Introdução à Análise Funcional

José Carlos de Sousa Oliveira Santos
Departamento de Matemática Pura
Faculdade de Ciências
Universidade do Porto

Porto — Julho de 2010

Índice

Introdução

Estas notas destinam-se aos alunos da cadeira Medida e Integração, frequentada por alunos do terceiro e do quarto anos da licenciatura em Matemática da Faculdade de Ciências da Universidade do Porto.

São usadas as seguintes notações:

$\overline{\mathbb{R}}$	$\mathbb{R} \cup \{\pm\infty\}$
$\mathbb{R}_+$	$\{\, x \in \mathbb{R} \mid x \geqslant 0 \,\}$
$\overline{\mathbb{R}_+}$	$\mathbb{R}_+ \cup \{+\infty\}$
$\mathbb{R}_+^*$	$\mathbb{R}_+ \setminus \{0\}$
$\mathcal{P}(E)$	$\{\text{partes de } E\}$

Vai-se considerar em $\overline{\mathbb{R}}$ a relação de ordem $\leqslant$ que prolonga a relação de ordem $\leqslant$ de $\mathbb{R}$ e para a qual se tem:

$$(\forall r \in \mathbb{R}) : -\infty \leqslant r \leqslant +\infty.$$

Sempre que se falar de supremo ou de ínfimo de uma parte de $\overline{\mathbb{R}}$ será relativamente a esta relação de ordem. Observe-se que, com esta convenção, qualquer parte P de $\overline{\mathbb{R}}$ tem supremo e ínfimo e que o supremo (respectivamente ínfimo) de P é um número real se e só se P não for vazia e se for majorada (resp. minorada) por algum número real.

Um conjunto C dir-se-á numerável quando C for finito ou quando existir alguma bijecção de $\mathbb{N}$ em C.

O símbolo $\blacksquare$ assinala o fim das demonstrações.

Teoria da medida

Na passagem do século XIX para o século XX, tornou-se claro que o integral de Riemann era insuficiente para as necessidades dos analistas. Por exemplo, se $[a, b]$ é um intervalo de $\mathbb{R}$ e se $(f_n)_{n \in \mathbb{N}}$ é uma sucessão de funções de $[a, b]$ em $\mathbb{R}$ integráveis segundo Riemann pontualmente convergente para uma função f de $[a, b]$ em $\mathbb{R}$, não é necessariamente verdade que f seja integrável segundo Riemann, mesmo que seja limitada. Além disso, se $f \colon [a, b] \longrightarrow \mathbb{R}$ for uma função para a qual haja funções integráveis segundo Riemann $f_1, f_2 \colon [a, b] \longrightarrow \mathbb{R}$ tais que $f_1 \leqslant f \leqslant f_2$, seria desejável que f fosse também integrável segundo Riemann, mas não é esse necessariamente o caso.

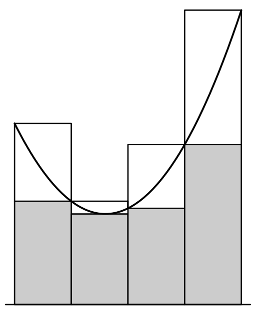

Figura 1.1: Ilustração geométrica do conceito de «integral de Riemann»

Examinemos com um pouco de detalhe o conceito de «integral de Riemann»; veja-se a figura 1.1. O integral de Riemann de uma função f de um intervalo $[a, b]$ em $\mathbb{R}_+$ não é mais do que a área da figura situada entre o gráfico de f e o eixo dos xx. A ideia do integral de Riemann consiste em enquadrar f entre duas funções, cada uma das quais é constante num conjunto finito de sub-intervalos de $[a, b]$ dois a dois disjuntos cuja reunião é precisamente $[a, b]$. Para uma tal função, dispomos de um conceito intuitivo de área abaixo do gráfico: se $[a, b]$ for a reunião disjunta dos intervalos $I_1, I_2, \ldots, I_n$ e se φ for uma função de $[a,b]$ em $\mathbb{R}$ cuja restrição a cada I_k seja constante e tome sempre o valor c_k, então a área abaixo do gráfico de φ será $\sum_{k=1}^{n} c_k \operatorname{comp}(I_k)$, onde $\operatorname{comp}(I_k)$ representa o comprimento de I_k. O integral de Riemann de f será então o valor para que tendem as áreas abaixo dos gráficos das funções φ à medida que tendem para a função f.

A ideia por trás do integral de Lebesgue, que é aquele que será abordado neste curso, é mais geral e consiste em substituir as famílias finitas de subintervalos de $[a,b]$ dois a dois disjuntos cuja reunião é igual a $[a,b]$ por famílias numeráveis de subconjuntos de $[a,b]$ dois a dois disjuntos cuja reunião é igual a $[a,b]$. Há dois problemas com esta abordagem. O primeiro reside no facto de envolver somas de famílias numeráveis de números reais não negativos ou, mais geralmente, de elementos de $\overline{\mathbb{R}_+}$. Este problema resolve-se facilmente empregando séries. O segundo problema é que surge a necessidade de se generalizar o conceito de comprimento de um intervalo a um conceito mais geral de medida de um subconjunto de $\mathbb{R}$ que seja aplicável a uma grande quantidade de partes de $\mathbb{R}$. É este conceito de medida que vai ser definido.

1.1 Álgebras e σ-álgebras

DEFINIÇÃO 1.1 Dado um conjunto X, diz-se que um subconjunto $\mathcal{A}$ de $\mathcal{P}(X)$ é uma *álgebra* se

1. $\emptyset \in \mathcal{A}$;
2. $(\forall A, B \in \mathcal{A}) : A \cup B \in \mathcal{A}$;
3. $(\forall A \in \mathcal{A}) : A^{\complement} \in \mathcal{A}$.

Resulta da terceira condição que a primeira condição pode ser substituída por «$X \in \mathcal{A}$». Por outro lado, deduz-se facilmente da segunda condição, usando indução, que se $\mathcal{A}$ for uma álgebra e se $A_1, \ldots, A_n \in \mathcal{A}$, então $\bigcup_{j=1}^{n} A_j \in \mathcal{A}$.

DEFINIÇÃO 1.2 Diz-se que uma álgebra $\mathcal{A}$ é uma *σ-álgebra* se, para cada sucessão $(A_n)_{n\in\mathbb{N}}$ de elementos de $\mathcal{A}$, $\bigcup_{n\in\mathbb{N}} A_n \in \mathcal{A}$.

Dado um conjunto X, para se verificar que um subconjunto $\mathcal{A}$ de $\mathcal{P}(X)$ é uma σ-álgebra, não é necessário verificar separadamente que é estável para reuniões finitas e para reuniões infinitas numeráveis. De facto, se se verificar que é estável para reuniões infinitas numeráveis, então resulta de se ter $\emptyset \in \mathcal{A}$ que $\mathcal{A}$ também é estável para reuniões finitas.

EXEMPLO 1.1 Dado um conjunto X, $\mathcal{P}(X)$ é uma σ-álgebra.

EXEMPLO 1.2 Dado um conjunto X, o conjunto das partes de X que são finitas ou têm complementar finito forma uma álgebra que não é uma σ-álgebra.

PROPOSIÇÃO 1.1 *Se $\mathcal{A}$ for uma álgebra (respectivamente uma σ-álgebra) e se $A_1, \ldots, A_n \in \mathcal{A}$ (resp. $(A_n)_{n\in\mathbb{N}}$ é uma sucessão de elementos de $\mathcal{A}$), então $\bigcap_{j=1}^n A_n \in \mathcal{A}$ (resp. $\bigcap_{n\in\mathbb{N}} A_n \in \mathcal{A}$).*

DEMONSTRAÇÃO: No caso das σ-álgebras, basta ver que

$$\left(\bigcap_{n\in\mathbb{N}} A_n\right)^{\complement} = \bigcup_{n\in\mathbb{N}} A_n^{\complement} \in \mathcal{A}.$$

No caso das álgebras, a demonstração é análoga. ∎

PROPOSIÇÃO 1.2 *Sejam X um conjunto e P um conjunto de partes de X. Então, das σ-álgebras de partes de X que contêm P, existe uma e uma só que está contida em todas as outras.*

DEMONSTRAÇÃO: A unicidade é imediata: se $\mathcal{A}, \mathcal{A}' \subset \mathcal{P}(X)$ forem σ--álgebras que contenham P e que estejam contidas em qualquer σ-álgebra que contenha P então, em particular, $\mathcal{A} \subset \mathcal{A}'$ e $\mathcal{A}' \subset \mathcal{A}$, pelo que $\mathcal{A} = \mathcal{A}'$.

Por outro lado, seja A o conjunto de todas as σ-álgebras contidas em $\mathcal{P}(X)$ que contenham P. O conjunto A não é vazio, pois $\mathcal{P}(X) \in A$. Verifica-se facilmente que a intersecção de todos os elementos de A é uma σ-álgebra, a qual necessariamente contém P e está contida em qualquer σ-álgebra em $\mathcal{P}(X)$ que contenha P. ∎

DEFINIÇÃO 1.3 Dados um conjunto X e um conjunto P de partes de X, designa-se por σ-álgebra *gerada* por P a única σ-álgebra de partes de X que contém P e que está contida em qualquer outra σ-álgebra nas mesmas condições.

Para o que se segue, é conveniente encarar $\overline{\mathbb{R}}$ como um espaço métrico. Para tal, considere-se a bijecção

$$\begin{array}{rccc} f\colon & \overline{\mathbb{R}} & \longrightarrow & [-1,1] \\ & x & \longmapsto & \begin{cases} \frac{x}{1+|x|} & \text{se } x \in \mathbb{R} \\ \pm 1 & \text{se } x = \pm\infty, \end{cases} \end{array}$$

e a distância

$$\begin{array}{rccc} d\colon & \overline{\mathbb{R}}\times\overline{\mathbb{R}} & \longrightarrow & \mathbb{R}_+ \\ & (x,y) & \mapsto & |f(x)-f(y)|. \end{array}$$

Relativamente a esta distância tem-se, por exemplo, que $\lim_{n\in\mathbb{N}} n = +\infty$ e que a função

$$\begin{array}{ccc} \overline{\mathbb{R}} & \longrightarrow & \overline{\mathbb{R}} \\ x & \mapsto & \begin{cases} 1/|x| & \text{se } x \neq 0, \pm\infty \\ 0 & \text{se } x = \pm\infty \\ +\infty & \text{se } x = 0 \end{cases} \end{array}$$

é contínua.

Vejamos como é possível determinar se uma parte A de $\overline{\mathbb{R}}$ é aberta sem usar a distância d.

PROPOSIÇÃO 1.3 *Dado um subconjunto A de $\overline{\mathbb{R}}$, são condições equivalentes:*

1. *A é um aberto de $\left(\overline{\mathbb{R}}, d\right)$;*
2. *$A \cap \mathbb{R}$ é um aberto de $\mathbb{R}$ (relativamente à métrica usual) e, além disso,*
   - *se $+\infty \in A$, então $]t, +\infty] \subset A$ para algum $t \in \mathbb{R}$;*
   - *se $-\infty \in A$, então $[-\infty, t[\subset A$ para algum $t \in \mathbb{R}$.*

DEMONSTRAÇÃO: Seja f a função que surge na definição da distância d. Então, pela definição de d, f é uma isometria de $\overline{\mathbb{R}}$ sobre $[-1, 1]$ (munido da métrica usual); em particular, é um homeomorfismo. Logo, $\mathbb{R}$ é um aberto de $\overline{\mathbb{R}}$, pois $\mathbb{R} = f^{-1}(]-1,1[)$. Resulta daqui que se A for um aberto de $\overline{\mathbb{R}}$, então $A \cap \mathbb{R}$ também o é.

Vejamos que a restrição de d a $\mathbb{R} \times \mathbb{R}$ é uma distância equivalente à distância usual em $\mathbb{R}$ (que será representada por d'), i. e. que d e d' dão origem aos mesmos abertos. Isto é o mesmo que dizer que a função $\mathrm{id}\colon (\mathbb{R}, d') \longrightarrow (\mathbb{R}, d|_{\mathbb{R}\times\mathbb{R}})$ é um homemomorfismo. Mas resulta novamente da definição de d que $f|_{\mathbb{R}}$ é um homeomorfismo de $\mathbb{R}$ sobre $]-1,1[$ (munido da métrica usual). Logo, $\mathrm{id}\colon (\mathbb{R}, d') \longrightarrow (\mathbb{R}, d|_{\mathbb{R}\times\mathbb{R}})$ é um homemomorfismo se e só se $f \circ \mathrm{id}$ for um homeomorfismo de $\mathbb{R}$ em $]-1,1[$ (relativamente à métrica usual em ambos os casos), o que claramente se verifica.

Seja A um aberto de $\overline{\mathbb{R}}$. Então, como já foi provado, $A \cap \mathbb{R}$ é também um aberto de $\overline{\mathbb{R}}$. Como, por outro lado, $A \cap \mathbb{R} \subset \mathbb{R}$, resulta do que se provou no parágrafo anterior que $A \cap \mathbb{R}$ é um aberto de $\mathbb{R}$ (relativamente à métrica usual). Caso $+\infty \in A$, seja $\varepsilon > 0$ tal que

$$(\forall x \in \mathbb{R}) : d(x, +\infty) < \varepsilon \Longrightarrow x \in A; \tag{1.1}$$

um tal ε existe necessariamente por se estar a supor que A é aberto e que contém $+\infty$. Pode-se (e vai-se) supor que $\varepsilon < 1$. Pela definição da distância d, (1.1) é equivalente a

$$(\forall x \in \mathbb{R}) : \left|\frac{x}{1+|x|} - 1\right| < \varepsilon \Longrightarrow x \in A$$

e, em particular,

$$(\forall x \in \mathbb{R}_+) : \frac{1}{1+x} < \varepsilon \Longrightarrow x \in A.$$

Então $\left]1/\varepsilon - 1, +\infty\right] \subset A$. Mostra-se analogamente que, se $-\infty \in A$, então existe algum $t \in \mathbb{R}$ tal que $[-\infty, t[\subset A$.

Suponha-se agora que $A \subset \overline{\mathbb{R}}$ satisfaz a segunda condição do enunciado; quer-se provar que A é um aberto de $\overline{\mathbb{R}}$, o que é o mesmo que afirmar que é vizinhança de todos os seus pontos. Seja então $a \in A$. Há três possibilidades:

$a \in \mathbb{R}$: então, como $A \cap \mathbb{R}$ é um aberto de $\overline{\mathbb{R}}$ e $a \in A \cap \mathbb{R}$, $A \cap \mathbb{R}$ é uma vizinhança de a, pelo que A é vizinhança de a;

$a = +\infty$: existe algum $t \in \mathbb{R}_+$ tal que $]t, +\infty] \subset A$, pelo que A contém $B(+\infty, 1/(1+t))$;

$a = -\infty$: análogo ao anterior. ∎

É claro que qualquer aberto de $\overline{\mathbb{R}}$ pode ser obtido como reunião de intervalos abertos, i. e. de intervalos do tipo $]x, y[$ (com $x, y \in \overline{\mathbb{R}}$ e $x < y$), do tipo $[-\infty, x[$ (com $x \in \mathbb{R}$) ou do tipo $]x, +\infty]$ (com $x \in \mathbb{R}$). Vejamos que é possível dizer mais sobre isto.

LEMA 1.1 *Qualquer aberto de $\overline{\mathbb{R}}$ pode ser escrito como reunião numerável de intervalos abertos dois a dois disjuntos.*

DEMONSTRAÇÃO: Seja A um aberto não vazio de $\overline{\mathbb{R}}$ e, para cada $x \in A$, seja I_x a reunião de todos os intervalos abertos de $\overline{\mathbb{R}}$ contido em A que contenham x. Então, como a reunião de intervalos com um ponto em comum é novamente um intervalo, a reunião de abertos é um aberto e a reunião de subconjuntos de A que contêm x é novamente um subconjunto de A que contém x, I_x é um intervalo aberto de $\overline{\mathbb{R}}$ contido em A e que contém x; de facto, é mesmo o maior intervalo aberto contido em A que contém x. Então $A = \bigcup_{x \in A} I_x$ e, por outro lado, é claro que, se $x, y \in A$, os intervalos I_x e I_y ou coincidem ou são disjuntos. Está então provado que A pode ser obtido sob a forma de reunião de uma família $(I_\lambda)_{\lambda\in\Lambda}$ de intervalos não vazios dois a dois disjuntos. Cada intervalo I_λ contém números racionais e, como os intervalos são dois a dois disjuntos, nenhum número racional está em dois intervalos distintos. Logo, $\left(I_\lambda \cap \mathbb{Q}\right)_{\lambda\in\Lambda}$ é uma família de conjuntos não vazios dois a dois disjuntos cuja reunião é igual a $A \cap \mathbb{Q}$, que é numerável. Resulta desta observação que Λ é numerável. ■

DEFINIÇÃO 1.4 Diz-se que um subconjunto X de $\overline{\mathbb{R}}$ é *boreliano* se pertencer à σ-álgebra gerada pelos abertos de $\overline{\mathbb{R}}$. Se $A \subset \overline{\mathbb{R}}$ então o conjunto dos borelianos contidos em A representa-se por $\mathcal{B}(A)$.

Alternativamente, poder-se-ia ter definido $\mathcal{B}(A)$ como sendo o a σ--álgebra gerada pelas partes abertas de A. Por outro lado, resulta do lema 1.1 que $\mathcal{B}(A)$ também é a σ-álgebra gerada pelos intervalos abertos de A.

Vejamos alguns exemplos de conjuntos borelianos.

1. Resulta imediatamente da definição que qualquer aberto de $\overline{\mathbb{R}}$ é boreliano.

2. Como as σ-álgebras são fechadas para a passagem ao complementar e como um conjunto é fechado se e só se o seu complementar é aberto, resulta da observação anterior que os conjuntos fechados são borelianos.

3. Qualquer conjunto finito ou numerável é boreliano, pois se C for um tal conjunto, então $C = \bigcup_{x\in C}\{x\}$ e cada conjunto $\{x\}$ é fechado e, portanto, boreliano.

4. Resulta da observação anterior que se $X \subset \overline{\mathbb{R}}$ for tal que $X^{\complement}$ seja finito ou numerável, então X é boreliano.

Em particular, $\mathbb{Q}$ e $\overline{\mathbb{R}} \setminus \mathbb{Q}$ são ambos borelianos.

Pode-se provar que há partes de $\overline{\mathbb{R}}$ que não são borelianas. De facto, se $\mathcal{A}$ for uma σ-álgebra de partes de um conjunto com o mesmo cardinal que $\mathbb{R}$ que seja gerada por um conjunto cujo cardinal não excede o de $\mathbb{R}$, então o cardinal de $\mathcal{A}$ também não excede o de $\mathbb{R}$; pode-se ver uma demonstração detalhada em [8, §1.6]. Como $\mathcal{B}(\overline{\mathbb{R}})$ é gerada pelos intervalos abertos e o conjunto dos intervalos abertos tem o mesmo cardinal que $\mathbb{R}$, isto prova que o cardinal de $\mathcal{B}(\overline{\mathbb{R}})$ é igual ao de $\mathbb{R}$, pelo que $\mathcal{B}(\overline{\mathbb{R}}) \subsetneq \mathcal{P}(\overline{\mathbb{R}})$. O mesmo argumento prova que $\mathcal{B}(\mathbb{R}) \subsetneq \mathcal{P}(\mathbb{R})$.

Um exemplo concreto de um conjunto não boreliano é o conjunto dos números reais x para os quais existe alguma sucessão $(a_n)_{n\in\mathbb{N}}$ de números naturais tal que

$$x = \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{a_3 + \cdots}}}$$

e que, para alguma sucessão estritamente crescente $(m_n)_{n\in\mathbb{N}}$ de números naturais, se tenha $(\forall j \in \mathbb{N}) : a_{m_j} | a_{m_{j+1}}$; veja-se [11, cap. III].

1.2 Medidas

DEFINIÇÃO 1.5 Dada uma σ-álgebra $\mathcal{A}$, diz-se que uma função m de $\mathcal{A}$ em $\overline{\mathbb{R}_+}$ é uma *medida* se tiver as seguintes propriedades:

1. $m(\emptyset) = 0$;
2. se $(A_n)_{n\in\mathbb{N}}$ é uma sucessão de elementos de $\mathcal{A}$ disjuntos dois a dois, então
$$m\left(\bigcup_{n\in\mathbb{N}} A_n\right) = \sum_{n=1}^{+\infty} m(A_n).$$

O termo da direita da igualdade anterior é $+\infty$ caso a série em questão seja uma sére divergente de números reais ou caso algum dos termos seja $+\infty$.

Observe-se que a segunda condição poderia ser substituída pela condição *a priori* mais forte:

2′ se $(A_n)_{n\in N}$ é uma família numerável de elementos de $\mathcal{A}$ disjuntos dois a dois, então
$$m\left(\bigcup_{n\in N} A_n\right) = \sum_{n\in N} m(A_n).$$

De facto, se $A_1, \ldots, A_n \in \mathcal{A}$ forem conjuntos disjuntos dois a dois e se se definir $A_k = \emptyset$ para cada $k > n$, então

$$m\left(\bigcup_{k=1}^{n} A_k\right) = m\left(\bigcup_{k=1}^{+\infty} A_k\right) = \sum_{k=1}^{+\infty} m(A_k) = \sum_{k=1}^{n} m(A_k),$$

uma vez que o conjunto vazio tem medida nula.

EXEMPLO 1.3 Seja X um conjunto. Então a função

$$\begin{aligned} m\colon \mathcal{P}(X) &\longrightarrow \overline{\mathbb{R}_+} \\ A &\mapsto \begin{cases} \#A & \text{se } A \text{ for finito} \\ +\infty & \text{caso contrário} \end{cases} \end{aligned}$$

é uma medida, que se designa por *medida de contagem*.

PROPOSIÇÃO 1.4 *Dadas uma σ-álgebra $\mathcal{A}$ e uma medida m definida em $\mathcal{A}$, tem-se*

1. *se $A, B \in \mathcal{A}$ e $A \subset B$, então $m(A) \leqslant m(B)$ e, além disso, se $m(A) < +\infty$, então $m(B \setminus A) = m(B) - m(A)$;*
2. *se $(A_n)_{n\in\mathbb{N}}$ é uma sucessão de elementos de $\mathcal{A}$, então*

$$m\left(\bigcup_{n\in\mathbb{N}} A_n\right) \leqslant \sum_{n=1}^{+\infty} m(A_n).$$

DEMONSTRAÇÃO: Para demonstrar a primeira alínea, basta ver que

$$m(B) = m(A \cup (B \setminus A)) = m(A) + m(B \setminus A) \geqslant m(A).$$

Caso $m(A) < +\infty$, resulta da igualdade $m(B) = m(A) + m(B \setminus A)$ que $m(B \setminus A) = m(B) - m(A)$.

Quanto à segunda alínea, considere-se a sucessão $(B_n)_{n\in\mathbb{N}}$ de elementos de $\mathcal{A}$ tal que $(\forall n \in \mathbb{N}) : B_n = A_n \setminus \bigcup_{k<n} A_k$. Então os conjuntos da forma B_n ($n \in \mathbb{N}$) são disjuntos dois a dois e a sua reunião é igual a $\bigcup_{n\in\mathbb{N}} A_n$, pelo que

$$m\left(\bigcup_{n\in\mathbb{N}} A_n\right) = m\left(\bigcup_{n\in\mathbb{N}} B_n\right) = \sum_{n=1}^{+\infty} m(B_n) \leqslant \sum_{n=1}^{+\infty} m(A_n),$$

pela alínea anterior. ∎

PROPOSIÇÃO 1.5 *Dadas uma σ-álgebra $\mathcal{A}$, uma medida m definida em $\mathcal{A}$ e uma sucessão monótona $(A_n)_{n\in\mathbb{N}}$ de elementos de $\mathcal{A}$, tem-se*

1. *se $A_1 \subset A_2 \subset A_3 \subset \cdots$, então*

$$m\left(\bigcup_{n\in\mathbb{N}} A_n\right) = \lim_{n\in\mathbb{N}} m(A_n);$$

2. *se* $A_1 \supset A_2 \supset A_3 \supset \cdots$ *e se* A_1 *tem medida finita, então*

$$m\left(\bigcap_{n\in\mathbb{N}} A_n\right) = \lim_{n\in\mathbb{N}} m(A_n).$$

DEMONSTRAÇÃO: No caso da primeira alínea, seja $(B_n)_{n\in\mathbb{N}}$ a sucessão de elementos de $\mathcal{A}$ tal que $B_1 = A_1$ e que $B_n = A_n \setminus A_{n-1}$ para cada número natural n maior do que 1. Então $(B_n)_{n\in\mathbb{N}}$ é uma sucessão de elementos de $\mathcal{A}$ disjuntos dois a dois e, para cada $n \in \mathbb{N}$,

$$A_n = \bigcup_{k=1}^{n} B_k,$$

pelo que

$$m(A_n) = \sum_{k=1}^{n} m(B_k).$$

Logo,

$$\begin{aligned} m\left(\bigcup_{n\in\mathbb{N}} A_n\right) &= m\left(\bigcup_{n\in\mathbb{N}} B_n\right) \\ &= \sum_{k=1}^{+\infty} m(B_k) \\ &= \lim_{n\in\mathbb{N}} \sum_{k=1}^{n} m(B_k) \\ &= \lim_{n\in\mathbb{N}} m(A_n). \end{aligned}$$

Para demonstrar a segunda alínea, basta ver que, pela primeira alínea

$$\begin{aligned} m\left(A_1 \Big\backslash \bigcap_{n\in\mathbb{N}} A_n\right) &= m\left(\bigcup_{n\in\mathbb{N}} (A_1 \setminus A_n)\right) \\ &= \lim_{n\in\mathbb{N}} m(A_1 \setminus A_n) \\ &= \lim_{n\in\mathbb{N}} \big(m(A_1) - m(A_n)\big) \text{ (pela proposição 1.4)} \\ &= m(A_1) - \lim_{n\in\mathbb{N}} m(A_n). \end{aligned}$$

Está então provado que

$$m(A_1) - m\left(\bigcap_{n\in\mathbb{N}} A_n\right) = m\left(A_1 \Big\backslash \bigcap_{n\in\mathbb{N}} A_n\right) = m(A_1) - \lim_{n\in\mathbb{N}} m(A_n),$$

de onde resulta que

$$m\left(\bigcap_{n\in\mathbb{N}} A_n\right) = \lim_{n\in\mathbb{N}} m(A_n).$$

■

DEFINIÇÃO 1.6 Se $A \in \mathcal{P}(\mathbb{R})$, define-se a *medida exterior de Lebesgue* de A e representa-se por $m^*(A)$ o ínfimo do conjunto dos números da forma $\sum_{n\in\mathbb{N}} \operatorname{comp}(I_n)$, sendo $(I_n)_{n\in\mathbb{N}}$ uma sucessão de intervalos abertos cuja reunião contém A.

É importante observar que, apesar do nome, a medida exterior de Lebesgue *não é* uma medida, como será visto mais à frente.

Observe-se que se a palavra «abertos» fosse omitida da definição da medida exterior de Lebesgue, a função assim definida seria a mesma. De facto, dado $A \in \mathcal{P}(\mathbb{R})$ represente-se provisoriamente por $\overline{m}(A)$ o ínfimo do conjunto dos números da forma $\sum_{n\in\mathbb{N}} \operatorname{comp}(I_n)$, sendo $(I_n)_{n\in\mathbb{N}}$ uma sucessão de intervalos cuja reunião contém A. É imediato que $\overline{m}(A) \leqslant m^*(A)$. Se se tivesse $\overline{m}(A) < m^*(A)$ para algum $A \in \mathcal{P}(\mathbb{R})$, então seja $r \in]0, m^*(A) - \overline{m}(A)[$. Pela definição de $\overline{m}(A)$, haveria alguma sucessão $(I_n)_{n\in\mathbb{N}}$ de intervalos cuja reunião conteria A e tal que a soma dos seus comprimentos seria menor do que $\overline{m}(A) + r/2$. Para cada $n \in \mathbb{N}$, seja J_n um intervalo aberto que contenha I_n tal que $\operatorname{comp}(J_n) \leqslant \operatorname{comp}(I_n) + 2^{-n-1}r$. Então $(J_n)_{n\in\mathbb{N}}$ é uma sucessão de intervalos abertos cuja reunião contém A, pelo que $\sum_{n=1}^{+\infty} \operatorname{comp}(J_n) \geqslant m^*(A)$. Mas, por outro lado,

$$\begin{aligned}\sum_{n=1}^{+\infty} \operatorname{comp}(J_n) &\leqslant \sum_{n=1}^{+\infty} \left(\operatorname{comp}(I_n) + \frac{r}{2^{n+1}}\right) \\ &< \overline{m}(A) + \frac{r}{2} + \frac{r}{2} \\ &= \overline{m}(A) + r \\ &< m^*(A),\end{aligned}$$

o que é absurdo.

Antes de se passar às propriedades da medida exterior de Lebesgue, vai-se demonstrar um resultado auxiliar. Convém começar por relembrar que uma partição de um intervalo $[a, b]$ é um subconjunto finito de $[a, b]$ que contém a e b. Se P for uma partição, então os intervalos da forma $[a', b']$ com a' e b' elementos consecutivos de P serão designados por intervalos da partição.[1]

[1]Naturalmente, isto é um abuso de linguagem visto que, de facto, nunca se tem $[a', b'] \subset P$.

LEMA 1.2 *Sejam $a, b \in \overline{\mathbb{R}}$ com $a < b$ e seja $(A_\alpha)_{\alpha \in J}$ uma família de abertos de $\overline{\mathbb{R}}$ cuja reunião contenha $[a, b]$. Existe então alguma partição do intervalo $[a, b]$ tal que cada um dos seus intervalos está contido em A_α, para algum $\alpha \in J$.*

DEMONSTRAÇÃO: Seja S o conjunto dos pontos $x \in [a, b]$ com a seguinte propriedade: existe alguma partição do intervalo $[a, x]$ tal que cada um dos seus intervalos está contido em A_α, para algum $\alpha \in J$. Quer-se então provar que $b \in S$.

É claro que $a \in S$ e S que é majorado por b, pelo que S tem supremo $s \in [a, b]$.

Vai-se começar por ver que $s \in S$. Visto que, como já foi afirmado acima, $a \in S$, pode-se supor que $s > a$. Seja β um elemento de J tal que $s \in I_\beta$ e seja t um elemento de $I_\beta \cap [a, s[$. Como $\sup S = s > t$, existe algum $t' \in S \cap [t, s]$. Por definição de S, existe então alguma partição P do intervalo $[a, t']$ tal que cada um dos seus intervalos está contido em A_α, para algum $\alpha \in J$. Mas então $s \in S$, pois $P \cup \{s\}$ é uma partição do intervalo $[a, s]$ nas condições desejadas.

Vai-se agora provar que $s = b$, o que terminará a demonstração do lema, uma vez que já se demonstrou que $s \in S$. Se não se tivesse $s = b$, então s seria menor do que b. Seja β como acima e seja t um elemento de $I_\beta \cap]s, b]$. Como $s \in S$, existe alguma partição P do intervalo $[a, s]$ tal que cada cada um dos seus intervalos está contido em A_α, para algum $\alpha \in J$. Mas então a partição $P \cup \{t\}$ do intervalo $[a, t]$ é tal que cada um dos seus intervalos está contido em A_α, para algum $\alpha \in J$. Isto quer dizer que $t \in S$, o que é absurdo, pois $t > s = \sup S$. ■

Como consequência do lema anterior, temos o

TEOREMA 1.1 (TEOREMA DE HEINE-BOREL) *Se $a, b \in \overline{\mathbb{R}}$ com $a < b$, então $[a, b]$ é compacto.*

DEMONSTRAÇÃO: Seja $(A_\alpha)_{\alpha \in J}$ uma cobertura aberta de $[a, b]$; quer-se provar que tem alguma subcobertura finita. Pelo lema 1.2, existe alguma partição $a = a_0 < a_1 < \cdots < a_n = b$ do intervalo $[a, b]$ tal que, para cada $k \in \{1, 2, \ldots, n\}$, $[a_{k-1}, a_k] \subset I_{\alpha(k)}$, para algum $\alpha(k) \in J$. Então

$$[a, b] = \bigcup_{k=1}^{n} [a_{k-1}, a_k] \subset \bigcup_{k=1}^{n} A_{\alpha(k)},$$

pelo que $(A_{\alpha(k)})_{k \in \{1,2,\ldots,n\}}$ é uma subcobertura finita de $(A_\alpha)_{\alpha \in J}$. ■

PROPOSIÇÃO 1.6 *A medida exterior de Lebesgue tem as seguintes propriedades:*

1. *se* $A \subset B \subset \mathbb{R}$*, então* $m^*(A) \leqslant m^*(B)$*;*

2. *se* $(A_n)_{n\in\mathbb{N}}$ *é uma sucessão de partes de* $\mathbb{R}$*, então*

$$m^*\left(\bigcup_{n\in\mathbb{N}} A_n\right) \leqslant \sum_{n=1}^{+\infty} m^*(A_n);$$

3. *se* I *é um intervalo de* $\mathbb{R}$*, então* $m^*(I) = \operatorname{comp}(I)$*;*

4. *se* $A \subset \mathbb{R}$ *e* $x \in \mathbb{R}$*, então* $m^*(A + x) = m^*(A)$*.*

DEMONSTRAÇÃO: A primeira alínea resulta de qualquer cobertura de B por intervalos abertos ser também uma cobertura de A por intervalos abertos.

A segunda alínea é trivial caso algum A_n tenha medida exterior infinita. Caso contrário, seja $r > 0$ e seja, para cada $k \in \mathbb{N}$, $(I_{k,n})_{n\in\mathbb{N}}$ uma sucessão de intervalos abertos cuja reunião contenha A_k e tal que $\sum_{n=1}^{+\infty} \operatorname{comp}(I_{k,n}) \leqslant m^*(A_k) + r/2^k$. Então a família $(I_{k,n})_{(k,n)\in\mathbb{N}^2}$ é uma família de intervalos abertos cuja reunião contém $\bigcup_{k\in\mathbb{N}} A_k$ e, portanto,

$$\begin{aligned} m^*\left(\bigcup_{k\in\mathbb{N}} A_k\right) &\leqslant \sum_{k=1}^{+\infty}\sum_{n=1}^{+\infty} \operatorname{comp}(I_{k,n}) \\ &\leqslant \sum_{k=1}^{+\infty}\left(m^*(A_k) + \frac{r}{2^k}\right) \\ &= \left(\sum_{k=1}^{+\infty} m^*(A_k)\right) + r. \end{aligned}$$

Como isto tem lugar para cada $r > 0$, está demonstrada a desigualdade da segunda alínea.

Seja agora I um intervalo de $\mathbb{R}$; quer-se provar que $m^*(I) = \operatorname{comp}(I)$. Comecemos pela desigualdade mais fácil de estabelecer, que é $m^*(I) \leqslant \operatorname{comp}(I)$. Esta é imediata caso I seja um intervalo aberto (pois nesse caso o conjunto $\{I\}$ já é uma família finita de intervalos abertos de $\mathbb{R}$ cuja reunião contém I e, portanto, $m^*(I) \leqslant \operatorname{comp}(I)$) ou quando I não é limitado. Caso I não seja aberto e seja limitado, é da forma $]a,b]$, da forma $[a,b[$ ou da forma $[a,b]$ ($a, b \in \mathbb{R}$). Veremos somente o primeiro caso, pois os outros dois são análogos. Se $\varepsilon > 0$, então $]a,b] \subset]a,b+\varepsilon[$ e, portanto, $m^*(]a,b]) \leqslant \operatorname{comp}(]a,b+\varepsilon[) = b - a + \varepsilon$. Como se tem esta desigualdade para qualquer $\varepsilon > 0$, tem-se $m^*(]a,b]) \leqslant b - a = \operatorname{comp}(]a,b])$.

Falta só ver que se tem sempre $m^*(I) \geqslant \mathrm{comp}(I)$. Vai-se começar por estabelecer esta desigualdade no caso em que I é um intervalo da forma $[a,b]$, com $a,b \in \mathbb{R}$ e $a \leqslant b$. Seja N um conjunto numerável e seja $(I_n)_{n\in N}$ uma família de intervalos abertos de $\mathbb{R}$ cuja reunião contenha $[a,b]$. Então, pelo lema 1.2, existe alguma partição $a = a_0 < a_1 < \cdots < a_n = b$ do intervalo $[a,b]$ tal que cada intervalo $[a_{k-1},a_k]$ está contido em algum $I_{n(k)}$ ($n(k) \in N$). Então

$$\begin{aligned}\sum_{n\in N} \mathrm{comp}(I_n) &\geqslant \sum_{k=1}^{n} \mathrm{comp}(I_{n(k)})\\ &\geqslant \sum_{k=1}^{n} \mathrm{comp}([a_{k-1},a_k])\\ &= \mathrm{comp}([a,b]).\end{aligned}$$

Vejamos agora o caso dos intervalos do tipo $]a,b]$, com $a,b \in \mathbb{R}$ e $a < b$. Se $\varepsilon \in]0,b-a[$, tem-se

$$\begin{aligned}[a+\varepsilon,b] \subset]a,b] \subset [a,b] &\Longrightarrow m^*([a+\varepsilon,b]) \leqslant m^*(]a,b]) \leqslant m^*([a,b])\\ &\Longleftrightarrow b-a-\varepsilon \leqslant m^*(]a,b]) \leqslant b-a.\end{aligned}$$

Mais uma vez, como isto tem lugar para qualquer $\varepsilon > 0$, $m^*(]a,b]) = b-a$. O mesmo argumento prova que $m^*([a,b[) = m^*(]a,b[) = b-a$.

Finalmente, se I for um intervalo não limitado, I pode ser obtido como reunião de uma sucessão $(I_n)_{n\in\mathbb{N}}$ de intervalos limitados cujo comprimento tende para $+\infty$, pelo que $m^*(I) \geqslant \lim_{n\in\mathbb{N}} m^*(I_n) = +\infty$.

Para demonstrar a quarta alínea, basta observar que se $(I_n)_{n\in\mathbb{N}}$ é uma sucessão de intervalos abertos cuja reunião contém A, então $(I_n + x)_{n\in\mathbb{N}}$ é uma sucessão de intervalos abertos cuja reunião contém $A + x$ e que se $(I_n)_{n\in\mathbb{N}}$ é uma sucessão de intervalos abertos cuja reunião contém $A + x$, então $(I_n - x)_{n\in\mathbb{N}}$ é uma sucessão de intervalos abertos cuja reunião contém A. Consequentemente, $m^*(A)$ e $m^*(A+x)$ são os ínfimos do mesmo conjunto e, portanto, são iguais. ∎

Se $X, A \in \mathcal{P}(\mathbb{R})$, então, uma vez que $X = (X \cap A) \cup (X \setminus A)$, sabe-se que

$$m^*(X) \leqslant m^*(X \cap A) + m^*(X \setminus A). \tag{1.2}$$

DEFINIÇÃO 1.7 Se $A \in \mathcal{P}(\mathbb{R})$, diz-se que A é *mensurável* se, para cada $X \in \mathcal{P}(\mathbb{R})$, $m^*(X) = m^*(X \cap A) + m^*(X \setminus A)$. O conjunto das partes mensuráveis de $\mathbb{R}$ representa-se por $\mathcal{M}(\mathbb{R})$.

Como será visto, $\mathcal{M}(\mathbb{R})$ é uma σ-álgebra e a restrição de m^* a $\mathcal{M}(\mathbb{R})$ é uma medida.

É conveniente observar que resulta da relação (1.2) que, a fim de se provar que um conjunto $A \in \mathcal{P}(\mathbb{R})$ é mensurável, basta provar que, para cada $X \in \mathcal{P}(\mathbb{R})$, $m^*(X \cap A) + m^*(X \setminus A) \leqslant m^*(X)$.

Resulta imediatamente da definição de $\mathcal{M}(\mathbb{R})$ que se $A \in \mathcal{M}(\mathbb{R})$, então $A^\complement \in \mathcal{M}(\mathbb{R})$. Como, obviamente, $\emptyset \in \mathcal{M}(\mathbb{R})$, a fim de se demonstrar que $\mathcal{M}(\mathbb{R})$ é uma σ-álgebra, só falta provar que é estável para reuniões numeráveis. Comecemos por ver que é estável para as reuniões finitas, ou seja, que $\mathcal{M}(\mathbb{R})$ é uma álgebra. Sejam então $A, B \in \mathcal{M}(\mathbb{R})$. Tem-se, para cada $X \in \mathcal{P}(\mathbb{R})$,

$$\begin{aligned} m^*(X) &\leqslant m^*(X \cap (A \cup B)) + m^*(X \setminus (A \cup B)) \\ &\leqslant m^*(X \cap A) + m^*\left(X \cap A^\complement \cap B\right) + m^*\left(X \cap A^\complement \cap B^\complement\right) \\ &= m^*(X \cap A) + m^*\left(X \cap A^\complement\right) \text{ (pois } B \in \mathcal{M}(\mathbb{R})) \\ &= m^*(X), \end{aligned}$$

pois A é mensurável. Está então provado que $\mathcal{M}(\mathbb{R})$ é uma álgebra.

Sejam agora $A_1, \ldots, A_n$ elementos de $\mathcal{M}(\mathbb{R})$ dois a dois disjuntos e $X \in \mathcal{P}(\mathbb{R})$; vai-se provar que

$$m^*\left(X \cap \bigcup_{j=1}^{n} A_j\right) = \sum_{j=1}^{n} m^*(X \cap A_j). \tag{1.3}$$

A demonstração será feita por indução. Naturalmente, caso $n = 1$ nada há a demonstrar. Vai-se então supor que $n > 1$ e que o resultado já está demonstrado para valores menores do que n. Então, uma vez que A_n é mensurável,

$$\begin{aligned} & m^*\left(X \cap \bigcup_{j=1}^{n} A_j\right) = \\ &= m^*\left(X \cap \left(\bigcup_{j=1}^{n} A_j\right) \cap A_n\right) + m^*\left(X \cap \left(\bigcup_{j=1}^{n} A_j\right) \setminus A_n\right) \\ &= m^*(X \cap A_n) + m^*\left(X \cap \bigcup_{j=1}^{n-1} A_j\right) \\ &= m^*(X \cap A_n) + \sum_{j=1}^{n-1} m^*(X \cap A_j) \end{aligned}$$

$$= \sum_{j=1}^{n} m^*(X \cap A_j).$$

Seja agora $(A_n)_{n\in\mathbb{N}}$ uma sucessão de partes mensuráveis de $\mathbb{R}$; quer-se provar que a sua reunião é mensurável. Para tal, seja $(B_n)_{n\in\mathbb{N}}$ a sucessão de partes de $\mathbb{R}$ tal que $B_1 = A_1$ e que

$$(foralln \in \mathbb{N} \setminus \{1\}) : B_n = A_n \setminus \bigcup_{j<n} A_j.$$

Então

1. como $\mathcal{M}(\mathbb{R})$ é uma algebra, cada B_n é mensurável;
2. os elementos da sucessão $(B_n)_{n\in\mathbb{N}}$ são dois a dois disjuntos;
3. $\bigcup_{n\in\mathbb{N}} B_n = \bigcup_{n\in\mathbb{N}} A_n$.

Tome-se agora $X \in \mathcal{P}(\mathbb{R})$. Então, para cada $n \in \mathbb{N}$,

$$\begin{aligned} m^*(X) &= m^*\left(X \cap \bigcup_{j=1}^{n} B_j\right) + m^*\left(X \Big\backslash \bigcup_{j=1}^{n} B_j\right) \\ &= \sum_{j=1}^{n} m^*(X \cap B_j) + m^*\left(X \Big\backslash \bigcup_{j=1}^{n} B_j\right) \text{ (por (1.3))} \\ &\geqslant \sum_{j=1}^{n} m^*(X \cap B_j) + m^*\left(X \Big\backslash \bigcup_{j=1}^{+\infty} B_j\right). \end{aligned}$$

Resulta de se ter esta desigualdade para cada $n \in \mathbb{N}$ que

$$\begin{aligned} m^*(X) &\geqslant \sum_{j=1}^{+\infty} m^*(X \cap B_j) + m^*\left(X \Big\backslash \bigcup_{j=1}^{+\infty} B_j\right) \\ &\geqslant m^*\left(X \cap \bigcup_{j=1}^{+\infty} B_j\right) + m^*\left(X \Big\backslash \bigcup_{j=1}^{+\infty} B_j\right). \end{aligned}$$

Está então provado que $\bigcup_{n\in\mathbb{N}} B_n$ $(= \bigcup_{n\in\mathbb{N}} A_n)$ é mensurável, o que conclui a demonstração da

PROPOSIÇÃO 1.7 *O conjunto $\mathcal{M}(\mathbb{R})$ forma uma σ-álgebra.*

PROPOSIÇÃO 1.8 *A restrição de m^* a $\mathcal{M}(\mathbb{R})$ é uma medida.*

DEMONSTRAÇÃO: Seja $(A_n)_{n\in\mathbb{N}}$ uma sucessão de partes mensuráveis de $\mathbb{R}$ disjuntas duas a duas; quer-se então provar que $m^*\left(\bigcup_{n\in\mathbb{N}} A_n\right) = \sum_{n=1}^{+\infty} m^*(A_n)$. Que se tem $m^*\left(\bigcup_{n\in\mathbb{N}} A_n\right) \leqslant \sum_{n=1}^{+\infty} m^*(A_n)$ resulta da proposição 1.6 (mesmo sem se estar a supor que os conjuntos A_n ($n \in \mathbb{N}$) são disjuntos dois a dois). Por outro lado, tem-se, para cada $N \in \mathbb{N}$,

$$m^*\left(\bigcup_{n\in\mathbb{N}} A_n\right) \geqslant m^*\left(\bigcup_{n=1}^{N} A_n\right) = \sum_{n=1}^{N} m^*(A_n),$$

por (1.3). Logo,

$$m^*\left(\bigcup_{n\in\mathbb{N}} A_n\right) \geqslant \lim_{N\in\mathbb{N}} \sum_{n=1}^{N} m^*(A_n) = \sum_{n=1}^{+\infty} m^*(A_n).$$

■

DEFINIÇÃO 1.8 Designa-se por *medida de Lebesgue* e representa-se por l a restrição a $\mathcal{M}(\mathbb{R})$ da medida exterior de Lebesgue.

Tem-se então uma medida l definida numa σ-álgebra de partes de $\mathbb{R}$ e quer-se mostrar que é adequada para definir um integral que generalize o de Riemann, mas ainda só se provou que se $\mathcal{M}(\mathbb{R})$ contém algum intervalo I, então $l(I) = \operatorname{comp}(I)$.

PROPOSIÇÃO 1.9 *Qualquer parte boreliana de* $\mathbb{R}$ *é mensurável.*

DEMONSTRAÇÃO: Visto que, como foi observado na página 6, $\mathcal{B}(\mathbb{R})$ é a σ-álgebra gerada pelos intervalos abertos de $\mathbb{R}$, basta provar que qualquer intervalo aberto de $\mathbb{R}$ é mensurável.

Seja I um intervalo não limitado de $\mathbb{R}$; quer-se provar que se $X \in \mathcal{P}(\mathbb{R})$, então $m^*(X) = m^*(X \cap I) + m^*(X \setminus I)$. Para demonstrar isso, vai-se recorrer à observação feita na página 10 segundo a qual a medida exterior de um conjunto $A \in \mathcal{P}(\mathbb{R})$ é o ínfimo do conjunto das somas $\sum_{n\in\mathbb{N}} \operatorname{comp}(I_n)$, onde $(I_n)_{n\in\mathbb{N}}$ é uma sucessão de intervalos cuja reunião contém A. Seja então $(I_n)_{n\in\mathbb{N}}$ uma sucessão de intervalos cuja reunião contenha X. Para cada $n \in \mathbb{N}$ sejam $I_n' = I_n \cap I$ e $I_n'' = I_n \setminus I$; então I_n' e I_n'' são intervalos[2] disjuntos cuja reunião é igual a I_n e, consequentemente, $\operatorname{comp}(I_n) = \operatorname{comp}(I_n') + \operatorname{comp}(I_n'')$. Resulta então da proposição 1.6 que $m^*(I_n) = m^*(I_n') + m^*(I_n'')$. Como, por outro lado,

$$m^*(X \cap I) \leqslant \sum_{n=1}^{+\infty} m^*(I_n') \quad \text{e} \quad m^*(X \setminus I) \leqslant \sum_{n=1}^{+\infty} m^*(I_n''),$$

[2]Que I_n'' é intervalo resulta de se estar a supor que I não é limitado, o que implica que $I^{\complement}$ é um intervalo.

deduz-se que

$$m^*(X \cap I) + m^*(X \setminus I) \leqslant \sum_{n=1}^{+\infty} \big(m^*(I'_n) + m^*(I''_n)\big) = \sum_{n=1}^{+\infty} m^*(I_n).$$

Como esta desigualdade tem lugar para cada sucessão $(I_n)_{n\in\mathbb{N}}$ de intervalos cuja reunião contenha X, isto prova que $m^*(X \cap I) + m^*(X \setminus I) \leqslant m^*(X)$.

Está então provado que $\mathcal{M}(\mathbb{R})$ contém todos os intervalos não limitados. Mas qualquer intervalo pode ser obtido como intersecção de dois intervalos não limitados, pelo que $\mathcal{M}(\mathbb{R})$ contém todos os intervalos, de onde se deduz, como já se viu, que contém todos os borelianos. ∎

Tem-se então $\mathcal{B}(\mathbb{R}) \subset \mathcal{M}(\mathbb{R}) \subset \mathcal{P}(\mathbb{R})$. Alguma destas inclusões será, de facto, uma igualdade? Será visto que a resposta é negativa em ambos os casos.

Se E for uma parte mensurável de $\mathbb{R}$ e se A for um aberto de $\mathbb{R}$ que contenha E, então A é mensurável (pela proposição 1.9) e $l(E) \leqslant l(A)$, pelo que

$$l(E) \leqslant \inf\{\, l(A) \mid A \text{ é aberto e } E \subset A \,\}.$$

De facto, a desigualdade anterior é uma igualdade. Isto é trivial caso $l(E) = +\infty$ e, caso contrário, sabe-se, pela definição da medida exterior de Lebesgue, que, para cada $\varepsilon > 0$, existe alguma sucessão $(I_n)_{n\in\mathbb{N}}$ de intervalos abertos cuja reunião contém E e tal que $\sum_{n=1}^{+\infty} \operatorname{comp}(I_n) < l(E) + \varepsilon$. Logo, se $A = \bigcup_{n\in\mathbb{N}} I_n$, então A é um aberto de $\mathbb{R}$ que contém E e

$$l(A) = l\left(\bigcup_{n\in\mathbb{N}} I_n\right) \leqslant \sum_{n=1}^{+\infty} l(I_n) = \sum_{n=1}^{+\infty} \operatorname{comp}(I_n) < l(E) + \varepsilon.$$

Vai-se demonstrar agora um resultado análogo.

PROPOSIÇÃO 1.10 *Se E for uma parte mensurável de $\mathbb{R}$, então*

$$l(E) = \sup\{\, l(K) \mid K \text{ é compacto e } E \supset K \,\}. \tag{1.4}$$

DEMONSTRAÇÃO: Comece-se por supor que E é limitado. Sejam $a, b \in \mathbb{R}$ tais que $E \subset [a,b]$ e seja $F = [a,b] \setminus E$. Vai-se provar que

$$l(F) = \inf\{\, l(A) \mid [a,b] \supset A \supset F \text{ e } [a,b] \setminus A \text{ é compacto} \,\}. \tag{1.5}$$

É imediato que (1.4) equivale a (1.5), pois, dados dois conjuntos $A, B \in \mathcal{M}(\mathbb{R})$ tais que $A \subset B$, tem-se $l(B) + l(A \setminus B) = l(A)$. Observe-se que se $[a,b] \supset A$ e $[a,b] \setminus A$ é compacto, então, em particular, $[a,b] \setminus A$ é um fechado de $[a,b]$, pelo que A é um aberto de $[a,b]$. Reciprocamente, se A é um aberto de $[a,b]$, então $[a,b] \setminus A$, sendo fechado e limitado, é compacto. Logo, (1.5) equivale a

$$l(F) = \inf \{ \, l(A) \mid A \supset F \text{ e } A \text{ é um aberto de } [a,b] \, \}.$$

Seja $\varepsilon > 0$ e seja $(I_n)_{n\in\mathbb{N}}$ uma família de intervalos abertos cuja reunião contenha F e tal que $\sum_{n=1}^{\infty} l(I_n) < l(F) + \varepsilon$. Então

$$l\left(\bigcup_{n\in\mathbb{N}} (I_n \cap [a,b])\right) \leqslant l\left(\bigcup_{n\in\mathbb{N}} I_n\right) \leqslant \sum_{n=1}^{\infty} l(I_n) < l(F) + \varepsilon.$$

Como $\bigcup_{n\in\mathbb{N}} (I_n \cap [a,b])$ é um aberto de $[a,b]$ que contém F, deduz-se que

$$\inf \{ \, l(A) \mid A \supset F \text{ e } A \text{ é um aberto de } [a,b] \, \} < l(F) + \varepsilon$$

e, como isto ocorre para cada $\varepsilon > 0$,

$$\inf \{ \, l(A) \mid A \supset F \text{ e } A \text{ é um aberto de } [a,b] \, \} \leqslant l(F).$$

A desigualdade oposta é imediata, pois se $A \supset F$, então $l(A) \geqslant l(F)$.

Suponha-se que $l(E) < +\infty$ mas E não é necessariamente limitado. Seja, para cada $n \in \mathbb{N}$, $E_n = E \cap [-n,n]$. Então, visto que a sucessão $(E_n)_{n\in\mathbb{N}}$ é crescente e a sua reunião é E, $l(E) = \lim_{n\in\mathbb{N}} l(E_n)$, pela proposição 1.5. Seja $\varepsilon > 0$ e seja $n \in \mathbb{N}$ tal que $l(E) - l(E_n) < \varepsilon/2$. Pela alínea anterior, existe algum compacto $K \subset E_n$ tal que $l(E_n) - l(K) < \varepsilon/2$. Logo, $l(E) < l(K) + \varepsilon$. Como isto tem lugar para cada $\varepsilon > 0$,

$$l(E) \leqslant \sup \{ \, l(K) \mid K \text{ é compacto e } E \supset K \, \}.$$

A desigualdade oposta é trivial.

Caso $l(E) = +\infty$, seja E_n como no parágrafo anterior. Como se tem $\lim_{n\in\mathbb{N}} l(E_n) = +\infty$ e como, para cada $n \in \mathbb{N}$, E_n contém compactos com medida de Lebesgue tão próxima quanto se queira de $l(E_n)$, há compactos contidos em E com medida de Lebesgue tão grande quanto se queira, pelo que $l(E) = +\infty = \sup \{ \, l(K) \mid K \subset E \text{ e } K \text{ é compacto} \, \}$. ∎

PROPOSIÇÃO 1.11 *Qualquer parte de $\mathbb{R}$ com medida exterior nula é mensurável.*

DEMONSTRAÇÃO: Se $A \in \mathcal{P}(\mathbb{R})$ for tal que $m^*(A) = 0$ e se $X \in \mathcal{P}(\mathbb{R})$, então, como $X \cap A \subset A$, $m^*(X \cap A) \leqslant m^*(A) = 0$; logo, $m^*(X \cap A) = 0$. Por outro lado, $X \setminus A \subset X$, pelo que $m^*(X \setminus A) \leqslant m^*(X)$. Então $m^*(X \cap A) + m^*(X \setminus A) \leqslant m^*(X)$. ∎

1.3 O conjunto de Cantor

1.3.1 Definição e propriedades básicas

DEFINIÇÃO 1.9 Para cada $\alpha \in]0,1]$, define-se o conjunto C_α como sendo a intersecção $\bigcap_{n\in\mathbb{Z}_+} I_n$, onde $(I_n)_{n\in\mathbb{Z}_+}$ é a sucessão de reuniões de um número finito de intervalos fechados dois a dois disjuntos tal que:

1. $I_0 = [0,1]$;

2. I_1 é o conjunto que se obtém retirando de I_0 o intervalo aberto central de comprimento $\alpha/3$, ou seja,
$$I_1 = [0,1] \setminus]1/2 - \alpha/6, 1/2 + \alpha/6[= [0, 1/2 - \alpha/6] \cup [1/2 + \alpha/6, 1];$$

3. I_2 é o conjunto que se obtém retirando de I_1 o intervalo aberto central de comprimento $\alpha/9$ de cada um dos intervalos que o formam;

4. mais geralmente, para cada $n \in \mathbb{N}$ o conjunto I_n é reunião disjunta de 2^n intervalos fechados e o conjunto I_{n+1} obtém-se de I_n retirando de cada um daqueles intervalos o intervalo central aberto de comprimento $\alpha/3^{n+1}$.

O conjunto C_1 designa-se por *conjunto de Cantor* e representa-se por C.

Veja-se a figura 1.2.

