Notas de Aula
Curso de Física-Matemática II
Instituto de Física da USP
2000

João Carlos Alves Barata. Notas de aula. Versão de 15 de setembro de 2000.

Espaços Vetoriais e Noções Básicas

Sumário

Como ler estas notas. A intenção destas notas é introduzir/recordar as noções básicas de corpos e espaços vetoriais, assim como de outros conceitos algébricos básicos. A primeira seção pretende recordar alguns conceitos básicos sobre teoria de conjuntos.

Pretende-se que estas notas sejam ser usadas para consulta e, por isso, talvez as mesmas não devam ser lidas sistematicamente do começo ao fim, mas recomenda-se ao leitor procurar primeiro os temas que mais lhe interessam.

Fazemos notar que nem todos os temas aqui tratados dizem respeito ao curso de Física-Matemática II, exceto os assuntos relacionados a espaços vetoriais, normas, produtos escalares etc. Os demais assuntos são de qualquer forma importantes para a formação matemática geral de qualquer estudante.

Para melhor orientação do leitor listamos as seções mais importantes, cuja leitura não deve ser omitida. São as Seções 2.2 (página 17), 2.2.1 (página 22), 2.2.2 (página 26), 3 (página 34), 3.3 (página 37) e 3.4 (página 41).

JCAB.

1 Noções Básicas

• **Conjuntos**

Partiremos do pressuposto de serem familiares as noções básicas envolvendo conjuntos e suas relações.

Sejam X e I conjuntos arbitrários e seja associado a cada $\alpha \in I$ um sub-conjunto A_α de X. O conjunto I será freqüentemente denominado conjunto ou família de índices. Vamos introduzir algumas notações a serem usadas adiante. Definimos

$$\bigcup_{\alpha\in I} A_\alpha \;:=\; \{x \in X \text{ tal que } x \in A_\alpha \text{ para algum } \alpha \in I\} \tag{1.1}$$

e

$$\bigcap_{\alpha\in I} A_\alpha \;:=\; \{x \in X \text{ tal que } x \in A_\alpha \text{ para todo } \alpha \in I\}. \tag{1.2}$$

Para $A, B \subset X$ denotamos por $A \setminus B$ a chamada diferença entre os conjuntos A e B, a saber

$$A \setminus B \;:=\; \{x \in X \text{ tal que } x \in A \text{ mas } x \notin B\}. \tag{1.3}$$

Por vezes usa-se a notação $A - B$ para $A \setminus B$. Para $A \subset X$ denota-se por A^c o chamado complemento de A em relação a X: $A^c := X \setminus A$. Note-se que ao usar-se o símbolo A^c deve estar subentendido qual o conjunto X ao qual o complemento se refere. É fácil ver que $A \setminus B = B^c \cap A$.

Dizemos que um conjunto $B \subset A$ é um subconjunto *próprio* de A se $A \setminus B \neq \emptyset$, ou seja, que houver elementos em A que não estão em B.

Se A e B são conjuntos e $A \cap B = \emptyset$ então $A \cup B$ é dita ser uma união disjunta de A e B. A mesma nomenclatura se aplica para uniões genéricas $\bigcup_{\alpha\in I} A_\alpha$ desde que $A_\alpha \cap A_\beta = \emptyset$ para $\alpha \neq \beta$.

Por $A\triangle B$ denota-se a chamada *diferença simétrica* entre A e B:

$$A\triangle B \;:=\; (A\cup B)\setminus(A\cap B). \tag{1.4}$$

Exercício. Mostre que $A\triangle B = B\triangle A$ e que $(A\triangle B)\triangle C = A\triangle(B\triangle C)$.

Exercício. Seja $\{A_\alpha, \; \alpha \in I\}$ uma coleção arbitrária de subconjuntos de X. Prove as seguintes relações:

$$X \setminus \left(\bigcup_{\alpha\in I} A_\alpha\right) = \bigcap_{\alpha\in I} (X \setminus A_\alpha), \tag{1.5}$$

$$X \setminus \left(\bigcap_{\alpha\in I} A_\alpha\right) = \bigcup_{\alpha\in I} (X \setminus A_\alpha). \tag{1.6}$$

Se X é um conjunto denota-se por $\mathbb{P}(X)$ a coleção de todos os subconjuntos de X. $\mathbb{P}(X)$ é por vezes chamado de "conjunto das partes de X". Por convenção adota-se sempre que $\emptyset \in \mathbb{P}(X)$.

• **Intervalos**

Ainda não introduzimos os números reais nem a relação de ordem entre eles mas, como essas noções são conhecidas, vamos colocar aqui uma palavra sobre a nomenclatura usada para descrever intervalos da reta real. Para $a < b \in \mathbb{R}$ o conjunto

$$(a, \; b) = \{x \in \mathbb{R}, \quad \text{com } a < x < b\}$$

é dito ser um intervalo aberto. Para $a \le b \in \mathbb{R}$ o conjunto

$$[a, \; b] = \{x \in \mathbb{R}, \quad \text{com } a \le x \le b\}$$

é dito ser um intervalo fechado. Para $a < b \in \mathbb{R}$ os conjuntos

$$[a, \; b) = \{x \in \mathbb{R}, \quad \text{com } a \le x < b\}$$

e

$$(a, \; b] = \{x \in \mathbb{R}, \quad \text{com } a < x \le b\}$$

são ditos ser intervalos semi-abertos (ou semi-fechados).

É importante dizer que a nomenclatura "aberto" ou "fechado" acima é usada independentemente da topologia usada em $\mathbb{R}$ (a noção de topologia será introduzida adiante).

• **Pares ordenados**

Outro conceito importante que introduziremos é o de par ordenado. O conceito de par ordenado $(a, \; b)$ formado por dois elementos genéricos $a, \; b \in X$ é intuitivo. A intuição é que entende-se como par ordenado uma lista de dois elementos sendo que um deles assume a posição de "primeiro" elemento da lista (no caso, a) e o outro a de "segundo" (no caso, b). Formalmente define-se $(a, \; b)$ como sendo o conjunto $\{a, \{b\}\}$. Esta definição formal corresponde à intuição pois, no conjunto $C = \{a, \{b\}\}$, há uma distinção entre o papel de a e de b, dado que a é um elemento do conjunto C, enquanto que b é um elemento de um subconjunto de C, a saber do conjunto $C \setminus \{a\}$. Apesar de existir a definição formal acima, recomenda-se ao estudante fiar-se inicialmente na intuição por trás do conceito.

Dados dois conjuntos A e B definimos por $A \times B$ o conjunto de todos os pares ordenados $(a, \; b)$ sendo $a \in A$ e $b \in B$. O conjunto $A \times B$ é chamado de produto cartesiano de A e B. Note que, em geral, $A \times B \neq B \times A$. Por que?

• **Relações**

Sejam A e B conjuntos e seja o produto cartesiano $A \times B$. Um subconjunto de $A \times B$ é dito ser uma relação binária, ou simplesmente relação entre A e B.

Exemplo. Seja A o conjunto de homens vivos e B o conjunto de mulheres vivas e seja $R \subset A \times B$ o conjunto $R := \{(a, \; b), \; a \text{ é irmão de } b\}$. R representa uma relação (de irmandade) entre homens e mulheres.

Outros exemplos virão abaixo.

Dada uma relação $G \subset A \times B$ entre conjuntos A e B há duas noções importantes associadas: a de domínio da relação e a de imagem da relação. Define-se por domínio de G o conjunto

$$\mathrm{Dom}(G) \;:=\; \{a \in A \text{ tal que } (a, \; b) \in G \text{ para algum } b \in B\}. \tag{1.7}$$

Define-se por imagem de G o conjunto

$$\mathrm{Im}(G) \;:=\; \{b \in B \text{ tal que } (a, \; b) \in G \text{ para algum } a \in A\}. \tag{1.8}$$

Note-se que $\mathrm{Dom}(G) \subset A$ e que $\mathrm{Im}(G) \subset B$.

• **Funções**

Este é talvez o mais importante exemplo de relação. Sejam A e B conjuntos e F uma relação entre A e B. Então a relação F é dita ser uma função de A em B se $\mathrm{Dom}(F) = A$ e se $(a, \; b) \in F$ e $(a, \; b') \in F$ só for possível caso $b = b'$. Em outras palavras, a cada elemento a de A a função associa um e apenas um elemento b de B que faz o papel de segundo elemento do par ordenado $(a, \; b)$. Este segundo elemento associado pela função F ao elemento a, é mais conveniente denotá-lo por $F(a)$. Assim, uma função é o conjunto de pares $\{(a, \; F(a)) \in A \times B, \; a \in A\}$. Freqüentemente denotamos uma função F de A em B por $F: \; A \to B$.

• **Funções Sobrejetoras, Injetoras e Bijetoras**

Uma função $F: \; A \to B$ é dita ser **sobrejetora** se $\mathrm{Im}(F) = B$. Uma função $F: \; A \to B$ é dita ser **injetora** ou **injetiva** se a cada $b \in \mathrm{Im}(F)$ existir um e somente um elemento $a \in \mathrm{Dom}(F)$ tal que $(a, \; b) \in F$. Uma função que for sobrejetora e injetora é dita ser **bijetora**.

Seja uma função bijetora $F \subset A \times B$. Então a relação $F^{-1} \subset B \times A$ dada por

$$F^{-1} \;=\; \{(b, \; a) \text{ tal que } (a, \; b) \in F\}$$

é em verdade uma função denominada *função inversa de* F. É claro que $(F^{-1})^{-1} = F$.

• **Extensões de Funções**

Seja $F : A \to B$ uma função e suponha que A seja subconjunto de um outro conjunto A'. Uma função $G : A' \to B$ é dita ser uma *extensão* de F se F e G coincidirem na parte comum de seus domínios, que vem a ser o conjunto A, ou seja, se $G(a) = F(a)$ para todo $a \in A$.

Se lembrarmos que uma função $F : A \to B$ é um subconjunto de $A \times B$ e que uma função $G : A' \to B$ é um subconjunto de $A' \times B$ e se lembrarmos que $A \times B \subset A' \times B$ caso $A \subset A'$, então uma definição alternativa de extensão seria seguinte: uma função G é uma extensão de uma função F se $F \subset G$, ambas entendidas como subconjuntos de $A' \times B$.

Exercício. Verifique a equivalência dessas duas definições do conceito de extensão de funções.

• **Relações de Equivalência**

Outro tipo importante de relação é formado pelas chamadas relações de equivalência. Uma relação $E \subset A \times A$ é dita ser uma relação de equivalência em A se os seguintes itens forem satisfeitos:

1. $(a, a) \in E$ para todo $a \in A$.
2. $(a, b) \in E$ implica que $(b, a) \in E$.
3. $(a, b) \in E$ e $(b, c) \in E$ implicam que $(a, c) \in E$.

Se o par (a, b) pertence a uma relação de equivalência E então a e b são ditos serem equivalentes segundo E. Quase sempre usa-se a notação $a \sim b$ para indicar que dois elementos são equivalentes segundo uma relação de equivalência dada.

Seja A um conjunto e $E \subset A \times A$ uma relação de equivalência em A. Para cada $a \in A$ podemos definir o conjunto

$$E(a) := \{a' \in A \text{ tal que } (a, a') \in E\}. \tag{1.9}$$

Esse conjunto é chamado de *classe de equivalência* de a (pela relação de equivalência E).

Exercício. Seja A é um conjunto e $E \subset A \times A$ é uma relação de equivalência em A. Suponha que $a, b \in A$ e que $a \sim b$ segundo E. Prove que $E(a) = E(b)$.

Exercício importante. Prove que se A é um conjunto e $E \subset A \times A$ é uma relação de equivalência em A então A é a união disjunta de classes de equivalência de seus elementos.

Exercício. Seja o conjunto dos números reais $\mathbb{R}$ e seja a relação $W \subset \mathbb{R} \times \mathbb{R}$ definida por

$$W := \{(x, y) \in \mathbb{R} \times \mathbb{R} \text{ tal que } x - y \in \mathbb{Q}\}, \tag{1.10}$$

onde $\mathbb{Q}$ é o conjunto dos números racionais. Prove que W é uma relação de equivalência.

• **Relações de Compatibilidade**

Seja P um conjunto. Uma *relação de compatibilidade* em P é um conjunto $C \subset P \times P$ com as seguintes propriedades:

1. Se γ e γ' são tais que $(\gamma, \gamma') \in C$, então $(\gamma', \gamma) \in C$.
2. Para todo $\gamma \in P$ vale $(\gamma, \gamma) \notin C$.

Para uma dada relação de compatibilidade C denotamos $\gamma \sim_C \gamma'$ caso $(\gamma, \gamma') \in C$ e dizemos que γ e γ' são *C-compatíveis*. Caso contrário, denotamos $\gamma \not\sim_C \gamma'$ se $(\gamma, \gamma') \notin C$ e dizemos que γ e γ' são *C-incompatíveis*.

Se uma dada relação C é subentendida, denotamos simplesmente $\gamma \sim \gamma'$ caso $(\gamma, \gamma') \in C$ e dizemos simplesmente que γ e γ' são *compatíveis*.

Relações de compatibilidade são importantes na Mecânica Estatística, especialmente nas chamadas expansões de polímeros e de "clusters".

Exemplo. Seja $P = \mathbb{P}(\mathbb{R}^n)$, a coleção de todos os subconjuntos de $\mathbb{R}^n$. Uma relação de compatibilidade em P é a seguinte: $A \sim B \iff A \cap B = \emptyset$. Verifique.

• **Relações de Ordem**

Seja X um conjunto. Uma relação $R \subset X \times X$ é dita ser uma relação de ordem parcial em X se as seguintes condições forem satisfeitas:

1. Para todo $a \in X$ tem-se que $(a, a) \in R$.
2. Se $(a, b) \in R$ e $(b, a) \in R$ então forçosamente $a = b$.
3. Se $(a, b) \in R$ e $(b, c) \in R$ então $(a, c) \in R$.

Se X possui uma ordem parcial R, X é chamado de conjunto parcialmente ordenado por R.

Exemplo. Seja X um conjunto e $\mathbb{P}(X)$ a coleção de todos os sub-conjuntos de X. Podemos estabelecer em $\mathbb{P}(X)$ uma relação R do seguinte tipo: para $A, B \subset X$ tem-se $(A, B) \in R$ se $A \subset B$. Como exercício deixamos ao estudante mostrar que esta é uma relação de ordem parcial de acordo com a definição acima. Este exemplo ilustra também por que chamar tal relação de ordem de "parcial". A razão é que nem todo par (A, B) é elemento de R pois, para dois conjuntos A e B arbitrários, nem sempre vale que $A \subset B$ ou que $B \subset A$ (por exemplo se $A \cap B = \emptyset$).

Em função da analogia com essa relação de ordem usual dos números reais é costume, dada uma relação de ordem R qualquer, indicar que $(a, b) \in R$ através da notação $a \leq b$. Note que o símbolo de "menor ou igual" ($\leq$) é usado nesse contexto mesmo que a e b não sejam números. Por vezes usaremos também o símbolo $\preceq$ para indicar uma relação de ordem.

• **Relações de Ordem Total**

Outro conceito importante é o de relação de ordem total. Uma ordem parcial R em um conjunto X é dita ser uma relação de ordem total se para todo $a, b \in X$ tem-se que $(a, b) \in R$ ou que $(b, a) \in R$. Se X possui uma relação de ordem total R então X é dito ser *totalmente ordenado* ou *linearmente ordenado*. Ou seja, se X é um conjunto dotado de uma relação de ordem parcial então dizemos que um sub-conjunto $A \subset X$ é linearmente ordenado se $a \leq b$ ou $b \leq a$ para todo $a, b \in A$.

• **Exemplos**

Exemplos. Seja $\mathbb{R}$ o conjunto de números reais e a relação de ordem $(x, y) \in R$ se $x - y$ for um número negativo ou nulo (ou seja, se $x \leq y$). Mostre que essa é uma relação de ordem total em $\mathbb{R}$.

Contra-exemplo Seja $\mathcal{C}$ é um conjunto não-vazio qualquer. Então $\mathbb{P}(\mathcal{C})$ é ordenado pela inclusão de conjuntos: $A \leq B$ se $A \subset B$. Porém $\mathbb{P}(\mathcal{C})$ não é linearmente ordenado pois se $A \cap B = \emptyset$ não podemos dizer que $A \leq B$ nem que $B \leq A$.

Exercício. Você consegue construir uma relação de ordem em $\mathbb{R}^2$ ou em $\mathbb{R}^3$? E uma relação de ordem total?

• **Mais Exemplos**

Seja o conjunto dos números naturais $\mathbb{N}$. Podemos estabelecer em $\mathbb{N}$ a relação de ordem usual onde dizemos que $x \leq y$ se $x - y$ for um número negativo ou nulo. Esta relação é uma relação de ordem total. O leitor não deve pensar que essa é a única relação de ordem total existente em $\mathbb{N}$. Um outro exemplo é o seguinte.

Vamos estabelecer uma relação de ordem em $\mathbb{N}$ que denotaremos pelo símbolo $\preceq_{p-i}$. Sejam $a, b \in \mathbb{N}$. Se a e b forem pares dizemos que $a \preceq_{p-i} b$ se $a \leq b$. Se a e b forem ímpares dizemos que $a \preceq_{p-i} b$ se $a \leq b$. Se a é par e b é ímpar então dizemos sempre que $a \preceq_{p-i} b$.

Exercício. Mostre que a relação $\preceq_{p-i}$ estabelece uma relação de ordem total em $\mathbb{N}$.

Um exemplo análogo pode ser construído em $\mathbb{R}$. Vamos estabelecer uma relação de ordem em $\mathbb{R}$ que denotaremos pelo símbolo $\preceq_{r-i}$. Sejam $x, y \in \mathbb{R}$. Se x e y forem racionais dizemos que $x \preceq_{r-i} y$ se $x \leq y$. Se x e y forem irracionais dizemos que $x \preceq_{r-i} y$ se $x \leq y$. Se x é racional e y é irracional então dizemos sempre que $x \preceq_{r-i} y$.

Exercício. Mostre que a relação $\preceq_{r-i}$ estabelece uma relação de ordem total em $\mathbb{R}$.

• **Conjuntos Dirigidos**

Um conjunto I é dito ser um *conjunto dirigido* ("directed set") se for dotado de uma relação de ordem parcial, que denotaremos por "$\preceq$", e se for dotado da seguinte propriedade: para quaisquer dois elementos a e b de I existe pelo menos um terceiro elemento $c \in I$ tal que $a \preceq c$ e $b \preceq c$.

Exemplo. $\mathbb{R}$ é um conjunto dirigido com a relação de ordem usual.

Exemplo. $\mathbb{R}$ é um conjunto dirigido com a relação de ordem $\preceq_{r-i}$ definida acima.

Exemplo. Seja o conjunto $\mathbb{R}^n$, $n = 1, 2, \ldots$, e seja I o conjunto de todos os abertos limitados de $\mathbb{R}^n$ (um conjunto é *limitado* se for subconjunto de alguma bola aberta de raio finito centrada na origem). Mostre que I é um conjunto dirigido pela relação de ordem de inclusão: $A \preceq B$ se $A \subset B$. Note que essa relação de ordem não é uma relação de ordem total.

Contra-Exemplo. Seja X um conjunto e seja $I = \mathbb{P}(X) \setminus \{X\}$, ou seja, I é a coleção de todos os subconjuntos de X, exceto o próprio X. Podemos ter em I uma relação de ordem (de inclusão) dizendo que $A \preceq B$ se $A \subset B$. Notemos, porém, que I não é um conjunto dirigido pois para $A \in I$, $A \neq \emptyset$ temos $X \setminus A \in I$ mas não existe em I nenhum conjunto que contenha A e $X \setminus A$ simultaneamente como subconjuntos.

Exemplo. Causalidade de Einstein. Seja $\mathbb{M}^4$ o espaço-tempo quadri-dimensional de Minkowski e sejam $E_0 = (t_0, x_0, y_0, z_0)$ e $E_1 = (t_1, x_1, y_1, z_1)$ dois eventos em $\mathbb{M}^4$. Dizemos que o evento E_0 precede causalmente o evento E_1, (em notação simbólica $E_0 \succeq_{Einstein} E_1$), se $t_0 \leq t_1$ e se

$$c^2(t_1 - t_0)^2 - (x_1 - x_0)^2 - (y_1 - y_0)^2 - (z_1 - z_0)^2 \geq 0,$$

onde c é a velocidade da luz.

Exercício. Mostre que $\succeq_{Einstein}$ é uma relação de ordem em $\mathbb{M}^4$ e que $\mathbb{M}^4$ é um conjunto dirigido por essa relação.

• Máximos e Mínimos

Se X é um conjunto dotado de uma relação de ordem parcial (que denotamos, então por $\leq$) diz-se que um elemento $z \in X$ é um *máximo* de X se $a \leq z$ para todo $a \in X$. Um elemento z é dito ser *maximal* se não existir $x \in X$, $x \neq z$ tal que $z \leq x$.

Analogamente, um elemento a é dito ser um *mínimo* de X se $a \leq x$ para todo $x \in X$. Um elemento a é dito ser *minimal* se não existir $x \in X$, $x \neq z$ tal que $z \leq x$.

Notação. Se A possuir um máximo e este for único denotamo-lo por $\max(A)$. Igualmente, se A possuir um mínimo e este for único denotamo-lo por $\min(A)$.

Um conjunto X dotado de uma relação parcial de ordem $\leq$ é dito ser um conjunto bem-ordenado se todo subconjunto A não vazio de X tem um elemento mínimo em A.

Exercício. Mostre que todo conjunto bem-ordenado segundo uma relação parcial de ordem é também totalmente ordenado segundo a mesma relação.

Exercício. A recíproca não é, entretanto, verdadeira. Mostre que $\mathbb{R}$ é totalmente ordenado pela relação usual de ordem entre números reais, mas não é um conjunto bem-ordenado.

Exercício. Mostre que o conjunto dos números naturais $\mathbb{N}$ é bem ordenado.

• Majorantes e Minorantes. Ínfimo e Supremo. Conjuntos Limitados

Seja X um conjunto dotado de uma ordem parcial e $A \subset X$. Se existe $t \in X$ tal que $a \leq t$ para todo $a \in A$ dizemos que t é um *majorante* de A.

O menor dos majorantes de A é dito ser o *supremo* de A e é indicado por $\sup(A)$. Assim, s é o supremo de A se $s \leq t$ para todo t que seja majorante de A.

Analogamente, se existe $h \in X$ tal que $h \leq a$ para todo $a \in A$ dizemos que h é um *minorante* de A.

O maior dos minorantes de A é dito ser o ínfimo de A e é indicado por $\inf(A)$. Assim, i é o ínfimo de A se $h \leq i$ para todo h que seja minorante de A.

Um conjunto $A \subset X$ que tenha pelo menos um majorante é dito ser um conjunto *limitado superiormente*. Um conjunto $A \subset X$ que tenha pelo menos um minorante é dito ser um conjunto *limitado inferiormente*.

É interessante notar o seguinte. Dado um conjunto X dotado de uma ordem parcial poderíamos nos perguntar se todo subconjunto limitado superiormente de X possui um supremo, ou , analogamente, se todo subconjunto de X limitado inferiormente possui um ínfimo. A validade ou não dessas propriedades depende de X.

Por exemplo, para $X = \mathbb{Q}$, o conjunto dos racionais com a relação de ordem usual, verifica-se que a propriedade não é valida. Tomemos $A = \{x \in \mathbb{Q},\ x^2 < 2\}$. Claramente esse conjunto é limitado inferior e superiormente mas não possui nem supremo nem ínfimo (por que?).

Para $X = \mathbb{N}$ e $X \in R$ (com as relações de ordem usuais) a propriedade é, porém, válida. Não demonstraremos esses fatos aqui e o leitor é remetido aos bons textos de Análise para as provas.

Exercício. Tome $X = \mathbb{R}$ com a relação de ordem usual. Mostre que $\inf((-1,\ 1)) = -1$ e que $\sup((-1,\ 1)) = 1$. Note que -1 e 1 não são elementos de $(-1,\ 1)$.

Exercício. Suponha que A e B sejam dois sub-conjuntos de um conjunto X dotado de uma ordem parcial e que $\inf(A)$ e $\inf(B)$ existam. Mostre então que

$$\inf(A \cup B) \ = \ \min\{\inf(A), \inf(B)\}.$$

Exercício. Suponha que A e B sejam dois sub-conjuntos de um conjunto X dotado de uma ordem parcial e que $\sup(A)$ e $\sup(B)$ existam. Mostre então que

$$\sup(A \cup B) \ = \ \max\{\sup(A), \sup(B)\}.$$

• O Lema de Zorn [1]

Uma das afirmativas fundamentais de toda a Matemática usual é o seguinte resultado, conhecido como lema de Zorn, em homenagem a seu formulador[2]:

Seja X um conjunto e $\leq$ uma relação de ordem parcial em X. Suponha que todo o sub-conjunto linearmente ordenado de X tenha pelo menos um majorante em X. Então todo sub-conjunto linearmente ordenado de X contém um majorante que é também um elemento maximal de X.

