ assim como para $n = 1$) e

$$\begin{aligned} \int_{\epsilon/2 \leq |z| \leq \epsilon} d^2 z \Delta(G_0\chi_\epsilon)(z) &= \left(\chi(r/\epsilon) + \ln \frac{r}{r_0} r \frac{\chi'(r/\epsilon)}{\epsilon} \right) \Big|_{\epsilon/2}^{\epsilon} \\ &= \left(\chi(1) - \chi(\frac{1}{2}) \right) + \left(\ln \frac{\epsilon}{r_0} \chi'(1) - \ln \frac{\epsilon}{2r_0} \chi'(\frac{1}{2}) \right) \\ &= -1 \end{aligned}$$

para $n = 2$.

- A integral de $|\Delta(G_0\chi_\epsilon)(z)|$ sobre a casca $\{z \in \mathbb{R}^n / \epsilon/2 \leq |z| \leq \epsilon\}$ satisfaz uma estimativa da forma

$$\int_{\epsilon/2 \leq |z| \leq \epsilon} d^n z |\Delta(G_0\chi_\epsilon)(z)| \leq C + C' \left| \ln \frac{\epsilon}{r_0} \right|,$$

onde C e C' são constantes não-negativas que independem de ϵ e tais que $C' = 0$ exceto quando $n = 2$.

De fato, pondo $s = r/\epsilon$, usamos novamente as fórmulas anteriores, concluindo que

$$\begin{aligned} & \int_{\epsilon/2 \leq |z| \leq \epsilon} d^n z |\Delta(G_0 \chi_\epsilon)(z)| \\ & \leq \int_{\epsilon/2}^{\epsilon} dr \left\{ \frac{n-3}{n-2} \frac{|\chi'(r/\epsilon)|}{\epsilon} + \frac{1}{n-2} \frac{r}{\epsilon} \frac{|\chi''(r/\epsilon)|}{\epsilon} \right\} \\ & = \int_{1/2}^1 ds \left\{ \frac{n-3}{n-2} |\chi'(s)| + \frac{1}{n-2} s |\chi''(s)| \right\} \\ & \leq C, \end{aligned}$$

para $n \neq 2$ (isto é, para $n > 2$ assim como para $n = 1$), com uma constante $C > 0$ que não depende de ϵ , e

$$\begin{aligned} & \int_{\epsilon/2 \leq |z| \leq \epsilon} d^2 z |\Delta(G_0 \chi_\epsilon)(z)| \\ & \leq \int_{\epsilon/2}^{\epsilon} dr \left\{ \left(2 + \left| \ln \frac{r}{r_0} \right| \right) \frac{|\chi'(r/\epsilon)|}{\epsilon} + \left| \ln \frac{r}{r_0} \right| \frac{r}{\epsilon} \frac{|\chi''(r/\epsilon)|}{\epsilon} \right\} \\ & = \int_{1/2}^1 ds \left\{ \left(2 + \left| \ln \frac{\epsilon s}{r_0} \right| \right) |\chi'(s)| + \left| \ln \frac{\epsilon s}{r_0} \right| s |\chi''(s)| \right\} \\ & \leq C + C' \left| \ln \frac{\epsilon}{r_0} \right|, \end{aligned}$$

para $n = 2$, com constantes $C, C' > 0$ que não dependem de ϵ .

Logo,

$$\begin{aligned} |\Delta u_\epsilon(x) - f(x)| & = \left| - \int_{\epsilon/2 \leq |z| \leq \epsilon} d^n z f(x-z) \Delta(G_0 \chi_\epsilon)(z) \right. \\ & \quad \left. + f(x) \int_{\epsilon/2 \leq |z| \leq \epsilon} d^n z \Delta(G_0 \chi_\epsilon)(z) \right| \\ & \leq \int_{\epsilon/2 \leq |z| \leq \epsilon} d^n z |f(x-z) - f(x)| |\Delta(G_0 \chi_\epsilon)(z)| \\ & \leq \sup_{|z| \leq \epsilon} |f(x-z) - f(x)| \int_{\epsilon/2 \leq |z| \leq \epsilon} d^n z |\Delta(G_0 \chi_\epsilon)(z)| \\ & \leq \sup_{|\tilde{z}| \leq \epsilon} |f'(\tilde{z})| \epsilon \left(C + C' \left| \ln \frac{\epsilon}{r_0} \right| \right) \end{aligned}$$

(sendo que $C' = 0$ para $n \neq 2$), onde usamos, no último passo, o teorema do valor médio para f . Tomando o limite $\epsilon \rightarrow 0$, obtemos finalmente $\Delta u(x) = f(x)$. \square

2.1.3 Funções de Green

Uma das origens da noção de função de Green é a observação de que, dado um domínio Ω em \mathbb{R}^n e uma função $f \in C^1(\mathbb{R}^n)$ com suporte compacto contido em Ω , as fórmulas (2.27) e (2.28) continuam valendo, pelo menos, dentro de Ω – isto é, continuam fornecendo uma solução $u \in C^2(\mathbb{R}^n)$ da equação de Poisson $\Delta u = f$ em Ω – se nelas substituirmos o potencial de Coulomb $G_0(x - y)$ por uma outra “função de Green”, isto é, por uma função $G(x, y)$ que difere de $G_0(x - y)$ por alguma função em $C^2(\mathbb{R}^n)$ que é harmônica sobre Ω em relação à variável y . A liberdade na escolha desta função harmônica pode ser utilizada para fixar o comportamento de G sobre a fronteira $\partial\Omega$ de Ω de tal forma que ela serve para resolver problemas de fronteira para o operador de Laplace em Ω , inclusive os de Dirichlet e de Neumann. O teorema básico que permite chegar a essas conclusões é uma modificação da fórmula de Poisson conhecida como a terceira identidade de Green.

Para formalizar estas idéias, consideramos daqui em diante o potencial de Coulomb como função de dois argumentos, dependendo apenas da sua diferença. Para tanto, definimos

$$(\mathbb{R}^n)^2 = \mathbb{R}^n \times \mathbb{R}^n \quad , \quad (\mathbb{R}^n)_{\text{md}}^2 = \{(x, y) \in \mathbb{R}^n \times \mathbb{R}^n \mid x \neq y\} \quad , \quad (2.35)$$

e mais geralmente, para qualquer domínio Ω em \mathbb{R}^n ,

$$\begin{aligned} \Omega^2 &= \Omega \times \Omega \quad , \quad \Omega_{\text{md}}^2 = \{(x, y) \in \Omega \times \Omega \mid x \neq y\} \quad , \\ \bar{\Omega}^2 &= \bar{\Omega} \times \bar{\Omega} \quad , \quad \bar{\Omega}_{\text{md}}^2 = \{(x, y) \in \bar{\Omega} \times \bar{\Omega} \mid x \neq y\} \quad , \end{aligned} \quad (2.36)$$

onde o sufixo “md” significa “menos a diagonal”. Também utilizaremos a notação ∇_x, Δ_x e ∇_y, Δ_y para denotar o operador vetorial ∇ e o operador Δ de Laplace formados por derivadas parciais em relação às variáveis x_1, \dots, x_n e às variáveis y_1, \dots, y_n , respectivamente:

$$\nabla_x = \left(\frac{\partial}{\partial x_1}, \dots, \frac{\partial}{\partial x_n} \right) \quad , \quad \nabla_y = \left(\frac{\partial}{\partial y_1}, \dots, \frac{\partial}{\partial y_n} \right) \quad . \quad (2.37)$$

$$\Delta_x = \sum_{i=1}^n \frac{\partial^2}{\partial x_i^2} \quad , \quad \Delta_y = \sum_{i=1}^n \frac{\partial^2}{\partial y_i^2} \quad . \quad (2.38)$$

Com essa notação, adotaremos a seguinte

Definição 2.3 *A função de Green coulombiana é a função $G_0 \in C^\infty((\mathbb{R}^n)_{\text{md}}^2)$ definida por $G_0(x, y) = G_0(x - y)$, ou seja,*

$$G_0(x, y) = \begin{cases} -\frac{1}{(n-2) \text{vol}(S^{n-1})} \frac{1}{|x-y|^{n-2}} & (n > 2) \\ \frac{1}{2\pi} \ln \frac{|x-y|}{r_0} & (n = 2) \\ \frac{1}{2} |x-y| & (n = 1) \end{cases} \quad . \quad (2.39)$$

Se Ω é um domínio em \mathbb{R}^n , uma **função de Green para o operador de Laplace em Ω** é uma função $G \in C^2(\overline{\Omega}_{\text{md}}^2)$ tal que $G - G_0 \in C^2(\overline{\Omega}^2)$ e, para todo $x \in \Omega$,

$$\Delta_y (G(x, y) - G_0(x, y)) = 0. \quad (2.40)$$

Veremos no próximo capítulo como esta definição pode ser reformulada na linguagem de distribuições, já utilizada no Capítulo 1.8.

Mostraremos agora qual é a fórmula que generaliza a representação integral (2.27) a domínios mais gerais do que \mathbb{R}^n e funções de Green mais gerais do que a coulombiana:

Teorema 2.6 (Terceira Identidade de Green): *Seja Ω um domínio limitado em \mathbb{R}^n cuja fronteira $\partial\Omega = \overline{\Omega} \setminus \Omega$ é uma hipersuperfície de classe C^1 em \mathbb{R}^n , e seja $G \in C^2(\overline{\Omega}_{\text{md}}^2)$ uma função de Green para o operador de Laplace em Ω . Se $u \in C^2(\overline{\Omega})$ é uma solução da equação de Poisson $\Delta u = f$ em Ω , então temos para todo $x \in \Omega$*

$$\begin{aligned} u(x) &= \int_{\Omega} d^n y f(y) G(x, y) \\ &\quad - \int_{\partial\Omega} d\sigma(y) \cdot (\nabla u(y) G(x, y) - u(y) \nabla_y G(x, y)). \end{aligned} \quad (2.41)$$

DEMONSTRAÇÃO: Seja $x \in \Omega$ fixo e seja $d(x, \partial\Omega)$ a distância de x à fronteira $\partial\Omega$ de Ω ; portanto, $B_\epsilon(x) \subset \Omega$ desde que $0 < \epsilon < d(x, \partial\Omega)$.

Observamos primeiro que a função $G(x, \cdot)$, que apresenta o mesmo tipo de singularidade no ponto x como a função $G_0(x, \cdot)$, é integrável sobre $\overline{\Omega}$ e que

$$\int_{\Omega} d^n y f(y) G(x, y) = \lim_{\epsilon \rightarrow 0} \int_{\Omega \setminus \overline{B}_\epsilon(x)} d^n y f(y) G(x, y).$$

Isto segue da estimativa

$$\begin{aligned} &\left| \int_{\Omega} d^n y f(y) G(x, y) - \int_{\Omega \setminus \overline{B}_\epsilon(x)} d^n y f(y) G(x, y) \right| \\ &= \left| \int_{B_\epsilon(x)} d^n y f(y) G(x, y) \right| \\ &= \left| \int_{B_\epsilon(x)} d^n y f(y) (G - G_0)(x, y) + \int_{B_\epsilon(x)} d^n y f(y) G_0(x, y) \right| \\ &\leq \int_{B_\epsilon(x)} d^n y |f(y)| |(G - G_0)(x, y)| + \int_{B_\epsilon(x)} d^n y |f(y)| |G_0(x, y)| \\ &\leq \sup_{y \in B_\epsilon(x)} |f(y)| \left\{ \sup_{y \in B_\epsilon(x)} |(G - G_0)(x, y)| \text{ vol}(B_\epsilon(x)) + g_n(\epsilon) \right\} \end{aligned}$$

(veja a equação (2.14), em conjunto com a equação (2.15)). Por outro lado, tendo em vista que $\Delta u = f$ e $\Delta_y G(x, y) = 0$ em $\Omega \setminus \bar{B}_\epsilon(x)$, temos pela segunda identidade de Green (B.45)

$$\begin{aligned}
& \int_{\Omega \setminus \bar{B}_\epsilon(x)} d^n y f(y) G(x, y) \\
&= \int_{\Omega \setminus \bar{B}_\epsilon(x)} d^n y (\Delta u(y) G(x, y) - u(y) \Delta_y G(x, y)) \\
&= \int_{\partial(\Omega \setminus \bar{B}_\epsilon(x))} d\sigma(y) \cdot (\nabla u(y) G(x, y) - u(y) \nabla_y G(x, y)) \\
&= \int_{\partial\Omega} d\sigma(y) \cdot (\nabla u(y) G(x, y) - u(y) \nabla_y G(x, y)) \\
&\quad - \int_{S_\epsilon(x)} d\sigma(y) \cdot (\nabla u(y) G(x, y) - u(y) \nabla_y G(x, y)) .
\end{aligned}$$

Precisamos então calcular a segunda integral de hipersuperfície, no limite $\epsilon \rightarrow 0$. Mostramos primeiro que

$$\lim_{\epsilon \rightarrow 0} \int_{S_\epsilon(x)} d\sigma(y) \cdot \nabla u(y) G(x, y) = 0 .$$

Isto segue da estimativa

$$\begin{aligned}
& \left| \int_{S_\epsilon(x)} d\sigma(y) \cdot \nabla u(y) G(x, y) \right| \\
&= \left| \int_{S_\epsilon(x)} d\sigma(y) \cdot \nabla u(y) (G - G_0)(x, y) \right. \\
&\quad \left. + \int_{S_\epsilon(x)} d\sigma(y) \cdot \nabla u(y) G_0(x, y) \right| \\
&\leq \int_{S_\epsilon(x)} d\sigma(y) |\nabla u(y)| |(G - G_0)(x, y)| \\
&\quad + \left| G_0(\epsilon) \int_{S_\epsilon(x)} d\sigma(y) \cdot \nabla u(y) \right| \\
&\leq \int_{S_\epsilon(x)} d\sigma(y) |\nabla u(y)| |(G - G_0)(x, y)| \\
&\quad + \left| G_0(\epsilon) \int_{\bar{B}_\epsilon(x)} d^n y f(y) \right| \\
&\leq \text{vol}(S^{n-1}) \sup_{|y-x|=\epsilon} (|\nabla u(y)| |(G - G_0)(x, y)|) \epsilon^{n-1} \\
&\quad + \frac{1}{n} \text{vol}(S^{n-1}) \sup_{|y-x| \leq \epsilon} |f(y)| \epsilon^n |G_0(\epsilon)| ,
\end{aligned}$$

onde no penúltimo passo, foi utilizada o teorema da divergência para a função u na bola $B_\epsilon(x)$, em conjunto com a equação $\Delta u = f$. Finalmente, mostramos que

$$\lim_{\epsilon \rightarrow 0} \int_{S_\epsilon(x)} d\sigma(y) \cdot u(y) \nabla_y G(x, y) = u(x) .$$

Isto segue da estimativa

$$\begin{aligned} & \left| \int_{S_\epsilon(x)} d\sigma(y) \cdot u(y) \nabla_y G(x, y) - u(x) \right| \\ &= \left| \int_{S_\epsilon(x)} d\sigma(y) \cdot u(y) \nabla_y (G - G_0)(x, y) \right. \\ & \quad \left. + \int_{S_\epsilon(x)} d\sigma(y) \cdot u(y) \nabla_y G_0(x, y) - u(x) \int_{S_\epsilon(x)} d\sigma(y) \cdot \nabla_y G_0(x, y) \right| \\ &\leq \int_{S_\epsilon(x)} d\sigma(y) |u(y)| |\nabla_y (G - G_0)(x, y)| \\ & \quad + \int_{S_\epsilon(x)} d\sigma(y) |u(y) - u(x)| |\nabla_y G_0(x, y)| \\ &\leq \text{vol}(S^{n-1}) \sup_{|y-x|=\epsilon} (|u(y)| |\nabla_y (G - G_0)(x, y)|) \epsilon^{n-1} \\ & \quad + \sup_{|y-x|=\epsilon} |u(y) - u(x)| , \end{aligned}$$

onde utilizamos a normalização (2.9). □

A terceira identidade de Green motiva a seguinte

Definição 2.4 *Seja Ω um domínio em \mathbb{R}^n . Uma função de Green $G_D \in C^2(\overline{\Omega}_{\text{md}}^2)$ para o operador de Laplace em Ω é chamada **função de Green para o problema de Dirichlet** em Ω se*

$$G_D(x, y) = 0 \quad \text{para } y \in \partial\Omega . \quad (2.42)$$

*De forma análoga, uma função de Green $G_N \in C^2(\overline{\Omega}_{\text{md}}^2)$ para o operador de Laplace em Ω é chamada **função de Green para o problema de Neumann** em Ω se*

$$\mathbf{n}(y) \cdot \nabla_y G_N(x, y) = \frac{1}{\text{vol}(\partial\Omega)} \quad \text{para } y \in \partial\Omega . \quad (2.43)$$

Observamos que seria inconsistente exigir que a expressão do lado esquerdo da equação (2.43) se anule, pois basta aplicar a terceira identidade de Green (2.41)

com $u = 1$ e $f = 0$ para concluir que se $G \in C^2(\overline{\Omega}_{\text{md}}^2)$ é uma função de Green para o operador de Laplace em Ω qualquer, temos

$$\int_{\partial\Omega} d\sigma(y) \cdot \nabla_y G(x, y) = 1. \quad (2.44)$$

Notamos também que no caso das funções de Green para o problema de Dirichlet ou de Neumann, a terceira identidade de Green se reduz a

$$u(x) = \int_{\Omega} d^n y f(y) G_D(x, y) + \int_{\partial\Omega} d\sigma(y) \cdot u(y) \nabla_y G_D(x, y) \quad (2.45)$$

no primeiro caso e

$$u(x) = \int_{\Omega} d^n y f(y) G_N(x, y) - \int_{\partial\Omega} d\sigma(y) \cdot \nabla u(y) G_N(x, y) + \langle u \rangle_{\partial\Omega} \quad (2.46)$$

no segundo caso, onde

$$\langle u \rangle_{\partial\Omega} = \frac{1}{\text{vol}(\partial\Omega)} \int_{\partial\Omega} d\sigma(y) u(y) \quad (2.47)$$

representa a média da função u sobre a fronteira $\partial\Omega$ de Ω , fornecendo uma solução explícita, em termos de fórmulas integrais, tanto do problema de Dirichlet (2.1),

$$u(x) = \int_{\Omega} d^n y f(y) G_D(x, y) + \int_{\partial\Omega} d\sigma(y) \cdot g_D(y) \nabla_y G_D(x, y), \quad (2.48)$$

como do problema de Neumann (2.2),

$$u(x) = \int_{\Omega} d^n y f(y) G_N(x, y) - \int_{\partial\Omega} d\sigma(y) g_N(y) G_N(x, y) + \langle u \rangle_{\partial\Omega}. \quad (2.49)$$

Estas equações constituem uma generalização da fórmula de Poisson (2.27) pois incorporam, além da contribuição provida da fonte f e representada pela primeira integral “de volume” (que no caso da fórmula de Poisson se reduz à convolução da fonte f com o potencial de Coulomb G_0), uma contribuição adicional provida dos dados g_D ou g_N prescritos na fronteira e representada pela segunda integral “de superfície”. Obviamente, a fórmula de Poisson é um caso especial, obtido quando (a) supomos que f tenha suporte compacto contido em Ω , (b) procuramos a solução u com $g_D = 0$ ou $g_N = 0$ e (c) tomamos o limite em que $\Omega = \mathbb{R}^n$.

Sendo assim, concluímos que, conforme prometido anteriormente, a resolução dos problemas de Dirichlet e de Neumann se reduz à resolução de um único problema desta natureza, para uma escolha específica da fonte e da condição de fronteira: basta calcular a função de Green correspondente para obter a solução geral, segundo as equações (2.48) ou (2.49). Invertendo a lógica anterior, podemos até usar estas fórmulas para definir a solução geral e assim provar a sua existência,³

³Note que o Teorema 2.6 como tal não inclui nenhuma afirmação sobre a existência da solução u , sendo que essa faz parte das suas hipóteses e não das suas conclusões.

isto é, provar o Teorema 2.2 assim como os teoremas análogos para o problema de Dirichlet não-homogêneo e o problema de Neumann homogêneo ou não-homogêneo, desde que a existência da função de Green correspondente seja garantida e as suas propriedades sejam suficientemente bem especificadas para podermos fazer as estimativas necessárias nas integrais pertinentes. Essa demonstração constitui uma versão mais sofisticada da demonstração do Teorema 2.5 dada acima, e não é nossa intenção apresentá-la aqui.

Isso posto, resta a questão como calcular essas funções de Green para um domínio Ω dado. Ressalta-se que para domínios gerais, trata-se de uma tarefa difícil, muitas vezes praticamente inviável. No entanto, existem alguns métodos que permitem tratar de casos especiais, entre eles (a seguir, nos restringiremos ao caso das funções de Green para o problema de Dirichlet):

- Métodos da teoria das funções analíticas no plano complexo:

é fácil provar que se Ω e Ω' são dois domínios no plano complexo, cada um com fronteira regular, e com funções de Green-Dirichlet G_D para Ω e G'_D para Ω' , e se $f : \Omega \rightarrow \Omega'$ é uma transformação conforme que se estende a um homeomorfismo $\bar{f} : \bar{\Omega} \rightarrow \bar{\Omega}'$, então⁴

$$G'_D(x', y') = G_D(x, y) \quad \text{onde } x' = f(x), y' = f(y).$$

Este fato permite calcular funções de Green-Dirichlet para domínios mais complicadas a partir de funções de Green-Dirichlet para domínios mais simples, lançando mão de transformações conformes entre elas.

- Métodos de reflexão na fronteira:

é fácil provar que uma combinação linear de funções de Green coulombianas, da forma

$$G(x, y) = G_0(x, y) + \sum_j \lambda_j G_0(x_j, y) \quad \text{para } x, y \in \bar{\Omega}, x \neq y, x_j \neq y$$

é uma função de Green para o operador de Laplace em Ω , desde que os pontos x_j (que podem depender do ponto x) não pertençam ao domínio Ω e, exceto possivelmente quando x está na fronteira $\partial\Omega$, nem ao seu fecho $\bar{\Omega}$:

$$x_j \notin \bar{\Omega} \quad , \quad x \in \Omega \implies x_j \notin \bar{\Omega}.$$

Este fato permite encontrar, para alguns domínios especiais com alto grau de simetria, λ_j e x_j tais que $G(x, y)$ se anula quando $y \in \partial\Omega$, e assim este G será igual a G_D .

A maneira mais didática de explicar estas técnicas é através de exemplos. Aqui, queremos discutir dois exemplos da aplicação do método de reflexão na fronteira:

⁴Tecnicamente, as hipóteses exatas são as seguintes: $\partial\Omega$ e $\partial\Omega'$ são curvas de classe C^1 no plano complexo, $f : \Omega \rightarrow \Omega'$ é bijetora, holomorfa e com inversa $f^{-1} : \Omega' \rightarrow \Omega$ holomorfa e $\bar{f} : \bar{\Omega} \rightarrow \bar{\Omega}'$ é bijetora, contínua e com inversa $\bar{f}^{-1} : \bar{\Omega}' \rightarrow \bar{\Omega}$ contínua.

o primeiro é tecnicamente o mais simples de todos, mas apresenta a desvantagem de que o domínio correspondente não é limitado, enquanto que o segundo é um pouco mais complicado, mas fornece a função de Green-Dirichlet para o domínio limitado mais simples possível: a bola.

Exemplo 2.2 (Reflexão num hiperplano): Seja Ω um semi-espaço em \mathbb{R}^n , então $\partial\Omega$ é um hiperplano em \mathbb{R}^n , e denotamos por σ a reflexão ortogonal neste hiperplano, isto é, σ é a transformação (linear) involutiva⁵ de \mathbb{R}^n que é a identidade sobre o subespaço $(n-1)$ -dimensional $\partial\Omega$ de \mathbb{R}^n e menos a identidade sobre o subespaço unidimensional $(\partial\Omega)^\perp$ de \mathbb{R}^n que é o seu complemento ortogonal. Notando que para $x \in \Omega$, temos $\sigma x \notin \Omega$, concluímos que

$$G_0(x, y) + \lambda G_0(\sigma x, y)$$

é uma função de Green para o operador de Laplace em Ω . Falta apenas determinar o fator λ tal que esta esta combinação se anula quando $y \in \partial\Omega$, i.e., quando $\sigma y = y$. Mas como $G_0(x, y)$ depende apenas de $|x - y|$ e σ é isométrica ($|\sigma z| = |z|$), é óbvio que a solução é $\lambda = -1$, e portanto a função de Green-Dirichlet para este domínio é

$$G_D(x, y) = G_0(|x - y|) - G_0(|\sigma x - y|), \quad (2.50)$$

com G_0 o potencial de Coulomb dado pela equação (2.11), por exemplo.

Exemplo 2.3 (Reflexão numa esfera): Seja Ω a bola aberta em \mathbb{R}^n de raio R em torno da origem, então $\partial\Omega$ é a esfera em \mathbb{R}^n de raio R em torno da origem, e denotamos por σ a reflexão conforme nesta esfera, isto é, σ é a transformação (não-linear) involutiva⁵ de $\mathbb{R}^n \setminus \{0\}$ dada por

$$\sigma x = \frac{R^2}{r^2} x \quad \text{para } x \in \mathbb{R}^n, x \neq 0,$$

onde $r = |x|$ denota a norma euclideana de x , como antes. Notando que para $x \in \Omega$, temos $\sigma x \notin \Omega$, concluímos que

$$G_0(x, y) + \lambda G_0(\sigma x, y)$$

é uma função de Green para o operador de Laplace em Ω . Falta apenas determinar o fator λ (que aqui dependerá de x) tal que esta esta combinação se anula quando $y \in \partial\Omega$, i.e., quando $|y| = R$, ou ainda, quando $\sigma y = y$. Mas neste caso, temos

$$\begin{aligned} |\sigma x - y|^2 &= \left| \frac{R^2}{|x|^2} x - y \right|^2 = \left(\frac{R^2}{|x|^2} \right)^2 |x|^2 - 2 \frac{R^2}{|x|^2} x \cdot y + |y|^2 \\ &= \frac{R^2}{|x|^2} \left(R^2 - 2x \cdot y + |x|^2 \frac{|y|^2}{R^2} \right) \\ &= \frac{R^2}{|x|^2} \left(|y|^2 - 2x \cdot y + |x|^2 \right) = \frac{R^2}{|x|^2} |x - y|^2. \end{aligned}$$

⁵Uma transformação é chamada involutiva se ela coincide com o seu próprio inverso.

e como $G_0(x, y)$ depende apenas de $|x - y|$, podemos reescrever o resultado final de tal maneira que a distinção de casos da equação (2.11) se torna desnecessária, chegando à conclusão de que a função de Green-Dirichlet para este domínio é

$$G_D(x, y) = G_0(|x - y|) - G_0\left(\frac{|x|}{R} |\sigma x - y|\right), \quad (2.51)$$

com G_0 o potencial de Coulomb dado pela equação (2.11), por exemplo.

Notamos que a função de Green-Dirichlet (2.51) para a bola acima deduzida também é chamada o **núcleo de Poisson**. Como já foi mencionado anteriormente, ela pode ser combinada com a equação (2.48) para provar, por construção explícita, a existência da solução geral do problema de Dirichlet, homogêneo ou não-homogêneo, no mais simples dos domínios limitados: a bola. A fórmula explícita que resulta desta combinação é devida a Poisson, o que explica a nomenclatura.

2.1.4 Teoremas da média esférica e de regularidade

Como corolário da terceira identidade de Green, podemos deduzir o teorema da média esférica para funções harmônicas. Para a definição do conceito da média esférica M_u de uma função u , veja o Apêndice B (equações (B.46) e (B.47)).

Teorema 2.7 (Teorema da média esférica para funções harmônicas): *Sejam Ω um domínio em \mathbb{R}^n e $u \in C(\Omega)$. Se a função u é de classe C^2 e é harmônica, a sua média esférica M_u independe da variável radial, i.e., vale*

$$M_u(x, r) = u(x) \quad \text{para } x \in \Omega, \quad 0 < r < d(x, \partial\Omega). \quad (2.52)$$

Reciprocamente, se a média esférica M_u da função u satisfaz esta condição, então u é de classe C^∞ e é harmônica.

DEMONSTRAÇÃO: Supondo primeiro que u seja de classe C^2 , notamos que $M_u(x, 0) = u(x)$ e que vale $\partial M_u / \partial r = 0$ se e somente se $\Delta\phi = 0$ (veja a equação (B.49) na demonstração da Proposição B.2). Portanto, para completar a demonstração, precisamos mostrar apenas que se $u \in C(\Omega)$ é tal que a sua média esférica M_u satisfaz a equação (2.52), então necessariamente $u \in C^\infty(\Omega)$. Inicialmente, observamos que é suficiente provar esta afirmação sob a hipótese adicional de que Ω é limitado e que $u \in C(\bar{\Omega})$. (No caso geral, recobrimos Ω por uma família de subdomínios limitados com fecho contido em Ω e concluímos que se a afirmação vale para cada um destes subdomínios, ela vale para Ω também.) A idéia da demonstração reside a) na introdução de uma regularização da função u definida e contínua sobre Ω através de uma família de funções u_ϵ definidas e de classe C^∞ sobre subdomínios apropriados Ω_ϵ de Ω ($\epsilon > 0$) tais que

$$\bigcup_{\epsilon > 0} \Omega_\epsilon = \Omega$$

e b) na observação de que se u satisfaz a equação (B.49), então

$$u|_{\Omega_\epsilon} = u_\epsilon .$$

Para efetuar a desejada regularização, utilizamos as funções reguladoras $\chi \in C^\infty(\mathbb{R})$ e $\chi_\epsilon \in C^\infty(\mathbb{R}^n)$ já introduzidas anteriormente (veja as equações (2.18)-(2.25)), e definimos

$$\begin{aligned} \Omega_\epsilon &= \{x \in \Omega \mid d(x, \partial\Omega) > \epsilon\} , \\ u_\epsilon(x) &= \frac{1}{l_n \epsilon^n} \int d^n y u(y) \chi_\epsilon(x-y) \quad \text{para } x \in \Omega_\epsilon . \end{aligned}$$

Intuitivamente, Ω_ϵ é obtido por remoção de uma camada de espessura ϵ em torno da fronteira $\partial\Omega$ de Ω . Sendo que a função $(x, y) \mapsto \chi_\epsilon(x-y)$ é de classe C^∞ sobre $\mathbb{R}^n \times \mathbb{R}^n$ e que para $x \in \Omega_\epsilon$, o suporte da função $\chi_\epsilon(x-y)$ (como função de y , para x fixo) é contido em Ω e portanto a integral na definição de u_ϵ se estende tão somente sobre o compacto fixo $\tilde{\Omega}$, a função u_ϵ será de classe C^∞ sobre Ω_ϵ .

Isso posto, calculamos, para $x \in \Omega_\epsilon$,

$$\begin{aligned} u_\epsilon(x) &= \frac{1}{l_n \epsilon^n} \int_{B_\epsilon(x)} d^n y u(y) \chi_\epsilon(x-y) \\ &= \frac{1}{l_n \epsilon^n} \int_0^\epsilon dr r^{n-1} \chi(r/\epsilon) \int_{S_r(x)} d\sigma(y) u(y) \\ &= \frac{\text{vol}(S^{n-1})}{l_n \epsilon^n} \int_0^\epsilon dr r^{n-1} \chi(r/\epsilon) M_u(x, r) , \end{aligned}$$

e observamos que se u satisfaz a equação (2.52), segue $u_\epsilon(x) = u(x)$, devido à equação (2.25). □

Combinando o teorema da média esférica com a fórmula de Poisson, obtemos o seguinte

Teorema 2.8 (Teorema de regularidade): *Sejam Ω um domínio em \mathbb{R}^n , f uma função contínua sobre Ω e u uma função de classe C^2 sobre Ω que é solução da equação de Poisson $\Delta u = f$. Então se f é de classe C^∞ , u também é de classe C^∞ .*

DEMONSTRAÇÃO: Em particular, o teorema afirma que funções harmônicas são automaticamente de classe C^∞ , o que é uma consequência imediata do Teorema 2.7. Para tratar do caso geral da equação de Poisson com uma fonte f de classe C^∞ não-trivial, observamos que é suficiente provar que cada ponto de Ω possui uma vizinhança aberta contida em Ω sobre a qual a solução u é de classe C^∞ . Sejam então x um ponto qualquer de Ω e $\epsilon > 0$ tal que a bola fechada $\bar{B}_\epsilon(x)$

de raio ϵ em torno de x esteja contido em Ω , e seja $f_{x,\epsilon}$ a função sobre \mathbb{R}^n definida por

$$f_{x,\epsilon}(y) = \begin{cases} \chi_\epsilon(y-x)f(y) & \text{para } y \in \Omega \\ 0 & \text{para } y \notin \Omega \end{cases},$$

onde, mais uma vez, $\chi \in C^\infty(\mathbb{R})$ e $\chi_\epsilon \in C^\infty(\mathbb{R}^n)$ são as funções reguladoras já introduzidas anteriormente (veja as equações (2.18)-(2.25)). Obviamente, $f_{x,\epsilon}$ é de classe C^∞ e de suporte compacto, e portanto podemos considerar a função $u_{x,\epsilon}$ sobre \mathbb{R}^n definida pela fórmula de Poisson

$$u_{x,\epsilon}(y) = \int d^n z f_{x,\epsilon}(z) G_0(y-z).$$

Conforme o Teorema 2.5, $u_{x,\epsilon}$ também é de classe C^∞ . Mas sobre a bola aberta $B_{\epsilon/2}(x)$ de raio $\epsilon/2$ em torno de x , $f_{x,\epsilon}$ coincide com f ; logo $u - u_{x,\epsilon}$ é harmônica e portanto tanto $u - u_{x,\epsilon}$ como u são de classe C^∞ .

□

Para uma melhor apreciação deste teorema, parece útil mencionar que a mesma afirmação (propriedades de regularidade da fonte implicam em propriedades análogas de regularidade da solução) é válida para qualquer operador elíptico, enquanto que isso não é o caso para operadores hiperbólicos como o operador de ondas, por exemplo. De fato, é de se esperar que para equações de evolução, a regularidade da solução dependerá não apenas da regularidade da fonte mas também da regularidade da condição inicial. No caso da equação de ondas, isso se concretiza através do princípio de “propagação de singularidades”: singularidades na condição inicial resultam em singularidades na solução, já para a equação homogênea.

2.1.5 Singularidades removíveis

Um outro teorema útil na teoria das funções harmônicas é o seguinte:

Teorema 2.9 (*Singularidades removíveis de funções harmônicas*): *Sejam Ω um domínio em \mathbb{R}^n , $x_0 \in \Omega$ e u uma função de classe C^2 e harmônica sobre $\Omega \setminus \{x_0\}$. Se $n > 1$ e u é limitada, então u admite uma única extensão a uma função de classe C^2 e harmônica sobre Ω .*⁶

Entre outras coisas, este teorema implica que a singularidade na origem exibida pela função G_0 (o potencial de Coulomb) é típica e inevitável: sendo uma função harmônica fora da singularidade, ela seria trivial se não apresentasse um pólo – isto é, uma singularidade em torno da qual não é limitada. (Portanto, não existe, para $n > 1$, a opção de uma simples descontinuidade.) O teorema também fornece

⁶Obviamente, para $n = 1$, este teorema é falso: podemos juntar dois segmentos de funções da forma $f(x) = ax + b$ de forma descontínua, porém limitada.

uma demonstração imediata da afirmação do contra-exemplo de Zaremba (veja a discussão no início desta seção), pois implica que uma função u de classe C^2 e harmônica sobre \dot{B}_r^n (a bola aberta de raio r em \mathbb{R}^n em torno da origem, com a origem removida) e contínua sobre \bar{B}_r^n (a bola fechada de raio r em \mathbb{R}^n em torno da origem) é de fato harmônica sobre B_r^n (a bola aberta de raio r em \mathbb{R}^n em torno da origem). Portanto, segundo o princípio do máximo/mínimo, o seu valor na origem é limitado (superiormente e inferiormente) por seus valores em S_r^{n-1} (a esfera de raio r em \mathbb{R}^n); em particular, é impossível que uma tal função se anule sobre S_r^{n-1} sem se anular identicamente, inclusive na origem.

DEMONSTRAÇÃO: Primeiro, observamos que é suficiente demonstrar este teorema para o caso onde o domínio Ω é a bola aberta B_R^n em torno da origem e a função u dada é contínua sobre $\bar{B}_R^n \setminus \{0\}$. (Basta escolher $R < d(x_0, \partial\Omega)$, onde $d(x_0, \partial\Omega)$ é a distância de x_0 à fronteira $\partial\Omega$ de Ω e aplicar uma translação pelo vetor $-x_0$). Segundo, notamos que devido aos teoremas de unicidade e existência da solução clássica do problema de Dirichlet homogêneo (Teoremas 2.1 e 2.2) na bola, é suficiente demonstrar o teorema para o caso onde a função u dada se anula na esfera S_R^{n-1} . (De fato, suponha que a afirmação do teorema vale neste caso. Se u é uma função de classe C^2 e harmônica sobre $B_R^n \setminus \{0\}$, contínua sobre $\bar{B}_R^n \setminus \{0\}$, existe uma única função v de classe C^2 e harmônica sobre B_R^n , contínua sobre \bar{B}_R^n , tal que v coincide com u sobre S_R^{n-1} . Assim, a afirmação do teorema para u segue da afirmação do teorema para $u - v$.)

Seja então u uma função de classe C^2 e harmônica sobre $B_R^n \setminus \{0\}$, contínua e limitada sobre $\bar{B}_R^n \setminus \{0\}$, tal que $u = 0$ sobre S_R^{n-1} . Pomos

$$M_+ = \sup_{0 < |x| \leq R} u(x) \quad , \quad M_- = \inf_{0 < |x| \leq R} u(x)$$

Agora definimos, para todo $0 < \epsilon < R$, as cascas

$$\begin{aligned} C_{\epsilon,R}^n &= \{x \in \mathbb{R}^n / \epsilon < |x| < R\} \\ \bar{C}_{\epsilon,R}^n &= \{x \in \mathbb{R}^n / \epsilon \leq |x| \leq R\} \end{aligned}$$

e pomos

$$u_\epsilon^\pm(x) = M_\pm \frac{G_0(|x|) - G_0(R)}{G_0(\epsilon) - G_0(R)} .$$

A função u , apropriadamente restrita, e as funções u_ϵ^\pm são de classe C^2 e harmônicas sobre $C_{\epsilon,R}^n$, contínuas sobre $\bar{C}_{\epsilon,R}^n$ e tais que

$$u_\epsilon^\pm(x) = \begin{cases} 0 & \text{se } |x| = R \\ M_\pm & \text{se } |x| = \epsilon \end{cases} ,$$

Destarte,

$$\begin{aligned} u - u_\epsilon^+ &\leq 0 \quad \text{sobre } \partial C_{\epsilon,R}^n = S_\epsilon^{n-1} \cup S_R^{n-1} , \\ u - u_\epsilon^- &\geq 0 \quad \text{sobre } \partial C_{\epsilon,R}^n = S_\epsilon^{n-1} \cup S_R^{n-1} . \end{aligned}$$

Aplicando o princípio do máximo à função $u - u_\epsilon^+$ e o princípio do mínimo à função $u - u_\epsilon^-$, concluímos que vale $u_\epsilon^- \leq u \leq u_\epsilon^+$ sobre $\bar{C}_{\epsilon,R}^n$. Portanto, dado um ponto $x \in \bar{B}_R^n \setminus \{0\}$ qualquer, temos $u_\epsilon^-(x) \leq u(x) \leq u_\epsilon^+(x)$, desde que $0 < \epsilon < |x|$, e consequentemente

$$\lim_{\epsilon \rightarrow 0} u_\epsilon^-(x) \leq u(x) \leq \lim_{\epsilon \rightarrow 0} u_\epsilon^+(x).$$

Mas para $R > 0$ e $x \in \bar{B}_R^n \setminus \{0\}$ fixos, temos

$$\lim_{\epsilon \rightarrow 0} u_\epsilon^\pm(x) = \lim_{\epsilon \rightarrow 0} M_\pm \frac{1/|x|^{n-2} - 1/R^{n-2}}{1/\epsilon^{n-2} - 1/R^{n-2}} = 0$$

se $n > 2$, e

$$\lim_{\epsilon \rightarrow 0} u_\epsilon^\pm(x) = \lim_{\epsilon \rightarrow 0} M_\pm \frac{\ln(R/|x|)}{\ln(R/\epsilon)} = 0$$

se $n = 2$, enquanto que

$$\lim_{\epsilon \rightarrow 0} u_\epsilon^\pm(x) = \lim_{\epsilon \rightarrow 0} M_\pm \frac{|x| - R}{\epsilon - R} = M_\pm \left(1 - \frac{|x|}{R}\right)$$

se $n = 1$. Nos primeiros dois casos, segue que $u = 0$, enquanto que no último caso, não há como concluir nada. \square

2.2 O método de separação de variáveis

Um dos métodos mais tradicionais para encontrar soluções de equações diferenciais parciais é por separação de variáveis. Grosso modo, este método se baseia na ideia de procurar soluções u que, como funções de n variáveis y_1, \dots, y_n , digamos, fatoram no produto de n funções u_1, \dots, u_n de apenas uma variável cada uma, conforme

$$u(y_1, \dots, y_n) = u_1(y_1) \dots u_n(y_n). \quad (2.53)$$

Claramente, tais soluções são muito especiais. Porém, soluções mais gerais podem ser construídas formando combinações lineares, pois para equações diferenciais parciais lineares (homogêneas), uma combinação linear de tais soluções ainda é uma solução que, no entanto, não terá mais a mesma forma. E tomando limites, pode-se passar a considerar séries convergentes, ao invés de combinações lineares finitas. Finalmente, em certas situações e mediante uso de uma noção adequada de convergência, é possível provar que toda solução (dentro de uma classe suficientemente ampla de funções) admitirá tal expansão em série.

A limitação principal do método provém das condições de fronteira, sendo que ele só serve para resolver problemas de fronteira em domínios Ω muito simples, frequentemente caracterizados por um alto grau de simetria, e isso mediante

uma escolha adequada das variáveis y_1, \dots, y_n , adaptada à geometria de Ω . Por exemplo, se Ω for um retângulo ($n = 2$) ou, para n geral, um hipercubo, podemos usar as coordenadas cartesianas usuais x_1, \dots, x_n , enquanto que se Ω for um disco ($n = 2$) ou, para n geral, uma bola, temos que passar para coordenadas esféricas, e assim por diante. Sem tal adaptação das coordenadas empregadas, simplesmente não existirá nenhuma solução que fatora e ainda satisfaz condições do tipo Dirichlet ou Neumann, por exemplo, sobre a fronteira $\partial\Omega$ de Ω . Portanto, em função dessa dificuldade, o método torna-se inútil para domínios com geometria mais complicada. Mas mesmo assim, os exemplos em que o método funciona bem são suficientemente interessantes para merecerem um tratamento com um certo grau de detalhe.

Antes de entrarmos na discussão de exemplos, já podemos resumir a estratégia do método, que consiste de vários passos.

- *Passo 1*: Escolher um sistema de coordenadas y_1, \dots, y_n no domínio Ω que permite decompor sua fronteira $\partial\Omega$ em um número finito de “pedaços”, cada um dos quais é dado por uma equação simples do tipo $y_i = 0$, para algum valor de i .
- *Passo 2*: Reescrever o operador diferencial pertinente em termos das coordenadas y_1, \dots, y_n .
- *Passo 3*: Determinar as equações diferenciais ordinárias e as condições de fronteira satisfeitas por cada um dos fatores u_i , como função de y_i .
- *Passo 4*: Resolver essas equações, em termos de funções conhecidas.
- *Passo 5*: Formar combinações lineares e séries das soluções fatoráveis e determinar qual classe de soluções admite uma expansão em série deste tipo.

Naturalmente, este roteiro pode ser modificado para adaptá-lo a outras situações, dependendo das circunstâncias – principalmente quando o método funciona apenas parcialmente, ou seja, quando o problema em questão sugere uma separação parcial das variáveis, dividindo-as em blocos.

Um exemplo típico e razoavelmente geral desta situação ocorre no estudo de equações de evolução – mais especificamente, da equação de difusão assim como da equação de ondas – com condições de fronteira que independem do tempo.

Concretamente, considere a equação de difusão homogênea

$$\left(\frac{\partial}{\partial t} - \lambda \Delta\right) u = 0 \quad (2.54)$$

e a equação de ondas homogênea

$$\square u \equiv \left(\frac{1}{c^2} \frac{\partial^2}{\partial t^2} - \Delta\right) u = 0 \quad (2.55)$$

em $n + 1$ dimensões, em um domínio Ω da forma $I \times \Omega_0$ onde I é um intervalo aberto em \mathbb{R} (que pode ser finito ou semi-infinito ou mesmo a reta inteira) e Ω_0 é um domínio em \mathbb{R}^n . Neste caso, procuramos soluções u da forma

$$u(t, x) = v(t) w(x) \quad \text{para } t \in I, x \in \Omega_0, \quad (2.56)$$

notando que, neste caso, precisamos impor condições suplementares apenas sobre uma parte da fronteira $\partial\Omega$ de Ω – condições que se dividem naturalmente em condições de fronteira sobre a “parte espacial” $I \times \partial\Omega_0$ de $\partial\Omega$ e condições iniciais sobre uma hipersuperfície de Cauchy do tipo $\{t_0\} \times \Omega_0$: claramente, tais condições são compatíveis com a hipótese de fatoração formulada na equação (2.56). De fato, para funções u deste tipo, a equação de difusão (2.54) toma a forma

$$\dot{v}(t) w(x) - \lambda v(t) \Delta w(x) = 0,$$

enquanto que a equação de ondas (2.55) toma a forma

$$\frac{1}{c^2} \ddot{v}(t) w(x) - v(t) \Delta w(x) = 0,$$

de modo que no subconjunto aberto de Ω onde $u \neq 0$ (que também é um produto cartesiano, a saber, do subconjunto aberto de I em que $v \neq 0$ com o subconjunto aberto de Ω_0 em que $w \neq 0$), obtemos, após divisão por u ,

$$\frac{\dot{v}(t)}{v(t)} = \lambda \frac{\Delta w(x)}{w(x)}$$

para a equação de difusão e

$$\frac{1}{c^2} \frac{\ddot{v}(t)}{v(t)} = \frac{\Delta w(x)}{w(x)}$$

para a equação de ondas. Agora, nestas equações, o lado esquerdo é apenas uma função de t , enquanto que o lado direito é apenas uma função de x , e portanto ambos os lados tem que ser constantes, ou seja, existe uma constante $\alpha \in \mathbb{R}$ tal que

$$\Delta w(x) = \alpha w(x) \quad (2.57)$$

e ainda

$$\dot{v}(t) = \alpha \lambda v(t) \quad (2.58)$$

para a equação de difusão e

$$\ddot{v}(t) = \alpha c^2 v(t) \quad (2.59)$$

para a equação de ondas. Evidencia-se então que as condições iniciais fixam a função w , a menos de um fator constante $v(t_0)$, enquanto que as condições de fronteira fixam o comportamento desta função w na fronteira $\partial\Omega_0$ de Ω_0 .

Sendo assim, somos levados a estudar, como passo intermediário, o problema de determinar o “espectro” do operador de Laplace, ou seja, encontrar as soluções da equação de Laplace modificada (2.57), no domínio Ω_0 , com uma constante α

a ser determinada, sujeitas a condições de fronteira do tipo Dirichlet ou Neumann, digamos. Tipicamente, para domínios Ω_0 limitados, o “espectro” dos valores α permitidos será discreto e pode ser enumerado por inteiros. Também é interessante notar que, para condições de fronteira do tipo Dirichlet ou Neumann homogêneas, conforme as quais a solução, no primeiro caso, ou a derivada normal da solução, no segundo caso, se anula na fronteira $\partial\Omega_0$ de Ω_0 , a constante α não pode ser positiva, pois neste caso, a primeira identidade de Green (B.44) implica

$$\int_{\Omega_0} d^n x |\nabla u|^2 = - \int_{\Omega_0} d^n x u \Delta u = -\alpha \int_{\Omega_0} d^n x |u|^2$$

e portanto devemos ter $\alpha \leq 0$.

O exemplo mais simples é o seguinte:

Exemplo 2.4 Considere a equação de Laplace modificada (2.57) unidimensional no intervalo $\Omega_0 =]0, L[$, com condições de fronteira de Dirichlet homogêneas

$$w^{(D)}(0) = 0 = w^{(D)}(L), \quad (2.60)$$

ou condições de fronteira de Neumann homogêneas

$$w^{(N)'}(0) = 0 = w^{(N)'}(L). \quad (2.61)$$

Então a constante α deve, de fato, ser negativa, pois se ela fosse positiva, a solução da equação seria uma combinação linear das duas funções exponenciais $\exp(\pm\sqrt{\alpha}x)$, e se ela fosse igual a zero, seria a soma de uma função linear e uma função constante; assim, tanto a função w como a sua derivada w' se anulariam, no máximo, em apenas um ponto da reta, o que é incompatível com as condições de fronteira estipuladas nas equações (2.60) ou (2.61). Portanto, concluímos que $\alpha < 0$ e que tanto a função w como a sua derivada w' são combinações lineares das funções $\sin(\sqrt{-\alpha}x)$ e $\cos(\sqrt{-\alpha}x)$. Obviamente, a condição de fronteira em $x=0$ implica que só aparece a função $\sin(\sqrt{-\alpha}x)$ (Dirichlet) ou a função $\cos(\sqrt{-\alpha}x)$ (Neumann), e a condição de fronteira em $x=L$ implica que $\sqrt{-\alpha}L$ deve ser um múltiplo inteiro e, sem perda de generalidade, positivo, de π . Assim, obtemos uma família, parametrizada por um inteiro não-negativo k , de soluções da equação (2.57) satisfazendo as condições de fronteira (2.60) de Dirichlet e (2.61) de Neumann, respectivamente, definidas por

$$\begin{aligned} w_k^{(D)}(x) &= \sin(k\pi x/L) \quad (k = 1, 2, 3, \dots), \\ w_k^{(N)}(x) &= \cos(k\pi x/L) \quad (k = 0, 1, 2, \dots), \end{aligned} \quad (2.62)$$

e tais que o correspondente valor de α é

$$\alpha_k = -(k\pi/L)^2. \quad (2.63)$$

Finalmente, formando combinações lineares e passando ao limite de um número infinito de termos, obtemos uma família de soluções w da equação de Laplace

modificada (2.57) no intervalo (a valores reais ou complexas), satisfazendo as condições de fronteira (2.60) de Dirichlet e (2.61) de Neumann, respectivamente, da forma

$$w^{(D/N)}(x) = \sum_{k=0}^{\infty} a_k w_k^{(D/N)}(x), \quad (2.64)$$

onde entendemos que $w_k^{(D)} \equiv 0$ para $k=0$, com coeficientes a_k (reais ou complexos) que devem apresentar um decaimento assintótico (i.e., no limite em que $k \rightarrow \infty$) suficiente para garantir a convergência desta série de Fourier.

Passando à discussão das equações de difusão (2.54) e de ondas (2.55) unidimensionais no intervalo, começamos por escrever suas soluções fatoráveis, conforme a equação (2.56), resolvendo as equações (2.58) e (2.59), respectivamente, com os valores de α dados pela equação (2.63): o resultado é

$$\begin{aligned} u_k^{(D)}(t, x) &= \exp(-(k\pi/L)^2 \lambda t) \sin(k\pi x/L) \quad (k = 1, 2, 3, \dots), \\ u_k^{(N)}(t, x) &= \exp(-(k\pi/L)^2 \lambda t) \cos(k\pi x/L) \quad (k = 0, 1, 2, \dots), \end{aligned} \quad (2.65)$$

para a equação de difusão e

$$\begin{aligned} u_k^{(D,0)}(t, x) &= \cos(k\pi ct/L) \sin(k\pi x/L) \quad (k = 1, 2, 3, \dots), \\ u_k^{(D,1)}(t, x) &= \sin(k\pi ct/L) \sin(k\pi x/L) \quad (k = 1, 2, 3, \dots), \\ u_k^{(N,0)}(t, x) &= \cos(k\pi ct/L) \cos(k\pi x/L) \quad (k = 0, 1, 2, \dots), \\ u_k^{(N,1)}(t, x) &= \sin(k\pi ct/L) \cos(k\pi x/L) \quad (k = 0, 1, 2, \dots), \end{aligned} \quad (2.66)$$

para a equação de ondas. A solução geral é obtida formando combinações lineares e, no limite, séries da forma

$$\begin{aligned} u^{(D)}(t, x) &= \sum_{k=1}^{\infty} a_k \exp(-(k\pi/L)^2 \lambda t) \sin(k\pi x/L), \\ u^{(N)}(t, x) &= \sum_{k=0}^{\infty} a_k \exp(-(k\pi/L)^2 \lambda t) \cos(k\pi x/L), \end{aligned} \quad (2.67)$$

para a equação de difusão, com coeficientes a_k a serem determinados pela expansão de Fourier da condição inicial

$$\begin{aligned} u^{(D)}(0, x) &= \sum_{k=1}^{\infty} a_k \sin(k\pi x/L), \\ u^{(N)}(0, x) &= \sum_{k=0}^{\infty} a_k \cos(k\pi x/L), \end{aligned} \quad (2.68)$$

e

$$\begin{aligned} u^{(D)}(t, x) &= \sum_{k=1}^{\infty} \left(a_k \cos(k \pi c t / L) + b_k \sin(k \pi c t / L) \right) \sin(k \pi x / L) , \\ u^{(N)}(t, x) &= \sum_{k=0}^{\infty} \left(a_k \cos(k \pi c t / L) + b_k \sin(k \pi c t / L) \right) \cos(k \pi x / L) , \end{aligned} \quad (2.69)$$

para a equação de ondas, com coeficientes a_k e b_k a serem determinados pelas expansões de Fourier das condições iniciais

$$\begin{aligned} u^{(D)}(0, x) &= \sum_{k=1}^{\infty} a_k \sin(k \pi x / L) , \\ \dot{u}^{(D)}(0, x) &= \sum_{k=1}^{\infty} b_k (k \pi c / L) \sin(k \pi x / L) , \\ u^{(N)}(0, x) &= \sum_{k=0}^{\infty} a_k \cos(k \pi x / L) , \\ \dot{u}^{(D)}(0, x) &= \sum_{k=0}^{\infty} b_k (k \pi c / L) \cos(k \pi x / L) . \end{aligned} \quad (2.70)$$

Os coeficientes devem ser escolhidos de tal modo a garantir a convergência das séries de Fourier (2.68) e (2.70) para as condições iniciais em um sentido apropriado (por exemplo, na norma do supremo, o que corresponde à convergência uniforme, ou na norma L^2), e então podemos garantir a convergência das correspondentes séries de soluções (2.67) e (2.69) neste mesmo sentido, uniformemente em t .

Para resolver o mesmo problema de determinar o “espectro” do operador de Laplace em duas dimensões espaciais ($n = 2$) e para domínios Ω_0 simples tais como o retângulo e o disco, podemos novamente aplicar o método de separação de variáveis.

Exemplo 2.5 Considere a equação de Laplace modificada (2.57) no retângulo $\Omega_0 =]0, a[\times]0, b[$, digamos, com condições de fronteira de Dirichlet homogêneas

$$\begin{aligned} w^{(D)}(0, y) &= 0 = w^{(D)}(a, y) \quad (0 \leq y \leq b) , \\ w^{(D)}(x, 0) &= 0 = w^{(D)}(x, b) \quad (0 \leq x \leq a) , \end{aligned} \quad (2.71)$$

ou condições de fronteira de Neumann homogêneas

$$\begin{aligned} \frac{\partial w^{(N)}}{\partial x}(0, y) &= 0 = \frac{\partial w^{(N)}}{\partial x}(a, y) \quad (0 \leq y \leq b) , \\ \frac{\partial w^{(N)}}{\partial y}(x, 0) &= 0 = \frac{\partial w^{(N)}}{\partial y}(x, b) \quad (0 \leq x \leq a) . \end{aligned} \quad (2.72)$$

Neste caso, os primeiros dois passos são triviais, de modo que podemos prosseguir imediatamente ao

Passo 3: Suponha que a função w fatora, i.e., vale

$$w(x, y) = X(x)Y(y) .$$

Então a equação (2.57) assume a forma

$$X''(x)Y(y) + X(x)Y''(y) - \alpha X(x)Y(y) = 0 ,$$

de modo que no subconjunto aberto de Ω onde $w \neq 0$ (que também é um retângulo, a saber, o produto cartesiano do subconjunto aberto de $]0, a[$ em que $X \neq 0$ com o subconjunto aberto de $]0, b[$ em que $Y \neq 0$), obtemos, após divisão por w ,

$$\frac{X''(x)}{X(x)} + \frac{Y''(y)}{Y(y)} - \alpha = 0 ,$$

ou seja,

$$\frac{X''(x)}{X(x)} = \alpha - \frac{Y''(y)}{Y(y)} .$$

Agora, nesta equação, o lado esquerdo é apenas uma função de x , enquanto que o lado direito é apenas uma função de y , e portanto ambos os lados devem ser constantes, ou seja, existem constantes $\beta, \gamma \in \mathbb{R}$ com $\beta + \gamma = \alpha$ tais que

$$X''(x) = \beta X(x) , \quad Y''(y) = \gamma Y(y) .$$

Passo 4: Considerando cada uma dessas duas equações, podemos aplicar o mesmo procedimento do exemplo anterior e combinar os resultados para obter uma família, parametrizada por dois inteiros não-negativos k e l , de soluções da equação (2.57) fatoráveis e satisfazendo as condições de fronteira (2.71) de Dirichlet e (2.72) de Neumann, respectivamente, definidas por

$$\begin{aligned} w_{k,l,\square}^{(D)}(x, y) &= \sin(k\pi x/a) \sin(l\pi y/b) \quad (k, l = 1, 2, 3, \dots) , \\ w_{k,l,\square}^{(N)}(x, y) &= \cos(k\pi x/a) \cos(l\pi y/b) \quad (k, l = 0, 1, 2, \dots) , \end{aligned} \quad (2.73)$$

e tais que o correspondente valor de α é

$$\alpha_{k,l,\square} = -((k\pi/a)^2 + (l\pi/b)^2) . \quad (2.74)$$

Passo 5: Formando combinações lineares e passando ao limite de um número infinito de termos, obtemos uma família de soluções w da equação de Laplace modificada (2.57) no retângulo (a valores reais ou complexas), satisfazendo as condições de fronteira (2.71) de Dirichlet e (2.72) de Neumann, respectivamente, da forma

$$w_{\square}^{(D/N)}(x, y) = \sum_{k,l=0}^{\infty} a_{k,l} w_{k,l,\square}^{(D/N)}(x, y) , \quad (2.75)$$

onde entendemos que $w_{k,l,\square}^{(D)} \equiv 0$ para $k = 0$ ou $l = 0$, com coeficientes $a_{k,l}$ (reais ou complexos) que devem apresentar um decaimento assintótico (i.e., no limite em que $k \rightarrow \infty$ e/ou $l \rightarrow \infty$) suficiente para garantir a convergência desta série de Fourier.

Exemplo 2.6 Considere a equação de Laplace modificada (2.57) no disco de raio a em torno da origem, digamos, e, mais uma vez, com condições de fronteira de Dirichlet ou Neumann homogêneas.