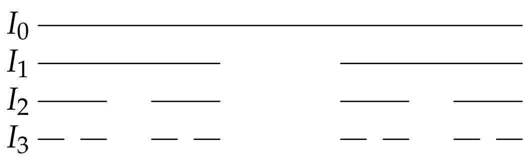

Figura 1.2: Construção do conjunto $C_{3/4}$

Há uma passagem nesta definição cuja legitimidade exige uma demonstração. Para que a última alínea faça sentido, é necessário demonstrar que o comprimento de cada um dos 2^n intervalos fechados cuja reunião disjunta forma I_n é maior de que $\alpha/3^{n+1}$; caso contrário, não se pode retirar deles um «intervalo aberto central de comprimento $\alpha/3^{n+1}$». Para justificar a passagem, repare-se que o conjunto I_1 é obtido retirando-se de $[0,1]$ um segmento de comprimento $\alpha/3$; logo, $l(I_1) = 1 - \alpha/3$. Em seguida, obtém-se I_2 retirando de I_1 dois segmentos de comprimento $\alpha/9$, pelo que $l(I_2) = 1 - \alpha/3 - (2\alpha)/9$. Mais geralmente, tem-se:

$$(\forall n \in \mathbb{Z}_+) : l(I_n) = 1 - \sum_{k=1}^{n} \frac{2^{k-1}\alpha}{3^k} = 1 - \alpha\left(1 - \frac{2^n}{3^n}\right) \tag{1.6}$$

e então o que se quer mostrar é que:

$$(\forall n \in \mathbb{Z}_+) : \frac{1}{2^n}\left(1 - \alpha\left(1 - \frac{2^n}{3^n}\right)\right) > \frac{\alpha}{3^{n+1}}.$$

Verifica-se facilmente que esta expressão equivale a

$$(\forall n \in \mathbb{Z}_+) : \frac{1-\alpha}{2^n} > -\frac{2\alpha}{3^{n+1}}$$

e esta última proposição é obviamente verdadeira.[3]

PROPOSIÇÃO 1.12 *Para cada $\alpha \in]0,1]$, o conjunto C_α é mensurável e a sua medida é $1 - \alpha$. Em particular, o conjunto de Cantor tem medida nula.*

DEMONSTRAÇÃO: Cada I_n ($n \in \mathbb{Z}_+$) é boreliano e, em particular, é mensurável. Logo, C_α é mensurável, por ser a intersecção de uma sucessão numerável de conjuntos mensuráveis. Resulta da segunda alínea da proposição 1.5 que

$$\begin{aligned} l(C_\alpha) &= \lim_{n\in\mathbb{N}} l(I_n) \\ &= \lim_{n\in\mathbb{N}} \left(1 - \alpha\left(1 - \frac{2^n}{3^n}\right)\right) \text{ por (1.6)} \\ &= 1 - \alpha. \end{aligned}$$

■

Os conjuntos C_α com $\alpha \in]0,1[$ designam-se geralmente por *conjuntos de Cantor gordos*. A proposição anterior explica esta terminologia; são «gordos» pelo facto de terem medida positiva.

PROPOSIÇÃO 1.13 *Para cada $\alpha \in]0,1]$, C_α é fechado e tem interior vazio.*

DEMONSTRAÇÃO: Que C_α ($\alpha \in]0,1]$) é fechado resulta de ser definido como a intersecção de uma sucessão de conjuntos fechados. Por outro lado, se o interior de C_α não fosse vazio, então C_α conteria algum intervalo $]a,b[$ ($a, b \in \mathbb{R}$, $a < b$). Um tal intervalo teria que estar contido em cada

[3] Também se deduz desta expressão que α não pode ser maior do que 1.

um dos 2^n intervalos dois a dois disjuntos cuja reunião forma I_n ($n \in \mathbb{Z}_+$) e o comprimento de cada um destes intervalos é $2^{-n}\left(1-\alpha\left(1-\left({}^{2^n}/_{3^n}\right)\right)\right)$, pelo que

$$b-a = l\big(]a,b[\big) \leqslant 2^{-n}\left(1-\alpha\left(1-\left(\frac{2^n}{3^n}\right)\right)\right).$$

Como isto tem lugar para cada $n \in \mathbb{Z}_+$

$$b-a \leqslant \lim_{n\in\mathbb{Z}_+} 2^{-n}\left(1-\alpha\left(1-\left(\frac{2^n}{3^n}\right)\right)\right) = 0,$$

o que é absurdo. ∎

Naturalmente, resulta da proposição anterior que cada C_α é compacto, visto que é fechado e limitado.

PROPOSIÇÃO 1.14 *O cardinal cada conjunto C_α é igual ao de $\mathbb{R}$.*

DEMONSTRAÇÃO: O conjunto I_1 é a reunião de dois intervalos fechados disjuntos, que serão representados por $I_{(0)}$ e $I_{(1)}$; estes intervalos ficam completamente determinados se se convencionar que cada elemento de $I_{(0)}$ é menor do que qualquer elemento de $I_{(1)}$.

Por sua vez, cada intervalo $I_{(i)}$ ($i \in \{0,1\}$) é reunião disjunta de dois intervalos fechados, $I_{(i,0)}$ e $I_{(i,1)}$; mais uma vez, estes intervalos ficam completamente determinados se se convencionar que cada elemento do primeiro é menor do que qualquer elemento do segundo.

Prosseguindo assim, define-se uma família $\{\, I_\beta \mid \beta \in \bigcup_{n\in\mathbb{N}}\{0,1\}^n \,\}$ de intervalos fechados. Resulta da definição desta família que se tem

$$I_{(a_1,\dots,a_n)} \supset I_{(b_1,\dots,b_m)}$$

se e só se $n \leqslant m$ e $(\forall i \in \{1,\dots,n\}) : a_i = b_i$.

Seja $s = (a_n)_{n\in\mathbb{N}}$ uma sucessão de elementos de $\{0,1\}$ e, para cada $n \in \mathbb{N}$, seja $s(n) = (a_1, a_2, \dots, a_n)$. Então

$$I_{s(1)} \supset I_{s(2)} \supset I_{s(3)} \supset \cdots,$$

e, como cada $I_{s(n)}$ é um intervalo fechado não vazio, $\bigcap_{n\in\mathbb{N}} I_{s(n)}$ não é vazio. De facto, aquele conjunto tem um único elemento, visto que, por (1.6), os comprimentos dos intervalos $I_{s(n)}$ tendem para 0. Seja então ψ a função de $\{0,1\}^{\mathbb{N}}$ (i. e. o conjunto das sucessões de elementos de $\{0,1\}$) em C_α tal que

$$(\forall s \in \{0,1\}^{\mathbb{N}}) : \bigcap_{n\in\mathbb{N}} I_{s(n)} = \{\psi(s)\}.$$

Então ψ é uma bijecção e, portanto, o cardinal de C_α é igual ao de $\{0,1\}^{\mathbb{N}}$. Este último conjunto tem o mesmo cardinal que $\mathbb{R}$ (veja-se o corolário A.1 do apêndice A). ■

Vejamos uma outra demonstração da proposição anterior no caso específico do conjunto de Cantor. Observe-se que neste caso

- I_1 é o conjunto dos números reais que podem ser escritos sob a forma ${}^{a_1}/_3 + r_1$ com $a_1 \in \{0,2\}$ e $r_i \in [0, {}^1/_3]$;
- I_2 é o conjunto dos números reais que podem ser escritos sob a forma ${}^{a_1}/_3 + {}^{a_2}/_9 + r_2$ com $a_1, a_2 \in \{0,2\}$ e $r_2 \in [0, {}^1/_9]$

e mais geralmente, que se $n \in \mathbb{N}$

$$I_n = \left\{ \left(\sum_{k=1}^{n} \frac{a_k}{3^k} \right) + r_n \;\middle|\; \{a_1, \ldots, a_n\} \subset \{0,2\} \text{ e } r_n \in \left[0, 3^{-n}\right] \right\}.$$

Além disso, se $x \in I_n$ e se $a_1, \ldots, a_n \in \{0,2\}$ e $r_n \in [0, 3^{-n}]$ são tais que $x = (\sum_{k=1}^{n} {}^{a_k}/_{3^k}) + r_n$, então os $(a_k)_{k \in \{1,\ldots,n\}}$ e r_n são únicos. De facto, isto é óbvio para $n = 1$ e se, dado $n > 1$, esta afirmação já estiver provada para números naturais inferiores a n, se se tivesse

$$\left(\sum_{k=1}^{n} \frac{a_k}{3^k} \right) + r_n = \left(\sum_{k=1}^{n} \frac{b_k}{3^k} \right) + s_n$$

para números $a_1, b_1, \ldots, a_n, b_n \in \{0,2\}$ e $r_n, s_n \in [0, 3^{-n}]$, então, como se tem

$$\left(\sum_{k=1}^{n-1} \frac{a_k}{3^k} \right) + \frac{a_n}{3^n} + r_n = \left(\sum_{k=1}^{n-1} \frac{b_k}{3^k} \right) + \frac{b_n}{3^n} + s_n$$

e como ${}^{a_n}/_{3^n} + r_n, {}^{b_n}/_{3^n} + s_n \in [0, 3^{-n+1}]$, resulta da hipótese de indução que $a_k = b_k$ quando $k < n$ e que ${}^{a_n}/_{3^n} + r_n = {}^{b_n}/_{3^n} + s_n$. Se $a_n \neq b_n$, então $a_n - b_n = \pm 2$, pelo que se teria $r_n - s_n = \pm {}^2/_{3^n}$, o que é impossível, uma vez que $r_n, s_n \in [0, 3^{-n}]$.

Sendo assim, C é formado pelos números reais que podem ser escritos sob a forma $\sum_{n=1}^{+\infty} {}^{a_n}/_{3^n}$, com cada a_n igual a 0 ou a 2. Além disso, cada elemento de C pode ser escrito sob aquela forma de uma só maneira.[4]

[4]Esta maneira de descrever C é a que foi originalmente empregue pelo próprio Cantor [9, §3.1] e é equivalente à afirmação de que C é formado pelos números do intervalo $[0,1]$ que podem ser escritos na base 3 usando unicamente os algarismos 0 e 2.

É agora fácil provar que C e $\mathbb{R}$ têm o mesmo cardinal. Basta ver que C tem o mesmo cardinal que o conjunto das sucessões de zeros e dois, o qual tem o mesmo cardinal que $\mathbb{R}$, pelo corolário A.1 do apêndice A. Recorrendo à demonstração daquele corolário, pode-se fazer uma construção mais explícita. Para cada $x \in C$, seja $(a_n(x))_{n\in\mathbb{N}}$ a sucessão de elementos de $\{0,2\}$ tal que $x = \sum_{n=1}^{+\infty} a_n(x)/3^n$; então a função

$$\begin{aligned} f\colon C &\longrightarrow [0,1] \\ x &\mapsto \sum_{n=1}^{+\infty} \frac{a_n(x)}{2^{n+1}} = \sum_{n=1}^{+\infty} \frac{a_n(x)/2}{2^n} \end{aligned}$$

é sobrejectiva, visto que cada elemento de $[0,1]$ pode ser escrito sob a forma $\sum_{n=1}^{+\infty} b_n/2^n$, com cada b_n igual a 0 ou a 1.[5] Além disso, para cada $x \in [0,1[$ isto pode ser feito exactamente de uma maneira se se excluirem as sucessões $(b_n)_{n\in\mathbb{N}}$ tais que $b_n = 1$ para cada n suficientemente pequeno, as quais formam um conjunto numerável. Mas isto é o mesmo que afirmar que C pode ser escrito como reunião de um conjunto numerável com um conjunto que tem o mesmo cardinal que $[0,1[$ (que tem o mesmo cardinal que $\mathbb{R}$) e resulta desta observação que têm C e $\mathbb{R}$ têm o mesmo cardinal.

1.3.2 Aplicações à medida de Lebesgue

PROPOSIÇÃO 1.15 *Os conjuntos $\mathcal{M}(\mathbb{R})$ e $\mathcal{P}(\mathbb{R})$ têm o mesmo cardinal.*

DEMONSTRAÇÃO: Visto que $\mathcal{M}(\mathbb{R}) \subset \mathcal{P}(\mathbb{R})$, é claro que o cardinal de $\mathcal{M}(\mathbb{R})$ é menor ou igual ao de $\mathcal{P}(\mathbb{R})$. Por outro lado, sabe-se, pela proposição 1.11 e pelo facto de o conjunto de Cantor ter medida nula, que $\mathcal{M}(\mathbb{R}) \supset \mathcal{P}(C)$. Mas, como C e $\mathbb{R}$ têm o mesmo cardinal, resulta que o cardinal de $\mathcal{M}(\mathbb{R})$ é maior ou igual ao de $\mathcal{P}(\mathbb{R})$. ∎

Esta proposição pode ser empregue para provar que $\mathcal{B}(\mathbb{R}) \subsetneq \mathcal{M}(\mathbb{R})$ pois, como foi mencionado na página 6, $\mathcal{B}(\mathbb{R})$ tem o mesmo cardinal que $\mathbb{R}$. Também há demonstrações de que $\mathcal{B}(\mathbb{R}) \subsetneq \mathcal{M}(\mathbb{R})$ que não se baseiam neste argumento de cardinalidade; veja-se [6, §8.2–4], por exemplo.

Ainda resta estabelecer que $\mathcal{M}(\mathbb{R}) \subsetneq \mathcal{P}(\mathbb{R})$.

PROPOSIÇÃO 1.16 *Não existe nenhuma medida $m\colon \mathcal{P}(\mathbb{R}) \longrightarrow \overline{\mathbb{R}_+}$ tal que:*

1. *se I é um intervalo de $\mathbb{R}$, então $m(I) = \operatorname{comp}(I)$;*

[5] Posto de outro modo, cada elemento de $[0,1]$ pode ser escrito na base 2 sob a forma $0{,}b_1b_2b_3\ldots$; veja-se o apêndice A.

2. $(\forall A \in \mathcal{P}(\mathbb{R}))(\forall x \in \mathbb{R}) : m(A+x) = m(A)$.

DEMONSTRAÇÃO: Considere-se no conjunto $[0,1]$ a relação de equivalência $\sim$ assim definida: $x \sim y$ se e só se $x - y \in \mathbb{Q}$. Seja $V \subset [0,1]$ um conjunto que contenha um e um só elemento de cada classe de equivalência. Vai-se provar que

$$[0,1] \subset \bigcup_{q \in \mathbb{Q} \cap [-1,1]} V + q \subset [-1,2] \tag{1.7}$$

e que a reunião é disjunta. Resultará destes factos que

$$1 = m([0,1]) \leqslant m\left(\bigcup_{q \in \mathbb{Q} \cap [-1,1]} V + q\right) \leqslant m([-1,2]) = 3$$

e, portanto, que

$$1 \leqslant \sum_{q \in \mathbb{Q} \cap [-1,1]} m(V+q) \leqslant 3 \iff 1 \leqslant \sum_{q \in \mathbb{Q} \cap [-1,1]} m(V) \leqslant 3.$$

Mas isto é impossível, pois $m(V) = 0$ ou $m(V) > 0$. No primeiro caso, $\sum_{q \in \mathbb{Q} \cap [-1,1]} m(V) = 0$ e, no segundo, $\sum_{q \in \mathbb{Q} \cap [-1,1]} m(V) = +\infty$.

Passemos então à demonstração de (1.7) e de que a reunião é disjunta. Esta última afirmação resulta do seguinte facto: se se tivesse $v_1 + q_1 = v_2 + q_2$ com $v_1, v_2 \in V$ e $q_1, q_2 \in \mathbb{Q}$, então ter-se-ia $v_1 - v_2 = q_2 - q_1 \in \mathbb{Q}$, pelo que $v_1 \sim v_2$ e, portanto, $v_1 = v_2$. Mas então $q_1 = q_2$.

Seja agora $x \in [0,1]$; quer-se provar que x pode ser escrito sob a forma $v + q$ para algum $v \in V$ e para algum $q \in \mathbb{Q} \cap [-1,1]$. Seja v o único elemento de V tal que $x \sim v$ e seja $q = x - v$. Então $q \in \mathbb{Q}$ e, uma vez que $x, v \in [0,1]$, $q \in [-1,1]$.

Finalmente, se $v \in V$ e se $q \in \mathbb{Q} \cap [-1,1]$, é claro que $v + q \in [-1,2]$, visto que $v \in [0,1]$. ∎

É consequência imediata que se tem o

COROLÁRIO 1.1 *Há partes de* $\mathbb{R}$ *não mensuráveis.*

Naturalmente, a proposição anterior mostra que isto não resulta de alguma deficiência na maneira como foi definido o conceito de conjunto mensurável, mas sim que resulta da natureza dos conceitos com que se está a trabalhar.

Resulta deste corolário que, tal como foi afirmado na página 10, a medida exterior de Lebesgue não é uma medida.

Se se examinar a demonstração da proposição 1.16, fica claro que o conjunto V que aí surge não é mensurável. É interessante observar que a definição de V envolveu o axioma da escolha. De facto, este pode ser enunciado do seguinte modo: dado um conjunto $\{ X_i \mid i \in I \}$, em que, para cada $i \in I$, X_i é um conjunto não vazio, existe algum conjunto $\{ x_i \mid i \in I \}$ tal que $(\forall i \in I) : x_i \in X_i$. É precisamente este facto que se usa implicitamente para garantir que existe um conjunto que tem um e um só ponto de cada classe de equivalência da relação de equivalência $\sim$. Convém assinalar que se pode provar que sem o axioma da escolha não se pode demonstrar a existência de conjuntos não mensuráveis; veja-se [15]. Para mais detalhes sobre este assunto, veja-se [5].

1.4 Aplicações ao integral de Riemann

1.4.1 Definição e propriedades elementares

Vai-se começar por definir de uma maneira bastante sucinta o integral de Riemann. Para uma abordagem mais detalhada e com exemplos, veja-se [13, cap. 6] ou [16, cap. 13].

DEFINIÇÃO 1.10 Sejam $a, b \in \mathbb{R}$ tais que $a < b$, seja f uma função limitada de $[a, b]$ em $\mathbb{R}$ e seja P uma partição de $[a, b]$. Se $P = \{a_0, a_1, \ldots, a_n\}$ com $a = a_0 < a_1 < a_2 < \cdots < a_n = b$, define-se então a *soma superior* e a *soma inferior* de f relativamente a P como sendo os números

$$\overline{\Sigma}(f, P) = \sum_{k=1}^{n} \sup f([a_{k-1}, a_k])(a_k - a_{k-1})$$

e

$$\underline{\Sigma}(f, P) = \sum_{k=1}^{n} \inf f([a_{k-1}, a_k])(a_k - a_{k-1})$$

respectivamente.

PROPOSIÇÃO 1.17 *Sejam $a, b \in \mathbb{R}$ com $a < b$, seja f uma função limitada de $[a, b]$ em $\mathbb{R}$ e sejam P e Q partições de $[a, b]$. Então:*

1. *se $P \subset Q$, tem-se $\underline{\Sigma}(f, P) \leqslant \underline{\Sigma}(f, Q)$ e $\overline{\Sigma}(f, P) \geqslant \overline{\Sigma}(f, Q)$;*
2. *$\underline{\Sigma}(f, P) \leqslant \overline{\Sigma}(f, Q)$.*

DEMONSTRAÇÃO: A primeira alínea demonstra-se por indução relativamente ao cardinal de $Q \setminus P$. Se for igual a 0, então $Q = P$ e nada há a demonstrar. Por outro lado, para demonstrar o passo de indução, basta observar que se $c, c', c'' \in [a, b]$ com $c < c' < c''$, então

$$\begin{aligned}\inf f([c,c''])(c''-c) &= \inf f([c,c''])(c'-c) + \inf f([c,c''])(c''-c')\\ &\leqslant \inf f([c,c'])(c'-c) + \inf f([c',c''])(c''-c')\end{aligned}$$

e

$$\begin{aligned}\sup f([c,c''])(c''-c) &= \sup f([c,c''])(c'-c) + \sup f([c,c''])(c''-c')\\ &\geqslant \sup f([c,c'])(c'-c) + \sup f([c',c''])(c''-c').\end{aligned}$$

A segunda alínea resulta de se ter

$$\underline{\Sigma}(f,P) \leqslant \underline{\Sigma}(f,P\cup Q) \leqslant \overline{\Sigma}(f,P\cup Q) \leqslant \overline{\Sigma}(f,Q).$$ ■

Sendo assim, qualquer soma inferior é menor ou igual que qualquer soma superior e, consequentemente, o supremo do conjunto das somas inferiores é menor ou igual ao ínfimo do conjunto das somas superiores.

DEFINIÇÃO 1.11 Sejam $a, b \in \mathbb{R}$ com $a < b$ e seja f uma função limitada de $[a, b]$ em $\mathbb{R}$. Diz-se que a função f é *integrável segundo Riemann* se o supremo do conjunto das somas inferiores for igual ao ínfimo do conjunto das somas superiores. Este número designa-se então por integral de Riemann da função f e representa-se por $\int_a^b f(x)\,dx$.

TEOREMA 1.2 (CRITÉRIO DE INTEGRABILIDADE DE RIEMANN) *Se $a, b \in \mathbb{R}$ com $a < b$ e se f é uma função limitada de $[a, b]$ em $\mathbb{R}$, então f é integrável segundo Riemann se e só se, para cada $\varepsilon > 0$, existir alguma partição P de $[a, b]$ tal que $\overline{\Sigma}(f,P) - \underline{\Sigma}(f,P) < \varepsilon$.*

DEMONSTRAÇÃO: Se f for integrável segundo Riemann e se $\varepsilon > 0$, então existem partições P e Q de $[a, b]$ tais que $\overline{\Sigma}(f,P) - \underline{\Sigma}(f,Q) < \varepsilon$. Mas resulta então da proposição 1.17 que $\overline{\Sigma}(f,P\cup Q) - \underline{\Sigma}(f,P\cup Q) < \varepsilon$.

Reciprocamente, se a condição de enunciado se verificar e se se representar por s (respectivamente S) o supremo do conjunto das somas inferiores (resp. superiores) então, dado $\varepsilon > 0$, como há alguma partição P tal que $\overline{\Sigma}(f,P) - \underline{\Sigma}(f,P) < \varepsilon$ e como $\underline{\Sigma}(f,P) \leqslant s \leqslant S \leqslant \overline{\Sigma}(f,P)$, tem-se que $S - s < \varepsilon$; como isto tem lugar para qualquer $\varepsilon > 0$, $s = S$. ■

1.4.2 Oscilação

DEFINIÇÃO 1.12 Se $A \subset \mathbb{R}$, se f é uma função de A em $\mathbb{R}$ e se $B \subset A$, define-se a *oscilação de f em B* e representa-se por $o(f,B)$ o número

$$\sup f(B) - \inf f(B) = \sup \{|f(x) - f(y)| \mid x, y \in B\}.$$

Se $x \in A$, define-se a *oscilação de f no ponto x* como sendo

$$o_f(x) = \inf \left\{ o\big(f, A \cap]a,b[\big) \;\middle|\; a, b \in \mathbb{R} \text{ e } a < x < b \right\}.$$

EXEMPLO 1.4 Se $f \colon \mathbb{R} \longrightarrow \mathbb{R}$ for a função identidade, então o_f é a função nula, pois se $x \in \mathbb{R}$ e se $\varepsilon > 0$, então $o(f,]x - \varepsilon/2, x + \varepsilon/2[) = \varepsilon$.

EXEMPLO 1.5 Considere-se a função

$$\begin{aligned} f \colon \mathbb{R} &\longrightarrow \mathbb{R} \\ x &\mapsto \begin{cases} 0 & \text{se } x < 0 \\ 1 & \text{caso contrário.} \end{cases} \end{aligned}$$

Então $o_f(x) = 0$ para cada $x \neq 0$ e $o_f(0) = 1$.

PROPOSIÇÃO 1.18 *Se $A \subset \mathbb{R}$ e f é uma função de A em $\mathbb{R}$, então os pontos de A onde f é contínua são exactamente aqueles onde a oscilação é nula.*

DEMONSTRAÇÃO: Seja $x \in A$. Se f for contínua no ponto x, seja $\varepsilon > 0$; vai-se provar que $o_f(x) < \varepsilon$. Seja $\delta > 0$ tal que

$$(\forall y \in A) : |y - x| < \delta \Longrightarrow |f(y) - f(x)| < \frac{\varepsilon}{2}.$$

Então $o(f, A \cap]x - \delta, x + \delta[) < \varepsilon$, pois se $y_1, y_2 \in A \cap]x - \delta, x + \delta[$, então $|f(y_1) - f(y_2)| \leqslant |f(y_1) - f(x)| + |f(x) - f(y_2)| < \varepsilon$. Resulta então da definição de $o_f(x)$ que $o_f(x) < \varepsilon$.

Reciprocamente, suponha-se que f tem oscilação nula num ponto $x \in A$. Seja $\varepsilon > 0$ e sejam $a, b \in \mathbb{R}$ tais que $a < x < b$ e que $o(f, A \cap]a,b[) < \varepsilon$. Se se tomar $\delta > 0$ tal que $]x - \delta, x + \delta[\subset]a,b[$, então, por maioria de razão, $o(f, A \cap]x - \delta, x + \delta[) < \varepsilon$ e, em particular, se $y \in A \cap]x - \delta, x + \delta[$, tem-se $|f(y) - f(x)| < \varepsilon$. ∎

Há uma relação importante entre o conceito de oscilação e a definição do integral de Riemann, que reside no facto de, se $a, b \in \mathbb{R}$ com $a < b$, se $f \colon [a,b] \longrightarrow \mathbb{R}$ é uma função limitada e se P é uma partição de $[a,b]$, então, se os pontos de P forem $a = a_0 < a_1 < \ldots < a_n = b$, tem-se

$$\overline{\Sigma}(f,P) - \underline{\Sigma}(f,P) = \sum_{k=1}^{n} o\big(f, [a_{k-1}, a_k]\big)(a_k - a_{k-1}). \tag{1.8}$$

TEOREMA 1.3 *Sejam $a, b \in \mathbb{R}$ tais que $a < b$ e seja f uma função limitada de $[a, b]$ em $\mathbb{R}$. Então f é integrável segundo Riemann se e só se o conjunto dos pontos de descontinuidade de f tiver medida nula.*

DEMONSTRAÇÃO: Seja D o conjunto dos pontos de descontinuidade de f e, para cada $\alpha > 0$, seja $D_\alpha = \{\, x \in [a, b] \mid o_f(x) \geqslant \alpha \,\}$. Resulta da proposição 1.18 que $(\forall \alpha > 0) : D_\alpha \subset D$ e que, por outro lado,

$$D = \{\, x \in [a, b] \mid o_f(x) > 0 \,\} = \bigcup_{n \in \mathbb{N}} D_{1/n};$$

consequentemente, afirmar que D tem medida nula equivale a afirmar que cada conjunto D_α tem medida nula. Vai-se então provar que f é integrável segundo Riemann se e só se cada D_α tem medida nula.

Comece-se por supor que f é integrável segundo Riemann e seja $\alpha > 0$; quer-se provar que D_α tem medida nula. Seja $\varepsilon > 0$; vai-se mostrar que existe alguma família finita de intervalos abertos cuja reunião contém D_α e tal que a soma dos seus comprimentos é inferior a ε. Sabe-se, pelo critério de integrabilidade de Riemann, que existe alguma partição P de $[a, b]$ tal que

$$\overline{\Sigma}(f, P) - \underline{\Sigma}(f, P) < \frac{\alpha\varepsilon}{2}.$$

Então $P = \{a_0, a_1, \ldots, a_n\}$ com $a = a_0 < a_1 < \cdots < a_n = b$. Tem-se $D_\alpha = (D_\alpha \cap P) \cup (D_\alpha \setminus P)$. O primeiro destes conjuntos é finito e, consequentemente, existe alguma família finita de intervalos abertos cuja reunião o contém e cuja soma dos comprimentos é menor do que $\varepsilon/2$; basta tomar, por exemplo, o conjunto de intervalos

$$\left\{ \left] a_n - \frac{\varepsilon}{4(n+1)}, a_n + \frac{\varepsilon}{4(n+1)} \right[\;\middle|\; k \in \{0, 1, \ldots, n\} \right\}.$$

Por outro lado, tem-se

$$\begin{aligned}
\frac{\alpha\varepsilon}{2} &> \overline{\Sigma}(f, P) - \underline{\Sigma}(f, P) \\
&= \sum_{k \in \{1, \ldots, n\}} o(f, [a_{k-1}, a_k])(a_k - a_{k-1}) \text{ (por (1.8))} \\
&\geqslant \sum_{\substack{k \in \{1, \ldots, n\} \\ ]a_{k-1}, a_k[\cap D_\alpha \neq \emptyset}} o(f, [a_{k-1}, a_k])(a_k - a_{k-1}) \\
&\geqslant \sum_{\substack{k \in \{1, \ldots, n\} \\ ]a_{k-1}, a_k[\cap D_\alpha \neq \emptyset}} \alpha(a_k - a_{k-1}),
\end{aligned}$$

pois cada um dos intervalos $]a_{k-1}, a_k[$ envolvidos contém algum ponto tal que a oscilação de f nesse ponto é maior ou igual a α. Mas resulta destas desigualdades que

$$\sum_{\substack{k\in\{1,\ldots,n\}\\ ]a_{k-1},a_k[\cap D_\alpha\neq\emptyset}}(a_k - a_{k-1}) < \frac{\varepsilon}{2},$$

ou seja, que a soma dos comprimentos dos intervalos $]a_{k-1}, a_k[$ que intersectam $D_\alpha \setminus P$ é menor do que $\varepsilon/2$.

Suponha-se agora que cada D_α tem medida nula; quer-se provar que f é integrável segundo Riemann. Antes de se prosseguir, é conveniente que se prove que D_α é compacto. De facto, trata-se de um conjunto limitado e por outro lado, $\mathbb{R} \setminus D_\alpha$ é um aberto, pois se $x \in \mathbb{R} \setminus D_\alpha$, há duas possibilidades:

$x \notin [a, b]$: então $x < a$ ou $x > b$; no primeiro caso, $x \in]-\infty, a[\subset \mathbb{R} \setminus D_\alpha$ e, no segundo, $x \in]b, +\infty[\subset \mathbb{R} \setminus D_\alpha$;

$x \in [a, b]$: então, como $o_f(x) < \alpha$, existe algum intervalo aberto I que contém x e tal que $o(f, [a, b] \cap I) < \alpha$ (pela definição de $o_f(x)$), o que implica que $I \subset \mathbb{R} \setminus D_\alpha$.

Seja então $\varepsilon > 0$; vai-se provar que existe alguma partição P de $[a, b]$ tal que $\overline{\Sigma}(f, P) - \underline{\Sigma}(f, P) < \varepsilon$. Sejam α e β números reais maiores do que 0 e tais que $(b - a)\alpha + o(f, [a, b])\beta < \varepsilon$. Por hipótese, $l(D_\alpha) = 0$ e, portanto, existe alguma sucessão $(I_n)_{n\in\mathbb{N}}$ de intervalos abertos de $\mathbb{R}$ tal que $D_\alpha \subset \bigcup_{n\in\mathbb{N}} I_n$ e que $\sum_{n=1}^{+\infty} \operatorname{comp}(I_n) < \beta$. Por outro lado, como D_α é compacto, existe algum $N \in \mathbb{N}$ tal que $D_\alpha \subset \bigcup_{k=1}^{N} I_k$. Para cada $k \in \{1, 2 \ldots, n\}$, seja $J_k = \overline{I_k} \cap [a, b]$. Então cada J_k é um intervalo fechado de $\mathbb{R}$ contido em $[a, b]$ e, portanto, $\bigcup_{k=1}^{n} J_k$ pode ser escrito sob a forma

$$[a_1, a_2] \cup [a_3, a_4] \cup \cdots \cup [a_{2n-1}, a_{2n}],$$

com $a \leqslant a_1 < a_2 < \cdots < a_n \leqslant b$ e

$$\begin{aligned}
\sum_{j=1}^{2n}(a_{2j} - a_{2j-1}) &= \sum_{j=1}^{n} \operatorname{comp}([a_{2j-1}, a_{2j}]) \\
&\leqslant \sum_{k=1}^{N} \operatorname{comp}(J_k) \\
&\leqslant \sum_{n=1}^{N} \operatorname{comp}(I_k) \\
&< \beta
\end{aligned}$$

Seja $P' = \{a,b\} \cup \{\, a_k \mid 1 \leqslant k \leqslant 2n \,\}$. Então P' é uma partição de $[a,b]$ e, pela maneira como foi definida, é possível dividir os seus intervalos em dois conjuntos disjuntos, de modo que a soma dos comprimentos dos intervalos do primeiro conjunto (os intervalos da forma $[a_{2j-1}, a_{2j}]$) é menor do que β e cada um dos restantes intervalos é formado unicamente por pontos x de $[a,b]$ tais que $o_f(x) < \alpha$. Observe-se que se $[a', b']$ é um tal intervalo, então existe alguma partição Q de $[a', b']$ tal que a oscilação de f em cada intervalo da partição Q é menor do que α. Isto pode ser demonstrado do seguinte modo: para cada $x \in [a', b']$, seja I_x algum intervalo aberto de $\mathbb{R}$ que contenha x e tal que $o(f, [a', b'] \cap I) < \alpha$; um tal intervalo existe necessariamente, pela definição de $o_f(x)$. Então $[a', b'] \subset \bigcup_{x \in [a', b']} I_x$ e, portanto, pelo lema 1.2, existe alguma partição Q de $[a', b']$ tal que cada intervalo da partição Q está contido em algum conjunto I_x ($x \in [a', b']$), pelo que a oscilação da função f em cada intervalo é menor do que α. Recorrendo a uma tal partição para cada um dos intervalos da partição P' que não sejam da forma $[a_{2j-1}, a_{2j}]$ ($j \in \{1, \ldots, n\}$), obtém-se uma partição P de $[a,b]$ formada por pontos $a = b_0 < b_1 < \cdots < b_m = b$ de modo a que o conjunto $\{1, 2, \ldots, m\}$ pode ser escrito como a reunião de dois conjuntos disjuntos I e I' tais que:

- a soma dos comprimentos dos intervalos $[b_{k-1}, b_k]$ ($k \in I$) é menor do que β;
- a oscilação de f em cada intervalo $[b_{k-1}, b_k]$ ($k \in I'$) é menor do que α.

Logo, tem-se

$$\begin{aligned}
&\overline{\Sigma}(f,P) - \underline{\Sigma}(f,P) = \\
&= \sum_{k=1}^{m} o\big(f, [b_{k-1}, b_k]\big)(b_k - b_{k-1}) \text{ (por (1.8))} \\
&= \sum_{k \in I} o\big(f, [b_{k-1}, b_k]\big)(b_k - b_{k-1}) + \sum_{k \in I'} o\big(f, [b_{k-1}, b_k]\big)(b_k - b_{k-1}) \\
&\leqslant \sum_{k \in I} o\big(f, [a,b]\big)(b_k - b_{k-1}) + \sum_{k \in I'} \alpha(b_k - b_{k-1}) \\
&< o(f, [a,b])\beta + (b-a)\alpha \\
&< \varepsilon,
\end{aligned}$$

pela escolha de α e de β. ∎

Recorrendo a este teorema, é bastante fácil demonstrar toda uma sére de resultados do tipo «se $a, b \in \mathbb{R}$ com $a < b$ e se f é uma função de $[a,b]$

em $\mathbb{R}$ que satisfaz a condição P, então f é integrável segundo Riemann». Vamos ver alguns. Em todos eles, a e b são números reais com $a < b$ e f é uma função de $[a,b]$ em $\mathbb{R}$.

COROLÁRIO 1.2 *Se f é contínua, então f é integrável segundo Riemann.*

DEMONSTRAÇÃO: Pelo teorema de Weirstrass (veja-se [13, cap. 4] ou [16, cap. 7]), f é limitada. Como o conjunto dos pontos de descontinuidade de f é o conjunto vazio, que tem medida nula, resulta do teorema 1.3 que f é integrável segundo Riemann. ■

COROLÁRIO 1.3 *Se f é monótona, então f é integrável segundo Riemann.*

DEMONSTRAÇÃO: Vai-se fazer a demonstração no caso em que f é crescente; o caso em que é decrescente é análogo.

É claro que f é limitada, pois $(\forall x \in [a,b]) : f(a) \leqslant f(x) \leqslant f(b)$. Basta então provar que o conjunto dos pontos de descontinuidade de f tem medida nula. De facto, vai-se mesmo provar que aquele conjunto é finito ou numerável. Se f for descontínua num ponto $c \in]a,b[$, então, como os limites $\lim_{x\to c^-} f(x)$ e $\lim_{x\to c^+} f(x)$ existem (são $\sup f([a,c[)$ e $\inf f(]c,b])$ respectivamente) e como $\lim_{x\to c^-} f(x) \leqslant f(c) \leqslant \lim_{x\to c^+} f(x)$, a descontinuidade de f em c só pode resultar de se ter

$$\lim_{x\to c^-} f(x) < f(c) \text{ ou } \lim_{x\to c^+} f(x) > f(c),$$

o que equivale a afirmar que $\lim_{x\to c^+} f(x) - \lim_{x\to c^-} f(x) > 0$. Seja D o conjunto dos pontos de descontinuidade de f e, para cada $c \in D$, sejam $a_c = \lim_{x\to c^-} f(x)$ e $b_c = \lim_{x\to c^+} f(x)$. Então os intervalos $]a_c, b_c[$ $(c \in D)$ são dois a dois disjuntos (pois se $c, d \in D$ e $c < d$, então cada elemento de $]a_c, b_c[$ é menor do que cada elemento de $]a_d, b_d[$) e cada um deles contém algum número racional. Visto que $\mathbb{Q}$ é numerável, resulta desta observação que D é numerável. ■

COROLÁRIO 1.4 *Sejam $c, d \in \mathbb{R}$ tais que $c < d$ e que $f([a,b]) \subset [c,d]$. Se f for integrável segundo Riemann e se g for uma função contínua de $[c,d]$ em $\mathbb{R}$, então $g \circ f$ é integrável segundo Riemann.*

DEMONSTRAÇÃO: Pelo teorema de Weierstrass, g é limitada e, consequentemente, $g \circ f$ é limitada. Por outro lado, em todos os pontos onde f é contínua, $g \circ f$ também o é; posto de outro modo, o conjunto dos pontos de descontinuidade de $g \circ f$ é um subconjunto do conjunto dos pontos de descontinuidade de f. Como este tem medida nula, aquele também tem.■

COROLÁRIO 1.5 *Se existir alguma sucessão* $(f_n)_{n \in \mathbb{N}}$ *de funções integráveis segundo Riemann de* $[a, b]$ *em* $\mathbb{R}$ *que convirja uniformemente para* f, *então* f *é integrável segundo Riemann.*

DEMONSTRAÇÃO: A função f é limitada, por ser o limite uniforme de uma sucessão de funções limitadas. Por outro lado, seja D_n o conjunto dos pontos de descontinuidade da função f_n ($n \in \mathbb{N}$). Se $x \in [a, b] \setminus \bigcup_{n \in \mathbb{N}} D_n$, então cada f_n é contínua no ponto x e, portanto, f é contínua no ponto x. Logo, o conjunto dos pontos de descontinuidade de f está contido em $\bigcup_{n \in \mathbb{N}} D_n$, o qual tem medida nula, visto que cada D_n tem medida nula.■

Integração

2.1 Funções mensuráveis

DEFINIÇÃO 2.1 Um *espaço de medida* é um tripleto $(X, \mathcal{A}, m)$, sendo X um conjunto, $\mathcal{A}$ uma σ-álgebra de partes de X e m uma medida de $\mathcal{A}$ em $\overline{\mathbb{R}_+}$. Dado um tal espaço de medida, diz-se que uma função $f\colon X \longrightarrow \overline{\mathbb{R}}$ é *mensurável* se $\left(\forall B \in \mathcal{B}\left(\overline{\mathbb{R}}\right)\right) : f^{-1}(B) \in \mathcal{A}$.

EXEMPLO 2.1 Qualquer função constante f é mensurável, pois se f toma sempre o valor $\lambda \in \mathbb{R}$, então, para cada $B \in \mathcal{B}\left(\overline{\mathbb{R}}\right)$,

$$f^{-1}(B) = \begin{cases} X & \text{caso } \lambda \in B \\ \emptyset & \text{caso contrário.} \end{cases}$$

Visto que $\emptyset, X \in \mathcal{A}$, está provado que f é mensurável.

Nos enunciados que se seguem, estará pressuposto que se estará a trabalhar num espaço de medida $(X, \mathcal{A}, m)$.

PROPOSIÇÃO 2.1 *Se f é uma função de X em $\overline{\mathbb{R}}$, são condições equivalentes:*

1. *f é mensurável;*
2. $(\forall a \in \mathbb{R}) : f^{-1}([-\infty, a]) \in \mathcal{A}$*;*
3. $(\forall a \in \mathbb{R}) : f^{-1}([-\infty, a[) \in \mathcal{A}$*;*
4. $(\forall a \in \mathbb{R}) : f^{-1}([a, +\infty]) \in \mathcal{A}$*;*
5. $(\forall a \in \mathbb{R}) : f^{-1}(]a, +\infty]) \in \mathcal{A}$*.*

DEMONSTRAÇÃO: Se a primeira condição se verifica e se $a \in \mathbb{R}$, então, como $[-\infty, a] \in \mathcal{B}(\overline{\mathbb{R}})$, $f^{-1}([-\infty, a]) \in \mathcal{A}$.

Se a segunda condição se verifica e se $a \in \mathbb{R}$, então

$$\begin{aligned} f^{-1}([-\infty, a[) &= f^{-1}\left(\bigcap_{n\in\mathbb{N}} \left[-\infty, a - \frac{1}{n}\right]\right) \\ &= \bigcap_{n\in\mathbb{N}} f^{-1}\left(\left[-\infty, a - \frac{1}{n}\right]\right), \end{aligned}$$

que pertence a $\mathcal{A}$, pois é a intersecção de uma sucessão de elementos de $\mathcal{A}$. Mostra-se de maneira análoga que a quarta condição implica a quinta.

Se a terceira condição se verifica e se $a \in \mathbb{R}$, então

$$f^{-1}([a, +\infty]) = X \setminus f^{-1}([-\infty, a[) \in \mathcal{A}.$$

Finalmente, suponha-se que a quinta condição se verifica. Vai-se provar que se A é um aberto de $\overline{\mathbb{R}}$, então $f^{-1}(A) \in \mathcal{A}$; isto basta para provar que f é mensurável, pois $\{\, P \subset \overline{\mathbb{R}} \mid f^{-1}(P) \in \mathcal{A} \,\}$ é obviamente uma σ-álgebra e, portanto, se se provar que contém os abertos de $\overline{\mathbb{R}}$, ficará provado que contém $\mathcal{B}(\overline{\mathbb{R}})$, o que é o mesmo que dizer que f é mensurável. Por outro lado, qualquer aberto de $\overline{\mathbb{R}}$ é uma reunião de um conjunto numerável de intervalos da forma $]a, b[$ ($a, b \in \mathbb{R}$ e $a < b$), $]a, +\infty]$ ($a \in \mathbb{R}$) ou $[-\infty, a[$ ($a \in \mathbb{R}$), pelo que basta provar que que a imagem recíproca por f de cada um dos intervalos deste tipo pertence a $\mathcal{A}$. No caso dos intervalos da forma $]a, +\infty[$ ($a \in \mathbb{R}$), está-se precisamente a supor isso. No caso dos intervalos da forma $[-\infty, a[$ ($a \in \mathbb{R}$), basta ver que

$$[-\infty, a[= \overline{\mathbb{R}} \setminus [a, +\infty[= \overline{\mathbb{R}} \setminus \bigcap_{n\in\mathbb{N}} \left]a - \frac{1}{n}, +\infty\right],$$

pelo que

$$f^{-1}([-\infty, a[) = X \setminus \bigcap_{n\in\mathbb{N}} f^{-1}\left(\left]a - \frac{1}{n}, +\infty\right]\right) \in \mathcal{A}.$$

Finalmente, se $a, b \in \mathbb{R}$ e $a < b$, então

$$]a, b[= [-\infty, b[\setminus[-\infty, a] = [-\infty, b[\setminus \bigcap_{n\in\mathbb{N}} \left[-\infty, a + \frac{1}{n}\right[$$

e então

$$f^{-1}(]a, b[) = f^{-1}([-\infty, b[) \setminus \bigcap_{n\in\mathbb{N}} f^{-1}\left(\left[-\infty, a + \frac{1}{n}\right[\right).$$

∎

EXEMPLO 2.2 Se $(X, \mathcal{A}, m) = (\mathbb{R}, \mathcal{M}(\mathbb{R}), l)$ e se f é a função definida por $f(x) = x^2$, então f é mensurável, pois, para cada $a \in \mathbb{R}$

$$f^{-1}([-\infty, a[) = \begin{cases} \emptyset & \text{caso } a \leqslant 0 \\ ]-\sqrt{a}, \sqrt{a}[& \text{caso contrário.} \end{cases}$$

Há um conceito que é conveniente introduzir agora, pois ajuda a dar um exemplo de uma função não mensurável e será útil posteriormente em outros contextos.

DEFINIÇÃO 2.2 Sejam X um conjunto e A uma parte de X. Designa-se por *função característica* do conjunto A a função

$$\begin{aligned} \chi_A \colon \ X & \longrightarrow \{0,1\} \\ x & \mapsto \begin{cases} 1 & \text{se } x \in A \\ 0 & \text{caso contrário.} \end{cases} \end{aligned}$$

EXEMPLO 2.3 Se $\mathcal{A} \neq \mathcal{P}(X)$, então há funções não mensuráveis de X em $\overline{\mathbb{R}}$, pois se $A \in \mathcal{P}(X) \setminus \mathcal{A}$, a função χ_A não é mensurável, uma vez que $\chi_A^{-1}([-\infty, 1[) = A^\complement \notin \mathcal{A}$.

PROPOSIÇÃO 2.2 *Seja N um conjunto numerável. Se $(f_n)_{n \in N}$ for uma família de funções mensuráveis, então as funções $\sup_{n \in N} f_n$ e $\inf_{n \in N} f_n$ são mensuráveis.*

DEMONSTRAÇÃO: Seja $a \in \mathbb{R}$; quer-se provar que

$$\left(\sup_{n \in N} f_n\right)^{-1}(]a, +\infty]) \in \mathcal{M}(\mathbb{R}).$$

Mas, para cada $x \in X$,

$$\begin{aligned} x \in \left(\sup_{n \in N} f_n\right)^{-1}(]a, +\infty]) & \iff \sup_{n \in N} f_n(x) > a \\ & \iff (\exists n \in N) : f_n(x) > a \\ & \iff x \in \bigcup_{n \in N} f_n^{-1}(]a, +\infty]). \end{aligned}$$

Logo,

$$\left(\sup_{n \in N} f_n\right)^{-1}(]a, +\infty]) = \bigcup_{n \in N} f_n^{-1}(]a, +\infty])$$

e este conjunto pertence a $\mathcal{A}$, pois N é finito ou numerável. Mostra-se de maneira análoga que $\inf_{n \in N} f_n$ é mensurável. ∎

Se o conjunto N do enunciado anterior for finito, então $\sup_{n\in N} f_n = \max_{n\in N} f_n$ e $\inf_{n\in N} f_n = \min_{n\in N} f_n$. Em particular, isto mostra que se f e g são mensuráveis, então $\max\{f, g\}$ e $\min\{f, g\}$ também o são.

Convém relembrar o que são o limite superior e o limite inferior de uma sucessão $(a_n)_{n\in\mathbb{N}}$ de elementos de $\overline{\mathbb{R}}$. Definem-se por

$$\limsup_{n\in\mathbb{N}} = \inf_{n\in\mathbb{N}}\left(\sup_{p\geqslant n} a_p\right) \quad \text{e por} \quad \liminf_{n\in\mathbb{N}} = \sup_{n\in\mathbb{N}}\left(\inf_{p\geqslant n} a_p\right)$$

respectivamente. A sucessão $(a_n)_{n\in\mathbb{N}}$ tem limite quando e só quando $\limsup_{n\in\mathbb{N}} a_n = \liminf_{n\in\mathbb{N}} a_n$ e, caso esta condição se verifique,

$$\limsup_{n\in\mathbb{N}} a_n = \lim_{n\in\mathbb{N}} a_n = \liminf_{n\in\mathbb{N}} a_n \tag{2.1}$$

(veja-se [12, ap. A]).

COROLÁRIO 2.1 *Seja $(f_n)_{n\in\mathbb{N}}$ uma sucessão de funções mensuráveis de X em $\overline{\mathbb{R}}$. Então:*

1. *as funções $\limsup_{n\in\mathbb{N}} f_n$ e $\liminf_{n\in\mathbb{N}} f_n$ são mensuráveis.*
2. *se, para cada $x \in X$, existir o limite $\lim_{n\in\mathbb{N}} f_n(x)$, a função $\lim_{n\in\mathbb{N}} f_n$ é mensurável.*

DEMONSTRAÇÃO: A primeira alínea resulta imediatamente da proposição 2.2 e das definições de limite superior e de limite inferior. Quanto à segunda, é consequência da primeira e de (2.1). ∎

TEOREMA 2.1 *Seja $n \in \mathbb{N}$, sejam $f_1, f_2, \ldots, f_n$ funções mensuráveis de X em $\mathbb{R}$ e seja F uma função contínua de $\mathbb{R}^n$ em $\mathbb{R}$. Então a função h de X em $\mathbb{R}$ definida por $h(x) = F(f_1(x), f_2(x), \ldots, f_n(x))$ é mensurável.*

DEMONSTRAÇÃO: Seja $a \in \mathbb{R}$; quer-se provar que $h^{-1}(]-\infty, a[) \in \mathcal{A}$. Como $]-\infty, a[$ é um aberto de $\mathbb{R}$ e F é contínua, $F^{-1}(]-\infty, a[)$ é um aberto de $\mathbb{R}^n$ e, portanto, pode ser escrito como reunião de produtos de intervalos abertos de $\mathbb{R}$. Recorrendo ao mesmo método que foi empregue na demonstração da proposição 1.9 (ou seja, considerando apenas intervalos com extremidades racionais), vê-se que $F^{-1}(]-\infty, a[)$ pode ser escrito sob a forma $\bigcup_{N\in\mathbb{N}}]a_{1,N}, b_{1,N}[\times]a_{2,N}, b_{2,N}[\times\cdots\times]a_{n,N}, b_{n,N}[$. Mas então

$$\begin{aligned}
&h^{-1}(]-\infty, a[) = \\
&= \left\{ x \in X \;\middle|\; (f_1(x), \ldots, f_n(x)) \in \bigcup_{N\in\mathbb{N}}]a_{1,N}, b_{1,N}[\times\cdots\times]a_{n,N}, b_{n,N}[\right\} \\
&= \bigcup_{N\in\mathbb{N}} \left(f_1^{-1}(]a_{1,N}, b_{1,N}[) \cap \cdots \cap f_n^{-1}(]a_{n,N}, b_{n,N}[) \right) \qquad (2.2)
\end{aligned}$$

Mas, para cada $n \in \mathbb{N}$ e cada $k \in \{1, 2, \ldots, n\}$,

$$f_k^{-1}(]a_{k,N}, b_{k,N}[) = f_k^{-1}(]-\infty, b_{k,N}[) \setminus f_k^{-1}(]-\infty, a_{k,N}]) \in \mathcal{A}.$$

Logo, (2.2) pertence a $\mathcal{A}$. ■

COROLÁRIO 2.2 *Sejam f e g funções mensuráveis de X em $\mathbb{R}$. Então as funções $f \pm g$, fg e $|f|$ são mensuráveis.*

DEMONSTRAÇÃO: Basta aplicar o teorema anterior à adição, à subtracção e à multiplicação, que são funções contínuas de $\mathbb{R}^2$ em $\mathbb{R}$, bem como à função valor absoluto, que é uma função contínua de $\mathbb{R}$ em $\mathbb{R}$. ■

Em particular, se f é mensurável e $\lambda \in \mathbb{R}$, então λf é mensurável, visto que as funções constantes são mensuráveis. Obviamente, é mais natural demonstrar este resultado directamente a partir da definição de função mensurável.

2.2 Integral: definição e propriedades elementares

DEFINIÇÃO 2.3 Diz-se que uma função real f é uma *função simples* se a sua imagem for um conjunto finito.

É imediato que as funções simples de um conjunto X em $\mathbb{R}$ formam um espaço vectorial, pois a soma de duas funções simples é novamente uma função simples e o produto de uma função simples por um escalar ainda é uma função simples. De facto, trata-se do espaço vectorial gerado pelas funções características (i. e. pelas funções do conjunto $\{ \chi_A \mid A \subset X \}$), pois se s for uma função simples e se a sua imagem for o conjunto $\{x_1, \ldots, x_n\} \subset \mathbb{R}$, então

$$s = \sum_{k=1}^{n} x_k \chi_{s^{-1}(\{x_k\})}.$$

Em particular, se $(X, \mathcal{A}, m)$ for um espaço de medida e se s for uma função mensurável simples de X em $\mathbb{R}$, s é combinação linear de funções características de elementos de $\mathcal{A}$, pois, para cada x pertencente à imagem de s, $s^{-1}(\{x\}) \in \mathcal{A}$.

É conveniente introduzir a seguinte notação: se f é uma função de um conjunto X em $\overline{\mathbb{R}}$, então definem-se as funções f^+ e f^- de X em $\overline{\mathbb{R}_+}$ por $f^+ = \max\{f, 0\}$ e $f^- = -\min\{f, 0\}$. É imediato que $f = f^+ - f^-$ e que $|f| = f^+ + f^-$. Além disso, resulta da observação feita logo após a demonstração da proposição 2.2 que, se f for mensurável, então f^+ e f^- também o são.

PROPOSIÇÃO 2.3 *Se f é uma função de X em $\overline{\mathbb{R}}$, então existe alguma sucessão $(s_n)_{n\in\mathbb{N}}$ de funções simples tal que $f = \lim_{n\in\mathbb{N}} s_n$. Além disso:*

1. *se $f \geqslant 0$, pode-se tomar $(s_n)_{n\in\mathbb{N}}$ monótona crescente e tal que, para cada $n \in \mathbb{N}$, $s_n \geqslant 0$;*
2. *se f for mensurável, pode-se tomar $(s_n)_{n\in\mathbb{N}}$ tal que cada s_n seja mensurável.*

DEMONSTRAÇÃO: Basta demonstrar o teorema no caso em que $f \geqslant 0$, pois, uma vez demonstrado isto, a validade do teorema no caso geral resultará de se ter $f = f^+ - f^-$ e de $f^+, f^- \geqslant 0$, pois se $(s_n)_{n\in\mathbb{N}}$ e $(s'_n)_{n\in\mathbb{N}}$ forem sucessões de funções simples tais que $f^+ = \lim_{n\in\mathbb{N}} s_n$ e $f^- = \lim_{n\in\mathbb{N}} s'_n$, então $f = \lim_{n\in\mathbb{N}}(s_n - s'_n)$.

Seja então $f\colon X \longrightarrow \overline{\mathbb{R}_+}$. Para cada $n \in \mathbb{N}$, seja

$$F_n = \{\, x \in X \mid f(x) \geqslant n \,\}$$

e, para cada $k \in \{1, 2, \ldots, n2^n\}$, seja

$$E_{n,k} = \left\{ x \in X \;\middle|\; \frac{k-1}{2^n} \leqslant f(x) < \frac{k}{2^n} \right\}.$$

Define-se então $s_n\colon X \longrightarrow \mathbb{R}_+$ por

$$s_n = \left(\sum_{k=1}^{n2^n} \frac{k-1}{2^n} \chi_{E_{n,k}} \right) + n\chi_{F_n}$$

e resulta de definição de $(s_n)_{n\in\mathbb{N}}$ que se f for mensurável, então cada função s_n ($n \in \mathbb{N}$) também o é e que a sucessão $(s_n)_{n\in\mathbb{N}}$ é monótona crescente. Veja-se na figura 2.1 o gráfico da função $f\colon [-2, 2] \longrightarrow \mathbb{R}$

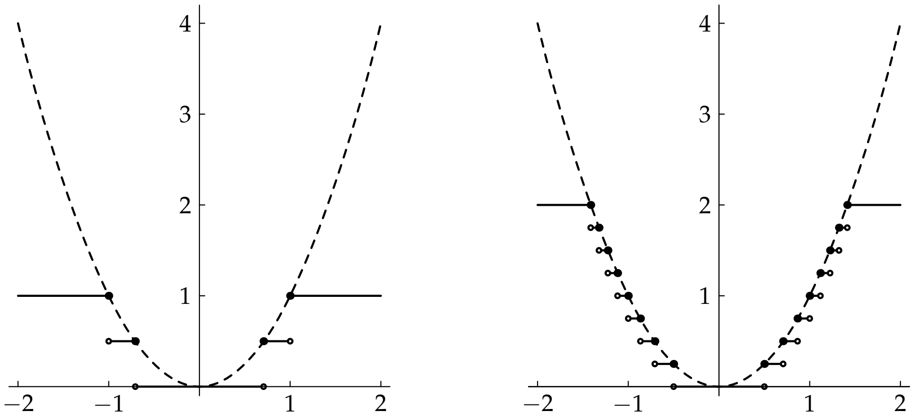

Figura 2.1: Aproximação de f por funções simples

definida por $f(x) = x^2$ (representado a tracejado) juntamente com os gráficos de s_1 e de s_2.