Exercício. Verifique que se $X = [0,\ 1]$ todo sub-conjunto de X é linearmente ordenado pela relação de ordem usual e tem um majorante em X e que 1 é um desses possíveis majorantes. Verifique que 1 é um elemento maximal de X.

Exercício. Verifique que se $X = [0,\ 1)$ nem todo sub-conjunto de X tem um majorante em X (tente, por exemplo, sub-conjuntos do tipo $[a,\ 1)$ com $0 \leq a < 1$). Verifique que X não tem um elemento maximal.

O Lema de Zorn é "equivalente" ao chamado Axioma da Escolha que enunciaremos abaixo, ou seja, admitir um como verdadeiro leva a demonstrar a validade do segundo. Essa equivalência não será provada aqui (vide bibliografia). Toda a Matemática usual é fundada na aceitação de um ou de outro como verdadeiro e, em princípio, uma nova Matemática pode ser construída (com resultados distintos dos da Matemática usual) se esses dois axiomas forem substituídos por um terceiro inequivalente. A relevância de tais Matemáticas em Física é uma questão em aberto.

• O Axioma da Escolha [3]

O Axioma da Escolha consiste na seguinte afirmativa simples:

Seja A_s, $s \in I$, uma família de conjuntos, onde I é um conjunto arbitrário de índices. Então podemos construir um conjunto A tomando ("escolhendo") um elemento a_s de cada conjunto A_s. Em termos mais técnicos o axioma diz que há funções $F:\ I \to \bigcup_{s \in I} A_s$ tais que $F(s) \in A_s$ para todo $s \in I$.

A primeira vista esse axioma parece constituir-se de uma obviedade. Sucede, porém, que, sobretudo pelo fato de o conjunto I de índices ser arbitrário (podendo ser até um conjunto infinito e não-contável), a afirmativa que o mesmo contém não pode ser derivada de princípios mais básicos. O axioma faz uma afirmação de existência (de uma função como a F, ou de um conjunto como A formado por elementos escolhidos de cada A_s) que, geralmente, não pode ser demonstrada construtivamente, ou seja, por exibição explícita de uma tal função F ou de um conjunto A.

Faremos uso explícito do Axioma da Escolha adiante quando exibirmos exemplos de conjuntos não-mensuráveis.

Uma típica situação na qual se faz uso do axioma da escolha ocorre quando são dados um conjunto X e uma uma relação de equivalência E em X e constrói-se um conjunto $A \subset$ tomando-se um representante de cada classe de equivalência de X por E.

Nem sempre é possível exibir explicitamente os elementos de A, mas assumimos (via axioma da escolha) que um tal conjunto existe. Para ter-se em mente um caso onde uma tal situação ocorre, tome-se o exemplo dado em (1.10), página 5.

• Cardinalidade

Seja $\mathcal{K}$ uma coleção de conjuntos. Dados dois conjuntos A e B da coleção $\mathcal{K}$, dizemos que A e B são equivalentes se houver uma função bijetora de A sobre B, ou seja, se houver uma função com domínio igual a A e imagem igual a B tal que a cada elemento $b \in B$ existe um único elemento $a \in A$ com $f(a) = b$.

[1] Este tópico pode ser omitido numa primeira leitura.
[2] Max Zorn (1906-1993).
[3] Este tópico pode ser omitido numa primeira leitura.

Exercício Mostre que essa é uma relação de equivalência entre os conjuntos da coleção $\mathcal{K}$.

Para dois conjunto que são equivalentes no sentido acima diz-se também que os mesmos têm a mesma *cardinalidade*. Ou seja, dois conjuntos têm a mesma cardinalidade se e somente se forem equivalentes.

Um conjunto A é dito ter n elementos (para um número natural n) se for equivalente ao conjunto $\{1, \ldots, n\}$.

Nota. Esta última definição pressupõe que o conceito de número natural já seja conhecido. Outra construção mais simples em termos de pressupostos é feita de modo informal como segue: diz-se que um conjunto tem um elemento se for equivalente ao conjunto $\{\emptyset\}$; que um conjunto tem dois elementos se for equivalente ao conjunto $\{\emptyset, \{\emptyset\}\}$; que tem três elementos se for equivalente ao conjunto $\{\emptyset, \{\emptyset, \{\emptyset\}\}$ e assim por diante. Em verdade essa construção permite produzir uma *definição* do conceito de número natural: o número "um" é, grosseiramente falando, o nome dado à classe de equivalência formada pelos conjuntos equivalentes ao conjunto $\{\emptyset\}$; o número "dois" é o nome dado à classe de equivalência do conjunto $\{\emptyset, \{\emptyset\}\}$; o número "três" é nome dado à classe de equivalência do conjunto $\{\emptyset, \{\emptyset, \{\emptyset\}\}$ e assim por diante. Alias, o número "zero" é o nome dado à classe de equivalência de $\emptyset$. O números naturais seriam então o conjunto de todas as classes de equivalência construídas dessa forma. Esta definição do conceito de número natural pressupõe apenas conhecidos conceitos primitivos como os de conjuntos, classes de equivalência e de conjunto vazio. A mesma é devida a von Neumann[4].

Diz-se que um conjunto A é finito se tiver a cardinalidade de $\{1, \ldots, n\}$ para algum $n \in \mathbb{N}$. A é dito ser infinito se não for finito.

Exercício. Seja A um conjunto finito com n elementos. Mostre que $\mathbb{P}(A)$ tem 2^n elementos.

- **Conjuntos Contáveis**

Um conjunto A é dito ser contável se for finito ou se tiver a cardinalidade do conjunto dos números naturais, ou seja, se for finito ou se existir uma função bijetora $f : \mathbb{N} \to A$ cujo domínio é $\mathbb{N}$ e cuja imagem é todo A.

Nota. Por vezes conjuntos contáveis que não são finitos são chamados de conjuntos enumeráveis. Não há, infelizmente, unidade nessa nomenclatura mas emprega-la-emos aqui se vier a ser necessário.

Vamos agora provar alguns teoremas fundamentais sobre conjuntos contáveis (cuja importância, apesar da aparente simplicidade dos enunciados, não pode ser sub-estimada pois seu alcance estende-se por toda a Matemática, em particular, por muito do que veremos no restante do curso).

Precisamos da seguinte proposição:

Proposição 1.1 *Um conjunto é contável se e somente se for equivalente a um subconjunto de* $\mathbb{N}$. □

Prova. Por definição todo conjunto contável A (finito ou não) é equivalente a algum subconjunto de $\mathbb{N}$ (no pior dos casos ao próprio $\mathbb{N}$).

Provemos então a recíproca. Seja A equivalente a um subconjunto Z de $\mathbb{N}$. Se Z for finito A também o será e portanto contável. Suponhamos então que Z não é finito. Vamos construir uma função bijetora $F : \mathbb{N} \to Z$. A mesma é dada da seguinte forma

$$F(1) = \min Z,$$

$$F(n) = \min\{Z \setminus \{F(1), F(2), \ldots, F(n-1)\}\} \text{ para } n = 2, 3, \ldots.$$

É fácil ver que F é bijetora e que sua imagem é Z (faça isso). Assim, Z é enumerável e, portanto, A também o é. ■

Esta proposição tem uma conseqüência simples:

Proposição 1.2 *Se A é um conjunto contável e $B \subset A$ então B é contável.* □

Prova. Se A é contável e $B \subset A$ então B é equivalente a um subconjunto de $\mathbb{N}$ e, portanto, pela proposição anterior, B é contável. ■

Chegamos um importante teorema:

[4]J. von Neumann (1903-1957).

Teorema 1.3 *O produto cartesiano $\mathbb{N} \times \mathbb{N}$ é contável.* □

Prova. Seja a função $G : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$ dada por $G(a, b) = 2^a 3^b$. A imagem dessa função é um subconjunto próprio de $\mathbb{N}$ mas essa função é bijetora: a cada elemento z de sua imagem há um e somente um par (a, b) de números naturais tais que $2^a 3^b = z$ (por que?). Assim, fica provado pela proposição 1.1 que $\mathbb{N} \times \mathbb{N}$ é contável. ■

Note que, como $\mathbb{N} \times \mathbb{N}$ não é finito (por que?) é um conjunto enumerável.

Esse último teorema tem uma conseqüência de grande importância:

Teorema 1.4 *O conjunto $\mathbb{Q}_+$ dos números racionais positivos é um conjunto contável.* □

Prova. Todo racional positivo é da forma p/q onde p e $q \in \mathbb{N}$ são irredutíveis ou primos entre si (ou seja, não há "cancelamentos" que permitam escrever $p/q = a/b$ com $a < p$ e $b < q$). Assim, há uma correspondência um-a-um entre $\mathbb{Q}_+$ e o subconjunto de $\mathbb{N} \times \mathbb{N}$ formado por todos os pares (p, q) onde p e q são primos entre si. Como $\mathbb{N} \times \mathbb{N}$ é contável, a proposição 1.2 diz então que $\mathbb{Q}_+$ é também contável. ■

Exercício. Prove que o conjunto dos números inteiros $\mathbb{Z}$ e o conjunto dos números racionais $\mathbb{Q}$ são conjuntos contáveis.

Um fato também importante é que há conjuntos de números que não são contáveis. O exemplo mais importante é o dos números reais.

Teorema 1.5 *O conjunto dos números reais não é contável.* □

Prova. Para provar isso basta mostrar que há um subconjunto de $\mathbb{R}$ que não é contável. Considere o conjunto U de todos os números reais do intervalo $[0, 1]$ tais que em sua representação decimal apareçam apenas os dígitos 0 ou 1. Por exemplo, números como $0,001101$ ou $0,1$ ou 0 ou $0,1011$ ou $1/9 = 0,11111\ldots$ são elementos de U. Tecnicamente todo elemento u de U pode ser escrito da forma

$$u = \sum_{n=1}^{\infty} \frac{d_n(u)}{10^n}.$$

Assim, $d_n(u)$ é o n-ésimo dígito do número u de U. Note que dois elementos u e v de U são iguais se e somente se $d_n(u) = d_n(v)$ para todo n.

Vamos provar que U não é um conjunto contável. Para isso vamos supor o oposto, ou seja, que U é contável e veremos que essa hipótese leva a um absurdo.

Vamos supor que haja uma função bijetora $f : \mathbb{N} \to U$ cuja imagem é U. Considere o número real a definido por

$$a = \sum_{n=1}^{\infty} \frac{1 - d_n(f(n))}{10^n}.$$

Como $1 - d_n(f(n))$ é igual a 0 ou a 1 (por que?), segue obviamente que a é um elemento de U.

Entretanto, é fácil ver que a não faz parte da imagem da função f. Para ver isso note que se a fosse um elemento da imagem de f haveria um inteiro m tal que $f(m) = a$. Mas isso significa então que o m-ésimo dígito de a seria $d_m(a) = d_m(f(m))$. Mas pela definição do próprio a, o seu m-ésimo dígito é $1 - d_m(f(m))$. Assim, teríamos que $d_m(f(m)) = 1 - d_m(f(m))$ o que não é possível.

Concluímos então que a é um elemento de U mas não pode ser um elemento da imagem da função f. Isso é uma contradição, pois supomos justamente que a imagem da f era todo o conjunto U. Portanto, U não é contável e, assim, $\mathbb{R}$ também não o é. ■

Nota. É fácil ver que, em verdade, poderíamos substituir a base decimal, usada na representação do conjunto U acima, por qualquer base $b \in \mathbb{N}$ com $b > 2$. Ou seja, se considerarmos o conjunto U_b de todos os reais u do intervalo

$[0,\ 1]$ representáveis na base b, $b \in \mathbb{N}$, $b > 2$, da forma

$$u \;=\; \sum_{n=1}^{\infty} \frac{d_n(u)}{b^n}.$$

onde $d_n(u) \in \{0,\ 1\}$, então, repetindo o que fizemos acima, veríamos que U_b não é contável. Claramente $U = U_{10}$.

Nota. O caso da base binária $b = 2$ foi excluído da última nota pois nele não vale a unicidade da representação dos elementos de U_2 na forma

$$u \;=\; \sum_{n=1}^{\infty} \frac{d_n(u)}{2^n}.$$

onde $d_n(u) \in \{0,\ 1\}$. Para ver isso, faça o exercício seguinte.

Exercício. Mostre que na base binária $0, 1$ e $0, 01111111\ldots$ representam o mesmo número, a saber, o número $1/2$. Sugestão: usa a fórmula da progressão geométrica infinita para calcular quanto vale $0, 01111111\ldots$.

Nota. Os conjuntos U_b, $b > 2$, são exemplos de uma classe de conjuntos chamados de *conjuntos de Cantor*. Tornaremos a reencontrar tais conjuntos quando falarmos de Teoria da Medida.

Finalizamos com um outro teorema de grande importância:

Teorema 1.6 *Se C_i, $i \in \mathbb{N}$, são conjuntos contáveis então $C = \bigcup_{i \in \mathbb{N}} C_i$ também o é.* □

Prova. Se cada C_i é contável então para cada $i \in \mathbb{N}$ há uma função bijetora $g_i :\ \mathbb{N} \to C_i$ cuja imagem é C_i. Defina-se então a função $G :\ (\mathbb{N} \times \mathbb{N}) \to C$ dada por $G(a,\ b) = g_a(b)$. Esta função não é, em geral, bijetora, pois podem existir elementos comuns entre conjuntos C_i e C_j com $i \neq j$ e teríamos $g_i(m) = g_j(n)$ para algum n e m. Entretanto, a imagem de G é C.

Considere então em $\mathbb{N} \times \mathbb{N}$ a seguinte relação de equivalência: o par $(a,\ b)$ é equivalente ao par $(c,\ d)$ se e somente se $g_a(b) = g_c(d)$. O conjunto $\mathbb{N} \times \mathbb{N}$ pode ser então, como já observamos, escrito como a união disjunta de suas classes de equivalência pela relação acima. Construamos então um subconjunto K de $\mathbb{N} \times \mathbb{N}$ tomando-se um e somente um elemento de cada classe de equivalência escolhido arbitrariamente (usamos aqui o axioma da escolha para afirmar que tal construção é possível).

Defina então agora a função $H :\ K \to C$ dada por $H(a,\ b) = g_a(b)$ para $(a,\ b) \in K$. Pela própria construção do conjunto K essa função H é bijetora e sua imagem é C. Como K é um subconjunto de $\mathbb{N} \times \mathbb{N}$ que é contável, temos que K também o é e, portanto, C é contável. ■

• O Produto Cartesiano de uma Família Arbitrária de Conjuntos

Já discutimos o conceito de produto cartesiano de dois conjuntos A e B: $A \times B$. Agora que já discutimos o conceito de função podemos apresentar uma outra visão do conceito de produto cartesiano de dois conjuntos, a saber, podemos dizer que $A \times B$ é o conjunto de todas as funções $f :\ \{1,\ 2\} \to A \cup B$ tais que $f(1) \in A$ e $f(2) \in B$. A idéia é dizer que cada par ordenado $(a,\ b)$ com $a \in A$ e $b \in B$ é uma função onde o primeiro membro do par é a imagem de 1 (por ser o primeiro) e o segundo a imagem de 2 (por ser o segundo). Essa idéia permite definir produtos cartesianos de um número finito n de conjuntos $A_1,\ A_2, \ldots, A_n$ denotado por $A_1 \times A_2 \times \ldots \times A_n$ como sendo o conjunto de todas as funções $f :\ \{1, 2, \ldots, n\} \to \bigcup_{j=1}^{n} A_j$ satisfazendo $f(j) \in A_j$ para todo $j \in \{1, \ldots, n\}$. A função f tem, por assim dizer, o papel de ordenar os elementos de $\bigcup_{j=1}^{n} A_j$ tomando-se sucessivamente um elemento de cada A_i por vez. O produto cartesiano $A_1 \times A_2 \times \ldots \times A_n$ é assim entendido como o conjunto formado todas por ênuplas ordenadas $(a_1, \ldots, a_n)$ com $a_i \in A_i$.

Essa idéia pode ser generalizada ainda mais. Sejam I um conjunto (não necessariamente finito ou contável) e A_i, $i \in I$, conjuntos indexados por elementos de I. Definimos então o produto cartesiano da família de conjuntos $\{A_i,\ i \in I\}$, denotado por

$$\prod_{i \in I} A_i$$

como sendo o conjunto de todas as funções $f :\ I \to \bigcup_{j \in I} A_j$ tais que $f(x) \in A_x$ para todo $x \in I$.

Se por ventura todos os conjuntos A_i forem idênticos então denota-se o produto cartesiano acima por A^I. Assim, A^I denota o conjunto de todas as funções de I em A.

Desta forma $\mathbb{N} \times \mathbb{N}$ e $\mathbb{N}^{\{1,\ 2\}}$ são duas notações distintas para o mesmo objeto, que também é denotado simplesmente por $\mathbb{N}^2$, como se sabe. Genericamente $\mathbb{N}^d$ designa $\mathbb{N}^{\{1, \ldots, d\}}$ para $d \in \mathbb{N}$, $d > 0$.

É interessante notar que produtos cartesianos contáveis de conjuntos contáveis **<u>não</u>** são, geralmente, conjuntos contáveis. Considere como exemplo o produto cartesiano

$$K \;:=\; \prod_{i \in \mathbb{N}} \{0,\ 1\}; \;=\; \{0,\ 1\}^{\mathbb{N}}$$

que é denominado espaço de Cantor (1845-1918). Podemos mostrar que K não é contável. Cada elemento de K é uma função $d :\ \mathbb{N} \to \{0,\ 1\}$. Podemos assim associar univocamente a cada d o número real

$$\sum_{n=1}^{\infty} \frac{d(n)}{10^n}$$

que é um elemento do conjunto $U \subset \mathbb{R}$ definido acima. Por outro lado, todo elemento de U pode ser escrito assim para um único $d \in K$. Assim, K e U têm a mesma cardinalidade e, portanto, K não é contável pois U, como já vimos, não o é.

Exercício. Mostre que todos os conjuntos U_b, definidos acima, com $b > 2$, tem a mesma cardinalidade de K (e, portanto, a mesma cardinalidade entre si).

• Ordem Lexicográfica [5]

É possível estabelecer uma relação de ordem <u>total</u> em $\mathbb{R}^2$ da seguinte forma: dizemos que $(x_1,\ x_2) <_L (y_1,\ y_2)$ se $x_1 < y_1$ ou se $x_1 = y_1$ e $x_2 < y_2$. Essa relação de ordem é denominada *relação de ordem lexicográfica* de $\mathbb{R}^2$.

Essa definição pode ser facilmente generalizada. Seja X um conjunto totalmente ordenado por uma relação de ordem total $\prec_X$. Então X^n pode ser totalmente ordenado dizendo-se $(x_1,\ \ldots,\ x_n) <_L (y_1,\ \ldots,\ y_n)$ se houver um j, $j \in \{1,\ \ldots,\ n\}$, tal que $x_i = y_i$ para todo $i < j$ e $x_j \prec_X y_j$.

Seja X um conjunto totalmente ordenado por uma relação de ordem total $\prec_X$ e seja Seja $\mathfrak{X} = \bigcup_{n=1}^{\infty} X^n$. Podemos estabelecer em $\mathfrak{X}$ uma ordem total $\prec_{\mathfrak{X}}$, também denominada lexicográfica, da seguinte maneira. Sejam $m,\ n \in \mathbb{N}$ e $p = \min\{m,\ n\}$. Então dizemos $(x_1,\ \ldots,\ x_m) \prec_{\mathfrak{X}} (y_1,\ \ldots,\ y_n)$ se $(x_1,\ \ldots,\ x_p) <_L (y_1,\ \ldots,\ y_p)$ no sentido dado no parágrafo anterior, ou se $(x_1,\ \ldots,\ x_p) = (y_1,\ \ldots,\ y_p)$, mas $m < n$.

Exercício. Por que essas relações de ordem são denominadas "lexicográficas"? Pense na maneira como palavras (de tamanho arbitrário!) são ordenados em um dicionário.

Podemos ainda estender a definição de ordem lexicográfica. Seja X um conjunto totalmente ordenado por uma relação de ordem total $\preceq_X$ e seja Y um conjunto totalmente ordenado por uma relação de ordem total $\preceq_Y$. Então X^Y pode ser totalmente ordenado dizendo-se $X^Y \ni x <_L y \in X^y$ se houver um $j \in Y$, tal que $x(i) = y(i)$ para todo $i \prec_Y j$ e $x(j) \prec_X y(j)$.

Exemplo. Sejam $f,\ g$, duas funções de $\mathbb{R}$ em $\mathbb{R}$. Dizemos que $f <_L g$ se existir $y \in \mathbb{R}$ tal que $f(x) = g(x)$ para todo $x < y$ mas $f(y) < g(y)$. Lembrando que o conjunto de todas as funções de $\mathbb{R}$ em $\mathbb{R}$ é $\mathbb{R}^{\mathbb{R}}$, vê-se que essa definição coincide com a dada acima.

• A União Disjunta de uma Família Arbitrária de Conjuntos [6]

Sejam, como acima, um conjunto I (não necessariamente finito ou contável) e A_i, $i \in I$, conjuntos indexados por elementos de I. Os conjuntos A_i podem eventualmente possuir elementos comuns, ou seja, podem haver elementos x que comparecem em vários conjuntos A_i. Porém, quando formamos a união usual dos conjuntos A_i, ou seja, $\bigcup_{i \in I} A_i$, cada elemento x comparece apenas uma vez, mesmo que pertença a vários A_i's. Por vezes estamos interessados em formar um outro tipo de união de conjuntos onde essa possível multiplicidade de cada elemento x possa ser levada em conta. A definição abaixo é, para tal, das mais adequadas.

[5] Este tópico pode ser omitido numa primeira leitura.
[6] Este tópico pode ser omitido numa primeira leitura.

Definimos a *união disjunta* da família de conjuntos A_i como sendo o conjunto, denotado por $\bigsqcup_{i\in I} A_i$, dado pela união de todos os pares ordenados $(a,\ i)$ com $i \in I$, $a \in A_i$, ou seja,

$$\bigsqcup_{i\in I} A_i \ := \ \bigcup_{i\in I} \bigcup_{a\in A_i} (a,\ i).$$

• Números Reais Algébricos e Transcendentes [7]

Na reta real diz-se que um número x é um número *algébrico* se x for raiz de um polinômio do tipo

$$P(t) = a_0 + a_1 t + a_2 t^2 + \cdots + a_n t^n,$$

para algum $n \in \mathbb{N}$, onde os coeficientes $a_0, \ldots, a_n$ são números racionais. Um tal polinômio é dito ser um *polinômio racional*.

Todo número racional p/q é também algébrico pois é raiz do polinômio racional $p - qt$. Há também muitos números irracionais que são algébricos. Por exemplo, o número $\sqrt{2}$ é raiz do polinômio racional $-2 + t^2$ e, portanto, é algébrico. Os números reais que não são algébricos são chamados de *transcendentes*.

Exercício. Prove que o conjunto de todos os números algébricos da reta real é um conjunto contável. Use para tal o fato de que os racionais formam um conjunto contável.

O exercício anterior pode ser usado para concluir que existem números transcendentes (que não são raiz de nenhum polinômio racional) pois os reais, como sabemos, não são contáveis enquanto, segundo o exercício, os algébricos o são. Deve, portanto, haver uma coleção não-contável de números transcendentes na reta real.

Sabe-se, por exemplo, que π e e são transcendentes. Provar que π é algébrico é equivalente a resolver o célebre *problema da quadratura do círculo*, que consiste em achar um método através do qual, “apenas com régua e compasso” constrói-se um quadrado cuja área é igual a de um círculo de raio 1, o que é possível caso haja meios de se construir um segmento de reta cujo comprimento seja $\sqrt{\pi}$.

Esse problema clássico da geometria Euclidiana ficou em aberto por cerca de dois mil anos (!), tendo sido resolvido negativamente em 1882 por Lindemann[8] quando o mesmo provou justamente que π não é um número algébrico, concluindo assim a impossibilidade da construção proposta. Uma demonstração de que e é transcendente foi obtida pela primeira vez por Hermite[9] em 1873. Não há atualmente, infelizmente, uma demonstração simples do fato que π é transcendente, mas isso não é, por sorte, o caso da demonstração que e não é algébrico. Para uma prova simples de que e é transcendente vide, por exemplo, “Calculus” de M. Spivak.

