

# Capítulo 3

## Formas Lineares e Normas em Espaços Vetoriais

### Conteúdo

---

<b>3.1</b>	<b>Formas Lineares, Sesquilineares e Produtos Escalares em Espaços Vetoriais . . . . .</b>	<b>220</b>
3.1.1	Formas Multilineares . . . . .	220
3.1.2	Formas Sesquilineares e as Desigualdades de Cauchy-Schwarz e Minkowski . . . . .	226
3.1.3	Produtos Escalares . . . . .	229
3.1.4	Formas Quadráticas em Espaços Vetoriais Reais . . . . .	231
3.1.5	Exemplos . . . . .	231
<b>3.2</b>	<b>Normas em Espaços Vetoriais . . . . .</b>	<b>233</b>
3.2.0.1	O Lema da Simetria . . . . .	241
<b>3.3</b>	<b>Ortogonalidade, Conjuntos Ortonormais e o Procedimento de Gram-Schmidt . . . . .</b>	<b>242</b>
<b>3.4</b>	<b>Formas Bilineares e Sesquilineares e Produtos Escalares em Espaços de Dimensão Finita</b>	<b>244</b>
<b>3.5</b>	<b>Estruturas Complexas sobre Espaços Vetoriais Reais . . . . .</b>	<b>247</b>
	<b>APÊNDICES . . . . .</b>	<b>254</b>
<b>3.A</b>	<b>Equivalência de Normas em Espaços Vetoriais de Dimensão Finita . . . . .</b>	<b>254</b>
<b>3.B</b>	<b>Prova do Teorema de Fréchet, von Neumann e Jordan . . . . .</b>	<b>255</b>

---



noção de espaço vetorial que introduzimos na Seção 2.1.5, página 100, é da maior importância na Física e na Matemática. Neste capítulo vamos estudá-la com mais detalhe. Particular atenção será dada às noções de forma multilinear, forma sesquilinear, produto escalar e norma em espaços vetoriais. As importantes desigualdades de Cauchy-Schwarz e de Minkowski serão demonstradas com bastante generalidade. Este capítulo trata quase exclusivamente de aspectos “algébricos” de espaços vetoriais, pondo de lado aspectos topológicos, os quais serão discutidos em capítulos futuros.

## 3.1 Formas Lineares, Sesquilineares e Produtos Escalares em Espaços Vetoriais

### 3.1.1 Formas Multilineares

Seja  $V$  um espaço vetorial sobre um corpo  $\mathbb{K}$  (que doravante suporemos ter característica diferente de 2, o caso, por exemplo, dos reais ou dos complexos) e  $n$  um número inteiro positivo. Uma  $n$ -forma multilinear<sup>1</sup> em  $V$  é uma função  $\omega : V^n \rightarrow \mathbb{K}$  que seja linear em cada um dos seus argumentos, ou seja, para todo  $\alpha, \beta \in \mathbb{K}$ , todos  $v_1, \dots, v_n \in V$ ,  $v'_i \in V$  e todo  $i = 1, \dots, n$  vale

$$\begin{aligned} \omega(v_1, \dots, v_{i-1}, (\alpha v_i + \beta v'_i), v_{i+1}, \dots, v_n) \\ = \alpha \omega(v_1, \dots, v_{i-1}, v_i, v_{i+1}, \dots, v_n) + \beta \omega(v_1, \dots, v_{i-1}, v'_i, v_{i+1}, \dots, v_n). \end{aligned} \quad (3.1)$$

O seguinte fato importante é consequência imediata da definição acima: se  $\omega$  é uma  $n$ -forma multilinear então

$$\omega(v_1, \dots, v_{i-1}, 0, v_{i+1}, \dots, v_n) = 0$$

para todo  $i$ , ou seja, se um dos argumentos é o vetor nulo a forma se anula.

---

<sup>1</sup>Também chamada  $n$ -forma linear ou simplesmente  $n$ -forma.

**E. 3.1** *Exercício.* Prove isso. Sugestão: o que acontece se escolhermos  $\alpha = \beta = 0$ ? ✦

Um fato importante é o seguinte: o conjunto de todas as  $n$ -formas lineares em um espaço vetorial  $V$  sobre um corpo  $\mathbb{K}$  é igualmente um espaço vetorial sobre  $\mathbb{K}$ , que denotaremos por  $\mathcal{M}_n(V, \mathbb{K})$ , ou simplesmente por  $\mathcal{M}_n(V)$ . Para tal procede-se da seguinte forma: para duas  $n$ -formas lineares  $\omega_1$  e  $\omega_2$  e dois escalares  $\alpha_1, \alpha_2 \in \mathbb{K}$  define-se a combinação linear  $\alpha_1\omega_1 + \alpha_2\omega_2$  como sendo a  $n$ -forma linear que a toda  $n$ -upla de vetores  $v_1, \dots, v_n \in V$  associa

$$(\alpha_1\omega_1 + \alpha_2\omega_2)(v_1, \dots, v_n) = \alpha_1\omega_1(v_1, \dots, v_n) + \alpha_2\omega_2(v_1, \dots, v_n).$$

**E. 3.2** *Exercício.* Complete os detalhes da prova que o conjunto de todas as  $n$ -formas lineares em um espaço vetorial  $V$  sobre um corpo  $\mathbb{K}$  forma um espaço vetorial sobre  $\mathbb{K}$ . ✦

• **Formas bilineares**

De particular interesse é o caso  $n = 2$ , em cujo caso as formas são denominadas *formas bilineares*: uma forma bilinear é uma função  $\omega : V^2 \rightarrow \mathbb{K}$  que seja linear em cada um dos seus dois argumentos, ou seja, para todo  $\alpha, \beta \in \mathbb{K}$ , todos  $u, v, w \in V$ , valem

$$\begin{aligned} \omega(u, (\alpha v + \beta w)) &= \alpha\omega(u, v) + \beta\omega(u, w), \\ \omega((\alpha u + \beta v), w) &= \alpha\omega(u, w) + \beta\omega(v, w). \end{aligned}$$

Um exemplo básico importante é o seguinte. Seja  $V = \mathbb{R}^n$  o espaço vetorial (sobre o corpo dos reais) formado por  $n$ -uplas de números reais:  $V = \{x = (x_1, \dots, x_n), x_i \in \mathbb{R}\}$ . Uma forma bilinear em  $V$  é dada por

$$\langle x, y \rangle_{\mathbb{R}} = \sum_{k=1}^n x_k y_k. \tag{3.2}$$

Outro exemplo é  $\omega_A(x, y) = \langle x, Ay \rangle_{\mathbb{R}}$ , onde  $A$  é uma matriz  $n \times n$  real qualquer.

• **Formas bilineares simétricas e antissimétricas**

Uma forma bilinear  $\omega$  é dita ser uma *forma bilinear simétrica* se satisfizer  $\omega(u, v) = \omega(v, u)$  para todos  $u, v \in V$ .

Uma forma bilinear  $\omega$  é dita ser uma *forma bilinear antissimétrica* se satisfizer  $\omega(u, v) = -\omega(v, u)$  para todos  $u, v \in V$ . A noção de forma bilinear antissimétrica será estendida logo abaixo com a introdução da noção de *forma alternante*.

Se  $\omega$  é uma forma bilinear, as formas  $\omega_r$  e  $\omega_a$  definidas por  $\omega_r(u, v) := \frac{1}{2}(\omega(u, v) + \omega(v, u))$  e  $\omega_a(u, v) := \frac{1}{2}(\omega(u, v) - \omega(v, u))$  são, respectivamente, simétrica e antissimétrica. Naturalmente,  $\omega = \omega_r + \omega_a$  e, portanto, toda forma bilinear pode ser escrita como soma de uma forma simétrica e de uma antissimétrica.

• **A identidade de polarização para forma bilineares simétricas**

Se  $\omega$  for uma forma bilinear simétrica em  $V$ , vale a seguinte relação, denominada *identidade de polarização de formas bilineares simétricas*:

$$\omega(u, v) = \frac{1}{4} [\omega(u + v, u + v) - \omega(u - v, u - v)], \tag{3.3}$$

válida para todos  $u, v \in V$ . Para verificá-la, basta expandir-se o lado direito e constatar-se, com uso da simetria de  $\omega$ , que se obtém o lado esquerdo.

**E. 3.3** *Exercício.* Verifique! ✦

A importância da identidade de polarização (3.3) é mostrar que, no caso de formas bilineares simétricas, o conhecimento dos valores  $\omega(u, u)$  para todos os  $u \in V$  determina os valores  $\omega(u, v)$  para todos os  $u, v \in V$ .

• **Formas bilineares não-degeneradas**

Uma forma bilinear simétrica ou antissimétrica  $\omega$  é dita ser uma *forma bilinear não-degenerada* se satisfizer a seguinte condição: se para todo vetor  $v$  valer  $\omega(v, u) = 0$ , então  $u = 0$ .

• **Formas bilineares não-singulares**

Seja  $V$  um espaço vetorial e  $\omega$  uma forma bilinear em  $V$ . Para  $u \in V$  fixo a aplicação  $l_u(v) = \omega(u, v)$  é um funcional linear em  $V$ , ou seja, um elemento do espaço dual  $V'$ . Se a aplicação  $l : V \rightarrow V'$  que associa cada  $u \in V$  ao funcional linear  $l_u$  acima for um isomorfismo de espaços vetoriais a forma bilinear  $\omega$  é dita ser uma *forma bilinear não-singular*.

Há vários outros tipos de formas multilineares que são importantes, como por exemplo as chamadas formas multilineares alternantes e, dentre estas, as formas simpléticas.

• **Formas simétricas**

Uma  $n$ -forma  $\omega$  em  $V$  é dita ser uma *forma simétrica* se para todo  $\pi \in S_n$ , o grupo de permutações de  $n$  elementos, valer

$$\omega(v_{\pi(1)}, \dots, v_{\pi(n)}) = \omega(v_1, \dots, v_n), \tag{3.4}$$

para quaisquer vetores  $v_1, \dots, v_n \in V$ .

• **Formas alternantes**

Uma  $n$ -forma linear  $\omega$  em um espaço vetorial  $V$  sobre um corpo  $\mathbb{K}$  é dita ser uma *forma alternante* (ou uma *forma antissimétrica*) se satisfizer

$$\omega(v_1, \dots, v_{i-1}, v_i, v_{i+1}, v_{i+2}, \dots, v_n) = -\omega(v_1, \dots, v_{i-1}, v_{i+1}, v_i, v_{i+2}, \dots, v_n) \tag{3.5}$$

para todos os vetores  $v_1, \dots, v_n \in V$  e todo  $i = 1, \dots, n - 1$ . Em palavras, quando trocamos de lugar dois argumentos vizinhos quaisquer a forma troca de sinal.

Deve ser bem claro que essa definição equivale à seguinte afirmação: se  $\omega$  é uma  $n$ -forma linear alternante, então para todo  $\pi \in S_n$ , o grupo de permutações de  $n$  elementos, vale

$$\omega(v_{\pi(1)}, \dots, v_{\pi(n)}) = (\text{sinal } \pi) \omega(v_1, \dots, v_n), \tag{3.6}$$

para todos os vetores  $v_1, \dots, v_n \in V$ , onde  $\text{sinal } \pi$  é o sinal da permutação  $\pi$  (definida na Seção 21.2.1.1, página 1037).

**E. 3.4** *Exercício*. Está claro? ✱

**Nomenclatura.** Se  $\omega$  é uma  $n$ -forma linear alternante,  $n$  é dito ser o *grau* de  $\omega$ .

O conjunto de todas as  $n$ -formas lineares alternantes em um espaço vetorial  $V$  sobre um corpo  $\mathbb{K}$  é igualmente um espaço vetorial sobre  $\mathbb{K}$ : para duas  $n$ -formas lineares alternantes  $\omega_1$  e  $\omega_2$  e dois escalares  $\alpha_1, \alpha_2 \in \mathbb{K}$  define-se a combinação linear  $\alpha_1\omega_1 + \alpha_2\omega_2$  como sendo a  $n$ -forma linear que a toda  $n$ -upla de vetores  $v_1, \dots, v_n \in V$  associa

$$(\alpha_1\omega_1 + \alpha_2\omega_2)(v_1, \dots, v_n) = \alpha_1\omega_1(v_1, \dots, v_n) + \alpha_2\omega_2(v_1, \dots, v_n).$$

É fácil constatar que a  $n$ -forma linear assim definida é também alternante.

**E. 3.5** *Exercício*. Complete os detalhes da prova que o conjunto de todas as  $n$ -formas lineares alternantes em um espaço vetorial  $V$  sobre um corpo  $\mathbb{K}$  forma um espaço vetorial sobre  $\mathbb{K}$ . ✱

• **Formas simpléticas**

Formas bilineares alternantes não-degeneradas são denominadas *formas simpléticas*<sup>2</sup>. Formas simpléticas são importantes em algumas áreas da Física, como por exemplo na Mecânica Clássica e no estudo de métodos de quantização.

Assim, uma forma simplética em um espaço vetorial  $V$  sobre um corpo  $\mathbb{K}$  é uma forma bilinear para a qual

$$\omega(u, v) = -\omega(v, u)$$

para todos os vetores  $u, v \in V$  e tal que se  $\omega(u, v) = 0$  para todo  $v$ , então  $u = 0$ .

<sup>2</sup>Do grego *symplektikós*: que serve para ligar, trançado, enlaçado.

Um exemplo básico importante de forma simplética no caso do espaço vetorial  $V = \mathbb{R}^n$  e que, como veremos na Seção 3.4, é o caso geral, vem a ser o seguinte:

$$\omega_A(x, y) = \langle x, Ay \rangle_{\mathbb{R}},$$

onde  $A$  é uma matriz  $n \times n$  real, inversível e *antissimétrica*, ou seja, que satisfaz  $A^T = -A$  (o que equivale a dizer que seus elementos de matriz satisfazem  $A_{ij} = -A_{ji}$ ). De fato,  $\omega_A$  é antissimétrica, pois

$$\omega_A(x, y) = \langle x, Ay \rangle_{\mathbb{R}} = \langle A^T x, y \rangle_{\mathbb{R}} = -\langle Ax, y \rangle_{\mathbb{R}} = -\langle y, Ax \rangle_{\mathbb{R}} = -\omega_A(y, x).$$

Fora isso,  $\omega_A$  é não-degenerada, pois se houver  $x \neq 0$  tal que  $\langle x, Ay \rangle_{\mathbb{R}} = 0$  para todo  $y$ , então  $0 = \langle x, Ay \rangle_{\mathbb{R}} = \langle A^T x, y \rangle_{\mathbb{R}} = -\langle Ax, y \rangle_{\mathbb{R}}$  para todo  $y$ , o que só é possível se  $Ax = 0$ . Como  $x \neq 0$ , isso implicaria que  $A$  não é injetora e, portanto, que  $A$  não seria inversível.

Uma consequência do fato de  $A$  ter de ser inversível é que  $n$  tem que ser par. De fato, a condição  $A^T = -A$  diz que  $\det(A) = \det(-A^T) = (-1)^n \det(A^T) = (-1)^n \det(A)$ . Portanto, se  $n$  é ímpar teríamos  $\det(A) = 0$ , uma contradição.

• Algumas propriedades básicas de formas lineares alternantes

É evidente pela definição que se  $\omega$  é uma  $n$ -forma alternante então  $\omega(v_1, \dots, v_n) = 0$  caso haja  $v_i = v_j$  para algum par  $i \neq j$ . Em particular, para formas simpléticas  $\omega(u, u) = 0$  para todo  $u \in V$ .

**E. 3.6 Exercício.** A propriedade mencionada no último parágrafo é equivalente à definição de forma linear alternante: se  $\omega$  é uma  $n$ -forma linear e  $\omega(v_1, \dots, v_n) = 0$  sempre que  $v_i = v_j$  para algum par  $i \neq j$ , então  $\omega$  é alternante. Prove isso. Sugestão: para  $i \neq j$  defina a forma bilinear  $\omega_{ij}(v_i, v_j) := \omega(v_1, \dots, v_n)$  onde todos os vetores  $v_1, \dots, v_n$  estão fixos exceto  $v_i$  e  $v_j$ . Usando agora que  $\omega_{ij}(x+y, x+y) = 0$ , mostre que  $\omega_{ij}(v_i, v_j) = -\omega_{ij}(v_j, v_i)$  para todo  $v_i$  e  $v_j$ . A afirmação principal segue disso (por quê?). \*

A seguinte proposição sobre formas lineares alternantes é importante:

**Proposição 3.1** Se  $\omega$  é uma  $n$ -forma linear alternante e  $v_1, \dots, v_n$  são vetores linearmente dependentes, então  $\omega(v_1, \dots, v_n) = 0$ . □

**E. 3.7 Exercício.** Prove isso. \*

• Formas alternantes maximais

A Proposição 3.1 tem uma consequência imediata: se  $V$  é um espaço vetorial de dimensão  $n$  e  $\omega$  é uma forma linear alternante de ordem  $m > n$ , então  $\omega = 0$ .

**E. 3.8 Exercício.** Por quê? \*

Assim, em um espaço de dimensão  $n$  o grau máximo de uma forma alternante é  $n$ . Formas alternantes de grau máximo são ditas *formas alternantes maximais*. Vamos mais adiante estudar como são essas formas maximais, mas antes, precisamos discutir alguns fatos importantes sobre formas alternantes em espaços de dimensão finita.

Em um espaço vetorial  $V$  de dimensão  $n$  o espaço vetorial das formas alternantes maximais é unidimensional. Para ver isso notemos o seguinte. Seja  $\{b_1, \dots, b_n\}$  uma base em  $V$ . Sejam agora  $\omega_1$  e  $\omega_2$  duas formas alternantes maximais em  $V$  e seja  $x_1, \dots, x_n$  uma  $n$ -upla de vetores de  $V$ . Como  $\{b_1, \dots, b_n\}$  é uma base, podemos sempre escrever

$$x_i = \sum_{j=1}^n \alpha_{ij} b_j,$$

para todo  $i = 1, \dots, n$ . Assim,

$$\omega_1(x_1, \dots, x_n) = \sum_{j_1=1}^n \cdots \sum_{j_n=1}^n \alpha_{1j_1} \cdots \alpha_{nj_n} \omega_1(b_{j_1}, \dots, b_{j_n})$$

e, analogamente,

$$\omega_2(x_1, \dots, x_n) = \sum_{j_1=1}^n \cdots \sum_{j_n=1}^n \alpha_{1j_1} \cdots \alpha_{nj_n} \omega_2(b_{j_1}, \dots, b_{j_n}).$$

Ocorre que  $\omega_1(b_{j_1}, \dots, b_{j_n})$  é zero caso ocorram dois índices  $j_k$  iguais. Por isso, podemos reescrever as expressões acima da seguinte forma:

$$\omega_1(x_1, \dots, x_n) = \sum_{j \in S_n} \alpha_{1j(1)} \cdots \alpha_{nj(n)} \omega_1(b_{j(1)}, \dots, b_{j(n)})$$

e, analogamente,

$$\omega_2(x_1, \dots, x_n) = \sum_{j \in S_n} \alpha_{1j(1)} \cdots \alpha_{nj(n)} \omega_2(b_{j(1)}, \dots, b_{j(n)}),$$

onde, acima,  $S_n$  é o conjunto de todas as bijeções de  $\{1, \dots, n\}$  em si mesmo (o chamado *grupo de permutações de  $n$  elementos*).

**E. 3.9** *Exercício.* Justifique. \*

Como  $\omega_1$  é uma forma alternante maximal, tem-se que

$$\omega_1(b_{j(1)}, \dots, b_{j(n)}) = \text{sinal}(j) \omega_1(b_1, \dots, b_n).$$

Assim,

$$\omega_1(x_1, \dots, x_n) = \left( \sum_{j \in S_n} \text{sinal}(j) \alpha_{1j(1)} \cdots \alpha_{nj(n)} \right) \omega_1(b_1, \dots, b_n)$$

e, analogamente,

$$\omega_2(x_1, \dots, x_n) = \left( \sum_{j \in S_n} \text{sinal}(j) \alpha_{1j(1)} \cdots \alpha_{nj(n)} \right) \omega_2(b_1, \dots, b_n).$$

Como se vê nessas últimas expressões,  $\omega_1(x_1, \dots, x_n)$  e  $\omega_2(x_1, \dots, x_n)$  diferem apenas pelos fatores  $\omega_1(b_1, \dots, b_n)$  e  $\omega_2(b_1, \dots, b_n)$ , respectivamente. Como esses fatores são apenas números (elementos do corpo  $\mathbb{K}$ ), são proporcionais um ao outro. Isso prova então que  $\omega_1(x_1, \dots, x_n)$  e  $\omega_2(x_1, \dots, x_n)$  são proporcionais um ao outro para toda  $n$ -upla  $x_1, \dots, x_n$  e isso era o que queríamos provar.

Com as observações acima chegamos ao importante conceito de forma determinante.

• **A forma determinante**

Como observamos acima, todas as  $n$ -formas lineares alternantes maximais de um espaço vetorial  $V$  de dimensão  $n$  são proporcionais umas às outras. Assim, o conhecimento de uma forma alternante maximal determina todas as outras.

A *forma determinante*<sup>3</sup>  $\omega_{det}$  em um espaço vetorial  $V$  de dimensão  $n$  é a  $n$ -forma linear alternante maximal tal que  $\omega_{det}(b_1, \dots, b_n) = 1$  no caso em que  $\{b_1, \dots, b_n\}$  é a base canônica de  $V$ :

$$b_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad b_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad \dots, \quad b_n = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}.$$

Assim,

$$\omega_{det}(x_1, \dots, x_n) = \sum_{j \in S_n} \text{sinal}(j) \alpha_{1j(1)} \cdots \alpha_{nj(n)},$$

onde  $\alpha_{ij}$  é a  $j$ -ésima componente do vetor  $x_i$  na base canônica.