Passo 1: Introduzimos coordenadas polares no plano, escrevendo

$$x = r \cos \varphi, \quad y = r \sin \varphi,$$

com $0 < r < \infty$ ($r = \sqrt{x^2 + y^2}$) e $0 < \varphi < 2\pi$. A matriz jacobiana desta transformação de coordenadas é dada por

$$\begin{pmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \varphi} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \varphi} \end{pmatrix} = \begin{pmatrix} \cos \varphi & -r \sin \varphi \\ \sin \varphi & r \cos \varphi \end{pmatrix} = \begin{pmatrix} x/r & -y \\ y/r & x \end{pmatrix},$$

que tem determinante igual a r e inversa igual a

$$\begin{pmatrix} \frac{\partial r}{\partial x} & \frac{\partial r}{\partial y} \\ \frac{\partial \varphi}{\partial x} & \frac{\partial \varphi}{\partial y} \end{pmatrix} = \begin{pmatrix} \cos \varphi & \sin \varphi \\ -\sin \varphi/r & \cos \varphi/r \end{pmatrix} = \begin{pmatrix} x/r & y/r \\ -y/r^2 & x/r^2 \end{pmatrix}.$$

Passo 2: Notando que

$$\begin{aligned} \frac{\partial w}{\partial x} &= \frac{\partial w}{\partial r} \frac{\partial r}{\partial x} + \frac{\partial w}{\partial \varphi} \frac{\partial \varphi}{\partial x} = \cos \varphi \frac{\partial w}{\partial r} - \frac{\sin \varphi}{r} \frac{\partial w}{\partial \varphi}, \\ \frac{\partial w}{\partial y} &= \frac{\partial w}{\partial r} \frac{\partial r}{\partial y} + \frac{\partial w}{\partial \varphi} \frac{\partial \varphi}{\partial y} = \sin \varphi \frac{\partial w}{\partial r} + \frac{\cos \varphi}{r} \frac{\partial w}{\partial \varphi}, \end{aligned}$$

e portanto

$$\begin{aligned} \frac{\partial^2 w}{\partial x^2} &= \cos \varphi \frac{\partial}{\partial r} \left(\cos \varphi \frac{\partial w}{\partial r} - \frac{\sin \varphi}{r} \frac{\partial w}{\partial \varphi} \right) - \frac{\sin \varphi}{r} \frac{\partial}{\partial \varphi} \left(\cos \varphi \frac{\partial w}{\partial r} - \frac{\sin \varphi}{r} \frac{\partial w}{\partial \varphi} \right) \\ &= \cos^2 \varphi \frac{\partial^2 w}{\partial r^2} + \frac{2 \cos \varphi \sin \varphi}{r^2} \frac{\partial w}{\partial \varphi} - \frac{2 \cos \varphi \sin \varphi}{r} \frac{\partial^2 w}{\partial r \partial \varphi} \\ &\quad + \frac{\sin^2 \varphi}{r} \frac{\partial w}{\partial r} + \frac{\sin^2 \varphi}{r^2} \frac{\partial^2 w}{\partial \varphi^2} \\ \frac{\partial^2 w}{\partial y^2} &= \sin \varphi \frac{\partial}{\partial r} \left(\sin \varphi \frac{\partial w}{\partial r} + \frac{\cos \varphi}{r} \frac{\partial w}{\partial \varphi} \right) + \frac{\cos \varphi}{r} \frac{\partial}{\partial \varphi} \left(\sin \varphi \frac{\partial w}{\partial r} + \frac{\cos \varphi}{r} \frac{\partial w}{\partial \varphi} \right) \\ &= \sin^2 \varphi \frac{\partial^2 w}{\partial r^2} - \frac{2 \cos \varphi \sin \varphi}{r^2} \frac{\partial w}{\partial \varphi} + \frac{2 \cos \varphi \sin \varphi}{r} \frac{\partial^2 w}{\partial r \partial \varphi} \\ &\quad + \frac{\cos^2 \varphi}{r} \frac{\partial w}{\partial r} + \frac{\cos^2 \varphi}{r^2} \frac{\partial^2 w}{\partial \varphi^2} \end{aligned}$$

concluimos que o operador de Laplace em coordenadas polares no plano tem a forma

$$\Delta w = \frac{\partial^2 w}{\partial r^2} + \frac{1}{r} \frac{\partial w}{\partial r} + \frac{1}{r^2} \frac{\partial^2 w}{\partial \varphi^2}. \quad (2.76)$$

Além disso, as condições de fronteira de Dirichlet e de Neumann assumem a forma

$$w(a, \varphi) = 0, \quad (2.77)$$

e

$$\frac{\partial w}{\partial r}(a, \varphi) = 0, \quad (2.78)$$

respectivamente.

Passo 3: Suponha que a função w fatora, i.e., vale

$$w(r, \varphi) = R(r) \Phi(\varphi).$$

Então conforme a equação (2.76), a equação (2.57) assume a forma

$$R''(r) \Phi(\varphi) + \frac{1}{r} R'(r) \Phi(\varphi) + \frac{1}{r^2} R(r) \Phi''(\varphi) - \alpha R(r) \Phi(\varphi) = 0,$$

de modo que no subconjunto aberto do disco onde $w \neq 0$, obtemos, após divisão por w e multiplicação por r^2 ,

$$\frac{r^2 R''(r)}{R(r)} + \frac{r R'(r)}{R(r)} + \frac{\Phi''(\varphi)}{\Phi(\varphi)} - \alpha r^2 = 0,$$

ou seja,

$$\frac{r^2 R''(r)}{R(r)} + \frac{r R'(r)}{R(r)} - \alpha r^2 = -\frac{\Phi''(\varphi)}{\Phi(\varphi)}.$$

Agora, nesta equação, o lado esquerdo é apenas uma função de r , enquanto que o lado direito é apenas uma função de φ , e portanto ambos os lados devem ser constantes, ou seja, existe uma constante $\beta \in \mathbb{R}$ tal que

$$R''(r) + \frac{1}{r} R'(r) - \left(\alpha - \frac{\beta}{r^2}\right) R(r) = 0, \quad \Phi''(\varphi) = \beta \Phi(\varphi).$$

Passo 4: Começando pela última dessas equações e usando o fato de que Φ deve ser uma função periódica de φ , com período 2π , obtemos uma família, parametrizada por um inteiro m , de soluções, definidas por

$$\Phi_m(\varphi) = \exp(im\varphi) \quad (m \in \mathbb{Z}),$$

e tais que o correspondente valor de β é

$$\beta_m = -m^2.$$

Assim, a primeira dessas equações assume a forma

$$R''(r) + \frac{1}{r} R'(r) - \left(\alpha + \frac{m^2}{r^2}\right) R(r) = 0.$$

Reescalando a variável radial, definimos

$$s = \sqrt{-\alpha} r, \quad S(s) = R(r),$$

onde usamos que a constante α deve ser negativa, e concluímos que S deve satisfazer a equação de Bessel

$$S''(s) + \frac{1}{s} S'(s) + \left(1 - \frac{m^2}{s^2}\right) S(s) = 0. \quad (2.79)$$

Esta equação tem duas soluções linearmente independentes, que podem ser encontradas utilizando uma expansão do tipo de série de Laurent em torno da origem, ou seja, fazendo um “Ansatz” da forma

$$S(s) = \sum_{p=0}^{\infty} c_p s^{p+q}$$

onde q é algum inteiro fixo e os c_p são coeficientes reais: basta inserir esta expansão na equação diferencial (2.79) e reagrupar os termos para obter

$$\sum_{p=0}^{\infty} c_p ((p+q)^2 - m^2) s^{p+q-2} + \sum_{p=2}^{\infty} c_{p-2} s^{p+q-2} = 0.$$

Considerando primeiro os dois termos de ordem mais baixa, proporcionais a s^{q-2} e a s^{q-1} , concluímos que esta equação somente poderá ter soluções se $q = \pm|m|$ e $c_1 = 0$; neste caso, os demais termos da equação exigem que os coeficientes c_p satisfaçam a seguinte relação de recorrência:

$$p(p+2q)c_p + c_{p-2} = 0 \quad \text{para } p \geq 2, \text{ com } q = \pm|m|.$$

Em particular, teremos $c_p = 0$ para p ímpar. Além disso, o caso de q ser negativo na verdade não existe, pois se tivermos $q = -|m| < 0$, pondo $p = -2q = 2|m|$ implica $c_{2|m|-2} = 0$ e portanto, recursivamente, segue que $c_p = 0$ para p par também, desde que $2 \leq p \leq 2|m| - 2$, de modo que estamos de volta no caso anterior de uma expansão que começa com a potência $|m|$, ao invés de $-|m|$. Assim, escolhendo $c_0 = 1/2^{|m|}|m|!$, chegamos à função de Bessel da primeira espécie, denotada por J_m , que tem um zero de ordem $|m|$ na origem, dada explicitamente por ⁷

$$J_{|m|}(s) = \sum_{r=0}^{\infty} \frac{(-1)^r}{r!(r+|m|)!} \left(\frac{s}{2}\right)^{2r+|m|}. \quad (2.80)$$

Como essa série converge absolutamente e uniformemente em cada compacto, segue que J_m é uma função analítica inteira, e é sabido que, ao longo do semi-eixo real positivo, ela possui uma sequência infinita de zeros e uma sequência infinita de extremos que denotaremos por $(z_{m,k}^0)_{k \in \mathbb{N}}$ e por $(z_{m,k}^1)_{k \in \mathbb{N}}$, respectivamente, de

⁷A segunda solução, linearmente independente da primeira, seria a função de Bessel da segunda espécie, denotada por Y_m , mas como decorre do argumento apresentado acima, esta não é nem meromorfa na origem; assim, ela levaria a uma solução que possui uma singularidade essencial no centro do disco e portanto deve ser descartada.

modo que temos

$$\begin{aligned}
J_0(z) = 0 &\iff z = z_{0,k}^0 \text{ para algum } k \in \mathbb{N} \\
J'_0(z) = 0 &\iff z = 0 \text{ ou } z = z_{0,k}^1 \text{ para algum } k \in \mathbb{N} \\
J_1(z) = 0 &\iff z = 0 \text{ ou } z = z_{1,k}^0 \text{ para algum } k \in \mathbb{N} \\
J'_1(z) = 0 &\iff z = z_{1,k}^1 \text{ para algum } k \in \mathbb{N}
\end{aligned} \tag{2.81}$$

e para $|m| > 1$

$$\begin{aligned}
J_{|m|}(z) = 0 &\iff z = 0 \text{ ou } z = z_{|m|,k}^0 \text{ para algum } k \in \mathbb{N} \\
J'_{|m|}(z) = 0 &\iff z = 0 \text{ ou } z = z_{|m|,k}^1 \text{ para algum } k \in \mathbb{N}
\end{aligned} \tag{2.82}$$

Combinando estes resultados, obtemos uma família, parametrizada por um inteiro m e um inteiro não-negativo k , de soluções da equação (2.57) fatoráveis e satisfazendo as condições de fronteira (2.77) de Dirichlet e (2.78) de Neumann, respectivamente, definidas por

$$\begin{aligned}
w_{m,k,\circ}^{(D)}(r, \varphi) &= J_{|m|}(z_{|m|,k}^0 r/a) \exp(im\varphi) \quad (m \in \mathbb{Z}, k = 1, 2, 3, \dots), \\
w_{m,k,\circ}^{(N)}(r, \varphi) &= J_{|m|}(z_{|m|,k}^1 r/a) \exp(im\varphi) \quad (m \in \mathbb{Z}, k = 1, 2, 3, \dots),
\end{aligned} \tag{2.83}$$

e tais que o correspondente valor de α é

$$\alpha_{|m|,k,\circ}^{(D)} = -(z_{|m|,k}^0/a)^2, \quad \alpha_{|m|,k,\circ}^{(N)} = -(z_{|m|,k}^1/a)^2. \tag{2.84}$$

Passo 5: Formando combinações lineares e passando ao limite de um número infinito de termos, obtemos uma família de soluções w da equação de Laplace modificada (2.57) no disco, satisfazendo as condições de fronteira (2.77) de Dirichlet e (2.78) de Neumann, respectivamente, da forma

$$w_{\circ}^{(D/N)}(r, \varphi) = \sum_{m=-\infty}^{\infty} \sum_{k=1}^{\infty} a_{m,k} w_{m,k,\circ}^{(D/N)}(r, \varphi), \tag{2.85}$$

com coeficientes $a_{m,k}$ que devem apresentar um decaimento assintótico (i.e., no limite em que $m \rightarrow \pm\infty$ e/ou $k \rightarrow \infty$) suficiente para garantir a convergência desta série.

Por final, a discussão das correspondentes equações de difusão (2.54) e de ondas (2.55) bidimensionais, tanto no retângulo como no disco, prossegue de maneira análoga como no caso unidimensional, começando por escrever suas soluções fatoráveis, conforme a equação (2.56), e resolvendo as equações (2.58) e (2.59), respectivamente, com os valores de α dados pelas equações (2.74) ou (2.84). Deixaremos os detalhes para o leitor como exercício.

2.3 O operador de difusão

2.3.1 A equação do calor na barra

Discutir o “setup” físico e completar a análise do sistema de equações (2.58), usando os resultados do Exemplo 2.4.

Comentar sobre a extensão para valores mais altos da dimensão espacial n ($n = 2$: difusão ou condução do calor em uma placa retangular ou circular, etc.)

2.3.2 Solução da equação de difusão unidimensional

Solução por convolução com a condição inicial a ($u(0, x) = a(x)$),

$$u(t, x) = \int dy K(t, x - y) a(y) ,$$

onde K denota o *núcleo de difusão* ou *núcleo do calor*:

$$K(t, x) = \frac{1}{\sqrt{4\pi\lambda t}} \exp\left(-\frac{|x|^2}{4\lambda t}\right) .$$

Derivação por transformação de Fourier na variável x . (Detalhes a serem elaborados.)

Generalização para n dimensões do espaço:

$$u(t, x) = \int d^n y K(t, x - y) a(y) ,$$

onde

$$K(t, x) = \frac{1}{(4\pi\lambda t)^{n/2}} \exp\left(-\frac{|x|^2}{4\lambda t}\right) .$$

2.3.3 O princípio do máximo/mínimo

Veja o livro [Jost], Cap. 4, Seção 1.

2.4 O operador de ondas

2.4.1 A equação da corda vibrante

Discutir o “setup” físico e completar a análise do sistema de equações (2.59), usando os resultados do Exemplo 2.4.

Comentar sobre a extensão para valores mais altos da dimensão espacial n ($n = 2$: placa retangular (trampolim) ou circular (tamborim) vibrante, etc.)

2.4.2 Solução da equação de ondas unidimensional

O nosso objetivo nesta seção é discutir um problema específico que admite uma solução clássica simples e satisfatória: a determinação da solução geral da equação de ondas homogênea unidimensional:

$$\frac{1}{c^2} \frac{\partial^2 u}{\partial t^2} - \frac{\partial^2 u}{\partial x^2} = 0. \quad (2.86)$$

Para tal fim, efetuamos uma transformação linear de coordenadas para as tal-chamadas *coordenadas do cone da luz* ou *coordenadas características*

$$x^+ = ct + x, \quad x^- = ct - x, \quad (2.87)$$

com transformação inversa dada por

$$ct = \frac{1}{2}(x^+ + x^-), \quad x = \frac{1}{2}(x^+ - x^-). \quad (2.88)$$

Utilizando as abreviações

$$\partial_+ \equiv \frac{\partial}{\partial x^+}, \quad \partial_- \equiv \frac{\partial}{\partial x^-}, \quad (2.89)$$

temos, pela regra da cadeia,

$$\frac{1}{c} \frac{\partial}{\partial t} = \partial_+ + \partial_-, \quad \frac{\partial}{\partial x} = \partial_+ - \partial_-, \quad (2.90)$$

e

$$\partial_+ = \frac{1}{2} \left(\frac{1}{c} \frac{\partial}{\partial t} + \frac{\partial}{\partial x} \right), \quad \partial_- = \frac{1}{2} \left(\frac{1}{c} \frac{\partial}{\partial t} - \frac{\partial}{\partial x} \right). \quad (2.91)$$

Portanto,

$$\square \equiv \frac{1}{c^2} \frac{\partial^2}{\partial t^2} - \frac{\partial^2}{\partial x^2} = 4\partial_+ \partial_- = 4\partial_- \partial_+. \quad (2.92)$$

Daí, obtemos o seguinte

Teorema 2.10 (*Solução clássica geral da equação de ondas unidimensional*): *Uma função u de classe C^2 sobre \mathbb{R}^2 é solução da equação de ondas homogênea se e somente se existem funções u_+ e u_- de classe C^2 sobre \mathbb{R} tais que*

$$u(t, x) \equiv u(x^+, x^-) = u_+(x^+) + u_-(x^-). \quad (2.93)$$

DEMONSTRAÇÃO: Por um lado, a equação (2.92) mostra que uma função de classe C^2 da forma (2.93) é solução da equação de ondas (2.86). Se, por outro lado, u é uma função de classe C^2 das variáveis x^+ e x^- tal que $\square u = 0$, podemos concluir que $\partial_- u$ não depende de x^+ e portanto fornece uma função de classe C^1 sobre \mathbb{R} tal que pondo

$$u_+(x^+) = u(x^+, 0) \quad \text{e} \quad u_-(x^-) = \int_0^{x^-} dy^- (\partial_- u)(y^-),$$

obtemos funções u^+ e u^- de classe C^2 sobre \mathbb{R} satisfazendo a equação (2.93). (É óbvio que o mesmo argumento funciona quando trocamos $+$ e $-$.) □

Como corolário imediato, obtemos

Teorema 2.11 (*Solução clássica geral do problema de Cauchy para a equação de ondas unidimensional*): Dadas duas funções $a \in C^2(\mathbb{R})$ e $b \in C^1(\mathbb{R})$, existe uma única solução $u_{a,b} \in C^2(\mathbb{R}^2)$ do problema de Cauchy para a equação de ondas homogênea unidimensional com estes dados iniciais, isto é, satisfazendo

$$u_{a,b}(0, x) = a(x) \quad , \quad \frac{\partial u_{a,b}}{\partial t}(0, x) = b(x) . \quad (2.94)$$

Explicitamente, ela é dada por

$$u_{a,b}(t, x) = \frac{1}{2}(a+v)(x+ct) + \frac{1}{2}(a-v)(x-ct) , \quad (2.95)$$

onde $v \in C^2(\mathbb{R})$ é uma primitiva de $b/c \in C^1(\mathbb{R})$:

$$v'(x) = \frac{1}{c} b(x) . \quad (2.96)$$

Através desta solução, podemos verificar o princípio de “propagação de singularidades” para a equação de ondas: Caso as condições iniciais apresentarem algum tipo de singularidade, que no instante t_0 está localizada, digamos, num ponto x_0 , esta vai propagar com velocidade c no espaço-tempo, ou seja, ao longo do cone da luz, e portanto num instante posterior t estará localizada nos pontos $x_0 + c(t - t_0)$ e $x_0 - c(t - t_0)$. (Mais exatamente, como as hipóteses dos Teoremas 2.10 e 2.11 excluem singularidades “stricto sensu”, temos que tratar tais situações como limites, aproximando por exemplo uma descontinuidade por um crescimento ou decréscimo abrupto, dentro de um intervalo de comprimento ϵ .)

O Teorema 2.10 e a sua demonstração apontam para a inutilidade de procurar condições mínimas de diferenciabilidade das funções envolvidas, dentro do contexto da análise clássica. Mais especificamente, quando deixamos de exigir que as soluções consideradas sejam de classe C^2 , podemos encontrar outras soluções, mas a questão se podemos realmente falar em “soluções da equação de ondas” constitui-se num problema mal posto, pois a resposta dependerá da forma como apresentamos o operador de ondas. Por exemplo, quando escrevemos a equação de ondas na forma (2.86), podemos relaxar as hipóteses sobre u no sentido de que só é necessário exigir que u seja duas vezes diferenciável em relação às variáveis t e x , sem que haja necessidade de exigir a existência de uma das segundas derivadas mistas

$$\frac{\partial^2 u}{\partial t \partial x} \quad \text{ou} \quad \frac{\partial^2 u}{\partial x \partial t} .$$

Por outro lado, quando escrevemos a equação de ondas na forma $\partial_+ \partial_- u = 0$, a solução geral é da forma (2.93) onde u^- é apenas uma função diferenciável e

u^+ é uma função arbitrária! E quando escrevemos a equação de ondas na forma $\partial_- \partial_+ u = 0$, a solução geral é da forma (2.93) onde u^+ é apenas uma função diferenciável e u^- é uma função arbitrária! Obviamente, estamos pagando um preço alto se insistimos em impor hipóteses mínimas sobre o grau de diferenciabilidade das funções envolvidas: além de invalidar a regra de Schwartz (sobre a permutabilidade de segundas derivadas parciais), perdemos a invariância do problema sob transformações lineares das variáveis.

Como veremos no próximo capítulo, a teoria das distribuições evita todos esses problemas e – apesar de abrir o caminho para aumentar o espaço das soluções de uma forma muito mais radical do que seria possível através de uma mera redução no grau de diferenciabilidade, levando a um tratamento rigoroso de soluções singulares – respeita a invariância de equações diferenciais parciais (com coeficientes constantes) sob transformações lineares das variáveis.

Generalização para n dimensões do espaço: mais complicado. Primeiro, trata-se o caso quando n é ímpar, usando o resultado do teorema da média esférica, e depois o caso quando n é par, subindo para uma dimensão a mais e depois reduzindo: isto é o método de abaixamento de Hadamard. Veja o livro [Folland].

2.4.3 Causalidade: domínios de dependência e de influência

3 Distribuições

O presente capítulo será dedicado ao desenvolvimento do cálculo de distribuições, que constitui uma extensão do cálculo (diferencial e integral) comum. Através dele, a definição de operações como diferenciação, integração, multiplicação, convolução ou transformação de Fourier pode ser estendida a funções singulares, onde as definições clássicas não se aplicam, ou mesmo a objetos que não são funções no sentido comum, tais como a “função delta” de Dirac formalmente introduzida no Capítulo 1.8. Estes objetos chamados de “distribuições” (por L. Schwartz) ou de “funções generalizadas” (por I. Gel’fand) são realmente funcionais, ao invés de funções, e uma definição correta requer certas ferramentas da análise funcional – uma área da matemática que combina álgebra linear com topologia. Não será nossa intenção desenvolver esta teoria em toda sua profundidade, sendo o interesse principal adquirir uma capacidade pragmática, porém matematicamente correta, de aplicá-la a problemas concretos. Neste sentido, o presente texto pretende apresentar o cálculo e não a teoria de distribuições.

No capítulo inteiro, o corpo base \mathbb{F} será o corpo \mathbb{R} dos números reais ou o corpo \mathbb{C} dos números complexos, todos os espaços vetoriais serão espaços vetoriais sobre \mathbb{F} , e linearidade (de aplicações) significará \mathbb{F} -linearidade. Como é sempre mais fácil trabalhar sobre um corpo algebricamente fechado, e como é sempre possível passar de um espaço vetorial real para um espaço vetorial complexo por “complexificação”, adota-se frequentemente a estratégia de trabalhar no âmbito complexo e incorporar o caso real estudando as possíveis “formas reais”. (Resumidamente, a complexificação V^c de um espaço vetorial real V é simplesmente o produto tensorial $V \otimes \mathbb{C}$, onde vale $\alpha(v \otimes \lambda) = v \otimes \alpha\lambda$, enquanto que uma forma real de um espaço vetorial complexo W é um subespaço real V de W tal que $W \cong V^c$. Tal subespaço real V pode sempre ser realizado como o conjunto dos pontos fixos de alguma conjugação $*$ em W – a saber, aquela que sob o isomorfismo $W \cong V^c$ corresponde à conjugação em V^c levando $v \otimes \lambda$ em $v \otimes \bar{\lambda}$.¹ Reciprocamente, qualquer conjugação em W induz

¹ Conforme já foi mencionado nas notações e convenções que podem ser encontradas no início deste livro, uma *involução* em um conjunto X qualquer é uma aplicação de X em X cujo quadrado é a identidade, ou seja, uma bijeção de X que é o seu próprio inverso, e uma *conjugação* em um espaço vetorial complexo W é uma involução antilinear em W .

uma decomposição direta de W , como espaço vetorial real, nos seus autoespaços associados aos seus dois autovalores, 1 e -1 , os quais são transformados entre si pela multiplicação por i , de modo que $W = V \oplus iV$.) Essa abordagem se mostra particularmente frutífera na área de álgebra, onde foi adotada originalmente por Élie Cartan para resolver o problema da classificação das álgebras de Lie semi-simples e depois aplicada a álgebras associativas, motivando o conceito de uma $*$ -álgebra já mencionado nas notações e convenções que podem ser encontradas no início deste livro.

3.1 Dualidade na álgebra linear

Na teoria de distribuições, as noções do espaço dual e da transposta de uma aplicação linear ocupam uma posição central na definição de quase todas as operações usuais do cálculo. Portanto, parece adequado apresentar, preliminarmente, o cerne algébrico desses conceitos.

Inicialmente, formalizamos a definição de espaço dual, já mencionada nas notações e convenções que podem ser encontradas no início deste livro.

Definição 3.1 *Seja V um espaço vetorial. O **espaço dual** V^* de V é o espaço das formas lineares sobre V , isto é, das aplicações lineares de V para o corpo base \mathbb{F} , ou seja, $V^* = L(V, \mathbb{F})$.*

Lembramos que a adição e a multiplicação por escalares em V^* são definidas “pontualmente”.

No que segue, usaremos o símbolo $\langle v^*, v \rangle$, ao invés de $v^*(v)$, para denotar o número em \mathbb{F} obtido pela aplicação da forma linear $v^* \in V^*$ ao vetor $v \in V$. Esta notação é baseada na ideia de tratar a avaliação de formas sobre vetores como definindo um “pareamento” entre o espaço original V e seu espaço dual V^* , ou seja, como uma aplicação bilinear

$$\begin{aligned} V^* \times V &\longrightarrow \mathbb{F} \\ (v^*, v) &\longmapsto \langle v^*, v \rangle \end{aligned} \tag{3.1}$$

que é não-degenerada em ambos os fatores:

$$\begin{aligned} \langle v^*, v \rangle = 0 \quad \text{para todo } v \in V &\implies v^* = 0. \\ \langle v^*, v \rangle = 0 \quad \text{para todo } v^* \in V^* &\implies v = 0. \end{aligned} \tag{3.2}$$

A primeira destas duas propriedades é uma consequência imediata da definição do espaço dual, enquanto que a segunda é bem mais profunda: dado um vetor $v \in V$ com $v \neq 0$, usa-se o teorema da base para concluir que existe uma base de V à qual ele pertence, o que permite construir uma forma linear $v^* \in V^*$ sobre V que satisfaz $\langle v^*, v \rangle = 1$ e se anula sobre todos os demais vetores da referida base.

A noção de dual existe não somente em nível de espaços vetoriais mas também em nível de aplicações lineares entre espaços vetoriais.

Definição 3.2 *Sejam V e W espaços vetoriais e $T : V \rightarrow W$ uma aplicação linear. A **aplicação linear dual** ou **transposta** T^* de T é a aplicação linear $T^* : W^* \rightarrow V^*$ definida por*

$$\langle T^*(w^*), v \rangle = \langle w^*, T(v) \rangle \quad \text{para } v \in V, w^* \in W^*. \quad (3.3)$$

Note que há várias afirmações embutidas nesta definição:

- Em primeiro lugar, observamos que para todo $w^* \in W^*$, a aplicação $V \rightarrow \mathbb{F}$ definida por $v \mapsto \langle w^*, T(v) \rangle$, como composição das duas aplicações lineares $T : V \rightarrow W$ e $w^* : W \rightarrow \mathbb{F}$, é linear. Portanto, a equação (3.3) realmente define uma aplicação T^* de W^* para V^* .
- Em segundo lugar, observamos que a aplicação T^* é linear, pois para todo $v \in V$ e para $\lambda_1, \lambda_2 \in \mathbb{F}$, $w_1^*, w_2^* \in W^*$, temos

$$\begin{aligned} \langle T^*(\lambda_1 w_1^* + \lambda_2 w_2^*), v \rangle &= \langle \lambda_1 w_1^* + \lambda_2 w_2^*, T(v) \rangle \\ &= \lambda_1 \langle w_1^*, T(v) \rangle + \lambda_2 \langle w_2^*, T(v) \rangle \\ &= \lambda_1 \langle T^*(w_1^*), v \rangle + \lambda_2 \langle T^*(w_2^*), v \rangle \\ &= \langle \lambda_1 T^*(w_1^*) + \lambda_2 T^*(w_2^*), v \rangle. \end{aligned}$$

No caso de espaços vetoriais de dimensão finita, estes conceitos podem ser reformulados em termos de componentes e de matrizes, relativos a bases apropriadas. Para tanto, precisamos primeiro introduzir a noção de bases duais.

Definição 3.3 *Seja V um espaço vetorial de dimensão n e $\{e_1, \dots, e_n\}$ uma base de V . Então os vetores $e_i^* \in V^*$ definidos por*

$$\langle e_i^*, e_j \rangle = \delta_{ij} \quad (1 \leq i, j \leq n) \quad (3.4)$$

formam uma base $\{e_1^, \dots, e_n^*\}$ de V^* , chamada a **base dual** à base original.*

Novamente, há duas afirmações embutidas nesta definição:

- Os vetores e_1^*, \dots, e_n^* geram V^* :
De fato, para $v^* \in V^*$ qualquer e $1 \leq j \leq n$, pomos $\lambda_j = \langle v^*, e_j \rangle$ e obtemos

$$\left\langle \sum_{i=1}^n \lambda_i e_i^*, e_j \right\rangle = \sum_{i=1}^n \lambda_i \langle e_i^*, e_j \rangle = \sum_{i=1}^n \lambda_i \delta_{ij} = \lambda_j = \langle v^*, e_j \rangle.$$

Tendo em vista que $\{e_1, \dots, e_n\}$ é uma base de V e que aplicações lineares são iguais quando coincidem sobre os vetores de uma base, concluímos que

$$\sum_{i=1}^n \lambda_i e_i^* = v^*.$$

- Os vetores e_1^*, \dots, e_n^* são linearmente independentes:

De fato, se $\lambda_1, \dots, \lambda_n$ são coeficientes tais que

$$\sum_{i=1}^n \lambda_i e_i^* = 0,$$

então para $1 \leq j \leq n$, vale

$$0 = \left\langle \sum_{i=1}^n \lambda_i e_i^*, e_j \right\rangle = \sum_{i=1}^n \lambda_i \langle e_i^*, e_j \rangle = \sum_{i=1}^n \lambda_i \delta_{ij} = \lambda_j.$$

Isso posto, podemos estabelecer que no caso de aplicações lineares entre espaços vetoriais de dimensão finita, a transposição de aplicações lineares, relativamente a bases duais, corresponde à transposição de matrizes. De fato, sejam V e W espaços vetoriais de dimensão finita e $T : V \rightarrow W$ uma aplicação linear. Introduzindo uma base $\{e_1, \dots, e_n\}$ de V e uma base $\{f_1, \dots, f_m\}$ de W , junto com as respectivas bases duais $\{e_1^*, \dots, e_n^*\}$ de V^* e $\{f_1^*, \dots, f_m^*\}$ de W^* , podemos representar a aplicação linear $T : V \rightarrow W$ por sua matriz (T_{ki}) e a aplicação linear transposta $T^* : W^* \rightarrow V^*$ por sua matriz (T_{ik}^*) , conforme segue:

$$T(e_i) = \sum_{l=1}^m T_{li} f_l \quad \text{para } 1 \leq i \leq n,$$

$$T^*(f_k^*) = \sum_{j=1}^n T_{jk}^* e_j^* \quad \text{para } 1 \leq k \leq m.$$

Então temos

$$T_{ki} = \sum_{l=1}^m T_{li} \delta_{kl} = \sum_{l=1}^m T_{li} \langle f_k^*, f_l \rangle = \left\langle f_k^*, \sum_{l=1}^m T_{li} f_l \right\rangle = \langle f_k^*, T(e_i) \rangle,$$

$$T_{ik}^* = \sum_{j=1}^n T_{jk}^* \delta_{ij} = \sum_{j=1}^n T_{jk}^* \langle e_j^*, e_i \rangle = \left\langle \sum_{j=1}^n T_{jk}^* e_j^*, e_i \right\rangle = \langle T^*(f_k^*), e_i \rangle,$$

o que nos leva a concluir que, independentemente das bases utilizadas, a matriz de T^* é exatamente a transposta da matriz de T , justificando assim a terminologia introduzida na Definição 3.2.

Essas observações implicam, em particular, que para um espaço vetorial de dimensão finita, $V^* \cong V$ e $\dim V^* = \dim V$. Porém, não existe nenhum isomorfismo preferido, ou canônico, entre V e V^* , sendo que a escolha de um tal isomorfismo depende da – e em alguns casos equivale à – escolha de uma estrutura adicional. O exemplo padrão de uma tal estrutura adicional é um produto escalar (\cdot, \cdot) sobre V , que induz um isomorfismo linear

$$\begin{aligned} V &\longrightarrow V^* \\ v &\longmapsto (v, \cdot) \end{aligned} \tag{3.5}$$

Utilizando este isomorfismo linear para identificar os dois espaços, uma aplicação linear $T : V \rightarrow V$ é chamada *simétrica* se $T^* = T$ e *antissimétrica* se $T^* = -T$.

A situação é diferente quando, para um espaço vetorial V de dimensão finita, consideramos o seu *espaço bidual* V^{**} , pois é possível construir, sem qualquer hipótese adicional, um isomorfismo linear canônico

$$\begin{aligned} i : V &\longrightarrow V^{**} \\ v &\longmapsto i(v) \end{aligned} \quad (3.6)$$

definido por

$$\langle i(v), v^* \rangle = \langle v^*, v \rangle \quad \text{para } v^* \in V^* . \quad (3.7)$$

Podemos utilizar e sempre utilizaremos este isomorfismo para identificar todo espaço vetorial V de dimensão finita com seu espaço bidual V^{**} . Também identificaremos toda aplicação linear $T : V \rightarrow W$ entre espaços vetoriais de dimensão finita com sua *aplicação linear bidual* $T^{**} : V^{**} \rightarrow W^{**}$.

A extensão das ideias acima expostas – e mais geralmente de conceitos e métodos provindo da álgebra linear – ao caso de espaços vetoriais de dimensão infinita constitui-se num problema altamente não-trivial e é um dos temas centrais da área da matemática conhecida como análise funcional. A título de exemplo, notamos que para espaços vetoriais de dimensão infinita, a aplicação (3.5), definida por um produto escalar (\cdot, \cdot) dado, fornece tão somente um mergulho de V em V^* , ao invés de um isomorfismo, e a aplicação (3.6), definida pela equação (3.7), fornece tão somente um mergulho de V em V^{**} , ao invés de um isomorfismo,² de tal forma que tomando o dual de uma aplicação linear simétrica e o bidual de uma aplicação linear qualquer, obtemos uma continuação ou extensão, ao invés de uma identificação.

Na análise funcional, os conceitos e métodos provindo da álgebra linear são complementados e enriquecidos por noções e técnicas de natureza topológica e analítica. Mais exatamente, os espaços vetoriais da análise funcional carregam, além de sua estrutura algébrica como espaços vetoriais, uma topologia, o que permite empregar noções topológicas tais como convergência, continuidade, conjuntos compactos, conjuntos conexos, etc.. Esta topologia deve ser compatível com a estrutura algébrica (o que significa simplesmente que a adição e a multiplicação por escalares devem ser operações contínuas), e todas as aplicações lineares relevantes devem ser contínuas. Isto significa, em particular, que o espaço V^* de todas as formas lineares sobre V , chamado o (*espaço*) *dual algébrico* de V , deve ser substituído pelo espaço V' de todas as formas lineares contínuas sobre V , chamado o (*espaço*) *dual topológico* de V .

Os espaços V que tipicamente aparecem na teoria de distribuições são *espaços de funções*. Portanto, os seus duais são *espaços de funcionais*, sendo que um *funcional* é, por definição, uma aplicação (não necessariamente linear) de um espaço de funções para o corpo base, ou seja, um funcional é uma “função” numérica cujo argumento é outra função, em vez de um ponto em \mathbb{F} ou em \mathbb{F}^n . Na teoria

² De fato, é possível mostrar que se a aplicação (3.6) definida pela equação (3.7) for um isomorfismo, então V é necessariamente de dimensão finita.

de distribuições, os funcionais relevantes são funcionais lineares e contínuos e os operadores relevantes são aplicações lineares e contínuas. Ademais, os espaços envolvidos são tais que a continuidade de aplicações lineares pode ser formulada através do critério de transformar sequências convergentes (para 0) em sequências convergentes (para 0).

Estes conceitos, que servem como fundamento teórico para o desenvolvimento do cálculo de distribuições, serão elaborados de maneira mais detalhada na próxima seção, porém sem demonstrações.

3.2 Espaços vetoriais topológicos e espaços localmente convexos

Começamos pela formalização de alguns conceitos mencionados no fim da seção anterior.

Definição 3.4 *Um espaço vetorial topológico E é simultaneamente um espaço vetorial e um espaço topológico tal que a adição*

$$\begin{aligned} E \times E &\longrightarrow E \\ (x, y) &\longmapsto x + y \end{aligned} \quad (3.8)$$

e a multiplicação por escalares

$$\begin{aligned} \mathbb{F} \times E &\longrightarrow E \\ (\lambda, x) &\longmapsto \lambda x \end{aligned} \quad (3.9)$$

são aplicações contínuas.

Definição 3.5 *Um espaço vetorial topológico localmente convexo ou simplesmente espaço localmente convexo é um espaço vetorial topológico que admite uma base de vizinhanças abertas da origem que são convexas.*

Para um melhor entendimento destas definições, faremos primeiro uma breve revisão de alguns conceitos básicos da topologia geral.

Inicialmente, lembremos que um espaço topológico pode ser definido como um conjunto X munido de um determinado conjunto \mathfrak{X} de subconjuntos de X , chamados os subconjuntos abertos ou simplesmente *abertos* de X , satisfazendo os seguintes axiomas:

- (i) O conjunto vazio \emptyset e o conjunto total X são abertos.
- (ii) A interseção $U_1 \cap \dots \cap U_N$ de qualquer família finita $(U_i)_{i=1 \dots N}$ de abertos é aberta.
- (iii) A união $\bigcup_{i \in I} U_i$ de qualquer família $(U_i)_{i \in I}$ de abertos é aberta.

O conjunto \mathfrak{X} de todos os abertos de X é chamado a *topologia* de X . No entanto, na maioria das situações, é suficiente considerar apenas os abertos pertencendo a uma *base* da topologia, isto é, a um subconjunto \mathfrak{B} de \mathfrak{X} tal que qualquer aberto U de X ($U \in \mathfrak{X}$) pode ser escrito como a união de uma família $(U_i)_{i \in I}$ de abertos U_i da base ($U_i \in \mathfrak{B}$). É fácil ver que isso é o caso se e somente se para todo aberto U de X ($U \in \mathfrak{X}$) e todo ponto $x \in U$, existe um aberto B da base ($B \in \mathfrak{B}$) tal que $x \in B \subset U$. Também é fácil ver que uma base \mathfrak{B} de \mathfrak{X} deve satisfazer os seguintes axiomas:

- (i) \mathfrak{B} é um recobrimento de X , i.e., todo elemento de X pertence a algum elemento B da base \mathfrak{B} .
- (ii) Para quaisquer dois elementos B_1 e B_2 da base \mathfrak{B} e qualquer ponto x de X na interseção $B_1 \cap B_2$, existe um elemento B_3 da base \mathfrak{B} tal que $x \in B_3 \subset B_1 \cap B_2$.

Reciprocamente, mostra-se que dado um conjunto X munido de um determinado conjunto \mathfrak{B} de subconjuntos de X satisfazendo estes dois axiomas, existe uma única topologia \mathfrak{X} sobre X que tenha \mathfrak{B} como base: é a *topologia gerada* por esta base.

Outros meios de caracterizar uma topologia são através dos subconjuntos fechados ou simplesmente *fechados* de X , ou através do sistema de *vizinhanças* de cada ponto de X . Por exemplo, um subconjunto A de X é chamado *fechado* se seu complemento em X é aberto. Portanto, o sistema dos subconjuntos fechados de X satisfaz seu próprio sistema de axiomas:

- (i) O conjunto vazio \emptyset e o conjunto total X são fechados.
- (ii) A união $A_1 \cup \dots \cup A_N$ de qualquer família finita $(A_i)_{i=1 \dots N}$ de fechados é fechada.
- (iii) A interseção $\bigcap_{i \in I} A_i$ de qualquer família $(A_i)_{i \in I}$ de fechados é fechada.

Este sistema de axiomas garante que, reciprocamente, a topologia de X pode ser reconstruída a partir do sistema dos subconjuntos fechados de X : dado um conjunto X munido de um determinado conjunto de subconjuntos de X chamados fechados e satisfazendo o referido sistema de axiomas, obtemos uma topologia se definirmos um subconjunto U de X como sendo aberto se seu complemento em X for fechado. Por outro lado, um subconjunto V_x de X é chamado uma *vizinhança* do ponto x de X se existe um aberto U_x de X tal que $x \in U_x \subset V_x$. (Observe que, segundo esta convenção, as vizinhanças dos pontos de X não são necessariamente abertas e portanto deve-se distinguir entre vizinhanças e vizinhanças abertas.) Novamente, o sistema das vizinhanças dos pontos de X satisfaz seu próprio sistema de axiomas que não detalharemos aqui mas que garante que, reciprocamente, a topologia de X pode ser reconstruída a partir dele: dado um conjunto X munido de um determinado conjunto de subconjuntos de X chamados vizinhanças e satisfazendo o sistema de axiomas pertinente, obtemos uma topologia se definirmos um subconjunto U de X como sendo aberto se ele for vizinhança de cada um dos seus pontos.

Continuando, lembremos que uma aplicação $f : X \rightarrow Y$ entre espaços topológicos é chamada *contínua no ponto* x de X se para qualquer vizinhança $V_{f(x)}$ de $f(x)$ em Y , existe uma vizinhança U_x de x em X tal que $f(U_x) \subset V_{f(x)}$, e é chamada *contínua* se é contínua em todo ponto de X ; isto é o caso se e somente se a imagem inversa de qualquer aberto V de Y é um aberto $f^{-1}(V)$ de X . Finalmente, um *homeomorfismo* entre espaços topológicos X e Y é uma aplicação bijetora contínua $f : X \rightarrow Y$ tal que a aplicação inversa $f^{-1} : Y \rightarrow X$ também seja contínua. (Cabe destacar que esta condição não é automática: existem aplicações bijetoras contínuas cuja inversa não é contínua.)

Além disso, a Definição 3.4 utiliza o conceito do produto de dois espaços topológicos, que é mais facilmente entendido em termos de bases: dados dois espaços topológicos X , com topologia \mathfrak{X} , e Y , com topologia \mathfrak{Y} , demonstra-se que o produto cartesiano $Z = X \times Y$ admite uma única topologia \mathfrak{Z} tal que para qualquer base \mathfrak{B}_X de X e qualquer base \mathfrak{B}_Y de Y , o conjunto

$$\mathfrak{B}_X \times \mathfrak{B}_Y = \{B_X \times B_Y \mid B_X \in \mathfrak{B}_X, B_Y \in \mathfrak{B}_Y\}$$

forma uma base de \mathfrak{Z} .

Com essa terminologia à nossa disposição, podemos afirmar que, em qualquer espaço vetorial topológico E , a translação por qualquer vetor a de E ,

$$T_a : E \rightarrow E \\ x \mapsto x + a \quad , \quad (3.10)$$

é uma aplicação contínua e, mais ainda, um homeomorfismo, já que a aplicação inversa é a translação pelo vetor $-a$ de E : $T_a^{-1} = T_{-a}$. Portanto, ela leva o sistema das vizinhanças (vizinhanças abertas) da origem 0 de E para o sistema das vizinhanças (vizinhanças abertas) do ponto a de E . Isso implica que a topologia de um espaço vetorial topológico sempre admite uma base que seja invariante sob translações e portanto é completamente determinada por uma base de vizinhanças abertas da origem 0 .

Finalmente, para entendermos o aspecto algébrico das definições apresentadas no início desta seção, precisamos apenas lembrar que um subconjunto C de um espaço vetorial é chamado *convexo* se para quaisquer dois pontos de C , o inteiro segmento da reta que une estes dois pontos pertence a C :

$$x_0, x_1 \in C \implies \lambda x_1 + (1 - \lambda)x_0 \in C \text{ para } 0 \leq \lambda \leq 1 .$$

Observe que a interseção de uma família de conjuntos convexos é um conjunto convexo e que translações e aplicações lineares transformam conjuntos convexos em conjuntos convexos.

Isso posto, o significado da Definição 3.5 torna-se mais claro, pois os conjuntos convexos são os que mais se aproximam da ideia de conjuntos “redondos”. De fato, as “bolas” definidas por uma norma ou, mais geralmente, um sistema de seminormas são convexas. Portanto, um método padrão de munir espaços vetoriais de uma topologia de espaço localmente convexo é através de sistemas de seminormas.

Definição 3.6 Seja E um espaço vetorial. Uma aplicação $s : E \rightarrow \mathbb{R}$ é chamada uma **seminorma** sobre E se possui as seguintes três propriedades:

- (i) *Positividade:* $s(x) \geq 0$ para $x \in E$;
- (ii) *Homogeneidade:* $s(\lambda x) = |\lambda| s(x)$ para $\lambda \in \mathbb{F}$ e $x \in E$;
- (iii) *Desigualdade triangular:* $s(x + y) \leq s(x) + s(y)$ para $x, y \in E$.

Ela é chamada uma **norma** sobre E se, ao invés de (i), vale

- (i') *Positividade definida:* $s(x) \geq 0$ e $s(x) = 0 \Leftrightarrow x = 0$ para $x \in E$.

Consideremos alguns exemplos, primeiro para espaços de dimensão finita e, em seguida, para alguns espaços de dimensão infinita.

Exemplo 3.1 No espaço vetorial \mathbb{F}^n , definimos a *norma da soma* ou *norma tipo l^1* , denotada por $\|\cdot\|_1$, e a *norma do máximo* ou *norma tipo l^∞* , denotada por $\|\cdot\|_\infty$, pondo

$$\|x\|_1 = \sum_{i=1}^n |x_i| \quad \text{para } x \in \mathbb{F}^n, \quad (3.11)$$

$$\|x\|_\infty = \max_{i=1 \dots n} |x_i| \quad \text{para } x \in \mathbb{F}^n. \quad (3.12)$$

Outros exemplos são a *norma euclidiana* ou *norma tipo l^2* , denotada por $\|\cdot\|_2$ e definida por

$$\|x\|_2^2 = \sum_{i=1}^n |x_i|^2 \quad \text{para } x \in \mathbb{F}^n, \quad (3.13)$$

e mais geralmente, para qualquer número real $p \geq 1$, a *norma tipo l^p* , denotada por $\|\cdot\|_p$ e definida por

$$\|x\|_p^p = \sum_{i=1}^n |x_i|^p \quad \text{para } x \in \mathbb{F}^n. \quad (3.14)$$

Para estas, é mais difícil provar a desigualdade triangular, sendo que a demonstração requer, preliminarmente, estabelecer a desigualdade de Cauchy-Schwarz (no caso $p = 2$) ou de Hölder (no caso geral $p > 1$).

Exemplo 3.2 Seja X um subconjunto qualquer de \mathbb{R}^n . Denotamos por $B(X)$ o espaço vetorial das funções limitadas sobre X , por $C(X)$ o espaço vetorial das funções contínuas sobre X e por $C_b(X)$ o espaço vetorial das funções contínuas e limitadas sobre X , a valores em \mathbb{F} , de modo que vale $C_b(X) = B(X) \cap C(X)$. Em $B(X)$, definimos a *norma do supremo*, denotada por $\|\cdot\|_X$, pondo

$$\|f\|_X = \sup_{x \in X} |f(x)| \quad \text{para } f \in B(X). \quad (3.15)$$

Em particular, para um subconjunto compacto K de \mathbb{R}^n , $C_b(K) = C(K)$ e

$$\|f\|_K = \sup_{x \in K} |f(x)| \quad \text{para } f \in C(K). \quad (3.16)$$

Por outro lado, para um subconjunto aberto Ω de \mathbb{R}^n , $C_b(\Omega) \subsetneq C(\Omega)$ e

$$\|f\|_\Omega = \sup_{x \in \Omega} |f(x)| \quad \text{para } f \in C_b(\Omega). \quad (3.17)$$

Também introduzimos os seguintes subespaços: o espaço vetorial $C_c(\Omega)$ das funções contínuas sobre Ω a valores em \mathbb{F} de suporte compacto contido em Ω e o espaço vetorial $C_0(\Omega)$ das funções contínuas sobre Ω a valores em \mathbb{F} que se anulam no bordo de Ω ou, quando $\Omega = \mathbb{R}^n$, se anulam no infinito, com as inclusões

$$C_c(\Omega) \subsetneq C_0(\Omega) \subsetneq C_b(\Omega) \subsetneq C(\Omega), \quad (3.18)$$

onde os primeiros três espaços podem ser munidos da norma do supremo, enquanto que $C(\Omega)$ não possui nenhuma norma natural. (Dizemos que uma função contínua f sobre Ω *se anula no bordo de Ω* ou, quando $\Omega = \mathbb{R}^n$, *se anula no infinito*, se para todo $\epsilon > 0$, existe um subconjunto compacto K_ϵ de Ω tal que $|f(x)| < \epsilon$ para $x \in \Omega \setminus K_\epsilon$.)

Exemplo 3.3 Denotamos por $\mathcal{L}^1(\mathbb{R}^n)$ o espaço vetorial das funções mensuráveis sobre \mathbb{R}^n a valores em \mathbb{F} que são integráveis (no sentido de Lebesgue), por $\mathcal{L}^2(\mathbb{R}^n)$ o espaço vetorial das funções mensuráveis sobre \mathbb{R}^n a valores em \mathbb{F} que são quadraticamente integráveis (no sentido de Lebesgue), e mais geralmente, para qualquer número real $p \geq 1$, por $\mathcal{L}^p(\mathbb{R}^n)$ o espaço vetorial das funções mensuráveis sobre \mathbb{R}^n a valores em \mathbb{F} cujos módulos elevados à p -ésima potência são integráveis (no sentido de Lebesgue). Em $\mathcal{L}^1(\mathbb{R}^n)$, definimos a *seminorma L^1* , denotada por $\|\cdot\|_1$, pondo

$$\|f\|_1 = \int d^n x |f(x)| \quad \text{para } f \in \mathcal{L}^1(\mathbb{R}^n), \quad (3.19)$$

em $\mathcal{L}^2(\mathbb{R}^n)$, definimos a *seminorma L^2* , denotada por $\|\cdot\|_2$, pondo

$$\|f\|_2^2 = \int d^n x |f(x)|^2 \quad \text{para } f \in \mathcal{L}^2(\mathbb{R}^n), \quad (3.20)$$

e mais geralmente, em $\mathcal{L}^p(\mathbb{R}^n)$, definimos a *seminorma L^p* , denotada por $\|\cdot\|_p$, pondo

$$\|f\|_p^p = \int d^n x |f(x)|^p \quad \text{para } f \in \mathcal{L}^p(\mathbb{R}^n). \quad (3.21)$$

Novamente, a demonstração da desigualdade triangular requer, preliminarmente, estabelecer a desigualdade de Cauchy-Schwarz (no caso $p = 2$) ou de Hölder (no caso geral $p > 1$). No entanto, os teoremas usuais sobre a integral de Lebesgue implicam que, para qualquer valor de p , a condição $\|f\|_p = 0$ implica apenas que $f = 0$ quase sempre. Portanto, $\|\cdot\|_p$ é realmente apenas uma seminorma e não uma norma.

Exemplo 3.4 Denotamos por $\mathcal{L}^\infty(\mathbb{R}^n)$ o espaço vetorial das funções mensuráveis sobre \mathbb{R}^n a valores em \mathbb{F} que são essencialmente limitadas, i.e., limitadas exceto sobre conjuntos de medida zero. Em $\mathcal{L}^\infty(\mathbb{R}^n)$, definimos a *seminorma do supremo essencial* ou *seminorma* L^∞ , denotada por $\|\cdot\|_\infty$, pondo

$$\|f\|_\infty = \inf_{\tilde{f}} \sup_{x \in \mathbb{R}^n} |\tilde{f}(x)| \quad \text{para } f \in \mathcal{L}^\infty(\mathbb{R}^n), \quad (3.22)$$

onde o ínfimo é sobre todas as funções \tilde{f} mensuráveis sobre \mathbb{R}^n a valores em \mathbb{F} que são limitadas e iguais a f exceto sobre um conjunto de medida zero. Novamente, temos que a condição $\|f\|_\infty = 0$ implica apenas que $f = 0$ quase sempre. Portanto, $\|\cdot\|_\infty$ é realmente apenas uma seminorma e não uma norma.

A solução do problema apontado nos últimos dois exemplos é bem conhecida: consiste da passagem de funções a classes de equivalência de funções, sendo que duas funções são declaradas equivalentes quando diferem apenas sobre um conjunto de medida zero. O conjunto das classes de equivalência formadas por funções pertencendo ao espaço $\mathcal{L}^p(\mathbb{R}^n)$ é denotado por $L^p(\mathbb{R}^n)$ e forma um espaço vetorial sobre o qual $\|\cdot\|_p$ define sim uma norma e não apenas uma seminorma. De fato, esta afirmação é um caso particular do seguinte fato geral.

Exercício 3.1 Sejam E um espaço vetorial e s uma seminorma sobre E . Prove que o conjunto

$$E_0 = \{x \in E \mid s(x) = 0\} \quad (3.23)$$

é um subespaço vetorial de E , chamado o *núcleo* de s , e que s induz uma norma $\|\cdot\|_s$ sobre o espaço quociente E/E_0 , dada por

$$\|x \bmod E_0\|_s = s(x) \quad \text{para } x \in E. \quad (3.24)$$

Observação 3.1 Cada um dos espaços vetoriais normados dos Exemplos 3.2-3.4 acima, com a exceção de $C_c(\Omega)$, é *completo*, no sentido de que toda sequência de Cauchy nele converge, e portanto, por definição, é um *espaço de Banach*. (Veja a Definição 3.10 logo abaixo.) Essa afirmação resume, de forma compacta, toda uma série de teoremas clássicos da análise, entre eles a afirmação de que o limite uniforme de uma sequência de funções limitadas ou de funções contínuas ou de funções contínuas que se anulam no bordo ou no infinito é, respectivamente, uma função limitada ou função contínua ou função contínua que se anula no bordo ou no infinito; por outro lado, o limite uniforme de uma sequência de funções de suporte compacto não é, em geral, uma função de suporte compacto. Assim, $B(X)$ é um espaço de Banach e $C_b(X)$ é um subespaço fechado de $B(X)$, portanto também um espaço de Banach, Ademais, $C_0(\Omega)$ é um subespaço fechado de $C_b(\Omega)$ e assim, novamente, também um espaço de Banach, enquanto que $C_c(\Omega)$ não é: na verdade, é fácil provar que $C_c(\Omega)$ é denso em $C_0(\Omega)$. Encontraremos afirmações semelhantes mais adiante e, por isso, deixaremos a prova dos fatos aqui citados como exercício para o leitor.

Passemos agora ao estudo de sistemas de seminormas. A ideia básica é que qualquer família $(s_i)_{i \in I}$ de seminormas sobre um espaço vetorial E fornece um sistema de “bolas” em E , definidas como segue. Para todo índice $i \in I$ e todo número positivo $\epsilon > 0$, definimos a bola $B_i(\epsilon)$ de raio ϵ em torno da origem por

$$B_i(\epsilon) = \{x \in E \mid s_i(x) < \epsilon\}. \quad (3.25)$$

Mais geralmente, para todo vetor $a \in E$, todo subconjunto finito $\{i_1, \dots, i_r\}$ do conjunto I de índices e todo número positivo $\epsilon > 0$, definimos a bola $B_{i_1, \dots, i_r}(a, \epsilon)$ de raio ϵ em torno de a por

$$B_{i_1, \dots, i_r}(a, \epsilon) = \{x \in E \mid s_{i_1}(x - a) < \epsilon, \dots, s_{i_r}(x - a) < \epsilon\}. \quad (3.26)$$

Obviamente, $B_{i_1, \dots, i_r}(a, \epsilon)$ é a imagem de

$$B_{i_1, \dots, i_r}(0, \epsilon) \equiv B_{i_1, \dots, i_r}(\epsilon) = B_{i_1}(\epsilon) \cap \dots \cap B_{i_r}(\epsilon) \quad (3.27)$$

sob a translação T_a .

Exercício 3.2 Sejam E um espaço vetorial e $(s_i)_{i \in I}$ uma família de seminormas sobre E . Usando a homogeneidade e a desigualdade triangular para cada uma delas, prove que as bolas $B_{i_1, \dots, i_r}(a, \epsilon)$ são convexas e formam uma base de uma topologia que torna E um espaço localmente convexo. Dizemos que esta topologia é *gerada* pela referida família de seminormas.

Agora, é um dos teoremas fundamentais da teoria que a afirmação recíproca também é verdade:

Teorema 3.1 *A topologia de qualquer espaço localmente convexo pode ser gerada por uma família apropriada de seminormas.*

Este fato explica o papel central desempenhado pelos espaços localmente convexos na análise funcional, pois permite transformar afirmações de natureza topológica em afirmações de caráter analítico, expressas em termos de estimativas envolvendo seminormas, e vice versa. Como exemplo deste tipo de transformação, citamos a seguinte

Proposição 3.1 *Sejam E e F espaços localmente convexos com topologias geradas por famílias $(s_i^E)_{i \in I}$ e $(s_j^F)_{j \in J}$ de seminormas, respectivamente, e seja $T : E \rightarrow F$ uma aplicação linear. Então as seguintes propriedades de T são equivalentes:*

- (i) T é contínua.
- (ii) T é contínua na origem.
- (iii) T é limitada no seguinte sentido: para todo $j \in J$, existem $i_1, \dots, i_p \in I$ e uma constante $C_{j; i_1, \dots, i_p} > 0$ tais que

$$s_j^F(Tx) \leq C_{j; i_1, \dots, i_p} \max_{1 \leq k \leq p} s_{i_k}^E(x) \quad \text{para } x \in E. \quad (3.28)$$

Outro exemplo é o seguinte critério que permite decidir se um espaço localmente convexo satisfaz ou não o axioma de Hausdorff sobre separação de pontos por vizinhanças: é uma simples generalização do fato de que uma única seminorma gera uma topologia de Hausdorff se e somente se for uma norma:

Proposição 3.2 *Seja E um espaço localmente convexo com topologia gerada por uma família $(s_i)_{i \in I}$ de seminormas. Então E é um espaço de Hausdorff se e somente se a interseção dos núcleos de todas as seminormas da família é trivial, i.e., se e somente se para todo $x \in E$ com $x \neq 0$, existe $i \in I$ tal que $s_i(x) \neq 0$.*

A demonstração é elementar, usando apenas a definição de um *espaço de Hausdorff* como um espaço topológico tal que quaisquer dois pontos distintos x e y podem ser separados por vizinhanças, i.e., para $x \neq y$ existem vizinhanças V_x de x e V_y de y tais que $V_x \cap V_y = \emptyset$.