Exercício. Seja $\mathbb{A}_0 = \mathbb{Q}$ e $\mathbb{A}_1$ o conjunto dos números algébricos, definidos como o conjunto de todos os zeros reais de polinômios com coeficientes racionais. Definimos $\mathbb{A}_2$ como o conjunto de todos os zeros reais de polinômios com coeficientes em $\mathbb{A}_1$. Sucessivamente, definimos $\mathbb{A}_n$, $n \geq 1$ como o conjunto de todos os zeros reais de polinômios com coeficientes em $\mathbb{A}_{n-1}$. Seja também $\mathbb{A} = \bigcup_{n=0}^{\infty} \mathbb{A}_n$. Mostre que todos os $\mathbb{A}_n$ e $\mathbb{A}$ são conjuntos contáveis e, portanto, subconjuntos próprios de $\mathbb{R}$.

• Ínfimo e Supremo de Famílias Contáveis de Conjuntos [10]

Seja I um conjunto arbitrário de índices e $\{A_i,\ i \in I\}$ uma coleção de conjuntos indexados por elementos de I. Chama-se por vezes o conjunto $\inf_{i\in I} A_i := \bigcap_{i\in I} A_i$ de ínfimo da coleção $\{A_i,\ i \in I\}$ e o conjunto $\sup_{i\in I} A_i := \bigcup_{i\in I} A_i$ de supremo da coleção $\{A_i,\ i \in I\}$.

Seja $\{A_, n \in \mathbb{N}\}$ uma coleção contável de subconjuntos de um conjunto X. Define-se um conjunto chamado de limite do ínfimo da coleção (denotado por $\underline{\lim} A_n$) como sendo o conjunto dado por

$$\underline{\lim} A_n \ := \ \bigcup_{n=1}^{\infty} \bigcap_{k=n}^{\infty} A_k.$$

[7] Este tópico pode ser omitido numa primeira leitura.
[8] C. L. F. Lindemann (1852-1939).
[9] C. Hermite (1822-1901).
[10] Este tópico pode ser omitido numa primeira leitura.

O chamado limite do supremo da coleção é o conjunto definido por

$$\overline{\lim} A_n \ := \ \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_k.$$

Exercício. Justifique a seguinte afirmativa: $\underline{\lim} A_n$ é o conjunto de todos os pontos x de X que pertencem a todos os conjuntos A_n exceto a no máximo um número finito deles. (Dizemos que x pertence a quase todos os A_n's).

Exercício. Justifique a seguinte afirmativa: $\overline{\lim} A_n$ é o conjunto de todos os pontos x de X que pertencem um número infinito de conjuntos A_n.

Chegamos a uma definição importante: dizemos que uma coleção contável de conjuntos $\{A_n,\ n \in \mathbb{N}\}$ converge a um conjunto A ($A = \lim_{n\to\infty} A_n$) se

$$\underline{\lim} A_n \ = \ \overline{\lim} A_n = A.$$

Exercício. Justifique a seguinte afirmativa: $\lim_{n\to\infty} A_n$ só existe se não há pontos $x \in X$ que, simultaneamente, pertençam a infinitos conjuntos A_n e não pertençam a infinitos conjuntos A_n.

Exercício Seja a família contável de subconjuntos de $\mathbb{R}$ dada por $A_n = [0,\ 10]$ se n for par e $A_n = [0,\ 5]$ se n for ímpar. Determine $\underline{\lim} A_n$ e $\overline{\lim} A_n$ e $\lim_{n\to\infty} A_n$ se este existir.

Exercício Seja a família contável de subconjuntos de $\mathbb{R}$ dada por $A_n = [0,\ 1]$ se n for par e $A_n = [2,\ 3]$ se n for ímpar. Determine $\underline{\lim} A_n$ e $\overline{\lim} A_n$ e $\lim_{n\to\infty} A_n$, se este existir.

Exercício. Seja a família contável de subconjuntos de $\mathbb{R}$ dada por

$$A_n \ = \ \left[-\frac{1}{n+1},\ 1 + \frac{1}{n+1}\right]$$

com $n \in \mathbb{N}$. Determine $\underline{\lim} A_n$, $\overline{\lim} A_n$ e $\lim_{n\to\infty} A_n$, se este existir.

Exercício. Seja a família contável de subconjuntos de $\mathbb{R}$ dada por

$$A_n \ = \ \left[\frac{1}{n+2},\ 1 - \frac{1}{n+2}\right]$$

com $n \in \mathbb{N}$. Determine $\underline{\lim} A_n$, $\overline{\lim} A_n$ e $\lim_{n\to\infty} A_n$, se este existir.

Exercício Crie seus próprios exemplos de famílias contáveis A_n de subconjuntos de $\mathbb{R}$ e estude seus $\underline{\lim} A_n$, $\overline{\lim} A_n$ e $\lim_{n\to\infty} A_n$, se este existir.

2 Noções Algébricas Básicas

Ainda atentos ao caráter introdutório apresentaremos aqui definições e exemplos das estruturas algébricas mais comuns.

• **Operações e Relações**

Sejam C e I dois conjuntos e consideremos o produto cartesiano C^I (o conceito de produto cartesiano arbitrário de conjuntos foi definido à página 11). Uma função $f : C^I \to C$ é por vezes dita ser uma *operação* sobre C. Se I é um conjunto finito, f é dita ser uma operação *finitária* sobre C.

Um conjunto $R \subset C^I$ é dito ser uma relação em C. Se I é um conjunto finito, R é dito ser uma relação *finitária* em C.

• **Funções Finitárias**

Sejam C e I dois conjuntos e consideremos funções $f : C^I \to C$. Se I é um conjunto finito $f : C^I \to C$ é dita ser uma função *finitária* sobre C ou operação *finitária* sobre C. Sem perda de generalidade consideraremos aqui funções finitárias do tipo $f : C^n \to C$ para algum $n \in \mathbb{N}$. Se f é uma função finitária para um dado n, f é dita ser uma função n-ária sobre C. Um exemplo de uma função não finitária seria uma função do tipo $f : C^{\mathbb{N}} \to C$ que a cada seqüência em C associa um elemento de C.

Funções 2-árias serão chamadas aqui de funções *binárias* e funções 1-árias são chamadas de funções *unárias*.

Por vezes iremos falar também de funções 0-árias sobre C, que consistem em funções $f : \{\emptyset\} \to C$. Uma tal função tem por imagem simplesmente um elemento fixo de C. Exemplos de funções 0-árias sobre $\mathbb{R}$ seriam $f(\emptyset) = 1$ ou $f(\emptyset) = 0$ ou $f(\emptyset) = \sqrt{2}$. Freqüentemente denotamos tais funções pelo elemento de C por ela associado. Nos três exemplos acima, poderíamos denotar as funções por 1, 0 ou $\sqrt{2}$, respectivamente.

• **Relações Finitárias**

Há uma nomenclatura análoga para o caso de relações. Sejam C e I dois conjuntos e consideremos relações $R \subset C^I$. Se I é um conjunto finito R é dita ser uma relação *finitária* sobre C. Sem perda de generalidade consideraremos aqui relações finitárias do tipo $R \subset C^n$ para algum $n \in \mathbb{N}$. Se R é uma relação finitária para um dado n, R é dita ser uma relação n-ária sobre C. Para o caso $n = 1$ as relações são também chamadas de unárias e para o caso $n = 2$ são ditas binárias.

Ainda um comentário sobre a nomenclatura.

Sejam C e I conjuntos e seja $\alpha : C^I \to C$ uma operação sobre o conjunto C. A cardinalidade de I é dita ser o *tipo* da operação α. Assim, uma função n-ária é também dita ser de tipo n. Analogamente, se $R \subset C^I$ é uma relação em C a cardinalidade de I é dita ser o *tipo* da relação R.

• **Comentário Sobre a Notação**

Antes de prosseguirmos, façamos uma observação sobre a notação que é costumeiramente adotada, especialmente quando se trata de funções binárias.

Dado um conjunto C e uma função binária denotada por um símbolo ϕ, a imagem de um par $(a,\ b) \in C^2$ é comummente denotada por $\phi(a,\ b)$. É muito prático, por vezes, usar uma outra notação e denotar $\phi(a,\ b)$ por $a\,\phi\,b$. Essa notação é denominada mesofixa. Um exemplo claro desse uso está na função soma, denotada pelo símbolo $+ : \mathbb{C}^2 \to \mathbb{C}$ de dois números complexos. Denotamos $+(z,\ w)$ por $z + w$. Outro exemplo está na função produto $\cdot : \mathbb{C}^2 \to \mathbb{C}$ de dois números complexos. Denotamos $\cdot(z,\ w)$ por $z \cdot w$.

Essa notação será usada adiante para outras funções binárias além das funções soma e produto de números ou matrizes.

Funções unárias também têm por vezes uma notação especial, freqüentemente do tipo exponencial. Tal é o caso da operação que associa a cada conjunto o seu complementar $A \mapsto A^c$. Ou ainda o caso da transposição de matrizes $M \mapsto M^T$, da conjugação de números complexos $z \mapsto z^*$ para o que usa-se também sabidamente a notação $z \mapsto \overline{z}$.

2.1 Semi-grupos, Monóides e Grupos

• **Semi-grupos**

Um semi-grupo é um conjunto G dotado de uma operação binária $G \times G \to G$ denotada por "$\cdot$" e denominada *produto* tal que a seguinte propriedade é satisfeita.

1. *Associatividade.* Para todos $a,\ b$ e $c \in G$ vale $(a \cdot b) \cdot c = a \cdot (b \cdot c)$.

• **Monóides**

Um monóide é um conjunto G dotado de uma operação binária $G \times G \to G$ denotada por "$\cdot$" e denominada *produto* tal que as seguintes propriedades são satisfeitas.

1. *Associatividade.* Para todos $a,\ b$ e $c \in G$ vale $(a \cdot b) \cdot c = a \cdot (b \cdot c)$.
2. *Elemento neutro.* Existe um (único) elemento $e \in G$, denominado elemento neutro, tal que $g \cdot e = e \cdot g = g$ para todo $g \in G$.

Observação A unicidade do elemento neutro é garantida pela observação que se houvesse e' tal que $g \cdot e' = e' \cdot g = g$ para todo $g \in G$ teríamos $e' = e' \cdot e = e$.

• **Grupos**

Uma das noções mais fundamentais de toda a Matemática é o conceito de *grupo*. Um grupo é um conjunto G dotado de uma operação binária $G \times G \to G$ denotada por "$\cdot$" e denominada *produto* e de uma operação unária $G \to G$ (bijetora) denominada inversa, denotada por "$\cdot^{-1}$", tais que as seguintes propriedades são satisfeitas.

1. *Associatividade.* Para todos $a,\ b$ e $c \in G$ vale $(a \cdot b) \cdot c = a \cdot (b \cdot c)$.
2. *Elemento neutro.* Existe um (único) elemento $e \in G$, denominado elemento neutro, tal que $g \cdot e = e \cdot g = g$ para todo $g \in G$.
3. *Inversa.* $g \cdot g^{-1} = g^{-1} \cdot g = e$ para todo $g \in G$.

Observações.

1. A unicidade do elemento neutro é garantida pela observação que se houvesse e' tal que $g \cdot e' = e' \cdot g = g$ para todo $g \in G$ teríamos $e' = e' \cdot e = e$.
2. Analogamente se estabelece a unicidade da inversa, pois se $g,\ h \in G$ são tais que $h \cdot g = g \cdot h = e$, teremos $g^{-1} = g^{-1} \cdot e = g^{-1} \cdot (g \cdot h) = (g^{-1} \cdot g) \cdot h = e \cdot h = h$.
3. Como $e \cdot e = e$ segue que $e^{-1} = e$.
4. Para todo $g \in G$ vale $(g^{-1})^{-1} = g$, pois $(g^{-1})^{-1} = (g^{-1})^{-1} \cdot e = (g^{-1})^{-1} \cdot (g^{-1} \cdot g) = ((g^{-1})^{-1} \cdot g^{-1}) \cdot g = e \cdot g = g$.

Um grupo é dito ser *comutativo* ou *Abeliano*[11] se $a \cdot b = b \cdot a$ para todos $a,\ b \in G$. Essa nomenclatura se aplica também a semi-grupos e monóides.

É evidente que todo grupo é um monóide e que todo monóide é um semi-grupo.

• **Exemplos**

O conjunto $S = \{1,\ 2,\ 3,\ \ldots\}$ é um semi-grupo em relação à operação de soma usual. O conjunto $M = \{0,\ 1,\ 2,\ 3,\ \ldots\}$ é um monóide em relação à operação de soma usual, sendo o elemento neutro $e = 0$. O conjunto $G = \mathbb{Z} = \{\ldots,\ -2,\ -1,\ 0,\ 1,\ 2,\ \ldots\}$ é um grupo em relação à operação de soma usual, sendo o elemento neutro $e = 0$ e a inversa $n^{-1} = -n$.

[11] Niels Henrik Abel (1802-1829).

$\mathbb{R}$ dotado da operação de multiplicação usual é um monóide onde o elemento neutro é o número 1. Não é um grupo, pois 0 não tem inversa multiplicativa. $\mathrm{Mat}(n)$, o conjunto das matrizes complexas $n \times n$ com o produto usual de matrizes é também apenas um monóide.

O conjunto $\{x \in \mathbb{R}, \ x > 0\}$ é um semi-grupo Abeliano em relação à operação de soma, mas não é um monóide.

O conjunto $\mathbb{R}_+ = \{x \in \mathbb{R}, \ x \geq 0\}$ é um monóide Abeliano em relação à operação de soma mas não um grupo.

Outro exemplo importante é o seguinte. Seja C um conjunto e tomemos $S = C^C$, o conjunto de todas as funções de C em C. Então S é um monóide com o produto formado pela *composição* de funções: $\circ$ e onde o elemento neutro é a função identidade $\mathrm{id}(s) = s$, $\forall s \in C$. O sub-conjunto de C^C formado pelas funções bijetoras é um grupo não-Abeliano, onde o elemento inverso de uma função $f : C \to C$ é a função inversa f^{-1}. Se S é um conjunto finito esse grupo é denominado *grupo de permutações* dos elementos de S.

- **$\mathbb{R}_+$ estendido** [12]

O conjunto $\mathbb{R}_+ = \{x \in \mathbb{R}, \ x \geq 0\}$ é um semi-grupo Abeliano em relação à operação de soma e em relação à operação de produto e vale ainda a propriedade distributiva $a(b + c) = ab + ac$. $\mathbb{R}_+$ é também, sabidamente, um conjunto linearmente ordenado pela relação de ordem usual.

Vamos abaixo descrever um outro conjunto linearmente ordenado que contém $\mathbb{R}_+$ e é também um semi-grupo Abeliano em relação à operação de soma e em relação à operação de produto e vale ainda a propriedade distributiva.

Definimos um conjunto, que denotaremos por $\mathscr{R}_+$, juntando a $\mathbb{R}_+$ um conjunto formado por um elemento, elemento esse que denotaremos provisoriamente por ω, com $\omega \notin \mathbb{R}_+$, para o qual certas relações algébricas serão definidas. Seja $\mathscr{R}_+ = \mathbb{R}_+ \cup \{\omega\}$ e definimos as operações de soma e produto em $\mathscr{R}_+$ da seguinte forma: se a e b são elementos de $\mathbb{R}_+$ suas soma e produto são definidos como usualmente. Fora isso, valem

1. $a + \omega = \omega + a = \omega$, para todo $a \in \mathbb{R}_+$.
2. $\omega + \omega = \omega$.
3. $a\omega = \omega a = \omega$, para todo $a \in \mathbb{R}_+$, $a \neq 0$.
4. $0\omega = \omega 0 = 0$.
5. $\omega\omega = \omega$.

Exercício. Verifique que $\mathscr{R}_+$ é um semi-grupo Abeliano em relação à operação de soma e em relação à operação de produto definidas acima e que vale ainda a propriedade distributiva.

$\mathscr{R}_+$ é linearmente ordenado tomando-se em $\mathbb{R}_+$ a relação de ordem usual e fixando-se $a < \omega$ para todo $a \in \mathbb{R}_+$.

É bastante claro que na definição abstrata acima o objeto representado pelo símbolo ω desempenha o papel formalmente desempenhado por um número infinito positivo. A construção das relações algébricas acima prescinde, porém, dessa noção, pois ω pode ser qualquer objeto (fora de $\mathbb{R}_+$).

Com um certo abuso de linguagem, é costume, substituir o símbolo ω pelo símbolo ∞, dando a entender que ω representa algo como um número infinito positivo. É comum também denotar-se $\mathscr{R}_+ = [0, \ \infty]$.

Exercício. Que problemas surgem quando se tenta estender a construção acima para o conjunto $\mathbb{R}$ de todos os reais?

2.2 Corpos e Espaços Vetoriais

- **Corpos**

Um *corpo* é um conjunto C dotado de duas operações binárias, denotadas por $+$ e $\cdot$, denominadas *soma* e *produto*, respectivamente, satisfazendo o seguinte:

1. A operação de soma tem as seguintes propriedades:
   (a) Comutatividade: $\alpha + \beta = \beta + \alpha$

[12] Este tópico pode ser omitido numa primeira leitura.

   (b) Associatividade: $\alpha + (\beta + \gamma) = (\alpha + \beta) + \gamma$
   (c) Elemento neutro: existe um elemento $0 \in C$, chamado de *zero*, tal que $\alpha + 0 = \alpha$ para todo $\alpha \in C$.
   (d) Para cada $\alpha \in C$ existe um único elemento denotado por β com a propriedade $\alpha + \beta = 0$. Esse elemento é mais comummente denotado por $-\alpha$.
2. A operação de produto tem as seguintes propriedades:
   (a) Comutatividade: $\alpha \cdot \beta = \beta \cdot \alpha$
   (b) Associatividade: $\alpha \cdot (\beta \cdot \gamma) = (\alpha \cdot \beta) \cdot \gamma$
   (c) Elemento neutro: existe um elemento $1 \in C$, chamado de *unidade*, tal que $\alpha \cdot 1 = \alpha$ para todo $\alpha \in C$.
   (d) Para cada $\alpha \in C$, $\alpha \neq 0$, existe um único elemento denotado por β com a propriedade $\alpha \cdot \beta = 1$. Esse elemento é mais comummente denotado por α^{-1}.
3. O produto é distributivo em relação à adição: $\alpha \cdot (\beta + \gamma) = \alpha \cdot \beta + \alpha \cdot \gamma$.

Os elementos de um corpo são por vezes denominados *escalares*.

Exemplos. É fácil verificar que $\mathbb{Q}$, $\mathbb{R}$ e $\mathbb{C}$ são corpos em relação às operações usuais de soma e produto. O conjunto das matrizes $n \times n$ para qualquer $n \geq 2$ com o produto usual de matrizes não é um corpo pois, entre outras razões, o produto não é comutativo.

Em um corpo C sempre vale que $\alpha \cdot 0 = 0$ para todo $\alpha \in C$. De fato, como $0 = 0 + 0$, segue que

$$\alpha \cdot 0 \ = \ \alpha \cdot (0 + 0) \ = \ \alpha \cdot 0 + \alpha \cdot 0 .$$

Somando-se a ambos os lados o elemento inverso $-\alpha \cdot 0$ teremos

$$\alpha \cdot 0 + (-\alpha \cdot 0) \ = \ \alpha \cdot 0 + \alpha \cdot 0 + (-\alpha \cdot 0),$$

ou seja,

$$0 \ = \ \alpha \cdot 0 + 0 \ = \ \alpha \cdot 0,$$

como queríamos provar. Pela comutatividade do produto vale também $0 \cdot \alpha = 0$ para todo $\alpha \in C$.

Vamos exibir de mais alguns exemplos menos triviais de corpos.

- **Os Corpos $\mathbb{Q}(\sqrt{p})$, com p Primo** [13]

Exercício. Mostre que o conjunto de todos os números reais da forma $a + b\sqrt{2}$, com a e b racionais, é um corpo.

O corpo do exemplo acima é denotado por $\mathbb{Q}(\sqrt{2})$.

Exercício. Seja p um número primo. Mostre que o conjunto de todos os números reais da forma $a + b\sqrt{p}$, com a e b racionais, é um corpo.

O corpo do exemplo acima é denotado por $\mathbb{Q}(\sqrt{p})$.

Exercício. Mostre que o conjunto de todos os números reais da forma $a + b\sqrt{2}$ com a e b inteiros **não** é um corpo.

- **Os Corpos $\mathbb{Z}_p$, com p Primo** [14]

O bem conhecido *algorítimo de Euclides* afirma que, dado $n \in \mathbb{N}$, $n > 0$, então todo número inteiro z pode ser escrito de maneira única na forma $z = qn + r$, onde $q \in \mathbb{Z}$ e $r \in \{0, \ 1, \ldots, \ n - 1\}$.

O numero r é denominado *resto da divisão de z por n* e é também denotado por $r = z \bmod n$.

Seja n um inteiro positivo maior ou igual a 2 e seja $\mathbb{Z}_n$ o conjunto $\{0, \ 1, \ \ldots, \ n - 1\}$. Vamos definir operações de soma e produto em $\mathbb{Z}_n$ da seguinte forma:

$$\alpha + \beta = [\alpha + \beta] \bmod n \qquad \text{e} \qquad \alpha \cdot \beta = [\alpha\beta] \bmod n .$$

Acima $[\alpha + \beta]$ e $[\alpha\beta]$ são a soma e o produto usuais em $\mathbb{Z}$.

Temos o seguinte teorema:

[13] Este tópico pode ser omitido numa primeira leitura.
[14] Este tópico pode ser omitido numa primeira leitura.

Teorema 2.1 *O conjunto $\mathbb{Z}_n$ é um corpo com as operações acima definidas se e somente se n é um número primo.* □

Prova. As operações de soma e produto definidas acima são automaticamente comutativas, associativas e distributivas (por que?). Fora isso sempre vale que $-\alpha = n - \alpha$ para todo $\alpha \in \mathbb{Z}_n$. Resta-nos estudar a existência de elementos inversos α^{-1}. Vamos supor que $\mathbb{Z}_n$ seja um corpo. Então, $a \in \{2, \ldots, n-1\}$ tem uma inversa em $\mathbb{Z}_n$, ou seja, um número $b \in \{1, \ldots, n-1\}$ tal que $a \cdot b = 1$. lembrando da definição de produto em $\mathbb{Z}_n$, isso significa que existe um inteiro r tal que $ab = rn + 1$. Mas isso implica

$$b - \frac{1}{a} = r\left(\frac{n}{a}\right).$$

Como o lado esquerdo não é um número inteiro, o lado direito também não pode ser. Isso diz então que n/a não pode ser inteiro para nenhum $a \in \{2, \ldots, n-1\}$, ou seja, n não tem divisores e é, portanto, um primo. Resta-nos mostrar que $\mathbb{Z}_p$ é efetivamente um corpo quando p é primo, o que agora se reduz a mostrar que para todo $a \in \mathbb{Z}_p$ existe um elemento inverso.

Para apresentar a demonstração, recordemos três conceitos da teoria de números. **1.** Sejam dois números inteiros f e g, dizemos que f *divide* g se $g/f \in \mathbb{Z}$. Se f *divide* g, denotamos esse fato por $f|g$. **2.** Sejam dois números inteiros f e g. O *máximo divisor comum* de f e g, denotado $\text{mdc}(f, g)$ é o maior inteiro m tal que $m|f$ e $m|g$. **3.** Dois números inteiros f e g são ditos ser *primos entre si* se $\text{mdc}(f, g) = 1$.

A demonstração da existência de inverso em $\mathbb{Z}_p$ será apresentada em partes. Vamos primeiro demonstrar a seguinte afirmativa.