Como observamos, todas as outras  $n$ -formas lineares alternantes maximais de  $V$  são proporcionais a  $\omega_{det}$ .

---

<sup>3</sup>Também chamada de *forma volume*, pois em  $\mathbb{R}^3$ ,  $\omega_{det}(x_1, x_2, x_3)$  é igual ao volume do paralelepípedo descrito pelos vetores  $x_1, x_2, x_3$ .

• **Determinante de matrizes**

Sejam  $a_1, \dots, a_n$  vetores, representados na base canônica por vetores-coluna

$$a_i = \begin{pmatrix} \alpha_{i1} \\ \vdots \\ \alpha_{in} \end{pmatrix} .$$

Denotamos por  $\llbracket a_1, \dots, a_n \rrbracket$  a matriz  $n \times n$  construída de forma que sua  $k$ -ésima coluna seja o vetor-coluna  $a_k$ , ou seja

$$\llbracket a_1, \dots, a_n \rrbracket = \begin{pmatrix} \alpha_{11} & \cdots & \alpha_{n1} \\ \vdots & \ddots & \vdots \\ \alpha_{1n} & \cdots & \alpha_{nn} \end{pmatrix} .$$

É evidente que toda matriz  $A \in \text{Mat}(\mathbb{C}, n)$  pode ser escrita na forma  $A = \llbracket a_1, \dots, a_n \rrbracket$  para algum conjunto de vetores  $a_1, \dots, a_n$  que representam suas colunas.

Define-se, então, o *determinante* da matriz  $A$  como sendo

$$\det(A) := \omega_{det}(a_1, \dots, a_n) , \tag{3.7}$$

ou seja,

$$\det(A) = \sum_{j \in S_n} \text{sinal}(j) \alpha_{1j(1)} \cdots \alpha_{nj(n)} . \tag{3.8}$$

Essa expressão é frequentemente denominada *fórmula de Leibniz*<sup>4</sup> para o determinante de uma matriz.

Cremos que o conceito de determinante de matrizes e suas propriedades básicas sejam bem conhecidos do estudante que tenha uma formação básica em Cálculo e Álgebra Linear, mas as mesmas serão (re)apresentadas e deduzidas na Seção 10.1, página 462. Vide, em particular, o Teorema 10.1, página 467.

Na Seção 11.7, página 623, com base no Teorema de Hadamard, Teorema 10.40, página 576, demonstraremos que o determinante de matrizes é uma função contínua.

• **Formas Multilineares em dimensão finita e produtos tensoriais**

Se  $U$  é um espaço de dimensão finita sobre o corpo  $\mathbb{K}$  estabelecemos na Seção 2.3.5.1, página 181, que o espaço vetorial  $\mathcal{M}_n(U)$  de todas as formas  $n$ -lineares sobre  $U$  é isomorfo ao produto tensorial  $(U')^{\otimes n}$  e dos espaços duais de  $U$  e, portanto, também a  $(U^{\otimes n})'$ .

Seja  $\{\mathbf{e}_1, \dots, \mathbf{e}_m\}$  uma base em  $U$  e  $\{\ell_1, \dots, \ell_m\}$  sua correspondente base dual canônica. Se  $\omega \in \mathcal{M}_n(U)$  e  $u_1, \dots, u_n \in U$  com  $u_k = \sum_{a=1}^m (u_k)_a \mathbf{e}_a$  para todo  $k = 1, \dots, n$ , temos

$$\omega(u_1, \dots, u_n) = \sum_{a_1=1}^m \cdots \sum_{a_n=1}^m (u_1)_{a_1} \cdots (u_n)_{a_n} \omega(\mathbf{e}_{a_1}, \dots, \mathbf{e}_{a_n}) .$$

Conforme discutimos na Seção 2.3.5.1, o isomorfismo entre  $\mathcal{M}_n(U)$  e  $(U')^{\otimes n}$  é dado pela aplicação  $\Psi : \mathcal{M}_n(U) \rightarrow (U')^{\otimes n}$  definida por

$$\Psi(\omega) := \sum_{a_1=1}^m \cdots \sum_{a_n=1}^m \omega(\mathbf{e}_{a_1}, \dots, \mathbf{e}_{a_n}) \ell_{a_1} \otimes \cdots \otimes \ell_{a_n} . \tag{3.9}$$

Conforme discutido na Seção 2.3.5.1, essa definição é natural: independe de escolhas de base.

<sup>4</sup>Gottfried Wilhelm von Leibniz (1646–1716).

### 3.1.2 Formas Sesquilineares e as Desigualdades de Cauchy-Schwarz e Min-kowski

• **Formas sesquilineares. Definições**

Seja  $V$  um espaço vetorial complexo. Uma *forma sesquilinear*<sup>5</sup> é uma função  $\omega : V \times V \rightarrow \mathbb{C}$  que satisfaz as seguintes propriedades:

1. Linearidade em relação à segunda variável:

$$\omega(u, \alpha v + \beta w) = \alpha \omega(u, v) + \beta \omega(u, w),$$

para todos os vetores  $u, v$  e  $w$  e para todos os números complexos  $\alpha$  e  $\beta$ .

2. Antilinearidade em relação à primeira variável:

$$\omega(\alpha u + \beta v, w) = \bar{\alpha} \omega(u, w) + \bar{\beta} \omega(v, w),$$

para todos os vetores  $u, v$  e  $w$  e para todos os números complexos  $\alpha$  e  $\beta$ .

É imediato pela definição que toda forma sesquilinear  $\omega$  se anula no vetor nulo, ou seja,

$$\omega(u, 0) = \omega(0, u) = 0,$$

para todo vetor  $u$ .

**E. 3.10** *Exercício.* Prove isso. \*

Uma forma sesquilinear é dita ser uma *forma sesquilinear Hermitiana* se satisfizer:

3. Simetria por conjugação complexa:

$$\omega(u, v) = \overline{\omega(v, u)},$$

para todos os vetores  $u$  e  $v$ .

Uma forma sesquilinear é dita ser uma *forma sesquilinear positiva* se satisfizer:

4. Positividade. Para todo  $u \in V$ ,

$$\omega(u, u) \geq 0.$$

Abaixo (Teorema 3.1, página 227) provaremos que toda forma sesquilinear positiva é automaticamente Hermitiana. Lá provaremos também que se  $\omega$  é uma forma sesquilinear positiva então vale que  $|\omega(u, v)|^2 \leq \omega(u, u)\omega(v, v)$  para todos os vetores  $u$  e  $v$ . Essa desigualdade é conhecida como *Desigualdade de Cauchy-Schwarz*.

Uma forma sesquilinear é dita ser uma *forma sesquilinear não-degenerada* se satisfizer:

5. Não-degenerescência. Se um vetor  $u$  é tal que vale  $\omega(u, v) = 0$  para todo vetor  $v$ , então  $u = 0$ .

*Nomenclatura.* Uma forma sesquilinear que não é não-degenerada é dita ser degenerada.

• **A identidade de polarização para formas sesquilineares**

Se  $\omega$  é uma forma sesquilinear em  $V$ , então vale a seguinte identidade, denominada *identidade de polarização de formas sesquilineares*:

$$\omega(u, v) = \frac{1}{4} \sum_{n=0}^3 i^{-n} \omega((u + i^n v), (u + i^n v)) = \frac{1}{4} \sum_{n=0}^3 i^n \omega((u + i^{-n} v), (u + i^{-n} v)), \tag{3.10}$$

válida para todos  $u, v \in V$ . Para verificá-la, basta expandir-se o lado direito e constatar-se, com uso da sesquilinearidade de  $\omega$ , que se obtém o lado esquerdo.

---

<sup>5</sup>Do radical grego *sesqui*: um e meio.

**E. 3.11** *Exercício.* Verifique! ✦

Como no caso da identidade de polarização para formas lineares simétricas, relação (3.3), a importância da identidade de polarização (3.10) é mostrar que, no caso de formas sesquilineares, o conhecimento dos valores  $\omega(u, u)$  para todos os  $u \in V$  determina os valores  $\omega(u, v)$  para todos os  $u, v \in V$ .

• **Formas sesquilineares não-singulares**

Seja  $V$  um espaço vetorial e  $\omega$  uma forma sesquilinear em  $V$ . Para  $u \in V$  fixo a aplicação  $l_u(v) = \omega(u, v)$  é um funcional linear em  $V$ , ou seja, um elemento do espaço dual  $V'$ . Se a aplicação antilinear  $l : V \rightarrow V'$  que associa cada  $u \in V$  ao funcional linear  $l_u$  acima for um anti-isomorfismo<sup>6</sup> de espaços vetoriais a forma sesquilinear  $\omega$  é dita ser uma *forma sesquilinear não-singular*.

• **A desigualdade de Cauchy-Schwarz**

De importância fundamental na teoria das formas sesquilineares é o seguinte teorema, que apresenta-nos a importante desigualdade de Cauchy<sup>7</sup>-Schwarz<sup>8</sup>.

**Teorema 3.1** *Se  $\omega$  é uma forma sesquilinear positiva, então é também Hermitiana, ou seja,*

$$\omega(u, v) = \overline{\omega(v, u)},$$

para todos os vetores  $u$  e  $v$ . Fora isso, vale a desigualdade de Cauchy-Schwarz: para todos os vetores  $u$  e  $v$ ,

$$|\omega(u, v)|^2 \leq \omega(u, u)\omega(v, v). \tag{3.11}$$

Por fim, se  $\omega$  é uma forma sesquilinear positiva e não-degenerada então  $\omega(u, u) = 0$  se e somente se  $u = 0$ . □

*Prova.* Faremos uso do fato que, para qualquer número complexo  $\lambda$  e quaisquer vetores  $u$  e  $v$  vale, pela hipótese de positividade,

$$\omega(u + \lambda v, u + \lambda v) \geq 0.$$

Escrevendo-se explicitamente o lado esquerdo temos a desigualdade

$$|\lambda|^2 \omega(v, v) + \lambda \omega(u, v) + \bar{\lambda} \omega(v, u) + \omega(u, u) \geq 0.$$

**E. 3.12** *Exercício.* Verifique isso. ✦

Vamos agora escrever  $\lambda$  na forma  $\lambda = x + iy$ , onde  $x$  é a parte real de  $\lambda$  e  $y$  sua parte imaginária. A última expressão fica  $f(x, y) \geq 0$ , onde

$$f(x, y) := (x^2 + y^2)\omega(v, v) + (x + iy)\omega(u, v) + (x - iy)\omega(v, u) + \omega(u, u).$$

**E. 3.13** *Exercício.* Verifique essa afirmação. ✦

Vamos decompor  $\omega(u, v)$  e  $\omega(v, u)$  nas suas partes reais e imaginárias, escrevendo

$$\omega(u, v) = \alpha + i\beta \quad \text{e} \quad \omega(v, u) = \gamma + i\delta, \tag{3.12}$$

onde  $\alpha, \beta, \gamma$  e  $\delta \in \mathbb{R}$ . Ficamos com

$$f(x, y) = (x^2 + y^2)\omega(v, v) + (x\alpha - y\beta) + i(x\beta + y\alpha) + (x\gamma + y\delta) + i(x\delta - y\gamma) + \omega(u, u) \geq 0. \tag{3.13}$$

<sup>6</sup>Definido à página 124.

<sup>7</sup>Augustin Louis Cauchy (1789–1857).

<sup>8</sup>Herman Amandus Schwarz (1843–1921).



Como  $f(x, y)$  tem de ser real ( $e \geq 0$ ), segue que a parte imaginária da expressão acima deve ser nula e, como  $\omega(v, v)$  e  $\omega(u, u)$  são reais, devemos ter

$$0 = (x\beta + y\alpha) + (x\delta - y\gamma) = x(\beta + \delta) + y(\alpha - \gamma).$$

Como isso deve valer para todos  $x, y \in \mathbb{R}$ , segue que  $\beta = -\delta$  e  $\alpha = \gamma$ . Comparando com (3.12), isso diz que

$$\omega(u, v) = \overline{\omega(v, u)},$$

provando que  $\omega$  é Hermitiano.

Com as relações  $\beta = -\delta$  e  $\alpha = \gamma$  a expressão (3.13) fica

$$f(x, y) = (x^2 + y^2)\omega(v, v) + 2(x\alpha - y\beta) + \omega(u, u). \tag{3.14}$$

Vamos agora considerar dois casos: um onde  $\omega(v, v) = 0$  e outro onde  $\omega(v, v) \neq 0$ . No primeiro

$$f(x, y) = 2(x\alpha - y\beta) + \omega(u, u).$$

Assim, como  $\omega(u, u) \geq 0$  pela positividade, a condição  $f(x, y) \geq 0$  é possível para todos  $x$  e  $y \in \mathbb{R}$  se e somente se  $\alpha = \beta = 0$ , ou seja, se e somente se  $\omega(u, v) = 0$  para todo  $u$ . Aqui a desigualdade de Cauchy-Schwarz (3.11) é trivialmente satisfeita, pois ambos os lados são iguais a zero.

Passemos ao caso  $\omega(v, v) \neq 0$ . Resta-nos provar a desigualdade de Cauchy-Schwarz (3.11) para esse caso. Podemos reescrever o lado direito de (3.14) como

$$f(x, y) = \omega(v, v) \left[ \left( x + \frac{\alpha}{\omega(v, v)} \right)^2 + \left( y - \frac{\beta}{\omega(v, v)} \right)^2 \right] + \omega(u, u) - \left( \frac{\alpha^2 + \beta^2}{\omega(v, v)} \right).$$

**E. 3.14** *Exercício.* Verifique. ✱

Daí, constatamos que  $f(x, y) \geq 0$  para todos  $x$  e  $y \in \mathbb{R}$  se e somente se

$$\omega(u, u) - \left( \frac{\alpha^2 + \beta^2}{\omega(v, v)} \right) \geq 0,$$

ou seja, se e somente se

$$\omega(u, u)\omega(v, v) \geq \alpha^2 + \beta^2.$$

O lado direito é, porém,  $|\omega(u, v)|^2$ , e a última desigualdade significa

$$|\omega(u, v)|^2 \leq \omega(u, u)\omega(v, v),$$

que é a desigualdade de Cauchy-Schwarz que queríamos demonstrar.

Finalmente, se  $\omega$  é uma forma sesquilinear positiva e não-degenerada e um certo vetor  $u$  é tal que  $\omega(u, u) = 0$ , segue pela desigualdade de Cauchy-Schwarz que  $\omega(u, v) = 0$  para todo  $v$ , o que implica  $u = 0$ , pois  $\omega$  é não-degenerada. ■

• **A desigualdade de Minkowski**

A desigualdade de Cauchy-Schwarz tem uma consequência de certa importância, a chamada *desigualdade de Minkowski*<sup>9</sup>: Se  $\omega$  é uma forma sesquilinear positiva (em particular, se  $\omega$  é um produto escalar), então, para todos os vetores  $u$  e  $v$ , vale

$$\omega(u - v, u - v)^{1/2} \leq \omega(u, u)^{1/2} + \omega(v, v)^{1/2}. \tag{3.15}$$

---

<sup>9</sup>Hermann Minkowski (1864–1909).

A demonstração é simples:

$$\begin{aligned} \omega(u - v, u - v) &= \omega(u, u) - \omega(u, v) - \omega(v, u) + \omega(v, v) \\ &= \omega(u, u) - 2\operatorname{Re}(\omega(u, v)) + \omega(v, v) \\ &\leq \omega(u, u) + 2|\omega(u, v)| + \omega(v, v) \\ &\leq \omega(u, u) + 2\omega(u, u)^{1/2}\omega(v, v)^{1/2} + \omega(v, v) \\ &= \left[ \omega(u, u)^{1/2} + \omega(v, v)^{1/2} \right]^2, \end{aligned}$$

que é o que se queria demonstrar. Acima, na passagem da primeira para a segunda linha usamos a Hermiticidade de  $\omega$  e na passagem da terceira para a quarta linha, usamos a desigualdade de Cauchy-Schwarz, ambos esses fatos sendo consequência do Teorema 3.1, página 227.

### 3.1.3 Produtos Escalares

#### • Produtos internos ou produtos escalares

Uma forma sesquilinear positiva  $\omega$  é dita ser um *produto escalar*, ou *produto interno*, se satisfizer:

6.  $\omega(u, u) = 0$  se e somente se  $u = 0$ .

A proposição seguinte apresenta uma definição alternativa de produto escalar.

**Proposição 3.2** *Uma forma sesquilinear positiva é um produto escalar se e somente se for não-degenerada.* □

*Prova.* Se  $\omega$  é um produto escalar, então se  $u$  é tal que  $\omega(u, v) = 0$  para todo  $v$ , vale em particular (tomando  $v = u$ ) que  $\omega(u, u) = 0$  e, portanto,  $u = 0$ . Assim, todo o produto escalar é não-degenerado. Reciprocamente, pelo Teorema 3.1, página 227, se  $\omega$  é uma forma sesquilinear positiva e não-degenerada, então vale automaticamente que  $\omega(u, u) = 0$  se e somente se  $u = 0$  ■

#### • Notações para produtos escalares

Seguindo a convenção, denotaremos frequentemente produtos escalares de dois vetores  $u$  e  $v$  não por  $\omega(u, v)$  mas por  $\langle u, v \rangle$ . É frequente também denotar um produto escalar de dois vetores  $u$  e  $v$  por  $(u, v)$ . Essa notação pode causar confusão com a de par ordenado e por isso a evitamos. Em textos de Física é comum encontrar também a chamada *notação de Dirac* para produtos escalares:  $\langle u|v \rangle$ . Por diversas razões não compartilhamos do entusiasmo de alguns com essa notação e também a evitamos.

#### • Detalhando a definição de produto escalar

Como o conceito de produto escalar é muito importante, vamos detalhá-lo um pouco mais antes de passarmos a exemplos.

Um *produto escalar*, ou *produto interno*, em um espaço vetorial  $V$  sobre o corpo dos complexos é uma função  $V \times V \rightarrow \mathbb{C}$ , denotada por  $\langle u, v \rangle$ , para  $u, v \in V$ , com as seguintes propriedades:

1. O produto escalar é linear na segunda variável:

$$\langle u, \alpha v + \beta w \rangle = \alpha \langle u, v \rangle + \beta \langle u, w \rangle$$

para todos  $u, v$  e  $w \in V$  e todos  $\alpha, \beta \in \mathbb{C}$ .

2. O produto escalar é antilinear na primeira variável:

$$\langle \alpha u + \beta v, w \rangle = \bar{\alpha} \langle u, w \rangle + \bar{\beta} \langle v, w \rangle$$

para todos  $u, v$  e  $w \in V$  e todos  $\alpha, \beta \in \mathbb{C}$ , onde  $\bar{\alpha}$  é o complexo conjugado de  $\alpha \in \mathbb{C}$ .

3. Conjugação complexa:

$$\langle u, v \rangle = \overline{\langle v, u \rangle}$$

para todos  $u, v \in V$ .

4. Para todo  $u \in V$

$$\langle 0, u \rangle = \langle u, 0 \rangle = 0.$$

5. Positividade. Para todo vetor  $u$  não-nulo

$$\langle u, u \rangle > 0.$$

*Nota.* Alguns postulados da definição de produto escalar acima são redundantes, pois nem todos são independentes. Nós os listamos apenas para ressaltar sua relevância individual. Por exemplo, o item 2 segue de 1 e 3 (por quê?). O item 4 segue de 1 e 2 (por quê?). Os itens 1, 2 e 5 implicam o item 3 (como veremos no Teorema 3.1). Independentes são apenas 1, 2 e 5 ou 1, 3 e 5. ♣

Para um produto escalar de dois vetores vale a seguinte e importantíssima desigualdade, conhecida como *Desigualdade de Cauchy-Schwarz*:

$$|\langle u, v \rangle|^2 \leq |\langle u, u \rangle| |\langle v, v \rangle|.$$

A demonstração (mais geral) é apresentada no Teorema 3.1, página 227.

**Advertência.** Em textos de Matemática a definição de produto escalar é, por vezes, apresentada de forma que se tenha linearidade na primeira variável e antilinearidade na segunda variável. A convenção que adotamos é oposta a essa e é seguida, felizmente, por 100% dos textos de Física.

• **Formas sesquilineares positivas e produtos escalares**

Se  $V$  é um espaço vetorial dotado de uma forma sesquilinear positiva  $\omega$ , existe uma maneira canônica de construir a partir de  $V$  e  $\omega$  um outro espaço vetorial dotado de um produto escalar.

Seja  $\omega$  uma forma sesquilinear positiva em um espaço vetorial  $V$ . Então, existe um espaço vetorial  $\tilde{V}$ , um produto escalar  $\tilde{\omega}$  e uma aplicação linear sobrejetora  $E : V \rightarrow \tilde{V}$  tais que

$$\tilde{\omega}(E(u), E(v)) = \omega(u, v)$$

e que  $E(u) = 0$  em  $\tilde{V}$  caso  $\omega(u, u) = 0$ .

