Capítulo 3
Formas Lineares e Normas em Espaços Vetoriais

Sumário

A noção de espaço vetorial que introduzimos na Seção 2.1.5, página 140, 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*[1] em V é uma função $\omega : V^n \to \mathbb{K}$ que seja linear em cada um dos seus argumentos, ou seja, para todo α, $\beta \in \mathbb{K}$, todos $v_1, \ldots, v_n \in V$, $v_i' \in V$

[1] Também chamada *forma n-linear*, *n-forma linear*, ou simplesmente *n-forma*.

e todo $i = 1, \ldots, n$ vale

$$\omega\Big(v_1, \ldots, v_{i-1}, (\alpha v_i + \beta v_i'), v_{i+1}, \ldots, v_n\Big)$$
$$= \alpha\omega\Big(v_1, \ldots, v_{i-1}, v_i, v_{i+1}, \ldots, v_n\Big) + \beta\omega\Big(v_1, \ldots, v_{i-1}, v_i', v_{i+1}, \ldots, v_n\Big) . \quad (3.1)$$

O seguinte fato importante é consequência imediata da definição acima: se ω é uma n-forma multilinear, então, para cada $i = 1, \ldots, n$,

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

ou seja, se um dos argumentos é o vetor nulo a forma se anula.

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

Um fato importante é o seguinte: o conjunto de todas as formas n-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 formas n-lineares ω_1 e ω_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 forma n-linear que a toda n-upla de vetores $v_1, \ldots, v_n \in V$ associa

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

E. 3.2 *Exercício*. Complete os detalhes da prova que o conjunto de todas as formas n-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 $\omega : V^2 \to \mathbb{K}$ 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

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

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: $V = \{x = (x_1, \ldots, x_n),\ x_i \in \mathbb{R}\}$. Uma forma bilinear em V é dada por

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

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

Uma forma bilinear ω 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 ω é uma forma bilinear, as formas ω_r e ω_a definidas por $\omega_r(u, v) := \frac{1}{2}\big(\omega(u, v) + \omega(v, u)\big)$ e $\omega_a(u, v) := \frac{1}{2}\big(\omega(u, v) - \omega(v, u)\big)$ 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.

• Formas bilineares não degeneradas

Uma forma bilinear simétrica ou antissimétrica ω 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 ω 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 $l : V \to V'$ que associa cada $u \in V$ ao funcional linear l_u acima for um isomorfismo de espaços vetoriais a forma bilinear ω 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.

• A identidade de polarização para forma bilineares simétricas

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

$$\omega\big(u,\ v\big) \;=\; \frac{1}{2}\Big[\omega\big(u+v,\ u+v\big) - \omega\big(u,\ u\big) - \omega\big(v,\ v\big)\Big]\ , \tag{3.3}$$

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

E. 3.3 *Exercício*. Verifique! ✠

E. 3.4 *Exercício*. Verifique também que

$$\omega\big(u,\ v\big) \;=\; \frac{1}{4}\Big[\omega\big(u+v,\ u+v\big) - \omega\big(u-v,\ u-v\big)\Big]\ , \tag{3.4}$$

também para formas bilineares simétricas. Essa identidade é também denominada *identidade de polarização*. ✠

As identidades (3.3) e (3.4) serão generalizadas para formas n-lineares simétricas, respectivamente, em (3.6) e (3.7), página 264.

A importância da identidade de polarização (3.3) ou (3.4) é mostrar que, no caso de formas bilineares simétricas, o conhecimento dos valores "diagonais" $\omega(u,\ u)$ para todos os $u \in V$ determina os valores $\omega(u,\ v)$ para todos $u,\ v \in V$. Essa afirmação será generalizada para formas multilineares simétricas na Proposição 3.1, página 264, com uso de (3.6) e (3.7).

No Apêndice 3.D, página 317, apresentamos duas versões explícitas da identidade de polarização para formas trilineares simétricas.

• Formas multilineares simétricas

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

$$\omega\left(v_{\pi(1)},\ \ldots,\ v_{\pi(n)}\right) \;=\; \omega\left(v_1,\ \ldots,\ v_n\right)\ , \tag{3.5}$$

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

• A identidade de polarização para formas multilineares simétricas

Vamos generalizar as considerações feitas acima sobre a identidade de polarização para formas bilineares simétricas (vide (3.4), página 264) para o caso de formas multilineares simétricas. Para $n \in \mathbb{N}$ com $n \geq 2$, seja ω uma forma n-linear e simétrica em um espaço vetorial real V, ou seja, tal que valha $\omega\big(v_1,\ \ldots,\ v_n\big) = \omega\big(v_{\pi(1)},\ \ldots,\ v_{\pi(n)}\big)$ para quaisquer vetores $v_k \in V$, $k = 1,\ \ldots,\ n$, e qualquer $\pi \in S_n$, o grupo de permutações de n elementos. Temos o seguinte:

Proposição 3.1 *Seja V um espaço vetorial real, seja $n \in \mathbb{N}$ com $n \geq 2$ e seja ω uma forma n-linear simétrica em V. Então, para quaisquer $u_l \in V$, $l = 1,\ \ldots,\ n$, vale a seguinte identidade:*

$$\omega\big(u_1,\ \ldots,\ u_n\big) \;=\; \frac{(-1)^n}{n!}\sum_{\epsilon_1=0}^{1}\cdots\sum_{\epsilon_n=0}^{1}(-1)^{\epsilon_1+\cdots+\epsilon_n}\omega\left(v_\epsilon,\ \ldots,\ v_\epsilon\right)\ , \quad \textit{onde} \quad v_\epsilon \;:=\; \sum_{k=1}^{n}\epsilon_k u_k\ . \tag{3.6}$$

Essa relação também pode ser expressa de forma alternativa como

$$\omega\big(u_1,\ \ldots,\ u_n\big) \;=\; \frac{1}{2^n n!}\sum_{\epsilon_1=0}^{1}\cdots\sum_{\epsilon_n=0}^{1}(-1)^{\epsilon_1+\cdots+\epsilon_n}\omega\left(w_\epsilon,\ \ldots,\ w_\epsilon\right)\ , \quad \textit{onde} \quad w_\epsilon \;:=\; \sum_{k=1}^{n}(-1)^{\epsilon_k}u_k\ . \tag{3.7}$$

Ambas as relações (3.6) e (3.7) generalizam a identidade de polarização (3.4) para formas multilineares simétricas e também são denominadas identidade de polarização*. Elas nos informam que no caso de formas multilineares simétricas, o conhecimento dos valores "diagonais" $\omega(u, \ldots, u)$ para todos os $u \in V$ determina os valores gerais $\omega(u_1, \ldots, u_n)$ para todos os $u_1, \ldots, u_n \in V$.*

Uma consequência imediata de (3.6) ou (3.7) é a seguinte afirmação relevante, que destacamos para futuro referenciamento: se ω é uma forma n-linear simétrica satisfazendo $\omega\big(u, \ldots, u\big) = 0$ para todo $u \in V$, então ω é identicamente nula: $\omega = 0$. □

Para não quebrarmos o ritmo da exposição, apresentamos a demonstração da Proposição 3.1 e das relações (3.6) e (3.7) no Apêndice 3.C, página 316. A Proposição 3.1 será evocada, por exemplo, na Seção 35.4.4.1, página 1983 (especificamente, no Apêndice 35.B, página 2025), quando tratarmos de coordenadas normais de Riemann.

E. 3.5 *Exercício.* Seja ω é uma forma bilinear simétrica, Verifique que (3.6) afirma no caso $n = 2$ que

$$\omega(u_1, u_2) \;=\; \frac{1}{2}\Big(\omega(u_1 + u_2, u_1 + u_2) - \omega(u_1, u_1) - \omega(u_2, u_2)\Big)$$

e que (3.7) afirma no caso $n = 2$ que

$$\begin{aligned}\omega(u_1, u_2) \;&=\; \frac{1}{8}\Big(\omega(u_1 + u_2, u_1 + u_2) - \omega(u_1 - u_2, u_1 - u_2) - \omega(-u_1 + u_2, -u_1 + u_2) + \omega(-u_1 - u_2, -u_1 - u_2)\Big)\\ &=\; \frac{1}{4}\Big(\omega(u_1 + u_2, u_1 + u_2) - \omega(u_1 - u_2, u_1 - u_2)\Big)\,,\end{aligned}$$

onde usamos adicionalmente a multilinearidade de ω. Compare essas relações à identidade de polarização (3.3) ou (3.4), página 264. ✣

No Apêndice 3.D, página 317, apresentamos uma versão explícita da identidade de polarização para formas trilineares simétricas.

• Formas multilineares alternantes

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

$$\omega\left(v_1, \ldots, v_{i-1}, v_i, v_{i+1}, v_{i+2}, \ldots, v_n\right) \;=\; -\omega\left(v_1, \ldots, v_{i-1}, v_{i+1}, v_i, v_{i+2}, \ldots, v_n\right) \tag{3.8}$$

para todos os vetores $v_1, \ldots, v_n \in V$ e todo $i = 1, \ldots, n-1$. 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 ω é uma forma n-linear alternante, então para todo $\pi \in S_n$, o grupo de permutações de n elementos, vale

$$\omega\left(v_{\pi(1)}, \ldots, v_{\pi(n)}\right) \;=\; (\mathrm{sinal}\,\pi)\,\omega\left(v_1, \ldots, v_n\right)\,, \tag{3.9}$$

para todos os vetores $v_1, \ldots, v_n \in V$, onde $\mathrm{sinal}\,\pi$ é o sinal da permutação π (definida na Seção 21.1.1.2, página 1164).

E. 3.6 *Exercício.* Está claro? ✣

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

O conjunto de todas as formas n-lineares alternantes em um espaço vetorial V sobre um corpo $\mathbb{K}$ é igualmente um espaço vetorial sobre $\mathbb{K}$: para duas formas n-lineares alternantes ω_1 e ω_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 forma n-linear que a toda n-upla de vetores $v_1, \ldots, v_n \in V$ associa

$$(\alpha_1\omega_1 + \alpha_2\omega_2)(v_1, \ldots, v_n) \;:=\; \alpha_1\omega_1(v_1, \ldots, v_n) + \alpha_2\omega_2(v_1, \ldots, v_n)\,.$$

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

E. 3.7 *Exercício.* Complete os detalhes da prova que o conjunto de todas as formas n-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*[2]. 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 $u = 0$.

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, ω_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, ω_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 ω é uma n-forma alternante então $\omega\,(v_1,\ \ldots,\ v_n) = 0$ 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. 3.8 *Exercício*. A propriedade mencionada no último parágrafo é equivalente à definição de forma linear alternante: se ω é uma forma n-linear e $\omega\,(v_1,\ \ldots,\ v_n) = 0$ sempre que $v_i = v_j$ para algum par $i \neq j$, então ω é alternante. Prove isso. Sugestão: para $i \neq j$ defina a forma bilinear $\omega_{ij}(v_i,\ v_j) := \omega\,(v_1,\ \ldots,\ v_n)$ onde todos os vetores $v_1,\ \ldots,\ v_n$ 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 3.2 *Se ω é uma forma n-linear alternante e $v_1,\ \ldots,\ v_n$ são vetores linearmente dependentes, então $\omega\,(v_1,\ \ldots,\ v_n) \ = \ 0$.* □

E. 3.9 *Exercício*. Prove isso. ✠

• **Formas alternantes maximais**

A Proposição 3.2 tem uma consequência imediata: se V é um espaço vetorial de dimensão n e ω é uma forma linear alternante de ordem $m > n$, então $\omega = 0$.

E. 3.10 *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 $\{b_1,\ \ldots,\ b_n\}$ uma base em V. Sejam agora ω_1 e ω_2 duas formas alternantes maximais

[2] Do grego *symplektikós*: que serve para ligar, trançado, enlaçado.

em V e seja $x_1, \ldots, x_n$ uma n-upla de vetores de V. Como $\{b_1, \ldots, b_n\}$ é uma base, podemos sempre escrever

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

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

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

e, analogamente,

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

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

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

e, analogamente,

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

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

E. 3.11 *Exercício*. Justifique. ✠

Como ω_1 é uma forma alternante maximal, tem-se que

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

Assim,

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

e, analogamente,

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

Como se vê nessas últimas expressões, $\omega_1(x_1, \ldots, x_n)$ e $\omega_2(x_1, \ldots, x_n)$ diferem apenas pelos fatores $\omega_1(b_1, \ldots, b_n)$ e $\omega_2(b_1, \ldots, b_n)$, respectivamente. Como esses fatores são apenas números (elementos do corpo $\mathbb{K}$), são proporcionais um ao outro. Isso prova então que $\omega_1(x_1, \ldots, x_n)$ e $\omega_2(x_1, \ldots, x_n)$ são proporcionais um ao outro para toda n-upla $x_1, \ldots, x_n$ 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 formas n-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*[3] ω_{det} em um espaço vetorial V de dimensão n é a forma n-linear alternante maximal tal que

[3] 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.

$\omega_{det}(b_1, \ldots, b_n) = 1$ no caso em que $\{b_1, \ldots, b_n\}$ é a base canônica de V:

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

Assim,

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

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

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

- **Determinante de matrizes**

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

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

Denotamos por $[\![a_1, \ldots, a_n]\!]$ a matriz $n \times n$ construída de forma que sua k-ésima coluna seja o vetor-coluna a_k, ou seja

$$[\![a_1, \ldots, a_n]\!] = \begin{pmatrix} \alpha_{11} & \cdots & \alpha_{n1} \\ \vdots & \ddots & \vdots \\ \alpha_{1n} & \cdots & \alpha_{nn} \end{pmatrix}.$$

É evidente que toda matriz $A \in \mathrm{Mat}\,(\mathbb{C}, n)$ pode ser escrita na forma $A = [\![a_1, \ldots, a_n]\!]$ para algum conjunto de vetores $a_1, \ldots, a_n$ que representam suas colunas.

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

$$\det(A) := \omega_{det}(a_1, \ldots, a_n) , \tag{3.10}$$

ou seja,

$$\det(A) = \sum_{j \in S_n} \mathrm{sinal}\,(j)\, \alpha_{1j(1)} \cdots \alpha_{nj(n)} . \tag{3.11}$$

Essa expressão é frequentemente denominada *fórmula de Leibniz*[4] 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 são (re)apresentadas e (re)deduzidas na Seção 10.1, página 569. Vide, em particular, o Teorema 10.1, página 574.

Na Seção 11.7, página 731, com base no Teorema de Hadamard, Teorema 10.40, página 684, demonstraremos que o determinante de matrizes é uma função contínua em certas topologias adequadas.

- **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 221, 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 $(U^{\otimes n})'$.

[4]Gottfried Wilhelm von Leibniz (1646–1716).

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

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

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

$$\Psi(\omega) := \sum_{a_1=1}^{m} \cdots \sum_{a_n=1}^{m} \omega(\mathbf{e}_{a_1}, \ldots, \mathbf{e}_{a_n}) \ell_{a_1} \otimes \cdots \otimes \ell_{a_n} . \tag{3.12}$$

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

3.1.2 Formas Sesquilineares e as Desigualdades de Cauchy-Schwarz e Minkowski

• **Formas sesquilineares. Definições**

Seja V um espaço vetorial complexo. Uma *forma sesquilinear*[5] é uma função $\omega : V \times V \to \mathbb{C}$ 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 α e β.

2. Antilinearidade em relação à primeira variável:

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

para todos os vetores u, v e w e para todos os números complexos α e β.

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

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

para todo vetor u.

E. 3.12 *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 $u \in V$,

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

[5] Do radical grego *sesqui:* um e meio.

Abaixo (Teorema 3.1, página 270) provaremos que toda forma sesquilinear positiva é automaticamente Hermitiana. Lá provaremos também que se ω 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 ω é uma forma sesquilinear em V, então vale a seguinte identidade, denominada *identidade de polarização de formas sesquilineares*:

$$\omega(u,\ v) \;=\; \frac{1}{4}\sum_{n=0}^{3} i^{-n}\omega\big((u+i^n v),\ (u+i^n v)\big) \;=\; \frac{1}{4}\sum_{n=0}^{3} i^{n}\omega\big((u+i^{-n} v),\ (u+i^{-n} v)\big)\ , \tag{3.13}$$

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

E. 3.13 *Exercício*. Verifique! ✠

Como no caso da identidade de polarização para formas lineares simétricas, relação (3.4), a importância da identidade de polarização (3.13) 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 ω 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 $l : V \to V'$ que associa cada $u \in V$ ao funcional linear l_u acima for um anti-isomorfismo[6] de espaços vetoriais a forma sesquilinear ω 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[7]-Schwarz[8].

Teorema 3.1 *Se ω 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,

$$\big|\omega(u,\ v)\big|^2 \;\leq\; \omega(u,\ u)\,\omega(v,\ v)\ . \tag{3.14}$$

Por fim, se ω 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 λ e quaisquer vetores u e v vale, pela hipótese de positividade,

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

Escrevendo-se explicitamente o lado esquerdo temos a desigualdade

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

[6] Definido à página 164.
[7] Augustin Louis Cauchy (1789–1857).
[8] Herman Amandus Schwarz (1843–1921).

E. 3.14 *Exercício*. Verifique isso. ✠

Vamos agora escrever λ na forma $\lambda = x + iy$, onde x é a parte real de λ 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. 3.15 *Exercício*. Verifique essa afirmação. ✠

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

$$\omega(u, v) \;=\; \alpha + i\beta \qquad \text{e} \qquad \omega(v, u) \;=\; \gamma + i\delta \;, \tag{3.15}$$

onde α, β, γ e $\delta \in \mathbb{R}$. Ficamos com

$$f(x, y) \;=\; (x^2 + y^2)\omega(v, v) + (x\alpha - y\beta) + i(x\beta + y\alpha) + (x\gamma + y\delta) + i(x\delta - y\gamma) + \omega(u, u) \;\geq\; 0 \;. \tag{3.16}$$

Como $f(x, y)$ tem de ser real (e ≥ 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 (3.15), isso diz que

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

provando que ω é Hermitiano.

Com as relações $\beta = -\delta$ e $\alpha = \gamma$ a expressão (3.16) fica

$$f(x, y) \;=\; (x^2 + y^2)\omega(v, v) + 2(x\alpha - y\beta) + \omega(u, u) \;. \tag{3.17}$$

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 (3.14) 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 (3.14) para esse caso. Podemos reescrever o lado direito de (3.17) 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. 3.16 *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 ω 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 ω é não degenerada. ■

• **A desigualdade de Minkowski**

A desigualdade de Cauchy-Schwarz tem uma consequência de certa importância, a chamada *desigualdade de Minkowski*[9]: se ω é uma forma sesquilinear positiva (em particular, se ω é um produto escalar), então, para todos os vetores u e v, vale

$$\omega(u-v,\ u-v)^{1/2} \ \leq\ \omega(u,\ u)^{1/2} + \omega(v,\ v)^{1/2} \ . \tag{3.18}$$

A demonstração é simples:

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

que é o que se queria demonstrar. Acima, na passagem da primeira para a segunda linha usamos a Hermiticidade de ω 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 270.

Versões mais gerais da desigualdade de Minkowski serão apresentadas no Teorema 25.4, página 1492, e na Seção 32.4.1, página 1721.

3.1.3 Produtos Escalares

• **Produtos internos ou produtos escalares**

Uma forma sesquilinear positiva ω 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 3.3 *Uma forma sesquilinear positiva é um produto escalar se e somente se for não degenerada.* □

Prova. Se ω 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 270, se ω 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

[9] Hermann Minkowski (1864–1909).