Lema 2.2 *Se f e g são dois números inteiros quaisquer então existem inteiros k' e l' tais que*

$$mdc(f, g) = k'f + l'g.$$

□

Prova. Seja $m = \text{mdc}(f, g)$. Seja M o conjunto de todos os números positivos que sejam da forma $kf + lg$ com k e l inteiros. Seja m' o menor elemento de M. Note que como os elementos de M são positivos, esse menor elemento existe. Claramente

$$m' = k'f + l'g \tag{2.1}$$

para algum k' e l'. Como, por definição, $m|f$ e $m|g$, segue que $m|m'$, o que só é possível se

$$m' \geq m. \tag{2.2}$$

Vamos agora demonstrar por contradição que $m'|f$. Se isso não fosse verdade, existiriam (pelo algoritmo de Euclides) inteiros α e β com

$$0 < \beta < m' \tag{2.3}$$

tal que

$$f = \alpha m' + \beta.$$

Usando (2.1) isso diz que

$$\beta = f - \alpha(k'f + l'g) = (1 - \alpha k')f + (-\alpha l')g.$$

Mas, como $\beta > 0$ isso diz que $\beta \in M$. Logo, $\beta \geq m'$, contradizendo (2.3). Logo $m'|f$. De maneira totalmente análoga prova-se que $m'|g$. Portanto $m' \leq \text{mdc}(f, g) = m$. Lembrando que havíamos provado (2.2), segue que $m = m'$ e, portanto $m = k'f + l'g$, demonstrando o Lema. ■

Corolário 2.3 *Se f e g são dois números inteiros primos entre si então existem inteiros k' e l' tais que*

$$1 = k'f + l'g.$$

□

Prova. Pela definição, como f e g são dois números inteiros primos entre si segue que $\text{mdc}(f, g) = 1$. ■

Para finalmente demonstrarmos a existência de inverso em $\mathbb{Z}_p$, com p primo, seja $a \in \{1, \ldots, p-1\}$. É óbvio que a e p são primos entre si (por que?). Assim, pelo corolário, existem inteiros r e s com

$$1 = sa - rp.$$

Isso diz que $sa = rp + 1$. Logo, definindo $b \in \mathbb{Z}_p$ como sendo $b = s \bmod p$ teremos

$$ba = (s \bmod p)a = (rp + 1) \bmod p = 1,$$

ou seja, $b = a^{-1}$, completando a demonstração. ■

- **Característica de um Corpo** [15]

Seja C um corpo e 1 sua unidade. Para um número natural n definimos $n \cdot 1 = \underbrace{1 + \cdots + 1}_{n \text{ vezes}}$.

Define-se a *característica* de C como sendo o menor número natural não-nulo n tal que $n \cdot 1 = 0$. Se um tal número não existir, diz-se que o corpo tem característica zero.

Exemplos. $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$, $\mathbb{Q}(\sqrt{2})$ têm característica zero. $\mathbb{Z}_p$, p primo, tem característica p. Mostre isso.

Exercício. Mostre que a característica de um corpo é ou igual a zero ou é um número primo. Sugestão: Mostre primeiro que $(nm) \cdot 1 = (n \cdot 1)(m \cdot 1)$ para quaisquer números naturais n e m. Use então o fato que todo natural pode ser decomposto em um produto de fatores primos e use o fato que, em um corpo, se $a \cdot b = 0$ então ou a ou b ou ambos são zero (ou seja, todo corpo é um anel de integridade: não tem divisores de zero).

- **Espaços Vetoriais**

Um espaço vetorial V sobre um corpo K é um conjunto de elementos chamados *vetores* dotado de uma operação "+" : $V \times V \to V$ denominada *soma* e também de um *produto por escalares* "·" : $K \times V \to V$ com as seguintes propriedades:

1. A cada par $u, v \in V$ de vetores é associado um elemento $u + v \in V$, denominado *soma* de u e v, com as seguintes propriedades:

   (a) A soma é comutativa:
   $$u + v = v + u$$
   para todos $u, v \in V$,

   (b) A soma é associativa:
   $$u + (v + w) = (u + v) + w$$
   para todos $u, v, w \in V$,

   (c) Existe um único vetor denotado por 0, denominado *vetor nulo*, tal que
   $$u + 0 = u$$
   para todo $u \in V$,

   (d) A cada $u \in V$ existe associado um único vetor denotado por $-u$ tal que
   $$u + (-u) = 0.$$

2. A cada par $\alpha \in K$, $u \in V$ existe associado um vetor denotado por $\alpha \cdot u \in V$, denominado *produto de u por α*, de forma que

   (a) O produto por escalares é associativo:
   $$\alpha \cdot (\beta \cdot u) = (\alpha\beta) \cdot u,$$
   para todos $\alpha, \beta \in K$ e $u \in V$, onde $\alpha\beta$ é o produto de α por β em K,

[15] Este tópico pode ser omitido numa primeira leitura.

(b) $1 \cdot u = u$ para todo $u \in V$, onde 1 é a unidade de K,

(c) O produto por escalares é distributivo em relação à soma de vetores:

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

para todo $\alpha \in K$ e todos $u, v \in V$,

(d) O produto por escalares é distributivo em relação à soma de escalares:

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

para todos $\alpha, \beta \in K$ e todo $u \in V$.

Exercício. Mostre usando os postulados acima que $0 \cdot u = 0$ para todo $u \in V$, onde, permitindo-nos um certo abuso de linguagem, o 0 do lado esquerdo representa o zero do corpo K e o do lado direito o vetor nulo de V.

Nomenclatura. Os elementos de um corpo sobre os quais um espaço vetorial se constitui são freqüentemente denominados *escalares*.

Notação. É frequente omitir-se o símbolo "·" de produto por escalares quando nenhuma confusão é possível.

Anti-exemplo. Tomemos o conjunto dos reais com a operação de soma usual, um corpo $\mathbb{Z}_p$ com p primo e o produto $\mathbb{Z}_p \times \mathbb{R} \to \mathbb{R}$, $\alpha \cdot x$, $\alpha \in \mathbb{Z}_p$ e $x \in \mathbb{R}$ dada pelo produto usual em $\mathbb{R}$. Essa estrutura **não** forma um espaço vetorial. A regra distributiva

$$(\alpha + \beta) \cdot x = \alpha \cdot x + \beta \cdot x$$

não é satisfeita para todo $\alpha, \beta \in \mathbb{Z}_p$. Acima, $\alpha \cdot x$ é o produto usual em $\mathbb{R}$.

• **Os Espaços Vetoriais $\mathbb{R}^n$ e $\mathbb{C}^n$**

Trataremos agora de dois importantes exemplos de espaços vetoriais de dimensão (algébrica) finita, os espaços $\mathbb{R}^n$ e $\mathbb{C}^n$.

O conjunto $\mathbb{R}^n$ de todas as n-uplas ordenadas de números reais é naturalmente um espaço vetorial sobre o corpo dos reais com as operações

$$x(a_1, \ldots, a_n) + y(b_1, \ldots, b_n) = (xa_1 + yb_1, \ldots, xa_n + yb_n).$$

definidas para todo x e $y \in \mathbb{R}$ e todos os vetores $(a_1, \ldots, a_n)$, $(b_1, \ldots, b_n) \in \mathbb{R}^n$. O vetor nulo é, naturalmente, $(0, \ldots, 0)$.

É muito frequente, especialmente quando se fala sobre matrizes, representar os vetores de $\mathbb{R}^n$ em forma de uma coluna:

$$\begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}.$$

Em uma tal notação o vetor é dito ser um vetor-coluna.

O conjunto $\mathbb{C}^n$ de todas as n-uplas ordenadas de números complexos é naturalmente um espaço vetorial sobre o corpo dos complexos com as operações

$$x(a_1, \ldots, a_n) + y(b_1, \ldots, b_n) = (xa_1 + yb_1, \ldots, xa_n + yb_n).$$

definidas para todo x e $y \in \mathbb{C}$ e todos os vetores $(a_1, \ldots, a_n)$, $(b_1, \ldots, b_n) \in \mathbb{C}^n$. O vetor nulo é, naturalmente, $(0, \ldots, 0)$.

Como no caso de $\mathbb{R}^n$, e muito frequente, especialmente quando se fala sobre matrizes, representar os vetores de $\mathbb{C}^n$ em forma de um vetor-coluna:

$$\begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}.$$

2.2.1 Espaços Vetoriais de Dimensão Infinita. Exemplos

Vamos apresentar aqui alguns exemplos de espaços vetoriais de dimensão infinita sobre o corpo dos complexos (ou dos reais). Trataremos de duas classes de exemplos: espaços de seqüências e espaços de funções.

• **Espaços de Seqüências**

Seja $\mathbb{C}^{\mathbb{N}}$ o conjunto de todas as seqüências de números complexos.

$$\mathbb{C}^{\mathbb{N}} = \{(a_1, a_2, a_3, \ldots), \ a_i \in \mathbb{C}\}.$$

Uma seqüência de números complexos nada mais é, lembremos, que uma função de $\mathbb{N}$ em $\mathbb{C}$. Assim, o conjunto $\mathbb{C}^{\mathbb{N}}$ pode ser interpretado como o conjunto de todas as funções de $\mathbb{N}$ em $\mathbb{C}$.

$\mathbb{C}^{\mathbb{N}}$ tem naturalmente uma estrutura de espaço vetorial sobre o corpo dos complexos. Define-se a operação de soma de vetores por

$$(a_1, a_2, a_3, \ldots) + (b_1, b_2, b_3, \ldots) = (a_1 + b_1, a_2 + b_2, a_3 + b_3, \ldots),$$

O produto por escalares (elementos do corpo $\mathbb{C}$) é dado por

$$\lambda(a_1, a_2, a_3, \ldots) = (\lambda a_1, \lambda a_2, \lambda a_3, \ldots), \quad \lambda \in \mathbb{C}.$$

O vetor nulo é a seqüência identicamente nula

$$0 = (0, 0, 0, \ldots).$$

$\mathbb{C}^{\mathbb{N}}$ possui vários subconjuntos próprios que também são espaços vetoriais sobre o corpo dos complexos com as mesmas operações de soma e produto por escalares. Vamos abaixo apresentar alguns deles. Os exemplos escolhidos são todos importantes, especialmente os espaços ℓ_p.

O primeiro exemplo é o conjunto de seqüências definido por

$$d = \left\{a \in \mathbb{C}^{\mathbb{N}} \,\middle|\, a_n = 0, \text{ exceto para um número finito de } n\text{'s}\right\}.$$

Os elementos de d são seqüências com a propriedade que $a_n = 0$ para todo n grande o suficiente.

Exercício (fácil). Mostre que o conjunto de seqüências d acima é um espaço vetorial sobre o corpo $\mathbb{C}$ com as mesmas operações definidas em $\mathbb{C}^{\mathbb{N}}$, ou seja, mostre que d é um sub-espaço de $\mathbb{C}^{\mathbb{N}}$.

Outro exemplo é o conjunto de seqüências com a propriedade de decairem a zero quando $n \to \infty$ mais rapidamente que qualquer potência de n:

$$s = \left\{a \in \mathbb{C}^{\mathbb{N}} \,\middle|\, \lim_{n\to\infty} n^k a_n = 0 \text{ para todo inteiro positivo } k\right\}.$$

A seqüência $a_n = e^{-n}$ é um elemento de s (por que?). A seqüência $a_n = e^{-n^2}$ é um elemento de s (por que?). A seqüência $a_n = n^{-2}$ <u>não</u> é um elemento de s (por que?). Para qualquer $l \in \mathbb{Z}$ a seqüência $a_n = n^{-l}$ <u>não</u> é um elemento de s (por que?).

Exercício (fácil). Mostre que o conjunto de seqüências s acima é um espaço vetorial sobre o corpo $\mathbb{C}$ com as mesmas operações definidas em $\mathbb{C}^{\mathbb{N}}$, ou seja, mostre que s é um sub-espaço de $\mathbb{C}^{\mathbb{N}}$.

O próximo exemplo é muito importante. Trata-se das seqüências p-somáveis (com $p \geq 1$). Para cada $p \geq 1$ defina-se

$$\ell_p = \left\{a \in \mathbb{C}^{\mathbb{N}} \,\middle|\, \sum_{n=1}^{\infty} |a_n|^p < \infty\right\}, \quad p \geq 1. \tag{2.4}$$

Esses exemplos são importantes pois os conjuntos ℓ_p, $p \geq 1$, são exemplos-protótipo de espaços de Banach. O espaço ℓ_2 é também um espaço de Hilbert.

Vamos mostrar que ℓ_1 é de fato um espaço vetorial. Sejam $a = (a_1, a_2, a_3, \ldots)$ e $b = (b_1, b_2, b_3, \ldots)$ elementos de ℓ_1, ou seja, tais que

$$\sum_{n=1}^{\infty} |a_n| < \infty \quad \text{e} \quad \sum_{n=1}^{\infty} |b_n| < \infty. \tag{2.5}$$

Desejamos provar que $\sum_{n=1}^{\infty} |a_n+b_n| < \infty$. Isso segue facilmente da observação que, para todo n, $|a_n+b_n| \leq |a_n|+|b_n|$.

Exercício. Complete os detalhes.

Mostrar que os conjuntos ℓ_p com $p > 1$ são espaços vetoriais é um pouco mais difícil. Isso equivale a provar que se $a = (a_1,\ a_2,\ a_3,\ \ldots)$ e $b = (b_1,\ b_2,\ b_3,\ \ldots)$ elementos de ℓ_p, ou seja, tais que

$$\sum_{n=1}^{\infty} |a_n|^p < \infty \quad \text{e} \quad \sum_{n=1}^{\infty} |b_n|^p < \infty \tag{2.6}$$

então $\sum_{n=1}^{\infty} |a_n+b_n|^p < \infty$. Esse fato é verdadeiro, mas sua demonstração é mais longa que o caso $p = 1$ visto acima. A prova é conseqüência das desigualdades de Hölder e de Minkowski, que serão demonstradas na seção 2.2.2 à página 26. Convidamos o leitor a estudar aquelas demonstrações.

Mais abaixo (Proposição 2.5, página 23) mostraremos que, como conjuntos, $\ell_p \subset \ell_{p'}$ se $1 \leq p \leq p'$.

Um exemplo ilustrativo da distinção entre os vários espaços ℓ_p é o da chamada seqüência harmônica: $a_n = \frac{1}{n}$. Sabidamente essa seqüência não está em ℓ_1 pois $\sum_{n=1}^{\infty} \frac{1}{n} = \infty$. No entanto, $\sum_{n=1}^{\infty} \frac{1}{n^p} < \infty$ para todo $p > 1$, mostrando que a seqüência harmônica pertence a cada espaço ℓ_p com $p > 1$.

Exercício. Construa um exemplo de uma seqüência que não pertence a ℓ_2 mas pertence a cada ℓ_p com $p > 2$.

Mais um exemplo é dado pelo conjunto de seqüências que decaem a zero no infinito:

$$c_0 \;=\; \left\{ a \in \mathbb{C}^{\mathbb{N}} \;\middle|\; \lim_{n\to\infty} a_n = 0 \right\}.$$

Exercício (fácil). Mostre que o conjunto de seqüências c_0 acima é um espaço vetorial sobre o corpo $\mathbb{C}$ com as mesmas operações definidas em $\mathbb{C}^{\mathbb{N}}$, ou seja, mostre que c_0 é um sub-espaço de $\mathbb{C}^{\mathbb{N}}$.

Nosso último exemplo é o das seqüências limitadas.

$$\ell_\infty \;=\; \left\{ a \in \mathbb{C}^{\mathbb{N}} \;\middle|\; \sup_{n\in\mathbb{N}} |a_n| < \infty \right\}.$$

Uma seqüência a_n pertence a ℓ_∞ se existir número $M \geq 0$ tal que $|a_n| \leq M$ para todo n. (Nota. um tal número M não precisa ser o mesmo para seqüências distintas em ℓ_∞).

Exercício (fácil). Mostre que o conjunto de seqüências ℓ_∞ acima é um espaço vetorial sobre o corpo $\mathbb{C}$ com as mesmas operações definidas em $\mathbb{C}^{\mathbb{N}}$, ou seja, mostre que ℓ_∞ é um sub-espaço de $\mathbb{C}^{\mathbb{N}}$.

Contra-Exemplo. O seguinte subconjunto de $\mathbb{C}^{\mathbb{N}}$ não é um espaço vetorial (com as operações definidas acima).

$$\left\{ a \in \mathbb{C}^{\mathbb{N}} \;\middle|\; a_n = 0 \text{ para um conjunto infinito de } n\text{'s} \right\}.$$

Exercício. Justifique. *Sugestão:* o que acontece quando se soma a seqüência $a = (2,\ 0,\ 2,\ 0,\ \ldots)$, ou seja, $a_n = 1 - (-1)^n$ com a seqüência $b = (0,\ 2,\ 0,\ 2,\ \ldots)$, ou seja $b_n = 1 + (-1)^n$?

Os espaços vetoriais d, s, ℓ_p, c_0, ℓ_∞ então relacionados entre si por relações de inclusão. A saber, tem-se as duas seguintes proposições.

Proposição 2.4 *Com as definições acima, tem-se* $d \subset s \subset \ell_p \subset c_0 \subset \ell_\infty \subset \mathbb{C}^{\mathbb{N}}$. □

Prova. *Exercício.*

Proposição 2.5 *Com as definições acima, tem-se* $\ell_p \subset \ell_{p'}$ *caso* $1 \leq p \leq p'$. □

Prova. Seja $a = (a_1,\ a_2,\ a_3,\ \ldots)$ um elemento de ℓ_p, ou seja, tal que

$$\sum_{n=1}^{\infty} |a_n|^p < \infty. \tag{2.7}$$

Queremos provar que também vale

$$\sum_{n=1}^{\infty} |a_n|^{p'} < \infty.$$

para todo $p' > p$. Se (2.7) vale, então $a_n \to 0$ para $n \to \infty$. Assim, existe algum N tal que $|a_n| < 1$ para todo $n > N$. Para esses n's tem-se obviamente que

$$|a_n|^{p'} \;\leq\; |a_n|^p$$

para todo $p' \geq p \geq 1$ (certo?). Portanto,

$$\sum_{n=1}^{\infty} |a_n|^{p'} \;=\; \sum_{n=1}^{N} |a_n|^{p'} + \sum_{n=N+1}^{\infty} |a_n|^{p'} \;\leq\; \sum_{n=1}^{N} |a_n|^{p'} + \sum_{n=N+1}^{\infty} |a_n|^{p} \;<\; \infty,$$

já que $\sum_{n=1}^{N} |a_n|^{p'}$ é uma soma finita e $\sum_{n=N+1}^{\infty} |a_n|^p \leq \sum_{n=1}^{\infty} |a_n|^p < \infty$ por hipótese. Isso termina a prova. ■

- **Espaços de Funções**

É bastante claro que o espaço de todas as funções de $\mathbb{R}$ em $\mathbb{C}$, que denotamos por $\mathbb{C}^{\mathbb{R}}$, é um espaço vetorial sobre o corpo dos complexos. As operações de soma e produto por escalares são familiares: se f e g são duas funções de $\mathbb{R}$ em $\mathbb{C}$ sua soma $f+g$ é a função que a cada $x \in \mathbb{R}$ associa $(f+g)(x) = f(x)+g(x)$. Analogamente, se $\lambda \in \mathbb{C}$, $(\lambda f)(x) = \lambda f(x)$. O vetor nulo é a função identicamente nula.

$\mathbb{C}^{\mathbb{R}}$ possui muitos sub-espaços de interesse. Vamos listar alguns exemplos.

Um subconjunto de $\mathbb{C}^{\mathbb{R}}$ de interesse é formado pelo conjunto de todas as funções limitadas:

$$\mathscr{B}(\mathbb{R}) \;=\; \{ f : \mathbb{R} \to \mathbb{C} |\;\; \text{existe } M \geq 0 \text{ tal que } |f(x)| \leq M \text{ para todo } x \in \mathbb{R} \}.$$

Exercício. Mostre que $\mathscr{B}(\mathbb{R})$ é um espaço vetorial sobre $\mathbb{C}$ com as operações usuais.

$$C(\mathbb{R}) \;=\; \{ f : \mathbb{R} \to \mathbb{C} |\; f \text{ é contínua} \}.$$

Exercício. Mostre que $C(\mathbb{R})$ é um espaço vetorial sobre $\mathbb{C}$ com as operações usuais.

Um subconjunto de $C(\mathbb{R})$ é $\mathcal{P}(\mathbb{R})$, o conjunto de todos os polinômios de uma variável real com coeficientes complexos: $P_n(t) \in \mathcal{P}(\mathbb{R})$,

$$P_n(t) = a_n t^n + \cdots + a_1 t + a_0$$

com $t \in \mathbb{R}$, $a_i \in \mathbb{C}$, é dito ser um polinômio de grau n se $a_n \neq 0$.

É claro que $\mathcal{P}(\mathbb{R})$ é um espaço vetorial sobre $\mathbb{C}$ com as operações usuais. Note se que, para um dado n, o conjunto de todos os polinômios de grau n não é um espaço vetorial, entre outras razões por não conter o vetor nulo. No entanto, o conjunto de todos os polinômios de grau igual a ou menor que n é um espaço vetorial de dimensão finita.

Podemos ainda definir os espaços $C^n(\mathbb{R})$, de todas as funções de $\mathbb{R}$ em $\mathbb{C}$ que são contínuas e têm suas n primeiras derivadas contínuas.

Exercício. Mostre que todos os $C^n(\mathbb{R})$ são espaços vetoriais sobre $\mathbb{C}$ com as operações usuais.

Temos também o espaços das funções infinitamente diferenciáveis:

$$C^\infty(\mathbb{R}) \;=\; \{ f : \mathbb{R} \to \mathbb{C} |\; f \text{ é contínua e infinitamente diferenciável} \}.$$

Todo polinômio pertence a $C^\infty(\mathbb{R})$, assim como as funções seno e co-seno, exponencial etc.

Claramente $\mathcal{P}(\mathbb{R}) \subset C^\infty(\mathbb{R}) \subset \cdots \subset C^3(\mathbb{R}) \subset C^2(\mathbb{R}) \subset C^1(\mathbb{R}) \subset C(\mathbb{R})$.

Exercício. Mostre que $C^\infty(\mathbb{R})$ é um espaço vetorial sobre $\mathbb{C}$ com as operações usuais.

Exercício. Mostre que a função

$$s(x) \;=\; \begin{cases} e^{-\frac{1}{x^2}}, & x \neq 0 \\ 0, & x = 0 \end{cases}$$

é contínua e infinitamente diferenciável, ou seja, é um elemento de $C^\infty(\mathbb{R})$. *Sugestão.* Estude o que acontece em $x = 0$ com a função e suas derivadas. Mostre que essa função e todas as suas derivadas se anulam em $x = 0$.

Também importante é o espaço $C_0(\mathbb{R})$ de todas as funções de $\mathbb{R}$ em $\mathbb{C}$ contínuas com suporte[16] compacto[17].

$C_0^n(\mathbb{R})$ denota o conjunto de todas as funções de $C^n(\mathbb{R})$ que têm suporte compacto. $C_0^\infty(\mathbb{R})$ denota o conjunto de todas as funções de $C^\infty(\mathbb{R})$ que têm suporte compacto.

Exercício. Mostre que todos esses conjuntos são também espaços vetoriais.

Exemplo. A função

$$h(x) \;=\; \begin{cases} e^{-\frac{1}{x^2-4}}, & x \in (-2,\ 2) \\ 0, & \text{de outra forma} \end{cases}$$

é um elemento de $C_0^\infty(\mathbb{R})$.

Exercício. Mostre que a função h acima é de fato infinitamente diferenciável. O que acontece com h e suas derivadas em $x = \pm 2$? Esboce um gráfico de h.

Outro exemplo importante é o *Espaço de Schwartz*[18], que desempenha um papel importante na teoria das distribuições e na teoria das transformadas de Fourier:

$$\mathscr{S}(\mathbb{R}) \;=\; \left\{ f : \mathbb{R} \to \mathbb{C} \;\middle|\; f \in C^\infty(\mathbb{R}) \text{ e } \lim_{x\to\pm\infty} |x|^m f^{(n)}(x) = 0 \text{ para todos inteiros positivos ou nulos } m \text{ e } n \right\}.$$

As funções de $\mathscr{S}(\mathbb{R})$ são infinitamente diferenciáveis e decaem a zero quando $x \to \pm\infty$ mais rapidamente que qualquer potência de x e todas as suas derivadas têm essa propriedade. Exemplos de funções de $\mathscr{S}$ são $f(x) = e^{-x^2}$, $f(x) = e^{-x^4}$ etc. Uma função como $g(x) = \frac{1}{1+x^2}$ é infinitamente diferenciável mas não pertence a $\mathscr{S}(\mathbb{R})$ pois não decai a zero com $x \to \pm\infty$ mais rapidamente que <u>qualquer</u> potência de x. Por exemplo, $\lim_{x\to\infty} x^3 g(x) = \infty$.

Exercício. Mostre que $\mathscr{S}(\mathbb{R})$ é um espaço vetorial sobre $\mathbb{C}$ com as operações usuais.

Exercício. Mostre que $C_0^\infty(\mathbb{R}) \subset \mathscr{S}(\mathbb{R}) \subset C^\infty(\mathbb{R})$, que $C_0^\infty(\mathbb{R}) \subset \mathscr{S}(\mathbb{R}) \subset \mathscr{B}(\mathbb{R})$ e que $C_0^\infty(\mathbb{R}) \subset \cdots \subset C_0^3(\mathbb{R}) \subset C_0^2(\mathbb{R}) \subset C_0^1(\mathbb{R}) \subset C_0(\mathbb{R}) \subset \mathscr{B}(\mathbb{R})$.