Para a mencionada construção, notemos em primeiro lugar que o conjunto de todos os vetores  $u$  com a propriedade que  $\omega(u, u) = 0$  formam um subespaço de  $V$ . De fato, se  $u$  e  $v$  são dois vetores desse tipo, teremos que

$$\omega(\alpha u + \beta v, \alpha u + \beta v) = |\alpha|^2 \omega(u, u) + \bar{\alpha} \beta \omega(u, v) + \alpha \bar{\beta} \omega(v, u) + |\beta|^2 \omega(v, v) = 0,$$

pois  $\omega(u, u) = \omega(v, v) = 0$ , por hipótese, e pois  $\omega(v, u) = \omega(u, v) = 0$  em função da condição de  $\omega$  ser positivo (pela desigualdade de Cauchy-Schwarz). Vamos denominar esse subespaço por  $Z$ . O espaço vetorial quociente  $\tilde{V} = V/Z$  (vide a construção da página 170) tem as propriedades desejadas. A aplicação  $E : V \rightarrow \tilde{V}$  é a aplicação que associa cada elemento de  $v$  de  $V$  à sua classe de equivalência  $[v] : E : V \ni v \mapsto [v] \in \tilde{V}$ . Definimos então  $\tilde{\omega}$  por

$$\tilde{\omega}([u], [v]) = \omega(u, v).$$

É um exercício simples (faça) mostrar que essa definição de fato independe dos representantes, no caso  $u$  e  $v$ , tomados nas classes  $[u]$  e  $[v]$ .

**E. 3.15** *Exercício.* Mostre que  $\tilde{\omega}$  é de fato um produto escalar em  $\tilde{V}$ . ♣

• **Produtos escalares e formas simpléticas reais**

Seja  $V$  um espaço vetorial complexo dotado de um produto escalar  $\langle \cdot, \cdot \rangle$ . Então, a expressão

$$\omega(u, v) := \text{Im}(\langle u, v \rangle),$$

$u, v \in V$ , define uma forma simplética real em  $V$ . As condições de antissimetria ( $\omega(u, v) = -\omega(v, u)$ ) e de linearidade por combinações lineares com escalares reais são elementares de se constatar. Que  $\omega$  é não-degenerada, segue do fato que se  $\omega(u, v) = 0$  para todo  $u$  valeria, tomando  $u = -iv$ ,  $0 = \text{Im}(\langle -iv, v \rangle) = \langle v, v \rangle$ , o que implica  $v = 0$ .

Na Seção 3.5, página 247, veremos que, sob hipóteses adequadas, toda forma simplética real é a parte imaginária de um produto escalar em um espaço complexo.

### 3.1.4 Formas Quadráticas em Espaços Vetoriais Reais

Seja  $V$  um espaço vetorial sobre os reais. Uma função  $Q : V \rightarrow \mathbb{R}$  que satisfaça

1.  $Q(\lambda v) = \lambda^2 Q(v)$ , para todos  $\lambda \in \mathbb{R}$  e todo  $v \in V$ ,
2. a aplicação  $(u, v) \in V^2 \mapsto Q(u + v) - Q(u) - Q(v)$  é bilinear,

é dita ser uma *forma quadrática (real) em  $V$* .

Um exemplo importante é o seguinte: seja  $B : V^2 \mapsto \mathbb{R}$  uma forma bilinear simétrica em  $V$  e defina-se

$$Q(u) = B(u, u), \quad u \in V$$

Então essa  $Q$  é uma forma quadrática em  $V$ . De fato, para  $\lambda \in \mathbb{R}$ ,  $Q(\lambda u) = B(\lambda u, \lambda u) = \lambda^2 B(u, u) = \lambda^2 Q(u)$  e  $Q(u + v) - Q(u) - Q(v) = B(u + v, u + v) - B(u, u) - B(v, v) = B(u, v) + B(v, u) = 2B(u, v)$  (devido à suposta simetria de  $B$ ). Agora,  $2B$  é, evidentemente, uma forma bilinear, se  $B$  o for.

A afirmação recíproca é igualmente verdadeira: se  $Q$  é uma forma quadrática, então  $B(u, v) := Q(u + v) - Q(u) - Q(v)$  é uma forma bilinear simétrica. A simetria é evidente e a bilinearidade segue das hipóteses sobre  $Q$ .

A conclusão é que toda forma quadrática real procede de uma forma bilinear simétrica e vice-versa.

Essa última propriedade inspira como definir a noção de forma quadrática no caso complexo. Se  $V$  é um espaço vetorial complexo e  $S : V^2 \rightarrow \mathbb{C}$  é uma forma sesquilinear Hermitiana, então  $Q : V \rightarrow \mathbb{C}$ , definida por  $Q(u) = S(u, u)$ ,  $u \in V$ , é dita ser forma quadrática sobre  $V$ .

### 3.1.5 Exemplos

Para ilustrar os conceitos apresentados acima, passemos a alguns exemplos.

• **Exemplos de formas sesquilineares e produtos escalares**

**Exemplo 3.1** Seja  $V = \mathbb{C}^n$ . Um exemplo de produto escalar é dado pelo produto escalar usual:

$$\omega(u, v) = \langle u, v \rangle_{\mathbb{C}} := \sum_{k=1}^n \overline{u_k} v_k, \tag{3.16}$$

onde  $u = (u_1, \dots, u_n)$  e  $v = (v_1, \dots, v_n)$ . □

**Exemplo 3.2** Seja  $V = \mathbb{C}^n$ . Um exemplo de produto escalar é dado por

$$\omega(u, v) = \langle Au, Av \rangle_{\mathbb{C}},$$

onde  $u = (u_1, \dots, u_n)$ ,  $v = (v_1, \dots, v_n)$  e onde  $A$  é uma matriz  $n \times n$  *invertível*. □

**Exemplo 3.3** Exemplo de uma forma sesquilinear Hermitiana que não é positiva. Seja  $V = \mathbb{C}^n$  e seja  $\omega$  dado por

$$\omega(u, v) = \langle u, Av \rangle_{\mathbb{C}} = \sum_{k, l=1}^n \overline{u_k} A_{kl} v_l,$$

onde  $A$  é uma matriz  $n \times n$  *autoadjunta*, ou seja, seus elementos de matriz satisfazem  $A_{kl} = \overline{A_{lk}}$ . Assim definida  $\omega$  é uma forma sesquilinear Hermitiana, mas em geral pode não ser positiva. Um caso concreto é o seguinte. Tomemos  $V = \mathbb{C}^2$  e  $A = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$ . Então, é fácil ver que  $\omega(u, u) = \langle u, Au \rangle_{\mathbb{C}} = i(u_1 \overline{u_2} - \overline{u_1} u_2) = -2\text{Im}(u_1 \overline{u_2})$ , que pode ser negativo ou mesmo nulo. Assim, essa  $\omega$  não é positiva. É fácil ver, porém, que essa  $\omega$  é não-degenerada (mostre isso!).  $\square$

**Exemplo 3.4** Exemplo de uma forma sesquilinear que não é Hermitiana. Seja  $V = \mathbb{C}^n$  e seja dado por

$$\omega(u, v) = \langle u, Av \rangle_{\mathbb{C}} = \sum_{k, l=1}^n \overline{u_k} A_{kl} v_l,$$

onde  $A$  é uma matriz  $n \times n$  que não é *autoadjunta*, ou seja,  $A_{kl} \neq \overline{A_{lk}}$  para pelo menos um elemento de matriz  $A_{kl}$ . Assim definida  $\omega$  é uma forma sesquilinear, mas em geral pode não ser Hermitiana. Um caso concreto é o seguinte. Tomemos  $V = \mathbb{C}^2$  e  $A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ . Então, é fácil ver que  $\omega(u, v) = \langle u, Av \rangle_{\mathbb{C}} = \overline{u_1} v_2$ , enquanto que  $\omega(v, u) = \overline{v_1} u_2$ . Logo,  $\omega(u, v)$  e  $\overline{\omega(v, u)}$  podem ser distintos e  $\omega$  não é Hermitiana. Fora isso, essa  $\omega$  também não é positiva e é degenerada (mostre isso!).  $\square$

**Exemplo 3.5** Exemplo de uma forma sesquilinear positiva mas que não é um produto escalar. Seja  $V = \mathbb{C}^n$  e seja  $\omega$  dado por

$$\omega(u, v) = \langle Au, Av \rangle_{\mathbb{C}}$$

onde  $A$  é uma matriz  $n \times n$  *não-inversível*. Então, existe  $u_0$  não-nulo tal que  $Au_0 = 0$ . Daí, segue que  $\omega(u_0, v) = \langle Au_0, Av \rangle_{\mathbb{C}} = 0$  para todo  $v$  e, portanto,  $\omega$  é degenerada e  $\omega(u_0, u_0) = 0$ .

Um caso concreto é o seguinte. Tomemos  $V = \mathbb{C}^2$  e  $A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ . Note que  $A$  não é inversível (por quê?). Aqui temos que  $\omega(u, v) = \overline{u_1} v_1$ . Note que todo vetor da forma  $u^b = \begin{pmatrix} 0 \\ u_2 \end{pmatrix}$  é tal que  $Au^b = 0$  e, portanto,  $\omega(u^b, v) = 0$  para todo  $v$ .  $\square$

Na Seção 3.4, página 244, mostraremos como é a forma geral de formas bilineares, sesquilineares e produtos escalares nos espaços de dimensão finita  $\mathbb{R}^n$  e  $\mathbb{C}^n$ . Tratemos agora de dois exemplos em espaços vetoriais de dimensão infinita.

**Exemplo 3.6** Seja  $V = C([a, b])$  o espaço vetorial das funções contínuas complexas de um intervalo fechado  $[a, b]$  da reta real ( $a < b$ ). Seja  $p$  uma função contínua estritamente positiva definida em  $[a, b]$ , ou seja,  $p(x) > 0$  para todo  $x \in [a, b]$ . Então, a expressão

$$\omega(f, g) = \int_a^b \overline{f(x)} g(x) p(x) dx,$$

para funções  $f$  e  $g$  de  $V$  define um produto escalar em  $V$  (justifique!).  $\square$

**Exemplo 3.7** Seja  $V = C([0, 1])$  o espaço vetorial das funções contínuas complexas de um intervalo fechado  $[0, 1]$  da reta real. Seja  $p$  uma função tal que  $p$  é contínua e estritamente positiva no intervalo  $[0, 1/2]$  e identicamente nula no intervalo  $[1/2, 1]$ . Então, a expressão

$$\omega(f, g) = \int_0^1 \overline{f(x)} g(x) p(x) dx,$$

para funções  $f$  e  $g$  de  $V$  define uma forma sesquilinear positiva em  $V$ , que não é um produto escalar (justifique!).  $\square$

**Exemplo 3.8** Considere o espaço vetorial  $\mathbb{C}^n$  e o produto escalar usual:  $\omega(u, v) = \langle u, v \rangle_{\mathbb{C}} = \sum_{i=1}^n \overline{u_i} v_i$ . A desigualdade de Cauchy-Schwarz implica

$$\left| \sum_{i=1}^n \overline{u_i} v_i \right|^2 \leq \left( \sum_{j=1}^n |u_j|^2 \right) \left( \sum_{k=1}^n |v_k|^2 \right). \tag{3.17}$$

$\square$

**E. 3.16** *Exercício.* Considere o espaço vetorial das funções contínuas no intervalo  $[0, 1]$  e o produto escalar  $\omega(f, g) = \int_0^1 \overline{f(x)}g(x) dx$ . Tomando as funções  $f(x) = x$  e  $g(x) = e^x$ , use a desigualdade de Cauchy-Schwarz para mostrar que  $e \geq \sqrt{7}$ . \*

**E. 3.17** *Exercício.* Tente livremente obter outras desigualdades interessantes do mesmo estilo usando esse método. \*

## 3.2 Normas em Espaços Vetoriais

Aqui trataremos sempre, exceto se mencionado de outra forma, de espaços vetoriais sobre o corpo dos complexos.

### • Seminormas

Uma *seminorma* é uma função  $V \rightarrow \mathbb{R}$  usualmente denotada por  $\| \cdot \|$ , com as seguintes propriedades:

1. Para todo  $v \in V$  tem-se  $\|v\| \geq 0$ .
2. Para qualquer  $\alpha \in \mathbb{C}$  e qualquer  $v \in V$  tem-se  $\|\alpha v\| = |\alpha| \|v\|$ .
3. Para quaisquer vetores  $u$  e  $v \in V$  tem-se  $\|u + v\| \leq \|u\| + \|v\|$ . Essa desigualdade é denominada *desigualdade triangular*.

### Notas.

- Note-se que, pelo item 2, vale para uma seminorma que  $\|0\| = 0$  (tome  $\alpha = 0$ ).
- Para uma seminorma vale a desigualdade

$$\|a\| \geq \left| \|a - b\| - \|b\| \right|, \tag{3.18}$$

para quaisquer  $a, b \in V$ . Como faremos uso da mesma no futuro, vamos apresentar sua demonstração aqui, que é uma consequência direta da desigualdade triangular. De fato, a desigualdade triangular diz-nos que

$$\|a - b\| \leq \|a\| + \|b\| \tag{3.19}$$

e que

$$\|b\| = \|a - (a - b)\| \leq \|a\| + \|a - b\|. \tag{3.20}$$

De (3.19) segue que

$$\|a\| \geq \|a - b\| - \|b\|$$

e de (3.20) que

$$\|a\| \geq -(\|a - b\| - \|b\|).$$

Quando dois números reais  $x$  e  $y$  são tais que  $x \geq y$  e  $x \geq -y$  então  $x \geq |y|$ . Assim, as duas últimas desigualdades dizem que

$$\|a\| \geq \left| \|a - b\| - \|b\| \right|,$$

que é o que queríamos provar.

Essa desigualdade diz, incidentalmente, que  $\|a\| \geq 0$  para todo vetor de  $V$ . Isso mostra que o item 1 da definição de seminorma (e de norma, vide abaixo) é supérfluo.

- Note-se também que se fizermos em (3.18) as substituições  $a \rightarrow a - b$ ,  $b \rightarrow -b$ , obtemos

$$\left| \|a\| - \|b\| \right| \leq \|a - b\|, \tag{3.21}$$

para quaisquer  $a, b \in V$ . Essa desigualdade será empregada diversas vezes neste texto.

- Pelos itens 2 e 3 da definição de seminorma, vale que

$$\|\alpha u + \beta v\| \leq |\alpha| \|u\| + |\beta| \|v\| \tag{3.22}$$

para quaisquer  $\alpha, \beta \in \mathbb{C}$  e quaisquer vetores  $u$  e  $v \in V$ .



### • Normas

Uma *norma* é uma função  $V \rightarrow \mathbb{R}$  usualmente denotada por  $\| \cdot \|$ , com as seguintes propriedades:

1. Para todo  $v \in V$  tem-se  $\|v\| \geq 0$ .

2.  $\|v\| = 0$  se e somente se  $v$  for o vetor nulo:  $v = 0$ .
3. Para qualquer  $\alpha \in \mathbb{C}$  e qualquer  $v \in V$  tem-se  $\|\alpha v\| = |\alpha|\|v\|$ .
4. Para quaisquer vetores  $u$  e  $v \in V$  tem-se  $\|u + v\| \leq \|u\| + \|v\|$ .

Notas.

- Como se percebe, uma norma é uma seminorma dotada também da propriedade que  $\|v\| = 0$  implica  $v = 0$ .
- Note também que, pelo item 3 acima, tem-se  $\|0\| = 0$  (tome  $\alpha = 0$ ).
- Pelos itens 3 e 4 da definição de norma, vale que

$$\|\alpha u + \beta v\| \leq |\alpha|\|u\| + |\beta|\|v\| \tag{3.23}$$

para quaisquer  $\alpha, \beta \in \mathbb{C}$  e quaisquer vetores  $u$  e  $v \in V$ .

- Como toda norma é uma seminorma, vale também a importante desigualdade

$$\left| \|a\| - \|b\| \right| \leq \|a - b\|, \tag{3.24}$$

para quaisquer  $a, b \in V$ . Essa desigualdade será empregada diversas vezes neste texto.

- As quatro condições da definição de norma, acima, não são, em verdade, logicamente independentes e listamo-as devido à sua importância individual. Assim, por exemplo, a condição de positividade 1, como no caso de seminormas, segue das condições 3 e 4 (mais precisamente, de (3.24)).
- A condição 4, acima, é de particular importância e é denominada *desigualdade triangular*.



Um espaço vetorial pode ter várias normas. Vide exemplos abaixo.

**Exemplos de normas em espaços vetoriais**

Seja  $\mathbb{C}^n = \{(z_1, \dots, z_n), \text{ com } z_1, \dots, z_n \in \mathbb{C}\}$ ,  $n \geq 1$ , o espaço vetorial das  $n$ -uplas de números complexos. Para  $z = (z_1, \dots, z_n) \in \mathbb{C}^n$ , a expressão

$$\|z\|_1 := \sum_{k=1}^n |z_k| \tag{3.25}$$

define uma norma em  $\mathbb{C}^n$ , denominada *norma  $\ell_1$* . Verifique! A expressão

$$\|z\|_\infty := \max\{|z_1|, \dots, |z_n|\} \tag{3.26}$$

também define uma norma em  $\mathbb{C}^n$ . Verifique!

A norma (3.25) pode ser generalizada. Para cada  $p \in \mathbb{R}$ ,  $p \geq 1$ , a expressão

$$\|z\|_p := \left[ \sum_{k=1}^n |z_k|^p \right]^{\frac{1}{p}} \tag{3.27}$$

também define uma norma em  $\mathbb{C}^n$ , denominada *norma  $\ell_p$* . A única dificuldade em provar isso reside em demonstrar a desigualdade triangular  $\|z + w\|_p \leq \|z\|_p + \|w\|_p$  para quaisquer  $z, w \in \mathbb{C}^n$ . Essa desigualdade, porém, é precisamente a *desigualdade de Minkowski*, demonstrada na Proposição 5.20, página 307. Vide também a Seção 24.5.1, página 1274 (especificamente, a expressão (24.42) do Teorema 24.4, página 1278).

Seja  $C([a, b], \mathbb{C})$  o espaço vetorial das funções complexas contínuas definidas no intervalo  $[a, b] \subset \mathbb{R}$ . A expressão

$$\|f\|_1 := \int_a^b |f(x)| dx, \tag{3.28}$$

$f \in C([a, b], \mathbb{C})$ , define uma norma em  $C([a, b], \mathbb{C})$ , denominada *norma  $L_1$* . Verifique! A expressão

$$\|f\|_\infty := \sup_{x \in [a, b]} |f(x)|, \tag{3.29}$$

$f \in C([a, b], \mathbb{C})$ , também define uma norma em  $C([a, b], \mathbb{C})$ , denominada *norma do supremo*. Verifique!

A norma (3.28) pode ser generalizada. Para cada  $p \in \mathbb{R}$ ,  $p \geq 1$ , a expressão

$$\|f\|_p := \left[ \int_a^b |f(x)|^p dx \right]^{\frac{1}{p}}, \tag{3.30}$$

$f \in C([a, b], \mathbb{C})$ , define uma norma em  $C([a, b], \mathbb{C})$ , denominada *norma  $L_p$* . A única dificuldade em provar isso reside em demonstrar a desigualdade triangular  $\|f + g\|_p \leq \|f\|_p + \|g\|_p$  para quaisquer  $f, g \in C([a, b], \mathbb{C})$ . Para uma demonstração, vide Seção 5.3.3.1, página 307, ou, com mais generalidade (para funções em espaços mensuráveis), vide a Seção 31.4.1, página 1489 (especificamente, vide a expressão (31.45) do Teorema 31.7, página 1489).

• **Equivalência de normas**

**Definição.** Duas normas  $\|\cdot\|_1$  e  $\|\cdot\|_2$  em um espaço vetorial  $V$  são ditas equivalentes se existirem duas constantes positivas  $c_1$  e  $c_2$ , com  $0 < c_1 \leq c_2$ , tais que

$$c_1\|v\|_1 \leq \|v\|_2 \leq c_2\|v\|_1$$

para todo vetor  $v \in V$ . A importância da noção de equivalência de normas se manifesta no fato que duas normas equivalentes geram a mesma topologia métrica. ♠

**E. 3.18 Exercício.** Mostre que a relação de equivalência entre normas é uma relação de equivalência. \*

**E. 3.19 Exercício.** Mostre que as normas  $\|\cdot\|_1$  e  $\|\cdot\|_\infty$  do espaço  $\mathbb{C}^n$ , definidas em (3.25) e (3.26), respectivamente, são equivalentes. \*

Em espaços vetoriais reais ou complexos de dimensão finita vale o seguinte resultado especial, cuja demonstração encontra-se no Apêndice 3.A, página 254:

**Teorema 3.2** *Em um espaço vetorial de dimensão finita sobre  $\mathbb{C}$  ou  $\mathbb{R}$  todas as normas são equivalentes.* □

A afirmação do Teorema 3.2 é frequentemente falsa em espaços de dimensão infinita. Isso é atestado nos exemplos do Exercício E. 3.20.

**E. 3.20 Exercício.** As normas  $\|\cdot\|_1$  e  $\|\cdot\|_\infty$  do espaço  $C([a, b], \mathbb{C})$ , definidas em (3.28) e (3.29), respectivamente, não são equivalentes. É fácil ver que  $\|f\|_1 \leq (b - a)\|f\|_\infty$  para toda  $f \in C([a, b], \mathbb{C})$  (faça!). Seja, porém, a família de funções  $f_\alpha(x) = e^{-\alpha(x-a)} \in C([a, b], \mathbb{C})$  com  $\alpha > 0$ . É fácil ver que  $\|f_\alpha\|_\infty = 1$  e  $\|f_\alpha\|_1 = \frac{1}{\alpha}(1 - e^{-\alpha(b-a)})$  (faça!). Mostre que não existe nenhuma constante  $c$  tal que  $\|f_\alpha\|_\infty \leq c\|f_\alpha\|_1$  para todo  $\alpha > 0$ . \*

• **Equivalência entre seminormas**

Há uma noção de equivalência entre seminormas análoga à de equivalência entre normas.