Para se provar que $f = \lim_{n\in\mathbb{N}} s_n$, tome-se $x \in X$. Se $f(x) = +\infty$, então $(\forall n \in \mathbb{N}) : s_n(x) = n$ e, portanto, $\lim_{n\in\mathbb{N}} s_n(x) = f(x)$. Caso $f(x) \in \mathbb{R}_+$, então, para cada $n \in \mathbb{N}$ tal que $f(x) \leqslant n$, tem-se que $f(x)$ pertence a um e um só intervalo da forma $\left[{}^{(k-1)}/_{2^n}, {}^{k}/_{2^n}\right[$ e que $s_n(x) = {}^{(k-1)}/_{2^n}$, pelo que

$$0 \leqslant f(x) - s_n(x) < \frac{1}{2^n}.$$

Logo, $\lim_{n\in\mathbb{N}} s_n(x) = f(x)$. ∎

Vai-se começar por definir o conceito de integral no caso das funções mensuráveis simples e usá-lo para definir o integral de uma função mensurável arbitrária. Se $(X, \mathcal{A}, m)$ for um espaço de medida e se s for uma função mensurável simples de X em $\mathbb{R}_+$, define-se

$$I_X(s) = \sum_{x\in s(X)} x.m\big(s^{-1}(\{x\})\big),$$

convencionando-se que $0.(+\infty) = (+\infty).0 = 0$. Convém observar que se $s = \sum_{k=1}^{n} x_k \chi_{A_k}$, com $A_1, \ldots, A_n$ elementos dois a dois disjuntos de $\mathcal{A}$, então

$$I_X(s) = \sum_{k=1}^{n} x_k m(A_k). \tag{2.3}$$

De facto, supondo que cada A_k é não vazio e que cada x_k é diferente de 0 (o que não altera nem a função s nem o valor da soma (2.3)), tem-se que cada x_k pertence à imagem de s e que $A_k \subset s^{-1}(\{x_k\})$. De facto, se

$x \in s(X)$, então $s^{-1}(\{x\})$ é a reunião dos A_k contidos em $s^{-1}(\{x\})$. Logo, se se representar por I_x o conjunto $\{\, k \in \{1, 2, \ldots, n\} \mid A_k \subset s^{-1}(\{x\}) \,\}$, tem-se

$$\begin{aligned}\sum_{k=1}^{n} x_k m(A_k) &= \sum_{x\in s(X)} \sum_{k\in I_x} x_k m(A_k) \\ &= \sum_{x\in s(X)} \sum_{k\in I_x} x.m(A_k) \\ &= \sum_{x\in s(X)} x \sum_{k\in I_x} m(A_k) \\ &= \sum_{x\in s(X)} x.m(s^{-1}(\{x\})).\end{aligned}$$

Será provado posteriormente que se continua a ter (2.3) mesmo quando não se está a supor que os conjuntos A_k são dois a dois disjuntos.

Observe-se que se $Y \in \mathcal{A}$, então $\mathcal{A}_Y = \{\, M \cap Y \mid M \in \mathcal{A} \,\}$ é uma σ-álgebra e que a restrição m_Y de m a $\mathcal{A}_Y$ ainda é uma medida, pelo que $(Y, \mathcal{A}_Y, m_Y)$ é outro espaço de medida. Se s for uma função mensurável simples de X em $\mathbb{R}$, então a $s|_Y$ é uma função mensurável simples de Y em $\mathbb{R}$, pelo que faz sentido falar de $I_Y(s|_Y)$. Por abuso de notação, este número será representado por $I_Y(s)$.

DEFINIÇÃO 2.4 Se $(X, \mathcal{A}, m)$ é um espaço de medida e f é uma função mensurável de X em $\overline{\mathbb{R}_+}$, então representa-se por $\int_X f\, dm$ o supremo do conjunto

$$\{\, I_X(s) \mid 0 \leqslant s \leqslant f \text{ e } s \text{ é mensurável e simples} \,\}. \tag{2.4}$$

Caso $\int_X f\, dm < +\infty$, o número $\int_X f\, dm$ designa-se por *integral de Lebesgue* de f e diz-se então que a função f é *integrável*.

Se f for uma função mensurável de X em $\overline{\mathbb{R}}$, então diz-se que f é integrável se as funções f^+ e f^- o forem. Nesse caso designa-se por integral de Lebesgue de f o número

$$\int_X f\, dm = \int_X f^+\, dm - \int_X f^-\, dm.$$

Quando $f\colon X \longrightarrow \overline{\mathbb{R}}$ é mensurável e, das funções f^+ e f^-, uma é integrável e a outra não, também se emprega a notação $\int_X f\, dm$ para representar $\int_X f^+\, dm - \int_X f^-\, dm$, embora neste caso não se diga que f é integrável.

Observe-se que, nas condições da definição anterior, se f somente tomar valores não negativos, então o integral de Lebesgue de f (que se designará simplesmente por integral de f, a menos que haja ambiguidade) foi definido de duas maneiras diferentes: como o supremo do conjunto (2.4) e como $\int_X f^+ \, dm - \int_X f^- \, dm$. Naturalmente, isto não leva a qualquer ambiguidade, pois neste caso $f = f^+$ e $f^- \equiv 0$, pelo que

$$\int_X f^+ \, dm - \int_X f^- \, dm = \int_X f \, dm - 0 = \int_X f \, dm.$$

Analogamente ao que foi feito após se introduzir a notação $I_X(s)$, se $(X, \mathcal{A}, m)$ for um espaço de medida, se $Y \in \mathcal{A}$ e se f for uma função de X em $\overline{\mathbb{R}}$ cuja restrição a Y seja integrável, então o integral desta restrição será representado por $\int_Y f \, dm$ e não por $\int_Y f|_Y \, dm_Y$.

Antes de se verem exemplos de funções integráveis e de funções não integráveis, será demonstrado que o integral de uma função mensurável simples é aquilo que se esperaria que fosse.

PROPOSIÇÃO 2.4 *Se s é uma função mensurável simples de X em $\mathbb{R}_+$, então*

$$\int_X s \, dm = I_X(s).$$

DEMONSTRAÇÃO: Comecemos por demonstrar o teorema no caso em que $s \geqslant 0$. Resulta imediatamente da definição de $\int_X s \, dm$ que este número é maior ou igual a $I_X(s)$, uma vez que $\int_X s \, dm$ é o supremo de um conjunto que contém $I_X(s)$. Seja agora s' uma função mensurável simples de X em $\mathbb{R}_+$ tal que $s' \leqslant s$; quer-se provar que $I_X(s') \leqslant I_X(s)$. Se se provar isto, então resultará de definição de $\int_X s \, dm$ que $\int_X s \, dm \leqslant I_X(s)$.

A função s pode ser escrita sob a forma

$$s = \sum_{k=1}^{n} x_k \chi_{A_k},$$

onde os conjuntos $A_1, A_2, \ldots, A_n$ são elementos dois a dois disjuntos de $\mathcal{A}$. Para cada $k \in \{1, 2, \ldots, n\}$, a restrição de s' a A_k pode ser escrita sob a forma

$$s'|_{A_k} = \sum_{j=1}^{n_k} y_{j,k} B_{j,k}$$

onde os conjuntos $B_1, B_2, \ldots, B_n$ são elementos dois a dois disjuntos de $\mathcal{A}$ cuja reunião é igual a A_k. Como $s' \leqslant s$, cada $y_{j,k}$ é menor ou igual a x_j.

Então

$$\begin{aligned}
I_X(s') &= \sum_{k=1}^{n} \sum_{j=1}^{n_k} y_{j,k}\chi_{B_{j,k}} \\
&\leqslant \sum_{k=1}^{n} \sum_{j=1}^{n_k} x_k\chi_{B_{j,k}} \\
&= \sum_{k=1}^{n} x_k \sum_{j=1}^{n_k} \chi_{B_{j,k}} \\
&= \sum_{k=1}^{n} x_k\chi_{A_k} \\
&= I_X(s).
\end{aligned}$$

■

Vão ser vistos exemplos de funções integráveis e de funções não integráveis. Em todos os casos, as funções terão por domínio sub-conjuntos mensuráveis A de $\mathbb{R}$ e, a menos que se diga explicitamente o contrário, a σ-álgebra em questão será $\mathcal{M}(\mathbb{R})_A$ e a medida será l_A.

EXEMPLO 2.4 Se $f \equiv 1$, então f não é integrável, pois $f = \chi_{\mathbb{R}}$ e, pela proposição 2.4, $\int_{\mathbb{R}} \chi_{\mathbb{R}}\, dl = I_{\mathbb{R}}(\chi_{\mathbb{R}}) = +\infty$.

EXEMPLO 2.5 A restrição a $[-1, 1]$ da função f do exemplo anterior é integrável, pois $\int_{[-1,1]} f\, dl = 2$, novamente pela proposição 2.4.

EXEMPLO 2.6 Se $f\colon \mathbb{R} \longrightarrow \mathbb{R}$ é a função definida por $f(x) = x$, então f não é integrável, pois

$$(\forall x \in \mathbb{R}) : f^+(x) = \begin{cases} x & \text{se } x \geqslant 0 \\ 0 & \text{caso contrário} \end{cases}$$

e então $f^+ \geqslant \chi_{[1,+\infty[}$, pelo que

$$\int_{\mathbb{R}} f^+\, dl \geqslant \int_{\mathbb{R}} \chi_{[1,+\infty[}\, dl = +\infty.$$

EXEMPLO 2.7 Sejam $a, b \in \mathbb{R}$ tais que $a < b$ e seja $f\colon [a, b] \longrightarrow \mathbb{R}$ a função definida por

$$f(x) = \begin{cases} 1 & \text{se } x \in \mathbb{Q} \\ 0 & \text{caso contrário.} \end{cases}$$

Então f é integrável, pois $f = \chi_{[a,b]\cap\mathbb{Q}}$, e $\int_{[a,b]} f\,dl = l([a,b]\cap\mathbb{Q}) = 0$. Este exemplo mostra que há funções integráveis segundo Lebesgue que não são integráveis segundo Riemann, apesar de serem limitadas.[1] Será visto mais à frente que qualquer função de $[a,b]$ em $\mathbb{R}$ integrável segundo Riemann também é integrável segundo Lebesgue e que $\int_a^b f(x)\,dx = \int_{[a,b]} f\,dl$.

EXEMPLO 2.8 Considere-se em $\mathbb{N}$ a medida de contagem m, definida no exemplo 1.3. Então a função $f\colon \mathbb{N} \longrightarrow \mathbb{R}$ definida por $f(n) = 1/n$ não é integrável. De facto, se se definir, para cada $n \in \mathbb{N}$,

$$\begin{array}{rccc} s_n\colon & \mathbb{N} & \longrightarrow & \mathbb{R} \\ & k & \mapsto & \begin{cases} 1/k & \text{se } k \leqslant n \\ 0 & \text{caso contrário,} \end{cases} \end{array}$$

então $(s_n)_{n\in\mathbb{N}}$ é uma sucessão de funções simples, mensuráveis e não negativas e, além diso, $(\forall n \in \mathbb{N}) : s_n \leqslant f$. Logo

$$(\forall n \in \mathbb{N}) : \int_{\mathbb{N}} f\,dm \geqslant \int_{\mathbb{N}} s_n\,dm = \sum_{k=1}^{n} \frac{1}{k},$$

pelo que $\int_{\mathbb{N}} f\,dl \geqslant \lim_{n\in\mathbb{N}} \sum_{k=1}^{n} \frac{1}{k} = +\infty$.

Se $(X, \mathcal{A}, m)$ for um espaço de medida, vai-se representar por $\mathcal{L}(X)$ o espaço das funções integráveis de X em $\overline{\mathbb{R}}$. Se $\lambda \in \mathbb{R}$ é claro o que se entende por λf, excepto se $\lambda = 0$ e se $\pm\infty$ estiver na imagem de f. Vai-se convencionar que, mesmo nesse caso, $0.f = 0$.

Vai-se agora introduzir uma expressão que é empregue frequentemente em teoria da medida, nomeadamente *quase sempre* e que significa «sempre, excepto num conjunto de medida nula». Assim por exemplo, no enunciado do teorema 1.3 a condição «se o conjunto dos pontos de descontinuidade de f tiver medida nula» poderia ter sido escrita sob a forma «se f for contínua quase sempre». É mesmo frequente empregar-se q. s. no lugar de «quase sempre».

PROPOSIÇÃO 2.5 *Seja $(X, \mathcal{A}, m)$ um espaço de medida.*

1. *Se f for uma função mensurável de X num intervalo $[a,b]$ de $\mathbb{R}$ e se $m(X) < +\infty$, então $f \in \mathcal{L}(X)$ e*

$$a.m(X) \leqslant \int_X f\,dm \leqslant b.m(X).$$

[1] Que a função f não é integrável segundo Riemann resulta do teorema 1.3 (a função f é descontínua em todos os pontos do seu domínio), mas é mais natural observar que, para cada partição P de $[a,b]$, $\underline{\Sigma}(f,P) = 0$ e $\overline{\Sigma}(f,P) = b - a$.

2. *Se* $f, g \in \mathcal{L}(X)$ *e se* $f \leqslant g$*, então*
$$\int_X f\,dm \leqslant \int_X g\,dm.$$
3. *Se* $f \in \mathcal{L}(X)$ *e se* $\lambda \in \mathbb{R}$*, então* λf *é integrável e*
$$\int_X \lambda f\,dm = \lambda \int_X f\,dm.$$
4. *Se* $m(X) = 0$ *e* $f\colon X \longrightarrow \overline{\mathbb{R}}$ *for mensurável, então* $\int_X f\,dm = 0$.
5. *Se* $f \in \mathcal{L}(X)$ *e se* $Y \in \mathcal{A}$*, então* $f|_Y \in \mathcal{L}(Y)$.
6. *Se* f *é uma função mensurável de* X *em* $\overline{\mathbb{R}_+}$ *e se* $\int_X f\,dm = 0$*, então* $f(x) = 0$ *q. s.*

DEMONSTRAÇÃO: Para demonstrar a primeira alínea, comece-se por supor que $a \geqslant 0$. Então $f \geqslant 0$, pelo que
$$\int_X f\,dm \geqslant I_X(a\chi_X) = a.m(X)$$
e por outro lado, se s for uma função simples e mensurável tal que $0 \leqslant s \leqslant f$, então $s \leqslant b$, pelo que $I_X(s) \leqslant b.m(X)$. Como isto tem lugar para cada s naquelas condições, tem-se que $\int_X f\,dm \leqslant b.m(X)$.

Se se tiver $b \leqslant 0$, então $f^+ \equiv 0$ e o argumento anterior mostra que
$$-b.m(X) \leqslant \int_X f^-\,dm \leqslant -a.m(X),$$
pelo que
$$\int_X f\,dm = \int_X f^+\,dm - \int_X f^-\,dm = -\int_X f^-\,dm \in [a.m(X), b.m(X)].$$

Finalmente, se $a < 0 < b$, então $0 \leqslant \int_X f^+\,dm \leqslant b.m(X)$ e $0 \leqslant \int_X f^-\,dm \leqslant -a.m(X)$, pelo que
$$\int_X f\,dm = \int_X f^+\,dm - \int_X f^-\,dm \in [a.m(x), b.m(X)].$$

A segunda alínea resulta imediatamente da definição do integral de Lebesgue quando $0 \leqslant f \leqslant g$. No caso geral, tem-se $f^+ \leqslant g^+$ e $f^- \geqslant g^-$, pelo que
$$\int_X f\,dm = \int_X f^+\,dm - \int_X f^-\,dm \geqslant \int_X g^+\,dm - \int_X g^-\,dm = \int_X g\,dm.$$

Tal como no caso da segunda alínea, a terceira resulta imediatamente da definição do integral de Lebesgue caso $0 \leqslant f$ e $0 \leqslant \lambda$. Caso $0 \leqslant f$ e $\lambda < 0$, então

$$\begin{aligned}\int_X \lambda f\, dm &= \int_X \overbrace{(\lambda f)^+}^{\equiv 0}\, dm - \int_X (\lambda f)^-\, dm\\ &= -\int_X -\lambda f\, dm\\ &= -(-\lambda)\int_X f\, dm\\ &= \lambda \int_X f\, dm.\end{aligned}$$

Caso $f \in \mathcal{L}(X)$ e $\lambda \geqslant 0$, então

$$\begin{aligned}\int_X \lambda f\, dm &= \int_X (\lambda f)^+\, dm - \int_X (\lambda f)^-\, dm\\ &= \int_X \lambda f^+\, dm - \int_X \lambda f^-\, dm\\ &= \lambda\int_X f^+\, dm - \lambda\int_X f^-\, dm\\ &= \lambda \int_X f\, dm\end{aligned}$$

Finalmente, caso $f \in \mathcal{L}(X)$ e $\lambda < 0$, tem-se

$$\begin{aligned}\int_X \lambda f\, dm &= \int_X (\lambda f)^+\, dm - \int_X (\lambda f)^-\, dm\\ &= \int_X -\lambda f^-\, dm - \int_X -\lambda f^+\, dm\\ &= (-\lambda)\int_X f^-\, dm - (-\lambda)\int_X f^+\, dm\\ &= (-\lambda).\left(-\int_X f\, dm\right)\\ &= \lambda \int_X f\, dm.\end{aligned}$$

Para demonstrar a quarta alínea, comece-se por supor que $f \geqslant 0$. Se s for uma função mensurável simples tal que $0 \leqslant s \leqslant f$, então $I_X(s) = 0$, pois se $s = \sum_{k=1}^n x_k \chi_{A_k}$ tem-se

$$I_X(s) = \sum_{k=1}^n x_k m(A_K) = 0,$$

uma vez que, para cada $k \in \{1, 2, \ldots, n\}$, $m(A_k) \leqslant m(X) = 0$. No caso geral, tem-se então

$$\int_X f\,dm = \int_X f^+\,dm - \int_X f^-\,dm = 0.$$

Para demonstrar a quinta alínea, basta fazê-lo no caso em que $f \geqslant 0$, uma vez que $f^+|_Y = f|_Y^+$ e que $f^-|_Y = f|_Y^-$. Mas se $f \geqslant 0$ e se s é uma função mensurável simples de Y em $\mathbb{R}$ tal que $0 \leqslant s \leqslant f$, então se se definir $s^*\colon X \longrightarrow \mathbb{R}$ como sendo o prolongamento de s a X que se anula em $X \setminus Y$, então s^* é uma função mensurável simples tal que $s^* \leqslant f$, pelo que

$$I_Y(s) = I_X(s^*) \leqslant \int_X f\,dm,$$

de onde resulta que

$$\int_Y f\,dm \leqslant \int_X f\,dm < +\infty.$$

Finalmente, vai-se demonstrar a sexta alínea. Para cada $n \in \mathbb{N}$, seja

$$D_n = \left\{ x \in X \;\middle|\; f(x) \geqslant \frac{1}{n} \right\}.$$

Afirmar que $f(x) = 0$ q. s. equivale a afirmar que $\{\, x \in X \mid f(x) \neq 0 \,\}$ tem medida nula e, visto que

$$\{\, x \in X \mid f(x) \neq 0 \,\} = \bigcup_{n \in \mathbb{N}} D_n,$$

isto é o mesmo que afirmar que $(\forall n \in \mathbb{N}) : m(D_n) = 0$. Se se tivesse $m(D_n) > 0$ para algum $n \in \mathbb{N}$, então a função $n^{-1}\chi_{D_n}$ seria uma função mensurável simples e $0 \leqslant n^{-1}\chi_{D_n} \leqslant f$, pelo que

$$\int_X f\,dm \geqslant \int_X \frac{1}{n}\chi_{D_n}\,dm = \frac{1}{n}m(D_n) > 0,$$

o que é absurdo. ∎

PROPOSIÇÃO 2.6 *Seja f uma função mensurável de X em $\overline{\mathbb{R}}$. Então, caso $f \geqslant 0$ ou caso $f \in \mathcal{L}(X)$, a função*

$$\begin{aligned} \phi\colon\ \mathcal{A} &\longrightarrow \overline{\mathbb{R}_+} \\ A &\longmapsto \int_A f\,dm \end{aligned}$$

é σ-aditiva, i. e. se $(A_n)_{n\in\mathbb{N}}$ é uma sucessão de elementos de $\mathcal{A}$ disjuntos dois a dois, então

$$\phi\left(\bigcup_{n\in\mathbb{N}} A_n\right) = \sum_{n=1}^{+\infty} \phi(A_n).$$

DEMONSTRAÇÃO: Naturalmente, basta fazer a demonstração no caso em que $f \geqslant 0$. A validade da proposição no caso em que $f \in \mathcal{L}(X)$ resultará então imediatamente da definição de $\int_X f\,dm$.

Vai-se começar por provar que ϕ é crescente, i. e. que se $A, B \in \mathcal{A}$ e se $A \subset B$, então $\phi(A) \leqslant \phi(B)$. Resulta das definições de ϕ e de integral de Lebesgue de uma função que, para tal, basta provar que se s é uma função mensurável simples de A em $\mathbb{R}$ tal que $0 \leqslant s \leqslant f|_A$, então $I_A(s) \leqslant \phi(B)$. Seja s^* o prolongamento de s a B que se anula em $B \setminus A$. Então s^* é uma função simples mensurável de B em $\mathbb{R}$ tal que $0 \leqslant s^* \leqslant f|_B$, pelo que

$$I_A(s) = I_B(s^*) \leqslant \int_B f\,dm.$$

Seja então $(A_n)_{n\in\mathbb{N}}$ uma sucessão de elementos de $\mathcal{A}$ disjuntos dois a dois. Caso $f = \chi_E$ para algum $E \in \mathcal{A}$, então

$$\begin{aligned}
\phi\left(\bigcup_{n\in\mathbb{N}} A_n\right) &= \int_{\bigcup_{n\in\mathbb{N}} A_n} \chi_E\,dm \\
&= m\left(\left(\bigcup_{n\in\mathbb{N}} A_n\right) \cap E\right) \\
&= m\left(\bigcup_{n\in\mathbb{N}} (A_n \cap E)\right) \\
&= \sum_{n=1}^{+\infty} m(A_n \cap E) \\
&= \sum_{n=1}^{+\infty} \int_{A_n} \chi_E\,dm \\
&= \sum_{n=1}^{+\infty} \phi(A_n).
\end{aligned}$$

Caso f seja uma função mensurável simples, então $f = \sum_{k=1}^{n} x_k \chi_{E_k}$, com $E_1, E_2, \ldots, E_n$ elementos de $\mathcal{A}$ dois a dois disjuntos. Então

$$\begin{aligned}
\phi\left(\bigcup_{n\in\mathbb{N}} A_n\right) &= \int_{\bigcup_{n\in\mathbb{N}} A_n} \sum_{k=1}^{n} x_k \chi_{E_k}\, dm \\
&= \sum_{k=1}^{n} x_k \int_{\bigcup_{n\in\mathbb{N}} A_n} \chi_{E_k}\, dm \\
&= \sum_{k=1}^{n} x_k \sum_{n=1}^{+\infty} \int_{A_n} \chi_{E_k}\, dm \\
&= \sum_{n=1}^{+\infty} \int_{A_n} \sum_{k=1}^{n} x_k \chi_{E_k}\, dm \\
&= \sum_{n=1}^{+\infty} \phi(A_n).
\end{aligned}$$

Passemos agora ao caso geral. Se s for uma função simples mensurável tal que $0 \leqslant s \leqslant f$, então

$$\int_{\bigcup_{n\in\mathbb{N}} A_n} s\, dm = \sum_{n=1}^{+\infty} \int_{A_n} s\, dm \leqslant \sum_{n=1}^{+\infty} \int_{A_n} f\, dm = \sum_{n=1}^{+\infty} \phi(A_n),$$

pelo que

$$\phi\left(\bigcup_{n\in\mathbb{N}} A_n\right) = \int_{\bigcup_{n\in\mathbb{N}} A_n} f\, dm \leqslant \sum_{n=1}^{+\infty} \phi(A_n).$$

Para demonstrar a desigualdade oposta, basta considerar o caso em que $(\forall n \in \mathbb{N}) : \phi(A_n) < +\infty$, pois se se tivesse $\phi(A_n) = +\infty$ para algum $n \in \mathbb{N}$, então, como ϕ é crescente, ter-se-ia

$$\phi\left(\bigcup_{n\in\mathbb{N}} A_n\right) = \sum_{n=1}^{+\infty} \phi(A_n) = +\infty.$$

Seja $n \in \mathbb{N}$ e seja $\varepsilon > 0$. Para cada $j \in \{1, 2, \ldots, n\}$, existe alguma função mensurável simples s tal que $0 \leqslant s \leqslant f$ e que

$$(\forall j \in \{1, 2, \ldots, n\}) : \int_{A_j} s\, dm \geqslant \int_{A_j} f\, dm - \frac{\varepsilon}{n}.$$

Então

$$\begin{aligned}
\phi\left(\bigcup_{j=1}^{n} A_j\right) &= \int_{\bigcup_{j=1}^{n} A_j} f\,dm \\
&\geqslant \int_{\bigcup_{j=1}^{n} A_j} s\,dm \\
&= \sum_{j=1}^{n} \int_{A_j} s\,dm \\
&\geqslant \sum_{j=1}^{n} \left(\int_{A_j} f\,dm - \frac{\varepsilon}{n}\right) \\
&= \left(\sum_{j=1}^{n} \phi(A_j)\right) - \varepsilon
\end{aligned}$$

e, como esta desigualdade é valida para qualquer $\varepsilon > 0$, tem-se que $\phi\left(\bigcup_{j=1}^{n} A_j\right) \geqslant \sum_{j=1}^{n} \phi(A_j)$. Como ϕ é crescente, isto prova que

$$(\forall n \in \mathbb{N}) : \phi\left(\bigcup_{j\in\mathbb{N}} A_j\right) \geqslant \sum_{j=1}^{n} \phi(A_j)$$

e, portanto, que

$$\phi\left(\bigcup_{n\in\mathbb{N}} A_n\right) \geqslant \sum_{n=1}^{+\infty} \phi(A_n).$$

■

Resulta desta proposição que, dado um espaço de medida $(X, \mathcal{A}, m)$ e dada uma função mensurável f de X em $\overline{\mathbb{R}_+}$, a função

$$\begin{array}{ccc}
\mathcal{A} & \longrightarrow & \overline{\mathbb{R}_+} \\
A & \mapsto & \displaystyle\int_A f\,dm
\end{array}$$

é uma medida. Segundo a quarta alínea da proposição 2.5 esta medida anula-se em todos os $A \in \mathcal{A}$ tais que $m(A) = 0$. Pode-se provar que se m for tal que X é reunião de alguma família finita ou numerável de elementos de $\mathcal{A}$ de medida finita[2] então, para cada medida μ de $\mathcal{A}$ em $\overline{\mathbb{R}_+}$ tal que

$$(\forall A \in \mathcal{A}) : m(A) = 0 \Longrightarrow \mu(A) = 0,$$

[2] É o caso, por exemplo, de $(\mathbb{R}, \mathcal{M}(\mathbb{R}), l)$.

existe alguma função $f \in \mathcal{L}(X)$ tal que $f \geqslant 0$ e que

$$(\forall A \in \mathcal{A}) : \mu(X) = \int_A f \, dm.$$

Isto é o teorema de Lebesgue-Radon-Nikodym; veja-se [14, cap. 6]. Sem a hipótese de que X é reunião de alguma família finita ou numerável de elementos de $\mathcal{A}$ de medida finita, o enunciado é falso. Tome-se, por exemplo, $X = \mathbb{R}$, $\mathcal{A} = \mathcal{M}(\mathbb{R})$, m a medida de contagem[3] e l a medida de Lebesgue. Se $A \in \mathcal{M}(\mathbb{R})$ for tal que $m(A) = 0$, então $A = \emptyset$, pelo que $l(A) = 0$. Mas não existe nenhuma função mensurável $f \colon \mathbb{R} \longrightarrow \overline{\mathbb{R}_+}$ tal que

$$(\forall A \in \mathcal{M}(\mathbb{R})) : l(A) = \int_A f \, dm \tag{2.5}$$

pois, caso existisse, ter-se-ia, para cada $x \in \mathbb{R}$,

$$0 = l(\{x\}) = \int_{\{x\}} f \, dm = f(x)m(\{x\}) = f(x).$$

Logo, $f \equiv 0$, o que é incompatível com (2.5).

COROLÁRIO 2.3 *Seja $f \in \mathcal{L}(X)$. Se $Y \in \mathcal{A}$ for tal que $m(X \setminus Y) = 0$, então $\int_Y f \, dm = \int_X f \, dm$.*

DEMONSTRAÇÃO: Pela proposição anterior e porque $X = Y \dot{\cup} (X \setminus Y)$, tem-se

$$\int_X f \, dm = \int_Y f \, dm + \int_{Y \setminus X} f \, dm = \int_Y f \, dm,$$

pela quarta alínea da proposição 2.5. ■

Resulta imediatamente do corolário 2.3 que se $(X, \mathcal{A}, m)$ for um espaço de medida e se $f, g \in \mathcal{L}(X)$ forem tais que $f(x) = g(x)$ q. s., então $\int_X f \, dm = \int_X g \, dm$, pois se $Y = \{ x \in X \mid f(x) = g(x) \}$, então

$$\int_X f \, dm = \int_Y f \, dm = \int_Y g \, dm = \int_X g \, dm.$$

PROPOSIÇÃO 2.7 *Seja $f \in \mathcal{L}(X)$. Então $|f| \in \mathcal{L}(X)$ e*

$$\left| \int_X f \, dm \right| \leqslant \int_X |f| \, dm.$$

[3] Trata-se da medida definida no exemplo 1.3.

DEMONSTRAÇÃO: Sejam

$$A = \{ x \in X \mid f(x) \geqslant 0 \} \quad \text{e} \quad B = \{ x \in X \mid f(x) < 0 \}.$$

Então $A, B \in \mathcal{A}$ e $X = A \dot{\cup} B$, pelo que, pela proposição 2.6,

$$\begin{aligned}\int_X |f|\, dm &= \int_A |f|\, dm + \int_B |f|\, dm \\ &= \int_A f^+\, dm + \int_B f^-\, dm \\ &< +\infty,\end{aligned}$$

visto que, por hipótese, $\int_A f^+\, dm, \int_B f^-\, dm < +\infty$. Está então visto que $|f|$ é integrável. Para terminar a demonstração basta observar que

$$-|f| \leqslant f \leqslant |f|,$$

pelo que, pela segunda alínea da proposição 2.5,

$$\begin{aligned}&\int_X -|f|\, dm \leqslant \int_X f\, dm \leqslant \int_X |f|\, dm \iff \\ \iff &-\int_X |f|\, dm \leqslant \int_X f\, dm \leqslant \int_X |f|\, dm \\ \iff &\left| \int_X f\, dm \right| \leqslant \int_X |f|\, dm.\end{aligned}$$

■

PROPOSIÇÃO 2.8 *Sejam f e g funções mensuráveis de X em $\overline{\mathbb{R}}$. Se g for integrável e se $|f| \leqslant g$, então f é integrável.*

DEMONSTRAÇÃO: Basta observar que resulta de se ter $|f| \leqslant g$ que $f^+ \leqslant g$ e $f^- \leqslant g$, pelo que

$$\int_X f^+\, dm, \int_X f^-\, dm < +\infty.$$

■

Resulta das proposições 2.7 e 2.8 que, dados um espaço de medida $(X, \mathcal{A}, m)$ e uma função mensurável f de X em $\overline{\mathbb{R}}$, f é integrável se e só se $|f|$ for integrável.

EXEMPLO 2.9 Seja $f = \sum_{n=1}^{+\infty} \frac{(-1)^n}{n} \chi_{[n-1,n[}$. Esta função não é integrável segundo Lebesgue, pois se o fosse então $|f|$ também o seria. Ora $|f| = \sum_{n=1}^{+\infty} n^{-1} \chi_{[n-1,1n[}$, pelo que, para cada $N \in \mathbb{N}$, $|f| \geqslant \sum_{n=1}^{N} n^{-1} \chi_{[n-1,1n[}$. Então

$$\int_{\mathbb{R}} f\, dl \geqslant \int_{\mathbb{R}} \sum_{n=1}^{N} \frac{1}{n} \chi_{[n-1,1n[}\, dl = \sum_{n=1}^{N} \frac{1}{n}$$

e, portanto, $\int_{\mathbb{R}} f\, dl \geqslant \lim_{N \in \mathbb{N}} \sum_{n=1}^{N} n^{-1} = +\infty$.

Este exemplo mostra que a observação feita no exemplo 2.7, segundo a qual qualquer função integrável segundo Riemann é integrável segundo Lebesgue, não é válida no caso dos integrais impróprios,[4] pois a função f do exemplo que se acabou de ver é tal que, para cada $a, b \in \mathbb{R}$ com $a \leqslant b$, $f|_{[a,b]}$ é integrável segundo Riemann e, além disso, o integral impróprio $\int_{-\infty}^{+\infty} f(x)\,dx$ converge (é igual a $\sum_{n=1}^{+\infty} {}^{(-1)^n}/{}_n = -\log(2)$).

2.3 Integração de limites de sucessões

TEOREMA 2.2 (TEOREMA DA CONVERGÊNCIA MONÓTONA) *Se $(f_n)_{n\in\mathbb{N}}$ for uma sucessão monótona crescente de funções mensuráveis de X em $\overline{\mathbb{R}_+}$, então*

$$\lim_{n\in\mathbb{N}} \int_X f_n\,dm = \int_X \lim_{n\in\mathbb{N}} f_n\,dm.$$

DEMONSTRAÇÃO: Antes de se passar à demonstração propriamente dita, convém explicar porque é que existem ambos os limites mencionados no enunciado. De facto, o limite da sucessão $\left(\int_X f_n\,dm\right)_{n\in\mathbb{N}}$ existe necessariamente, pois trata-se de uma sucessão monótona, pela segunda alínea da proposição 2.5. Para cada $x \in X$, a sucessão $(f_n(x))_{n\in\mathbb{N}}$ também é monótona, pelo que, mais uma vez, o limite $\lim_{n\in\mathbb{N}} f_n(x)$ existe necessariamente.

Seja $f = \lim_{n\in\mathbb{N}} f_n$. Uma vez que $(\forall n \in \mathbb{N}) : f_n \leqslant f$, aplicando mais uma vez a segunda alínea da proposição 2.5 deduz-se que

$$(\forall n \in \mathbb{N}) : \int_X f_n\,dm \leqslant \int_X f\,dm,$$

pelo que

$$\lim_{n\in\mathbb{N}} \int_X f_n\,dm \leqslant \int_X f\,dm.$$

Vai-se agora demonstrar a desigualdade oposta. Seja s uma função mensurável simples de X em $\mathbb{R}$ tal que $0 \leqslant s \leqslant f$. Se se mostrar que $\int_X s\,dm \leqslant \lim_{n\in\mathbb{N}} \int_X f_n\,dm$, resulta da definição de $\int_X f\,dm$ que este número é menor ou igual ao limite $\lim_{n\in\mathbb{N}} \int_X f_n\,dm$. Seja $\alpha \in]0,1[$ e seja, para cada $n \in \mathbb{N}$

$$E_n = \{\, x \in X \mid f_n(x) \geqslant \alpha s(x) \,\}.$$

[4]No entanto, será visto mais à frente que continua válida no caso das funções não-negativas.

Se $x \in E_n$ para algum $n \in \mathbb{N}$, então $\alpha s(x) \leqslant f_n(x) \leqslant f_{n+1}(x)$, pelo que $x \in E_{n+1}$; logo, $E_n \subset E_{n+1}$. Por outro lado, se $x \in X$, então, uma vez que $\lim_{n\in\mathbb{N}} f_n(x) \geqslant s(x) > \alpha s(x)$, tem-se necessariamente que $f_n(x) \geqslant \alpha s(x)$ para algum $n \in \mathbb{N}$, pelo que $x \in E_n$. Está então provado que a sucessão $(E_n)_{n\in\mathbb{N}}$ é crescente e que a reunião dos seus elementos é X. Logo, resulta da proposição 2.6 e de se ter $s \geqslant 0$ que a função

$$\begin{array}{rcl} \mathcal{A} & \longrightarrow & \overline{\mathbb{R}_+} \\ A & \mapsto & \displaystyle\int_A s\,dm \end{array}$$

é uma medida e, portanto, a proposição 1.5 permite deduzir que

$$\int_X s\,dm = \int_{\bigcup_{n\in\mathbb{N}} E_n} s\,dm = \lim_{n\in\mathbb{N}} \int_{E_n} s\,dm.$$

Mas, por outro lado,

$$(\forall n \in \mathbb{N}) : \int_X f\,dm \geqslant \int_{E_n} f\,dm \geqslant \alpha \int_{E_n} s\,dm.$$

Resulta destas duas observações que $\int_X f\,dm \geqslant \alpha \int_X s\,dm$. Como esta desigualdade tem lugar para cada $\alpha \in]0,1[$, deduz-se que $\int_X f\,dm \geqslant \int_X s\,dm$, como se pretendia demonstrar. ■

Se se retirar do enunciado do teorema a hipótese de que a sucessão é monótona, então o enunciado é falso, mesmo que se possa garantir que os limites existem.

EXEMPLO 2.10 Considere-se em $\mathbb{N}$ a medida de contagem e, para cada $n \in \mathbb{N}$, seja f_n a função de $\mathbb{N}$ em $\mathbb{R}_+$ definida por

$$f_n(m) = \begin{cases} 1 & \text{se } m = n \\ 0 & \text{caso contrário.} \end{cases}$$

Então $\lim_{n\in\mathbb{N}} f_n = 0$ e então $\int_{\mathbb{N}} \lim_{n\in\mathbb{N}} f_n\,dm = 0$ mas, por outro lado, $(\forall n \in \mathbb{N}) : \int_{\mathbb{N}} f_n\,dm = \int_{\mathbb{N}} \chi_{\{n\}}\,dm = 1$, pelo que $\lim_{n\in\mathbb{N}} \int_{\mathbb{N}} f_n\,dm = 1$.

TEOREMA 2.3 *Se f e g forem funções integráveis de X em $\mathbb{R}$, então $f+g$ também é integrável e*

$$\int_X f+g\,dm = \int_X f\,dm + \int_X g\,dm.$$

DEMONSTRAÇÃO: Caso f e g sejam funções mensuráveis simples então, pela proposição 2.4, basta provar que $I_X(f+g) = I_X(f) + I_X(g)$. Suponha-se então que $f = \sum_{k=1}^{m} x_k \chi_{X_k}$ e que $g = \sum_{l=1}^{n} y_l \chi_{Y_l}$, sendo tanto os conjuntos $X_1, \ldots, X_m$ como os conjuntos $Y_1, \ldots, Y_n$ dois a dois disjuntos. Para cada $k \in \{1, \ldots, m\}$ e cada $l \in \{1, \ldots, n\}$, seja $Z_{kl} = X_k \cap Y_l$. Então os conjuntos Z_{kl} são dois a dois disjuntos e

$$f + g = \sum_{k=1}^{m} \sum_{l=1}^{n} (x_k + y_l) Z_{kl},$$

pelo que

$$\begin{aligned}
I_X(f+g) &= \sum_{k=1}^{m} \sum_{l=1}^{n} (x_k + y_l) m(Z_{kl}) \\
&= \sum_{k=1}^{m} \sum_{l=1}^{n} x_k m(Z_{kl}) + \sum_{l=1}^{n} \sum_{k=1}^{m} y_l m(Z_{kl}) \\
&= \sum_{k=1}^{m} x_k \sum_{l=1}^{n} m(Z_{kl}) + \sum_{l=1}^{n} y_l \sum_{k=1}^{m} m(Z_{kl}) \\
&= \sum_{k=1}^{m} x_k m(X_k) + \sum_{l=1}^{n} y_l m(Y_l) \\
&= I_X(f) + I_X(g).
\end{aligned}$$

Se $f, g \geqslant 0$, existem, pela proposição 2.3, sucessões monótonas crescentes $(s_n)_{n\in\mathbb{N}}$ e $(s'_n)_{n\in\mathbb{N}}$ de funções mensuráveis simples tais que se $n \in \mathbb{N}$, então $0 \leqslant s_n \leqslant f$ e $0 \leqslant s'_n \leqslant g$. Então $(s_n + s'_n)_{n\in\mathbb{N}}$ é uma sucessão monótona crescente de funções mensuráveis simples tal que $(\forall n \in \mathbb{N}) : 0 \leqslant s_n + s'_n \leqslant f + g$. Decorre então da proposição 2.4 e do teorema da convergência monótona que

$$\begin{aligned}
\int_X f + g \, dm &= \lim_{n\in\mathbb{N}} \int_X s_n + s'_n \, dm \\
&= \lim_{n\in\mathbb{N}} \int_X s_n \, dm + \lim_{n\in\mathbb{N}} \int_X s'_n \, dm \\
&= \int_X f \, dm + \int_X g \, dm;
\end{aligned}$$

em particular, $f + g$ é integrável.

Isto permite provar que sempre que f e g são funções integráveis de X em $\mathbb{R}$, $f + g$ também é integrável. De facto, basta provar que $|f + g|$ é integrável, mas isto resulta de se ter

$$\int_X |f+g| \, dm \leqslant \int_X |f| + |g| \, dm = \int_X |f| \, dm + \int_X |g| \, dm < +\infty.$$

Para terminar, se f e g forem funções integráveis de X em $\mathbb{R}$, quer-se provar que $\int_X f+g\,dm = \int_X f\,dm + \int_X g\,dm$. Seja $h = f+g$. Então

$$h^+ - h^- = h = f+g = f^+ - f^- + g^+ - g^-,$$

pelo que

$$h^+ + f^- + g^- = h^- + f^+ + g^+.$$

Logo, uma vez que todas as funções envolvidas são não-negativas,

$$\int_X h^+\,dm + \int_X f^-\,dm + \int_X g^-\,dm = \int_X h^-\,dm + \int_X f^+\,dm + \int_X g^+\,dm,$$

pelo que

$$\begin{aligned}\int_X h\,dm &= \int_X h^+\,dm - \int_X h^-\,dm\\ &= \int_X f^+\,dm - \int_X f^-\,dm + \int_X g^+\,dm - \int_X g^-\,dm\\ &= \int_X f\,dm + \int_X g\,dm.\end{aligned}$$ ■

Resulta deste teorema que, como foi mencionado na página 40, se $s = \sum_{k=1}^n x_k\chi_{A_k}$, como $A_1,\ldots,A_n \in \mathcal{A}$, então $I_X(s) = \sum_{k=1}^n x_k m(A_k)$, mesmo não se estando a supor que os conjuntos $A_1,\ldots,A_n$ são disjuntos dois a dois. De facto,

$$\begin{aligned}I_X(s) &= \int_X s\,dm \text{ (pela proposição 2.4)}\\ &= \int_X \sum_{k=1}^n x_k\chi_{A_k}\\ &= \sum_{k=1}^n \int_X x_k\chi_{A_k}\\ &= \sum_{k=1}^n I_X(x_k\chi_{A_k})\\ &= \sum_{k=1}^n x_k m(A_k).\end{aligned}$$

COROLÁRIO 2.4 *Se $(f_n)_{n\in\mathbb{N}}$ é uma sucessão de funções mensuráveis de X em $\overline{\mathbb{R}_+}$, então*

$$\int_X \sum_{n=1}^{+\infty} f_n\,dm = \sum_{n=1}^{+\infty} \int_X f_n\,dm.$$

DEMONSTRAÇÃO: Seja $f = \sum_{n=1}^{+\infty} f_n$. Se alguma função f_n não for integrável, então $\int_X f_n \, dm = +\infty$ e, como $f \geqslant f_n$, $\int_X f \, dm = +\infty$. Logo, a igualdade do enunciado reduz-se a $+\infty = +\infty$.

Caso $(f_n)_{n\in\mathbb{N}}$ seja uma sucessão de funções integráveis, então, pelo teorema anterior, cada função $\sum_{n=1}^{N} f_n$ $(N \in \mathbb{N})$ é integrável e

$$\int_X \sum_{n=1}^{N} f_n \, dm = \sum_{n=1}^{N} \int_X f_n \, dm,$$

pelo que, pelo teorema da convergência monótona

$$\begin{aligned}\int_X \sum_{n=1}^{+\infty} f_N \, dm &= \int_X \lim_{N\in\mathbb{N}} \sum_{n=1}^{N} f_n \, dm \\ &= \lim_{N\in\mathbb{N}} \int_X \sum_{n=1}^{N} f_n \, dm \\ &= \lim_{N\in\mathbb{N}} \sum_{n=1}^{N} \int_X f_n \, dm \\ &= \sum_{n=1}^{+\infty} \int_X f_n \, dm. \qquad \blacksquare\end{aligned}$$

Sabe-se, pelo exemplo 2.10, que o integral do limite de uma sucessão $(f_n)_{n\in\mathbb{N}}$ de funções integráveis não é necessariamente igual ao limite dos integrais, mesmo que todos os limites envolvidos existam. No entanto, como se vai ver, a desigualdade do exemplo 2.10 é a única que pode ter lugar quando as funções envolvidas são não-negativas.

TEOREMA 2.4 (TEOREMA DE FATOU) *Seja $(f_n)_{n\in\mathbb{N}}$ uma sucessão de funções mensuráveis de X em $\overline{\mathbb{R}_+}$. Então*

$$\int_X \liminf_{n\in\mathbb{N}} f_n \, dm \leqslant \liminf_{n\in\mathbb{N}} \int_X f_n \, dm.$$

DEMONSTRAÇÃO: Para cada $n \in \mathbb{N}$, seja $\varphi_n = \inf_{p\geqslant n} f_p$. Então $(\varphi_n)_{n\in\mathbb{N}}$ é uma sucessão monótona crescente de funções mensuráveis não-negativas. Além disso, $\lim_{n\in\mathbb{N}} \varphi_n = \liminf_{n\in\mathbb{N}} f_n$ (por definição de limite inferior) e $(\forall n \in \mathbb{N}) : \varphi_n \leqslant f_n$. Decorre então do teorema da convergência monótona que

$$\begin{aligned}\int_X \liminf_{n\in\mathbb{N}} f_n\,dm &= \int_X \lim_{n\in\mathbb{N}} \varphi_n\,dm\\ &= \lim_{n\in\mathbb{N}} \int_X \varphi_n\,dm\\ &= \liminf_{n\in\mathbb{N}} \int_X \varphi_n\,dm\\ &\leqslant \liminf_{n\in\mathbb{N}} \int_X f_n\,dm. \qquad \blacksquare\end{aligned}$$

Se as hipóteses do teorema se verificam então, em particular, tem-se

$$\int_X \lim_{n\in\mathbb{N}} f_n\,dm \leqslant \lim_{n\in\mathbb{N}} \int_X f_n\,dm$$

sempre que todos os limites envolvidos existam.

TEOREMA 2.5 (TEOREMA DA CONVERGÊNCIA DOMINADA) *Seja φ uma função integrável de X em $\overline{\mathbb{R}_+}$. Se $(f_n)_{n\in\mathbb{N}}$ é uma sucessão de funções mensuráveis de X em $\overline{\mathbb{R}}$, se o limite $\lim_{n\in\mathbb{N}} f_n(x)$ existe para cada $x \in X$ e se $(\forall n \in \mathbb{N}) : |f_n| \leqslant \varphi$, então cada função f_n é integrável, a função $\lim_{n\in\mathbb{N}} f_n$ é integrável e*

$$\int_X \lim_{n\in\mathbb{N}} f_n\,dm = \lim_{n\in\mathbb{N}} \int_X f_n\,dm.$$

DEMONSTRAÇÃO: Seja $f = \lim_{n\in\mathbb{N}} f_n$. Que a função f bem como cada função f_n são integráveis é uma consequência da proposição 2.8.

Uma vez que $(\forall n \in \mathbb{N}) : f_n + \varphi \geqslant 0$, resulta do teorema de Fatou que

$$\int_X f + \varphi\,dm = \int_X \lim_{n\in\mathbb{N}} f_n + \varphi\,dm \leqslant \liminf_{n\in\mathbb{N}} \int_X f_n + \varphi\,dm,$$

pelo que

$$\int_X f\,dm \leqslant \liminf_{n\in\mathbb{N}} \int_X f_n\,dm. \tag{2.6}$$

Por outro lado, como $(\forall n \in \mathbb{N}) : -f_n + \varphi \geqslant 0$, tem-se, pelo mesmo motivo, que

$$\int_X -f + \varphi\,dm = \int_X \lim_{n\in\mathbb{N}} -f_n + \varphi\,dm \leqslant \liminf_{n\in\mathbb{N}} \int_X -f_n + \varphi\,dm,$$

pelo que

$$-\int_X f\,dm = \int_X -f\,dm \leqslant \liminf_{n\in\mathbb{N}} \int_X -f_n\,dm = -\limsup_{n\in\mathbb{N}} \int_X f_n\,dm,$$

ou seja,

$$\int_X f\,dm \geqslant \limsup_{n\in\mathbb{N}} \int_X f_n\,dm. \tag{2.7}$$

Mas o limite superior de uma sucessão real é sempre maior ou igual ao limite inferior da mesma sucessão. Logo, decorre de (2.6) e de (2.7) que

$$\int_X f\,dm = \limsup_{n\in\mathbb{N}} \int_X f_n\,dm = \liminf_{n\in\mathbb{N}} \int_X f_n\,dm,$$

ou seja, que

$$\int_X f\,dm = \lim_{n\in\mathbb{N}} \int_X f_n\,dm.$$

■

COROLÁRIO 2.5 *Se $m(X) < +\infty$ e se $(f_n)_{n\in\mathbb{N}}$ é uma sucessão uniformemente limitada de funções integráveis de X em $\mathbb{R}$ que converge pontualmente para alguma função de X em $\mathbb{R}$, então $\lim_{n\in\mathbb{N}} f_n$ é integrável e*

$$\int_X \lim_{n\in\mathbb{N}} f_n\,dm = \lim_{n\in\mathbb{N}} \int_X f_n\,dm.$$

DEMONSTRAÇÃO: Seja $M \in \mathbb{R}_+$ um majorante de todas as funções f_n ($n \in \mathbb{N}$). Então pode-se aplicar o teorema da convergência dominada com $\varphi \equiv M$. ■

2.4 Integral de Riemann e integral de Lebesgue

Sejam $a, b \in \mathbb{R}$ tais que $a < b$ e seja f uma função integrável segundo Riemann de $[a, b]$ em $\mathbb{R}$. Quer-se provar que f é integrável segundo Lebesgue e, para o fazer, vai-se aproximar f por funções que são integráveis segundo Lebesgue. A maneira de definir tais funções é sugerida pela definição do integral de Riemann: a cada partição P de $[a, b]$, associam-se duas funções $s, S\colon [a, b] \longrightarrow P$ tais que $s \leqslant f \leqslant S$ do modo a que, em cada intervalo $[a', b']$ da partição, a restrição de s a $[a', b'[$ seja o ínfimo da restrição de f a $[a', b']$ e, analogamente, a restrição de S a $[a', b'[$ seja o supremo da restrição de f a $[a', b'[$. Veja-se a figura 2.2.

TEOREMA 2.6 *Sejam $a, b \in \mathbb{R}$ com $a < b$ e seja f uma função integrável segundo Riemann de $[a, b]$ em $\mathbb{R}$. Então f é integrável segundo Lebesgue e*

$$\int_a^b f(x)\,dx = \int_{[a,b]} f\,dl. \tag{2.8}$$

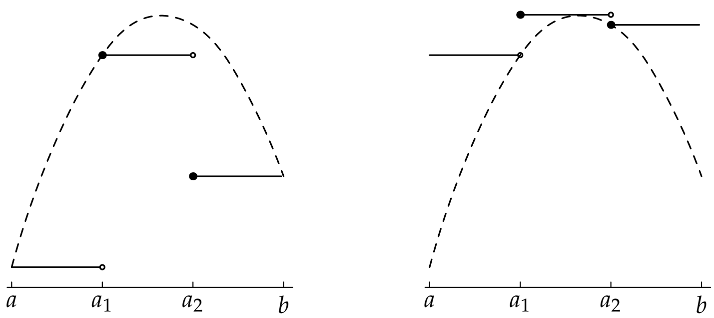

Figura 2.2: Aproximação de f por funções constantes nos intervalos de uma partição

DEMONSTRAÇÃO: Para cada $n \in \mathbb{N}$, seja P_n' uma partição de $[a,b]$ tal que

$$\overline{\Sigma}(f,P_n') - \underline{\Sigma}(f,P_n') < \frac{1}{n}. \tag{2.9}$$

Define-se, para cada $n \in \mathbb{N}$, $P_n = \bigcup_{k=1}^n P_n'$. Então cada P_n é uma partição de $[a,b]$, a sucessão $(P_n)_{n\in\mathbb{N}}$ é crescente e resulta da primeira alínea da proposição 1.17, de (2.9) e de se ter $P_n \supset P_n'$ que

$$\overline{\Sigma}(f,P_n) - \underline{\Sigma}(f,P_n) < \frac{1}{n}.$$

Definem-se as funções $s_n, S_n\colon [a,b] \longrightarrow \mathbb{R}$ do seguinte modo: se os pontos de P_n forem $a = a_{0,n} < a_{1,n} < \cdots < a_{m_n,n} = b$, então

$$s_n = \left(\sum_{k=1}^{m_n-1} \inf f([a_{k-1},a_k])\chi_{[a_{k-1},a_k[}\right) + \inf f([a_{m_n-1},a_{m_n}])\chi_{[a_{m_n-1},a_{m_n}]}$$

e

$$S_n = \left(\sum_{k=1}^{m_n-1} \sup f([a_{k-1},a_k])\chi_{[a_{k-1},a_k[}\right) + \sup f([a_{m_n-1},a_{m_n}])\chi_{[a_{m_n-1},a_{m_n}]}.$$

Então a sucessão $(s_n)_{n\in\mathbb{N}}$ é crescente, a sucessão $(S_n)_{n\in\mathbb{N}}$ é decrescente e

$$(\forall n \in \mathbb{N}) : s_n \leqslant f \leqslant S_n. \tag{2.10}$$

Sejam $s = \sup_{n\in\mathbb{N}} s_n$ e $S = \inf_{n\in\mathbb{N}} S_n$. Então, por (2.10), $s \leqslant f \leqslant S$. Por outro lado, as funções s e S são mensuráveis, pelo corolário 2.1. Se $n \in \mathbb{N}$, resulta das definições de s_n e de S_n que

$$\int_X s_n\,dl = \underline{\Sigma}(f,P_n) \quad \text{e que} \quad \int_X S_n\,dl = \overline{\Sigma}(f,P_n),$$

pelo que
$$\int_X S_n - s_n \, dl = \overline{\Sigma}(f, P_n) - \underline{\Sigma}(f, P_n) < \frac{1}{n}.$$
Logo, pelo corolário 2.5,
$$0 \leqslant \int_X S - s \, dl = \int_X \lim_{n \in \mathbb{N}} S_n - s_n \, dl = \lim_{n \in \mathbb{N}} \int_X S_n - s_n \, dl \leqslant 0,$$
pelo que $\int_X S - s \, dl = 0$. Resulta então da sexta alínea da proposição 2.5 que $s(x) = S(x)$ q. s. e, como $s \leqslant f \leqslant S$, $s(x) = f(x) = S(x)$ q. s., pelo que f é mensurável. De facto, como, para cada $t \in \mathbb{R}$
$$\{ x \in X \mid s(x) > t \} \subset \{ x \in X \mid f(x) > t \} \subset \{ x \in X \mid S(x) > t \}$$
e como o primeiro e o terceiro destes três conjuntos são mensuráveis e têm a mesma medida, o segundo também é mensurável.

Finalmente, aplicando novamente o corolário 2.5, vê-se que
$$\int_{[a,b]} f \, dl = \lim_{n \in \mathbb{N}} \int_{[a,b]} s_n \, dl = \lim_{n \in \mathbb{N}} \underline{\Sigma}(f, P_n) \leqslant \int_a^b f(x) \, dx$$
e que
$$\int_{[a,b]} f \, dl = \lim_{n \in \mathbb{N}} \int_{[a,b]} S_n \, dl = \lim_{n \in \mathbb{N}} \overline{\Sigma}(f, P_n) \geqslant \int_a^b f(x) \, dx,$$
pelo que se tem (2.8). ∎

Foi afirmado numa nota de rodapé na página 52 que o teorema anterior continua válido no contexto dos integrais impróprios desde que f seja não-negativa; que não é válido no caso geral, é algo que já fora visto no exemplo 2.9. Para simplificar a exposição, vai-se supor que se está a trabalhar com uma função real f definida num intervalo da forma $[a, +\infty[$ ($a \in \mathbb{R}$); os outros tipos de integrais impróprios são análogos. Afirmar que o integral impróprio $\int_a^{+\infty} f(x) \, dx$ converge é afirmar que a restrição de f a qualquer intervalo $[a, M]$, com $M > a$, é integrável segundo Riemann e que o limite $\lim_{M \to +\infty} \int_a^M f(x) \, dx$ existe (em $\mathbb{R}$); se for esse o caso, define--se
$$\int_a^{+\infty} f(x) \, dx = \lim_{M \to +\infty} \int_a^M f(x) \, dx.$$
Visto que se está a supor que f é não-negativa, a função
$$\begin{array}{ccc}]a, +\infty[& \longrightarrow & \mathbb{R} \\ M & \mapsto & \displaystyle\int_a^M f(x) \, dx \end{array}$$

é crescente e então o limite $\lim_{M\to+\infty}\int_a^M f(x)\,dx$ existe se e só se o limite $\lim_{n\in\mathbb{N}}\int_a^{a+n} f(x)\,dx$ existir e, caso ambos os limites existam, têm o mesmo valor. Seja $(f_n)_{n\in\mathbb{N}}$ a sucessão de funções de $[a,+\infty[$ em $\mathbb{R}$ tal que, para cada $n\in\mathbb{N}$ e cada $x\in[a,+\infty[$,

$$f_n(x)=\begin{cases} f(x) & \text{se } x\leqslant a+n\\ 0 & \text{caso contrário;}\end{cases}$$

posto de outro modo, $f_n=f.\chi_{[a,a+n]}$. Mas então, se $n\in\mathbb{N}$,

$$\int_{[a,+\infty[} f_n\,dl=\int_{[a,a+n]} f_n\,dl+\int_{]a+n,+\infty[} f_n\,dl=\int_a^{a+n} f_n(x)\,dx,$$

pelo teorema anterior e porque f_n se anula em $]a+n,+\infty[$. Por outro lado, a sucessão $(f_n)_{n\in\mathbb{N}}$ é crescente e $f=\lim_{n\in\mathbb{N}} f_n$, pelo que, pelo teorema da convergência monótona,

$$\int_{[a,+\infty[} f\,dl=\lim_{n\in\mathbb{N}}\int_{[a,+\infty[} f_n\,dl=\lim_{n\in\mathbb{N}}\int_a^{a+n} f_n(x)\,dx=\int_a^{+\infty} f\,dx.$$

A partir deste ponto, se $a,b\in\mathbb{R}$ com $a\leqslant b$ e se f for uma função integrável segundo Lebesgue de $[a,b]$ em $\overline{\mathbb{R}}$, o integral de Lebesgue de f vai ser representado pela notação tradicional

$$\int_a^b f(x)\,dx; \tag{2.11}$$

o teorema 2.6 garante que isto não leva a qualquer ambiguidade. Também se empregará a notação (2.11) se o domínio de f for um intervalo não compacto. Como foi visto, isto só poderá conduzir a ambiguidades caso o contradomínio de f contenha tanto valores maiores do que 0 como valores menores do que 0, mas nesses casos será declarado claramente qual é o integral com que se está a trabalhar.