Quase todos os exemplos acima tratam de espaços de funções definidos por propriedades das funções e suas derivadas. Os exemplos que veremos agora tratam de espaços de funções definidos por propriedades de suas integrais.

Para cada $p \geq 1$ definimos os espaços

$$\mathscr{L}_p(\mathbb{R},\ dx) \;=\; \left\{ f : \mathbb{R} \to \mathbb{C} \;\middle|\; \int_{-\infty}^{\infty} |f(x)|^p\, dx < \infty \right\}.$$

Na definição é preciso requerer, em verdade, que as funções sejam mensuráveis no sentido de Lebesgue, mas não será necessário discutir isso aqui.

É fácil provar que o conjunto $\mathscr{L}_1(\mathbb{R},\ dx)$ é um espaço vetorial (*Exercício!*). Para $p > 1$ a demonstração é um pouco mais trabalhosa, como no caso dos espaços de seqüências ℓ_p, e requer demonstrar generalizações para integrais das desigualdades de Hölder e Minkowski.

Exercício. Estude as demonstrações das desigualdades de Hölder e Minkowski para seqüências apresentadas na Seção 2.2.2 e generalize-as para integrais.

Comentamos que os espaços $\mathscr{L}_2(\mathbb{R},\ dx)$ desempenham um papel importantíssimo na Mecânica Quântica.

Ao contrário do que ocorre com os espaços de seqüências ℓ_p, não há entre os espaços $\mathscr{L}_p(\mathbb{R},\ dx)$ relações de inclusão óbvias. Por exemplo, a função

$$f_1(x) \;=\; \begin{cases} \dfrac{1}{x}, & x \geq 1 \\[2mm] 0, & \text{de outra forma} \end{cases}$$

pertence a $\mathscr{L}_2(\mathbb{R},\ dx)$ mas não a $\mathscr{L}_1(\mathbb{R},\ dx)$.

Exercício. Justifique essas afirmativas sobre f_1.

[16] O suporte de uma função f, denotado por $\operatorname{supp}(f)$ é o fecho do conjunto de todos os pontos de seu domínio onde a função é não-nula.
[17] Na reta real um conjunto é compacto se for fechado e limitado.
[18] Laurent Schwartz (1915-).

Por outro lado, a função

$$f_2(x) \;=\; \begin{cases} \dfrac{1}{\sqrt{x}}, & 0 < x \leq 1 \\[2mm] 0, & \text{de outra forma} \end{cases}$$

pertence a $\mathscr{L}_1(\mathbb{R},\ dx)$ mas não a $\mathscr{L}_2(\mathbb{R},\ dx)$.

Exercício. Justifique essas afirmativas sobre f_2.

Assim, não é verdade que $\mathscr{L}_1(\mathbb{R},\ dx) \subset \mathscr{L}_2(\mathbb{R},\ dx)$ nem que $\mathscr{L}_2(\mathbb{R},\ dx) \subset \mathscr{L}_1(\mathbb{R},\ dx)$.

O exemplo da função f_2 mostra também que as funções de $\mathscr{L}_1(\mathbb{R},\ dx)$ não são necessáriamente limitadas. O mesmo vale para os demais espaços $\mathscr{L}_p(\mathbb{R},\ dx)$, como mostra o seguinte exercício:

Exercício. Seja $p \geq 1$ fixo. Mostre que

$$f_3(x) \;=\; \begin{cases} x^{-\frac{1}{2p}}, & 0 < x \leq 1 \\[2mm] 0, & \text{de outra forma} \end{cases}$$

pertence a $\mathscr{L}_p(\mathbb{R},\ dx)$.

Os espaços $\mathscr{L}_p(\mathbb{R},\ dx)$ acima têm uma generalização. Seja $r : \mathbb{R} \to \mathbb{R}$ uma função positiva (e mensurável por Lebesgue). Define-se

$$\mathscr{L}_p(\mathbb{R},\ r(x)dx) \;=\; \left\{ f : \mathbb{R} \to \mathbb{C} \;\middle|\; \int_{-\infty}^{\infty} |f(x)|^p\, r(x)dx < \infty \right\}.$$

Para funções definidas em um intervalo $[a,\ b] \subset \mathbb{R}$ tem-se também

$$\mathscr{L}_p([a,\ b],\ r(x)dx) \;=\; \left\{ f : \mathbb{R} \to \mathbb{C} \;\middle|\; \int_a^b |f(x)|^p\, r(x)dx < \infty \right\}.$$

Encontraremos esses espaços quando tratarmos do problema de Sturm-Liouville.

Por fim comentamos que há versões de todos os espaços aqui tratados para funções de várias variáveis.

2.2.2 As Desigualdades de Hölder e de Minkowski para Seqüências

• **<u>A Desigualdade de Hölder (caso discreto). Seqüências ℓ_p</u>**

Vamos aqui demonstrar duas desigualdades importantes as quais são conhecidas como desigualdade de Hölder[19] e de Minkowski[20].

Antes de tratarmos dessas desigualdades, vamos demonstrar um resultado preparatório que usaremos adiante.

Sejam a e b dois números reais, ambos maiores ou iguais a zero e sejam p e q ambos tais que $1 < p < \infty$ e $1 < q < \infty$, mas tais que

$$\frac{1}{p} + \frac{1}{q} = 1.$$

Vamos então mostrar que para todo $a,\ b \geq 0$

$$a^{1/p}\, b^{1/q} \;\;\leq\;\; \frac{a}{p} + \frac{b}{q},$$

sendo que a igualdade só é válida caso $a = b$.

Em primeiro lugar note que se $a = 0$ e/ou $b = 0$ a desigualdade acima é trivialmente satisfeita pois o lado esquerdo é sempre zero, enquanto que o lado direito é sempre maior ou igual a zero.

[19] Otto L. Hölder (1859-1937).
[20] Hermann Minkowski (1864-1909).

Vamos estão supor que a e b são ambos não nulos. Tudo o que queremos é provar que $-a^{1/p}b^{1/q} + \frac{a}{p} + \frac{b}{q}$ é sempre maior ou igual a zero. Podemos escrever a última expressão como

$$b\left(-t^{\alpha} + \alpha t + \frac{1}{q}\right)$$

onde $\alpha = 1/p$ e $t = a/b$. Como $1 < p < \infty$, temos que $0 < \alpha < 1$ enquanto que $t \geq 0$. Note-se que a função

$$f(x) \;=\; -x^{\alpha} + \alpha x + \frac{1}{q},$$

é contínua para $x \in [0,\ \infty)$ e que, para $x > 0$, tem-se

$$f'(x) \;=\; \alpha\left(1 - x^{\alpha-1}\right)$$

e

$$f''(x) \;=\; \alpha(1-\alpha)x^{\alpha-2} > 0.$$

Assim, $f(x)$ tem um único mínimo local em $x = 1$, onde $f(1) = 0$ (verifique). Fora isso, $f(0) = \frac{1}{q} > 0$ e $\lim_{x\to\infty} f(x) = +\infty$. Desses fatos concluímos facilmente que $f(x) \geq 0$ para todo $x \geq 0$, a igualdade só se dando caso $x = 1$. Isso fecha o que queríamos provar.

Vamos agora então enunciar e provar a desigualdade de Hölder. Sejam $x_1, \ldots, x_n$ e $y_1, \ldots, y_n$ duas seqüências finitas arbitrárias de números complexos ($n \in \mathbb{N}$). A desigualdade de Hölder afirma que

$$\sum_{i=1}^{n} |x_i||y_i| \;\leq\; \left(\sum_{i=1}^{n} |x_i|^p\right)^{1/p} \left(\sum_{i=1}^{n} |y_i|^q\right)^{1/q},$$

para quaisquer p, q com $1 < p < \infty$ e $1 < q < \infty$ e tais que $\frac{1}{p} + \frac{1}{q} = 1$.

Vamos provar a desigualdade de Hölder. Em primeiro lugar, note que a desigualdade é trivialmente verdadeira caso todos os x_i ou todos os y_i sejam nulos, pois nesse caso tanto o lado direito quanto o lado esquerdo da desigualdade são iguais a zero.

Vamos então considerar o caso em que os x_i e os y_i não são todos identicamente nulos. Seja, para um j fixo

$$a = \frac{|x_j|^p}{\sum_{i=1}^{n} |x_i|^p}$$

e

$$b = \frac{|y_j|^q}{\sum_{i=1}^{n} |y_i|^q}.$$

Usando a desigualdade $a^{1/p}b^{1/q} \leq \frac{a}{p} + \frac{b}{q}$ mostrada acima, temos que

$$\frac{|x_j||y_j|}{\left(\sum_{i=1}^{n} |x_i|^p\right)^{1/p}\left(\sum_{i=1}^{n} |y_i|^q\right)^{1/q}} \leq \frac{1}{p}\,\frac{|x_j|^p}{\sum_{i=1}^{n} |x_i|^p} + \frac{1}{q}\,\frac{|y_j|^q}{\sum_{i=1}^{n} |y_i|^q}.$$

Somando ambos os lados da desigualdade para todo j entre 1 e n, teremos

$$\frac{\sum_{j=1}^{n} |x_j||y_j|}{\left(\sum_{i=1}^{n} |x_i|^p\right)^{1/p}\left(\sum_{i=1}^{n} |y_i|^q\right)^{1/q}} \;\leq\; \frac{1}{p}\,\frac{\sum_{j=1}^{n} |x_j|^p}{\sum_{i=1}^{n} |x_i|^p} + \frac{1}{q}\,\frac{\sum_{j=1}^{n} |y_j|^q}{\sum_{i=1}^{n} |y_i|^q}.$$

$$\begin{aligned} &= \frac{1}{p} + \frac{1}{q} \\ &= 1, \end{aligned} \tag{2.8}$$

que é o que queríamos provar.

Vamos agora para finalizar generalizar a desigualdade de Hölder para seqüências infinitas.

Uma seqüência infinita $\{x_i\}_{i\in\mathbb{N}}$ de números complexos é dita ser do tipo ℓ_p, $p > 0$, se

$$\sum_{i=1}^{\infty} |x_i|^p \;<\; \infty.$$

Seja $\{x_i\}_{i\in\mathbb{N}}$ uma seqüência do tipo ℓ_p e seja $\{y_i\}_{i\in\mathbb{N}}$ uma seqüência do tipo ℓ_q com $1 < p < \infty$, $1 < q < \infty$ e $1/p + 1/q = 1$.

Como vimos, temos para qualquer $n \in \mathbb{N}$ a desigualdade

$$\sum_{i=1}^{n} |x_i||y_i| \;\leq\; \left(\sum_{i=1}^{n} |x_i|^p\right)^{1/p} \left(\sum_{i=1}^{n} |y_i|^q\right)^{1/q},$$

Assim, segue que

$$\sum_{i=1}^{n} |x_i||y_i| \;\leq\; \left(\sum_{i=1}^{\infty} |x_i|^p\right)^{1/p} \left(\sum_{i=1}^{\infty} |y_i|^q\right)^{1/q} < \infty.$$

Essa desigualdade vale para todo n e diz, em particular, que a seqüência $s_n \;=\; \sum_{i=1}^{n} |x_i||y_i|$, $n \in \mathbb{N}$, é monótona crescente e limitada. Assim, existe $\lim_{n\to\infty} s_n$ e vale

$$\sum_{i=1}^{\infty} |x_i||y_i| \;\leq\; \left(\sum_{i=1}^{\infty} |x_i|^p\right)^{1/p} \left(\sum_{i=1}^{\infty} |y_i|^q\right)^{1/q} < \infty.$$

Essa última relação é a desigualdade de Hölder para seqüências infinitas $\{x_i\}_{i\in\mathbb{N}}$ do tipo ℓ_p e $\{y_i\}_{i\in\mathbb{N}}$ do tipo ℓ_q.

Para finalizar esse ponto vamos a mais uma pequena mas significativa extensão da desigualdade de Hölder.

A desigualdade de Hölder apresentada acima envolve seqüências dos tipos ℓ_p e ℓ_q com $1/p + 1/q = 1$, $1 < p < \infty$ e $1 < q < \infty$. É de se notar que os casos $p = 1$ ou $q = 1$ foram excluídos.

Uma seqüência de números complexos $\{a_i\}_{i\in\mathbb{N}}$ é dita ser do tipo ℓ_∞ se for limitada, ou seja, se existir um número $M \geq 0$ tal que $|a_i| \leq M$ para todo $i \in \mathbb{N}$.

Há também uma desigualdade como a de Hölder envolvendo a seqüências do tipo ℓ_1 e ℓ_∞, que passaremos a descrever a seguir.

Sejam $\{x_i\}_{i\in\mathbb{N}}$ uma seqüência do tipo ℓ_1 e $\{y_i\}_{i\in\mathbb{N}}$ uma seqüência do tipo ℓ_∞.

Então é bem fácil de se verificar que a seguinte relação é válida:

$$\sum_{i=1}^{\infty} |x_i||y_i| \;\leq\; \left(\sum_{i=1}^{\infty} |x_i|\right)\left(\sup_{i\in\mathbb{N}} |y_i|\right).$$

Essa é a versão prometida de desigualdade de Hölder para pares de seqüências de tipo ℓ_1 e ℓ_∞.

A desigualdade de Hölder pode ser generalizada ainda mais, para espaços de funções. Nessa forma mais geral a desigualdade de Hölder é de grande importância. Aqui usaremos a desigualdade mais simples acima para introduzirmos abaixo mais um tipo de métrica em espaços $\mathbb{R}^n$ (ou $\mathbb{C}^n$).

Para isso vamos provar uma das conseqüências da desigualdade de Hölder, conhecida como desigualdade de Minkowski.

• **A Desigualdade de Minkowski. Caso Discreto.**

Sejam $x_1, \ldots, x_n$ e $y_1, \ldots, y_n$ duas seqüências finitas arbitrárias de números complexos ($n \in \mathbb{N}$). A desigualdade de Minkowski afirma que

$$\left(\sum_{i=1}^{n} |x_i + y_i|^p\right)^{1/p} \leq \left(\sum_{i=1}^{n} |x_i|^p\right)^{1/p} + \left(\sum_{i=1}^{n} |y_i|^p\right)^{1/p}$$

para qualquer $p \geq 1$.

Vamos demonstrá-la. O caso $p = 1$ é trivial (por que?). Consideremos então $p > 1$. Teremos que

$$\begin{aligned} \sum_{i=1}^{n} |x_i + y_i|^p &= \sum_{i=1}^{n} |x_i + y_i||x_i + y_i|^{p-1} \\ &\leq \sum_{i=1}^{n} |x_i||x_i + y_i|^{p-1} + \sum_{i=1}^{n} |y_i||x_i + y_i|^{p-1}. \end{aligned} \tag{2.9}$$

Usando a desigualdade de Hölder podemos dizer que

$$\sum_{i=1}^{n} |x_i||x_i + y_i|^{p-1} \leq \left(\sum_{i=1}^{n} |x_i|^p\right)^{1/p} \left(\sum_{i=1}^{n} |x_i + y_i|^{q(p-1)}\right)^{1/q},$$

onde $1/p + 1/q = 1$, ou seja, $p = q(p-1)$. A última desigualdade diz então que

$$\sum_{i=1}^{n} |x_i||x_i + y_i|^{p-1} \leq \left(\sum_{i=1}^{n} |x_i|^p\right)^{1/p} \left(\sum_{i=1}^{n} |x_i + y_i|^{p}\right)^{1/q}$$

e, analogamente,

$$\sum_{i=1}^{n} |y_i||x_i + y_i|^{p-1} \leq \left(\sum_{i=1}^{n} |y_i|^p\right)^{1/p} \left(\sum_{i=1}^{n} |x_i + y_i|^{p}\right)^{1/q}.$$

Substituindo estas duas últimas relações em (2.9), teremos

$$\sum_{i=1}^{n} |x_i + y_i|^p \leq \left(\left(\sum_{i=1}^{n} |x_i|^p\right)^{1/p} + \left(\sum_{i=1}^{n} |y_i|^p\right)^{1/p}\right) \left(\sum_{i=1}^{n} |x_i + y_i|^{p}\right)^{1/q},$$

donde tiramos que

$$\left(\sum_{i=1}^{n} |x_i + y_i|^p\right)^{1/p} \leq \left(\sum_{i=1}^{n} |x_i|^p\right)^{1/p} + \left(\sum_{i=1}^{n} |y_i|^p\right)^{1/p},$$

que é o que queríamos provar.

Assim como a desigualdade de Hölder, a desigualdade de Minkowski pode ser generalizada para seqüências infinitas. Sejam $\{x_i\}_{i\in\mathbb{N}}$ e $\{y_i\}_{i\in\mathbb{N}}$ seqüências infinitas de de números complexos, ambas do tipo ℓ_p.

Temos que, para qualquer $n \in \mathbb{N}$,

$$\left(\sum_{i=1}^{n} |x_i + y_i|^p\right)^{1/p} \leq \left(\sum_{i=1}^{n} |x_i|^p\right)^{1/p} + \left(\sum_{i=1}^{n} |y_i|^p\right)^{1/p} \leq \left(\sum_{i=1}^{\infty} |x_i|^p\right)^{1/p} + \left(\sum_{i=1}^{\infty} |y_i|^p\right)^{1/p} < \infty$$

Como a desigualdade vale para qualquer n, segue que a seqüência $s_n = \left(\sum_{i=1}^{n} |x_i + y_i|^p\right)^{1/p}$, $n \in \mathbb{N}$, é monótona crescente e limitada e, portanto, converge. Fora isso, vale

$$\left(\sum_{i=1}^{\infty} |x_i + y_i|^p\right)^{1/p} \leq \left(\sum_{i=1}^{\infty} |x_i|^p\right)^{1/p} + \left(\sum_{i=1}^{\infty} |y_i|^p\right)^{1/p} < \infty.$$

Essa é a desigualdade de Minkowski para seqüências infinitas de números complexos $\{x_i\}_{i\in\mathbb{N}}$ e $\{y_i\}_{i\in\mathbb{N}}$, ambas do tipo ℓ_p com $p \geq 1$.

Um fato que extraímos da desigualdade de Minkowski é o seguinte: se as seqüências $\{x_i\}_{i\in\mathbb{N}}$ e $\{y_i\}_{i\in\mathbb{N}}$ são ambas do tipo ℓ_p, $p \geq 1$, então a seqüência $\{x_i + y_i\}_{i\in\mathbb{N}}$ também o é. Fora isso, é claro também que se $\{x_i\}_{i\in\mathbb{N}}$ é do tipo ℓ_p então a seqüência $\{\alpha x_i\}_{i\in\mathbb{N}}$ também é do tipo ℓ_p para qualquer $\alpha \in \mathbb{C}$. Esses dois fatos juntos dizem-nos que para cada $p \geq 1$ a coleção de todas as seqüências do tipo ℓ_p é um *espaço vetorial* sobre os complexos, um fato de grande importância. Por isso passaremos a chamar a coleção de todas as seqüências do tipo ℓ_p, $p \geq 1$, de espaço ℓ_p, sempre entendido como um espaço vetorial sobre os complexos.

A coleção de todas as seqüências de tipo ℓ_∞ é também claramente um espaço vetorial (por que?).

Veremos em outro momento que os espaços ℓ_p, $p \geq 1$, são exemplos de *espaços de Banach* e que o espaço ℓ_2 é, em particular, um *espaço de Hilbert*.

A desigualdade de Minkowski pode também ser bastante generalizada, e estendida a espaços de funções.

• **A Desigualdade de Cauchy-Schwarz para Seqüências**

A desigualdade de Hölder tem um caso particular bastante especial. Sejam $\{x_i\}_{i\in\mathbb{N}}$ e $\{y_i\}_{i\in\mathbb{N}}$ duas seqüências de números complexos complexos do tipo ℓ_2. Então a desigualdade de Hölder nos diz que

$$\sum_{i=1}^{\infty} |x_i||y_i| \leq \left(\sum_{i=1}^{\infty} |x_i|^2\right)^{1/2} \left(\sum_{i=1}^{\infty} |y_i|^2\right)^{1/2}. \tag{2.10}$$

Essa desigualdade é conhecida como *desigualdade de Cauchy-Schwarz*[21] (para seqüências) e é, sem exagero, uma das desigualdades mais importantes. Muitos resultados importantes são extraídos dela, alguns dos quais iremos tratar adiante. Por ora limitamo-nos a mencioná-la, dado que trataremos dela com mais detalhe adiante.

*

É quase desnecessário mencionar o quão importantes espaços vetoriais são no contexto da Física, onde, porém, quase somente espaços vetoriais sobre o corpo dos reais ou dos complexos aparecem. Discutiremos mais aspectos básicos da teoria dos espaços vetoriais na Seção 3, página 34.

Vamos agora retomar a apresentação de algumas estruturas algébricas básicas.

*

2.3 Anéis, Álgebras e Módulos

• **Anéis**

Um *anel* é um conjunto A dotado de duas operações binárias denotadas por "+" e "·" e denominadas *soma* e *produto*, respectivamente, tais que A é um grupo Abeliano em relação à operação de soma e um semi-grupo em relação à operação de produto. Por fim, a operação de produto é distributiva em relação à soma: para quaisquer a, b e $c \in A$ valem $a \cdot (b + c) = a \cdot b + a \cdot c$ e $(a + b) \cdot c = a \cdot c + b \cdot c$.

Como usual, denotamos por $-a$ a inversa aditiva do elemento a de um anel.

Se 0 é o elemento neutro de um anel A em relação à operação de soma, então $a \cdot 0 = 0$ pois, como $0 = 0 + 0$, tem-se pela propriedade distributiva $a \cdot 0 = a \cdot 0 + a \cdot 0$, que implica $0 = a \cdot 0 - (a \cdot 0) = a \cdot 0 + a \cdot 0 - (a \cdot 0) = a \cdot 0$.

• **Álgebras**

Uma álgebra é um espaço vetorial V sobre um corpo K dotado de uma operação de produto binária "·" dita *produto* da álgebra, de modo que as seguintes propriedades são satisfeitas

1. O produto da álgebra é distributivo em relação a soma vetorial: para todos a, b e $c \in V$ valem

$$a \cdot (b + c) = a \cdot b + a \cdot c \quad \text{e} \quad (a + b) \cdot c = a \cdot c + b \cdot c.$$

[21] Augustin Louis Cauchy (1789-1857). Karl Herman Amandus Schwarz (1843-1921).

2. O produto por escalares comuta com o produto da álgebra e é distributivo em relação a ele: para todos $a,\ b \in V$ e $\alpha \in K$ vale
$$\alpha(a \cdot b) \ = \ (\alpha a) \cdot b \ = \ a \cdot (\alpha b).$$

Uma álgebra V é dita ser uma *álgebra comutativa* ou *Abeliana*[22] se para todos $a,\ b \in V$ tivermos
$$a \cdot b \ = \ b \cdot a.$$

Uma álgebra V é dita ser uma *álgebra associativa* se para todos $a,\ b$ e $c \in V$ tivermos
$$a \cdot (b \cdot c) \ = \ (a \cdot b) \cdot c.$$

Álgebras associativas são anéis.

Notação. Se A é uma álgebra associativa, podemos sem ambigüidade denotar o produto de dois de seus elementos $a,\ b \in A$ simplesmente por por ab.

Devemos dizer que há muitas álgebras importantes encontradas na Física que não são nem comutativas nem associativas. Por exemplo, a álgebras do produto vetorial em $\mathbb{R}^3$ não é nem comutativa nem associativa.

• **Álgebras de Lie**

Uma classe especialmente importante de álgebras não-comutativas e não-associativas é formada pelas chamadas *álgebras de Lie*.

Uma álgebra L é dita ser uma *álgebra de Lie* se satisfizer

1. *Anti-comutatividade.* Para todos $a, b \in L$ vale $a \cdot b = -b \cdot a$.
2. *Identidade de Jacobi.* Para todos $a,\ b$ e $c \in L$ vale
$$a \cdot (b \cdot c) + c \cdot (a \cdot b) + b \cdot (c \cdot a) \ = \ 0. \tag{2.11}$$

Por razões históricas o produto de dois elementos de uma álgebra de Lie é denotado pelo símbolo $[a,\ b]$ em lugar de $a \cdot b$.

Seja A uma álgebra associativa. Podemos associar a A uma álgebra de Lie definindo o produto $[a,\ b] = ab - ba$ para $a,\ b \in A$. A anti-comutatividade é óbvia e a identidade de Jacobi segue do fato que
$$\begin{aligned}
&[a,\ [b,\ c]] + [c,\ [a,\ b]] + [b,\ [c,\ a]] \ = \\
&\qquad a(bc - cb) - (bc - cb)a + c(ab - ba) - (ab - ba)c + b(ca - ac) - (ca - ac)b \ = \\
&\qquad abc - acb - bca + cba + cab - cba - abc + bac + bca - bac - cab + acb \ = \\
&\qquad 0,
\end{aligned}$$
como se verifica facilmente.