• **A norma associada a um produto escalar**

Se  $\omega$  é um produto escalar em um espaço vetorial  $V$  existe associada a  $\omega$  uma norma  $\|\cdot\|_\omega$  dada por

$$\|v\|_\omega = \omega(v, v)^{1/2},$$

$v \in V$ .

**E. 3.21 Exercício.** Mostre que os postulados da definição de norma são de fato satisfeitos. \*

• **Invariância de normas associadas a produtos escalares**

Se uma norma em um espaço vetorial  $V$  é produzida por um produto escalar, como acima, existe naturalmente um grupo de transformações lineares de  $V$  em  $V$  que mantém essa norma invariante. Esse grupo é discutido na Seção 21.3.3,



página 1051. Por exemplo, a chamada *norma Euclidiana* de  $\mathbb{R}^n$ , definida por  $\|x\| = \sqrt{\langle x, x \rangle_{\mathbb{R}}}$  para  $x \in \mathbb{R}^n$ , é invariante pelo grupo  $O(n)$  das matrizes ortogonais, ou seja, das matrizes  $R$ , reais  $n \times n$ , que satisfazem  $R^T R = \mathbb{1}$ . Isso significa que  $\|Rx\| = \|x\|$  para toda  $R \in O(n)$ . O grupo  $O(n)$  e seus amigos são discutidos na Seção 21.3.3.1, página 1055 e seguintes.

• **A desigualdade triangular**

Talvez a principal consequência da desigualdade de Minkowski (3.15) seja a seguinte. Vamos supor que  $\omega$  seja um produto escalar. Então, podemos definir<sup>10</sup> uma *métrica* ou *distância* entre dois vetores  $a$  e  $b$  por

$$d_\omega(a, b) := \|a - b\|_\omega = \omega(a - b, a - b)^{1/2}.$$

Como  $\omega$  é um produto escalar, segue que  $d_\omega(a, b) = 0$  se e somente se  $a = b$  (por quê?). É também claro que  $d_\omega(a, b) = d_\omega(b, a)$  (por quê?). Fora isso, segue da desigualdade de Minkowski que para quaisquer vetores  $a, b$  e  $c$  vale

$$d_\omega(a, b) \leq d_\omega(a, c) + d_\omega(c, b).$$

Para ver isso, note que

$$\begin{aligned} d_\omega(a, b) &= \omega(a - b, a - b)^{1/2} \\ &= \omega((a - c) - (b - c), (a - c) - (b - c))^{1/2} \\ &\leq \omega(a - c, a - c)^{1/2} + \omega(b - c, b - c)^{1/2} \\ &= d_\omega(a, c) + d_\omega(c, b). \end{aligned}$$

Acima, na passagem da segunda à terceira linha, usamos a desigualdade de Minkowski com  $u = a - b$  e  $v = b - c$ .

A desigualdade  $d_\omega(a, b) \leq d_\omega(a, c) + d_\omega(c, b)$  é importante no estudo de propriedades topológicas de espaços vetoriais e é denominada *desigualdade triangular* (pergunta ao estudante: de onde vem esse nome?).

Note que a desigualdade triangular vale também se  $\omega$  não for um produto escalar, mas apenas uma forma sesquilinear positiva (por quê?). Nesse caso é também verdade que  $d_\omega(a, b) = d_\omega(b, a)$ , porém, não é mais verdade que  $d_\omega(a, b) = 0$  se e somente se  $a = b$  e, por isso,  $d_\omega$  é dita ser uma *pseudométrica*.

• **Norma e produto escalar**

Se um espaço vetorial  $V$  possuir um produto escalar então, como observamos, é possível definir nele uma norma da seguinte forma:  $\|u\| = \sqrt{\langle u, u \rangle}$ ,  $u \in V$ .

A norma assim definida possui duas propriedades importantes que mencionamos aqui: a *identidade do paralelogramo* e a *identidade de polarização*.

**Identidade do paralelogramo:** Para todos os vetores  $u, v \in V$  vale

$$\|u + v\|^2 + \|u - v\|^2 = 2\|u\|^2 + 2\|v\|^2. \tag{3.31}$$

*Prova.* Tem-se simplesmente pelas definições que

$$\|u + v\|^2 = \langle u + v, u + v \rangle = \|u\|^2 + \langle u, v \rangle + \langle v, u \rangle + \|v\|^2$$

e

$$\|u - v\|^2 = \langle u - v, u - v \rangle = \|u\|^2 - \langle u, v \rangle - \langle v, u \rangle + \|v\|^2.$$

Somando-se ambas tem-se o resultado desejado.

**E. 3.22 Exercício.** Por que (3.31) é chamada “identidade do paralelogramo”? \*

---

<sup>10</sup>As noções de métrica e de espaços métricos serão discutidas no Capítulo 24, página 1245.

**E. 3.23** *Exercício*. Usando a identidade do paralelogramo demonstre a *identidade de Apolônio*<sup>11</sup>:

$$\|z - x\|^2 + \|z - y\|^2 = \frac{1}{2}\|x - y\|^2 + 2\left\|z - \frac{(x + y)}{2}\right\|^2,$$

válida para todos os vetores  $x, y, z \in V$ . \*

**Identidade de polarização:** Para todos os vetores  $u, v$  de um espaço vetorial complexo  $V$  vale

$$\langle u, v \rangle = \frac{1}{4} \sum_{n=0}^3 i^{-n} \|u + i^n v\|^2, \tag{3.32}$$

$$\langle u, v \rangle = \frac{1}{4} \sum_{n=0}^3 i^n \|u + i^{-n} v\|^2, \tag{3.33}$$

ou seja,

$$4\langle u, v \rangle = \|u + v\|^2 - \|u - v\|^2 - i\|u + iv\|^2 + i\|u - iv\|^2.$$

Prova. *Exercício*. Expanda o lado direito e verifique a igualdade.

**E. 3.24** *Exercício*. Por que essa relação é chamada “identidade de polarização”? \*

Notemos que, com a definição dada acima de norma associada a um produto escalar, a desigualdade de Cauchy-Schwarz fica

$$|\langle u, v \rangle| \leq \|u\| \|v\|.$$

• **A identidade de polarização**

A *identidade de polarização* mencionada acima é um caso especial de uma outra ligeiramente mais geral, também denominada identidade de polarização. Seja  $A$  um operador linear em um espaço vetorial  $V$  sobre os complexos e sejam  $u$  e  $v$  elementos de seu domínio. Então, vale que

$$\langle u, Av \rangle = \frac{1}{4} \sum_{n=0}^3 i^{-n} \langle (u + i^n v), A(u + i^n v) \rangle, \tag{3.34}$$

$$\langle u, Av \rangle = \frac{1}{4} \sum_{n=0}^3 i^n \langle (u + i^{-n} v), A(u + i^{-n} v) \rangle, \tag{3.35}$$

**E. 3.25** *Exercício*. Mostre isso. Sugestão: expanda o lado direito das igualdades acima e constate as igualdades. \*

Tomando-se  $A$  como o operador identidade reobtem-se as identidades (3.32)-(3.33).

A relação (3.34) mostra que se para um operador linear  $A$  conhecermos todas as quantidades  $\langle \psi, A\psi \rangle$  para todos os vetores  $\psi \in V$ , então conhecemos também todas as quantidades  $\langle u, Av \rangle$  para todos  $u, v \in V$ .

Para a física quântica a identidade de polarização (3.34) diz que se  $A$  for um observável (operador autoadjunto), então o conhecimento de todos os valores esperados de  $A$ , ou seja, das quantidades  $\langle \psi, A\psi \rangle$  com  $\|\psi\| = 1$  e dos produtos escalares  $\langle u, v \rangle$  para vetores com  $\|u\| = \|v\| = 1$ , fixa todas as probabilidades de transição  $|\langle u, Av \rangle|^2$ , pois

$$\langle u, Av \rangle = \frac{1}{4} \sum_{n=0}^3 i^{-n} \langle \psi_n, A\psi_n \rangle (2 + i^n \langle u, v \rangle + i^{-n} \langle v, u \rangle), \tag{3.36}$$

onde

$$\psi_n = \frac{1}{\|u + i^n v\|} (u + i^n v) = \frac{1}{\sqrt{2 + i^n \langle u, v \rangle + i^{-n} \langle v, u \rangle}} (u + i^n v).$$

---

<sup>11</sup>Apolônio de Perga (ci. 261 A.C. – ci. 190 A.C.).

• Uma consequência da identidade de polarização

A relação (3.34) permite-nos facilmente provar a seguinte afirmação, frequentemente empregada:

**Proposição 3.3** *Se um operador linear  $A$  agindo em um espaço vetorial complexo  $V$  satisfaz  $\langle u, Au \rangle = 0$  para todo vetor  $u \in V$ , então  $A = 0$ . □*

Para matrizes reais em espaços vetoriais reais não vale uma afirmativa tão forte. Por exemplo, se  $V = \mathbb{R}^n$  e  $A$  for uma matriz antissimétrica, ou seja  $A^T = -A$ , então vale automaticamente que  $\langle x, Ax \rangle_{\mathbb{R}} = \sum_{a,b=1}^n x_a A_{ab} x_b = 0$ , pois  $A_{ab} = -A_{ba}$  para todo  $x \in \mathbb{R}^n$ . Porém,  $A$  pode ser não-nula.

Todavia, para matrizes simétricas vale o seguinte:

**Proposição 3.4** *Seja  $M \in \text{Mat}(\mathbb{R}, n)$  uma matriz simétrica (ou seja, tal que  $M^T = M$ ) para a qual valha que  $\langle x, Mx \rangle_{\mathbb{R}} = 0$  para todo  $x \in \mathbb{R}^n$ . Então,  $M = 0$ . □*

*Prova.* Se  $M$  é uma matriz simétrica, é fácil verificar que para quaisquer vetores  $u$  e  $v \in \mathbb{R}^n$  tem-se

$$\langle u, Mv \rangle_{\mathbb{R}} = \frac{1}{4} \left[ \langle (u+v), M(u+v) \rangle_{\mathbb{R}} - \langle (u-v), M(u-v) \rangle_{\mathbb{R}} \right].$$

(Para provar isso, expanda o lado direito e use que  $\langle u, Mv \rangle_{\mathbb{R}} = \langle v, Mu \rangle_{\mathbb{R}}$ , pois  $M$  é simétrica). Logo, da hipótese sobre  $M$ , segue que  $\langle u, Mv \rangle_{\mathbb{R}} = 0$  para todos  $u$  e  $v \in \mathbb{R}^n$  e, portanto,  $M = 0$ . ■

• Obtendo produtos escalares a partir de normas

Nas últimas páginas vimos que podemos obter uma norma a partir de um produto escalar e que essa norma satisfaz a identidade do paralelogramo, expressão (3.31). Podemos nos perguntar: se uma norma for dada em um espaço vetorial complexo, seria possível obter um produto escalar a partir dessa norma?

A resposta a essa questão é fornecida por um teorema devido a Fréchet<sup>12</sup>, von Neumann<sup>13</sup> e Jordan<sup>14</sup>, teorema esse sugerido pela identidade de polarização, expressão (3.32), página 237.

**Teorema 3.3 (Teorema de Fréchet, von Neumann e Jordan)** *Seja  $V$  um espaço vetorial complexo, normado com norma  $\| \cdot \|$  e vamos supor que essa norma satisfaça a identidade do paralelogramo*

$$\|a+b\|^2 + \|a-b\|^2 = 2\|a\|^2 + 2\|b\|^2 \tag{3.37}$$

para todos  $a, b \in V$ . Defina-se, para  $u, v \in V$ ,

$$\omega(u, v) := \frac{1}{4} \sum_{n=0}^3 i^{-n} \|u + i^n v\|^2. \tag{3.38}$$

Então,  $\omega$  é um produto escalar em  $V$ .

Com essa definição, vale  $\omega(u, u) = \|u\|^2$  para todo  $u \in V$  e, portanto, a norma associada ao produto escalar  $\omega$  é a própria norma  $\| \cdot \|$ . Com isso, reconhecemos que (3.38) coincide com a identidade de polarização para o produto escalar  $\omega$ .

Concluí-se, então, que uma norma é associada a um produto escalar se e somente se satisfizer a identidade do paralelogramo. □

<sup>12</sup>Maurice René Fréchet (1878–1973).

<sup>13</sup>János von Neumann (1903–1957). Von Neumann também adotou os nomes de Johann von Neumann e John von Neumann.

<sup>14</sup>Ernst Pascual Jordan (1902–1980).

A demonstração do Teorema de Fréchet, von Neumann e Jordan encontra-se no Apêndice 3.B, página 255. Vide também [429] ou [227] para outras demonstrações essencialmente idênticas.

A demonstração do Teorema 3.3 é engenhosa e a principal dificuldade consiste em demonstrar que (3.38) é uma forma sesquilinear, um fato um tanto surpreendente se observarmos que o lado direito de (3.38) contém uma soma de normas, que não são sequer funções lineares, satisfazendo apenas  $\|\alpha u\| = |\alpha| \|u\|$  e  $\|\alpha u + \beta v\| \leq |\alpha| \|u\| + |\beta| \|v\|$  para todos  $\alpha, \beta \in \mathbb{C}$  e todos  $u, v \in V$ .

Mencionemos, por fim, que nem toda norma satisfaz a identidade do paralelogramo e, portanto, nem toda norma é associada a um produto escalar e, assim, nem sempre é possível definir um produto escalar a partir de uma norma. Os Exercícios E. 3.26 e E. 3.27, servem como exemplo de tais situações.

**E. 3.26 Exercício.** Seja o espaço vetorial  $V = C([0, 1], \mathbb{C})$  das funções contínuas do intervalo  $[0, 1]$  assumindo valores complexos e seja a norma  $\|f\|_\infty = \sup_{x \in [0, 1]} |f(x)|$ . Mostre que a identidade do paralelogramo não é satisfeita para as funções  $f(x) = x$  e  $g(x) = 1$ ,  $x \in [0, 1]$ , que são elementos de  $V$ . \*

**E. 3.27 Exercício.** Seja o espaço vetorial  $V = \mathbb{C}^n$ , com  $n \geq 2$ . Para  $a = (a_1, \dots, a_n) \in \mathbb{C}^n$  a expressão  $\|a\|_p := [|a_1|^p + \dots + |a_n|^p]^{1/p}$ , define uma norma em  $V = \mathbb{C}^n$ , caso  $p \geq 1$ . Mostre que essa norma viola a identidade do paralelogramo para todo  $p \neq 2$ . Para tal considere os vetores  $u = (1, 0, 0, \dots, 0)$  e  $v = (0, 1, 0, \dots, 0)$ . A norma  $\|\cdot\|_p$  será discutida com mais detalhe no Capítulo 24, página 1245. \*

• **Generalizando a identidade do paralelogramo**

O Exercício E. 3.27 ensina-nos que a identidade do paralelogramo não é válida para a norma  $\|\cdot\|_p$ , caso  $p \geq 1$  mas  $p \neq 2$ . Em tais casos, porém, há desigualdades que em muito se assemelham à identidade do paralelogramo.

Sabemos do Corolário 5.6, página 307, que para  $z, w \in \mathbb{C}$ , valem as desigualdades  $|z + w|^p + |z - w|^p \leq 2(|z|^p + |w|^p)$  para  $0 < p < 2$  e  $|z + w|^p + |z - w|^p \leq 2^{p-1}(|z|^p + |w|^p)$ , para  $p \geq 2$ . É imediato por essas desigualdades que para todos  $u, v \in \mathbb{C}^n$  valem

$$\|u + v\|_p^p + \|u - v\|_p^p \leq 2(\|u\|_p^p + \|v\|_p^p), \quad 1 \leq p < 2, \tag{3.39}$$

e

$$\|u + v\|_p^p + \|u - v\|_p^p \leq 2^{p-1}(\|u\|_p^p + \|v\|_p^p), \quad p \geq 2. \tag{3.40}$$

Note-se que no caso  $p = 2$  (e somente nesse caso), (3.40) não é apenas uma desigualdade, mas sim uma igualdade, a identidade do paralelogramo.

**E. 3.28 Exercício.** Mostre isso! \*

As desigualdades (3.39) e (3.40) substituem em certos casos a identidade do paralelogramo. Veremos isso quando discutirmos a propriedade de convexidade uniforme na Seção 24.6, página 1287.

• **Normas assimétricas**

Seja  $V$  um espaço vetorial real. Uma função  $p : V \rightarrow [0, \infty)$  é dita ser uma *norma assimétrica* (também denominada *norma de Finsler*<sup>15</sup>, ou *métrica de Finsler*) se satisfizer:

1. Para todo  $v \in V$  tem-se  $p(v) \geq 0$ , sendo que  $p(v) = 0$  se e somente se  $v = 0$ .
2. Para qualquer  $\alpha \geq 0$  e qualquer  $v \in V$  tem-se  $p(\alpha v) = \alpha p(v)$ .
3. Para quaisquer vetores  $u$  e  $v \in V$  tem-se  $p(u + v) \leq p(u) + p(v)$  (propriedade triangular).

Como se percebe, uma norma assimétrica não necessariamente satisfaz a condição  $p(-v) = p(v)$ , para todo  $v \in V$ , válida para normas.

A noção de seminorma assimétrica é definida analogamente, sem a exigência, porém, que  $p(v) = 0$  implique  $v = 0$ .

---

<sup>15</sup>Paul Finsler 1894–1970).

Há dois exemplos elementares de normas assimétricas. Se  $V = \mathbb{R}$ , defina-se

$$p(x) = \begin{cases} |x| & \text{se } x < 0, \\ 2|x| & \text{se } x \geq 0. \end{cases}$$

É fácil ver que essa função define uma norma assimétrica, mas não uma norma em  $\mathbb{R}$ .

Um outro exemplo, este mais importante, é o chamado *funcional de Minkowski*<sup>16</sup>, relevante na Análise Convexa e na Teoria das Distribuições (Espaços Localmente Convexos). Seja  $V$  um espaço vetorial real e  $C \subset V$  um subconjunto convexo de  $V$  tal que  $0 \in C$ . Definamos  $\lambda C := \{\lambda c, c \in C\}$  e suponhamos adicionalmente que  $C$  seja um *conjunto absorvente*, ou seja, que satisfaça  $\bigcup_{\lambda \geq 0} \lambda C = V$ .

Defina-se  $p_C : V \rightarrow [0, \infty)$  por

$$p_C(v) = \inf \{ \lambda \in [0, \infty) \mid v \in \lambda C \}.$$

A condição de  $C$  ser absorvente garante que  $p_C(v)$  seja finito para todo  $v \in V$ .

Mostremos que todo funcional de Minkowski é uma seminorma assimétrica. É claro pela definição que  $p_C(v) \geq 0$  para todo  $v \in V$ . Para a propriedade triangular, sejam  $v_1, v_2 \in V$  e sejam  $\lambda_1 > p_C(v_1)$  e  $\lambda_2 > p_C(v_2)$ . Observe-se que ambos  $\lambda_1$  e  $\lambda_2$  são positivos. Então,  $v_1 \in \lambda_1 C$  e  $v_2 \in \lambda_2 C$  e, portanto, existem  $c_1$  e  $c_2 \in C$  tais que  $v_1 = \lambda_1 c_1$  e  $v_2 = \lambda_2 c_2$ . Logo,  $\frac{\lambda_1}{\lambda_1 + \lambda_2} c_1 + \frac{\lambda_2}{\lambda_1 + \lambda_2} c_2$  é um elemento de  $C$ , por ser uma combinação linear convexa de elementos de  $C$ . Como

$$v_1 + v_2 = (\lambda_1 + \lambda_2) \left[ \frac{\lambda_1}{\lambda_1 + \lambda_2} c_1 + \frac{\lambda_2}{\lambda_1 + \lambda_2} c_2 \right]$$

temos que  $\lambda_1 + \lambda_2 \geq p_C(v_1 + v_2)$ . Concluimos disso que  $p_C(v_1 + v_2) \leq \lambda_1 + \lambda_2$  para todos  $\lambda_1 \in (p_C(v_1), \infty)$  e  $\lambda_2 \in (p_C(v_2), \infty)$ . Isso implica  $p_C(v_1 + v_2) \leq p_C(v_1) + p_C(v_2)$ , como desejávamos estabelecer.

Note que  $p_C(v)$  só coincide com  $p_C(-v)$  se  $C$  coincidir com  $-C$ . A condição que  $p_C(v) = 0$  implica  $v = 0$  é verdadeira se  $V$  tem dimensão finita, mas pode falhar em dimensão infinita.



A noção de norma assimétrica é relevante nas chamadas *variedades de Finsler*, uma variedade diferenciável onde o tensor métrico Riemanianno é substituído por uma norma assimétrica em cada ponto do espaço tangente, também denominada *métrica de Finsler*.