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 $V \times V \to \mathbb{C}$, 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:
$$\langle \alpha u + \beta v,\ w\rangle \ = \ \overline{\alpha}\langle u,\ w\rangle + \overline{\beta}\langle v,\ w\rangle$$
para todos $u,\ v$ e $w \in V$ e todos $\alpha,\ \beta \in \mathbb{C}$, onde $\overline{\alpha}$ 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*:
$$\big|\langle u,\ v\rangle\big|^2 \ \leq \ \big|\langle u,\ u\rangle\big|\,\big|\langle v,\ v\rangle\big|\ .$$
A demonstração (mais geral) é apresentada no Teorema 3.1, página 270.

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 ω, existe uma maneira canônica de construir a partir de V e ω um outro espaço vetorial dotado de um produto escalar.

Seja ω 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 $E : V \to \tilde{V}$ tais que
$$\tilde{\omega}\big(E(u),\ E(v)\big) \ = \ \omega(u,\ v)$$
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
$$\omega(\alpha u + \beta v,\ \alpha u + \beta v) \ = \ |\alpha|^2\omega(u,\ u) + \overline{\alpha}\beta\omega(u,\ v) + \alpha\overline{\beta}\omega(v,\ u) + |\beta|^2\omega(v,\ v) \ = \ 0\ ,$$

pois $\omega(u,\ u) = \omega(v,\ v) = 0$, por hipótese, e pois $\omega(v,\ u) = \omega(u,\ v) = 0$ em função da condição de ω 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 211) tem as propriedades desejadas. A aplicação $E : V \to \tilde{V}$ 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

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

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. 3.17 *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

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

$u,\ v \in V$, 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 ω é não degenerada, segue do fato que se $\omega(u,\ v) = 0$ para todo u valeria, tomando $u = -iv$, $0 = \mathrm{Im}(\langle -iv,\ v\rangle) = \langle v,\ v\rangle$, o que implica $v = 0$.

Na Seção 3.6, página 304, 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

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

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

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 $S : V^2 \to \mathbb{C}$ é uma forma sesquilinear, então $Q : V \to \mathbb{C}$, definida por $Q(u) = S(u,\ u)$, $u \in V$, é dita ser forma quadrática sobre V. Segundo a identidade de polarização para formas sesquilineares (3.13), página 270, temos para todos $u,\ v \in V$,

$$S(u,\ v) \ = \ \frac{1}{4}\sum_{k=0}^{3} i^{-k} Q\big(u + i^k v\big)\ ,$$

o que similarmente indica que também em espaços vetoriais sobre os complexos toda forma quadrática procede de uma forma sesquilinear e Hermitiana e vice-versa. De acordo com essa definição, uma forma quadrática Q associada a uma forma sesquilinear ω somente será real se ω for Hermitiana, adicionalmente. Também de acordo com essa definição, vale $Q(\lambda u) = |\lambda|^2 Q(u)$, para todo $u \in V$ e $\lambda \in \mathbb{C}$.

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:

$$\omega(u,\ v) \ = \ \langle u,\ v\rangle_{\mathbb{C}} \ := \ \sum_{k=1}^{n} \overline{u_k} v_k \ , \tag{3.19}$$

onde $u = (u_1,\ \ldots, u_n)$ e $v = (v_1,\ \ldots, v_n)$. ♦

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 $u = (u_1,\ \ldots, u_n)$, $v = (v_1,\ \ldots, v_n)$ e onde A é uma matriz $n \times n$ *inversível.* ♦

Exemplo 3.3 Exemplo de uma forma sesquilinear Hermitiana que não é positiva. Seja $V = \mathbb{C}^n$ e seja ω 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 $A_{kl} = \overline{A_{lk}}$. A assim definida ω 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 $\omega(u,\ u) = \langle u,\ Au\rangle_{\mathbb{C}} = i(u_1\overline{u_2} - \overline{u_1}u_2) = -2\mathrm{Im}(u_1\overline{u_2})$, que pode ser negativo ou mesmo nulo. Assim, essa ω não é positiva. É fácil ver, porém, que essa ω é não degenerada (mostre isso!). ♦

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 <u>não</u> é *autoadjunta*, ou seja, $A_{kl} \neq \overline{A_{lk}}$ para pelo menos um elemento de matriz A_{kl}. A assim definida ω 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 ω não é Hermitiana. Fora isso, essa ω também não é positiva e é degenerada (mostre isso!). ♦

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

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

onde A é uma matriz $n \times n$ *<u>não</u>-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, ω 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. ♦

Na Seção 3.4, página 288, 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!). ♦

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!). ♦

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

$$\left| \sum_{i=1}^n \overline{u_i} v_i \right|^2 \ \leq \ \left(\sum_{j=1}^n |u_j|^2 \right) \left(\sum_{k=1}^n |v_k|^2 \right) \ . \tag{3.20}$$

♦

E. 3.18 *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. 3.19 *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 $V \to \mathbb{R}$ 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

$$\|a\| \ \geq \ \Big| \|a - b\| - \|b\| \Big| \ , \tag{3.21}$$

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

$$\|a - b\| \ \leq \ \|a\| + \|b\| \tag{3.22}$$

e que

$$\|b\| \ = \ \|a - (a - b)\| \ \leq \ \|a\| + \|a - b\| \ . \tag{3.23}$$

De (3.22) segue que

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

e de (3.23) 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

$$\|a\| \;\geq\; \Big| \|a-b\| - \|b\| \Big| ,$$

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 (3.21) as substituições $a \to a-b$, $b \to -b$, obtemos

$$\Big| \|a\| - \|b\| \Big| \;\leq\; \|a-b\| , \tag{3.24}$$

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

$$\|\alpha u + \beta v\| \;\leq\; |\alpha|\|u\| + |\beta|\|v\| \tag{3.25}$$

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

♣

• Normas

Uma *norma* é uma função $V \to \mathbb{R}$ 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

$$\|\alpha u + \beta v\| \;\leq\; |\alpha|\|u\| + |\beta|\|v\| \tag{3.26}$$

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

$$\Big| \|a\| - \|b\| \Big| \;\leq\; \|a-b\| , \tag{3.27}$$

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-las 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 (3.27)).
- 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 $\mathbb{C}^n = \{(z_1, \; \ldots, \; z_n), \text{ com } z_1, \; \ldots, \; z_n \in \mathbb{C}\}$, $n \geq 1$, o espaço vetorial das n-uplas de números complexos. Para $z = (z_1, \; \ldots, \; z_n) \in \mathbb{C}^n$, a expressão

$$\|z\|_1 \;:=\; \sum_{k=1}^{n} |z_k| \tag{3.28}$$

define uma norma em $\mathbb{C}^n$, denominada *norma* ℓ_1. Verifique! A expressão

$$\|z\|_\infty \;:=\; \max\big\{|z_1|, \; \ldots, \; |z_n|\big\} \tag{3.29}$$

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

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

$$\|z\|_p := \left[\sum_{k=1}^{n} |z_k|^p\right]^{\frac{1}{p}} \tag{3.30}$$

também define uma norma em $\mathbb{C}^n$, denominada *norma* ℓ_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 5.22, página 372. Vide também a Seção 25.5.1, página 1489 (especificamente, a expressão (25.50) do Teorema 25.4, página 1492).

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

$$\|f\|_1 := \int_a^b |f(x)| \, dx , \tag{3.31}$$

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

$$\|f\|_\infty := \sup_{x \in [a, \ b]} |f(x)| , \tag{3.32}$$

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

A norma (3.31) pode ser generalizada. Para cada $p \in \mathbb{R}$, $p \geq 1$, a expressão

$$\|f\|_p := \left[\int_a^b |f(x)|^p \, dx\right]^{\frac{1}{p}} , \tag{3.33}$$

$f \in C([a, \ b], \ \mathbb{C})$, 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 372, ou, com mais generalidade (para funções em espaços mensuráveis), vide a Seção 32.4.1, página 1721 (especificamente, vide a expressão (32.46) do Teorema 32.7, página 1721).

• 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. 3.20 *Exercício*. Mostre que a relação de equivalência entre normas é uma relação de equivalência. ✣

E. 3.21 *Exercício*. Mostre que as normas $\|\cdot\|_1$ e $\|\cdot\|_\infty$ do espaço $\mathbb{C}^n$, definidas em (3.28) e (3.29), 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 312:

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. 3.22.

E. 3.22 *Exercício*. As normas $\| \cdot \|_1$ e $\| \cdot \|_\infty$ do espaço $C([a,\ b],\ \mathbb{C})$, definidas em (3.31) e (3.32), 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 ω é um produto escalar em um espaço vetorial V existe associada a ω uma norma $\| \cdot \|_\omega$ dada por

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

$v \in V$.

E. 3.23 *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 21.2.3, página 1178. 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 $\mathrm{O(n)}$ 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 $R \in \mathrm{O(n)}$. O grupo $\mathrm{O(n)}$ e seus amigos são discutidos na Seção 21.2.3.1, página 1181 e seguintes.

• A desigualdade triangular

Talvez a principal consequência da desigualdade de Minkowski (3.18) seja a seguinte. Vamos supor que ω seja um produto escalar. Então, podemos definir[10] 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 ω 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 ω 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_ω é dita ser uma *pseudométrica*.

[10] As noções de métrica e de espaços métricos serão discutidas no Capítulo 25, página 1454.

• **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

$$\|u+v\|^2 + \|u-v\|^2 \ = \ 2\|u\|^2 + 2\|v\|^2 \ . \tag{3.34}$$

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. 3.24 *Exercício*. Por que (3.34) é chamada "identidade do paralelogramo"? ✠

E. 3.25 *Exercício*. Usando a identidade do paralelogramo demonstre a *identidade de Apolônio*[11]:

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

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

$$\langle u,\ v\rangle \ = \ \frac{1}{4}\sum_{n=0}^{3} i^{-n}\|u+i^n v\|^2 \ , \tag{3.35}$$

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

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. 3.26 *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

$$\langle u,\ Av\rangle \ = \ \frac{1}{4}\sum_{n=0}^{3} i^{-n}\Big\langle (u+i^n v),\ A(u+i^n v)\Big\rangle \ , \tag{3.37}$$

$$\langle u,\ Av\rangle \ = \ \frac{1}{4}\sum_{n=0}^{3} i^{n}\Big\langle (u+i^{-n} v),\ A(u+i^{-n} v)\Big\rangle \ , \tag{3.38}$$

[11] Apolônio de Perga (ci. 261 A.C. – ci. 190 A.C.).

E. 3.27 *Exercício.* Mostre isso. Sugestão: expanda o lado direito das igualdades acima e constate as igualdades. ✠

Tomando-se A como o operador identidade reobtém-se as identidades (3.35)-(3.36).

A relação (3.37) 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 (3.37) 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

$$\langle u, \ Av \rangle \ = \ \frac{1}{4} \sum_{n=0}^{3} i^{-n} \langle \psi_n, \ A\psi_n \rangle \left(2 + i^n \langle u, \ v \rangle + i^{-n} \langle v, \ u \rangle \right) , \tag{3.39}$$

onde

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

• Uma consequência da identidade de polarização

A relação (3.37) permite-nos facilmente provar a seguinte afirmação, frequentemente empregada:

Proposição 3.4 *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 3.5 *Seja $M \in \mathrm{Mat}(\mathbb{R}, \ n)$ 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

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

(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, $M = 0$. ■

A Proposição 3.5 pode ser generalizada para formas multilineares simétricas. Vide Proposição 3.1, página 264.

• 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 (3.34). 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[12], von Neumann[13] e Jordan[14][15] , teorema esse sugerido pela identidade de polarização, expressão (3.35), página 280.

[12] Maurice René Fréchet (1878–1973).
[13] János von Neumann (1903–1957). Von Neumann também adotou os nomes de Johann von Neumann e John von Neumann.
[14] Ernst Pascual Jordan (1902–1980).
[15] P. Jordan and J. von Neumann, "On Inner Product in Linear Metric Spaces", Ann. of Math. **36**, no. 2, 719-723 (1935).

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*

$$\|a+b\|^2 + \|a-b\|^2 \ = \ 2\|a\|^2 + 2\|b\|^2 \tag{3.40}$$

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

$$\omega(u,\ v) \ := \ \frac{1}{4}\sum_{n=0}^{3} i^{-n}\|u + i^n v\|^2 \ . \tag{3.41}$$

Então, ω 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 ω é a própria norma $\|\cdot\|$. Com isso, reconhecemos que (3.41) coincide com a identidade de polarização para o produto escalar ω.

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

A demonstração do Teorema de Fréchet, von Neumann e Jordan encontra-se no Apêndice 3.B, página 313. Vide também [604] ou [316] para outras demonstrações essencialmente idênticas.

A demonstração do Teorema 3.3 é engenhosa e sua principal dificuldade consiste em demonstrar que (3.41) é uma forma sesquilinear, um fato um tanto surpreendente se observarmos que o lado direito de (3.41) 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 α, $\beta \in \mathbb{C}$ 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. 3.28 e E. 3.29, oferecem exemplos de tais situações.

E. 3.28 *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. 3.29 *Exercício*. Seja o espaço vetorial $V = \mathbb{C}^n$, com $n \geq 2$. Para $a = (a_1,\ \ldots,\ a_n) \in \mathbb{C}^n$ a expressão $\|a\|_p := \left[|a_1|^p + \cdots + |a_n|^p\right]^{1/p}$, 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 $u = (1,\ 0,\ 0,\ \ldots, 0)$ e $v = (0,\ 1,\ 0,\ \ldots, 0)$. A norma $\|\cdot\|_p$ será discutida com mais detalhe no Capítulo 25, página 1454. ✠

• Generalizando a identidade do paralelogramo

O Exercício E. 3.29 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 372, que para z, $w \in \mathbb{C}$, valem as desigualdades $|z+w|^p + |z-w|^p \leq 2\big(|z|^p + |w|^p\big)$ para $0 < p < 2$ e $|z+w|^p + |z-w|^p \leq 2^{p-1}\big(|z|^p + |w|^p\big)$, para $p \geq 2$. É imediato por essas desigualdades que para todos u, $v \in \mathbb{C}^n$ valem

$$\big\|u+v\big\|_p^p + \big\|u-v\big\|_p^p \ \leq \ 2\Big(\big\|u\big\|_p^p + \big\|v\big\|_p^p\Big) \ , \qquad 1 \ \leq \ p \ < \ 2 \ , \tag{3.42}$$

e

$$\big\|u+v\big\|_p^p + \big\|u-v\big\|_p^p \ \leq \ 2^{p-1}\Big(\big\|u\big\|_p^p + \big\|v\big\|_p^p\Big) \ , \qquad p \ \geq \ 2 \ . \tag{3.43}$$

Note-se que no caso $p = 2$ (e somente nesse caso), (3.43) não é apenas uma desigualdade, mas sim uma igualdade, a identidade do paralelogramo.

E. 3.30 *Exercício*. Mostre isso! ✠

As desigualdades (3.42) e (3.43) substituem em certos casos a identidade do paralelogramo. Veremos isso quando discutirmos a propriedade de convexidade uniforme na Seção 25.6, página 1502.

• Normas assimétricas

Seja V um espaço vetorial real. Uma função $p : V \to [0,\ \infty)$ é dita ser uma *norma assimétrica* (também denominada *norma de Finsler*[16], 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 $v = 0$.

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 $\mathbb{R}$.

Uma outro exemplo, este mais importante, é o chamado *funcional de Minkowski*[17], 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 $p_C : V \to [0,\ \infty)$ por

$$p_C(v) \ = \ \inf\left\{\lambda \in [0,\ \infty) \middle|\ v \in \lambda C\right\} .$$

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 λ_1 e λ_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 $\lambda_1 + \lambda_2 \geq p_C(v_1 + v_2)$. Concluímos 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*, classe de variedades diferenciáveis onde o tensor métrico Riemanniano 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

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

[16] Paul Finsler 1894–1970).
[17] Hermann Minkowski (1864–1909).

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

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

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,

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

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 $\big\|\lambda x + (1 - \lambda)y - z\big\| = \big\|\lambda(x - z) + (1 - \lambda)(y - z)\big\| \leq \lambda\|x - z\| + (1 - \lambda)\|y - z\| \leq \lambda r + (1 - \lambda)r = r$, 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 $\|x\|_p := \Big(|x_1|^p + |x_2|^p\Big)^{1/p}$, com $p \geq 1$. Temos também a norma $\|x\|_\infty := \max\big\{|x_1|, \; |x_2|\big\}$. Na Figura 3.1, página 285, 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.

• **Uma desigualdade útil**

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

$$1 + \|x\| \; \leq \; \Big(1 + \|y\|\Big)\Big(1 + \|x - y\|\Big) \tag{3.44}$$

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

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

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

3.2.1 O Lema da Simetria

O elegante lema que segue encontra aplicações na Eletrostática, na Análise Harmônica, no estudo da chamada Transformação de Kelvin[18] e no estudo de aplicações conformes.

Lema 3.1 (Lema da Simetria) *Seja V um espaço vetorial (real ou complexo) e ω uma forma bilinear ou sesquilinear*[19]*. Suponhamos também que ω 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

$$\omega\Big(\eta(x, \; y), \; \eta(x, \; y)\Big) \; = \; \omega\Big(\eta(y, \; x), \; \eta(y, \; x)\Big) \; , \tag{3.45}$$

ou seja, de forma explícita,

$$\omega\left(\frac{y}{\sqrt{\omega(y, \; y)}} - \sqrt{\omega(y, \; y)}\, x, \; \frac{y}{\sqrt{\omega(y, \; y)}} - \sqrt{\omega(y, \; y)}\, x\right) \; = \; \omega\left(\frac{x}{\sqrt{\omega(x, \; x)}} - \sqrt{\omega(x, \; x)}\, y, \; \frac{x}{\sqrt{\omega(x, \; x)}} - \sqrt{\omega(x, \; x)}\, y\right) \; . \tag{3.46}$$

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

$$\left\|\frac{y}{\|y\|} - \|y\|x\right\| \; = \; \left\|\frac{x}{\|x\|} - \|x\|y\right\| \tag{3.47}$$

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

[18] William Thomson, 1st Baron Kelvin, (1824–1907).

[19] A demonstração mostra que basta que ω seja bilinear real.