A título de exemplo de como os resultados demonstrados neste capítulo podem ajudar ao cálculo de integrais de funções integráveis segundo Riemann, considerem-se as funções

$$\begin{array}{rccc}\Gamma\colon & \mathbb{R}_+^* & \longrightarrow & \mathbb{R}\\ & x & \mapsto & \displaystyle\int_0^{+\infty} e^{-t}t^{x-1}\,dt\end{array}\quad\text{e}\quad\begin{array}{rccc}\zeta\colon &]1,+\infty[& \longrightarrow & \mathbb{R}\\ & x & \mapsto & \displaystyle\sum_{n=1}^{+\infty}\frac{1}{n^x};\end{array}$$

vai-se provar que

$$(\forall x\in]1,+\infty[):\zeta(x)\Gamma(x)=\int_0^{+\infty}\frac{x^{t-1}}{e^t-1}\,dt.$$

Convém começar por observar que se $n \in \mathbb{N}$ então a substituição $t = nu$ permite concluir que

$$\Gamma(x) = n^x \int_0^{+\infty} e^{-nt} t^{x-1}\, dt, \tag{2.12}$$

pelo que, se $x > 1$,

$$\begin{aligned}
\zeta(x)\Gamma(x) &= \sum_{n=1}^{+\infty} \frac{\Gamma(x)}{n^x} \\
&= \sum_{n=1}^{+\infty} \int_0^{+\infty} e^{-nt} t^{x-1}\, dt \text{ (por (2.12))} \\
&= \int_0^{+\infty} \sum_{n=1}^{+\infty} e^{-nt} t^{x-1}\, dt \text{ (pelo corolário 2.4)} \\
&= \int_0^{+\infty} \frac{e^{-t}}{1-e^{-t}} t^{x-1}\, dt \\
&= \int_0^{+\infty} \frac{t^{x-1}}{e^t - 1}\, dt.
\end{aligned}$$

Derivação

Neste capítulo, vai-se estudar a derivabilidade de funções no contexto do integral de Lebesgue. Em particular, vai-se ver até que ponto é válido o teorema fundamental do Cálculo neste contexto. Adaptar as demontrações do contexto do integral de Riemann (veja-se [13, cap. 6] ou [16, cap. 14]) não apresenta qualquer dificuldade. De facto, demonstra-se facilmente, recorrendo ao teorema da convergência dominada, que é válido o seguinte

TEOREMA 3.1 *Se $a, b \in \mathbb{R}$ com $a < b$ e se f é uma função integrável de $[a, b]$ em $\mathbb{R}$, então a função*

$$\begin{array}{rccc} F\colon & [a,b] & \longrightarrow & \mathbb{R} \\ & t & \mapsto & \int_a^t f(x)\,dx \end{array}$$

é contínua e é mesmo derivável em cada ponto $x \in [a, b]$ onde f seja contínua, tendo-se então que $F'(x) = f(x)$.

Resulta deste facto, mesmo sem recorrer ao teorema 2.6 (i. e. ao facto de o integral de Lebesgue ser uma generalização do de Riemann) que se $f\colon [a, b] \longrightarrow \mathbb{R}$ for uma função contínua e se F for uma primitiva de f, então

$$\int_a^b f(x)\,dx = F(b) - F(a).$$

O que se vai estudar é o que acontece se se enfraquecerem as hipóteses referentes à função f.

3.1 O teorema da derivação de Lebesgue

LEMA 3.1 (LEMA DA COBERTURA DE VITALI) *Seja X uma parte de $\mathbb{R}$ tal que $m^*(X) < +\infty$ e seja $\mathcal{C}$ uma família de intervalos de $\mathbb{R}$ com mais que um* *ponto tal que, para cada $x \in X$ e cada $\varepsilon > 0$, exista algum $I \in \mathcal{C}$ tal que $x \in I$ e que* $\operatorname{comp}(I) < \varepsilon$. *Então, para cada $\varepsilon > 0$, existem elementos $I_1, \ldots, I_n$ de $\mathcal{C}$ dois a dois disjuntos tais que*

$$m^*\left(X \Big\backslash \bigcup_{k=1}^{n} I_k\right) < \varepsilon. \tag{3.1}$$

DEMONSTRAÇÃO: Basta fazer a demonstração no caso em que todos os elementos de $\mathcal{C}$ são intervalos fechados, pois então, no caso geral, considera-se o conjunto

$$\mathcal{C}' = \left\{\, \overline{I} \mid I \in \mathcal{C} \,\right\},$$

o qual satisfaz as hipóteses do lema. Logo, haverá, para cada $\varepsilon > 0$, elementos $I_1, \ldots, I_n$ de $\mathcal{C}$ dois a dois disjuntos tais que

$$l\left(X \Big\backslash \bigcup_{k=1}^{n} \overline{I_k}\right) < \varepsilon,$$

de onde resulta que se tem (3.1), pois os conjuntos

$$X \Big\backslash \bigcup_{k=1}^{n} \overline{I_k} \quad \text{e} \quad X \Big\backslash \bigcup_{k=1}^{n} I_k$$

têm a mesma medida.

Seja A um aberto de $\mathbb{R}$ de medida finita que contenha X; um tal aberto existe necessariamente pela definição da medida exterior de Lebesgue e por se estar a supor que X tem medida exterior finita. Seja

$$\mathcal{C}^* = \{\, I \in \mathcal{C} \mid I \subset A \,\}.$$

O conjunto $\mathcal{C}^*$ satisfaz então as hipóteses do lema. Por outro lado, se $I \in \mathcal{C}^*$, então $\operatorname{comp}(I) \leqslant l(A)$. Logo, o conjunto $\{\, \operatorname{comp}(I) \mid I \in \mathcal{C}^* \,\}$ é majorado (em $\mathbb{R}$); seja s_0 o seu supremo. Fixemos $I_1 \in \mathcal{A}^*$ tal que $\operatorname{comp}(I_1) > s_0/2$. Caso $X \subset I_1$, então $X \backslash I_1 = \emptyset$ e o lema está demonstrado. Caso contrário, consideram-se os intervalos $I \in \mathcal{A}^*$ que não intersectam I_1; existem tais intervalos, pelas hipóteses do lema e porque I_1 é fechado. Seja s_1 o supremo das medidas dos intervalos de $\mathcal{C}^*$ que não intersectam I_1 e seja I_2 um intervalo de $\mathcal{C}^*$ que não intersecta I_1 tal que $\operatorname{comp}(I_2) > s_1/2$. Caso $X \subset I_1 \cup I_2$, então $X \backslash (I_1 \cup I_2) = \emptyset$ e o lema estará demonstrado. Caso contrário, recomeça-se o processo: seja s_2 o supremo das medidas dos intervalos de $\mathcal{C}^*$ que não intersectam $I_1 \cup I_2$ e seja I_3 um intervalo de

$\mathcal{C}^*$ que não intersecta $I_1 \cup I_2$ tal que $\text{comp}(I_3) > s_2/2$. Podem dar-se dois casos: ou ao fim de um número finito de passos, obtiveram-se intervalos $I_1, I_2, \ldots, I_n$ tais que $X \subset \bigcup_{k=1}^n I_k$ e então $X \setminus \left(\bigcup_{k=1}^n I_k\right) = \emptyset$, ou o processo não acaba e leva nesse caso a uma sucessão $(I_n)_{n\in\mathbb{N}}$ de intervalos dois a dois disjuntos tal que $(\forall n \in \mathbb{N}) : \text{comp}(I_n) > s_{n-1}/2$. No primeiro caso, nada haverá a demonstrar, pelo que se vai supor que se está no segundo.

É conveniente observar que, visto que os intervalos I_n ($n \in \mathbb{N}$) são dois a dois disjuntos e visto que a sua reunião está contida em A, então

$$\sum_{n=1}^{+\infty} \text{comp}(I_n) = m^*\left(\bigcup_{n\in\mathbb{N}} I_n\right) \leqslant m^*(A) < +\infty;$$

em particular, $\lim_{n\in\mathbb{N}} \text{comp}(I_n) = 0$. Seja $\varepsilon > 0$ e seja $n \in \mathbb{N}$ tal que

$$\sum_{m=n+1}^{+\infty} \text{comp}(I_m) < \frac{\varepsilon}{5}.$$

Vai-se provar que se tem (3.1) para esta escolha de $I_1, I_2, \ldots, I_n$ e isso será feito provando que, se se definir, para cada $m \in \mathbb{N}$, J_m como sendo o intervalo aberto centrado no centro de I_m tal que $\text{comp}(J_m) = 5\,\text{comp}(I_m)$, então

$$X \Big\backslash \bigcup_{k=1}^{n} I_k \subset \bigcup_{m>n} J_m.$$

Resultará daqui que

$$m^*\left(X \Big\backslash \bigcup_{k=1}^{n} I_k\right) \leqslant m^*\left(\bigcup_{m>n} J_m\right) \leqslant \sum_{m>n} 5\,\text{comp}(I_m) < \varepsilon.$$

Seja $x \in X \big\backslash \left(\bigcup_{k=1}^n I_k\right)$; quer-se provar que $x \in J_m$ para algum algum número natural $m > n$. Seja $I \in \mathcal{C}^*$ tal que $x \in I$ e $I \cap \left(\bigcup_{k=1}^n I_k\right) = \emptyset$. Se $m \in \mathbb{N}$ e se $I \cap \left(\bigcup_{k=1}^m I_k\right) = \emptyset$, então, pela definição de s_m,

$$\text{comp}(I) \leqslant s_m < 2\,\text{comp}(I_{m+1}). \tag{3.2}$$

Como I é um intervalo com mais que um ponto, $\text{comp}(I) > 0$, pelo que não se pode ter (3.2) para qualquer $m \in \mathbb{N}$. Consequentemente, existe algum $m \in \mathbb{N}$ tal que I intersecta a reunião $\bigcup_{k=1}^{m+1} I_k$, mas não a reunião $\bigcup_{k=1}^m I_k$ e é óbvio, pela escolha de I, que $m \geqslant n$. Mas então I intersecta I_{m+1} e $\text{comp}(I) \leqslant s_m < 2\,\text{comp}(I_{m+1})$.

Seja y o centro do intervalo I_{m+1}; vai-se provar que

$$|x-y| < \frac{5}{2}\operatorname{comp}(I_{m+1}),$$

de onde resulta que $x \in J_{m+1}$. Para tal, seja $z \in I \cap I_{m+1}$; então

$$\begin{aligned}|x-y| &\leqslant |x-z| + |z-y| \\ &\leqslant \operatorname{comp}(I) + \frac{\operatorname{comp}(I_{m+1})}{2} \text{ (pois } x, z \in I) \\ &< \frac{5}{2}\operatorname{comp}(I_{m+1}),\end{aligned}$$

por (3.2). ■

Seja I uma parte de $\mathbb{R}$, seja f uma função de I em $\mathbb{R}$ e seja $a \in I$. O limite superior e o limite inferior da função f no ponto a são os elementos de $\overline{\mathbb{R}}$ definidos por

$$\limsup_{x\to a} f(x) = \lim_{r\to 0^+} \sup\{\, f(x) \mid x \in I\cap]-r,r[\,\}$$

e por

$$\liminf_{x\to a} f(x) = \lim_{r\to 0^+} \inf\{\, f(x) \mid x \in I\cap]-r,r[\,\}$$

respectivamente. É imediato que o limite superior da função f no ponto a é maior ou igual ao limite inferior de f nesse ponto. Por outro lado, tal como no caso do limite superior e do limite inferior de sucessões, pode-se provar que o limite $\lim_{x\to a} f(x)$ existe se e só se $\limsup_{x\to a} f(x) = \liminf_{x\to a} f(x)$ e que, caso estas condições se verifiquem,

$$\limsup_{x\to a} f(x) = \lim_{x\to a} f(x) = \liminf_{x\to a} f(x).$$

Foi visto, no decorrer da demonstração do corolário 1.3, que se $a, b \in \mathbb{R}$ com $a < b$ e se f é uma função monótona de $[a,b]$ em $\mathbb{R}$, então o conjunto dos pontos de descontinuidade de f é finito ou numerável; em particular, f é contínua q. s. Vai ser agora demonstrado um teorema de Lebesgue que generaliza este resultado.

TEOREMA 3.2 (TEOREMA DA DERIVAÇÃO DE LEBESGUE) *Se f é uma função monótona de um intervalo $[a,b]$ em $\mathbb{R}$, então f é derivável q. s.*

DEMONSTRAÇÃO: Vai-se supor que f é monótona crescente. Se f for monótona decrescente, a demonstração é análoga, ou então pode-se empregar o facto de os pontos onde f é derivável são exactamente os pontos onde $-f$ é derivável.

Se $c \in [a, b]$ sejam

$$\overline{D}f(c) = \limsup_{h \to 0} \frac{f(c+h) - f(c)}{h} \quad \text{e} \quad \underline{D}f(c) = \liminf_{h \to 0} \frac{f(c+h) - f(c)}{h}.$$

Então f é derivável em c se e só se $\overline{D}f(c) = \underline{D}f(c) \in \mathbb{R}$ e, nesse caso, $\overline{D}f(c) = f'(c) = \underline{D}f(c)$. Então, para demonstrar o teorema basta provar que ambos os conjuntos

$$\{\, c \in]a, b[\mid \overline{D}f(c) > \underline{D}f(c) \,\}, \tag{3.3}$$

e

$$\{\, c \in]a, b[\mid \overline{D}f(c) = +\infty \,\} \tag{3.4}$$

têm medida exterior nula.[1]

Vai-se começar por provar que o conjunto (3.3) tem medida exterior nula. Se $\alpha, \beta \in \mathbb{R}_+^*$ são tais que $\alpha < \beta$, seja

$$X_{\alpha,\beta} = \{\, c \in]a, b[\mid \overline{D}f(c) > \beta > \alpha > \underline{D}f(c) \,\};$$

vai-se provar que todos os conjuntos desta forma têm medida exterior nula. Resultará deste facto que o conjunto (3.3) tem medida exterior nula, visto que poderá ser escrito como reunião numerável de conjuntos com medida nula, pois tem-se

$$\{\, c \in]a, b[\mid \overline{D}f(c) > \underline{D}f(c) \,\} = \bigcup_{\substack{\alpha,\beta \in \mathbb{Q}_+^* \\ \alpha < \beta}} X_{\alpha,\beta}.$$

Seja $\varepsilon > 0$. Resulta da definição de medida exterior de Lebesgue que existe algum aberto A de $\mathbb{R}$ tal que $X_{\alpha,\beta} \subset A$ e que $m^*(A) \leqslant m^*(X_{\alpha,\beta}) + \varepsilon$. Se $c \in X_{\alpha,\beta}$, então $\alpha > \underline{D}f(c)$ pelo que há números $h \neq 0$ arbitrariamente pertos de 0 e tais que $\alpha > (f(c+h) - f(c))/h$. Consideremos todos os intervalos contidos em $]a, b[\cap A$ de uma das seguintes formas:

- $[c, c+h]$, com $c \in X_{\alpha,\beta}$ e com $h > 0$ tal que $\alpha > \dfrac{f(c+h) - f(c)}{h}$;

[1] De facto, prova-se facilmente que as funções $\overline{D}f$ e $\underline{D}f$ são ambas mensuráveis, de onde resulta que os conjuntos (3.3) e (3.4) são mensuráveis. Consequentemente, afirmar que têm medida exterior nula é o mesmo que afirmar que têm medida nula.

- $[c+h,c]$, com $c \in X_{\alpha,\beta}$ e com $h < 0$ tal que $\alpha > \dfrac{f(c+h)-f(c)}{h}$.

O conjunto $\mathcal{C}_{\alpha,\beta}$ de todos estes intervalos está nas condições do lema da cobertura de Vitali. Existem então intervalos $I_1, \ldots, I_N \in \mathcal{C}_{\alpha,\beta}$ dois a dois disjuntos tais que $m^*\big(X_{\alpha,\beta} \setminus (I_1 \cup \cdots \cup I_N)\big) \leqslant \varepsilon$. Se d pertence a $X_{\alpha,\beta} \cap \big(\mathring{I}_1 \cup \cdots \cup \mathring{I}_N\big)$, então, visto que $\overline{D}f(d) > \beta$, pode-se tomar $k \neq 0$ arbitrariamente perto de 0 tal que $(f(d+k)-f(d))/k > \beta$ e que $[d,d+k]$ (respectivamente $[d+k,d]$) esteja contido em algum intervalo I_l, caso $k > 0$ (resp. $k < 0$). Temos assim uma família $\mathcal{D}_{\alpha,\beta}$ de intervalos que satisfaz as condições do lema da cobertura de Vitali, relativamente ao conjunto $X_{\alpha,\beta} \cap \big(\mathring{I}_1 \cup \cdots \cup \mathring{I}_N\big)$. Portanto, existem intervalos $J_1, \ldots, J_M \in \mathcal{D}_{\alpha,\beta}$ dois a dois disjuntos tais que

$$m^*\Big(X_{\alpha,\beta} \cap \big(\mathring{I}_1 \cup \cdots \cup \mathring{I}_N\big) \setminus \big(J_1 \cup \cdots \cup J_M\big)\Big) \leqslant \varepsilon. \tag{3.5}$$

Mas

$$\begin{aligned} m^*(X_{\alpha,\beta}) &= m^*\Big(X_{\alpha,\beta} \setminus \big(I_1 \cup \cdots \cup I_N\big)\Big) + m^*\Big(X_{\alpha,\beta} \cap \big(I_1 \cup \cdots \cup I_N\big)\Big) \\ &\leqslant \varepsilon + m^*\Big(X_{\alpha,\beta} \cap \big(\mathring{I}_1 \cup \cdots \cup \mathring{I}_N\big)\Big) \end{aligned}$$

Por outro lado, $m^*\Big(X_{\alpha,\beta} \cap \big(\mathring{I}_1 \cup \cdots \cup \mathring{I}_N\big)\Big)$ é menor ou igual à soma de

$$m^*\Big(X_{\alpha,\beta} \cap \big(\mathring{I}_1 \cup \cdots \cup \mathring{I}_N\big) \setminus \big(J_1 \cup \cdots \cup J_M\big)\Big) \tag{3.6}$$

com

$$m^*\Big(X_{\alpha,\beta} \cap \big(\mathring{I}_1 \cup \cdots \cup \mathring{I}_N\big) \cap \big(J_1 \cup \cdots \cup J_M\big)\Big). \tag{3.7}$$

Mas (3.6) é menor ou igual a ε (por (3.5)) e (3.7) é menor ou igual a

$$m^*\big(J_1 \cup \cdots \cup J_M\big) \leqslant |k_1| + \cdots + |k_M|.$$

Isto prova então que $m^*(X_{\alpha,\beta}) \leqslant 2\varepsilon + \sum_{i=1}^M |k_i|$. Por outro lado, uma vez que

$$(\forall i \in \{1,2,\ldots,M\}) : \frac{|f(d_i+k_i)-f(d_i)|}{|k_i|} = \frac{f(d_i+k_i)-f(d_i)}{k_i} > \beta,$$

tem-se que

$$\beta m^*(X_{\alpha,\beta}) \leqslant 2\varepsilon\beta + \sum_{i=1}^{M} \beta|k_i| \leqslant 2\varepsilon\beta + \sum_{i=1}^{M} |f(d_i+k_i)-f(d_i)|. \tag{3.8}$$

Mas cada intervalo $[d_i, d_i + k_i]$ (caso $k_i > 0$) ou $[d_i + k_i, d_i]$ (caso $k_i < 0$) estão contido em algum intervalo $[c_j, c_j + h_j]$ ou $[c_j + h_j, c_j]$, pelo que

$$\begin{aligned}\sum_{i=1}^{M} |f(d_i + k_i) - f(d_i)| &\leqslant \sum_{j=1}^{N} |f(c_i + h_i) - f(c_i)| \\ &< \alpha \sum_{j=1}^{N} |h_j| \\ &\leqslant \alpha m^*(A) \\ &\leqslant \alpha\big(m^*(X_{\alpha,\beta}) + \varepsilon\big).\end{aligned}$$

Resulta de (3.8) e desta série de desigualdades que

$$\beta m^*(X_{\alpha,\beta}) \leqslant 2\varepsilon\beta + \alpha m^*(X_{\alpha,\beta}) + \alpha\varepsilon,$$

o que equivale a afirmar que

$$m^*(X_{\alpha,\beta}) \leqslant \varepsilon\frac{2\beta + \alpha}{\beta - \alpha}.$$

Como ε é arbitrário, resulta desta desigualdade que $m^*(X_{\alpha,\beta}) = 0$.

Vejamos agora porque é que o conjunto (3.4) tem medida exterior nula, o que terminará a demonstração. Seja $\varepsilon > 0$. Analogamente ao que foi feito atrás, define-se, para cada $\beta \in \mathbb{R}_+^*$,

$$X_\beta = \left\{\, c \in]a, b[\;\middle|\; \overline{D}f(c) > \beta \,\right\}.$$

Como o conjunto (3.4) é a intersecção de todos os conjuntos da forma X_β, se se provar que $m^*(X_\beta) < \varepsilon$ para β suficientemente grande, estará provado que o conjunto (3.4) também tem medida exterior menor do que ε; como ε é arbitrário, resultará então que o conjunto (3.4) tem medida exterior nula.

Por um processo análogo à demonstração de que o conjunto (3.3) tem medida exterior nula, pode-se mostrar que há intervalos $I_1, I_2, \ldots, I_N$ dois a dois disjuntos e todos da forma $[c_i, c_i + h_i]$ ou da forma $[c_i + h_i, c_i]$ tais que

$$m^*\big(X_\beta \backslash (I_1 \cup \cdots \cup I_N)\big) < \frac{\varepsilon}{2}.$$

Então

$$\begin{aligned} m^*(X_\beta) &\leqslant m^*\big(X_\beta\backslash(I_1\cup\cdots I_N)\big)+m^*\big(X_\beta\cap(I_1\cup\cdots I_N)\big)\\ &\leqslant \frac{\varepsilon}{2}+\sum_{j=1}^N \operatorname{comp}(I_j)\\ &= \frac{\varepsilon}{2}+\frac{1}{\beta}\sum_{j=1}^N \beta|h_i|\\ &\leqslant \frac{\varepsilon}{2}+\frac{1}{\beta}\sum_{j=1}^N |f(c_j+h_j)-f(c_j)|\\ &\leqslant \frac{\varepsilon}{2}+\frac{f(b)-f(a)}{\beta}. \end{aligned}$$

Logo, $m^*(X_\beta) < \varepsilon$ quando β for suficientemente grande. ∎

Pode ser vista em [3] uma demonstração deste teorema que não emprega nem o lema da cobertura de Vitali nem qualquer teoria da medida para lá de conjuntos de medida nula.[2]

Dado que, se se verificarem as hipóteses deste teorema, o conjunto dos pontos de descontinuidade de f é finito ou numerável, é natural que se ponha a questão de se saber se o mesmo será ou não verdade para o conjunto dos pontos onde f não é derivável. Vai-se ver que a resposta é negativa.

EXEMPLO 3.1 Seja C o conjunto de Cantor e considere-se a função f de C no intervalo fechado $[0,1]$ definida na página 23. Vai-se prolongar f a uma função (que também será representada por f) monótona crescente de $[0,1]$ em $[0,1]$. Comecemos por ver como definir $f(x)$ quando $x \in [1/3, 2/3]$. De facto, há apenas uma maneira de o fazer: visto que $f(1/3) = f(2/3) = 1/2$ e visto que se pretende que f seja crescente, tem-se então necessariamente $f(x) = 1/2$ para cada $x \in [1/3, 2/3]$. O mesmo argumento aplica-se aos intervalos $[1/9, 2/9]$ e $[7/9, 8/9]$; pelo argumento anterior, f toma sempre o valor $1/4$ no primeiro destes intervalos e $3/4$ no segundo. Prosseguindo deste modo, obtém-se um função crescente de $[0,1]$ em $[0,1]$, a qual é obviamente derivável (com derivada nula) em todos os pontos de $[0,1]\backslash C$ e pode-se provar que o conjunto dos pontos onde f não é derivável é precisamente o conjunto de Cantor. Esta função designa-se por *função de*

[2]Observe-se que a demonstração que foi feita pode ser facilmente reescrita numa linguagem ligeiramente diferente de maneira a serem eliminadas quaisquer referências à teoria da medida, com excepção do conceito de conjunto de medida nula.

Cantor e o seu gráfico está esboçado na figura 3.1. Observe-se que a função de Cantor é contínua, pois a sua imagem é $[0,1]$ e, dada uma função real monótona definida num intervalo de $\mathbb{R}$, ela é contínua se e só se a sua imagem for um intervalo.

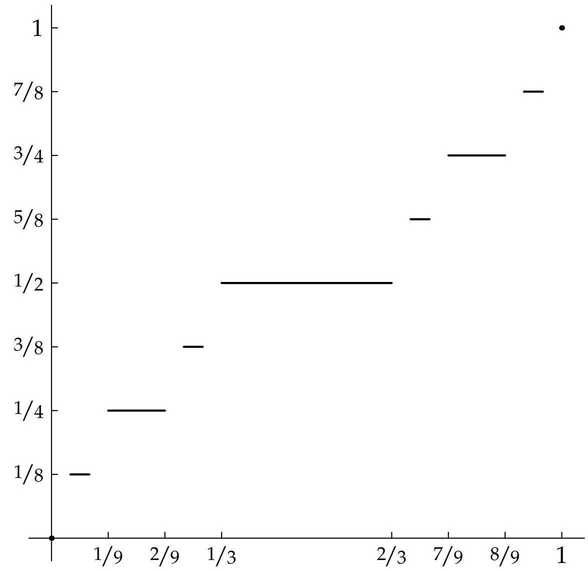

Figura 3.1: Esboço do gráfico da função de Cantor

TEOREMA 3.3 (TEOREMA DA DERIVAÇÃO DE FUBINI) *Seja I um intervalo de $\mathbb{R}$, seja $\sum_{n=1}^{+\infty} f_n$ uma série pontualmente convergente de funções monótonas crescentes de I em $\mathbb{R}$ e seja $f = \sum_{n=1}^{+\infty} f_n$. Então*

$$f'(x) = \sum_{n=1}^{+\infty} f_n'(x) \text{ q. s.} \tag{3.9}$$

DEMONSTRAÇÃO: Sejam $a, b \in I$ tais que $a < b$; vai-se provar que a relação (3.9) é válida em $]a,b[$, de onde se deduz que é válida em I.

Vai-se supor que $(\forall n \in \mathbb{N}) : f_n(a) = 0$, o que implica que $f(a) = 0$. Se se demonstrar o teorema sob esta hipótese, então o teorema ficará demonstrado no caso geral, visto que $f' = (f - f(a))'$ e que $(\forall n \in \mathbb{N}) : f_n' = (f_n - f_n(a))'$.

Uma vez que $\sum_{n=1}^{+\infty} f_n(b) = f(b)$, existe, para cada $k \in \mathbb{N}$, algum $n(k) \in \mathbb{N}$ tal que

$$\sum_{m=n(k)}^{+\infty} f_m(b) < 2^{-k}.$$

Naturalmente, podem-se escolher os números $n(k)$ de modo a que a sucessão $(n(k))_{k\in\mathbb{N}}$ seja estritamente crescente. Define-se então, para cada $k \in \mathbb{N}$,

$$\begin{aligned} t_k\colon\ ]a,b[&\longrightarrow \mathbb{R} \\ x &\mapsto \sum_{m=n(k)}^{+\infty} f_m(x). \end{aligned}$$

Cada função t_k é monótona crescente e

$$(\forall x \in [a,b]) : 0 \leqslant t_k(x) \leqslant 2^{-k}.$$

Faz então sentido definir a função

$$\begin{aligned} t\colon\ ]a,b[&\longrightarrow \mathbb{R} \\ x &\mapsto \sum_{k=1}^{+\infty} t_k(x), \end{aligned}$$

a qual é monótona crescente.

Seja $E \subset]a,b[$ um conjunto de medida nula tal que cada uma das funções f, f_m ($m \in \mathbb{N}$), t e t_k ($k \in \mathbb{N}$) seja derivável em cada ponto de $]a,b[\setminus E$. Para cada $n \in \mathbb{N}$, seja

$$s_n = \sum_{k=n+1}^{+\infty} t_k = t - \sum_{k=1}^{n} t_k,$$

a qual também é derivável em todos os pontos de $]a,b[\setminus E$. Se $x \in]a,b[\setminus E$, então, para cada $n \in \mathbb{N}$

$$t'(x) - s_n'(x) = \left(\sum_{k=1}^{n} t_k\right)'(x) = \sum_{k=1}^{n} t_k'(x)$$

e, visto que $s_n'(x) \geq 0$, tem-se $t'(x) \geqslant \sum_{k=1}^{n} t_k'(x)$. Logo, visto que a série $\sum_{n=1}^{+\infty} t_k'(x)$ é uma série de números maiores ou iguais a 0, ela converge e, portanto, $\lim_{k\in\mathbb{N}} t_k'(x) = 0$. Isto prova então que

$$\begin{aligned} 0 &= \lim_{k\in\mathbb{N}} t_k'(x) \\ &= \lim_{k\in\mathbb{N}} \left(f - \sum_{m=1}^{n(k)-1} f_m\right)'(x) \\ &= f'(x) - \lim_{k\in\mathbb{N}} \sum_{m=1}^{n(k)-1} f_m'(x). \end{aligned}$$

Logo, a sucessão $\left(\sum_{m=1}^{n(k)-1} f'_m(x)\right)_{k\in\mathbb{N}}$ converge para $f'(x)$. Mas trata-se de uma sub-sucessão da sucessão crescente $\left(\sum_{n=1}^{N} f'_m(x)\right)_{N\in\mathbb{N}}$, pelo que esta última também converge para $f'(x)$; por outras palavras,

$$f'(x) = \sum_{n=1}^{+\infty} f'_m(x).$$

■

3.2 O teorema fundamental do Cálculo

Se $A, B \in \mathcal{M}(\mathbb{R})$, com $B \subset A$ e $l(A \setminus B) = 0$, e se f é uma função integrável de B em $\overline{\mathbb{R}}$, vai-se empregar a notação:

$$\int_A f\,dl = \int_B f\,dl.$$

Observe-se que se F for um prolongamento qualquer da função f a A, então $\int_A F\,dl = \int_A f\,dl$, pois

$$\begin{aligned}\int_A F\,dl &= \int_{B\,\dot\cup(A\setminus B)} F\,dl\\ &= \int_B f\,dl \text{ (pois } F|_B = f \text{ e } l(B\setminus A) = 0)\\ &= \int_A f\,dl.\end{aligned}$$

Com esta notação, faz sentido considerar o seguinte problema: se $a, b \in \mathbb{R}$ (com $a < b$) e $f\colon [a,b] \longrightarrow \mathbb{R}$ é uma função contínua e derivável q. s., tem-se ou não necessariamente

$$\int_a^b f'(x)\,dx = f(b) - f(a)?$$

A resposta é negativa, pois viu-se (exemplo 3.1) que é possível definir uma função não constante de $[0,1]$ em $[0,1]$ cuja derivada é nula q. s.. Aliás, este não é o único problema que pode surgir.

EXEMPLO 3.2 A função

$$\begin{aligned} f\colon\ [0,1] &\longrightarrow \mathbb{R}\\ x &\mapsto \begin{cases} 0 & \text{se } x = 0\\ x^2 \operatorname{sen}\left(\frac{1}{x^2}\right) & \text{caso contrário}\end{cases}\end{aligned}$$

é derivável em todos os pontos, mas f' não é integrável, pois verifica-se facilmente que $|f'|$ não é integrável.

Apesar destes exemplos, é possível, sob hipóteses adicionais, relacionar $\int_a^b f'(x)\,dx$ com $f(b) - f(a)$ quando f é derivável q. s. Por exemplo, se as hipóteses do teorema da derivação de Lebesgue se verificarem, então tem-se

$$\int_a^b f'(x)\,dx \leqslant f(b) - f(a). \tag{3.10}$$

De facto, prolongue-se f a uma função (que também será representada por f) de $[a, b+1]$ em $\mathbb{R}$ definindo-se $f(x) = f(b)$ quando $x \geqslant b$. Para cada $n \in \mathbb{N}$ e para cada $x \in [a, b]$ define-se então

$$\varphi_n(x) = n\big(f(x + 1/n) - f(x)\big).$$

Então

1. $(\forall n \in \mathbb{N}) : \varphi_n \geqslant 0$;
2. $\lim_{n\in\mathbb{N}} \varphi_n(x) = f'(x)$ q. s.

Observe-se que se tem então, para cada $n \in \mathbb{N}$:

$$\begin{aligned}
\int_a^b \varphi_n(x)\,dx &= n \int_a^b f(x + 1/n) - f(x)\,dx \\
&= n\Big(\int_a^b f(x + 1/n)\,dx - \int_a^b f(x)\,dx\Big) \\
&= n\Big(\int_{a+1/n}^{b+1/n} f(x)\,dx - \int_a^b f(x)\,dx\Big) \\
&= n\Big(\int_b^{b+1/n} f(x)\,dl - \int_a^{a+1/n} f(x)\,dx\Big) \\
&\leqslant f(b) - f(a),
\end{aligned}$$

pois $f(x) = f(b)$ em todos os pontos de $[b, b + 1/n]$ e $f(x) \geqslant f(a)$ em todos os pontos de $[a, a + 1/n]$. Então, pelo teorema de Fatou:

$$\begin{aligned}
\int_a^b f'(x)\,dx &= \int_a^b \lim_{n\in\mathbb{N}} \varphi_n(x)\,dx \\
&\leqslant \liminf_{n\in\mathbb{N}} \int_a^b \varphi_n(x)\,dx \\
&\leqslant f(b) - f(a).
\end{aligned}$$

DEFINIÇÃO 3.1 Seja $A \subset \mathbb{R}$ e seja f uma função de A em $\overline{\mathbb{R}}$. Diz-se que a função f é *localmente integrável* se cada $x \in A$ possuir alguma vizinhança V tal que a restrição de f a $A \cap V$ seja integrável.

EXEMPLO 3.3 A função identidade de $\mathbb{R}$ em $\mathbb{R}$ é localmente integrável, pois a sua restrição a qualquer intervalo limitado é integrável e qualquer ponto tem vizinhanças que são intervalos limitados.

EXEMPLO 3.4 A função

$$\begin{array}{ccc} \mathbb{R}_+ & \longrightarrow & \overline{\mathbb{R}} \\ x & \mapsto & \begin{cases} 1/x & \text{se } x > 0 \\ +\infty & \text{se } x = 0 \end{cases} \end{array}$$

não é localmente integrável, pois não existe qualquer vizinhança V de 0 tal que a restrição de f a $V \cap \mathbb{R}_+$ seja integrável.

PROPOSIÇÃO 3.1 *Sejam I um intervalo de $\mathbb{R}$ e f uma função mensurável de I em $\overline{\mathbb{R}}$. São então condições equivalentes:*

1. *a função f é localmente integrável;*
2. *se $a, b \in I$ e $a < b$, então $f|_{[a,b]}$ é integrável.*

DEMONSTRAÇÃO: É claro que a segunda condição implica a primeira. Suponha-se que a primeira condição se verifica e sejam a e b elementos de I tais que $a < b$; quer-se provar que $f|_{[a,b]}$ é integrável. Para cada $x \in [a, b]$, seja I_x um intervalo aberto que contenha x e tal que a restrição de f a $I_x \cap I$ seja integrável. Então $[a, b] \subset \bigcup_{x\in[a,b]} I_x$ e então, pelo lema 1.2, existe alguma partição P de $[a, b]$ tal que cada intervalo da partição está contido em I_x, para algum $x \in [a, b]$; em particular, a restrição de f a cada intervalo da partição é integrável. Mas então, uma vez que $[a, b]$ é a reunião daquele conjunto finito de intervalos, a restrição de f a $[a, b]$ é integrável. ∎

Sendo assim, se f for uma função localmente integrável de um intervalo I de $\mathbb{R}$ em $\overline{\mathbb{R}}$ e se $a \in I$, então faz sentido considerar, para cada $x \in I$, o integral $\int_a^x f(t)\,dt$ (caso $x \geqslant a$) bem como o integral $\int_x^a f(t)\,dt$ (caso $x \leqslant a$). Vai-se empregar a notação $\int_a^x f(t)\,dt$ para representar o número $-\int_x^a f(t)\,dt$ quando $x \leqslant a$. Com esta convenção, tem-se, para quaisquer $a, b, c \in I$,

$$\int_a^c f(x)\,dx = \int_a^b f(x)\,dx + \int_b^c f(x)\,dx. \tag{3.11}$$

TEOREMA 3.4 *Seja I um intervalo de $\mathbb{R}$, seja $c \in I$, seja f uma função localmente integrável de I em $\overline{\mathbb{R}}$ e seja*

$$\begin{array}{rccc} F\colon & I & \longrightarrow & \mathbb{R} \\ & x & \mapsto & \displaystyle\int_c^x f(t)\,dt. \end{array}$$

Então tem-se $F'(x) = f(x)$ q. s.

DEMONSTRAÇÃO: Para simplificar, vai-se supor que $I = \mathbb{R}$. A demonstração do teorema neste caso particular implica que ele é válido no caso geral, pois se $I \neq \mathbb{R}$, pode-se prolongar f a $\mathbb{R}$ pondo $f(x) = 0$ se $x \in \mathbb{R} \setminus I$. A «nova» função F obtida a partir deste prolongamento é um prolongamento a $\mathbb{R}$ da função F original (que é constante em $\mathbb{R} \setminus I$). Então $F'(x) = f(x)$ para quase todos os $x \in \mathbb{R}$ e, em particular, $F'(x) = f(x)$ para quase todos os $x \in I$.

Vai-se começar por demonstrar este teorema no caso particular em que $F \equiv 0$. Nesse caso, $F' \equiv 0$ e o que se quer demonstrar é então que $f(x) = 0$ q. s. Observe-se que afirmar que $F \equiv 0$ é o mesmo que afirmar que, para qualquer $a \in \mathbb{R}$, $\int_c^a f(x)\,dx = 0$. Mas então, se $a, b \in \mathbb{R}$ tem-se

$$\begin{aligned} \int_a^b f(x)\,dx &= \int_c^b f(x)\,dx - \int_c^a f(x)\,dx \text{ (por (3.11))} \\ &= F(b) - F(a) \\ &= 0. \end{aligned}$$

Vai-se provar que

$$A \in \mathcal{M}(\mathbb{R}) \Longrightarrow \int_A f\,dl = 0, \tag{3.12}$$

de onde resulta que $f(x) = 0$ q. s., pois se $A = \{\, x \in \mathbb{R} \mid f(x) \geqslant 0 \,\}$, tem-se

$$\int_A f\,dl = 0 \Longleftrightarrow \int_{\mathbb{R}} f^+\,dl = 0 \Longrightarrow f^+(x) = 0 \text{ q. s.}$$

e, analogamente, $f^-(x) = 0$ q. s., pelo que $f(x) = 0$ q. s., pois $f = f^+ - f^-$.

Comece-se por supor que A é limitado. Para cada $n \in \mathbb{N}$ existe, pela definição de medida exterior, algum aberto A_n que contém A e tal que $m(A_n \setminus A) < 1/n$; naturalmente, visto que se está a supor que A é limitado, pode-se supor que cada A_n é limitado. O aberto A_n pode ser escrito como uma reunião finita ou numerável de intervalos limitados de $\mathbb{R}$ dois a dois disjuntos, pelo lema 1.1. Mas resulta então de se estar a supor que se tem sempre $\int_a^b f(x)\,dx = 0$ e da proposição 2.6 que $\int_{A_n} f\,dl$ é a soma dos integrais de f em cada um daqueles intervalos e, portanto, é igual a 0.

Seja $A^* = \bigcap_{n\in\mathbb{N}} A_n$. Então $A^* \supset A$ e $l(A^* \setminus A) = 0$. Por outro lado, as funções

$$\begin{array}{ccc} \mathcal{M}(\mathbb{R}) & \longrightarrow & \overline{\mathbb{R}_+} \\ B & \mapsto & \int_B f^+\,dl \end{array} \quad \text{e} \quad \begin{array}{ccc} \mathcal{M}(\mathbb{R}) & \longrightarrow & \overline{\mathbb{R}_+} \\ B & \mapsto & \int_B f^-\,dl \end{array}$$

são medidas e, portanto, pela proposição 1.5,

$$\begin{aligned} \int_{A^*} f\,dl &= \int_{A^*} f^+\,dl - \int_{A^*} f^-\,dl \\ &= \lim_{n\in\mathbb{N}} \left(\int_{A_n} f^+\,dl - \int_{A_n} f^-\,dl \right) \\ &= 0 \end{aligned}$$

Mas, por outro lado,

$$0 = \int_{A^*} f\,dl = \int_A f\,dl + \int_{A^*\setminus A} f\,dl = \int_A f\,dl,$$

pois $l(A^* \setminus A) = 0$.

Finalmente, se A não for limitado é sempre possível escrever A como reunião disjunta uma sucessão de partes mensuráveis e limitadas de $\mathbb{R}$. Como o integral de f em cada uma destas partes é igual a 0, $\int_A f\,dl = 0$, pela proposição 2.6.

Passemos agora ao caso geral. Basta demonstrar o teorema no caso em que $f \geqslant 0$. De facto, uma vez demonstrado neste caso, pode-se demonstrar o teorema no caso geral definindo $F_+, F_- : \mathbb{R} \longrightarrow \mathbb{R}$ por

$$F_+(x) = \int_c^x f^+(t)\,dt \quad \text{e por} \quad F_-(x) = \int_c^x f^-(t)\,dt.$$

Então, para cada $x \in \mathbb{R}$,

$$F(x) = \int_c^x f(t)\,dt = \int_c^x f^+(t)\,dt - \int_c^x f^-(t)\,dt = F_+(x) - F_-(x)$$

e, como se tem $F_+'(x) = f^+(x)$ e $F_-'(x) = f^-(x)$ q. s., tem-se

$$F'(x) = F_+'(x) - F_-'(x) = f^+(x) - f^-(x) = f(x) \text{ q. s.}$$

Suponha-se então que $f \geqslant 0$ (o que implica, pelo teorema da derivação de Lebesgue, que F é derivável em quase todos os pontos de $[a,b]$) e suponha-se também que f é majorada por algum $M \in \mathbb{R}_+$. Quer-se provar que $F'(x) = f(x)$ q. s. e, para tal, vai-se provar que se $a, b \in \mathbb{R}$ são tais que $a < b$, então $F'(x) = f(x)$ em quase todos os pontos de $[a, b]$, o que equivale a afirmar, pelo caso particular do teorema que foi demonstrado no início, que

$$(\forall a, b \in \mathbb{R}) : a < b \Longrightarrow \int_a^b F'(t) - f(t)\, dt = 0. \tag{3.13}$$

Vei-se supor que $c = a$, o que faz somente com que a função F seja substituída por uma outra função (que também será representada por F) que difere daquela por uma constante; consequentemente, têm a mesma derivada. Definem-se

$$\begin{aligned} f^*\colon \quad [a, b+1] & \longrightarrow \quad \mathbb{R} \\ x \quad & \mapsto \begin{cases} f(x) & \text{se } x \leqslant b \\ f(b) & \text{caso contrário.} \end{cases} \end{aligned}$$

e

$$\begin{aligned} F^*\colon \quad [a, b+1] & \longrightarrow \quad \mathbb{R} \\ x \quad & \mapsto \int_a^x f^*(t)\, dt; \end{aligned}$$

define-se ainda, para cada $n \in \mathbb{N}$ e cada $x \in [a, b]$,

$$\varphi_n(x) = n\big(F^*(x + 1/n) - F^*(x)\big) = n \int_x^{x+1/n} f^*(t)\, dt.$$

Resulta então de uma nova aplicação do teorema da derivação de Lebesgue que cada $\varphi_n'(x)$ existe q. s. e resulta de se estar a supor que $0 \leqslant f \leqslant M$ que $(\forall n \in \mathbb{N}) : 0 \leqslant \varphi_n \leqslant M$. Mas então

$$\begin{aligned} \int_a^b F'(x)\, dx &= \int_a^b \lim_{n \in \mathbb{N}} \varphi_n(x)\, dx \\ &= \lim_{n \in \mathbb{N}} \int_a^b \varphi_n(x)\, dx \text{ (pelo corolário 2.5)} \\ &= \lim_{n \in \mathbb{N}} \frac{\int_b^{b+1/n} F^*(x)\, dx}{1/n} - \lim_{n \in \mathbb{N}} \frac{\int_a^{a+1/n} F^*(x)\, dx}{1/n} \\ &= F(b) - F(a), \end{aligned}$$

pois F é contínua, pelo teorema 3.1. Está então provado que

$$\int_a^b F'(x)\, dx = F(b) - F(a) = \int_a^b f(x)\, dx.$$

Isto foi provado para cada $a, b \in \mathbb{R}$ tais que $a < b$; portanto, tem-se (3.13).

Passemos agora ao caso geral; por outras palavras, não se está a supor que f é majorada. Sejam $a, b \in \mathbb{R}$ tais que $a < b$. Para cada $n \in \mathbb{N}$, seja

$$\begin{aligned} g_n\colon \ [a,b] &\longrightarrow \mathbb{R} \\ x &\mapsto \begin{cases} f(x) & \text{se } f(x) \leqslant n \\ n & \text{caso contrário.} \end{cases} \end{aligned}$$

e seja

$$\begin{aligned} G_n\colon \ [a,b] &\longrightarrow \mathbb{R} \\ x &\mapsto \int_c^x g_n(t)\,dt. \end{aligned}$$

Pelo que já foi visto, tem-se $G_n'(x) = g_n(x)$ q. s. Para cada $x \in [a,b]$, tem-se

$$F(x) = G_n(x) + \int_c^x f(t) - g_n(t)\,dt \tag{3.14}$$

e, pelo teorema da derivação de Lebesgue, a função

$$\begin{aligned} [a,b] &\longrightarrow \mathbb{R} \\ x &\mapsto \int_c^x f(t) - g_n(t)\,dt \end{aligned}$$

é derivável em quase todos os pontos; além disso, nos pontos onde for derivável a derivada é maior ou igual a 0, visto que se trata de uma função crescente. Mas resulta então de (3.14) que

$$F'(x) \geqslant G_n'(x) = g_n(x) \text{ q. s.}$$

Como isto tem lugar para qualquer $n \in \mathbb{N}$, deduz-se que $F'(x) \geqslant f(x)$ q. s. e, portanto,

$$\int_a^b F'(x)\,dx \geqslant \int_a^b f(x)\,dx = F(b) - F(a).$$

Mas esta desigualdade é, de facto, uma igualdade, pois tem-se sempre a desigualdade oposta (que é a relação (3.10)), como já foi visto. Está então provado que

$$\int_a^b F'(x)\,dx = \int_a^b f(x)\,dx$$

sempre que $a, b \in \mathbb{R}$ e $a < b$; logo, $F'(x) = f(x)$ q. s. ∎

Como já foi mencionado, se $a, b \in \mathbb{R}$ com $a < b$ e se f for uma função de $[a, b]$ em $\mathbb{R}$ derivável em quase todos os pontos de I e tal que f' seja integrável, não é necessariamente verdade que se tenha

$$\int_a^b f'(t)\,dt = f(b) - f(a), \tag{3.15}$$

mesmo supondo que f é contínua. Haverá alguma condição mais forte do que a continuidade que garanta que se tem (3.15)? A resposta é afirmativa.

DEFINIÇÃO 3.2 Sejam $a, b \in \mathbb{R}$ tais que $a < b$ e seja f uma função de $[a, b]$ em $\mathbb{R}$. Diz-se que a função f é *absolutamente contínua* se para cada $\varepsilon > 0$ existir algum $\delta > 0$ tal que para qualquer família finita de intervalos abertos dois a dois disjuntos $]a_1, b_1[,]a_2, b_2[, \ldots,]a_n, b_n[$ de $[a, b]$, se

$$\sum_{k=1}^{n} (b_k - a_k) < \delta,$$

então

$$\sum_{k=1}^{n} |f(b_k) - f(a_k)| < \varepsilon.$$

É imediato que qualquer função absolutamente contínua é contínua; basta tomar $n = 1$. Em contrapartida, há funções, tal como a função de Cantor, que são contínuas mas não absolutamente contínuas.

TEOREMA 3.5 (TEOREMA FUNDAMENTAL DO CÁLCULO) *Sejam $a, b \in \mathbb{R}$ tais que $a < b$ e seja f uma função de $[a, b]$ em $\mathbb{R}$. São então condições equivalentes:*

1. *a função f é absolutamente contínua;*
2. *a função f é derivável q. s., f' é integrável e*

$$(\forall x \in [a, b]) : f(x) - f(a) = \int_a^x f'(t)\,dt.$$

Veja-se [1, §7.3] ou [14, cap. 7] para uma demonstração deste teorema.

Finalmente, convém observar que se $f\colon [a, b] \longrightarrow \mathbb{R}$ for derivável em *todos* os pontos de $[a, b]$ e se se supuser que f' é integrável, então é verdade que a relação (3.15) é válida, sem haver necessidade de qualquer hipótese adicional; veja-se [14, cap. 7] para a demonstração.

Espaços L^p

4.1 Funções convexas

DEFINIÇÃO 4.1 Se $a, b \in \overline{\mathbb{R}}$, $a < b$ e φ é uma função de $]a, b[$ em $\mathbb{R}$, diz-se que φ é *convexa* se

$$(\forall x, y \in]a, b[)(\forall t \in [0, 1]) : \varphi\big((1-t)x + ty\big) \leqslant (1-t)\varphi(x) + t\varphi(y).$$

Ao estudar-se uma função convexa $\varphi\colon]a, b[\longrightarrow \mathbb{R}$, empregar-se-á a seguinte convenção: se um elemento de $]a, b[$ for representado por uma letra minúscula (x, por exemplo), então o ponto do gráfico de φ cuja abcissa é aquele elemento (ou seja, o ponto $(x, \varphi(x))$ no caso do exemplo dado) será representado pela letra maiúscula correspondente (X, no caso do exemplo dado). Com esta convenção, o significado geométrico da convexidade á o seguinte: dados quaisquer dois pontos x e y do domínio de φ, com $x < y$, o gráfico da restrição de φ ao intervalo $[x, y]$ está abaixo do segmento de recta que une X a Y; veja-se a figura 4.1.

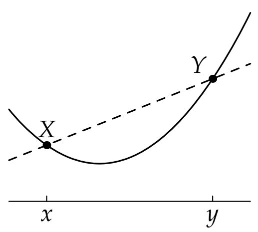

Figura 4.1: Função convexa

Também se pode observar na figura 4.1 outro comportamento típico das funções convexas, nomeadamente o facto de, dados quaisquer dois pontos x e y do domínio de φ, com $x < y$, não só o gráfico da restrição de φ ao intervalo $[x, y]$ estar *abaixo* do segmento de recta que une X a

Y como também o gráfico da restrição de φ a $]a,b[\setminus[x,y]$ estar *acima* da recta definida por aqueles dois pontos. Para ver porquê, seja $z \in]y,b[$; o caso em que $z \in]a,x[$ é análogo. Geometricamente, é imediato que se o ponto Y está *abaixo* da recta definida por X e por Z então o ponto Z está necessariamente *acima* da recta definida por X e por Y. Também é fácil demonstrar analiticamente aquela propriedade. De facto, se $t \in]0,1[$ for tal que $y = (1-t)x + tz$, i. e. se $t = (y-x)/(z-x)$, então

$$\begin{aligned}\varphi(y) &= \varphi\big((1-t)x + tz\big)\\ &\leqslant (1-t)\varphi(x) + t\varphi(z)\\ &= \frac{z-y}{z-x}\varphi(x) + \frac{y-x}{z-x}\varphi(z),\end{aligned}$$

pelo que

$$\varphi(z) \geqslant \frac{\varphi(y) - \frac{z-y}{z-x}\varphi(x)}{\frac{y-x}{z-x}} = \frac{z-x}{y-x}\varphi(y) - \frac{z-y}{y-x}\varphi(x).$$

EXEMPLO 4.1 A função

$$\begin{array}{ccc}\mathbb{R} & \longrightarrow & \mathbb{R}\\ x & \mapsto & |x|\end{array}$$

é convexa, pois se $x, y \in \mathbb{R}$ e se $t \in [0,1]$, tem-se

$$\big|(1-t)x + ty\big| \leqslant |(1-t)x| + |ty| = (1-t)|x| + t|y|.$$

PROPOSIÇÃO 4.1 *Se $a, b \in \overline{\mathbb{R}}$, $a < b$ e φ é uma função de $]a,b[$ em $\mathbb{R}$,*

1. *φ é convexa se e só se, sempre que s, t e u forem números reais tais que $a < s < t < u < b$, se tiver*

$$\frac{\varphi(t) - \varphi(s)}{t-s} \leqslant \frac{\varphi(u) - \varphi(t)}{u-t}; \tag{4.1}$$

2. *se φ for derivável, então é convexa se e só se φ' for crescente.*

DEMONSTRAÇÃO: Se a função φ for convexa e se s, t e u forem números reais tais que $a < s < t < u < b$, seja

$$\alpha = \frac{t-s}{u-s}\cdot$$

Então $t - s = \alpha(u - s)$, o que equivale a afirmar que $t = (1 - \alpha)s + \alpha u$. Resulta então da definição de função convexa que

$$\varphi(t) \leqslant (1-\alpha)\varphi(s) + \alpha\varphi(u) \iff (u-s)\varphi(t) \leqslant (u-t)\varphi(s) + (t-s)\varphi(u). \tag{4.2}$$

Recorrendo à igualdade $u - s = (u - t) + (t - s)$, deduz-se de (4.2) que $(u - t)(\varphi(t) - \varphi(s)) \leqslant (t - s)(\varphi(u) - \varphi(t))$, o que equivale a ter-se (4.1).

Reciprocamente, se a condição (4.1) se verifica sempre que $a < s < t < u < b$, se $x, y \in]a, b[$ e se $t \in [0, 1]$ então, supondo que $x < y$, como se tem $a < x < (1 - t)x + ty < y < b$, os cálculos do parágrafo anterior mostram que

$$\varphi\big((1-t)x + ty\big) \leqslant (1-t)\varphi(x) + t\varphi(y).$$

A demontração é análoga caso $x > y$ e, caso $x = y$, nada há a demonstrar.

Passemos agora à segunda alínea. Vai-se começar por provar que se φ for convexa e derivável, então φ' é crescente. Sejam então $x, y \in]a, b[$ tais que $x < y$; quer-se provar que $\varphi'(x) \leqslant \varphi'(y)$. Se $0 < h \leqslant y - x$, então

$$\frac{\varphi(x+h) - \varphi(x)}{h} = \frac{\varphi(x+h) - \varphi(x)}{(x+h) - x} \leqslant \frac{\varphi(y) - \varphi(x+h)}{y - (x+h)},$$

pelo que

$$\varphi'(x) = \lim_{h\to 0} \frac{\varphi(x+h) - \varphi(x)}{h} \leqslant \lim_{h\to 0} \frac{\varphi(y) - \varphi(x+h)}{y - (x+h)} = \frac{\varphi(y) - \varphi(x)}{y - x} \tag{4.3}$$

e, analogamente,

$$\frac{\varphi(y) - \varphi(x)}{y - x} \leqslant \varphi'(y), \tag{4.4}$$

pelo que $\varphi'(x) \leqslant \varphi'(y)$. Em termos geométricos, o que as desigualdades (4.3) e (4.4) significam é que o declive do segmento de recta que une X a $(y, \varphi(y))$ é maior ou igual ao declive da recta tangente ao gráfico de φ no ponto X e é menor ou igual ao declive da recta tangente ao gráfico de φ no ponto Y; veja-se a figura 4.2.