• **Bolas em espaços normados**

Seja  $V$  um espaço vetorial dotado de uma norma  $\| \cdot \|$ . Definimos a *bola aberta* de raio  $r > 0$  centrada em  $z \in V$ , denotada por  $\mathcal{B}_r(z)$ , por

$$\mathcal{B}_r(z) := \{ x \in V : \|x - z\| < r \}.$$

Definimos a *bola fechada* de raio  $r > 0$  centrada em  $z \in V$ , denotada por  $\overline{\mathcal{B}_r(z)}$ , por

$$\overline{\mathcal{B}_r(z)} := \{ x \in V : \|x - z\| \leq r \}.$$

O *bordo* de uma bola aberta  $\mathcal{B}_r(z)$  ou fechada  $\overline{\mathcal{B}_r(z)}$ , denotado por  $\partial \mathcal{B}_r(z)$ , é definido por  $\partial \mathcal{B}_r(z) := \overline{\mathcal{B}_r(z)} \setminus \mathcal{B}_r(z)$ , ou seja,

$$\partial \mathcal{B}_r(z) := \{ x \in V : \|x - z\| = r \}.$$

Uma bola fechada  $\overline{\mathcal{B}_r(z)}$  é sempre um subconjunto convexo de  $V$ , ou seja, para todo  $x, y \in \overline{\mathcal{B}_r(z)}$  e para todo  $\lambda \in [0, 1]$  tem-se  $\lambda x + (1 - \lambda)y \in \overline{\mathcal{B}_r(z)}$ . De fato, pelas propriedades definidoras de uma norma, temos para  $x, y \in \overline{\mathcal{B}_r(z)}$  e  $\lambda \in [0, 1]$  que  $\| \lambda x + (1 - \lambda)y - z \| = \| \lambda(x - z) + (1 - \lambda)(y - z) \| \leq \lambda \|x - z\| + (1 - \lambda) \|y - z\| \leq \lambda r + (1 - \lambda)r = r$ , estabelecendo que  $\lambda x + (1 - \lambda)y \in \overline{\mathcal{B}_r(z)}$ . De forma totalmente análoga, prova-se que uma bola aberta  $\mathcal{B}_r(z)$  é também sempre um subconjunto convexo de  $V$ .

No espaço  $\mathbb{R}^2$  podem ser definidas várias normas. Vejamos alguns exemplos. Para  $x \equiv (x_1, x_2) \in \mathbb{R}^2$ , temos as normas  $\|x\|_p := \left( |x_1|^p + |x_2|^p \right)^{1/p}$ , com  $p \geq 1$ . Temos também a norma  $\|x\|_\infty := \max \{ |x_1|, |x_2| \}$ . Na Figura 3.1, página 241, exibimos o aspecto das bolas fechadas  $\overline{\mathcal{B}_1(0)}$ , de raio 1 centradas em 0, relacionadas às normas  $\| \cdot \|_p$  e  $\| \cdot \|_\infty$  em  $\mathbb{R}^2$ . O estudante deve perceber que, de fato, todas essas bolas são conjuntos convexos.

<sup>16</sup>Hermann Minkowski (1864–1909).

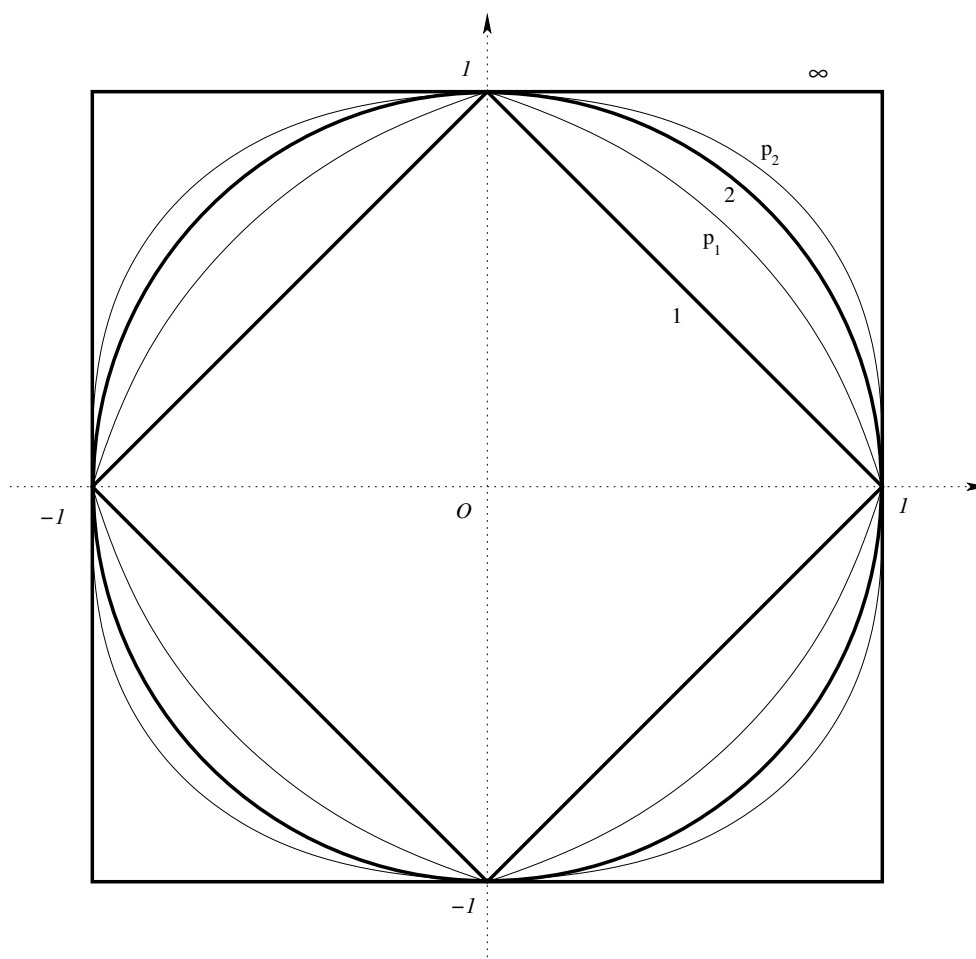


Figura 3.1: As bolas  $\overline{\mathcal{B}_1(0)}$ , de raio 1 centradas em  $0$ , relacionadas às normas  $\|\cdot\|_p$  e  $\|\cdot\|_\infty$  em  $\mathbb{R}^2$ . As linhas sólidas indicam os bordos  $\partial\mathcal{B}_1(0)$ . O índice 1 indica a bola  $\overline{\mathcal{B}_1(0)}$  para a norma  $\|\cdot\|_1$ . Trata-se de um quadrado obliquo com arestas de comprimento  $\sqrt{2}$  centrado na origem  $O$ . O índice  $p_1$  indica as bolas  $\overline{\mathcal{B}_1(0)}$  para as normas  $\|\cdot\|_{p_1}$  quando  $1 < p_1 < 2$ . O índice 2 indica a bola  $\overline{\mathcal{B}_1(0)}$  para a norma  $\|\cdot\|_2$ . Essa é a única bola que coincide com o disco de raio 1 centrado na origem  $O$ . O índice  $p_2$  indica as bolas  $\overline{\mathcal{B}_1(0)}$  para as normas  $\|\cdot\|_{p_2}$  quando  $p_2 > 2$ . O índice  $\infty$  indica a bola  $\overline{\mathcal{B}_1(0)}$  para a norma  $\|\cdot\|_\infty$ . Trata-se de um quadrado com arestas de comprimento 2 centrado na origem  $O$ .

• **Uma desigualdade útil**

Seja  $V$  um espaço vetorial dotado de uma norma  $\|\cdot\|$ . Então, vale a desigualdade

$$1 + \|x\| \leq (1 + \|y\|)(1 + \|x - y\|) \tag{3.41}$$

para todos  $x, y \in V$ . A prova é muito simples:

$$1 + \|x\| = 1 + \|x - y + y\| \leq 1 + \|x - y\| + \|y\| \leq 1 + \|x - y\| + \|y\| + \|x - y\| \|y\| = (1 + \|y\|)(1 + \|x - y\|).$$

A desigualdade (3.2) é usada em estimativas de produtos de convolução.

**3.2.0.1 O Lema da Simetria**

O elegante lema que segue encontra aplicações na Eletrostática, na Análise Harmônica, no estudo da chamada Transformação de Kelvin<sup>17</sup> e no estudo de aplicações conformes.

<sup>17</sup>William Thomson, 1st Baron Kelvin, (1824–1907).

**Lema 3.1 (Lema da Simetria)** *Seja  $V$  um espaço vetorial (real ou complexo) e  $\omega$  uma forma bilinear ou sesquilinear<sup>18</sup>. Suponhamos também que  $\omega$  seja positiva em  $V$ . Para  $u, v \in V$  tais que  $\omega(u, u)$  e  $\omega(v, v)$  sejam não nulas, definamos a expressão*

$$\eta(u, v) := \frac{v}{\sqrt{\omega(v, v)}} - \sqrt{\omega(v, v)}u \in V.$$

Então, se  $x, y \in V$  são tais que  $\omega(x, x)$  e  $\omega(y, y)$  são não nulas, vale

$$\omega(\eta(x, y), \eta(x, y)) = \omega(\eta(y, x), \eta(y, x)), \tag{3.42}$$

ou seja, de forma explícita,

$$\omega\left(\frac{y}{\sqrt{\omega(y, y)}} - \sqrt{\omega(y, y)}x, \frac{y}{\sqrt{\omega(y, y)}} - \sqrt{\omega(y, y)}x\right) = \omega\left(\frac{x}{\sqrt{\omega(x, x)}} - \sqrt{\omega(x, x)}y, \frac{x}{\sqrt{\omega(x, x)}} - \sqrt{\omega(x, x)}y\right). \tag{3.43}$$

Em particular, se  $V = \mathbb{R}^n$  ou  $\mathbb{C}^n$  e  $\omega$  é o produto escalar usual nesses espaços, então vale

$$\left\| \frac{y}{\|y\|} - \|y\|x \right\| = \left\| \frac{x}{\|x\|} - \|x\|y \right\| \tag{3.44}$$

para todos os vetores  $x, y$  não nulos. □

*Prova.* Para demonstrar (3.42), basta expandir o lado esquerdo de (3.43):

$$\begin{aligned} \omega(\eta(x, y), \eta(x, y)) &= \omega\left(\frac{y}{\sqrt{\omega(y, y)}} - \sqrt{\omega(y, y)}x, \frac{y}{\sqrt{\omega(y, y)}} - \sqrt{\omega(y, y)}x\right) \\ &= 1 - \omega(y, x) - \omega(x, y) + \omega(y, y)\omega(x, x). \end{aligned}$$

A última expressão é manifestamente invariante pela troca  $x \leftrightarrow y$ . Logo,  $\omega(\eta(x, y), \eta(x, y)) = \omega(\eta(y, x), \eta(y, x))$ , como desejávamos. A relação (3.44) é uma mera tradução disso para o produto escalar usual de  $\mathbb{R}^n$  ou  $\mathbb{C}^n$ . ■

### 3.3 Ortogonalidade, Conjuntos Ortonormais e o Procedimento de Gram-Schmidt

Seja  $V$  é um espaço vetorial dotado de um produto escalar  $\langle \cdot, \cdot \rangle$ , e seja  $\| \cdot \|$  a norma associada a esse produto escalar, ou seja, para  $v \in V$  tem-se  $\|v\| := \sqrt{\langle v, v \rangle}$ , tal como definido acima.

• **Normalização. Raio associado a um vetor**

Um vetor  $e \in V$  é dito ser um *vetor unitário*, ou um *vetor normalizado*, em relação ao produto escalar em questão (e à norma a este associada) se  $\|e\| = 1$ .

Se  $u \in V$  é um vetor não-nulo, podemos transformá-lo em um vetor unitário se o multiplicarmos por  $1/\|u\|$ . Esse procedimento é por vezes denominado *normalização* do vetor  $u$ . Se  $u$  é um vetor não-nulo, o vetor  $\frac{1}{\|u\|}u$  é normalizado, assim como todos os vetores da forma  $\frac{\lambda}{\|\lambda u\|}u$  onde  $\lambda$  é um número complexo de módulo um, i.e.,  $|\lambda| = 1$  (aqui estamos supondo que  $V$  seja um espaço vetorial sobre o corpo dos complexos). Para um vetor não-nulo  $u$ , fixo, o conjunto de vetores normalizados  $\mathfrak{R}(u) := \left\{ \frac{\lambda}{\|\lambda u\|}u, \lambda \in \mathbb{C} \text{ com } |\lambda| = 1 \right\}$  é dito ser o *raio associado ao vetor  $u$* .

• **Projeção de um vetor na direção de outro vetor. Ortogonalidade**

<sup>18</sup>A demonstração mostra que basta que  $\omega$  seja bilinear real.

Se  $u$  e  $v$  são dois vetores de  $V$  com  $v \neq 0$ , o vetor  $\frac{\langle v, u \rangle}{\|v\|^2}v$  é dito ser a *componente de  $u$  na direção de  $v$* , ou a *projeção de  $u$  na direção de  $v$* . Essa nomenclatura tem origem na bem-conhecida e familiar interpretação geométrica do produto escalar usual em  $\mathbb{R}^2$  ou em  $\mathbb{R}^3$ , mas a usamos mesmo no caso de  $V$  ser um espaço vetorial complexo ou ter dimensão infinita.

Dados dois vetores  $u, v \in V$ , dizemos que  $u$  é *ortogonal* a  $v$  em relação ao produto escalar em questão se  $\langle u, v \rangle = 0$ . Evidentemente,  $u$  é ortogonal a  $v$  se e somente se  $v$  for ortogonal a  $u$ . Caso  $v$  seja não-nulo, dizer que  $u$  é ortogonal a  $v$  significa dizer que  $u$  tem uma componente nula na direção  $v$  (e vice-versa). O vetor nulo é o único vetor de  $V$  que é ortogonal a todos os vetores de  $V$ .

Sejam  $u$  e  $v$  dois vetores linearmente independentes. Se subtrairmos de  $u$  sua componente na direção de  $v$  obtemos o vetor  $w = u - \frac{\langle v, u \rangle}{\|v\|^2}v$ . Esse vetor é não-nulo (pois  $u$  e  $v$  são linearmente independentes) e é ortogonal a  $v$  (pois  $\langle v, w \rangle = 0$ , como facilmente se constata). Com isso, obtivemos dois vetores ortogonais,  $w$  e  $v$ , a partir de dois vetores linearmente independentes,  $u$  e  $v$ . Essa ideia será generalizada logo adiante quando falarmos do procedimento de ortogonalização de Gram-Schmidt.

Comentamos, finalmente, que a noção de ortogonalidade é uma *relação de compatibilidade* na coleção dos vetores unitários de  $V$ . Vide definição na Seção 1.1.1.2, página 47.

• **Conjunto ortonormal de vetores**

Seja  $E = \{e_\alpha, \alpha \in \Lambda\}$  um conjunto não-vazio de vetores distintos de  $V$ ,  $\Lambda$  sendo um conjunto arbitrário não-vazio de índices (podendo ser finito, enumerável ou não). O conjunto  $E$  é dito ser um *conjunto ortonormal de vetores* em relação ao produto escalar em questão se para todo  $\alpha \in \Lambda$  tivermos  $\|e_\alpha\| = 1$  e se para todos  $\alpha, \beta \in \Lambda$  com  $\alpha \neq \beta$  valer  $\langle e_\alpha, e_\beta \rangle = 0$ . Assim,  $E$  é um conjunto ortonormal se todos os seus elementos forem vetores unitários e se quaisquer dois vetores distintos de  $E$  forem ortogonais entre si em relação ao produto escalar em questão.

**E. 3.29 Exercício.** Se  $E = \{e_\alpha, \alpha \in \Lambda\}$  é um conjunto ortonormal de vetores, mostre que  $\|e_\alpha - e_\beta\| = \sqrt{2}$  sempre que  $\alpha \neq \beta$ . ✱

• **Procedimento de ortogonalização de Gram-Schmidt**

Dado um conjunto finito  $B = \{b_1, \dots, b_n\}$  composto por  $n$  vetores não-nulos e linearmente independentes de  $V$ , podemos construir um conjunto ortonormal  $E = \{e_1, \dots, e_n\}$  no subespaço  $n$ -dimensional gerado pelos vetores de  $B$  por um procedimento conhecido como *procedimento de ortogonalização de Gram-Schmidt*<sup>1920</sup>, que passaremos a descrever<sup>21</sup>.

O procedimento começa escolhendo-se um vetor de  $B$  e normalizando-se esse vetor, definindo assim o primeiro vetor  $e_1$  de  $E$ . Para simplificar a notação, escolhemos começar com o vetor  $b_1$  e definimos, assim,  $e_1 := \frac{1}{\|b_1\|}b_1$ . No segundo passo, tomamos o vetor  $b_2$ , subtraímos do mesmo sua componente na direção  $e_1$  e em seguida normalizamos o vetor disso resultante, definindo assim o vetor  $e_2$ :

$$e_2 := \frac{1}{\|b_2 - \langle e_1, b_2 \rangle e_1\|} \left( b_2 - \langle e_1, b_2 \rangle e_1 \right).$$

Pela construção, é evidente que  $\|e_1\| = \|e_2\| = 1$  e é fácil verificar que  $\langle e_1, e_2 \rangle = 0$ .

A ideia do procedimento é prosseguir com os demais vetores de forma análoga, tomando na  $k$ -ésima etapa o vetor  $b_k$ , subtraindo do mesmo suas componentes nas direções  $e_1, \dots, e_{k-1}$  e normalizando o vetor assim resultante. Obtemos

$$e_1 := \frac{b_1}{\|b_1\|}, \quad e_k := \frac{1}{\left\| b_k - \sum_{l=1}^{k-1} \langle e_l, b_k \rangle e_l \right\|} \left( b_k - \sum_{l=1}^{k-1} \langle e_l, b_k \rangle e_l \right), \quad k = 2, \dots, n.$$

Observe-se que cada  $e_k$  depende apenas dos vetores  $e_1, \dots, e_{k-1}$ , definidos nas  $k - 1$  etapas anteriores. Como é fácil verificar, valem as relações  $\langle e_i, e_j \rangle = \delta_{ij}$  para todos  $i, j = 1, \dots, n$ , atestando que  $E = \{e_1, \dots, e_n\}$  é um conjunto ortonormal de vetores.

<sup>19</sup>Jørgen Pedersen Gram (1850–1916).

<sup>20</sup>Erhard Schmidt (1876–1959).

<sup>21</sup>Seria mais adequado chamar o procedimento de *procedimento de ortogonalização de Gram-Schmidt*, pois o conjunto de vetores resultante é ortonormal, mas aquela nomenclatura é adotada amplamente.



Note que a construção acima descrita não é única, pois podemos reordenar os elementos de  $B$ , obtendo assim uma nova sequência de vetores  $e_j, j = 1, \dots, n$ .

Comentamos, por fim, que o procedimento de Gram-Schmidt, descrito acima, aplica-se sem qualquer modificação ao caso de  $B$  ser um conjunto contável (não necessariamente finito) de vetores não-nulos com a propriedade que qualquer subconjunto finito de  $B$  seja composto por vetores linearmente independentes (a existência de uma tal  $B$  requer, naturalmente, que  $V$  seja um espaço vetorial de dimensão infinita). O conjunto  $E$  assim produzido será igualmente um conjunto ortonormal contável. Nesse contexto, o procedimento de Gram-Schmidt tem aplicações no estudo e na construção de famílias de polinômios ortogonais, como os de Legende, os de Tchebychev etc. Vide Capítulo 16, página 810, e referências lá citadas.

### 3.4 Formas Bilineares e Sesquilineares e Produtos Escalares em Espaços de Dimensão Finita

É possível estabelecer a forma geral de uma forma bilinear ou sesquilinear em certos espaços vetoriais, como os espaços de dimensão finita  $\mathbb{R}^n$  ou  $\mathbb{C}^n$ . É o que discutiremos nesta seção.

Faremos uso do chamado Teorema da Representação de Riesz, que afirma o seguinte.

**Teorema 3.4 (Teorema da Representação de Riesz)** *Seja  $l$  um funcional linear contínuo em um espaço de Hilbert  $\mathcal{H}$  (com um produto escalar  $\langle \cdot, \cdot \rangle_{\mathcal{H}}$ ). Então, existe  $\phi \in \mathcal{H}$ , único, tal que*

$$l(x) = \langle \phi, x \rangle_{\mathcal{H}}, \quad \forall x \in \mathcal{H}.$$

□

A demonstração desse importante teorema pode ser encontrada na Seção 38.2.1.1, página 1984. Notemos que esse teorema se aplica aos espaços vetoriais  $\mathbb{R}^n$  ou  $\mathbb{C}^n$ , pois os mesmos são espaços de Hilbert em relação aos produtos escalares  $\langle \cdot, \cdot \rangle_{\mathbb{R}}$  e  $\langle \cdot, \cdot \rangle_{\mathbb{C}}$ , respectivamente, definidos em (3.2) e (3.16) (páginas 221 e 231).

• **Continuidade**

Vamos provar a seguinte afirmação: toda forma bilinear em  $\mathbb{R}^n$  é contínua (em ambas as variáveis), o mesmo valendo para formas bilineares ou sesquilineares em  $\mathbb{C}^n$ .

Vamos provar a afirmação para as formas sesquilineares em  $\mathbb{C}^n$ . Os outros casos são idênticos. Seja  $\omega$  uma forma sesquilinear em  $\mathbb{C}^n$ . Para vetores  $x, y \in \mathbb{C}^n, y \neq 0$ , escrevemos

$$\omega(x, y) = \|y\| \omega(x, y/\|y\|), \tag{3.45}$$

onde  $\|y\| = \sqrt{\langle y, y \rangle_{\mathbb{C}}}$ . Notemos, então, que se  $v$  é um vetor de norma igual a 1 e  $\{b_1, \dots, b_n\}$  é uma base ortonormal em  $\mathbb{C}^n$ , então  $v = v_1 b_1 + \dots + v_n b_n$  com  $|v_j| \leq 1$  para todo  $j$ . Assim,

$$\omega(x, v) = v_1 \omega(x, b_1) + \dots + v_n \omega(x, b_n)$$

e, portanto,

$$|\omega(x, v)| \leq |\omega(x, b_1)| + \dots + |\omega(x, b_n)|.$$

Para cada  $x$  fixo o lado direito é uma constante  $K_x$  e não depende de  $v$ . Aplicando isso a (3.45), teremos

$$|\omega(x, y)| \leq \|y\| K_x.$$

Isso mostra que

$$\lim_{y \rightarrow 0} |\omega(x, y)| = 0$$

para todo  $x$  fixo. Como  $\omega(x, y)$  é linear na segunda variável, segue que

$$\lim_{y \rightarrow y_0} \omega(x, y) = \omega(x, y_0),$$

para todo  $y_0 \in \mathbb{C}^n$ , provando a continuidade de  $\omega$  na segunda variável. A prova para a primeira variável é idêntica. Os casos em que  $\omega$  é bilinear em  $\mathbb{R}^n$ , ou em  $\mathbb{C}^n$ , é análogo.