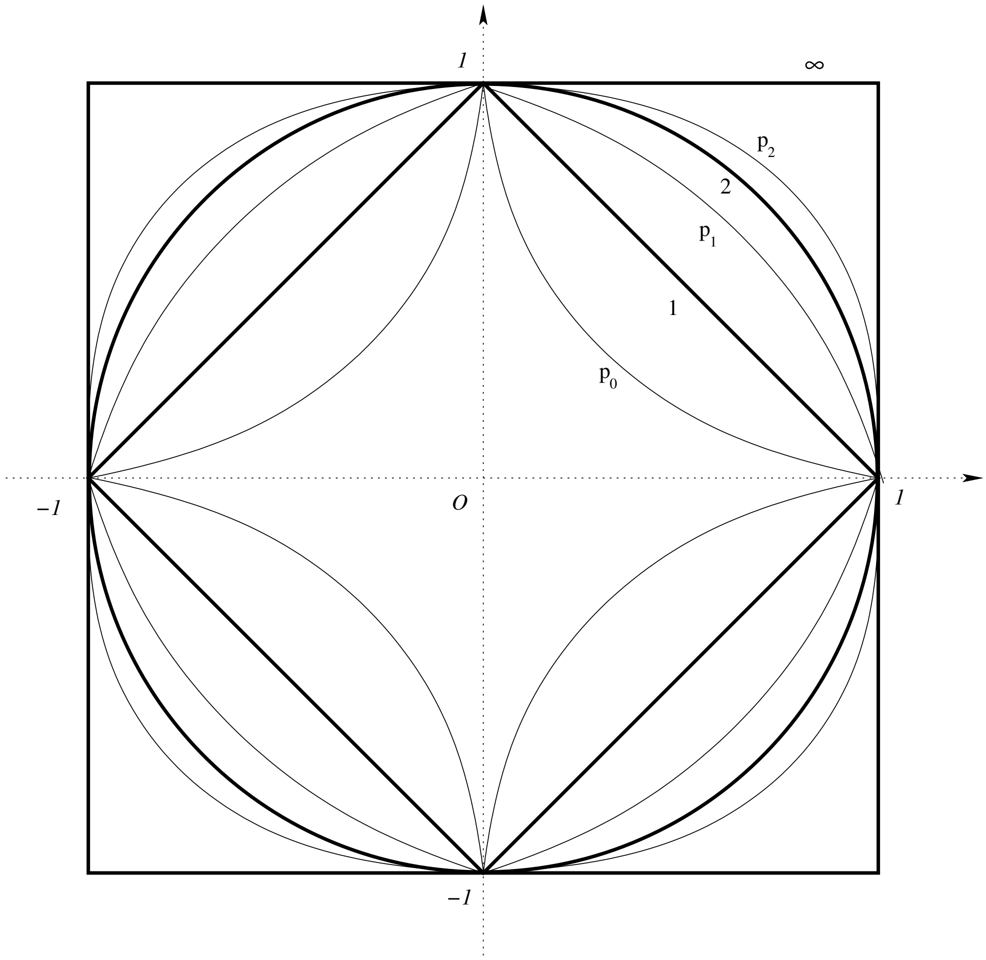

Figura 3.1: As bolas $\overline{\mathcal{B}_1(0)}$, de raio 1 centradas em 0, relacionadas às normas $\|\cdot\|_p$ e $\|\cdot\|_\infty$ em $\mathbb{R}^2$. As linhas sólidas indicam os bordos $\partial\mathcal{B}_1(0)$. O índice 1 indica a bola $\overline{\mathcal{B}_1(0)}$ para a norma $\|\cdot\|_1$. Trata-se de um quadrado oblíquo 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 ∞ 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.

Prova. Para demonstrar (3.45), basta expandir o lado esquerdo de (3.46):

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

A última expressão é manifestamente invariante pela troca $x \leftrightarrow y$. Logo, $\omega\big(\eta(x,\, y),\, \eta(x,\, y)\big) = \omega\big(\eta(y,\, x),\, \eta(y,\, x)\big)$, como desejávamos. A relação (3.47) 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 $\mathbf{e} \in V$ é dito ser um *vetor unitário*, ou um *vetor normalizado*, em relação ao produto escalar em questão (e à norma a este associada) se $\|\mathbf{e}\| = 1$.

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 $\frac{\lambda}{\|u\|}u$ onde λ 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 $\mathfrak{R}(u) := \left\{ \frac{\lambda}{\|u\|}u, \ \lambda \in \mathbb{C} \text{ com } |\lambda| = 1 \right\}$ é dito ser o *raio associado ao vetor u*.

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

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 1.1.2.7, página 73.

• Conjunto ortonormal de vetores

Seja $E = \{\mathbf{e}_\alpha, \ \alpha \in \Lambda\}$ um conjunto não vazio de vetores distintos de V, Λ 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 $\|\mathbf{e}_\alpha\| = 1$ e se para todos $\alpha, \beta \in \Lambda$ com $\alpha \neq \beta$ valer $\langle \mathbf{e}_\alpha, \mathbf{e}_\beta\rangle = 0$. 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. 3.31 *Exercício*. Se $E = \{\mathbf{e}_\alpha, \ \alpha \in \Lambda\}$ é um conjunto ortonormal de vetores, mostre que $\|\mathbf{e}_\alpha - \mathbf{e}_\beta\| = \sqrt{2}$ sempre que $\alpha \neq \beta$. ✠

• Procedimento de ortogonalização de Gram-Schmidt

Dado um conjunto finito $B = \{b_1, \ldots, b_n\}$ composto por n vetores não nulos e linearmente independentes de V, podemos construir um conjunto ortonormal $E = \{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ no subespaço n-dimensional gerado pelos vetores de B por um procedimento conhecido como *procedimento de ortogonalização de Gram-Schmidt*[20][21], que passaremos a descrever[22].

[20] Jørgen Pedersen Gram (1850–1916).
[21] Erhard Schmidt (1876–1959).
[22] Seria mais adequado chamar o procedimento de *procedimento de ortonormalização de Gram-Schmidt*, pois o conjunto de vetores resultante é ortonormal, mas aquela nomenclatura é adotada amplamente.

O procedimento começa escolhendo-se um vetor de B e normalizando-se esse vetor, definindo assim o primeiro vetor $\mathbf{e}_1$ de E. Para simplificar a notação, escolhemos começar com o vetor b_1 e definimos, assim, $\mathbf{e}_1 := \frac{1}{\|b_1\|} b_1$. No segundo passo, tomamos o vetor b_2, subtraímos do mesmo sua componente na direção $\mathbf{e}_1$ e em seguida normalizamos o vetor disso resultante, definindo assim o vetor $\mathbf{e}_2$:

$$\mathbf{e}_2 \ := \ \frac{1}{\|b_2 - \langle \mathbf{e}_1, \ b_2 \rangle \, \mathbf{e}_1\|} \Big(b_2 - \langle \mathbf{e}_1, \ b_2 \rangle \, \mathbf{e}_1 \Big) \ .$$

Pela construção, é evidente que $\|\mathbf{e}_1\| = \|\mathbf{e}_2\| = 1$ e é fácil verificar que $\langle \mathbf{e}_1, \ \mathbf{e}_2 \rangle = 0$.

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 $\mathbf{e}_1, \ \ldots, \ \mathbf{e}_{k-1}$ e normalizando o vetor assim resultante. Obtemos

$$\mathbf{e}_1 \ := \ \frac{b_1}{\|b_1\|} \ , \qquad \mathbf{e}_k \ := \ \frac{1}{\left\| b_k - \sum_{l=1}^{k-1} \langle \mathbf{e}_l, \ b_k \rangle \, \mathbf{e}_l \right\|} \left(b_k - \sum_{l=1}^{k-1} \langle \mathbf{e}_l, \ b_k \rangle \, \mathbf{e}_l \right) \ , \qquad k \ = \ 2, \ \ldots, \ n \ .$$

Observe-se que cada $\mathbf{e}_k$ depende apenas dos vetores $\mathbf{e}_1, \ \ldots, \ \mathbf{e}_{k-1}$, definidos nas $k-1$ etapas anteriores. Como é fácil verificar, valem as relações $\langle \mathbf{e}_i, \ \mathbf{e}_j \rangle = \delta_{i\,j}$ para todos $i, \ j = 1, \ldots, n$, atestando que $E = \big\{ \mathbf{e}_1, \ \ldots, \ \mathbf{e}_n \big\}$ é um conjunto ortonormal de vetores.

Note que a construção acima descrita não é única, pois podemos reordenar os elementos de B, obtendo assim uma nova sequência de vetores $\mathbf{e}_j$, $j = 1, \ \ldots, \ n$.

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 Legendre, os de Tchebychev etc. Vide Capítulo 16, página 927, e referências lá citadas.

3.3.1 Conjuntos Ortogonais para Formas Bilineares Simétricas Não Degeneradas

O procedimento de Gram-Schmidt descrito acima possui uma extensão parcial para formas bilineares simétricas e não degeneradas em espaços vetoriais reais. Trata-se da Proposição 3.6, abaixo. Faremos uso desse resultado no Capítulo 24, página 1397, especificamente na demonstração do Teorema de Liouville sobre aplicações conformes em $\mathbb{R}^n$, Teorema 24.4, página 1422.

O procedimento de Gram-Schmidt, descrito anteriormente, é uma implementação concreta da afirmação da Proposição 3.6 no caso de ω ser um produto escalar.

Na demonstração da Proposição 3.6 empregaremos o seguinte lema, cuja afirmação é de interesse por si só:

Lema 3.2 *Se V é um espaço vetorial real e ω é uma forma bilinear simétrica e não degenerada em V, então existe $w \in V$ tal que $\omega(w, \ w) \neq 0$.* □

Prova. Por ω ser bilinear e simétrica, temos pela identidade de polarização (3.4), página 264, que para quaisquer $u, \ v \in V$ vale $\omega(u, \ v) = \frac{1}{4}\big[\omega(u+v, \ u+v) - \omega(u-v, \ u-v)\big]$. Fixado um $u \in V$ com $u \neq 0$, sabemos que o fato de ω ser não degenerado significa que se $\omega(u, \ v) = 0$, então $v = 0$. Logo, como a forma bilinear ω é suposta ser não degenerada, existe ao menos um $v \in V$, não nulo, tal que $\omega(u, \ v) \neq 0$. Portanto, pela identidade de polarização, ou $\omega\big(u+v, \ u+v\big)$ é não nulo ou $\omega\big(u-v, \ u-v\big)$ é não nulo (ou ambos). Assim, a afirmação é válida para $w = u+v$ ou para $w = u-v$ (ou ambos). ■

Seja V um espaço vetorial real e seja ω uma forma bilinear simétrica e não degenerada em V. Dois vetores u e v são ditos ω-ortogonais se $\omega(u, \ v) = 0$. A proposição a seguir afirma que se V tem dimensão finita é sempre possível encontrar um conjunto de $n \equiv \dim(V)$ vetores não nulos e linearmente independentes em V que são mutuamente ω-ortogonais.

Para a formulação dessa proposição é conveniente introduzirmos a nomenclatura e a notação apropriadas. Para $k \in \{1, \ldots, n\}$ seja $\{v_1, \ldots, v_k\} \subset V$. Denotamos por $[v_1, \ldots, v_k]$ o subespaço gerado por $v_1, \ldots, v_k$ (ou seja, o conjunto de todas as combinações lineares de $v_1, \ldots, v_k$) e denotamos por $[v_1, \ldots, v_k]^{\perp_\omega}$ o complemento ω-ortogonal de $[v_1, \ldots, v_k]$, ou seja

$$\big[v_1, \ldots, v_k\big]^{\perp_\omega} := \big\{w \in V | \, \omega(w, v_j) = 0 \text{ para todo } j = 1, \ldots, k\big\} .$$

Em palavras, $\big[v_1, \ldots, v_k\big]^{\perp_\omega}$ é o conjunto de todos os vetores $w \in V$ que são ω-ortogonais a todos os vetores de $[v_1, \ldots, v_k]$.

Proposição 3.6 *Seja V um espaço vetorial real de dimensão $n \in \mathbb{N}$ e seja ω uma forma bilinear simétrica e não degenerada em V. Então, existem n vetores linearmente independentes $w_1, \ldots, w_n$ tais que $\omega(w_i, w_j) = 0$ para todos $i \neq j$ e tais que $\omega(w_i, w_i) \neq 0$ para cada $i = 1, \ldots, n$.*

Um ponto a ser notado, e que ficará claro na demonstração, é que dos dois primeiros vetores w_1 e w_2 exige-se apenas que $\omega(w_1, w_1) \neq 0$, que $\omega(w_1, w_2) \neq 0$ e que $\omega(w_1, w_2) = 0$. □

Prova. Seja $V^{(n)} \equiv V$. De acordo com o Lema 3.2, página 287, existe $w_1 \in V^{(n)}$ com $\omega(w_1, w_1) \neq 0$. Seja agora $V^{(n-1)} := [w_1]^{\perp_\omega}$. O espaço $V^{(n-1)}$ é claramente um espaço vetorial real de dimensão $n-1$. Afirmamos que ω restrita a $V^{(n-1)}$ é igualmente uma forma bilinear simétrica e não degenerada. A bilinearidade e a simetria são afirmações evidentes e suponha-se que exista $v \in V$, com $v \neq 0$, tal que $\omega(u, v) = 0$ para todo $u \in V^{(n-1)}$. Como $\omega(w_1, v) = 0$ (pois $v \in [w_1]^{\perp_\omega}$), concluiríamos também que $\omega(u, v) = 0$ para todo $u \in V^{(n)} \equiv V$, o que contraria a suposta não degenerescência de ω em V.

Assim, novamente evocando o Lema 3.2, podemos afirmar que existe $w_2 \in [w_1]^{\perp_\omega}$ com $\omega(w_2, w_2) \neq 0$. Como $w_2 \in [w_1]^{\perp_\omega}$, vale também $\omega(w_1, w_2) = 0$.

Podemos agora proceder por indução, supondo existirem, para algum $k \in \{1, \ldots, n-1\}$, vetores $w_1, \ldots, w_k$ com $\omega(w_i, w_j) = 0$ para todos $i, j \in \{1, \ldots, k\}$ com $i \neq j$ e tais que $\omega(w_i, w_i) \neq 0$ para cada $i = 1, \ldots, k$. O subespaço $[w_1, \ldots w_k]$ é k dimensional e seja $V^{(n-k)} := \big[w_1, \ldots w_k\big]^{\perp_\omega}$, que é $(n-k)$-dimensional. Afirmamos que a forma ω restrita a $V^{(n-k)}$ é bilinear, simétrica e não degenerada. Como antes, a bilinearidade e a simetria são afirmações evidentes e suponha-se que exista $v \in V^{(n-k)}$, com $v \neq 0$, tal que $\omega(u, v) = 0$ para todo $u \in V^{(n-k)}$. Como $\omega(w_j, v) = 0$ para todo $j = 1, \ldots, k$ (pois $v \in [w_1, \ldots w_k]^{\perp_\omega}$), concluiríamos também que $\omega(u, v) = 0$ para todo $u \in V^{(n)} \equiv V$, o que contraria a suposta não degenerescência de ω em V.

Assim, novamente evocando o Lema 3.2, podemos afirmar que existe $w_{k+1} \in \big[w_1, \ldots w_k\big]^{\perp_\omega}$ com $\omega(w_{k+1}, w_{k+1}) \neq 0$ com $\omega(w_{k+1}, w_j) = 0$ para todo $j = 1, \ldots, k$. O Princípio de Indução Finita completa a demonstração. ■

3.4 Formas Bilineares, 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.

• **Continuidade**

Vamos provar a seguinte afirmação: toda forma linear ou bilinear em $\mathbb{R}^n$ é contínua (em ambas as variáveis), o mesmo valendo para formas lineares, 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 ω uma forma sesquilinear em $\mathbb{C}^n$. Para vetores $x, y \in \mathbb{C}^n$, $y \neq 0$, escrevemos

$$\omega(x, y) = \|y\| \, \omega\big(x, y/\|y\|\big) , \tag{3.48}$$

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

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

e, portanto,

$$\left|\omega(x,\ v)\right| \ \leq\ \left|\omega(x,\ b_1)\right| + \cdots + \left|\omega(x,\ b_n)\right| .$$

Para cada x fixo o lado direito é uma constante K_x e não depende de v. Aplicando isso a (3.48), teremos

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

Isso mostra que

$$\lim_{y\to 0} \left|\omega(x,\ y)\right| \ =\ 0$$

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

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

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

• O Teorema da Representação de Riesz

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

Teorema 3.4 (Teorema da Representação de Riesz em dimensão finita) *Seja l um funcional linear (portanto, contínuo) em $\mathbb{C}^n$ (dotado do produto escalar usual $\langle\cdot,\ \cdot\rangle_{\mathbb{C}}$). Então, existe $\phi \in \mathbb{C}^n$, único, tal que*

$$l(x) \ =\ \langle\phi,\ x\rangle_{\mathbb{C}}, \qquad \forall x \in \mathbb{C}^n .$$

□

Comentário. Esse importante teorema pode ser generalizado a espaços de Hilbert, tal como discutido na Seção 41.2.2.1, página 2347. Vide Teorema 41.4, página 2348. Para benefício do leitor apresentamos a demonstração desse caso particular de dimensão finita, que pouco se distingue da demonstração geral. ♣

Prova do Teorema 3.4. Seja l um funcional linear contínuo em $\mathbb{C}^n$. Seja $N \subset \mathbb{C}^n$ o núcleo de l, ou seja, o conjunto de todos os vetores de $\mathbb{C}^n$ que são anulados por l:

$$N \ :=\ \Big\{y \in \mathbb{C}^n \big|\ l(y) \ =\ 0\Big\} .$$

Vamos mostrar que N é um subespaço linear fechado de $\mathbb{C}^n$. Que N é um subespaço é elementar pois, se $x,\ y \in N$, então $l(\alpha x + \beta y) = \alpha l(x) + \beta l(y) = \alpha 0 + \beta 0 = 0$. Que N é fechado pode ser visto pelo fato que podemos caracterizar N como a imagem inversa do número 0 de $\mathbb{C}$ por l: $N = l^{-1}(\{0\})$. O conjunto $\{0\}$, constituído por um único ponto, é fechado em $\mathbb{C}$ e funções contínuas são tais que sua imagem inversa mapeia fechados em fechados (vide página 1677). Logo, N é fechado.

Caso N seja idêntico a $\mathbb{C}^n$, isso significa que $l(x) = 0$ para todo $x \in \mathbb{C}^n$ e o teorema estaria provado, adotando-se para tal $\phi = 0$.

Vamos supor que $N \neq \mathbb{C}^n$. Como N é fechado, todo $x \in \mathbb{C}^n$ é da forma $x = y + z$ com $y \in N$ e $z \in N^\perp$. Como $N \neq \mathbb{C}^n$, devem existir elementos não nulos em $N^\perp$, doutra forma teríamos $x = y \in N$ para todo $x \in \mathbb{C}^n$.

Seja, então, z_0 um vetor não-nulo de $N^\perp$. É obvio que $l(z_0) \neq 0$. Para qualquer vetor $u \in \mathbb{C}^n$ vale que $l(z_0)u - l(u)z_0$ é um elemento de N, pois

$$l\Big(l(z_0)u - l(u)z_0\Big) \ =\ l(z_0)l(u) - l(u)l(z_0) \ =\ 0 .$$

Assim, como $l(z_0)u - l(u)z_0$ é um elemento de N e z_0 é um elemento de $N^\perp$, ambos são ortogonais entre si, ou seja, $0 = \big\langle z_0,\ l(z_0)u - l(u)z_0\big\rangle_{\mathbb{C}}$. Isso diz, porém, que $0 = l(z_0)\langle z_0,\ u\rangle_{\mathbb{C}} - l(u)\|z_0\|^2$, ou seja,

$$l(u) \ =\ \frac{l(z_0)}{\|z_0\|^2} \langle z_0,\ u\rangle_{\mathbb{C}} \ =\ \left\langle \frac{\overline{l(z_0)}}{\|z_0\|^2} z_0,\ u \right\rangle_{\mathbb{C}} .$$

Definindo

$$\phi = \frac{\overline{l(z_0)}}{\|z_0\|^2} z_0 ,$$

fica provado que para todo $u \in \mathbb{C}^n$ vale $l(u) = \langle\phi, u\rangle_{\mathbb{C}}$, como desejávamos estabelecer.