Exercício. Mostre que $\mathbb{R}^3$ dotado do produto vetorial usual é uma álgebra de Lie.

• **Módulos**

Seja A um anel. Um *A-módulo à esquerda* é um grupo abeliano M dotado de uma função $A \times M \to M$ que a cada par $a \in A,\ m \in M$ associa um elemento de M denotado por $a \cdot m$ com as seguintes propriedades: para todos $a,\ b \in A$ e todos $m,\ n \in M$

1. $a \cdot (m + n) = a \cdot m + a \cdot n$,
2. $(a + b) \cdot m = a \cdot m + b \cdot m$,
3. $a \cdot (b \cdot m) = (ab) \cdot m$,

[22] Niels Henrik Abel (1802-1829).

4. Se A possuir uma identidade e, então $e \cdot m = m$.

Seja A um anel. Um *A-módulo à direita* é um grupo abeliano M dotado de uma função $M \times A \to M$ que a cada par $a \in A,\ m \in M$ associa um elemento de M denotado por $m \cdot a$ com as seguintes propriedades: para todos $a,\ b \in A$ e todos $m,\ n \in M$

1. $(m + n) \cdot a = m \cdot a + n \cdot a$,
2. $m \cdot (a + b) = m \cdot a + m \cdot b$,
3. $(m \cdot b) \cdot a = m \cdot (ba)$,
4. Se A possuir uma identidade e, então $m \cdot e = m$.

Sejam A e B dois anéis. Um *bimódulo* em relação a A e B é um grupo abeliano M dotado de duas função $A \times M \to M$ e $M \times B \to M$ que a cada $a \in A,\ b \in B$ e $m \in M$ associam elementos de M denotados por $a \cdot m$ e $m \cdot b$, respectivamente, de modo que M seja um A-módulo à esquerda e um B-módulo à direita e de modo que valha

1. $a \cdot (m \cdot b) = (a \cdot m) \cdot b$ para todos $a \in A,\ b \in B,\ m \in M$.

2.4 Ações e Representações

• **Ações**

Seja M um conjunto e G um grupo. Uma função $\alpha : G \times M \to M$ é dita ser uma *ação à esquerda de G sobre M* se as seguintes condições foram satisfeitas

1. Para todo $g \in G$ a função $\alpha(g,\ \cdot) : M \to M$ é bijetora.
2. Se e é a identidade de G então $\alpha(e,\ \cdot) : M \to M$ é a função identidade: $\alpha(e,\ x) = x$ para todo $x \in M$.
3. Para todos $g,\ h \in G$ e todo $x \in M$ vale
$$\alpha(g,\ \alpha(h,\ x)) \ = \ \alpha(gh,\ x). \tag{2.12}$$

Uma função $\beta : G \times M \to M$ é dita ser uma *ação à direita de G sobre M* se as seguintes condições foram satisfeitas

1. Para todo $g \in G$ a função $\beta(g,\ \cdot) : M \to M$ é bijetora.
2. Se e é a identidade de G então $\beta(e,\ \cdot) : M \to M$ é a função identidade: $\beta(e,\ x) = x$ para todo $x \in M$.
3. Para todos $g,\ h \in G$ e todo $x \in M$ vale
$$\beta(g,\ \beta(h,\ x)) \ = \ \beta(hg,\ x). \tag{2.13}$$

Note-se que a distinção básica entre (2.12) e (2.13) é a ordem do produto no grupo. Se G é abeliano não há distinção entre uma ação à direita ou à esquerda.

Exercício. Seja $\alpha : G \times M \to M$ uma ação à esquerda de um grupo G em um conjunto M. Mostre que $\beta : G \times M \to M$ definida por $\beta(g,\ x) = \alpha(g^{-1},\ x)$ é uma ação à direita de G em M.

É frequente encontrar-se outras notações para designar ações de grupos em conjuntos. Uma ação à esquerda $\alpha(g,\ x)$ é frequentemente denotada por $\alpha_g(x)$, de modo que a relação (2.12) fica $\alpha_g(\alpha_h(x)) = \alpha_{gh}(x)$. Para uma ação à esquerda, (2.13) fica $\beta_g(\beta_h(x)) = \beta_{hg}(x)$.

Talvez a notação mais conveniente seja denotar uma ação à esquerda $\alpha(g,\ x)$ simplesmente por $g \cdot x$ ou apenas gx. A relação (2.12) fica $g(hx) = (gh)x$. Para uma ação à direita $\beta(g,\ x)$ a notação fica $x \cdot g$, ou apenas xg, de modo que (2.13) fica $(xh)g = x(hg)$. Essa notação justifica o uso da nomenclatura *à direita* ou *à esquerda* para classificar as ações.

Seja $\mathcal{F}$ uma coleção de funções bijetoras de um conjunto M em si mesmo. Uma ação $\alpha : G \to M$ é dita ser uma *ação de G em M pela família $\mathcal{F}$* se para todo $g \in G$ as funções $\alpha(g,\ \cdot) : M \to M$ forem elementos do conjunto $\mathcal{F}$.

• **Representações de Grupos** [23]

Uma representação de um grupo é uma ação a esquerda de um grupo em um espaço vetorial pela família das aplicações lineares invertíveis agindo nesse espaço vetorial.

Sejam G um grupo e V um espaço vetorial sobre um corpo K. Uma representação de G em V é uma função $\pi : G \times V \to V$ tal que para todo $g \in G$ as funções $\pi(g, \cdot) : V \to V$ sejam lineares e bijetivas e satisfazem $\pi(e, v) = v$ e $\pi(g, \pi(h, v)) = \pi(gh, v)$ para todos $g, h \in V$ e todo $v \in \mathbb{V}$.

Devido à linearidade é conveniente denotar $\pi(g, v)$ por $\pi(g)v$. Uma representação satisfaz assim:

1. Para todo $g \in G$, $\pi(g)$ é uma aplicação linear bijetora de V em V:
$$\pi(g)(\alpha u + \beta v) = \alpha\pi(g)u + \beta\pi(g)v$$
para todos $\alpha, \beta \in K$ e todos $u, v \in V$.
2. $\pi(e) = \mathbb{1}$, o operador identidade em V.
3. Para todos $g, h \in G$ vale
$$\pi(g)\pi(h) = \pi(gh).$$

• **Representações de Álgebras** [24]

Seja A uma álgebra sobre um corpo K e V um espaço vetorial sobre o mesmo corpo. Uma representação de A em V é uma família de funções lineares de V em V, $\{\pi(a), a \in A\}$, satisfazendo

1. Para todo $a \in A$, $\pi(a) : V \to V$ é uma aplicação linear, ou seja
$$\pi(a)(\alpha u + \beta v) = \alpha\pi(a)u + \beta\pi(a)v$$
para todos $\alpha, \beta \in K$ e todos $u, v \in V$.
2. Para todos $\alpha, \beta \in K$ e todos $a, b \in A$ vale
$$\pi(\alpha a + \beta b) = \alpha\pi(a) + \beta\pi(b).$$
3. Para todos $a, b \in A$
$$\pi(ab) = \pi(a)\pi(b).$$

Uma representação π de uma álgebra A em um espaço vetorial V é dita ser uma *representação fiel* se $\pi(a) = 0$ só ocorrer para $a = 0$.

Uma representação π de uma álgebra A em um espaço vetorial V é dita ser uma *representação não-degenerada* se $\pi(a)v = 0$ para todo $a \in A$ só ocorrer para $v = 0$.

[23] Este tópico pode ser omitido numa primeira leitura.
[24] Este tópico pode ser omitido numa primeira leitura.

3 Espaços Vetoriais

Vamos aqui desenvolver com um pouco mais de detalhe o assunto espaços vetoriais.

3.1 Sub-Espaços e Espaços Quocientes

• **Sub-espaços**

Seja V um espaço vetorial sobre um corpo K. Um subconjunto W de V é dito ser um subespaço de V (sobre o mesmo corpo K) se para todo $\alpha, \beta \in K$ e todo $u, v \in W$ valer que $\alpha u + \beta v \in W$. É evidente que um subespaço de um espaço vetorial é por si só um espaço vetorial.

• **Quocientes**

Se W é um subespaço de um espaço vetorial V sobre um corpo K, então é possível definir em V uma relação de equivalência $E_W \subset V \times V$ da seguinte forma: dizemos que $(u, v) \in V \times V$ pertence a E_W se $u - v \in W$.

Exercício. Mostre que isso de fato define uma relação de equivalência em V.

Seguindo a notação usual denotaremos também essa relação de equivalência pelo símbolo $\sim_W$: $u \sim_W v$ se $u - v \in W$.

Denotemos por V/W o conjunto das classes de equivalência de V pela relação E_W. Denotaremos por $[u] \in V/W$ a classe de equivalência que contem o vetor $u \in V$.

Com esses ingredientes podemos transformar V/W em um espaço vetorial sobre K. Isso se dá definindo em V/W uma soma e um produto por escalares. O vetor nulo será a classe de equivalência $[0]$ que contém o vetor 0. Como subconjunto de V, a classe $[0]$, alias, vem a ser o conjunto W (por que?).

Se $[u]$ e $[v]$ são as classes de equivalência que contém os elementos u e v, respectivamente, de V, então definimos
$$[u] + [v] = [u + v].$$

Exercício. Mostre que essa definição é coerente, no sentido que independe dos representantes (u e v) escolhidos nas classes.

Exercício. Mostre que essa operação de soma é comutativa e associativa.

Exercício. Mostre que $[u] + [0] = [u]$ para todo $u \in V$.

Analogamente, a operação de multiplicação por escalares é definida por
$$\alpha[u] = [\alpha u],$$
para todo $u \in V$.

Exercício. Mostre que essa definição é coerente, no sentido que independe do representante (u) escolhido nas classe.

Exercício. Mostre que o conjunto V/W é, portanto, um espaço vetorial sobre o corpo K com as operações definidas acima.

O espaço vetorial V/W assim obtido é denominado *espaço quociente* de V por W.

3.2 Bases Algébricas de um Espaço Vetorial

• **Dependência Linear**

Um conjunto finito $u_1, \ldots, u_n \in V$ de vetores é dito ser *linearmente dependente* se existir um conjunto de escalares $\alpha_1, \ldots, \alpha_n \in V$, nem todos nulos, tais que
$$\alpha_1 u_1 + \cdots + \alpha_n u_n = 0.$$

Um conjunto arbitrário de vetores $\{u_a, a \in I\}$ é dito ser *linearmente independente* se não possuir nenhum subconjunto finito que seja linearmente dependente.

- **Combinações Lineares**

Para um conjunto finito de vetores $\{u_1, \ldots, u_n\} \subset V$ e de escalares $\{\alpha_1, \ldots, \alpha_n\} \subset K$, uma expressão como

$$\alpha_1 u_1 + \cdots + \alpha_n u_n$$

é dita ser uma *combinação linear* dos vetores $u_1, \ldots, u_n$.

- **Varredura Linear**

Seja $C \subset V$ um conjunto de vetores. A *varredura linear* ("linear span") de C, denotado por span (C) é o conjunto de todos os vetores de V que podem ser escritos como uma combinação linear finita de elementos de C.

- **Bases Algébricas em Espaços Vetoriais**

Uma *base algébrica*[25] em um espaço vetorial V é um conjunto $B = \{b_i, \ i \in I\}$ de vetores linearmente independentes tais que span $(B) = V$ e tais que qualquer vetor u de V pode ser escrito de modo único como uma combinação linear finita de elementos de B.

Se B é uma base algébrica, então para cada $u \in V$ existem univocamente definidos $\alpha_1, \ldots, \alpha_n \in K$ e $i_1, \ldots, i_n \in I$ tais que:

$$u = \alpha_1 b_{i_1} + \cdots + \alpha_n b_{i_n}.$$

Os seguintes teoremas podem ser demonstrados com uso do Lema de Zorn (omitiremos as demonstrações aqui).

Teorema 3.1 *Todo espaço vetorial V possui uma base algébrica, exceto o espaço vetorial trivial $V = \{0\}$.* □

Teorema 3.2 *Dado um espaço vetorial V (não trivial), todas as bases algébricas em V tem a mesma cardinalidade.* □

- **Dimensão Algébrica**

Um espaço vetorial é dito ser de *dimensão algébrica finita* se possuir uma base algébrica finita. Se um espaço vetorial V tem dimensão algébrica finita, sua *dimensão algébrica*, ou simplesmente *dimensão* é definida como sendo o número de elementos de sua base.

Nem todo espaço vetorial tem uma base algébrica finita (vide exemplos abaixo). De modo geral, se um espaço vetorial possui uma base algébrica, sua dimensão algébrica é definida como sendo a cardinalidade de suas bases algébricas (pelo teorema 3.2 acima são todas iguais).

Exemplo 1. $V = \mathbb{C}^n$ sobre o corpo dos complexos ou $V = \mathbb{R}^n$ sobre o corpo dos reais. Tais são bem conhecidos exemplos-protótipo de espaços vetoriais de dimensão finita ($= n$).

Seja $\mathcal{P}$ = conjunto de todos os polinômios de uma variável real com coeficientes complexos: $P_n(t) \in \mathcal{P}$,

$$P_n(t) = a_n t^n + \cdots + a_1 t + a_0$$

com $t \in \mathbb{R}$, $a_i \in \mathbb{C}$, é dito ser um polinômio de grau n se $a_n \neq 0$.

Exemplo 2. $V = \mathcal{P}$ sobre o corpo dos complexos. Este é claramente um espaço vetorial de dimensão infinita. V possui uma base algébrica, a saber, o conjunto de todos os polinômios da forma $b_n = t^n$, $n = 0, 1, 2, \ldots$.

Exemplo 3. $V = \mathbb{R}$ sobre o corpo dos reais. O conjunto dos reais sobre o corpo dos reais é também um espaço vetorial de dimensão 1, a saber, uma possível base é formada pelo elemento 1: $B = \{1\}$, já que, obviamente, qualquer elemento $x \in \mathbb{R}$ pode ser escrito como $x = x \cdot 1$, com x no corpo dos reais.

Esse exemplo pode parecer banal, e de fato o é, mas leva a um anti-exemplo curioso que mostra que a dimensão algébrica de um espaço vetorial é também fortemente dependente do corpo de escalares utilizado.

Exemplo 4. $V = \mathbb{R}$ sobre o corpo dos racionais.

A surpresa aqui é que este **não** é um espaço vetorial de dimensão finita: não existe um conjunto finito $\{x_1, \ldots, x_m\}$ de números reais tais que todo $x \in \mathbb{R}$ possa ser escrito como

$$x = r_1 x_1 + \cdots + r_m x_m,$$

[25] Também denominada "base de Hamel". Georg Hamel (1877-1954)

onde os números r_i são racionais. A razão é que, como $\mathbb{Q}$ é um conjunto contável, a coleção de números que se deixam escrever como o lado direito é uma coleção contável (tem a mesma cardinalidade de $\mathbb{Q}^m$). O conjunto $\mathbb{R}$, porém, não é contável.

Um resultado surpreendente diz, porém, que esse espaço vetorial possui uma base algébrica, ou seja, existe um conjunto $H \subset \mathbb{R}$ tal que para cada $x \in \mathbb{R}$ existe uma conjunto finito $h_1, \ldots, h_n$ de elementos de H e um conjunto finito de racionais $r_1, \ldots, r_n$ tais que $x = r_1 h_1 + \cdots + r_n h_n$. A demonstração da existência de uma tal base faz uso do Lema de Zorn. Essa base é denominada *base de Hamel* de $\mathbb{R}$.

Curiosidade. Seja a equação $f(x+y) = f(x) + f(y)$ para todo $x, y \in \mathbb{R}$. Não é difícil demonstrar que as únicas funções contínuas que a satisfazem são da forma $f(x) = cx$, c constante real. Uma conseqüência curiosa da existência da base de Hamel é que existem também funções não-contínuas que satisfazem essa equação.

Outros exemplos menos dramáticos que mostram a dependência da dimensão com o corpo utilizado são os seguintes: sejam $V_1 = \mathbb{C}$ sobre o corpo dos complexos e $V_2 = \mathbb{C}$ sobre o corpo dos reais. V_1 tem dimensão 1, mas V_2 tem dimensão 2.

Mais adiante faremos uso do seguinte resultado:

Teorema 3.3 *Se em um espaço vetorial V existir um conjunto $\{v_1, \ldots, v_n\}$ de n vetores linearmente independentes, então a dimensão algébrica de V é maior ou igual a n.* □

Prova. A demonstração é feita por absurdo. Suponhamos que haja uma base $B = \{b_1, \ldots, b_k\}$ em V com $k < n$. Então podemos escrever

$$v_1 = \alpha_1 b_1 + \cdots + \alpha_k b_k.$$

pois B é uma base. Nem todos os α_i podem ser nulos. Supondo que α_k seja um elemento não nulo, podemos escrever

$$b_k = (\alpha_k)^{-1}(v_1 - \alpha_1 b_1 - \cdots - \alpha_{k-1} b_{k-1}) \tag{3.1}$$

Analogamente, temos que

$$v_2 = \beta_1 b_1 + \cdots + \beta_k b_k$$

e, usando (3.1), podemos escrever

$$v_2 = \gamma_1 b_1 + \cdots + \gamma_{k-1} b_{k-1} + \lambda_1 v_1.$$

Os γ_i não podem ser todos nulos, pois de outra forma teríamos $v_2 = \lambda_1 v_1$, contrariando a hipótese de os v_i's serem linearmente independentes. Suponhamos que γ_{k-1} seja o elemento não nulo, podemos escrever b_{k-1} como uma combinação linear envolvendo $\{b_1, \ldots, b_{k-2}\}$ e os vetores v_1 e v_2. Prosseguindo, concluiremos após k passos que

$$v_{k+1} = \lambda'_1 v_1 + \cdots + \lambda'_k v_k$$

contrariando a hipótese de que os v_i's são linearmente independentes. ■

- **Bases Topológicas em Espaços Vetoriais** [26]

O conceito de base algébrica não deve ser confundido com o de *base topológica*, conceito esse pertencente ao contexto dos espaços vetoriais topológicos:

Uma *base topológica* em um espaço vetorial topológico V é um conjunto $B = \{b_i, \ i \in I\}$ de vetores linearmente independentes tais que span (B) é um conjunto denso em V, ou seja, o fecho de span (B) é V.

A *dimensão topológica* de um espaço vetorial é então definida como sendo a cardinalidade das bases topológicas de V.

Para ilustrar como os conceitos de base algébrica e base topológica são diferentes, consideremos novamente o seguinte Exemplo 4 acima:

Exemplo 5. $V = \mathbb{R}$ sobre o corpo dos racionais, com a topologia usual sobre $\mathbb{R}$, tem uma base topológica de dimensão finita: $B = \{1\}$. De fato, o conjunto $\{r \cdot 1, \ r \in \mathbb{Q}\}$ é denso em $\mathbb{R}$. Esse espaço vetorial possui então uma dimensão topológica igual a um.

Definição. Um espaço vetorial topológico sobre o corpo dos reais ou dos complexos é dito ser separável se possuir uma base topológica contável.

[26] Este tópico pode ser omitido numa primeira leitura.

3.3 Normas e Produtos Escalares em Espaços Vetoriais

Aqui trataremos exclusivamente de espaços vetoriais sobre o corpo dos complexos.

• **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 $\|u + v\| \leq \|u\| + \|v\|$.

Note-se que, pelo item 3 acima tem-se $\|0\| = 0$. Fora isso, vale que

$$\|\alpha u + \beta v\| \leq |\alpha|\|u\| + |\beta|\|v\|$$

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

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

• **Exemplos**

Seja o espaço $\mathbb{C}^n$ e seja $a = (a_1,\ \ldots,\ a_n) \in \mathbb{C}^n$. Para todo $p \geq 1$ a expressão

$$\|a\|_p \ := \ \left(\sum_{k=1}^{n} |a_k|^p\right)^{1/p}$$

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

Exercício. Verifique isso. Para a propriedade $\|u+v\|_p \leq \|u\|_p + \|v\|_p$ você vai precisar da desigualdade de Minkowski tratada na Seção 2.2.2 e à página 29.

O caso $p = 2$ é bem conhecido e corresponde à chamada norma Euclideana.

Seja o espaço ℓ_p, $p \geq 1$, definido em (2.4), página 22, e seja $a = (a_1,\ a_2,\ a_3,\ \ldots) \in \ell_p$. A expressão

$$\|a\|_p \ := \ \left(\sum_{k=1}^{\infty} |a_k|^p\right)^{1/p}$$

define uma norma em ℓ_p.

Exercício. Verifique isso. Para a propriedade $\|u+v\|_p \leq \|u\|_p + \|v\|_p$ você vai precisar da desigualdade de Minkowski tratada na Seção 2.2.2 e à página 29.

• **Equivalência entre 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 $v \in V$.

Exercício. Mostre que a relação de equivalência entre normas é uma relação de equivalência.

Tem-se o seguinte teorema, cuja demonstração pode ser encontrada em todos os bons livros sobre Álgebra Linear ou sobre Espaços Vetoriais de Dimensão Finita.

Teorema 3.4 *Em um espaço vetorial de dimensão finita sobre $\mathbb{C}$ todas as normas são equivalentes.* □

A afirmação desse teorema é freqüentemente falsa em espaços de dimensão infinita. A importância da noção de equivalência de normas se manifesta no fato que duas normas equivalentes geram a mesma topologia.

• **Semi-Normas**

Uma *semi-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. 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 $\|u + v\| \leq \|u\| + \|v\|$.

Note-se que, pelo item 2, vale para uma semi-norma que $\|0\| = 0$. É evidente pelas definições que toda norma é uma semi-norma. A diferença entre norma e semi-norma é que para uma semi-norma a relação $\|v\| = 0$ não implica $v = 0$.

Para uma semi-norma vale a desigualdade

$$\|a\| \ \geq \ \big|\, \|a-b\| - \|b\| \,\big|, \tag{3.2}$$

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

A desigualdade triangular diz-nos que

$$\|a-b\| \ \leq \ \|a\| + \|b\| \tag{3.3}$$

e que

$$\|b\| \ = \ \|a-(a-b)\| \ \leq \ \|a\| + \|a-b\|. \tag{3.4}$$

De (3.3) segue que

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

e de (3.3) 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 os itens 1 da definição de semi-norma e de norma são supérfluos.

• **Equivalência entre Semi-Normas**

Há uma noção de equivalência entre semi-normas que é idêntica à de equivalência entre normas.

• **Produtos Escalares**

Um *produto escalar* ou *produto interno* em um espaço vetorial V é uma função $V \times V \to \mathbb{C}$, denotada por $\langle u,\ v\rangle$ com[27] $u,\ v \in V$ 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 anti-linear na segunda 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 $\alpha \in \mathbb{C}$.

[27] É frequente também denotar um produto escalar de dois vetores u e v por $(u,\ v)$. Essa notação pode causar confusão com a de par ordenado.

3. Conjugação complexa:
$$\langle u,\, v\rangle = \overline{\langle v,\, u\rangle}$$
para todos $u,\, v \in V$.

4. Para todo $u \in V$
$$\langle u,\, 0\rangle = 0.$$

5. Positividade. Para todo vetor u não nulo
$$\langle u,\, u\rangle > 0.$$

Os postulados da definição acima não são necessariamente independentes e colocamos alguns apenas para ressaltar sua relevância individual. Por exemplo, o item 2 segue de 1 e 3 (por que?). Os itens 1, 2 e 5 implicam o item 3. O item 4 segue de 1 (por que?). Independentes são apenas 1, 3 e 5 ou 1, 2 e 5.

O conceito de produto escalar é um caso particular do conceito um pouco mais geral de *forma sesquilinear*, que estudaremos com um pouco mais de detalhe na seção 3.4, página 41.

Para um produto escalar de dois vetores vale a seguinte e importantíssima desigualdade, conhecida como *Desigualdade de Cauchy-Schwarz*:
$$|\langle u,\, v\rangle|^2 \;\leq\; |\langle u,\, u\rangle||\langle v,\, v\rangle|.$$
A demonstração será apresentada na seção 3.4.

• **Exemplos**

Seja o espaço $\mathbb{C}^n$ e sejam $a = (a_1,\, \ldots,\, a_n)$, $b = (b_1,\, \ldots,\, b_n) \in \mathbb{C}^n$. A expressão
$$\langle a,\, b\rangle \;:=\; \sum_{k=1}^{n} \overline{a_k} b_k$$
define um produto escalar, dito produto escalar Euclideano, em $\mathbb{C}^n$.

Exercício. Verifique isso.