• **Formas sesquilineares em  $\mathbb{C}^n$**

Seja  $\omega$  uma forma sesquilinear em  $\mathbb{C}^n$ . Então, pelo que acabamos de ver, para cada  $x \in \mathbb{C}^n$

$$l_x : \mathbb{C}^n \rightarrow \mathbb{C}, \quad l_x(y) = \omega(x, y)$$

é um funcional linear e contínuo. Pelo Teorema da Representação de Riesz existe um único vetor  $\eta_x \in \mathbb{C}^n$  tal que  $l_x(y) = \langle \eta_x, y \rangle_{\mathbb{C}}$  para todo  $y \in \mathbb{C}^n$ , ou seja,

$$\omega(x, y) = \langle \eta_x, y \rangle_{\mathbb{C}}.$$

Seja  $A$  a função que a cada  $x \in \mathbb{C}^n$  associa o (único!) vetor  $\eta_x$  com a propriedade acima:  $A(x) = \eta_x$ . Tem-se,

$$\omega(x, y) = \langle A(x), y \rangle_{\mathbb{C}}. \tag{3.46}$$

Afirmamos que  $A$  é um operador linear, ou seja,  $A(\alpha_1 x_1 + \alpha_2 x_2) = \alpha_1 A(x_1) + \alpha_2 A(x_2)$  para todos os números complexos  $\alpha_1$  e  $\alpha_2$  e todos os vetores  $x_1$  e  $x_2$ . De fato, por (3.46),

$$\begin{aligned} \langle A(\alpha_1 x_1 + \alpha_2 x_2), y \rangle_{\mathbb{C}} &= \omega(\alpha_1 x_1 + \alpha_2 x_2, y) \\ &= \overline{\alpha_1} \omega(x_1, y) + \overline{\alpha_2} \omega(x_2, y) \\ &= \overline{\alpha_1} \langle A(x_1), y \rangle_{\mathbb{C}} + \overline{\alpha_2} \langle A(x_2), y \rangle_{\mathbb{C}} \\ &= \langle \alpha_1 A(x_1) + \alpha_2 A(x_2), y \rangle_{\mathbb{C}}. \end{aligned}$$

Assim, para todo  $y \in \mathbb{C}^n$  tem-se

$$\left\langle \left[ A(\alpha_1 x_1 + \alpha_2 x_2) - \alpha_1 A(x_1) - \alpha_2 A(x_2) \right], y \right\rangle_{\mathbb{C}} = 0,$$

o que implica

$$A(\alpha_1 x_1 + \alpha_2 x_2) = \alpha_1 A(x_1) + \alpha_2 A(x_2),$$

que é o que queríamos provar. Assim,  $A$  é em verdade um operador linear. Resumimos esses fatos no seguinte teorema:

**Teorema 3.5** *Para toda forma sesquilinear  $\omega$  em  $\mathbb{C}^n$  existe uma matriz  $n \times n$  complexa  $A_\omega$  tal que*

$$\omega(x, y) = \langle A_\omega x, y \rangle_{\mathbb{C}},$$

para todos  $x, y \in \mathbb{C}^n$ . □

Esse teorema estabelece assim a forma geral das formas sesquilineares em  $\mathbb{C}^n$ .

• **Formas bilineares em  $\mathbb{R}^n$**

Seja  $\omega$  uma forma bilinear em  $\mathbb{R}^n$ . Então, para cada  $x \in \mathbb{R}^n$

$$l_x : \mathbb{R}^n \rightarrow \mathbb{R} : \quad l_x(y) = \omega(x, y)$$

é um funcional linear e contínuo. Pelo Teorema da Representação de Riesz existe um único vetor  $\eta_x \in \mathbb{R}^n$  tal que  $l_x(y) = \langle \eta_x, y \rangle_{\mathbb{R}}$ , ou seja,

$$\omega(x, y) = \langle \eta_x, y \rangle_{\mathbb{R}}.$$

Seja  $A$  a função que a cada  $x \in \mathbb{R}^n$  associa o (único!) vetor  $\eta_x$  com a propriedade acima:  $A(x) = \eta_x$ . De maneira análoga ao que fizemos acima podemos provar que  $A$  é um operador linear, ou seja, uma matriz  $n \times n$  real e  $\omega(x, y) = \langle Ax, y \rangle_{\mathbb{R}}$ .

Resumimos esses fatos no seguinte teorema:

**Teorema 3.6** Para toda forma bilinear  $\omega$  em  $\mathbb{R}^n$  existe uma matriz  $n \times n$  real  $A_\omega$  tal que

$$\omega(x, y) = \langle A_\omega x, y \rangle_{\mathbb{R}},$$

para todos  $x, y \in \mathbb{R}^n$ . □

Esse teorema estabelece assim a forma geral das formas bilineares em  $\mathbb{R}^n$ .

• **Formas bilineares em  $\mathbb{C}^n$**

Seja  $\omega$  uma forma bilinear em  $\mathbb{C}^n$ . Então,

$$\omega_s(x, y) = \omega(\bar{x}, y)$$

define uma forma sesquilinear em  $\mathbb{C}^n$ , onde  $\bar{x} = (\bar{x}_1, \dots, \bar{x}_n)$  para  $x = (x_1, \dots, x_n) \in \mathbb{C}^n$ . Pelo que provamos acima, portanto, existe uma matriz complexa  $A_\omega$  tal que

$$\omega_s(x, y) = \langle A_\omega x, y \rangle_{\mathbb{C}},$$

para todos  $x, y \in \mathbb{C}^n$ , ou seja,

$$\omega(x, y) = \langle A_\omega \bar{x}, y \rangle_{\mathbb{C}},$$

para todos  $x, y \in \mathbb{C}^n$ .

Note que isso também diz que

$$\omega(x, y) = \langle \overline{A_\omega} x, y \rangle_{\mathbb{R}},$$

onde  $\overline{A_\omega}$  é o complexo conjugado da matriz  $A_\omega$ .

Resumimos esses fatos no seguinte teorema:

**Teorema 3.7** Para toda forma bilinear  $\omega$  em  $\mathbb{C}^n$  existe uma matriz  $n \times n$  complexa  $A_\omega$  tal que

$$\omega(x, y) = \langle A_\omega x, y \rangle_{\mathbb{R}}$$

para todos  $x, y \in \mathbb{C}^n$ . □

Esse teorema estabelece assim a forma geral das formas bilineares em  $\mathbb{C}^n$ .

• **Formas simpléticas**

Se  $\omega$  é uma forma bilinear alternante em  $\mathbb{R}^n$  ou  $\mathbb{C}^n$ , ou seja,  $\omega$  é bilinear e  $\omega(x, y) = -\omega(y, x)$ , então  $\omega$  é da forma  $\omega(x, y) = \langle Ax, y \rangle_{\mathbb{R}}$  onde  $A$  é uma matriz antissimétrica, ou seja,  $A^T = -A$ . De fato, como  $\langle x, y \rangle_{\mathbb{R}} = \langle y, x \rangle_{\mathbb{R}}$  e como  $\omega(x, y) = -\omega(y, x)$ , segue que

$$\langle Ax, y \rangle_{\mathbb{R}} = -\langle Ay, x \rangle_{\mathbb{R}} = -\langle y, A^T x \rangle_{\mathbb{R}} = -\langle A^T x, y \rangle_{\mathbb{R}}.$$

Como isso vale para todo  $x, y \in \mathbb{R}^n$  (ou  $\mathbb{C}^n$ ), tem-se  $A^T = -A$ .

Isso determina a forma geral de uma forma bilinear alternante em  $\mathbb{R}^n$  ou  $\mathbb{C}^n$ .

Se  $\omega$  é uma forma simplética, ou seja,  $\omega$  é uma forma bilinear alternante não-degenerada, então  $A$  tem de ser também inversível. De fato, se  $\langle Ax, y \rangle_{\mathbb{R}} = 0$  para todo  $y$ , então  $Ax = 0$ . Se  $A$  é inversível isso só é possível se  $x = 0$ .

Uma consequência do fato de  $A$  ter de ser inversível é que  $n$  tem que ser par. De fato, a condição  $A^T = -A$  diz que  $\det(A) = \det(-A^T) = (-1)^n \det(A^T) = (-1)^n \det(A)$ . Portanto, se  $n$  é ímpar teríamos  $\det(A) = 0$ .

A conclusão é que formas simpléticas só ocorrem nos espaços de dimensão finita  $\mathbb{R}^n$  ou  $\mathbb{C}^n$  se a dimensão  $n$  for par, e nesse caso, têm a forma  $\omega(x, y) = \langle Ax, y \rangle_{\mathbb{R}}$ , onde  $A$  é inversível e satisfaz  $A^T = -A$ .

• **Formas sesquilineares Hermitianas em  $\mathbb{C}^n$**

Se  $\omega$  é uma forma sesquilinear Hermitiana em  $\mathbb{C}^n$ , tem-se  $\omega(x, y) = \overline{\omega(y, x)}$ . Se  $A$  é a matriz tal que  $\langle Ax, y \rangle_{\mathbb{C}} = \omega(x, y)$ , então

$$\langle Ax, y \rangle_{\mathbb{C}} = \overline{\langle Ay, x \rangle_{\mathbb{C}}} = \langle x, Ay \rangle_{\mathbb{C}} = \langle A^* x, y \rangle_{\mathbb{C}},$$

onde  $A^* := \overline{A^T}$  é a adjunta de  $A$ . Como a última relação vale para todo  $x, y \in \mathbb{C}^n$ , tem-se  $A = A^*$ , ou seja,  $A$  é uma matriz autoadjunta.

Portanto, a forma geral de uma forma sesquilinear Hermitiana em  $\mathbb{C}^n$  é  $\langle Ax, y \rangle_{\mathbb{C}}$ , onde  $A$  é uma matriz autoadjunta.

• **Produtos escalares em  $\mathbb{C}^n$**

Se  $\omega$  é um produto escalar em  $\mathbb{C}^n$ ,  $\omega$  é sesquilinear Hermitiana e  $\omega(x, x) > 0$  se  $x \neq 0$ . Se  $A$  é a matriz tal que  $\langle Ax, y \rangle_{\mathbb{C}} = \omega(x, y)$ , então

$$\langle Ax, x \rangle_{\mathbb{C}} > 0 \tag{3.47}$$

se  $x \neq 0$ . Uma consequência disso é o seguinte: se  $v_i$  é um dos autovetores de  $A$  com autovalor  $\lambda_i$ , então  $\lambda_i > 0$ . De fato, tomando  $x = v_i$  em (3.47), teremos<sup>22</sup>  $0 < \langle Av_i, v_i \rangle_{\mathbb{C}} = \lambda_i \langle v_i, v_i \rangle_{\mathbb{C}}$ , o que implica  $\lambda_i > 0$ . Esse fato, em particular, nos diz que  $A$  é inversível (pois o determinante de  $A$  é o produto de seus autovalores).

Outra consequência dessas observações é a seguinte. É bem sabido que os autovetores  $v_i$  de uma matriz autoadjunta  $A$  podem ser escolhidos de modo a formar uma base ortonormal (vide Teorema 10.15, página 510). Vamos definir uma matriz  $B$  de modo que  $Bv_i = \sqrt{\lambda_i}v_i$  para todos os autovetores  $v_i$  de  $A$ . Isso define a ação de  $B$  nos vetores de uma base e, portanto,  $B$  fica definida em toda parte<sup>23</sup>.

É fácil provar que  $B$  assim definida é também autoadjunta,  $B^* = B$ , e que  $B^2 = A$ . Claramente  $B$  é também inversível e tem autovalores  $> 0$ .

**E. 3.30** *Exercício.* Mostre esses fatos. ✦

Disso concluímos que

$$\omega(x, y) = \langle Ax, y \rangle_{\mathbb{C}} = \langle Bx, By \rangle_{\mathbb{C}}.$$

Para futura referência reunimos nossas conclusões sobre produtos escalares em espaços  $\mathbb{C}^m$  na seguinte proposição:

**Proposição 3.5** *Se  $\omega$  é um produto escalar em  $\mathbb{C}^n$ , então existe uma única matriz  $M_\omega \in \text{Mat}(\mathbb{C}, n)$  autoadjunta e de autovalores positivos (e, portanto, inversível) tal que  $\omega(x, y) = \langle x, M_\omega y \rangle_{\mathbb{C}}$  para todos  $x, y \in \mathbb{C}^n$ .*

*Igualmente, se  $\omega$  é um produto escalar em  $\mathbb{C}^n$ , então existe uma (única) matriz autoadjunta  $B_\omega$ , inversível e com autovalores  $> 0$  tal que  $\omega(x, y) = \langle B_\omega x, B_\omega y \rangle_{\mathbb{C}}$  para todo  $x, y \in \mathbb{C}^n$ .* □

## 3.5 Estruturas Complexas sobre Espaços Vetoriais Reais

Seja  $V$  um espaço vetorial real. Em  $V$  está, portanto, definido um produto por escalares reais:  $xv \in V$ , onde  $x \in \mathbb{R}$  e  $v \in V$ . Sob certas circunstâncias é possível transformar  $V$  em um espaço vetorial complexo definindo um produto por escalares complexos:  $z \cdot v \in V$  para  $z \in \mathbb{C}$  e  $v \in V$ . Também sob hipóteses, um produto escalar complexo pode ser definido em  $V$ .

Suponha que exista um operador linear  $J : V \rightarrow V$ , agindo em  $V$ , com a propriedade  $J^2 = -\mathbb{1}$ , onde  $\mathbb{1}$  denota o operador identidade. Se  $z \in \mathbb{C}$  é da forma  $z = x + iy$  com  $x, y \in \mathbb{R}$ , defina-se em  $V$  o produto por escalares complexos por

$$(x + iy) \cdot v := xv + yJv. \tag{3.48}$$

As seguintes propriedades poder ser facilmente verificadas como exercício:

1. O produto por escalares complexos (3.48) é associativo:

$$\alpha \cdot (\beta \cdot u) = (\alpha\beta) \cdot u,$$

para todos  $\alpha, \beta \in \mathbb{C}$  e  $u \in V$ , onde  $\alpha\beta$  é o produto de  $\alpha$  por  $\beta$  em  $\mathbb{C}$ ,

2.  $1 \cdot u = u$  para todo  $u \in V$ .

<sup>22</sup>Lembre-se que os autovalores de uma matriz autoadjunta são sempre números reais.

<sup>23</sup>Para o estudante mais avançado: aqui poderíamos usar também o teorema espectral, Teorema 10.7.

3. O produto por escalares complexos (3.48) é distributivo em relação à soma de vetores:

$$\alpha \cdot (u + v) = \alpha \cdot u + \alpha \cdot v ,$$

para todo  $\alpha \in \mathbb{C}$  e todos  $u, v \in V$ .

4. O produto por escalares complexos (3.48) é distributivo em relação à soma de escalares:

$$(\alpha + \beta) \cdot u = \alpha \cdot u + \beta \cdot u ,$$

para todos  $\alpha, \beta \in \mathbb{C}$  e todo  $u \in V$ .

Portanto, pela definição da Seção 2.1.5, página 100,  $V$  é um espaço vetorial complexo com o produto definido acima. Vamos denotar por  $V_J$  esse espaço vetorial complexo, para não confundí-lo com  $V$ , que é um espaço vetorial real. Note que os vetores de  $V$  e de  $V_J$  são os mesmos, mas  $V$  e  $V_J$  representam estruturas diferentes.  $V_J$  é dito ser uma *estrutura complexa* sobre o espaço vetorial real  $V$ .

Uma questão de grande interesse, especialmente no contexto das chamadas álgebras CAR e CCR (vide [62]) que descrevem as álgebras de comutação e anticomutação canônicas da Mecânica Quântica e das Teorias Quânticas de Campos (que descrevem modelos fermiônicos<sup>24</sup> e bosônicos<sup>25</sup>), é saber se é possível introduzir um produto escalar complexo no espaço complexo  $V_J$ . Como veremos no que segue, tal é possível se houver em  $V$  uma forma simplética real ou um produto escalar real satisfazendo certas hipóteses. Desenvolveremos primeiro as ideias gerais e apresentaremos exemplos posteriormente, à página 250.

### • Formas simpléticas reais e produtos escalares reais

Para mostrar como construir produtos escalares complexos no espaço complexo  $V_J$  precisamos do seguinte resultado preparatório, que tem interesse por si só, por estabelecer uma relação entre formas simpléticas<sup>26</sup> reais e produtos escalares reais.

**Lema 3.2** *Seja  $V$  um espaço vetorial real e suponha que exista um operador linear  $J : V \rightarrow V$  satisfazendo  $J^2 = -\mathbb{1}$ . Valem as seguintes afirmações*

**I.** *Se  $\varepsilon : V \times V \rightarrow \mathbb{R}$  é um produto escalar real em  $V$  satisfazendo*

$$\varepsilon(Ju, v) = -\varepsilon(u, Jv)$$

*para todos  $u, v \in V$ , então  $\sigma : V \times V \rightarrow \mathbb{R}$  definida para todos  $u, v \in V$  por*

$$\sigma(u, v) := \varepsilon(Ju, v) = -\varepsilon(u, Jv) \tag{3.49}$$

*é uma forma simplética real e satisfaz*

(a)  $\sigma(Ju, v) = -\sigma(u, Jv)$  para todos  $u, v \in V$ ,

(b)  $\sigma(u, Ju) \geq 0$  para todo  $u \in V$ .

**II.** *Se  $\sigma : V \times V \rightarrow \mathbb{R}$  é uma forma simplética real em  $V$  satisfazendo*

(a)  $\sigma(Ju, v) = -\sigma(u, Jv)$  para todos  $u, v \in V$ ,

(b)  $\sigma(u, Ju) \geq 0$  para todo  $u \in V$ ,

*então  $\varepsilon : V \times V \rightarrow \mathbb{R}$  definida para todos  $u, v \in V$  por*

$$\varepsilon(u, v) := \sigma(u, Jv) = -\sigma(Ju, v) \tag{3.50}$$

*é um produto escalar real e satisfaz*

(a)  $\varepsilon(Ju, v) = -\varepsilon(u, Jv)$  para todos  $u, v \in V$ .

<sup>24</sup>Enrico Fermi (1901–1954).

<sup>25</sup>Satyendra Nath Bose (1894–1974).

<sup>26</sup>Para a definição, vide página 222.

□

Prova da parte I. Pelas hipóteses,  $\varepsilon$  é um produto escalar real e, portanto, é uma forma bilinear real, positiva, simétrica e não-degenerada. Que  $\sigma$  definida em (3.49) é uma forma bilinear é evidente. Para todos  $u, v \in V$  tem-se

$$\sigma(u, v) = \varepsilon(Ju, v) = -\varepsilon(u, Jv) \stackrel{\text{simetria}}{=} -\varepsilon(Jv, u) = -\sigma(v, u),$$

provando que  $\sigma$  é uma forma alternante. Se  $\sigma(u, v) = 0$  para todo  $v \in V$ , então  $\varepsilon(Ju, v) = 0$  para todo  $v \in V$ . Mas como  $\varepsilon$  é não-degenerada, segue que  $Ju = 0$ , o que implica  $u = 0$ , pois  $J^2 = -\mathbb{1}$ . Isso provou que  $\sigma$  é não degenerada e, portanto, é uma forma simplética. Note-se agora que

$$\sigma(u, Jv) = \varepsilon(Ju, Jv) = -\varepsilon(u, J^2v) = \varepsilon(u, v) = -\sigma(Ju, v).$$

Por fim,  $\sigma(u, Ju) = \varepsilon(Ju, Ju) \geq 0$ , pois  $\varepsilon$  é um produto escalar. Pelo mesmo motivo,  $\varepsilon(Ju, Ju) = 0$  se e somente se  $Ju = 0$ . Como  $J^2 = -\mathbb{1}$ , isso implica  $u = 0$ . Isso provou as afirmações da parte I.