Por fim, para demonstrar que tal ϕ é único, suponhamos que exista um outro ϕ' tal que também valha $l(u) = \langle\phi', u\rangle_{\mathbb{C}}$ para todo $u \in \mathbb{C}^n$. Teríamos, então, $\langle\phi, u\rangle_{\mathbb{C}} = \langle\phi', u\rangle_{\mathbb{C}}$, ou seja, $\langle\phi - \phi', u\rangle_{\mathbb{C}} = 0$ para todo $u \in \mathbb{C}^n$. Como essa relação vale para todo $u \in \mathbb{C}^n$, vale também para $u = \phi - \phi'$. Logo, $0 = \langle\phi - \phi', \phi - \phi'\rangle_{\mathbb{C}} = \|\phi - \phi'\|^2$ e, portanto, $\phi = \phi'$. ■

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

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

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

é um funcional linear e contínuo. Pelo Teorema da Representação de Riesz, Teorema 3.4, página 289, 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 η_x com a propriedade acima: $A(x) = \eta_x$. Tem-se,

$$\omega(x, y) = \langle A(x), y\rangle_{\mathbb{C}} . \tag{3.49}$$

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 α_1 e α_2 e todos os vetores x_1 e x_2. De fato, por (3.49),

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

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

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

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 ω em $\mathbb{C}^n$ existe uma matriz $n \times n$ complexa A_ω 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 $\mathbb{C}^n$.

• **Forma geral de formas bilineares em $\mathbb{R}^n$**

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

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

é um funcional linear e contínuo. Pelo Teorema da Representação de Riesz, Teorema 3.4, página 289, 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 η_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 ω em $\mathbb{R}^n$ existe uma matriz $n \times n$ real A_ω 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 $\mathbb{R}^n$.

• Dois resultados especiais sobre formas bilineares

O resultado a seguir, e o corolário que se lhe segue, é útil e será evocado nestas Notas, particularmente na demonstração do Teorema de Liouville sobre aplicações conformes nos espaços $\mathbb{R}^n$, Teorema 24.4, página 1422.

Proposição 3.7 *Seja ω uma forma bilinear simétrica em $\mathbb{R}^n$ com a propriedade que $\omega(u, \; v) = 0$ sempre que $\langle u, \; v \rangle_{\mathbb{R}} = 0$. Então, existe uma constante γ tal que*

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

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

Prova. Tomemos $w \in \mathbb{R}^n$ com $w \neq 0$, fixo, e defina-se, $\rho_w : \mathbb{R}^n \to \mathbb{R}$ por

$$\rho_w(x) \; = \; \omega(w, \; x) - \frac{\omega(w, \; w)}{\langle w, \; w \rangle_{\mathbb{R}}} \langle w, \; x \rangle_{\mathbb{R}} \; , \qquad x \in \mathbb{R}^n \; .$$

Claramente ρ_w é linear. Pela definição, $\rho_w(w) = 0$ e para $x \in [w]^\perp$ (o complemento ortogonal de w segundo o produto escalar $\langle \cdot, \; \cdot \rangle_{\mathbb{R}}$) temos também $\rho_w(x) = 0$, devido à hipótese de $\omega(u, \; v)$ ser nulo sempre que $\langle u, \; v \rangle_{\mathbb{R}} = 0$. Assim, ρ_w é identicamente nulo e, portanto, para todo $x \in \mathbb{R}^n$,

$$\omega(w, \; x) \; = \; \frac{\omega(w, \; w)}{\langle w, \; w \rangle_{\mathbb{R}}} \langle w, \; x \rangle_{\mathbb{R}} \; . \tag{3.50}$$

Se considerarmos agora um outro vetor $z \in \mathbb{R}^n$ com $z \neq 0$ e definirmos

$$\rho_z(x) \; = \; \omega(z, \; x) - \frac{\omega(z, \; z)}{\langle z, \; z \rangle_{\mathbb{R}}} \langle z, \; x \rangle_{\mathbb{R}} \; , \qquad x \in \mathbb{R}^n \; ,$$

concluiremos analogamente que ρ_z é identicamente nulo e, portanto, para todo $x \in \mathbb{R}^n$,

$$\omega(z, \; x) \; = \; \frac{\omega(z, \; z)}{\langle z, \; z \rangle_{\mathbb{R}}} \langle z, \; x \rangle_{\mathbb{R}} \; . \tag{3.51}$$

Tomando $x = z$ em (3.50) e $x = w$ (3.51) e usando a hipótese de simetria de ω (que afirma que $\omega(w, \; z) = \omega(z, \; w)$), obtemos

$$\frac{\omega(w, \; w)}{\langle w, \; w \rangle_{\mathbb{R}}} \langle w, \; z \rangle_{\mathbb{R}} \; = \; \frac{\omega(z, \; z)}{\langle z, \; z \rangle_{\mathbb{R}}} \langle z, \; w \rangle_{\mathbb{R}} \; .$$

Como $\langle w, \; z \rangle_{\mathbb{R}} = \langle z, \; w \rangle_{\mathbb{R}}$, temos que

$$\left(\frac{\omega(w, \; w)}{\langle w, \; w \rangle_{\mathbb{R}}} - \frac{\omega(z, \; z)}{\langle z, \; z \rangle_{\mathbb{R}}} \right) \langle z, \; w \rangle_{\mathbb{R}} \; = \; 0 \; . \tag{3.52}$$

Como essa relação vale para todos w e z não nulos, concluímos que $\frac{\omega(w,w)}{\langle w,\, w\rangle_{\mathbb{R}}} = \frac{\omega(z,z)}{\langle z,\, z\rangle_{\mathbb{R}}}$ também para quaisquer w e z não nulos, desde que z e w não sejam ortogonais. Assim, fixando $z \neq 0$, a igualdade $\frac{\omega(w,w)}{\langle w,\, w\rangle_{\mathbb{R}}} = \frac{\omega(z,z)}{\langle z,\, z\rangle_{\mathbb{R}}}$ vale para todo w no conjunto complementar de $[z]^\perp$, que é um conjunto aberto de $\mathbb{R}^n$ e é constante nesse conjunto. Por continuidade, deve valer também em $[z]^\perp$.

Comentário. Para entender esse argumento de continuidade, tomemos w da forma $w = w_0 + \epsilon z$ com $w_0 \in [z]^\perp$, arbitrário, e $\epsilon \neq 0$. Teremos $\langle z,\, w\rangle_{\mathbb{R}} = \epsilon \|z\|^2 \neq 0$ e, portanto, por (3.52), vale a igualdade $\frac{\omega(w,w)}{\langle w,\, w\rangle_{\mathbb{R}}} = \frac{\omega(z,z)}{\langle z,\, z\rangle_{\mathbb{R}}}$, que significa que

$$\frac{\omega(w_0,\, w_0) + \epsilon^2 \omega(z,\, z)}{\|w_0\|^2 + \epsilon^2 \|z\|^2} = \frac{\omega(z,\, z)}{\langle z,\, z\rangle_{\mathbb{R}}}$$

para todo $\epsilon > 0$, pois $\omega(w_0,\, z) = 0$ já que $\langle w_0,\, z\rangle_{\mathbb{R}} = 0$. Podemos agora passar ao limite $\epsilon \to 0$, que implica $\frac{\omega(w_0,\, w_0)}{\langle w_0,\, w_0\rangle_{\mathbb{R}}} = \frac{\omega(z,z)}{\langle z,\, z\rangle_{\mathbb{R}}}$ para quaisquer w_0 e $z \neq 0$ com $w_0 \in [z]^\perp$. ♣

Retomando, concluímos que a razão $\frac{\omega(w,w)}{\langle w,\, w\rangle_{\mathbb{R}}}$ é constante entre os vetores não nulos e, denotando essa constante por γ, a relação (3.50) diz-nos que

$$\omega(w,\, x) = \gamma \langle w,\, x\rangle_{\mathbb{R}}$$

para todo $x \in \mathbb{R}^n$ e todo $w \in \mathbb{R}^n$ com $w \neq 0$. A igualdade acima, porém, é também válida se $w = 0$ e, portanto, ela vale sempre. ■

Corolário 3.1 *No espaço vetorial $\mathbb{R}^{n+p}$, com $n,\, p \in \mathbb{N}_0$, $n + p \geq 1$, seja a forma bilinear simétrica $\omega_N : \mathbb{R}^{n+p} \to \mathbb{R}$ definida por $\omega_N(u,\, v) := \langle u,\, Nv\rangle_{\mathbb{R}}$, onde $N := \operatorname{diag}\big(\underbrace{1,\, \ldots,\, 1}_{n \text{ vezes}},\, \underbrace{-1,\, \ldots,\, -1}_{p \text{ vezes}}\big)$. Ou seja, para $u = \big(u^1,\, \ldots,\, u^{n+p}\big)$ e $v = \big(v^1,\, \ldots,\, v^{n+p}\big)$, temos*

$$\omega_N(u,\, v) = u^1 v^1 + \cdots + u^n v^n - u^{n+1} v^{n+1} - \cdots - u^{n+p} v^{n+p} . \tag{3.53}$$

Seja $\varphi : \mathbb{R}^{n+p} \to \mathbb{R}$ for outra forma bilinear simétrica com a propriedade que $\varphi(u,\, v) = 0$ sempre que

$$\omega_N(u,\, v) = 0 \quad \text{com} \quad \omega_N(u,\, u) \neq 0 \ \text{ e } \ \omega_N(v,\, v) \neq 0 .$$

Então, existe uma constante γ tal que

$$\varphi(u,\, v) = \gamma \omega_N(u,\, v)$$

para todos $u,\, v \in \mathbb{R}^{n+p}$. □

Prova. Se $n = 0$ ou $p = 0$ não há o que se provar, pois nesse caso a afirmação segue diretamente da Proposição 3.7. Tomemos, assim, $n \geq 1$ e $p \geq 1$.

Seja $\{\mathbf{e}_1,\, \ldots,\, \mathbf{e}_{n+p}\}$ a base ortogonal canônica de $\mathbb{R}^{n+p}$, de sorte que todo $x \in \mathbb{R}^{n+p}$ se escreve como $x = \sum_{k=1}^{n+p} x^k \mathbf{e}_k$ e, pela definição de ω_N,

$$\omega_N\big(\mathbf{e}_i,\, \mathbf{e}_j\big) = \begin{cases} \delta_{ij} , & \text{para } i,\, j \in \{1,\, \ldots,\, n\} , \\ -\delta_{ij} , & \text{para } i,\, j \in \{n+1,\, \ldots,\, n+p\} , \\ 0 , & \text{de outra forma} . \end{cases} \tag{3.54}$$

Temos $\mathbb{R}^{n+p} = V_1 \oplus V_2$, onde $V_1 \simeq \mathbb{R}^n$ é o subespaço gerado por $\mathbf{e}_1,\, \ldots,\, \mathbf{e}_n$ e onde $V_2 \simeq \mathbb{R}^p$ é o subespaço gerado por $\mathbf{e}_{n+1},\, \ldots,\, \mathbf{e}_{n+p}$.

Consideremos a restrição $\varphi_1 \equiv \varphi \restriction_{V_1 \oplus V_2}$ de φ a $V_1 \oplus V_1$. Naturalmente, aplica-se também a φ_1 a premissa: $\varphi_1(u,\, v) = 0$ sempre que $\omega_N(u,\, v) = 0$ para $u,\, v \in V_1$. Como $\omega_N(u,\, v) = \langle u,\, v\rangle_{\mathbb{R}}$ para $u,\, v \in V_1$, temos pela Proposição 3.7 que existe constante γ_1 tal que $\varphi_1(u,\, v) = \gamma_1 \langle u,\, v\rangle_{\mathbb{R}}$ para todos $u,\, v \in V_1$.

De forma análoga concluímos também que existe constante γ_2 tal que $\varphi_2(u,\, v) = \gamma_2 \langle u,\, v\rangle_{\mathbb{R}}$ para todos $u,\, v \in V_2$, onde $\varphi_2 \equiv \varphi \restriction_{V_2 \oplus V_2}$ é a restrição de φ a $V_2 \oplus V_2$.

Pelas hipóteses, temos também que $\varphi(u,\, v) = 0$ para $u \in V_1$ e $v \in V_2$, pois nesse caso $\langle u,\, v\rangle_{\mathbb{R}} = 0$, já que V_1 e V_2 são ortogonais em relação ao produto escalar $\langle \cdot,\, \cdot\rangle_{\mathbb{R}}$.

Tomemos agora dois elementos gerais $u,\ v \in \mathbb{R}^{n+p}$ decompostos da forma $u = u_1 + u_2$ e $v = v_1 + v_2$, com $u_k,\ v_k \in V_k$, $k = 1,\ 2$. Teremos, do exposto acima,

$$\varphi(u,\ v) \ = \ \gamma_1 \langle u_1,\ v_1 \rangle_{\mathbb{R}} + \gamma_2 \langle u_2,\ v_2 \rangle_{\mathbb{R}} \tag{3.55}$$

e

$$\omega_N(u,\ v) \ = \ \langle u_1,\ v_1 \rangle_{\mathbb{R}} - \langle u_2,\ v_2 \rangle_{\mathbb{R}} \ . \tag{3.56}$$

Para obter disso uma relação entre γ_1 e γ_2, tomemos o caso particular em que

$$u \ = \ 2\mathbf{e}_1 + \mathbf{e}_{n+1} \qquad \text{e} \qquad v \ = \ \mathbf{e}_1 + 2\mathbf{e}_{n+1} \ .$$

Temos, $u_1 = 2\mathbf{e}_1 \in V_1$, $u_2 = \mathbf{e}_{n_1} \in V_2$, $v_1 = \mathbf{e}_1 \in V_1$, $v_2 = 2\mathbf{e}_{n+1} \in V_2$ e assim,

$$\langle u_1,\ v_1 \rangle_{\mathbb{R}} \ = \ 2 \qquad \text{e} \qquad \langle u_2,\ v_2 \rangle_{\mathbb{R}} \ = \ 2 \ .$$

Disso obtemos,

$$\omega_N(u,\ v) \ = \ 0 \ , \quad \text{sendo que} \quad \omega_N(u,\ u) \ = \ 3 \ , \quad \text{e} \quad \omega_N(v,\ v) \ = \ -3 \ . \tag{3.57}$$

Dessa forma a condição (3.53) é satisfeita e, pela hipótese sobre φ, temos que $\varphi(u,\ v) = 0$, ou seja,

$$0 \ = \ \varphi(u,\ v) \ = \ \gamma_1 \langle u_1,\ v_1 \rangle_{\mathbb{R}} + \gamma_2 \langle u_2,\ v_2 \rangle_{\mathbb{R}} \ = \ 2(\gamma_1 + \gamma_2) \ .$$

Assim, concluímos que para as constantes γ_1 e γ_2 vale $\gamma_1 = -\gamma_2$ e, portanto, para quaisquer $u,\ v \in \mathbb{R}^{n+p}$ temos

$$\varphi(u,\ v) \ \stackrel{(3.55)}{=} \ \gamma_1 \Big(\langle u_1,\ v_1 \rangle_{\mathbb{R}} - \langle u_2,\ v_2 \rangle_{\mathbb{R}} \Big) \ \stackrel{(3.56)}{=} \ \gamma_1 \omega_N(u,\ v) \ ,$$

completando a prova. ■

• Formas bilineares em $\mathbb{C}^n$

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

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

define uma forma sesquilinear em $\mathbb{C}^n$, onde $\overline{x} = (\overline{x_1},\ \ldots,\ \overline{x_n})$ para $x = (x_1,\ \ldots,\ x_n) \in \mathbb{C}^n$. Pelo que provamos acima, portanto, existe uma matriz complexa A_ω tal que

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

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

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

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_ω.

Resumimos esses fatos no seguinte teorema:

Teorema 3.7 *Para toda forma bilinear ω em $\mathbb{C}^n$ existe uma matriz $n \times n$ complexa A_ω 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 ω é uma forma bilinear alternante em $\mathbb{R}^n$ ou $\mathbb{C}^n$, ou seja, ω é bilinear e $\omega(x,\ y) = -\omega(y,\ x)$, então ω é da forma $\omega(x,\ y) = \langle A\,x,\ y\rangle_{\mathbb{R}}$ 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

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

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 ω é uma forma simplética, ou seja, ω 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 ω 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 ω é um produto escalar em $\mathbb{C}^n$, ω 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

$$\langle Ax,\ x\rangle_{\mathbb{C}} > 0 \tag{3.58}$$

se $x \neq 0$. Uma consequência disso é o seguinte: se v_i é um dos autovetores de A com autovalor λ_i, então $\lambda_i > 0$. De fato, tomando $x = v_i$ em (3.58), teremos[23] $0 < \langle Av_i,\ v_i\rangle_{\mathbb{C}} = \lambda_i \langle v_i,\ v_i\rangle_{\mathbb{C}}$, 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 618). 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[24].

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. 3.32 *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 3.8 *Se ω é um produto escalar em $\mathbb{C}^n$, então existe uma única matriz $M_\omega \in \mathrm{Mat}\,(\mathbb{C},\ n)$ autoadjunta e de autovalores positivos (e, portanto, inversível) tal que $\omega(x,\ y) = \langle x,\ M_\omega y\rangle_{\mathbb{C}}$ para todos $x,\ y \in \mathbb{C}^n$.*

[23] Lembre-se que os autovalores de uma matriz autoadjunta são sempre números reais.
[24] Para o estudante mais avançado: aqui poderíamos usar também o teorema espectral, Teorema 10.7.

Igualmente, se ω é um produto escalar em $\mathbb{C}^n$, então existe uma (única) matriz autoadjunta B_ω, 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 Isometrias, o Teorema de Mazur-Ulam e Generalizações

A noção de isometria em espaços métricos é discutida na página 1472 do Capítulo 25. Aqui trataremos dessa noção no contexto de espaços vetoriais normados.

3.5.1 Isometrias em Espaços Vetoriais Normados

Como já comentamos, em um espaço normado dados dois vetores u, v a norma de sua diferença, $\|u - v\|$ pode ser interpretada como uma medida de distância entre eles. Efetivamente, como discutimos no Capítulo 25, página 1454, essa expressão define uma métrica nesse espaço.

Sejam $\mathcal{V}_1$ e $\mathcal{V}_2$ espaços vetoriais dotados de normas $\|\cdot\|_{\mathcal{V}_1}$ e $\|\cdot\|_{\mathcal{V}_2}$, respectivamente. Uma *isometria* entre $\mathcal{V}_1$ e $\mathcal{V}_2$ em relação às normas $\|\cdot\|_{\mathcal{V}_1}$ e $\|\cdot\|_{\mathcal{V}_2}$ é uma função $A : \mathcal{V}_1 \to \mathcal{V}_2$ (não necessariamente suposta linear!) tal que para todos $u,\ v \in \mathcal{V}_1$ vale

$$\big\|A(u) - A(v)\big\|_{\mathcal{V}_2} \ = \ \|u - v\|_{\mathcal{V}_1} \ . \tag{3.59}$$

Assim, uma isometria, como o nome já sugere, preserva as distância entre os vetores, definidas pela normas $\|\cdot\|_{\mathcal{V}_1}$ e $\|\cdot\|_{\mathcal{V}_2}$.

No que segue estabeleceremos alguns fatos importantes sobre isometrias desse tipo. Um deles é que toda isometria é injetora. De fato, se $A : \mathcal{V}_1 \to \mathcal{V}_2$ é uma isometria e houver $u,\ v \in \mathcal{V}_1$ tais que $A(u) = A(v)$, a relação (3.59) informa que $u = v$, demonstrando a injetividade.