Seja o espaço $\mathbb{C}^n$ e sejam $p_1,\; p_2,\; \ldots,\; p_n$ números positivos não-nulos ($p_k > 0$, $k = 1,\; \ldots,\; n$). Para $a = (a_1,\, \ldots,\, a_n)$, $b = (b_1,\, \ldots,\, b_n) \in \mathbb{C}^n$ expressão
$$\langle a,\, b\rangle_{p_1,\ldots,p_n} \;:=\; \sum_{k=1}^{n} p_k \overline{a_k} b_k$$
define um produto escalar em $\mathbb{C}^n$.

Exercício. Verifique isso.

O caso $p_1 = \cdots = p_n = 1$ corresponde ao produto escalar Euclideano acima.

Seja o espaço ℓ_2, definido em (2.4), página 22, e sejam $a = (a_1,\; a_2,\; a_3,\; \ldots) \in \ell_2$, $b = (b_1,\; b_2,\; b_3,\; \ldots) \in \ell_2$. A expressão
$$\langle a,\, b\rangle \;:=\; \sum_{k=1}^{\infty} \overline{a_k} b_k$$
define uma produto escalar em ℓ_2.

Exercício. Verifique isso. Para provar a finitude da soma do lado direito use a desigualdade de Cauchy-Schwarz para seqüências (2.10), tratada à página 30.

Mais alguns exemplos serão vistos quando tratarmos de formas sesquilineares na Seção 3.4.

• **Norma e Produto Escalar**

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

Exercício. Verifique que os postulados que definem norma são de fato satisfeitos por essa definição.

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
$$\|u+v\|^2 + \|u-v\|^2 = 2\|u\|^2 + 2\|v\|^2. \tag{3.5}$$

Prova. Tem-se simplesmente pelas definições que
$$\|u+v\|^2 = \|u\|^2 + \langle u,\, v\rangle + \langle v,\, u\rangle + \|v\|^2$$
e
$$\|u-v\|^2 = \|u\|^2 - \langle u,\, v\rangle - \langle v,\, u\rangle + \|v\|^2.$$
Somando-se ambas tem-se o resultado.

Exercício. Por que essa relação é chamada “identidade do paralelogramo”?

Identidade de polarização: Para todos os vetores $u,\, v \in V$ vale
$$\langle u,\, v\rangle \;=\; \frac{1}{4}\sum_{n=0}^{3} i^{-n}\|u + i^n v\|^2,$$
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.

Exercício. Por que essa relação é chamada “identidade de polarização”?

Notemos que, com a definição dada acima de norma associada a um produto escalar, a desigualdade de Cauchy-Schwarz fica
$$|\langle u,\, v\rangle| \leq \|u\|\|v\|.$$

• **A Identidade de Polarização**

A identidade de polarização mencionada acima é um caso especial de uma outra ligeiramente mais geral, também denominada identidade de polarização. Seja A um operador linear em um espaço vetorial V sobre os complexos e sejam u e v elementos de seu domínio. Então vale que
$$\langle u,\, Av\rangle \;=\; \frac{1}{4}\sum_{n=0}^{3} i^{-n}\langle (u + i^n v),\; A(u + i^n v)\rangle.$$

Exercício. Mostre isso.

Tomando-se A como o operador identidade reobtem-se a identidade anterior.

• **Obtendo Produtos Escalares a Partir de Normas**

Nas últimas páginas vimos que podemos obter uma norma a partir de um produto escalar. Podemos nos perguntar: se uma norma for dada em um espaço vetorial, seria possível obter um produto escalar a partir dessa norma?

A chave para responder isso é sugerida pelas identidades do paralelogramo e de polarização, ambas válidas para normas definidas a partir de produtos escalares: Se uma norma satisfaz a identidade do paralelogramo, ou seja, se
$$\|u+v\|^2 + \|u-v\|^2 = 2\|u\|^2 + 2\|v\|^2.$$
para todos os vetores $u,\, v \in V$, então um produto escalar pode ser definido por
$$\langle u,\, v\rangle \;=\; \frac{1}{4}\sum_{n=0}^{3} i^{-n}\|u + i^n v\|^2.$$
A demonstração que o lado direito define de fato um produto escalar é engenhosa, a principal dificuldade consiste em demonstrar a linearidade do produto escalar (item 1 da definição de produto escalar). Omitiremos a demonstração aqui.

Mencionemos por fim que nem toda a norma satisfaz a identidade do paralelogramo e, portanto, nem sempre é possível definir um produto escalar a partir de uma norma.

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.

3.4 Formas Sesquilineares, Produtos Escalares e a Desigualdade de Cauchy-Schwarz

• **Formas Sesquilineares. Definições**

Seja V um espaço vetorial complexo. Uma *forma sesquilinear* 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. Anti-linearidade 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.

Exercício. Prove isso.

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

3. 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 $u \in V$,

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

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

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

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

• **A Desigualdade de Cauchy-Schwarz**

Teorema 3.5 *Toda forma sesquilinear positiva ω é Hermitiana, ou seja,*

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

para todos os vetores u e v e não-degenerada, ou seja, se $\omega(v, v) = 0$ então $\omega(v, u) = 0$ para todo u. Fora isso vale a desigualdade de Cauchy-Schwarz[28]*: para todos os vetores u e v,*

$$|\omega(u, v)|^2 \leq \omega(u, u)\,\omega(v, v). \tag{3.6}$$

□

[28] Augustin Louis Cauchy (1789-1857). Karl Herman Amandus Schwarz (1843-1921).

Prova.

Faremos uso do fato que, para qualquer número complexo λ e quaisquer vetores u e v vale, pela hipótese de positividade,

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

Escrevendo-se explicitamente o lado esquerdo temos a desigualdade

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

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

$$f(x, y) := (x^2 + y^2)\omega(v, v) + (x + iy)\omega(u, v) + (x - iy)\omega(v, u) + \omega(u, u) \geq 0.$$

Exercício. Verifique isso.

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

$$\omega(u, v) = \alpha + i\beta \qquad \text{e} \qquad \omega(v, u) = \gamma + i\delta, \tag{3.7}$$

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

$$f(x, y) = (x^2 + y^2)\omega(v, v) + (x\alpha - y\beta) + i(x\beta + y\alpha) + (x\gamma + y\delta) + i(x\delta - y\gamma) + \omega(u, u) \geq 0. \tag{3.8}$$

Como $f(x, y)$ tem que ser real (e ≥ 0) seque 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 (3.7), 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 (3.8) fica

$$f(x, y) = (x^2 + y^2)\omega(v, v) + 2(x\alpha - y\beta) + \omega(u, u). \tag{3.9}$$

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. Isso estabelece que ω é não-degenerada. Fora isso, a desigualdade de Cauchy-Schwarz (3.6) 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 (3.6) para esse caso. Podemos reescrever o lado direito de (3.9) como

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

Exercício. Verifique.

Dai, 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. ■

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 $e \geq \sqrt{5}$.

Exercício. Tente livremente obter outras desigualdades interessantes do mesmo estilo usando esse método.

• **Produtos Internos ou Produtos Escalares**

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

6. Se u é tal que $\omega(u,\ u) = 0$, então $u = 0$.

• **Exemplos**

Exemplo 1. Seja $V = \mathbb{C}^n$. Um exemplo de produto escalar é dado pelo produto escalar usual:

$$\omega(u,\ v) = \langle u,\ v\rangle := \sum_{k=1}^{n} \overline{u_k}v_k$$

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

Exemplo 2. Seja $V = \mathbb{C}^n$. Um exemplo de produto escalar é dado por

$$\omega(u,\ v) = \langle Au,\ Av\rangle$$

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

Exemplo 3. Exemplo de uma forma sesquilinear Hermitiana que não é positiva. Seja $V = \mathbb{C}^n$ e seja dado por

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

onde A é uma matriz $n \times n$ *auto-adjunta*, 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 $\omega(u,\ u) = \langle u,\ Au\rangle = i(u_1\overline{u_2} - \overline{u_1}u_2) = -2\mathrm{Im}(u_1\overline{u_2})$. Assim, essa ω não é positiva e é degenerada (por que?).

Exemplo 4. Exemplo de uma forma sesquilinear que não é Hermitiana. Seja $V = \mathbb{C}^n$ e seja dado por

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

onde A é uma matriz $n \times n$ que <u>não</u> é *auto-adjunta*, 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

$$\omega(u,\ v) = \langle u,\ Au\rangle = \overline{u_1}v_2,$$

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.

Exemplo 5. Exemplo de uma forma sesquilinear positiva mas que não é um produto escalar. Seja $V = \mathbb{C}^n$ e seja dado por

$$\omega(u,\ v) = \langle Au,\ Av\rangle$$

onde A é uma matriz $n \times n$ <u>não</u>*-invertível*. Então, se $\omega(u,\ u) = 0$, temos $\langle Au,\ Au\rangle = 0$, ou seja, $Au = 0$. Dai, seque que $\omega(u,\ v) = \langle Au,\ Av\rangle = 0$ para todo v e, portanto, ω é não-degenerada. No entanto, não é necessariamente verdade que se $Au = 0$ então $u = 0$ pois u, pode ser não-nulo e estar no núcleo de A, pois A não é invertível.

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 invertível (por que?). 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.

Exemplo 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 $x \in [a,\ b]$. Defina-se, para duas funções f e g de V, o seguinte produto escalar:

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

Exemplo 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 fechado $[0,\ 1/2]$ e identicamente nula no intervalo $(1/2,\ 1]$. Defina-se, para duas funções f e g de V, a seguinte forma sesquilinear positiva:

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

• **A Desigualdade de Minkowski**

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

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

A demonstração é simples:

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

que é o que se queria demonstrar. Acima, na passagem da terceira para a quarta, linha usamos a desigualdade de Cauchy-Schwarz.

• **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

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

$v \in V$.

Exercício. Mostre que os postulados da definição de norma são de fato satisfeitos.

• **A Desigualdade Triangular**

A talvez principal importância da desigualdade de Minkowski é a seguinte. Vamos supor que ω seja um produto escalar. Então podemos definir 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 que?). É também claro que $d_\omega(a,\ b)=d_\omega(b,\ a)$ (por que?). 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 que?). 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 *pseudo-métrica*.

• **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, por uma construção canônica, existe um espaço vetorial $\tilde{V}$, um produto escalar $\tilde{\omega}$ e uma aplicação linear sobrejetiva $E: V\to\tilde{V}$ tais que

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

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

Para a mencionada construção, notemos em primeiro lugar que o conjunto de todos os vetores u com a propriedade que $\omega(u,\ u)=0$ formam um sub-espaço 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 $\omega(u,\ u)=\omega(u,\ v)=0$ em função da condição de ω ser positivo e, portanto, não-degenerado (ou, se quiserem, em função desigualdade de Cauchy-Schwarz). Vamos denominar esse sub-espaço por Z. O espaço vetorial quociente $\tilde{V}=V/Z$ (vide a construção da página 34) 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

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

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

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

3.5 O Dual Algébrico de um Espaço Vetorial

Seja V um espaço vetorial sobre um corpo K (por exemplo o corpo $\mathbb{C}$). Uma aplicação $l: V\to K$, definida sobre todo V, é dita ser um *funcional linear* se

$$l(\alpha x+\beta y)\ =\ \alpha l(x)+\beta l(y)$$

para todo $x,\ y\in V$ e todo $\alpha,\ \beta\in K$.

Exercício. Mostre que, de acordo com a definição acima, vale para qualquer funcional linear l que $l(0)=0$.

O conjunto de todas os funcionais lineares de V em K é denominado espaço dual algébrico de V e denotado V'. O conjunto V' é feito um espaço vetorial (sobre K), através da seguinte relação:

$$(\alpha l+\beta m)(x)\ =\ l(\alpha x)+m(\beta x),$$

para todo l e $m\in V'$; $\alpha,\ \beta\in K$ e todo $x\in V$. O vetor nulo de V' é o funcional linear que associa trivialmente todo vetor de V a zero: $l(x)=0,\ \forall x\in V$.

O seguinte teorema é verdadeiro e será implicitamente usado várias vezes no que segue. Sua demonstração é, como veremos, elementar mas instrutiva.

Teorema 3.6 *Seja um espaço vetorial V sobre um corpo K. Se um vetor v tem a propriedade que $l(v)=0$ para todo $l\in V'$ então $v=0$.* □

Prova. Seja B uma base algébrica em V. Para cada elemento $b\in B$ podemos associar um funcional linear l_b, definido da seguinte forma. Como todo $w\in V$ pode ser escrito como uma combinação linear finita de elementos de B, podemos sempre escrever

$$w=w_b b+w',$$

onde w' é uma combinação linear finita de elementos de $B\setminus\{b\}$ e $w_b\in K$. (É claro que $w_b=0$ caso b não comparece na decomposição de w em uma soma finita de elementos de B).

Definimos então

$$l_b(w)\ =\ w_b,$$

para todo vetor $w\in V$. É um exercício simples mostrar que, para cada $b\in B$, a aplicação $l_b: V\to K$ dada acima é um funcional linear.

Exercício. Mostre isso.

Seja então v um vetor como no enunciado do teorema. Se $l(v)=0$ para todo $l\in V'$, vale obviamente que $l_b(v)=0$ para todo $b\in B$. Isso, porém, trivialmente implica que $v=0$. completando a demonstração.

Notação. Para $x\in V$ e $l\in V'$ é muito freqüente, e graficamente conveniente, usar-se a notação $\langle l,\ x\rangle$ em lugar de $l(x)$.

Se A e B são espaços vetoriais e $A\subset B$ então $B'\subset A'$.

Exercício. Justifique essa última afirmativa.

• **O Dual Topológico de um Espaço Vetorial**

Seja V um espaço vetorial topológico. O conjunto de todos os funcionais lineares contínuos sobre V é dito ser o *dual topológico* de V. O dual topológico será denotado nestas notas por $V^\dagger$. Note-se que $V^\dagger\subset V'$.

• **Exemplos de Funcionais Lineares**

Exemplo 1. Seja $V=\mathbb{C}^n$, sobre o corpo dos complexos. Seja $a_1,\ \ldots,\ a_n$ um conjunto fixo de números complexos. Para qualquer vetor $z=(z_1,\ \ldots,\ z_n)\in\mathbb{C}^n$ defina-se

$$l(z)\ =\ \overline{a_1}z_1+\cdots+\overline{a_n}z_n.$$

Então l é um funcional linear em $\mathbb{C}^n$.

Exercício. Verifique.

Em verdade, é possível demonstrar a recíproca: em $\mathbb{C}^n$ todo funcional linear é da forma acima para algum conjunto $\{a_1,\ \ldots,\ a_n\}$. Essa afirmativa é um caso particular de um teorema importante conhecido como "Lema de Riesz", que é válido no contexto mais geral dos chamados espaços de Hilbert, dos quais $\mathbb{C}^n$ é um exemplo.

Seja $\mathcal{P}$ = conjunto de todos os polinômios de uma variável real com coeficientes complexos: $P_n(t)\in\mathcal{P}$,

$$P_n(t)=a_nt^n+\cdots+a_1t+a_0$$

com $t \in \mathbb{R}$, $a_i \in \mathbb{C}$, é dito ser um polinômio de grau n se $a_n \neq 0$. O conjunto $\mathcal{P}$ é claramente um espaço vetorial sobre os complexos.

Exemplo 2. Para cada $t_0 \in \mathbb{R}$ e $p \in \mathcal{P}$,

$$l(p) = p(t_0)$$

é um funcional linear em $\mathcal{P}$.

Exercício. Verifique.

Esse exemplo pode ser generalizado:

Exemplo 3. Sejam $t_1, \ldots, t_n \in \mathbb{R}$, distintos, e $a_1, \ldots, a_n$ números complexos. Para todo $p \in \mathcal{P}$, definamos

$$l(p) = a_1 p(t_1) + \cdots + a_n p(t_n).$$

Então l é um funcional linear em $\mathcal{P}$.

Exercício. Verifique.

O último exemplo pode ser fortemente generalizado nos dois exemplos que seguem.

Exemplo 3. Seja (a, b) um intervalo finito de $\mathbb{R}$ e h uma função integrável nesse intervalo (ou seja, $\int_a^b |h(t)| dt \leq \infty$). Então,

$$l(p) = \int_a^b \overline{h(t)}\, p(t)\, dt$$

está definida para todo $p \in \mathcal{P}$ e define um funcional linear em $\mathcal{P}$.

Exercício. Justifique as duas últimas afirmativas.

Exemplo 4. Seja a função $g(x) = e^{-x^2}$. Então

$$l(p) = \int_{-\infty}^{\infty} g(t)\, p(t)\, dt.$$

está definida para todo $p \in \mathcal{P}$ e define um funcional linear em $\mathcal{P}$.

Exercício. Justifique as duas últimas afirmativas.

• A relação entre V e V' [29]

Vamos aqui discutir o fato que sempre existe uma maneira (não-canônica, vide abaixo) de associar vetores de um espaço vetorial V com elementos de seu dual algébrico V'.

Seja V um espaço vetorial sobre um corpo K e $B \subset V$ uma base algébrica em V. Seja $\mathcal{F}_B$ a coleção de todas as funções de B em K. Afirmamos que existe uma bijeção de $\mathcal{F}_B$ sobre V', ou seja, esses dois conjuntos podem ser identificados nesse sentido.

Para tal, seja $f \in \mathcal{F}_B$. Definimos uma aplicação $I : \mathcal{F}_B \to V'$ da seguinte forma. Como todo $x \in V$ pode ser escrito como uma combinação linear finita de elementos de B, digamos, $x = \alpha_1 b_{i_1} + \cdots + \alpha_n b_{i_n}$, escrevemos

$$I(f)(x) = \alpha_1 f(b_{i_1}) + \cdots + \alpha_n f(b_{i_n}).$$

$I(f)$ é um funcional linear pois, se escrevemos $y = \alpha_{n+1} b_{i_{n+1}} + \cdots + \alpha_{n+m} b_{i_{n+m}}$, teremos

$$\begin{aligned} I(f)(x+y) &= \alpha_1 f(b_{i_1}) + \cdots + \alpha_{n+m} f(b_{i_{n+m}}) \\ &= \alpha_1 f(b_{i_1}) + \cdots + \alpha_n f(b_{i_n}) + \alpha_{n+1} f(b_{i_{n+1}}) + \cdots + \alpha_{n+m} f(b_{i_{n+m}}) \\ &= I(f)(x) + I(f)(y). \end{aligned} \tag{3.10}$$

Isso então mostrou que $I(f)$ é de fato um elemento de V' para cada $f \in \mathcal{F}_B$. Vamos mostrar o reverso: que a cada elemento l de V' há um elemento g_l de $\mathcal{F}_B$ associado e que $I(g_l) = l$. Seja novamente $x = \alpha_1 b_{i_1} + \cdots + \alpha_n b_{i_n} \in V$ e seja l um elemento de V'. Tem-se

$$l(x) = \alpha_1 l(b_{i_1}) + \cdots + \alpha_n l(b_{i_n}).$$

[29] Este tópico pode ser omitido numa primeira leitura.

Definimos então $g_l : B \to K$ por

$$g_l(b) = l(b)$$

para todo $b \in K$. Pela definição

$$\begin{aligned} I(g_l) &= \alpha_1 g_l(b_{i_1}) + \cdots + \alpha_n g_l(b_{i_n}) \\ &= \alpha_1 l(b_{i_1}) + \cdots + \alpha_n l(b_{i_n}) \\ &= l(x) \end{aligned} \tag{3.11}$$

para todo $x \in V$. Logo $I(g_l) = l$ como queríamos.

A aplicação $I : \mathcal{F}_B \to V'$ é, portanto, uma bijeção entre esses dois conjuntos. Notemos, porém, que essa bijeção não é canônica no sentido que a mesma depende da base adotada. Se trocarmos B por outra base a bijeção altera-se.

De posse desses fatos podemos entender a relação entre V e V' da seguinte forma. Seja o subconjunto $\mathcal{G}_B$ de $\mathcal{F}_B$ formado por todas as funções que assumem valores não nulos (no corpo K) apenas para um conjunto finito de B, ou seja, para para $g \in \mathcal{G}_B$ existe um conjunto finito $B_g = \{b_1, \ldots, b_n\} \subset B$ tal que g é não-nula nos elementos de B_g, mas é nula em $B \setminus B_g$.

Os conjuntos $\mathcal{G}_B$ e V podem ser identificados no seguinte sentido. Afirmamos que existe uma bijeção $J : \mathcal{G}_B \to V$. Tal é fácil de ver se lembrarmos que os elementos de V podem ser escritos como uma combinação linear finita de elementos de B. De fato, para $g \in \mathcal{G}_B$ definimos

$$J(g) = g(b_1)b_1 + \cdots + g(b_n)b_n \in V$$

onde $\{b_1, \ldots, b_n\} = B_g$. Reciprocamente, se $x \in V$ e $x = \alpha_1 b_{i_1} + \cdots + \alpha_n b_{i_n}$, definimos $g_x \in \mathcal{G}_B$ por

$$g_x(b_{i_a}) = \alpha_a, \qquad a = 1, \ldots, n$$

e

$$g_x(b) = 0,$$

se $b \notin \{b_{i_1}, \ldots, b_{i_n}\}$. É fácil ver então que

$$\begin{aligned} J(g_x) &= g(b_{i_1})b_{i_1} + \cdots + g(b_i n)b_{i_n} \\ &= \alpha_1 b_{i_1} + \cdots + \alpha_n b_{i_n} \\ &= x, \end{aligned} \tag{3.12}$$

o que mostra que J é bijetora. Notemos novamente que essa bijeção também não é canônica, no sentido que a mesma depende da base adotada. Se trocarmos B por outra base a bijeção altera-se.

Exercício importante. Mostre agora que $J^{-1} : V \to \mathcal{G}_b$ é linear, ou seja, $J^{-1}(\alpha x + \beta y) = \alpha J^{-1}(x) + \beta J^{-1}(y)$ para todos $x, y \in V$ e todos $\alpha, \beta \in K$.

Juntando o discutido acima, concluímos que $\phi_1 = I \circ J^{-1}$ é uma aplicação linear injetora de V em V'. A mesma, porém, não é "natural", pois depende da base algébrica B escolhida.

Assim, fixada uma base B em V há uma maneira de associar todos os elementos de V com elementos do seu dual algébrico. Notemos porém que pode haver elementos de V' aos quais não correspondem tais identificações, ou seja, a imagem de $\phi_1 = I \circ J^{-1}$ é tipicamente (especialmente em dimensão infinita) um subconjunto próprio de V'.

Exemplo. Seja $\mathcal{P}$ o espaço vetorial dos polinômios em $\mathbb{R}$ definido acima. Seja $T = \{t_i \in \mathbb{R},\ i \in \mathbb{N}\}$, um conjunto contável de pontos distintos da reta real e seja $q(t) = q_0 + q_1 t + \cdots + q_n t^n$, polinômio. Definamos $l_q \in V'$ por

$$l_q(p) = q_0 p(t_0) + q_1 p(t_1) + \cdots + q_n p(t_n).$$

Exercício. Mostre que a aplicação $\mathcal{P} \ni q \to l_q \in V'$ é linear e injetora.

Exercício. Será que com o conjunto T fixado todo elemento de V' seria da forma l_q para algum q?. Pense. Inspire-se nos exemplos 3 e 4 da página 47. O que acontece para conjuntos T diferentes?

Comentário. Mais interessante que a relação entre V e V', é a relação de V com o *dual algébrico* de V', o chamado bi-dual algébrico de V e denotado por $(V')'$, assunto que discutiremos agora. A razão é que, ao contrário do que tipicamente ocorre entre V e V', há sempre uma aplicação linear injetora entre V e $(V')'$ que é natural, ou seja, independente de escolhas de bases.

Outro interesse na relação entre V e $(V')'$ reside no fato que a mesma revela-nos, como veremos, uma profunda distinção entre espaços vetoriais de dimensão finita e infinita.

• **O Bi-dual Algébrico de um Espaço Vetorial** [30]

Se V é um espaço vetorial sobre um corpo K já observamos que V' é também um espaço vetorial sobre o mesmo corpo. Assim, V' tem também seu dual algébrico que é denominado bi-dual algébrico de V.

O bi-dual algébrico de um espaço vetorial V é o espaço $(V')'$. Como vimos nas páginas anteriores, existe pelo menos uma aplicação linear injetiva de V em V'. Chamemos esta aplicação de ϕ_1. Analogamente, existe pelo menos uma aplicação linear injetiva ϕ_2 de V' em $(V')'$. A composição $\phi_2 \circ \phi_1$ fornece uma aplicação linear injetiva de V em $(V')'$. Como ϕ_1 e ϕ_2 dependem de escolhas de base, a composição $\phi_2 \circ \phi_1$ também depende, não sendo, assim, natural.