Prova da parte II. Pelas hipóteses,  $\sigma$  é uma forma simplética real e, portanto, é uma forma bilinear real, alternante e não-degenerada. Que  $\varepsilon$  definida em (3.50) é uma forma bilinear é evidente. Para todos  $u, v \in V$  tem-se

$$\varepsilon(u, v) = \sigma(u, Jv) = -\sigma(Ju, v) \stackrel{\text{alternância}}{=} \sigma(v, Ju) = \varepsilon(v, u),$$

provando que  $\varepsilon$  é uma forma simétrica. Se  $\varepsilon(u, v) = 0$  para todo  $v \in V$ , então  $\sigma(u, Jv) = 0$  para todo  $v \in V$ . Mas como  $\sigma$  é não-degenerada, segue que  $u = 0$ , provando que  $\varepsilon$  é uma forma não-degenerada. Para todo  $u$  tem-se também  $\varepsilon(u, u) = \sigma(u, Ju) \geq 0$ , por hipótese, provando que  $\varepsilon$  é uma forma positiva. Assim, pela Proposição 3.2, página 229,  $\varepsilon$  é um produto escalar. Note-se agora que, por definição,  $\varepsilon(u, v) = -\sigma(Ju, v)$  para todos  $u, v \in V$ . Disso segue que  $\sigma(u, v) = \varepsilon(Ju, v)$  e que

$$\varepsilon(u, Jv) = -\sigma(Ju, Jv) = \sigma(u, J^2v) = -\sigma(u, v) = -\varepsilon(Ju, v).$$

Isso provou as afirmações da parte II. ■

### • Produtos escalares complexos sobre estruturas complexas

A proposição que segue mostra como se pode construir em  $V_J$  um produto escalar complexo se for fornecida uma forma simplética real ou um produto escalar real em  $V$  satisfazendo certas hipóteses.

**Proposição 3.6** *Suponhamos que  $V$  seja um espaço vetorial real e que exista  $J : V \rightarrow V$ , um operador linear em  $V$ , satisfazendo  $J^2 = -\mathbb{1}$ . Então, valem as seguintes afirmações:*

**A.** *Se existir uma forma simplética real  $\sigma : V \times V \rightarrow \mathbb{R}$  satisfazendo*

(a)  $\sigma(Ju, v) = -\sigma(u, Jv)$  para todos  $u, v \in V$ ,

(b)  $\sigma(u, Ju) \geq 0$  para todo  $u \in V$ <sup>27</sup>,

então,  $V \times V \ni (u, v) \mapsto \langle u, v \rangle_{J, \sigma} \in \mathbb{C}$  definida por

$$\langle u, v \rangle_{J, \sigma} := \sigma(u, Jv) + i\sigma(u, v)$$

para todos  $u, v \in V$ , é um produto escalar complexo sobre a estrutura complexa  $V_J$ .

**B.** *Se existir um produto escalar real  $\varepsilon : V \times V \rightarrow \mathbb{R}$  satisfazendo*

(a)  $\varepsilon(Ju, v) = -\varepsilon(u, Jv)$  para todos  $u, v \in V$ ,

então,  $V \times V \ni (u, v) \mapsto \langle u, v \rangle_{J, \varepsilon} \in \mathbb{C}$  definida por

$$\langle u, v \rangle_{J, \varepsilon} := \varepsilon(u, v) + i\varepsilon(Ju, v)$$

para todos  $u, v \in V$ , é um produto escalar complexo sobre a estrutura complexa  $V_J$ .

---

<sup>27</sup>Em [62] essa última condição não é mencionada, mas ela é necessária.

□

Prova. Mostremos em primeiro lugar que as hipóteses das partes **A** e **B** são equivalentes. Pelo Lema 3.2, página 248, a existência de uma forma simplética real  $\sigma$  satisfazendo as hipóteses da parte **A** implica a existência de um produto escalar real  $\varepsilon$  dado por  $\varepsilon(u, v) := \sigma(u, Jv) = -\sigma(Ju, v)$  satisfazendo as hipóteses da parte **B**, sendo que, por essa definição de  $\varepsilon$ ,

$$\sigma(u, Jv) + i\sigma(u, v) = \varepsilon(u, v) + i\varepsilon(Ju, v). \quad (3.51)$$

Reciprocamente, também pelo Lema 3.2, página 248, a existência de um produto escalar real  $\varepsilon$  satisfazendo as hipóteses da parte **B** implica a existência de uma forma simplética real  $\sigma$  dada por  $\sigma(u, v) := \varepsilon(Ju, v) = -\varepsilon(u, Jv)$  satisfazendo as hipóteses da parte **A**, sendo que, por essa definição de  $\sigma$ , a igualdade (3.51) é também válida. Assim, é suficiente provarmos, digamos, a parte **A**.

Prova da parte **A**. É evidente que para quaisquer  $u, v, w \in V$  valem

$$\langle (u+v), w \rangle_{J,\sigma} = \langle u, w \rangle_{J,\sigma} + \langle v, w \rangle_{J,\sigma}, \quad \langle u, (v+w) \rangle_{J,\sigma} = \langle u, v \rangle_{J,\sigma} + \langle u, w \rangle_{J,\sigma}.$$

Além disso,

$$\langle v, u \rangle_{J,\sigma} = \sigma(v, Ju) + i\sigma(v, u) = -\sigma(Ju, v) - i\sigma(u, v) = \sigma(u, Jv) - i\sigma(u, v) = \overline{\langle u, v \rangle_{J,\sigma}}. \quad (3.52)$$

Para  $x, y \in \mathbb{R}$  tem-se também

$$\begin{aligned} \langle u, (x+iy) \cdot v \rangle_{J,\sigma} &= \langle u, xv + yJv \rangle_{J,\sigma} \\ &= \langle u, xv \rangle_{J,\sigma} + \langle u, yJv \rangle_{J,\sigma} \\ &= \sigma(u, xJv) + i\sigma(u, xv) + \sigma(u, yJ^2v) + i\sigma(u, yJv) \\ &\stackrel{J^2=-1}{=} \sigma(u, xJv) + i\sigma(u, xv) + \sigma(u, -yv) + i\sigma(u, yJv) \\ &= x(\sigma(u, Jv) + i\sigma(u, v)) + iy(\sigma(u, Jv) + i\sigma(u, v)) \\ &= (x+iy)\langle u, v \rangle_{J,\sigma}. \end{aligned}$$

Pela propriedade (3.52), isso implica também  $\langle (x+iy) \cdot u, v \rangle_{J,\sigma} = (x-iy)\langle u, v \rangle_{J,\sigma}$ , mostrando que  $\langle \cdot, \cdot \rangle_{J,\sigma}$  é uma forma sesquilinear.

Pelas hipóteses, tem-se  $\langle u, u \rangle_{J,\sigma} = \sigma(u, Ju) \geq 0$ , mostrando que  $\langle \cdot, \cdot \rangle_{J,\sigma}$  é positiva. Se  $0 = \langle u, v \rangle_{J,\sigma} = \sigma(u, Jv) + i\sigma(u, v)$  para todo  $u$ , segue que  $\sigma(u, v) = 0$  para todo  $u$ , o que implica que  $v = 0$ , pois  $\sigma$  é não-degenerada (pela nossa definição de forma simplética). Isso mostrou que  $\langle \cdot, \cdot \rangle_{J,\sigma}$  é não-degenerada. Assim,  $\langle \cdot, \cdot \rangle_{J,\sigma}$  é uma forma sesquilinear positiva e não-degenerada e pelo Teorema 3.1, página 227, segue que  $\langle u, u \rangle_{J,\sigma} = 0$  se e somente se  $u = 0$ . Isso mostrou que  $\langle \cdot, \cdot \rangle_{J,\sigma}$  é um produto escalar complexo em  $V_J$ . ■

### • Exemplos

Vamos primeiramente estudar o caso de espaços de dimensão finita. Vale a seguinte proposição:

**Proposição 3.7** *Um espaço vetorial real  $V$  de dimensão finita admite uma estrutura complexa (não necessariamente única) se e somente se tiver dimensão par.* □

Prova. Se  $J$  é um operador linear agindo no espaço vetorial real de dimensão finita  $V$ , podemos representá-lo como uma matriz. Se  $J^2 = -1$  então, tomando-se o determinante de ambos os lados, temos  $(\det(J))^2 = (-1)^n$ , onde  $n$  é a dimensão de  $V$ . Como o lado esquerdo é positivo,  $n$  tem de ser par. Reciprocamente, vamos supor que  $V$  tenha dimensão par, digamos  $2m$ . Desejamos mostrar que existe um operador linear agindo em  $V$  satisfazendo  $J^2 = -1$ . Uma possível

escolha é a seguinte. Como  $V$  tem dimensão par podemos encontrar dois subespaços  $V_1$  e  $V_2$ , ambos de dimensão  $m$ , com  $V = V_1 \oplus V_2$ . Como  $V_1$  e  $V_2$  têm a mesma dimensão, são isomorfos, e existe um operador linear  $\mathcal{A} : V_1 \rightarrow V_2$  que é bijetivo (o Exemplo 3.9, abaixo, deixará isso mais claro. Um tal operador não é necessariamente único, mas isso não representa um problema). Todo elemento  $v \in V$  pode ser escrito da forma  $v = v_1 \oplus v_2$  com  $v_1 \in V_1$  e  $v_2 \in V_2$ . Podemos definir  $Jv = J(v_1 \oplus v_2) := (-\mathcal{A}v_2) \oplus (\mathcal{A}v_1)$ . É trivial, então, verificar que  $J^2 = -\mathbb{1}$ , como desejado. ■

**Exemplo 3.9** Seja  $V$  um espaço vetorial real de dimensão  $2m$ . Em alguma base, podemos representar  $v \in V$  na forma de um vetor-coluna:

$$v = \begin{pmatrix} v_1 \\ \vdots \\ v_m \\ v_{m+1} \\ \vdots \\ v_{2m} \end{pmatrix}. \quad \text{Defina-se, então,} \quad Jv := \begin{pmatrix} -v_{m+1} \\ \vdots \\ -v_{2m} \\ v_1 \\ \vdots \\ v_m \end{pmatrix}, \quad (3.53)$$

ou seja, em forma matricial, na mesma base,

$$J = \begin{pmatrix} 0_m & -\mathbb{1}_m \\ \mathbb{1}_m & 0_m \end{pmatrix}$$

sendo  $0_m$  e  $\mathbb{1}_m$  matrizes  $m \times m$ . É elementar verificar que  $J^2 = -\mathbb{1}_{2m}$ , como desejado.

A escolha de  $J$  indicada acima dependeu de uma particular decomposição de  $V$  em dois subespaços de dimensão  $m$ . Há várias outras decomposições possíveis, que fornecem outros operadores  $J$  e, portanto, outras estruturas complexas. Permanecendo no exemplo acima, é fácil ver que, se  $x, y \in \mathbb{R}$ , então o produto por escalares complexos fica

$$(x + iy) \cdot \begin{pmatrix} v_1 \\ \vdots \\ v_m \\ v_{m+1} \\ \vdots \\ v_{2m} \end{pmatrix} := (x + yJ) \begin{pmatrix} v_1 \\ \vdots \\ v_m \\ v_{m+1} \\ \vdots \\ v_{2m} \end{pmatrix} = \begin{pmatrix} xv_1 - yv_{m+1} \\ \vdots \\ xv_m - yv_{2m} \\ xv_{m+1} + yv_1 \\ \vdots \\ xv_{2m} + yv_m \end{pmatrix}. \quad (3.54)$$

Seguindo ainda o exemplo de (3.53) e (3.54) para  $V = \mathbb{R}^{2m}$ , vamos ilustrar a Proposição 3.6 e o produto escalar complexo para  $(\mathbb{R}^{2m})_J$ . Adotemos para  $\varepsilon$  o produto escalar usual:

$$\varepsilon(u, v) := \sum_{k=1}^{2m} u_k v_k = u_1 v_1 + \cdots + u_{2m} v_{2m}.$$

Temos que

$$\varepsilon(Ju, v) = -u_{m+1} v_1 - \cdots - u_{2m} v_m + u_1 v_{m+1} + \cdots + u_m v_{2m}$$

e que

$$\varepsilon(u, Jv) = -u_1 v_{m+1} - \cdots - u_m v_{2m} + u_{m+1} v_1 + \cdots + u_{2m} v_m$$

Logo,  $\varepsilon(Ju, v) = -\varepsilon(u, Jv)$  e podemos aplicar a Proposição 3.6, obtendo em  $(\mathbb{R}^{2m})_J$  o produto escalar

$$\begin{aligned} \langle u, v \rangle_{J, \varepsilon} &= \varepsilon(u, v) + i\varepsilon(Ju, v) \\ &= \left( u_1 v_1 + \cdots + u_{2m} v_{2m} \right) + i \left( -u_{m+1} v_1 - \cdots - u_{2m} v_m + u_1 v_{m+1} + \cdots + u_m v_{2m} \right) \\ &= u_1 (v_1 + i v_{m+1}) + \cdots + u_m (v_m + i v_{2m}) + u_{m+1} (v_{m+1} - i v_1) + \cdots + u_{2m} (v_{2m} - i v_m) \\ &= \overline{(u_1 + i u_{m+1})} (v_1 + i v_{m+1}) + \cdots + \overline{(u_m + i u_{2m})} (v_m + i v_{2m}). \end{aligned}$$

**E. 3.31** *Exercício.* Verifique que  $\langle u, \lambda \cdot v \rangle_{J, \varepsilon} = \lambda \langle u, v \rangle_{J, \varepsilon}$  para todo  $\lambda \in \mathbb{C}$ .

★



Entendemos, assim, que a estrutura complexa que estudamos consiste nesse caso em identificar bijetivamente  $\mathbb{R}^{2m}$  e  $\mathbb{C}^m$  por

$$\mathbb{R}^{2m} \ni \begin{pmatrix} v_1 \\ \vdots \\ v_m \\ v_{m+1} \\ \vdots \\ v_{2m} \end{pmatrix} \longleftrightarrow \begin{pmatrix} v_1 + iv_{m+1} \\ \vdots \\ v_m + iv_{2m} \end{pmatrix} \in \mathbb{C}^m$$

e adotar em  $\mathbb{C}^m$  o produto escalar complexo  $\langle \cdot, \cdot \rangle_{\mathbb{C}}$  usual (definido à página 33). □

Vejamos como as ideias acima podem ser generalizadas e de modo a incluir espaços de dimensão infinita.

**Exemplo 3.10** Se  $V$  é um espaço vetorial real de (dimensão finita ou não) é sempre possível encontrar um operador linear  $J$  satisfazendo  $J^2 = -\mathbb{1}$  se  $V$  possuir dois subespaços  $V_1$  e  $V_2$  com  $V = V_1 \oplus V_2$  e tais que existe  $\mathcal{A} : V_1 \rightarrow V_2$ , linear e bijetora (em dimensão finita isso requer que  $V_1$  e  $V_2$  tenham a mesma dimensão e, portanto, que  $V$  tenha dimensão par, como mencionado na Proposição 3.7). De fato, para  $v \in V$  da forma  $v = v_1 \oplus v_2$  com  $v_1 \in V_1$  e  $v_2 \in V_2$ , definindo  $Jv := (-\mathcal{A}^{-1}v_2) \oplus (\mathcal{A}v_1)$  é fácil constatar que  $J^2 = -\mathbb{1}$ .

Para um tal  $J$  o produto por um escalar complexo  $\lambda = x + iy$ , com  $x, y \in \mathbb{R}$ , fica definido por

$$\lambda \cdot (v_1 \oplus v_2) := (x + yJ)(v_1 \oplus v_2) = x(v_1 \oplus v_2) + y((-\mathcal{A}^{-1}v_2) \oplus (\mathcal{A}v_1)) = (xv_1 - y\mathcal{A}^{-1}v_2) \oplus (xv_2 + y\mathcal{A}v_1).$$

Se  $V$  é um espaço de Hilbert real separável com uma base  $\{\phi_k, k \in \mathbb{N}\}$ , podemos tomar  $V_1$  e  $V_2$  como os espaço gerados por  $\{\phi_k, k \in \mathbb{N}, k \text{ par}\}$  e  $\{\phi_k, k \in \mathbb{N}, k \text{ ímpar}\}$ , respectivamente. Uma possível escolha para a bijeção linear  $\mathcal{A} : V_1 \rightarrow V_2$  seria

$$\mathcal{A} \left( \sum_{m=0}^{\infty} a_{2m} \phi_{2m} \right) = \sum_{m=0}^{\infty} a_{2m} \phi_{2m+1},$$

para a qual

$$\mathcal{A}^{-1} \left( \sum_{m=0}^{\infty} a_{2m+1} \phi_{2m+1} \right) = \sum_{m=0}^{\infty} a_{2m+1} \phi_{2m},$$

ou seja, em termos de elementos da base,  $\mathcal{A}\phi_{2m} = \phi_{2m+1}$  e  $\mathcal{A}^{-1}\phi_{2m+1} = \phi_{2m}$  para todo  $m \geq 0$ . Com essa definição, teríamos

$$J \left[ \left( \sum_{m=0}^{\infty} a_{2m} \phi_{2m} \right) \oplus \left( \sum_{m=0}^{\infty} a_{2m+1} \phi_{2m+1} \right) \right] = \left[ \left( -\sum_{m=0}^{\infty} a_{2m+1} \phi_{2m} \right) \oplus \left( \sum_{m=0}^{\infty} a_{2m} \phi_{2m+1} \right) \right].$$

O produto com escalares complexos  $\lambda = x + iy$ , com  $x, y \in \mathbb{R}$ , fica definido por

$$(x + iy) \cdot \sum_{m=0}^{\infty} a_m \phi_m = \left( \sum_{m=0}^{\infty} (xa_{2m} - ya_{2m+1}) \phi_{2m} \right) \oplus \left( \sum_{m=0}^{\infty} (xa_{2m+1} + ya_{2m}) \phi_{2m+1} \right).$$

Para um tal  $J$  o produto por um escalar complexo  $\lambda = x + iy$  com  $x, y \in \mathbb{R}$  fica definido por

$$\lambda \cdot (v_1 \oplus v_2) := (x + yJ)(v_1 \oplus v_2) = x(v_1 \oplus v_2) + y((-\mathcal{A}^{-1}v_2) \oplus (\mathcal{A}v_1)) = (xv_1 - y\mathcal{A}^{-1}v_2) \oplus (xv_2 + y\mathcal{A}v_1).$$

Para  $\alpha, \beta \in V$  da forma  $\alpha = \sum_{m=0}^{\infty} \alpha_m \phi_m$ ,  $\beta = \sum_{m=0}^{\infty} \beta_m \phi_m$  e  $\varepsilon(\alpha, \beta) := \sum_{m=0}^{\infty} \alpha_m \beta_m$ , o produto escalar real usual, constatamos que

$$\varepsilon(\alpha, J\beta) = -\sum_{m=0}^{\infty} \alpha_{2m} \beta_{2m+1} + \sum_{m=0}^{\infty} \alpha_{2m+1} \beta_{2m} \quad \text{e que} \quad \varepsilon(J\alpha, \beta) = -\sum_{m=0}^{\infty} \alpha_{2m+1} \beta_{2m} + \sum_{m=0}^{\infty} \alpha_{2m} \beta_{2m+1}.$$

Assim,  $\varepsilon(\alpha, J\beta) = -\varepsilon(J\alpha, \beta)$  e pela parte **B** da Proposição 3.6, página 249,  $\langle \alpha, \beta \rangle_{J, \varepsilon} := \varepsilon(\alpha, \beta) + i\varepsilon(J\alpha, \beta)$  é um produto escalar complexo. Explicitamente, tem-se

$$\langle \alpha, \beta \rangle_{J, \varepsilon} = \sum_{m=0}^{\infty} \overline{(\alpha_{2m} + i\alpha_{2m+1})} (\beta_{2m} + i\beta_{2m+1}).$$

**E. 3.32** *Exercício.* Verifique! Verifique também que  $\langle \alpha, \lambda \cdot \beta \rangle_{J, \varepsilon} = \lambda \langle \alpha, \beta \rangle_{J, \varepsilon}$  para todo  $\lambda \in \mathbb{C}$ . ✱

A forma simplética real associada a  $\varepsilon$  pela parte **I** do Lema 3.2, página 248, é

$$\sigma(\alpha, \beta) = -\varepsilon(\alpha, J\beta) = \sum_{m=0}^{\infty} \alpha_{2m} \beta_{2m+1} - \sum_{m=0}^{\infty} \alpha_{2m+1} \beta_{2m}.$$

□

**Exemplo 3.11** Uma situação que não se deve deixar de comentar é a seguinte. Se  $V$  é um espaço vetorial complexo com um produto escalar complexo  $\langle \cdot, \cdot \rangle$ ,  $V$  é naturalmente também um espaço vetorial real, sendo que, como comentamos à página 231,  $\sigma(u, v) := \text{Im}(\langle u, v \rangle)$   $u, v \in V$ , define uma forma simplética real em  $V$ . Definindo em  $V$  o operador linear  $Ju = iu$ , tem-se  $J^2 = -\mathbb{1}$ . A multiplicação por escalares complexos não apresenta novidades: para  $x, y \in \mathbb{R}$  e  $u \in V$  vale, pela definição,  $(x + iy) \cdot u = xu + yJu = (x + iy)u$ .

É fácil constatar que  $\sigma(u, Jv) = \text{Im}(\langle u, iv \rangle) = -\text{Im}(\langle iu, v \rangle) = -\sigma(Ju, v)$  e que  $\sigma(u, Ju) = \text{Im}(\langle u, iu \rangle) = \langle u, u \rangle \geq 0$ . Assim, pela parte **A** da Proposição 3.6, página 249,  $\langle u, v \rangle_{J, \sigma} := \sigma(u, Jv) + i\sigma(u, v)$  é um produto escalar complexo em  $V$ . No entanto, é fácil ver que nesse caso  $\langle u, v \rangle_{J, \sigma} = \text{Im}(\langle u, iv \rangle) + i\text{Im}(\langle u, v \rangle) = \text{Re}(\langle u, v \rangle) + i\text{Im}(\langle u, v \rangle) = \langle u, v \rangle$ .