Uma das utilidades principais do axioma de separação de Hausdorff é que ele garante a unicidade de limites, em particular de limites de sequências convergentes. Para uso posterior, damos aqui um critério explícito de convergência de sequências:

Proposição 3.3 *Seja E um espaço localmente convexo com topologia gerada por uma família $(s_i)_{i \in I}$ de seminormas. Uma sequência $(x_k)_{k \in \mathbb{N}}$ de vetores em E converge para um vetor x em E se e somente se para todo $i \in I$, a sequência numérica $s_i(x_k - x)_{k \in \mathbb{N}}$ converge para 0, ou seja, se e somente se vale*

$$\forall i \in I \quad \forall \epsilon > 0 \quad \exists k_0 \in \mathbb{N} \quad \forall k > k_0 : \quad s_i(x_k - x) < \epsilon . \quad (3.29)$$

De forma análoga, temos:

Definição 3.7 *Seja E um espaço localmente convexo com topologia gerada por uma família $(s_i)_{i \in I}$ de seminormas. Uma sequência $(x_k)_{k \in \mathbb{N}}$ de vetores em E é uma sequência de Cauchy se e somente se vale*

$$\forall i \in I \quad \forall \epsilon > 0 \quad \exists k_0 \in \mathbb{N} \quad \forall k, l > k_0 : \quad s_i(x_k - x_l) < \epsilon . \quad (3.30)$$

Para completar a discussão do mecanismo descrito no Exercício 3.2 que, segundo o Teorema 3.1, caracteriza a classe dos espaços localmente convexos, dentro da classe dos espaços vetoriais topológicos, cabe mencionar que é possível construir exemplos de espaços vetoriais topológicos que não sejam localmente convexos, mas que esses exemplos são um tanto artificiais e parecem não ter nenhuma utilidade concreta. Por outro lado, embora a topologia de qualquer espaço localmente convexo seja unicamente determinada por algum sistema de seminormas, observa-se que tal sistema está longe de ser fixado pela topologia, o que motiva a seguinte

Definição 3.8 *Seja E um espaço vetorial. Duas famílias $(s_i)_{i \in I}$ e $(s'_i)_{i' \in I'}$ de seminormas sobre E são ditas **equivalentes** se geram a mesma topologia sobre E .*

Não é difícil encontrar um critério explícito para decidir se dois sistemas de seminormas sobre o mesmo espaço vetorial são equivalentes ou não: ele segue diretamente da Proposição 3.1, aplicada à identidade de E , considerando E como o mesmo espaço vetorial mas como dois espaços localmente convexos diferentes, cada um com a topologia gerada pelo sistema de seminormas pertinente:

Proposição 3.4 *Seja E um espaço vetorial. Duas famílias $(s_i)_{i \in I}$ e $(s'_{i'})_{i' \in I'}$ de seminormas sobre E são equivalentes se e somente se*

- para todo $i \in I$, existem $i'_1, \dots, i'_q \in I'$ e uma constante $C_{i; i'_1, \dots, i'_q} > 0$ tais que

$$s_i(x) \leq C_{i; i'_1, \dots, i'_q} \max_{1 \leq l \leq q} s'_{i'_l}(x) \quad \text{para } x \in E, \quad (3.31)$$

- para todo $i' \in I'$, existem $i_1, \dots, i_p \in I$ e uma constante $C_{i'; i_1, \dots, i_p} > 0$ tais que

$$s'_{i'}(x) \leq C_{i'; i_1, \dots, i_p} \max_{1 \leq k \leq p} s_{i_k}(x) \quad \text{para } x \in E. \quad (3.32)$$

Uma aplicação concreta deste critério é a seguinte

Proposição 3.5 *Seja E um espaço localmente convexo com topologia gerada por uma família **finita** $(s_i)_{i=1 \dots r}$ de seminormas. Então pondo*

$$\|x\|_\infty = \max_{i=1 \dots r} s_i(x) \quad \text{para } x \in E, \quad (3.33)$$

ou

$$\|x\|_1 = \sum_{i=1}^r s_i(x) \quad \text{para } x \in E, \quad (3.34)$$

ou

$$\|x\|_2 = \left(\sum_{i=1}^r s_i(x)^2 \right)^{1/2} \quad \text{para } x \in E, \quad (3.35)$$

ou, para qualquer número real $p \geq 1$,

$$\|x\|_p = \left(\sum_{i=1}^r s_i(x)^p \right)^{1/p} \quad \text{para } x \in E, \quad (3.36)$$

obtemos uma única seminorma sobre E que é equivalente à família original de seminormas, ou seja, gera a mesma topologia. Em particular, um espaço localmente convexo de Hausdorff com topologia gerada por uma família finita de seminormas é um **espaço normável** e, se fixarmos uma norma equivalente, um **espaço normado**.

Uma afirmação análoga relaciona espaços localmente convexos de Hausdorff com topologia gerada por um sistema enumerável de seminormas e espaços vetoriais topológicos metrizáveis:

Definição 3.9 *Seja E um espaço vetorial. Uma aplicação $d : E \times E \rightarrow \mathbb{R}$ é chamada uma **semi-métrica** sobre E **invariante por translações** se possui as seguintes propriedades:*

- (i) *Positividade:* $d(x, y) \geq 0$ e $d(x, x) = 0$ para $x, y \in E$;
- (ii) *Simetria:* $d(y, x) = d(x, y)$ para $x, y \in E$;
- (iii) *Desigualdade triangular:* $d(x, z) \leq d(x, y) + d(y, z)$ para $x, y, z \in E$;
- (iv) *Invariância sob translações:* $d(x + a, y + a) = d(x, y)$ para $a, x, y \in E$.

Ela é chamada uma **métrica** sobre E **invariante por translações** se, ao invés de (i), vale

- (i') *Positividade definida:* $d(x, y) \geq 0$ e $d(x, y) = 0 \Leftrightarrow x = y$ para $x, y \in E$.

Um espaço vetorial topológico cuja topologia decorre de alguma métrica invariante por translações é chamado **metrizável** e, se fixarmos essa métrica, **metrizado**.

Temos então a seguinte

Proposição 3.6 *Seja E um espaço localmente convexo com topologia gerada por uma família **enumerável** $(s_i)_{i \in \mathbb{N}}$ de seminormas. Então a fórmula*

$$d(x, y) = \sum_{i=1}^{\infty} \frac{1}{2^i} \frac{s_i(x - y)}{1 + s_i(x - y)} \quad \text{para } x, y \in E \quad (3.37)$$

define uma semi-métrica sobre E invariante por translações que gera a mesma topologia sobre E que a família original de seminormas. Em particular, um espaço localmente convexo de Hausdorff com topologia gerada por uma família enumerável de seminormas é um espaço vetorial topológico metrizável.

Reciprocamente, temos:

Teorema 3.2 *A topologia de qualquer espaço vetorial topológico metrizável pode ser gerada por uma família enumerável de seminormas.*

Em particular, isso inclui a afirmação de que um espaço vetorial topológico metrizável é automaticamente localmente convexo. Portanto, daqui em diante, nos referiremos a este tipo de espaço como um *espaço localmente convexo metrizável* e, se fixarmos a métrica, *metrizado*.

Definição 3.10 *Um espaço localmente convexo E normável ou metrizável é dito **completo** se toda sequência de Cauchy em E é convergente. Um espaço localmente convexo normável completo é chamado um **espaço de Banach**, e um espaço localmente convexo metrizável completo é chamado um **espaço de Fréchet**.*

Cabe salientar que a noção de um espaço localmente convexo completo (e até de um espaço vetorial topológico completo) pode ser introduzida de maneira absolutamente geral, mesmo quando a topologia de um tal espaço não pode ser gerada por algum sistema enumerável de seminormas. Neste caso, no entanto, a definição requer substituir sequências por redes ou filtros (convergentes ou de Cauchy, respectivamente) – conceitos que não utilizaremos neste livro.

Finalmente, mesmo no caso geral de espaços localmente convexos arbitrários, podemos usar a liberdade de passar de um sistema de seminormas a qualquer outro que seja equivalente para impor condições adicionais sobre os sistemas de seminormas a serem utilizados. Um exemplo trivial dessa liberdade é que sistemas de seminormas obtidos um do outro através de alguma permutação dos índices são sempre equivalentes, ou seja, dada qualquer família $(s_i)_{i \in I}$ de seminormas e qualquer bijeção $\sigma : I \rightarrow I$, a família permutada $(s_{\sigma(i)})_{i \in I}$ gera a mesma topologia localmente convexa que a família original $(s_i)_{i \in I}$. Sendo assim, a topologia localmente convexa gerada por uma família de seminormas depende apenas do conjunto de seminormas que dela fazem parte e não da maneira como este conjunto é parametrizado, o que abre a opção de ordenar este conjunto da maneira mais adequada ao problema específico sob consideração, como veremos a seguir.

Para organizar a análise da situação, começamos por lembrar que um conjunto X é chamado de *ordenado* se vem munido de uma *relação de ordem* \leq entre os seus elementos, sujeita aos axiomas de *reflexividade* (para qualquer $x \in X$, vale $x \leq x$), *antissimetria* (para quaisquer $x, x' \in X$, se vale $x \leq x'$ e $x' \leq x$, então $x = x'$) e *transitividade* (para quaisquer $x, x', x'' \in X$, se vale $x \leq x'$ e $x' \leq x''$, então $x \leq x''$); dizemos então que dois elementos x e x' de X são *comparáveis* se vale uma das duas relações, $x \leq x'$ ou $x' \leq x$, e chamamos a ordem de *total* se quaisquer dois elementos de X são comparáveis, enquanto que a chamamos de *parcial* para indicar a possibilidade de existirem elementos que não são comparáveis.³ Mencionamos também que uma aplicação $f : X \rightarrow Y$ entre conjuntos ordenados que preserva a ordem (i.e., tal que para quaisquer $x, x' \in X$, se vale $x \leq x'$, então $f(x) \leq f(x')$) costuma ser chamada de *monótona*. Continuando, lembramos ainda que um conjunto (parcialmente) ordenado X é chamado de *direcionado* se vale a seguinte condição adicional: para quaisquer $x, x' \in X$, existe $x'' \in X$ tal que $x \leq x''$ e $x' \leq x''$: podemos chamar tal elemento x'' de cota superior do conjunto $\{x, x'\}$ e, por indução, concluir que a propriedade de X ser direcionado significa que qualquer subconjunto finito de X possui uma cota superior. A título de exemplo que é relevante no presente contexto, notamos que dado um espaço vetorial E , o conjunto $\text{SN}(E)$ de todas as seminormas sobre E , munido da relação de ordem natural definida por

$$s \leq s' \iff \left(s(x) \leq s'(x) \text{ para todo } x \in E \right) \quad \text{para } s, s' \in \text{SN}(E), \quad (3.38)$$

é um conjunto direcionado, sendo que uma cota superior para $\{s, s'\}$ é $\max\{s, s'\}$, ou $s + s'$. Combinando essas noções, chegamos ao seguinte conceito:

³ Em inglês, costuma-se usar o termo “poset”, abreviando a expressão “partially ordered set”.

Definição 3.11 *Seja E um espaço vetorial. Um **sistema direcionado** de seminormas sobre E é uma família $(s_i)_{i \in I}$ de seminormas sobre E onde I é um conjunto direcionado e a família, como aplicação de I em $\text{SN}(E)$, é monótona, ou seja, vale*

$$s_i(x) \leq s_{i'}(x) \quad \text{para todo } x \in E \quad \text{se } i \leq i' . \quad (3.39)$$

Obviamente, sistemas finitos e enumeráveis de seminormas são automaticamente direcionados (e até totalmente ordenados, pelo menos tacitamente, ou seja, mediante a escolha de uma enumeração), mas mesmo no caso completamente geral, vale a seguinte afirmação, reforçando o Teorema 3.1 acima.

Teorema 3.3 *A topologia de qualquer espaço localmente convexo pode ser gerada por um sistema direcionado de seminormas.*

De fato, se a topologia de E for gerada por alguma família $(s_i)_{i \in I}$ de seminormas sobre E , podemos considerar uma nova família $(s_{i'})_{i' \in I'}$ de seminormas sobre E assim definida: I' é o conjunto dos subconjuntos finitos de I , que é um conjunto direcionado pela inclusão, e para $i' \in I'$, $s_{i'}(x) = \max_{i \in i'} s_i(x)$; então é claro que esta nova família é um sistema direcionado equivalente à família original.

Como exemplo da utilidade de se usar sistemas direcionados de seminormas, formulamos as seguintes versões simplificadas das Proposições 3.1 e 3.4, que as torna quase literalmente iguais às afirmações correspondentes para espaços normados:

Proposição 3.7 *Sejam E e F espaços localmente convexos com topologias geradas por sistemas direcionados $(s_i^E)_{i \in I}$ e $(s_j^F)_{j \in J}$ de seminormas, respectivamente, e seja $T : E \rightarrow F$ uma aplicação linear. Então as seguintes propriedades de T são equivalentes:*

- (i) T é contínua.
- (ii) T é contínua na origem.
- (iii) T é limitada no seguinte sentido: para todo $j \in J$, existem $i \in I$ e uma constante $C_{j;i} > 0$ tais que

$$s_j^F(Tx) \leq C_{j;i} s_i^E(x) \quad \text{para } x \in E . \quad (3.40)$$

Proposição 3.8 *Seja E um espaço vetorial. Dois sistemas direcionados $(s_i)_{i \in I}$ e $(s'_{i'})_{i' \in I'}$ de seminormas sobre E são equivalentes se e somente se*

- para todo $i \in I$, existem $i' \in I'$ e uma constante $C_{i;i'} > 0$ tais que

$$s_i(x) \leq C_{i;i'} s'_{i'}(x) \quad \text{para } x \in E , \quad (3.41)$$

- para todo $i' \in I'$, existem $i \in I$ e uma constante $C_{i';i} > 0$ tais que

$$s'_{i'}(x) \leq C_{i';i} s_i(x) \quad \text{para } x \in E . \quad (3.42)$$

A utilidade dos conceitos introduzidos acima para a teoria de distribuições decorre do fato de que, ao contrário dos espaços introduzidos nos Exemplos 3.2-3.4 acima, alguns dos espaços mais importantes dessa teoria são espaços de Fréchet, mas não de Banach. A seguir, apresentaremos três classes de exemplos de tais espaços de funções, com apenas ligeiras variações, indicando em cada caso o tipo de convergência de seqüências, conforme o critério formulado na Proposição 3.3.

Exemplo 3.5 Sejam Ω um aberto de \mathbb{R}^n e r um inteiro não-negativo. Denotamos por $C_b^r(\Omega)$ o espaço vetorial das funções de classe C^r sobre Ω a valores em \mathbb{F} que, assim como todas as suas derivadas até ordem r , são limitadas sobre Ω . Neste espaço vetorial, definimos a *norma do supremo das derivadas até ordem r* , denotada por $\|\cdot\|_{r,\Omega}$, pondo⁴

$$\|f\|_{r,\Omega} = 2^r \max_{|\alpha| \leq r} \sup_{x \in \Omega} |\partial_\alpha f(x)| \quad \text{para } f \in C_b^r(\Omega). \quad (3.43)$$

Também introduzimos os seguintes subespaços: o espaço vetorial $C_c^r(\Omega)$ das funções de classe C^r sobre Ω a valores em \mathbb{F} de suporte compacto contido em Ω e o espaço vetorial $C_0^r(\Omega)$ das funções de classe C^r sobre Ω a valores em \mathbb{F} que, assim como todas as suas derivadas até ordem r , se anulam no bordo de Ω ou, quando $\Omega = \mathbb{R}^n$, se anulam no infinito,⁵ com as inclusões

$$C_c^r(\Omega) \subsetneq C_0^r(\Omega) \subsetneq C_b^r(\Omega) \subsetneq C^r(\Omega), \quad (3.44)$$

onde os primeiros três espaços podem ser munidos da norma do supremo das derivadas até ordem r , enquanto que $C^r(\Omega)$ não possui nenhuma norma natural. Obviamente, para $r = 0$, recuperamos o caso do Exemplo 3.2, notando que a topologia gerada pela norma $\|\cdot\|_\Omega$ é geralmente conhecida como a *topologia da convergência uniforme* sobre Ω . Mais geralmente, para $r > 0$, a topologia gerada pela norma $\|\cdot\|_{r,\Omega}$ é chamada a *topologia da convergência uniforme das derivadas até ordem r* sobre Ω . Finalmente, denotamos por $C_b^\infty(\Omega)$ o espaço vetorial das funções de classe C^∞ sobre Ω a valores em \mathbb{F} que, assim como todas as suas derivadas, são limitadas sobre Ω , e introduzimos os correspondentes subespaços: o espaço vetorial $C_c^\infty(\Omega)$ das funções de classe C^∞ sobre Ω a valores em \mathbb{F} de suporte compacto contido em Ω e o espaço vetorial $C_0^\infty(\Omega)$ das funções de classe C^∞ sobre Ω a valores em \mathbb{F} que, assim como todas as suas derivadas, se anulam no bordo de Ω ou, quando $\Omega = \mathbb{R}^n$, se anulam no infinito,⁵ com as inclusões

$$C_c^\infty(\Omega) \subsetneq C_0^\infty(\Omega) \subsetneq C_b^\infty(\Omega) \subsetneq C^\infty(\Omega). \quad (3.45)$$

Obviamente, $C_b^\infty(\Omega)$, $C_c^\infty(\Omega)$ e $C_0^\infty(\Omega)$ é, respectivamente, a interseção de todos os espaços $C_b^r(\Omega)$, $C_c^r(\Omega)$ e $C_0^r(\Omega)$ ($r = 0, 1, 2, \dots$), e todos eles são espaços localmente convexos metrízáveis em relação à topologia gerada pela seqüência das normas $\|\cdot\|_{r,\Omega}$ ($r = 0, 1, 2, \dots$), chamada a *topologia da convergência uniforme de todas as derivadas* sobre Ω .

⁴ O motivo para a inclusão do fator 2^r , que nesta altura é meramente convencional, será esclarecido na Observação 3.3 logo abaixo.

⁵ Lembramos que, por definição, uma função f sobre \mathbb{R}^n se anula no infinito se para todo $\epsilon > 0$, existe um compacto $K \subset \mathbb{R}^n$ tal que $|f(x)| < \epsilon$ para todo $x \in \mathbb{R}^n \setminus K$, e de forma semelhante, uma função f sobre um subconjunto aberto Ω de \mathbb{R}^n se anula no bordo de Ω se para todo $\epsilon > 0$, existe um compacto $K \subset \Omega$ tal que $|f(x)| < \epsilon$ para todo $x \in \Omega \setminus K$.

Exemplo 3.6 Sejam K um subconjunto compacto de \mathbb{R}^n cujo interior K^0 é denso em K e r um inteiro não-negativo. Denotamos por $C^r(K)$ o espaço vetorial das funções de classe C^r sobre K a valores em \mathbb{F} . (Lembramos aqui a convenção, já adotada anteriormente, de que uma função f de classe C^r sobre K é uma função f de classe C^r sobre o interior K^0 de K tal que cada uma das suas derivadas $\partial_\alpha f$ de ordem $|\alpha| \leq r$ admite uma extensão contínua de K^0 para K , que é necessariamente única e, para simplificar a notação, também é denotada por $\partial_\alpha f$.) Neste espaço vetorial, definimos a *norma do supremo das derivadas até ordem r* , denotada por $\|\cdot\|_{r,K}$, pondo⁴

$$\|f\|_{r,K} = 2^r \max_{|\alpha| \leq r} \sup_{x \in K} |\partial_\alpha f(x)| \quad \text{para } f \in C^r(K). \quad (3.46)$$

Obviamente, para $r = 0$, recuperamos o caso do Exemplo 3.2, notando que a topologia gerada pela norma $\|\cdot\|_K$ é geralmente conhecida como a *topologia da convergência uniforme* sobre K . Mais geralmente, para $r > 0$, a topologia gerada pela norma $\|\cdot\|_{r,K}$ é chamada a *topologia da convergência uniforme das derivadas até ordem r* sobre K . Finalmente, denotamos por $C^\infty(K)$ o espaço vetorial das funções de classe C^∞ sobre K a valores em \mathbb{F} : obviamente, $C^\infty(K)$ é a interseção de todos os espaços $C^r(K)$ ($r = 0, 1, 2, \dots$) e é um espaço localmente convexo metrizable em relação à topologia gerada pela sequência das normas $\|\cdot\|_{r,K}$ ($r = 0, 1, 2, \dots$), chamada a *topologia da convergência uniforme de todas as derivadas* sobre K .

Exemplo 3.7 Sejam Ω um aberto de \mathbb{R}^n e r um inteiro não-negativo. Denotamos por $C^r(\Omega)$ o espaço vetorial das funções de classe C^r sobre Ω a valores em \mathbb{F} . Para cada subconjunto compacto K de Ω , a fórmula do exemplo anterior⁵

$$\|f\|_{r,K} = 2^r \max_{|\alpha| \leq r} \sup_{x \in K} |\partial_\alpha f(x)| \quad \text{para } f \in C^r(\Omega) \quad (3.47)$$

define uma seminorma $\|\cdot\|_{r,K}$ sobre $C^r(\Omega)$. Para $r = 0$, $C(\Omega)$ é um espaço localmente convexo em relação à topologia gerada pela família das seminormas $\|\cdot\|_K$ ($K \subset \Omega$, K compacto), geralmente conhecida como a *topologia da convergência uniforme sobre subconjuntos compactos* de Ω . Mais geralmente, para $r > 0$, $C^r(\Omega)$ é um espaço localmente convexo em relação à topologia gerada pela família das seminormas $\|\cdot\|_{r,K}$ ($K \subset \Omega$, K compacto), chamada a *topologia da convergência uniforme das derivadas até ordem r sobre subconjuntos compactos* de Ω . Finalmente, denotamos por $C^\infty(\Omega)$ o espaço vetorial das funções de classe C^∞ sobre Ω a valores em \mathbb{F} : obviamente, $C^\infty(\Omega)$ é a interseção de todos os espaços $C^r(\Omega)$ ($r = 0, 1, 2, \dots$) e é um espaço localmente convexo em relação à topologia gerada pela família das seminormas $\|\cdot\|_{r,K}$ ($r = 0, 1, 2, \dots$, $K \subset \Omega$, K compacto), chamada a *topologia da convergência uniforme de todas as derivadas sobre subconjuntos compactos* de Ω . Cabe ressaltar que todos estes espaços localmente convexos são metrizable, pois suas topologias podem ser geradas por famílias enumeráveis de seminormas: as famílias acima descritas são equivalentes a famílias onde o conjunto de todos os subconjuntos compactos K de Ω é substituído por alguma (na verdade, qualquer) sequência crescente $(K_k)_{k \in \mathbb{N}}$ de subconjuntos compactos K_k

de Ω que preenchem Ω , i.e., que satisfazem

$$K_1 \subset \dots \subset K_k \subset K_{k+1} \subset \dots \subset \Omega \quad \text{e} \quad \bigcup_{k \in \mathbb{N}} K_k = \Omega,$$

de modo que todo subconjunto compacto de Ω é contido em algum K_k .

Observação 3.2 Cada um dos espaços localmente convexos metrizáveis dos Exemplos 3.5-3.7 acima, com a exceção de $C_c^r(\Omega)$ e $C_c^\infty(\Omega)$, é *completo*, no sentido de que toda sequência de Cauchy nele converge, e portanto, por definição, é um *espaço de Fréchet* e, no caso dos espaços dos Exemplos 3.5 e 3.6 acima com $r < \infty$, até um *espaço de Banach*. Mais uma vez, essa afirmação resume, de forma compacta, toda uma série de teoremas clássicos da análise.

Observação 3.3 Uma propriedade adicional de todos os espaços de funções introduzidos nos Exemplos 3.2 e 3.5-3.7 acima, com $\mathbb{F} = \mathbb{C}$, é que são não apenas *espaços vetoriais* localmente convexos mas são **-álgebras* comutativas localmente convexas, em relação ao produto de funções comum e à conjugação complexa comum que, assim como a adição e a multiplicação por escalares, são definidos “pontualmente”, ou seja, por

$$(fg)(x) = f(x)g(x) \quad \text{e} \quad \overline{f}(x) = \overline{f(x)} \quad (3.48)$$

Sem querer entrar na extensa teoria de **-álgebras* topológicas e de **-álgebras* localmente convexas, entre as quais figuram com destaque as Banach **-álgebras* e as C^* -*álgebras*, vamos aqui apenas registrar algumas convenções que são pertinentes para entender corretamente a terminologia. Primeiro, uma **-álgebra* topológica é definida de maneira óbvia: é uma **-álgebra* munida de uma topologia tal que, além da adição e a multiplicação por escalares, o produto e a conjugação são contínuos.⁶ Segundo, uma **-álgebra* localmente convexa é uma **-álgebra* cuja topologia pode ser gerada por uma família, e portanto até por um sistema direcionado, de **-seminormas* submultiplicativas:

Definição 3.12 *Seja A uma **-álgebra*. Uma **-seminorma submultiplicativa* sobre A é uma *seminorma* s sobre A com as seguintes propriedades adicionais:*

- (i) *Invariância sob conjugação:* $s(a^*) = s(a)$ para $a \in A$;
- (ii) *Submultiplicatividade:* $s(ab) \leq s(a)s(b)$ para $a, b \in A$.

*Ela é chamada uma **-norma submultiplicativa* se for positiva definida.*

Finalmente, notando que qualquer **-seminorma* submultiplicativa s sobre uma **-álgebra* A satisfaz as desigualdades

$$s(a^*a) \leq s(a)^2 \quad \text{e} \quad s(aa^*) \leq s(a)^2 \quad \text{para } a \in A, \quad (3.49)$$

podemos introduzir a noção de uma C^* -*seminorma*:

⁶Cabe mencionar aqui que frequentemente, exige-se do produto apenas que ele seja separadamente contínuo.

Definição 3.13 *Seja A uma $*$ -álgebra. Uma C^* -seminorma sobre A é uma $*$ -seminorma submultiplicativa s sobre A para a qual as desigualdades anteriores se tornam uma igualdade:*

$$s(a^*a) = s(aa^*) = s(a)^2 \quad \text{para } a \in A. \quad (3.50)$$

Ela é chamada uma C^ -norma se for positiva definida.*

E com isso chegamos à seguinte

Definição 3.14 *Uma **Banach $*$ -álgebra** é uma $*$ -álgebra munida de uma $*$ -norma submultiplicativa em relação à qual ela se torna um espaço de Banach. Uma C^* -álgebra é uma $*$ -álgebra munida de uma C^* -norma em relação à qual ela se torna um espaço de Banach.*

Para justificar a afirmação feita no início desta observação, vamos mostrar que todas as seminormas $\|\cdot\|_{r,X}$ dos Exemplos 3.2 e 3.5-3.7 acima são $*$ -seminormas submultiplicativas. De fato, a invariância destas seminormas sob conjugação é evidente, enquanto que sua submultiplicatividade, ou seja, a validade da desigualdade

$$\|fg\|_{r,X} \leq \|f\|_{r,X} \|g\|_{r,X} \quad (3.51)$$

para $X = \Omega$ e $f, g \in C_b^r(\Omega)$ ou $X = K \subset \Omega$ e $f, g \in C^r(K)$ ou $C^r(\Omega)$, segue do seguinte cálculo, baseado na regra do produto (1.57),

$$\begin{aligned} \|fg\|_{r,X} &= 2^r \max_{|\alpha| \leq r} \sup_{x \in X} |\partial_\alpha(fg)(x)| \\ &\leq 2^r \max_{|\alpha| \leq r} \sum_{\beta \leq \alpha} \binom{\alpha}{\beta} \sup_{x \in X} |\partial_\beta f(x)| \sup_{x \in X} |\partial_{\alpha-\beta} g(x)| \\ &\leq (2^r)^2 \max_{|\beta| \leq r} \sup_{x \in X} |\partial_\beta f(x)| \max_{|\gamma| \leq r} \sup_{x \in X} |\partial_\gamma g(x)| \\ &= \|f\|_{r,X} \|g\|_{r,X} \end{aligned}$$

onde usamos o fato de que a soma dos coeficientes binomiais $\binom{k}{l}$, de $l = 0$ até $l = k$, vale 2^k e portanto

$$\sum_{\beta \leq \alpha} \binom{\alpha}{\beta} = \prod_{i=1}^n \sum_{\beta_i=0}^{\alpha_i} \binom{\alpha_i}{\beta_i} = \prod_{i=1}^n 2^{\alpha_i} = 2^{|\alpha|}$$

que é $\leq 2^r$ para todo $\alpha \in \mathbb{N}^n$ com $|\alpha| \leq r$: foi para garantir a validade desta desigualdade, sem fator adicional, que incluímos o fator 2^r na definição das seminormas nas equações (3.43)-(3.47). Por outro lado, entre as seminormas $\|\cdot\|_{r,X}$ dos Exemplos 3.2 e 3.5-3.7 acima, apenas as com $r=0$ são C^* -seminormas, pois

$$\begin{aligned} \|\bar{f}f\|_X &= \sup_{x \in X} |(\bar{f}f)(x)| = \sup_{x \in X} |f(x)|^2 \\ &= \left(\sup_{x \in X} |f(x)| \right)^2 = \|f\|_X^2 \end{aligned}$$

com $X = \Omega$ e $f, g \in C_b(\Omega)$ ou $X = K \subset \Omega$ e $f, g \in C(K)$ ou $C(\Omega)$, enquanto que a mesma igualdade falha assim que $r > 0$, como mostra o seguinte contra-exemplo para o caso mais simples onde $r = 1$ e $n = 1$, tomando $\Omega = \mathbb{R}$ e considerando as funções oscilatórias $f_k \in C_b^1(\mathbb{R})$ definidas por $f_k(x) = \exp(ikx)$, para $k \in \mathbb{R}$ tal que $k > 1$: com estas escolhas, vale $\|\tilde{f}_k f_k\|_{1, \mathbb{R}} = 1$, enquanto que $\|f_k\|_{1, \mathbb{R}} = 2k$, mostrando que a discrepância entre os dois lados das desigualdades (3.49) pode até se tornar arbitrariamente grande, na medida em que $k \rightarrow \infty$.

Finalmente, precisamos abordar brevemente a questão de como definir o dual de um espaço vetorial topológico e, em particular, de um espaço localmente convexo:

Definição 3.15 *Seja E um espaço vetorial topológico. O dual algébrico de E é o espaço E^* de todas as formas lineares sobre E , e o dual topológico de E é o espaço E' de todas as formas lineares contínuas sobre E ; obviamente, temos*

$$E' \subset E^* .$$

Nesta definição, a palavra “espaço” utilizada para caracterizar E' (ou E^*) significa apenas “espaço vetorial”, pois não especificamos nenhuma topologia sobre E' (ou E^*). Ocorre que existem várias opções para munir o espaço E' de uma topologia localmente convexa, entre as quais se destaca uma, inclusive pelo fato de que ela ocupa uma posição importante na teoria de distribuições. Sua definição baseia-se na observação elementar de que, dado um espaço localmente convexo E , cada vetor x em E fornece uma seminorma s_x sobre o dual topológico E' de E , a saber o valor absoluto da forma linear sobre E' definida pela avaliação dos elementos de E' no vetor x :

$$s_x(x') = |\langle x', x \rangle| \quad \text{para } x' \in E' . \quad (3.52)$$

A topologia sobre E' gerada pela família $(s_x)_{x \in E}$ de seminormas é chamada a *topologia fraca* sobre E' (em relação a E). Ela também pode ser interpretada como a *topologia da convergência pontual sobre E* , já que segundo a Proposição 3.3, uma sequência $(x'_k)_{k \in \mathbb{N}}$ de formas lineares contínuas $x'_k \in E'$ sobre E converge para uma forma linear contínua $x' \in E'$ sobre E nesta topologia se e somente para todo vetor x em E , a sequência numérica $(\langle x'_k, x \rangle)_{k \in \mathbb{N}}$ converge para $\langle x', x \rangle$. Este ponto de vista facilita o entendimento de outras topologias localmente convexas em E' , pois todas elas podem ser definidas como topologias de convergência uniforme sobre certas classes de subconjuntos de E ; um exemplo é a topologia padrão de espaço normado no dual topológico E' de um espaço normado E , que pode ser vista como a topologia de convergência uniforme sobre a bola unitária de E .

3.3 Espaços de funções teste

A seguir, detalhamos a definição dos três mais importantes espaços de funções encontrados na teoria de distribuições: seus elementos são chamados *funções teste*, por motivos a serem esclarecidos mais adiante; uma definição precisa do conceito de um espaço de funções teste será apresentada na Seção 3.8 (veja a Definição 3.24).

Definição 3.16 *Os mais importantes espaços de funções teste da teoria de distribuições são*

- o espaço \mathcal{E} de todas as funções lisas,

$$\mathcal{E} = C^\infty(\mathbb{R}^n), \quad (3.53)$$

- o espaço \mathcal{D} das funções lisas de suporte compacto,

$$\mathcal{D} = C_c^\infty(\mathbb{R}^n) = \{ \varphi \in C^\infty(\mathbb{R}^n) \mid \text{supp } \varphi \text{ compacto} \}, \quad (3.54)$$

com os subespaços

$$\mathcal{D}(K) = \{ \varphi \in C^\infty(\mathbb{R}^n) \mid \text{supp } \varphi \subset K \}, \quad (3.55)$$

onde K é um subconjunto compacto de \mathbb{R}^n ,

- o espaço \mathcal{S} das funções lisas que, assim como tidas as suas derivadas, são rapidamente decrescentes,

$$\mathcal{S} = \{ \varphi \in C^\infty(\mathbb{R}^n) \mid \sup_{x \in \mathbb{R}^n} |x^\beta \partial_\alpha \varphi(x)| < \infty \text{ para } \alpha, \beta \in \mathbb{N}^n \}. \quad (3.56)$$

Para todo aberto Ω de \mathbb{R}^n , definimos ainda

$$\mathcal{E}(\Omega) = C^\infty(\Omega), \quad (3.57)$$

e

$$\mathcal{D}(\Omega) = \{ \varphi \in C^\infty(\mathbb{R}^n) \mid \text{supp } \varphi \text{ compacto, } \text{supp } \varphi \subset \Omega \}. \quad (3.58)$$

Note que o espaço $\mathcal{E}(\Omega)$ não é subespaço nem espaço quociente de \mathcal{E} , pois não há nenhuma relação direta entre estes dois espaços.⁷ Quanto ao espaço $\mathcal{D}(\Omega)$, pode-se perguntar se não seria mais natural adotar a seguinte definição:

$$\mathcal{D}(\Omega) = C_c^\infty(\Omega) = \{ \varphi \in C^\infty(\Omega) \mid \text{supp } \varphi \text{ compacto, } \text{supp } \varphi \subset \Omega \}. \quad (3.59)$$

Ocorre que as duas definições acabam sendo equivalentes; porém, a prova desta afirmação requer um argumento mais cuidadoso devido a uma sutil diferença entre as noções de suporte usadas: pondo $N = \{x \in \Omega \mid \varphi(x) \neq 0\}$ (que para qualquer função contínua φ sobre \mathbb{R}^n ou sobre Ω é um subconjunto aberto de Ω), o suporte na equação (3.58) é o fecho de N em \mathbb{R}^n enquanto que o suporte na equação (3.59) é o fecho de N em Ω . No entanto, como se exige que estes fechos sejam compactos, ambos serão fechados tanto em Ω como em \mathbb{R}^n e portanto acabam sendo iguais, mantendo uma distância > 0 do bordo $\partial\Omega = \bar{\Omega} \setminus \Omega$ de Ω , de modo que $\mathcal{D}(\Omega)$ pode e deve sempre ser considerado como um subespaço de \mathcal{D} , com a convenção de que

⁷ A projeção de \mathcal{E} para $\mathcal{E}(\Omega)$ dada pela restrição de uma função sobre \mathbb{R}^n a Ω obviamente não é injetora, e também não é sobrejetora, pois existem funções de classe C^∞ sobre Ω que divergem na fronteira de Ω de tal forma a não admitirem nenhuma extensão ao \mathbb{R}^n inteiro que seja de classe C^∞ .

uma função teste definida sobre Ω e com suporte compacto contido em Ω deve ser estendida ao \mathbb{R}^n inteiro, atribuindo-lhe o valor 0 fora de Ω . O mesmo argumento mostra que para quaisquer dois abertos Ω_1 e Ω_2 de \mathbb{R}^n , vale

$$\mathcal{D}(\Omega_1) \subset \mathcal{D}(\Omega_2) \quad \text{se} \quad \Omega_1 \subset \Omega_2 . \quad (3.60)$$

Também temos as inclusões

$$\mathcal{D}(\Omega) \subsetneq \mathcal{E}(\Omega) \quad \text{e} \quad \mathcal{D} \subsetneq \mathcal{S} \subsetneq \mathcal{E} , \quad (3.61)$$

sendo que a função Gaussiana sobre \mathbb{R}^n , definida por

$$\text{Gs}_a(x) = \exp(-|x|^2/a^2) ,$$

onde a é uma constante, pertence a \mathcal{S} mas não a \mathcal{D} . Observamos, de fato, que qualquer polinômio e, mais geralmente, qualquer função analítica sobre \mathbb{R}^n tem suporte igual ao \mathbb{R}^n inteiro, devido ao princípio dos zeros isolados, e portanto não pertence ao espaço \mathcal{D} .

Destarte, torna-se um problema não-trivial mostrar que o espaço \mathcal{D} não se reduz ao espaço trivial $\{0\}$ e, mais do que isso, que os espaços $\mathcal{D}(\Omega)$ contêm um “número suficiente” de funções não-triviais para serem úteis.

A construção de funções não-triviais nos espaços $\mathcal{D}(\Omega)$ é baseada no seguinte fato, que é fundamental para toda a teoria:

Proposição 3.9 *A função $\varphi_0 : \mathbb{R} \rightarrow \mathbb{R}$ definida por*

$$\varphi_0(x) = \begin{cases} \exp(-1/x) & \text{para } x > 0 \\ 0 & \text{para } x \leq 0 \end{cases} \quad (3.62)$$

é de classe C^∞ sobre \mathbb{R} , assim como as suas derivadas, dadas por

$$\varphi_0^{(r)}(x) = \begin{cases} \frac{P_{r-1}(x)}{x^{2r}} \exp(-1/x) & \text{para } x > 0 \\ 0 & \text{para } x \leq 0 \end{cases} \quad (3.63)$$

onde P_{r-1} é um polinômio de grau $r-1$.

Exercício 3.3 *Demonstre essa proposição, mostrando primeiro a validade da fórmula (3.63) para $r \geq 1$ e $x > 0$, por indução sobre r , deduzindo uma relação recursiva entre P_r e P_{r-1} . Em seguida, argumente que esta fórmula permite reduzir a afirmação da proposição ao seguinte*

Lema 3.1 (Singularidades removíveis): *Sejam I um intervalo aberto de \mathbb{R} , $a \in I$, $r \geq 0$ e φ uma função de classe C^r sobre $I \setminus \{a\}$ tal que para $0 \leq s \leq r$, os limites*

$$\varphi_-^{(s)}(a) = \lim_{\substack{x \rightarrow a \\ x < a}} \varphi^{(s)}(x) \quad \text{e} \quad \varphi_+^{(s)}(a) = \lim_{\substack{x \rightarrow a \\ x > a}} \varphi^{(s)}(x) \quad (3.64)$$

existem e são iguais. Então φ admite uma única extensão a uma função de classe C^r sobre I , que também denotaremos por φ , tal que para $0 \leq s \leq r$,

$$\varphi^{(s)}(a) = \varphi_{\pm}^{(s)}(a) . \quad (3.65)$$

DEMONSTRAÇÃO: Novamente, a demonstração utiliza indução sobre r . Como o caso $r = 0$ é familiar, precisamos mostrar que se a afirmação vale para r , ela também vale para $r + 1$.

Seja então φ uma função de classe C^{r+1} sobre $I \setminus \{a\}$ tal que para $0 \leq s \leq r + 1$, os limites na equação (3.64) existem e são iguais. Pela hipótese de indução, φ é de classe C^r sobre I . Além disto, a r -ésima derivada $\varphi^{(r)}$ de φ é diferenciável não somente em $I \setminus \{a\}$, mas também no ponto a , pois pelo teorema fundamental do cálculo, temos para $x \in I$ com $x < y < a$

$$\frac{\varphi^{(r)}(y) - \varphi^{(r)}(x)}{y - x} = \int_0^1 dt \varphi^{(r+1)}(x + t(y - x)) ,$$

e para $x \in I$ com $a < y < x$

$$\frac{\varphi^{(r)}(x) - \varphi^{(r)}(y)}{x - y} = \int_0^1 dt \varphi^{(r+1)}(y + t(x - y)) .$$

Usando a hipótese de que os limites na equação (3.64) existem e assim a função $\varphi^{(r+1)}$ é contínua e portanto uniformemente contínua sobre os subintervalos compactos $[x, a]$ e $[a, x]$ de I , podemos tomar o limite $y \rightarrow a$ para concluir que, para $x \in I$ com $x < a$

$$\frac{\varphi^{(r)}(x) - \varphi^{(r)}(a)}{x - a} = \int_0^1 dt \varphi^{(r+1)}(x + t(a - x)) ,$$

e para $x \in I$ com $a < x$

$$\frac{\varphi^{(r)}(x) - \varphi^{(r)}(a)}{x - a} = \int_0^1 dt \varphi^{(r+1)}(a + t(x - a)) .$$

Da mesma forma, tomando agora o limite $x \rightarrow a$, vem

$$\lim_{\substack{x \rightarrow a \\ x < a}} \frac{\varphi^{(r)}(x) - \varphi^{(r)}(a)}{x - a} = \varphi_-^{(r+1)}(a) ,$$

e

$$\lim_{\substack{x \rightarrow a \\ x > a}} \frac{\varphi^{(r)}(x) - \varphi^{(r)}(a)}{x - a} = \varphi_+^{(r+1)}(a) .$$

Disto, decorre imediatamente que φ é de classe C^{r+1} sobre I , com

$$\varphi^{(r+1)}(a) = \varphi_{\pm}^{(r+1)}(a) .$$

□

Usando a função φ_0 definida pela equação (3.62), podemos construir um grande número de funções teste nos espaços $\mathcal{D}(\Omega)$. Por exemplo, para qualquer intervalo fechado $[a, b]$ de \mathbb{R} , a fórmula

$$\varphi_{a,b}(x) = \begin{cases} \exp\left(\frac{1}{(x-a)(x-b)}\right) & \text{para } a < x < b \\ 0 & \text{para } x \leq a \text{ ou } x \geq b \end{cases} \quad (3.66)$$

fornece uma função $\varphi_{a,b} \in \mathcal{D}$ com suporte $\text{supp } \varphi_{a,b} = [a, b]$, pois

$$\varphi_{a,b}(x) = \varphi_0((b-a)(x-a)) \varphi_0((b-a)(b-x)). \quad (3.67)$$

Integrando, obtemos uma função $\phi_{a,b} \in \mathcal{E}$ tal que $\phi_{a,b}(x) = 0$ para $x \leq a$ e $\phi_{a,b}(x) = 1$ para $x \geq b$:

$$\phi_{a,b}(x) = \int_{-\infty}^x dy \varphi_{a,b}(y) / \int_{-\infty}^{+\infty} dy \varphi_{a,b}(y). \quad (3.68)$$

Finalmente, dado qualquer intervalo fechado $[a, b]$ de \mathbb{R} e qualquer $\epsilon > 0$ ($\epsilon \ll b-a$), podemos definir uma função $\chi_{a,b,\epsilon} \in \mathcal{D}$ tal que $\chi_{a,b,\epsilon} = 1$ sobre o intervalo fechado $[a, b]$ e $\text{supp } \chi_{a,b,\epsilon} = [a-\epsilon, b+\epsilon]$, pondo

$$\chi_{a,b,\epsilon}(x) = \phi_{a-\epsilon,a}(x) (1 - \phi_{b,b+\epsilon}(x)). \quad (3.69)$$

Em n dimensões ($n > 1$), dado qualquer vetor $a \in \mathbb{R}^n$ e quaisquer $r > 0$ e $\epsilon > 0$ ($\epsilon \ll r$), podemos definir uma função $\chi_{a,r,\epsilon} \in \mathcal{D}$ tal que $\chi_{a,r,\epsilon} = 1$ sobre a bola fechada $\bar{B}_r(a)$ de raio r em torno de a e $\text{supp } \chi_{a,r,\epsilon} = \bar{B}_{r+\epsilon}(a)$, pondo

$$\chi_{a,r,\epsilon}(x) = 1 - \phi_{r,r+\epsilon}(|x-a|). \quad (3.70)$$

Este resultado pode ser generalizado, substituindo a bola fechada $\bar{B}_r(a)$ por qualquer subconjunto compacto K de \mathbb{R}^n e a bola aberta $B_{r+2\epsilon}(a)$, digamos, por qualquer subconjunto aberto Ω de \mathbb{R}^n tal que $K \subset \Omega$:

Teorema 3.4 *Para qualquer compacto $K \subset \mathbb{R}^n$ e qualquer aberto $\Omega \subset \mathbb{R}^n$ com $K \subset \Omega$, existe uma função $\chi_{K,\Omega} \in \mathcal{D}(\Omega)$ tal que $0 \leq \chi_{K,\Omega} \leq 1$ e $\chi_{K,\Omega} = 1$ sobre K (e até sobre uma vizinhança de K , de modo que $\text{supp}(1 - \chi_{K,\Omega}) \cap K = \emptyset$).*

DEMONSTRAÇÃO: Inicialmente, para todo ponto $a \in K$, escolhemos $r(a) > 0$ e $\epsilon(a) > 0$ tais que a bola fechada $\bar{B}_{r(a)+2\epsilon(a)}(a)$ de raio $r(a) + 2\epsilon(a)$ em torno de a seja contida em Ω , e consideramos a função $\chi_{a,r(a),\epsilon(a)} \in \mathcal{D}$ definida acima. Como K é compacto, o recobrimento aberto $(B_{r(a)}(a))_{a \in K}$ de K possui um subrecobrimento finito $(B_{r(a_i)}(a_i))_{i=1 \dots N}$. Definimos

$$\chi_{K,\Omega}(x) = 1 - \prod_{i=1}^N (1 - \chi_{a_i,r(a_i),\epsilon(a_i)}(x)).$$

Então $\chi_{K,\Omega}$ possui as propriedades desejadas, tendo como suporte a união finita de bolas fechadas

$$\text{supp } \chi_{K,\Omega} = \bigcup_{i=1}^N \bar{B}_{r(a_i)+\epsilon(a_i)}(a_i) .$$

□

Um argumento semelhante pode ser empregado para provar o Teorema B.2 sobre a existência de partições da unidade, pelo menos no caso mais simples de recobrimentos abertos finitos de compactos de \mathbb{R}^n , do qual o Teorema 3.4 é um caso especial.

Tendo demonstrado a existência de um “número suficiente” de funções teste em cada um dos espaços $\mathcal{D}(\Omega)$, resta como última tarefa nesta seção a de especificar a estrutura topológica dos espaços $\mathcal{D}(\Omega)$, $\mathcal{E}(\Omega)$ e \mathcal{S} . Em particular, queremos formular explicitamente os critérios de convergência para sequências $(\varphi_k)_{k \in \mathbb{N}}$ de funções em cada um deles.

Começamos pelo espaço $\mathcal{E}(\Omega)$, que é localmente convexo em relação à topologia gerada pelas seminormas

$$\|\varphi\|_{r,K} = 2^r \max_{|\alpha| \leq r} \sup_{x \in K} |\partial_\alpha \varphi(x)| \quad \text{para } \varphi \in \mathcal{E}(\Omega) , \quad (3.71)$$

onde $r = 0, 1, 2, \dots$ e K percorre os subconjuntos compactos de Ω : conforme explicado no Exemplo 3.7, esta é a topologia da convergência uniforme de todas as derivadas sobre subconjuntos compactos de Ω , e segundo a Proposição 3.3, temos para qualquer sequência $(\varphi_k)_{k \in \mathbb{N}}$ de funções $\varphi_k \in \mathcal{E}(\Omega)$

$$\varphi_k \rightarrow 0 \text{ em } \mathcal{E}(\Omega) \iff \begin{array}{l} \partial_\alpha \varphi_k \rightarrow 0 \text{ uniformemente sobre } K \\ \text{para todo } \alpha \in \mathbb{N}^n \text{ e todo compacto } K \text{ de } \Omega \end{array} . \quad (3.72)$$

De forma semelhante, o espaço \mathcal{S} é localmente convexo em relação à topologia gerada pelas seminormas

$$\|\varphi\|_{\alpha,\beta} = \sup_{x \in \mathbb{R}^n} |x^\beta \partial_\alpha \varphi(x)| \quad \text{para } \varphi \in \mathcal{S} , \quad (3.73)$$

onde $\alpha, \beta \in \mathbb{N}^n$, e segundo a Proposição 3.3, temos para qualquer sequência $(\varphi_k)_{k \in \mathbb{N}}$ de funções $\varphi_k \in \mathcal{S}$

$$\varphi_k \rightarrow 0 \text{ em } \mathcal{S} \iff \begin{array}{l} x^\beta \partial_\alpha \varphi_k(x) \rightarrow 0 \text{ uniformemente sobre } \mathbb{R}^n \\ \text{para todo } \alpha, \beta \in \mathbb{N}^n \end{array} . \quad (3.74)$$

Para definir a topologia do espaço $\mathcal{D}(\Omega)$, precisamos primeiro especificar a topologia dos espaços $\mathcal{D}(K)$, onde K percorre os subconjuntos compactos de Ω : cada um destes é localmente convexo em relação à topologia gerada pelas normas

$$\|\varphi\|_{r,K} = 2^r \max_{|\alpha| \leq r} \sup_{x \in K} |\partial_\alpha \varphi(x)| \quad \text{para } \varphi \in \mathcal{D}(K) , \quad (3.75)$$

onde $r = 0, 1, 2, \dots$. Note que para quaisquer dois subconjuntos compactos K e K' de Ω , temos

$$K \subset K' \implies \mathcal{D}(K) \subset \mathcal{D}(K')$$

sendo que, neste caso, $\mathcal{D}(K)$ é um subespaço fechado de $\mathcal{D}(K')$ tal que a topologia de $\mathcal{D}(K)$ coincide com a topologia induzida pela topologia de $\mathcal{D}(K')$, uma vez que

$$K \subset K' \implies \|\varphi\|_{r,K} = \|\varphi\|_{r,K'} \quad \text{para } \varphi \in \mathcal{D}(K).$$

O espaço $\mathcal{D}(\Omega)$ é então definido como o “limite indutivo” da família dos subespaços $\mathcal{D}(K)$, onde K percorre os subconjuntos compactos de Ω . Tecnicamente, esta topologia é definida como sendo a mais fina topologia localmente convexa tal que para todo subconjunto compacto K de Ω , a inclusão de $\mathcal{D}(K)$ em $\mathcal{D}(\Omega)$ é contínua. Cabe mencionar que esta topologia não pode ser gerada por um sistema enumerável de seminormas e portanto, ao contrário dos espaços $\mathcal{E}(\Omega)$, \mathcal{S} e $\mathcal{D}(K)$, o espaço $\mathcal{D}(\Omega)$ não é metrizável. No entanto, o critério de convergência de seqüências em $\mathcal{D}(\Omega)$ é fácil de manuseiar, devido ao fato de que uma seqüência em $\mathcal{D}(\Omega)$ converge nesta topologia se e somente se ela for contida e convergente em algum dos subespaços $\mathcal{D}(K)$, ou seja,

$$\varphi_k \rightarrow 0 \text{ em } \mathcal{D}(\Omega) \iff \begin{array}{l} \text{Existe um subconjunto compacto } K \text{ de } \Omega \\ \text{tal que } \text{supp } \varphi_k \subset K \text{ para todo } k \in \mathbb{N} \\ \text{e } \partial_\alpha \varphi_k \rightarrow 0 \text{ uniformemente sobre } K \\ \text{para todo } \alpha \in \mathbb{N}^n \end{array} \quad (3.76)$$

Os critérios de convergência de seqüências são úteis por serem suficientes para provar a continuidade de aplicações lineares, devido à seguinte

Proposição 3.10 (*1º critério de continuidade para aplicações lineares*): *Se E é um dos espaços $\mathcal{D}(\Omega)$, $\mathcal{E}(\Omega)$ ou \mathcal{S} e F é qualquer espaço localmente convexo, então uma aplicação linear $f : E \rightarrow F$ é contínua se e somente se f transforma seqüências convergentes (para 0) em E em seqüências convergentes (para 0) em F :*

$$\varphi_k \rightarrow 0 \text{ em } E \implies f(\varphi_k) \rightarrow 0 \text{ em } F.$$

Observemos que para os espaços $\mathcal{E}(\Omega)$ e \mathcal{S} , a afirmação é uma consequência direta do fato de que são metrizáveis, enquanto que para os espaços $\mathcal{D}(\Omega)$, ela decorre da construção explícita do “limite indutivo”, em conjunto com o fato de que os subespaços $\mathcal{D}(K)$ são metrizáveis.

Na prática, no entanto, é frequentemente mais conveniente provar a continuidade de uma aplicação linear por estimativas em termos de seminormas. Novamente, para os espaços $\mathcal{E}(\Omega)$ e \mathcal{S} , isto pode ser feito diretamente usando as seminormas (3.71) e (3.73) que definem as suas respectivas topologias, enquanto que para os espaços $\mathcal{D}(\Omega)$, temos que usar propriedades da topologia do “limite indutivo”, resumidas nas duas proposições enunciadas logo abaixo, para então podermos trabalhar com as seminormas (3.75).

Proposição 3.11 (*2º critério de continuidade para aplicações lineares*): *Seja Ω um aberto de \mathbb{R}^n . Se F é qualquer espaço localmente convexo, então uma aplicação linear $f : \mathcal{D}(\Omega) \rightarrow F$ é contínua se e somente se a sua restrição a cada um dos subespaços $\mathcal{D}(K)$, onde K é qualquer subconjunto compacto de Ω , for contínua.*

Uma demonstração explícita da equivalência entre os critérios formulados nas Proposições 3.10 e 3.11 no caso especial em que F é o corpo base (\mathbb{R} ou \mathbb{C}) será apresentada logo adiante (veja a Proposição 3.13 abaixo), sendo que o mesmo argumento funciona para qualquer espaço F localmente convexo.

Proposição 3.12 (*3º critério de continuidade para aplicações lineares*): *Sejam Ω um aberto de \mathbb{R}^n , Ω' um aberto de \mathbb{R}^m e $f : \mathcal{D}(\Omega) \rightarrow \mathcal{D}(\Omega')$ uma aplicação linear tal que para todo subconjunto compacto K de Ω , existe um subconjunto compacto K' de Ω' com $f(\mathcal{D}(K)) \subset \mathcal{D}(K')$. Então f é contínua se e somente se para todo subconjunto compacto K de Ω , existe um subconjunto compacto K' de Ω' com $f(\mathcal{D}(K)) \subset \mathcal{D}(K')$ tal que a restrição de f a $\mathcal{D}(K)$ é uma aplicação linear contínua de $\mathcal{D}(K)$ em $\mathcal{D}(K')$.*

Como exemplo elementar, notamos que esta proposição pode ser aplicada para mostrar que se $\tilde{\Omega}$ e Ω são abertos de \mathbb{R}^n , com $\tilde{\Omega} \subset \Omega$, então a inclusão de $\mathcal{D}(\tilde{\Omega})$ em $\mathcal{D}(\Omega)$ é uma aplicação linear contínua.

3.4 O conceito de distribuição

Definição 3.17 Uma **distribuição** sobre um aberto Ω de \mathbb{R}^n é um funcional linear contínuo sobre o espaço de funções teste $\mathcal{D}(\Omega)$. O espaço das distribuições sobre Ω será denotado por $\mathcal{D}'(\Omega)$. Em particular, o espaço $\mathcal{D}'(\mathbb{R}^n)$ das distribuições sobre \mathbb{R}^n será denotado por \mathcal{D}' , quando não houver perigo de confusão.

Portanto, uma distribuição $T \in \mathcal{D}'(\Omega)$ sobre Ω é uma aplicação $T : \mathcal{D}(\Omega) \rightarrow \mathbb{C}$, cujo valor sobre uma função teste $\varphi \in \mathcal{D}(\Omega)$ será denotado por $\langle T, \varphi \rangle$,

$$\begin{aligned} T : \mathcal{D}(\Omega) &\longrightarrow \mathbb{C} \\ \varphi &\longmapsto \langle T, \varphi \rangle \end{aligned} \quad (3.77)$$

e que satisfaz as propriedades de linearidade, i.e.,

$$\langle T, \lambda_1 \varphi_1 + \lambda_2 \varphi_2 \rangle = \lambda_1 \langle T, \varphi_1 \rangle + \lambda_2 \langle T, \varphi_2 \rangle, \quad (3.78)$$

e continuidade, para a qual existem dois critérios úteis (e equivalentes):

Proposição 3.13 *Sejam Ω um aberto de \mathbb{R}^n e $T : \mathcal{D}(\Omega) \rightarrow \mathbb{C}$ um funcional linear sobre $\mathcal{D}(\Omega)$. Então T é contínuo e portanto define uma distribuição sobre Ω se e somente se valem os dois seguintes critérios, que são equivalentes:*

- Para todo subconjunto compacto K de Ω , existem $r \in \mathbb{N}$ e uma constante $C_K > 0$ tais que

$$|\langle T, \varphi \rangle| \leq C_K \|\varphi\|_{r,K} \quad \text{para } \varphi \in \mathcal{D}(K). \quad (3.79)$$

- $\varphi_k \rightarrow 0$ em $\mathcal{D}(\Omega) \implies \langle T, \varphi_k \rangle \rightarrow 0$.

DEMONSTRAÇÃO: Levando em conta a definição da topologia dos espaços $\mathcal{D}(K)$ em termos do sistema de seminormas (3.75), o primeiro critério é apenas uma reformulação da Proposição 3.11. Também é claro que o primeiro critério implica o segundo. Para provar o recíproco, suponha que o primeiro critério não esteja satisfeito. Neste caso, devem existir um subconjunto compacto K de Ω e, para todo $k \in \mathbb{N}$, uma função teste $\varphi_k \in \mathcal{D}(K)$ tal que $|\langle T, \varphi_k \rangle| > k \|\varphi_k\|_{k,K}$. Mas então, pondo $\tilde{\varphi}_k = \varphi_k / k \|\varphi_k\|_{k,K}$, temos que, para todo $r \in \mathbb{N}$ fixo, vale $\|\tilde{\varphi}_k\|_{r,K} \leq \|\tilde{\varphi}_k\|_{k,K} = 1/k$ quando $k \geq r$, implicando que $\tilde{\varphi}_k \rightarrow 0$ em $\mathcal{D}(\Omega)$, enquanto que $|\langle T, \tilde{\varphi}_k \rangle| > 1$. □

Note que $\mathcal{D}'(\Omega)$ é o dual topológico de $\mathcal{D}(\Omega)$, um subespaço do dual algébrico $\mathcal{D}^*(\Omega)$ de $\mathcal{D}(\Omega)$, que é o espaço de todos os funcionais lineares sobre $\mathcal{D}(\Omega)$, contínuos ou não.

Passamos a apresentar exemplos de distribuições, começando pelo fato de que funções comuns podem ser reinterpretadas como distribuições. Mais exatamente, existe uma grande classe de funções que podem ser reinterpretadas como distribuições, a saber, a das funções localmente integráveis:

Definição 3.18 *Uma função mensurável f sobre um aberto Ω de \mathbb{R}^n a valores complexos é dita **localmente integrável** se para todo subconjunto compacto K de Ω , $f|_K \in \mathcal{L}^1(K)$, ou seja,*

$$\int_K d^n x |f(x)| < \infty.$$

O espaço das funções localmente integráveis sobre Ω será denotado por $\mathcal{L}_{loc}^1(\Omega)$. Lembramos também que na teoria da integral de Lebesgue, costuma-se identificar funções que diferem apenas sobre um conjunto de medida zero. O espaço quociente, formado pelas classes de equivalência de funções em $\mathcal{L}_{loc}^1(\Omega)$, neste sentido, será denotado por $L_{loc}^1(\Omega)$.

Cabe ressaltar que o espaço $\mathcal{L}_{loc}^1(\Omega)$ é grande no sentido de conter todos os espaços tradicionais de funções: por exemplo, $\mathcal{L}_{loc}^1(\Omega)$ contém o espaço $C(\Omega)$ das funções

contínuas sobre Ω (veja o Exemplo 3.7) e $\mathcal{L}_{loc}^1(\mathbb{R}^n)$ contém todos os espaços $\mathcal{L}^p(\mathbb{R}^n)$, para $1 \leq p \leq \infty$ (veja os Exemplos 3.3 e 3.4). Na verdade, para que uma função f sobre Ω pertença a $\mathcal{L}_{loc}^1(\Omega)$, basta que f seja mensurável e, exceto sobre um conjunto de medida zero, limitada sobre compactos K de Ω .