Finalmente, suponha-se que φ é derivável com derivada crescente; quer-se provar que φ é convexa. Sejam então $s, t, u \in]a, b[$ tais que $s < t < u$; quer-se provar que se tem (4.1). Vai-se começar por demonstrar isto no caso em que $\varphi(u) = \varphi(s)$, caso em que (4.1) equivale a afirmar que $\varphi(t) \leqslant \varphi(s) = \varphi(u)$. De facto, pelo teorema da média, há números $c \in]s, t[$ e $d \in]t, u[$ tais que

$$\varphi'(c) = \frac{\varphi(t) - \varphi(s)}{t - s} \quad \text{e que} \quad \varphi'(d) = \frac{\varphi(u) - \varphi(t)}{u - t},$$

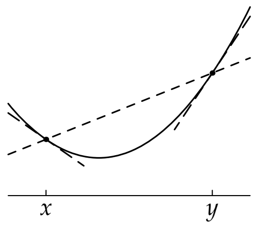

Figura 4.2: Significado geométrico das desigualdades (4.3) e (4.4)

pelo que, se se tivesse $\varphi(t) > \varphi(s) = \varphi(u)$, ter-se-ia $\varphi'(c) > 0 > \varphi'(d)$, o que é absurdo, pois está-se a supor que φ' é crescente.

Finalmente, vai-se demonstrar que se tem (4.1) no caso geral. Para cada $x \in]a,b[$, seja

$$\eta(x) = \varphi(x) - \frac{\varphi(u) - \varphi(s)}{u - s}(x - s).$$

Visto que η' e φ' têm diferença constante e visto que φ' é crescente, η' também é crescente. Por outro lado, $\eta(s) = \eta(u)(= \varphi(s))$. Mas então, pelo que foi visto no parágrafo anterior, tem-se que

$$\begin{aligned}\eta(t) < \varphi(s) &\iff \varphi(t) - \frac{\varphi(u) - \varphi(s)}{u - s}(t - s) < \varphi(s)\\ &\iff \frac{\varphi(t) - \varphi(s)}{t - s} \leqslant \frac{\varphi(u) - \varphi(s)}{u - s}.\end{aligned}$$ ■

Resulta da segunda alínea desta proposição que, por exemplo, a função exponencial é convexa, bem como, para cada $n \in \mathbb{N}$ a função

$$\begin{array}{ccc}\mathbb{R}_+^* & \longrightarrow & \mathbb{R}\\ x & \mapsto & x^n.\end{array}$$

PROPOSIÇÃO 4.2 *Se $a, b \in \overline{\mathbb{R}}$ são tais que $a < b$, então qualquer função convexa de $]a,b[$ em $\mathbb{R}$ é contínua.*

DEMONSTRAÇÃO: Se $x \in]a,b[$, vai-se mostrar que $\lim_{y \to x^+} \varphi(y) = \varphi(x)$; pode-se mostrar de maneira análoga que $\lim_{y \to x^-} \varphi(y) = \varphi(x)$. Sejam s e t pontos do intervalo $]a,b[$ tais que $s < x < t$ e seja $y \in]x,t[$. Como φ é convexa, Y está *acima* da recta definida por S e por X e *abaixo* da recta definida por X e por T, i. e. Y está na região angular marcada a tracejado na figura 4.3.

Há então funções afins $r_1, r_2\colon]x,t[\longrightarrow \mathbb{R}$ tais que $r_1(x) = r_2(x) = \varphi(x)$ e que $(\forall y \in]x,t[) : r_1(y) \leqslant \varphi(y) \leqslant r_2(y)$. Como $\lim_{y \to x^+} r_1(x) = \lim_{y \to x^+} r_2(x) = \varphi(x)$, resulta que $\lim_{y \to x^+} \varphi(x) = \varphi(x)$. ■

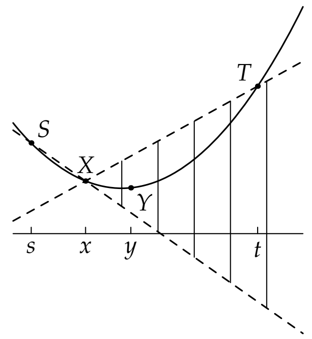

Figura 4.3: Justificação geométrica da continuidade de φ

4.2 Desigualdades de Jensen, Hölder e Minkovski

TEOREMA 4.1 (DESIGUALDADE DE JENSEN) *Seja* $(X, \mathcal{A}, m)$ *um espaço de medida tal que* $m(X) = 1$, *sejam* $a, b \in \overline{\mathbb{R}}$ *tais que* $a < b$, *seja* f *uma função mensurável de* X *em* $]a, b[$ *e seja* φ *uma função convexa de* $]a, b[$ *em* $\mathbb{R}$. *Então* $\int_X f\,dm \in]a, b[$, *a função* $\varphi \circ f$ *é mensurável e*

$$\varphi\left(\int_X f\,dm\right) \leqslant \int_X \varphi \circ f\,dm.$$

DEMONSTRAÇÃO: Que $\varphi \circ f$ é mensurável resulta imediatamente do teorema 2.1.

Sabe-se, pela primeira alínea da proposição 2.5 e por se estar a supor que $m(X) = 1$, que $a \leqslant \int_X f\,dm \leqslant b$. Se se tivesse $\int_X f\,dm = a$, então ter-se-ia $\int_X f - a\,dm = 0$, o que, pela sexta alínea da proposição 2.5, implicaria que $f(x) = a$ q. s., o que é impossível. O mesmo argumento mostra que $\int_X f\,dm < b$.

Seja $\iota = \int_X f\,dm$ e seja

$$s = \sup\left\{ \frac{\varphi(\iota) - \varphi(t)}{\iota - t} \;\middle|\; t \in]a, \iota[\right\}$$

Resulta imediatamente da definição de s que se $t \in]a, \iota[$, então

$$\frac{\varphi(\iota) - \varphi(t)}{\iota - t} \leqslant s \iff \varphi(\iota) - \varphi(t) \leqslant s(\iota - t) \iff \varphi(t) \geqslant \varphi(\iota) + s(t - \iota).$$

Por outro lado, sabe-se, pela primeira alínea da proposição 4.1 e pela definição de s que se $t \in]\iota, b[$, então

$$\frac{\varphi(t) - \varphi(\iota)}{t - \iota} \geqslant s \iff \varphi(t) - \varphi(\iota) \geqslant s(\iota - t) \iff \varphi(t) \geqslant \varphi(\iota) + s(t - \iota).$$

Então tem-se

$$(\forall t \in]a,b[) : \varphi(t) \geqslant \varphi(\iota) + s(t-\iota)$$

e, portanto,

$$(\forall x \in X) : \varphi(f(x)) \geqslant \varphi(\iota) + s(f(x)-\iota).$$

Resulta desta desigualdade que

$$\begin{aligned}\int_X \varphi \circ f \, dm &\geqslant \int_X \varphi(\iota) + s(f-\iota)\, dm \\ &= \varphi(\iota) + s\left(\int_X f\, dm - \iota\right) \\ &= \varphi\left(\int_X f\, dm\right).\end{aligned}$$

■

Observe-se que o enunciado anterior não afirma que $\varphi \circ f$ é integrável. No entanto, resulta da parte final da demonstração que se não for integrável, então $\int_X \varphi \circ f\, dm = +\infty$.

EXEMPLO 4.2 Seja $n \in \mathbb{N}$, seja $X = \{1,2,\ldots,n\}$ e considere-se a medida definida em $\mathcal{P}(X)$ tal que $(\forall k \in X) : m(\{k\}) = 1/n$. Se f for uma função de X em $\mathbb{R}$, seja, para cada $k \in X$, $x_k = f(k)$. Visto que a função exponencial é convexa, a desigualdade de Jensen afirma que:

$$e^{\int_X f\, dm} \leqslant \int_X \exp \circ f\, dm,$$

o que equivale a afirmar que

$$e^{\left(\sum_{k=1}^n x_k\right)/n} \leqslant \frac{1}{n}\sum_{k=1}^n e^{x_k},$$

ou seja, que

$$\sqrt[n]{e^{x_1}e^{x_2}\cdots e^{x_n}} \leqslant \frac{e^{x_1}+e^{x_2}+\cdots+e^{x_n}}{n}.$$

Designando, para cada $k \in X$, e^{x_k} por a_k, esta desigualdade escreve-se sob a forma

$$\sqrt[n]{a_1a_2\cdots a_n} \leqslant \frac{a_1+a_2+\cdots+a_n}{n},$$

que é a desigualdade entre a média aritmética e a média geométrica.

EXEMPLO 4.3 No exemplo anterior, não havia necessidade de todas as partes de X formadas por um único ponto terem a mesma medida. Repetindo o mesmo raciocínio, conclui-se que se $\alpha_1,\ldots,\alpha_n \in \mathbb{R}_+$ forem tais que $\alpha_1+\cdots+\alpha_n = 1$ e se $a_1,\ldots,a_n \in \mathbb{R}_+^*$, então

$$a_1^{\alpha_1}\cdots a_n^{\alpha_n} \leqslant \alpha_1a_1+\cdots+\alpha_na_n.$$

DEFINIÇÃO 4.2 Diz-se que dois elementos p e q de $[1,+\infty]$ são *expoentes conjugados* se ${}^1/_p + {}^1/_q = 1$.

Assim, por exemplo, $p = 1$ e $q = +\infty$ são expoentes conjugados, bem como $p = 2$ e $q = 2$. Um cálculo simples revela que para cada $p \in [1,+\infty]$ há um e um só $q \in [1,+\infty]$ tal que p e q sejam expoentes conjugados. Diz-se então que q é *o* expoente conjugado de p.

PROPOSIÇÃO 4.3 (DESIGUALDADE DE HÖLDER) *Sejam $(X, \mathcal{A}, m)$ um espaço de medida e $p, q \in]1,+\infty[$ dois expoentes conjugados. Se f e g são funções mensuráveis de X em $\overline{\mathbb{R}_+}$, então*

$$\int_X fg\,dm \leqslant \left(\int_X f^p\,dm\right)^{1/p}\left(\int_X g^q\,dm\right)^{1/q}.$$

DEMONSTRAÇÃO: Se $\int_X f^p\,dm = 0$ ou se $\int_X g^q\,dm = 0$, então, pela sexta alínea da proposição 2.5, $f(x) = 0$ q. s. ou $g(x) = 0$ q. s. e, portanto, $f(x)g(x) = 0$ q. s., pelo que a desigualdade de Hölder se reduz neste caso a $0 \leqslant 0$. Caso um dos integrais $\int_X f^p\,dm$ e $\int_X g^q\,dm$ seja maior do que zero e o outro seja igual a $+\infty$, a desigualdade de Hölder reduz-se a $\int_X fg\,dm \leqslant +\infty$.

Vai-se agora demonstrar a desigualdade de Hölder no caso em que $0 < \int_X f^p\,dm, \int_X g^q\,dm < +\infty$. Sejam

$$\varphi = \frac{f}{\left(\int_X f^p\,dm\right)^{1/p}} \quad \text{e} \quad \gamma = \frac{g}{\left(\int_X g^q\,dm\right)^{1/q}};$$

resulta destas definições que

$$\int_X \varphi^p\,dm = \int_X \gamma^q\,dm = 1. \tag{4.5}$$

Seja $x \in X$. Se $\varphi(x), \gamma(x) \in]0,+\infty[$, então há números reais y e z tais que $\varphi(x) = e^{y/p}$ e $\gamma(x) = e^{z/q}$. Visto que se está a supor que ${}^1/_p + {}^1/_q = 1$ e como a função exponencial é convexa, deduz-se que

$$e^{y/p}e^{z/q} = e^{y/p+z/q} \leqslant \frac{e^y}{p} + \frac{e^z}{q},$$

o que é o mesmo que afirmar que

$$\varphi(x)\gamma(x) \leqslant \frac{\varphi(x)^p}{p} + \frac{\gamma(x)^q}{q}. \tag{4.6}$$

e esta desigualdade também se verifica se algum dos valores $\varphi(x)$ ou $\gamma(x)$ for igual a 0 ou a $+\infty$. Integrando, deduz-se de (4.6), de (4.5) e de se ter $1/p + 1/q = 1$ que

$$\int_X \varphi\gamma \, dm \leqslant 1,$$

o que equivale à desigualdade de Hölder. ∎

PROPOSIÇÃO 4.4 (DESIGUALDADE DE MINKOWSKI) *Sejam $(X, \mathcal{A}, m)$ um espaço de medida e $p \in]1, +\infty[$. Se f e g são funções mensuráveis de X em $\overline{\mathbb{R}_+}$, então*

$$\left(\int_X (f+g)^p \, dm\right)^{1/p} \leqslant \left(\int_X f^p \, dm\right)^{1/p} + \left(\int_X g^p \, dm\right)^{1/p}.$$

DEMONSTRAÇÃO: Seja q o expoente conjugado de p. Resulta de desigualdade de Hölder que

$$\int_X f(f+g)^{p-1} \, dm \leqslant \left(\int_X f^p \, dm\right)^{1/p} \left(\int_X (f+g)^{(p-1)q} \, dm\right)^{1/q} \tag{4.7}$$

e que

$$\int_X g(f+g)^{p-1} \, dm \leqslant \left(\int_X g^p \, dm\right)^{1/p} \left(\int_X (f+g)^{(p-1)q} \, dm\right)^{1/q}. \tag{4.8}$$

Somando as desigualdades (4.7) e (4.8) e usando o facto de se ter $p + q = pq \iff (p-1)q = p$, vem que

$$\begin{aligned} &\int_X (f+g)^p \, dm \leqslant \\ &\qquad \leqslant \left(\left(\int_X f^p \, dm\right)^{1/p} + \left(\int_X g^p \, dm\right)^{1/p}\right)\left(\int_X (f+g)^p \, dm\right)^{1/q}. \end{aligned} \tag{4.9}$$

Para terminar a demonstração, vai-se supor que $\int_X (f+g)^p \, dm > 0$ e que ambos os integrais $\int_X f^p \, dm$ e $\int_X g^q \, dm$ são finitos; no caso contrário, a desigualdade de Minkowski é trivial. Visto que a função de $\mathbb{R}_+^*$ em $\mathbb{R}_+^*$ definida por $x \mapsto x^p$ é convexa, tem-se que

$$\left(\frac{f+g}{2}\right)^p \leqslant \frac{f^p}{2} + \frac{g^p}{2}$$

e, portanto, $0 < \int_X (f+g)^p \, dm < +\infty$. Então a desigualdade de Minkovski resulta de se dividirem ambos os membros da desigualdade (4.9) por

$$\left(\int_X (f+g)^p \, dm\right)^{1/q} = \left(\int_X (f+g)^p \, dm\right)^{1-1/p}. \quad \blacksquare$$

4.3 Espaços de funções integráveis

DEFINIÇÃO 4.3 Seja $(X, \mathcal{A}, m)$ um espaço de medida e seja f uma função mensurável de X em $\overline{\mathbb{R}}$. Diz-se que a função f é *essencialmente limitada* se existir algum $M \in \mathbb{R}_+$ tal que

$$m\left(\{\, x \in X \mid |f(x)| > M \,\}\right) = 0. \tag{4.10}$$

Se f for essencialmente limitada, então o ínfimo do conjunto dos números $M \in \mathbb{R}_+$ para os quais se tem (4.10) designa-se por *supremo essencial* da função f e representa-se por $\|f\|_\infty$.

PROPOSIÇÃO 4.5 *Se f for uma função essencialmente limitada de X em $\overline{\mathbb{R}}$, então o conjunto $\{\, x \in X \mid |f(x)| > \|f\|_\infty \,\}$ tem medida nula.*

DEMONSTRAÇÃO: Basta ver que, pela definição de $\|f\|_\infty$, se tem

$$(\forall n \in \mathbb{N}) : m\left(\left\{\, x \in X \,\middle|\, |f(x)| > \|f\|_\infty + \frac{1}{n} \,\right\}\right) = 0$$

e que

$$\{\, x \in X \mid |f(x)| > \|f\|_\infty \,\} = \bigcup_{n \in \mathbb{N}} \left\{\, x \in X \,\middle|\, |f(x)| > \|f\|_\infty + \frac{1}{n} \,\right\}. \quad \blacksquare$$

Se $p \in [1, +\infty[$ e se a função $|f|^p$ for integrável, então emprega-se a notação

$$\|f\|_p = \left(\int_X |f|^p \, dm\right)^{1/p}.$$

EXEMPLO 4.4 A função

$$\begin{array}{rccc} f\colon & [1, +\infty[& \longrightarrow & \overline{\mathbb{R}} \\ & x & \longmapsto & \begin{cases} 1/x & \text{se } x > 0 \\ +\infty & \text{se } x = 0 \end{cases} \end{array}$$

não é essencialmente limitada, pois se $M > 0$,

$$l(\{x \in X \mid |f(x)| > M\}) = l([0, 1/M[) = 1/M.$$

Por outro lado,

$$(\forall p \in]1, +\infty[) : \|f\|_p = \left(\frac{1}{p-1}\right)^{1/p} = \frac{1}{(p-1)^{1/p}}.$$

A notação $\|\cdot\|_p$ ($p \in [1, +\infty]$) sugere que se está a falar de normas. De facto, poder-se-ia pensar que, dado $p \in [1, +\infty]$, se se considerasse o conjunto E_p de todas as funções f para as quais o número $\|f\|_p$ está definido, então E_p seria um espaço vectorial e que a função $\|\cdot\|_p$ seria uma norma definida em E_p. De facto, em geral não é esse o caso, pois se $f \in E_p$ e se $\|\cdot\|_p$ fosse uma norma, então ter-se-ia $\|f\|_p = 0$ se e só $f \equiv 0$. Mas o que se tem efectivamente é que $\|f\|_p = 0$ se e só se $f(x) = 0$ q. s.; isto resulta da sexta alínea da proposição 2.5 se $p \in [1, +\infty[$ e da proposição 4.5 caso $p = +\infty$.

Vai-se ver como obter uma norma a partir de E_p e de $\|\cdot\|_p$. Para já, convém observar que estamos efectivamente a trabalhar com espaços vectoriais e que, por outro lado, o obstáculo mencionado no parágrafo anterior a que $\|\cdot\|_p$ seja uma norma é de facto o único obstáculo.

PROPOSIÇÃO 4.6 *Sejam f e g funções mensuráveis de X em $\mathbb{R}$ e seja $\lambda \in \mathbb{R}$.*

1. *Se $p \in [1, +\infty[$ e se as funções $|f|^p$ e $|g|^p$ forem integráveis, então as funções $|f+g|^p$ e $|\lambda f|^p$ também o são e tem-se*

$$\|f+g\|_p \leqslant \|f\|_p + \|g\|_p \quad e \quad \|\lambda f\|_p = |\lambda| \|f\|_p.$$

2. *Se as funções f e g forem essencialmente limitadas, então as funções $f+g$ e λf também o são e tem-se*

$$\|f+g\|_\infty \leqslant \|f\|_\infty + \|g\|_\infty \quad e \quad \|\lambda f\|_\infty = |\lambda| \|f\|_\infty.$$

DEMONSTRAÇÃO: Se as hipóteses da primeira alínea se verificarem, então resulta da desigualdade de Minkovski que $|f+g|^p$ é integrável e que $\|f+g\|_p \leqslant \|f\|_p + \|g\|_p$. É imediato que se tem $\|\lambda f\|_p = |\lambda| \|f\|_p$.

Se as hipóteses da segunda alínea se verificarem então, pela proposição 4.5, tem-se $|f(x)| \leqslant \|f\|_\infty$ e $|g(x)| \leqslant \|g\|_\infty$ q. s., pelo que se tem

$$|f(x) + g(x)| \leqslant |f(x)| + |g(x)| \leqslant \|f\|_\infty + \|g\|_\infty \text{ q. s.}$$

Isto não só mostra que $f+g$ é essencialmente limitada, como também mostra que $\|f+g\|_\infty \leqslant \|f\|_\infty + \|g\|_\infty$. Mais uma vez, é trivial que $\|\lambda f\|_\infty = |\lambda| \|f\|_\infty$. ∎

DEFINIÇÃO 4.4 Sejam V um espaço vectorial real e p uma função de V em $\mathbb{R}_+$. Diz-se que p é uma *semi-norma* se

1. $(\forall \lambda \in \mathbb{R})(\forall v \in V) : p(\lambda v) = |\lambda| p(v)$;
2. $(\forall v, w \in V) : p(v + w) \leqslant p(v) + p(w)$.

É claro que, se V e p forem como na definição anterior, então p é uma norma se e só se $(\forall v \in V) : p(v) = 0 \Longleftrightarrow v = 0$. De facto, até se poderia ter empregue o símbolo $\Longrightarrow$ em vez de $\Longleftrightarrow$, pois a outra implicação verifica-se automaticamente, uma vez que

$$p(0) = p(0.0) = |0| p(0) = 0.$$

A proposição 4.6 afirma então que, dado um espaço de medida X,

1. para cada $p \in [1, +\infty[$, o espaço das funções mensuráveis f de X em $\mathbb{R}$ tais que f^p é integrável forma um subespaço vectorial do espaço das funções de X em $\mathbb{R}$ e, além disso, $\| \cdot \|_p$ é uma semi-norma neste espaço;
2. o espaço das funções mensuráveis e essencialmente limitadas de X em $\mathbb{R}$ forma um subespaço vectorial do espaço das funções de X em $\mathbb{R}$ e, além disso, $\| \cdot \|_\infty$ é uma semi-norma neste espaço.

PROPOSIÇÃO 4.7 *Sejam V um espaço vectorial real, p uma semi-norma definida em V e $W = \{ v \in V \mid p(v) = 0 \}$. Então*

1. *W é um subespaço vectorial de V;*
2. *para cada $v \in V$, a restrição de p a $v + W$ é constante;*
3. *a função*
$$\begin{array}{rccc} \| \cdot \| \colon & V/W & \longrightarrow & \mathbb{R}_+ \\ & v + W & \longmapsto & p(v) \end{array}$$
é uma norma.

DEMONSTRAÇÃO: Sejam $w, w' \in W$ e $\lambda, \mu \in \mathbb{R}$. Então

$$0 \leqslant p(\lambda w + \mu w') \leqslant p(\lambda w) + p(\mu w') = |\lambda| p(w) + |\mu| p(w') = 0,$$

pelo que $\lambda w + \mu w' \in W$. Isto prova que W é um espaço vectorial.

Seja $v \in V$. Se $w \in W$, então

$$\begin{aligned} p(v) &= p\big((v+w)-w\big) \\ &\leqslant p(v+w)+p(-w) \\ &= p(v+w)+p((-1).w) \\ &= p(v+w)+|-1|p(w) \\ &= p(v+w) \\ &\leqslant p(v)+p(w) \\ &= p(v), \end{aligned}$$

o que mostra que $p(v+w)=p(v)$

Vejamos agora que $\|\cdot\|$ é uma norma. Se $v \in V$, então

$$\|v+W\|=0 \iff p(v)=0 \iff v \in W \iff v+W=0+W.$$

Se $\lambda \in \mathbb{R}$, então

$$\|\lambda v+W\|=p(\lambda v)=|\lambda|p(v)=|\lambda|\|v+W\|.$$

Finalmente, se $v' \in V$ então

$$\begin{aligned} \big\|(v+W)+(v'+W)\big\| &= \big\|(v+v')+W\big\| \\ &= p(v+v') \\ &\leqslant p(v)+p(v') \\ &= \|v+W\|+\|v'+W\|. \end{aligned}$$

■

Seja $p \in [1,+\infty]$ e aplique-se a proposição anterior ao conjunto de todas as funções mensuráveis $f\colon X \longrightarrow \mathbb{R}$ tais que $|f|^p$ seja integrável (caso $p<+\infty$) ou que f seja essencialmente limitada (caso $p=+\infty$) e à semi-norma $\|\cdot\|_p$. Representa-se por $L^p(X)$ o quociente do espaço vectorial em questão pelo subespaço das funções f tais que $\|f\|_p=0$, i. e. o espaço das funções que se anulam q. s. Resulta desta definição e das proposições 4.6 e 4.7 que $L^p(X)$ é um espaço vectorial real relativamente às operações

$$\begin{array}{ccc} L^p(X)\times L^p(X) & \longrightarrow & L^p(X) \\ ([f],[g]) & \mapsto & [f+g] \end{array} \quad \text{e} \quad \begin{array}{ccc} \mathbb{R}\times L^p(X) & \longrightarrow & L^p(X) \\ (\lambda,[f]) & \mapsto & [\lambda f], \end{array}$$

que faz sentido considerar a função

$$\begin{array}{ccc} L^p(X) & \longrightarrow & \mathbb{R} \\ [f] & \mapsto & \|f\|_p; \end{array}$$

e que esta função é uma norma definida em $L^p(X)$. Esta norma será novamente representada por $\|\cdot\|_p$.

No parágrafo anterior, foi empregue a notação $[f]$ para representar um elemento genérico de $L^p(X)$, o que está correcto do ponto de vista lógico, pois cada elemento de $L^p(X)$ é, por definição, a classe de equivalência de alguma função f. No entanto, é usual e é mais natural representar-se um elemento genérico de $L^p(X)$ por f e não por $[f]$ e é assim que se fará daqui para a frente. De facto, é usual pensar-se nos elementos de $L^p(X)$ como sendo funções. É apenas preciso ter-se em mente que, ao proceder-se assim, duas das funções com que estaremos a trabalhar serão consideradas idênticas sempre que diferirem apenas num conjunto de medida nula.

TEOREMA 4.2 *Sejam $p, q \in [1, +\infty]$ expoentes conjugados. Se $f \in L^p(X)$ e se $g \in L^q(X)$, então $fg \in L^1(X)$ e*

$$\|fg\|_1 \leqslant \|f\|_p \|g\|_q.$$

DEMONSTRAÇÃO: Se $p, q \in]1, +\infty[$, isto não é mais do que a desigualdade de Hölder. Caso $p = 1$ e $q = +\infty$, então tem-se, pela proposição 4.5,

$$|g(x)| \leqslant \|g\|_\infty \text{ q. s.},$$

pelo que

$$|f(x)g(x)| \leqslant |f(x)|\|g\|_\infty \text{ q. s.}$$

Logo, $fg \in L^1(X)$ (pois $f \in L^1(X)$) e

$$\|fg\|_1 \leqslant \int_X |f|\|g\|_\infty \, dm = \left(\int_X |f| \, dm\right)\|g\|_\infty = \|f\|_1 \|g\|_\infty.$$

A demonstração é análoga no caso em que $p = +\infty$ e $q = 1$. ∎

Visto que $\|\cdot\|_p$ é uma norma definida no espaço vectorial $L^p(X)$, existe uma distância natural neste espaço, definida por

$$\begin{array}{ccc} L^p(X) \times L^p(X) & \longrightarrow & \mathbb{R}_+ \\ (f, g) & \mapsto & \|f - g\|_p. \end{array}$$

Está-se então a lidar com um espaço métrico, o que leva naturalmente a algumas questões topológicas.

TEOREMA 4.3 *Cada espaço $L^p(X)$ é completo.*

DEMONSTRAÇÃO: Seja $(f_n)_{n\in\mathbb{N}}$ uma sucessão de Cauchy de elementos de $L^p(X)$; quer-se mostrar que converge para algum $f \in L^p(X)$. Afirmar que a sucessão $(f_n)_{n\in\mathbb{N}}$ é de Cauchy é afirmar que

$$(\forall\varepsilon \in \mathbb{R}_+^*)(\exists N \in \mathbb{N})(\forall m,n \in \mathbb{N}) : m,n \geqslant N \Longrightarrow \|f_m - f_n\|_p < \varepsilon.$$

Suponha-se que $p < +\infty$. Existe então alguma subsucessão $(f_{n_k})_{k\in\mathbb{N}}$ da sucessão $(f_n)_{n\in\mathbb{N}}$ tal que

$$(\forall k \in N) : \|f_{n_{k+1}} - f_{n_k}\|_p < \frac{1}{2^k}. \tag{4.11}$$

Definem-se então as funções φ_k ($k \in \mathbb{N}$) e φ de X em $\mathbb{R}$ por

$$\varphi_k = \sum_{i=1}^{k} |f_{n_{i+1}} - f_{n_i}| \quad \text{e por} \quad \varphi = \lim_{k\in\mathbb{N}} \varphi_k = \sum_{i=1}^{+\infty} |f_{n_{i+1}} - f_{n_i}|.$$

Como $\|\cdot\|_p$ é uma norma e se tem (4.11), então

$$(\forall k \in \mathbb{N}) : \|\varphi_k\|_p \leqslant 1 \Longleftrightarrow (\forall k \in \mathbb{N}) : \int_X |\varphi_k|^p \, dm \leqslant 1.$$

Aplicando o teorema de Fatou à sucessão $(|\varphi_k|^p)_{k\in\mathbb{N}}$, vem que

$$\int_X |\varphi|^p \, dm \leqslant \liminf_{k\in\mathbb{N}} \int_X |\varphi_k|^p \, dm \leqslant 1,$$

o que equivale a afirmar que $\|\varphi\|_p \leqslant 1$. Mas então tem-se $|\varphi(x)| < +\infty$ q. s. e, nos pontos $x \in X$ tais que $|\varphi(x)| < +\infty$, a série

$$\sum_{k=1}^{+\infty} \left(f_{n_{k+1}}(x) - f_{n_k}(x)\right) \tag{4.12}$$

é absolutamente convergente; define-se então $f\colon X \longrightarrow \mathbb{R}$ por

$$f(x) = f_{n_1}(x) + \sum_{k=1}^{+\infty} \left(f_{n_{k+1}}(x) - f_{n_k}(x)\right)$$

quando a série (4.12) for absolutamente convergente e por $f(x) = 0$ nos restantes $x \in X$. Então tem-se

$$f(x) = f_{n_1}(x) + \sum_{k=1}^{+\infty} \left(f_{n_{k+1}}(x) - f_{n_k}(x)\right) \text{ q. s.},$$

o que é o mesmo que afirmar que

$$f(x) = \lim_{k\in\mathbb{N}} f_{n_k} \text{ q. s.}$$

Vai-se provar que $f \in L^p(X)$ e que, em $L^p(X)$, $f = \lim_{n\in\mathbb{N}} f_n$. Seja então $\varepsilon \in \mathbb{R}_+^*$; quer-se provar que $\|f - f_n\|_p \leqslant \varepsilon$ se n for suficientemente grande. Para tal, tome-se $N \in \mathbb{N}$ tal que $\|f_m - f_n\|_p < \varepsilon$ quando $m, n \geqslant N$. Se $m \geqslant N$, decorre do teorema de Fatou que

$$\begin{aligned}
\int_X |f - f_m|^p \, dm &= \int_X \lim_{k\in\mathbb{N}} |f_{n_k} - f_m|^p \, dm \\
&\leqslant \liminf_{k\in\mathbb{N}} \int_X |f_{n_k} - f_m|^p \, dm \\
&\leqslant \varepsilon^p,
\end{aligned}$$

o que é o mesmo que afirmar que $\|f - f_m\|_p \leqslant \varepsilon$. Resulta também daqui que $f \in L^p(X)$, pois

$$n \geqslant N \Longrightarrow \left(\int_X |f|^p \, dm\right)^{1/p} \leqslant \|f - f_n\|_p + \|f_n\|_p \leqslant \varepsilon + \|f_n\|_p < +\infty.$$

Falta apenas demonstrar o teorema no caso em que $p = +\infty$. Seja $(f_n)_{n\in\mathbb{N}}$ uma sucessão de Cauchy de elementos de $L^\infty(X)$ e seja Y o conjunto de todos os pontos $x \in X$ para os quais se tem simultaneamente:

1. $(\forall n \in \mathbb{N}) : |f_n(x)| \leqslant \|f_n\|_\infty$;
2. $(\forall m, n \in \mathbb{N}) : |f_m(x) - f_n(x)| \leqslant \|f_m - f_n\|_\infty$.

Então $X \setminus Y$ tem medida nula, pois é a reunião de todos os conjuntos de um dos tipos

$$\{\, x \in X \mid |f_n(x)| > \|f_n\|_\infty \,\} \quad (n \in \mathbb{N})$$

ou

$$\{\, x \in X \mid |f_m(x) - f_n(x)| > \|f_m - f_n\|_\infty \,\} \quad (m, n \in \mathbb{N})$$

e o conjuntos de todos os conjuntos de algum destes tipos é um conjunto numerável de partes de X de medida nula.

Se $x \in Y$, então, como se tem $|f_m(x) - f_n(x)| \leqslant \|f_m - f_n\|_\infty$ para cada m e n naturais e como a sucessão $(f_n)_{n\in\mathbb{N}}$ é de Cauchy em $L^\infty(X)$, a sucessão $(f_n(x))_{n\in\mathbb{N}}$ é uma sucessão de Cauchy de números reais e, como tal, converge; seja $f(x)$ o seu limite. Prolonga-se f a X definindo $f(x) = 0$ se $x \in X \setminus Y$. Vai-se provar que $f \in L^\infty(X)$ e que, em $L^\infty(X)$,

$f = \lim_{n\in\mathbb{N}} f_n$. Para tal, tome-se $\varepsilon \in \mathbb{R}_+^*$ e tome-se $N \in \mathbb{N}$ tal que se $m, n \geqslant N$, então $\|f_m - f_n\|_\infty < \varepsilon$. Então, para cada $x \in Y$,

$$n \geqslant N \Longrightarrow |f(x) - f_n(x)| = \lim_{m\in\mathbb{N}} |f_m(x) - f_n(x)| \leqslant \varepsilon. \tag{4.13}$$

Então, para cada $x \in Y$ e cada $n \geqslant N$, $|f(x)| \leqslant |f(x) - f_n(x)| + |f_n(x)| \leqslant \varepsilon + \|f_n\|_\infty$, pelo que $f|_Y$ é limitada e, portanto, $f \in L^\infty(X)$. Por outro lado, resulta de (4.13) que

$$n \geqslant N \Longrightarrow \|f - f_n\|_\infty \leqslant \varepsilon.$$ ∎

Resulta da demonstração do teorema anterior que se tem.

TEOREMA 4.4 *Se $p \in [1, +\infty[$ e se $(f_n)_{n\in\mathbb{N}}$ é uma sucessão de Cauchy de elementos de $L^p(X)$, então existe alguma subsucessão $(f_{n_k})_{k\in\mathbb{N}}$ tal que a sucessão $(f_{n_k}(x))_{k\in\mathbb{N}}$ converge q. s.*

DEFINIÇÃO 4.5 Seja E um espaço métrico. Dada uma função f de E em $\mathbb{R}$, o *suporte* de f é a aderência do conjunto $\{\, x \in E \mid f(x) \neq 0 \,\}$.

Posto de outro modo, o suporte de uma função f é o menor fechado de E que contém todos os pontos de E onde f não se anula.

Observe-se que se $A \in \mathcal{M}(\mathbb{R})$ e se f for uma função contínua de A em $\mathbb{R}$ com suporte compacto, então f pertence a todos os espaços $L^p(A)$ ($p \in [1, +\infty]$). De facto, seja K o suporte de f. Então $f(A \setminus K) = \{0\}$ e $f|_K$ é uma função limitada, pois f é contínua e K é compacto; seja $M = \sup |f|$. Então

- $\sup |f| = M < +\infty$, pelo que $f \in L^\infty(A)$;
- se $p \in [1, +\infty[$, então

  $$\int_A |f|^p \, dm = \int_K |f|^p \, dm + \overbrace{\int_{A\setminus K} |f|^p \, dm}^{=0} \leqslant M^p m(K) < +\infty,$$

  pelo que $f \in L^p(A)$.

O conjunto de todas as funções contínuas com suporte compacto de E em $\mathbb{R}$ representa-se por $C_c(E)$.

TEOREMA 4.5 *Seja I um intervalo de $\mathbb{R}$ e seja $p \in [1, +\infty[$. O conjunto $C_c(I)$ é então uma parte densa de $L^p(I)$.*

DEMONSTRAÇÃO: Seja $f \in L^p(I)$; quer-se mostrar que há funções $g \in C_c(I)$ arbitrariamente próximas de f. Vai-se começar por ver que isto é verdade quando $f = \chi_X$ para alguma parte compacta X de I. Seja A um aberto de I tal que $X \subset A$ e que $\overline{A}$ seja novamente um compacto de I. Que um tal aberto existe, resulta da seguinte observação: para cada $x \in X$ considera-se algum intervalo aberto $]a_x, b_x[$ que contenha x e que $[a_x, b_x] \cap I$ é um intervalo fechado de $\mathbb{R}$; como X é compacto, existe uma parte finita $\{x_1, \ldots, x_n\}$ de X tal que o conjunto $A = \bigcup_{j=1}^{n}]a_{x_j}, b_{x_j}[\cap I$ contém X e

$$\overline{A} = \bigcup_{j=1}^{n} [a_{x_j}, b_{x_j}] \cap I \subset I.$$

Seja $\varepsilon \in \mathbb{R}_+^*$. Para cada $n \in \mathbb{N}$, seja A_n um aberto de $\mathbb{R}$ contido em A que contenha X e tal que $l\big(A_n \setminus X\big) < 2^{-n-1}\varepsilon^p$; para se ver que um tal aberto existe, basta ver que resulta da definição de medida de Lebesgue que existe algum aberto A_n^* de $\mathbb{R}$ que contém X e tal que $l\left(A_n^* \setminus X\right) < 2^{-n}\varepsilon^p$ e que se pode então tomar $A_n = A_n^* \cap A$. Seja $g_n \colon I \longrightarrow [0,1]$ uma função contínua que tome o valor 1 nos pontos de X e o valor 0 nos pontos de $I \setminus A_n$; pode-se, por exemplo, definir g_n por

$$\begin{array}{rccc} g_n\colon & I & \longrightarrow & [0,1] \\ & x & \mapsto & \dfrac{\inf_{y \in I \setminus A_n} |x-y|}{\inf_{y \in I \setminus A_n} |x-y| + \inf_{k \in X} |x-k|}. \end{array}$$

Seja $g = \sum_{n=1}^{\infty} 2^{-n} g_n$. A função g é contínua, por ser soma de uma série uniformemente convergente de funções contínuas. Além disso, cada g_n anula-se em $I \setminus A$ e, portanto, o suporte de g está contido em $\overline{A}$; em particular, g tem suporte compacto. Se $x \in I$, então $g(x)$ só poderá ser diferente de $\chi_X(x)$ se $x \in \bigcup_{n \in \mathbb{N}} A_n \setminus X$ e este conjunto tem medida inferior a ε^p. Por isto e porque as funções χ_X e g só tomam valores no intervalo $[0,1]$,

$$\left\|\chi_X - g\right\|_p = \left(\int_I |\chi_X - g|^p \, dl\right)^{1/p} \leqslant \varepsilon.$$

Continue-se a supor que $f = \chi_X$, mas agora vai-se supor apenas que $l(X) < +\infty$ (i. e. que f é integrável). Seja $\varepsilon \in \mathbb{R}_+^*$. Sabe-se, pela proposição 1.10, que X contém algum compacto K tal que $l(X \setminus K) < 2^{-p}\varepsilon^p$, o que implica que $\left\|\chi_X - \chi_K\right\|_p < \varepsilon/2$. Mas sabe-se, pela primeira parte da demonstração, que há alguma função $g \in C_c(I)$ tal que $\left\|\chi_K - g\right\|_p < \varepsilon/2$, pelo que $\left\|\chi_X - g\right\|_p < \varepsilon$.

Suponha-se agora que f é uma função simples integrável que só toma valores não negativos. Então $f = \sum_{k=1}^{n} a_k \chi_{X_k}$, onde cada X_k é uma parte mensurável de I com medida finita e $a_1, \ldots, a_n \in \mathbb{R}_+^*$. Seja $\varepsilon \in \mathbb{R}_+^*$ e, para cada $k \in \{1, 2, \ldots, n\}$, seja $g_k \in C_c(I)$ tal que $\left\|\chi_{X_k} - g_k\right\|_p < \varepsilon/(a_k n)$. Então $\left\|a_k\chi_{X_k} - a_k g_k\right\|_p < \varepsilon/n$ e, portanto, se se definir $g = \sum_{k=1}^{n} a_k g_k$,

$$\|s - g\|_p = \left\|\sum_{k=1}^{n} a_k\chi_{X_k} - \sum_{k=1}^{n} a_k g_k\right\|_p \leqslant \sum_{k=1}^{n} \left\|a_k\chi_{X_k} - a_k g_k\right\|_p < \varepsilon.$$

Seja agora $f \in L^p(I)$ uma função que só tome valores não negativos. Seja $(s_n)_{n\in\mathbb{N}}$ uma sucessão monótona crescente de funções simples mensuráveis de I em $\mathbb{R}_+$ que convirja pontualmente para a função f; uma tal sucessão existe, pela proposição 2.3. Como se tem $|f - s_n|^p \leqslant f^p$ para cada $n \in \mathbb{N}$, decorre do teorema da convergência dominada que $\lim_{n\in\mathbb{N}} \|f - s_n\|_p = 0$; em particular, se $\varepsilon \in \mathbb{R}_+^*$, então existe algum $n \in \mathbb{N}$ tal que $\|f - s_n\|_p < \varepsilon/2$. Mas já foi visto que existe alguma função $g \in C_c(I)$ tal que $\|s_n - g\|_p < \varepsilon/2$ e então $\|f - g\| < \varepsilon$.

Finalmente, se $f \in L^p(I)$ e $\varepsilon \in \mathbb{R}_+^p$, sejam $g_1, g_2 \in C_c(I)$ tais que $\|f^+ - g_1\|_p, \|f^- - g_2\|_p < \varepsilon/2$. Então, se $g = g_1 - g_2$,

$$\|f - g\|_p = \left\|(f^+ - g_1) - (f^- - g_2)\right\|_p < \varepsilon.$$ ∎

Em particular, se $a, b \in \mathbb{R}$ e $a < b$ então o espaço de todas as funções contínuas de $[a, b]$ em $\mathbb{R}$ é denso em $L^1([a, b])$ e pode-se deduzir daqui que, de certo modo, o integral de Lebesgue está para o de Riemann tal como os números racionais estão para os números racionais. De facto, o espaço métrico dos números reais contém os números racionais juntamente com exactamente aqueles pontos que é necessário acrescentar para que fique um espaço métrico completo, nem mais (como seria o caso dos números complexos) nem menos (como seria o caso do conjunto dos números algébricos reais). Podia-se ter definido o espaço $R^1([a, b])$ de maneira análoga ao espaço $L^1([a, b])$, empregando o integral de Riemann em vez do de Lebesgue, e definir naquele espaço uma norma análoga à norma $\|\cdot\|_1$ (a qual seria, de facto, a restrição da norma $\|\cdot\|_1$ àquele espaço, pelo teorema 2.6). No entanto, este espaço não é completo! De facto, se fosse completo, então seria um subespaço completo de $L^1([a, b])$; em particular, seria um fechado de $L^1([a, b])$. Como, por outro lado, $L^1([a, b])$ contém todas as funções contínuas de $[a, b]$ em $\mathbb{R}$, teria que conter a aderência deste conjunto de funções, o qual, como foi mencionado, é todo o espaço $L^1([a, b])$. Logo, ter-se-ia $R^1([a, b]) = L^1([a, b])$, mas isto não é verdade.

Convém observar que afirmar que $R^1([a,b]) \subsetneq L^1([a,b])$ *não é* o mesmo que afirmar que há funções integráveis segundo Lebesgue que não são integráveis segundo Riemann[1], pois $L^1([a,b])$ e $R^1([a,b])$ não são espaços de funções mas sim espaços de classes de equivalência de funções. O que se quer então é provar que há funções $f\colon [a,b] \longrightarrow \mathbb{R}$ integráveis segundo Lebesgue tais que se $g\colon [a,b] \longrightarrow \mathbb{R}$ é mensurável e $f(x) = g(x)$ q. s., então g não é integrável segundo Riemann. Para tal, tome-se, por exemplo, um conjunto mensurável $A \subset [a,b]$ tal que

1. A é fechado;
2. $\mathring{A} = \emptyset$;
3. $l(A) > 0$;

pelas proposições 1.12 e 1.13, os conjuntos de Cantor gordos estão nestas condições caso $[a,b] = [0,1]$ e, naturalmente, é possível definir de maneira análoga conjuntos com as mesmas propriedades em $[a,b]$ ou então considerar a imagem de um conjunto de Cantor gordo pela função

$$\begin{array}{ccc} [0,1] & \longrightarrow & [a,b] \\ t & \mapsto & a + t(b-a). \end{array}$$

Se $f\colon [a,b] \longrightarrow \mathbb{R}$ for uma função mensurável tal que $f(x) = \chi_A(x)$ q. s., vai-se provar que f não é integrável segundo Riemann. De facto, se $D = \{\, x \in [a,b] \mid f(x) \neq \chi_A(x) \,\}$ então, por hipótese, D tem medida nula e, em particular, $A \cap D$ tem medida nula. Mas então $l(A \setminus D) = l(A) > 0$ e, para cada $x \in A \setminus D$, tem-se que

- $f(x) = \chi_A(x) = 1$;
- qualquer vizinhança aberta V de x em $[a,b]$ tem pontos de $[a,b] \setminus A$ e, sendo o conjunto $\big([a,b] \setminus A\big) \cap V$ um aberto de $[a,b]$, tem medida não nula e, portanto, não está contido em D, pelo que há algum $y \in V$ tal que $f(y) = 0$.

Logo, o conjunto dos pontos de descontinuidade de f contém $A \setminus D$ que tem medida não nula e, portanto, f não é integrável segundo Riemann. Isto mostra que $\chi_A \in L^1([a,b]) \setminus R^1([a,b])$.

[1] Para tal, veja-se o exemplo 2.7 na página 42.

Espaços vectoriais normados

5.1 Complementos de Álgebra Linear

5.1.1 Famílias livres, famílias geradoras e bases

Vai-se trabalhar com espaços vectoriais reais e é conveniente rever alguns conceitos de Álgebra Linear. Sempre que não se disser qual é o corpo sobre o qual se está a trabalhar, isso quererá dizer que o que se está a fazer é válido para qualquer corpo.

DEFINIÇÃO 5.1 Seja V um espaço vectorial e seja B uma família de elementos de V.

1. Diz-se que a família B é *livre* se, para cada $n \in \mathbb{N}$, para quaisquer n elementos $v_1, \ldots, v_n$ de B e quaisquer números reais $\lambda_1, \ldots, \lambda_n$, se tiver
$$\lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_n v_n = 0 \Longrightarrow \lambda_1 = \lambda_2 = \cdots = \lambda_n = 0.$$
2. Diz-se que a família B *gera* o espaço V se o único subespaço vectorial W de V que contém B é o próprio V.
3. Diz-se que a família B é uma *base* de V se for uma família livre que gera V.

Verifica-se facilmente que, dado um subconjunto B de um espaço vectorial V, afirmar que B é livre (respectivamente gera V) equivale a afirmar que, para cada vector de V, há, no máximo (resp. mínimo) uma maneira de o escrever como combinação linear de elementos de B. Consequentemente, B é uma base de V se e só cada vector de V puder ser escrito como combinação linear de elementos de B de uma e uma só maneira.

Vai-se provar que qualquer parte livre de um espaço vectorial está contida em alguma base. Visto que qualquer espaço vectorial contém alguma parte livre (nomeadamente $\emptyset$), isto prova que qualquer espaço vectorial tem uma base. A demonstração será baseada no lema de Zorn, que é enunciado e demonstrado no apêndice B.

TEOREMA 5.1 *Qualquer parte livre de um espaço vectorial V está contida numa base de V. Além disso, quaisquer duas bases têm o mesmo cardinal.*

DEMONSTRAÇÃO: Seja V um espaço vectorial, seja L uma parte livre de V, seja

$$\mathcal{L} = \{\, B \subset V \mid B \text{ é livre e } B \supset L \,\}$$

e considere-se em $\mathcal{L}$ a relação de ordem induzida pela inclusão, i. e. a relação de ordem $\leqslant$ tal que $A \leqslant B$ se e só se $A \subset B$. Seja $(B_i)_{i\in I}$ uma família de elementos de $\mathcal{L}$ totalmente ordenada e seja $B = \bigcup_{i\in I} B_i$. Vai-se provar que $B \in \mathcal{L}$, de onde resultará imediatamente que B é um majorante de $(B_i)_{i\in I}$ em $\mathcal{L}$. É claro que $B \supset L$. Além disso, B é livre, pois, se $v_1, \ldots, v_n \in B$ e se, para cada $k \in \{1, \ldots, n\}$, $i(k) \in I$ for tal que $v_k \in B_{i(k)}$, então, como a família $(B_i)_{i\in I}$ é totalmente ordenada, existe algum $j \in \{1, \ldots, n\}$ tal que $B_{i(j)} \supset B_{i(k)}$, para cada $k \in \{1, \ldots, n\}$. Mas então $B_{i(j)} \supset \{v_1, \ldots, v_n\}$ e, portanto, se $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$, $\sum_{k=1}^{n} \lambda_k v_k = 0 \Longrightarrow \lambda_1 = \cdots = \lambda_n = 0$, pois $B_{i(j)}$ é livre.

Uma vez que qualquer parte de $\mathcal{L}$ totalmente ordenada por inclusão tem um majorante, resulta do lema de Zorn que $\mathcal{L}$ tem algum elemento maximal M, o qual é livre (por definição de $\mathcal{L}$). Vejamos que gera V. Se não fosse esse o caso, haveria algum $v \in V$ que não seria combinação linear de elementos de M. Mas então $M \cup \{v\}$ seria livre, i. e. $M \in \mathcal{L}$, o que é absurdo, pois M é maximal e $M \subsetneq M \cup \{v\}$.

Vejamos agora que quaisquer duas bases têm o mesmo cardinal. Comece-se por supor que V tem alguma base finita $\{v_1, \ldots, v_n\}$ ($n \in \mathbb{N}$). Vai-se provar que se $\{w_1, \ldots, w_m\}$ ($m \in \mathbb{N}$) for uma família livre de elementos de V, então $m \leqslant n$, de onde resulta que qualquer base de V tem, no máximo, n elementos, de onde se pode deduzir, por simetria, que qualquer base tem exactamente n elementos. O vector w_1 pode ser escito sob a forma

$$w_1 = \lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_n v_n, \tag{5.1}$$

e, como $w_1 \neq 0$, algum λ_i é diferente de 0. Suponha-se que se trata de λ_1. Então v_1 pertence ao espaço vectorial gerado por

$$\{w_1, v_2, v_3, \ldots, v_n\}, \tag{5.2}$$

o qual é necessariamente igual a V. Além disso, a família (5.2) é necessariamente livre, pois se $\eta_1, \eta_2, \ldots, \eta_n \in \mathbb{R}$ forem tais que

$$\eta_1 w_1 + \eta_2 v_2 + \cdots + \eta_n v_n = 0, \tag{5.3}$$

então, caso $\eta_1 \neq 0$, tem-se simultaneamente (5.1) e

$$w_1 = -\frac{\eta_2}{\eta_1}v_2 - \frac{\eta_3}{\eta_1}v_3 + \cdots - \frac{\eta_n}{\eta_1}v_n,$$

o que é absurdo, pois $\{v_1, \ldots, v_n\}$ é livre. Logo, tem-se $\eta_1 = 0$ e, portanto, os restantes η_i ($i \in \{2, \ldots, n\}$) são nulos, por (5.3) e porque $\{v_1, \ldots, v_n\}$ é livre.

Pode-se agora recomeçar o processo e ir mostrando sucessivamente que os conjuntos $\{w_1, w_2, v_3, v_4, \ldots, v_n\}$, $\{w_1, w_2, w_3, v_4, \ldots, v_n\}$, etc. são bases de V. Mas então, caso se tivesse $m > n$ chegava-se à conclusão que $\{w_1, \ldots, w_n\}$ seria uma base de V e, portanto, que w_m seria uma combinação linear dos elementos daquela base, o que é absurdo, pois está-se a supor que o conjunto $\{w_1, w_2, \ldots, w_m\}$ é livre.

Suponha-se agora que $\{\, v_i \mid i \in I \,\}$ e $\{\, w_j \mid j \in J \,\}$ são bases de V com um número infinito de elementos cada uma; quer-se provar que I e J têm o mesmo cardinal. Para cada $j \in J$, w_j pode ser escrito sob a forma $\sum_{i \in I} \lambda_{ij} v_i$, com $\lambda_{ij} \neq 0$, sendo $I_j = \{\, i \in I \mid \lambda_{ij} \neq 0 \,\}$ um conjunto finito e não vazio. Vejamos que $I = \bigcup_{j \in J} I_j$. Por um lado, é claro que $I \supset \bigcup_{j \in J} I_j$. Por outro lado, qualquer $v \in V$ pode ser escrito como combinação linear de vectores da forma w_j ($j \in \mathbb{N}$) e, portanto, como combinação linear de vectores v_i ($i \in \bigcup_{j \in J} I_j$). Logo, V é gerado por $\{\, v_i \mid i \in \bigcup_{j \in J} I_j \,\}$ e, como $\{\, v_i \mid i \in I \,\}$ é uma base de V, $I = \bigcup_{j \in J} I_j$.

Visto que $I = \bigcup_{j \in J} I_j$ e como cada I_j é finito, o cardinal de I não excede o de $J \times \mathbb{N}$, que é igual ao cardinal de J. Pelo mesmo argumento, o cardinal de J não excede o de I e, portanto, I e J têm o mesmo cardinal.■

Convém observar que o uso do lema de Zorn na demonstração anterior é dispensável caso se esteja a supor que o espaço vectorial V é gerado por um conjunto finito F de vectores de V. De facto, se M for uma parte de F que gera V e com o menor número possível de elementos, prova-se facilmente que M é uma base de V. Seja $n = \#M$. Então, como foi visto no decorrer da demonstração, qualquer parte livre de V tem, no máximo, n elementos e, em particular, qualquer elemento de $\mathcal{L}$ tem, no máximo, n elementos, pelo que $\mathcal{L}$ tem algum elemento maximal.

Visto que qualquer espaço vectorial tem alguma base e que todas as bases têm o mesmo cardinal, faz sentido fazer a seguinte

DEFINIÇÃO 5.2 Designa-se por *dimensão* de um espaço vectorial o cardinal das suas bases.

Relembre-se que, dados dois subespaços vectoriais W_1 e W_2 de um espaço vectorial V, se diz que são suplementares se $W_1 \cap W_2 = \{0\}$ e se

$W_1 + W_2 = V$, o que equivale a afirmar que a aplicação linear

$$\begin{array}{ccc} W_1 \times W_2 & \longrightarrow & V \\ (v, w) & \mapsto & v + w \end{array}$$

é bijectiva. Quando W_1 e W_2 são suplementares diz-se que V é a soma directa interna de W_1 e W_2 e escreve-se $V = W_1 \bigoplus W_2$.

PROPOSIÇÃO 5.1 *Qualquer subespaço vectorial de um espaço vectorial tem algum suplementar.*

DEMONSTRAÇÃO: Sejam V um espaço vectorial e W_1 um subespaço vectorial de V; quer-se mostrar que existe algum subespaço vectorial W_2 de V tal que $V = W_1 \bigoplus W_2$. Seja B uma base de W_1. Então B é uma parte livre de V e, portanto, pelo teorema 5.1, existe alguma base B' de V que contém B. Então o subespaço vectorial de V gerado por $B' \setminus B$ está nas condições pretendidas. ■

5.1.2 Hiperplanos

DEFINIÇÃO 5.3 Sejam V um espaço vectorial e W um subespaço vectorial de V. Diz-se que W é um *hiperplano* se $\dim(V/W) = 1$.

Alternativamente, poder-se-ia definir um hiperplano como sendo um subespaço próprio maximal. De facto, é claro que se W for um hiperplano de V, W tem que ser um subespaço próprio (pois, caso contrário, $\dim(V/W) = 0$) e caso não fosse maximal, i. e. caso houvesse algum subespaço vectorial W' de V tal que $W \subsetneq W' \subsetneq V$, então a aplicação natural

$$\begin{array}{ccc} V/W & \longrightarrow & V/W' \\ v + W & \mapsto & v + W' \end{array}$$

seria sobrejectiva mas não injectiva e então

$$\dim(V/W') \geqslant 1 \Longrightarrow \dim(V/W) \geqslant 2.$$

Reciprocamente, se W for um subespaço de um espaço vectorial V que não seja um hiperplano, então $\dim(V/W) = 0$ ou $\dim(V/W) \geqslant 2$. No primeiro caso, $W = V$, pelo que W não é um subespaço próprio de V. No segundo caso, seja U um subespaço de V/W de dimensão 1 e seja π a projecção natural de V sobre V/W. Então $\pi^{-1}(U)$ é um subespaço vectorial de V tal que $W \subsetneq \pi^{-1}(U) \subsetneq V$, pelo que W não é um subespaço próprio maximal.

EXEMPLO 5.1 Seja V o espaço vectorial real das sucessões convergentes de números reais e seja H o subespaço vectorial formado pelas sucessões de números reais convergentes para 0. Comece-se por observar que o conjunto das sucessões convergentes de números reais forma um espaço vectorial (relativamente às operações usuais de soma de duas sucessões e de produto de uma sucessão por um número real) porque, dadas duas sucessões convergentes de números reais $(a_n)_{n\in\mathbb{N}}$ e $(b_n)_{n\in\mathbb{N}}$ e dado $\lambda \in \mathbb{R}$, se ter

$$\lim_{n\in\mathbb{N}}(a_n + b_n) = \lim_{n\in\mathbb{N}} a_n + \lim_{n\in\mathbb{N}} b_n \quad \text{e} \quad \lim_{n\in\mathbb{N}} \lambda a_n = \lambda \lim_{n\in\mathbb{N}} a_n. \tag{5.4}$$

Resulta também deste argumento que H é efectivamente um subespaço vectorial de V. Vejamos que se trata de um hiperplano. Seja H' um subespaço vectorial de V que contenha estrictamente H. Então existe alguma sucessão convergente de números reais $(a_n)_{n\in\mathbb{N}}$ que não converge para 0 e que pertence a H'; seja l o seu limite. Dada qualquer sucessão convergente de números reais $(b_n)_{n\in\mathbb{N}}$, se o seu limite for l', então a sucessão

$$\left(b_n - \frac{l'}{l}a_n\right)_{n\in\mathbb{N}}$$

converge para 0, ou seja, pertence a H e, portanto, a H'. Mas então $(b_n)_{n\in\mathbb{N}} \in H'$, pois H' é um espaço vectorial e

$$(b_n)_{n\in\mathbb{N}} = \frac{l'}{l}\overbrace{(a_n)_{n\in\mathbb{N}}}^{\in H'} + \overbrace{(b_n)_{n\in\mathbb{N}} - \frac{l'}{l}(a_n)_{n\in\mathbb{N}}}^{\in H'}.$$

Isto prova que $H' = V$ e, consequentemente, que H é um subespaço próprio maximal.