Como consequência, podemos afirmar que se uma isometria $A : \mathcal{V}_1 \to \mathcal{V}_2$ é sobrejetora, então ela é bijetora e, portanto, possui uma inversa, $A^{-1} : \mathcal{V}_2 \to \mathcal{V}_1$. Em tal caso, segue também de (3.59) (tomando-se $u = A^{-1}(x)$ e $v = A^{-1}(y)$) que

$$\|x - y\|_{\mathcal{V}_2} \ = \ \big\|A^{-1}(x) - A^{-1}(y)\big\|_{\mathcal{V}_1}$$

para todos $x,\ y \in \mathcal{V}_2$, o que revela que $A^{-1} : \mathcal{V}_2 \to \mathcal{V}_1$ também é uma isometria.

Isometrias são também contínuas em relação às topologias métricas produzidas em $\mathcal{V}_1$ e $\mathcal{V}_2$ pelas normas $\|\cdot\|_{\mathcal{V}_1}$ e $\|\cdot\|_{\mathcal{V}_2}$, respectivamente, pois é claro por (3.59) que $\big\|A(u) - A(v)\big\|_{\mathcal{V}_2}$ tende a zero se e somente se $\|u - v\|_{\mathcal{V}_1}$ tende a zero. Assim, caso uma isometria $A : \mathcal{V}_1 \to \mathcal{V}_2$ seja sobrejetora (e, portanto, bijetora), ela é um exemplo de homeomorfismo: uma aplicação contínua com inversa contínua.

Um outro fato simples, porém relevante, a se observar é que se $\mathcal{V}_1$, $\mathcal{V}_2$ e $\mathcal{V}_3$ são espaços vetoriais reais normados e se $A : \mathcal{V}_1 \to \mathcal{V}_2$ e $B : \mathcal{V}_2 \to \mathcal{V}_3$ são isometrias, então a composição $B \circ A : \mathcal{V}_1 \to \mathcal{V}_3$ é também uma isometria. A prova é elementar e deixada ao estudante.

Vejamos alguns exemplos elementares de isometrias. Seja $\mathcal{V}$ um espaço vetorial normado e tomemos $t \in \mathcal{V}$, fixo. A aplicação $T_t : \mathcal{V} \to \mathcal{V}$ definida por $T(u) = u + t$ é uma isometria, como facilmente se verifica. Trata-se de uma translação de todos os vetores de $\mathcal{V}$ por $t \in \mathcal{V}$, fixo. Um outro exemplo desse mesmo tipo, com alguma relevância, são as reflexões: seja $b \in \mathcal{V}$ um vetor fixo e defina-se $\rho_b : \mathcal{V} \to \mathcal{V}$ por $\rho_b(u) = b - u$ para cada $u \in \mathcal{V}$. Essa aplicação ρ_b representa uma "reflexão" em torno de b e é elementar constatar que se trata de uma isometria bijetora. Outro exemplo é aquele em que $A = U$, sendo $U : \mathcal{V} \to \mathcal{V}$ um operador linear satisfazendo $\|Uv\| = \|v\|$ para todo $v \in \mathcal{V}$. Por fim, temos também um exemplo que combina esse último e as translações: seja $U : \mathcal{V}_1 \to \mathcal{V}_2$, linear e satisfazendo $\|Uv\|_{\mathcal{V}_2} = \|v\|_{\mathcal{V}_1}$ para todo $v \in \mathcal{V}_1$ e seja também $t \in \mathcal{V}_2$, fixo. Então, $A : \mathcal{V}_1 \to \mathcal{V}_2$ dada por $A(v) := Uv + t$ é também uma isometria, como facilmente se constata. Um dos resultados que estabeleceremos adiante, o *Teorema de Mazur-Ulam*, afirma que no caso de espaços vetoriais *reais* normados esse último exemplo é o mais geral possível.

No caso de espaços vetoriais sobre os complexos, porém, podem ocorrer outras complicações. Por exemplo, seja $\mathcal{V} = \mathbb{C}^n$, com a norma $\big\|(z_1,\ \ldots,\ z_n)\big\|_2 := \sqrt{|z_1|^2 + \cdots + |z_n|^2}$. A aplicação $A\big((z_1,\ \ldots,\ z_n)\big) := \big(\overline{z_1} + t_1,\ \ldots,\ \overline{z_n} + t_n\big)$, sendo $(t_1,\ \ldots,\ t_n) \in \mathbb{C}^n$ um vetor fixo, é claramente uma isometria, mas não combina um operador linear com uma translação. Notar que a conjugação complexa é aditiva, mas antilinear.

Ainda no caso de espaços vetoriais complexos, outras complicações são também possíveis: podemos, no exemplo acima, considerar transformações nas quais a conjugação complexa é feita apenas em algumas das componentes. Por exemplo, em $\mathbb{C}^2$, a transformação $A\big((z_1,\, z_2)\big) := \big(\overline{z_1},\, z_2\big)$ é uma isometria na norma $\big\|(z_1,\, z_2)\big\|_2 := \sqrt{|z_1|^2 + |z_2|^2}$, mas não é nem linear nem antilinear, apenas aditiva.

●●● ●● ● ●● ●●●

O Teorema de Mazur-Ulam, que passamos a apresentar, é o início de uma série de resultados, similares entre si na natureza de suas afirmações, que influenciam áreas como a Teoria de Grupos e mesmo a Mecânica Quântica.

3.5.2 O Teorema de Mazur-Ulam

O Teorema de Mazur[25]-Ulam[26], que apresentamos e demonstramos a seguir, contém uma informação bastante poderosa: toda isometria sobrejetora entre espaços métricos reais e normados $\mathcal{V}_1$ e $\mathcal{V}_2$ é uma transformação afim (para a definição, vide Seção 5.2.2, página 348), e portanto, por força da Proposição 5.8, página 348, é uma combinação de uma isometria linear entre $\mathcal{V}_1$ e $\mathcal{V}_2$ e de uma translação em $\mathcal{V}_2$.

Podemos, dessa forma, afirmar que o grupo das isometrias bijetoras de um espaço vetorial real normado $\mathcal{V}$ em si mesmo é o *produto semidireto* (vide Seção 2.2.4.2, página 186) do grupo de isometrias bijetoras lineares de $\mathcal{V}$ e o grupo de translações de $\mathcal{V}$.

Apresentamos, com algumas melhorias, uma demonstração obtida por Bogdan Nica em [412], que cita outras fontes prévias. Essa demonstração de [412] tem a virtude de sistematizar os detalhes algébricos da demonstração original e de outras posteriores. A demonstração original de Mazur e Ulam[27] é a apresentada em [339].

• A deficiência afim de uma isometria

Para a demonstração do Teorema de Mazur-Ulam, a seguir, é relevante introduzirmos a noção de deficiência afim de uma isometria.

Sejam $\mathcal{U}$ e $\mathcal{V}$ dois espaços vetoriais reais dotados de normas $\|\cdot\|_{\mathcal{U}}$ e $\|\cdot\|_{\mathcal{V}}$, respectivamente, e seja $C : \mathcal{U} \to \mathcal{V}$ uma isometria. Fixemos $x,\, y \in \mathcal{V}$ e definamos

$$\mathrm{Def}_{x,y}(C) \; := \; \left\| C\left(\frac{1}{2}(x+y)\right) - \frac{1}{2}\big(C(x) + C(y)\big) \right\|_{\mathcal{V}} . \tag{3.60}$$

Essa quantidade é denominada *deficiência afim*, ou *índice de deficiência afim*, de C.

Observemos primeiramente que podemos trivialmente escrever

$$C\left(\frac{1}{2}(x+y)\right) - \frac{1}{2}\big(C(x)+C(y)\big) \; = \; \frac{1}{2}\left[C\left(\frac{1}{2}(x+y)\right) - C(x)\right] + \frac{1}{2}\left[C\left(\frac{1}{2}(x+y)\right) - C(y)\right] .$$

Assim,

$$\begin{aligned} \mathrm{Def}_{x,y}(C) \quad & \leq \quad \frac{1}{2}\left\|C\left(\frac{1}{2}(x+y)\right) - C(x)\right\|_{\mathcal{V}} + \frac{1}{2}\left\|C\left(\frac{1}{2}(x+y)\right) - C(y)\right\|_{\mathcal{V}} \\ & \overset{(C \text{ é isometria})}{=} \quad \frac{1}{2}\left\|\frac{1}{2}(x+y) - x\right\|_{\mathcal{U}} + \frac{1}{2}\left\|\frac{1}{2}(x+y) - y\right\|_{\mathcal{U}} , \end{aligned}$$

estabelecendo que

$$\mathrm{Def}_{x,y}(C) \; \leq \; \frac{1}{2}\big\|x-y\big\|_{\mathcal{U}} . \tag{3.61}$$

[25]Stanisław Mieczysław Mazur (1905–1981).
[26]Stanisław Marcin Ulam (1909–1984).
[27]Para a referência original: Stanisław Mazur; Stanisław Ulam, "Sur les transformations isométriques d'espaces vectoriels normés". C. R. Acad. Sci. Paris. **194**: 946-948 (1932).

O ponto relevante dessa desigualdade é estabelecer um limitante superior para $\mathrm{Def}_{x,y}(C)$ que é independente da isometria C. Observe-se também que a majoração do lado direito é também independente do espaço $\mathcal{V}$, onde a imagem de C é definida, e da norma definida nesse espaço.

• **O Teorema de Mazur-Ulam. Enunciado e demonstração**

Teorema 3.8 (Teorema de Mazur-Ulam) *Sejam $\mathcal{V}_1$ e $\mathcal{V}_2$ espaços vetoriais reais e normados, com normas $\|\cdot\|_{\mathcal{V}_1}$ e $\|\cdot\|_{\mathcal{V}_2}$, respectivamente, e seja $A : \mathcal{V}_1 \to \mathcal{V}_2$ uma isometria sobrejetora (e, portanto, bijetora). Então, A é uma transformação afim, ou seja, para todos $x, y \in \mathcal{V}_1$ e todo $\lambda \in [0, 1]$ vale*

$$A\big(\lambda x + (1-\lambda)y\big) \;=\; \lambda A(x) + (1-\lambda)A(y) \tag{3.62}$$

e, por ser contínua, a aplicação $U : \mathcal{V}_1 \to \mathcal{V}_2$ definida por $U(x) := A(x) - A(0)$ é real-linear (pela Proposição 5.8, página 348), é também uma isometria, ou seja, $\|Ux\|_{\mathcal{V}_2} = \|x\|_{\mathcal{V}_1}$ para todo $x \in \mathcal{V}_1$, e é também bijetora.

Em outras palavras, toda isometria sobrejetora $A : \mathcal{V}_1 \to \mathcal{V}_2$ entre espaços vetoriais reais normados é da forma $A(x) = Ux + b$ para todo $x \in \mathcal{V}_1$, onde $b \in \mathcal{V}_2$ é constante e $U : \mathcal{V}_1 \to \mathcal{V}_2$ é real-linear, bijetora e satisfaz $\|Ux\|_{\mathcal{V}_2} = \|x\|_{\mathcal{V}_1}$ para todo $x \in \mathcal{V}_1$. □

Provado Teorema 3.8. Se A é uma isometria, então A é contínua, como já comentamos. O ponto essencial da demonstração é provar que A é uma transformação afim, tema desenvolvido na Seção 5.2.2, página 348, ou seja, que todo $x, y \in \mathcal{V}_1$ e todo $\lambda \in [0, 1]$, vale $A\big(\lambda x + (1-\lambda)y\big) = \lambda A(x) + (1-\lambda)A(y)$.

Para provar isso, fixemos $x, y \in \mathcal{V}_1$, com $x \neq y$ (caso $x = y$ a igualdade (3.62) é evidente), tomemos provisoriamente o caso particular em que $\lambda = \frac{1}{2}$ (o caso geral será tratado ao final) e consideremos a deficiência afim $\mathrm{Def}_{x,y}(A)$ tal como na definição (3.60).

Nosso desejo é provar que se A é uma isometria, então $\mathrm{Def}_{x,y}(A) = 0$ para todos $x, y \in \mathcal{V}_1$, pois disso seguirá facilmente que A é uma transformação afim.

Ainda com x e y fixos, vamos denotar A por $A^{(1)} : \mathcal{V}_1 \to \mathcal{V}_2$ e definir, $A^{(n+1)} : \mathcal{V}_1 \to \mathcal{V}_1$ para cada $n \in \mathbb{N}$, por

$$A^{(n+1)} \;:=\; \big(A^{(n)}\big)^{-1} \circ \rho^{(n)} \circ A^{(n)} \;,$$

onde $\rho^{(1)} : \mathcal{V}_2 \to \mathcal{V}_2$ e $\rho^{(n)} : \mathcal{V}_1 \to \mathcal{V}_1$, $n \in \mathbb{N}$, $n \geq 2$, são as reflexões definidas por

$$\begin{aligned} \rho^{(1)}(z) \;&:=\; \Big(A^{(1)}(x) + A^{(1)}(y)\Big) - z \;=\; \big(A(x) + A(y)\big) - z \;, \text{ com } z \in \mathcal{V}_2 \;, \\ \rho^{(n)}(z) \;&:=\; \Big(A^{(n)}(x) + A^{(n)}(y)\Big) - z \;, \text{ para } n \in \mathbb{N},\; n \geq 2 \;, \text{ com } z \in \mathcal{V}_1 \;. \end{aligned}$$

Observe-se que para quaisquer $z, w \in \mathcal{V}_2$ vale $\big\|\rho^{(1)}(z) - \rho^{(1)}(w)\big\|_{\mathcal{V}_2} = \big\|z - w\big\|_{\mathcal{V}_2}$ e, para quaisquer $z, w \in \mathcal{V}_1$ e $n \geq 2$, vale $\big\|\rho^{(n)}(z) - \rho^{(n)}(w)\big\|_{\mathcal{V}_1} = \big\|z - w\big\|_{\mathcal{V}_1}$. Portanto, todas as aplicações $\rho^{(n)}$, $n \in \mathbb{N}$, são isometrias bijetoras.

Segue dessas definições que $A^{(2)} = A^{-1} \circ \rho^{(1)} \circ A$ é uma isometria bijetora de $\mathcal{V}_1$ em si mesmo e, portanto, por indução, concluímos que todas as aplicações $A^{(n)}$, $n \neq 2$, são também isometrias bijetoras de $\mathcal{V}_1$ em si mesmo. (Apenas $A^{(1)} \equiv A$ é uma isometria bijetora de $\mathcal{V}_1$ em $\mathcal{V}_2$).

Observe-se agora que para todo $n \in \mathbb{N}$ tem-se

$$\begin{aligned} \big(\rho^{(n)} \circ A^{(n)}\big)(x) \;&=\; \rho^{(n)}\Big(A^{(n)}(x)\Big) \;=\; A^{(n)}(x) + A^{(n)}(y) - A^{(n)}(x) \;=\; A^{(n)}(y) \quad \text{e} \\ \big(\rho^{(n)} \circ A^{(n)}\big)(y) \;&=\; \rho^{(n)}\Big(A^{(n)}(y)\Big) \;=\; A^{(n)}(x) + A^{(n)}(y) - A^{(n)}(y) \;=\; A^{(n)}(x) \;. \end{aligned}$$

Portanto, para todo $n \in \mathbb{N}$ tem-se

$$A^{(n+1)}(x) \;=\; y \qquad \text{e} \qquad A^{(n+1)}(y) \;=\; x \;. \tag{3.63}$$

Seguindo a definição (3.60), consideremos, para $n \geq 1$,

$$
\begin{aligned}
\mathrm{Def}_{x,y}\big(A^{(n+1)}\big) \quad &= \quad \left\| A^{(n+1)}\left(\frac{1}{2}(x+y)\right) - \frac{1}{2}\Big(A^{(n+1)}(x) + A^{(n+1)}(y)\Big) \right\|_{\mathcal{V}_1} \\
&\overset{(3.63)}{=} \quad \left\| A^{(n+1)}\left(\frac{1}{2}(x+y)\right) - \frac{1}{2}(y+x) \right\|_{\mathcal{V}_1} \\
&\overset{\text{def. de } A^{(n+1)}}{=} \quad \left\| \big(A^{(n)}\big)^{-1}\left(\big(\rho^{(n)} \circ A^{(n)}\big)\left(\frac{1}{2}(x+y)\right)\right) - \frac{1}{2}(x+y) \right\|_{\mathcal{V}_1} \\
&\overset{\text{isometria}}{=} \quad \left\| \big(\rho^{(n)} \circ A^{(n)}\big)\left(\frac{1}{2}(x+y)\right) - A^{(n)}\left(\frac{1}{2}(x+y)\right) \right\|_{\mathcal{V}_1} .
\end{aligned}
\tag{3.64}
$$

(Nota: no caso $n = 1$ a norma ao lado direito da última linha acima tem de ser a norma em $\mathcal{V}_2$).

Agora, tem-se, porém,

$$
\rho^{(n)}\left(A^{(n)}\left(\frac{1}{2}(x+y)\right)\right) \;=\; A^{(n)}(x) + A^{(n)}(y) - A^{(n)}\left(\frac{1}{2}(x+y)\right)
$$

e retornando com isso ao lado direito de (3.64), obtemos

$$
\mathrm{Def}_{x,y}\big(A^{(n+1)}\big) \;=\; \left\| A^{(n)}(x) + A^{(n)}(y) - 2A^{(n)}\left(\frac{1}{2}(x+y)\right) \right\|_{\mathcal{V}_1} \;=\; 2\,\mathrm{Def}_{x,y}\Big(A^{(n)}\Big) .
$$

Disso segue que para todo $n \in \mathbb{N}$ vale

$$
\mathrm{Def}_{x,y}\big(A^{(n+1)}\big) \;=\; 2^n\,\mathrm{Def}_{x,y}\big(A^{(1)}\big) \;\equiv\; 2^n\,\mathrm{Def}_{x,y}\big(A\big) . \tag{3.65}
$$

O problema com essa igualdade é que ela indica que, caso $\mathrm{Def}_{x,y}\big(A\big)$ seja não nulo, $\mathrm{Def}_{x,y}\big(A^{(n+1)}\big)$ cresce com n. Porém, quando obtivemos (3.61) observamos que aquela majoração independe do particular isomorfismo considerado. Assim, deve sempre valer $\mathrm{Def}_{x,y}\big(A^{(n)}\big) \leq \|x-y\|_{\mathcal{V}_1}$ para todo n, condição essa que, por (3.65), é inevitavelmente violada para algum n suficientemente grande. A única escapatória é concluir que $\mathrm{Def}_{x,y}\big(A\big) = 0$, ou seja,

$$
A\left(\frac{1}{2}(x+y)\right) \;=\; \frac{1}{2}A(x) + \frac{1}{2}A(y) . \tag{3.66}
$$

Vide (3.60). Até aqui x e y eram arbitrários, porém distintos, mas (3.66) trivialmente vale também se $x = y$. Portanto, concluímos que ela vale para todos $x, y \in \mathcal{V}_1$.