Para mostrar que a reinterpretção de funções localmente integráveis como distribuições a ser apresentada no próximo exemplo não acarreta nenhuma perda de informação relevante, precisaremos do seguinte lema provindo da teoria da medida e da integral de Lebesgue:

Lema 3.2 *Seja f uma função localmente integrável sobre um aberto Ω de \mathbb{R}^n tal que*

$$\int_K d^n x f(x) = 0$$

para qualquer subconjunto compacto K de Ω . Então

$$f = 0 \quad \text{quase sempre .}$$

Exemplo 3.8 (Funções localmente integráveis como distribuições): Dada uma função localmente integrável f sobre Ω , definimos para toda função teste $\varphi \in \mathcal{D}(\Omega)$

$$\langle T_f, \varphi \rangle = \int d^n x f(x) \varphi(x) . \quad (3.80)$$

Então T_f é uma distribuição, pois é linear,

$$\begin{aligned} & \langle T_f, \lambda_1 \varphi_1 + \lambda_2 \varphi_2 \rangle \\ &= \int d^n x f(x) (\lambda_1 \varphi_1(x) + \lambda_2 \varphi_2(x)) \\ &= \lambda_1 \int d^n x f(x) \varphi_1(x) + \lambda_2 \int d^n x f(x) \varphi_2(x) \\ &= \lambda_1 \langle T_f, \varphi_1 \rangle + \lambda_2 \langle T_f, \varphi_2 \rangle , \end{aligned}$$

e também é contínua, sendo que para todo subconjunto compacto K de Ω e toda função teste $\varphi \in \mathcal{D}(K)$, temos

$$\begin{aligned} |\langle T_f, \varphi \rangle| &= \left| \int_K d^n x f(x) \varphi(x) \right| \\ &\leq \int_K d^n x |f(x)| |\varphi(x)| \leq \int_K d^n x |f(x)| \cdot \sup_{x \in K} |\varphi(x)| , \end{aligned}$$

ou seja

$$|\langle T_f, \varphi \rangle| \leq \|f\|_K \| \varphi \|_{0,K} . \quad (3.81)$$

Desta forma, obtemos uma aplicação linear

$$\begin{aligned} \mathcal{L}_{loc}^1(\Omega) &\longrightarrow \mathcal{D}'(\Omega) \\ f &\longmapsto T_f \end{aligned} , \quad (3.82)$$

cujo núcleo é igual a

$$\{ f \in \mathcal{L}_{loc}^1(\Omega) \mid f = 0 \text{ quase sempre} \},$$

De fato, é óbvio que se $f = 0$ quase sempre, então $T_f = 0$. Para demonstrar a afirmação recíproca, observamos que segundo o Lema 3.2, é suficiente provar que a integral de f sobre qualquer subconjunto compacto K de Ω se anula. Suponha portanto que K é um subconjunto compacto de \mathbb{R}^n contido em Ω , e escolha um subconjunto aberto Ω' e um subconjunto compacto K' de \mathbb{R}^n tais que $K \subset \Omega' \subset K' \subset \Omega$. Para todo $k \in \mathbb{N}$, ponha $\Omega'_k = \{x \in \Omega' \mid d(x, K) < 1/k\}$ e aplique o Teorema 3.4 para encontrar uma função $\chi_k \in \mathcal{D}(\Omega'_k)$ tal que $0 \leq \chi_k \leq 1$ e $\chi_k = 1$ sobre K . Denotando a função característica de um compacto qualquer L (que vale 1 sobre L e 0 fora de L) por χ_L , concluímos que na medida em que $k \rightarrow \infty$, $f\chi_k \rightarrow f\chi_K$ pontualmente e $|f\chi_k| \leq |f|\chi_{K'}$ com $|f|\chi_{K'} \in \mathcal{L}^1(\mathbb{R}^n)$; assim, segundo o teorema da convergência dominada,

$$0 = \langle T_f, \chi_k \rangle = \int d^n x f(x) \chi_k(x) \longrightarrow \int_K d^n x f(x).$$

Assim, obtemos uma aplicação linear injetora

$$\begin{array}{ccc} L_{loc}^1(\Omega) & \longrightarrow & \mathcal{D}'(\Omega) \\ [f] & \longmapsto & T_f \end{array} \quad (3.83)$$

Frequentemente, e sempre quando for conveniente, utilizaremos esta correspondência como uma identidade, identificando T_f com f (ou mais exatamente, com a classe de equivalência $[f]$ de f). Assim, $L_{loc}^1(\Omega)$ pode ser considerado como um subespaço de $\mathcal{D}'(\Omega)$.

Como resultado desta construção, torna-se importante dispor de um critério prático para decidir se uma determinada função é localmente integrável ou não. Mais concretamente, a questão se coloca quando consideramos funções sobre \mathbb{R}^n , ou mais geralmente sobre um aberto Ω de \mathbb{R}^n , que são definidas e contínuas sobre Ω exceto em algum subconjunto S que é de medida zero e que interpretamos como sendo o conjunto das suas singularidades. Então é claro que uma tal função é localmente integrável sobre Ω se ela for limitada sobre $\Omega \setminus S$; um exemplo típico seria o seguinte:

Exemplo 3.9 (Função de Heaviside): A função de Heaviside é a função localmente integrável θ sobre \mathbb{R} definida por

$$\theta(x) = \begin{cases} 1 & \text{para } x > 0 \\ 0 & \text{para } x < 0 \end{cases} \quad (3.84)$$

Note que não é preciso especificar o valor de θ em $x = 0$ pois para qualquer cálculo no contexto de distribuições, este valor é irrelevante.

Assim, uma questão interessante é o que pode ser dito sobre uma função que, no complemento do conjunto das suas singularidades, é contínua mas não é limitada. Uma resposta é dada pela seguinte

Proposição 3.14 *Sejam Ω um aberto de \mathbb{R}^n e f uma função mensurável sobre Ω que é contínua sobre $\Omega \setminus S$ onde S é um subconjunto de Ω de medida zero. Suponha que $(\Omega_k)_{k \in \mathbb{N}}$ é uma sequência decrescente de subconjuntos abertos de Ω cuja interseção é S , i.e., que satisfazem*

$$\Omega \supset \Omega_1 \supset \dots \supset \Omega_k \supset \Omega_{k+1} \supset \dots \supset S \quad \text{e} \quad \bigcap_{k \in \mathbb{N}} \Omega_k = S,$$

e tal que, para todo subconjunto compacto K de Ω , vale

$$\lim_{k \rightarrow \infty} \text{vol}(K \cap \Omega_k) = 0.$$

Então f é localmente integrável sobre Ω se e somente se, para todo subconjunto compacto K de Ω , a sequência crescente $(I_{K,k})_{k \in \mathbb{N}}$ de integrais

$$I_{K,k} = \int_{K \setminus \Omega_k} d^n x |f(x)|$$

for limitada (e portanto convergente).

DEMONSTRAÇÃO: Obviamente, se f for localmente integrável sobre Ω , cada uma das integrais $I_{K,k}$ é limitada pela integral de $|f|$ sobre K . Para provar a afirmação recíproca, fixamos um subconjunto compacto qualquer K de Ω e, denotando a função característica de um compacto qualquer L (que vale 1 sobre L e 0 fora de L) por χ_L , como antes, pomos $g_{K,k} = \chi_{K \setminus \Omega_k} |f|$ para obter uma sequência $(g_{K,k})_{k \in \mathbb{N}}$ crescente de funções g_k reais não-negativas e integráveis sobre K que converge pontualmente para a função $|f|$ sobre $K \setminus S$ (e se anula identicamente sobre $K \cap S$). Então se o critério de limitação enunciado na proposição é satisfeito, segue do teorema da convergência monótona, em conjunto com a hipótese de que S tenha medida zero, que essa convergência se dá também na norma do espaço $\mathcal{L}^1(K)$ e que o limite é uma função que pertence a $\mathcal{L}^1(K)$. □

Na prática, a sequência $(\Omega_k)_{k \in \mathbb{N}}$ de abertos que figura nesta proposição é frequentemente obtida como parte de uma família $(\Omega_\epsilon(S))_{\epsilon > 0}$ de vizinhanças abertas do conjunto S de singularidades de f definidas por

$$\Omega_\epsilon(S) = \{x \in \Omega \mid d(x, S) < \epsilon\}$$

pondo $\Omega_k = \Omega_{\epsilon_k}(S)$, onde $(\epsilon_k)_{k \in \mathbb{N}}$ é uma sequência numérica decrescente qualquer que converge para 0.

De qualquer modo, a reinterpretção de funções como distribuições sugere considerar distribuições como “funções generalizadas” (e de fato foi este o termo adotado pela “escola russa” liderada por Israel Gel’fand) e assim serve de justificativa para

uma convenção, bastante utilizada na física, segundo a qual o valor $\langle T, \varphi \rangle$ de uma distribuição $T \in \mathcal{D}'(\Omega)$ sobre uma função teste $\varphi \in \mathcal{D}(\Omega)$ é denotado por

$$\langle T, \varphi \rangle \equiv \int d^n x T(x) \varphi(x), \quad (3.85)$$

mesmo quando T não corresponde a nenhuma função localmente integrável e, portanto, a integral que aparece nesta fórmula não é uma integral no sentido de Lebesgue. Também escrevemos

$$\langle T, \varphi \rangle \equiv \langle T(x), \varphi(x) \rangle, \quad (3.86)$$

onde x é apenas um rótulo para indicar qual é o argumento de T e de φ , ou seja, a variável independente: portanto, podemos empregar qualquer símbolo ao invés de x , desde que utilizamos o mesmo símbolo para o argumento de T e o argumento de φ . (É a mesma convenção que rege a utilização de índices em somatórias ou de variáveis de integração.) A rigor, trata-se de um abuso de notação, que não deve ser mal entendido: não sugere que uma distribuição pudesse assumir um valor definido em algum ponto do seu domínio.

Passamos a apresentar alguns exemplos de distribuições que não correspondem a funções localmente integráveis.

Exemplo 3.10 (Distribuição de Dirac e suas derivadas): A distribuição de Dirac, também chamada de *função delta de Dirac*, é definida como o funcional que associa a cada função teste $\varphi \in \mathcal{D}$ o seu valor na origem:

$$\langle \delta, \varphi \rangle = \varphi(0). \quad (3.87)$$

As derivadas da distribuição de Dirac ou *derivadas da função delta de Dirac* são definidas como os funcionais que, a menos de um sinal, associam a cada função teste $\varphi \in \mathcal{D}$ o valor das derivadas correspondentes na origem:

$$\langle \delta^{(\alpha)}, \varphi \rangle = (-1)^{|\alpha|} \partial_\alpha \varphi(0). \quad (3.88)$$

Deixamos ao leitor a tarefa de provar que δ e $\delta^{(\alpha)}$ de fato são distribuições.

Na Seção 3.6, veremos em qual sentido $\delta^{(\alpha)}$ é realmente a derivada parcial de ordem α de δ .

Exemplo 3.11 (Valor principal de $1/x$): Inicialmente, observa-se que, conforme o critério da Proposição 3.14, a função $1/x$ sobre \mathbb{R} , que é contínua exceto na origem, não é localmente integrável, pois sua integral sobre o intervalo $[\epsilon, 1]$, onde $0 < \epsilon < 1$, vale

$$\int_\epsilon^1 \frac{dx}{x} = -\ln \epsilon$$

e portanto diverge no limite $\epsilon \rightarrow 0$. Pela mesma razão, a tentativa de definir uma distribuição $1/x$ pela expressão

$$\left\langle \frac{1}{x}, \varphi(x) \right\rangle = \int dx \frac{\varphi(x)}{x} \quad \text{para } \varphi \in \mathcal{D}$$

falha, pois (como veremos logo adiante) esta integral só existe quando $\varphi(0) = 0$. Portanto, encontrar uma definição da função $1/x$ como distribuição é uma tarefa não-trivial, que requer “regularizar” esta expressão, ou seja: queremos encontrar uma distribuição em \mathcal{D}' cuja restrição ao subespaço fechado de codimensão 1

$$\ker \delta = \{ \varphi \in \mathcal{D} \mid \varphi(0) = 0 \}$$

de \mathcal{D} seja igual ao funcional dado por esta expressão. Isso pode ser feito por subtração do termo constante $\varphi(0)$,

$$\left\langle \text{v.p.} \frac{1}{x}, \varphi(x) \right\rangle = \int dx \frac{\varphi(x) - \varphi(0)}{x}, \quad (3.89)$$

desde que a expressão no lado direito desta fórmula seja entendida como o limite

$$\int dx \frac{\varphi(x) - \varphi(0)}{x} = \lim_{a \rightarrow \infty} \int_{|x| \leq a} dx \frac{\varphi(x) - \varphi(0)}{x}.$$

Para justificar esta fórmula, observamos que, para toda função φ em \mathcal{D} (ou mesmo em \mathcal{E}), existe uma única função ψ em \mathcal{E} tal que

$$\varphi(x) = \varphi(0) + x\psi(x).$$

De fato, podemos definir ψ pela fórmula integral

$$\psi(x) = \int_0^1 ds \varphi'(sx),$$

pois esta implica

$$x\psi(x) = \int_0^1 ds x \varphi'(sx) = \int_0^1 ds \frac{d}{ds} \varphi(sx) = \varphi(x) - \varphi(0).$$

Portanto, se $a \gg \epsilon_+, \epsilon_- > 0$, teremos

$$\begin{aligned} & \int_{-a}^{-\epsilon_-} dx \frac{\varphi(x)}{x} + \int_{\epsilon_+}^a dx \frac{\varphi(x)}{x} \\ &= \varphi(0) \left(\int_{-a}^{-\epsilon_-} \frac{dx}{x} + \int_{\epsilon_+}^a \frac{dx}{x} \right) + \int_{-a}^{-\epsilon_-} dx \psi(x) + \int_{\epsilon_+}^a dx \psi(x) \\ &= \varphi(0) (\ln \epsilon_- - \ln \epsilon_+) + \int_{-a}^{-\epsilon_-} dx \psi(x) + \int_{\epsilon_+}^a dx \psi(x). \end{aligned}$$

No limite $\epsilon_+ \rightarrow 0$ e $\epsilon_- \rightarrow 0$, os últimos dois termos são convergentes mas o termo logarítmico é divergente, exceto quando impomos uma relação de tipo $\epsilon_-/\epsilon_+ = \exp(\lambda)$ onde λ é uma constante. Escolhendo $\lambda = 0$ e a suficientemente

grande para que valha $\text{supp } \varphi \subset [-a, a]$, obtemos a seguinte fórmula alternativa para a definição do valor principal de $1/x$:

$$\left\langle \text{v.p.} \frac{1}{x}, \varphi(x) \right\rangle = \lim_{\epsilon \rightarrow 0} \int_{|x| \geq \epsilon} dx \frac{\varphi(x)}{x}. \quad (3.90)$$

Deixamos ao leitor a tarefa de provar que $\text{v.p.}(1/x)$ de fato é uma distribuição.

Obviamente, a construção do valor principal de $1/x$, que remonta a Cauchy, proporciona uma regularização específica da função $1/x$. Ademais, é claro que qualquer outra regularização da mesma função deve ter a forma

$$\text{v.p.} \frac{1}{x} + \lambda \delta(x),$$

com alguma constante $\lambda \in \mathbb{C}$ (a mesma que antes, caso ela for real).

Exemplo 3.12 (Parte finita de $1/x^k$, $k > 1$): Como no exemplo anterior, observa-se que, conforme o critério da Proposição 3.14, a função $1/x^k$ sobre \mathbb{R} , que é contínua exceto na origem, não é localmente integrável, pois sua integral sobre o intervalo $[\epsilon, 1]$, onde $0 < \epsilon < 1$, vale

$$\int_{\epsilon}^1 \frac{dx}{x^k} = \frac{1}{k-1} \left(\frac{1}{\epsilon^{k-1}} - 1 \right)$$

e portanto diverge no limite $\epsilon \rightarrow 0$. Pela mesma razão, a tentativa de definir uma distribuição $1/x^k$ pela expressão

$$\left\langle \frac{1}{x^k}, \varphi(x) \right\rangle = \int dx \frac{\varphi(x)}{x^k} \quad \text{para } \varphi \in \mathcal{D}$$

falha, sendo que esta integral em geral não existe: novamente, encontrar uma definição da função $1/x^k$ como distribuição requer “regularizar” esta expressão. Aqui, isso pode ser feito por subtração do polinômio de Taylor de φ de ordem $k-1$ na origem:

$$\left\langle \text{p.f.} \frac{1}{x^k}, \varphi(x) \right\rangle = \int dx \frac{1}{x^k} \left(\varphi(x) - \sum_{l=0}^{k-1} \frac{\varphi^{(l)}(0)}{l!} x^l \right), \quad (3.91)$$

desde que, novamente, a expressão no lado direito desta fórmula seja entendida como o limite

$$\int dx \frac{1}{x^k} \left(\varphi(x) - \sum_{l=0}^{k-1} \frac{\varphi^{(l)}(0)}{l!} x^l \right) = \lim_{a \rightarrow \infty} \int_{|x| \leq a} dx \frac{1}{x^k} \left(\varphi(x) - \sum_{l=0}^{k-1} \frac{\varphi^{(l)}(0)}{l!} x^l \right).$$

Para justificar esta fórmula, observamos que, para $k \geq 1$ e para toda função φ em \mathcal{D} , existe uma única função $\psi_k[\varphi]$ em \mathcal{E} tal que

$$\varphi(x) = \sum_{l=0}^{k-1} \frac{\varphi^{(l)}(0)}{l!} x^l + x^k \psi_k[\varphi](x).$$

De fato, segundo o teorema de Taylor, podemos definir $\psi_k[\varphi]$ pela fórmula integral

$$\psi_k[\varphi](x) = \frac{1}{(k-1)!} \int_0^1 ds \varphi^{(k)}(sx) (1-s)^{k-1},$$

pois esta implica, por integração explícita,

$$x^{k+1} \psi_{k+1}[\varphi](x) - x^k \psi_k[\varphi](x) = - \frac{\varphi^{(k)}(0)}{k!} x^k,$$

e assim, por indução sobre k ,

$$\psi_k[\varphi](x) = \frac{1}{x^k} \left(\varphi(x) - \sum_{l=0}^{k-1} \frac{\varphi^{(l)}(0)}{l!} x^l \right).$$

Portanto, se $a \gg \epsilon_+, \epsilon_- > 0$, teremos

$$\begin{aligned} & \int_{-a}^{-\epsilon_-} dx \frac{\varphi(x)}{x^k} + \int_{\epsilon_+}^a dx \frac{\varphi(x)}{x^k} \\ &= \sum_{l=0}^{k-1} \frac{\varphi^{(l)}(0)}{l!} \left(\int_{-a}^{-\epsilon_-} \frac{dx}{x^{k-l}} + \int_{\epsilon_+}^a \frac{dx}{x^{k-l}} \right) \\ & \quad + \int_{-a}^{-\epsilon_-} dx \psi_k[\varphi](x) + \int_{\epsilon_+}^a dx \psi_k[\varphi](x) \\ &= \sum_{l=0}^{k-2} \frac{\varphi^{(l)}(0)}{l!(k-l-1)} \left(\frac{1}{\epsilon_+^{k-l-1}} - \frac{1}{(-\epsilon_-)^{k-l-1}} - \frac{1}{a^{k-l-1}} + \frac{1}{(-a)^{k-l-1}} \right) \\ & \quad - \frac{\varphi^{(k-1)}(0)}{(k-1)!} (\ln \epsilon_+ - \ln \epsilon_-) \\ & \quad + \int_{-a}^{-\epsilon_-} dx \psi_k[\varphi](x) + \int_{\epsilon_+}^a dx \psi_k[\varphi](x). \end{aligned}$$

No limite $\epsilon_+ \rightarrow 0$ e $\epsilon_- \rightarrow 0$, os últimos dois termos são convergentes mas o termo logarítmico é divergente, exceto quando impomos uma relação de tipo $\epsilon_-/\epsilon_+ = \exp(\lambda)$ onde λ é uma constante. Escolhendo $\lambda = 0$ e tomando o limite $a \rightarrow \infty$, obtemos a seguinte fórmula alternativa para a definição da parte finita de $1/x^k$:

$$\left\langle \text{p.f.} \frac{1}{x^k}, \varphi(x) \right\rangle = \lim_{\epsilon \rightarrow 0} \left(\int_{|x| \geq \epsilon} dx \frac{\varphi(x)}{x^k} - \sum_{\substack{l=0 \\ k-l \text{ ímpar}}}^{k-2} \frac{2 \varphi^{(l)}(0)}{l!(k-l-1)} \frac{1}{\epsilon^{k-l-1}} \right). \quad (3.92)$$

Deixamos ao leitor a tarefa de provar que p.f. $(1/x^k)$ de fato é uma distribuição.

Novamente, a construção da parte finita de $1/x^k$ proporciona uma regularização específica da função $1/x^k$ e qualquer outra regularização da mesma função deve ter a forma

$$\text{p.f. } \frac{1}{x^k} + \sum_{l=0}^{k-1} \lambda_l \delta^{(l)}(x),$$

com constantes $\lambda_0, \dots, \lambda_{k-1} \in \mathbb{C}$.

Voltando à notação introduzida pelas equações (3.85) e (3.86), ressaltamos mais uma vez que para uma distribuição T qualquer sobre um aberto Ω de \mathbb{R}^n , não existe, em geral, a noção do “valor” de T em um ponto do seu domínio Ω . O que existe sim é a “média ponderada” de T sobre qualquer região finita em Ω com volume > 0 . De fato, para uma função localmente integrável f sobre Ω , a média de f sobre um subconjunto compacto K de Ω é a expressão

$$\bar{f}_K = \frac{1}{\text{vol}(K)} \int_K d^n x f(x) = \frac{1}{\text{vol}(K)} \int d^n x f(x) \chi_K(x),$$

onde χ_K é a função característica de K (que vale 1 sobre K e vale 0 fora de K), enquanto que a média ponderada sobre K , com uma função peso φ , seria

$$\frac{1}{\text{vol}(K)} \int_K d^n x f(x) \varphi(x).$$

Portanto, podemos dizer que distribuições são entidades matemáticas para as quais existe a média ponderada sobre qualquer região finita com volume > 0 e qualquer função peso de classe C^∞ .

3.5 Suporte e suporte singular

A nossa meta nesta seção será generalizar o conceito de suporte de funções para distribuições. Veremos que, como no caso de funções, o suporte de uma distribuição indica a região na qual ela não se anula. Veremos também que praticamente a mesma construção permite definir o conceito de suporte singular de uma distribuição, que indica a região na qual ela é singular, no sentido de não se reduzir a uma função lisa.

Ambas as definições baseiam-se na observação de que, apesar do fato de que distribuições não assumem valores específicos nos pontos do seu domínio, a afirmação de que uma distribuição se anula ou se reduz a uma função lisa sobre um determinado subconjunto de \mathbb{R}^n faz sentido se (e em geral somente se) este for um aberto.

Definição 3.19 *Sejam $\tilde{\Omega}$ e Ω abertos de \mathbb{R}^n , com $\tilde{\Omega} \subset \Omega$. Então a inclusão*

$$\mathcal{D}(\tilde{\Omega}) \longrightarrow \mathcal{D}(\Omega), \quad (3.93)$$

sendo uma aplicação linear contínua, induz uma projeção

$$\begin{aligned} \mathcal{D}'(\Omega) &\longrightarrow \mathcal{D}'(\tilde{\Omega}) \\ T &\longmapsto T|_{\tilde{\Omega}} \end{aligned} \quad (3.94)$$

A distribuição $T|_{\tilde{\Omega}}$ é chamada a **restrição** da distribuição T a $\tilde{\Omega}$. Dizemos que T **se anula** ou **é zero sobre** ou **em** $\tilde{\Omega}$ se $T|_{\tilde{\Omega}} = 0$ e que T é **regular sobre** ou **em** $\tilde{\Omega}$ se existe uma função lisa $f \in \mathcal{E}(\tilde{\Omega})$ tal que $T|_{\tilde{\Omega}} = T_f$ em $\mathcal{D}'(\tilde{\Omega})$.

A primeira parte desta definição se justifica pelo fato de que para distribuições T_f provindas de funções contínuas f , temos

Proposição 3.15 *Sejam $\tilde{\Omega}$ e Ω abertos de \mathbb{R}^n , com $\tilde{\Omega} \subset \Omega$, e seja f uma função contínua sobre Ω . Então f se anula em $\tilde{\Omega}$ se e somente se T_f se anula em $\tilde{\Omega}$.*

DEMONSTRAÇÃO: Se f se anula em $\tilde{\Omega}$, então é óbvio que para qualquer função teste $\varphi \in \mathcal{D}(\tilde{\Omega})$ com suporte em $\tilde{\Omega}$, vale $\langle T_f, \varphi \rangle = 0$. Reciprocamente, suponhamos que T_f se anula em $\tilde{\Omega}$, e vamos mostrar que para qualquer ponto $x \in \tilde{\Omega}$ e qualquer $\epsilon > 0$, vale $|f(x)| \leq \epsilon$. Para tanto, usando a continuidade de f no ponto x , escolhemos $\delta > 0$ tal que $\bar{B}_x(\delta) \subset \tilde{\Omega}$ e $|f(y) - f(x)| \leq \epsilon$ para $y \in \bar{B}_x(\delta)$, assim como uma função teste $\varphi \in \mathcal{D}$ tal que $\varphi \geq 0$, $\int d^n y \varphi(y) = 1$ e $\text{supp } \varphi \subset \bar{B}_x(\delta)$. Então $\langle T_f, \varphi \rangle = 0$ e portanto

$$\begin{aligned} |f(x)| &= \left| f(x) \int d^n y \varphi(y) \right| \\ &= \left| \int d^n y f(y) \varphi(y) - \int d^n y (f(y) - f(x)) \varphi(y) \right| \\ &\leq \langle T_f, \varphi \rangle + \int d^n y |f(y) - f(x)| \varphi(y) \\ &\leq \sup_{y \in \bar{B}_x(\delta)} |f(y) - f(x)| \cdot \int d^n y \varphi(y) \\ &\leq \epsilon. \end{aligned}$$

□

Para a segunda parte da definição acima, observamos que a função f que aparece na condição de regularidade precisa ser de classe C^∞ apenas em $\tilde{\Omega}$, podendo desenvolver singularidades (de qualquer natureza, integráveis ou não) na fronteira de $\tilde{\Omega}$. A proposição acima mostra então que se T é regular em $\tilde{\Omega}$, a função $f \in \mathcal{E}(\tilde{\Omega})$ tal que $T|_{\tilde{\Omega}} = T_f$ em $\mathcal{D}'(\tilde{\Omega})$ é unicamente determinada por T .

Obviamente, se uma distribuição se anula/é regular em algum aberto, ela também se anula/é regular em qualquer subconjunto aberto deste. Na direção oposta, temos a seguinte

Proposição 3.16 *Sejam $\tilde{\Omega}$ e Ω abertos de \mathbb{R}^n , com $\tilde{\Omega} \subset \Omega$, e tais que $\tilde{\Omega}$ é a união de uma família $(\tilde{\Omega}_i)_{i \in I}$ de abertos de \mathbb{R}^n . Se uma distribuição T sobre Ω se anula / é regular em cada um dos $\tilde{\Omega}_i$, então ela também se anula / é regular em $\tilde{\Omega}$.*

DEMONSTRAÇÃO: Primeiro, observamos que podemos demonstrar as duas afirmações contidas nesta proposição simultaneamente, pois uma distribuição que se anula em um determinado aberto é um caso particular de uma distribuição que é regular neste aberto: corresponde à escolha $f \equiv 0$. Ademais, tendo em vista a observação anterior, podemos supor sem perda de generalidade que a família $(\tilde{\Omega}_i)_{i \in I}$ constitui um recobrimento localmente finito de $\tilde{\Omega}$. (Caso contrário, podemos substituí-la por um refinamento localmente finito.) Sendo assim, podemos escolher uma partição da unidade $(\chi_i)_{i \in I}$ subordinada ao recobrimento $(\tilde{\Omega}_i)_{i \in I}$ de $\tilde{\Omega}$.

Isso posto, observamos que se para todo $i \in I$, T é regular em $\tilde{\Omega}_i$, i.e., se para todo $i \in I$, existe $f_i \in \mathcal{E}(\tilde{\Omega}_i)$ tal que $T|_{\tilde{\Omega}_i} = T_{f_i}$ em $\mathcal{D}'(\tilde{\Omega}_i)$, então para quaisquer dois índices $i, j \in I$, vale

$$T_{f_i}|_{\tilde{\Omega}_i \cap \tilde{\Omega}_j} = T|_{\tilde{\Omega}_i \cap \tilde{\Omega}_j} = T_{f_j}|_{\tilde{\Omega}_i \cap \tilde{\Omega}_j} \quad \text{em } \mathcal{D}'(\tilde{\Omega}_i \cap \tilde{\Omega}_j).$$

Pela propriedade de unicidade mencionada acima, segue que $f_i = f_j$ em $\tilde{\Omega}_i \cap \tilde{\Omega}_j$. Logo, existe $f \in \mathcal{E}(\tilde{\Omega})$ tal que, para todo $i \in I$, f_i é a restrição de f a $\tilde{\Omega}_i$. Falta mostrar então que para qualquer função teste $\varphi \in \mathcal{D}(\tilde{\Omega})$, vale $\langle T|_{\tilde{\Omega}} - T_f, \varphi \rangle = 0$.

Seja portanto $\varphi \in \mathcal{D}(\tilde{\Omega})$ qualquer função teste. Para todo ponto x em $\text{supp } \varphi$, existem uma vizinhança aberta $\Omega_x \subset \tilde{\Omega}$ de x e um subconjunto finito I_x de I tal que $\tilde{\Omega}_i \cap \Omega_x = \emptyset$ se $i \notin I_x$. Como $\text{supp } \varphi$ é compacto, o recobrimento aberto $(\Omega_x)_{x \in \text{supp } \varphi}$ de $\text{supp } \varphi$ possui um subrecobrimento finito $(\Omega_{x_k})_{k=1, \dots, N}$. Seja

$$I_0 = \bigcup_{k=1}^N I_{x_k} \quad \text{e} \quad \Omega_0 = \bigcup_{k=1}^N \Omega_{x_k};$$

então I_0 é um subconjunto finito de I e Ω_0 é um aberto de \mathbb{R}^n tal que $\text{supp } \varphi \subset \Omega_0 \subset \tilde{\Omega}$ e

$$i \notin I_0 \implies \tilde{\Omega}_i \cap \Omega_0 = \emptyset \implies \chi_i \varphi = 0.$$

Assim obtemos

$$\sum_{i \in I} \chi_i = 1 \implies \varphi = \sum_{i \in I} \chi_i \varphi = \sum_{i \in I_0} \chi_i \varphi$$

e portanto

$$\langle T - T_f, \varphi \rangle = \sum_{i \in I_0} \langle T - T_{f_i}, \chi_i \varphi \rangle = 0,$$

pois $\text{supp } (\chi_i \varphi) = \text{supp } \chi_i \cap \text{supp } \varphi \subset \tilde{\Omega}_i$. □

Esta proposição mostra que distribuições e funções comuns compartilham uma propriedade que para funções é completamente óbvia: a existência de um maior aberto onde se anulam. De forma análoga, uma distribuição admite um maior aberto onde é regular. Portanto, podemos estender a Definição B.5 como segue:

Definição 3.20 *O suporte de uma distribuição T sobre um aberto Ω de \mathbb{R}^n , denotado por $\text{supp } T$, é o complemento, em Ω , do maior aberto de Ω onde T se anula:*

$$\text{supp } T = \Omega \setminus \left(\text{maior aberto de } \Omega \text{ onde } T \text{ se anula} \right). \quad (3.95)$$

O suporte singular de uma distribuição T sobre um aberto Ω de \mathbb{R}^n , denotado por $\text{sing supp } T$, é o complemento, em Ω , do maior aberto de Ω onde T é regular:

$$\text{sing supp } T = \Omega \setminus \left(\text{maior aberto de } \Omega \text{ onde } T \text{ é regular} \right). \quad (3.96)$$

Note que, conforme esta definição, o suporte e o suporte singular de uma distribuição sobre um aberto Ω de \mathbb{R}^n é fechado em Ω , mas não necessariamente em \mathbb{R}^n .

A seguinte propriedade é óbvia:

Proposição 3.17 *Para qualquer distribuição T sobre um aberto Ω de \mathbb{R}^n , vale*

$$\text{supp } T = \emptyset \iff T = 0. \quad (3.97)$$

$$\text{sing supp } T = \emptyset \iff T = T_f \text{ com } f \in \mathcal{E}(\Omega). \quad (3.98)$$

Como corolário da Proposição 3.15, temos a primeira parte da seguinte

Proposição 3.18 *Para qualquer função contínua f sobre um aberto Ω de \mathbb{R}^n , vale*

$$\text{supp } T_f = \text{supp } f. \quad (3.99)$$

$$\text{sing supp } T_f = \text{sing supp } f, \quad (3.100)$$

onde o suporte singular de f é definido como sendo o complemento do conjunto dos pontos x de Ω tais que f é de classe C^∞ em alguma vizinhança aberta de x .

Exemplo 3.13 O suporte e o suporte singular da distribuição de Dirac e de suas derivadas é a origem:

$$\text{supp } \delta = \{0\}, \quad \text{sing supp } \delta = \{0\}. \quad (3.101)$$

$$\text{supp } \delta^{(\alpha)} = \{0\}, \quad \text{sing supp } \delta^{(\alpha)} = \{0\}. \quad (3.102)$$

O suporte e o suporte singular do valor principal de $1/x$ e da parte finita de $1/x^k$ são a reta inteira e a origem, respectivamente:

$$\text{supp v.p. } \frac{1}{x} = \mathbb{R}, \quad \text{sing supp v.p. } \frac{1}{x} = \{0\}. \quad (3.103)$$

$$\text{supp p.f. } \frac{1}{x^k} = \mathbb{R}, \quad \text{sing supp p.f. } \frac{1}{x^k} = \{0\}. \quad (3.104)$$

De fato, v.p. $(1/x)$ e p.f. $(1/x^k)$ são obviamente de classe C^∞ sobre $\mathbb{R} \setminus \{0\}$, mas não sobre \mathbb{R} .

Um teorema interessante devido a Laurent Schwartz que apenas mencionamos aqui, sem demonstração, afirma que a distribuição de Dirac e as suas derivadas são completamente caracterizadas pela propriedade de suporte enunciada no exemplo anterior:

Teorema 3.5 (Schwartz): *Se $T \in \mathcal{D}'$ é uma distribuição com suporte $\{0\}$, então existe $r \geq 0$ tal que T é uma combinação linear da função delta de Dirac e de suas derivadas até ordem r :*

$$T = \sum_{|\alpha| \leq r} c_\alpha \delta^{(\alpha)}. \quad (3.105)$$

Distribuições de suporte pontual são os exemplos mais simples dentro de uma classe maior, a das distribuições de suporte compacto, que têm papel importante em equações diferenciais parciais e suas aplicações⁸ e para as quais daremos uma interpretação alternativa na Seção 3.8.

3.6 Diferenciação

A definição de derivadas parciais para distribuições é baseada na seguinte

Proposição 3.19 *Para todo aberto Ω de \mathbb{R}^n , o operador linear de diferenciação parcial $\partial_\alpha : \mathcal{D}(\Omega) \rightarrow \mathcal{D}(\Omega)$ é contínuo.*

DEMONSTRAÇÃO: Para todo subconjunto compacto K de Ω , a restrição do operador linear $\partial_\alpha : \mathcal{D}(\Omega) \rightarrow \mathcal{D}(\Omega)$ ao subespaço $\mathcal{D}(K)$ é um operador linear $\partial_\alpha : \mathcal{D}(K) \rightarrow \mathcal{D}(K)$ que é contínuo, pois para $r = 0, 1, 2, \dots$, temos

$$\|\partial_\alpha \varphi\|_{r,K} \leq \|\varphi\|_{r+|\alpha|,K} \quad \text{para } \varphi \in \mathcal{D}(K),$$

o que prova a afirmação, segundo a Proposição 3.12. \square

Definição 3.21 *Para todo aberto Ω de \mathbb{R}^n e qualquer distribuição T sobre Ω , define-se sua derivada parcial de ordem α como sendo a distribuição $\partial_\alpha T$ sobre Ω dada pela fórmula*

$$\langle \partial_\alpha T, \varphi \rangle = (-1)^{|\alpha|} \langle T, \partial_\alpha \varphi \rangle \quad \text{para } \varphi \in \mathcal{D}(\Omega). \quad (3.106)$$

⁸Por exemplo, a fonte de uma equação diferencial parcial tal como a equação (1.44), quando representa uma componente de uma distribuição (no sentido da Física) de cargas e correntes localizadas em uma região finita do espaço, como acontece normalmente em problemas de eletromagnetismo, é uma distribuição (no sentido da Matemática) de suporte compacto.

Observa-se que $\partial_\alpha T$ é realmente uma distribuição sobre Ω , i.e., um funcional linear contínuo sobre $\mathcal{D}(\Omega)$, pois (a menos de um sinal) é a composição do funcional linear contínuo T sobre $\mathcal{D}(\Omega)$ com o operador linear contínuo da proposição anterior.

Afim de motivar esta definição da derivada para distribuições (inclusive o sinal), mostramos que para distribuições regulares, ela coincide com a noção clássica de derivada para funções. Mais exatamente, temos

Proposição 3.20 *Seja Ω um aberto de \mathbb{R}^n e seja f uma função de classe C^r sobre Ω . Então para todo multi-índice $\alpha \in \mathbb{N}^n$ com $|\alpha| \leq r$, vale*

$$\partial_\alpha T_f = T_{\partial_\alpha f}. \quad (3.107)$$

DEMONSTRAÇÃO: Usando indução sobre o grau $|\alpha|$ de α , podemos reduzir a demonstração da afirmação geral à demonstração do caso onde $|\alpha| = 1$, no qual o operador ∂_α se reduz a uma simples derivada parcial ∂_j e podemos utilizar integração por partes na coordenada x_j : para qualquer função teste $\varphi \in \mathcal{D}(\Omega) \subset \mathcal{D}$, temos

$$\begin{aligned} \langle \partial_j T_f, \varphi \rangle &= -\langle T_f, \partial_j \varphi \rangle = -\int d^n x f(x) \partial_j \varphi(x) \\ &= -\int dx_1 \dots dx_n f(x_1, \dots, x_n) \frac{\partial \varphi}{\partial x_j}(x_1, \dots, x_n) \\ &= -\int dx_1 \dots dx_{j-1} dx_{j+1} \dots dx_n \\ &\quad \times \left(\int dx_j \frac{\partial}{\partial x_j} (f\varphi)(x_1, \dots, x_j, \dots, x_n) \right) \\ &\quad + \int dx_1 \dots dx_n \frac{\partial f}{\partial x_j}(x_1, \dots, x_n) \varphi(x_1, \dots, x_n) \\ &= \langle T_{\partial_j f}, \varphi \rangle, \end{aligned}$$

pois φ tendo suporte compacto, o termo de bordo

$$\begin{aligned} &\int dx_j \frac{\partial}{\partial x_j} (f\varphi)(x_1, \dots, x_j, \dots, x_n) \\ &= \lim_{a \rightarrow \infty} \left((f\varphi)(x_1, \dots, x_{j-1}, a, x_{j+1}, \dots, x_n) \right. \\ &\quad \left. - (f\varphi)(x_1, \dots, x_{j-1}, -a, x_{j+1}, \dots, x_n) \right) \end{aligned}$$

se anula. □

Destacamos uma propriedade de derivadas parciais que vale para funções comuns (mais exatamente, funções de classe C^2) e que continua valendo para distribuições: é o lema de Schwartz sobre comutatividade das derivadas parciais, que constitui a justifica para podermos utilizar a notação dos multi-índices para derivadas parciais de ordem superior.

Proposição 3.21 (*Lema de Schwartz*): Para todo aberto Ω de \mathbb{R}^n e qualquer distribuição T sobre Ω , vale

$$\frac{\partial}{\partial x_i} \frac{\partial}{\partial x_j} T = \frac{\partial}{\partial x_j} \frac{\partial}{\partial x_i} T. \quad (3.108)$$

DEMONSTRAÇÃO: Para $T \in \mathcal{D}'(\Omega)$ e $\varphi \in \mathcal{D}(\Omega)$, vale

$$\begin{aligned} \left\langle \frac{\partial}{\partial x_i} \frac{\partial}{\partial x_j} T, \varphi \right\rangle &= - \left\langle \frac{\partial T}{\partial x_j}, \frac{\partial \varphi}{\partial x_i} \right\rangle = \left\langle T, \frac{\partial}{\partial x_j} \frac{\partial}{\partial x_i} \varphi \right\rangle \\ &= \left\langle T, \frac{\partial}{\partial x_i} \frac{\partial}{\partial x_j} \varphi \right\rangle = - \left\langle \frac{\partial T}{\partial x_i}, \frac{\partial \varphi}{\partial x_j} \right\rangle = \left\langle \frac{\partial}{\partial x_j} \frac{\partial}{\partial x_i} T, \varphi \right\rangle. \end{aligned}$$

□

Passamos a apresentar alguns exemplos. Primeiro, comparando as equações (3.87), (3.88) e (3.106), percebe-se que a distribuição $\delta^{(\alpha)}$ introduzida anteriormente realmente é a derivada parcial de ordem α de δ :

$$\delta^{(\alpha)} = \partial_\alpha \delta. \quad (3.109)$$

Exercício 3.4 Prove que, no sentido de distribuições sobre \mathbb{R} ,

$$\frac{d}{dx} \theta(x) = \delta(x). \quad (3.110)$$

$$\frac{d}{dx} |x| = 2\theta(x) - 1. \quad (3.111)$$

Com essas ferramentas à nossa disposição, já podemos iniciar um estudo, no âmbito da teoria de distribuições, de equações diferenciais parciais (lineares) a coeficientes constantes: são equações da forma

$$P(\partial)S = T, \quad (3.112)$$

onde S e T devem ser distribuições ($S, T \in \mathcal{D}'$) e

$$P(\partial) = \sum_{|\alpha| \leq r} a_\alpha \partial_\alpha. \quad (3.113)$$

Em particular, podemos dar uma definição rigorosa da noção de função de Green⁹ ou solução fundamental:

Definição 3.22 Uma **função de Green** ou **solução fundamental** de uma equação diferencial parcial (linear) a coeficientes constantes, definida por um operador diferencial parcial $P(\partial)$, é uma distribuição $G \in \mathcal{D}'$ tal que

$$P(\partial)G = \delta. \quad (3.114)$$

⁹ Na Física, o termo “função de Green” também é utilizado para caracterizar certas soluções da correspondente equação homogênea $P(\partial)G = 0$.

Por enquanto, esta definição é válida para operadores a coeficientes constantes, mas ela se estende facilmente a operadores cujos coeficientes são funções lisas, usando a definição do produto de uma função lisa com uma distribuição a ser dada na próxima seção. No entanto, o seguinte teorema fundamental, que mencionamos sem demonstração, é válido apenas para operadores a coeficientes constantes:

Teorema 3.6 (Ehrenpreis-Malgrange): *Qualquer operador diferencial parcial a coeficientes constantes admite uma função de Green, ou solução fundamental.*

Naturalmente, esta função de Green ou solução fundamental não é única, pois pode ser modificada pela adição de qualquer solução da equação homogênea. Portanto, costuma-se impor condições adicionais (condições de fronteira e, no caso de equações de evolução, condições iniciais) para determinar uma função de Green ou solução fundamental específica.

Neste livro, determinaremos funções de Green para alguns dos operadores diferenciais parciais mais importantes que aparecem na Física, ilustrando ao mesmo tempo diversas técnicas utilizadas para a solução de equações diferenciais parciais. Na presente seção, apresentamos uma abordagem direta ao problema de determinar uma função de Green para o operador de Laplace em \mathbb{R}^n .

No caso $n = 1$, a solução deste problema é simples, para não dizer trivial, pois combinando as equações (3.110) e (3.111), concluímos imediatamente que

$$G(x) = \frac{1}{2} |x|. \quad (3.115)$$

Esta função de Green é unicamente determinada pelas condições adicionais de que $G(-x) = G(x)$ e $G(0) = 0$.

No caso $n > 1$, a ideia básica consiste em utilizar a invariância do operador de Laplace e da função δ de Dirac sob o grupo $O(n)$ de rotações em \mathbb{R}^n (veja a equação (3.164) abaixo) para justificar a procura por uma função de Green G que seja apenas uma “função” da variável radial r , o que reduz o problema à solução de uma equação diferencial ordinária. Naturalmente, esta é apenas a parte clássica da construção de uma função de Green, pois é restrita à região fora da origem. Contudo, o procedimento deve fornecer uma solução que desenvolve uma singularidade na origem: caso contrário, obteríamos uma solução da equação homogênea em \mathbb{R}^n , ao invés de uma solução fundamental. A construção completa procede em três passos:

O primeiro passo consiste em resolver a equação

$$\Delta G = \delta \quad (3.116)$$

apenas no aberto $\mathbb{R}^n \setminus \{0\}$, onde a função delta de Dirac se anula; portanto, tentamos resolver a equação

$$\Delta G = 0 \quad \text{em } \mathbb{R}^n \setminus \{0\}$$

sob a hipótese de que G seja uma função comum (pelo menos de classe C^2) da variável radial r . Este é um problema puramente clássico que já foi resolvido na

Seção 2.1.2: sabemos que a solução geral é

$$G(x) = \begin{cases} -\frac{c}{(n-2)r^{n-2}} + c' & (n > 2) \\ c \ln r + c' & (n = 2) \end{cases}$$

onde c e c' são constantes de integração. Obviamente, para a solução da equação (3.116), mesmo em \mathbb{R}^n , o valor da constante aditiva c' é irrelevante e portanto podemos supor sem perda de generalidade que G é um múltiplo do potencial de Coulomb G_0 introduzido na Seção 2.1.2: Temos

$$G = k G_0, \quad (3.117)$$

com

$$k = \int_{S_a^{n-1}} d\sigma(x) \cdot \nabla G(x). \quad (3.118)$$

No segundo passo, observa-se que, conforme o critério da Proposição 3.14, a função G assim definida, que é contínua exceto na origem, é localmente integrável, pois como já foi demonstrado na Seção 2.1.2, a integral de $|G|$ sobre uma bola B_a^n de raio a em torno da origem vale

$$\int_{B_a^n} d^n x |G(x)| = k g_n(a) < \infty,$$

com $g_n(a)$ dado pela equação (2.15), e portanto permanece limitada quando $a \rightarrow 0$. Assim, G define uma distribuição, que também denotaremos por G , e a distribuição ΔG está bem definida. Ademais, as mesmas estimativas mostram que para qualquer função contínua f sobre \mathbb{R}^n de suporte compacto, temos

$$\int d^n x G(x) f(x) = \lim_{\epsilon \rightarrow 0} \int_{|x| \geq \epsilon} d^n x G(x) f(x),$$

pois para $\epsilon > 0$ finito, a diferença pode ser estimada por

$$\begin{aligned} & \left| \int d^n x G(x) f(x) - \int_{|x| \geq \epsilon} d^n x G(x) f(x) \right| \\ & \leq \int_{\bar{B}_\epsilon} d^n x |G(x)| |f(x)| \\ & \leq \sup_{x \in \text{supp } f} |f(x)| \cdot \text{vol}(S^{n-1}) k g_n(\epsilon), \end{aligned}$$

mostrando que ela se anula no limite $\epsilon \rightarrow 0$.

O terceiro passo consiste em calcular, para toda função teste $\varphi \in \mathcal{D}$ com $\text{supp } \varphi \subset B_R(0)$, a expressão $\langle \Delta G, \varphi \rangle$, usando a segunda identidade de Green (B.45):

$$\begin{aligned} \langle \Delta G, \varphi \rangle &= \langle G, \Delta \varphi \rangle = \int d^n x G(x) (\Delta \varphi)(x) \\ &= \lim_{\epsilon \rightarrow 0} \int_{|x| \geq \epsilon} d^n x G(x) (\Delta \varphi)(x) \\ &= \lim_{\epsilon \rightarrow 0} \int_{\epsilon \leq |x| \leq R} d^n x \left(G(x) (\Delta \varphi)(x) - \varphi(x) (\Delta G)(x) \right) \\ &= \int_{|x|=R} d\sigma(x) \cdot \left(G(x) (\nabla \varphi)(x) - (\nabla G)(x) \varphi(x) \right) \\ &\quad - \lim_{\epsilon \rightarrow 0} \int_{|x|=\epsilon} d\sigma(x) \cdot \left(G(x) (\nabla \varphi)(x) - (\nabla G)(x) \varphi(x) \right). \end{aligned}$$

A primeira integral se anula pois φ e todas as suas derivadas são zero quando $|x| \geq R$, enquanto que para $\epsilon > 0$ finito, a primeira contribuição à segunda integral pode ser estimada por

$$\begin{aligned} \left| \int_{|x|=\epsilon} d\sigma(x) \cdot G(x) (\nabla \varphi)(x) \right| \\ \leq \text{vol}(S^{n-1}) \epsilon^{n-1} |G(\epsilon)| \cdot \sup_{x \in \text{supp } \varphi} |\nabla \varphi(x)|, \end{aligned}$$

mostrando ela se anula no limite $\epsilon \rightarrow 0$. Finalmente, a segunda contribuição à segunda integral é dividida em duas partes:

$$\begin{aligned} \langle \Delta G, \varphi \rangle &= \lim_{\epsilon \rightarrow 0} \left\{ \int_{|x|=\epsilon} d\sigma(x) \cdot (\nabla G)(x) (\varphi(x) - \varphi(0)) \right. \\ &\quad \left. + \int_{|x|=\epsilon} d\sigma(x) \cdot (\nabla G)(x) \varphi(0) \right\}. \end{aligned}$$

Utilizando a equação (3.118), concluímos que para $\epsilon > 0$ finito, a primeira contribuição pode ser estimada por

$$\left| \int_{|x|=\epsilon} d\sigma(x) \cdot (\nabla G)(x) (\varphi(x) - \varphi(0)) \right| \leq k \sup_{|x|=\epsilon} |\varphi(x) - \varphi(0)|.$$

mostrando que ela se anula no limite $\epsilon \rightarrow 0$, devido à continuidade de φ na origem, enquanto que a segunda contribuição dá um resultado finito:

$$\langle \Delta G, \varphi \rangle = k \varphi(0).$$

Assim, provamos o seguinte

Teorema 3.7 A função de Green para o operador de Laplace em \mathbb{R}^n invariante sob o grupo de rotações $O(n)$ coincide com o potencial de Coulomb G_0 introduzido no Capítulo 2, dado por

$$G_0(x) = \begin{cases} -\frac{1}{(n-2)\text{vol}(S^{n-1})} \frac{1}{r^{n-2}} & (n > 2) \\ \frac{1}{2\pi} \ln \frac{r}{r_0} & (n = 2) , \\ \frac{1}{2} r & (n = 1) \end{cases} \quad (3.119)$$

onde $r = |x| = (\sum_{i=1}^n x_i^2)^{1/2}$ e $\text{vol}(S^{n-1}) = 2\pi^{n/2}/\Gamma(n/2)$.

3.7 Multiplicação

A definição da multiplicação entre distribuições e funções lisas é baseada na primeira parte da seguinte

Proposição 3.22 Para todo aberto Ω de \mathbb{R}^n , o operador bilinear de multiplicação

$$\begin{aligned} \mathcal{E}(\Omega) \times \mathcal{D}(\Omega) &\longrightarrow \mathcal{D}(\Omega) \\ (f, \varphi) &\longmapsto f\varphi \end{aligned} \quad (3.120)$$

é separadamente contínuo, o que significa que

- Para toda função $f \in \mathcal{E}(\Omega)$, a aplicação linear

$$\begin{aligned} \mathcal{D}(\Omega) &\longrightarrow \mathcal{D}(\Omega) \\ \varphi &\longmapsto f\varphi \end{aligned} \quad (3.121)$$

é contínua.

- Para toda função $\varphi \in \mathcal{D}(\Omega)$, a aplicação linear

$$\begin{aligned} \mathcal{E}(\Omega) &\longrightarrow \mathcal{D}(\Omega) \\ f &\longmapsto f\varphi \end{aligned} \quad (3.122)$$

é contínua.

DEMONSTRAÇÃO: As relações (B.9) e (3.51) com $X = K$ mostram que dada uma função $f \in \mathcal{E}(\Omega)$ e qualquer subconjunto compacto K de Ω , a restrição do operador linear (3.121) de multiplicação por f ao subespaço $\mathcal{D}(K)$ é um operador linear contínuo

$$\begin{aligned} \mathcal{D}(K) &\longrightarrow \mathcal{D}(K) \\ \varphi &\longmapsto f\varphi \end{aligned} ,$$

e, de forma semelhante, que dada uma função $\varphi \in \mathcal{D}(\Omega)$ com suporte $K = \text{supp } \varphi \subset \Omega$, o operador linear (3.122) de multiplicação por φ define um operador linear contínuo

$$\begin{aligned} \mathcal{E}(\Omega) &\longrightarrow \mathcal{D}(K) \\ f &\longmapsto f\varphi \end{aligned} .$$

□

Definição 3.23 Para todo aberto Ω de \mathbb{R}^n e para qualquer função lisa f sobre Ω e qualquer distribuição T sobre Ω , define-se o seu produto como sendo a distribuição fT sobre Ω dada pela fórmula

$$\langle fT, \varphi \rangle = \langle T, f\varphi \rangle \quad \text{para } \varphi \in \mathcal{D}(\Omega) . \quad (3.123)$$

Observa-se que fT é realmente uma distribuição sobre Ω , i.e., um funcional linear contínuo sobre $\mathcal{D}(\Omega)$, pois é a composição do funcional linear contínuo T sobre $\mathcal{D}(\Omega)$ com a aplicação linear contínua (3.121).

Afim de motivar esta definição da multiplicação entre distribuições e funções lisas, mostramos que para distribuições regulares, ela coincide com a noção clássica de multiplicação entre funções. Mais exatamente, temos

Proposição 3.23 Seja Ω um aberto de \mathbb{R}^n e sejam f uma função lisa sobre Ω e g uma função localmente integrável sobre Ω . Então vale

$$fT_g = T_{fg} . \quad (3.124)$$

DEMONSTRAÇÃO: Para qualquer função teste $\varphi \in \mathcal{D}(\Omega)$, temos

$$\begin{aligned} \langle T_{fg}, \varphi \rangle &= \int d^n x (fg)(x) \varphi(x) \\ &= \int d^n x g(x) (f\varphi)(x) = \langle T_g, f\varphi \rangle = \langle fT_g, \varphi \rangle . \end{aligned}$$

□

Continua valendo a regra do produto:

Proposição 3.24 Para todo aberto Ω de \mathbb{R}^n , as derivadas parciais do produto $fT \in \mathcal{D}'(\Omega)$ entre uma função lisa $f \in \mathcal{E}(\Omega)$ sobre Ω e uma distribuição $T \in \mathcal{D}'(\Omega)$ sobre Ω satisfazem a **regra do produto**

$$\partial_\alpha(fT) = \sum_{\beta \leq \alpha} \binom{\alpha}{\beta} \partial_\beta f \partial_{\alpha-\beta} T . \quad (3.125)$$

DEMONSTRAÇÃO: Usando indução sobre o grau $|\alpha|$ de α , podemos reduzir a demonstração da afirmação geral à demonstração do caso onde $|\alpha| = 1$, no qual o operador ∂_α se reduz a uma simples derivada parcial ∂_j , exatamente como no caso clássico do produto de duas funções lisas. Assim, basta notar que para qualquer função teste $\varphi \in \mathcal{D}(\Omega)$, temos

$$\begin{aligned} \langle \partial_j(fT), \varphi \rangle &= -\langle fT, \partial_j\varphi \rangle = -\langle T, f\partial_j\varphi \rangle \\ &= \langle T, \partial_j f\varphi - \partial_j(f\varphi) \rangle = \langle \partial_j f T, \varphi \rangle + \langle \partial_j T, f\varphi \rangle \\ &= \langle \partial_j f T + f\partial_j T, \varphi \rangle. \end{aligned}$$

□

Exemplo 3.14 Para qualquer função lisa $f \in \mathcal{E}$ sobre \mathbb{R} , vale $f\delta = f(0)\delta$, ou seja,

$$f(x)\delta(x) = f(0)\delta(x). \quad (3.126)$$

De forma semelhante, temos $f\delta' = f(0)\delta' - f'(0)\delta$, ou seja,

$$f(x)\delta'(x) = f(0)\delta'(x) - f'(0)\delta(x), \quad (3.127)$$

e mais geralmente

$$f(x)\delta^{(k)}(x) = \sum_{l=0}^k (-1)^l \binom{k}{l} f^{(l)}(0)\delta^{(k-l)}(x). \quad (3.128)$$

Também temos

$$x \text{ v.p. } \frac{1}{x} = 1, \quad (3.129)$$

e mais geralmente

$$x^k \text{ p.f. } \frac{1}{x^k} = 1. \quad (3.130)$$

Exercício 3.5 Prove estas fórmulas.

Exercício 3.6 Mostre que a função de Green

(i) do operador $\frac{d}{dx} - \lambda$ é $\theta(x) \exp(\lambda x)$.

(ii) do operador $\frac{d^2}{dx^2} + k^2$ é $\frac{1}{k} \theta(x) \sin(kx)$.

Um dos maiores problemas, para não dizer o maior problema, da teoria de distribuições é a impossibilidade de definir um produto geral entre duas distribuições arbitrárias que satisfaça propriedades elementares de um produto, tais como associatividade ou comutatividade. De fato, se houvesse a possibilidade de definir um produto geral entre duas distribuições, teríamos, por exemplo,

$$\delta(x) \left(x \text{ v.p. } \frac{1}{x} \right) = \delta(x), \quad (3.131)$$

enquanto que

$$\left(\delta(x) x \right) \text{ v.p. } \frac{1}{x} = 0. \quad (3.132)$$

De forma geral, para poder definir um produto de duas distribuições S e T que seja novamente uma distribuição ST , precisamos ter um certo equilíbrio entre as singularidades de S e as singularidades de T : quanto mais singular um dos fatores, mais regular deve ser o outro. Um exemplo típico, que estabelece simetria entre os dois fatores, é obtido pela afirmação de que o produto de duas distribuições S e T é bem definido quando seus suportes singulares são disjuntos:

$$S, T \in \mathcal{D}'(\Omega) \quad , \quad \text{sing supp } S \cap \text{sing supp } T = \emptyset \quad \implies \quad ST \in \mathcal{D}'(\Omega) .$$

Existe uma extensão da noção de suporte singular chamada de “conjunto de frente de ondas” que, em certas circunstâncias, permite multiplicar distribuições mesmo quando seus suportes singulares se interceptam, e que abordaremos na Seção 3.15.

3.8 Algumas classes especiais de distribuições

Nesta seção, introduzimos algumas classes de distribuições especiais, entre elas as distribuições de suporte compacto, as distribuições temperadas e as distribuições integráveis. Cada uma dessas classes constitui um subespaço do espaço \mathcal{D}' (ou $\mathcal{D}'(\Omega)$) que pode ser identificado com o dual de um outro “espaço de funções teste”, que é maior do que \mathcal{D} (ou $\mathcal{D}(\Omega)$).

O exemplo mais simples é obtido considerando, para qualquer aberto Ω de \mathbb{R}^n , o espaço $\mathcal{E}(\Omega)$ das funções lisas sobre Ω e observando que a inclusão $\mathcal{D}(\Omega) \subset \mathcal{E}(\Omega)$ é contínua, ou seja, a topologia de $\mathcal{D}(\Omega)$ é mais fina do que a topologia induzida pela topologia de $\mathcal{E}(\Omega)$. (Isto segue da Proposição 3.11, tendo em vista que para todo subconjunto compacto K de Ω , o sistema de seminormas definindo a topologia de $\mathcal{D}(K)$ é exatamente o mesmo que o sistema de seminormas definindo a topologia induzida em $\mathcal{D}(K)$ pela topologia de $\mathcal{E}(\Omega)$, o que se verifica comparando as equações (3.71) e (3.75).) Além disso, vale a seguinte

Proposição 3.25 *Para todo aberto Ω de \mathbb{R}^n , $\mathcal{D}(\Omega)$ é denso em $\mathcal{E}(\Omega)$.*

DEMONSTRAÇÃO: Considere qualquer uma das seminormas $\|\cdot\|_{r,K}$ (definidas na equação (3.71)) que geram a topologia de $\mathcal{E}(\Omega)$, onde $r \in \mathbb{N}$ e K é um subconjunto compacto de Ω , e note que escolhendo

uma função teste $\chi \in \mathcal{D}(\Omega)$ tal que $0 \leq \chi \leq 1$ e $\chi = 1$ sobre uma vizinhança de K , conforme o Teorema 3.4, obtém-se que, para toda função $\varphi \in \mathcal{E}(\Omega)$, vale $\chi\varphi \in \mathcal{D}(\Omega)$ e $\|\varphi - \chi\varphi\|_{r,K} = 0$. \square

Assim, passando aos espaços duais, segue que a inclusão contínua $\mathcal{D}(\Omega) \subset \mathcal{E}(\Omega)$ define uma aplicação linear de $\mathcal{E}'(\Omega)$ para $\mathcal{D}'(\Omega)$ que associa a cada funcional linear contínuo sobre o espaço $\mathcal{E}(\Omega)$ sua restrição ao subespaço $\mathcal{D}(\Omega)$, e a Proposição 3.25 é crucial na medida em que garante que essa aplicação linear é injetora e portanto pode ser vista como proporcionando uma inclusão $\mathcal{E}'(\Omega) \subset \mathcal{D}'(\Omega)$: deste ponto de vista, uma distribuição $T \in \mathcal{D}'(\Omega)$ pertence ao subespaço $\mathcal{E}'(\Omega)$ se e somente se, como funcional linear sobre o espaço de funções teste $\mathcal{D}(\Omega)$, ela é contínua não apenas na topologia padrão de $\mathcal{D}(\Omega)$ mas também na topologia mais grossa induzida pela topologia padrão de $\mathcal{E}(\Omega)$, e se isso for o caso, ela possui uma única extensão ao espaço de funções teste maior $\mathcal{E}(\Omega)$.