DEFINIÇÃO 5.4 Seja V um espaço vectorial sobre um corpo k. O *dual algébrico* de V é o conjunto V^* das formas lineares de V em k.

O dual algébrico[1] de um espaço vectorial tem uma estrutura natural de espaço vectorial, assim definida: se $f, g \in V^*$, se $\alpha, \beta \in \mathbb{R}$ e se $v \in V$, então $(\alpha f + \beta g)(v) = \alpha f(v) + \beta g(v)$.

Não é óbvio *a priori* que o dual algébrico de um espaço vectorial $V \neq \{0\}$ não seja reduzido à função nula. De facto, isso nunca acontece.

[1]Num contexto puramente algébrico, o dual algébrico designa-se unicamente por «dual», mas iremos trabalhar mais à frente com outro tipo de dual.

PROPOSIÇÃO 5.2 *Se V é um espaço vectorial e v e w são dois elementos distintos de V, então existe algum $f \in V^*$ tal que $f(v) \neq f(w)$.*

DEMONSTRAÇÃO: Visto que V^* é um conjunto de formas lineares, afirmar, para algum $f \in V^*$, que $f(v) \neq f(w)$ é o mesmo que afirmar que $f(v - w) \neq 0$. Como $\{v - w\}$ é um conjunto livre, é prolongável a uma base de V. Logo, existe uma e uma só forma linear $f \colon V \longrightarrow k$ tal que $f(v - w) = 1$ e que se anula nos restantes elementos da base. Em particular, $f(v - w) \neq 0$. ■

Se X é um conjunto e $\mathcal{F}$ é um conjunto de funções de X num conjunto Y, diz-se que o conjunto $\mathcal{F}$ separa os pontos de X se, dados dois pontos distintos x e y de X, existir alguma função $f \in \mathcal{F}$ tal que $f(x) \neq f(y)$. Com esta terminologia, a proposição anterior afirma que o dual algébrico de um espaço vectorial V separa os pontos de V.

PROPOSIÇÃO 5.3 *Seja V um espaço vectorial sobre um corpo k. Os hiperplanos V são os núcleos dos elementos não nulos de V^*. Além disso, se $f, g \in V^* \setminus \{0\}$, são condições equivalentes:*

1. $\ker f = \ker g$;
2. $f = \lambda g$, *para algum* $\lambda \in k \setminus \{0\}$.

DEMONSTRAÇÃO: Seja H um hiperplano de V, seja π a projecção natural de V sobre V/H e seja φ um isomorfismo linear V/H em k, o qual existe necessariamente, visto que $\dim(V/H) = 1$. Então, se se definir $f \in V^*$ por $f = \varphi \circ \pi$, tem-se, para cada $v \in V$,

$$\begin{aligned} v \in \ker f &\iff f(v) = 0 \\ &\iff \varphi(\pi(v)) = 0 \\ &\iff \pi(v) = 0 \text{ (pois } \varphi \text{ é injectiva)} \\ &\iff v \in H. \end{aligned}$$

Está então provado que $H = \ker f$. Como $H \neq V$, $f \not\equiv 0$.

Reciprocamente, se $f \in V^* \setminus \{0\}$ e se $H = \ker f$, quer-se provar que H é um hiperplano. A aplicação linear f induz um isomorfismo entre $V/\ker(f)$ e a imagem de f. Mas $V/\ker(f) = V/H$ e a imagem de f é k; logo, está provado que o espaço vectorial V/H é isomorfo a k, o que equivale a afirmar que $\dim(V/H) = 1$, ou seja, que H é um hiperplano.

Sejam agora $g \in V^* \setminus \{0\}$ tal que $\ker g = H = \ker f$; quer-se provar que $f = \lambda g$ para algum $\lambda \in \mathbb{R}$. Visto que $g \neq 0$, existe algum vector v

tal que $g(v) = 1$. Seja $\lambda = f(v)$. A forma linear $f - \lambda g$ anula-se em v e, obviamente, anula-se em H, pois tanto f como g se anulam em H. Mas então $\ker(f - \lambda g) \supsetneq H$, o que implica, visto que H é um hiperplano, que $\ker(f - \lambda g) = V$ e isto equivale a afirmar que $f - \lambda g \equiv 0$. ■

EXEMPLO 5.2 Reexaminemos o exemplo 5.1. Com as notações empregues nesse exemplo, H é um hiperplano de V pois é o núcleo da forma linear

$$\begin{array}{ccc} V & \longrightarrow & \mathbb{R} \\ (a_n)_{n\in\mathbb{N}} & \mapsto & \lim\limits_{n\in\mathbb{N}} a_n. \end{array}$$

Que se trata realmente de uma forma linear resulta de uma nova aplicação das relações (5.4).

5.2 Normas: exemplos e propriedades elementares

DEFINIÇÃO 5.5 Um *espaço vectorial normado* é um par ordenado $(V, \|\cdot\|)$, onde V é um espaço vectorial real e $\|\cdot\|$ é uma norma de V em $\mathbb{R}_+$.

EXEMPLO 5.3 Seja $n \in \mathbb{N}$. A norma usual em $\mathbb{R}^n$ é a norma $\|\cdot\|_2$, que se define por

$$\|(x_1, \ldots, x_n)\|_2 = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}.$$

Mais geralmente, se $p \in [1, +\infty[$, pode-se definir em $\mathbb{R}^n$ a norma $\|\cdot\|_p$, que se define por

$$\|(x_1, \ldots, x_n)\|_p = \left(|x_1|^p + |x_2|^p + \cdots + |x_n|^p\right)^{1/p}$$

e a norma $\|\cdot\|_\infty$, que se define por

$$\|(x_1, \ldots, x_n)\|_\infty = \max\{|x_1|, |x_2|, \ldots, |x_n|\}.$$

Se $X = \{1, 2, \ldots, n\}$, então o conjunto $\mathcal{F}(X)$ das funções de X em $\mathbb{R}$ identifica-se naturalmente a $\mathbb{R}^n$ através da bijecção

$$\begin{array}{ccc} \mathcal{F}(X) & \longrightarrow & \mathbb{R}^n \\ f & \mapsto & \big(f(1), f(2), \ldots, f(n)\big). \end{array}$$

Seja $p \in [1, +\infty]$. Se se considerar em X a medida de contagem, então $(L^p(X), \|\cdot\|_p)$ identifica-se naturalmente a $(\mathbb{R}^n, \|\cdot\|_p)$.

EXEMPLO 5.4 Seja $C([0,1])$ o espaço das funções contínuas de $[0,1]$ em $\mathbb{R}$. Duas normas com que se trabalha frequentemente neste espaço são a *norma do supremo*

$$\|f\|_\infty = \sup|f|$$

e a *norma do integral*

$$\|f\|_1 = \int_0^1 |f|.$$

Observe-se que, dada uma função $f \in C([0,1])$ e dada uma sucessão $(f_n)_{n\in\mathbb{N}}$ de elementos de $C([0,1])$, afirmar que $(f_n)_{n\in\mathbb{N}}$ converge para f relativamente à norma do supremo é o mesmo que afirmar que a sucessão $(f_n)_{n\in\mathbb{N}}$ converge uniformemente para a função f.

EXEMPLO 5.5 Seja X um conjunto e seja $p \in [1,+\infty]$. Se se considerar em X a medida de contagem m, então $(X, \mathcal{P}(X), m)$ é um espaço de medida e faz então sentido considerar o espaço vectorial normado $(L^p(X), \|\cdot\|_p)$. Este espaço é geralmente representado por $l^p(X)$. Em particular, $l^p(\mathbb{N})$ identifica-se naturalmente, caso $p < +\infty$, ao conjunto das séries $\sum_{n=1}^{+\infty} a_n$ de números reais tais que a série $\sum_{n=1}^{+\infty} |a_n|^p$ converge, munido da norma

$$\left\| \sum_{n=1}^{+\infty} a_n \right\|_p = \left(\sum_{n=1}^{+\infty} |a_n|^p \right)^{1/p}.$$

Por outro lado o espaço $l^\infty(X)$ não é mais do que o conjunto das funções limitadas de X em $\mathbb{R}$ e, dada uma tal função f, $\|f\|_\infty = \sup|f|$.

Seja $(V, \|\cdot\|)$ um espaço vectorial normado. Pelo mesmo motivo que foi apresentado na página 93 no caso dos espaços $L^p(X)$, é possível definir uma métrica natural em V induzida pela norma $\|\cdot\|$.

Se (E_1, d_1) e (E_2, d_2) são espaços métricos, é natural considerar no produto cartesiano $E_1 \times E_2$ a métrica produto d, que se define por

$$d\big((x_1,x_2),(y_1,y_2)\big) = \max\big\{d_1(x_1,y_1), d_2(x_2,y_2)\big\}.$$

PROPOSIÇÃO 5.4 *Seja $(V, \|\cdot\|)$ um espaço vectorial normado e considerem-se nos produtos cartesianos $V \times V$ e $\mathbb{R} \times V$ as métricas produto. Então as seguintes funções são contínuas:*

1. $$\begin{array}{ccc} V & \longrightarrow & \mathbb{R} \\ v & \longmapsto & \|v\|; \end{array}$$
2. $$\begin{array}{ccc} V \times V & \longrightarrow & V \\ (v,w) & \longmapsto & v+w; \end{array}$$

3. $$\begin{array}{ccc} \mathbb{R}\times V & \longrightarrow & V \\ (\lambda, v) & \mapsto & \lambda v. \end{array}$$

As duas primeiras são mesmo uniformemente contínuas.

DEMONSTRAÇÃO: Se $v, w \in V$, então

$$\|v\| = \|(v-w)+w\| \leqslant \|v-w\| + \|w\|,$$

ou seja

$$\|v\| - \|w\| \leqslant \|v-w\|. \tag{5.5}$$

Trocando v com w, conclui-se que

$$\|w\| - \|v\| \leqslant \|v-w\|. \tag{5.6}$$

Mas afirmar que se tem simultaneamente (5.5) e (5.6) é o mesmo que afirmar que $\big|\|v\| - \|w\|\big| \leqslant \|v-w\|$. Então, dado $\varepsilon > 0$, se $\delta = \varepsilon$, tem-se, para cada $v, w \in V$,

$$\|v-w\| < \delta \iff \|v-w\| < \varepsilon \Longrightarrow \big|\|v\| - \|w\|\big| < \varepsilon,$$

pelo que $\|\cdot\|$ é uniformemente contínua.

Passemos agora à adição de vectores. Se $v_1, v_2, w_1, w_2 \in V$, então

$$\begin{aligned} \big\|(v_1+v_2)-(w_1+w_2)\big\| &= \big\|(v_1-w_1)+(v_2-w_2)\big\| \\ &\leqslant \|v_1-w_1\| + \|v_2-w_2\|. \end{aligned}$$

Logo, dado $\varepsilon > 0$, se $\delta = \varepsilon/2$, tem-se, para cada $(v_1, v_2), (w_1, w_2) \in V \times V$

$$\|v_1-w_1\|, \|v_2-w_2\| < \delta \Longrightarrow \big\|(v_1+v_2)-(w_1+w_2)\big\| < \varepsilon,$$

pelo que a adição de vectores é uniformemente contínua.

Sejam agora $\lambda_0 \in \mathbb{R}$ e $v_0 \in V$; quer-se provar que a multiplicação de um escalar por um vector é contínua no ponto (λ_0, v_0). Se $\lambda \in \mathbb{R}$ e $v \in V$, então

$$\lambda v - \lambda_0 v_0 = (\lambda-\lambda_0)(v-v_0) + (\lambda-\lambda_0)v_0 + \lambda_0(v-v_0),$$

pelo que

$$\|\lambda v - \lambda_0 v_0\| \leqslant |\lambda-\lambda_0|\|v-v_0\| + |\lambda-\lambda_0|\|v_0\| + |\lambda_0|\|v-v_0\|.$$

Então, se $\delta \leqslant 1$ e se $v \in V$ e $\lambda \in \mathbb{R}$ forem tais que $\|v - v_0\| < \delta$ e que $|\lambda - \lambda_0| < \delta$, tem-se

$$\begin{aligned}\|\lambda v - \lambda_0 v_0\| &< \delta^2 + \delta\|v_0\| + |\lambda_0|\delta \\ &\leqslant \delta + \delta\|v_0\| + |\lambda_0|\delta \\ &= \delta(1 + \|v_0\| + |\lambda_0|).\end{aligned}$$

Logo, dado $\varepsilon > 0$, se se tomar $\delta = \min\{1, \varepsilon/(1+\|v_0\|+|\lambda_0|)\}$, tem-se que

$$\|v - v_0\|, |\lambda - \lambda_0| < \delta \Longrightarrow \|\lambda v - \lambda_0 v_0\| < \varepsilon.$$

ou seja, a função é contínua no ponto (λ_0, v_0). ■

A terceira função do enunciado anterior só é uniformemente contínua quando $V = \{0\}$.

5.3 Aplicações lineares contínuas

TEOREMA 5.2 *Sejam* $(V, \|\cdot\|)$ *e* $(W, \|\cdot\|)$ *espaços vectoriais normados e seja* f *uma aplicação linear de* V *em* W*. São então condições equivalentes:*

1. *a função* f *é uniformemente contínua;*
2. *a função* f *é contínua;*
3. *a função* f *é contínua em* 0;
4. *o conjunto*
$$\{\ \|f(v)\| \mid \|v\| \leqslant 1\ \} \tag{5.7}$$
é majorado;
5. *existe algum* $K \geqslant 0$ *tal que* $(\forall v \in V) : \|f(v)\| \leqslant K\|v\|$*; por outras palavras, o conjunto*
$$\{\ K \geqslant 0 \mid (\forall v \in V) : \|f(v)\| \leqslant K\|v\|\ \} \tag{5.8}$$
não é vazio.

Além disso, caso se verifiquem então o conjunto (5.8) *é o conjunto dos majorantes do conjunto* (5.7)*; em particular, o supremo do conjunto* (5.7) *coincide com o ínfimo do conjunto* (5.8).

DEMONSTRAÇÃO: É imediato que a primeira condição implica a segunda e que a segunda implica a terceira.

Se a terceira condição se verificar, então existe algum $r > 0$ tal que $\|v\| \leqslant r \Longrightarrow \|f(v)\| < 1$. Então, se $\|v\| \leqslant 1$,

$$\|f(v)\| = \frac{1}{r}\|rf(v)\| = \frac{1}{r}\|f(rv)\| < \frac{1}{r},$$

pois $\|rv\| < r$.

Suponha-se agora que a quarta condição se verifica e seja K um majorante do conjunto (5.7). Então, para cada $v \in V$, ou $v = 0$, caso em que $\|f(v)\| = 0 \leqslant K\|v\| = 0$, ou então $v \neq 0$, pelo que $\big\|v/\|v\|\big\| = 1$ e então

$$\left\|f\left(\frac{v}{\|v\|}\right)\right\| \leqslant K \Longleftrightarrow \|f(v)\| \leqslant K\|v\|.$$

Finalmente, suponha-se que a quinta condição se verifica, e seja $K \geqslant 0$ tal que $(\forall v \in V) : \|f(v)\| \leqslant K\|v\|$. Se $K = 0$, então $f \equiv 0$ e é então claro que é uniformemente contínua. Caso contrário, se $\varepsilon > 0$, toma-se $\delta = \varepsilon/K$. Então, se $v, w \in V$,

$$\begin{aligned}\|v - w\| < \delta \Longleftrightarrow \|v - w\| &< \frac{\varepsilon}{K}\\ \Longrightarrow \|f(v - w)\| &\leqslant K\|v - w\| < \varepsilon\\ \Longleftrightarrow \|f(v) - f(w)\| &< \varepsilon.\end{aligned}$$

Por outro lado, tem-se, para cada $v \in V$ tal que $\|v\| \leqslant 1$, que

$$\|f(v)\| \leqslant K\|v\| \leqslant K.$$

pelo que qualquer elemento do conjunto (5.8) é maior ou igual a qualquer número da forma $\|f(v)\|$ com $\|v\| \leqslant 1$, i. e. é majorante do conjunto (5.7). Como, ao demonstrar-se que a quarta condição implica a quinta, se provou que qualquer majorante do conjunto (5.7) pertence a (5.8), isto conclui a demonstração. ∎

EXEMPLO 5.6 Seja $c_{00}(\mathbb{N})$ o espaço das sucessão $(a_n)_{n\in\mathbb{N}}$ de números reais tais que $a_n = 0$ se n for suficientemente grande,[2] considere-se neste espaço a norma $\|\cdot\|_\infty$ e a forma linear

$$\begin{aligned} f\colon \quad c_{00}(\mathbb{N}) &\longrightarrow \mathbb{R}\\ (a_n)_{n\in\mathbb{N}} &\longmapsto \sum_{n=1}^{+\infty} a_n.\end{aligned}$$

[2] A notação $c_{00}(\mathbb{N})$ pode parecer estranha, mas provém do facto de se empregar a notação $c_0(\mathbb{N})$ para o espaço das sucessão de números reais convergentes para 0.

Esta forma linear não é contínua, pois se $n \in \mathbb{N}$ e se $(s_n)_{n\in\mathbb{N}}$ é a sucessão de elementos de $c_{00}(\mathbb{N})$ tal que, para cada $n \in \mathbb{N}$

$$s_n = (\overbrace{1/n, 1/n, \ldots, 1/n}^{n \text{ vezes}}, 0, 0, \ldots)$$

então $(\forall n \in \mathbb{N}) : \|s_n\|_\infty = 1/n$, pelo que $\lim_{n\in\mathbb{N}} s_n = 0$. Mas, por outro lado, $(\forall n \in \mathbb{N}) : f(s_n) = 1$. Logo, como $f(0) = 0$, a função f não é contínua no ponto 0.

DEFINIÇÃO 5.6 Se $(V, \|\cdot\|)$ e $(W, \|\cdot\|)$ forem espaços vectoriais normados e f for uma aplicação linear contínua de V em W, define-se a *norma* de f e representa-se por $\|f\|$ o número

$$\|f\| = \sup\{\ \|f(v)\| \mid \|v\| \leqslant 1\}.$$

Observe-se que, pelo teorema 5.2, nas condições da definição anterior tem-se que

$$\|f\| = \inf\{K \geqslant 0 \mid (\forall v \in V) : \|f(v)\| \leqslant K\|v\|\}. \tag{5.9}$$

Este ínfimo é mesmo um mínimo, pois, ainda pelo teorema 5.2, o conjunto (5.8) é o conjunto dos majorantes de outro conjunto e, portanto, contém o seu ínfimo. Está assim demonstrado o

COROLÁRIO 5.1 *Se V e W forem espaços vectoriais normados e f for uma aplicação linear contínua de V em W, então*

$$(\forall v \in V) : \|f(v)\| \leqslant \|f\|\|v\|.$$

Não se deve pensar que, analogamente ao que foi visto no corolário anterior, o supremo que surge na definição de norma de uma aplicação linear contínua é sempre um máximo.

EXEMPLO 5.7 Seja $c_0(\mathbb{N})$ o espaço das sucessões convergentes para 0 de números reais, munido da norma $\|\cdot\|_\infty$ e considere-se a aplicação linear

$$\begin{array}{rccc} f\colon & c_0(\mathbb{N}) & \longrightarrow & \mathbb{R} \\ & (a_n)_{n\in\mathbb{N}} & \mapsto & \displaystyle\sum_{n=1}^{+\infty} 2^{-n} a_n. \end{array}$$

Para já, observe-se que esta definição faz sentido, i. e. que se $(a_n)_{n\in\mathbb{N}}$ é uma sucessão de números reais convergente para 0, a série $\sum_{n=1}^{+\infty} 2^{-n} a_n$

converge; basta ver que uma tal sucessão é necessáriamente limitada e que se $L \geqslant 0$ for um majorante de $\left(|a_n|\right)_{n\in\mathbb{N}}$, então

$$\sum_{n=1}^{+\infty}\left|2^{-n}a_n\right| \leqslant \sum_{n=1}^{+\infty} 2^{-n}L = L,$$

pelo que a série $\sum_{n=1}^{+\infty} 2^{-n}a_n$ é mesmo absolutamente convergente. Mas os mesmo cálculos mostram que se tem

$$\left(\forall (a_n)_{n\in\mathbb{N}} \in c_{00}(\mathbb{N})\right) : \left|f\left((a_n)_{n\in\mathbb{N}}\right)\right| \leqslant \left\|(a_n)_{n\in\mathbb{N}}\right\|_\infty,$$

pelo que $\|f\| \leqslant 1$. De facto, $\|f\| = 1$, pois se $n \in \mathbb{N}$ e se $(a_n)_{n\in\mathbb{N}}$ é a sucessão

$$(\overbrace{1, 1, \ldots, 1}^{n \text{ vezes}}, 0, 0, \ldots)$$

então $\left\|(a_n)_{n\in\mathbb{N}}\right\|_\infty = 1$ e $\left|f\left((a_n)_{n\in\mathbb{N}}\right)\right| = 1 - 2^{-n}$, pelo que

$$\|f\| \geqslant \sup\left\{\, 1 - 2^{-n} \mid n \in \mathbb{N} \,\right\} = 1.$$

Mas nunca se tem $\left|f\left((a_n)_{n\in\mathbb{N}}\right)\right| = 1$ para uma sucessão $(a_n)_{n\in\mathbb{N}}$ tal que $\left\|(a_n)_{n\in\mathbb{N}}\right\|_\infty = 1$ pois, como se trata de uma sucessão convergente para 0, existe algum $N \in \mathbb{N}$ tal que $|a_n| \leqslant 1/2$ quando $n > N$, pelo que

$$\left|f\left((a_n)_{n\in\mathbb{N}}\right)\right| \leqslant \sum_{n=1}^{N} 2^{-n} + \frac{1}{2}\sum_{n=N+1}^{+\infty} 2^{-n} < 1 = \sum_{n=1}^{+\infty} 2^{-n}.$$

Se V e W forem espaços vectoriais normados vai-se representar por $\mathcal{L}(V, W)$ o espaço das aplicações lineares contínuas de V em W. Prova-se facilmente que é um subespaço vectorial do espaço de todas as aplicações lineares de V em W.

PROPOSIÇÃO 5.5 *Se V e W forem espaços vectoriais normados, então $\|\cdot\|$ é uma norma no espaço vectorial $\mathcal{L}(V, W)$.*

DEMONSTRAÇÃO: É imediato que $\|0\| = 0$ e, por outro lado, se $f \in \mathcal{L}(V, W)$ é tal que $\|f\| = 0$, então, pelo corolário 5.1,

$$(\forall v \in V) : \|f(v)\| \leqslant \|f\|\|v\| = 0,$$

pelo que $f \equiv 0$.

Se $f \in \mathcal{L}(V, W)$ e se $\lambda \in \mathbb{R}$, quer-se provar que $\|\lambda f\| = |\lambda| \|f\|$. Basta observar que

$$\begin{aligned}
\|\lambda f\| &= \sup \{ \|\lambda f(v)\| \mid \|v\| = 1 \} \\
&= \sup \{ |\lambda| \|f(v)\| \mid \|v\| = 1 \} \\
&= |\lambda| \sup \{ \|f(v)\| \mid \|v\| = 1 \} \\
&= |\lambda| \|f\|.
\end{aligned}$$

Finalmente, é preciso provar que se $f, g \in \mathcal{L}(V, W)$, então $\|f + g\| \leqslant \|f\| + \|g\|$. Se $v \in V$, então

$$\begin{aligned}
\|(f+g)(v)\| &= \|f(v) + g(v)\| \\
&\leqslant \|f(v)\| + \|g(v)\| \\
&\leqslant \|f\|\|v\| + \|g\|\|v\| \\
&= (\|f\| + \|g\|)\|v\|,
\end{aligned}$$

o que mostra, pela relação (5.9), que $\|f + g\| \leqslant \|f\| + \|g\|$. ■

PROPOSIÇÃO 5.6 *Sejam V_1, V_2 e V_3 espaços vectoriais normados, seja $f \in \mathcal{L}(V_2, V_3)$ e seja $g \in \mathcal{L}(V_1, V_2)$. Então $\|f \circ g\| \leqslant \|f\|\|g\|$.*

DEMONSTRAÇÃO: Seja $v \in V_1$. Então

$$\big\|(f \circ g)(v)\big\| = \big\|f(g(v))\big\| \leqslant \|f\|\big\|g(v)\big\| \leqslant \|f\|\|g\|\|v\|,$$

pelo que $\|f \circ g\| \leqslant \|f\|\|g\|$, pela relação (5.9). ■

Segundo a proposição 5.3, os hiperplanos de um espaço vectorial V são os núcleos das formas lineares não nulas. Se V for um espaço vectorial normado e se se considerarem somente as formas lineares contínuas, que hiperplanos se obtêm?

DEFINIÇÃO 5.7 Seja V um espaço vectorial normado. O *dual topológico* de V é o conjunto V' das formas lineares contínuas de V em $\mathbb{R}$.

PROPOSIÇÃO 5.7 *Se V for um espaço vectorial normado, então os núcleos dos elementos de $V' \setminus \{0\}$ são os hiperplanos fechados de V.*

DEMONSTRAÇÃO: Se $f \in V'$, então $\ker f$ é um fechado de V pois é igual a $f^{-1}(\{0\})$.

Reciprocamente, seja W um hiperplano fechado de V e seja $f \in V^*$ tal que $H = \ker f$; quer-se provar que f é contínua. Seja $v \in V$ tal que

$f(v) = 1$. Então $f^{-1}(\{1\}) = v + W$, que é um fechado de V. Existe então alguma bola aberta centrada em 0 que não intersecta $v + W$; seja r o raio de uma bola nessas condições. Vejamos que

$$(\forall w \in V) : \|w\| \leqslant 1 \Longrightarrow \big|f(w)\big| \leqslant \frac{1}{r}. \tag{5.10}$$

Se assim não fosse, i. e. se existisse algum vector $w \in V$ tal que $\|w\| \leqslant 1$ e que $\big|f(w)\big| > 1/r$, então

$$\left\|\frac{w}{f(w)}\right\| < r \quad \text{e} \quad f\left(\frac{w}{f(w)}\right) = 1,$$

o que é impossível, pela escolha de r. Está então provado que se verifica a relação (5.10) e resulta então da quarta alínea do teorema 5.2 que f é contínua. ■

A proposição anterior sugere que um subespaço vectorial de um espaço vectorial V não é necessariamente um fechado de V, pois resulta do enunciado que uma maneira de obter um tal subespaço consiste em tomar o núcleo de uma forma linear descontínua.

EXEMPLO 5.8 Foi vista, no exemplo 5.6, uma forma linear descontínua definida em $\big(c_{00}(\mathbb{N}), \|\cdot\|_\infty\big)$. O seu núcleo é

$$\left\{ (a_n)_{n\in\mathbb{N}} \in c_{00}(\mathbb{N}) \;\middle|\; \sum_{n=1}^{+\infty} a_n = 0 \right\},$$

que então é um subespaço vectorial V de $c_{00}(\mathbb{N})$ que não é fechado. Também se pode ver directamente que V não é um fechado de $c_{00}(\mathbb{N})$ dando um exemplo de alguma sucessão $(s_n)_{n\in\mathbb{N}}$ de elementos de V que seja convergente para algum elemento $s \in c_{00}(\mathbb{N})$ que não pertença a V. Por exemplo, seja $(s_n)_{n\in\mathbb{N}}$ a sucessão de elementos de V tal que, para cada $n, m \in \mathbb{N}$,

$$s_n(m) = \begin{cases} 1 & \text{se } m = 1 \\ -1/n & \text{se } 2 \leqslant m \leqslant n+1 \\ 0 & \text{nos restantes casos;} \end{cases}$$

posto de outro modo, tem-se

$$\begin{aligned} s_1 &= (1, -1, 0, 0, 0, 0, \ldots), \\ s_2 &= (1, -1/2, -1/2, 0, 0, 0, \ldots), \\ s_3 &= (1, -1/3, -1/3, -1/3, 0, 0, \ldots) \end{aligned}$$

e assim sucessivamente. Então $\lim_{n\in\mathbb{N}} s_n = (1, 0, 0, 0, \ldots) \notin V$.

EXEMPLO 5.9 Seja $D([0,1])$ o espaço das funções deriváveis de $[0,1]$ em $\mathbb{R}$, o qual é claramente um subespaço vectorial de $C([0,1])$. Considere-se neste último espaço a norma do supremo. É visto nos cursos de Análise Real que há exemplos de sucessões $(f_n)_{n\in\mathbb{N}}$ de funções deriváveis que convergem uniformemente para funções f não deriváveis; por exemplo, pode-se tomar

$$\begin{array}{rccc} f_n\colon & [0,1] & \longrightarrow & \mathbb{R} \\ & x & \mapsto & \sqrt{\left(x-\dfrac{1}{2}\right)^2+\dfrac{1}{n}}, \end{array}$$

e

$$\begin{array}{rccc} f\colon & [0,1] & \longrightarrow & \mathbb{R} \\ & x & \mapsto & |x-1/2|. \end{array}$$

Mas afirmar que existe uma sucessão de elementos de $D([0,1])$ que converge uniformemente para uma função $f \in C([0,1]) \setminus D([0,1])$ é o mesmo que afirmar que $D([0,1])$ não é um subespaço fechado de $C([0,1])$.

Será visto mais à frente que qualquer subespaço vectorial de dimensão finita de um espaço vectorial normado V é um fechado de V.

Se V for um espaço vectorial real e p e q forem normas definidas em V, então as métricas induzidas por p e por q são distintas se e só se $p \neq q$. E no que se refere às topologias, i. e. aos abertos obtidos a partir de cada métrica? *A priori*, nada impede que normas distintas induzam as mesmas topologias. Para estudar este assunto, vai-se introduzir a seguinte terminologia: se V é um espaço vectorial real e p é uma norma definida em V, dir-se-á que um conjunto $A \subset V$ é p-aberto se for aberto para a métrica induzida por p. Comecemos por estudar o seguinte problema: quando é que qualquer p-aberto é q-aberto? Observe-se que afirmar que qualquer p-aberto é q-aberto equivale a afirmar que a função $\mathrm{id}\colon (V,q) \longrightarrow (V,p)$ é contínua, pois, mais geralmente, afirmar que uma função f de um espaço métrico E num espaço métrico F é contínua equivale a afirmar que se A for um aberto de F, então $f^{-1}(A)$ é um aberto de E. Mas, pelo teorema 5.1, $\mathrm{id}\colon (V,q) \longrightarrow (V,p)$ é contínua se e só se existir algum $k > 0$ tal que $(\forall v \in V) : p(v) \leqslant kq(v)$ ou, posto de uma maneira mais simples, se e só se existir algum $k > 0$ tal que $p \leqslant kq$. Está então demonstrada a

PROPOSIÇÃO 5.8 *Seja V um espaço vectorial real e sejam p e q normas definidas em V. São então condições equivalentes:*

1. $(\exists k \in \mathbb{R}_+^*) : p \leqslant kq$ *(i. e. $(\forall v \in V) : p(v) \leqslant kq(v)$);*

2. *qualquer p-aberto é q-aberto.*

EXEMPLO 5.10 Se $f \in C([0,1])$, então $\int_0^1 |f| \leqslant \sup|f|$, i. e. $\|f\|_1 \leqslant \|f\|_\infty$. Então, pela proposição anterior, qualquer parte de $C([0,1])$ que seja aberta relativamente à norma do integral também é aberta relativamente à norma do supremo. O recíproco não é verdadeiro. Isto pode ser visto de vários modos.

1. Sejam s e i a norma do supremo e a norma do integral respectivamente. A bola aberta unitária relativamente à norma do supremo,

$$B_s(0,1) = \{\, f \in C([0,1]) \mid |f| < 1 \}$$

   não é um aberto relativamente à norma do integral. De facto, para qualquer $\varepsilon > 0$ é possível encontrar $f \in B_i(0,\varepsilon) \setminus B_s(0,1)$; por exemplo, toma-se $r < \min\{2\varepsilon, 1\}$ e define-se

$$\begin{array}{rccc} f\colon & [0,1] & \longrightarrow & \mathbb{R} \\ & x & \longmapsto & \begin{cases} 1 - x/r & \text{se } x < r \\ 0 & \text{caso contrário.} \end{cases} \end{array}$$

   Veja-se o gráfico de f na figura 5.1; a distância de f à função nula (relativamente à norma do integral) é a área a sombreado. Então $\sup|f| = 1$ e $\int_0^1 |f| = r/2 < \varepsilon$. Isto mostra que, seja qual for $\varepsilon > 0$, $B_s(0,1) \not\supset B_i(0,\varepsilon)$; logo, $B_s(0,1)$ não é um aberto para a norma do integral.

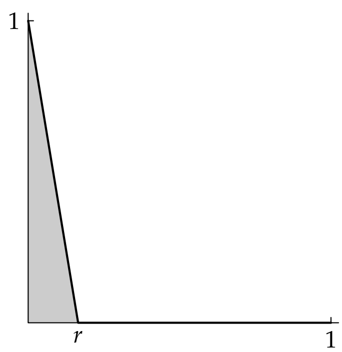

Figura 5.1: Gráfico de f

2. Para cada $n \in \mathbb{N}$, seja $(f_n)_{n\in\mathbb{N}}$ a sucessão de elementos de $C([0,1])$ tal que, para cada $n \in \mathbb{N}$ e cada $x \in [0,1]$, $f_n(x) = x^n$. Então, para

cada $n \in \mathbb{N}$, $\|f_n\|_1 = (n+1)^{-1}$ e $\|f_n\|_\infty = 1$. Logo, não existe nenhum $k > 0$ tal que $\|\cdot\|_\infty \leqslant k\|\cdot\|_1$, pois se existisse tinha-se

$$(\forall n \in \mathbb{N}) : 1 = \|f_n\|_\infty \leqslant k\|f_n\|_1 = \frac{k}{n+1},$$

o que é impossível.

Se V for um espaço vectorial real e p e q forem normas definidas em V, quando é que p e q induzem em V as mesmas topologias? Pela proposição anterior, isto acontece quando e só quando houver números $k_1, k_2 > 0$ tais que $p \leqslant k_1 q$ e que $q \leqslant k_2 p$.

DEFINIÇÃO 5.8 Se V for um espaço vectorial real, diz-se que duas normas p e q definidas em V são *equivalentes* se houver números $k_1, k_2 > 0$ tais que $p \leqslant k_1 q$ e que $q \leqslant k_2 p$.

Com esta terminologia, resulta então da proposição 5.8 e do que foi observado antes da definição anterior que é válido o

COROLÁRIO 5.2 *Seja V um espaço vectorial real e sejam p e q normas definidas em V. Então as normas p e q induzem as mesmas topologias se e só se forem equivalentes.*

EXEMPLO 5.11 Pelo que foi visto no exemplo 5.10, no espaço $C([0,1])$ a norma do supremo e a norma do integral não são equivalentes.

5.4 Espaços vectoriais normados de dimensão finita

No âmbito de exemplos de normas equivalentes, poder-se-ia ter mostrado também que, em $\mathbb{R}^n$ ($n \in \mathbb{N}$), as normas $\|\cdot\|_p$ ($p \in [1, +\infty]$) são todas equivalentes entre si. Mas, de facto, pode-se provar mais do que isso.

TEOREMA 5.3 *Num espaço vectorial real de dimensão finita, todas as normas são equivalentes.*

DEMONSTRAÇÃO: Para demonstrar este teorema, basta demonstrar que em $\mathbb{R}^n$ ($n \in \mathbb{N}$) qualquer norma $p \colon V \longrightarrow \mathbb{R}_+$ é equivalente à norma $\|\cdot\|_2$.

Seja $\{e_1, e_2, \ldots, e_n\}$ a base canónica de $\mathbb{R}^n$. Então, se $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$,

$$\begin{aligned} p\big(\lambda_1 e_1 + \cdots + \lambda_n e_n\big) &\leqslant |\lambda_1| p(e_1) + \cdots + |\lambda_n| p(e_n) \\ &\leqslant \sum_{k=1}^{n} p(e_k) \max\{|\lambda_1|, \ldots, |\lambda_n|\} \\ &= \sum_{k=1}^{n} p(e_k) \sqrt{\max\{|\lambda_1|^2, \ldots, |\lambda_n|^2\}} \\ &\leqslant \sum_{k=1}^{n} p(e_k) \sqrt{\sum_{k=1}^{n} |\lambda_k|^2} \\ &= \sum_{k=1}^{n} p(e_k) \big\|(\lambda_1 e_1 + \cdots + \lambda_n e_n)\big\|_2. \end{aligned}$$

Logo, se $k_1 = \sum_{k=1}^{n} p(e_k)$, está provado que $p \leqslant k_1 \| \cdot \|_2$.

Falta agora provar que existe $k_2 > 0$ tal que $\| \cdot \|_2 \leqslant k_2 p$. O conjunto $\{\, v \in \mathbb{R}^n \mid \|v\|_2 = 1 \,\}$ é um compacto de $(V, \| \cdot \|_2)$ (pois é fechado e limitado) e, por outro lado, resulta da proposição 5.8 e de se ter $p \leqslant k_1 \| \cdot \|_2$ que a função identidade de $(V, \| \cdot \|_2)$ em (V, p) é contínua. Logo, o conjunto $\{\, v \in V \mid \|v\|_2 = 1 \,\}$, que é um compacto de $(V, \| \cdot \|_2)$ (pois é uma parte fechada e limitada de $\mathbb{R}^n$, relativamente à métrica usual), também é um compacto de (V, p) e, portanto, a sua imagem por p, i. e.

$$\{\, p(v) \in V \mid \|v\|_2 = 1 \,\} \tag{5.11}$$

é um compacto de $\mathbb{R}_+$; seja m o seu ínfimo. Então $m > 0$, pois pertence ao conjunto (5.11), ao qual 0 não pertence. Vejamos que $k_2 = 1/m$ está nas condições pretendidas. Para tal, seja $v \in \mathbb{R}^n$; quer-se provar que $\|v\|_2 \leqslant k_2 p(v)$. Se $v = 0$, isto é óbvio. Caso contrário, $\big\| v / \|v\|_2 \big\|_2 = 1$ e, portanto, pelas definições de m e de k_2,

$$m \leqslant p\left(\frac{v}{\|v\|_2}\right) \iff \|v\|_2 \leqslant k_2 p(v).$$

■

Se (V, p) for um espaço vectorial normado e se W for um subespaço vectorial de V então, a menos que seja dito explicitamente o contrário, vai-se considerar em W a norma $p|_W$.

COROLÁRIO 5.3 *Qualquer aplicação linear de um espaço vectorial normado de dimensão finita num espaço vectorial normado é contínua.*

DEMONSTRAÇÃO: Basta provar que qualquer aplicação linear f de $\mathbb{R}^n$ ($n \in \mathbb{N}$) num espaço vectorial normado W é contínua se se considerar em $\mathbb{R}^n$ a norma $\|\cdot\|_1$ pois, como qualquer norma em $\mathbb{R}^n$ é equivalente à norma $\|\cdot\|_1$, resulta daqui que qualquer f é contínua seja qual for a norma que se considere em $\mathbb{R}^n$.

Seja $\{e_1, e_2, \ldots, e_n\}$ a base canónica de $\mathbb{R}^n$. Então, se $(x_1, \ldots, x_n) \in \mathbb{R}^n$,

$$\begin{aligned}
\|f(x_1, \ldots, x_n)\| &= \left\| f\left(\sum_{k=1}^{n} x_k e_k\right)\right\| \\
&= \left\|\sum_{k=1}^{n} x_k f(e_k)\right\| \\
&\leqslant \sum_{k=1}^{n} |x_k| \|f(e_k)\| \\
&\leqslant \left(\max_{1\leqslant k\leqslant n} \|f(e_k)\|\right) \sum_{k=1}^{n} |x_k| \\
&= \left(\max_{1\leqslant k\leqslant n} \|f(e_k)\|\right) \big\|(x_1, x_2, \ldots, x_n)\big\|_1.
\end{aligned}$$

■

Outra consequência do teorema 5.3 já foi enunciada na página 116.

COROLÁRIO 5.4 *Quelquer subespaço vectorial de dimensão finita de um espaço vectorial normado é fechado.*

DEMONSTRAÇÃO: Seja $(V, \|\cdot\|)$ um espaço vectorial normado e seja W um subespaço vectorial de dimensão finita; quer-se provar que W é um fechado de V. Para tal, vai-se provar que se $v \in V \setminus W$, então $v \notin \overline{W}$, o que equivale a afirmar que $W = \overline{W}$, i. e. que W é fechado. Considere-se a função

$$\begin{array}{rccc}
p\colon & W \bigoplus \mathbb{R}v & \longrightarrow & \mathbb{R}_+ \\
& w + \lambda v & \mapsto & \|w\| + |\lambda|.
\end{array}$$

Então p é uma norma e, como a dimensão de $W \bigoplus \mathbb{R}v$ é finita, p é equivalente à restrição a $W \bigoplus \mathbb{R}v$ da norma $\|\cdot\|$. Em particular, existe algum $k > 0$ tal que

$$(\forall w \in W)(\forall \lambda \in \mathbb{R}) : p(w + \lambda v) \leqslant k\|w + \lambda v\|$$

e, portanto, se $w \in W$ tem-se $1 \leqslant \|w\| + 1 = p(w - v) \leqslant k\|w - v\|$. Mas então tem-se $(\forall w \in W) : \|w - v\| \geqslant 1/k$. Logo, $v \notin \overline{W}$. ■

Antes de se passar ao próximo teorema, vejamos o que significa a compacidade local no contexto dos espaços vectoriais normados. Em geral, um espaço métrico E diz-se localmente compacto se qualquer vizinhança V de qualquer $p \in E$ contiver alguma vizinhança compacta de p, o que equivale a afirmar que qualquer ponto de E tem alguma vizinhança compacta.

PROPOSIÇÃO 5.9 *Se V for um espaço vectorial normado, são condições equivalentes:*

1. *V é localmente compacto;*
2. *a bola fechada unitária de V é compacta;*
3. *qualquer parte fechada e limitada de V é compacta.*

DEMONSTRAÇÃO: É conveniente começar por observar que todas a bolas fechadas são homeomorfas entre si. De facto, se $v \in V$ e $r > 0$, então a função

$$\begin{array}{ccc} \overline{B(0,1)} & \longrightarrow & \overline{B(v,r)} \\ w & \mapsto & v + rw \end{array}$$

é um homeomorfismo.

Vai-se começar por mostrar que a primeira condição implica a terceira. Seja então L uma parte fechada e limitada de V e seja K uma vizinhança compacta de 0. Visto que K é uma vizinhança de 0, existe algum $r > 0$ tal que $\overline{B(0,r)} \subset K$ e, como $\overline{B(0,r)}$ é fechado e K é compacto, $\overline{B(0,r)}$ é compacto. Seja $R \geqslant 0$ tal que $L \subset \overline{B(0,R)}$. Se $f\colon \overline{B(0,R)} \longrightarrow \overline{B(0,r)}$ for um homeomorfismo, então $f(L)$ é um fechado de $\overline{B(0,r)}$ e, portanto, é compacto. Logo, L é compacto.

É trivial que a terceira condição implica a segunda.

Finalmente, se a segunda condição se verifica, então todas as bolas fechadas são compactas e, portanto, V é localmente compacto. ∎

TEOREMA 5.4 (TEOREMA DE RIESZ) *Um espaço vectorial normado é localmente compacto se e só se tiver dimensão finita.*

DEMONSTRAÇÃO: O espaço $\mathbb{R}^n$ ($n \in \mathbb{N}$) munida da norma usual é localmente compacto. Como todas as normas em $\mathbb{R}^n$ são equivalentes à usual, resulta desta observação que $(\mathbb{R}^n, \|\cdot\|)$ é localmente compacto, seja qual for a norma que se esteja a considerar e, consequentemente, que qualquer espaço vectorial normado de dimensão finita é localmente compacto.

Seja agora $(V, \|\cdot\|)$ um espaço vectorial normado localmente compacto; vai-se provar que tem dimensão finita. Pela proposição 5.9, a bola fechada unitária de V, que vai ser representada por B, é compacta. Uma vez que

$$B \subset \bigcup_{v \in B} \left(v + \frac{1}{2}\mathring{B} \right),$$

existem vectores $v_1, \ldots, v_n \in B$ tais que

$$B \subset \left(v_1 + \frac{1}{2}\mathring{B} \right) \cup \left(v_2 + \frac{1}{2}\mathring{B} \right) \cup \cdots \cup \left(v_n + \frac{1}{2}\mathring{B} \right).$$

Seja M o espaço vectorial gerado pelos vectores $v_1, \ldots, v_n$. Então

$$\begin{aligned} B &\subset M + \frac{1}{2}B \\ &\subset M + \frac{1}{2}\left(M + \frac{1}{2}B \right) \\ &= M + \frac{1}{4}B \\ &\subset M + \frac{1}{4}\left(M + \frac{1}{2}B \right) \\ &= M + \frac{1}{8}B \\ &\ldots \end{aligned}$$

Isto prova que $(\forall n \in \mathbb{N}) : B \subset M + 2^{-n}B$ e, portanto,

$$B \subset \bigcap_{n \in \mathbb{N}} \left(M + 2^{-n}B \right). \tag{5.12}$$

Vejamos que $\bigcap_{n \in \mathbb{N}} \left(M + 2^{-n}B \right) = \overline{M}$. Se $v \in \overline{M}$ e se $n \in \mathbb{N}$ então, pela definição de aderência, existe algum $v_n \in M$ tal que $\|v - v_n\| \leqslant 2^{-n}$ e então, como $v = v_n + (v - v_n)$ e $v - v_n \in 2^{-n}B$, $v \in M + 2^{-n}B$; como isto tem lugar para qualquer $n \in \mathbb{N}$, $v \in \bigcap_{n \in \mathbb{N}} \left(M + 2^{-n}B \right)$. Reciprocamente, se $v \in \bigcap_{n \in \mathbb{N}} \left(M + 2^{-n}B \right)$ e se $\varepsilon > 0$, tome-se $n \in \mathbb{N}$ tal que $2^{-n} < \varepsilon$ e tome-se $w \in M$ tal que $v \in w + 2^{-n}B$. Então $\|v - w\| = 2^{-n} < \varepsilon$; logo, $v \in \overline{M}$.

Resulta então de (5.12) e do que foi demonstrado no parágrafo anterior que $B \subset \overline{M}$. Como M tem dimensão finita, tem-se, pelo corolário 5.4, que $\overline{M} = M$. Mas então provou-se que $B \subset M$, pelo que $V \subset M$; logo, $V = M$. ∎

5.5 O teorema de Hahn-Banach

TEOREMA 5.5 (TEOREMA DE HAHN-BANACH) *Sejam V um espaço vectorial normado, W um subespaço vectorial de V e $f \in W'$. Existe então alguma forma linear $F \in V'$ que prolonga f e tal que $\|F\| = \|f\|$.*

DEMONSTRAÇÃO: Convém começar por observar que se F for um prolongamento de f a um subespaço vectorial U de V que contenha W e se F for linear e contínua, então resulta da definição de norma de uma aplicação linear contínua que $\|F\| \geqslant \|f\|$; logo, provar que $\|F\| = \|f\|$ equivale a provar que $\|F\| \leqslant \|f\|$.

Seja $\mathcal{V}$ o conjunto de todos os pares ordenados (U, F) tais que

1. U é um subespaço vectorial de V que contém W;
2. $F \in U'$ e é um prolongamento de f tal que $\|F\| = \|f\|$

e considere-se em $\mathcal{V}$ a relação de ordem $\leqslant$ assim definida: $(W_1, F_1) \leqslant (W_2, F_2)$ se e só se $W_1 \subset W_2$ e $F_1 = F_2|_{W_1}$. Se se provar que existe algum elemento $(U, F) \in \mathcal{V}$ com $U = V$, o teorema estará demonstrado. A demonstração de que existe efectivamente um tal elemento de $\mathcal{V}$ será feita em dois passos. Primeiro demonstrar-se-á, recorrendo ao lema de Zorn, que $(\mathcal{V}, \leqslant)$ tem algum elemento maximal. Em seguida, será visto que se (Z, F) for um tal elemento maximal, então tem-se necessariamente $Z = V$.

A fim de se poder aplicar o lema de Zorn ao conjunto $\mathcal{V}$ e à relação de ordem $\leqslant$, considere-se uma família $\{\, (U_i, F_i) \mid i \in I \,\}$ de $\mathcal{V}$ totalmente ordenada; quer-se provar que é majorada. Seja $U = \bigcup_{i \in I} U_i$. Se $v \in U$ e se $i, j \in I$ são tais que $v \in U_i$ e $v \in U_j$, então tem-se $(U_i, F_i) \leqslant (U_j, F_j)$ ou $(U_j, F_j) \leqslant (U_i, F_i)$; em ambos os casos, resulta da definição da relação de ordem $\leqslant$ que $F_i(v) = F_j(v)$. Sendo assim, faz sentido definir $F \colon U \longrightarrow \mathbb{R}$ do seguinte modo: se $v \in V$, seja $i \in I$ tal que $v \in U_i$; então $f(v) = F_i(v)$. Vejamos que U é um subespaço vectorial de V e que $F \in U^*$. Se $v, w \in U$ e se $\alpha, \beta \in \mathbb{R}$, sejam $i, j \in I$ tais que $v \in U_i$ e que $w \in U_j$. Então $(U_i, F_i) \leqslant (U_j, F_j)$ ou vice-versa. No primeiro caso, então $v, w \in U_j$, pelo que $\alpha v + \beta w \in U_j$ e

$$F(\alpha v + \beta w) = F_j(\alpha v + \beta w) = \alpha F_j(v) + \beta F_j(w) = \alpha F(v) + \beta F(w);$$

o segundo caso é análogo. Finalmente, vejamos que F é contínua e que $\|F\| \leqslant \|f\|$. Se $v \in U$ é tal que $\|v\| \leqslant 1$, seja $i \in I$ tal que $v \in U_i$. Então

$$\big|F(v)\big| = \big|F_i(v)\big| \leqslant \|F_i\| = \|f\|.$$

É óbvio que $(\forall i \in I) : (U_i, F_i) \leqslant (U, F)$. Verificam-se então as hipóteses do lema de Zorn, pelo que existe algum elemento maximal (Z, F) de $\mathcal{V}$.

Suponha-se que $Z \subsetneq V$ e seja então $v \in V \setminus Z$. Para cada $\alpha \in \mathbb{R}$, seja

$$\begin{aligned} F_\alpha \colon \quad Z \bigoplus \mathbb{R}v & \longrightarrow \quad \mathbb{R} \\ w + \lambda v & \mapsto \quad F(w) + \lambda\alpha. \end{aligned}$$

Então $F_\alpha \in \left(Z \bigoplus \mathbb{R}v\right)^*$ e é um prolongamento de F. Quer-se mostrar que existe algum $\alpha \in \mathbb{R}$ tal que F_α seja contínua e que $\|F_\alpha\| = \|F\|$, o que estará em contradição com a maximalidade do par (Z, F). Afirmar que F_α é contínua e que $\|F_\alpha\| = \|F\|$ equivale a

$$\begin{aligned} & (\forall w \in Z)(\forall \lambda \in \mathbb{R}) : |F_\alpha(w + \lambda v)| \leqslant \|F\| \|w + \lambda v\| \iff \\ & \iff (\forall w \in Z)(\forall \lambda \in \mathbb{R}) : |F(w) + \lambda\alpha| \leqslant \|F\| \|w + \lambda v\| \\ & \iff (\forall w \in Z)(\forall \lambda \in \mathbb{R}) : |F(-\lambda w) + \lambda\alpha| \leqslant \|F\| \| - \lambda w + \lambda v\| \\ & \iff (\forall w \in Z) : |F(w) - \alpha| \leqslant \|F\| \|w - v\| \\ & \iff (\forall w \in Z) : F(w) - \alpha \leqslant \|F\| \|w - v\| \wedge -F(w) + \alpha \leqslant \|w - v\| \\ & \iff (\forall w \in Z) : F(w) - \|F\| \|w - v\| \leqslant \alpha \leqslant F(w) + \|F\| \|w - v\|. \end{aligned}$$

É necessário então que se mostre que o conjunto

$$\left\{ \left[F(w) - \| F\| \|w - v\|, F(w) + \|F\| \|w - v\|\right] | w \in Z \right\} \tag{5.13}$$

tem intersecção não vazia. Para tal, veja-se que se $w_1, w_2 \in Z$, então

$$\begin{aligned} F(w_1) - F(w_2) &= F(w_1 - w_2) \\ &\leqslant \left|F(w_1 - w_2)\right| \\ &\leqslant \|F\| \|w_1 - w_2\| \\ &\leqslant \|F\| \|w_1 - v\| + \|F\| \|v - w_2\| \end{aligned}$$

o que equivale a afirmar que

$$F(w_1) - \|F\| \|w_1 - v\| \leqslant F(w_2) + \|F\| \|w_2 - v\|.$$

Logo,

$$\sup_{w \in Z} \left(F(w) - \|F\| \|w - v\|\right) \leqslant \inf_{w \in Z} \left(F(w) + \|F\| \|w - v\|\right)$$

e, naturalmente, qualquer número situado entre estes dois números pertence a todos os intervalos do conjunto (5.13). ■

COROLÁRIO 5.5 *Se V é um espaço vectorial normado e v e w são dois elementos distintos de V, então existe algum $f \in V'$ tal que $f(v) \neq f(w)$.*

DEMONSTRAÇÃO: Tal como na demonstração da proposição 5.2, basta demonstrar que existe alguma forma linear $f \in V'$ tal que $f(v-w) \neq 0$. Seja $W = \mathbb{R}(v-w)$ e seja

$$\begin{array}{rccc} \eta\colon & W & \longrightarrow & \mathbb{R} \\ & \lambda(v-w) & \mapsto & \lambda. \end{array}$$

Então $\eta \in W'$ e é portanto prolongável a alguma forma linear contínua $f \in V'$. Naturalmente, $f(v-w) = \eta(v-w) = 1 \neq 0$. ∎

COROLÁRIO 5.6 *Seja W um subespaço vectorial de um espaço vectorial normado V e seja $v \in V$. São então condições equivalentes:*

1. *$v \in \overline{W}$;*
2. *para cada $f \in V'$ tal que $f(W) = \{0\}$, $f(v) = 0$.*

DEMONSTRAÇÃO: Se $v \in \overline{W}$ e se $f \in V'$ se anula em W então, visto que f é contínua, f também se anula em $\overline{W}$, pelo que $f(v) = 0$.

Seja agora $v \in V \setminus \overline{W}$; quer-se mostrar que existe algum $f \in V'$ que se anula em W e tal que $f(v) \neq 0$. Seja

$$m = \inf\{\ \| v - w\| | w \in W\ \}.$$

Então $m > 0$, pois se se tivesse $m = 0$ isso implicaria que haveria alguma sucessão $(w_n)_{n\in\mathbb{N}}$ de pontos de W tal que $\lim_{n\in\mathbb{N}}(v - w_n)$, i. e. que $v \in \overline{W}$. Mas então

$$(\forall \lambda \in \mathbb{R})(\forall w \in W) : \|\lambda v + w\| \geqslant m|\lambda|; \tag{5.14}$$

isto é trivial se $\lambda = 0$ e, caso contrário,

$$\|\lambda v + w\| = |\lambda| \big\|v - (-\lambda^{-1}w)\big\| \geqslant m|\lambda|.$$

Considere-se a forma linear $f \in \left(W \bigoplus \mathbb{R}v\right)^*$ tal que, para cada $\lambda \in \mathbb{R}$ e cada $w \in W$, $f(\lambda v + w) = \lambda$. Então f é contínua, pois se $\lambda \in \mathbb{R}$ e $w \in W$, então

$$\big|f(\lambda v + w)\big| = |\lambda| \leqslant m^{-1}\|\lambda v + w\|$$

por (5.14). Logo, é prolongável a alguma forma linear contínua $F \in V'$ e então, visto que F é um prolongamento de f, $F(W) = \{0\}$ e $F(v) = 1$. ∎

Afirmar que as condições do enunciado anterior são equivalentes para cada $v \in V$ é o mesmo que dizer que

$$\overline{W} = \bigcap_{\substack{f \in V' \\ f(W)=\{0\}}} \ker f.$$

Espaços de Banach

6.1 Definição e propriedades elementares

DEFINIÇÃO 6.1 Designa-se por *espaço de Banach* um espaço vectorial normado completo.

EXEMPLO 6.1 Se $\|\cdot\|$ for a norma usual em $\mathbb{R}^n$ ($n \in \mathbb{N}$), então $(\mathbb{R}^n, \|\cdot\|)$ é um espaço de Banach.[1]

EXEMPLO 6.2 Sejam $p \in [1, +\infty]$ e $(X, \mathcal{A}, m)$ um espaço de medida. Então o teorema 4.3 afirma que $L^p(X)$ é um espaço de Banach.