É fácil concluir disso, evocando a continuidade de A, que $A\big(\lambda x + (1-\lambda)y\big) = \lambda A(x) + (1-\lambda)A(y)$ para todo $\lambda \in [0,\, 1]$. O raciocínio para tal, também empregado na demonstração da Proposição 5.4, página 344, é o seguinte. Dados dois vetores u e v em um espaço vetorial, o vetor $\frac{1}{2}(u+v)$ representa o ponto intermediário entre eles. Assim, como (3.60) é válida para quaisquer pontos $x, y \in \mathcal{V}_1$, ela nos informa que A leva os pontos intermediários de qualquer par de vetores no ponto intermediário de suas imagens. Se assim é, isso se aplica também ao par de pontos x e $z_1 \equiv \frac{1}{2}(x+y)$ e ao par de pontos $z_1 \equiv \frac{1}{2}(x+y)$ e y:

$$
A\left(\frac{1}{2}(x+z_1)\right) \;=\; \frac{1}{2}A(x) + \frac{1}{2}A(z_1) \qquad \text{e} \qquad A\left(\frac{1}{2}(z_1+y)\right) \;=\; \frac{1}{2}A(z_1) + \frac{1}{2}A(y) .
$$

Os pontos intermediários entre x e z_1 e entre y e z_1 são $z_2 \equiv \frac{3}{4}x + \frac{1}{4}y$ e $z_3 \equiv \frac{1}{4}x + \frac{3}{4}y$. Essa argumento pode ser repetido para o ponto intermediário entre x e z_2, z_2 e z_1, z_1 e z_3 e z_3 e y, resultando em mais pontos situados entre x e y para os quais (3.66) vale.

Prosseguindo por indução, esse argumento pode ser repetido um conjunto enumerável de vezes e informa que há um conjunto enumerável <u>denso</u> na corda aberta $\{\lambda x + (1-\lambda)y,\ \lambda \in (0,\, 1)\}$ para os quais vale a relação (3.62), isto é, $A\big(\lambda x + (1-\lambda)y\big) = \lambda A(x) + (1-\lambda)A(y)$, especificamente, para todos os λ da forma

$$
\lambda_{n,\,k} \;=\; \frac{k}{2^n} \quad \text{com} \quad k \in \{0,\ 1\ \ldots,\ 2^n\}\,, \quad n \in \mathbb{N} .
$$

Como esse conjunto é denso em [0, 1], pela *continuidade* de A concluímos que todos os pontos dessa corda aberta satisfazem (3.62) e como x e y são elementos arbitrários de $\mathcal{V}_1$, ela vale em toda parte a para todo $\lambda \in [0,\, 1]$, estabelecendo que A é uma transformação afim.

Evocando agora a Proposição 5.8, página 348, concluímos que $U : \mathcal{V}_1 \to \mathcal{V}_2$, definida por $U(x) = A(x) - A(0)$, com $x \in \mathcal{V}_1$, é linear, e escrevemos $U(x) = Ux$. Assim, se $x \in \mathcal{V}$ vale $\|Ux\|_{\mathcal{V}_2} = \big\|A(x) - A(0)\big\|_{\mathcal{V}_2} \stackrel{\text{Isometria}}{=} \|x - 0\|_{\mathcal{V}_1} = \|x\|_{\mathcal{V}_1}$, o que significa que $U : \mathcal{V}_1 \to \mathcal{V}_2$ é uma isometria linear. U deve ser também sobrejetora, pois se sua imagem $U\mathcal{V}_1$ fosse um subespaço próprio de $\mathcal{V}_2$, a imagem de A também seria um subconjunto próprio de $\mathcal{V}_2$, pois $\mathrm{Im}\,(A) = U\mathcal{V}_1 + A(0)$, que é uma translação de $U\mathcal{V}_1$ por $A(0)$. Isso completa o que desejávamos estabelecer no Teorema 3.8. ■

3.5.3 Generalizações para Formas Bilineares

Vamos agora discutir como afirmações como as do Teorema de Mazur-Ulam, Teorema 3.8, página 297, podem ser estendidas para formas bilineares simétricas em espaços vetoriais reais. O caso de espaços sobre os complexos é tratado na Seção 3.5.4, página 301. Essas extensões são de grande relevância na Teoria de Grupos - como teremos a oportunidade de discutir na Seção 21.2.3, página 1178 - e também na Física Quântica.

Lema 3.3 *Seja V um espaço vetorial real dotado de uma forma bilinear simétrica e não degenerada $\omega : V \times V \to \mathbb{R}$. Seja q a forma quadrática associada a ω, definida por $q(u) := \omega(u,\, u)$, para todo $u \in V$ (para a definição geral de forma quadrática real, vide Seção 3.1.4, página 274). Seja $R : V \to V$ uma função sobrejetora. Então, R satisfaz as condições*

$$R(0) = 0 \qquad e \qquad q\big(R(u) - R(v)\big) \;=\; q(u - v) \quad \textit{para todos } u,\, v \in V \tag{3.67}$$

se e somente se satisfizer

$$\omega\big(R(u),\; R(v)\big) \;=\; \omega(u,\, v)\ , \tag{3.68}$$

também para todos $u,\, v \in V$.

Se ω for um produto escalar a condição de $R : V \to V$ ser sobrejetora pode ser omitida. □

Prova do Lema 3.3. A prova é dividida em duas partes:

I. Vamos supor a validade de (3.67). Tomando-se $v = 0$ na segunda igualdade de (3.67) e usando a primeira ($R(0) = 0$), tem-se $q\big(R(u)\big) = q(u)$ para todo u. Agora, por um lado,

$$\begin{aligned} q\big(R(u)\big) + q\big(R(v)\big) - q\big(R(u) - R(v)\big) \;&=\; q\big(u\big) + q\big(v\big) - q\big(u - v\big) \;=\; \omega\big(u,\; u\big) + \omega\big(v,\; v\big) - \omega\big(u - v,\; u - v\big) \\ &=\; \omega\big(u,\; v\big) + \omega\big(v,\; u\big) \stackrel{\text{simetria}}{=} 2\omega\big(u,\; v\big)\ . \end{aligned}$$

Por outro lado,

$$\begin{aligned} q\big(R(u)\big) + q\big(R(v)\big) - q\big(R(u) - R(v)\big) \;&=\; \omega\big(R(u),\; R(u)\big) + \omega\big(R(v),\; R(v)\big) - \omega\big(R(u) - R(v),\; R(u) - R(v)\big) \\ &=\; \omega\big(R(u),\; R(v)\big) + \omega\big(R(v),\; R(u)\big) \stackrel{\text{simetria}}{=} 2\omega\big(R(u),\; R(v)\big)\ . \end{aligned}$$

Isso estabeleceu que

$$\omega\big(R(u),\; R(v)\big) \;=\; \omega\big(u,\; v\big)\ .$$

II. Vamos supor a validade de (3.68). Tomando-se nela $v = 0$, temos que $\omega\big(R(u),\; R(0)\big) = 0$ para todo $u \in V$, o que implica $R(0) = 0$ pois ω é não degenerada e R é sobrejetora (se ω for um produto escalar tomamos $u = v = 0$ e (3.68) diz-nos que $\omega\big(R(0),\; R(0)\big) = 0$, o que igualmente implica $R(0) = 0$). Além disso,

$$\begin{aligned} q\big(R(u) - R(v)\big) \;&=\; \omega\big(R(u) - R(v),\; R(u) - R(v)\big) \;=\; \omega\big(R(u),\; R(u)\big) + \omega\big(R(v),\; R(v)\big) - 2\omega\big(R(u),\; R(v)\big) \\ &\stackrel{(3.68)}{=} \omega\big(u,\; u\big) + \omega\big(v,\; v\big) - 2\omega\big(u,\; v\big) \;=\; \omega(u - v,\; u - v) \;=\; q(u - v)\ , \end{aligned} \tag{3.69}$$

estabelecendo a validade de (3.67), completando a prova do Lema 3.3. ■

O teorema que segue é muito importante no estudo de funções que mantêm formas quadráticas invariantes.

Teorema 3.9 *Seja V um espaço vetorial real dotado de uma forma bilinear simétrica e não degenerada $\omega : V \times V \to \mathbb{R}$. Seja q a forma quadrática associada a ω, definida por $q(u) := \omega(u, u)$, para todo $u \in V$. Seja $A : V \to V$ uma função sobrejetora definida em V e tal que $q\big(A(u) - A(v)\big) = q\big(u - v\big)$ para todos $u, v \in V$. Então, existem: 1º um vetor constante $c \in V$ e 2º um operador linear sobrejetor $R : V \to V$ satisfazendo $\omega\big(Ru, Rv\big) = \omega(u, v)$ para todos $u, v \in V$, tais que*

$$A(u) \;=\; Ru + c$$

para todo $u \in V$. O operador linear R e a constante c são univocamente determinados.

Se ω for um produto escalar a condição de $A : V \to V$ ser sobrejetora pode ser omitida. □

Prova do Teorema 3.9. Seja V um espaço vetorial real dotado de uma forma bilinear simétrica e não degenerada $\omega : V \times V \to \mathbb{R}$. Seja q a forma quadrática associada a ω, definida por $q(u) := \omega(u, u)$, para todo $u \in V$. Seja $A : V \to V$ uma função sobrejetora definida em V tal que $q\big(A(u) - A(v)\big) = q\big(u - v\big)$ para todos $u, v \in V$.

Defina-se $c := A(0) \in V$ e defina-se uma nova função $R : V \to V$ dada por $R(v) := A(v) - c = A(v) - A(0)$ para todo $v \in V$. É claro que R é sobrejetora, por A o ser. É claro também que $R(0) = 0$ e que $q\big(R(u) - R(v)\big) = q\big(u - v\big)$ para todos $u, v \in V$.

Vamos agora mostrar que se $R : V \to V$ satisfaz (3.67) (e, portanto, (3.68), pelo Lema 3.3), então R é linear. Desejamos provar que para todos $\alpha_1, \alpha_2 \in \mathbb{R}$ e todos $u_1, u_2 \in V$ vale $R(\alpha_1 u_1 + \alpha_2 u_2) - \alpha_1 R(u_1) - \alpha_2 R(u_2) = 0$. Temos que, para $w \in V$, arbitrário,

$$\begin{aligned}
&\omega\Big(R(w),\;\; R(\alpha_1 u_1 + \alpha_2 u_2) - \alpha_1 R(u_1) - \alpha_2 R(u_2)\Big) \\
&\qquad = \;\omega\Big(R(w),\;\; R(\alpha_1 u_1 + \alpha_2 u_2)\Big) - \alpha_1 \omega\big(R(w),\;\; R(u_1)\big) - \alpha_2 \omega\big(R(w),\;\; R(u_2)\big) \\
&\qquad \overset{(3.68)}{=} \;\omega\Big(w,\; \alpha_1 u_1 + \alpha_2 u_2\Big) - \alpha_1 \omega\big(w,\; u_1\big) - \alpha_2 \omega\big(w,\; u_2\big) \\
&\qquad = \;\omega\Big(w,\; \underbrace{(\alpha_1 u_1 + \alpha_2 u_2) - \alpha_1 u_1 - \alpha_2 u_2}_{=0}\Big) \;=\; 0\;.
\end{aligned} \tag{3.70}$$

Como R é sobrejetora e w é arbitrário, a não degenerescência de ω implica que $R(\alpha_1 u_1 + \alpha_2 u_2) - \alpha_1 R(u_1) - \alpha_2 R(u_2) = 0$ para todos $\alpha_1, \alpha_2 \in \mathbb{R}$ e todos $u_1, u_2 \in V$, significando que $R : V \to V$ é linear.

Se ω for um produto escalar, prova-se que $R(\alpha_1 u_1 + \alpha_2 u_2) - \alpha_1 R(u_1) - \alpha_2 R(u_2) = 0$ mostrando-se que

$$\omega\Big(R(\alpha_1 u_1 + \alpha_2 u_2) - \alpha_1 R(u_1) - \alpha_2 R(u_2),\;\; R(\alpha_1 u_1 + \alpha_2 u_2) - \alpha_1 R(u_1) - \alpha_2 R(u_2)\Big) \;=\; 0\;,$$

o que é feito expandindo-se o lado esquerdo em nove termos do tipo $\omega\big(R(x),\; R(y)\big) = \omega(x,\; y)$, como feito em (3.70), e reagrupando-os para obter $\omega\Big((\alpha_1 u_1 + \alpha_2 u_2) - \alpha_1 u_1 - \alpha_2 u_2,\;\; (\alpha_1 u_1 + \alpha_2 u_2) - \alpha_1 u_1 - \alpha_2 u_2\Big) \;=\; 0.$

Resta apenas demonstrar a unicidade de R e c. Vamos supor que para todo $u \in V$ tenhamos $Ru + c = R'u + c'$. Tomando-se $u = 0$ concluiríamos que $c = c'$. A relação restante $Ru = R'u$, válida para todo $u \in V$, significa que $R = R'$. Isso completa a prova do Teorema 3.9. ■

Uma generalização do Lema 3.3 e do Teorema 3.9 para o caso complexo pode ser encontrada na Seção 3.5.4, página 301.

3.5.4 Generalizações para Formas Sesquilineares

O Lema 3.3 e o Teorema 3.9, páginas 299 e 300, respectivamente, possuem versões para formas sesquilineares não degeneradas em espaços vetoriais complexos. Trataremos disso na seção corrente.

• **Dois lemas preparatórios**

Lema 3.4 *Seja V um espaço vetorial complexo dotado de uma forma sesquilinear e não degenerada $\omega : V \times V \to \mathbb{C}$. Seja q a forma quadrática associada a ω, definida por $q(u) := \omega(u, u)$, para todo $u \in V$. Seja $R : V \to V$ uma função (não necessariamente linear) tal que*

$$q\big(R(u) - R(v)\big) \;=\; q(u - v) \tag{3.71}$$

para todos $u, v \in V$. Valem as seguintes afirmações:

1. *Se $R(iv) = +iR(v)$ para todo $v \in V$, então*

$$R(\gamma v) \;=\; \gamma R(v) \,, \qquad \forall v \in V \;\; e \;\; \forall \gamma \in \{-1, 1, -i, i\} \,. \tag{3.72}$$

   Fora isso, $\omega\big(R(u),\, R(v)\big) = \omega(u, v)$, para todo $u, v \in V$.

2. *Se $R(iv) = -iR(v)$ para todo $v \in V$, então*

$$R(\gamma v) \;=\; \overline{\gamma} R(v) \,, \qquad \forall v \in V \;\; e \;\; \forall \gamma \in \{-1, 1, -i, i\} \,. \tag{3.73}$$

   Fora isso, $\omega\big(R(u),\, R(v)\big) = \omega(v, u)$, para todo $u, v \in V$.

Se, em particular, ω for uma forma sesquilinear e Hermitiana, temos, no caso 1, $\omega\big(R(u),\, R(v)\big) = \omega(u, v)$ e no caso 2 $\omega\big(R(u),\, R(v)\big) = \overline{\omega(u, v)}$. □

Prova. Se $R(iv) = +iR(v)$ para todo $v \in V$, então para todo $v \in V$ vale $R(-v) = R(i^2 v) = +iR(iv) = i^2 R(v) = -R(v)$. Disso segue, pela mesma hipótese, que $R(-iv) = -R(iv) = -iR(v)$. Isso estabeleceu (3.72).

Se $R(iv) = -iR(v)$ para todo $v \in V$, então para todo $v \in V$ vale $R(-v) = R(i^2 v) = -iR(iv) = i^2 R(v) = -R(v)$. Disso segue, pela mesma hipótese, que $R(-iv) = -R(iv) = iR(v)$. Isso estabeleceu (3.73).

Consideremos o caso $R(iv) = +iR(v)$ para todo $v \in V$ e suponhamos (3.71). Pela identidade de polarização (3.13), página 270,

$$
\begin{aligned}
\omega\big(R(u),\, R(v)\big) \;&\stackrel{(3.13)}{=}\; \frac{1}{4}\sum_{k=0}^{3} i^k \omega\Big(R(u) + i^{-k}R(v),\; R(u) + i^{-k}R(v)\Big) \\
&\stackrel{(3.72)}{=}\; \frac{1}{4}\sum_{k=0}^{3} i^k \omega\Big(R(u) - R\big(-i^{-k}v\big),\; R(u) - R\big(-i^{-k}v\big)\Big) \\
&=\; \frac{1}{4}\sum_{k=0}^{3} i^k q\Big(R(u) - R\big(-i^{-k}v\big)\Big) \\
&\stackrel{(3.71)}{=}\; \frac{1}{4}\sum_{k=0}^{3} i^k q\big(u + i^{-k}v\big) \\
&=\; \frac{1}{4}\sum_{k=0}^{3} i^k \omega\big(u + i^{-k}v,\; u + i^{-k}v\big) \\
&\stackrel{(3.13)}{=}\; \omega(u,\, v) \,.
\end{aligned}
$$

Consideremos agora o caso $R(iv) = -iR(v)$ para todo $v \in V$ e suponhamos (3.71). Pela identidade de polarização (3.13), página 270,

$$
\begin{aligned}
\omega\big(R(u),\ R(v)\big) &\stackrel{(3.13)}{=} \frac{1}{4}\sum_{k=0}^{3} i^k\omega\Big(R(u)+i^{-k}R(v),\ R(u)+i^{-k}R(v)\Big)\\
&\stackrel{(3.73)}{=} \frac{1}{4}\sum_{k=0}^{3} i^k\omega\Big(R(u)-R\big(-i^k v\big),\ R(u)-R\big(-i^k v\big)\Big)\\
&= \frac{1}{4}\sum_{k=0}^{3} i^k q\Big(R(u)-R\big(-i^k v\big)\Big)\\
&\stackrel{(3.71)}{=} \frac{1}{4}\sum_{k=0}^{3} i^k q\big(u+i^k v\big)\\
&= \frac{1}{4}\sum_{k=0}^{3} i^k\omega\big(u+i^k v,\ u+i^k v\big)\\
&\stackrel{\text{sesquilinearidade}}{=} \frac{1}{4}\sum_{k=0}^{3} i^k\omega\big(i^{-k}u+v,\ i^{-k}u+v\big)\\
&\stackrel{(3.13)}{=} \omega(v,\ u)\ .
\end{aligned}
$$

Isso completa a prova. ■

Com alguns ingredientes adicionais as afirmações do Lema 3.4 pode ser mais apuradas.

Lema 3.5 *Seja V um espaço vetorial complexo dotado de uma forma sesquilinear e não degenerada $\omega : V \times V \to \mathbb{C}$. Seja q a forma quadrática associada a ω, definida por $q(u) := \omega(u, u)$, para todo $u \in V$. Seja $R : V \to V$ uma função sobrejetora satisfazendo ou $R(iv) = +iR(v)$ para todo $v \in V$ ou $R(iv) = -iR(v)$ para todo $v \in V$. Então, R satisfaz as condições*

$$
R(0) \;=\; 0 \qquad e \qquad q\big(R(u)-R(v)\big) \;=\; q(u-v) \qquad \text{para todos } u,\, v \in V \tag{3.74}
$$

se e somente se satisfizer a condição (caso "+")

$$
\omega\big(R(u),\ R(v)\big) \;=\; \omega(u,\, v) \tag{3.75}
$$

para todos $u,\, v \in V$, ou a condição (caso "−")

$$
\omega\big(R(u),\ R(v)\big) \;=\; \omega(v,\, u)\ , \tag{3.76}
$$

também para todos $u,\, v \in V$. □

Prova do Lema 3.5. A prova é dividida em duas partes.

I. Vamos supor a validade de qualquer um dos dois casos (3.75) ou (3.76). Tomando-se nela $v = 0$, temos que $\omega\big(R(u),\ R(0)\big) = 0$ para todo $u \in V$, o que implica $R(0) = 0$ pois ω é não degenerada e R é sobrejetora. Além disso,

$$
\begin{aligned}
q\big(R(u)-R(v)\big) &= \omega\big(R(u)-R(v),\ R(u)-R(v)\big) = \omega\big(R(u),\ R(u)\big)+\omega\big(R(v),\ R(v)\big)-\omega\big(R(u),\ R(v)\big)-\omega\big(R(v),\ R(u)\big)\\
&\stackrel{(3.75)\text{ ou }(3.76)}{=} \omega\big(u,\ u\big)+\omega\big(v,\ v\big)-\omega\big(u,\ v\big)-\omega\big(v,\ u\big) = \omega(u-v,\ u-v) = q(u-v)\ ,
\end{aligned} \tag{3.77}
$$

que estabelece (3.74).