Para lidar com outros exemplos, introduzimos então a seguinte terminologia, estendendo e complementando a Definição 3.16:

Definição 3.24 *Para todo aberto Ω de \mathbb{R}^n , um **espaço de distribuições sobre Ω** é o espaço dual $\mathcal{F}'(\Omega)$ de um **espaço de funções teste sobre Ω** , que é um espaço localmente convexo completo $\mathcal{F}(\Omega)$ contido no espaço $\mathcal{E}(\Omega)$ e contendo o espaço $\mathcal{D}(\Omega)$ tal que as inclusões*

$$\mathcal{D}(\Omega) \subset \mathcal{F}(\Omega) \subset \mathcal{E}(\Omega) \quad (3.133)$$

são contínuas e tal que $\mathcal{D}(\Omega)$ é denso em $\mathcal{F}(\Omega)$ (sendo que devido à Proposição 3.25, $\mathcal{F}(\Omega)$ é automaticamente denso em $\mathcal{E}(\Omega)$), de modo que a restrição de funcionais lineares a subespaços proporciona inclusões

$$\mathcal{E}'(\Omega) \subset \mathcal{F}'(\Omega) \subset \mathcal{D}'(\Omega) \quad (3.134)$$

Novamente, a continuidade das inclusões na equação (3.133) significa que a topologia padrão de $\mathcal{D}(\Omega)$ é mais fina do que a topologia induzida pela topologia de $\mathcal{F}(\Omega)$ e que a topologia de $\mathcal{F}(\Omega)$ é mais fina do que a topologia induzida pela topologia padrão de $\mathcal{E}(\Omega)$; portanto, por restrição, todo funcional linear contínuo sobre $\mathcal{F}(\Omega)$ define um funcional linear contínuo sobre $\mathcal{D}(\Omega)$ e todo funcional linear contínuo sobre $\mathcal{E}(\Omega)$ define um funcional linear contínuo sobre $\mathcal{F}(\Omega)$, o que proporciona as inclusões na equação (3.134), devido às condições de densidade.

Quanto à propriedade de completude de $\mathcal{F}(\Omega)$ que exigimos nesta definição (e que alguns autores omitem), notamos o seguinte:

- Este requerimento é consistente com a terminologia já adotada anteriormente, segundo a qual os próprios espaços $\mathcal{D}(\Omega)$ e $\mathcal{E}(\Omega)$ devem ser espaços de funções teste sobre Ω e seus respectivos duais $\mathcal{D}'(\Omega)$ e $\mathcal{E}'(\Omega)$ devem ser espaços de distribuições sobre Ω , pois $\mathcal{D}(\Omega)$ e $\mathcal{E}(\Omega)$ realmente são completos. De fato,

a completude de $\mathcal{E}(\Omega)$ na sua topologia padrão (a saber, a topologia da convergência uniforme de todas as derivadas sobre subconjuntos compactos de Ω) já foi notada na Observação 3.2, enquanto que a completude de $\mathcal{D}(\Omega)$ na sua topologia padrão (a saber, a topologia do “limite indutivo” da família dos subespaços $\mathcal{D}(K)$, onde K percorre os subconjuntos compactos de Ω) segue diretamente da completude de cada um destes subespaços, mas apresentar os detalhes deste argumento foge do escopo do presente texto.

- Este requerimento pode sempre ser imposto sem perda de generalidade, uma vez que se não fosse satisfeito, podemos substituir $\mathcal{F}(\Omega)$ por seu completamento $\bar{\mathcal{F}}(\Omega)$, preservando todas as condições estipuladas na Definição 3.24 (continuidade das inclusões e densidade).

Supondo então que o espaço $\mathcal{F}(\Omega)$ é completo em sua própria topologia localmente convexa, é natural considerar $\mathcal{F}(\Omega)$ como o completamento de $\mathcal{D}(\Omega)$ nesta topologia, que pode ser definida por uma família própria de seminormas sobre $\mathcal{D}(\Omega)$. A seguir, consideraremos vários exemplos onde tal família própria de seminormas é definido por um espaço $\mathcal{W}(\Omega)$ de funções peso que também deve satisfazer

$$\mathcal{D}(\Omega) \subset \mathcal{W}(\Omega) \subset \mathcal{E}(\Omega) \quad (3.135)$$

da seguinte forma: escolhendo uma família qualquer $\chi = (\chi_\alpha)_{\alpha \in \mathbb{N}^n}$ de funções peso $\chi_\alpha \in \mathcal{W}(\Omega)$ a valores reais e tais que $\chi_\alpha \geq 0$, defina

$$s_{r,\chi}(\varphi) = \max_{|\alpha| \leq r} \sup_{x \in \Omega} |(\chi_\alpha \partial_\alpha \varphi)(x)| \quad \text{para } \varphi \in \mathcal{D}(\Omega), \quad (3.136)$$

onde $\|\cdot\|_{r,\Omega}$ denota a norma do supremo das derivadas até ordem r sobre Ω já introduzida anteriormente; veja as equações (3.43), (3.46) e (3.47). Obviamente, a topologia localmente convexa $\mathcal{T}_{\mathcal{W}}$ gerada em $\mathcal{D}(\Omega)$ pela família de seminormas $s_{r,\chi}$ ($r = 0, 1, 2, \dots$, $\chi \in \mathcal{W}(\Omega)$, $\chi \geq 0$) depende da escolha do espaço $\mathcal{W}(\Omega)$, e podemos comparar as topologias geradas por duas escolhas diferentes: se $\mathcal{W}_1(\Omega) \subset \mathcal{W}_2(\Omega)$, então $\mathcal{T}_{\mathcal{W}_1}$ é mais grossa do que $\mathcal{T}_{\mathcal{W}_2}$. Também notamos que para qualquer escolha do espaço $\mathcal{W}(\Omega)$ e qualquer compacto $K \subset \Omega$, a topologia $\mathcal{T}_{\mathcal{W}}$ induz a topologia padrão sobre o subespaço $\mathcal{D}(K)$ de $\mathcal{D}(\Omega)$,

A título de exemplo, consideramos os dois exemplos extremos; outros serão investigados logo adiante:

- $\mathcal{W}(\Omega) = \mathcal{D}(\Omega)$: essa é a topologia mais grossa obtida por este método, e é fácil ver que ela coincide com a topologia da convergência uniforme de todas as derivadas sobre subconjuntos compactos de Ω , de modo que neste caso, $\mathcal{F}(\Omega)$ é maximal: $\mathcal{F}(\Omega) = \mathcal{E}(\Omega)$.

Além disso, na maioria dos casos de importância na prática, $\mathcal{F}(\Omega)$ não é apenas um espaço vetorial localmente convexo mas, assim como $\mathcal{D}(\Omega)$ e $\mathcal{E}(\Omega)$, é uma $*$ -álgebra comutativa localmente convexa, conforme já discutido na Observação 3.3. Nestas circunstâncias, a questão de provar densidade de $\mathcal{D}(\Omega)$ em $\mathcal{F}(\Omega)$ costuma

ser abordada usando uma sequência crescente $(K_k)_{k \in \mathbb{N}}$ de subconjuntos compactos K_k de Ω que preenchem Ω , i.e., que satisfazem

$$K_1 \subset \dots \subset K_k \subset K_{k+1} \subset \dots \subset \Omega \quad \text{e} \quad \bigcup_{k \in \mathbb{N}} K_k = \Omega,$$

de modo que, na verdade, cada um destes compactos é contido no interior do compacto seguinte e que todo subconjunto compacto de Ω é contido em algum K_k , e uma sequência $(\chi_k)_{k \in \mathbb{N}}$ de funções teste $\chi_k \in \mathcal{D}(K_{k+1})$ tais que $0 \leq \chi_k \leq 1$, $\chi_k = 1$ sobre uma vizinhança de K_k contida no interior de K_{k+1} , conforme o Teorema 3.4; então a questão é se, para toda função $\varphi \in \mathcal{F}(\Omega)$, $\chi_k \varphi \rightarrow \varphi$ quando $k \rightarrow \infty$, na topologia de $\mathcal{F}(\Omega)$. Neste caso, dizemos que a sequência $(\chi_k)_{k \in \mathbb{N}}$ constitui uma *aproximação da unidade* em $\mathcal{F}(\Omega)$, em $\mathcal{F}(\Omega)$, o que

A SER CONTINUADO/COMPLETADO

Sistemas de seminormas gerados por funções peso

Em muitos casos, queremos definir $\mathcal{F}(\Omega)$ através de um sistema de seminormas

3.8.1 Distribuições de suporte compacto

Começamos por apresentar uma construção geral que se mostra útil em diversas circunstâncias, das quais o tratamento de distribuições de suporte compacto é apenas um caso entre vários – se bem que talvez o mais importante.

Proposição 3.26 *Para todo aberto Ω de \mathbb{R}^n e para qualquer função lisa φ sobre Ω e qualquer distribuição T sobre Ω , a expressão $\langle T, \varphi \rangle$ ainda está bem definida desde que a interseção dos suportes de T e de φ seja um compacto. Explicitamente, ela é definida por*

$$\langle T, \varphi \rangle = \langle \chi T, \varphi \rangle = \langle T, \chi \varphi \rangle, \quad (3.137)$$

onde $\chi \in \mathcal{D}(\Omega)$ é qualquer função teste tal que $0 \leq \chi \leq 1$ e $\chi = 1$ sobre uma vizinhança de $\text{supp } T \cap \text{supp } \varphi$, conforme o Teorema 3.4, sendo que, nestas condições, a expressão $\langle T, \chi \varphi \rangle$ não depende da escolha da “função de corte” χ .

DEMONSTRAÇÃO: Para mostrar que a expressão $\langle T, \chi \varphi \rangle$ realmente não depende da escolha de χ , considere duas funções teste χ_1 e χ_2 em $\mathcal{D}(\Omega)$ com as propriedades requeridas ($0 \leq \chi_i \leq 1$ e $\chi_i = 1$ sobre vizinhanças de $\text{supp } T \cap \text{supp } \varphi$). Então a diferença $\chi_1 - \chi_2 \in \mathcal{D}(\Omega)$ se anula sobre uma vizinhança de $\text{supp } T \cap \text{supp } \varphi$ (a interseção das duas anteriores) e portanto T se anula sobre a função teste $(\chi_1 - \chi_2) \varphi \in \mathcal{D}(\Omega)$, pois $\text{supp } T \cap \text{supp } ((\chi_1 - \chi_2) \varphi)$ está contido em

$$\text{supp } T \cap \text{supp } \varphi \cap \text{supp } (\chi_1 - \chi_2) = \emptyset.$$

□

Em geral, o domínio natural da extensão definida pela equação (3.137) não é um espaço de funções muito natural. Para identificá-lo, introduzimos, para todo aberto Ω de \mathbb{R}^n e todo subconjunto fechado F de Ω ,¹⁰ o subespaço $\mathcal{E}_F(\Omega)$ do espaço $\mathcal{E}(\Omega)$ constituído de todas as funções lisas sobre Ω cujo suporte intersecta F em algum compacto, ou seja,

$$\mathcal{E}_F(\Omega) = \{ \varphi \in C^\infty(\Omega) \mid F \cap \text{supp } \varphi \text{ compacto} \}, \quad (3.138)$$

munido da topologia induzida pela topologia de $\mathcal{E}(\Omega)$, ou seja, da convergência uniforme de todas as derivadas sobre subconjuntos compactos de Ω . Vale observar que, com esta topologia, $\mathcal{E}_F(\Omega)$ nem sempre será completo. Também notamos os casos extremos

$$\begin{aligned} \mathcal{E}_F(\Omega) &= \mathcal{E}(\Omega) & \text{se } F \text{ compacto,} \\ \mathcal{E}_F(\Omega) &= \mathcal{D}(\Omega) & \text{se } F = \Omega, \end{aligned} \quad (3.139)$$

sendo que a topologia do espaço $\mathcal{E}_F(\Omega)$ coincide com a do espaço $\mathcal{E}(\Omega)$, no primeiro caso, mas não com a do espaço $\mathcal{D}(\Omega)$, no segundo caso. Com esta notação, a construção dada pela Proposição 3.26 proporciona, para qualquer distribuição T sobre Ω tal que $\text{supp } T \subset F$, uma extensão natural de T , como funcional linear contínuo sobre $\mathcal{D}(\Omega)$, para um funcional linear contínuo sobre $\mathcal{E}_F(\Omega)$ que, por abuso de notação, continuaremos denotando por T . De fato, é óbvio que a prescrição dada pela equação (3.137) define um funcional linear sobre $\mathcal{E}_F(\Omega)$, cuja continuidade segue da estimativa (3.51) (substituindo f por χ , g por φ e X por $\text{supp } \chi$).

¹⁰ Note que aqui F pode ser fechado apenas em Ω mas não em \mathbb{R}^n , sendo apenas a interseção de Ω com algum subconjunto fechado de \mathbb{R}^n , mas se for compacto, então será necessariamente fechado tanto em Ω como em \mathbb{R}^n .

Escolhendo $F = \Omega$, obtemos como corolário a primeira parte da seguinte

Proposição 3.27 *Para todo aberto Ω de \mathbb{R}^n , uma distribuição T sobre Ω pertence ao espaço $\mathcal{E}'(\Omega)$ se e somente se tiver suporte compacto, e neste caso sua extensão a um funcional linear contínuo sobre $\mathcal{E}(\Omega)$ é explicitamente dada pela construção da Proposição 3.26, onde $\chi \in \mathcal{D}(\Omega)$ é qualquer função teste tal que $0 \leq \chi \leq 1$ e $\chi = 1$ sobre uma vizinhança do suporte de T .*

DEMONSTRAÇÃO: Seja $T \in \mathcal{D}'(\Omega)$ uma distribuição sobre Ω . Já foi demonstrado que se T tiver suporte compacto, então T admite a extensão pleiteada. Portanto, suponha agora que T não tenha suporte compacto; então se escolhermos uma sequência crescente qualquer $(K_k)_{k \in \mathbb{N}}$ de subconjuntos compactos K_k de Ω que preencham Ω , i.e., tal que

$$K_1 \subset \dots \subset K_k \subset K_{k+1} \subset \dots \subset \Omega \quad \text{e} \quad \bigcup_{k \in \mathbb{N}} K_k = \Omega,$$

de modo que todo subconjunto compacto de Ω é contido em algum K_k , o suporte de T terá interseção não-trivial com o complemento $\Omega \setminus K_k$ de cada um dos K_k e portanto podemos construir uma sequência $(\varphi_k)_{k \in \mathbb{N}}$ de funções teste $\varphi_k \in \mathcal{D}(\Omega)$ com as seguintes propriedades: (a) vale $\text{supp } \varphi_k \subset \Omega \setminus K_k$, o que implica que a sequência $(\varphi_k)_{k \in \mathbb{N}}$ converge para 0 em $\mathcal{E}(\Omega)$ porque para todo compacto K de Ω , existe $k_0 \in \mathbb{N}$ tal que $K \subset K_{k_0}$ e portanto $\text{supp } \varphi_k \cap K = \emptyset$ assim que $k \geq k_0$, e (b) vale $\langle T, \varphi_k \rangle \neq 0$ e sem perda de generalidade, podemos ainda normalizar cada uma das funções φ_k tal que $\langle T, \varphi_k \rangle = 1$. Então é claro que T não pode ser obtido por restrição de um funcional linear contínuo sobre o espaço $\mathcal{E}(\Omega)$ ao subespaço $\mathcal{D}(\Omega)$. □

Resumindo, concluímos que

Corolário 3.1 *Para todo aberto Ω de \mathbb{R}^n , o espaço $\mathcal{E}(\Omega)$ das funções lisas sobre Ω é um espaço de funções teste sobre Ω , e o seu dual $\mathcal{E}'(\Omega)$ é um espaço de distribuições sobre Ω cujos elementos são exatamente as **distribuições de suporte compacto** sobre Ω .*

3.8.2 Distribuições temperadas

Tomando $\Omega = \mathbb{R}^n$, consideramos o espaço \mathcal{S} introduzido na Definição 3.16 e observamos que a) \mathcal{S} é completo e b) as inclusões $\mathcal{D} \subset \mathcal{S}$ e $\mathcal{S} \subset \mathcal{E}$ são contínuas, ou seja, a topologia de \mathcal{D} é mais fina do que a topologia induzida pela topologia de \mathcal{S} e a topologia de \mathcal{S} é mais fina do que a topologia induzida pela topologia de \mathcal{E} (deixamos os detalhes do argumento como exercício para o leitor). Além disso, vale a seguinte

Proposição 3.28 \mathcal{D} é denso em \mathcal{S} .

DEMONSTRAÇÃO: Considerando qualquer uma das seminormas $\|\cdot\|_{\alpha,\beta}$ que geram a topologia de \mathcal{S} , onde $\alpha, \beta \in \mathbb{N}^n$ (veja a equação (3.73)), note que, para toda função $\varphi \in \mathcal{S}$, a função $x^\beta \partial_\alpha \varphi(x)$ se anula no infinito, i.e., para todo $\rho > 0$ existe um compacto $K_{\alpha,\beta}^\varphi(\rho)$ tal que

$$|x^\beta \partial_\alpha \varphi(x)| < \rho \quad \text{para } x \in \mathbb{R}^n \setminus K_{\alpha,\beta}^\varphi(\rho).$$

Seja então $\varphi \in \mathcal{S}$ e seja χ uma função teste fixa em \mathcal{D} tal que $0 \leq \chi \leq 1$ e, digamos, $\chi(x) = 1$ quando $|x| \leq 1$ e $\chi(x) = 0$ quando $|x| \geq 2$. Então a sequência $(\varphi_k)_{k \in \mathbb{N}}$ de funções teste φ_k em \mathcal{D} definidas por

$$\varphi_k(x) = \chi(x/k) \varphi(x)$$

converge para φ , na topologia do espaço \mathcal{S} . De fato, suponha que

$$C = \sum_{0 < \gamma \leq \alpha} \binom{\alpha}{\gamma} \sup_{x \in \mathbb{R}^n} |\partial_\gamma \chi(x)|$$

e, dado $\epsilon > 0$, que

$$K = K_{\alpha,\beta}^\varphi(\epsilon/2) \cup \bigcup_{0 < \gamma \leq \alpha} K_{\alpha-\gamma,\beta}^\varphi(\epsilon/2C)$$

e seja $k_0 \in \mathbb{N}$, $k_0 > 0$ tal que este compacto K é contido na bola aberta de raio k_0 em torno da origem e portanto φ_k coincide com φ em uma vizinhança aberta de K , para todo $k \in \mathbb{N}$ tal que $k \geq k_0$. Então usando a regra do produto (1.57) e a regra da cadeia para calcular $\partial_\alpha \varphi_k$, concluímos que para $k \geq k_0$,

$$\begin{aligned} \|\varphi - \varphi_k\|_{\alpha,\beta} &= \sup_{x \in \mathbb{R}^n \setminus K} |x^\beta \partial_\alpha (\varphi - \varphi_k)(x)| \\ &= \sup_{x \in \mathbb{R}^n \setminus K} \left| \sum_{\gamma \leq \alpha} \binom{\alpha}{\gamma} x^\beta \partial_\gamma (1 - \chi(x/k)) \partial_{\alpha-\gamma} \varphi(x) \right| \\ &\leq \sup_{x \in \mathbb{R}^n \setminus K} |(1 - \chi(x/k)) x^\beta \partial_\alpha \varphi(x)| \\ &\quad + \sum_{0 < \gamma \leq \alpha} \binom{\alpha}{\gamma} \frac{1}{k^{|\gamma|}} \sup_{x \in \mathbb{R}^n \setminus K} |(\partial_\gamma \chi)(x/k) x^\beta \partial_{\alpha-\gamma} \varphi(x)| \\ &< \epsilon \end{aligned}$$

(onde ainda usamos que $1/k^{|\gamma|} \leq 1$). □

Sendo assim, concluímos que

Definição 3.25 O espaço \mathcal{S} das funções lisas de decaimento rápido sobre \mathbb{R}^n é um espaço de funções teste sobre \mathbb{R}^n , e o seu dual \mathcal{S}' é um espaço de distribuições sobre \mathbb{R}^n cujos elementos são chamados de **distribuições temperadas** sobre \mathbb{R}^n , e temos as inclusões naturais

$$\mathcal{D} \subset \mathcal{S} \subset \mathcal{E}, \quad \mathcal{E}' \subset \mathcal{S}' \subset \mathcal{D}'. \quad (3.140)$$

3.8.3 Distribuições integráveis

Começamos por introduzir alguns espaços de funções adicionais, que na verdade já foram contemplados no Exemplo 3.5:

Definição 3.26 Além dos espaços de funções teste já amplamente utilizados, e apresentados na Definição 3.16, introduzimos

- o espaço \mathcal{B} das funções lisas que, assim como todas as suas derivadas, são limitadas,¹¹

$$\mathcal{B} = C_b^\infty(\mathbb{R}^n) = \left\{ \varphi \in C^\infty(\mathbb{R}^n) \mid \begin{array}{l} \partial_\alpha \varphi \text{ é limitada} \\ \text{para } \alpha \in \mathbb{N}^n \end{array} \right\}, \quad (3.141)$$

- o espaço \mathcal{B}_0 das funções lisas que, assim como todas as suas derivadas, se anulam no infinito,

$$\mathcal{B}_0 = C_0^\infty(\mathbb{R}^n) = \left\{ \varphi \in C^\infty(\mathbb{R}^n) \mid \begin{array}{l} \partial_\alpha \varphi \text{ se anula no infinito} \\ \text{para } \alpha \in \mathbb{N}^n \end{array} \right\}. \quad (3.142)$$

Para todo aberto Ω de \mathbb{R}^n , definimos ainda

$$\mathcal{B}(\Omega) = C_b^\infty(\Omega) = \left\{ \varphi \in C^\infty(\Omega) \mid \begin{array}{l} \partial_\alpha \varphi \text{ é limitada} \\ \text{para } \alpha \in \mathbb{N}^n \end{array} \right\}, \quad (3.143)$$

e

$$\mathcal{B}_0(\Omega) = C_0^\infty(\Omega) = \left\{ \varphi \in C^\infty(\Omega) \mid \begin{array}{l} \partial_\alpha \varphi \text{ se anula no bordo de } \Omega \\ \text{para } \alpha \in \mathbb{N}^n \end{array} \right\}. \quad (3.144)$$

A estrutura topológica padrão destes espaços também já foi especificada no Exemplo 3.5: é a *topologia uniforme*, ou seja, a topologia da convergência uniforme de todas as derivadas sobre Ω , gerada pela sequência de normas $\|\cdot\|_{r,\Omega}$ ($r = 0, 1, 2, \dots$) definidas na equação (3.43).

Proposição 3.29 Para todo aberto Ω de \mathbb{R}^n , em relação à topologia uniforme, $\mathcal{B}_0(\Omega)$ é fechado em $\mathcal{B}(\Omega)$.

¹¹ Tais funções costumam ser chamadas de totalmente limitadas.

DEMONSTRAÇÃO: Seja $(\varphi_k)_{k \in \mathbb{N}}$ uma sequência de funções $\varphi_k \in \mathcal{B}_0(\Omega)$ que converge para uma função $\varphi \in \mathcal{B}(\Omega)$ na topologia uniforme. Então dado $\alpha \in \mathbb{N}^n$ e $\epsilon > 0$, existem $k \in \mathbb{N}$ tal que $\|\varphi - \varphi_k\|_{|\alpha|, \Omega} < 2^{|\alpha|-1}\epsilon$ e um compacto $K \subset \Omega$ tal que $|\partial_\alpha \varphi_k(x)| < \epsilon/2$ para todo $x \in \Omega \setminus K$, o que implica $|\partial_\alpha \varphi(x)| < \epsilon$ para todo $x \in \Omega \setminus K$, mostrando que $\varphi \in \mathcal{B}_0(\Omega)$.

□

Agora, observamos que, em relação a esta topologia uniforme, a) tanto $\mathcal{B}(\Omega)$ como $\mathcal{B}_0(\Omega)$ são completos e b) as inclusões $\mathcal{D}(\Omega) \subset \mathcal{B}_0(\Omega)$ e $\mathcal{B}(\Omega) \subset \mathcal{E}(\Omega)$ são contínuas, ou seja, a topologia de $\mathcal{D}(\Omega)$ é mais fina do que a topologia induzida pela topologia de $\mathcal{B}_0(\Omega)$ e a topologia de $\mathcal{B}(\Omega)$ é mais fina do que a topologia induzida pela topologia de $\mathcal{E}(\Omega)$ (novamente, deixamos os detalhes do argumento como exercício para o leitor). Além disso, vale a seguinte

Proposição 3.30 Para todo aberto Ω de \mathbb{R}^n , em relação à topologia uniforme, $\mathcal{D}(\Omega)$ é denso em $\mathcal{B}_0(\Omega)$.

DEMONSTRAÇÃO: Considere qualquer uma das normas $\|\cdot\|_{r, \Omega}$ (definidas na equação (3.43)) que geram a topologia uniforme de $\mathcal{B}_0(\Omega)$, onde $r \in \mathbb{N}$, e note que, para toda função $\varphi \in \mathcal{B}_0(\Omega)$ e todo $\epsilon > 0$, existe um compacto $K \subset \Omega$ (que depende de φ , r e ϵ) tal que, para todo $\alpha \in \mathbb{N}^n$ com $|\alpha| \leq r$, $|\partial_\alpha \varphi(x)| < 2^{-r}\epsilon$ para $x \in \Omega \setminus K$ e, logo, $\|\varphi\|_{r, \Omega \setminus K} \leq \epsilon$.

Portanto, escolhendo uma função teste $\chi \in \mathcal{D}(\Omega)$ tal que $0 \leq \chi \leq 1$ e $\chi = 1$ sobre uma vizinhança de K , conforme o Teorema 3.4, obtém-se $\chi\varphi \in \mathcal{D}(\Omega)$ e $\|\varphi - \chi\varphi\|_{r, \Omega} = \|(1 - \chi)\varphi\|_{r, \Omega \setminus K} \leq \|(1 - \chi)\|_{r, \Omega \setminus K} \|\varphi\|_{r, \Omega}$.

□

Usando um índice u para indicar que se usa a topologia uniforme, as últimas duas proposições implicam que $\mathcal{B}_{0,u}(\Omega)$ é um espaço de funções teste sobre Ω e portanto seu dual $\mathcal{B}'_{0,u}(\Omega)$ é um espaço de distribuições sobre Ω , mas o mesmo não vale para o espaço $\mathcal{B}_u(\Omega)$ e seu dual $\mathcal{B}'_u(\Omega)$.

O problema que se evidencia a partir desta observação é que nem $\mathcal{B}'_u(\Omega)$ nem $\mathcal{B}'_{0,u}(\Omega)$ serve como candidato para o que seria o espaço das distribuições integráveis sobre Ω , uma vez que $\mathcal{B}_u(\Omega)$ não é um espaço de funções teste enquanto que $\mathcal{B}_{0,u}(\Omega)$ não contém a função 1, e isso é uma obstrução fatal pois como sugere a equação (3.80), a integral de uma distribuição integrável deveria ser o seu valor sobre a função 1 e portanto o espaço das distribuições integráveis deveria ser o dual de um espaço de funções teste que contém a função 1.

A solução para este problema consiste em uma mudança de topologia, passando da topologia uniforme para uma nova topologia localmente convexa mais grossa chamada a *topologia estrita*, cuja definição é inspirada em uma construção análoga na teoria das C^* -álgebras – a topologia estrita na álgebra dos multiplicadores de uma C^* -álgebra (e que funciona mesmo para álgebras que não são comutativas). Essa topologia é definida de forma implícita no trabalho original de Schwartz [?, p. 203], porém sem indicação de uma família de seminormas que a gere – uma lacuna que preencheremos a seguir.

Proposição 3.31 Para todo aberto Ω de \mathbb{R}^n , em relação à topologia estrita, $\mathcal{D}(\Omega)$ é denso em $\mathcal{B}(\Omega)$.

DEMONSTRAÇÃO: Considerando qualquer uma das seminormas $\|\cdot\|_{r,K}$ que geram a topologia de $\mathcal{E}(\Omega)$, onde $r \in \mathbb{N}$ e K é um subconjunto compacto de Ω (veja a equação (3.71)), note que escolhendo uma função

Dado $\varphi \in \mathcal{B}(\Omega)$ e $\epsilon > 0$,

□

Corolário 3.2 Para todo aberto Ω de \mathbb{R}^n , o espaço $\mathcal{B}_s(\Omega)$ das funções lisas totalmente limitadas sobre Ω , munido da topologia estrita, é um espaço de funções teste sobre Ω , e o seu dual $\mathcal{B}'_s(\Omega)$ é um espaço de distribuições sobre Ω cujos elementos são chamados de **distribuições integráveis** sobre Ω , sendo que a **integral** de uma tal distribuição é o seu valor sobre a função $1 \in \mathcal{B}(\Omega)$.

3.9 Convergência de sequências e séries

Definição 3.27 Para todo aberto Ω de \mathbb{R}^n , uma sequência $(T_k)_{k \in \mathbb{N}}$ de distribuições T_k sobre Ω converge em $\mathcal{D}'(\Omega)$ para uma distribuição T sobre Ω se, para toda função teste φ sobre Ω , vale

$$\langle T_k, \varphi \rangle \longrightarrow \langle T, \varphi \rangle \quad \text{quando } k \rightarrow \infty .$$

Observamos que esta noção de convergência se refere à topologia fraca no dual topológico $\mathcal{D}'(\Omega)$ de $\mathcal{D}(\Omega)$, introduzida no final da Seção 3.2. Trata-se de uma convergência fraca no sentido de que a mera convergência pontual sobre cada função teste constitui uma condição fraca (ou seja, pouco restritiva) para que uma sequência $(T_k)_{k \in \mathbb{N}}$ de distribuições $T_k \in \mathcal{D}'(\Omega)$ seja convergente em $\mathcal{D}'(\Omega)$. Portanto, deve ser fácil encontrar tais sequências e assim, deve haver muitas. E de fato, como veremos mais adiante, existe um grande número de sequências de funções que, embora sejam divergentes em relação a qualquer noção de convergência da análise clássica, convergem no sentido de distribuições.

A grande utilidade da noção de convergência fraca de sequências de distribuições decorre do seguinte teorema fundamental, devido a Laurent Schwartz:

Teorema 3.8 (Schwartz): Seja $(T_k)_{k \in \mathbb{N}}$ uma sequência de distribuições T_k sobre um aberto Ω de \mathbb{R}^n tal que, para toda função teste φ sobre Ω , a sequência numérica $(\langle T_k, \varphi \rangle)_{k \in \mathbb{N}}$ seja convergente. Então o funcional T sobre $\mathcal{D}(\Omega)$ definido por

$$\langle T, \varphi \rangle = \lim_{k \rightarrow \infty} \langle T_k, \varphi \rangle \quad \text{para } \varphi \in \mathcal{D}(\Omega)$$

é uma distribuição sobre Ω : $T \in \mathcal{D}'(\Omega)$.

Observamos que é fácil provar que o funcional T sobre $\mathcal{D}(\Omega)$ definido por esta fórmula é linear, sendo que a parte profunda do teorema, cuja demonstração requer ferramentas avançadas de análise funcional, consiste na afirmação de que ele é contínuo.

Os seguintes dois exemplos tratam de duas situações diferentes em que há convergência fraca, no sentido de distribuições, de seqüências de funções localmente integráveis:

Exemplo 3.15 Seja $(f_k)_{k \in \mathbb{N}}$ uma seqüência de funções localmente integráveis f_k sobre \mathbb{R}^n que converge em $\mathcal{L}_{loc}^1(\mathbb{R}^n)$ para uma função localmente integrável f sobre \mathbb{R}^n : isso significa que

$$f_k|_K \rightarrow f|_K \text{ em } \mathcal{L}^1(K) \quad \text{para todo compacto } K \text{ de } \mathbb{R}^n .$$

Então

$$T_{f_k} \rightarrow T_f \text{ em } \mathcal{D}' .$$

DEMONSTRAÇÃO: Segue diretamente da seguinte simples estimativa, válida para qualquer função teste $\varphi \in \mathcal{D}$, pondo $K = \text{supp } \varphi$:

$$\begin{aligned} |\langle T_{f_k}, \varphi \rangle - \langle T_f, \varphi \rangle| &= \left| \int d^n x (f_k(x) - f(x)) \varphi(x) \right| \\ &\leq \int_K d^n x |f_k(x) - f(x)| \cdot \sup_{x \in K} |\varphi(x)| \\ &= \|(f_k - f)|_K\|_1 \|\varphi\|_{0,K} \\ &\rightarrow 0 \quad \text{quando } k \rightarrow \infty . \end{aligned}$$

□

Exemplo 3.16 Seja $(f_k)_{k \in \mathbb{N}}$ uma seqüência de funções localmente integráveis f_k sobre \mathbb{R}^n com as seguintes propriedades:

- (i) $f_k \rightarrow 0$ quase sempre sobre \mathbb{R}^n ;
- (ii) para todo compacto K de \mathbb{R}^n com $0 \notin K$, $f_k|_K \rightarrow 0$ em $\mathcal{L}^1(K)$, ou seja,

$$\int_K d^n x f_k(x) \rightarrow 0 \quad \text{e} \quad \int_K d^n x |f_k(x)| \rightarrow 0 ;$$

- (iii) para todo compacto K de \mathbb{R}^n com $0 \in K$,

$$\int_K d^n x f_k(x) \rightarrow 1 \quad \text{e} \quad \int_K d^n x |f_k(x)| \leq C_K ,$$

onde $C_K > 0$ é alguma constante que independe de k .

Então

$$T_{f_k} \rightarrow \delta \text{ em } \mathcal{D}' .$$

DEMONSTRAÇÃO: Inicialmente, fixamos uma função teste $\varphi \in \mathcal{D}$ e escolhemos um compacto K de \mathbb{R}^n contendo o suporte de φ e também uma bola fechada \bar{B} em torno da origem, digamos de raio 1. Então para qualquer $\rho > 0$ com $\rho < 1$, escrevemos K como a união de dois subconjuntos compactos, a saber, a bola fechada \bar{B}_ρ de raio ρ em torno da origem e o complemento K_ρ em K da bola aberta B_ρ de raio ρ em torno da origem. Tendo em vista que a interseção destes dois compactos é a esfera de raio ρ em torno da origem, que é de medida 0, obtemos a seguinte estimativa:

$$\begin{aligned} |\langle T_{f_k} - \delta, \varphi \rangle| &= \left| \int_K d^n x f_k(x) \varphi(x) - \varphi(0) \right| \\ &= \left| \int_{K_\rho} d^n x f_k(x) \varphi(x) + \int_{\bar{B}_\rho} d^n x f_k(x) (\varphi(x) - \varphi(0)) \right. \\ &\quad \left. + \left(\int_{\bar{B}_\rho} d^n x f_k(x) - 1 \right) \varphi(0) \right| \\ &\leq \sup_{x \in K_\rho} |\varphi(x)| \cdot \int_{K_\rho} d^n x |f_k(x)| \\ &\quad + \sup_{x \in \bar{B}_\rho} |\varphi(x) - \varphi(0)| \cdot \int_{\bar{B}_\rho} d^n x |f_k(x)| \\ &\quad + |\varphi(0)| \left| \int_{\bar{B}_\rho} d^n x f_k(x) - 1 \right| \\ &\leq \|\varphi\|_K \cdot \int_{K_\rho} d^n x |f_k(x)| + C_K \sup_{x \in \bar{B}_\rho} |\varphi(x) - \varphi(0)| \\ &\quad + \|\varphi\|_K \left| \int_{\bar{B}_\rho} d^n x f_k(x) - 1 \right|. \end{aligned}$$

Isso posto, podemos usar a continuidade de φ na origem para concluir que, dado qualquer $\epsilon > 0$, existe $\rho > 0$ com $\rho < 1$ tal que

$$\sup_{x \in \bar{B}_\rho} |\varphi(x) - \varphi(0)| \leq \frac{\epsilon}{3C_K} .$$

As hipóteses (ii) e (iii) garantem então que existe $k_0 \in \mathbb{N}$ tal que para $k \geq k_0$, vale

$$\int_{K_\rho} d^n x |f_k(x)| \leq \frac{\epsilon}{3\|\varphi\|_K}$$

e

$$\left| \int_{\bar{B}_\rho} d^n x f_k(x) - 1 \right| \leq \frac{\epsilon}{3 \|\varphi\|_K}$$

e portanto

$$|\langle T_{f_k} - \delta, \varphi \rangle| \leq \epsilon .$$

□

Para funções reais não-negativas e integráveis, podemos simplificar as hipóteses:

Exemplo 3.17 Seja $(f_k)_{k \in \mathbb{N}}$ uma seqüência de funções reais não-negativas e integráveis f_k sobre \mathbb{R}^n normalizadas por

$$\int d^n x f_k(x) = 1 \quad \text{para todo } k \in \mathbb{N}$$

e tais que para todo compacto K de \mathbb{R}^n com $0 \notin K$, $f_k|_K \rightarrow 0$ em $\mathcal{L}^1(K)$. Então

$$T_{f_k} \rightarrow \delta \text{ em } \mathcal{D}' .$$

Outros exemplos podem ser deduzidos da seguinte

Proposição 3.32 Seja Ω um aberto de \mathbb{R}^n e seja $L : \mathcal{D}(\Omega) \rightarrow \mathcal{D}(\Omega)$ um operador linear contínuo. Então seu dual ou transposto $L' : \mathcal{D}'(\Omega) \rightarrow \mathcal{D}'(\Omega)$, definido por

$$\langle L'(T), \varphi \rangle = \langle T, L(\varphi) \rangle \quad \text{para } T \in \mathcal{D}'(\Omega), \varphi \in \mathcal{D}(\Omega)$$

também é um operador linear contínuo.

DEMONSTRAÇÃO : Daremos duas demonstrações diferentes, a primeira em termos de seqüências convergentes e a segunda usando seminormas.

Primeiro, seja $(T_k)_{k \in \mathbb{N}}$ uma seqüência de distribuições T_k sobre Ω . Conforme a Definição 3.27, temos que $T_k \rightarrow T$ em $\mathcal{D}'(\Omega)$ se e somente se para qualquer função teste φ sobre Ω , vale $\langle T_k, \varphi \rangle \rightarrow \langle T, \varphi \rangle$ e portanto

$$\langle L'(T_k), \varphi \rangle = \langle T_k, L(\varphi) \rangle \rightarrow \langle T, L(\varphi) \rangle = \langle L'(T), \varphi \rangle ,$$

o que significa que $L'(T_k) \rightarrow L'(T)$ em $\mathcal{D}'(\Omega)$.

Segundo, usando que a topologia fraca sobre $\mathcal{D}'(\Omega)$ é definida pela família de seminormas $(s_\varphi)_{\varphi \in \mathcal{D}(\Omega)}$ dadas por $s_\varphi(T) = |\langle T, \varphi \rangle|$ para $T \in \mathcal{D}'(\Omega)$, a Proposição 3.1 mostra que continuidade de L' é uma simples consequência da igualdade

$$s_\varphi(L'(T)) = s_{L(\varphi)}(T) \quad \text{para } T \in \mathcal{D}'(\Omega) .$$

□

Note que o segundo método, além de mais simples, demonstra um fato mais geral do que o primeiro, tendo em vista que o espaço $\mathcal{D}'(\Omega)$ com a topologia fraca não é metrizável; portanto, o segundo método é que demonstra continuidade de L' , enquanto que o primeiro demonstra apenas continuidade sequencial de L' . No entanto, esta será suficiente para as aplicações.

Tendo em vista a Proposição 3.19 e a Definição 3.21, temos como corolário

Proposição 3.33 *Para todo aberto Ω de \mathbb{R}^n , o operador linear de diferenciação parcial $\partial_\alpha : \mathcal{D}'(\Omega) \rightarrow \mathcal{D}'(\Omega)$ é contínuo. Em particular, para qualquer sequência $(T_k)_{k \in \mathbb{N}}$ de distribuições T_k sobre Ω , vale*

$$T_k \rightarrow T \text{ em } \mathcal{D}'(\Omega) \quad \implies \quad \partial_\alpha T_k \rightarrow \partial_\alpha T \text{ em } \mathcal{D}'(\Omega) .$$

Outro corolário, baseado na Proposição 3.22 e na Definição 3.23, é a seguinte

Proposição 3.34 *Para todo aberto Ω de \mathbb{R}^n , o operador bilinear de multiplicação*

$$\begin{aligned} \mathcal{E}(\Omega) \times \mathcal{D}'(\Omega) &\longrightarrow \mathcal{D}'(\Omega) \\ (f, T) &\longmapsto fT \end{aligned} \tag{3.145}$$

é separadamente contínuo, o que significa que

- Para toda função $f \in \mathcal{E}(\Omega)$, a aplicação linear

$$\begin{aligned} \mathcal{D}'(\Omega) &\longrightarrow \mathcal{D}'(\Omega) \\ T &\longmapsto fT \end{aligned} \tag{3.146}$$

é contínua. Em particular, para qualquer sequência $(T_k)_{k \in \mathbb{N}}$ de distribuições T_k sobre Ω , vale

$$T_k \rightarrow T \text{ em } \mathcal{D}'(\Omega) \quad \implies \quad fT_k \rightarrow fT \text{ em } \mathcal{D}'(\Omega) .$$

- Para toda distribuição $T \in \mathcal{D}'(\Omega)$, a aplicação linear

$$\begin{aligned} \mathcal{E}(\Omega) &\longrightarrow \mathcal{D}'(\Omega) \\ f &\longmapsto fT \end{aligned} \tag{3.147}$$

é contínua. Em particular, para qualquer sequência $(f_k)_{k \in \mathbb{N}}$ de funções lisas f_k sobre Ω , vale

$$f_k \rightarrow f \text{ em } \mathcal{E}(\Omega) \quad \implies \quad f_k T \rightarrow fT \text{ em } \mathcal{D}'(\Omega) .$$

Outros exemplos serão apresentados nas seções seguintes.

3.10 Distribuições periódicas e séries de Fourier

Nossa primeira meta nesta seção será estender a noção de uma função periódica para a noção de uma distribuição periódica. Trata-se de um problema que leva à questão de como efetuar transformações das variáveis independentes em distribuições. Conforme foi mencionado no início do Capítulo 2, transformações de coordenadas gerais são descritas por difeomorfismos, e não é muito difícil se convencer que o espaço \mathcal{D} das funções teste e seus subespaços $\mathcal{D}(\Omega)$ associados aos abertos Ω de \mathbb{R}^n , assim como os correspondentes espaços \mathcal{D}' e $\mathcal{D}'(\Omega)$ de distribuições, são invariantes em relação à ação de difeomorfismos: este fato constitui o ingrediente crucial para a construção de uma teoria de distribuições sobre variedades. Não pretendemos nos aprofundar nesta direção aqui e portanto nos restringiremos a discutir o caso de transformações afins.

Uma aplicação afim de \mathbb{R}^n em \mathbb{R}^n é uma aplicação

$$\begin{aligned} (a, A) : \mathbb{R}^n &\longrightarrow \mathbb{R}^n \\ x &\longmapsto (a, A) \cdot x \end{aligned}$$

da forma

$$(a, A) \cdot x = Ax + a \quad \text{para } x \in \mathbb{R}^n, \quad (3.148)$$

onde a é um vetor em \mathbb{R}^n e A é uma matriz ($n \times n$). Dizemos que (a, A) é uma transformação afim quando for uma aplicação afim bijetora, o que é equivalente a exigir que a matriz A seja não-singular, i.e. $\det A \neq 0$. Obviamente, as transformações afins de \mathbb{R}^n formam um grupo, que denotaremos por $GA(n, \mathbb{R})$, em relação ao produto dado pela lei de composição, que também pode ser visto como decorrendo da condição de que o símbolo \cdot na equação (3.148) represente uma ação do grupo afim sobre o espaço \mathbb{R}^n , ou seja, que valha a regra

$$((a, A)(b, B)) \cdot x = (a, A) \cdot ((b, B) \cdot x) \quad \text{para } x \in \mathbb{R}^n. \quad (3.149)$$

Explicitamente,

$$(a, A)(b, B) = (a + Ab, AB), \quad (3.150)$$

pois

$$\begin{aligned} ((a, A)(b, B)) \cdot x &= (a, A) \cdot (Bx + b) = A \cdot (Bx + b) + a \\ &= A(Bx) + (Ab + a) = (a + Ab, AB) \cdot x. \end{aligned}$$

A unidade neste grupo é $(0, 1)$, e o inverso de (a, A) é $(a, A)^{-1} = (-A^{-1}a, A^{-1})$. Isso significa que o grupo das transformações afins de \mathbb{R}^n é o produto semi-direto do grupo das translações, que pode ser identificado com o próprio \mathbb{R}^n , com o grupo $GL(n, \mathbb{R})$ das matrizes ($n \times n$) inversíveis:

$$GA(n, \mathbb{R}) = GL(n, \mathbb{R}) \ltimes \mathbb{R}^n.$$

(As letras “GA” e “GL” significam, respectivamente, “general affine” e “general linear”.)

Isso posto, podemos definir como o grupo afim age sobre funções teste e sobre distribuições. Primeiro, para $(a, A) \in GA(n, \mathbb{R})$ e $\varphi \in \mathcal{D}$, definimos $(a, A) \cdot \varphi \in \mathcal{D}$ por

$$((a, A) \cdot \varphi)(x) = \varphi((a, A)^{-1} \cdot x) = \varphi(A^{-1}(x - a)) \quad \text{para } x \in \mathbb{R}^n. \quad (3.151)$$

A necessidade de aplicar a transformação inversa ao argumento da função φ decorre da exigência de que o símbolo \cdot represente uma ação do grupo afim sobre o espaço \mathcal{D} , ou seja, que valha a regra

$$((a, A)(b, B)) \cdot \varphi = (a, A) \cdot ((b, B) \cdot \varphi) \quad \text{para } \varphi \in \mathcal{D}. \quad (3.152)$$

De fato, a equação (3.151) implica que para $\varphi \in \mathcal{D}$ e $x \in \mathbb{R}^n$,

$$\begin{aligned} ((a, A) \cdot ((b, B) \cdot \varphi))(x) &= ((b, B) \cdot \varphi)((a, A)^{-1} \cdot x) \\ &= \varphi((b, B)^{-1} \cdot ((a, A)^{-1} \cdot x)) \\ &= \varphi(((b, B)^{-1}(a, A)^{-1}) \cdot x) \\ &= \varphi(((a, A)(b, B))^{-1} \cdot x) \\ &= (((a, A)(b, B)) \cdot \varphi)(x). \end{aligned}$$

É fácil provar

Proposição 3.35 *Para qualquer transformação afim $(a, A) \in GA(n, \mathbb{R})$ de \mathbb{R}^n , o operador linear*

$$\begin{aligned} (a, A) : \mathcal{D} &\longrightarrow \mathcal{D} \\ \varphi &\longmapsto (a, A) \cdot \varphi \end{aligned} \quad (3.153)$$

é contínuo.

Usando o princípio geral da dualização para transferir operações do âmbito das funções teste ao das distribuições, chegamos à

Definição 3.28 *Para qualquer distribuição $T \in \mathcal{D}'$, define-se sua transformada sob a transformação afim $(a, A) \in GA(n, \mathbb{R})$ como sendo a distribuição $(a, A) \cdot T \in \mathcal{D}'$ dada pela fórmula*

$$\langle (a, A) \cdot T, \varphi \rangle = |\det A| \langle T, (a, A)^{-1} \cdot \varphi \rangle \quad \text{para } \varphi \in \mathcal{D}. \quad (3.154)$$

Observa-se que $(a, A) \cdot T$ é realmente uma distribuição, i.e., um funcional linear contínuo sobre \mathcal{D} , pois (a menos de uma inversão) é a composição do funcional linear contínuo T sobre \mathcal{D} com o operador linear contínuo da proposição anterior. Ademais, a Proposição 3.32 implica

Proposição 3.36 *Para qualquer transformação afim $(a, A) \in GA(n, \mathbb{R})$ de \mathbb{R}^n , o operador linear*

$$\begin{aligned} (a, A) : \mathcal{D}' &\longrightarrow \mathcal{D}' \\ T &\longmapsto (a, A) \cdot T \end{aligned} \quad (3.155)$$

é contínuo.

Voltando à fórmula (3.154), observamos que, novamente, a necessidade de aplicar a transformação inversa ao argumento do funcional T decorre da exigência de que o símbolo \cdot represente uma ação do grupo afim sobre o espaço \mathcal{D}' , ou seja, que valha a regra

$$((a, A)(b, B)) \cdot T = (a, A) \cdot ((b, B) \cdot T) \quad \text{para } T \in \mathcal{D}' . \quad (3.156)$$

De fato, a equação (3.154) implica que para $T \in \mathcal{D}'$ e $\varphi \in \mathcal{D}$,

$$\begin{aligned} \langle (a, A) \cdot ((b, B) \cdot T), \varphi \rangle &= |\det A| \langle (b, B) \cdot T, (a, A)^{-1} \cdot \varphi \rangle \\ &= |\det A| |\det B| \langle T, (b, B)^{-1} \cdot ((a, A)^{-1} \cdot \varphi) \rangle \\ &= |\det(AB)| \langle T, ((b, B)^{-1} (a, A)^{-1}) \cdot \varphi \rangle \\ &= |\det(AB)| \langle T, ((a, A)(b, B))^{-1} \cdot \varphi \rangle \\ &= \langle ((a, A)(b, B)) \cdot T, \varphi \rangle . \end{aligned}$$

Finalmente, verificamos que a definição da transformação afim do argumento para distribuições (inclusive o determinante) dada pela equação (3.154), quando aplicada a distribuições regulares, coincide com a noção clássica de transformação afim do argumento para funções. Mais exatamente, temos

Proposição 3.37 *Para funções $f \in \mathcal{L}_{loc}^1(\mathbb{R}^n)$, vale*

$$(a, A) \cdot T_f = T_{(a, A) \cdot f} . \quad (3.157)$$

DEMONSTRAÇÃO: Para qualquer função teste $\varphi \in \mathcal{D}$, temos

$$\begin{aligned} \langle (a, A) \cdot T_f, \varphi \rangle &= |\det A| \langle T_f, (a, A)^{-1} \cdot \varphi \rangle \\ &= |\det A| \int d^n x f(x) ((a, A)^{-1} \cdot \varphi)(x) \\ &= |\det A| \int d^n x f(x) \varphi(Ax + a) \\ &= \int d^n y f(A^{-1}(y - a)) \varphi(y) \\ &= \langle T_{(a, A) \cdot f}, \varphi \rangle . \end{aligned}$$

□

Este resultado sugere complementar a notação introduzida nas equações (3.85) e (3.86), da seguinte forma: escrevendo

$$\langle (a, A)^{-1} \cdot T, \varphi \rangle \equiv \int d^n x T(Ax + a) \varphi(x) , \quad (3.158)$$

a equação (3.154) assume a forma

$$\int d^n x T(Ax + a) \varphi(x) = \frac{1}{|\det A|} \int d^n y T(y) \varphi(A^{-1}(y - a)) , \quad (3.159)$$

e escrevendo

$$\langle (a, A)^{-1} \cdot T, \varphi \rangle \equiv \langle T(Ax + a), \varphi(x) \rangle, \quad (3.160)$$

a equação (3.154) assume a forma

$$\langle T(Ax + a), \varphi(x) \rangle = \frac{1}{|\det A|} \langle T(y), \varphi(A^{-1}(y - a)) \rangle. \quad (3.161)$$

Exercício 3.7 Prove que

$$\langle \delta(x - a), \varphi(x) \rangle = \varphi(a), \quad (3.162)$$

$$\langle \delta^{(\alpha)}(x - a), \varphi(x) \rangle = (-1)^{|\alpha|} \partial_\alpha \varphi(a), \quad (3.163)$$

e

$$\delta(Ax) = \frac{1}{|\det A|} \delta(x). \quad (3.164)$$

No restante desta seção, restringiremo-nos a considerar a situação unidimensional $n = 1$. Mencionamos apenas que a noção de distribuição periódica e a técnica de séries de Fourier para a análise de tais distribuições podem ser generalizadas ao caso multidimensional $n > 1$, onde a periodicidade se refere a translações por vetores pertencendo a um reticulado $\Lambda \cong \mathbb{Z}^n$ em \mathbb{R}^n . Também efetuaremos uma pequena mudança de notação, usando a letra t (ao invés de x) para indicar a variável independente e a letra τ (ao invés de a) para indicar o período básico (que será fixo durante toda a discussão que segue); a correspondente frequência angular é $\omega = 2\pi/\tau$.

Definição 3.29 Uma distribuição $T \in \mathcal{D}'$ sobre \mathbb{R} é chamada **periódica, de período** τ , se

$$T(t + \tau) = T(t), \quad (3.165)$$

i.e., se para toda função teste $\varphi \in \mathcal{D}$ sobre \mathbb{R} ,

$$\langle T(t), \varphi(t - \tau) \rangle = \langle T(t), \varphi(t) \rangle. \quad (3.166)$$

Obviamente, a distribuição T_f associada a uma função f localmente integrável será periódica se e somente se a função f é periódica no sentido usual (exceto sobre um conjunto de medida zero), i.e., satisfaz

$$f(t + \tau) = f(t) \quad \text{para quase todo } t \in \mathbb{R}. \quad (3.167)$$

Uma tal função é completamente determinada (exceto sobre um conjunto de medida zero) por seus valores em qualquer intervalo de comprimento τ , por exemplo no intervalo $[0, \tau]$. A seguir, frequentemente identificaremos funções localmente integráveis com as distribuições que geram.

As funções periódicas elementares de período τ são as funções trigonométricas

$$\cos(k\omega t) \quad \text{e} \quad \sin(k\omega t) \quad , \quad k = 1, 2, 3, \dots ,$$

em conjunto com a função constante 1. Uma forma equivalente de representar estas funções é através das exponenciais

$$\exp(ik\omega t) = \cos(k\omega t) + i \sin(k\omega t) \quad , \quad k \in \mathbb{Z} .$$

Observa-se a seguinte relação de ortogonalidade, que é fundamental:

$$\frac{1}{\tau} \int_0^\tau dt \exp(-ik\omega t) \exp(il\omega t) = \delta_{kl} . \quad (3.168)$$

Definição 3.30 *Uma série de Fourier é uma série de funções periódicas sobre \mathbb{R} , de período τ , da forma*

$$\sum_{k=-\infty}^{\infty} a_k \exp(ik\omega t) , \quad (3.169)$$

com coeficientes complexos $a_k \in \mathbb{C}$.

Um dos problemas principais do qual trata a teoria clássica das séries de Fourier é estabelecer noções e critérios de convergência para tais séries, ou seja, decidir em qual sentido e sob quais hipóteses existe o limite

$$f(t) = \lim_{m \rightarrow \infty} f_m(t) \quad (3.170)$$

das somas parciais

$$f_m(t) = \sum_{k=-m}^m a_k \exp(ik\omega t) \quad (3.171)$$

e quais seriam as suas propriedades, isto é, qual seria o espaço de funções periódicas ao qual pertence esse limite. As seguintes duas proposições fornecem exemplos simples de afirmações desta natureza.

Proposição 3.38 *Seja $(a_k)_{k \in \mathbb{Z}}$ uma sequência de números complexos satisfazendo a seguinte condição de decaimento para $|k| \rightarrow \infty$: existem constantes $\epsilon > 0$, $C > 0$ e $k_0 \in \mathbb{N}$ tais que*

$$|a_k| \leq \frac{C}{|k|^{1+\epsilon}} \quad \text{para } k \in \mathbb{Z} \text{ com } |k| \geq k_0 . \quad (3.172)$$

Então a série de Fourier (3.169) converge absoluta e uniformemente sobre $[0, \tau]$ para uma função periódica contínua:

$$f(t) = \sum_{k=-\infty}^{\infty} a_k \exp(ik\omega t) \quad \text{em } C([0, \tau]) . \quad (3.173)$$

DEMONSTRAÇÃO: Usando o fato de que

$$\sum_{k=0}^{\infty} \frac{1}{k^{1+\epsilon}} < \infty ,$$

vemos que a série de Fourier (3.169) converge absoluta e uniformemente, pois para $m, n \in \mathbb{N}$ com $n \geq m \geq k_0$ e $t \in [0, \tau]$, temos

$$\begin{aligned} |f_n(t) - f_m(t)| &\leq \sum_{k=-n}^{-m-1} |a_k \exp(ik\omega t)| + \sum_{k=m+1}^n |a_k \exp(ik\omega t)| \\ &= \sum_{k=-n}^{-m-1} |a_k| + \sum_{k=m+1}^n |a_k| \leq \sum_{k=m+1}^n \frac{2C}{k^{1+\epsilon}}, \end{aligned}$$

mostrando que a sequência das somas parciais da série de Fourier (3.169) é uma sequência de Cauchy na topologia da convergência uniforme sobre o intervalo compacto $[0, \tau]$, ou seja, no espaço de Banach $C([0, \tau])$, e portanto converge nesta topologia. \square

Proposição 3.39 *Seja $(a_k)_{k \in \mathbb{Z}}$ uma sequência de números complexos satisfazendo a seguinte condição de decaimento para $|k| \rightarrow \infty$: existem um número natural $r \in \mathbb{N}$ e constantes $\epsilon > 0$, $C > 0$ e $k_0 \in \mathbb{N}$ tais que*

$$|a_k| \leq \frac{C}{|k|^{r+1+\epsilon}} \quad \text{para } k \in \mathbb{Z} \text{ com } |k| \geq k_0. \quad (3.174)$$

Então a série de Fourier (3.169) converge absolutamente, na topologia da convergência uniforme das derivadas até ordem r sobre $[0, \tau]$, para uma função periódica de classe C^r :

$$f(t) = \sum_{k=-\infty}^{\infty} a_k \exp(ik\omega t) \quad \text{em } C^r([0, \tau]). \quad (3.175)$$

Ademais, para $0 \leq p \leq r$, a série de Fourier obtida diferenciando a série de Fourier original (3.169) p vezes, termo a termo, converge absoluta e uniformemente sobre $[0, \tau]$ para a p -ésima derivada $f^{(p)}$ de f :

$$f^{(p)}(t) = \sum_{k=-\infty}^{\infty} a_k (ik\omega)^p \exp(ik\omega t) \quad \text{em } C([0, \tau]). \quad (3.176)$$

DEMONSTRAÇÃO: Conforme a proposição anterior, as hipóteses desta proposição garantem que para $0 \leq p \leq r$, a série de Fourier

$$\sum_{k=-\infty}^{\infty} a_k (ik\omega)^p \exp(ik\omega t)$$

obtida diferenciando a série de Fourier original (3.169) p vezes, termo a termo, converge absoluta e uniformemente sobre $[0, \tau]$. Portanto, segundo um teorema de análise clássica, f é de classe C^r e, para $0 \leq p \leq r$, sua p -ésima derivada $f^{(p)}$ é igual à função representada por esta série. \square

Reciprocamente, dada uma função periódica, surge a questão se e em qual sentido ela pode ser representada por uma série de Fourier. Nesta direção, observamos primeiro que se a função f é representada por uma série de Fourier, a relação de ortogonalidade (3.168) permite calcular os coeficientes a_k em termos de f , através da fórmula

$$a_k = \frac{1}{\tau} \int_0^\tau dt \exp(-ik\omega t) f(t), \quad (3.177)$$

desde que seja permitida a troca entre a integração e o limite. (Isto ocorre, por exemplo, quando a série (3.169) converge uniformemente sobre $[0, \tau]$ e portanto f é contínua.) Mas de qualquer modo, os coeficientes a_k definidos pela equação (3.177) são bem definidos para qualquer função periódica e localmente integrável $f \in \mathcal{L}_{loc}^1(\mathbb{R})$ e são chamados os *coeficientes de Fourier* de f ; quando substituídos em (3.169), eles definem a *série de Fourier* de f . Infelizmente, esta série pode ser divergente, mesmo sob a hipótese de que f seja contínua. Em outras palavras, é necessário impor condições mais restritivas sobre f para poder garantir que sua série de Fourier seja absoluta e uniformemente convergente. Uma condição suficiente é que f seja de classe C^2 , pois esta hipótese permite integrar a definição (3.177) duas vezes por partes para chegar a uma estimativa da forma

$$|a_k| \leq \frac{C}{|k|^2} \quad \text{para } k \in \mathbb{Z} \text{ com } |k| \geq k_0$$

que, por sua vez, permite aplicar a Proposição 3.38.