EXEMPLO 6.3 Considere-se no espaço $l^1(\mathbb{N})$ das séries absolutamente convergentes a norma

$$\begin{array}{rccc} \|\cdot\|_2\colon & l^1(\mathbb{N}) & \longrightarrow & \mathbb{R}_+ \\ & \displaystyle\sum_{n=1}^{+\infty} a_n & \mapsto & \sqrt{\displaystyle\sum_{n=1}^{+\infty} |a_n|^2}. \end{array}$$

Então $\big(l^1(\mathbb{N}), \|\cdot\|_2\big)$ não é um espaço de Banach. Para ver porquê, considere-se a seguinte sucessão $(s_n)_{n\in\mathbb{N}}$ de elementos de $l^1(\mathbb{N})$:

$$\begin{aligned} s_1 &= (1, 0, 0, 0, 0, 0, \ldots) \\ s_2 &= (1, 1/2, 0, 0, 0, 0, \ldots) \\ s_3 &= (1, 1/2, 1/3, 0, 0, 0, \ldots) \\ s_4 &= (1, 1/2, 1/3, 1/4, 0, 0, \ldots) \\ &\ldots \end{aligned}$$

[1] De facto, será visto mais à frente que $(\mathbb{R}^n, \|\cdot\|)$ é um espaço de Banach seja qual for a norma que esteja a ser considerada em $\mathbb{R}^n$.

Trata-se de uma sucessão de Cauchy de elementos de $\left(l^1, \|\cdot\|_2\right)$, pois

$$(\forall p, m, n \in \mathbb{N}) : m, n \geqslant p \Longrightarrow \|s_m - s_n\|_2 \leqslant \sqrt{\sum_{k=p}^{+\infty} \frac{1}{k^2}}.$$

No entanto, a sucessão $(s_n)_{n\in\mathbb{N}}$ não é convergente. Se o fosse e se convergisse para $s = (a_1, a_2, a_3, \ldots)$, então ter-se-ia necessariamente $a_n = 1/n$ para cada $n \in \mathbb{N}$, pois

$$(\forall m \in \mathbb{N}) : m \geqslant n \Longrightarrow |a_n - 1/n| \leqslant \|s - s_m\|_2$$

e $\lim_{m\in\mathbb{N}} \|s - s_m\|_2 = 0$. Mas $(1/n)_{n\in\mathbb{N}} \notin l^1(\mathbb{N})$.

O facto de $\mathbb{R}^n$ ou $\mathbb{C}^n$ ($n \in \mathbb{N}$) serem completos permite-nos saber que certas sucessões de elemetos destes espaços convergem sem se ter que efectivamente determinar os respectivos limites. Nos espaços de Banach a situação é a mesma. Vejamos um exemplo deste tipo de ideias, para o qual é necessário começar por introduzir um conceito.

DEFINIÇÃO 6.2 Seja V um espaço vectorial normado. Diz-se que uma série $\sum_{n=1}^{+\infty} v_n$ de elementos de V é *absolutamente convergente* se a série $\sum_{n=1}^{+\infty} \|v_n\|$ for convergente.

PROPOSIÇÃO 6.1 *Num espaço de Banach, qualquer série absolutamente convergente é convergente.*

DEMONSTRAÇÃO: Seja $\sum_{n=1}^{+\infty} v_n$ uma série absolutamente convergente de vectores de um espaço de Banach V e sejam $m, n, p \in \mathbb{N}$ tais que $m > n \geqslant p$. Então

$$\left\|\sum_{k=1}^{m} v_k - \sum_{k=1}^{n} v_k\right\| = \left\|\sum_{k=n+1}^{m} v_k\right\| \leqslant \sum_{k=n+1}^{m} \|v_k\| \leqslant \sum_{k=p}^{+\infty} \|v_k\|. \tag{6.1}$$

Como $\lim_{p\in\mathbb{N}} \sum_{k=p}^{+\infty} \|v_k\| = 0$, resulta de (6.1) que a sucessão das somas parciais da série $\sum_{n=1}^{+\infty} v_n$ é de Cauchy e, portanto, converge. ∎

PROPOSIÇÃO 6.2 *Um subespaço vectorial W de um espaço de Banach V é um espaço de Banach se e só se W for um fechado de V.*

DEMONSTRAÇÃO: Este resultado é uma consequência imediata de, dado um espaço métrico completo E, um subespaço de E ser completo se e só se for fechado. ∎

Este resultado é muitas vezes empregue para estudar o problema de saber se um dado espaço vectorial normado é ou não um espaço de Banach. Considere-se novamente o exemplo 6.3. Aquilo que se demonstrou de facto foi que o espaço $l^1(\mathbb{N})$ das séries absolutamente convergentes de números reais não é um subespaço fechado de $\left(l^2(\mathbb{N}), \|\cdot\|_2\right)$; para tal, foi visto um exemplo de uma sucessão de pontos de $l^1(\mathbb{N})$ que converge para um ponto de $l^2(\mathbb{N})$ que não pertence a $l^1(\mathbb{N})$. Será demonstrado que, num certo sentido, é sempre possível fazer isto, ou seja, que é possível «mergulhar» qualquer espaço vectorial normado não completo V num espaço de Banach W e, naturalmente, V não será um subespaço fechado de W.

PROPOSIÇÃO 6.3 *Sejam V um espaço vectorial normado p e q duas normas equivalentes definidas em V. Então (V, p) é um espaço de Banach se e só se (V, q) o for.*

DEMONSTRAÇÃO: Afirmar que as normas p e q são equivalentes equivale a afirmar, pela proposição 5.8, que os abertos de (V, p) e de (V, q) são os mesmos e isto é o mesmo que dizer que $\mathrm{id}\colon (V, p) \longrightarrow (V, q)$ é um homeomorfismo. Mas então, pelo teorema 5.2, trata-se de uma função uniformemente contínua com inversa uniformemente contínua, pelo que, dada uma sucessão $(v_n)_{n \in \mathbb{N}}$ de elementos de V, ela é uma sucessão de Cauchy em (V, p) se e só se for uma sucessão de Cauchy em (V, q). Logo, (V, p) e (V, q) têm as mesmas sucessões de Cauchy e, como p e q são equivalentes, as mesmas sucessões convergentes. Logo, afirmar que (V, p) é completo equivale a afirmar que (V, q) é completo. ∎

Resulta desta proposição, do teorema 5.3 e do que foi mencionado no exemplo 6.1 que $\mathbb{R}^n$ ($n \in \mathbb{N}$) é completo para qualquer norma.

PROPOSIÇÃO 6.4 *Sejam V e W espaços vectoriais normados. Se W for um espaço de Banach, então $\mathcal{L}(V, W)$ também o é.*

DEMONSTRAÇÃO: Seja $(f_n)_{n \in \mathbb{N}}$ uma sucessão de Cauchy de elementos de $\mathcal{L}(V, W)$; quer-se mostrar que é convergente.

Seja $v \in V$. Se $m, n \in \mathbb{N}$, então

$$\left\|f_m(v) - f_n(v)\right\| \leqslant \|f_m - f_n\| \|v\|$$

e, portanto, a sucessão $\left(f_n\right)_{n \in \mathbb{N}}$ é uma sucessão de Cauchy de elementos de W; logo, converge. Seja $f(w) = \lim_{n \in \mathbb{N}} f_n(w)$. Vai-se mostrar que $f \in \mathcal{L}(V, W)$ e que $\lim_{n \in \mathbb{N}} f_n = f$.

A função f é linear, pois se $v, w \in V$ e se $\alpha, \beta \in \mathbb{R}$, tem-se

$$\begin{aligned} f(\alpha v + \beta w) &= \lim_{n \in \mathbb{N}} f_n(\alpha v + \beta w) \\ &= \lim_{n \in \mathbb{N}} \big(\alpha f_n(v) + \beta f_n(w)\big) \\ &= \alpha f(v) + \beta f(w). \end{aligned}$$

Vejamos agora que f é contínua. Seja $\varepsilon > 0$. Existe algum $p \in \mathbb{N}$ tal que

$$(\forall m, n \in \mathbb{N}) : m, n \geqslant p \Longrightarrow \|f_m - f_n\| < \varepsilon.$$

Logo, se $v \in V$ e se $m, n \in \mathbb{N}$ são tais que $m, n \geqslant p$,

$$\|f_m(v) - f_n(v)\| = \big\|(f_m - f_n)(v)\big\| \leqslant \|f_m - f_n\|\|v\| < \varepsilon\|v\|$$

e, portanto,

$$(\forall n \in \mathbb{N}) : n \geqslant p \Longrightarrow \big\|(f - f_n)(v)\big\| \leqslant \varepsilon\|v\|. \tag{6.2}$$

Em particular, $f - f_p$ é contínua e, como $f = (f - f_p) + f_p$, f é contínua. Por outro lado, resulta de (6.2) que se $n \geqslant p$, então $\|f - f_n\| \leqslant \varepsilon$, pelo que $\lim_{n \in \mathbb{N}} f_n = f$. ∎

Seja V um espaço de Banach. Resulta da proposição anterior que $\mathcal{L}(V, V)$ também é um espaço de Banach e este facto, juntamente com a proposição 6.1, permite definir, por exemplo, $\exp(f)$ $(f \in \mathcal{L}(V, V))$ por

$$\exp(f) = \sum_{n=0}^{+\infty} \frac{f^n}{n!}.$$

Esta soma converge pois, pela proposição 5.6, tem-se $\big\|f^n\big\| \leqslant \|f\|^n$ para cada $n \in \mathbb{N}$, de onde resulta que a série $\sum_{n=0}^{+\infty} f^n/n!$ é absolutamente convergente e, em particular, é convergente. Analogamente, poder-se-iam definir $\operatorname{sen}(f)$, $\cos(f)$, etc.

COROLÁRIO 6.1 *Se V é um espaço vectorial normado, então V' é um espaço de Banach.*

DEMONSTRAÇÃO: Basta ver que $\mathbb{R}$ é completo. ∎

PROPOSIÇÃO 6.5 *Sejam V um espaço vectorial normado, W um subespaço vectorial de V e f uma aplicação linear contínua de W num espaço de Banach U. Então f é prolongável a uma e uma só aplicação linear contínua de $\overline{W}$ em U e a norma de um tal prolongamento é igual à de f.*

DEMONSTRAÇÃO: A unicidade é muito simples de demonstrar e não depende de U ser um espaço de Banach. Basta observar que se F_1 e F_2 forem prolongamentos de f a $\overline{W}$ nas condições do enunciado, então, visto que F_1 e F_2 são contínuas, o conjunto $\{\, w \in \overline{W} \mid F_1(w) = F_2(w) \,\}$ é um fechado de $\overline{W}$. Mas por outro lado, este conjunto contém W e o único fechado de $\overline{W}$ que contém W é $\overline{W}$.

Passemos à existência. Seja $w \in \overline{W}$ e seja $(w_n)_{n\in\mathbb{N}}$ uma sucessão de elementos de W que converge para w. Então $(w_n)_{n\in\mathbb{N}}$ é uma sucessão de Cauchy e, como f é uniformemente contínua (pelo teorema 5.2), a sucessão $\big(f(w_n)\big)_{n\in\mathbb{N}}$ também o é. Logo, visto que se está a supor que U é completo, a sucessão converge. O limite não depende da escolha da sucessão $(w_n)_{n\in\mathbb{N}}$, pois se $(w'_n)_{n\in\mathbb{N}}$ for outra sucessão de elementos de W que converge para w, então $\lim_{n\in\mathbb{N}} w_n - w'_n = 0$ e, portanto, pela continuidade e pela linearidade de f, $\lim_{n\in\mathbb{N}} f(w_n) - f(w'_n) = 0$, o que é o mesmo que dizer que $\lim_{n\in\mathbb{N}} f(w_n) = \lim_{n\in\mathbb{N}} f(w'_n)$. Faz então sentido definir $F\colon \overline{W} \longrightarrow U$ do seguinte modo: se $w \in \overline{W}$, seja $(w_n)_{n\in\mathbb{N}}$ uma sucessão de elementos de W que converge para w; define-se então $F(w)$ como sendo o limite da sucessão $\big(f(w_n)\big)_{n\in\mathbb{N}}$. É claro que F é um prolongamento de f, pois se $w \in W$, toma-se a sucessão $(w_n)_{n\in\mathbb{N}}$ tal que $(\forall n \in \mathbb{N}) : w_n = w$ e então

$$F(w) = \lim_{n\in\mathbb{N}} f(w) = f(w).$$

Vejamos que F é linear. Sejam $\alpha, \beta \in \mathbb{R}$ e sejam $v, w \in \overline{W}$; quer-se provar que $F(\alpha v + \beta w) = \alpha F(v) + \beta F(w)$. Sejam $(v_n)_{n\in\mathbb{N}}$ e $(w_n)_{n\in\mathbb{N}}$ sucessões de elementos de W convergentes para v e para w respectivamente. Então $\lim_{n\in\mathbb{N}} \alpha v_n + \beta w_n = \alpha v + \beta w$ e

$$\begin{aligned} F(\alpha v + \beta w) &= f\left(\lim_{n\in\mathbb{N}} \big(\alpha v_n + \beta w_n\big)\right) \\ &= \alpha \lim_{n\in\mathbb{N}} f(v_n) + \beta \lim_{n\in\mathbb{N}} f(w_n) \\ &= \alpha F(v) + \beta F(w). \end{aligned}$$

A função F é contínua, pois se $w \in W$ e se $(w_n)_{n\in\mathbb{N}}$ é uma sucessão de elementos de W que converge para w, então

$$\begin{aligned} \|F(w)\| &= \big\|\lim_{n\in\mathbb{N}} f(w_n)\big\| \\ &= \lim_{n\in\mathbb{N}} \big\|f(w_n)\big\| \\ &\leqslant \lim_{n\in\mathbb{N}} \|f\|\|w_n\| \\ &= \|f\|\|w\|. \end{aligned}$$

Resulta também daqui que $\|F\| \leqslant \|f\|$ e então, tal como na demonstração do teorema de Hahn-Banach, $\|F\| = \|f\|$. ■

DEFINIÇÃO 6.3 Se V e W são espaços vectoriais normados, diz-se que uma função f de V em W é uma *isometria* se

$$(\forall v, w \in V) : \|f(v) - f(w)\| = \|v - w\|. \tag{6.3}$$

Naturalmente, se, nas condições da definição anterior, f for linear, então a condição (6.3) equivale à condição

$$(\forall v \in V) : \|f(v)\| = \|v\|.$$

Por outro lado, observe-se que a condição (6.3) afirma simplesmente que a distância entre dois pontos de V é igual à distância entre as suas imagens; é precisamente assim que se define o conceito de isometria entre espaços métricos.

Verifica-se facilmente que qualquer isometria é injectiva e que, caso seja sobrejectiva, a inversa é também uma isometria.

TEOREMA 6.1 *Seja V um espaço vectorial normado. Existe então um espaço de Banach W e uma isometria linear $\iota\colon V \longrightarrow W$ tal que $\iota(V)$ é um subespaço denso de W. Além disso, se $\widetilde{W}$ for outro espaço de Banach e se $\tilde{\iota}$ for uma isometria linear de V em $\widetilde{W}$ tal que $\tilde{\iota}(V)$ seja um subespaço denso de $\widetilde{W}$, então existe uma e uma só aplicação linear contínua $\psi\colon W \longrightarrow \widetilde{W}$ tal que $\psi \circ \iota = \tilde{\iota}$. Esta aplicação linear é uma isometria bijectiva.*

DEMONSTRAÇÃO: Seja S o espaço das sucessões de Cauchy de elementos de V. O espaço S tem uma estrutura natural de espaço vectorial, assim definida: se $(v_n)_{n\in\mathbb{N}}, (w_n)_{n\in\mathbb{N}} \in S$ e $\alpha, \beta \in \mathbb{R}$, então

$$\alpha(v_n)_{n\in\mathbb{N}} + \beta(w_n)_{n\in\mathbb{N}} = (\alpha v_n + \beta w_n)_{n\in\mathbb{N}};$$

verifica-se facilmente que esta última sucessão também é de Cauchy. Se $(v_n)_{n\in\mathbb{N}} \in S$, então $(\|v_n\|)_{n\in\mathbb{N}}$ é uma sucessão de Cauchy de números reais, visto que, pela proposição 5.4, $\|\cdot\|$ é uniformemente contínua. Pode--se então definir

$$\begin{array}{cccc} p\colon & S & \longrightarrow & \mathbb{R} \\ & (v_n)_{n\in\mathbb{N}} & \mapsto & \lim\limits_{n\in\mathbb{N}} \|v_n\|. \end{array}$$

Cálculos simples mostram que p é uma semi-norma. Se $(W, \|\cdot\|_W)$ for o espaço vectorial normado obtido a partir de S e de p pelo processo descrito na proposição 4.7 então, por definição, W é o quociente de S pelo espaço vectorial formado pelas sucessões $(v_n)_{n\in\mathbb{N}} \in S$ tais que $\lim_{n\in\mathbb{N}} \|v\| = 0$, i. e. pelo espaço das sucessões convergentes para 0. Seja $\iota\colon V \longrightarrow W$ a função assim definida: se $v \in V$, então $\iota(v)$ é a classe de equivalência da sucessão constante que toma sempre o valor v. É simples verificar que ι é uma isometria linear.

Vejamos que $\iota(V)$ é um subespaço denso de W. Para tal, considere-se a classe de equivalência $\big[(v_n)_{n\in\mathbb{N}}\big]$ de um elemento de S; vai-se mostrar que

$$\lim_{n\in\mathbb{N}} \big[\iota(v_n)\big] = \big[(v_n)_{n\in\mathbb{N}}\big]. \tag{6.4}$$

Seja então $\varepsilon > 0$ e seja $N \in \mathbb{N}$ tal que

$$(\forall r, s \in \mathbb{N}) : r, s \geqslant N \Longrightarrow \|v_r - v_s\| < \varepsilon.$$

Então, se $r \in \mathbb{N}$ é tal que $r \geqslant N$,

$$\big\|\big[(v_n)_{n\in\mathbb{N}}\big] - \big[\iota(v_r)\big]\big\|_W = \lim_{n\in\mathbb{N}} \|v_n - v_r\| \leqslant \varepsilon,$$

pois $n \geqslant N \Longrightarrow \|v_n - v_r\| < \varepsilon$.

Vejamos agora que W é um espaço de Banach. Seja $(w_n)_{n\in\mathbb{N}}$ uma sucessão de Cauchy de elementos de W; quer-se mostrar que converge. Para cada $n \in \mathbb{N}$, seja $v_n \in V$ tal que

$$\big\|w_n - \iota(v_n)\big\|_W < \frac{1}{n}; \tag{6.5}$$

um tal vector existe visto que $\iota(V)$ é denso em W. Então a sucessão $(v_n)_{n\in\mathbb{N}}$ é uma sucessão de Cauchy de elementos de V, pois se $\varepsilon > 0$ e se se tomar $N \in \mathbb{N}$ tal que

- $(\forall r, s \in \mathbb{N}) : r, s \geqslant N \Longrightarrow \big\|w_r - w_s\big\|_W < \dfrac{\varepsilon}{3}$;
- $\dfrac{1}{N} < \dfrac{\varepsilon}{3}$,

então, se $r, s \in \mathbb{N}$ forem tais que $r, s \geqslant N$,

$$\begin{aligned}
\|v_r - v_s\| &= \big\|\iota(v_r) - \iota(v_s)\big\|_W \\
&\leqslant \big\|\iota(v_r) - w_r\big\|_W + \big\|w_r - w_s\big\|_W + \big\|w_s - \iota(v_s)\big\|_W \\
&< \frac{1}{r} + \frac{\varepsilon}{3} + \frac{1}{s} \\
&< \frac{2}{N} + \frac{\varepsilon}{3} \\
&< \varepsilon.
\end{aligned}$$

Resulta da relação (6.5) que as sucessões $(w_n)_{n\in\mathbb{N}}$ são ambas convergentes ou ambas divergentes e que, no primeiro caso, têm o mesmo limite. Mas então deduz-se da relação (6.4) que $\left[(v_n)_{n\in\mathbb{N}}\right] \in W$ é o limite da sucessão $(w_n)_{n\in\mathbb{N}}$.

Sejam agora $\widetilde{W}$ e $\tilde{\iota}$ como no enunciado. Considere-se a aplicação linear

$$\begin{array}{rccc} \tilde{\iota}\circ\iota^{-1}\colon & \iota(V) & \longrightarrow & \widetilde{W} \\ & \iota(v) & \mapsto & \tilde{\iota}(v); \end{array}$$

esta definição faz sentido, visto que ι é injectiva. Seja $\psi\colon W \longrightarrow \widetilde{W}$ um prolongamento linear e contínuo de $\tilde{\iota}\circ\iota^{-1}$ a W; este prolongamento existe e é único pela proposição 6.5 e porque $W = \overline{\iota(V)}$. É óbvio que $\psi\circ\iota = \tilde{\iota}$ e tudo o que falta demonstrar é que ψ é uma isometria e uma bijecção. Se $w \in W$ então existe alguma sucessão $(v_n)_{n\in\mathbb{N}}$ de elementos de V tal que $w = \lim_{n\in\mathbb{N}} \iota(v_n)$ e então

$$\begin{aligned} \|\psi(w)\|_{\widetilde{W}} &= \left\|\psi\left(\lim_{n\in\mathbb{N}} \iota(v_n)\right)\right\|_{\widetilde{W}} \\ &= \lim_{n\in\mathbb{N}} \|\psi(\iota(v_n))\|_{\widetilde{W}} \\ &= \lim_{n\in\mathbb{N}} \left\|\tilde{\iota}(v_n)\right\|_{\widetilde{W}} \\ &= \lim_{n\in\mathbb{N}} \|v_n\| \text{ (pois } \tilde{\iota} \text{ é uma isometria)} \\ &= \|w\|. \end{aligned}$$

Como ψ é uma isometria, ψ e ψ^{-1} são funções uniformemente contínuas, pelo que $\psi(W)$ é completo; logo, é um subespaço fechado de $\widetilde{W}$, pela proposição 6.2. Mas, como se tem $\psi\circ\iota = \tilde{\iota}$, $\psi(W) \supset \tilde{\iota}(V)$ e, portanto, visto que $\psi(W)$ é fechado e $\tilde{\iota}(V)$ é denso,

$$\psi(W) \supset \overline{\tilde{\iota}(V)} = \widetilde{W}.$$

■

6.2 Espaços de Hilbert

Nesta secção e na próxima, os espaços de Banach pouco aparecem. Elas servem apenas para introduzir alguns problemas que serão resolvidos empregando teoremas relativos àqueles espaços.

DEFINIÇÃO 6.4 Seja V um espaço vectorial real. Diz-se que uma função

$$\begin{array}{rccc} \langle\cdot,\cdot\rangle\colon & V\times V & \longrightarrow & \mathbb{R} \\ & (v,w) & \mapsto & \langle v,w\rangle \end{array}$$

é um *produto escalar* se

1. for bilinear;
2. simétrica (i. e. $(\forall v, w \in V) : \langle v, w\rangle = \langle w, v\rangle$);
3. definida positiva (i. e. $(\forall v \in V) : \langle v, v\rangle \geqslant 0$ e, além disso, tem-se $\langle v, v\rangle = 0$ se e só se $v = 0$).

EXEMPLO 6.4 Se $n \in \mathbb{N}$, o produto escalar usual em $\mathbb{R}^n$ define-se por

$$\begin{array}{ccc} \mathbb{R}^n \times \mathbb{R}^n & \longrightarrow & \mathbb{R} \\ \big((x_1, \ldots, x_n), (y_1, \ldots, y_n)\big) & \mapsto & \sum_{k=1}^{n} x_k y_k. \end{array}$$

EXEMPLO 6.5 Seja $(X, \mathcal{A}, m)$ um espaço de medida. Então, pela desigualdade de Hölder, se $f, g \in L^2(X)$, $fg \in L^1(X)$. Faz então sentido definir o seguinte produto escalar:

$$\begin{array}{ccc} L^2(X) \times L^2(X) & \longrightarrow & \mathbb{R} \\ (f, g) & \mapsto & \int_X fg\, dm. \end{array}$$

DEFINIÇÃO 6.5 Um *espaço pré-hilbertiano* é um par ordenado $(V, \langle\cdot,\cdot\rangle)$, onde V é um espaço vectorial real e $\langle\cdot,\cdot\rangle$ é um produto escalar de V em $\mathbb{R}$.

PROPOSIÇÃO 6.6 (DESIGUALDADE DE CAUCHY-SCHWARZ) *Se $(V, \langle\cdot,\cdot\rangle)$ for um espaço pré-hilbertiano, então*

$$(\forall v, w \in V) : |\langle v, w\rangle| \leqslant \sqrt{\langle v, v\rangle}\sqrt{\langle w, w\rangle}.$$

DEMONSTRAÇÃO: Basta fazer a domonstração para $v, w \neq 0$, pois o enunciado é trivial se $v = 0$ ou $w = 0$.

Sejam $v, w \in V \setminus \{0\}$ e sejam $\lambda, \mu \in \mathbb{R}$. Então

$$\begin{aligned} 0 &\leqslant \langle \lambda v + \mu w, \lambda v + \mu w\rangle \\ &= \lambda^2\langle v, v\rangle + 2\lambda\mu\langle v, w\rangle + \mu^2\langle w, w\rangle \\ &= \left(\lambda\sqrt{\langle v, v\rangle} + \mu\sqrt{\langle w, w\rangle}\right)^2 + \lambda\mu\left(\langle v, w\rangle - \sqrt{\langle v, v\rangle}\sqrt{\langle w, w\rangle}\right). \end{aligned}$$

Em particular, se se tomar $\lambda = \pm\sqrt{\langle w, w\rangle}$ e $\mu = \mp\sqrt{\langle v, v\rangle}$, tem-se

$$0 \leqslant -\sqrt{\langle w, w\rangle}\sqrt{\langle v, v\rangle}\left(\langle v, w\rangle - \sqrt{\langle v, v\rangle}\sqrt{\langle w, w\rangle}\right).$$

Como $\sqrt{\langle v, v\rangle}, \sqrt{\langle w, w\rangle} > 0$, resulta daqui que

$$\langle v, w\rangle \leqslant \sqrt{\langle v, v\rangle}\sqrt{\langle w, w\rangle}.$$

Visto que isto tem lugar para cada v e cada w,

$$\begin{aligned} |\langle v, w\rangle| &= \pm\langle v, w\rangle \\ &= \langle \pm v, w\rangle \\ &\leqslant \sqrt{\langle \pm v, \pm v\rangle}\sqrt{\langle w, w\rangle} \\ &= \sqrt{\langle v, v\rangle}\sqrt{\langle w, w\rangle}. \end{aligned}$$

■

Resulta da desigualdade de Cauchy-Schwarz que se $(V, \langle\cdot,\cdot\rangle)$ for um espaço pré-hilbertiano, a função

$$\begin{array}{rccc} \|\cdot\|\colon & V & \longrightarrow & \mathbb{R} \\ & v & \mapsto & \sqrt{\langle v, v\rangle} \end{array}$$

é uma norma. A única propriedade que não é imediata é a desigualdade triangular, mas se $v, w \in V$, tem-se

$$\begin{aligned} \|v+w\|^2 &= \langle v+w, v+w\rangle \\ &= \langle v, v\rangle + 2\langle v, w\rangle + \langle w, w\rangle \\ &\leqslant \|v\|^2 + 2\|v\|\|w\| + \|w\|^2 \\ &= \big(\|v\| + \|w\|\big)^2. \end{aligned}$$

Logo, cada espaço pré-hilbertiano tem uma estrutura natural de espaço vectorial normado. Alguns destes espaços vectoriais normados serão espaços de Banach, i. e. serão completos.

DEFINIÇÃO 6.6 Designa-se por *espaço de Hilbert* um espaço pré-hilbertiano completo.

Num espaço pré-hilbertiano, o produto escalar permite definir o conceito de vectores ortogonais.

DEFINIÇÃO 6.7 Seja V um espaço pré-hilbertiano. Diz-se que dois vectores $v, w \in V$ são *ortogonais*[2] se $\langle v, w\rangle = 0$. Diz-se que uma família $(v_i)_{i\in I}$ de vectores de um espaço pré-hilbertiano é *ortogonal* se, para quaisquer dois elementos distintos $i, j \in I$, v_i e v_j forem perpendiculares. Se, além disso, todos os elementos da família tiverem norma 1, diz-se que a família é *ortonormal*.

[2]Tembém se emprega aqui o termo *perpendiculares*.

PROPOSIÇÃO 6.7 *Sejam V um espaço pré-hilbertiano, W um subespaço vectorial de V e $v \in V$. Se existir algum vector $v_0 \in W$ tal que $v - v_0$ seja perpendicular a qualquer elemento de W, então a função*

$$\begin{array}{ccc} W & \longrightarrow & \mathbb{R}_+ \\ w & \mapsto & \|w - v\| \end{array}$$

atinge o seu valor mínimo no ponto v_0 e apenas nesse ponto.

DEMONSTRAÇÃO: Basta ver que se $w \in W$, então

$$\|w - v\|^2 = \|(w - v_0) - (v - v_0)\|^2 = \|w - v_0\|^2 + \|v - v_0\|^2,$$

pois $w - v_0 \in W$ e, portanto, $\langle w - v_0, v - v_0 \rangle = 0$. ■

Pode-se provar (veja-se [7, §I.2] ou [14, cap. 4]) que se V for um espaço de Hilbert e W for fechado, então existe necessariamente um vector v_0 nas condições do enunciado.

6.3 Séries de Fourier

Vai-se representar por $C(T)$ o espaço das funções contínuas de $\mathbb{R}$ em $\mathbb{R}$ que são periódicas de período 2π. Esta notação tem origem no facto de o conjunto das funções de $\mathbb{R}$ em $\mathbb{R}$ periódicas de período 2π se identificar naturalmente com o conjunto de todas as funções de $\mathbb{R}/(2\pi\mathbb{Z})$ em $\mathbb{R}$ e de T ser a notação usualmente empregue para representar o quociente $\mathbb{R}/(2\pi\mathbb{Z})$. Além disso representa-se por $L^p(T)$ ($p \in [1, +\infty]$) o conjunto das classes de equivalência das funções de $\mathbb{R}$ em $\mathbb{R}$ periódicas de período 2π tais que $f|_{[-\pi,\pi]} \in L^p([-\pi, \pi])$, relativamente à relação de equivalência usual (i. e. duas funções são equivalentes quando e só quando são iguais q. s.). No que se segue, somente irão intervir os espaços $L^1(T)$ e $L^2(T)$. A norma que se vai considerar no primeiro destes espaços será a norma $\|\cdot\|_1$, definida por

$$\|f\|_1 = \frac{1}{2\pi} \int_{-\pi}^{\pi} |f(x)|\, dx.$$

Naturalmente, $L^1(T)$ é um espaço de Banach, pois a função

$$\begin{array}{ccc} L^1(T) & \longrightarrow & L^1([-\pi, \pi]) \\ f & \mapsto & \frac{1}{2\pi} f|_{[-\pi,\pi]} \end{array}$$

é uma isometria sobrejectiva. Analogamente, o espaço $L^2(T)$ é um espaço de Hilbert relativamente ao produto escalar definido por

$$\langle f, g\rangle = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)g(x)\,dx.$$

DEFINIÇÃO 6.8 Se $f \in L^2(T)$, designa-se por *série de Fourier* da função f a série de funções

$$\frac{a_0}{2} + \sum_{n=1}^{+\infty}\big(a_n\cos(nt) + b_n\operatorname{sen}(nt)\big), \tag{6.6}$$

sendo

$$(\forall n \in \mathbb{Z}_+) : a_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(t)\cos(nt)\,dt$$

e

$$(\forall n \in \mathbb{N}) : b_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(t)\operatorname{sen}(nt)\,dt.$$

Naturalmente, os números da forma a_n ($n \in \mathbb{Z}_+$) e da forma b_n ($n \in \mathbb{N}$) dependem de f, pelo que talvez fosse preferível escrevê-los sob a forma $a_n(f)$ e $b_n(f)$ respectivamente mas, para simplificar as notações, isso só será feito quando houver mais de uma função envolvida. No entanto, convém observar que se se encararem as sucessões $(a_n)_{n\in\mathbb{Z}_+}$ e $(b_n)_{n\in\mathbb{N}}$ como sucessões de funções de $L^1(T)$ em $\mathbb{R}$, então tratam-se de aplicações lineares contínuas.

Os coeficientes $(a_n)_{n\in\mathbb{Z}_+}$ e $(b_n)_{n\in\mathbb{N}}$ da definição anterior estão relacionados com a proposição 6.7. De facto, um cálculo simples mostra que as funções de $\mathbb{R}$ em $\mathbb{R}$ definidas por

$$t \mapsto \cos(nt) \quad (n \in \mathbb{Z}_+)$$

e

$$t \mapsto \operatorname{sen}(nt) \quad (n \in \mathbb{N})$$

formam uma família ortogonal de elementos de $L^2(T)$; além disso, todos os elementos desta família têm norma $1/2$ excepto, naturalmente, a função constante que toma sempre o valor 1, cuja norma é igual a 1. Seja $\mathcal{T}$ o conjunto dos elementos de $L^2(T)$ que se podem exprimir como a soma de uma série convergente

$$\alpha_0 + \sum_{n=1}^{+\infty}\big(\alpha_n\cos(nt) + \beta_n\operatorname{sen}(nt)\big);$$

as séries deste tipo designam-se por *séries trigonométricas*. A proposição 6.7 afirma então que, caso a série (6.6) seja convergente (em $L^2(T)$), a sua soma é a série trigonométrica mais próxima da função f.

Em 1822, no seu livro *Théorie analytique de la chaleur*, Fourier afirmou que *qualquer* função periódica de $\mathbb{R}$ em $\mathbb{R}$ de período 2π se pode exprimir como soma da sua série de Fourier, i. e. que a série (6.6) converge pontualmente para a função f. Inicialmente, esta ideia foi recebida com bastante cepticismo mas, à medida que o século XIX ia avançando, foi-se constatando que não se encontravam excepções a esta regra. Dirichlet demonstrou que a hipótese de Fourier é válida em muitos casos. Por exemplo, resulta dos trabalhos de Dirichlet que se f for contínua e se, para algum subconjunto finito F de $[-\pi,\pi]$, f for derivável em todos os pontos de $[-\pi,\pi]\setminus F$ e se a derivada for uma função contínua e limitada, então a hipótese é verdadeira e a convergência é mesmo uniforme nesse caso; veja-se [10, cap. 15] uma demonstração deste resultado. Uma questão que se põe naturalmente neste contexto é a seguinte: *haverá alguma função contínua de $\mathbb{R}$ em $\mathbb{R}$ periódica de período 2π que não seja a soma da sua série de Fourier?*

Vejamos outra questão, que tem origem no seguinte enunciado:

TEOREMA 6.2 (LEMA DE RIEMANN-LEBESGUE) *Se $f \in L^1(T)$, então*

$$\lim_{n\in\mathbb{Z}_+} a_n = \lim_{n\in\mathbb{N}} b_n = 0.$$

DEMONSTRAÇÃO: Vai-se demonstrar somente que $\lim_{n\in\mathbb{Z}_+} a_n = 0$; a demonstração de que $\lim_{n\in\mathbb{N}} b_n = 0$ é análoga. Quer-se então provar que

$$\lim_{n\in\mathbb{Z}_+} \int_{-\pi}^{\pi} f(t)\cos(nt)\,dt = 0 \tag{6.7}$$

Sejam $a, b \in [-\pi,\pi]$ tais que $a<b$. Vai-se provar que se tem (6.7) para $f=\chi_{[a,b]}$. Isto resulta de um cálculo directo: se $n\in\mathbb{N}$, então

$$\begin{aligned}
\left|\int_{-\pi}^{\pi} \chi_{[a,b]}(t)\cos(nt)\,dt\right| &= \left|\int_a^b \cos(nt)\,dt\right| \\
&= \left|\left[\frac{\operatorname{sen}(nt)}{n}\right]_{t=a}^{t=b}\right| \\
&= \left|\frac{\operatorname{sen}(nb)-\operatorname{sen}(na)}{n}\right| \\
&\leqslant \frac{1}{n}.
\end{aligned}$$

Resulta da validade do lema para as funções do tipo $\chi_{[a,b]}$ e da linearidade de a_n ($n \in \mathbb{Z}_+$) que se f for combinação linear de funções do tipo $\chi_{[a,b]}$, então continua-se a ter (6.7).

Suponha-se agora que f é contínua. Seja $\varepsilon > 0$. Visto que f é contínua, a sua restrição a $[-\pi, \pi]$ é uniformemente contínua. Existe então algum $\delta > 0$ tal que

$$(\forall x, y \in [-\pi, \pi]) : |x - y| < \delta \Longrightarrow \big|f(x) - f(y)\big| < \varepsilon.$$

Seja $-\pi = \alpha_0 < \alpha_1 < \cdots < \alpha_n = \pi$ uma partição do intervalo $[-\pi, \pi]$ tal que cada intervalo da partição tenha comprimento inferior a δ. Seja

$$g = \sum_{k=1}^{n} f(\alpha_k)\chi_{[\alpha_{k-1},\alpha_k]}.$$

Então $\sup|f - g| < \varepsilon$. Por outro lado, se n for suficientemente grande $|a_n(g)| < \varepsilon$, pelo que,

$$\begin{aligned}\big|a_n(f)\big| &= \big|a_n(f-g) + a_n(g)\big| \\ &\leqslant \big|a_n(f-g)\big| + \big|a_n(g)\big| \\ &\leqslant 2\varepsilon,\end{aligned}$$

pois

$$\begin{aligned}\big|a_n(f-g)\big| &= \left|\frac{1}{2\pi}\int_{-\pi}^{\pi}(f(t) - g(t))\cos(nt)\,dt\right| \\ &\leqslant \frac{1}{2\pi}\int_{-\pi}^{\pi}|f(t) - g(t)|\,dt \\ &\leqslant \frac{1}{2\pi}\int_{-\pi}^{\pi}\varepsilon\,dt \\ &\leqslant \varepsilon.\end{aligned}$$

A validade do teorema no caso geral decorre do teorema 4.5, pois este implica que que $C(T)$ é denso em $L^1(T)$. ∎

Isto leva à seguinte pergunta: *dadas duas sucessões de números reais* $(a_n)_{n\in\mathbb{Z}_+}$ *e* $(b_n)_{n\in\mathbb{N}}$ *convergentes para* 0, *existe necessariamente alguma função* $f \in L^1(T)$ *tal que* $(\forall n \in \mathbb{Z}_+) : a_n(f) = a_n$ *e que* $(\forall n \in \mathbb{N}) : b_n(f) = b_n$?

Será visto que os próximos dois teoremas sobre espaços de Banach permitirão mostrar que a resposta a ambas as perguntas é negativa. Antes de se prosseguir, é conveniente reformular o conceito de séries de Fourier de uma maneira que leva a cálculos mais simples. Para isso, é necessário ver o que se entende por integral de uma função com valores em $\mathbb{C}$.

DEFINIÇÃO 6.9 Se f for uma função de X em $\mathbb{C}$, então diz-se que f é *integrável* se as funções $\operatorname{Re} f$ e $\operatorname{Im} f$ o forem. Nesse caso designa-se por integral de Lebesgue de f o número

$$\int_X f\,dm = \int_X \operatorname{Re} f\,dm + i\int_X \operatorname{Im} f\,dm.$$

É também conveniente alargar o conceito de série de maneira a englobar expressões da forma $\sum_{n\in\mathbb{Z}} c_n$. No contexto em que se está a trabalhar, dir-se-á que uma tal série converge se o limite $\lim_{N\in\mathbb{N}}\sum_{n=-N}^{N} c_n$ existir e, se for esse o caso, o limite em questão será designado por soma da série.

DEFINIÇÃO 6.10 Se $f\in L^2(T)$, designa-se por *série de Fourier* da função f a série de funções

$$\sum_{n\in\mathbb{Z}} \hat{f}(n)e^{int} \tag{6.8}$$

sendo

$$(\forall n\in\mathbb{Z}) : \hat{f}(n) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(t)e^{-int}\,dt.$$

Assim sendo, o conceito de série de Fourier de uma função $f\in L^2(T)$ foi definido de duas maneiras distintas. No entanto, a diferença é superficial, uma vez que, para cada $n\in\mathbb{N}$

$$\begin{aligned}\hat{f}(n) &= \frac{1}{2\pi}\int_{-\pi}^{\pi} f(t)e^{-int}\,dt\\ &= \frac{1}{2\pi}\int_{-\pi}^{\pi} f(t)\cos(nt)\,dt - \frac{i}{2\pi}\int_{-\pi}^{\pi} f(t)\operatorname{sen}(nt)\,dt\\ &= \frac{a_n - ib_n}{2},\end{aligned}$$

para cada $n\in -\mathbb{N}$

$$\hat{f}(n) = \frac{a_{-n}+ib_{-n}}{2} = \overline{\hat{f}(-n)} \tag{6.9}$$

e $\hat{f}(0) = {}^{a_0}\!/_2$. Por outro lado,

$$(\forall N\in\mathbb{N}) : \sum_{n=-N}^{N} \hat{f}(n)e^{int} = \frac{a_0}{2} + \sum_{n=1}^{N}\big(a_n\cos(nt) + b_n\operatorname{sen}(nt)\big).$$

A primeira das perguntas atrás formuladas pode então ser reformulada do seguinte modo: se $f \in C(T)$, tem-se ou não necessariamente:

$$(\forall t \in \mathbb{R}) : f(t) = \lim_{N \in \mathbb{N}} \sum_{n=-N}^{N} \hat{f}(n) e^{int}?$$

A segunda admite a seguinte reformulação: dada uma famíla $(z_n)_{n \in \mathbb{Z}}$ de números complexos tal que $(\forall n \in \mathbb{Z}) : \overline{z_{-n}} = z_n$ e que $\lim_{n \in \mathbb{N}} z_n = 0$, terá que existir alguma função $f \in L^1(T)$ tal que $(\forall n \in \mathbb{Z}) : c_n = \hat{f}(n)$?

Como já foi afirmado, irá ver-se que ambas as perguntas têm resposta negativa.

6.4 O teorema de Banach-Steinhaus

O próximo teorema sobre espaços de Banach que vai ser demonstrado vai ser o primeiro em cuja demonstração a completude daqueles espaços não vai ser empregue directamente, i. e. a demonstração não vai recorrer explicitamente ao facto de se estar a trabalhar num espaço no qual qualquer sucessão de Cauchy é convergente. Em vez disso, vai-se recorrer ao seguinte resultado:

TEOREMA 6.3 (TEOREMA DE BAIRE) *Num espaço métrico completo, qualquer família numerável de abertos densos tem intersecção densa.*

DEMONSTRAÇÃO: Sejam E o espaço métrico em questão. Se $x \in E$ e $r > 0$, vai-se representar por $B'(x,r)$ a bola fechada de centro x e raio r.

Seja $(A_n)_{n \in \mathbb{N}}$ uma família numerável de abertos densos. Se $x \in E$ e $\varepsilon > 0$, quer-se mostrar que algum elemento de $B(x,\varepsilon)$ pertence a todos os A_n simultaneamente.

Sejam $x_1 = x$ e $\varepsilon_1 = \varepsilon$. Como A_1 é um aberto denso, $B(x_1,\varepsilon) \cap A_1$ é um aberto não vazio. Seja $x_2 \in B(x_1,\varepsilon) \cap A_1$ e seja $\varepsilon_2 \in]0, \varepsilon_1/2]$ tal que $B'(x_2,\varepsilon_2) \subset B(x_1,\varepsilon_1) \cap A_1$. Como A_2 é um aberto denso, $B(x_2,\varepsilon_2) \cap A_2$ é um aberto não vazio. Seja $x_3 \in B(x_2,\varepsilon_2) \cap A_2$ e seja $\varepsilon_3 \in]0, \varepsilon_2/2]$ tal que $B'(x_3,\varepsilon_3) \subset B(x_2,\varepsilon_2) \cap A_2$. Continuando este processo, obtêm-se sucessões $(x_n)_{n \in \mathbb{N}}$ e $(\varepsilon_n)_{n \in \mathbb{N}}$ tais que, para cada $n \in \mathbb{N}$,

1. $0 < \varepsilon_n \leqslant \varepsilon/2^n$;
2. $B'(x_{n+1},\varepsilon_{n+1}) \subset B(x_n,\varepsilon_n) \cap A_n$.

Vejamos que $\bigcap_{n\in\mathbb{N}} B'(x_n,\varepsilon_n) \neq \emptyset$. Resulta de se ter

$$(\forall n \in \mathbb{N}) : \left|x_{n+1} - x_n\right| < \varepsilon_n \leqslant \frac{\varepsilon}{2^n}$$

que

$$(\forall m,n,p \in \mathbb{N}) : m \geqslant n \geqslant p \Longrightarrow \left|x_m - x_n\right| \leqslant \frac{\varepsilon}{2^{p-1}}$$

e, portanto, que a sucessão $(x_n)_{n\in\mathbb{N}}$ é de Cauchy. Então converge e o seu limite pertence a todas as bolas $B'(x_n,\varepsilon_n)$. Por outro lado,

$$\begin{aligned}\bigcap_{n\in\mathbb{N}} B'(x_n,\varepsilon_n) &\subset \bigcap_{n\in\mathbb{N}} \big(B(x_n,\varepsilon_n) \cap A_n\big)\\ &\subset B(x,\varepsilon) \cap \left(\bigcap_{n\in\mathbb{N}} A_n\right).\end{aligned}$$ ∎

Sejam V e W espaços vectoriais normados e $(f_i)_{i\in I}$ uma família de aplicações lineares contínuas de V em W. Considerem-se as seguintes possibilidades:

1. o conjunto $\{\, \|f_i\| \mid i \in I \,\}$ é majorado;
2. existe algum $v \in V$ tal que o conjunto $\{\, \|f_i(v)\| \mid i \in I \,\}$ não é majorado.

Estas duas possibilidades são mutualmente incompatíveis. De facto, se a primeira se verificar e se M for majorante do conjunto de todas as normas $\|f_i\|$ ($i \in I$), então, caso $v \in V$ e $i \in I$, $\left\|f_i(v)\right\| \leqslant \|f_i\|\|v\| \leqslant M\|v\|$, pelo que a segunda condição não se verifica.

Por outro lado, é possível que nenhuma das possibilidades tenha lugar. Por exemplo, seja $\big(c_{00}(\mathbb{N}), \|\cdot\|_\infty\big)$ o espaço do exemplo 5.6 e seja, para $n \in \mathbb{N}$,

$$\begin{aligned} f_n\colon \quad c_{00}(\mathbb{N}) &\longrightarrow \mathbb{R}\\ (a_n)_{n\in\mathbb{N}} &\mapsto \sum_{k=1}^{n} a_k. \end{aligned}$$

Então a primeira possibilidade não se verifica, pois $(\forall n \in \mathbb{N}) : \|f_n\| = n$. Mas a segunda também não se verifica, pois se $(a_n)_{n\in\mathbb{N}} \in c_{00}(\mathbb{N})$ e se $N \in \mathbb{N}$ for tal que $a_k = 0$ quando $k > N$, então todos os números da forma $f_k\big((a_n)_{n\in\mathbb{N}}\big)$ ($k \geqslant N$) são iguais a $\sum_{k=1}^{N} a_k$ e, portanto, o conjunto $\{\, |f_n\left((a_n)_{n\in\mathbb{N}}\right)| \mid n \in \mathbb{N} \,\}$ é majorado, visto que é finito.

O próximo teorema mostra que a situação muda totalmente se se estiver a supor que V é um espaço de Banach.

TEOREMA 6.4 (TEOREMA DE BANACH-STEINHAUS) *Sejam V um espaço de Banach, W um espaço vectorial normado e $(f_i)_{i\in I}$ uma família de aplicações lineares contínuas de V em W. Então tem lugar uma e só uma das seguintes possibilidades:*

1. *o conjunto $\{\, \|f_i\| \mid i \in I \,\}$ é majorado;*
2. *o conjunto $\{\, v \in V \mid \{\, \|f_i(v)\| \mid i \in I \,\}$ não é majorado $\}$ é denso em V.*

DEMONSTRAÇÃO: Para cada $n \in \mathbb{N}$, seja

$$A_n = \{\, v \in V \mid (\exists i \in I) : \|f_i(v)\| > n \,\}.$$

Este conjunto é aberto, pois é a reunião das imagens recíprocas de $]n, +\infty[$ pelas funções contínuas

$$\begin{array}{ccc} V & \longrightarrow & \mathbb{R}_+ \\ v & \longmapsto & \|f_i(v)\| \end{array}$$

($i \in I$). Há duas possibilidades: ou todos os A_n são densos em V ou há algum que não o é.

Caso algum A_n não seja uma parte densa de V, $V \setminus \overline{A_n}$ é um aberto não vazio de V; existem então $v \in V$ e $r > 0$ tais que $\overline{B(v,r)} \subset V \setminus \overline{A_n}$. Em particular,

$$\|w - v\| \leqslant r \Longrightarrow w \notin A_n \Longleftrightarrow (\forall i \in I) : \|f_i(w)\| \leqslant n.$$

Então, se $\|w\| \leqslant 1$, tem-se, para cada $i \in I$,

$$\begin{aligned} \|f_i(w)\| &= \frac{1}{r}\|f_i(rw)\| \\ &\leqslant \frac{1}{r}\|f_i(v + rw) - f_i(v)\| \\ &\leqslant \frac{1}{r}\big(\|f_i(v + rw)\| + \|f_i(v)\|\big) \\ &\leqslant \frac{2n}{r} \end{aligned}$$

e, portanto, verifica-se a primeira possibilidade do enunciado.

Caso todos os conjuntos A_n sejam densos em V, a sua intersecção também o é, pelo teorema de Baire. Mas se v pertencer à intersecção, resulta da definição de $(A_n)_{n\in\mathbb{N}}$ que o conjunto $\{\, \|f_i(v)\| \mid i \in I \,\}$ não é majorado. ∎

Pode-se ver em [7, §III.14] uma demonstração deste teorema que emprega directamente a definição de espaço completo e não o teorema de Baire.

Vai-se ver agora como deduzir do teorema de Banach-Steinhaus que há funções em $C(T)$ cuja série de Fourier não converge em todos os pontos, uma consequência que já é extraída no artigo onde este teorema foi publicado originalmente; veja-se [2]. Observe-se que se se definir, para cada $N \in \mathbb{N}$ e cada $t \in \mathbb{R}$,

$$K_N(t) = \sum_{n=-N}^{N} e^{int}, \tag{6.10}$$

então

$$\begin{aligned}\sum_{n=-N}^{N} \hat{f}(n)e^{int} &= \sum_{n=-N}^{N} \frac{1}{2\pi}\int_{-\pi}^{\pi} f(u)e^{-inu}e^{int}\,du \\ &= \frac{1}{2\pi}\int_{-\pi}^{\pi} f(u)\sum_{n=-N}^{N} e^{in(t-u)}\,du \\ &= \frac{1}{2\pi}\int_{-\pi}^{\pi} f(u)K_N(t-u)\,du.\end{aligned}$$

Quer-se então saber se se tem

$$(\forall f \in C(T))(\forall t \in \mathbb{R}) : f(t) = \lim_{n\in\mathbb{N}} \frac{1}{2\pi}\int_{-\pi}^{\pi} f(u)K_n(t-u)\,du.$$

Vai-se estudar um caso particular deste problema, nomeadamente o de saber se

$$(\forall f \in C(T)) : f(0) = \lim_{n\in\mathbb{N}} \frac{1}{2\pi}\int_{-\pi}^{\pi} f(u)K_n(-u)\,du$$

Para cada $n \in \mathbb{N}$, considere-se a aplicação linear

$$\begin{array}{rccc} F_n\colon & C(T) & \longrightarrow & \mathbb{R} \\ & f & \mapsto & \dfrac{1}{2\pi}\displaystyle\int_{-\pi}^{\pi} f(u)K_n(-u)\,du. \end{array}$$

Vai-se considerar em $C(T)$ a norma $\|\cdot\|_\infty$, ou seja, a norma definida por $\|f\|_\infty = \sup|f|$. Então $(C(T), \|\cdot\|_\infty)$ é um espaço de Banach, pois $C(T)$ é um subespaço fechado de $(L^\infty(\mathbb{R}), \|\cdot\|_\infty)$. Relativamente a esta norma, cada função F_n é contínua, pois se $\|f\|_\infty \leqslant 1$

$$\big|F_n(f)\big| \leqslant \frac{1}{2\pi}\int_{-\pi}^{\pi}\big|K_n(-u)\big|\,du = \|K_n\|_1.$$

Vejamos que a desigualdade anterior é, de facto, uma igualdade. A função K_n, apesar da maneira como está definida, só toma valores reais. Isto pode ser demonstrado provando que $(\forall t \in \mathbb{R}) : \overline{K_n(t)} = K_n(t)$ ou então aproveitando o facto de $K_n(t)$ se exprimir como a soma de números em progressão geométrica para provar que

$$(\forall t \in \mathbb{R}) : K_n(t) = \begin{cases} \frac{\operatorname{sen}((n+1/2)t)}{\operatorname{sen}(t/2)} & \text{se } t \notin 2\pi\mathbb{Z} \\ 2n+1 & \text{caso contrário.} \end{cases} \tag{6.11}$$

Seja

$$\begin{aligned} \varphi\colon \quad \mathbb{R} &\longrightarrow \mathbb{R} \\ t &\mapsto \begin{cases} 1 & \text{se } K_n(-t) \geqslant 0 \\ -1 & \text{caso contrário.} \end{cases} \end{aligned}$$

Então é possível encontrar uma sucessão $(\varphi_m)_{m\in\mathbb{N}}$ de elementos de $C(T)$ tais que $(\forall m \in \mathbb{N}) : \|\varphi_m\|_\infty \leqslant 1$ e que $\lim_{m\in\mathbb{N}} \|\varphi - \varphi_m\|_1 = 0$. Logo, decorre do teorema da convergência dominada que

$$\begin{aligned} \lim_{m\in\mathbb{N}} F_n(\varphi_m) &= \frac{1}{2\pi}\int_{-\pi}^{\pi} \varphi(u)K_n(-u)\,du \\ &= \frac{1}{2\pi}\int_{-\pi}^{\pi} \left|K_n(-u)\right|du \\ &= \|K_n\|_1. \end{aligned}$$

Vejamos agora que $\lim_{n\in\mathbb{N}} \|F_n\| = +\infty$ o que equivale, pelo que foi visto no parágrafo anterior, a afirmar que $\lim_{n\in\mathbb{N}} \|K_n\|_1 = +\infty$. Se $x \in \mathbb{R}$, então $|\operatorname{sen}(x)| \leqslant |x|$; resulta desta observação e da relação (6.11) que

$$\begin{aligned} \|K_n\|_1 &> \frac{2}{\pi}\int_0^{\pi} \left|\operatorname{sen}\left(n+\frac{1}{2}\right)t\right| \frac{dt}{t} \\ &= \frac{2}{\pi}\int_0^{(n+\frac{1}{2})\pi} |\operatorname{sen} t| \,\frac{dt}{t} \\ &> \frac{2}{\pi}\sum_{k=1}^{n}\frac{1}{k\pi}\int_{(k-1)\pi}^{k\pi} |\operatorname{sen} t|\,dt \\ &= \frac{4}{\pi}\sum_{k=1}^{n}\frac{1}{k}. \end{aligned}$$

Visto que $\lim_{n\in\mathbb{N}} \sum_{k=1}^{n} 1/k = +\infty$, está provado que $\lim_{n\in\mathbb{N}} \|K_n\|_1 = +\infty$.

Pode-se agora aplicar o teorema de Banach-Steinhaus. Já foi visto que o conjunto $\{\, \|F_n\| \mid n \in \mathbb{N} \,\}$ não é majorado. Logo, há funções $f \in$

$C(T)$ (de facto, todo um conjunto denso de funções) para as quais a sucessão $\left(|F_n(f)|\right)_{n\in\mathbb{N}}$ não é majorada. Em particular, não existe o limite $\lim_{n\in\mathbb{N}} F_n(f)$ e, portanto, não se pode ter $f(0) = \lim_{n\in\mathbb{N}} F_n(f)$. Isto é o mesmo que afirmar que $f(0)$ não é limite da sua série de Fourier no ponto 0.

Naturalmente, o ponto 0 nada tem de especial; o mesmo argumento aplica-se a qualquer ponto de $\mathbb{R}$.

Está então provado que é possível encontrar alguma função $f \in C(T)$ tal que, para cada $x \in \mathbb{R}$, a série de Fourier de f no ponto x não converge para $f(x)$. Uma demonstração mais elementar deste resultado pode ser vista em [10, cap. 18]. Por outro lado, pode-se provar que a série de Fourier de uma função $f \in C(T)$ no ponto x converge para $f(x)$ q. s.; veja-se [4].

Uma questão que se pode pôr agora é a seguinte: será possível obter uma função $f \in C(T)$ a partir da sua série de Fourier? A resposta (afirmativa) é dada pelo próximo teorema. Se $f \in C(T)$, se $N \in \mathbb{N}$ e se $t \in \mathbb{R}$, seja

$$s_N(f,t) = \sum_{n=-N}^{N} \hat{f}(n)e^{int}.$$

Já foi visto que não se tem necessariamente $f(t) = \lim_{N\in\mathbb{N}} s_N(f,t)$ para cada $t \in \mathbb{R}$.