II. Vamos supor a validade de (3.74). Tomando-se $v = 0$ na segunda igualdade de (3.74) e usando a primeira ($R(0) = 0$), tem-se $q\big(R(u)\big) = q(u)$ para todo u. Agora, por um lado,

$$q\big(R(u)\big)+q\big(R(v)\big)-q\big(R(u)-R(v)\big) \stackrel{(3.74)}{=} q(u)+q(v)-q(u-v) = \omega(u,\ u)+\omega(v,\ v)-\omega(u-v,\ u-v) = \omega(u,\ v)+\omega(v,\ u) .$$

Por outro lado,

$$\begin{aligned} q\big(R(u)\big) + q\big(R(v)\big) - q\big(R(u) - R(v)\big) &= \omega\big(R(u),\ R(u)\big) + \omega\big(R(v),\ R(v)\big) - \omega\big(R(u) - R(v),\ R(u) - R(v)\big) \\ &= \omega\big(R(u),\ R(v)\big) + \omega\big(R(v),\ R(u)\big) . \end{aligned}$$

Isso estabeleceu que

$$\omega\big(u,\ v\big) + \omega\big(v,\ u\big) = \omega\big(R(u),\ R(v)\big) + \omega\big(R(v),\ R(u)\big) . \tag{3.78}$$

Trocando-se v por iv em (3.78), usando-se a sesquilinearidade de ω, obtemos

$$\omega\big(u,\ v\big) - \omega\big(v,\ u\big) = -i\Big[\omega\big(R(u),\ R(iv)\big) + \omega\big(R(iv),\ R(u)\big)\Big]$$

e adicionando-se isso à própria (3.78), obtemos

$$2\omega\big(u,\ v\big) = \omega\big(R(u),\ R(v)\big) + \omega\big(R(v),\ R(u)\big) - i\Big[\omega\big(R(u),\ R(iv)\big) + \omega\big(R(iv),\ R(u)\big)\Big] . \tag{3.79}$$

Se tivermos válida a condição $R(iv) = +iR(v)$, a relação (3.79) e a sesquilinearidade de ω implicam

$$\omega\big(u,\ v\big) = \omega\big(R(u),\ R(v)\big) .$$

Se tivermos válida a condição $R(iv) = -iR(v)$, a relação (3.79) e a sesquilinearidade de ω implicam

$$\omega\big(u,\ v\big) = \omega\big(R(v),\ R(u)\big) .$$

Isso provou a validade de (3.75) ou de (3.76), completando a prova do Lema 3.5. ■

• O resultado principal

Podemos agora oferecer a prometida generalização do Teorema 3.9, página 300, para formas sesquilineares.

Teorema 3.10 *Seja V um espaço vetorial complexo dotado de uma forma sesquilinear e não degenerada $\omega : V \times V \to \mathbb{C}$. Seja q a forma quadrática associada a ω, definida por $q(u) := \omega(u,\ u)$, para todo $u \in V$. Seja $A : V \to V$ uma função sobrejetora definida em V e tal que $q\big(A(u) - A(v)\big) = q\big(u - v\big)$ para todos $u,\ v \in V$ e satisfazendo $A(iu) - A(0) = \pm i\big(A(u) - A(0)\big)$ também para todo $u \in V$. Então, existem: 1º um vetor constante $c \in V$ e, 2º um operador sobrejetor $R : V \to V$ satisfazendo*

- *(caso "+"): $R : V \to V$ é linear e $\omega\big(Ru,\ Rv\big) = \omega(u,\ v)$ para todos $u,\ v \in V$;*
- *(caso "−"): $R : V \to V$ é antilinear e $\omega\big(Ru,\ Rv\big) = \omega(v,\ u)$ para todos $u,\ v \in V$;*

e tais que

$$A(u) = Ru + c$$

para todo $u \in V$. O operador R e a constante c são univocamente determinados. □

Prova do Teorema 3.10. Seja V um espaço vetorial complexo dotado de uma forma sesquilinear e não degenerada $\omega : V \times V \to \mathbb{C}$. Seja q a forma quadrática associada a ω, definida por $q(u) := \omega(u, u)$, para todo $u \in V$. Seja $A : V \to V$ uma função sobrejetora definida em V tal que $q\big(A(u) - A(v)\big) = q\big(u - v\big)$ para todos $u, v \in V$.

Defina-se $c := A(0) \in V$ e defina-se uma nova função $R : V \to V$ dada por $R(v) := A(v) - c$ para todo $v \in V$. É claro que R é sobrejetora, por A o ser. É claro também que $R(0) = 0$ e que $q\big(R(u) - R(v)\big) = q\big(u - v\big)$ para todos $u, v \in V$.

Além disso, pelas hipóteses, $R(iu) = A(iu) - A(0) = \pm i\big(A(u) - A(0)\big) = \pm iR(u)$, para todo $u \in V$. Similarmente, vale também, sob as mesmas hipóteses, $R(-u) = R(i^2u) = \pm iR(iu) = i^2R(u) = -R(u)$.

Vamos agora mostrar que se $R : V \to V$ satisfaz (3.74) (e, portanto, (3.75) ou (3.76), pelo Lema 3.5) e as demais hipóteses então R é linear ou antilinear. Tomemos $u_1, u_2 \in V$ e $\alpha_1, \alpha_2 \in \mathbb{C}$, arbitrários. Temos que, para $w \in V$, arbitrário, e assumindo o caso $R(iv) = +iR(v)$,

$$
\begin{aligned}
&\omega\Big(R(w),\ \ R(\alpha_1 u_1 + \alpha_2 u_2) - \alpha_1 R(u_1) - \alpha_2 R(u_2)\Big) \\
&\qquad = \ \omega\Big(R(w),\ \ R(\alpha_1 u_1 + \alpha_2 u_2)\Big) - \alpha_1\omega\big(R(w),\ \ R(u_1)\big) - \alpha_2\omega\big(R(w),\ \ R(u_2)\big) \\
&\qquad \overset{(3.75)}{=} \ \omega\Big(w,\ \alpha_1 u_1 + \alpha_2 u_2\Big) - \alpha_1\omega\big(w,\ u_1\big) - \alpha_2\omega\big(w,\ u_2\big) \\
&\qquad = \ \omega\Big(w,\ \underbrace{(\alpha_1 u_1 + \alpha_2 u_2) - \alpha_1 u_1 - \alpha_2 u_2}_{=0}\Big) \ = \ 0\ .
\end{aligned}
\tag{3.80}
$$

Como R é sobrejetora e w é arbitrário, a não degenerescência de ω implica que $R(\alpha_1 u_1 + \alpha_2 u_2) - \alpha_1 R(u_1) - \alpha_2 R(u_2) = 0$ para todos $\alpha_1,\ \alpha_2 \in \mathbb{C}$ e todos $u_1,\ u_2 \in V$, significando que $R : V \to V$ é linear.

Analogamente, assumindo o caso $R(iv) = -iR(v)$,

$$
\begin{aligned}
&\omega\Big(R(w),\ \ R(\alpha_1 u_1 + \alpha_2 u_2) - \overline{\alpha_1} R(u_1) - \overline{\alpha_2} R(u_2)\Big) \\
&\qquad = \ \omega\Big(R(w),\ \ R(\alpha_1 u_1 + \alpha_2 u_2)\Big) - \overline{\alpha_1}\omega\big(R(w),\ \ R(u_1)\big) - \overline{\alpha_2}\omega\big(R(w),\ \ R(u_2)\big) \\
&\qquad \overset{(3.76)}{=} \ \omega\Big(\alpha_1 u_1 + \alpha_2 u_2,\ w\Big) - \overline{\alpha_1}\omega\big(u_1,\ w\big) - \overline{\alpha_2}\omega\big(u_2,\ w\big) \\
&\qquad = \ \omega\Big(\underbrace{(\alpha_1 u_1 + \alpha_2 u_2) - \alpha_1 u_1 - \alpha_2 u_2}_{=0},\ w\Big) \ = \ 0\ .
\end{aligned}
\tag{3.81}
$$

Como R é sobrejetora e w é arbitrário, a não degenerescência de ω implica que $R(\alpha_1 u_1 + \alpha_2 u_2) - \overline{\alpha_1} R(u_1) - \overline{\alpha_2} R(u_2) = 0$ para todos $\alpha_1,\ \alpha_2 \in \mathbb{C}$ e todos $u_1,\ u_2 \in V$, significando que $R : V \to V$ é antilinear.

Resta apenas demonstrar a unicidade de R e c. Vamos supor que para todo $u \in V$ tenhamos $Ru + c = R'u + c'$. Tomando-se $u = 0$, concluiríamos que $c = c'$. A relação restante $Ru = R'u$, válida para todo $u \in V$, significa que $R = R'$. Isso completa a prova do Teorema 3.10. ■

3.6 Estruturas Complexas sobre Espaços Vetoriais Reais

Seja V um espaço vetorial real. Em V está, portanto, definido um produto por escalares reais: $x\,v \in V$, 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 $J : V \to V$, 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

$$(x + iy) \cdot v \ := \ xv + yJv \ . \tag{3.82}$$

As seguintes propriedades poder ser facilmente verificadas como exercício:

1. O produto por escalares complexos (3.82) 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 α por β em $\mathbb{C}$,

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

3. O produto por escalares complexos (3.82) 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 (3.82) 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 140, V é um espaço vetorial complexo com o produto definido acima. Vamos denotar por V_J esse espaço vetorial complexo, para não confundi-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 [83]) que descrevem as álgebras de comutação e anticomutação canônicas da Mecânica Quântica e das Teorias Quânticas de Campos (que descrevem modelos fermiônicos[28] e bosônicos[29]), 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 308.

• 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[30] reais e produtos escalares reais.

Lema 3.6 *Seja V um espaço vetorial real e suponha que exista um operador linear $J : V \to V$ satisfazendo $J^2 = -\mathbb{1}$. Valem as seguintes afirmações:*

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

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

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

$$\sigma(u, \ v) \ := \ \varepsilon(Ju, \ v) \ = \ -\varepsilon(u, \ Jv) \tag{3.83}$$

é uma forma simplética real e satisfaz

[28] Enrico Fermi (1901–1954).
[29] Satyendra Nath Bose (1894–1974).
[30] Para a definição, vide página 266.

(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* $\sigma : V \times V \to \mathbb{R}$ *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 $\varepsilon : V \times V \to \mathbb{R}$ *definida para todos* $u,\ v \in V$ *por*

$$\varepsilon(u,\ v) \ := \ \sigma(u,\ Jv) \ = \ -\sigma(Ju,\ v) \tag{3.84}$$

é um produto escalar real e satisfaz

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

□

Prova da parte I. Pelas hipóteses, ε é um produto escalar real e, portanto, é uma forma bilinear real, positiva, simétrica e não degenerada. Que σ definida em (3.83) 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 σ 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 ε é não degenerada, segue que $Ju = 0$, o que implica $u = 0$, pois $J^2 = -\mathbb{1}$. Isso provou que σ é não degenerada e, portanto, é uma forma simplética. Note-se agora que

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

Por fim, $\sigma(u,\ Ju) = \varepsilon(Ju,\ Ju) \geq 0$, pois ε 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, σ é uma forma simplética real e, portanto, é uma forma bilinear real, alternante e não degenerada. Que ε definida em (3.84) 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 ε 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 σ é não degenerada, segue que $u = 0$, provando que ε é uma forma não degenerada. Para todo u tem-se também $\varepsilon(u,\ u) = \sigma(u,\ Ju) \geq 0$, por hipótese, provando que ε é uma forma positiva. Assim, pela Proposição 3.3, página 272, ε 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

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

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

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

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

(b) $\sigma(u,\ Ju) \geq 0$ *para todo* $u \in V$[31]*,*

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*. Com isso, vale também*

$$\langle u,\ Jv\rangle_{J,\,\sigma} \ = \ -\langle Ju,\ v\rangle_{J,\,\sigma} \tag{3.85}$$

para todos $u,\ v \in V$*, indicando que* J *é antiautoadjunto em relação ao produto escalar* $\langle\cdot,\ \cdot\rangle_{J,\,\sigma}$*. Em verdade, na estrutura complexa* V_J *tem-se a identificação* $J = i$ *o que torna a relação (3.85) um tanto redundante.*

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

□

Prova. Mostremos em primeiro lugar que as hipóteses das partes **A** e **B** são equivalentes. Pelo Lema 3.6, página 305, a existência de uma forma simplética real σ satisfazendo as hipóteses da parte **A** implica a existência de um produto escalar real ε dado por $\varepsilon(u,\ v) := \sigma(u,\ Jv) = -\sigma(Ju,\ v)$ satisfazendo as hipóteses da parte **B**, sendo que, por essa definição de ε,

$$\sigma(u,\ Jv) + i\sigma(u,\ v) \ = \ \varepsilon(u,\ v) + i\varepsilon(Ju,\ v)\ . \tag{3.86}$$

Reciprocamente, também pelo Lema 3.6, página 305, a existência de um produto escalar real ε satisfazendo as hipóteses da parte **B** implica a existência de uma forma simplética real σ dada por $\sigma(u,\ v) := \varepsilon(Ju,\ v) = -\varepsilon(u,\ Jv)$ satisfazendo as hipóteses da parte **A**, sendo que, por essa definição de σ, a igualdade (3.86) 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

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

Além disso,

$$\langle v,\ u\rangle_{J,\,\sigma} \ = \ \sigma(v,\ Ju) + i\sigma(v,\ u) \ = \ -\sigma(Ju,\ v) - i\sigma(u,\ v) \ = \ \sigma(u,\ Jv) - i\sigma(u,\ v) \ = \ \overline{\langle u,\ v\rangle_{J,\,\sigma}}\ . \tag{3.87}$$

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

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

[31]Em [83] essa última condição não é mencionada, mas ela é necessária.

Pela propriedade (3.87), 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 σ 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 270, 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.

Para (3.85), observemos que, para todos u e v, vale $\sigma\big(Ju,\ .Jv\big) = \sigma\big(u,\ v\big)$, onde usamos o item A-(a) e o fato que $J^2 = -\mathbb{1}$. Assim,

$$\langle Ju,\ v\rangle_{J,\,\sigma} \;=\; \sigma\big(Ju,\ Jv\big) + i\sigma\big(Ju,\ v\big) \;=\; \sigma\big(u,\ v\big) + i\sigma\big(Ju,\ v\big)\ , \tag{3.88}$$

mas

$$\langle u,\ Jv\rangle_{J,\,\sigma} \;=\; -\sigma\big(u,\ v\big) + i\sigma(u,\ Jv) \;=\; -\sigma\big(u,\ v\big) - i\sigma(Ju,\ v)\ , \tag{3.89}$$

usando novamente o item A-(a) na última passagem. Comparando-se as duas últimas igualdades, concluímos que $\langle Ju,\ v\rangle_{J,\,\sigma} = -\langle u,\ Jv\rangle_{J,\,\sigma}$, como afirmado em (3.85). ■

Comentário. A relação $J = i$ é evidente na estrutura complexa V_J, mas pode ser vista pela igualdade

$$\langle u,\ Jv\rangle_{J,\,\sigma} \overset{(3.89)}{=} -\sigma\big(u,\ v\big) + i\sigma(u,\ Jv) \;=\; i\Big(\sigma(u,\ Jv) + i\sigma\big(u,\ v\big)\Big) \;=\; i\langle u,\ v\rangle_{J,\,\sigma}\ .$$

Logo, $\langle u,\ (J-i)v\rangle_{J,\,\sigma} = 0$ para todos $u,\ v \in V$, o que implica $J = i$ em V_J, por $\langle \cdot,\ \cdot\rangle_{J,\,\sigma}$ ser um produto escalar. ♣

• **Exemplos**

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

Proposição 3.10 *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 $J^2 = -\mathbb{1}$ 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 $J^2 = -\mathbb{1}$. 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 $\mathcal{A} : V_1 \to V_2$ 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:

$$v \;=\; \begin{pmatrix} v_1 \\ \vdots \\ v_m \\ v_{m+1} \\ \vdots \\ v_{2m} \end{pmatrix} . \qquad \text{Defina-se, então,} \qquad Jv \;:=\; \begin{pmatrix} -v_{m+1} \\ \vdots \\ -v_{2m} \\ v_1 \\ \vdots \\ v_m \end{pmatrix} , \tag{3.90}$$

ou seja, em forma matricial, na mesma base,

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

sendo $\mathbb{0}_m$ 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

$$(x+iy)\cdot\begin{pmatrix} v_1 \\ \vdots \\ v_m \\ v_{m+1} \\ \vdots \\ v_{2m}\end{pmatrix} := (x+yJ)\begin{pmatrix} v_1 \\ \vdots \\ v_m \\ v_{m+1} \\ \vdots \\ v_{2m}\end{pmatrix} = \begin{pmatrix} xv_1 - yv_{m+1} \\ \vdots \\ xv_m - yv_{2m} \\ xv_{m+1} + yv_1 \\ \vdots \\ xv_{2m} + yv_m\end{pmatrix} . \tag{3.91}$$

Seguindo ainda o exemplo de (3.90) e (3.91) para $V = \mathbb{R}^{2m}$, vamos ilustrar a Proposição 3.9 e o produto escalar complexo para $(\mathbb{R}^{2m})_J$. Adotemos para ε o produto escalar usual:

$$\varepsilon(u,\ v) := \sum_{k=1}^{2m} u_k v_k = u_1v_1 + \cdots + u_{2m}v_{2m} .$$

Temos que

$$\varepsilon(Ju,\ v) = -u_{m+1}v_1 - \cdots - u_{2m}v_m + u_1v_{m+1} + \cdots + u_mv_{2m}$$

e que

$$\varepsilon(u,\ Jv) = -u_1v_{m+1} - \cdots - u_mv_{2m} + u_mv_1 + \cdots + u_{2m}v_m$$

Logo, $\varepsilon(Ju,\ v) = -\varepsilon(u,\ Jv)$ e podemos aplicar a Proposição 3.9, obtendo em $(\mathbb{R}^{2m})_J$ o produto escalar

$$\begin{aligned}
\langle u,\ v\rangle_{J,\,\varepsilon} &= \varepsilon(u,\ v) + i\varepsilon(Ju,\ v) \\
&= \Big(u_1v_1 + \cdots + u_{2m}v_{2m}\Big) + i\Big(-u_{m+1}v_1 - \cdots - u_{2m}v_m + u_1v_{m+1} + \cdots + u_mv_{2m}\Big) \\
&= u_1(v_1 + iv_{m+1}) + \cdots + u_m(v_m + iv_{2m}) + u_{m+1}(v_{m+1} - iv_1) + \cdots + u_{2m}(v_{2m} - iv_m) \\
&= \overline{(u_1 + iu_{m+1})}(v_1 + iv_{m+1}) + \cdots + \overline{(u_m + iu_{2m})}(v_m + iv_{2m}) .
\end{aligned}$$

E. 3.33 *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 38). ♦

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 $\mathcal{A} : V_1 \to V_2$, 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 3.10). 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

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

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 $\mathcal{A} : V_1 \to V_2$ 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

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

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 3.9, página 306, $\langle\alpha,\ \beta\rangle_{J,\,\varepsilon} := \varepsilon(\alpha,\ \beta) + i\varepsilon(J\alpha,\ \beta)$ 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. 3.34 *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 ε pela parte **I** do Lema 3.6, página 305, 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 274, $\sigma(u,\ v) := \mathrm{Im}(\langle u,\ v\rangle)$ $u,\ v \in V$, 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 $\sigma(u,\ Jv) = \mathrm{Im}(\langle u,\ iv\rangle) = -\mathrm{Im}(\langle iu,\ v\rangle) = -\sigma(Ju,\ v)$ e que $\sigma(u,\ Ju) = \mathrm{Im}(\langle u,\ iu\rangle) = \langle u,\ u\rangle \geq 0$. Assim, pela parte **A** da Proposição 3.9, página 306, $\langle u,\ v\rangle_{J,\,\sigma} := \sigma(u,\ Jv) + i\sigma(u,\ v)$ é um produto escalar complexo em V. No entanto, é fácil ver que nesse caso $\langle u,\ v\rangle_{J,\,\sigma} = \mathrm{Im}(\langle u,\ iv\rangle) + i\mathrm{Im}(\langle u,\ v\rangle) = \mathrm{Re}(\langle u,\ v\rangle) + i\mathrm{Im}(\langle u,\ v\rangle) = \langle u,\ v\rangle$.