O análogo das Proposições 3.38 e 3.39 para distribuições requer hipóteses muito menos restritivas para garantir que o limite

$$T = \lim_{m \rightarrow \infty} f_m \quad (3.178)$$

exista em \mathcal{D}' :

Teorema 3.9 *Seja $(a_k)_{k \in \mathbb{Z}}$ uma sequência de números complexos polinomialmente limitada para $|k| \rightarrow \infty$, ou seja, tal que existem constantes $C > 0$ e $N, k_0 \in \mathbb{N}$ tais que*

$$|a_k| \leq C |k|^N \quad \text{para } k \in \mathbb{Z} \text{ com } |k| \geq k_0. \quad (3.179)$$

Então a série de Fourier (3.178) converge em \mathcal{D}' para uma distribuição periódica:

$$T(t) = \sum_{k=-\infty}^{\infty} a_k \exp(ik\omega t) \quad \text{em } \mathcal{D}'. \quad (3.180)$$

Ademais, para $p \in \mathbb{N}$, a série de Fourier obtida diferenciando a série de Fourier original (3.178) p vezes, termo a termo, converge em \mathcal{D}' para a p -ésima derivada $T^{(p)}$ de T :

$$T^{(p)}(t) = \sum_{k=-\infty}^{\infty} a_k (ik\omega)^p \exp(ik\omega t) \quad \text{em } \mathcal{D}'. \quad (3.181)$$

DEMONSTRAÇÃO : Considere a sequência $(b_k)_{k \in \mathbb{Z}}$ de números complexos definidos por

$$b_0 = 0 \quad , \quad b_k = \frac{a_k}{(ik\omega)^{N+2}} \quad \text{para } k \in \mathbb{Z} \setminus \{0\} .$$

Então conforme a Proposição 3.38, a correspondente série de Fourier converge absoluta e uniformemente sobre $[0, \tau]$ para uma função periódica contínua f :

$$f(t) = \sum_{k=-\infty}^{\infty} b_k \exp(ik\omega t) \quad \text{em } C([0, \tau]) .$$

Considerando f como distribuição e usando a Proposição 3.33, podemos diferenciar esta série $N + 2$ vezes, termo a termo, para obter a equação (3.180):

$$T(t) = a_0 + f^{(N+2)}(t) = \sum_{k=-\infty}^{\infty} a_k \exp(ik\omega t) \quad \text{em } \mathcal{D}' .$$

Finalmente, diferenciando mais p vezes, termo a termo, obtemos a equação (3.181). □

Reciprocamente, dada uma distribuição periódica, surge a questão se ela pode ser representada por uma série de Fourier. O problema principal aqui é dar uma definição adequada dos coeficientes de Fourier de uma distribuição periódica. Para abordar esta questão de uma maneira sistemática, mostra-se conveniente introduzir uma interpretação alternativa de distribuições periódicas, a saber como funcionais lineares contínuos sobre um espaço de funções teste periódicas.

Começamos por introduzir este espaço de funções teste periódicas, mais exatamente o espaço \mathcal{E}_τ das funções lisas que são periódicas de período τ :

$$\mathcal{E}_\tau = \{ \varphi \in \mathcal{E}(\mathbb{R}) \mid \varphi(t + \tau) = \varphi(t) \text{ para } t \in \mathbb{R} \} . \quad (3.182)$$

O espaço \mathcal{E}_τ é localmente convexo e metrizável em relação à topologia gerada pelas seminormas

$$\|\varphi\|_r = \sum_{s=0}^r \sup_{t \in \mathbb{R}} |\varphi^{(s)}(t)| \quad \text{para } \varphi \in \mathcal{E}_\tau , \quad (3.183)$$

onde $r = 0, 1, 2, \dots$: é a topologia da convergência uniforme de todas as derivadas sobre \mathbb{R} , ou equivalentemente, sobre qualquer intervalo compacto I contendo pelo menos um intervalo de periodicidade. Portanto, segundo a Proposição 3.3, temos para qualquer sequência $(\varphi_k)_{k \in \mathbb{N}}$ de funções $\varphi_k \in \mathcal{E}_\tau$

$$\varphi_k \rightarrow 0 \text{ em } \mathcal{E}_\tau \iff \varphi_k^{(r)} \rightarrow 0 \text{ uniformemente sobre } \mathbb{R} \text{ para todo } r \in \mathbb{N} . \quad (3.184)$$

Agora, consideramos o operador linear contínuo

$$\begin{aligned} \mathcal{D} &\longrightarrow \mathcal{E}_\tau \\ \varphi &\longmapsto \varphi_\tau \end{aligned} \quad (3.185)$$

de *continuação periódica* que associa a cada função teste φ de suporte compacto a função teste φ_τ periódica de período τ definida por

$$\varphi_\tau(t) = \sum_{n=-\infty}^{\infty} \varphi(t - n\tau). \quad (3.186)$$

Para justificar esta construção, observamos que o suporte de φ sendo compacto, a restrição de φ_τ a qualquer intervalo finito I se reduz a uma soma finita de translados de φ ; explicitamente, para qualquer compacto K de \mathbb{R} ,

$$\mathbb{Z}(I, K) = \{n \in \mathbb{Z} \mid (I - n\tau) \cap K \neq \emptyset\}$$

é um subconjunto finito de \mathbb{Z} tal que

$$\varphi_\tau|_I = \sum_{n \in \mathbb{Z}(I, K)} ((n\tau, 1) \cdot \varphi)|_I \quad \text{se } \text{supp } \varphi \subset K,$$

o que demonstra que a função φ_τ é bem definida e lisa, além de periódica. Linearidade e continuidade do operador (3.185) assim definido também são facilmente verificadas a partir das definições, pois se I é um intervalo compacto de periodicidade $([0, \tau]$, por exemplo), K é um compacto de \mathbb{R} qualquer e $|\mathbb{Z}(I, K)|$ é a cardinalidade de $\mathbb{Z}(I, K)$, temos

$$\|\varphi_\tau\|_r \leq |\mathbb{Z}(I, K)| \|\varphi\|_{r, K} \quad \text{para } \varphi \in \mathcal{D}(K).$$

O operador de continuação periódica não é injetor (existem funções teste de suporte compacto cuja continuação periódica se anula identicamente), mas é sobrejetor. Isso segue trivialmente do fato de que ele admite um inverso a direita. Para construir este, note que é fácil encontrar uma função teste χ de suporte compacto tal que $\chi_\tau \equiv 1$: basta escolher qualquer função teste ϕ de suporte compacto tal que $\phi > 0$ sobre um intervalo compacto qualquer contendo pelo menos um intervalo de periodicidade, e por $\chi = \phi/\phi_\tau$. Agora considere o operador linear contínuo

$$\begin{aligned} \mathcal{E}_\tau &\longrightarrow \mathcal{D} \\ \psi &\longmapsto \chi\psi \end{aligned} \quad (3.187)$$

de multiplicação por χ (veja a Proposição 3.22). Tendo em vista que, para $\psi \in \mathcal{E}_\tau$, $(\chi\psi)_\tau = \chi_\tau\psi$, a condição $\chi_\tau \equiv 1$ significa que ele é de fato um inverso a direita do operador de continuação periódica. Este inverso a direita está longe de ser único, pois existem muitas funções teste χ de suporte compacto tais que $\chi_\tau \equiv 1$, mas esta ambiguidade torna-se irrelevante quando consideramos distribuições periódicas.

De fato, sejam $T \in \mathcal{D}'$ uma distribuição periódica de período τ e χ uma função teste de suporte compacto tal que $\chi_\tau \equiv 1$. Esta condição pode ser reformulada como afirmando que a série

$$\sum_{n=-\infty}^{\infty} \chi(t - n\tau)$$

converge em \mathcal{E} para a função 1. Portanto, conforme a Proposição 3.34, a série

$$\sum_{n=-\infty}^{\infty} \chi(t - n\tau) T(t)$$

converge em \mathcal{D}' para a distribuição $T(t)$, o que significa que para toda função teste φ de suporte compacto, vale

$$\sum_{n=-\infty}^{\infty} \langle T(t), \chi(t - n\tau) \varphi(t) \rangle = \langle T(t), \varphi(t) \rangle .$$

Agora, sejam χ_1 e χ_2 funções teste de suporte compacto tais que

$$(\chi_1)_\tau \equiv 1 \equiv (\chi_2)_\tau .$$

Então para toda função teste ψ periódica de período τ , vale

$$\begin{aligned} \langle T(t), \chi_1(t) \psi(t) \rangle &= \sum_{n=-\infty}^{\infty} \langle T(t), \chi_2(t - n\tau) \chi_1(t) \psi(t) \rangle \\ &= \sum_{n=-\infty}^{\infty} \langle T(t + n\tau), \chi_2(t) \chi_1(t + n\tau) \psi(t + n\tau) \rangle \\ &= \sum_{m=-\infty}^{\infty} \langle T(t), \chi_1(t - m\tau) \chi_2(t) \psi(t) \rangle \\ &= \langle T(t), \chi_2(t) \psi(t) \rangle . \end{aligned}$$

Proposição 3.40 *O espaço das distribuições periódicas de período τ pode ser identificado com o espaço \mathcal{E}'_τ , o dual topológico de \mathcal{E}_τ , da seguinte maneira. Cada distribuição periódica $T \in \mathcal{D}'$ de período τ define um funcional linear contínuo $T_\tau \in \mathcal{E}'_\tau$ sobre \mathcal{E}_τ conforme*

$$\langle T_\tau(t), \psi(t) \rangle = \langle T(t), \chi(t) \psi(t) \rangle \quad \text{para } \psi \in \mathcal{E}_\tau , \quad (3.188)$$

onde χ é uma função teste de suporte compacto qualquer tal que $\chi_\tau \equiv 1$, e cada funcional linear contínuo $T_\tau \in \mathcal{E}'_\tau$ sobre \mathcal{E}_τ define uma distribuição periódica $T \in \mathcal{D}'$ de período τ conforme

$$\langle T(t), \varphi(t) \rangle = \langle T_\tau(t), \varphi_\tau(t) \rangle \quad \text{para } \varphi \in \mathcal{D} . \quad (3.189)$$

DEMONSTRAÇÃO: Falta apenas mostrar que as equações (3.188) e (3.189) definem operações mutuamente inversas. Para tanto, notamos que para toda distribuição periódica $T \in \mathcal{D}'$ de período τ e quaisquer duas funções teste χ e φ de suporte compacto, vale

$$\begin{aligned} \langle T(t), \chi(t) \varphi_\tau(t) \rangle &= \sum_{n=-\infty}^{\infty} \langle T(t), \chi(t) \varphi(t - n\tau) \rangle \\ &= \sum_{n=-\infty}^{\infty} \langle T(t + n\tau), \chi(t + n\tau) \varphi(t) \rangle \\ &= \sum_{m=-\infty}^{\infty} \langle T(t), \chi(t - m\tau) \varphi(t) \rangle \\ &= \langle T(t), \chi_\tau(t) \varphi(t) \rangle . \end{aligned}$$

Portanto, se $\psi \in \mathcal{E}_\tau$ é da forma $\psi = \varphi_\tau$ com $\varphi \in \mathcal{D}$ (o que sempre é o caso) e $\chi \in \mathcal{D}$ é tal que $\chi_\tau \equiv 1$, então temos $\langle T, \chi\psi \rangle = \langle T, \chi\varphi_\tau \rangle = \langle T, \chi_\tau\varphi \rangle = \langle T, \varphi \rangle$, e assim as duas equações coincidem. Da mesma maneira, se $\varphi \in \mathcal{D}$ é da forma $\varphi = \chi\psi$ com $\chi \in \mathcal{D}$ tal que $\chi_\tau \equiv 1$ e $\psi \in \mathcal{E}_\tau$, então temos $\varphi_\tau = \psi$, logo $\langle T_\tau, \psi \rangle = \langle T_\tau, \varphi_\tau \rangle$, e novamente as duas equações coincidem. \square

Se $f \in \mathcal{L}_{loc}^1(\mathbb{R})$ é uma função localmente integrável e periódica de período τ (exceto sobre um conjunto de medida zero), então para $\chi \in \mathcal{D}$ tal que $\chi_\tau \equiv 1$ e $\psi \in \mathcal{E}_\tau$, vale

$$\begin{aligned} \langle (T_f)_\tau, \psi \rangle &= \langle T_f, \chi\psi \rangle = \int dt f(t) \chi(t) \psi(t) \\ &= \sum_{n=-\infty}^{\infty} \int_{n\tau}^{(n+1)\tau} dt f(t) \chi(t) \psi(t) \\ &= \sum_{n=-\infty}^{\infty} \int_0^\tau dt f(t - n\tau) \chi(t - n\tau) \psi(t - n\tau) \\ &= \sum_{n=-\infty}^{\infty} \int_0^\tau dt f(t) \chi(t - n\tau) \psi(t) \\ &= \int_0^\tau dt f(t) \chi_\tau(t) \psi(t) = \int_0^\tau dt f(t) \psi(t) , \end{aligned}$$

ou seja,

$$\langle (T_f)_\tau, \psi \rangle = \int_0^\tau dt f(t) \psi(t) \quad \text{para } \psi \in \mathcal{E}_\tau . \quad (3.190)$$

Isso posto, definimos os *coeficientes de Fourier* de uma distribuição $T \in \mathcal{D}'$ periódica de período τ por

$$a_k = \frac{1}{\tau} \langle T_\tau(t), \exp(-ik\omega t) \rangle , \quad (3.191)$$

e podemos formular o teorema final, cuja demonstração não será apresentada aqui.

Teorema 3.10 *Seja $T \in \mathcal{D}'$ uma distribuição periódica de período τ . Então os coeficientes de Fourier a_k de T , conforme definidos pela equação (3.191), formam uma sequência $(a_k)_{k \in \mathbb{Z}}$ de números complexos polinomialmente limitada para $|k| \rightarrow \infty$ (ou seja, existem constantes $C > 0$ e $N, k_0 \in \mathbb{N}$ tais que vale a estimativa (3.179)), e a série de Fourier correspondente converge em \mathcal{D}' para T .*

3.11 Produto tensorial

Inicialmente, lembramos a definição do produto tensorial de uma função f sobre um aberto U de \mathbb{R}^m e uma função g sobre um aberto V de \mathbb{R}^n : é a função $f \otimes g$ sobre o aberto $U \times V$ de \mathbb{R}^{m+n} definida por

$$(f \otimes g)(x, y) = f(x)g(y). \quad (3.192)$$

Observamos que se f e g são localmente integráveis, $f \otimes g$ também o será, pois todo compacto C de $U \times V$ é contido no produto cartesiano $K \times L$ de um compacto K de U e um compacto L de V ,¹² e vale

$$\int_{K \times L} d^m x d^n y |(f \otimes g)(x, y)| = \int_K d^m x |f(x)| \int_L d^n y |g(y)|.$$

Note ainda que o teorema de Fubini garante que, para toda função teste $\varphi \in \mathcal{D}(U \times V)$ sobre $U \times V$, vale

$$\begin{aligned} & \int_{K \times L} d^m x d^n y f(x)g(y)\varphi(x, y) \\ &= \int_K d^m x f(x) \int_L d^n y g(y)\varphi(x, y) \\ &= \int_L d^n y g(y) \int_K d^m x f(x)\varphi(x, y). \end{aligned}$$

Em particular, para funções teste sobre $U \times V$ que são decomponíveis no sentido de poderem ser escritas na forma $\varphi \otimes \psi$ com $\varphi \in \mathcal{D}(U)$ e $\psi \in \mathcal{D}(V)$, temos

$$\begin{aligned} & \int_{K \times L} d^m x d^n y f(x)g(y)(\varphi \otimes \psi)(x, y) \\ &= \int_K d^m x f(x)\varphi(x) \int_L d^n y g(y)\psi(y), \end{aligned}$$

ou seja,

$$\langle T_{f \otimes g}, \varphi \otimes \psi \rangle = \langle T_f, \varphi \rangle \langle T_g, \psi \rangle.$$

Queremos usar esta fórmula como ponto de partida para estender o conceito de produto tensorial às distribuições. Para tanto, precisamos da seguinte

¹²Basta tomar K e L como a imagem de C pelas respectivas projeções de \mathbb{R}^{m+n} sobre \mathbb{R}^m e \mathbb{R}^n .

Proposição 3.41 *Sejam U um aberto de \mathbb{R}^m e V um aberto de \mathbb{R}^n . Então o espaço*

$$\left\{ \sum_k \varphi_k \otimes \psi_k \mid \varphi_k \in \mathcal{D}(U), \psi_k \in \mathcal{D}(V) \right\}$$

das combinações lineares de funções teste decomponíveis sobre $U \times V$ é denso em $\mathcal{D}(U \times V)$.

A demonstração desta proposição usa a convolução clássica de funções e portanto será adiada para a seção seguinte.

Teorema 3.11 *Sejam U um aberto de \mathbb{R}^m , V um aberto de \mathbb{R}^n , S uma distribuição sobre U e T uma distribuição sobre V . Então existe uma única distribuição $S \otimes T$ sobre $U \times V$, chamada o **produto tensorial** de S e T , tal que para toda função teste decomponível $\varphi \otimes \psi$ sobre $U \times V$, vale*

$$\langle S \otimes T, \varphi \otimes \psi \rangle = \langle S, \varphi \rangle \langle T, \psi \rangle. \quad (3.193)$$

A unicidade da distribuição $S \otimes T \in \mathcal{D}'(U \times V)$ que satisfaz a equação (3.193) é uma consequência imediata da proposição anterior, enquanto que a demonstração de sua existência requer um procedimento diferente, com a vantagem de revelar outras propriedades importantes da sua construção.

Proposição 3.42 *Sejam U um aberto de \mathbb{R}^m , V um aberto de \mathbb{R}^n , T uma distribuição sobre V e φ uma função teste sobre $U \times V$. Então a função*

$$\langle T(y), \varphi(x, y) \rangle$$

que a cada ponto x de U associa o valor de $T \in \mathcal{D}'(V)$ sobre a função teste $\varphi(x, \cdot) \in \mathcal{D}(V)$ é uma função teste sobre U ,

$$\langle T(y), \varphi(x, y) \rangle \in \mathcal{D}(U),$$

tal que para todo multi-índice $\alpha \in \mathbb{N}^m$, vale¹³

$$\partial_\alpha \langle T(y), \varphi(x, y) \rangle = \langle T(y), \partial_\alpha \varphi(x, y) \rangle. \quad (3.194)$$

Finalmente, a aplicação linear

$$\begin{aligned} \mathcal{D}(U \times V) &\longrightarrow \mathcal{D}(U) \\ \varphi &\longmapsto \langle T(y), \varphi(x, y) \rangle \end{aligned} \quad (3.195)$$

assim definida é contínua.

¹³A seguir, frequentemente identificaremos multi-índices α em \mathbb{N}^m e multi-índices β em \mathbb{N}^n com os multi-índices $(\alpha, 0)$ e $(0, \beta)$ em \mathbb{N}^{m+n} e escrevemos multi-índices gerais em \mathbb{N}^{m+n} na forma (α, β) com $\alpha \in \mathbb{N}^m$ e $\beta \in \mathbb{N}^n$; assim, por exemplo, $\partial_{(\alpha, \beta)} = \partial_\alpha \partial_\beta$.

DEMONSTRAÇÃO: Sejam K um compacto de U e L um compacto de V tais que $\text{supp } \varphi \subset K \times L$. Então para todo ponto x de U , $\varphi(x, \cdot)$ é uma função teste sobre V com suporte contido em L ; portanto, o número $\langle T(y), \varphi(x, y) \rangle$ é bem definido. Obviamente, se $x \notin K$, vale $\varphi(x, \cdot) \equiv 0$ e portanto $\langle T(y), \varphi(x, y) \rangle = 0$. Para mostrar que a função $\langle T(y), \varphi(x, y) \rangle$ é contínua, considere, para $x \in U$ e $h \in B_\epsilon(x)$ (com $\epsilon > 0$ tal que a bola aberta $B_\epsilon(x)$ de raio ϵ em torno de x esteja contida em U), a seguinte estimativa, obtida por aplicação do teorema do valor médio,

$$\begin{aligned} |\partial_\beta \varphi(x+h, y) - \partial_\beta \varphi(x, y)| &\leq \sum_{i=1}^m |h_i| \sup_{0 \leq s \leq 1} |\partial_i \partial_\beta \varphi(x+sh, y)| \\ &\leq \sum_{i=1}^m |h_i| \sup_{x' \in K, y' \in L} |\partial_i \partial_\beta \varphi(x', y')|, \end{aligned}$$

onde $\beta \in \mathbb{N}^n$ e $\partial_i = \partial/\partial x_i$, mostrando que

$$\lim_{h \rightarrow 0} \varphi(x+h, \cdot) = \varphi(x, \cdot) \quad \text{em } \mathcal{D}(V),$$

e portanto

$$\lim_{h \rightarrow 0} \langle T(y), \varphi(x+h, y) \rangle = \langle T(y), \varphi(x, y) \rangle.$$

Para mostrar que a função $\langle T(y), \varphi(x, y) \rangle$ é de classe C^∞ e, ao mesmo tempo, que vale a equação (3.194), suponha que já tenhamos demonstrado que ela é de classe C^r e que a equação (3.194) é válida para todo multi-índice $\alpha \in \mathbb{N}^m$ com $|\alpha| \leq r$. Considere então, para $x \in U$ e $h \in B_\epsilon(x)$ (com $\epsilon > 0$ tal que a bola aberta $B_\epsilon(x)$ de raio ϵ em torno de x esteja contida em U), a seguinte estimativa, obtida por aplicação do teorema de Taylor,

$$\begin{aligned} &\frac{1}{t} \left| \partial_\beta \partial_\alpha \varphi(x+th, y) - \partial_\beta \partial_\alpha \varphi(x, y) - t \sum_{i=1}^m h_i \partial_\beta \partial_i \partial_\alpha \varphi(x, y) \right| \\ &\leq \frac{1}{2} t \sum_{i,j=1}^m |h_i| |h_j| \sup_{0 \leq s \leq 1} |\partial_\beta \partial_i \partial_j \partial_\alpha \varphi(x+sh, y)| \\ &\leq \frac{1}{2} t \sum_{i,j=1}^m |h_i| |h_j| \sup_{x' \in K, y' \in L} |\partial_\beta \partial_i \partial_j \partial_\alpha \varphi(x', y')|, \end{aligned}$$

onde $0 < t < 1$, $\beta \in \mathbb{N}^n$ e $\partial_i = \partial/\partial x_i$, mostrando que

$$\begin{aligned} \lim_{t \rightarrow 0} \frac{1}{t} \left(\partial_\alpha \varphi(x+th, \cdot) - \partial_\alpha \varphi(x, \cdot) \right) &= \sum_{i=1}^m h_i \partial_i \partial_\alpha \varphi(x, \cdot) \\ &\text{em } \mathcal{D}(V), \end{aligned}$$

e portanto, devido à hipótese de indução,

$$\begin{aligned} & \lim_{t \rightarrow 0} \frac{1}{t} \left(\partial_\alpha \langle T(y), \varphi(x + th, y) \rangle - \partial_\alpha \langle T(y), \varphi(x, y) \rangle \right) \\ &= \sum_{i=1}^m h_i \langle T(y), \partial_i \partial_\alpha \varphi(x, y) \rangle, \end{aligned}$$

i.e., a função $\partial_\alpha \langle T(y), \varphi(x, y) \rangle = \langle T(y), \partial_\alpha \varphi(x, y) \rangle$ é diferenciável e as suas derivadas parciais são funções contínuas, dadas por

$$\partial_i \partial_\alpha \langle T(y), \varphi(x, y) \rangle = \langle T(y), \partial_i \partial_\alpha \varphi(x, y) \rangle,$$

o que fecha a indução. Finalmente, para mostrar a continuidade da aplicação linear (3.195), é suficiente mostrar que para qualquer compacto de $U \times V$ da forma $K \times L$ onde K é um compacto de U e L é um compacto de V , a aplicação linear

$$\begin{array}{ccc} \mathcal{D}(K \times L) & \longrightarrow & \mathcal{D}(K) \\ \varphi & \longmapsto & \langle T(y), \varphi(x, y) \rangle \end{array}$$

é contínua. (Aqui, usamos que para qualquer compacto C de $U \times V$, a topologia de $\mathcal{D}(C)$ coincide com a topologia induzida pela topologia de $\mathcal{D}(K \times L)$, onde K e L são as respectivas projeções de C .) Mas como T é uma distribuição, existem uma constante $C_L > 0$ e um inteiro não-negativo s tais que

$$|\langle T, \psi \rangle| \leq C_L \|\psi\|_{s,L} \quad \text{para } \psi \in \mathcal{D}(L),$$

e portanto temos, para $\varphi \in \mathcal{D}(K \times L)$,

$$\begin{aligned} \|\langle T(y), \varphi(x, y) \rangle\|_{r,K} &= \sum_{|\alpha| \leq r} \sup_{x \in K} |\partial_\alpha \langle T(y), \varphi(x, y) \rangle| \\ &= \sum_{|\alpha| \leq r} \sup_{x \in K} |\langle T(y), \partial_\alpha \varphi(x, y) \rangle| \\ &\leq C_L \sum_{|\alpha| \leq r} \sup_{x \in K} \|\partial_\alpha \varphi(x, \cdot)\|_{s,L} \\ &\leq C_L \|\varphi\|_{r+s, K \times L}. \end{aligned}$$

□

Obviamente, esta proposição demonstra o enunciado de existência do teorema acima: basta compor a aplicação linear contínua (3.195) com a aplicação linear contínua $S : \mathcal{D}(U) \rightarrow \mathbb{C}$. Ademais, é evidente que na construção desta proposição, podemos trocar S e T , e usando a propriedade de unicidade do teorema acima, chegamos ao seguinte teorema, como corolário:

Teorema 3.12 (Teorema de Fubini para distribuições): *Sejam U um aberto de \mathbb{R}^m , V um aberto de \mathbb{R}^n , S uma distribuição sobre U e T uma distribuição sobre V . Então para toda função teste φ sobre $U \times V$, vale*

$$\langle S(x), \langle T(y), \varphi(x, y) \rangle \rangle = \langle S \otimes T, \varphi \rangle = \langle T(y), \langle S(x), \varphi(x, y) \rangle \rangle. \quad (3.196)$$

Para concluir, recordamos algumas propriedades úteis do produto tensorial de distribuições.

Proposição 3.43 (Suporte do produto tensorial): *Sejam U um aberto de \mathbb{R}^m , V um aberto de \mathbb{R}^n , S uma distribuição sobre U e T uma distribuição sobre V . Então vale*

$$\text{supp}(S \otimes T) = \text{supp } S \times \text{supp } T. \quad (3.197)$$

DEMONSTRAÇÃO: Para mostrar a igualdade dos complementos, considere um ponto qualquer $(x, y) \in U \times V$.

“ \subset ”: Suponha que $(x, y) \notin \text{supp } S \times \text{supp } T$, ou seja, $x \notin \text{supp } S$ ou $y \notin \text{supp } T$, o que significa que existem uma vizinhança aberta $U_x \subset U$ de x tal que S se anula sobre U_x ou uma vizinhança aberta $V_y \subset V$ de y tal que T se anula sobre V_y : vamos mostrar que $S \otimes T$ se anula sobre $U_x \times V$, no primeiro caso, e sobre $U \times V_y$, no segundo caso, o que implica $(x, y) \notin \text{supp}(S \otimes T)$. Considere então uma função teste qualquer $\varphi \in \mathcal{D}(U \times V)$ e

– no primeiro caso, suponha que $\text{supp } \varphi \subset U_x \times V$: então para todo $y' \in V$, vale $\text{supp } \varphi(\cdot, y') \subset U_x$ e portanto $\langle S, \varphi(\cdot, y') \rangle = 0$, o que implica $\langle T(y), \langle S(x), \varphi(x, y) \rangle \rangle = 0$;

– no segundo caso, suponha que $\text{supp } \varphi \subset U \times V_y$: então para todo $x' \in U$, vale $\text{supp } \varphi(x', \cdot) \subset V_y$ e portanto $\langle T, \varphi(x', \cdot) \rangle = 0$, o que implica $\langle S(x), \langle T(y), \varphi(x, y) \rangle \rangle = 0$.

De qualquer modo, segue que $\langle S \otimes T, \varphi \rangle = 0$.

“ \supset ”: Suponha que $(x, y) \notin \text{supp}(S \otimes T)$, o que significa que existem vizinhanças abertas $U_x \subset U$ de x e $V_y \subset V$ de y tais que $S \otimes T$ se anula sobre $U_x \times V_y$. Se $x \in \text{supp } S$, existe uma função teste $\varphi \in \mathcal{D}(U_x)$ tal que $\langle S, \varphi \rangle \neq 0$ e então para qualquer função teste $\psi \in \mathcal{D}(V_y)$, vale $\langle S, \varphi \rangle \langle T, \psi \rangle = \langle S \otimes T, \varphi \otimes \psi \rangle = 0$ e portanto $\langle T, \psi \rangle = 0$, mostrando que $y \notin \text{supp } T$. De modo análogo, se $y \in \text{supp } T$, existe uma função teste $\psi \in \mathcal{D}(V_y)$ tal que $\langle T, \psi \rangle \neq 0$ e então para qualquer função teste $\varphi \in \mathcal{D}(U_x)$, vale $\langle S, \varphi \rangle \langle T, \psi \rangle = \langle S \otimes T, \varphi \otimes \psi \rangle = 0$ e portanto $\langle S, \varphi \rangle = 0$, mostrando que $x \notin \text{supp } S$. □

Proposição 3.44 (Associatividade do produto tensorial): *Sejam U um aberto de \mathbb{R}^m , V um aberto de \mathbb{R}^n , W um aberto de \mathbb{R}^p , R uma distribuição sobre U , S uma distribuição sobre V e T uma distribuição sobre W . Então vale*

$$(R \otimes S) \otimes T = R \otimes (S \otimes T). \quad (3.198)$$

DEMONSTRAÇÃO: De fato, ambos os lados são distribuições sobre $U \times V \times W$ tais que para funções teste decomponíveis $\varphi \otimes \psi \otimes \chi$ sobre $U \times V \times W$, com $\varphi \in \mathcal{D}(U)$, $\psi \in \mathcal{D}(V)$, $\chi \in \mathcal{D}(W)$, vale

$$\begin{aligned} \langle (R \otimes S) \otimes T, \varphi \otimes \psi \otimes \chi \rangle &= (\langle R, \varphi \rangle \langle S, \psi \rangle) \langle T, \chi \rangle \\ &= \langle R, \varphi \rangle (\langle S, \psi \rangle \langle T, \chi \rangle) = \langle R \otimes (S \otimes T), \varphi \otimes \psi \otimes \chi \rangle. \end{aligned}$$

□

Proposição 3.45 *Sejam U um aberto de \mathbb{R}^m , V um aberto de \mathbb{R}^n , S uma distribuição sobre U , T uma distribuição sobre V , $\alpha \in \mathbb{N}^m$, $\beta \in \mathbb{N}^n$, f uma função lisa sobre U e g uma função lisa sobre V . Então vale*

$$\partial_{(\alpha, \beta)}(S \otimes T) = \partial_\alpha S \otimes \partial_\beta T, \quad (3.199)$$

e

$$(f \otimes g)(S \otimes T) = fS \otimes gT. \quad (3.200)$$

DEMONSTRAÇÃO: Mais uma vez, em cada uma das duas equações, ambos os lados são distribuições sobre $U \times V$ tais que para funções teste decomponíveis $\varphi \otimes \psi$ sobre $U \times V$, com $\varphi \in \mathcal{D}(U)$, $\psi \in \mathcal{D}(V)$, vale

$$\begin{aligned} \langle \partial_{(\alpha, \beta)}(S \otimes T), \varphi \otimes \psi \rangle &= (-1)^{|\alpha, \beta|} \langle S \otimes T, \partial_{(\alpha, \beta)}(\varphi \otimes \psi) \rangle \\ &= (-1)^{|\alpha| + |\beta|} \langle S \otimes T, \partial_\alpha \varphi \otimes \partial_\beta \psi \rangle \\ &= (-1)^{|\alpha|} (-1)^{|\beta|} \langle S, \partial_\alpha \varphi \rangle \langle T, \partial_\beta \psi \rangle \\ &= \langle \partial_\alpha S, \varphi \rangle \langle \partial_\beta T, \psi \rangle = \langle \partial_\alpha S \otimes \partial_\beta T, \varphi \otimes \psi \rangle, \end{aligned}$$

e

$$\begin{aligned} \langle (f \otimes g)(S \otimes T), \varphi \otimes \psi \rangle &= \langle S \otimes T, (f \otimes g)(\varphi \otimes \psi) \rangle \\ &= \langle S \otimes T, f\varphi \otimes g\psi \rangle \\ &= \langle S, f\varphi \rangle \langle T, g\psi \rangle \\ &= \langle fS, \varphi \rangle \langle gT, \psi \rangle = \langle fS \otimes gT, \varphi \otimes \psi \rangle. \end{aligned}$$

□

Proposição 3.46 *Para quaisquer dois abertos U de \mathbb{R}^m e V de \mathbb{R}^n , o operador bilinear de produto tensorial*

$$\begin{aligned} \mathcal{D}'(U) \times \mathcal{D}'(V) &\longrightarrow \mathcal{D}'(U \times V) \\ (S, T) &\longmapsto S \otimes T \end{aligned} \quad (3.201)$$

é separadamente contínuo, o que significa que

- Para toda distribuição $S \in \mathcal{D}'(U)$, a aplicação linear

$$\begin{aligned} \mathcal{D}'(V) &\longrightarrow \mathcal{D}'(U \times V) \\ T &\longmapsto S \otimes T \end{aligned} \quad (3.202)$$

é contínua. Em particular, para qualquer sequência $(T_k)_{k \in \mathbb{N}}$ de distribuições T_k sobre V , vale

$$T_k \rightarrow T \text{ em } \mathcal{D}'(V) \quad \implies \quad S \otimes T_k \rightarrow S \otimes T \text{ em } \mathcal{D}'(U \times V) .$$

- Para toda distribuição $T \in \mathcal{D}'(V)$, a aplicação linear

$$\begin{aligned} \mathcal{D}'(U) &\longrightarrow \mathcal{D}'(U \times V) \\ S &\longmapsto S \otimes T \end{aligned} \quad (3.203)$$

é contínua. Em particular, para qualquer sequência $(S_k)_{k \in \mathbb{N}}$ de distribuições S_k sobre U , vale

$$S_k \rightarrow S \text{ em } \mathcal{D}'(U) \quad \implies \quad S_k \otimes T \rightarrow S \otimes T \text{ em } \mathcal{D}'(U \times V) .$$

DEMONSTRAÇÃO: As afirmações são consequências imediatas da Proposição 3.32, tendo em vista que os operadores lineares (3.202) e (3.203) são os duais ou transpostos dos operadores lineares

$$\begin{aligned} \mathcal{D}(U \times V) &\longrightarrow \mathcal{D}(V) \\ \varphi &\longmapsto \langle S(x), \varphi(x, y) \rangle \end{aligned}$$

e

$$\begin{aligned} \mathcal{D}(U \times V) &\longrightarrow \mathcal{D}(U) \\ \varphi &\longmapsto \langle T(y), \varphi(x, y) \rangle \end{aligned}$$

respectivamente, e estes são contínuos (veja as equações (3.195) e (3.196)). □

Finalmente, notamos que podemos iterar a construção do produto tensorial e assim, fazendo uso da sua associatividade, chegar a uma definição do produto tensorial de um número arbitrário (finito) de distribuições. A título de exemplo, notamos que a distribuição δ de Dirac sobre \mathbb{R}^n é o produto tensorial de n cópias da distribuição δ de Dirac sobre \mathbb{R} .

3.12 Convolução

Inicialmente, lembramos a definição da convolução clássica para funções integráveis:

Definição 3.31 Para quaisquer duas funções integráveis f e g sobre \mathbb{R}^n , a **convolução** de f e g é a função integrável $f * g$ sobre \mathbb{R}^n definida por

$$(f * g)(x) = \int d^n y f(y) g(x - y) . \quad (3.204)$$

Para justificar esta definição, observamos que se as funções f e g são integráveis sobre \mathbb{R}^n , então a função $f(y) g(x - y)$ é integrável sobre \mathbb{R}^{2n} e portanto o teorema de Fubini garante que a integral na equação (3.204) existe para quase todo $x \in \mathbb{R}^n$ e define $f * g$ como uma função integrável sobre \mathbb{R}^n tal que

$$\int d^n x |(f * g)(x)| \leq \int d^n x \int d^n y |f(y)| |g(x - y)| ,$$

de modo que obtemos a desigualdade

$$\|f * g\|_1 \leq \|f\|_1 \|g\|_1 . \quad (3.205)$$

Assim, identificando funções que coincidem exceto sobre conjuntos de medida zero, como sempre, a convolução clássica pode ser vista como um produto sobre o espaço de Banach $L^1(\mathbb{R}^n)$, isto é, uma aplicação bilinear contínua

$$\begin{aligned} L^1(\mathbb{R}^n) \times L^1(\mathbb{R}^n) &\longrightarrow L^1(\mathbb{R}^n) \\ ([f], [g]) &\longmapsto [f * g] \end{aligned} , \quad (3.206)$$

que satisfaz as propriedades de comutatividade e associatividade:

$$[g * f] = [f * g] , \quad (3.207)$$

$$[(f * g) * h] = [f * (g * h)] . \quad (3.208)$$

Isso significa que o espaço de Banach $L^1(\mathbb{R}^n)$, munido do produto de convolução, é uma álgebra de Banach comutativa e associativa, porém sem unidade.

As equações (3.207) e (3.208) são provadas por uma mudança de variáveis de integração e, no segundo caso, uma aplicação do teorema de Fubini. Explícitamente, temos

$$\begin{aligned} (g * f)(x) &= \int d^n y g(y) f(x - y) \\ &= \int_{z=x-y} d^n z f(z) g(x - z) = (f * g)(x) , \end{aligned}$$

e

$$\begin{aligned} ((f * g) * h)(x) &= \int d^n y (f * g)(y) h(x - y) \\ &= \int d^n y \int d^n z f(z) g(y - z) h(x - y) \\ &= \int_{\substack{u=z \\ v=y-z}} d^n u \int d^n v f(u) g(v) h(x - u - v) \\ &= \int d^n u f(u) (g * h)(x - u) = (f * (g * h))(x) . \end{aligned}$$

FALTA: Na verdade, se entendermos que as funções envolvidas assumam seus valores em \mathbb{C} , $L^1(\mathbb{R}^n)$ se torna uma Banach $*$ -álgebra em relação a qualquer uma das duas seguintes conjugações: a “conjugação pontual” (veja notações e convenções), mas também a “conjugação CPT” definida por

$$f^*(x) = f(-x)^*$$

A conveniência de complementar a conjugação pontual pela reflexão na origem do espaço \mathbb{R}^n tornará-se mais evidente na próxima seção, no contexto do estudo da transformação de Fourier. Tem a ver com o fato que esta é a conjugação em relação à qual as funções oscilatórias $\exp(i\xi \cdot x)$ (para qualquer $\xi \in \mathbb{R}^n$) se tornam reais.

Notamos ainda que, aplicando mais uma vez uma mudança de variáveis em conjunto com o teorema de Fubini, obtemos, para toda função teste φ sobre \mathbb{R}^n ,

$$\begin{aligned} \int d^n x (f * g)(x) \varphi(x) &= \int d^n x \int d^n y f(y) g(x - y) \varphi(x) \\ &= \int d^n y \int d^n z f(y) g(z) \varphi(y + z), \end{aligned}$$

ou seja,

$$\langle T_{f * g}, \varphi \rangle = \langle T_f \otimes T_g, \varphi^\Delta \rangle, \quad (3.209)$$

onde φ^Δ é a função sobre \mathbb{R}^{2n} definida por

$$\varphi^\Delta(x, y) = \varphi(x + y). \quad (3.210)$$

Essa fórmula sugere uma definição do produto de convolução $S * T$ entre duas distribuições S e T em termos do seu produto tensorial $S \otimes T$. Para tanto, a dificuldade a ser superada decorre do fato de que φ^Δ não é uma função teste sobre \mathbb{R}^{2n} : ela pertence ao espaço $\mathcal{E}(\mathbb{R}^{2n})$, pois é óbvio que se a função φ for lisa, a função φ^Δ também será, mas ela não pertence ao espaço $\mathcal{D}(\mathbb{R}^{2n})$, pois φ^Δ é constante ao longo de cada um dos subespaços afins “paralelos à antidiagonal” de \mathbb{R}^{2n} , dados por $\{(x, y) \in \mathbb{R}^{2n} \mid x + y = a\}$, onde a é algum vetor fixo em \mathbb{R}^n , e portanto nunca tem suporte compacto. No entanto, podemos tentar interpretar a fórmula (3.209) no sentido da Proposição 3.26, o que é possível desde que a interseção

$$(\text{supp } S \times \text{supp } T) \cap \text{supp } \varphi^\Delta = \text{supp } (S \otimes T) \cap \text{supp } \varphi^\Delta$$

seja compacta. Esta observação nos leva diretamente à seguinte

Definição 3.32 Para todo subconjunto fechado C de \mathbb{R}^n , definimos a **faixa** $F_C^{(2)}$ em \mathbb{R}^{2n} gerada por C como sendo o subconjunto fechado de \mathbb{R}^{2n} definido por

$$F_C^{(2)} = \{(x, y) \in \mathbb{R}^{2n} \mid x + y \in C\}. \quad (3.211)$$

Dizemos que dois subconjuntos fechados A e B de \mathbb{R}^n satisfazem a **condição da faixa** se para todo subconjunto compacto K de \mathbb{R}^n , a interseção do produto cartesiano $A \times B$ com a faixa $F_K^{(2)}$ gerada por K é um subconjunto compacto de \mathbb{R}^{2n} .¹⁴

$$K \text{ compacto} \implies (A \times B) \cap F_K^{(2)} \text{ compacto.} \quad (3.212)$$

Finalmente, dizemos que duas distribuições S e T sobre \mathbb{R}^n satisfazem a **condição da faixa** se os seus suportes satisfazem a condição da faixa:

$$K \text{ compacto} \implies (\text{supp } S \times \text{supp } T) \cap F_K^{(2)} \text{ compacto.} \quad (3.213)$$

Notamos que a condição da faixa é simétrica no sentido de ser invariante sob a troca de A e B , ou de S e T . Também notamos que a definição da faixa $F_C^{(2)}$ gerada por um subconjunto fechado C de \mathbb{R}^n se motiva pela observação de que se C é o suporte da função φ , então $F_C^{(2)}$ é o suporte da função φ^Δ . Além disso, temos a seguinte

Proposição 3.47 *A aplicação linear*

$$\begin{aligned} \mathcal{D}(\mathbb{R}^n) &\longrightarrow \mathcal{E}(\mathbb{R}^{2n}) \\ \varphi &\longmapsto \varphi^\Delta \end{aligned} \quad (3.214)$$

definida pela equação (3.210) é contínua.

DEMONSTRAÇÃO: Para quaisquer dois compactos K e L de \mathbb{R}^n e quaisquer dois multi-índices $\alpha, \beta \in \mathbb{N}^n$, temos $\partial_{(\alpha, \beta)} \varphi^\Delta = (\partial_{\alpha + \beta} \varphi)^\Delta$ e portanto

$$\sup_{(x, y) \in K \times L} |\partial_{(\alpha, \beta)} \varphi^\Delta(x, y)| \leq \sup_{z \in K + L} |\partial_{\alpha + \beta} \varphi(z)|.$$

□

Agora, dadas duas distribuições S e T sobre \mathbb{R}^n que satisfazem a condição da faixa, podemos considerar o subconjunto fechado $F = \text{supp } S \times \text{supp } T = \text{supp } (S \otimes T)$ de \mathbb{R}^{2n} e, usando a notação introduzida na equação (3.138), concluir que a aplicação linear contínua (3.214) é de fato uma aplicação linear contínua

$$\begin{aligned} \mathcal{D}(\mathbb{R}^n) &\longrightarrow \mathcal{E}_F(\mathbb{R}^{2n}) \\ \varphi &\longmapsto \varphi^\Delta \end{aligned}, \quad (3.215)$$

cujas composição com a distribuição $S \otimes T$ sobre \mathbb{R}^{2n} proporciona uma distribuição sobre \mathbb{R}^n :

¹⁴Essa condição é não-trivial pois, mesmo quando K for compacto, a faixa $F_K^{(2)}$ gerada por K , por si só, é apenas fechada, mas nunca é compacta.

Definição 3.33 Para duas distribuições S e T sobre \mathbb{R}^n que satisfazem a condição da faixa, define-se a sua **convolução** como sendo a distribuição $S * T$ sobre \mathbb{R}^n definida por

$$\langle S * T, \varphi \rangle = \langle S \otimes T, \varphi^\Delta \rangle \quad \text{para } \varphi \in \mathcal{D}(\mathbb{R}^n), \quad (3.216)$$

ou mais explicitamente,

$$\langle S * T, \varphi \rangle = \langle (S \otimes T)(x, y), \varphi(x + y) \rangle \quad \text{para } \varphi \in \mathcal{D}(\mathbb{R}^n), \quad (3.217)$$

ou mais explicitamente ainda,

$$\langle S * T, \varphi \rangle = \langle (S \otimes T)(x, y), \chi(x, y) \varphi(x + y) \rangle \quad \text{para } \varphi \in \mathcal{D}(K), \quad (3.218)$$

onde K é qualquer compacto de \mathbb{R}^n e $\chi \in \mathcal{D}(\mathbb{R}^{2n})$ é uma função teste tal que $\chi \equiv 1$ sobre uma vizinhança aberta do compacto $(\text{supp } S \times \text{supp } T) \cap F_K^{(2)}$ de \mathbb{R}^{2n} .

Claramente, o produto de convolução ainda pode ser definido em outras circunstâncias. Por exemplo, na construção da convolução de funções integráveis apresentada no início desta seção, não foi imposta nenhuma restrição sobre os suportes dos dois fatores. Assim, coloca-se naturalmente a pergunta se seria possível introduzir uma noção de “distribuição integrável” e definir a convolução de distribuições integráveis. Isso é de fato possível, mas a construção é tecnicamente exigente.

O passo inicial é simples, pois existe um candidato óbvio para um espaço de funções teste cujo dual seria o espaço das distribuições integráveis – semelhante ao espaço de funções teste cujo dual seria o espaço das medidas integráveis: o segundo é o espaço $B^0(\mathbb{R}^n)$ das funções contínuas e limitadas sobre \mathbb{R}^n , munido da topologia da convergência uniforme, enquanto que o primeiro é o espaço $B^\infty(\mathbb{R}^n)$ das funções lisas e limitadas sobre \mathbb{R}^n (com todas as derivadas parciais também limitadas), munido da topologia da convergência uniforme de todas as derivadas. (Veja o Exemplo 3.5.) A integral de uma medida ou distribuição integrável seria então o valor do funcional sobre a função constante 1, em consonância com o Exemplo 3.8. Também é fácil definir a convolução neste âmbito, uma vez que estes espaços são invariantes sob a aplicação que leva uma função φ sobre \mathbb{R}^n na função φ^Δ sobre \mathbb{R}^{2n} .

O problema principal reside na dificuldade de interpretar funcionais lineares contínuos sobre estes espaços como medidas ou distribuições que são, respectivamente, funcionais lineares contínuos sobre os espaços $C_c^0(\mathbb{R}^n)$ das funções contínuas e $C_c^\infty(\mathbb{R}^n)$ ($= \mathcal{D}$) das funções lisas de suporte compacto. Conforme o procedimento padrão, definem-se medidas e distribuições como sendo integráveis se são, respectivamente, contínuas em relação à topologia sobre $C_c^0(\mathbb{R}^n)$ induzida por $B^0(\mathbb{R}^n)$ e contínuas em relação à topologia sobre $C_c^\infty(\mathbb{R}^n)$ induzida por $B^\infty(\mathbb{R}^n)$. Infelizmente, $C_c^0(\mathbb{R}^n)$ não é denso em $B^0(\mathbb{R}^n)$ e $C_c^\infty(\mathbb{R}^n)$ não é denso em $B^\infty(\mathbb{R}^n)$. Assim, medidas e distribuições integráveis permitem uma extensão canônica por continuidade somente aos respectivos fechados, sendo que o fecho de $C_c^0(\mathbb{R}^n)$ em $B^0(\mathbb{R}^n)$ é o espaço $C_0^0(\mathbb{R}^n)$ das funções contínuas sobre \mathbb{R}^n que “vão a zero no infinito” e o fecho de $C_c^\infty(\mathbb{R}^n)$ em $B^\infty(\mathbb{R}^n)$ é o espaço $C_0^\infty(\mathbb{R}^n)$ das funções lisas sobre \mathbb{R}^n que “vão a zero no infinito” (com todas as derivadas parciais também “indo a

zero no infinito”), e a função constante 1 não pertence a nenhum destes fechos. Uma segunda extensão até $B^0(\mathbb{R}^n)$ ou $B^\infty(\mathbb{R}^n)$ sempre é possível, devido ao teorema de Hahn-Banach, mas perde-se a unicidade: têm um espaço de dimensão infinita de tais extensões! Encontrar uma entre elas que seja privilegiada, sob algum ponto de vista, constitui o maior desafio desta construção.

No caso das medidas, o conceito central que permite tal extensão privilegiada é o de um funcional positivo. ... Mas para distribuições, isso não funciona. Aqui, tem que representar distribuições como derivadas de medidas. ...

Agora um conceito que é útil para discutir convoluções iteradas e, em particular, formular as condições sob as quais a convolução define um produto associativo.

Definição 3.34 *Se S e T são distribuições sobre \mathbb{R}^n , dizemos que o par (S, T) satisfaz a **condição da faixa** se para todo compacto K de \mathbb{R}^n , a interseção do produto cartesiano $\text{supp } S \times \text{supp } T$ dos suportes de S e T com a faixa $F_K^{(2)}$ gerada por K é um subconjunto compacto de \mathbb{R}^{2n} : De modo análogo, se T_1, \dots, T_k são distribuições sobre \mathbb{R}^n , dizemos que a k -upla (T_1, \dots, T_k) satisfaz a **condição da faixa** se para todo compacto K de \mathbb{R}^n , a interseção do produto cartesiano $\text{supp } T_1 \times \dots \times \text{supp } T_k$ dos suportes dos T_i com a faixa $F_K^{(k)}$ gerada por K é um subconjunto compacto de \mathbb{R}^{kn} :*

$$K \text{ compacto} \implies (\text{supp } T_1 \times \dots \times \text{supp } T_k) \cap F_K^{(k)} \text{ compacto} . \quad (3.219)$$

A SER CONTINUADO

Assntos a serem tratados:

- Propriedades elementares da convolução:
- Regularização de distribuições por convolução com funções teste.
- Demonstração da Proposição 3.41.

3.13 Transformação de Fourier

Inicialmente, recordamos a definição da transformação de Fourier clássica para funções integráveis sobre \mathbb{R}^n :

Definição 3.35 Para qualquer função integrável f sobre \mathbb{R}^n , a **transformada de Fourier** de f é a função contínua e limitada \hat{f} sobre \mathbb{R}^n definida por

$$\hat{f}(\xi) = \int d^n x f(x) \exp(-i\xi \cdot x). \quad (3.220)$$

Frequentemente, far-se-á necessário indicar explicitamente os argumentos de cada uma das funções envolvidas; neste caso, será conveniente escrever

$$\hat{f}(\xi) = \widehat{f(x)}. \quad (3.221)$$

Para provar as duas afirmações já incorporadas nesta definição, observamos que se a função f é integrável sobre \mathbb{R}^n , então o teorema da convergência dominada garante que para qualquer sequência $(\xi_k)_{k \in \mathbb{N}}$ de pontos $\xi_k \in \mathbb{R}^n$ que converge para um ponto $\xi \in \mathbb{R}^n$, a sequência de funções $f(x) \exp(-i\xi_k \cdot x)$ converge em $\mathcal{L}^1(\mathbb{R}^n)$ para a função $f(x) \exp(-i\xi \cdot x)$, mostrando que \hat{f} é contínua. Obviamente, \hat{f} também é limitada, pois temos $|\hat{f}(\xi)| \leq \|f\|_1$ para todo $\xi \in \mathbb{R}^n$. Na verdade, vale uma afirmação um pouco mais forte:

Teorema 3.13 (Lema de Riemann-Lebesgue): Para qualquer função integrável f sobre \mathbb{R}^n , sua transformada de Fourier é uma função contínua \hat{f} sobre \mathbb{R}^n que se anula no infinito.

DEMONSTRAÇÃO: Isso pode ser provado de várias maneiras, mas talvez a mais simples é usar o fato de que o espaço \mathcal{S} (e até o espaço \mathcal{D}) é um subespaço denso de $\mathcal{L}^1(\mathbb{R}^n)$ (na norma L^1), assim como, aliás, de $C_0(\mathbb{R}^n)$ (na norma do supremo) que, como segue da Proposição 3.50 logo abaixo, é invariante sob a transformação de Fourier. \square

Note, no entanto, que o espaço $\mathcal{L}^1(\mathbb{R}^n)$ não é invariante sob a transformação de Fourier, pois $f \in \mathcal{L}^1(\mathbb{R}^n)$ não implica $\hat{f} \in \mathcal{L}^1(\mathbb{R}^n)$.

Tendo em vista que a função \hat{f} permanece inalterada quando a função f é substituída por qualquer outra que coincide com f exceto sobre um conjunto de medida zero, concluímos que a transformação de Fourier clássica pode ser vista como uma aplicação linear contínua

$$\begin{aligned} \mathcal{F}: L^1(\mathbb{R}^n) &\longrightarrow C_0(\mathbb{R}^n) \\ [f] &\longmapsto \hat{f} \end{aligned} \quad (3.222)$$

do espaço de Banach $L^1(\mathbb{R}^n)$ das (classes de equivalência de) funções integráveis, com a norma L^1 (veja o Exemplo 3.3 e a Observação 3.1) para o espaço de Banach

$C_0(\mathbb{R}^n)$ das funções contínuas que se anulam no infinito, com a norma do supremo (veja o Exemplo 3.2 e a Observação 3.1), de norma ≤ 1 , pois satisfaz a desigualdade

$$\|\hat{f}\| \leq \|f\|_1. \quad (3.223)$$

Mas na verdade, também é um homomorfismo de álgebras:

Proposição 3.48 *Para quaisquer duas funções integráveis f e g sobre \mathbb{R}^n , vale*

$$\widehat{f * g} = \hat{f} \hat{g}. \quad (3.224)$$

DEMONSTRAÇÃO: É um cálculo direto, usando o teorema de Fubini e uma simples mudança de variáveis:

$$\begin{aligned} \widehat{f * g}(\xi) &= \int d^n x (f * g)(x) \exp(-i\xi \cdot x) \\ &= \int d^n x \int d^n y f(y) \exp(-i\xi \cdot y) g(x - y) \exp(-i\xi \cdot (x - y)) \\ &= \int d^n y f(y) \exp(-i\xi \cdot y) \int d^n z g(z) \exp(-i\xi \cdot z) \\ &= \hat{f}(\xi) \hat{g}(\xi). \end{aligned}$$

□

Para encontrarmos um espaço de funções (e posteriormente, de distribuições) que seja invariante sob a transformação de Fourier, notamos a seguinte

Proposição 3.49 *Seja f uma função integrável sobre \mathbb{R}^n e \hat{f} sua transformada de Fourier.*

(a) *Suponha que os produtos de f com monômios x^α de ordem $\leq r$ são integráveis. Então \hat{f} é de classe C^r , com derivadas parciais de ordem $\leq r$ dadas por*

$$\partial_\alpha \hat{f}(\xi) = (-i)^{|\alpha|} \widehat{x^\alpha f(x)}. \quad (3.225)$$

(b) *Suponha que f é de classe C^r , com derivadas parciais de ordem $\leq r$ que são integráveis e se anulam no infinito. Então os produtos de \hat{f} com monômios ξ^α de ordem $\leq r$ são dadas por*

$$\xi^\alpha \hat{f}(\xi) = (-i)^{|\alpha|} \widehat{\partial_\alpha f(x)}. \quad (3.226)$$

DEMONSTRAÇÃO: Obviamente, a validade destas afirmações para qualquer valor de r segue da sua validade para $r = 1$, por indução. E quando $r = 1$, a primeira é uma consequência imediata do teorema sobre diferenciação debaixo do sinal de integração, enquanto que a segunda decorre de uma integração por partes, na qual a hipótese de que f deve se anular no infinito serve para eliminar o termo de bordo.