O produto escalar real  $\varepsilon$  associado a  $\sigma$  pela parte **II** do Lema 3.2, página 248, é

$$\varepsilon(u, v) = \sigma(u, Jv) = \text{Im}(\langle u, iv \rangle) = \text{Re}(\langle u, v \rangle).$$

É interessante notar também que se tivéssemos adotado  $Ju = -iu$ ,  $u \in V$ , teríamos ainda para  $\sigma(u, v) = \text{Im}(\langle u, v \rangle)$  que  $\sigma(u, Jv) = -\sigma(Ju, v)$ . Porém,  $\sigma(u, Ju) = -\langle u, u \rangle \leq 0$ , violando a condição de positividade. □

**Exemplo 3.12** Uma situação um pouco diferente é a seguinte. Seja  $V$  um espaço vetorial complexo dotado de um produto escalar complexo  $\langle \cdot, \cdot \rangle$ . Sejam  $V_1$  e  $V_2$  dois subespaços ortogonais de  $V$  (ortogonais segundo o produto escalar  $\langle \cdot, \cdot \rangle$ ). Encarando  $V$  como um espaço real, definamos o operador linear  $J : V \rightarrow V$  por  $J(v_1 \oplus v_2) = i(v_1 \oplus (-v_2))$ , onde  $v_1 \in V_1$  e  $v_2 \in V_2$ . É claro que  $J^2 = -\mathbb{1}$ . A multiplicação por escalares complexos  $x + iy$ , com  $x, y \in \mathbb{R}$ , fica

$$(x + iy) \cdot (v_1 \oplus v_2) = x(v_1 \oplus v_2) + yJ(v_1 \oplus v_2) = ((x + iy)v_1) \oplus ((x - iy)v_2),$$

ou seja,  $\lambda \cdot (v_1 \oplus v_2) = (\lambda v_1) \oplus (\bar{\lambda} v_2)$ , para todos  $\lambda \in \mathbb{C}$ ,  $v_1 \in V_1$  e  $v_2 \in V_2$ .

É também fácil constatar que para o produto escalar real  $\varepsilon(u, v) = \text{Re}(\langle u, v \rangle)$  vale a relação  $\varepsilon(u, Jv) = -\varepsilon(Ju, v)$  (para isso é essencial que  $V_1$  e  $V_2$  sejam ortogonais segundo  $\langle \cdot, \cdot \rangle$ ).

A forma simplética real  $\sigma$  associada a  $\varepsilon$  pela parte **I** do Lema 3.2, página 248, é, tomando  $u = u_1 \oplus u_2$ ,  $v = v_1 \oplus v_2$ , com  $u_1, v_1 \in V_1$  e  $u_2, v_2 \in V_2$ ,

$$\sigma(u, v) := \varepsilon(Ju, v) = \text{Im}(\langle u_1, v_1 \rangle) - \text{Im}(\langle u_2, v_2 \rangle),$$

como facilmente se verifica.

Pela parte **B** da Proposição 3.6, página 249,  $\langle u, v \rangle_{J, \varepsilon} := \varepsilon(u, v) + i\varepsilon(Ju, v)$  é um produto escalar complexo. Por essa definição, tem-se, tomando  $u = u_1 \oplus u_2$ ,  $v = v_1 \oplus v_2$ , com  $u_1, v_1 \in V_1$  e  $u_2, v_2 \in V_2$ ,

$$\begin{aligned} \langle u, v \rangle_{J, \varepsilon} &= \langle (u_1 \oplus u_2), (v_1 \oplus v_2) \rangle_{J, \varepsilon} \\ &= \text{Re}(\langle u_1, v_1 \rangle) + \text{Re}(\langle u_2, v_2 \rangle) + i(\text{Re}(\langle iu_1, v_1 \rangle) + \text{Re}(\langle -iu_2, v_2 \rangle)) \\ &= \text{Re}(\langle u_1, v_1 \rangle) + \text{Re}(\langle u_2, v_2 \rangle) + i\text{Im}(\langle u_1, v_1 \rangle) - i\text{Im}(\langle u_2, v_2 \rangle) \\ &= \langle u_1, v_1 \rangle + \overline{\langle u_2, v_2 \rangle}. \end{aligned}$$

**E. 3.33** *Exercício.* Verifique também que  $\langle u, \lambda \cdot v \rangle_{J, \varepsilon} = \lambda \langle u, v \rangle_{J, \varepsilon}$  para todo  $\lambda \in \mathbb{C}$ . ★

□

# Apêndices

## 3.A Equivalência de Normas em Espaços Vetoriais de Dimensão Finita

Apresentamos aqui a demonstração do Teorema 3.2, página 235, que afirma que todas as normas em um espaço vetorial de dimensão finita sobre  $\mathbb{C}$  ou  $\mathbb{R}$  são equivalentes.

A demonstração que segue faz uso de algumas noções e resultados elementares sobre topologias métricas. O leitor interessado deve seguir as referências dadas abaixo aos pontos destas Notas onde tais noções e resultados são tratados.

**Prova do Teorema 3.2.** Seja  $V$  um espaço vetorial de dimensão finita, de sorte que existe uma base  $B = \{b_1, \dots, b_n\}$  de vetores linearmente independentes de  $V$  tais que todo  $u \in V$  pode ser escrito de modo único como uma combinação linear  $u = \alpha_1 b_1 + \dots + \alpha_n b_n$  dos vetores de  $B$ , onde os coeficientes  $\alpha_k$  são reais ou complexos (dependendo de  $V$  ser um espaço vetorial sobre  $\mathbb{R}$  ou sobre  $\mathbb{C}$ ). Fixada uma base  $B$ , podemos definir uma norma  $\|\cdot\|_E$  em  $V$  por

$$\|u\|_E = \sqrt{\sum_{k=1}^n |\alpha_k|^2},$$

onde, como acima,  $u = \alpha_1 b_1 + \dots + \alpha_n b_n$ .

Seja agora  $\|\cdot\|$  uma outra norma definida em  $V$ . Temos que

$$\|u\| = \left\| \alpha_1 b_1 + \dots + \alpha_n b_n \right\| \stackrel{(3.23)}{\leq} |\alpha_1| \|b_1\| + \dots + |\alpha_n| \|b_n\| \stackrel{(3.17)}{\leq} \sqrt{\sum_{k=1}^n |\alpha_k|^2} \sqrt{\sum_{k=1}^n \|b_k\|^2}.$$

Assim, estabelecemos que para todo  $u \in V$  vale

$$\|u\| \leq M_1 \|u\|_E, \tag{3.A.1}$$

com  $M_1 := \sqrt{\sum_{k=1}^n \|b_k\|^2}$  sendo uma constante positiva independente de  $u$ .

Para todos  $u, v \in V$  vale  $\left| \|u\| - \|v\| \right| \stackrel{(3.24)}{\leq} \|u - v\| \stackrel{(3.A.1)}{\leq} M_1 \|u - v\|_E$ . Essa relação estabelece que a função  $\varphi : \mathbb{R}^n \rightarrow \mathbb{R}$  (ou  $\varphi : \mathbb{C}^n \rightarrow \mathbb{R}$ ) definida por

$$\varphi(\alpha_1, \dots, \alpha_n) := \left\| \alpha_1 b_1 + \dots + \alpha_n b_n \right\|$$

é contínua na topologia métrica usual de  $\mathbb{R}^n$  (ou  $\mathbb{C}^n$ ), pois mostra (com  $u = \alpha_1 b_1 + \dots + \alpha_n b_n$  e  $v = \beta_1 b_1 + \dots + \beta_n b_n$ ) que

$$\left| \varphi(\alpha_1, \dots, \alpha_n) - \varphi(\beta_1, \dots, \beta_n) \right| \leq M_1 \sqrt{\sum_{k=1}^n |\alpha_k - \beta_k|^2},$$

provando que se  $(\alpha_1, \dots, \alpha_n)$  converge a  $(\beta_1, \dots, \beta_n)$  na topologia métrica usual de  $\mathbb{R}^n$  (ou  $\mathbb{C}^n$ ), então  $\varphi(\alpha_1, \dots, \alpha_n)$  converge a  $\varphi(\beta_1, \dots, \beta_n)$ .

Seja  $B_1$  a bola aberta centrada em 0 e de raio 1 em  $\mathbb{R}^n$  (ou  $\mathbb{C}^n$ ) na topologia métrica usual:

$$B_1 := \left\{ (\alpha_1, \dots, \alpha_n) \in \mathbb{R}^n \text{ (ou } \mathbb{C}^n) \mid \sqrt{\sum_{k=1}^n |\alpha_k|^2} < 1 \right\},$$

e seja  $\partial B_1$  seu bordo<sup>28</sup>:

$$\partial B_1 := \left\{ (\alpha_1, \dots, \alpha_n) \in \mathbb{R}^n \text{ (ou } \mathbb{C}^n) \mid \sqrt{\sum_{k=1}^n |\alpha_k|^2} = 1 \right\}.$$

<sup>28</sup>Para a definição da noção de bordo e para a observação que todo bordo é fechado, vide página 1374.

$\partial B_1$  é fechado e limitado e, portanto (pelo *Teorema de Heine-Borel*, Teorema 32.14, página 1553), é compacto na topologia métrica usual. Logo, pelo Teorema 32.16, página 1554, a função contínua  $\varphi$  assume em  $\partial B_1$  um mínimo  $M_2 \geq 0$  e, portanto,

$$\varphi(\alpha_1, \dots, \alpha_n) \geq M_2 \tag{3.A.2}$$

para toda  $n$ -upla  $(\alpha_1, \dots, \alpha_n)$  com  $\sqrt{\sum_{k=1}^n |\alpha_k|^2} = 1$ .

Seja  $(\gamma_1, \dots, \gamma_n)$  um ponto de  $\partial B_1$  onde o mínimo de  $\varphi$  é assumido e seja  $v_0 = \gamma_1 b_1 + \dots + \gamma_n b_n$ . O fato que  $(\gamma_1, \dots, \gamma_n) \in \partial B_1$  significa, evidentemente, que  $\|v_0\|_E = 1$ . A constante  $M_2$  não pode ser nula, pois se o fosse teríamos  $\|v_0\| = 0$ , ou seja,  $v_0 = 0$ , o que contraria  $\|v_0\|_E = 1$ .

Segue de (3.A.2) que  $\|\alpha_1 b_1 + \dots + \alpha_n b_n\| \geq M_2$  para todo vetor  $u = \alpha_1 b_1 + \dots + \alpha_n b_n$  com  $\|u\|_E = 1$ . Como para todo  $v \in V, v \neq 0$ , tem-se, evidentemente, que  $\left\| \frac{1}{\|v\|_E} v \right\|_E = 1$ , segue que

$$\left\| \frac{1}{\|v\|_E} v \right\| \geq M_2, \quad \text{ou seja,} \quad \|v\| \geq M_2 \|v\|_E,$$

sendo que a última desigualdade vale também, evidentemente, para  $v = 0$ . Provamos, portanto, que existem constantes  $M_1$  e  $M_2$  com  $M_2 > 0$  tais que para todo vetor  $v \in V$ ,

$$M_2 \|v\|_E \leq \|v\| \leq M_1 \|v\|_E,$$

estabelecendo que toda norma  $\|\cdot\|$  é equivalente à norma  $\|\cdot\|_E$ . Como a equivalência de normas é uma relação de equivalência, segue que todas as normas em  $V$  são equivalentes. ■

### 3.B Prova do Teorema de Fréchet, von Neumann e Jordan

Nesta Seção apresentamos a demonstração do Teorema de Fréchet, von Neumann e Jordan, Teorema 3.3, página 238.

Vamos supor que  $\|\cdot\|$  seja uma norma em um espaço vetorial complexo  $V$  e que satisfaça a identidade do paralelogramo

$$\|a + b\|^2 + \|a - b\|^2 = 2\|a\|^2 + 2\|b\|^2 \tag{3.B.3}$$

para todos  $a, b \in V$ . Defina-se, para  $u, v \in V$ ,

$$\omega(u, v) := \frac{1}{4} \sum_{n=0}^3 i^{-n} \|u + i^n v\|^2,$$

ou seja, escrevendo os termos da soma explicitamente,

$$\omega(u, v) := \frac{1}{4} \left[ \|u + v\|^2 - \|u - v\|^2 - i \left( \|u + iv\|^2 - \|u - iv\|^2 \right) \right]. \tag{3.B.4}$$

Vale a propriedade Hermitiana

$$\overline{\omega(u, v)} = \omega(v, u) \tag{3.B.5}$$

para todos  $u, v \in V$  pois, como  $\|a\| = \|-a\|$  e  $\|a\| = \|ia\|$  para todo  $a \in V$ , segue que

$$\begin{aligned} \overline{\omega(u, v)} &= \frac{1}{4} \left[ \|u + v\|^2 - \|u - v\|^2 + i \left( \|u + iv\|^2 - \|u - iv\|^2 \right) \right] \\ &= \frac{1}{4} \left[ \|v + u\|^2 - \|v - u\|^2 + i \left( \|iu - v\|^2 - \|iu + v\|^2 \right) \right] \\ &= \frac{1}{4} \left[ \|v + u\|^2 - \|v - u\|^2 - i \left( \|v + iu\|^2 - \|v - iu\|^2 \right) \right] \\ &= \omega(v, u). \end{aligned}$$

É importante observar que, por (3.B.4),

$$\omega(u, u) := \frac{1}{4} \left[ \|2u\|^2 - \|u - u\|^2 - i \left( \|(1+i)u\|^2 - \|(1-i)u\|^2 \right) \right] = \|u\|^2, \quad (3.B.6)$$

já que, do fato que  $|1+i| = |1-i|$ , segue pelas propriedades definidoras de uma norma que  $\|(1+i)u\| = |1+i| \|u\| = |1-i| \|u\| = \|(1-i)u\|$ .

Assim, estabelecemos que para todo  $u \in V$  vale  $\omega(u, u) = \|u\|^2$ , o que implica, pelas propriedades definidoras de uma norma, que  $\omega(u, u) \geq 0$ , sendo que  $\omega(u, u) = 0$  se e somente se  $u = 0$ .

Para provar que  $\omega$  é um produto escalar, resta-nos provar que  $\omega$  é uma forma sesquilinear. Como  $\omega$  tem a propriedade Hermitiana (3.B.5), é suficiente provar que  $\omega$  é linear na segunda variável. De fato, esse é o único ponto não-trivial da demonstração do Teorema 3.3 e o único em que a identidade do paralelogramo é usada. O leitor verá que a demonstração de que  $\omega$  é linear na segunda variável é engenhosa, sendo feita, sucessivamente, primeiro para números inteiros, depois para racionais, depois para números reais e, por fim, para números complexos.

Definindo-se, para  $u, v \in V$ ,

$$f(u, v) := \|u + v\|^2 - \|u - v\|^2, \quad (3.B.7)$$

podemos escrever, por (3.B.4),

$$\omega(u, v) := \frac{1}{4} \left[ f(u, v) - i f(u, iv) \right]. \quad (3.B.8)$$

Segue facilmente da definição (3.B.7) que

$$f(u, v) = f(v, u), \quad (3.B.9)$$

$$f(u, -v) = -f(u, v), \quad (3.B.10)$$

$$f(u, 0) = 0. \quad (3.B.11)$$

A seguinte proposição é fundamental para a prova de que  $\omega$  é uma forma sesquilinear e em sua demonstração é feito uso da identidade do paralelogramo.

**Proposição 3.8** Para todos  $u, v$  e  $w \in V$  vale

$$f(u, v + w) = f(u, v) + f(u, w). \quad (3.B.12)$$

Por (3.B.9), segue que  $f(u + v, w) = f(u, w) + f(v, w)$ , também para todos  $u, v$  e  $w \in V$ . □

*Prova.* Precisamos apenas provar (3.B.12), o que é feito com uso da identidade do paralelogramo (3.B.3). Por (3.B.3) com  $a = u + v$  e  $b = w$ , vê-se que

$$\|u + v + w\|^2 = 2\|u + v\|^2 + 2\|w\|^2 - \|u + v - w\|^2.$$

Trocando-se  $v \rightarrow -v$  e  $w \rightarrow -w$ , segue disso que

$$\|u - v - w\|^2 = 2\|u - v\|^2 + 2\|w\|^2 - \|u - v + w\|^2.$$

Logo, como  $f(u, v + w) = \|u + v + w\|^2 - \|u - v - w\|^2$ , segue que

$$f(u, v + w) = 2\|u + v\|^2 - 2\|u - v\|^2 + \|u - v + w\|^2 - \|u + v - w\|^2.$$

Assim, provamos que

$$f(u, v + w) = 2f(u, v) + f(u, w - v). \quad (3.B.13)$$

Trocando  $v \leftrightarrow w$ , isso fica  $f(u, v + w) = 2f(u, w) + f(u, v - w)$  e, por (3.B.10), concluímos que vale também

$$f(u, v + w) = 2f(u, w) - f(u, w - v). \quad (3.B.14)$$

Somando (3.B.13) e (3.B.14), obtemos  $f(u, v + w) = f(u, v) + f(u, w)$ , que é o que queríamos. ■

Tomando  $v = w$ , (3.B.12) implica que  $f(u, 2v) = 2f(u, v)$ . Vamos assumir que para algum  $n \in \mathbb{N}$ , valha  $f(u, nv) = nf(u, v)$ . Isso é verdadeiro para  $n = 0$  (por (3.B.11)) e  $n = 1$  (trivialmente) e vale também, como vimos, para  $n = 2$ . Então,

$$f(u, (n+1)v) = f(u, v + nv) \stackrel{(3.B.12)}{=} f(u, v) + f(u, nv) \stackrel{\text{hipótese}}{=} f(u, v) + nf(u, v) = (n+1)f(u, v).$$

Com isso, provamos por indução que

$$f(u, nv) = nf(u, v) \tag{3.B.15}$$

para todo  $n \in \mathbb{N}$  e todos  $u, v \in V$ . Substituindo  $v$  por  $\frac{1}{n}v$ , isso está também dizendo que

$$f\left(u, \frac{1}{n}v\right) = \frac{1}{n}f(u, v), \tag{3.B.16}$$

também para todo  $n \in \mathbb{N}$ , e todos  $u, v \in V$ . Assim, se  $p$  e  $q$  são inteiros positivos  $q \neq 0$ , vale

$$f\left(u, \frac{p}{q}v\right) \stackrel{(3.B.15)}{=} pf\left(u, \frac{1}{q}v\right) \stackrel{(3.B.16)}{=} \frac{p}{q}f(u, v).$$

Por (3.B.10) e por (3.B.11), segue disso que

$$f(u, rv) = rf(u, v) \tag{3.B.17}$$

para todo  $r \in \mathbb{Q}$  e todos  $u, v \in V$ .

Seja  $x \in \mathbb{R}$  e seja  $r_k, k \in \mathbb{N}$ , uma sequência de números racionais tal que  $\lim_{k \rightarrow \infty} r_k = x$ . Então, usando a desigualdade (3.18), página 233, com  $a = (r_k - x)v$  e  $b = -u - xv$ , tem-se que  $\left| \|u + r_kv\| - \|u + xv\| \right| \leq \|(r_k - x)v\| = |r_k - x| \|v\|$  e como  $\lim_{k \rightarrow \infty} |r_k - x| = 0$ , segue que  $\lim_{k \rightarrow \infty} \left| \|u + r_kv\| - \|u + xv\| \right| = 0$ , ou seja,  $\lim_{k \rightarrow \infty} \|u + r_kv\| = \|u + xv\| = \left\| u + \lim_{k \rightarrow \infty} r_kv \right\|$ . Isso implica imediatamente que

$$\lim_{k \rightarrow \infty} f(u, r_kv) = f\left(u, \lim_{k \rightarrow \infty} r_kv\right) \tag{3.B.18}$$

e, portanto, provamos que

$$f(u, xv) = xf(u, v), \tag{3.B.19}$$

para todo  $x \in \mathbb{R}$  e todos  $u, v \in V$ , pois

$$f(u, xv) = f\left(u, \lim_{k \rightarrow \infty} r_kv\right) \stackrel{(3.B.18)}{=} \lim_{k \rightarrow \infty} f(u, r_kv) \stackrel{(3.B.17)}{=} \lim_{k \rightarrow \infty} r_k f(u, v) = xf(u, v).$$

Sejam agora  $x, y \in \mathbb{R}$ . Tem-se, pelo exposto acima,

$$f(u, (x + iy)v) = f(u, xv + iyv) \stackrel{(3.B.12)}{=} f(u, xv) + f(u, iyv) \stackrel{(3.B.19)}{=} xf(u, v) + yf(u, iv). \tag{3.B.20}$$

Por (3.B.8), segue que

$$\begin{aligned} \omega(u, (x + iy)v) &= \frac{1}{4} \left[ f(u, (x + iy)v) - if(u, (x + iy)iv) \right] \\ &\stackrel{(3.B.20)}{=} \frac{1}{4} \left[ (xf(u, v) + yf(u, iv)) - i(xf(u, iv) + yf(u, -v)) \right] \\ &\stackrel{(3.B.10)}{=} \frac{1}{4} \left[ (xf(u, v) + yf(u, iv)) - i(xf(u, iv) - yf(u, v)) \right] \\ &= x \frac{1}{4} [f(u, v) - if(u, iv)] + iy \frac{1}{4} [f(u, v) - if(u, iv)] \\ &= (x + iy)\omega(u, v). \end{aligned}$$

Com isso, provamos que para todo  $z \in \mathbb{C}$  e todos  $u, v \in V$  vale  $\omega(u, zv) = z\omega(u, v)$ . Pela propriedade Hermitiana (3.B.5), segue também que  $\omega(zu, v) = \bar{z}\omega(u, v)$ . Isso estabeleceu que  $\omega$  é uma forma sesquilinear, completando a prova do Teorema 3.3. ■



## Parte II

# Tópicos de Análise Real e Complexa