O produto escalar real ε associado a σ pela parte **II** do Lema 3.6, página 305, é

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

É interessante notar também que se tivéssemos adotado $Ju = -iu$, $u \in V$, teríamos ainda para $\sigma(u,\ v) = \mathrm{Im}(\langle u,\ v\rangle)$ 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 $J : V \to V$ 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

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

ou seja, $\lambda \cdot (v_1 \oplus v_2) = (\lambda v_1) \oplus (\overline{\lambda} v_2)$, 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 $\varepsilon(u,\ v) = \mathrm{Re}\big(\langle u,\ v\rangle\big)$ 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 σ associada a ε pela parte **I** do Lema 3.6, página 305, é, tomando $u = u_1 \oplus u_2$, $v = v_1 \oplus v_2$, com $u_1,\ v_1 \in V_1$ e $u_2,\ v_2 \in V_2$,

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

como facilmente se verifica.

Pela parte **B** da Proposição 3.9, página 306, $\langle u,\ v\rangle_{J,\,\varepsilon} := \varepsilon(u,\ v) + i\varepsilon(Ju,\ v)$ 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 $u_2,\ v_2 \in V_2$,

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

E. 3.35 *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 278, 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 $B = \{b_1, \ldots, b_n\}$ de vetores linearmente independentes de V tais que todo $u \in V$ pode ser escrito de modo único como uma combinação linear $u = \alpha_1 b_1 + \cdots + \alpha_n b_n$ dos vetores de B, onde os coeficientes α_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, $u = \alpha_1 b_1 + \cdots + \alpha_n b_n$.

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

$$\|u\| = \Big\|\alpha_1 b_1 + \cdots + \alpha_n b_n\Big\| \overset{(3.26)}{\leq} |\alpha_1|\,\|b_1\| + \cdots + |\alpha_n|\,\|b_n\| \overset{(3.20)}{\leq} \sqrt{\sum_{k=1}^{n} |\alpha_k|^2}\sqrt{\sum_{k=1}^{n} \|b_k\|^2} .$$

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 $\Big|\|u\| - \|v\|\Big| \overset{(3.27)}{\leq} \|u - v\| \overset{(3.A.1)}{\leq} M_1 \|u - v\|_E$. Essa relação estabelece que a função $\varphi : \mathbb{R}^n \to \mathbb{R}$ (ou $\varphi : \mathbb{C}^n \to \mathbb{R}$) definida por

$$\varphi(\alpha_1, \ldots, \alpha_n) := \Big\|\alpha_1 b_1 + \cdots + \alpha_n b_n\Big\|$$

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

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

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

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:

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

e seja ∂B_1 seu bordo[32]:

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

[32] Para a definição da noção de bordo e para a observação que todo bordo é fechado, vide página 1593.

∂B_1 é fechado e limitado e, portanto (pelo *Teorema de Heine-Borel*, Teorema 33.14, página 1790), é compacto na topologia métrica usual. Logo, pelo Teorema 33.16, página 1791, a função contínua φ assume em ∂B_1 um mínimo $M_2 \geq 0$ e, portanto,

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

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

Seja $(\gamma_1, \ldots, \gamma_n)$ um ponto de ∂B_1 onde o mínimo de φ é assumido e seja $v_0 = \gamma_1 b_1 + \cdots + \gamma_n b_n$. O fato que $(\gamma_1, \ldots, \gamma_n) \in \partial B_1$ 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 $\left\|\alpha_1 b_1 + \cdots + \alpha_n b_n\right\| \geq M_2$ para todo vetor $u = \alpha_1 b_1 + \cdots + \alpha_n b_n$ 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

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

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 282.

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,

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

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

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

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

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

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 ω é um produto escalar, resta-nos provar que ω é uma forma sesquilinear. Como ω tem a propriedade Hermitiana (3.B.5), é suficiente provar que ω 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 ω 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 $u,\ v \in V$,

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

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

$$\omega(u,\ v) \ := \ \frac{1}{4}\Big[f(u,\ v) - if(u,\ iv)\Big] \ . \tag{3.B.8}$$

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

$$f(u,\ v) \ = \ f(v,\ u) \ , \tag{3.B.9}$$

$$f(u,\ -v) \ = \ -f(u,\ v) \ , \tag{3.B.10}$$

$$f(u,\ 0) \ = \ 0 \ . \tag{3.B.11}$$

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

Proposição 3.11 *Para todos $u,\ v$ e $w \in V$ vale*

$$f(u,\ v+w) \ = \ f(u,\ v) + f(u,\ w) \ . \tag{3.B.12}$$

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 $v \to -v$ e $w \to -w$, 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

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

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

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

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 $\lim\limits_{k\to\infty} r_k = x$. Então, usando a desigualdade (3.21), página 276, com $a = (r_k - x)v$ e $b = -u - xv$, tem-se que $\Big|\|u + r_kv\| - \|u + xv\|\Big| \leq \|(r_k - x)v\| = |r_k - x|\,\|v\|$ e como $\lim\limits_{k\to\infty} |r_k - x| = 0$, segue que $\lim\limits_{k\to\infty} \Big|\|u+r_kv\| - \|u+xv\|\Big| = 0$, ou seja, $\lim\limits_{k\to\infty} \|u+r_kv\| = \|u+xv\| = \Big\|u + \lim\limits_{k\to\infty} r_kv\Big\|$. Isso implica imediatamente que

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

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

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

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

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

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 $\omega(zu,\ v) = \overline{z}\omega(u,\ v)$. Isso estabeleceu que ω é uma forma sesquilinear, completando a prova do Teorema 3.3. ■

3.C Formas Multilineares Simétricas. Prova da Proposição 3.1

Apresentamos aqui uma demonstração da Proposição 3.1, página 264. É suficiente para tal provarmos as relações (3.6) e (3.7)[33]. As demais afirmações são imediatas a partir dessas relações.

Prova da Proposição 3.1. Para demonstrar (3.6) usemos a multilinearidade de ω e escrevamos o lado direito como

$$\frac{(-1)^n}{n!}\sum_{\epsilon_1=0}^{1}\cdots\sum_{\epsilon_n=0}^{1}(-1)^{\epsilon_1+\cdots+\epsilon_n}\omega\Big(v_\epsilon,\ \ldots,\ v_\epsilon\Big)$$

$$= \frac{(-1)^n}{n!}\sum_{\epsilon_1=0}^{1}\cdots\sum_{\epsilon_n=0}^{1}(-1)^{\epsilon_1+\cdots+\epsilon_n}\left[\sum_{k_1=1}^{n}\cdots\sum_{k_n=1}^{n}\omega\Big(\epsilon_{k_1}u_{k_1},\ \ldots,\ \epsilon_{k_n}u_{k_n}\Big)\right]$$

$$= \frac{(-1)^n}{n!}\sum_{k_1=1}^{n}\cdots\sum_{k_n=1}^{n}\left[\sum_{\epsilon_1=0}^{1}\cdots\sum_{\epsilon_n=0}^{1}(-1)^{\epsilon_1+\cdots+\epsilon_n}\omega\Big(\epsilon_{k_1}u_{k_1},\ \ldots,\ \epsilon_{k_n}u_{k_n}\Big)\right]. \tag{3.C.21}$$

Façamos agora algumas considerações sobre a expressão

$$\sum_{\epsilon_1=0}^{1}\cdots\sum_{\epsilon_n=0}^{1}(-1)^{\epsilon_1+\cdots+\epsilon_n}\omega\Big(\epsilon_{k_1}u_{k_1},\ \ldots,\ \epsilon_{k_n}u_{k_n}\Big) \tag{3.C.22}$$

que ocorre acima. Se o conjunto de índices $\{k_1,\ \ldots,\ k_n\}$ diferir de $\{1,\ \ldots,\ n\}$ (o que ocorre se alguns dos valores dos k's repetirem-se), então, para algum $j\in\{1,\ \ldots,\ n\}$ o índice k_j não comparecerá em (3.C.22). Assim, $\omega\big(\epsilon_{k_1}u_{k_1},\ \ldots,\ \epsilon_{k_n}u_{k_n}\big)$ será também independente de ϵ_{k_j}. Como a soma $\sum_{\epsilon_{k_j}=0}^{1}(-1)^{\epsilon_{k_j}}$ é nula, vemos que a toda a expressão (3.C.22) anula-se. Dessa forma, somente contribuem em (3.C.21) os conjuntos de índices com $\{k_1,\ \ldots,\ k_n\}=\{1,\ \ldots,\ n\}$, ou seja, as permutações do conjunto $\{1,\ \ldots,\ n\}$. Portanto, devemos ter $k_j=\pi(j)$ para cada $j\in\{1,\ \ldots,\ n\}$, sendo $\pi\in S_n$. Assim, o lado direito de (3.C.21) se escreve como

$$\frac{(-1)^n}{n!}\sum_{\pi\in S_n}\sum_{\epsilon_1=0}^{1}\cdots\sum_{\epsilon_n=0}^{1}(-1)^{\epsilon_1+\cdots+\epsilon_n}\omega\Big(\epsilon_{\pi(1)}u_{\pi(1)},\ \ldots,\ \epsilon_{\pi(n)}u_{\pi(n)}\Big).$$

Há de se observar agora que, como ω é linear em cada um de seus argumentos, a expressão $\omega\big(\epsilon_{k_1}u_{k_1},\ \ldots,\ \epsilon_{k_n}u_{k_n}\big)$ será nula sempre que um dos ϵ_l anular-se. Assim, só sobrevivem os termos com $\epsilon_1=\cdots=\epsilon_n=1$ e a última expressão fica

$$\frac{(-1)^n}{n!}\sum_{\pi\in S_n}(-1)^n\omega\Big(u_{\pi(1)},\ \ldots,\ u_n\Big) = \frac{1}{n!}\sum_{\pi\in S_n}\omega\Big(u_{\pi(1)},\ \ldots,\ u_n\Big).$$

Por fim, como ω é simétrica, todos os termos da soma em $\pi\in S_n$ são iguais a $\omega\big(u_1,\ \ldots,\ u_n\big)$ e esses termos são em número de $n!$, cancelando esse fator. O resultado é, portanto, $\omega\big(u_1,\ \ldots,\ u_n\big)$, estabelecendo (3.6).

A demonstração de (3.7) é apenas ligeiramente distinta. Para provar (3.7), usemos a multilinearidade de ω e escre-

[33]Para a prova de (3.6) seguimos parcialmente: Erik G. F. Thomas, "A polarization identity for multilinear maps", Indagationes Mathematicae **25**, Issue 3, p. 468–474 (2014), https://doi.org/10.1016/j.indag.2013.11.003. Vide também os trabalhos de Edward Nelson (1932–2014) and Martin Schetzen (1928–2017), lá citados.

vamos o lado direito como

$$
\begin{aligned}
&\frac{1}{2^n n!}\sum_{\epsilon_1=0}^{1}\cdots\sum_{\epsilon_n=0}^{1}(-1)^{\epsilon_1+\cdots+\epsilon_n}\omega\Big(w_\epsilon,\ \ldots,\ w_\epsilon\Big)\\
&\qquad = \frac{1}{2^n n!}\sum_{\epsilon_1=0}^{1}\cdots\sum_{\epsilon_n=0}^{1}(-1)^{\epsilon_1+\cdots+\epsilon_n}\left[\sum_{k_1=1}^{n}\cdots\sum_{k_n=1}^{n}\omega\Big((-1)^{\epsilon_{k_1}}u_{k_1},\ \ldots,\ (-1)^{\epsilon_{k_n}}u_{k_n}\Big)\right]\\
&\qquad = \frac{1}{2^n n!}\sum_{k_1=1}^{n}\cdots\sum_{k_n=1}^{n}\left[\sum_{\epsilon_1=0}^{1}\cdots\sum_{\epsilon_n=0}^{1}(-1)^{\epsilon_1+\cdots+\epsilon_n}\omega\Big((-1)^{\epsilon_{k_1}}u_{k_1},\ \ldots,\ (-1)^{\epsilon_{k_n}}u_{k_n}\Big)\right]. \qquad (3.C.23)
\end{aligned}
$$

Consideremos agora a expressão, que ocorre acima,

$$
\sum_{\epsilon_1=0}^{1}\cdots\sum_{\epsilon_n=0}^{1}(-1)^{\epsilon_1+\cdots+\epsilon_n}\omega\Big((-1)^{\epsilon_{k_1}}u_{k_1},\ \ldots,\ (-1)^{\epsilon_{k_n}}u_{k_n}\Big). \qquad (3.C.24)
$$

Pelo mesmo argumento apresentado na prova de (3.6), esse termo é nulo se o conjunto de índices $\{k_1,\ \ldots,\ k_n\}$ diferir de $\{1,\ \ldots,\ n\}$ e, portanto, os únicos termos relevantes no lado direito de (3.C.23) são aqueles para os quais $k_j = \pi(j)$ para cada $j \in \{1,\ \ldots,\ n\}$, sendo $\pi \in S_n$. Retornando com essas informações a (3.C.23), o lado direito aquela expressão fica

$$
\begin{aligned}
&\frac{1}{2^n n!}\sum_{\pi\in S_n}^{n}\left[\sum_{\epsilon_1=0}^{1}\cdots\sum_{\epsilon_n=0}^{1}(-1)^{\epsilon_1+\cdots+\epsilon_n}\omega\Big((-1)^{\epsilon_{\pi(1)}}u_{\pi(1)},\ \ldots,\ (-1)^{\epsilon_{\pi(n)}}u_{\pi(n)}\Big)\right]\\
&\qquad = \frac{1}{2^n n!}\sum_{\pi\in S_n}^{n}\left[\sum_{\epsilon_1=0}^{1}\cdots\sum_{\epsilon_n=0}^{1}(-1)^{\epsilon_1+\cdots+\epsilon_n+\epsilon_{\pi(1)}+\cdots+\epsilon_{\pi(n)}}\omega\Big(u_{\pi(1)},\ \ldots,\ u_{\pi(n)}\Big)\right]\\
&\qquad = \frac{1}{2^n n!}\sum_{\pi\in S_n}^{n}\left[\sum_{\epsilon_1=0}^{1}\cdots\sum_{\epsilon_n=0}^{1}\omega\Big(u_{\pi(1)},\ \ldots,\ u_{\pi(n)}\Big)\right]. \qquad (3.C.25)
\end{aligned}
$$

Na segunda igualdade, usamos que $\epsilon_{\pi(1)} + \cdots + \epsilon_{\pi(n)} = \epsilon_1 + \cdots + \epsilon_n$, pois π é uma bijeção de $\{1,\ \ldots,\ n\}$ em si mesmo.

Na última expressão em (3.C.25), as somas sobre os ϵ's são mudas e contribuem com um fator 2^n. Como ω é simétrica, os termos na soma em $\pi \in S_n$ independem de π (e valem $\omega\big(u_1,\ \ldots,\ u_n\big)$) e ganhamos mais um fator $n!$. O resultado é, portanto, $\omega\big(u_1,\ \ldots,\ u_n\big)$, demonstrando (3.7). ■

3.D A Identidade de Polarização para Formas Trilineares Simétricas

A identidade de polarização para formas bilineares simétricas, relação (3.4), página 264, pode ser estendida para formas trilineares simétricas (em verdade, também para formas n-lineares simétricas. Vide Proposição 3.1, página 264). Fazemos uso da mesma na Seção 35.4.4.1, página 1983, quando tratarmos de coordenadas normais.

Seja V um espaço vetorial real e $\sigma : V \times V \times V \to \mathbb{R}$ uma forma trilinear simétrica, ou seja, uma forma trilinear que satisfaz $\sigma\big(u^1,\ u^2,\ u^3\big) = \sigma\big(u^{\pi(1)},\ u^{\pi(2)},\ u^{\pi(3)}\big)$ para qualquer permutação (aplicação bijetora) $\pi : \{1,\ 2,\ 3\} \to \{1,\ 2,\ 3\}$

e quaisquer vetores $u^1, u^2, u^3 \in V$. Então, para quaisquer $u, v, w \in V$ vale

$$\begin{aligned} \sigma(u, v, w) = \frac{1}{24}\Big[& \sigma(u+v+w, u+v+w, u+v+w) \\ & -\sigma(u-v+w, u-v+w, u-v+w) - \sigma(u+v-w, u+v-w, u+v-w) \\ & +\sigma(u-v-w, u-v-w, u-v-w) \Big]. \end{aligned} \tag{3.D.26}$$

A relação (3.D.26) é a *identidade de polarização para formas trilineares simétricas*. Ela segue diretamente de (3.7), página 264. Uma maneira pedestre de demostrá-la consiste em expandir o lado direito e constatar que equivale ao lado esquerdo. Isso requer trabalho e paciência (há 108 termos na expansão!). A simetria de σ deve ser usada para a obtenção de cancelamentos.

E. 3.36 *Exercício*. Usando (3.6), página 264, em lugar de (3.7), obtenha uma outra expressão equivalente para a identidade de polarização de uma forma trilinear simétrica $\sigma(u, v, w)$:

$$\begin{aligned} \sigma(u, v, w) = \frac{1}{6}\Big[& \sigma(u+v+w, u+v+w, u+v+w) \\ & -\sigma(u+v, u+v, u+v) - \sigma(u+w, u+w, u+w) - \sigma(v+w, v+w, v+w) \\ & +\sigma(u, u, u) + \sigma(v, v, v) + \sigma(w, w, w) \Big]. \end{aligned} \tag{3.D.27}$$

✠

E. 3.37 *Exercício*. Verifique explicitamente a validade de (3.D.26) e de (3.D.27). ✠

Uma consequência imediata de (3.D.26) ou (3.D.27) é a afirmação que para uma forma trilinear simétrica σ o conhecimento dos valores "diagonais" $\sigma(r, r, r)$ para todo $r \in V$ determina os valores de $\sigma(u, v, w)$ para quaisquer $u, v, w \in V$. Temos a seguinte proposição, que dispensa demonstração em face de (3.D.26) ou (3.D.27), e que particulariza uma afirmação da Proposição 3.1 da página 264.

Proposição 3.12 *Seja σ uma forma trilinear simétrica satisfazendo $\sigma(r, r, r) = 0$ para todo $r \in V$. Então, $\sigma = 0$.* □

Parte II

Tópicos de Análise Real e Complexa