□

Disto, segue imediatamente a seguinte¹⁵

Proposição 3.50 *No espaço de Schwartz, a transformação de Fourier define uma aplicação linear contínua*

$$\begin{aligned} \mathcal{F} : \mathcal{S} &\longrightarrow \mathcal{S} \\ \varphi &\longmapsto \hat{\varphi} \end{aligned} \quad (3.227)$$

DEMONSTRAÇÃO: O fato de que \mathcal{F} mapeia \mathcal{S} em \mathcal{S} segue imediatamente da proposição anterior, e a linearidade de \mathcal{F} é óbvia. Resta provar que \mathcal{F} é contínua. Para tanto, lembramos que a topologia de \mathcal{S} é definida pela família de seminormas dada pela equação (3.73) e usamos as equações (3.225), (3.226) e (1.57), em conjunto com a identidade

$$\partial_{\beta} x^{\alpha} = \begin{cases} \frac{\alpha!}{\beta!} x^{\alpha-\beta} & \text{se } \beta \leq \alpha \\ 0 & \text{caso contrário} \end{cases}$$

e a estimativa

$$\int \frac{d^n x}{(1+|x|^2)^n} \leq \prod_{i=1}^n \int_{-\infty}^{\infty} \frac{dx_i}{1+x_i^2} = \pi^n$$

para calcular, para quaisquer dois multi-índices $\alpha, \beta \in \mathbb{N}^n$,

$$\begin{aligned} \|\hat{\varphi}\|_{\alpha, \beta} &= \sup_{\xi \in \mathbb{R}^n} |\xi^{\beta} \partial_{\alpha} \hat{\varphi}(\xi)| = \sup_{\xi \in \mathbb{R}^n} |\xi^{\beta} \mathcal{F}(x^{\alpha} \varphi(x))(\xi)| \\ &= \sup_{\xi \in \mathbb{R}^n} |\mathcal{F}(\partial_{\beta}(x^{\alpha} \varphi(x)))(\xi)| \leq \|\partial_{\beta}(x^{\alpha} \varphi(x))\|_1 \\ &\leq \sum_{\gamma \leq \alpha, \gamma \leq \beta} \binom{\beta}{\gamma} \frac{\alpha!}{\gamma!} \|x^{\alpha-\gamma} \partial_{\beta-\gamma} \varphi(x)\|_1 \\ &\leq \sum_{\gamma \leq \alpha, \gamma \leq \beta} \pi^n \binom{\beta}{\gamma} \frac{\alpha!}{\gamma!} \sup_{x \in \mathbb{R}^n} |(1+|x|^2)^n x^{\alpha-\gamma} \partial_{\beta-\gamma} \varphi(x)|. \end{aligned}$$

□

Antes de prosseguir, notamos, para uso posterior, a fórmula para a integral Gaussiana, válida para qualquer $\lambda > 0$,

$$\int d^n x \exp(-\lambda x^2) = \left(\frac{\pi}{\lambda}\right)^{n/2}, \quad (3.228)$$

¹⁵A seguir, simplificamos a notação utilizada no Exemplo 3.1, usando o símbolo $|\cdot|$, ao invés do símbolo $\|\cdot\|_2$, para denotar a norma euclidiana em \mathbb{R}^n assim como em \mathbb{C}^n , dada pela equação (3.13).

que se demonstra pelo seguinte cálculo elementar,

$$\begin{aligned} \int d^n x \exp(-\lambda x^2) &= \left(\int_{-\infty}^{\infty} dx \exp(-\lambda x^2) \right)^n \\ &= \left(\int d^2 x \exp(-\lambda x^2) \right)^{n/2} = \left(\int_0^{\infty} dr \int_0^{2\pi} d\theta r \exp(-\lambda r^2) \right)^{n/2} \\ &= \left(\frac{\pi}{\lambda} \int_0^{\infty} du \exp(-u) \right)^{n/2} = \left(\frac{\pi}{\lambda} \right)^{n/2} \end{aligned}$$

assim como o seguinte lema, que formalmente pode ser vista como uma extensão da invariância da medida de Lebesgue sob translações, no sentido de incluir translações complexas:

Lema 3.3 *Seja f uma função analítica sobre \mathbb{C}^n tal que para todo vetor fixo $y \in \mathbb{R}^n$, a restrição de f ao subespaço real $\{z = x + iy \mid x \in \mathbb{R}^n\}$ de \mathbb{C}^n é integrável e se anula no infinito. Então vale*

$$\int d^n x f(x + iy) = \int d^n x f(x) . \quad (3.229)$$

DEMONSTRAÇÃO : Primeiro, segundo o teorema de Fubini, as integrais na equação (3.229) podem ser escritas como integrais iteradas, de modo que a validade do lema para $n = 1$ implica a sua validade para qualquer n . Agora, no plano complexo, essas integrais podem ser escritas como limites

$$\begin{aligned} \int_{-\infty}^{\infty} dx f(x + iy) &= \lim_{a \rightarrow \infty} \int_{-a}^a dx f(x + iy) . \\ \int_{-\infty}^{\infty} dx f(x) &= \lim_{a \rightarrow \infty} \int_{-a}^a dx f(x) . \end{aligned}$$

Mas como f é analítica, o teorema integral de Cauchy implica que

$$\begin{aligned} \int_{-a}^a dx f(x + iy) - \int_{-a}^a dx f(x) \\ = \int_0^y du f(a + iu) - \int_0^y du f(-a + iu) , \end{aligned}$$

e esta expressão se anula no limite $a \rightarrow \infty$, pela hipótese de que a restrição de f a cada subespaço real se anula no infinito. \square

Retornando à análise da transformação de Fourier, talvez o aspecto mais interessante e importante de sua restrição ao espaço \mathcal{S} seja o fato de que nele, ela é um automorfismo e quase uma involução.

Teorema 3.14 (*Teorema de inversão de Fourier*): No espaço de Schwartz, a transformação de Fourier define um isomorfismo linear contínuo

$$\begin{aligned} \mathcal{F} : \mathcal{S} &\longrightarrow \mathcal{S} \\ \varphi &\longmapsto \mathcal{F}\varphi \end{aligned} \quad (3.230)$$

definido pela fórmula

$$(\mathcal{F}\varphi)(\xi) = \int d^n x \varphi(x) \exp(-i\xi \cdot x), \quad (3.231)$$

cujos inverso

$$\begin{aligned} \mathcal{F}^{-1} : \mathcal{S} &\longrightarrow \mathcal{S} \\ \psi &\longmapsto \mathcal{F}^{-1}\psi \end{aligned} \quad (3.232)$$

é dado pela fórmula

$$(\mathcal{F}^{-1}\psi)(x) = \frac{1}{(2\pi)^n} \int d^n \xi \psi(\xi) \exp(i\xi \cdot x). \quad (3.233)$$

DEMONSTRAÇÃO: Temporariamente, definimos uma transformação de Fourier modificada

$$\begin{aligned} \tilde{\mathcal{F}} : \mathcal{S} &\longrightarrow \mathcal{S} \\ \psi &\longmapsto \tilde{\mathcal{F}}\psi \end{aligned}$$

pela fórmula

$$(\tilde{\mathcal{F}}\psi)(x) = \frac{1}{(2\pi)^n} \int d^n \xi \psi(\xi) \exp(i\xi \cdot x).$$

Então é claro que todas as propriedades da transformação de Fourier já demonstradas valem igualmente para esta transformação modificada. Assim, resta apenas mostrar que a transformação composta $\tilde{\mathcal{F}}\mathcal{F}$ é a identidade em \mathcal{S} , ou seja, que para $\varphi \in \mathcal{S}$ e todo $x \in \mathbb{R}^n$, vale

$$\varphi(x) = \frac{1}{(2\pi)^n} \int d^n \xi \int d^n y \varphi(y) \exp(i\xi \cdot (x - y)). \quad (3.234)$$

(De modo completamente análogo, poderemos provar que a transformação composta $\mathcal{F}\tilde{\mathcal{F}}$ também é a identidade em \mathcal{S} .) A dificuldade aqui provém do fato de que, nesta fórmula, a integração só pode ser executada na ordem indicada (primeiro sobre a variável y e depois sobre a variável ξ , pois o integrando $\varphi(y) \exp(i\xi \cdot (x - y))$, como função de $(y, \xi) \in \mathbb{R}^{2n}$, não é uma função integrável e assim o teorema de Fubini não se aplica diretamente. Contudo, este problema pode ser contornado pelo seguinte truque: para todo $\epsilon > 0$, considere a integral

$$\varphi_\epsilon(x) = \frac{1}{(2\pi)^n} \int d^n \xi \int d^n y \varphi(y) \exp(i\xi \cdot (x - y)) \exp(-\epsilon \xi^2). \quad (3.235)$$

Aqui, o integrando $\varphi(y) \exp(i\xi \cdot (x - y)) \exp(-\epsilon \xi^2)$, como função de $(y, \xi) \in \mathbb{R}^{2n}$, é uma função integrável e portanto o teorema de Fubini implica que também vale

$$\varphi_\epsilon(x) = \frac{1}{(2\pi)^n} \int d^n y \int d^n \xi \varphi(y) \exp(i\xi \cdot (x - y)) \exp(-\epsilon \xi^2). \quad (3.236)$$

Agora, na equação (3.235), podemos executar a integral sobre y para obter

$$\varphi_\epsilon(x) = \frac{1}{(2\pi)^n} \int d^n \xi (\mathcal{F}\varphi)(\xi) \exp(i\xi \cdot x) \exp(-\epsilon \xi^2).$$

Mas como $\mathcal{F}\varphi \in \mathcal{S}$ e $\exp(-\epsilon|\xi|^2) \leq 1$ para todo $\epsilon > 0$ e todo $\xi \in \mathbb{R}^n$, segue do teorema da convergência dominada que

$$\lim_{\epsilon \rightarrow 0} \varphi_\epsilon(x) = (\tilde{\mathcal{F}}\mathcal{F}\varphi)(x). \quad (3.237)$$

Por outro lado, na equação (3.236), podemos executar a integral sobre ξ usando complementação quadrática do expoente, conforme a qual

$$\begin{aligned} & \int d^n \xi \exp(i\xi \cdot (x - y)) \exp(-\epsilon \xi^2) \\ &= \exp\left(-\frac{(x - y)^2}{4\epsilon}\right) \int d^n \xi \exp\left(-\epsilon\left(\xi^2 - \frac{i}{\epsilon}\xi \cdot (x - y) - \frac{(x - y)^2}{4\epsilon^2}\right)\right) \\ &= \exp\left(-\frac{(x - y)^2}{4\epsilon}\right) \int d^n \xi \exp\left(-\epsilon\left(\xi - \frac{i}{2\epsilon}(x - y)\right)^2\right) \\ &= \exp\left(-\frac{(x - y)^2}{4\epsilon}\right) \int d^n \xi \exp(-\epsilon \xi^2) \\ &= \left(\frac{\pi}{\epsilon}\right)^{n/2} \exp\left(-\frac{(x - y)^2}{4\epsilon}\right), \end{aligned}$$

onde usamos a equação (3.228) no último passo e o Lema 3.3, aplicado à função analítica $\exp(-\zeta^2)$ ($\zeta = \xi + i\eta$) no penúltimo passo, para obter

$$\varphi_\epsilon(x) = \frac{1}{(4\pi\epsilon)^{n/2}} \int d^n y \varphi(y) \exp\left(-\frac{(x - y)^2}{4\epsilon}\right).$$

Usando novamente a equação (3.228) e aplicando o critério estabelecido no Exemplo 3.17, concluímos que

$$\lim_{\epsilon \rightarrow 0} \varphi_\epsilon(x) = \varphi(x). \quad (3.238)$$

□

A expressão “quase uma involução” utilizada acima se refere ao fato de que $(2\pi)^{-n/2}\mathcal{F}$ é uma involução a menos de conjugação com a conjugação complexa, ou seja, vale

$$\mathcal{F}^{-1}\varphi = \frac{1}{(2\pi)^n} \overline{\mathcal{F}\varphi}. \quad (3.239)$$

Agora estamos prontos para estender a definição da transformação de Fourier de funções para distribuições, mais exatamente, para distribuições temperadas:

Definição 3.36 Para qualquer distribuição temperada T sobre \mathbb{R}^n , define-se a sua **transformada de Fourier** como sendo a distribuição temperada $\mathcal{F}T$ sobre \mathbb{R}^n dada pela fórmula

$$\langle \mathcal{F}T, \varphi \rangle = \langle T, \mathcal{F}\varphi \rangle \quad \text{para } \varphi \in \mathcal{S}. \quad (3.240)$$

Observa-se que $\mathcal{F}T$ é realmente uma distribuição temperada, i.e., um funcional linear contínuo sobre \mathcal{S} , pois é a composição do funcional linear contínuo T sobre \mathcal{S} com a aplicação linear contínua (3.227).

Mais uma vez, notamos que para distribuições regulares, esta definição coincide com a definição clássica:

Proposição 3.51 Seja f uma função integrável sobre \mathbb{R}^n e \hat{f} sua transformada de Fourier. Então vale

$$\mathcal{F}T_f = T_{\mathcal{F}f}. \quad (3.241)$$

DEMONSTRAÇÃO: Para qualquer função teste $\varphi \in \mathcal{D}$ ou mesmo $\varphi \in \mathcal{S}$, podemos aplicar o teorema de Fubini para mostrar que

$$\begin{aligned} \int d^n x f(x) \mathcal{F}\varphi(x) &= \int d^n x \int d^n \xi f(x) \varphi(\xi) \exp(-i\xi \cdot x) \\ &= \int d^n \xi \int d^n x f(x) \varphi(\xi) \exp(-i\xi \cdot x) = \int d^n \xi \mathcal{F}f(\xi) \varphi(\xi), \end{aligned}$$

e portanto

$$\langle \mathcal{F}T_f, \varphi \rangle = \langle T_f, \mathcal{F}\varphi \rangle = \langle T_{\mathcal{F}f}, \varphi \rangle.$$

□

Usando a primeira igualdade neste cálculo, para o caso especial em que $f \in \mathcal{S}$, com $\varphi \in \mathcal{S}$ substituído por $\psi \in \mathcal{S}$ e f dada por $f = (2\pi)^n \mathcal{F}^{-1}\bar{\varphi} = \overline{\mathcal{F}\varphi}$ com $\varphi \in \mathcal{S}$ (veja a equação (3.239)), chegamos à fórmula de Plancherel:

$$\int d^n \xi \bar{\varphi}(\xi) \hat{\psi}(\xi) = (2\pi)^n \int d^n x \bar{\varphi}(x) \psi(x). \quad (3.242)$$

Em um primeiro momento, esta fórmula vale para $\varphi, \psi \in \mathcal{S}$, mas como \mathcal{S} é denso no espaço de Hilbert $L^2(\mathbb{R}^n)$, ela continua valendo, por extensão contínua, para $\varphi, \psi \in \mathcal{L}^2(\mathbb{R}^n)$, ou seja, $(2\pi)^{-n/2}\mathcal{F}$ é uma isometria do espaço de Hilbert $L^2(\mathbb{R}^n)$.

Exemplo 3.18 Vale

$$\mathcal{F}\delta = 1, \quad \mathcal{F}1 = (2\pi)^n \delta. \quad (3.243)$$

De fato, para $\varphi \in \mathcal{S}$,

$$\begin{aligned} \langle \mathcal{F}\delta, \varphi \rangle &= \langle \delta, \mathcal{F}\varphi \rangle = \int d^n x \varphi(x) = \langle 1, \varphi \rangle, \\ \langle \mathcal{F}^{-1}\delta, \varphi \rangle &= \langle \delta, \mathcal{F}^{-1}\varphi \rangle = \frac{1}{(2\pi)^n} \int d^n x \varphi(x) = \frac{1}{(2\pi)^n} \langle 1, \varphi \rangle, \end{aligned}$$

3.14 O teorema de Payley-Wiener

Na seção anterior, vimos que a transformada de Fourier

3.15 O conjunto de frente de ondas

A SER CONTINUADO

3.16 Aplicações a equações diferenciais parciais

A SER ELABORADO

A O teorema do fluxo

Neste apêndice, queremos formular, de maneira precisa, porém sem demonstrações, o teorema do fluxo para equações diferenciais ordinárias, mais exatamente para sistemas autônomos de equações diferenciais ordinárias de primeira ordem – lineares ou não.

Um sistema autônomo de equações diferenciais ordinárias de primeira ordem pode ser definido fixando um domínio Ω em \mathbb{R}^n e um campo vetorial X definido em Ω , isto é, uma aplicação $X : \Omega \rightarrow \mathbb{R}^n$. A equação diferencial associada com estes dados é

$$\frac{du}{dt}(t) = X(u(t)) . \quad (\text{A.1})$$

O teorema de existência e unicidade local de soluções afirma que quando X é de classe C^1 ou pelo menos satisfaz uma condição de Lipschitz, então para qualquer ponto u_0 de Ω , existe uma única solução local u da equação (A.1) com condição inicial u_0 , ou seja, uma única curva $u :]a, b[\rightarrow \Omega$ (onde $a < 0$ e $b > 0$ são tais que $u(]a, b[) \subset \Omega$), de classe C^1 , tal que

$$\dot{u}(t) = X(u(t)) \quad \text{para } a < t < b \quad (\text{A.2})$$

e

$$u(0) = u_0 . \quad (\text{A.3})$$

Por diferenciação repetida da equação (A.2), podemos concluir que para qualquer inteiro $k > 0$, u será de classe C^{k+1} quando X é de classe C^k , e portanto que u será de classe C^∞ quando X é de classe C^∞ .

Quanto à questão do domínio de definição da solução, enfrentamos em primeiro lugar o problema que os números a e b não são univocamente determinados. Este problema se resolve através da observação que, devido à unicidade, duas soluções definidas em intervalos diferentes devem coincidir na interseção destes dois intervalos e fornecem assim uma única solução definida na união dos dois intervalos. Portanto, existem um único número $a_{\min} < 0$ (possivelmente $-\infty$) e um único número $b_{\max} > 0$ (possivelmente $+\infty$) tais que a solução u seja definida sobre o intervalo $]a_{\min}, b_{\max}[$ mas não admita nenhuma extensão para um intervalo maior:

falamos então da solução maximal (ou maximalmente estendida). Observamos que em geral, os valores de a_{\min} e b_{\max} ainda dependem da condição inicial u_0 .

O próximo passo nesta linha de argumentação é formar o subconjunto

$$\mathcal{D}_X = \bigcup_{u_0 \in \Omega}]a_{\min}(u_0), b_{\max}(u_0)[\times \{u_0\}, \quad (\text{A.4})$$

do produto cartesiano $\mathbb{R} \times \Omega$, e considerar a aplicação

$$\mathcal{F}_X : \mathcal{D}_X \longrightarrow \Omega \quad (\text{A.5})$$

definida de modo que para qualquer ponto u_0 em Ω , a aplicação

$$\mathcal{F}(\cdot, u_0) :]a_{\min}(u_0), b_{\max}(u_0)[\rightarrow \Omega$$

seja a solução maximal da equação diferencial sob consideração com condição inicial u_0 , ou seja, vale

$$\begin{aligned} \frac{\partial \mathcal{F}_X}{\partial t}(t, u_0) &\equiv \dot{\mathcal{F}}_X(t, u_0) = X(\mathcal{F}(t, u_0)) \\ &\text{para } a_{\min}(u_0) < t < b_{\max}(u_0) \end{aligned} \quad (\text{A.6})$$

e

$$\mathcal{F}_X(0, u_0) = u_0. \quad (\text{A.7})$$

A aplicação \mathcal{F} se chama *fluxo* do campo vetorial X , e \mathcal{D}_X é o seu *domínio*.

Com estas noções, podemos formular o teorema do fluxo, que inclui e unifica o teorema da existência e unicidade das soluções e o teorema da dependência das soluções em relação às condições iniciais, como afirmando que \mathcal{D}_X é um subconjunto aberto de $\mathbb{R} \times \Omega$ e que para qualquer inteiro $k > 0$, \mathcal{F}_X é uma aplicação de classe C^k quando X é de classe C^k , e portanto que \mathcal{F}_X será de classe C^∞ quando X é de classe C^∞ .

Na mesma direção, mencionamos ainda que um campo vetorial é chamado *completo* quando o domínio \mathcal{D}_X do seu fluxo \mathcal{F}_X é o produto cartesiano $\mathbb{R} \times \Omega$ inteiro. Isto significa que todas as soluções da equação diferencial definida por X podem ser estendidas à reta real inteira.

Finalmente, é útil observar que o caso de sistemas não-autônomos, ou seja, de sistemas de equações diferenciais ordinárias de primeira ordem onde o campo vetorial depende explicitamente do tempo, pode ser reduzido ao caso de sistemas autônomos através de um aumento de dimensão do sistema por uma unidade.

Passamos agora à identificação das equações diferenciais lineares. A equação diferencial associada ao campo vetorial X será linear se e somente se X é uma função linear, ou mais geralmente, uma função afim: no primeiro caso, falamos de um sistema linear homogêneo e no segundo, de um sistema linear não-homogêneo. Isto significa que podemos tomar $\Omega = \mathbb{R}^n$ e que X será uma aplicação linear de \mathbb{R}^n em \mathbb{R}^n , dada por uma matriz A , ou mais geralmente, uma aplicação afim de \mathbb{R}^n em \mathbb{R}^n , dada por uma matriz A e um vetor a .

B Resumo do cálculo vetorial

Neste apêndice, apresentamos alguns resultados do cálculo vetorial em \mathbb{R}^n que são amplamente utilizados na teoria clássica de equações diferenciais parciais, a saber, o teorema de Gauss ou teorema da divergência e as identidades de Green, que constituem ferramentas indispensáveis para o tratamento de problemas de fronteira, tais como o problema de Dirichlet ou de Neumann para o laplaciano em domínios limitados de \mathbb{R}^n .

B.1 Integração sobre subvariedades de \mathbb{R}^n : teoria local

Começamos com uma breve descrição da construção da integral de funções sobre hipersuperfícies em \mathbb{R}^n e, mais geralmente, sobre subvariedades de \mathbb{R}^n . Para evitar complicações desnecessárias, suporemos que integramos apenas funções contínuas, para as quais as integrais de Riemann e de Lebesgue coincidem. O cerne da questão é encontrar uma definição que não depende das coordenadas (locais) utilizadas para parametrizar a hipersuperfície ou subvariedade sob consideração. Para tanto, precisamos da fórmula de mudança de variáveis em integrais múltiplas, que rege o comportamento da integral em \mathbb{R}^n sob transformações de coordenadas (locais), abstratamente representadas por difeomorfismos (locais).

Definição B.1 *Sejam U um aberto de \mathbb{R}^n e V um aberto de \mathbb{R}^m . Uma aplicação $\varphi : U \rightarrow V$ é chamada um **difeomorfismo** de classe C^r se 1) φ é bijetora e 2) $\varphi : U \rightarrow V$ e $\varphi^{-1} : V \rightarrow U$ são de classe C^r .*

Aqui e a seguir, r pode assumir qualquer um dos valores $0, 1, \dots, \infty, \omega$. Outras observações são as seguintes:

- Segundo a terminologia padrão em topologia, os difeomorfismos de classe C^0 são chamados de homeomorfismos.

- A hipótese de que além da aplicação dada φ , a aplicação inversa φ^{-1} também seja de classe C^r é importante e não pode ser dispensada. Considere, por exemplo, a aplicação $\varphi : \mathbb{R} \rightarrow \mathbb{R}$ definida por $\varphi(x) = x^3$: ela é bijetora e de classe C^∞ , enquanto que a aplicação inversa $\varphi^{-1} : \mathbb{R} \rightarrow \mathbb{R}$, dada por $\varphi^{-1}(y) = \sqrt[3]{y}$, é contínua, mas não é diferenciável na origem: portanto, φ é um homeomorfismo, mas não é um difeomorfismo de classe C^1 .
- Quando $r \geq 1$, uma condição necessária e suficiente para que uma aplicação $\varphi : U \rightarrow V$ bijetora e de classe C^r seja um difeomorfismo de classe C^r é que a sua derivada $\varphi'(x)$ (i.e., a matriz jacobiana) seja não-degenerada em todo ponto $x \in U$: esta é exatamente a afirmação do teorema das funções inversas.
- A existência de um difeomorfismo φ entre um aberto U de \mathbb{R}^n e um aberto V de \mathbb{R}^m implica que $m = n$, ou seja: a dimensão é invariante sob difeomorfismos. Quando $r \geq 1$, isto é imediato pelo item anterior, enquanto que para $r = 0$, é garantido por um teorema não-trivial de topologia geral.

O conceito matemático de difeomorfismo (entre abertos de \mathbb{R}^n) confere um significado preciso à noção intuitiva de uma “transformação de coordenadas” ou “mudança de variáveis”, na medida em que identifica as exigências matemáticas a serem impostas para que tal transformação ou mudança seja admissível. De fato, através de um difeomorfismo $\varphi : U \rightarrow V$, as coordenadas cartesianas x_1, \dots, x_n dos pontos em U podem ser consideradas como funções das coordenadas cartesianas y_1, \dots, y_n dos pontos em V , de modo que os y_i se tornam um sistema de coordenadas (geralmente não-cartesianas) para os pontos em U , e vice versa.

A título de exemplo, e como subsídio para o desenvolvimento subsequente, formulamos aqui o teorema clássico sobre o comportamento da integral de Riemann-Lebesgue sob difeomorfismos:

Teorema B.1 *Sejam U e V abertos de \mathbb{R}^n e $\varphi : U \rightarrow V$ um difeomorfismo de classe C^1 . Então, para toda função f contínua sobre V e todo compacto K em U , vale a fórmula de mudança de variáveis em integrais múltiplas*

$$\int_{\varphi(K)} d^n y f(y) = \int_K d^n x |\det(\varphi'(x))| f(\varphi(x)). \quad (\text{B.1})$$

O próximo passo será dar uma definição do conceito de uma subvariedade de \mathbb{R}^n : tais subvariedades serão os domínios de integração.

Definição B.2 *Uma carta ou um sistema de coordenadas de classe C^r de \mathbb{R}^n (em torno de um ponto $x^{(0)}$ de \mathbb{R}^n) é uma tripla (U, φ, U') formada por uma vizinhança aberta U de $x^{(0)}$ em \mathbb{R}^n , uma vizinhança aberta U' de 0 em \mathbb{R}^n e um difeomorfismo $\varphi : U \rightarrow U'$ de classe C^r (tal que $\varphi(x^{(0)}) = 0$). O aberto U é chamado o domínio da carta ou do sistema de coordenadas. Quando especificamos tão somente o difeomorfismo φ , sem fixar os abertos U e U' , falamos de uma carta local ou de um sistema de coordenadas locais (em torno de $x^{(0)}$).*

Definição B.3 Uma **subvariedade** p -dimensional de \mathbb{R}^n de classe C^r é um subconjunto X de \mathbb{R}^n tal que para cada ponto $x^{(0)}$ de X , existe uma carta (U, φ, U') de classe C^r de \mathbb{R}^n em torno de $x^{(0)}$ tal que $\varphi(U \cap X) = U' \cap \mathbb{R}^p$, onde \mathbb{R}^p é identificado com um subespaço de \mathbb{R}^n ; por exemplo,

$$\mathbb{R}^p \cong \left\{ \left(\begin{array}{c} x_1 \\ \vdots \\ x_p \\ 0 \\ \vdots \\ 0 \end{array} \right) \in \mathbb{R}^n / \left(\begin{array}{c} x_1 \\ \vdots \\ x_p \end{array} \right) \in \mathbb{R}^p \right\} \quad (\text{B.2})$$

Tais cartas ou (sistemas de) coordenadas de \mathbb{R}^n serão ditas **adaptadas** à subvariedade X .

Um exemplo simples é exibido na Figura B.1.

Fig. B.1: Coordenadas adaptadas a uma curva no plano, como exemplo simples de uma subvariedade unidimensional de \mathbb{R}^2

De forma geral, notamos que quando a tripla (U, φ, U') é uma carta de classe C^r de \mathbb{R}^n em torno de $x^{(0)}$ adaptada a X , a tripla $(U \cap X, \varphi|_{U \cap X}, U' \cap \mathbb{R}^p)$ fornecerá uma carta para X em torno de $x^{(0)}$, no sentido de que a aplicação inversa $(\varphi|_{U \cap X})^{-1} = \varphi^{-1}|_{U' \cap \mathbb{R}^p} : U' \cap \mathbb{R}^p \rightarrow U \cap X$ pode ser vista como uma *parametrização* dos pontos de X em torno de $x^{(0)}$. Mais explicitamente, para pontos

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in U \quad \text{e} \quad y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \in U',$$

podemos considerar a correspondência bijetora dada por

$$y = \varphi(x) \quad \text{e} \quad x = \varphi^{-1}(y)$$

como fornecendo um sistema de coordenadas y_1, \dots, y_p ao longo de X complementado por um sistema de coordenadas y_{p+1}, \dots, y_n transversais a X , pois os pontos de U pertencendo a $U \cap X$ são caracterizados pelo conjunto de equações

$$y_{p+1} = 0, \dots, y_n = 0.$$

De fato, $U \cap X$ é tão somente um membro de uma família de subvariedades $(U \cap X)(a_{p+1}, \dots, a_n)$ de \mathbb{R}^n , definidas pelo conjunto de equações

$$y_{p+1} = a_{p+1}, \dots, y_n = a_n,$$

onde a_{p+1}, \dots, a_n são constantes. A versão global desta observação leva ao conceito de uma *folheação* por subvariedades, ao invés de uma única subvariedade isolada.

Isso posto, podemos estender a noção de integração de funções contínuas sobre compactos de \mathbb{R}^n a funções contínuas sobre compactos de subvariedades de \mathbb{R}^n . Inicialmente, consideramos compactos “pequenos”, contidos em domínios de cartas.

Definição B.4 *Sejam X uma subvariedade p -dimensional de \mathbb{R}^n de classe C^1 , f uma função contínua sobre X e K um compacto de X contido no domínio $U \cap X$ de alguma carta de X provindo de uma carta (U, φ, U') de \mathbb{R}^n adaptada a X . Considerando-se as coordenadas cartesianas x_1, \dots, x_n dos pontos de $U \cap X$ como funções das coordenadas y_1, \dots, y_p dos pontos de $U' \cap \mathbb{R}^p$, funções dadas pela aplicação φ^{-1} , introduzimos a função matricial*

$$g^\varphi(y) = \begin{pmatrix} g_{11}^\varphi(y) & \cdots & g_{1p}^\varphi(y) \\ \vdots & \ddots & \vdots \\ g_{p1}^\varphi(y) & \cdots & g_{pp}^\varphi(y) \end{pmatrix}, \quad (\text{B.3})$$

onde

$$g_{ij}^\varphi(y) = \frac{\partial x}{\partial y_i}(y) \cdot \frac{\partial x}{\partial y_j}(y) \equiv \sum_{m=1}^n \frac{\partial x_m}{\partial y_i}(y) \frac{\partial x_m}{\partial y_j}(y) \quad (1 \leq i, j \leq p), \quad (\text{B.4})$$

e definimos a **integral de f sobre K** pela fórmula

$$\int_K d\sigma(x) f(x) = \int_{\varphi(K)} d^p y \sqrt{\det(g^\varphi(y))} f(\varphi^{-1}(y)). \quad (\text{B.5})$$

A função matricial g^φ é chamada a **métrica induzida** e a função $\det g^\varphi$, o **determinante métrico**, em relação ao sistema de coordenadas utilizado.

A motivação para introduzir o determinante métrico nesta fórmula reside na observação de que a integral assim definida não depende do sistema de coordenadas escolhido.

Para mostrar isso, suponhamos que (U, φ, U') e (V, ψ, V') sejam duas cartas de \mathbb{R}^n adaptadas a X tais que $K \subset U \cap V$. Queremos mostrar que

$$\int_{\varphi(K)} d^p y \sqrt{\det(g^\varphi(y))} f(\varphi^{-1}(y)) = \int_{\psi(K)} d^p z \sqrt{\det(g^\psi(z))} f(\psi^{-1}(z)).$$

Devido ao teorema B.1, aplicado ao difeomorfismo

$$\chi = \psi|_{U \cap V \cap X} \circ (\varphi|_{U \cap V \cap X})^{-1} : (U \cap V) \cap \mathbb{R}^p \longrightarrow \psi(U \cap V) \cap \mathbb{R}^p,$$

temos

$$\begin{aligned} \int_{\psi(K)} d^p z \sqrt{\det(g^\psi(z))} f(\psi^{-1}(z)) \\ = \int_{\varphi(K)} d^p y |\det(\chi'(y))| \sqrt{\det(g^\psi(\chi(y)))} f(\varphi^{-1}(y)) . \end{aligned}$$

Destarte, é suficiente que valha

$$\det(g^\varphi(y)) = (\det(\chi'(y)))^2 \cdot \det(g^\psi(\chi(y))) .$$

Mas pela regra da cadeia,

$$(\varphi^{-1})'(y) = (\psi^{-1})'(\chi(y)) \cdot \chi'(y)$$

ou seja,

$$\frac{\partial x_m}{\partial y_i}(y) = \sum_{k=1}^p \frac{\partial x_m}{\partial z_k}(\chi(y)) \frac{\partial z_k}{\partial y_i}(y)$$

e portanto

$$\begin{aligned} g_{ij}^\varphi(y) &= \sum_{m=1}^n \frac{\partial x_m}{\partial y_i} \frac{\partial x_m}{\partial y_j} = \sum_{m=1}^n \sum_{k,l=1}^p \frac{\partial x_m}{\partial z_k}(\chi(y)) \frac{\partial z_k}{\partial y_i}(y) \frac{\partial x_m}{\partial z_l}(\chi(y)) \frac{\partial z_l}{\partial y_j}(y) \\ &= \sum_{k,l=1}^p \frac{\partial z_k}{\partial y_i}(y) g_{kl}^\psi(\chi(y)) \frac{\partial z_l}{\partial y_j}(y) . \end{aligned}$$

Usando que o determinante de um produto de matrizes é igual ao produto dos determinantes dos fatores e que o determinante de uma matriz é igual ao da sua transposta, obtemos o resultado desejado.

B.2 Integração sobre subvariedades de \mathbb{R}^n : teoria global

A definição da integral no caso geral – isto é, quando o compacto K não está contido em um único domínio de carta de X – pode ser reduzida ao caso anterior mediante a utilização de uma partição da unidade. Partições da unidade constituem uma das ferramentas mais importantes e poderosas de análise, providenciando um embasamento técnico para a idéia de construir um objeto global (neste caso, a integral de uma função sobre um compacto “grande”) a partir de uma coleção de objetos locais (neste caso, uma coleção de integrais da mesma função sobre compactos “pequenos”), através de um processo de colagem.

Para poder introduzir o conceito de uma partição da unidade, precisamos de algumas definições preliminares. A primeira é a do suporte de uma função.

Definição B.5 O *suporte* de uma função f sobre um aberto Ω de \mathbb{R}^n , denotado por $\text{supp } f$, é o fecho, em Ω , do conjunto dos pontos nos quais ela não se anula:

$$\text{supp } f = \overline{\{x \in \Omega / f(x) \neq 0\}} . \quad (\text{B.6})$$

Em outras palavras, o suporte de uma função f sobre Ω pode ser caracterizado como o complemento, em Ω , do maior aberto de Ω onde f se anula:

$$\text{supp } f = \Omega \setminus \left(\text{maior aberto de } \Omega \text{ onde } f \text{ se anula} \right) . \quad (\text{B.7})$$

Note que, conforme esta definição, o suporte de uma função sobre um aberto Ω de \mathbb{R}^n é fechado em Ω , mas não necessariamente em \mathbb{R}^n .

A notação “supp” se explica pela ortografia da palavra “support” em inglês e pela necessidade de evitar uma possível confusão entre as noções do suporte e do supremo de uma função. Note as inclusões óbvias

$$\text{supp } (\lambda f + \mu g) \subset \text{supp } f \cup \text{supp } g , \quad (\text{B.8})$$

$$\text{supp } (fg) \subset \text{supp } f \cap \text{supp } g , \quad (\text{B.9})$$

que implicam que dado um espaço ou uma álgebra de funções de um certo tipo (por exemplo, funções contínuas ou funções lisas) com suporte contido em algum subconjunto fixo de \mathbb{R}^n é um subespaço deste espaço ou subálgebra desta álgebra.

As próximas duas definições referem-se a recobrimentos abertos de subconjuntos de \mathbb{R}^n .

Definição B.6 Sejam X um subconjunto de \mathbb{R}^n e $(\Omega_i)_{i \in I}$ um recobrimento aberto de X , isto é, uma família de subconjuntos abertos Ω_i de \mathbb{R}^n tal que

$$X \subset \bigcup_{i \in I} \Omega_i .$$

Este é dito

- um **recobrimento finito** se o conjunto I é finito;
- um **recobrimento pontualmente finito** se para todo ponto x em X , o conjunto $I(x) = \{i \in I / x \in \Omega_i\}$ é finito;
- um **recobrimento localmente finito** se todo ponto x em X possui uma vizinhança aberta Ω_x tal que o conjunto $I(\Omega_x) = \{i \in I / \Omega_x \cap \Omega_i \neq \emptyset\}$ é finito.

Definição B.7 Seja X um subconjunto de \mathbb{R}^n . Um recobrimento (aberto) $(\Omega_i)_{i \in I}$ de X é chamado um **refinamento** de um recobrimento (aberto) $(\Omega'_{i'})_{i' \in I'}$ de X se existe uma aplicação de refinamento $\rho : I \rightarrow I'$ tal que para todo $i \in I$, vale $\Omega_i \subset \Omega'_{\rho(i)}$. Quando a aplicação ρ é injetora e tal que $\Omega_i = \Omega'_{\rho(i)}$, dizemos que $(\Omega_i)_{i \in I}$ é um **sub-recobrimento** de $(\Omega'_{i'})_{i' \in I'}$.

Uma versão simplificada desta definição consiste em afirmar que um recobrimento (aberto) $(\Omega_i)_{i \in I}$ de X é um refinamento de um outro recobrimento (aberto) $(\Omega'_{i'})_{i' \in I'}$ de X se e somente se para todo $i \in I$, existe $i' \in I'$ tal que $\Omega_i \subset \Omega'_{i'}$. Como um dado aberto Ω_i pode estar contido em vários dos abertos $\Omega'_{i'}$, a construção de uma aplicação de refinamento requer então o uso do axioma da escolha. De maneira semelhante, observa-se que, em geral, a aplicação de refinamento não será injetora, sendo que um dado aberto $\Omega'_{i'}$ pode conter vários dos abertos Ω_i .

Com estas noções à nossa disposição, podemos definir dois conceitos de topologia: um que é muito bem conhecido e outro menos conhecido, mas também útil:

Definição B.8 *Um subconjunto X de \mathbb{R}^n é dito*

- (i) **compacto** quando qualquer recobrimento aberto de X admite um sub-recobrimento finito,
- (ii) **paracompacto** quando qualquer recobrimento aberto de X admite um refinamento localmente finito.

De passagem, mencionamos que nestas últimas três definições, podemos substituir o espaço \mathbb{R}^n por qualquer espaço topológico de Hausdorff. Notemos também a seguinte

Proposição B.1 *Todo subconjunto X de \mathbb{R}^n é paracompacto.*

Muito mais geralmente, mostra-se que qualquer espaço métrico é paracompacto.

Finalmente, voltando para o caso de subconjuntos X de \mathbb{R}^n , estamos em condições de formular a definição de uma partição da unidade subordinada a um recobrimento aberto localmente finito de X :

Definição B.9 *Uma **partição da unidade** (de classe C^r) subordinada a um recobrimento aberto localmente finito $(\Omega_i)_{i \in I}$ de um subconjunto X de \mathbb{R}^n é uma família $(\chi_i)_{i \in I}$ de funções (de classe C^r) sobre \mathbb{R}^n a valores reais com as seguintes propriedades:*

$$0 \leq \chi_i \leq 1 \quad \text{para todo } i \in I, \quad (\text{B.10})$$

$$\text{supp } \chi_i \subset \Omega_i, \quad \text{supp } \chi_i \text{ compacto} \quad \text{para todo } i \in I, \quad (\text{B.11})$$

$$\sum_{i \in I} \chi_i = 1 \quad \text{sobre } X. \quad (\text{B.12})$$

Note que a soma na equação (B.12) é bem definida devido à condição (B.11) e à condição de que o recobrimento $(\Omega_i)_{i \in I}$ seja localmente finito, pois essas duas hipóteses implicam que todo ponto x de X possui uma vizinhança aberta Ω_x sobre a qual a soma na equação (B.12) se estende apenas sobre o subconjunto finito $I(\Omega_x) = \{i \in I / \Omega_x \cap \Omega_i \neq \emptyset\}$ de I . Ademais, a Proposição B.1 mostra que se partirmos de um recobrimento aberto qualquer, podemos sempre passar para um refinamento que seja localmente finito e, sem perda de generalidade, formado por

abertos limitados, e depois procurar uma partição da unidade subordinada a este refinamento.

O teorema fundamental sobre partições da unidade afirma sua existência:

Teorema B.2 *Para qualquer recobrimento localmente finito $(\Omega_i)_{i \in I}$ de qualquer subconjunto fechado X de \mathbb{R}^n por abertos limitados Ω_i de \mathbb{R}^n , sempre existe uma partição da unidade $(\chi_i)_{i \in I}$ de classe C^∞ a ele subordinada.*

A demonstração é tecnicamente exigente e não será apresentada aqui. Observamos apenas que o caso mais importante para aplicações na prática é o de um recobrimento finito $(\Omega_i)_{i \in I}$ de um compacto K em \mathbb{R}^n . Mas mesmo no caso mais simples onde o recobrimento se reduz a um único aberto Ω contendo K , a afirmação é não-trivial: esta versão encontra-se no Capítulo 3 (veja o Teorema 3.4) e constitui um ingrediente importante para a demonstração do caso geral.

Isso posto, podemos indicar como se define a integral de uma função contínua f sobre um subconjunto compacto K de uma subvariedade X de classe C^r de \mathbb{R}^n mesmo quando K não está contido no domínio de uma única carta. Neste caso, escolhemos um recobrimento aberto finito $(U_i)_{i \in I}$ de K por domínios U_i de cartas (U_i, φ_i, U'_i) de \mathbb{R}^n adaptadas a X e uma partição da unidade $(\chi_i)_{i \in I}$ subordinada a este recobrimento e pomos

$$\int_K d\sigma(x) f(x) = \sum_{i \in I} \int_{\text{supp } \chi_i \cap K} d\sigma(x) \chi_i(x) f(x). \quad (\text{B.13})$$

É possível mostrar que o resultado não depende da escolha do recobrimento $(U_i)_{i \in I}$ de K ou da partição da unidade $(\chi_i)_{i \in I}$, mas não entraremos nos detalhes da demonstração deste fato.

Apesar da importância dessa definição da integral do ponto de vista teórico, vale salientar que, na prática, não é dessa maneira que se calculam integrais de funções sobre subvariedades de \mathbb{R}^n . De fato, em vez de lançarmos mão de uma partição de unidade, podemos na grande maioria dos casos que ocorrem na prática usar o fato de que a subvariedade X em questão, apesar de não poder ser recoberta por um único domínio de carta, contém algum domínio de carta Y que é “grande” no sentido de ser aberto e denso, de modo que podemos substituir a integral sobre X pela integral sobre Y e calcular essa usando um sistema de coordenadas conveniente, já que o complemento $X \setminus Y$ de Y é de medida zero e portanto não contribui para a integral. A única precaução que precisa ser tomada é que, em geral, Y não será compacto, mesmo quando X for compacto. Como exemplo típico, consideraremos logo abaixo o caso em que X é a esfera unitária S^{n-1} em \mathbb{R}^n e Y é o domínio obtido removendo um único ponto de X , por exemplo o “polo norte” $(0, \dots, 0, 1)$ ou o “polo sul” $(0, \dots, 0, -1)$; coordenadas convenientes para essa escolha são as obtidas pela projeção estereográfica.

B.3 Integração de campos vetoriais

As integrais que tipicamente aparecem no cálculo vetorial são integrais de campos vetoriais \mathbf{A} ao longo de curvas γ e sobre hipersuperfícies Σ em \mathbb{R}^n , devidamente orientadas. Para explicar a sua definição, lembramos primeiro a seguinte

Definição B.10 *Um campo escalar (real) ϕ sobre um aberto Ω de \mathbb{R}^n ou de uma subvariedade X de \mathbb{R}^n é uma aplicação $\phi : \Omega \rightarrow \mathbb{R}$. Um campo vetorial \mathbf{A} sobre um aberto Ω de \mathbb{R}^n ou de uma subvariedade X de \mathbb{R}^n é uma aplicação $\mathbf{A} : \Omega \rightarrow \mathbb{R}^n$.*

Por outro lado, de um ponto de vista puramente geométrico, uma **curva** em \mathbb{R}^n é simplesmente uma subvariedade γ de dimensão 1, enquanto que uma **hipersuperfície** S em \mathbb{R}^n é simplesmente uma subvariedade de codimensão 1, ou seja, de dimensão $n - 1$. Supondo que estas subvariedades sejam de classe C^r ($r \geq 1$) e sejam *orientadas*, podemos definir, de maneira natural, o **campo unitário \mathbf{t} tangente a γ** e o **campo unitário \mathbf{n} normal a S** . De fato, para cada ponto x de uma curva γ de classe C^r ($r \geq 1$), existem exatamente dois vetores $\pm \mathbf{t}(x)$ de comprimento 1 tangentes a γ em x , e para cada ponto x de uma hipersuperfície S de classe C^r ($r \geq 1$), existem exatamente dois vetores $\pm \mathbf{n}(x)$ de comprimento 1 ortogonais a S em x . Por definição, a escolha de uma orientação de γ ou S equivale à escolha de uma dessas duas possibilidades, exigindo-se porém que $\mathbf{t}(x)$ e $\mathbf{n}(x)$, como funções de x , sejam de classe C^r . (No caso das curvas, é sempre possível escolher uma orientação: curvas são sempre orientáveis. No caso das hipersuperfícies, isso já não acontece: existem hipersuperfícies que não são orientáveis, tais como a notória faixa de Möbius. Geralmente, as noções de orientabilidade e de orientação podem ser formuladas para subvariedades de \mathbb{R}^n de qualquer dimensão, e mostra-se que subvariedades orientáveis admitem exatamente duas orientações opostas.)

Destarte, para todo campo vetorial \mathbf{A} definido e contínuo sobre uma curva orientada γ de classe C^1 , fica bem definida a componente de \mathbf{A} tangencial a γ , $A_{\parallel} = \mathbf{t} \cdot \mathbf{A}$, assim como a integral de A_{\parallel} ao longo de γ ,

$$\int_{\gamma} d\mathbf{x} \cdot \mathbf{A} = \int_{\gamma} d\sigma(x) A_{\parallel}(x) = \int_{\gamma} d\sigma(x) \mathbf{t}(x) \cdot \mathbf{A}(x), \quad (\text{B.14})$$

chamada **circulação de \mathbf{A} ao longo de γ** caso a curva γ for fechada, enquanto que para todo campo vetorial \mathbf{A} definido e contínuo sobre uma hipersuperfície orientada S de classe C^1 , fica bem definida a componente de \mathbf{A} normal a S , $A_{\perp} = \mathbf{n} \cdot \mathbf{A}$, assim como a integral de A_{\perp} sobre S ,

$$\int_S d\boldsymbol{\sigma} \cdot \mathbf{A} = \int_S d\sigma(x) A_{\perp}(x) = \int_S d\sigma(x) \mathbf{n}(x) \cdot \mathbf{A}(x), \quad (\text{B.15})$$

chamada **fluxo de \mathbf{A} através de S** ; veja a Figura B.2. Note que invertendo-se a orientação, as integrais assim definidas mudam de sinal.

A definição das integrais de campos vetoriais ao longo de curvas (conexas) pode ser simplificada pela introdução de uma parametrização explícita. De fato,

Fig. B.2: Integração de campos vetoriais ao longo de curvas e sobre superfícies

como subvariedade unidimensional de \mathbb{R}^n de classe C^r , γ sempre admite uma parametrização global, ou seja, pode ser escrita como imagem de uma aplicação $c : I \rightarrow \mathbb{R}^n$ de classe C^r , definida num intervalo $I = [a, b]$ e compatível com a orientação de γ no sentido de que

$$\mathbf{t}(c(t)) = \frac{\dot{c}(t)}{|\dot{c}(t)|} \quad (a < t < b) .$$

Usando que, neste caso, o determinante métrico que aparece nas equações (B.3-B.5) é simplesmente igual a $|\dot{c}(t)|^2$, obtemos então

$$\int_{\gamma} d\mathbf{x} \cdot \mathbf{A} = \int_a^b dt \dot{c}(t) \cdot \mathbf{A}(c(t)) ; \quad (\text{B.16})$$

veja a Figura B.3. Assim, também é fácil verificar diretamente que quando γ é escrita como imagem de uma outra aplicação $\tilde{c} : \tilde{I} \rightarrow \mathbb{R}^n$ de classe C^r , definida num outro intervalo $\tilde{I} = [\tilde{a}, \tilde{b}]$ e compatível com a orientação de γ , esta difere da primeira tão somente por uma reparametrização, ou seja, temos $\tilde{c} = c \circ r$, onde r é uma aplicação bijetora e monótona crescente $r : \tilde{I} \rightarrow I$ de classe C^r , e portanto

$$\int_{\tilde{a}}^{\tilde{b}} d\tilde{t} \dot{\tilde{c}}(\tilde{t}) \cdot \mathbf{A}(\tilde{c}(\tilde{t})) = \int_a^b dx \dot{c}(t) \cdot \mathbf{A}(c(t)) ,$$

mostrando, mais uma vez, que a definição da integral de um campo vetorial ao longo de uma curva independe da parametrização da curva.

B.4 Exemplo: Coordenadas esféricas

Definição B.11 *Um sistema de coordenadas (U, φ, U') de \mathbb{R}^n se chama um sistema de coordenadas esféricas se*

Fig. B.3: Integração de campos vetoriais ao longo de curvas parametrizadas

- $U \subset \mathbb{R}^n \setminus \{0\}$ é um **cone** (isto é, $x \in U$ implica $\lambda x \in U$ para todo $\lambda > 0$) e é **denso** (isto é, $\bar{U} = \mathbb{R}^n$);
- $U' = V' \times \mathbb{R}^+$ onde V' é um aberto de \mathbb{R}^{n-1} ;
- a última coordenada é a coordenada radial:

$$y_n = r \equiv |x| \quad \text{onde} \quad r^2 \equiv |x|^2 = \sum_{i=1}^n x_i^2; \quad (\text{B.17})$$

- mediante a aplicação $\varphi^{-1} : U' \rightarrow U$, as coordenadas cartesianas originais x_1, \dots, x_n dependem linearmente da coordenada radial:

$$x_i = r \varphi_i^{(1)}(y_1, \dots, y_{n-1}) \quad (1 \leq i \leq n), \quad (\text{B.18})$$

sendo a aplicação $\varphi^{(1)} : V' \rightarrow U \cap S^{n-1}$ a inversa da restrição de φ à esfera unitária: $\varphi^{(1)}(y_1, \dots, y_{n-1}) = \varphi^{-1}(y_1, \dots, y_{n-1}, 1)$.

Um sistema de coordenadas esféricas de \mathbb{R}^n providencia um sistema de coordenadas para a esfera unitária

$$S^{n-1} = \{x \in \mathbb{R}^n / r^2 \equiv |x|^2 = 1\}, \quad (\text{B.19})$$

assim como para as esferas de qualquer raio a ,

$$S_a^{n-1} = \{x \in \mathbb{R}^n / r^2 \equiv |x|^2 = a^2\}. \quad (\text{B.20})$$

Reciprocamente, um sistema de coordenadas para a esfera unitária definido sobre um domínio denso de S^{n-1} pode ser estendido, de forma única, a um sistema de coordenadas esféricas de \mathbb{R}^n .

A título de exemplo, mencionamos os sistemas de coordenadas $(U_{\text{PN}}, \varphi_{\text{PN}}, U')$ e $(U_{\text{PS}}, \varphi_{\text{PS}}, U')$ obtidos por projeção estereográfica a partir do polo norte $(0, \dots, 0, +1)$ ou do polo sul $(0, \dots, 0, -1)$ da esfera unitária S^{n-1} , definidos por

$$x_i = r \frac{2y_i}{1 + |y|^2} \quad (1 \leq i \leq n-1), \quad (\text{B.21})$$

e

$$x_n = \begin{cases} -r \frac{1 - |y|^2}{1 + |y|^2} & \text{no caso da projeção estereográfica} \\ & \text{a partir do polo norte,} \\ +r \frac{1 - |y|^2}{1 + |y|^2} & \text{no caso da projeção estereográfica} \\ & \text{a partir do polo sul,} \end{cases} \quad (\text{B.22})$$

onde

$$|y|^2 = \sum_{i=1}^{n-1} y_i^2, \quad (\text{B.23})$$

e por

$$\begin{aligned} U_{\text{PN}} &= \mathbb{R}^n \setminus \{(0, \dots, 0, x_n) / x_n \geq 0\}, \\ U_{\text{PS}} &= \mathbb{R}^n \setminus \{(0, \dots, 0, x_n) / x_n \leq 0\}, \end{aligned}$$

enquanto que, em ambos os casos, $V' = \mathbb{R}^{n-1}$ e $U' = \mathbb{R}^{n-1} \times \mathbb{R}^+$.

Voltando ao caso geral, vamos deduzir algumas fórmulas para calcular integrais em coordenadas esféricas.

- Para qualquer função contínua f sobre a esfera S_a^{n-1} de raio a e qualquer subconjunto compacto K da esfera unitária S^{n-1} , $f(a \cdot)$ é uma função contínua sobre a esfera unitária S^{n-1} e aK é um subconjunto compacto da esfera S_a^{n-1} de raio a tal que vale

$$\int_{aK} d\sigma(x) f(x) = a^{n-1} \int_K d\sigma(x) f(ax). \quad (\text{B.24})$$

- Para qualquer função contínua f sobre $\mathbb{R}^n \setminus \{0\}$ e qualquer subconjunto compacto K da esfera unitária S^{n-1} , a integral de f sobre a “parte da concha entre as esferas de raio a e b gerada por K ” ($a < b$), ou seja, o subconjunto compacto $K(a, b)$ de $\mathbb{R}^n \setminus \{0\}$ definido por

$$K(a, b) = \bigcup_{a \leq r \leq b} rK$$

vale

$$\int_{K(a,b)} d^n x f(x) = \int_a^b dr r^{n-1} \int_K d\sigma(x) f(rx). \quad (\text{B.25})$$

Em particular, escolhendo $K = S^{n-1}$ e tomando o limite $a \rightarrow 0$ e $b \rightarrow \infty$ (com hipóteses adicionais adequadas para garantir que este limite existe), podemos escrever

$$\int d^n x f(x) = \int_0^\infty dr r^{n-1} \int_{S^{n-1}} d\sigma(x) f(rx). \quad (\text{B.26})$$

Para a demonstração, escolhamos um sistema de coordenadas esféricas (U, φ, U') de \mathbb{R}^n qualquer e comparamos a métrica induzida g_a^φ na esfera S_a^{n-1} de raio a com a métrica induzida g_1^φ na esfera unitária S^{n-1} , usando a equação (B.18):

$$(g_a^\varphi)_{ij} = \sum_{m=1}^n \frac{\partial(a\varphi_m^{(1)})}{\partial y_i} \frac{\partial(a\varphi_m^{(1)})}{\partial y_j} = \sum_{m=1}^n a^2 \frac{\partial\varphi_m^{(1)}}{\partial y_i} \frac{\partial\varphi_m^{(1)}}{\partial y_j} = a^2 (g_1^\varphi)_{ij}$$

$$(1 \leq i, j \leq n-1).$$

Pela definição (B.5), concluímos que, se K é um compacto contido em $U \cap S^{n-1}$, então aK é um compacto contido em $U \cap S_a^{n-1}$ e a equação (B.24) é válida. Para a demonstração da equação (B.25), introduzimos as funções matriciais

$$A = (\varphi^{-1})' = \begin{pmatrix} \frac{\partial x_1}{\partial y_1} & \cdots & \frac{\partial x_1}{\partial y_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial x_n}{\partial y_1} & \cdots & \frac{\partial x_n}{\partial y_n} \end{pmatrix}$$

e

$$B = \begin{pmatrix} B_{11}^\varphi & \cdots & B_{1n}^\varphi \\ \vdots & \ddots & \vdots \\ B_{n1}^\varphi & \cdots & B_{nn}^\varphi \end{pmatrix}$$

onde

$$B_{ij} = \frac{\partial x}{\partial y_i} \cdot \frac{\partial x}{\partial y_j} \equiv \sum_{m=1}^n \frac{\partial x_m}{\partial y_i} \frac{\partial x_m}{\partial y_j} \quad (1 \leq i, j \leq n).$$

Temos então

$$B_{ij} = \sum_{m=1}^n A_{mi} A_{mj},$$

ou seja,

$$B = A^T A.$$

Logo,

$$\det(B) = \det(A^T) \det(A) = (\det(A))^2.$$

Por outro lado, as relações (B.17) e (B.18) implicam que, para $1 \leq i \leq n-1$,

$$B_{in} = \sum_{m=1}^n \frac{\partial x_m}{\partial y_i} \frac{\partial x_m}{\partial y_n} = \sum_{m=1}^n \frac{\partial x_m}{\partial y_i} \frac{x_m}{y_n} = \frac{1}{2y_n} \frac{\partial}{\partial y_i} \left(\sum_{m=1}^n x_m^2 \right) = \frac{1}{2y_n} \frac{\partial y_n^2}{\partial y_i} = 0,$$

$$B_{nn} = \sum_{m=1}^n \left(\frac{\partial x_m}{\partial y_n} \right)^2 = \sum_{m=1}^n \frac{x_m^2}{y_n^2} = 1,$$

enquanto que, para $1 \leq i, j \leq n-1$,

$$B_{ij}(y_1, \dots, y_{n-1}, r) = (g_r^\varphi)_{ij}(y_1, \dots, y_{n-1}) = r^2 (g_1^\varphi)_{ij}(y_1, \dots, y_{n-1}).$$

Logo,

$$\det(B)(y_1, \dots, y_{n-1}, r) = r^{2n-2} \det(g_1^\varphi)(y_1, \dots, y_{n-1}) .$$

Combinando estes resultados com a fórmula (B.1) de mudança de variáveis em integrais múltiplas e a definição (B.5) da integral sobre subvariedades, concluímos que se K é um compacto contido em $U \cap S^{n-1}$, a equação (B.25) também é válida. O caso geral pode ser tratado usando a definição (B.13), em termos de uma partição da unidade, ou então notando que só existe uma única possibilidade de K não ser contido em nenhum domínio de carta de S^{n-1} , a saber quando $K = S^{n-1}$, mas neste caso, podemos escolher U tal que o complemento de $U \cap S^{n-1}$ em S^{n-1} se reduz a um único ponto, e a argumentação anterior ainda se aplica.

Como aplicação destas fórmulas gerais, calculamos o volume da esfera S_a^{n-1} e da bola aberta

$$B_a^n = \{ x \in \mathbb{R}^n / r \equiv |x| < a \} \quad (\text{B.27})$$

ou fechada

$$\bar{B}_a^n = \{ x \in \mathbb{R}^n / r \equiv |x| \leq a \} \quad (\text{B.28})$$

em \mathbb{R}^n , utilizando a função Gaussiana f_n sobre \mathbb{R}^n dada por

$$f_n(x) = \exp(-r^2) .$$

Devido à equação (B.26), temos

$$\begin{aligned} \int d^n x f_n(x) &= \int_0^\infty dr r^{n-1} \int_{S^{n-1}} d\sigma(x) f_n(rx) \\ &= \text{vol}(S^{n-1}) \int_0^\infty dr r^{n-1} \exp(-r^2) \\ &= \frac{1}{2} \text{vol}(S^{n-1}) \int_0^\infty dt t^{(n/2-1)} \exp(-t) \\ &= \frac{1}{2} \Gamma(n/2) \text{vol}(S^{n-1}) , \end{aligned}$$

onde Γ é a **função gama** de Euler, definida por

$$\Gamma(z) = \int_0^\infty dt t^{z-1} \exp(-t) \quad \text{para } z \in \mathbb{C} , \text{ Re } z > 0 . \quad (\text{B.29})$$

A propriedade fundamental desta função, facilmente demonstrada por uma simples integração por partes, é que

$$\Gamma(z+1) = z \Gamma(z) . \quad (\text{B.30})$$

Também temos

$$\Gamma(1) = 1 , \quad \Gamma(2) = 1 , \quad \Gamma(3) = 2 , \quad \dots \quad (\text{B.31})$$

e geralmente, para qualquer inteiro $n \geq 1$,

$$\Gamma(n) = (n-1)! . \quad (\text{B.32})$$

Por outro lado

$$f_n(x) = \prod_{i=1}^n \exp(-x_i^2)$$

e portanto

$$\begin{aligned} \int d^n x f_n(x) &= \left(\int dx f_1(x) \right)^n = \left(\int d^2 x f_2(x) \right)^{n/2} \\ &= \left(\frac{1}{2} \Gamma(1) \text{vol}(S^1) \right)^{n/2} = \pi^{n/2}. \end{aligned}$$

Concluimos que

$$\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}, \quad \Gamma\left(\frac{3}{2}\right) = \frac{1}{2}\sqrt{\pi}, \quad \Gamma\left(\frac{5}{2}\right) = \frac{1}{2}\frac{3}{2}\sqrt{\pi}, \quad \dots \quad (\text{B.33})$$

e geralmente, para qualquer inteiro $k \geq 1$,

$$\Gamma\left(k - \frac{1}{2}\right) = \frac{(2k-1)!!}{2^k} \sqrt{\pi}. \quad (\text{B.34})$$

Também concluimos que

$$\text{vol}(S^{n-1}) = \frac{2\pi^{n/2}}{\Gamma(n/2)}, \quad (\text{B.35})$$

ou seja,

$$\text{vol}(S^{n-1}) = \left\{ \begin{array}{ll} \frac{2\pi^k}{(k-1)!} & \text{para } n \text{ par, } n = 2k \\ \frac{2^{k+1}\pi^k}{(2k-1)!!} & \text{para } n \text{ ímpar, } n = 2k+1 \end{array} \right\}. \quad (\text{B.36})$$

Finalmente, notando que

$$\text{vol}(B_a^n) = \int_{B_a^n} d^n x = \text{vol}(S^{n-1}) \int_0^a dr r^{n-1} = \text{vol}(S^{n-1}) \frac{a^n}{n},$$

obtemos

$$\text{vol}(S_a^{n-1}) = a^{n-1} \text{vol}(S^{n-1}), \quad \text{vol}(B_a^n) = \frac{a^n}{n} \text{vol}(S^{n-1}). \quad (\text{B.37})$$

B.5 Teoremas integrais do cálculo vetorial

Lembramos, antes de mais nada, a definição do gradiente, da divergência e, em $n=3$ dimensões, do rotacional.