Ao contrário do que ocorre na relação entre V e V', podemos sempre encontrar uma aplicação linear injetiva de V em $(V')'$ e que é natural: independente de base. Vamos denotá-la por λ. Definimos $\lambda : V \to (V')'$ da seguinte forma: para $x \in \mathbb{V}$, $\lambda(x)$ é o elemento de $(V')'$ que associa a cada $l \in V'$ o valor $l(x)$:

$$\lambda(x)(l) \;=\; l(x).$$

Exercício. Mostre que $\lambda : V \to (V')'$ é linear.

Exercício. Mostre que $\lambda : V \to (V')'$ é injetora. Sugestão: use o teorema 3.6, enunciado e demonstrado na página 46.

É transparente pela definição de λ que a mesma é independente de bases e, portanto, "natural". A relação entre $x \in V$ e um elemento de $(V')'$ mostra acima é tão direta que quase poderíamos dizer que V é um subconjunto de $(V')'$: $V \subset (V')'$. Alguns autores, abusando um pouco da linguagem, chegam mesmo a escrever uma tal relação de inclusão. Mais correta, no entanto é a relação $\lambda(V) \subset (V')'$.

Poderíamos nesse momento nos perguntar: quando podemos eventualmente ter $\lambda(V) = (V')'$? Para o caso de espaços vetoriais sobre o corpo dos reais ou dos complexos resposta é simples e um tanto surpreendente e se expressa no seguinte teorema.

Teorema 3.7 *Seja V um espaço vetorial sobre o corpo dos reais ou dos complexos. Então $\lambda(V) = (V')'$ se e somente se V é um espaço vetorial de dimensão finita.* □

Este teorema revela uma importante distinção entre espaços de dimensão finita e infinita. Em dimensão finita todos os funcionais lineares do dual algébrico de V' são da forma $\lambda(x)$ para algum vetor x. Em dimensão infinita, porém, há certamente elementos em $(V')'$ que não são dessa forma. Assim, ou tomarmos duais duplos em dimensão infinita sempre obtemos espaços vetoriais "maiores", o que não ocorre em dimensão finita.

Prova. Seja V um espaço vetorial sobre um corpo $K = \mathbb{C}$ ou $\mathbb{R}$.

Caso de dimensão finita. Vamos em primeiro lugar supor que V é de dimensão finita e denotemos por dim V sua dimensão. Seja também $B = \{b_1, \ldots, b_n\}$ uma base de V. É claro que o número de elementos de B é $n = \dim V$.

É fácil mostrar que o conjunto $\{\lambda(b_1), \ldots, \lambda(b_n)\}$ é linearmente independente em $(V')'$. De fato, se existirem escalares α_i tais que

$$\alpha_1\lambda(b_1) + \cdots + \alpha_1\lambda(b_n) = 0$$

ou seja,

$$\lambda(\alpha_1 b_1 + \cdots + \alpha_n b_n) = 0$$

teríamos para todo $l \in V'$

$$\lambda(w)(l) = l(w) = 0$$

onde $w = \alpha_1 b_1 + \cdots + \alpha_1 b_n$. Isso, porém, implica $w = 0$ (pelo teorema 3.6, página 46), o que implica $\alpha_1 = \cdots = \alpha_n = 0$.

[30] Este tópico pode ser omitido numa primeira leitura.

Isso claramente diz que dim $(V')' \geq \dim V$. Afirmamos que a igualdade só se dá se $\lambda(V) = (V')'$. De fato, se $\lambda(V) = (V')'$ então todo elemento de $(V')'$ é da forma

$$\lambda(\alpha_1 b_1 + \cdots + \alpha_n b_n) = \alpha_1\lambda(b_1) + \cdots + \alpha_1\lambda(b_n)$$

e, portanto $\{\lambda(b_1), \ldots, \lambda(b_n)\}$ é uma base em $(V')'$ e dim $(V')' = \dim V$. Se, por outro lado, $\lambda(V)$ é um subconjunto próprio de $(V')'$, existem elementos $v'' \in (V')'$ tais que $v'' - \alpha_1\lambda(b_1) - \cdots - \alpha_n\lambda(b_n) \neq 0$ para todos $\alpha_i \in K$. Portanto, $\{v'', \lambda(b_1), \ldots, \lambda(b_n)\}$ é um conjunto de $n+1$ vetores linearmente independentes. Logo dim $(V')' > n = \dim V$, pelo teorema 3.3, página 36.

Vamos então mostrar que obrigatoriamente tem-se que dim $(V')' = \dim V$, provando o teorema.

Como vimos quando discutimos a relação entre V e V' à página 47, V' é equivalente ao conjunto $\mathcal{F}_B$ de todas as funções de B em K, enquanto que V é equivalente ao conjunto $\mathcal{G}_B$ formado por todas as funções que assumem valores não nulos (no corpo K) apenas para um conjunto finito de B. Como B tem um número finito de elementos, sucede $\mathcal{G}_B = \mathcal{F}_B$ (por que?). Logo V e V' são equivalentes: existe uma bijeção linear φ_1 entre ambos.

A aplicação φ_1 leva a base B em uma base $\varphi_1(B)$ em V'. Para ver isso, notemos que todo elemento $l \in V'$ é da forma $l = \varphi_1(v)$, para algum $v \in V$. Como todo $v \in V$ é da forma $v = \alpha_1 b_1 + \cdots + \alpha_n b_n$, segue que todo elemento $l \in V'$ é da forma $\alpha_1\varphi_1(b_1) + \cdots + \alpha_n\varphi_1(b_n)$. Como φ_1 é bijetora, $\{\varphi_1(b_1), \ldots, \varphi_1(b_n)\}$ é um conjunto de vetores linearmente independentes pois se existirem escalares $\beta_1, \ldots, \beta_n$ tais que

$$\beta_1\varphi_1(b_1) + \cdots + \beta_n\varphi_1(b_n) = 0$$

teríamos $\varphi_1(\beta_1 b_1 + \cdots + \beta_n b_n) = 0$ o que implica $\beta_1 b_1 + \cdots + \beta_n b_n = 0$, pois φ_1 é bijetora. Isso porém implica $\beta_1 = \cdots = \beta_n = 0$, pois $\{b_1, \ldots, b_n\}$ é uma base. Assim, $\varphi_1(B) = \{\varphi_1(b_1), \ldots, \varphi_1(b_n)\}$ é uma base em V' e, portanto, dim $V' = n = \dim V$.

Analogamente, tem-se que V' e $(V')'$ são equivalentes e, portanto, existe uma bijeção linear φ_2 entre ambos que leva a base $\varphi_1(B)$ em uma base $\varphi_2 \circ \varphi_1(B)$ em $(V')'$. Portanto, dim $V' = \dim (V')'$.

Logo dim $V = \dim V' = \dim (V')'$, como queríamos provar.

Caso de dimensão infinita. No caso de dimensão infinita desejamos mostrar que sempre há elementos em $(V')'$ que não são da forma $\lambda(x)$ para algum $x \in V$.

Abaixo K é o corpo dos reais ou dos complexos.

Vamos primeiro delinear a estratégia a ser seguida. Seja B uma base em V (fixa daqui por diante). Como sabemos, existe uma aplicação linear bijetora $S : \mathcal{F}_B \to V'$. Uma função $s : B \to K$, $s \in \mathcal{F}_B$ é dita ser *limitada* se existir um $M > 0$ tal que $|s(b)| < M$ para todo $b \in B$. Seja $\mathcal{L}_B$ o conjunto de todas as funções limitadas de B em K. É claro que $\mathcal{L}_B \subset \mathcal{F}_B$. Vamos mostrar o seguinte: não existe nenhum vetor não nulo $v \in V$ com a propriedade que

$$\lambda(v)(\beta) = 0$$

para todo $\beta \in \phi(\mathcal{L}_B)$. Seja $v = \alpha_1 b1 + \cdots + \alpha_m b_m$ um tal vetor para o qual $\lambda(v)(\beta) = 0$. Isso significa que para todo $\beta \in \phi(\mathcal{L}_B)$

$$0 = \lambda(v)(\beta) = \beta(v) = \alpha_1\beta(b1) + \cdots + \alpha_m\beta(b_m).$$

Tomemos funcionais β_i's da forma

$$\beta_i(b) = \begin{cases} 1, & \text{se } b = b_i \\ 0, & \text{de outra forma} \end{cases}$$

para $i = 1, \ldots, m$. Como todo β_i é um elemento de $\phi(\mathcal{L}_B)$ (por que?), teríamos $0 = \beta_i(v) = \alpha_i$ para todo i, o que implica $v = 0$.

A conclusão é que nenhum elemento de $(V')'$ que seja da forma $\lambda(v)$ para algum $v \in V$ não nulo pode anular todos os elementos de $\phi(\mathcal{L}_B) \subset V'$. A estratégia que seguiremos será a de exibir um elemento de $(V')'$ que tem precisamente a propriedade de anular todos os elementos de $\phi(\mathcal{L}_B)$. Um tal elemento não pode pertencer, portanto, a $\lambda(V)$, o que mostra que $\lambda(V)$ é um subconjunto próprio de $(V')'$ no caso de dimensão infinita.

Seja $u \in V' \setminus \phi(\mathcal{L}_B)$ e U o sub-espaço de V' gerado por u. Todo elemento $l \in V'$ pode ser escrito de modo único na forma

$$l = au + y$$

onde $a \in K$ e y pertence ao sub-espaço complementar de U. Definamos $\alpha(l) = a$. É claro que $\alpha \in (V')'$ e que α aniquila todo elemento de $\phi(\mathcal{L}_B)$, pois estes pertencem ao sub-espaço complementar de U (por que?). Assim, $\alpha \in (V')'$ mas $\alpha \notin \lambda(V)$. ■

4 Somas Diretas e Produtos Tensoriais

Nota. Toda esta seção pode ser omitida numa primeira leitura.

4.1 Algumas Noções Preliminares

• **Grupo Normal e Grupo Quociente**

Seja G um grupo. Um sub-grupo N de G é dito ser *normal* se valer $gng^{-1} \in N$ para todo $n \in N$ e todo $g \in G$. Note-se que todo subgrupo de um grupo Abeliano é normal.

A presença de um subgrupo normal N de um grupo G permite-nos definir um grupo, que será chamado de quociente de G por N, através do seguinte procedimento. Primeiramente definimos em G uma relação de equivalência da seguinte maneira: dizemos que dois elementos x e y de G são equivalentes, $x \sim y$, se $xy^{-1} \in N$. Denotemos por G/N a coleção das classes de equivalência de G por essa relação. Seja $[\cdot]_N$ a aplicação $G \to G/N$ que associa a cada elemento de G a classe de equivalência a qual o elemento pertence. Note-se que $[\cdot]_N$ é sobrejetora. $[\cdot]_N$ é denominada aplicação quociente (associada a N). Podemos fazer de G/N um grupo definindo o produto como $[g]_N[h]_N = [gh]_N$. É fácil verificar que, pelo fato de N ser normal, essa noção de produto dá de fato a G/N uma estrutura de grupo, cujo elemento neutro é $[e]_N$, sendo e o elemento neutro de G.

Essa construção do grupo quociente será usada várias vezes no que segue.

• **Grupo Abeliano Livremente Gerado por um Conjunto**

Outra noção importante é a do *grupo Abeliano livremente gerado por um conjunto* X. Seja X um conjunto. Seja $F(X)$ a coleção de todas as funções de suporte finito de X em $\mathbb{Z}$. É fácil ver que $F(X)$ tem naturalmente uma estrutura de grupo Abeliano, definindo, para $f, f' \in F(X)$ o produto de f e f' como sendo o elemento $ff' = (f + f')$ de $F(X)$ dado por

$$(f + f')(x) = f(x) + f'(x). \tag{4.1}$$

para todo $x \in X$. É claro que esse $(f + f')$ tem suporte finito. O elemento neutro e de $F(X)$ é claramente a função identicamente nula. Pelo fato de $F(X)$ ter essa estrutura natural de grupo $F(X)$ é denominado grupo Abeliano livremente gerado pelo conjunto X.

Para $x \in X$ vamos denotar por δ_x a função característica de x:

$$\delta_x(y) := \begin{cases} 1, & \text{se } y = x \\ 0, & \text{se } y \neq x \end{cases} . \tag{4.2}$$

Claramente $\delta_x \in F(X)$. Dado que cada $f \in F(X)$ tem suporte finito, pode-se escrever-lo da forma

$$f = \sum_{n=1}^{N} a_n \, \delta_{x_n}, \tag{4.3}$$

para valores de N e dos a_n's dependentes de f, com $\{x_1, \ldots, x_N\} = \operatorname{supp} f$ e com $a_i \in \mathbb{Z}$ para $i = 1, \ldots, N$.

Com um um flagrante abuso de linguagem é costume escrever (4.3) da forma

$$f = \sum_{n=1}^{N} a_n \, x_n, \tag{4.4}$$

onde fica, por assim dizer, subentendido que aqui os x_n's representam não os elementos de X mas sim suas funções características (X pode ser um conjunto qualquer, de modo que operações como soma de elementos de X ou multiplicação de elementos de X por um inteiro podem não serem sequer definidas).

É fácil verificar que $F(X)$ é um grupo Abeliano livre (dai seu nome), o que quer dizer que não há em $F(X)$ nenhuma relação não trivial entre seus elementos, a não ser aquela que lhe confere Abelianidade: $ff'f^{-1}f'^{-1} = e$.

• **Relações e Grupos Gerados Módulo Relações**

Vamos passar agora a uma construção muito importante, a de grupo Abeliano livremente gerado por um conjunto módulo relações. Vamos apresentar essa construção de forma bem geral.

Seja J um conjunto (em princípio arbitrário) de índices e sejam então, para cada $j \in J$, elementos de $F(X)$ dados por

$$r_j = \sum_{i=1}^{n(j)} \alpha_{j,\,i} \, x_{j,\,i} \tag{4.5}$$

onde, para cada $j \in J$, $n(j) \in \mathbb{N}$ e, para todo $j \in J$ e $i \in \{1, \ldots, n(j)\}$, tem-se $\alpha_{j,\,i} \in \mathbb{Z}$ e $x_{j,\,i} \in X$ com $x_{j,\,i} \neq x_{j,\,i'}$ se $i \neq i'$. Denotamos $\mathcal{R} := \{r_j, \; j \in J\}$. Os elementos de $\mathcal{R}$ serão chamados "relações".

Seja então R o subgrupo de $F(X)$ formado por todos os elementos de $F(X)$ que são combinações lineares finitas de r_j's com coeficientes em $\mathbb{Z}$:

$$s \in R \iff s = s_1 r_{j_1} + \cdots + s_m r_{j_m}, \tag{4.6}$$

para certos $s_i \in \mathbb{Z}$ e $m \in \mathbb{N}$, que dependem de s. R é dito ser o subgrupo de $F(X)$ gerado pelos r_j's.

Por ser um subgrupo de um grupo Abeliano, R é normal. Assim, podemos definir o *"grupo Abeliano livremente gerado por X, módulo as relações $\mathcal{R}$"* como sendo o grupo $F(X)/R$. Note-se que $[R]_R = e$, o que equivale a dizer os elementos de $\mathcal{R}$ são identificados como zero (dai serem chamados de relações, pois refletem identidades que não existiam em $F(X)$ e que estão sendo agora impostas em $F(X)/R$).

4.2 Produtos Tensoriais

• **O Produto Tensorial de dois Grupos Abelianos**

Vamos ilustrar as definições da última sub-seção com uma definição importante, a de produto tensorial de dois grupos Abelianos.

Sejam A e B dois grupos Abelianos cujo produto de grupo denotaremos aditivamente: com o símbolo $+$. Seja $X = A \times B$. Seja em $F(X) = F(A \times B)$ o conjunto $\mathcal{R}$ de relações dado por

$$\begin{aligned} \mathcal{R} \;:=\; & \{r \in F(X) |\; r = (a + a',\, b) - (a,\, b) - (a',\, b) \\ & \text{ou } r = (a,\, b + b') - (a,\, b) - (a,\, b + b'),\; \text{com } a,\, a' \in A \text{ e } b,\, b' \in B\}. \end{aligned} \tag{4.7}$$

Seja $R = R(A \times B)$ o subgrupo de $F(A \times B)$ gerado por $\mathcal{R}$. Chegamos assim a definição do grupo Abeliano $A \otimes B$, o produto tensorial de A e B, que é definido como $A \otimes B \; := \; F(A \times B)/R(A \times B)$.

• **O Produto Tensorial de dois Espaços Vetoriais**

Sejam U e V dois espaços vetoriais (sobre $\mathbb{C}$). Como U e V são dois grupos Abelianos, o grupo Abeliano $U \otimes V$ está definido pelo procedimento da última sub-seção. Isso, entretanto, ainda não faz de $U \otimes V$ um espaço vetorial. Para isso tomemos $X = U \otimes V$ e consideremos o sub-espaço de $F(X)$ definido por

$$\mathcal{R} := \{r \in F(U \otimes V) |\; r = (\alpha u) \otimes v - u \otimes (\alpha v),\; \text{com } \alpha \in \mathbb{C},\; u \in U,\; v \in V\}. \tag{4.8}$$

Como antes, seja $R = R(U \otimes V)$ o subgrupo gerado por $\mathcal{R}$. Definimos agora um novo grupo Abeliano $U \otimes_{\mathbb{C}} V$ como $U \otimes_{\mathbb{C}} V \; := \; F(U \otimes V)/R(U \otimes V)$. $U \otimes_{\mathbb{C}} V$ não é apenas mais um grupo Abeliano, mas possui naturalmente uma estrutura de espaço vetorial. Para isso é preciso definir o produto de um escalar por um elemento do espaço. Definimos então o produto de um escalar $\alpha \in \mathbb{C}$ por um elemento $u \otimes_{\mathbb{C}} v \in U \otimes_{\mathbb{C}} V$ como sendo o elemento $(\alpha u) \otimes_{\mathbb{C}} v = u \otimes_{\mathbb{C}} (\alpha v)$ (ambos são idênticos pela definição de $U \otimes_{\mathbb{C}} V$).

O assim definido espaço vetorial $U \otimes_{\mathbb{C}} V$ é o produto tensorial dos espaços vetoriais U e V sobre o corpo $\mathbb{C}$.

• **O Produto Tensorial de dois Módulos sobre uma Álgebra**

Vamos aqui a uma definição que nos será importante. Sejam M e N dois bimódulos sobre uma álgebra A, ambos supostos serem espaços vetoriais sobre o corpo dos complexos. Conforme a sub-seção anterior podemos definir o espaço vetorial $M \otimes_{\mathbb{C}} N$. Entanto, em muitos casos é necessário definir um outro tipo de produto tensorial entre M e N.

Para tal seja $X = M \otimes_{\mathbb{C}} N$ e definamos em $F(X)$ o conjunto de relações

$$\mathcal{R} \; := \; \{r \in F(X) | \; r = (ma) \otimes_{\mathbb{C}} n - m \otimes_{\mathbb{C}} (an), \; \text{com } a \in A, \; m \in M, \; n \in N\}. \tag{4.9}$$

Definamos então $R = R(M \otimes_{\mathbb{C}} N)$ como o subgrupo gerado por $\mathcal{R}$ e o produto tensorial

$$M \otimes_A N \; := \; F(M \otimes_{\mathbb{C}} N)/R(M \otimes_{\mathbb{C}} N). \tag{4.10}$$

Podemos fazer de $M \otimes_A N$ um módulo, digamos a direita, sobre A tomando o produto

$$a \cdot (m \otimes_A n) \; := \; (ma) \otimes_A n = m \otimes_A (an). \tag{4.11}$$

Faremos uso freqüente desse produto tensorial adiante. O mais importante para nós será a identidade $(ma) \otimes_A n = m \otimes_A (an)$ válida em todo $M \otimes_A N$ para todo $a \in A$.

4.3 Produtos Diretos e Somas Diretas

• **Produto Direto e Soma Direta de Grupos**

Seja J um conjunto arbitrário de índices e $\mathcal{G} \; := \; \{G_i, \; i \in J\}$ uma coleção de grupos. Seja o produto cartesiano $\mathbb{G} \; := \; \times_{i \in J} G_i$. Podemos fazer de $\mathbb{G}$ um grupo definindo o produto de dois elementos $\mathbb{G} \ni \underline{g} = \times_{a \in J} g_a$, $\mathbb{G} \ni \underline{h} = \times_{b \in J} h_b$ como $\underline{g} \cdot \underline{h} = \times_{a \in J} (g_a h_a)$. Com essa estrutura $\mathbb{G}$ é dito ser o *produto direto* dos grupos G_i, $i \in J$ e será denotado por $\mathbb{G}_p = \prod_{i \in J} G_i$.

$\mathbb{G}_p$ possui um subgrupo importante, aquele formado por elementos $\times_{a \in J} g_a \in \mathbb{G}_p$ onde apenas um número finito de g_a's é distinto da identidade e_a do respectivo grupo G_a. Esse subgrupo é dito ser a soma direta dos G_i's , $i \in J$ e é denotado por $\mathbb{G}_s = \bigoplus_{i \in J} G_i$.

• **Soma Direta de Espaços Vetoriais**

Se $\{V_i, \; i \in J\}$ é uma coleção de espaços vetoriais que, em particular, são grupos Abelianos, cai definida, pelo apresentado na sub-seção anterior, a soma direta $\mathbb{V}_s \; := \; \bigoplus_{i \in J} V_i$, definida primeiramente como grupo Abeliano. $\mathbb{V}_s$ pode ser feito um espaço vetorial definindo-se, para um escalar genérico $\alpha \in \mathbb{C}$,

$$\alpha \cdot (\times_{a \in J} v_a) \; := \; \times_{a \in J} (\alpha v_a), \tag{4.12}$$

para todo $\times_{a \in J} v_a \in \mathbb{V}_s$.

Um caso especial que irá nos interessar é o seguinte: seja M um bimódulo sobre uma álgebra A e tomemos $J = \mathbb{N}$ e $V_n = M^{\otimes_A n}$. Teremos então definida a soma direta $\bigoplus_{n \in \mathbb{N}} M^{\otimes_A n}$.

4.4 Módulos e Derivações

Seja A uma álgebra sobre $\mathbb{C}$ com identidade e e seja $\mathcal{M}$ um bimódulo sobre A. Uma aplicação linear $\delta : \; A \to \mathcal{M}$ é dita ser uma *derivação* de A em $\mathcal{M}$ se satisfaz a regra de Leibnitz:

$$\delta(ab) \; = \; a\delta(b) + \delta(a)b, \tag{4.13}$$

para todos $a, \, b \in A$.

Vamos a alguns exemplos.

Exemplo 1. Seja A álgebra sobre $\mathbb{C}$ com unidade e e $\mathcal{M} = A \otimes_{\mathbb{C}} A$ com os seguintes produtos de bimódulo:

$$a \cdot (b \otimes c) \;\; := \;\; (ab) \otimes c, \tag{4.14}$$

$$(b \otimes c) \cdot a \;\; := \;\; b \otimes (ca). \tag{4.15}$$

Deixa-se ao leitor verificar a associatividade dos produtos de bimódulo nesse caso. Defina-se

$$\delta(a) \; := \; a \otimes e - e \otimes a. \tag{4.16}$$

Deixa-se ao leitor verificar a validade da regra de Leibniz nesse exemplo. Note-se também que, por essa definição, $\delta(e) = 0$.

Exemplo 2. Seja A álgebra sobre $\mathbb{C}$ com unidade e e $\mathcal{M} = A \otimes_{\mathbb{C}} A$ com os seguintes produtos de bimódulo:

$$a \cdot (b \otimes c) \;\; := \;\; (ab) \otimes c, \tag{4.17}$$

$$(b \otimes c) \cdot a \;\; := \;\; b \otimes (ca) - (bc) \otimes a. \tag{4.18}$$

Deixa-se ao leitor verificar a associatividade dos produtos de bimódulo nesse caso. Defina-se

$$\delta(a) \; := \; e \otimes a. \tag{4.19}$$

Deixa-se ao leitor verificar a validade da regra de Leibniz nesse exemplo. Note-se também que, por essa definição, $\delta(e) = e \otimes e \neq 0$.

Exemplo 3. Exemplo importante de derivações pode ser vistos em álgebras de Lie. Seja $\mathcal{A}$ uma álgebra de Lie vista como um bimódulo sobre si mesma. Seja z um elemento fixo da álgebra e seja a aplicação $d_z : \mathcal{A} \to \mathcal{A}$ dada por $d_z(a) = [z, \; a]$. É fácil verificar (faça!) usando a identidade de Jacobi (2.11) que

$$d_z([a, \; b]) \; = \; [d_z(a), \; b] + [a, \; d_z(b)]$$

para todo $a, \; b \in \mathcal{A}$. Assim, tem-se que a cada $z \in \mathcal{A}$ é associada uma derivação d_z.