TEOREMA 6.5 (TEOREMA DE FÉJER) *Se $f \in C(T)$ então a sucessão de funções $(\sigma_n(f,\cdot))_{n\in\mathbb{N}}$ definida por*

$$(\forall n \in \mathbb{N})(\forall t \in \mathbb{R}) : \sigma_n(f,t) = \frac{1}{n+1}\sum_{k=0}^{n} s_n(f,t)$$

converge uniformemente para a função f.

Veja-se [10, cap. 2] para a demonstração.

Visto que, pela relação (6.9), se tem $\overline{\hat{f}(n)} = \hat{f}(-n)$ quando $n \in -\mathbb{N}$ e como $\hat{f}(0) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(t)\,dt \in \mathbb{R}$, cada função do tipo $\sigma_n(f,\cdot)$ ($n \in \mathbb{N}$) é combinação linear da função constante que toma sempre o valor 1 e de funções do tipo $t \mapsto \cos(nt)$ e $t \mapsto \operatorname{sen}(nt)$ ($n \in \mathbb{N}$); tais combinações lineares designam-se por *polinómios trigonométricos*. Resulta imediatamente do teorema de Féjer que se tem o seguinte

COROLÁRIO 6.2 *Os polinómios trigonométricos formam uma parte densa de $C(T)$.*

Para uma demonstração deste corolário que não recorre ao teorema de Féjer, veja-se [14, cap. 4].

6.5 O teorema da aplicação aberta

DEFINIÇÃO 6.11 Sejam E e F espaços métricos. Diz-se que uma função f de E em F é *aberta* se, para cada aberto A de E, $f(A)$ for um aberto de F.

EXEMPLO 6.6 A função id: $\left(C([0,1]), \|\cdot\|_\infty\right) \longrightarrow \left(C([0,1]), \|\cdot\|_1\right)$ não é aberta pois, como foi visto na primeira alínea do exemplo 5.10, a imagem da bola aberta unitária de $\left(C([0,1]), \|\cdot\|_\infty\right)$ não é um aberto de $\left(C([0,1]), \|\cdot\|_1\right)$.

EXEMPLO 6.7 A projecção

$$\begin{array}{rccc} \pi\colon & \mathbb{R}^2 & \longrightarrow & \mathbb{R} \\ & (x,y) & \mapsto & x \end{array}$$

é aberta. De facto, qualquer aberto A de $\mathbb{R}^2$ pode ser escrito sob a forma

$$A = \bigcup_{i\in I}]a_{i1}, a_{i2}[\times]b_{i1}, b_{i2}[,$$

pelo que $\pi(A) = \bigcup_{i\in I}]a_{i1}, a_{i2}[$, que é um aberto de $\mathbb{R}$.

Na demonstração do próximo teorema vai ser empregue o teorema de Baire, mas sob uma forma ligeiramente distinta daquela que foi empregue na demonstração do teorema de Banach-Steinhaus. O teorema de Baire afirma que, num espaço métrico completo E, qualquer sucessão de abertos densos tem interesecção densa. Seja agora $(F_n)_{n\in\mathbb{N}}$ uma sucessão de fechados de E com interior vazio. Então, para cada $n \in \mathbb{N}$, afirmar que F_n é fechado equivale a afirmar que $F_n^{\complement}$ é aberto e afirmar que F_n tem interior vazio equivale a afirmar que $F_n^{\complement}$ é denso. Mas então $\left(F_n^{\complement}\right)_{n\in\mathbb{N}}$ é uma sucessão de abertos densos, pelo que a sua intersecção é densa, o que equivale a afirmar que $\bigcup_{n\in\mathbb{N}} F_n$ tem interior vazio. Está então visto que resulta do teorema de Baire que, num espaço métrico completo, dada uma sucessão de fechados com interior vazio, a sua reunião também tem interior vazio. De facto, vê-se facilmente que todo o raciocínio empregue é válido nos dois sentidos, i. e. que, dado um espaço métrico, as condições

- qualquer sucessão de abertos densos tem intersecção densa;
- a reunião de uma sucessão de fechados com interior vazio tem interior vazio

são equivalentes.

TEOREMA 6.6 (TEOREMA DA APLICAÇÃO ABERTA) *Qualquer aplicação linear contínua e sobrejectiva de um espaço de Banach noutro espaço de Banach é aberta.*

DEMONSTRAÇÃO: Sejam B e B' as bolas abertas unitárias de V e de W respectivamente. Se se mostrar que $f(B) \supset rB'$ para algum $r > 0$, o teorema estará demonstrado, pois resulta daqui que f envia vizinhanças de 0 em vizinhanças de 0, o que, por sua vez, implica que f é aberta.

Como f é sobrejectiva, tem-se

$$W = f(V) = f\left(\bigcup_{k\in\mathbb{N}} kB\right) = \bigcup_{k\in\mathbb{N}} f(kB) \subset \bigcup_{k\in\mathbb{N}} \overline{f(kB)} \subset W,$$

pelo que

$$W = \bigcup_{k\in\mathbb{N}} \overline{f(kB)}.$$

Resulta então do teorema de Baire e do que foi observado antes da demonstração do teorema que os conjuntos da forma $\overline{f(kB)}$ não podem ter todos o interior vazio.[3] Fixe-se então algum $k \in \mathbb{N}$ tal que $\overline{f(kB)}$ tenha interior não vazio e seja A um aberto não vazio contido em $\overline{f(kB)}$.

Seja $w_0 \in A$ e seja $r > 0$ tal que $\overline{B(w_0, r)} \subset A$. Se $w \in \overline{B(0, r)}$, então $w_0, w_0 + w \in A \subset \overline{f(kB)}$ e há, portanto, sucessões $(v'_n)_{n\in\mathbb{N}}$ e $(v''_n)_{n\in\mathbb{N}}$ de elementos de kB tais que

$$\lim_{n\in\mathbb{N}} v'_n = w_0 \quad \text{e que} \quad \lim_{n\in\mathbb{N}} v''_n = w_0 + w.$$

Para cada $n \in \mathbb{N}$, seja $v_n = v'_n - v''_n$. Então $(\forall n \in \mathbb{N}) : \|v_n\| < 2k$ e $\lim_{n\in\mathbb{N}} f(v_n) = w$. Isto tem lugar para cada $w \in W$ tal que $\|w\| \leqslant r$, o que equivale a afirmar que se $\varepsilon > 0$ e se $w \in W$ for tal que $\|w\| \leqslant r$, então existe algum $v \in V$ tal que $\|v\| \leqslant 2k$ e que $\|f(v) - w\| \leqslant \varepsilon$.

Sejam agora $\varepsilon > 0$ e $w \in W \setminus \{0\}$. Como $\left\|\frac{r}{\|w\|}w\right\| = r$, existe algum $v' \in V$ tal que $\|v'\| \leqslant 2k$ e que $\left\|f(v') - \frac{r}{\|w\|}w\right\| \leqslant \frac{r}{\|w\|}\varepsilon$. Seja $v \in V$ tal que $v' = \frac{r}{\|w\|}v$; então

$$\left\|f\left(\frac{r}{\|w\|}v\right) - \frac{r}{\|w\|}w\right\| < \frac{r}{\|w\|}\varepsilon \iff \|f(v) - w\| \leqslant \varepsilon$$

[3]De facto, não é difícil provar que, dado $k \in \mathbb{N}$, a função $p_k : W \longrightarrow W$ definida por $p_k(v) = kv$ é um homeomorfismo que envia $\overline{f(B)}$ em $\overline{f(kB)}$. Consequentemente, afirmar que *algum* conjunto da forma $\overline{f(kB)}$ tem interior não vazio é o mesmo que afirmar que *todos* os conjuntos da forma $\overline{f(kB)}$ têm interior não vazio.

e $\|v\| \leqslant \frac{2k\|w\|}{r}$. Isto prova que, para cada $\varepsilon > 0$ e cada $w \in W \setminus \{0\}$, existe algum $v \in V$ tal que $\|v\| \leqslant \frac{2k\|w\|}{r}$ e que $\|f(v) - w\| \leqslant \varepsilon$ e, obviamente, isto continua a ser verdade se $w = 0$. Logo, se $\varepsilon > 0$ e se $c = r/(2k)$, tem-se

$$(\forall w \in W)(\exists v \in V) : \|f(v) - w\| \leqslant \varepsilon \wedge \|v\| \leqslant c^{-1}\|w\|. \tag{6.12}$$

Seja $w \in W$ tal que $\|w\| < c/2$. Sabe-se, por (6.12), que existe algum $v_1 \in V$ tal que $\|v_1\| < 1/2$ e que

$$\big\|f(v_1) - w\big\| \leqslant \frac{c}{4}.$$

Aplicando (6.12) novamente (desta vez ao vector $w - f(v_1)$), sabe-se que existe algum vector $v_2 \in V$ tal que $\|v_2\| < 1/4$ e que

$$\big\|f(v_2) + f(v_1) - w\big\| \leqslant \frac{c}{8}.$$

Pelo mesmo motivo (considerando desta vez o vector $w - f(v_1) - f(v_2)$), existe algum vector $v_3 \in V$ tal que $\|v_3\| < 1/8$ e que

$$\big\|f(v_3) + f(v_2) + f(v_1) - w\big\| \leqslant \frac{c}{16}$$

e assim sucessivamente. Por isto, e pela linearidade de f, existe alguma sucessão $(v_n)_{n\in\mathbb{N}}$ de elementos de V tal que

$$(\forall n \in \mathbb{N} \setminus \{1\}) : \big\|v_n\big\| < \frac{1}{2^n} \tag{6.13}$$

e que

$$(\forall n \in \mathbb{N}) : \left\|f\left(\sum_{k=1}^{n} v_k\right) - w\right\| \leqslant \frac{c}{2^{n+1}}. \tag{6.14}$$

Por (6.13), a série $\sum_{n=1}^{\infty} v_n$ é absolutamente convergente e, portanto, convergente, pela proposição 6.1; seja v a sua soma. Decorre de (6.14) que $\big\|f(v) - w\big\| = 0$, i. e. que $f(v) = w$. Por outro lado,

$$\|v\| \leqslant \sum_{n=1}^{\infty} \|v_n\| < \sum_{n=1}^{\infty} \frac{1}{2^n} = 1.$$

Como tudo o que se supôs relativamente a W foi que $\|w\| < c/2$, está então provado que

$$f(B) \supset \frac{c}{2}B'.$$

∎

COROLÁRIO 6.3 *Se V e W são espaços de Banach e se f é uma bijecção linear contínua de V em W, então f é um homeomorfismo e existe algum $M > 0$ tal que*

$$(\forall v \in V) : \|f(v)\| \geqslant M\|v\|.$$

DEMONSTRAÇÃO: Para provar que f é um homeomorfismo, falta apenas provar que f^{-1} é contínua, o que equivale a afirmar que se A é um aberto de V, então $\left(f^{-1}\right)^{-1}(A)$ é um aberto de W. Mas $\left(f^{-1}\right)^{-1}(A) = f(A)$, que é efectivamente um aberto de W, pelo teorema da aplicação aberta.

É claro que f^{-1} é uma aplicação linear. Então, para cada $w \in W$, $\|f^{-1}(w)\| \leqslant \|f^{-1}\|\|w\|$ e então, se $v \in V$

$$\|v\| = \|f^{-1}(f(v))\| \leqslant \|f^{-1}\|\|f(v)\| \Longleftrightarrow \|f(v)\| \geqslant \|f^{-1}\|^{-1}\|v\|. \quad \blacksquare$$

EXEMPLO 6.8 A bijecção $\mathrm{id}\colon \left(C([0,1]), \|\cdot\|_\infty\right) \longrightarrow \left(C([0,1]), \|\cdot\|_1\right)$ é linear e contínua, mas não é um homeomorfismo, pois ser um homeomorfismo equivale a afirmar que a norma do supremo e a norma do integral são equivalentes, o que não é o caso (pelo que foi visto no exemplo 5.10 na página 117).

Vai-se agora responder à pergunta enunciada na página 140: dadas duas sucessões de números reais $(a_n)_{n\in\mathbb{Z}_+}$ e $(b_n)_{n\in\mathbb{N}}$ convergentes para 0, terá que existir uma função $f \in L^1(T)$ tal que $(\forall n \in \mathbb{Z}_+) : a_n(f) = a_n$ e que $(\forall n \in \mathbb{N}) : b_n(f) = b_n$? Considere-se a função

$$\begin{array}{rccc} \Psi\colon & L^1(T) & \longrightarrow & c_0(\mathbb{N}) \\ & f & \mapsto & (a_0, b_1, a_1, b_2, a_2, \ldots). \end{array}$$

Pelo lema de Riemann-Lebesgue, esta definição faz sentido, i. e. se $f \in L^1(T)$, então efectivamente $\Psi(f)$ é uma sucessão convergente para 0. Quer-se provar que Ψ não é sobrejectiva. Claramente, Ψ é uma aplicação linear. Se se considerar em $c_0(\mathbb{N})$ a norma $\|\cdot\|_\infty$, Ψ é contínua, pois se $f \in L^1([0,1])$ for tal que $\|f\|_1 = 1$, então

$$(\forall n \in \mathbb{Z}_+) : |a_n| \leqslant \frac{1}{\pi}\int_{-\pi}^{\pi} |f(t)|\,dt = \|f\|_1 = 2$$

e

$$(\forall n \in \mathbb{N}) : |b_n| \leqslant \frac{1}{\pi}\int_{-\pi}^{\pi} |f(t)|\,dt = \|f\|_1 = 2.$$

Observe-se que Ψ é injectiva. De facto, seja $f \in L^1(T)$ tal que a sua série de Fourier tenha todos os coeficientes iguais a 0; quer-se provar que $f = 0$, i. e. que $f(x) = 0$ q. s. Afirmar que a série de Fourier de f tem todos os coeficientes iguais a 0 é o mesmo que afirmar que se $P\colon \mathbb{R} \longrightarrow \mathbb{R}$ for um polinómio trigonométrico, então $\int_{-\pi}^{\pi} f(x)P(x)\,dx = 0$. Seja agora $g \in C(T)$. Então, pelo corolário 6.2, existe, para cada $\varepsilon > 0$, algum polinómio trigonométrico P tal que $\sup|f - P| < \varepsilon$, pelo que

$$\begin{aligned}
&\left|\frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)g(x)\,dx\right| \leqslant \\
&\leqslant \left|\frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)(g(x) - P(x))\,dx\right| + \left|\frac{1}{2\pi}\overbrace{\int_{-\pi}^{\pi} f(x)P(x)\,dx}^{=0}\right| \\
&\leqslant \|f\|_1\varepsilon;
\end{aligned}$$

como isto tem lugar para cada $\varepsilon > 0$, $\int_{-\pi}^{\pi} f(x)g(x)\,dx = 0$. Pode-se agora concluir que $f(x) = 0$ q. s. de várias maneiras. Uma delas consiste em observar que resulta do teorema 4.5 que $C(T)$ é denso em $L^1(T)$ e que, portanto, a função

$$\begin{aligned}
\varphi\colon \mathbb{R} &\longrightarrow \mathbb{R} \\
x &\mapsto \begin{cases} f(x)/|f(x)| & \text{se } f(x) \neq 0 \\ 0 & \text{caso contrário} \end{cases}
\end{aligned}$$

é limite (em $L^1(T)$) de uma sucessão $(g_n)_{n\in\mathbb{N}}$ de elementos de $C(T)$ e, como φ é limitada, pode-se mesmo tomar $(g_n)_{n\in\mathbb{N}}$ uniformemente limitada. Logo, pelo teorema da convergência dominada,

$$\begin{aligned}
\int_{-\pi}^{\pi} |f(x)|\,dx &= \int_{-\pi}^{\pi} f(x)\varphi(x)\,dx \\
&= \int_{-\pi}^{\pi} f(x) \lim_{n\in\mathbb{N}} g_n(x)\,dx \\
&= \lim_{n\in\mathbb{N}} \int_{-\pi}^{\pi} f(x)g(x)\,dx \\
&= 0.
\end{aligned}$$

Logo, f anula-se em quase todos os pontos de $[-\pi, \pi]$ e resulta então da periodicidade de f que $f(x) = 0$ q. s. Outra maneira de se provar isto, que não emprega o teorema 4.5, consiste em, dados $a, b \in [-\pi, \pi]$ tais que $a < b$, tomar-se uma sucessão $(g_n)_{n\in\mathbb{N}}$ de funções contínuas de $[-\pi,\pi]$ em $[0,1]$ tal que $(\forall x \in [-\pi,\pi]) : \lim_n g_n(x) = \chi_{[a,b]}(x)$; veja--se na figura 6.1 o gráfico de $\chi_{[a,b]}$ (a cheio) juntamente com o gráfico (a tracejado) de uma função contínua de $[-\pi,\pi]$ em $[0,1]$ próximo do de $\chi_{[a,b]}$. Se $x \in [-\pi,\pi]$, $f(x)\chi_{[a,b]}(x) = \lim_{n\in\mathbb{N}} f(x)g_n(x)$ e, pelo teorema da convergência dominada,

$$\int_a^b f(x)\,dx = \int_{-\pi}^{\pi} f(x)\chi_{[a,b]}(x)\,dx = \lim_{n\in\mathbb{N}} \int_{-\pi}^{\pi} f(x)g_n(x)\,dx = 0.$$

Resulta então do teorema 3.4 que $f(x) = 0$ q. s., como se quer demonstrar.

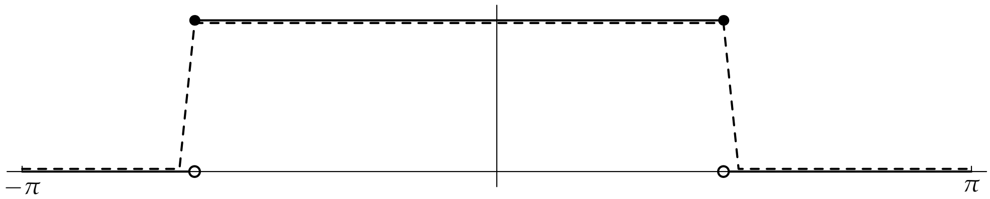

Figura 6.1: Gráficos de $\chi_{[a,b]}$ e de uma função contínua próxima

Sendo Ψ uma aplicação linear injectiva e contínua de um espaço de Banach noutro espaço de Banach, decorre do corolário 6.3 que, se fosse sobrejectiva, haveria algum $M > 0$ tal que

$$(\forall f \in L^1(T)) : \|\Psi(f)\|_\infty \geqslant M\|f\|_1.$$

Acontece que isto é impossível, pois se se definir, para cada $N \in \mathbb{N}$, a função K_N por (6.10), então um cálculo directo mostra que

$$\Psi(K_N) = (1,2,0,2,0,2,0,\ldots,2,0,0,0,0,\ldots)$$

(com o número 2 a surgir N vezes), pelo que $\|\Psi(K_N)\|_\infty = 2$, mas foi visto na página 146 que $\lim_{N\in\mathbb{N}} \|K_N\|_1 = +\infty$.

Se (E,d_E) e (F,d_F) forem espaços métricos e se f for uma função contínua de E em F, então o seu gráfico é um fechado de $E \times F$, se se considerar neste produto cartesiano a métrica d definida por $d\big((x_1,y_1),(x_2,y_2)\big) = \max\{d_E(x_1,y_1), d_F(x_2,y_2)\}$. O recíproco não é verdadeiro pois, por exemplo, a função

$$\begin{array}{rccc} \iota\colon & \mathbb{R} & \longrightarrow & \mathbb{R} \\ & x & \mapsto & \begin{cases} x^{-1} & \text{se } x \neq 0 \\ 0 & \text{se } x = 0 \end{cases} \end{array}$$

é descontínua, embora o seu gráfico (que está representado na figura 6.2) seja um fechado de $\mathbb{R}^2$.

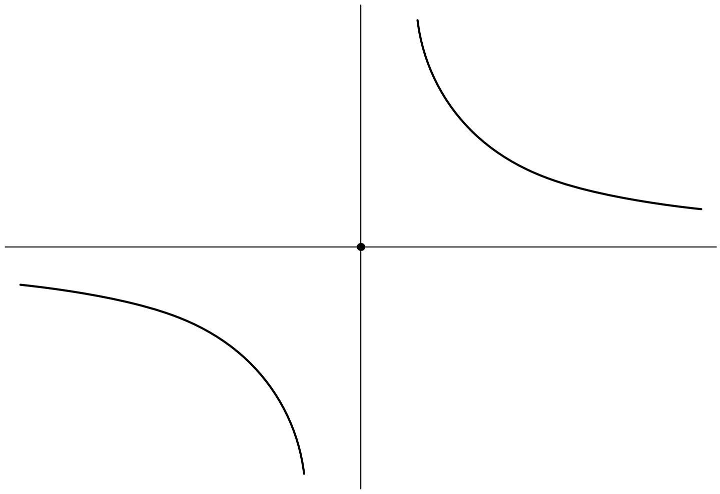

Figura 6.2: Gráfico da função ι, que é um fechado de $\mathbb{R}^2$

De facto, o recíproco nem sequer é verdadeiro, em geral, para aplicações lineares contínuas entre espaços vectoriais normados. Com efeito, se $C^1([0,1])$ for o subespaço de $(C([0,1]), \|\cdot\|_\infty)$ formado pelas funções de classe C^1, então a função

$$\begin{array}{rccc} D\colon & C^1([0,1]) & \longrightarrow & C([0,1]) \\ & f & \mapsto & f' \end{array}$$

é uma aplicação linear descontínua; basta ver que se se considerar, para cada $n \in \mathbb{N}$, a função

$$\begin{array}{rccc} \varphi_n\colon & [0,1] & \longrightarrow & \mathbb{R} \\ & x & \mapsto & x^n, \end{array}$$

então, para cada $n \in \mathbb{N}$, $\|\varphi_n\|_\infty = 1$ e $\|D(\varphi_n)\|_\infty = n$. Mas o seu gráfico é um fechado de $C^1([0,1]) \times C([0,1])$, pois se $(f_n, f_n')_{n\in\mathbb{N}}$ for uma sucessão de pontos do gráfico convergente para $(f,g) \in C^1([0,1]) \times C([0,1])$, então

$$(\forall n \in \mathbb{N})(\forall x \in [0,1]) : f_n(x) = f_n(0) + \int_0^x f_n'(t)\, dt$$

pelo que, se $x \in [0,1]$,

$$\begin{aligned} f(x) &= \lim_{n\in\mathbb{N}} f_n(x) \\ &= \lim_{n\in\mathbb{N}} \left(f_n(0) + \int_0^x f_n'(t)\, dt \right) \end{aligned}$$

$$\begin{aligned}
&= f(0) + \lim_{n\in\mathbb{N}} \int_0^x f_n'(t)\,dt \\
&= f(0) + \int_0^x \lim_{n\in\mathbb{N}} f_n'(t)\,dt \text{ (a convergência é uniforme)} \\
&= f(0) + \int_0^x g(t)\,dt.
\end{aligned}$$

Logo $g = f'$, i. e. (f, g) pertence ao gráfico de D.

TEOREMA 6.7 (TEOREMA DO GRÁFICO FECHADO) *Sejam E e F espaços de Banach e seja f uma aplicação linear de E em F cujo gráfico seja um fechado de $E \times F$. Então f é contínua.*

DEMONSTRAÇÃO: Seja $\mathrm{Gr}(f)$ o gráfico de f. Como f é uma aplicação linear, $\mathrm{Gr}(f)$ é um subespaço vectorial de $E \times F$, o qual, sendo fechado, é um espaço de Banach, pois resulta do facto de E e F serem espaços de Banach que $E \times F$ também é um espaço de Banach.

Considere-se a aplicação

$$\begin{array}{rccc} \psi\colon & E & \longrightarrow & \mathrm{Gr}(f) \\ & v & \mapsto & (v, f(v)), \end{array}$$

que é claramente uma bijecção linear. A sua inversa é a função

$$\begin{array}{ccc} \mathrm{Gr}(f) & \longrightarrow & E \\ (v, f(v)) & \mapsto & v, \end{array}$$

que é contínua; logo, ψ é contínua, pelo corolário 6.3. Por outro lado, a função

$$\begin{array}{rccc} \pi\colon & \mathrm{Gr}(f) & \longrightarrow & F \\ & (v, f(v)) & \mapsto & f(v), \end{array}$$

também é contínua. Logo, f é contínua, pois $f = \pi \circ \psi$. ■

A título de aplicação do teorema do gráfico fechado, vai-se demonstrar o seguinte

TEOREMA 6.8 *Seja $(X, \mathcal{A}, m)$ um espaço de medida. São então condições equivalentes:*

1. $(\forall p_1, p_2 \in [1, +\infty]) : p_1 < p_2 \Longrightarrow L^{p_1}(X) \subset L^{p_2}(X)$;
2. $(\exists p_1, p_2 \in [1, +\infty]) : p_1 < p_2 \wedge L^{p_1}(X) \subset L^{p_2}(X)$;
3. $\inf\{\, m(A) \mid A \in \mathcal{A} \wedge m(A) > 0 \,\} > 0$.

DEMONSTRAÇÃO: É claro que a primeira condição implica a segunda.

Se a segunda condição se verificar, sejam $p_1, p_2 \in [1, +\infty]$ tais que $p_1 < p_2$ e que $L^{p_1}(X) \subset L^{p_2}(X)$. Observe-se que a inclusão

$$\begin{array}{rccc} \iota\colon & L^{p_1}(X) & \longrightarrow & L^{p_2}(X) \\ & f & \mapsto & f \end{array}$$

é contínua. De facto, seja $\left(f_n, f_n\right)_{n\in\mathbb{N}}$ uma sucessão de pontos do gráfico de ι convergente para $(f, g) \in L^{p_1}(X) \times L^{p_2}(X)$; quer-se mostrar que (f, g) pertence ao gráfico de ι, i. e. que $g = f$. Sabe-se, pelo teorema 4.4, que existe alguma subsucessão $(f_{n_k})_{k\in\mathbb{N}}$ da sucessão $(f_n)_{n\in\mathbb{N}}$ para a qual se tem

$$\lim_{k\in\mathbb{N}} f_{n_k}(x) = f(x) \text{ q. s.} \tag{6.15}$$

Naturalmente, continua-se a ter $\lim_{k\in\mathbb{N}} f_{n_k} = g$ em $L^{p_2}(X)$, pelo que, pelo mesmo motivo, existe alguma subsucessão $(f_{m_k})_{k\in\mathbb{N}}$ da sucessão $(f_{n_k})_{k\in\mathbb{N}}$ tal que se tem

$$\lim_{k\in\mathbb{N}} f_{m_k}(x) = g(x) \text{ q. s.} \tag{6.16}$$

Mas resulta então de (6.15) e de (6.16) que $f(x) = g(x)$ q. s., ou seja, que $f = g$.

Seja $A \in \mathcal{A}$ tal que $0 < m(A) < +\infty$. Observe-se que se $p \in [1, +\infty]$, então

$$m(A)^{1/p} = \left\|\chi_A\right\|_p;$$

isto é trivial se $p = +\infty$ e, caso contrário, $\chi_A = \chi_A^p$, pelo que

$$m(A)^{1/p} = \left(\int_X \chi_A \, dm\right)^{1/p} = \left(\int_X \chi_A^p \, dm\right)^{1/p} = \left\|\chi_A\right\|_p.$$

Resulta desta observação que

$$m(A)^{1/p_2} = \left\|\chi_A\right\|_{p_2} = \left\|\iota(\chi_A)\right\|_{p_2} \leqslant \|\iota\| \left\|\chi_A\right\|_{p_1} = \|\iota\| m(A)^{1/p_1},$$

o que equivale a afirmar que

$$m(A)^{(p_1-p_2)/(p_1p_2)} \leqslant \|\iota\|. \tag{6.17}$$

Como, por hipótese, $p_1 < p_2$, resulta de (6.17) e de se estar a supor que $m(A) > 0$ que

$$m(A) \geqslant \|\iota\|^{p_1p_2/(p_1-p_2)}. \tag{6.18}$$

Por maioria de razão, também se tem (6.18) se $m(A) = +\infty$, o que mostra que a terceira condição se verifica.

Finalmente, suponha-se que a terceira condição se verifica. Sejam $p_1, p_2 \in [1, +\infty]$ tais que $p_1 < p_2$ e seja $f \in L^{p_1}(X)$; quer-se provar que $f \in L^{p_2}(X)$. Para cada $n \in \mathbb{N}$, seja

$$A_n = \{\, x \in X \mid |f(x)| > n \,\}.$$

Naturalmente, $\bigcap_{n\in\mathbb{N}} A_n = \emptyset$; em particular $m\big(\bigcap_{n\in\mathbb{N}} A_n\big) = 0$. Resulta de $|f|^{p_1}$ ser integrável que cada A_n tem medida finita e então, pela proposição 1.5,

$$0 = m\left(\bigcap_{n\in\mathbb{N}} A_n\right) = \lim_{n\in\mathbb{N}} m(A_n).$$

Mas está-se a supor que não há elementos de $\mathcal{A}$ com medida positiva arbitrariamente pequena, pelo que $m(A_n) = 0$ se n for suficientemente grande. Seja então $n \in \mathbb{N}$ tal que $m(A_n) = 0$, i. e. tal que $|f(x)| \leqslant n$ q. s. Então

$$\begin{aligned}\int_X |f|^{p_2}\,dm &= \int_X |f|^{p_1}|f|^{p_2-p_1}\,dm \\ &\leqslant \int_X |f|^{p_1} n^{p_2-p_1}\,dm \\ &= n^{p_2-p_1}\int_X |f|^{p_1}\,dm \\ &< +\infty,\end{aligned}$$

ou seja, $f \in L^{p_2}(X)$. ∎

EXEMPLO 6.9 O espaço de medida $(\mathbb{N}, \mathcal{P}(\mathbb{N}), m)$, onde m é a medida de contagem, está nas condições do enunciado.

Analogamente, pode-se demonstrar o

TEOREMA 6.9 *Seja $(X, \mathcal{A}, m)$ um espaço de medida. São então condições equivalentes:*

1. $(\forall p_1, p_2 \in [1, +\infty]) : p_1 < p_2 \Longrightarrow L^{p_1}(X) \supset L^{p_2}(X)$;
2. $(\exists p_1, p_2 \in [1, +\infty]) : p_1 < p_2 \wedge L^{p_1}(X) \supset L^{p_2}(X)$;
3. $\sup\{\, m(A) \mid A \in \mathcal{A} \wedge m(A) < +\infty \,\} < +\infty$.

EXEMPLO 6.10 O espaço de medida $([0,1], \mathcal{M}_{[0,1]}, l)$, está nas condições do enunciado.

Números reais e bases

Considere-se a igualdade

$$\frac{20}{9} = 2{,}2222222222\dots$$

O que isto significa é que $^{20}/_{9}$ é a soma da série

$$2.10^0 + 2.10^{-1} + 2.10^{-2} + 2.10^{-3} + \cdots$$

Mais geralmente, como se sabe, cada número $x \in \mathbb{R}_+$ pode ser escrito sob a forma $x = a_k a_{k-1} a_{k-2} \dots a_1 a_0{,}a_{-1} a_{-2} a_{-3} \dots$ com cada a_n ($n \in \mathbb{Z}$ e $n \leqslant k$) pertencente a $\{0, 1, \dots, 9\}$, querendo isto dizer que se tem $x = \sum_{n=-\infty}^{k} a_n.10^n$. É provavelmente menos conhecido que esta maneira de representar o número x não é necessariamente única, pois, por exemplo, $1 = 0{,}9999999999\dots$

Segundo o próximo teorema, aquilo que foi mencionado no parágrafo anterior relativamente ao número 10 pode ser feito com qualquer número natural $b > 1$. Antes de se enunciar o teorema, convém introduzir a seguinte notação: se b for um tal número natural, seja $\mathcal{S}_b$ o conjunto das famílias $(a_n)_{n\in\mathbb{N}}$ de elementos de $\{0, 1, \dots, b-1\}$ tais que $a_n = 0$ se n for suficientemente grande. Observe-se que, para cada elemento de $\mathcal{S}_b$, se se escolher $N \in \mathbb{Z}$ tal que $(\forall n \in \mathbb{Z}) : n > N \Longrightarrow a_n = 0$ (um tal N existe, por hipótese), então a série $\sum_{n=-\infty}^{N} a_n b^n$ converge e a sua soma é independente da esolha de N; aquela soma será representada por $\sum_{n=-\infty}^{+\infty} a_n b^n$.

TEOREMA A.1 *Seja b um número natural maior do que* 1. *É possível escrever cada número $x \in \mathbb{R}_+$ sob a forma*

$$x = \sum_{n=-\infty}^{+\infty} a_n b^n \tag{A.1}$$

para algum elemento $(a_n)_{n\in\mathbb{Z}} \in \mathcal{S}_b$. Mais precisamente, se $N \in \mathbb{Z}$ for tal que $x \leqslant b^N$, então tem-se (A.1) *para alguma família $(a_n)_{n\in\mathbb{Z}}$ com $a_n = 0$ sempre* *que $n \geqslant N$. Por outro lado, se não se tomarem em consideração os elementos de $(a_n)_{n\in\mathbb{Z}} \in \mathcal{S}_b$ para os quais $a_n = b-1$ para todo n suficientemente pequeno, então existe, para cada $x \in \mathbb{R}_+$, uma e uma só família $(a_n)_{n\in\mathbb{Z}} \in \mathcal{S}_b$ para a qual se tem* (A.1).

DEMONSTRAÇÃO: Seja $x \in \mathbb{R}_+$ e seja $N \in \mathbb{Z}$ tal que $x \leqslant b^N$; quer-se provar que se tem (A.1) para algum elemento de $(a_n)_{n\in\mathbb{Z}} \in \mathcal{S}_b$ tal que $a_n = 0$ sempre que $n \geqslant N$. Seja $(a_n)_{n<N}$ uma sucessão de elementos de $\{0, 1, \ldots, b-1\}$ tal que:

1. $a_{N-1} \in \{0, 1, \ldots, b-1\}$ é tal que
$$0 \leqslant xb^{-N} - a_{N-1}b^{-1} \leqslant b^{-1} (\Longleftrightarrow 0 \leqslant x - a_{N-1}b^{N-1} \leqslant b^{N-1});$$
2. se $a_N, a_{N-1}, \ldots, a_n$ ($n \leqslant N$) são tais que $0 \leqslant x - \sum_{k=n}^{N} a_k b^k \leqslant b^n$, a_{n-1} é tal que
$$0 \leqslant \left(x - \sum_{k=n}^{N} a_k b^k\right) - a_{n-1}b^{n-1} = x - \sum_{k=n-1}^{N} a_k b^k \leqslant b^{n-1}.$$

É então claro que se tem (A.1) se se definir $a_n = 0$ para cada $n \geqslant N$.

Vejamos agora quando se pode escrever $x \in \mathbb{R}_+$ sob a forma (A.1) de mais de uma maneira. Suponha-se então que

$$x = \sum_{n=-\infty}^{+\infty} a_n b^n = \sum_{n=-\infty}^{+\infty} \alpha_n b^n \tag{A.2}$$

para dois elementos distintos $(a_n)_{n\in\mathbb{Z}}, (\alpha_n)_{n\in\mathbb{Z}} \in \mathcal{S}_b$. Visto que $a_n = \alpha_n = 0$ se n for suficientemente grande e visto que se está a supor que $(a_n)_{n\in\mathbb{Z}} \neq (\alpha_n)_{n\in\mathbb{Z}}$, existe algum $N \in \mathbb{N}$ tal que $a_N \neq \alpha_N$ que $(\forall n \in \mathbb{Z}) : n > N \Longrightarrow a_n = \alpha_n$. Suponha-se que $a_N > \alpha_N$; o caso em que $a_N < \alpha_N$ é análogo. Então

$$\begin{aligned}
\sum_{n=-\infty}^{+\infty} a_n b^n = \sum_{n=-\infty}^{+\infty} \alpha_n b^n \Longleftrightarrow & \sum_{n=-\infty}^{N} a_n b^n = \sum_{n=-\infty}^{N} \alpha_n b^n \\
\Longleftrightarrow & \; a_N - \alpha_N = \sum_{n=-\infty}^{-1} (\alpha_{N+n} - a_{N+n}) b^n \\
\Longrightarrow & \sum_{n=-\infty}^{-1} (\alpha_{N+n} - a_{N+n}) b^n \geqslant 1 \\
\Longleftrightarrow & \sum_{n=-\infty}^{-1} (\alpha_{N+n} - a_{N+n}) b^n \geqslant \sum_{n=-\infty}^{-1} (b-1) b^n.
\end{aligned}$$

Mas todos os números da forma $\alpha_{N+n} - a_{N+n}$ ($n \in \mathbb{Z}_-^*$) são menores ou iguais a $b-1$, pelo que se tem a desigualdade anterior quando e só quando são todos iguais a $b-1$. Por outro lado se forem efectivamente todos iguais a $b-1$ (i. e. se cada α_{N+n} for igual a $b-1$ e cada a_{N+n} for igual a 0 para cada $n < 0$), então $\sum_{n=-\infty}^{-1}(\alpha_{N+n} - a_{N+n})b^n = 1$, pelo que só se poderá ter (A.2) se $a_N = \alpha_N + 1$. Está então provado que se os números $x \in \mathbb{R}_+$ que podem ser escritos sob a forma (A.1) de duas maneiras distintas são aqueles que podem ser escritos sob a forma

$$x = \sum_{n=N}^{+\infty} a_n b^n \quad (N \in \mathbb{N})$$

com $a_N, a_{N+1}, \ldots \in \{0, 1, \ldots, b-1\}$, $a_n = 0$ se $n \gg N$ e $a_N \neq 0$; um tal x pode também ser escrito sob a forma

$$x = \sum_{n=-\infty}^{+\infty} \alpha_n b^n$$

com

$$(\forall n \in \mathbb{N}) : \alpha_n = \begin{cases} a_n & \text{se } n > N \\ a_N - 1 & \text{se } n = N \\ b-1 & \text{se } n < N. \end{cases}$$

Reciprocamente, se $x \in \mathbb{R}_+$ puder ser escrito sob esta última forma, então também pode ser escrito sob a forma anterior. ∎

O enunciado do teorema pode transmitir a impressão de haver algo de defeituoso nas representações de um número $x \in \mathbb{R}_+$ sob a forma (A.1) para as quais se tem $a_n = b-1$ para cada n suficientemente grande mas, de facto, vê-se pela parte final da demontração que há uma simetria entre as representações daquela forma e aquelas para as quais se tem $a_n = 0$ para cada n suficientemente pequeno.

Dado um número $x \in \mathbb{R}_+$ e dado um número natural b tal que $1 < b \leqslant 10$, se se tem (A.1) é então usual representar-se x sob a forma

$$x = \alpha_k \alpha_{k-1} \alpha_{k-2} \ldots \alpha_1 \alpha_0, \alpha_{-1} \alpha_{-2} \alpha_{-3} \ldots$$

onde $k \in \mathbb{Z}_+$ é tal que $n \geqslant k \Longrightarrow \alpha_n = 0$ e onde α_n ($n \leqslant k$) é o algarismo correspodente ao número a_n; diz-se então que $\alpha_k \alpha_{k-1} \ldots \alpha_0, \alpha_{-1} \alpha_{-2} \ldots$ representa o número x na base b. Se existir um inteiro $m < 0$ tal que $\alpha_n = 0$ quando $n < m$, x também se representa por $\alpha_k \alpha_{k-1} \ldots \alpha_0, \alpha_{-1} \ldots \alpha_m$. Finalmente, se todos os α_n com $n < 0$ forem nulos (o que equivale a afirmar que $x \in \mathbb{Z}_+$), x representa-se por $\alpha_k \alpha_{k-1} \ldots \alpha_0$.

Caso $b > 10$, introduzem-se novos algarismos para representar os números $10, 11, \ldots, b-1$. Assim, por exemplo, para se representarem números em base 16 é corrente empregarem-se os símbolos A, B, C, D, E e F para representar 10, 11, 12 13, 14 e 15, respectivamente. Com estas notações, $3/4$ e $2/3$ representam-se na base 16 por $0{,}C$ e por $0{,}AAAAAAAAAA\ldots$ respectivamente.

Seja $\{0,1\}^{\mathbb{N}}$ o conjunto das sucessões de zeros e uns.

COROLÁRIO A.1 *Os conjuntos $\mathcal{P}(\mathbb{N})$, $\{0,1\}^{\mathbb{N}}$ e $\mathbb{R}$ têm o mesmo cardinal.*

DEMONSTRAÇÃO: O conjunto das partes de $\mathbb{N}$ tem o mesmo cardinal que o conjunto $\{0,1\}^{\mathbb{N}}$ das sucessões de zeros e uns, pois a função

$$\begin{array}{ccc} \{0,1\}^{\mathbb{N}} & \longrightarrow & \mathcal{P}(\mathbb{N}) \\ (a_n)_{n\in\mathbb{N}} & \mapsto & \{\, n \in \mathbb{N} \mid a_n = 1 \,\} \end{array}$$

é bijectiva. Por outro lado, resulta do teorema anterior (com $b = 2$) que a função

$$\begin{array}{rccc} \sigma\colon & \{0,1\}^{\mathbb{N}} & \longrightarrow & [0,1] \\ & (a_n)_{n\in\mathbb{N}} & \mapsto & \displaystyle\sum_{n=1}^{+\infty} \frac{a_n}{2^n} \end{array}$$

é sobrejectiva e que se $U = \{\, (a_n)_{n\in\mathbb{N}} \in \{0,1\}^{\mathbb{N}} \mid a_n = 1 \text{ se } n \gg 0 \,\}$, então a restrição de σ a $U^{\complement}$ é injectiva e tem por imagem $[0,1[$, que tem o mesmo cardinal que $\mathbb{R}$. Observe-se que U é numerável. Logo, $\{0,1\}^{\mathbb{N}}$ é a reunião de um conjunto com o mesmo cardinal que $\mathbb{R}$ (nomeadamente $U^{\complement}$) com um conjunto numerável (nomeadamente U), pelo que $\{0,1\}^{\mathbb{N}}$ tem o mesmo cardinal que $\mathbb{R}$. ∎

Lema de Zorn

Comecemos por ver a definição geral de «relação de ordem».

DEFINIÇÃO B.1 Diz-se que uma relação binária $\leqslant$ definida num conjunto C é uma *relação de ordem* se:

1. $(\forall a, b, c \in C) : a \leqslant b \wedge b \leqslant c \Longrightarrow a \leqslant c$;
2. $(\forall a \in C) : a \leqslant a$;
3. $(\forall a, b \in C) : a \leqslant b \wedge b \leqslant a \Longrightarrow a = b$.

Designa-se por *conjunto ordenado* um par ordenado $(C, \leqslant)$ onde C é um conjunto e $\leqslant$ é uma relação de ordem definida em C.

EXEMPLO B.1 Se P for um conjunto de partes de um conjunto, é usual considerar-se em P a relação de ordem $\leqslant$ induzida pela inclusão: $A \leqslant B$ se e só se $A \subset B$.

Num conjunto ordenado $(C, \leqslant)$, definem-se os termos *majorante* e *minorante* tal como em $\overline{\mathbb{R}}$ com a relação de ordem usual.

DEFINIÇÃO B.2 Seja $(C, \leqslant)$ um conjunto ordenado. Diz-se que um elemento $m \in C$ é um *elemento maximal* (respectivamente *minimal*) se, para cada $c \in C$, se tiver $c \leqslant m$ (resp. $m \leqslant c$).

É claro que $(\overline{\mathbb{R}}, \leqslant)$ não tem nenhum elemento maximal e também não tem nenhum elemento minimal. Em contrapartida, um conjunto ordenado pode ter vários elementos minimais bem como vários elementos maximais. Por exemplo, se C for um conjunto com mais de que um ponto e se se definir no conjunto P das partes não vazias de C a relação de ordem induzida pela inclusão, então os elementos minimais de $(P, \leqslant)$ são os conjuntos da forma $\{c\}$, com $c \in C$. Por outro lado, há um e um só elemento maximal, que é o próprio C.

DEFINIÇÃO B.3 Caso uma relação de ordem definida num conjunto C seja tal que, dados $a, b \in C$, se tenha $a \leqslant b$ ou $b \leqslant a$, então diz-se que se trata de uma *relação de ordem total*. Diz-se então que $(C, \leqslant)$ é um *conjunto totalmente ordenado*.

Por exemplo, a relação de ordem usual $\leqslant$ em $\overline{\mathbb{R}}$ é uma relação de ordem total. Em contrapartida, se C for um conjunto com mais do que um ponto e se se considerar no conjunto das partes de C a relação de ordem induzida pela inclusão, então não se trata de uma relação de ordem total.

DEFINIÇÃO B.4 Se $(C, \leqslant)$ for um conjunto ordenado, diz-se que um elemento c de C é *primeiro elemento* de C se

$$(\forall d \in C) : c \leqslant d.$$

Se um conjunto tiver primeiro elemento, este é necessariamente único, pois se a e a' forem ambos primeiro elemento de um conjunto A, então $a \leqslant a'$ e $a' \leqslant a$, pelo que $a = a'$.

Se $(C, \leqslant)$ for um conjunto ordenado e se $S \subset C$, então, a menos que seja dito explicitamente o contrário, considerar-se-á em S a relação de ordem induzida por $\leqslant$.

DEFINIÇÃO B.5 Caso uma relação de ordem definida num conjunto C seja tal que qualquer parte não vazia de C tenha primeiro elemento, então diz-se que se trata de uma *boa ordenação*. Diz-se então que $(C, \leqslant)$ é um *conjunto bem ordenado*.

Em $\mathbb{N}$, por exemplo, a relação de ordem usual é uma boa ordenação. Mas a relação de ordem usual em $\mathbb{Z}$ já não o é, pois o próprio conjunto $\mathbb{Z}$ não tem primeiro elemento.

Observe-se que qualquer conjunto bem ordenado $(C, \leqslant)$ é totalmente ordenado, pois se $a, b \in C$ e se c for o primeiro elemento de $\{a, b\}$, então $c = a$ ou $c = b$ e, por outro lado, $c \leqslant a$ e $c \leqslant b$.

LEMA B.1 (LEMA DE ZORN) *Seja $(C, \leqslant)$ um conjunto ordenado tal que qualquer subconjunto totalmente ordenado tenha algum majorante. Então $(C, \leqslant)$ tem algum elemento maximal.*

DEMONSTRAÇÃO: Vai-se considerar em C a relação binária $<$ assim definida: se $a, b \in C$, então $a < b$ se e só se $a \leqslant b$ e $a \neq b$.

Suponha-se que C não tem qualquer elemento maximal. Se A for uma parte totalmente ordenada de C então, por hipótese, A tem algum majorante m. Suponha-se que $m \in A$. Como se está a supor que C não tem elementos maximais então, em particular, m não é um elemento maximal e, portanto, existe algum $m' \in C$ tal que $m < m'$. Logo, $m' \notin A$ pois se se tivesse $m' \in A$ então, como m é majorante de A, decorreria que $m' \leqslant m$, o que é impossível, visto que $m \leqslant m'$ e $m' \neq m$. Está então provado que cada subconjunto A de C totalmente ordenado tem algum majorante que não pertence a A.

Seja $\mathcal{T}$ o conjunto das partes totalmente ordenadas de C e, para cada $A \in \mathcal{T}$, seja M_A o conjunto dos majorantes de A que não pertencem a A. Pelo que foi visto atrás, $\{\, M_A \mid A \in \mathcal{T} \,\}$ é um conjunto de conjuntos não vazios. Logo, pelo axioma da escolha, existe alguma função $f\colon \mathcal{T} \longrightarrow C$ tal que, para cada $A \in \mathcal{T}$, $f(A)$ é um majorante de A que não pertence a A.

Se $A \subset C$ e se $a \in A$, seja $S(A,a) = \{\, a' \in A \mid a' < a \,\}$. Dada uma parte B de C bem ordenada, dir-se-á que está subordinada a f caso se tenha:

$$(\forall b \in B) : f\big(S(B,b)\big) = b.$$

Vai-se provar que a reunião U de todas as partes bem ordenadas de C subordinadas a f é novamente uma parte bem ordenada de C subordinada a f. Uma vez provado isto, a demonstração estará praticamente concluída, pois se $u = f(U)$, é claro que $U \cup \{u\}$ é uma parte totalmente ordenada de C subordinada a f, pelo que $U \cup \{u\} \subset U$, ou seja, $u \in U$, o que é absurdo, pois $u = f(U)$ e $f(U)$ é um majorante de U que não pertence a U. Para se provar que U é uma parte bem ordenada de C subordinada a f, vai-se provar que se $u \in U$ e que se B for uma parte bem ordenada de C subordinada a f tal que $u \in B$, então $S(U,u) = S(B,u)$. Uma vez provado isto, estará provado que U é totalmente ordenado, pois se P for uma parte não vazia de U, se $u \in P$ e se B for uma parte totalmente ordenada de C subordinada a f tal que $u \in B$, há duas possibilidades:

1. ou $S(U,u) \cap P = \emptyset$ e então u é o primeiro elemento de P;
2. ou $S(U,u)$ intersecta P e então $S(U,u) \cap P = S(B,u) \cap P \subset B$ e, como B é bem ordenado, $S(U,u) \cap P$ tem necessariamente primeiro elemento, o qual terá então que ser o primeiro elemento de P.

Além disso, U é uma parte bem ordenada de C subordinada a f, pois se $u \in U$ e se B for uma parte totalmente ordenada de C subordinada a f tal que $u \in B$, então

$$f\big(S(U,u)\big) = f\big(S(B,u)\big) = u.$$

Sejam então $u \in U$ e B uma parte totalmente ordenada de C subordinada a f tal que $u \in B$; quer-se provar que $S(U,u) = S(B,u)$ o que equivale a afirmar que $S(U,u) \subset S(B,u)$. Por outras palavras, quer-se provar que se $a \in U$ for tal que $a < u$, então $a \in B$. Suponha-se que $a \notin B$. Afirmar que $a \in U$ é afirmar que $a \in A$ para alguma parte totalmente ordenada de C subordinada a f tal que $a \in A$. Por hipótese, $a \in A \setminus B$ e, portanto, $A \setminus B \neq \emptyset$. Seja α o primeiro elemento de $A \setminus B$. Vai-se provar que $B = S(A,\alpha)$. Resultará daqui que $u < \alpha \leqslant a < u$, o que é absurdo.

Como α é o primeiro elemento de $A \setminus B$, é claro que

$$S(A,\alpha) \subset B. \tag{B.1}$$

Se $S(A,\alpha) \subsetneq B$, então seja b o primeiro elemento de $B \setminus S(A,\alpha)$. Se $u \in S(B,b)$ e se $v \in A$ for tal que $v < u$, então $v < b$ (pois $u < b$). Por outro lado, pela definição de b tem-se que $u \in S(A,\alpha)$, de onde resulta que $v \in S(A,\alpha)$ (pois, por hipótese, $v \in A$). Mas então $v \in S(B,b)$, por (B.1) e porque $v < b$. Está então provado que

$$(\forall u \in S(B,b))(\forall v \in S(A,\alpha)) : v < u \Longrightarrow v \in S(B,b). \tag{B.2}$$

Como $A \setminus B$ não é vazio então $A \setminus S(B,b)$ não é vazio; seja z o seu primeiro elemento. É trivial que $S(A,z) \subset S(B,b)$. Vejamos que, de facto,

$$S(B,b) = S(A,z); \tag{B.3}$$

para o demonstrar falta somente provar que $S(B,b) \subset S(A,z)$. Se $u \in S(B,b)$ então, pela definição de b, $u \in S(A,\alpha)$; como $S(A,\alpha) \subset A$, a fim de provar que $u \in S(A,z)$ só falta provar que $u < z$. Mas z é o primeiro elemento de $A \setminus S(B,b)$ e $u \in S(B,b)$, pelo que $u \neq z$. Como $u, z \in A$, que é bem ordenado, resulta desta observação que $z < u$ ou $u < z$. Por (B.2), se se tivesse $z < u$, então ter-se-ia que $z \in S(B,b)$, o que não é o caso. Está então provado que $u < z$, ou seja que $u \in S(A,z)$ como se queria demonstrar. Da igualdade (B.3) resulta agora que

$$b = f\big(S(B,b)\big) = f\big(S(A,z)\big) = z.$$

Para terminar a demonstração, comparemos α com z. Uma vez que $\alpha \in A \setminus B \subset A \setminus S(B,b)$ e que z é o primeiro elemento de $A \setminus S(B,b)$, é claro que $z \leqslant \alpha$. Mas não se pode ter $z = \alpha$, pois $z = b \in B$, enquanto que $\alpha \notin B$. Logo, $z < \alpha$ e, como $z \in A$, isto é o mesmo que dizer que $z \in S(A,\alpha)$, ou seja, $b \in S(A,\alpha)$. Isto é impossível, pois b é o primeiro elemento de $B \setminus S(A,\alpha)$. ∎

A demonstração anterior empregou o axioma da escolha. É interessante observar que, de facto, o axioma da escolha é *equivalente* ao lema de

Zorn, no sentido de que deste é possível deduzir aquele. Vai-se ver como é que se pode fazer tal demonstração, pois é curta e é também um bom exemplo de como aplicar o lema de Zorn.

Seja então $\{ X_i \mid i \in I \}$ um conjunto de conjuntos não vazios; quer--se provar que existe algum conjunto $\{ x_i \mid i \in I \}$ tal que, para cada $i \in I$, $x_i \in X_i$, o que é o mesmo que afirmar que existe alguma função $f\colon I \longrightarrow \bigcup_{i\in I} X_i$ tal que, para cada $i \in$, $f(i) \in X_i$. Seja $\mathcal{S}$ o conjunto de todas as funções $f\colon J \longrightarrow \bigcup_{i\in I} X_i$ tais que $J \subset I$ e que, para cada $i \in J$, $f(i) \in X_i$. Naturalmente, o que se pretende provar é que $\mathcal{S}$ tem algum elemento cujo domínio seja I. Considere-se em $\mathcal{S}$ a relação de ordem $\leqslant$ assim definida: dada uma função $f \in \mathcal{S}$ de domínio J e dada uma função $g \in \mathcal{S}$ de domínio J', tem-se $f \leqslant g$ se e só se $J \subset J'$ e $f = g|_J$.

Seja $A = \left(f_k\colon J_K \longrightarrow \bigcup_{i\in I} X_i\right)_{k\in K}$ uma parte totalmente ordenada de $\mathcal{S}$; quer-se provar que tem algum majorante em $\mathcal{S}$. Seja $J = \bigcup_{k\in K} J_k$. Se $i \in J$, então $i \in J_k$, para algum $k \in K$. Caso i também pertenca a algum $J_{k'}$ ($k' \in K$), então, visto que $\mathcal{S}$ é totalmente ordenado, tem-se $f_k \leqslant f_{k'}$ ou $f_{k'} \leqslant f_k$. Suponha-se que $f_k \leqslant f_{k'}$. Então, pela definição de $\leqslant$, $f_k(i) = f_{k'}(i)$. Pelo mesmo argumento, se $f_{k'} \leqslant f_k$ então também se tem $f_k(i) = f_{k'}(i)$. Pode-se então definir $f\colon J \longrightarrow \bigcup_{i\in I} X_i$ do seguinte modo: se $i \in J_k$ ($k \in K$), então $f(i) = f_k(i)$. É claro que f é um majorante de A.

Estão então satisfeitas as condições do lema de Zorn, pelo que $\mathcal{S}$ tem algum elemento maximal $f\colon J \longrightarrow \bigcup_{i\in I} X_i$. Caso não $J \subsetneq I$, seja $i \in I \setminus J$, seja x_i um elemento de X_i e considere-se a função $F\colon J \cup \{i\} \longrightarrow \bigcup_{i\in I} X_i$ que prolonga f e tal que $F(i) = x_i$. Então $f \leqslant F$ e $f \neq F$, o que contradiz o facto de f ser um elemento maximal. Logo, $J = I$.

Bibliografia

[1] Asplund, E. and L. Bungart: *A first course in integration*. Holt, Rinehart and Winston, 1966.

[2] Banach, S. et H. Steinhaus: *Sur le principe de la condensation de singularités*. Fund. Math., 9 :50–61, 1927.

[3] Botsko, M. W.: *An elementary proof of Lebesgue's differentiation theorem*. Amer. Math. Monthly, 110:834–838, 2003.

[4] Carleson, L.: *On convergence and growth of partial sums of Fourier series*. Acta Math., 116:135–157, 1966.

[5] Ciesielski, K.: *How good is Lebesgue measure?* Math. Intelligencer, 11(2):54–58, 1989.

[6] Cohn, D. L.: *Measure Theory*. Birkhäuser, 1980.

[7] Conway, J. B.: *A course in Functional Analysis*. Springer-Verlag, 1985.

[8] Folland, G. B.: *Real Analysis: Modern Techniques and their Applications*. John Wiley & Sons, 1984.

[9] Hawkins, T.: *Lebesgue's theory of integration: Its origins and developments*. Chelsea, 2nd edition, 1975.

[10] Körner, T. W.: *Fourier Analysis*. Cambridge University Press, 1995.

[11] Lusin, N.: *Les ensembles analytiques*. Fund. Math., 10 :1–95, 1927.

[12] Matos, Coimbra de e J. C. Santos: *Análise Complexa*. Dinternal, 2000.

[13] Rudin, W.: *Principles of Mathematical Analysis*. McGraw-Hill, 3rd edition, 1976.

[14] Rudin, W.: *Real and Complex Analysis*. McGraw-Hill, 3rd edition, 1986.

[15] Solovay, R. M.: *A model of set theory in which every set of reals is Lebesgue measurable*. Ann. of Math. (2), 92:1–56, 1970.

[16] Spivak, M.: *Calculus*. Publish or Perish, 3rd edition, 1994.

Índice remissivo