Definição B.12 *Sejam ϕ um campo escalar e \mathbf{A} um campo vetorial de classe C^1 sobre um aberto Ω de \mathbb{R}^n . Então o **gradiente** de ϕ é o campo vetorial $\nabla\phi$ sobre Ω definido por*

$$\nabla\phi = \left(\frac{\partial\phi}{\partial x_1}, \dots, \frac{\partial\phi}{\partial x_n} \right), \quad (\text{B.38})$$

e a **divergência** de \mathbf{A} é o campo escalar $\nabla \cdot \mathbf{A}$ sobre Ω definido por

$$\nabla \cdot \mathbf{A} = \sum_{i=1}^n \frac{\partial A_i}{\partial x_i}. \quad (\text{B.39})$$

Em $n=3$ dimensões, pode-se introduzir também o **rotacional** de \mathbf{A} como o campo vetorial $\nabla \times \mathbf{A}$ sobre Ω definido por

$$\begin{aligned} (\nabla \times \mathbf{A})_1 &= \frac{\partial A_3}{\partial x_2} - \frac{\partial A_2}{\partial x_3}, \\ (\nabla \times \mathbf{A})_2 &= \frac{\partial A_1}{\partial x_3} - \frac{\partial A_3}{\partial x_1}, \\ (\nabla \times \mathbf{A})_3 &= \frac{\partial A_2}{\partial x_1} - \frac{\partial A_1}{\partial x_2}. \end{aligned} \quad (\text{B.40})$$

Podemos, agora, formular os teoremas integrais do cálculo vetorial; veja também a Figura B.4.

Teorema B.3 (Teorema de Gauss ou teorema da divergência): *Seja Ω um domínio limitado de \mathbb{R}^n cuja fronteira $\partial\Omega = \overline{\Omega} \setminus \Omega$ é uma hipersuperfície de classe C^1 em \mathbb{R}^n , orientada de tal forma que o campo unitário \mathbf{n} normal a $\partial\Omega$ aponta para fora de Ω , e seja \mathbf{A} um campo vetorial de classe C^1 sobre $\overline{\Omega}$. Então vale a fórmula*

$$\int_{\Omega} d^n x \nabla \cdot \mathbf{A} = \int_{\partial\Omega} d\sigma \cdot \mathbf{A}. \quad (\text{B.41})$$

Teorema B.4 (Teorema de Stokes): *Seja Σ um domínio limitado de uma superfície orientada S de classe C^1 em \mathbb{R}^3 cuja fronteira $\partial\Sigma = \overline{\Sigma} \setminus \Sigma$ é uma curva de classe C^1 em \mathbb{R}^3 , orientada de tal forma que o seu sentido de circulação e o campo unitário normal a S satisfazem a regra da mão direita, e seja \mathbf{A} um campo vetorial de classe C^1 sobre $\overline{\Sigma}$. Então vale a fórmula*

$$\int_{\Sigma} d\sigma \cdot (\nabla \times \mathbf{A}) = \int_{\partial\Sigma} d\mathbf{x} \cdot \mathbf{A}. \quad (\text{B.42})$$

Observe que a definição da componente do rotacional de \mathbf{A} normal à superfície S envolve tão somente os valores de \mathbf{A} sobre S ou, mais exatamente, as primeiras derivadas de \mathbf{A} em direções tangenciais a S .

Notamos que os teoremas de Gauss e de Stokes admitem uma generalização comum, o teorema de Stokes geral, cuja formulação requer

Fig. B.4: Teoremas de Gauss e Stokes

- o conceito de formas diferenciais de grau p , como extensão do conceito de campos escalares ($p = 0$) e de campos vetoriais ($p = 1$);
- a definição da integral de uma forma diferencial de grau p sobre (subconjuntos compactos de) subvariedades p -dimensionais de \mathbb{R}^n ;
- a definição do operador d de diferenciação exterior e do operador $*$ da dualidade de Hodge que, em conjunto, permitem unificar os operadores gradiente, divergência e rotacional sob um ponto de vista comum.

Nestes termos, o teorema geral de Stokes afirma que para todo domínio limitado Σ em uma subvariedade p -dimensional orientada S de classe C^1 em \mathbb{R}^n cuja fronteira $\partial\Sigma = \bar{\Sigma} \setminus \Sigma$ é uma subvariedade $(p-1)$ -dimensional de classe C^1 de \mathbb{R}^n , adequadamente orientada, e para toda forma diferencial ω de grau $p-1$ e de classe C^1 sobre $\bar{\Sigma}$, vale a fórmula

$$\int_{\Sigma} d\omega = \int_{\partial\Sigma} \omega. \quad (\text{B.43})$$

Como não utilizaremos esta versão geral do teorema, não entraremos em maiores detalhes.

Uma aplicação imediata do teorema da divergência são as identidades de Green:

Teorema B.5 (Primeira Identidade de Green): *Seja Ω um domínio limitado em \mathbb{R}^n cuja fronteira $\partial\Omega = \bar{\Omega} \setminus \Omega$ é uma hipersuperfície de classe C^1 em \mathbb{R}^n , orientada de tal forma que o campo unitário \mathbf{n} normal a $\partial\Omega$ aponta para fora de Ω , e sejam φ um campo escalar de classe C^1 e ψ um campo escalar de classe C^2 sobre $\bar{\Omega}$. Então*

vale a fórmula

$$\int_{\Omega} d^n x (\varphi \Delta \psi + \nabla \varphi \cdot \nabla \psi) = \int_{\partial \Omega} d\sigma \cdot \varphi \nabla \psi . \quad (\text{B.44})$$

DEMONSTRAÇÃO: Aplique o teorema da divergência com $\mathbf{A} = \varphi \nabla \psi$ e use a regra de Leibniz. \square

Teorema B.6 (Segunda Identidade de Green): Seja Ω um domínio limitado em \mathbb{R}^n cuja fronteira $\partial \Omega = \bar{\Omega} \setminus \Omega$ é uma hipersuperfície de classe C^1 em \mathbb{R}^n , orientada de tal forma que o campo unitário \mathbf{n} normal a $\partial \Omega$ aponta para fora de Ω , e sejam φ e ψ dois campos escalares de classe C^2 sobre $\bar{\Omega}$. Então vale a fórmula

$$\int_{\Omega} d^n x (\varphi \Delta \psi - \psi \Delta \varphi) = \int_{\partial \Omega} d\sigma \cdot (\varphi \nabla \psi - \psi \nabla \varphi) . \quad (\text{B.45})$$

DEMONSTRAÇÃO: Antisimetrize a primeira identidade de Green em φ e ψ . \square

B.6 Médias esféricas

Na análise das propriedades de funções harmônicas assim como na solução do problema de Cauchy para a equação de ondas, a noção da média esférica de uma função desempenha um papel importante.

Definição B.13 Seja ϕ uma função contínua sobre um aberto Ω de \mathbb{R}^n . A **média esférica** de ϕ é a função M_ϕ que associa a cada ponto x de Ω e cada número positivo r com $0 < r < d(x, \partial \Omega)$ o valor médio de ϕ na esfera de raio r em torno de x , isto é,

$$M_\phi(x, r) = \frac{1}{\text{vol}(S_r(x))} \int_{S_r(x)} d\sigma(y) \phi(y) . \quad (\text{B.46})$$

Uma simples transformação de variáveis mostra que

$$M_\phi(x, r) = \frac{1}{\text{vol}(S^{n-1})} \int_{S^{n-1}} d\sigma(z) \phi(x + rz) . \quad (\text{B.47})$$

Note que a segunda fórmula implica que a expressão $M_\phi(x, r)$ está bem definida quando $-d(x, \partial \Omega) < r < d(x, \partial \Omega)$ e que M_ϕ é uma função contínua sobre o aberto

$$\Omega_M = \bigcup_{x \in \Omega} \{x\} \times]-d(x, \partial \Omega), d(x, \partial \Omega)[$$

de \mathbb{R}^{n+1} tal que

$$M_\phi(x, -r) = M_\phi(x, r) \quad , \quad M_\phi(x, 0) = \phi(x)$$

sendo que M_ϕ será de classe C^p quando ϕ for de classe C^p . Ademais, M_ϕ satisfaz uma equação diferencial interessante:

Proposição B.2 *Seja ϕ uma função de classe C^2 sobre um aberto Ω de \mathbb{R}^n . Então sobre o aberto Ω_M de \mathbb{R}^{n+1} , a média esférica M_ϕ de ϕ satisfaz a equação diferencial parcial*

$$\Delta_x M_\phi = \frac{1}{r^{n-1}} \frac{\partial}{\partial r} \left(r^{n-1} \frac{\partial}{\partial r} \right) M_\phi \equiv \left(\frac{\partial^2}{\partial r^2} + \frac{n-1}{r} \frac{\partial}{\partial r} \right) M_\phi . \quad (\text{B.48})$$

DEMONSTRAÇÃO: Basta provar a equação (B.48) para $r > 0$, pois $M_\phi(x, r)$ sendo uma função par de r , ambos os lados desta equação também são funções pares de r . Mas para $x \in \Omega$ e $0 < r < d(x, \partial\Omega)$, obtemos por diferenciação da equação (B.47)

$$\begin{aligned} \Delta_x M_\phi(x, r) &= \frac{1}{\text{vol}(S^{n-1})} \int_{S^{n-1}} d\sigma(z) \Delta\phi(x + rz) \\ &= \frac{1}{\text{vol}(S_r(x))} \int_{S_r(x)} d\sigma(y) \Delta\phi(y) , \end{aligned}$$

ou seja, $\Delta_x M_\phi = M_{\Delta\phi}$, e por outro lado,

$$\begin{aligned} \frac{\partial M_\phi}{\partial r}(x, r) &= \frac{1}{\text{vol}(S^{n-1})} \int_{S^{n-1}} d\sigma(z) \cdot \nabla\phi(x + rz) \\ &= \frac{1}{\text{vol}(S_r(x))} \int_{S_r(x)} d\sigma(y) \cdot \nabla\phi(y) . \end{aligned}$$

Aplicando a segunda identidade de Green (B.45) à bola $B_r(x)$, com $\varphi = 1$ e $\psi = \phi$, e usando a fórmula (B.37), obtemos

$$r^{n-1} \frac{\partial M_\phi}{\partial r}(x, r) = \frac{1}{\text{vol}(S^{n-1})} \int_{B_r(x)} d^n y \Delta\phi(y) . \quad (\text{B.49})$$

Usando a fórmula (B.25), podemos reescrever esta equação na forma

$$r^{n-1} \frac{\partial M_\phi}{\partial r}(x, r) = \frac{1}{\text{vol}(S^{n-1})} \int_0^r d\rho \rho^{n-1} \int_{S^{n-1}} d\sigma(z) \Delta\phi(x + \rho z) ,$$

ou seja,

$$r^{n-1} \frac{\partial M_\phi}{\partial r}(x, r) = \int_0^r d\rho \rho^{n-1} M_{\Delta\phi}(x, \rho) .$$

A afirmação segue diferenciando novamente em relação a r . \square

C Equações diferenciais parciais clássicas na teoria dos campos

Neste apêndice, mostraremos como as três equações diferenciais parciais clássicas surgem na física: as equações de Laplace e de Poisson assim como a equação de ondas no eletromagnetismo e a equação de difusão na dinâmica dos fluidos.

C.1 Eletrodinâmica: equações de Maxwell, equação de Poisson e equação de ondas

Áreas importantes da física clássica onde as equações de Laplace e de Poisson ocupam uma posição central na solução de qualquer problema são a eletrostática e a magnetostática. Outra área, igualmente importante, onde a equação de ondas desempenha um papel semelhante é a teoria das ondas eletromagnéticas, sua geração e propagação.

De maneira geral, a tarefa principal da eletrodinâmica é calcular o **campo elétrico** \mathbf{E} e o **campo magnético** \mathbf{B} a partir da **densidade de carga (elétrica)** ρ e da **densidade de corrente (de carga elétrica)** \mathbf{j} , que são dadas. As leis que determinam como cargas e correntes produzem campos eletromagnéticos constituem-se na forma de um sistema de equações diferenciais parciais de primeira ordem para os campos, no qual as cargas e correntes aparecem como fontes (isto é, como termos não-homogêneos) – um sistema conhecido como as **equações de Maxwell**. Estas equações (no vácuo) são:

- a lei de Gauss:

$$\nabla \cdot \mathbf{E} = \frac{\rho}{\epsilon_0}, \quad (\text{C.1-a})$$

- a lei de Faraday ou lei da indução:

$$\nabla \times \mathbf{E} = -\kappa \frac{\partial \mathbf{B}}{\partial t}, \quad (\text{C.1-b})$$

- a lei da ausência de cargas magnéticas:

$$\nabla \cdot \mathbf{B} = 0, \quad (\text{C.1-c})$$

- a lei de Ampère-Maxwell:

$$\nabla \times \mathbf{B} = \kappa\mu_0 \left(\mathbf{j} + \epsilon_0 \frac{\partial \mathbf{E}}{\partial t} \right), \quad (\text{C.1-d})$$

onde ϵ_0 , μ_0 e κ são constantes satisfazendo a condição

$$\kappa^2 \epsilon_0 \mu_0 = 1/c^2, \quad (\text{C.2})$$

e c é a velocidade da luz (no vácuo). O valor atribuído à constante κ depende do sistema de unidades utilizado. Existem aqui duas principais alternativas: nos sistemas assimétricos, como o SI (sistema internacional), temos $\kappa=1$, enquanto que nos sistemas simétricos, como os sistemas de Gauss e de Heaviside, temos $\kappa = 1/c$.

Inicialmente, suponhamos que todos os campos – as densidades de carga e de corrente assim como o campo elétrico e o campo magnético – não dependem explicitamente do tempo t . Neste caso, as equações de Maxwell dividem-se em dois sistemas separados de equações: as equações básicas da eletrostática,

$$\nabla \cdot \mathbf{E} = \frac{\rho}{\epsilon_0}, \quad (\text{C.3-a})$$

$$\nabla \times \mathbf{E} = 0, \quad (\text{C.3-b})$$

e as equações básicas da magnetostática,

$$\nabla \cdot \mathbf{B} = 0, \quad (\text{C.4-a})$$

$$\nabla \times \mathbf{B} = \kappa\mu_0 \mathbf{j}. \quad (\text{C.4-b})$$

A última destas equações é conhecida como **lei de Ampère**. Em particular, é na situação estática – e apenas nela – que a eletrodinâmica pode ser separada em eletricidade e magnetismo.

Na eletrostática, o rotacional do campo elétrico é zero, o que implica a existência de um campo escalar ϕ , chamado *potencial escalar* ou *potencial eletrostático*, tal que o campo elétrico seja o seu gradiente (a menos de um sinal, que é meramente uma questão de conveniência, mas tradicional):

$$\mathbf{E} = -\nabla\phi. \quad (\text{C.5})$$

Observamos que o potencial escalar é univocamente determinado a menos de uma constante aditiva. Notamos também que o teorema utilizado aqui (o lema de Poincaré) – afirmando que um campo vetorial com rotacional igual a zero é o gradiente de um campo escalar – é válido apenas quando se impõe a hipótese de

que o domínio de definição dos campos considerado (sobre o qual o campo vetorial dado deve ser, digamos, de classe C^1 , para que o campo escalar a ser construído seja, pelo menos, de classe C^2) tenha cohomologia trivial em dimensão 1. (Isto é o caso, por exemplo, para domínios contraíveis, ou mais geralmente, domínios simplesmente conexos.) Substituindo a equação (C.5) na outra equação básica da eletrostática, a lei de Gauss (C.3-a), podemos concluir que ϕ deve satisfazer a equação de Poisson:

$$\Delta\phi = -\frac{\rho}{\epsilon_0}. \quad (\text{C.6})$$

Na magnetostática, a divergência do campo magnético é zero, o que implica a existência de um campo vetorial \mathbf{A} , chamado *potencial vetorial* ou *potencial magnetostático*, tal que o campo magnético seja o seu rotacional:

$$\mathbf{B} = \nabla \times \mathbf{A}. \quad (\text{C.7})$$

Observamos que o potencial vetorial não é completamente determinado por esta condição, pois podemos lhe adicionar o gradiente de um campo escalar qualquer, sem modificar o campo magnético. Podemos usufruir da liberdade de efetuar uma tal *transformação de calibre* para submeter o potencial vetorial a uma *condição de calibre* ou *escolha de calibre*, isto é, a um vínculo adicional, como por exemplo o *calibre de Coulomb*

$$\nabla \cdot \mathbf{A} = 0. \quad (\text{C.8})$$

Notamos também que o teorema utilizado aqui (o lema de Poincaré) – afirmando que um campo vetorial com divergência igual a zero é o rotacional de um outro campo vetorial – é válido apenas quando se impõe a hipótese de que o domínio de definição dos campos considerado (sobre o qual o campo vetorial dado deve ser, digamos, de classe C^1 , para que o campo vetorial a ser construído seja, pelo menos, de classe C^2) tenha cohomologia trivial em dimensão 2. (Isto é o caso, por exemplo, para domínios contraíveis.) Substituindo as equações (C.7) e (C.8) na outra equação básica da magnetostática, a lei de Ampère (C.4-b), podemos concluir que \mathbf{A} deve satisfazer uma versão vetorial da equação de Poisson:

$$\Delta\mathbf{A} = -\kappa\mu_0 \mathbf{j}. \quad (\text{C.9})$$

O mesmo tipo de argumentação pode ser utilizado para analisar as equações de Maxwell em geral: Inicialmente, resolvemos a equação (C.1-c) como antes, introduzindo o **potencial vetorial** \mathbf{A} , que satisfaz

$$\mathbf{B} = \nabla \times \mathbf{A}. \quad (\text{C.10})$$

Substituindo esta equação na lei de Faraday (C.1-b), concluímos que, ao invés do próprio campo elétrico \mathbf{E} , como na eletrostática, é a combinação $\mathbf{E} + \kappa \partial\mathbf{A}/\partial t$ que tem rotacional nulo e que portanto pode ser escrita (a menos de um sinal) como o gradiente de um campo escalar. Assim, chegamos a introduzir o **potencial escalar** ϕ , que satisfaz

$$\mathbf{E} = -\nabla\phi - \kappa \frac{\partial\mathbf{A}}{\partial t}. \quad (\text{C.11})$$

Mais uma vez, os potenciais \mathbf{A} e ϕ estão longes de serem univocamente determinados pelos campos \mathbf{E} e \mathbf{B} . De fato, outros potenciais \mathbf{A}' e ϕ' geram os mesmos campos físicos \mathbf{E} e \mathbf{B} se e somente se podem ser obtidos dos potenciais originais \mathbf{A} e ϕ através de uma **transformação de calibre**

$$\mathbf{A} \longrightarrow \mathbf{A}' = \mathbf{A} + \nabla\chi \quad , \quad \phi \longrightarrow \phi' = \phi - \kappa \frac{\partial\chi}{\partial t} \quad , \quad (\text{C.12})$$

onde χ é uma função arbitrária de t e \mathbf{x} . Para remover essa ambiguidade, costuma-se impor uma **condição de calibre** ou efetuar uma **escolha de calibre**: isso significa que os potenciais são submetidos a condições suplementares apropriadas. Geralmente, contudo, tal procedimento restringe mas não elimina a referida ambiguidade, o que motiva a definição de uma **transformação de calibre residual** como uma transformação de calibre que é compatível com uma determinada escolha de calibre. As duas escolhas de calibre mais importantes na prática são:

- **Calibre de Coulomb:**

$$\nabla \cdot \mathbf{A} = 0 \quad . \quad (\text{C.13})$$

As transformações de calibre residuais são as transformações de calibre (C.12) sujeitas ao vínculo

$$\Delta\chi = 0 \quad . \quad (\text{C.14})$$

- **Calibre de Lorentz:**

$$\nabla \cdot \mathbf{A} + \frac{1}{\kappa c^2} \frac{\partial\phi}{\partial t} = 0 \quad . \quad (\text{C.15})$$

As transformações de calibre residuais são as transformações de calibre (C.12) sujeitas ao vínculo

$$\square\chi \equiv \frac{1}{c^2} \frac{\partial^2\chi}{\partial t^2} - \Delta\chi = 0 \quad . \quad (\text{C.16})$$

Agora, podemos substituir as equações (C.10) e (C.11) nas equações de Maxwell não-homogêneas (C.1-a) e (C.1-d), obtendo, após um cálculo elementar,

$$\Delta\phi + \kappa \nabla \cdot \frac{\partial\mathbf{A}}{\partial t} = -\frac{\rho}{\epsilon_0} \quad , \quad (\text{C.17})$$

e

$$\Delta\mathbf{A} - \nabla(\nabla \cdot \mathbf{A}) - \frac{1}{c^2} \frac{\partial^2\mathbf{A}}{\partial t^2} - \frac{1}{\kappa c^2} \nabla \left(\frac{\partial\phi}{\partial t} \right) = -\kappa\mu_0 \mathbf{j} \quad , \quad (\text{C.18})$$

onde utilizamos a relação (C.2) e a identidade

$$\nabla \times (\nabla \times \mathbf{A}) = \nabla(\nabla \cdot \mathbf{A}) - \Delta\mathbf{A} \quad .$$

No calibre de Coulomb, estas equações se reduzem a

$$\Delta\phi = -\frac{\rho}{\epsilon_0} \quad (\text{C.19})$$

e

$$\square \mathbf{A} \equiv \frac{1}{c^2} \frac{\partial^2 \mathbf{A}}{\partial t^2} - \Delta \mathbf{A} = \kappa \mu_0 \mathbf{j}_t, \quad (\text{C.20})$$

onde

$$\mathbf{j}_t = \mathbf{j} - \epsilon_0 \nabla \left(\frac{\partial \phi}{\partial t} \right) \quad (\text{C.21})$$

é a parte *transversal* da corrente, isto é, sua parte com divergência nula:

$$\nabla \cdot \mathbf{j}_t = 0. \quad (\text{C.22})$$

(Isso se comprova, por exemplo, calculando a divergência da equação (C.20).) A equação (C.19) é idêntica à equação análoga da eletrostática: o tempo t só aparece como um parâmetro adicional. Portanto, neste caso, o potencial escalar ϕ é denominado **potencial escalar instantâneo**. O calibre de Coulomb se mostra particularmente útil quando $\rho \equiv 0$, pois, neste caso, pode-se por $\phi \equiv 0$.

No calibre de Lorentz, contudo, obtemos

$$\square \phi \equiv \frac{1}{c^2} \frac{\partial^2 \phi}{\partial t^2} - \Delta \phi = \frac{\rho}{\epsilon_0}, \quad (\text{C.23})$$

$$\square \mathbf{A} \equiv \frac{1}{c^2} \frac{\partial^2 \mathbf{A}}{\partial t^2} - \Delta \mathbf{A} = \kappa \mu_0 \mathbf{j}. \quad (\text{C.24})$$

Portanto, em termos dos potenciais no calibre de Lorentz, as equações não-homogêneas de Maxwell assumem uma forma particularmente simétrica, pois tornam-se um sistema de equações de onda, com a densidade de carga e de corrente como fontes.

As equações de Maxwell acima apresentadas são válidas no vácuo, isto é, na ausência de matéria. Porém, isto é uma situação pouco realística, pois sempre há alguma forma de matéria presente em alguma região do espaço. Em princípio, é claro, a influência desta matéria pode ser considerada como uma contribuição às densidades de carga e de corrente que entram nas equações de Maxwell – só que esta contribuição provém de uma distribuição complicadíssima de cargas e correntes¹ que não conhecemos e nem queremos conhecer em detalhe. Portanto, esta seria, na prática, uma maneira inviável de abordar o problema; precisamos procurar outros métodos para calcular uma **interação efetiva** entre as cargas e correntes presentes na matéria e os campos eletromagnéticos.

O problema da interação efetiva entre as cargas e correntes presentes na matéria e os campos eletromagnéticos se impõe particularmente quando as cargas estão – pelo menos parcialmente – livremente móveis (o que é uma condição prévia para a existência de correntes), ou seja, na presença de condutores elétricos. Neste caso,

¹Num pedaço de metal macroscópico, os elétrons livres formam um gás com pelo menos umas 10^{20} partículas em movimentação caótica.

a estratégia para resolver o problema é de simplesmente excluir o(s) volume(s) ocupado(s) por estes materiais da região em que deve-se calcular os campos e de impor *condições de fronteira* sobre os campos na região assim definida, ou seja, na superfície dos materiais: é através dessas condições de fronteira que se codifica a supracitada interação efetiva. Enfrentamos portanto a tarefa de deduzir quais deveriam ser as condições de fronteira adequadas para a eletrodinâmica.

Para começar, consideramos a situação idealizada de uma superfície S carregando uma densidade superficial ω de carga e uma densidade superficial \mathbf{k} de corrente. Localmente, pelo menos, a superfície S separa o espaço em duas regiões V_1 e V_2 . Suponhamos também que os campos \mathbf{E} e \mathbf{B} e as suas derivadas parciais em relação ao tempo sejam de classe C^∞ em V_1 e V_2 e apresentem, no máximo, descontinuidades na superfície separadora S . Então, valem as seguintes condições de contato:

$$\mathbf{n}_{12} \cdot (\mathbf{E}_1 - \mathbf{E}_2) = \frac{\omega}{\epsilon_0} \quad , \quad \mathbf{n}_{12} \times (\mathbf{E}_1 - \mathbf{E}_2) = 0 \quad , \quad (\text{C.25})$$

$$\mathbf{n}_{12} \cdot (\mathbf{B}_1 - \mathbf{B}_2) = 0 \quad , \quad \mathbf{n}_{12} \times (\mathbf{B}_1 - \mathbf{B}_2) = \kappa\mu_0 \mathbf{k} \quad , \quad (\text{C.26})$$

onde \mathbf{E}_i e \mathbf{B}_i são os valores de \mathbf{E} e de \mathbf{B} na superfície S obtidos tomando o limite provindo do interior de V_i ($i = 1, 2$), enquanto que \mathbf{n}_{12} é o vetor normal à superfície, direcionado de V_2 a V_1 .

Em outras palavras, quando atravessamos a superfície separadora, passando de uma região para a outra, a componente normal do campo elétrico e a componente tangencial do campo magnético apresentam descontinuidades, sendo a primeira proporcional à densidade de carga superficial e a segunda, à densidade de corrente superficial, enquanto que a componente tangencial do campo elétrico e a componente normal do campo magnético são contínuas.

Para provar estas afirmações, utilizamos as equações de Maxwell e os teoremas de Gauss e Stokes. De fato, as equações de contato para as componentes normais,

$$\mathbf{E}_1^\perp - \mathbf{E}_2^\perp = \mathbf{n}_{12} \cdot (\mathbf{E}_1 - \mathbf{E}_2) \quad \text{e} \quad \mathbf{B}_1^\perp - \mathbf{B}_2^\perp = \mathbf{n}_{12} \cdot (\mathbf{B}_1 - \mathbf{B}_2)$$

resultam das equações de Maxwell (C.1-a) e (C.1-c) por integração sobre a superfície do pequeno volume \tilde{V} na figura abaixo (o adjetivo “pequeno” se refere à extensão d de \tilde{V} na direção da normal \mathbf{n}_{12}):

$$\begin{aligned} \int_{S \cap \tilde{V}} d\sigma \mathbf{n}_{12} \cdot (\mathbf{E}_1 - \mathbf{E}_2) &= \lim_{d \rightarrow 0} \int_{\partial \tilde{V}} d\sigma \cdot \mathbf{E} = \lim_{d \rightarrow 0} \int_{\tilde{V}} d^3x \nabla \cdot \mathbf{E} \\ &= \frac{1}{\epsilon_0} \lim_{d \rightarrow 0} \int_{\tilde{V}} d^3x \rho = \frac{1}{\epsilon_0} \int_{S \cap \tilde{V}} d\sigma \omega \quad , \\ \int_{S \cap \tilde{V}} d\sigma \mathbf{n}_{12} \cdot (\mathbf{B}_1 - \mathbf{B}_2) &= \lim_{d \rightarrow 0} \int_{\partial \tilde{V}} d\sigma \cdot \mathbf{B} = \lim_{d \rightarrow 0} \int_{\tilde{V}} d^3x \nabla \cdot \mathbf{B} = 0 \quad . \end{aligned}$$

Fig. C.1: Cálculo da descontinuidade na componente normal do campo elétrico e magnético, em uma superfície separando duas regiões, coberta por uma densidade de carga e de corrente dada, utilizando o teorema de Gauss: veja texto

De maneira análoga, as equações de contato para as componentes tangenciais,

$$\mathbf{E}_1^t - \mathbf{E}_2^t = \mathbf{t} \cdot (\mathbf{E}_1 - \mathbf{E}_2) \quad \text{e} \quad \mathbf{B}_1^t - \mathbf{B}_2^t = \mathbf{t} \cdot (\mathbf{B}_1 - \mathbf{B}_2)$$

(onde \mathbf{t} percorre os vetores tangentes à superfície S) resultam das equações de Maxwell (C.1-b) e (C.1-d) por integração sobre o bordo da pequena superfície $\tilde{\Sigma}$ na figura abaixo (o adjetivo “pequeno” se refere à extensão d de $\tilde{\Sigma}$ na direção da normal \mathbf{n}_{12}):

$$\begin{aligned} \int_{S \cap \tilde{\Sigma}} dx (\mathbf{t} \times \mathbf{n}_{12}) \cdot (\mathbf{E}_1 - \mathbf{E}_2) &= \lim_{d \rightarrow 0} \int_{\partial \tilde{\Sigma}} d\mathbf{x} \cdot \mathbf{E} = \lim_{d \rightarrow 0} \int_{\tilde{\Sigma}} d\boldsymbol{\sigma} \cdot (\nabla \times \mathbf{E}) \\ &= -\kappa \lim_{d \rightarrow 0} \int_{\tilde{\Sigma}} d\boldsymbol{\sigma} \cdot \frac{\partial \mathbf{B}}{\partial t} = 0, \\ \int_{S \cap \tilde{\Sigma}} dx (\mathbf{t} \times \mathbf{n}_{12}) \cdot (\mathbf{B}_1 - \mathbf{B}_2) &= \lim_{d \rightarrow 0} \int_{\partial \tilde{\Sigma}} d\mathbf{x} \cdot \mathbf{B} = \lim_{d \rightarrow 0} \int_{\tilde{\Sigma}} d\boldsymbol{\sigma} \cdot (\nabla \times \mathbf{B}) \\ &= \kappa \mu_0 \lim_{d \rightarrow 0} \int_{\tilde{\Sigma}} d\boldsymbol{\sigma} \cdot \left(\mathbf{j} + \epsilon_0 \frac{\partial \mathbf{E}}{\partial t} \right) = \kappa \mu_0 \int_{S \cap \tilde{\Sigma}} dx \mathbf{k} \cdot \mathbf{t}. \end{aligned}$$

Podemos agora formular as condições de fronteira que enfrentamos na eletrodinâmica na presença de condutores.

Num condutor elétrico, as cargas eletrônicas (ou mais exatamente, uma parte das cargas eletrônicas) são livremente móveis relativamente aos íons positivos. Portanto, no interior de um condutor, uma voltagem (diferença de potencial) externa será quase imediatamente compensada por uma redistribuição destas cargas, e assim sucumbirá quase instantaneamente. Portanto, no interior de um condutor, temos $\mathbf{E} \equiv 0$ e, devido à lei de Gauss (C.1-a), também temos $\rho \equiv 0$, ou seja, as cargas eletrônicas (ou mais exatamente, um excesso ou uma falta de cargas eletrônicas, relativamente ao valor médio) só podem ser localizadas na sua superfície. A distribuição dessas cargas sobre a superfície será devidamente descrita por uma densidade de carga superficial ω , e qualquer reajuste na distribuição

Fig. C.2: Cálculo da descontinuidade na componente tangencial do campo elétrico e magnético, em uma superfície separando duas regiões, coberta por uma densidade de carga e de corrente dada, utilizando o teorema de Stokes: veja texto

de cargas que se torne necessário para manter $\mathbf{E} \equiv 0$ no interior do condutor se passará ao longo da sua superfície, ou seja, será um reajuste de ω .

Num condutor elétrico *perfeito* ou *ideal*, isto é, sem resistência de Ohm, encontramos uma situação análoga com respeito ao campo magnético. No seu interior, qualquer campo magnético externo será quase imediatamente compensado por uma redistribuição de correntes eletrônicas, e assim sucumbirá quase instantaneamente. Portanto, no interior de um condutor perfeito, temos (além de $\mathbf{E} \equiv 0$) $\mathbf{B} \equiv 0$ e, devido à lei de Ampère (C.1-d), também temos (além de $\rho \equiv 0$) $\mathbf{j} \equiv 0$, pelo menos na magnetostática (e numa boa aproximação para campos com dependência lenta do tempo), ou seja, correntes só podem ser localizadas na sua superfície. A distribuição dessas correntes sobre a superfície será devidamente descrita por uma densidade de corrente superficial \mathbf{k} , e qualquer reajuste na distribuição de correntes que se torne necessário para manter $\mathbf{B} \equiv 0$ no interior do condutor se passará ao longo da sua superfície, ou seja, será um reajuste de \mathbf{k} .

Existem condutores perfeitos na prática: os supercondutores, nos quais uma corrente pode correr durante anos, sem atenuação perceptível. Neles, a supressão de campos magnéticos adquire o nível de uma equação de estado para a fase supercondutora, como pode ser demonstrado, de maneira impressionante, pelo efeito de Meissner-Ochsenfeld.

Combinando estes fatos com as condições de contato (C.25) e (C.26), podemos concluir que na superfície de um condutor, a componente tangencial do campo elétrico, sendo contínua, se anula, enquanto que a componente normal é dada pela densidade de carga superficial:

$$\mathbf{E}_{\text{int}} = 0, \quad (\text{C.27})$$

e

$$\mathbf{n} \cdot \mathbf{E}_{\text{ext}} = \frac{\omega}{\epsilon_0}, \quad (\text{C.28})$$

$$\mathbf{n} \times \mathbf{E}_{\text{ext}} = 0, \quad (\text{C.29})$$

Além disto, na superfície de um condutor perfeito, a componente normal do campo magnético, sendo contínua, se anula, enquanto que a componente tangencial é dada

pela densidade de corrente superficial:

$$\mathbf{B}_{\text{int}} = 0, \quad (\text{C.30})$$

e

$$\mathbf{n} \cdot \mathbf{B}_{\text{ext}} = 0, \quad (\text{C.31})$$

$$\mathbf{n} \times \mathbf{B}_{\text{ext}} = \kappa \mu_0 \mathbf{k}, \quad (\text{C.32})$$

Nestas fórmulas, \mathbf{n} é o vetor normal à superfície, direcionado para fora do condutor.

Isso posto, podemos formular os dois problemas básicos da eletrostática na presença de condutores.

Consideramos como dados um sistema de condutores C_i e uma densidade de carga ρ no espaço exterior aos condutores, i.e., na região $\Omega = \mathbb{R}^3 \setminus (C_1 \cup \dots \cup C_n)$. Também sejam dados

1. os valores Q_1, \dots, Q_n da carga total sobre cada condutor (condutores isolados); ou
2. os valores V_1, \dots, V_n do potencial sobre cada condutor (condutores aterrados).

A tarefa, nos dois casos, é calcular o potencial eletrostático ϕ em Ω , como solução da equação de Poisson (C.6), sujeito às condições de fronteira apropriadas. No segundo caso, identificando o bordo $\partial\Omega$ de Ω com a união disjunta das superfícies ∂C_i (porém com orientação oposta), essa condição de fronteira é

$$\phi|_{\partial C_i} = V_i. \quad (\text{C.33})$$

Trata-se de um caso particular do *problema de Dirichlet*, já formulado no Capítulo 1.7. Quanto ao primeiro caso, vamos simplesmente mencionar aqui que ele pode ser reduzido ao segundo, mediante a solução de um sistema de n equações lineares (inversão de uma matriz).

C.2 Hidrodinâmica: equações de balanço e equação de difusão

A hidrodinâmica é a área da física que trata do comportamento dos fluidos. (A noção de fluido é um conceito geral que abrange líquidos e gases.) É uma teoria de campos particularmente simples e plástica, na qual o campo fundamental é o **campo de escoamento** \mathbf{v} , cujo valor $\mathbf{v}(t, \mathbf{x})$ no instante t e no local \mathbf{x} é a velocidade de fluxo do fluido neste instante e neste local. As curvas integrais deste campo vetorial, isto é, as soluções da equação diferencial ordinária

$$\frac{d\mathbf{x}}{d\tau}(\tau) = \mathbf{v}(t, \mathbf{x}(\tau)),$$

chamam-se as **linhas de fluxo** do fluido. (Em terminologia matemática, são as curvas integrais do campo \mathbf{v} congelado no instante t .) Quando \mathbf{v} não depende explicitamente do tempo, são idênticas às trajetórias das partículas do fluido.

Além do campo vetorial \mathbf{v} , existem na hidrodinâmica outros campos, como a densidade de massa ρ^m , a pressão p , a temperatura T ou a concentração c^S de alguma substância S dissolvida no fluido, por exemplo. A tarefa da hidrodinâmica é estabelecer equações diferenciais parciais que descrevam corretamente a evolução temporal de todos esses campos a partir de condições iniciais dadas.

O método padrão para deduzir tais equações diferenciais parciais, entre elas as famosas equações de Navier-Stokes e a equação de condução de calor ou de difusão, é estabelecer e analisar equações de balanço, particularmente equações de continuidade para quantidades conservadas.

Um exemplo típico de uma quantidade conservada é a massa. Para formular a lei de conservação da massa precisamos introduzir dois campos: a **densidade de massa** ρ^m e a **densidade de corrente de massa** \mathbf{j}^m , com a seguinte interpretação física:

- ρ^m é um campo escalar tal que para cada volume V , a massa $m_V(t)$ contida no volume V no instante t é

$$m_V(t) = \int_V d^3x \rho^m(t, \mathbf{x}) .$$

- \mathbf{j}^m é um campo vetorial tal que para cada superfície S , a massa $m_S(t_2, t_1)$ que passa por S entre os instantes t_1 e t_2 ($t_1 < t_2$) é

$$m_S(t_2, t_1) = \int_{t_1}^{t_2} dt \mu_S^m(t) ,$$

onde

$$\mu_S^m(t) = \int_S d\boldsymbol{\sigma}(\mathbf{x}) \cdot \mathbf{j}^m(t, \mathbf{x})$$

é a taxa de fluxo de massa através de S no instante t .

A lei de conservação da massa afirma então que a massa contida num volume V só pode mudar quando há uma corrente de massa passando pela sua superfície ∂V : temos

$$\frac{d}{dt} m_V(t) = -\mu_{\partial V}^m(t) , \quad (\text{C.34})$$

ou seja,

$$\frac{d}{dt} \int_V d^3x \rho^m(t, \mathbf{x}) = - \int_{\partial V} d\boldsymbol{\sigma}(\mathbf{x}) \cdot \mathbf{j}^m(t, \mathbf{x}) , \quad (\text{C.35})$$

ou ainda, utilizando o teorema de Gauss,

$$\int_V d^3x \frac{\partial \rho^m}{\partial t}(t, \mathbf{x}) = - \int_V d^3x (\boldsymbol{\nabla} \cdot \mathbf{j}^m)(t, \mathbf{x}) .$$

É importante observar que esta afirmação vale para qualquer volume V : a massa é uma quantidade *localmente* conservada. Este fato nos permite transformar a

fomulação integral (C.34) ou (C.35) em uma formulação diferencial:

$$\frac{\partial \rho^m}{\partial t} + \nabla \cdot \mathbf{j}^m = 0. \quad (\text{C.36})$$

Esta **equação de continuidade** é a expressão matemática da **lei de conservação (local) da massa**.

Um outro exemplo importante de uma quantidade conservada é a carga elétrica. De fato, temos na eletrodinâmica a mesma **equação de continuidade** relacionando a densidade de carga (elétrica) ρ e a densidade de corrente (elétrica) \mathbf{j} , expressando a **lei de conservação (local) da carga (elétrica)**:

$$\frac{\partial \rho}{\partial t} + \nabla \cdot \mathbf{j} = 0. \quad (\text{C.37})$$

Observa-se que esta equação é uma consequência imediata das equações de Maxwell (C.1-a) e (C.1-d):

$$\frac{\partial \rho}{\partial t} + \nabla \cdot \mathbf{j} = \epsilon_0 \frac{\partial}{\partial t} (\nabla \cdot \mathbf{E}) - \epsilon_0 \nabla \cdot \frac{\partial \mathbf{E}}{\partial t} + \frac{1}{\kappa \mu_0} \nabla \cdot (\nabla \times \mathbf{B}) = 0.$$

Em outras palavras, a equação de continuidade (C.37) constitui uma condição de consistência para as equações de Maxwell. Historicamente, foi através de uma análise das possibilidades de sanar a inconsistência de algumas das equações estabelecidas do eletromagnetismo, ou seja, a lei de Gauss

$$\nabla \cdot \mathbf{E} = \frac{\rho}{\epsilon_0}$$

em conjunto com a lei de Ampère

$$\nabla \times \mathbf{B} = \kappa \mu_0 \mathbf{j}$$

e a lei de conservação da carga, que Maxwell chegou à proposta de modificar a lei de Ampère através do seu famoso termo adicional, proporcional à derivada temporal do campo elétrico,

$$\nabla \times \mathbf{B} = \kappa \mu_0 \left(\mathbf{j} + \epsilon_0 \frac{\partial \mathbf{E}}{\partial t} \right),$$

chegando assim à formulação definitiva das leis fundamentais da eletrodinâmica.

De forma análoga à massa e à carga elétrica, podem-se formular equações de continuidade ou, no mínimo, equações de balanço para todas as grandezas extensivas da física. Para explicar esta afirmação, damos algumas definições:

Definição C.1 *As grandezas físicas dividem-se em duas categorias: grandezas extensivas, ou “quantidades”, e grandezas intensivas, ou “qualidades”:*

- **Grandezas extensivas** são grandezas associadas a sistemas que se duplicam, triplicam, ... conforme duplicamos, triplicamos, ... o sistema.

Exemplos de grandezas extensivas são massa, carga, energia, momento, momento angular, entropia, volume, número de partículas (de uma certa espécie), ...

- **Grandezas intensivas** são grandezas associadas a sistemas para as quais isto não ocorre.

Exemplos de grandezas intensivas são temperatura, pressão, potenciais tais como o potencial escalar e o potencial vetorial da eletrodinâmica ou os potenciais químicos,

A cada grandeza extensiva, rotulada por um índice a , digamos, podemos associar três campos:

- a sua **densidade**, um campo escalar ρ^a ,
- a sua **densidade de corrente**, um campo vetorial \mathbf{j}^a ,
- a sua **densidade de produção**, um campo escalar q^a .

Então, a **equação de balanço** para a quantidade a é

$$\frac{\partial \rho^a}{\partial t} + \nabla \cdot \mathbf{j}^a = q^a . \quad (\text{C.38})$$

A existência de fontes ($q^a > 0$) e/ou ralos ($q^a < 0$) geralmente indica que a quantidade a não é conservada ou, ainda, que o sistema sob consideração não é fechado.

Definição C.2 *Um sistema fechado é um sistema físico que, na sua interação com outros sistemas físicos, não troca nenhuma grandeza extensiva com eles.*

Definição C.3 *Uma quantidade conservada é uma grandeza extensiva tal que, em sistemas fechados, a densidade de produção correspondente é nula.*

Exemplos clássicos e bem conhecidos de quantidades conservadas são massa, carga, energia, momento linear e momento angular, enquanto que o exemplo padrão de uma grandeza extensiva não-conservada é a entropia S . Para ela, vale uma “semi-conservação”: isto é o conteúdo do famoso **segundo teorema fundamental da termodinâmica**: A densidade de produção da entropia em sistemas fechados é sempre não-negativa: $q^S \geq 0$.

Para melhor entender o motivo da restrição a sistemas fechados feita nestas afirmações, é útil observar que a troca de uma grandeza extensiva a entre dois sistemas físicos 1 e 2 pode contribuir não apenas às densidades de corrente, \mathbf{j}_1^a e \mathbf{j}_2^a , mas também às densidades de produção, q_1^a e q_2^a , de ambos os sistemas. A primeira contribuição corresponde ao caso onde os dois sistemas estão espacialmente separados e a troca se realiza por fluxo através da interface entre eles, enquanto que o segundo caso pode se realizar quando os dois sistemas se sobrepõem no espaço.

Considere, por exemplo, a troca de energia e momento entre um sistema de cargas elétricas em movimento (1) e o campo eletromagnético (2) que estas cargas geram e ao qual estão sujeitas. A energia e o momento só das partículas ou só do campo eletromagnético não são conservadas. O que é conservado, isso sim, são a energia e o momento do sistema total (1+2), que é um sistema fechado.

Este exemplo mostra que, em sistemas abertos, mesmo uma quantidade conservada pode apresentar uma densidade de produção não nula, mas as fontes e os ralos correspondentes sempre podem ser identificados com os ralos e as fontes da mesma quantidade associada a um outro sistema aberto que se encontra em interação com o primeiro.

Além das equações de balanço para as grandezas extensivas, há um outro ingrediente que é de suma importância na derivação das equações diferenciais parciais da hidrodinâmica: são as relações constitutivas que regem a interdependência entre as diversas grandezas físicas que aparecem na teoria.

Um exemplo elementar de uma tal relação constitutiva é a relação entre a densidade de massa, a densidade de corrente de massa e o campo de escoamento:

$$\mathbf{j}^m = \rho^m \mathbf{v} . \quad (\text{C.39})$$

Para outras grandezas extensivas a , a relação entre a densidade, a densidade de corrente e o campo de escoamento é mais complicada, pois o transporte da grandeza a pode-se realizar através de dois processos distintos: *convecção* e *condução*. Portanto, a densidade de corrente se decompõe em duas contribuições,

$$\mathbf{j}^a = \mathbf{j}_{\text{conv}}^a + \mathbf{j}_{\text{cond}}^a , \quad (\text{C.40})$$

onde a **parte convectiva** $\mathbf{j}_{\text{conv}}^a$ descreve o transporte da grandeza a devido ao escoamento do fluido,

$$\mathbf{j}_{\text{conv}}^a = \rho^a \mathbf{v} , \quad (\text{C.41})$$

enquanto que a **parte condutiva** $\mathbf{j}_{\text{cond}}^a$ descreve o transporte da grandeza a devido a outros processos e portanto poderá estar presente mesmo quando o fluido estiver em repouso (isto é, quando $\mathbf{v} \equiv 0$). Por exemplo, é um fenômeno geral na termodinâmica que quando houver uma distribuição não-homogênea de alguma grandeza extensiva a no espaço, ou seja, quando a densidade ρ^a da quantidade a não for constante, o sistema estará fora de equilíbrio, e portanto iniciar-se-ão processos dissipativos ou compensatórios, tentando (re)estabelecer o equilíbrio. A lei básica que rege um tal processo de dissipação de uma grandeza extensiva a (processo durante o qual ρ^a tende para uma função constante, na medida em que $t \rightarrow \infty$) é uma relação constitutiva que fornece $\mathbf{j}_{\text{cond}}^a$ como função dada e explícita das derivadas da densidade ρ^a (e possivelmente das densidades de outras grandezas extensivas envolvidas), sendo a hipótese mais simples neste contexto que $\mathbf{j}_{\text{cond}}^a$ depende linearmente do gradiente de ρ^a , ou seja

$$\mathbf{j}_{\text{cond}}^a = -c_a \nabla \rho^a , \quad (\text{C.42})$$

onde c_a é uma constante positiva (o sinal negativo indica que a corrente sempre estará orientada de regiões de alta densidade de a para regiões de baixa densidade). Sob esta hipótese e na ausência de escoamento ($\mathbf{v} \equiv 0$), a equação de balanço (C.38) para a quantidade a assume a forma de uma equação de difusão:

$$\frac{\partial \rho^a}{\partial t} - c_a \Delta \rho^a = q^a . \quad (\text{C.43})$$

Consideramos, por exemplo, a dissolução de uma substância num solvente (como açúcar em água). Temos então um sistema com dois tipos de partículas: as moléculas do solvente (água) e as moléculas da substância dissolvida (açúcar). Na ausência de reações químicas, o número de moléculas de cada espécie é uma quantidade conservada. Além disto, a densidade de moléculas n_{sol} do solvente será, tipicamente, uma constante (e $\mathbf{j}^{n_{\text{sol}}} = 0$), de forma que podemos substituir a densidade de moléculas n_{sub} da substância dissolvida pela concentração local $c = n_{\text{sub}}/n_{\text{sol}}$ da solução, obtendo assim a equação de difusão

$$\frac{\partial c}{\partial t} - \lambda \Delta c = 0, \quad (\text{C.44})$$

onde λ é uma constante positiva, a constante de difusão.

De maneira análoga, deduz-se a equação de condução do calor, substituindo o número de partículas pela energia calórica, que na ordem mais baixa de aproximação é proporcional à temperatura T (com o calor específico como coeficiente de proporcionalidade), levando assim à equação de condução do calor

$$\frac{\partial T}{\partial t} - \lambda \Delta T = 0, \quad (\text{C.45})$$

onde λ é uma constante positiva, a condutibilidade térmica.

Bibliografia

- Cartan, H.:** *Calcul Différentiel*, Hermann, Paris 1967.
- de Figueiredo, D.G.:** *Análise de Fourier e Equações Diferenciais Parciais*, Projeto Euclides, IMPA, Rio de Janeiro 1977.
- Dieudonné, J.:** *Foundations of Modern Analysis*, Academic Press, New York 1960.
- Duff, G., R.L. Walker, R.L.:** *Differential Equations of Applied Mathematics*, John Wiley, New York 1966.
- Folland, G.:** *Introduction to Partial Differential Equations*, 2nd edition, Princeton University Press, Princeton 1995.
- John, F.:** *Partial Differential Equations*, 4th edition, Springer, New York 1982.
- Jost, J.:** *Partial Differential Equations*, 2nd edition, Springer, New York 2007.
- Lang, S.:** *Analysis I*, Addison-Wesley, Reading 1968.
- Lang, S.:** *Analysis II*, Addison-Wesley, Reading 1969.
- Schwartz, L.:** *Mathematics for the Physical Sciences*, Hermann, Paris 1966.
- Schwartz, L.:** *Théorie des Distributions*, 3ème edition, Hermann, Paris 1966.

Referências Bibliográficas

- [1] Thomas, E.G.F., *A Polarization Identity for Multilinear Maps*, Indag. Math. **25** (2014) 468-474.

Índice Remissivo

- álgebra de Clifford, 26
- aplicação
 - contínua, 84
- aplicação linear
 - antissimétrica, 81
 - bidual, 81
 - dual ou transposta, 79
 - simétrica, 81
- aproximação da unidade, 130
- base de uma topologia, 83
- base dual, 79
- calibre
 - condição de, 198
 - de Coulomb, 198
 - de Lorentz, 198
 - escolha de, 198
 - transformação de, 198
 - residual, 198
- campo
 - de escoamento, 203
 - elétrico, 195
 - escalar, 183
 - magnético, 195
 - unitário
 - normal, 183
 - tangente, 183
 - vetorial, 183
- característico
 - covetor, 11, 17
 - hiperplano, 11, 17
- carta, 176
 - adaptada, 177
 - domínio, 176
 - local, 176
- circulação de campo vetorial ao longo de
 - uma curva fechada, 183
- coeficientes de Fourier
 - de distribuições periódicas, 151
 - de uma função periódica, 147
- conjunto
 - aberto, 82
 - compacto, 181
 - convexo, 84
 - direcionado, 92
 - fechado, 83
 - ordenado, 92
 - paracompacto, 181
- convolução
 - de distribuições, 162
 - de funções integráveis, 159
- coordenadas, 176
 - adaptadas, 177
 - características em \mathbb{R}^2 , 74
 - do cone da luz em \mathbb{R}^2 , 74
 - domínio, 176
 - esféricas em \mathbb{R}^n , 184
 - locais, 176
- curva, 183
- d'alembertiano, 23
- densidade
 - de carga (elétrica), 195
 - de grandeza extensiva, 206
 - de massa, 204
- densidade de corrente
 - de carga (elétrica), 195
 - de grandeza extensiva, 206
 - de massa, 204
 - parte condutiva, 207
 - parte convectiva, 207
- densidade de produção

- de grandeza extensiva, 206
- determinante
 - métrico, 178
- difeomorfismo, 175
- distribuição
 - de Dirac, 110
 - derivadas, 110
 - definição, 105
 - periódica, 144
 - regular sobre um aberto, 115
 - restrição a um aberto, 115
 - zero sobre um aberto, 115
- divergência, 190
- equação
 - característica, 20
 - de balanço, 206
 - de continuidade, 205
 - de evolução, 28
 - estática, 28
- equação de difusão, 28
- equação de Laplace, 22, 28
- equação de ondas, 27
 - bidimensional, 27
 - tridimensional, 27
 - unidimensional, 27
 - solução geral, 74, 75
- equação de Poisson, 22, 28
- equação diferencial
 - explícita, 5
 - geral, 2
 - implícita, 5
 - linear, 2, 13
 - fonte, 14
 - homogênea, 13
 - não-homogênea, 13
 - solução, 14
 - solução fraca, 14
 - termo não-homogêneo, 14
- ordem ou grau, 2
- quasilinear, 3
- semilinear, 3
- total, 6
 - condição de curvatura zero, 10
 - condição de integrabilidade de Frobenius, 9
 - integrável, 8
- equações de Maxwell, 195
- espaço
 - de funções
 - \mathcal{B} , \mathcal{B}_0 , \mathcal{D}_{LP} , 134
 - $\mathcal{B}(\Omega)$, $\mathcal{B}_0(\Omega)$, $\mathcal{D}_{LP}(\Omega)$, 134
 - \mathcal{D} , \mathcal{E} , \mathcal{S} , 99
 - $\mathcal{D}(\Omega)$, $\mathcal{E}(\Omega)$, 99
 - espaço de Banach, 91
 - espaço de distribuições, 128
 - espaço de Fréchet, 91
 - espaço de funções teste, 128
 - espaço localmente convexo, 82
 - completo, 91
 - normável, normado, 90
 - espaço topológico, 82
 - de Hausdorff, 89
 - espaço vetorial
 - bidual, 81
 - dual, 78
 - algébrico, 81, 98
 - topológico, 81, 98
 - topológico, 82
 - localmente convexo, 82
 - metrizável, metrizado, 91
- espaço localmente convexo
 - metrizável, metrizado, 91
- fórmula de mudança de variáveis
 - em integrais múltiplas, 176
- fórmula de Plancherel, 170
- fórmula de Poisson, 45
- faixa, 160
 - condição da, 161, 163
- fluxo de campo vetorial através
 - de uma superfície, 183
- forma quadrática
 - assinatura, 19
- função
 - de Green
 - como distribuição, 120
 - coulombiana, 51
 - do problema de Dirichlet, 54
 - do problema de Neumann, 54
 - para o operador de Laplace, 52
 - de Heaviside, 108

- delta de Dirac, 110
 - derivadas, 110
- gama de Euler, 188
- harmônica, 22
- localmente integrável
 - como distribuição, 107
 - definição, 106
- teste, 98
- funcional, 81
- gradiente, 190
- grandeza
 - intensiva, 206
- grandeza extensiva, 205
- hipersuperfície, 183
- homeomorfismo, 84
- identidades de Green
 - primeira, 191
 - segunda, 192
 - terceira, 52
- interação efetiva, 199
- laplaciano, 22
- lei
 - da indução, 195
 - da inexistência de
 - cargas magnéticas, 196
 - de Ampère, 196
 - de Ampère-Maxwell, 196
 - de Faraday, 195
 - de Gauss, 195
- lei de conservação (local)
 - da carga (elétrica), 205
 - da massa, 205
- lema
 - de Riemann-Lebesgue, 164
- lema de Schwartz, 120
- linhas de fluxo, 203
- média esférica, 192
- métrica
 - induzida, 178
 - invariante por translações, 91
- multi-índice, 3
 - ordem ou grau, 3
- núcleo
 - de Poisson, 58
- núcleo de difusão, 73
- núcleo do calor, 73
- norma, 85
 - da soma ou tipo l^1 , 85
 - do máximo ou tipo l^∞ , 85
 - do supremo, 85
 - das derivadas até ordem r , 94, 95
 - euclideana ou tipo l^2 , 85
 - tipo l^p , 85
- operador
 - de Laplace, 22
- operador de difusão, 22
- operador de Dirac, 26
- operador de Helmholtz, 24
- operador de Klein-Gordon, 24
- operador de ondas, 23
- operador de Schrödinger, 24
- operador de transporte, 22
- operador diferencial
 - a coeficientes constantes, 16
 - elíptico, 17, 20
 - formal, 14, 15
 - grau, 15
 - hiperbólico, 20
 - parabólico, 20
 - símbolo, 16
 - símbolo principal, 17
- operadores
 - associados ao cone de luz, 25
 - de Cauchy-Riemann, 24
- parte finita de $1/x^k$, 112
- partições da unidade
 - definição, 181
 - teorema de existência, 182
- polinômio, 15
 - grau, 15
- potencial
 - de Coulomb, 43
 - escalar, 197
 - instantâneo, 199
 - vetorial, 197
- princípio do máximo/mínimo, 41

- problema de Cauchy, 12, 29
- problema de Dirichlet, 30
 - solução clássica
 - existência, 40
 - unicidade, 39
- problema de Neumann, 30
- produto
 - tensorial de distribuições, 153
- quantidade conservada, 206
- recobrimento aberto
 - finito, 180
 - localmente finito, 180
 - pontualmente finito, 180
 - refinamento, 180
 - sub-recobrimento, 180
- regra do produto, 125
- relação de ordem, 92
 - antissimetria, 92
 - parcial, 92
 - reflexividade, 92
 - total, 92
 - transitividade, 92
- rotacional, 190
- série de Fourier, 145
 - de uma função periódica, 147
- semi-métrica
 - invariante por translações, 91
- seminorma, 85
 - L^1 , 86
 - L^2 , 86
 - L^p , 86
 - do supremo essencial ou L^∞ , 87
 - núcleo, 87
 - sistema direcionado, 93
 - sistemas equivalentes, 89
- sistema de equações diferenciais
 - explícito, 5
 - geral, 4
 - implícito, 5
 - total, 7
- sistema fechado, 206
- solução fundamental, 120
- subvariedade, 177
- suporte
 - de uma distribuição, 117
 - de uma função, 180
- suporte singular, 117
- teorema
 - da divergência, 190
 - da média esférica, 58
 - das singularidades removíveis, 60
 - de Cauchy-Kovalevski, 12
 - de Ehrenpreis-Malgrange, 121
 - de Frobenius, 9
 - de Fubini
 - para distribuições, 156
 - de Gauss, 190
 - de inversão de Fourier, 168
 - de regularidade, 59
 - de Schwartz
 - de convergência, 136
 - do suporte pontual, 118
 - de Stokes, 190
- termodinâmica
 - segundo teorema fundamental, 206
- topologia, 83
 - da convergência uniforme sobre subconjuntos compactos, 95
 - de todas as derivadas, 95
 - da convergência pontual, 98
 - da convergência uniforme, 94, 95
 - das derivadas até ordem r , 94, 95
 - de todas as derivadas, 94, 95
 - da convergência uniforme sobre subconjuntos compactos das derivadas até ordem r , 95
 - fraca, 98
 - gerada por um sistema de seminormas, 88
 - gerada por uma base, 83
- transformação de Fourier
 - de distribuições temperadas, 170
 - de funções integráveis, 164
- valor principal de $1/x$, 110
- vizinhança, 83